Optimal Uncertainty Quantification via Convex Optimization and Relaxation

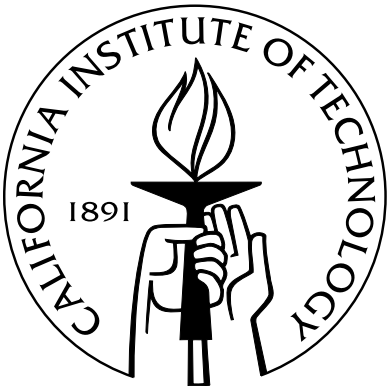
Thesis by

Shuo Han

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2014

(Defended September 26, 2013)

© 2014
Shuo Han
All Rights Reserved

To my parents, Shiquan Han and Yibing Liu.

Acknowledgments

I am grateful to my advisor Richard Murray for his endless support throughout my doctoral study. I cannot imagine surviving this stressful yet ultimately rewarding journey without his constant encouragement and enthusiasm. He has not only taught me the strategy for selecting important research directions, but also how to stay detail-oriented and get things done. By giving me a lot of freedom in research, he has allowed me to explore and learn from failures so that I could eventually settle down on a subject that I truly enjoy. His professional attitude towards teaching and mentoring is also unparalleled and has undoubtedly set up a model for my future academic career. I always feel that I still have a lot to learn from him.

I would like to thank Venkat Chandrasekaran, Babak Hassibi, Steven Low, and Houman Owhadi for serving on my thesis defense committee. In particular, I am indebted to Houman for introducing me to the area of optimal uncertainty quantification, which eventually leads to the development of this thesis, and to Steven, for spending time with me identifying research problems in power systems. I would also like to thank Yaser Abu-Mostafa, Michael Dickinson, John Doyle, and Pietro Perona for serving on my candidacy committee.

I would also like to extend my gratitude to my collaborators: to Andrew Straw, for helping me get started on my very first research project and teaching me insect biology, to Andrea Censi, for sharing his great vision on robotics and endless computer tips, to Sawyer Fuller, for his knack of engineering, to Molei Tao, for various insightful discussions and being one of the best roommates over the years, and to Ufuk Topcu, for advice on research directions and help on my transition to the next stage of academic career.

I feel very fortunate to have two amazing groups of people around me: people from the Control and Dynamical Systems option and from the Rigorous Systems Sesearch Group (RSRG). They have greatly enriched both my academic and social life at Caltech. I would like to especially thank Steven Low and Adam Wierman for allowing me to attend the RSRG meetings regularly.

I would like to acknowledge the generous financial support from these funding sources: the Institute for Collaborative Biotechnology (through the Army Research Office, W911NF-09-0001), Boeing, the National Science Foundation (CNS-0931746), the Air Force Office of Scientific Research (FA9550-12-1-0389), and the Predictive Science Academic Alliance Program (through the Department of Energy).

I would like to extend my appreciation to my parents, who are an ocean away back in China, for everything they have done in my life and being there for me. I would also like to thank my uncle and his family for their warm accommodations during my stay in Los Angeles. Last but not least, I would like to thank my girlfriend, Sijia, for staying at my side during some of my most difficult times and sharing many of my happiest moments.

Abstract

Many engineering applications face the problem of bounding the expected value of a quantity of interest (performance, risk, cost, etc.) that depends on stochastic uncertainties whose probability distribution is not known exactly. Optimal uncertainty quantification (OUQ) is a framework that aims at obtaining the best bound in these situations by explicitly incorporating available information about the distribution. Unfortunately, this often leads to non-convex optimization problems that are numerically expensive to solve.

This thesis emphasizes on efficient numerical algorithms for OUQ problems. It begins by investigating several classes of OUQ problems that can be reformulated as convex optimization problems. Conditions on the objective function and information constraints under which a convex formulation exists are presented. Since the size of the optimization problem can become quite large, solutions for scaling up are also discussed. Finally, the capability of analyzing a practical system through such convex formulations is demonstrated by a numerical example of energy storage placement in power grids.

When an equivalent convex formulation is unavailable, it is possible to find a convex problem that provides a meaningful bound for the original problem, also known as a convex relaxation. As an example, the thesis investigates the setting used in Hoeffding's inequality. The naive formulation requires solving a collection of non-convex polynomial optimization problems whose number grows doubly exponentially. After structures such as symmetry are exploited, it is shown that both the number and the size of the polynomial optimization problems can be reduced significantly. Each polynomial optimization problem is then bounded by its convex relaxation using sums-of-squares. These bounds are found to be tight in all the numerical examples tested in the thesis and are significantly better than Hoeffding's bounds.

Contents

Chapter 1

Introduction

Stochastic factors are prevalent in modern engineering systems, and often need to be accounted for explicitly during system design. For example, when designing a data center, how should the servers be scheduled optimally when the user demands can be random over time? When designing a bridge, how to assess the probability of collapse under random ambient disturbances such as vibrations and temperature changes? When designing a power grid, how to guarantee that the random generation from the renewable will not disrupt the operation of the system?

Before answering any of these questions, one first needs to face the question of how to choose the corresponding probabilistic model. As said by George E. P. Box, "all models are wrong, but some are useful". One common practice is to choose some simple distribution, and hope it will capture the stochasticities well. Fortunately, blessed by the central limit theorem, many stochastic phenomena can be well modeled by the universal and all-mighty Gaussian distribution. This has led to a myriad of simple yet powerful algorithms, including the method of least squares [12] and the Kalman filter [37] as two of the most well known examples. Even when Gaussian distributions fail, there are cases where other simple distributions would model the complex stochastic phenomena relatively well. One important generalization of the Gaussian distribution is the exponential family, which includes not only the Gaussian distribution, but also other useful distributions, such as the exponential, Poisson, beta, and Dirichlet distributions. Distributions in the exponential family have various properties that are amenable to numerical computation. In particular, composition of multiple distributions in the exponential family can lead to more sophisticated yet still tractable models. These are known as probabilistic graphical models, which are used extensively in machine learning [6, 13, 38].

On the flip side, the consequence of the mismatch between any of these *a priori* models and actual stochasticities is usually difficult to characterize without extensive test of the system in practice. This can become problematic for most critical applications where the design needs to be certified before deployment of the system. Perhaps one of the most famous example is the Space Shuttle Columbia disaster in 2003. After investigation of the disaster has been carried out, several memos and email communications were revealed (Fig. 1.1). As a matter of fact, on the day before reentry, Bob Doremus and David Paternostro, two managers in mission operations at the Johnson Space Center, "expressed some skepticism" during a discussion with two engineers about a simulation, which showed that the landing would be survivable with two flat tires. Nevertheless, no one expected the worst case to happen and assumed a safe entry. While there were many factors that contributed to the loss of Columbia, this shows how overly optimistic treatment of uncertainties in the model can potentially become disastrous.

Written summary provided by Bob Doremus regarding a conversation held on Jan. 31 between Carlisle Campbell, Robert Doremus and David Paternostro about STS-107:

> Carlisle Campbell phoned DF52/Bob Doremus. DF53/David Paternostro was also in the office. Carlisle brought in Bob Daugherty and the 4 discussed the possibility of landing with 2 flat tires. Carlisle said that Howard Law had done an entry sim at Ames (the sim was evidently done on Friday) and that sim showed that the landing with 2 flat tires was survivable. Bob Doremus and David Paternostro expressed some skepticism as to the accuracy of the Ames sim in light of other data (Convair 990 testing), but appreciated the information. All four agreed at the end of the discussion that we were doing a "what-if" discussion and that we all expected a safe entry on Saturday.

Figure 1.1: An excerpt of email during the last mission of Space Shuttle Columbia regarding a simulation of landing. (Source: NASA [2])

Analysis on the effects of uncertainties is highly dependent on the information available on the underlying distribution. In one extreme regime, where the amount of data is sufficient to sample the underlying distribution, the problem becomes conceptually easy, since traditional statistical estimation methods such as bootstrapping is expected to work well. Aside from computational challenges, however, this approach usually does not apply for more complex systems, where sampled data are often difficult or expensive to obtain so that construction of the actual distribution is often beyond reach.

Still, there is hope for analyzing the effect of uncertainty in the model at least to a certain level, since it is often possible to obtain partial knowledge about the underlying distribution. In the scheduling of queueing systems, for example, certain types of analysis

on the system performance can be made by only knowing whether the distribution of job sizes is heavy-tailed or light-tailed, despite the fact that the actual distribution is unknown (cf. [47]). Another case is when certain functionals (e.g., mean, variance, or any moments) of the underlying distribution are known. These functionals will be referred to as *information constraints*. Since the information constraints imposed by these functionals can be satisfied by multiple distributions, one popular approach is to choose the distribution by adopting the *principle of maximum entropy* [35, 36]. The idea of this principle is to incorporate as little prior knowledge as possible into the distribution. Although the principle of maximum entropy has been proved in many cases as a good guideline for choosing the distribution, one should keep in mind that the ultimate objective in many engineering applications is to obtain a good estimate of some quantity of interest (e.g, performance or cost) other than the distribution itself. From this viewpoint, the principle of maximum entropy uses the entropy as the "quantity of interest", which can be meaningless for most engineering systems.

In particular, one useful quantity is the upper and lower bounds of the expected performance, which can serve as certificates of the design. Formally, this can be obtained by solving an infinite-dimensional optimization problem that maximizes or minimizes the expected performance over all probability distributions that satisfy the information constraints. Study of these kinds of problems has become a research direction named *optimal uncertainty quantification* (OUQ), which was originally formulated in [53]. Despite the fact that the optimization problem to be solved is infinite-dimensional, it has been shown that the problem can be reduced to a finite-dimensional yet equivalent one, for which numerical solutions are available. OUQ has been applied in various engineering problems, include seismic safety assessment of structures and transport in porous media [53].

Organization of the thesis

This thesis emphasizes on developing efficient computational methods for solving a number of problems within the OUQ framework. As R. Rockafellar pointed out in 1993, "*the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity*" [60]. By virtue of this, Chapter 2 seeks cases for which an equivalent convex formulation exists. Specially, it proceeds by viewing the OUQ problem from two different perspectives: the primal form and dual form of the corresponding optimization problem. Conditions on the objective function and information constraints under which a convex for-

mulation exists are presented. Compared to previous results, the new cases presented in this chapter allow more freedom in incorporating knowledge about the unknown probability distribution and can potentially provide better quantification, as demonstrated in a simple example using the Gaussian distribution.

As a case study, Chapter 3 applies the results in Chapter 2 to a problem that arises in power systems: evaluating the effect of placing energy storage in power grids when renewable generation is present. Despite the existence of an equivalent convex formulation, the corresponding convex optimization problem becomes challenging to solve numerically due to the size of the system. Two approaches are presented to address the scaling issues. Although these approaches are developed specifically for the problem of energy storage placement, the form of the problem (derived in Section 2.5) is general enough so that they can be applied to other examples as well. One approach focuses on solving a number of smaller problems in order to obtain the solution iteratively. This is made possible by exploiting the special form of the objective function in the optimization problem named *polytopic canonical form.* Another approach focuses on solving the large original problem through massive parallelization. After manipulation, the optimization problem can be converted to the standard form solvable using the alternating direction method of multipliers, which is a parallelizable first-order method. In the end, these approaches are demonstrated through numerical examples using standard test cases used in power systems research.

When a convex formulation in unavailable, the next best option is to find a convex problem that gives a meaningful numerical bound for the original problem. This procedure is commonly referred to as convex relaxation. Relaxation fits naturally in the context of OUQ since it is aligned with the original purpose of obtaining bounds for some quantity of interest. In Chapter 4, we study a setting that is used in Hoeffding's inequality. In this setting, one is interested in obtaining a bound for the probability that the sum of independent random variables deviates from its expected sum. Hoeffding's inequality gives a bound in simple expressions, but this bound is not tight. As it turns out, this problem falls into the OUQ framework and a tight bound can be obtained by solving a series of non-convex polynomial optimization problems.

Chapter 2

OUQ via Convex Optimization: Theory

The purpose of this chapter is to introduce the optimal uncertainty quantification (OUQ) problem and several special cases in which the problem can be solved efficiently using convex optimization. The chapter begins with the formulation of the OUQ problem as an infinite-dimensional optimization problem and its equivalent finite-dimensional formulation. In general, the finite-dimensional problem is non-convex and can still be difficult to solve. The chapter attempts to derive convex formulation for several cases from two different perspectives: the primal form and dual form of the optimization problem. Compared to previous results, the new cases presented in this chapter allow more freedom in incorporating knowledge about the unknown probability distribution and can potentially provide better quantification, which is demonstrated in a simple example.

2.1 Optimal uncertainty quantification and finite reduction

The optimal uncertainty quantification problem is an optimization problem in the form:

$$\begin{aligned}
\underset{\mathcal{D}}{\text{maximize}} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}\left[f(\theta)\right] && (2.1)\\
\text{subject to} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}[g(\theta)] \preceq 0 && (2.2)\\
& \mathbb{E}_{\theta\sim\mathcal{D}}[h(\theta)] = 0 && (2.3)\\
& \theta \in \Theta \quad \text{almost surely,} && (2.4)
\end{aligned}$$

where θ is a random variable defined on $\mathbb{R}^d$ whose probability distribution is $\mathcal{D}$, and f, g, h are measurable functions defined on $\mathbb{R}^d$. The distribution is constrained to have support on a given set $\Theta \subseteq \mathbb{R}^d$. The inequality in (2.2) denotes generalized inequality with respect to a certain cone: in the simplest case, if the cone is the positive orthant, then the generalized inequality becomes entry-wise inequality. Note that the condition that $\mathcal{D}$ is a probability distribution automatically implies that

$$\mathbb{E}_{\theta\sim\mathcal{D}}[1] = 1, \qquad \mathcal{D} \geq 0.$$

The functions g and h are used to incorporate the available information of $\mathcal{D}$. One important class of g and h is powers of θ, which are used to represent information on the moments of θ. The use of inequality constraints arises when, for example, one does not have perfect estimates and would rather use a confidence interval. For example, we can choose to incorporate bounds on the mean of θ:

$$\hat{\mu}_{\text{lb}} \preceq \mathbb{E}[\theta] \preceq \hat{\mu}_{\text{ub}}.$$

In the following, we give a few examples of optimal uncertainty quantification problems.

Example 2.1 (Support constraints)**.** Suppose the random variable $\theta \in \mathbb{R}^d$ and let f be an arbitrary function and $\Theta \subseteq \mathbb{R}^d$. Then the problem with constraint on the support of θ

$$\begin{aligned} \underset{\mathcal{D}}{\text{maximize}} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}\left[f(\theta)\right] \\ \text{subject to} \quad & \theta \in \Theta \quad \text{almost surely} \end{aligned} \tag{2.5}$$

is an optimal uncertainty quantification problem with only the support constraint. In fact, problem (2.5) can be reduced to an optimization problem over a finite-dimensional decision variable in this case:

$$\begin{aligned} \underset{\theta}{\text{maximize}} \quad & f(\theta) \\ \text{subject to} \quad & \theta \in \Theta. \end{aligned}$$

This equivalence has appeared in many applications, including moment-based relaxations

of polynomial optimization problems [39] and the concept of least favorable prior [77]. To show this, note that

$$\max_{\mathcal{D}} \mathbb{E}_{\theta\sim\mathcal{D}}[f(\theta)] \leq \max_{\theta\in\Theta} f(\theta),$$

since the average can never exceed the maximum regardless of the distribution $\mathcal{D}$. On the other hand,

$$\max_{\mathcal{D}} \mathbb{E}_{\theta\sim\mathcal{D}}[f(\theta)] \geq \max_{\theta\in\Theta} f(\theta),$$

since the right-hand side can be achieved by a Dirac-delta distribution concentrated at $\theta^* = \arg\max_{\theta\in\Theta} f(\theta)$.

Example 2.2 (Probability with moment constraints)**.** Suppose we only have access to the mean and covariance of a certain random variable θ defined on $\mathbb{R}^d$, but we are interested in $\mathbb{P}(\theta \in \Theta)$ for some set $\Theta \subseteq \mathbb{R}^d$. In many cases, the set Θ may correspond to some undesired event and we would like to quantify its worst-case probability by solving the following problem:

$$\begin{array}{ll} \underset{\mathcal{D}}{\text{maximize}} & \mathbb{P}(\theta \in \Theta) \\ \text{subject to} & \mathbb{E}_{\theta\sim\mathcal{D}}[\theta] = \hat{\mu}, \qquad \text{cov}[\theta] = \widehat{\Sigma}, \end{array}$$

where $\hat{\mu}$ and $\widehat{\Sigma}$ are the mean and covariance of θ, respectively. The problem can be converted into an optimal uncertainty quantification problem by defining f and h as follows:

$$f(\theta) = I(\theta \in \Theta) = \begin{cases} 1 & \theta \in \Theta \\ 0 & \theta \notin \Theta, \end{cases}$$

and

$$h(\theta) = \begin{bmatrix} \theta - \hat{\mu} \\ \mathbf{vec}(\theta\theta^T - \widehat{\Sigma} - \hat{\mu}\hat{\mu}^T) \end{bmatrix},$$

where **vec** denotes the vectorization of a matrix.

Example 2.3 (Probability on disjoint sets)**.** In this more advanced example, similar to

Example 2.2, we want to compute the maximum of

$$\mathbb{P}\left[\bigvee_{i=1}^{d}(\theta_i \geq a_i)\right]$$

for some given constant $a \in \mathbb{R}^d$ where the distribution of θ is subject to the same moment constraints as in Example 2.2. Here, θ_i and a_i denote the i-th element of θ and a, respectively. The problem can be converted into an optimal uncertainty quantification problem by defining f as follows:

$$f(\theta) = \max\left\{I^1_{-\infty}(\theta_1 \geq a_1), \ldots, I^1_{-\infty}(\theta_d \geq a_d), 0\right\},$$

where

$$I^1_{-\infty}(A) = \begin{cases} 1 & A \text{ is true} \\ -\infty & A \text{ is false} \end{cases}$$

is the $-\infty$-1 indicator function for a given set A. The definition of h is the same as in Example 2.2.

An optimal uncertainty quantification problem is an optimization problem over the infinite-dimensional space of probability distributions with objective functions and constraints linear in the distribution $\mathcal{D}$. Since the space of probability distributions is convex and the constraints are also convex, the optimization problem is convex. However, the fact that an optimization problem is convex does not immediately imply that it is numerically tractable. In fact, any non-convex optimization problem can always be rewritten as a convex problem either with an infinite number of constraints or over an infinite-dimensional space [39]. The perhaps surprising fact, as shown previously by other researchers (cf. [65, 53]), is that any optimal uncertainty quantification problem can always be reduced to an equivalent finite-dimensional optimization problem in the sense that the reduced problem yields the same optimal value.

Theorem 2.4 (Finite reduction property)**.** *The (finite-dimensional) problem*

$$\underset{\{p_i,\theta_i\}_{i=1}^n}{\text{maximize}} \quad \sum_{i=1}^{n} p_i f(\theta_i) \tag{2.6}$$

$$\begin{aligned} \text{subject to} \quad & \sum_{i=1}^{n} p_i = 1, \qquad p \succeq 0 \\ & \sum_{i=1}^{n} p_i g(\theta_i) \preceq 0 \\ & \sum_{i=1}^{n} p_i h(\theta_i) = 0 \end{aligned}$$

achieves the same optimal value as problem (2.1)*. Here, n equals 1 plus the total number of independent scalar equalities encoded in g and h.*

Example 2.5 ("Seesaw")**.** Consider the following uncertainty quantification problem for a scalar random variable θ:

$$\begin{aligned} \underset{\mathcal{D}}{\text{maximize}} \quad & \mathbb{P}(\theta \geq \gamma) \\ \text{subject to} \quad & \mathbb{E}_{\theta \sim \mathcal{D}}[\theta] = 0, \qquad a \leq \theta \leq b, \end{aligned}$$

where a, b, and γ are constants satisfying $a < 0 \leq \gamma < b$. In order to maximize $\mathbb{P}(\theta \geq \gamma)$, we would want to assign as much probability as possible on the right side of γ. However, the condition $\mathbb{E}_{\theta \sim \mathcal{D}}[\theta] = 0$ requires that the probability on both sides must be balanced around 0. This is analogous to a seesaw pivoted at 0 with two end points at a and b, respectively (Fig. 2.1). It is not difficult to see that the best assignment is to put all the probability on the right side at γ (for least leverage) and all the probability on the left side at a (for most leverage). This implies that the optimal distribution can be achieved with a discrete distribution consisting of two Dirac masses at a and γ, respectively. Indeed, since there is only one scalar constraint, we have $n = 1$ and the total number of Dirac masses predicted by Theorem 2.6 is $n + 1 = 2$.

Example 2.6 (Probability with moment constraints, revisited)**.** Consider again Example 2.2. Suppose $\theta \in \mathbb{R}^d$, then the constraint $(\theta - \hat{\mu})$ will give d equality constraints, and $\mathbf{vec}(\theta\theta^T - \widehat{\Sigma} - \hat{\mu}\hat{\mu}^T)$ will give $d(d+1)/2$ independent equality constraints (rather than d^2, since $\widehat{\Sigma}$ is symmetric). Therefore, $n = d + d(d+1)/2$ in this example.

Note that problem (2.1) can also be written in a form without the equality constraints

$$\mathbb{E}_{\theta \sim \mathcal{D}}[h(\theta)] = 0$$

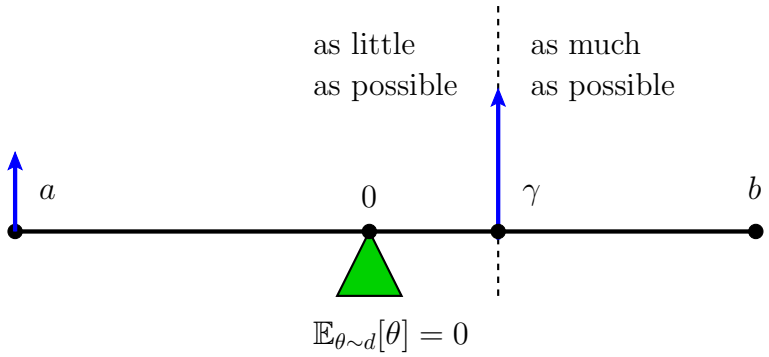


Figure 2.1: "Seesaw" analogy in Example 2.5

by introducing inequality constraints

$$\mathbb{E}_{\theta\sim\mathcal{D}}[h(\theta)] \preceq 0, \qquad \mathbb{E}_{\theta\sim\mathcal{D}}[-h(\theta)] \preceq 0.$$

However, we will still use the form as in (2.1) to distinguish pure equalities from pure inequalities in order to define n properly.

A proof of Theorem 2.4 can be found in, e.g., [53]. Here we would like to present an informal proof to give some intuition. We start by approximating the probability distribution $\mathcal{D}$ in the original problem (2.1) using a discrete distribution in which the masses are located at $\{\theta_i\}_{i=1}^M$ with weights $\{q_i\}_{i=1}^M$, where M is potentially a very large number. Under such approximation, problem (2.1) becomes

$$\underset{\{q_i,\theta_i\}_{i=1}^M}{\text{maximize}} \quad \sum_{i=1}^M q_i f(\theta_i) \tag{2.7}$$

$$\begin{aligned} \text{subject to} \quad & \sum_{i=1}^M q_i = 1, \qquad q \succeq 0 \\ & \sum_{i=1}^M q_i g(\theta_i) \preceq 0 \\ & \sum_{i=1}^M q_i h(\theta_i) = 0. \end{aligned} \tag{2.8}$$

Note that this problem is similar to problem (2.6) in Theorem 2.4, except that the number of Dirac masses is M. Suppose the optimal solution of problem (2.7) is $\{q_i^*, \theta_i^*\}_{i=1}^M$, and

define the set

$$\mathcal{I} = \{i \colon q_i^* \neq 0\} \subseteq \{1, 2, \ldots, M\}.$$

Since $\{q_i^*\}_{i=1}^M$ satisfies equation (2.8), it implies that

$$\sum_{i=1}^M q_i^* g(\theta_i^*) \in \operatorname{conv}\{g(\theta_i^*)\}_{i\in\mathcal{I}}, \qquad \sum_{i=1}^M q_i^* h(\theta_i^*) \in \operatorname{conv}\{h(\theta_i^*)\}_{i\in\mathcal{I}},$$

or, written in vector form,

$$\begin{bmatrix} \sum_{i=1}^M q_i^* g(\theta_i^*) \\ \sum_{i=1}^M q_i^* h(\theta_i^*) \end{bmatrix} \in \operatorname{conv}\left\{ \begin{bmatrix} g(\theta_i^*) \\ h(\theta_i^*) \end{bmatrix} \right\}_{i\in\mathcal{I}},$$

where $\operatorname{conv}(P)$ denotes the convex hull of a given set P. We will then make use of Carathéodory's theorem (cf. pp. 126, [67]) from convex geometry, which is stated below:

Theorem 2.7 (Carathéodory)**.** *If a point $x \in \mathbb{R}^n$ lies in the convex hull of a set $P \subseteq \mathbb{R}^n$, then there is a subset $P' \subseteq P$ consisting of $n+1$ or fewer points such that x lies in the convex hull of P'.*

From Carathéodory's theorem, we know that there exists a subset $\mathcal{A} \subseteq \mathcal{I}$ whose size is at most 1 plus the total number of independent components of g and h such that

$$\begin{bmatrix} \sum_{i=1}^M q_i^* g(\theta_i^*) \\ \sum_{i=1}^M q_i^* h(\theta_i^*) \end{bmatrix} \in \operatorname{conv}\left\{ \begin{bmatrix} g(\theta_i^*) \\ h(\theta_i^*) \end{bmatrix} \right\}_{i\in\mathcal{A}},$$

i.e., there exists nonnegative $\{p_i^*\}_{i\in\mathcal{A}}$ satisfying $\sum_{i\in\mathcal{A}} p_i^* = 1$ such that

$$\sum_{i=1}^M q_i^* g(\theta_i^*) = \sum_{i\in\mathcal{A}} p_i^* g(\theta_i^*), \qquad \sum_{i=1}^M q_i^* h(\theta_i^*) = \sum_{i\in\mathcal{A}} p_i^* h(\theta_i^*).$$

On the other hand, from the Lagrangian of problem (2.7),

$$L = \sum_{i=1}^M q_i f(\theta_i) - \lambda^T \sum_{i=1}^M q_i g(\theta_i) + \nu^T \sum_{i=1}^M q_i h(\theta_i) + \sum_{i=1}^M \mu_i q_i + \mu_0 \left(1 - \sum_{i=1}^M q_i\right),$$

where $\lambda, \nu, \{\mu_i\}_{i=0}^M$ are the dual variables, and the Karush-Kuhn-Tucker condition, we know

that

$$\frac{\partial L}{\partial q_i} = 0$$

at the optimum, i.e.,

$$f(\theta_i^*) - \lambda^{*T} g(\theta_i^*) + \nu^{*T} h(\theta_i^*) + \mu_i^* - \mu_0^* = 0, \quad i = 1, 2, \ldots, M$$

for dual optimal $\lambda^*, \nu^*, \{\mu_i^*\}_{i=0}^M$. Thereore, we have

$$\begin{aligned}\sum_{i\in\mathcal{A}} p_i^* f(\theta_i^*) &= \sum_{i\in\mathcal{A}} p_i^* \left[\lambda^{*T} g(\theta_i^*) - \nu^{*T} h(\theta_i^*) - \mu_i^* + \mu_0^*\right] \\ &= \lambda^{*T} \sum_{i\in\mathcal{A}} p_i^* g(\theta_i^*) - \nu^{*T} \sum_{i\in\mathcal{A}} p_i^* h(\theta_i^*) - \sum_{i\in\mathcal{A}} p_i^* \mu_i^* + \mu_0^*.\end{aligned}$$

On the other hand, from complementary slackness, we know that

$$\mu_i^* q_i^* = 0, \quad i = 1, 2, \ldots, M,$$

which implies that $\mu_i^* = 0$ for all $i \in \mathcal{I}$ (hence for all $i \in \mathcal{A}$). Therefore, we have

$$\begin{aligned}\sum_{i\in\mathcal{A}} p_i^* f(\theta_i^*) &= \lambda^{*T} \sum_{i=1}^M q_i^* g(\theta_i^*) - \nu^{*T} \sum_{i=1}^M q_i^* h(\theta_i^*) + \mu_0^* \\ &= \lambda^{*T} \sum_{i=1}^M q_i^* g(\theta_i^*) - \nu^{*T} \sum_{i=1}^M q_i^* h(\theta_i^*) - \sum_{i=1}^M \mu_i^* q_i^* + \mu_0^* \\ &= \sum_{i=1}^M q_i^* \left[\lambda^{*T} g(\theta_i^*) - \nu^{*T} h(\theta_i^*) - \mu_i^* + \mu_0^*\right] \\ &= \sum_{i=1}^M q_i^* f(\theta_i^*),\end{aligned}$$

which is the optimal value of problem (2.7). Note that $\{p_i^*, \theta_i^*\}_{i\in\mathcal{A}}$ is a feasible solution to the reduced problem

$$\begin{aligned}\underset{\{p_i,\theta_i\}_{i\in\mathcal{A}}}{\text{maximize}} \quad & \sum_{i\in\mathcal{A}} p_i f(\theta_i) \\ \text{subject to} \quad & \sum_{i\in\mathcal{A}} p_i = 1, \qquad p_i \geq 0, \quad i \in \mathcal{A}\end{aligned} \tag{2.9}$$

$$\sum_{i\in\mathcal{A}} p_i g(\theta_i) \preceq 0$$
$$\sum_{i\in\mathcal{A}} p_i h(\theta_i) = 0,$$

and hence the optimal value of problem (2.7) is no larger than the optimal value of problem (2.9). On the other hand, it is easy to verify that the optimal value of problem (2.9) cannot be larger than that of problem (2.7) due to the fact that the former has fewer decision variables. Therefore, both problem (2.7) and the reduced problem (2.9) achieve the same optimal value. It is worth pointing out that the proof is incomplete, since it starts by approximating the distribution using a discrete distribution. Therefore, it remains unclear how the results will extend as $M \to \infty$, i.e., when the discrete distribution approaches the continuous distribution $\mathcal{D}$. This issue has been ignored in this informal proof.

2.2 Convex reformulation via primal form

Although the optimal uncertainty quantification problem adopts a finite reduction, it is still unclear whether it can be solved efficiently, i.e., in polynomial time. In this section, we will show that this is true if both of the following conditions hold:

1. The function f that appears in the objective is *piecewise concave*, i.e., it can be written as
$$f(\theta) = \max_{k=1,2,\ldots,K} f^{(k)}(\theta), \tag{2.10}$$
where each function $f^{(k)}$ is concave.

2. The function h that appears in the constraints is affine:
$$h(\theta) = A^T\theta + b. \tag{2.11}$$

3. The function g, which needs to be defined from $\mathbb{R}^d$ to $\mathbb{R}^p$, is entry-wise piecewise convex, i.e., each entry g_i $(i = 1, 2, \ldots, p)$ can be written as
$$g_i(\theta) = \min_{l_i=1,2,\ldots,L_i} g_i^{(l_i)}(\theta), \tag{2.12}$$
where each function $g_i^{(l_i)}$ is convex.

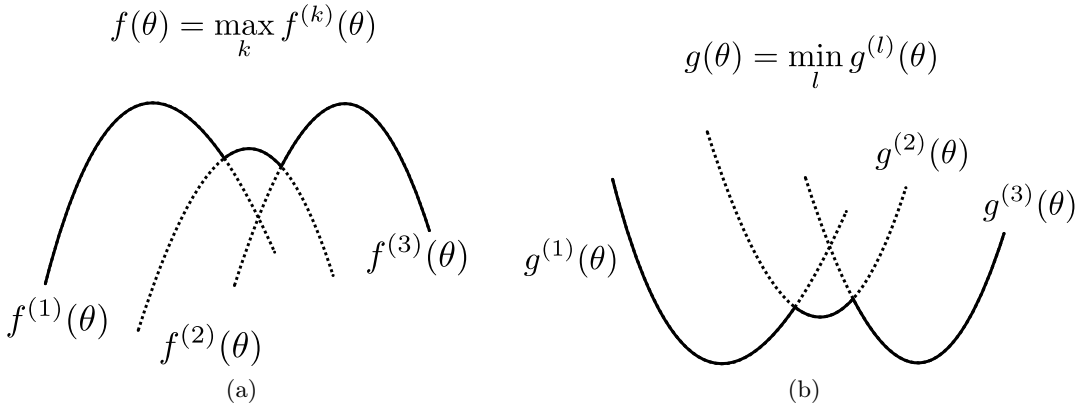


Figure 2.2: (a) Piecewise concave and (b) piecewise convex functions in one dimension.

Fig. 2.2 illustrates how piecewise concave and piecewise convex functions would look like in one dimension. In general, these functions are neither concave nor convex. Nevertheless, we list in the following several cases that can be expressed in piecewise concave/convex form. We begin by the examples where the function f appearing in the objective is piecewise concave:

Example 2.8 (Concave functions)**.** The function f itself is concave. In this case, $K = 1$ and $f = f^{(1)}$ is concave.

Example 2.9 (Piecewise affine and convex)**.** The function f is piecewise affine and convex. In this case, $f^{(k)}$ is affine (hence concave) for each k. This case will be discussed later in greater detail.

Example 2.10 (Tail probability)**.** The random variable θ is univariate and the function f is the 0-1 indicator function:

$$f(\theta) = I(\theta \geq a) = \begin{cases} 1 & \theta \geq a \\ 0 & \theta < a \end{cases}$$

for some constant $a \in \mathbb{R}$. In this case, the function f can be written as

$$f(\theta) = \max\{0, I^1_{-\infty}(\theta \geq a)\}.$$

It can be readily verified that both 0 and $I^1_{-\infty}(\theta \geq a)$ are concave.

Next we list several cases where the function g appearing in the constraints is piecewise convex.

Example 2.11 (Even-order moments)**.** The random variable θ is univariate and $g_i(\theta) = \theta^{2q}$ for some positive integer q. This is a special case in which the function g_i itself is convex.

Example 2.12 (Tail probability)**.** This is similar to Example 2.10 in that we define $g(\theta) = I(\theta \geq a)$ for some constant $a \in \mathbb{R}$. In this case, the function g can be written in a different way as

$$g(\theta) = \min\{1, I_0^{\infty}(\theta \geq a)\}.$$

Again, it can be verified that both functions inside are convex.

We will now show that under such conditions on f, g, and h (i.e., conditions (2.10)–(2.12)) the optimal uncertainty quantification problem can be reformulated as a (finite-dimensional) convex optimization problem. For notational simplicity, we will first present the case in which g is defined from $\mathbb{R}^d$ to $\mathbb{R}$ (i.e., $p = 1$):

$$g(\theta) = \min_{l=1,2,\ldots,L} g^{(l)}(\theta), \qquad g^{(l)} \text{ is convex.} \tag{2.13}$$

This can be easily generalized to the case of $p > 1$. From the finite reduction property, we know that it suffices to use a finite number of Dirac masses to represent the optimal distribution. In this case, due to the special form of the objective function and constraints, these Dirac masses satisfy a useful property as given in the following lemma:

Lemma 2.13. *If the functions f and g can be expressed as*

$$f(\theta) = \max_{k=1,2,\ldots,K} f^{(k)}(\theta), \qquad g(\theta) = \min_{l=1,2,\ldots,L} g^{(l)}(\theta)$$

where $f^{(k)}$ $(k = 1, 2, \ldots, K)$ is concave and $g^{(l)}$ $(l = 1, 2, \ldots, L)$ is convex, and h is affine, then the optimal distribution can be achieved by a discrete distribution that contains at most $K \cdot L$ Dirac masses located at $\{\theta_{kl}\}$ $(k = 1, 2, \ldots, K;\ l = 1, 2, \ldots, L)$. In addition, each θ_{kl} satisfies

$$f(\theta_{kl}) = f^{(k)}(\theta_{kl}), \qquad g(\theta_{kl}) = g^{(l)}(\theta_{kl}),$$

i.e., it achieves maximum at $f^{(k)}$ and minimum at $g^{(l)}$.

Proof. Suppose for certain k and l, the optimal distribution contains two Dirac masses located at ϕ_1 and ϕ_2 with probabilities q_1 and q_2, respectively, whereas both ϕ_1 and ϕ_2 achieve maximum at $f^{(k)}$ and minimum at $g^{(l)}$, i.e.,

$$\begin{aligned} f(\phi_1) &= f^{(k)}(\phi_1), \qquad f(\phi_2) = f^{(k)}(\phi_2), \\ g(\phi_1) &= g^{(l)}(\phi_1), \qquad g(\phi_2) = g^{(l)}(\phi_2). \end{aligned}$$

Consider a new Dirac mass whose probability q and location ϕ are given by

$$q = q_1 + q_2, \qquad \phi = \frac{q_1\phi_1 + q_2\phi_2}{q_1 + q_2}.$$

It can be verified that replacing the two previous Dirac masses (q_1, ϕ_1) and (q_2, ϕ_2) with this new Dirac mass (q, ϕ) will still yield a valid probability distribution. Moreover, the new distribution will give an objective $\mathbb{E}[f(\theta)]$ that is no smaller than the previous one, since

$$qf(\phi) \geq qf^{(k)}(\phi) \geq q_1 f^{(k)}(\phi_1) + q_2 f^{(k)}(\phi_2) = q_1 f(\phi_1) + q_2 f(\phi_2), \tag{2.14}$$

where the second inequality is an application of Jensen's inequality and last equality uses the fact that ϕ_1 and ϕ_2 achieves maximum at $f^{(k)}$. On the other hand, the new distribution will remain as a feasible solution. The equality constraint on $\mathbb{E}[h(\theta)]$ remains feasible, because

$$\begin{aligned} qh(\phi) &= q(A^T\phi + b) = A^T(q_1 + q_2)\phi + b(q_1 + q_2) \\ &= A^T(q_1\phi_1 + q_2\phi_2) + b(q_1 + q_2) = q_1 h(\phi_1) + q_2 h(\phi_2). \end{aligned}$$

The feasibility of the inequality constraint on $\mathbb{E}[g(\theta)]$ can be proved by using a similar argument as in (2.14) by observing that $\mathbb{E}[g(\theta)]$ evaluated at the new distribution will be no larger than that at the original distribution, because

$$qg(\phi) \leq qg^{(l)}(\phi) \leq q_1 g^{(l)}(\phi_1) + q_2 g^{(l)}(\phi_2) = q_1 g(\phi_1) + q_2 g(\phi_2).$$

Therefore, the two old Dirac masses can be replaced by the new single one without affecting optimality, from which the uniqueness of θ_{kl} follows. □

The number of Dirac masses given by Lemma 2.13 is independent from the one given by the finite reduction property. From Lemma 2.13, we can obtain the equivalent convex optimization problem for the original problem:

Theorem 2.14. *The (convex) optimization problem*

$$\begin{aligned}
\underset{\{p_{kl},\gamma_{kl}\}_{k,l}}{\text{maximize}} \quad & \sum_{k,l} p_{kl} f^{(k)}(\gamma_{kl}/p_{kl}) \\
\text{subject to} \quad & \sum_{k,l} p_{kl} = 1 && (2.15) \\
& p_{kl} \geq 0, \qquad \forall k,l && (2.16) \\
& \sum_{k,l} p_{kl} h(\gamma_{kl}/p_{kl}) = A^T \left(\sum_{k,l} \gamma_{kl} \right) + b = 0 \\
& \sum_{k,l} p_{kl} g^{(l)}(\gamma_{kl}/p_{kl}) \leq 0
\end{aligned}$$

achieves the same optimal value as problem (2.1) *if the functions* f*,* g*, and* h *satisfy* (2.10)*,* (2.13)*, and* (2.11)*, respectively.*

Proof. According to Lemma 2.13, we can optimize over a new set of Dirac masses whose probability weights and locations are $\{p_{kl}, \theta_{kl}\}$. The requirement that the set of Dirac masses forms a valid probability distribution imposes the constraints (2.15) and (2.16). Under the new set of Dirac masses, the objective function can be rewritten as

$$\mathbb{E}\left[f(\theta)\right] = \sum_{k,l} p_{kl} f(\theta_{kl}) = \sum_{k,l} p_{kl} f^{(k)}(\theta_{kl}),$$

where the second equality uses the fact that θ_{kl} achieves maximum at $f^{(k)}$. As will be shown later, this step is critical since f is generally not concave, but $\sum_{k,l} p_{kl} f^{(k)}(\theta_{kl})$ is concave. Similarly, the constraint can be rewritten as

$$\mathbb{E}[g(\theta)] = \sum_{k,l} p_{kl} g(\theta_{kl}) = \sum_{k,l} p_{kl} g^{(l)}(\theta_{kl}).$$

The final form can be obtained by introducing new variables $\gamma_{kl} = p_{kl}\theta_{kl}$ for all k, l and choosing to optimize over $\{p_{kl}, \gamma_{kl}\}$ instead of $\{p_{kl}, \theta_{kl}\}$. Each term in the sum in the

objective function

$$\sum_{k,l} p_{kl} f^{(k)}(\gamma_{kl}/p_{kl})$$

is a perspective transform of $f^{(k)}$ and hence is concave since $f^{(k)}$ is concave. Therefore, the objective function is concave because it is a sum of concave functions. Likewise, the term

$$\sum_{k,l} p_{kl} g^{(l)}(\gamma_{kl}/p_{kl})$$

is convex and resulting constraint is also convex. The rest of the constraints does not affect convexity since they are affine inequality or equality constraints. In conclusion, the final optimization problem is a (finite-dimensional) convex problem and is equivalent to the original problem (2.1) due to Lemma 2.13. □

In addition, there are a couple of straightforward extensions to this formulation.

Multiple inequality constraints The formulation above corresponds to the case of $p = 1$. In the case of $p > 1$, the number of Dirac masses in the new set needs to be expanded to $K \cdot \prod_{i=1,2,\ldots,p} L_i$. Among these, there is at most one Dirac mass θ^* that achieves maximum at $f^{(k)}$ and minimum at $g_1^{(l_1)}, g_2^{(l_2)}, \ldots, g_p^{(l_p)}$ for any given k, $\{l_i\}_{i=1}^p$, i.e., the location θ^* satisfies

$$f(\theta^*) = f^{(k)}(\theta^*), \quad g_1(\theta^*) = g_1^{(l_1)}(\theta^*), \quad \ldots, \quad g_p(\theta^*) = g_p^{(l_p)}(\theta^*).$$

The corresponding convex optimization problem can then be formed by following a similar procedure given in the proof of Theorem 2.14.

Polytopic support constraints It is also possible to impose certain types of constraints on the support of distribution without affecting convexity. Specifically, the support of distribution can be constrained to be a polytope $\mathcal{P}_\theta$, i.e,

$$\theta \in \mathcal{P}_\theta \quad \text{almost surely.} \tag{2.17}$$

It is known that any polytope can always be represented as the intersection of several affine halfspaces, i.e., $\mathcal{P}_\theta = \{\theta \colon A\theta \preceq b\}$ for some given constants A and b (cf. [11]). After finite reduction using the Dirac masses $\{\theta_{kl}, p_{kl}\}$, the support constraint (2.17) becomes $K \cdot L$

separate constraints, each of which corresponds to a certain Dirac mass θ_{kl}:

$$A\theta_{kl} \preceq b. \tag{2.18}$$

Substitute $\gamma_{kl} = p_{kl}\theta_{kl}$ into (2.18) so that it becomes an affine inequality constraint in (γ_{kl}, p_{kl}):

$$A\gamma_{kl} \preceq p_{kl}b. \tag{2.19}$$

The final optimization problem will optimize over $\{p_{kl}, \gamma_{kl}\}$ with the extra constraints (2.19). Because (2.19) is affine, the resulting optimization problem remains convex.

2.3 Convex reformulation via dual form

Another convex formulation of optimal uncertainty quantification problems can be derived from the Lagrange dual problem of (2.1). In this section, it is assumed that the generalized inequality present in (2.2) is entry-wise inequality. In this case, the Lagrangian of problem (2.1) can be written as

$$\begin{aligned} L = &\int f(\theta)\mathcal{D}(\theta)\,d\theta - \lambda^T \int g(\theta)\mathcal{D}(\theta)\,d\theta - \nu^T \int h(\theta)\mathcal{D}(\theta)\,d\theta \\ &+ \int \lambda_p(\theta)\mathcal{D}(\theta)\,d\theta + \mu\left(1 - \int \mathcal{D}(\theta)\,d\theta\right). \end{aligned}$$

The latter two terms are due to fact that $\mathcal{D}$ is a probability distribution and hence $\mathcal{D} \geq 0$ and

$$\mathbb{E}[1] = \int \mathcal{D}(\theta)\,d\theta = 1.$$

The Lagrange dual can be derived as

$$\sup_{\mathcal{D}} L = \begin{cases} \mu & f(\theta) - \lambda^T g(\theta) - \nu^T h(\theta) + \lambda_p(\theta) - \mu = 0 \text{ for all } \theta \\ \infty & \text{otherwise.} \end{cases}$$

Combining the conditions on the Lagrange multipliers, i.e.,

$$\lambda \succeq 0, \qquad \lambda_p(\theta) \geq 0, \quad \forall\theta,$$

we can obtain the the dual problem as follows:

$$\begin{aligned} \underset{\lambda,\nu,\mu}{\text{minimize}} \quad & \mu && (2.20)\\ \text{subject to} \quad & f(\theta) - \lambda^T g(\theta) - \nu^T h(\theta) - \mu \le 0, \quad \forall \theta && (2.21)\\ & \lambda \succeq 0, \end{aligned}$$

which is a linear program with an infinite number of constraints (also known as a semi-infinite program). The inequality constraint (2.21) implies that the optimal solution $(\lambda^*, \nu^*, \mu^*)$ must satisfy

$$\mu^* = \max_{\theta} \left[f(\theta) - \lambda^{*T} g(\theta) - \nu^{*T} h(\theta) \right],$$

so that problem (2.20) can be rewritten by eliminating the inequality constraint (2.21) as

$$\begin{aligned} \underset{\lambda,\nu}{\text{minimize}} \quad & \max_{\theta} \left[f(\theta) - \lambda^T g(\theta) - \nu^T h(\theta) \right] && (2.22)\\ \text{subject to} \quad & \lambda \succeq 0. \end{aligned}$$

It turns out that Theorem 2.14 can also be proved from the dual form (2.22). Similar to Section 2.2, we will only prove for the case of $p = 1$ for notational convenience. Before proceeding to the proof, we present the following lemma that will be used later.

Lemma 2.15. *Given a set of real-valued functions $\{f^{(k)}\}_{k=1}^K$, the optimal value of the optimization problem*

$$\begin{aligned} \underset{\{p_k,\theta_k\}_{k=1}^K}{\text{maximize}} \quad & \sum_{k=1}^K p_k f^{(k)}(\theta_k) && (2.23)\\ \text{subject to} \quad & \sum_{k=1}^K p_k = 1, \qquad p_k \ge 0, \quad k = 1, 2, \ldots, K \end{aligned}$$

is $\max_\theta \max_{k=1,2,\ldots,K} \{f^{(k)}(\theta)\}$.

Proof. Denote the optimal value of problem (2.23) as OPT and

$$\theta^* = \arg\max_{\theta} \max_{k=1,2,\ldots,K} \{f^{(k)}(\theta)\}, \qquad k^* = \arg \max_{k=1,2,\ldots,K} \{f^{(k)}(\theta^*)\}.$$

Then we have OPT $\geq f^{(k^*)}(\theta^*) = \max_\theta \max_{k=1,2,\ldots,K}\{f^{(k)}(\theta)\}$, since

$$p_k = \begin{cases} 1 & k = k^* \\ 0 & \text{otherwise,} \end{cases} \qquad \theta_k = \theta^*, \quad \forall k$$

is a feasible solution of problem (2.23), and its corresponding objective value is $f^{(k^*)}(\theta^*)$. On the other hand, suppose $\{p_k^*, \theta_k^*\}_{k=1}^K$ is the optimal solution of problem (2.23). Then we have

$$\begin{aligned} \text{OPT} &= \sum_{k=1}^K p_k^* f^{(k)}(\theta_k^*) \\ &\leq \sum_{k=1}^K \left[p_k^* \max_{k=1,2,\ldots,K} \left\{ f^{(k)}(\theta_k^*) \right\} \right] \\ &= \left(\sum_{k=1}^K p_k^* \right) \cdot \max_{k=1,2,\ldots,K} \left\{ f^{(k)}(\theta_k^*) \right\} \\ &= \max_{k=1,2,\ldots,K} \left\{ f^{(k)}(\theta_k^*) \right\} \\ &\leq \max_{k=1,2,\ldots,K} \left\{ \max_\theta f^{(k)}(\theta) \right\} \\ &= \max_\theta \max_{k=1,2,\ldots,K} \left\{ f^{(k)}(\theta) \right\}. \end{aligned}$$

Therefore, we have OPT $= \max_\theta \max_{k=1,2,\ldots,K}\{f^{(k)}(\theta)\}$. □

We are now ready to prove Theorem 2.14 from the dual form.

Proof. (Theorem 2.14) For convenience, we define the objective function in (2.22) as

$$L(\lambda, \nu) = \max_\theta \left[f(\theta) - \lambda g(\theta) - \nu^T h(\theta) \right],$$

where λ is now reduced to a scalar because $p = 1$. Recall that

$$f(\theta) = \max_{k=1,2,\ldots K} f^{(k)}(\theta), \qquad g(\theta) = \min_{l=1,2,\ldots L} g^{(l)}(\theta).$$

Because $\lambda \geq 0$, we have

$$L(\lambda, \nu) = \max_\theta \max_{k,l} \left\{ f^{(k)}(\theta) - \lambda g^{(l)}(\theta) - \nu^T h(\theta) \right\},$$

and, by Lemma 2.15,

$$L(\lambda,\nu) = \max_{\{p_{kl},\theta_{kl}\}_{k,l}} \sum_{k,l} p_{kl}\left[f^{(k)}(\theta_{kl}) - \lambda g^{(l)}(\theta_{kl}) - \nu^T h(\theta_{kl})\right],$$

where $\{p_{kl}\}$ need to satisfy $\sum_{k,l} p_{kl} = 1$ and $p_{kl} \geq 0$ for all k and l. Similar to the previous proof in Section 2.2, we introduce new variables $\gamma_{kl} = p_{kl}\theta_{kl}$, so that

$$L(\lambda,\nu) = \max_{\{p_{kl},\gamma_{kl}\}_{k,l}} \sum_{k,l} \left[p_{kl} f^{(k)}(\gamma_{kl}/p_{kl}) - \lambda p_{kl} g^{(l)}(\gamma_{kl}/p_{kl}) - \nu^T p_{kl} h(\gamma_{kl}/p_{kl})\right].$$

Next, because $f^{(k)}$ is concave and $g^{(l)}$ is convex for all k and l, and h is affine, if problem (2.22) is feasible, then the optimal solution is a saddle point of

$$\sum_{k,l} \left[p_{kl} f^{(k)}(\gamma_{kl}/p_{kl}) - \lambda p_{kl} g^{(l)}(\gamma_{kl}/p_{kl}) - \nu^T p_{kl} h(\gamma_{kl}/p_{kl})\right].$$

Therefore, problem (2.22) achieves the same optimal value as the following problem obtained by exchanging the order of maximizing and minimizing:

$$\begin{array}{ll} \underset{\{p_{kl},\gamma_{kl}\}}{\text{maximize}} & \min\limits_{\lambda\geq 0,\nu} \sum\limits_{k,l} \left[p_{kl} f^{(k)}(\gamma_{kl}/p_{kl}) - \lambda p_{kl} g^{(l)}(\gamma_{kl}/p_{kl}) - \nu^T p_{kl} h(\gamma_{kl}/p_{kl})\right] \\ \text{subject to} & \sum\limits_{k,l} p_{kl} = 1, \qquad p_{kl} \geq 0, \quad \forall k,l. \end{array} \tag{2.24}$$

Using the fact that

$$\min_{\lambda\geq 0,\nu} \sum_{k,l} \left[p_{kl} f^{(k)}(\gamma_{kl}/p_{kl}) - \lambda p_{kl} g^{(l)}(\gamma_{kl}/p_{kl}) - \nu^T p_{kl} h(\gamma_{kl}/p_{kl})\right]$$
$$= \begin{cases} \sum_{k,l} p_{kl} f^{(k)}(\gamma_{kl}/p_{kl}) & \sum_{k,l} p_{kl} g^{(l)}(\gamma_{kl}/p_{kl}) \leq 0 \text{ and } \sum_{k,l} p_{kl} h(\gamma_{kl}/p_{kl}) = 0 \\ -\infty & \text{otherwise,} \end{cases}$$

we can further rewrite problem (2.24) as the problem in Theorem 2.14. □

From the semi-infinite dual form (2.20), it is also straightforward to discover another case that permits a convex reformulation. This happens when $\theta \in \mathbb{R}$, and the reformulated problem corresponds to a sum-of-squares (SOS) optimization problem. More details on SOS optimization can be found in, e.g, [54].

Theorem 2.16. *If $\theta \in \mathbb{R}$, both g and h are polynomials in θ, and $f = \max_{k=1,2,\ldots,K} f^{(k)}(\theta)$ where each $f^{(k)}$ is a polynomial in θ, then the problem*

$$\begin{aligned} \underset{\lambda,\nu,\mu}{\text{minimize}} \quad & \mu && (2.25)\\ \text{subject to} \quad & -f^{(k)}(\theta) + \lambda^T g(\theta) + \nu^T h(\theta) + \mu \text{ is SOS in } \theta, \quad k = 1, 2, \ldots, K && (2.26)\\ & \lambda \succeq 0 \end{aligned}$$

achieves the same optimal value as problem (2.20).

Proof. Use the fact that a univariate polynomial is nonnegative if and only if it can be written as a sum of squares (cf. [14]). □

The new problem (2.25) is an SOS optimization problem, which can be converted into a semidefinite program (hence convex).

2.4 A simple example on the Gaussian distribution

The theory presented in the previous sections only requires that the function g in the constraint should be piecewise convex. Due to its greater flexibility in incorporating knowledge about the distribution, the new formulation is expected to provide better quantification results than previous approaches that only incorporates information on the moments, e.g., by Bertsimas and Popescu [10]. It is difficult to quantify exactly the improvement given by the new formulation since the answer will depend on the objective f, the constraint g, and the true (but unknown) probability distribution. In this section, we will present a case-specific comparison using a simple example.

In the moment-based formulation, the constraint g is a vector consisting of powers of the random variable θ, i.e.,

$$g(\theta) = \begin{bmatrix} \theta^{p_1} & \theta^{p_2} & \ldots & \theta^{p_m} \end{bmatrix} \tag{2.27}$$

for certain integers $p_1, p_2, \ldots, p_m$. In our new formulation, we are allowed to use any piecewise concave functions. Therefore, one way to compare the two formulations is to add our piecewise concave function of choice, denoted as g_{pwc}, into g, so that the the function in

the inequality constraint becomes $[\ g(\theta) \quad g_{\text{pwc}}(\theta)\ ]$, and ask how much improvement can be obtained on the optimal value of the optimal uncertainty quantification problem (2.1) for some objective f. However, we may lose convexity when solving the new optimal uncertainty quantification problem because not all power functions are (piecewise) convex.

Instead, the approach used in this section is as follows. First choose a probability distribution $\mathcal{D}^*$ from which the information constraints will be generated and revealed to the optimal uncertainty quantification algorithm. Second, choose a *single* function that is piecewise affine and convex (therefore it is both piecewise convex and piecewise concave) and compute $\mathbb{E}_{\theta\sim\mathcal{D}^*}[g_{\text{pwc}}(\theta)]$. Finally, solve the optimization problem over the probability distribution $\mathcal{D}$

$$\begin{aligned} \underset{\mathcal{D}}{\text{maximize}} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}[g_{\text{pwc}}(\theta)] \\ \text{subject to} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}[g(\theta)] = \hat{g}, \end{aligned} \tag{2.28}$$

where g only incorporates moment information (i.e., in the form (2.27)). Note that the notation used here is slightly different than that in the original optimal uncertainty quantification (2.1) in that g is used in the equality constraints. If the optimal value of problem (2.28) is very close to the true value $\mathbb{E}_{\theta\sim\mathcal{D}^*}[g_{\text{pwc}}(\theta)]$, then it implies that g_{pwc} will not provide much additional information over g when used as an information constraint for any optimal uncertainty quantification problem.

In our example, the distribution is the standard Gaussian distribution $\mathcal{N}(0,1)$ and the piecewise affine function is $g_{\text{pwc}}(\theta) = |\theta|$. These choices are rather arbitrary and the sole purpose is to show that there is indeed a noticeable difference between the new formulation and previous ones. First we compute the true value $\mathbb{E}_{\theta\sim\mathcal{D}^*}[g_{\text{pwc}}(\theta)]$. In this case, it can be computed analytically as

$$\mathbb{E}_{\theta\sim\mathcal{D}^*}[g_{\text{pwc}}(\theta)] = 2\cdot\frac{1}{\sqrt{2\pi}}\int_0^\infty \theta e^{-\frac{\theta^2}{2}}\,d\theta = \sqrt{\frac{2}{\pi}} \approx 0.799.$$

The odd-order moments of $\mathcal{N}(0,1)$ are all zero and several even-order moments of $\mathcal{N}(0,1)$ are listed below:

p	1	2	3	4	5
$\mathbb{E}[\theta^{2p}]$	1	3	15	105	945

m	2	4	6	8	10	12
Optimal $\mathbb{E}[g_{\text{pwc}}(\theta)]$	1.0000	1.0000	0.8881	0.8881	0.8561	0.8561

Table 2.1: Optimal value of problem (2.28) for different m with $g(\theta) = \begin{bmatrix} \theta & \theta^2 & \cdots & \theta^m \end{bmatrix}$. For comparison, the true value of $\mathbb{E}[g_{\text{pwc}}(\theta)] = \sqrt{\frac{2}{\pi}} \approx 0.799$.

The optimal value of problem (2.28) can be obtained via convex optimization by making use of Theorem 2.16. In this case, the objective function $g_{\text{pwc}}(\theta) = |\theta| = \max\{-\theta, \theta\}$ and the corresponding optimization problem is

$$\begin{aligned} \underset{\nu,\mu}{\text{minimize}} \quad & \mu \\ \text{subject to} \quad & \theta + (g(\theta) - \hat{g})^T \nu + \mu \quad \text{is SOS in } \theta \\ & -\theta + (g(\theta) - \hat{g})^T \nu + \mu \quad \text{is SOS in } \theta. \end{aligned}$$

From the above optimization problem, we can see that odd-order moments do not affect the optimal value and therefore we can restrict the highest order in $g(\theta)$ to be even in problem (2.28). In particular, we let

$$g(\theta) = \begin{bmatrix} \theta & \theta^2 & \cdots & \theta^m \end{bmatrix}$$

and compute the optimal value for different choices of m. It can be seen from Table 2.1 that, even with the information of up to the 12th moment incorporated, the optimal bound is still somewhat far away from the true value $\mathbb{E}_{\theta\sim\mathcal{D}^*}[g_{\text{pwc}}(\theta)]$. Therefore, we can expect that incorporating constraint such as $g_{\text{pwc}}(\theta) = |\theta|$ and, more generally, any piecewise convex functions will provide additional benefits over solely moment constraints when solving optimal uncertainty quantification problems.

2.5 Piecewise affine objective with first and second moment constraints

In this section, we will focus on an important class of OUQ problems that fall into the form appeared in Theorem 2.14. In particular, we require that: (1) the function f should be piecewise affine and convex; (2) the constraints should only consist of first and second

moments, i.e., the problem has the form:

$$\begin{aligned} \underset{\mathcal{D}}{\text{maximize}} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}\left[f(\theta)\right] \\ \text{subject to} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}[\theta] = \hat{\mu}, \quad \text{cov}_{\theta\sim\mathcal{D}}[\theta] \preceq \widehat{\Sigma}, \end{aligned} \tag{2.29}$$

where $f(\theta) = \max_{k=1,2,\ldots,K}(a_k^T\theta + b_k)$ for some $a_k, b_k \in \mathbb{R}^n$ $(k = 1, 2, \ldots, K)$. The inequality in the constraint is the usual partial ordering on matrices: for any $A, B \in \mathbb{R}^{n\times n}$, we have $A \preceq B$ if and only if $B - A$ is positive semidefinite. By introducing the function

$$g_v(\theta) = v^T(\theta\theta^T - \hat{\mu}^T\hat{\mu} - \widehat{\Sigma})v,$$

where v is any given vector in $\mathbb{R}^n$, we are able to convert the constraint $\text{cov}_{\theta\sim\mathcal{D}}[\theta] \preceq \widehat{\Sigma}$ to an infinite number of constraints parameterized by v, so that problem (2.29) becomes

$$\begin{aligned} \underset{\mathcal{D}}{\text{maximize}} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}\left[f(\theta)\right] \\ \text{subject to} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}[\theta] = \hat{\mu}, \\ & \mathbb{E}_{\theta}[g_v(\theta)] \leq 0, \quad v \in \mathbb{R}^n. \end{aligned} \tag{2.30}$$

Since $g_v(\theta)$ is convex (hence piecewise convex) in θ, problem (2.30) satisfies the requirements in Theorem 2.14. Therefore, we can derive the corresponding convex optimization problem by making use of Theorem 2.14. It should be mentioned that this form has also been extensively studied by, e.g., Delage and Ye [19], but is derived in a different way as given by Theorem 2.14.

From the definition of f, we know that the total number of Dirac masses is at most K and therefore problem (2.29) can be reformulated as:

$$\begin{aligned} \underset{\{p_k,\gamma_k\}}{\text{maximize}} \quad & \sum_{k=1}^{K}(a_k^T\gamma_k + b_k p_k) \\ \text{subject to} \quad & \sum_{k=1}^{K} p_k = 1, \qquad p_k \geq 0, \quad k = 1, 2, \ldots, K \\ & \sum_{k=1}^{K}\gamma_k = \hat{\mu} \end{aligned}$$

$$\sum_{k=1}^{K} p_k g_v(\gamma_k/p_k) \leq 0, \quad v \in \mathbb{R}^n.$$

The last constraint, which is equivalent to

$$v^T \left[\sum_{k=1}^{K} \gamma_k \gamma_k^T/p_k - (\hat{\mu}\hat{\mu}^T + \widehat{\Sigma}) \right] v \leq 0, \quad v \in \mathbb{R}^n,$$

can be rewritten back using matrix inequalities as

$$\sum_{k=1}^{K} \gamma_k \gamma_k^T/p_k \preceq \hat{\mu}\hat{\mu}^T + \widehat{\Sigma}. \tag{2.31}$$

Finally, we introduce slack variables Γ_k and rewrite (2.31) as two constraints

$$\sum_{k=1}^{K} \Gamma_k = \hat{\mu}\hat{\mu}^T + \widehat{\Sigma}, \qquad \begin{bmatrix} \Gamma_k & \gamma_k \\ \gamma_k^T & p_k \end{bmatrix} \succeq 0, \quad k = 1, 2, \ldots, K.$$

To summarize, problem (2.29) achieves the same optimal value as the problem

$$\begin{aligned} \underset{\{p_k, \gamma_k, \Gamma_k\}}{\text{maximize}} \quad & \sum_{k=1}^{K} (a_k^T \gamma_k + b_k p_k) \\ \text{subject to} \quad & \sum_{k=1}^{K} p_k = 1, \qquad p_k \geq 0, \quad k = 1, 2, \ldots, K \\ & \sum_{k=1}^{K} \gamma_k = \hat{\mu}, \qquad \sum_{k=1}^{K} \Gamma_k = \hat{\mu}\hat{\mu}^T + \widehat{\Sigma} \\ & \begin{bmatrix} \Gamma_k & \gamma_k \\ \gamma_k^T & p_k \end{bmatrix} \succeq 0, \quad k = 1, 2, \ldots, K. \end{aligned} \tag{2.32}$$

It can be verified that this is a convex optimization problem and, in fact, a semidefinite program. As will be seen later, it is sometimes useful to solve instead the Lagrange dual problem of (2.32), which can be derived by following the standard procedure. Since the semidefinite cone is self-dual, the dual problem is also a semidefinite program:

$$\underset{Q,q,r}{\text{minimize}} \quad \text{tr}((\widehat{\Sigma} + \hat{\mu}\hat{\mu}^T)Q) + \hat{\mu}^T q + r \tag{2.33}$$

$$\text{subject to} \quad \begin{bmatrix} Q & (q-a_k)/2 \\ (q-a_k)^T/2 & r-b_k \end{bmatrix} \succeq 0, \quad k=1,2,\ldots,K, \tag{2.34}$$

where Q is a $n \times n$ symmetric matrix, $q_k \in \mathbb{R}^n$, and $r \in \mathbb{R}$. In fact, the matrix

$$\begin{bmatrix} Q & (q-a_k)/2 \\ (q-a_k)^T/2 & r-b_k \end{bmatrix}$$

is the Lagrange multiplier of

$$\begin{bmatrix} \Gamma_k & \gamma_k \\ \gamma_k^T & p_k \end{bmatrix}$$

in the primal problem (2.32).

For notational convenience, we introduce a index set $\mathcal{K}$ and rewrite the dual problem (2.33) as

$$\underset{Q,q,r}{\text{minimize}} \quad \text{tr}((\widehat{\Sigma}+\hat{\mu}\hat{\mu}^T)Q)+\hat{\mu}^T q + r \tag{2.35}$$

$$\text{subject to} \quad \begin{bmatrix} Q & (q-a_k)/2 \\ (q-a_k)^T/2 & r-b_k \end{bmatrix} \succeq 0, \quad k \in \mathcal{K}. \tag{2.36}$$

For any given index set $\mathcal{K}$ and $\mathcal{C} = \{(a_k, b_k)\}_{k\in\mathcal{K}}$, we denote the optimization problem (2.35) as $\text{COUQ}(\mathcal{K}, \mathcal{C})$ and its optimal value as $\text{COUQ}^*(\mathcal{K}, \mathcal{C})$. The dependence of the problem on $\hat{\mu}$ and $\widehat{\Sigma}$ is omitted. This notation also applies to any subset $\mathcal{A} \subset \mathcal{K}$. That is, $\text{COUQ}(\mathcal{A}, \mathcal{C})$ denotes the optimization problem

$$\underset{Q,q,r}{\text{minimize}} \quad \text{tr}((\widehat{\Sigma}+\hat{\mu}\hat{\mu}^T)Q)+\hat{\mu}^T q + r$$

$$\text{subject to} \quad \begin{bmatrix} Q & (q-a_k)/2 \\ (q-a_k)^T/2 & r-b_k \end{bmatrix} \succeq 0, \quad k \in \mathcal{A}.$$

2.6 Related work

The earliest origin of OUQ, or similar problems under different names, can be traced back to the work on generalization of Chebyshev-type inequalities in the 1950s and 1960s by Isii [32, 33, 34], Mulholland and Rogers [46], Godwin [25], Marshall and Olkin [45], and Olkin and Pratt [51], among others. Aside from the formulation mentioned in the previous

sections, there exist a few other cases for which convex optimization can be applied, as recently shown by other researchers, including Bertsimas and Popescu [10], Popescu [55], Lasserre [40], and Vandenberghe et al. [74]. In the following, we list two most representative formulations.

Polynomial objective with moment constraints [16] (pp. 170) In this formulation, the random variable $\theta \in \mathbb{R}$ is univariate, the objective function f is a polynomial of even order in θ, i.e., $f(\theta) = \sum_{i=1}^{2p} c_i \theta^i$ for some integer p, and the constraints are bounds on the moments of θ:

$$\underline{m}_i \leq \mathbb{E}[\theta^i] \leq \overline{m}_i, \quad i = 1, 2, \ldots, 2p.$$

In this case, problem (2.1) becomes

$$\begin{aligned} \underset{\mathcal{D}}{\text{maximize}} \quad & \mathbb{E}_{\theta \sim \mathcal{D}} \left[\sum_{i=1}^{2p} c_i \theta^i \right] \\ \text{subject to} \quad & \underline{m}_i \leq \mathbb{E}_{\theta \sim \mathcal{D}}[\theta^i] \leq \overline{m}_i, \quad i = 1, 2, \ldots, 2p. \end{aligned} \tag{2.37}$$

Let $x_i = \mathbb{E}[\theta^i]$ $(i = 0, 1, 2, \ldots, 2p)$ be the moments and define a $(2p+1) \times (2p+1)$ Hankel matrix $H(x_0, x_1, \ldots, x_{2p})$ such that $H_{ij} = x_{i+j-2}$. It can be shown that $\{x_i\}_{i=1}^{2p}$ corresponds to the moments of some distribution (or the limit of a sequence of distributions) if and only if $x_0 = 1$ and the Hankel matrix $H(x_0, x_1, \ldots, x_{2p}) \succeq 0$. Therefore, the optimization problem (2.37) can be cast as a semidefinite program

$$\begin{aligned} \underset{\{x_i\}_{i=1}^{2p}}{\text{maximize}} \quad & \sum_{i=1}^{2p} c_i x_i \\ \text{subject to} \quad & \underline{m}_i \leq x_i \leq \overline{m}_i, \quad i = 1, 2, \ldots, 2p \\ & H(1, x_1, \ldots, x_{2p}) \succeq 0. \end{aligned}$$

Probability bound with moment constraints [10] In this formulation, the random variable $\theta \in \mathbb{R}$, the objective function $f = I(\theta \geq a)$ is the 0-1 indicator function for some constant a, and the constraints are hard constraints on the moments of θ:

$$\mathbb{E}[\theta^i] = m_i, \quad i = 1, 2, \ldots, p.$$

In this case, problem (2.1) becomes

$$\begin{aligned} \underset{\mathcal{D}}{\text{maximize}} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}\left[I(\theta \geq a)\right] = \mathbb{P}(\theta \geq a) \\ \text{subject to} \quad & \mathbb{E}_{\theta\sim\mathcal{D}}[\theta^i] = m_i, \quad i = 1, 2, \ldots, p. \end{aligned} \tag{2.38}$$

Bertsimas and Popescu have shown that the optimal value of problem (2.38) can be obtained from the following semidefinite program over $y = \{y_r\}_{r=0}^{p}$ and $X, Z \in \mathbb{S}_+^{p+1}$:

$$\begin{aligned} \underset{y,X,Z}{\text{minimize}} \quad & \sum_{r=0}^{p} m_r y_r \\ \text{subject to} \quad & (y_0 - 1) + \sum_{r=1}^{p} a^r y_r = x_{00} \\ & \sum_{i,j\,:\, i+j=2l-1} x_{ij} = 0, \quad l = 1, 2, \ldots, p \\ & \sum_{i,j\,:\, i+j=2l-1} z_{ij} = 0, \quad l = 1, 2, \ldots, p \\ & (-1)^l \sum_{r=l}^{p} y_r \binom{r}{l} a^{r-l} = \sum_{i,j\,:\, i+j=2l} x_{ij}, \quad l = 0, 1, \ldots, p \\ & \sum_{r=0}^{l} y_r \binom{p-r}{l-r} a^r = \sum_{i,j\,:\, i+j=2l} z_{ij}, \quad l = 0, 1, \ldots, p \\ & X, Z \succeq 0. \end{aligned}$$

Other variants of this, e.g., when $f = I(a \leq \theta \leq b)$ (for some constants a, b) and/or $\theta \in \mathbb{R}^+$, can also be found in [10].

2.7 Conclusions

The chapter begins by introducing the formulation of the OUQ problem as an optimization problem. Although the OUQ problem is infinite-dimensional and not immediately amenable to numerical solution, previous work has shown that it is possible to solve instead an equivalent finite-dimensional formulation that will yield the same optimal value. The main focus of this chapter is to investigate cases in which the equivalent finite-dimensional problem is not only solvable numerically, but can also be solved efficiently.

In particular, we show that convex formulation exists for several cases using two different

approaches: from the primal form and dual form of the optimization problem. From both the primal and the dual form, we show that the OUQ problem adopts a convex formulation if: (1) the objective is piecewise concave, (2) the inequality constraint is piecewise convex, and (3) the equality constraint is affine (Theorem 2.14). Support constraints can be incorporated as well if the support is constrained to be a certain polytope. In the univariate case, we also show from the dual form that the OUQ problem adopts a convex formulation if: (1) the objective is piecewise polynomial and (2) the constraints are polynomial (Theorem 2.16).

A simple example using Gaussian distributions shows that the new cases presented in this chapter allow more freedom in incorporating knowledge about the unknown probability distribution and can potentially provide better quantification. In the end, we apply the theoretical results to a very useful case of piecewise affine objective with first and second moment constraints, whose computational aspect and applications will be discussed in detail in the next chapter.

Chapter 3

OUQ via Convex Optimization: Computational Issues and Applications

Despite the fact that there exists a class of optimal uncertainty quantification problems where convex optimization can be applied, as seen from the previous chapter, there can still be computational issues when the scale of the given problem becomes large. This chapter addresses some of these computational issues. We restrict our discussion to problems with piecewise affine objective and first and second moment constraints (introduced in Section 2.5). There are two important measures on the scale of the problem: the number of required Dirac masses and the dimension of the random variable. In the following, two different ways of addressing the issues of scaling will be presented. One focuses on solving many smaller problems to obtain the solution iteratively; another focuses on solving the large original problem through massive parallelization. In the end, these efforts are demonstrated by an application of energy storage placement evaluation in power grids.

Material from Section 3.1 and 3.3 has also been published in [29].

3.1 Iterative methods for polytopic canonical form

In this section, we generalize the piecewise affine objective function in the OUQ problem from Section 2.5 to what will later be called the polytopic canonical form, which is equivalent to piecewise affine functions except that the number of affine functions can potentially be extremely large. Exact methods for solving problems in this form can be prohibitively expensive. In order to partially alleviate this difficulty, we propose an iterative approximate

method that only requires solving smaller problems at each iteration. The method is guaranteed to converge, and it often converges close to the true optimum for problems we have tested.

3.1.1 The polytopic canonical form (PCF): Motivation and definition

In many applications, the objective function f in the OUQ problem is defined as the optimal value of another optimization problem. This can happen, for example, in two-stage stochastic programming problems, where there exist two decisions made at different time instances. Formally, a two-stage stochastic programming problem is the one that can be expressed in the following form for some cost function J:

$$\underset{u_1 \in \mathcal{U}_1}{\text{minimize}} \quad \mathbb{E}_\theta[\min_{u_2 \in \mathcal{U}_2(\theta)} J(u_1, u_2, \theta)].$$

For example, in the energy storage placement problem presented later in this chapter, we need to deal with the problem of deciding the assignment of the amount of energy storage at different nodes in a power grid. However, the corresponding objective function, in this case the total savings of generation, will also depend on the particular choice of power flow during the day *after* the assignment of storage is chosen. In this example, the assignment of energy storage corresponds to u_1, the decision that happens earlier in time, whereas the choice of power flow corresponds to u_2, the decision that happens later. The decision u_2 is sometimes called the *recourse action* and such problems are also referred to as stochastic programming problems *with recourse* [66].

Throughout this chapter, we will not attempt to solve the two-stage stochastic programming problem, but will rather focus on quantifying

$$\mathbb{E}_\theta[\min_{u_2 \in \mathcal{U}_2(\theta)} J(u_1, u_2, \theta)]$$

for given u_1, i.e., the objective function f in the OUQ problem is defined as

$$f(\theta) = \min_{u_2 \in \mathcal{U}_2(\theta)} J(u_1, u_2, \theta).$$

Sometimes the dependence of u_1 in J (and consequently in f) is dropped when it is clear from the context. In particular, we will consider the case where the minimization of J over u_2

corresponds to a linear program. This implies that J is linear in u_2 and the constraint set $\mathcal{U}_2$ is a polytope. In particular, we restrict the linear program to the following form:

$$\begin{aligned} \underset{u_2}{\text{minimize}} \quad & c^T u_2 \\ \text{subject to} \quad & Au_2 + B\theta = d \\ & u_2 \succeq 0. \end{aligned}$$

This is the same as the canonical form for a linear program, except that the equality constraint depends also on θ. Its Lagrange dual problem is

$$\begin{aligned} \underset{\nu}{\text{maximize}} \quad & \nu^T d - \nu^T B\theta \\ \text{subject to} \quad & A^T \nu \preceq c. \end{aligned}$$

If the primal problem is feasible, then strong duality holds [16], which implies that f can also be defined by the optimal value of the dual problem:

$$f(\theta) = \max_{\nu:\ A^T\nu \preceq c} (\nu^T d - \nu^T B\theta).$$

Note that the function inside the maximum is affine in θ and the coefficients belong to a certain polytope. This motivates a more general definition for this type of f called the *polytopic canonical form*.

Definition 3.1 (Polytopic canonical form)**.** A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be in the *polytopic canonical form* (PCF) if it can be written as

$$f(\theta) = \max_{(a,b)\in\mathcal{P}} \{a^T\theta + b\}, \quad a \in \mathbb{R}^n, b \in \mathbb{R} \tag{3.1}$$

for some polytope $\mathcal{P}$ of dimension $(n+1)$.

Under such definition, f can be regarded as the optimal value of a family of linear programs (LP) parameterized by θ:

$$\underset{a,b}{\text{maximize}} \quad a^T\theta + b \qquad \text{subject to} \quad (a,b) \in \mathcal{P}. \tag{3.2}$$

The PCF (3.1) subsumes the piecewise affine form

$$f(\theta) = \max_{(a_k,b_k)\in\mathcal{C}} \{a_k^T\theta + b_k\} \qquad \text{for some } \mathcal{C} = \{(a_k, b_k)\}_{k\in\mathcal{K}}. \tag{3.3}$$

For any f in the form (3.3) with $\mathcal{C} = \{(a_k, b_k)\}_{k\in\mathcal{K}}$, we can choose $\mathcal{P}$ to be the convex hull of $\mathcal{C}$. This implies $\mathcal{C} \subset \mathcal{P}$, and hence

$$f(\theta) = \max_{(a_k,b_k)\in\mathcal{C}} \{a_k^T\theta + b_k\} \leq \max_{(a,b)\in\mathcal{P}} \{a^T\theta + b\}. \tag{3.4}$$

The last inequality is always an equality, which can be shown by using a basic property of linear programs as follows. Denote the vertices (extreme points) of $\mathcal{P}$ as $\mathcal{V}$. We have $\mathcal{V} \subseteq \mathcal{C}$, hence

$$\max_{(a_k,b_k)\in\mathcal{V}} \{a_k^T\theta + b_k\} \leq \max_{(a_k,b_k)\in\mathcal{C}} \{a_k^T\theta + b_k\}. \tag{3.5}$$

From the optimality of the extreme points, we know that any optimum for the linear program (3.2) can always be attained at some $(a_k, b_k) \in \mathcal{V}$, no matter how θ is chosen (cf. [11]), i.e.,

$$\max_{(a,b)\in\mathcal{P}} \{a^T\theta + b\} = \max_{(a_k,b_k)\in\mathcal{V}} \{a_k^T\theta + b_k\}. \tag{3.6}$$

Therefore, from (3.4)–(3.6), the equality

$$\max_{(a,b)\in\mathcal{P}} \{a^T\theta + b\} = \max_{(a_k,b_k)\in\mathcal{C}} \{a_k^T\theta + b_k\}$$

must hold and $f(\theta) = \max_{(a,b)\in\mathcal{P}}\{a^T\theta + b\}$, i.e., any f in the form (3.3) can be rewritten in PCF. On the other hand, given any function f in PCF, we can also rewrite it in the form (3.3) by setting $\mathcal{C}$ as the vertices of $\mathcal{P}$. The benefit of using PCF is its flexibility. In PCF, $\mathcal{P}$ can be defined either by its vertices, in which case it reduces to the form (3.3), or by the intersection of half-spaces. The latter representation can sometimes be more compact, e.g., for the storage placement problem in Section 3.3.

3.1.2 Exact iterative method method for PCF

For any OUQ problems in which f is in PCF, there is at least one practical issue in directly solving the corresponding convex optimization problem (2.32) or its dual problem (2.33) after rewriting f in the form (3.3). Obtaining the vertices $\mathcal{V}$, usually through vertex enumeration

algorithms (e.g. [8]), can be computationally demanding when the dimension of $\mathcal{P}$ is high or the number of its composing constraints is large. In general, the cardinality of $\mathcal{V}$, denoted as $|\mathcal{V}|$, grows exponentially with the dimension of $\mathcal{P}$. This becomes prohibitively expensive even for a moderate dimension and a moderate number of constraints. Even if $\mathcal{V}$ could be obtained, solving the SDP (2.35) would also be expensive when $|\mathcal{V}|$ (hence $|\mathcal{K}|$) is large.

To this end, we seek iterative methods that solve a smaller problem at each iteration. Recall the definition of $\mathrm{COUQ}(\mathcal{K}, \mathcal{V})$ for a certain index set $\mathcal{K}$ and coefficients $\mathcal{V} = \{(a_k, b_k)\}_{k\in\mathcal{K}}$:

$$\begin{aligned} \underset{Q,q,r}{\text{minimize}} \quad & \mathrm{tr}((\widehat{\Sigma} + \hat{\mu}\hat{\mu}^T)Q) + \hat{\mu}^T q + r \\ \text{subject to} \quad & \begin{bmatrix} Q & (q-a_k)/2 \\ (q-a_k)^T/2 & r - b_k \end{bmatrix} \succeq 0, \quad k \in \mathcal{K}. \end{aligned}$$

In general, if we choose an arbitrary subset $\mathcal{A} \subset \mathcal{K}$ and solve the problem $\mathrm{COUQ}(\mathcal{A}, \mathcal{V})$ to obtain its optimal value $\mathrm{COUQ}^*(\mathcal{A}, \mathcal{V})$, we are only guaranteed to obtain a lower bound for the optimal value of the original problem, i.e., we have $\mathrm{COUQ}^*(\mathcal{A}, \mathcal{V}) \leq \mathrm{COUQ}^*(\mathcal{K}, \mathcal{V})$ since the constraints corresponding to $k \in \mathcal{K}\backslash\mathcal{A}$ have been ignored. The inequality is tight if and only if the optimal solution (Q^*, q^*, r^*) for $\mathrm{COUQ}(\mathcal{A}, \mathcal{V})$ also satisfies the constraints for $k \in \mathcal{K}\backslash\mathcal{A}$, i.e.,

$$\begin{bmatrix} Q^* & (q^*-a_k)/2 \\ (q^*-a_k)^T/2 & r^* - b_k \end{bmatrix} \succeq 0, \quad \forall k \in \mathcal{K}\backslash\mathcal{A}. \tag{3.7}$$

Based on this fact, one can use the following procedure to obtain $\mathrm{COUQ}^*(\mathcal{K}, \mathcal{V})$, without including all the constraints in $\mathcal{K}$ in the optimization problem at first. When the procedure finishes, the optimal solution satisfies condition (3.7) and hence the corresponding optimal value is $\mathrm{COUQ}^*(\mathcal{K}, \mathcal{V})$.

1. Start with an initial index set $\mathcal{A} \subset \mathcal{K}$.
2. Obtain (Q^*, q^*, r^*) for the problem $\mathrm{COUQ}(\mathcal{A}, \mathcal{V})$.
3. If (Q^*, q^*, r^*) satisfies (3.7), report (Q^*, q^*, r^*) as the solution to $\mathrm{COUQ}(\mathcal{K}, \mathcal{V})$ and terminate. Otherwise, there must exist a set $\mathcal{B} \subset \mathcal{K}\backslash\mathcal{A}$ such that the condition (3.7) is violated for $k \in \mathcal{B}$. Set $\mathcal{A} := \mathcal{A} \cup \mathcal{B}$ and repeat steps 2–3.

3.1.3 Approximate iterative method for PCF

There are two issues with the procedure presented in the previous section. One issue is that checking the condition (3.7) can be difficult, because the number of constraints to be checked is $|\mathcal{K}| - |\mathcal{A}|$ and is usually extremely large (on the same order as $|\mathcal{K}|$ assuming $|\mathcal{A}|$ is small). The other issue is that, in the worst case, the index set $\mathcal{A}$ may continue to grow until $\mathcal{A} = \mathcal{K}$ so that the final problem to solve has the same complexity as the original problem.

Fortunately, when f can be expressed in PCF, we have a theorem that finds a violating constraint in step 3 without exhaustively checking all the constraints in $\mathcal{K}\backslash\mathcal{A}$. Moreover, Corollary 3.5 will show that, once such a constraint is found, it can replace an existing constraint in $\mathcal{A}$ without affecting convergence of the method. This prevents $\mathcal{A}$ from growing and avoids the possibility of solving a problem as large as $\mathcal{A} = \mathcal{K}$. This method of finding a violating constraint uses an important property of the solution to the problem $\mathrm{COUQ}(\mathcal{A}, \mathcal{V})$. Recall from Section 2.5 that, when we obtain the optimal solution (Q^*, q^*, r^*) to $\mathrm{COUQ}(\mathcal{A}, \mathcal{V})$, we will also automatically obtain the corresponding optimal solution $\{p_k^*, \gamma_k^*\}_{k\in\mathcal{A}}$ to the primal problem, from which the optimal (discrete) probability distribution $\mathcal{D}^*$ of the OUQ problem can be computed: for every (p_k^*, γ_k^*), there is a Dirac mass located at $\theta_k^* = \gamma_k^*/p_k^*$ with probability p_k^*.

Remark 3.2. Recall from the finite reduction property (Theorem 2.4) that the number of Dirac masses required for realizing $\mathcal{D}^*$ is at most the number of independent scalar equalities in the constraint plus 1. In the case of problem (2.29), the number of independent scalar equalities is $N = n + n(n+1)/2$ (the factor $1/2$ is due to the symmetry of $\widehat{\Sigma}$). Therefore, we know the maximum number of required Dirac masses is $\min(|\mathcal{K}|, N+1)$. In practice, depending on the problem, the actual number of nonzero Dirac masses can be even smaller than $\min(|\mathcal{K}|, N+1)$.

The primal problem gives us another way to compute $\mathrm{COUQ}^*(\mathcal{A}, \mathcal{V})$, i.e.,

$$\mathrm{COUQ}^*(\mathcal{A}, \mathcal{V}) = \sum_{k\in\mathcal{A}} p_k^*(a_k^T\theta_k^* + b_k).$$

By using this alternative expression, Theorem 3.3 shows that the locations of the Dirac masses $\{\theta_k^*\}_{k\in\mathcal{A}}$ corresponding to a suboptimal solution (Q^*, q^*, r^*) can be used for finding a violating constraint in $\mathcal{K}\backslash\mathcal{A}$.

Theorem 3.3. *For a given set $\mathcal{A}$, suppose (Q^*, q^*, r^*) is the optimal solution for* $\mathrm{COUQ}(\mathcal{A}, \mathcal{V})$ *and the set of Dirac masses of the optimal distribution is $\{\theta_k^*\}_{k\in\mathcal{A}}$. If for any $u \in \mathcal{A}$, there exists some $v \in \mathcal{K}$ such that*

$$a_v^T\theta_u^* + b_v > a_u^T\theta_u^* + b_u, \tag{3.8}$$

then the constraint

$$\begin{bmatrix} Q^* & (q^* - a_v)/2 \\ (q^* - a_v)^T/2 & r^* - b_v \end{bmatrix} \succeq 0 \tag{3.9}$$

is violated.

Proof. We prove the theorem by contradiction. Consider the problem $\mathrm{COUQ}(\mathcal{A} \cup \{v\}, \mathcal{V})$. Suppose the condition (3.9) is not violated, then (Q^*, q^*, r^*) would also be the optimal solution for $\mathrm{COUQ}(\mathcal{A} \cup \{v\}, \mathcal{V})$, which implies that $\mathrm{COUQ}^*(\mathcal{A} \cup \{v\}, \mathcal{V})$ is

$$\sum_{k\in\mathcal{A}} p_k^*(a_k^T\theta_k^* + b_k), \tag{3.10}$$

for $f(\theta) = \max_{k\in\mathcal{A}\cup\{v\}}\{a_k^T\theta + b_k\}$. On the other hand, $\mathrm{COUQ}^*(\mathcal{A} \cup \{v\}, \mathcal{V})$ should be at least

$$p_u^*(a_v^T\theta_u^* + b_v) + \sum_{k\in\mathcal{A}\setminus\{u\}} p_k^*(a_k^T\theta_k^* + b_k), \tag{3.11}$$

which is attained under the same discrete distribution consisting of $\{(\theta_k^*, p_k^*)\}_{k\in\mathcal{A}}$. The quantity (3.11) will always be greater than (3.10), hence a contradiction. □

Remark 3.4. Condition (3.8) is only sufficient. Hence, it is not guaranteed to find all the violating constraints.

If f is in PCF, such (a_v, b_v) for any given θ_u can be found by solving the LP

$$\underset{a,b}{\text{maximize}} \quad a^T\theta_u + b \qquad \text{subject to} \quad (a, b) \in \mathcal{P}.$$

If the optimal solution (a^*, b^*) for this LP satisfies

$$(a^*)^T\theta_u + b^* > a_u^T\theta_u + b_u,$$

then we have successfully found $(a_v, b_v) = (a^*, b^*)$. Otherwise, no such (a_v, b_v) exists. Another useful by-product of this new way of finding a violating constraint is that the

constraint corresponding to u can be removed from $\mathcal{A}$ in the next iteration while still ensuring that $\text{COUQ}^*(\mathcal{A}, \mathcal{V})$ is increasing.

Corollary 3.5. *For $\mathcal{A}$, $\mathcal{V}$, u and v defined in Theorem 3.3, let $\mathcal{A}'(u,v) = (\mathcal{A}\backslash\{u\}) \cup \{v\}$. Then*

$$\text{COUQ}^*(\mathcal{A}'(u,v), \mathcal{V}) > \text{COUQ}^*(\mathcal{A}, \mathcal{V}).$$

Proof. For $\{\theta_k\}_{k\in\mathcal{A}}$ in the proof of Theorem 3.3,

$$\begin{aligned}\text{COUQ}^*(\mathcal{A}'(u,v), \mathcal{V}) \geq\ & p_u(a_v^T\theta_u + b_v) \\ & + \sum_{k\in\mathcal{A}\backslash\{u\}} p_k(a_k^T\theta_k + b_k).\end{aligned}$$

The proof of Theorem 3.3 has shown that the right hand side is strictly greater than

$$\sum_{k\in\mathcal{A}} p_k(a_k^T\theta_k + b_k) = \text{COUQ}^*(\mathcal{A}, \mathcal{V}),$$

which completes the proof. □

Due to Corollary 3.5, we can use a modified iterative method than the one proposed at the beginning of this section. In particular, Step 3 can be changed to:

3') Obtain $\{\theta_k\}_{k\in\mathcal{A}}$ and check if for any $u \in \mathcal{A}$, there exists $v \in \mathcal{K}$ such that (a_v, b_v) satisfies (3.8). If not, then report (Q^*, q^*, r^*) as the optimal solution to the problem $\text{COUQ}(\mathcal{K}, \mathcal{V})$ and terminate. Otherwise, for every (u, v) satisfying (3.8), set $\mathcal{A} := \mathcal{A}'(u,v)$ and repeat steps 2 and 3'.

This approximate method is guaranteed to converge. At each iteration, the new index set $\mathcal{A}$ will give a non-decreasing optimal value for the corresponding optimization problem. Therefore, this sequence of optimal values is monotone and, at the same time, must be bounded by $\text{COUQ}^*(\mathcal{K}, \mathcal{V})$. By the monotone convergence theorem [18], this sequence, consisting of real numbers, must have a limit, i.e., the method converges. This method is not, in general, guaranteed to converge to the true optimum since there may still be violating constraints when the algorithm exits (see Remark 3.4). However, the result will always be a lower bound of the true optimal value, since some constraints in $\mathcal{K}$ have been removed from the minimization problem $\text{COUQ}(\mathcal{K}, \mathcal{V})$. Therefore, we can run the same optimization

problem multiple times with different initial assignments of $\mathcal{A}$ and choose the highest among all the results to get an improved approximation.

Choosing the size of $\mathcal{A}$ can be potentially important for this method to work properly, since $|\mathcal{A}|$ remains constant over iterations. If its size is too small, $\mathcal{A}$ may not be capable of including all the Dirac masses necessary for realizing the optimal distribution. One possible choice of $|\mathcal{A}|$ is the maximum number of necessary Dirac masses, although this can be conservative for a particular problem (see Remark 3.2). It remains an open question whether knowing such conservatism *a priori* can help speed up the optimization procedure.

We now use a simple example to test this approximate method on small problems. In these examples, we arbitrarily generate $\hat{\mu} \in \mathbb{R}^n$, $\widehat{\Sigma} \in \mathbb{S}^n_+$, and choose $\mathcal{P}$ as the $(n+1)$-dimensional hypercube

$$\{(a, b) : \mathbf{0} \preceq a \preceq \mathbf{1},\, 0 \leq b \leq 1\},$$

where $\mathbf{1}$ and $\mathbf{0}$ denote vectors in $\mathbb{R}^n$ containing all ones and all zeros, respectively. For each n, we compare the relative error between the exact solution from (2.35) and the approximate solution. When computing the approximate solution, we choose $|\mathcal{A}|$ to be the maximum number of necessary Dirac masses. Fig. 3.1 shows the results for n from 1 to 16. The choice of n is limited by the computational time of the exact method (for $n = 16$, it takes about 18.6 hours on an Intel Xeon 3.00 GHz workstation). To obtain statistics about the approximate method, we perform 100 trials for each n, and compute the 10% and 90% quantile of the errors. It can been seen that most of the errors are within 5%.

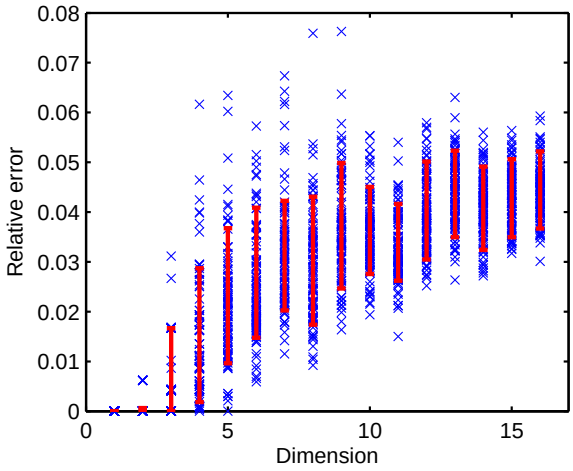


Figure 3.1: Relative errors of the approximate method. Blue crosses: relative errors. Red bars: 10% and 90% quantiles of the relative errors.

3.2 Parallel solution via alternating direction method of multipliers (ADMM)

This section continues the effort on scaling up the optimization problem (2.32) and/or (2.33) for the optimal uncertainty quantification problem presented in Section 2.5. Rather than exploiting the structure in the cost function (i.e, polytopic canonical form), this section will focus on the structure in the optimization problem itself and attempts to devise algorithms that are inherently parallelizable in order to mitigate the scaling issues.

Aside from the issue of having a large number of Dirac masses (e.g., large $\mathcal{K}$), the dimension of the random variable can also be problematic as it becomes large. As mentioned earlier, both the optimization problem (2.32) and its dual (2.33) are semidefinite programs. Many optimization solvers, both commercial and free (e.g., SeDuMi [70] and SDPT3 [71, 73]), use second-order methods such as Newton's method or its variants and their complexity for each iteration grows as $O(n^6)$ if no structure in the problem is exploited, where n is the dimension of the semidefinite matrix. This complexity has to do with the fact that an $n \times n$ matrix has $O(n^2)$ variables and the complexity of an iteration in Newton's method is cubic in the number of variables since it requires solving a system of linear equations in order to obtain the descent direction [50]. Moreover, there can potentially be insufficient memory as the number of variable grows if the linear equations are stored in a naive form (as done by many optimization solvers).

To this end, this section investigates how first-order methods can be applied in this case. First-order methods, despite their slower convergence rate, are less expensive per iteration and can be parallelized more easily than second-order methods. In particular, the alternating direction method of multipliers (ADMM) is chosen due to its mild requirement on the cost function and better convergence guarantees.

3.2.1 The alternating direction method of multipliers (ADMM)

The alternating direction method of multipliers (ADMM), or similar methods under different names, can be traced back to the 1970s [21], or even as early as the 1950s. It has been recently revived due to the demand in solving problems in machine learning and statistics with large data sets. Interested readers can refer to Boyd et al. [15] for a recent comprehensive review on ADMM.

ADMM solves optimization problems with equality constraints in the form

$$\begin{aligned} \underset{x,z}{\text{minimize}} \quad & f(x) + g(z) \\ \text{subject to} \quad & Ax + Bz = c. \end{aligned} \tag{3.12}$$

We require that both f and g should be convex functions so that the optimization problem (3.12) is convex. ADMM uses the augmented Lagrangian of problem (3.12) as given below:

$$L_\rho(x, z, y) = f(x) + g(z) + y^T(Ax + Bz - c) + (\rho/2)\left\|Ax + Bz - c\right\|^2$$

where y is the dual variable corresponding to the equality constraints and $\rho > 0$ is the penalty parameter. The augmented Lagrangian L_ρ can be viewed as the (regular) Lagrangian of a modified optimization problem with a slightly different objective function:

$$\begin{aligned} \underset{x,z}{\text{minimize}} \quad & f(x) + g(z) + (\rho/2)\left\|Ax + Bz - c\right\|^2 \\ \text{subject to} \quad & Ax + Bz = c. \end{aligned} \tag{3.13}$$

The term $(\rho/2)\left\|Ax + Bz - c\right\|^2$ vanishes when the equality constraint is satisfied. Therefore, problem (3.13) achieves the same optimum as the original problem (3.12). This implies that we can find the optimal solution to problem (3.12) by working with the augmented Lagrangian. This has several numerical benefits including better convergence guarantees (see [15] for details).

From optimization theory, we know that problem (3.13) (and hence problem (3.12)) can be solved by solving the unconstrained problem that minimizes $L_\rho(x, z, y)$ when the value of the dual variable y is chosen as y^*, the optimal value of the dual problem. ADMM proceeds by searching for y^* iteratively using dual ascent: At iteration $t + 1$, the primal variables x and z are first updated by minimizing $L_\rho(x, z, y)$ for the current value of y^t. In particular, ADMM updates x and z separately (hence the name alternating direction) by solving two sequential unconstrained optimization problems:

$$x^{t+1} := \arg\min_x L_\rho(x, z^t, y^t), \tag{3.14}$$

$$z^{t+1} := \arg\min_z L_\rho(x^{t+1}, z, y^t). \tag{3.15}$$

After the primal variables x and z are updated, ADMM updates the dual variable y by using dual ascent:

$$y^{t+1} := y^t + \rho(Ax^{t+1} + Bz^{t+1} - c).$$

Note that the step size is the same as the penalty parameter ρ.

Partitioning the primal variables into two groups x and z gives flexibility in exploiting structures in the optimization problem. For many applications, the x-update (3.14) and/or the z-update (3.15) can be often be parallelized if the primal variables are grouped properly. In certain cases, these updates can even be computed in closed form (e.g., when the minimizations in (3.14) and/or (3.15) are least-squares). Readers can refer to [15] for examples on partitioning techniques. In general, the way of partitioning is problem dependent and usually requires careful thoughts. In many cases, the given problem is not even in the form (3.12) and additional steps, such as introducing slack variables, are needed before applying ADMM.

3.2.2 ADMM on the convex optimal uncertainty quantification problem

In the optimal uncertainty quantification problem presented in Section 2.5, we are given the estimated mean $\hat{\mu}$ and covariance $\widehat{\Sigma}$, and cost function

$$f(\theta) = \max_{k \in \mathcal{K}} \{a_k^T \theta + b_k\},$$

and the task is to solve the optimization problem (2.33), which is rewritten below in a more compact notation:

$$\begin{aligned} \underset{P}{\text{minimize}} \quad & \text{tr}(\widetilde{\Sigma} P) && (3.16) \\ \text{subject to} \quad & P - S_k \succeq 0, \quad k \in \mathcal{K}, && (3.17) \end{aligned}$$

where

$$\widetilde{\Sigma} = \begin{bmatrix} \widehat{\Sigma} + \hat{\mu}\hat{\mu}^T & \hat{\mu} \\ \hat{\mu} & 1 \end{bmatrix}, \quad S_k = \begin{bmatrix} 0 & a_k/2 \\ a_k^T/2 & b_k \end{bmatrix}.$$

In order to apply ADMM, we introduce slack variables $\{Q_k\}_{k\in\mathcal{K}}$ and rewrite problem (3.16) as a problem with equality constraints:

$$\begin{aligned} \underset{P,\{Q_k\}}{\text{minimize}} \quad & \operatorname{tr}(\widetilde{\Sigma}P) + \sum_k I(Q_k, S_k) \\ \text{subject to} \quad & Q_k = P, \quad k \in \mathcal{K}, \end{aligned}$$

where

$$I(Q_k, S_k) = \begin{cases} 0 & Q_k - S_k \succeq 0, \\ \infty & \text{otherwise} \end{cases}$$

is the indicator function for incorporating the inequality constraints (3.17). We group the variables such that $x = \{Q_k\}_{k\in\mathcal{K}}$ and $z = P$. In this case, the augmented Lagrangian of this problem becomes

$$L_\rho(\{Q_k\}, P, \{Z_k\}) = \operatorname{tr}(\widetilde{\Sigma}P) + \sum_{k\in\mathcal{K}} I(Q_k, S_k) + \sum_{k\in\mathcal{K}} \operatorname{tr}(Z_k(Q_k - P)) + (\rho/2) \sum_{k\in\mathcal{K}} \|Q_k - P\|^2,$$

where $\{Z_k\}_{k\in\mathcal{K}}$ are the dual variables. For every iteration of ADMM, the update steps are:

x-update During the x-update, the variables $P = P^t$ and $\{Z_k^t\}$ are fixed and $\{Q_k^{t+1}\}$ is obtained by solving the minimization problem

$$\{Q_k^{t+1}\} := \arg\min_{\{Q_k\}} L_\rho(\{Q_k\}, P^t, \{Z_k^t\}). \tag{3.18}$$

In this case, the augmented Lagrangian decomposes over $\{Q_k\}$:

$$L_\rho(\{Q_k\}, P^t, \{Z_k^t\}) = \sum_{k\in\mathcal{K}} L_{\rho,k}(Q_k, P^t, Z_k^t) + r(P^t, \{Z_k^t\}),$$

where

$$\begin{aligned} L_{\rho,k}(Q_k, P^t, Z_k^t) &= I(Q_k, S_k) + \operatorname{tr}(Z_k^t Q_k) + (\rho/2) \left\|Q_k - P^t\right\|^2, \\ r(P^t, \{Z_k^t\}) &= \operatorname{tr}(\widetilde{\Sigma}P^t) - \sum_{k\in\mathcal{K}} \operatorname{tr}(Z_k^t P^t). \end{aligned}$$

Therefore, the minimization step (3.18) can be computed separately for each Q_k^{t+1} as

$$Q_k^{t+1} := \arg\min_{Q_k} L_{\rho,k}(Q_k, P^t, Z_k^t).$$

This is equivalent to solving the constrained least-squares problem

$$\begin{array}{ll} \underset{Q_k}{\text{minimize}} & \operatorname{tr}(Z_k^t Q_k) + (\rho/2)\left\|Q_k - P^t\right\|^2 \\ \text{subject to} & Q_k - S_k \succeq 0. \end{array}$$

Its solution Q_k^{t+1} can be obtained by projecting the solution $\widetilde{Q}_k^{t+1}$ of the unconstrained least-squares problem

$$\underset{Q_k}{\text{minimize}} \quad \operatorname{tr}(Z_k^t Q_k) + (\rho/2)\left\|Q_k - P^t\right\|^2$$

onto the constraint set $\{Q_k \colon Q_k - S_k \succeq 0\}$, i.e.,

$$Q_k^{t+1} = S_k + \Pi_{\mathbb{S}_+^{n+1}}(\widetilde{Q}_k^{t+1} - S_k),$$

whereas solution of the unconstrained least-squares problem can obtained in closed form as

$$\widetilde{Q}_k^{t+1} = P^t - Z_k^t/\rho.$$

By combining the results, we have

$$Q_k^{t+1} = S_k + \Pi_{\mathbb{S}_+^{n+1}}(P^t - Z_k^t/\rho - S_k)$$

as the rule for x-update.

z-update During the z-update, P^{t+1} is obtained by solving the minimization problem

$$\begin{aligned} P^{t+1} &:= \arg\min_P L_\rho(\{Q_k^{t+1}\}, P, \{Z_k^t\}) \\ &= \arg\min_P \left[\operatorname{tr}(\widetilde{\Sigma}P) - \sum_{k\in\mathcal{K}} \operatorname{tr}(Z_k^t P) + (\rho/2)\sum_{k\in\mathcal{K}} \left\|Q_k^{t+1} - P\right\|^2\right]. \end{aligned}$$

This is an unconstrained least-squares problem over P and its solution can be computed in closed form as

$$P^{t+1} = \frac{1}{|\mathcal{K}|}\left[\sum_{k\in\mathcal{K}} Q_k^{t+1} + \Big(\sum_{k\in\mathcal{K}} Z_k^t - \widetilde{\Sigma}\Big)/\rho\right],$$

Dual ascent Finally, the dual ascent update for the dual variables $\{Z_k\}_{k\in\mathcal{K}}$ can also be carried out separately for each Z_k:

$$Z_k^{t+1} := Z_k^t + \rho(Q_k^{t+1} - P^{t+1}).$$

The step that dominates computational complexity is the projection onto $\mathbb{S}_+^{n+1}$ when updating $\{Q_k\}$. It requires a total number of $|\mathcal{K}|$ eigenvalue decompositions of an $(n+1)$-dimensional matrix, each of which has the complexity of $O(n^3)$ [72]. However, note that the updates on $\{Q_k\}$ are independent for different k and hence can be carried out in parallel, which would give potential speed improvement when parallel computing is available.

Table 3.1 compares the computational time of SeDuMi (interfaced via YALMIP [43]) and ADMM for solving problem (3.16). The computation is done in MATLAB on a personal laptop equipped with an Intel Core 2 Duo 2.4 GHz processor and 4 GB of memory. In the case of ADMM, no parallelization is implemented. All data, including the coefficients of the cost function $\{a_k, b_k\}_{k\in\mathcal{K}}$, mean $\hat{\mu}$, and covariance $\widehat{\Sigma}$ are chosen arbitrarily. Note that when the problem size is extremely small, e.g., when $n = 2$ and $|\mathcal{K}| = 3$, the CPU time measurements may not reflect the true time complexity of the optimization algorithm due to other overhead involved in computation. It can be seen from the comparison that, for small to medium sized problems, SeDuMi outperforms ADMM in computational time due to the fact that it is a second-order method and requires fewer iterations to converge. However, as problem size grows, ADMM becomes more favorable, especially for cases where $|\mathcal{K}|$ is small, but n is large. In the case of $n = 150$, SeDuMi even failed to compute the optimal solution due to insufficient memory. It should also be noted that, in the extremely case of unlimited number of computational units, a factor of $1/|\mathcal{K}|$ reduction in computational time of ADMM can be expected if parallelization is implemented.

n	$\lvert\mathcal{K}\rvert$	SeDuMi CPU time (s)	ADMM CPU time (s)
2	3	0.38	0.08
20	30	2.96	5.56
50	100	318.9	318.2
100	2	304.9	0.45
150	2	out of memory	1.26

Table 3.1: Computational time of SeDuMi vs. ADMM for solving convex optimal uncertainty quantification problems with first and second moment constraints.

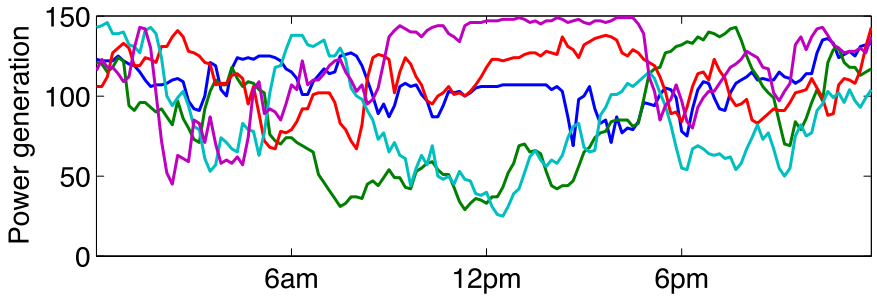


Figure 3.2: Time traces of wind generation on five different days. (Source: AESO)

3.3 Application: Energy storage placement evaluation in power grids

In this section, we introduce the storage placement evaluation problem in power grids as one application of OUQ. One of the major efforts in power systems research is to increase the penetration level of renewable energy in power grids. Unlike conventional generation, renewable energy sources (e.g., wind, see Fig. 3.2) tend to suffer from random fluctuations over time and can lead to reliability issues. Placing storage devices in power grids (Fig. 3.3) is considered a promising solution to mitigating the effect of random fluctuations in the renewables [59], and related problems were recently studied in, e.g., [24, 68]. In this context, it is important to evaluate the ramifications of a given storage placement plan [31]. It is shown that this evaluation problem can be converted into PCF using Lagrange duality and solved within the framework of convex OUQ. Later in this section, numerical results for the storage placement problem are presented, where a total of three scenarios are considered. For the first two scenarios, we use simple network configurations, in particular, 1-bus and 2-bus networks with synthetic renewable generation data. The purpose of these examples is to show some insight into the differences between deterministic analysis and the OUQ analysis. For the third scenario, we use the IEEE 14-bus test case as a more practical configuration

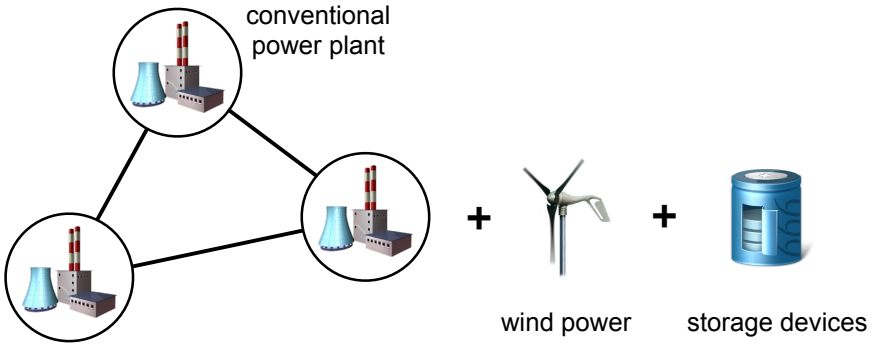


Figure 3.3: Integration of wind generation into power grids. Adding storage devices is considered as a promising solution to mitigating the random fluctuation of wind generation.

and data from real renewable generation. Through this example, we aim to demonstrate that the method is capable of analyzing a practical system.

3.3.1 A simple power grid model with energy storage

We model a power grid as a discrete-time dynamical system on a finite graph $(\mathcal{N}, \mathcal{E})$ with time indices $\mathcal{T} = \{1, 2, \ldots, T\}$, which is illustrated in Fig. 3.4. The vertices $\mathcal{N}$ are also called *buses*. For simplicity, we use the shorthand notation x to denote the vectorization of any set of variables $\{x_i(t)\}_{i\in\mathcal{N}, t\in\mathcal{T}}$. At time t, we refer to $g_i(t)$, $d_i(t)$, and $r_i(t)$ as power generation from renewables, user consumption, and charge rate of storage devices at bus i. As a convention, if the storage devices are being charged, then $r_i(t) > 0$. Under this convention, the total local net power consumption becomes

$$d_i(t) - g_i(t) + r_i(t).$$

Due to physical constraints, the storage level at bus i must stay between 0 and the maximum capacity E_i, i.e.,

$$0 \le \sum_{\tau=0}^{t} r_i(\tau) \le E_i, \quad \forall i \in \mathcal{N},\ t \in \mathcal{T} \cup \{0\}.$$

In an abuse of notation, we use $r_i(0)$ to denote the initial level of storage at bus i. Aside from local generation and consumption, power can also flow between adjacent buses. For

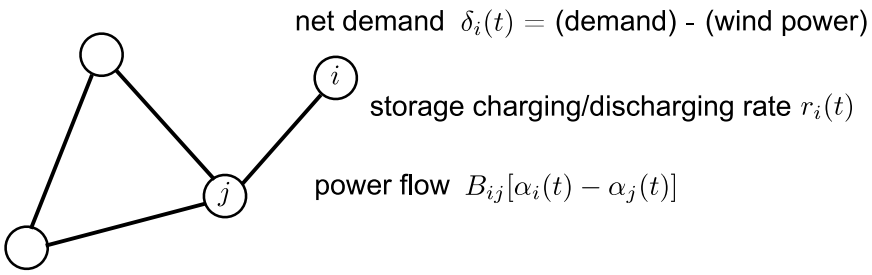


Figure 3.4: Model of power grid with energy storage. At any time t and for each bus (node) i, there is an associated net demand $\delta_i(t)$, and the storage device can be charged/discharged with rate $r_i(t)$. Between any two buses i and j, the power flow is proportional to the difference in their voltage angles $[\alpha_i(t) - \alpha_j(t)]$.

any neighboring buses i and j (i.e., $(i,j) \in \mathcal{E}$), the power flow from i to j is given by

$$B_{ij}(\alpha_i(t) - \alpha_j(t)),$$

where B_{ij} is the susceptance of the transmission line between i and j, and $\alpha_i(t)$ is the voltage angle of bus i. Here we use a DC power flow model for simplicity (cf. [56] for its applicability). A transmission line can only support a limited amount power flow $Q_{ij} \geq 0$, which imposes the constraint

$$|B_{ij}[\alpha_i(t) - \alpha_j(t)]| \leq Q_{ij}, \quad \forall (i,j) \in \mathcal{E},\ t \in \mathcal{T}.$$

In summary, the total net power consumption at bus i is

$$P_i(t) = \delta_i(t) + r_i(t) + \sum_{(i,j)\in\mathcal{E}} B_{ij}[\alpha_i(t) - \alpha_j(t)],$$

where $\delta_i(t) \triangleq d_i(t) - g_i(t)$. If $P_i(t) \leq 0$, the consumption is covered by all the sources, including local sources and and power flow from adjacent buses. However, if $P_i(t) > 0$, the unmet portion must be matched by additional power sources, usually from the so-called *spinning reserves* in the form of conventional generation.

For simplicity, we assume that the operating cost only depends on the amount of power drawn from spinning reserves. All the other factors, including renewable usage, charging/discharging, and power transmission are assumed to incur no cost. This simplification can potentially be crude. For example, storage devices such as chemical batteries often have

a finite number of charging cycles, so charging/discharging cannot be treated as entirely free. These potential refinements will be left for future work. Under this assumption, at time t, the cost for bus i can be modeled as a hinge cost

$$J_i(t) = [P_i(t)]^+ \triangleq \max\{P_i(t), 0\},$$

and the operating cost for the entire grid over time is

$$J = \sum_{i \in \mathcal{N}} \sum_{t=1}^{T} J_i(t).$$

Suppose $\delta_i(t)$ is known, for a given placement of storage $\{E_i\}_{i \in \mathcal{N}}$, one can choose how to operate the storage devices and transmit power over the network to minimize the operating cost by solving the problem

$$\begin{aligned} \underset{r,\alpha}{\text{minimize}} \quad & J(\delta, r, \alpha) \\ \text{subject to} \quad & |B_{ij}[\alpha_i(t) - \alpha_j(t)]| \le Q_{ij}, \quad \forall (i,j) \in \mathcal{E},\ t \in \mathcal{T} \\ & 0 \le \sum_{\tau=0}^{t} r_i(\tau) \le E_i, \quad i \in \mathcal{N},\ t \in \mathcal{T} \cup \{0\} \\ & \sum_{\tau=1}^{T} r_i(\tau) \ge 0, \quad i \in \mathcal{N}. \end{aligned} \tag{3.19}$$

The last constraint is added in order to prevent one from minimizing the operating cost by setting a large initial level of charge (which in practice will incur cost). This optimization problem is always feasible, since $r = 0$ and $\alpha = 0$ will satisfy all the constraints.

We would like to quantify the worst-case operating cost under a given placement of storage $\{E_i\}$. We treat $\delta_i(t)$ as the uncertainties for capturing the stochasticity in both renewable generation and user demand. There are two candidate formulations due to the extra freedom in optimizing the power flow by choosing r and α.

- $\max_{\delta \sim d} \mathbb{E}_\delta[\min_{r,\alpha} J(\delta, r, \alpha)]$: This is the "clairvoyant" worst-case analysis. It assumes that power flow optimization will have full knowledge about the actual instantiation of δ.

- $\min_{r,\alpha}[\max_{\delta \sim d} \mathbb{E}_\delta[J(\delta, r, \alpha)]]$: This is the "conservative" worst-case analysis. It assumes

a fixed plan for power flow, independent of the actual instantiation of δ.

In this work, we choose the first formulation because the time horizon under consideration will be 24 hours, and one normally has good knowledge about δ within this horizon (into the future) from forecast, which has been a common practice for many system operators. The second formulation seems too conservative by abandoning any real-time control on the power flow. Formally, the OUQ problem becomes

$$\begin{aligned} \underset{d}{\text{maximize}} \quad & \mathbb{E}_{\delta\sim d}\left[G(\delta)\right] \\ \text{subject to} \quad & \mathbb{E}_{\delta\sim d}[\delta] = \hat{\mu}, \quad \operatorname{cov}_{\delta\sim d}[\delta] \preceq \widehat{\Sigma}, \end{aligned}$$

where $G(\delta)$ is the optimal value of the optimization problem (3.19) for a given δ.

3.3.2 Conversion into PCF

Unfortunately, the function G is not in PCF. However, it is possible to convert G into PCF through Lagrange duality. By introducing slack variables, the optimization problem (3.19) can be rewritten as a linear program, i.e.,

$$\begin{aligned} \underset{r,\alpha,J_i(t)}{\text{minimize}} \quad & \sum_{i\in\mathcal{N}}\sum_{t=1}^{T} J_i(t) \\ \text{subject to} \quad & B_{ij}[\alpha_i(t)-\alpha_j(t)] \le Q_{ij}, \quad \forall (i,j)\in\mathcal{E},\ t\in\mathcal{T} \\ & B_{ij}[\alpha_i(t)-\alpha_j(t)] \ge -Q_{ij}, \quad \forall (i,j)\in\mathcal{E},\ t\in\mathcal{T} \\ & 0 \le \sum_{\tau=0}^{t} r_i(\tau) \le E_i, \quad i\in\mathcal{N},\ t\in\mathcal{T}\cup\{0\} \\ & \sum_{\tau=1}^{T} r_i(\tau) \ge 0, \quad i\in\mathcal{N} \\ & J_i(t) \ge 0 \\ & J_i(t) \ge \delta_i(t) + r_i(t) + \sum_{j\in N(i)} B_{ij}[\theta_i(t)-\theta_j(t)], \end{aligned}$$

whose Lagrange dual problem

$$\underset{\lambda,\nu}{\text{maximize}} \quad \sum_{i\in\mathcal{N}}\sum_{t=1}^{T}\lambda_i^{(1)}(t)\delta_i(t) - \sum_{i\in\mathcal{N}}\sum_{t=0}^{T}\lambda_i^{(2)}(t)E_i - \frac{\theta_{\max}}{2}\sum_{(i,j)\in\mathcal{E}}\sum_{t=1}^{T}(\lambda_{ij}^{(3)}(t)+\lambda_{ji}^{(3)}(t))$$

$$- \frac{1}{2} \sum_{(i,j)\in\mathcal{E}} \sum_{t=1}^{T} Q_{ij}(\lambda_{ij}^{(4)}(t) + \lambda_{ji}^{(4)}(t)) \tag{3.20}$$

$$\text{subject to} \quad 0 \le \lambda_i^{(1)}(t) \le 1, \quad \lambda_i^{(2)}(t) \ge 0, \quad i \in \mathcal{N},\ t \in \mathcal{T} \tag{3.21}$$

$$\lambda_{ij}^{(3)}(t), \lambda_{ij}^{(4)}(t) \ge 0, \quad (i,j) \in \mathcal{E},\ t \in \mathcal{T} \tag{3.22}$$

$$\lambda_i^{(2)}(t) \ge \lambda_i^{(1)}(t+1) - \lambda_i^{(1)}(t) \quad i \in \mathcal{N},\ t \in \mathcal{T}\backslash\{T\} \tag{3.23}$$

$$\lambda_i^{(2)}(T) \ge -\lambda_i^{(1)}(T) - \nu, \quad i \in \mathcal{N} \tag{3.24}$$

$$\lambda_i^{(2)}(0) \ge \lambda_i^{(1)}(1) + \nu, \quad i \in \mathcal{N} \tag{3.25}$$

$$\sum_{(i,j)\in\mathcal{E}} \Big[B_{ij}(\lambda_i^{(1)}(t) - \lambda_j^{(1)}(t) - \lambda_{ij}^{(4)}(t) + \lambda_{ji}^{(4)}(t)) - \lambda_{ij}^{(3)}(t) + \lambda_{ji}^{(3)}(t) \Big] = 0, \quad i \in \mathcal{N},\ t \in \mathcal{T}. \tag{3.26}$$

is also a linear program. It can be seen that the dual (3.20) has the form (3.2) for $a = \lambda^{(1)}$,

$$b = -\sum_{i\in\mathcal{N}} \sum_{t=0}^{T} \lambda_i^{(2)}(t) E_i - \frac{\theta_{\max}}{2} \sum_{(i,j)\in\mathcal{E}} \sum_{t=1}^{T} (\lambda_{ij}^{(3)}(t) + \lambda_{ji}^{(3)}(t)) - \frac{1}{2} \sum_{(i,j)\in\mathcal{E}} \sum_{t=1}^{T} Q_{ij}(\lambda_{ij}^{(4)}(t) + \lambda_{ji}^{(4)}(t)),$$

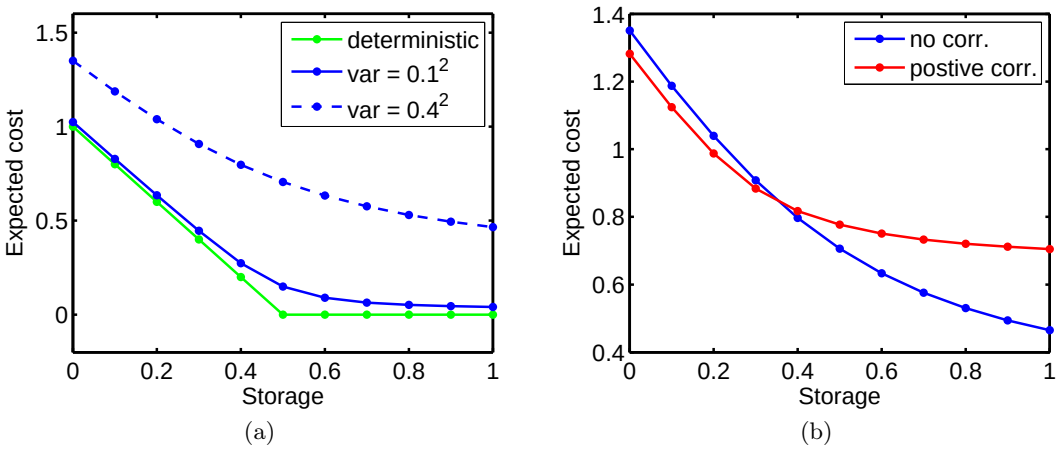
and the polytope $\mathcal{P}$ defined by constraints (3.21)-(3.26). Since the primal problem is a linear program and always feasible, we know that strong duality holds, which implies that the dual problem gives the same optimal value as the primal problem. In other words, G can be redefined by the dual problem and hence can be rewritten in PCF.

3.3.3 Numerical results: 1-bus and 2-bus networks

First we consider a network consisting of one isolated bus, i.e., $|\mathcal{N}| = 1$. This setting has the benefit of isolating any influence by power transmission. We will fix $\hat{\mu}$ and focus on the effect of $\widehat{\Sigma}$. The number of time slices is chosen as 5 so that the exact method can be used. Fig. 3.5a compares the results from (1) deterministic analysis, which assumes that δ follows $\hat{\mu}$ deterministically, (2) OUQ analysis with $\widehat{\Sigma} = (0.1)^2 I$ (I is the identity matrix), and (3) OUQ analysis with $\widehat{\Sigma} = (0.4)^2 I$. All the curves follow the law of diminishing returns, i.e., adding storage will become less helpful in reducing the operating cost if some storage has already been in place. The differences are in the slope of the curves. For the deterministic

analysis, there is a hard threshold after which adding storage will have zero reduction on the cost, whereas the same hard threshold does not appear for the OUQ analysis. This trend is not difficult to understand for the deterministic case: the operating cost cannot be made lower than the cumulative net demand over the entire time horizon, since adding storage does not contribute to power generation. For the results from the OUQ analysis, lower variance will cause a steeper slope. This can be understood by treating the case with lower variance as closer to the deterministic case, which has the steepest slope among all the curves.

The cost-storage curve is not only affected by the variance (diagonal entries of $\widehat{\Sigma}$), but also by the (time) correlation (off-diagonal entries of $\widehat{\Sigma}$). Fig. 3.5b compares the results of no correlation and positive correlation, where $\widehat{\Sigma}$ is generated from a Laplace covariance function (also known as covariance kernel): $\widehat{\Sigma}_{ij} = \exp(|i-j|/\tau)$ for some constant τ. It can be seen that the presence of positive correlation leads to a slower decrease in the cost. This is expected, since the cost is dominated by the "bad event" during which the net demand at all time instances becomes higher than normal simultaneously, and this is more likely to happen with positive time correlation.



(a) (b)

Figure 3.5: Results for 1-bus network. (a) Effect of variance. (b) Effect of (positive) time correlation.

The purpose of the 2-bus example is to examine the effect of power flow, which can potentially make the operating cost less sensitive to the locations of storage. In the extreme case, if an infinite amount of power is allowed to flow across a fully connected network, then any storage placement will give the same operating cost. For a 2-bus network, there can be

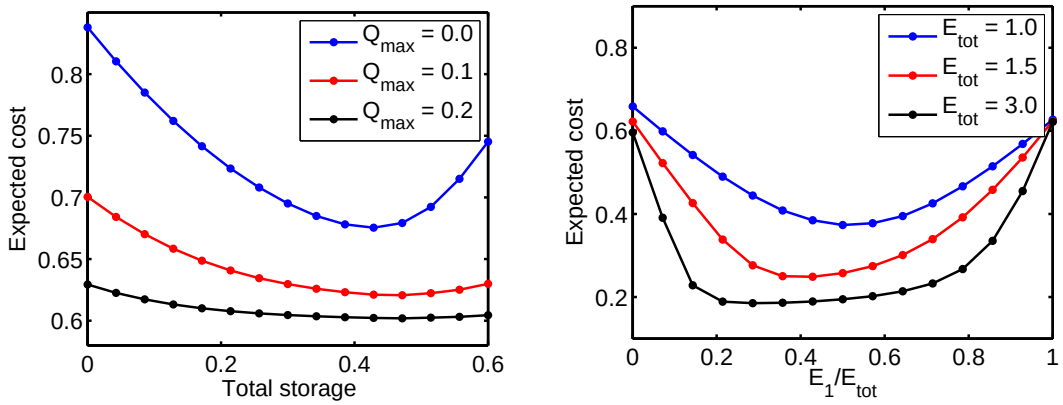


Figure 3.6: Results for 2-bus network. (a) Effect of transmission capacity $Q_{\max}$. (b) Effect of total storage E_{tot}.

only one transmission path, and we study how the maximum power flow $Q_{\max}$ of this path affects the operating cost. The two buses are set to be identical, except for their covariance matrix: $\widehat{\Sigma}_1 = (0.1)^2 I$ and $\widehat{\Sigma}_2 = (0.4)^2 I$. Fig. 3.6a compares the results for three power flow limits: $Q_{\max} = 0$ (the two buses are isolated), 0.1, and 0.2. In the simulation, the total storage E_{tot} is fixed, and the operating cost is plotted against E_1, the storage assigned to bus 1. As expected, as $Q_{\max}$ becomes larger, the distribution of storage between the two buses becomes less important.

We also study the effect of total storage E_{tot} on the distribution between the two buses. Fig. 3.5b shows the operating cost as a function of E_1/E_{tot}, the relative portion of storage for bus 1. As E_{tot} increases, assigning more portion to bus 2 becomes more beneficial. This can be understood from the diminishing return curves in Fig. 3.5a. Recall that bus 1, whose local demand has a lower variance, enters the diminishing return regime more quickly than bus 2. Therefore, when there has already been enough storage for bus 1, i.e., E_{tot} is large enough, it starts to become more helpful to assign more storage to bus 2, which has not yet entered the diminishing return regime.

3.3.4 Results: IEEE 14-bus network with renewable generation

In this more practical example, we choose the IEEE 14-bus test case [1] as the network model. The IEEE 14-bus system, which is shown in Fig. , can be viewed as an abstraction of a portion of the Midwestern US transmission grid. It consists of 5 generator buses and 9

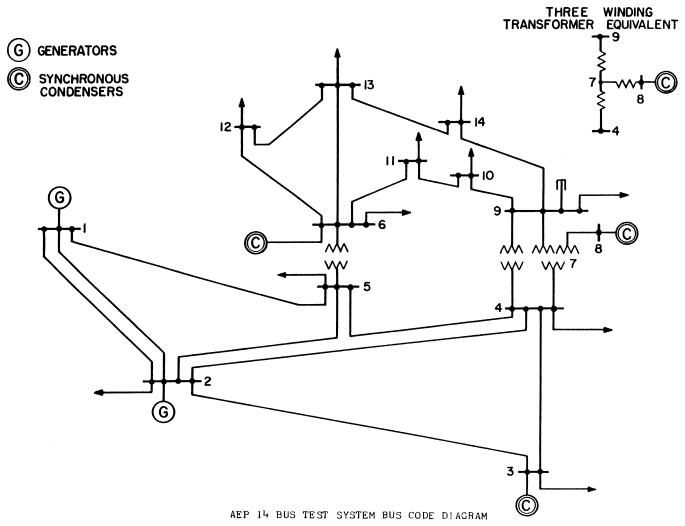


Figure 3.7: The IEEE 14-bus power systems test case. It consists of 5 generator buses and 9 load-only buses.

load-only buses. Daily load and generation profiles are created using the data set from [68]. For simplicity, we treat user demand as deterministic and assume that uncertainty only comes from generation, since uncertainty in generation often dominates that in user demand. The time horizon is set to be 24 hours and divided into 8 time slices, which gives 40 random variables in total. The sample mean $\hat{\mu}$ is also obtained from the data set in [68], whereas the sample covariance $\widehat{\Sigma}$ is computed from the historical data provided by the Alberta Electric System Operator (AESO) [3], since the number of samples in the former data set is insufficient to compute $\widehat{\Sigma}$. For compatibility, the generation data from AESO is scaled accordingly to match the data from [68].

Given $\hat{\mu}$, we can solve for the optimal storage placement strategy when there is no random fluctuation in generation, i.e., the generation profile always follows $\hat{\mu}$. This particular placement is then evaluated using the OUQ analysis. Due to the size of the problem, the approximate method in Section 3.1.2 is used. Similar to the 1-bus and 2-bus examples, correlation affects the result in the 14-bus example as well. Fig. 3.8a shows the results for (1) deterministic analysis assuming no random fluctuation, (2) OUQ analysis, and (3) sample average approximation (with 2σ error bars). Since the strategy under evaluation is the optimal placement in the absense of random fluctuation, the deterministic analysis gives the most optimistic prediction, i.e., a lower bound for the expected cost. The OUQ analysis, despite the fact that it considers the worst-case distribution, is surprisingly close to the

sample average. Although the OUQ analysis should in principle give an upper bound for the expected cost, its curves does not always stay above the sample averge, especially for large total storage. This is presumerably due to two reasons. First, the OUQ analysis is performed using the approximate method, which does not guarantee to yield an upper bound. This is a major weakness of the approximate method, and needs further investigation in the future. Second, the sample average can deviate from the true expected value due to finite samples. Since results from the OUQ analysis still fall within the confidence interval of the sample averge, it is believed that the OUQ analysis may still be close to the true upper bound of the expected cost. On the other hand, Fig. 3.8b compares the results with the deterministic worst-case analysis, which computes the cost under the worst single deterministic event by ignoring all the moment constraints. For this example, the worst case corresponds to constant zero renewable generation, since generation must stay nonnegative. The OUQ results are considerably less conservative than the deterministic worst-case analysis, which gives a constant cost of 48.13.

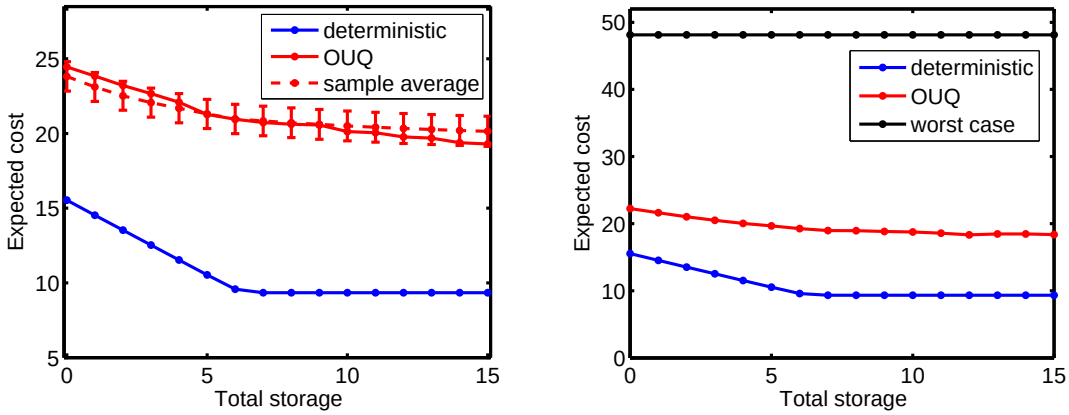


Figure 3.8: Cost evaluation using different models of uncertainty for the IEEE 14-bus case with real wind generation.

3.4 Conclusions

This chapter addresses some computational issues in solving optimal uncertainty quantification problems with piecewise affine objective and first and second moment constraints, which were introduced in Section 2.5. For analyzing practical systems, the scale of the

problem may become large: either the number of required Dirac masses or the dimension of the random variable. Two different approaches to the scaling issues are presented. One focuses on the case where the objective function can be expressed in polytopic canonical form (PCF), where the number of Dirac masses becomes exponentially large. This corresponds to solving a semidefinite program with exponentially many constraints. Since a majority of the constraints are inactive, our method is to search (locally) for the inactive constraints iteratively with the purpose of solving a much smaller problem with only the active constraints. Although updating the candidate set for the inactive constraints is generally difficult, it has been shown that this procedure only requires solving a linear program when the objective function is in PCF. Another focus is on solving the large original problem through massive parallelization. In particular, we investigate the application of the alternating direction method of multipliers (ADMM), which is a parallelizable first-order method. It has been shown that ADMM outperforms popular second-order optimization solvers such as SeDuMi when the problem size becomes large. In the end, we illustrate the application of these numerical methods in an example of energy storage placement evaluation in power grids. It shows that optimal uncertainty quantification can be readily applied to medium-sized practical systems using convex optimization with proper numerical implementations.

Chapter 4

OUQ via Convex Relaxation: An Example on Hoeffding's Inequality

This chapter applies the optimal uncertainty quantification framework to a setting that is usually presented along with Hoeffding's inequality [30], which is an important concentration inequality used in many areas. In this setting, the task is to obtain a tight bound for the probability that the sum of independent random variables deviates from the sum of their means. Although a tight bound is often nontrivial to obtain, Hoeffding's inequality is capable of giving a loose bound using a simple expression. In certain cases, however, especially when the number of samples is small, it is often desirable to obtain a tight bound to make the best use of the limited samples. This problem falls into the optimal uncertainty quantification framework and can be converted into an equivalent finite-dimensional problem through reduction. Unlike the problems presented in previous chapters, the corresponding optimization problem is equivalent to a number of non-convex polynomial optimization problems, for which there are no efficient numerical algorithms. However, using convex relaxation techniques such as sums-of-squares, it is still possible to obtain a valid bound that may be better than what is given by Hoeffding's inequality. In fact, numerical results show that the bound obtained from convex relaxation is tight in many cases. Another difficulty is that the number of polynomial optimization problems is usually quite large. By carefully exploiting structures in the problem such as symmetry, we show that it is possible to greatly reduce the number of polynomial optimization problems.

4.1 Hoeffding's inequality and its related OUQ problem

Consider a collection of n bounded independent random variables $X_1, X_2, \ldots, X_n$. Without loss of generality, we assume that each $X_i \in [0,1]$ almost surely. We are interested in obtaining an upper bound for

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} \mathbb{E}X_i + t\right). \tag{4.1}$$

Such quantity can be used to obtain an upper limit of the confidence interval of the sample mean $\sum_{i=1}^{n} X_i/n$, which widely used in many areas such as adaptive stochastic optimization [7, 17]. One such upper bound can be obtained through Hoeffding's inequality as

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} \mathbb{E}X_i + t\right) \leq \exp(-2t^2/n) = \exp(-2n\bar{t}^2), \qquad \bar{t} = t/n. \tag{4.2}$$

If $X_1, X_2, \ldots, X_n$ are not only independent, but also identically distributed, then as $n \to \infty$, we know from the central limit theorem that the random variable

$$\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathbb{E}X_i\right)$$

converges in distribution to the normal distribution $\mathcal{N}(0, \sigma^2)$, where σ^2 is the variance of X_i (for any i). In other words, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathbb{E}X_i\right) \geq z\right) = 1 - \Phi(z/\sigma),$$

where Φ is the cumulative probability distribution of the standard normal distribution $\mathcal{N}(0,1)$. Letting $z = \sqrt{n}\bar{t}$, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathbb{E}X_i\right) \geq \bar{t}\right) = 1 - \Phi(\sqrt{n}\bar{t}/\sigma).$$

Using the asymptotic approximation of $(1 - \Phi)$ [5]:

$$\lim_{x\to\infty} [1 - \Phi(x)] = \frac{1}{\sqrt{2\pi}x} \exp(-x^2/2),$$

we have

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathbb{E}X_i\right) \geq \bar{t}\right) = \frac{\sigma}{\bar{t}\sqrt{2\pi n}} \exp(-n\bar{t}^2/2\sigma^2).$$

The right-hand side increases with σ. For $X_i \in [0,1]$, the maximum achievable variance is $\sigma^2 = 1/4$, which leads to the following asymptotic upper bound

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n}\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathbb{E}X_i\right) \geq \bar{t}\right) \leq \frac{1}{2\bar{t}\sqrt{2\pi n}} \exp(-2n\bar{t}^2). \tag{4.3}$$

We can see that the bound (4.2) given by Hoeffding's inequality is similar to the asymptotic bound (4.3) from the central limit theorem as $n \to \infty$, except for the factor $1/2\bar{t}\sqrt{2\pi n}$, which decreases much more slowly than $\exp(-2n\bar{t}^2)$ as n grows. However, the bound is not tight in general for any finite n, and can be quite loose in particular for small n. For certain applications where samples are expensive to obtain (e.g., in stochastic simulation of complex systems), it can be of great interest to obtain the best bound possible in order to fully use the limited samples. Formally, the procedure of seeking the best bound can be cast as the following optimization problem over probability distributions

$$\begin{aligned} \underset{\nu,\{\nu_i\}_{i=1}^n}{\text{maximize}} \quad & \mathbb{P}_\nu\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} \mathbb{E}_{\nu_i}\left[X_i\right] + t\right) \\ \text{subject to} \quad & \nu = \nu_1 \otimes \nu_2 \otimes \cdots \otimes \nu_n. \end{aligned} \tag{4.4}$$

Each ν_i is a probability measure on $[0,1]$. The measure ν, defined on $[0,1]^n$, is the joint probability measure of the random vector $(X_1, X_2, \ldots, X_n)$ and the operation $\otimes$ denotes the product measure. This particular form of ν written as the product measure of individual measures $\{\nu_i\}_{i=1}^n$ encodes independence among the random variables $\{X_i\}_{i=1}^n$. For clarity, we indicate the dependence of $\mathbb{P}_\nu$ and $\mathbb{E}_{\nu_i}$ on their corresponding probability measures in the subscripts. After the introduction of slack variables $\mu_i = \mathbb{E}X_i$ $(i = 1, 2, \ldots, n)$, problem (4.4) can be rewritten as

$$\underset{\nu,\{\nu_i,\mu_i\}_{i=1}^n}{\text{maximize}} \quad \mathbb{P}_\nu\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} \mu_i + t\right) \tag{4.5}$$

$$\text{subject to} \quad \nu = \nu_1 \otimes \nu_2 \otimes \cdots \otimes \nu_n \tag{4.6}$$

$$\mathbb{E}_{\nu_i}[X_i] = \mu_i, \quad i = 1, 2, \ldots, n. \tag{4.7}$$

Both the objective function and constraint (4.7) are linear in the distributions ν and ν_i, respectively. If we ignore the independence constraint (4.6), problem (4.5) is an optimal uncertainty quantification problem with constraints on the first moment of each ν_i. In fact, as will be shown in the next section, the inclusion of the independence constraint (4.6) does not affect the application of finite reduction, and therefore problem (4.5) can be considered as a generalized optimal uncertainty quantification problem.

Analytical expression for the optimal solution of problem (4.5) is generally unavailable except for a few cases. In the case of $n = 2$ and $n = 3$, analytical expressions are given by Owhadi et al. [53]:

$$n = 2: \quad \max_{\nu} \mathbb{P}_\nu \left(\sum_{i=1}^{2} X_i \geq \sum_{i=1}^{2} \mathbb{E}X_i + t \right) = (1 - t/2)^2,$$
$$n = 3: \quad \max_{\nu} \mathbb{P}_\nu \left(\sum_{i=1}^{3} X_i \geq \sum_{i=1}^{3} \mathbb{E}X_i + t \right) = (1 - t/3)^3.$$

If the upper bound for X_i is removed, i.e., $X_i \in [0, \infty)$ almost surely, but the mean $\mathbb{E}X_i$ is fixed, then an analytical expression is conjectured by Samuels [61]. Therefore, in order to obtain a tight bound of (4.1) for arbitrary n, we need to solve problem (4.5) numerically.

4.2 Finite reduction

Recall from Section 2.1 that the infinite-dimensional problem (4.5) can be reduced to a finite-dimensional one whose optimal value remain unchanged, since the optimal distribution ν^* can always be achieved by a certain discrete distribution. The independence constraint (4.6) implies that each compositing ν_i^* can also be achieved by a discrete distribution whose number of compositing Dirac masses is determined by its own information constraint [53]. In this case, the only information constraint for ν_i is its mean, which implies that the optimal distribution ν_i^* contains at most two Dirac masses. In the following, we will denote the locations and weights of the two Dirac masses in each ν_i as $(x_{i,0}, x_{i,1})$ and $(p_i, 1 - p_i)$, respectively. After finite reduction, the reformulated optimization problem over the locations and weights of all the Dirac masses becomes

$$\underset{\{p_i, x_{i,0}, x_{i,1}\}_{i=1}^n, \mu}{\text{maximize}} \quad p_1 p_2 \dots p_n I\left(x_{1,0} + x_{2,0} + \cdots + x_{n,0} \geq \mu + t\right)$$

$$
\begin{aligned}
&+ (1-p_1)p_2 \dots p_n I\left(x_{1,1} + x_{2,0} + \cdots + x_{n,0} \geq \mu + t\right) \\
&+ \dots \\
&+ (1-p_1)(1-p_2)\dots(1-p_n) I\left(x_{1,1} + x_{2,1} + \cdots + x_{n,1} \geq \mu + t\right) \\
\text{subject to} \quad & 0 \leq p_i \leq 1, \quad 0 \leq x_{i,0} \leq x_{i,1} \leq 1, \quad i = 1, 2, \dots, n \\
& \mu = \sum_{i=1}^{n} \left(p_i x_{i,0} + (1-p_i) x_{i,1}\right),
\end{aligned}
$$

where I denotes the 0-1 indicator function. The constraint $x_{i,0} \leq x_{i,1}$ is added for convenience and will not affect the optimal solution due to symmetry in the problem. By introducing new variables $\delta_i \triangleq x_{i,1} - x_{i,0}$, we can rewrite the problem as

$$
\begin{aligned}
\underset{\{p_i, \delta_i\}_{i=1}^n}{\text{maximize}} \quad & p_1 p_2 \dots p_n I\left((p_1 - 1)\delta_1 + (p_2 - 1)\delta_2 + \cdots + (p_n - 1)\delta_n \geq t\right) \\
& + (1-p_1)p_2 \dots p_n I\left(p_1\delta_1 + (p_2 - 1)\delta_2 + \cdots + (p_n - 1)\delta_n \geq t\right) \\
& + \dots \\
& + (1-p_1)(1-p_2)\dots(1-p_n) I\left(p_1\delta_1 + p_2\delta_2 + \cdots + p_n\delta_n \geq t\right) \\
\text{subject to} \quad & 0 \leq p_i \leq 1, \quad 0 \leq \delta_i \leq 1, \quad i = 1, 2, \dots, n.
\end{aligned}
$$

To simplify notation, we define

$$
\begin{aligned}
f(p, \alpha) &\triangleq \prod_{i=1}^{n} p_i^{1-\alpha_i} (1-p_i)^{\alpha_i}, \\
g(p, \alpha, \delta) &\triangleq \sum_{i=1}^{n} (p_i - 1)^{1-\alpha_i} p_i^{\alpha_i} \delta_i = \sum_{i=1}^{n} (p_i - 1 + \alpha_i)\delta_i = (p - 1 + \alpha)^T \delta,
\end{aligned} \tag{4.8}
$$

where $p = (p_1, p_2, \dots, p_n) \in [0,1]^n$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \{0,1\}^n$, and $\delta = (\delta_1, \delta_2, \dots, \delta_n) \in [0,1]^n$. In this way, we are able to rewrite the problem as

$$
\begin{aligned}
\underset{p, \delta}{\text{maximize}} \quad & \sum_{\alpha \in \{0,1\}^n} f(p, \alpha) I\left(g(p, \alpha, \delta) \geq t\right) \\
\text{subject to} \quad & 0 \preceq p \preceq 1, \qquad 0 \preceq \delta \preceq 1,
\end{aligned} \tag{4.9}
$$

where $\preceq$ represents entry-wise inequality: for any $a, b \in \mathbb{R}^n$, we write $a \preceq b$ if and only if $a_i \leq b_i$ for all $i = 1, 2, \dots, n$. By introducing a new decision variable $A \subseteq \{0,1\}^n$ that

denotes which indicator functions are active, we can move the indicator functions in the objective into the constraints and rewrite problem (4.9) as

$$\begin{aligned} \underset{p,\delta,A}{\text{maximize}} \quad & \sum_{\alpha\in A} f(p,\alpha) \\ \text{subject to} \quad & 0 \preceq p \preceq 1, \qquad 0 \preceq \delta \preceq 1, \qquad A \subseteq \{0,1\}^n \\ & g(p,\alpha,\delta) \geq t, \quad \alpha \in A \\ & g(p,\alpha,\delta) < t, \quad \alpha \notin A. \end{aligned} \tag{4.10}$$

If A is given, then problem (4.10) becomes

$$\begin{aligned} \underset{p,\delta}{\text{maximize}} \quad & \sum_{\alpha\in A} f(p,\alpha) \\ \text{subject to} \quad & 0 \preceq p \preceq 1, \qquad 0 \preceq \delta \preceq 1 \\ & g(p,\alpha,\delta) \geq t, \quad \alpha \in A \\ & g(p,\alpha,\delta) < t, \quad \alpha \notin A, \end{aligned} \tag{4.11}$$

which is a polynomial optimization problem since both $f(p,\alpha)$ and $g(p,\alpha,\delta)$ are polynomials in p and δ. Therefore, the optimal value of problem (4.10) can be obtained by enumerating all possible choices of A and solving the corresponding polynomial optimization problem (4.11) for each A. The issue with this approach, however, is that the number of choices of A is 2^{2^n}, and thus computationally intractable even for small n.

4.3 Removal of redundant enumerations

Although the doubly exponentially many enumerations may seem daunting at first glance, many of them can be eliminated in the first place without affecting the optimality. For instance, some choices of A will render problem (4.11) infeasible: if we take $n = 2$ and $A = \{(0,0)\}$, then part of the constraints in problem (4.11) will become

$$\begin{aligned} (p-1)^T\delta \geq t \qquad & \text{for } \alpha = (0,0) \in A, \\ p^T\delta < t \qquad & \text{for } \alpha = (1,1) \notin A, \end{aligned}$$

which is infeasible since $\delta \succeq 0$. In this section, we investigate further along this direction and show that the number of enumerations can be reduced to something much less than 2^{2^n} by exploiting various structures in the optimization problem (4.10).

We first show that the last constraint in problem (4.10) can be removed without affecting the optimal value of problem (4.10) in the following lemma:

Lemma 4.1. *The following optimization problem*

$$\begin{aligned} \underset{p,\delta,A}{\text{maximize}} \quad & \sum_{\alpha\in A} f(p,\alpha) \\ \text{subject to} \quad & 0 \preceq p \preceq 1, \qquad 0 \preceq \delta \preceq 1, \qquad A \subseteq \{0,1\}^n \\ & g(p,\alpha,\delta) \geq t, \quad \alpha \in A. \end{aligned} \tag{4.12}$$

achieves the same optimal value as problem (4.10).

Proof. Denote the optimal value of problem (4.10) as OPT_1 and that of problem (4.12) as OPT_2. We have $\mathrm{OPT}_1 \leq \mathrm{OPT}_2$ since problem (4.12) has a larger constraint set. To show that $\mathrm{OPT}_1 \geq \mathrm{OPT}_2$, we first denote the optimal solution of problem (4.12) as (p^*, δ^*, A^*) and define

$$\widetilde{A}^* \triangleq \{\alpha \in \{0,1\}^n \colon g(p^*,\alpha,\delta^*) \geq t\}.$$

It is not difficult to verify that $A^* \subseteq \widetilde{A}^*$, which also implies that

$$\sum_{\alpha\in\widetilde{A}^*} f(p^*,\alpha) \geq \sum_{\alpha\in A^*} f(p^*,\alpha) = \mathrm{OPT}_2. \tag{4.13}$$

On the other hand, we have

$$\mathrm{OPT}_1 \geq \sum_{\alpha\in\widetilde{A}^*} f(p^*,\alpha) \tag{4.14}$$

since $(p^*, \delta^*, \widetilde{A}^*)$ is a feasible solution to problem (4.10). By combining (4.13) and (4.14) together, we have $\mathrm{OPT}_1 \geq \mathrm{OPT}_2$ and the lemma is proved. □

Next, we will show that the optimal solution δ^* to problem (4.10) can be chosen from a finite set instead of the original set $\{\delta \colon 0 \preceq \delta \preceq 1\}$ appeared in the constraint. This will not only cut down the number of variables in the corresponding polynomial optimization

problem (4.11) by removing δ from (4.11), but, perhaps surprisingly, also the number of enumerations.

Lemma 4.2. *Let $\Delta(A)$ be a set such that $\delta \in \Delta(A)$ if and only if (δ, γ) is a vertex of the polytope defined by the inequalities*

$$0 \preceq \delta \preceq 1, \qquad \alpha^T \delta - \gamma \geq 0, \quad \alpha \in A.$$

Then the following optimization problem

$$\begin{aligned} \underset{p,\delta,A}{\text{maximize}} \quad & \sum_{\alpha \in A} f(p, \alpha) && (4.15) \\ \text{subject to} \quad & 0 \preceq p \preceq 1, \qquad \delta \in \Delta(A), \qquad A \subseteq \{0,1\}^n \\ & g(p, \alpha, \delta) \geq t, \quad \alpha \in A && (4.16) \end{aligned}$$

achieves the same optimal value as problem (4.12).

Proof. Denote the optimal value of problem (4.12) as OPT_2 and that of problem (4.15) as OPT_3. We have $\text{OPT}_2 \geq \text{OPT}_3$ since $\Delta(A) \subset \{\delta \colon 0 \preceq \delta \preceq 1\}$. To show that $\text{OPT}_2 \leq \text{OPT}_3$ also holds, we denote the optimal solution to problem (4.12) as (p^*, δ^*, A^*). It then follows that

$$g(p^*, \alpha, \delta^*) \geq t, \quad \alpha \in A^*. \tag{4.17}$$

Using the definition of g in (4.8), we can rewrite (4.17) as

$$(p^* - 1 + \alpha)^T \delta^* \geq t, \quad \alpha \in A^*,$$

or

$$(p^* - 1)^T \delta^* + \min_{\alpha \in A^*} \alpha^T \delta^* \geq t.$$

In other words, the optimal value of the optimization problem (with given p^* and A^*)

$$\begin{aligned} \underset{\delta}{\text{maximize}} \quad & (p^* - 1)^T \delta + \min_{\alpha \in A^*} \alpha^T \delta \\ \text{subject to} \quad & 0 \preceq \delta \preceq 1 \end{aligned}$$

must be greater than t since $\delta = \delta^*$ is a feasible solution to this problem. By introducing a slack variable

$$\gamma = \min_{\alpha \in A^*} \alpha^T \delta,$$

we can rewrite the problem as a linear program:

$$\begin{aligned} \underset{\delta,\gamma}{\text{maximize}} \quad & (p^* - 1)^T \delta + \gamma \\ \text{subject to} \quad & 0 \preceq \delta \preceq 1, \\ & \alpha^T \delta - \gamma \geq 0, \quad \alpha \in A^*. \end{aligned} \tag{4.18}$$

For convenience, we also define $\gamma^* = \min_{\alpha \in A^*} \alpha^T \delta^*$. It is not difficult to check that $\Delta(A^*)$ is the projection of the set of vertices of the constraint polytope

$$0 \preceq \delta \preceq 1, \qquad \alpha^T \delta - \gamma \geq 0, \quad \alpha \in A^*$$

onto the coordinate δ. Recall that the optimal value of a linear program can always be achieved at a vertex of the constraint polytope (cf. [11]). Denote the δ and γ this vertex as $\widetilde{\delta}^*$ and $\widetilde{\gamma}^*$, respectively. We have $\widetilde{\delta}^* \in \Delta(A^*)$ and

$$(p^* - 1)^T \widetilde{\delta}^* + \widetilde{\gamma}^* = (p^* - 1)^T \delta^* + \gamma^*.$$

Since $\widetilde{\delta}^*$ and $\widetilde{\gamma}^*$ satisfy the constraint (4.18), we have $\min_{\alpha \in A^*} \alpha^T \widetilde{\delta}^* \geq \widetilde{\gamma}^*$ and hence

$$\begin{aligned} (p^* - 1)^T \widetilde{\delta}^* + \min_{\alpha \in A^*} \alpha^T \widetilde{\delta}^* &\geq (p^* - 1)^T \widetilde{\delta}^* + \widetilde{\gamma}^* \\ &= (p^* - 1)^T \delta^* + \gamma^* \\ &= (p^* - 1)^T \delta^* + \min_{\alpha \in A^*} \alpha^T \delta^* \geq t. \end{aligned}$$

As a result, we know that $(p^*, \widetilde{\delta}^*, A^*)$ is a feasible solution to problem (4.15), which implies that

$$\mathrm{OPT}_3 \geq \sum_{\alpha \in A^*} f(p^*, \alpha) = \mathrm{OPT}_2.$$

Here we have used the fact that the objective function of problem (4.15) does not depend on δ. □

The optimization problem (4.15) in Lemma 4.2 contains the constraint $\delta \in \Delta(A)$, which is a joint constraint on both δ and A. As will be seen later, decoupling this constraint between δ and A will lead to a simpler problem. One such way of decoupling is to replace $\Delta(A)$ in the constraint with $\bigcup_{A\subseteq\{0,1\}^n} \Delta(A)$. This does not affect the optimal value of the problem, since $\Delta(A) \subseteq \bigcup_{A\subseteq\{0,1\}^n} \Delta(A) \subseteq \{\delta\colon 0 \preceq \delta \preceq 1\}$.

Corollary 4.3. *The following optimization problem*

$$\begin{aligned} \underset{p,\delta,A}{\text{maximize}} \quad & \sum_{\alpha\in A} f(p,\alpha) && (4.19)\\ \text{subject to} \quad & 0 \preceq p \preceq 1, \qquad \delta \in \bigcup_{A\subseteq\{0,1\}^n} \Delta(A), \qquad A \subseteq \{0,1\}^n \\ & g(p,\alpha,\delta) \geq t, \quad \alpha \in A && (4.20) \end{aligned}$$

achieves the same optimal value as problem (4.12).

Remark 4.4. The set $\Delta(A)$ is a finite set since the number of vertices of a finite-dimensional polytope is finite.

At this point, we have successfully reduced the constraint set of δ from a continuum to a finite set. However, the difficulty that the number of choices for A is doubly exponential still remains unsolved. In the following, we show that once the choice of δ becomes finite, the corresponding optimal solution A^* (when δ is given) can be chosen from a set whose size is much smaller than 2^{2^n}. From the proof of Lemma 4.2, we know that the constraint (4.16) can be rewritten as

$$\delta^T p \geq t + 1^T\delta - \min_{\alpha\in A} \alpha^T\delta.$$

Define $j(\delta, A) \triangleq 1^T\delta - \min_{\alpha\in A} \alpha^T\delta$. It can be verified that $j(\delta, A_1)$ and $j(\delta, A_2)$ can be the same for a given δ even $A_1 \neq A_2$. In fact, the number of different $j(\delta, A)$ (for a given δ) is much smaller than the number of different choices of A. Therefore, we can choose to enumerate over all possible values of $j(\delta, A)$ instead of A itself, which is the key to the this reduction. For convenience, we define the set

$$M(\delta) \triangleq \{j(\delta, A)\colon A \subseteq \{0,1\}^n\} = \{1^T\delta - \alpha^T\delta\colon \alpha \in \{0,1\}^n\}.$$

to capture the possible values of $j(\delta, A)$, which is a finite set for any given δ (and is much

smaller than 2^{2^n} as will be seen later). Computing $M(\delta)$ is not numerically expensive for moderate n since it contains at most 2^n elements. Using the newly introduced notations, we are able to rewrite problem (4.12) as the one given in the following theorem, which is the main result of this section.

Theorem 4.5. *Let* $\bar{A}(\delta, m) \triangleq \{\alpha \in \{0,1\}^n \colon 1^T\delta - \alpha^T\delta \leq m\}$. *The optimization problem*

$$\begin{aligned} \underset{p,\delta,m}{\text{maximize}} \quad & \sum_{\alpha \in \bar{A}(\delta,m)} f(p,\alpha) \\ \text{subject to} \quad & 0 \preceq p \preceq 1, \qquad \delta \in \bigcup_{A \subseteq \{0,1\}^n} \Delta(A), \qquad m \in M(\delta) \\ & \delta^T p \geq t + m \end{aligned} \tag{4.21}$$

achieves the same optimal value as problem (4.10).

Proof. According to Lemma 4.1 and Corollary 4.3, it suffices to prove that problem (4.21) achieves the same optimal value as problem (4.15). We first define an optimization problem over A as follows:

$$\begin{aligned} \underset{A}{\text{maximize}} \quad & \sum_{\alpha \in A} f(p,\alpha) \\ \text{subject to} \quad & A \subseteq \{0,1\}^n \\ & \delta^T p \geq t + 1^T\delta - \min_{\alpha \in A} \alpha^T\delta, \end{aligned} \tag{4.22}$$

whose optimal value is denoted as $\mathrm{OPT}_A(p,\delta)$. In this way, problem (4.15) can be rewritten as

$$\begin{aligned} \underset{p,\delta}{\text{maximize}} \quad & \mathrm{OPT}_A(p,\delta) \\ \text{subject to} \quad & 0 \preceq p \preceq 1, \qquad \delta \in \bigcup_{A \subseteq \{0,1\}^n} \Delta(A). \end{aligned} \tag{4.23}$$

Using the definition of $j(\delta, A)$, we can further rewrite problem (4.22) by introducing a slack variable m as

$$\underset{A,m}{\text{maximize}} \quad \sum_{\alpha \in A} f(p,\alpha) \tag{4.24}$$

$$\begin{aligned} \text{subject to} \quad & A \subseteq \{0,1\}^n, \qquad m \in M(\delta) \\ & \delta^T p \geq t + m, \qquad j(\delta, A) = m. \end{aligned}$$

Since $f(p, \alpha) \geq 0$, the optimal solution of A in problem (4.22) is

$$\bar{A}(\delta, m) \triangleq \{\alpha \in \{0,1\}^n \colon 1^T\delta - \alpha^T\delta \leq m\},$$

which is the largest subset for which the constraint $j(\delta, A) = m$ holds, i.e., if A satisfies $j(\delta, A) = m$ for some given δ and m then $A \subseteq \bar{A}(\delta, m)$. In other words, the following problem,

$$\begin{aligned} \underset{m}{\text{maximize}} \quad & \sum_{\alpha \in \bar{A}(\delta,m)} f(p, \alpha) \\ \text{subject to} \quad & m \in M(\delta), \qquad \delta^T p \geq t + m, \end{aligned} \tag{4.25}$$

achieves the same optimal value as problem (4.24). Combining (4.23) and (4.25) completes the proof. □

Remark 4.6. Define the following optimization problem (that depends on δ and m):

$$\begin{aligned} \underset{p}{\text{maximize}} \quad & \sum_{\alpha \in \bar{A}(\delta,m)} f(p, \alpha) \\ \text{subject to} \quad & 0 \preceq p \preceq 1 \\ & \delta^T p \geq t + m. \end{aligned} \tag{4.26}$$

Note that in problem (4.21), both δ and m are chosen from finite sets. Therefore, we can solve problem (4.21) by enumerating all possible combinations of δ and m and solve the corresponding subproblem (4.26). Compared to the subproblem (4.11) of the original problem, the subproblem (4.26) only has n variables (instead of $2n$ in (4.11)). Moreover, as we will show later, the number of subproblems (i.e., the number of combinations of δ and m) is significantly fewer than that generated by (4.10).

We can further reduce the number of enumerations by exploiting symmetry in the problem. This gives the final form used in our actual numerical implementation. For a set

$\mathcal{S} \subseteq \mathbb{R}^n$, define its *sorted set* sorted($\mathcal{S}$) as

$$\text{sorted}(\mathcal{S}) \triangleq \{s\colon s_1 \geq s_2 \geq \cdots \geq s_n,\ \Pi s \in \mathcal{S} \text{ for some permutation matrix } \Pi\}.$$

Corollary 4.7. *The optimization problem*

$$\begin{aligned} \underset{p,\delta,m}{\text{maximize}} \quad & \sum_{\alpha \in \bar{A}(\delta,m)} f(p,\alpha) \\ \text{subject to} \quad & 0 \preceq p \preceq 1, \qquad \delta \in \bigcup_{A \subseteq \{0,1\}^n} \text{sorted}\left(\Delta(A)\right), \qquad m \in M(\delta) \\ & \delta^T p \geq t + m. \end{aligned} \tag{4.27}$$

achieves the same optimal value as problem (4.10).

Proof. Denote the optimal value of problem (4.21) as OPT_4 and that of problem (4.27) as OPT_5. To prove that $\text{OPT}_4 \geq \text{OPT}_5$, it suffices to show that

$$\bigcup_{A \subseteq \{0,1\}^n} \text{sorted}(\Delta(A)) \subseteq \bigcup_{A \subseteq \{0,1\}^n} \Delta(A). \tag{4.28}$$

Recall the definition of $\Delta(A)$. If $\bar{\delta} \in \text{sorted}(\Delta(A))$ for some $A \subseteq \{0,1\}^n$, then there must exist $\bar{\gamma}$ and a permutation matrix Π such that $(\Pi\bar{\delta}, \bar{\gamma})$ is a vertex of the polytope defined by the inequalities (in (δ, γ))

$$0 \preceq \delta \preceq 1, \qquad \alpha^T\delta - \gamma \geq 0, \quad \alpha \in A.$$

It is not difficult to verify that $(\bar{\delta}, \bar{\gamma})$ must be a vertex of the polytope defined by

$$0 \preceq \delta \preceq 1, \qquad \alpha^T\delta - \gamma \geq 0, \quad \alpha \in \Pi^T A.$$

It then follows that $\bar{\delta} \in \Delta(\Pi^T A)$ and hence $\bar{\delta} \in \bigcup_{A \subseteq \{0,1\}^n} \Delta(A)$, which implies (4.28).

On the other hand, suppose the optimal solution to problem (4.21) is (p^*, δ^*, m^*). First of all, there must exist a permutation matrix Π such that the entries of $\Pi\delta^*$ are in descending order. We will show that $(\Pi p^*, \Pi\delta^*, m^*)$ is a feasible solution to problem (4.27) and its corresponding objective is the same as OPT_4. For checking feasibility, we will only show that $M(\Pi\delta^*) = M(\delta^*)$ and hence $m^* \in M(\Pi\delta^*)$ since other constraints are not difficult to

verify. From the definition of $M(\delta)$, it can be shown that

$$\begin{aligned} M(\Pi\delta^*) &= \{j(\Pi\delta^*, A)\colon A \subseteq \{0,1\}^n\} = \{1^T\Pi\delta^* - \min_{\alpha\in A} \alpha^T\Pi\delta^*\colon A \subseteq \{0,1\}^n\} \\ &= \{1^T\delta^* - \min_{\bar{\alpha}\in\Pi^T A} \bar{\alpha}^T\delta^*\colon A \subseteq \{0,1\}^n\} = \{1^T\delta^* - \min_{\bar{\alpha}\in A} \bar{\alpha}^T\delta^*\colon A \subseteq \{0,1\}^n\} \\ &= M(\delta^*). \end{aligned}$$

It remains to show that $(\Pi p^*, \Pi\delta^*, m^*)$ yields the same objective as OPT_4. Note that

$$\begin{aligned} \bar{A}(\Pi\delta^*, m^*) &= \{\alpha \in \{0,1\}^n\colon 1^T\Pi\delta^* - \alpha^T\Pi\delta^* \le m^*\} \\ &= \{\alpha \in \{0,1\}^n\colon 1^T\delta^* - (\Pi^T\alpha)^T\delta^* \le m^*\} \\ &= \Pi^T\bar{A}(\delta^*, m^*). \end{aligned}$$

and $f(\Pi p^*, \alpha) = f(p^*, \Pi^T\alpha)$. Here, for any matrix $\Pi \in \mathbb{R}^{n\times n}$ and set $A \subseteq \mathbb{R}^n$ the notation ΠA represents the set $\{\Pi a\colon a \in A\}$, i.e., the element-wise action of Π on A. It then follows that

$$\mathrm{OPT}_5 \ge \sum_{\alpha\in\bar{A}(\Pi\delta^*,m^*)} f(\Pi p^*, \alpha) = \sum_{\alpha\in\Pi\bar{A}(\Pi\delta^*,m^*)} f(p^*, \Pi^T\alpha) = \sum_{\bar{\alpha}\in\bar{A}(\Pi\delta^*,m^*)} f(p^*, \bar{\alpha}) = \mathrm{OPT}_4,$$

which completes the proof. □

4.4 Additional computational issues

4.4.1 Generating the enumerations

In order to enumerate over all possible choices of δ and m in problem (4.27), one needs to compute the set

$$\bigcup_{A\subseteq\{0,1\}^n} \mathrm{sorted}(\Delta(A)). \tag{4.29}$$

One naive method is to enumerate over all 2^{2^n} subsets of $\{0,1\}^n$ and perform vertex enumeration (required for obtaining $\Delta(A)$) on each subset. However, this can be done more efficiently by exploiting structures in $\Delta(A)$. The first property that can be exploited is

permutation invariance of sorted($\Delta(A)$), i.e.,

$$\text{sorted}(\Delta(\Pi A)) = \text{sorted}(\Delta(A)) \quad \text{for any permutation } \Pi.$$

Note that the set of all permutations forms a group (the symmetric group). Ideas of reduction by exploiting invariance under group actions are abundant, and can be found in many areas such as polynomial equation solving [78, 22], optimization [23], numerical solution of partial differential equations [20], geometric mechanics [44], and dynamical systems [28]. According to this property, if sorted($\Delta(A)$) has already been included in (4.29) for some A, then it will be redundant to compute sorted($\Delta(\Pi A)$) for any permutation Π. Before dealing with this permutation invariance, we define an equivalence relationship on the power sets of $\{0,1\}^n$ as

$$A \sim A' \iff A = \Pi A' \quad \text{for some permutation } \Pi.$$

and the corresponding equivalence class

$$[A] \triangleq \{\Pi A\}.$$

If we can devise a representation ϕ that is permutation invariant, i.e., for any equivalent class $[A]$, there exists a set $\bar{A} \in [A]$ named the *canonical set* such that

$$\phi(A) = \bar{A}, \quad \forall A \in [A],$$

then we can enumerate over all canonical sets $\{\phi(A)\colon A \subseteq \{0,1\}^n\}$ and only compute sorted($\Delta(\phi(A))$) to avoid any redundant computation for permuted versions of A, since

$$\bigcup_{A \subseteq \{0,1\}^n} \text{sorted}(\Delta(A)) = \bigcup_{A \subseteq \{0,1\}^n} \text{sorted}(\Delta(\phi(A))).$$

One such representation can be obtained as follows. First of all, define the matrix representation of a set $A = \{\alpha_i\}_{i=1}^k \subseteq \{0,1\}^n$ as

$$\text{mat}(A) \triangleq \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k \end{bmatrix}.$$

Without loss of generality, we assume that $\alpha_1, \alpha_2, \ldots, \alpha_k$ are arranged in lexicographical order. Later, it will become clear that this particular arrangement does not affect the resulting permutation invariant representation. The matrix representation of two equivalent sets satisfies the following property:

$$A \sim A' \iff \text{mat}(A) = P\text{mat}(A')Q \quad \text{for some permutation matrices } P \text{ and } Q.$$

For example, consider two sets

$$A = \{(0,0,1),(0,1,0),(1,0,1)\}, \qquad A' = \{(0,0,1),(0,1,1),(1,0,0)\}.$$

They are equivalent since

$$A = \Pi A' \quad \text{for } \Pi = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

On the other hand, their matrix representations,

$$\text{mat}(A) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \qquad \text{mat}(A') = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

satisfy

$$\text{mat}(A) = P\text{mat}(A')Q^T \quad \text{for } P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } Q^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

A permutation invariant representation of a set A can be obtained from its matrix representation $\text{mat}(A)$ according to Algorithm 4.1 using singular value decomposition.

To understand Algorithm 4.1, consider two sets A_1 and A_2 that only differ by a permutation, i.e, $\text{mat}(A_1) = P\text{mat}(A_2)Q^T$. The singular value decompositions of their matrix

Algorithm 4.1 Permutation invariant representation

1. Obtain the singular value decomposition of $\mathrm{mat}(A) = U\Sigma V^T$, where the singular values appearing in Σ are arranged in descending order, i.e., $\sigma_1 \geq \sigma_2 \geq \dots$.
2. Find permutation matrices P and Q such that the columns of PU and QV are arranged in ascending lexicographical order.
3. Output the set consisting the columns of $P\mathrm{mat}(A)Q^T$ as the invariant representation of A.

representations obtained from step 1 of Algorithm 4.1,

$$\mathrm{mat}(A_1) = U_1\Sigma_1 V_1^T, \qquad \mathrm{mat}(A_2) = U_2\Sigma_2 V_2^T,$$

must satisfy $\Sigma_1 = \Sigma_2$, $U_1 = PU_2$, and $V_1 = QV_2$. This implies that there can be freedom in choosing U and V in the corresponding singular value decomposition up to permutations for sets in the same equivalence class. Step 2 in Algorithm 4.1 removes this freedom by enforcing the ordering of elements in U and V. It is worth noting that the approach will not guarantee a unique canonical representation in the presence of repeated singular values. However, in our numerical examples, this non-uniqueness is not found to be a major issue, and this method has reduced the number of enumerations satisfactorily.

Another property that can be exploited is to only enumerate over the sets in which all elements are minimal elements, which we will call *minimal sets*. Recall that an element $\alpha \in A$ is called a minimal element of A if for any $\alpha' \in A$, we have

$$\alpha' \leq \alpha \implies \alpha' = \alpha.$$

In other words, if a set $A \subseteq \{0,1\}^n$ is not minimal, i.e., it contains an element $\tilde{\alpha}$ that is not a minimal element, then there must exist $\alpha \in A$ such that $\alpha \leq \tilde{\alpha}$. For such set A, the inequality $\tilde{\alpha}^T\delta - \gamma \geq 0$ is redundant among all the inequalities that define $\Delta(A)$, since this inequality is implied by the fact that $\alpha^T\delta - \gamma \geq 0$ (by the definition of $\Delta(A)$), and $\tilde{\alpha}^T\delta - \gamma \geq \alpha^T\delta - \gamma$ for any $0 \preceq \delta \preceq 1$. Enumeration over such sets can be aided by a directed graph $G = (V, E)$ constructed as follows.

Definition 4.8. The set of vertices is the power set of $\{0,1\}^n$, i.e., $V = 2^{\{0,1\}^n}$. For any $A_1, A_2 \in V$, we have $(A_1, A_2) \in E$ if and only if A_2 is a minimal set and $A_2 = A_1 \sqcup \{\alpha\}$ for

some $\alpha \in \{0,1\}^n$. Here the symbol $\sqcup$ denotes disjoint union.

Note that the condition $(A_1, A_2) \in E$ implies that A_1 is also a minimal set, since removing any element from a minimal set preserves minimality. On the other hand, for any minimal A_1, in order to find all the edges originated from A_1, we can enumerate over all elements in $\{0,1\}^n \backslash A_1$ and check for each element α whether $A_2 = A_1 \sqcup \{\alpha\}$ is minimal. Checking this condition is not difficult, since it is equivalent to checking that α is a minimal element of A_2 and α does not make any element in A_1 non-minimal, which can be done by comparing α with every element in A_1.

It is worth noting that the permutation invariance mentioned previously can be naturally incorporated when constructing the graph. Specifically, the modified directed graph $G' = (V', E')$ needs to satisfy:

Definition 4.9. The set of vertices is the power set of $\{0,1\}^n$, i.e., $V' = 2^{\{0,1\}^n}$. For any $A_1', A_2' \in V'$, we have $(A_1', A_2') \in E'$ if any only if A_2' is a minimal set and $A_2' = \phi(A_1' \sqcup \{\alpha\})$ for some $\alpha \in \{0,1\}^n$.

It can be seen that the only modification is that we now require that A_2' is a canonical set in addition to being minimal. The following theorem guarantees that all minimal sets, most of which (modulo the ambiguities caused by repeated singular values) being canonical, can be found by traversing G' starting from the empty set $\emptyset$.

Theorem 4.10. *Let $G' = (V', E')$ be a graph that satisfies Definition* (4.9)*. Then a set A' is both minimal and canonical if and only if there exists a path from the empty set $\emptyset$ to A' in G'.*

Proof. ($\Longleftarrow$) Suppose there exists a path from $\emptyset$ to A', then there exists a set B' such that $(B', A') \in E'$. From Definition (4.9), we know that A' is minimal and also canonical due to the action of ϕ.

($\Longrightarrow$) Suppose A' is both minimal and canonical. Choose any $\alpha \in A'$ and define $B = A' \backslash \{\alpha\}$. It follows that B is also minimal since A is minimal. On the other hand, we have $A' = B \sqcup \{\alpha\}$ and

$$B \sqcup \{\alpha\} = \phi(B \sqcup \{\alpha\}),$$

since A' is canonical. Consider the set $\phi(B)$ and a permutation P such that $\phi(B) = PB$.

n	naive method (2^{2^n})	graph traversal
2	16	4
3	256	9
4	$65,536$	29
5	$4,294,967,296$	217
6	$\approx 1.8 \times 10^{19}$	$18,186$

Table 4.1: The number of sets involved in computing the union set (4.29) using two different methods. The actual number of sets needed (i.e., the number of minimal canonical sets) may be smaller than what is obtained by the graph traversal method, since the canonical representation obtained from singular value decomposition may be non-unique (see previous discussions for details).

We then have

$$A' = B \sqcup \{\alpha\} = \phi(B \sqcup \{\alpha\}) = \phi(PB \sqcup \{P\alpha\}) = \phi(\phi(B) \sqcup \{P\alpha\}).$$

In other words, there exists a canonical set $\phi(B)$ whose cardinality is $|A'| - 1$ such that

$$(\phi(B), A') \in E'.$$

Repeat the same argument on $\phi(B)$ until the cardinality becomes 0, which implies that the preceding set becomes the empty set $\emptyset$. This proves that there exists a path from $\emptyset$ to A'. $\square$

Table 4.1 shows the number of minimal canonical sets obtained by running a graph traversal algorithm (cf. [42]) on the graph $G' = (V', E')$ constructed as per Definition 4.9. For comparison, the number of sets obtained from the naive method is also listed. It can be seen that the number of enumerations has been greatly reduced. This shows the effectiveness of reduction using Corollary 4.7, even not counting the fact that each new subproblem has fewer variables and is less expensive to solve computationally.

Another thing to note is that the cardinality of the union set (4.29) is actually quite small even compared to the number of minimal canonical sets, which has already been greatly reduced from 2^{2^n}. Table 4.2 lists the cardinality of (4.29) as well as the total number of enumerations (over δ and m) generated by problem (4.27), the latter of which is also the number of polynomial optimization problems.

We do not have results for $n \geq 7$ in Table 4.1 and 4.2 because the graph G' becomes too large to traverse within reasonable time. For $n = 7$, the graph traversal has not stopped

n	$\left\lvert\bigcup_{A\subseteq\{0,1\}^n} \text{sorted}(\Delta(A))\right\rvert$	$\left\lvert\left\{(\delta, M(\delta)\colon \delta \in \bigcup_{A\subseteq\{0,1\}^n} \text{sorted}(\Delta(A))\right\}\right\rvert$ (i.e., number of subproblems)
2	3	3
3	4	6
4	6	15
5	12	57
6	43	420

Table 4.2: Cardinality of the union set (4.29) and the number of enumerations in problem (4.27).

#	element			
1	0	0	0	0
2	1.0000	0	0	0
3	1.0000	1.0000	0	0
4	1.0000	1.0000	1.0000	0
5	1.0000	0.5000	0.5000	0.5000
6	1.0000	1.0000	1.0000	1.0000

Table 4.3: Contents of the union set (4.29) for $n = 4$. Each row corresponds to one element in the set.

after more than one month on an Intel Xeon 3.0 GHz workstation (running on a single core). However, for $n \leq 6$, the union set (4.29) shows interesting patterns where all the entries of the elements in the set appear to be "simple" fractions up to numerical precision, as can be seen from Table 4.3, 4.4, and 4.5, in which the cases of $n = 4, 5, 6$ are listed (for example, 0.3333 can be considered as 1/3, which is a "simple" fraction). Therefore, it is possible that there might be a simpler method for computing the union set (4.29) if, for example, an expression underlying such patterns can be found.

4.4.2 Solving the polynomial optimization problem

In general, polynomial optimization problems are NP-hard to solve. However, these problems can be relaxed as sums-of-squares (SOS) optimization problems, which give a lower (upper) bound for the original minimization (maximization) problem. We now briefly introduce the SOS relaxation technique for polynomial optimization problems. More details can be found in [54, 39, 41, 14]. A polynomial optimization problem is one that has the form

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \geq 0, \quad i = 1, 2, \ldots, n_{in} \end{aligned} \tag{4.30}$$

#	element				
1	0	0	0	0	0
2	1.0000	0	0	0	0
3	1.0000	1.0000	0	0	0
4	1.0000	1.0000	1.0000	0	0
5	1.0000	0.5000	0.5000	0.5000	0
6	1.0000	1.0000	1.0000	1.0000	0
7	1.0000	0.3333	0.3333	0.3333	0.3333
8	1.0000	0.6667	0.3333	0.3333	0.3333
9	1.0000	0.6667	0.6667	0.3333	0.3333
10	1.0000	0.5000	0.5000	0.5000	0.5000
11	1.0000	1.0000	0.5000	0.5000	0.5000
12	1.0000	1.0000	1.0000	1.0000	1.0000

Table 4.4: Contents of the union set (4.29) for $n = 5$. Each row corresponds to one element in the set.

$$h_i(x) = 0, \quad i = 1, 2, \ldots, n_{eq},$$

where f, $\{g_i\}$, and $\{h_i\}$ are all polynomials. A polynomial p is called SOS if it can be written as $p(x) = \sum_i p_i^2(x)$ for certain polynomials $\{p_i\}$. Equivalently, an SOS polynomial p of given degree $2d$ can always be written as the following quadratic form:

$$p(x) = v^T(x) Q v(x), \tag{4.31}$$

where Q is a positive semidefinite matrix and $v(x)$ is the vector of all the monomials with degree at most d.

If we are able to find a scalar γ, polynomials $\{\mu_i\}$, and SOS polynomials $\sigma_0, \{\sigma_i\}_{i=1}^{n_{in}}$ such that

$$f(x) - \gamma = \sigma_0(x) + \sum_{i=1}^{n_{eq}} h_i(x)\mu_i(x) + \sum_{i=1}^{n_{in}} g_i(x)\sigma_i(x),$$

then γ is guaranteed to be a lower bound for the optimal value of problem (4.30), since $f(x) - \gamma \geq 0$ whenever $g_i(x) \geq 0$ for all $i = 1, 2, \ldots, n_{in}$ and $h_i(x) = 0$ for all $i = 1, 2, \ldots, n_{eq}$. The lower bound will continue to approach the actual optimal value as the degrees of σ_0, $\{\sigma_i\}$, and $\{\mu_i\}$ grows, although the gap is not guaranteed to become zero at any finite degree. Under certain conditions (see Putinar's *positivstellenstaz* [57]), the gap will shrink down to zero at some finite degree. Therefore, we can restrict the search for polynomials to some maximum degree $2d$ and keep increasing d until the relaxation is exact. For any fixed d,

#	element					
1	0	0	0	0	0	0
2	1.0000	0	0	0	0	0
3	1.0000	1.0000	0	0	0	0
4	1.0000	1.0000	1.0000	0	0	0
5	1.0000	0.5000	0.5000	0.5000	0	0
6	1.0000	1.0000	1.0000	1.0000	0	0
7	1.0000	0.3333	0.3333	0.3333	0.3333	0
8	1.0000	0.6667	0.3333	0.3333	0.3333	0
9	1.0000	0.6667	0.6667	0.3333	0.3333	0
10	1.0000	0.5000	0.5000	0.5000	0.5000	0
11	1.0000	1.0000	0.5000	0.5000	0.5000	0
12	1.0000	1.0000	1.0000	1.0000	1.0000	0
13	1.0000	0.8000	0.4000	0.4000	0.2000	0.2000
14	1.0000	0.8000	0.6000	0.4000	0.2000	0.2000
15	1.0000	0.4000	0.4000	0.4000	0.2000	0.2000
16	1.0000	0.6000	0.4000	0.4000	0.2000	0.2000
17	1.0000	0.6000	0.6000	0.2000	0.2000	0.2000
18	1.0000	0.8000	0.6000	0.4000	0.4000	0.2000
19	1.0000	0.6000	0.6000	0.4000	0.2000	0.2000
20	1.0000	0.4000	0.4000	0.2000	0.2000	0.2000
21	1.0000	0.6000	0.4000	0.2000	0.2000	0.2000
22	1.0000	0.7500	0.5000	0.5000	0.2500	0.2500
23	1.0000	0.5000	0.5000	0.2500	0.2500	0.2500
24	1.0000	0.7500	0.7500	0.5000	0.5000	0.2500
25	1.0000	0.7500	0.5000	0.2500	0.2500	0.2500
26	1.0000	0.7500	0.7500	0.2500	0.2500	0.2500
27	1.0000	0.2500	0.2500	0.2500	0.2500	0.2500
28	1.0000	0.5000	0.2500	0.2500	0.2500	0.2500
29	1.0000	0.7500	0.2500	0.2500	0.2500	0.2500
30	1.0000	0.7500	0.7500	0.5000	0.2500	0.2500
31	1.0000	0.6667	0.6667	0.3333	0.3333	0.3333
32	1.0000	1.0000	0.6667	0.3333	0.3333	0.3333
33	1.0000	1.0000	0.3333	0.3333	0.3333	0.3333
34	1.0000	0.6667	0.3333	0.3333	0.3333	0.3333
35	1.0000	1.0000	0.6667	0.6667	0.3333	0.3333
36	1.0000	0.3333	0.3333	0.3333	0.3333	0.3333
37	1.0000	0.6667	0.6667	0.6667	0.3333	0.3333
38	1.0000	1.0000	0.6667	0.6667	0.6667	0.3333
39	1.0000	1.0000	0.5000	0.5000	0.5000	0.5000
40	1.0000	1.0000	1.0000	0.5000	0.5000	0.5000
41	1.0000	0.5000	0.5000	0.5000	0.5000	0.5000
42	1.0000	1.0000	0.6667	0.6667	0.6667	0.6667
43	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 4.5: Contents of the union set (4.29) for $n = 6$. Each row corresponds to one element in the set.

we wish to obtain the best possible lower bound γ. Formally, this can be written as the following optimization problem

$$\begin{aligned} \underset{\gamma}{\text{maximize}} \quad & \gamma \\ \text{subject to} \quad & f(x) - \gamma = \sigma_0(x) + \sum_{i=1}^{n_{eq}} h_i(x)\mu_i(x) + \sum_{i=1}^{n_{in}} g_i(x)\sigma_i(x) \\ & \sigma_0, \sigma_1, \ldots, \sigma_{n_{in}} \text{ are SOS.} \end{aligned} \tag{4.32}$$

Note that the last constraint in problem (4.32) corresponds to a series of positive semidefinite constraints given by (4.31) and hence this optimization problem is a semidefinite program (SDP). As a (free) by-product of the relaxation, a solution that achieves the relaxed objective value can also be obtained. In general, this solution is infeasible since it yields a smaller (larger) objective value than the optimal value of the corresponding minimization (maximization) problem. However, if the solution turns out to be feasible, it implies that the relaxation is exact and the solution is an optimal solution of the original optimization problem.

In this work, the SOS relaxation is numerically carried out by SparsePOP [75, 76], which is a free software package that not only forms and solves the SOS relaxation (up to a given order) of a polynomial optimization problem, but also checks the exactness of the relaxation. Empirically, we find that it suffices to restrict the order to be $\lfloor n/2 + 1 \rfloor$ in order to obtain an exact SOS relaxation when solving problem (4.26) for any given n. Recall that the order in SOS relaxation is defined as the maximum degree of the monomial basis. The number of monomial basis for n variables with degree at most d is

$$N = \binom{n+d}{d}.$$

We list the number of monomials basis for solving problem (4.26) for different n in Table 4.6. If no structure in the polynomial optimization problem and its SOS relaxation is exploited, the size of the corresponding SDP (i.e., dimension of the positive semidefinite matrices) is determined by the number of monomial basis. In our case, the SOS relaxation becomes quite expensive to solve starting from $n = 7$ and prohibitive beyond $n = 9$. Aside from the quick growth in problem size, another major reason behind this difficulty in solving larger problems

n	2	3	4	5	6	7	8	9
$N = \binom{n+d}{d}$	6	10	35	56	210	330	1287	2002

Table 4.6: The number of monomial basis used in the sums-of-squares relaxation of problem (4.26). The degree $d = \lfloor n/2 + 1 \rfloor$.

is that the solvers (SeDuMi [70] and SDPT3 [71, 73]) use iterative second-order methods, where the complexity of each iteration is $O(N^6)$ (cubic in the number of variables and the number of variables is N^2). It is possible that first-order methods may partially alleviate this difficulty by reducing the complexity of each iteration in order to solve problems with larger n, although it remains unclear since first-order methods often take a lot more steps to converge.

4.5 Results

We now show some numerical results on the upper bound for the probability of deviation (4.1) obtained by solving the optimization problem (4.9) using the enumeration method from Section 4.3 and sums-of-squares relaxation from Section 4.4. For all numerical results presented below, the corresponding SOS relaxations are exact and therefore all numerical bounds are tight. In addition, recall that any feasible solution of problem (4.15) will yield a lower bound for (4.1). In particular, if we choose $A = \{(1, 1, \ldots, 1)\}$, $\delta = (1, 1, \ldots, 1)$ (it is not difficult to verify that $\delta \in \Delta(A)$), and optimize over p, the optimization problem (4.15) becomes

$$\begin{aligned} \underset{p}{\text{maximize}} \quad & \prod_{i=1}^{n}(1 - p_i) \\ \text{subject to} \quad & 0 \preceq p \preceq 1, \qquad \sum_{i=1}^{n} p_i \geq t, \end{aligned} \tag{4.33}$$

whose optimal value can be determined analytically as $(1 - t/n)^n$ from the inequality of arithmetic and geometric means. Therefore, we have a lower bound given by

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} \mathbb{E}X_i + t\right) \geq (1 - t/n)^n. \tag{4.34}$$

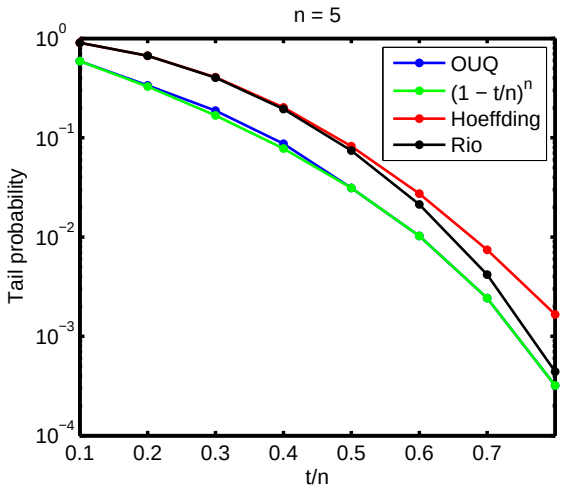


Figure 4.1: Comparison of bounds under different $\bar{t} = t/n$ for $n = 5$.

In fact, this lower bound has been shown to be tight for $n = 2$ and $n = 3$ [53]. Fig. 4.1 and 4.2 compare the bounds obtained using different approaches, including the upper bounds from OUQ, the upper bounds from Hoeffding's inequality, the lower bounds given by (4.34), and another analytical upper bound obtained by Rio [58]:

$$\mathbb{P}\left(\frac{1}{n}\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathbb{E}X_i\right) \geq \bar{t}\right) \leq (1-\bar{t})^{n(2\bar{t}-\bar{t}^2)}, \tag{4.35}$$

which holds for any $\bar{t} \in [0, 1]$. This bound is obtained using martingale decomposition, and is tighter than Hoeffding's inequality. Fig. 4.1 shows how the bounds change under different (normalized) deviation $\bar{t} = t/n$. Since each $X_i \in [0, 1]$, we have $\sum_{i=1}^{n} X_i/n \leq 1$ and $\sum_{i=1}^{n} \mathbb{E}X_i/n \geq 0$, and hence

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} \mathbb{E}X_i + t\right) = 0$$

for $t \geq n$ (i.e., $\bar{t} \geq 1$). Therefore, the lower bound (4.34) is tight at $t/n = 1$, whereas Hoeffding's upper bound is nonzero and hence not tight at $t/n = 1$. Moreover, the numerical bound matches the lower bound (4.34) surprisingly well for $0 < t/n < 1$, although we do not have an explanation for this at this moment. Fig. 4.2 shows how the bounds change with the number of samples n. Initially, at least for $n = 2$ and $n = 3$, the numerical bound matches the lower bound (4.34) as expected. Starting from $n = 5$, the difference

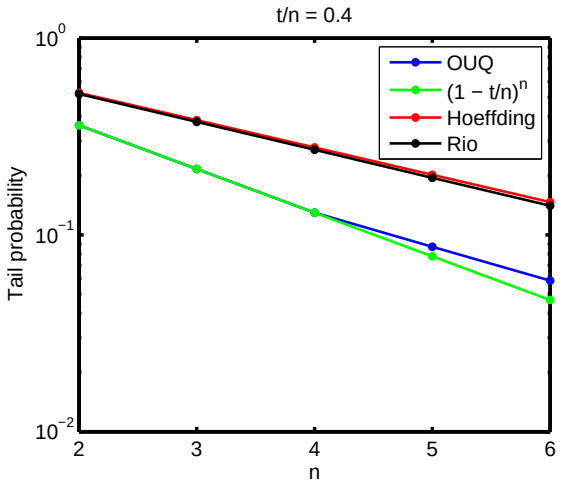


Figure 4.2: Comparison of bounds under different n for $\bar{t} = t/n = 0.4$.

between the numerical bound and the lower bound (4.34) becomes visible. Note that the logarithms of both Hoeffding's bound and the lower bound (4.34) vary linearly with n, only with different slopes (the lower bound curve is steeper). As n grows, from the central limit theorem, we know that Hoeffding's bound will eventually give the correct slope (modulo the slowly varying factor that can be ignored for large n), it is expected that the slope of the numerical bound will gradually decrease until it is approximately the same as that of Hoeffding's bound (i.e., the two curves become approximately parallel). In all cases, OUQ gives significantly better bounds than either Hoeffding's inequality or Rio's formula.

4.6 Conclusions

This chapter applies the optimal uncertainty quantification framework to obtain a bound for the probability that the sum of independent random variables deviates from the sum of their means, a setting traditionally presented along with Hoeffding's inequality. In particular, the optimal uncertainty quantification framework attempts to obtain a better bound than the one given by Hoeffding's inequality. After finite reduction, the corresponding optimization problem is equivalent to doubly exponentially many non-convex polynomial optimization problems. Several structures are exploited to cut down the number of problems without introducing much computational burden. For solving the polynomial optimization problems, sums-of-squares relaxation techniques are used. Although these relaxation techniques can only guarantee a upper bound for the optimal value, they are found to give exact solutions in

all cases we tested and hence tight bounds for the probability of interest. Numerical results show that the optimal uncertainty quantification framework tend to give significantly better bound than Hoeffding's inequality in the regime of small number of independent random variables.

Chapter 5

Concluding Remarks

5.1 Summary

This thesis emphasizes on developing efficient computational methods for optimal uncertainty analysis (OUQ) using convex optimization and relaxation. It begins with cases for which there exists an equivalent convex formulation. By viewing the OUQ problem from either its primal form or dual form, we are able to derive conditions on the objective function and information constraints under which a convex formulation exists. It is shown that a convex formulation exists if the objective function is piecewise convex and the inequality information constraints are piecewise concave. In addition, from the dual form, it is shown that a different convex formulation also exists in the univariate case if the objective is piecewise polynomial and the information constraints are polynomial. The new results subsume some of the existing work by others and can potentially provide better quantification results by allowing more freedom in incorporating knowledge about the distribution.

For the purpose of demonstrating the application of convex optimization in OUQ, the thesis proceeds by presenting a case study using the example of energy storage placement in power grids. After reformulation, the problem becomes a semidefinite program with (exponentially) many linear matrix inequality constraints that are mostly inactive. Due its size, such a convex optimization problem is still difficult to solve numerically. Motivated by this challenge, the thesis attempts to address the scaling issues by exploiting specific structures in the problem. We notice that the objective function is defined by the optimal value of a linear program and can be rewritten in the *polytopic canonical form*. This special property allows quick elimination of inactive constraints in the original large problem through iterations. As a result, the optimal solution may be obtained by solving a sequence of smaller problems,

each of which only contains the candidate active constraints. We also investigate cheaper numerical solutions to the semidefinite program when the size of the matrix becomes large, since second-order methods used by most off-the-shelf solvers are too expensive to apply. By exploiting the special form of the constraints, we are able to rewrite the problem in a form that is readily solvable using first-order methods such as the alternating direction method of multipliers (ADMM). Various numerical experiments show that ADMM can greatly reduce the running time, particularly when the size of the matrix becomes large.

Lastly, the thesis investigates the application of convex relaxations in OUQ. We choose Hoeffding's inequality as the example due to its wide usage in many areas such as adaptive stochastic optimization. Relaxation fits naturally in the context of OUQ since it is aligned with the original purpose of obtaining bounds for some quantity of interest. After finite reduction, the optimization problem becomes equivalent to doubly exponentially many non-convex polynomial optimization problems. Before applying convex relaxation to each polynomial optimization problem, a number of technical steps are also applied in order to reduce the number of problems for tractability. Bounds for the polynomial optimization problems are then obtained using sums-of-squares relaxation. Surprisingly, for all the cases tested, the relaxations are found to be exact and hence tight bounds are actually obtained. Numerical results show that the OUQ framework tend to give significantly better bound than Hoeffding's inequality in the regime of small number of independent random variables.

5.2 Future directions

Optimal uncertainty quantification and, more generally, the field of uncertainty quantification is a rich area for research. This section lists a few future directions as immediate extensions of this thesis.

From analysis to decision making

Aa an extension of OUQ, another related question to consider is how to make the optimal decision under stochastic uncertainties whose probability distribution is only partially known. Formally, this corresponds to the following optimization problem:

$$\underset{u}{\text{minimize}} \quad \max_{\mathcal{D}\in\Delta} \mathbb{E}_{\theta\sim\mathcal{D}}\left[f(u,\theta)\right], \tag{5.1}$$

where the set Δ imposes the information constraints on the distribution $\mathcal{D}$ such as (2.2) and (2.3) in the OUQ problem (2.1). Therefore, the OUQ problem can be viewed as the inner-loop optimization for problem (5.1). In the special case where the underlying stochasticity can be modeled by a (discrete) Markov process whose transition transition probability is not exactly known, the problem has been studied by Nilim et al. [49] and Xu et al. [79].

Problems in the form of (5.1) are often referred to as *distributionally robust stochastic optimization*. It is a generalization of robust optimization: when the set Δ only imposes constraints on the support of $\mathcal{D}$, problem (5.1) is reduced to robust optimization. This type of robust optimization problem was first proposed in the 1950s by Scarf [62] in the context of inventory optimization, and has recently attracted attention of many researchers [65, 19, 26, 9, 80]. In particular, Delage and Ye [19] have shown that, when the objective function f is bi-affine in u and θ, and the set Δ only imposes constraints on the first and second moments of $\mathcal{D}$, problem (5.1) can be reformulated as a semidefinite program. It remains an open question whether this can be combined with the computational speed-ups studied in Chapter 3 of this thesis.

Further acceleration of optimization algorithms

One nice consequence of having a convex formuation of an optimization problem is that every possible effort trying to speed up the optimization algorithm will never affect the quality of the solution, since the algorithm will always return the global optimum once it converges. As mentioned in Section 3.2, despite their savings on computational complexity for each iteration, first-order methods such as ADMM often suffer from long convergence time. Normally, these methods converge at the rate of $O(1/t)$, where t is the number of iterations. Here the convergence rate is defined as the dependence of error on the number of iterations (as opposed to the relationship between error at the current iteration with the one at the previous iteration, which is another commonly used concept).

For gradient descent, Nesterov proposed an algorithm that is able to speed up the convergence rate to $O(1/t^2)$, which is often referred to as Nesterov's accelerated gradient descent [48]. Goldstein et al. have applied Nesterov's accelerated scheme to ADMM and shown speed improvement for quadratic programs from numerical experiments [27]. However, such acceleration does not offer strong theoretical convergence guarantees. In fact, based on our numerical experiments on solving the OUQ problem, this accelerated ADMM often con-

verges slower than the original ADMM and sometimes fails to converge. This might be due to the fact that the optimization problem for OUQ is a semidefinite program rather than a quadratic program. Since semidefinite programs are ubiquitous, applying acceleration to ADMM should be an iteresting subject for further investigation.

Dealing with identical distributions

Hoeffding's inequality does not assume that the independent variables $\{X_i\}_{i=1}^n$ are indentically distributed. In the case where the random variables are i.i.d., Hoeffding's inequality is still often used since it has already been a loose bound even without the assumption of identical distribution, and the bound expression is simple. Computing a tight bound within the OUQ framework, however, turns out to be much more difficult if such assumption is incoporated. This is because imposing the constraint of identical distribution will destroy the finite reduction property of OUQ. One workaround is to replace the constraint of identical distribution with the constraint of identical moments (up to a certain order) [52]. This is expected to yield a better bound than only imposing the independence constraint. Since the number of Dirac masses will change as new constraints are added, the whole procedure of removing the redundant enumerations as presented in Section 4.3 needs to be revisited, and it is unclear whether the resulting optimization problem is still tractable.

Connection with existing work in information theory

In the 1970s, Smith studied the problem of the information capacity of amplitude- and variance-constrained scalar Gaussian channels [69]. To determine this, it requires solving an optimization problem where one tries to maximize the mutual information over all valid input probability distributions that satisfy the amplitude and variance constraints. Later, similar problems are also studied, with different assumptions on the channel noise and/or constraints on the input distribution [64, 4, 63]. For many instances of this type of problem, it has been shown that the optimal input distribution that achieves the capacity is always a discrete distribution. Unlike the finite reduction property that is present in OUQ, the number of Dirac masses in the capacity-achieving distribution cannot be determined directly from the constraints imposed on the distribution. It remains an interesting topic to search for a unifying framework that connects these two similar results.

Bibliography

[1] Power systems test case archive. `http://www.ee.washington.edu/research/pstca/pf14/pg_tca14bus.htm` (retrieved: October 10, 2013).

[2] Various memos and e-mails. `http://www.jsc.nasa.gov/news/columbia/var_memos_emails.pdf` (retrieved: October 15, 2013).

[3] Wind power / AIL data: Historical wind data (January 1, 2004 - December 31, 2009). `http://www.aeso.ca/gridoperations/20544.html` (retrieved: October 10, 2013).

[4] I. C. Abou-Faycal, M. D. Trott, and S. Shamai. The capacity of discrete-time memoryless rayleigh-fading channels. *Information Theory, IEEE Transactions on*, 47(4):1290–1301, 2001.

[5] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables (9th printing)*, chapter 7.2, pages 299–300. Dover, 1972.

[6] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski. A learning algorithm for boltzmann machines. *Cognitive science*, 9(1):147–169, 1985.

[7] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3):235–256, 2002.

[8] D. Avis and K. Fukuda. A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. *Discrete & Computational Geometry*, 8(1):295–313, 1992.

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

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

[11] D. Bertsimas and J. N. Tsitsiklis. *Introduction to linear optimization*. Athena Scientific, Belmont, Massachusetts, 1997.

[12] Å. Björck. *Numerical methods for least squares problems*. SIAM, 1996.

[13] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.

[14] G. Blekherman, P. A. Parrilo, and R. R. Thomas. *Semidefinite optimization and convex algebraic geometry*. SIAM, 2013.

[15] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011.

[16] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[17] S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, 5(1):1–122, 2012.

[18] N. L. Carothers. *Real analysis*. Cambridge University Press, 2000.

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

[20] A. Fässler and E. Stiefel. *Group theoretical methods and their applications*. Birkhäuser, Boston, 1992.

[21] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2(1):17–40, 1976.

[22] K. Gatermann. *Computer algebra methods for equivariant dynamical systems*, volume 1728 of *Lecture Notes in Mathematics*. Springer, 2000.

[23] K. Gatermann and P. A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1):95–128, 2004.

[24] D. Gayme and U. Topcu. Optimal power flow with distributed energy storage dynamics. In *American Control Conference*, 2011.

[25] H. J. Godwin. On generalizations of tchebychef's inequality. *Journal of the American Statistical Association*, 50(271):923–945, 1955.

[26] J. Goh and M. Sim. Distributionally robust optimization and its tractable approximations. *Operations Research*, 58(4):902–917, 2010.

[27] T. Goldstein, B. O'Donoghue, and S. Setzer. Fast alternating direction optimization methods. Technical report, UCLA Computational and Applied Mathematics, 2012.

[28] M. Golubitsky and D. Schaeffer. *Singularities and Groups in Bifurcation Theory II*, volume 69 of *Applied Mathematical Sciences*. Springer, 1985.

[29] S. Han, U. Topcu, M. Tao, H. Owhadi, and R. M. Murray. Convex optimal uncertainty quantification: Algorithms and a case study in energy storage placement for power grids. In *American Control Conference*, 2013.

[30] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[31] A. Ipakchi and F. Albuyeh. Grid of the future. *Power and Energy Magazine, IEEE*, 7(2):52–62, 2009.

[32] K. Isii. On a method for generalizations of Tchebycheff's inequality. *Annals of the Institute of Statistical Mathematics*, 10(2):65–88, 1959.

[33] K. Isii. On sharpness of Tchebycheff-type inequalities. *Annals of the Institute of Statistical Mathematics*, 14(1):185–197, 1962.

[34] K. Isii. Inequalities of the types of chebyshev and cramér-rao and mathematical programming. *Annals of the Institute of Statistical Mathematics*, 16(1):277–293, 1964.

[35] E. T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620, 1957.

[36] E. T. Jaynes. Information theory and statistical mechanics. II. *Physical Review*, 108(2):171, 1957.

[37] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[38] D. Kollar and N. Friedman. *Probabilistic graphical models: principles and techniques*. The MIT Press, 2009.

[39] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[40] J. B. Lasserre. Bounds on measures satisfying moment conditions. *The Annals of Applied Probability*, 12(3):1114–1137, 2002.

[41] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging applications of algebraic geometry*, pages 157–270. Springer, 2009.

[42] C. E. Leiserson, R. L. Rivest, C. Stein, and T. H. Cormen. *Introduction to algorithms*. The MIT Press, 2001.

[43] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *IEEE International Symposium on Computer Aided Control Systems Design*, pages 284–289. IEEE, 2004.

[44] J. E. Marsden and T. S. Ratiu. *Introduction to mechanics and symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer, 1999.

[45] A. W. Marshall and I. Olkin. Multivariate Chebyshev inequalities. *The Annals of Mathematical Statistics*, pages 1001–1014, 1960.

[46] H. P. Mulholland and C. A. Rogers. Representation theorems for distribution functions. *Proceedings of the London Mathematical Society*, 3(2):177–223, 1958.

[47] J. U. Nair. *Scheduling for heavy-tailed and light-tailed workloads in queueing systems*. PhD thesis, California Institute of Technology, 2012.

[48] Y. Nesterov. A method of solving a convex programming problem with convergence rate $O(1/k^2)$. In *Soviet Mathematics Doklady*, volume 27, pages 372–376, 1983.

[49] A. Nilim and L. El Ghaoui. Robust control of markov decision processes with uncertain transition matrices. *Operations Research*, pages 780–798, 2005.

[50] J. Nocedal and S. Wright. *Numerical optimization*. Springer, 2006.

[51] I. Olkin and J. W. Pratt. A multivariate tchebycheff inequality. *The Annals of Mathematical Statistics*, 29(1):226–234, 1958.

[52] H. Owhadi. personal communication.

[53] H. Owhadi, C. Scovel, T. Sullivan, M. McKerns, and M. Ortiz. Optimal uncertainty quantification. *SIAM Review*, 55(2):271–345, 2013.

[54] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, 2000.

[55] I. Popescu. A semidefinite programming approach to optimal-moment bounds for convex classes of distributions. *Mathematics of Operations Research*, 30(3):632–657, 2005.

[56] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans. Usefulness of DC power flow for active power flow analysis. In *Power Engineering Society General Meeting*, 2005.

[57] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.

[58] E. Rio. On mcdiarmid's concentration inequality. *Electronic Communications in Probability*, 18:1–11, 2013.

[59] B. P. Roberts and C. Sandberg. The role of energy storage in development of smart grids. *Proceedings of the IEEE*, 99(6):1139–1144, 2011.

[60] R. T. Rockafellar. Lagrange multipliers and optimality. *SIAM Review*, pages 183–238, 1993.

[61] S. M. Samuels. On a chebyshev-type inequality for sums of independent random variables. *The Annals of Mathematical Statistics*, pages 248–259, 1966.

[62] H. Scarf, K. J. Arrow, and S. Karlin. A min-max solution of an inventory problem. *Studies in the Mathematical Theory of Inventory and Production*, 10:201–209, 1958.

[63] V. Sethuraman and B. Hajek. Capacity per unit energy of fading channels with a peak constraint. *Information Theory, IEEE Transactions on*, 51(9):3102–3120, 2005.

[64] S. Shamai and I. Bar-David. The capacity of average and peak-power-limited quadrature gaussian channels. *Information Theory, IEEE Transactions on*, 41(4):1060–1071, 1995.

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

[66] A. Shapiro and A. P. Ruszczyński. *Stochastic programming*. Elsevier, 2003.

[67] B. Simon. *Convexity: an analytic viewpoint*, volume 187. Cambridge University Press, 2011.

[68] A. E. Sjödin, D. F. Gayme, and U. Topcu. Risk-mitigated optimal power flow with high wind penetration. In *American Control Conference*, 2012.

[69] J. G. Smith. The information capacity of amplitude- and variance-constrained scalar gaussian channels. *Information and Control*, 18(3):203–219, 1971.

[70] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11(1-4):625–653, 1999.

[71] K.-C. Toh, M. J. Todd, and R. H. Tütüncü. SDPT3—a matlab software package for semidefinite programming, version 1.3. *Optimization Methods and Software*, 11(1-4):545–581, 1999.

[72] L. N. Trefethen and D. Bau III. *Numerical linear algebra*. SIAM, 1997.

[73] R. H. Tütüncü, K. C. Toh, and M. J. Todd. Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical Programming*, 95(2):189–217, 2003.

[74] L. Vandenberghe, S. Boyd, and K. Comanor. Generalized Chebyshev bounds via semidefinite programming. *SIAM Review*, 49(1):52, 2007.

[75] H. Waki, S. Kim, M. Kojima, and M. Muramatsu. Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, 17(1):218–242, 2006.

[76] H. Waki, S. Kim, M. Kojima, M. Muramatsu, and H. Sugimoto. SparsePOP—a sparse semidefinite programming relaxation of polynomial optimization problems. *ACM Transactions on Mathematical Software*, 35(2):15, 2008.

[77] A. Wald. Statistical decision functions. 1950.

[78] P. A. Worfolk. Zeros of equivariant vector fields: Algorithms for an invariant approach. *Journal of Symbolic Computation*, 17(6):487–511, 1994.

[79] H. Xu and S. Mannor. Distributionally robust markov decision processes. *Mathematics of Operations Research*, 2012.

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