DISTRIBUTIONALLY ROBUST OPTIMIZATION IN CONTEXT OF DATA-DRIVEN PROBLEMS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Erick Hans Delage
June 2009

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Yinyu Ye) Principal Advisor

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Stephen P. Boyd)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Shie Mannor)

Approved for the University Committee on Graduate Studies.

Abstract

A decision maker is often confronted with situations where he has incomplete knowledge about some of the parameters of the problem that he is addressing: *e.g.*, market demand, future cost of merchandise, or quality of the available resources. While a simple approach that lets these parameters take their most likely values might lead to good solutions, it can easily lead to unexpected results in general. To avoid unnecessary deceptions, one should instead modify his decision model so that it accounts for the uncertainty that is present.

At a conceptual level, stochastic programming is an effective approach since it leads to decisions that directly trades off risk and performance. This is done by requiring that the decision maker formulate a distribution describing the probability that the parameters take any given value. The framework then offers an exhaustive array of risk measures to choose from for evaluating the performance of a candidate solution. Unfortunately, while this decision problem can generally be solved efficiently, the resulting "optimal risk-sensitive solution" can be misleading in applications where there is ambiguity in the choice of a distribution for representing the uncertain parameters. This is for instance the case in data-driven problems, where information about the distribution of parameters is mostly derived from the observation of historical realizations. After a limited amount of observations, a decision maker is often unable to fully determine the underlying distribution. This leads to three important questions for making a stochastic programming approach practical:

1. How accurate is a distribution estimate that is based only on a finite set of samples?
2. Can we account for distribution uncertainty when solving a stochastic program?
3. What is the computational cost associated with taking this uncertainty into account?

This thesis develops new theoretical foundations for a quantitative methodology that provides answers to these questions and can be used in a wide range of applications.

We first address these questions from a frequentist point of view. In doing so, we initially derive a new confidence region for the covariance matrix of a random vector given a finite set of independent and identically distributed samples. Unlike typical confidence regions, this new one constrains the terms of the matrix jointly with respect to their effect on the variance of any one dimensional projection of the random vector. The result allows us to define a "well structured" set of distributions that is guaranteed with high probability to contain the distribution from which the samples were drawn.

With the objective of accounting for such distribution uncertainty, we analyze the computational difficulty related to solving the distributionally robust form of the stochastic program. This model suggests making decisions according to the worst distribution in the uncertainty set. Although the model has been widely studied since Scarf introduced it in 1958, we are the first to use duality theory to show that for cases where the objective is convex in the decision variables and "piecewise concave" in the parameters, there exists a polynomial time algorithm that can solve the problem to any level of accuracy. In fact, we show that, for some of these models, the distributionally robust form of the stochastic program is much easier to solve than its original form. This analysis leads to a framework that provides reliable solutions to data-driven problems. Many applications can benefit from this framework, including both a fleet mix optimization problem and a portfolio selection problem. In particular, experiments with real stock market data confirm that this approach accounts more effectively for uncertainty in the future value of financial assets.

In the final part of this work, we choose to adopt a Bayesian approach. In this case, it is well known that one can choose to represent distribution uncertainty with a hyper-distribution (*i.e.*, a distribution over distributions). After formulating the distributionally robust problem in this context, we contrast the robust approach to a percentile criterion (*a.k.a.*, value-at-risk) which is more natural to consider. A study of Markov decision processes leads to interesting new results on the computational difficulties related to solving this percentile optimization problem. More specifically, we show that while this problem can be solved efficiently for a class of cost uncertainty, it can also lead to NP-hard problems. After providing an approximation algorithm for the intractable form, we perform a comparative study of how the robust and percentile based methods exploit the distribution information present in a machine replacement problem.

To my wife – Joyce Bernas Delage.
And to my parents – Jean Delage and Ginette Prud'homme.
Their unconditional love and support made it all possible.

Acknowledgements

First, I would like to thank my advisor, Yinyu Ye, for his invaluable guidance and constant enthusiasm for this research. I cannot overstate my gratitude and appreciation for his encouragement, support and accessibility. Secondly, I want to express my gratitude to Shie Mannor who has been a precious mentor through my years as a graduate student. Our discussions also played a determinant role in developing my passion for uncertainty and its influences on decision making.

I also want to thank Stephen Boyd for sharing his passion and knowledge of convex optimization with such generosity. I am thankful to him and Sunil Kumar for being on my defense committee. I also need to thank John G. Carlsson, Amir Dembo, Anthony Man-Cho So, Alexander Shapiro, Melvyn Sim, Benjamin Armbruster, Shipra Agrawal, Zizhuo Wang, and Zhisu Zhu for valuable discussions on topics related to this thesis.

This thesis would not have been possible without the financial support provided by the Fonds Québécois de la Recherche sur la Nature et les Technologies and by the The Boeing Company. I wish to specifically thank Sharon Arroyo for her interest in this work and her efforts in helping me demonstrate that the theory that I developed can benefit real and important practical problems.

I am grateful to my friends, Laura Diaz, Grace Gao, Themis Mastorides, Girish Mulye, David Varodayan, Ocie Ward, Robbie Yan, Argyris Zymnis, and many others, who made my stay at Stanford an extremely enjoyable period of my life. And, I am just as grateful to my friends from Montreal, especially Pierre-Luc Beauchesne, who always made sure I would not feel too far from home.

Finally, I would like to thank my wife, Joyce, my parents, Jean and Ginette, and my sister Geneviève, for their unconditional love, support, and encouragement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Stochastic programming can effectively describe many decision making problems in uncertain environments. For instance, given that one is interested in solving an optimization problem of the type

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad h(\boldsymbol{x},\boldsymbol{\xi}) \ \ ,$$

where $\mathcal{X}$ is a **convex** set of feasible solutions and $h(\boldsymbol{x},\boldsymbol{\xi})$ is a **convex** cost function in $\boldsymbol{x}$ that depends on some vector of parameters $\boldsymbol{\xi}$, it is often the case that at the time of optimization, the parameters have not yet been fully resolved. For example, an investment manager cannot know the exact return for all available securities, or in a different context, a manufacturing manager might not know the exact amount of future demand.

If one chooses to represent his uncertainty about $\boldsymbol{\xi}$ through a distribution F, then he can instead resort to minimizing the expected cost. This leads to solving a stochastic program:

$$\text{(SP)} \qquad \underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \mathbb{E}\left[h(\boldsymbol{x},\boldsymbol{\xi})\right] \ \ ,$$

where the expectation is taken with respect to the random parameters $\boldsymbol{\xi} \in \mathbb{R}^m$. Thus, based on a well formulated stochastic model, our investment banker can now choose a portfolio of stocks which maximizes long-term expected return, or similarly our manufacturing company can take an early decision which leads to the highest expected profits. Unfortunately, although the SP is a convex optimization problem, in order to solve it one must often resort to Monte Carlo approximations (see [55]), which can be computationally challenging. A more challenging difficulty that arises in practice is the need to commit to a distribution F given only limited information about the stochastic parameters.

In an effort to address these issues, a robust formulation for stochastic programming was proposed in [52]. In this model, after defining a set $\mathcal{D}$ of probability distributions that is assumed to include the true distribution F, the objective function is reformulated with respect to the worst case expected cost over the choice of a distribution in this set. Hence, this leads to solving the Distributionally Robust Stochastic Program:

$$\text{(DRSP)} \qquad \underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \left(\max_{F\in\mathcal{D}} \mathbb{E}_F[h(\boldsymbol{x},\boldsymbol{\xi})] \right) ,$$

where $\mathbb{E}_F[\cdot]$ is the expectation taken with respect to the random vector $\boldsymbol{\xi}$ given that it follows the probability distribution F. Although the DRSP is a convex problem, it is well known that some choices of $\mathcal{D}$ lead to instances of the problem that are NP-hard to solve. Hence, while the DRSP is one of the most general model for addressing uncertainty in decision making, the computational difficulties that are related to dealing with $\boldsymbol{\xi}$ of large dimension have limited the practical application of this model. Furthermore, it is the case that current DRSP approaches can lead to a false sense of security because they assume exact knowledge of the moments of the stochastic parameters. In practice, moment estimates are inherently noisy due to the fact that distribution information is often limited to historical data.

This thesis studies methods for formulating an accurate distributionally robust optimization model in the context of such data-driven problems and for solving this problem effectively. Focusing mainly on a frequentist approach which makes minimal assumptions about the stochastic process, we show how the historical observations can be used to construct an uncertainty set for the distribution which is guaranteed with high probability to contain F. The structure of this uncertainty set also ensures that the DRSP model can be solved efficiently for a large class of problems. This leads to providing solutions with tangible guarantees for a number of practical problems where uncertainty cannot simply be ignored. These include an aircraft fleet composition problem with random flight demand, a portfolio selection problem with random future stock values, and a machine replacement problem with random life span and replacement costs. We also briefly consider the case where one can commit to a Bayesian prior, and we compare the solution of a DRSP approach which accounts for this Bayesian prior to the solution of a percentile optimization model which is conceptually more appropriate yet can lead to an intractable problem. Overall, we believe that after reading through this thesis, a decision maker should have in hand the necessary tools to address large scale data-driven problems effectively and in a way that provides him with valuable guarantees about the future performance of his decision.

1.1 Prior Work in Distributionally Robust Optimization

Since its introduction by H. Scarf in 1958, the DRSP model, which is also known as the minimax stochastic program, has gained a lot of interest in the scientific community and can serve many purposes. In its simplest form, it can be used to compute upper bounds on the moment of a random vector (*i.e.*, the moment problem as reviewed in [34]). Alternatively, when estimating optimality of a candidate solution to a stochastic program, one can formulate the DRSP to compute an upper bound on the problem's optimal value (*e.g.*, in [9] and in [32]). Finally, in this work we are mainly interested by its capacity to provide robust decisions in contexts where distribution information is limited (e.g. in [16] and in [56]).

Interestingly, if one uses the distributional set that contains the set of single point distributions (*i.e.*, that put all of their weight at a single point) over the support $\mathcal{S}$, then the DRSP reduces to the so-called robust optimization problem [2] which takes the form

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \left(\max_{\boldsymbol{\xi}\in\mathcal{S}} \; h(\boldsymbol{x},\boldsymbol{\xi})\right)$$

Despite the fact that this model entirely disregards the stochastic structure of $\boldsymbol{\xi}$, it has become very popular because it is often the case that accounting for uncertainty this way comes at a small overhead cost in computation time (see [4] for a review).

More generally, authors have considered a wide range of forms for the distributional set $\mathcal{D}$. In [33] and in [54], the authors consider a set containing unimodal distributions that satisfy some given support constraints. Under some conditions on $h(\boldsymbol{x},\boldsymbol{\xi})$, they characterize the worst distribution as being a uniform distribution. The most popular type of distributional set $\mathcal{D}$ imposes linear constraints on moments of the distribution as is discussed in [52, 16, 46, 6]. While many different forms of the distributional set can be found in the literature (see [17] and reference therein), computational difficulties that arise when dealing with $\boldsymbol{\xi}$ of large dimension or with a rich objective function have limited the practical application of this model. More specifically, the worst case moment expression can rarely be simplified analytically, as in the linear chance constraint problem considered in [12]. Instead, it is more common that the model is intractable and that only global optimization methods can be employed to get an optimal solution (*e.g.*, in [20] and in [23]).

In order to make the DRSP model tractable, approaches that consider moment constraints typically assumed that the first and second moments are known exactly. For instance, in his original model, Scarf considered a newsvendor problem with a one dimensional decision variable x and a single parameter ξ, which represented respectively how much inventory the vendor should hold and a random amount, with known mean and variance, of demand for the newspapers. The cost function had the form $h(x, \xi) = \max\{c(x-\xi)\ ,\ r(\xi - x)\}$. To solve this model, Scarf exploited the fact that the worst case distribution of demand could be chosen to be one with all of its weight on two points. This idea was reused in [65, 67, 45] where, although the objective functions take more interesting forms, the authors make the same assumption. This line of work also relies on characterizing the worst case distribution as a point distribution. Unfortunately, the assumption of exact knowledge of these moments already fails to generate tractable models when a constraint on support is also considered (see [6] which is revisited in Section 4.1).

Although it is rarely the case in practice that a random vector's moments are exactly known, only a few authors have considered the challenge of accounting precisely for the only distribution information that is present in most decision problems: *i.e.*, historical samples drawn from the distribution. In [43], the authors considered a discrete random variable with a polynomial number of states and proposed using a likelihood based approach for constructing the distributional set based on past observations. They show that the resulting distributionally robust problem is tractable. Unfortunately, these results do not generalize easily to problems that involve continuous variables. In fact, prior to this work, the idea that an optimization model should account for a confidence region of moments had only been discussed briefly in two special forms of portfolio selection problems (see [19] and [27]). In contrast, this thesis provides a unifying framework for formulating and solving a well-justified distributionally robust optimization model in the context of data-driven problems. The details of our contribution are presented next.

1.2 Contribution of this Thesis

In Chapter 2, we address the question of distribution uncertainty from a frequentist point of view. Using the historical data that is available, we derive new confidence regions for the mean and covariance matrix of a random vector. While the typical confidence region for a covariance matrix is expressed using sum of square differences, the one that we propose constrains the terms directly with respect to their effect on the variance of projections of the random vector. This new confidence region appeared for the first time in our article [15] and should be of value to the field of applied statistics.

In Chapter 3, we analyze the computational difficulties related to solving the distributionally robust stochastic program with a distributional set that constrains the support and accounts for moment uncertainty. Although the model has been studied heavily since Scarf introduced it in 1958, we are the first to show that for decision models where the objective is convex in the decision variables and ”piecewise concave” in the parameters, there exists a polynomial time algorithm that can solve the distributionally robust form to any level of accuracy. In fact, we also show that in some cases the distributionally robust stochastic program is much easier to solve than the original stochastic programming form. In conjunction with our new results on confidence regions, this analysis leads to proposing a complete framework for providing reliable risk-sensitive solutions to data-driven problems.

In Chapter 4, we present three applications of the distributionally robust framework: the problem of bounding a probability inequality, the problem of composing a fleet of aircraft when demand for the scheduled flights is unknown, and a problem of portfolio selection. In each of these three problems, given that the distribution information comes from historical samples, the new theory allows the formulation of an optimization model that is tractable and is guaranteed with high probability to account for the true unknown distribution. Furthermore, our experiments use real stock data to demonstrate that in the portfolio selection problem there are heavy risks related to disregarding one's uncertainty in either the form of the distribution or in the quality of some measurements of its moments.

In Chapter 5, we adopt the Bayesian perspective to address distribution uncertainty in stochastic decision problems. Under this framework, it is well known that distribution uncertainty can be described by a hyper-distribution (*i.e.*, a distribution over distributions). After adapting the distributionally robust approach to this context, we compare it to a percentile optimization model in terms of accuracy of representation and computational tractability. Our study focuses more specifically on Markov decision processes and leads to interesting new theoretical results for the problem of percentile optimization. More specifically, while the percentile criterion can be solved efficiently in some special cases, we show that it can easily lead to NP-hard problems. Fortunately, in the case of data-driven Markov decision processes, we propose the first approximation of this criterion whose accuracy can be measured in terms of the number of observations.

Chapter 2

Moment Uncertainty in Data-driven Problems

We first adopt the frequentist point of view which defines the probability distribution of our random parameters as the limit of the relative frequency of occurrence in a large number of trials. Our first objective is to answer the question: What is the distribution information present in historical data for a vector of uncertain parameters? More specifically, based on a set of M samples $\{\boldsymbol{\xi}_i\}_{i=1}^M$ identically and independently drawn from the distribution associated with $\boldsymbol{\xi}$, we are interested in determining useful statistical characteristics of the distribution. In this chapter, we focus on the task of defining confidence regions for the mean and covariance matrix of this random vector solely based on basic knowledge of the support of this random vector and a finite set of samples. The main contribution of this work is to derive a new form of confidence region for a covariance matrix which takes the shape of the difference between two positive semi-definite cones. This new confidence region appeared for the first time in our article [15].

2.1 Introduction

In what follows, we will show how one can define a confidence region for the mean and the covariance matrix of a random vector $\boldsymbol{\xi}$ such that it is assured with high probability to contain the mean and covariance matrix of the distribution of $\boldsymbol{\xi}$. This will be achieved without making any structural assumptions about the distribution of $\boldsymbol{\xi}$ beside that its support is bounded by a ball of known finite radius. In order to simplify the derivations, we start by reformulating the random vector $\boldsymbol{\xi}$ in terms of a linear combination of uncorrelated random variables. More specifically, given the random vector $\boldsymbol{\xi} \in \mathbb{R}^m$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma} \succ 0$, let us define $\boldsymbol{\zeta} \in \mathbb{R}^m$ to be the normalized random vector $\boldsymbol{\zeta} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\xi} - \boldsymbol{\mu})$ such that $\mathbb{E}[\boldsymbol{\zeta}] = 0$ and $\mathbb{E}[\boldsymbol{\zeta}\boldsymbol{\zeta}^{\mathsf{T}}] = \boldsymbol{I}$. Also, let us make the following assumption about $\boldsymbol{\zeta}$:

Assumption 2.1.1 : There exists a ball of radius R that contains the entire support of the unknown distribution of $\boldsymbol{\zeta}$. More specifically, there exist an $R \geq 0$ such that

$$\mathbb{P}\left((\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) \leq R^2\right) = 1 \ .$$

In practice, even when one does not have information about $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, we believe that one can often still make an educated and conservative guess about the magnitude of R. We will also revisit this issue in Section 2.4 where R will be derived based on the bounded support of $\boldsymbol{\xi}$ and a set of samples $\{\boldsymbol{\xi}_i\}_{i=1}^M$. In what follows, a confidence region for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ will be resolved based on Assumption 2.1.1 and on an inequality known as the "independent bounded differences inequality", which was popularized by McDiarmid.[1] In fact, this inequality can be seen as a generalized version of Hoeffding's inequality.

Theorem 2.1.2 : *(McDiarmid [39]) Let $\{\boldsymbol{\xi}_i\}_{i=1}^M$ be a set of independent random vectors $\boldsymbol{\xi}_i$ taking values in a set $\mathcal{S}_i$ for each i. Suppose that the real-valued function $g(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_M)$ defined on $\mathcal{S}_1 \times \mathcal{S}_2 \times ... \times \mathcal{S}_M$ satisfies*

$$|g(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_M) - g(\boldsymbol{\xi}'_1, \boldsymbol{\xi}'_2, ..., \boldsymbol{\xi}'_M)| \leq c_j \tag{2.1}$$

whenever the vector sets $\{\boldsymbol{\xi}_i\}_{i=1}^M$ and $\{\boldsymbol{\xi}'_i\}_{i=1}^M$ differ only in the j-th vector. Then for any $t \geq 0$,

$$\mathbb{P}\left(g(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_M) - \mathbb{E}\left[g(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_M)\right] \leq -t\right) \leq \exp\left(\frac{-2t^2}{\sum_{j=1}^M c_j^2}\right) \ .$$

[1] If the support of $\boldsymbol{\xi}$ is unbounded, then one can derive similar results by considering that the support is bounded with high probability, or by using some knowledge of the higher moments of $\boldsymbol{\xi}$, as confirmed in [59].

2.2 A Confidence Region for the Mean

A first use of McDiarmid's theorem leads to defining an ellipsoidal constraint relating the empirical estimate $\hat{\boldsymbol{\mu}} = M^{-1}\sum_{i=1}^{M}\boldsymbol{\xi}_i$ to the true mean and true covariance matrix of the random vector $\boldsymbol{\xi}$.

The following result is an interesting consequence of McDiarmid's theorem.

Lemma 2.2.1: *(Shawe-Taylor & Cristianini [57]) Let $\{\boldsymbol{\zeta}_i\}_{i=1}^{M}$ be a set of M samples generated independently at random according to the distribution of $\boldsymbol{\zeta}$. If $\boldsymbol{\zeta}$ satisfies Assumption 2.1.1 then with probability at least $(1-\delta)$ over the choice of samples $\{\boldsymbol{\zeta}_i\}_{i=1}^{M}$, we have that*

$$\left\|\frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\zeta}_i\right\|^2 \leq \frac{R^2}{M}\left(2+\sqrt{2\ln(1/\delta)}\right)^2 .$$

This result can in turn be used to derive a similar statement about the random vector $\boldsymbol{\xi}$.

Corollary 2.2.2: *Let $\{\boldsymbol{\xi}_i\}_{i=1}^{M}$ be a set of M samples generated independently at random according to the distribution of $\boldsymbol{\xi}$. If $\boldsymbol{\xi}$ satisfies Assumption 2.1.1, then with probability greater than $1-\delta$, we have that*

$$(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu}) \leq \beta(\delta) , \tag{2.2}$$

where $\hat{\boldsymbol{\mu}} = \frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\xi}_i$ and $\beta(\delta) = (R^2/M)\left(2+\sqrt{2\ln(1/\delta)}\right)^2$.

Proof: This generalization for a $\boldsymbol{\xi}$ with arbitrary mean and covariance matrix is quite straightforward:

$$\begin{aligned}\mathbb{P}\left((\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu}) \leq \beta(\delta)\right) &= \mathbb{P}\left(\left\|\boldsymbol{\Sigma}^{-1/2}\left(\frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\xi}_i-\boldsymbol{\mu}\right)\right\|^2 \leq \beta(\delta)\right) \\ &= \mathbb{P}\left(\left\|\sum_{i=1}^{M}\boldsymbol{\zeta}_i\right\|^2 \leq \beta(\delta)\right) \geq 1-\delta . \quad \square\end{aligned}$$

Since $\boldsymbol{\Sigma}$ is non-singular, the inequality of Equation (2.2) constrains the vector $\boldsymbol{\mu}$ and matrix $\boldsymbol{\Sigma}$ to a convex set. This set can be represented by the following linear matrix inequality after applying the principles of Schur's complement:

$$\begin{bmatrix} \boldsymbol{\Sigma} & (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) \\ (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^{\mathsf{T}} & \beta(\delta) \end{bmatrix} \succeq 0 \ .$$

2.3 New Confidence Region for the Covariance Matrix

In order for Constraint (2.2) to describe a bounded set, one must be able to bound the uncertainty in $\boldsymbol{\Sigma}$. While confidence regions for the covariance matrix are typically defined in terms of bounding the sum of square differences between each term of the matrix and its estimate (see for example [57]), we favor the structure imposed by two linear matrix inequalities bounding $\boldsymbol{\Sigma}$ around its empirical estimate $\hat{\boldsymbol{\Sigma}} = M^{-1}\sum_{i=1}^{M}(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})^{\mathsf{T}}$:

$$\mathbb{P}\left(c_{\text{min}}\hat{\boldsymbol{\Sigma}} \preceq \boldsymbol{\Sigma} \preceq c_{\text{max}}\hat{\boldsymbol{\Sigma}}\right) \geq 1 - \delta \ . \tag{2.3}$$

Note that the difficulty of this task is mostly due to the fact that one needs to derive a confidence interval for the eigenvalues of the stochastic matrix $\boldsymbol{\Sigma}^{-1/2}\hat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1/2}$. For the case that interests us, where $M \gg m$ with M finite and m fixed, prior work usually assumes $\boldsymbol{\xi}$ is a normally distributed random vector (see [1] and [18]). Under the Gaussian assumption, the empirical covariance matrix follows the Wishart distribution, thus one can formulate the distribution of eigenvalues in a closed form expression and derive such percentile bounds. In the case where $\boldsymbol{\xi}$ takes a non-normal form, the asymptotic distribution of eigenvalues was studied in [62] and [22] among others. However, to the best of our knowledge, our work is the first to formulate an uncertainty set with the characteristics presented in Equation (2.3) for a sample set of finite size. In what follows, we start by demonstrating how a confidence region of the form presented in Equation (2.3) can be defined around $\hat{\boldsymbol{I}} = M^{-1}\sum_i \boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^{\mathsf{T}}$ for the covariance matrix of $\boldsymbol{\zeta}$. Next, we will assume that the mean of $\boldsymbol{\xi}$ is exactly known and we will formulate the confidence region for $\boldsymbol{\Sigma}$ in terms of $\hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) = M^{-1}\sum_{i=1}^{M}(\boldsymbol{\xi}_i - \boldsymbol{\mu})(\boldsymbol{\xi}_i - \boldsymbol{\mu})^{\mathsf{T}}$. We conclude this section with our main result about a confidence region for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ which relies solely on the M samples and on support information about the random vector $\boldsymbol{\xi}$.

Lemma 2.3.1 : *Let $\{\boldsymbol{\zeta}_i\}_{i=1}^M$ be a set of M samples generated independently at random according to the distribution of $\boldsymbol{\zeta}$. If $\boldsymbol{\zeta}$ satisfies Assumption 2.1.1, then with probability greater than $1-\delta$, we have that*

$$\frac{1}{1+\alpha(\delta/2)}\hat{\boldsymbol{I}} \preceq \boldsymbol{I} \preceq \frac{1}{1-\alpha(\delta/2)}\hat{\boldsymbol{I}} \quad , \tag{2.4}$$

where $\alpha(\delta/2) = (R^2/\sqrt{M})\left(\sqrt{1-m/R^4}+\sqrt{\ln(2/\delta)}\right)$, provided that

$$M > R^4\left(\sqrt{1-m/R^4}+\sqrt{\ln(2/\delta)}\right)^2 \quad . \tag{2.5}$$

Proof: The proof of this theorem relies on applying Theorem 2.1.2 twice to show that both $\frac{1}{1+\alpha(\delta/2)}\hat{\boldsymbol{I}} \preceq \boldsymbol{I}$ and $\boldsymbol{I} \preceq \frac{1}{1-\alpha(\delta/2)}\hat{\boldsymbol{I}}$ occur with probability greater than $1-\delta/2$. Our statement then simply follows by the union bound. However, for the sake of conciseness, this proof will focus on deriving the upper bound since the steps that we follow can easily be modified for the derivation of the lower bound.

When applying Theorem 2.1.2 to show that $\boldsymbol{I} \preceq \frac{1}{1-\alpha(\delta/2)}\hat{\boldsymbol{I}}$ occurs with probability greater than $1-\delta/2$, the main step consists of defining $g(\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_2,...,\boldsymbol{\zeta}_M) = \min_{\|\boldsymbol{z}\|=1} \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z}$ and finding a lower bound for $\mathbb{E}\left[g(\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_2,...,\boldsymbol{\zeta}_M)\right]$. One can start by showing that Constraint (2.1) is met when $c_j = R^2/M$ for all j.

$$|g(\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_2,...,\boldsymbol{\zeta}_M) - g(\boldsymbol{\zeta}_1',\boldsymbol{\zeta}_2',...,\boldsymbol{\zeta}_M')| = \left|\min_{\|\boldsymbol{z}\|=1} \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z} - \min_{\|\boldsymbol{z}\|=1} \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}'\boldsymbol{z}\right| \quad ,$$

where $\hat{\boldsymbol{I}}' = \frac{1}{M}\sum_{i=1}^M \boldsymbol{\zeta}_i'\boldsymbol{\zeta}_i'^\mathsf{T} = \hat{\boldsymbol{I}} + \frac{1}{M}(\boldsymbol{\zeta}_j'\boldsymbol{\zeta}_j'^\mathsf{T} - \boldsymbol{\zeta}_j\boldsymbol{\zeta}_j^\mathsf{T})$ since $\{\boldsymbol{\zeta}_i\}_{i=1}^M$ and $\{\boldsymbol{\zeta}_i'\}_{i=1}^M$ only differ in the j-th vector.

Now assume that $\min_{\|\boldsymbol{z}\|=1} \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z} \geq \min_{\|\boldsymbol{z}\|=1} \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}'\boldsymbol{z}$. Then, $\forall \boldsymbol{z}^* \in \arg\min_{\|\boldsymbol{z}\|=1} \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}'\boldsymbol{z}$, one can show that

$$\begin{aligned}|g(\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_2,...,\boldsymbol{\zeta}_M) - g(\boldsymbol{\zeta}_1',\boldsymbol{\zeta}_2',...,\boldsymbol{\zeta}_M'))| &= \min_{\|\boldsymbol{z}\|=1} \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z} - \boldsymbol{z}^{*\mathsf{T}}\hat{\boldsymbol{I}}'\boldsymbol{z}^* \leq \boldsymbol{z}^{*\mathsf{T}}(\hat{\boldsymbol{I}}-\hat{\boldsymbol{I}}')\boldsymbol{z}^* \\ &= \boldsymbol{z}^{*\mathsf{T}}\frac{1}{M}(\boldsymbol{\zeta}_j\boldsymbol{\zeta}_j^\mathsf{T} - \boldsymbol{\zeta}_j'\boldsymbol{\zeta}_j'^\mathsf{T})\boldsymbol{z}^* = \frac{1}{M}\left((\boldsymbol{\zeta}_j^\mathsf{T}\boldsymbol{z}^*)^2 - (\boldsymbol{\zeta}_j'^\mathsf{T}\boldsymbol{z}^*)^2\right) \leq \frac{\|\boldsymbol{z}^*\|^2\|\boldsymbol{\zeta}_j\|^2}{M} \leq \frac{R^2}{M} \quad .\end{aligned}$$

Otherwise, for the case that $\min_{\|\boldsymbol{z}\|=1} \ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z} \leq \min_{\|\boldsymbol{z}\|=1} \ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}'\boldsymbol{z}$, the same argument applies using $\boldsymbol{z}^* \in \arg\min_{\|\boldsymbol{z}\|=1} \ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z}$.

The task of bounding $\mathbb{E}\left[g(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, ..., \boldsymbol{\zeta}_M)\right]$ is a bit harder. We can instead start by finding an upper bound on the expected maximum eigenvalue of $(\boldsymbol{I} - \hat{\boldsymbol{I}})$ since

$$\mathbb{E}\left[\max_{\|\boldsymbol{z}\|=1} \ \boldsymbol{z}^\mathsf{T}(\boldsymbol{I} - \hat{\boldsymbol{I}})\boldsymbol{z}\right] = 1 - \mathbb{E}\left[\min_{\|\boldsymbol{z}\|=1} \ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z}\right] \ , \tag{2.6}$$

where the expectation is taken with respect to the random matrix $\hat{\boldsymbol{I}}$. Using Jensen's inequality and basic linear algebra, one can show that

$$\begin{aligned}
\left(\mathbb{E}\left[\max_{\|\boldsymbol{z}\|=1} \ \boldsymbol{z}^\mathsf{T}(\boldsymbol{I} - \hat{\boldsymbol{I}})\boldsymbol{z}\right]\right)^2 &\leq \mathbb{E}\left[\left(\max_{\|\boldsymbol{z}\|=1} \ \boldsymbol{z}^\mathsf{T}(\boldsymbol{I} - \hat{\boldsymbol{I}})\boldsymbol{z}\right)^2\right] \\
&\leq \mathbb{E}\left[\sum_{i=1}^{m} \sigma_i^2(\boldsymbol{I} - \hat{\boldsymbol{I}})\right] = \mathbb{E}\left[\mathbf{trace}\left(\left(\boldsymbol{I} - \hat{\boldsymbol{I}}\right)^2\right)\right] \\
&= \mathbb{E}\left[\mathbf{trace}\left(\left(\frac{1}{M}\sum_{i=1}^{M}\left(\boldsymbol{I} - \boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^\mathsf{T}\right)\right)^2\right)\right] \\
&= \mathbf{trace}\left(\frac{1}{M^2}\sum_{i=1}^{M}\mathbb{E}\left[\boldsymbol{I} - 2\boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^\mathsf{T} + \left(\boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^\mathsf{T}\right)^2\right]\right) \\
&= \frac{1}{M}\left(\mathbf{trace}\left(\mathbb{E}\left[\left(\boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^\mathsf{T}\right)^2\right]\right) - \mathbf{trace}\left(\boldsymbol{I}\right)\right) \\
&= \frac{\mathbb{E}\left[\|\boldsymbol{\zeta}_i\|^4\right] - m}{M} \leq \frac{R^4 - m}{M} \ ,
\end{aligned}$$

where $\sigma_i(\boldsymbol{I} - \hat{\boldsymbol{I}})$ refers to the i-th singular value of $\boldsymbol{I} - \hat{\boldsymbol{I}}$. The derivation above uses the fact that each $\boldsymbol{\zeta}_i$ is drawn independently thus making $\mathbb{E}\left[(\boldsymbol{I} - \boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^\mathsf{T})(\boldsymbol{I} - \boldsymbol{\zeta}_j\boldsymbol{\zeta}_j^\mathsf{T})\right] = \mathbb{E}\left[\boldsymbol{I} - \boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^\mathsf{T}\right]\mathbb{E}\left[\boldsymbol{I} - \boldsymbol{\zeta}_j\boldsymbol{\zeta}_j^\mathsf{T}\right] = 0$.

By replacing this lower bound in Equation (2.6), we have that $\mathbb{E}\left[g(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, ..., \boldsymbol{\zeta}_M)\right] \geq 1 - (R^2/\sqrt{M})\sqrt{1 - m/R^4}$. Thus, Theorem 2.1.2 allows us to confirm the proposed upper bound using the following argument. Since the statement

$$\mathbb{P}\left(\min_{\|\boldsymbol{z}\|=1} \ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z} - \mathbb{E}\left[\min_{\|\boldsymbol{z}\|=1} \ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z}\right] \leq -\epsilon\right) \leq \exp\left(\frac{-2\epsilon^2}{\sum_{j=1}^{M}(R^4/M^2)}\right) \ ,$$

implies that

$$\mathbb{P}\left(\min_{\|\boldsymbol{z}\|=1}\ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z} - \mathbb{E}\left[\min_{\|\boldsymbol{z}\|=1}\ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z}\right] \geq -\frac{R^2\sqrt{\ln(2/\delta)}}{\sqrt{M}}\right) \geq 1-\delta/2\ \ ,$$

and since relaxing $\mathbb{E}\left[\min_{\|\boldsymbol{z}\|=1}\ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z}\right]$ to its lower bound can only include more random events, we necessarily have that

$$\mathbb{P}\left(\min_{\|\boldsymbol{z}\|=1}\ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z} \geq 1 - \frac{R^2}{\sqrt{M}}\left(\sqrt{1-m/R^4}+\sqrt{\ln(2/\delta)}\right)\right) \geq 1-\delta/2\ \ .$$

Thus, given that M is large enough such that $1-\alpha(\delta/2) > 0$, we conclude that

$$\mathbb{P}\left(\boldsymbol{I} \preceq \frac{1}{1-\alpha(\delta/2)}\hat{\boldsymbol{I}}\right) \geq 1-\delta/2\ \ .$$

The task of showing that $1/(1+\alpha(\delta/2))\hat{\boldsymbol{I}} \preceq \boldsymbol{I}$ also occurs with probability $1-\delta/2$ is very similar. One needs to apply Theorem 2.1.2, now defining $g(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, ..., \boldsymbol{\zeta}_M) = -\min_{\|\boldsymbol{z}\|=1}\ \boldsymbol{z}^\mathsf{T}\hat{\boldsymbol{I}}\boldsymbol{z}$, and to demonstrate that $\mathbb{E}\left[g(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, ..., \boldsymbol{\zeta}_M)\right] \geq -1-\alpha(\delta/2)$. The rest follows easily. □

Remark 2.3.2 : Considering the case where ζ is actually a random variable and where $\hat{\sigma}^2 = \sum_{i=1}^M \zeta_i^2$, one can easily verify that the parameter $\alpha(\delta)$ in Lemma 2.3.1 is asymptotically of the right order in terms of M and R. Since $\mathbb{E}\left[\zeta^4\right]$ is bounded by R^4, the central limit theorem guarantees that $\sqrt{M}(\hat{\sigma}^2-\mathbb{E}\left[\zeta^2\right])$ converges in distribution to $\mathcal{N}(0, \mathbb{E}\left[\zeta^4\right]-1)$. Thus, it follows that the distribution of $\left(M/(\mathbb{E}\left[\zeta^4\right]-1)\right)\left(\hat{\sigma}^2-\mathbb{E}\left[\zeta^2\right]\right)^2$ converges to a χ^2-distribution with degree 1. For any $\delta > 0$, one can find $c(\delta)$ such that $|\hat{\sigma}^2 - \mathbb{E}\left[\zeta^2\right]| \leq \frac{c(\delta)\sqrt{\mathbb{E}\left[\zeta^4\right]-1}}{\sqrt{M}}$ is satisfied with probability greater than $1-\delta$. Hence, asymptotically the confidence region $-\left(1+\frac{c(\delta)R^2}{\sqrt{M}}\right)^{-1}\hat{\sigma}^2 \leq 1 \leq \left(1-\frac{c(\delta)R^2}{\sqrt{M}}\right)^{-1}\hat{\sigma}^2$ is tight.

We are now interested in extending Lemma 2.3.1 to a random vector with general mean and covariance matrix. Given the random event that Constraint (2.4) is satisfied, then:

$$\boldsymbol{I} \preceq \frac{1}{1-\alpha(\delta/2)}\hat{\boldsymbol{I}} \;\Rightarrow\; \boldsymbol{\Sigma}^{1/2}\boldsymbol{I}\boldsymbol{\Sigma}^{1/2} \preceq \frac{1}{1-\alpha(\delta/2)}\boldsymbol{\Sigma}^{1/2}\hat{\boldsymbol{I}}\boldsymbol{\Sigma}^{1/2}$$

$$\Rightarrow\; \boldsymbol{\Sigma} \preceq \frac{1}{1-\alpha(\delta/2)}\frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^{\mathsf{T}}\boldsymbol{\Sigma}^{1/2} \;\Rightarrow\; \boldsymbol{\Sigma} \preceq \frac{1}{1-\alpha(\delta/2)}\hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) ,$$

and similarly,

$$\frac{1}{1+\alpha(\delta/2)}\hat{\boldsymbol{I}} \preceq \boldsymbol{I} \quad\Rightarrow\quad \frac{1}{1+\alpha(\delta/2)}\hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) \preceq \boldsymbol{\Sigma} \ .$$

Since Constraint (2.4) is satisfied with probability greater than $1-\delta$, the following corollary follows easily.

Corollary 2.3.3 : *Let $\{\boldsymbol{\xi}_i\}_{i=1}^{M}$ be a set of M samples generated independently at random according to the distribution of $\boldsymbol{\xi}$. If $\boldsymbol{\xi}$ satisfies Assumption 2.1.1 and M satisfies Constraint* (2.5)*, then with probability greater then $1-\delta$, we have that*

$$\frac{1}{1+\alpha(\delta/2)}\hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) \preceq \boldsymbol{\Sigma} \preceq \frac{1}{1-\alpha(\delta/2)}\hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) \ ,$$

where $\hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) = \frac{1}{M}\sum_{i=1}^{M}(\boldsymbol{\xi}_i-\boldsymbol{\mu})(\boldsymbol{\xi}_i-\boldsymbol{\mu})^{\mathsf{T}}$ and $\alpha(\delta/2)$ is defined as in Lemma 2.3.1.

Combined with Corollary 2.2.2, this statement leads to the description of a convex set which is constructed using empirical estimates of the mean and covariance matrix, and yet is guaranteed to contain the true mean and covariance matrix of $\boldsymbol{\xi}$ with high probability.

Theorem 2.3.4 : *Let $\{\boldsymbol{\xi}_i\}_{i=1}^{M}$ be a set of M samples generated independently at random according to the distribution of $\boldsymbol{\xi}$. If $\boldsymbol{\xi}$ satisfies Assumption 2.1.1 and M satisfies Equation* (2.5)*, then with probability greater than $1-\delta$ over the choice of $\{\boldsymbol{\xi}_i\}_{i=1}^{M}$, the following* *set of constraints are met:*

$$(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) \leq \beta(\delta/2) \tag{2.7a}$$

$$\boldsymbol{\Sigma} \preceq \frac{1}{1 - \alpha(\delta/4) - \beta(\delta/2)}\hat{\boldsymbol{\Sigma}} \tag{2.7b}$$

$$\boldsymbol{\Sigma} \succeq \frac{1}{1 + \alpha(\delta/4)}\hat{\boldsymbol{\Sigma}} \ , \tag{2.7c}$$

where $\hat{\boldsymbol{\Sigma}} = \frac{1}{M}\sum_{i=1}^{M}(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})^{\mathsf{T}}$, $\alpha(\delta/4) = (R^2/\sqrt{M})\left(\sqrt{1 - m/R^4} + \sqrt{\ln(4/\delta)}\right)$, $\beta(\delta/2) = (R^2/M)\left(2 + \sqrt{2\ln(2/\delta)}\right)^2$.

Proof: By applying Corollary 2.2.2, 2.3.3, and Lemma 2.3.1, the union bound guarantees us with probability greater than $1 - \delta$ that the following constraints are met:

$$(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) \leq \beta(\delta/2)$$

$$\frac{1}{1 + \alpha(\delta/4)}\hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) \preceq \boldsymbol{\Sigma} \preceq \frac{1}{1 - \alpha(\delta/4)}\hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) \ .$$

Note that our result is not proven yet since, although the first constraint is exactly Constraint (2.7a), the second and third constraints actually refer to a covariance matrix estimate that uses the true mean of the distribution instead of its empirical estimate. The following steps will convince us that these conditions are sufficient for Constraint (2.7b) and Constraint (2.7c) to hold:

$$\begin{aligned}
(1 - \alpha(\delta/4))\boldsymbol{\Sigma} \ &\preceq \ \hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) \ = \ \frac{1}{M}\sum_{i=1}^{M}(\boldsymbol{\xi}_i - \boldsymbol{\mu})(\boldsymbol{\xi}_i - \boldsymbol{\mu})^{\mathsf{T}} \\
&= \ \frac{1}{M}\sum_{i=1}^{M}(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\mu}} - \boldsymbol{\mu})(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^{\mathsf{T}} \\
&= \ \frac{1}{M}\sum_{i=1}^{M}(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})^{\mathsf{T}} + (\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^{\mathsf{T}} \\
&\qquad\qquad + (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})^{\mathsf{T}} + (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^{\mathsf{T}} \\
&= \ \hat{\boldsymbol{\Sigma}} + (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^{\mathsf{T}} \ \preceq \ \hat{\boldsymbol{\Sigma}} + \beta(\delta/2)\boldsymbol{\Sigma} \ ,
\end{aligned}$$

where the last semi-definite inequality of the derivation can be explained using the fact that for any $\boldsymbol{x} \in \mathbb{R}^m$,

$$\begin{aligned}\boldsymbol{x}^\mathsf{T}(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})^\mathsf{T}\boldsymbol{x} &= (\boldsymbol{x}^\mathsf{T}(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu}))^2 = \left(\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}^{1/2}\,\boldsymbol{\Sigma}^{-1/2}(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})\right)^2 \\ &\leq \|\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}^{1/2}\|^2\|\boldsymbol{\Sigma}^{-1/2}(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})\|^2 \leq \beta(\delta/2)\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x} \ .\end{aligned}$$

Thus we can conclude that Constraint (2.7b) is met. Similar steps can be used to show that Constraint (2.7c) also holds:

$$(1+\alpha(\delta/4))\boldsymbol{\Sigma} \succeq \hat{\boldsymbol{\Sigma}}(\boldsymbol{\mu}) = \frac{1}{M}\sum_{i=1}^{M}(\boldsymbol{\xi}_i-\boldsymbol{\mu})(\boldsymbol{\xi}_i-\boldsymbol{\mu})^\mathsf{T} = \hat{\boldsymbol{\Sigma}} + (\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})^\mathsf{T} \succeq \hat{\boldsymbol{\Sigma}} \,.$$

□

2.4 Bounding the Support of ζ using Empirical Data

The above derivations assumed that one could describe a ball containing the support of the "fictitious" random vector $\boldsymbol{\zeta}$. In fact, this assumption can be replaced by an assumption on the support of the more tangible random vector $\boldsymbol{\xi}$ as is presented in the following corollary.

Corollary 2.4.1 : *Let $\{\boldsymbol{\xi}_i\}_{i=1}^M$ be a set of M samples generated independently at random according to the distribution of $\boldsymbol{\xi}$. Given that the support of the distribution of $\boldsymbol{\xi}$ is known to be contained in $\mathcal{S}_{\boldsymbol{\xi}}$, let*

$$\hat{R} = \sup_{\boldsymbol{\xi}\in\mathcal{S}_{\boldsymbol{\xi}}} \|\hat{\boldsymbol{\Sigma}}^{-1/2}(\boldsymbol{\xi}-\hat{\boldsymbol{\mu}})\|_2$$

be a stochastic approximation of R and for any $\delta > 0$, let

$$\bar{R} = \left(1-(\hat{R}^2+2)\frac{2+\sqrt{2\ln(4/\bar{\delta})}}{\sqrt{M}}\right)^{-1/2}\hat{R} \ ,$$

where $\bar{\delta} = 1 - \sqrt{1-\delta}$. *If*

$$M \;>\; \max\left\{ (\hat{R}^2+2)^2\left(2+\sqrt{2\ln(4/\bar{\delta})}\right)^2 \;,\; \frac{\left(8+\sqrt{32\ln(4/\bar{\delta})}\right)^2}{\left(\sqrt{\hat{R}+4}-\hat{R}\right)^4} \right\} \;, \tag{2.8}$$

then with probability greater than $1-\delta$, *constraints* (2.7a), (2.7b), *and* (2.7c) *are satisfied with* $\alpha(\delta/4)$ *and* $\beta(\delta/2)$ *replaced with* $\bar{\alpha}(\bar{\delta}/4) = (\bar{R}^2/\sqrt{M})\left(\sqrt{1-m/\bar{R}^4}+\sqrt{\ln(4/\bar{\delta})}\right)$ *and* $\bar{\beta}(\bar{\delta}/2) = (\bar{R}^2/M)\left(2+\sqrt{2\ln(2/\bar{\delta})}\right)^2$ *respectively.*

Proof: Since we assumed that $\boldsymbol{\Sigma}$ was non-singular, the fact that the support of $\boldsymbol{\xi}$ is bounded by a ball of radius $R_{\boldsymbol{\xi}}$ implies that $\boldsymbol{\zeta}$ is also bounded. Thus, there exists an R such that $\mathbb{P}(\|\boldsymbol{\zeta}\| \leq R) = 1$. Given that $\boldsymbol{\zeta}$ has a bounded support and given Condition (2.8), Theorem 3.4.2 guarantees us that with probability greater than $1-\bar{\delta}$, constraints (2.7a), (2.7b), and (2.7c) are met. Thus

$$\begin{aligned} R &= \sup_{\boldsymbol{\zeta}\in\mathcal{S}_{\boldsymbol{\zeta}}} \|\boldsymbol{\zeta}\| = \sup_{\boldsymbol{\xi}\in\mathcal{S}_{\boldsymbol{\xi}}} \|\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\xi}-\boldsymbol{\mu})\| = \sup_{\boldsymbol{\xi}\in\mathcal{S}_{\boldsymbol{\xi}}} \|\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\xi}-\boldsymbol{\mu}+\hat{\boldsymbol{\mu}}-\hat{\boldsymbol{\mu}})\| \\ &\leq \sup_{\boldsymbol{\xi}\in\mathcal{S}_{\boldsymbol{\xi}}} \|\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\xi}-\hat{\boldsymbol{\mu}})\| + \|\boldsymbol{\Sigma}^{-1/2}(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})\| \\ &\leq \sup_{\boldsymbol{\xi}\in\mathcal{S}_{\boldsymbol{\xi}}} \sqrt{1+\alpha(\bar{\delta}/4)}\|\hat{\boldsymbol{\Sigma}}^{-1/2}(\boldsymbol{\xi}-\hat{\boldsymbol{\mu}})\| + \sqrt{\beta(\bar{\delta}/2)} \\ &\leq \sqrt{1+\alpha(\bar{\delta}/4)}\hat{R} + \sqrt{\beta(\bar{\delta}/2)} \;\leq\; \hat{R}\sqrt{1+cR^2}+cR \;, \end{aligned}$$

where $c = \left(2+\sqrt{2\ln(4/\bar{\delta})}\right)/\sqrt{M}$.

A careful analysis of the function $\psi(R,\hat{R}) = \hat{R}\sqrt{1+cR^2}+cR$ leads to the observation that if M satisfies Condition (2.8) then the fact that $R \leq \psi(R,\hat{R})$ necessarily implies that $R \leq \bar{R}$. We can therefore conclude that $\mathbb{P}(R \leq \bar{R}) \geq 1-\bar{\delta}$.

Given the event that $R \leq \bar{R}$ occurs, since

$$\begin{aligned} \alpha(\bar{\delta}/4) &= (R^2/\sqrt{M})\left(\sqrt{1-m/R^4}+\sqrt{2\ln(4/\bar{\delta})}\right) \\ &\leq (\bar{R}^2/\sqrt{M})\left(\sqrt{1-m/\bar{R}^4}+\sqrt{2\ln(4/\bar{\delta})}\right) = \bar{\alpha}(\bar{\delta}/4) \end{aligned}$$

and since

$$\beta(\bar{\delta}/2) = (R^2/M)\left(2+\sqrt{2\ln(2/\bar{\delta})}\right)^2 \leq (\bar{R}^2/M)\left(2+\sqrt{2\ln(2/\bar{\delta})}\right)^2 = \bar{\beta}(\bar{\delta}/2) \ ,$$

we can conclude with a second application of Theorem 2.3.4 that with probability greater than $1-\bar{\delta}$ the following statements are satisfied:

$$\begin{aligned}
&(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\mu}}-\boldsymbol{\mu}) \leq \beta(\bar{\delta}/2) \leq \bar{\beta}(\bar{\delta}/2) \ , \\
&\boldsymbol{\Sigma} \preceq \frac{1}{1-\alpha(\delta/4)-\beta(\delta/2)}\hat{\boldsymbol{\Sigma}} \preceq \frac{1}{1-\bar{\alpha}(\bar{\delta}/4)-\bar{\beta}(\bar{\delta}/2)}\hat{\boldsymbol{\Sigma}} \ , \\
&\boldsymbol{\Sigma} \succeq \frac{1}{1-\alpha(\delta/4)}\hat{\boldsymbol{\Sigma}} \succeq \frac{1}{1-\bar{\alpha}(\delta/4)}\hat{\boldsymbol{\Sigma}} \ .
\end{aligned}$$

It follows that Theorem 2.3.4 applies with $\bar{\alpha}(\bar{\delta}/4)$ and $\bar{\beta}(\bar{\delta}/4)$ because the probability that the event $\mathcal{E}$ that constraints (2.7a), (2.7b), and (2.7c) equipped with $\bar{\alpha}(\bar{\delta}/4)$ and $\bar{\beta}(\bar{\delta}/4)$ are met is necessarily greater than $1-\delta$:

$$\mathbb{P}(\mathcal{E}) \geq \mathbb{P}(\mathcal{E}|R \leq \bar{R})\mathbb{P}(R \leq \bar{R}) \geq (1-\bar{\delta})(1-\bar{\delta}) = 1-\delta \ . \ \square$$

2.5 Conclusion

In this chapter, we derived confidence regions for the mean and covariance matrix of a random vector which relied on the observation of a finite set of i.i.d. samples and the knowledge of a ball containing the support of the random vector. This result allows us to have confidence that the unknown distribution F associated with $\boldsymbol{\xi}$ must be such that it satisfies the following three conditions:

$$\begin{aligned}
&\mathbb{P}(\boldsymbol{\xi} \in \mathcal{S}) = 1 \\
&(\mathbb{E}[\boldsymbol{\xi}]-\hat{\boldsymbol{\mu}})^{\mathsf{T}}\hat{\boldsymbol{\Sigma}}^{-1}(\mathbb{E}[\boldsymbol{\xi}]-\hat{\boldsymbol{\mu}}) \leq \gamma_1 \\
&\mathbb{E}\left[(\boldsymbol{\xi}-\mathbb{E}[\boldsymbol{\xi}])(\boldsymbol{\xi}-\mathbb{E}[\boldsymbol{\xi}])^{\mathsf{T}}\right] \preceq \gamma_2\hat{\boldsymbol{\Sigma}} \ .
\end{aligned}$$

In what follows, we will use this fact to formulate and solve a version of the stochastic program which is robust with respect to any distribution that meets these conditions. The results that were just presented will for the first time enable the possibility of taking decisions that have strong probabilistic guarantees with respect to the true distribution behind the historical data instead of an estimate of this distribution as is typically done.

Chapter 3

Distributionally Robust Optimization under Moment Uncertainty

It is a common belief that the distributionally robust stochastic program presents a real computational challenge when applied to a rich decision problem. In fact, it has been shown that if one imposes that the distribution must satisfy some known support and moments values then the problem easily becomes NP-hard. In this chapter, we show that a well-justified relaxation of this distributional set actually leads to a version of the distributionally robust optimization model that can be solved in polynomial time for a wide range of problems. Leveraging on the new confidence regions that were presented in Section 2, we can now provide solutions to data-driven problems which have strong probabilistic guarantees with respect to performance. Thus, the main contribution of this chapter is to present theory that permits the formulation of tractable versions of the DRSP which account accurately for the available information about the distribution of uncertain parameters.

3.1 Introduction

It is often the case in practice that one has limited information about the distribution F driving the uncertain parameters which are involved in the decision making process. In such situations, it is commonly considered safer to only rely on estimates of the mean $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}_0$ of the random vector: *e.g.*, using empirical estimates. However, as discussed in Chapter 2, it is actually rarely the case that one is entirely confident in these estimates. For this reason, in this work we propose representing this uncertainty using two constraints parameterized by $\gamma_1 \geq 0$ and $\gamma_2 \geq 1$:

$$(\mathbb{E}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0)^\mathsf{T} \boldsymbol{\Sigma}_0^{-1} (\mathbb{E}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0) \leq \gamma_1 \tag{3.1a}$$

$$\mathbb{E}[(\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\mathsf{T}] \preceq \gamma_2 \boldsymbol{\Sigma}_0 \ . \tag{3.1b}$$

While Constraint (3.1a) assumes that the mean of $\boldsymbol{\xi}$ lies in an ellipsoid of size γ_1 centered at the estimate $\boldsymbol{\mu}_0$, Constraint (3.1b) forces $\mathbb{E}[(\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\mathsf{T}]$, the centered second moment matrix of $\boldsymbol{\xi}$, to lie in a positive semi-definite cone defined with a matrix inequality. In other words, it describes how likely $\boldsymbol{\xi}$ is to be close to $\boldsymbol{\mu}_0$ in terms of the correlations expressed in $\boldsymbol{\Sigma}_0$. Finally, the parameters γ_1 and γ_2 provide natural means of quantifying one's confidence in $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ respectively.

In what follows, we will study the DRSP model under the distributional set

$$\mathcal{D}_1(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2) = \left\{ F \in \mathcal{M} \;\middle|\; \begin{array}{l} \mathbb{P}(\boldsymbol{\xi} \in \mathcal{S}) = 1 \\ (\mathbb{E}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0)^\mathsf{T} \boldsymbol{\Sigma}_0^{-1} (\mathbb{E}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0) \leq \gamma_1 \\ \mathbb{E}[(\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\mathsf{T}] \preceq \gamma_2 \boldsymbol{\Sigma}_0 \end{array} \right\} ,$$

where $\mathcal{M}$ is the set of all probability measures on the measurable space $(\mathbb{R}^m, \mathcal{B})$, with $\mathcal{B}$ the Borel σ-algebra on $\mathbb{R}^m$, and $\mathcal{S} \subseteq \mathbb{R}^m$ is any closed convex set known to contain the support of F. The set $\mathcal{D}_1(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$, which will also be referred to in short-hand notation as $\mathcal{D}_1$, can be seen as a generalization of many previously proposed sets. For example, $\mathcal{D}_1(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{I}, 0, \infty)$ imposes exact mean and support constraints as is studied in [16] and in [6]. Similarly, $\mathcal{D}_1(\mathbb{R}^m, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, 0, 1)$ relates closely to the exact mean and covariance matrix constraints considered in [52, 65, 45]. We will soon show that there is a lot to be gained, both on a theoretical and practical point of view, by formulating the DRSP model using the set $\mathcal{D}_1(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$ which constrains all three types of statistics: support, mean and centered second moment matrix.

Remark 3.1.1 : While our proposed uncertainty model cannot be used to express an arbitrarily large confidence in the second-order statistics of $\boldsymbol{\xi}$, in Section 3.4 we will use the results presented in Chapter 2 to propose a natural way of assigning $\boldsymbol{\mu}_0$, $\boldsymbol{\Sigma}_0$, γ_1 and γ_2 based on historical data. Of course, in some situations it might be interesting to include the following constraint in the description of $\mathcal{D}_1$:

$$\gamma_3 \boldsymbol{\Sigma}_0 \preceq \mathbb{E}\left[(\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^{\mathsf{T}}\right] \ , \tag{3.2}$$

where $0 \leq \gamma_3 \leq \gamma_2$. Unfortunately, this leads to important computational difficulties for the DRSP model. Furthermore, in most applications of our model, we expect the worst case distribution to actually achieve maximum variance, thus making Constraint (3.2) irrelevant. For example, both an instance of a stochastic linear program and an instance of the portfolio optimization problem, presented in Section 3.5 and Section 4.3 respectively, will have this characteristic.

We first show that with this distributional set the distributionally robust optimization model can be solved in polynomial time for a large range of objective functions. In fact, the structure of our distribution set allows us to solve instances of the DRSP that are known to be intractable under a moment matching approach (see sections 3.5 and 4.1 for more details). Using the results presented in Chapter 2, we then show that when the problem is data-driven, the solution of our proposed optimization model actually has strong probabilistic guarantees with respect to its performance measured using the true distribution that generates the data. We further describe the implications of these results for three forms of stochastic programs: a two-stage stochastic linear program with cost uncertainty, an expected utility maximization model, and a conditional value-at-risk minimization model. We finally close this chapter by extending the type of distribution information that can be accounted for in $\mathcal{D}_1$ without losing the property that the resulting optimization model be tractable.

3.2 Complexity of the Inner Moment Problem

We start by considering the question of solving the inner maximization problem of a DRSP that uses the set $\mathcal{D}_1$.

Definition 3.2.1 : Given any fixed $\boldsymbol{x} \in \mathcal{X}$, let $\Psi(\boldsymbol{x};\gamma_1,\gamma_2)$ be the optimal value of the moment problem:

$$\underset{F \,\in\, \mathcal{D}_1}{\text{maximize}} \quad \mathbb{E}_F[h(\boldsymbol{x},\boldsymbol{\xi})] \ , \tag{3.3}$$

where $\mathbb{E}_F[\cdot]$ is the expectation taken with respect to the random vector $\boldsymbol{\xi}$ given that it follows the probability distribution $F \in \mathcal{D}_1$.

Since F is a probability measure on $(\mathbb{R}^m, \mathcal{B})$, Problem (3.3) can be described as the conic linear problem:

$$\begin{aligned}
&\underset{F}{\text{maximize}} && \int_{\mathcal{S}} h(\boldsymbol{x},\boldsymbol{\xi})dF(\boldsymbol{\xi}) && \text{(3.4a)}\\
&\text{subject to} && \int_{\mathcal{S}} dF(\boldsymbol{\xi}) = 1 && \text{(3.4b)}\\
& && \int_{\mathcal{S}} (\boldsymbol{\xi}-\boldsymbol{\mu}_0)(\boldsymbol{\xi}-\boldsymbol{\mu}_0)^{\mathsf{T}} dF(\boldsymbol{\xi}) \preceq \gamma_2 \boldsymbol{\Sigma}_0 && \text{(3.4c)}\\
& && \int_{\mathcal{S}} \begin{bmatrix} \boldsymbol{\Sigma}_0 & (\boldsymbol{\xi}-\boldsymbol{\mu}_0) \\ (\boldsymbol{\xi}-\boldsymbol{\mu}_0)^{\mathsf{T}} & \gamma_1 \end{bmatrix} dF(\boldsymbol{\xi}) \succeq 0 && \text{(3.4d)}\\
& && F \in \mathcal{M} \ . && \text{(3.4e)}
\end{aligned}$$

As it is often done with moment problems of this form, we can circumvent the difficulty of finding the optimal value of this problem by making use of duality theory. For more information on duality theory, we refer the reader to [10] for finite dimensional problems, to [50, 51] for infinite dimensional problems, and to [30, 53] specifically for moment problems like Problem (3.4).

Lemma 3.2.2 : *For a fixed $\boldsymbol{x} \in \mathbb{R}^n$, suppose that $\gamma_1 \geq 0$, $\gamma_2 \geq 1$, $\boldsymbol{\Sigma}_0 \succ 0$, and that $h(\boldsymbol{x},\boldsymbol{\xi})$ is F-integrable for all $F \in \mathcal{D}_1$. Then, $\Psi(\boldsymbol{x};\gamma_1,\gamma_2)$ must be equal to the optimal value of the problem:*

$$\begin{aligned}
&\underset{\boldsymbol{Q},\boldsymbol{q},r,t}{\text{minimize}} && r + t && \text{(3.5a)}\\
&\text{subject to} && r \geq h(\boldsymbol{x},\boldsymbol{\xi}) - \boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{Q}\boldsymbol{\xi} - \boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{q} \ \ \forall\, \boldsymbol{\xi} \in \mathcal{S} && \text{(3.5b)}\\
& && t \geq \left(\gamma_2\boldsymbol{\Sigma}_0 + \boldsymbol{\mu}_0\boldsymbol{\mu}_0^{\mathsf{T}}\right) \bullet \boldsymbol{Q} + \boldsymbol{\mu}_0^{\mathsf{T}}\boldsymbol{q} + \sqrt{\gamma_1}\,\|\boldsymbol{\Sigma}_0^{1/2}(\boldsymbol{q} + 2\boldsymbol{Q}\boldsymbol{\mu}_0)\| && \text{(3.5c)}\\
& && \boldsymbol{Q} \succeq 0 \ , && \text{(3.5d)}
\end{aligned}$$

where $(\boldsymbol{A} \bullet \boldsymbol{B})$ *refers to the Frobenius inner product between matrices,* $\boldsymbol{Q} \in \mathbb{R}^{m \times m}$ *is a symmetric matrix, the vector* $\boldsymbol{q} \in \mathbb{R}^m$*, and* $r, t \in \mathbb{R}$*. In addition, if* $\Psi(\boldsymbol{x}; \gamma_1, \gamma_2)$ *is finite then the set of optimal solutions to Problem* (3.5) *must be non-empty.*

We defer the proof of this Lemma to Appendix A.1 since it is the result of applying some well established concepts in duality theory.

To show that there exists a tractable solution method for solving Problem (3.5), we employ a famous equivalence between convex optimization and separation of a convex set from a point.

Lemma 3.2.3 : *(Grötschel* et al. *[29]) Consider a convex optimization problem of the form*

$$\underset{\boldsymbol{z} \in \mathcal{Z}}{\text{minimize}} \quad \boldsymbol{c}^{\mathsf{T}} \boldsymbol{z}$$

with linear objective and convex feasible set $\mathcal{Z}$*. Given that the set of optimal solutions is non-empty, the problem can be solved to any precision* ϵ *in time polynomial in* $\log(1/\epsilon)$ *and in the size of the problem by using the ellipsoid method if and only if* $\mathcal{Z}$ *satisfies the following two conditions :*

1. *for any* $\bar{\boldsymbol{z}}$*, it can be verified whether* $\bar{\boldsymbol{z}} \in \mathcal{Z}$ *or not in time polynomial in the dimension of* $\boldsymbol{z}$*;*

2. *for any infeasible* $\bar{\boldsymbol{z}}$*, a hyperplane that separates* $\bar{\boldsymbol{z}}$ *from the feasible set* $\mathcal{Z}$ *can be generated in time polynomial in the dimension of* $\boldsymbol{z}$*.*

A first application of this lemma leads to quantifying the difficulty of solving the feasibility problem associated with Constraint (3.5b).

Assumption 3.2.4 : The support set $\mathcal{S} \subset \mathbb{R}^m$ is convex and compact (closed and bounded), and it is equipped with an oracle that can for any $\boldsymbol{\xi} \in \mathbb{R}^m$ either confirm that $\boldsymbol{\xi} \in \mathcal{S}$ or provide a hyperplane that separates $\boldsymbol{\xi}$ from $\mathcal{S}$ in time polynomial in m.

Lemma 3.2.5 : *Let function* $h(\boldsymbol{x}, \boldsymbol{\xi})$ *be concave in* $\boldsymbol{\xi}$ *and be such that one can provide a super-gradient of* $\boldsymbol{\xi}$ *in time polynomial in* m*. Then, under Assumption 3.2.4, for any fixed* *assignment* $\boldsymbol{x}$, $\boldsymbol{Q} \succeq 0$, *and* $\boldsymbol{q}$, *one can find an assignment* $\boldsymbol{\xi}_*$ *that is* ϵ*-optimal with respect to the problem*

$$\underset{t,\boldsymbol{\xi}}{\text{maximize}} \quad t \tag{3.6a}$$

$$\text{subject to} \quad t \leq h(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{\xi}^\mathsf{T} \boldsymbol{Q} \boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T} \boldsymbol{q} \tag{3.6b}$$

$$\boldsymbol{\xi} \in \mathcal{S} \ , \tag{3.6c}$$

in time polynomial in $\log(1/\epsilon)$ *and the size of the problem.*

Proof: First, the feasible set of the problem is convex since $\boldsymbol{Q} \succeq 0$ so that $h(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{\xi}^\mathsf{T} \boldsymbol{Q} \boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T} \boldsymbol{q}$ is a concave function in $\boldsymbol{\xi}$. Because $\mathcal{S}$ is compact, the set of optimal solutions for Problem (3.6) is therefore non-empty. By Assumption 3.2.4, Condition (1), and Condition (2) in Lemma 3.2.3 are met for Constraint (3.6c). On the other hand, feasibility of Constraint (3.6b) can be verified directly after the evaluation of $h(\boldsymbol{x}, \boldsymbol{\xi})$; furthermore, for an infeasible assignment $(\bar{\boldsymbol{\xi}}, \bar{t})$, the following separating hyperplane can be generated in polynomial time:

$$t - (\nabla_{\boldsymbol{\xi}} h(\boldsymbol{x}, \bar{\boldsymbol{\xi}}) - 2\boldsymbol{Q}\bar{\boldsymbol{\xi}} - \boldsymbol{q})^\mathsf{T} \boldsymbol{\xi} \leq h(\boldsymbol{x}, \bar{\boldsymbol{\xi}}) - \nabla_{\boldsymbol{\xi}} h(\boldsymbol{x}, \bar{\boldsymbol{\xi}})^\mathsf{T} \bar{\boldsymbol{\xi}} + \bar{\boldsymbol{\xi}}^\mathsf{T} \boldsymbol{Q} \bar{\boldsymbol{\xi}} \ ,$$

where $\nabla_{\boldsymbol{\xi}} h(\boldsymbol{x}, \boldsymbol{\xi})$ is a super-gradient of $h(\boldsymbol{x}, \cdot)$. It follows from Lemma 3.2.3 that the ellipsoid method will converge to an ϵ-optimal solution in polynomial time. □

We are now able to derive an important result about the complexity of solving Problem (3.5) under a general form for $h(\boldsymbol{x}, \boldsymbol{\xi})$.

Assumption 3.2.6 : The function $h(\boldsymbol{x}, \boldsymbol{\xi})$ has the form $h(\boldsymbol{x}, \boldsymbol{\xi}) = \max_{k \in \{1,\ldots,K\}} h_k(\boldsymbol{x}, \boldsymbol{\xi})$ such that for each k, $h_k(\boldsymbol{x}, \boldsymbol{\xi})$ is concave in $\boldsymbol{\xi}$. In addition, given a pair $(\boldsymbol{x}, \boldsymbol{\xi})$, it is assumed that one can in polynomial time:

1. evaluate the value of $h_k(\boldsymbol{x}, \boldsymbol{\xi})$
2. find a super-gradient of $h_k(\boldsymbol{x}, \boldsymbol{\xi})$ in $\boldsymbol{\xi}$.

Furthermore, for any $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{q} \in \mathbb{R}^m$, and positive semi-definite $\boldsymbol{Q} \in \mathbb{R}^{m \times m}$, the set $\{y \in \mathbb{R} \mid \exists \boldsymbol{\xi} \in \mathcal{S},\ y \leq h(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{q}^\mathsf{T} \boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T} \boldsymbol{Q} \boldsymbol{\xi}\}$ is closed.

Proposition 3.2.7 : *Given that $\mathcal{S}$ satisfies Assumption 3.2.4 and that $h(\boldsymbol{x},\boldsymbol{\xi})$ satisfies Assumption 3.2.6 and satisfies the condition of Lemma 3.2.2, then Problem* (3.5) *is a convex optimization problem whose optimal value is finite and equal to $\Psi(\boldsymbol{x};\gamma_1,\gamma_2)$. Moreover, Problem* (3.5) *can be solved to any precision ϵ in time polynomial in $\log(1/\epsilon)$ and the size of the problem.*

Proof: First, the constraints of Problem (3.5) describe a convex set since for any $\boldsymbol{\xi}\in\mathcal{S}$ the function $h(\boldsymbol{x},\boldsymbol{\xi})-\boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{Q}\boldsymbol{\xi}-\boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{q}$ is linear in $(\boldsymbol{Q},\boldsymbol{q})$, and the function $\left(\gamma_2\boldsymbol{\Sigma}_0+\boldsymbol{\mu}_0\boldsymbol{\mu}_0^{\mathsf{T}}\right)\bullet\boldsymbol{Q}+\boldsymbol{\mu}_0^{\mathsf{T}}\boldsymbol{q}+\sqrt{\gamma_1}\,\|\boldsymbol{\Sigma}_0^{1/2}(\boldsymbol{q}+2\boldsymbol{Q}\boldsymbol{\mu}_0)\|$ is convex in $(\boldsymbol{Q},\boldsymbol{q})$. The feasible set is also non-empty since the assignment $\boldsymbol{Q}=\boldsymbol{I}$, $\boldsymbol{q}=0$, $t=\mathbf{trace}\left(\gamma_2\boldsymbol{\Sigma}_0+\boldsymbol{\mu}_0\boldsymbol{\mu}_0^{\mathsf{T}}\right)+2\sqrt{\gamma_1}\,\|\boldsymbol{\Sigma}_0^{1/2}\boldsymbol{\mu}_0\|$, $r=\sup_{\boldsymbol{\xi}\in\mathcal{S}}\ \max_{k\in\{1,\ldots,K\}}\ h_k(\boldsymbol{x},\boldsymbol{\xi})-\|\boldsymbol{\xi}\|^2$ is necessarily feasible. Note that the assumption that each $h_k(\boldsymbol{x},\boldsymbol{\xi})$ is concave ensures that such an assignment for r exists. Based on Lemma 3.2.2 and the fact that the optimal value of Problem (3.4) is bounded below by $h(\boldsymbol{x},\boldsymbol{\mu}_0)$ since the Dirac measure[1] $\delta_{\boldsymbol{\mu}_0}$ is in $\mathcal{D}_1$, we can conclude that $\Psi(\boldsymbol{x};\gamma_1,\gamma_2)$ is finite and that the optimal solution set of Problem (3.5) is non-empty.

We can now use Lemma 3.2.3 to show that Problem (3.5) can be solved efficiently given that we verify the two conditions for each of its constraints. In the case of Constraint (3.5d), feasibility can be verified in $O(m^3)$ arithmetic operations. Moreover, a separating hyperplane can be generated, if necessary, based on the eigenvector corresponding to the lowest eigenvalue. The feasibility of Constraint (3.5c) is also easily verified. Based on an infeasible assignment $(\bar{\boldsymbol{Q}},\bar{\boldsymbol{q}},\bar{r},\bar{t})$, a separating hyperplane can be constructed in polynomial time:

$$\begin{aligned}\left(\gamma_2\boldsymbol{\Sigma}_0+\boldsymbol{\mu}_0\boldsymbol{\mu}_0^{\mathsf{T}}+\nabla_{\boldsymbol{Q}}g(\bar{\boldsymbol{Q}},\bar{\boldsymbol{q}})\right)\bullet\boldsymbol{Q}+&(\boldsymbol{\mu}_0+\nabla_{\boldsymbol{q}}g(\bar{\boldsymbol{Q}},\bar{\boldsymbol{q}}))^{\mathsf{T}}\boldsymbol{q}-t\\ &\le\nabla_{\boldsymbol{q}}g(\bar{\boldsymbol{Q}},\bar{\boldsymbol{q}})^{\mathsf{T}}\bar{\boldsymbol{q}}+\nabla_{\boldsymbol{Q}}g(\bar{\boldsymbol{Q}},\bar{\boldsymbol{q}})\bullet\bar{\boldsymbol{Q}}-g(\bar{\boldsymbol{Q}},\bar{\boldsymbol{q}})\ ,\end{aligned}$$

where $g(\boldsymbol{Q},\boldsymbol{q})=\sqrt{\gamma_1}\,\|\boldsymbol{\Sigma}_0^{1/2}(\boldsymbol{q}+2\boldsymbol{Q}\boldsymbol{\mu}_0)\|$ and where $\nabla_{\boldsymbol{Q}}g(\boldsymbol{Q},\boldsymbol{q})$ and $\nabla_{\boldsymbol{q}}g(\boldsymbol{Q},\boldsymbol{q})$ are the gradients of $g(\boldsymbol{Q},\boldsymbol{q})$ in $\boldsymbol{Q}$ and $\boldsymbol{q}$ respectively. Finally, given the assumed structure of $h(\boldsymbol{x},\boldsymbol{\xi})$, Constraint (3.5b) can be decomposed into K sub-constraints

$$r\ge h_k(\boldsymbol{x},\boldsymbol{\xi})-\boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{Q}\boldsymbol{\xi}-\boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{q}\ \ \forall\,\boldsymbol{\xi}\in\mathcal{S}\ \ \forall\,k\in\{1,2,...,K\}$$

[1] Recall that the Dirac measure $\delta_{\boldsymbol{a}}$ is the measure of mass one at the point $\boldsymbol{a}$.

Considering the k-th sub-constraint, Lemma 3.2.5 states that $\sup_{\boldsymbol{\xi}\in\mathcal{S}} h_k(\boldsymbol{x},\boldsymbol{\xi})-\boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{Q}\boldsymbol{\xi}-\boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{q}$ can be solved to any ϵ-precision in polynomial time. Given that the optimal value is found to be above $r+\epsilon$, one can conclude infeasibility of the constraint and generate an associated separating hyperplane using any optimal solution $\boldsymbol{\xi}_*$ as follows:

$$(\boldsymbol{\xi}_*\boldsymbol{\xi}_*^{\mathsf{T}} \bullet \boldsymbol{Q}) + \boldsymbol{\xi}_*^{\mathsf{T}}\boldsymbol{q} + r \geq h_{k^*}(\boldsymbol{x},\boldsymbol{\xi}_*) \ \ .$$

Since K is finite, the conditions derived from [29] are necessarily met by Problem (3.5). We therefore conclude that $\Psi(\boldsymbol{x};\gamma_1,\gamma_2)$ can be computed up to any precision ϵ in polynomial time using the ellipsoid method. □

3.3 Complexity of Distributionally Robust Optimization

Based on our result with the inner moment problem, we can now address the existence of a tractable solution method for the DRSP model under the distributional set $\mathcal{D}_1$:

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \left(\max_{F\,\in\,\mathcal{D}_1} \ \mathbb{E}_F[h(\boldsymbol{x},\boldsymbol{\xi})]\right) \ \ . \tag{3.7}$$

Assumption 3.3.1 : The set $\mathcal{X} \subset \mathbb{R}^n$ is convex and compact (closed and bounded), and it is equipped with an oracle that can for any $\boldsymbol{x} \in \mathbb{R}^n$ either confirm that $\boldsymbol{x} \in \mathcal{X}$ or provide a hyperplane that separates $\boldsymbol{x}$ from $\mathcal{X}$ in time polynomial in n.

Assumption 3.3.2 : The function $h(\boldsymbol{x},\boldsymbol{\xi})$ is convex in $\boldsymbol{x}$. In addition, it is assumed that one can find in polynomial time a sub-gradient of $h(\boldsymbol{x},\boldsymbol{\xi})$ in $\boldsymbol{x}$.

Based on these new assumptions, the proposition that follows states that the distributionally robust optimization model is tractable. It is actually the case that the model is tractable under weaker conditions on the separation oracles of $\mathcal{S}$ and $\mathcal{X}$, yet for clarity of exposure we defer such technical derivations to Appendix A.2.

Proposition 3.3.3 : *Given that assumptions 3.2.4, 3.2.6, 3.3.1, and 3.3.2 hold, then the DRSP model presented in Problem* (3.7) *can be solved to any precision* ϵ *in time polynomial in* $\log(1/\epsilon)$ *and the sizes of* $\boldsymbol{x}$ *and* $\boldsymbol{\xi}$.

Proof: The proof of this theorem follows similar lines as the proof for Proposition 3.2.7. We first reformulate the inner moment problem in its dual form and use the fact that min-min operations can be performed jointly and that the constraint involving $h(\boldsymbol{x}, \boldsymbol{\xi})$ decomposes. This leads to an equivalent convex optimization form for Problem (3.7):

$$\begin{aligned}
\underset{\boldsymbol{x},\boldsymbol{Q},\boldsymbol{q},r,t}{\text{minimize}} \quad & r + t && \text{(3.8a)}\\
\text{subject to} \quad & r \geq h_k(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{Q}\boldsymbol{\xi} - \boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{q} \ , \ \ \forall\, \boldsymbol{\xi} \in \mathcal{S}, \ \ k \in \{1, ..., K\} && \text{(3.8b)}\\
& t \geq \left(\gamma_2 \boldsymbol{\Sigma}_0 + \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^{\mathsf{T}}\right) \bullet \boldsymbol{Q} + \boldsymbol{\mu}_0^{\mathsf{T}} \boldsymbol{q} + \sqrt{\gamma_1}\, \|\boldsymbol{\Sigma}_0^{1/2}(\boldsymbol{q} + 2\boldsymbol{Q}\boldsymbol{\mu}_0)\| && \text{(3.8c)}\\
& \boldsymbol{Q} \succeq 0 && \text{(3.8d)}\\
& \boldsymbol{x} \in \mathcal{X} \ . && \text{(3.8e)}
\end{aligned}$$

As in the proof of Proposition 3.2.7, we need to show that the ellipsoid method can be successfully applied. Because $\mathcal{X}$ is compact and non-empty, similar arguments to those presented in Proposition 3.2.7 ensure that the optimal solution set is once again non-empty. Regarding Condition (1) and Condition (2), the arguments that were presented in the proof of Proposition 3.2.7 still apply for Constraint (3.8c) and Constraint (3.8d). However, the argument for Constraint (3.8b) needs to be revisited since $\boldsymbol{x}$ is now considered an optimization variable. Feasibility of an assignment $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{Q}}, \bar{\boldsymbol{q}}, \bar{r})$ can still be verified in polynomial time because of Lemma 3.2.5 and of the fact that K is finite. However, in the case that one of the indexed constraints, say the k^*-th one, is found to be infeasible, one now needs to generate a separating hyperplane using the worst case $\boldsymbol{\xi}_*$ and $\nabla_{\boldsymbol{x}} h_k(\bar{\boldsymbol{x}}, \boldsymbol{\xi}_*)$, a sub-gradient of $h_{k^*}(\cdot, \boldsymbol{\xi}_*)$ at $\bar{\boldsymbol{x}}$:

$$(\boldsymbol{\xi}_* {\boldsymbol{\xi}_*}^{\mathsf{T}} \bullet \boldsymbol{Q}) + {\boldsymbol{\xi}_*}^{\mathsf{T}} \boldsymbol{q} + r - \nabla_{\boldsymbol{x}} h_{k^*}(\bar{\boldsymbol{x}}, \boldsymbol{\xi}_*)^{\mathsf{T}} \boldsymbol{x} \geq h_{k^*}(\bar{\boldsymbol{x}}, \boldsymbol{\xi}_*) - \nabla_{\boldsymbol{x}} h_{k^*}(\bar{\boldsymbol{x}}, \boldsymbol{\xi}_*)^{\mathsf{T}} \bar{\boldsymbol{x}} \ .$$

Since by Assumption 3.3.2, a sub-gradient $\nabla_{\boldsymbol{x}} h_k(\bar{\boldsymbol{x}}, \boldsymbol{\xi}_*)$ can be obtained in polynomial time and since, by Assumption 3.3.1, the conditions are met for Constraint (3.8e), we can conclude that Lemma 3.2.3 can be applied. Problem (3.8) can therefore be solved to any precision in polynomial time. □

We believe this result should be of high significance for both theoreticians and practitioners as it indicates that, if $\min_{\boldsymbol{x}} \max_{\boldsymbol{\xi}\in\mathcal{S}} h(\boldsymbol{x},\boldsymbol{\xi})$ is a tractable robust optimization problem (as discussed in [2] and [4]), then the less conservative DRSP with $\mathcal{D}_1$ is also tractable. In some cases, the inner moment problem might even be reducible (see Section 4.3 for an example). Moreover, one also has access to the wide spectrum of methods for robust optimization problems: ranging from methods that use cutting planes more efficiently than the ellipsoid method such as in [26, 64, 8], to methods that approximate the feasible set with a finite number of sampled constraints such as in [14, 11].

Remark 3.3.4 : The constraint $\boldsymbol{Q} \succeq 0$ plays an important role in making Problem (3.7) solvable in polynomial time. This constraint corresponds to the second moment matrix inequality in the construction of our distribution set $\mathcal{D}_1$. If the inequality is replaced by an equality, then $\boldsymbol{Q}$ becomes "free" and Problem (3.6) is no longer ensured to be a convex optimization problem. This explains why many DRSP problems under the "exact covariance constraint" actually become intractable.

Remark 3.3.5 : We also remark that the bounded condition on $\mathcal{S}$ in Assumption 3.2.4 is imposed in order to simplify the exposition of our results. In the case that $\mathcal{S}$ is unbounded, Proposition 3.2.7 and Proposition 3.3.3 will hold as long as feasibility with respect to Constraint (3.5b) can be verified in polynomial time. In fact, given an infeasible assignment $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{Q}}, \bar{\boldsymbol{q}}, \bar{r})$, one can interrupt the solution process for Problem (3.6) when the achieved maximum is higher than $\bar{r}$. This is guaranteed to occur in polynomial time since either the problem is unbounded above or the set of optimal t^* for Problem (3.6) is non-empty due to the technical condition in Assumption 3.2.6.

3.4 Reliable Solutions for Data-driven Problems

In some practical situations where one wishes to formulate a DRSP model, it might not be clear how to define an uncertainty set for the mean and second moment matrix of the random vector of parameters $\boldsymbol{\xi}$. It is more likely the case that one only has in hand a set of independent samples, $\{\boldsymbol{\xi}_i\}_{i=1}^M$, drawn according to the distribution of $\boldsymbol{\xi}$ and wishes to guarantee that the solution of the DRSP model is robust with respect to the unknown distribution of the random vector $\boldsymbol{\xi}$.

We first use our recent result to define, based on the samples $\{\boldsymbol{\xi}_i\}_{i=1}^M$, a set of distributions which is known to contain the distribution of $\boldsymbol{\xi}$ with high probability, given that M is sufficiently large.

Definition 3.4.1 : Given a set $\{\boldsymbol{\xi}_i\}_{i=1}^M$ of M samples, for any $\delta > 0$ let $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$, $\bar{\gamma}_1$ and $\bar{\gamma}_2$ be defined as

$$\hat{\boldsymbol{\mu}} = \frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\xi}_i \ , \qquad \hat{\boldsymbol{\Sigma}} = \frac{1}{M}\sum_{i=1}^{M}(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})^{\mathsf{T}}$$
$$\bar{\gamma}_1 = \frac{\bar{\beta}(\bar{\delta}/2)}{1-\bar{\alpha}(\bar{\delta}/4)-\bar{\beta}(\bar{\delta}/2)} \ , \qquad \bar{\gamma}_2 = \frac{1+\bar{\beta}(\bar{\delta}/2)}{1-\bar{\alpha}(\bar{\delta}/4)-\bar{\beta}(\bar{\delta}/2)} \ .$$

where $\bar{\alpha}(\bar{\delta}/4) = O(1/\sqrt{M})$ and $\bar{\beta}(\bar{\delta}/2) = O(1/M)$ are constants defined in Corollary 2.4.1.

Note that it follows from Definition 3.4.1 that $\bar{\gamma}_1 \rightarrow 0$ and $\bar{\gamma}_2 \rightarrow 1$ as M goes to infinity.

Corollary 3.4.2 : *Let $\{\boldsymbol{\xi}_i\}_{i=1}^M$ be a set of M samples generated independently at random according to the distribution of $\boldsymbol{\xi}$. If M satisfies Constraint* (2.8) *and $\boldsymbol{\xi}$ has a support contained in a bounded set $\mathcal{S}$, then with probability greater than $1-\delta$ over the choice of $\{\boldsymbol{\xi}_i\}_{i=1}^M$, the distribution of $\boldsymbol{\xi}$ lies in the set $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \bar{\gamma}_1, \bar{\gamma}_2)$.*

Proof: This result can be derived from Corollary 2.4.1. One can show that given any estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ that satisfy both constraints (2.7a) and (2.7b) equipped with $\bar{\alpha}(\bar{\delta}/4)$ and $\bar{\beta}(\bar{\delta}/2)$, these estimates should also satisfy constraints (3.1a) and (3.1b). First, Constraint (3.1a) is necessarily met since for such $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$,

$$\begin{aligned}(1-\bar{\alpha}(\bar{\delta}/4)-\bar{\beta}(\bar{\delta}/2))(\hat{\boldsymbol{\mu}}-\mathbb{E}[\boldsymbol{\xi}])^{\mathsf{T}}\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\mu}}-\mathbb{E}[\boldsymbol{\xi}]) &\leq (\hat{\boldsymbol{\mu}}-\mathbb{E}[\boldsymbol{\xi}])^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\mu}}-\mathbb{E}[\boldsymbol{\xi}]) \\ &\leq \bar{\beta}(\bar{\delta}/2) \ ,\end{aligned}$$

where we used the fact that Constraint (2.7a) implies that $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{x} \geq (1-\bar{\alpha}(\bar{\delta}/4) - \bar{\beta}(\bar{\delta}/2))\boldsymbol{x}^{\mathsf{T}}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{x}$ for any $\boldsymbol{x} \in \mathbb{R}^m$. Similarly, the same $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ can be shown to satisfy

Constraint (3.1b):

$$\begin{aligned}\frac{1}{1-\bar{\alpha}(\bar{\delta}/4)-\bar{\beta}(\bar{\delta}/2)}\hat{\boldsymbol{\Sigma}} &\succeq \boldsymbol{\Sigma} = \mathbb{E}[\boldsymbol{\xi}\boldsymbol{\xi}^\mathsf{T}] - \mathbb{E}[\boldsymbol{\xi}]\mathbb{E}[\boldsymbol{\xi}]^\mathsf{T} \\ &\succeq \mathbb{E}[(\boldsymbol{\xi}-\hat{\boldsymbol{\mu}})(\boldsymbol{\xi}-\hat{\boldsymbol{\mu}})^\mathsf{T}] - \frac{\bar{\beta}(\bar{\delta}/2)}{1-\bar{\alpha}(\bar{\delta}/4)-\bar{\beta}(\bar{\delta}/2)}\hat{\boldsymbol{\Sigma}} \ ,\end{aligned}$$

since for all $\boldsymbol{x} \in \mathbb{R}^m$,

$$\begin{aligned}\boldsymbol{x}^\mathsf{T}\mathbb{E}[\boldsymbol{\xi}]\mathbb{E}[\boldsymbol{\xi}]^\mathsf{T}\boldsymbol{x} &= \left(\boldsymbol{x}^\mathsf{T}(\mathbb{E}[\boldsymbol{\xi}]-\hat{\boldsymbol{\mu}}+\hat{\boldsymbol{\mu}})\right)^2 \\ &= \left(\boldsymbol{x}^\mathsf{T}(\mathbb{E}[\boldsymbol{\xi}]-\hat{\boldsymbol{\mu}})\right)^2 + 2\boldsymbol{x}^\mathsf{T}(\mathbb{E}[\boldsymbol{\xi}]-\hat{\boldsymbol{\mu}})\hat{\boldsymbol{\mu}}^\mathsf{T}\boldsymbol{x} + \left(\boldsymbol{x}^\mathsf{T}\hat{\boldsymbol{\mu}}\right)^2 \\ &= \mathbf{trace}\left(\boldsymbol{x}^\mathsf{T}(\mathbb{E}[\boldsymbol{\xi}]-\hat{\boldsymbol{\mu}})(\mathbb{E}[\boldsymbol{\xi}]-\hat{\boldsymbol{\mu}})^\mathsf{T}\boldsymbol{x}\right) + 2\boldsymbol{x}^\mathsf{T}\mathbb{E}[\boldsymbol{\xi}]\hat{\boldsymbol{\mu}}^\mathsf{T}\boldsymbol{x} - (\boldsymbol{x}^\mathsf{T}\hat{\boldsymbol{\mu}})^2 \\ &\leq (\mathbb{E}[\boldsymbol{\xi}]-\hat{\boldsymbol{\mu}})^\mathsf{T}\boldsymbol{\Sigma}^{-1}(\mathbb{E}[\boldsymbol{\xi}]-\hat{\boldsymbol{\mu}})\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x} + 2\boldsymbol{x}^\mathsf{T}\mathbb{E}[\boldsymbol{\xi}]\hat{\boldsymbol{\mu}}^\mathsf{T}\boldsymbol{x} - (\boldsymbol{x}^\mathsf{T}\hat{\boldsymbol{\mu}})^2 \\ &\leq \boldsymbol{x}^\mathsf{T}\left(\frac{\bar{\beta}(\bar{\delta}/2)}{1-\bar{\alpha}(\bar{\delta}/4)-\bar{\beta}(\bar{\delta}/2)}\hat{\boldsymbol{\Sigma}} + \mathbb{E}[\boldsymbol{\xi}]\hat{\boldsymbol{\mu}}^\mathsf{T} + \hat{\boldsymbol{\mu}}\mathbb{E}[\boldsymbol{\xi}]^\mathsf{T} - \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\mathsf{T}\right)\boldsymbol{x} \\ &= \boldsymbol{x}^\mathsf{T}\left(\frac{\bar{\beta}(\bar{\delta}/2)}{1-\bar{\alpha}(\bar{\delta}/4)-\bar{\beta}(\bar{\delta}/2)}\hat{\boldsymbol{\Sigma}} + \mathbb{E}[\boldsymbol{\xi}\boldsymbol{\xi}^\mathsf{T}] - \mathbb{E}[(\boldsymbol{\xi}-\hat{\boldsymbol{\mu}})(\boldsymbol{\xi}-\hat{\boldsymbol{\mu}})^\mathsf{T}]\right)\boldsymbol{x} \ .\end{aligned}$$

By Corollary 2.4.1, the random variables $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are known to satisfy constraints (2.7a) and (2.7b) with probability greater than $1-\delta$, therefore they also satisfy constraints (3.1a) and (3.1b) with probability greater than $1-\delta$. □

We can now extend the results presented in sections 3.2 and 3.3 to a data-driven framework where moments of the distribution are estimated using independent samples. Based on the computational argument of Proposition 3.3.3 and the probabilistic guarantees provided by Corollary 3.4.2, we present an important result for data-driven problems.

Theorem 3.4.3 : *Let $\{\boldsymbol{\xi}_i\}_{i=1}^M$ be a set of M samples generated independently at random according to the distribution F whose support is contained in the set $\mathcal{S}$. For any $\delta > 0$, if assumptions 3.2.4, 3.2.6, 3.3.1, and 3.3.2 are satisfied then, given the set $\{\boldsymbol{\xi}_i\}_{i=1}^M$, one can solve Problem* (3.7) *in polynomial time under the set $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \bar{\gamma}_1, \bar{\gamma}_2)$ where $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$, $\bar{\gamma}_1$, and $\bar{\gamma}_2$ are assigned as in Definition 3.4.1. Furthermore, if M satisfies Constraint* (2.8), *then with probability greater than $1-\delta$ over the choice of $\{\boldsymbol{\xi}_i\}_{i=1}^M$, we have that any optimal* *solution $\boldsymbol{x}^*$ of the DRSP formed using these samples will satisfy the constraint*

$$\mathbb{E}\left[h(\boldsymbol{x}^*, \boldsymbol{\xi})\right] \leq \Psi(\boldsymbol{x}^*; \bar{\gamma}_1, \bar{\gamma}_2) \ \ ,$$

where $\mathbb{E}\left[\cdot\right]$ is the expectation with respect to the true distribution of $\boldsymbol{\xi}$.

Since we believe the moment problem to be interesting in its own right, we wish to mention a simple consequence of the above result for moment problems in a data-driven framework.

Corollary 3.4.4: *Let $\delta > 0$ and let $\{\boldsymbol{\xi}_i\}_{i=1}^M$ be a set of M samples generated independently at random according to the distribution F which support is contained in the set $\mathcal{S}$. For any $\delta > 0$ and function $g(\boldsymbol{\xi})$, if $\mathcal{S}$ satisfies Assumption 3.2.4 and the function $h(\boldsymbol{x}, \boldsymbol{\xi}) = g(\boldsymbol{\xi})$ satisfies Assumption 3.2.6 then, given the set $\{\boldsymbol{\xi}_i\}_{i=1}^M$, one can solve in polynomial time the moment problem*

$$\underset{F \,\in\, \mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \bar{\gamma}_1, \bar{\gamma}_2)}{\text{maximize}} \quad \mathbb{E}_F[g(\boldsymbol{\xi})] \ \ ,$$

where $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$, $\bar{\gamma}_1$ and $\bar{\gamma}_2$ are assigned as in Definition 3.4.1. Furthermore, if M satisfies Constraint (2.8)*, then with probability greater than $1 - \delta$ over the choice of $\{\boldsymbol{\xi}_i\}_{i=1}^M$, we have that*

$$\mathbb{E}\left[g(\boldsymbol{\xi})\right] \ \leq \ \Psi(0; \bar{\gamma}_1, \bar{\gamma}_2) \ \ ,$$

where $\mathbb{E}\left[\cdot\right]$ is the expectation with respect to the true distribution of $\boldsymbol{\xi}$.

3.5 Lessons from a Stochastic Linear Program

We consider the special case of a two-stage stochastic linear program:

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \mathbb{E}\left[h_{\text{SLP}}(\boldsymbol{x}, \boldsymbol{\xi})\right] \\ \text{subject to} & \boldsymbol{A}_1 \boldsymbol{x} \leq \boldsymbol{b}_1 \ \ , \end{array}$$

where $\boldsymbol{\xi}$ is a random vector of parameters and $h(\boldsymbol{x},\boldsymbol{\xi})$ is the optimal value of a second stage linear optimization problem of the form

$$\begin{array}{rll} h_{\mathrm{SLP}}(\boldsymbol{x},\boldsymbol{\xi}) := & \underset{\boldsymbol{y}}{\text{minimize}} & \boldsymbol{c}_1^{\mathsf{T}}\boldsymbol{x} + \boldsymbol{\xi}^{\mathsf{T}}\boldsymbol{y} \\ & \text{subject to} & \boldsymbol{A}_2\boldsymbol{x} + \boldsymbol{B}_2\boldsymbol{y} \leq \boldsymbol{b}_2 \end{array} .$$

Here, $\boldsymbol{x} \in \mathbb{R}^{n_1}$ represents a set of variables which must be chosen prior to gathering the knowledge about $\boldsymbol{\xi}$. At a later time, once $\boldsymbol{\xi}$ is observed, some recourse actions $\boldsymbol{y}$ can be taken in order to minimize the actual incurred cost. This model is said to only have cost uncertainty since the feasible set for $\boldsymbol{y} \in \mathbb{R}^{n_2}$ is fully determined given $\boldsymbol{x}$ and does not depend on the particular instance of $\boldsymbol{\xi}$.

We can easily verify that in this problem $h_{\mathrm{SLP}}(\boldsymbol{x},\boldsymbol{\xi})$ is convex in $\boldsymbol{x}$, since it is the optimal value of a minimization problem (over $\boldsymbol{y}$) that is convex jointly in $\boldsymbol{x}$ and $\boldsymbol{y}$. It is also concave in $\boldsymbol{\xi}$, since it is the minimum of a set of linear functions of $\boldsymbol{\xi}$. Based on Proposition 3.3.3, we know that the distributionally robust version of this problem can be solved efficiently.

Corollary 3.5.1 : *The distributionally robust stochastic linear program with uncertainty in the second stage cost and with distributional set $\mathcal{D}_1$:*

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \max_{F \in \mathcal{D}_1(\mathcal{S},\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0,\gamma_1,\gamma_2)} \mathbb{E}_F[h_{SLP}(\boldsymbol{x},\boldsymbol{\xi})] \tag{3.9a}$$

$$\text{subject to} \quad \boldsymbol{A}_1\boldsymbol{x} \leq \boldsymbol{b}_1 \tag{3.9b}$$

can be solved in polynomial time.

Because we are now informed that this problem is a tractable one, it becomes interesting to look for solution methods that can exploit the structure in order to reduce computational requirements. The following proposition states that Problem (3.9) can be reduced to a second-order cone program (see [36]).

Proposition 3.5.2 : *Let $\mathcal{E}_{\boldsymbol{\mu}}$ be the ellipsoid $\mathcal{E}_{\boldsymbol{\mu}} = \{\boldsymbol{\mu} \in \mathbb{R}^{n_2} | (\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathsf{T}}\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu}-\boldsymbol{\mu}_0) \leq \gamma_1\}$*

and satisfy $\mathcal{E}_{\boldsymbol{\mu}} \subseteq \mathcal{S}$*. Then Problem* (3.9) *is equivalent to the second-order cone program*

$$\begin{array}{ll} \underset{\boldsymbol{x},\boldsymbol{y}}{\text{minimize}} & \boldsymbol{c}_1^{\mathsf{T}}\boldsymbol{x} + \hat{\boldsymbol{\mu}}^{\mathsf{T}}\boldsymbol{y} + \sqrt{\gamma_1}\|\hat{\boldsymbol{\Sigma}}^{1/2}\boldsymbol{y}\| \\ \text{subject to} & \boldsymbol{A}_1\boldsymbol{x} \leq \boldsymbol{b}_1 \\ & \boldsymbol{A}_2\boldsymbol{x} + \boldsymbol{B}_2\boldsymbol{y} \leq \boldsymbol{b}_2 \end{array} .$$

Proof: This proposition is a consequence of the fact that $h_{\text{SLP}}(\boldsymbol{x},\boldsymbol{\xi})$ is a concave function of $\boldsymbol{\xi}$. Thus, by Jensen's inequality we know that for any distribution F, it is the case that $\mathbb{E}_F[h_{\text{SLP}}(\boldsymbol{x},\boldsymbol{\xi})] \leq h_{\text{SLP}}(\boldsymbol{x},\mathbb{E}_F[\boldsymbol{\xi}])$, and equality is achieved if the distribution of $\boldsymbol{\xi}$ is the Dirac measure $\delta_{\mathbb{E}_F[\boldsymbol{\xi}]}$. If we can show that for all $\boldsymbol{\mu} \in \mathcal{E}_{\boldsymbol{\mu}}$ the Dirac measure $\delta_{\boldsymbol{\mu}}$ is necessarily in $\mathcal{D}_1(\mathcal{S},\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0,\gamma_1,\gamma_2)$, then we have demonstrated that Problem (3.9) is equivalent to:

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \underset{\boldsymbol{\mu}\in\mathcal{E}_{\boldsymbol{\mu}}}{\max} \left\{ \underset{\boldsymbol{y}\in\{\boldsymbol{y}|\boldsymbol{A}_2\boldsymbol{x}+\boldsymbol{B}_2\boldsymbol{y}\leq\boldsymbol{b}_2\}}{\min} \boldsymbol{c}_1^{\mathsf{T}}\boldsymbol{x} + \boldsymbol{\mu}^{\mathsf{T}}\boldsymbol{y} \right\} \\ \text{subject to} & \boldsymbol{A}_1\boldsymbol{x} \leq \boldsymbol{b}_1 \end{array} .$$

Since $\mathcal{E}_{\boldsymbol{\mu}}$ is compact and convex, by Sions minimax theorem [58], inverting the order of $\max_{\boldsymbol{\mu}}$ and $\min_{\boldsymbol{y}}$ leads to an equivalent problem:

$$\begin{array}{ll} \underset{\boldsymbol{x},\boldsymbol{y}}{\text{minimize}} & \underset{\boldsymbol{\mu}\in\mathcal{E}_{\boldsymbol{\mu}}}{\max} \left\{ \boldsymbol{c}_1^{\mathsf{T}}\boldsymbol{x} + \boldsymbol{\mu}^{\mathsf{T}}\boldsymbol{y} \right\} \\ \text{subject to} & \boldsymbol{A}_1\boldsymbol{x} \leq \boldsymbol{b}_1 \\ & \boldsymbol{A}_2\boldsymbol{x} + \boldsymbol{B}_2\boldsymbol{y} \leq \boldsymbol{b}_2 \end{array} .$$

Finally, the reduction of a robust linear objective with ellipsoidal uncertainty to a second-order cone objective is a well known one in the field of robust optimization (see [2]).

It remains to show that for all $\boldsymbol{\mu} \in \mathcal{E}_{\boldsymbol{\mu}}$, it is true that $\delta_{\boldsymbol{\mu}} \in \mathcal{D}_1(\mathcal{S},\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}},\gamma_1,\gamma_2)$. By our assumption that $\mathcal{E}_{\boldsymbol{\mu}} \subseteq \mathcal{S}$, $\delta_{\boldsymbol{\mu}}$ necessarily satisfies the support and mean constraints imposed on all distributions in $\mathcal{D}_1$. Assuming without loss of generality that $0 < \gamma_1 \leq \gamma_2$, the second moment condition can be verified using the fact that $\mathbb{E}_{\delta_{\mu}}[\boldsymbol{\xi}] = \boldsymbol{\mu}$. Specifically,

$$\begin{aligned}
(\mathbb{E}_{\delta_{\mu}}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0)^{\mathsf{T}} \boldsymbol{\Sigma}_0^{-1} (\mathbb{E}_{\delta_{\mu}}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0) \leq \gamma_1 \quad &\leftrightarrow \quad \begin{bmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\mu} - \boldsymbol{\mu}_0 \\ \boldsymbol{\mu}^{\mathsf{T}} - \boldsymbol{\mu}_0^{\mathsf{T}} & \gamma_1 \end{bmatrix} \succeq 0 \\
&\leftrightarrow \quad (\boldsymbol{\mu} - \boldsymbol{\mu}_0)(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^{\mathsf{T}} \preceq \gamma_1 \boldsymbol{\Sigma}_0 \\
&\rightarrow \quad \mathbb{E}\left[(\delta_{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)(\delta_{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)^{\mathsf{T}}\right] \preceq \gamma_2 \boldsymbol{\Sigma}_0 \ .
\end{aligned}$$

This completes our proof. □

In proving Proposition 3.5.2, we used the fact that the distribution that puts all of its weight on a single point is a worst case distribution for Problem (3.9). This means that if the mean of the distribution is known exactly, our robust approach reduces to the well known deterministic counter part of the stochastic program,

$$\begin{array}{ll}
\underset{\boldsymbol{x},\boldsymbol{y}}{\text{minimize}} & \boldsymbol{c}_1^{\mathsf{T}} \boldsymbol{x} + \mathbb{E}[\boldsymbol{\xi}]^{\mathsf{T}} \boldsymbol{y} \\
\text{subject to} & \boldsymbol{A}_1 \boldsymbol{x} \leq \boldsymbol{b}_1 \\
& \boldsymbol{A}_2 \boldsymbol{x} + \boldsymbol{B}_2 \boldsymbol{y} \leq \boldsymbol{b}_2 \ .
\end{array}$$

This is interesting since, in practice, this deterministic counter part is often used as a heuristic to provide a quick solution to a stochastic program. Proposition 3.5.2 confirms that this heuristic is well-justified when one is more confident in the mean of the random vector of uncertain parameters than he is about the form of its distribution.

It is important to mention that the phenomenon that was just described is not simply caused by the fact that $\mathcal{D}_1$ imposes no lower bound on the covariance matrix. In fact, in the case that the support is unbounded and that the mean and covariance matrix are exactly known, one can show that solving the DRSP with distributional set $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 0, 1)$ leads to the same solution as the DRSP which imposes equality constraints.

Proposition 3.5.3 : *Given that for any $\boldsymbol{x} \in \mathcal{X}$, the set $\mathcal{Y}(\boldsymbol{x}) = \{\boldsymbol{y} | \boldsymbol{A}_2 \boldsymbol{x} + \boldsymbol{B}_2 \boldsymbol{y} \leq \boldsymbol{b}_2\}$ is bounded by a ball centered at 0 of finite radius $R(\boldsymbol{x})$, Problem* (3.9) *with the distribution set $\mathcal{D}_1(\mathbb{R}^{n_2}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, 0, 1)$ is equivalent to Problem* (3.9) *with the distributional set*

$$\bar{\mathcal{D}}(\mathbb{R}^{n_2}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = \left\{ F \in \mathcal{M} \middle| \mathbb{P}(\boldsymbol{\xi} \in \mathbb{R}^{n_2}) = 1, \, \mathbb{E}[\boldsymbol{\xi}] = \boldsymbol{\mu}_0, \, \mathbb{E}\left[(\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^{\mathsf{T}}\right] = \boldsymbol{\Sigma}_0 \right\}.$$

Proof: Using Lemma 3.2.2, we can state that the inner maximization problem

$$\max_{F \in \mathcal{D}_1(\mathbb{R}^{n_2}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, 0, 1)} \boldsymbol{c}_1^\mathsf{T} \boldsymbol{x} + \mathbb{E}_F[h_{\text{SLP}}(\boldsymbol{x}, \boldsymbol{\xi})] \tag{3.10}$$

is equivalent to the solution of its dual problem

$$\begin{aligned}
&\underset{\boldsymbol{Q}, \boldsymbol{q}, r, t}{\text{minimize}} && r + t && (3.11a)\\
&\text{subject to} && r \geq h_{\text{SLP}}(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{\xi}^\mathsf{T} \boldsymbol{Q} \boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T} \boldsymbol{q} \;\; \forall\, \boldsymbol{\xi} \in \mathbb{R}^{n_2} && (3.11b)\\
& && t \geq \left(\boldsymbol{\Sigma}_0 + \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\mathsf{T}\right) \bullet \boldsymbol{Q} + \boldsymbol{\mu}_0^\mathsf{T} \boldsymbol{q} && (3.11c)\\
& && \boldsymbol{Q} \succeq 0 \;\;, && (3.11d)
\end{aligned}$$

where $(\boldsymbol{A} \bullet \boldsymbol{B})$ refers to the Frobenius inner product between matrices, $\boldsymbol{Q} \in \mathbb{R}^{n_2 \times n_2}$ is a symmetric matrix, the vector $\boldsymbol{q} \in \mathbb{R}^{n_2}$ and $r, t \in \mathbb{R}$. We now show that, because $\mathcal{Y}(\boldsymbol{x})$ is a bounded set for all $\boldsymbol{x} \in \mathcal{X}$, Constraint (3.11d) is actually redundant. In fact, any real valued assignment for $\boldsymbol{Q}, \boldsymbol{q}, r, t$ that is feasible with respect to Constraint (3.11b) necessarily has that $\boldsymbol{Q} \succeq 0$. Let's assume that $\boldsymbol{Q} \not\succeq 0$, then there exists a vector $\boldsymbol{z} \in \mathbb{R}^{n_2}$ such that $\|\boldsymbol{z}\| = 1$ and $\boldsymbol{z}^\mathsf{T} \boldsymbol{Q} \boldsymbol{z} = -\delta$ for some $\delta > 0$. Yet,

$$\begin{aligned}
r \;&\geq\; \min_{\boldsymbol{y} \in \mathcal{Y}(\boldsymbol{x})} \boldsymbol{y}^\mathsf{T} \boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T} \boldsymbol{Q} \boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T} \boldsymbol{q} \;\; \forall\, \boldsymbol{\xi} \in \mathbb{R}^{n_2}\\
&\geq\; \max_{\alpha > 0} \left(\min_{\boldsymbol{y} \in \mathcal{Y}(\boldsymbol{x})} \boldsymbol{y}^\mathsf{T} \alpha \boldsymbol{z} - (\alpha \boldsymbol{z})^\mathsf{T} \boldsymbol{Q} (\alpha \boldsymbol{z}) - (\alpha \boldsymbol{z})^\mathsf{T} \boldsymbol{q} \right)\\
&\geq\; \max_{\alpha > 0} \left(-\alpha R(\boldsymbol{x}) + \alpha^2 \delta - \alpha \|\boldsymbol{q}\| \right)\\
&=\; \infty \;\;,
\end{aligned}$$

But since r is real valued, we have a contradiction. Thus, feasibility with respect to Constraint (3.11b) implies feasibility with respect to Constraint (3.11d) thus Constraint (3.11d) is redundant in Problem (3.11).

Applying duality theory to Problem (3.11) with Constraint (3.11d) removed, we get the problem

$$\begin{aligned} \underset{F}{\text{maximize}} \quad & \mathbb{E}_F[h_{\text{SLP}}(\boldsymbol{x}, \boldsymbol{\xi})] \\ \text{subject to} \quad & \mathbb{E}_F[\boldsymbol{\xi}] = \boldsymbol{\mu}_0 \\ & \mathbb{E}_F[(\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\mathsf{T}] = \boldsymbol{\Sigma}_0 \ , \end{aligned}$$

which must achieve the same optimal value as Problem (3.10). Since the problem above is the inner maximization problem considered in Problem (3.9) with the distributional set $\bar{\mathcal{D}}(\mathbb{R}^{n_2}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, we can therefore conclude that the two problems are equivalent. □

This last theorem indicates that, as we expected in Remark 3.1.1, in many problems the lower bound on the covariance matrix is redundant. On the other hand, it is actually the case that imposing such a constraint can cause the optimization model to become intractable as is shown in the next proposition. This is the main reason why we consider $\mathcal{D}_1$ to be a more practical distributional set to use. Based on the cases that we study in this work, using $\mathcal{D}_1$ instead of $\bar{\mathcal{D}}$ either leads to the same solution or gives a solution which is slightly more robust than needed to a problem that is actually NP-hard to solve.

Proposition 3.5.4: *The DRSP with $\bar{\mathcal{D}}(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ is NP-hard in general.*

Proof: We will show that the special case where $\mathcal{S} = \mathbb{R}^{n_2+} = \{\boldsymbol{\xi} \in \mathbb{R}^{n_2} | \boldsymbol{\xi} \geq 0\}$ is already hard for the form

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \max_{F \in \bar{\mathcal{D}}(\mathbb{R}^{n_2+}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)} \mathbb{E}_F[h_{\text{SLP}}(\boldsymbol{x}, \boldsymbol{\xi})] & \text{(3.12a)} \\ \text{subject to} \quad & \boldsymbol{x} \geq 0 \ , \ \ \mathbf{1}_{n_1}^\mathsf{T} \boldsymbol{x} \leq 2 \ , & \text{(3.12b)} \end{aligned}$$

where $\mathbf{1}_{n_1}$ is the vector of all ones in $\mathbb{R}^{n_1}$, and $h_{\text{SLP}}(\boldsymbol{x}, \boldsymbol{\xi})$ is the optimal value of the second stage linear program

$$\begin{aligned} h_{\text{SLP}}(\boldsymbol{x}, \boldsymbol{\xi}) := \quad \underset{\boldsymbol{y}}{\text{minimize}} \quad & \boldsymbol{\xi}^\mathsf{T} \boldsymbol{y} \\ \text{subject to} \quad & \sum_{i=1}^{n_2} y_i \geq -1, \ , \ -\boldsymbol{x} \leq \boldsymbol{y} \leq 0 \ . \end{aligned}$$

We showed in the proof of Theorem 3.5.3 that this problem is equivalent to Problem (3.11) with Constraint (3.11d) removed. Following similar lines that were used in the proof of Proposition 4.2 in [6], we reduce the NP-hard problem of checking if a given matrix is copositive to the separation problem associated with Constraint (3.11b). By the equivalence of optimization and separation (see [29]) solving Problem (3.12) is therefore NP-hard.

Problem COPOS: Given a matrix $\boldsymbol{H} \in \mathbb{R}^{n\times n}$ with rational entries, does there exists a $\boldsymbol{z} \geq 0$ such that $\boldsymbol{z}^\mathsf{T}\boldsymbol{H}\boldsymbol{z} < 0$?

Consider verifying feasibility of Constraint (3.11b) for an assignment

$$\boldsymbol{Q} = \begin{bmatrix} 0 & 0 & \mathbf{0}_n^\mathsf{T} \\ 0 & 0 & \mathbf{0}_n^\mathsf{T} \\ \mathbf{0}_n & \mathbf{0}_n & \boldsymbol{H} \end{bmatrix} \;, \qquad \boldsymbol{q} = \mathbf{0}_{n+2} \;, \qquad r = 0 \;, \qquad \boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \\ \mathbf{0}_n \end{bmatrix} \;,$$

where $\mathbf{0}_n$ is the vector of all zeros in $\mathbb{R}^n$. In this case, Constraint (3.11b) can be reformulated as

$$\begin{aligned} 0 \geq \max_{\boldsymbol{\xi}\geq 0} \left\{ h_{\mathrm{SLP}}(\boldsymbol{x},\boldsymbol{\xi}) - \boldsymbol{\xi}_{3:n+2}^\mathsf{T}\boldsymbol{H}\boldsymbol{\xi}_{3:n+2} \right\} & \\ \leftrightarrow \;\; 0 \geq \max_{\boldsymbol{\xi}\geq 0} \left\{ \min\{-\xi_1, -\xi_2) - \boldsymbol{\xi}_{3:n+2}^\mathsf{T}\boldsymbol{H}\boldsymbol{\xi}_{3:n+2} \right\} & \\ \leftrightarrow \;\; 0 \geq \; -\boldsymbol{\xi}_{3:n+2}^\mathsf{T}\boldsymbol{H}\boldsymbol{\xi}_{3:n+2} \;\;, \; \forall\, \boldsymbol{\xi}_{3:n+2} \geq 0 \;\;, & \end{aligned}$$

where $\boldsymbol{\xi}_{3:n+2}$ is short-hand notation for $[\xi_3, \xi_4, ..., \xi_{n+2}]^\mathsf{T}$. Since verifying the last statement is equivalent to showing that $\boldsymbol{H}$ is copositive, Problem (3.9) with $\bar{\mathcal{D}}(\mathbb{R}^{n_2+}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ is necessarily NP-hard. □

Although this last proof is similar in nature to the proof presented by Bertsimas et al. [6], here it addresses the problem of finding a distributionally robust solution instead of a tight bound for a probability inequality. It actually confirms more rigorously the intuition that Problem (3.9) is at least as hard as the optimal inequality problem.

3.6 Implications for Expected Utility Maximization

When dealing with uncertainty, different decision makers naturally exhibit different attitudes toward the risks related to the performance of their decision. This is the case even when the probability distribution of the random variables involved is well known. For this reason, it is common to use an expected utility measure as the objective of a stochastic program. Specifically, the decision maker can formulate

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{maximize}} & \mathbb{E}\left[u(-h(\boldsymbol{x},\boldsymbol{\xi}))\right] \\ \text{subject to} & \boldsymbol{x} \in \mathcal{X} \quad , \end{array}$$

where $u(\cdot)$ is a non-decreasing concave function that encodes his relative valuation of the different reachable levels of his revenue, and $h(\boldsymbol{x},\boldsymbol{\xi})$ is a cost function, or equivalently $-h(\boldsymbol{x},\boldsymbol{\xi})$ is a revenue function. This form implements risk averseness; for instance, losing some level α in performance reduces utility more than is gained by an equivalent increase of α in performance:

$$\begin{aligned} (u(y+\alpha)-u(y))+(u(y-\alpha)-u(y)) &= 2\left(\frac{1}{2}u(y+\alpha)+\frac{1}{2}u(y-\alpha)\right)-2u(y) \\ &\leq 2u(y)-2u(y) \;=\; 0 \;. \end{aligned}$$

In this section, we briefly discuss how the distributionally robust framework can be applied to an expected utility maximization problem with piecewise linear concave utility of the form $u(y) = \min_{k\in\{1,2,...,K\}} a_k y + b_k$ with $a_k \geq 0$ for all $k \in \{1, 2, ..., K\}$.

Because the expected utility framework relies on the knowledge of the true distribution of outcomes, the distributionally robust optimization model becomes attractive in contexts where one does not hold enough information about the uncertain parameters to commit to a choice of distribution. The case of a data-driven problem continues to be an excellent example of such a situation. For this reason, we take the time to mention implications of our results to this problem.

More specifically, a simple application of the theory developed in Section 3.4 leads to the following interesting corollary about the distributionally robust expected utility maximization problem.

Corollary 3.6.1 : *Let $\{\boldsymbol{\xi}_i\}_{i=1}^M$ be a set of M samples generated independently at random according to the distribution F whose support is contained in the set $\mathcal{S}$. For any $\delta > 0$, if assumptions 3.2.4, 3.2.6, 3.3.1, and 3.3.2 are satisfied then, given the set $\{\boldsymbol{\xi}_i\}_{i=1}^M$, one can solve in polynomial time the problem*

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{maximize}} \quad \min_{F\in\mathcal{D}_1(\mathcal{S},\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}},\bar{\gamma}_1,\bar{\gamma}_2)} \mathbb{E}_F[u(-h(\boldsymbol{x},\boldsymbol{\xi}))] \ ,$$

where $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$, $\bar{\gamma}_1$, and $\bar{\gamma}_2$ are assigned as in Definition 3.4.1. Furthermore, if M satisfies Constraint (2.8)*, then with probability greater than $1-\delta$ over the choice of $\{\boldsymbol{\xi}_i\}_{i=1}^M$, we have that any optimal solution $\boldsymbol{x}^*$ of the DRSP formed using these samples will satisfy the constraint*

$$\mathbb{E}\left[u(-h(\boldsymbol{x}^*,\boldsymbol{\xi}))\right] \geq \min_{F\in\mathcal{D}_1(\mathcal{S},\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}},\bar{\gamma}_1,\bar{\gamma}_2)} \mathbb{E}_F[u(-h(\boldsymbol{x}^*,\boldsymbol{\xi}))] \ ,$$

where $\mathbb{E}[\cdot]$ is the expectation with respect to the true distribution of $\boldsymbol{\xi}$.

Proof: After reformulating the expected utility maximization problem in its minimization form,

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \max_{F\in\mathcal{D}_1(\mathcal{S},\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}},\gamma_1,\gamma_2)} \mathbb{E}\left[-u(-h(\boldsymbol{x},\boldsymbol{\xi}))\right] \ ,$$

the validity of this corollary is confirmed by verifying that the future cost function $\bar{h}(\boldsymbol{x},\boldsymbol{\xi}) = -u(-h(\boldsymbol{x},\boldsymbol{\xi}))$ satisfies assumptions 3.2.6 and 3.3.2 given that $h(\boldsymbol{x},\boldsymbol{\xi})$ does. Assumption 3.2.6 is satisfied since

$$\bar{h}(\boldsymbol{x},\boldsymbol{\xi}) \;=\; \max_{k\in\{1,2,\ldots,K\}} a_k h(\boldsymbol{x},\boldsymbol{\xi}) - b_k \;=\; \max_{k\in\{1,2,\ldots,K\},l\in\{1,2,\ldots,L\}} a_k h_l(\boldsymbol{x},\boldsymbol{\xi}) - b_k \ ,$$

with each piece $a_k h_l(\boldsymbol{x},\boldsymbol{\xi}) - b_k$ of $\bar{h}(\boldsymbol{x},\boldsymbol{\xi})$ being concave in $\boldsymbol{\xi}$ since $a_k \geq 0$. A super-gradient for each of them can also be generated efficiently using a super-gradient for $h_l(\boldsymbol{x},\boldsymbol{\xi})$. Similarly for Assumption 3.3.2, we can confirm that, being the maximum of convex functions, $\bar{h}(\boldsymbol{x},\boldsymbol{\xi})$ is convex in $\boldsymbol{x}$ and that a sub-gradient in $\boldsymbol{x}$ can be generated by scaling the sub-gradient of $h(\boldsymbol{x}, \boldsymbol{\xi})$ by a_{k^*} where k^* is the index that achieves the maximum at $(\boldsymbol{x}, \boldsymbol{\xi})$. $\square$

This corollary reveals two interesting facts about the problem of expected utility maximization under distribution uncertainty. First, its distributionally robust form can often be solved in polynomial time using a cutting plane method. Second, the solution of this robust form has tangible guarantees in a data-driven problem given that the confidence regions developed in Chapter 2 are used to construct $\mathcal{D}_1$.

3.7 Implications for Conditional Value-at-Risk Minimization

Conditional value-at-risk, also called mean excess loss, was recently introduced in the mathematical finance community as a risk measure for decision making. It is closely related to the more common value-at-risk measure, which for a risk tolerance level of $\vartheta \in (0, 1)$ evaluates the lowest amount τ such that with probability $1 - \vartheta$, the loss does not exceed τ. Instead, CVaR evaluates the conditional expectation of loss above the value-at-risk. In order to keep the focus of our discussion on the topic of DRSP models, we refer the reader to [49] for technical details on this subject. CVaR has gained a lot of interest in the community because of its attractive computational properties. For instance, [49] demonstrated that one can evaluate the ϑ-CVaR$[c(\boldsymbol{x}, \boldsymbol{\xi})]$ of a cost function $c(\boldsymbol{x}, \boldsymbol{\xi})$, where the random vector $\boldsymbol{\xi}$ is distributed according to F, by solving a convex minimization problem:

$$\vartheta\text{-CVaR}_F[c(\boldsymbol{x}, \boldsymbol{\xi})] = \min_{\lambda \in \mathbb{R}} \ \lambda + \frac{1}{\vartheta} \mathbb{E}_F \left[(c(\boldsymbol{x}, \boldsymbol{\xi}) - \lambda)^+ \right] \ ,$$

where $(y)^+ = \max\{y \ , \ 0\}$ and the notation ϑ-CVaR$_F[\cdot]$ emphasizes the fact that the measure depends on the distribution of $\boldsymbol{\xi}$.

While CVaR is an interesting risk measure, it still requires the decision maker to commit to a distribution F. This is a step that can be difficult to take in practice; thus, justifying the introduction of a distributionally robust version of the criterion such as in [61] and in [66]. Using the results presented earlier, we can derive new conclusions for the general form of robust conditional value at risk. Given that the distribution is known to lie in a distributional set $\mathcal{D}_1$, let the Distributionally Robust ϑ-CVaR Problem be expressed as:

$$\text{(DR } \vartheta\text{-CVaR)} \qquad \underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \left(\max_{F\in\mathcal{D}_1} \vartheta\text{-CVaR}_F[c(\boldsymbol{x},\boldsymbol{\xi})] \right) \ .$$

By the equivalence statement presented above, this problem is equivalent to the form

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \left(\max_{F\in\mathcal{D}_1} \ \left(\min_{\lambda\in\mathbb{R}} \ \lambda + \frac{1}{\vartheta}\mathbb{E}_F\left[(c(\boldsymbol{x},\boldsymbol{\xi})-\lambda)^+\right] \right) \right) \ .$$

Given that $c(\boldsymbol{x},\boldsymbol{\xi})$ meets the conditions of Assumption 3.2.6 and Assumption 3.3.2, one can show that the minimax theorem holds, for $\max_{F\in\mathcal{D}_1}\min_{\lambda\in\mathbb{R}}$, since the function $\lambda + \frac{1}{\vartheta}\mathbb{E}_F\left[(c(\boldsymbol{x},\boldsymbol{\xi})-\lambda)^+\right]$ is real valued, convex in λ and concave (actually linear) in F, and since $\mathcal{D}_1$ is weakly compact (see [53]). Thus, interchanging the $\max_F$ and $\min_\lambda$ operators leads to an equivalent formulation of the DR ϑ-CVaR Problem:

$$\underset{\boldsymbol{x}\in\mathcal{X},\lambda\in\mathbb{R}}{\text{minimize}} \quad \left(\max_{F\in\mathcal{D}_1} \ \mathbb{E}_F[h(\boldsymbol{x},\lambda,\boldsymbol{\xi})] \right) \ ,$$

where $h(\boldsymbol{x},\lambda,\boldsymbol{\xi}) = \lambda + \frac{1}{\vartheta}(c(\boldsymbol{x},\boldsymbol{\xi})-\lambda)^+$. Since

$$\begin{aligned} h(\boldsymbol{x},\lambda,\boldsymbol{\xi}) &= \lambda + \frac{1}{\vartheta}\max\{\ 0\ ,\ c(\boldsymbol{x},\boldsymbol{\xi})-\lambda\ \} \\ &= \max\left\{\ \lambda\ ,\ \max_k \left(1-\frac{1}{\vartheta}\right)\lambda + \frac{1}{\vartheta}c_k(\boldsymbol{x},\boldsymbol{\xi})\ \right\} \ , \end{aligned}$$

it is clear that $h(\boldsymbol{x},\lambda,\boldsymbol{\xi})$ meets Assumption 3.2.6 and Assumption 3.3.2. Hence, Proposition 3.3.3 allows us to conclude that finding an optimal $\boldsymbol{x}$ (and its associated λ) with respect to the worst case conditional value-at-risk over the set of distributions $\mathcal{D}_1$ can be done in polynomial time.

Remark 3.7.1 : In Theorem 1 of [5], the authors independently derived a similar result for robust expected utility maximization, robust conditional value at risk minimization, and an extension to a special family of risk measures known as optimized certainty equivalents. That work specifically addresses the more restrictive class of two-stage stochastic linear problems discussed in Section 3.5 and assumes that the support of the distribution is unbounded and that the first and second moments are exactly known. Their result improves on ours from a computational point of view since they reduce the problem to a semi-definite program. On the other hand, our result extends the range of applications for which the distributionally robust model can be used. This is achieved both in terms of richness in the formulation of the objective and of richness of the distribution information that can be accounted for, as will be shown next.

3.8 Beyond First and Second Moment Information

Solutions to distributionally robust optimization problem can be overly conservative when the distributional set that is used only accounts for support, and first and second moment information. We are about to present a class of distributional sets which can account for a wider range of moment information while ensuring that the resulting decision model is still tractable. The ultimate objective here is to allow the distributional set used by the DRSP to account for as much of the information that is available and consequently obtain better performing robust solutions.

In this section, we consider the family of distributional sets that can be represented in the form

$$\mathcal{D}_2(\mathcal{S}, \{\boldsymbol{\psi}_i, \boldsymbol{b}_i, \mathcal{K}_i\}_{i=1}^{L}) = \left\{ F \in \mathcal{M} \left| \begin{array}{l} \mathbb{P}(\boldsymbol{\xi} \in \mathcal{S}) = 1 \\ \boldsymbol{b}_i - \mathbb{E}\left[\boldsymbol{\psi}_i(\boldsymbol{\xi})\right] \in \mathcal{K}_i^* \ , \ \forall\, i \in \{1, ..., L\} \end{array} \right. \right\}$$

where $\mathcal{S} \subseteq \mathbb{R}^m$ is a closed convex set, $\mathcal{M}$ is the set of all probability measures on the measurable space $(\mathcal{S}, \mathcal{B})$, with $\mathcal{B}$ the Borel σ-algebra on $\mathcal{S}$, and, for each $i \in \{1, ..., L\}$, $\boldsymbol{\psi}_i : \mathbb{R}^m \to \mathbb{R}^{p_i}$ is a real valued mapping, $\boldsymbol{b}_i \in \mathbb{R}^{p_i}$, $\mathcal{K}_i$ is a closed convex cone, and $\mathcal{K}_i^*$ is its dual cone.[2] We make the following assumption about each pair $(\boldsymbol{\psi}_i(\boldsymbol{\xi}), \mathcal{K}_i)$.

Assumption 3.8.1 : For each pair $(\boldsymbol{\psi}_i(\boldsymbol{\xi}), \mathcal{K}_i)$ involved in the description of $\mathcal{D}_2$, the real valued function $g_i(\boldsymbol{z}_i, \boldsymbol{\xi}) = \boldsymbol{z}_i^\mathsf{T} \boldsymbol{\psi}_i(\boldsymbol{\xi})$ is convex in $\boldsymbol{\xi}$ for any $\boldsymbol{z}_i \in \mathcal{K}_i$. Furthermore, one can

[2] The dual of a cone is defined as $\mathcal{K}^* = \{\boldsymbol{z} \in \mathbb{R}^p | \boldsymbol{z}^\mathsf{T} \boldsymbol{y} \geq 0 \ \forall \boldsymbol{y} \in \mathcal{K}\}$.

in polynomial time :

1. evaluate the function $g_i(\boldsymbol{z}_i, \boldsymbol{\xi})$
2. find a sub-gradient of $g_i(\boldsymbol{z}_i, \boldsymbol{\xi})$ in $\boldsymbol{\xi}$
3. verify that $\boldsymbol{y} \in \mathcal{K}_i$ and if not generate a hyperplane that separates $\boldsymbol{y}$ from $\mathcal{K}_i$
4. verify that $\boldsymbol{z} \in \mathcal{K}_i^*$ and if not generate a hyperplane that separates $\boldsymbol{z}$ from $\mathcal{K}_i^*$.

We first confirm that $\mathcal{D}_2$ is well behaved in the sense that one can efficiently find a distribution in its interior.

Lemma 3.8.2: *Given that $\mathcal{D}_2(\mathcal{S}, \{\boldsymbol{\psi}_i, \boldsymbol{b}_i, \mathcal{K}_i\}_{i=1}^L)$ satisfies Assumption 3.8.1, one can find a distribution F_0 that is in the strict interior of $\mathcal{D}_2$ or conclude that $\mathcal{D}_2$ is empty in polynomial time in m.*

The proof of this lemma is deferred to Appendix A.3 in order to describe right away our main result. Interestingly enough, one can actually extend the computational results of Section 3.3 to a distributionally robust optimization problem that uses a set in this very general family of distributional sets.

Proposition 3.8.3 : *Given that assumptions 3.2.4, 3.2.6, 3.3.1, 3.3.2, and 3.8.1 are satisfied, then the DRSP with $\mathcal{D}_2(\{\boldsymbol{\psi}_i, \boldsymbol{b}_i, \mathcal{K}_i\}_{i=1}^L)$ can be solved in polynomial time.*

Proof: After verifying that $\mathcal{D}_2$ is non-empty with the help of Lemma 3.8.2, the proof of this proposition follows similar steps as were used in proving Proposition 3.3.3. We first derive the dual problem for the infinite dimensional problem,

$$\begin{array}{ll}\underset{F \in \mathcal{M}}{\text{maximize}} & \displaystyle\int_{\mathcal{S}} h(\boldsymbol{x}, \boldsymbol{\xi}) dF(\boldsymbol{\xi}) \\ \text{subject to} & \displaystyle\int_{\mathcal{S}} dF(\boldsymbol{\xi}) = 1 \\ & \boldsymbol{b}_i - \displaystyle\int_{\mathcal{S}} \boldsymbol{\psi}_i(\boldsymbol{\xi}) dF(\boldsymbol{\xi}) \in \mathcal{K}_i^* \ , \ \forall\, i \in \{1, ..., L\} \ ,\end{array}$$

which takes the form of a semi-infinite linear conic problem

$$\begin{aligned}
\underset{r,\boldsymbol{z}_1,\boldsymbol{z}_2,...,\boldsymbol{z}_L}{\text{minimize}} \quad & \sum_{i=1}^{L} \boldsymbol{b}_i^{\mathsf{T}}\boldsymbol{z}_i + r \\
\text{subject to} \quad & r \;\geq\; \max_{\boldsymbol{\xi}\in\mathcal{S}} \; h_k(\boldsymbol{x},\boldsymbol{\xi}) - \sum_{i=1}^{L} \boldsymbol{z}_i^{\mathsf{T}}\boldsymbol{\psi}_i(\boldsymbol{\xi}) \;,\; \forall\, k \in \{1,2,...,K\} \\
& \boldsymbol{z}_i \;\in\; \mathcal{K}_i \;,\; \forall\, i \in \{1,...,L\} \;\;,
\end{aligned}$$

where $r \in \mathbb{R}$ is the dual variable associated with the support constraint, and for each i the vector $\boldsymbol{z}_i \in \mathbb{R}^{p_i}$ is the dual vector associated with the ith moment constraint.

We can easily show that $h(\boldsymbol{x},\boldsymbol{\xi})$ and all functions $\boldsymbol{\psi}_i(\boldsymbol{\xi})$ are F-measurable for all $F \in \mathcal{M}$ since they are real valued functions and the support set $\mathcal{S}$ is bounded. Since there exists an F_0 in the interior of $\mathcal{D}_2$, the weaker version of Proposition 3.4 in [53] does apply and we can conclude that there is no duality gap between the primal problem and its dual, and that if the optimal value of the primal is finite then the set of optimal solutions to its dual must be non-empty. The next step is to make the replacement of the inner problem in the DRSP.

$$\begin{aligned}
\underset{\boldsymbol{x},r,\boldsymbol{z}_1,\boldsymbol{z}_2,...,\boldsymbol{z}_L}{\text{minimize}} \quad & \sum_{i=1}^{L} \boldsymbol{b}_i^{\mathsf{T}}\boldsymbol{z}_i + r \\
\text{subject to} \quad & r \;\geq\; h_k(\boldsymbol{x},\boldsymbol{\xi}) - \sum_{i=1}^{L} \boldsymbol{z}_i^{\mathsf{T}}\boldsymbol{\psi}_i(\boldsymbol{\xi}) \;\;,\; \forall\, \boldsymbol{\xi} \in \mathcal{S} \;\;,\; \forall\, k \in \{1,2,...,K\} \\
& \boldsymbol{z}_i \;\in\; \mathcal{K}_i \;,\; \forall\, i \in \{1,...,L\} \;\;.
\end{aligned}$$

Similarly as was shown in the proof of Proposition 3.3.3, the ellipsoid method can be applied to this optimization problem and is guaranteed to produce an ϵ-optimal solution in polynomial time. This again relies on verifying feasibility of a candidate solution in two steps. The first one consists of verifying the feasibility of all the $\boldsymbol{z}$'s with respect to their respective cone $\mathcal{K}_i$ by calling the oracles described in Assumption 3.8.1. Then, given that the $\boldsymbol{z}$'s are feasible, one can verify feasibility for the first set of constraints by solving K convex maximization problems of finite dimension using again the ellipsoid method. □

This last result allows a decision maker the opportunity to consider collecting a much wider range of distribution information for the problem that he studies while preserving the computational advantages of a model that accounts only for the mean and the covariance matrix. The following examples should provide valuable insights on the type of distribution information that can be encoded through this framework.

Example 3.8.4: Any twice differentiable convex function $\psi(\boldsymbol{\xi})$ with $\mathcal{K}_1 = \{y \in \mathbb{R} | y \geq 0\}$ leads to imposing an upper bound on $\mathbb{E}[\psi(\boldsymbol{\xi})] \leq b$. The pair $(\psi(\boldsymbol{\xi}), \mathcal{K}_1)$ satisfies Assumption 3.8.1 since $\nabla^2(y\,\psi(\boldsymbol{\xi})) = y\nabla^2\psi(\boldsymbol{\xi}) \geq 0 \;\forall\, y \in \mathcal{K}_1$. In particular, $\psi_1(\boldsymbol{\xi}) = (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^{\mathsf{T}}\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\xi} - \boldsymbol{\mu}_0)$ can be paired with $\mathcal{K}_1$ and any $b > 0$.

Example 3.8.5: Consider pairing $\boldsymbol{\psi}_2(\boldsymbol{\xi}) = \mathrm{vec}\left(\boldsymbol{\xi}\boldsymbol{\xi}^{\mathsf{T}}\right)$ with $\mathcal{K}_2 = \left\{\boldsymbol{y} \in \mathbb{R}^{m^2} \middle| \mathrm{mat}(\boldsymbol{y}) \succeq 0\right\}$, where $\mathrm{vec}(\boldsymbol{A}) : \mathbb{R}^{m \times m} \to \mathbb{R}^{m^2}$ and $\mathrm{mat}(\boldsymbol{x}) : \mathbb{R}^{m^2} \to \mathbb{R}^{m \times m}$ are the matrix to vector and vector to matrix conversion functions respectively. We can verify that this pair satisfies Assumption 3.8.1 since $\nabla^2(\boldsymbol{y}^{\mathsf{T}}\boldsymbol{\psi}(\boldsymbol{\xi})) = \mathrm{mat}(\boldsymbol{y}) \succeq 0 \;\forall\, \boldsymbol{y} \in \mathcal{K}_2$. Effectively, when combined to a positive semi-definite matrix $\boldsymbol{B}$, this pairing leads to a distributional set that imposes a second moment constraint: $\mathbb{E}[\boldsymbol{\xi}\boldsymbol{\xi}^{\mathsf{T}}] \preceq \boldsymbol{B}$.

These two examples demonstrate that the set $\mathcal{D}_1$ is a special case of this new general family of distributional set; specifically,

$$\mathcal{D}_1(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, 0, 1) = \mathcal{D}_2(\mathcal{S}, \{(\psi_1(\boldsymbol{\xi}), 0, \mathcal{K}_1), (\boldsymbol{\psi}_2(\boldsymbol{\xi}), \boldsymbol{\Sigma}_0 + \boldsymbol{\mu}_0\boldsymbol{\mu}_0^{\mathsf{T}}, \mathcal{K}_2)\}) \;\; .$$

We now show examples that extend the theory to bounding higher orders of $\boldsymbol{\xi}$.

Example 3.8.6 : For any parameter ξ_j, we can pair $\boldsymbol{\psi}(\boldsymbol{\xi}) = \begin{bmatrix} \xi_j^4/12 \\ \xi_j^3/3 \\ \xi_j^2/2 \end{bmatrix}$ with the cone $\mathcal{K} = \left\{\boldsymbol{y} \in \mathbb{R}^3 \middle| \begin{bmatrix} y_1 & y_2 \\ y_2 & y_3 \end{bmatrix} \succeq 0\right\}$. This pairing satisfies Assumption 3.8.1 since

$$\frac{\partial^2}{\partial \xi_j^2}\boldsymbol{y}^{\mathsf{T}}\boldsymbol{\psi}(\boldsymbol{\xi}) = y_1\xi_j^2 + 2y_2\xi_j + y_3 = \left(\begin{bmatrix} \xi_j \\ 1 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} y_1 & y_2 \\ y_2 & y_3 \end{bmatrix} \begin{bmatrix} \xi_j \\ 1 \end{bmatrix}\right) \geq 0 \;\forall\, \boldsymbol{y} \in \mathcal{K} \; .$$

Such a choice for $\boldsymbol{\psi}(\boldsymbol{\xi})$ and $\mathcal{K}$ leads to imposing a moment constraint of the form

$$\mathbb{E}\left[\begin{bmatrix} \xi_j^4/12 & \xi_j^3/6 \\ \xi_j^3/6 & \xi_j^2/2 \end{bmatrix}\right] \preceq \boldsymbol{B} \ .$$

Note that this moment constraint actually carries a lot of information about the random variable ξ_j. For instance, if the mean is also known to lie in $[a_j, b_j]$, then for any fixed $\bar{\xi} \in \mathbb{R}^m$ it implies that

$$\begin{aligned} \mathbb{E}\left[(\xi_j - \bar{\xi})^4\right] &= \mathbb{E}\left[\xi_j^4 - 4\bar{\xi}\xi_j^3 + 6\bar{\xi}^2\xi_j^2 - 4\bar{\xi}^3\xi_j + \bar{\xi}^4\right] \\ &= \begin{bmatrix} 1 \\ -\bar{\xi} \end{bmatrix}^{\mathsf{T}} \mathbb{E}\left[\begin{bmatrix} \xi_j^4 & 2\xi_j^3 \\ 2\xi_j^3 & 6\xi_j^2 \end{bmatrix}\right] \begin{bmatrix} 1 \\ -\bar{\xi} \end{bmatrix} - 4\bar{\xi}^3\mathbb{E}\left[\xi_j\right] + \bar{\xi}^4 \\ &\leq 12 \begin{bmatrix} 1 \\ -\bar{\xi} \end{bmatrix}^{\mathsf{T}} \boldsymbol{B} \begin{bmatrix} 1 \\ -\bar{\xi} \end{bmatrix} - 4a_j\bar{\xi}^3 + \bar{\xi}^4 \end{aligned}$$

Example 3.8.7 : For any parameter ξ_j, we can impose a constraint on its first to $2K$th moments as follows. We pair $\boldsymbol{\psi} : \mathcal{S} \rightarrow \mathbb{R}^{2k-1}$ such that $\psi_i(\boldsymbol{\xi}) = \xi_j^{i+1}$ with

$$\mathcal{K} = \left\{ \boldsymbol{y} \in \mathbb{R}^{2K-1} \middle| \exists \boldsymbol{X} \in \mathbb{R}^{K\times K}, \boldsymbol{X} \succeq 0 \ , \ y_k = \sum_{i,j: i+j=k+1} X_{i,j}/((k+1)k) \,\forall\, k \right\} \ .$$

Such a pairing satisfies Assumption 3.8.1 since

$$\frac{\partial^2}{\partial \xi_j^2}\boldsymbol{y}^{\mathsf{T}}\boldsymbol{\psi}(\boldsymbol{\xi}) = \sum_{k=2}^{2K} k(k-1)y_{k-1}\xi_j^{k-2} = \left(\begin{bmatrix} 1 \\ \xi_j \\ \xi_j^2 \\ \ldots \\ \xi_j^{K-1} \end{bmatrix}^{\mathsf{T}} \boldsymbol{X} \begin{bmatrix} 1 \\ \xi_j \\ \xi_j^2 \\ \ldots \\ \xi_j^{K-1} \end{bmatrix} \right) \geq 0 \ \forall\, \boldsymbol{y} \in \mathcal{K} \ .$$

Specifically, such a choice for $\boldsymbol{\psi}(\boldsymbol{\xi})$ and $\mathcal{K}$ leads to imposing a moment constraint of the form

$$\mathbb{E}\left[\begin{bmatrix} \boldsymbol{\xi}^2 \\ \boldsymbol{\xi}^3 \\ \ldots \\ \boldsymbol{\xi}^{2K} \end{bmatrix}\right] \preceq_{\mathcal{K}^*} \boldsymbol{B} \ ,$$

where $\mathcal{K}^*$ is the set of all $\boldsymbol{y} \in \mathbb{R}^{2K-1}$ such that

$$\begin{bmatrix} \frac{y_1}{2} & \frac{y_2}{6} & \frac{y_3}{12} & \ldots & \frac{y_K}{(K+1)K} \\ \frac{y_2}{6} & \frac{y_3}{12} & \frac{y_4}{20} & \ldots & \frac{y_{K+1}}{(K+2)(K+1)} \\ \frac{y_3}{12} & \frac{y_4}{20} & \frac{y_5}{30} & \ldots & \frac{y_{K+2}}{(K+3)(K+2)} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{y_K}{(K+1)K} & \frac{y_{K+1}}{(K+2)(K+1)} & \frac{y_{K+2}}{(K+3)(K+2)} & \ldots & \frac{y_{2K-1}}{2K(2K-1)} \end{bmatrix} \succeq 0 \ .$$

In general, given that one is interested in accounting for the moment of a mapping $\boldsymbol{\psi} : \mathbb{R}^m \to \mathbb{R}^p$, Assumption 3.8.1 should be interpreted as imposing that the uncertainty set for $\mathbb{E}[\boldsymbol{\psi}(\boldsymbol{\xi})]$ needs to be relaxed to the dual of a cone $\mathcal{K}$ that satisfies two properties: first, the cone $\mathcal{K}$ must have an efficient feasibility oracle, and second, the cone $\mathcal{K}$ must be a subset of the cone $\{\boldsymbol{z} \in \mathbb{R}^p | \boldsymbol{z}^\mathsf{T}\boldsymbol{\psi}(\boldsymbol{\xi}) \text{ is convex in } \boldsymbol{\xi}\}$. This was achieved in the last example for constraining the moments of each variable. We conjecture that one should be able to exploit further the theory of sum of squares to define such a cone in the more general case where the mapping involves a joint moment of order higher than two (see [48] for more information on this theory).

3.9 Conclusion

In this chapter, we described the computational difficulties related to solving the DRSP problem. We first showed that, when the distributional set takes the form described by $\mathcal{D}_1$, the problem can be solved in polynomial time for a large range of objective functions. In light of the new confidence regions presented in Chapter 2, we were able to propose a framework which generates solutions with strong probabilistic guarantees to problems where the knowledge of the parameters is limited to a set of historical samples drawn identically and independently from the same distribution. In our discussion about two-stage stochastic linear problems, we also confirmed that one could exploit the structure of the problem he is addressing to reduce the complexity of finding this robust solution. The analysis also seems to indicate that, when one truly seeks the solution that accounts for exact knowledge of the moments, using $\mathcal{D}_1$ with $\gamma_1 = 0$ and $\gamma_2 = 1$ either leads to the right solution if the true problem is in fact tractable, and otherwise finds in polynomial time a solution that is slightly more robust than needed. We then discussed the implications of our results for two commonly used risk-adjusted objectives that are centered around a subjective distribution and allowed these to be used in a context where one cannot commit to such a distribution. Our final theoretical contribution was to extend our distributional set to account for a much larger range of statistics. We believe this should allow practitioners to account more accurately for the distribution information that they have in hand when formulating a distributionally robust optimization problem.

Chapter 4

Applications of Distributionally Robust Optimization

Since the distributionally robust framework discussed in Chapter 3 only imposes weak conditions on the objective function of the decision problem, it is possible to revisit many interesting forms of stochastic programs and reformulate them taking into account distribution and moment uncertainty. In this chapter, we consider three specific problems that can benefit from the application of our method: a Chebyshev inequality type problem, a stochastic fleet mix optimization problem and a portfolio selection problem. In the case of portfolio selection, we will provide empirical evidence that our distributionally robust approach addresses more effectively the true uncertainty that is present in the daily return of stocks.

4.1 Optimal Inequalities in Probability Theory

Consider the problem of finding a tight upper bound on $\mathbb{P}(\boldsymbol{\xi} \in \mathcal{C})$ for a random vector $\boldsymbol{\xi}$ with known mean and covariance matrix, and some closed set $\mathcal{C}$. By formulating this problem as a semi-infinite linear program:

$$\underset{F \in \mathcal{D}}{\text{maximize}} \quad \int_{\mathcal{S}} \mathbb{1}\{\boldsymbol{\xi} \in \mathcal{C}\} dF(\boldsymbol{\xi}) \ , \tag{4.1}$$

many have proposed methods to provide extensions to the popular Chebyshev inequality (see [38] and [6]). However, these methods fail when dealing with support constraints. More specifically, if $\mathcal{C}$ is a finite union of disjoint convex sets, it is known that for Problem (4.1) with unconstrained support, $\mathcal{S} = \mathbb{R}^m$, the worst case value can be found in polynomial time. But, if the support is constrained, such as $\mathcal{S} = \mathbb{R}^+$, then the problem is known to potentially be NP-hard. In fact, the hardness of this problem arises already in finding a distribution that is feasible.

Our framework recommends to relax the restrictions on the covariance matrix of $\boldsymbol{\xi}$ and consider the distributional set $\mathcal{D}_1(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$. Such a distributional set constrains all three types of statistics: support, mean and second moment matrix. If $\mathcal{C}$ is a finite union of disjoint convex sets $\mathcal{C}_k$ (equipped with their respective feasibility oracle), and if for each k, $\mathcal{C}_k \cap \mathcal{S} \neq \emptyset$, then our framework leads to a new Chebyshev inequality that can be evaluated in polynomial time. First, in our framework the problem of finding a $F \in \mathcal{D}_1$ is already resolved using the Dirac measure $\delta_{\boldsymbol{\mu}_0}$. We can also construct an $h(\boldsymbol{x}, \boldsymbol{\xi})$ that satisfies Assumption 3.2.6 and Assumption 3.3.2 by letting $h_0(\boldsymbol{x}, \boldsymbol{\xi}) = 0$, and for all $k \neq 0$ letting $h_k(\boldsymbol{x}, \boldsymbol{\xi}) = 1$ if $\boldsymbol{\xi} \in \mathcal{C}_k$ and $h_k(\boldsymbol{x}, \boldsymbol{\xi}) = -\infty$ otherwise. Therefore, if the distribution of $\boldsymbol{\xi}$ is known to be a member of $\mathcal{D}_1$, then clearly

$$\mathbb{P}(\boldsymbol{\xi} \in \mathcal{C}) \;=\; \mathbb{E}\left[\mathbb{1}\{\boldsymbol{\xi} \in \mathcal{C}\}\right] \;=\; \mathbb{E}\left[\max_k \; h_k(\boldsymbol{x}, \boldsymbol{\xi})\right] \;=\; \mathbb{E}\left[h(\boldsymbol{x}, \boldsymbol{\xi})\right] \;\leq\; \max_{F \in \mathcal{D}_1} \; \mathbb{E}_F[h(\boldsymbol{x}, \boldsymbol{\xi})] \;,$$

and a tight Chebyshev bound can be found in polynomial time. Note that by using the form $\mathcal{D}_1(\mathbb{R}^+, \boldsymbol{\mu}, \boldsymbol{\Sigma}, 0, 1)$ one can also provide useful upper bounds to the mentioned NP-hard version of the problem which assumes a known covariance matrix.

4.2 Stochastic Fleet Mix Optimization Problems

As discussed in [35], a stochastic fleet mix optimization problem considers the decisions that an airline company must take when composing a fleet of aircraft for the service of some future flights. As these decisions are typically made ten to twenty years ahead of time, it involves a lot of uncertainty with respect to the profits that will be generated from any given portfolio of aircraft. It is therefore natural to formulate this problem as a stochastic mixed integer linear program with recourse actions. More specifically, the problem takes the following shape

$$\underset{\boldsymbol{x}\geq 0}{\text{maximize}} \quad \mathbb{E}\left[\rho(\boldsymbol{x},\tilde{\boldsymbol{p}},\tilde{\boldsymbol{c}},\tilde{\boldsymbol{l}})\right] \ ,$$

where $\boldsymbol{x} \in \mathbb{R}^{n_1}$ is a vector describing how many aircraft of each type is acquired by the airline, $\rho(\boldsymbol{x},\tilde{\boldsymbol{p}},\tilde{\boldsymbol{c}},\tilde{\boldsymbol{l}})$ is a function that computes the future weekly profits generated from the chosen fleet, and $\tilde{\boldsymbol{p}}$, $\tilde{\boldsymbol{c}}$, $\tilde{\boldsymbol{l}}$ are some random profit parameters which will be described shortly. In estimating future revenues and costs, the model assumes that the airline company allocates flights to aircraft in a way that is optimal with respect to actual expenses and revenues. This gives rise to a second stage optimization model of the form

$$\begin{aligned}
&\rho(\boldsymbol{x},\tilde{\boldsymbol{p}},\tilde{\boldsymbol{c}},\tilde{\boldsymbol{l}}) := \\
&\underset{\boldsymbol{z},\boldsymbol{w},\boldsymbol{y}}{\text{maximize}} \quad \sum_k \left(-o_k x_k + \sum_i \tilde{p}_i^k w_i^k - \tilde{c}_k (z_k - x_k)^+ + \tilde{l}_k (x_k - z_k)^+\right) \\
&\text{subject to} \quad \sum_k w_i^k = 1 \ , \ \forall i \\
&\qquad \sum_{g\in\text{in}(v)} y_g^k + \sum_{i\in\text{arr}(v)} w_i^k = \sum_{g\in\text{out}(v)} y_g^k + \sum_{i\in\text{dep}(v)} w_i^k \ , \ \forall k \ , \ \forall v \\
&\qquad z_k = \sum_{v\in\{v|\text{time}(v)=0\}} \left(y_{g\in\text{out}(v)}^k + \sum_{i\in\text{dep}(v)} w_i^k \right) \ , \ \forall k \\
&\qquad z_k \geq 0 \ , \ y_g^k \geq 0 \ , \ w_i^k \in \{0,1\} \ , \ \forall k \ , \ \forall g \ , \ \forall i \ .
\end{aligned}$$

where each variable $w_i^k = 1$ describes whether or not an aircraft of type k is assigned to flight i, the vector $\boldsymbol{z} \in \mathbb{R}^{n_1}$ counts how many aircraft of each type are actually used by the optimal flight assignment. The cost decomposes into the following sets of terms: the first set denotes ownership cost, the second set denotes profits made from each flight, the third set denotes rental costs for extra aircraft that are required by the assignment, and the last set denotes profits made from leasing out the part of the fleet which is unused. Specifically, o_k is the deterministic ownership cost for an aircraft of type k, $\tilde{p}_i^k$ is the profit generated by using aircraft of type k for flight i, $\tilde{c}_k$ is the cost of renting an aircraft of type k, and $\tilde{l}_k$ is the revenue per aircraft of type k leased out. The general form of this problem considers

$\tilde{\boldsymbol{p}}$, $\tilde{\boldsymbol{c}}$, and $\tilde{\boldsymbol{l}}$ to be uncertain at the time of composing the original mix since they depend on factors that can not be determined at the time of composing the fleet: demand for the flight, price of gas, ticket price, crew cost, etc. Note that these factors are assumed to be resolved when comes the time of making the allocation of aircraft.

Table 4.1: Example of flight schedule for a fleet mix optimization problem.

Flight ID	From	Departure Time	To	Arrival Time
#1	A	t_0	B	t_1
#2	B	t_1	A	t_2
#3	B	t_1	C	t_2
#4	A	t_2	B	t_3
#5	C	t_2	A	t_3

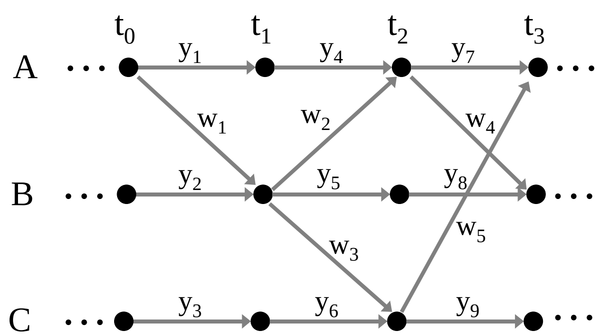


Figure 4.1: Flow graph associated with the flight schedule presented in Table 4.1.

The second stage aircraft allocation must satisfy some basic constraints. Typically, it is enforced that each flight is serviced by an aircraft. More importantly, one must ensure that in an allocation scheme the aircraft needed for a flight is actually present in the airport at departure time. This gives rise to some constraints reminiscent of flow constraints in network problems. The graph of flow constraints is derived from the schedule of flights. Some variables $\boldsymbol{y}$ are added to the problem formulation to allow an aircraft to stay in an airport garage between flights. Figure 4.1 presents the graph corresponding to the simple schedule presented in Table 4.1. The nodes of this graph are indexed by v. The sets of flight legs arriving to or departing from a node v are referred to as "arr(v)" and "dep(v)" respectively while the sets of ground legs incoming to and outgoing from an airport at time v are referred to as "in(v)" and "out(v)" respectively

The stochastic fleet mix optimization problem is an excellent example of a two-stage decision model where the uncertainty is limited to the second stage cost, while the second stage feasible set is deterministic. This is due to the assumption that the scheduling of flights has been completed prior to composing the original fleet mix. To the best of our knowledge, the following results consider for the first time the distributionally robust version of this problem. Note that for ease of notation we will consider $\tilde{\boldsymbol{p}}$ and $\boldsymbol{w}$ to be in vector form such that $\tilde{\boldsymbol{p}}^\mathsf{T}\boldsymbol{w} = \sum_{i,k} \tilde{p}_i^k w_i^k$. For clarity, we also construct the aggregated vector of uncertain parameters $\boldsymbol{\xi} = [\tilde{\boldsymbol{p}}^\mathsf{T}, \tilde{\boldsymbol{c}}^\mathsf{T}, \tilde{\boldsymbol{l}}^\mathsf{T}]^\mathsf{T}$ and assume that our uncertainty in the distribution of these parameters is represented using the distributional set $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \gamma_1, \gamma_2)$, for some $\hat{\boldsymbol{\mu}} = [\hat{\boldsymbol{p}}^\mathsf{T}, \hat{\boldsymbol{c}}^\mathsf{T}, \hat{\boldsymbol{l}}^\mathsf{T}]^\mathsf{T}$ which uses some estimates $\hat{\boldsymbol{p}}$, $\hat{\boldsymbol{c}}$, and $\hat{\boldsymbol{l}}$ of the mean of each parameters, and an estimate $\hat{\boldsymbol{\Sigma}}$ of the covariance matrix of $\boldsymbol{\xi}$. When $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are point estimates, the theory developed in Section 3.4 can be used to justify the size of γ_1 and γ_2. The distributional set $\mathcal{D}_2$ can also allow the decision maker to account more specifically for the information that he has about future outcomes. By using our framework, he gets the added benefit that the DRSP is often much simpler to solve than its stochastic version.

Proposition 4.2.1 : *The DRSP for fleet mix optimization problem with known mean*

$$\underset{\boldsymbol{x}\geq 0}{\text{maximize}} \quad \min_{F\in\mathcal{D}_1(\mathcal{S},\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}},0,\gamma_2)} \mathbb{E}_F[\rho(\boldsymbol{x}, \tilde{\boldsymbol{p}}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{l}})] \ ,$$

where F refers to the joint distribution of $\boldsymbol{\xi} = [\tilde{\boldsymbol{p}}^\mathsf{T}, \tilde{\boldsymbol{c}}^\mathsf{T}, \tilde{\boldsymbol{l}}^\mathsf{T}]^\mathsf{T}$ reduces to the deterministic mixed integer linear program

$$\begin{array}{ll}
\underset{\boldsymbol{x},\boldsymbol{w},\boldsymbol{y}}{\text{maximize}} & \displaystyle\sum_k -o_k x_k + \sum_{i,k} \hat{p}_i^k w_i^k \\
\text{subject to} & \displaystyle\sum_k w_i^k = 1 \ , \ \ \forall i \\
& \displaystyle\sum_{g\in in(v)} y_g^k + \sum_{i\in arr(v)} w_i^k = \sum_{g\in out(v)} y_g^k + \sum_{i\in dep(v)} w_i^k \ , \ \ \forall k \ , \ \ \forall v \\
& x_k = \displaystyle\sum_{v\in\{v|time(v)=0\}} \left(y^k_{g\in out(v)} + \sum_{i\in dep(v)} w_i^k \right) \ , \ \ \forall k \\
& x_k \geq 0 \ , \ \ y_g^k \geq 0 \ , \ \ w_i^k \in \{0,1\} \ , \ \ \forall k \ , \ \ \forall g \ , \ \ \forall i \ .
\end{array}$$

Proof: To show this result one simply needs to first realize that $\rho(\boldsymbol{x}, \tilde{\boldsymbol{p}}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{l}})$ is jointly convex in $\tilde{\boldsymbol{p}}$, $\tilde{\boldsymbol{c}}$, and $\tilde{\boldsymbol{l}}$ thus $h(\boldsymbol{x}, \tilde{\boldsymbol{p}}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{l}}) = -\rho(\boldsymbol{x}, \tilde{\boldsymbol{p}}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{l}})$ is jointly concave in those same parameters. One can then use the arguments presented in the proof of Proposition 3.5.2 to show that it is also the case here that

$$\min_{F \in \mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 0, \gamma_2)} \mathbb{E}_F[\rho(\boldsymbol{x}, \tilde{\boldsymbol{p}}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{l}})] \quad = \quad \rho(\boldsymbol{x}, \hat{\boldsymbol{p}}, \hat{\boldsymbol{c}}, \hat{\boldsymbol{l}}) \;\; .$$

Thus, we can join the two stages of maximization problems. Under the assumption that $\hat{l}_k \leq o_k \leq \hat{c}_k$ for all k, one can show that $\boldsymbol{z}^*$ is necessarily equal to $\boldsymbol{x}^*$ since there is no more need to take recourse actions. The problem therefore reduces to the form presented in the proposition. □

In practice, it is also common to consider a relaxation of the second stage problem which replaces the binary constraints on the terms of $\boldsymbol{w}$ by the constraint that they lie in the interval from zero to one:

$$\forall\, i, \forall\, k \;\; : \quad w_i^k \in \{0, 1\} \;\rightarrowtail\; w_i^k \in [0, 1] \;\; .$$

This leads to an approximation of the profit function, referred to as $\bar{\rho}(\boldsymbol{x}, \tilde{\boldsymbol{p}}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{l}})$, that is amenable to convex optimization solution techniques. Yet, minimizing $\mathbb{E}\left[\bar{\rho}(\cdot, \tilde{\boldsymbol{p}}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{l}})\right]$ remains computationally challenging because of the size of the second stage problem. A corollary of Proposition 3.5.2 describes how the computational burden can be reduced by considering distributional robustness.

Corollary 4.2.2 : *Let $\mathcal{E}_{\boldsymbol{\mu}}$ be the ellipsoid $\mathcal{E}_{\boldsymbol{\mu}} = \{\boldsymbol{\mu} \in \mathbb{R}^{n_2} | (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})^{\mathsf{T}} \hat{\boldsymbol{\Sigma}}^{-1} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}) \leq \gamma_1\}$ and satisfy $\mathcal{E}_{\boldsymbol{\mu}} \subseteq \mathcal{S}$. Then the distributionally robust problem*

$$\underset{\boldsymbol{x} \geq 0}{\text{maximize}} \quad \min_{F \in \mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \gamma_1, \gamma_2)} \mathbb{E}_F[\bar{\rho}(\boldsymbol{x}, \tilde{\boldsymbol{p}}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{l}})]$$

reduces to a second-order cone program.

It is also possible to encode risk averseness in the stochastic fleet mix optimization problem by solving an expected utility maximization problem. Based on Corollary 3.6.1, we know that if the utility function is piecewise linear, then the problem is necessarily tractable. Moreover, as mentioned in Remark 3.7.1, the authors of [5] showed in some work performed independently from ours how the distributionally robust form of a problem like the stochastic fleet mix optimization model with relaxed binary constraints and with an expected utility objective can be reduced to a semi-definite program. This connection between robust expected utility and semi-definite programming is actually directly related to the reductions that we present in the next section which addresses portfolio selection problems.

4.3 Portfolio Selection Problems

We now apply our framework to an instance of portfolio optimization. In such a problem, one is interested in maximizing his expected utility obtained from the single step return of his investment portfolio. Given that n investment options are available, expected utility can be defined as $\mathbb{E}\left[u(\boldsymbol{\xi}^\mathsf{T}\boldsymbol{x})\right]$, where $u(\cdot)$ is a non-decreasing utility function and $\boldsymbol{\xi} \in \mathbb{R}^n$ is a random vector of rates of return on investment for the different options. In the robust approach to this problem, one defines a distributional set $\mathcal{D}$ that is known to contain the distribution F and chooses the portfolio which is optimal according to the following Distributionally Robust Portfolio Optimization model:

$$\text{(DRPO)} \qquad \underset{\boldsymbol{x}}{\text{maximize}} \quad \min_{F \in \mathcal{D}} \ \mathbb{E}_F[u(\boldsymbol{\xi}^\mathsf{T}\boldsymbol{x})] \tag{4.2a}$$

$$\text{subject to} \quad \sum_{i=1}^{n} x_i = 1 \ , \ \ \boldsymbol{x} \geq 0 \ . \tag{4.2b}$$

In [45], the author addressed the case of Problem (4.2) where $\mathbb{E}\left[\boldsymbol{\xi}\right]$ and where $\mathbb{E}\left[\boldsymbol{\xi}\boldsymbol{\xi}^\mathsf{T}\right]$ are known exactly and one considers $\mathcal{D}$ to be the set of all distributions with such first and second moments. Based on these assumptions, the author presents a parametric quadratic programming algorithm that is efficient for a large family of utility functions $u(\cdot)$. This approach is interesting since it provides the means to take into account uncertainty in the form of the distribution of returns. However, our experiments show that in practice it is highly sensitive to the noise in the empirical estimation of these moments (see Section 4.4). The proposed algorithm also relies on solving a one-dimensional non-convex mathematical program; thus, there are no guarantees of finding a near optimal solution in polynomial time. Although the approach that we are about to propose addresses a smaller family of utility functions, it will take into account moment uncertainty and will lead to the formulation of a semi-definite program, which can be solved efficiently using interior point methods.

In [27], the authors attempt to account for moment uncertainty in Markowitz models. Their motivation is closely aligned with ours and many of the techniques that they propose can be applied in our context: *e.g.*, the use of factor models to reduce the dimensionality of $\boldsymbol{\xi}$. Similarly, the results presented in Chapter 2 for a data-driven framework should extend easily to the context of Markowitz models. Because Problem (4.2) reduces to a Markowitz model when the utility function is quadratic and concave, we consider our model to be richer than the one considered in [27]. On the other hand, a robust Markowitz model typically gives rise to a problem that is simpler to solve.

4.3.1 Portfolio Optimization with Moment Uncertainty

In order to apply our framework, we make the assumption that the utility function is piecewise linear and concave, such that $u(y) = \min_{k\in\{1,2,\ldots,K\}}\ a_k y + b_k$ with $a_k \geq 0$. This assumption is not very limiting since most interesting utility functions are concave and can usually be approximated accurately using simple piecewise linear functions. We use historical knowledge of investment returns $\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_M\}$ to define a distributional uncertainty set for F. This is done using the set $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \gamma_1, \gamma_2)$ where $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are assigned as the empirical estimates of the mean $\hat{\boldsymbol{\mu}} = M^{-1}\sum_{i=1}^{M}\boldsymbol{\xi}_i$ and covariance matrix $\hat{\boldsymbol{\Sigma}} = M^{-1}\sum_{i=1}^{M}(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})(\boldsymbol{\xi}_i - \hat{\boldsymbol{\mu}})^\mathsf{T}$ of $\boldsymbol{\xi}$ respectively.[1] We consider two options for the choice of $\mathcal{S}$: either $\mathcal{S} = \mathbb{R}^n$, or an "ellipsoidal" set $\mathcal{S} = \{\boldsymbol{\xi} | (\boldsymbol{\xi} - \boldsymbol{\xi}_0)^\mathsf{T}\boldsymbol{\Theta}(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \leq 1\}$, with $\boldsymbol{\xi}^\mathsf{T}\boldsymbol{\Theta}\boldsymbol{\xi} > 0$ for some $\boldsymbol{\xi} \in \mathbb{R}^n$.

Building on the results presented in Chapter 3, one can make the following statement about the tractability of the DRPO model.

Theorem 4.3.1 : *Given that $u(\cdot)$ is piecewise linear concave and that $\mathcal{X}$ satisfies Assumption 3.3.1, finding an optimal solution $\boldsymbol{x} \in \mathbb{R}^n$ to the DRPO model, Problem* (4.2)*, equipped*

[1] One should also verify that $\hat{\boldsymbol{\Sigma}} \succ 0$.

with the set of distributions $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \gamma_1, \gamma_2)$ *can be done in* $O(n^{6.5})$.

Proof: We first reformulate Problem (4.2) as a minimization problem :

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \left(\max_{F \,\in\, \mathcal{D}_1(\mathcal{S},\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}},\gamma_1,\gamma_2)} \mathbb{E}_F[\max_k \; -a_k \boldsymbol{\xi}^\mathsf{T}\boldsymbol{x} - b_k] \right) \; .$$

After confirming that $\mathcal{S}$, with $\boldsymbol{\Theta} \succeq 0$, satisfies the weaker version of Assumption 3.2.4 (see Remark 3.3.5) and that $h(\boldsymbol{x},\boldsymbol{\xi}) = \max_k \; -a_k\boldsymbol{\xi}^\mathsf{T}\boldsymbol{x} - b_k$ satisfies Assumption 3.2.6 and Assumption 3.3.2, a straightforward application of Proposition 3.3.3 shows that Problem (4.2) can be solved in polynomial time. In order to get a more precise computational bound, one needs to take a closer look at the dual formulation presented in Lemma 3.2.2 and its proof in Appendix A.1, and exploit the special structure of $h(\boldsymbol{x},\boldsymbol{\xi})$ in Problem (4.2):

$$\underset{\boldsymbol{x},\boldsymbol{Q},\boldsymbol{q},r,\boldsymbol{P},\boldsymbol{p},s}{\text{minimize}} \quad \gamma_2(\hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{Q}) - \hat{\boldsymbol{\mu}}^\mathsf{T}\boldsymbol{Q}\hat{\boldsymbol{\mu}} + r + (\hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{P}) - 2\hat{\boldsymbol{\mu}}^\mathsf{T}\boldsymbol{p} + \gamma_1 s \tag{4.3a}$$

$$\text{subject to} \quad \begin{bmatrix} \boldsymbol{P} & \boldsymbol{p} \\ \boldsymbol{p}^\mathsf{T} & s \end{bmatrix} \succeq 0 \;,\; \boldsymbol{p} = -\boldsymbol{q}/2 - \boldsymbol{Q}\hat{\boldsymbol{\mu}} \;,\; \boldsymbol{Q} \succeq 0 \tag{4.3b}$$

$$\boldsymbol{\xi}^\mathsf{T}\boldsymbol{Q}\boldsymbol{\xi} + \boldsymbol{\xi}^\mathsf{T}\boldsymbol{q} + r \geq -a_k\boldsymbol{\xi}^\mathsf{T}\boldsymbol{x} - b_k \;,\; \forall \boldsymbol{\xi} \in \mathcal{S}, k \in \{1,2,...,K\} \tag{4.3c}$$

$$\sum_{i=1}^n \boldsymbol{x}_i = 1 \;,\; \boldsymbol{x} \geq 0 \;. \tag{4.3d}$$

Given that $\mathcal{S} = \mathbb{R}^n$, one can use Schur's complement to replace Constraint (4.3c) by an equivalent linear matrix inequality:

$$\begin{aligned}
\underset{\boldsymbol{x},\boldsymbol{Q},\boldsymbol{q},r,\boldsymbol{P},\boldsymbol{p},s}{\text{minimize}} \quad & \gamma_2(\hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{Q}) - \hat{\boldsymbol{\mu}}^\mathsf{T}\boldsymbol{Q}\hat{\boldsymbol{\mu}} + r + (\hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{P}) - 2\hat{\boldsymbol{\mu}}^\mathsf{T}\boldsymbol{p} + \gamma_1 s \\
\text{subject to} \quad & \begin{bmatrix} \boldsymbol{P} & \boldsymbol{p} \\ \boldsymbol{p}^\mathsf{T} & s \end{bmatrix} \succeq 0 \;,\; \boldsymbol{p} = -\boldsymbol{q}/2 - \boldsymbol{Q}\hat{\boldsymbol{\mu}} \\
& \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{q}/2 + a_k\boldsymbol{x}/2 \\ \boldsymbol{q}^\mathsf{T}/2 + a_k\boldsymbol{x}^\mathsf{T}/2 & r + b_k \end{bmatrix} \succeq 0 \;,\; \forall k \\
& \sum_{i=1}^n x_i = 1 \;,\; \boldsymbol{x} \geq 0 \;.
\end{aligned}$$

On the other hand, if $\mathcal{S}$ is "ellipsoidal" and Θ has at least one positive eigenvalue, then the S-Lemma (*cf*., Theorem 2.2 in [44]) can be used for any given $k \in \{1, ..., K\}$ to replace Constraint (4.3c) and reformulate the problem as a semi-definite program:

$$\begin{array}{ll} \underset{\boldsymbol{x},\boldsymbol{Q},\boldsymbol{q},r,\boldsymbol{P},\boldsymbol{p},s,\tau}{\text{minimize}} & \gamma_2(\hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{Q}) - \hat{\boldsymbol{\mu}}^\mathsf{T}\boldsymbol{Q}\hat{\boldsymbol{\mu}} + r + (\hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{P}) - 2\hat{\boldsymbol{\mu}}^\mathsf{T}\boldsymbol{p} + \gamma_1 s \\ \text{subject to} & \begin{bmatrix} \boldsymbol{P} & \boldsymbol{p} \\ \boldsymbol{p}^\mathsf{T} & s \end{bmatrix} \succeq 0 \;,\; \boldsymbol{p} = -\boldsymbol{q}/2 - \boldsymbol{Q}\hat{\boldsymbol{\mu}} \;,\; \boldsymbol{Q} \succeq 0 \\ & \begin{bmatrix} \boldsymbol{Q} & \frac{1}{2}(\boldsymbol{q} + a_k\boldsymbol{x}) \\ \frac{1}{2}(\boldsymbol{q} + a_k\boldsymbol{x})^\mathsf{T} & r + b_k \end{bmatrix} \succeq -\tau_k \begin{bmatrix} \boldsymbol{\Theta} & -\boldsymbol{\Theta}\boldsymbol{\xi}_0 \\ -\boldsymbol{\xi}_0^\mathsf{T}\boldsymbol{\Theta} & \boldsymbol{\xi}_0^\mathsf{T}\boldsymbol{\Theta}\boldsymbol{\xi}_0 - 1 \end{bmatrix} \;,\; \forall k \\ & \tau_k \geq 0 \;\; \forall k \\ & \sum_{i=1}^{n} x_i = 1 \;,\; \boldsymbol{x} \geq 0 \;, \end{array}$$

where $\tau \in \mathbb{R}^K$ is an extra vector of optimization variables.

In both cases, the optimization problem that needs to be solved is a semi-definite program. It is well known that an interior point algorithm can be used to solve an SDP of the form

$$\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^{\tilde{n}}}{\text{minimize}} & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}_i(\boldsymbol{x}) \succeq 0 \;,\; \forall\, i \in \{1, 2, ..., \tilde{K}\} \end{array}$$

in $O\left(\left(\sum_{i=1}^{\tilde{K}} \tilde{m}_i\right)^{0.5} \left(\tilde{n}^2 \sum_{i=1}^{\tilde{K}} \tilde{m}_i^2 + \tilde{n} \sum_{i=1}^{\tilde{K}} \tilde{m}_i^3\right)\right)$, where $\tilde{m}_i$ stands for the dimension of the positive semi-definite cone (*i.e.*, $\boldsymbol{A}_i(\boldsymbol{x}) \in \mathbb{R}^{\tilde{m}_i \times \tilde{m}_i}$) (see [42]). In both SDPs that interest us, one can show that $\tilde{n} \leq n^2 + 4n + 2 + K$ and that both problems can be solved in $O(K^{3.5}n^{6.5})$ operations, with K being the number of pieces in the utility function $u(\cdot)$. We conclude that the portfolio optimization problem can be solved in $O(n^{6.5})$. □

The results presented in Theorem 4.3.1 are related to [7] where the authors proposed semi-definite programming models for solving moment problems that are similar to the one present in the objective of the DRPO. However, notice how our SDP models actually address the more involved problem of making optimal robust decisions and do not result in a heavier computational load. It is also the case that our proposed SDP models consider a more practical set of distributions which accounts for mean and covariance matrix uncertainty (in the form of a linear matrix inequality) and support information.

Remark 4.3.2 : The computational complexity presented here is based on general theory for solving semi-definite programs. Based on an implementation that uses SeDuMi ([60]), we actually observed empirically that complexity grows in the order of $O(n^5)$ for dense problems. In practice, one may also be able to exploit structure in problems where subsets (or linear combinations) of assets are known to behave independently from each other.

Remark 4.3.3 : Recently, we became aware of independent work presented in [40], which also addresses the computational difficulties related to the method proposed by Popescu. Their work is closely related to this result. Actually, for the case of unbounded support, their derivations lead to a further reduction of the DRPO model with known moments to the form of a second-order cone program. On the other hand, they do not consider support constraints and do not study the effect of moment uncertainty on the performance of a portfolio. Their approach is therefore susceptible, in practice, to the same deficiencies as Popescu's method when the moments are estimated using historical data.

4.3.2 A Case of Worst Distribution with Largest Second Moment Matrix

When presenting our distributionally robust framework, we argued in Remark 3.1.1 that a positive semi-definite lower bound on the centered second moment matrix was uninteresting. Actually, in the case of a portfolio optimization problem with piecewise linear concave utility function, the argument can be made more formally. The proof of the following proposition also provides valuable insights on the structure of a worst case distribution for the distributionally robust portfolio optimization problem.

Proposition 4.3.4 : *The distributionally robust portfolio optimization problem with piecewise linear concave utility and no support constraint on the distribution is an instance of Problem* (3.7) *where Constraint* (3.4c) *of its inner moment problem is tight for a worst case distribution.*

Proof: Consider the inner problem of our robust portfolio optimization with unconstrained support for the distribution:

$$\max_{F \in \mathcal{D}_1(\mathbb{R}^n, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 0, \gamma_2)} \mathbb{E}_F[\max_k \; -a_k \boldsymbol{\xi}^\mathsf{T} \boldsymbol{x} - b_k] \;. \tag{4.4}$$

For simplicity, we consider that there is no uncertainty in the mean of the distribution (*i.e.*, $\gamma_1 = 0$); thus, Constraint (3.4c) reduces to an upper bound on the covariance matrix of $\boldsymbol{\xi}$. The dual of this problem can be shown to reduce to:

$$\begin{array}{ll} \underset{\boldsymbol{Q}, \boldsymbol{q}, r}{\text{minimize}} & (\hat{\boldsymbol{\Sigma}} \bullet \boldsymbol{Q}) + \hat{\boldsymbol{\mu}}^\mathsf{T} \boldsymbol{Q} \hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\mu}}^\mathsf{T} \boldsymbol{q} + r \\ \text{subject to} & \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{q}/2 + a_k \boldsymbol{x}/2 \\ \boldsymbol{q}^\mathsf{T}/2 + a_k \boldsymbol{x}^\mathsf{T}/2 & r + b_k \end{bmatrix} \succeq 0 \;, \; \forall\, k \in \{1, 2, ..., K\} \;. \end{array}$$

Applying duality theory a second time leads to formulating a new equivalent version of the primal problem, which by strong duality achieves the same optimum:

$$\underset{\{(\boldsymbol{\Lambda}_k, \boldsymbol{\lambda}_k, \nu_k)\}_{k=1}^K}{\text{maximize}} \quad \sum_{k=1}^K a_k \boldsymbol{x}^\mathsf{T} \boldsymbol{\lambda}_k + \nu_k b_k \tag{4.5a}$$

$$\text{subject to} \quad \sum_{k=1}^K \boldsymbol{\Lambda}_k \preceq \gamma_2 \hat{\boldsymbol{\Sigma}} + \hat{\boldsymbol{\mu}} \hat{\boldsymbol{\mu}}^\mathsf{T} \tag{4.5b}$$

$$\sum_{k=1}^K \boldsymbol{\lambda}_k = \hat{\boldsymbol{\mu}} \;\; , \;\; \sum_{k=1}^K \nu_k = 1 \tag{4.5c}$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_k & \boldsymbol{\lambda}_k \\ \boldsymbol{\lambda}_k^\mathsf{T} & \nu_k \end{bmatrix} \succeq 0 \quad \forall\, k \in \{1, 2, ..., K\} \;. \tag{4.5d}$$

We can show that there always exists an optimal solution such that Constraint (4.5b) is satisfied with equality. Given an optimal assignment $X^* = \{(\boldsymbol{\Lambda}_k^*, \boldsymbol{\lambda}_k^*, \nu_k^*)\}_{k=1}^K$ such that $\boldsymbol{\Delta} = \gamma_2 \hat{\boldsymbol{\Sigma}} + \hat{\boldsymbol{\mu}} \hat{\boldsymbol{\mu}}^\mathsf{T} - \sum_{k=1}^K \boldsymbol{\Lambda}_k^* \succeq 0$, consider an alternate solution $X' = \{(\boldsymbol{\Lambda}_k', \boldsymbol{\lambda}_k', \nu_k')\}_{k=1}^K$ which is exactly the same as the original solution X^* except for $\boldsymbol{\Lambda}_1' = \boldsymbol{\Lambda}_1^* + \boldsymbol{\Delta}$. Obviously the two solutions achieve the same objective values since $\{(\boldsymbol{\lambda}_k^*, \nu_k^*)\}_{k=1}^K$ and $\{(\boldsymbol{\lambda}_k', \nu_k)\}_{k=1}^K$ are the same. If we can show that X' is also feasible then it is necessarily optimal. The only feasibility constraint that needs to be verified is the following:

$$\begin{bmatrix} \boldsymbol{\Lambda}_1' & \boldsymbol{\lambda}_1' \\ \boldsymbol{\lambda}_1'^{\mathsf{T}} & \nu_1' \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}_1^* & \boldsymbol{\lambda}_1^* \\ \boldsymbol{\lambda}_1^{*\mathsf{T}} & \nu_1^* \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Delta} & 0 \\ 0 & 0 \end{bmatrix} \succeq 0 \ ,$$

and is necessarily satisfied since by definition X^* is feasible and by construction $\boldsymbol{\Delta}$ is positive semi-definite. It is therefore the case that there exists a solution X^* that is optimal with respect to Problem (4.5) and satisfies Constraint (4.5b) with equality. Furthermore, one is assured that $\sum_{k=1}^{K} a_k \boldsymbol{x}^{\mathsf{T}} \boldsymbol{\lambda}_k^* + \nu_k^* b_k$ is equal to the optimal value of Problem (4.4).

After assuming without loss of generality that all $\nu_k^* > 0$, let us now construct K random vectors $\{\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, ..., \boldsymbol{\zeta}_K\}$ that satisfy the following conditions:

$$\mathbb{E}\left[\boldsymbol{\zeta}_k\right] = \frac{1}{\nu_k^*}\boldsymbol{\lambda}_k^* \ , \ \mathbb{E}\left[\boldsymbol{\zeta}_k\boldsymbol{\zeta}_k^{\mathsf{T}}\right] = \frac{1}{\nu_k^*}\boldsymbol{\Lambda}_k^* \ .$$

Note that since X^* satisfies Constraint (4.5d), we are assured that

$$\begin{aligned} \mathbb{E}\left[\boldsymbol{\zeta}_k\boldsymbol{\zeta}_k^{\mathsf{T}}\right] - \mathbb{E}\left[\boldsymbol{\zeta}_k\right]\mathbb{E}\left[\boldsymbol{\zeta}_k\right]^{\mathsf{T}} &= \begin{bmatrix} \boldsymbol{I} \\ -\mathbb{E}\left[\boldsymbol{\zeta}_k\right]^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \mathbb{E}\left[\boldsymbol{\zeta}_k\boldsymbol{\zeta}_k^{\mathsf{T}}\right] & \mathbb{E}\left[\boldsymbol{\zeta}_k\right] \\ \mathbb{E}\left[\boldsymbol{\zeta}_k\right]^{\mathsf{T}} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{I} \\ -\mathbb{E}\left[\boldsymbol{\zeta}_k\right]^{\mathsf{T}} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{I} \\ -\mathbb{E}\left[\boldsymbol{\zeta}_k\right]^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \frac{1}{\nu_k}\boldsymbol{\Lambda}_k^* & \frac{1}{\nu_k}\boldsymbol{\lambda}_k^* \\ \frac{1}{\nu_k}\boldsymbol{\lambda}_k^{*\mathsf{T}} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{I} \\ -\mathbb{E}\left[\boldsymbol{\zeta}_k\right]^{\mathsf{T}} \end{bmatrix} \\ &= \frac{1}{\nu_k^*}\begin{bmatrix} \boldsymbol{I} \\ -\mathbb{E}\left[\boldsymbol{\zeta}_k\right]^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \boldsymbol{\Lambda}_k^* & \boldsymbol{\lambda}_k^* \\ \boldsymbol{\lambda}_k^{*\mathsf{T}} & \nu_k^* \end{bmatrix} \begin{bmatrix} \boldsymbol{I} \\ -\mathbb{E}\left[\boldsymbol{\zeta}_k\right]^{\mathsf{T}} \end{bmatrix} \ \succeq \ 0 \ . \end{aligned}$$

Hence, such random vectors $\{\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, ..., \boldsymbol{\zeta}_K\}$ exist. For instance, if $\mathbb{E}\left[(\boldsymbol{\zeta}_k - \mathbb{E}\left[\boldsymbol{\zeta}_k\right])(\boldsymbol{\zeta}_k - \mathbb{E}\left[\boldsymbol{\zeta}_k\right])^{\mathsf{T}}\right] \succ 0$, then $\boldsymbol{\zeta}_k$ can take the form of a multivariate Gaussian distribution with such mean and covariance matrix. Otherwise, one could construct a lower dimensional random vector; in particular, if $\mathbb{E}\left[(\boldsymbol{\zeta}_k - \mathbb{E}\left[\boldsymbol{\zeta}_k\right])(\boldsymbol{\zeta}_k - \mathbb{E}\left[\boldsymbol{\zeta}_k\right])^{\mathsf{T}}\right] = 0$ then the random vector could have the Dirac measure $\delta_{\mathbb{E}\left[\boldsymbol{\zeta}_k\right]}$ as a distribution.

Let $\tilde{k}$ be a discrete random variable that follows a distribution with parameters $(\nu_1^*, \nu_2^*, ..., \nu_K^*)$, such that $\mathbb{P}(\tilde{k} = i) = \nu_i^*$, and use it to construct the random vector $\boldsymbol{\xi} = \boldsymbol{\zeta}_{\tilde{k}}$. Since X^* satisfies Constraint (4.5b) and Constraint (4.5c) tightly, one can show that the distribution function of $\boldsymbol{\xi}^*$ lies in $\mathcal{D}(\mathbb{R}^n, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 0, \gamma_2)$ and has largest covariance.

$$\mathbb{E}\left[\boldsymbol{\xi}^*\right] = \sum_{k=1}^{K} \mathbb{E}\left[\boldsymbol{\zeta}_k | \tilde{k} = k\right] \mathbb{P}(\tilde{k} = l) = \sum_{k=1}^{K} \frac{1}{\nu_k^*} \boldsymbol{\lambda}_k^* \nu_k^* = \hat{\boldsymbol{\mu}}$$

$$\mathbb{E}\left[\boldsymbol{\xi}^* \boldsymbol{\xi}^{*\mathsf{T}}\right] = \sum_{k=1}^{K} \mathbb{E}\left[\boldsymbol{\zeta}_k \boldsymbol{\zeta}_k^{\mathsf{T}} | \tilde{k} = k\right] \mathbb{P}(\tilde{k} = l) = \sum_{k=1}^{K} \frac{1}{\nu_k^*} \boldsymbol{\Lambda}_k^* \nu_k^* = \gamma_2 \hat{\boldsymbol{\Sigma}} + \hat{\boldsymbol{\mu}} \hat{\boldsymbol{\mu}}^{\mathsf{T}}$$

Moreover, when used as a candidate distribution in Problem (4.4) it actually achieves the maximum since we can show that it must be greater or equal to it.

$$\begin{aligned}
\mathbb{E}\left[\max_l \; -a_l \boldsymbol{x}^{\mathsf{T}} \boldsymbol{\xi}^* - b_l\right] &= \sum_{k=1}^{K} \mathbb{E}\left[\max_l \; -a_l \boldsymbol{x}^{\mathsf{T}} \boldsymbol{\zeta}_{\tilde{k}} - b_l \middle| \tilde{k} = k\right] \mathbb{P}(\tilde{k} = k) \\
&\geq \sum_{k=1}^{K} \mathbb{E}\left[-a_k \boldsymbol{x}^{\mathsf{T}} \boldsymbol{\zeta}_k - b_k\right] \mathbb{P}(\tilde{k} = k) \;=\; \sum_{k=1}^{K} -a_k \boldsymbol{x}^{\mathsf{T}} \boldsymbol{\lambda}_k^* - b_k \nu_k^* \\
&= \max_{F \,\in\, \mathcal{D}_1(\mathbb{R}^n, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 0, \gamma_2)} \mathbb{E}_F[\max_k \; -a_k \boldsymbol{x}^{\mathsf{T}} \boldsymbol{\xi} - b_k]
\end{aligned}$$

We therefore constructed a worst case distribution that has largest covariance. □

An interesting consequence of Proposition 4.3.4 is that in the framework considered in [45], if the utility function is piecewise linear concave, one can find the optimal portfolio in polynomial time using our semi-definite programming formulation with the distributional set $\mathcal{D}_1(\mathbb{R}^n, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 0, 1)$. Theoretically, our formulation is more tractable than the method proposed in [45]. However, it is true that our framework does not provide a polynomial time algorithm for the larger range of utility functions considered in [45].

4.4 Experiments with Historical Stock Data

We evaluate our portfolio optimization framework using a historical data set of 30 assets over a horizon of 15 years (1992-2007), obtained from the Yahoo! Finance web site.[2] Each experiment consists of randomly choosing 4 assets, and building a dynamic portfolio with

[2]The list of assets that is used in our experiments was inspired by [27]. More specifically, the 30 assets are: AAR Corp., Boeing Corp., Lockheed Martin, United Technologies, Intel Corp., Hitachi, Texas Instruments, Dell Computer Corp., Palm Inc., Hewlett Packard, IBM Corp., Sun Microsystems, Bristol-Myers-Squibb,

these assets through the years 2001-2007. At any given day of the experiment, the algorithms are allowed to use a period of 30 days from the most recent history to select the portfolio. All methods assume that in this period the samples are independent and identically distributed. Note that 30 samples of data might be insufficient to generate good empirical estimates of the mean and covariance matrix of returns; however, using a larger history would make the assumption of independent and identical samples somewhat unrealistic.

In implementing our method, referred to as the DRPO model, the distributional set is formulated as $\mathcal{D}_1(\mathbb{R}^4, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, 1.35, 8.32)$, where $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are the empirical estimates of the mean and covariance of $\boldsymbol{\xi}$ respectively. Due to the sample size being too small to use $\bar{\gamma}_1$ and $\bar{\gamma}_2$ from Definition 3.4.1, instead these parameters are chosen based on a simple statistical analysis of the amount of noise present in the estimation of mean and covariance matrix during the years 1992-2001.[3] We compare our DRPO model to the one proposed by [45], where the mean and covariance of the distribution F is assumed to be equal to the empirical estimates measured on the last 30 days. The method is also compared to using a naive implementation of the stochastic program, referred to as the SP model, which maximizes the average utility obtained over the last 30 days of returns. We believe that the statistics measured over a set of 300 experiments demonstrate how much there is to gain in terms of average performance and risk reduction by considering an optimization model that accounts for both distribution and moment uncertainty.

First, from the analysis of the daily returns generated by each method, one observes that they achieve comparable average daily utility. However, our DRPO model stands out as being more reliable. For instance, the lower 1st percentile of the utility distribution is 0.8% higher then the two competing methods. Also, this difference in reliability becomes more obvious when considering the respective long term performances. Figure 4.2 presents

Applera Corp.-Celera Group, Eli Lilly and Co., Merck and Co., Avery Denison Corp., Du Pont, Dow Chemical, Eastman Chemical Co., AT&T, Nokia, Motorola, Ariba, Commerce One Inc., Microsoft, Oracle, Akamai, Cisco Systems, Northern Telecom, Duke Energy Company, Exelon Corp., Pinnacle West, FMC Corp., General Electric, 'Honeywell, Ingersoll Rand.

[3] More specifically, given that one chooses 4 stocks randomly and selects a period of 60 days between 1992 and 2001 randomly, the values for γ_1 and γ_2 are chosen such that when using the first 30 days of the period to center the set $\mathcal{D}_1(\mathbb{R}^4, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \gamma_1, \gamma_2)$, the distributional set contains, with 99% probability, distributions with moments equal to the moments estimated from the last 30 days of the period.

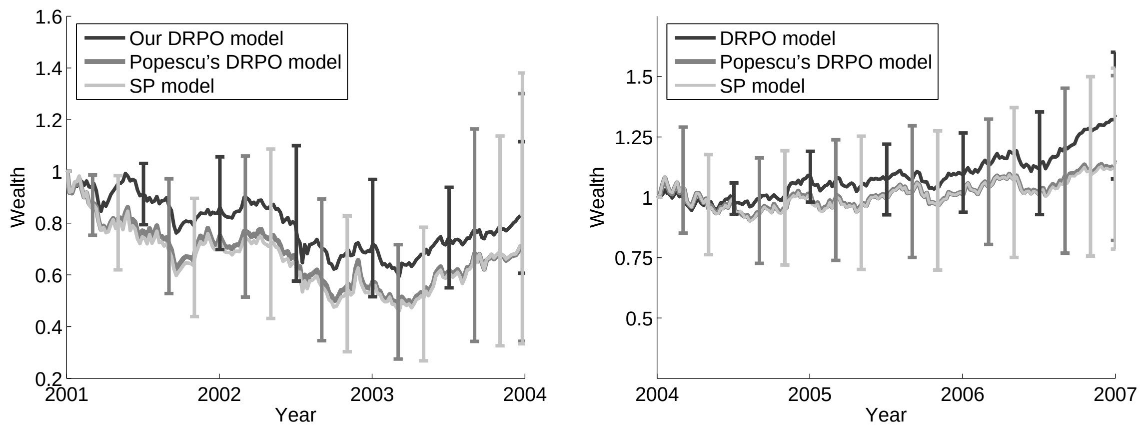


Figure 4.2: Comparison of wealth evolution in 300 experiments conducted over the years 2001-2007 using three different portfolio optimization models. For each model, the figures indicate periodically the 10th and 90th percentile of the wealth distribution in the set of experiments.

the average evolution of wealth on a six-year period when managing a portfolio of 4 assets on a daily basis with any of the three methods. In Table 4.2, the performances over the years 2001-2004 are presented separately from the performances over the years 2004-2007 in order to measure how they are affected by different levels of economic growth. The figures also periodically indicate the 10th and 90th percentile of the wealth distribution over the set of 300 experiments. The statistics of the long term experiments demonstrate empirically that our method significantly outperforms the two other ones in terms of average return and risks during both the years of economic growth and the years of decline. More

Table 4.2: Comparison of immediate and long term performance on 6 years of real stock market data.

Method	Single Day Utility (2001-2007)		Yearly Return (2001-2004)		Yearly Return (2004-2007)	
	Avg.	1st perc.	Avg.	10th perc.	Avg.	10th perc.
Our DRPO model	1.000	0.983	0.944	0.846	1.102	1.025
Popescu's DRPO model	1.000	0.975	0.700	0.334	1.047	0.936
SP model	1.000	0.973	0.908	0.694	1.045	0.923

specifically, our DRPO model outperformed Popescu's DRPO model in terms of total return accumulated over the period 2001-2007 in 79.2% of our experiments (a total set of 300 experiments). Also, it performed on average at least 1.67 times better than any competing model. Note that these experiments are purely illustrative of the strengths and weaknesses of the different models. For instance, the returns obtained in each experiment do not take into account transaction fees. The realized returns are also biased due to the fact that the assets involved in our experiments were known to be major assets in their category in January 2007. On the other hand, the realized returns were also negatively biased due to the fact that in each experiment the models were managing a portfolio of only four assets. Overall we believe that these biases affected all methods equally.

4.5 Conclusion

In this chapter, we presented three applications that can benefit from our contribution to the theory of distributionally robust optimization. In each of them, our work allowed the models to account more accurately for the uncertainty that is inherently associated with the problem. Specifically, in the optimal inequality problem we provided a tractable representation of distribution uncertainty that allows for the first time to account for general support information. On the other hand, in the fleet mix and portfolio selection problems, we showed how accounting for uncertainty in the distribution can reduce both the computational burden of solving the stochastic program and the risks related to errors in the design or estimation of the probabilistic model. Our experiments with the portfolio selection problem clearly exposed the risks and the responsibility related to choosing a probabilistic model in practice. The performance of our framework with respect to the true underlying distribution of daily returns on the stock market emphasized the need of using a decision model that accounts for both uncertainty in the distributional form as well as uncertainty in the estimation of moments. While in this chapter we restricted our analysis to cases where only first and second moments are used, we believe that important improvements could be obtained in all three applications by including more statistics of the joint distribution as proposed in Section 3.8.

Chapter 5

A Bayesian View on Distribution Uncertainty

In this chapter, we discuss methods for dealing with distribution uncertainty under the Bayesian framework. We first describe how Bayesian inference can be used to represent the decision maker's subjective belief of what is the true distribution that generates the evidence that is observed. We show how the distributionally robust model should be formulated in order to provide a solution with the probabilistic guarantees that we now expect from it. An important contribution of this work is to study the benefits and computational difficulties related to using a percentile optimization problem that addresses more directly these statistical objectives in this context. This analysis is actually performed on a special class of stochastic programs known as Markov decision processes.

5.1 Introduction

An alternative framework for representing distribution uncertainty is to use Bayesian inference. In this framework, the decision maker needs to lay down a set of assumptions about the random vector $\boldsymbol{\xi}$. In general terms, he first needs to describe a set of hypotheses about the nature of this distribution and then quantifies his level of belief in each of these hypotheses before gathering any evidence. This is done through the formulation of a "prior" distribution over the set of hypotheses. Assuming that the process that generates observations of this random vector is well defined, the evidence can then be used directly to infer a "posterior" distribution over the set of hypotheses. Specifically, one must define a set of distribution hypotheses $\{F_\theta\}$, indexed by $\theta \in \Theta$, and a prior measure $\mathcal{G}$ on the measurable space $(\Theta, \mathcal{B})$ where $\mathcal{B}$ is the Borel σ-algebra on Θ. Given that the probability of making observation $\mathcal{O}$ given the true form of F is also defined, then the posterior distribution can be obtained using Baye's rule:

$$\mathcal{G}(\theta|\mathcal{O}) = \mathbb{P}(F = F_\theta|\mathcal{O}) = \frac{\mathbb{P}(\mathcal{O}|F = F_\theta)\mathbb{P}(F = F_\theta)}{\mathbb{P}(\mathcal{O})} = \frac{\mathbb{P}(\mathcal{O}|F = F_\theta)\mathcal{G}(\theta)}{\int_\Theta \mathbb{P}(\mathcal{O}|F = F_\theta)d\mathcal{G}(\theta)} .$$

In the context of data-driven problems, the evidence is typically a set of M samples $\mathcal{O} = \{\boldsymbol{\xi}_i\}_{i=1}^M$ drawn independently and identically from F. In this case, the inference equation takes the form

$$\mathcal{G}(\theta|\mathcal{O}) = \frac{\prod_{i=1}^M F_\theta(\boldsymbol{\xi}_i)\mathcal{G}(\theta)}{\int_\Theta \prod_{i=1}^M F_\theta(\boldsymbol{\xi}_i)d\mathcal{G}(\theta)} .$$

In practice, it is convenient to use a set of hypotheses for which the conjugate prior is known. This is a set $\{F_\theta\}$ for which we know of a form $\mathcal{G}_\phi$, indexed by $\phi \in \Phi$, for the prior distribution that is preserved through the inference operation:

$$\mathcal{G}(\theta|\mathcal{O}) = \frac{\prod_{i=1}^M F_\theta(\boldsymbol{\xi}_i)\mathcal{G}_{\phi_1}(\theta)}{\int_\Theta \prod_{i=1}^M F_\theta(\boldsymbol{\xi}_i)d\mathcal{G}_{\phi_1}(\theta)} = \mathcal{G}_{\phi_2}(\theta) ,$$

for some $\phi_2 \in \Phi$. We present an example which involves a Bernoulli random variable and refer the reader to [24] for more information on this topic.

Example 5.1.1: We consider the case of a Bernoulli random variable with unknown probability of success $\theta \in [0, 1]$ and its conjugate prior, the Beta distribution $\mathcal{G}_{(\alpha,\beta)}(\theta) = \kappa(\alpha, \beta)\,\theta^{\alpha-1}(1-\theta)^{\beta-1}$ for some α and β greater or equal to 1, and some normalization function $\kappa(\alpha, \beta)$ which is known as one over the Beta function. One can verify that these are actually conjugate distributions since after observing k successes and $M-k$ failures,

$$
\begin{aligned}
\mathcal{G}(\theta|\mathcal{O}) &= \frac{\theta^k(1-\theta)^{M-k}\ \mathcal{G}_{(\alpha,\beta)}(\theta)}{\int_0^1 \theta^k(1-\theta)^{M-k}\ d\mathcal{G}_{(\alpha,\beta)}(\theta)} \\
&= \frac{\theta^k(1-\theta)^{M-k}\ \kappa(\alpha,\beta)\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\int_0^1 \theta^k(1-\theta)^{M-k}\ \kappa(\alpha,\beta)\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta} \\
&= \kappa(\alpha+k,\beta+M-k)\theta^{\alpha+k-1}(1-\theta)^{\beta+M-k-1} \\
&= \mathcal{G}_{(\alpha+k,\beta+M-k)}\ .
\end{aligned}
$$

Figure 5.1 provides the reader with more insights on how the knowledge of the distribution is improved as more samples are drawn from this Bernoulli random variable.

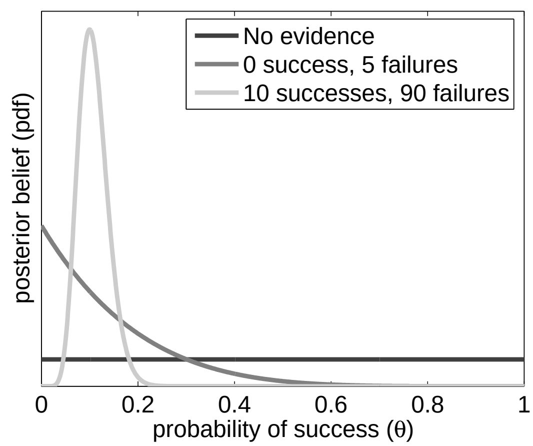


Figure 5.1: Posterior belief for the probability of success of a Bernoulli random variable with 10% probability of success.

In what follows, we study both the formulation of a distributionally robust optimization problem and the formulation of a percentile optimization problem which makes better use of the Bayesian prior.

5.2 Distributionally Robust Optimization with a Bayesian Prior

Using Bayesian inference, we just saw how one can use evidence about a random vector to formulate mathematically his posterior belief about the true distribution of this random vector. In particular, after observing M samples drawn identically and independently from F, this evidence $\mathcal{O} = \{\boldsymbol{\xi}_i\}_{i=1}^M$ is used to derive the posterior distribution $\mathcal{G}(\theta|\mathcal{O})$. In order to obtain equivalent guarantees as presented in Theorem 3.4.3 for the solution of our distributionally robust model, we need the distributional set $\mathcal{D}_1$ to contain the true distribution F with high probability according to $\mathcal{G}(\theta|\mathcal{O})$. While different methods can be used to do so, we suggest the following one.

We first resolve the shape of the distributional set. This is done by drawing N samples, $N \gg M$, $\{\theta_i\}_{i=1}^N$ independently from the posterior distribution $\mathcal{G}(\theta|\mathcal{O})$. For each of these samples, we compute $\boldsymbol{\mu}_i = \mathbb{E}_{F_{\theta_i}}[\boldsymbol{\xi}]$ and $\boldsymbol{\Sigma}_i = \mathbb{E}_{F_{\theta_i}}[(\boldsymbol{\xi} - \boldsymbol{\mu}_i)(\boldsymbol{\xi} - \boldsymbol{\mu}_i)^\mathsf{T}]$ to high accuracy either analytically or through Monte-Carlo simulation. One can then formulate an estimate of the expected mean of the distribution and expected covariance matrix: $\hat{\boldsymbol{\mu}} = \frac{1}{N}\sum_{i=1}^N \boldsymbol{\mu}_i$ and $\hat{\boldsymbol{\Sigma}} = \frac{1}{N}\sum_{i=1}^N \boldsymbol{\Sigma}_i$. These are used to center the distributional set $\mathcal{D}_1$ and give the shapes of the different moment uncertainty sets. It remains to choose the size of these sets. Using a new set of samples $\{\theta_i\}_{i=1}^N$, one can re-evaluate a new random set of triplets $\{\theta_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^N$ which are independent of our estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$. If N is large enough then choosing

$$\begin{aligned}
\bar{\gamma}_1 &= \inf_t \left\{ t \left| \frac{1}{N}\sum_{i=1}^N \mathbb{1}\left\{ (\boldsymbol{\mu}_i - \hat{\boldsymbol{\mu}})^\mathsf{T} \hat{\boldsymbol{\Sigma}}^{-1} (\boldsymbol{\mu}_i - \hat{\boldsymbol{\mu}}) \leq t \right\} \geq 1 - \delta/2 \right. \right\} \\
\bar{\gamma}_2 &= \inf_t \left\{ t \left| \frac{1}{N}\sum_{i=1}^N \mathbb{1}\left\{ (\boldsymbol{\xi} - \hat{\boldsymbol{\mu}})(\boldsymbol{\xi} - \hat{\boldsymbol{\mu}})^\mathsf{T} \leq t\hat{\boldsymbol{\Sigma}} \right\} \geq 1 - \delta/2 \right. \right\} \\
&= \inf_t \left\{ t \left| \frac{1}{N}\sum_{i=1}^N \mathbb{1}\left\{ \boldsymbol{\Sigma}_i + (\boldsymbol{\mu}_i - \hat{\boldsymbol{\mu}})(\boldsymbol{\mu}_i - \hat{\boldsymbol{\mu}})^\mathsf{T} \leq t\hat{\boldsymbol{\Sigma}} \right\} \geq 1 - \delta/2 \right. \right\} \quad ,
\end{aligned}$$

and choosing $\mathcal{S}$ such that $\mathbb{P}_{F_\theta}(\boldsymbol{\xi} \in \mathcal{S}) = 1, \forall\, \theta \in \Theta$ leads to a distributional set $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \bar{\gamma}_1, \bar{\gamma}_2)$ that is known with high probability to contain the true distribution F that is associated with $\boldsymbol{\xi}$.

Proposition 5.2.1: *Given that* $\forall\,\theta \in \Theta$, $\mathcal{G}(\theta)$ *is the probability that* F *is equal to* F_θ *before any evidence is gathered about the distribution* F *and let* $\mathcal{O} = \{\boldsymbol{\xi}_i\}_{i=1}^M$ *be* M *samples generated independently at random according to the distribution* F. *For any* $\delta > 0$, *if assumptions 3.2.4, 3.2.6, 3.3.1, and 3.3.2 are satisfied then one can solve in polynomial time the DRSP with* $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}; \bar{\gamma}_1, \bar{\gamma}_2)$ *as defined above. Furthermore, with probability greater than* $1 - \delta$, *we have that any optimal solution* $\boldsymbol{x}^*$ *of this DRSP will satisfy the constraint*

$$\mathbb{E}\left[h(\boldsymbol{x}^*, \boldsymbol{\xi})\right] \leq \Psi(\boldsymbol{x}^*; \bar{\gamma}_1, \bar{\gamma}_2) \ ,$$

where $\mathbb{E}\left[\cdot\right]$ *is the expectation with respect to the true distribution* F *of* $\boldsymbol{\xi}$, *and* $\Psi(\boldsymbol{x}^*; \bar{\gamma}_1, \bar{\gamma}_2)$ *is the optimal value of the DRSP.*

Proof: This result follows simply from Proposition 3.3.3 and the fact that by the union bound

$$\begin{aligned}
\mathbb{P}_{\mathcal{G}}\left(F \in \mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \bar{\gamma}_1, \bar{\gamma}_2)\middle|\mathcal{O}\right) &= 1 - \mathbb{P}_{\mathcal{G}}\left(F \notin \mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \bar{\gamma}_1, \bar{\gamma}_2)\middle|\mathcal{O}\right) \\
&\geq 1 - \mathbb{P}_{\mathcal{G}}\left(\|\hat{\boldsymbol{\Sigma}}^{-1/2}(\mathbb{E}_F[\boldsymbol{\mu}] - \hat{\boldsymbol{\mu}})\|^2 > \bar{\gamma}_1\middle|\mathcal{O}\right) - \mathbb{P}_{\mathcal{G}}\left(\mathbb{E}_F[(\boldsymbol{\xi} - \hat{\boldsymbol{\mu}})(\boldsymbol{\xi} - \hat{\boldsymbol{\mu}})^{\mathsf{T}}] > \bar{\gamma}_2\hat{\boldsymbol{\Sigma}}\middle|\mathcal{O}\right) \\
&\geq 1 - (1 - (1 - \delta/2 - \epsilon)) - (1 - (1 - \delta/2 - \epsilon)) = 1 - \delta - 2\epsilon \ ,
\end{aligned}$$

for any $\epsilon > 0$ with high probability given that N is large enough when selecting $\bar{\gamma}_1$ and $\bar{\gamma}_2$. □

The distributionally robust optimization problem can therefore easily be modified to account for a Bayesian prior. Unfortunately, a distributional set $\mathcal{D}_1$ that accounts simply for support, mean and covariance matrix information might not fully exploit the richness of this prior. For instance, it is common that a Bayesian prior will only consider distributions from a given family (*e.g.*, Gaussian distributions) yet this information is lost when formulating $\mathcal{D}_1$. While this concern justifies the use of a richer distributional set (*e.g.*, $\mathcal{D}_2$ presented in Section 3.8), it is important to realize that exploiting a Bayesian prior to its full potential might come at a heavy price if the prior is actually inaccurate. We therefore encourage practitioners to use distributionally robust optimization to evaluate the sensitivity of the solution to the level of assumptions that are made. In the remainder of this chapter, we assume that the prior is accurate and propose a model that actually makes the optimal choice of distributional set with respect to both the prior knowledge and the objective that is considered.

5.3 The Percentile Optimization Problem

Unlike the frequentist approach, the Bayesian framework makes the assumption that the distribution F associated with the random vector $\boldsymbol{\xi}$ was drawn from a probability measure $\mathcal{G}$ that is known. Intuitively, this means that we know that the probability that F takes the form F_θ is exactly $\mathcal{G}(\theta|\mathcal{O})$. For this reason, one might question if $\mathcal{D}_1$ is the right distributional set to use in order to achieve the desired probabilistic guarantees on the performance of the solution. In fact, the Bayesian framework allows us to conceive the possibility of choosing both $\mathcal{D}$ and $\boldsymbol{x}$ optimally through an optimization problem of the form

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathcal{X},\mathcal{D}\subseteq\mathcal{M}}{\text{minimize}} \quad & \max_{F\in\mathcal{D}} \mathbb{E}_F[h(\boldsymbol{x},\boldsymbol{\xi})] & \text{(5.1a)} \\ \text{subject to} \quad & \mathbb{P}_{\mathcal{G}}\left(F_\theta \in \mathcal{D}|\mathcal{O}\right) \geq 1-\delta \quad , & \text{(5.1b)} \end{aligned}$$

where $\mathbb{P}_{\mathcal{G}}(F_\theta \in \mathcal{D}|\mathcal{O})$ is measured with respect to drawing θ according to the distribution $\mathcal{G}(\theta|\mathcal{O})$. This chance constraint ensures that $\mathcal{D}$ is chosen to satisfy the same probabilistic properties as were satisfied by the set $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \bar{\gamma}_1, \bar{\gamma}_2)$ described in Section 5.2. After presenting how Problem (5.1) can be reformulated as a percentile optimization problem, we describe more formally how its solution improves upon the solution from the DRSP that was discussed earlier.

Lemma 5.3.1 : *Problem* (5.1) *is equivalent to the following percentile optimization problem*

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathcal{X},y\in\mathbb{R}}{\text{minimize}} \quad & y & \text{(5.2a)} \\ \text{subject to} \quad & \mathbb{P}_{\mathcal{G}}\left(\, y \geq \mathbb{E}_{F_\theta}[h(\boldsymbol{x},\boldsymbol{\xi})] \,|\, \mathcal{O}\,\right) \geq 1-\delta \quad , & \text{(5.2b)} \end{aligned}$$

where $\mathbb{P}_{\mathcal{G}}(\cdot|\mathcal{O})$ is measured with respect to drawing θ according to the posterior distribution $\mathcal{G}(\theta|\mathcal{O})$.

Proof: To prove the equivalence of the two problems, we first insert the slack variable y in Problem (5.1) to have its objective take the same form as the percentile optimization problem:

$$\begin{array}{ll} \underset{\boldsymbol{x}\in\mathcal{X},\mathcal{D}\in\mathcal{M},y\in\mathbb{R}}{\text{minimize}} & y \\ \text{subject to} & y \geq \max_{F\in\mathcal{D}} \mathbb{E}_{F_\theta}[h(\boldsymbol{x},\boldsymbol{\xi})] \\ & \mathbb{P}_{\mathcal{G}}\left(F\in\mathcal{D}|\mathcal{O}\right) \geq 1-\delta \end{array}$$

We can then show that any pair $(\boldsymbol{x}, y)$ that is feasible with some $\mathcal{D}$ according to this problem is also feasible in the percentile optimization problem. Conversely, given any feasible pair $(\boldsymbol{x}, y)$ for the percentile optimization problem, one can find a set $\mathcal{D}$ which makes the triplet $(\boldsymbol{x}, \mathcal{D}, y)$ feasible according to Problem (5.1). Since the two objective functions are the same then the two problems are equivalent.

Thus, we only need to confirm the feasibility equivalence. First, given some triplet $(\boldsymbol{x}, \mathcal{D}, y)$ that we know is feasible for Problem (5.1), we know that there exists a set $\bar{\Theta} \subseteq \Theta$ such that $F_\theta \in \mathcal{D}$ for all $\theta \in \bar{\Theta}$ and that $\mathbb{P}_{\mathcal{G}}(\theta \in \bar{\Theta}|\mathcal{O}) \geq 1-\delta$. One can then verify that

$$\begin{aligned} \mathbb{P}_{\mathcal{G}}\left(y \geq \mathbb{E}_{F_\theta}[h(\boldsymbol{x},\boldsymbol{\xi})]|\mathcal{O}\right) &\geq \mathbb{P}_{\mathcal{G}}\left(y \geq \mathbb{E}_{F_\theta}[h(\boldsymbol{x},\boldsymbol{\xi})]\middle|\theta\in\bar{\Theta},\mathcal{O}\right)\mathbb{P}_{\mathcal{G}}(\theta\in\bar{\Theta}|\mathcal{O}) \\ &\geq \mathbb{P}_{\mathcal{G}}\left(y \geq \max_{\theta\in\bar{\Theta}} \mathbb{E}_{F_\theta}[h(\boldsymbol{x},\boldsymbol{\xi})]\middle|\theta\in\bar{\Theta},\mathcal{O}\right)(1-\delta) \\ &\geq \mathbb{P}_{\mathcal{G}}\left(y \geq \max_{F\in\mathcal{D}} \mathbb{E}_{F}[h(\boldsymbol{x},\boldsymbol{\xi})]\middle|\theta\in\bar{\Theta},\mathcal{O}\right)(1-\delta) \;=\; 1-\delta\,. \end{aligned}$$

On the other hand, if we have a pair $(\boldsymbol{x}, y)$ that is feasible according to the percentile optimization problem, then it must be that there is a set $\bar{\Theta}$ such that $y \geq \max_{\theta\in\bar{\Theta}} \mathbb{E}_{F_\theta}[h(\boldsymbol{x},\boldsymbol{\xi})]$ and $\mathbb{P}_{\mathcal{G}}(\theta \in \bar{\Theta}|\mathcal{O}) \geq 1-\delta$. We can therefore construct a set $\bar{\mathcal{D}} = \{F_\theta|\theta \in \bar{\Theta}\}$ and be assured that $\mathbb{P}_{\mathcal{G}}(F \in \bar{\mathcal{D}}|\mathcal{O}) \geq 1-\delta$. The set $\bar{\mathcal{D}}$ is therefore feasible according to Problem (5.1) and so are y and $\boldsymbol{x}$ since

$$y \;\geq\; \max_{\theta\in\bar{\Theta}} \mathbb{E}_{F_\theta}[h(\boldsymbol{x},\boldsymbol{\xi})] \;=\; \max_{F\in\bar{\mathcal{D}}} \mathbb{E}_{F_\theta}[h(\boldsymbol{x},\boldsymbol{\xi})]\;. \quad \square$$

Proposition 5.3.2: *Given that $\forall\,\theta \in \Theta$, $\mathcal{G}(\theta)$ is the probability that F is equal to F_θ before* *any evidence is gathered about the distribution* F *and that* $\mathcal{O} = \{\boldsymbol{\xi}_i\}_{i=1}^M$ *are* M *samples generated independently at random according to the distribution* F*, then, with probability greater than* $1-\delta$ *we have that an optimal solution pair* $(\boldsymbol{x}^*, y^*)$ *of Problem* (5.2) *satisfies the constraint*

$$\mathbb{E}\left[h(\boldsymbol{x}^*, \boldsymbol{\xi})\right] \leq y^* \leq \Psi(\boldsymbol{x}^*_{\mathit{DRSP}}; \bar{\gamma}_1, \bar{\gamma}_2) \ ,$$

where $\mathbb{E}\left[\cdot\right]$ *is the expectation with respect to the true distribution* F*, and* $\Psi(\boldsymbol{x}^*_{\mathit{DRSP}}; \bar{\gamma}_1, \bar{\gamma}_2)$ *is the optimal value of the DRSP with* $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \bar{\gamma}_1, \bar{\gamma}_2)$.

Proof: The probabilistic guarantees are a direct consequence of the chance constraint which enforces that $\mathbb{P}_{\mathcal{G}}(\mathbb{E}\left[h(\boldsymbol{x}^*, \boldsymbol{\xi})\right] \leq y^*) \geq 1-\delta$ is satisfied. To show that the bounds obtained from the percentile optimization problem improves on the bound obtained from the distributionally robust model, one simply needs to observe that Problem (5.1) is a relaxation of the DRSP with $\mathcal{D}_1(\mathcal{S}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \bar{\gamma}_1, \bar{\gamma}_2)$ and exploit the relation presented in Lemma 5.3.1. □

The percentile optimization problem presented in Problem (5.2) can actually be recognized as a member of chance-constrained problems which have been widely studied for single-period optimization problems (*e.g.*, in [13, 46]). Unfortunately, although conceptually these problems can lead to better performing solutions, they present a much harder computational challenge. In addition to being infinite dimensional and non-convex, problems with chance constraints are suspected to be "severely computationally intractable" [41]. In what follows, we study in more depth how the computational challenge balances against the performance benefits of this model for the special case of a Markov decision process with uncertainty in the model's parameters.

5.4 Markov Decision Processes with Model Uncertainty

We consider an infinite horizon Markov Decision Process (MDP) that can be described as follows: a finite state space S with $|S|$ states, a finite action space A with $|A|$ actions, a transition probability matrix $\boldsymbol{P} \in \mathbb{R}^{|S|\times|A|\times|S|}$ with $P_{(s,a,s')} = \mathbb{P}(s'|s,a)$, an initial distribution on states $\mathbb{P}(x_0 = i) = q_i$ for some $\boldsymbol{q} \in \mathbb{R}^{|S|}$, and a reward vector $\boldsymbol{r} \in \mathbb{R}^{|S|}$. In the context of an infinite horizon MDP, one can choose to apply a mixed policy π, which is a mapping from the set of states S to the probability simplex over the available actions. For reasons of tractability, we limit our attention to the set of stationary Markov policies, which is denoted by Υ. When considering an infinite horizon, an optimal discounted reward stationary policy π is the solution to the following optimization problem:

$$\text{(Nominal MDP)} \qquad \underset{\pi \in \Upsilon}{\text{maximize}} \;\; \mathbb{E}_{\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \alpha^t r(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi\right] ,$$

where $\alpha \in [0, 1)$ is the discount factor, and $r(i)$ is the reward obtained in state i, *i.e.*, $r(i) = r_i$. This problem is known to be easily solvable using value iteration (see [47] and [3] for more details on the subject). However, this model does not take into account any uncertainty in the choice of the parameters $\boldsymbol{P}$ and $\boldsymbol{r}$. Since the parameters of an MDP are typically either estimated from data or learned from experience, it is not surprising that, in some applications, unavoidable modeling uncertainty often causes the long term performance of a strategy to significantly differ from the model's predictions (refer to experiments in [37]).

To date, most efforts have focused on the study of robust (*i.e.*, distributionally robust) MDPs (*e.g.*, in [43, 31, 25]). In this context, under the assumption that the reward parameters and the transition probability parameters lie in their respective uncertainty sets, one considers choosing the best strategy for the worst case realization of these parameters. When the reward vector $\boldsymbol{r}$ and the transition probability matrix $\boldsymbol{P}$ are known to lie in a set R and P respectively, the robust problem thus becomes:

$$\text{(Robust MDP)} \qquad \underset{\pi \in \Upsilon}{\text{maximize}} \;\; \min_{\boldsymbol{P} \in \mathsf{P}, \boldsymbol{r} \in \mathsf{R}} \mathbb{E}_{\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \alpha^t r(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi\right] .$$

The type of reward uncertainty that is considered here is also known as *fixed uncertainty* since the reward vector is drawn once and remains fixed for all time steps. It is a well known fact that, if R and P are convex sets then the optimal policy π^* for the Robust MDP can be found efficiently (see [3] and [43]).

Given that one has formulated a Bayesian prior for the MDP, after performing Bayesian inference, he can actually consider that the reward vector $\tilde{\boldsymbol{r}}$ and transition probability matrix

$\tilde{\boldsymbol{P}}$ are randomly drawn from their respective posterior probability distribution $\mathcal{G}_{\tilde{\boldsymbol{r}}}$ and $\mathcal{G}_{\tilde{\boldsymbol{P}}}$. Although confidence regions can be used to formulate the robust problem, it is more natural to use a percentile based approach since the real concern is about the level of confidence in the total cumulative reward and not in the location of the parameters themselves. In this context, the percentile optimization problem takes the following Percentile MDP form:

$$\underset{y\in\mathbb{R},\pi\in\Upsilon}{\text{maximize}} \quad y \tag{5.3a}$$

$$\text{subject to} \quad \mathbb{P}_{\mathcal{G}}\left(\mathbb{E}_{\tilde{\boldsymbol{P}}}\left[\textstyle\sum_{t=0}^{\infty}\alpha^t\tilde{r}(x_t)|x_0 \propto \boldsymbol{q},\pi\right] \geq y\right) \geq 1-\delta\ , \tag{5.3b}$$

where the probability $\mathbb{P}_{\mathcal{G}}$ is the probability of drawing the reward vector $\tilde{\boldsymbol{r}}$ from $\mathcal{G}_{\tilde{\boldsymbol{r}}}$ and the transition probability matrix $\tilde{\boldsymbol{P}}$ from $\mathcal{G}_{\tilde{\boldsymbol{P}}}$, and where $\mathbb{E}_{\tilde{\boldsymbol{P}}}[\sum_{t=0}^{\infty}\alpha^t\tilde{r}(x_t)|x_0 \propto \boldsymbol{q},\pi]$ is the expectation with respect to the random trajectory of x_t given a concrete realization of $\tilde{\boldsymbol{r}}$ and $\tilde{\boldsymbol{P}}$, a policy π, and a distribution $\boldsymbol{q}$ for the initial state x_0. For a given policy π, the above percentile optimization problem gives us a $1-\delta$ guarantee under the distribution of $\tilde{\boldsymbol{r}}$ and $\tilde{\boldsymbol{P}}$ that π will perform better than y^*, the optimal value of Problem (5.3).[1]

In what follows, we will describe the spectrum of computational difficulties related to solving the percentile criterion when it is applied to an MDP with model uncertainty. For completeness, we start by considering uncertainty in the rewards and show that, although in one special case the Percentile MDP can be solved using a deterministic "second-order cone program", in general finding the optimal solution of the Percentile MDP is NP-hard. We then address the more interesting question of uncertainty in the transition probabilities of the Markov chain and present an approximation method for finding an optimal policy of the Percentile MDP.

Notation

In this section, the following notation is used. $\mathbf{1}_K$ is the vector of all ones in $\mathbb{R}^K$. For clarity, $Q_{(i,j)}$ will refer to the i-th row, j-th column term of a matrix $\boldsymbol{Q}$. Also, for the sake of simpler linear manipulations, we will present a policy π under its matrix form $\boldsymbol{\Pi} \in \mathbb{R}^{|S|\times|S|\times|A|}$, such that $\Pi_{(s_1,s_2,a)} = \pi(s_1,a)\mathbb{1}\{s_1 = s_2\}$. When this three dimensional matrix will be

[1] Note that [21] introduced the percentile criterion as a risk-adjusted performance measure for "average reward" MDPs. However, their study did not address the question of parameter uncertainty.

multiplied to another matrix $\boldsymbol{Q} \in \mathbb{R}^{|S|\times|A|\times K}$, this will refer to a matrix multiplication carried along $\mathbb{R}^{|S|\times(|S|\,|A|)} \times \mathbb{R}^{(|S|\,|A|)\times K}$, such that $(\boldsymbol{\Pi Q})_{(i,j)} = \sum_{(k,a)} \Pi_{(i,k,a)} Q_{(k,a,j)}$. Note that this formulation explicitly denotes the linear relation between the decision variable $\boldsymbol{\Pi}$ and the inferred transition probability $\boldsymbol{P}_\pi$, such that $(\boldsymbol{\Pi P})_{(i,j)} = (\boldsymbol{P}_\pi)_{(i,j)} = \mathbb{P}(s' = j|s = i, a \propto \pi(i))$.

5.4.1 Decision making under uncertain reward parameters

First, the problem of reward uncertainty is addressed for a common family of distribution functions, the multivariate Gaussian distribution $\tilde{\boldsymbol{r}} \propto \mathcal{N}(\boldsymbol{\mu}_{\tilde{\boldsymbol{r}}}, \boldsymbol{\Sigma}_{\tilde{\boldsymbol{r}}})$. Under the assumption of Gaussian rewards, solving the Percentile MDP is not considerably harder than solving the Nominal MDP. However, we later prove that there exist distributions over the parameters for which the percentile optimization problem becomes intractable.

Reward uncertainty with Gaussian distribution

The Gaussian assumption is standard in many applications as it allows to model the correlation between the reward obtained in different states. Also, in the Bayesian framework it is common to use a Gaussian prior, with parameters $(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, for $\tilde{\boldsymbol{r}}$. Then, based on new independent measurements of the rewards $\{\hat{\boldsymbol{r}}_1, \hat{\boldsymbol{r}}_2, ..., \hat{\boldsymbol{r}}_M\}$ that are corrupted with Gaussian noise such that $\hat{\boldsymbol{r}} = \tilde{\boldsymbol{r}} + \boldsymbol{\nu}$ with $\boldsymbol{\nu} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{\nu}})$, one can obtain an analytical posterior over $\tilde{\boldsymbol{r}}$. This posterior distribution also has the Gaussian shape with parameters (see [24] for more details):

$$\boldsymbol{\mu}_1 = \boldsymbol{\Sigma}_1 \left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1} \sum_{i=1}^{M} \hat{\boldsymbol{r}}_i \right), \qquad \boldsymbol{\Sigma}_1 = \left(\boldsymbol{\Sigma}_0^{-1} + M\boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1} \right)^{-1}.$$

Lemma 5.4.1: *(Prékopa [46], Theorem 10.4.1) Suppose $\boldsymbol{\xi} \in \mathbb{R}^n$ has a multivariate Gaussian distribution. Then, the set of $\boldsymbol{z} \in \mathbb{R}^n$ vectors satisfying*

$$\mathbb{P}(\boldsymbol{z}^\mathsf{T}\boldsymbol{\xi} \leq 0) \geq 1 - \delta$$

is the same as those satisfying

$$\boldsymbol{z}^{\mathsf{T}}\boldsymbol{\mu}_{\boldsymbol{\xi}} + \Phi^{-1}(1-\delta)\sqrt{\boldsymbol{z}^{\mathsf{T}}\boldsymbol{\Sigma}_{\boldsymbol{\xi}}\boldsymbol{z}} \;\; \leq \;\; 0 \,,$$

where $\boldsymbol{\mu}_{\boldsymbol{\xi}} = \mathbb{E}\left[\boldsymbol{\xi}\right]$, $\boldsymbol{\Sigma}_{\boldsymbol{\xi}}$ *is the covariance matrix of the random vector* $\boldsymbol{\xi}$, δ *is a fixed probability such that* $0 \leq \delta \leq 1$, *and* Φ *is the cumulative distribution function of* $\mathcal{N}(0,1)$.

Lemma 5.4.1 is an important result in the field of stochastic programming. In our specific context, the lemma allows us to show that finding an optimal stationary policy for the problem of maximizing the $(1-\delta)$-percentile criterion under Gaussian uncertainty can be expressed as a second-order cone program.

Theorem 5.4.2 : *For any* $\delta \in (0, 0.5]$, *the Percentile MDP with Gaussian uncertainty in the rewards*

$$\begin{aligned} \underset{y\in\mathbb{R},\pi\in\Upsilon}{\text{maximize}} \quad & y && \text{(5.4a)} \\ \text{subject to} \quad & \mathbb{P}_{\mathcal{G}}(\mathbb{E}_{\boldsymbol{P}}[\textstyle\sum_{t=0}^{\infty} \alpha^t \tilde{r}(x_t) | x_0 \propto \boldsymbol{q}, \pi] \geq y) \geq 1-\delta \,, && \text{(5.4b)} \end{aligned}$$

where the expectation is taken with respect to the random trajectory of x_t *when following stationary policy* π, *and where* $\tilde{\boldsymbol{r}} \propto \mathcal{N}(\boldsymbol{\mu}_{\tilde{\boldsymbol{r}}}, \boldsymbol{\Sigma}_{\tilde{\boldsymbol{r}}})$, *is equivalent to the convex second-order cone program*

$$\begin{aligned} \underset{\boldsymbol{\rho}_1,\boldsymbol{\rho}_2,\ldots,\boldsymbol{\rho}_{|A|}\in\mathbb{R}^{|S|}}{\text{maximize}} \quad & \textstyle\sum_a \boldsymbol{\rho}_a^{\mathsf{T}}\boldsymbol{\mu}_{\tilde{\boldsymbol{r}}} - \Phi^{-1}(1-\delta)\|\sum_a \boldsymbol{\rho}_a^{\mathsf{T}}\boldsymbol{\Sigma}_{\tilde{\boldsymbol{r}}}^{\frac{1}{2}}\|_2 && \text{(5.5a)} \\ \text{subject to} \quad & \textstyle\sum_a \boldsymbol{\rho}_a^{\mathsf{T}} = \boldsymbol{q}^{\mathsf{T}} + \sum_a \alpha\boldsymbol{\rho}_a^{\mathsf{T}}\boldsymbol{P}_a && \text{(5.5b)} \\ & \boldsymbol{\rho}_a^{\mathsf{T}} \geq 0 \,, \quad \forall \;\, a \in A \,, && \text{(5.5c)} \end{aligned}$$

where $\boldsymbol{P}_a$ *is the transition probability matrix when action* a *is taken in each state. Given an optimal assignment for* $\boldsymbol{\rho}_1^*, \boldsymbol{\rho}_2^*, \ldots, \boldsymbol{\rho}_{|A|}^*$, *an optimal policy* π^* *to Problem* (5.4) *can be retrieved using:*

$$\pi^*(s,a) = \begin{cases} \frac{1}{|A|} & \text{if } \sum_a \rho^*_{(a,s)} = 0 \\ \frac{\rho^*_{(a,s)}}{\sum_a \rho^*_{(a,s)}} & \text{otherwise.} \end{cases}$$

Proof: We first use the fact that Constraint (5.4b) can be expressed in the form

$$\mathbb{P}_{\mathcal{G}}(\boldsymbol{v}^\mathsf{T}\tilde{\boldsymbol{r}} \geq y) \geq 1-\delta \tag{5.6a}$$

$$\boldsymbol{q}^\mathsf{T}\sum_{t=0}^{\infty}(\alpha\boldsymbol{\Pi}\boldsymbol{P})^t = \boldsymbol{v}^\mathsf{T}\ . \tag{5.6b}$$

Using a change of variables that is commonly used in the MDP literature (see [47]), Constraint (5.6b) is equivalent to:

$$\boldsymbol{v}^\mathsf{T} = \boldsymbol{q}^\mathsf{T} + \alpha\textstyle\sum_a \boldsymbol{\rho}_a^\mathsf{T}\boldsymbol{P}_a$$
$$\boldsymbol{v}^\mathsf{T} = \textstyle\sum_{a\in A}\boldsymbol{\rho}_a^\mathsf{T}\ ,\ \ \boldsymbol{\rho}_a^\mathsf{T} \geq 0\ ,\quad \forall\ a\in A\ ,$$

with some set of new decision variables: $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, ..., \boldsymbol{\rho}_{|A|} \in \mathbb{R}^{|S|}$. From a feasible assignment $(\boldsymbol{v}, \boldsymbol{\rho}_1, , \boldsymbol{\rho}_2, ..., \boldsymbol{\rho}_{|A|})$, an equivalent pair $(\boldsymbol{v}, \boldsymbol{\Pi})$ feasible according to Constraint (5.6b) can be retrieved using:

$$\Pi_{(s,s',a)} = \begin{cases} 0 & \text{if } v_{s'} = 0 \\ \frac{\rho_{(a,s')}}{v_{s'}}\mathbb{1}\{s=s'\} & \text{otherwise.} \end{cases}$$

Given that $\delta \leq 0.5$, one can use Lemma 5.4.1 to convert Constraint (5.6a) into an equivalent deterministic convex constraint. Theorem 5.4.2 follows naturally. □

Complexity of the solution

It is important to note that second-order cone programming (SOCP) is a well developed field of optimization for which a number of polynomial time algorithms have been proposed. We refer the reader to [36, 10] for background on the subject and algorithms for solving this family of problems.[2] Based on a primal-dual interior point method presented in [36], we can show the following.

Theorem 5.4.3: *Given an MDP with $|S|$ states, $|A|$ actions, and Gaussian uncertainty in the reward vector, the Percentile MDP Problem* (5.4) *can be solved in time* $O(|A|^{\frac{7}{2}}|S|^{\frac{7}{2}})$.

[2] In our implementation, we used the CVX toolbox developed for Matlab by Michael Grant *et al*. [28].

Proof: Based on the work presented in [36], the computations involved in solving an SOCP to any precision can be bounded above by $O\left(\sqrt{K}(k^2\sum_{i=1}^{K}k_i + k^3)\right)$, where K is the number of constraints, k is the number of variables, and k_i is the size of the vector in the norm operator of constraint i. These results lead to a bound of $O(|A|^{\frac{7}{2}}|S|^{\frac{7}{2}})$ for Problem (5.5) and consequently for Problem (5.4) since the transformation from one problem to the other does not depend on the size of the MDP. □

Unfortunately, one can also show that some uncertainty models for the reward parameters actually lead to intractable forms of the Percentile MDP Problem (5.4).

Theorem 5.4.4 : *Solving the Percentile MDP Problem* (5.4) *with* ***general uncertainty*** *in the reward parameters is NP-hard.*

A detailed proof of this theorem is presented in Appendix B.1 where we show that the NP-complete 3SAT problem can be reduced to solving Problem (5.4) for an MDP with discrete reward uncertainty.

5.4.2 Decision making under uncertain transition probabilities

We now focus on the problem of transition probability uncertainty. This type of uncertainty is present in applications where one does not have a physical model of the dynamics of the system. In this case, the transition probability matrix $\boldsymbol{P}$ must be estimated from experimentation and is therefore inherently uncertain. Since the Bayesian framework allows us to formulate a distribution over $\boldsymbol{P}$, we consider a Percentile MDP problem with transition probability uncertainty:

$$\underset{y\in\mathbb{R},\pi\in\Upsilon}{\text{maximize}} \quad y \tag{5.7a}$$

$$\text{subject to} \quad \mathbb{P}_{\mathcal{G}}\left(\mathbb{E}_{\tilde{\boldsymbol{P}}}[\textstyle\sum_{t=0}^{\infty}\alpha^t r(x_t)|x_0 \propto \boldsymbol{q},\pi] \geq y\right) \geq 1-\delta\ , \tag{5.7b}$$

where the probability $\mathbb{P}_{\mathcal{G}}(\cdot)$ is the probability of drawing the transition probability matrix $\tilde{\boldsymbol{P}}$ from a posterior distribution $\mathcal{G}_{\tilde{\boldsymbol{P}}}$ and where $\mathbb{E}_{\tilde{\boldsymbol{P}}}[\cdot|x_0 \propto \boldsymbol{q},\pi]$ is the expectation with respect to the random trajectory of x_t given a concrete realization of $\tilde{\boldsymbol{P}}$, deterministic rewards $\boldsymbol{r}$, a policy π, and a distribution $\boldsymbol{q}$ for the initial state . As was the case for reward uncertainty, this problem is hard to solve in general. However, in Section 5.4.2 we use the Dirichlet prior to suggest a method that generates a near optimal policy given a sufficient number of observations of transitions drawn from $\tilde{\boldsymbol{P}}$.

Computational complexity of uncertainty in the transition probabilities

Finding an optimal policy for a Percentile MDP problem is NP-hard even when there is no uncertainty in the reward parameters.

Corollary 5.4.5: *Solving the Percentile MDP Problem* (5.7) *for general uncertainty in the transition probability parameters is NP-hard.*

Following similar lines as for proving Theorem 5.4.4, given an instance of the NP-complete 3SAT Problem, one can easily construct in polynomial time an MDP with discrete uncertainty in the transition probabilities. Solving Problem (5.7) for this uncertain MDP is equivalent to determining if the 3SAT instance is satisfiable. A sketch of this proof is included in Appendix B.2.

The Dirichlet prior on transition probabilities

Since we cannot expect to solve the Percentile MDP Problem (5.7) for a general distribution, we use, for each state-action pair (i, a), independent Dirichlet priors to model the uncertainty in the parameter $P_{(i,a,j)}$ associated with each j. This assumption is very convenient for describing prior knowledge about transition probability parameters due to the fact that, after gathering new transition observations, one can easily evaluate a posterior distribution over these parameters. More specifically, for a vector of transition probability parameters $\tilde{\boldsymbol{p}} = [\tilde{p}_1, ..., \tilde{p}_{|S|}]^\mathsf{T}$, the Dirichlet distribution over $\tilde{\boldsymbol{p}}$ follows the density function $\mathcal{G}(\boldsymbol{p}) = (1/Z(\beta)) \prod_{j=1}^{|S|} p_j^{\beta_j - 1}$, where β are modeling parameters for the Dirichlet prior and $Z(\beta)$ is a normalization factor. Given a set of observed transition observations $\{j^{(1)}, j^{(2)}, ..., j^{(M)}\}$ drawn from the multinomial distribution $F_{\boldsymbol{p}}(j) = p_j$, one can analytically resolve the posterior distribution over $\tilde{\boldsymbol{p}}$. This distribution conveniently takes the same Dirichlet form $\mathcal{G}(\boldsymbol{p}|j^{(1)}, j^{(2)}, ..., j^{(M)}) = (1/Z(\beta, M_1, ..., M_{|S|})) \prod_{j=1}^{|S|} p_j^{\beta_j + M_j - 1}$, where M_j

is the number of times that a transition to j was observed. It is also known that the covariance between different terms of $\tilde{\boldsymbol{p}}$ is (see [24] for details):

$$\Sigma_{(j,k)} = \begin{cases} \frac{(\beta_j+M_j)(\beta_0+M-\beta_j-M_j)}{(\beta_0+M)^2(\beta_0+M+1)} & \text{if } k=j \\ -\frac{(\beta_k+M_k)(\beta_j+M_j)}{(\beta_0+M)^2(\beta_0+M+1)} & \text{otherwise} \end{cases},$$

where $\beta_0 = \sum_j \beta_j$ and $M = \sum_j M_j$.

Expected return approximation using a Dirichlet prior

Even with the Dirichlet assumption we are confronted with the following difficulty in solving the Percentile MDP Problem (5.7). Unlike in the case of reward uncertainty, finding a policy that simply minimizes the expected return $\mathbb{E}_{\mathcal{G}}[\sum_{t=0}^{\infty} \alpha^t r(x_t)|x_0 \propto \boldsymbol{q}, \pi]$ under transition probability uncertainty $\tilde{\boldsymbol{P}}$ is already hard. More specifically, the expected return can be expressed as

$$\begin{aligned} \mathbb{E}_{\mathcal{G}}\left[\sum_{t=0}^{\infty} \alpha^t r(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi\right] &= \mathbb{E}_{\mathcal{G}}\left[\mathbb{E}_{\tilde{\boldsymbol{P}}}\left[\sum_{t=0}^{\infty} \alpha^t r(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi\right]\right] \\ &= \mathbb{E}_{\mathcal{G}}\left[\boldsymbol{q}^\mathsf{T}(\boldsymbol{I} - \alpha\boldsymbol{\Pi}\tilde{\boldsymbol{P}})^{-1}\boldsymbol{r}\right] \\ &= \mathbb{E}_{\mathcal{G}}\left[\boldsymbol{q}^\mathsf{T}\left(\boldsymbol{I} - \alpha\boldsymbol{\Pi}(\mathbb{E}_{\mathcal{G}}[\tilde{\boldsymbol{P}}] + \Delta\tilde{\boldsymbol{P}})\right)^{-1}\boldsymbol{r}\right] \\ &= \mathbb{E}_{\mathcal{G}}\left[\boldsymbol{q}^\mathsf{T}\left((\boldsymbol{X}^\pi)^{-1} - (\boldsymbol{X}^\pi)^{-1}\alpha\boldsymbol{X}^\pi\boldsymbol{\Pi}\Delta\tilde{\boldsymbol{P}}\right)^{-1}\boldsymbol{r}\right] \\ &= \mathbb{E}_{\mathcal{G}}\left[\boldsymbol{q}^\mathsf{T}(\boldsymbol{I} - \alpha\boldsymbol{X}^\pi\boldsymbol{\Pi}\Delta\tilde{\boldsymbol{P}})^{-1}\boldsymbol{X}^\pi\boldsymbol{r}\right] \\ &= \mathbb{E}_{\mathcal{G}}\left[\boldsymbol{q}^\mathsf{T}\sum_{k=0}^{\infty}\alpha^k(\boldsymbol{X}^\pi\boldsymbol{\Pi}\Delta\tilde{\boldsymbol{P}})^k\boldsymbol{X}^\pi\boldsymbol{r}\right], \end{aligned}$$

where $\Delta\tilde{\boldsymbol{P}} = \tilde{\boldsymbol{P}} - \mathbb{E}_{\mathcal{G}}[\tilde{\boldsymbol{P}}]$, and $\boldsymbol{X}^\pi = (\boldsymbol{I} - \alpha\boldsymbol{\Pi}\mathbb{E}_{\mathcal{G}}[\tilde{\boldsymbol{P}}])^{-1}$. The matrix $\boldsymbol{X}^\pi$ is always well defined since $\tilde{\boldsymbol{P}}$ is modeled with the Dirichlet distribution, thus ensuring that $\mathbb{E}_{\mathcal{G}}[\tilde{\boldsymbol{P}}]$ is a valid transition probability matrix and that $\boldsymbol{I} - \alpha\boldsymbol{\Pi}\mathbb{E}_{\mathcal{G}}[\tilde{\boldsymbol{P}}]$ is nonsingular. The expression $\mathbb{E}_{\mathcal{G}}[\sum_{t=0}^{\infty} \alpha^t r(x_t)|x_0 \propto \boldsymbol{q}, \pi]$ therefore depends on all the moments of the uncertainty in $\tilde{\boldsymbol{P}}$. Following similar lines as in [37], we focus on finding a stationary policy that performs well according to the second-order approximation of the expected return. We expect the norm of higher order moments of $\Delta\tilde{\boldsymbol{P}}$ to decay with the number of observed transitions.

$$\begin{aligned}
&\mathbb{E}_{\mathcal{G}}\left[\sum_{t=0}^{\infty}\alpha^t r(x_t)\middle| x_0 \propto \boldsymbol{q}, \pi, \tilde{\boldsymbol{P}}\right] \\
&\quad = \boldsymbol{q}^\top \boldsymbol{X}^\pi \boldsymbol{r} + \alpha \boldsymbol{q}^\top \boldsymbol{X}^\pi \boldsymbol{\Pi} \mathbb{E}_{\mathcal{G}}[\Delta\tilde{\boldsymbol{P}}]\boldsymbol{X}^\pi \boldsymbol{r} + \alpha^2 \boldsymbol{q}^\top \boldsymbol{X}^\pi \boldsymbol{\Pi} \mathbb{E}_{\mathcal{G}}[\Delta\tilde{\boldsymbol{P}}\boldsymbol{X}^\pi \boldsymbol{\Pi}\Delta\tilde{\boldsymbol{P}}]\boldsymbol{X}^\pi \boldsymbol{r} + L_{\text{exp}} \\
&\quad \approx \boldsymbol{q}^\top \boldsymbol{X}^\pi \boldsymbol{r} + \alpha^2 \boldsymbol{q}^\top \boldsymbol{X}^\pi \boldsymbol{\Pi} \boldsymbol{Q} \boldsymbol{X}^\pi \boldsymbol{r} \; ,
\end{aligned}$$

where $L_{\text{exp}} = \sum_{k=3}^{\infty} \alpha^k \boldsymbol{q}^\top \mathbb{E}_{\mathcal{G}}\left[(\boldsymbol{X}^\pi \boldsymbol{\Pi}\Delta\tilde{\boldsymbol{P}})^k\right] \boldsymbol{X}^\pi \boldsymbol{r}$, and where $\boldsymbol{Q} \in \mathbb{R}^{|S|\times|A|\times|S|}$, such that

$$\begin{aligned}
Q_{(i,a,j)} &= \left(\mathbb{E}_{\mathcal{G}}[(\Delta\tilde{\boldsymbol{P}}\boldsymbol{X}^\pi \boldsymbol{\Pi}\Delta\tilde{\boldsymbol{P}}]\right)_{(i,a,j)} \\
&= \sum_{k,l,a'} (\boldsymbol{X}^\pi \boldsymbol{\Pi})_{(k,l,a')} \mathbb{E}_{\mathcal{G}}[\Delta\tilde{P}_{(i,a,k)}\Delta\tilde{P}_{(l,a',j)}] \\
&= \sum_{k} X^\pi_{(k,i)} \pi(i,a) \mathbb{E}_{\mathcal{G}}[\Delta\tilde{P}_{(i,a,k)}\Delta\tilde{P}_{(i,a,j)}] \\
&= \pi(i,a) \sum_k \Sigma^{(i,a)}_{(j,k)} X^\pi_{(k,i)} \; .
\end{aligned}$$

This is under the assumption that, for each action a, the rows of $\tilde{\boldsymbol{P}}_a$ are independent from each other and using $\boldsymbol{\Sigma}^{(i,a)}$ to represent the covariance between the terms of the transition probability vector from state i with action a. We are now interested in the second-order approximation of $\mathbb{E}_{\mathcal{G}}[\sum_{t=0}^{\infty} \alpha^t r(x_t) | x_0 \propto \boldsymbol{q}, \pi]$.

Definition 5.4.6 : $\mathbb{F}(\pi)$ is the second-order approximation of the expected return under transition probability uncertainty, such that

$$\mathbb{F}(\pi) = \boldsymbol{q}^\top \boldsymbol{X}^\pi \boldsymbol{r} + \alpha^2 \boldsymbol{q}^\top \boldsymbol{X}^\pi \boldsymbol{\Pi} \boldsymbol{Q} \boldsymbol{X}^\pi \boldsymbol{r} \; .$$

Remark 5.4.7 : One should note that the approximation $\mathbb{F}(\pi)$ depends on the first two moments of random matrix $\tilde{\boldsymbol{P}}$. It can therefore efficiently be evaluated for any policy. Although $\mathbb{F}(\pi)$ is still non-concave in π, in practice, global optimization techniques will lead to useful solutions as presented in Section 5.5.2.

Before studying the usefulness of maximizing $\mathbb{F}(\pi)$, we first introduce the definition of $(1-\delta)$-percentile performance for a policy in this context and present a lemma that constrains the range of possible solutions for any percentile optimization problem.

Definition 5.4.8 : For a fixed policy π, the $(1-\delta)$-percentile performance of policy π under transition probability uncertainty $\tilde{\boldsymbol{P}}$, $\mathcal{Y}(\pi,\delta)$, is the solution to:

$$\begin{array}{rcll} \mathcal{Y}(\pi,\delta) & = & \underset{y\in\mathbb{R}}{\text{maximize}} & y \\ & & \text{subject to} & \mathbb{P}_{\mathcal{G}}\left(\mathbb{E}_{\tilde{\boldsymbol{P}}}[\sum_{t=0}^{\infty}\alpha^t r(x_t)|x_0 \propto \boldsymbol{q},\pi] \geq y\right) \geq 1-\delta \ , \end{array}$$

where the probability $\mathbb{P}_{\mathcal{G}}(\cdot)$ is the probability of drawing $\tilde{\boldsymbol{P}}$ from the posterior Dirichlet distribution, and where the expectation is taken with respect to the random trajectory of x_t when following stationary policy π and given a concrete realization of $\tilde{\boldsymbol{P}}$.

Lemma 5.4.9: *Given any random variable $\tilde{z}$ with mean μ and standard deviation σ, then the optimal value y^* of the optimization problem*

$$\underset{y\in\mathbb{R}}{\text{maximize}} \quad y \tag{5.9a}$$

$$\text{subject to} \quad \mathbb{P}(\tilde{z} \geq y) \geq 1-\delta \ , \tag{5.9b}$$

is assured to be in the range $y^* \in [\mu - \frac{\sigma}{\sqrt{\delta}}, \mu + \frac{\sigma}{\sqrt{1-\delta}}]$.

The proof is given in Appendix B.3. One can now derive the following theorem.

Theorem 5.4.10: *Suppose that we are given the state transition observations $\{(s_1,a_1,s'_1),(s_2,a_2,s'_2),...,(s_M,a_M,s'_M)\}$, and let the minimum number of transitions observed from any state using any action be $M^* := \min_{i,a}\sum_j M_j^{(i,a)}$, and $\delta \in (0,0.5]$. Then, a policy*

$$\hat{\pi} \in \underset{\pi}{\text{argmax}} \ \ \mathbb{F}(\pi) \tag{5.10}$$

is $O(1/\sqrt{\delta M^*})$ optimal with respect to the Percentile MDP problem

$$\underset{\pi\in\Upsilon}{\text{maximize}} \quad \mathcal{Y}(\pi,\delta) \ . \tag{5.11}$$

Proof: Using Lemma 5.4.9 with $\tilde{z}$ replaced by $g(\pi, \boldsymbol{P}) = \mathbb{E}_{\boldsymbol{P}}[\sum_{t=0}^{\infty} \alpha^t r(x_t) | x_0 \propto \boldsymbol{q}, \pi]$, one can easily show that for any policy π

$$\begin{aligned}
\mathcal{Y}(\pi, \delta) - \mathbb{F}(\pi) &\leq \mathbb{E}_{\mathcal{G}}[g(\pi, \tilde{\boldsymbol{P}})] + \frac{1}{\sqrt{1-\delta}} \sqrt{\mathbb{E}_{\mathcal{G}}\left[\left(g(\pi, \tilde{\boldsymbol{P}}) - \mathbb{E}_{\mathcal{G}}[g(\pi, \tilde{\boldsymbol{P}})]\right)^2\right]} - \mathbb{F}(\pi) \\
&= L_{\text{exp}}(\pi) + \sqrt{\frac{L_{\text{var}}(\pi)}{1-\delta}} \, ,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{Y}(\pi, \delta) - \mathbb{F}(\pi) &\geq \mathbb{E}_{\mathcal{G}}[g(\pi, \tilde{\boldsymbol{P}})] - \frac{1}{\sqrt{\delta}} \sqrt{\mathbb{E}_{\mathcal{G}}\left[\left(g(\pi, \tilde{\boldsymbol{P}}) - \mathbb{E}_{\mathcal{G}}[g(\pi, \tilde{\boldsymbol{P}})]\right)^2\right]} - \mathbb{F}(\pi) \\
&= L_{\text{exp}}(\pi) - \sqrt{\frac{L_{\text{var}}(\pi)}{\delta}} \, ,
\end{aligned}$$

where

$$\begin{aligned}
L_{\text{exp}}(\pi) &= \sum_{k=3}^{\infty} \alpha^k \boldsymbol{q}^{\mathsf{T}} \mathbb{E}_{\mathcal{G}}\left[(\boldsymbol{X}^{\pi} \boldsymbol{\Pi} \Delta \tilde{\boldsymbol{P}})^k\right] \boldsymbol{X}^{\pi} \boldsymbol{r} = O\left(\frac{1}{(M^*)^2}\right) \\
L_{\text{var}}(\pi) &= \mathbb{E}_{\mathcal{G}}\left[\left(g(\pi, \tilde{\boldsymbol{P}}) - \mathbb{E}_{\mathcal{G}}[g(\pi, \tilde{\boldsymbol{P}})]\right)^2\right] \\
&= \sum_{k \geq 1, l \geq 1} \mathbb{E}_{\mathcal{G}}\left[\left(g_k(\pi, \tilde{\boldsymbol{P}}) - \mathbb{E}_{\mathcal{G}}[g_k(\pi, \tilde{\boldsymbol{P}})]\right)\left(g_l(\pi, \tilde{\boldsymbol{P}}) - \mathbb{E}_{\mathcal{G}}[g_l(\pi, \tilde{\boldsymbol{P}})]\right)\right] \\
&= \sum_{k+l \geq 2}^{\infty} \mathbb{E}_{\mathcal{G}}\left[g_k(\pi, \tilde{\boldsymbol{P}}) g_l(\pi, \tilde{\boldsymbol{P}})\right] - \sum_{k \geq 2, l \geq 2}^{\infty} \mathbb{E}_{\mathcal{G}}\left[g_k(\pi, \tilde{\boldsymbol{P}})\right] \mathbb{E}_{\mathcal{G}}\left[g_l(\pi, \tilde{\boldsymbol{P}})\right] \\
&= O\left(\frac{1}{M^*}\right) \, ,
\end{aligned}$$

where $g_k(\pi, \tilde{\boldsymbol{P}}) = \alpha^k \boldsymbol{q}^{\mathsf{T}} (\boldsymbol{X}^{\pi} \boldsymbol{\Pi} \Delta \tilde{\boldsymbol{P}})^k \boldsymbol{X}^{\pi} \boldsymbol{r}$. The bounds $O(\frac{1}{(M^*)^2})$ and $O(\frac{1}{M^*})$ were derived from the rate of decay for each moment of a Dirichlet distribution (see [63] for details on these moments).

This gives us a bound between the optimal $(1-\delta)$-percentile performance obtained from policy $\pi^* = \mathrm{argmax}_\pi \, \mathcal{Y}_{\tilde{\boldsymbol{P}}}(\pi, \delta)$ and $\hat{\pi}$ returned by Problem (5.10).

$$\begin{aligned}
\mathcal{Y}_{\tilde{\boldsymbol{P}}}(\pi^*, \delta) - \mathcal{Y}_{\tilde{\boldsymbol{P}}}(\hat{\pi}, \delta) &= \mathcal{Y}_{\tilde{\boldsymbol{P}}}(\pi^*, \delta) - \mathbb{F}(\pi^*) + \mathbb{F}(\pi^*) - \mathcal{Y}_{\tilde{\boldsymbol{P}}}(\hat{\pi}, \delta) \\
&\leq \mathcal{Y}_{\tilde{\boldsymbol{P}}}(\pi^*, \delta) - \mathbb{F}(\pi^*) + \mathbb{F}(\hat{\pi}) - \mathcal{Y}_{\tilde{\boldsymbol{P}}}(\hat{\pi}, \delta) \\
&\leq L_{\text{exp}}(\pi^*) + \frac{\sqrt{L_{\text{var}}(\pi^*)}}{\sqrt{1-\delta}} - L_{\text{exp}}(\hat{\pi}) + \frac{\sqrt{L_{\text{var}}(\hat{\pi})}}{\sqrt{\delta}} \\
&= O\left(\frac{1}{\sqrt{\delta M^*}}\right) . \quad \square
\end{aligned}$$

Next, we show how the percentile criterion studied in this chapter outperforms the nominal and robust criterion on instances of the machine replacement problem with either reward or transition probability uncertainty.

5.5 Experiments on Machine Replacement Problems

Let us consider a very practical instance of an MDP with inherent model uncertainty. A factory owner wants to design a replacement policy for a line of machines. This problem is known to be well modeled with an MDP with states representing reachable aging phases and actions describing different repair or replacement alternatives. Although the parameters used in such a model can typically be estimated from historical data (experienced repair costs and decreases in production due to failures), one can rarely fully resolve them. For example, there is inherent uncertainty in future fluctuations for the cost of new equipment. Moreover, one often doesn't have access to enough historical data to adequately assess the probability of a machine breaking down at a given aging stage. There is therefore a need for incorporating this uncertainty in the performance evaluation of any given repair policy.

Consider more specifically the repair cost that is incurred by a factory that holds a high number of machines, given that each of these machines are modeled with the same underlying MDP. Since the parameters of this MDP are not known with certainty, it is natural in this setting to apply a repair policy uniformly on all the machines with the hope that, with probability higher than $1-\delta$, this policy will have a low average maintenance cost on the set of machines. This is specifically what the percentile criterion presented earlier

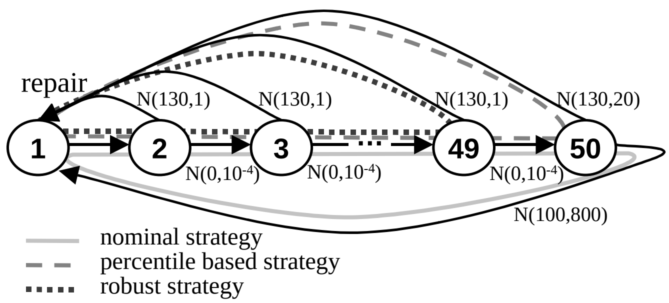


Figure 5.2: Instance of a machine replacement problem with uncertainty in the rewards. The optimal paths followed for three strategies are drawn.

quantifies. We now present two instances of this problem with either reward or transition probability uncertainty. Note that we have selected simple instances of this problem in order to compare the three approaches that were presented in this Bayesian framework: first the Nominal MDP as estimated from the samples, second the Robust MDP which constructs a DRSP model from the posterior distribution, and finally the Percentile MDP.

5.5.1 Machine replacement as an MDP with Gaussian rewards

We study a simple version of the machine replacement problem with 50 states, 2 actions, deterministic transitions, a discount factor of 0.8, and Gaussian uncertainty in the rewards (see Figure 5.2). Our model develops as follow: after the policy is chosen by the agent, the environment is created according to a predefined joint Gaussian distribution over the rewards, and the policy is applied on this environment which is solely deterministic thereafter. For each of the first 49 steps, repairs have a cost independently distributed as $\mathcal{N}(130, 1)$. The 50th state of the machine's life was designed to be a more risky state: not repairing incurs a highly uncertain cost $\mathcal{N}(100, 800)$, while repairing is a more secure but still uncertain option $\mathcal{N}(130, 20)$.

The performance of policies obtained using nominal, robust and percentile formulations, with a confidence of 99%, are presented in Figure 5.3.[3] These results describe what

[3]Implementation details: the robust problem was constructed using a 99% confidence ellipsoid of the random cost vector as the uncertainty set. Also, all second-order cone programming was implemented in Matlab using the CVX software [28].

one would typically expect from the three solution concepts. While the nominal strategy, blind to any form of risk, finds no advantage in ever repairing, the robust strategy ends up following a more conservative policy (repairing the machine in state #49 to avoid state #50). On the other hand, the percentile based strategy handles the risk more efficiently by waiting until state #50 to apply a mixed strategy that repairs the machine with 90% probability. This strategy performed better than its robust alternative while preserving small variance in performance over the 10000 different sampled environments.

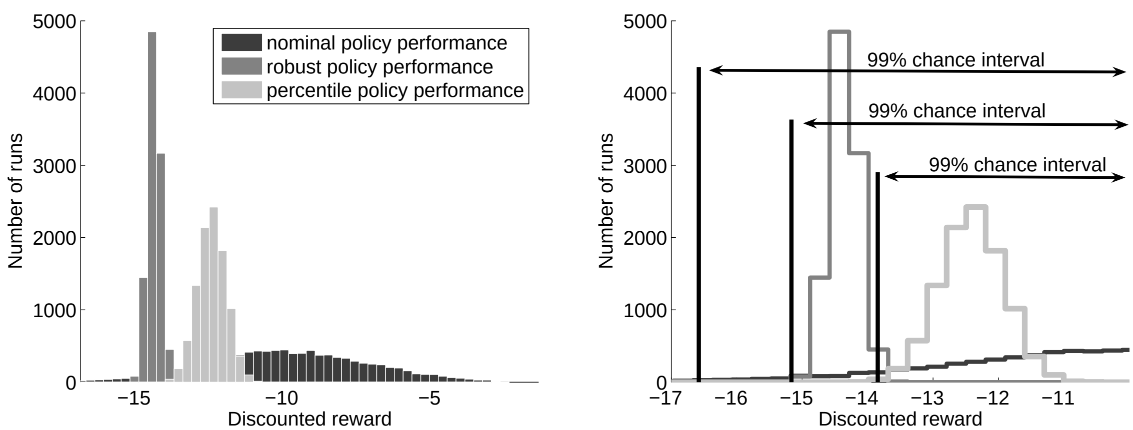


Figure 5.3: Performance comparison between nominal, robust and percentile based policies on 10000 runs of the machine replacement problem. The right figure focuses on the interval $[-17, -10]$.

5.5.2 Machine Replacement as an MDP with Dirichlet prior on transition probabilities

In this experiment, we use a version of the machine replacement problem with 10 states, 4 actions, a discount factor of 0.8, a uniform initial state distribution and transition probability uncertainty modeled with a Dirichlet distribution. States 1 to 8 describe the normal aging of the machine, while states $R1$ and $R2$ represent two possible stages of repairs: $R1$ being normal repairs on the machine with cost of 2, and $R2$ standing for a harder one with a cost of 10. Letting the machine reach the age of 8 is penalized with a cost of 20. In each of these states, one has access to three repair services for the machine. We designed a

Dirichlet model for transitions occurring when no repairs are done. In the case of each of the three repair options, we used slightly perturbed versions of a reference Dirichlet model that is presented in Figure 5.4. In this figure, the expected transition probability parameters are presented given that M transitions were observed from each state. The parameter M acts as a control for the amount of transition probability uncertainty present in the model.

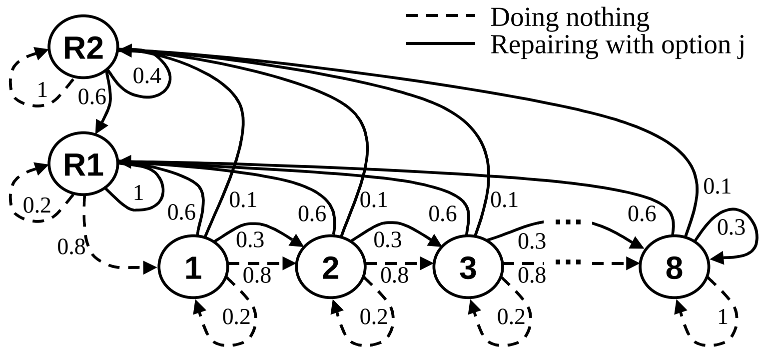
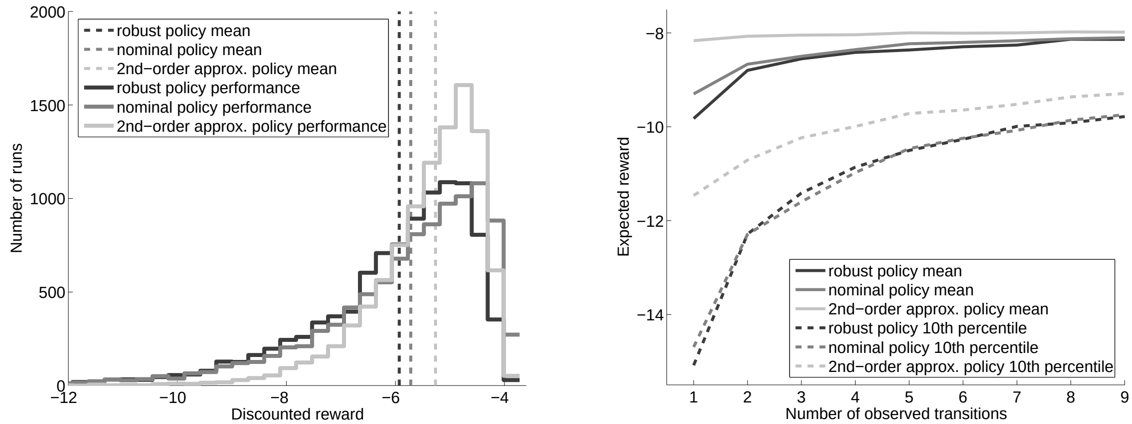


Figure 5.4: Instance of a machine replacement problem with Dirichlet uncertainty in the transition probabilities. The diagram presents the expected transition probabilities for the two types of actions (repairing, or not). The three repair options lead to slightly perturbed versions of the Dirichlet model presented here.

We applied three solution methods to this decision problem. First, the Nominal MDP was formulated using the expected transition probabilities based on the posterior distribution. Since we chose to evaluate the 10th percentile performance of policies, the Robust MDP is constructed using a 90% confidence box in $\mathbb{R}^{|S|\times|A|\times|S|}$ for the random matrix $\tilde{\boldsymbol{P}}$.[4] Finally, we used the "2nd-order approximation" performance measure presented in Section 5.4.2 to find a solution to the Percentile MDP Problem. To do so, we were required to solve a non-convex optimization problem using a gradient descent algorithm (applied on $-\mathbb{F}(\pi)$). Although gradient descent techniques provide no guarantees of reaching a global optimum, by taking as an initial point the policy returned by the Nominal MDP, we were assured to find a policy that performs better than the nominal one with respect to $\mathbb{F}(\pi)$.

[4]Specifically, using 10000 samples drawn from $\mathcal{G}_{\tilde{\boldsymbol{P}}}$ and a given τ ratio, for each parameter $P_{(i,a,j)}$ we chose $A_{(i,a,j)}$ and $B_{(i,a,j)}$ so that they included a ratio of τ of the random samples. A search over τ was done to find the minimal τ that led to a box $A_{(i,a,j)} \leq P_{(i,a,j)} \leq B_{(i,a,j)}$ containing 90% of the samples drawn



(a) (b)

Figure 5.5: (a) presents a performance comparison between nominal, robust and percentile based policies on 10000 runs of the machine replacement problem with $M = 1$. (b) presents the effect of decreasing the uncertainty in the transitions probabilities.

Figure 5.5(a) shows the histogram of expected discounted rewards obtained using the different methods on 10000 instances of the described uncertain machine replacement problem (with $M = 1$). We also indicated the mean and the 10th percentile of the different methods. It is interesting to see that although the 2nd-order approximation method and the nominal method do not directly address the percentile criterion, they both outperform the policy obtained using the robust method in terms of 10th percentile performance for large parameter uncertainty. When having a look at the different policies returned by the methods, we realize that the robust policy acts more conservatively by applying repairs earlier. On the other hand, the nominal strategy does not make any use of the fact that 3 repair options are available. The 2nd-order approximation method returns a policy that uses a mixed strategy over the repair options in states $R1$ (*i.e.*, heavy repair state) in order to reduce the variance of transition probabilities and, indirectly, the overall expected cost. In Figure 5.5(b), we show how these results evolve with the number of observed transitions (quantified by M in the Dirichlet model). As expected, when more transitions are observed, the 2nd-order approximation policy slowly converges to the nominal policy, due

from $\tilde{\boldsymbol{P}}$. This approach is similar to the method for constructing the DRSP proposed in Section 5.2.

to the vanishing second term of $\mathbb{F}(\pi)$.

5.6 Conclusion

In this chapter, we presented how distribution uncertainty can be addressed through the use of a Bayesian approach. After describing how to construct a distributional set that accounts for the prior information that is available in this framework, we presented a percentile optimization problem that can potentially lead to better solutions although it presents a heavier computational challenge. In the context of a Markov decision process with parameter uncertainty, we showed that, although the percentile optimization problem reduces, in the special case of an MDP with Gaussian reward uncertainty, to a second-order cone program, in general the model is intractable to solve and requires the use of approximation algorithms. Our experiments also demonstrated that, given a preferred level of risk, the percentile criterion performs better than policies obtained using an estimate of the distribution or obtained using a distributionally robust model. Overall, we consider that research on new approximation methods for percentile optimization problems has the potential to lead to solutions that improve significantly on the DRSP approach when addressing Bayesian data-driven problems. Yet, because the DRSP is less sensitive to the prior and can be solved more easily, we believe it to be a very effective tool to use in practice.

Chapter 6

Conclusions

This thesis developed new foundations for a quantitative methodology that accounts for distribution uncertainty in a wide range of decision problems where the information about the model is extracted from historical data. We initially considered a frequentist point of view and showed how confidence regions for the first two moments of a random vector could be derived simply using the historical samples and some knowledge about the support of the distribution. These results allowed us to define a set of distributions that is guaranteed with high probability to contain the actual distribution from which the samples were drawn. We then analyzed the computational difficulties related to solving the distributionally robust stochastic program with such a distributional set. Our conclusions were that as long as the objective is convex in the decision variables and "piecewise concave" in the parameters, then one can find a solution to this model in polynomial time. We even demonstrated that, although the problem is NP-hard in general, there is a rich family of distributional sets for which the DRSP model is tractable. We finally showed how to construct the DRSP after committing to a Bayesian prior. In this context, we emphasized how the shape of the distributional set in a DRSP could influence the quality of robust solution. The percentile optimization problem was designed in order to choose this set optimally yet incurred an important price in terms of computational tractability. The study of Markov decision processes illustrated this trade-off quite vividly.

The contribution of this thesis to the study of the complexity and application of the DRSP is significant because, after a number of negative results in this field, the DRSP had commonly been considered to present too much of a computational challenge in practice. Instead, our results lead us to believe that in many large data-driven problems the distributionally robust framework is actually more practical to use than a stochastic program since it provides the decision maker with risk-sensitive solutions that are more reliable and tractable. Many applications have already benefited from our work including a machine replacement problem, a fleet mix optimization problem and a portfolio selection problem. Interestingly enough, the empirical experiments on portfolio selection clearly indicated that an approach which assumes to know exactly the distribution or some of the moments of the random parameters, like the future returns of assets, is bound to be deceptive in the long run. In light of these observations, we strongly believe that a distributionally robust approach would have been better prepared for the sudden change in valuation of securitized mortgages which caused the global financial crisis of 2008-2009. Our hope is that methods such as the one presented in this thesis will soon become common practice in the field of decision making under uncertainty.

Appendix A

Extended Proofs for Chapter 3

A.1 Proof of Lemma 3.2.2

We first establish the primal-dual relationship between Problem (3.4) and Problem (3.5). In a second step, we demonstrate that the conditions for strong duality to hold are met.

Step 1 : One can first show by formulating the Lagrangian of Problem (3.3) that the dual can take the following form

$$\begin{aligned}
\underset{r,\boldsymbol{Q},\boldsymbol{P},\boldsymbol{p},s}{\text{minimize}} \quad & (\gamma_2\boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0\boldsymbol{\mu}_0^\mathsf{T}) \bullet \boldsymbol{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \boldsymbol{P}) - 2\boldsymbol{\mu}_0^\mathsf{T}\boldsymbol{p} + \gamma_1 s && \text{(A.1a)}\\
\text{subject to} \quad & \boldsymbol{\xi}^\mathsf{T}\boldsymbol{Q}\boldsymbol{\xi} - 2\boldsymbol{\xi}^\mathsf{T}(\boldsymbol{p} + \boldsymbol{Q}\boldsymbol{\mu}_0) + r - h(\boldsymbol{x},\boldsymbol{\xi}) \geq 0 \ , \ \forall\, \boldsymbol{\xi} \in \mathcal{S} && \text{(A.1b)}\\
& \boldsymbol{Q} \succeq 0 && \text{(A.1c)}\\
& \begin{bmatrix} \boldsymbol{P} & \boldsymbol{p} \\ \boldsymbol{p}^\mathsf{T} & s \end{bmatrix} \succeq 0 \ , && \text{(A.1d)}
\end{aligned}$$

where $r \in \mathbb{R}$, $\boldsymbol{Q} \in \mathbb{R}^{m\times m}$ are the dual variables for Constraint (3.4b) and Constraint (3.4c) respectively, while $\boldsymbol{P} \in \mathbb{R}^{m\times m}$, $\boldsymbol{p} \in \mathbb{R}^m$ and $s \in \mathbb{R}$ form together a matrix which is the dual variable associated with Constraint (3.4d).

We can further simplify this dual problem by solving analytically for the variables $(\boldsymbol{P}, \boldsymbol{p}, s)$, while keeping $(\boldsymbol{Q}, r)$ fixed. Because of Constraint (A.1d), we can consider two cases for the variable s^*: either $s^* = 0$ or $s^* > 0$. Assuming that $s^* = 0$, then it must be that $\boldsymbol{p}^* = 0$ otherwise ${\boldsymbol{p}^*}^\mathsf{T}\boldsymbol{p}^* > 0$ and

$$\begin{bmatrix} \boldsymbol{p}^* \\ y \end{bmatrix}^\mathsf{T} \begin{bmatrix} \boldsymbol{P}^* & \boldsymbol{p}^* \\ {\boldsymbol{p}^*}^\mathsf{T} & s^* \end{bmatrix} \begin{bmatrix} \boldsymbol{p}^* \\ y \end{bmatrix} = {\boldsymbol{p}^*}^\mathsf{T}\boldsymbol{P}^*\boldsymbol{p}^* - 2{\boldsymbol{p}^*}^\mathsf{T}\boldsymbol{p}^* y < 0 \ , \text{ for } y > \frac{{\boldsymbol{p}^*}^\mathsf{T}\boldsymbol{P}^*\boldsymbol{p}^*}{2{\boldsymbol{p}^*}^\mathsf{T}\boldsymbol{p}^*} \ ,$$

which contradicts Constraint (A.1d). Similarly, $\boldsymbol{P}^* = 0$ is an optimal solution since it minimizes the objective. We conclude that if $s^* = 0$ then, after replacing $\boldsymbol{q} = -2(\boldsymbol{p} + \boldsymbol{Q}\boldsymbol{\mu}_0)$, Problem (A.1)'s objective does indeed reduce to

$$\gamma_2(\boldsymbol{\Sigma}_0 \bullet \boldsymbol{Q}) - \boldsymbol{\mu}_0^\mathsf{T}\boldsymbol{Q}\boldsymbol{\mu}_0 + r \;=\; r + \gamma_2(\boldsymbol{\Sigma}_0 \bullet \boldsymbol{Q}) + \boldsymbol{\mu}_0^\mathsf{T}\boldsymbol{Q}\boldsymbol{\mu}_0 + \boldsymbol{\mu}_0^\mathsf{T}\boldsymbol{q} + \sqrt{\gamma_1}\|\boldsymbol{\Sigma}_0^{1/2}(\boldsymbol{q} + 2\boldsymbol{Q}\boldsymbol{\mu}_0)\| \ .$$

If instead one assumes that $s^* > 0$, then using Schur's complement, Constraint (A.1d) can be shown equivalent to $\boldsymbol{P} \succeq \frac{1}{s}\boldsymbol{p}\boldsymbol{p}^\mathsf{T}$. Since $\boldsymbol{\Sigma}_0 \succeq 0$, $\boldsymbol{P}^* = \frac{1}{s}\boldsymbol{p}\boldsymbol{p}^\mathsf{T}$ is a valid optimal solution and can be replaced in the objective. It remains to solve for $s^* > 0$, which reduces to solving the one dimensional convex optimization problem $\text{minimize}_{s>0} \frac{1}{s}\boldsymbol{p}^\mathsf{T}\boldsymbol{\Sigma}_0\boldsymbol{p} + \gamma_1 s$. By setting the derivative of the objective function to zero, we obtain that $s^* = \sqrt{\frac{1}{\gamma_1}\boldsymbol{p}^\mathsf{T}\boldsymbol{\Sigma}_0\boldsymbol{p}}$. Thus, once again, after replacing $\boldsymbol{q} = -2(\boldsymbol{p} + \boldsymbol{Q}\boldsymbol{\mu}_0)$, the optimal value of Problem (A.1) reduces to the form of Problem (3.5):

$$r + \gamma_2(\boldsymbol{\Sigma}_0 \bullet \boldsymbol{Q}) + \boldsymbol{\mu}_0^\mathsf{T}\boldsymbol{Q}\boldsymbol{\mu}_0 + \boldsymbol{\mu}_0^\mathsf{T}\boldsymbol{q} + \sqrt{\gamma_1}\|\boldsymbol{\Sigma}_0^{1/2}(\boldsymbol{q} + 2\boldsymbol{Q}\boldsymbol{\mu}_0)\| \ .$$

Step 2 : One can easily show that the conditions on γ_1, γ_2 and $\boldsymbol{\Sigma}_0$ are sufficient to ensure that the Dirac measure $\delta_{\boldsymbol{\mu}_0}$ lies in the relative interior of the feasible set of Problem (3.3). Based on the weaker version of Proposition 3.4 in [53], we can conclude that there is no duality gap between the two problems and that if $\Psi(\boldsymbol{x}; \gamma_1, \gamma_2)$ is finite then the set of optimal solutions to Problem (3.5) must be non-empty. □

A.2 Proof of Proposition 3.3.3 with Weak Feasibility Oracles

In this section, we discuss an extension of the results presented in Proposition 3.3.3 to problems where the feasibility oracles can only verify that an instance is close enough from the interior of the feasible set and can only generate approximate cutting planes. We also use this as an opportunity to expose a more rigorous and detailed proof of polynomial solvability of the DRSP with $\mathcal{D}_1$. In what follows, we will refer to a notion of distance from a set as defined next.

Definition A.2.1: Given a set $\mathcal{Z} \subseteq \mathbb{R}^m$, for any $\boldsymbol{z} \in \mathbb{R}^m$ let $\mathbf{dist}(\boldsymbol{z}, \mathcal{Z}) = \inf_{\bar{\boldsymbol{z}} \in \mathcal{Z}} \|\boldsymbol{z} - \bar{\boldsymbol{z}}\|$.

We first make the concept of weak feasibility oracles more precise through two weaker assumptions about the set $\mathcal{S}$ and $\mathcal{X}$.

Assumption A.2.2: The support set $\mathcal{S} \subset \mathbb{R}^m$ is convex and compact (closed and bounded), contains $\boldsymbol{\mu}_0$, and is equipped with an oracle that can for any $\boldsymbol{\xi} \in \mathbb{R}^m$ and $\epsilon > 0$ either confirm that $\mathbf{dist}(\boldsymbol{\xi}, \mathcal{S}) \leq \epsilon$ or provide a vector $\boldsymbol{a}$ such that $\|\boldsymbol{a}\| \geq 1$ and that, for any $\boldsymbol{\xi}' \in \mathcal{S}$, $\boldsymbol{a}^\mathsf{T}\boldsymbol{\xi}' \leq \boldsymbol{a}^\mathsf{T}\boldsymbol{\xi} + \epsilon$ in time polynomial in m and $\log(1/\epsilon)$. Furthermore, there exist two balls centered at some $\boldsymbol{\xi}_0$ and of radius $r_{\boldsymbol{\xi}}$ and $R_{\boldsymbol{\xi}}$ respectively such that $\text{Ball}(\boldsymbol{\xi}_0, r_{\boldsymbol{\xi}}) \subseteq \mathcal{S} \subseteq \text{Ball}(\boldsymbol{\xi}_0, R_{\boldsymbol{\xi}})$.

Assumption A.2.3: The set $\mathcal{X} \subset \mathbb{R}^n$ is convex and compact (closed and bounded), and it is equipped with an oracle that can for any $\boldsymbol{x} \in \mathbb{R}^n$ and $\epsilon > 0$ either confirm that $\mathbf{dist}(\boldsymbol{x}, \mathcal{X}) \leq \epsilon$ or provide a vector $\boldsymbol{a}$ such that $\|\boldsymbol{a}\| \geq 1$ and that for any $\boldsymbol{x}' \in \mathcal{X}$, $\boldsymbol{a}^\mathsf{T}\boldsymbol{x}' \leq \boldsymbol{a}^\mathsf{T}\boldsymbol{x} + \epsilon$ in time polynomial in n and $\log(1/\epsilon)$. Furthermore, there exist two balls centered at some $\boldsymbol{x}_0$ and of radius $r_{\boldsymbol{x}}$ and $R_{\boldsymbol{x}}$ respectively such that $\text{Ball}(\boldsymbol{x}_0, r_{\boldsymbol{x}}) \subseteq \mathcal{X} \subseteq \text{Ball}(\boldsymbol{x}_0, R_{\boldsymbol{x}})$.

Given that the oracles associated with $\mathcal{X}$ and $\mathcal{S}$ satisfy the weaker assumptions above, the solvability of the DRSP with $\mathcal{D}_1$ can be described as follows.

Proposition A.2.4: *Given that assumptions 3.2.6, 3.3.2, A.2.2, and A.2.3 hold, and that $h(\boldsymbol{x}, \boldsymbol{\xi})$ is Lipschitz continuous in $\boldsymbol{\xi}$, then for any $\epsilon > 0$, one can find in time polynomial in $\log(1/\epsilon)$ and the sizes of $\boldsymbol{x}$ and $\boldsymbol{\xi}$ an assignment $\hat{\boldsymbol{x}}$ such that $\mathbf{dist}(\hat{\boldsymbol{x}}, \mathcal{X}) \leq \epsilon$ and $\hat{\boldsymbol{x}}$ is ϵ-optimal with respect to the objective of the DRSP with $\mathcal{D}_1$, also known as Problem* (3.7).

Before providing a detailed proof of Proposition A.2.4, we first need to present the computational equivalence between optimization and weak separation. We then present a short lemma which will help us provide a weak feasibility oracle for Constraint (3.8b), which involves an internal maximization that can only be done approximately.

Lemma A.2.5 : *(Grötschel* et al. *[29]) Consider a convex optimization problem of the form*

$$\underset{\boldsymbol{z}\in\mathcal{Z}}{\text{minimize}} \qquad \boldsymbol{c}^\mathsf{T}\boldsymbol{z}$$

with linear objective and convex feasible set $\mathcal{Z}$. Given that there exist two balls centered at some $\boldsymbol{z}_0$ and of radius r and R respectively such that Ball$(\boldsymbol{z}_0, r) \subseteq \mathcal{Z} \subseteq$ Ball$(\boldsymbol{z}_0, R)$, for any $\epsilon > 0$ one can find in time polynomial in $\log(1/\epsilon)$ and in the size of the problem an assignment $\hat{\boldsymbol{z}}$ such that $\mathbf{dist}(\hat{\boldsymbol{z}}, \mathcal{Z}) \leq \epsilon$ and $\boldsymbol{c}^\mathsf{T}\boldsymbol{z} \geq \boldsymbol{c}^\mathsf{T}\hat{\boldsymbol{z}} - \epsilon$ for all $\boldsymbol{z} \in \mathcal{Z}$ if and only if for any assignment $\hat{\boldsymbol{z}}$ one can in time polynomial in the dimension of $\boldsymbol{z}$ either assert that $\mathbf{dist}(\hat{\boldsymbol{z}}, \mathcal{Z}) \leq \epsilon$ or find a vector $\boldsymbol{a}$ such that $\|\boldsymbol{a}\| \geq 1$ and that for any $\boldsymbol{z} \in \mathcal{Z}$, $\boldsymbol{a}^\mathsf{T}\boldsymbol{z} \leq \boldsymbol{a}^\mathsf{T}\hat{\boldsymbol{z}} + \epsilon$.

The following lemma relies on a first application of Grötschel *et al.*'s equivalence result in order to account for the weak feasibility oracle associated to $\mathcal{S}$.

Lemma A.2.6 : *Let function $h(\boldsymbol{x}, \boldsymbol{\xi})$ be concave in $\boldsymbol{\xi}$ and be such that one can evaluate $h(\boldsymbol{x}, \boldsymbol{\xi})$ and provide a super-gradient of $\boldsymbol{\xi}$ in time polynomial in m. Then, under Assumption A.2.2, for any fixed assignment $\boldsymbol{x}$, $\boldsymbol{Q} \succeq 0$, $\boldsymbol{q}$, and r, and for any $\epsilon > 0$ one can find in time polynomial in $\log(1/\epsilon)$ and the size of $\boldsymbol{\xi}$ an assignment $\hat{\boldsymbol{\xi}}$ such that $\mathbf{dist}(\hat{\boldsymbol{\xi}}, \mathcal{S}) \leq \epsilon$ and*

$$h(\boldsymbol{x}, \hat{\boldsymbol{\xi}}) - \hat{\boldsymbol{\xi}}^\mathsf{T}\boldsymbol{Q}\hat{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^\mathsf{T}\boldsymbol{q} - r \;\geq\; h(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{\xi}^\mathsf{T}\boldsymbol{Q}\boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T}\boldsymbol{q} - r - \epsilon \;\; \forall \boldsymbol{\xi} \in \mathcal{S} \; .$$

Proof: We can apply Lemma A.2.5 to the following optimization problem:

$$\underset{t,\boldsymbol{\xi}}{\text{maximize}} \qquad t \tag{A.2a}$$

$$\text{subject to} \qquad t \leq h(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{\xi}^\mathsf{T}\boldsymbol{Q}\boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T}\boldsymbol{q} - r \tag{A.2b}$$

$$\boldsymbol{\xi} \in \mathcal{S} \; . \tag{A.2c}$$

First, the feasible set of the problem is convex since $\boldsymbol{Q} \succeq 0$ so that $h(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{\xi}^\mathsf{T}\boldsymbol{Q}\boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T}\boldsymbol{q} - r$ is a concave function in $\boldsymbol{\xi}$ and $\mathcal{S}$ is assumed to be a convex set.

Secondly, we assumed to know of a ball of radius $R_{\boldsymbol{\xi}}$ such that $\mathcal{S} \subseteq \text{Ball}(\boldsymbol{\xi}_0, R_{\boldsymbol{\xi}})$. We can therefore verify that

$$\begin{aligned} h(\boldsymbol{x}, \boldsymbol{\xi}_0) - \boldsymbol{\xi}_0^\mathsf{T}\boldsymbol{Q}\boldsymbol{\xi}_0 - \boldsymbol{\xi}_0^\mathsf{T}\boldsymbol{q} - r \ \leq \ t^* \ &= \quad \max_{\boldsymbol{\xi} \in \mathcal{S}} h(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{\xi}^\mathsf{T}\boldsymbol{Q}\boldsymbol{\xi} - \boldsymbol{\xi}^\mathsf{T}\boldsymbol{q} - r \\ &\leq \quad \max_{\boldsymbol{\xi} \in \mathcal{S}} h(\boldsymbol{x}, \boldsymbol{\xi}) + \|\boldsymbol{q}\|(\|\boldsymbol{\xi}_0\| + R_{\boldsymbol{\xi}}) - r \ < \ \infty \ , \end{aligned}$$

by the fact that $\mathcal{S}$ is compact and that $h(\boldsymbol{x}, \boldsymbol{\xi})$ is finite everywhere. Since t^* lies in some interval $[t_{\text{min}}, t_{\text{max}}]$, we can reduce the search over t to $\text{Ball}(t_0, R_t)$ with $t_0 = t_{\text{min}} - 1$ and $R_t = t_{\text{max}} - t_0 + 1$. It then becomes trivial to show that $\mathcal{S} \times \text{Ball}(t_0, R_t)$ is a subset of a ball centered at $(\boldsymbol{\xi}_0, t_0)$ of radius $\sqrt{R_{\boldsymbol{\xi}}^2 + R_t^2}$. Also, since the point $(\boldsymbol{\xi}_0, t_0)$ satisfies all constraints strictly, we can also conclude that it is in the interior of the feasible set and therefore that there exists a small ball centered at $(\boldsymbol{\xi}_0, t_0)$ which is a subset of the feasible set of Problem (A.2).

Thirdly, regarding the weak separation problem, we confirm that for any $\epsilon > 0$ and any assignment $(\bar{\boldsymbol{\xi}}, \bar{t})$ separation from $\mathcal{S}$ is a tractable problem. Furthermore, letting $g(\boldsymbol{\xi}, t) = t - h(\boldsymbol{x}, \boldsymbol{\xi}) + \boldsymbol{\xi}^\mathsf{T}\boldsymbol{Q}\boldsymbol{\xi} + \boldsymbol{\xi}^\mathsf{T}\boldsymbol{q} + r$, we then confirm that $g(\bar{\boldsymbol{\xi}}, \bar{t}) \leq \epsilon$ can be verified directly after the evaluation of $h(\boldsymbol{x}, \bar{\boldsymbol{\xi}})$, and that if $g(\bar{\boldsymbol{\xi}}, \bar{t}) > \epsilon$ then the following separating hyperplane can be generated in polynomial time:

$$(-\nabla_{\boldsymbol{\xi}} h(\boldsymbol{x}, \bar{\boldsymbol{\xi}}) + 2\boldsymbol{Q}\bar{\boldsymbol{\xi}} + \boldsymbol{q})^\mathsf{T}\boldsymbol{\xi} + t \leq (-\nabla_{\boldsymbol{\xi}} h(\boldsymbol{x}, \bar{\boldsymbol{\xi}}) + 2\boldsymbol{Q}\bar{\boldsymbol{\xi}} + \boldsymbol{q})^\mathsf{T}\bar{\boldsymbol{\xi}} + \bar{t} \ ,$$

where $\nabla_{\boldsymbol{\xi}} h(\boldsymbol{x}, \boldsymbol{\xi})$ is a super-gradient of $h(\boldsymbol{x}, \cdot)$. This is a consequence of the convexity of $g(\boldsymbol{\xi}, t)$ in both $\boldsymbol{\xi}$ and t:

$$\nabla_{\boldsymbol{\xi}} g(\boldsymbol{\xi}, t)^\mathsf{T}(\boldsymbol{\xi} - \bar{\boldsymbol{\xi}}) + \nabla_t g(\boldsymbol{\xi}, t)(t - \bar{t}) \leq g(\boldsymbol{\xi}, t) - g(\bar{\boldsymbol{\xi}}, \bar{t}) \leq 0 - \epsilon \leq 0 \ \ ,$$

for all feasible pair $(\boldsymbol{\xi}, t)$. One can also verify that the cutting plane vector always has its norm larger than one since it is equal to one in the space of t.

Finally, by Lemma A.2.5, these properties guarantee us that the ellipsoid method will converge in polynomial time to an assignment $\hat{\boldsymbol{\xi}}$ and $\hat{t}$ such that $\mathbf{dist}(\hat{\boldsymbol{\xi}}, \mathcal{S}) \leq \epsilon/2$ and such that

$$t^* - \epsilon/2 \;\leq\; \hat{t} \;\leq\; h(\boldsymbol{x}, \hat{\boldsymbol{\xi}}) - \hat{\boldsymbol{\xi}}^{\mathsf{T}} \boldsymbol{Q} \hat{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^{\mathsf{T}} \boldsymbol{q} - r + \epsilon/2$$

$$\Rightarrow \;\; h(\boldsymbol{x}, \hat{\boldsymbol{\xi}}) - \hat{\boldsymbol{\xi}}^{\mathsf{T}} \boldsymbol{Q} \hat{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}^{\mathsf{T}} \boldsymbol{q} - r \geq t^* - \epsilon \; .$$

This concludes our proof. □

We are now ready to prove the statement presented in Proposition A.2.4.

Proof: We first reformulate the inner moment problem in its dual form using Lemma 3.2.2 and use the fact that min-min operations can be performed jointly and the fact that the constraint involving $h(\boldsymbol{x}, \boldsymbol{\xi})$ decomposes. This leads to an equivalent convex optimization form for Problem (3.7):

$$\begin{aligned} \underset{\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{q}, r, t}{\text{minimize}} \quad & r + t \\ \text{subject to} \quad & r \geq h_k(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{\xi}^{\mathsf{T}} \boldsymbol{Q} \boldsymbol{\xi} - \boldsymbol{\xi}^{\mathsf{T}} \boldsymbol{q} \;\; , \;\; \forall \, \boldsymbol{\xi} \in \mathcal{S}, \;\; k \in \{1, ..., K\} \\ & t \geq \left(\gamma_2 \boldsymbol{\Sigma}_0 + \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^{\mathsf{T}} \right) \bullet \boldsymbol{Q} + \boldsymbol{\mu}_0^{\mathsf{T}} \boldsymbol{q} + \sqrt{\gamma_1} \, \| \boldsymbol{\Sigma}_0^{1/2} (\boldsymbol{q} + 2 \boldsymbol{Q} \boldsymbol{\mu}_0) \| \\ & \boldsymbol{Q} \succeq 0 \\ & \boldsymbol{x} \in \mathcal{X} \;\; . \end{aligned}$$

After further algebraic manipulation the problem can take the simpler form:

$$\underset{\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{q}, r, t}{\text{minimize}} \quad r + t \tag{A.4a}$$

$$\text{subject to} \quad r \geq h_k(\boldsymbol{x}, \boldsymbol{\xi}) - (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^{\mathsf{T}} \boldsymbol{Q} (\boldsymbol{\xi} - \boldsymbol{\mu}_0) - (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^{\mathsf{T}} \boldsymbol{q} \; , \; \forall \, \boldsymbol{\xi} \in \mathcal{S} \; , \; \forall k \tag{A.4b}$$

$$t \geq (\gamma_2 \boldsymbol{\Sigma}_0) \bullet \boldsymbol{Q} + \sqrt{\gamma_1} \, \| \boldsymbol{\Sigma}_0^{1/2} \boldsymbol{q} \| \tag{A.4c}$$

$$\boldsymbol{Q} \succeq 0 \tag{A.4d}$$

$$\boldsymbol{x} \in \mathcal{X} \;\; . \tag{A.4e}$$

It remains to show that the ellipsoid method can be successfully applied.

First, we can easily confirm that in Problem (A.4) the objective is linear and that the feasible set is convex: Constraint (A.4b) is convex in $\boldsymbol{x}$ since it is the sum of a convex function $h_k(\boldsymbol{x}, \boldsymbol{\xi})$ and of a linear function, Constraint (A.4c) is a linear constraint, and finally both the positive semidefinite cone and the set $\mathcal{X}$ are known to be convex sets. This feasible set is also non-empty since the assignment $\bar{\boldsymbol{x}} = \boldsymbol{x}_0$, $\bar{\boldsymbol{Q}} = 0$, $\bar{\boldsymbol{q}} = 0$, $\bar{t} = 0$, $\bar{r} = \sup_{\boldsymbol{\xi}\in\mathcal{S}} \max_{k\in\{1,\ldots,K\}} h_k(\boldsymbol{x},\boldsymbol{\xi})$ is feasible. Based on Lemma 3.2.2, we know that for any fixed $\boldsymbol{x}$ the objective value of Problem (A.4) is bounded below by $h(\boldsymbol{x},\boldsymbol{\mu}_0)$ (see Appendix A.1 for details) and above by $\sup_{\boldsymbol{\xi}\in\mathcal{S}} \max_{k\in\{1,\ldots,K\}} h_k(\boldsymbol{x},\boldsymbol{\xi})$. and can conclude that the optimal value, OPT, is finite; thus, the set of optimal solutions is non-empty.

Secondly, while $\mathcal{X}$ is assumed to be a subset of some ball of radius $R_{\boldsymbol{x}}$ centered at $\boldsymbol{x}_0$, this is not known for $\boldsymbol{Q}$, $\boldsymbol{q}$, r and t. However, we can still construct the sets Ball$(\boldsymbol{Q}_0, R_{\boldsymbol{Q}})$, Ball$(\boldsymbol{q}_0, R_{\boldsymbol{q}})$, Ball$(r_0, R_r)$, and Ball$(t_0, R_t)$ which respectively contain an optimal assignment $\boldsymbol{x}^*$, $\boldsymbol{Q}^*$, $\boldsymbol{q}^*$, r^*, and t^*. Based on our assumption on the sets $\mathcal{S}$ and $\mathcal{X}$, we can use the fact that $\boldsymbol{\mu}_0 \in \mathcal{S}$ and that $\boldsymbol{\Sigma}_0 \succ 0$ and $\boldsymbol{Q}^* \succeq 0$ to make the following statements:

$$
\begin{aligned}
&h(\boldsymbol{x},\boldsymbol{\mu}_0) \leq \text{OPT} \leq h(\boldsymbol{x}_0,\boldsymbol{\xi})\,,\ \forall\, \boldsymbol{x}\in\mathcal{X},\ \boldsymbol{\xi}\in\mathcal{S} \ \Rightarrow\ \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\boldsymbol{\mu}_0) \leq \text{OPT} \leq \max_{\boldsymbol{\xi}\in\mathcal{S}} h(\boldsymbol{x}_0,\boldsymbol{\xi})\\
&r^* \geq \max_k \left\{ h_k(\boldsymbol{x}^*,\boldsymbol{\mu}_0) - (\boldsymbol{\mu}_0-\boldsymbol{\mu}_0)^\mathsf{T}\boldsymbol{Q}^*(\boldsymbol{\mu}_0-\boldsymbol{\mu}_0) - (\boldsymbol{\mu}_0-\boldsymbol{\mu}_0)^\mathsf{T}\boldsymbol{q}^*\right\} \geq \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\boldsymbol{\mu}_0)\\
&t^* \geq (\gamma_2\boldsymbol{\Sigma}_0)\bullet\boldsymbol{Q}^* + \sqrt{\gamma_1}\,\|\boldsymbol{\Sigma}_0^{1/2}\boldsymbol{q}^*\| \geq 0 \quad .
\end{aligned}
$$

These statements imply the following bounds on r^*, t^*, $\boldsymbol{Q}^*$, and $\boldsymbol{q}^*$:

$$
\begin{aligned}
&r^* = \text{OPT} - t^* \leq \max_{\boldsymbol{\xi}\in\mathcal{S}} h(\boldsymbol{x}_0,\boldsymbol{\xi})\\
&t^* = \text{OPT} - r^* \leq \max_{\boldsymbol{\xi}\in\mathcal{S}} h(\boldsymbol{x}_0,\boldsymbol{\xi}) - \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\boldsymbol{\mu}_0)\\
&(\gamma_2\boldsymbol{\Sigma}_0)\bullet\boldsymbol{Q}^* = t^* - \sqrt{\gamma_1}\,\|\boldsymbol{\Sigma}_0^{1/2}\boldsymbol{q}^*\| \leq \max_{\boldsymbol{\xi}\in\mathcal{S}} h(\boldsymbol{x}_0,\boldsymbol{\xi}) - \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\boldsymbol{\mu}_0)\\
&\sqrt{\gamma_1}\,\|\boldsymbol{\Sigma}_0^{1/2}\boldsymbol{q}^*\| = t^* - (\gamma_2\boldsymbol{\Sigma}_0)\bullet\boldsymbol{Q}^* \leq \max_{\boldsymbol{\xi}\in\mathcal{S}} h(\boldsymbol{x}_0,\boldsymbol{\xi}) - \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\boldsymbol{\mu}_0) \quad .
\end{aligned}
$$

Thus, considering $\boldsymbol{Q}_0 = \boldsymbol{I}$ and $\boldsymbol{q}_0 = 0$, we have that

$$
\begin{aligned}
\|\boldsymbol{Q}-\boldsymbol{Q}_0\| &= \|\boldsymbol{Q}^*-\boldsymbol{I}\| \leq \|\boldsymbol{I}\| + \|\boldsymbol{Q}^*\|\\
&\leq R_{\boldsymbol{Q}} := \sqrt{m}\left(1 + \frac{1}{\sigma_{\min}(\hat{\boldsymbol{\Sigma}})}(\max_{\boldsymbol{\xi}\in\mathcal{S}} h(\boldsymbol{x}_0,\boldsymbol{\xi}) - \min_{\boldsymbol{x}\in\mathcal{X}} h(\boldsymbol{x},\boldsymbol{\mu}_0))\right) \quad ,
\end{aligned}
$$

where $\sigma_{\min}(\hat{\boldsymbol{\Sigma}})$ refers to the smallest singular value of $\hat{\boldsymbol{\Sigma}}$. We also have that

$$\|\boldsymbol{q}^* - \boldsymbol{q}_0\| = \|\boldsymbol{q}^*\| \leq R_{\boldsymbol{q}} := \frac{1}{\sqrt{\gamma_1 \sigma_{\min}(\hat{\boldsymbol{\Sigma}})}} (\max_{\boldsymbol{\xi} \in \mathcal{S}} h(\boldsymbol{x}_0, \boldsymbol{\xi}) - \min_{\boldsymbol{x} \in \mathcal{X}} h(\boldsymbol{x}, \boldsymbol{\mu}_0)) \ ,$$

Both $\mathrm{Ball}(\boldsymbol{Q}_0, R_{\boldsymbol{Q}})$ and $\mathrm{Ball}(\boldsymbol{q}_0, R_{\boldsymbol{q}})$ can be added as new constraints to Problem (A.4) without affecting the solution. A similar argument applies for r and t, thus we can intersect the feasible set with the following Cartesian product of sets:

$$\mathrm{Ball}(\boldsymbol{x}_0, R_{\boldsymbol{x}}) \times \mathrm{Ball}(\boldsymbol{Q}_0, R_{\boldsymbol{Q}}) \times \mathrm{Ball}(\boldsymbol{q}_0, R_{\boldsymbol{q}}) \times \mathrm{Ball}(r_0, R_r) \times \mathrm{Ball}(t_0, R_t) \ ,$$

where

$$\begin{aligned}
r_0 &= \max_{\boldsymbol{\xi} \in \mathcal{S}} h(\boldsymbol{x}_0, \boldsymbol{\xi}) - (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\mathsf{T} \boldsymbol{I} (\boldsymbol{\xi} - \boldsymbol{\mu}_0) + 1 \ , \\
R_r &= \max \left\{ \max_{\boldsymbol{\xi} \in \mathcal{S}} h(\boldsymbol{x}_0, \boldsymbol{\xi}) - r_0 \ , \ r_0 - \min_{\boldsymbol{x} \in \mathcal{X}} h(\boldsymbol{x}, \boldsymbol{\mu}_0) \right\} \ , \\
t_0 &= \gamma_2 \, \mathbf{trace}(\boldsymbol{\Sigma}_0) + 1 \ , \\
R_t &= \max \left\{ \max_{\boldsymbol{\xi} \in \mathcal{S}} h(\boldsymbol{x}_0, \boldsymbol{\xi}) - \min_{\boldsymbol{x} \in \mathcal{X}} h(\boldsymbol{x}, \boldsymbol{\mu}_0) - t_0 \ , \ t_0 \right\} \ .
\end{aligned}$$

After the addition of this constraint, we are guaranteed that this new feasible set for Problem (A.4) is a subset of a ball of radius $R = \sqrt{R_{\boldsymbol{x}}^2 + R_{\boldsymbol{Q}}^2 + R_{\boldsymbol{q}}^2 + R_r^2 + R_t^2}$ centered at $(\boldsymbol{x}_0, \boldsymbol{Q}_0, \boldsymbol{q}_0, r_0, t_0)$. Also, since the center of this ball is known to lie in the strict interior of this feasible set, we are guaranteed that this set includes a small ball centered at $(\boldsymbol{x}_0, \boldsymbol{Q}_0, \boldsymbol{q}_0, r_0, t_0)$.

Thirdly, we confirm that the weak separation problem is tractable for any $\epsilon > 0$ and any assignment $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{Q}}, \bar{\boldsymbol{q}}, \bar{r}, \bar{t})$. By Assumption A.2.3, separation of $\boldsymbol{x}$ from $\mathcal{X}$ is known to be tractable. Considering Constraint (A.4d), it is well known that the minimum eigenvalue of a symmetric matrix can be estimated efficiently to any $\epsilon > 0$ precision using singular value decomposition. Thus, one can confirm that $\bar{\boldsymbol{Q}} \succeq -\epsilon$ by verifying that this estimate is greater than zero, if not then the associated eigenvector $\boldsymbol{z}$, normalized to have norm one, must satisfy $\boldsymbol{z}^\mathsf{T} \bar{\boldsymbol{Q}} \boldsymbol{z} < 0$ and one can create a separating hyperplane $(\boldsymbol{z}\boldsymbol{z}^\mathsf{T}) \bullet \boldsymbol{Q} \geq (\boldsymbol{z}\boldsymbol{z}^\mathsf{T}) \bullet \bar{\boldsymbol{Q}}$ for all feasible $\boldsymbol{Q}$ since $\boldsymbol{z}^\mathsf{T} \boldsymbol{Q} \boldsymbol{z} \geq 0$. The task is not more difficult for Constraint (A.4c) since we can evaluate $\bar{t} \geq (\gamma_2\boldsymbol{\Sigma}_0) \bullet \bar{\boldsymbol{Q}} + \sqrt{\gamma_1}\,\|\boldsymbol{\Sigma}_0^{1/2}\bar{\boldsymbol{q}}\|$ directly and if not, then generate a separating hyperplane

$$\gamma_2\boldsymbol{\Sigma}_0 \bullet \boldsymbol{Q} + \nabla_{\boldsymbol{q}} g(\bar{\boldsymbol{q}}))^{\mathsf{T}}\boldsymbol{q} - t \leq \gamma_2\boldsymbol{\Sigma}_0 \bullet \bar{\boldsymbol{Q}} + \nabla_{\boldsymbol{q}} g(\bar{\boldsymbol{q}}))^{\mathsf{T}}\bar{\boldsymbol{q}} - \bar{t} \ ,$$

where $g(\boldsymbol{q}) = \sqrt{\gamma_1}\,\|\boldsymbol{\Sigma}_0^{1/2}\boldsymbol{q}\|$ and where the norm of the cutting plane vector is greater than one since it is already equal to one in the direction of t. Finally, concerning any indexed element of Constraint (A.4b), after verifying that $\bar{\boldsymbol{Q}} \succeq 0$, by Lemma A.2.6 we have that in polynomial time an algorithm can generate an assignment $\hat{\boldsymbol{\xi}}$ such that $\mathbf{dist}(\hat{\boldsymbol{\xi}}, \mathcal{S}) < \epsilon$ and that

$$g_k(\bar{\boldsymbol{x}}, \bar{\boldsymbol{Q}}, \bar{\boldsymbol{q}}, \bar{r}, \hat{\boldsymbol{\xi}}) \ \geq \ \max_{\boldsymbol{\xi}\in\mathcal{S}} g_k(\bar{\boldsymbol{x}}, \bar{\boldsymbol{Q}}, \bar{\boldsymbol{q}}, \bar{r}, \boldsymbol{\xi}) - \epsilon \quad .$$

where $g_k(\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{q}, r, \boldsymbol{\xi}) = h_k(\boldsymbol{x}, \boldsymbol{\xi}) - (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^{\mathsf{T}}\boldsymbol{Q}(\boldsymbol{\xi} - \boldsymbol{\mu}_0) - (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^{\mathsf{T}}\boldsymbol{q} - r$. We can state that the assignment is approximately feasible if $g_k(\bar{\boldsymbol{x}}, \bar{\boldsymbol{Q}}, \bar{\boldsymbol{q}}, \bar{r}, \hat{\boldsymbol{\xi}}) \leq 0$. Otherwise we use the fact that for any assignment $(\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{q}, r, t)$, we have that

$$g_k(\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{q}, r, \hat{\boldsymbol{\xi}}) \leq \max_{\boldsymbol{\xi}\in\mathcal{S}} g_k(\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{q}, r, \boldsymbol{\xi}) + C_k\,\mathbf{dist}(\hat{\boldsymbol{\xi}}, \mathcal{S}) \leq C_k\epsilon \ ,$$

where C_k is the Lipschitz constant of $g_k(\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{q}, r, \cdot)$. Using the convexity of $g_k(\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{q}, r, \boldsymbol{\xi})$ in $(\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{q}, r)$ we can generate the following separating hyperplane:

$$\begin{aligned}\nabla_{\boldsymbol{x}} h_k(\bar{\boldsymbol{x}}, \hat{\boldsymbol{\xi}})^{\mathsf{T}}\boldsymbol{x} - &(\hat{\boldsymbol{\xi}} - \boldsymbol{\mu}_0)(\hat{\boldsymbol{\xi}} - \boldsymbol{\mu}_0)^{\mathsf{T}} \cdot \boldsymbol{Q} - (\hat{\boldsymbol{\xi}} - \boldsymbol{\mu}_0)^{\mathsf{T}}\boldsymbol{q} - r \\ &\leq \nabla_{\boldsymbol{x}} h_k(\bar{\boldsymbol{x}}, \hat{\boldsymbol{\xi}})^{\mathsf{T}}\bar{\boldsymbol{x}} - (\hat{\boldsymbol{\xi}} - \boldsymbol{\mu}_0)(\hat{\boldsymbol{\xi}} - \boldsymbol{\mu}_0)^{\mathsf{T}} \cdot \bar{\boldsymbol{Q}} - (\hat{\boldsymbol{\xi}} - \boldsymbol{\mu}_0)^{\mathsf{T}}\bar{\boldsymbol{q}} - \bar{r} + C_k\epsilon \ ,\end{aligned}$$

where once again the norm of the cutting plane vector is greater than one since it is already of size one in the direction of r.

Finally, by Lemma A.2.5, these properties guarantee us that the ellipsoid method will converge in polynomial time to an assignment $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{Q}}, \hat{\boldsymbol{q}}, \hat{r}, \hat{t})$ such that $\mathbf{dist}(\hat{\boldsymbol{x}}, \mathcal{X}) \leq \epsilon$ and that $\hat{r} + \hat{t} \leq \text{OPT} + \epsilon$. Yet, one still needs to account for constraints (A.4b), (A.4c), and (A.4d) being approximately satisfied when comparing the performance of $\hat{\boldsymbol{x}}$ to the optimal performance with respect to the original DRSP problem. In fact, we can show that $\hat{\boldsymbol{Q}}' = \hat{\boldsymbol{Q}} + \epsilon\boldsymbol{I}$, $\hat{\boldsymbol{q}}' = \hat{\boldsymbol{q}}$, $\hat{t}' = \hat{t} + (1 + \gamma_2\,\mathbf{trace}(\boldsymbol{\Sigma}_0))\epsilon$, and $\hat{r}' = \hat{r} + \epsilon$ are necessarily feasible given $\hat{\boldsymbol{x}}$.

$$\Psi(\hat{\boldsymbol{x}};\gamma_1,\gamma_2) \leq \hat{t}' + \hat{r}' \leq \hat{r} + \hat{t} + (2 + \gamma_2\,\mathbf{trace}(\boldsymbol{\Sigma}_0))\epsilon \leq \text{OPT} + (3 + \gamma_2\,\mathbf{trace}(\boldsymbol{\Sigma}_0))\epsilon\ .$$

We conclude that $\hat{\boldsymbol{x}}$, which was obtained in time polynomial in $\log(1/\epsilon)$, is $\alpha\epsilon$-optimal according to Problem (3.7), where $\alpha = 3 + \gamma_2\,\mathbf{trace}(\boldsymbol{\Sigma}_0)$. This completes our proof. □

A.3 Proof of Lemma 3.8.2

The problem of finding a strictly feasible distribution can be formulated as finding an optimal pair (F, t) for the following infinite dimensional convex problem

$$\begin{aligned} \underset{F\in\mathcal{M},t\in\mathbb{R}}{\text{maximize}} \quad & t && \text{(A.5a)}\\ \text{subject to} \quad & \boldsymbol{b}_i - \mathbb{E}_F[\boldsymbol{\psi}_i(\boldsymbol{\xi})] - t\boldsymbol{r}_i \in \mathcal{K}_i^*\ ,\ \forall\, i \in \{1, 2, ..., L\}\ , && \text{(A.5b)}\end{aligned}$$

where $\boldsymbol{r}_i$ is a vector of norm one known to lie in the strict interior of $\mathcal{K}_i^*$, and verifying that $t^* > 0$. We first show that this problem actually reduces to a simpler convex problem which has finite number of decision variables

$$\begin{aligned} \underset{\boldsymbol{\mu}\in\mathcal{S},t\in\mathbb{R}}{\text{maximize}} \quad & t && \text{(A.6a)}\\ \text{subject to} \quad & \boldsymbol{b}_i - \boldsymbol{\psi}_i(\boldsymbol{\mu}) - t\boldsymbol{r}_i \in \mathcal{K}_i^*\ ,\ \forall\, i \in \{1, 2, ..., L\}\ . && \text{(A.6b)}\end{aligned}$$

One can simply see that for any feasible solution $(\bar{\boldsymbol{\mu}}, \bar{t})$ for Problem (A.6), the distribution $\delta_{\bar{\boldsymbol{\mu}}}$ will satisfy Constraint (A.5) with $\bar{t}$ and therefore achieve the same objective since $\mathbb{E}_{\delta_{\bar{\boldsymbol{\mu}}}}[\boldsymbol{\psi}_i(\boldsymbol{\xi})] = \boldsymbol{\psi}_i(\bar{\boldsymbol{\mu}})$. We must then verify that the converse is true. Given a pair $(\bar{F}, \bar{t})$, we can in fact show that the pair $(\mathbb{E}_{\bar{F}}[\boldsymbol{\xi}], \bar{t})$ is feasible according to Problem (A.6) and leads to the same objective value. This is demonstrated using the argument that for any

$i \in \{1, 2, ...L\}$,

$$\begin{aligned}
\boldsymbol{b}_i - \mathbb{E}_{\bar{F}}[\boldsymbol{\psi}_i(\boldsymbol{\xi})] - t\boldsymbol{r}_i \in \mathcal{K}_i^* \quad &\Leftrightarrow \quad \boldsymbol{y}^\mathsf{T}\left(\boldsymbol{b}_i - \mathbb{E}_{\bar{F}}[\boldsymbol{\psi}_i(\boldsymbol{\xi})] - t\boldsymbol{r}_i\right) \geq 0 \;\forall\, \boldsymbol{y} \in \mathcal{K}_i \\
&\Leftrightarrow \quad \boldsymbol{y}^\mathsf{T}(\boldsymbol{b}_i - t\boldsymbol{r}_i) - \mathbb{E}_{\bar{F}}[\boldsymbol{y}^\mathsf{T}\boldsymbol{\psi}_i(\boldsymbol{\xi})] \geq 0 \;\forall\, \boldsymbol{y} \in \mathcal{K}_i \\
&\Rightarrow \quad \boldsymbol{y}^\mathsf{T}(\boldsymbol{b}_i - t\boldsymbol{r}_i) - \boldsymbol{y}^\mathsf{T}\boldsymbol{\psi}_i(\mathbb{E}_{\bar{F}}[\boldsymbol{\xi}]) \geq 0 \;\forall\, \boldsymbol{y} \in \mathcal{K}_i \\
&\Leftrightarrow \quad \boldsymbol{b}_i - \boldsymbol{\psi}_i(\mathbb{E}_{\bar{F}}[\boldsymbol{\xi}]) - t\boldsymbol{r}_i \in \mathcal{K}_i^* \;\;,
\end{aligned}$$

where we used the fact that $\boldsymbol{y}^\mathsf{T}\boldsymbol{\psi}_i(\cdot)$ was assumed to be convex for all $\boldsymbol{y} \in \mathcal{K}_i$.

Because we have shown that the feasibility problem reduces to a convex semi-infinite problem, it is now possible to apply the ellipsoid method. The method will converge in polynomial time since we assumed that we had access to a feasibility oracle for each dual cone $\mathcal{K}_i^*$. Given that the optimal t^* that is returned is positive, then we know that the Dirac distribution $\delta_{\boldsymbol{\mu}^*}$ is strictly feasible with respect to the set $\mathcal{D}_2$. Otherwise, we can conclude that $\mathcal{D}_2$ is empty. □

Appendix B

Extended Proofs for Chapter 5

B.1 Proof of Theorem 5.4.4

We reduce the NP-complete 3SAT problem to solving Problem (5.4) with general reward uncertainty in the reward parameters.

3SAT Problem: Let W be a collection of disjunctive clauses $W = \{w_1, w_2, ..., w_M\}$ on a finite set of variables $V = \{v_1, v_2, ..., v_N\}$ such that $|w_m| = 3\ \forall m \in \{1, ..., M\}$. Let each clause be of the form $w = v_i \vee v_j \vee \neg v_k$, where $\neg v_k$ is the negation of v_k. Is there a truth assignment for V that satisfies all the clauses in W?

Given an instance of the 3SAT Problem, we can construct an MDP with uncertainty in the rewards such that 3SAT is satisfiable if and only if the optimal value for the Percentile MDP Problem (5.4) is greater than 0. After describing a 2-action uncertain MDP and its associated percentile optimization problem in steps 1 and 2, we will demonstrate, in Step 3, that feasible policies must necessarily be deterministic on a set of states. In Step 4, this fact will be used to build from such a feasible policy an assignment for the variables that satisfies all the clauses of the 3SAT problem. Step 5 will confirm that if the original 3SAT problem is satisfiable then the constructed percentile optimization problem has a feasible solution. The final step demonstrates that the transformation involved can be done in polynomial time with a polynomial amount of memory. Note that our notation will use $\mathbb{1}$ to refer to the indicator operator over a deterministic clause such that $\mathbb{1}\{v\} = 1 \iff v$ is true.

Step 1 : Let $W = \{w_1, w_2, ..., w_M\}$ and $V = \{v_1, v_2, ..., v_N\}$ be an instance of 3SAT, we first create a set of $N+1$ states $S_0 = \{s_{(0,0)}, s_{(1,0)}, ..., s_{(N,0)}\}$ with no reward and with two actions a_1, a_2 available. Then we create two sets of states $S_1 = \{s_{(0,1)}, s_{(1,1)}, ..., s_{(N,1)}\}$ and $S_2 = \{s_{(0,2)}, s_{(1,2)}, ..., s_{(N,2)}\}$ for which the rewards will be uncertain and, finally, we create an absorption state s_3 with no reward. The transition matrix for our uncertain MDP is described in Figure B.1. Specifically, if action a_j is taken in state $s_{(i,0)}$, the system transitions to state $s_{(i,j)}$ in the next time step. All remaining states, $s \in S_1 \cup S_2$, lead to s_3 for all actions. We set the initial distribution $\boldsymbol{q}$ to be uniform over the states in S_0, and the discount factor to any value $\alpha > 0$. It remains to describe the uncertainty over the rewards for states $S_1 \cup S_2$.

We create a discrete probability function $\mathcal{G}$ for the random reward matrix $\tilde{\boldsymbol{R}} \in \mathbb{R}^{N\times 2}$, with $\tilde{r}(s_{(i,j)}) = \tilde{R}_{(i,j)}$. To each variable v_n, $n \in \{1, ..., N\}$, we associate two events ($\mathcal{E}^{(v_n,1)}$ and $\mathcal{E}^{(v_n,2)}$) that each have probability $0.25/N$ of occurring. If event $\mathcal{E}^{(v_n,1)}$ occurs, it leads to using the reward matrix $\boldsymbol{R}^{(v_n,1)}$, while event $\mathcal{E}^{(v_n,2)}$ leads to using $\boldsymbol{R}^{(v_n,2)}$. The total probability of events $\{\mathcal{E}^{(v_1,1)}, \mathcal{E}^{(v_1,2)}, ..., \mathcal{E}^{(v_N,1)}, \mathcal{E}^{(v_N,2)}\}$ is $1/2$. Next, for each clause w_m in the original 3SAT problem, we create an event $\mathcal{E}^{(w_m)}$ that occurs with probability $0.5/M$. Drawing $\mathcal{E}^{(w_m)}$ leads to using the reward matrix $\boldsymbol{R}^{(w_m)}$. The matrices $\boldsymbol{R}^{(v_n,1)}$,$\boldsymbol{R}^{(v_n,2)}$ and $\boldsymbol{R}^{(w_m)}$ are described as follow

$$R^{(v_n,1)}_{(i,j)} = \begin{cases} -1 & \text{if } i=n \text{ and } j=1 \\ 0 & \text{otherwise} \end{cases}, \qquad R^{(v_n,2)}_{(i,j)} = \begin{cases} -1 & \text{if } i=n \text{ and } j=2 \\ 0 & \text{otherwise} \end{cases},$$

$$R^{(w_m)}_{(0,1)} = -1 \qquad R^{(w_m)}_{(0,2)} = -1$$

$$R^{(w_m)}_{(i,1)} = \begin{cases} 1 & \text{if } v_i \in w_m \\ 0 & \text{otherwise} \end{cases} \quad \forall\, 1 \le i \le N\,, \quad R^{(w_m)}_{(i,2)} = \begin{cases} 1 & \text{if } \neg v_i \in w_m \\ 0 & \text{otherwise} \end{cases} \quad \forall\, 1 \le i \le N\,.$$

where $\neg v$ means negating the boolean variable. It is clear that these events form a distribution

$$\sum_{n=1}^{N} \left(\mathbb{P}_{\mathcal{G}}(\mathcal{E}^{(v_n,1)}) + \mathbb{P}_{\mathcal{G}}(\mathcal{E}^{(v_n,2)}) \right) + \sum_{m=1}^{M} \mathbb{P}_{\mathcal{G}}(\mathcal{E}^{(w_m)}) = 1.$$

Step 2 : We now focus on the feasibility of

$$\mathbb{P}_{\mathcal{G}} \left(\mathbb{E}_{\boldsymbol{P}} \left[\sum_{t=0}^{\infty} \alpha^t \tilde{r}(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi \right] \ge 0 \right) \ge 0.75\,, \tag{B.1}$$

which is equivalent to showing that Problem (5.4)'s optimal value is higher than 0 when $\delta = 0.25$. With the constructed uncertain MDP, this constraint is equivalent to

$$\sum_{n=1}^{N} \mathbb{P}_{\mathcal{G}}(\mathcal{E}^{(v_n,1)}) \left(\mathbb{1}^{(v_n,1)} + \mathbb{1}^{(v_n,2)}\right) + \sum_{m=1}^{M} \mathbb{P}_{\mathcal{G}}(\mathcal{E}^{(w_m)}) \mathbb{1}^{(w_m)} \geq 0.75 \,, \tag{B.2}$$

where we used the fact that $\mathbb{P}_{\mathcal{G}}(\mathcal{E}^{(v_n,2)}) = \mathbb{P}_{\mathcal{G}}(\mathcal{E}^{(v_n,1)})$ and where

$$\begin{aligned}
\mathbb{1}^{(v_n,1)} &= \mathbb{1}\left\{\mathbb{E}_{\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \alpha^t \tilde{r}(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi, \mathcal{E}^{(v_n,1)}\right] \geq 0\right\} \\
\mathbb{1}^{(v_n,2)} &= \mathbb{1}\left\{\mathbb{E}_{\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \alpha^t \tilde{r}(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi, \mathcal{E}^{(v_n,2)}\right] \geq 0\right\} \\
\mathbb{1}^{(w_m)} &= \mathbb{1}\left\{\mathbb{E}_{\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \alpha^t \tilde{r}(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi, \mathcal{E}^{(w_m)}\right] \geq 0\right\} \,.
\end{aligned}$$

Step 3 : In the described uncertain MDP, we now outline why a policy that is feasible according to Constraint (B.1) must be deterministic for states $\{s_{(1,0)}, \ldots s_{(N,0)}\}$. First we show that $\mathbb{1}^{(v_n,1)} + \mathbb{1}^{(v_n,2)}$ is 1 if the policy for state $s_{(n,0)}$ is deterministic and 0 otherwise. For any v_n in this MDP,

$$\begin{aligned}
\mathbb{1}^{(v_n,1)} &= \mathbb{1}\left\{\mathbb{E}_{\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \alpha^t \tilde{r}(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi, \mathcal{E}^{(v_n,1)}\right] \geq 0\right\} \\
&= \mathbb{1}\left\{0 + \sum_{i=0}^{N} \alpha q_i \left(\pi(i,1) R^{(v_n,1)}_{(i,1)} + \pi(i,2) R^{(v_n,1)}_{(i,2)}\right) + \sum_{t=2}^{\infty} 0 \geq 0\right\} \\
&= \mathbb{1}\left\{\frac{\alpha}{N+1} \sum_{i=1}^{N} -\mathbb{1}\{i = n \wedge 1 = 1\}\pi(i,1) - \mathbb{1}\{i = n \wedge 1 = 2\}\pi(i,2) \geq 0\right\} \\
&= \mathbb{1}\left\{\frac{-\alpha\pi(n,1)}{N+1} \geq 0\right\} = \mathbb{1}\{\pi(n,1) = 0\} = \mathbb{1}\{\pi(n,2) = 1\} \,,
\end{aligned}$$

where we started by expanding the expectation term, then used the definition of $\boldsymbol{R}^{(v_n,1)}$, and finally the fact that $\alpha > 0$, $\pi(n,1) > 0$, and that $\pi(n,1) + \pi(n,2) = 1$. By symmetry, it is also clear that $\mathbb{1}^{(v_n,2)} = \mathbb{1}\{\pi(n,1) = 1\}$.

It remains to show that, because $\mathbb{1}^{(v_n,1)} + \mathbb{1}^{(v_n,2)} \leq 1$, Constraint (B.2) can only be met with equality, which occurs if and only if the policy is deterministic for states $\{s_{(1,0)}, ..., s_{(N,0)}\}$.

$$\begin{aligned}
&\sum_{n=1}^{N} \mathbb{P}_{\mathcal{G}}(\mathcal{E}^{(v_n,1)})(\mathbb{1}^{(v_n,1)} + \mathbb{1}^{(v_n,2)}) + \sum_{m=1}^{M} \mathbb{P}_{\mathcal{G}}(\mathcal{E}^{(w_m)})\mathbb{1}^{(w_m)} \geq 0.75 \\
&\qquad \leftrightarrow \frac{0.25}{N}\sum_{n=1}^{N} \mathbb{1}\{\pi(n,\cdot) \text{ is deterministic}\} + \frac{0.5}{M}\sum_{m=1}^{M} \mathbb{1}^{(w_m)} \geq 0.75 \\
&\qquad \leftrightarrow (\pi(n,\cdot) \text{ is deterministic}) \ \wedge \ (\mathbb{1}^{(w_m)} = 1) \ , \ \forall\, n \in \{1, ...N\}, \forall\, m \in \{1, ...M\} \ .
\end{aligned}$$

Step 4 : With the described uncertain MDP, given a policy that is feasible according to Constraint (B.1), the assignment $v_i = \mathbb{1}\{\pi(i,1) = 1\}$ for the variables in V satisfies all the clauses in W. We already showed in Step 3, that for a policy π to be feasible according to Constraint (B.1), it must be deterministic for states $\{s_{(1,0)}, ..., s_{(N,0)}\}$ and satisfy $(\mathbb{1}^{(w_m)} = 1)$ for all $m \in \{1, ..., M\}$. Now given a clause $w \in W$, for example $w_1 = v_1 \vee v_2 \vee \neg v_3$, then by construction of event $\mathcal{E}^{(w_1)}$,

$$\begin{aligned}
(\mathbb{1}^{(w_m)} = 1) \ &\leftrightarrow \ \mathbb{E}_{\boldsymbol{P}}\left[\sum_{t=0}^{\infty} \alpha^t \tilde{r}(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi, \mathcal{E}^{(w_1)}\right] \ \geq \ 0 \\
&\leftrightarrow \ 0 + \sum_{i=0}^{N} \alpha q_i \left(\pi(i,1)R^{(w_1)}_{(i,1)} + \pi(i,2)R^{(w_1)}_{(i,2)}\right) + \sum_{t=2}^{\infty} 0 \ \geq \ 0 \\
&\leftrightarrow \ \frac{\alpha}{N+1}\left(-1 + \pi(1,1)\cdot 1 + \pi(2,1)\cdot 1 + \pi(3,2)\cdot 1\right) \ \geq \ 0 \\
&\leftrightarrow \ \pi(1,1) + \pi(2,1) + \pi(3,2) \ \geq \ 1 \\
&\rightarrow \ (\pi(1,1) = 1) \ \vee \ (\pi(2,1) = 1) \ \vee \ (\pi(3,1) = 0) \\
&\rightarrow \ v_1 \vee v_2 \vee \neg v_3 \ ,
\end{aligned}$$

given that π is deterministic. This can be shown for any clause $w \in W$ and it allows us to conclude that, given that the optimal value of Problem (5.4) is greater or equal to 0, the 3SAT problem is satisfiable and we can construct a satisfying assignment from the optimal point of the optimization problem.

Step 5 : It is also easy to show that if the optimal value for Problem (5.4) is smaller than 0, then there is no satisfying assignment for the 3SAT problem. This can be demonstrated by showing that given any satisfying assignment for the variables in V, there exists a policy π that is feasible according to Constraint (B.1). Using the satisfying assignment for the variables in V, we test the feasibility of policy

$$\begin{aligned}
&\pi(0,1) = 1 \\
&\pi(i,1) = \mathbb{1}\{v_i\}\ \forall i \in \{1, ..., N\} \\
&\pi(i,2) = \mathbb{1}\{\neg v_i\}\ \forall i \in \{1, ..., N\}\ ,
\end{aligned}$$

which is obviously a valid deterministic policy. But also, since, for example, clause $w_1 = (v_1 \vee v_2 \neg v_3)$ is satisfied by the variable assignment, then $\pi(1,1) + \pi(2,1) + \pi(3,2) \geq 1$ is necessarily satisfied. Thus, $\mathbb{1}^{(w_m)} = 1$ is satisfied for all $w_m \in W$. From the statements presented in steps 3 and 4, we get that Constraint (B.1) is satisfied by π.

$$\begin{aligned}
&\mathbb{P}_{\mathcal{G}}\left(\mathbb{E}_{\boldsymbol{P}}\left[\sum_{t=0}^{\infty}\alpha^t \tilde{r}(x_t)\middle| x_0 \propto \boldsymbol{q}, \pi\right] \geq 0\right) \\
&\qquad = \frac{0.25}{N}\sum_{n=1}^{N}\mathbb{1}\{\pi(n,\cdot) \text{ is deterministic}\} + \frac{0.5}{M}\sum_{m=1}^{M}\mathbb{1}^{(w_m)} \ = \ 0.75\ \ .
\end{aligned}$$

Step 6 : The uncertain MDP that is used to solve the 3SAT problem can be constructed in polynomial time. First, the MDP presented in Figure B.1 has $3(N+1)+1$ states and 2 actions, the transition matrix has therefore $2|S|^2$ entries which are either 0 or 1. Then, each of the $2N+M$ events can be described by its probability, $0.25/N$ or $0.5/M$, and its associated $|S| \times |A|$ reward matrix, with entries as either -1, 0, or 1. Overall, the problem can obviously be constructed in polynomial time and polynomial space. □

Remark B.1.1 : We want to note the fact that the proof did not require the assumption of stationarity in the reward uncertainty, or the stationarity of the policy π. In fact, the proof is valid for both types of uncertainty and strategies (*i.e.*, stationary or non-stationary).

B.2 Proof of Corollary 5.4.5

The proof of Corollary 5.4.5 follows similar lines as what was presented in the proof of Theorem 5.4.4. Given an instance of the NP-complete 3SAT Problem, one can construct in polynomial time the MDP with discrete transition uncertainty presented in Figure B.2. Based on this instance of the 3SAT Problem, the same events, $\mathcal{E}^{(v_n,1)}$, $\mathcal{E}^{(v_n,2)}$ and $\mathcal{E}^{(w_m)}$, can be created as described in the proof of Theorem 5.4.4. However, in this proof, drawing $\mathcal{E}^{(v_n,1)}$, $\mathcal{E}^{(v_n,2)}$ or , $\mathcal{E}^{(w_m)}$ from the posterior $\mathcal{G}$ will lead to using transition parameters $\boldsymbol{P}^{(v_n,1)}$, $\boldsymbol{P}^{(v_n,2)}$, or $\boldsymbol{P}^{(w_m)}$ respectively in the uncertain MDP. These parameters are defined as follows:

$$P_{(i,j)}^{(v_n,1)}(k) = \begin{cases} 1 & \text{if } r(k) = R_{(i,j)}^{(v_n,1)} \\ 0 & \text{otherwise} \end{cases}, \qquad P_{(i,j)}^{(v_n,2)}(k) = \begin{cases} 1 & \text{if } r(k) = R_{(i,j)}^{(v_n,2)} \\ 0 & \text{otherwise} \end{cases},$$

$$P_{(i,j)}^{(w_m)}(k) = \begin{cases} 1 & \text{if } r(k) = R_{(i,j)}^{(w_m)} \\ 0 & \text{otherwise} \end{cases},$$

where $R_{(i,j)}^{(v_n,1)}$, $R_{(i,j)}^{(v_n,2)}$ and $R_{(i,j)}^{(w_m)}$ refer to the definitions in the proof of Theorem 5.4.4.

Clearly, the MDP that is associated with each event is equivalent to the MDP that was originally associated with that event in the proof of Theorem 5.4.4. Therefore, in the constructed Percentile MDP where $\tilde{\boldsymbol{P}}$ is drawn from $\mathcal{G}$, one can use similar arguments to show that the feasibility of

$$\mathbb{P}_{\mathcal{G}}\left(\mathbb{E}_{\tilde{\boldsymbol{P}}}\left[\sum_{t=0}^{\infty} \alpha^t r(x_t) \middle| x_0 \propto \boldsymbol{q}, \pi\right] \geq 0\right) \geq 0.75 \,,$$

is equivalent to determining the satisfiability of the instance of the 3SAT Problem. □

B.3 Proof of Lemma 5.4.9

Take the assignment $y_1 = \mu - \frac{\sigma}{\sqrt{\delta}}$, Markov's inequality can be used to show that y_1 is always feasible according to Constraint (5.9b):

$$\begin{aligned} \mathbb{P}(\tilde{z} \geq y_1) &= \mathbb{P}\left(\tilde{z} \geq \mu - \frac{\sigma}{\sqrt{\delta}}\right) = \mathbb{P}\left(|\tilde{z} - \mu| \leq \frac{\sigma}{\sqrt{\delta}}\right) + \mathbb{P}\left(\tilde{z} - \mu \geq \frac{\sigma}{\sqrt{\delta}}\right) \\ &\geq 1 - \mathbb{P}\left(|\tilde{z} - \mu| \geq \frac{\sigma}{\sqrt{\delta}}\right) \geq 1 - \delta \,. \end{aligned}$$

On the other hand, for any $\eta \in (0, 1 - \delta)$, $y_2 = \mu + \frac{\sigma}{\sqrt{(1-\delta-\eta)}}$ is on the contrary always assured to be infeasible:

$$\begin{aligned} \mathbb{P}(\tilde{z} \geq y_2) &= \mathbb{P}\left(\tilde{z} \geq \mu + \frac{\sigma}{\sqrt{(1-\delta-\eta)}}\right) \leq \mathbb{P}\left(|\tilde{z} - \mu| \geq \frac{\sigma}{\sqrt{(1-\delta-\eta)}}\right) \\ &\leq 1 - \delta - \eta \,. \end{aligned}$$

Therefore, $\mu - \frac{\sigma}{\sqrt{\delta}} \leq y^* \leq \mu + \frac{\sigma}{\sqrt{1-\delta}}$. $\square$

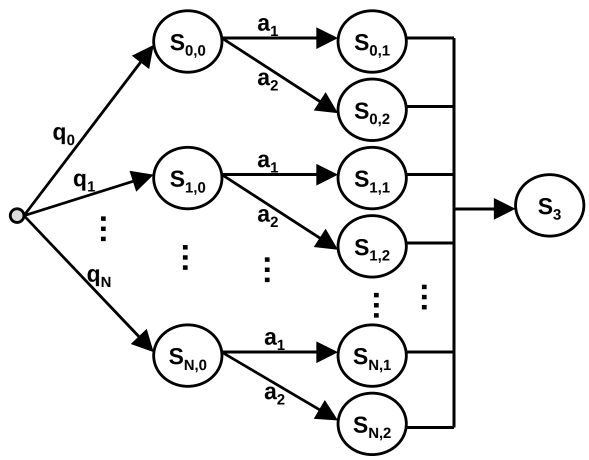


Figure B.1: Structure for the MDP with uncertainty in the rewards used in the proof of Theorem 5.4.4.

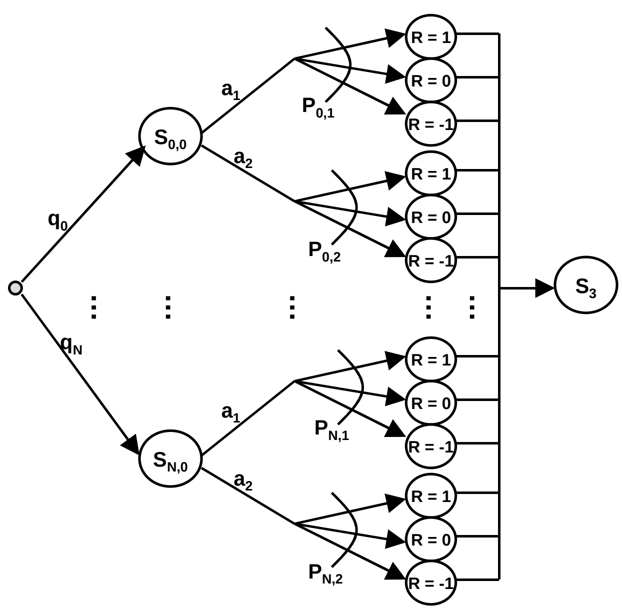


Figure B.2: Structure for the MDP with uncertainty in the transition probabilities used in the proof of Corollary 5.4.5.

Bibliography

[1] T. W. Anderson. *An Introduction to Multivariate Analysis*. John Wiley & Sons, 1984.

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

[3] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[4] D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. Working paper, 2008.

[5] D. Bertsimas, X. V. Doan, K. Natarajan, and C. P. Teo. Models for minimax stochastic linear optimization problems with risk aversion. Submitted to *Mathematics of Operations Research*, 2009.

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

[7] D. Bertsimas, I. Popescu, and J. Sethuraman. Moment problems and semidefinite programming. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *Handbook of Semidefinite Programming*, pages 469–510. Kluwer Academic Publishers, 2000.

[8] D. Bertsimas and S. Vempala. Solving convex programs by random walks. *Journal of the ACM*, 51:540–556, 2004.

[9] J. R. Birge and R. J.-B. Wets. Computing bounds for stochastic programming problems by means of a generalized moment problem. *Mathematics of Operations Research*, 12(1):149–162, 1987.

[10] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[11] G. Calafiore and M. C. Campi. Uncertain convex programs: Randomized solutions and confidence levels. *Mathematical Programming*, 102:25–46, 2005.

[12] G. Calafiore and L. El Ghaoui. On distributionally robust chance-constrained linear programs. *Optimization Theory and Applications*, 130(1):1–22, 2006.

[13] A. Charnes and W. W. Cooper. Chance constrained programming. *Management Science*, 6:73–79, 1959.

[14] D. P. de Farias and B. Van Roy. On constraint sampling for the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research*, 29:2004, 2001.

[15] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. Accepted in *Operations Research*, 2009.

[16] J. Dupacová. The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20:73–88, 1987.

[17] J. Dupacová. Stochastic programming: Minimax approach. *Encyclopedia of Optimization*, 5:327–330, 2001.

[18] A. Edelman. *Eigenvalues and Condition Numbers of Random Matrices*. PhD thesis, MIT, 1989.

[19] L. El Ghaoui, M. Oks, and F. Oustry. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556, 2003.

[20] Y. Ermoliev, A. Gaivoronski, and C. Nedeva. Stochastic optimization problems with partially known distribution functions. *Journal on Control and Optimization*, 23:696–716, 1985.

[21] J. A. Filar, D. Krass, and K. W. Ross. Percentile performance criteria for limiting average Markov control problems. In *IEEE Transactions on Automatic Control*, volume 40, pages 2–10, 1995.

[22] Y. Fujikoshi. Asymptotic expansions for the distributions of the sample roots under nonnormality. *Biometrika*, 67(1):45–51, 1980.

[23] A. A. Gaivoronski. A numerical method for solving stochastic programming problems with moment constraints on a distribution function. *Annals of Operations Research*, 31:347–370, 1991.

[24] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis, Second Edition*. Chapman & Hall/CRC, 2003.

[25] R. Givan, S. M. Leach, and T. Dean. Bounded-parameter Markov decision processes. *Artificial Intelligence*, 122(1-2):71–109, 2000.

[26] J. L. Goffin and J. P. Vial. On the computation of weighted analytic centers and dual ellipsoids with the projective algorithm. *Mathematical Programming*, 60(1):81–92, 1993.

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

[28] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, 2009. Software available at: http://stanford.edu/ boyd/cvx.

[29] M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1:169–197, 1981.

[30] K. Isii. On the sharpness of Chebyshev-type inequalities. *Annals of the Institute of Statistical Mathematics*, 14:185–197, 1963.

[31] G. Iyengar. Robust dynamic programming. *Mathematics of Operations Research*, 2002.

[32] P. Kall. Stochastic programming with recourse: Upper bounds and moment problems A review. In *Advances in Mathematical Optimization*, pages 86–103, 1988.

[33] C. M. Lagoa and B. R. Barmish. Distributionally robust Monte Carlo simulation: A tutorial survey. In *Proceedings of the International Federation of Automatic Control World Congress*, pages 1–12, 2002.

[34] H. J. Landau. *Moments in Mathematics: Lecture Notes Prepared for the AMS Short Course*. American Mathematical Society, 1987.

[35] O. Listes and R. Dekker. A scenario aggregation-based approach for determining a robust airline fleet composition for dynamic capacity allocation. *Transportation Science*, 39(3):367–382, 2005.

[36] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284:193–228, 1998.

[37] S. Mannor, D. Simester, P. Sun, and J. N. Tsitsiklis. Bias and variance approximation in value function estimates. *Management Science*, 53(2):308–322, 2007.

[38] A. Marshall and I. Olkin. Multivariate Chebyshev inequalities. *Annals of Mathematical Statistics*, 31:1001–1024, 1960.

[39] C. McDiarmid. Concentration. In M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, and B. Reed, editors, *Probabilistic Methods for Algorithmic Discrete Mathematics*, pages 195–248. Springer, 1998.

[40] K. Natarajan, M. Sim, and J. Uichanco. Tractable robust expected utility and risk models for portfolio optimization. Accepted in *Mathematical Finance*, 2008.

[41] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2006.

[42] Y. Nesterov and A. Nemirovski. *Interior-point polynomial methods in convex programming*, volume 13. Studies in Applied Mathematics, 1994.

[43] A. Nilim and L. El Ghaoui. Robust Markov decision processes with uncertain transition matrices. *Operations Research*, 2005.

[44] I. Pólik and T. Terlaky. A survey of the S-Lemma. *SIAM Review*, 49(3):371–418, 2007.

[45] I. Popescu. Robust mean-covariance solutions for stochastic optimization. *Operations Research*, 55(1):98–112, 2007.

[46] A. Prékopa. *Stochastic Programming*. Kluwer Academic Publishers, 1995.

[47] M. L. Putterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994.

[48] A. Rajwade. *Squares*, volume 171 of *London Mathematical Society Lecture Notes Series*. Cambridge University Press, 1993.

[49] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41, 2000.

[50] R. T. Rockafeller. *Convex Analysis*. Princeton University Press, 1970.

[51] R. T. Rockafeller. *Conjugate Duality and Optimization*, Regional Conference Series in Applied Mathematics, volume 16. SIAM, 1974.

[52] H. Scarf. A min-max solution of an inventory problem. *Studies in The Mathematical Theory of Inventory and Production*, pages 201–209, 1958.

[53] A. Shapiro. On duality theory of conic linear problems. In M. A. Goberna and M. A. López, editors, *Semi-Infinite Programming: Recent Advances*, pages 135–165. Kluwer Academic Publishers, 2001.

[54] A. Shapiro. Worst-case distribution analysis of stochastic programs. *Mathematical Programming*, 107(1):91–96, 2006.

[55] A. Shapiro and T. Homem-de-Mello. On the rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. *SIAM Journal on Optimization*, 11(1):70–86, 2000.

[56] A. Shapiro and A. J. Kleywegt. Minimax analysis of stochastic problems. *Optimization Methods and Software*, 17:523–542, 2002.

[57] J. Shawe-Taylor and N. Cristianini. Estimating the moments of a random vector with applications. In J. Siemons, editor, *Proceedings of GRETSI 2003 Conference*, pages 47–52. Cambridge University Press, 2003.

[58] M. Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1):171–176, 1958.

[59] A. M.-C. So. Private communication, 2008.

[60] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999.

[61] J. Čerbáková. Worst-case VaR and CVaR. In H. Haasis, H. Kopfer, and J. Schnberger, editors, *Operations Research Proceedings*, pages 817–822. Springer, 2005.

[62] C. Waternaux. Asymptotic distribution of the sample roots for the nonnormal population. *Biometrika*, 63(3):639–645, 1976.

[63] S. S. Wilks. *Mathematical Statistics*. John Wiley & Sons, 1962.

[64] Y. Ye. Complexity analysis of the analytic center cutting plane method that uses multiple cuts. *Mathematical Programming*, 78(1):85–104, 1997.

[65] J. Yue, B. Chen, and M.-C. Wang. Expected value of distribution information for the newsvendor problem. *Operations Research*, 54(6):1128–1136, 2006.

[66] S. S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. Technical report, Kyoto University, 2005.

[67] Z. Zhu, J. Zhang, and Y. Ye. Newsvendor optimization with limited distribution information. Technical report, Stanford University, 2006.