GOAL DRIVEN OPTIMIZATION

CHEN WENQING

NATIONAL UNIVERSITY OF SINGAPORE

2007

GOAL DRIVEN OPTIMIZATION

CHEN WENQING

(*B.Eng., Shanghai Jiao Tong University, China*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF DECISION SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

2007

ACKNOWLEDGMENT

I would like to express my sincere thanks to those who have consistently been helping me with my research work.

I am grateful to my supervisors, Professor Sun Jie and Professor Melvyn Sim, for their encouragement, support and valuable advice on my research work, all along the way of improving both my mind set in research and my skills to overcome problems.

I want to thank all my colleagues and friends in National University of Singapore and in my hometown Shanghai who have given me much support during my four-year Ph.D. studies. I also thank my parents and my husband for their support, love and care, which I can never live without.

Last but not least, I am grateful to National University of Singapore for providing me with the environment and facilities needed to complete my study.

CONTENTS

1. Introduction . . . 1

1.1 Decision Criterion . . . 2

1.2 Safeguarding Constraint . . . 6

1.3 Purpose of the Thesis . . . 10

2. Shortfall Aspiration Level Criterion and Goal Driven Model . . . 11

2.1 Aspiration Level Criterion . . . 11

2.2 Shortfall Aspiration Level Criterion . . . 13

2.3 Example: Single Product Newsvendor Problem . . . 22

2.4 Reduction to Stochastic Optimization Problems with CVaR Objectives . . . 27

3. Deterministic Approximations for Goal Driven Model . . . 32

3.1 Assumption on Data Structure . . . 32

3.2 Approximation of $\mathrm{E}((y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+)$ and $CVaR_{1-\gamma}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})$ . . . 39

3.3 Decision Rule Approximation of Recourse . . . 48

3.4 Example: Multi-product Newsvendor Problem . . . 55

4. Goal Driven Model with Probabilistic Constraint . . . 67

4.1 Individual probabilistic Constraint . . . 67

4.2 Joint probabilistic Constraint . . . 79

4.3 Optimizing over α . . . 88

4.4 Example: Emergency Resource Allocation . . . 93

4.5 Goal Driven Model with Probabilistic Constraint . . . 99

5. Applications 103
5.1 Project Management 103
5.2 Case Study: NFL Replica Jerseys 111
5.2.1 Full postponement strategy 114
5.2.2 Partial postponement strategy 117
5.2.3 Tradeoff between profit and service level 127

6. Conclusion 131
6.1 Summary of Results 131
6.2 Future Studies. 133

Appendix 143
.1 Proof of Theorem 5 144
.2 Approximation of a conic exponential quadratic constraint . . 149

SUMMARY

Achieving a target objective, goal or aspiration level are relevant aspects of decision making under uncertainties. We develop a goal driven stochastic optimization model that takes into account an aspiration level. Our model maximizes the *shortfall aspiration level criterion*, which encompasses the probability of success in achieving the goal and an expected level of underperformance or shortfall.

The key advantage of the proposed model is its tractability. We show that proposed model is reduced to solving a small collection of stochastic linear optimization problems with objectives evaluated under the popular conditional-value-at-risk (CVaR) measure. Using techniques in robust optimization, we propose a decision rule based deterministic approximation of the goal driven optimization problem by solving a polynomial number of subproblems, with each subproblem being a second order cone problem (SOCP).

As an extension, we consider the probabilistic constrained problem where a system of linear inequalities with stochastic entries is required to remain feasible with high probability. We review SOCP approximations for the individual probabilistic constrained problem. Moreover, a new formulation is proposed for approximating joint probabilistic constrained problem. Improvement of the new method upon the standard approach is shown.

We apply the goal driven model to project management and inventory planning problems and show experimentally that the proposed algorithms are computationally efficient.

LIST OF TABLES

4.1 Comparisons among Worst case solution Z^W, Solution using Bonferroni's inequality Z^B and Solution using new approximation Z^N. 96

5.1 Comparison of the deterministic and sampling models on $(1 - S\hat{A}LC)$. 110

5.2 Demand prediction for New England Patriots of the 2003 season 114

LIST OF FIGURES

3.1 Plot of $\rho_i(\gamma)$ against γ for $i = 3, 4$ and 5, defined in Proposition 5. 42
3.2 Goal driven optimization versus maximizing expected profit ($m = 10$) 63
3.3 Shortfall aspiration level criteria evaluated on shifted exponential distribution with sampling approximation using the same distribution. 65
3.4 Shortfall aspiration level criteria evaluated on normal distribution with sampling approximation using the same distribution. 66
3.5 Shortfall aspiration level criteria evaluated on normal distribution with sampling approximation using the shifted exponential distribution. 66

4.1 Inventory allocation: 15 nodes, 50 arcs 97
4.2 Convergence of the heuristic: 15 nodes, 50 arcs 98

5.1 Activity grid 6 by 4 107
5.2 Comparison of the deterministic and sampling models on $(1 - \hat{SALC})$ 108
5.3 Comparison of the deterministic and sampling models on Probability of violation and worst case completion time. 109
5.4 Supply chain 112
5.5 Shortfall expected profit vs Purchasing quantity 115

5.6 Full postponement: $SALC$ vs Target profit 116
5.7 Partial postponement: Frequency of the profit – Goal driven model vs Maximizing expected profit 121
5.8 Partial postponement: Frequency of the profit – Different target levels 121
5.9 Partial postponement vs Full postponement 122
5.10 Test on exponential distribution 123
5.11 Test on normal distribution 124
5.12 Test on uniform distribution 125
5.13 Test on two point discrete distribution 126
5.14 Tradeoff between risk and service level 130

.1 Evaluation of approximation of $\inf_{\mu>0} \mu \exp\left(\frac{a}{\mu} + \frac{1}{\mu^2}\right)$. 152

1. INTRODUCTION

Data uncertainties are present in many real world applications. In a supply chain, the demand, capacity, and resource potential are always unknown and can only be predicted in some precision. In finance, the security return and exchange rate fluctuate frequently. Even in engineering or science, the existence of measurement errors leads to uncertainties in the data. To handle the uncertainties, the real problem can be modeled as a mathematical programming problem in which some of the unknown parameters are taken as random variables. The mathematical programming problem is known as the stochastic optimization problem.

Obviously, the objective and the constraint functions of a stochastic optimization problem might be affected by the random parameters. If the objective function includes random parameters, it cannot be simply minimized or maximized, so it is necessary to specify a criterion for making decisions. The decision criterion takes the statistical features of the objective, so the random objective can be transformed to a deterministic equivalent. On the other hand, the random parameters often cause the constraint infeasibility when the solutions are obtained using nominal data values, so we also want to protect the constraints from this infeasibility. We classify all constraints that handle uncertainties as safeguarding constraints. The next sections will review some decision criteria and safeguarding constraints.

Notations We denote a regular face letter as a scalar or function. E and P represent the expectation function and the probability function respectively. Bold face lower case letters such as $\boldsymbol{x}$ represent vectors and the corresponding upper case letters such as $\boldsymbol{A}$ denote matrices. We denote random variable with the tilde sign, such as $\tilde{x}$. In addition, $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$. The same operations can be made on vectors, such as $\boldsymbol{y}^+$ and $\boldsymbol{z}^-$ in which corresponding operations are performed componentwise.

1.1 Decision Criterion

In a classical stochastic optimization problem, one seeks to minimize the aggregated expected cost over a multiperiod planning horizon, which corresponds to decision makers who are risk neutral; see for instance, Birge and Louveaux [16]. However, optimization of an expectation implicitly assumes that the decision can be repeated a great number of times under identical conditions. Such assumptions may not be widely applicable in practice. The framework of stochastic optimization can also be adopted to address downside risk by optimizing over an expected utility or more recently, a mean risk objective; see chapter 2 of Birge and Louveaux [16], Ahmed [1] and Ogryczak and Ruszczynski [45]. In such a model, the onus is on the decision maker to articulate his/her utility function or to determine the right parameter for the mean-risk functional. This can be rather subjective and difficult to obtain in practice.

Recent research in decision theory suggests a way of comprehensively and rigorously discussing decision theory without using utility functions; see Castagnoli and LiCalzi [20] and Bordley and LiCalzi [15]. With the introduction of an *aspiration level* or the targeted objective, the decision risk analysis focuses on making decisions so as to maximize the probability of reaching the aspiration level. As a matter of fact, the aspiration level plays an important role in daily decision making. Lanzillotti's study [34], which interviewed the officials of 20 large companies, verified that the managers are more concerned about a target return on investment. In another study, Payne et al. [46, 47] illustrated that managers tend to disregard investment possibilities that are likely to under perform against their target. Simon [58] also argued that most firms' goals are not maximizing profit but attaining a target profit. In an empirical study by Mao [39], managers were asked to define what they considered as risk. From their responses, Mao concluded that "risk is primarily considered to be the prospect of not meeting some target rate of return".

In this thesis, we study a two stage stochastic optimization model that takes into account an aspiration level. This work is closely related to Charnes et al.'s P-model [21, 22] and Bereanu's [12] optimality criterion of maximizing the probability of getting a profit above a targeted level. However, maximizing the probability of achieving a target is generally not a computationally tractable model. As such, studies along this objective have been confined to simple problems such as the Newsvendor problem; see Sankarasubramanian and Kumaraswamy [56], Lau and Lau [36], Li et al. [38] and Parlar and

Weng [48].

Besides its computational intractability, maximizing the success probability assumes that the modeler is indifferent to the level of losses. It does not address how catastrophic these losses can be expected when the "bad" events of small probability occur. However, studies have suggested that subjects are not completely insensitive to these losses; see for instance Payne et al [46]. Diecidue and van de Ven [25] argue that a model that solely maximizes the success probability is "too crude to be normatively or descriptively relevant." They suggested an objective that takes into account of a weighted combination of the success probability as well as an expected utility. However, such a model remains computationally intractable when applied to the stochastic optimization framework.

Our goal driven optimization model maximizes the *shortfall aspiration level criterion*, which takes into account of the probability of success in achieving the goal and an expected level of under-performance or shortfall. A key advantage of the proposed model over maximizing the success probability is its tractability. We show that the proposed model is reduced to solving a small collections of stochastic optimization problems with objectives evaluated under the popular Conditional-Value-at-Risk (CVaR) measure proposed by Rockafellar and Uryasev [54]. This class of stochastic optimization problems with mean risk objectives have recently been studied by Ahmed [1] and Riis and Schultz [52]. They proposed decomposition methods that facilitate sampling approximations.

The quality of sampling approximation of a stochastic optimization problem depends on several issues; the confidence of the approximation around the desired accuracy, the size of the problem, the type of recourse and the variability of the objective; see Shaprio and Nemirovski [57]. Even in a two stage model, the number of sampled scenarios required to approximate the solution to reasonable accuracy can be astronomically large, for instance, in the presence of rare but catastrophic scenarios or in the absence of relatively complete recourse. Moreover, sampling approximation of stochastic optimization problems requires complete probability descriptions of the underlying uncertainties, which are almost never available in real world environments. Hence, it is conceivable that the models that are heavily tuned to an assumed distribution may perform poorly in practice.

Recently, a new methodology dealing with uncertainties, called robust optimization, attracts a lot of attentions. Robust optimization makes mild distributional assumptions, such as the knowledge of the support or deviation measure, to approximate the stochastic optimization problems. The simplest approximation scheme of this type was proposed independently by Ben-Tal et al. [6, 7, 8] and El-Ghaoui et al. [28]. They showed that under the ellipsoidal uncertainty set, the robust counterpart of an LP becomes an SOCP. A more computationally convenient method was proposed by Bertsimas and Sim [14]. They used a polyhedral uncertainty set, with which the robust counterpart of an LP remains an LP. Chen, Sim and Sun [23] introduced the idea of *forward and backward deviation measures* to construct an asymmetric uncertainty set, with which the new robust counterpart successfully captures the asymmetry of random parameters. Motivated from recent development in robust optimization involving multiperiod decision process, we propose a new decision rule based deterministic approximation of the stochastic optimization problems with CVaR objectives. In line with robust optimization, we require only modest assumptions on distributions, such as known means, bounded supports, standard deviations and the *forward and backward deviations* introduced in [23]. We adopt a comprehensive model of uncertainty that incorporates both models in [23] and [24]. We also introduce new bounds on the CVaR measures and expected positivity of a weighted sum of random variables, both of which are integral in achieving a tractable approximation in the form of second order cone optimization problem (SOCP); see Ben-Tal and Nemirovski [10]. This allows us to leverage on the state-of-the-art SOCP solvers, which are increasingly more powerful, efficient and robust.

1.2 Safeguarding Constraint

All the constraints that handle uncertainties can be classified to safeguarding constraints. The simplest one is the worst case models, in which the constraints should be satisfied for all realizations of the random parameters. However, this strategy may be overconservative and even leads to an infeasible problem. Hence some violation allowances can provide more reasonable solutions and decisions. For example, a firm is willing to provide a relatively high level of product availability with an additional cost, because offering high service level not only keeps the current customers, but attracts new customers as well. However, the cost usually increases rapidly as the service level increases. It is impractical to require one hundred percent service level.

The tradeoff between the profit and the service level is also an important issue when making decisions. In this thesis, we use a goal driven model with constraints that allow some violations to describe such kind of problems.

Those constraints with violation allowances are called probabilistic constraints. Probabilistic constraints were first introduced by Charnes, Cooper, and Symonds [21]. A general way to express the probabilistic constraint is

$$\mathrm{P}\left(f_i(\boldsymbol{x}, \tilde{\boldsymbol{d}}) \leq 0,\ i = 1, \ldots, m\right) \geq 1 - \epsilon, \tag{1.1}$$

where $\epsilon \in (0, 1)$ is a given risk requirement, $f_i(\boldsymbol{x}, \tilde{\boldsymbol{d}})$ are known functions of the decision vector $\boldsymbol{x}$ and the random parameters $\tilde{\boldsymbol{d}}$. probabilistic constraints can be classified to two different types: individual ($m = 1$) and joint probabilistic constraints ($m > 1$).

Generally, probabilistic constrained problems are computationally intractable. The difficulties are as follows: first, with random parameters, it is difficult to evaluate the probability of the constraint satisfaction, which makes the whole problem computationally intractable. A possible way is to use Monte-Carlo simulation. However, it is too costly if the probability requirement ϵ is very small. It can be seen that the required sample size increases dramatically as the dimension of the problem increases or the probability requirement ϵ decreases. As given in [18], it can be concluded that the sample size should be at least inversely proportional to the probability requirement ϵ.

Second, even in the nice case that each f_i is affine with $\boldsymbol{x}$, probabilistic constraints are usually non-convex. If the random parameters and the decision variables can be separated, individual probabilistic constraint can be easily transformed to an equivalent linear constraint, but this property does not apply to joint probabilistic constraint. Joint probabilistic constraint is convex only when the separated random parameters follow logconcave distribution, a wide family of distributions such that the logarithm of cumulative density function is concave. On the other hand, if the random parameters and the decision variables cannot be separated, then the convexity holds only for some special cases, such as individual probabilistic constraint with normally distributed parameters [40]. Joint probabilistic constraints are generally non-convex.

For the convex problems, there are some beautiful methods in the literature of stochastic programming, such as supporting hyperplane, central cutting plane and reduced gradient method [49] [41]. However, for the general nonconvex cases, the efficiency of these methods is very low.

A natural way to deal with probabilistic constraints is to seek for convex conservative approximations, in the sense that if the approximation holds, the probabilistic constraint is satisfied. Nemirovski and Shapiro [44] proposed a special class of conservative approximations for the individual probabilistic constraint. They also proposed a beautiful convex formulation called Bernstein approximation. Although this approximation does not depend on any simulation or scenarios, it requires full knowledge of the moments information, which may not be easy to know. Moreover, the formulation involves some exponential cone, which may not be easy to solve. As for joint probabilistic constraint, they propose to use Bonferroni's inequality to approximate as follows

$$\begin{cases} \mathrm{P}\left(f_i(\boldsymbol{x}, \tilde{\boldsymbol{d}}) \leq 0\right) \geq 1 - \epsilon_i \\ \sum_{i=1}^{m} \epsilon_i = \epsilon. \end{cases}$$

Since the probability requirement for each component ϵ_i is no longer known, the approximation model becomes nonconvex. To simplify the problem, they choose the probability requirement for each component ϵ_i as ϵ/m.

Robust optimization methodologies can also be applied to consider the individual probabilistic constrained problems (see [6, 7, 8, 28, 14]). In [23], Chen et.al applied robust optimization to a project management network, in which a joint probabilistic constraint was formulated, but they also use Bonferroni's inequality and simply divided the probability requirement equally among the constraints to achieve the feasibility.

In this thesis, we show that with different definitions of the uncertainty set, we can approximate the individual probabilistic constraint to a second order cone formulation in different ways. For the problems with joint probabilistic constraints, we also show that Bonferroni's inequality may destroy the quality of the solutions, especially when the constraints are correlated with each other. We propose a new formulation to approximate the joint probabilistic constraint. The new formulation can be proved at least as good as the approximations using Bonferroni's inequality.

1.3 Purpose of the Thesis

This thesis analyzes the stochastic optimization problem in both the objective and the constraint aspects. To handle the random objective, a new decision criterion and the corresponding solution methodology are proposed. In addition, to protect the constraints from infeasibility, efficient methods are proposed to solve the probabilistic constrained problem. The aims of this thesis are as follows.

- To propose a new decision criterion, shortfall aspiration level criterion, which takes into account of the probability of success in achieving the goal and an expected level of under-performance or shortfall.
- To propose methods for improving solutions of models with probabilistic constraints.
- To apply goal driven models to project management and inventory planning problems.

It is recognized that among various stochastic optimization problems, the linear problem is the most widely used. Hence this thesis focuses on this case rather than general nonlinear problems.

2. SHORTFALL ASPIRATION LEVEL CRITERION AND GOAL DRIVEN MODEL

2.1 Aspiration Level Criterion

We consider a two stage decision process in which the decision maker first selects a feasible solution $\boldsymbol{x} \in \Re^{n_1}$, or so-called *here-and-now* solution in the face of uncertain outcomes that may influence the optimization model. Upon realization of $\tilde{\boldsymbol{z}}$, which denotes the vector of N random variables whose realizations correspond to the various scenarios, we select an optimal *wait-and-see* solution or recourse action. We also refer to $\tilde{\boldsymbol{z}}$ as the vector of primitive uncertainties, which consolidates all underlying uncertainties in the stochastic model. Given the solution, $\boldsymbol{x}$ and a realization of scenario, $\boldsymbol{z}$, the optimal wait-and-see objective we consider is given by

$$
\begin{aligned}
f(\boldsymbol{x}, \boldsymbol{z}) = \boldsymbol{c}(\boldsymbol{z})'\boldsymbol{x} + \quad \min_{\boldsymbol{u}, \boldsymbol{y}} \quad & \boldsymbol{d_u}'\boldsymbol{u} + \boldsymbol{d_y}'\boldsymbol{y} \\
\text{s.t.} \quad & \boldsymbol{B}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u} + \boldsymbol{Y}\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{z}) \\
& \boldsymbol{y} \geq \boldsymbol{0},
\end{aligned}
\tag{2.1}
$$

where $\boldsymbol{d_u} \in \Re^{n_2}$ and $\boldsymbol{d_y} \in \Re^{n_3}$ are known vectors, $\boldsymbol{U} \in \Re^{m_2 \times n_2}$ and $\boldsymbol{Y} \in \Re^{m_2 \times n_3}$ are known matrices, $\boldsymbol{c}(\tilde{\boldsymbol{z}}) \in \Re^{n_1}$, $\boldsymbol{B}(\tilde{\boldsymbol{z}}) \in \Re^{m_2 \times n_1}$ and $\boldsymbol{h}(\tilde{\boldsymbol{z}}) \in \Re^{m_2}$ are random data as function mapping of $\tilde{\boldsymbol{z}}$. In the language of stochastic optimization, this is a fixed recourse model in which the matrices $\boldsymbol{U}$ and $\boldsymbol{Y}$ associated with the recourse actions are not influenced by uncertainties; see Birge and Louveaux [16]. The model (2.1) represents a rather general fixed recourse framework characterized in classical stochastic optimization formulations. Using the convention of stochastic optimization, if the model (2.1) is infeasible, the function $f(\boldsymbol{x}, \boldsymbol{z})$ will be assigned an infinite value.

We denote by $\tau(\tilde{\boldsymbol{z}})$ the target level or aspiration level, which, in the most general setting, depends on the primitive uncertainties, $\tilde{\boldsymbol{z}}$; see Bordley and LiCalzi [15]. The wait-and-see objective $f(\boldsymbol{x}, \tilde{\boldsymbol{z}})$ is a random variable with probability distribution as a function of $\boldsymbol{x}$. Under the *aspiration level criterion*, which we will subsequently define, we examine the following model

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & ALC\Big(f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})\Big) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned} \tag{2.2}$$

where $\boldsymbol{b} \in \Re^{m_1}$ and $\boldsymbol{A} \in \Re^{m_1 \times n_1}$ are known. We use the phrase *aspiration level prospect* to represent the random variable, $f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})$. Hence, an aspiration level prospect taking a positive value denotes a shortfall of the wait-and-see objective against the target level. The functional $ALC(\cdot)$ is the aspiration level criterion, which evaluates the chance of exceeding the target level of performance.

Definition 1. Given an aspiration level prospect, $\tilde{v}$, the aspiration level criterion is defined as

$$ALC(\tilde{v}) \triangleq \mathrm{P}(\tilde{v} \leq 0). \tag{2.3}$$

We adopt the same definition as used in Diecidue and van de Ven [25] and in Canada et al. [19], chapter 5. We can equivalently express the aspiration level criterion as

$$ALC(\tilde{v}) = 1 - \mathrm{P}(\tilde{v} > 0) = 1 - \mathrm{E}(\mathcal{H}(\tilde{v})) \tag{2.4}$$

where $\mathcal{H}(\cdot)$ is a heavy-side utility function defined as

$$\mathcal{H}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

2.2 Shortfall Aspiration Level Criterion

The aspiration level criterion has several drawbacks from the computational and modeling perspectives. The lack of any form of structural convexity leads to computational intractability. Moreover, it is evident from Equation (2.4) that the aspiration level criterion does not take into account the shortfall level and may equally value a catastrophic event with low probability over a mild violation with the same probability. In view of the deficiencies of the aspiration level criterion, we introduce the shortfall aspiration level criterion.

Definition 2. Given an aspiration level prospect, $\tilde{v}$ with the following conditions:

$$\begin{aligned} \mathrm{E}(\tilde{v}) &< 0 \\ \mathrm{P}(\tilde{v} > 0) &> 0, \end{aligned} \tag{2.5}$$

the shortfall aspiration level criterion is defined as

$$SALC(\tilde{v}) \triangleq 1 - \inf_{a>0} \left(\mathrm{E} \left(\mathcal{S}(\tilde{v}/a) \right) \right) \tag{2.6}$$

where we define the shortfall utility function as follows:

$$\mathcal{S}(x) = (x+1)^{+}.$$

We present the properties of the shortfall aspiration level criterion in the following theorem.

Theorem 1. Let $\tilde{v}$ be an aspiration level prospect satisfying the inequalities (2.5). The shortfall aspiration level criterion has the following properties

(a)

$$SALC(\tilde{v}) \leq ALC(\tilde{v}).$$

(b)

$$SALC(\tilde{v}) \in (0,1).$$

Moreover, there exists a finite $a^* > 0$, such that

$$SALC(\tilde{v}) = 1 - \mathrm{E}\left(\mathcal{S}\left(\tilde{v}/a^*\right)\right)$$

(c)

$$SALC(\tilde{v}) = \sup_{\gamma}\{1-\gamma \ : \ CVaR_{1-\gamma}(\tilde{v}) \le 0, \gamma \in (0,1)\}$$

where

$$CVaR_{1-\gamma}(\tilde{v}) \triangleq \min_{\beta}\left(\beta + \frac{\mathrm{E}\left((\tilde{v}-\beta)^+\right)}{\gamma}\right) \tag{2.7}$$

is the risk measure known as Conditional-Value-at-Risk (CVaR) popularized by Rockafellar and Uryasev [54].

(d) Suppose for all $\boldsymbol{x} \in X$, $\tilde{v} = \tilde{v}(\boldsymbol{x})$ is normally distributed. Then the feasible solution that maximizes the shortfall aspiration level criterion also maximizes the aspiration level criterion.

Proof : (a) Observe that for all $a > 0$, $\mathcal{S}(x/a) \ge \mathcal{H}(x)$, hence, we have

$$\begin{aligned} \mathrm{P}(\tilde{v} > 0) &= \mathrm{E}\left(\mathcal{H}(\tilde{v})\right) \\ &\le \inf_{a>0} \mathrm{E}\left(\mathcal{S}(\tilde{v}/a)\right) \\ &= 1 - SALC(\tilde{v}). \end{aligned}$$

Therefore,

$$ALC(\tilde{v}) = \mathrm{P}(\tilde{v} \leq 0) = 1 - \mathrm{P}(\tilde{v} > 0) \geq SALC(\tilde{v}).$$

(b) Since $\mathrm{P}(\tilde{v} > 0) > 0$, from (a), we have $SALC(\tilde{v}) \leq 1 - \mathrm{P}(\tilde{v} > 0) < 1$. To show that $SALC(\tilde{v}) > 0$, it suffices to find a $b > 0$ such that $\mathrm{E}(\mathcal{S}(\tilde{v}/b)) < 1$. Observe that

$$\mathrm{E}(\mathcal{S}(\tilde{v}/a)) = 1 + \frac{\mathrm{E}(\tilde{v}) + \mathrm{E}\left((\tilde{v}+a)^-\right)}{a}.$$

As $\mathrm{E}(\tilde{v}) < 0$ and $\mathrm{E}\left((\tilde{v}+a)^-\right)$ is nonnegative, continuous in a and converges to zero as a approaches infinity, there exists a $b > 0$, such that $\mathrm{E}(\tilde{v}) + \mathrm{E}\left((\tilde{v}+b)^-\right) < 0$. Hence,

$$SALC(\tilde{v}) = 1 - \inf_{a>0} \frac{\mathrm{E}\left((\tilde{v}+a)^+\right)}{a} \geq 1 - \frac{\mathrm{E}\left((\tilde{v}+b)^+\right)}{b} > 0.$$

Since $\mathrm{P}(\tilde{v} > 0) > 0$ implies $\mathrm{E}(\tilde{v}^+) > 0$, we also observe that

$$\lim_{a \downarrow 0} \mathrm{E}(\mathcal{S}(\tilde{v}/a)) = \lim_{a \downarrow 0} \frac{\mathrm{E}\left((\tilde{v}+a)^+\right)}{a} \geq \lim_{a \downarrow 0} \frac{\mathrm{E}\left(\tilde{v}^+\right)}{a} = \infty.$$

Moreover,

$$\lim_{a \to \infty} \mathrm{E}(\mathcal{S}(\tilde{v}/a)) = 1.$$

We have also shown that $\inf_{a>0} \mathrm{E}(\mathcal{S}(\tilde{v}/a)) \in (0,1)$, hence, the infimum cannot be achieved at the limits of $a = 0$ and $a = \infty$. Moreover, due to the continuity of the function $\mathrm{E}(\mathcal{S}(\tilde{v}/a))$ over $a > 0$, the infimum is achieved at a finite $a > 0$.

(c) Using the observations in (b), we have

$$\begin{aligned}
& 1-\inf_{a>0}\mathrm{E}(\mathcal{S}(\tilde{v}/a)) \\
= \ & \sup_{v<0}\left(1+\frac{\mathrm{E}\big((\tilde{v}-v)^+\big)}{v}\right) \\
= \ & \sup_{\gamma,v}\left\{1-\gamma \ : \ 1-\gamma\le 1+\frac{\mathrm{E}\big((\tilde{v}-v)^+\big)}{v}, v<0,\gamma\in(0,1)\right\} \\
= \ & \sup_{\gamma,v}\left\{1-\gamma \ : \ v+\frac{\mathrm{E}\big((\tilde{v}-v)^+\big)}{\gamma}\le 0, v<0,\gamma\in(0,1)\right\} \\
= \ & \sup_{\gamma,v}\left\{1-\gamma \ : \ v+\frac{\mathrm{E}\big((\tilde{v}-v)^+\big)}{\gamma}\le 0,\gamma\in(0,1)\right\} \quad \text{With } \mathrm{E}(\tilde{v}^+)>0,\ v<0 \text{ is implied} \\
= \ & \sup_{\gamma}\{1-\gamma \ : \ CVaR_{1-\gamma}(\tilde{v})\le 0,\gamma\in(0,1)\}.
\end{aligned}$$

(d) Observe that

$$\max_{\boldsymbol{x}}\left\{ALC\Big(\tilde{v}(\boldsymbol{x})\Big) \ : \ \boldsymbol{x}\in\mathcal{X}\right\} \tag{2.8}$$

is equivalent to

$$\max_{\boldsymbol{x},\gamma}\left\{1-\gamma \ : \ \mathrm{P}\Big(\tilde{v}(\boldsymbol{x})\le 0\Big)\ge 1-\gamma, \boldsymbol{x}\in\mathcal{X}\right\}.$$

Let $\mu(\boldsymbol{x})$ and $\sigma(\boldsymbol{x})$ be the mean and standard deviation of $\tilde{v}(\boldsymbol{x})$. The constraint $\mathrm{P}\Big(\tilde{v}(\boldsymbol{x})\le 0\Big)\ge 1-\gamma$ is equivalent to

$$-\mu(\boldsymbol{x})\ge\Phi^{-1}(1-\gamma)\sigma(\boldsymbol{x}),$$

where $\Phi(\cdot)$ is the distribution function of a standard normal. Since $\mathrm{E}(\tilde{v}(\boldsymbol{x}))<0$, the optimal objective satisfies $1-\gamma>1/2$ and hence, $\Phi^{-1}(1-\gamma)>0$. Noting that $\Phi^{-1}(1-\gamma)$ is a decreasing function in γ, the optimal solution in

Model (2.8) corresponds to maximizing the following ratio:

$$\begin{aligned} \max \quad & \frac{-\mu(\boldsymbol{x})}{\sigma(\boldsymbol{x})} \\ \text{s.t.} \quad & \boldsymbol{x} \in \mathcal{X}. \end{aligned} \tag{2.9}$$

This relation was observed by Dragomirescu [26]. Using the result in (c), we can express the maximization of the shortfall aspiration level criterion as follows:

$$\begin{aligned} \max \quad & 1-\gamma \\ \text{s.t.} \quad & CVaR_{1-\gamma}(\tilde{v}(\boldsymbol{x})) \le 0 \\ & \boldsymbol{x} \in \mathcal{X}, \gamma \in (0,1) \end{aligned} \tag{2.10}$$

Under normal distribution, we can also evaluate the CVaR measure in closed form as follows:

$$CVaR_{1-\gamma}(\tilde{v}(\boldsymbol{x})) = \mu(\boldsymbol{x}) + \underbrace{\frac{\phi(\Phi^{-1}(\gamma))}{\gamma}}_{\xi(\gamma)} \sigma(\boldsymbol{x})$$

where $\phi(\cdot)$ is the density of a standard normal. Moreover, $\xi(\gamma)$ is also a decreasing function in γ. Therefore, the optimum solution of Model (2.10) is identical to Model (2.9). ■

We now propose the following goal driven optimization problem.

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & SALC\Big(f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})\Big) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned} \tag{2.11}$$

Theorem 1(a) implies that an optimal solution of Model (2.11), $\boldsymbol{x}^*$ can achieve the following success probability,

$$\mathrm{P}(f(\boldsymbol{x}^*, \tilde{\boldsymbol{z}}) \leq \tau(\tilde{\boldsymbol{z}})) \geq SALC\Big(f(\boldsymbol{x}^*, \tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})\Big).$$

The optimal parameter, a^* within the shortfall aspiration level criterion is chosen to attain the tightest bound in meeting the success probability. The aspiration level criterion of (2.4) penalizes the shortfall with an heavy-side utility function that is insensitive to the magnitude of violation. In contract, the shortfall aspiration level criterion,

$$SALC(f(\boldsymbol{x}^*, \tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})) = 1 - \frac{1}{a^*}\mathrm{E}\left((f(\boldsymbol{x}^*, \tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}}) + a^*)^+\right) \qquad \text{for some } a^* > 0$$

has an expected utility component that penalizes an expected level of "near" shortfall when the aspiration level prospect raises above $-a^*$. Speaking intuitively, given two aspiration level prospects, $\tilde{v}_1$ and $\tilde{v}_2$ with the same aspiration level criteria defined in (2.3), suppose $\tilde{v}_2$ incurs greater expected shortfall, the shortfall aspiration level criterion will rank $\tilde{v}_1$ higher than $\tilde{v}_2$. Nevertheless, Theorem 1(d) suggests that if the distribution of the objective is "fairly normally distributed", we expect the solution that maximizes the shortfall aspiration level criterion to also maximize the aspiration level criterion.

We now discuss the conditions of (2.5) with respect to the goal driven optimization model. The first condition implies that the aspiration level should be strictly achievable in expectation. Hence, the goal driven optimization model appeals to decision makers who are risk averse and are not unrealistic in setting their goals. The second condition implies that there does not exist a feasible solution, which always achieves the aspiration level. In other words, the goal driven optimization model is used in problem instances where the risk of under-performance is inevitable. Hence, it appeals to decision makers who are not too apathetic in setting their goals.

Theorem 1(c) shows the connection between the shortfall aspiration level criterion with the CVaR measure. The CVaR measure satisfies four desirable properties of financial risk measures known as *coherent risk*. A coherent risk measure or functional, $\varphi(\cdot)$ satisfies the following *Axioms of coherent risk measure*:

(i) Translation invariance: For all $a \in \Re$, $\varphi(\tilde{v} + a) = \varphi(\tilde{v}) + a$.

(ii) Subadditivity: For all random variables $\tilde{v}_1, \tilde{v}_2$, $\varphi(\tilde{v}_1 + \tilde{v}_2) \leq \varphi(\tilde{v}_1) + \varphi(\tilde{v}_2)$.

(iii) Positive homogeneity: For all $\lambda \geq 0$, $\varphi(\lambda\tilde{v}) = \lambda\varphi(\tilde{v})$.

(iv) Monotonicity: For all $\tilde{v} \leq \tilde{w}$, $\varphi(\tilde{v}) \leq \varphi(\tilde{w})$.

The four axioms were presented and justified in Artzner et al. [3]. The first axiom ensures that $\varphi(\tilde{v} - \varphi(\tilde{v})) = 0$, so that the risk of $\tilde{v}$ after compensation with $\varphi(\tilde{v})$ is zero. It means that reducing the cost by a fixed amount of a simply reduces the risk measure by a. The subadditivity axiom states that the risk associated with the sum of two financial instruments is not more than the sum of their individual risks. It appears naturally in finance - one can think equivalently of the fact that "a merger does not create extra risk," or of the "risk pooling effects" observed in the sum of random variables. The positive homogeneity axiom implies that the risk measure scales proportionally with its size. The final axiom is an obvious criterion, but it rules out the classical mean-standard deviation risk measure.

A byproduct of a risk measure that satisfies these axioms is the preservation of convexity; see for instance Ruszczynski and Shapiro [55]. Hence, the function $CVaR_{1-\gamma}(f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}}))$ is convex in $\boldsymbol{x}$. Using the connection with the CVaR measure, we express the goal driven optimization model (2.11), equivalently as follows:

$$\begin{aligned}\max_{\gamma,\boldsymbol{x}} \quad & 1-\gamma \\ \text{s.t.} \quad & CVaR_{1-\gamma}(f(\boldsymbol{x},\tilde{\boldsymbol{z}})-\tau(\tilde{\boldsymbol{z}})) \leq 0 \\ & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \\ & \gamma \in (0,1).\end{aligned} \tag{2.12}$$

2.3 Example: Single Product Newsvendor Problem

The classical single-product Newsvendor model maximizes the expected profit to help the decision makers to balance between the holding cost of excess inventory and the penalty for stockouts. In this section, we use the shortfall aspiration level criterion as objective to model the problem and we show that the goal driven model can be solved efficiently. We define

p : Unit selling price;

c : Unit purchasing cost;

s : Unit salvage value;

R : Target profit;

$\tilde{d}$: Demand;

x : Order quantity (Decision variable).

We formulate the problem as follows.

$$\max \quad SALC(-g(x,\tilde{d})+R), \tag{2.13}$$

where

$$\begin{aligned} g(x,d) &\triangleq (p-c)x+(s-p)(x-d)^+ \\ &= \begin{cases} (s-c)x+(p-s)d & \text{if } d<x \\ (p-c)x & \text{otherwise.} \end{cases} \end{aligned} \tag{2.14}$$

From the definition of the CVaR measure and the translation invariance property, we know that Model (2.13) is equivalent to

$$\begin{aligned} \max \quad & 1-\gamma \\ \text{s.t.} \quad & \min_{\beta}\Big(\frac{\mathrm{E}\Big((\beta-g(x,\tilde{d}))^+\Big)}{\gamma}-\beta\Big)+R\le 0. \end{aligned} \tag{2.15}$$

To obtain the optimal solution analytically, first, we have the following lemma.

Lemma 2. For any $0\le x_a<x_b$,

$$\mathrm{E}\Big(((p-c)x_a-g(x_a,\tilde{d}))^+\Big)<\mathrm{E}\Big(((p-c)x_a-g(x_b,\tilde{d}))^+\Big).$$

Proof : We let

$$\hat{d} = x_a + \frac{c-s}{p-s}(x_b - x_a).$$

Note that the nondecreasing piecewise linear function g has the following property:

$$(p-c)x_a > g(x_a, d) > g(x_b, d), \qquad \text{if } d < \hat{d};$$

$$(p-c)x_a = g(x_a, d) = g(x_b, d), \qquad \text{if } d = \hat{d};$$

$$(p-c)x_a < g(x_a, d) < g(x_b, d), \qquad \text{if } d > \hat{d}.$$

This property directly implies the result. ∎

Theorem 3. Assume that and there exists x satisfying

$$\mathrm{P}(g(x, \tilde{d}) < R) > 0, \tag{2.16}$$

$$\mathrm{E}(g(x, \tilde{d})) > R. \tag{2.17}$$

Then the model (2.15) is feasible. Moreover, the optimal solution x^*, β^* and γ^* satisfy

$$\text{(i) } \beta^* = (p-c)x^* \tag{2.18}$$

$$\text{(ii) } \gamma^* = \mathrm{P}(\tilde{d} < x^*) \tag{2.19}$$

$$\text{(iii) } \mathrm{E}\Big(g(x^*, \tilde{d}) \mid \tilde{d} < x^*\Big) = R. \tag{2.20}$$

Proof : With the assumptions (2.16) and (2.17), there exists x, such that the shortfall aspiration level criterion $SALC(-g(x,\tilde{d})+R) \in (0,1)$ (See Theorem 1). This also guarantees the feasibility of the model (2.15).

From the definition of the CVaR measure, we know that

$$\beta^* = \text{argmin}_\beta \frac{\mathrm{E}\Big((\beta - g(x^*,\tilde{d}))^+\Big)}{\gamma^*} - \beta.$$

From the first order condition, we have

$$\gamma^* = \mathrm{P}(g(x^*,\tilde{d}) < \beta^*). \tag{2.21}$$

Hence,

$$\beta^* \leq \max_d \Big\{ g(x^*,d) \Big\} = (p-c)x^*,$$

otherwise, $\gamma^* = 1$ contradicting the solution $SALC(-g(x,\tilde{d}) + R) \in (0,1)$. To show that $\beta^* = (p-c)x^*$, there remains to prove that $\beta^* \geq (p-c)x^*$. Suppose $\beta^* < (p-c)x^*$. Then there exists a $\delta > 0$ such that $\beta^* = (p-c)(x^* - \delta)$. From Lemma 2, we also notice that

$$\mathrm{E}\Big((\beta^* - g(x^* - \delta,\tilde{d}))^+\Big) < \mathrm{E}\Big((\beta^* - g(x^*,\tilde{d}))^+\Big).$$

Hence

$$\begin{aligned} R &= \beta^* - \frac{\mathrm{E}\Big((\beta^* - g(x^*, \tilde{d}))^+\Big)}{\gamma^*} \\ &< \beta^* - \frac{\mathrm{E}\Big(((\beta^* - g(x^* - \delta, \tilde{d}))^+\Big)}{\gamma^*}. \end{aligned}$$

If we define

$$\gamma' \triangleq \frac{\mathrm{E}\Big(((\beta^* - g(x^* - \delta, \tilde{d}))^+\Big)}{\beta^* - R},$$

which is also feasible in the model (2.15). It is obvious that $\gamma' < \gamma^*$, contradicting that γ^* is the optimal solution.

Substitute $\beta^* = (p - c)x^*$ into the equation (2.21). Since function g is non-decreasing, we have

$$\gamma^* = \mathrm{P}\Big(g(x^*, \tilde{d}) < (p - c)x^*\Big) = \mathrm{P}(\tilde{d} < x^*).$$

Also, substituting γ^* and β^* into the constraint of the model (2.15), we have

$$\mathrm{E}\Big(g(x^*, d) \mid d < x^*\Big) = R.$$

■

Theorem 3 implies that we can decide the optimal purchasing quantity and calculate the shortfall aspiration level criterion efficiently if the distribution of the demand is known. However, this result does not apply to more complicated problems.

In the next section, we show that for the general problems, the goal driven model can be reduced to solving a small collections of stochastic linear optimization problems with objectives evaluated under the popular conditional-value-at-risk (CVaR) measure.

2.4 Reduction to Stochastic Optimization Problems with CVaR Objectives

For a fixed γ, the first constraint in Model (2.12) is convex in the decision variable $\boldsymbol{x}$. However, the Model is not jointly convex in γ and $\boldsymbol{x}$. Nevertheless, we can still obtain the optimal solution by solving a sequence of subproblems in the form of stochastic optimization problems with CVaR objectives as follows:

$$\begin{aligned} Z(\gamma) = \quad \min_{\boldsymbol{x}} \quad & CVaR_{1-\gamma}\Big(f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})\Big) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned} \tag{2.22}$$

or equivalently,

$$\begin{aligned} Z(\gamma) = \min_{\boldsymbol{x},\boldsymbol{u}(\cdot),\boldsymbol{y}(\cdot)} \quad & CVaR_{1-\gamma}\Big(\boldsymbol{c}(\tilde{\boldsymbol{z}})'\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}(\tilde{\boldsymbol{z}}) + \boldsymbol{d_y}'\boldsymbol{y}(\tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})\Big) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{B}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}(\tilde{\boldsymbol{z}}) + \boldsymbol{Y}\boldsymbol{y}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}(\tilde{\boldsymbol{z}}) \\ & \boldsymbol{y}(\tilde{\boldsymbol{z}}) \geq \boldsymbol{0} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned} \tag{2.23}$$

where $\boldsymbol{u}(\tilde{\boldsymbol{z}})$ and $\boldsymbol{y}(\tilde{\boldsymbol{z}})$ correspond to the second stage or recourse variables in the space of measurable function.

Algorithm 1. (Binary Search)

Input: A routine that solves Model (2.22) optimally and $\zeta > 0$

Output: $\boldsymbol{x}$

1. Set $\gamma_1 := 0$ and $\gamma_2 := 1$.
2. If $\gamma_2 - \gamma_1 < \zeta$, stop. Output: $\boldsymbol{x}$
3. Let $\gamma := \frac{\gamma_1+\gamma_2}{2}$. Compute $Z(\gamma)$ from Model (2.22) and obtain the corresponding optimal solution $\boldsymbol{x}$.
4. If $Z(\gamma) \leq 0$, update $\gamma_2 := \gamma$. Otherwise, update $\gamma_1 := \gamma$
5. Go to Step 2.

Proposition 1. Suppose Model (2.12) is feasible. Algorithm 1 finds a solution, $\boldsymbol{x}$ with objective $1-\gamma^{\dagger}$ satisfying $|\gamma^{\dagger}-\gamma^{*}|<\zeta$ in at most $\lceil \log_2(1/\zeta) \rceil$ computations of the subproblem (2.22), where $1-\gamma^{*}$ being the optimal objective of Model (2.12).

Proof : Observe that each looping in Algorithm 1 reduces the gap between γ_2 and γ_1 by half. We now show the correctness of the binary search. Suppose $Z(\gamma) \le 0$, γ is feasible in Model (2.12), hence, $\gamma^{*} \le \gamma$. Otherwise, γ would be infeasible in Model (2.12). In this case, we claim that the optimal feasible solution, γ^{*} must be greater than γ. Suppose not, we have $\gamma^{*} \le \gamma$. We know the optimal solution $\boldsymbol{x}^{*}$ of Model (2.12) satisfies

$$CVaR_{1-\gamma^{*}}\Big(f(\boldsymbol{x}^{*},\tilde{\boldsymbol{z}})-\tau(\tilde{\boldsymbol{z}})\Big) \le 0.$$

However, since $\gamma^{*} \le \gamma$, we have

$$Z(\gamma) \le CVaR_{1-\gamma}\Big(f(\boldsymbol{x}^{*},\tilde{\boldsymbol{z}})-\tau(\tilde{\boldsymbol{z}})\Big) \le CVaR_{1-\gamma^{*}}\Big(f(\boldsymbol{x}^{*},\tilde{\boldsymbol{z}})-\tau(\tilde{\boldsymbol{z}})\Big) \le 0,$$

contradicting that $Z(\gamma) > 0$. ■

If $\tilde{\boldsymbol{z}}$ takes values from $\boldsymbol{z}^k$, $k = 1, \ldots, K$ with probability p_k, we can formulate the subproblem of (2.22) as a linear optimization problem as follows:

$$\begin{array}{lll}
\min\limits_{\beta,\boldsymbol{s},\boldsymbol{x},\boldsymbol{y}^k,\boldsymbol{y}^k} & \beta + \dfrac{1}{\gamma}\displaystyle\sum_{k=1}^{K} s_k p_k & \\
\text{s.t.} & s_k \geq \boldsymbol{c}(\boldsymbol{z}^k)'\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}^k + \boldsymbol{d_y}'\boldsymbol{y}^k - \tau(\boldsymbol{z}^k) - \beta & k = 1, \ldots, K \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& \boldsymbol{B}(\boldsymbol{z}^k)\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}^k + \boldsymbol{Y}\boldsymbol{y}^k = \boldsymbol{h}(\boldsymbol{z}^k) & k = 1, \ldots, K \\
& \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{s} \geq \boldsymbol{0} & \\
& \boldsymbol{y}^k \geq \boldsymbol{0} & k = 1, \ldots, K
\end{array}$$

Unfortunately, the number of possible recourse decisions increases proportionally with the number of possible realization of the random vector $\tilde{\boldsymbol{z}}$, which could be extremely large or even infinite. Nevertheless, under relatively complete recourse, the two stage stochastic optimization model can be solved rather effectively using sampling approximation. In such problems, the second stage problem is always feasible regardless of the choice of feasible first stage variables. Decomposition techniques has been studied in Ahmed [1] and Riis and Schultz [52] to enable efficient computations of the stochastic optimization problem with CVaR objective.

In the absence of relatively complete recourse, the solution obtained from sampling approximation may not be meaningful. Even though the objective of the sampling approximation could be finite, in the actual performance, the second stage problem can be infeasible, in which case the actual objective is infinite. Indeed, a two stage stochastic optimization is generally intractable. For instance, checking whether the first stage decision $\boldsymbol{x}$ gives rise to feasible recourse for all realization of $\tilde{\boldsymbol{z}}$ is already an NP-hard problem; see Ben-Tal et al. [5]. Moreover, with the assumption that the stochastic parameters are independently distributed, Dyer and Stougie [27] show that two-stage stochastic programming problems are NP-hard. Under the same assumption they show that certain multi-stage stochastic programming problems are PSPACE-hard. We therefore pursue an alterative method of approximating the stochastic optimization problem, that could at least guarantee the feasibility of the solution, and determine an upper bound of the objective function.

3. DETERMINISTIC APPROXIMATIONS FOR GOAL DRIVEN MODEL

We have shown that solving the goal driven optimization model (2.11) involves solving a sequence of stochastic optimization problems with CVaR objectives in the form of Model (2.23). Hence, we devote this section to formulating a tractable deterministic approximation of Model (2.23).

3.1 Assumption on Data Structure

One of the central problems in stochastic models is how to properly account for data uncertainty. Unfortunately, complete probability descriptions are almost never available in real world environments. Following the recent development of robust optimization such as Ben-Tal et al. [5], Bertsimas and Sim [14], Chen, Sim and Sun [23] and Chen et al. [24], we relax the assumption of full distributional knowledge and modify the representation of data uncertainties with the aim of producing a computationally tractable model.

We adopt the parametric uncertainty model in which the data uncertainties are affinely dependent on the primitive uncertainties.

Affine Parametric Uncertainty: We assume that the uncertain input data to the model $\boldsymbol{c}(\tilde{\boldsymbol{z}})$, $\boldsymbol{B}(\tilde{\boldsymbol{z}})$, $\boldsymbol{h}(\tilde{\boldsymbol{z}})$ and $\tau(\tilde{\boldsymbol{z}})$ are affinely dependent on the primitive uncertainties $\tilde{\boldsymbol{z}}$ as follows:

$$\begin{aligned}
\boldsymbol{c}(\tilde{\boldsymbol{z}}) &= \boldsymbol{c}^0 + \sum_{j=1}^{N} \boldsymbol{c}^j \tilde{z}_j, \\
\boldsymbol{B}(\tilde{\boldsymbol{z}}) &= \boldsymbol{B}^0 + \sum_{j=1}^{N} \boldsymbol{B}^j \tilde{z}_j, \\
\boldsymbol{h}(\tilde{\boldsymbol{z}}) &= \boldsymbol{h}^0 + \sum_{j=1}^{N} \boldsymbol{h}^j \tilde{z}_j, \\
\tau(\tilde{\boldsymbol{z}}) &= \tau^0 + \sum_{j=1}^{N} \tau^j \tilde{z}_j.
\end{aligned}$$

Note that this parametric uncertainty representation is useful for relating multivariate random variables across different data entries through the shared primitive uncertainties.

Since the assumption of having exact probability distributions of the primitive uncertainties is unrealistic, as in the spirit of robust optimization, we adopt a modest distributional assumption on the primitive uncertainties, such as known means, supports, subset of independently distributed random variables and some aspects of deviations. Under the affine parametric uncertainty, we can translate the primitive uncertainties so that their means are zeros. For the subset of independently distributed primitive uncertainties, we will use the forward and backward deviations, which were recently introduced by Chen, Sim and Sun [23].

Definition 3. Given a random variable $\tilde{z}$ with zero mean, the forward deviation is defined as

$$\sigma_f(\tilde{z}) \triangleq \sup_{\beta>0} \left\{ \sqrt{2\ln(\mathrm{E}(\exp(\beta\tilde{z})))/\beta^2} \right\} \tag{3.1}$$

and backward deviation is defined as

$$\sigma_b(\tilde{z}) \triangleq \sup_{\beta>0} \left\{ \sqrt{2\ln(\mathrm{E}(\exp(-\beta\tilde{z})))/\beta^2} \right\}. \tag{3.2}$$

Given a sequence of independent samples, we can essentially estimate the magnitude of the deviation measures from (3.1) and (3.2). Some of the properties of the deviation measures include:

Proposition 2. (Chen, Sim and Sun [23])

Let σ, p and q be respectively the standard, forward and backward deviations of a random variable, $\tilde{z}$ with zero mean.

(a) Then $p \geq \sigma$ and $q \geq \sigma$. If $\tilde{z}$ is normally distributed, then $p = q = \sigma$.

(b)

$$\mathrm{P}(\tilde{z} \geq \beta p) \leq \exp(-\beta^2/2);$$

$$\mathrm{P}(\tilde{z} \leq -\beta q) \leq \exp(-\beta^2/2).$$

(c) For all $\beta \geq 0$,

$$\ln \mathrm{E}(\exp(\beta \tilde{z})) \leq \tfrac{\beta^2 p^2}{2};$$

$$\ln \mathrm{E}(\exp(-\beta \tilde{z})) \leq \tfrac{\beta^2 q^2}{2}.$$

Proposition 2(a) shows that the forward and backward deviations are no less than the standard deviation of the underlying distribution, and under normal distribution, these two values coincide with the standard deviation. As exemplified in Proposition 2(b), the deviation measures provide an easy bound on the distributional tails. Chen, Sim and Sun ([23]) show that the new deviation measures provide tighter approximation of probabilistic bounds compared to standard deviations. This information, whenever available, enable us to improve upon the solutions of the approximation.

When only the support of the distributions are available, Chen, Sim and Sun [23] show how to obtain upper bounds of the forward and backward deviation measures.

Theorem 4. (Chen, Sim and Sun [23]) If $\tilde{z}$ has zero mean and distributed in

$[-\underline{z}, \bar{z}]$, $\underline{z}, \bar{z} > 0$, then

$$\sigma_f(\tilde{z}) \leq \bar{\sigma}_f(\tilde{z}) = \frac{\underline{z} + \bar{z}}{2}\sqrt{g\left(\frac{\underline{z} - \bar{z}}{\underline{z} + \bar{z}}\right)}$$

and

$$\sigma_b(\tilde{z}) \leq \bar{\sigma}_b(\tilde{z}) = \frac{\underline{z} + \bar{z}}{2}\sqrt{g\left(\frac{\bar{z} - \underline{z}}{\underline{z} + \bar{z}}\right)},$$

where

$$g(\mu) = 2\max_{s>0}\left\{\frac{\phi_\mu(s) - \mu s}{s^2}\right\},$$

and

$$\phi_\mu(s) = \ln\left(\frac{e^s + e^{-s}}{2} + \frac{e^s - e^{-s}}{2}\mu\right).$$

Moreover the bounds are tight.

Note that the forward and backward deviations may be infinite for heavier tailed distributions. Despite the stringent assumption, the advantage of using the forward and backward deviations is the ability to capture distributional asymmetry and stochastic independence, while keeping the resultant optimization model computationally amicable. The interested reader may refer to Natarajan et al. [43] for the computational experience of using the forward and backward deviations in minimizing the Value-at-Risk of a portfolio, which gives surprisingly good out-of-sample performance on real data.

Assumption 1. We assume that the uncertainties $\{\tilde{z}_j\}_{j=1:N}$ are zero mean random variables, with finite positive definite covariance matrix, $\boldsymbol{\Sigma}$ and support $\mathcal{W} = [-\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}}]$, $\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}} \in (0, \infty]^N$. Of the N primitive uncertainties, the first I random variables, that is, $\tilde{z}_j$, $j = 1, \ldots, I$ are stochastically independent. Moreover, the corresponding forward and backward deviations are finite and given by $p_j = \sigma_f(\tilde{z}_j) > 0$ and $q_j = \sigma_b(\tilde{z}_j) > 0$ respectively for $j = 1, \ldots, I$. We may also use the deviation bounds in Theorem 4. We denote $\boldsymbol{P} = \mathrm{diag}(p_1, \ldots, p_I)$ and $\boldsymbol{Q} = \mathrm{diag}(q_1, \ldots, q_I)$.

In practice, these parameters are, at best, estimated values. Moreover, the forward and backward deviations are harder to estimate compared to standard deviations in the sense that we may require more samples to achieve the same relative accuracy. It is fair to say that the effect of their estimation errors on the optimization problem has not been fully understood. As proposed in classical robust optimization, one possibility to address these estimation errors is to build uncertainty sets around these parameters. See for instance, Ben-Tal and Nemirovski [6], Bertsimas and Sim [14] and Goldfarb and Iyengar [32]. For simplicity, we assume in this thesis that the exact parameters are given.

Similar uncertainty models have been defined in Chen, Sim and Sun [23] and Chen et al. [24]. While the uncertainty model proposed in the former focuses on only independent primitive uncertainties with known support, forward and backward deviation measures, the uncertainty model proposed in the latter discards independence and assumes known support and covariance of the primitive uncertainties. Hence, Assumption 1 encompasses both models discussed in Chen, Sim and Sun [23] and Chen et al. [24].

Under Assumption 1, it is evident that $\boldsymbol{h}^0$, for instance, represents the mean of $\boldsymbol{h}(\tilde{\boldsymbol{z}})$ and $\boldsymbol{h}^j$ represents the magnitude and direction associated with the primitive uncertainty, $\tilde{z}_j$. Assumption 1, provides a flexibility of incorporating a subset of mutually independent random variables, which can lead better evaluation of the objective function. For instance, if $\tilde{\boldsymbol{h}}$ is multivariate normally distributed with mean $\boldsymbol{h}^0$ and covariance, $\boldsymbol{\Sigma}$, then we can decompose $\tilde{\boldsymbol{h}}$ into primitive uncertainties that are stochastically independent as follows

$$\tilde{\boldsymbol{h}} = \boldsymbol{h}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}^0 + \boldsymbol{\Sigma}^{1/2}\tilde{\boldsymbol{z}}.$$

To fit into the affine parametric uncertainty and Assumption 1, we can assign the vector $\boldsymbol{h}^j$ to the jth column of $\boldsymbol{\Sigma}^{1/2}$. Moreover, $\tilde{\boldsymbol{z}}$ has stochastically independent entries with covariance equal to the identity matrix, infinite support and unit forward and backward deviations; see Proposition 2(a).

3.2 Approximation of $\mathrm{E}((y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+)$ and $CVaR_{1-\gamma}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})$

Although the CVaR measure,

$$CVaR_{1-\gamma}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \min_{\beta} \left(\beta + \frac{\mathrm{E}\left((y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} - \beta)^+\right)}{\gamma}\right)$$

is convex in the variable $(y^0, \boldsymbol{y})$, it does not necessarily lead to a tractable optimization problem. The key difficulty lies in the evaluation of the expectation, $\mathrm{E}((\cdot)^+)$, which involves multi-dimension integration. Such evaluation is typically analytically prohibitive when the dimension of the integration exceeds four. Hence, providing bounds on $\mathrm{E}((y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+)$ is pivotal in developing tractable approximations of the CVaR measure. We next present various ways of bounding $\mathrm{E}((y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+)$ and $CVaR_{1-\gamma}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})$ as follows:

Theorem 5. Assuming $\tilde{\boldsymbol{z}}$ follows Assumption 1, the following functions $\pi^i(y^0, \boldsymbol{y})$, $i \in \{1, \ldots, 5\}$ are upper bounds of $\mathrm{E}\left((y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right)$. Likewise, the following functions,

$$\eta^i_{1-\gamma}(y^0, \boldsymbol{y}) \triangleq \min_{\beta} \left(\beta + \frac{1}{\gamma}\pi^i(y^0 - \beta, \boldsymbol{y})\right) \qquad i \in \{1, \ldots, 5\}$$

are the upper bounds of $CVaR_{1-\gamma}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}))$.

(a)

$$\begin{aligned}
\pi^1(y^0, \boldsymbol{y}) &\triangleq \left(y^0 + \max_{\boldsymbol{z}\in\mathcal{W}} \boldsymbol{z}'\boldsymbol{y}\right)^+ \\
&= \min_{r,\boldsymbol{s},\boldsymbol{t}} \left\{r \mid r \geq y^0 + \boldsymbol{s}'\overline{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}}, \boldsymbol{s} - \boldsymbol{t} = \boldsymbol{y}, \boldsymbol{s}, \boldsymbol{t} \geq \boldsymbol{0}, r \geq 0\right\}, \\
\eta^1_{1-\gamma}(y^0, \boldsymbol{y}) &\triangleq y^0 + \max_{\boldsymbol{z}\in\mathcal{W}} \boldsymbol{y}'\boldsymbol{z} \\
&= y^0 + \min_{\boldsymbol{s},\boldsymbol{t}\geq\boldsymbol{0}} \left\{\boldsymbol{s}'\overline{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}} \mid \boldsymbol{s} - \boldsymbol{t} = \boldsymbol{y}\right\}.
\end{aligned}$$

The bound $\pi^1(y^0, \boldsymbol{y})$ is tight whenever $y^0 + \boldsymbol{y}'\boldsymbol{z} \leq 0$ for all $\boldsymbol{z} \in \mathcal{W}$.

(b)

$$\begin{aligned}
\pi^2(y^0, \boldsymbol{y}) &\triangleq y^0 + \left(-y^0 + \max_{\boldsymbol{z}\in\mathcal{W}} (-\boldsymbol{y})'\boldsymbol{z}\right)^+ \\
&= \min_{r,\boldsymbol{s},\boldsymbol{t}} \left\{r \mid r \geq \boldsymbol{s}'\overline{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}}, \boldsymbol{s} - \boldsymbol{t} = -\boldsymbol{y}, \boldsymbol{s}, \boldsymbol{t} \geq \boldsymbol{0}, r \geq y^0\right\}, \\
\eta^2_{1-\gamma}(y^0, \boldsymbol{y}) &\triangleq y^0 + (1/\gamma - 1) \max_{\boldsymbol{z}\in\mathcal{W}} (-\boldsymbol{y})'\boldsymbol{z} \\
&= y^0 + (1/\gamma - 1) \min_{\boldsymbol{s},\boldsymbol{t}\geq\boldsymbol{0}} \left\{\boldsymbol{s}'\overline{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}} \mid \boldsymbol{s} - \boldsymbol{t} = -\boldsymbol{y}\right\}.
\end{aligned}$$

The bound $\pi^2(y^0, \boldsymbol{y})$ is tight whenever $y^0 + \boldsymbol{y}'\boldsymbol{z} \geq 0$ for all $\boldsymbol{z} \in \mathcal{W}$.

(c)

$$\begin{aligned}
\pi^3(y^0, \boldsymbol{y}) &\triangleq \tfrac{1}{2}y^0 + \tfrac{1}{2}\sqrt{y^{0\,2} + \boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}}, \\
\eta^3_{1-\gamma}(y^0, \boldsymbol{y}) &\triangleq y^0 + \sqrt{\frac{1-\gamma}{\gamma}}\sqrt{\boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}}
\end{aligned}$$

(d)

$$\pi^4(y^0, \boldsymbol{y}) \triangleq \begin{cases} \inf_{\mu>0}\left\{\frac{\mu}{e}\exp\left(\frac{y^0}{\mu} + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2}\right)\right\} & \text{if } y_j = 0\ \forall j = I+1,\ldots,N \\ +\infty & \text{otherwise} \end{cases}$$

$$\eta^4_{1-\gamma}(y^0, \boldsymbol{y}) \triangleq \begin{cases} y^0 + \sqrt{-2\ln\gamma}\|\boldsymbol{u}\|_2 & \text{if } y_j = 0\ \forall j = I+1,\ldots,N \\ +\infty & \text{otherwise} \end{cases},$$

where $u_j = \max\{p_j y_j, -q_j y_j\}$, $j = 1, \ldots, I$.

(e)

$$\pi^5(y^0, \boldsymbol{y}) \triangleq \begin{cases} y^0 + \inf_{\mu>0}\left\{\frac{\mu}{e}\exp\left(-\frac{y^0}{\mu} + \frac{\|\boldsymbol{v}\|_2^2}{2\mu^2}\right)\right\} & \text{if } y_j = 0\ \forall j = I+1,\ldots,N \\ +\infty & \text{otherwise} \end{cases}$$

$$\eta^5_{1-\gamma}(y^0, \boldsymbol{y}) \triangleq \begin{cases} y^0 + \frac{1-\gamma}{\gamma}\sqrt{-2\ln(1-\gamma)}\|\boldsymbol{v}\|_2 & \text{if } y_j = 0\ \forall j = I+1,\ldots,N \\ +\infty & \text{otherwise} \end{cases},$$

where $v_j = \max\{-p_j y_j, q_j y_j\}$, $j = 1, \ldots, I$.

The proof is shown in Appendix .1.

Remark : The first and second bounds in Proposition 5 are derived from the support of the primitive uncertainties. Observe that the first bound is independent of the parameter γ. The third bound is derived from the covariance of the primitive uncertainties. The last two bounds act upon primitive uncertainties that are stochastically independent.

To understand the conservativeness of the approximation, we compare the bounds of $CVaR_{1-\gamma}(\tilde{z})$, where $\tilde{z}$ is standard normally distributed. Figure

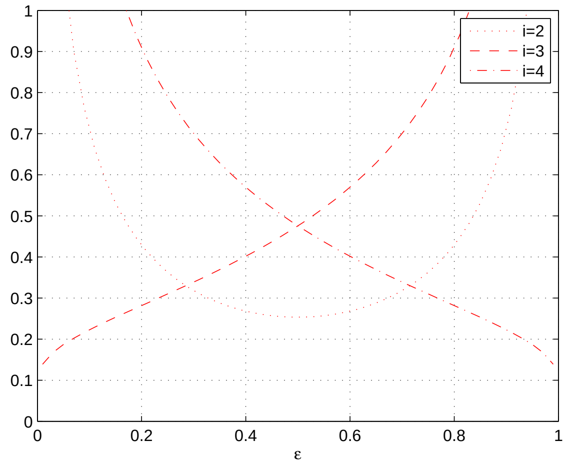


Fig. 3.1: Plot of $\rho_i(\gamma)$ against γ for $i = 3, 4$ and 5, defined in Proposition 5.

3.1 compares the approximation ratios given by

$$\rho_i(\gamma) = \frac{\eta^i_{1-\gamma}(0,1) - CVaR_{1-\gamma}(\tilde{z})}{CVaR_{1-\gamma}(\tilde{z})}, \qquad i = 3, 4, 5$$

It is clear that none of the bounds dominate another across $\gamma \in (0, 1)$. For small values of γ, the bound $\eta^4_{1-\gamma}(0, 1)$ is the tightest, while at high values, $\eta^5_{1-\gamma}(0, 1)$ dominates. At mid-range, $\eta^3_{1-\gamma}(0, 1)$ gives the best bound. Hence, this motivate us to integrate the best of all bounds to achieve the tightest approximation. The unified approximation in Figure 3.1 achieve a worst case approximation error of 33% at $\gamma = 0.2847$ and $\gamma = 0.7153$. We next show how to unify these bounds.

Theorem 6. **(a)** Let $\mathcal{L} \subset \{1, 2, \ldots, 5\}$. Define

$$\begin{aligned} \pi^{\mathcal{L}}(y^0, \boldsymbol{y}) \triangleq \; & \min_{y_i^0, \boldsymbol{y}_i} \; \sum_{i \in \mathcal{L}} \pi^i(y_i^0, \boldsymbol{y}_i) \\ & \text{s.t.} \quad \sum_{i \in \mathcal{L}} y_i^0 = y^0 \\ & \qquad \sum_{i \in \mathcal{L}} \boldsymbol{y}_i = \boldsymbol{y}. \end{aligned}$$

Then for all $(y^0, \boldsymbol{y})$

$$\mathrm{E}\left((y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) \leq \pi^{\mathcal{L}}(y^0, \boldsymbol{y}) \leq \min_{i \in \mathcal{L}} \{\pi^i(y^0, \boldsymbol{y}))\} \tag{3.3}$$

(b) Let

$$\eta^{\mathcal{L}}_{1-\gamma}(y^0, \boldsymbol{y}) \triangleq \min_{\beta} \left(\beta + \frac{1}{\gamma} \pi^{\mathcal{L}}(y^0 - \beta, \boldsymbol{y})\right)$$

or equivalently

$$\begin{aligned} \eta^{\mathcal{L}}_{1-\gamma}(y^0, \boldsymbol{y}) \triangleq \; & \min_{y_i^0, \boldsymbol{y}_i} \; \sum_{i \in \mathcal{L}} \eta^i_{1-\gamma}(y_i^0, \boldsymbol{y}_i) \\ & \text{s.t.} \quad \sum_{i \in \mathcal{L}} y_i^0 = y^0 \\ & \qquad \sum_{i \in \mathcal{L}} \boldsymbol{y}_i = \boldsymbol{y}. \end{aligned}$$

Then for all $(y^0, \boldsymbol{y})$ and $\gamma \in (0, 1)$

$$CVaR_{1-\gamma}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) \leq \eta^{\mathcal{L}}_{1-\gamma}(y^0, \boldsymbol{y}) \leq \min_{i \in \mathcal{L}} \{\eta^i_{1-\gamma}(y^0, \boldsymbol{y})\} \tag{3.4}$$

Proof : (a) To show the upper bound, we note that

$$\begin{aligned}
& \sum_{i\in\mathcal{L}} \pi^i(y_i^0, \boldsymbol{y}_i) \\
\geq\ & \sum_{i\in\mathcal{L}} \mathrm{E}\left((y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^+\right) && \text{Theorem 5} \\
\geq\ & \mathrm{E}\left(\left(\textstyle\sum_{i\in\mathcal{L}}(y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})\right)^+\right) && \text{Subadditvity} \\
=\ & \mathrm{E}\left((y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right).
\end{aligned}$$

Finally, to show that $\pi^{\mathcal{L}}(y^0, \boldsymbol{y}) \leq \pi^i(y^0, \boldsymbol{y})$, $i = 1, \ldots, 5$, let

$$(y_r^0, \boldsymbol{y}_r) = \begin{cases} (y^0, \boldsymbol{y}) & \text{if } r = i \\ (0, \mathbf{0}) & \text{otherwise} \end{cases} \qquad \text{for } r \in \mathcal{L}.$$

Hence,

$$\pi^r(y_r^0, \boldsymbol{y}_r) = \begin{cases} \pi^r(y^0, \boldsymbol{y}) & \text{if } r = i \\ 0 & \text{otherwise} \end{cases} \qquad \text{for } r \in \mathcal{L},$$

and therefore

$$\pi^{\mathcal{L}}(y^0, \boldsymbol{y}) \leq \sum_{i\in\mathcal{L}} \pi^i(y_i^0, \boldsymbol{y}_i) = \pi^i(y^0, \boldsymbol{y}).$$

(b) Observe that

$$\begin{aligned}
& \eta^{\mathcal{L}}_{1-\gamma}(y^0, \boldsymbol{y}) \\
= & \min_{\beta} \left(\beta + \frac{\pi^{\mathcal{L}}(y^0 - \beta, \boldsymbol{y})}{\gamma} \right) \\
= & \min_{\beta, \boldsymbol{\beta}, y_i^0, \boldsymbol{y}_i, \forall i} \left(\beta + \sum_{i \in \mathcal{L}} \left(\frac{\pi^i(y_i^0 - \beta_i, \boldsymbol{y}_i)}{\gamma} \right) \mid \sum_{i \in \mathcal{L}} \boldsymbol{y}_i = \boldsymbol{y}, \sum_{i \in \mathcal{L}} y_i^0 = y^0, \sum_{i \in \mathcal{L}} \beta_i = \beta \right) \\
= & \min_{y_i^0, \boldsymbol{y}_i, \forall i} \left(\sum_{i \in \mathcal{L}} \underbrace{\min_{\beta_i} \left(\beta_i + \frac{\pi^i(y_i^0 - \beta_i, \boldsymbol{y}_i)}{\gamma} \right)}_{= \eta^i_{1-\gamma}(y_i^0, \boldsymbol{y}_i)} \mid \sum_{i \in \mathcal{L}} \boldsymbol{y}_i = \boldsymbol{y}, \sum_{i \in \mathcal{L}} y_i^0 = y^0 \right).
\end{aligned}$$

Finally, the inequalities (3.4) are trivial consequence of the inequalities (3.3).

■

Remark : Note that in the presence of stochastically dependent primitive uncertainties and unbounded support, all the bounds, except for the third, of Theorem 5 can become infinite. However, such trivial bound is avoided in the unified bound.

From Theorem 5(a), the epigraph of the unified bound of $\mathrm{E}((y^0+\boldsymbol{y}'\tilde{\boldsymbol{z}})^+)$, $\pi^{\{1,2,\ldots,5\}}(y^0,\boldsymbol{y}) \leq \omega$ can be expressed as follows:

$$
\begin{array}{ll}
\exists r_i, y_i^0 \in \Re, \boldsymbol{y}_i, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{d}, \boldsymbol{h} \in \Re^N, i=1,\ldots,5, \boldsymbol{u}, \boldsymbol{v} \in \Re^I, \text{ such that} & \\
r_1+r_2+r_3+r_4+r_5 \leq \omega & \\
y_1^0 + \boldsymbol{s}'\bar{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}} \leq r_1 & \\
0 \leq r_1 & \\
\boldsymbol{s}-\boldsymbol{t} = \boldsymbol{y}_1 & \\
\boldsymbol{s}, \boldsymbol{t} \geq 0 & \\
\boldsymbol{d}'\bar{\boldsymbol{z}} + \boldsymbol{h}'\underline{\boldsymbol{z}} \leq r_2 & \\
y_2^0 \leq r_2 & \\
\boldsymbol{d}-\boldsymbol{h} = -\boldsymbol{y}_2 & \\
\boldsymbol{d}, \boldsymbol{h} \geq 0 & \\
\frac{1}{2}y_3^0 + \frac{1}{2}\|(y_3^0, \boldsymbol{\Sigma}^{1/2}\boldsymbol{y}_3)\|_2 \leq r_3 & \\
\inf_{\mu>0} \frac{\mu}{e}\exp\left(\frac{y^0}{\mu} + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2}\right) \leq r_4 & \\
u_j \geq p_j y_4^j, u_j \geq -q_j y_4^j & \forall j=1,\ldots,I \\
y_4^j = 0 & \forall j = I+1,\ldots,N \\
y^0 + \inf_{\mu>0} \frac{\mu}{e}\exp\left(-\frac{y^0}{\mu} + \frac{\|\boldsymbol{v}\|_2^2}{2\mu^2}\right) \leq r_5 & \\
v_j \geq q_j y_5^j, v_j \geq -p_j y_5^j & \forall j=1,\ldots,I \\
y_5^j = 0 & \forall j = I+1,\ldots,N \\
y_1^0+y_2^0+y_3^0+y_4^0+y_5^0 = y^0 & \\
\boldsymbol{y}_1+\boldsymbol{y}_2+\boldsymbol{y}_3+\boldsymbol{y}_4+\boldsymbol{y}_5 = \boldsymbol{y}. &
\end{array}
\tag{3.5}
$$

Due to the presence of the constraint, $\inf_{\mu>0}\mu\exp\left(\frac{a}{\mu}+\frac{b^2}{\mu^2}\right) \leq c$, the set of constraints in (3.5) is not exactly second order cone representable (see

Ben-Tal and Nemirovski [10]). Fortunately, using a few number second order cones, we can accurately approximate such constraint to within the precision of the solver. We present the second order cone approximation in Appendix .2.

Similarly, from Theorem 5(b), the epigraph of the unified CVaR approximation, $\eta_{1-\gamma}^{\{1,2,\ldots,5\}}(y^0, \boldsymbol{y}) \leq \omega$ is second order cone representable as follows:

$$
\begin{array}{ll}
\exists r_i, y_i^0 \in \Re, \boldsymbol{y}_i, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{d}, \boldsymbol{h} \in \Re^N, i = 1, \ldots, 5, \boldsymbol{u}, \boldsymbol{v} \in \Re^I \text{ such that} & \\
r_1 + r_2 + r_3 + r_4 + r_5 \leq \omega & \\
y_1^0 + \boldsymbol{s}'\bar{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}} \leq r_1 & \\
\boldsymbol{s}, \boldsymbol{t} \geq 0 & \\
\boldsymbol{s} - \boldsymbol{t} = \boldsymbol{y}_1 & \\
y_2^0 + (1/\gamma - 1)\boldsymbol{d}'\bar{\boldsymbol{z}} + (1/\gamma - 1)\boldsymbol{h}'\underline{\boldsymbol{z}} \leq r_2 & \\
\boldsymbol{d} - \boldsymbol{h} = -\boldsymbol{y}_2 & \\
\boldsymbol{d}, \boldsymbol{h} \geq 0 & \\
y_3^0 + \sqrt{\frac{1-\gamma}{\gamma}} \|\boldsymbol{\Sigma}^{1/2}\boldsymbol{y}_3\|_2 \leq r_3 & \\
y_4^0 + \sqrt{-2\ln(\gamma)} \|\boldsymbol{u}\|_2 \leq r_4 & \\
u_j \geq p_j y_4^j, u_j \geq -q_j y_4^j & \forall j = 1, \ldots, I \\
y_4^j = 0 & \forall j = I+1, \ldots, N \\
y_5^0 + \frac{1-\gamma}{\gamma}\sqrt{-2\ln(1-\gamma)} \|\boldsymbol{v}\|_2 \leq r_5 & \\
v_j \geq q_j y_5^j, v_j \geq -p_j y_5^j & \forall j = 1, \ldots, I \\
y_5^j = 0 & \forall j = I+1, \ldots, N \\
y_1^0 + y_2^0 + y_3^0 + y_4^0 + y_5^0 = y^0 & \\
\boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 + \boldsymbol{y}_4 + \boldsymbol{y}_5 = \boldsymbol{y}. &
\end{array}
$$

It is rather surprising to note that while the epigraph of the function $\pi^{\mathcal{L}}(\cdot,\cdot)$ is approximately second-order cone representable, the epigraph of $\eta^{\mathcal{L}}_{1-\gamma}(\cdot,\cdot)$, is fully second-order cone representable.

3.3 Decision Rule Approximation of Recourse

Depending on the distribution of $\tilde{\boldsymbol{z}}$, the second stage recourse decisions, $\boldsymbol{u}(\tilde{\boldsymbol{z}})$ and $\boldsymbol{y}(\tilde{\boldsymbol{z}})$ can be very large or even infinite. Moreover, since we do not specify the exact distributions of the primitive uncertainties, it would not be possible to obtain an optimal recourse decision. To enable us to formulate a tractable problem in which we could derive an upper bound of Model (2.23), we first adopt the linear decision rule used in Ben-Tal et al. [5] and Chen, Sim, and Sun [23]. We restrict $\boldsymbol{u}(\tilde{\boldsymbol{z}})$ and $\boldsymbol{y}(\tilde{\boldsymbol{z}})$ to be affinely dependent on the primitive uncertainties, that is

$$\boldsymbol{u}(\tilde{\boldsymbol{z}}) = \boldsymbol{u}^0 + \textstyle\sum_{j=1}^N \boldsymbol{u}^j \tilde{z}_j, \ \boldsymbol{y}(\tilde{\boldsymbol{z}}) = \boldsymbol{y}^0 + \textstyle\sum_{j=1}^N \boldsymbol{y}^j \tilde{z}_j. \tag{3.6}$$

Under linear decision rule, the following constraint

$$\boldsymbol{B}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}^j + \boldsymbol{Y}\boldsymbol{y}^j = \boldsymbol{h}^j \qquad j = 0,\ldots,N$$

is a sufficient condition to satisfy the affine constraint involving recourse variables in Model (2.23). Moreover, since the support of $\tilde{\boldsymbol{z}}$ is $\mathcal{W} = [-\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}}]$, an inequality constraint $y_i(\tilde{\boldsymbol{z}}) \geq 0$ in Model (2.23) is the same as the robust counterpart

$$y_i^0 + \sum_{j=1}^{N} y_i^j z_j \geq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W},$$

which is representable by the following linear inequalities

$$y_i^0 \geq \sum_{j=1}^{N} (\underline{z}_j s_j^i + \bar{z}_j t_j^i)$$

for some $\boldsymbol{s}^i, \boldsymbol{t}^i \geq \boldsymbol{0}$ satisfying $s_j^i - t_j^i = y_i^j$, $j = 1, \ldots, N$. As for the aspiration level prospect, we let

$$w(\tilde{\boldsymbol{z}}) = w_0 + \sum_{j=1}^{N} w_j \tilde{z}_j, \tag{3.7}$$

where

$$w_j = \boldsymbol{c}^{j\prime}\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}^j + \boldsymbol{d_y}'\boldsymbol{y}^j - \tau^j \qquad j = 0, \ldots, N, \tag{3.8}$$

so that

$$w(\tilde{\boldsymbol{z}}) = \boldsymbol{c}(\tilde{\boldsymbol{z}})'\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}(\tilde{\boldsymbol{z}}) + \boldsymbol{d_y}'\boldsymbol{y}(\tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}}).$$

Hence, applying the bound on the CVaR measure at the objective function, we have

$$CVaR_{1-\gamma}(w(\tilde{\boldsymbol{z}})) \leq \eta_{1-\gamma}^{\mathcal{L}}(w_0, \boldsymbol{w})$$

where we use $\boldsymbol{w}$ to denote the vector with elements $w_j,\ j = 1, \ldots, N$. Putting these together, we solve the following problem, which is an SOCP.

$$
\begin{array}{llll}
Z_{LDR}(\gamma) = & \min\limits_{\boldsymbol{x},\boldsymbol{u}^j,\boldsymbol{y}^j,w_0,\boldsymbol{w}} & \eta^{\mathcal{L}}_{1-\gamma}(w_0, \boldsymbol{w}) & \\
& \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& & w_j = \boldsymbol{c}^{j\prime}\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}^j + \boldsymbol{d_y}'\boldsymbol{y}^j - \tau^j & j = 0, \ldots, N. \\
& & \boldsymbol{B}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}^j + \boldsymbol{Y}\boldsymbol{y}^j = \boldsymbol{h}^j & j = 0, \ldots, N. \\
& & y_i^0 + \sum_{j=1}^N y_i^j z_j \geq \boldsymbol{0} & \forall \boldsymbol{z} \in \mathcal{W}, i = 1, \ldots, n_3 \\
& & \boldsymbol{x} \geq 0. &
\end{array}
\tag{3.9}
$$

Theorem 7. Let $(\boldsymbol{x}, \boldsymbol{u}^0, \ldots, \boldsymbol{u}^N, \boldsymbol{y}^0, \ldots, \boldsymbol{y}^N)$ be an optimal solution of Model (3.9). The solution $\boldsymbol{x}$ and linear decision rules $\boldsymbol{u}(\tilde{\boldsymbol{z}})$ and $\boldsymbol{y}(\tilde{\boldsymbol{z}})$ defined in the equations (3.6), are feasible in the subproblem (2.23). Moreover,

$$Z(\gamma) \leq Z_{LDR}(\gamma).$$

Deflected linear decision rule

The most common type of stochastic optimization problems is one of complete recourse, which is defined on the matrix $(\boldsymbol{U}, \boldsymbol{Y})$ such that for any $\boldsymbol{t}$, there exists $(\boldsymbol{u}, \boldsymbol{y})$, $\boldsymbol{y} \geq \boldsymbol{0}$ satisfying $\boldsymbol{U}\boldsymbol{u} + \boldsymbol{Y}\boldsymbol{y} = \boldsymbol{t}$. It is easy to see in Model (2.23) that complete recourse problem always admits a feasible recourse, however, it may not necessarily be one of linear decision rule. Although linear decision rule leads to a tractable approximation of the recourse, Chen et al. [24] show that linear decision rules can be inadequate and can lead to infeasible instances even in complete recourse problems. To resolve such infeasibility, we adopt the *deflected linear decision rules* proposed by Chen et al. [24] as an improvement over linear decision rules. We first define the vector $\bar{\boldsymbol{d}}$ with elements

$$\begin{aligned} \bar{d}_i = \quad \min_{\boldsymbol{u},\boldsymbol{y}} \quad & \boldsymbol{d_u}'\boldsymbol{u} + \boldsymbol{d_y}'\boldsymbol{y} \\ \text{s.t.} \quad & \boldsymbol{U}\boldsymbol{u} + \boldsymbol{Y}\boldsymbol{y} = \boldsymbol{0} \\ & y_i = 1 \\ & \boldsymbol{y} \geq \boldsymbol{0}, \end{aligned} \tag{3.10}$$

where we denote $\bar{d}_i = \infty$ if the corresponding optimization problem is infeasible. For notational convenience, we define the sets

$$\mathcal{C} \triangleq \{i \;:\; \bar{d}_i < \infty, i = 1, \ldots, n_3\}, \qquad \bar{\mathcal{C}} \triangleq \{i = 1, \ldots, n_3\} \backslash \mathcal{C}.$$

For $i \in \mathcal{C}$, we define $(\bar{\boldsymbol{u}}^i, \bar{\boldsymbol{y}}^i)$ as the optimal solution of the corresponding optimization problem.

Note that if $\bar{d}_i < 0$, then given any feasible solution $\boldsymbol{u}$ and $\boldsymbol{y}$, the solution $\boldsymbol{u} + \kappa\bar{\boldsymbol{u}}^i$, and $\boldsymbol{y} + \kappa\bar{\boldsymbol{y}}^i$ will also be feasible, and that the objective will be reduced by $|\kappa\bar{d}_i|$. Hence, whenever a second stage decision is feasible, its objective will be unbounded from below. Therefore, it is reasonable to assume that $\bar{\boldsymbol{d}} \geq \mathbf{0}$.

Next, we present the model that achieves a better bound than Model (3.9). Let $\{1\} \subset \mathcal{L} \subset \{1, 2, \ldots, 5\}$. We define

$$
\begin{array}{rlll}
Z_{DLDR}(\gamma) = & \displaystyle\min_{\boldsymbol{x}, \boldsymbol{u}^j, \boldsymbol{y}^j, w_0, \boldsymbol{w}} & \eta^{\mathcal{L}}_{1-\gamma}(w_0, \boldsymbol{w}) + \frac{1}{\gamma} \sum_{i \in \mathcal{C}} \pi^{\mathcal{L}}(-y_i^0, -\boldsymbol{y}_i) \bar{d}_i & \\
& \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& & w_j = \boldsymbol{c}^{j\prime}\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}^j + \boldsymbol{d_y}'\boldsymbol{y}^j - \tau^j & j = 0, \ldots, N. \\
& & \boldsymbol{B}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}^j + \boldsymbol{Y}\boldsymbol{y}^j = \boldsymbol{h}^j & j = 0, \ldots, N. \\
& & y_i^0 + \sum_{j=1}^N y_i^j z_j \geq 0 & \forall \boldsymbol{z} \in \mathcal{W}, i \in \bar{\mathcal{C}} \\
& & \boldsymbol{x} \geq \mathbf{0}, &
\end{array}
\tag{3.11}
$$

in which $\boldsymbol{y}_i$ denotes the vector with elements y_i^j, $j = 1, \ldots, N$.

Theorem 8. Let $(\boldsymbol{x}, \boldsymbol{u}^0, \ldots, \boldsymbol{u}^N, \boldsymbol{y}^0, \ldots, \boldsymbol{y}^N)$ be an optimal solution of Model (3.11). The solution $\boldsymbol{x}$ and the corresponding deflected linear decision rule

$$
\begin{aligned}
\boldsymbol{u}(\tilde{\boldsymbol{z}}) &= \boldsymbol{u}^0 + \sum_{j=1}^N \boldsymbol{u}^j \tilde{z}_j + \sum_{i \in \mathcal{C}} \bar{\boldsymbol{u}}^i (y_i^0 + \boldsymbol{y}_i' \tilde{\boldsymbol{z}})^- \\
\boldsymbol{y}(\tilde{\boldsymbol{z}}) &= \boldsymbol{y}^0 + \sum_{j=1}^N \boldsymbol{y}^j \tilde{z}_j + \sum_{i \in \mathcal{C}} \bar{\boldsymbol{y}}^i (y_i^0 + \boldsymbol{y}_i' \tilde{\boldsymbol{z}})^-,
\end{aligned}
\tag{3.12}
$$

are feasible in the subproblem (2.23). Moreover,

$$
Z(\gamma) \leq Z_{DLDR}(\gamma) \leq Z_{LDR}(\gamma).
$$

Proof : Noting that

$$\boldsymbol{U}\bar{\boldsymbol{u}}^i + \boldsymbol{Y}\bar{\boldsymbol{y}}^i = \boldsymbol{0},$$

it is straightforward to verify that the recourse with deflected linear decision rule satisfies the affine constraints in Model (2.23). For $i \in \mathcal{C}$, we have $\bar{y}_i^i = 1$, hence, the nonnegativity condition holds at every i element of $\boldsymbol{y}(\tilde{\boldsymbol{z}})$. Besides, for $i \in \bar{\mathcal{C}}$, we have $y_i^0 + \sum_{j=1}^N y_i^j \tilde{z}_j \geq 0$. Therefore, since $\bar{\boldsymbol{y}}^j \geq \boldsymbol{0}$ for all $j \in \mathcal{C}$, the nonnegativity condition of $\boldsymbol{y}(\tilde{\boldsymbol{z}})$ holds at every i element, $i \in \bar{\mathcal{C}}$ as well. To show the bound, $Z(\gamma) \leq Z_{DLDR}(\gamma)$, we note that $\bar{d}_i = \boldsymbol{d}_u \bar{\boldsymbol{u}}_i + \boldsymbol{d}_y \bar{\boldsymbol{y}}_i$, $i \in \mathcal{C}$. Under the deflected linear decision rule, the aspiration level prospect becomes

$$\begin{aligned} & \boldsymbol{c}(\tilde{\boldsymbol{z}})'\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}(\tilde{\boldsymbol{z}}) + \boldsymbol{d_y}'\boldsymbol{y}(\tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}}) \\ = \quad & w(\tilde{\boldsymbol{z}}) + \sum_{i \in \mathcal{C}} \bar{d}_i (y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^-, \end{aligned}$$

where $w(\tilde{\boldsymbol{z}})$ is defined in Equations (3.7) and (3.8). We now evaluate the objective of Model (2.23) under the deflected linear decision rule as follows:

$$
\begin{aligned}
& CVaR_{1-\gamma}\Big(w(\tilde{\boldsymbol{z}}) + \sum_{i\in\mathcal{C}} \bar{d}_i (y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^-\Big) \\
= \;& \min_{\beta}\left\{\beta + \frac{1}{\gamma}\mathrm{E}\left(\Big(w(\tilde{\boldsymbol{z}}) + \sum_{i\in\mathcal{C}} \bar{d}_i (y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^- - \beta\Big)^+\right)\right\} \\
= \;& \min_{\beta}\left\{\beta + \frac{1}{\gamma}\mathrm{E}\left(\left(w(\tilde{\boldsymbol{z}}) + \sum_{i\in\mathcal{C}} \bar{d}_i \left((-y_i^0 - \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^+\right) - \beta\right)^+\right)\right\} \\
\le \;& \min_{\beta}\left\{\beta + \frac{1}{\gamma}\mathrm{E}\left(\Big(w(\tilde{\boldsymbol{z}}) - \beta\Big)^+\right) + \sum_{i\in\mathcal{C}} \frac{1}{\gamma}\mathrm{E}\left(\left(-y_i^0 - \boldsymbol{y}_i'\tilde{\boldsymbol{z}}\right)^+\right)\bar{d}_i\right\} \qquad (3.13) \\
= \;& CVaR_{1-\gamma}\left(w(\tilde{\boldsymbol{z}})\right) + \frac{1}{\gamma}\sum_{i\in\mathcal{C}} \mathrm{E}\left(\left(-y_i^0 - \boldsymbol{y}_i'\tilde{\boldsymbol{z}}\right)^+\right)\bar{d}_i \\
\le \;& \eta_{1-\gamma}^{\mathcal{L}}(w_0, \boldsymbol{w}) + \frac{1}{\gamma}\sum_{i\in\mathcal{C}} \pi^{\mathcal{L}}(-y_i^0, -\boldsymbol{y}_i)\bar{d}_i \\
= \;& Z_{DLDR}(\gamma),
\end{aligned}
$$

where the first inequality are due to $(x+a)^+ \le (x)^+ + a$, for all $a \ge 0$, and that $\bar{\boldsymbol{d}} \ge \boldsymbol{0}$. The last inequality is due to Theorems 6.

To prove the improvement over Model (3.9), we now consider an optimal solution of Model (3.9), $(\boldsymbol{x}, \boldsymbol{u}^0, \ldots, \boldsymbol{u}^N, \boldsymbol{y}^0, \ldots, \boldsymbol{y}^N)$. Clearly, the solution is feasible in the constraints of Model (3.11). From Theorems 5(a) and 6, the constraint $y_i^0 + \sum_{j=1}^N y_i^j z_j \ge 0,\ \forall \boldsymbol{z} \in \mathcal{W}$ enforced in Model (3.9) ensures that

$$
0 \le \pi^{\mathcal{L}}(-y_i^0, -\boldsymbol{y}_i) \le \pi^1(-y_i^0, -\boldsymbol{y}_i) = 0,
$$

for all $i \in \mathcal{C}$. Therefore, the solution of Model (3.9) yields the same objective as Model (3.11). Hence, $Z_{DLDR}(\gamma) \leq Z_{LDR}(\gamma)$. ■

Remark : Chen et al. [24] show that for complete recourse problems, $\bar{d}_i$ is finite for all $i = 1, \ldots, n_3$. Therefore, in such problems, there always exist a feasible recourse in the form of deflected linear decision rule. As such, the magnitude of improvement of deflected linear rule over linear decision rule can be arbitrarily large.

3.4 Example: Multi-product Newsvendor Problem

In our computation studies, we compare the solutions obtained from sampling approximation and deterministic approximation using robust optimization. In particular, we test whether our approach has the ability of finding meaningful solutions even in the absence of complete distribution information.

We consider a multi-product Newsvendor problem evaluated under the goal driven optimization framework. The classical multi-product Newsvendor problem was first introduced by Hadley and Whitin [33] and was extended by Ben-Daya and Raouf [17] and Lau and Lau [37]. These models utilize the risk-neutral objectives that maximize expected profits. Given a set of m products, we consider a simple risk-neutral multi-product Newsvendor problem,

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \sum_{i=1}^{m} \left\{ (p_i - c_i)x_i - (p_i - s_i)\mathrm{E}\left((x_i - \tilde{h}_i)^+ \right) \right\} \\ \text{s.t.} \quad & \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned} \tag{3.14}$$

where the terms are defined as follows:

c_i : unit purchasing cost

p_i : unit selling price

s_i : unit salvage value

$\tilde{h}_i$: stochastic demand

x_i : order quantity,

with $p_i > c_i > s_i$ for all products. Note that regardless of the dependency of products' demands, we can easily decompose Model (3.14) into m independent Newsvendor problems. Hence, we can analytically obtain the optimal solution of Model (3.14). Note that the formulation of Model (3.14) tacitly contains the following recourse problem

$$(x_i - \tilde{h}_i)^+ = \min_{y_i} \{ y_i : y_i \geq 0, y_i \geq x_i - \tilde{h}_i) \}.$$

Hence, putting it in standard stochastic optimization framework, we have

$$\begin{aligned}
\max_{\boldsymbol{x},\boldsymbol{y}(\cdot)} \quad & (\boldsymbol{p}-\boldsymbol{c})'\boldsymbol{x} - \sum_{i=1}^{m} \mathrm{E}(y_i(\tilde{\boldsymbol{h}})) \\
\text{s.t.} \quad & y_i(\tilde{\boldsymbol{h}}) - y_{m+i}(\tilde{\boldsymbol{h}}) = (p_i - s_i)(x_i - \tilde{h}_i) \quad i = 1,\ldots,m \\
& y_i(\tilde{\boldsymbol{h}}) \geq 0 \quad i = 1,\ldots,2m \\
& \boldsymbol{x} \geq \boldsymbol{0},
\end{aligned}$$

However, not all decision makers are comfortable with implementing the risk neutral solution. Given a target profit, τ, Sankarasubramanian and Kumaraswamy [56] proposed a single-product model that maximizes the probability of attaining the target. Likewise, Lau and Lau [36] and Li et al. [38] extended the model to only two products. These approaches rely on full assumption of demand distribution and are not analytically tractable for multi-products. Moreover, as we have discussed, maximizing probability does not take into account of the level of shortfall against the target objective.

We consider the goal driven optimization model as follows:

$$
\begin{aligned}
\max_{\gamma,\boldsymbol{x},\boldsymbol{y}(\cdot)} \quad & 1-\gamma \\
\text{s.t.} \quad & CVaR_{1-\gamma}\Big(\tau - (\boldsymbol{p}-\boldsymbol{c})'\boldsymbol{x} + \sum_{i=1}^{m} y_i(\tilde{\boldsymbol{h}})\Big) \le 0 \\
& y_i(\tilde{\boldsymbol{h}}) - y_{m+i}(\tilde{\boldsymbol{h}}) = (p_i - s_i)(x_i - \tilde{h}_i) \qquad i = 1,\ldots,m \\
& y_i(\tilde{\boldsymbol{h}}) \ge 0 \qquad i = 1,\ldots,2m \\
& \boldsymbol{x} \ge \boldsymbol{0}.
\end{aligned}
\tag{3.15}
$$

Using Algorithm 1, we reduce the problem (3.15) to solving a sequence of subproblems in the form of stochastic optimization problems with CVaR objectives as follows:

$$
\begin{aligned}
Z(\gamma) = \min_{\boldsymbol{x},\boldsymbol{y}(\cdot)} \quad & CVaR_{1-\gamma}\Big(\tau - (\boldsymbol{p}-\boldsymbol{c})'\boldsymbol{x} + \sum_{i=1}^{m} y_i(\tilde{\boldsymbol{h}})\Big) \\
\text{s.t.} \quad & y_i(\tilde{\boldsymbol{h}}) - y_{m+i}(\tilde{\boldsymbol{h}}) = (p_i - s_i)(x_i - \tilde{h}_i) \qquad i = 1,\ldots,m \\
& y_i(\tilde{\boldsymbol{h}}) \ge 0 \qquad i = 1,\ldots,2m \\
& \boldsymbol{x} \ge \boldsymbol{0}.
\end{aligned}
\tag{3.16}
$$

In the nominal test problem, we choose $c_i = 3, p_i = 5, s_i = 2$ for all products. The demands across products are uncorrelated. The distribution of each demand is unknown except for being a nonnegative random variable with mean $\mu_i = 100$ and standard deviation $\sigma_i = 10$. Hence,

$$\tilde{\boldsymbol{h}} = \boldsymbol{h}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}^0 + \sum_{j=1}^{m} \boldsymbol{h}^j \tilde{z}_j,$$

where $\boldsymbol{h}^0$ is a vector of 100s, and $\boldsymbol{h}^j$ is a vector with the jth element taking the value of ten and zero otherwise. Therefore, the primitive uncertainties, $\tilde{\boldsymbol{z}}$ have covariance being the identity matrix and support of $\tilde{z}_i$ being $[-10, \infty)$. Note that we do not utilize the forward and backward deviations in this experiment. To apply deflected linear decision rule, need to obtain $\bar{\boldsymbol{d}} \in \Re^{2m}$ as follows

$$\begin{aligned} \bar{d}_i = \min_{\boldsymbol{y}} \quad & \sum_{j=1}^{m} y_j \\ \text{s.t.} \quad & y_j - y_{m+j} = 0 \quad j = 1, \ldots, m \\ & y_i = 1 \\ & y_j \geq 0 \qquad\qquad i = 1, \ldots, 2m. \end{aligned}$$

Clearly, $\bar{d}_i = 1$ for all $i = 1, \ldots, 2m$. Hence, using deflected linear decision rule, we can obtain the upper bound of the subproblems (3.16) by solving the following problem:

$$
\begin{aligned}
Z_{DLDR}(\gamma) = \min_{\boldsymbol{x}, w_0, \boldsymbol{w}, \boldsymbol{y}^j} \quad & \eta_{1-\gamma}^{\{1,2,3\}}(w_0, \boldsymbol{w}) + \frac{1}{\gamma} \sum_{i=1}^{2m} \pi^{\{1,2,3\}}(-y_i^0, -\boldsymbol{y}_i) \\
\text{s.t.} \quad & y_i^0 - y_{m+i}^0 = (p_i - s_i)(x_i - h_i^0) & i = 1, \ldots, m \\
& y_i^j - y_{m+i}^0 = (p_i - s_i)(-h_i^j) & i = 1, \ldots, m, j = 1, \ldots, m \\
& w_0 = \tau - (\boldsymbol{p} - \boldsymbol{c})'\boldsymbol{x} + \sum_{i=1}^{m} y_i^0 \\
& w_j = \sum_{i=1}^{m} y_i^j & j = 1, \ldots, m \\
& \boldsymbol{x} \geq \boldsymbol{0},
\end{aligned}
\tag{3.17}
$$

where

$$
\begin{aligned}
\eta_{1-\gamma}^{\{1,2,3\}}(w_0, \boldsymbol{w}) = \min_{s, \boldsymbol{r}, y_i^0, \boldsymbol{y}_i} \quad & s \\
\text{s.t.} \quad & r_1 + r_2 + r_3 \leq s \\
& y_1^0 - \boldsymbol{y}_1' \underline{\boldsymbol{z}} \leq r_1 \\
& -\boldsymbol{y}_1 \geq 0 \\
& y_2^0 + (1/\gamma - 1)\boldsymbol{y}_2' \underline{\boldsymbol{z}} \leq r_2 \\
& \boldsymbol{y}_2 \geq 0 \\
& y_3^0 + \sqrt{\tfrac{1-\gamma}{\gamma}} \|\boldsymbol{y}_3\|_2 \leq r_3 \\
& y_1^0 + y_2^0 + y_3^0 = w_0 \\
& \boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 = \boldsymbol{w},
\end{aligned}
$$

and

$$
\begin{aligned}
\pi^{\{1,2,3\}}(y^0, \boldsymbol{y}) = \min_{s,\boldsymbol{r},y_i^0,\boldsymbol{y}_i} \quad & s \\
\text{s.t.} \quad & r_1 + r_2 + r_3 \le s \\
& y_1^0 - \boldsymbol{y}_1' \underline{\boldsymbol{z}} \le r_1 \\
& 0 \le r_1 \\
& -\boldsymbol{y}_1 \ge 0 \\
& \boldsymbol{y}_2' \underline{\boldsymbol{z}} \le r_2 \\
& y_2^0 \le r_2 \\
& \boldsymbol{y}_2 \ge 0 \\
& \tfrac{1}{2} y_3^0 + \tfrac{1}{2} \|(y_3^0, \boldsymbol{y}_3)\|_2 \le r_3 \\
& y_1^0 + y_2^0 + y_3^0 = y^0 \\
& \boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 = \boldsymbol{y},
\end{aligned}
$$

and $\underline{z}_j = 10$ for $j = 1, \ldots, m$. Therefore, the deterministic approximation of the subproblem using robust optimization has $2m$ second order cones in dimension $m + 2$ and one second order cone of dimension $m + 1$.

After obtaining the robust solution of the goal driven optimization model, we generate the profit profile on a sample size of $M = 500,000$ using various assumed distributions with the same mean and standard deviations. After obtaining the profit profiles, $u_1, \ldots, u_M$, we can estimate the shortfall aspiration level criterion as follows:

$$\hat{SALC} = 1 - \inf_{a>0} \frac{1}{aM} \sum_{k=1}^{M} (\tau - u_k + a)^+.$$

In our experiment, we consider two types of distributions: a normal distribution and a shifted exponential distribution with density function

$$f_{\tilde{h}_i}(x; \mu_i, \sigma_i) = \begin{cases} \frac{1}{\sigma_i} \exp\left(-\frac{1}{\sigma_i}(x - (\mu_i - \sigma_i))\right) & \text{if } x \geq \mu_i - \sigma_i \\ 0 & \text{otherwise,} \end{cases}$$

in which the mean and standard deviation are given by μ_i and σ_i respectively. While keeping the target profit τ proportional to m, we analyze the profit profile as we vary the number of products, m. After some experiments, we choose $\tau = 183m$ in order to obtain reasonably interesting profiles for m ranging from 5 to 30.

Figure 3.2 shows the profit profiles of two solutions: one that maximizes the expected profit and the other maximizes the shortfall aspiration level criterion. Indeed, the classical risk neutral model obtains a higher expected profit than the goal driven model. However, its risk of under performing against the target profit is substantially higher.

We next investigate the conservativeness of the solution obtained by

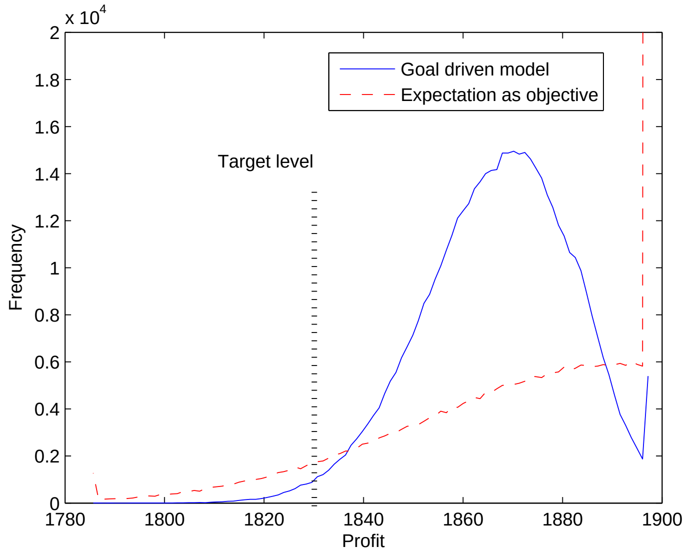


Fig. 3.2: Goal driven optimization versus maximizing expected profit ($m = 10$)

robust optimization against the solution obtained by sampling approximation using 1000 samples of the exact distribution. We formulate the problems using an in-house developed software, *PROF* (Platform for Robust Optimization Formulation). The Matlab based software is essentially an SOCP modeling environment that contains reusable functions for modeling multiperiod robust optimization using decision rules. We have implemented bounds for the CVaR measure and expected positivity of a weighted sum of random variables. The software calls upon CPLEX 10.0 to solve the underlying SOCP. It takes less than 0.5 seconds to solve Problem (3.17) of the size, $m = 30$. In contrast, it takes about 30 seconds to obtain the solution by sampling approximation using 1000 samples.

Since the stochastic optimization problem is one of complete recourse, and that the demand variances are relatively small, we expect sampling approximation to outperform the robust solution. In Figure 3.3, where the demands follows the shifted exponential distribution, the solution obtained by sampling approximation achieves higher shortfall aspiration level criterion. However, the gap against the robust solution tapered off as the number of products increases. In contrast, Figure 3.4, where the demands are normally distributed, shows that the shortfall aspiration level criterion obtained by the robust solution is only marginally lower than that of the solution obtained by sampling approximation. We observe that in these examples, the shortfall aspiration level criterion increases as the number of products, m increases. It is probably due to the increased risk pooling effect, which is consistent with our intuitions.

We have seen in this example that the solution obtained by sampling approximations is likely to outperform the robust solution if the demand distribution is correctly assumed. However, we find another interesting phenomenon. We use the solution obtained by sampling approximation based on the shifted exponential distribution and evaluate the shortfall aspiration level criteria based on a different distribution, in this case, a normal distribution with the same mean and standard deviation. Figure 3.5 suggests that the

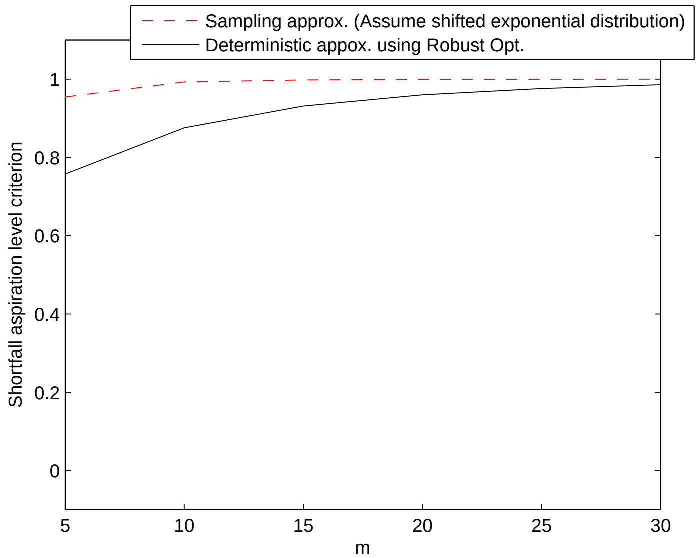


Fig. 3.3: Shortfall aspiration level criteria evaluated on shifted exponential distribution with sampling approximation using the same distribution.

robust solution can grossly outperform the solution obtained by sampled approximation using a different distribution with identical mean and standard deviation.

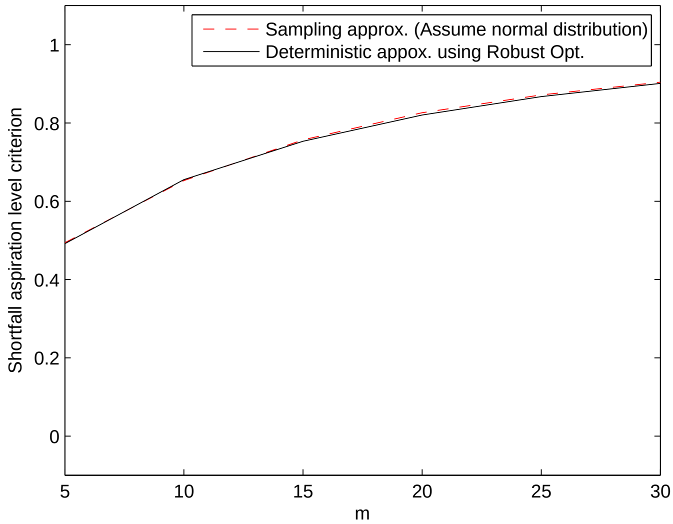


Fig. 3.4: Shortfall aspiration level criteria evaluated on normal distribution with sampling approximation using the same distribution.

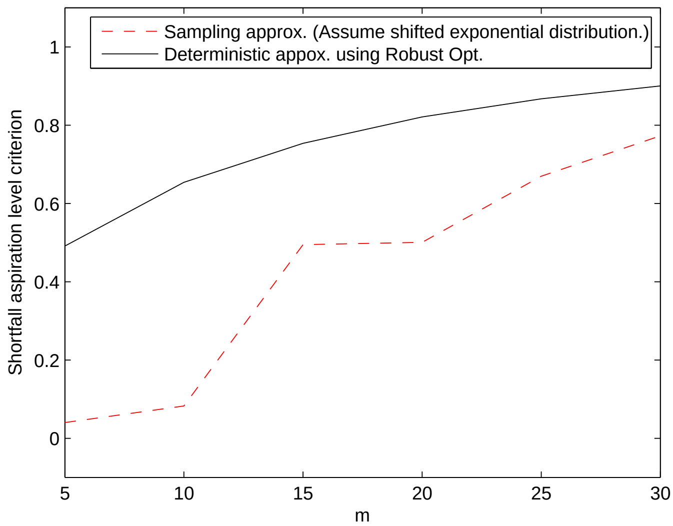


Fig. 3.5: Shortfall aspiration level criteria evaluated on normal distribution with sampling approximation using the shifted exponential distribution.

4. GOAL DRIVEN MODEL WITH PROBABILISTIC CONSTRAINT

4.1 Individual probabilistic Constraint

In this section, we review some of the tractable approximations of individual probabilistic constraint problems found in the literature, which are in the form of second order cone. For simplicity, we consider a linear individual probabilistic constraint

$$\mathrm{P}\Big(y(\tilde{\boldsymbol{z}}) \leq 0)\Big) \geq 1 - \epsilon, \tag{4.1}$$

where $y(\tilde{\boldsymbol{z}})$ are affinely dependent of $\tilde{\boldsymbol{z}}$,

$$y(\tilde{\boldsymbol{z}}) = y^0 + \sum_{j=1}^{N} y^j \tilde{z}_j,$$

where $(y^0, y^1, \ldots, y^N)$ are decision variables and $\epsilon \in (0, 1)$ being the given risk requirement. To illustrate the generality, we can represent the following probabilistic constraint problem

$$\mathrm{P}\Big(\boldsymbol{a}(\tilde{\boldsymbol{z}})'\boldsymbol{x} \geq b(\tilde{\boldsymbol{z}})\Big)\Big) \geq 1-\epsilon,$$

where

$$\begin{aligned} \boldsymbol{a}(\tilde{\boldsymbol{z}}) &= \boldsymbol{a}^0 + \sum_{j=1}^{N} \boldsymbol{a}^j \tilde{z}_j \\ b(\tilde{\boldsymbol{z}}) &= b^0 + \sum_{j=1}^{N} b^j \tilde{z}_j, \end{aligned}$$

by enforcing the following affine relations

$$y^j = -\boldsymbol{a}^{j\prime}\boldsymbol{x} + b^j \qquad \forall j = 0, \ldots, N.$$

Clearly, the constraint (4.1) is not necessarily convex in its decision variables, $(y^0, y^1, \ldots, y^N)$. For notational convenience, we denote $\boldsymbol{y} = (y^1, \ldots, y^N)$, so $y(\tilde{\boldsymbol{z}}) = y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}$. A step towards tractability is convexifying the probabilistic constraint (4.1) using the CVaR measure. The CVaR measure has been established by Shapiro and Nemirovski [44] as the tightest convex approximation of an individual probabilistic constraint problem. It has been well established that that suppose $(y^0, \boldsymbol{y})$ satisfies

$$CVaR_{1-\epsilon}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) \leq 0 \tag{4.2}$$

it also satisfies the probabilistic constraint (4.1). Moreover, the safeguarding constraint of (4.2) is convex in its decision variables, $(y^0, \boldsymbol{y})$. However, evaluation of the CVaR measure requires full knowledge of the underlying distribution, $\tilde{\boldsymbol{z}}$. Moreover, despite its convexity, even if the distributions of $\tilde{\boldsymbol{z}}$ is completely specified, it remains unclear how we can evaluate the CVaR measure precisely. To simplify the problem, we made the same assumptions of the uncertainties as the previous chapter.

Assumption 1: We assume that the primitive uncertainties $\{\tilde{z}_j\}_{j=1:N}$ are zero mean random variables, with covariance $\boldsymbol{\Sigma}$ and support $\mathcal{W} = [-\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}}]$. Of the N primitive uncertainties, the first I random variables, that is, $\tilde{z}_j$, $j = 1, \ldots, I$ are stochastically independent. Moreover, the corresponding forward and backward deviations given by $p_j = \sigma_f(\tilde{z}_j)$ and $q_j = \sigma_b(\tilde{z}_j)$ respectively for $j = 1, \ldots, I$, and we denote $\boldsymbol{P} = \text{diag}(p_1, \ldots, p_I)$ and $\boldsymbol{Q} = \text{diag}(q_1, \ldots, q_I)$.

There are several attractive proposals of robust optimization that approximates individual probabilistic constraint (see [6, 7, 8, 14, 23])). In such a proposal, $(y^0, \boldsymbol{y})$ satisfying the following robust counterpart

$$y^0 + \max_{\boldsymbol{z} \in \mathcal{U}} \boldsymbol{y}'\boldsymbol{z} \leq 0$$

guarantees that

$$\mathrm{P}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} \leq 0) \geq 1 - \epsilon. \tag{4.3}$$

Clearly, the choice of uncertainty set depends on the underlying assumption of primitive uncertainty.

Another approach of approximating the probabilistic constraint problem is to provide an upper bound of $CVaR_{1-\epsilon}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})$, so that if the bound is nonnegative, the probabilistic constraint (4.3) will be satisfied. The key difficulty lies in the evaluation of the expectation of a positive component of a random variable, $\mathrm{E}((\cdot)^+)$, which can be viewed as a multi-dimension integration. From Theorem 5 and 6, we know that for a given $\mathcal{L} \subseteq \{1, \ldots, 5\}$, $\pi^{\mathcal{L}}(y^0, \boldsymbol{y})$ upper bounds $\mathrm{E}((y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+)$. We define

$$\eta^{\mathcal{L}}_{1-\epsilon}(y^0, \boldsymbol{y}) \triangleq \min_{\beta} \left\{ \beta + \frac{1}{\epsilon} \pi^{\mathcal{L}}(y^0 - \beta, \boldsymbol{y}) \right\}.$$

Clearly,

$$CVaR_{1-\epsilon}(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) \leq \eta^{\mathcal{L}}_{1-\epsilon}(y^0, \boldsymbol{y})$$

and a sufficient condition for satisfying (4.3) is

$$\eta^{\mathcal{L}}_{1-\epsilon}(y^0, \boldsymbol{y}) \leq 0. \tag{4.4}$$

Since the epigraph of $\pi^{\mathcal{L}}(y^0 - \beta, \boldsymbol{y})$ is second order cone representable, the constraint (4.4) is also second order cone representable.

Before we show the connection between the robust optimization and approximation of CVaR, we need the following result.

Proposition 3. Let $\mathcal{U}_i$, $i \in \mathcal{L}$, be compact uncertainty sets such that their intersections

$$\mathcal{U}_{\mathcal{L}} = \bigcap_{i \in \mathcal{L}} \mathcal{U}_i,$$

has a non-empty interior. Then

$$\max_{\boldsymbol{z} \in \mathcal{U}_{\mathcal{L}}} \boldsymbol{y}'\boldsymbol{z} = \min_{\boldsymbol{y}_i, i \in \mathcal{L}} \left(\sum_{i \in \mathcal{L}} \max_{\boldsymbol{z}_i \in \mathcal{U}_i} \boldsymbol{y}_i'\boldsymbol{z}_i \mid \sum_{i \in \mathcal{L}} \boldsymbol{y}_i = \boldsymbol{y} \right).$$

Proof : We observe that the problem

$$\begin{aligned} \max \quad & \boldsymbol{y}'\boldsymbol{z} \\ \text{s.t.} \quad & \boldsymbol{z} \in \mathcal{U}_{\mathcal{L}} \end{aligned}$$

is equivalently

$$\begin{aligned} \max \quad & \boldsymbol{y}'\boldsymbol{z} \\ \text{s.t.} \quad & \boldsymbol{z}_i = \boldsymbol{z} \\ & \boldsymbol{z}_i \in \mathcal{U}_i \qquad \forall i \in \mathcal{L}. \end{aligned} \tag{4.5}$$

By strong duality, we have

$$
\begin{aligned}
& \max_{\boldsymbol{z}}\{\boldsymbol{y}'\boldsymbol{z} \ : \ \boldsymbol{z} = \boldsymbol{z}_i, i \in \mathcal{L}\} \\
= & \min_{\boldsymbol{y}_i, i\in\mathcal{L}} \left\{ \sum_{i\in\mathcal{L}} \boldsymbol{y}_i'\boldsymbol{z}_i \ : \ \sum_{i\in\mathcal{L}} \boldsymbol{y}_i = \boldsymbol{y} \right\}.
\end{aligned}
$$

Hence, the problem (4.5) is equivalent to

$$
\max_{\boldsymbol{z}\in\mathcal{U}_{\mathcal{L}}} \boldsymbol{y}'\boldsymbol{z} = \max_{\boldsymbol{z}_i\in\mathcal{U}_i, i\in\mathcal{L}} \left\{ \min_{\boldsymbol{y}_i, i\in\mathcal{L}} \left\{ \sum_{i\in\mathcal{L}} \boldsymbol{y}_i'\boldsymbol{z}_i \mid \sum_{i\in\mathcal{L}} \boldsymbol{y}_i = \boldsymbol{y} \right\} \right\}.
$$

Observe the set $\mathcal{U}_{\mathcal{L}}$ is a compact set with nonempty interior. Hence, $\max_{\boldsymbol{z}\in\mathcal{U}} \boldsymbol{y}'\boldsymbol{z}$ is therefore finite. Furthermore, there exists finite optimal primal and dual solutions $\boldsymbol{z}_i$ and $\boldsymbol{y}_i$, $i \in \mathcal{L}$ that satisfy strong duality. Hence, we can exchange "max" with "min", so that

$$
\begin{aligned}
\max_{\boldsymbol{z}\in\mathcal{U}} \boldsymbol{y}'\boldsymbol{z} &= \min_{\boldsymbol{y}_i, i\in\mathcal{L}} \left\{ \max_{\boldsymbol{z}_i\in\mathcal{U}_i, i\in\mathcal{L}} \sum_{i\in\mathcal{L}} \boldsymbol{y}_i'\boldsymbol{z}_i \mid \sum_{i\in\mathcal{L}} \boldsymbol{y}_i = \boldsymbol{y} \right\} \\
&= \min_{\boldsymbol{y}_i, i\in\mathcal{L}} \left\{ \sum_{i\in\mathcal{L}} \max_{\boldsymbol{z}_i\in\mathcal{U}_i} \boldsymbol{y}_i'\boldsymbol{z}_i \mid \sum_{i\in\mathcal{L}} \boldsymbol{y}_i = \boldsymbol{y} \right\}.
\end{aligned}
$$

■

Theorem 9. Suppose $\tilde{\boldsymbol{z}}$ follows the Model of Uncertainty, U. Let $\mathcal{L} \subseteq \{1, \ldots, 5\}$ and define

$$
\mathcal{U}_{\mathcal{L}}(\epsilon) \triangleq \bigcap_{l\in\mathcal{L}} \mathcal{U}_l(\epsilon)
$$

where

$$
\begin{aligned}
&\mathcal{U}_1(\epsilon) \triangleq \mathcal{W}\\
&\mathcal{U}_2(\epsilon) \triangleq \{\boldsymbol{z} \mid \boldsymbol{z} = (1-1/\epsilon)\boldsymbol{\zeta}, \text{for some } \boldsymbol{\zeta} \in \mathcal{W}\}\\
&\mathcal{U}_3(\epsilon) \triangleq \left\{\boldsymbol{z} \mid \|\boldsymbol{z}\|_2 \le \sqrt{\frac{1-\epsilon}{\epsilon}}\right\}\\
&\mathcal{U}_4(\epsilon) \triangleq \left\{\boldsymbol{z} \mid \exists \boldsymbol{s}, \boldsymbol{t} \in \Re^I, (z_1, \ldots, z_I) = \boldsymbol{s} - \boldsymbol{t}, \|\boldsymbol{P}^{-1}\boldsymbol{s} + \boldsymbol{Q}^{-1}\boldsymbol{t}\| \le \sqrt{-2\ln\epsilon}\right\}\\
&\mathcal{U}_5(\epsilon) \triangleq \left\{\boldsymbol{z} \mid \exists \boldsymbol{s}, \boldsymbol{t} \in \Re^I, (z_1, \ldots, z_I) = \boldsymbol{s} - \boldsymbol{t}, \|\boldsymbol{Q}^{-1}\boldsymbol{s} + \boldsymbol{P}^{-1}\boldsymbol{t}\| \le \frac{1-\epsilon}{\epsilon}\sqrt{-2\ln(1-\epsilon)}\right\}.
\end{aligned}
$$

Then

$$\eta^{\mathcal{L}}_{1-\epsilon}(y^0, \boldsymbol{y}) = y^0 + \max_{\boldsymbol{z} \in \mathcal{U}_{\mathcal{L}}(\epsilon)} \boldsymbol{y}'\boldsymbol{z}.$$

Proof :

For notational convenience, we ignore the representation of uncertainty sets as functions of ϵ. Observe that for any $\epsilon \in (0,1)$, the sets, $\mathcal{U}_1, \ldots, \mathcal{U}_1$ are compact and contain $\mathbf{0}$ in their interiors.

Uncertainty Set $\mathcal{U}_1$:

$$
\begin{aligned}
\eta^1_{1-\epsilon}(y^0, \boldsymbol{y}) &= \min_{\beta} \left(\beta + \frac{\pi^1(y^0 - \beta, \boldsymbol{y})}{\epsilon}\right)\\
&= \min_{\beta} \left(\beta + \frac{1}{\epsilon}(y^0 - \beta + \max_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{y}'\boldsymbol{z})^+\right)\\
&= y^0 + \max_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{y}'\boldsymbol{z} + \min_{\beta} \left(\beta + \frac{1}{\epsilon}(-\beta)^+\right)\\
&= y^0 + \max_{\boldsymbol{z} \in \mathcal{U}_1} \boldsymbol{y}'\boldsymbol{z}.
\end{aligned}
$$

Uncertainty Set $\mathcal{U}_2$:

$$\begin{aligned}
\eta^2_{1-\epsilon}(y^0,\boldsymbol{y}) &= \min_\beta \left(\beta + \frac{\pi^2(y^0-\beta,\boldsymbol{y})}{\epsilon}\right) \\
&= y^0 + \min_\beta \left(\beta + \frac{\pi^2(-\beta,\boldsymbol{y})}{\epsilon}\right) \\
&= y^0 + \min_\beta \left\{\beta + \frac{1}{\epsilon}\left(\left(\max_{\boldsymbol{z}\in\mathcal{W}}(-\boldsymbol{y})'\boldsymbol{z}+\beta\right)^+ - \beta\right)\right\} \\
&= y^0 + \min_\beta \left\{\beta(1-1/\epsilon) + \frac{1}{\epsilon}\left((\max_{\boldsymbol{z}\in\mathcal{W}}(-\boldsymbol{y})'\boldsymbol{z}+\beta)^+\right)\right\} \\
&= y^0 + (1/\epsilon-1)\min_\beta \left\{-\beta + \frac{1}{1-\epsilon}\left((\max_{\boldsymbol{z}\in\mathcal{W}}(-\boldsymbol{y})'\boldsymbol{z}+\beta)^+\right)\right\} \\
&= y^0 + (1/\epsilon-1)\max_{\boldsymbol{z}\in\mathcal{W}}\boldsymbol{y}'(-\boldsymbol{z}) + (1/\epsilon-1)\min_\beta\left(-\beta+\frac{1}{1-\epsilon}(\beta)^+\right) \\
&= y^0 + \max_{\boldsymbol{z}\in\mathcal{U}_2}\boldsymbol{y}'\boldsymbol{z}.
\end{aligned}$$

Uncertainty Set $\mathcal{U}_3$:

$$\begin{aligned}
\eta^3_{1-\epsilon}(y^0,\boldsymbol{y}) &= \min_\beta \left(\beta + \frac{\pi^3(y^0-\beta,\boldsymbol{y})}{\epsilon}\right) \\
&= \min_\beta \left(\beta + \frac{y^0-\beta+\sqrt{(y^0-\beta)^2+\boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}}}{2\epsilon}\right) \\
&= y^0 + \sqrt{\frac{1-\epsilon}{\epsilon}}\sqrt{\boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}} \\
&= y^0 + \max_{\boldsymbol{z}\in\mathcal{U}_3}\boldsymbol{y}'\boldsymbol{z},
\end{aligned}$$

where the second equality follows from choosing the optimum β,

$$\beta^* = y^0 + \frac{\sqrt{\boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}}(1-2\epsilon)}{2\sqrt{\epsilon(1-\epsilon)}}.$$

Uncertainty Set $\mathcal{U}_4$:

For notational convenience, we denote

$$\boldsymbol{y}_{\mathcal{I}} = (y_1, \ldots, y_I)$$

$$\boldsymbol{y}_{\bar{\mathcal{I}}} = (y_{I+1}, \ldots, y_N).$$

$$\begin{aligned}
\eta^4_{1-\epsilon}(y^0, \boldsymbol{y}) &= \min_{\beta} \left(\beta + \frac{\pi^4(y^0 - \beta, \boldsymbol{y})}{\epsilon} \right) \\
&= \min_{\beta,\mu,\boldsymbol{u}} \left(\beta + \frac{\frac{\mu}{e}\exp(\frac{y^0-\beta}{\mu} + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2})}{2\epsilon} \mid \boldsymbol{u} \geq \boldsymbol{P}\boldsymbol{y}_{\mathcal{I}}, \boldsymbol{u} \geq -\boldsymbol{Q}\boldsymbol{y}_{\mathcal{I}}, \boldsymbol{y}_{\bar{\mathcal{I}}} = \boldsymbol{0} \right) \\
&= \min_{\mu,\boldsymbol{u}} \left(y^0 + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2} - \mu \ln \epsilon \mid \boldsymbol{u} \geq \boldsymbol{P}\boldsymbol{y}_{\mathcal{I}}, \boldsymbol{u} \geq -\boldsymbol{Q}\boldsymbol{y}_{\mathcal{I}}, \boldsymbol{y}_{\bar{\mathcal{I}}} = \boldsymbol{0} \right) \\
&= \min_{\boldsymbol{u}} \left(y^0 + \sqrt{-2\ln\epsilon}\, u_0 \mid \boldsymbol{P}^{-1}\boldsymbol{u} \geq \boldsymbol{y}_{\mathcal{I}}, \boldsymbol{Q}^{-1}\boldsymbol{u} \geq -\boldsymbol{y}_{\mathcal{I}}, \boldsymbol{y}_{\bar{\mathcal{I}}} = \boldsymbol{0}, \|\boldsymbol{u}\|_2 \leq u_0 \right) \\
&= y^0 + \max_{\boldsymbol{z} \in \mathcal{U}_4} \boldsymbol{y}'\boldsymbol{z},
\end{aligned}$$

where the second and third equalities follow from choosing the tightest β^* and μ^*, that is

$$\beta^* = y^0 + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2} - \mu \ln \epsilon - \mu,$$

$$\mu^* = \frac{\|\boldsymbol{u}\|_2}{\sqrt{-2\ln\epsilon}}.$$

The last equality is the result of strong conic duality and has been derived in Chen, Sim and Sun [23].

Uncertainty Set $\mathcal{U}_5$:

Following from the above exposition,

$$
\begin{aligned}
&\eta^5_{1-\epsilon}(y^0,\boldsymbol{y})\\
=\;& \min_{\beta}\left(\beta+\frac{\pi^5(y^0-\beta,\boldsymbol{y})}{\epsilon}\right)\\
=\;& \min_{\beta,\mu,\boldsymbol{v}}\left(\beta+\frac{y^0-\beta+\frac{\mu}{e}\exp(-\frac{y^0-\beta}{\mu}+\frac{\|\boldsymbol{v}\|_2^2}{2\mu^2})}{2\epsilon}\mid \boldsymbol{v}\geq-\boldsymbol{P}\boldsymbol{y}_{\mathcal{I}},\boldsymbol{v}\geq\boldsymbol{Q}\boldsymbol{y}_{\mathcal{I}},\boldsymbol{y}_{\bar{\mathcal{I}}}=\boldsymbol{0}\right)\\
=\;& \min_{\mu,\boldsymbol{v}}\left(y^0+(\frac{1}{\epsilon}-1)(\frac{\|\boldsymbol{v}\|_2^2}{2\mu^2}-\mu\ln(1-\epsilon))\mid \boldsymbol{v}\geq-\boldsymbol{P}\boldsymbol{y}_{\mathcal{I}},\boldsymbol{v}\geq\boldsymbol{Q}\boldsymbol{y}_{\mathcal{I}},\boldsymbol{y}_{\bar{\mathcal{I}}}=\boldsymbol{0}\right)\\
=\;& \min_{\boldsymbol{v}}\left(y^0+\frac{1-\epsilon}{\epsilon}\sqrt{-2\ln(1-\epsilon)}\|\boldsymbol{v}\|\mid\;\mid \boldsymbol{P}^{-1}\boldsymbol{v}\geq-\boldsymbol{y}_{\mathcal{I}},\boldsymbol{Q}^{-1}\boldsymbol{v}\geq\boldsymbol{y}_{\mathcal{I}},\boldsymbol{y}_{\bar{\mathcal{I}}}=\boldsymbol{0}\right)\\
=\;& y^0+\max_{\boldsymbol{z}\in\mathcal{U}_5}\boldsymbol{y}'\boldsymbol{z}.
\end{aligned}
$$

Uncertainty Set $\mathcal{U}_{\mathcal{L}}$:

$$
\begin{aligned}
&\eta^{\mathcal{L}}(y^0,\boldsymbol{y})\\
=\;& \min_{\beta}\left(\beta+\frac{\pi^{\mathcal{L}}(y^0-\beta,\boldsymbol{y})}{\epsilon}\right)\\
=\;& \min_{\boldsymbol{\beta},y_{l0},\boldsymbol{y}_l,l\in\mathcal{L}}\left(\beta+\sum_{l\in\mathcal{L}}\left(\frac{\pi^l(y_{l0}-\beta_l,\boldsymbol{y}_l)}{\epsilon}\right)\mid\sum_{l\in\mathcal{L}}\boldsymbol{y}_l=\boldsymbol{y},\sum_{l\in\mathcal{L}}y_{l0}=y^0,\sum_{l\in\mathcal{L}}\beta_l=\beta\right)\\
=\;& \min_{y_{l0},\boldsymbol{y}_l,l\in\mathcal{L}}\left(\sum_{l\in\mathcal{L}}\min_{\beta_l}\left(\beta_l+\frac{\pi^l(y_{l0}-\beta_l,\boldsymbol{y}_l)}{\epsilon}\right)\mid\sum_{l\in\mathcal{L}}\boldsymbol{y}_l=\boldsymbol{y},\sum_{l\in\mathcal{L}}y_{l0}=y^0\right)\\
=\;& \min_{y_{l0},\boldsymbol{y}_l,l\in\mathcal{L}}\left(\sum_{l\in\mathcal{L}}\left(y_{l0}+\max_{\boldsymbol{z}\in\mathcal{U}_l}\boldsymbol{y}_l'\boldsymbol{z}\right)\mid\sum_{l\in\mathcal{L}}\boldsymbol{y}_l=\boldsymbol{y},\sum_{l\in\mathcal{L}}y_{l0}=y^0\right)\\
=\;& y^0+\min_{\boldsymbol{y}_l,l\in\mathcal{L}}\left(\sum_{l\in\mathcal{L}}\left(\max_{\boldsymbol{z}\in\mathcal{U}_l}\boldsymbol{y}_l'\boldsymbol{z}\right)\mid\sum_{l\in\mathcal{L}}\boldsymbol{y}_l=\boldsymbol{y}\right)\\
=\;& y^0+\max_{\boldsymbol{z}\in\mathcal{U}_{\mathcal{L}}}\boldsymbol{y}'\boldsymbol{z},
\end{aligned}
$$

where the last inequality is due to Proposition 3. ■

Hence, the different approximations of individual chance constrained problems using robust optimization are the consequences of applying different bounds on $\mathrm{E}((\cdot)^+)$. Notably, when the primitive uncertainties are characterized only by their means and covariance, the corresponding uncertainty set is an ellipsoid of the form $\mathcal{U}_3$. See, for instance, Bertsimas et al. [14] and El-Ghaoui et al. [28]. When $I = N$, that is all the primitive uncertainties are independently distributed, Chen, Sim and Sun [23] proposed the asymmetrical uncertainty set

$$\mathcal{U}_A(\epsilon) = \underbrace{\mathcal{W}}_{=\mathcal{U}_1(\epsilon)} \bigcap \mathcal{U}_4(\epsilon),$$

which generalizes the uncertainty set proposed by Ben-Tal and Nemirovski [8]. Noting that $\mathcal{U}_A(\epsilon) \subseteq \mathcal{U}_{\{1,2,4,5\}}(\epsilon)$, we can therefore improve upon the approximation using the uncertainty set $\mathcal{U}_{\{1,2,4,5\}}(\epsilon)$. However, in most application of chance constrained problems, the safety factor, ϵ is relatively small. In which case, the uncertainty sets of $\mathcal{U}_2(\epsilon)$ and $\mathcal{U}_5(\epsilon)$ are usually exploded to engulf the uncertainty sets of $\mathcal{W}$ and $\mathcal{U}_4(\epsilon)$, respectively . For instance, under symmetric distributions, that is $\boldsymbol{P} = \boldsymbol{Q}$ and $\bar{\boldsymbol{z}} = \underline{\boldsymbol{z}}$, it is easy to establish that for $\epsilon < 0.5$, we have

$$\mathcal{U}_{\{1,2,4,5\}}(\epsilon) = \underbrace{\mathcal{U}_1(\epsilon)}_{=\mathcal{W}} \bigcap \underbrace{\mathcal{U}_2(\epsilon)}_{\supseteq\mathcal{W}} \bigcap \; \mathcal{U}_4 \bigcap \underbrace{\mathcal{U}_5}_{\supseteq\mathcal{U}_4} = \mathcal{U}_A(\epsilon).$$

For $\boldsymbol{\mathcal{L}} = \{1, \dots, 5\}$, the constraint $\eta^{\mathcal{L}}_{1-\epsilon}(y^0, \boldsymbol{y}) \leq 0$ can be expressed as follows:

$$
\begin{array}{ll}
\exists \delta_i, y_{0i} \in \Re, \boldsymbol{y}_i, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{d}, \boldsymbol{h} \in \Re^N, i = 1, 2, \ldots, 5, \boldsymbol{u}, \boldsymbol{v} \in \Re^I \text{ such that} & \\
\delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5 \leq 0 & \\
y_{10} + \boldsymbol{s}'\bar{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}} \leq \delta_1 & \\
\boldsymbol{s}, \boldsymbol{t} \geq 0 & \\
\boldsymbol{s} - \boldsymbol{t} = \boldsymbol{y}_1 & \\
y_{20} + (1/\epsilon - 1)\boldsymbol{d}'\bar{\boldsymbol{z}} + (1/\epsilon - 1)\boldsymbol{h}'\underline{\boldsymbol{z}} \leq \delta_2 & \\
\boldsymbol{d} - \boldsymbol{h} = -\boldsymbol{y}_2 & \\
\boldsymbol{d}, \boldsymbol{h} \geq 0 & \\
y_{30} + \sqrt{\frac{1-\gamma}{\gamma}} \|\boldsymbol{\Sigma}^{1/2}\boldsymbol{y}_3\|_2 \leq \delta_3 & \\
y_{40} + \sqrt{-2\ln(\gamma)} \|\boldsymbol{u}\|_2 \leq \delta_4 & \\
u_j \geq p_j y_{4j}, u_j \geq -q_j y_{4j} & \forall j = 1, \ldots, I \\
y_{4j} = 0 & \forall j = I+1, \ldots, N \\
y_{50} + \frac{1-\gamma}{\gamma} \sqrt{-2\ln(1-\gamma)} \|\boldsymbol{v}\|_2 \leq \delta_5 & \\
v_j \geq q_j y_{5j}, v_j \geq -p_j y_{5j} & \forall j = 1, \ldots, I \\
y_{5j} = 0 & \forall j = I+1, \ldots, N \\
y_{10} + y_{20} + y_{30} + y_{40} + y_{40} = y^0 & \\
\boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 + \boldsymbol{y}_4 + \boldsymbol{y}_5 = \boldsymbol{y}. &
\end{array}
$$

It is interesting to note that while the epigraph of the function $\pi^{\mathcal{L}}(\cdot, \cdot)$ is approximately second-order cone representable, the epigraph of $\eta^{\mathcal{L}}(\cdot, \cdot)$, is fully second-order cone representable.

4.2 Joint probabilistic Constraint

We now extend the result in the previous section to the joint linear probabilistic constraint. Similarly, we consider a linear joint probabilistic constraint

$$\mathrm{P}\Big(y_i(\tilde{\boldsymbol{z}}) \leq 0, i \in \mathcal{M})\Big) \geq 1 - \epsilon, \tag{4.6}$$

where $\mathcal{M} = \{1, \ldots, m\}$. $y_i(\tilde{\boldsymbol{z}})$ are affinely dependent of $\tilde{\boldsymbol{z}}$,

$$y_i(\tilde{\boldsymbol{z}}) = y_i^0 + \sum_{j=1}^{N} y_i^j \tilde{z}_j \qquad i \in \mathcal{M},$$

where $(y_1^0, \ldots, y_1^N, \ldots, y_m^0, \ldots, y_m^N)$ are the decision variables. For notational convenience, we use

$$\boldsymbol{y}_i = (y_i^1, \ldots, y_i^N),$$

so $y_i(\tilde{\boldsymbol{z}}) = y_i^0 + \boldsymbol{y}_i' \tilde{\boldsymbol{z}}$. Moreover, we use

$$\boldsymbol{Y} = (y_1^0, \ldots, y_1^N, \ldots, y_m^0, \ldots, y_m^N),$$

to represent all the decision variables in the joint probabilistic constraint. It is straight forward to see that by suitable affine constraints imposing the decision variables $\boldsymbol{Y}$ and $\boldsymbol{x}$, we can express the probabilistic constraint

$$\mathrm{P}\Big(\boldsymbol{a}_i(\tilde{\boldsymbol{z}})'\boldsymbol{x} \leq b_i(\tilde{\boldsymbol{z}}), i \in \mathcal{M}\Big) \geq 1 - \epsilon, \tag{4.7}$$

as the form of constraint (4.6).

It is not surprising that a joint probabilistic constraint is more difficult to solve than an individual one. The standard approach proposed in the literatures [44, 23] approximates the problem using Bonferroni's inequality, so the joint constraint can be decomposed into m individual constraints in the form of

$$\mathrm{P}\Big(y_i(\tilde{\boldsymbol{z}}) \leq 0\Big) \geq 1 - \epsilon_i,\ i \in \mathcal{M}, \tag{4.8}$$

in which

$$\sum_{i=1}^{m} \epsilon_i \leq \epsilon. \tag{4.9}$$

Consequently, using the techniques discussed in the previous section, we can approximate the constraints (4.8) as follows

$$\eta^{\mathcal{L}}_{1-\epsilon_i}(y_i^0, \boldsymbol{y}_i) \leq 0,\ i \in \mathcal{M}. \tag{4.10}$$

The main issue with using Bonferroni's inequality is the choice of ϵ_i. Unfortunately, the problem becomes non-convex and possibly intractable if

ϵ_i are made variables and enforcing the constraint (4.9) as part of the optimization model. As such, it is natural to choose, $\epsilon_i = \epsilon/m$.

In some instances, Bonferroni's inequality may be rather conservative even for an optimal choice of ϵ_i. For instance, suppose $y_i(\tilde{\boldsymbol{z}})$ are completely correlated, such as with $a^0 \in \Re, \boldsymbol{a} \in \Re^N$,

$$y_i(\tilde{\boldsymbol{z}}) = \delta_i(a^0 + \boldsymbol{a}'\tilde{\boldsymbol{z}}), \qquad i = 1, \ldots, m \tag{4.11}$$

for some $\delta_i > 0$. Clearly, the least conservative choice of ϵ_i is $\epsilon_i = \epsilon$ for all $i \in \mathcal{M}$, which would violate the condition (4.9) imposed by Bonferroni's inequality. As a matter of fact, it is easy to see that the least conservative choice of ϵ_i while satisfying Bonferroni's inequality is $\epsilon_i = \epsilon/m$ for all $i = 1, \ldots, m$. Hence, if $y_i(\tilde{\boldsymbol{z}})$ are correlated, the efficiency of Bonferroni's inequality would possibly diminish.

We propose a new tractable way for approximating the joint probabilistic constraint. Given a set of positive constants, $\alpha_i \in (0, \infty]$, $i \in \mathcal{M}$, we define $\mathcal{J}$ as the index set of finite constants, that is

$$\mathcal{J} \triangleq \{i \ : \ \alpha_i < \infty\}$$

and its compliment index set,

$$\hat{\mathcal{J}} \triangleq \mathcal{M} \backslash \mathcal{J}.$$

Define

$$\psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \triangleq \min_{\beta, w_0, \boldsymbol{w}} \left(\beta + \frac{1}{\epsilon} \left\{ \pi^{\mathcal{L}}(w_0 - \beta, \boldsymbol{w}) + \sum_{i \in \mathcal{J}} \pi^{\mathcal{L}}(\alpha_i y_i^0 - w_0, \alpha_i \boldsymbol{y}_i - \boldsymbol{w}) \right\} \right).$$

The next result shows how we can use the function $\psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J})$ to approximate a joint probabilistic constraint.

Theorem 10. Under Assumption 1, the joint probabilistic constraint (4.6) is satisfied if

$$\psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leq 0 \tag{4.12}$$

and

$$y_i^0 + \max_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{y}_i' \boldsymbol{z} \leq 0 \qquad \forall i \in \hat{\mathcal{J}}. \tag{4.13}$$

Proof : Under Assumption 1, the set $\mathcal{W}$ is the support of the primitive uncertainty, $\tilde{\boldsymbol{z}}$, hence, the robust counterpart (4.13) implies

$$\mathrm{P}(y_i^0 + \boldsymbol{y}_i' \tilde{\boldsymbol{z}} > 0) = 0, \qquad \forall i \in \hat{\mathcal{J}}.$$

Hence, since $\boldsymbol{\alpha} > 0$, we have

$$\mathrm{P}\Big(y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}} \leq 0, i \in \mathcal{M}\Big) = \mathrm{P}\Big(y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}} \leq 0, i \in \mathcal{J}\Big) = \mathrm{P}\Big(\max_{i\in\mathcal{J}}\{\alpha_i y_i^0 + \alpha_i \boldsymbol{y}_i'\tilde{\boldsymbol{z}}\} \leq 0\Big).$$

Therefore, it suffices to show that if $\boldsymbol{Y}$ is feasible in the constraint (4.12), then the CVaR measure,

$$CVaR_{1-\epsilon}\left(\max_{i\in\mathcal{J}}\{\alpha_i y_i(\tilde{\boldsymbol{z}})\}\right) \leq 0.$$

We first claim that for any $y_1, ..., y_m$ and w,

$$w + \sum_i (y_i - w)^+ \geq \max_i\{y_i\}. \tag{4.14}$$

Indeed, for any index j,

$$w + \sum_i (y_i - w)^+ = w + (y_j - w)^+ + \sum_{i\neq j}(y_i - w)^+ \geq y_j + \sum_{i\neq j}(y_i - w)^+ \geq y_j.$$

Therefore,

$$
\begin{aligned}
& CVaR_{1-\epsilon}\Big(\max_{i\in\mathcal{J}}\{\alpha_i(y_i^0+\boldsymbol{y}_i'\tilde{\boldsymbol{z}})\}\Big) \\
= \;& \min_{\beta}\left\{\beta+\frac{1}{\epsilon}\mathrm{E}\left(\Big(\max_{i\in\mathcal{J}}\{\alpha_i(y_i^0+\boldsymbol{y}_i'\tilde{\boldsymbol{z}})\}-\beta\Big)^+\right)\right\} \\
\le \;& \min_{\beta,w_0,\boldsymbol{w}}\left\{\beta+\frac{1}{\epsilon}\mathrm{E}\left(\Big(w_0+\boldsymbol{w}'\tilde{\boldsymbol{z}}+\sum_{i\in\mathcal{J}}(\alpha_i y_i^0+\alpha_i\boldsymbol{y}_i'\tilde{\boldsymbol{z}}-(w_0+\boldsymbol{w}'\tilde{\boldsymbol{z}}))^+-\beta\Big)^+\right)\right\} \\
\le \;& \min_{\beta,w_0,\boldsymbol{w}}\left\{\beta+\frac{1}{\epsilon}\left(\mathrm{E}\left((w_0-\beta+\boldsymbol{w}'\tilde{\boldsymbol{z}})^+\right)+\sum_{i\in\mathcal{J}}\mathrm{E}\left((\alpha y_i^0-w_0+(\alpha_i\boldsymbol{y}_i-\boldsymbol{w})'\tilde{\boldsymbol{z}})^+\right)\right)\right\} \\
\le \;& \min_{\beta,w_0,\boldsymbol{w}}\left\{\beta+\frac{1}{\epsilon}\left(\pi^{\mathcal{L}}(w_0-\beta,\boldsymbol{w})+\sum_{i\in\mathcal{J}}\pi^{\mathcal{L}}(\alpha y_i^0-w_0,\alpha_i\boldsymbol{y}_i-\boldsymbol{w})\right)\right\} \\
= \;& \psi_{1-\epsilon}^{\mathcal{L}}(\boldsymbol{Y},\boldsymbol{\alpha},\mathcal{J})\le 0,
\end{aligned}
$$

where the first inequality is due to Inequality (4.14), the second inequality follows from

$$(a+b^+)^+\le a^+ +(b^+)^+ = a^+ + b^+$$

and the last inequality is the application of Theorem 6. ■

For a given $\boldsymbol{\alpha}$, the function $\psi_{1-\epsilon}^{\mathcal{L}}(\boldsymbol{Y},\boldsymbol{\alpha},\mathcal{J})$ is convex in $\boldsymbol{Y}$. Moreover, the corresponding epigraph is also second order cone representable. However, the function is not jointly convex in $\psi_{1-\epsilon}^{\mathcal{L}}(\boldsymbol{Y},\boldsymbol{\alpha},\mathcal{J})$. Nevertheless, for a given $\boldsymbol{Y}$, we note that the function, $\psi_{1-\epsilon}^{\mathcal{L}}(\boldsymbol{Y},\boldsymbol{\alpha},\mathcal{J})$ is convex with respect to $\boldsymbol{\alpha}$ and the corresponding epigraph is also second order cone representable. We will later exploit this property for improving the choice of $\boldsymbol{\alpha}$.

In the example (4.11) in which $y_i(\tilde{\boldsymbol{z}})$ is completely correlated, suppose we have

$$\eta^{\mathcal{L}}_{1-\epsilon_i}(a^0, \boldsymbol{a}) \leq 0$$

it is sufficient to guarantee feasibility in the joint probabilistic constraint problem. Choosing $\alpha_i = 1/\delta_i > 0$, we see that

$$\begin{aligned}
& \psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \\
= \; & \min_{\beta, w_0, \boldsymbol{w}} \left(\beta + \frac{1}{\epsilon} \left\{ \pi^{\mathcal{L}}(w_0 - \beta, \boldsymbol{w}) + \sum_{i \in \mathcal{J}} \pi^{\mathcal{L}}(\alpha_i y_i^0 - w_0, \alpha_i \boldsymbol{y}_i - \boldsymbol{w}) \right\} \right) \\
= \; & \min_{\beta, w_0, \boldsymbol{w}} \left(\beta + \frac{1}{\epsilon} \left\{ \pi^{\mathcal{L}}(w_0 - \beta, \boldsymbol{w}) + \sum_{i \in \mathcal{J}} \pi^{\mathcal{L}}(\alpha_i \delta_i a^0 - w_0, \alpha_i \delta_i \boldsymbol{a} - \boldsymbol{w}) \right\} \right) \\
\leq \; & \min_{\beta} \left(\beta + \frac{1}{\epsilon} \left\{ \pi^{\mathcal{L}}(a^0 - \beta, \boldsymbol{a}) + \sum_{i \in \mathcal{J}} \pi^{\mathcal{L}}(a^0 - a_0, \boldsymbol{a} - \boldsymbol{a}) \right\} \right) \\
= \; & \min_{\beta} \left(\beta + \frac{1}{\epsilon} \left\{ \pi^{\mathcal{L}}(a^0 - \beta, \boldsymbol{a}) \right\} \right) \\
= \; & \eta^{\mathcal{L}}_{1-\epsilon}(a^0, \boldsymbol{a}) \leq 0.
\end{aligned}$$

Therefore, we see that the new bound is potentially better than the application of Bonferroni's inequality on individual probabilistic constraints. We prove a stronger result as follows.

Theorem 11. Let $\epsilon_i \in (0,1), i \in \mathcal{M}$ and $\sum_{i=1}^m \epsilon_i \leq \epsilon$. Suppose $\boldsymbol{Y}$ satisfies

$$\eta^{\mathcal{L}}_{1-\epsilon_i}(y_i^0, \boldsymbol{y}_i) \leq 0 \qquad \forall i \in \mathcal{M},$$

then there exists $\alpha_i \in (0, \infty]$, $i = 1, \ldots, m$ such that $(\boldsymbol{Y}, \boldsymbol{\alpha})$ are feasible in the constraints (4.12) and (4.13).

Proof : Let β_i be the optimal solution to the model with constraints

$$\underbrace{\min_{\beta}\left(\beta + \frac{1}{\epsilon_i}\left(\pi^{\mathcal{L}}(y_i^0 - \beta, \boldsymbol{y}_i)\right)\right)}_{=\eta^{\mathcal{L}}_{1-\epsilon_i}(y_i^0, \boldsymbol{y}_i)} \leq 0.$$

Since

$$\pi^{\mathcal{L}}(y_i^0 - \beta_i, \boldsymbol{y}_i) \geq \mathrm{E}\left((y_i^0 - \beta_i + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^+\right) \geq 0,$$

we must have $\beta_i \leq 0$. Let $\mathcal{J} = \{i | \beta_i < 0\}$,

$$\alpha_j = -\frac{1}{\beta_j} \qquad \forall j \in \mathcal{J},$$

and correspondingly,

$$\alpha_j = \infty \qquad \forall j \in \underbrace{\{1, \ldots, m\} \backslash \mathcal{J}}_{=\hat{\mathcal{J}}}.$$

Since $\beta_j = 0$ for all $j \in \hat{\mathcal{J}}$, the following condition

$$0 \leq \mathrm{E}\left((y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^+\right) \leq \pi^{\mathcal{L}}(y_i^0, \boldsymbol{y}_i) \leq 0 \qquad \forall i \in \hat{\mathcal{J}}$$

implies that $\mathrm{E}\left((y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^+\right) = 0$ for all $i \in \hat{\mathcal{J}}$. Since $\mathcal{W}$ is the support of $\tilde{\boldsymbol{z}}$, this could only occur when

$$y_i^0 + \boldsymbol{y}_i'\boldsymbol{z} \leq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}, \qquad \forall i \in \hat{\mathcal{J}}$$

which satisfies the set of inequalities in (4.13).

For $i \in \mathcal{J}$, the constraint $\eta^{\mathcal{L}}_{1-\epsilon_i}(y_i^0, \boldsymbol{y}_i) \leq 0$ is equivalent to

$$\frac{1}{-\beta_i}\pi^{\mathcal{L}}(y_i^0 - \beta_i, \boldsymbol{y}_i) \leq \epsilon_i$$

Since the function $\pi^{\mathcal{L}}(\cdot,\cdot)$ is positive homogenous, we have

$$\begin{aligned} & \frac{1}{-\beta_i}\pi^{\mathcal{L}}(y_i^0 - \beta_i, \boldsymbol{y}_i) \\ = \;& \pi^{\mathcal{L}}\left(\tfrac{1}{-\beta_i} y_i^0 + 1, \tfrac{1}{-\beta_i}\boldsymbol{y}_i\right) \\ = \;& \pi^{\mathcal{L}}\left(\alpha_i y_i^0 + 1, \alpha_i \boldsymbol{y}_i\right) \leq \epsilon_i \qquad \forall i \in \mathcal{J}. \end{aligned}$$

Finally,

$$
\begin{aligned}
& \psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y},\boldsymbol{\alpha},\mathcal{J}) \\
= \;& \min_{\beta,w_0,\boldsymbol{w}} \left(\beta + \frac{1}{\epsilon}\left\{\pi^{\mathcal{L}}(w_0-\beta,\boldsymbol{w}) + \sum_{i\in\mathcal{J}} \pi^{\mathcal{L}}(\alpha_i y_i^0 - w_0, \alpha_i \boldsymbol{y}_i - \boldsymbol{w})\right\}\right) \\
\leq \;& -1 + \frac{1}{\epsilon}\left\{\pi^{\mathcal{L}}(-1+1,\boldsymbol{0}) + \sum_{i\in\mathcal{J}} \pi^{\mathcal{L}}(\alpha_i y^0 + 1, \alpha_i \boldsymbol{y} - \boldsymbol{0})\right\} \\
= \;& -1 + \frac{1}{\epsilon}\sum_{i\in\mathcal{J}} \pi^{\mathcal{L}}(\alpha_i y^0 + 1, \alpha_i \boldsymbol{y}) \\
\leq \;& -1 + \frac{1}{\epsilon}\sum_{i\in\mathcal{J}} \epsilon_i \leq 0,
\end{aligned}
$$

where the first inequality is due to the choice of $\beta = -1,\ w_0 = -1,\ \boldsymbol{w} = \boldsymbol{0}$ and the last inequality follows from $\sum_{i=1}^m \epsilon_i \leq \epsilon$. ■

4.3 Optimizing over α

In this section, we propose a method to choose coefficients $\boldsymbol{\alpha}$ such that the solutions of models with Constraint (4.12) and (4.13) can be improved. Consider an optimization model with a joint probabilistic constraint as follows

$$
\begin{aligned}
Z_\epsilon = \min \quad & \boldsymbol{c}'\boldsymbol{x} \\
\text{s.t.} \quad & \mathrm{P}(y_i(\tilde{\boldsymbol{z}}) \leq 0, \quad i \in \mathcal{M}) \geq 1-\epsilon \\
& (\boldsymbol{x},\boldsymbol{Y}) \in X,
\end{aligned}
\tag{4.15}
$$

in which X is efficiently computable convex set, such as a polyhedron or a second order cone representable set. Given a set of constant, $\boldsymbol{\alpha} > 0$ and a set $\mathcal{J}$, we consider the following optimization model.

$$\begin{aligned} Z_\epsilon^1(\boldsymbol{\alpha}, \mathcal{J}) = \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \psi_{1-\epsilon}^{\mathcal{L}}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leq 0 \\ & y_i^0 + \max_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{y}_i'\boldsymbol{z} \leq 0 \qquad \forall i \in \mathcal{M} \backslash \mathcal{J} \\ & (\boldsymbol{x}, \boldsymbol{Y}) \in X. \end{aligned} \tag{4.16}$$

Under Assumption 1, suppose Model (4.16) is feasible, the solution $\boldsymbol{x}, \boldsymbol{Y}$ is also feasible in Model (4.15), albeit more conservatively.

The main concern here is how to choose $\boldsymbol{\alpha}$ and $\mathcal{J}$. A likely choice, is say $\alpha_i = 1$, for all $i \in \mathcal{M}$ and $\mathcal{J} = \mathcal{M}$. Alternatively, we may use the classical approach by decomposing into m individual probabilisticconstraint problem with $\epsilon_i = \epsilon/m$. Base on Theorem 11, we can find a feasible $\boldsymbol{\alpha} > \mathbf{0}$ and set $\mathcal{J}$ such that Model (4.16) is also feasible.

Our aim is to improve upon the objective by minimizing $\psi_{1-\epsilon}^{\mathcal{L}}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J})$ over α_i, $i \in \mathcal{J}$, resulting in greater slack in the model (4.16). Hence, this approach will lead to improvement in the objective, or at least will not increase the value. We consider the following optimization problem over α_i, $i \in \mathcal{J}$,

$$
\begin{aligned}
Z^1_\alpha(\boldsymbol{Y}, \mathcal{J}) = \min \quad & \psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \\
\text{s.t.} \quad & \textstyle\sum_{i \in \mathcal{J}} \alpha_i = 1 \\
& \alpha_i \geq 0 \qquad \forall i \in \mathcal{J}.
\end{aligned} \tag{4.17}
$$

Since the feasible region of Model (4.17) is compact, the optimal solution for α_i, $i \in \mathcal{J}$ is therefore achievable. Suppose we obtain an initial feasible $(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J})$ satisfying $\psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leq 0$, due to the positive homogenous property, we can scale $\boldsymbol{\alpha}$ with any positive constraint without affecting its feasibility. Therefore, we can infer that $Z^1_\alpha(\boldsymbol{Y}, \mathcal{J}) \leq 0$.

However, it is possible that the optimum solution of Model (4.17) contains some element $\alpha^*_k = 0$ for some index $k \in \mathcal{J}$. This will require an update of the set $\mathcal{J}$ and reevaluation of Model (4.17). The following suggests how we should perform the updates.

Proposition 4. Assume there exists $(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J})$, $\boldsymbol{\alpha} > 0$, such that $\psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leq 0$.

(a) Let $\boldsymbol{\alpha}^*$ be the optimum solution to Model (4.17) and suppose there exists a nonempty set $\mathcal{K} \subset \mathcal{J}$ such that $\alpha_i = 0, \forall i \in \mathcal{K}$. Then

$$
y^0_i + \max_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{y}'_i \boldsymbol{z} \leq 0 \quad \forall i \in \mathcal{J} \backslash \mathcal{K}.
$$

(b) Moreover,

$$Z^1_\alpha(\boldsymbol{Y}, \mathcal{K}) \le 0.$$

Proof : (a) We have argued that $Z^1_\alpha(\boldsymbol{Y}, \mathcal{J}) \le 0$. Let $k \in \mathcal{K}$, that is, $\alpha^*_k = 0$. Observe that

$$\begin{aligned}
0 \ge\ & \psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}^*, \mathcal{J}) \\
=\ & \beta + \frac{1}{\epsilon}\left\{\pi^{\mathcal{L}}(w_0 - \beta, \boldsymbol{w}) + \sum_{i\in\mathcal{J}} \pi^{\mathcal{L}}(\alpha^*_i y^0_i - w_0, \alpha^*_i \boldsymbol{y}_i - \boldsymbol{w})\right\} \\
=\ & \beta + \frac{1}{\epsilon}\left\{\pi^{\mathcal{L}}(w_0 - \beta, \boldsymbol{w}) + \pi^{\mathcal{L}}(-w_0, -\boldsymbol{w})\right\} + \frac{1}{\epsilon}\sum_{i\in\mathcal{J}\backslash\{k\}} \pi^{\mathcal{L}}(\alpha^*_i y^0_i - w_0, \alpha^*_i \boldsymbol{y}_i - \boldsymbol{w}) \\
\ge\ & \beta + \frac{1}{\epsilon}\left\{\pi^{\mathcal{L}}(w_0 - \beta, \boldsymbol{w}) + \pi^{\mathcal{L}}(-w_0, -\boldsymbol{w})\right\} \\
\ge\ & \beta + \frac{1}{\epsilon}\pi^{\mathcal{L}}(-\beta, \boldsymbol{0}) \\
\ge\ & \beta + \frac{1}{\epsilon}\mathrm{E}\left((-\beta)^+\right) \\
=\ & \beta + \tfrac{1}{\epsilon}(-\beta)^+,
\end{aligned}$$

where the second equality is due to $\alpha^*_k = 0$ and the second inequality is due to convexity of the function, $\pi^{\mathcal{L}}(\cdot,\cdot)$. Since, $\epsilon \in (0, 1)$, the equality $\beta + \frac{1}{\epsilon}(-\beta)^+ = 0$ is satisfied if and only if $\beta = 0$ and the inequality $\pi^{\mathcal{L}}(w_0, \boldsymbol{w}) + \pi^{\mathcal{L}}(-w_0, -\boldsymbol{w}) = 0$ is satisfied if and only if $w_0 = 0,\ \boldsymbol{w} = 0$. Hence, we now conclude that

$$\pi^{\mathcal{L}}(y^0_i, \boldsymbol{y}_i) = 0 \quad \forall i \in \mathcal{J}\backslash\mathcal{K} \tag{4.18}$$

which implies

$$0 \le \mathrm{E}((y^0_i + \boldsymbol{y}'_i\tilde{\boldsymbol{z}})^+) \le \pi^{\mathcal{L}}(y^0_i, \boldsymbol{y}_i) = 0, \quad \forall i \in \mathcal{J}\backslash\mathcal{K}.$$

Since $\mathcal{W}$ is the support of $\tilde{\boldsymbol{z}}$, this could only occur when

$$y_i^0 + \max_{\boldsymbol{z}\in\mathcal{W}} \boldsymbol{y}_i'\boldsymbol{z} \leq 0 \quad \forall i \in \mathcal{J}\backslash\mathcal{K}.$$

(b) Under the assumption that there exists $(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J})$, $\boldsymbol{\alpha} > 0$, such that $\psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leq 0$. Since $\mathcal{K} \subset \mathcal{J}$ and using the same $\boldsymbol{\alpha}$, we observe that

$$\psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{K}) \leq \psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leq 0.$$

Again, due to the positive homogenous property of Theorem 10(b), we scale $\boldsymbol{\alpha}$ by a positive constant so that it is feasible in Problem (4.17). Hence, the result follows. ■

We propose an algorithm for improving the choice of $\boldsymbol{\alpha}$ and the set $\mathcal{J}$. Again, we assume we can find an initial feasible solution of Model (4.16).

Algorithm 2.

Input: $(\boldsymbol{Y}, \mathcal{J})$

1. Solve Problem (4.17) with Input $(\boldsymbol{Y}, \mathcal{J})$. Obtain optimal solution $\boldsymbol{\alpha}^*$
2. Set $\mathcal{K} := \{i | \alpha_i^* = 0, i \in \mathcal{J}\}$ and $\boldsymbol{\alpha} := \boldsymbol{\alpha}^*$.
3. If $\mathcal{K} \neq \emptyset$ Then Set $\mathcal{J} := \mathcal{K}$. Goto Step 1.

4. Else Solve Model (4.16) with Input $(\boldsymbol{\alpha}, \mathcal{J})$. Obtain optimal solution $(\boldsymbol{x}^*, \boldsymbol{Y}^*)$. Set $\boldsymbol{Y} = \boldsymbol{Y}^*$.

5. Repeat Step 1 until termination criteria is met.

6. Output solution $(\boldsymbol{x}^*, \boldsymbol{Y}^*)$.

Theorem 12. In Algorithm 2, the sequence of objectives obtained by solving Model (4.16) is non-increasing.

Proof : Starting with a feasible solution of Model (4.16), we are assured that there exists $(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J})$, $\boldsymbol{\alpha} > 0$, such that $\psi^{\mathcal{L}}_{1-\epsilon}(\boldsymbol{Y}, \boldsymbol{\alpha}, \mathcal{J}) \leq 0$. The condition in Step 3 ensures that $\alpha^*_i > 0$ for all $i \in \mathcal{J}$. Moreover, Proposition 4(a,b) ensure that the updates on $\boldsymbol{\alpha}$ and $\mathcal{J}$ do not affect the feasibility of the Model (4.16). ■

4.4 Example: Emergency Resource Allocation

We use an emergency resource allocation problem to test our algorithm solving joint probabilistic constrained problem. It is a two stage problem. The resources are allocated to multi facilities with different locations before the emergent event occurs. In the second stage, that is, after the emergent event occurs, the resources are reallocated through transshipment. The difference from the classical transshipment problem is that the resources can only be transshipped between two locations whose distance is less than a tolerance. Moreover, the transshipment cost can be ignored, so the objective of an emergency resource allocation problem is only the first stage cost. The constraint is to achieve a high confidence level that there is no deficiency when emergent events occurs. We use a directed network with m nodes and n arcs to denote the transshipment network. $\mathcal{E}$ represents the arc set. If arc $(i, j) \in \mathcal{E}$, then the resources can be transshipped from node i to j. Moreover, we define

c_i : Unit purchasing cost;

$\tilde{d}_i$: Demand;

x_i : Storage quantity (First stage decision variable);

w_{ij} : Transshipment quantity (Recourse Decision variable).

The problem can be formulated as a joint probabilistic constrained problem as follows.

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \mathrm{P}\left(\begin{array}{ll} x_i + \sum_{j:(j,i)\in\mathcal{E}} w_{ji}(\tilde{\boldsymbol{z}}) - \sum_{j:(i,j)\in\mathcal{E}} w_{ij}(\tilde{\boldsymbol{z}}) \geq d_i(\tilde{\boldsymbol{z}}) & i = 1,\dots,m \\ x_i \geq \sum_j w_{ij}(\tilde{\boldsymbol{z}}) & i = 1,\dots,m \\ \boldsymbol{w}(\tilde{\boldsymbol{z}}) \geq \boldsymbol{0} & \end{array} \right) \geq 1-\epsilon \\ & \boldsymbol{x} \geq \boldsymbol{0}. \end{array} \tag{4.19}$$

We assume that the demand $\boldsymbol{d}$ are affinely dependent on the uncertainties as follows.

$$\boldsymbol{d}(\tilde{\boldsymbol{z}}) = \boldsymbol{d}^0 + \sum_{j=1}^{N} \boldsymbol{d}^j \tilde{z}_j.$$

In addition, we restrict recourse variables $\boldsymbol{w}(\tilde{\boldsymbol{z}})$ to follow linear decision rule, that is

$$\boldsymbol{w}(\tilde{\boldsymbol{z}}) = \boldsymbol{w}^0 + \sum_{j=1}^{N} \boldsymbol{w}^j \tilde{z}_j.$$

With introduced recourse variables $\boldsymbol{r}(\tilde{\boldsymbol{z}}), \boldsymbol{s}(\tilde{\boldsymbol{z}}), \boldsymbol{t}(\tilde{\boldsymbol{z}}), \boldsymbol{y}(\tilde{\boldsymbol{z}})$, we can transform the model (4.19) to the standard form.

$$\begin{array}{lll}
\min & \boldsymbol{c}'\boldsymbol{x} & \\
\text{s.t.} & x_i + \sum_{j:(j,i)\in\mathcal{E}} w_{ji}(\tilde{\boldsymbol{z}}) - \sum_{j:(i,j)\in\mathcal{E}} w_{ij}(\tilde{\boldsymbol{z}}) + \boldsymbol{r}(\tilde{\boldsymbol{z}}) = d_i(\tilde{\boldsymbol{z}}) & i = 1,\ldots,m \\
& x_i + \boldsymbol{s}(\tilde{\boldsymbol{z}}) = \sum_j w_{ij}(\tilde{\boldsymbol{z}}) & i = 1,\ldots,m \\
& \boldsymbol{w}(\tilde{\boldsymbol{z}}) + \boldsymbol{t}(\tilde{\boldsymbol{z}}) = \boldsymbol{0} & \\
& \boldsymbol{y}(\tilde{\boldsymbol{z}}) = \begin{pmatrix} \boldsymbol{r}(\tilde{\boldsymbol{z}}) \\ \boldsymbol{s}(\tilde{\boldsymbol{z}}) \\ \boldsymbol{t}(\tilde{\boldsymbol{z}}) \end{pmatrix} & \\
& \mathrm{P}(\boldsymbol{y}(\tilde{\boldsymbol{z}}) \le \boldsymbol{0}) \ge 1 - \epsilon & \\
& \boldsymbol{x} \ge \boldsymbol{0}. &
\end{array} \tag{4.20}$$

It is easy to see that the recourse variables $\boldsymbol{r}(\tilde{\boldsymbol{z}}), \boldsymbol{s}(\tilde{\boldsymbol{z}}), \boldsymbol{t}(\tilde{\boldsymbol{z}}), \boldsymbol{y}(\tilde{\boldsymbol{z}})$ also

Nodes	Arcs	Z^W	Z^B	Z^N	$(Z^W - Z^N)/Z^W$	$(Z^B - Z^N)/Z^B$
15	50	1500	1158.1	1043.3	30.45%	9.91%
15	60	1500	1059.7	968.1	35.46%	8.64%
15	70	1500	1027.3	929.5	38.03%	9.52%
15	80	1500	1009.3	890.1	40.66%	11.81%
15	90	1500	989.1	865.7	42.29%	12.48%

Tab. 4.1: Comparisons among Worst case solution Z^W, Solution using Bonferroni's inequality Z^B and Solution using new approximation Z^N.

follow linear decision rule.

$$\boldsymbol{r}(\tilde{\boldsymbol{z}}) = \boldsymbol{r}^0 + \textstyle\sum_{j=1}^{N} \boldsymbol{r}^j \tilde{z}_j$$

$$\boldsymbol{s}(\tilde{\boldsymbol{z}}) = \boldsymbol{s}^0 + \textstyle\sum_{j=1}^{N} \boldsymbol{s}^j \tilde{z}_j$$

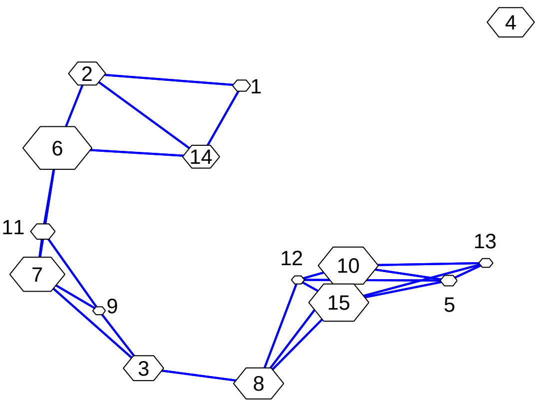
$$\boldsymbol{t}(\tilde{\boldsymbol{z}}) = \boldsymbol{t}^0 + \textstyle\sum_{j=1}^{N} \boldsymbol{t}^j \tilde{z}_j$$

$$\boldsymbol{y}(\tilde{\boldsymbol{z}}) = \boldsymbol{y}^0 + \textstyle\sum_{j=1}^{N} \boldsymbol{y}^j \tilde{z}_j.$$

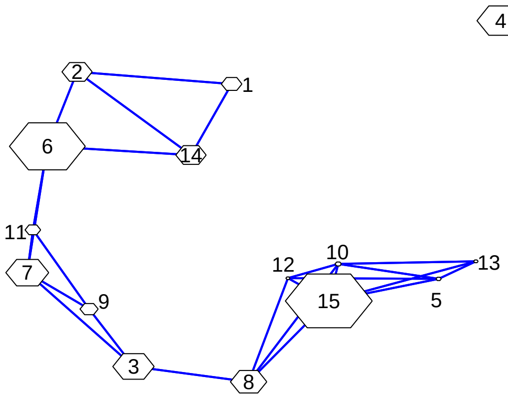
Therefore, we can apply Algorithm 2 to solve the model (4.20). We randomly generate m facilities and assume that the purchasing cost $c_i = 1$, the demand for each facility follows two point distribution

$$\begin{cases} \mathrm{P}(\tilde{d}_i = 0) = 0.9 \\ \mathrm{P}(\tilde{d}_i = 100) = 0.1 \end{cases} \quad \forall i.$$

Figure 4.1 shows the solutions for 15 facilities. The area of the hexagon on each location denotes the optimal storage quantity. We compare the solution of the new method Z^N with the solution using Bonferroni's inequality



Solution using Bonferroni's inequality



Solution using New Method

Fig. 4.1: Inventory allocation: 15 nodes, 50 arcs

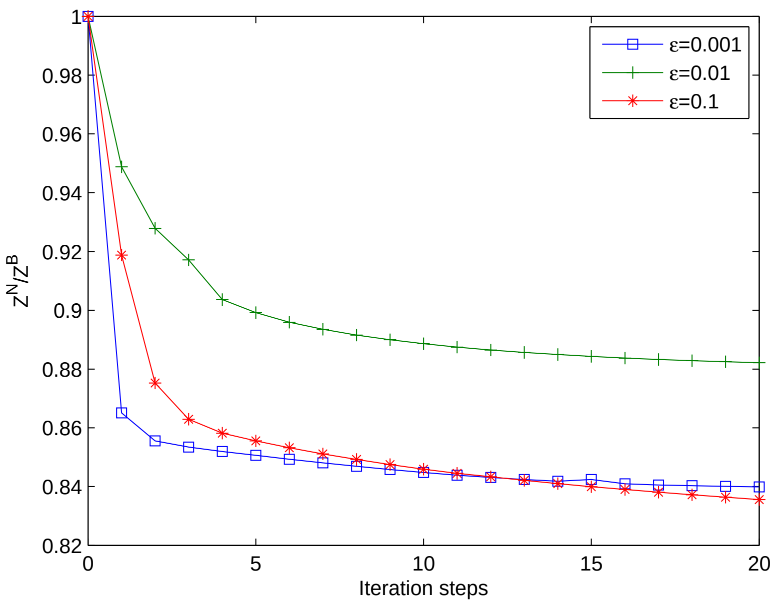


Fig. 4.2: Convergence of the heuristic: 15 nodes, 50 arcs

Z^B and the worst case solution Z^W. Table 4.1 shows the comparison results. The new method has $8 - 12\%$ improvement compared with Bonferroni's inequality and $30 - 42\%$ improvement compared with the worst case method. This experiment shows that the new method solves the joint probabilistic constrained problem efficiently. Moreover, we tested the convergence rate of Algrithm 2. Figure 4.2 shows that the improvement is mostly in the first several steps.

4.5 Goal Driven Model with Probabilistic Constraint

The aim of this section is to provide a tractable approximation for goal driven model with probabilistic constraint. We consider the following program.

$$\begin{aligned}
\max \quad & SALC\Big(\boldsymbol{c}(\tilde{\boldsymbol{z}})'\boldsymbol{x} + \boldsymbol{d}_u'\boldsymbol{u}(\tilde{\boldsymbol{z}}) + \boldsymbol{d}_y'\boldsymbol{y}(\tilde{\boldsymbol{z}}) + \boldsymbol{d}_r'\boldsymbol{r}(\tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})\Big) \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{B}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}(\tilde{\boldsymbol{z}}) + \boldsymbol{Y}\boldsymbol{y}(\tilde{\boldsymbol{z}}) + \boldsymbol{R}\boldsymbol{r}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}(\tilde{\boldsymbol{z}}) \\
& \mathrm{P}\Big(\boldsymbol{r}(\tilde{\boldsymbol{z}}) \leq \boldsymbol{0}\Big) \geq 1 - \epsilon \\
& \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{y}(\tilde{\boldsymbol{z}}) \geq 0,
\end{aligned} \tag{4.21}$$

where $\boldsymbol{c} \in \Re^{n_1}, \boldsymbol{b} \in \Re^{m_1}, \boldsymbol{d_u} \in \Re^{n_2}, \boldsymbol{d_y} \in \Re^{n_3}, \boldsymbol{d_r} \in \Re^{n_4}, \boldsymbol{A} \in \Re^{m_1\times n_1}, \boldsymbol{U} \in \Re^{m_2\times n_2}, \boldsymbol{Y} \in \Re^{m_2\times n_3}, \boldsymbol{R} \in \Re^{m_2\times n_4}$ are known parameters, $\boldsymbol{h}(\tilde{\boldsymbol{z}}) \in \Re^{m_2}, \boldsymbol{B}(\tilde{\boldsymbol{z}}) \in \Re^{m_2\times n_1}$ are random parameters as function mapping of the primitive uncertainties $\tilde{\boldsymbol{z}}$, $\tau(\tilde{\boldsymbol{z}})$ is the target level also depending on the primitive uncertainties $\tilde{\boldsymbol{z}}$, $\boldsymbol{x} \in \Re^{n_1}$ is the first stage decision variables, and $\boldsymbol{u}(\cdot) \in \Re^{n_2}, \boldsymbol{y}(\cdot) \in \Re^{n_3}, \boldsymbol{r}(\cdot) \in \Re^{n_4}$ are the second stage decision variables, also as function mapping of the realization of the primitive uncertainties $\tilde{\boldsymbol{z}}$. Note that the optimal solution of the goal driven model (4.21) can be obtained by solving a sequence of subproblems as follows.

$$\begin{aligned}
\min \quad & CVaR_{1-\gamma}\Big(\boldsymbol{c}(\tilde{\boldsymbol{z}})'\boldsymbol{x} + \boldsymbol{d}_u'\boldsymbol{u}(\tilde{\boldsymbol{z}}) + \boldsymbol{d}_y'\boldsymbol{y}(\tilde{\boldsymbol{z}}) + \boldsymbol{d}_r'\boldsymbol{r}(\tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})\Big) \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{B}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}(\tilde{\boldsymbol{z}}) + \boldsymbol{Y}\boldsymbol{y}(\tilde{\boldsymbol{z}}) + \boldsymbol{R}\boldsymbol{r}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}(\tilde{\boldsymbol{z}}) \\
& \mathrm{P}\Big(\boldsymbol{r}(\tilde{\boldsymbol{z}}) \le \boldsymbol{0}\Big) \ge 1 - \epsilon \\
& \boldsymbol{x} \ge \boldsymbol{0}, \boldsymbol{y}(\tilde{\boldsymbol{z}}) \ge 0,
\end{aligned} \tag{4.22}$$

We assume *Affine Parametric Perturbation* and *Model of Primitive Uncertainty*, U, as follows:

$$\boldsymbol{h}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}^0 + \sum_{j=1}^{N} \boldsymbol{h}^j \tilde{z}_j,$$

$$\boldsymbol{B}(\tilde{\boldsymbol{z}}) = \boldsymbol{B}^0 + \sum_{j=1}^{N} \boldsymbol{B}^j \tilde{z}_j,$$

$$\tau(\tilde{\boldsymbol{z}}) = \tau^0 + \sum_{j=1}^{N} \tau^j \tilde{z}_j.$$

Note that the number of the second stage vector $\boldsymbol{u}(\tilde{\boldsymbol{z}}), \boldsymbol{y}(\tilde{\boldsymbol{z}}), \boldsymbol{r}(\tilde{\boldsymbol{z}})$ can be very large or even infinite depending on the distribution of $\tilde{\boldsymbol{z}}$ Then the model (4.22) is generally intractable. As an approximation, we use the linear decision rule used in Ben-Tal et al. [11] and Chen, Sim, Sun [23], which limits the space of recourse solutions as follows,

$$\boldsymbol{u}(\tilde{\boldsymbol{z}}) = \boldsymbol{u}^0 + \sum_{j=1}^{N} \boldsymbol{u}^j \tilde{z}_j,$$

$$\boldsymbol{y}(\tilde{\boldsymbol{z}}) = \boldsymbol{y}^0 + \sum_{j=1}^{N} \boldsymbol{y}^j \tilde{z}_j,$$

$$\boldsymbol{r}(\tilde{\boldsymbol{z}}) = \boldsymbol{r}^0 + \sum_{j=1}^{N} \boldsymbol{r}^j \tilde{z}_j.$$

We define the vector $\bar{\boldsymbol{d}}$ with elements

$$\begin{aligned} \bar{d}_i = \min \quad & \boldsymbol{d_u}'\boldsymbol{u} + \boldsymbol{d_y}'\boldsymbol{y} + \boldsymbol{d_r}'\boldsymbol{r} \\ \text{s.t.} \quad & \boldsymbol{U}\boldsymbol{u} + \boldsymbol{Y}\boldsymbol{y} + \boldsymbol{R}\boldsymbol{r} = \boldsymbol{0} \\ & w_i = 1 \\ & \boldsymbol{y} \geq \boldsymbol{0}, \boldsymbol{u}, \boldsymbol{r} \text{ free}, \end{aligned}$$

where we denote $\bar{d}_i = \infty$ if the corresponding optimization problem is infeasible. For notational convenience, we define the sets

$$\mathcal{C} \triangleq \{i \ : \ \bar{d}_i < \infty, i = 1, \ldots, n_3\}, \qquad \bar{\mathcal{C}} \triangleq \{i = 1, \ldots, n_3\} \backslash \mathcal{C}.$$

For $i \in \mathcal{C}$, we define $(\bar{\boldsymbol{u}}^i, \bar{\boldsymbol{y}}^i, \bar{\boldsymbol{r}}^i)$ as the optimal solution of the corresponding optimization problem.

Note that if $\bar{d}_i < 0$, then given any feasible solution $\boldsymbol{u}$, $\boldsymbol{y}$ and $\boldsymbol{r}$, the solution $\boldsymbol{u}+\kappa\bar{\boldsymbol{u}}^i$, , $\boldsymbol{y}+\kappa\bar{\boldsymbol{y}}^i$ and $\boldsymbol{r}+\kappa\bar{\boldsymbol{r}}^i$ will also be feasible, and that the objective will be reduced by $|\kappa\bar{d}_i|$. Hence, whenever a second stage decision is feasible, its objective will be unbounded from below. Therefore, it is reasonable to assume that $\bar{\boldsymbol{d}} \geq \boldsymbol{0}$. Therefore, let $\{1\} \subset \mathcal{L} \subset \{1, 2, \ldots, 5\}$. Then under the deflected linear decision rule, we can approximate the problem (4.22) as

$$
\begin{array}{lll}
\min & \beta + \dfrac{1}{\gamma}\pi^{\mathcal{L}}(\xi^0 - \beta, \boldsymbol{\xi}) + \dfrac{1}{\gamma}\displaystyle\sum_{i \in \mathcal{C}} \pi^{\mathcal{L}}(-y_i^0, -\boldsymbol{y}_i)\bar{d}_i & \\
\text{s.t.} & \xi^j = \boldsymbol{c}^{j\prime}\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}^j + \boldsymbol{d_y}'\boldsymbol{y}^j + \boldsymbol{d_r}'\boldsymbol{r}^j - \tau^j & j = 0, \ldots, N. \\
& \boldsymbol{B}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}^j + \boldsymbol{Y}\boldsymbol{y}^j + \boldsymbol{R}\boldsymbol{r}^j = \boldsymbol{h}^j & j = 0, \ldots, N \\
& y_i^0 + \sum_{j=1}^N \boldsymbol{y}_i^j z_j \geq \boldsymbol{0} & \forall \boldsymbol{z} \in \mathcal{W}, i \in \bar{\mathcal{C}} \\
& \psi_{1-\epsilon}^{\mathcal{L}}(\boldsymbol{r}^0, \ldots, \boldsymbol{r}^N, \boldsymbol{\alpha}, \mathcal{J}) \leq 0 & \\
& r_i^0 + \max_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{r}_i'\boldsymbol{z} \leq 0 \qquad \forall & i \in \hat{\mathcal{J}}.
\end{array}
\tag{4.23}
$$

Remark : Algorithm 2 can be applied to solve the model (4.23).

5. APPLICATIONS

5.1 Project Management

We apply the goal driven optimization model to a project management problem with uncertain activity completion time. Project management is a well known problem which can be described with a directed graph having m arcs and n nodes. The arc set is denoted as $\mathcal{E}$, $|\mathcal{E}| = m$. Each arc (i, j) represents an activity which has uncertain completion time $\tilde{t}_{ij}$. It is affinely dependent on the additional amount of resource $x_{ij} \in [0, \bar{x}_{ij}]$ and a primitive uncertainty $\tilde{z}_{ij}$, as follows:

$$\tilde{t}_{ij} = (1 + \tilde{z}_{ij})b_{ij} - a_{ij}x_{ij}$$

where $\tilde{z}_{ij} \in [-\underline{z}_{ij}, \overline{z}_{ij}], \underline{z}_{ij} \leq 1,\ (i, j) \in \mathcal{E}$ is an independent random variable with zero mean, standard deviation σ_{ij}, forward and backward deviations, p_{ij} and q_{ij} respectively. The completion time adheres to precedent constraints. For instance, activity e_1 precedes activity e_2 if activity e_1 must be completed before activity e_2. Each node on the graph represents an event marking the completion of a particular subset of activities. For simplicity, we use node 1 as the start event and node n as the end event. The cost of using each unit of resource on activity (i, j) is c_{ij} and the total cost is limited to a budget B. Our goal is to find a resource allocation to each activity $(i, j) \in \mathcal{E}$ that maximize the shortfall aspiration level criterion in achieving a fixed targeted completion time, τ. We formulate the goal driven optimization model as follows.

$$
\begin{aligned}
\max \quad & SALC(u_n(\tilde{\boldsymbol{z}}) - \tau) \\
& u_j(\tilde{\boldsymbol{z}}) - u_i(\tilde{\boldsymbol{z}}) - w_{ij}(\tilde{\boldsymbol{z}}) = (1 + \tilde{z}_{ij})b_{ij} - a_{ij}x_{ij} \quad \forall (i,j) \in \mathcal{E} \\
& u_1(\tilde{\boldsymbol{z}}) = 0 \\
& \boldsymbol{c}'\boldsymbol{x} \leq B \\
& \boldsymbol{0} \leq \boldsymbol{x} \leq \overline{\boldsymbol{x}}, \boldsymbol{w}(\tilde{\boldsymbol{z}}) \geq \boldsymbol{0} \\
& \boldsymbol{x} \in \Re^m, \ \boldsymbol{u}(\cdot), \boldsymbol{w}(\cdot) \in \mathcal{Y},
\end{aligned}
\tag{5.1}
$$

where $u_i(\tilde{\boldsymbol{z}})$ is the second stage decision vector, representing the completion time at node i when the uncertain parameters $\tilde{\boldsymbol{z}}$ are realized. The recourse $w_{ij}(\tilde{\boldsymbol{z}})$ represents the slack at the arc (i, j). Using Algorithm 1, we reduce the problem (5.1) to solving a sequence of subproblems in the form of stochastic optimization problems with CVaR objectives. Since the project management problem has complete recourse, accordingly, we use sampling approximation to obtain solutions to the subproblem as follows.

$$
\begin{aligned}
\tilde{Z}^s_K(\gamma) = \quad \min \quad & \omega + \tfrac{1}{\gamma K}\textstyle\sum_{k=1}^K t_k \\
\text{s.t.} \quad & t_k \geq u^k_n - \tau - \omega && \forall k = 1,\ldots,K \\
& u^k_j - u^k_i \geq (1+\tilde{z}^k_{ij})b_{ij} - a_{ij}x_{ij} && \forall (i,j) \in \mathcal{E}, k = 1,\ldots,K \\
& u^k_1 = 0 && \forall k = 1,\ldots,K \\
& \boldsymbol{c}'\boldsymbol{x} \leq B \\
& \boldsymbol{t} \geq \boldsymbol{0}, \boldsymbol{0} \leq \boldsymbol{x} \leq \overline{\boldsymbol{x}} \\
& \boldsymbol{x} \in \Re^m,\ \boldsymbol{u} \in \Re^{n\times K},\ \boldsymbol{t} \in \Re^K
\end{aligned}
\tag{5.2}
$$

where $\tilde{\boldsymbol{z}}^1, \ldots, \tilde{\boldsymbol{z}}^K$ are K independent samples of $\tilde{\boldsymbol{z}}$. We use the same samples throughout the iterations of Algorithm 1.

To derive a deterministic approximation of Model (3.11), we note that the following linear program

$$
\begin{aligned}
\bar{d}_{ij} = \min \quad & u_n \\
\text{s.t.} \quad & u_j - u_i - w_{ij} = 0 && \forall (i,j) \\
& u_1 = 0, w_{ij} = 1 \\
& \boldsymbol{w} \geq 0, \boldsymbol{u} \in \Re^n, \boldsymbol{w} \in \Re^m.
\end{aligned}
$$

achieves the optimum value at $\bar{d}_{ij} = 1$. Accordingly, given a set $\mathcal{L} = \{1,2,3,4\}$, we formulate the deterministic approximation of the subproblem as follows.

$$
\begin{aligned}
Z^d(\gamma) = \quad \min \quad & \beta + \frac{\pi^{\mathcal{L}}(u_n^0 - \tau - \beta, \boldsymbol{u}_n)}{\gamma} + \sum_{(i,j)\in\mathcal{E}} \left(\frac{1}{\gamma} \pi^{\mathcal{L}}(-w_{ij}^0, -\boldsymbol{w}_{ij}) \right) & \\
\text{s.t.} \quad & u_j^0 - u_i^0 - b_{ij} + a_{ij}x_{ij} - w_{ij}^0 = 0 & \forall (i,j) \in \mathcal{E} \\
& u_j^{kl} - u_i^{kl} - b_{ij} + a_{ij}x_{ij} - w_{ij}^{kl} = 0 & \forall (i,j),(k,l) \in \mathcal{E} \\
& u_1^0 = 0, u_1^{kl} = 0 & \forall (k,l) \in \mathcal{E} \\
& \boldsymbol{c}'\boldsymbol{x} \le B & \\
& \boldsymbol{0} \le \boldsymbol{x} \le \overline{\boldsymbol{x}}. &
\end{aligned}
\tag{5.3}
$$

We formulate Model (5.3) using an in-house developed software, *PROF* (Platform for Robust Optimization Formulation). The Matlab based software is essentially an SOCP modeling environment that contains reusable functions for modeling multiperiod robust optimization using decision rules. We have implemented bounds for the CVaR measure and expected positivity of a weighted sum of random variables. The software calls upon CPLEX 9.1 to solve the underlying SOCP.

We use the fictitious project introduced in [23] as an experiment. We create a 6 by 4 grid (See Figure 5.1) as the activity network. There are in total 24 nodes and 38 arcs in the activity network. The first node lies at the bottom left corner and the last node lies at the right upper corner. Each arc proceeds either towards the right node or the upper node. Every activity $(i,j) \in \mathcal{E}$ has independent and identically distributed completion time with

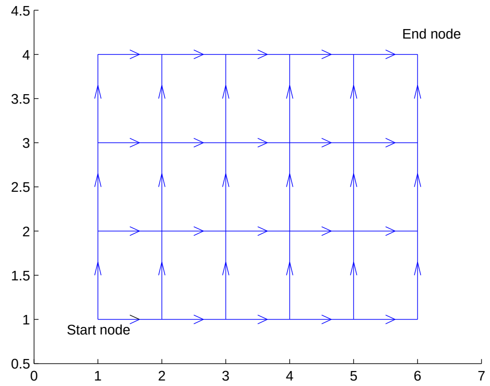


Fig. 5.1: Activity grid 6 by 4

distribution at

$$\mathrm{P}(\tilde{z}_{ij} = z) = \begin{cases} 0.9 & \text{if } z = -25/900 \\ 0.1 & \text{if } z = 25/100. \end{cases}$$

From the distribution of each arc, we can easily calculate the support and deviation information, that is, $\underline{z}_{ij} = 25/900$, $\overline{z}_{ij} = 25/100$, $\sigma_{ij} = 0.0833$, $p_{ij} = 0.1185$, $q_{ij} = 0.0833$. For all activities, we let $a_{ij} = c_{ij} = 1$, $\overline{x}_{ij} = 24$ and $b_{ij} = 100$. We choose an aspiration level of $\tau = 800$. The total cost of resource is kept under the budget B. We compare the performance of the sampling approximation model (5.2) against the deterministic approximation model (5.3). After deciding the allocation of the resource, we use $M = 500,000$ samples to obtain a sampled distribution of the actual completion time $u_n^1, \ldots, u_n^M$. Using these samples we determine the sampled

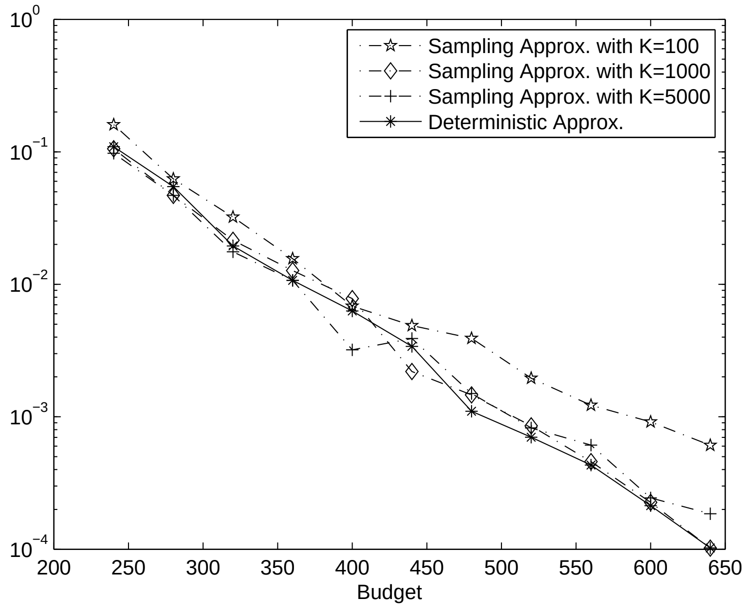


Fig. 5.2: Comparison of the deterministic and sampling models on $(1 - S\hat{ALC})$

shortfall aspiration level criterion as follows.

$$S\hat{ALC} = 1 - \inf_{s>0} \frac{1}{sM} \sum_{k=1}^{M} (u_n^M - \tau + s)^+.$$

We denote $S\hat{ALC}_K^s$ as the sampled shortfall aspiration level criterion when Model (5.2) is used to approximate the subproblem. Likewise, we denote $S\hat{ALC}^d$ as the sampled shortfall aspiration level criterion when Model (5.3) is used in the approximation. By adjusting the budget level, B from 240 to 640, we show the results in Table 5.1 and Figure 5.2, 5.3. In both the

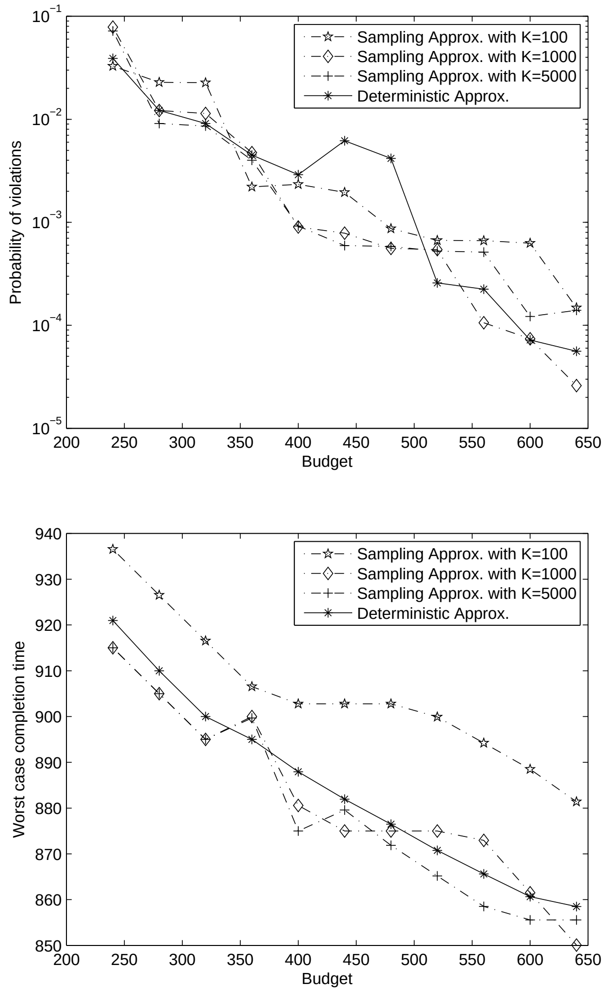


Fig. 5.3: Comparison of the deterministic and sampling models on Probability of violation and worst case completion time.

B	$1-\hat{SALC}_K^d$	$1-\hat{SALC}_{100}^s$	$1-\hat{SALC}_{1000}^s$	$1-\hat{SALC}_{5000}^s$
240	0.1094	0.1602	0.1055	0.0977
280	0.0547	0.0625	0.0469	0.0469
320	0.0195	0.0322	0.0215	0.0176
360	0.0107	0.0156	0.0127	0.0107
400	0.0063	0.0068	0.0078	0.0032
440	0.0034	0.0049	0.0022	0.0039
480	0.0011	0.0039	0.0015	0.0015
520	7.02×10^{-4}	0.0020	8.54×10^{-4}	8.24×10^{-4}
560	4.31×10^{-4}	0.0012	4.58×10^{-4}	6.10×10^{-4}
600	2.14×10^{-4}	9.16×10^{-4}	2.26×10^{-4}	2.44×10^{-4}
640	1.02×10^{-4}	6.10×10^{-4}	1.02×10^{-4}	1.86×10^{-4}

Tab. 5.1: Comparison of the deterministic and sampling models on $(1-\hat{SALC})$.

deterministic and the sampling approximations, we observe that γ decreases with increasing budget levels. We also see that the probability of violation does not exactly represent the risk. For instance, we compare the results for the deterministic method and the sampling method with 100 sample size. The former one has higher probability of violation than the latter one for budget= 240, 360, 400, 440 and 480. However the former one has a shorter worst case completion time than the latter one for all budgets.

It is evident that when the number of samples are limited, sampling approximation can perform poorly. Moreover, due to the variability of sampling approximation, the performance does not necessarily improve with more samples; see Table 5.1 with $B = 440, 560, 600, 640$. We note that despite the modest distributional assumption and the non-optimal recourse, the performance of the deterministic approximation is rather comparable with the performance of the sampling approximation where sufficient number of samples are used.

5.2 Case Study: NFL Replica Jerseys

We adopt the case addressing the inventory planning for the National Football League (NFL) replica jerseys from John C. W. Parsons' thesis. We relax the contract requirement between the retailers and the distributor and focus on an optimal postponement strategy.

NFL is the premier professional league for American football. It consists of 32 teams. The football season is between September and January, with 16 regular games per team. During this period, the football fans have high demand for the replica jerseys of their favorite players. In December 2000, Reebok signed a 10 year contract with NFL to provide the replica jerseys. Since the demand of the jerseys is driven by the fans feel for the game, it is influenced by many uncontrollable factors. The long lead time (See Figure 5.4) makes it impossible for Reebok to determine the purchasing quantity after the demand is exactly predicted.

Each team's jersey has a distinct combination of style, colors, cuts and team logo, but different players' jerseys of a same team are the same except

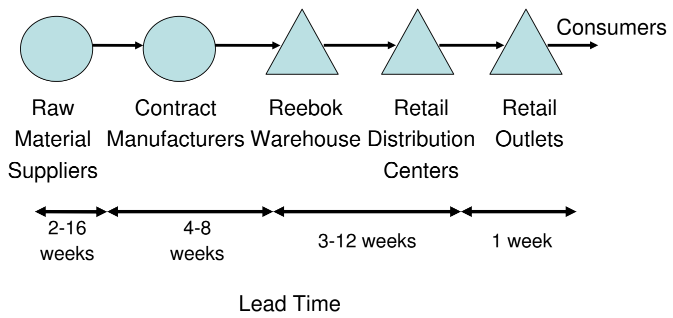


Fig. 5.4: Supply chain

for the name and number. However, one player's jersey is not substituted by another player's jerseys due to the customer preference. It may happen that for one player's jerseys, there's overstock, but for another player's jerseys, there's under stock. To avoid this kind of waste, Reebok has two options to purchase the jerseys from international contract manufacturers: blank jerseys and dressed jerseys. A blank jersey is a jersey with only team markings and without player's name and number. A dressed jersey is a completed jersey with specific player's name and number. Reebok can transform the blank jerseys to the dressed jerseys in its distribution center (Indianapolis) with a higher cost than the international contract manufacturers. This provides a valuable postponement opportunity. The problem is what kind of strategy Reebok should use to decide the purchasing quantity of the blank and dressed jerseys.

The demand of the replica jerseys is very sensitive to the game performance, so it fluctuates every year and even the demand distribution changes every year. Here, we consider the risk that the profit is less than a given target profit R and apply the goal driven model to decide the optimal postponement strategy. We use the planning problem for New England Patriots of the 2003 season as an example. The notations and data are as follows.

$n = 7$	: Number of products;
$p = \$24$	: Unit selling price for dressed jerseys;
$c = \$10.9$	: Unit purchasing cost for dressed jerseys;
$c_0 = \$9.5$	: Unit purchasing cost for blank jerseys;
$s = \$7$	: Unit salvage value for dressed jerseys;
$h = \$8.46$	: Unit salvage value for blank jerseys;
$\tilde{d}_i$	: Demand for the ith player's jerseys;
$\tilde{d}$	: Total demand of the replica jerseys.

	Player	Mean μ	Stdev σ
	New Eng Patriot Total	87680	19211
1	Brady, Tom #12	30763	13843
2	Law, TY #24	10569	4756
3	Brown, Troy #80	8159	3671
4	Vinatieri, Adam #04	7270	4362
5	Bruschi, Tedy #54	5526	3316
6	Smith, Antowain #32	2118	1271
7	Other players	23275	10474

Tab. 5.2: Demand prediction for New England Patriots of the 2003 season

5.2.1 Full postponement strategy

There are various postponement strategies to help Reebok to decide the purchasing quantity of the blank and dressed jerseys. One intuitive strategy is full postponement, in other words, purchasing only blank jerseys. Then the problem reduces to a single period newsvendor problem. We show that the goal driven model can be solved exactly in this case. We denote $Q_0 \in \Re_+$ as the purchasing quantity of blank jerseys and formulate the problem as follows.

$$\max \quad SALC(-f(Q_0, \tilde{d}) + R), \tag{5.4}$$

where

$$\begin{aligned} f(Q_0, d) &\triangleq (p - c_0 - e)Q_0 + (h - p)(Q_0 - d)^+ \\ &= \begin{cases} (h - c_0 - e)Q_0 + (p - h)d & \text{if } d < Q_0 \\ (p - c_0 - e)Q_0 & \text{otherwise.} \end{cases} \end{aligned} \tag{5.5}$$

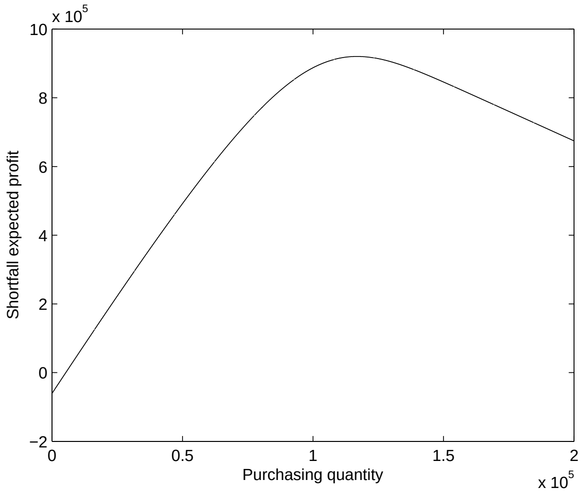


Fig. 5.5: Shortfall expected profit vs Purchasing quantity

From the definition of the CVaR measure and the translation invariance property, we know that Model (5.4) is equivalent to

$$
\begin{aligned}
\max \quad & 1-\gamma \\
\text{s.t.} \quad & \min_v \left(\frac{\mathrm{E}\Big((v - f(Q_0, \tilde{d}))^+ \Big)}{\gamma} - v \right) + R \leq 0.
\end{aligned}
\tag{5.6}
$$

Theorem 3 implies that we can decide the optimal purchasing quantity and calculate the shortfall aspiration level criterion efficiently if the distribution of the demand is known. We use New England Patriots (2003 season) as an illustrative example and assume that the total demand follows normal

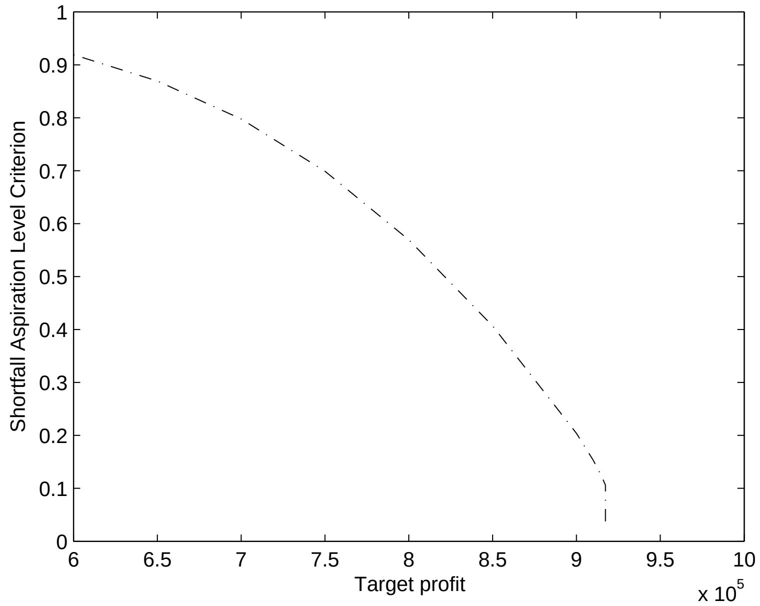


Fig. 5.6: Full postponement: $SALC$ vs Target profit

distribution. We plot

$$\mathrm{E}\Big(f(Q_0,\tilde{d}) \mid \tilde{d} < Q_0\Big)$$

for different value of Q_0 (See Figure 5.5). Then for a given target profit value R, we can calculate the optimal Q_0^* using binary search. With Q_0^*, the equation

$$\gamma^* = \mathrm{P}(\tilde{d} < Q_0^*)$$

provides a close form to calculate the shortfall aspiration level criterion. Figure 5.6 shows the shortfall aspiration level criterion $SALC$ for different target profit values. We see that $SALC$ decreases as the target profit increases. This coincides with our intuition.

5.2.2 Partial postponement strategy

Although full postponement strategy is easy to perform, it may not be the optimal strategy to maximize the shortfall aspiration level criterion. In this section, we consider the partial postponement strategy. We denote Q_i as the purchasing quantity of ith player's jerseys and formulate the problem as follows.

$$\max \quad SALC\Big(-w(Q_0, \boldsymbol{Q}, \tilde{\boldsymbol{d}}) + R\Big) \tag{5.7}$$

where

$$\begin{aligned} & w(Q_0, \boldsymbol{Q}, \boldsymbol{d}) \triangleq \\ \max \quad & \sum_i \Big((p-c)Q_i + (p-c_0-e)q_i + (s-p)(Q_i+q_i-d_i)^+\Big) + (h-c_0)(Q_0 - \sum_i q_i)^+ \\ \text{s.t.} \quad & \sum_i q_i \le Q_0 \\ & q_i \ge 0, \end{aligned} \tag{5.8}$$

where $q_i(\boldsymbol{d})$ is the recourse variable representing the quantity of blank jerseys transformed into ith player's jerseys when the demand $\boldsymbol{d}$ is realized. Using Algorithm 1, we reduce the problem (5.1) to solving a sequence of subproblems in the form of stochastic optimization problems with CVaR objectives. Since we do not have the full knowledge of the demand distribution, first we use the deterministic model to solve the problem. To derive a deterministic approximation of the subproblem, we assume that the demand $\boldsymbol{d}$ are affinely dependent on some primitive uncertainties $\tilde{\boldsymbol{z}}$, which has zero mean and standard deviation σ_i, that is

$$\boldsymbol{d}(\tilde{\boldsymbol{z}}) = \boldsymbol{d}^0 + \sum_{j=1}^{N} \boldsymbol{d}^j \tilde{z}_j,$$

where

$$\begin{aligned} &d_i^0 = \mu_i,\ d_i^i = 1, && \forall i = 1, \ldots, n \\ &d_i^j = 0, && \forall i \neq j. \end{aligned}$$

We note that the constraints of the model (5.8) are just the limit on the postponement quantity and transshipment. We notice that one extra unit of postponement quantity may bring $(p - c_0 - e)$ extra profit, one unit of transshipment from one dressed jersey to another dressed jersey may bring $(p - s)$ extra profit and one unit of transshipment from dressed jersey to blank jersey may bring $(h + e - s)$ extra profit. Therefore, we deduct the extra profit from the objective and have the following equivalent formulation.

$$\begin{aligned} & w(Q_0, \boldsymbol{Q}, \boldsymbol{d}) \\ = \ \max \ & \sum_i \Big((p-c)Q_i + (p - c_0 - e)q_i + (s-p)(Q_i + q_i - d_i)^+ - (p-s)q_i^- \Big) \\ & +(h - c_0)(Q_0 - \textstyle\sum_i q_i)^+ - (p - c_0 - e)(\textstyle\sum_i q_i - Q_0)^+ \\ = \ \max \ & \sum_i \Big((p-c)Q_i + (p - h - e)q_i + (s-p)(Q_i + q_i - d_i)^+ - (p-s)q_i^- \Big) \\ & +(h - c_0)Q_0 + (h + e - p)(\textstyle\sum_i q_i - Q_0)^+. \end{aligned}$$

We assume that $\boldsymbol{q}$ follows linear decision rule, that is

$$\boldsymbol{q}(\tilde{\boldsymbol{z}}) = \boldsymbol{q}^0 + \sum_{j=1}^{N} \boldsymbol{q}^j \tilde{z}_j.$$

Let $\mathcal{L} = \{1, 3, 4\}$. Then the subproblem can be approximated as

$$\begin{aligned} Z_d(\gamma) &\triangleq \\ \min \quad & \beta + \frac{1}{\gamma}\pi^{\mathcal{L}}\left((c_0 - h)Q_0 + \sum_i \left((c-p)Q_i + (h+e-p)q_i^0\right) - \beta, (h+e-p)\sum_i \boldsymbol{q}_i\right) \\ & + \frac{1}{\gamma}(p-s)\sum_i \pi^{\mathcal{L}}(Q_i + q_i^0 - d_i^0, \boldsymbol{q}_i - \boldsymbol{d}_i) + \frac{1}{\gamma}(p-s)\sum_i \pi^{\mathcal{L}}(-q_i^0, -\boldsymbol{q}_i) \\ & + \frac{1}{\gamma}(p-h-e)\pi^{\mathcal{L}}(\sum_i q_i^0 - Q_0, \sum_i \boldsymbol{q}_i). \end{aligned} \tag{5.9}$$

Since the demand for "other players" jerseys is hard to predict, we use blank jerseys to satisfy this part of demand, by adding one more constraint to the model.

$$Q_7 = 0.$$

After deciding the purchasing quantity of the blank and dressed jerseys, we use $M = 500,000$ samples following a test distribution to obtain the frequency of the profit in each interval. If the interval is small enough, the frequency almost represents the distribution of the profit. We test the solutions on normal distribution. Figure 5.7 shows the comparison between the goal driven model and the model maximizing the expected profit. It can be seen that the goal driven model results in lower risk attaining the target level. Figure 5.8 shows the solutions for target level 650,000 and 900,000. We see that the risk attaining the target profit increases as the target level increases.

Based on the simulated profit, we can also estimate the shortfall aspiration level criterion. We compare the full postponement strategy and the partial postponement strategy for different target levels (See Figure 5.9). Although we do not assume the demand distribution when applying the partial postponement strategy, it outperforms the full postponement strategy, especially for higher target levels.

We also notice that this problem has relatively complete recourse, that is, for any given Q_0 and $\boldsymbol{Q}$, there always exists a feasible $\boldsymbol{q}$. Therefore, we did another test to see whether we can use an assumed distribution and apply the sampling method to decide the purchasing quantity. We assume the demand follows independent exponential distribution and the subproblem

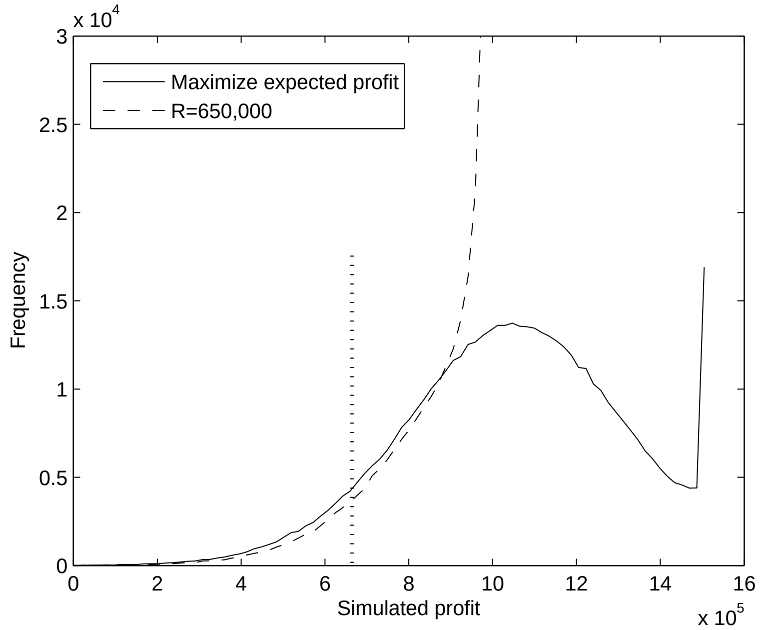


Fig. 5.7: Partial postponement: Frequency of the profit – Goal driven model vs Maximizing expected profit

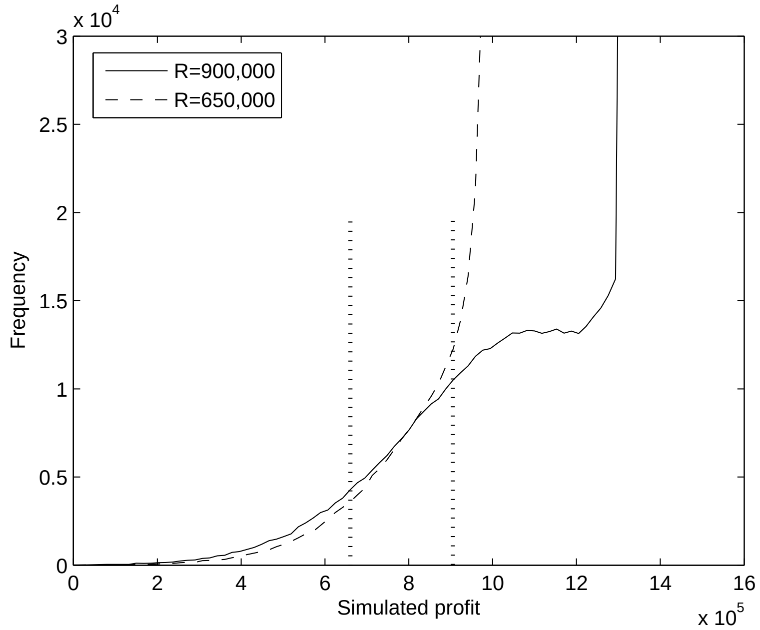


Fig. 5.8: Partial postponement: Frequency of the profit – Different target levels

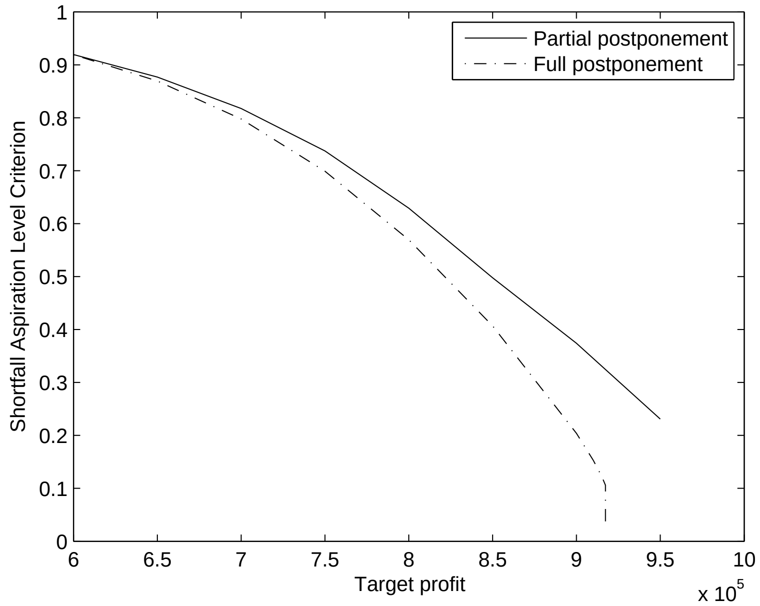


Fig. 5.9: Partial postponement vs Full postponement

can be approximated as follows.

$$
\begin{aligned}
& \tilde{Z}^s_K(\gamma) = \\
\min \quad & \tfrac{1}{\gamma K}\textstyle\sum_{k=1}^K t_k - v \\
\text{s.t.} \quad & t_k \geq R + v - \omega_k && \forall k = 1, \ldots, K \\
& \omega_k = \textstyle\sum_i \Big((p-c)Q_i + (p - c_0 - e)q_i^k && +(s-p)y_i^k \Big) + (h - c_0)(Q_0 - \textstyle\sum_i q_i^k) \\
& && \forall k = 1, \ldots, K \\
& y_i^k \geq Q_i + q_i^k - d_i^k && \forall i, \forall k = 1, \ldots, K \\
& \sum_i q_i^k \leq Q_0 && \forall k = 1, \ldots, K \\
& \boldsymbol{q} \geq 0,\ \boldsymbol{t} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0},
\end{aligned}
\tag{5.10}
$$

where $\boldsymbol{d}^1, \ldots, \boldsymbol{d}^K$ are K independent samples of $\tilde{\boldsymbol{d}}$.

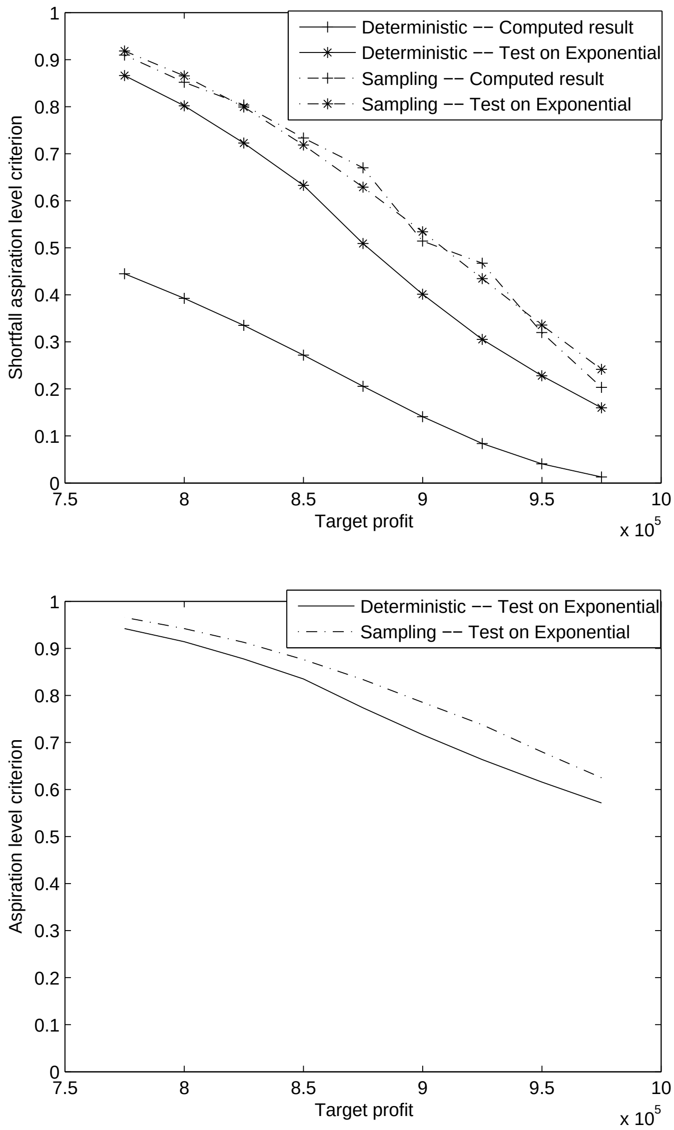


Fig. 5.10: Test on exponential distribution

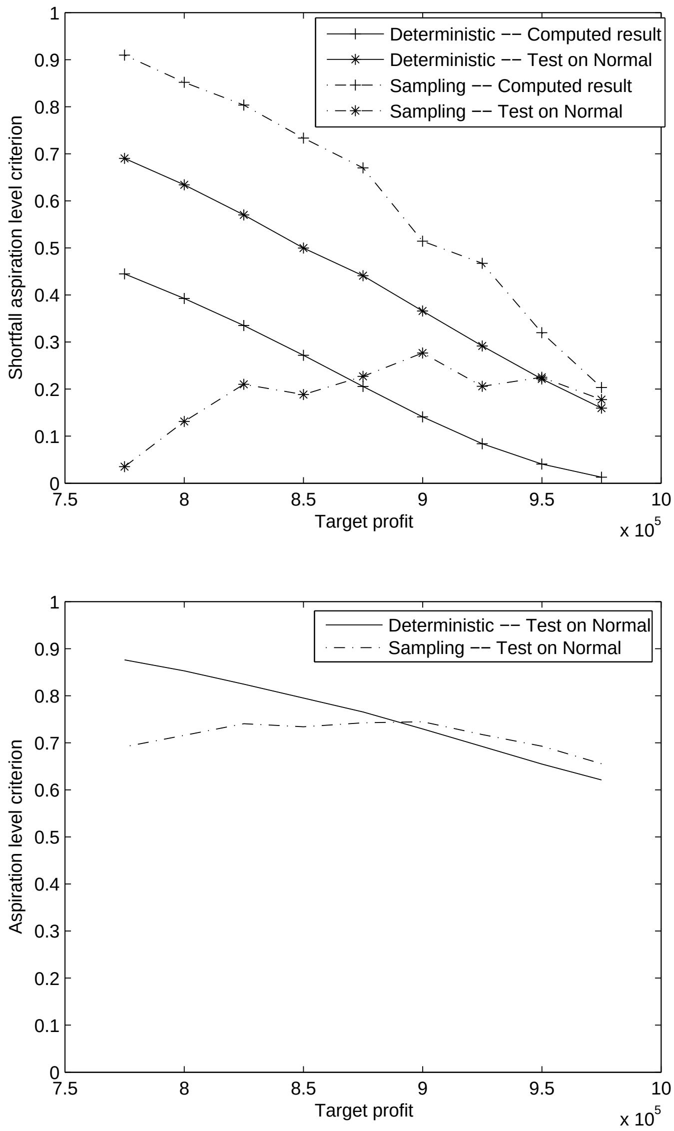


Fig. 5.11: Test on normal distribution

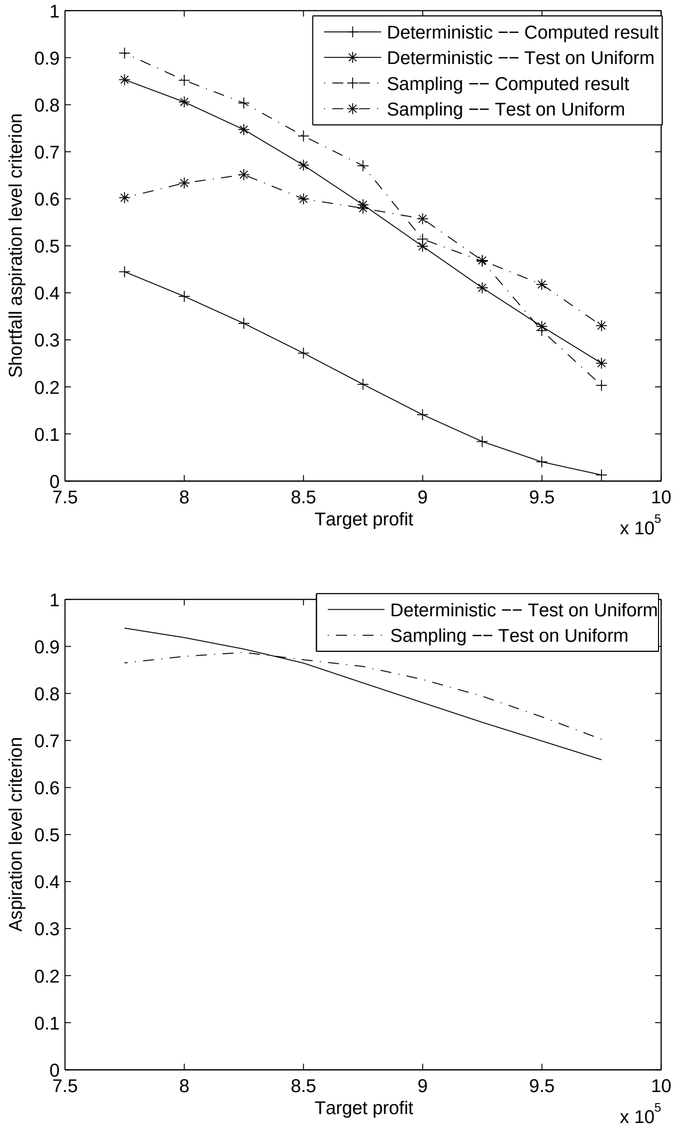


Fig. 5.12: Test on uniform distribution

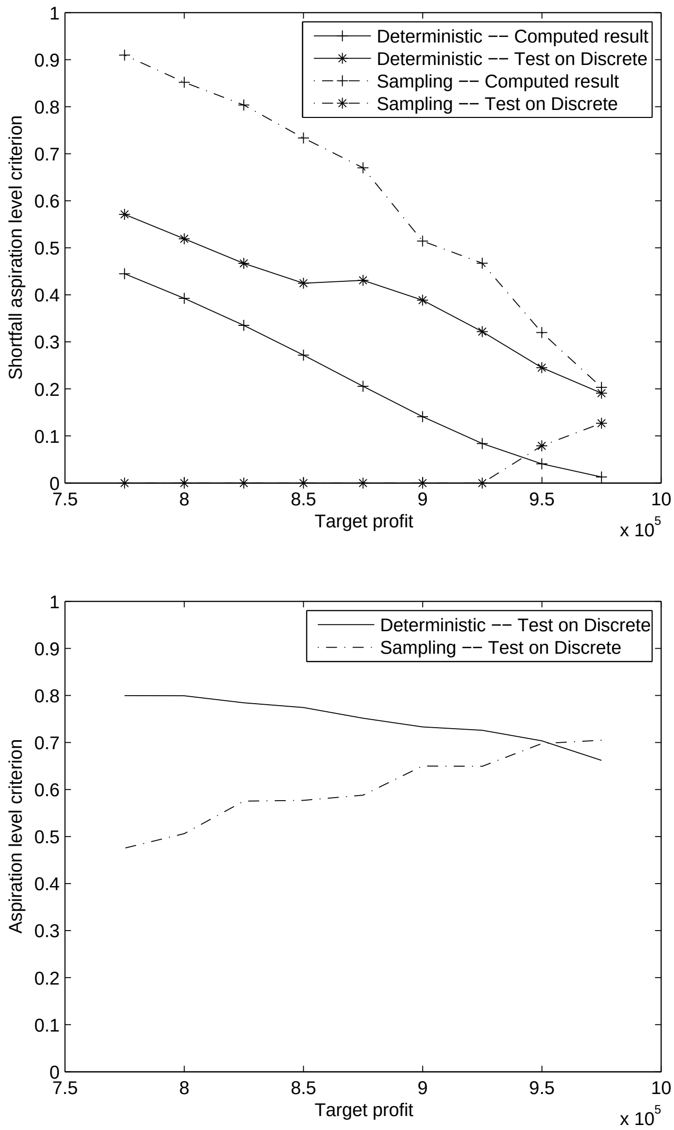


Fig. 5.13: Test on two point discrete distribution

We compare the performances of the sampling approximation model (5.10) against the deterministic approximation model (5.9). First, we test the two models on exponential distribution (See Figure 5.10). It can be seen that the sampling method outperforms the deterministic method in both shortfall aspiration level criterion and the aspiration level criterion. Second, we test two models on other distributions: normal, uniform and two point discrete distribution (See Figure 5.11, 5.12, 5.13). It can be seen that with a wrong assumed distribution, the sampling methods performs poorly compared with the deterministic method. Besides, the objective of the deterministic model provides a lower bound of the shortfall aspiration level criterion for all demand distributions with the same mean and deviation value.

5.2.3 Tradeoff between profit and service level

The previous models only consider the profit when deciding the postponement strategy and the purchasing quantity. In practice, achieving a high profit is not the only objective when making decisions. Another consideration is the service level, which influences the customer demand in the future. However, to achieve a higher service level, it is usual that the risk attaining a target profit also increases. We propose a model to tradeoff between these two considerations.

$$\begin{aligned} \max \quad & SALC\Big(- w(Q_0, \boldsymbol{Q}, \tilde{\boldsymbol{d}}) + R\Big) \\ \text{s.t.} \quad & \mathrm{P}\Big(\sum_i (\tilde{d}_i - Q_i)^+ \leq Q_0\Big) \geq 1 - \epsilon, \end{aligned} \tag{5.11}$$

where ϵ is a given risk requirement and the probabilistic constraint guarantees the service level $1 - \epsilon$. We introduce a recourse variable $\boldsymbol{v}(\tilde{\boldsymbol{z}})$, which is also a function of the primitive uncertainties $\tilde{\boldsymbol{z}}$ and reformulate the probabilistic constraint as a joint probabilistic constraint.

$$\mathrm{P}\begin{pmatrix} \boldsymbol{d}(\tilde{\boldsymbol{z}}) - \boldsymbol{Q} \geq \boldsymbol{v}(\tilde{\boldsymbol{z}}) \\ \boldsymbol{v}(\tilde{\boldsymbol{z}}) \geq \boldsymbol{0} \\ \sum_i v_i(\tilde{\boldsymbol{z}}) \leq Q_0 \end{pmatrix} \geq 1 - \epsilon \tag{5.12}$$

To simplify the problem, we let $\boldsymbol{v}$ follows linear decision rule, that is

$$\boldsymbol{v}(\tilde{\boldsymbol{z}}) = \boldsymbol{v}^0 + \sum_{j=1}^{N} \boldsymbol{v}^j z_j.$$

With introduced recourse variables $\boldsymbol{r}(\tilde{\boldsymbol{z}}), \boldsymbol{t}(\tilde{\boldsymbol{z}}), u(\tilde{\boldsymbol{z}}), \boldsymbol{y}(\tilde{\boldsymbol{z}})$, which also follow linear decision rule, that is

$$\begin{aligned}
\boldsymbol{r}(\tilde{\boldsymbol{z}}) &= \boldsymbol{r}^0 + \textstyle\sum_{j=1}^N \boldsymbol{r}^j z_j \\
\boldsymbol{t}(\tilde{\boldsymbol{z}}) &= \boldsymbol{t}^0 + \textstyle\sum_{j=1}^N \boldsymbol{t}^j z_j \\
u(\tilde{\boldsymbol{z}}) &= u^0 + \textstyle\sum_{j=1}^N u^j z_j \\
\boldsymbol{y}(\tilde{\boldsymbol{z}}) &= \boldsymbol{y}^0 + \textstyle\sum_{j=1}^N \boldsymbol{y}^j z_j,
\end{aligned}$$

we can transform the problem to the standard form as follows.

$$\begin{aligned}
\max \quad & SALC\Big(-w(Q_0, \boldsymbol{Q}, \tilde{\boldsymbol{d}}) + R\Big) \\
\text{s.t.} \quad & \boldsymbol{d}(\tilde{\boldsymbol{z}}) - \boldsymbol{Q} - \boldsymbol{r}(\tilde{\boldsymbol{z}}) = \boldsymbol{v}(\tilde{\boldsymbol{z}}) \\
& \boldsymbol{v}(\tilde{\boldsymbol{z}}) - \boldsymbol{t}(\tilde{\boldsymbol{z}}) = \boldsymbol{0} \\
& \sum_i v_i(\tilde{\boldsymbol{z}}) + u(\tilde{\boldsymbol{z}}) = Q_0 \\
& \boldsymbol{y}(\tilde{\boldsymbol{z}}) = \begin{pmatrix} \boldsymbol{r}(\tilde{\boldsymbol{z}}) \\ \boldsymbol{t}(\tilde{\boldsymbol{z}}) \\ u(\tilde{\boldsymbol{z}}) \end{pmatrix} \\
& \mathrm{P}(\boldsymbol{y}(\tilde{\boldsymbol{z}}) \leq \boldsymbol{0}) \geq 1 - \epsilon.
\end{aligned} \tag{5.13}$$

Therefore we can apply the methodologies proposed in Chapter 4 to solve the problem. After deciding the purchasing quantity, we simulate $M = 500,000$ scenarios following normal distribution to estimate the short-

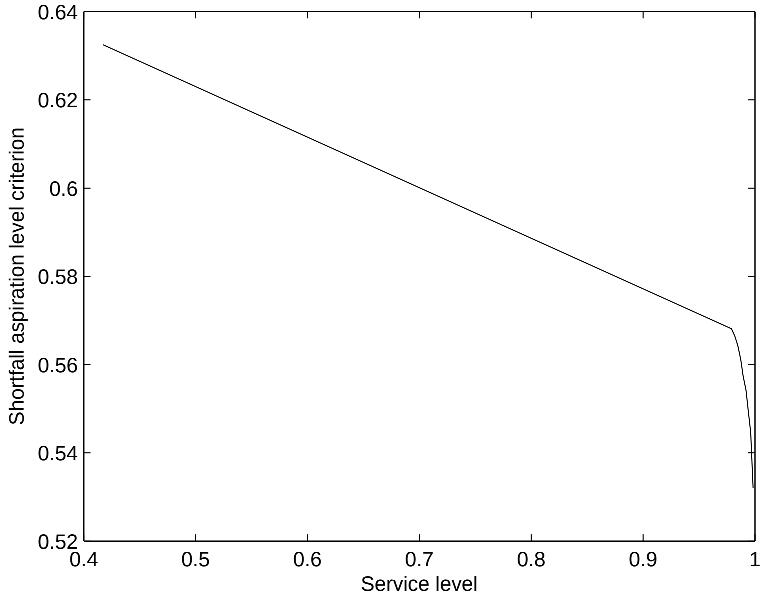


Fig. 5.14: Tradeoff between risk and service level

fall aspiration level criterion and service level. Figure 5.14 shows the tradeoff between the shortfall aspiration level criterion and the service level for target level $800,000$. It can be seen that as the service level increases, the shortfall aspiration level decreases, which implies that the risk attaining the target level increases. This coincides with our intuition and the relation between the service level and the risk provides a useful tool for aiding in making decisions.

6. CONCLUSION

6.1 Summary of Results

This thesis proposed a mathematical model, goal driven stochastic optimization model, which helps the decision maker to achieve a target level, or an aspiration level, with low risk. Specifically, the main results are as follows.

- **Shortfall aspiration level criterion:** The new criterion incorporates both the probability of success in achieving the target level and an expected level of under-performance or shortfall. The goal driven model applies the shortfall aspiration level criterion as its objective. The key advantage is its tractability. We showed that the goal driven model can be exactly solved for single product newsvendor problem. For more complicated problems, we showed that the proposed model is reduced to solving a small collections of stochastic linear optimization problems with objectives evaluated under the CVaR measure.

- **Deterministic approximation for goal driven model:** Using techniques in robust optimization, we proposed a decision rule based deterministic approximation of the goal driven optimization problem by solving a polynomial number of second order cone optimization problems (SOCP) with respect to the desired accuracy. The advantages of this approximation over the sampling approximation are: (1) it requires mild distributional assumptions, such as mean, support and deviation measures; (2) the size of the problem does not increase exponentially as the dimension of the problem.

- **Methodology to solve probabilistic constrained problem:** We reviewed the SOCP approximations of the individual probabilistic constraint and show that the the approximation of the CVaR measure is related to robust optimization. For the joint probabilistic constraint, we showed that Bonferroni's inequality may be rather poor in approximating constraints with uncertainties that are correlated with each other. We proposed a new formulation to approximate the joint probabilistic constraint and investigated its properties. In particular, we showed that it outperforms any solution obtained by Bonferroni's inequality.

The methodologies proposed in this thesis were applied to project management and inventory planning problems to test the tractability. The comparison between the goal driven model and the classical model shows that the new decision criterion can help the decision maker to minimize the risk attaining a target level. The comparison between the sampling approximation and the deterministic approximation shows that the latter is more robust and stable when the decision maker has no full knowledge of the distribution of the random data.

Moreover, we applied the goal driven model with joint probabilistic constraint to tradeoff between the risk achieving a target profit and the service level when deciding the inventory level. This idea helps to make decision more practically.

6.2 Future Studies.

This thesis only considers the linear structure of the stochastic optimization model. In the future, it is worthwhile to consider other cases and derive more efficient methodologies. Some possible theoretical researches are as follows:

- Use other risk measure to consider the risk attaining the target level.
- Extend the model to the multi-period problem.

Apart from the above possibilities, it is worthwhile to apply the methodologies to other areas, such as portfolio management, control in engineering, and so on. This may contribute to a better understanding of the merit and weakness of the methodologies.

BIBLIOGRAPHY

[1] S. Ahmed (2006): Convexity and decomposition of mean-risk stochastic programs, *Mathematical Progamming*, 106, 433-446.

[2] F. Alizadeh, D. Goldfarb, (2003): Second-order cone programming, *Mathematical Programming*, 95, 3-51.

[3] Ph. Artzner,F. Delbaen, J.M. Eber, D. Heath (1999): Coherent Risk Measures, *Mathematical Finance*, 9(3), 203-228.

[4] M. Ben-Daya, A. Raouf (1993): On the constrained multi-item single-period inventory problem, *International Journal of Operations and Production Management*, 13, 104-112.

[5] A. Ben-Tal, A. Goryashko, E. Guslitzer and A. Nemirovski (2002): Adjusting Robust Solutions of Uncertain Linear Programs, *Working Paper*, Technion.

[6] Ben-Tal, A. Nemirovski (1998): Robust convex optimization, *Math. operations research*, 23, 769-805.

[7] Ben-Tal, A. Nemirovski (1999): Robust solutions to uncertain programs, *Operations research letters*, 25, 1-13.

[8] A. Ben-Tal, A. Nemirovski (2000): Robust solutions of linear programming problems contaminated with uncertain data, *Mathmatical programming*, 88, 411-424.

[9] A. Ben-Tal, A. Nemirovski, (2000): Robust solutions of linear programming problems contaminated with uncertain data, *Mathmatical programming*, 88, 411-424.

[10] A. Ben-Tal, A. Nemirovski (2001): Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications, *MPR-SIAM Series on Optimization*, SIAM, Philadelphia.

[11] A. Ben-Tal, A. Goryashko, E. Guslitzer, A. Nemirovski, (2002): Adjusting robust solutions of uncetain linear programs, *working paper*.

[12] B. Bereanu (1964): Programme de risque minimal en programmation linéaire stochastique, *C.R. Acad. Sci. (Paris)*, tom 259, nr. 5, 981-983.

[13] Dimitri P. Bertsekas, (1999): Nonlinear programming, *Imprint Belmont, Mass. : Athena Scientific*.

[14] D. Bertsimas, M. Sim, (2004): Price of robustness, *Operations research*, 52, 35-53.

[15] R. Bordley, M. LiCalzi (2000): Decision analysis using targets instead of utility functions. *Decisions in Economics anf Finance*, 23, 53-74.

[16] J.R. Birge, F. Louveaux (1997): Introduction to Stochastic Programming, *Springer*, New York.

[17] M. Ben-Daya, A. Raouf (1993): On the constrained multi-item single-period inventory problem, *International Journal of Operations and Production Management*, 13, 104-112.

[18] G. Calafiore, M. C. Campi, (2004): Decision making in an uncertain environment: the scenario-based optimization approach, *working paper.*

[19] J.R. Canada, W.G. Sullivan, D.J. Kulonda, J.A. White (2005): Capital investment analysis for engineering and management, *Prentice Hall* 3rd ed.

[20] E.R. Castagnoli, M. LiCalzi (1996): Expected utility without utility, *Theory and Decision*, 41, 281-301.

[21] A. Charnes, W.W. Cooper, G.H. Symonds (1958): Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil, *Management Science* 4, 235-263.

[22] A. Charnes, W.W. Cooper (1963): Deterministic equivalents for opti-

mizing and satisficing under chance constraints, *Operations research*, 11, 18-39.

[23] X. Chen, M. Sim, P. Sun (2006): A robust optimization perspective of stochastic programming, accepted in *Operations Research*.

[24] X. Chen, M. Sim, P. Sun, J. Zhang (2006): A tractable approximation of stochastic programming via robust optimization *Working Paper, Revised June 2006*, National University of Singapore.

[25] E. Diecidue, J. van de Ven (2005): Aspiration level, probability of success, and expected Utility. *Working Paper*, INSEAD.

[26] M. Dragomirescu (1972): An algorithm for the minimum-risk problem of stochastic programming, *Operations Research*, 20(1), 154-164.

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

[28] El-Ghaoui, L. F. Oustry, H. Lebret, (1998): Robust solutions to uncertain semidefinite programs, *SIAM journal of optimization*, 9, 33-52.

[29] H. I. Gassmann, I. Deak, T. Szantai, (2002): Computing multivariate normal probabilities: a new look, *Journal of computational and graphical statistics*, 11, 920-949.

[30] A. Genz, (1992): Numerical computation of the multivariate normal probabilities, *Journal of computational and graphical statics*, 1, 141-150.

[31] D. Goldfarb, K. Scheinberg, (1999): On parametric semidefinite programming, *Applied Numerical Mathematics*, 29, 361-377.

[32] D. Goldfarb, G. Iyengar (2003): Robust quadratically constrained programs, *Math. Prog. Ser. B*, 97(3), 495-515

[33] G. Hadley, T. Whitin (1963): Analysis of Inventory Systems, *Prentice Hall Englewood Cliffs, NJ.*

[34] R.F. Lanzillotti (1958): Pricing objectives in large companies, *American Economic Review*, 48,921-940.

[35] H.S. Lau, A.H.L. Lau (1988): Maximizing the probability of achieving a target profit in a two-product newsboy problem, *Decision Sciences*, 19, 392-408.

[36] H.S. Lau, A.H.L. Lau (1988): Maximizing the probability of achieving a target profit in a two-product newsboy problem, *Decision Sciences*, 19, 392-408.

[37] H.S. Lau, A.H.L. Lau (1996): The newsstand problem: A capacitated multiple-product single-period inventory problem, *European Journal of Operational Research* 94, 29-42.

[38] J. Li, H.S. Lau, A.H.L. Lau (1991): A two-product newsboy problem with satisficing objective and independent exponential demands, *IIE Transactions*, 23(1), 29-39.

[39] J.C.T. Mao (1970): Survey of Capital Budgeting: Theory and Practice, *Journal of Finance*, 25, 349-360.

[40] K. Marti, Konvexitatsaussagen zum linearen stochastischen optimierungsproblem, Z. Wahrsch. theorie u. verw. Geb., 18: 159-166, 1971

[41] J. Mayer, (1998): Stochastic linear programming algorithms, *Gordon and Breach, Amsterdam.*

[42] L. B. Miller, H. Wagner, (1965): Chance-constrained programming with joint constraints, *Operations research*, 13, 930-945.

[43] K. Natarajan, D. Pachamanova, M. Sim (2006): A Tractable Parametric Approach to Value-at-Risk Optimization, Working Paper, NUS Business School.

[44] A. Nemirovski, A. Shapiro (2006): Convex approximation of chance constrained programs, *SIAM Journal of Optimization*, 17(4), 969-996.

[45] W. Ogryczak, A. Ruszczynski (2002): Dual stochastic dominance and related mean-risk model, *SIAM Journal of Optimization*, 13, 60-78.

[46] Payne, J.W., Laughhunn, D.J. and Crum, R.(1980): Tanslation of gambles and aspiration level effects in risky choice behaviour, *Management Science*, 26, 1039-1060.

[47] J.W. Payne, D.J. Laughhunn, R. Crum (1980): Further tests of aspiration level effects in risky choice behaviour, *Management Science*, 27, 953-958.

[48] M. Parlar, K. Weng (2003): Balancing desirable but conflicting objectives in the newsvendor problem, *IIE Transactions*, 35, 131142.

[49] A. Prekopa, (1995): Stochastic programming, *Kluwer, Dordrecht*.

[50] A. Prekopa, (1970): On probabilistic constrained programming, *Mathematical programming study*, 28, 113-138.

[51] A. Prekopa, (1971): Logarithmic concave measures with applications to stochastic programming, *Acta. Sci. Math*, 32, 301-316.

[52] M. Riis, R. Schultz (2003): Applying the minimum risk criterion in stochastic recourse programs, *Computational Optimization and Applications*, 24, 267-287.

[53] T. Rockafellar, S. Uryasev (2000): Optimization of Conditional Value-at-Risk, *Journal of risk* 2, 21-41.

[54] T. Rockafellar, S. Uryasev (2002): Conditional value-at-risk for general loss distributuions, *Journal of Banking and Finance*, 26.

[55] A. Ruszczynski, A. Shapiro (2004): Optimization of convex risk functions, Available at SSRN: http://ssrn.com/abstract=675461.

[56] E. Sankarasubramanian, S. Kumaraswamy (1983): Optimal ordering quantity for pre-determined level of profit, *Management Science*, 29, 512-514.

[57] A. Shapiro, A. Nemirovski (2006): On complexity of stochastic programming problems, *Applied Optimization, Springer, US*, 99, 111-146.

[58] H.A. Simon (1959): Theories of decision-making in economics and behavioral science, *The American Economic Review*, 49(3),253-283.

APPENDIX

.1 Proof of Theorem 5

(a) Since $\mathcal{W}$ is the support set of $\tilde{\boldsymbol{z}}$, we have

$$\mathrm{E}\left((y_0+\boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) \le \underbrace{(y_0+\max_{\boldsymbol{z}\in\mathcal{W}}\boldsymbol{y}'\boldsymbol{z})^+}_{=\pi^1(y_0,\boldsymbol{y})}.$$

Note that whenever, $y_0+\max_{\boldsymbol{z}\in\mathcal{W}}\boldsymbol{y}'\boldsymbol{z} \le 0$, it is trivial to see that $\mathrm{E}\left((y_0+\boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) = 0 = \pi^1(y_0,\boldsymbol{y})$.

Hence,

$$\begin{aligned}
\psi_{1-\gamma}(y_0+\boldsymbol{y}'\tilde{\boldsymbol{z}}) &\le \min_\theta\left(\theta+\frac{\pi^1(y_0-\theta,\boldsymbol{y})}{\gamma}\right) \\
&= \min_\theta\left(\theta+\frac{1}{\gamma}(y_0-\theta+\max_{\boldsymbol{z}\in\mathcal{W}}\boldsymbol{y}'\boldsymbol{z})^+\right) \\
&= y_0+\max_{\boldsymbol{z}\in\mathcal{W}}\boldsymbol{y}'\boldsymbol{z}+\min_\theta\left(\theta+\frac{1}{\gamma}(-\theta)^+\right) \\
&= y_0+\max_{\boldsymbol{z}\in\mathcal{W}}\boldsymbol{y}'\boldsymbol{z} \\
&= \eta^1_{1-\gamma}(y_0,\boldsymbol{y}),
\end{aligned}$$

where the last equality is due to $\min_\theta\left(\theta+\frac{1}{\gamma}(-\theta)^+\right)=0$ for all $\gamma\in(0,1)$.

(b) Since $w^+ = w+(-w)^+$, we have

$$\mathrm{E}\left((y_0+\boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) = y_0+\mathrm{E}\left((-y_0-\boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) \le \underbrace{y_0+\left(-y_0+\max_{\boldsymbol{z}\in\mathcal{W}}(-\boldsymbol{y})'\boldsymbol{z}\right)^+}_{=\pi^2(y_0,\boldsymbol{y})}.$$

Note that whenever $y_0+\boldsymbol{y}'\boldsymbol{z} \geq 0, \forall \boldsymbol{z} \in \mathcal{W}$, or equivalently, $-y_0+\max_{\boldsymbol{z}\in\mathcal{W}}(-\boldsymbol{y})'\boldsymbol{z} \leq 0$, it is trivial to see that $\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) = y_0 = \pi^2(y_0, \boldsymbol{y})$. Therefore,

$$
\begin{aligned}
\psi_{1-\gamma}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) &\leq \min_\theta \left(\theta + \frac{\pi^2(y_0-\theta, \boldsymbol{y})}{\gamma}\right) \\
&= y_0 + \min_\theta \left(\theta + \frac{\pi^2(-\theta, \boldsymbol{y})}{\gamma}\right) \\
&= y_0 + \min_\theta \left\{\theta + \frac{1}{\gamma}\left(\left(\max_{\boldsymbol{z}\in\mathcal{W}}(-\boldsymbol{y})'\boldsymbol{z} + \theta\right)^+ - \theta\right)\right\} \\
&= y_0 + \min_\theta \left\{\theta(1-1/\gamma) + \frac{1}{\gamma}\left((\max_{\boldsymbol{z}\in\mathcal{W}}(-\boldsymbol{y})'\boldsymbol{z} + \theta)^+\right)\right\} \\
&= y_0 + (1/\gamma - 1)\min_\theta \left\{-\theta + \frac{1}{1-\gamma}\left((\max_{\boldsymbol{z}\in\mathcal{W}}(-\boldsymbol{y})'\boldsymbol{z} + \theta)^+\right)\right\} \\
&= y_0 + (1/\gamma - 1)\max_{\boldsymbol{z}\in\mathcal{W}} \boldsymbol{y}'(-\boldsymbol{z}) + (1/\gamma - 1)\min_\theta \left(-\theta + \frac{1}{1-\gamma}(\theta)^+\right) \\
&= y_0 + (1/\gamma - 1)\max_{\boldsymbol{z}\in\mathcal{W}} \boldsymbol{y}'(-\boldsymbol{z}) \\
&= \eta^2_{1-\gamma}(y_0, \boldsymbol{y}),
\end{aligned}
$$

(c) Using Jensen's inequality and the relation, $w^+ = (w + |w|)/2$, we have

$$
\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) = \frac{1}{2}(y_0 + +\mathrm{E}(|y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}|)) \leq \underbrace{\frac{1}{2}\left(y_0 + \sqrt{y_0^2 + \|\boldsymbol{\Sigma}\boldsymbol{y}\|_2^2}\right)}_{=\pi^2(y_0,\boldsymbol{y})}.
$$

Hence,

$$\begin{aligned}
\psi_{1-\gamma}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) &\leq \min_{\theta} \left(\theta + \frac{\pi^3(y_0 - \theta, \boldsymbol{y})}{\gamma}\right) \\
&= \min_{\theta} \left(\theta + \frac{y_0 - \theta + \sqrt{(y_0 - \theta)^2 + \boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}}}{2\gamma}\right) \\
&= y_0 + \sqrt{\frac{1-\gamma}{\gamma}}\sqrt{\boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}} \\
&= \eta^3_{1-\gamma}(y_0, \boldsymbol{y})
\end{aligned}$$

where the second equality follows from choosing the optimum θ,

$$\theta^* = y_0 + \frac{\sqrt{\boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}}(1-2\gamma)}{2\sqrt{\gamma(1-\gamma)}}.$$

(d) The bound is trivially true if there exists $y_j \neq 0$ for any $j > I$. Henceforth, we assume $y_j = 0, \forall j = I+1, \ldots, N$. The key idea of the inequality comes from the observation that

$$w^+ \leq \mu \exp(w/\mu - 1) \qquad \forall \mu > 0.$$

Since $\tilde{z}_j$, $j = 1, \ldots, I$ are stochastically independent, we have

$$\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) \leq \mu \mathrm{E}(\exp((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})/\mu - 1)) = \mu \exp(y_0/\mu - 1) \prod_{j=1}^{I} \mathrm{E}(\exp(y_j \tilde{z}_j/\mu)) \qquad \forall \mu > 0. \tag{.1}$$

This relation was first shown in Nemirovski and Shapiro [44]. Using the deviation measures of Chen, Sim and Sun [23], and Proposition 2(c), we have

$$\ln(\mathrm{E}(\exp(y_j\tilde{z}_j/\mu))) \leq \begin{cases} {y_j}^2{p_j}^2/(2\mu^2) & \text{if } y_j \geq 0 \\ {y_j}^2{q_j}^2/(2\mu^2) & \text{otherwise.} \end{cases} \quad (.2)$$

Since p_j and q_j are nonnegative, we have

$$\ln(\mathrm{E}(\exp(y_j\tilde{z}_j/\mu))) \leq \frac{(\max\{y_jp_j, -y_jq_j\})^2}{2\mu^2} = \frac{u_j^2}{2\mu^2}. \quad (.3)$$

Substituting this in the inequality (.1), we have

$$\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) \leq \inf_{\mu>0}\left\{\mu\exp(y_0/\mu - 1)\prod_{j=1}^{I}\mathrm{E}(\exp(y_j\tilde{z}_j/\mu))\right\} \leq \underbrace{\inf_{\mu>0}\left\{\frac{\mu}{e}\exp\left(\frac{y_0}{\mu} + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2}\right)\right\}}_{=\pi^4(y_0,\boldsymbol{y})}.$$

Hence,

$$\begin{aligned} \psi_{1-\gamma}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) &\leq \min_{\theta}\left(\theta + \frac{\pi^4(y_0-\theta, \boldsymbol{y})}{\gamma}\right) \\ &= \min_{\theta,\mu}\left(\theta + \frac{\frac{\mu}{e}\exp(\frac{y_0-\theta}{\mu} + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2})}{2\gamma}\right) \\ &= \min_{\mu}\left(y_0 + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2} - \mu\ln\gamma\right) \\ &= y_0 + \sqrt{-2\ln\gamma}\|\boldsymbol{u}\|_2 \\ &= \eta^4_{1-\gamma}(y_0, \boldsymbol{y}) \end{aligned}$$

where the second and third equalities follow from choosing the minimizers θ^* and μ^* as follows

$$\theta^* = y_0 + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2} - \mu \ln \gamma - \mu,$$

$$\mu^* = \frac{\|\boldsymbol{u}\|_2}{\sqrt{-2\ln\gamma}}.$$

(e) Again, we assume $y_j = 0, \forall j = I+1,\ldots,N$. Note that

$$\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) = y_0 + \mathrm{E}\left((-y_0 - \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) \leq \underbrace{y_0 + \inf_{\mu>0}\left\{\frac{\mu}{e}\exp\left(-\frac{y_0}{\mu} + \frac{\|\boldsymbol{v}\|_2^2}{2\mu^2}\right)\right\}}_{=\pi^5(y_0,\boldsymbol{y})}.$$

where $v_j = \max\{-p_j y_j, q_j y_j\}$, $j = 1,\ldots,I$. Hence, following from the above exposition, we have

$$\begin{aligned}
\psi_{1-\gamma}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) &\leq \min_\theta \left(\theta + \frac{\pi^5(y_0 - \theta, \boldsymbol{y})}{\gamma}\right) \\
&= \min_{\theta,\mu} \left(\theta + \frac{y_0 - \theta + \frac{\mu}{e}\exp(-\frac{y_0-\theta}{\mu} + \frac{\|\boldsymbol{v}\|_2^2}{2\mu^2})}{2\gamma}\right) \\
&= \min_\mu \left(y_0 + (\frac{1}{\gamma} - 1)(\frac{\|\boldsymbol{v}\|_2^2}{2\mu^2} - \mu\ln(1-\gamma))\right) \\
&= y_0 + \frac{1-\gamma}{\gamma}\sqrt{-2\ln(1-\gamma)}\|\boldsymbol{v}\|_2 \\
&= \eta^5_{1-\gamma}(y_0, \boldsymbol{y}).
\end{aligned}$$

■

.2 Approximation of a conic exponential quadratic constraint

Our aim to is show that the following conic exponential quadratic constraint,

$$\mu \exp\left(\frac{a}{\mu} + \frac{b^2}{\mu^2}\right) \leq c$$

for some $\mu > 0, a, b$ and c, can be approximately represented in the form of second order cones. Note with $\mu > 0$, the constraint

$$\mu \exp\left(\frac{a}{\mu} + \frac{b^2}{\mu^2}\right) \leq c$$

is equivalent to

$$\mu \exp\left(\frac{x}{\mu}\right) \leq c$$

for some variables x and d satisfying

$$b^2 \leq \mu d$$

$$a + d \leq x.$$

To approximate the conic exponential constraint, we use the method described in Ben-Tal and Nemirovski [10]. Using Taylor's series expansion, we have

$$\exp(x) = \exp\left(\frac{x}{2^L}\right)^{2^L} \approx \left(1 + \frac{x}{2^L} + \frac{1}{2}\left(\frac{x}{2^L}\right)^2 + \frac{1}{6}\left(\frac{x}{2^L}\right)^3 + \frac{1}{24}\left(\frac{x}{2^L}\right)^4\right)^{2^L},$$

where L is a positive integer. Observe that the approximation improves with larger values of L. Using the approximation, the following constraint

$$\mu\left(1 + \frac{x/\mu}{2^L} + \frac{1}{2}\left(\frac{x/\mu}{2^L}\right)^2 + \frac{1}{6}\left(\frac{x/\mu}{2^L}\right)^3 + \frac{1}{24}\left(\frac{x/\mu}{2^L}\right)^4\right)^{2^L} \leq c$$

is equivalent to

$$\mu\left(\frac{1}{24}\left(23 + 20\frac{x/\mu}{2^L} + 6\left(\frac{x/\mu}{2^L}\right)^2 + \left(1 + \frac{x/\mu}{2^L}\right)^4\right)\right)^{2^L} \leq c,$$

which is equivalent to the following set of constraints

$$
\begin{aligned}
& y = \tfrac{x}{2^L} \\
& z = \mu + \tfrac{x}{2^L} \\
& y^2 \leq \mu f, \quad z^2 \leq \mu g, \quad g^2 \leq \mu h \\
& \tfrac{1}{24}(23\mu + 20y + 6f + h) \leq v_1 \\
& v_i^2 \leq \mu v_{i+1} && \forall i = 1, \ldots, L-1 \\
& v_L^2 \leq \mu c
\end{aligned}
$$

for some variables $y, z \in \Re, f, g, h \in \Re_+, \boldsymbol{v} \in \Re_+^L$. Finally, using the well known result that

$$w^2 \leq st, \quad s, t \geq 0$$

is second order cone representable as

$$\left\| \begin{bmatrix} w \\ (s-t)/2 \end{bmatrix} \right\|_2 \leq \frac{s+t}{2},$$

we obtain an approximation of the conic exponential quadratic constraint that is second order cone representable.

To test the approximation, we plot in Figure .1, the exact and approximated values of the function $f(a)$ defined as follows:

$$f(a) = \inf_{\mu > 0} \mu \exp\left(\frac{a}{\mu} + \frac{1}{\mu^2}\right).$$

We obtain the exact solution by substituting $\mu^* = \frac{a+\sqrt{a^2+8}}{2}$ and the approximated solution by solving the SOCP approximation with $L = 4$. We solve the SOCP using CPLEX 9.1, with precision level of 10^{-7}. The relative errors for $a \geq -3$ is less than 10^{-7}. The approximation is poor when the actual value of $f(a)$ falls below the precision level, which is probably not a major concern in practice.

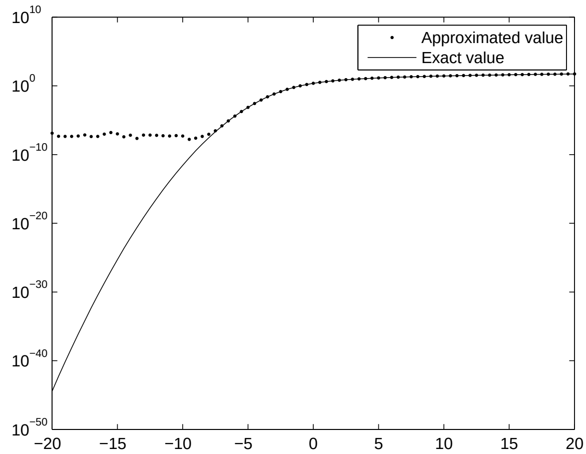


Fig. .1: Evaluation of approximation of $\inf_{\mu>0} \mu \exp\left(\frac{a}{\mu} + \frac{1}{\mu^2}\right)$.