Goal Driven Optimization

Wenqing Chen[*] Melvyn Sim[†]

Submitted: May 2006
Revised: December 2006

Abstract

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 under-performance or shortfall. The key advantage of the proposed model is its tractability. We show that proposed model is reduced to solving a small collections 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 respect to the accuracy, with each subproblem being a second order cone optimization problem (SOCP). We compare the numerical performance of the deterministic approximation with sampling approximation and report the computational insights on a multi-product Newsvendor problem.

[*]NUS Business School, National University of Singapore. Email: chenwenqing@gmail.com

[†]NUS Business School, National University of Singapore and Singapore MIT Alliance (SMA). Email: dscsimm@nus.edu.sg. The research of the author is supported by SMA, NUS academic research grants R-314-000-066-122 and R-314-000-068-122.

1 Introduction

Stochastic optimization is an adopted framework for modeling optimization problems that involve uncertainty. In a standard 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 [9]. However, optimizing an expectation 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 [9], Ahmed [1] and Ogryczak and Ruszczynski [29]. 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 [12] and Bordley and LiCalzi [8]. With the introduction of an *aspiration level* or the target 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, aspiration level plays an important role in daily decision making. Lanzillotti's study [23], 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. [30, 31] illustrated that managers tend to disregard investment possibilities that are likely to under perform against their target. Simon [38] also argued that most firms' goals are not maximizing profit but attaining a target profit. In an empirical study by Mao [26], 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".

Based on the motivations from decision analysis, 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 [13, 14] and Bereanu's [6] optimality criterion of maximizing the probability of getting a profit above a target 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 [36], Lau and Lau [22], Li et al. [24] and Parlar and Weng [32].

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", small probability events occur. However, studies have suggested that subjects are not completely insensitive to these losses; see for instance Payne et al [30]. Diecidue and van de Ven [17] 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 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 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 proposed model is reduced to solving a small collections of stochastic optimization problems with objectives evaluated under the Conditional-Value-at-Risk (CVaR) measure popularized by Rockafellar and Uryasev [34]. This class of stochastic optimization problems with mean risk objectives have recently been studied by Ahmed [1] and Riis and Schultz [33]. 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 [37]. Even in a two stage model, the number of sampled scenarios required to approximate the solution to reasonable accuracy can be astronomical 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 models that are heavily tuned to an assumed distribution may perform poorly in practice.

Motivated by recent development in robust optimization involving multiperiod decision process (see Ben-Tal et al. [3], Chen, Sim and Sun [15] and Chen et al. [16]), 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 and bounded supports, standard deviations and the *forward and backward deviations* introduced by Chen, Sim and Sun [15]. We adopt a comprehensive model of uncertainty that incorporates both models of Chen, Sim and Sun [15] and Chen et al. [16]. 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 [5]. This allows us to leverage on the state-of-the-art SOCP solvers, which are increasingly becoming more powerful, efficient and robust. Finally, we compare the performance of the deterministic approximation with a sampling approximation on a class of multi-product Newsvendor problem that maximizes the shortfall aspiration level criterion.

The structure of the paper is as follows. In Section 2, we introduce the goal driven model and propose the shortfall aspiration level criteria. We show that the goal driven optimization problem can be reduced to solving a sequence of stochastic optimization problems with CVaR objectives. Using techniques in robust optimization, we develop in Section 3, a deterministic approximation of the stochastic optimization problem with CVaR objective. In Section 4, we report some computational results and insights on a multi-product Newsvendor problem. Finally, Section 5 concludes this paper.

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

2 A Goal Driven Optimization Model

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

$$
\begin{array}{rll}
f(\boldsymbol{x},\boldsymbol{z}) = \boldsymbol{c}(\boldsymbol{z})'\boldsymbol{x} + & \displaystyle\min_{\boldsymbol{u},\boldsymbol{y}} & \boldsymbol{d_u}'\boldsymbol{u} + \boldsymbol{d_y}'\boldsymbol{y} \\
& \text{s.t.} & \boldsymbol{T}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u} + \boldsymbol{Y}\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{z}) \\
& & \boldsymbol{y} \geq \boldsymbol{0},
\end{array}
\tag{1}
$$

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, $\boldsymbol{c}(\tilde{\boldsymbol{z}}) \in \Re^{n_1}$, $\boldsymbol{T}(\tilde{\boldsymbol{z}}) \in \Re^{m_2 \times n_1}$ 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 [9]. The model (1) represents a rather general fixed recourse framework characterized in classical stochastic optimization formulations. Using the convention of stochastic optimization, if the model (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 [8]. 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:

$$
\begin{array}{ll}
\displaystyle\max_{\boldsymbol{x}} & \beta\Big(f(\boldsymbol{x},\tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})\Big) \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \geq \boldsymbol{0}
\end{array}
\tag{2}
$$

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 $\beta(\cdot)$ 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*

$$\beta(\tilde{v}) \triangleq \mathrm{P}(\tilde{v} \leq 0). \tag{3}$$

We adopt the same definition as used in Diecidue and van de Ven [17] and in Canada et al. [11], chapter 5. We can equivalently express the aspiration level criterion as

$$\beta(\tilde{v}) = 1 - \mathrm{P}(\tilde{v} > 0) = 1 - \mathrm{E}(\mathcal{H}(\tilde{v})) \tag{4}$$

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

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

2.1 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 (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:*

$$\begin{array}{c} \mathrm{E}(\tilde{v}) < 0 \\ \mathrm{P}(\tilde{v} > 0) > 0, \end{array} \tag{5}$$

the shortfall aspiration level criterion is defined as

$$\alpha(\tilde{v}) \triangleq 1 - \inf_{a>0} \left(\mathrm{E}\left(\mathcal{S}(\tilde{v}/a)\right)\right) \tag{6}$$

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 (5). The shortfall aspiration level criterion has the following properties*

(a)

$$\alpha(\tilde{v}) \leq \beta(\tilde{v})$$

(b)

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

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

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

(c)

$$\alpha(\tilde{v}) = \sup_{\gamma}\{1-\gamma \ : \ \psi_{1-\gamma}(\tilde{v}) \leq 0, \gamma \in (0,1)\}$$

where

$$\psi_{1-\gamma}(\tilde{v}) \triangleq \min_{\theta} \left(\theta + \frac{\mathrm{E}\left((\tilde{v}-\theta)^+\right)}{\gamma}\right) \tag{7}$$

is the risk measure known as Conditional-Value-at-Risk (CVaR) popularized by Rockafellar and Uryasev [34].

(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 $a > 0$, $\mathcal{S}(x/a) \geq \mathcal{H}(x)$, hence, we have

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

Therefore,

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

(b) Since $\mathrm{P}(\tilde{v} > 0) > 0$, from (a), we have $\alpha(\tilde{v}) \leq 1 - \mathrm{P}(\tilde{v} > 0) < 1$. To show that $\alpha(\tilde{v}) > 0$, 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,

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

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

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

(d) Observe that

$$\max_{\boldsymbol{x}} \left\{ \beta\Big(\tilde{v}(\boldsymbol{x})\Big) \ : \ \boldsymbol{x} \in \mathcal{X} \right\} \tag{8}$$

is equivalent to

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

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

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

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 (8) corresponds to maximizing the following ratio:

$$\begin{array}{ll} \max & \dfrac{-\mu(\boldsymbol{x})}{\sigma(\boldsymbol{x})} \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X}. \end{array} \tag{9}$$

This relation was observed by Dragomirescu [18]. Using the result in (c), we can express the maximization of the shortfall aspiration level criterion as follows:

$$\begin{array}{ll} \max & 1-\gamma \\ \text{s.t.} & \psi_{1-\gamma}(\tilde{v}(\boldsymbol{x})) \leq 0 \\ & \boldsymbol{x} \in \mathcal{X}, \gamma \in (0,1) \end{array} \tag{10}$$

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

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

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 (10) is identical to Model (9). ■

We now propose the following goal driven optimization problem.

$$\begin{array}{ll} \max\limits_{\boldsymbol{x}} & \alpha\Big(f(\boldsymbol{x},\tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}})\Big) \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array} \tag{11}$$

Theorem 1(a) implies that an optimal solution of Model (11), $\boldsymbol{x}^*$ can achieve the following success probability,

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

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 (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,

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

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 (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 (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. [2]. 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 [35]. Hence, the function $\psi_{1-\gamma}(f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) - \tau(\tilde{\boldsymbol{z}}))$ is convex in $\boldsymbol{x}$. Using the connection with the CVaR measure, we express the goal driven optimization model (11), equivalently as follows:

$$
\begin{aligned}
\max_{\gamma,\boldsymbol{x}} \quad & 1-\gamma \\
\text{s.t.} \quad & \psi_{1-\gamma}(f(\boldsymbol{x},\tilde{\boldsymbol{z}})-\tau(\tilde{\boldsymbol{z}}))\leq 0 \\
& \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \\
& \boldsymbol{x}\geq \boldsymbol{0} \\
& \gamma\in(0,1).
\end{aligned}
\tag{12}
$$

2.2 Reduction to stochastic optimization problems with CVaR objectives

For a fixed γ, the first constraint in Model (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:

$$
\begin{aligned}
Z(\gamma)= \min_{\boldsymbol{x}} \quad & \psi_{1-\gamma}\Big(f(\boldsymbol{x},\tilde{\boldsymbol{z}})-\tau(\tilde{\boldsymbol{z}})\Big) \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \\
& \boldsymbol{x}\geq \boldsymbol{0},
\end{aligned}
\tag{13}
$$

or equivalently,

$$
\begin{aligned}
Z(\gamma)= \min_{\boldsymbol{x},\boldsymbol{u}(\cdot),\boldsymbol{y}(\cdot)} \quad & \psi_{1-\gamma}\Big(\boldsymbol{c}(\tilde{\boldsymbol{z}})'\boldsymbol{x}+\boldsymbol{d_u}'\boldsymbol{u}(\tilde{\boldsymbol{z}})+\boldsymbol{d_y}'\boldsymbol{y}(\tilde{\boldsymbol{z}})-\tau(\tilde{\boldsymbol{z}})\Big) \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \\
& \boldsymbol{T}(\tilde{\boldsymbol{z}})\boldsymbol{x}+\boldsymbol{U}\boldsymbol{u}(\tilde{\boldsymbol{z}})+\boldsymbol{Y}\boldsymbol{y}(\tilde{\boldsymbol{z}})=\boldsymbol{h}(\tilde{\boldsymbol{z}}) \\
& \boldsymbol{y}(\tilde{\boldsymbol{z}})\geq \boldsymbol{0} \\
& \boldsymbol{x}\geq \boldsymbol{0}
\end{aligned}
\tag{14}
$$

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 (13) 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 (13) 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 (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 (13), where* $1-\gamma^*$ *being the optimal objective of Model (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 $Z(\gamma) \leq 0$, γ is feasible in Model (12), hence, $\gamma^* \leq \gamma$. Otherwise, γ would be infeasible in Model (12). In this case, we claim that the optimal feasible solution, γ^* must be greater than γ. Suppose not, we have $\gamma^* \leq \gamma$. We know the optimal solution $\boldsymbol{x}^*$ of Model (12) satisfies

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

However, since $\gamma^* \leq \gamma$, we have

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

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 (13) as a linear optimization problem as follows:

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

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 [33] 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. [3]. Moreover, with the assumption that the stochastic parameters are independently distributed, Dyer and Stougie [19] show that two-stage stochastic programming problems are #P-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 via Robust Optimization

We have shown that solving the goal driven optimization model (11) involves solving a sequence of stochastic optimization problems with CVaR objectives in the form of Model (14). Hence, we devote this section to formulating a tractable deterministic approximation of Model (14).

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. [3], Bertsimas and Sim [7], Chen, Sim and Sun [15] and Chen et al. [16], 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 $\boldsymbol{c}(\tilde{\boldsymbol{z}})$, $\boldsymbol{T}(\tilde{\boldsymbol{z}})$, $\boldsymbol{h}(\tilde{\boldsymbol{z}})$ and $\tau(\tilde{\boldsymbol{z}})$ are affinely dependent on the primitive uncertainties $\tilde{\boldsymbol{z}}$ as follows:

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

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 [15].

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

$$\sigma_f(\tilde{z}) \triangleq \sup_{\theta>0} \left\{ \sqrt{2\ln(\mathrm{E}(\exp(\theta\tilde{z})))/\theta^2} \right\} \tag{15}$$

and backward deviation is defined as

$$\sigma_b(\tilde{z}) \triangleq \sup_{\theta>0} \left\{ \sqrt{2\ln(\mathrm{E}(\exp(-\theta\tilde{z})))/\theta^2} \right\}. \tag{16}$$

Given a sequence of independent samples, we can essentially estimate the magnitude of the deviation measures from (15) and (16). Some of the properties of the deviation measures include:

Proposition 2 *(Chen, Sim and Sun [15])*
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)

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

(c) *For all $\theta \geq 0$,*

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

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 ([15]) 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 [15] show how to obtain upper bounds of the forward and backward deviation measures.

Theorem 2 *(Chen, Sim and Sun [15]) 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. [27] 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.

Model of Uncertainty, U: 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 2. We denote $\boldsymbol{P} = \text{diag}(p_1, \ldots, p_I)$ and $\boldsymbol{Q} = \text{diag}(q_1, \ldots, q_I)$.

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 [4], Bertsimas and Sim [7] and Goldfarb and Iyengar [21]. For simplicity, we assume in this paper that the exact parameters are given.

Similar uncertainty models have been defined in Chen, Sim and Sun [15] and Chen et al. [16]. 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, the Model of Uncertainty, U encompasses both models discussed in Chen, Sim and Sun [15] and Chen et al. [16].

Under the Model of Uncertainty, U, 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, $\tilde{z}_j$. The Model of Uncertainty, U, 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 the Model of Uncertainty, U, 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).

3.1 Approximations of $\mathrm{E}((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+)$ and $\psi_{1-\gamma}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})$

Although the CVaR measure,

$$\psi_{1-\gamma}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) = \min_{\theta} \left(\theta + \frac{\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} - \theta)^+\right)}{\gamma}\right)$$

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 $\psi_{1-\gamma}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})$ as follows:

Theorem 3 *Assuming $\tilde{\boldsymbol{z}}$ follows the Model of Uncertainty, U, 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,*

$$\eta^i_{1-\gamma}(y_0, \boldsymbol{y}) \triangleq \min_{\theta} \left(\theta + \frac{1}{\gamma}\pi^i(y_0 - \theta, \boldsymbol{y})\right) \qquad i \in \{1, \ldots, 5\}$$

are the upper bounds of $\psi_{1-\gamma}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})$).

(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)

$$\begin{aligned}
\pi^4(y_0, \boldsymbol{y}) &\triangleq \begin{cases} \inf_{\mu > 0} \left\{\frac{\mu}{e} \exp\left(\frac{y_0}{\mu} + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2}\right)\right\} & \textit{if } y_j = 0 \; \forall j = I+1, \ldots, N \\ +\infty & \textit{otherwise} \end{cases} \\
\eta^4_{1-\gamma}(y_0, \boldsymbol{y}) &\triangleq \begin{cases} y_0 + \sqrt{-2\ln\gamma}\|\boldsymbol{u}\|_2 & \textit{if } y_j = 0 \; \forall j = I+1, \ldots, N \\ +\infty & \textit{otherwise} \end{cases},
\end{aligned}$$

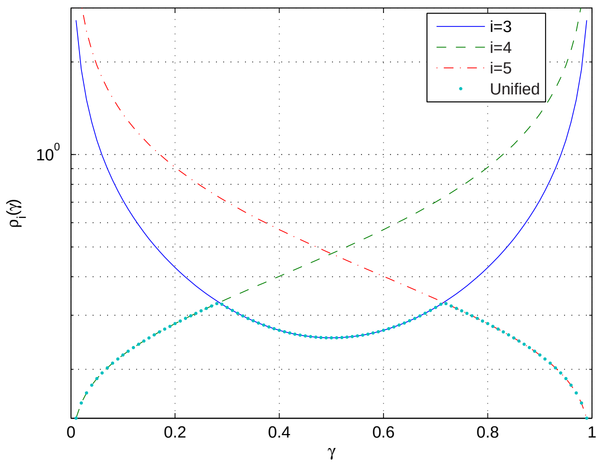

Figure 1: Plot of $\rho_i(\gamma)$ against γ for $i = 3, 4$ and 5, defined in Proposition 3.

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

(e)

$$\pi^5(y_0, \boldsymbol{y}) \triangleq \begin{cases} y_0 + \inf_{\mu>0}\left\{\frac{\mu}{e}\exp\left(-\frac{y_0}{\mu} + \frac{\|\boldsymbol{v}\|_2^2}{2\mu^2}\right)\right\} & \textit{if } y_j = 0\ \forall j = I+1, \ldots, N \\ +\infty & \textit{otherwise} \end{cases}$$

$$\eta^5_{1-\gamma}(y_0, \boldsymbol{y}) \triangleq \begin{cases} y_0 + \frac{1-\gamma}{\gamma}\sqrt{-2\ln(1-\gamma)}\|\boldsymbol{v}\|_2 & \textit{if } y_j = 0\ \forall j = I+1, \ldots, N \\ +\infty & \textit{otherwise} \end{cases},$$

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

The proof is shown in Appendix A.

Remark : The first and second bounds in Proposition 3 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 $\psi_{1-\gamma}(\tilde{z})$, where $\tilde{z}$ is standard normally distributed. Figure 1 compares the approximation ratios given by

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

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 1 achieves a worst case approximation error of 33% at $\gamma = 0.2847$ and $\gamma = 0.7153$. We next show how to unify these bounds.

Theorem 4 **(a)** *Let*

$$\begin{array}{rll}\pi(y_0, \boldsymbol{y}) \triangleq & \min_{y_{i0}, \boldsymbol{y}_i} & \sum_{i=1}^{5} \pi^i(y_{i0}, \boldsymbol{y}_i) \\ & \text{s.t.} & \sum_{i=1}^{5} y_{i0} = y_0 \\ & & \sum_{i=1}^{5} \boldsymbol{y}_i = \boldsymbol{y}. \end{array}$$

Then for all $(y_0, \boldsymbol{y})$

$$\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) \leq \pi(y_0, \boldsymbol{y}) \leq \min_{i \in \{1,\dots,5\}} \{\pi^i(y_0, \boldsymbol{y})\} \qquad (17)$$

(b) *Let*

$$\eta_{1-\gamma}(y_0, \boldsymbol{y}) \triangleq \min_{\theta} \left(\theta + \frac{1}{\gamma}\pi(y_0 - \theta, \boldsymbol{y})\right)$$

or equivalently

$$\begin{array}{rll}\eta_{1-\gamma}(y_0, \boldsymbol{y}) \triangleq & \min_{y_{i0}, \boldsymbol{y}_i} & \sum_{i=1}^{5} \eta^i_{1-\gamma}(y_{i0}, \boldsymbol{y}_i) \\ & \text{s.t.} & \sum_{i=1}^{5} y_{i0} = y_0 \\ & & \sum_{i=1}^{5} \boldsymbol{y}_i = \boldsymbol{y}. \end{array}$$

Then for all $(y_0, \boldsymbol{y})$ *and* $\gamma \in (0, 1)$

$$\psi_{1-\gamma}(y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}) \leq \eta_{1-\gamma}(y_0, \boldsymbol{y}) \leq \min_{i \in \{1,\dots,5\}} \{\eta^i_{1-\gamma}(y_0, \boldsymbol{y})\} \qquad (18)$$

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

$$\begin{array}{lll} & \sum_{i=1}^{5} \pi^i(y_{i0}, \boldsymbol{y}_i) & \\ \geq & \sum_{i=1}^{5} \mathrm{E}\left((y_{i0} + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^+\right) & \text{Proposition 3} \\ \geq & \mathrm{E}\left(\left(\sum_{i=1}^{5}(y_{i0} + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})\right)^+\right) & \text{Subadditvity} \\ = & \mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right). & \end{array}$$

Finally, to show that $\pi(y_0, \boldsymbol{y}) \leq \pi^i(y_0, \boldsymbol{y})$, $i = 1, \dots, 5$, let

$$(y_{r0}, \boldsymbol{y}_r) = \begin{cases} (y_0, \boldsymbol{y}) & \text{if } r = i \\ (0, \mathbf{0}) & \text{otherwise} \end{cases} \qquad \text{for } r = 1, \dots, 5.$$

Hence,

$$\pi^r(y_{r0}, \boldsymbol{y}_r) = \begin{cases} \pi^r(y_0, \boldsymbol{y}) & \text{if } r = i \\ 0 & \text{otherwise} \end{cases} \qquad \text{for } r = 1, \dots, 5,$$

and therefore

$$\pi(y_0, \boldsymbol{y}) \le \sum_{i=1}^{5} \pi^i(y_{i0}, \boldsymbol{y}_i) = \pi^i(y_0, \boldsymbol{y}).$$

(b) Observe that

$$\begin{aligned}
\eta_{1-\gamma}(y_0, \boldsymbol{y}) = & \min_{\theta} \left(\theta + \frac{\pi(y_0 - \theta, \boldsymbol{y})}{\gamma} \right) \\
= & \min_{\theta, \boldsymbol{\theta}, y_{i0}, \boldsymbol{y}_i, \forall i} \left(\theta + \sum_{i=1}^{5} \left(\frac{\pi^i(y_{i0} - \theta_i, \boldsymbol{y}_i)}{\gamma} \right) \mid \sum_{i=1}^{5} \boldsymbol{y}_i = \boldsymbol{y}, \sum_{i=1}^{5} y_{i0} = y_0, \sum_{i=1}^{5} \theta_i = \theta \right) \\
= & \min_{y_{i0}, \boldsymbol{y}_i, \forall i} \left(\sum_{i=1}^{5} \underbrace{\min_{\theta_i} \left(\theta_i + \frac{\pi^i(y_{i0} - \theta_i, \boldsymbol{y}_i)}{\gamma} \right)}_{= \eta^i_{1-\gamma}(y_{i0}, \boldsymbol{y}_i)} \mid \sum_{i=1}^{5} \boldsymbol{y}_i = \boldsymbol{y}, \sum_{i=1}^{5} y_{i0} = y_0 \right).
\end{aligned}$$

Finally, the inequalities (18) are trivial consequence of the inequalities (17). ∎

Remark : Note that in the presence of stochastically dependent primitive uncertainties and unbounded support, all the bounds, except for the third, of Theorem 3 can become infinite. However, such trivial bound is avoided in the unified bound.

From Theorem 3(a), the epigraph of the unified bound of $\mathrm{E}((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+)$, $\pi(y_0, \boldsymbol{y}) \le s$ can be expressed as follows:

$$\begin{array}{ll}
\exists r_i, y_{i0} \in \Re, \boldsymbol{y}_i, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{d}, \boldsymbol{h} \in \Re^N, i = 1, \ldots, 5, \boldsymbol{u}, \boldsymbol{v} \in \Re^I, \text{ such that} & \\
r_1 + r_2 + r_3 + r_4 + r_5 \le s & \\
y_{10} + \boldsymbol{s}'\bar{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}} \le r_1 & \\
0 \le r_1 & \\
\boldsymbol{s} - \boldsymbol{t} = \boldsymbol{y}_1 & \\
\boldsymbol{s}, \boldsymbol{t} \ge 0 & \\
\boldsymbol{d}'\bar{\boldsymbol{z}} + \boldsymbol{h}'\underline{\boldsymbol{z}} \le r_2 & \\
y_{20} \le r_2 & \\
\boldsymbol{d} - \boldsymbol{h} = -\boldsymbol{y}_2 & \\
\boldsymbol{d}, \boldsymbol{h} \ge 0 & \\
\frac{1}{2} y_{30} + \frac{1}{2} \|(y_{30}, \boldsymbol{\Sigma}^{1/2} \boldsymbol{y}_3)\|_2 \le r_3 & \\
\inf_{\mu > 0} \frac{\mu}{e} \exp\left(\frac{y_0}{\mu} + \frac{\|\boldsymbol{u}\|_2^2}{2\mu^2} \right) \le r_4 & \\
u_j \ge p_j y_{4j}, u_j \ge -q_j y_{4j} & \forall j = 1, \ldots, I \\
y_{4j} = 0 & \forall j = I+1, \ldots, N \\
y_0 + \inf_{\mu > 0} \frac{\mu}{e} \exp\left(-\frac{y_0}{\mu} + \frac{\|\boldsymbol{v}\|_2^2}{2\mu^2} \right) \le r_5 & \\
v_j \ge q_j y_{5j}, v_j \ge -p_j y_{5j} & \forall j = 1, \ldots, I \\
y_{5j} = 0 & \forall j = I+1, \ldots, N \\
y_{10} + y_{20} + y_{30} + y_{40} + y_{50} = y_0 & \\
\boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 + \boldsymbol{y}_4 + \boldsymbol{y}_5 = \boldsymbol{y}. &
\end{array} \tag{19}$$

Due to the presence of the constraint, $\inf_{\mu>0} \mu \exp\left(\frac{a}{\mu} + \frac{b^2}{\mu^2}\right) \le c$, the set of constraints in (19) is not exactly second order cone representable (see Ben-Tal and Nemirovski [5]). 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 B.

Similarly, from Theorem 3(b), the epigraph of the unified CVaR approximation, $\eta_{1-\gamma}(y_0, \boldsymbol{y}) \leq s$ is second order cone representable as follows:

$$
\begin{array}{ll}
\exists r_i, y_{i0} \in \Re, \boldsymbol{y}_i, \boldsymbol{s}, \boldsymbol{t}, \boldsymbol{d}, \boldsymbol{h} \in \Re^N, i = 1, \ldots, 5, \boldsymbol{u}, \boldsymbol{v} \in \Re^I \text{ such that} & \\
r_1 + r_2 + r_3 + r_4 + r_5 \leq s & \\
y_{10} + \boldsymbol{s}'\bar{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}} \leq r_1 & \\
\boldsymbol{s}, \boldsymbol{t} \geq 0 & \\
\boldsymbol{s} - \boldsymbol{t} = \boldsymbol{y}_1 & \\
y_{20} + (1/\gamma - 1)\boldsymbol{d}'\bar{\boldsymbol{z}} + (1/\gamma - 1)\boldsymbol{h}'\underline{\boldsymbol{z}} \leq r_2 & \\
\boldsymbol{d} - \boldsymbol{h} = -\boldsymbol{y}_2 & \\
\boldsymbol{d}, \boldsymbol{h} \geq 0 & \\
y_{30} + \sqrt{\frac{1-\gamma}{\gamma}} \|\boldsymbol{\Sigma}^{1/2}\boldsymbol{y}_3\|_2 \leq r_3 & \\
y_{40} + \sqrt{-2\ln(\gamma)}\|\boldsymbol{u}\|_2 \leq r_4 & \\
u_j \geq p_j y_{4j}, u_j \geq -q_j y_{4j} & \forall j = 1, \ldots, I \\
y_{4j} = 0 & \forall j = I+1, \ldots, N \\
y_{50} + \frac{1-\gamma}{\gamma}\sqrt{-2\ln(1-\gamma)}\|\boldsymbol{v}\|_2 \leq r_5 & \\
v_j \geq q_j y_{5j}, v_j \geq -p_j y_{5j} & \forall j = 1, \ldots, I \\
y_{5j} = 0 & \forall j = I+1, \ldots, N \\
y_{10} + y_{20} + y_{30} + y_{40} + y_{50} = y_0 & \\
\boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 + \boldsymbol{y}_4 + \boldsymbol{y}_5 = \boldsymbol{y}. &
\end{array}
$$

It is rather surprising to note that while the epigraph of the function $\pi(\cdot, \cdot)$ is approximately second-order cone representable, the epigraph of $\eta(\cdot, \cdot)$, is fully second-order cone representable.

3.2 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 (14), we first adopt the linear decision rule used in Ben-Tal et al. [3] and Chen, Sim, and Sun [15]. We restrict $\boldsymbol{u}(\tilde{\boldsymbol{z}})$ and $\boldsymbol{y}(\tilde{\boldsymbol{z}})$ to be affinely dependent on the primitive uncertainties, that is

$$\boldsymbol{u}(\tilde{\boldsymbol{z}}) = \boldsymbol{u}^0 + \textstyle\sum_{j=1}^N \boldsymbol{u}^j \tilde{z}_j, \ \boldsymbol{y}(\tilde{\boldsymbol{z}}) = \boldsymbol{y}^0 + \textstyle\sum_{j=1}^N \boldsymbol{y}^j \tilde{z}_j. \tag{20}$$

Under linear decision rule, the following constraint

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

is a sufficient condition to satisfy the affine constraint involving recourse variables in Model (14). 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 (14) 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

$$w(\tilde{\boldsymbol{z}}) = w_0 + \sum_{j=1}^{N} w_j \tilde{z}_j, \tag{21}$$

where

$$w_j = \boldsymbol{c}^{j\prime}\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}^j + \boldsymbol{d_y}'\boldsymbol{y}^j - \tau^j \qquad j = 0, \ldots, N, \tag{22}$$

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

$$\psi_{1-\gamma}(w(\tilde{\boldsymbol{z}})) \leq \eta_{1-\gamma}(w_0, \boldsymbol{w})$$

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.

$$\begin{array}{rlll} Z_{LDR}(\gamma) = & \min\limits_{\boldsymbol{x},\boldsymbol{u}^j,\boldsymbol{y}^j,w_0,\boldsymbol{w}} & \eta_{1-\gamma}(w_0, \boldsymbol{w}) & \\ & \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\ & & w_j = \boldsymbol{c}^{j\prime}\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}^j + \boldsymbol{d_y}'\boldsymbol{y}^j - \tau^j & j = 0, \ldots, N. \\ & & \boldsymbol{T}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}^j + \boldsymbol{Y}\boldsymbol{y}^j = \boldsymbol{h}^j & j = 0, \ldots, N. \\ & & y_i^0 + \sum_{j=1}^{N} y_i^j z_j \geq \boldsymbol{0} & \forall \boldsymbol{z} \in \mathcal{W}, i = 1, \ldots, n_3 \\ & & \boldsymbol{x} \geq 0. & \end{array} \tag{23}$$

Theorem 5 *Let* $(\boldsymbol{x}, \boldsymbol{u}^0, \ldots, \boldsymbol{u}^N, \boldsymbol{y}^0, \ldots, \boldsymbol{y}^N)$ *be an optimal solution of Model (23). The solution* $\boldsymbol{x}$ *and linear decision rules* $\boldsymbol{u}(\tilde{\boldsymbol{z}})$ *and* $\boldsymbol{y}(\tilde{\boldsymbol{z}})$ *defined in the equations (20), are feasible in the subproblem (14). 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 (14) 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. [16] 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. [16] as an improvement over linear decision rules. We first define the vector $\bar{\boldsymbol{d}}$ with elements

$$\begin{array}{rll} \bar{d}_i = & \min_{\boldsymbol{u},\boldsymbol{y}} & \boldsymbol{d_u}'\boldsymbol{u} + \boldsymbol{d_y}'\boldsymbol{y} \\ & \text{s.t.} & \boldsymbol{U}\boldsymbol{u} + \boldsymbol{Y}\boldsymbol{y} = \boldsymbol{0} \\ & & y_i = 1 \\ & & \boldsymbol{y} \geq \boldsymbol{0}, \end{array} \tag{24}$$

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 $\bar{\boldsymbol{d}} \geq \boldsymbol{0}$.

Next, we present the model that achieves a better bound than Model (23).

$$\begin{array}{rlll} Z_{DLDR}(\gamma) = & \min\limits_{\boldsymbol{x},\boldsymbol{u}^j,\boldsymbol{y}^j,w_0,\boldsymbol{w}} & \eta_{1-\gamma}(w_0, \boldsymbol{w}) + \dfrac{1}{\gamma}\sum\limits_{i\in\mathcal{C}} \pi(-y_i^0, -\boldsymbol{y}_i)\bar{d}_i & \\ & \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\ & & w_j = \boldsymbol{c}^{j\prime}\boldsymbol{x} + \boldsymbol{d_u}'\boldsymbol{u}^j + \boldsymbol{d_y}'\boldsymbol{y}^j - \tau^j & j = 0, \ldots, N. \\ & & \boldsymbol{T}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{u}^j + \boldsymbol{Y}\boldsymbol{y}^j = \boldsymbol{h}^j & j = 0, \ldots, N. \\ & & y_i^0 + \sum_{j=1}^N y_i^j z_j \geq 0 & \forall \boldsymbol{z} \in \mathcal{W}, i \in \bar{\mathcal{C}} \\ & & \boldsymbol{x} \geq \boldsymbol{0}, & \end{array} \tag{25}$$

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

Theorem 6 *Let $(\boldsymbol{x}, \boldsymbol{u}^0, \ldots, \boldsymbol{u}^N, \boldsymbol{y}^0, \ldots, \boldsymbol{y}^N)$ be an optimal solution of Model (25). The solution $\boldsymbol{x}$ and the corresponding* deflected linear decision rule

$$\begin{array}{rcl} \boldsymbol{u}(\tilde{\boldsymbol{z}}) & = & \boldsymbol{u}^0 + \sum\limits_{j=1}^N \boldsymbol{u}^j \tilde{z}_j + \sum\limits_{i\in\mathcal{C}} \bar{\boldsymbol{u}}^i (y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^- \\ \boldsymbol{y}(\tilde{\boldsymbol{z}}) & = & \boldsymbol{y}^0 + \sum\limits_{j=1}^N \boldsymbol{y}^j \tilde{z}_j + \sum\limits_{i\in\mathcal{C}} \bar{\boldsymbol{y}}^i (y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^-, \end{array} \tag{26}$$

are feasible in the subproblem (14). 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 (14). 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 (21) and (22). We now evaluate the objective of Model (14) under the deflected linear decision rule as follows:

$$\begin{aligned}
& \psi_{1-\gamma}\Big(w(\tilde{\boldsymbol{z}}) + \sum_{i\in\mathcal{C}} \bar{d}_i(y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^-\Big) \\
= \quad & \min_\theta \left\{\theta + \frac{1}{\gamma}\mathrm{E}\left(\Big(w(\tilde{\boldsymbol{z}}) + \sum_{i\in\mathcal{C}} \bar{d}_i(y_i^0 + \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^- - \theta\Big)^+\right)\right\} \\
= \quad & \min_\theta \left\{\theta + \frac{1}{\gamma}\mathrm{E}\left(\left(w(\tilde{\boldsymbol{z}}) + \sum_{i\in\mathcal{C}} \bar{d}_i\left((-y_i^0 - \boldsymbol{y}_i'\tilde{\boldsymbol{z}})^+\right) - \theta\right)^+\right)\right\} \\
\leq \quad & \min_\theta \left\{\theta + \frac{1}{\gamma}\mathrm{E}\left(\Big(w(\tilde{\boldsymbol{z}}) - \theta\Big)^+\right) + \sum_{i\in\mathcal{C}} \frac{1}{\gamma}\mathrm{E}\left(\Big(-y_i^0 - \boldsymbol{y}_i'\tilde{\boldsymbol{z}}\Big)^+\right)\bar{d}_i\right\} \\
= \quad & \psi_{1-\gamma}\left(w(\tilde{\boldsymbol{z}})\right) + \frac{1}{\gamma}\sum_{i\in\mathcal{C}} \mathrm{E}\left(\Big(-y_i^0 - \boldsymbol{y}_i'\tilde{\boldsymbol{z}}\Big)^+\right)\bar{d}_i \\
\leq \quad & \eta_{1-\gamma}(w_0, \boldsymbol{w}) + \frac{1}{\gamma}\sum_{i\in\mathcal{C}} \pi(-y_i^0, -\boldsymbol{y}_i)\bar{d}_i \\
= \quad & Z_{DLDR}(\gamma),
\end{aligned} \tag{27}$$

where the first inequality are due to $(x+a)^+ \leq (x)^+ + a$, for all $a \geq 0$, and that $\bar{\boldsymbol{d}} \geq \boldsymbol{0}$. The last inequality is due to Theorems 4.

To prove the improvement over Model (23), we now consider an optimal solution of Model (23), $(\boldsymbol{x}, \boldsymbol{u}^0, \ldots, \boldsymbol{u}^N, \boldsymbol{y}^0, \ldots, \boldsymbol{y}^N)$. Clearly, the solution is feasible in the constraints of Model (25). From Theorems 3(a) and 4, the constraint $y_i^0 + \sum_{j=1}^N y_i^j z_j \geq 0$, $\forall \boldsymbol{z} \in \mathcal{W}$ enforced in Model (23) ensures that

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

for all $i \in \mathcal{C}$. Therefore, the solution of Model (23) yields the same objective as Model (25). Hence, $Z_{DLDR}(\gamma) \leq Z_{LDR}(\gamma)$. ■

Remark : Chen et al. [16] 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.

4 Computation Studies: 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 [20] and was extended by Ben-Daya and Raouf [10] and Lau and Lau [25]. 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,

$$\begin{array}{ll} \max\limits_{\boldsymbol{x}} & \sum\limits_{i=1}^{m} \left\{ (p_i - c_i)x_i - (p_i - s_i)\mathrm{E}\left((x_i - \tilde{h}_i)^+ \right) \right\} \\ \text{s.t.} & \boldsymbol{x} \geq \boldsymbol{0}, \end{array} \tag{28}$$

where the terms are defined as follows:

$$\begin{array}{ll} c_i & : \text{unit purchasing cost} \\ p_i & : \text{unit selling price} \\ s_i & : \text{unit salvage value} \\ \tilde{h}_i & : \text{stochastic demand} \\ x_i & : \text{order quantity,} \end{array}$$

with $p_i > c_i > s_i$ for all products. Note that regardless of the dependency of products' demands, we can easily decompose Model (28) into m independent Newsvendor problems. Hence, we can analytically obtain the optimal solution of Model (28). Note that the formulation of Model (28) 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

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

However, not all decision makers are comfortable with implementing the risk neutral solution. Given a target profit, τ, Sankarasubramanian and Kumaraswamy [36] proposed a single-product model that maximizes the probability of attaining the target. Likewise, Lau and Lau [22] and Li et al. [24] 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:

$$
\begin{array}{rll}
\max\limits_{\gamma,\boldsymbol{x},\boldsymbol{y}(\cdot)} & 1-\gamma & \\
\text{s.t.} & \psi_{1-\gamma}\Big(\tau - (\boldsymbol{p}-\boldsymbol{c})'\boldsymbol{x} + \sum\limits_{i=1}^{m} y_i(\tilde{\boldsymbol{h}})\Big) \leq 0 & \\
& y_i(\tilde{\boldsymbol{h}}) - y_{m+i}(\tilde{\boldsymbol{h}}) = (p_i - s_i)(x_i - \tilde{h}_i) & i = 1,\ldots,m \\
& y_i(\tilde{\boldsymbol{h}}) \geq 0 & i = 1,\ldots,2m \\
& \boldsymbol{x} \geq \boldsymbol{0}. &
\end{array} \tag{29}
$$

Using Algorithm 1, we reduce the problem (29) to solving a sequence of subproblems in the form of stochastic optimization problems with CVaR objectives as follows:

$$
\begin{array}{rll}
Z(\gamma) = \min\limits_{\boldsymbol{x},\boldsymbol{y}(\cdot)} & \psi_{1-\gamma}\Big(\tau - (\boldsymbol{p}-\boldsymbol{c})'\boldsymbol{x} + \sum\limits_{i=1}^{m} y_i(\tilde{\boldsymbol{h}})\Big) & \\
\text{s.t.} & y_i(\tilde{\boldsymbol{h}}) - y_{m+i}(\tilde{\boldsymbol{h}}) = (p_i - s_i)(x_i - \tilde{h}_i) & i = 1,\ldots,m \\
& y_i(\tilde{\boldsymbol{h}}) \geq 0 & i = 1,\ldots,2m \\
& \boldsymbol{x} \geq \boldsymbol{0}. &
\end{array} \tag{30}
$$

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, we need to obtain $\bar{\boldsymbol{d}} \in \Re^{2m}$ as follows

$$
\begin{array}{rll}
\bar{d}_i = \min\limits_{\boldsymbol{y}} & \sum\limits_{j=1}^{m} y_j & \\
\text{s.t.} & y_j - y_{m+j} = 0 & j = 1,\ldots,m \\
& y_i = 1 & \\
& y_j \geq 0 & i = 1,\ldots,2m.
\end{array}
$$

Clearly, $\bar{d}_i = 1$ for all $i = 1,\ldots,2m$. Hence, using deflected linear decision rule, we can obtain an upper bound of the subproblems (30) by solving the following problem:

$$
\begin{aligned}
Z_{DLDR}(\gamma) = \min_{\boldsymbol{x},w_0,\boldsymbol{w},\boldsymbol{y}^j} \quad & \eta_{1-\gamma}(w_0,\boldsymbol{w}) + \frac{1}{\gamma}\sum_{i=1}^{2m}\pi(-y_i^0,-\boldsymbol{y}_i) \\
\text{s.t.} \quad & y_i^0 - y_{m+i}^0 = (p_i - s_i)(x_i - h_i^0) && i = 1,\dots,m \\
& y_i^j - y_{m+i}^0 = (p_i - s_i)(-h_i^j) && i = 1,\dots,m, j = 1,\dots,m \\
& w_0 = \tau - (\boldsymbol{p}-\boldsymbol{c})'\boldsymbol{x} + \sum_{i=1}^{m} y_i^0 \\
& w_j = \sum_{i=1}^{m} y_i^j && j = 1,\dots,m \\
& \boldsymbol{x} \geq \boldsymbol{0},
\end{aligned}
\tag{31}
$$

where

$$
\begin{aligned}
\eta_{1-\gamma}(w_0,\boldsymbol{w}) = \min_{s,\boldsymbol{r},y_{i0},\boldsymbol{y}_i} \quad & s \\
\text{s.t.} \quad & r_1 + r_2 + r_3 \leq s \\
& y_{10} - \boldsymbol{y}_1'\underline{\boldsymbol{z}} \leq r_1 \\
& -\boldsymbol{y}_1 \geq 0 \\
& y_{20} + (1/\gamma - 1)\boldsymbol{y}_2'\underline{\boldsymbol{z}} \leq r_2 \\
& \boldsymbol{y}_2 \geq 0 \\
& y_{30} + \sqrt{\tfrac{1-\gamma}{\gamma}}\|\boldsymbol{y}_3\|_2 \leq r_3 \\
& y_{10} + y_{20} + y_{30} = w_0 \\
& \boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 = \boldsymbol{w},
\end{aligned}
$$

$$
\begin{aligned}
\pi(y_0,\boldsymbol{y}) = \min_{s,\boldsymbol{r},y_{i0},\boldsymbol{y}_i} \quad & s \\
\text{s.t.} \quad & r_1 + r_2 + r_3 \leq s \\
& y_{10} - \boldsymbol{y}_1'\underline{\boldsymbol{z}} \leq r_1 \\
& 0 \leq r_1 \\
& -\boldsymbol{y}_1 \geq 0 \\
& \boldsymbol{y}_2'\underline{\boldsymbol{z}} \leq r_2 \\
& y_{20} \leq r_2 \\
& \boldsymbol{y}_2 \geq 0 \\
& \tfrac{1}{2}y_{30} + \tfrac{1}{2}\|(y_{30},\boldsymbol{y}_3)\|_2 \leq r_3 \\
& y_{10} + y_{20} + y_{30} = y_0 \\
& \boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 = \boldsymbol{y},
\end{aligned}
$$

and $\underline{z}_j = 10$ for $j = 1,\dots,m$. Therefore, the deterministic approximation of the subproblem using robust optimization has $2m$ second order cones of 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, $u_1,\dots,u_M$, we can estimate the shortfall

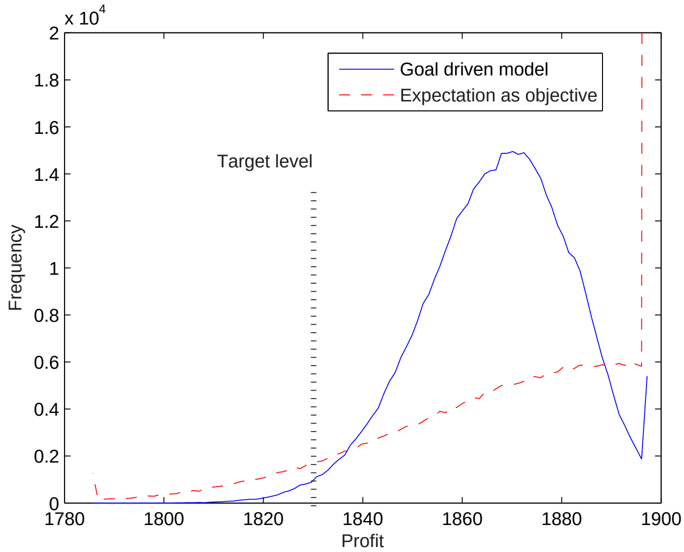

Figure 2: Goal driven optimization versus maximizing expected profit ($m = 10$)

aspiration level criterion as follows:

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

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 five to 30.

Figure 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 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.

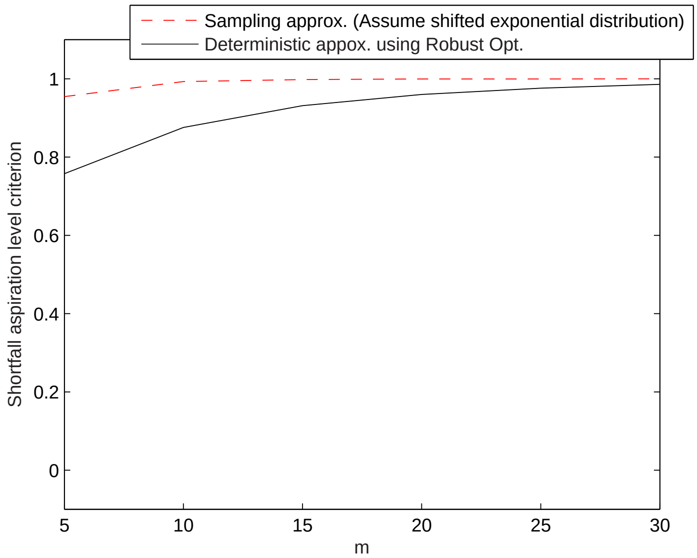

Figure 3: Shortfall aspiration level criteria evaluated on shifted exponential distribution with sampling approximation using the same distribution.

The software calls upon CPLEX 10.0 to solve the underlying SOCP. It takes less than 0.5 seconds to solve Problem (31) 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, 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 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 5 suggests that the robust solution can grossly outperform the solution obtained by sampled approximation using a different distribution with identical mean and standard deviation.

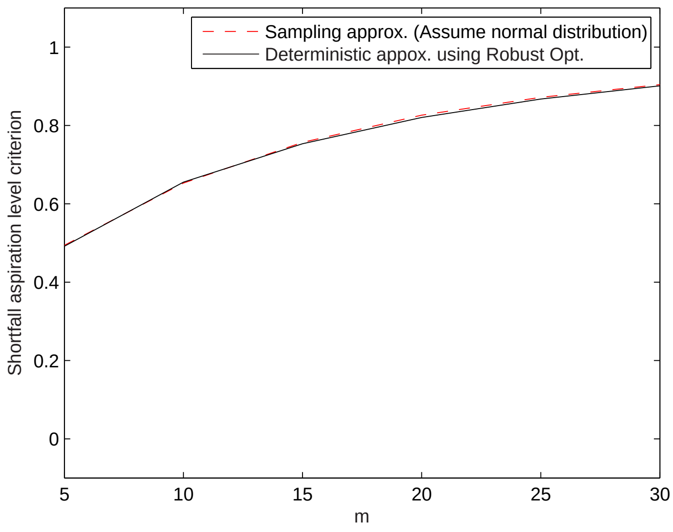

Figure 4: Shortfall aspiration level criteria evaluated on normal distribution with sampling approximation using the same distribution.

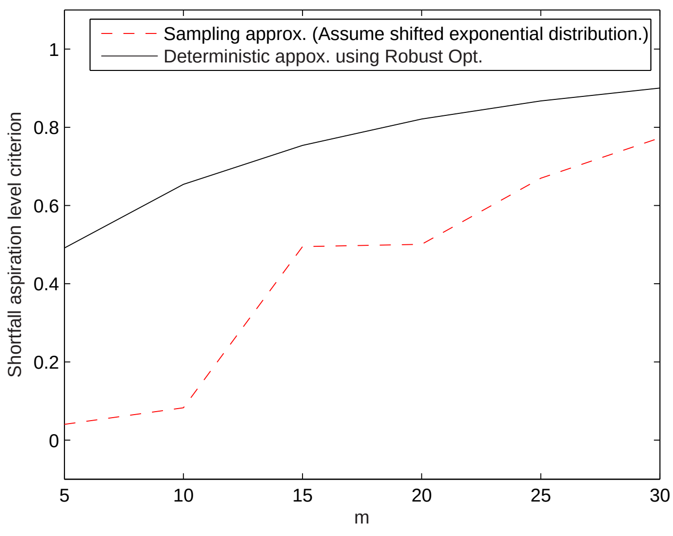

Figure 5: Shortfall aspiration level criteria evaluated on normal distribution with sampling approximation using the shifted exponential distribution.

5 Conclusions

We propose a new framework for modeling stochastic optimization problem that takes into account of an aspiration level. We also introduce the shortfall aspiration level criterion, which factors into the success probability and the adversity of under-performance. Moreover, the goal driven optimization model that maximizes the shortfall aspiration level criteria is analytically tractable.

We also propose two methods of solving the goal driven optimization problem, one using sampling approximations, while the other using deterministic approximations. Although the exposition in this paper is confined to a two period model, the deterministic approximation via decision rule can easily be extended to multiperiod modeling; see for instance Chen et al. [16]. This has immense advantage over sampling approximation.

Acknowledgements

We would like to thank the reviewers of the paper for valuable insightful comments and editorial suggestions.

A Proof of Theorem 3

(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

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

Hence,

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

where the second equality follows from choosing the optimum θ,

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

(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

$$
\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^+\right) \leq \mu \mathrm{E}(\exp((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})/\mu - 1)) = \mu \exp(y_0/\mu - 1)\prod_{j=1}^{I} \mathrm{E}(\exp(y_j\tilde{z}_j/\mu)) \qquad \forall \mu > 0. \quad (32)
$$

This relation was first shown in Nemirovski and Shapiro [28]. Using the deviation measures of Chen, Sim and Sun [15], and Proposition 2(c), we have

$$\ln(\mathrm{E}(\exp(y_j\tilde{z}_j/\mu))) \leq \begin{cases} {y_j}^2{p_j}^2/(2\mu^2) & \text{if } y_j \geq 0 \\ {y_j}^2{q_j}^2/(2\mu^2) & \text{otherwise.} \end{cases} \tag{33}$$

Since p_j and q_j are nonnegative, we have

$$\ln(\mathrm{E}(\exp(y_j\tilde{z}_j/\mu))) \leq \frac{(\max\{y_jp_j, -y_jq_j\})^2}{2\mu^2} = \frac{u_j^2}{2\mu^2}. \tag{34}$$

Substituting this in the inequality (32), 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 $v_j = \max\{-p_jy_j, q_jy_j\}$, $j = 1,\ldots,I$. Hence, following from the above exposition, we have

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

■

B 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

$$\begin{aligned} & b^2 \leq \mu d \\ & a + d \leq x. \end{aligned}$$

To approximate the conic exponential constraint, we use the method described in Ben-Tal and Nemirovski [5]. 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

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

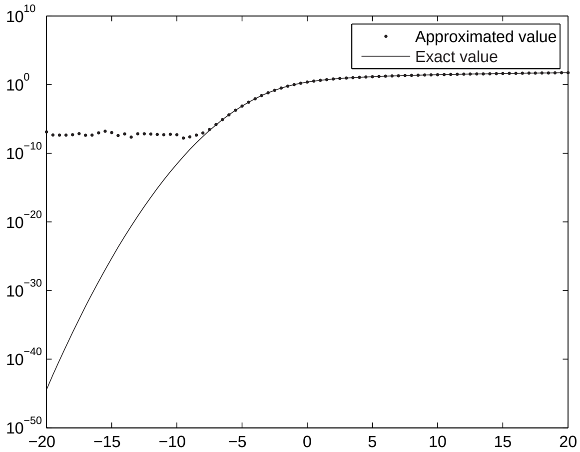

Figure 6: Evaluation of approximation of $\inf_{\mu>0} \mu \exp\left(\frac{a}{\mu} + \frac{1}{\mu^2}\right)$.

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 6, 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.

References

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

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

[3] Ben-Tal, A., A. Goryashko, E. Guslitzer and A. Nemirovski. (2004): Adjusting robust Solutions of uncertain linear programs. *Mathematical Programming*, 99, 351-376.

[4] Ben-Tal, A., Nemirovski, A. (1998): Robust convex optimization, *Math. Oper. Res.*, 23, 769-805.

[5] Ben-Tal, A., Nemirovski, A. (2001): Lectures on modern convex optimization: analysis, algorithms, and engineering applications, *MPR-SIAM Series on Optimization*, SIAM, Philadelphia.

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

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

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

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

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

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

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

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

[14] Charnes, A.,Cooper, W. W. (1963): Deterministic equivalents for optimizing and satisficing under chance constraints, *Operations Research*, 11, 18-39.

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

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

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

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

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

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

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

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

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

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

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

[26] Mao, J. C. T. (1970): Survey of capital budgeting: theory and practice, *Journal of Finance*, 25, 349-360.

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

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

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

[30] 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.

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

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

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

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

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

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

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

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