A Tractable Approximation of Stochastic Programming via Robust Optimization

Xin Chen[*], Melvyn Sim[†], Peng Sun[‡] and Jiawei Zhang[§]

February 2006

Abstract

Stochastic programming, despite its immense modeling capabilities, is well known to be computationally excruciating. In this paper, we introduce a unified framework of approximating multiperiod stochastic programming from the perspective of robust optimization. Specifically, we propose a framework that integrates multistage modeling with safeguarding constraints. The framework is computationally tractable in the form of second order cone programming (SOCP) and scalable across periods. We compare the computational performance of our proposal with classical stochastic programming approach using sampling approximations and report very encouraging results for a class of project management problems.

[*]Department of Mechanical and Industrial Engineering, University of Illinois at Urbana-Champaign. Email: xinchen@uiuc.edu

[†]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 and NUS academic research grants R-314-000-066-122 and R-314-000-068-112

[‡]Fuqua School of Business, Duke University Box 90120, Durham, NC 27708, USA. Email: psun@duke.edu

[§]Stern School of Business, New York University, New York, NY, 10012. Email: jzhang@stern.nyu.edu

1 Introduction

The study of stochastic programming dates back to Beale [1] and Dantzig [18]. In a typical two-stage stochastic program, decisions are made in the first stage in the face of uncertainty. Once the uncertainties are realized, the optimal second stage decisions or recourse decisions are carried out. Such "stochastic programs" attempt to integrate optimization and stochastic modeling that could potentially solve a large class of important practical problems ranging from engineering control to supply chain management. (see, e.g. Ruszczynski and Shapiro [25], Birge and Louveaux [14].) Despite the immense modeling potential, stochastic programs, especially multistage ones, are notoriously difficult to solve to optimality (see Shapiro and Nemirovski [27], Dyer and Stougie [19]). Quite often, finding a feasible solution is already a hard problem. It is therefore important to develop a tractable and scalable methodology that could reasonably approximate stochastic programs.

Besides multiperiod modeling, another aspect of stochastic programming deals with the constraints' feasibility under parameter uncertainties, which we generally term as safeguarding constraints. Typically, we either ignore parameter variability, which could lead to massive violations of constraints (see Ben-Tal and Nemirovski [4]), or require the constraints to be satisfied for all possible realizations of uncertainties, which can be excessively conservative. The "middle path" of these extremes is to enforce safeguarding constraints to within some prescribed levels of feasibility. In robust optimization, we control the level of feasibility of the constraints by adjusting the sizes of the uncertainty sets (see Ben-Tal and Nemirovski [2, 3, 4], El-Ghaoui et al. [20, 21], Iyangar and Goldfarb [22], Bertsimas and Sim [9, 10, 11, 12] and Chen, Sim and Sun [16]). Charnes and Cooper [15] introduced the chance-constrained formulation, in which feasibility level is measured by the probability of constraint violations. Unfortunately, such constraints are generally non-convex and intractable. Bernstein approximation (see Nemirovski and Shapiro [23]) provides reasonably good approximation to these problems in the form of tractable convex optimization problems. Under mild distributional assumptions, the robust optimization framework of Chen, Sim and Sun [16] provides approximations to chance-constrained problems in the from of second order cone optimization problems (SOCP), which have the benefits of greater tractability both in theory and in practice.

Literatures on multistage stochastic programs with safeguarding constraints are rather limited, perhaps due to the lack of tractable solution methodologies. To the best of our knowledge, it is not until recently that Chen, Sim and Sun [16] proposed tractable methodologies that lead to approximate solutions to such models. A closely related approach is Ben-Tal et al. [6], which propose an adjustable robust counterpart to handle dynamic decision making under uncertainty. We note that the uncertainties addressed by the model of Ben-Tal et al. [6] is non-stochastic, while the uncertainties considered in Chen, Sim and Sun [16] require mild distributional assumptions such as known mean, support and some deviation measures of the random data. In these models, *linear decision rule* is the key enabling mechanism that permits scalability to mutistage models. Interesting applications of such models include supplier-retailer flexible commitments contracts (Ben-Tal et al. [7]), project crashing with uncertain activity time (Chen, Sim and Sun [16]) and analyzing distribution systems with transhipment (Chou, Sim and So [17]).

In this paper, we propose a framework for approximating multistage stochastic optimization with chance constraints and *semi-complete recourse*. We adopt the phrase *semi-complete recourse* from stochastic programming terminology as a less restrictive condition compared to complete recourse problems. When hard constraints in the model are inevitable, we show that linear decision rule can lead to infeasible instances even under complete recourse, which motivates our proposal of a new *deflected linear decision rule* suited for stochastic models with semi-complete recourse. We introduce "computationally friendly" models in the form of second order cone program (SOCP), which could be solved efficiently both in theory and in practice. In our approach, the distributions of the primitive uncertainties do not have to be fully specified. Overall, the most important feature of our proposal is the scalability to multistage stochastic programs.

The structure of the paper is as follows. In Section 2, we describe the general stochastic model with chance constraints and semi-complete recourse. In Section 3 we propose the deflected linear decision rule and use it to approximate the general model. Section 4 discusses methods of approximating chance constraints and the objective function. We then summarizes the previous discussions in Section 5 and provide a second order cone formulation to approximate the original problem. In Section 6 we present encouraging preliminary computational results. Section 7 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 Stochastic Programming Model with Chance Constraints and Semi-complete Recourse

A classical two-stage stochastic program with fixed recourse can be formulated as follows (see, e.g. Ruszczynski and Shapiro [25]).

$$\begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} + \mathrm{E}(Q(\boldsymbol{x}, \tilde{\boldsymbol{z}})) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned} \tag{1}$$

where

$$\begin{aligned} Q(\boldsymbol{x}, \boldsymbol{z}) = \quad \min \quad & \boldsymbol{q}'\boldsymbol{y} \\ \text{s.t.} \quad & \boldsymbol{T}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{D}\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{z}) \\ & \boldsymbol{y} \geq \boldsymbol{0}. \end{aligned} \tag{2}$$

We define $\tilde{\boldsymbol{z}} \in \Re^N$ as the vector of primitive uncertainties that consolidate all underlying uncertainties in the stochastic model and E to represent the expectation on random variable $\tilde{\boldsymbol{z}}$. The stochastic model represents a sequence of events. Here vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are the first and the second stage decision variables, respectively. The second stage decision (a.k.a recourse decision), $\boldsymbol{y}$ is made after applying decision $\boldsymbol{x}$ and after the actual value of $\tilde{\boldsymbol{z}}$ is realized. For given $(\boldsymbol{x}, \boldsymbol{z})$, the second stage cost $Q(\boldsymbol{x}, \boldsymbol{z})$ is set to be $+\infty$ if the feasible set of (2) is empty, and $-\infty$ if problem (2) is unbounded from below. It can be shown (see e.g. Ruszczynski and Shapiro (2003)) that, under very general conditions, problem (1) is equivalent to

$$\begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} + \mathrm{E}(\boldsymbol{q}'\boldsymbol{y}(\tilde{\boldsymbol{z}})) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{T}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \boldsymbol{D}\boldsymbol{y}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}(\tilde{\boldsymbol{z}}) \\ & \boldsymbol{x} \geq 0 \\ & \boldsymbol{y}(\tilde{\boldsymbol{z}}) \geq \boldsymbol{0} \\ & \boldsymbol{y}(\cdot) \in \mathcal{Y}, \end{aligned} \tag{3}$$

where $\mathcal{Y}$ is a space of measurable functions. The vector of functions, $\boldsymbol{y}(\cdot)$, corresponds to the second stage decision vector or *recourse variables* as functions of the realization of $\tilde{\boldsymbol{z}}$. There are several types of recourse in the context of stochastic programming. In fixed recourse problems, the matrix $\boldsymbol{D}$ is not subject to uncertainty. The stochastic program (1) is said to have *relatively complete recourse* if for any $\boldsymbol{x} \in \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$, $\mathrm{E}(Q(\boldsymbol{x}, \tilde{\boldsymbol{z}})) < +\infty$. Relatively complete recourse problems ensure that the second stage problem is always feasible for any choice of feasible first stage decision vector, $\boldsymbol{x}$. It is generally not easy to identify conditions of relatively complete recourse (see Birge and Louveaux [14]). An important class of relatively complete recourse problems is known as complete recourse, which is defined on the matrix $\boldsymbol{D}$ such that for any $\boldsymbol{t}$, there exists $\boldsymbol{y} \geq \boldsymbol{0}$, s.t. $\boldsymbol{D}\boldsymbol{y} = \boldsymbol{t}$. Hence, complete recourse problems depend on the structure of the matrix $\boldsymbol{D}$, which is an easier condition to identify. Moreover, many stochastic programming problems have complete recourse. A special case of complete recourse is simple recourse, where $\boldsymbol{D} = [\boldsymbol{I} - \boldsymbol{I}]$.

Another aspect of stochastic programming problems is probabilistic constraints or chance constraints introduced by Charnes and Cooper [15], which is almost independently addressed from multiperiod models. Chance constrained problems in the form of a single stage problem are as follows

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \mathrm{P}(\boldsymbol{a}_i(\tilde{\boldsymbol{z}})'\boldsymbol{x} \leq b_i(\tilde{\boldsymbol{z}})) \geq 1 - \epsilon_i. \end{array} \tag{4}$$

The ith constraint is allowed to be violated within probability level less than ϵ_i. Therefore, we can view the usual nonnegative constraint or hard constraint as a special case of chance constraint in which ϵ_i is zero.

Based on these frameworks, we propose a two-stage stochastic optimization model with fixed recourse and chance constraints as follows:

$$\begin{array}{lll} Z_{STOC} = \min & \boldsymbol{c}'\boldsymbol{x} + \mathrm{E}(\boldsymbol{d}'\boldsymbol{v}(\tilde{\boldsymbol{z}})) + \mathrm{E}(\boldsymbol{f}'\boldsymbol{w}(\tilde{\boldsymbol{z}})) & \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\ & \boldsymbol{U}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}(\tilde{\boldsymbol{z}}) + \boldsymbol{W}\boldsymbol{w}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}(\tilde{\boldsymbol{z}}) & \\ & \mathrm{P}(v_j(\tilde{\boldsymbol{z}}) \geq 0) \geq 1 - \epsilon_j & \forall j \in \{1, \ldots, n_2\} \\ & w_i(\tilde{\boldsymbol{z}}) \geq 0 & \forall i \in I \subseteq \{1, \ldots, n_3\} \\ & \boldsymbol{x} \geq 0 & \\ & \boldsymbol{v}(\cdot), \boldsymbol{w}(\cdot) \in \mathcal{Y} & \end{array} \tag{5}$$

where $\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{f}$ and $\boldsymbol{b}$ are known vectors in $\Re^{n_1}, \Re^{n_2}, \Re^{n_3}$ and $\Re^{m_1}$ respectively, $\boldsymbol{A}, \boldsymbol{V}$ and $\boldsymbol{W}$ are known matrices in $\Re^{m_1 \times n_1}, \Re^{m_2 \times n_2}, \Re^{m_2 \times n_3}$. We assume the following affine data dependency on $\boldsymbol{U}(\tilde{\boldsymbol{z}}) \in \Re^{m_2 \times n_1}$ and $\boldsymbol{h}(\tilde{\boldsymbol{z}}) \in \Re^{m_2}$, that is,

$$\boldsymbol{U}(\tilde{\boldsymbol{z}}) = \boldsymbol{U}^0 + \sum_{k=1}^{N} \boldsymbol{U}^k \tilde{z}_k,$$
$$\boldsymbol{h}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}^0 + \sum_{k=1}^{N} \boldsymbol{h}^k \tilde{z}_k.$$

In the proposed stochastic optimization model, we make distinctions between the recourse variables $\boldsymbol{v}(\tilde{\boldsymbol{z}})$ and $\boldsymbol{w}(\tilde{\boldsymbol{z}})$.

(a) The recourse variables $v_i(\tilde{\boldsymbol{z}})$ are permitted to be negative with low probability, that is $v_i(\tilde{\boldsymbol{z}}) < 0$ with probability at most $\epsilon_i > 0$.

(b) The recourse variable $w_i(\tilde{\boldsymbol{z}})$, $i \in I$ must satisfy the inequality constraint for all outcomes.

We propose a new notion of *semi-complete recourse*, which is a characteristic of the matrix $\boldsymbol{W}$.

Definition 1 *We say that the stochastic program (5) has semi-complete recourse if there exists vector* $\boldsymbol{r}$ *with* $r_i > 0$ *for all* $i \in I$, *such that* $\boldsymbol{W}\boldsymbol{r} = \boldsymbol{0}$.

To motivate the model, we next present some examples to illustrate the generality of Model (5).

Example 1: Single stage chance constrained problems

With $\boldsymbol{W} = \boldsymbol{0}$, which is the simplest case of semi-complete recourse, $\boldsymbol{V} = \boldsymbol{I}$, $\boldsymbol{d} = \boldsymbol{0}$ and $\boldsymbol{f} = \boldsymbol{0}$, Model (5) is essentially the same as Model (4).

Example 2: Complete recourse problems

In complete recourse problems, we have $\boldsymbol{V} = 0$ and $\boldsymbol{d} = \boldsymbol{0}$, and the matrix $\boldsymbol{W}$ satisfies the condition of complete recourse, which implies that for any $\boldsymbol{t}$, there exists $w_i \geq 0$, $i \in I$, such that $\boldsymbol{D}\boldsymbol{w} = \boldsymbol{t}$. We next show that if a matrix $\boldsymbol{W}$ satisfies the condition of complete recourse, it also satisfies the condition of semi-complete.

Proposition 1 *Under complete recourse, there exists* $\boldsymbol{r}$ *with* $r_i > 0$ *for all* $i \in I$, *such that* $\boldsymbol{W}\boldsymbol{r} = \boldsymbol{0}$.

Proof : By definition of complete recourse, for any $\boldsymbol{v} > \boldsymbol{0}$, we can find a vector $\boldsymbol{s} \geq \boldsymbol{0}$ such that $\boldsymbol{W}\boldsymbol{s} = -\boldsymbol{W}\boldsymbol{v}$. Clearly, $\boldsymbol{r} := \boldsymbol{s} + \boldsymbol{v} > \boldsymbol{0}$ and $\boldsymbol{W}\boldsymbol{r} = \boldsymbol{0}$. ∎

We refer to *strict semi-complete* recourse for the case that the matrix $\boldsymbol{W}$ satisfies the condition of semi-complete recourse but not complete recourse. A very simple example of strict semi-complete recourse is $\boldsymbol{W} = \boldsymbol{0}$.

Example 3: Chance constraints with violation penalties

Consider the following two stage stochastic optimization problem with chance constraints and penalties of constraint violations.

$$
\begin{array}{lll}
\min & \boldsymbol{c}'\boldsymbol{x} + \mathrm{E}(\boldsymbol{d}'\boldsymbol{v}(\tilde{\boldsymbol{z}})) + \mathrm{E}(\bar{\boldsymbol{d}}'\boldsymbol{v}(\tilde{\boldsymbol{z}})^-) & \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& \boldsymbol{U}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}(\tilde{\boldsymbol{z}}) & \\
& \mathrm{P}(v_j(\tilde{\boldsymbol{z}}) \geq 0) \geq 1 - \epsilon_j & \forall j \\
& \boldsymbol{x} \geq 0, &
\end{array} \tag{6}
$$

where $\bar{\boldsymbol{d}}$ is a non-negative vector associated with violation of the chance constraints. The effective cost contribution associated with the recourse action is

$$
d_i v_i(\tilde{\boldsymbol{z}}) + \bar{d}_i v_i(\tilde{\boldsymbol{z}})^-,
$$

where $\bar{d}_i(v_i(\tilde{\boldsymbol{z}}))^-$ is the cost penalty for constraint violation. Clearly, we can linearized the objective function of Model (6) as follows:

$$
\begin{array}{lll}
\min & \boldsymbol{c}'\boldsymbol{x} + \mathrm{E}(\boldsymbol{d}'\boldsymbol{v}(\tilde{\boldsymbol{z}})) + \mathrm{E}(\bar{\boldsymbol{d}}'\boldsymbol{w}_1(\tilde{\boldsymbol{z}})) & \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& \boldsymbol{U}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}(\tilde{\boldsymbol{z}}) & \\
& \boldsymbol{v}(\tilde{\boldsymbol{z}}) + \boldsymbol{w}_1(\tilde{\boldsymbol{z}}) - \boldsymbol{w}_2(\tilde{\boldsymbol{z}}) = \boldsymbol{0} & \\
& \mathrm{P}(v_j(\tilde{\boldsymbol{z}}) \geq 0) \geq 1 - \epsilon_j & \forall j \\
& \boldsymbol{x} \geq 0 & \\
& \boldsymbol{w}_1(\tilde{\boldsymbol{z}}), \boldsymbol{w}_2(\tilde{\boldsymbol{z}}) \geq \boldsymbol{0}, &
\end{array}
$$

and the associated matrix corresponding to $\boldsymbol{W}$ in Model (5) has the form of

$$
\boldsymbol{W} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{I} & -\boldsymbol{I} \end{bmatrix},
$$

which trivially satisfies the condition of strict semi-complete recourse.

Example 4: Distributional systems with transshipment

This example is adopted from Chou, Sim and So [17] to motivate the framework. Consider a distribution system of n retailers and arc set $\mathcal{E}$ facing exogenous random demand $\boldsymbol{d}(\tilde{\boldsymbol{z}})$. At the first stage, each retailer i decides its inventory level x_i, which incurs unit holding cost, h_i. At the second stage, demands are realized. When shortages at one location occur, we can cover from available inventories at other retail locations through possible lateral transshipment. Assuming a fully connected network, let c_{ij} be the unit transportation cost between retailers i and j for all $(i,j) \in \mathcal{E}$. We assume cost symmetry, that is $c_{ij} = c_{ji}$, and denote $w_{ij}(\tilde{\boldsymbol{z}})$ as the recourse decision related to transhipment quantity from retailer i to j after the realization of demands. Hence, the final inventory balance at retailer i is

$$v_i(\tilde{\boldsymbol{z}}) = x_i - d_i(\tilde{\boldsymbol{z}}) - \sum_{j:(i,j)\in\mathcal{E}} w_{ij}(\tilde{\boldsymbol{z}}) + \sum_{j:(j,i)\in\mathcal{E}} w_{ji}(\tilde{\boldsymbol{z}}),$$

which takes negative values when shortages occur. The goal is to find an initial inventory allocation so as to minimize the total inventory holding cost and expected transhipment cost while ensuring certain service levels, i.e., the demand at each retailer will be satisfied with high probability. Hence, we have the following formulation

$$\begin{array}{lll} \min & \boldsymbol{h}'\boldsymbol{x} + \sum_{(i,j)\in\mathcal{E}} c_{ij}\mathrm{E}\left(w_{ij}(\tilde{\boldsymbol{z}})\right) & \\ \text{s.t.} & v_i(\tilde{\boldsymbol{z}}) = x_i - d_i(\tilde{\boldsymbol{z}}) - \sum_{j:(i,j)\in\mathcal{E}} w_{ij}(\tilde{\boldsymbol{z}}) + \sum_{j:(j,i)\in\mathcal{E}} w_{ji}(\tilde{\boldsymbol{z}}) & \forall i \\ & \mathrm{P}(v_i(\tilde{\boldsymbol{z}}) \geq 0) \geq 1 - \epsilon_i & \forall i \in I \\ & \boldsymbol{w}(\tilde{\boldsymbol{z}}) \geq \boldsymbol{0} & \\ & \boldsymbol{x} \geq \boldsymbol{0}. & \end{array} \qquad (7)$$

To check semi-complete recourse, we send a unit flow across every arc. Since the network is fully connected, flow conservation is maintained at every node, satisfying the condition of semi-complete recourse. Note that when $n = 2$, the recourse matrix,

$$W = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix},$$

does not satisfy the condition for complete recourse. As a matter of fact, the total flow into all the nodes must be maintained at zero. Hence, Model (7) has strict semi-complete recourse.

3 Recourse Approximation via Decision Rules

Modern solution methodology for stochastic optimization model (see for instance, Birge and Louveaux [14]) focuses on solving multistage problems with hard constraints (as in the case of Model (2) and Model (5) in which $\epsilon_j = 0$). 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. Indeed, under 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.

Due to the astronomical number of scenarios, Monte Carlo sampling methods are often used to obtain approximate solutions to stochastic optimization problems. Despite the wide adoption of such sampling approximation in stochastic optimization, the theoretical performance of the method has only been studied recently, for instance, by Shapiro and Nemirovski ([27]). They concluded that the number of samples required to approximate multistage stochastic programs to reasonable accuracy grows exponentially with the number of stages. Moreover, the problem remains hard even under complete recourse.

In view of the hardness results, we propose a tractable approximation for Model (5) by restricting the recourse decisions to specified decision rules. In Ben-Tal et al. [7], and Chen, Sim and Sun [16], linear decision rules have been used and promising computational results have been reported. We extend the linear decision rule and propose a deflected linear decision rule in order to tackle problems with semi-complete recourse.

3.1 Linear decision rules

The extremely large number of recourse variables leads to the computational intractability of stochastic optimization. To resolve the issue, we adopt linear decision rule proposed in Ben-Tal et al [6] and Chen, Sim and Sun [16]. Using linear decision rule, we restrict recourse variables, say $\boldsymbol{w}(\tilde{\boldsymbol{z}})$ and $\boldsymbol{v}(\tilde{\boldsymbol{z}})$, to be affinely dependent on the primitive uncertainties. We denote $\mathcal{L}$ as the space of linear decision functions. Hence, $v(\cdot) \in \mathcal{L} \subseteq \mathcal{Y}$ implies that there exists a set of vectors $\boldsymbol{v}^0, \ldots, \boldsymbol{v}^N$ such that

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

We can approximate the stochastic model (5) as follows:

$$
\begin{array}{lll}
Z_{LDR} = \min & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{d}'\boldsymbol{v}^0 + \boldsymbol{f}'\boldsymbol{w}^0 & \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& \boldsymbol{U}^k\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}^k + \boldsymbol{W}\boldsymbol{w}^k = \boldsymbol{h}^k & \forall k \in \{0, \ldots, N\} \\
& \mathrm{P}(v_j(\tilde{\boldsymbol{z}}) \geq 0) \geq 1 - \epsilon_j & \forall j \in \{1, \ldots, n_2\} \\
& w_i(\boldsymbol{z}) \geq 0 & \forall \boldsymbol{z} \in \mathcal{W}, \forall i \in I \subseteq \{1, \ldots, n_3\} \\
& \boldsymbol{x} \geq 0 & \\
& \boldsymbol{v}(\cdot), \boldsymbol{w}(\cdot) \in \mathcal{L}. &
\end{array}
\tag{8}
$$

Since any feasible solution of Model (8) is also feasible in (5) and the objectives coincide, we have $Z_{STOC} \leq Z_{LDR}$. We will defer the discussion of approximating the constraints in Model (8) to Section 4.

3.2 Deflected linear decision rules

3.2.1 On linear decision rules and hard constraints

Although linear decision rule may not be optimal, Chen, Sim and Sun [16] show encouraging computational results for models with chance constraints. However, linear decision rules under "hard" inequality constraints such as,

$$\boldsymbol{w}(\tilde{\boldsymbol{z}}) \geq 0,$$

may perform poorly. As an illustration, suppose the primitive uncertainties, $\tilde{\boldsymbol{z}}$ have infinite support, then the following nonnegativity constraints

$$\boldsymbol{w}(\tilde{\boldsymbol{z}}) = \boldsymbol{w}^0 + \sum_{k=1}^{N} \boldsymbol{w}^k \tilde{z}_k \geq \boldsymbol{0}$$

imply that

$$\boldsymbol{w}^k = \boldsymbol{0} \quad \forall k \in \{1, \ldots, N\},$$

and the decision rule is reduced to $\boldsymbol{w}(\tilde{\boldsymbol{z}}) = \boldsymbol{w}^0$, and hence, independent of the primitive uncertainties. This may lead to infeasible instance even in the case of complete recourse. For example, consider the following stochastic optimization model that determine $\mathrm{E}(|b(\tilde{\boldsymbol{z}})|)$:

$$\min \left\{ \mathrm{E}\left(w_1(\tilde{\boldsymbol{z}}) + w_2(\tilde{\boldsymbol{z}})\right) : w_1(\tilde{\boldsymbol{z}}) - w_2(\tilde{\boldsymbol{z}}) = b(\tilde{\boldsymbol{z}}), w_1(\tilde{\boldsymbol{z}}) \geq 0, w_2(\tilde{\boldsymbol{z}}) \geq 0 \right\}, \tag{9}$$

which is one with simple recourse. Suppose $\tilde{\boldsymbol{z}}$ has infinite support, we must have $w_1(\tilde{\boldsymbol{z}}) = w_1^0$ and $w_2(\tilde{\boldsymbol{z}}) = w_2^0$, and hence, it would be impossible to satisfy the equality constraint. Furthermore, one should note that in many such problems, it is meaningless to replace the hard constraints with "soft" constraints.

Our next goal is to improve upon the decision rules involving semi-complete recourse in the presence of hard constraints. The conditions of the semi-complete recourse implies that even if there is a constraint violation that is due to $\boldsymbol{w}(\tilde{\boldsymbol{z}})$, we can still steer back the solution towards feasibility by paying a finite price. We will illustrate this idea next.

3.2.2 Deflected linear decision rules

Under semi-complete recourse, for each $i \in I$, the following linear program

$$\begin{aligned} \bar{f}_i = \min \quad & \boldsymbol{f}'\boldsymbol{p} \\ \text{s.t.} \quad & \boldsymbol{W}\boldsymbol{p} = 0 \\ & p_i = 1 \\ & \boldsymbol{p} \geq 0 \end{aligned} \tag{10}$$

is feasible. We assume that the linear program (10) has an optimal solution, denoted as $\bar{\boldsymbol{p}}^i$; otherwise the problem (5) is unbounded. Then we define, for each i, $\bar{f}_i = \boldsymbol{f}'\bar{\boldsymbol{p}}^i$. Therefore, for any decision rules $\boldsymbol{r}(\tilde{\boldsymbol{z}})$ (which may not necessarily be nonnegative) and $\boldsymbol{v}(\tilde{\boldsymbol{z}})$ satisfying

$$\boldsymbol{U}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}(\tilde{\boldsymbol{z}}) + \boldsymbol{W}\boldsymbol{r}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}(\tilde{\boldsymbol{z}}), \tag{11}$$

we let

$$\boldsymbol{w}(\tilde{\boldsymbol{z}}) = \boldsymbol{r}(\tilde{\boldsymbol{z}}) + \sum_{i \in I} (r_i(\tilde{\boldsymbol{z}})^-)\bar{\boldsymbol{p}}^i. \tag{12}$$

It can be easily verified that

$$\begin{aligned} w_i(\tilde{\boldsymbol{z}}) \geq 0 \qquad & \forall i \in I \\ \boldsymbol{W}\boldsymbol{w}(\tilde{\boldsymbol{z}}) = \boldsymbol{W}\boldsymbol{r}(\tilde{\boldsymbol{z}}). & \end{aligned}$$

Therefore, for any given $\boldsymbol{x}$, as long as there exist $\boldsymbol{r}(\tilde{\boldsymbol{z}})$ and $\boldsymbol{v}(\tilde{\boldsymbol{z}})$ satisfying (11), we can find a feasible decision rule (referred to as a *deflected decision rule*), $\boldsymbol{w}(\tilde{\boldsymbol{z}})$. We note that the feasibility of (11) depends on the solution in the first stage, i.e., $\boldsymbol{x}$.

For the case of complete recourse, we can obtain stronger results. Not only does there exist a feasible deflected decision rule, there exists a feasible *deflected linear decision rule*, that is, Equation (12) with $\boldsymbol{r}(\cdot)$ being linear functions. We summarize the results in the next two propositions.

Proposition 2 *Under complete recourse, for any $\boldsymbol{x}$ and $\boldsymbol{v}(\cdot) \in \mathcal{L}$, there exists $\boldsymbol{r}(\cdot) \in \mathcal{L}$ such that*

$$\boldsymbol{U}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}(\tilde{\boldsymbol{z}}) + \boldsymbol{W}\boldsymbol{r}(\tilde{\boldsymbol{z}}) = \boldsymbol{h}(\tilde{\boldsymbol{z}}).$$

Proof : Let

$$v(\tilde{\boldsymbol{z}}) = \boldsymbol{v}^0 + \sum_{k=1}^{N} \boldsymbol{v}^k \tilde{z}_k$$

and

$$r(\tilde{\boldsymbol{z}}) = \boldsymbol{r}^0 + \sum_{k=1}^{N} \boldsymbol{r}^k \tilde{z}_k$$

By the assumption of complete recourse, there exist $\boldsymbol{r}^0, \boldsymbol{r}^1, \cdots, \boldsymbol{r}^N$ such that

$$\boldsymbol{U}^k \boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}^k + \boldsymbol{W}\boldsymbol{r}^k = \boldsymbol{h}^k \qquad \forall k \in \{0, \ldots, N\}.$$

This implies the desired result. ■

Proposition 3 *Under complete recourse, for any given $\boldsymbol{x}$ satisfying $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \geq \boldsymbol{0}$, there exist $\boldsymbol{v}(\cdot), \boldsymbol{r}(\cdot) \in \mathcal{L}$ such that (12) is feasible to problem (5).*

Proof : Notice that $\boldsymbol{v}(\tilde{\boldsymbol{z}}) = 0$ always satisfies the chance constraints. Proposition 2 together with equation (12) suggests that $\boldsymbol{v}(\tilde{\boldsymbol{z}}) = 0$ and (12) is feasible to problem (5). ■

From (12), we have

$$\boldsymbol{f}'\boldsymbol{w}(\tilde{\boldsymbol{z}}) = \boldsymbol{f}'\boldsymbol{r}(\tilde{\boldsymbol{z}}) + \bar{\boldsymbol{f}}'(\boldsymbol{r}(\tilde{\boldsymbol{z}})^-).$$

Therefore, using the deflected linear decision rule $\boldsymbol{w}(\cdot)$ and linear decision rule $\boldsymbol{v}(\cdot)$, we can approximate problem (5) as

$$\begin{array}{lll}
Z_{DLDR} = \min & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{d}'\boldsymbol{v}^0 + \boldsymbol{f}'\boldsymbol{r}^0 + \mathrm{E}\left[\bar{\boldsymbol{f}}'\boldsymbol{r}(\tilde{\boldsymbol{z}})^-\right] & \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& \boldsymbol{U}^k\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}^k + \boldsymbol{W}\boldsymbol{r}^k = \boldsymbol{h}^k & \forall k \in \{0, \ldots, N\} \\
& \mathrm{P}(v_j(\tilde{\boldsymbol{z}}) \geq 0) \geq 1 - \epsilon_j & \forall j \in \{1, \ldots, n_2\} \\
& \boldsymbol{x} \geq 0 & \\
& \boldsymbol{r}(\cdot), \boldsymbol{v}(\cdot) \in \mathcal{L}. &
\end{array} \tag{13}$$

Notice that in formulating and solving the above model, we do not directly need $\bar{\boldsymbol{p}}^i$ defined in (10). In fact, what we really need is just $\bar{f}_i$, $i \in I$.

Since any feasible solution $(\boldsymbol{x}, \boldsymbol{v}(\tilde{\boldsymbol{z}}), \boldsymbol{w}(\tilde{\boldsymbol{z}}))$ to Model (13), in which $\boldsymbol{w}(\tilde{\boldsymbol{z}}) = \boldsymbol{r}(\tilde{\boldsymbol{z}}) + \sum_{i\in I}(\boldsymbol{r}(\tilde{\boldsymbol{z}}))^- \bar{\boldsymbol{p}}^i$, is feasible in (5), and the objectives coincide, we have $Z_{STOC} \leq Z_{DLDR}$. Moreover, given any feasible solution, $(\boldsymbol{x}, \boldsymbol{v}(\tilde{\boldsymbol{z}}), \boldsymbol{w}(\tilde{\boldsymbol{z}}))$, of problem (8), we observe that

$$\mathrm{E}\left[\bar{\boldsymbol{f}}'\boldsymbol{w}(\tilde{\boldsymbol{z}})^-\right] = 0.$$

Hence by letting $\boldsymbol{r}(\tilde{\boldsymbol{z}}) = \boldsymbol{w}(\tilde{\boldsymbol{z}})$, we obtain a feasible solution of Model (13) with the same objective. Hence, $Z_{STOC} \leq Z_{DLDR} \leq Z_{LDR}$.

Unfortunately, Model (13) is still hard to solve because of the chance constraint $\mathrm{P}(v_j(\tilde{\boldsymbol{z}}) \geq 0) \geq 1-\epsilon_j$ and the nonlinear term $\mathrm{E}\left[\bar{\boldsymbol{f}}'\boldsymbol{r}(\tilde{\boldsymbol{z}})^-\right]$ in the objective function. In the sequel, we approximate (13) via techniques from robust optimization so that the resulting model is in the form of second-order cone programming, which can be solved efficiently both in theory and in practice.

4 Approximation via Robust Optimization

In this section, we first present approximation to the chance constraints in problems (13) and (8). Then we discuss the approximation to the nonlinear term in the objective function in formulation (13).

4.1 Approximating chance constraints

Even in the absence of recourse, the chance constraint model (4) is already computationally challenging. Indeed, although chance constraint is well known to be tractable for normal distributions (see Prekopa [24]), such models usually lead to intractable models even for simple distributions such as uniform distributions (see Nemirovski and Shapiro [27]).

Another caveat with stochastic optimization models is the need to assume exact distributions in all the unknown parameters; without which, it would be impossible to proceed with Monte Carlo sampling approximations. However, as complete distributional knowledge is rarely available in practice, solutions that are highly tuned to a particular distributional assumption can perform badly in practice (see for instance computational results in Bertsimas and Thiele [13]). As a "practical approach" to specifying uncertainties, we only require partial description of the primitive uncertainties, $\tilde{\boldsymbol{z}}$ in Model (5). Specifically, we assume that each random variable $\tilde{z}_k$ has mean zero and support in $\mathcal{W} = [-\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}}]$ with $\underline{z}_k, \bar{z}_k > 0$. We also assume that the standard deviations of $\tilde{z}_k$ are known and equal to σ_k.

Consider a single linear chance constraint

$$\mathrm{P}(v(\tilde{\boldsymbol{z}}) \geq 0) \geq 1 - \epsilon, \tag{14}$$

in which

$$v(\tilde{\boldsymbol{z}}) = v_0 + \sum_{k=1}^{N} v_k \tilde{z}_k,$$

where $v_0, \ldots, v_N$ are the decision variables. Let $\mathcal{G}_\Omega \subset \mathcal{W}$ denotes a compact uncertainty set parameterized by an uncertainty budget $\Omega \geq 0$, where the uncertainty budget determines the size of the uncertainty set and, as we shall see, is essentially related to the probability of violating the constraint. In robust optimization, we denote the set of all feasible solutions $v_0, \ldots, v_N$ satisfying

$$v(\boldsymbol{z}) \geq 0 \quad \forall \boldsymbol{z} \in \mathcal{G}_\Omega, \tag{15}$$

as the “robust counterpart” with respect to the uncertainty budget Ω. Such models were usually “misinterpreted” as worst case models suggesting that robust optimization are being over-conservative in the treatment of uncertainties. In fact, the level of conservativeness depends on the choice of uncertainty set. For instance, if the set $\mathcal{G}_\Omega$ only contains the nominal value, we ignore data perturbation and recover the nominal constraint. On the other extreme, if $\mathcal{G}_\Omega = \mathcal{W}$, we require the solution to remain feasible for all possible data realization, which is indeed the worst case model addressed by Soyster, [28].

Recently, Chen, Sim and Sun ([16]) introduce new deviation measures to construct second order conic representable uncertainty set from the underlying probability distributions. They propose the following asymmetric uncertainty set (as illustrated in Figure 1),

$$\mathcal{G}_\Omega = \underbrace{\left\{\boldsymbol{z} : \exists \boldsymbol{v}, \boldsymbol{w} \in \Re^N_+, \boldsymbol{z} = \boldsymbol{v} - \boldsymbol{w}, \|\boldsymbol{P}^{-1}\boldsymbol{v} + \boldsymbol{Q}^{-1}\boldsymbol{w}\|_2 \leq \Omega\right\}}_{\mathcal{A}_\Omega} \cap \underbrace{\{-\underline{\boldsymbol{z}} \leq \boldsymbol{z} \leq \bar{\boldsymbol{z}}\}}_{\mathcal{W}}, \tag{16}$$

with $\boldsymbol{P} = \text{diag}(p_1, \ldots, p_N)$ and $\boldsymbol{Q} = \text{diag}(q_1, \ldots, q_N)$, where, p_j, q_j are deviation measures related to the random variable $\tilde{z}_j$. They define the forward deviation as

$$p_j^* = \sup_{\theta > 0} \left\{ \sqrt{2\ln(\text{E}(\exp(\theta \tilde{z}_j)))/\theta^2} \right\} \tag{17}$$

and backward deviation as

$$q_j^* = \sup_{\theta > 0} \left\{ \sqrt{2\ln(\text{E}(\exp(-\theta \tilde{z}_j)))/\theta^2} \right\}. \tag{18}$$

Given a sequence of independent samples, we can essentially estimate the magnitude of the deviation measures from (17) and (18). Some of the properties of the deviation measures include:

Proposition 4 *(Chen, Sim and Sun [16])*

(a) *If the random variable $\tilde{z}$ has mean zero and standard deviation, σ, then $p^* \geq \sigma$ and $q^* \geq \sigma$. If in addition, $\tilde{z}$ is normally distributed, then $p^* = q^* = \sigma$.*

(b)

$$\text{P}(\tilde{z} \geq \Omega p^*) \leq \exp(-\Omega^2/2)$$

$$\text{P}(\tilde{z} \leq -\Omega q^*) \leq \exp(-\Omega^2/2)$$

Proposition 4(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 4(b), the deviation measures provide an easy bound on the distributional tails. Chen, Sim and Sun ([16]) show that new deviation measures provide tighter approximation of probabilistic bounds compared to standard deviations. More interestingly, even if only the mean and the support of the random variables are known, Chen, Sim and Sun ([16]) give a tight bound to the deviation measures.

Theorem 1 *(Chen, Sim and Sun [16]) If $\tilde{z}$ has zero mean and distributed in $[-\underline{z}, \bar{z}]$, $\underline{z}, \bar{z} > 0$, then*

$$p^* \leq \bar{p} = \frac{\underline{z} + \bar{z}}{2} \sqrt{f\left(\frac{\underline{z} - \bar{z}}{\underline{z} + \bar{z}}\right)} \in \mathcal{P}(\tilde{z})$$

and

$$q^* \leq \bar{q} = \frac{\underline{z} + \bar{z}}{2} \sqrt{f\left(\frac{\bar{z} - \underline{z}}{\underline{z} + \bar{z}}\right)} \in \mathcal{Q}(\tilde{z}),$$

where

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

and

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

With respect to the construction of $\mathcal{G}_\Omega$, suppose the distribution of $\tilde{z}_j$ is available, it would be natural to choose $p_j = p_j^*$ and $q_j = q_j^*$. Otherwise, Theorem 1 suggests that we can also bound the deviation measures from the distributional support and choose $p_j = \bar{p}_j$ and $q_j = \bar{q}_j$. For instance, if $\tilde{z}_j$ is arbitrarily distributed over $[-1, 1]$, then we set $p_j = q_j = 1$. Suppose the distribution is uniform over $[-1, 1]$, we have $p_j = q_j = 0.58$, which leads to a smaller uncertainty set compared to the one with arbitrary distribution. More importantly, Theorem 1 establishes the fact that the deviation measures are finite for all random variables with finite support, which is the case for most problems in practice.

The next result shows how we can incorporate the uncertainty set to formulate the robust counterpart as a tractable convex optimization model.

Theorem 2 *(Chen, Sim and Sun [16])*
The robust counterpart of

$$v(\boldsymbol{z}) \geq 0 \quad \forall \boldsymbol{z} \in \mathcal{G}_\Omega,$$

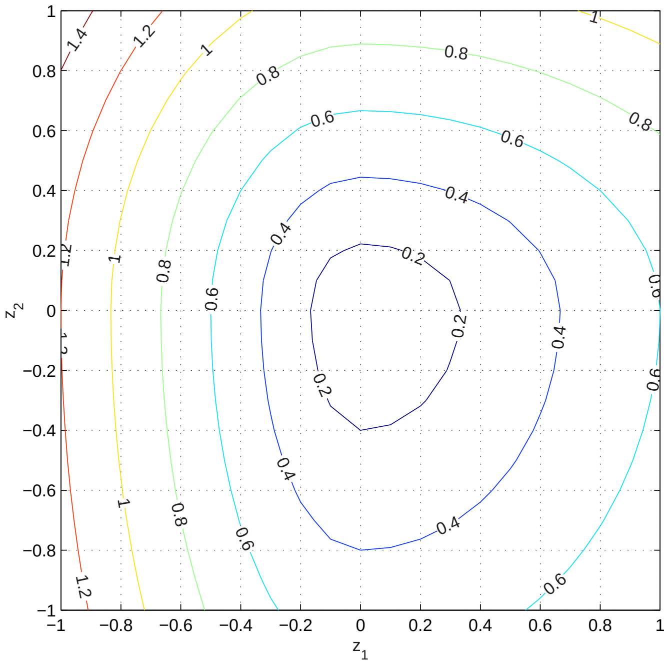

Figure 1: An uncertainty set represented by $\mathcal{G}_\Omega$ as Ω varies for $N = 2$.

(a) *is equivalent to the set of feasible solutions* $(v_0, ..., v_N)$ *satisfying the following set of constraints*

$$\left\{\begin{array}{ll} v_0 \geq \Omega u_0 + \boldsymbol{s}'\underline{\boldsymbol{z}} + \boldsymbol{t}'\overline{\boldsymbol{z}} & \\ u_k \geq q_k(v_k - s_k + t_k) & \forall k \in \{1,\dots,N\} \\ u_k \geq -p_k(v_k - s_k + t_k) & \forall k \in \{1,\dots,N\} \\ \|\boldsymbol{u}\|_2 \leq u_0 & \\ u_0 \geq 0, \boldsymbol{u} \in \Re^N, \boldsymbol{s}, \boldsymbol{t} \in \Re^N_+. & \end{array}\right\} \tag{19}$$

(b) *Moreover,*

$$\mathrm{P}\left(v(\boldsymbol{z}) < 0\right) \leq \exp(-\Omega^2/2)$$

It is easy to see that even for high reliability factor, say violation probability of less than 10^{-6}, the budget of uncertainty Ω is practically a small constant of 5.3. In contrast, Chen, Sim and Sun [16] show that the worst case budget $\Omega_{\max}$ is at least $\sqrt{N}$ in order for $\mathcal{G}_{\Omega_{\max}}$ to be as conservative as the worst case uncertainty set $\mathcal{W}$. Hence, the benefit of robust optimization increases with the number of primitive uncertainties in the model, which is essentially achieving the effect of risk pooling.

4.2 Bound on objective function

Now we deal with the objective function. Unfortunately, we are not aware of any tractable way of evaluating $\mathrm{E}\left(r_i(\tilde{\boldsymbol{z}})^-\right)$ exactly. Moreover, since the distributions of the primitive uncertainties are not fully specified, it would not be possible to evaluate its exact value. Given a random variable $\tilde{r}$ with mean μ and standard deviation, σ, a well known tight bound is,

$$\mathrm{E}(\tilde{r}^-) \le \frac{1}{2}\left(-\mu + \sqrt{\mu^2+\sigma^2}\right) \tag{20}$$

(see Scarf [26] for instance). Therefore, suppose $y(\cdot) \in \mathcal{L}$, we have

$$\mathrm{E}\left(y(\tilde{\boldsymbol{z}})^-\right) \le \frac{1}{2}\left(-y_0 + \sqrt{y_0^2 + \|\Sigma^{1/2}\boldsymbol{y}\|_2^2}\right),$$

where Σ is the covariance matrix of $\tilde{\boldsymbol{z}}$ and $\boldsymbol{y} = (y_1, \ldots, y_N)$. The bound does not take into account the distributional support, which could degrade the quality of the approximation. For instance, if $y(\tilde{\boldsymbol{z}}) \ge 0$, it follows trivially that $\mathrm{E}(y(\tilde{\boldsymbol{z}})^-) = 0$. Likewise, if $y(\tilde{\boldsymbol{z}}) \le 0$, we have $\mathrm{E}(y(\tilde{\boldsymbol{z}})^-) = -y_0$. Under these circumstances, the bound would be weak. Hence, we propose the following tighter bound that resolves these issues while still preserving the benefits of being second order cone representable.

Theorem 3 *Let $\tilde{\boldsymbol{z}} \in \Re^N$ be a vector of zero mean random variables with covariance matrix Σ and support in $\mathcal{W} = [-\underline{\boldsymbol{z}}, \bar{\boldsymbol{z}}]$.*

(a)

$$\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^-\right) \le h(y_0, \boldsymbol{y})$$

where

$$h(y_0, \boldsymbol{y}) = \min_{\boldsymbol{s},\boldsymbol{t},\boldsymbol{u},\boldsymbol{v} \ge \boldsymbol{0}} \left\{ \frac{1}{2}\Bigg(-y_0 + (\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + \sqrt{(-y_0 + (\boldsymbol{s}-\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}-\boldsymbol{v})'\underline{\boldsymbol{z}})^2 + \|\Sigma^{1/2}(-\boldsymbol{y} - \boldsymbol{s} + \boldsymbol{t} + \boldsymbol{u} - \boldsymbol{v})\|_2^2}\Bigg)\right\}.$$

(b) *Moreover,*

$$h(y_0, \boldsymbol{y}) \le \frac{1}{2}\left(-y_0 + \sqrt{y_0^2 + \|\Sigma^{1/2}\boldsymbol{y}\|_2^2}\right).$$

(c) *Suppose*

$$y(\boldsymbol{z}) \le 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}$$

then

$$\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^-\right) = h(y_0, \boldsymbol{y}) = -y_0.$$

Likewise, if

$$y(\boldsymbol{z}) \geq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W},$$

then

$$\mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^-\right) = h(y_0, \boldsymbol{y}) = 0.$$

Proof : (a) Since $-\underline{\boldsymbol{z}} \leq \tilde{\boldsymbol{z}} \leq \bar{\boldsymbol{z}}$, we observe that

$$(\bar{\boldsymbol{z}} - \tilde{\boldsymbol{z}})'\boldsymbol{s} \geq 0$$
$$(\underline{\boldsymbol{z}} + \tilde{\boldsymbol{z}})'\boldsymbol{t} \geq 0$$
$$(\bar{\boldsymbol{z}} - \tilde{\boldsymbol{z}})'\boldsymbol{u} \geq 0$$
$$(\underline{\boldsymbol{z}} + \tilde{\boldsymbol{z}})'\boldsymbol{v} \geq 0$$

for all $\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{u}, \boldsymbol{v} \geq \boldsymbol{0}$. Therefore,

$$\begin{aligned}
& \mathrm{E}((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^-) \\
\leq\ & \mathrm{E}((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}} - (\underline{\boldsymbol{z}} + \tilde{\boldsymbol{z}})'\boldsymbol{t} - (\bar{\boldsymbol{z}} - \tilde{\boldsymbol{z}})'\boldsymbol{s})^-) \\
=\ & \mathrm{E}((y_0 - (-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s})'\tilde{\boldsymbol{z}} - \underline{\boldsymbol{z}}'\boldsymbol{t} - \bar{\boldsymbol{z}}'\boldsymbol{s})^-) \\
=\ & \mathrm{E}\left(-y_0 + (-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s})'\tilde{\boldsymbol{z}} + \underline{\boldsymbol{z}}'\boldsymbol{t} + \bar{\boldsymbol{z}}'\boldsymbol{s} + (y_0 - (-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s})'\tilde{\boldsymbol{z}} - \underline{\boldsymbol{z}}'\boldsymbol{t} - \bar{\boldsymbol{z}}'\boldsymbol{s})^+\right) && (21) \\
=\ & \mathrm{E}\left(-y_0 + \underline{\boldsymbol{z}}'\boldsymbol{t} + \bar{\boldsymbol{z}}'\boldsymbol{s} + (y_0 - (-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s})'\tilde{\boldsymbol{z}} - \underline{\boldsymbol{z}}'\boldsymbol{t} - \bar{\boldsymbol{z}}'\boldsymbol{s})^+\right) \\
\leq\ & \mathrm{E}\left(-y_0 + \underline{\boldsymbol{z}}'\boldsymbol{t} + \bar{\boldsymbol{z}}'\boldsymbol{s} + (y_0 - (-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s})'\tilde{\boldsymbol{z}} - \underline{\boldsymbol{z}}'\boldsymbol{t} - \bar{\boldsymbol{z}}'\boldsymbol{s} + (\underline{\boldsymbol{z}} + \tilde{\boldsymbol{z}})'\boldsymbol{v} + (\bar{\boldsymbol{z}} - \tilde{\boldsymbol{z}})'\boldsymbol{u})^+\right) \\
=\ & \mathrm{E}\left(-y_0 + \underline{\boldsymbol{z}}'\boldsymbol{t} + \bar{\boldsymbol{z}}'\boldsymbol{s} + (y_0 - (-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s} + \boldsymbol{u} - \boldsymbol{v})'\tilde{\boldsymbol{z}} - \underline{\boldsymbol{z}}'(\boldsymbol{t} - \boldsymbol{v}) - \bar{\boldsymbol{z}}'(\boldsymbol{s} - \boldsymbol{u}))^+\right) \\
=\ & \underline{\boldsymbol{z}}'\boldsymbol{v} + \bar{\boldsymbol{z}}'\boldsymbol{u} + \mathrm{E}\left((-y_0 + (-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s} + \boldsymbol{u} - \boldsymbol{v})'\tilde{\boldsymbol{z}} + \underline{\boldsymbol{z}}'(\boldsymbol{t} - \boldsymbol{v}) + \bar{\boldsymbol{z}}'(\boldsymbol{s} - \boldsymbol{u}))^+\right) && (22) \\
=\ & \underline{\boldsymbol{z}}'\boldsymbol{v} + \bar{\boldsymbol{z}}'\boldsymbol{u} + \mathrm{E}\left((y_0 - (-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s} + \boldsymbol{u} - \boldsymbol{v})'\tilde{\boldsymbol{z}} - \underline{\boldsymbol{z}}'(\boldsymbol{t} - \boldsymbol{v}) - \bar{\boldsymbol{z}}'(\boldsymbol{s} - \boldsymbol{u}))^-\right) \\
\leq\ & \underline{\boldsymbol{z}}'\boldsymbol{v} + \bar{\boldsymbol{z}}'\boldsymbol{u} + \frac{1}{2}\Bigg(- y_0 + \underline{\boldsymbol{z}}'(\boldsymbol{t} - \boldsymbol{v}) + \bar{\boldsymbol{z}}'(\boldsymbol{s} - \boldsymbol{u})) + \\
& \qquad \sqrt{(-y_0 + \underline{\boldsymbol{z}}'(\boldsymbol{t} - \boldsymbol{v}) + \bar{\boldsymbol{z}}'(\boldsymbol{s} - \boldsymbol{u}))^2 + \|\Sigma^{1/2}(-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s} + \boldsymbol{u} - \boldsymbol{v})\|_2^2}\Bigg) && (23) \\
=\ & \frac{1}{2}\Bigg(- y_0 + \underline{\boldsymbol{z}}'(\boldsymbol{t} + \boldsymbol{v}) + \bar{\boldsymbol{z}}'(\boldsymbol{s} + \boldsymbol{u})) + \\
& \qquad \sqrt{(-y_0 + \underline{\boldsymbol{z}}'(\boldsymbol{t} - \boldsymbol{v}) + \bar{\boldsymbol{z}}'(\boldsymbol{s} - \boldsymbol{u}))^2 + \|\Sigma^{1/2}(-\boldsymbol{y} + \boldsymbol{t} - \boldsymbol{s} + \boldsymbol{u} - \boldsymbol{v})\|_2^2}\Bigg)
\end{aligned}$$

where the equalities of (21) and (22) follows from the fact that $x^+ = x + \underbrace{(-x)^+}_{=x^-}$. The inequality (23) is due to the bound (20).

(b) Note that with $\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{u}, \boldsymbol{v} = \mathbf{0}$, we have

$$\frac{1}{2}\left(- y_0 + (\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + \sqrt{(-y_0 + (\boldsymbol{s}-\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}-\boldsymbol{v})'\underline{\boldsymbol{z}})^2 + \|\Sigma^{1/2}(-\boldsymbol{y}-\boldsymbol{s}+\boldsymbol{t}+\boldsymbol{u}-\boldsymbol{v})\|_2^2}\right)$$
$$= \frac{1}{2}\left(-y_0 + \sqrt{y_0^2 + \|\Sigma^{1/2}\boldsymbol{y}\|_2^2}\right)$$

Therefore,

$$h(y_0, \boldsymbol{y}) \le \frac{1}{2}\left(-y_0 + \sqrt{y_0^2 + \|\Sigma^{1/2}\boldsymbol{y}\|_2^2}\right).$$

(c) Suppose

$$y_0 + \boldsymbol{y}'\boldsymbol{z} \le 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}$$

then let $\boldsymbol{s} = \boldsymbol{t} = \mathbf{0}$, $u_k = (y_k)^+$, $v_k = (-y_k)^+$ for $k = 1, \ldots, N$ and

$$z_k^* = \begin{cases} \bar{z}_k & \text{if } y_k > 0 \\ -\underline{z}_k & \text{otherwise} \end{cases}$$

Since $\boldsymbol{z}^* \in \mathcal{W}$, we have $y_0 + \boldsymbol{y}'\boldsymbol{z}^* \le 0$. Furthermore, it is easy to verify that

$$\boldsymbol{y} = \boldsymbol{u} - \boldsymbol{v}$$

and

$$y_0 + \boldsymbol{u}'\bar{\boldsymbol{z}} + \boldsymbol{v}'\underline{\boldsymbol{z}} = y_0 + \boldsymbol{y}'\boldsymbol{z}^* \le 0.$$

We have

$$\frac{1}{2}\Bigg(- y_0 + (\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + \sqrt{(-y_0 + (\boldsymbol{s}-\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}-\boldsymbol{v})'\underline{\boldsymbol{z}})^2 + \|\Sigma^{1/2}(-\boldsymbol{y}-\boldsymbol{s}+\boldsymbol{t}+\boldsymbol{u}-\boldsymbol{v})\|_2^2}\Bigg) = -y_0.$$

Hence,

$$-y_0 = \mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^-\right) \le h(y_0, \boldsymbol{y}) \le -y_0.$$

Similarly, if

$$y_0 + \boldsymbol{y}'\boldsymbol{z} \ge 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}$$

then let $\boldsymbol{v} = \boldsymbol{u} = \mathbf{0}$, $s_k = (-y_k)^+$, $t_k = (y_k)^+$ for $k = 1, \ldots, N$ and

$$z_k^* = \begin{cases} \bar{z}_k & \text{if } y_k < 0 \\ -\underline{z}_k & \text{otherwise} \end{cases}$$

Since $\boldsymbol{z}^* \in \mathcal{W}$, we have $y_0 + \boldsymbol{y}'\boldsymbol{z}^* \geq 0$. Furthermore, it is easy to verify that

$$\boldsymbol{y} = \boldsymbol{t} - \boldsymbol{s}$$

and

$$y_0 - \boldsymbol{s}'\bar{\boldsymbol{z}} - \boldsymbol{t}'\underline{\boldsymbol{z}} = y_0 + \boldsymbol{y}'\boldsymbol{z}^* \geq 0.$$

Hence, we have

$$\begin{aligned}
& \frac{1}{2}\Bigg(- y_0 + (\boldsymbol{s}+\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}+\boldsymbol{v})'\underline{\boldsymbol{z}} + \\
& \qquad \sqrt{(-y_0 + (\boldsymbol{s}-\boldsymbol{u})'\bar{\boldsymbol{z}} + (\boldsymbol{t}-\boldsymbol{v})'\underline{\boldsymbol{z}})^2 + \|\Sigma^{1/2}(-\boldsymbol{y} - \boldsymbol{s} + \boldsymbol{t} + \boldsymbol{u} - \boldsymbol{v})\|_2^2}\Bigg) \\
= \; & \frac{1}{2}\Bigg(- y_0 + \boldsymbol{s}'\bar{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}} + | - y_0 + \boldsymbol{s}'\bar{\boldsymbol{z}} + \boldsymbol{t}'\underline{\boldsymbol{z}}|\Bigg) \\
= \; & 0.
\end{aligned}$$

Therefore,

$$0 = \mathrm{E}\left((y_0 + \boldsymbol{y}'\tilde{\boldsymbol{z}})^-\right) \leq h(y_0, \boldsymbol{y}) \leq 0.$$

■

5 The Second-Order Cone Programming Model

Finally, we put the pieces together and propose the approximation for problem (5). Recall that in Section 3, we proposed two approximations of problem (5); using linear decision rule in problem (8) and using deflected linear decision rule in problem (13), both remain computationally challenging. Using the techniques of Theorem 2 to approximate the chance constraints, we can formulate a tractable approximation of Model (8) as follows:

$$\begin{array}{lll}
Z_{LDR^*} = \min & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{d}'\boldsymbol{v}^0 + \boldsymbol{f}'\boldsymbol{w}^0 & \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& \boldsymbol{U}^k\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}^k + \boldsymbol{W}\boldsymbol{w}^k = \boldsymbol{h}^k & \forall k \in \{0, \ldots, N\} \\
& v_j(\tilde{\boldsymbol{z}}) \geq 0 & \forall \boldsymbol{z} \in \mathcal{G}_{\Omega_i}, \forall j \in \{1, \ldots, n_2\} \\
& w_i(\boldsymbol{z}) \geq 0 & \forall \boldsymbol{z} \in \mathcal{W}, \forall i \in I \subseteq \{1, \ldots, n_3\} \\
& \boldsymbol{x} \geq 0 & \\
& \boldsymbol{v}(\cdot), \boldsymbol{w}(\cdot) \in \mathcal{L}, &
\end{array} \qquad (24)$$

where $\Omega_j = \sqrt{-2\ln(\epsilon_j)}$. In addition, using Theorem 3 presented in Section 4, we can also approximate the objective function and obtain the following tractable approximation of Model (13):

$$
\begin{array}{lll}
Z_{DLDR^*} = \min & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{d}'\boldsymbol{v}^0 + \boldsymbol{f}'\boldsymbol{r}^0 + \bar{\boldsymbol{f}}'\boldsymbol{g} & \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& \boldsymbol{U}^k\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}^k + \boldsymbol{W}\boldsymbol{r}^k = \boldsymbol{h}^k & \forall k \in \{0,\ldots,N\} \\
& v_j(\boldsymbol{z}) \geq 0 & \forall \boldsymbol{z} \in \mathcal{G}_{\Omega_i}, \forall j \in \{1,\ldots,n_2\} \\
& g_i \geq h(r_i^0, (r_i^1,\ldots,r_i^N)) & \forall i \in I \\
& \boldsymbol{x} \geq 0 & \\
& \boldsymbol{r}(\cdot), \boldsymbol{v}(\cdot) \in \mathcal{L}. &
\end{array}
\qquad (25)
$$

More explicitly, we have the following second order cone optimization formulation

$$
\begin{array}{lll}
Z_{DLDR^*} = \min & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{d}'\boldsymbol{v}^0 + \boldsymbol{f}'\boldsymbol{r}^0 + \bar{\boldsymbol{f}}'\boldsymbol{g} & \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & \\
& \boldsymbol{U}^k\boldsymbol{x} + \boldsymbol{V}\boldsymbol{v}^k + \boldsymbol{W}\boldsymbol{r}^k = \boldsymbol{h}^k & \forall k \in \{0,\ldots,N\} \\
& v_j^0 \geq \Omega_i u_0^{1j} + \boldsymbol{t}^{1j\prime}\bar{\boldsymbol{z}} + \boldsymbol{s}^{1j\prime}\underline{\boldsymbol{z}} & \forall j \in \{1,\ldots,n_2\} \\
& u_k^{1j} \geq q_k(v_j^k - s_k^{1j} + t_k^{1j}) & \forall j \in \{1,\ldots,n_2\}, k \in \{1,\ldots,N\} \\
& u_k^{1j} \geq -p_k(v_j^k - s_k^{1j} + t_k^{1j}) & \forall j \in \{1,\ldots,n_2\}, k \in \{1,\ldots,N\} \\
& \|\boldsymbol{u}^{1j}\|_2 \leq u_0^{1j} & \forall j \in \{1,\ldots,n_2\} \\
& \boldsymbol{t}^{1j}, \boldsymbol{s}^{1j} \geq \boldsymbol{0} & \forall j \in \{1,\ldots,n_2\} \\
& 2g_i + r_i^0 - (\boldsymbol{s}^{2i} + \boldsymbol{u}^{2i})'\bar{\boldsymbol{z}} - (\boldsymbol{t}^{2i} + \boldsymbol{v}^{2i})'\underline{\boldsymbol{z}} \geq & \\
& \quad \Big((-r_i^0 + (\boldsymbol{s}^{2i} - \boldsymbol{u}^{2i})'\bar{\boldsymbol{z}} + (\boldsymbol{t}^{2i} - \boldsymbol{v}^{2i})'\underline{\boldsymbol{z}})^2 + & \\
& \qquad \|\Sigma^{1/2}(-\boldsymbol{r}_i - \boldsymbol{s}^{2i} + \boldsymbol{t}^{2i} + \boldsymbol{u}^{2i} - \boldsymbol{v}^{2i})\|_2^2 \Big)^{\frac{1}{2}} & \forall i \in I \\
& \boldsymbol{s}^{2i}, \boldsymbol{t}^{2i}, \boldsymbol{u}^{2i}, \boldsymbol{v}^{2i} \geq \boldsymbol{0} & \forall i \in I \\
& \boldsymbol{x} \geq \boldsymbol{0}, &
\end{array}
\qquad (26)
$$

where $\boldsymbol{r}_i = (r_i^1, \ldots, r_i^N)$.

Notice that as input to the above formulation, the distributional information of the uncertainty $\tilde{\boldsymbol{z}}$ includes the forward and backward deviation measures p_i and q_i, as well as the covariance matrix Σ.

Theorem 4 *Under semi-complete recourse, deflected linear decision rule gives an objective value that is at least as good as linear decision rule, that is,* $Z_{DLDR^*} \leq Z_{LDR^*}$.

Proof : Given any feasible solution, $(\boldsymbol{x}, \boldsymbol{v}(\tilde{\boldsymbol{z}}), \boldsymbol{w}(\tilde{\boldsymbol{z}}))$ of problem (24), we observe that

$$w_i(\boldsymbol{z}) \geq 0 \qquad \forall \boldsymbol{z} \in \mathcal{W}.$$

Hence, from Theorem 3 (c), we have

$$h(w_i^0, (w_i^1, \ldots, w_i^N)) = 0.$$

Therefore, by letting $\boldsymbol{r}(\tilde{\boldsymbol{z}}) = \boldsymbol{w}(\tilde{\boldsymbol{z}})$, we obtain a feasible solution of Model (25) that yields the same objective as problem (24). Hence, $Z_{DLDR^*} \leq Z_{LDR^*}$. ■

We give a simple illustration of the modeling steps.

Example 5: Newsvendor problem

We can model the newsvendor problem as a stochastic programming problem (see, for instance, Birge and Louveaux [14]). A single retailer faces random demand. The retailer places an order to an outside supplier before knowing the actual demand. Per unit ordering cost is c and the selling price to the customers is $p > c$. Let the demand be $d(\tilde{z}) = \mu + \tilde{z}$, in which $\tilde{z}$ has zero mean, standard deviation σ and unknown support. Scarf [26] first studied such a model with ambiguity, that is, not knowing the exact demand distribution. For simplicity of exposition, we assume that unsatisfied demand is lost and leftover inventory has zero salvage value. Finally, we assume that the retailer's objective is to maximize the expected profit (or minimize expected negative profit). Let x denote the ordering quantity and $\tilde{d}$ denote the random demand. The stochastic optimization formulation of the newsvendor model is as follows:

$$\begin{array}{rl} \min & cx + p\mathrm{E}(w(\tilde{z})) \\ \text{s.t.} & w(\tilde{z}) \geq -x \\ & w(\tilde{z}) \geq -d(\tilde{z}) \\ & x \geq 0 \\ & w(\cdot) \in \mathcal{Y}, \end{array}$$

in which function $-w(\tilde{z})$ represents the amount sold after observing the demand $\tilde{z}$. Equivalently,

$$\begin{array}{rl} \min & cx + p\mathrm{E}(w_3(\tilde{z})) \\ \text{s.t.} & x + w_3(\tilde{z}) - w_1(\tilde{z}) = 0 \\ & w_3(\tilde{z}) - w_2(\tilde{z}) = -d(\tilde{z}) \\ & w_1(\tilde{z}), w_2(\tilde{z}) \geq 0 \\ & x \geq 0 \\ & w_1(\cdot), w_2(\cdot), w_3(\cdot) \in \mathcal{Y}. \end{array}$$

It is obvious that the associated recourse matrix

$$\boldsymbol{W} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix},$$

satisfies semi-complete recourse and that $\bar{f}_1 = \bar{f}_2 = p$. Using the approach of Model (25), we solve the following problem

$$\begin{aligned} \min \quad & cx + p(w_3^0 + g_1 + g_2) \\ \text{s.t.} \quad & x + w_3^0 - r_1^0 = 0 \\ & w_3^1 - r_1^1 = 0 \\ & w_3^0 - r_2^0 = -\mu \\ & w_3^1 - r_2^1 = -1 \\ & 2g_1 \geq -r_1^0 + \sqrt{(r_1^0)^2 + (\sigma r_1^1)^2} \\ & 2g_2 \geq -r_2^0 + \sqrt{(r_2^0)^2 + (\sigma r_2^1)^2} \\ & x \geq 0. \end{aligned}$$

We can show that the solution is identical to the famous result in Scarf [26]. To be specific, after simplification we have

$$\begin{aligned} \min \quad & cx + \tfrac{1}{2}p\left((-x-\mu) + \sqrt{(x + w_3^0)^2 + (\sigma w_3^1)^2} + \sqrt{(\mu + w_3^0)^2 + (\sigma(w_3^1 + 1))^2}\right) \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

From triangle inequality, we have

$$\sqrt{(x + w_3^0)^2 + (\sigma w_3^1)^2} + \sqrt{(\mu + w_3^0)^2 + (\sigma(w_3^1 + 1))^2} \geq \sqrt{(x-\mu)^2 + \sigma^2}.$$

The equality is achieved at $w_3^0 = -\mu$ and $w_3^1 = -1$. Therefore we have

$$\begin{aligned} \min \quad & cx + \tfrac{1}{2}p\left((-x-\mu) + \sqrt{(x-\mu)^2 + \sigma^2}\right) \\ \text{s.t.} \quad & x \geq 0, \end{aligned}$$

which yields the same min-max solution of the newsvendor problem studied by Scarf [26].

5.1 Multiperiod modeling

Our approach can be easily extended to deal with multiperiod stochastic programming problems as follows.

$$
\begin{array}{ll}
\min & \boldsymbol{c}'\boldsymbol{x} + \sum_{t=1}^{T}(\mathrm{E}(\boldsymbol{d}_t'\boldsymbol{v}_t(\tilde{\boldsymbol{\xi}}_t)) + \mathrm{E}(\boldsymbol{f}_t'\boldsymbol{w}_t(\tilde{\boldsymbol{\xi}}_t))) \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \geq 0 \\
& \boldsymbol{U}_t(\tilde{\boldsymbol{\xi}}_t)\boldsymbol{x} + \sum_{\tau=1}^{t}(\boldsymbol{V}_{t\tau}\boldsymbol{v}_\tau(\tilde{\boldsymbol{\xi}}_\tau) + \boldsymbol{W}_{t\tau}\boldsymbol{w}_\tau(\tilde{\boldsymbol{\xi}}_\tau)) = \boldsymbol{b}_t(\tilde{\boldsymbol{\xi}}_t), t = 1, 2, \ldots, T \\
& \mathrm{P}(v_{ti}(\tilde{\boldsymbol{\xi}}_t) \geq 0) \geq 1 - \epsilon_{ti}, \forall\ t, i \\
& w_{ti}(\tilde{\boldsymbol{\xi}}_t) \geq 0, \forall\ i \in I_t,
\end{array}
\tag{27}
$$

where $\tilde{\boldsymbol{\xi}}_t = (\tilde{\boldsymbol{z}}_1, .., \tilde{\boldsymbol{z}}_t)$, and the underlying uncertainties, $\tilde{\boldsymbol{z}}_1 \in \Re^{N_1}, \ldots, \tilde{\boldsymbol{z}}_T \in \Re^{N_T}$, unfold progressively from the first period to the last period and $\tilde{\boldsymbol{z}}_t$ is the vector of primitive uncertainties that is only available at the t^{th} period. We also assume that $\boldsymbol{U}_t(\tilde{\boldsymbol{\xi}}_t)$ and $\boldsymbol{b}_t(\tilde{\boldsymbol{\xi}}_t)$ are affine in $\tilde{\boldsymbol{\xi}}_t$.

There are a variety of ways to model chance constraints in a multi-stage stochastic programming problem. In the above model we take the unconditional probability from the perspective of the beginning of the planning horizon.

The multiperiod stochastic programming problem is said to have semi-complete recourse if for any t, there exists a vector $\boldsymbol{r}_t$ with $r_{ti} > 0$ for all $i \in I_t$ such that

$$\sum_{k=t}^{\tau} \boldsymbol{W}_{\tau k}\boldsymbol{r}_k = 0, \quad \tau = t, \ldots, T,$$

for some $\boldsymbol{r}_{t+1}, \ldots, \boldsymbol{r}_T$ satisfying $r_{ki} \geq 0$, $i \in I_k$, $k = t+1, \ldots, T$. Therefore, similar to problem (10), we can define

$$
\begin{array}{lll}
\bar{f}_t^i = \min & \sum_{\tau=t}^{T} \boldsymbol{f}'_\tau \boldsymbol{p}_{t\tau}^i & \\
\text{s.t.} & \sum_{k=t}^{\tau} \boldsymbol{W}_{\tau k}\boldsymbol{p}_{tk}^i = 0 & \forall \tau = t, \ldots, T \\
& p_{tt}^i = 1 & \\
& \boldsymbol{p}_{t\tau}^i \geq 0 & \forall \tau = t, \ldots, T,
\end{array}
\tag{28}
$$

which is feasible. Without loss of generality, we assume that the optimal solution $(\bar{\boldsymbol{p}}_{tt}^i, \ldots, \bar{\boldsymbol{p}}_{tT}^i)$ exists, denoted as for any t.

Similar to the two stage model, the deflected linear decision rule can be defined as follows.

$$\boldsymbol{v}_t(\tilde{\boldsymbol{\xi}}_t) = \boldsymbol{v}_t^0 + \sum_{\tau=1}^{t}\sum_{j=1}^{N_\tau} \boldsymbol{v}_\tau^j z_\tau^j$$
$$\boldsymbol{r}_t(\tilde{\boldsymbol{\xi}}_t) = \boldsymbol{r}_t^0 + \sum_{\tau=1}^{t}\sum_{j=1}^{N_\tau} \boldsymbol{r}_\tau^j z_\tau^j$$
$$\boldsymbol{w}_t(\tilde{\boldsymbol{\xi}}_t) = \boldsymbol{r}_t(\tilde{\boldsymbol{\xi}}_t) + \sum_{k=1}^{t}\sum_{i\in I_k} (r_{ki}(\tilde{\boldsymbol{\xi}}_k))^- \bar{\boldsymbol{p}}_{tk}^i.$$

Observe that the above decision rule fulfills the nonanticipativity requirement. Essentially, we end up with a formulation for the multi-period model similar to the one for the two-period model we have presented. Using the deflected linear decision rule, the size of the model does not explode exponentially with the number of periods. Such decision rules can be viewed as a first order approximation of the expected future costs, so that we can determine the first stage or 'here-and-now' decision. In practice, we do not use the decision rule as the responding actions in subsequent stages. Instead, we adopt the rolling horizon approach, that is, we resolve subsequent stages upon realizations of uncertainties at earlier stages.

6 Computational Experiment

In this section, we illustrate our approach in a preliminary computational experiment. Since classical stochastic optimization does not handle chance constraints with recourse, we restrict the comparison to a two stage stochastic optimization problem with complete recourse. Specifically, on the two stage problem, we demonstrate that when compared with sampling based stochastic programming approach, our proposed framework achieves similar performance in objective values. In contrast, the size of our model increases polynomially with the number of stages in the model, while the sample sizes according to traditional stochastic programming approaches may increase exponentially (see, e.g., Shapiro and Nemirovski [27], and Dyer and Stougie [19]). Therefore, we believe our proposed model is promising in addressing large-scale multiperiod stochastic optimization models.

In our experiment, we consider a project management example of several activities. Each activity has random completion time, $\tilde{t}_{ij}$. The completion of activities must adhere to precedent constraints. For instance, activity e_1 precedes activity e_2 if activity e_1 must be completed before activity e_2. In our computational experiments, we assume that the random completion time $t_{ij}(\tilde{\boldsymbol{z}})$ is independent of the completion times of other activities. In addition, the completion time also depends on the additional amount of resource, $x_{ij} \in [0,1]$, committed on the activity as follows

$$\tilde{t}_{ij} = b_{ij} + a_{ij}(1 - x_{ij})\tilde{z}_{ij}, \tag{29}$$

where $\tilde{z}_{ij} \in [-\underline{z}_{ij}, \bar{z}_{ij}]$, $(i,j) \in \mathcal{E}$ are independent random variables with zero means and standard deviations σ_{ij}.

Denote c_{ij} to be the cost of using each unit of resource for the activity on the arc (i,j). Our goal is to minimize the expected completion time subject to the constraint that the total resource is under a budget C.

A stochastic programming model to address the above project management problem is as follows.

$$\begin{aligned}
\min \quad & \mathrm{E}(y_n(\tilde{\boldsymbol{z}})) \\
\text{s.t.} \quad & \boldsymbol{c}'\boldsymbol{x} \le C \\
& y_j(\tilde{\boldsymbol{z}}) - y_i(\tilde{\boldsymbol{z}}) - w_{ij}(\tilde{\boldsymbol{z}}) = b_{ij} + a_{ij}(1 - x_{ij})\tilde{z}_{ij} \quad \forall (i,j) \in \mathcal{E} \\
& y_1(\tilde{\boldsymbol{z}}) = 0 \\
& w_{ij}(\tilde{\boldsymbol{z}}) \ge 0 \qquad \forall (i,j) \in \mathcal{E} \\
& \boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{1} \\
& \boldsymbol{x} \in \Re^{|\mathcal{E}|} \\
& \boldsymbol{w}(\cdot), \boldsymbol{y}(\cdot) \in \mathcal{Y}.
\end{aligned} \tag{30}$$

To check semi-complete recourse and model the deflected linear decision rules, we first take a look at the following problem

$$\begin{aligned}
\bar{f}_{kl} = \min \quad & y_n \\
\text{s.t.} \quad & y_j - y_i - w_{ij} = 0 \quad \forall (i,j) \in \mathcal{E} \\
& y_1 = 0 \\
& w_{kl} = 1 \\
& \boldsymbol{w} \ge \boldsymbol{0},
\end{aligned}$$

for each $(k,l) \in \mathcal{E}$. Here, $\bar{f}_{kl}$ corresponds to the longest duration path from node 1 to node n, when one of the arc (k,l) has unit time while the rest of the arcs have zero completion time. Clearly, $\bar{f}_{kl} = 1$ for all $(k,l) \in \mathcal{E}$.

We compare solutions from two approaches. The first is stochastic optimization using sampling approximation as follows:

$$
\begin{aligned}
Z_1(K) = \min \quad & \frac{1}{K}\sum_{k=1}^{K} y_n^k \\
\text{s.t.} \quad & \boldsymbol{c}'\boldsymbol{x} \le C \\
& y_j^k - y_i^k \ge b_{ij} + a_{ij}(1 - x_{ij})z_{ij}^k && \forall (i,j) \in \mathcal{E}, k \in \{1,\ldots,K\} \\
& y_1^k = 0 && k \in \{1,\ldots,K\} \\
& \boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{1} \\
& \boldsymbol{x} \in \Re^{|\mathcal{E}|} \\
& \boldsymbol{y}^k \in \Re^n && \forall k \in \{1,\ldots,K\},
\end{aligned}
\tag{31}
$$

where z_{ij}^k, $k = 1, \ldots, K$ are independent samples of $\tilde{z}_{ij}$.

For the second method, we adopt the framework of (26) as follows:

$$
\begin{aligned}
Z_2 = \min \quad & y_n^0 + \sum_{e\in\mathcal{E}} g_e \\
\text{s.t.} \quad & \boldsymbol{c}'\boldsymbol{x} \le C \\
& y_j^0 - y_i^0 - r_{(i,j)}^0 = b_{ij} && \forall (i,j) \in \mathcal{E} \\
& y_j^{(i,j)} - y_i^{(i,j)} - r_{(i,j)}^{(i,j)} = a_{ij}(1 - x_{ij}) && \forall (i,j) \in \mathcal{E} \\
& y_j^e - y_i^e - r_{(i,j)}^e = 0 && \forall e, (i,j) \in \mathcal{E}, e \ne (i,j) \\
& y_1^e = 0 && \forall e \in \mathcal{E} \\
& 2g_e + r_e^0 - (\boldsymbol{s}^e + \boldsymbol{u}^e)'\overline{\boldsymbol{z}} - (\boldsymbol{t}^e + \boldsymbol{v}^e)'\underline{\boldsymbol{z}} \ge \\
& \quad \Big((-r_e^0 + (\boldsymbol{s}^e - \boldsymbol{u}^e)'\overline{\boldsymbol{z}} + (\boldsymbol{t}^e - \boldsymbol{v}^e)'\underline{\boldsymbol{z}})^2 + \\
& \qquad \|\Sigma^{1/2}(-\boldsymbol{r}_e - \boldsymbol{s}^e + \boldsymbol{t}^e + \boldsymbol{u}^e - \boldsymbol{v}^e)\|_2^2 \Big)^{\frac{1}{2}} && \forall e \in \mathcal{E} \\
& \boldsymbol{s}^e, \boldsymbol{t}^e, \boldsymbol{u}^e, \boldsymbol{v}^e \ge \boldsymbol{0} && \forall e \in \mathcal{E} \\
& \boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{1} \\
& \boldsymbol{x} \in \Re^{|\mathcal{E}|} \\
& \boldsymbol{y}^0, \boldsymbol{y}^e \in \Re^n && \forall e \in \mathcal{E} \\
& \boldsymbol{r}^0, \boldsymbol{r}^e \in \Re^{|\mathcal{E}|} && \forall e \in \mathcal{E} \\
& \boldsymbol{s}^e, \boldsymbol{t}^e, \boldsymbol{u}^e, \boldsymbol{v}^e \in \Re^{|\mathcal{E}|},
\end{aligned}
\tag{32}
$$

where $\boldsymbol{r}_e = (r_e^a : \forall a \in \mathcal{E})$.

For our computational experiment, we create a fictitious project with the activity network in the form of H by W grid (see Figure 2). There are a total of $H \times W$ nodes, with the first node at the

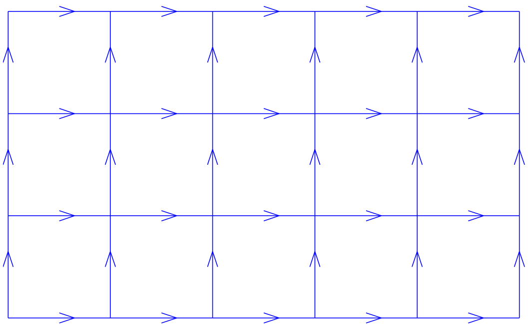

Figure 2: Project management "grid" with height, $H = 4$ and width $W = 6$.

left bottom corner and the last node at the right upper corner. Each arc on the graph proceeds either towards the right node or the upper node. We assume that every activity on arc has independent and identical completion time. In particular, for all arcs (i, j),

$$\mathrm{P}(\tilde{z}_{ij} = z) = \begin{cases} \beta & \text{if } z = \frac{1}{2\beta} \\ (1-\beta) & \text{if } z = -\frac{1}{2(1-\beta)}. \end{cases}$$

The parameter β controls the variance of the activity completion time, which increases to ∞ as β decreases to zero. We set $b_{ij} = 3, a_{ij} = 3$ for all $(i, j) \in \mathcal{E}$ and $\boldsymbol{c} = \mathbf{1}$. The project grid is fixed to $H = 4$ by $W = 6$. We compare the performances of models (31) and (32) in Table 1. In the table, $Z_1(K)$ and Z_2 are the optimal objective values of Model (31) using K samples, and Model (32), respectively. We then use Monte Carlo simulation with 100,000 samples to estimate the actual objective function values achieved by the first stage solutions derived from Models (31) and (32). The corresponding estimated objected function values are recorded in columns $\hat{Z}_1(K)$ and $\hat{Z}_2$. The computation experiment is conducted on an 800 MHz Labtop with 1G RAM using CPLEX version 9.1.

From Table 1, we can see how the performance of stochastic optimization of Model (31) changes as we change the sample sizes. We observe that the optimal objective value of Model (31) underestimates the expected completion time derived from the solutions of Model (31), which is due to the artifact of sampling approximation of stochastic optimization. When the parameter β becomes very small, the variances of the primitive uncertainties $\tilde{z}_{ij}$ increases dramatically and the gap between Z_1 and $\hat{Z}_1$ increases significantly. On the other hand, under the same circumstances, Model (32) provides a very consistent bound of the expected completion times and significantly outperforms the sampling method using 1000 samples. When the variances of $\tilde{z}_{ij}$ are moderate, the sampling approximation with 1000 samples slightly outperforms our approximation method. Given the approximation nature of our approach, the quality of the solutions generated by Model (32) is encouraging.

7 Conclusions

Although we only solve the stochastic optimization model approximately, we feel that the key advantage of our approach is the scalability to multistage models without suffering from the "curse of dimensionality" experienced by most dynamic and stochastic programs.

We see that being able to formulate as a standard mathematical programming model such as a second order cone programming problem is a definite advantage. It enables us to exploit specific structures for computationally efficiency suited for large scale implementations.

References

[1] Beale, E.M.L. (1955): On minimizing a convex function subject to linear inequalities, *Journal of the Royal Statistical Society Series B*, 17, 173-184.

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

[3] Ben-Tal, A., Nemirovski, A. (1999): Robust solutions to uncertain programs, *Oper. Res. Let.*, 25, 1-13.

[4] Ben-Tal, A., Nemirovski, A. (2000): Robust solutions of Linear Programming problems contaminated with uncertain data, *Math. Progr.*, 88, 411-424.

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

[7] Ben-Tal, A., B. Golany, A Nemirovski., J. Vial. (2003): Supplier-Retailer Flexible Commitments Contracts: A Robust Optimization Approach. *Working Paper*, Technion.

C	β	Z_2	$Z_1(1000)$	$Z_1(500)$	$Z_1(100)$	$Z_1(50)$
8	0.0001	58.50	13.50	13.50	12.00	12.00
8	0.001	58.53	34.50	33.00	15.00	13.49
8	0.005	58.67	52.60	48.50	49.33	17.97
8	0.01	58.83	55.60	49.06	46.51	26.96
8	0.1	54.34	42.54	43.54	39.30	41.39
8	0.2	48.73	38.29	38.50	36.47	38.24
8	0.3	45.30	36.27	36.31	35.38	36.54
8	0.4	41.90	35.11	35.14	34.24	35.08
19	0.0001	44.25	13.50	13.50	12.00	12.00
19	0.001	44.27	22.49	19.50	14.99	13.49
19	0.005	44.35	39.03	34.75	20.99	17.97
19	0.01	44.45	41.42	36.38	29.92	20.97
19	0.1	42.67	35.73	36.25	32.85	33.43
19	0.2	39.32	33.01	32.99	31.42	32.60
19	0.3	36.26	31.59	31.59	30.87	31.65
19	0.4	33.38	30.88	30.82	30.14	30.86
C	β	$\hat{Z}_2$	$\hat{Z}_1(1000)$	$\hat{Z}_1(500)$	$\hat{Z}_1(100)$	$\hat{Z}_1(50)$
8	0.0001	59.06	69.60	70.50	70.36	70.36
8	0.001	58.22	59.58	59.65	65.72	68.99
8	0.005	56.83	58.39	58.99	59.53	63.95
8	0.01	55.59	58.32	57.53	58.27	59.64
8	0.1	43.95	42.99	43.74	44.68	44.75
8	0.2	38.98	38.55	38.75	39.42	40.25
8	0.3	37.09	36.50	36.52	36.94	37.39
8	0.4	35.97	35.40	35.45	35.92	36.07
19	0.0001	43.90	69.60	70.50	70.36	70.36
19	0.001	44.01	47.84	55.41	65.72	68.99
19	0.005	43.59	45.96	45.79	48.73	63.95
19	0.01	43.31	45.99	44.65	45.21	58.46
19	0.1	37.15	36.10	36.67	37.65	39.33
19	0.2	33.38	33.13	33.26	34.02	34.74
19	0.3	33.75	31.76	31.81	32.28	32.97
19	0.4	31.84	31.06	31.07	31.46	31.85

Table 1: Objective values of models (32) and (31) for $K = 50, 100, 500$ and 1000.

[8] Bertsimas, D., Lauprete, G. J., Samarov, A. (2003): Shortfall as a risk measure: properties, optimization and applications. *Journal of Economics Dynamics and Control.*

[9] Bertsimas, D., Sim, M. (2003): Robust Discrete Optimization and Network Flows, *Math. Progr.*, 98, 49-71.

[10] Bertsimas, D., Sim, M. (2004): Price of Robustness, *Oper. Res.*, 52(1), 35-53.

[11] Bertsimas, D., Sim, M. (2004): Robust Conic Opimization, *Working Paper*, MIT.

[12] Bertsimas, D., Sim, M. (2004): Robust Discrete Opimization under Ellipsoidal Uncertainty Set, *Working Paper*, MIT.

[13] Bertsimas, D., A. Thiele. (2003): A Robust Optimization Approach to Supply Chain Management. Working Paper, MIT.

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

[15] Charnes, A. and Cooper, W.W. (1959): Uncertain Convex Programs: Randomize solutions and confidence level, *Management Science*, 6, 73-79.

[16] Chen, X. Sim, M., Sun, P. (2005): A Robust Optimization Perspective of Stochastic Programming, *Working Paper*, National University of Singapore.

[17] Chou, M. Sim, M., So, R. (2005): A Robust Optimization Framework for Analyzing Distribution Systems with Transshipments, *Working Paper*, National University of Singapore.

[18] Dantzig, G.B. (1955): Linear programming under uncertainty, *Management Science*, 1, 197-206.

[19] Dyer, M., Stougie, L. (2005): Computational complexity of stochastic programming problems, *Mathematical Programming*, to appear.

[20] El-Ghaoui, Lebret, H. (1997): Robust solutions to least-square problems to uncertain data matrices, *SIAM J. Matrix Anal. Appl.*, 18, 1035-1064.

[21] El-Ghaoui, L., Oustry, F., Lebret, H. (1998): Robust solutions to uncertain semidefinite programs, *SIAM J. Optim.*, 9, 33-52.

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

[23] Nemirovski, A. and Shapiro, A. (2004): Convex Approximation of Chance Constrained Programs, *Working Paper*, Georgia Tech, Atlanta.

[24] Prékopa, A. (1995): *Stochastic Programming*, Kluwer Academic Publishers, Dordrecht, Boston.

[25] Ruszczynski, A., Shapiro, A. (eds.) (2003): *Stochastic Programming, Handbook in Operations Research and Management Science*, Elsevier Science, Amsterdam.

[26] Scarf, H. (1958): A min-max solution of an inventory problem, *Studies in The Mathematical Theory of Inventory and Production*, 201-209 Stanford University Press, California, 1958.

[27] Shapiro, A. and Nemirovski, A. (2004): On Complexity of stochastic programming Problems, *Working Paper*, Georgia Tech, Atlanta.

[28] Soyster, A.L. (1973): Convex programming with set-inclusive constraints and applications to inexact linear programming, *Oper. Res.*, 21, 1154-1157.