

# Robust Two-stage Stochastic Linear Optimization with Risk Aversion <sup>\*</sup>

Ai-fan Ling <sup>†</sup>   Jie Sun <sup>‡</sup>   Nai-hua Xiu <sup>§</sup>   Xiao-guang, Yang <sup>¶</sup>

**Abstract:** We study a two-stage stochastic linear optimization problem where the recourse function is risk-averse rather than risk neutral. In particular, we consider the mean-conditional value-at-risk objective function in the second stage. The model is robust in the sense that the distribution of the underlying random variable is assumed to belong to a certain family of distributions rather than to be exactly known. We start from analyzing a simple case where uncertainty arises only in the objective function, and then explore the general case where uncertainty also arises in the constraints. We show that the former problem is equivalent to a semidefinite program and the latter problem is generally NP-hard. Applications to two-stage portfolio optimization and material order problems are considered. Numerical results show that the proposed robust risk-averse two-stage stochastic programming model can effectively control the risk with solutions of acceptable good quality.

**Keywords:** Uncertainty modeling, stochastic programming, robustness, conditional value-at-risk, semidefinite programming

**MR(2010)Classification** 90C15, 90B50

## 1 Introduction

Consider the classical two-stage stochastic linear optimization problem with fixed recourse (see, e.g., Birge and Louveaux, 1997):

$$\min_{\mathbf{x} \in \mathcal{X}} \{ \mathbf{c}^T \mathbf{x} + \mathbb{E}[\mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}(\omega))] \} \quad (1.1)$$

where

$$\begin{aligned} \mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}(\omega)) = & \min \mathbf{q}(\omega)^T \mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}(\omega)\mathbf{x} + \mathbf{D}\mathbf{y} = \mathbf{b}(\omega), \quad \mathbf{y} \geq 0, \end{aligned}$$

---

<sup>\*</sup>This work is partially supported by National Natural Science Foundations of China (71371090), Natural Science Foundation of Jiangxi Province (20114BAB211008), and the Fund for Outstanding Youths in Jiangxi University of Finance and Economics.

<sup>†</sup>School of Finance, Jiangxi University Finance & Economics, Nanchang, China. E-mail: aifanling@jxufe.edu.cn

<sup>‡</sup>Department of Mathematics and Statistics, Curtin University, Perth, Australia, and School of Business, National University of Singapore. E-mail: jie.sun@curtin.edu.au

<sup>§</sup>School of Science, Beijing Jiaotong University, Beijing, China, E-mail: nhxiu@center.njtu.edu.cn

<sup>¶</sup>Academy of Mathematics & Systems Science, CAS, Beijing, China. E-mail: xgyang@iss.ac.cn

is the second-stage recourse problem.  $\mathbf{x} \in \mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{U}\mathbf{x} = \mathbf{u}, \mathbf{x} \geq 0\}$  is the vector of first-stage decision variables which must be determined before the exact value of the random parameters  $\boldsymbol{\xi}(\omega) = (\mathbf{q}(\omega), \mathbf{A}(\omega), \mathbf{b}(\omega))$  are realized.  $\mathbf{A}(\omega) \in \mathbb{R}^{l \times n}$ ,  $\mathbf{q}(\omega) \in \mathbb{R}^m$ , and  $\mathbf{b}(\omega) \in \mathbb{R}^l$  are random parameters dependent on the event  $\omega$  in the sample space  $\Omega$ . Vectors  $\mathbf{c}, \mathbf{u}$ , matrix  $\mathbf{U}$ , and the fixed recourse matrix  $\mathbf{D} \in \mathbb{R}^{l \times m}$  are assumed to be deterministic and are given in advance.

As a modeling tool, (1.1) has appeared in a wide range of applications, see for example [6, 20, 31] and references therein. It should be noted that the expected recourse function in the objective of problem (1.1) is only suitable for a risk-neutral decision maker and therefore is called a risk-neutral approach. However, most of decision makers are risk-averse and the risk-neutral approach is not appropriate to them. We therefore turn to the following revision to model (1.1). Let  $\mathcal{F}$  be a  $\sigma$ -algebra on  $\Omega$  and  $\mathbb{P}$  be a probability measure on  $\Omega$ . The two-stage stochastic programming with mean-risk aversion can be expressed as

$$\min_{\mathbf{x} \in \mathcal{X}} \{ \mathbf{c}^T \mathbf{x} + \mathbb{E}[\mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}(\omega))] + \lambda \rho(\mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}(\omega))) \} \quad (1.2)$$

where  $\rho : \mathfrak{Z} \rightarrow \mathbb{R}$  is the risk measure function defined on a linear space  $\mathfrak{Z}$  of  $\mathcal{F}$ -measurable functions on the probability space  $\{\Omega, \mathcal{F}, \mathbb{P}\}$ ,  $\lambda \geq 0$  is a trade-off coefficient representing the risk-averse rate of the decision maker. The larger the coefficient  $\lambda$  is, the more risk-averse the decision maker is. A special case of (1.2), called the Mean-variance recourse has been considered by Ahmed (2006) [1]. It has been found that the mean-variance criterion can lead to an NP-Hard problem since the second-stage cost  $\mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}(\omega))$  is nonlinear with respect to  $\mathbf{x}$  despite the convexity of the variance operator.

Different from Ahmed (2006) [1], Schultz and Tiedemann (2006) [37] focus on two-stage mixed-integer stochastic programming involving mean-CVaR (conditional value-at-risk) criterion and use a split-variable formulation. They develop an algorithm based on the Lagrangian relaxation of non-anticipativity. Similar research can be found in [14, 15, 22]. Recently, Miller and Ruszczyński (2011) [27] formulate a risk-averse two-stage stochastic linear programming problem in which unresolved uncertainty remains after the second stage. The objective function in [27] is a composition of risk measures. Two decomposition methods for solving the problem are proposed, one is based on the generic cutting plane approach, while another exploits the composite structure of the objective function. Noyan (2012) [29] also considers a risk-averse two-stage stochastic programming model, where CVaR is used as the risk measure and constructs two decomposition algorithms based on the generic Benders-decomposition approach for solving such problems.

In summary, two-stage stochastic programming with risk-averse is a fairly recent and intensive development. It has found significant applications in different areas such as chemical engineering (Schultz and Tiedemann, 2006, [37]), energy optimization (see, e.g., Schultz and Neise, 2007, [36]), transportation network protection (Liu et al., 2009, [24]), and dynamic sampling algorithms for stochastic programs with risk aversion (Philpott and de Matos, 2012 [30]).

In this paper, we consider a two-stage stochastic linear programming model with the

mean-CVaR recourse function. A major consideration for the choice of mean-CVaR criterion is that mean-CVaR is a coherent risk measure while the mean-variance is not, see Rockafellar (2007) [33]. In addition, our model is different from other mean-CVaR models above because it is computationally tractable. In fact, the deterministic equivalence of our model is a semidefinite programming problem if the uncertainty does not show up in the second-stage constraints. The model can be expressed in the following form:

$$\min_{\mathbf{x} \in \mathcal{X}} \left\{ \mathbf{c}^T \mathbf{x} + \sup_{\mathbb{P} \in \mathcal{P}} \{ \mathbb{E}[\mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}(\omega))] + \lambda \text{CVaR}_\alpha(\mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}(\omega))) \} \right\}. \quad (1.3)$$

In this model we assume that the probability distribution  $\mathbb{P}$  for the random parameters  $\boldsymbol{\xi}(\omega)$  is not known precisely. It is then “robust” to hedge against ambiguity in probability distributions by using the maximum mean-CVaR of second-stage cost over a set  $\mathcal{P}$  of possible probability distributions. The set  $\mathcal{P}$  is typically described by a set of known moments or bounded moments. Some useful bounds on the expected second-stage cost with first-order moment information includes the Jensen bound [19] and the Edmundson-Madansky bound [13, 25], see also [12, 20] for more discussions. Applications on stochastic portfolio optimization and material order problems are considered. Numerical results show that the proposed robust (or “minimax”) two-stage stochastic programming model can effectively control the risk while keep solutions in acceptable good quality.

The general minimax stochastic programming problem was first considered by Žáčková (1966) [42] and Dupačová (1987) [11]. The related algorithms include subgradient-based methods [7], cutting plane algorithms [32], the sample-average approximation methods [38, 39, 40]. Ang, Meng and Sun (2014) [2] considered a two-stage stochastic linear programs with incomplete information on uncertainty in their recent research and found a second-order conic expression for their problems. The recent developments for optimization with uncertainty have been provided by Gabrel, Murat and Thiele (2014) [16]. In an earlier paper, Bertsimas et. al. (2010) [5] considered a minimax stochastic programming model with the first two order moments, i.e., their probability family satisfies

$$\mathcal{D}^{BDNT}(\mathcal{M}^\xi, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = \left\{ \mathbb{P} \in \mathcal{U}^\xi : \begin{array}{l} \mathbb{P}\{\boldsymbol{\xi}(\omega) \in \mathcal{M}^\xi\} = 1 \\ \mathbb{E}[\boldsymbol{\xi}(\omega)] = \boldsymbol{\mu}_0 \\ \mathbb{E}[\boldsymbol{\xi}(\omega)\boldsymbol{\xi}(\omega)^T] = \boldsymbol{\Sigma}_0 + \boldsymbol{\mu}_0\boldsymbol{\mu}_0^T \end{array} \right\}, \quad (1.4)$$

where  $\mathcal{U}^\xi$  is the set of all probability measures on the measurable space  $(\mathbb{R}^{m^\xi}, \mathcal{B})$  with  $\mathcal{B}$  the  $\sigma$ -algebra on  $\mathbb{R}^{m^\xi}$  and  $m^\xi$  the number of dimensions of  $\boldsymbol{\xi}(\omega)$ ,  $\mathcal{M}^\xi$  is any closed convex set known to contain the support of random vector  $\boldsymbol{\xi}$ , and  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\Sigma}_0$  are respectively the estimation of  $\mathbb{E}[\boldsymbol{\xi}(\omega)]$  and  $\text{Cov}(\boldsymbol{\xi}(\omega), \boldsymbol{\xi}(\omega))$ . The following utility function is used in [5] to measure risk-averse preference:

$$U(\mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}(\omega))) = \max_k (\alpha_k \mathcal{Q}(\mathbf{x}, \boldsymbol{\xi}(\omega)) + \beta_k),$$

where  $\alpha_k \geq 0, \beta_k \in \mathbb{R}$  are constants. In a single-stage stochastic programming setting, Delage and Ye (2010) [9] proposed a more general bounded moments uncertainty set with exact

moment uncertainty (1.4) as a special case, i.e.,

$$\mathcal{D}^{DY}(\mathcal{M}^\xi, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2) = \left\{ \mathbb{P} \in \mathcal{U}^\xi : \begin{array}{l} \mathbb{P}\{\boldsymbol{\xi}(\omega) \in \mathcal{M}^\xi\} = 1 \\ (\mathbb{E}[\boldsymbol{\xi}(\omega)] - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}_0^{-1} (\mathbb{E}[\boldsymbol{\xi}(\omega)] - \boldsymbol{\mu}_0) \leq \gamma_1 \\ \mathbb{E} \left[ (\boldsymbol{\xi}(\omega) - \boldsymbol{\mu}_0) (\boldsymbol{\xi}(\omega) - \boldsymbol{\mu}_0)^T \right] \leq \gamma_2 \boldsymbol{\Sigma}_0 \end{array} \right\}, \quad (1.5)$$

where  $\gamma_1 \geq 0, \gamma_2 \geq 1$  are two controlled parameters. In a sense, our model combines the features of (1.4) and (1.5). The probability family of ours is obtained by the bounded moments information and has separable mean value uncertainty and ellipsoidal second-order uncertainty. We explore the semidefinite programming formulation for the proposed robust two-stage stochastic programming with mean-CVaR by considering first the special case where uncertainties are only in objective and then the general case where the uncertainties are in the constraints. For the case of uncertainty in constraints, we use the affine uncertainty of Nemirovski and Shapiro [28] in the analysis, which was not considered by Bertsimas et. al. (2010) [5] and Delage and Ye [9].

This paper is organized as follows. We start by considering the case of uncertainty in objective and convert it into a semidefinite programming (SDP for short) problem in Section 2. Section 3 discusses the case of uncertainty in constraints with affine uncertainty and gives an approximate algorithms based on SDP. Two applications of the proposed models, the two-stage portfolio optimization and material ordering problem, are considered in Section 4.

**Notations.** We denote a random variable by  $x(\omega)$  or  $\tilde{x}$ . Boldface lowercase letters, such as  $\mathbf{a}$ , represent vector, and the corresponding uppercase letters, such as  $\mathbf{A}$ , denote matrices. The corresponding  $\mathbf{x}(\omega)$  or  $\tilde{\mathbf{x}}$ , and  $\mathbf{A}(\omega)$  or  $\tilde{\mathbf{A}}$  denote random vector and random matrices, respectively. Conventional symbols such as  $\mathbb{R}^n, \mathbb{S}_+^n$  and  $\mathbb{S}_{++}^n$  are used to express respectively the space of  $n$  dimensional real vectors,  $n$  dimensional symmetric positive semidefinite matrices, and  $n$  dimensional symmetric positive definite matrices. Two matrices  $\mathbf{A} - \mathbf{B} \in \mathbb{S}_+^n (\mathbb{S}_{++}^n)$  are denoted by  $\mathbf{A} \succeq (\succ) \mathbf{B}$  or  $\mathbf{A} - \mathbf{B} \succeq 0 (\succ 0)$ . The inner product of two matrices  $\mathbf{A}, \mathbf{B}$  is denoted by  $\mathbf{A} \cdot \mathbf{B} = \text{trace}(\mathbf{A}\mathbf{B}) = \sum_{i,j} a_{ij}b_{ij}$ .

## 2 Uncertainty in the objective

CVaR is popularized by the work of Rockafellar and Uryasev (2000, 2002) [34, 35] and is a coherent risk measure due to its monotonicity, translation invariance, convexity and positive homogeneity (see [4] for detail), and therefore is getting more popular than the value-at-risk (VaR for short), which does not satisfy the convexity [35]. Let  $F_\xi(\cdot)$  be the cumulative distribution function of a random variable  $\tilde{\xi}$ . For any given confidence level  $\alpha \in (0, 1)$ , the  $\alpha$ -VaR (denoted by  $\text{VaR}_\alpha(\tilde{\xi})$ ) is defined as the following  $\alpha$ -quantile

$$\text{VaR}_\alpha(\tilde{\xi}) = \inf\{v \in \mathbb{R} : F_{\tilde{\xi}}(v) \geq \alpha\}.$$

The  $\alpha$ -CVaR is defined as mean of the  $\alpha$ -tail distribution of  $\tilde{\xi}$ , that is

$$\text{CVaR}_\alpha(\tilde{\xi}) = \mathbb{E}\{\tilde{\xi} | \tilde{\xi} \geq \text{VaR}_\alpha(\tilde{\xi})\}.$$

Rockafellar and Uryasev (see, e.g. [34]) have shown that  $\text{CVaR}_\alpha(\tilde{\xi})$  can be expressed as the following one-variable optimization problem

$$\text{CVaR}_\alpha(\tilde{\xi}) = \min_{v \in \mathbb{R}} \left\{ v + \frac{1}{1-\alpha} \mathbb{E}[(\tilde{\xi} - v)_+] \right\}, \quad (2.1)$$

where  $(a)_+ = \max\{a, 0\}$ . Consider the mean-CVaR two-stage stochastic programming problem (1.3) with random parameter  $\tilde{\mathbf{q}}$  and deterministic input data in constraints:

$$\min_{\mathbf{x} \in \mathcal{X}} \left\{ \mathbf{c}^T \mathbf{x} + \sup_{\mathbb{P} \in \mathcal{P}} \left\{ \mathbb{E}_{\mathbb{P}} \left[ \mathcal{Q}(\mathbf{x}, \tilde{\xi}) \right] + \lambda \text{CVaR}_{\alpha, \mathbb{P}} \left( \mathcal{Q}(\mathbf{x}, \tilde{\xi}) \right) \right\} \right\} \quad (2.2)$$

with  $\tilde{\xi} = \tilde{\mathbf{q}}$ ,

$$\mathcal{Q}(\mathbf{x}, \tilde{\xi}) = \mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) = \min \left\{ \tilde{\mathbf{q}}^T \mathbf{y} : \mathbf{A}\mathbf{x} + \mathbf{D}\mathbf{y} = \mathbf{b}, \mathbf{y} \geq 0 \right\}. \quad (2.3)$$

It follows from (2.1) that

$$\begin{aligned} \mathbb{E}_{\mathbb{P}} [\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}})] + \lambda \text{CVaR}_{\alpha, \mathbb{P}} (\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}})) &= \mathbb{E}_{\mathbb{P}} [\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}})] + \lambda \min_{v \in \mathbb{R}} \left\{ v + \frac{1}{1-\alpha} \mathbb{E}_{\mathbb{P}} [(\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) - v)_+] \right\} \\ &= \min_{v \in \mathbb{R}} \left\{ \lambda v + \mathbb{E}_{\mathbb{P}} \left[ \mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) + \frac{\lambda}{1-\alpha} (\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) - v)_+ \right] \right\}. \end{aligned} \quad (2.4)$$

Then, problem (2.2) can be rewritten as

$$\min_{\mathbf{x} \in \mathcal{X}} \left\{ \mathbf{c}^T \mathbf{x} + \sup_{\mathbb{P} \in \mathcal{P}} \min_{v \in \mathbb{R}} \left\{ \lambda v + \mathbb{E}_{\mathbb{P}} \left[ \mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) + \frac{\lambda}{1-\alpha} (\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) - v)_+ \right] \right\} \right\}. \quad (2.5)$$

Notice that  $\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}})$  is convex in  $\mathbf{x}$  and is concave in  $\tilde{\mathbf{q}}$ . Assume that distribution family  $\mathcal{P}$  is weakly compact. Then, the minimax theorem holds for  $\sup_{\mathbb{P} \in \mathcal{P}} \min_{v \in \mathbb{R}}$  since the right-hand side of (2.5) is convex in variable  $v$ , and is concave in  $\mathbb{P}$  (see Shapiro 2001 [38] for detail). Thus, (2.5) is equal to

$$\begin{aligned} &\min_{\mathbf{x} \in \mathcal{X}} \left\{ \mathbf{c}^T \mathbf{x} + \min_{v \in \mathbb{R}} \left\{ \lambda v + \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}} \left[ \mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) + \frac{\lambda}{1-\alpha} (\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) - v)_+ \right] \right\} \right\} \\ &= \min_{\mathbf{x} \in \mathcal{X}, v \in \mathbb{R}} \left\{ \mathbf{c}^T \mathbf{x} + \lambda v + \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}} \left[ \mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) + \frac{\lambda}{1-\alpha} (\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) - v)_+ \right] \right\}. \end{aligned} \quad (2.6)$$

Suppose we do not know the distribution of  $\tilde{\mathbf{q}}$  exactly, but we have the first two-order moments information of  $\tilde{\mathbf{q}}$ , that is, the distribution of  $\tilde{\mathbf{q}}$  is assumed to belong to the following distribution family:

$$\mathcal{P} = \mathcal{D}^q(\mathcal{M}, \boldsymbol{\mu}^q, \boldsymbol{\Sigma}^q, \boldsymbol{\gamma}^q, \gamma_0^q) = \left\{ \mathbb{P} \in \mathcal{U} : \begin{array}{l} \mathbb{P}\{\tilde{\mathbf{q}} \in \mathcal{M}\} = 1, \\ |\mathbb{E}_{\mathbb{P}}[\tilde{q}_i - \mu_i^q]| \leq \sigma_i^q \gamma_i^q, i = 1, \dots, m, \\ \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{q}}\tilde{\mathbf{q}}^T] \preceq \gamma_0^q \boldsymbol{\Sigma}^q + \boldsymbol{\mu}^q(\boldsymbol{\mu}^q)^T \end{array} \right\},$$

where  $\mathcal{U}$  is the set of all probability measures on the measurable space  $(\mathbb{R}^m, \mathcal{B})$  with  $\mathcal{B}$  the  $\sigma$ -algebra on  $\mathbb{R}^m$ ,  $\mathcal{M}$  is any closed convex set known to contain the support of random vector  $\tilde{\mathbf{q}}$ ,  $\boldsymbol{\gamma}^q = (\gamma_1^q, \dots, \gamma_m^q)^T \geq 0$  and  $\boldsymbol{\gamma}_\sigma^q = (\sigma_1 \gamma_1^q, \dots, \sigma_m \gamma_m^q)^T$  with  $\sigma_i$  the standard deviation of  $\tilde{q}_i$  and  $\gamma_0^q \geq 1$  are parameters controlling the size of the uncertain set;  $\boldsymbol{\mu}^q = (\mu_1^q, \dots, \mu_m^q)^T$  and

$\Sigma^p$  are respectively certain estimates of the mean value and covariance matrix of the random vector  $\tilde{\mathbf{q}}$ . For example,  $\boldsymbol{\mu}^q$  and  $\Sigma^q$  may be the sample mean value and samples covariance matrix based on the historical data. We assume that  $\Sigma^q \succ 0$  holds throughout the paper.

If  $\gamma_1 = 0$  in set  $\mathcal{D}^{DY}$  of (1.5), then  $\mathcal{D}^{DY}$  becomes a subset of  $\mathcal{D}^q$ . Thus, generally,  $\mathcal{D}^q \neq \mathcal{D}^{DY}$ , and  $\mathcal{D}^q \cap \mathcal{D}^{DY} \neq \emptyset$ . If, moreover,  $\gamma_1 = 0$  and  $\gamma_0^q = 1$ , then the set  $\mathcal{D}^q$  contains the exact moment uncertainty set (1.4) in [5] as a special case. Hence,  $\mathcal{D}^q$  is more general than  $\mathcal{D}^{BDNT}$  and is comparable to  $\mathcal{D}^{DY}$ .

Let  $\bar{\boldsymbol{\mu}}^q = (\bar{\mu}_1^q, \dots, \bar{\mu}_m^q)^T = \boldsymbol{\mu}^q + \boldsymbol{\gamma}_\sigma^q$  and  $\underline{\boldsymbol{\mu}}^q = (\underline{\mu}_1^q, \dots, \underline{\mu}_m^q)^T = \boldsymbol{\mu}^q - \boldsymbol{\gamma}_\sigma^q$ . Then, inequalities  $|\mathbb{E}_{\mathbb{P}}[\tilde{q}_i - \mu_i^q]| \leq \sigma_i^q \gamma_i^q (i = 1, \dots, m)$  are equivalent to

$$\mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{q}}] \leq \bar{\boldsymbol{\mu}}^q \text{ and } \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{q}}] \geq \underline{\boldsymbol{\mu}}^q. \quad (2.7)$$

Assume that  $\mathcal{M} = \mathbb{R}^m$  in this paper. Then, the inner subproblem sup in (2.6) is an infinite-dimensional optimization problem with probability distribution  $\mathbb{P} \in \mathcal{P}$  and can be expressed as

$$U^q(\mathbf{x}) = \sup_{\mathbb{P} \in \mathcal{D}^q} \int_{\mathbb{R}^m} \left[ \mathcal{Q}(\mathbf{x}, \mathbf{q}) + \frac{\lambda}{1 - \alpha} (\mathcal{Q}(\mathbf{x}, \mathbf{q}) - v)_+ \right] \mathbb{P}(d\mathbf{q}), \quad (2.8)$$

subject to the probability, the given first- and second-order moments inequality constraints,

$$\int_{\mathbb{R}^m} \mathbb{P}(d\mathbf{q}) = 1, \int_{\mathbb{R}^m} \mathbf{q} \times \mathbb{P}(d\mathbf{q}) \leq \bar{\boldsymbol{\mu}}^q, \int_{\mathbb{R}^m} (-\mathbf{q}) \times \mathbb{P}(d\mathbf{q}) \leq -\underline{\boldsymbol{\mu}}^q, \int_{\mathbb{R}^m} \mathbf{q}\mathbf{q}^T \times \mathbb{P}(d\mathbf{q}) \preceq \gamma_0^q \Sigma^q + \boldsymbol{\mu}^q(\boldsymbol{\mu}^q)^T,$$

and nonnegative constraint

$$\mathbb{P}(\tilde{\mathbf{q}} \leq \mathbf{a}) \geq 0, \text{ for any } \mathbf{a} \in \mathbb{R}^m.$$

The nonnegative constraint indicates that sup in problem (2.8) is taken in the cone of nonnegative Borel measures on  $\mathbb{R}^m$ . Consider the dual of problem (2.8), we have [38]

$$\begin{aligned} \min_{z, \bar{\mathbf{z}}, \underline{\mathbf{z}}, \mathbf{Z}} \quad & z + \bar{\mathbf{z}}^T \bar{\boldsymbol{\mu}}^q - \underline{\mathbf{z}}^T \underline{\boldsymbol{\mu}}^q + \mathbf{Z} \cdot (\gamma_0^q \Sigma^q + \boldsymbol{\mu}^q(\boldsymbol{\mu}^q)^T), \\ \text{s.t.} \quad & z + \bar{\mathbf{z}}^T \mathbf{q} - \underline{\mathbf{z}}^T \mathbf{q} + \mathbf{Z} \cdot (\mathbf{q}\mathbf{q}^T) \geq \mathcal{R}(\mathbf{x}, \mathbf{q}), \quad \forall \mathbf{q} \in \mathbb{R}^m, \\ & \bar{\mathbf{z}}, \underline{\mathbf{z}} \geq \mathbf{0}, \mathbf{Z} \succeq \mathbf{0}, z \in \mathbb{R}. \end{aligned} \quad (2.9)$$

where  $z$  is the dual variable of equality constraint,  $\bar{\mathbf{z}} \in \mathbb{R}_+^m$  and  $\underline{\mathbf{z}} \in \mathbb{R}_+^m$  are respectively the dual variables of the second and third inequalities constraints, and  $\mathbf{Z} \in \mathbb{S}_+^m$  is the dual variable of the second-order moment inequality constraint of (2.8), and

$$\mathcal{R}(\mathbf{x}, \mathbf{q}) = \mathcal{Q}(\mathbf{x}, \mathbf{q}) + \frac{\lambda}{1 - \alpha} (\mathcal{Q}(\mathbf{x}, \mathbf{q}) - v)_+.$$

Introducing two new variables  $\mathbf{w}_1 = \bar{\mathbf{z}} - \underline{\mathbf{z}} \in \mathbb{R}^m$ ,  $\mathbf{w}_2 = \bar{\mathbf{z}} + \underline{\mathbf{z}} \in \mathbb{R}_+^m$ , then, combining  $\bar{\boldsymbol{\mu}}^q = \boldsymbol{\mu}^q + \boldsymbol{\gamma}_\sigma^q$  and  $\underline{\boldsymbol{\mu}}^q = \boldsymbol{\mu}^q - \boldsymbol{\gamma}_\sigma^q$ , (2.9) can be expressed equivalently as

$$\begin{aligned} \min_{z, \mathbf{w}_1, \mathbf{w}_2, \mathbf{Z}} \quad & z + \mathbf{w}_1^T \boldsymbol{\mu}^q + \mathbf{w}_2^T \boldsymbol{\gamma}_\sigma^q + \mathbf{Z} \cdot (\gamma_0^q \Sigma^q + \boldsymbol{\mu}^q(\boldsymbol{\mu}^q)^T), \\ \text{s.t.} \quad & z + \mathbf{w}_1^T \mathbf{q} + \mathbf{Z} \cdot (\mathbf{q}\mathbf{q}^T) \geq \mathcal{R}(\mathbf{x}, \mathbf{q}), \quad \forall \mathbf{q} \in \mathbb{R}^m, \\ & \mathbf{w}_2 \geq \mathbf{0}, \mathbf{Z} \succeq \mathbf{0}, z \in \mathbb{R}, \mathbf{w}_1 \in \mathbb{R}^m. \end{aligned} \quad (2.10)$$

From our assumptions, matrix  $\Sigma^q$  is positive definite. This means that strong duality holds [18]. Hence, for given  $\mathbf{x}$ , the objective function value  $U(\mathbf{x})$  of problem (2.8) is equal to the objective value of problem (2.10). We define for simplicity a matrix variable

$$\mathbf{M} = \begin{bmatrix} \mathbf{Z} & \frac{1}{2}\mathbf{w}_1 \\ \frac{1}{2}\mathbf{w}_1^T & z \end{bmatrix} \in \mathbb{S}^{m+1}.$$

Then the dual problem (2.10) can be rewritten as

$$\begin{aligned} \min_{z, \mathbf{w}_1, \mathbf{w}_2, \mathbf{Z}} \quad & z + \mathbf{w}_1^T \boldsymbol{\mu}^q + \mathbf{w}_2^T \boldsymbol{\gamma}^q + \mathbf{Z} \cdot (\gamma_0^q \Sigma^q + \boldsymbol{\mu}^q (\boldsymbol{\mu}^q)^T) \\ \text{s.t.} \quad & (\mathbf{q}^T, 1) \mathbf{M} (\mathbf{q}^T, 1)^T \geq \mathcal{R}(\mathbf{x}, \mathbf{q}), \quad \forall \mathbf{q} \in \mathbb{R}^m, \end{aligned} \quad (2.11)$$

Observe that, when  $\tilde{\mathbf{q}}$  is realized, (2.3) is

$$\mathcal{Q}(\mathbf{x}, \mathbf{q}) = \min_{\mathbf{y}} \{ \mathbf{q}^T \mathbf{y} : \mathbf{A}\mathbf{x} + \mathbf{D}\mathbf{y} = \mathbf{b}, \mathbf{y} \geq 0 \}.$$

Thus, the inequality constraint of problem (2.11) requires that there exists at least a  $\hat{\mathbf{y}} \in \mathcal{X}(\mathbf{x}) = \{ \mathbf{y} : \mathbf{A}\mathbf{x} + \mathbf{D}\mathbf{y} = \mathbf{b}, \mathbf{y} \geq 0 \}$ , such that inequality

$$(\mathbf{q}^T, 1) \mathbf{M} (\mathbf{q}^T, 1)^T \geq \mathbf{q}^T \hat{\mathbf{y}} + \frac{\lambda}{1-\alpha} (\mathbf{q}^T \hat{\mathbf{y}} - v)_+ \quad (2.12)$$

for all  $\mathbf{q} \in \mathbb{R}^m$  holds. Clearly, inequality (2.12) is equivalent to the following two inequalities under the conditions that these two inequalities hold simultaneously for all  $\mathbf{q} \in \mathbb{R}^m$ ,

$$\left. \begin{aligned} (\mathbf{q}^T, 1) \mathbf{M} (\mathbf{q}^T, 1)^T &\geq \mathbf{q}^T \hat{\mathbf{y}} + \frac{\lambda}{1-\alpha} (\mathbf{q}^T \hat{\mathbf{y}} - v), \\ (\mathbf{q}^T, 1) \mathbf{M} (\mathbf{q}^T, 1)^T &\geq \mathbf{q}^T \hat{\mathbf{y}}. \end{aligned} \right\} \quad (2.13)$$

Notice that  $\frac{\lambda}{1-\alpha} > 0$ , thus, if  $\mathbf{q}^T \hat{\mathbf{y}} - v < 0$ , then the first inequality in (2.13) is redundant; otherwise, the second inequality in (2.13) is redundant. Obviously, these two inequalities in (2.13) can be rewritten as the following matrix forms:

$$\begin{aligned} (\mathbf{q}^T, 1) \left[ \begin{array}{cc} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \hat{\mathbf{y}}) \\ \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \hat{\mathbf{y}})^T & \frac{\lambda v}{1-\alpha} + z \end{array} \right] (\mathbf{q}^T, 1)^T \geq 0, \quad \text{for all } \mathbf{q} \in \mathbb{R}^m, \\ (\mathbf{q}^T, 1) \left[ \begin{array}{cc} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - \hat{\mathbf{y}}) \\ \frac{1}{2} (\mathbf{w}_1 - \hat{\mathbf{y}})^T & z \end{array} \right] (\mathbf{q}^T, 1)^T \geq 0, \quad \text{for all } \mathbf{q} \in \mathbb{R}^m. \end{aligned}$$

This is equivalent to two linear matrix inequality constraints,

$$\left. \begin{aligned} \left[ \begin{array}{cc} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \hat{\mathbf{y}}) \\ \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \hat{\mathbf{y}})^T & \frac{\lambda v}{1-\alpha} + z \end{array} \right] \succeq 0, \\ \left[ \begin{array}{cc} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - \hat{\mathbf{y}}) \\ \frac{1}{2} (\mathbf{w}_1 - \hat{\mathbf{y}})^T & z \end{array} \right] \succeq 0. \end{aligned} \right\} \quad (2.14)$$

Then, for any given  $\mathbf{x} \in \mathcal{X}, v \in \mathbb{R}$ , the dual problem (2.11) can be expressed as

$$\begin{aligned}
& \min_{z, \mathbf{w}_1, \mathbf{w}_2, \mathbf{Z}; \mathbf{y}} && z + \mathbf{w}_1^T \boldsymbol{\mu}^q + \mathbf{w}_2^T \boldsymbol{\gamma}_\sigma^q + \mathbf{Z} \cdot (\gamma_0^q \boldsymbol{\Sigma}^q + \boldsymbol{\mu}^q (\boldsymbol{\mu}^q)^T) \\
& \text{s.t.} && \begin{bmatrix} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \mathbf{y}) \\ \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \mathbf{y})^T & \frac{\lambda v}{1-\alpha} + z \end{bmatrix} \succeq 0, \\
& && \begin{bmatrix} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - \mathbf{y}) \\ \frac{1}{2} (\mathbf{w}_1 - \mathbf{y})^T & z \end{bmatrix} \succeq 0, \\
& && \mathbf{A}\mathbf{x} + \mathbf{D}\mathbf{y} = \mathbf{b}, \quad \mathbf{y} \geq 0, \mathbf{w}_2 \geq \mathbf{0}, \mathbf{Z} \succeq 0, z \in \mathbb{R}, \mathbf{w}_1 \in \mathbb{R}^m.
\end{aligned} \tag{2.15}$$

In summary, for the two-stage stochastic programming problem (2.6) with bounded moment uncertainty, we have the following results.

**Theorem 2.1** *Assume that  $\boldsymbol{\Sigma}^q \succ 0$ . The mean-CVaR minimax two-stage stochastic programming problem (2.6) with random objective coefficients and deterministic constraints is equivalent to the following SDP:*

$$\begin{aligned}
& \min_{\mathbf{x}, v; z, \mathbf{w}_1, \mathbf{w}_2, \mathbf{Z}; \mathbf{y}} && \mathbf{c}^T \mathbf{x} + \lambda v + z + \mathbf{w}_1^T \boldsymbol{\mu}^q + \mathbf{w}_2^T \boldsymbol{\gamma}_\sigma^q + \mathbf{Z} \cdot (\gamma_0^q \boldsymbol{\Sigma}^q + \boldsymbol{\mu}^q (\boldsymbol{\mu}^q)^T) \\
& \text{s.t.} && \begin{bmatrix} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \mathbf{y}) \\ \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \mathbf{y})^T & \frac{\lambda v}{1-\alpha} + z \end{bmatrix} \succeq 0, \\
& && \begin{bmatrix} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - \mathbf{y}) \\ \frac{1}{2} (\mathbf{w}_1 - \mathbf{y})^T & z \end{bmatrix} \succeq 0, \\
& && \mathbf{A}\mathbf{x} + \mathbf{D}\mathbf{y} = \mathbf{b}, \quad \mathbf{y} \geq 0, \\
& && \mathbf{U}\mathbf{x} = \mathbf{u}, \quad \mathbf{x} \geq 0, \mathbf{w}_2 \geq \mathbf{0}, \mathbf{Z} \succeq 0, v, z \in \mathbb{R}, \mathbf{w}_1 \in \mathbb{R}^m.
\end{aligned} \tag{2.16}$$

### 3 Uncertainty in constraints with affine uncertainty

In this section, we will consider the uncertainty of coefficients  $\mathbf{A}, \mathbf{b}$  in constraint  $\mathbf{A}\mathbf{x} + \mathbf{D}\mathbf{y} = \mathbf{b}$  with deterministic objective coefficients. The general case where uncertainty arise in both the objective and in constraints can be converted into this case by a simple reformulation. Notice that the uncertainty of matrix  $\mathbf{A}$  will lead a random matrix with  $l \times m$  random elements. For simplicity, let us assume that the random coefficients  $\tilde{\mathbf{A}}, \tilde{\mathbf{b}}$  be affinely dependent on an  $r$  dimensional random vector  $\tilde{\boldsymbol{\zeta}} = (\tilde{\zeta}_1, \dots, \tilde{\zeta}_r)^T$ , namely,

$$\tilde{\mathbf{b}} = b_0 + \sum_{i=1}^r \mathbf{b}_i \tilde{\zeta}_i, \quad \tilde{\mathbf{A}} = \mathbf{A}_0 + \sum_{i=1}^r \mathbf{A}_i \tilde{\zeta}_i, \tag{3.1}$$

where  $\mathbf{b}_i \in \mathbb{R}^l, \mathbf{A}_i \in \mathbb{R}^{l \times n} (i = 0, 1, \dots, r)$  are constants vectors and matrices. Then, the second stage problem of the mean-CVaR two-stage stochastic programming (1.3) with constant objective  $\mathbf{q}$  can be expressed as

$$\begin{aligned}
\mathcal{Q}(\mathbf{x}, \tilde{\boldsymbol{\zeta}}) &= \min_{\mathbf{y}} \left\{ \mathbf{q}^T \mathbf{y} : \tilde{\mathbf{A}}\mathbf{x} + \mathbf{D}\mathbf{y} = \tilde{\mathbf{b}} \right\} \\
&= \min_{\mathbf{y}} \left\{ \mathbf{q}^T \mathbf{y} : \left( \mathbf{A}_0 + \sum_{i=1}^r \mathbf{A}_i \tilde{\zeta}_i \right) \mathbf{x} + \mathbf{D}\mathbf{y} = \mathbf{b}_0 + \sum_{i=1}^r \mathbf{b}_i \tilde{\zeta}_i, \boldsymbol{\zeta} \in \mathcal{M}^\zeta \right\},
\end{aligned} \tag{3.2}$$



where  $\mathcal{M}^\zeta$  is the support of random vector  $\zeta$  and is a known closed convex set. Notice that

$$\mathcal{R}(\mathbf{x}, \zeta) = \mathcal{Q}(\mathbf{x}, \zeta) + \frac{\lambda}{1-\alpha} (\mathcal{Q}(\mathbf{x}, \zeta) - v)_+$$

is still convex in  $v$  and

$$\sup_{\mathbb{P} \in \mathcal{P}} \min_{v \in \mathbb{R}} \mathbb{E}_{\mathbb{P}} \left[ \mathcal{R}(\mathbf{x}, \tilde{\zeta}) \right]$$

is concave in  $\mathbb{P}$ . Thus, the mean-CVaR two-stage stochastic programming (1.3) can be expressed as

$$\min_{\mathbf{x} \in \mathcal{X}, v \in \mathbb{R}} \left\{ \mathbf{c}^T \mathbf{x} + \lambda v + \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}} \left[ \mathcal{Q}(\mathbf{x}, \tilde{\zeta}) + \frac{\lambda}{1-\alpha} (\mathcal{Q}(\mathbf{x}, \tilde{\zeta}) - v)_+ \right] \right\}. \quad (3.3)$$

We assume that the probability distribution of  $\tilde{\zeta}$  belongs to the distribution family:

$$\mathcal{P} = \mathcal{D}^\zeta(\mathcal{M}^\zeta, \underline{\boldsymbol{\mu}}^\zeta, \boldsymbol{\Sigma}^\zeta, \boldsymbol{\gamma}^\zeta, \gamma_0^\zeta) = \left\{ \mathbb{P} \in \mathcal{U}^\zeta : \begin{array}{l} \mathbb{P}\{\tilde{\zeta} \in \mathcal{M}^\zeta\} = 1, \\ \left| \mathbb{E}_{\mathbb{P}}[\tilde{\zeta}_i - \mu_i^\zeta] \right| \leq \sigma_i^\zeta \gamma_i^\zeta, i = 1, \dots, r, \\ \mathbb{E}_{\mathbb{P}}[\tilde{\zeta} \tilde{\zeta}^T] \preceq \gamma_0^\zeta \boldsymbol{\Sigma}^\zeta + \underline{\boldsymbol{\mu}}^\zeta (\underline{\boldsymbol{\mu}}^\zeta)^T, \end{array} \right\},$$

where  $\mathcal{U}^\zeta$  is the set of all probability measures on the measurable space  $(\mathbb{R}^r, \mathcal{B})$  with  $\mathcal{B}$  the  $\sigma$ -algebra on  $\mathbb{R}^r$ ,  $\sigma_i^\zeta$  is the standard deviation of  $\tilde{\zeta}_i$  and  $\boldsymbol{\gamma}_\sigma^\zeta = (\sigma_1 \gamma_1^\zeta, \dots, \sigma_m \gamma_m^\zeta)^T$ . Let  $\bar{\boldsymbol{\mu}}^\zeta = (\bar{\mu}_1^\zeta, \dots, \bar{\mu}_r^\zeta)^T = \underline{\boldsymbol{\mu}}^\zeta + \boldsymbol{\gamma}_\sigma^\zeta$  and  $\underline{\boldsymbol{\mu}}^\zeta = (\underline{\mu}_1^\zeta, \dots, \underline{\mu}_r^\zeta)^T = \underline{\boldsymbol{\mu}}^\zeta - \boldsymbol{\gamma}_\sigma^\zeta$ . Then, inequalities  $\left| \mathbb{E}_{\mathbb{P}}[\tilde{b}_i - \mu_i^\zeta] \right| \leq \sigma_i^\zeta \gamma_i^\zeta$  ( $i = 1, \dots, r$ ) are equivalent to

$$\mathbb{E}_{\mathbb{P}}[\tilde{\zeta}] \leq \bar{\boldsymbol{\mu}}^\zeta \text{ and } \mathbb{E}_{\mathbb{P}}[\tilde{\zeta}] \geq \underline{\boldsymbol{\mu}}^\zeta. \quad (3.4)$$

Assume that  $\mathcal{M}^\zeta = \mathbb{R}^r$ , then, the inner subproblem  $\sup_{\mathbb{P} \in \mathcal{P}}$  in (3.3) with probability distribution  $\mathbb{P}$  can be expressed as

$$U^\zeta(\mathbf{x}) = \sup_{\mathbb{P} \in \mathcal{D}^\zeta} \int_{\mathbb{R}^r} \left[ \mathcal{Q}(\mathbf{x}, \zeta) + \frac{\lambda}{1-\alpha} (\mathcal{Q}(\mathbf{x}, \zeta) - v)_+ \right] \mathbb{P}(d\zeta), \quad (3.5)$$

subject to the probability and given first- and second-order moments inequality constraints:

$$\int_{\mathbb{R}^r} \mathbb{P}(d\zeta) = 1, \int_{\mathbb{R}^r} \zeta \times \mathbb{P}(d\zeta) \leq \bar{\boldsymbol{\mu}}^\zeta, \int_{\mathbb{R}^r} (-\zeta) \times \mathbb{P}(d\zeta) \leq -\underline{\boldsymbol{\mu}}^\zeta, \int_{\mathbb{R}^r} \zeta \zeta^T \times \mathbb{P}(d\zeta) \preceq \gamma_0^\zeta \boldsymbol{\Sigma}^\zeta + \underline{\boldsymbol{\mu}}^\zeta (\underline{\boldsymbol{\mu}}^\zeta)^T$$

and nonnegative constraint. Similar to (2.9), the dual of problem (3.5) is

$$\begin{aligned} U_D^\zeta(\mathbf{x}) = & \min_{z^\zeta, \bar{\mathbf{z}}^\zeta, \underline{\mathbf{z}}^\zeta, \mathbf{Z}^\zeta} z^\zeta + (\bar{\mathbf{z}}^\zeta)^T \bar{\boldsymbol{\mu}}^\zeta - (\underline{\mathbf{z}}^\zeta)^T \underline{\boldsymbol{\mu}}^\zeta + \mathbf{Z}^\zeta \cdot (\gamma_0^\zeta \boldsymbol{\Sigma}^\zeta + \underline{\boldsymbol{\mu}}^\zeta (\underline{\boldsymbol{\mu}}^\zeta)^T), \\ & \text{s.t. } z^\zeta + (\bar{\mathbf{z}}^\zeta)^T \zeta - (\underline{\mathbf{z}}^\zeta)^T \zeta + \mathbf{Z}^\zeta \cdot (\zeta \zeta^T) \geq \mathcal{R}(\mathbf{x}, \zeta), \quad \forall \zeta \in \mathbb{R}^r, \\ & \bar{\mathbf{z}}^\zeta, \underline{\mathbf{z}}^\zeta \geq \mathbf{0}, \mathbf{Z}^\zeta \succeq \mathbf{0}, z^\zeta \in \mathbb{R}, \end{aligned} \quad (3.6)$$

where  $z^\zeta$  is the dual variable of equality constraint,  $\bar{\mathbf{z}}^\zeta \in \mathbb{R}_+^r$  and  $\underline{\mathbf{z}}^\zeta \in \mathbb{R}_+^r$  are respectively the dual variables of the second and third inequalities constraints, and  $\mathbf{Z}^\zeta \in \mathbb{S}_+^r$  is the dual

variable of the second order moment inequality constraint. The strong dual theorem means that

$$U_D^\zeta(\mathbf{x}) = U^\zeta(\mathbf{x}), \forall \mathbf{x} \in \mathcal{X}.$$

An equivalent expression of problem (3.6) can be written as, by introducing two new variables  $\mathbf{w}_1^\zeta = \bar{\mathbf{z}}^\zeta - \mathbf{z}^\zeta \in \mathbb{R}^r$ ,  $\mathbf{w}_2^\zeta = \bar{\mathbf{z}}^\zeta + \mathbf{z}^\zeta \in \mathbb{R}_+^r$ ,

$$\begin{aligned} U_D^\zeta(\mathbf{x}) = \min_{z^\zeta, \mathbf{w}_1^\zeta, \mathbf{w}_2^\zeta, \mathbf{Z}^\zeta} \quad & z^\zeta + (\mathbf{w}_1^\zeta)^T \boldsymbol{\mu}^\zeta + (\mathbf{w}_2^\zeta)^T \boldsymbol{\gamma}_\sigma^\zeta + \mathbf{Z}^\zeta \cdot \left( \gamma_0^\zeta \boldsymbol{\Sigma}^\zeta + \boldsymbol{\mu}^\zeta (\boldsymbol{\mu}^\zeta)^T \right), \\ \text{s.t.} \quad & z^\zeta + (\mathbf{w}_1^\zeta)^T \boldsymbol{\zeta} + \mathbf{Z}^\zeta \cdot (\boldsymbol{\zeta} \boldsymbol{\zeta}^T) \geq \mathcal{R}(\mathbf{x}, \boldsymbol{\zeta}), \quad \forall \boldsymbol{\zeta} \in \mathbb{R}^r, \\ & \mathbf{w}_2^\zeta \geq \mathbf{0}, \mathbf{Z}^\zeta \succeq 0, z^\zeta \in \mathbb{R}, \mathbf{w}_1^\zeta \in \mathbb{R}^r. \end{aligned} \quad (3.7)$$

The construction of problem (3.7) is the same as (2.10). *Can the original problem (3.3) with uncertainty set  $\mathcal{D}^\zeta$  be formulated as a semidefinite program and therefore be solved by a polynomial-time algorithm?* This is the main question what we will discuss in the rest of this subsection. To this end, consider the dual of problem (3.2), we have

$$\begin{aligned} \mathcal{Q}(\mathbf{x}, \tilde{\boldsymbol{\zeta}}) &= \max_{\mathbf{z}} \left\{ (\tilde{\mathbf{b}} - \tilde{\mathbf{A}}\mathbf{x})^T \mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q} \right\} \\ &= \max_{\mathbf{z}} \left\{ \left[ (\mathbf{b}_0 - \mathbf{A}_0\mathbf{x}) + \sum_{i=1}^r (\mathbf{b}_i - \mathbf{A}_i\mathbf{x}) \tilde{\zeta}_i \right]^T \mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q} \right\} \end{aligned} \quad (3.8)$$

For any realized  $\boldsymbol{\zeta} \in \mathbb{R}^r$ , of random vector  $\tilde{\boldsymbol{\zeta}}$ , the inequality constraint of problem (3.7) means that for any  $\mathbf{z} \in \{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\}$ , the following inequality

$$\begin{aligned} z^\zeta + (\mathbf{w}_1^\zeta)^T \boldsymbol{\zeta} + \mathbf{Z}^\zeta \cdot (\boldsymbol{\zeta} \boldsymbol{\zeta}^T) &\geq \left[ (\mathbf{b}_0 - \mathbf{A}_0\mathbf{x}) + \sum_{i=1}^r (\mathbf{b}_i - \mathbf{A}_i\mathbf{x}) \zeta_i \right]^T \mathbf{z} \\ &\quad + \frac{\lambda}{1-\alpha} \left( \left[ (\mathbf{b}_0 - \mathbf{A}_0\mathbf{x}) + \sum_{i=1}^r (\mathbf{b}_i - \mathbf{A}_i\mathbf{x}) \zeta_i \right]^T \mathbf{z} - v \right)_+, \end{aligned} \quad (3.9)$$

holds for any  $\boldsymbol{\zeta} \in \mathbb{R}^r$ . Notice that the essential differences of inequality (2.12) and (3.9) lies in that inequality (3.9) must be satisfied for all  $\mathbf{z} \in \{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\}$ . Let

$$\mathcal{S}(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta; \boldsymbol{\zeta}, \mathbf{z}) = \left\{ \begin{array}{l} z^\zeta + (\mathbf{w}_1^\zeta)^T \boldsymbol{\zeta} + \mathbf{Z}^\zeta \cdot (\boldsymbol{\zeta} \boldsymbol{\zeta}^T) \\ - \left[ (\mathbf{b}_0 - \mathbf{A}_0\mathbf{x}) + \sum_{i=1}^r (\mathbf{b}_i - \mathbf{A}_i\mathbf{x}) \zeta_i \right]^T \mathbf{z} \\ - \frac{\lambda}{1-\alpha} \left( \left[ (\mathbf{b}_0 - \mathbf{A}_0\mathbf{x}) + \sum_{i=1}^r (\mathbf{b}_i - \mathbf{A}_i\mathbf{x}) \zeta_i \right]^T \mathbf{z} - v \right)_+ \end{array} \right\}.$$

Then the inequality (3.9) is equivalent to

$$\mathcal{S}(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta; \boldsymbol{\zeta}, \mathbf{z}) \geq 0, \quad (3.10)$$

for all  $\boldsymbol{\zeta} \in \mathbb{R}^r$ ,  $\mathbf{z} \in \{\mathbf{z} : \mathbf{D}\mathbf{z} \leq \mathbf{q}\}$ . Consider the following minimization problem

$$\mathcal{S}^*(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta) = \min_{\substack{\boldsymbol{\zeta} \in \mathbb{R}^r \\ \mathbf{z} \in \{\mathbf{z} : \mathbf{D}\mathbf{z} \leq \mathbf{q}\}}} \mathcal{S}(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta; \boldsymbol{\zeta}, \mathbf{z}), \quad (3.11)$$

and make the following assumption.

**Assumption 1** For any given feasible  $\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta$ , the optimal solutions of problem (3.11) exist and its optimal value is finite.

Then, under Assumption 1, inequality (3.9) is equivalent to

$$\mathcal{S}^*(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta) \geq 0. \quad (3.12)$$

**Assumption 2** The set

$$\Xi = \{(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta) : \mathcal{S}^*(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta) \geq 0, \mathbf{U}\mathbf{x} = \mathbf{u}, \mathbf{Z}^\zeta \succeq 0, v, z^\zeta \in \mathbb{R}, \mathbf{w}_1^\zeta \in \mathbb{R}^r\}$$

is nonempty.

Assumption 2 ensures that problem (3.3) with uncertainty set  $\mathcal{P}$  is feasible. We next show that problem (3.12) is generally NP-hard. Consider the following two problems.

**Problem 1** Let  $\mathbf{x} \in \mathcal{X}, v, z^\zeta \in \mathbb{R}, \mathbf{w}_1^\zeta \in \mathbb{R}^r, \mathbf{Z}^\zeta \succeq 0$  be given, then for any  $\boldsymbol{\zeta} \in \mathbb{R}^r, \mathbf{z} \in \{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\}$ , check whether or not the inequality constraint (3.12) is satisfied. If not, find a  $\hat{\boldsymbol{\zeta}} \in \mathbb{R}^r, \hat{\mathbf{z}} \in \{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\}$ , such that,

$$\mathcal{S}(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta; \hat{\boldsymbol{\zeta}}, \hat{\mathbf{z}}) < 0.$$

**Problem 2** For any matrix  $\mathbf{C}$  and vector  $\mathbf{d}$  with rational entries, and a nonzero rational number  $\epsilon$ , is there vector  $\mathbf{z}$ , such that

$$\mathbf{C}\mathbf{z} \leq \mathbf{d}, \quad \sqrt{\mathbf{z}^T \mathbf{z}} \geq \epsilon?$$

Problem 2 is related to the 2-norm maximization problem and is shown to be NP-complete by Mangasarian and Shiau [26]. The following results show the relationship of Problem 1 and Problem 2.

**Lemma 3.1** Problem 2 is a special case of Problem 1 and therefore Problem 1 is NP-Hard.

**Proof.** Consider the case of

$$\frac{\lambda}{1-\alpha} \left( \left[ \left[ (\mathbf{b}_0 - \mathbf{A}_0 \mathbf{x}) + \sum_{i=1}^r (\mathbf{b}_i - \mathbf{A}_i \mathbf{x}) \zeta_i \right]^T \mathbf{z} - v \right]_+ = 0.$$

Take

$$z^\zeta = \epsilon^2/4, \quad \mathbf{w}_1^\zeta = \mathbf{0}, \quad \mathbf{Z}^\zeta = \mathbf{I}, \quad \mathbf{b}_0 - \mathbf{A}_0 \mathbf{x} = \mathbf{0}, \\ \mathbf{A}_i \mathbf{x} = \mathbf{0}, \quad \mathbf{b}_i = \mathbf{e}_i, \quad i = 1, \dots, r,$$

where  $\mathbf{I}$  is the identity matrix and  $\mathbf{e}_i$  is a vector that its the  $i^{\text{th}}$  entry is one and other entries are zeros. Then,  $\mathcal{S}(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta; \boldsymbol{\zeta}, \mathbf{z})$  is degenerated into

$$\mathcal{S}_0(\boldsymbol{\zeta}, \mathbf{z}) = \boldsymbol{\zeta}^T \boldsymbol{\zeta} - \boldsymbol{\zeta}^T \mathbf{z} + \frac{\epsilon^2}{4}, \quad \mathbf{z} \in \{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{d}\}, \quad \boldsymbol{\zeta} \in \mathbb{R}^r.$$

It is not hard to find that, for any  $\mathbf{z} \in \{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\}$ .

$$\zeta_z = \frac{\mathbf{z}}{2} = \arg \min_{\zeta \in \mathbb{R}^r} \mathcal{S}_0(\zeta, \mathbf{z}) = \arg \min_{\zeta \in \mathbb{R}^r} \left\{ \zeta^T \zeta - \zeta^T \mathbf{z} + \frac{\zeta^2}{4} \right\},$$

and

$$\mathcal{S}_0(\zeta_z, \mathbf{z}) = -\frac{\mathbf{z}^T \mathbf{z}}{4} + \frac{\zeta^2}{4}, \quad \mathbf{z} \in \{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\}.$$

Thus, as a special case of Problem 1, find a  $\mathbf{z} \in \{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\}$ , such that  $\mathcal{S}_0(\zeta_z, \mathbf{z}) = -\frac{\mathbf{z}^T \mathbf{z}}{4} + \frac{\zeta^2}{4} < 0$  is equivalent to Problem 2 under the case of  $\mathbf{C} = \mathbf{D}^T$ ,  $\mathbf{d} = \mathbf{q}$ . This indicates that Problem 1 is NP-Hard.  $\square$

The analysis above leads to the following conclusion.

**Theorem 3.2** *The mean-CVaR two-stage stochastic programming problem (3.3) with random  $\tilde{\mathbf{A}}, \tilde{\mathbf{b}}$  is generally NP-Hard.*

In spite of the somewhat disappointing theorem, we can still develop an effective solution method for the problem in view of the special structure of the problem. Assume the set  $\{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\}$  have a total of  $L$  extreme points  $\{\mathbf{z}^1, \dots, \mathbf{z}^L\}$ . Note that the function  $\mathcal{S}(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta; \zeta, \mathbf{z})$  is linear in  $\mathbf{z}$  if other variables are fixed. Thus, inequality (3.10) is equivalent to

$$\mathcal{S}(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta; \zeta, \mathbf{z}^\ell) \geq 0, \quad \ell = 1, \dots, L, \forall \zeta \in \mathbb{R}^r, \quad (3.13)$$

which in turn are equivalent to

$$\min_{\zeta \in \mathbb{R}^r} \mathcal{S}(\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{Z}^\zeta; \zeta, \mathbf{z}^\ell) \geq 0, \quad \ell = 1, \dots, L. \quad (3.14)$$

Notice that for any  $\ell = 1, \dots, L$ ,

$$\begin{aligned} \left[ \sum_{i=1}^r (\mathbf{b}_i - \mathbf{A}_i \mathbf{x}) \zeta_i \right]^T \mathbf{z}^\ell &= ((\mathbf{z}^\ell)^T \mathbf{b}_1, \dots, (\mathbf{z}^\ell)^T \mathbf{b}_r) \zeta - (\mathbf{z}^\ell)^T (\mathbf{A}_1 \mathbf{x}, \dots, \mathbf{A}_r \mathbf{x}) \zeta \\ &= [((\mathbf{z}^\ell)^T \mathbf{b}_1, \dots, (\mathbf{z}^\ell)^T \mathbf{b}_r) - \mathbf{x}^T (\mathbf{A}_1^T \mathbf{z}^\ell, \dots, \mathbf{A}_r^T \mathbf{z}^\ell)] \zeta. \end{aligned} \quad (3.15)$$

Let  $\phi(\mathbf{x}, \mathbf{z}) = [(\mathbf{z}^T \mathbf{b}_1, \dots, \mathbf{z}^T \mathbf{b}_r) - \mathbf{x}^T (\mathbf{A}_1^T \mathbf{z}, \dots, \mathbf{A}_r^T \mathbf{z})]$ . Clearly,  $\phi(\mathbf{x}, \mathbf{z})$  is a linear function of  $\mathbf{x}$ . Then, inequality constraint (3.9), from (3.14) and (3.15), can be expressed equivalently as two inequality constraints:

$$\begin{aligned} \mathbf{Z}^\zeta \cdot (\zeta \zeta^T) + \left[ (\mathbf{w}_1^\zeta)^T - \left( 1 + \frac{\lambda}{1-\alpha} \right) \phi(\mathbf{x}, \mathbf{z}^\ell) \right] \zeta \\ + z^\zeta - \left( 1 + \frac{\lambda}{1-\alpha} \right) (\mathbf{b}_0 - \mathbf{A}_0 \mathbf{x})^T \mathbf{z}^\ell - \frac{\lambda v}{1-\alpha} \geq 0, \forall \zeta \in \mathbb{R}^r, \end{aligned} \quad (3.16)$$

$$\mathbf{Z}^\zeta \cdot (\zeta \zeta^T) + [(\mathbf{w}_1^\zeta)^T - \phi(\mathbf{x}, \mathbf{z}^\ell)] \zeta + z^\zeta - (\mathbf{b}_0 - \mathbf{A}_0 \mathbf{x})^T \mathbf{z}^\ell \geq 0, \forall \zeta \in \mathbb{R}^r. \quad (3.17)$$

Similar to the derivation in Section 2, two inequality constraints (3.16) and (3.17) are equivalent to the following two linear matrix inequalities:

$$\mathcal{C}(\mathbf{z}^\ell) : \begin{cases} \begin{bmatrix} \mathbf{Z} & \frac{1}{2} \left( \mathbf{w}_1^\zeta - \left( 1 + \frac{\lambda}{1-\alpha} \right) \phi^T(\mathbf{x}, \mathbf{z}^\ell) \right) \\ \frac{1}{2} \left( (\mathbf{w}_1^\zeta)^T - \left( 1 + \frac{\lambda}{1-\alpha} \right) \phi(\mathbf{x}, \mathbf{z}^\ell) \right) & \phi_1(\mathbf{x}, z^\zeta, \mathbf{z}^\ell) \end{bmatrix} \succeq 0, \\ \begin{bmatrix} \mathbf{Z} & \frac{1}{2} \left( \mathbf{w}_1^\zeta - \phi^T(\mathbf{x}, \mathbf{z}^\ell) \right) \\ \frac{1}{2} \left( (\mathbf{w}_1^\zeta)^T - \phi(\mathbf{x}, \mathbf{z}^\ell) \right) & z^\zeta - (\mathbf{b}_0 - \mathbf{A}_0 \mathbf{x})^T \mathbf{z}^\ell \end{bmatrix} \succeq 0, \end{cases}$$

where

$$\phi_1(\mathbf{x}, z^\zeta, \mathbf{z}^\ell) = z^\zeta - \left( 1 + \frac{\lambda}{1-\alpha} \right) (\mathbf{b}_0 - \mathbf{A}_0 \mathbf{x})^T \mathbf{z}^\ell + \frac{\lambda v}{1-\alpha}.$$

The following results are therefore evident.

**Theorem 3.3** *Assume that the set  $\{\mathbf{z} : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\}$  have a total of  $L$  extreme points  $\{\mathbf{z}^1, \dots, \mathbf{z}^L\}$ . Then the mean-CVaR two-stage stochastic programming (3.3) can be expressed equivalently by the following SDP with  $2L$  linear matrix inequalities constraints:*

$$\begin{aligned} \min_{\mathbf{x}, v, z^\zeta, \mathbf{w}_1^\zeta, \mathbf{w}_2^\zeta, \mathbf{Z}^\zeta} \quad & \mathbf{c}^T \mathbf{x} + \lambda v + z^\zeta + (\mathbf{w}_1^\zeta)^T \boldsymbol{\mu}^\zeta + (\mathbf{w}_2^\zeta)^T \boldsymbol{\gamma}_\sigma^\zeta + \mathbf{Z}^\zeta \cdot \left( \gamma_0^\zeta \boldsymbol{\Sigma}^\zeta + \boldsymbol{\mu}^\zeta (\boldsymbol{\mu}^\zeta)^T \right), \\ \text{s.t.} \quad & \mathcal{C}(\mathbf{z}^\ell), \ell = 1, \dots, L, \quad \mathbf{U} \mathbf{x} = \mathbf{u}, \\ & \mathbf{x} \geq 0, \mathbf{w}_2^\zeta \geq \mathbf{0}, \mathbf{Z}^\zeta \succeq 0, v, z^\zeta \in \mathbb{R}, \mathbf{z}_1^\zeta \in \mathbb{R}^r. \end{aligned} \quad (3.18)$$

## 4 Applications

In this section, we will consider four applications of the proposed models. Section 4.1 is devoted to the application in two-stage portfolio optimization problem, which, generally speaking, requires to maximize a risk-averse expected return and the random variable only arises in the objective function. We propose to use CVaR in dealing with the randomness of two-stage portfolio optimization problem. Section 4.2 gives another application in material ordering in which the random variable arises in the constraints. The demands of products are uncertain, which requires that the ordering plans must be made in view of uncertain market. We assume that the decision maker is risk-averse and therefore there is a CVaR term in the penalty costs of the second stage. Section 4.3 presents a stochastic production-transportation model, In Section 4.4, we report our results on a single facility minimax distance problem.

### 4.1 Application to two-stage portfolio optimization

#### 4.1.1 Problem formulation

An investor can invest  $n$  risky assets in two time stages. Denote the random net returns of risky assets in the two stages by  $\tilde{\mathbf{r}}^t = (\tilde{r}_1^t, \dots, \tilde{r}_n^t)^T \in \mathbb{R}^n (t = 1, 2)$ . Let  $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$  be the dollar amount of invested in the  $n$  risky assets at the first-stage and for simplicity,

assume that the initial budget wealth for the investment is one. Then the budget constraint of the first-stage is

$$\mathbf{e}^T \mathbf{x} = 1, \mathbf{x} \geq 0, \quad (4.1)$$

where  $\mathbf{e}$  is an all one vector with the suitable dimensions. At  $t = 1$ , the investor will rebalance his/her first-stage portfolio position  $\mathbf{x}$ . After balancing, assume that the second-stage position is denoted by  $\mathbf{y} = (y_1, \dots, y_n)^T \in \mathbb{R}^n$ . Then, under the self-finance and proportional transaction cost, we have

$$y_i = x_i(1 + \tilde{r}_i^1) + (1 - \theta)\Delta_i^b - (1 + \theta)\Delta_i^s, i = 1, \dots, n, \quad (4.2)$$

where  $\theta$  is proportional transaction cost of each unit risk asset,  $\Delta_i^b$  and  $\Delta_i^s$  are respectively the dollar amount of buying and selling the  $i$ -th risky asset. The wealth of portfolio, denoted by  $W_P^t$ , at the end of  $t$ -th-stage ( $t=1, 2$ ) is

$$W_P^1 = (\mathbf{e} + \tilde{\mathbf{r}}^1)^T \mathbf{x}, \quad W_P^2 = (\mathbf{e} + \tilde{\mathbf{r}}^2)^T \mathbf{y}.$$

The objective of investor is to maximize the expected wealth of the first- and second-stage and minimize the loss measured by CVaR of portfolio. Thus, the portfolio optimization of the investor can be expressed as the following minimization two-stage stochastic programming:

$$\min_{\mathbf{x} \text{ satisfies (4.1)}} \left\{ \mathbb{E}[\tilde{\mathbf{c}}^T \mathbf{x}] + \mathbb{E}[\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{r}}^2)] + \lambda \text{CVaR}_\alpha(\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{r}}^2)) \right\} \quad (4.3)$$

where

$$\begin{aligned} \mathcal{Q}(\mathbf{x}, \tilde{\mathbf{r}}^2) = \min & \quad -(\mathbf{e} + \tilde{\mathbf{r}}^2)^T \mathbf{y} = -(\mathbf{e} + \tilde{\mathbf{r}}^2, \mathbf{0}, \mathbf{0})^T (\mathbf{y}, \Delta^s, \Delta^b) \\ \text{s.t.} & \quad (4.2), \mathbf{y} \geq 0, \Delta^s \geq 0, \Delta^b \geq 0, \end{aligned}$$

and  $\mathbf{x}$  is the first-stage decision variable, and  $\mathbf{y}, \Delta^s, \Delta^b$  are the second-stage decision variables and  $\tilde{\mathbf{c}} = -\mathbf{e} - \tilde{\mathbf{r}}^1$ .

Notice that the coefficient in the first-stage of problem (4.3) is random and is different from the original problem (1.3) whose coefficient of the first-stage is deterministic. The distribution  $\tilde{\mathbf{r}}^1$  is not actually known and is possibly different from that of  $\tilde{\mathbf{r}}^2$ . We do not consider this complex case and assume that the distribution  $\tilde{\mathbf{r}}^1$  can be obtained by scenario generating approach.<sup>1</sup> Let there be  $K$  scenarios for return  $\tilde{\mathbf{r}}^1$  with values  $\mathbf{r}^{1i} (i = 1, \dots, K)$  at time  $t = 1$ . The probability that each scenario occurs is  $p_i (i = 1, \dots, K)$  with  $\sum_{i=1}^K p_i = 1$ . Assume that the distribution of  $\tilde{\mathbf{r}}^2$  is uncertainty and belongs to a distribution family  $\mathcal{D}^r$  with partial moment information. Then, the scenario + robust version of problem (4.3) can be expressed as

$$\min_{\mathbf{x} \text{ satisfies (4.1)}} \left\{ \hat{\mathbf{c}}^T \mathbf{x} + \max_{\mathbb{P} \in \mathcal{D}^r} \left\{ \mathbb{E}_{\mathbb{P}}[\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{r}}^2)] + \lambda \text{CVaR}_{\alpha, \mathbb{P}}(\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{r}}^2)) \right\} \right\} \quad (4.4)$$

---

<sup>1</sup>Gülten and Ruszczyński (2012) [17] consider a two-stage portfolio problem based on the scenario generating approach. Note that there are  $K^K$  scenarios for  $\tilde{\mathbf{r}}^2$  at  $t = 2$  and the curse of dimension will happen if  $K$  is large. Thus, we use the scenario generating approach for the uncertainty in the first-stage and robust approach for uncertainty in the second-stage.

where

$$\begin{aligned} \mathcal{Q}(\mathbf{x}, \tilde{\mathbf{r}}^2) = \min \quad & -(\mathbf{e} + \tilde{\mathbf{r}}^2)^T \mathbf{y} \\ \text{s.t.} \quad & \mathbf{y} = \mathbf{A}\mathbf{x} + (1 - \theta)\mathbf{\Delta}^b - (1 + \theta)\mathbf{\Delta}^s, \quad \mathbf{y} \geq 0, \quad \mathbf{\Delta}^s \geq 0, \quad \mathbf{\Delta}^b \geq 0, \end{aligned}$$

and  $\hat{\mathbf{c}} = -\mathbf{e} - \sum_{k=1}^K p_k \mathbf{r}^{1k}$ ,  $\mathbf{A} = \text{Diag}(\mathbf{e} + \sum_{i=1}^K p_k \mathbf{r}^{1k})$ , and  $\text{Diag}(\mathbf{a})$  is a diagonal matrix with entries of vector  $\mathbf{a}$  as its diagonal elements. Clearly, problem (4.4) is a standard instance of general two-stage robust optimization problem (2.2). Let  $\mathcal{D}^r = \mathcal{D}^q(\mathcal{M}, \boldsymbol{\mu}^q, \boldsymbol{\Sigma}^q, \boldsymbol{\gamma}^q, \gamma_0^q)$  as defined in Section 2. Then problem (4.4) is completely consistent with problem (2.2) and therefore can be formulated as the SDP by Theorem 2.1.

In order to obtain a solution of problem (4.4), we need more constraints on problem (4.4). Notice that the investor will not buy and sell the same risky asset simultaneously. This means  $\Delta_i^b \Delta_i^s = 0$ , for  $i = 1, \dots, n$  which will result in an NP-Hard problem. In order to eliminate this constraint, we introduce two variables

$$\Delta_i^+ = \Delta_i^b + \Delta_i^s \geq 0, \quad \Delta_i^- = \Delta_i^b - \Delta_i^s, \quad i = 1, \dots, n.$$

Then,

$$\Delta_i^b \Delta_i^s = 0 \Leftrightarrow (\Delta_i^+)^2 - (\Delta_i^-)^2 = 0 \Leftrightarrow (\Delta_i^-)^2 = (\Delta_i^+)^2 \Leftrightarrow |\Delta_i^-| = |\Delta_i^+| = \Delta_i^+$$

which can be relaxed into a two dimensional second order cone constraint, namely,

$$|\Delta_i^-| \leq \Delta_i^+, \quad i = 1, \dots, n. \quad (4.5)$$

The following two linear constraints are clear from the non-negativity of  $\Delta_i^b, \Delta_i^s$

$$\Delta_i^b = \frac{\Delta_i^+ + \Delta_i^-}{2} \geq 0, \quad \Delta_i^s = \frac{\Delta_i^+ - \Delta_i^-}{2} \geq 0. \quad (4.6)$$

Notice that the short-sale is not allowed, then the dollar amount of buying asset  $i$  at the end of first stage can not greater than that of holding all other assets, i.e.,

$$(1 + \theta)\Delta_i^b \leq \sum_{k \neq i}^n (1 + \mathbb{E}[\tilde{r}_k^1])x_k, \quad i = 1, \dots, n,$$

and the dollar amounts of selling asset  $i$  is not greater than that of assets  $i$  itself,

$$(1 + \theta)\Delta_i^s \leq (1 + \mathbb{E}[\tilde{r}_i^1])x_i, \quad i = 1, \dots, n.$$

This gets two inequalities for variables  $\Delta_i^+, \Delta_i^-$  from (4.6) that

$$\left. \begin{aligned} \frac{1}{2}(1 + \theta)(\Delta_i^+ + \Delta_i^-) &\leq \sum_{k \neq i}^n (1 + \mathbb{E}[\tilde{r}_k^1])x_k, \quad i = 1, \dots, n, \\ \frac{1}{2}(1 + \theta)(\Delta_i^+ - \Delta_i^-) &\leq (1 + \mathbb{E}[\tilde{r}_i^1])x_i, \quad i = 1, \dots, n. \end{aligned} \right\} \quad (4.7)$$

Under the case of self-finance, the transaction cost satisfies the budget constraint after the rebalance of portfolio is done, namely,

$$\sum_{i=1}^n \theta(\Delta_i^b + \Delta_i^s) = \sum_{i=1}^n [(1 + \mathbb{E}[\tilde{r}_i^1])x_i - y_i] \Leftrightarrow \sum_{i=1}^n \theta \Delta_i^+ = \sum_{i=1}^n [(1 + \mathbb{E}[\tilde{r}_i^1])x_i - y_i]. \quad (4.8)$$

Obviously, all constraints (4.5), (4.6), (4.7), (4.8) with respect to the new variables  $\Delta_i^+, \Delta_i^-$  are convex. Thus, problem (4.4) is still solved by the proposed in Section 2 under the new constraints. We solve problem (4.4) in next subsection using the software SDPT3[41].<sup>2</sup>

### 4.1.2 Numerical results

Consider a portfolio whose risky assets consist of four indices: (1) Dow Jones Industrial Average Index (DJI), (2) Dow Jones Transportation Average Index (DJT), (3) Dow Jones Composite Average Index (DJA) and (4) Dow Jones Utility Average (DJU). The daily returns of all indices are collected from Jan. 2nd, 2005 to Dec. 31th, 2012 and the partition of parameters estimated period and investment period can be found in Table 1.

The uncertainty set is taken as  $\mathcal{D}^q(\mathbb{R}^4, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \gamma^q \mathbf{e}, \gamma_0)$ , where  $\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}$  are the empirical estimates of the mean and covariance matrix of  $\tilde{\mathbf{q}}$ . Using the data in the estimated period of

Table 1: The partition of parameters estimated period and investment period for four Dow Jones indices risky assets. All data are from RESSET database(<http://www.resset.cn/cn/>).

Estimated period (Number of trading days)	Investment period	
	The first-stage (Number of trading days)	The second-stage (Number of trading days)
01/03/2005~ 12/31/2010 (1507)	01/02/2011~ 12/31/2011 (252)	01/02/2012~ 12/31/2012 (249)

Table 2: The empirical estimates of expected return and covariance matrix based on the data in Estimated period.

Asset	$\hat{\boldsymbol{\mu}}$	$\hat{\boldsymbol{\Sigma}}$			
		DJI	DJT	DJA	DJU
DJI	0.000130	0.000179	-0.000019	-0.000019	-0.000022
DJT	0.000263	-	0.000277	0.000150	0.000153
DJA	0.000367	-	-	0.000340	0.000235
DJU	0.000217	-	-	-	0.000193

Table 1, we obtain  $\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}$ , see Table 2 for the detail reports.

In order to obtain an estimate  $\hat{\mathbf{r}}^1$  of  $\mathbf{r}^1$ , the returns of risky assets at the end first-stage, we use the  $K$ -mean scenario tree generation method (see, e.g. Section 4.2 of reference [17]) based on the historical observations in estimated period of Table 1. Take  $K = 30$  and compute the estimated return  $\hat{\mathbf{r}}^1$  with 500, 1000 and 1500 observations. We find that the estimation of  $\hat{\mathbf{r}}^1$  based on  $K$ -mean scenario tree generation method with 1500 observations is closest to the

<sup>2</sup>We thank Professor Toh, Kim-Chuan of National University of Singapore for his kind helps on implementing SDPT3.



real average return  $\mathbf{r}^1$ . This motivates that we use only  $\widehat{\mathbf{r}}^1$  based on 1500 observations for the following numerical experiments.

For simplicity, we denote  $\mathcal{D}^q(\mathbb{R}^4, \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}, \gamma^q \mathbf{e}, \gamma_0)$  by  $\mathcal{D}(\gamma^q, \gamma_0)$ . In our proposed setting, there exist four controlled parameters,  $\gamma^q, \gamma_0$  and  $\alpha, \lambda$ . How to affect the optimal portfolio for these risk parameters will be discussed in the numerical results. The optional parameters  $\gamma^q, \gamma_0$  can control the size of uncertainty distribution family  $\mathcal{D}(\gamma^q, \gamma_0)$ . Generally speaking, we have

$$\mathcal{D}(\gamma^q, \gamma_0) \subseteq \mathcal{D}(\widehat{\gamma}^q, \gamma_0), \text{ if } \gamma^q \leq \widehat{\gamma}^q, \text{ and } \mathcal{D}(\gamma^q, \gamma_0) \subseteq \mathcal{D}(\gamma^q, \widehat{\gamma}_0), \text{ if } \gamma_0 \leq \widehat{\gamma}_0.$$

The distribution uncertainty set, on one hand, can also reveal the risk aversion of the investor, that is, a larger uncertainty distribution family, in theory, is often corresponding to a more conservative portfolio, and on the other hand, a larger uncertainty distribution family can capture also the more uncertainty from the practice and can be a truer response for the all uncertainties. As stated in the section of Section 2, when  $\gamma^q = 0$  and  $\gamma_0 = 1$ , the corresponding to uncertainty set  $\mathcal{D}(0, 1)$  containing the exact bounded moment uncertainty set considered by Bertsimas et.al.(2010)[5]. Additionally, when  $\gamma^q = 0$ ,  $\mathcal{D}(0, \gamma_0)$  is a subclass of uncertainty sets considered by Delage, Ye(2010)[9]. Figure 1 gives the variation tendency of  $\text{CVaR}_\alpha$  as  $\lambda$  increasing under the different values of  $\alpha$  and three uncertainty sets:  $\mathcal{D}(0, 1), \mathcal{D}(0, 12.5)$  and  $\mathcal{D}(2.25, 12.5)$ <sup>3</sup>. Because  $\lambda$  is a risk-averse coefficient, this implies that increasing the parameter  $\lambda$  will lead to a higher level of risk aversion. But,  $\text{CVaR}_\alpha$  does not increase as  $\lambda$  increases in our models since there is a trade-off between the total objective value and the  $\text{CVaR}_\alpha$  term: Larger  $\lambda$  can provide us with a higher total objective value and a lower  $\text{CVaR}_\alpha$  value when other parameters are fixed.

We find also that a larger distribution family will provide us a higher  $\text{CVaR}_\alpha$  for given  $\lambda$  and  $\alpha$ . This is because the worst-case of a larger distribution family will lead to a worse case than the relatively small distribution family. Notice that the specified  $\alpha$  often represents the risk preference in percentage terms, i.e.,  $\text{CVaR}_\alpha$  quantifies the mean value of the worst  $(1 - \alpha)$  of the total loss. Thus, larger  $\alpha$  values would lead usually to more conservative portfolio and a larger  $\text{CVaR}_\alpha$  since a larger  $\alpha$  means that a larger loss will happen if an extreme loss of a small probability event happens. Figure 1(a), (b) and (c) give this fact.

Figure 2 gives the function curves of the first-stage annual expected return with respect to risk aversion coefficient  $\lambda$  and indicates that the first-stage expected return decreases as the risk aversion coefficient  $\lambda$  increases. As stated above, a larger  $\lambda$  will leads to a conservative policy that inversely has a lower expected return. The similar case happens for uncertain distribution  $\mathcal{D}(\gamma^q, \gamma_0)$ , namely, the portfolio with larger uncertainty set will obtain the poorer expected return. This argument seems to be inconsistent with the implication of Figure 1 in which  $\text{CVaR}_\alpha$  increases as the uncertain distribution family becomes large. In fact, there is no inconsistency in Figure 2 and Figure 1. As we know,  $\text{CVaR}_\alpha$  gives the mean value of an extreme loss when an extreme event happens. A larger uncertainty set means that the extreme

---

<sup>3</sup>The values of parameters  $\gamma^q = 2.25, \gamma_0 = 12.5$  are set based on a simple statistical analysis of the amount of noise present in the estimation of mean and covariance matrix during the years 2005-2010, see also Delage, Ye(2010)[9], pp. 611, for the similar discussion.

Figure 1: The comparison of CVaR under different uncertainty sets and  $\alpha$ .

loss is large even though the probability of the loss is small. This means that, on average, the extreme loss will increase as the uncertain distribution family becomes large. The expected return of first-stage as described in Figure 2 is the normal return of portfolio without extreme event happening, which decreases clearly as the portfolio becomes more conservative. This fact can also be found by the varying  $\alpha$  values in Figure 2. The first-stage annual expected return drops to 7% at the close to  $\lambda = 10$  for  $\mathcal{D}(0, 12.5)$  and  $\mathcal{D}(2.25, 12.5)$  when  $\alpha = 0.90$ . The annual expected return drops to 7% at the close to  $\lambda = 3$  for  $\mathcal{D}(0, 12.5)$  and  $\mathcal{D}(2.25, 12.5)$  and  $\lambda = 7$  for  $\mathcal{D}(0, 1)$  when  $\alpha = 0.99$ .

How does the uncertain distribution family  $\mathcal{D}(\gamma^q, \gamma_0)$  affect the performance of the model? Figure 3 gives the answer to this problem. It can be found from Figure 3 (a) and (b) that the first-stage annual expected return decreases slightly, and CVaR increases slightly as parameter  $\gamma^q$  increases. The relative sensitive case happens to parameter  $\gamma_0$  from (c) and (d) in this figure. This phenomenon can be understood easily since, as explained above, the worst-case will become worse when uncertainty set is larger. This leads to worst-case CVaR increasing. On the other hand, a larger uncertainty set means a more conservative portfolio obtained and the usual (non-extreme case) expected return will be lower. However, even though the in-sample expected return decreases as uncertainty set becomes large, the out-of-sample performance of portfolio with the larger uncertainty set become more satisfactory, see, e.g., Figure 4, where we compare the accumulated wealth of the obtained optimal portfolios for three uncertainty sets,  $\mathcal{D}(0, 1)$ ,  $\mathcal{D}(0, 12.5)$  and  $\mathcal{D}(2.25, 12.5)$  with the ‘‘Equal-weighted’’ (EW, for short) strategy [10]<sup>4</sup>. Two unexpected results are found from Figure 4. One is that the out-of-sample accumulated

---

<sup>4</sup>Called also Naive Diversification or ‘ $1/N$ ’ portfolio in literature. DeMiguel, Garlappi and Uppal (2009) find in [10] that the estimation window needed for the sample-based mean-variance strategy and its extensions to outperform the EW strategy is around 3000 months and think that an optimal portfolio choice is still

Figure 2: The variation tendency of the first-stage annual expected return with respect to  $\lambda$  and comparisons under different distribution families and  $\alpha$ 's.

wealth of optimal portfolios based on uncertainty set  $\mathcal{D}(0, 12.5)$  and  $\mathcal{D}(2.25, 12.5)$  outperforms that of optimal portfolio of uncertainty set  $\mathcal{D}(0, 1)$ . Another is that the accumulated wealths of optimal portfolios based on uncertainty set  $\mathcal{D}(0, 12.5)$  and  $\mathcal{D}(2.25, 12.5)$  are also greater than that of EW strategy.

## 4.2 Application to two-stage material order problem

### 4.2.1 Problem formulation

A manufacturer plans to produce  $m$  products. There exist  $n$  raw materials in the market and each raw material can be used to produce the  $m$  products. Suppose each unit raw material  $j$  can produce  $a_{ij}$  units products  $i$  ( $i = 1, \dots, m$ ). The unit costs of the raw materials are  $\mathbf{c} = (c_1, \dots, c_n)^T$ . Assume that the market demand for each product  $i$  is  $b_i$ . The manufacturer's objective is to find the optimal order  $x_j$  ( $j = 1, \dots, n$ ), such that the cost is minimum and all productions satisfy the market demands, namely, the manufacturer needs solve the linear programming:

$$\min \left\{ \mathbf{c}^T \mathbf{x} : \sum_{j=1}^n a_{ij} x_j \geq b_i, (i = 1, \dots, m), \quad \sum_{j=1}^n x_j \leq u, \quad x_j \geq 0, (j = 1, \dots, n) \right\} \quad (4.9)$$

where  $u$  is the capacity of processing all the  $n$  raw materials. Notice that in the order model, on one hand, the manufacturer must order the raw materials before market demands  $b_i$  is realized; on the other hand, market demand  $b_i$  varies as the time goes. Hence, the manufacturer must give an estimation of market demand  $b_i$  before the problem (4.9) is solved. This leads to that

---

difficult to be obtained to outperform the EW strategy.

Figure 3: The variation tendencies of the first-stage annual expected return and the second-stage CVaR with respect to distribution uncertainty set.

products from the ordered raw materials can not satisfy the real market demands. Thus, the uncertainty of  $b_i$  introduced will be more popular than a fixed estimation of  $b_i$ . Let there be a “penalty” cost if the real demands cannot be covered by the products since the amount of shortage in products has to be produced by buying the raw materials from the market. Assume that these penalties are proportional to the corresponding shortage in products and denoted by  $q_i (i = 1, \dots, m)$ , the penalty cost of per unit of undeliverable production  $i$ . We call  $q_i$  the *recourse costs*. Additionally, it is not hard to find that the estimate error of coefficient  $a_{ij}$  is also not avoided since the fineness of raw material always varies. Then the manufacturer’s objective will become to minimize the cost of raw materials and recourse costs, namely,

$$\min_{\sum_{j=1}^n x_j \leq u, \mathbf{x} \geq 0} \mathbf{c}^T \mathbf{x} + \mathbb{E}[\mathcal{Q}(\mathbf{x}, \tilde{\boldsymbol{\xi}})], \text{ with } \mathcal{Q}(\mathbf{x}, \tilde{\boldsymbol{\xi}}) = \min_{\mathbf{y}} \left\{ \mathbf{q}^T \mathbf{y} : \tilde{\mathbf{A}} \mathbf{x} + \mathbf{y} \geq \tilde{\mathbf{b}}, \mathbf{y} \geq \mathbf{0} \right\} \quad (4.10)$$

and  $y_i$  is the recourse variable that represents the shortage amount of product  $i$ ,  $\tilde{\mathbf{A}} = (\tilde{a}_{ij})_{i=1:m}^{j=1:n}$  is an  $n$ -by- $m$  matrix, and  $\tilde{\boldsymbol{\xi}} = (\tilde{\mathbf{A}}, \tilde{\mathbf{b}})$  is the uncertainty parameter. This is a standard example of model established in Section 3 with deterministic objective and uncertainty in constraints.

Let the random coefficients  $\tilde{\mathbf{A}}, \tilde{\mathbf{b}}$  be driven by an  $r$  dimensional random vector  $\tilde{\boldsymbol{\zeta}} = (\tilde{\zeta}_1, \dots, \tilde{\zeta}_r)^T$  and satisfy equality (3.1). Then the distribution family of  $\tilde{\boldsymbol{\zeta}}$  is the uncertainty

Figure 4: The comparisons of the out-of-sample accumulated wealths of the optimal portfolios with uncertainty sets  $\mathcal{D}(0, 1)$ ,  $\mathcal{D}(0, 12.5)$  and  $\mathcal{D}(2.25, 12.5)$ , and Equal-weighted (EW) strategy.

set  $\mathcal{D}^\zeta(\mathcal{M}^\zeta, \boldsymbol{\mu}^\zeta, \boldsymbol{\Sigma}^\zeta, \boldsymbol{\gamma}^\zeta, \gamma_0^\zeta)$  defined in Section 3. The robust version of problem (4.10) with mean-CVaR as the loss measure in the second-stage can be expressed as

$$\min_{\mathbf{U}\mathbf{x} \leq u, \mathbf{x} \geq 0} \mathbf{c}^T \mathbf{x} + \max_{\mathbb{P} \in \mathcal{D}^\zeta} \left\{ \mathbb{E}_{\mathbb{P}}[\mathcal{Q}(\mathbf{x}, \tilde{\boldsymbol{\zeta}})] + \lambda \text{CVaR}_{\alpha, \mathbb{P}}(\mathcal{Q}(\mathbf{x}, \tilde{\boldsymbol{\zeta}})) \right\}, \quad (4.11)$$

where  $\mathbf{U} = \mathbf{e}^T$ . Applying the results of Theorem 3.3, we can solve the robust two-stage stochastic ordering problem (4.11) as an SDP.

## 4.2.2 Numerical results

In order to compare the numerical results, we consider an example from [20] (see, e.g., pp. 9-15). Let a refinery supply weekly the gasoline (product 1: P1) and fuel oil (product 2: P2) for a big company. The material of producing the two products is the crude oil that is from two different countries, which can be viewed as the two raw materials (raw 1(R1) and raw 2(R2)). From [20], (see, problem (2.10) of P11 in [20]), in this example, the coefficients of problem (4.11) are that

$$\mathbf{c} = (2, 3)^T, \mathbf{q} = (7, 12)^T, \mathbf{U} = (1, 1)^T, u = 100, \tilde{\mathbf{A}} = \begin{pmatrix} 2 + \tilde{\zeta}_1 & 3 \\ 6 & 3.4 + \tilde{\zeta}_2 \end{pmatrix}, \tilde{\mathbf{b}} = \begin{pmatrix} 180 + \tilde{\zeta}_3 \\ 162 + \tilde{\zeta}_4 \end{pmatrix},$$

where  $\tilde{\boldsymbol{\zeta}} = (\tilde{\zeta}_1, \tilde{\zeta}_2, \tilde{\zeta}_3, \tilde{\zeta}_4)^T$  are a random vector with uncertain distribution. For simplicity, for uncertainty set  $\mathcal{D}^\zeta$  with respect to random vector  $\tilde{\boldsymbol{\zeta}}$ , we take that  $\mu_i^\zeta = 0, i = 1, \dots, 4$  and

$\Sigma^\zeta = \text{Diag}(9, 12, 0.21, 0.16)$ , a diagonal matrix.<sup>5</sup> Then  $\mathcal{D}^\zeta$  is reduced as

$$\mathcal{D}^\zeta(\mathcal{M}^\zeta, \boldsymbol{\mu}^\zeta, \Sigma^\zeta, \boldsymbol{\gamma}^\zeta, \gamma_0^\zeta) = \left\{ \mathbb{P} : \begin{array}{l} \mathbb{P}\{\tilde{\boldsymbol{\zeta}} \in \mathcal{M}^\zeta\} = 1, \\ \left| \mathbb{E}_{\mathbb{P}}[\tilde{\zeta}_i] \right| \leq \sigma_i^\zeta \gamma_i^\zeta, i = 1, \dots, r, \\ \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{\zeta}}\tilde{\boldsymbol{\zeta}}^T] \preceq \gamma_0^\zeta \Sigma^\zeta, \end{array} \right\}. \quad (4.12)$$

Then, under our mean-CVaR setting, the weekly ordering plans of the refinery can be expressed the following two-stage stochastic problem:

$$\begin{aligned} \min_{\mathbf{x}_{raw}} \quad & 2x_{R1} + 3x_{R2} + \max_{\mathbb{P} \in \mathcal{D}^\zeta} \left\{ \mathbb{E}_{\mathbb{P}}[\mathcal{Q}(\mathbf{x}_{raw}, \tilde{\boldsymbol{\zeta}})] + \lambda \text{CVaR}_{\alpha, \mathbb{P}}(\mathcal{Q}(\mathbf{x}_{raw}, \tilde{\boldsymbol{\zeta}})) \right\}, \\ \text{s.t.} \quad & x_{R1} + x_{R2} \leq 100, \mathbf{x}_{raw} = (x_{R1}, x_{R2})^T \geq 0, \end{aligned} \quad (4.13)$$

where

$$\begin{aligned} \mathcal{Q}(\mathbf{x}_{raw}, \tilde{\boldsymbol{\zeta}}) = \min_{\mathbf{y}_{prod}} \quad & 7y_{P1} + 12y_{P2} \\ \text{s.t.} \quad & \begin{pmatrix} 2 + \tilde{\zeta}_1 & 3 \\ 6 & 3.4 + \tilde{\zeta}_2 \end{pmatrix} \begin{pmatrix} x_{R1} \\ x_{R2} \end{pmatrix} + \mathbf{D} \begin{pmatrix} y_{P1} \\ y_{P2} \end{pmatrix} \geq \begin{pmatrix} 180 + \tilde{\zeta}_3 \\ 162 + \tilde{\zeta}_4 \end{pmatrix}, \\ & \mathbf{y}_{prod} = (y_{P1}, y_{P2})^T \geq 0, \end{aligned} \quad (4.14)$$

where  $\mathbf{D} = \mathbf{I}_2$  is a two-dimension identity matrix. Notice that  $\tilde{\mathbf{A}}$  and  $\tilde{\mathbf{b}}$  can be expressed further as

$$\tilde{\mathbf{A}} = \underbrace{\begin{pmatrix} 2 & 3 \\ 6 & 3.4 \end{pmatrix}}_{\mathbf{A}_0} + \tilde{\zeta}_1 \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{\mathbf{A}_1} + \tilde{\zeta}_2 \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{\mathbf{A}_2} + \tilde{\zeta}_3 \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\mathbf{A}_3} + \tilde{\zeta}_4 \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\mathbf{A}_4},$$

and

$$\tilde{\mathbf{b}} = \underbrace{\begin{pmatrix} 180 \\ 162 \end{pmatrix}}_{\mathbf{b}_0} + \tilde{\zeta}_1 \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{\mathbf{b}_1} + \tilde{\zeta}_2 \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{\mathbf{b}_2} + \tilde{\zeta}_3 \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{\mathbf{b}_3} + \tilde{\zeta}_4 \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{\mathbf{b}_4}.$$

Thus, problem (4.13)-(4.14) can be solved using the SDPs established in Section 3. It is not hard to find that the set  $\{\mathbf{z} \geq 0 : \mathbf{D}^T \mathbf{z} \leq \mathbf{q}\} = \{\mathbf{z} \in \mathbb{R}_+^2 : z_1 \leq 7, z_2 \leq 12\}$  is a box and has four extreme points  $\mathbf{z}_1^* = (7, 12)^T$ ,  $\mathbf{z}_2^* = (0, 0)^T$ ,  $\mathbf{z}_3^* = (0, 12)^T$  and  $\mathbf{z}_4^* = (7, 0)^T$  and each extreme point is corresponding to two linear matrix inequality constraints. The resulted SDP is solved by SDPT3 [41] under the case of different risk aversion coefficient  $\lambda$ .<sup>6</sup>

<sup>5</sup>Here, we do not make any assumptions on the distribution of random vector  $\boldsymbol{\zeta}$ . But, in [20], the authors assumed that the components of  $\boldsymbol{\zeta}$  are distributed normally, uniformly and exponentially.

<sup>6</sup>In order to obtain a high equality and comparable solution of the resulted SDP, the suitable choices of parameters  $\gamma^q$  and  $\gamma_0$  are important. Toward this end, we assume that the estimations  $\boldsymbol{\mu}^\zeta$  and  $\Sigma^\zeta$  are obtained based on a large sample data that satisfies the condition in [9] (see, e.g. (17) in Delage and Ye [9]). Using the results of Corollary 4 in [9], we take confidence levels,  $\delta = 0.05$ . Then  $\gamma_1 = 0.1350$ ,  $\gamma_2 = 1.3861$  in  $\mathcal{D}^{DY}$ . We take for simplicity,  $\gamma_i^\zeta = \gamma_i$ , ( $i = 1, \dots, 4$ ),  $\gamma_0 = 2\gamma_2$ . We can show from the results in Delage and Ye [9] that  $\mathbb{P}\{\mathbb{P}_\zeta \in \mathcal{D}(\gamma^\zeta, \gamma_0)\} \geq 1 - \delta = 0.95$  approximately holds using a similar approach.

Table 3 gives some numerical results for different risk parameter  $\lambda$ . As stated by Kall and Wallace [20], two classes of stochastic scenarios for this example are considered. One is that they assume that there exist  $15 \times 15 = 225$  realizations (scenarios) for random vector  $\tilde{\zeta}$  and 225 blocks will be decomposed in their stochastic programming algorithm, by which the first-stage solution and cost are respectively  $\mathbf{x}_{raw} = (38.5390, 20.5390)^T$  and cost= 138.6940. Another is that they discretize the given distribution to 15 and 18 subintervals for the uniform and exponential distribution respectively, getting  $15 \times 18 = 270$  blocks that will be decomposed in their stochastic programming algorithm, by which they obtain another pair of first-stage solution and cost:  $\mathbf{x}_{raw} = (38.5660, 22.1410)^T$  and cost= 141.5560. Some interesting observations can be found from Table 3. (1) our models can easily obtain many first-stage order policies  $\mathbf{x}_{raw}$  with different risk reference  $\lambda$ . (2) The first-stage order policy with less cost than Kall and Wallace's [20] stochastic programming order policy can be obtained when  $\lambda = 6$  for  $\alpha = 0.90$  and  $\lambda = 12$  for  $\alpha = 0.95$ . (3) The first-stage order cost is convex in  $\lambda$ , that is, the cost is first decreases and then increases as  $\lambda$  increases. This is an interesting result. Generally speaking, the order cost will be decreasing as  $\lambda$  increases, like the numerical example in the previous subsection. The increase of the order cost is also easily understood. Due to that the cost of buying an unit raw material is much less than that of penalty caused by the shortage of an unit raw material. Thus, as  $\lambda$  attains a threshold and increases continuously that leads to decision makers' risk aversion level is raised, more conservative order policies will be favorite.

Finally we mention that O.F.V., in the fourth and seventh column of Table 3 stands for the objective function value (O.F.V.) of resulted SDP problem. One can find that O.F.V. are greater than the total cost (140.7470 for the first scenario algorithm or 144.179 for the second scenario algorithm in Kall and Wallace [20]). This is due to that our O.F.V. includes three terms, the first-stage cost, second-stage penalty, and CVaR. That explains why they are generally greater than the total cost in [20] with only the first two terms.

## 4.3 Stochastic production-transportation problem

### 4.3.1 Problem formulation

Suppose there are  $m$  facilities and  $n$  customer locations. Assume that each facility has a normalized production capacity of one. The production cost per unit at each facility  $i$  is  $c_i$ . The demand from each customer location  $j$  is  $b_j$  and known beforehand. We assume that  $\sum_j b_j < m$ . The transportation cost between facility  $i$  and customer location  $j$  is  $q_{ij}$ . The goal is to minimize the total production and transportation cost while satisfying all the customer orders. If we define  $x_i \geq 0$  to be the amount produced at facility  $i$  and  $y_{ij}$  to be the amount transported from  $i$  to  $j$ , the deterministic production-transportation problem is formulated as

Table 3: The optimal solutions and costs for different risk parameter  $\lambda$  when  $\alpha = 0.90$  and  $\alpha = 0.95$  with  $\gamma_i^\zeta = 0.1350 (i = 1, \dots, 4)$  and  $\gamma_0 = 2.7722$ .

$\lambda$	$\alpha = 0.90$			$\alpha = 0.95$		
	$(x_{R1}, x_{R2})$	Cost of first-stage	O.F.V.	$(x_{R1}, x_{R2})$	Cost of first-stage	O.F.V.
0	(42.6210, 31.7355)	180.4486	183.7536	(43.1633, 32.1724)	182.8438	186.4987
1	(40.3561, 28.5789)	166.4489	172.3759	(42.3750, 32.3595)	181.8287	188.1057
2	(39.5902, 27.0032)	160.1913	164.6120	(42.9759, 31.2196)	179.6108	184.3828
3	(38.3577, 25.2278)	152.3988	157.9372	(41.1759, 30.6541)	174.3143	180.2028
4	(38.0034, 23.1137)	145.3479	151.1803	(40.5778, 29.5328)	169.7542	175.9366
5	(37.4697, 22.4790)	142.3764	148.5539	(39.9808, 28.4257)	165.2389	171.7664
6	<b>(37.2108, 21.0046)</b>	<b>137.4354</b>	144.9260	(39.3849, 27.8780)	162.4038	168.2445
7	(37.5905, 22.2363)	141.8901	146.8060	(38.7902, 25.7308)	154.7729	160.0387
8	(38.1823, 22.7520)	144.6207	150.4359	(38.1968, 25.2060)	152.0118	158.1771
9	(38.7756, 23.2612)	147.3348	153.6232	(37.6048, 24.1726)	147.7278	154.3661
10	(39.3701, 23.7640)	150.0323	153.8005	(37.0144, 23.1621)	143.5154	147.6336
11	(39.9658, 24.2605)	152.7134	158.6063	(36.4256, 22.6662)	140.8499	147.0929
12	(40.5627, 24.7509)	155.3782	159.9173	<b>(37.3095, 21.7562)</b>	<b>139.8878</b>	146.7768
13	(41.1606, 25.2352)	158.0270	162.9270	(38.1969, 22.7950)	144.7788	150.0288
14	(41.7595, 25.7135)	160.6597	163.998	(39.0874, 23.3046)	148.0888	151.7778
15	(42.3594, 26.1859)	163.2767	166.9711	(41.7750, 24.3048)	156.4646	160.5089

follows:

$$\begin{aligned}
 \min \quad & \sum_{i=1}^m c_i x_i + \sum_{i=1}^m \sum_{j=1}^n q_{ij} y_{ij} \\
 \text{s.t.} \quad & \sum_{i=1}^m y_{ij} = b_j, \quad \text{for any } j, \\
 & \sum_{j=1}^n y_{ij} = x_i, \quad \text{for any } i, \\
 & 0 \leq x_i \leq 1, \quad y_{ij} \geq 0, y_{ii} = y_{jj} = 0, \text{ for any } i, j.
 \end{aligned} \tag{4.15}$$

We notice first that the transportation cost  $q_{ij}$  is in fact uncertainty. Then from Section 2 the two stage version of problem (4.15) with stochastic  $\tilde{q}_{ij}$  can be expressed as follows.

$$\min_{0 \leq \mathbf{x} \leq \mathbf{1}} \left\{ \mathbf{c}^T \mathbf{x} + \sup_{\mathbb{P} \in \mathcal{P}} \{ \mathbb{E}_{\mathbb{P}} [\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}})] + \lambda \text{CVaR}_{\alpha, \mathbb{P}} (\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}})) \} \right\} \tag{4.16}$$

with

$$\mathcal{Q}(\mathbf{x}, \tilde{\mathbf{q}}) = \min \left\{ \sum_{i=1}^m \sum_{j=1}^n \tilde{q}_{ij} y_{ij} : \text{with constraints in (4.15)} \right\}. \tag{4.17}$$

We use the symbol  $\text{vec}(\cdot)$  of matlab that stacks the columns of a matrix, and denoted by  $\mathbf{Y} = (y_{ij})$ , then  $\text{vec}(\mathbf{Y})$  is a  $mn$ -dimensional vector. Then the SDP expression of problem



(4.16) is from Theorem 2.1,

$$\begin{aligned}
& \min_{\mathbf{x}, v, z, \mathbf{w}_1, \mathbf{w}_2, \mathbf{Z}; \mathbf{Y}} \mathbf{c}^T \mathbf{x} + \lambda v + z + \mathbf{w}_1^T \boldsymbol{\mu}^q + \mathbf{w}_2^T \boldsymbol{\gamma}_\sigma^q + \mathbf{Z} \cdot (\gamma_0^q \boldsymbol{\Sigma}^q + \boldsymbol{\mu}^q (\boldsymbol{\mu}^q)^T) \\
& \text{s.t.} \quad \begin{bmatrix} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \text{vec}(\mathbf{Y})) \\ \frac{1}{2} (\mathbf{w}_1 - (1 + \frac{\lambda}{1-\alpha}) \text{vec}(\mathbf{Y}))^T & \frac{\lambda v}{1-\alpha} + z \end{bmatrix} \succeq 0, \\
& \quad \begin{bmatrix} \mathbf{Z} & \frac{1}{2} (\mathbf{w}_1 - \text{vec}(\mathbf{Y})) \\ \frac{1}{2} (\mathbf{w}_1 - \text{vec}(\mathbf{Y}))^T & z \end{bmatrix} \succeq 0, \\
& \quad \sum_{i=1}^m y_{ij} = b_j, \quad \text{for any } j, \\
& \quad \sum_{j=1}^n y_{ij} = x_i, \quad \text{for any } i, \\
& \quad 0 \leq x_i \leq 1, \quad y_{ij} \geq 0, y_{ii} = 0, \text{ for any } i, j, \\
& \quad \mathbf{x} \geq 0, \mathbf{w}_2 \geq \mathbf{0}, \mathbf{Z} \succeq 0, v, z \in \mathbb{R}, \mathbf{w}_1 \in \mathbb{R}^m.
\end{aligned} \tag{4.18}$$

### 4.3.2 Numerical results

In this subsection, we will report the numerical results of the two stage problems (4.15). We check the proposed model by taking the pair of  $(m, n)$  with  $(5, 20), (5, 50), (10, 100)$ . We generate randomly  $m$  facilities and  $n$  customer locations within the unit square and compute  $\bar{q}_{ij}$  for each pair of  $(n, m)$ . We generate randomly 5,000 uniform cost vectors  $\tilde{\mathbf{q}}$  from independent uniform distributions on intervals  $[1.5\bar{q}_{ij}, 1.5\bar{q}_{ij}]$  for all  $i, j$ . The production cost  $c_i$  is randomly generated from a uniform distribution on the interval  $[1.5\bar{c}, 1.5\bar{c}]$ , where  $\bar{c}$  is the average transportation cost. Similarly, the demand  $h_j$  is randomly generated from the uniform distribution on the interval  $[0.5m/n, m/n]$  so that the constraint  $\sum_j h_j < m$  is satisfied.

Similar to Subsection 4.1.2, we consider three uncertainty sets,  $\mathcal{D}(0, 1), \mathcal{D}(0, 12.5)$  and  $\mathcal{D}(2.25, 12.5)$  and compare their cost under different risk aversion parameter  $\lambda$  and  $\alpha = 0.90, 0.95$ .<sup>7</sup> We compare the proposed robust approach with the data-driven (or sample) approach with the risky neutral utility  $U(x) = x$ , see [5] for the similar statement. The data-driven model is solved using 10,000 samples drawn from the normal distribution with given first and second moments. For the obtained first stage optimal solution  $\mathbf{x}$  and second stage optimal solution  $\mathbf{Y}$ , we compute the first stage production cost (PC) and mean transportation cost (TC) and total cost of two stages for all test problems and approaches, see Table 4 for our numerical results and comparisons of this example.

[Insert Table 4 about here]

We find first that the size of uncertainty sets can impact on the optimal cost and a larger uncertainty set, such as  $\mathcal{D}(2.25, 12.5)$ , can obtain less total cost. This result is expected and consistent with Section 4.1 because that a large uncertainty set can provide more choices for

<sup>7</sup>In this example and next example, we obtain the optimal value under different risk aversion parameter  $\lambda$ , but we give the best optimal value for certain risk aversion parameter  $\lambda$  in reports due to the space limit. The reader can request the detail results with different  $\lambda$  at any time from the correspond author.

decision maker and include into the optimal decision in small uncertainty set as a feasible case. This shows also that a large uncertainty set of distributions with few information can work better than the moment-known uncertainty set  $\mathcal{D}(0, 1)$  or  $\mathcal{D}^{BDNT}$ . Some other observations are similar to Section 4.1.2.

We argue further that the proposed robust approach has better performance than data-driven method as uncertainty set from  $\mathcal{D}(0, 1)$  to  $\mathcal{D}(2.25, 12.5)$ . Indeed, from the total cost of all test instances, the robust model with uncertainty set  $\mathcal{D}(0, 1)$  has similar total cost to the data-driven approach, and the robust model with uncertainty set  $\mathcal{D}(0, 12.5)$  has slightly better total cost, and for the uncertainty set  $\mathcal{D}(2.25, 12.5)$ , the robust model has the least total cost. The explanations are two folds: one is that a large uncertainty set will provide more choices for decision maker, and the other is that combining the CVaR risk factor into objective is important for reducing the total cost and improving the model performance.

## 4.4 Single facility minimax distance problem

### 4.4.1 Problem formulation

Let  $(x_i, y_i)(i = 1, \dots, n)$  denote  $n$  customer locations on a plane. The single facility optimal distance problem is to identify a facility location  $(x, y)$  that minimizes the maximum distance from the facility to the customers. Assuming a rectilinear or Manhattan distance metric, the problem with certainty case is formulated as

$$\min_{x,y}(\{\max_{i=1,\dots,n} |x_i - x| + |y_i - y|\}).$$

Notice that the locations of  $n$  customers are possibly random and expressed as  $(\tilde{x}_i, \tilde{y}_i)(i = 1, \dots, n)$ . Carbone and Mehrez [8] considered the stochastic coordinates  $(\tilde{x}_i, \tilde{y}_i)(i = 1, \dots, n)$  with mean 0 and variance 1, identical, pairwise independent and normal distributed cases. In the current paper, we assume that

$$\tilde{x}_i = \bar{x}_i + \tilde{\zeta}_i^x, \quad \tilde{y}_i = \bar{y}_i + \tilde{\zeta}_i^y, \quad (i = 1, \dots, n),$$

where  $\bar{x}_i$  and  $\bar{y}_i$  are the mean value of  $\tilde{x}_i$  and  $\tilde{y}_i$ . Then, minimax distance problem with stochastic customers locations can be expressed as follows two stage problem under risky neutral case.

$$\min_{x,y} \mathbb{E}[\mathcal{Q}(x, y; \tilde{\zeta})],$$

where

$$\begin{aligned} \mathcal{Q}(x, y; \tilde{\zeta}) = & - \min_z -z \\ \text{s.t.} & \quad z + x + y \geq \tilde{x}_i + \tilde{y}_i, \quad \forall i, \\ & \quad z - x - y \geq -\tilde{x}_i - \tilde{y}_i, \quad \forall i, \\ & \quad z + x - y \geq \tilde{x}_i - \tilde{y}_i, \quad \forall i, \\ & \quad z - x + y \geq -\tilde{x}_i + \tilde{y}_i, \quad \forall i. \end{aligned}$$

This corresponds to a special case in Section 3 with  $\mathbf{c} = \mathbf{0}$  and determinate coefficient matrix  $\mathbf{A}$ , but random right hand side constraints. We assume further that  $\tilde{\boldsymbol{\zeta}} = (\zeta^x; \zeta^y)$  follows the distribution family of  $\mathcal{D}^\zeta$  given by (4.12). Then we can write problem above as the robust two stage programming form with risky aversion.

$$\min_{\mathbf{x}, \mathbf{y} \in \mathcal{X}} \left\{ \sup_{\mathbb{P} \in \mathcal{D}^\zeta} \left\{ \mathbb{E}[\mathcal{Q}(\mathbf{x}, \mathbf{y}; \tilde{\boldsymbol{\zeta}})] + \lambda \text{CVaR}_\alpha(\mathcal{Q}(\mathbf{x}, \mathbf{y}; \tilde{\boldsymbol{\zeta}})) \right\} \right\}. \quad (4.19)$$

Notice that coefficient matrix  $\mathbf{A}$  is constant and  $\tilde{\boldsymbol{\zeta}}$  can be expressed as the form (3.1) using the similar to Section 4.2.2. It is not hard to find that the dual feasible set has  $4n + 1$  extreme points in  $\mathbb{R}^{4n}$  that can be expressed as follows.

$$\{\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, -\mathbf{e}_{n+1}, \dots, -\mathbf{e}_{2n}, \mathbf{e}_{2n+1}, \dots, \mathbf{e}_{3n}, -\mathbf{e}_{3n+1}, \dots, -\mathbf{e}_{4n}\},$$

where  $\mathbf{e}_i (i = 1, \dots, 4n)$  is the  $4n$  dimensional vector whose  $i$ -th element is one and other elements are all zeros. Thus problem (4.19) can be solved using the SDPs approach developed in Section 3.

#### 4.4.2 Numerical results

In our numerical tests of this example, we randomly generate  $n = 10, 20, 40, 50, 60, 80, 100$  customer locations within the unit square respectively. Assume that each customer location is perturbed from its original position by a random distance in a random direction. We estimate the first and second moments  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  using 5,000 such random perturbations. Similarly, we compare the proposed robust approach with data-driven model to find the optimal facility locations. The data-driven model is still solved using 10,000 samples drawn from the normal distribution with given first and second moments, and risk-neutral case of  $U(x) = x$ . For simplicity, we consider still three uncertainty sets,  $\mathcal{D}(0, 1)$ ,  $\mathcal{D}(0, 12.5)$  and  $\mathcal{D}(2.25, 12.5)$  in this example and compare their maximum distance under different risk aversion parameter  $\lambda$  and  $\alpha = 0.90, 0.95$ . For each test problems, we compute the possible optimal facility location  $(x, y)$  and the maximum distance from  $n$  customer locations to  $(x, y)$ , see Table 5 for detail.

[Insert Table 5 about here]

Some similar observations are made from Table 5. The robust model with large uncertainty sets can find the facility location with smaller distance for all  $n$ . Data-driven works better when the number of customers  $n$  is not large, such as,  $n = 10, 20$ . However, the proposed robust approaches work better than the data-driven method as the number of customers  $n$  increases. This is to be expected because the stochastic risk will be increasing and it will be difficultly to control it by a risk-neutral model as  $n$  increases. Therefore, it is worthy to combine a risk controlling function into the objective for a large  $n$ . A shortcoming is that the CPU time increases rapidly as  $n$  gets very large. There are  $4n + 1$  extreme points for dual feasible set and  $4n + 1$  linear matrices inequalities constrains.

## 5 Concluding remarks

A robust two-stage stochastic linear programming model with mean-CVaR recourse is considered. We find that, by considering the uncertainty either in objective function or constraints, the model can be solved by using an SDP approach. However, the model with uncertainty in constraints is more difficult and is shown to be an NP-Hard problem. Four practical examples, the two-stage portfolio selection, the material order problem, the stochastic production-transportation problem, and the single facility minimax distance problem are solved by the proposed algorithms. Numerical results and comparisons with a data-driven method indicate that the model is efficient and capable for solving practical problems.

### Acknowledgment

The authors are thankful to the associate editor and three anonymous referees. Their comments and suggestions have greatly helped improve this paper.

### References

- [1] Ahmed, S. (2006). Convexity and decomposition of mean-risk stochastic programs. *Mathematical Programming*, 106, 433-46.
- [2] Ang, J., Meng, F. & Sun, J. (2014). Two-stage stochastic linear programs with incomplete information on uncertainty. *European Journal of Operational Research*, 233, 16-22.
- [3] Ariyawansa, K. A. and Felt, A. J.(2004). On a new collection of stochastic linear programming test problems, *INFORMS J. Comput.*, 16(3), pp. 291-299.
- [4] Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D.(1999). Coherent measures of risk. *Mathematical Finance*, 9, 203-228.
- [5] Bertsimas, D., Doan, X.V., Natarajan, K. & Teo, C.P.(2010). Models for Minimax Stochastic Linear Optimization Problems with Risk Aversion. *Mathematics of Operations Research*, 35, 580-602.
- [6] Birge, J., & Louveaux, F.(1997). *Introduction to stochastic programming*. New York: Springer-Verlag.
- [7] Breton, M. & El Hachem, S. (1995). Algorithms for the solution of stochastic dynamic minimax problems. *Computational Optimization and Applications*, 4, 317-345.
- [8] Carbone, R., Mehrez, A., 1980. The single facility minimax distance problem under stochastic location demand. *Management Sci.* 26(1) 113-115.
- [9] Delage, E. & Ye, Y. (2010). Distributionally Robust Optimization Under Moment Uncertainty with Application to Data-Driven Problems. *Operations Research*, 58, 595-612.
- [10] DeMiguel, V., Garlappi, L. & Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies*, 22, 1915-1953.
- [11] Dupačová, J. (1987). The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20, 73-88.

- [12] Dokov, S. P. & Morton, D. P. (2005). Second-order lower bounds on the expectation of a convex function. *Mathematics of Operations Research*, 30, 662-677.
- [13] Edmundson, H. P. (1956). Bounds on the expectation of a convex function of a random variable. *Acta Mathematica*, 30, 175-193.
- [14] Eichorn, A. & Romisch, W. (2005). Polyhedral risk measures in stochastic programming. *SIAM Journal on Optimization*, 16, 69-95.
- [15] Fabiań, C.I. (2008). Handling CVAR objectives and constraints in two-stage stochastic models. *European Journal of Operational Research*, 191, 888-911.
- [16] Gabrel, V., Murat C. & Thiele, A. (2014). Recent advances in robust optimization: An overview. *European Journal of Operational Research*, 235, 471-483.
- [17] Gülten, S. & Ruszczyński, A. (2012). Two-Stage Portfolio Optimization with Higher-Order Conditional Measures of Risk. *Manuscript, available at SSRN: <http://ssrn.com/abstract=2022787>*
- [18] Isii, K. (1960). The extrema of probability determined by generalized moments (I) bounded random variables. *Annals of the Institute of Statistical Mathematics*, 12, 119-134.
- [19] Jensen, J. L. (1960). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30, 175-193.
- [20] Kall, P. & Wallace, S. (1994). *Stochastic Programming*. Chichester UK: John Wiley & Sons.
- [21] Koberstein, A. Lucas, C. Wolf, C. and König, D.(2011). Modeling and optimizing risk in the strategic gas-purchase planning problem of local distribution companies. *J. Energy Markets*, 4(1), pp. 47-68.
- [22] Kúnzi-Bay, A. & Mayer, J. (2006). Computational aspects of minimizing conditional value-at-risk. *Computational Management Science*, 3, 3-27.
- [23] Linderoth, J., Shapiro, A. and Wright, S. J.(2006). The empirical behavior of sampling methods for stochastic programming. *Ann. Oper. Res.*, 142(1), pp. 215-241.
- [24] Liu, C., Fan, Y. & Ordóñez, F. (2009). A two-stage stochastic programming model for transportation network protection. *Computers and Operations Research*, 36, 1582-1590.
- [25] Madansky, A. (1959). Bounds on the expectation of a convex function of a multivariate random variable. *The Annals of Mathematical Statistics*, 30, 743-746.
- [26] Mangasarian, O. J. & Shiao, T. H. (1986). A variable-complexity norm maximization problem. *SIAM Journal on Algebraic Discrete Methods*, 7, 455-461.
- [27] Miller, N. & Ruszczyński, A. (2011). Risk-averse two-stage stochastic linear programming: modeling and decomposition. *Operations Research*, 59, 125-132.
- [28] Nemirovski, A. & Shapiro, A. (2006). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17, 969-996.
- [29] Noyan, N. (2012). Two-Stage Stochastic Programming Involving CVaR with an Application to Disaster Management. *Computers & Operations Research*, 39, 541-559.

- [30] Philpott, A. B. & de Matos, V. L. (2012). Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *European Journal of Operational Research*, 218, 470-483.
- [31] Prékopa, A. (1995). *Stochastic programming*. Dordrecht, Boston: Kluwer Academic.
- [32] Riis, M. & Andersen, K. A. (2005). Applying the minimax criterion in stochastic recourse programs. *European Journal of Operational Research*, 165, 569-584.
- [33] Rockafellar, R.T. (2007). Coherent approaches to risk in optimization under uncertainty. *Tutorials in Operations Research INFORMS*, 38-61.
- [34] Rockafellar, R.T. & Uryasev, S. (2000). Optimization of conditional value at risk. *The Journal of Risk*, 2, 21-41.
- [35] Rockafellar, R.T. & Uryasev, S. (2002). Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking and Finance*, 26, 1443-1471.
- [36] Schultz, R. & Neise, F. (2007). Algorithms for mean-risk stochastic integer programs in energy. *Revista Investigacion Operacional*, 28, 4-16.
- [37] Schultz, R. & Tiedemann, S. (2006). Conditional value-at-risk in stochastic programs with mixed-integer recourse. *Mathematical Programming*, 105, 365-86.
- [38] Shapiro, A. (2001). *On duality theory of conic linear problems. Semi-infinite programming: recent advances*, M.A. Goberna and M.A. Lopez, eds., Norwell USA: Kluwer Academic Publishers.
- [39] Shapiro, A. & Kleywegt, A. (2002). Minimax analysis of stochastic problems. *Optimization Methods & Software*, 17, 523-542.
- [40] Sun, H. & Xu, H. (2014). Convergence analysis of stationary points in sample average approximation of stochastic programs with second order stochastic dominance constraints. *Mathematical Programming*, 143, 31-59.
- [41] Toh, K. C., Tütüncü, R. H. & Todd, M. J. (2006). On the implementation and usage of SDPT3 - a Matlab software package for semidefinite-quadratic-linear programming. *Version 4.0*, available: <http://www.math.nus.edu.sg/~mattohkc/sdpt3/guide4-0-draft.pdf>
- [42] Žáčková, J. (1966). On minimax solutions of stochastic linear programming problems. *Časopis pro Peštování Matematiky*, 91, 423-430.
- [43] Zverovich, V., Fabiàn, C. I., Ellison, E. F. D. and Mitra, G. (2012). A computational study of a solver system for processing two-stage stochastic LPs with enhanced Benders decomposition. *Math. Programming Comput.*, 4(3), pp. 211-238.