European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

Robust optimization with order statistic uncertainty set

Pengfei Zhang[a,*], Diwakar Gupta[b]

[a] School of Management, University of Science and Technology of China, Hefei 230026, China
[b] McCombs School of Business, The University of Texas at Austin, 2110 Speedway Stop B6500, Austin, TX 78712-1750, United States

ARTICLE INFO

Article history:
Received 12 April 2022
Accepted 15 May 2023
Available online xxx

Keywords:
Uncertainty modelling
Robust optimization
Order statistics

ABSTRACT

In this paper, we propose a new uncertainty set for robust models of linear optimization problems. We first study data-free and distribution-free statistical properties of continuous and independent random variables using the Probability Integral Transform. Based on these properties, we construct a new uncertainty set by placing constraints on the order statistics of random variables. We utilize the quantiles of random variables to depict the uncertainties and then adopt the formulation of the assignment problem to develop a tractable formulation for the order statistic uncertainty set. We show that the robust optimization models with the interval uncertainty set, the budget uncertainty set, and the demand uncertainty set can be obtained as special cases of the robust optimization model with the order statistic uncertainty set. Finally, using a robust portfolio construction problem as an example, we show via numerical experiments that the order statistic uncertainty set has better performance than other uncertainty sets when the sample size is small and the correlation between random variables is low.

1. Introduction

In many optimization problems, the decision maker needs to make decisions in the presence of uncertainty. Stochastic optimization has long been used to find optimal solutions in such settings. Specifically, the random quantities are assumed to follow some probability distributions, which leads to either a random objective function, or random constraints, or both. Stochastic optimization models can impose significant computational burden, and as a result, approximation procedures are often used — see Birge & Louveaux (2011, Chapter 8–10).

Another popular approach is robust optimization, which has drawn increasingly more attention in recent years. Rather than model random quantities as having known distributions, the robust optimization model aims to find a solution that achieves the best objective performance while remaining feasible for any realization (scenario) of the uncertain quantities within an uncertainty set. If the solution is not feasible for some scenarios excluded from the uncertainty set, then the decision maker may find it more economical to develop contingency plans to deal with those scenarios. Robust Optimization (RO) is particularly attractive when uncertainty characterization via a probability distribution is unreliable. In this paper, we focus on robust optimization and propose a new uncertainty set.

1.1. The robust optimization model

Consider the standard deterministic linear optimization problem given below:

$$\max_{\mathbf{x}} \ \mathbf{c}'\mathbf{x} \tag{1a}$$

$$\text{s.t.} \ \sum_{j\in J_i} a_{ij}x_j \le b_i, \ \ \forall i \in I \tag{1b}$$

$$\underline{\mathbf{x}} \le \mathbf{x} \le \overline{\mathbf{x}}, \tag{1c}$$

where J_i is the index set of x_j's for the ith constraint, and the cardinality of the set J_i is denoted as $|J_i|$. The set of indexes for constraints is denoted as I, which is assumed to be a finite set. In the above model, each coefficient a_{ij} is known and fixed. Suppose now the decision maker is uncertain about the values of a_{ij}'s, and models them by random variables A_{ij}'s. The decision maker aims to find a solution that not only has a high objective value but also ensures the feasibility of constraint (1b) with a particular probability as specified by the following chance constraint.

$$\text{Prob}\left(\sum_{j\in J_i} A_{ij}x_j \le b_i\right) \ge p_i, \ \ \forall i. \tag{2}$$

* Corresponding author.
E-mail addresses: wtzhang@ustc.edu.cn (P. Zhang), diwakar.gupta@mccombs.utexas.edu (D. Gupta).

https://doi.org/10.1016/j.ejor.2023.05.024
0377-2217/

Table 1
Summary of uncertainty sets.

Uncertainty set	Definition	Distributional information	Parameter to control the size
Interval uncertainty set	$\mathcal{U}^I = \{\mathbf{Z} : 0 \le Z_j \le 1, \forall j\}$	range	None
Budget uncertainty set	$\mathcal{U}^B = \left\{\mathbf{Z} : \sum_{j=1}^{\lvert J\rvert} Z_j \le \Gamma, 0 \le Z_j \le 1, \forall j\right\}$	range	Γ
Ellipsoidal uncertainty set	$\mathcal{U}^Q = \left\{\mathbf{Z} \in \mathbb{R}^{\lvert J\rvert} : \mathbf{Z}'\mathbf{\Sigma}^{-1}\mathbf{Z} \le \Omega^2\right\}$	variance & covariance	Ω
Demand uncertainty set	$\mathcal{U}^D = \left\{\mathbf{Z} : \left\lvert \frac{\sum_{j\in S} Z_j}{\lvert S\rvert^{1/\alpha}} \right\rvert \le \gamma, \forall S \subseteq J\right\}$	variance	α, γ
Tail uncertainty set	$\mathcal{U}^T = \left\{\mathbf{Z} : \exists \mathbf{q} \in \mathbb{R}_+^N \text{ s.t. } \mathbf{Z} = \sum_{n=1}^N q_n \mathbf{z}^n, \right.$ $\left. \mathbf{1}'\mathbf{q} = 1, q_n \le \frac{1}{N(1-\alpha')}, n = 1, \ldots, N\right\}$	tail average	α'

The optimization problem involving chance constraints is generally hard to solve (see Yang & Xu, 2016). The probabilistic feasibility of constraints can also be achieved with the RO model – see Ben-Tal & Nemirovski (2000) and Bertsimas et al. (2011).

Uncertainty modelling in the RO model. As in the robust optimization framework presented in Bertsimas & Sim (2004), we assume that the random variable A_{ij} follows an unknown but symmetric distribution, and A_{ij} can take any value in the range $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$. We transform the random variable A_{ij} into the random variable $Z_{ij} \in [0, 1]$ such that $Z_{ij} = |A_{ij} - a_{ij}|/\hat{a}_{ij}$, and we let $\mathbf{Z}_i = (Z_{i1}, \ldots, Z_{ij})$. Henceforth, whenever random variables are mentioned, we mean the random variables Z_{ij}'s.

The robust model that guarantees feasibility of the constraint i for any realization of $\mathbf{Z}_i$ that lies within the uncertainty set $\mathcal{U}_i$ can be written as follows:

$$\max_{\mathbf{x}} \ \mathbf{c}'\mathbf{x} \tag{3a}$$

$$\text{s.t.} \ \sum_{j\in J_i} a_{ij}x_j + \max_{\mathbf{Z}_i\in\mathcal{U}_i} \sum_{j\in J_i} \hat{a}_{ij} \cdot |x_j| \cdot Z_{ij} \le b_i, \ \ \forall i \in I \tag{3b}$$

$$\underline{\mathbf{x}} \le \mathbf{x} \le \overline{\mathbf{x}}, \tag{3c}$$

We denote the subproblem $\max_{\mathbf{Z}_i\in\mathcal{U}_i} \sum_{j\in J_i} \hat{a}_{ij} \cdot |x_j| \cdot Z_{ij}$ as $\beta_i(\mathbf{x}, \mathcal{U}_i)$. We have assumed that the uncertainty sets in the above model to be "constraint-wise". Note that this is without loss of generality because we can always reformulate a joint uncertainty set $\mathcal{U}$ across constraints to be "constraint-wise" (see Section 1.2.1 in Ben-Tal et al., 2009). Therefore, we will drop the constraint index i and focus on an arbitrary constraint. As we will discuss in Section 3.1, the above robust formulation (3) is consistent with the framework presented in Bertsimas & Sim (2004). In the following, we review previous works on uncertainty set characterization, and then discuss uncertainty set design and our contribution.

1.2. Common uncertainty sets

The uncertainty set is an essential component of the RO approach. Table 1 summarizes some of the most common uncertainty sets that have been studied in the RO literature. The table also includes the distributional information that each uncertainty set utilizes and the parameter of each uncertainty set that may be used to adjust its size. We briefly discuss each of the uncertainty sets in the following.

The interval uncertainty set (also known as the box uncertainty set) can be found in Ben-Tal & Nemirovski (2000). It offers a high protection level, but tends to be conservative because it finds the best solution for the worst possible realization of the unknown parameters, i.e., all the random variables Z_j's in the optimal solution are set to 1. The budget uncertainty set, introduced in Bertsimas & Sim (2004), is the first polyhedral uncertainty set that can control the level of conservativeness for the RO model. The idea is to impose the budget constraint on the sum of all random variables Z_j's, which prevents all random variables from taking the extreme value of 1. The ellipsoidal uncertainty set (Ben-Tal & Nemirovski, 1998; El Ghaoui et al., 1998) is motivated by the standard deviation formula, which results in the quadratic form. The matrix $\mathbf{\Sigma}^{-1}$ is the variance-covariance matrix for random variables Z_j's. The demand uncertainty set is inspired by the generalized central limit theorem (Bandi et al., 2015; Bandi & Gupta, 2020). In Table 1, the parameter α is the tail coefficient, and $|S|$ stands for the cardinality of the set S, which is an arbitrary subset of the set J. Note that if we impose constraints on all possible subsets of J, then there will be $2^{|J|} - 1$ constraints for $|J|$ random variables. The tail uncertainty set consists of the convex hull of all the centroids of any $N(1-\alpha')$ points out of N points in the sampled data $\mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^N$ (Bertsimas et al., 2011). The tail uncertainty set is an attractive way to characterize uncertainty when the decision maker's risk preference corresponds to the conditional value-at-risk measure.

In addition to the above-mentioned linear or quadratic uncertainty sets, there are other data-driven approaches to design uncertainty sets. Bertsimas et al. (2018) construct uncertainty sets using statistical hypothesis tests in a data-driven manner. Shang et al. (2017) propose a novel data-driven uncertainty set for solving robust optimization problems based on a piecewise linear kernel. Cheramin et al. (2021) propose data-driven polyhedral uncertainty sets, which can capture correlation information between uncertain variables using principal component analysis.

1.3. The trade-off between the objective value and constraint feasibility

The choice of the uncertainty set is a key consideration in utilizing the RO approach. The uncertainty set in the RO model determines the trade-off between two conflicting goals: good objective performance and a high probability of constraint feasibility. The balance between the objective value and probability of constraint feasibility depends on two aspects of the uncertainty set. The first aspect is the *size* of the uncertainty set, which is chosen by the decision maker depending on his level of conservatism. For a particular uncertainty set, if its size gets smaller, then the objective value improves but the probability of constraint feasibility declines; the improvement in one is always at the expense of the other. The second aspect of the uncertainty set is the *geometric flexibility*. We may be able to improve the performance of both the objective value and constraint protection if the uncertainty set contains regions of more likely uncertain scenarios and excludes the extremely unlikely ones. To achieve this, we need to design the uncertainty set with greater geometric flexibility so that we can adjust its shape to contain regions of high probability.

Besides having the uncertainty set that possesses geometric flexibility, we also need to identify characterizations of uncertainties that can guide us to adjust the shape of the uncertainty set. Such characterization may stem from two sources. One source can be the data-free and distribution-free properties of random

variables. Typically, the properties emanate from general statistical *knowledge* of random variables, which requires no input from any sample data, and as few assumptions as possible about the data generating process. The second source is the specific *information* relevant to the particular problem setting, which may be derived from either historical data or institutional knowledge. Existing uncertainty sets have utilized different kinds of information, which often appear as parameters in the formulations; for example, the range in the budget uncertainty set, or the mean and covariance in the ellipsoidal uncertainty set. However, these statistics contain limited information and may lose some useful distributional information. Therefore, it is important to conceive uncertainty sets that can incorporate richer information from data.

1.4. Our contribution

The focus of this study is to explore the characterization of random variables and utilize it to design a new uncertainty set. We seek to answer the following questions: What are the data-free and distribution-free statistical characteristics of the collective behavior of random variables that may be utilized to refine the uncertainty set? How can we design an uncertainty set that captures rich distributional information while still resulting in a tractable linear programming formulation? Is it possible to construct an uncertainty set that offers the ability to adjust the level of uncertainty in each dimension separately rather than a single parameter that affects all dimensions in the same way? If we can construct a new uncertainty set, then under what conditions can the new uncertainty set outperform other existing approaches? Our main results are as follows:

1. We use the Probability Integral Transform to show that if the random variables Z_j's are continuous and mutually independent of each other, then the order statistics of the cumulative distribution functions (CDFs) of Z_j's follow the Beta distribution. For a given probability, each order statistic of the CDFs of random variables Z_j's has a high-density interval within the range [0,1]. Based on this data-free distribution-free property of CDFs of random variables Z_j's, we construct an order statistic uncertainty set by imposing constraints on order statistics of the CDFs of random variables Z_j's.
2. To embed the CDFs of random variables in the RO model with the order statistic uncertainty set, we utilize the quantiles of random variables, which carry rich distributional information of random variables. Because the order statistics of the CDFs of Z_j's have $|J|!$ possible outcomes, the constraints for them imply $|J|!$ implicit linear constraints. In order to develop a tractable linear formulation for the $|J|!$ implicit linear constraints, we adopt the formulation of the assignment problem.
3. We demonstrate the geometric flexibility of the order statistic uncertainty set and prove that the RO model with the order statistic uncertainty set reduces to the RO model with the interval uncertainty set, or the budget uncertainty set, or the demand uncertainty set if its parameters are selected suitably. This shows that the order statistic uncertainty set has a greater modeling power because the RO model with the order statistic uncertainty set provides a framework that incorporates these three uncertainty sets as special cases.
4. The order statistic uncertainty set captures rich information about distributions because it utilizes the quantiles of distributions to characterize uncertainties. Different quantile values for different random variables are used in the uncertainty set, and the uncertainties for different random variables can be depicted separately. The quantile levels in the order statistic uncertainty set determine the probabilistic guarantee for the constraint feasibility. We illustrate how we can achieve different trade-offs between the objective performance and constraint feasibility by choosing different quantile levels.
5. We apply the order statistic uncertainty set and several competing characterizations of the uncertainty set to a robust portfolio construction problem and compare and contrast their relative performance. Results of our numerical experiments show that when the correlation of portfolio returns is low, the order statistic uncertainty set outperforms the budget uncertainty set, the interval uncertainty set and the convex hull uncertainty set. Additionally, if the correlation of portfolio returns is low and the sample size is small, then the order statistic uncertainty set has better performance than the tail uncertainty set and the ellipsoidal uncertainty set as well.

The outline of the paper is as follows. In Section 2, we present the motivation to construct the order statistic uncertainty set and provide a linear formulation of the RO model with the order statistic uncertainty set. In Section 3, we show that the RO models with the interval uncertainty set, or the budget uncertainty set, or the demand uncertainty set can be obtained as special cases of the RO model with the order statistic uncertainty set. In Section 4, we derive the probabilistic bound for constraint feasibility for the RO model with the order statistic uncertainty set, and discuss how to determine the parameters for the order statistic uncertainty set. In Section 5, we solve a robust portfolio construction problem with the order statistic uncertainty set and other existing uncertainty sets, and compare their relative performance. Finally, Section 6 summaries our results.

2. The order statistic uncertainty set

In this section, we use the Probability Integral Transformation to derive a distribution-free and data-free property of random variables Z_j's, and then use the property to construct the order statistic uncertainty set. In Section 2.2, we present a linear formulation of the RO model with the new uncertainty set. All mathematical proofs can be found in the appendix.

2.1. Motivation: order statistics

Suppose the random variables Z_j's are continuous and independently distributed in the range [0,1], each following an arbitrary continuous distribution with an unknown cumulative distribution function F_j. Denote random variables $U_j = F_j(Z_j), \forall j \in J$, and each U_j can be shown to be uniformly distributed over [0,1] (see Roussas, 1997, Section 9.4). Denote the order statistics of U_j's as $U_{(1)}, \ldots, U_{(|J|)}$, which is the rearranged sequence of U_j's with the k-th order statistic $U_{(k)}$ being the k-th smallest among them. Different from the random variable U_j's, the order statistic random variable $U_{(k)}$ follows $Beta(k, |J| + 1 - k)$ distribution (see Gut, 2009, Chapter 4.1) instead of the uniform distribution. The mapping from $\{Z_1, \ldots, Z_{|J|}\}$ to $\{U_{(1)}, \ldots, U_{(|J|)}\}$ is illustrated in Fig. 1.

To motivate our approach, we discuss an example with $|J| = 20$ random variables. Fig. 2 shows the $Beta(k, |J| + 1 - k)$ distribution of different $U_{(k)}$'s, $\forall k = 1, \ldots, 20$. There are two observations worth noting.

1. If k is small, the distribution of $U_{(k)}$ tends to be right skewed, which means the $U_{(k)}$ variable tends to be small. As k increases, the distribution of $U_{(k)}$ gets more skewed to the left. Most $U_{(k)}$'s are extremely unlikely to be either 0 or 1.
2. Each order statistic $U_{(k)}$ has a high-density interval strictly smaller than [0,1], over which the area under its probability density function is close to 1. For example, the area un-

Variables	Distribution
$Z_1, Z_2, \ldots, Z_{\|J\|}$	Z_js are independent with arbitrary continuous distributions
via cdf transformation	
$U_1 = F_1(Z_1), U_2 = F_2(Z_2), \ldots, U_{\|J\|} = F_{\|J\|}(Z_{\|J\|})$	$U_j \sim$ Uniform(0,1), $\forall j$
via ordering from smallest to largest	
$U_{(1)}, U_{(2)}, \ldots, U_{(\|J\|)}$	$U_{(k)} \sim Beta(k, \|J\| + 1 - k)$

Fig. 1. Transformations of variables.

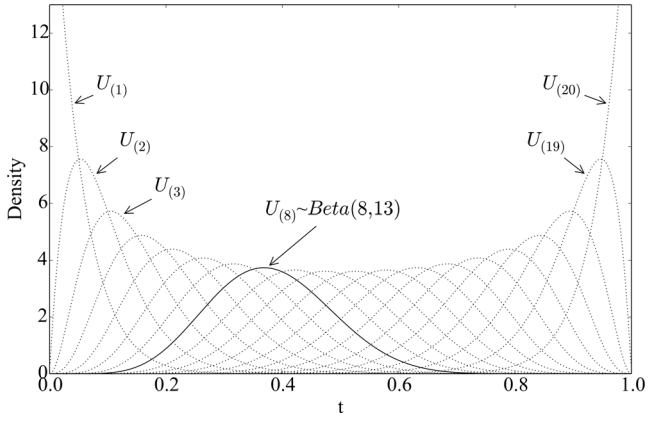

Fig. 2. Probability density functions of order statistics $U_{(k)}$'s for $|J| = 20$.

der the 8th order statistic's probability density function (the solid line) over the interval [0.05,0.85] is 0.999997, which is almost 1. This illustrates that the uncertainty characterization with either the box or the budget uncertainty sets may be too extreme because in the robust solution with these two uncertainty sets, at least $|J| - 1$ random variables Z_j's are equal to either 0 or 1 and such scenarios are not contained in the high-density interval [0.05,0.85].

From above, we see that regardless of the unknown cumulative distributions F_j of the random variable Z_j, the order statistics of $F_j(Z_j)$'s always follow the Beta distribution. Denote the cumulative distribution function for Beta$(k, |J| + 1 - k)$ distribution as $I_t(k, |J| + 1 - k)$, and its quantile function $Q_k^t = \inf\{\tau : I_\tau(k, |J| + 1 - k) = t\}$. Therefore, given any $\varepsilon'_k \in [0, 1]$, we can find the lower limit $Q_k^{\varepsilon'_k}$ for $U_{(k)}$ such that $Prob(U_{(k)} \le Q_k^{\varepsilon'_k}) = \varepsilon'_k$. Similarly, we can find the upper limit $Q_k^{1-\varepsilon_k}$ for $U_{(k)}$ such that $Prob(U_{(k)} \le Q_k^{1-\varepsilon_k}) = 1 - \varepsilon_k$. Denote $\boldsymbol{\varepsilon'}$ as the vector of values $\varepsilon'_1, \varepsilon'_2, \ldots, \varepsilon'_{|J|}$, and $\boldsymbol{\varepsilon}$ of values $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{|J|}$, where $0 \le \varepsilon'_j, \varepsilon_j \le 1, \forall j \in J$. We then construct the uncertainty set $\mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon})$ based on the order statistics of the CDFs of random variables as follows:

$$\mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon}) = \Big\{\mathbf{Z} : F_j(Z_j) = U_j,\ \forall j \in J,\ \text{and } Q_k^{\varepsilon'_k} \le U_{(k)} \le Q_k^{(1-\varepsilon_k)},\ \forall k \in J\Big\}. \tag{4}$$

In the above uncertainty set, the random variables Z_j's are restricted so that the order statistics $U_{(k)}$'s belong to the high-density regions. Note that $U_{(1)} \le U_{(2)} \le \ldots \le U_{(|J|)}$ is always implied by definition. Although ε'_k can be any value in the range [0,1], we only need to consider ε'_k's such that $Q_{k-1}^{\varepsilon'_{k-1}} \le Q_k^{\varepsilon'_k}$, for $k = 2, 3, \ldots, |J|$. To argue for this, we first assume that there exists a k_0 such that $Q_{k_0-1}^{\varepsilon'_{k_0-1}} > Q_{k_0}^{\varepsilon'_{k_0}}$. Then for any $U_{(k_0)}$ that satisfies $Q_{k_0}^{\varepsilon'_{k_0}} \le U_{(k_0)} \le Q_{k_0}^{(1-\varepsilon_{k_0})}$ and $U_{(k_0)} < Q_{k_0-1}^{\varepsilon'_{k_0-1}}$, we would have $U_{(k_0-1)} \ge Q_{k_0-1}^{\varepsilon'_{k_0-1}} > U_{(k_0)}$, which violates $U_{(k_0-1)} \le U_{(k_0)}$. For a similar reason, we also assume $Q_{k-1}^{(1-\varepsilon_{k-1})} \le Q_k^{(1-\varepsilon_k)}$, for $k = 2, 3, \ldots, |J|$.

We emphasize that all the properties that we have utilized above only rely on the assumption that the continuous distributions of Z_j's are independent. These properties are distribution-free because they hold regardless of the distributions F_j's of the random variables of Z_j's. These properties are also data-free because they are not based on any information extracted from data.

2.2. Robust model with order statistic uncertainty set

To use the uncertainty set $\mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon})$ in the robust model (3), we need to study the subproblem $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon}))$, which is $\max_{\mathbf{Z} \in \mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon})} \sum_{j \in J} \hat{a}_j \cdot |x_j| \cdot Z_j$. The following characterization of $\mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon})$ helps to reformulate the subproblem $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon}))$.

Proposition 2.1. *Given $\boldsymbol{\varepsilon'}$, $\boldsymbol{\varepsilon}$, and a fixed $\mathbf{x}$, $U_{(k)} = Q_k^{(1-\varepsilon_k)}, \forall k$ hold in the optimal solution to $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon}))$.*

Proposition 2.1 states that the order statistics of $F_j(Z_j)$'s are equal to their upper bounds in the optimal solution to $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon}))$. Note that in the optimal solution to $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon}))$, the variable Z_j should be as large as possible because Z_j's coefficient $\hat{a}_j \cdot |x_j|$ is non-negative. Because F_j is non-decreasing, then $F_j(Z_j)$ should also be as large as possible. As a result, each order statistic $U_{(k)}$ of $F_j(Z_j)$'s should be as large as its upper bound $Q_k^{(1-\varepsilon_k)}$. Based on Proposition 2.1, we know that the uncertainty set $\mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon})$ can be substituted by the following order statistic uncertainty set.

$$\mathcal{U}^{OS}(\boldsymbol{\varepsilon}) = \Big\{\mathbf{Z} : F_j(Z_j) = U_j,\ \forall j \in J,\ \text{and } U_{(k)} \le Q_k^{(1-\varepsilon_k)},\ \forall k \in J\Big\}. \tag{5}$$

Proposition 2.1 shows that the lower bounds in $\mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon})$ are irrelevant to the optimal solution for $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon}))$, and the lower bounds can be dropped in the definition of $\mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon})$. In the optimal solution for $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$, we must also have $U_{(k)} = Q_k^{(1-\varepsilon_k)}$. Therefore, we know the optimal objective value for $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon'}, \boldsymbol{\varepsilon}))$ must be equal to the optimal objective value for $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$.

The order statistic uncertainty set $\mathcal{U}^{OS}(\boldsymbol{\varepsilon})$ is intractable in its current form for three reasons. The first reason is that the uncertainty set $\mathcal{U}^{OS}(\boldsymbol{\varepsilon})$ is not directly defined on the random variable Z_j, but is constructed with constraints on the cumulative distribution functions of Z_j's. Another reason is that there are $|J|!$ permutations of $F_j(Z_j)$'s for all possible outcomes of $U_{(k)}$'s, which makes reformulating $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$ challenging. The third reason has to do

with the nonconvexity of the order statistic uncertainty set $\mathcal{U}^{OS}(\boldsymbol{\varepsilon})$ as stated in the following proposition.

Proposition 2.2. *If there exist k_1 and k_2 ($1 \le k_1 < k_2 \le |J|$) such that $Q_{k_1}^{(1-\varepsilon_{k_1})} \ne Q_{k_2}^{(1-\varepsilon_{k_2})}$, then the uncertainty set $\mathcal{U}^{OS}(\boldsymbol{\varepsilon})$ is not convex.*

To overcome the difficulty in reformulating $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$, we utilize the assignment formulation. The technique that we use is similar to the averaging function for the ordered weighted averaging approach (see Chassein & Goerigk, 2015). Let q_{jk} be Z_j's quantile of order $Q_k^{(1-\varepsilon_k)}$, i.e., $q_{jk} = \inf\{x : F_j(x) \ge Q_k^{(1-\varepsilon_k)}\}, \forall j, k \in J$. The following proposition provides a tractable formulation for the problem $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$.

Proposition 2.3. *For a fixed $\mathbf{x}$, the optimal objective value for $\beta\left(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon})\right)$ is equal to the optimal objective value for the following linear optimization problem:*

$$\max_{\eta} \sum_{j\in J} \hat{a}_j |x_j| \cdot \left(\sum_{k\in J} q_{jk}\eta_{jk} \right) \tag{6a}$$

$$\text{s.t.} \sum_{k} \eta_{jk} = 1, \forall j \in J \tag{6b}$$

$$\sum_{j} \eta_{jk} = 1, \forall k \in J \tag{6c}$$

$$\eta_{jk} \ge 0, \forall j, k \in J. \tag{6d}$$

The problem (6) in Proposition 2.3 is the linear relaxation of the maximum weight assignment problem, which is known to have an integer optimal solution. If $\eta_{jk} = 1$, then $\hat{a}_j|x_j|$ is assigned to q_{jk}, which implies $Z_j = q_{jk}$ and $F_j(Z_j) = Q_k^{(1-\varepsilon_k)}$. The integer optimal solution of problem (6) will map the set $\{F_j(Z_j), \forall j \in J\}$ to the set $\{Q_k^{(1-\varepsilon_k)}, \forall k \in J\}$.

The RO Model (3) with uncertainty set $\mathcal{U}^{OS}(\boldsymbol{\varepsilon})$ becomes the following problem (7). We follow the procedure in Bertsimas & Sim (2004) to reformulate the following model to a linear optimization model (we add back the index i for the constraints).

$$\max_{\mathbf{x}} \ \mathbf{c}'\mathbf{x} \tag{7a}$$

$$\text{s.t.} \sum_{j\in J_i} a_{ij}x_j + \max_{\mathbf{Z}_i \in \mathcal{U}_i^{OS}(\boldsymbol{\varepsilon})} \sum_{j\in J_i} \hat{a}_{ij} \cdot |x_j| \cdot Z_{ij} \le b_i, \ \forall i \in I \tag{7b}$$

$$\underline{\mathbf{x}} \le \mathbf{x} \le \bar{\mathbf{x}}. \tag{7c}$$

Theorem 2.4. *Model (7) is equivalent to the following linear programming problem:*

$$\max \sum_{j} c_j x_j \tag{8a}$$

$$\text{s.t.} \sum_{j\in J_i} a_{ij}x_j + \sum_{j\in J_i} (\theta_{ij} + \phi_{ij}) \le b_i, \ \forall i \in I \tag{8b}$$

$$-y_j \le x_j \le y_j, \forall j \tag{8c}$$

$$\underline{x}_j \le x_j \le \bar{x}_j, \forall j \tag{8d}$$

$$\theta_{ij} + \phi_{ik} \ge \hat{a}_{ij} q_{ijk} y_j, \forall j, k \in J_i, \forall i \in I \tag{8e}$$

$$y_j \ge 0, \forall j \tag{8f}$$

We leverage the strong duality to obtain the linear formulation (8) by replacing the maximizing problem in constraints (7b) with the dual of problem (6). Because Model (8) has more variables and constraints than the Model (7) in Bertsimas & Sim (2004), the computational complexity of the RO Model (8) is generally higher than the RO model with the budget uncertainty set.

3. Special cases

Using the order statistic uncertainty set in the RO model has several advantages. First, the quantile in the order statistic uncertainty set is a robust statistic and is less sensitive to extreme observations. The quantiles also contain richer information about the uncertainty of a random variable than some other statistics, e.g., the range.

In the following, we demonstrate the modeling power of the order statistic uncertainty set by showing that the RO model with the order statistic uncertainty set incorporates three existing uncertainty sets as special cases. Specifically, we show that with suitable parameters, the RO model with the order statistic uncertainty set reduces to the RO models with the interval uncertainty set, the budget uncertainty set, and the demand uncertainty set.

3.1. Special cases: the interval and the budget uncertainty set

Although motivated by different statistical properties, the order statistic uncertainty set has a close relationship with the interval and the budget uncertainty sets. We illustrate it with a numerical example with $|J| = 7$. The typical structures of the Z_j values in the optimal solutions for different uncertainty sets are shown in Fig. 3. In each figure, the values of Z_j's are ordered from the smallest to the largest. In the optimal solution of the RO model with the interval uncertainty set, all Z_j's are equal to 1. In the optimal solution of the RO model with the budget uncertainty set, up to one of Z_j's is fractional and all others are either 0 or 1.

In the optimal solution of the RO model with the order statistic uncertainty set, the values of Z_j's are fractions $q_{j_1,1}, q_{j_2,2}, \ldots, q_{j_{|J|},|J|}$, where $j_1, j_2, \ldots, j_{|J|}$ is a permutation of $1, 2, \ldots, |J|$. Each of the quantile values determines the level of uncertainty for the corresponding random variable. The $|J|$ fractional values $q_{j_1,1}, q_{j_2,2}, \ldots, q_{j_{|J|},|J|}$ can be completely different from each other. Moreover, for any particular k, the fractional value $q_{j_k,k}$ has up to $|J|$ possible outcomes because it depends on the index j_k. The values of $q_{j_1,1}, q_{j_2,2}, \ldots, q_{j_{|J|},|J|}$ altogether have up to $|J|!$ outcomes, depending on the sequence $j_1, j_2, \ldots, j_{|J|}$. The geometric shape of the order statistic uncertainty set can be flexibly adjusted by the $|J|^2$ parameters.

The RO models with the interval and the budget uncertainty sets are both special cases of the RO model with the order statistic uncertainty set. If we choose suitable values for the parameters in the RO model with the order statistic uncertainty set, we can obtain the RO models with the other two uncertainty sets.

1. The RO model with the order statistic uncertainty set reduces to the RO model with the interval uncertainty set if we choose $q_{jk} = 1, \forall k, j \in J$. The shape of the interval uncertainty set is not adjustable.
2. The RO model with the order statistic uncertainty set reduces to the RO model with the budget uncertainty set with budget Γ if we choose q_{jk} as follows: $q_{jk} = 0$, if $1 \le k \le |J| - \lfloor\Gamma\rfloor - 1, \forall j \in J$; $q_{jk} = \Gamma - \lfloor\Gamma\rfloor$, if $k = |J| - \lfloor\Gamma\rfloor, \forall j \in J$; $q_{jk} = 1$, if $|J| - \lfloor\Gamma\rfloor + 1 \le k \le |J|, \forall j \in J$. In Appendix E, we prove that the RO model with the order statistic uncertainty set with such a choice of q_{jk} values is equivalent to the RO model with the budget uncertainty set with budget Γ (the

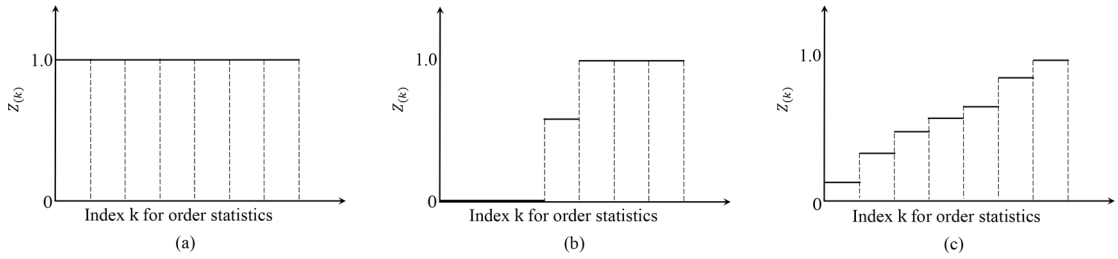

Fig. 3. Order statistics of Z_j's for different uncertainty sets – (a) interval uncertainty set, (b) budget uncertainty set, (c) order statistic uncertainty set.

Problem (4) in Bertsimas & Sim, 2004). In the optimal solution of the RO model with the budget uncertainty set, at most one of Z_j's can be a fraction, and all other Z_j's are either 0 or 1 (Bertsimas & Sim, 2004, Section 3). The geometric flexibility of the budget uncertainty set is limited because a small change in the value of the budget will only change the value of the fractional Z_j, but not any other Z_j's.

3.2. Special case: the demand uncertainty set

Next, we show that the RO model with the demand uncertainty set that has $2^{|J|}-1$ constraints for the Z_j's can be obtained from the RO model of the order statistic uncertainty set. Recall that the demand uncertainty set is as follows:

$$\mathcal{U}^D = \left\{ \mathbf{Z} : \left| \frac{\sum_{j\in S} Z_j}{|S|^{1/\alpha}} \right| \le \gamma, \forall S \subseteq J \right\}. \tag{9}$$

In the literature, α is often assumed to be in the range (1,2], and $\gamma \ge 0$. The following proposition characterizes the optimal solution to the problem $\beta(\mathbf{x}, \mathcal{U}^D)$ and connects the demand uncertainty set with the order statistic uncertainty set.

Proposition 3.1. *For a fixed* $\mathbf{x}$, $Z^*_{(k)} = \gamma \cdot (|J|+1-k)^{1/\alpha} - \gamma \cdot (|J|-k)^{1/\alpha}, \forall k \in J$ *hold in the optimal solution for* $\beta(\mathbf{x}, \mathcal{U}^D)$.

We explain the key idea of the above result. Notice that the objective function of $\beta(\mathbf{x}, \mathcal{U}^D)$ is the sum of the pairwise products of two sequences $\hat{a}_j|x_j|$'s and Z_j's. Because of the rearrangement inequality (see Cvetkovski 2012, Theorem 6.1), in the optimal solution of $\beta(\mathbf{x}, \mathcal{U}^D)$, the two sequences should be in the same order (both non-decreasing or non-increasing). Thus the largest $\hat{a}_j|x_j|$ should be paired with the largest Z_j, and the second largest of the two sequences should also be paired with each other, and so on. Because the $\hat{a}_j|x_j|$'s are non-negative, Z_j's should be as large as possible in the optimal solution for $\beta(\mathbf{x}, \mathcal{U}^D)$. Moreover, we should make the largest Z_j to be as large as possible because it is paired with the largest $\hat{a}_j|x_j|$. The largest of Z_j is restricted by the constraints with $|S|=1$, so the largest of Z_j should be equal to γ. The second largest Z_j is restricted by the constraints with $|S|=2$, so the second largest Z_j should be equal to $\gamma \cdot (2^{1/\alpha}-1)$. The remaining Z_j's can be analyzed similarly.

For the demand uncertainty set, the value of $Z^*_{(k)} = \gamma \cdot (|J|+1-k)^{1/\alpha} - \gamma \cdot (|J|-k)^{1/\alpha}$ only depends on the index k for the order statistic but does not depend on the index j of Z_j's. In contrast, the values of Z_j's for the order statistic uncertainty set (i.e., $q_{j_1,1}, q_{j_2,2}, \ldots, q_{j_{|J|},|J|}$) depend on both indices. This shows that the demand uncertainty set has less geometric flexibility than the order statistic uncertainty set because it does not capture the potential heterogeneity of different distributions of Z_j's.

The following corollary provides an equivalent formulation for the demand uncertainty set. It requires $|J|^2$ continuous variables, and $|J|^2 + 2\cdot|J|$ constraints, which is much less than the $2^{|J|}-1$ constraints in $\mathcal{U}^D$ if $|J|$ is large.

Corollary 3.2. *For a fixed* $\mathbf{x}$, *the optimal objective value for problem* $\beta\left(\mathbf{x}, \mathcal{U}^D\right)$ *is equal to the optimal objective value for the following linear optimization problem:*

$$\max_{\eta} \sum_{j\in J} \hat{a}_j|x_j| \cdot \left(\sum_{k\in J} \gamma \cdot \left(k^{1/\alpha} - (k-1)^{1/\alpha}\right) \cdot \eta_{jk} \right) \tag{10a}$$

$$\text{s.t.} \sum_k \eta_{jk} = 1, \forall j \in J \tag{10b}$$

$$\sum_j \eta_{jk} = 1, \forall k \in J \tag{10c}$$

$$\eta_{jk} \ge 0, \forall j,k \in J \tag{10d}$$

For the problem $\beta(\mathbf{x}, \mathcal{U}^D)$, we need to find the maximum sum of the pairwise products of the sequence $\hat{a}_j|x_j|, \forall j \in J$ and the sequence $\gamma \cdot (k^{1/\alpha} - (k-1)^{1/\alpha}), \forall k \in J$. Corollary 3.2 ensures that the elements of the two sequences are one-to-one paired with the assignment formulation, and the two sequences must be paired in the same order in the optimal solution of the problem (10). The robust optimization model with the demand uncertainty set, i.e., Model (3) with $\mathcal{U}^D_i$ (instead of $\mathcal{U}_i$), can be reformulated by taking duality of Model (10). The reformulated model is as follows:

$$\max \sum_j c_j x_j \tag{11a}$$

$$\text{s.t.} \sum_{j\in J_i} a_{ij}x_j + \sum_{j\in J_i} (\theta_{ij} + \phi_{ij}) \le b_i, \ \forall i \tag{11b}$$

$$-y_j \le x_j \le y_j, \forall j \tag{11c}$$

$$\underline{x}_j \le x_j \le \bar{x}_j, \forall j \tag{11d}$$

$$\theta_{ij} + \phi_{ik} \ge \hat{a}_{ij} \cdot y_j \cdot \gamma \cdot \left(k^{1/\alpha} - (k-1)^{1/\alpha}\right), \forall j,k \in J_i, \forall i \tag{11e}$$

$$y_j \ge 0, \forall j. \tag{11f}$$

4. Further analysis of the order statistic uncertainty set

In this section, we first derive the probabilistic guarantee of the order statistic uncertainty set to ensure the constraint feasibility under uncertainty. We then discuss how one may estimate the quantile values (q_{jk}'s) that we introduced in Section 2.2 from available data.

4.1. Probability of constraint feasibility

Denote the optimal solution to Model (7) as $\mathbf{x}^*$; the constraint index i is dropped in this section for the ease of exposition, i.e., we study a single constraint in (3b). We are interested in the probabilistic guarantee of constraint feasibility that the order statistic uncertainty set can provide, i.e., $\text{Prob}\left(\sum_{j\in J} A_j \cdot x_j^* \le b\right)$. In the following, we prove a probabilistic guarantee for the case when the random variables are independently and symmetrically distributed. The probabilistic guarantee is expressed by a formula derived in Steck (1971), which gives the probability of order statistics of the uniform distribution lying in a multi-dimensional rectangle.

Proposition 4.1 (This is a restatement of the first theorem in Steck, 1971)**.** *Let Δ be the $|J| \times |J|$ matrix whose (i, j)th element is given as:*

$$\Delta_{ij} = \begin{cases} \left(Q_i^{(1-\varepsilon_i)}\right)^{j-i+1}/(j-i+1)!, & j-i+1 \ge 0 \\ 0, & j-i+1 < 0. \end{cases} \tag{12}$$

We have

$$\text{Prob}\left(U_{(k)} \le Q_k^{(1-\varepsilon_k)}, k = 1, \ldots, |J|\right) = |J|!\det[\Delta]. \tag{13}$$

Theorem 4.2. *If the continuous variable A_j is independently and symmetrically distributed in $[a_j - \hat{a}_j, a_j + \hat{a}_j], \forall j \in J$, then the order statistic uncertainty set $\mathcal{U}^{OS}(\boldsymbol{\varepsilon})$ implies a probabilistic guarantee of at least $\frac{1}{2} + \frac{1}{2} \cdot |J|!\det[\Delta]$ for the constraint feasibility, i.e.,* $\text{Prob}(\sum_{j\in J} A_j \cdot x_j^* \le b) \ge \frac{1}{2} + \frac{1}{2} \cdot |J|!\det[\Delta]$.

According to the above result, the upper limits of the order statistics of the CDFs of random variables $Q_k^{(1-\varepsilon_k)}$'s determine the probabilistic guarantee of the order statistic uncertainty set. If we need a high probabilistic guarantee, then we should set ε_k's to be small because smaller ε_k's lead to larger $Q_k^{(1-\varepsilon_k)}$ values, which increases $\text{Prob}(U_{(k)} \le Q_k^{(1-\varepsilon_k)}, k = 1, \ldots, |J|)$.

4.2. Estimating parameters in the order statistic uncertainty set

The RO model (8) with the order statistic uncertainty set requires the quantiles q_{jk}'s as inputs. In practice, if there is no historical data, decision makers may choose these parameters based on institutional knowledge. If there is data, then the quantiles q_{jk}'s can be specified based on data with the following procedure.

We first need to specify the values of $\varepsilon_1, \ldots, \varepsilon_{|J|}$ in (5). For convenience, we can let all ε_k's be equal to each other and denote them as $\varepsilon_1 = \ldots = \varepsilon_{|J|} = \varepsilon$, which can range between 0 and 1. As we will show in our numerical experiments, such parameter configuration can lead to good performance. The ε value controls the size of the uncertainty set, which then determines the trade-off between the objective value and the probability of constraint feasibility. If ε is large (small), then the size of the uncertainty set is small (large, respectively), and consequently, we would get a large (small, respectively) objective value and low (high, respectively) probability of constraint feasibility. With the trial-and-error search, we can find an ε value such that the corresponding optimal solution achieves a particular target probability of constraint feasibility.

For a given ε value, we can determine $Q_1^{1-\varepsilon}, Q_2^{1-\varepsilon}, \ldots, Q_{|J|}^{1-\varepsilon}$ based on the quantile function $Q_k^t = \inf\{\tau : I_\tau(k, |J| + 1 - k) = t\}$. Then we can estimate each random variable Z_j's quantile of order $Q_k^{(1-\varepsilon_k)}$ according to the definition $q_{jk} = \inf\{x : F_j(x) \ge Q_k^{(1-\varepsilon_k)}\}, \forall j, k \in J$. Suppose we have N samples of Z_j denoted as $z_{j1}, \ldots, z_{jN}$, and we can use the simple random sampling to get the following estimation of q_{jk}'s.

$$q_{jk} = \min\left\{z_{jm} : \frac{\sum_{n=1}^{N} \mathbb{1}_{z_{jn} \le z_{jm}}}{N} \ge Q_k^{(1-\varepsilon)}, \forall m \in N\right\}. \tag{14}$$

The simple random sampling estimator is asymptotically normal, and the asymptotic variance could be reduced by various variance reduction approaches (see Glasserman et al., 2000), including stratified sampling, importance sampling, etc. Note that the above method provides quantile estimations with discontinuities. To resolve this issue, we can apply interpolation or smoothing techniques (see Dielman et al., 1994).

5. Numerical experiments

In this section, we apply our method to a portfolio construction problem and compare the performance of the order statistic uncertainty set with other existing uncertainty sets. All models were implemented in Python and solved with CPLEX 22.1. The computational experiments were performed on a Unix PC equipped with 2.4 GHz dual-Core Intel Core i5 processors and 8 GB memory. All problem instances discussed in this section were solved optimally using CPLEX's default setting.

5.1. A robust portfolio construction problem

Suppose we have one unit of asset to invest among $|J|$ portfolios $1, \ldots, |J|$. We model each portfolio j's return as an independent random variable distributed symmetrically in an interval $[r_j - \hat{r}_j, r_j + \hat{r}_j]$. The portfolio j's return can be denoted as $\tilde{r}_j = r_j + \rho_j \hat{r}_j$ with $-1 \le \rho_j \le 1$. Denote $Z_j = |\rho_j|, \forall j \in J$. Suppose we allocate x_j of the unit asset in portfolio j, and the goal is to keep the return high and make the associated risk low.

Suppose we have N samples $\rho_{j1}, \ldots, \rho_{jN}$ for each ρ_j. The following problem maximizes the worst-case return V with respect to an uncertainty set $\mathcal{U}$ while enforcing the expected return to be a given level r.

$$\max_{\boldsymbol{x}} \; V \tag{15a}$$

$$\text{s.t.} \; \sum_{j\in J} x_j\left(r_j - Z_j\hat{r}_j\right) \ge V, \; \forall \, \mathbf{Z} \in \mathcal{U}, \tag{15b}$$

$$\sum_{n=1}^{N}\sum_{j\in J} x_j\left(r_j + \rho_{jn}\hat{r}_j\right)/N = r, \tag{15c}$$

$$\sum_{j\in J} x_j = 1, \tag{15d}$$

$$0 \le x_j \le 1, \forall j \in J. \tag{15e}$$

Denote the optimal solution of problem (15) as $\boldsymbol{x}^*$ and V^*. For any realization of $\boldsymbol{\rho}$ within the uncertainty set $\mathcal{U}$ in the constraint (15b), the corresponding portfolio return will be no less than V^*. For realizations of $\boldsymbol{\rho}$ that are not in the uncertainty set $\mathcal{U}$, the portfolio return could be less than V^*. Therefore, there exists an $\omega \in [0, 1]$ such that

$$\textbf{Prob}\left(\sum_{j\in J} x_j^* \cdot \tilde{r}_j \le V^*\right) = \omega. \tag{16}$$

According to the relation (16), the ω value is the probability that the portfolio return is no greater than V^*. This means that the negative of the optimal objective value (i.e., $-V^*$) is the Value-at-Risk (VaR) of level ω because the VaR of level ω is defined as the minimum value such that the probability of a loss exceeds VaR is at most ω. Therefore, the problem (15) can be viewed as the

Robust Mean-Value at Risk (Robust-MVaR) portfolio optimization model. Using the trial-and-error procedure that we discussed in Section 4.2, we can solve the problem (15) with different parameters for the uncertainty set $\mathcal{U}$ and find the one such that the level ω for the corresponding optimal solution is equal to our target risk level (say, $\omega = 0.1$ or 0.05). For example, if the level ω for the optimal solution is larger (smaller) than what we want, then we can reduce (increase, respectively) it by increasing (reducing, respectively) the size of the uncertainty set $\mathcal{U}$ so that the optimal solution can ensure a higher (lower, respectively) probabilistic guarantee on the portfolio return being no less than the objective value.

In our numerical experiment, we solve problem (15) with the order statistic uncertainty set, the budget uncertainty set, the interval uncertainty set (see Chassein et al., 2019), the tail uncertainty set, the convex hull uncertainty set (see Chassein et al., 2019) and the ellipsoidal uncertainty set. We generate portfolio returns based on the value-weighted portfolio dataset in the 17 Industry Portfolios Daily category from Kenneth French's website (see French, 2021). This dataset has daily returns for 17 portfolios from July 01 1926 to October 29, 2021. With the dataset, we can estimate the sample mean μ_j and sample standard deviation σ_j for each portfolio j $(j = 1, \ldots, 17)$.

In the following numerical experiments, we aim to compare the performances of the order statistic uncertainty set and the other five uncertainty sets, as well as investigate how the sample size and the correlation of random variables affect the performances of uncertainty sets.

5.2. Probability bound of constraint feasibility

In this section, we evaluate the probability bound of constraint feasibility in Theorem 4.2 by comparing it with the posteriori empirical probability of constraint feasibility. We let $|J|$ take 4 different values 2, 4, 6 and 8. For each of the $|J|$ values, we let ε take different values between 0 and 1 and compare the probability bound of constraint feasibility and the posteriori empirical probability of constraint feasibility.

For each $|J|$ and each ε, we can calculate the probability bound of constraint feasibility using $\frac{1}{2} + \frac{1}{2} \cdot |J|! \det[\Delta]$ based on Theorem 4.2. In order to evaluate the posteriori empirical probability of constraint feasibility, we make use of Model (15). Specifically, for each $|J|$ and each ε, we use the first $|J|$ portfolios of the 17 portfolios (introduced in Section 5.1) to evaluate the posteriori empirical probability of constraint feasibility with 100 repetitions. In each repetition, we generate the in-sample dataset consisting of 1000 samples of the $|J|$ portfolios' returns drawn from the multivariate normal distribution with mean $(\mu_1, \ldots, \mu_{|J|})$ and covariance matrix whose (i, j)th entry is σ_i^2 if $i = j$ and 0 if $i \neq j$. With this in-sample dataset, we solve Model (15) to obtain the optimal solution $\boldsymbol{x}^*$ and V^*. Then we generate the out-of-sample dataset consisting of 10^6 samples of the $|J|$ portfolios' returns drawn from the same distribution that was used to generate the in-sample dataset. With the optimal solution $\boldsymbol{x}^*$, V^* and the out-of-sample dataset, we use the formula (16) to calculate the ω value. Then the posteriori empirical probability of constraint feasibility can be obtained as $1 - \omega$. The mean out-of-sample posteriori empirical probabilities of constraint feasibility (over the 100 repetitions) are shown in Fig. 4. The figure also has the probability bound of constraint feasibility for each $|J|$ and each ε.

For the case $|J| = 2$, the gap between the probability bound and the posteriori empirical probability of constraint feasibility is relatively small. As $|J|$ increases to 4, 6 and 8, the gap gets larger. Therefore, the probability bound in Theorem 4.2 is more useful for the cases with small $|J|$ values. Due to the gap between the probability bound and the posteriori empirical probability of constraint feasibility, we use the posteriori empirical probability of constraint feasibility to evaluate the probability level of constraint feasibility in the following numerical experiments.

5.3. The effect of sample size and correlation

In this section, we compare the performances of the order statistic uncertainty set and five other uncertainty sets for different sample sizes and different degrees of correlation between portfolio returns. We let $\omega = 0.1$ and set the target return r in constraint (15c) to be the median of $\mu_1, \ldots, \mu_{17}$.

We conduct the experiment with 400 repetitions. In each repetition, we generate the in-sample dataset consisting of N samples of 17 portfolios' returns drawn from the multivariate normal distribution with mean $(\mu_1, \ldots, \mu_{17})$ and covariance matrix whose (i, j)th entry is σ_i^2 if $i = j$ and $\overline{\rho} \cdot \sigma_i \cdot \sigma_j$ if $i \neq j$. We will let $\overline{\rho} = 0.05, 0.5$ and 0.95, corresponding to low-correlation, medium-correlation and high-correlation regimes. Using this in-sample dataset, we estimate the parameters for each of the six uncertainty sets, e.g., the range for the budget uncertainty set, the correlation matrix for the ellipsoidal uncertainty set, etc. Note that we used linear interpolation to avoid the discontinuity of the quantile estimator in (14). Then by adjusting the parameter of each uncertainty set that controls the size of the uncertainty set (e.g., adjusting the ε value for $\mathcal{U}^{OS}$, Γ for $\mathcal{U}^{B}$, α' for $\mathcal{U}^{T}$, and Ω for $\mathcal{U}^{Q}$), we can find the solution for each uncertainty set such that the corresponding objective value is equal to the in-sample VaR of level $\omega = 0.1$. Then we evaluate the out-of-sample returns and the out-of-sample VaR values of level $\omega = 0.1$ for the six solutions using another dataset consisting of 10^6 samples of 17 portfolios' returns drawn from the same distribution that is used to generate the in-sample dataset. As a result, for each repetition, each uncertainty set would have a solution with its out-of-sample return and out-of-sample VaR of level $\omega = 0.1$. Note that with the above procedure, the in-sample returns for different uncertainty sets are the same.

We have tested for 4 different in-sample sample sizes $N = 100,\ 200,\ 500,\ 1000$ for each of the case with $\overline{\rho} = 0.05,\ 0.5,\ 0.95$. To achieve the in-sample VaR of level $\omega = 0.1$, we used binary search to find the parameter (with precision up to 14 decimal places) for each uncertainty set. This requires repeatedly solving instances of Model (15). Computation time for cases with different sample sizes ranges from 1 to 6 hours. Note that we have discarded the instances for which we were not able to find a parameter for some uncertainty set(s) such that the in-sample VaR of level $\omega = 0.1$ can be achieved. We believe this does not affect the relative performance of different uncertainty sets.

Fig. 5 shows the mean out-of-sample return and the mean out-of-sample VaR of level $\omega = 0.1$ (over 400 repetitions) for different uncertainty sets when $\overline{\rho} = 0.5$ and $N = 100$. We say that one uncertainty set dominates another uncertainty set if the first uncertainty set has higher return and lower VaR value. The interval uncertainty set has lower return and higher VaR values than all other uncertainty sets, so the interval uncertainty set is dominated by all other uncertainty sets. The performances of the tail uncertainty set and the ellipsoidal uncertainty set are very close to each other. No uncertainty set dominates the order statistic uncertainty set because the order statistic uncertainty set either has higher return or has lower VaR value than other uncertainty sets. To check whether these findings are statistically significant, we can plot the confidence intervals for different uncertainty sets. To more clearly show the confidence intervals, we plot the confidence intervals of returns and VaR values separately in Fig. 6. We use the 95% level of confidence for all the confidence interval analyses in this paper. Note that the portfolio returns and VaR values reported in this paper are percentage values, i.e., we need to divide them by 100 and then add 1 to get the actual portfolio return and VaR. The "CH" in

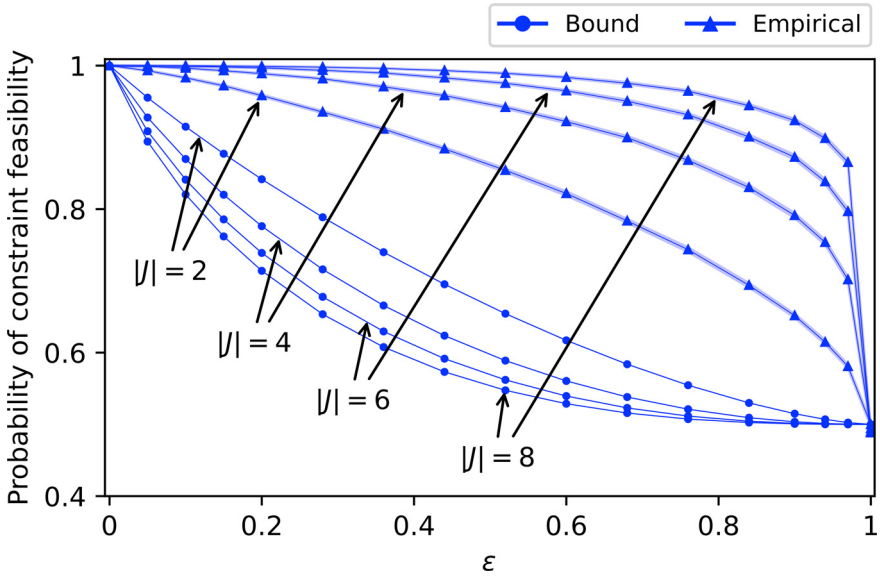

Fig. 4. Probability bound v.s. the posteriori empirical probability of constraint feasibility.

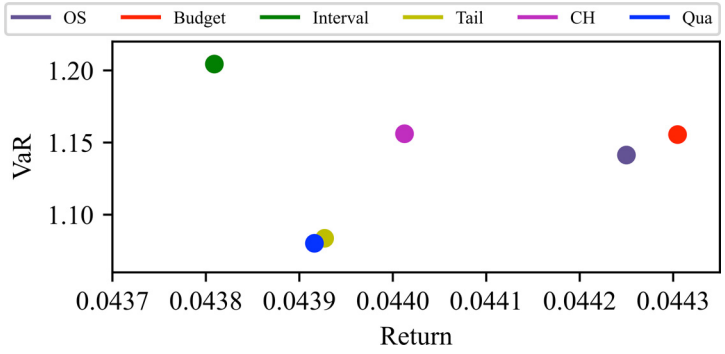

Fig. 5. The out-of-sample return and out-of-sample VaR for the case of $\overline{\rho} = 0.5$ and $N = 100$.

the figure represents the convex hull uncertainty set. The "Qua" in the figure represents the ellipsoidal uncertainty set.

5.3.1. The effect of sample size.

In Fig. 6, the out-of-sample returns for each of the six uncertainty sets do not seem to be affected by the sample size, no matter when the correlation is low, medium, or high. For the out-of-sample VaR, the performances of the order statistic uncertainty set and the budget uncertainty set do not seem to be affected by the sample size for different correlation regimes. The out-of-sample VaR values of the interval uncertainty set, the tail uncertainty set, the convex hull uncertainty set and the ellipsoidal uncertainty set improve when the sample size increases from 100 to 1000 for both the low-correlation and the medium-correlation cases, but not for the high-correlation cases. This can be explained by the effect of sample size, i.e., for low-correlation and medium-correlation regimes, the estimation of the parameters for the four uncertainty sets improves with the sample size. Moreover, we can see that the effect of sample size on the out-of-sample VaR is more significant when the correlation of random variables is lower.

Next, we compare the performance of the order statistic uncertainty set with the other five uncertainty sets. Note that in Fig. 6, the confidence intervals of the order statistic uncertainty set and other uncertainty sets may overlap. Because we cannot determine whether the out-of-sample returns or VaR values are statistically significantly different from each other if the confidence intervals overlap, we test whether the difference between the out-of-sample returns or VaR values is significantly greater (or less) than 0. We will always use this approach for significance tests whenever we have overlapping confidence intervals in this paper.

We say one uncertainty set is better than another uncertainty set if either of the following condition is satisfied: (1) the VaR of the first uncertainty set is lower than that of the second uncertainty set, and the return of the first uncertainty set is higher than or not significantly different from that of the second uncertainty set; (2) the return of the first uncertainty set is higher than that of the second uncertainty set, and the VaR of the first uncertainty set is lower than or not significantly different from that of the second uncertainty set. If neither of these two conditions is satisfied, then we know that one of the two uncertainty sets has significantly higher VaR value and higher return than the other uncertainty set, and in this case we say that the two uncertainty sets offer different trade-offs between the VaR and return, and the two uncertainty sets cannot be ranked.

In Fig. 7, we summarize the comparison between the order statistic uncertainty set and each of the other five uncertainty sets based on the results in Fig. 6. The "Ret" stands for the out-of-sample return. The sign "$>$" (or "$<$", or "$\approx$") denotes that the VaR value or the return of the corresponding uncertainty set is greater than (or smaller than, or not significantly different from) the order statistic uncertainty set. The cells with green color (or diagonal lines) mean that the performance of the order statistic uncertainty set is better (or worse) than the corresponding uncertainty set. The cells without green color or diagonal lines mean that the

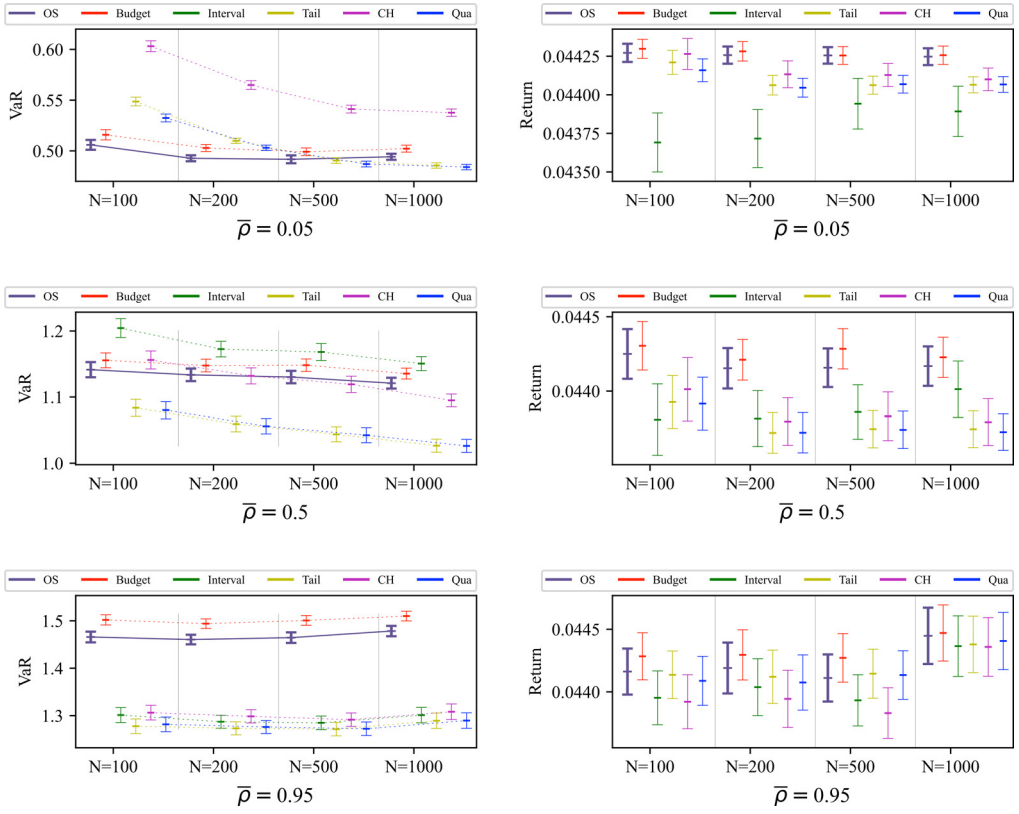

Fig. 6. The out-of-sample performance for different N sample sizes and different correlation regimes. The target return r is the median of the 17 portfolios returns; $\omega = 0.1$. For the case of $\bar{\rho} = 0.05$, the mean VaR values for the interval uncertainty set are 1.048, 1.019, 1.011, 0.995.

	$\overline{\rho} = 0.05$								$\overline{\rho} = 0.5$								$\overline{\rho} = 0.95$							
	N=100		N=200		N=500		N=1000		N=100		N=200		N=500		N=1000		N=100		N=200		N=500		N=1000	
	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret
Budget	$>$	$\approx$	$>$	$\approx$	$>$	$\approx$	$>$	$\approx$	$>$	$\approx$	$>$	$\approx$	$>$	$>$	$>$	$\approx$	$>$	$>$	$>$	$>$	$>$	$>$	$>$	$\approx$
Interval	$>$	$<$	$>$	$<$	$>$	$<$	$>$	$<$	$>$	$<$	$>$	$<$	$>$	$<$	$>$	$\approx$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$\approx$
Tail	$>$	$<$	$>$	$<$	$\approx$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$\approx$	$<$	$\approx$	$<$	$\approx$	$<$	$\approx$
CH	$>$	$\approx$	$>$	$<$	$>$	$<$	$>$	$<$	$>$	$<$	$\approx$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$\approx$
Qua	$>$	$<$	$>$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$<$	$\approx$	$<$	$\approx$	$<$	$\approx$	$<$	$\approx$

The sign $<$ ($>$ or $\approx$) means the corresponding VaR or Return is significantly less (significantly greater or not significantly different) than that of the order statistic uncertainty set.

: the order statistic uncertainty set is better. : the order statistic uncertainty set is worse.

Fig. 7. Compare the out-of-sample return and out-of-sample VaR of the order statistic uncertainty set with each of the other five uncertainty sets.

order statistic uncertainty set and the corresponding uncertainty set achieve different trade-offs between the VaR value and the return, and the performances of the two uncertainty sets can not be ranked.

For example, when $\bar{\rho} = 0.05$ and $N = 100$ in Fig. 6, the budget uncertainty set has larger VaR value than the order statistic uncertainty set and the returns of the two uncertainty sets are not significantly different from each other, and so we can conclude that the order statistic uncertainty set has better performance. Correspondingly, when $\bar{\rho}$=0.05 and $N = 100$ in Fig. 7, we add the sign "$>$" in the cell for the row "Budget" and the column "VaR", and add the sign "$\approx$" in the cell for the row "Budget" and the column "Ret". Because the order statistic uncertainty set is better than the budget uncertainty set in this case, we fill the two cells with color green. All other cells in Fig. 7 can be interpreted in the same way.

5.3.2. Comparision between the order statistic uncertainty set and other uncertainty sets

Comparison with the budget uncertainty set. For all the four cases when $\overline{\rho} = 0.05$, Fig. 7 shows that the out-of-sample returns of the order statistic uncertainty set and the budget uncertainty set are not significantly different from each other, and the out-of-sample VaR of the order statistic uncertainty set is significantly lower than the budget uncertainty set. This shows that the order statistic uncertainty set has better performance than the budget uncertainty set. Similarly, when $\overline{\rho} = 0.5$ (except for the case $N = 500$), the order statistic uncertainty set also has better performance than the budget uncertainty set. For the case $\overline{\rho} = 0.95$, the order statistic uncertainty set has better performance when $N = 1000$. For the other three cases with $N = 100$, 200 and 500, the order statistic uncertainty set has significantly lower VaR values and lower return

	N=100	N=200	N=500	N=1000	N=2000	N=5000	N=10000	N=15000
OS	0.109	0.133	0.214	0.284	0.484	0.957	1.866	2.782
Budget	0.113	0.118	0.220	0.259	0.423	0.912	2.162	2.482
Interval	0.057	0.085	0.160	0.223	0.404	0.849	1.982	2.483
Tail	0.091	0.243	0.374	0.633	1.262	3.065	5.625	8.862
CH	0.082	0.174	0.344	0.566	1.059	2.518	5.064	7.460
Qua	0.080	0.134	0.159	0.247	0.466	0.891	1.951	2.396

Fig. 8. The computation time (in seconds) of different uncertainty sets. The target return r is the median of the 17 portfolios returns; $\omega = 0.1$.

values than the budget uncertainty set. This means that for these three cases, the solutions of the two uncertainty sets offer different trade-offs between the return and VaR value, and the performances of the two uncertainty sets cannot be ranked.

Comparison with the interval/convex hull uncertainty set. For all cases when $\bar{\rho} = 0.05$, the order statistic uncertainty set has better performance than the interval and the convex hull uncertainty set. The interval and the convex hull uncertainty set has better performance than the order statistic uncertainty set only when $\overline{\rho} = 0.95$ and $N = 1000$. For all other cases, either the order statistic uncertainty set has better performance than the interval/convex hull uncertainty set, or the performance of the order statistic and the interval/convex hull uncertainty set cannot be ranked.

Comparison with the tail/ellipsoidal uncertainty set. For the low-correlation regime $\overline{\rho} = 0.05$ when the sample size is small (e.g., $N = 100$ or 200), the order statistic uncertainty set has better performance than the tail/ellipsoidal uncertainty set. For the high-correlation regime $\overline{\rho} = 0.95$, the tail/ellipsoidal uncertainty set has better performance than the order statistic uncertainty set. For the medium-correlation regime $\overline{\rho} = 0.5$, the order statistic uncertainty set has significantly higher out-of-sample returns and significantly higher VaR values than the tail/ellipsoidal uncertainty set. Therefore, for the medium-correlation case, we cannot draw a definitive conclusion regarding the relative performance of the uncertainty sets because they provide different balances of portfolio return and VaR.

In summary, when the correlation of the portfolio returns is low and the sample size is small, the order statistic uncertainty set tends to have a better out-of-sample performance than the other five uncertainty sets. For the medium-correlation case, the order statistic uncertainty set can outperform some uncertainty sets (e.g., the interval uncertainty set), but for some other uncertainty sets, the order statistic uncertainty set does not have a superior performance. For example when the correlation is 0.5, the solution of the tail/ellipsoidal uncertainty set offers different trade-offs compared with the solution of the order statistic uncertainty set. For the high-correlation regime, the performance of the order statistic uncertainty set deteriorates even more because it does not outperform the other five uncertainty sets for most cases.

From above, we can conclude that the order statistic uncertainty set tends to have better performance when the sample size is small and the correlations among the portfolio returns are low. When the correlation increases from $\bar{\rho} = 0.05$ to 0.5 and 0.95, the performance of the order statistic uncertainty set gets worse. This can be explained by the fact that the correlation between random variables is not incorporated in the order statistic uncertainty set, and thus the performance of the order statistic uncertainty set declines as the correlation grows.

5.3.3. Computation time

In this section, we evaluate the computational performances of different uncertainty sets. We first generate a sample dataset that consists of N samples of 17 portfolios' returns drawn from the multivariate normal distribution with mean $(\mu_1, \ldots, \mu_{17})$ and covariance matrix whose (i, j)th entry is σ_i^2 if $i = j$ and $\overline{\rho} \cdot \sigma_i \cdot \sigma_j$ if $i \neq j$. Because the results of the computation time for the case $\overline{\rho} = 0.5$ or the case $\overline{\rho} = 0.95$ are very similar to the case with $\overline{\rho} = 0.05$, we only present results for the case $\overline{\rho} = 0.05$. We set the target return r in constraint (15c) to be the median of $\mu_1, \ldots, \mu_{17}$. By trial-and-error, we can find the parameter for each uncertainty set such that the corresponding objective value is equal to the in-sample VaR of level $\omega = 0.1$. Then we use the calibrated parameter for each uncertainty set to solve problem (15) with 100 out-of-sample datasets. Each out-of-sample dataset consists of N samples of 17 portfolios' returns drawn from the same distribution as that used for the in-sample dataset. The mean computation time for the 100 out-of-sample datasets measures the time complexity of each uncertainty set. We have tested for 8 different sample sizes. The mean computation time for different uncertainty sets are shown in Fig. 8.

For each uncertainty set, the computation time increases with the sample size N. The computation time for the order statistic uncertainty set, the budget uncertainty set, the interval uncertainty set and the ellipsoidal uncertainty set are very close to each other. The tail uncertainty set and the convex hull uncertainty set consume the most computation time (except for the case with $N = 100$). Moreover, the gap between the tail/convex hull uncertainty set and other uncertainty sets grows as the sample size N increases. This can be explained by the model complexity of different uncertainty sets. The number of variables in the tail uncertainty set or the convex hull uncertainty set is proportional to the sample size N. The number of variables and constraints in the RO model with the order statistic uncertainty set, the budget uncertainty set, the interval uncertainty set or the ellipsoidal uncertainty set is proportional to the number of portfolios $|J|$ or $|J|^2$ but is not related to N. So when the sample size N increases from 100 to 15,000, the computation time for the tail uncertainty set and the convex hull uncertainty set increases faster than the other four uncertainty sets.

Among the order statistic uncertainty set, the budget uncertainty set, the interval uncertainty set and the ellipsoidal uncertainty set, the computation time of the order statistic uncertainty set is slightly more than the other three uncertainty sets for most cases. This can be explained by the fact that the number of variables and constraints in the RO model with the order statistic uncertainty set is more than those in the RO model with the other three uncertainty sets. That being said, the differences are small and not practically significant.

5.3.4. Robustness test.

We evaluate the performances of the order statistic uncertainty set and the other five uncertainty sets using a different set of parameters. Specifically, we set the target return r in constraint (15c) to be the 75th percentile of $\mu_1, \ldots, \mu_{17}$, and set ω to be 0.05 while keeping all other parameters to be the same. We follow the same procedure described at the beginning of Section 5.3 to find

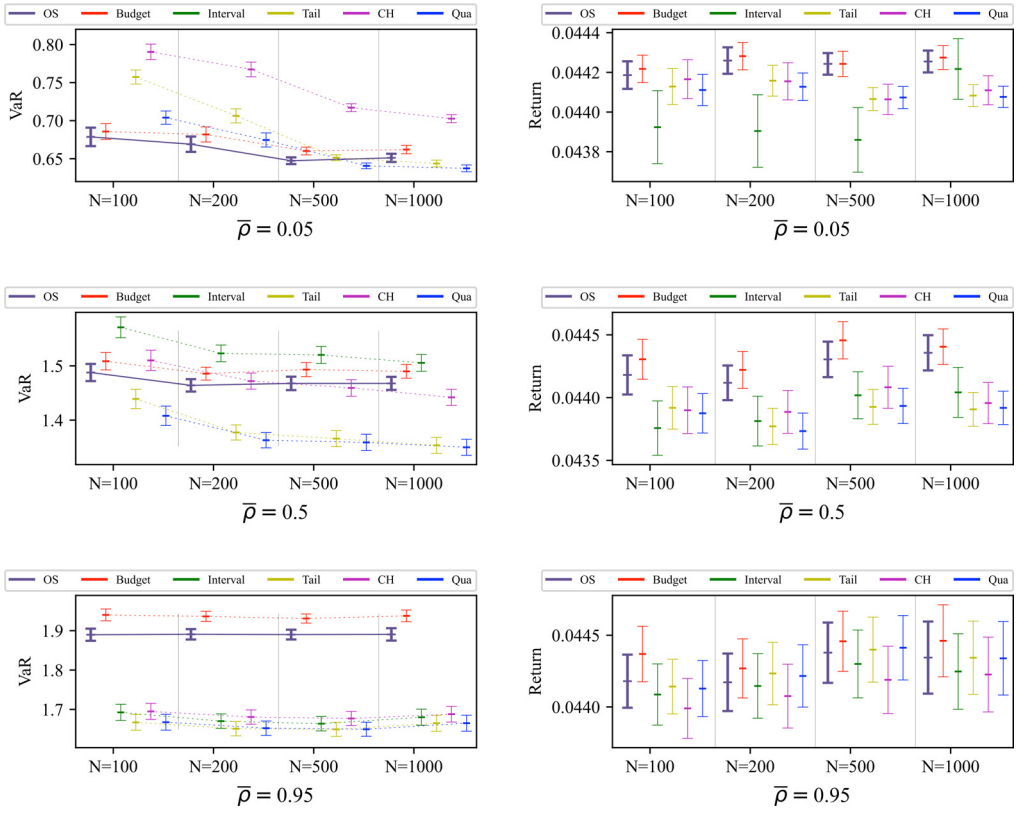

Fig. 9. The out-of-sample performance for different N sample sizes and different correlation regimes. The target return r is the 75th percentile of the 17 portfolios returns; $\omega = 0.05$. For the case of $\bar{\rho} = 0.05$, the mean VaR values for the interval uncertainty set are 1.357, 1.337, 1.299, 1.287.

	$\overline{\rho} = 0.05$								$\overline{\rho} = 0.5$								$\overline{\rho} = 0.95$							
	N=100		N=200		N=500		N=1000		N=100		N=200		N=500		N=1000		N=100		N=200		N=500		N=1000	
	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret	VaR	Ret
Budget	>	≈	>	≈	>	≈	>	≈	>	>	>	>	>	>	>	≈	>	>	>	>	>	>	>	>
Interval	>	<	>	<	>	<	>	≈	>	<	>	<	>	<	>	<	<	≈	<	≈	<	≈	<	≈
Tail	>	≈	>	<	>	<	<	<	<	<	<	<	<	<	<	<	<	≈	<	≈	<	≈	<	≈
CH	>	≈	>	<	>	<	>	<	>	<	≈	<	<	<	<	<	<	<	<	≈	<	<	<	≈
Qua	>	<	>	<	<	<	<	<	<	<	<	<	<	<	<	<	<	≈	<	≈	<	≈	<	≈

The sign < (> or ≈) means the corresponding VaR or Return is significantly less (significantly greater or not significantly different) than that of the order statistic uncertainty set.

: the order statistic uncertainty set is better. : the order statistic uncertainty set is worse.

Fig. 10. Compare the out-of-sample return and out-of-sample VaR of the order statistic uncertainty set with each of the other five uncertainty sets.

and evaluate solutions for different uncertainty sets. The results for VaR and return are shown in Fig. 9. The comparison analysis between the order statistic uncertainty set and each of the other five uncertainty sets is shown in Fig. 10.

The results can be interpreted in the same way as in Figs. 6 and 7. The relative performance of the order statistic uncertainty set and the other five uncertainty sets is qualitatively the same as what we observe in Section 5.3.2. When the correlation is low ($\bar{\rho} = 0.05$), the order statistic uncertainty set has better performance than the budget uncertainty set, the interval uncertainty set and the convex hull uncertainty set. When the correlation is low ($\bar{\rho} = 0.05$) and the sample size is small, e.g., $N = 100$ or 200, the order statistic uncertainty set has better performance than the tail uncertainty set and the ellipsoidal uncertainty set. As the correlation increases to 0.95, the performance of the order statistic uncertainty set decreases.

We also follow the same procedure in Section 5.3.3 to evaluate the computational performances of different uncertainty sets, except that we now make the target return r in constraint (15c) be the 75th percentile of $\mu_1, \ldots, \mu_{17}$, and make ω be 0.05 (all other parameters are kept to be the same). The results are shown in Fig. 11. The relative computational performance of different uncertainty sets is very similar to the results in Fig. 8. For example, the computation time of the order statistic uncertainty set is slightly higher than the budget uncertainty set, the interval uncertainty set and the ellipsoidal uncertainty set. The computation time of the tail/convex hull uncertainty set increases the most as the sample size N increases.

	N=100	N=200	N=500	N=1000	N=2000	N=5000	N=10000	N=15000
OS	0.178	0.195	0.275	0.281	0.433	0.997	2.018	2.873
Budget	0.105	0.164	0.198	0.263	0.414	1.048	1.653	2.749
Interval	0.054	0.107	0.143	0.218	0.376	1.033	1.616	2.533
Tail	0.089	0.209	0.332	0.631	1.119	3.979	5.591	10.104
CH	0.080	0.190	0.347	0.545	0.996	2.530	5.042	8.143
Qua	0.086	0.116	0.172	0.271	0.395	0.997	1.930	2.488

Fig. 11. The computation time (in seconds) of different uncertainty sets. The target return r is the 75th percentile of the 17 portfolios returns; $\omega = 0.05$.

6. Conclusion

In this paper, we develop and analyze the order statistic uncertainty set for robust linear optimization models. We use the Probability Integral Transform to study data-free and distribution-free properties of random variables, which are then embedded in the design of the order statistic uncertainty set. To depict uncertainties, the order statistic uncertainty set utilizes quantiles of random variables, which contain rich information of distributions. We demonstrate the geometric flexibility of the order statistic uncertainty set and show that the RO models with the interval uncertainty set, the budget uncertainty set, and the demand uncertainty set are all special cases of the RO model with the order statistic uncertainty set. Numerical experiments on a portfolio construction problem show that the order statistic uncertainty set outperforms five other existing uncertainty sets when the sample size is small and the correlation of random variables is low.

Appendix A. Proof of Proposition 2.1

The proof is by contradiction. First, assume that the statement of the proposition is not true. Then in the optimal solution to $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon}', \boldsymbol{\varepsilon}))$, there must exist at least one k such that $U_{(k)} < Q_k^{(1-\varepsilon_k)}$. Let $k' = arg\,max_k\{k : U_{(k)} < Q_k^{(1-\varepsilon_k)}\}$ and denote the corresponding random variable to be $Z_{j(k')}$. Specifically, $j(k')$ is a mapping from order statistics' index k' to the index j of random variable Z_j. We must have that $F_{j(k')}(Z_{j(k')}) = U_{(k')} < Q_{k'}^{(1-\varepsilon_{k'})}$. According to the definition of k', we have $U_{(k)} = Q_k^{(1-\varepsilon_k)}, \forall k > k'$.

Let $F_{j(k')}(Z_{j(k')} + \delta) = Q_{k'}^{(1-\varepsilon_{k'})}$. Such δ exists because $0 \le Q_{k'}^{(1-\varepsilon_{k'})} \le 1$; and $\delta > 0$ because $F_{j(k')}$ is non-decreasing. We construct a new solution to $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon}', \boldsymbol{\varepsilon}))$ by only modifying $Z_{j(k')}$ to be $Z_{j(k')} + \delta$. The new solution to $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon}', \boldsymbol{\varepsilon}))$ becomes $Z_1, \ldots, Z_{j(k')-1}, Z_{j(k')} + \delta, Z_{j(k')+1}, \ldots, Z_{|J|}$.

In the old solution, the order statistics of $F_1(Z_1), \ldots, F_{j(k')-1}(Z_{j(k')-1}), F_{j(k')}(Z_{j(k')}), F_{j(k')+1}(Z_{j(k')+1}), \ldots, F_{|J|}(Z_{|J|})$ are $U_{(1)}, \ldots, U_{(k'-1)}, U_{(k')}, U_{(k'+1)}, \ldots, U_{(|J|)}$. Next we show that in the new solution, the order statistics of $F_1(Z_1), \ldots, F_{j(k')-1}(Z_{j(k')-1}), F_{j(k')}(Z_{j(k')} + \delta), F_{j(k')+1}(Z_{j(k')+1}), \ldots, F_{|J|}(Z_{|J|})$ are $U_{(1)}, \ldots, U_{(k'-1)}, F_{j(k')}(Z_{j(k')} + \delta), U_{(k'+1)}, \ldots, U_{(|J|)}$. This will hold if we can prove that $U_{(k'-1)} \le F_{j(k')}(Z_{j(k')} + \delta) \le U_{(k'+1)}$, because the only difference between the old solution and the new solution is that $Z_{j(k')}$ becomes $Z_{j(k')} + \delta$. We have $U_{(k'-1)} \le F_{j(k')}(Z_{j(k')} + \delta)$ because $U_{(k'-1)} \le U_{(k')} < Q_{k'}^{(1-\varepsilon_{k'})} = F_{j(k')}(Z_{j(k')} + \delta)$. We have $F_{j(k')}(Z_{j(k')} + \delta) \le U_{(k'+1)}$ because $F_{j(k')}(Z_{j(k')} + \delta) = Q_{k'}^{(1-\varepsilon_{k'})} \le Q_{k'+1}^{(1-\varepsilon_{k'+1})} = U_{(k'+1)}$. Therefore, we have proven that $U_{(k'-1)} \le F_{j(k')}(Z_{j(k')} + \delta) \le U_{(k'+1)}$ holds, and this means that the order statistics of $F_1(Z_1), \ldots, F_{j(k')-1}(Z_{j(k')-1}), F_{j(k')}(Z_{j(k')} + \delta), F_{j(k')+1}(Z_{j(k')+1}), \ldots, F_{|J|}(Z_{|J|})$ are $U_{(1)}, \ldots, U_{(k'-1)}, F_{j(k')}(Z_{j(k')} + \delta), U_{(k'+1)}, \ldots, U_{(|J|)}$. Based on this conclusion, we can tell that the new solution $Z_1, \ldots, Z_{j(k')-1}, Z_{j(k')} + \delta, Z_{j(k')+1}, \ldots, Z_{|J|}$ is feasible to $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon}', \boldsymbol{\varepsilon}))$.

The new solution increases the objective value by $\hat{a}_{j(k')}|x_{j(k')}| \cdot \delta \ge 0$, and this means that the old solution to $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon}', \boldsymbol{\varepsilon}))$ was not a maximizer of $\beta(\mathbf{x}, \mathcal{U}'(\boldsymbol{\varepsilon}', \boldsymbol{\varepsilon}))$, and we can improve the objective value by increasing $Z_{j(k')}$ to be $Z_{j(k')} + \delta$. Using this same logic, we continue to modify our solution if there exist a k such that $U_{(k)} < Q_k^{(1-\varepsilon_k)}$, and eventually we must have $U_{(k)} = Q_k^{(1-\varepsilon_k)}, \forall k$.

Appendix B. Proof of Proposition 2.2

Without loss of generality, assume $k_1 < k_2$. As we stated in Section 2.1, if $k_1 < k_2$, then we have $Q_{k_1}^{(1-\varepsilon_{k_1})} \le Q_{k_2}^{(1-\varepsilon_{k_2})}$. Because $Q_{k_1}^{(1-\varepsilon_{k_1})} \ne Q_{k_2}^{(1-\varepsilon_{k_2})}$, we have $Q_{k_1}^{(1-\varepsilon_{k_1})} < Q_{k_2}^{(1-\varepsilon_{k_2})}$.

We further assume $Q_{k_1}^{(1-\varepsilon_{k_1})} < Q_{k_1+1}^{(1-\varepsilon_{k_1+1})}$, and the reason for it is in the following. Because $Q_{k_1}^{(1-\varepsilon_{k_1})} < Q_{k_2}^{(1-\varepsilon_{k_2})}$, there must exist an integer m $(k_1 \le m < k_2)$ such that $Q_m^{(1-\varepsilon_m)} < Q_{k_2}^{(1-\varepsilon_{k_2})}$ and $Q_m^{(1-\varepsilon_m)} < Q_{m+1}^{(1-\varepsilon_{m+1})}$. If $m = k_1$, then $Q_{k_1}^{(1-\varepsilon_{k_1})} < Q_{k_1+1}^{(1-\varepsilon_{k_1+1})}$ holds; if $m > k_1$, then we make m to be the new k_1, and then it holds that $Q_{k_1}^{(1-\varepsilon_{k_1})} < Q_{k_1+1}^{(1-\varepsilon_{k_1+1})}$. Note that because $m < k_2$, we have $Q_{m+1}^{(1-\varepsilon_{m+1})} \le Q_{k_2}^{(1-\varepsilon_{k_2})}$, i.e., for the new k_1, we also have $Q_{k_1+1}^{(1-\varepsilon_{k_1+1})} \le Q_{k_2}^{(1-\varepsilon_{k_2})}$. Therefore, in the following, we will assume $Q_{k_1}^{(1-\varepsilon_{k_1})} < Q_{k_1+1}^{(1-\varepsilon_{k_1+1})} \le Q_{k_2}^{(1-\varepsilon_{k_2})}$.

Next we choose a $\mathbf{Z} = (Z_1, \ldots, Z_j) \in \mathcal{U}^{OS}(\boldsymbol{\varepsilon})$, such that there exist j_1, j_2 $(j_1 \ne j_2)$ that satisfy $F_{j_1}(Z_{j_1}) = U_{j_1} = Q_{k_1}^{(1-\varepsilon_{k_1})}$ and $F_{j_2}(Z_{j_2}) = U_{j_2} = Q_{k_2}^{(1-\varepsilon_{k_2})}$. Since $0 \le Q_{k_1}^{(1-\varepsilon_{k_1})} \le 1$ and $0 \le Q_{k_2}^{(1-\varepsilon_{k_2})} \le 1$, we can find W_1 and W_2, such that $F_{j_1}(W_1) = Q_{k_2}^{(1-\varepsilon_{k_2})}$ and $F_{j_2}(W_2) = Q_{k_1}^{(1-\varepsilon_{k_1})}$. Then we construct $\mathbf{Z}'$ by replacing Z_{j_1} and Z_{j_2} in $\mathbf{Z}$ with W_1 and W_2, respectively. So we have $Z'_{j_1} = W_1, Z'_{j_2} = W_2$ and $Z'_j = Z_j$, for $j \ne j_1, j_2$. Since $F_{j_1}(Z'_{j_1}) = F_{j_2}(Z_{j_2})$, $F_{j_2}(Z'_{j_2}) = F_{j_1}(Z_{j_1})$ and $F_j(Z'_j) = F_j(Z_j)$, for $j \ne j_1, j_2$, we have $\{F_1(Z'_1), \ldots, F_{|J|}(Z'_{|J|})\} = \{F_1(Z_1), \ldots, F_{|J|}(Z_{|J|})\}$. This shows that $\mathbf{Z}' \in \mathcal{U}^{OS}(\boldsymbol{\varepsilon})$.

Because $Q_{k_1}^{(1-\varepsilon_{k_1})} < Q_{k_2}^{(1-\varepsilon_{k_2})}$, we have $Z_{j_1} < W_1$ and $Z_{j_2} > W_2$. Then for any $\lambda \in (0, 1)$, we must have $F_{j_1}((1-\lambda)Z_{j_1} + \lambda W_1)) > Q_{k_1}^{(1-\varepsilon_{k_1})}$ and $F_{j_2}((1-\lambda)Z_{j_2} + \lambda W_2)) > Q_{k_1}^{(1-\varepsilon_{k_1})}$.

For any $\lambda \in (0, 1)$, denote $\mathbf{Z}^\lambda = (1-\lambda)\mathbf{Z} + \lambda\mathbf{Z}'$. We next show $\mathbf{Z}^\lambda \notin \mathcal{U}^{OS}(\boldsymbol{\varepsilon})$. Since $Z_j^\lambda = (1-\lambda)Z_j + \lambda Z'_j = Z_j$, for $j \ne j_1, j_2$, we have $\{F_j(Z_j^\lambda) : \forall j \ne j_1, j_2\} = \{F_j(Z_j) : \forall j \ne j_1, j_2\} = \{Q_k^{(1-\varepsilon_k)} : \forall k \ne k_1, k_2\}$. Therefore, the k_1-th order statistic of

$\{F_j(Z_j^\lambda) : \forall j = 1, \ldots, |J|\} = \min\{Q_{k_1+1}^{[1-\varepsilon_{(k_1+1)}]}, F_{j_1}((1-\lambda)Z_{j_1} + \lambda W_1)), F_{j_2}((1-\lambda)Z_{j_2} + \lambda W_2)\}$ is strictly greater than $Q_{k_1}^{(1-\varepsilon_{k_1})}$. This proves that $\mathbf{Z}^\lambda \notin \mathcal{U}^{OS}(\boldsymbol{\varepsilon})$, and our proof is completed. □

Appendix C. Proof of Proposition 2.3

First we prove that the optimal solution to problem (6) is feasible to $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$. Note that the problem (6) is essentially a linear relaxation of the maximum weight assignment problem. It is well-known that there is always an optimal solution with all the η variables taking integer values. In the optimal solution to problem (6), for every $j \in J$, there exists a unique $k \in J$ such that $\eta_{jk} = 1$. If $\eta_{jk} = 1$, we let $Z_j = \sum_{k' \in J} q_{jk'} \eta_{jk'} = q_{jk}$, and we have $F_j(Z_j) = F_j(q_{jk}) = Q_k^{(1-\varepsilon_k)}$. Then we have $\{F_j(Z_j), \forall j \in J\} = \{Q_k^{(1-\varepsilon_k)}, \forall k \in J\}$. Therefore, for any $1 \le m \le |J|$, we must have the mth smallest element in the set $\{F_j(Z_j), \forall j \in J\}$ should be no greater than the mth smallest element in the set $\{Q_k^{(1-\varepsilon_k)}, \forall k \in J\}$, which is essentially $U_{(m)} \le Q_m^{(1-\varepsilon_m)}, \forall m \in J$. This proves that the optimal solution to problem (6) is indeed feasible to the problem $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$, and so the optimal objective value for problem (6) is less than or equal to the optimal objective value for the problem $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$.

Next we prove that the optimal solution to $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$ is feasible to problem (6). Assume in the optimal solution to $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$, the order statistics of $F_j(Z_j)$'s are $F_{j_1}(Z_{j_1}), F_{j_2}(Z_{j_2}), \ldots, F_{j_{|J|}}(Z_{j_{|J|}})$, where the sequence $j_1, j_2, \ldots, j_{|J|}$ are a permutation of the set $\{1, 2, \ldots, |J|\}$. We then have $F_{j_k}(Z_{j_k}) = Q_k^{(1-\varepsilon_k)}, \forall k \in J$, i.e., $Z_{j_k} = q_{j_k,k}$. Such a solution is feasible to problem (6) because we can construct the corresponding solution to problem (6) as follows:

$$\eta_{j_k,m} = \begin{cases} 1, & k = m, \\ 0, & k \ne m, \end{cases} \qquad \forall k, m \in J. \tag{C.1}$$

This proves that the optimal solution to $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$ is feasible to problem (6), and so the optimal objective value for $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$ is less than or equal to the optimal objective value for the problem (6).

In the above, we have proven that the optimal objective values for problem (6) and problem $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$ are no greater than each other, and so the optimal objective values of problem (6) and problem $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{\varepsilon}))$ must be equal to each other.

Appendix D. Proof of Theorem 2.4

Similar to Bertsimas & Sim (2004), we apply the strong duality to reformulate Model (7). For fixed $\mathbf{x}$, we first take dual of the maximizing problem in constraints (7b), and we get:

$$\min \sum_{j \in J_i} (\theta_{ij} + \phi_{ij}) \tag{D.1a}$$

$$\text{s.t. } \theta_{ij} + \phi_{ik} \ge \hat{a}_{ij} |x_j| q_{ijk}, \forall j, k \in J_i, \forall i \tag{D.1b}$$

Because the maximizing problem in constraints (7b) is feasible and bounded, we must have that the formulation (D.1) is also feasible and bounded due to strong duality. And their optimal objective values are equal. Substituting formulation (D.1) into Model (7), we can get the linear programming formulation (8). Hence proven. □

Appendix E. Proof: the equivalence of the RO models with the budget uncertainty set and the order statistic uncertainty set

We prove that the RO model with the budget uncertainty set with budget Γ (the Problem (4) in Bertsimas & Sim, 2004) is equivalent to the RO model with the order statistic uncertainty set with properly chosen q_{jk} values.

The Problem (4) in Bertsimas & Sim (2004) is essentially the following problem. We assume that there is only one constraint that has the budget uncertainty set, and so we have removed the i index.

$$\max_{\mathbf{x}} \ \mathbf{c}'\mathbf{x} \tag{E.1a}$$

$$\text{s.t. } \sum_j a_j x_j + \beta\left(\mathbf{x}, \mathcal{U}^B(\Gamma)\right) \le b, \tag{E.1b}$$

$$\underline{\mathbf{x}} \le \mathbf{x} \le \bar{\mathbf{x}}, \tag{E.1c}$$

where $\beta(\mathbf{x}, \mathcal{U}^B(\Gamma))$ is the following problem

$$\max_{\{S \cup \{t\} | S \subset J, |S| = \lfloor \Gamma \rfloor, t \in J \setminus S\}} \left\{ \sum_{j \in S} \hat{a}_j \cdot |x_j| + (\Gamma - \lfloor \Gamma \rfloor) \cdot \hat{a}_t \cdot |x_t| \right\}. \tag{E.2}$$

Note that we have used $\underline{\mathbf{x}}$ and $\bar{\mathbf{x}}$ instead of $\mathbf{l}$ and $\mathbf{u}$ for lower and upper bounds. We need to prove that the problem (E.1) is equivalent to the following RO model with the order statistic uncertainty set.

$$\max_{\mathbf{x}} \ \mathbf{c}'\mathbf{x} \tag{E.3a}$$

$$\text{s.t. } \sum_j a_j x_j + \beta\left(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{q})\right) \le b, \tag{E.3b}$$

$$\underline{\mathbf{x}} \le \mathbf{x} \le \bar{\mathbf{x}}, \tag{E.3c}$$

where $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{q}))$ is the following problem

$$\max_{\eta} \ \sum_{j \in J} \hat{a}_j |x_j| \cdot \left(\sum_{k \in J} q_{jk} \eta_{jk} \right) \tag{E.4a}$$

$$\text{s.t. } \sum_k \eta_{jk} = 1, \forall j \in J \tag{E.4b}$$

$$\sum_j \eta_{jk} = 1, \forall k \in J \tag{E.4c}$$

$$0 \le \eta_{jk} \le 1, \forall j, k \in J. \tag{E.4d}$$

and $\boldsymbol{q}$ satisfies $q_{jk} = 0$, if $1 \le k \le |J| - \lfloor \Gamma \rfloor - 1, \forall j \in J$; $q_{jk} = \Gamma - \lfloor \Gamma \rfloor$, if $k = |J| - \lfloor \Gamma \rfloor, \forall j \in J$; $q_{jk} = 1$, if $|J| - \lfloor \Gamma \rfloor + 1 \le k \le |J|, \forall j \in J$.

We just need to prove that the optimal objective value for $\beta(\mathbf{x}, \mathcal{U}^B(\Gamma))$ is equal to the optimal objective value for $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{q}))$.

The problem (E.4) is the linear relaxation of the maximum weight assignment problem, which is known to have an integer optimal solution. For every $j \in J$, there exists a unique $k \in J$ such that $\eta_{jk} = 1$. If $\eta_{jk} = 1$, then $\hat{a}_j |x_j|$ is paired to q_{jk}. Therefore, if $1 \le k \le |J| - \lfloor \Gamma \rfloor - 1$, then $\hat{a}_j |x_j|$ is paired with 0; if $k = |J| - \lfloor \Gamma \rfloor$, then $\hat{a}_j |x_j|$ is paired with $\Gamma - \lfloor \Gamma \rfloor$; if $|J| - \lfloor \Gamma \rfloor + 1 \le k \le |J|$, then $\hat{a}_j |x_j|$ is paired with 1. So for all $\hat{a}_1 |x_1|, \hat{a}_2 |x_2|, \ldots, \hat{a}_{|J|} |x_{|J|}|$, we know that $\lfloor \Gamma \rfloor$ of them will be paired with 1, and one of them will be paired with $\Gamma - \lfloor \Gamma \rfloor$, and the rest will be paired with 0. Therefore, the problem (E.4) is essentially the same as the problem (E.2), and so we have that the optimal objective values for $\beta(\mathbf{x}, \mathcal{U}^B(\Gamma))$ and $\beta(\mathbf{x}, \mathcal{U}^{OS}(\boldsymbol{q}))$ are equal to each other. Hence proven. □

Appendix F. Proof of Proposition 3.1

The constraints in $\mathcal{U}^D$ are equivalent to the following:

$$-\gamma \cdot m^{1/\alpha} \le \sum_{j\in S} Z_j \le \gamma \cdot m^{1/\alpha}, \forall S \subseteq J, |S|=m, m=1,\ldots,|J|, \quad \text{(F.1a)}$$

$$\text{or } -\gamma \cdot m^{1/\alpha} \le \min_{S\subseteq J,|S|=m} \sum_{j\in S} Z_j, \max_{S\subseteq J,|S|=m} \sum_{j\in S} Z_j \le \gamma \cdot m^{1/\alpha}, \quad \forall m = 1,\ldots,|J|. \quad \text{(F.1b)}$$

Because the kth order statistic of Z_j's is denoted as $Z_{(k)}$, we have the following:

$$\min_{S\subseteq J,|S|=m} \sum_{j\in S} Z_j = \sum_{k=1}^{m} Z_{(k)}, \quad \max_{S\subseteq J,|S|=m} \sum_{j\in S} Z_j = \sum_{k=|J|+1-m}^{|J|} Z_{(k)} \quad \text{(F.2)}$$

Then the demand uncertainty set $\mathcal{U}^D$ can be rewritten in terms of the order statistics of Z_j's as follows:

$$\mathcal{U}^D_{os} = \left\{ \mathbf{Z} : -\gamma \cdot j^{1/\alpha} \le \sum_{k=1}^{j} Z_{(k)}, \sum_{k=|J|+1-j}^{|J|} Z_{(k)} \le \gamma \cdot j^{1/\alpha}, \forall j \in J \right\}. \quad \text{(F.3)}$$

Since the demand uncertainty set is equivalent to $\mathcal{U}^D_{os}$, we just need to prove that $\mathbf{Z}^*$ maximizes $\beta(\mathbf{x}, \mathcal{U}^D_{os})$. We first check $\mathbf{Z}^*$ is feasible to $\mathcal{U}^D_{os}$, and then prove it is optimal to the problem $\beta(\mathbf{x}, \mathcal{U}^D_{os})$.

To prove the feasibility of $\mathbf{Z}^*$, there are two steps: (1) we need to prove that $(|J|+1-k)^{1/\alpha} - (|J|-k)^{1/\alpha}$ is increasing in k. That is $[(|J|-k)^{1/\alpha} - (|J|-k-1)^{1/\alpha}] - [(|J|+1-k)^{1/\alpha} - (|J|-k)^{1/\alpha}] \ge 0, \forall 1 \le k \le |J|-1$, or equivalently $2\cdot(|J|-k)^{1/\alpha} - (|J|-k-1)^{1/\alpha} - (|J|+1-k)^{1/\alpha} \ge 0$, which is evidenced by the fact that the function $x^{1/\alpha}$ is concave for $\alpha \ge 1$. (2) We check that $\mathbf{Z}^*$ satisfies all the constraints in $\mathcal{U}^D_{os}$. Note that $Z^*_{(k)} \ge 0, \forall k \in J$, so $-\gamma \cdot j^{1/\alpha} \le \sum_{k=1}^{j} Z^*_{(k)}, \forall j \in J$, is satisfied; for other constraints, $\sum_{k=|J|+1-j}^{|J|} Z^*_{(k)} = \sum_{k=|J|+1-j}^{|J|} \gamma \cdot (|J|+1-k)^{1/\alpha} - \gamma \cdot (|J|-k)^{1/\alpha} = \gamma \cdot j^{1/\alpha}$. Therefore, $\mathbf{Z}^*$ is feasible to $\mathcal{U}^D_{os}$.

We next prove that $\mathbf{Z}^*$ is the optimal solution to the problem $\beta(\mathbf{x}, \mathcal{U}^D_{os})$. Denote the ordered sequence of $\hat{a}_1|x_1|, \hat{a}_2|x_2|, \ldots, \hat{a}_{|J|}|x_{|J|}|$ as $[\hat{a}|x|]_{(1)}, [\hat{a}|x|]_{(2)}, \ldots, [\hat{a}|x|]_{(|J|)}$, i.e., $[\hat{a}|x|]_{(k)}$ is the kth smallest among all $\hat{a}_j|x_j|$'s. Because of the rearrangement inequality (Cvetkovski, 2012, Theorem 6.1), the objective value for any feasible $\mathbf{Z}$ in $\mathcal{U}^D_{os}$ should be no greater than $[\hat{a}|x|]_{(1)} \cdot Z_{(1)} + [\hat{a}|x|]_{(2)} \cdot Z_{(2)} + \ldots + [\hat{a}|x|]_{(|J|)} \cdot Z_{(|J|)}$. Then the difference of the optimal objective value of our $\mathbf{Z}^*$ and that of any feasible $\mathbf{Z}$ in $\mathcal{U}^D_{os}$ should be no less than:

$$[\hat{a}|x|]_{(1)} \cdot (Z^*_{(1)} - Z_{(1)}) + [\hat{a}|x|]_{(2)} \cdot (Z^*_{(2)} - Z_{(2)}) + \ldots + [\hat{a}|x|]_{(|J|)} \cdot (Z^*_{(|J|)} - Z_{(|J|)}) \quad \text{(F.4)}$$

Since $\mathbf{Z}$ satisfies $\sum_{k=|J|+1-j}^{|J|} Z_{(k)} \le \gamma \cdot j^{1/\alpha} = \sum_{k=|J|+1-j}^{|J|} Z^*_{(k)}, \forall j \in J$, we then have $\sum_{k=j}^{|J|} (Z^*_{(k)} - Z_{(k)}) \ge 0, \forall j \in J$, i.e., $Z^*_{(j)} - Z_{(j)} \ge -\sum_{k=j+1}^{|J|} (Z^*_{(k)} - Z_{(k)}), \forall 1 \le j \le |J|-1$. We apply these constraints to (F.4) one at a time repeatedly, and we can get:

$$\begin{aligned}
(F.4) &\ge -[\hat{a}|x|]_{(1)} \cdot \sum_{k=2}^{|J|} (Z^*_{(k)} - Z_{(k)}) + [\hat{a}|x|]_{(2)} \cdot \left(Z^*_{(2)} - Z_{(2)}\right) + \ldots \\
&\quad + [\hat{a}|x|]_{(|J|)} \cdot \left(Z^*_{(|J|)} - Z_{(|J|)}\right) \\
&= \left([\hat{a}|x|]_{(2)} - [\hat{a}|x|]_{(1)}\right) \cdot \left(Z^*_{(2)} - Z_{(2)}\right) \\
&\quad + \left([\hat{a}|x|]_{(3)} - [\hat{a}|x|]_{(1)}\right) \cdot \left(Z^*_{(3)} - Z_{(3)}\right) + \ldots \\
&\quad + \left([\hat{a}|x|]_{(|J|)} - [\hat{a}|x|]_{(1)}\right) \cdot \left(Z^*_{(|J|)} - Z_{(|J|)}\right) \\
&\ge -\left([\hat{a}|x|]_{(2)} - [\hat{a}|x|]_{(1)}\right) \cdot \sum_{k=3}^{|J|} \left(Z^*_{(k)} - Z_{(k)}\right) \\
&\quad + \left([\hat{a}|x|]_{(3)} - [\hat{a}|x|]_{(1)}\right) \cdot \left(Z^*_{(3)} - Z_{(3)}\right) \\
&\quad + \ldots + \left([\hat{a}|x|]_{(|J|)} - [\hat{a}|x|]_{(1)}\right) \cdot \left(Z^*_{(|J|)} - Z_{(|J|)}\right) \\
&= \left([\hat{a}|x|]_{(3)} - [\hat{a}|x|]_{(2)}\right) \cdot \left(Z^*_{(3)} - Z_{(3)}\right) \\
&\quad + \left([\hat{a}|x|]_{(4)} - [\hat{a}|x|]_{(2)}\right) \cdot \left(Z^*_{(4)} - Z_{(4)}\right) + \ldots \\
&\quad + \left([\hat{a}|x|]_{(|J|)} - [\hat{a}|x|]_{(2)}\right) \cdot \left(Z^*_{(|J|)} - Z_{(|J|)}\right) \ldots \\
&\ge \left([\hat{a}|x|]_{(|J|)} - [\hat{a}|x|]_{(|J|-1)}\right) \cdot \left(Z^*_{(|J|)} - Z_{(|J|)}\right) \\
&\ge 0
\end{aligned}$$

The last inequality holds because $Z_{(|J|)} \le \gamma = Z^*_{(|J|)}$, and $[\hat{a}|x|]_{(|J|)} - [\hat{a}|x|]_{(|J|-1)} \ge 0$ holds by definition. This concludes the proof. □

Appendix G. Proof of Corollary 3.2

Similar to the problem (6), the problem (10) is also a relaxed maximum weight assignment problem. Therefore, there always exists an optimal solution where all the η variables take integer values. For any $j \in J$, there exists a unique $k \in J$ such that $\eta_{jk} = 1$. As a result, problem (10) is equivalent to $\max_{j_1,\ldots,j_{|J|}} \sum_{k\in J} \hat{a}_k|x_k| \cdot \gamma \cdot ({j_k}^{1/\alpha} - (j_k-1)^{1/\alpha})$, where ${j_k}^{1/\alpha} - (j_k-1)^{1/\alpha}$ is assigned to $\hat{a}_k|x_k|$, and $\{j_1, j_2, \ldots, j_{|J|}\}$ is a permutation of the set $J = \{1,2,\ldots,|J|\}$. Due to the rearrangement inequality, the optimal solution to problem (10) must be $\sum_{k\in J} [\hat{a}|x|]_{(k)} \cdot \gamma \cdot ((|J|+1-k)^{1/\alpha} - (|J|-k)^{1/\alpha})$, which is exactly the optimal solution to the problem $\beta(\mathbf{x}, \mathcal{U}^D)$ as we derived in the proof for Proposition 3.1. □

Appendix H. Proof of Theorem 4.2

Denote the optimal solution to Model (7) as $\mathbf{x}^*$, and denote the corresponding value of the subproblem in constraint (7b) as β^*. Then β^* is equal to the optimal objective value of problem (6) when we fix $\mathbf{x}$ in problem (6) to be $\mathbf{x}^*$. Suppose the continuous variable A_j is independently and symmetrically distributed in $[a_j - \hat{a}_j, a_j + \hat{a}_j], \forall j \in J$, and we need to prove $\text{Prob}(\sum_{j\in J} A_j \cdot x^*_j \le b) \ge \frac{1}{2} + \frac{1}{2} \cdot |J|! \det[\Delta]$. Denote $\rho_j = (A_j - a_j)/\hat{a}_j \in [-1,1]$ and let $Z_j = |\rho_j|$. We then have

$$\text{Prob}\left(\sum_j A_j x^*_j \le b\right) \quad \text{(H.1a)}$$

$$= \text{Prob}\left(\sum_j a_j x^*_j + \sum_j \hat{a}_j |x^*_j| \cdot \rho_j \le b\right) \quad \text{(H.1b)}$$

$$= \text{Prob}\left(\sum_j \hat{a}_j |x^*_j| \cdot \rho_j \le b - \sum_j a_j x^*_j\right) \quad \text{(H.1c)}$$

$$\ge \text{Prob}\left(\sum_j \hat{a}_j |x^*_j| \cdot \rho_j \le \beta^*\right) \quad \text{(H.1d)}$$

Because ρ_j is symmetrically distributed in $[-1,1], \forall j \in J$, we must have $\text{Prob}(\sum_j \hat{a}_j|x^*_j| \cdot \rho_j \le 0) = \frac{1}{2}$ and that

$$\text{Prob}\left(0 \le \sum_j \hat{a}_j |x^*_j| \cdot \rho_j \le \beta^*\right)$$

$$= \text{Prob}\left(-\beta^* \le \sum_j \hat{a}_j |x^*_j| \cdot \rho_j \le 0\right)$$

$$= \frac{1}{2} \cdot \text{Prob}\left(\left| \sum_j \hat{a}_j |x_j^*| \cdot \rho_j \right| \le \beta^* \right)$$

Further, we have

$$\text{Prob}\left(\sum_j \hat{a}_j |x_j^*| \cdot \rho_j \le \beta^* \right) \tag{H.2a}$$

$$= \text{Prob}\left(\sum_j \hat{a}_j |x_j^*| \cdot \rho_j \le 0 \right) + \text{Prob}\left(0 \le \sum_j \hat{a}_j |x_j^*| \cdot \rho_j \le \beta^* \right) \tag{H.2b}$$

$$= \frac{1}{2} + \frac{1}{2} \cdot \text{Prob}\left(\left| \sum_j \hat{a}_j |x_j^*| \cdot \rho_j \right| \le \beta^* \right) \tag{H.2c}$$

$$\ge \frac{1}{2} + \frac{1}{2} \cdot \text{Prob}\left(\sum_j \hat{a}_j |x_j^*| \cdot |\rho_j| \le \beta^* \right) \tag{H.2d}$$

$$= \frac{1}{2} + \frac{1}{2} \cdot \text{Prob}\left(\sum_j \hat{a}_j |x_j^*| \cdot Z_j \le \beta^* \right) \tag{H.2e}$$

Note that $Z_1, \ldots, Z_{|J|}$ are independent with each other, so we know $F_j(Z_j)$'s are independent with each other and each $F_j(Z_j)$ follows Unif(0,1) distribution. Suppose $j_1, j_2, \ldots, j_{|J|}$ is an arbitrary permutation of $1, 2, \ldots, |J|$, and because of symmetry, we have $\text{Prob}[F_{j_1}(Z_{j_1}) \le F_{j_2}(Z_{j_2}) \le \ldots \le F_{j_{|J|}}(Z_{j_{|J|}})] = \frac{1}{|J|!}$. Then we have

$$\text{Prob}\left(\sum_j \hat{a}_j |x_j^*| \cdot Z_j \le \beta^* \right) \tag{H.3a}$$

$$= \sum_{\{j_1, j_2, \ldots, j_{|J|}\}:\ \{j_1, j_2, \ldots, j_{|J|}\} = \{1,2,\ldots,|J|\}} \text{Prob}\left[F_{j_1}(Z_{j_1}) \le F_{j_2}(Z_{j_2}) \le \ldots \le F_{j_{|J|}}(Z_{j_{|J|}}) \right]. \tag{H.3b}$$

$$\text{Prob}\left(\sum_{k=1}^{|J|} \hat{a}_{j_k} |x_{j_k}^*| \cdot Z_{j_k} \le \beta^* \middle| F_{j_1}(Z_{j_1}) \le F_{j_2}(Z_{j_2}) \le \ldots \le F_{j_{|J|}}(Z_{j_{|J|}}) \right) \tag{H.3c}$$

$$= \sum_{\{j_1, j_2, \ldots, j_{|J|}\}:\ \{j_1, j_2, \ldots, j_{|J|}\} = \{1,2,\ldots,|J|\}} \tag{H.3d}$$

$$\frac{1}{|J|!} \cdot \text{Prob}\left(\sum_{k=1}^{|J|} \hat{a}_{j_k} |x_{j_k}^*| \cdot Z_{j_k} \le \beta^* \middle| F_{j_1}(Z_{j_1}) \le F_{j_2}(Z_{j_2}) \le \ldots \le F_{j_{|J|}}(Z_{j_{|J|}}) \right) \tag{H.3e}$$

Denote $U_j = F_j(Z_j), \forall j\ 1, \ldots, |J|$, and then $U_j \sim \text{Unif}(0,1)$. Recall that β^* is equal to the optimal objective value of problem (6) when we fix $\mathbf{x}$ in problem (6) to be $\mathbf{x}^*$. Then for any permutation $j_1, j_2, \ldots, j_{|J|}$ of $1, 2, \ldots, |J|$, we have $\sum_{k=1}^{|J|} \hat{a}_{j_k} |x_{j_k}^*| \cdot q_{j_k,k} \le \beta^*$. Note that $\hat{a}_{j_k} |x_{j_k}^*| \ge 0$, so if $Z_{j_k} \le q_{j_k,k}, \forall k \in J$ hold, then $\sum_{k=1}^{|J|} \hat{a}_{j_k} |x_{j_k}^*| \cdot Z_{j_k} \le \beta^*$ holds. So we have

$$(H.3e) \tag{H.4a}$$

$$\ge \sum_{\{j_1, j_2, \ldots, j_{|J|}\}:\ \{j_1, j_2, \ldots, j_{|J|}\} = \{1,2,\ldots,|J|\}} \tag{H.4b}$$

$$\frac{1}{|J|!} \cdot \text{Prob}\left[Z_{j_k} \le q_{j_k,k}, \forall k \in J \ \middle|\ F_{j_1}(Z_{j_1}) \le F_{j_2}(Z_{j_2}) \le \ldots \le F_{j_{|J|}}(Z_{j_{|J|}}) \right] \tag{H.4c}$$

$$= \sum_{\{j_1, j_2, \ldots, j_{|J|}\}:\ \{j_1, j_2, \ldots, j_{|J|}\} = \{1,2,\ldots,|J|\}} \tag{H.4d}$$

$$\frac{1}{|J|!} \text{Prob}\left[F_{j_k}(Z_{j_k}) \le F_{j_k}(q_{j_k,k}), \forall k \in J \ \middle|\ F_{j_1}(Z_{j_1}) \le F_{j_2}(Z_{j_2}) \le \ldots \le F_{j_{|J|}}(Z_{j_{|J|}}) \right] \tag{H.4e}$$

$$= \sum_{\{j_1, j_2, \ldots, j_{|J|}\}:\ \{j_1, j_2, \ldots, j_{|J|}\} = \{1,2,\ldots,|J|\}} \tag{H.4f}$$

$$\frac{1}{|J|!} \text{Prob}\left[F_{j_k}(Z_{j_k}) \le Q_k^{1-\varepsilon_k}, \forall k \in J \ \middle|\ F_{j_1}(Z_{j_1}) \le F_{j_2}(Z_{j_2}) \le \ldots \le F_{j_{|J|}}(Z_{j_{|J|}}) \right] \tag{H.4g}$$

$$= \sum_{\{j_1, j_2, \ldots, j_{|J|}\}:\ \{j_1, j_2, \ldots, j_{|J|}\} = \{1,2,\ldots,|J|\}} \tag{H.4h}$$

$$\frac{1}{|J|!} \text{Prob}\left[U_{j_k} \le Q_k^{1-\varepsilon_k}, \forall k \in J \ \middle|\ U_{j_1} \le U_{j_2} \le \ldots \le U_{j_{|J|}} \right] \tag{H.4i}$$

Recall that $U_j \sim \text{Unif}(0,1)$, and because of symmetry, we have $\text{Prob}\ (U_{j_1} \le U_{j_2} \le \ldots \le U_{j_{|J|}}) = \frac{1}{|J|!}$ for any permulation $j_1, j_2, \ldots, j_{|J|}$. Then we have

$$\text{Prob}\left[\cap_{k=1}^{|J|} \left(U_{(k)} \le Q_k^{1-\varepsilon_k} \right) \right] \tag{H.5a}$$

$$= \sum_{\{j_1, j_2, \ldots, j_{|J|}\}:\ \{j_1, j_2, \ldots, j_{|J|}\} = \{1,2,\ldots,|J|\}} \tag{H.5b}$$

$$\frac{1}{|J|!} \text{Prob}\left[U_{j_k} \le Q_k^{1-\varepsilon_k}, \forall k \in J \ \middle|\ U_{j_1} \le U_{j_2} \le \ldots \le U_{j_{|J|}} \right], \tag{H.5c}$$

which is exactly (H.4i). So we have

$$\text{Prob}\left(\sum_j \hat{a}_j |x_j^*| \cdot Z_j \le \beta^* \right) \tag{H.6a}$$

$$= (H.3e) \tag{H.6b}$$

$$\ge (H.4i) \tag{H.6c}$$

$$= \text{Prob}\left[\cap_{k=1}^{|J|} \left(U_{(k)} \le Q_k^{1-\varepsilon_k} \right) \right] \tag{H.6d}$$

$$= |J|! \det[\Delta] \tag{H.6e}$$

Therefore, we have

$$\text{Prob}\left(\sum_j A_j x_j^* \le b \right) \tag{H.7a}$$

$$\ge \text{Prob}\left(\sum_j \hat{a}_j |x_j^*| \cdot \rho_j \le \beta^* \right) \tag{H.7b}$$

$$\geq \frac{1}{2} + \frac{1}{2} \cdot \text{Prob}\left(\sum_j \hat{a}_j |x_j^*| \cdot Z_j \leq \beta^* \right) \tag{H.7c}$$

$$\geq \frac{1}{2} + \frac{1}{2} \cdot |J|! \det[\Delta] \tag{H.7d}$$

References

Bandi, C., Bertsimas, D., & Youssef, N. (2015). Robust queueing theory. *Operations Research, 63*(3), 676–700.
Bandi, C., & Gupta, D. (2020). Operating room staffing and scheduling. *Manufacturing and Service Operations Management, 22*(5), 958–974.
Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust optimization*: vol. 28. Princeton University Press.
Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research, 23*(4), 769–805.
Ben-Tal, A., & Nemirovski, A. (2000). Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming, 88*(3), 411–424.
Bertsimas, D., Brown, D. B., & Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM Review, 53*(3), 464–501.
Bertsimas, D., Gupta, V., & Kallus, N. (2018). Data-driven robust optimization. *Mathematical Programming, 167*(2), 235–292.
Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research, 52*(1), 35–53.
Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media.
Chassein, A., Dokka, T., & Goerigk, M. (2019). Algorithms and uncertainty sets for data-driven robust shortest path problems. *European Journal of Operational Research, 274*(2), 671–686.
Chassein, A., & Goerigk, M. (2015). Alternative formulations for the ordered weighted averaging objective. *Information Processing Letters, 115*(6–8), 604–608.
Cheramin, M., Chen, R. L.-Y., Cheng, J., & Pinar, A., 2021. Data-driven robust optimization using scenario-induced uncertainty sets. arXiv preprint arXiv:2107.04977
Cvetkovski, Z. (2012). *Inequalities: Theorems, techniques and selected problems*. Springer Science & Business Media.
Dielman, T., Lowry, C., & Pfaffenberger, R. (1994). A comparison of quantile estimators. *Communications in Statistics-Simulation and Computation, 23*(2), 355–371.
El Ghaoui, L., Oustry, F., & Lebret, H. (1998). Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization, 9*(1), 33–52.
French, K. R., 2021. Data library. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Accessed 01 December 2021.
Glasserman, P., Heidelberger, P., & Shahabuddin, P. (2000). Variance reduction techniques for estimating value-at-risk. *Management Science, 46*(10), 1349–1364.
Gut, A. (2009). *An intermediate course in probability*. Springer-Verlag New York.
Roussas, G. G. (1997). *A course in mathematical statistics*. Elsevier.
Shang, C., Huang, X., & You, F. (2017). Data-driven robust optimization based on kernel learning. *Computers and Chemical Engineering, 106*, 464–479.
Steck, G. (1971). Rectangle probabilities for uniform order statistics and the probability that the empirical distribution function lies between two distribution functions. *The Annals of Mathematical Statistics, 42*(1), 1–11.
Yang, W., & Xu, H. (2016). Distributionally robust chance constraints for non-linear uncertainties. *Mathematical Programming, 155*, 231–265.