European Journal of Operational Research 311 (2023) 1159–1172

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Interfaces with Other Disciplines

Distortion risk measure under parametric ambiguity

Hui Shao[a], Zhe George Zhang[b,c,*]

[a] Zhejiang University, Haining 314499, China
[b] Western Washington University, Bellingham, WA 98225, USA
[c] Simon Fraser University, Burnaby V1M 2J2, Canada

ARTICLE INFO

Article history:
Received 7 April 2022
Accepted 15 May 2023
Available online 27 May 2023

Keywords:
Risk analysis
Extreme-case risk measure
Distortion function
Moment uncertainty
Robust optimization

ABSTRACT

This study develops closed-form solutions for distortion risk measures (DRM) in extreme cases by utilizing the first two moments and the symmetry of underlying distributions. The resultant extreme-case distributions, encompassing the worst- and best-case distributions, are identified by the envelopes of the distortion functions. The findings of this study extend previous research on worst-case risk measures such as worst-case VaR, worst-case CVaR, worst-case RVaR, and worst-case spectral risk measure, by presenting a unified framework. Furthermore, the compact solutions enhance tractability in optimization problems involving these risk measures, particularly when the true underlying distribution is unknown, and the first two moments are uncertain. The application of the extreme-case DRMs is illustrated with real data sets through numerical examples.

1. Introduction

The tradeoff between risk and return is a fundamental issue in most practical decision-making situations under uncertainty. The commonly used downside risk measures regarding time and money are usually based on the knowledge of probability distributions of relevant random variables. For example, the value-at-risk (hereafter, VaR) as the industry-standard reports the risk level of loss by computing an extremal quantile and its alternative conditional value at risk (hereafter, CVaR) provides the average loss exceeding an extremal quantile. All these risk measures require the specific probability distributions of the corresponding random variables. However, in practice, we may not have complete information about the probability distribution of a particular random variable. Instead, the first two moments of the random variable may be estimated based on an actual data set. This fact motivates researchers to develop extreme case risk measures which require only the first two moments information (we refer to Pichler & Xu, 2022; Popescu, 2007 for more detailed discussions). The paper aims to derive some closed-form solutions for the more general extreme-case risk measures when only the first two moments and symmetry of the underlying distributions are known. Furthermore, another critical issue in practice is that the data set for estimating the first two moments can be prone to error or bias. Thus, we should also investigate the performance of the extreme-case risk measures under moment uncertainty. Such an issue, called parametric ambiguity, is also addressed in this paper. The most famous example using the industry-standard risk measures or their extreme-case counterparts is the optimal portfolio selection problem. In a portfolio optimization model, the risk measures are regarding the investment return and the decision variables are the investment weights on a set of investment opportunities (e.g., stocks). In fact, other practical situations exist where the decision has to be made based on limited or incomplete information about random factors. For example, a manufacturer receives parts and components from several suppliers. The lead times (from ordering to receiving) for these suppliers are random variables that are correlated. The probability distributions for these random variables are unknown and their first two moments can be computed based on real data sets. If the on-time delivery time of some key parts is critical for uninterrupted supply-chain production, then determining the optimal proportion of orders for each supplier under extreme-case risk measures is of interest to the manufacturer. While the first example is about utilizing the extreme-case risk measures for the money, the second is for the time.

This paper attempts to derive the closed-form solutions for the extreme-case distortion risk measures by solving the stochastic optimization problems based on the partial information of the underlying distributions. Specifically, such a problem can be formulated

* Corresponding author at: Western Washington University, Bellingham, WA 98225, USA.

E-mail addresses: mathshao@gmail.com (H. Shao), george.zhang@wwu.edu (Z.G. Zhang).

https://doi.org/10.1016/j.ejor.2023.05.025
0377-2217/

as

$$\begin{aligned} &\inf_X \text{ or } \sup_X \qquad M_\phi(X) := \int_0^1 F_X^{-1}(u)\,\mathrm{d}\phi(u) \\ &\text{subject to} \qquad X \in \mathcal{P}, \end{aligned} \tag{1}$$

where $\mathcal{P}$ is the class of random variables with given mean, variance, and possibly other shape information, such as the symmetry of the underlying distributions. Here F_X^{-1} is the general inverse distribution function of the random variable X and ϕ is the distortion function that adjusts the true probability measure to give new weights to risk events (Wang, 2000). Notably, the optimization problem (1) can be seen as a dual version of the utility optimization problem, i.e., minimize or maximize $\mathbb{E}[u(X)]$ over a family $\mathcal{P}$ of random variables, by viewing the distribution function as its inverse (Yaari, 1987); we refer to Popescu (2007) for more details.

The optimization problem (1) is closely related to financial risk management since the objective function $\int_0^1 F^{-1}(u)\,\mathrm{d}\phi(u)$ is generally referred to as the distortion risk measure in the risk theory. Risk measures are introduced in return-risk tradeoff analysis that assigns real numbers to the loss distributions. Value-at-Risk, one of the most popular risk measures, determines the potential loss in the distributions being assessed at a given probability level. Despite the popularity of VaR, it fails to be a coherent measure. Conditional Value-at-Risk (also called Expected Shortfall), defined as the average of the tail distribution exceeding VaR, is a popular coherent risk measure and has attracted much attention in this area (Artzner et al., 1999). As a bridge between VaR and CVaR, Range Value-at-Risk (RVaR) is first proposed by Cont et al. (2010), which is regarded as a robust risk measure in the sense that it is continuous with respect to weak convergence of random variables. Spectral risk measures (hereafter, SRM) are more general coherent risk measures introduced by Acerbi (2002), which models the risk aversion through weights given to the quantiles of different levels. At the top of the listed risk measures, distortion risk measures are more general and are developed from the research on premium principles by Wang (1995), which are accompanied by their associated distortion functions. By choosing appropriate distortion functions, distortion risk measures become some popular risk measures such as VaR, CVaR, RVaR, and SRM as special cases.

The optimization problem (1) has been partially studied in the optimization literature, and it is generally called the worst-case problem. Ghaoui et al. (2003) consider the worst-case VaR problem when the first two moments of the underlying distributions are known and derive the closed-form solution to it. Chen et al. (2011) and Natarajan et al. (2010) obtain the closed-form solutions for worst-case CVaR when the first two moments are known; also see Toumazis & Kwon (2015) for using the worst-case CVaR in the problem of transporting hazardous materials. Peposcu (2005) incorporates the shape constraints such as symmetry and unimodality into the stochastic programming problems in addition to the moment constraints on distribution functions. Li et al. (2017) studies the closed-form solutions for worst-case RVaR with the first two moments and other shape information such as symmetry and unimodality. More recently, Guo & Xu (2021) considers the worst-case law invariant coherent risk measure when the ambiguous set is composed of robust spectral risk measures. Chen et al. (2020) studies the moment problem in distribution-free robust optimization, where the goal is to find a worst-case distribution that satisfies a given set of moments. Notably, the metrics in probability spaces describe another critical kind of worst-case risk measure problem. Pichler (2014) studies the worst-case risk measure problem when the underlying distributions are provided by the Wasserstein distance and the results turn out to be useful in pricing insurance contracts. Interested readers are referred to Pichler (2013) and Pichler & Xu (2022) for more risk measure problems associated with Wasserstein distance. For more related worst-case problems, see Berkhouch (2021); Bertazzi & Secomandi (2020); Bertsimas et al. (2010); Das et al. (2021); Guo et al. (2022); Li (2018); Wiesemann et al. (2014); Zymler et al. (2013), and Natarajan et al. (2008). For more recent and relevant work on robust portfolio management, we refer to Chen et al. (2019); Sahamkhadam et al. (2022), and Staden et al. (2021), among others.

Motivated by the closed-form solutions for the worst-case risk measures reviewed above, it is natural to consider whether the closed-form solutions exist for the more general distortion risk measures that generalize and unify the existing results. In this paper, we give an affirmative answer to this question. It is worth noting that we approach the optimization problem in a unique way. While most of the literature uses the standard optimization techniques to derive the results (see Popescu, 2007; Wiesemann et al., 2014; Zhu & Fukushima, 2009), our study is based on the Cauchy–Schwarz inequality combined with some calculus rules applied strategically. The main contribution of this work is twofold:

(1) With the first two moments and symmetry of the underlying distributions, we derive the closed-form solutions for the extreme-case distortion risk measures and characterize their corresponding extreme-case distributions by the envelopes of the distortion functions. More importantly, the closed-form solutions generalize many well-known worst-case risk measures in a unified framework, including the worst-case VaR, worst-case CVaR, worst-case RVaR, and worst-case SRM.
(2) With closed-form solutions for the worst-case distortion risk measures, we further investigate the impact of the moment uncertainties on the robust optimization problem through an empirical study.

The rest of the paper is structured as follows. The next section is focused on the main analytical results (Theorems 2.1 and 2.2). Section 3 presents the numerical illustrations to demonstrate the application of the results. Finally, Section 4 concludes the paper with some remarks. Some of the lengthy technical proofs are relegated to the Appendices.

2. Main results

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space and assume that all random variables considered in the paper are in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Denote by $F_X(x)$ the distribution function of random variable X, i.e., $F_X(t) = \mathbb{P}(X \le t)$, and denote by $F_X^{-1}(\alpha) = \inf\{x : F_X(x) \ge \alpha\}$ the inverse distribution function of X for $\alpha \in [0, 1]$. Meanwhile, we denote by $F^{-1+}(\alpha)$ the right-continuous inverse distribution function, i.e., $F_X^{-1+}(\alpha) = \inf\{x : F_X(x) > \alpha\}$. Note that the left- and right-continuous inverse distribution function are identical except on countable many points in [0,1]. Further, we let $X > 0$ representing loss and $X < 0$ representing gain. For notation convenience, we may alternatively denote by $q(u)$ the inverse distribution functions $F_X^{-1}(u)$ or $F_X^{-1+}(u)$ in the following development when there is no ambiguity. Denote by $\|f(u)\|_p$ the p-norm of f on the unit interval, i.e., $\|f(u)\|_p = (\int_0^1 |f(u)|^p\,du)^{\frac{1}{p}}$ for $p \ge 1$. In particular, all integrals are assumed to be finite from a practical point of view. Next we start with the definition of distortion risk measure (Wang, 2000; Yaari, 1987).

Definition 1 (Distortion risk measure)**.** The distortion risk measure is defined as

$$M_\phi(X) = -\int_{-\infty}^{0} \phi(F_X(x))\,dx + \int_0^{+\infty} 1 - \phi(F_X(x))\,dx,$$

where ϕ refers to a distortion function belonging to the following set:

$$\mathcal{D} = \{\phi : [0,1] \to [0,1] \mid \phi \text{ is non-decreasing, continuous on } 0 \text{ and } 1, \text{ and } \phi(0) = 0,\ \phi(1) = 1\}.$$

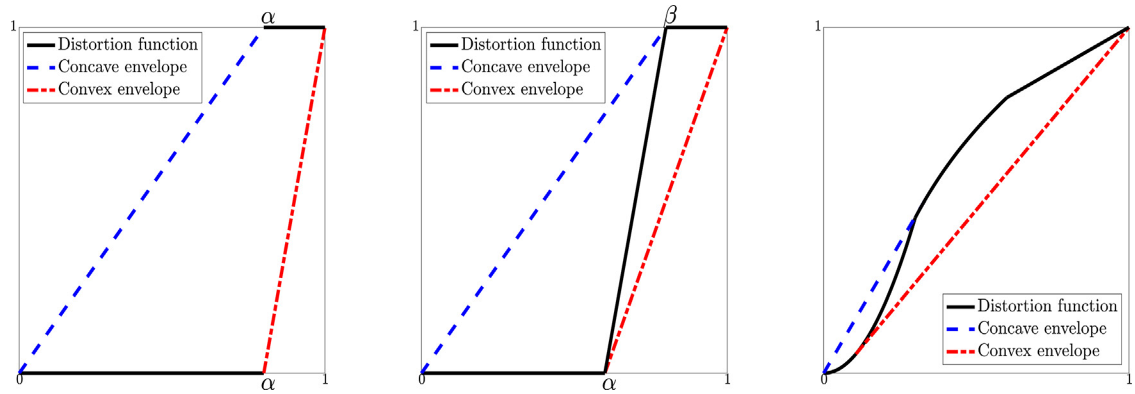

Fig. 1. Three examples of distortion functions, the convex envelope functions, and the concave envelope functions. The left panel presents the distortion function of VaR_α. The middle panel presents the distortion function of $\mathrm{RVaR}_{\alpha,\beta}$. The right panel presents an S-shape distortion function.

In particular, we call $\phi(u) = u$ for all $u \in [0, 1]$ as a trivial distortion function.

Notably, any distortion risk measure can be written as a Lebesgue–Stieltjes integral when its associated distortion function is left- or right-continuous (see Lemma B.1 in the appendices), and the Lebesgue–Stieltjes integral form plays a key role in the development of main results.

Definition 2. A random variable X is symmetrical if there is a constant m such that $\mathbb{P}(X \le x) = \mathbb{P}(X \ge 2m - x)$ for all $x \in \mathbb{R}$.

It is direct to see that $F_X^{-1}(u) + F_X^{-1+}(1-u) = 2\mathbb{E}[X]$ holds for all $u \in (0, 1)$ when the random variable X is symmetrical. In particular, we see that $F_X^{-1}(u) + F_X^{-1}(1-u) = 2\mathbb{E}[X]$ and $F_X^{-1+}(u) + F_X^{-1+}(1-u) = 2\mathbb{E}[X]$ hold a.e. (almost everywhere) when $u \in (0, 1)$ since $F_X^{-1+}(u) = F^{-1}(u)$ a.e. for all $u \in (0, 1)$.

Typical examples for the symmetrical random variables include the normal, the uniform, and the student-t, etc. More notations are introduced here. For a given pair of mean and variance $(\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+$, denote by $\mathcal{P}(\mu, \sigma^2)$ the family of random variables with mean μ and variance σ^2, and denote by $\mathcal{P}_S(\mu, \sigma^2)$ the family of symmetrical random variables with mean μ and variance σ^2, namely

$$\mathcal{P}(\mu, \sigma^2)=\{X \in L^2(\Omega, \mathcal{F}, \mathbb{P}) \mid \mathbb{E}[X] = \mu \text{ and } \mathrm{Var}[X] = \sigma^2\},$$
$$Y\mathcal{P}_S(\mu, \sigma^2)=\{X \in L^2(\Omega, \mathcal{F}, \mathbb{P}) \mid \mathbb{E}[X] = \mu, \mathrm{Var}[X] = \sigma^2, \text{ and } X \text{ is symmetrical}\};$$

Definition 3. For a distortion function ϕ, the convex and concave envelopes of ϕ are defined, respectively, by

$$\phi_* = \sup\{g : [0, 1] \to [0, 1] \mid g(u) \le \phi(u) \text{ for } u \in [0, 1] \text{ and } g \text{ is convex}\},$$
$$\phi^* = \inf\{g : [0, 1] \to [0, 1] \mid g(u) \ge \phi(u) \text{ for } u \in [0, 1] \text{ and } g \text{ is concave}\}.$$

Note that the convex and concave envelopes of a function ϕ satisfy the relation $(-\phi)_* = -\phi^*$. Such a relation plays a key role in the following derivation of the closed-form solutions for the extreme-case distortion risk measures (Fig. 1).

2.1. DRM with convex distortion functions

In this section, we develop the extreme-case DRM when the first two moments as well as the symmetry of the underlying distributions are known. We start with the univariate case and then extend the results to the multivariate case. Coherent risk measures play a crucial role in the modern risk theory. Indeed, DRM reduces a coherent one if its associated distortion function is convex (Acerbi, 2002; Artzner et al., 1999). The following two propositions present the closed-form solutions for worst-case DRMs with convex distortion functions. Similar to Peposcu (2005)'s framework, the proposition also considers the symmetry of the underlying distributions into account.

It is obvious that $M_\phi(X) = \mathbb{E}(X)$ holds for all X when the distortion function ϕ is trivial, then we have $\sup_{X\in\mathcal{P}(\mu,\sigma^2)} M_\phi(X) = \mu$, and the supremum can be attained at any $X \in \mathcal{P}(\mu, \sigma^2)$. Therefore, it suffices to consider non-trivial distortion functions only in the following proposition.

Proposition 2.1 (Worst-case convex DRM under. $\mathcal{P}(\mu, \sigma^2)$*) Let* $(\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+$ *and* ϕ *be a non-trivial convex distortion function, then*

$$\sup_{X\in\mathcal{P}(\mu,\sigma^2)} M_\phi(X) = \mu + \sigma\|\phi'(u) - 1\|_2, \tag{2}$$

where ϕ' *denotes the right derivative function of* ϕ*. Moreover, the supremum in (2) is attained if and only if the worst-case distribution* F *satisfies* $F^{-1}(u) = \mu + \sigma(\phi'(u) - 1)/\|\phi'(u) - 1\|_2$ *a.e.*

Proof. By changing of variables, $X \in \mathcal{P}(\mu, \sigma^2)$ implies that

$$\int_0^1 q(u)\,\mathrm{d}u = \mu \quad\text{and}\quad \int_0^1 (q(u) - \mu)^2\,\mathrm{d}u = \sigma^2. \tag{3}$$

It is obvious that ϕ is absolutely continuous since it is convex on [0,1], thus by Lemma B.1 we have $M_\phi(X) = \int_0^1 q(u)\mathrm{d}\phi(u)$ where $q(u) = F^{-1}(u)$, then

$$M_\phi(X) = \int_0^1 q(u)\phi'(u)\,\mathrm{d}u = \mu + \int_0^1 \big(q(u) - \mu\big)\big(\phi'(u) - 1\big)\,\mathrm{d}u. \tag{4}$$

Applying the Cauchy–Schwarz inequality to the right-hand side of Eq. (4) gives

$$M_\phi(X) \le \mu + \|q(u) - \mu\|_2 \cdot \|\phi'(u) - 1\|_2 = \mu + \sigma\|\phi'(u) - 1\|_2, \tag{5}$$

where the equality holds if and only if $q(u) - \mu = k_1(\phi'(u) - 1)$ *a.e.* for some $k_1 > 0$ by the Cauchy–Schwarz inequality. By the variance constraint in (3) and non-triviality of ϕ, we have $k_1 = \sigma/\|\phi'(u) - 1\|_2$, thus $q(u) = \mu + \sigma(\phi'(u) - 1)/\|\phi'(u) - 1\|_2$ *a.e.* Now the function $q(u)$ is obviously increasing as $\phi'(u)$ is increasing. Moreover, it also satisfies the mean and variance constraints in Eq. (3), which is the optimal solution to the worst-case problem $\sup_{X\in\mathcal{P}(\mu,\sigma^2)} M_\phi(X)$. □

Proposition 2.2 (Worst-case convex DRM under. $\mathcal{P}_S(\mu, \sigma^2)$*) Let* $(\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+$ *and* ϕ *be a convex distortion function, then*

$$\sup_{X\in\mathcal{P}_S(\mu,\sigma^2)} M_\phi(X) = \mu + \frac{\sigma}{2}\|\phi'(u) - \phi'(1-u)\|_2. \tag{6}$$

Moreover, the supremum in (6) is attained if and only if the worst-case distribution F satisfies

$$F^{-1}(u) = \mu + \sigma \frac{\phi'(u) - \phi'(1-u)}{\|\phi'(u) - \phi'(1-u)\|_2} \text{ a.e.},$$

when $\|\phi'(u) - \phi'(1-u)\|_2 \neq 0$.

Proof. It is direct that $q(u) + q(1-u) = 2\mu$ *a.e.* for all $u \in (0,1)$ when $X \in \mathcal{P}_S(\mu, \sigma^2)$, then similar with Proposition 2.1 we have

$$M_\phi(X) = \int_0^1 q(u)\phi'(u)\,du = \int_0^1 q(1-u)\phi'(1-u)\,du$$
$$= \int_0^1 (2\mu - q(u))\phi'(1-u)\,du = 2\mu - \int_0^1 q(u)\phi'(1-u)\,du.$$

Rearranging the above equation gives

$$M_\phi(X) = \mu + \frac{1}{2}\int_0^1 q(u)\big(\phi'(u) - \phi'(1-u)\big)$$
$$du = \mu + \frac{1}{2}\int_0^1 \big(q(u) - \mu\big)\big(\phi'(u) - \phi'(1-u)\big)\,du. \quad (7)$$

Applying the Cauchy–Schwarz inequality to the right-hand side of Eq. (7) then yields

$$M_\phi(X) \le \mu + \frac{1}{2}\|q(u) - \mu\|_2 \cdot \|\phi'(u) - \phi'(1-u)\|_2$$
$$= \mu + \frac{\sigma}{2}\|\phi'(u) - \phi'(1-u)\|_2, \quad (8)$$

where the equality holds if and only if $q(u) - \mu = k_2(\phi'(u) - \phi'(1-u))$ *a.e.* for some $k_2 > 0$ by the Cauchy–Schwarz inequality. By the variance constraints in $\mathcal{P}_S(\mu, \sigma^2)$ together with $\|\phi'(u) - \phi'(1-u)\|_2 \neq 0$, we have $k_2 = \sigma / \|\phi'(u) - \phi'(1-u)\|_2$, therefore $q(u) = \mu + \frac{\sigma}{2}\|\phi'(u) - \phi'(1-u)\|_2$ *a.e.* Now $q(u)$ is obviously increasing as $\phi'(u)$ is increasing, and it also satisfies the mean, variance, and symmetry constraints in $\mathcal{P}_S(\mu, \sigma^2)$, which is exactly the optimal solution to the worst-case optimization problem $\sup_{X \in \mathcal{P}_S(\mu,\sigma^2)} M_\phi(X)$. □

Remark 1. Note that when $\|\phi'(u) - \phi'(1-u)\|_2 = 0$, which is equivalent to $\phi'(u) - \phi'(1-u) = 0$ *a.e.*, we have $M_\phi(X) = \mathbb{E}[X]$ for all X by (7). Thus $\sup_{X \in \mathcal{P}_S(\mu,\sigma^2)} M_\phi(X) = \mu$, and the supremum can be attained at any $X \in \mathcal{P}_S(\mu, \sigma^2)$.

Different from the most literature on the worst-case problems, e.g., Popescu (2007); Zhu & Fukushima (2009), and Wiesemann et al. (2014), which generally focus on the duality method, we solve the optimization problem simply by the well-known Cauchy–Schwarz inequality and some fundamental calculus techniques.

The most notable feature of DRM with convex distortion functions is perhaps its connection with SRM. Recall that SRM is defined as $\rho_g(X) = \int_0^1 F_X^{-1}(u)g(u)\,du$, where $g : [0,1] \to [0,1]$ is called an "admissible risk spectrum" that is increasing and satisfies $\int_0^1 g(u)\,du = 1$. Applying Proposition 2.1 with the distortion function $\phi(t) = \int_0^t g(u)\,du$, it is straightforward to obtain the results in the following corollary.

Corollary 2.1 (Worst-case SRM under. $\mathcal{P}(\mu, \sigma^2)$ *and* $\mathcal{P}_S(\mu, \sigma^2)$**)** *The worst cases of SRM,* $S_g(X) = \int_0^1 F^{-1}(u)g(u)\,du$*, are given by*

$$\sup_{X \in \mathcal{P}(\mu,\sigma^2)} S_g(X) = \mu + \sigma\|g(u) - 1\|_2,$$

$$\sup_{X \in \mathcal{P}_S(\mu,\sigma^2)} S_g(X) = \mu + \frac{\sigma}{2}\|g(u) - g(1-u)\|_2\,.$$

Moreover, the associated worst-case distributions can be obtained explicitly by the above two propositions.

To the best of our knowledge, Li (2018) is the first study on the worst-case SRM based on the first two moments and presents the closed-form solutions. Applying Proposition 2.1 with the distortion function $\phi(t) = \int_0^t g(u)\,du$, it is straightforward to see that the results in the above proposition are in agreement with Li (2018).

2.2. DRM with general distortion functions

In this section we proceed to study the extreme-case DRM with general distortion functions. For technical considerations, we restrict the distortion functions to the following two families:

$$\mathcal{D}_+ = \{\phi \in \mathcal{D} \mid \ \phi \text{ is right-continuous and and } \phi'_* \in L^2((0,1))\},$$

$$\mathcal{D}_- = \{\phi \in \mathcal{D} \mid \ \phi \text{ is left-continuous and and } \phi'_* \in L^2((0,1))\}.$$

Note that for a general distortion function $\phi \in \mathcal{D}$, the left limit is pointwise less than the right limit of ϕ, i.e., $\phi_- \le \phi \le \phi_+$, where $\phi_-(u) = \lim_{v \to u-}\phi(v)$ and $\phi_+(u) = \lim_{v \to u+}\phi(v)$ for all $u \in (0,1)$. Then by definition we have

$$M_{\phi_+}(X) \le M_\phi(X) \le M_{\phi_-}(X), \quad \text{for all } \phi \in \mathcal{D}. \quad (9)$$

Therefore, based on inequality (9) and without losing generality, we derive the worst-case DRM under left-continuous distortion function, and derive the best-case DRM under right-continuous distortion function in the following development.

Theorem 2.1 (Extreme-case DRM under. $\mathcal{P}(\mu, \sigma^2)$**)** *Let* $(\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+$ *and* ϕ *be a distortion function with non-trivial envelope, then*

$$\sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_\phi(X) = \mu + \sigma\|\phi'_*(u) - 1\|_2, \quad \text{if } \phi \in \mathcal{D}_-, \quad (10)$$

$$\inf_{X \in \mathcal{P}(\mu,\sigma^2)} M_\phi(X) = \mu - \sigma\|\phi^{*\prime}(u) - 1\|_2, \quad \text{if } \phi \in \mathcal{D}_+. \quad (11)$$

Moreover, the supremum in (10) is attained if and only if the worst-case distribution F satisfies

$$F^{-1}(u) = \mu + \sigma \frac{\phi'_*(u) - 1}{\|\phi'_*(u) - 1\|_2} \text{ a.e.},$$

and the infimum in (11) is attained if and only if the best-case distribution F satisfies

$$F^{-1}(u) = \mu - \sigma \frac{\phi^{*\prime}(u) - 1}{\|\phi^{*\prime}(u) - 1\|_2} \text{ a.e.}$$

Proof. This proof will only focus on the worst-case DRM, since the best-case DRM can be directly obtained by the results of the worst case, noting the simple relationship $M_\phi(X) = -M_\psi(-X)$, where $\psi(u) = 1 - \phi(1-u)$. Taking infimum on both sides yields $\inf_X M_\phi(X) = -\sup_X M_\psi(-X)$. Thus, it suffices for us to consider the worst case. Due to limited space, some preliminaries needed are provided in Appendix B such as Lemmas B.2, B.3, B.4, and Proposition B.1, which are referred in the rest of this proof.

For the distortion function $\phi \in \mathcal{D}_-$, let $\{\phi_n\}_{n \ge 1} \subseteq \mathcal{D}_-$ be any sequence of piecewise constant distortion functions satisfying

$$\phi_n \ge \phi, \ \ n \ge 1, \ \text{ and } \lim_{n \to +\infty}\|\phi_n - \phi\|_\infty = 0.$$

By results (*i*) and (*ii*) in Lemma B.3, we obtain that $(\phi_n)_*(u) \ge \phi_*(u)$ and $\lim_{n \to +\infty}(\phi_n)_*(u) = \phi_*(u)$ for all $u \in [0,1]$. Next we prove that the following limit holds for all $X \in \mathcal{P}(\mu, \sigma^2)$,

$$\lim_{n \to +\infty} M_{(\phi_n)_*}(X) = M_{\phi_*}(X). \quad (12)$$

Indeed, for any $s \in (0, \frac{1}{2})$ we have

$$\left|M_{(\phi_n)_*}(X) - M_{\phi_*}(X)\right| = \left|\int_0^1 q(u)\,d\big((\phi_n)_*(u) - \phi_*(u)\big)\right|$$
$$\le \left|\int_0^s q(u)\big((\phi_n)'_*(u) - \phi'_*(u)\big)\,du\right| + \left|\int_s^{1-s} q(u)\,d\big((\phi_n)_*(u) - \phi_*(u)\big)\right|$$
$$+ \left|\int_{1-s}^1 q(u)\big((\phi_n)'_*(u) - \phi'_*(u)\big)\,du\right|$$

$$\leq\sqrt{\int_0^s q^2(u)\,\mathrm{d}u\int_0^s\left((\phi_n)'_*(u)-\phi'_*(u)\right)^2\mathrm{d}u}$$
$$+\left|q(u)\left((\phi_n)_*(u)-\phi_*(u)\right)\Big|_s^{1-s}-\int_s^{1-s}\left((\phi_n)_*(u)-\phi_*(u)\right)\mathrm{d}q(u)\right|$$
$$+\sqrt{\int_{1-s}^1 q^2(u)\mathrm{d}u\int_{1-s}^1\left((\phi_n)'_*(u)-\phi'_*(u)\right)^2\mathrm{d}u},\qquad(13)$$

where the last inequality holds due to the Cauchy–Schwarz inequality and integration by parts formula.

Note that for any convex distortion function f, we always have $f'(u)\leq\frac{f(1)-f(u)}{1-u}\leq\frac{1}{1-u}$. Applying it to $(\phi_n)_*$ and ϕ_* as both of them are increasing and convex, we obtain that for any $u\in[0,s]$,

$$0\leq(\phi_n)'_*(u)\leq\frac{1}{1-s}\quad\text{and}\quad 0\leq\phi'_*(u)\leq\frac{1}{1-s}.$$

Thus we have $\left((\phi_n)'_*(u)-\phi'_*(u)\right)^2\leq 1/(1-s)^2$ for all $u\in[0,s]$. Integrating both sides with respect to u then yields the following inequality,

$$\int_0^s\left((\phi_n)'_*(u)-\phi'_*(u)\right)^2\mathrm{d}u\leq\frac{s}{(1-s)^2}\leq 4s.\qquad(14)$$

On the other hand, it is direct to see that

$$\int_{1-s}^1\left((\phi_n)'_*(u)-\phi'_*(u)\right)^2\mathrm{d}u\leq\int_{1-s}^1\left(\left((\phi_n)'_*(u)\right)^2\right.$$
$$\left.+\left(\phi'_*(u)\right)^2\right)\mathrm{d}u\leq 2\int_{1-s}^1\left(\phi'_*(u)\right)^2\mathrm{d}u,\qquad(15)$$

where the first inequality holds since both $(\phi_n)'_*$ and ϕ'_* are nonnegative, and the second inequality holds due to the fact that $(\phi_n)_*(u)\geq\phi_*(u)$ for all $n\geq1$ and $u\in[0,1]$ by Lemma B.4.

Since $X\in L^2(\Omega,\mathcal{F},\mathbb{P})$ and $\phi'_*\in L^2([0,1])$, inequalities (14) and (15) indicate that for any arbitrary $\epsilon>0$, there exists $\delta\in(0,\frac{1}{2})$ such that

$$\sqrt{\int_0^\delta q^2(u)\mathrm{d}u\int_0^\delta\left((\phi_n)'_*(u)-\phi'_*(u)\right)^2\mathrm{d}u}$$
$$+\sqrt{\int_{1-\delta}^1 q^2(u)\,\mathrm{d}u\int_{1-\delta}^1\left((\phi_n)'_*(u)-\phi'_*(u)\right)^2\mathrm{d}u}<\frac{\epsilon}{2}.\qquad(16)$$

For this specific δ and then by the convergence of $(\phi_n)_*$, we assert that there is a sufficient large $N>0$ such that for all $n\geq N$,

$$\left|q(u)\left((\phi_n)_*(u)-\phi_*(u)\right)\Big|_\delta^{1-\delta}-\int_\delta^{1-\delta}\left((\phi_n)_*(u)-\phi_*(u)\right)\mathrm{d}q(u)\right|<\frac{\epsilon}{2}.\qquad(17)$$

Indeed, for the first term in (17) we have

$$\left|q(u)\left((\phi_n)_*(u)-\phi_*(u)\right)\Big|_\delta^{1-\delta}\right|\leq\max\{|q(\delta)|,|q(1-\delta)|\}$$
$$\times\left(\left|(\phi_n)_*(1-\delta)-\phi_*(1-\delta)\right|+\left|(\phi_n)_*(\delta)-\phi_*(\delta)\right|\right).$$

Letting $n\to+\infty$ yields the conclusion that $q(u)\left((\phi_n)_*(u)-\phi_*(u)\right)\Big|_\delta^{1-\delta}\to0$. For the second term in (17) we have

$$\left|\int_\delta^{1-\delta}\left((\phi_n)_*(u)-\phi_*(u)\right)\mathrm{d}q(u)\right|\leq\|(\phi_n)_*-\phi_*\|_\infty\cdot\left|q(1-\delta)-q(\delta)\right|,$$

thus by results (*ii*) Lemma B.3 we obtain that $\lim_{n\to+\infty}\int_\delta^{1-\delta}\left((\phi_n)_*(u)-\phi_*(u)\right)\mathrm{d}q(u)=0$. Therefore, by combining the above two analyses we see that (17) holds.

Substituting two inequalities (16) and (17) into inequality (13), we obtain the limit in (12), i.e., $\lim_{n\to+\infty}M_{(\phi_n)_*}(X)=M_{\phi_*}(X)$.

Next we focus on the upper bound of the distortion risk measure. Lemma B.2 indicates that $M_{\phi_n}(X)\leq M_\phi(X)\leq M_{\phi_*}(X)$ since $\phi_*(u)\leq\phi(u)\leq\phi_n(u)$ for $u\in[0,1]$. Taking supremum for the joint inequality then yields

$$\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_{\phi_n}(X)\leq\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_\phi(X)\leq\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_{\phi_*}(X).\qquad(18)$$

On the other hand, by applying Propositions 2.1 and B.1, we have

$$\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_{\phi_n}(X)=\mu+\sigma\|(\phi_n)'_*(u)-1\|_2=\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_{(\phi_n)_*}(X).\qquad(19)$$

Hence, it follows from (18), (19) that

$$M_{(\phi_n)_*}(X_*)\leq\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_{(\phi_n)_*}(X)=\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_{\phi_n}(X)\leq\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_{\phi_*}(X),\qquad(20)$$

where X_* is the worst-case random variable with respect to the measure $M_{\phi^*}(X)$, i.e., X_* satisfies $M_{\phi_*}(X_*)=\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_{\phi_*}(X)$ (see Proposition 2.1 for the existence). Finally, letting $n\to\infty$ in (20) and then applying (12) yields the conclusion that $\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_\phi(X)=\sup_{X\in\mathcal{P}(\mu,\sigma^2)}M_{\phi_*}(X)$, which completes the proof. □

As an interesting application, the distortion function is chosen to be the one that corresponds to symmetrical quantile average, then the following result is obtained.

Proposition 2.3 (The distance between the symmetrical quantile average and the mean)**.** *Let $(\mu,\sigma^2)\in\mathbb{R}\times\mathbb{R}_+$ and $\alpha\in[\frac{1}{2},1)$, then for any $X\in\mathcal{P}(\mu,\sigma^2)$ we have*

$$\left|\frac{F^{-1}(\alpha)+F^{-1}(1-\alpha)}{2}-\mu\right|\leq\sigma B(\alpha),\qquad(21)$$

where

$$B(\alpha)=\begin{cases}\sqrt{\frac{\alpha}{4(2\alpha-1)(1-\alpha)}}-1, & \frac{2}{3}\leq\alpha<1,\\ \sqrt{\frac{1-\alpha}{\alpha}}, & \frac{1}{2}\leq\alpha<\frac{2}{3}.\end{cases}$$

Moreover, the bounds in (21) is sharpe due to the explicit results in Theorem 2.1. In particular when $\alpha=\frac{1}{2}$, we obtain the maximal distance between the median and mean of a random variable with fixed mean μ and variance σ^2:

$$\left|F^{-1}(\frac{1}{2})-\mu\right|\leq\sigma,\qquad(22)$$

showing that the maximal distance is at most its standard deviation. The equality holds in (22) when the random variable X with distribution function F satisfies $\mathbb{P}(X=\mu+\sigma)=\mathbb{P}(X=\mu-\sigma)=\frac{1}{2}$.

The proof of this proposition is provided in Appendix A.

The extreme-case DRM for the symmetrical case is treated similarly and presented in the following theorem.

Theorem 2.2 (Extreme-case DRM under. $\mathcal{P}_S(\mu,\sigma^2)$**)** *Let $(\mu,\sigma^2)\in\mathbb{R}\times\mathbb{R}_+$ and ϕ be a distortion function, then*

$$\sup_{X\in\mathcal{P}_S(\mu,\sigma^2)}M_\phi(X)=\mu+\frac{\sigma}{2}\|\phi'_*(u)-\phi'_*(1-u)\|_2,\quad\text{if }\phi\in\mathcal{D}_-,\qquad(23)$$

$$\inf_{X\in\mathcal{P}_S(\mu,\sigma^2)}M_\phi(X)=\mu-\frac{\sigma}{2}\|\phi^{*\prime}(u)-\phi^{*\prime}(1-u)\|_2,\quad\text{if }\phi\in\mathcal{D}_+.\qquad(24)$$

Moreover, the supremum in (23) is attained if and only if the worst-case distribution F satisfies

$$F^{-1}(u)=\mu+\sigma\frac{\phi'_*(u)-\phi'_*(1-u)}{\|\phi'_*(u)-\phi'_*(1-u)\|_2}\ \textit{a.e.}$$

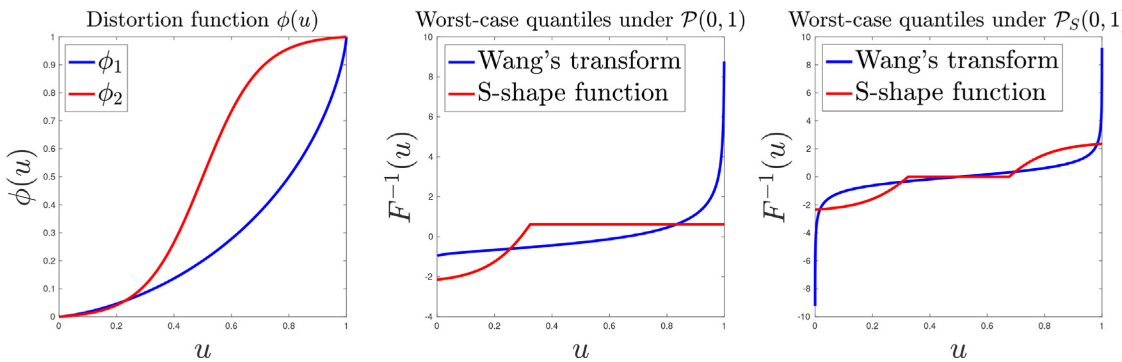

Fig. 2. The left panel presents two distortion functions: $\phi_1(u)=\Phi(\Phi^{-1}(u)+\Phi^{-1}(\alpha))$ (Wang's transform distortion function) and $\phi(u)=(\frac{1}{1+e^{-\beta(2u-1)}}-\frac{1}{1+e^{\beta}})/(\frac{1}{1+e^{-\beta}}-\frac{1}{1+e^{\beta}})$ (S-shaped distortion function), where $\alpha=0.2$ and $\beta=5$. The middle panel presents the worst-case quantile functions under the general constraints $\mathcal{P}(0,1)$. The right panel presents the worst-case quantile functions under the symmetrical constraints $\mathcal{P}_S(0,1)$.

when $\|\phi_*'(u)-\phi_*'(1-u)\|_2\neq 0$, *and the infimum in (24) is attained if and only if the best-case distribution F satisfies*

$$F^{-1}(u)=\mu-\sigma\frac{\phi^{*\prime}(u)-\phi^{*\prime}(1-u)}{\|\phi^{*\prime}(u)-\phi^{*\prime}(1-u)\|_2}\ a.e.$$

when $\|\phi^{*\prime}(u)-\phi^{*\prime}(1-u)\|_2\neq 0$.

Proof. The proof is similar to that for the Theorem 2.1 if the distortion function ϕ is replaced by its "symmetrical function" $\bar{\phi}(u)=(\phi(u)+\phi(1-u))/2$ (see the proof in Proposition B.1 in the Appendix B). Thus, the detailed derivation is omitted. □

The two theorems above indicate that both worst-case and best-case distributions can be characterized explicitly by the convex and concave envelopes of the corresponding distortion functions, with a pair of location and scale parameters (Fig. 2).

The main difficulties in the proof for the worst case are to show that the worst-case DRM is equivalent to that with the convex envelope of the corresponding distortion function, that is $\sup_X M_\phi(X)=\sup_X M_{\phi_*}(X)$. To this end, we take three steps to complete the proof. Firstly, we prove that it holds for convex distortion functions, i.e., Proposition 2.1. Secondly, we prove that it holds for piecewise constant distortion functions, i.e., Proposition B.1. Lastly, we prove that it holds for general distortion functions by approximation methods. The technical details of this three-step proof are provided in the Appendices.

The optimal bounds and extreme-case distribution formulas in Theorems 2.1 and 2.2 are extremely powerful, as they include many well-known worst-case risk measures as special cases. Table 1 below summarizes these cases.

Remark 2. The worst-case VaR (Ghaoui et al., 2003; Peposcu, 2005), worst-case CVaR (Chen et al., 2011; Natarajan et al., 2010; Zhu & Fukushima, 2009), and worst-case RVaR (Li et al., 2017) when the first two moments of the underlying distributions are available have been well understood in literature. In particular, we find that the worst-case values for these three risk measures are all the same as $\mu+\sigma\sqrt{\frac{\alpha}{1-\alpha}}$ (and $\mu+\sigma\sqrt{\frac{1}{2(1-\alpha)}}$ for the symmetrical case). However, there is no literature providing an explanation for such a result so far (i.e., why they are the same?). By using the main results of this paper, we are able to explain such a phenomenon clearly by noting that these worst-case risk measures have the same convex envelopes for their associated distortion functions.

Obviously, the worst-case values subject to the constraint $\mathcal{P}(\mu,\sigma^2)$ are generally greater than that subject to the symmetrical constraint $\mathcal{P}_S(\mu,\sigma^2)$. Figure 3 illustrates the upper bounds for the VaR, the DRM with exponential utility, and the DRM with power utility distortion function, subject to general and symmetrical constraints numerically, respectively.

2.3. Some corollaries: multivariate case

Theorem 2.1 can be easily extended to the multivariate case.

Corollary 2.2. *Let* $(\mu_i,\sigma_i^2)\in\mathbb{R}\times\mathbb{R}^+$ *for* $i=1,2,\ldots,n$ *and* ϕ *be a general distortion function, then*

1) $\sup_{X_i\in\mathcal{P}(\mu_i,\sigma_i^2),\ i=1,2,\ldots,n} M_\phi(\sum_{i=1}^n X_i)=\sum_{i=1}^n \sup_{X_i\in\mathcal{P}(\mu_i,\sigma_i^2)} M_{\phi_*}(X_i)$.
2) *The equality in the above equation holds if (1)* $(X_1,X_2,\ldots,X_n)$ *are comonotonic, i,e., there exists a uniform [0,1] random variable U such that* $X_i=F_{X_i}^{-1}(U)$.

Proof. Since $\phi_*(u)\le\phi(u)$ for any $u\in[0,1]$, Lemma B.2 implies that

$$\sup_{X_i\in\mathcal{P}(\mu_i,\sigma_i^2),\ i=1,2,\ldots,n} M_\phi\left(\sum_{i=1}^n X_i\right)\le\sup_{X_i\in\mathcal{P}(\mu_i,\sigma_i^2),\ i=1,2,\ldots,n} M_{\phi_*}\left(\sum_{i=1}^n X_i\right). \tag{25}$$

On the other hand, the distortion risk measure is sub-additive when the distortion function is convex. Therefore,

$$\sup_{X_i\in\mathcal{P}(\mu_i,\sigma_i^2),\ i=1,2,\ldots,n} M_{\phi_*}\left(\sum_{i=1}^n X_i\right)\le\sum_{i=1}^n\sup_{X_i\in\mathcal{P}(\mu_i,\sigma_i^2)} M_{\phi_*}(X_i), \tag{26}$$

where the equality holds if $(X_1,X_2,\ldots,X_n)$ are comonotonic. The proof is the completed by combining two inequalities (25) and (26). □

The above corollary does not take the dependence constraint among the random variables into account. Generally, the closed-form solutions for the worst-case DRM may be unavailable if constraints on the dependence are additionally proposed. However, in some special case with mild conditions, the closed-form solutions are still available.

Definition 4. For a mean vector $\mu\in\mathbb{R}^n$ and a positive semidefinite matrix $\Sigma\in\mathbb{R}^{n\times n}$, the family of random vectors with mean μ and covariance Σ is defined as the set

$$\mathcal{P}(\mu,\Sigma):=\{X\in\mathbb{R}^n\ \mid\ \mathbb{E}(X)=\mu \text{ and } \mathrm{COV}(X)=\Sigma\},$$

where $\mathrm{COV}(X)=\mathbb{E}(X-\mu)(X-\mu)^\top$ is the covariance.

We consider $\sup_{X\in\mathcal{P}(\mu,\Sigma)} M_\phi(w^\top X)$ and $\inf_{X\in\mathcal{P}(\mu,\Sigma)} M_\phi(w^\top X)$, where $w\in\mathbb{R}^n$ denotes a constant vector $w\in\mathbb{R}^n$ and $X\in\mathbb{R}^n$ denotes a random vector. The following proposition is the multivariate version of the Theorem 2.1.

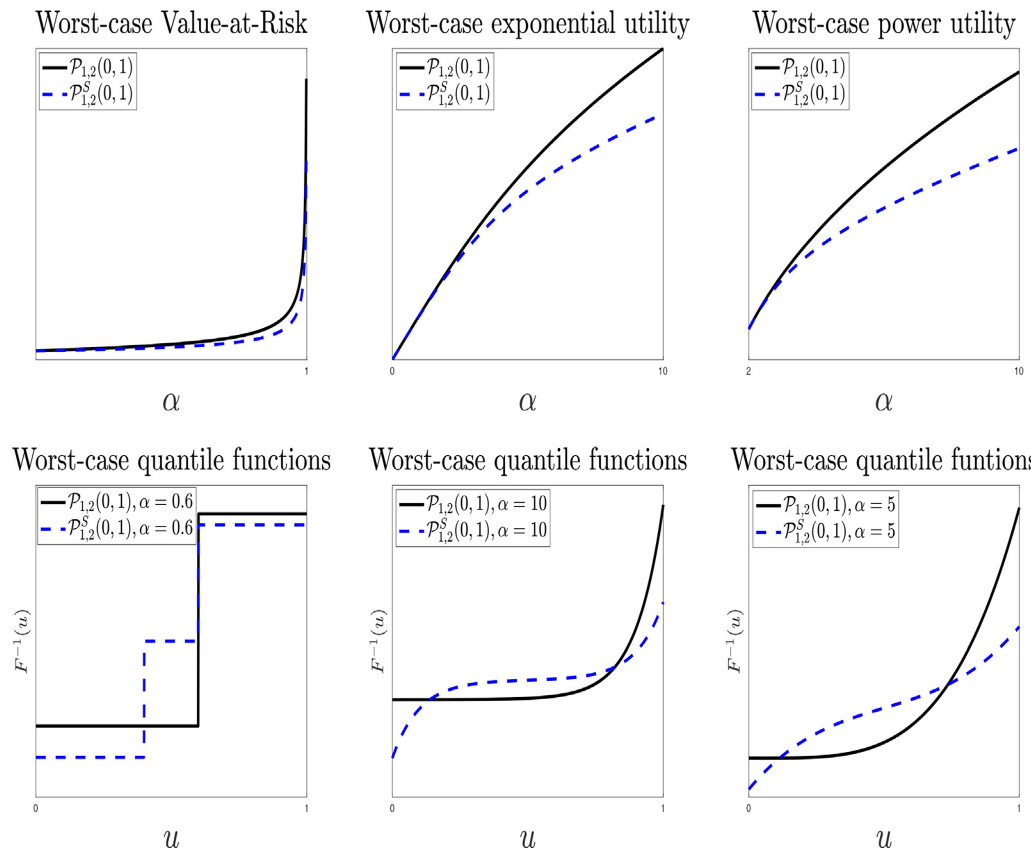

Fig. 3. Worst-case values and quantile functions of three distortion risk measures: value-at-risk (left), exponential utility distortion (middle), and power utility distortion (right). The top panels display the worst-case values for the general and symmetrical cases as a function the parameter in distortion functions (see Table 1); the bottom panels display the worst-case quantile functions for the general and symmetrical cases with preselected parameters.

Table 1
The table presents several special cases of DRMs, including well-known risk measures like VaR and CVaR. The worst and best cases for them are also reported in the last two columns. In particular for VaR_α, CVaR_α, and $\text{RVaR}_{\alpha,\beta}$, the assumption that $\alpha > \frac{1}{2}$ makes sense since confidence level is always greater than 90% from a practical point of view. Of course one can also easily obtains the explicit results for other domains, for example, $0 \le \alpha \le \frac{1}{2}$.

Risk measure	Distortion function	Constraint	Best case	Worst case
VaR_α	$1_{\{u\ge\alpha\}},\ \frac{1}{2} < \alpha < 1$	$\mathcal{P}(\mu,\sigma^2)$	$\mu - \sigma\sqrt{\frac{1-\alpha}{\alpha}}$	$\mu + \sigma\sqrt{\frac{\alpha}{1-\alpha}}$
		$\mathcal{P}_S(\mu,\sigma^2)$	$\mu - \sigma\sqrt{\frac{1-\alpha}{2\alpha^2}}$	$\mu + \sigma\sqrt{\frac{1}{2(1-\alpha)}}$
CVaR_α	$\frac{u-\alpha}{1-\alpha}1_{\{u\ge\alpha\}},\ \frac{1}{2} < \alpha < 1$	$\mathcal{P}(\mu,\sigma^2)$	μ	$\mu + \sigma\sqrt{\frac{\alpha}{1-\alpha}}$
		$\mathcal{P}_S(\mu,\sigma^2)$	μ	$\mu + \sigma\sqrt{\frac{1}{2(1-\alpha)}}$
$\text{RVaR}_{\alpha,\beta}$	$\min\{\frac{u-\alpha}{\beta-\alpha},1\}1_{\{u\ge\alpha\}},\ \frac{1}{2} < \alpha < \beta < 1$	$\mathcal{P}(\mu,\sigma^2)$	$\mu - \sigma\sqrt{\frac{1-\beta}{\beta}}$	$\mu + \sigma\sqrt{\frac{\alpha}{1-\alpha}}$
		$\mathcal{P}_S(\mu,\sigma^2)$	$\mu - \sigma\sqrt{\frac{1-\beta}{2\beta^2}}$	$\mu + \sigma\sqrt{\frac{1}{2(1-\alpha)}}$
Power Utility	$u^\alpha,\ \alpha \ge 1$	$\mathcal{P}(\mu,\sigma^2)$	μ	$\mu + \sigma\frac{\alpha-1}{\sqrt{2\alpha-1}}$
		$\mathcal{P}_S(\mu,\sigma^2)$	μ	$\mu + \frac{\sigma}{2}\sqrt{\frac{2\alpha^2}{2\alpha-1} - \frac{4^{1-\alpha}\alpha\sqrt{\pi}\Gamma(1+\alpha)}{\Gamma(\frac{1}{2}+\alpha)}}$
Exponential Utility	$\frac{e^{\alpha u}-1}{e^\alpha-1},\ \alpha > 0$	$\mathcal{P}(\mu,\sigma^2)$	μ	$\mu + \sigma\sqrt{-1 + \frac{\alpha}{2}\frac{e^\alpha+1}{e^\alpha-1}}$
		$\mathcal{P}_S(\mu,\sigma^2)$	μ	$\mu + \frac{\sigma}{2}\sqrt{\frac{\alpha(e^\alpha-e^{-\alpha}-2\alpha)}{e^\alpha+e^{-\alpha}-2}}$
Wang's Transform	$\Phi(\Phi^{-1}(u) - \Phi(\alpha)),\ \alpha \in \mathbb{R}$	$\mathcal{P}(\mu,\sigma^2)$	μ	$\mu + \sigma\sqrt{e^{\lambda^2}} - 1$
		$\mathcal{P}_S(\mu,\sigma^2)$	μ	No analytical expression

Corollary 2.3 (Multivariate case)**.** *Let $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n\times n}$ be a positive semidefinite matrix, then for any vector $w \in \mathbb{R}^n$, we have*

$$\sup_{X\in\mathcal{P}(\mu,\Sigma)} M_\phi(w^\top X) = \mu^\top w + \sqrt{w^\top \Sigma w}\cdot\|\phi_*'(u) - 1\|_2, \tag{27}$$

$$\inf_{X\in\mathcal{P}(\mu,\Sigma)} M_\phi(w^\top X) = \mu^\top w - \sqrt{w^\top \Sigma w}\cdot\|\phi^{*\prime}(u) - 1\|_2. \tag{28}$$

Moreover, the worst-case distributions can be obtained analytically by Theorem 2.1 and the projection method provided in Appendix C.

The extreme-case values in the proposition follows from Theorem 2.1 by observing the fact that $X \in \mathcal{P}(\mu,\Sigma)$ yields $w^\top X \in \mathcal{P}(\mu^\top w, w^\top\Sigma w)$ and the projection method. It is worthy mentioning that we are not the first one to propose the projection method (Popescu, 2007; Yu et al., 2009). Instead, we propose another more efficient Algorithm in Appendix C in which the closed-form solutions are available.

The closed-form solutions in Corollary 2.3 provide a unified framework for the extreme-case DRMs in multivariate cases. In particular, they generalize the popular worst-case VaR (Ghaoui et al., 2003) and worst-case CVaR (Chen et al., 2011; Natarajan et al., 2010) for multivariate cases. More importantly, the closed-form solutions provide a greater tractability for the robust portfolio optimization problem with any distortion risk measures.

3. Applications and numerical illustrations

The extreme-case DRM is closely related to the robust portfolio optimization problem in the portfolio management theory, which is one of the most popular applications of the worst-case DRM. Therefore, we consider numerical examples with parametric ambiguity based on real data sets to illustrate the application. Specifically, if we denote the losses of n risky assets by a random vector $X = (X_1, X_2, \dots, X_n)$, then the robust portfolio optimization problem seeks a portfolio weight $w = (w_1, w_2, \dots, w_n)$ that minimizes the worst-case DRM, i.e., $w = \arg\min_{w\in\mathcal{W}} \sup_{X\in\mathcal{P}} M_\phi(w^\top X)$, where $\mathcal{W}$ is the set of admissible portfolio weights (e.g., $w \geq 0$ if short selling is not allowed), and $\mathcal{P}$ is the uncertainty set.

However, such a formulation only takes the risk into account. As suggested by Brandtner (2013, 2016), we consider the trade-off between risk and return instead of restricting to a limited risk analysis only. To this end, we reproduce a portfolio selection approach which is well-established in the mean-variance framework, i.e., $(1-\lambda)\mathbb{E}(-X) - \lambda M_\phi(X)$, called the preference function, for $\lambda \in [0,1]$ where $-X$ represents the return. Obviously, such a function represents a utility. However, this paper looks from the loss perspective, thus the opposite preference function is adopted in the following, namely $\pi_\phi(X) = \lambda M_\phi(X) + (1-\lambda)\mathbb{E}(X)$ for $\lambda \in [0,1]$, which will be minimized and may represent dis-utility. Formally, a robust portfolio optimization problem in this paper seeks the portfolio weight w minimizing the worst-case risk $\pi_\phi(w^\top X)$,

$$w = \arg\min_{w\in\mathcal{W}} \sup_{X\in\mathcal{P}} \pi_\phi(w^\top X). \tag{29}$$

Notably, the above π_ϕ still belongs to the family of DRM since $\pi_\phi(w^\top X) = M_\psi(w^\top X)$ with $\psi(u) = \lambda\phi(u) + (1-\lambda)u$ for all $u \in [0,1]$.

3.1. Data, moment uncertainty, and feasible sets

The sample mean and sample covariance of the assets loss are prone to errors due to imperfect sampling process, which indicates that they may not be accurate. To this end, this paper considers the parameter ambiguity under the following box uncertainty:

$$\mathcal{P}(\mu_-, \mu_+; \Sigma_-, \Sigma_+) = \left\{X \in \mathbb{R}^n : \ \mu_- \leq \mathbb{E}(X) \leq \mu_+ \text{ and } \Sigma_- \leq \mathrm{COV}(X) \leq \Sigma_+\right\}, \tag{30}$$

where μ_- and μ_+ are componentwise lower and upper bounds for $\mathbb{E}(X)$ respectively; $\Sigma_- = \{\sigma_{XY-}\}_{XY}$ and $\Sigma_+ = \{\sigma_{XY+}\}_{XY}$ are componentwise lower and upper bounds for $\mathrm{COV}(X)$ respectively. Since the component interval $[\Sigma_-, \Sigma_+]$ may not contain a positive semidefinite matrix, we also assume that there exists a positive semidefinite Σ_0 such that $\Sigma_- \leq \Sigma_0 \leq \Sigma_+$.

To construct the box uncertainty set $\mathcal{P}(\mu_-, \mu_+; \Sigma_-, \Sigma_+)$, we make use of the "standard error" of estimators. Intuitively, the standard error (SE) measures the average distance between the estimator $T(X)$ and true values θ of parameters, thus it is reasonable to assume that $| T(X) - \theta | \leq \mathrm{SE}(T(X))$, or equivalently,

$$T(X) - \mathrm{SE}(T(X)) \leq \theta \leq T(X) + \mathrm{SE}(T(X)), \tag{31}$$

where $\mathrm{SE}(T(X))$ denotes the standard error of $T(X)$.

For the estimator of sample mean $\sum_{i=1}^n X_i/n$, the standard error is $\hat{\sigma}_X/\sqrt{n}$ where $\hat{\sigma}_X$ is the sample standard deviation. For the sake of simplicity, we assume that X and Y are bivariate normal, thus the standard error of sample covariance is $\sqrt{(\hat{\sigma}_X^2\hat{\sigma}_Y^2 + \hat{\sigma}_{XY}^2)/(n-1)}$ where $\hat{\sigma}_{XY}$ is the sample covariance between X and Y. Thus by (31), we obtain the bounds in uncertainty set (30) as

$$\mu_- = \hat{\mu}_X - \frac{\hat{\sigma}_X}{\sqrt{n}}, \qquad \sigma_{XY-} = \hat{\sigma}_{XY} - \sqrt{\frac{\hat{\sigma}_X^2\hat{\sigma}_Y^2 + \hat{\sigma}_{XY}^2}{n-1}}; \tag{32}$$

$$\mu_+ = \hat{\mu}_X + \frac{\hat{\sigma}_X}{\sqrt{n}}, \qquad \sigma_{XY+} = \hat{\sigma}_{XY} + \sqrt{\frac{\hat{\sigma}_X^2\hat{\sigma}_Y^2 + \hat{\sigma}_{XY}^2}{n-1}}, \tag{33}$$

where $\hat{\mu}_X$ ($\hat{\mu}_Y$) and $\hat{\sigma}_{XY}$ are sample mean and sample covariance, respectively.

3.2. GlueVaR

To illustrate the application of DRM, an intermediate of two most commonly used risk measures, VaR and CVaR, will be used in the numerical analysis for the sake of simplicity and flexibility. Specifically, we use a convex combination of VaR and CVaR with different confidence levels as the DRM, i.e.,

$$M_\phi(X) = \omega \mathrm{VaR}_\alpha(X) + (1-\omega)\mathrm{CVaR}_\beta(X), \tag{34}$$

where the weight $\omega \in [0,1]$ and $0 < \alpha \leq \beta < 1$. Such kind of risk measures are also called the GuleVaR in actuarial science, which can incorporate more information about decision makers' risk attitudes and retain the advantages of sub-additivity (Belles-Sampera et al., 2013; 2016). By assuming $\omega \leq \frac{\beta-\alpha}{1-\alpha}$, the distortion function and its corresponding convex envelope are given by

$$\phi(u) = \omega 1_{\{\alpha<u\leq\beta\}} + \frac{(1-\omega)u + \omega - \beta}{1-\beta} 1_{\{\beta<u\leq 1\}},$$

$$\phi_*(u) = \frac{\omega(u-\alpha)}{\beta-\alpha} 1_{\{\alpha<u\leq\beta\}} + \frac{(1-\omega)u + \omega - \beta}{1-\beta} 1_{\{\beta<u\leq 1\}},$$

respectively. Thus it is straightforward to see that $\psi_*(u) = \lambda\phi_*(u) + (1-\lambda)u$, and then the integral term in objective function (27) can be evaluated analytically as

$$\int_0^1 (\psi_*'(u) - 1)^2 \, du = \frac{\lambda^2((\beta-\omega)^2 - \alpha(\beta - \omega(2-\omega)))}{(\beta-\alpha)(1-\beta)}.$$

The distortion function is crucial to the distortion risk measure. GlueVaR is simply a mixture of two popular risk measures VaR and CVaR. It is therefore natural to see that the mixture method is a direct approach to construct new distortion functions. In addition to the mixture method, the composite method is another approach. Specifically, if ϕ_1 and ϕ_2 are two distortion functions, then the composite function $\phi_1 \circ \phi_2$ is also a distortion function by the composite method.

Below are two examples of the mixture and composite methods, respectively.

Example 3.1 (Inter-quantile measure and inter-RVaR-CVaR measure)**.** For $0 < \alpha < \beta < 1$, we let $\phi_1(u) = 1_{\{\alpha<u\leq 1\}}$ and $\phi_2(u) = 1_{\{\beta<u\leq 1\}}$, $u \in [0,1]$. Define a new distortion function $\phi(u) = c\phi_1(u) + (1-c)\phi_2(u)$ for $c \in [0,1]$, then

$$M_\phi(X) = c\mathrm{VaR}_\alpha(X) + (1-c)\mathrm{VaR}_\beta(X).$$

On the other hand, we have $\phi_*(u) = \frac{c(u-\alpha)}{\beta-\alpha} 1_{\{\alpha\leq u<\beta\}} + \frac{(1-c)u+c-\beta}{1-\beta} 1_{\{\beta\leq u\leq 1\}}$ when $0 \leq c \leq \frac{\beta-\alpha}{1-\alpha}$ for all $u \in [0,1]$, then

$$M_{\phi_*}(X) = c\mathrm{RVaR}_{\alpha,\beta}(X) + (1-c)\mathrm{CVaR}_\beta(X).$$

The distortion risk measure M_ϕ is called an inter-quantile measure, and M_{ϕ_*} is called an inter-RVaR-CVaR measure.

Table 2
Bounds in box uncertainty set $\mathcal{P}(\mu_-, \mu_+; \Sigma_-, \Sigma_+)$ based on standard errors of estimators. Note that the data presented are all daily frequency measurements and are expressed as percentages.

(μ_-, μ_+) (%)	AAPL (−0.2245, 0.0644)	ALGN (−0.1506, 0.2262)	SBUX (−0.0515, 0.2320)	EBAY (−0.0693, 0.2112)	M (−0.0752, 0.3473)
$(\hat{\sigma}_{XY-}, \hat{\sigma}_{XY+})$ (%)	AAPL	ALGN	SBUX	EBAY	M
AAPL	(0.0983, 0.1116)	(0.0418,0.0548)	(0.0454, 0.0556)	(0.0506, 0.0610)	(0.0644,0.0795)
ALGN	(0.0418,0.0548)	(0.1672, 0.1898)	(0.0509, 0.0640)	(0.0489, 0.0617)	(0.0746,0.0940)
SBUX	(0.0454, 0.0556)	(0.0509,0.0640)	(0.0947, 0.1074)	(0.0533, 0.0636)	(0.0904, 0.1064)
EBAY	(0.0506, 0.0610)	(0.0489, 0.0617)	(0.0533, 0.0636)	(0.0927, 0.1052)	(0.0680, 0.0829)
M	(0.0644,0.0795)	(0.0746,0.0940)	(0.0904, 0.1064)	(0.0680, 0.0829)	(0.2103, 0.2386)

Example 3.2 (Distorted-CVaR)**.** Let $\phi_1(u) = u^\beta$ and $\phi_2(u) = \frac{u-\alpha}{1-\alpha} 1_{\{\alpha \le u \le 1\}}$. Define a new distortion function $\phi(u) = \phi_1 \circ \phi_2(u)$ for $0 < \alpha < 1$ and $\beta > 0$, $u \in [0, 1]$, then

$$M_\phi(X) = \frac{\beta}{(1-\alpha)^\beta} \int_\alpha^1 \mathrm{VaR}_u(X) \cdot (u-\alpha)^{\beta-1} \mathrm{d}u.$$

In particular, $M_{\phi_*}(X) = M_\phi(X)$ holds for all $X \in \mathcal{P}(\mu, \sigma^2)$ when $\beta \ge 1$. We call M_ϕ as a distorted-CVaR. In particular, the distorted-CVaR reduces to CVaR when $\beta = 1$.

Both of the above two examples can be used as alternatives in the numerical illustrations.

3.3. A short numerical example

For the numerical analysis, we randomly select five companies in the U.S. stock market, which are Apple, Align Technology, Starbucks, Ebay, and Macy's, and the company codes are AAPL, ALGN, SBUX, EBAY, and M, respectively. Then we collect a sample of historical daily close prices of them from Yahoo Finance over a period of 2 years from June 2007 to June 2009 with total 503 daily samples. Based on the sample mean and sample covariance of five stock daily losses (i.e., the negative stock returns), we compute the bounds (32) and (33) of uncertainty set $\mathcal{P}(\mu_-, \mu_+; \Sigma_-, \Sigma_+)$ and report them in Table 2.

It is obvious that the optimization problem in (29) under the box uncertainty can be formulated as

$$\begin{aligned} &\inf_w \sup_{\mu, \Sigma} && \mu^\top w + \sqrt{w^\top \Sigma w} \cdot \|\psi_*'(u) - 1\|_2, \\ &\text{s.t.} && \mu_- \le \mu \le \mu_+, \ \Sigma_- \le \Sigma \le \Sigma_+, \ \Sigma \succeq 0, \ w \in \mathcal{W}, \end{aligned} \tag{35}$$

in which the inner "sup" part can be regarded as a semidefinite programming problem (SDP).

With sample mean and sample covariance, we can compute a "nominal portfolio" by setting $\mu_- = \mu_+ = \hat{\mu}$ and $\Sigma_- = \Sigma_+ = \hat{\Sigma}$, against which we can compute a "robust portfolio" under box uncertainty in Table 2 (see, e.g., Grant & Boyd, 2014 for the disciplined convex programming toolbox). In particular, the admissible set $\mathcal{W} = \{w : w \ge 0\}$ can be proposed if short selling is not allowed.

Table 3 displays the optimal portfolio weights for both nominal and robust settings, with and without short selling, respectively. The study reveals several noteworthy findings. Firstly, the decision maker will assume a short position in stock "M" if short selling is permitted, and will avoid taking any positions in this stock if short selling is prohibited. This behavior may be attributed to the fact that the objective function incorporates the return aspect while stock "M" has the lowest return in this portfolio. Secondly, both nominal and robust portfolio strategies assign the highest weight to the stock with the lowest risk, namely stock "AAPL," and the lowest weight to the stock with the highest risk, namely stock "M." Thirdly, the differences between the results of the robust and nominal portfolio strategies are relatively small, which may be explained by the use of daily returns in this study. The use of daily return leads to a relatively limited box uncertainty set $\mathcal{P}(\mu_-, \mu_+; \Sigma_-, \Sigma_+)$ since μ_- is close to μ_+ and Σ_- is close to Σ_+ as evidenced in Table 2.

Table 3
Optimal portfolio weights based on model (29) with $\omega = 0.7$, $\alpha = 0.95$, $\beta = 0.99$, and $\lambda = 0.7$. Types "I" and "II" denote portfolios with short selling allowed and not allowed, respectively.

Optimal weight	AAPL w_1	ALGN w_2	SBUX w_3	EBAY w_4	M w_5
Nominal porfolio (I)	0.4205	0.1385	0.3373	0.2476	−0.1438
Nominal portfolio (II)	0.4695	0.1226	0.2071	0.2008	0.0000
Robust portfolio (I)	0.4269	0.1257	0.3319	0.2495	−0.1340
Robust portfolio (II)	0.4727	0.1102	0.2110	0.2061	0.0000

4. Conclusions

In this paper, we have derived the closed-form solutions for the extreme DRMs (both worst-case and best-case) based on only the first two moments and symmetry of the underlying distributions. In addition, we show that the corresponding extreme-case distributions can be characterized by the envelopes of the distortion function. These results are important from the perspectives of both making theoretical contributions to the risk management literature and promoting the applications based on extreme-case risk measures. In the first aspect, our results generalize several well-known extreme-case risk measures with closed-form solutions by pushing the envelope. In the second aspect, with the closed-form solutions to the extreme-case DRM, we demonstrate that the robust optimization procedure based on the extreme-case DRMs developed in this paper can generate the optimal solutions for situations where only the imperfect estimations for the first two moments are available. Exploring more real-world applications of using closed-form solutions for the extreme-case DRMs can be a future research topic.

Acknowledgment

Dr. Hui Shao greatly acknowledges the support from the Yangtze River Delta Science and Technology Innovation Community Joint Research Program (No. 2022CSJGG0800). Dr. Zhe George Zhang received support from the Natural Sciences and Engineering Research Council of Canada [Grant RGPIN-2019-06364].

Appendics

In the section of appendices, we present the detailed proofs of the analytical results of this paper. We will take three steps to prove the main theorems. In the first step, we will prove that

Theorems 2.1 and 2.2 hold for convex distortion functions, which is just the Proposition 2.1 presented in the paper. In the second step, we will prove that these results hold for piecewise constant distortion functions, which is Proposition B.1 established in this section. Finally, in step 3, we will prove that these results hold for general distortion functions by approximation methods. The entire proof process is successive in nature and is divided into two appendices. The first and third steps are presented in the main text of the paper (see Propositions 2.1 and two theorems), and the second step is reported in Appendix B (see Proposition B.1).

Appendix A. Proof of Proposition 2.3

Proof. Note that the distortion function $\phi(u)$ associated with the symmetrical quantile average $M_\phi(X) = (F^{-1}(\alpha) + F^{-1}(1-\alpha))/2$ takes the following form:

$$\phi(u) = \begin{cases} 0, & \text{if } 0 \le u < 1-\alpha, \\ \frac{1}{2}, & \text{if } 1-\alpha \le u < \alpha, \\ 1, & \text{if } \alpha \le u \le 1. \end{cases}$$

Then the convex and concave envelopes for the distortion function $\phi(u)$ when $\frac{2}{3} \le \alpha < 1$ are

$$\phi_*(u) = \begin{cases} 0, & \text{if } 0 \le u < 1-\alpha, \\ \frac{u+\alpha-1}{2(2\alpha-1)}, & \text{if } 1-\alpha \le u < \alpha, \\ \frac{u-2\alpha+1}{2(1-\alpha)}, & \text{if } \alpha \le u \le 1. \end{cases} \quad \text{and}$$

$$\phi^*(u) = \begin{cases} \frac{u}{2(1-\alpha)}, & \text{if } 0 \le u < 1-\alpha, \\ \frac{u+3\alpha-2}{2(2\alpha-1)}, & \text{if } 1-\alpha \le u < \alpha, \\ 1, & \text{if } \alpha \le u \le 1, \end{cases}$$

respectively; when $\frac{1}{2} \le \alpha < \frac{2}{3}$, they are

$$\phi_*(u) = \begin{cases} 0, & \text{if } 0 \le u < 1-\alpha, \\ \frac{u+\alpha-1}{\alpha}, & \text{if } 1-\alpha \le u \le 1, \end{cases} \quad \text{and}$$

$$\phi^*(u) = \begin{cases} \frac{u}{\alpha}, & \text{if } 0 \le u < \alpha, \\ 1, & \text{if } \alpha \le u \le 1, \end{cases}$$

respectively.

By closed-form solutions in (10) and (11) we obtain that

$$\mu - \sigma\sqrt{\frac{\alpha}{4(2\alpha-1)(1-\alpha)} - 1} \le M_\phi(X)$$
$$\le \mu + \sigma\sqrt{\frac{\alpha}{4(2\alpha-1)(1-\alpha)} - 1}, \qquad \frac{2}{3} \le \alpha < 1,$$
$$\mu - \sigma\sqrt{\frac{1-\alpha}{\alpha}} \le M_\phi(X) \le \mu + \sigma\sqrt{\frac{1-\alpha}{\alpha}}, \qquad \frac{1}{2} \le \alpha < \frac{2}{3}.$$

Therefore, we obtain that

$$\left|\frac{F^{-1}(\alpha) + F^{-1}(1-\alpha)}{2} - \mu\right| \le \sigma B(\alpha), \tag{A.1}$$

where

$$B(\alpha) = \begin{cases} \sqrt{\frac{\alpha}{4(2\alpha-1)(1-\alpha)} - 1}, & \frac{2}{3} \le \alpha < 1, \\ \sqrt{\frac{1-\alpha}{\alpha}}, & \frac{1}{2} \le \alpha < \frac{2}{3}. \end{cases}$$

By letting $\alpha = \frac{1}{2}$, the inequality (A.1) implies that $\left|F^{-1}(\frac{1}{2}) - \mu\right| \le \sigma$, which completes the proof. □

Appendix B. Preliminaries and extensions

To prove that the main results, we first present some preliminaries with three lemmas.

Algorithm 1 The projection method.

INPUT $\mu \in \mathbb{R}^{n\times 1}$, $\omega \in \mathbb{R}^{n\times 1}$, $\Sigma \in \mathbb{R}^{n\times n}$, $\phi(u)$

1) $\Sigma_1 = \omega^\top \Sigma \omega$;
2) Substitute μ_1 and Σ_1 into Corollary 2.3 to generate the worst-case random variable X^*;
3) Use Gram–Schmidt othogonalization method to find $n-1$ non-correlated $\{\delta_1, \delta_2, \ldots, \delta_{n-1}\}$ that are also non-correlated with X^*. Moreover, let $\mathbb{E}(\delta_i) = 0$ and $\mathrm{Var}(\delta_i) = 1$;
4) By orthogonal diagonalization, $\Sigma_1 = P \cdot \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot P^\top$, where P is an orthogonal matrix;
5) $A_0 = P \cdot \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$;
6) Find $n-1$ elements from $\{e_1, e_2, \ldots, e_n\}$, e.g., $\{e_2, e_3, \ldots, e_n\}$, such that $\{A_1, e_2, e_3, \ldots, e_n\}$ is linearly independent, where $A_1 = ((1, 1, \ldots, 1) \cdot A_0)^\top$.
7) $\{A_1, e_2, e_3, \ldots, e_n\} \to \{A_1, \tilde{e_2}, \tilde{e_3}, \ldots, \tilde{e_n}\}$ by Gram–Schmidt orthogonalization.
8) $A = A_0 \cdot (A1, \tilde{e_2}, \tilde{e_3}, \ldots, \tilde{e_n})$.

OUTPUT $(X_1, X_2, \ldots, X_n) = (X^*, \delta_1, \delta_2, \ldots, \delta_{n-1}) \cdot A^\top \cdot \mathrm{diag}(\frac{1}{\omega_1}, \frac{1}{\omega_2}, \ldots, \frac{1}{\omega_n}) + \mu^\top$

B1. Some preliminaries

Lemma B.1 (Dhaene et al. (2012)). *When the distortion function ϕ is right-continuous or left-continuous, the distortion risk measure $M_\phi(X)$ has the following Lebesgue–Stieltjes integral representation:*

$$M_\phi(X) = \int_0^1 F_X^{-1}(u)\mathrm{d}\phi(u), \quad \text{when } \phi \text{ is right-continuous;} \tag{B.1}$$

$$M_\phi(X) = \int_0^1 F_X^{-1+}(u)\mathrm{d}\phi(u), \quad \text{when } \phi \text{ is left-continuous.} \tag{B.2}$$

In particular, $\int_0^1 F_X^{-1}(u)\mathrm{d}\phi(u) = \int_0^1 F_X^{-1+}(u)\mathrm{d}\phi(u)$ when ϕ is continuous.

Lemma B.2. *Let f and g be two functions defined on [0,1] with $f(0) = g(0)$ and $f(1) = g(1)$. If $g(u) \ge f(u)$ for $u \in [0,1]$, then for any random variable whose quantile function $q(u)$ satisfies $\int_0^1 q(u)\,\mathrm{d}f(u) < \infty$ or $\int_0^1 q(u)\,\mathrm{d}g(u) > -\infty$, we have $\int_0^1 q(u)\,\mathrm{d}g(u) \le \int_0^1 q(u)\,\mathrm{d}f(u)$.*

Proof. It is straightforward to see that

$$\int_0^1 q(u)\,\mathrm{d}g(u) - \int_0^1 q(u)\,\mathrm{d}f(u) = \int_0^1 q(u)\,\mathrm{d}(g(u) - f(u))$$
$$= -\int_0^1 (g(u) - f(u))\,\mathrm{d}q(u) \le 0,$$

where the last inequality holds since $q(u)$ is increasing. □

Next lemma presents some properties of convex and concave envelopes, which will be used in the proof of Proposition B.1, Theorems 2.1, and 2.2. More specifically, (*iii*), (*iv*), and (*vi*) will be used in the proof of Proposition B.1, as they describe the forms of the convex or concave envelopes of piecewise constant functions. (*i*) and (*ii*) together with Propositions 2.1 and B.1 will be used in the proof of Theorems 2.1 and 2.2.

Lemma B.3. *For convex and concave envelopes, we have following conclusions.*

(i) *For f and g defined on $[a,b]$ with $f \le g$, we have $f_* \le g_*$ and $f^* \le g^*$.*
(ii) *For f and g defined on $[a,b]$, if $\|f - g\|_\infty < \epsilon$ for some $\epsilon > 0$, we have $\|f_* - g_*\|_\infty < \epsilon$ and $\|f^* - g^*\|_\infty < \epsilon$.*

(iii) *If f defined on $[a, b]$ is an increasing function, then both f_* and f^* are increasing functions.*
(iv) *For f defined on $[a, b]$, we have $f_*(a) = f^*(a) = f(a)$ and $f_*(b) = f^*(b) = f(b)$.*
(v) *Let f and g be two functions defined on $[a, b]$ and $[b, c]$ respectively. Define*

$$h(x) = f(x)1_{\{a\le x<b\}} + g(x)1_{\{b\le x\le c\}},$$
$$\widetilde{h}(x) = f_*(x)1_{\{a\le x<b\}} + g_*(x)1_{\{b\le x\le c\}},$$

where 1_A denotes the indicator function of set A. Then $\widetilde{h}_*(x) = h_*(x)$ for all $x \in [a, b]$.
(vi) *Let f be a piecewise constant function defined on $[a, b]$, then both f_* and f^* are piecewise linear continuous on the interior of the domain. Moreover, all salient points of f_* and f^* are in the set of the discontinuous points of f. (For piecewise linear continuous function g defined on $[a, b]$, $x_0 \in (a, b)$ is called a salient point of g if $g'_-(x_0) \neq g'_+(x_0)$)*

Proof. We only prove the case for convex envelopes, since the proof for concave envelopes is similar. In particular, we will prove (i), (iii), (iv), and (v) by contradiction and prove (vi) by induction.

(i) Suppose there exists $x_0 \in [a, b]$ such that $f_*(x_0) > g_*(x_0)$. Define a new function $\widetilde{g}$ on $[a, b]$ as

$$\widetilde{g}(x) = \max\{g_*(x), f_*(x)\}. \tag{B.3}$$

Note that the maximum of any two convex functions is still convex, thus $\widetilde{g}$ is convex on $[a, b]$.

By (B.3), we have following three direct conclusions:

$$\widetilde{g}(x) \le \max\{g(x), f(x)\} = g(x), \tag{B.4}$$

$$\widetilde{g}(x_0) = \max\{g_*(x_0), f_*(x_0)\} > g_*(x_0), \tag{B.5}$$

$$\widetilde{g}(x) \ge g_*(x); \tag{B.6}$$

The Eq. (B.4) implies that the function $\widetilde{g}$ is a convex function that is dominated by g. On the other hand, Eqs. (B.5) and (B.6) imply that the convex function $\widetilde{g}$ is strictly greater than the envelope of g. Thus by definition of the convex envelope, a contradiction happens. The proof for the convex envelope is then completed.

For two functions f and g satisfy $f(x) \le g(x)$, which is equivalent to that $-g(x) \le -f(x)$, we have $(-g)_*(x) \le (-f)_*(x)$ by the above results (i), which further implies that $-g^*(x) \le -f^*(x)$, hence $f^*(x) \le g^*(x)$ for all $x \in [a, b]$, which completes the proof.

(ii) Note that $\|f - g\|_\infty < \epsilon$ implies that $f(x) - \epsilon < g(x) < f(x) + \epsilon$ for all $x \in [a, b]$. Applying result (i) then yields

$$f_*(x) - \epsilon = (f - \epsilon)_*(x) \le g_*(x) \le (f + \epsilon)_*(x) = f_*(x) + \epsilon$$

for all $x \in [a, b]$, which follows that $\|f_* - g_*\|_\infty < \epsilon$.

(iii) Suppose that f_* is not increasing, then there exist $x_0, x_1 \in [a, b]$ with $x_0 < x_1$ such that $f_*(x_0) > f_*(x_1)$. Noting that f_* is convex, we obtain that

$$\frac{f_*(a) - f_*(x_1)}{a - x_1} \le \frac{f_*(x_0) - f_*(x_1)}{x_0 - x_1} < 0.$$

Thus $f_*(a) > f_*(x_1)$ since $a - x_1 < 0$. Define a new function on $[a, b]$ as $\widetilde{f}(x) = \max\{f_*(x), f_*(a)\}$, which is a convex function since the maximum of any two convex functions is also convex. Now it is straightforward to see that for all $x \in [a, b]$,

$$\widetilde{f}(x) \le \max\{f(x), f(a)\} \le f(x), \tag{B.7}$$

$$\tilde{f}(x_1) = f_*(a) > f_*(x_1), \tag{B.8}$$

$$\tilde{f}(x) \ge f_*(x). \tag{B.9}$$

Similar to the proof for the result (i), Eq. (B.7) implies that $\widetilde{f}$ is a convex function dominated by f. Equations (B.8) and (B.9) imply $\widetilde{f}$ is a convex function strictly greater than f_*, which contradicts to the definition of convex envelope. The proof is competed.

(iv) Suppose that $f_*(b) \neq f(b)$. Then by definition of the convex envelope, $f_*(b) < f(b)$ must hold. Define a function $\widetilde{f}$ on $[a, b]$ as

$$\widetilde{f}(x) = f_*(x)1_{\{a\le x<b\}} + f(b)1_{\{x=b\}},$$

which is a convex function since f_* is convex and $f_*(b) < f(b)$. Now it is straightforward to see that

$$\widetilde{f}(x) \le f(x), \quad \widetilde{f}(x) \ge f_*(x), \quad \text{and} \quad \widetilde{f}(b) > f_*(b).$$

which leads to a contradiction since f_* is the convex envelope of f.

Similarly, we can prove that $f_*(a) = f(a)$. Therefore, the proof is terminated here.

(v) By definition of the convex envelope, we have $\widetilde{h}(x) \le h(x)$ for all $x \in [a, b]$. Applying the result in (i) gives

$$\widetilde{h}_*(x) \le h_*(x), \ \text{for all } x \in [a, b]. \tag{B.10}$$

Next we will prove the reverse direction of it by contradiction, i.e., we prove that

$$h_*(x) \le \widetilde{h}(x), \ \text{for all } x \in [a, b]. \tag{B.11}$$

Suppose that there exists $x_0 \in [a, b]$ such that $h_*(x_0) > \widetilde{h}(x_0)$. Without losing generality, we can assume that $x_0 \in [b, c]$. Define a new function $\hat{h}(x) = \max\{h_*(x), \widetilde{h}(x)\}$ and also define a convex function on $[b, c]$ by $\hat{g}(x) = \hat{h}(x)$. Therefore, on the domain of function $\hat{g}$ we have

$$\begin{aligned}\hat{g}(x) \le g(x) \quad \text{and} \quad \hat{g}(x) = \max\{h_*(x), \widetilde{h}(x)\}\\ \ge \widetilde{h}(x) = g_*(x), \ \text{for all } x \in [b, c].\end{aligned} \tag{B.12}$$

On the other hand, we note that $\hat{g}(x_0) = \max\{h_*(x_0), \widetilde{h}(x_0)\} > \widetilde{h}(x_0) = g_*(x_0)$, hence $\hat{g}(x) \neq g_*(x)$ for some $x \in [a, b]$. Combining with (B.12), a contradiction occurs since g_* is the convex envelope of g.

At last we combine Eqs. (B.10) and (B.11), thus $\widetilde{h}_*(x) = h_*(x)$ holds for all $x \in [a, b]$, which completes the proof.

(vi) Let f_n be a piecewise constant function defined on $[a, b]$, where $n - 1$ is the number of discontinuity points in the interior of the domain, then f_n must take the following form:

$$f_n(x) = \sum_{i=1}^{n} y_i 1_{\{x_{i-1}<x<x_i\}} + \sum_{i=0}^{n} z_i 1_{\{x=x_i\}}, \tag{B.13}$$

where $a = x_0 < x_1 < \ldots < x_n = b$ and $z_i \in \{y_i, y_{i+1}\}$ for $1 \le i \le n - 1$.

We will prove the result by induction on n. When $n = 1$, it is easy to check that the conclusion holds. Now suppose that the conclusion holds for all $n \le m - 1$, where $m \ge 2$ is a positive integer. Now consider the case $n = m$. Define a new function $\hat{f}(x)$ on $[x_m, x_{m+1}]$ as

$$\hat{f}(x) = z_{m-1}1_{\{x=x_{m-1}\}} + y_m 1_{\{x_{m-1}<x<x_m\}} + z_m 1_{\{x=x_m\}}, \ \ x \in [x_m, x_{m+1}]. \tag{B.14}$$

Hence by Eqs. (B.13) and (B.14), we can rewrite the function as

$$f_m(x) = f_{m-1}(x)1_{\{x_0\le x<x_{m-1}\}} + \hat{f}(x)1_{\{x_{m-1}\le x\le x_m\}},$$

where $f_{m-1}(x) = \sum_{i=1}^{m-1} y_i 1_{\{x_{i-1}<x<x_i\}} + \sum_{i=0}^{m-1} z_i 1_{\{x=x_i\}}$ is a piecewise constant increasing function with $m - 2$ discontinuity points on the domain (x_0, x_{m-1}).

For the function $\hat{f}$ in (B.14), there are four different cases. Specifically, they are (1) $z_{m-1} \le y_m \le z_m$, (2) $z_m \le y_m \le z_{m-1}$, (3) $\max\{z_{m-1}, z_m\} \le y_m$, and (4) $\min\{z_{m-1}, z_m\} \ge y_m$. We only consider the first case since the other three cases are similar. When $z_{m-1} \le$

$y_m \le z_m$, it is direct to calculate the convex envelope of the function $\hat{f}$ as follows:

$$\hat{f}_*(x) = \left(z_{m-1} + \frac{y_m - z_{m-1}}{x_m - x_{m-1}}(x - x_{m-1})\right)1_{\{x_{m-1} \le x < x_m\}} + z_m 1_{\{x = x_m\}}\ x \in [x_{m-1}, x_m]. \tag{B.15}$$

With the convex envelope function (B.15), we define a function $\widetilde{f}_m(x)$ as

$$\widetilde{f}_m(x) = (f_{m-1})_*(x)1_{\{x_0 \le x < x_m\}} + \hat{f}_*(x)1_{\{x_m \le x \le x_{m+1}\}},\ \ x \in [x_0, x_{m+1}]. \tag{B.16}$$

By the inductive assumption, $(f_{m-1})_*$ is a piecewise linear convex function and its salient points are contained in the set $A_{m-1} = \{x_1, \ldots, x_{m-2}\}$. Without losing generality, we assume that the set of salient points is just A_{m-1}. Therefore by letting

$$k = \sup\left\{ i \le m-1 \;\middle|\; \frac{y_m - (f_{m-1})_*(x_i)}{x_m - x_i} \ge \frac{(f_{m-1})_*(x_i) - (f_{m-1})_*(x_{i-1})}{x_i - x_{i-1}} \right\},$$

and by the definition (B.16), we have

$$(\widetilde{f}_m)_*(x) = (f_{m-1})_*(x)1_{\{x_0 \le x < x_k\}} + (f_{m-1})_*(x_k) + \frac{y_m - (f_{m-1})_*(u_k)}{x_m - x_k}(x - x_k)1_{\{x_k \le x < x_m\}} + z_m 1_{\{x = x_m\}}.$$

Therefore, $(\widetilde{f}_m)_*$ is piecewise linear on the interior of its domain with salient points contained in a subset of $A_m = A_{m-1} \cup \{x_{m-1}\}$, where all the elements are the discontinuity points of the function f_m. Finally by the result (v), we obtain that $(f_m)_*(x) = (\widetilde{f}_m)_*(x)$ for $x \in [a, b]$, which completes the proof. □

B2. When the distortion function is piecewise constant

Proposition B.1. *Theorems 2.1 and 2.2 hold when the distortion function ϕ is piecewise constant in $\mathcal{D}_-$ or $\mathcal{D}_+$.*

Proof. We will only prove the result for the upper bound, since the lower bound can be directly obtained by the result of the upper bound: $\inf_X M_\phi(X) = -\sup_X M_\psi(-X)$, where $\psi(u) = 1 - \phi(1-u)$ is also a distortion function.

Let distortion function $\phi \in \mathcal{D}_-$ and the distortion function ϕ be piecewise constant on (0,1), then ϕ can be expressed as

$$\phi(u) = \sum_{i=1}^{n} y_i 1_{\{u_{i-1} < u \le u_i\}},\ \text{for all } u \in [0,1]. \tag{B.17}$$

where $0{=}u_0 < u_1 < \ldots < u_n = 1$ and $0 = y_1 < y_2 < \ldots < y_n = 1$.

(1) We will first prove the case $X \in \mathcal{P}(\mu, \sigma^2)$. By results (iii), (iv) and (vi) in Lemma B.3, we know that the convex envelope function $\phi_*(u)$ is piecewise linear, increasing, and continuous on (0,1) with $\phi_*(0) = 0$ and $\phi_*(1) = 1$. Denote the salient points of ϕ_* by $u_{i_1} < u_{i_2} < \ldots < u_{i_{j-1}}$, and let $u_{i_0} = 0, u_{i_j} = 1$, then by (B.17), we have

$$\phi_*(u) = \sum_{m=1}^{j-1}\left(k_m(u - u_{i_m}) + y_{i_m}\right)1_{\{u_{i_m} \le u \le u_{i_{m+1}}\}},\ \text{for all } u \in [0,1], \tag{B.18}$$

where $k_m = (y_{i_{m+1}} - y_{i_m})/(u_{i_{m+1}} - u_{i_m})$ for $1 \le m \le j-1$. Now we construct a distortion function $\hat{\phi}(u)$ by

$$\hat{\phi}(u) = \sum_{m=1}^{j} y_{i_m} 1_{\{u_{i_{m-1}} < u \le u_{i_m}\}},\ \text{for all } u \in [0,1]. \tag{B.19}$$

Then it is straightforward to see that $\phi_*(u) \le \phi(u) \le \hat{\phi}(u)$ for all $u \in [0,1]$. Applying Lemma B.2 then yields

$$\sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_{\hat{\phi}}(X) \le \sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_\phi(X) \le \sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_{\phi_*}(X). \tag{B.20}$$

Recalling the result when the distortion function is convex (Proposition 2.1), we have

$$\sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_{\phi_*}(X) = \mu + \sigma\sqrt{\sum_{m=1}^{j} k_m^2 (u_{i_m} - u_{i_{m-1}}) - 1}, \tag{B.21}$$

and the equality holds if and only if the worst-case distribution $F(u)$ satisfies

$$F^{-1}(u) = \mu + \sigma\frac{k_m - 1}{\sqrt{\sum_{m=1}^{j} k_m^2 (u_{i_m} - u_{i_{m-1}}) - 1}}\ a.e.,$$

for all $u_{i_{m-1}} \le u < u_{i_m}$.

By (B.19), (B.20), and (B.21), we obtain that

$$\sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_\phi(X) \ge \sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_{\hat{\phi}}(X) \ge M_{\hat{\phi}}(X_*) = \mu + \sigma\sqrt{\sum_{m=1}^{j} k_m^2 (u_{i_m} - u_{i_{m-1}}) - 1} = \sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_{\phi_*}(X). \tag{B.22}$$

Combing (B.22) and (B.20) yields $\sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_\phi(X) = \sup_{X \in \mathcal{P}(\mu,\sigma^2)} M_{\phi_*}(X)$, which completes the proof.

(2) Define $\bar{\phi}(u) = (\phi(u) + \phi(1-u))/2$. It is easy to see that $\bar{\phi}$ is piecewise constant on (0,1) with $\bar{\phi}(0) = \bar{\phi}(1) = 1$. Hence $\bar{\phi}$ can be expressed by the following form:

$$\bar{\phi}(u) = \sum_{i=1}^{s} y_i' 1_{\{v_{i-1} < u < v_i\}} + \sum_{i=0}^{s} z_i' 1_{\{u = v_i\}},\ \text{for all } u \in [0,1], \tag{B.23}$$

where $0{=}v_0 < v_1 < \ldots < v_s = 1$, $z_0' = z_s' = 1$, and $z_i' \in \{y_i', y_{i+1}'\}$ for $1 \le i \le s-1$. Again by (iii), (iv) and (vi) in Lemma B.3 and Eq. (B.23), we know that $\bar{\phi}_*$ is piecewise linear continuous on (0,1) with $\bar{\phi}_*(0) = 1$ and $\bar{\phi}_*(1) = 1$.

Let $v_{j_0} = 0$ and $v_{j_l} = 1$. Denote the salient points of $\bar{\phi}_*$ by $v_{j_1} < v_{j_2} < \ldots < v_{j_{l-1}}$, then by direct calculation we have

$$\bar{\phi}_*(u) = \sum_{m=1}^{l-1} k_m(u - v_{j_m})1_{\{v_{j_{m-1}} \le u < v_{j_m}\}} + k_l(u - v_{j_l})1_{\{v_{j_{l-1}} \le u \le v_{j_l}\}}, \tag{B.24}$$

where $k_m = (y_{j_m}' - y_{j_{m-1}}')/(v_{j_m} - v_{j_{m-1}})$ for $1 \le m \le l$ and $y_{j_0}' = 1$.

Since $\bar{\phi} \ge \bar{\phi}_*$, $\bar{\phi}(0) = \bar{\phi}_*(0)$, and $\bar{\phi}(1) = \bar{\phi}_*(1)$, by equation (B.24), Proposition 2.2, and Lemma B.2, we obtain that

$$M_\phi(X) \le \mu + \frac{1}{2}\int_0^1 q(u)\,d\bar{\phi}_*(u) = \mu + \frac{1}{2}\int_0^1 (q(u) - \mu)\bar{\phi}_*'(u)\,du \le \mu + \frac{\sigma}{2}\|\bar{\phi}_*'(u)\|_2 = \mu + \frac{\sigma}{2}\sqrt{\sum_{m=1}^{l} k_m^2 (v_{j_m} - v_{j_{m-1}})}. \tag{B.25}$$

By Proposition 2.2, the equality in the second inequality in (B.25) holds if and only if the worst-case distribution F_{X^*}

satisfies

$$F_{X^*}^{-1}(u) = \mu + \sigma \frac{\bar{\phi}'_*(u)}{\|\bar{\phi}'_*(u)\|_2}$$
$$= \mu + \sigma \frac{k_m}{\sqrt{\sum_{m=1}^{l} k_m^2 (v_{j_m} - v_{j_{m-1}})}}, \quad \text{for } v_{j_{m-1}} \le u < v_{j_m}. \tag{B.26}$$

Thus by Eq. (B.25), we have

$$\sup_{X\in\mathcal{P}_S(\mu,\sigma^2)} M_\phi(X) \le \mu + \frac{\sigma}{2}\|\bar{\phi}'_*(u)\|_2 = \mu + \frac{\sigma}{2}\sqrt{\sum_{m=1}^{l} k_m^2 (v_{j_m} - v_{j_{m-1}})}. \tag{B.27}$$

On the other hand,

$$\sup_{X\in\mathcal{P}_S(\mu,\sigma^2)} M_\phi(X) \ge M_\phi(X_*) = \mu + \frac{1}{2}\int_0^1 (F_{X_*}^{-1}(u) - \mu)\,\mathrm{d}\bar{\phi}(u)$$
$$= \mu + \frac{\sigma}{2}\sqrt{\sum_{m=1}^{l} k_m^2 (v_{j_m} - v_{j_{m-1}})}. \tag{B.28}$$

Summarizing (B.27) and (B.28), we obtain that $\sup_{X\in\mathcal{P}_S(\mu,\sigma^2)} M_\phi(X) = \mu + \frac{\sigma}{2}\|\bar{\phi}'_*(u)\|_2$, which completes the proof. □

B3. Another lemma

To prove Theorems 2.1 and 2.2 by using the piecewise constant function to approximate the given function, we need to establish another lemma. More specifically, this lemma will be used to give an estimation for the corresponding convex envelopes of piecewise constant function via the convex envelope of the given function.

Lemma B.4. *Let f and g be two increasing, convex, and continuous functions defined on $[a,b]$. Suppose that g is piecewise linear with $g(x) \ge f(x)$ for $x \in [a,b]$ and $g(b) = f(b)$, then*

$$\int_a^b (g'(x))^2\,\mathrm{d}x \le \int_a^b (f'(x))^2\,\mathrm{d}x.$$

Proof. Let $C^1([a,b]) :=\{h \mid h$ is differentiable almost everywhere on $(a,b)\}$, and define a functional l on $C^1([a,b])$ by

$$l(h)(x) = \frac{h(b)-h(a)}{b-a}(x-a) + h(a), \;\; h \in C^1[a,b].$$

It is obvious that $l(h)$ is a linear function on $[a,b]$ with $l(h)(a) = h(a)$ and $l(h)(b) = h(b)$ for any $h \in C^1([a,b])$. We begin by proving that

$$\int_a^b (h'(x))^2\,\mathrm{d}x \ge \int_a^b (l(h)'(x))^2\,\mathrm{d}x, \;\; h \in C^1([a,b]). \tag{B.29}$$

To this end, we first define function $\hat{h}(x)$ on [0,1] as $\hat{h}(x) = \frac{h((b-a)x+a)-h(a)}{h(b)-h(a)}$, then arrangement gives

$$h(x) = (h(b)-h(a))\hat{h}\Big(\frac{x-a}{b-a}\Big) + h(a). \tag{B.30}$$

Differentiating on both sides of Eq. (B.30) yields

$$\int_a^b \big(h'(x)\big)^2\,\mathrm{d}x = \frac{(h(b)-h(a))^2}{b-a}\int_0^1 \big(\hat{h}'(x)\big)^2\,\mathrm{d}x$$
$$\ge \frac{(h(b)-h(a))^2}{b-a}\Big(\int_0^1 \hat{h}'(x)\,\mathrm{d}x\Big)^2 = \int_a^b \big(l(h)'(x)\big)^2\,\mathrm{d}x,$$

where the first inequality holds due to the Hölder's inequality.

Next we prove the conclusion by induction. Denote n the number of salient points of the function $g(x)$. First we note that the conclusion holds for $n = 0$ ($x = a$ and $x = b$). Indeed, $g(x)$ is linear in this case and by (B.29) we have

$$\int_a^b \big(f'(x)\big)^2\,\mathrm{d}x \ge \frac{(f(b)-f(a))^2}{b-a} \ge \frac{(g(b)-g(a))^2}{b-a} = \int_a^b \big(g'(x)\big)^2\,\mathrm{d}x,$$

where the first inequality holds due to the Hölder's inequality.

Suppose the conclusion is true for $n \le m$ where m is a positive integer, and we then proceed to prove the case $n = m+1$. Let x_0 be the first salient point of g on (a,b) and denote by

$$x_1 = \inf\Big\{x_0 \le x \le b \mid f(x) = \frac{g(x_0)-g(a)}{x_0-a}(x-a) + g(a)\Big\}.$$

Note that x_1 always exists by the conditions restricted on f and g. Denote by

$$f_{x_0}(x) = \begin{cases} \frac{g(x_0)-g(a)}{x_0-a}(x-a) + g(a), & x \in [a,x_0), \\ f(x), & x \in [x_0,b]. \end{cases} \tag{B.31}$$

Obviously, the number of salient points of g on (x_0,b) is not greater than m, then by (B.29), (B.31), and the inductive assumption, we obtain that

$$\int_a^b \big(f'(x)\big)^2\,\mathrm{d}x = \int_a^{x_1} \big(f'(x)\big)^2\,\mathrm{d}x + \int_{x_1}^b \big(f'(x)\big)^2\,\mathrm{d}x$$
$$\ge \int_a^{x_1} \big(f'_{x_0}(x)\big)^2\,\mathrm{d}x + \int_{x_1}^b \big(f'_{x_0}(x)\big)^2\,\mathrm{d}x$$
$$= \int_a^{x_0} \big(f'_{x_0}(x)\big)^2\,\mathrm{d}x + \int_{x_0}^b \big(f'_{x_0}(x)\big)^2\,\mathrm{d}x \ge \int_a^{x_0} \big(g'(x)\big)^2\,\mathrm{d}x$$
$$+ \int_{x_0}^b \big(g'(x)\big)^2\,\mathrm{d}x = \int_a^b \big(g'(x)\big)^2\,\mathrm{d}x,$$

which completes the proof. □

Appendix C. The projection method

The algorithm below presents how to find the marginal random variables $\{X_i\}_{i=1}^n$ explicitly such that $M_\phi(\omega^\top X)$ attains its maximum, i.e.,

$$\sup_{X\in\mathcal{P}(\mu,\Sigma)} M_\phi(w^\top X) = \mu^\top w + \sqrt{w^\top \Sigma w}\sqrt{\int_0^1 (\phi'_*(u)-1)^2\,du},$$

which has been proved in Proposition 2.3.

In the above algorithm, $\{e_1, e_2, \dots, e_n\}$ in the sixth step means a standard basis for $\mathbb{R}^n$. Namely, $e_i = (0, 0\dots, 0, 1, 0, \dots, 0)$ with 1 appears at the ith position.

Example C.1. Suppose $\omega = (0.2, 0.3, 0.5)$, $\mu = (-30, 10, 6)$, and $\Sigma = \begin{pmatrix} 75 & \frac{50}{3} & -300 \\ \frac{50}{3} & 100 & -\frac{160}{3} \\ -300 & -\frac{160}{3} & 36 \end{pmatrix}$. Then the worst-case DRM can be evaluated explicitly and the worst-case distribution for $X^* = \sum_{i=1}^3 \omega_i X_i$ can be obtained by Corollary 2.3. By applying the above algorithm, we obtain that

$$X_1 = 5(X^* - \delta_1 + \delta_2 - 6),$$
$$X_2 = \frac{10}{3}(2X^* - \delta_1 - 2\delta_2 + 3),$$
$$X_3 = 2(-2X^* + 2\delta_1 + \delta_2 + 3).$$

Appendix D. Moments uncertainty

The assumption about known first two moments requires a complete or a perfect data set, which is generally impossible to obtain in practice. Therefore, we assume that the mean and covariance, a vector-matrix pair, belongs to a given subset $\mathcal{P}$ of $\mathbb{R}^n \times \mathscr{U}_n$, where $\mathscr{U}_n$ could describe the set of positive semidefinite matrices and possibly with other requirements. $\mathcal{P}$ is called an uncertainty set. Therefore, the worst-case DRM can be obtained by solving the following optimization problem:

$$\begin{aligned} \sup_{\mu,\Sigma} \quad & \mu^\top w + \sqrt{w^\top \Sigma w} \cdot \|\phi'_*(u) - 1\|_2 \\ \text{s.t.} \quad & (\mu, \Sigma) \in \mathcal{P}, \end{aligned} \tag{D.1}$$

where $\mathcal{P}$ belongs to the box uncertainty in this paper.

D1. Box uncertainty

Suppose that the uncertainty set $\mathscr{S}$ belongs to a box, i.e., the mean and covariance of the underlying distributions have componentwise bounds:

$$\mathscr{S} = \left\{(\mu, \Sigma): \ \mu_- \leq \mu \leq \mu_+, \ \Sigma_- \leq \Sigma \leq \Sigma_+, \ \Sigma \succeq 0\right\},$$

where μ_- and μ_+ are componentwise lower and upper bounds for the mean vector μ respectively; Σ_- and Σ_+ are componentwise lower and upper bounds for the covariance matrix Σ respectively. Since the component interval $[\Sigma_-, \Sigma_+]$ may not contain a positive semidefinite matrix, we assume that there exists a positive semidefinite Σ_0 such that $\Sigma_- \leq \Sigma_0 \leq \Sigma_+$. See Ghaoui et al. (2003) for more details about the box uncertainty set.

Now the optimization problem (D.1) under the box uncertainty for the worst-case DRM becomes

$$\begin{aligned} \sup_{\mu,\Sigma} \quad & \mu^\top w + \sqrt{w^\top \Sigma w} \cdot \|\phi'_*(u) - 1\|_2 \\ \text{s.t.} \quad & \mu_- \leq \mu \leq \mu_+, \ \Sigma_- \leq \Sigma \leq \Sigma_+, \ \Sigma \succeq 0, \end{aligned} \tag{D.2}$$

Similarly, it is obvious that, the optimization problem (D.2) reduces to $\sup_{X\in\mathcal{P}(\mu,\Sigma)} M_\phi(w^\top X)$ (see Proposition 2.3) when the mean and covariance are exactly known, i.e., $\mu_- = \mu_+$ and $\Sigma_- = \Sigma_+$.

The optimization problem above can also be solved as an SDP problem. With the SDP duality method we can show that the problem (D.2) is equivalent to the following minimization SDP problem

$$\begin{aligned} \inf_{\lambda_\pm,\Lambda_\pm,\nu} \quad & \langle \Lambda_+, \Gamma_+\rangle - \langle \Lambda_-, \Gamma_-\rangle + \lambda_+^\top \mu_+ - \lambda_-^\top \mu_- \\ & +\nu \int_0^1 (\phi'_*(u) - 1)^2 \mathrm{d}u \\ \text{s.t.} \quad & \lambda_+ \geq 0, \ \lambda_- \geq 0, \ \Lambda_+ \geq 0, \ \Lambda_- \geq 0, \\ & \begin{pmatrix} \Lambda_+ - \Lambda_- & w/2 \\ w^\top/2 & \nu \end{pmatrix} \succeq 0, \ w = \lambda_+ - \lambda_-. \end{aligned} \tag{D.3}$$

Therefore, problem (D.3) can be treated as a semidefinite programming (SDP) problem, which can be solved efficiently by numerical methods (Grant & Boyd, 2014).

References

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance, 26*(7), 1505–1518.

Artzner, P., Delbaen, F., & Eber, J. (1999). Coherent measures of risk. *Mathematical Finance, 9*(3), 203–228.

Belles-Sampera, J., Guillén, M., & Santolino, M. (2013). Beyond value-at-risk: GlueVaR distortion risk measures. *Risk Analysis, 34*(1), 121–134.

Belles-Sampera, J., Guillén, M., & Santolino, M. (2016). What attitudes to risk underlie distortion risk measure choices? *Insurance: Mathematics and Economics, 68*, 101–109.

Berkhouch, M. (2021). Distributional uncertainty for spectral risk measures. In N. Teixeira, & I. Lisboa (Eds.), *Handbook of research on financial management during economic downturn and recovery* (pp. 476–494). IGI Global.

Bertazzi, L., & Secomandi, N. (2020). Worst-case benefit of restocking for the vehicle routing problem with stochastic demands. *Operations Research, 68*(3), 671–675.

Bertsimas, D., Doan, X., Natarajan, K., & Teo, C. (2010). Models for minimax stochastic linear optimization problems with risk aversion. *Mathematics of Operations Research, 35*(3), 580–602.

Brandtner, M. (2013). Conditional value-at-risk, spectral risk measures and (non-)diversification in portfolio selection problems—A comparison with mean–variance analysis. *Journal of Banking and Finance, 37*(12), 5526–5537.

Brandtner, M. (2016). Spectral risk measures: Properties and limitations: Comment on dowd, cotter, and sorwar. *Journal of Financial Services Research, 49*(1), 121–131.

Chen, L., He, S., & Zhang, S. (2011). Tight bounds for some risk measures, with applications to robust portfolio selection. *Operations Research, 59*(4), 847–865.

Chen, X., He, S., Jiang, B., Ryan, C., & Zhang, T. (2020). The discrete moment problem with nonconvex shape constraints. *Operations Research, 69*(1), 279–296.

Chen, Z., Sim, M., & Xu, H. (2019). Distributionally robust optimization with infinitely constrained ambiguity sets. *Operations Research, 67*(5), 1328–1344.

Cont, R., Deguest, R., & Scandolo, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance, 10*(6), 593–606.

Das, B., Dhara, A., & Natarajan, K. (2021). On the heavy-tail behavior of the distributionally robust newsvendor. *Operations research, 69*(4), 1077–1099.

Dhaene, J., Kukush, A., Linders, D., & Tang, Q. (2012). Remarks on quantiles and distortion risk measures. *European Actuarial Journal, 2*, 319–328.

Ghaoui, L., Oks, M., & Oustry, F. (2003). Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research, 51*, 543–556.

Grant, M., & Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.0 beta. http://www.cvxr.com/cvx.

Guo, J., He, S., Jiang, B., & Wangm, Z. (2022). A unified framework for generalized moment problems: A novel primal-dual approach. arXiv:2201.01445.

Guo, S., & Xu, H. (2021). Robust spectral risk optimization when the subjective risk aversion is ambiguous: A moment-type approach. *Mathematical Programming*. https://doi.org/10.1007/s10107-021-01630-5.

Li, J. (2018). Closed-form solutions for worst-case law invariant risk measures with application to robust portfolio optimization. *Operations Research, 66*(6), 1533–1541.

Li, L., Shao, H., Wang, R., & Yang, J. (2017). Worst-case range value-at-risk with partial information. *SIAM Journal on Financial Mathematics, 9*(1), 190–218.

Natarajan, K., Pachamanova, D., & Sim, M. (2008). Incorporating asymmetric distributional information in robust value-at-risk optimization. *Management Science, 54*(3), 573–585.

Natarajan, K., Sim, M., & Uichanco, J. (2010). Tractable robust expected utility and risk models for portfolio optimization. *Mathematical Finance, 20*(4), 695–731.

Peposcu, I. (2005). A semidefinite programming approach to optimal-moment bounds for convex classes of distributions. *Mathematics of Operation Research, 30*(3), 632–657.

Pichler, A. (2013). Evaluations of risk measures for different probability measures. *SIAM Journal on Optimization, 23*(1), 530–551.

Pichler, A. (2014). Insurance pricing under ambiguity. *European Actuarial Journal, 4*(2), 335–364.

Pichler, A., & Xu, H. (2022). Quantitative stability analysis for minimax distributionally robust risk optimization. *Mathematical Programming, 191*(1), 44–77.

Popescu, I. (2007). Robust mean-covariance solutions for stochastic optimization. *Operations Research, 55*(1), 98–112.

Sahamkhadam, M., Stephan, A., & Östermarkc, R. (2022). Copula-based Black–Litterman portfolio optimization. *European Journal of Operational Research, 297*(3), 1055–1070.

Staden, P., Dang, D., & Forsyth, P. (2021). The surprising robustness of dynamic mean-variance portfolio optimization to model misspecification errors. *European Journal of Operational Research, 289*(2), 774–792.

Toumazis, I., & Kwon, C. (2015). Worst-case conditional value-at-risk minimization for hazardous materials transportation. *Transportation Science, 50*(4), 1174–1187.

Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics, 17*(1), 43–54.

Wang, S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance, 67*(1), 15–36.

Wiesemann, W., Kuhn, D., & Sim, M. (2014). Distributionally robust convex optimization. *Operations Research, 62*(6), 1358–1376.

Yaari, M. (1987). The dual theory of choice under risk. *Econometrica, 55*(1), 95–115.

Yu, Y., Li, Y., Dale, S., & Csaba, S. (2009). A general projection property for distribution families. *Advances in Neural Information Processing Systems, 22*, 2232–2240.

Zhu, S., & Fukushima, M. (2009). Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research, 57*(5), 1155–1168.

Zymler, S., Kuhn, D., & Rustem, B. (2013). Worst-case value at risk of nonlinear portfolios. *Management Science, 59*(1), 172–188.