Robust Optimization of Rank-Dependent Models with Uncertain Probabilities*

Guanyu Jin[a], Roger J. A. Laeven[a,c,d], and Dick den Hertog[b]

[a]Dept. of Quantitative Economics, University of Amsterdam,
1001 NJ Amsterdam, The Netherlands
[b]Dept. of Business Analytics, University of Amsterdam,
1001 NJ Amsterdam, The Netherlands
[c]EURANDOM, 5600 MB Eindhoven, The Netherlands
[d]CentER, Tilburg University, 5000 LE Tilburg, The Netherlands

April 14, 2025

Abstract

This paper studies distributionally robust optimization for a rich class of risk measures with ambiguity sets defined by ϕ-divergences. The risk measures are allowed to be non-linear in probabilities, are represented by Choquet integrals possibly induced by a probability weighting function, and encompass many well-known examples. Optimization for this class of risk measures is challenging due to their rank-dependent nature. We show that for various shapes of probability weighting functions, including concave, convex and inverse S-shaped, the robust optimization problem can be reformulated into a rank-independent problem. In the case of a concave probability weighting function, the problem can be reformulated further into a convex optimization problem that admits explicit conic representability for a collection of canonical examples. While the number of constraints in general scales exponentially with the dimension of the state space, we circumvent this dimensionality curse and develop two types of algorithms. They yield tight upper and lower bounds on the exact optimal value and are formally shown to converge asymptotically. This is illustrated numerically in a robust newsvendor problem and a robust portfolio choice problem.

JEL Classification: C61, D81.

OR/MS Classification: Programming: Stochastic. Decision analysis: Risk.

Keywords: Distributionally robust optimization; rank dependence; ϕ-divergence; probability weighting functions; inverse S-shape.

1 Introduction

Many stochastic optimization problems in management science, operations research, engineering, economics, and finance arise from decisions involving risk (probabilities given) and ambiguity (prob-

*We are very grateful to Ronny Ben-Tal, Henry Lam and conference and seminar participants at EURANDOM, the 2023 SIAM Conference on Optimization, the Tinbergen Institute, the University of Amsterdam and the University of Vienna for their comments and suggestions. `Python` code to implement the optimization procedures developed in this paper is available from `https://github.com/GuanJinNL/ROptRDU.github.io.git`. This research was funded in part by the Netherlands Organization for Scientific Research under grant NWO VICI 2020–2027 (Jin, Laeven).
Email addresses: `G.Jin@uva.nl`, `R.J.A.Laeven@uva.nl`, and `D.denHertog@uva.nl`.

abilities unknown).

A variety of models for decision under risk have been proposed. Among the most popular and empirically viable models is the rank-dependent utility (RDU) model of Quiggin (1982). In RDU, the utility loss associated to a random variable X under the probabilistic model $\mathbb{P}$ is measured by a rank-dependent evaluation ρ with respect to a non-additive, distorted measure $h \circ \mathbb{P}$:[1]

$$\rho_{u,h,\mathbb{P}}(X) \triangleq \int -u(X)\mathrm{d}(h \circ \mathbb{P}),$$

where $u : \mathbb{R} \to \mathbb{R}$ is a utility function, assumed to be non-decreasing, and $h : [0,1] \to [0,1]$, with $h(0) = 1 - h(1) = 0$ and non-decreasing, is a distortion or probability weighting function. RDU serves as a pivotal building block of prospect theory (Tversky and Kahneman, 1992) and encompasses expected utility (von Neumann and Morgenstern, 1944) when h is linear and the dual theory of Yaari (1987) when u is affine. It accommodates Allais (1953) type phenomena that are incompatible with expected utility. The utility function captures attitude toward wealth and the shape of the distortion function, e.g., concave, convex or (inverse) S-shaped, dictates attitude toward risk. Importantly, under RDU probabilities of outcomes are weighted according to $h(F_X(x)) \triangleq h(\mathbb{P}(X \leq x))$, leading to an evaluation that is non-linear in probabilities, which depend on the ranking of outcomes. These non-linear and rank-dependent features bring major computational challenges.

In the aforementioned theories for decision under risk, $\mathbb{P}$ is assumed to be given. When ambiguity is present, $\mathbb{P}$ is unknown. Ambiguity is often treated via a worst-case approach that is robust against malevolent nature. For example, Gilboa and Schmeidler (1989) introduce maxmin expected utility, a.k.a. multiple priors, under which risks are evaluated according to their worst-case expected utility taken over a set of probabilistic models. Hansen and Sargent (2001, 2007) propose multiplier preferences under which probabilistic models far away from a reference model are penalized according to the Kullback-Leibler divergence (a.k.a. relative entropy). The variational preferences of Maccheroni et al. (2006) admit a general penalty function, thus significantly generalizing the maxmin and multiplier preferences models. Laeven and Stadje (2023) develop a rank-dependent theory for decision under risk and ambiguity that encompasses both the dual and rank-dependent counterparts of Gilboa and Schmeidler (1989) and Maccheroni et al. (2006).

Parallel to these developments, the field of (distributionally) robust optimization has been studying risk and ambiguity from a computational perspective. In robust optimization, ambiguity sets are often constructed exogenously from data, as a confidence region of the true underlying distribution, rather than endogenously based on preferences. A widely used family of statistical estimators for distributional uncertainty is given by ϕ-divergences (Csiszár, 1975; Ben-Tal and Teboulle, 1986, 1987; Pardo, 2006). Many distributionally robust optimization problems with these types of ambiguity sets can be reformulated into tractable robust counterparts, which can then be

[1]See also Schmeidler (1986, 1989) for related pioneering developments.

solved efficiently using standard optimization algorithms (see e.g., Ben-Tal et al., 2013, Wiesemann et al., 2014, and Esfahani and Kuhn, 2018).

Although a variety of distributionally robust optimization techniques have been developed for optimization under uncertainty, most of the literature is concerned with stochastic models in which the probabilities appear linearly in the optimization problems, such as (possibly generalized) variants of expected value or expected utility maximization. Despite the growing interest in and applications of rank-dependent models across a wide variety of fields (see e.g., Denneberg, 1994, Wakker, 2010, Föllmer and Schied, 2016, and the references therein), optimization of these models—whether ambiguity is present or not—is still relatively underdeveloped. The main difficulties lie in both the non-linearity in probabilities and the rank-dependence: for each value of a decision vector, the rank-dependent evaluation of an uncertain objective or constraint can be permuted, since the ranking of the outcomes depends upon the decision vector.

In this paper, we develop an efficient approach for optimizing rank-dependent models, with and without uncertain probabilities. More precisely, we study the following nominal and robust minimization problems, in a discrete probability space setting:

$$\inf_{\mathbf{a}\in\mathcal{A}} \rho_{u,h,\mathbf{p}}(f(\mathbf{a},\mathbf{X})), \tag{1}$$

$$\inf_{\mathbf{a}\in\mathcal{A}} \sup_{\mathbf{q}\in\mathcal{D}_\phi(\mathbf{p},r)} \rho_{u,h,\mathbf{q}}(f(\mathbf{a},\mathbf{X})), \tag{2}$$

where $\mathbf{a} \in \mathbb{R}^{n_a}$ is the decision vector, $\mathcal{A}$ is a set of constraints, $\mathbf{X} \in \mathbb{R}^l$ is a random vector, $f : \mathbb{R}^{n_a} \times \mathbb{R}^l \to \mathbb{R}$ is a deterministic function, $\mathcal{D}_\phi(\mathbf{p}, r)$ is a ϕ-divergence ambiguity set (formally defined in (8)), and $\rho_{u,h,\mathbf{q}}(.)$ is a rank-dependent evaluation (formally defined in (5)), for some probability vectors $\mathbf{p}, \mathbf{q} \in \mathbb{R}^m$.

Our main contributions can be summarized as follows:

- We show that we can reformulate problems (1) and (2), as well as their epigraph formulations, into rank-independent, tractable (robust) counterparts, for the various empirically relevant shapes of the probability weighting function h, including concave, convex, and inverse S-shaped functions.

- For concave probability weighting functions, we show that the reformulated robust counterpart admits a conic representation, provided h and ϕ are conic representable functions. For a list of canonical examples of h and ϕ, we provide explicit epigraph representations of the reformulated robust counterpart that can directly be implemented into standard conic optimization programs such as CVXPY.

- While the reformulated robust counterparts are more tractable than (1) and (2), their multiplicity of constraints increases exponentially in the dimension of the underlying discrete probability space. We provide two types of algorithms to circumvent this curse of dimensionality. Each of them, or a combination thereof in the case of convex and inverse S-shaped probability weighting functions, yields tight upper and lower bounds that, as we formally establish, converge to the optimal objective value of the exact problem. Moreover, we show that one of our algorithms can also be applied to a more general type of rank-dependent model, namely Choquet expected utility (CEU, Schmeidler, 1989).[2]

- We provide numerical examples of our approach in applications of robust optimization with rank-dependent models and uncertain probabilities, in the context of inventory planning and portfolio optimization. We efficiently obtain optimal decisions that are robust against uncertainty, and yield accurate optimal objective values. Concrete examples with codes are provided on `https://github.com/GuanJinNL/ROptRDU.github.io.git`.

To our best knowledge, our approach is the first that can handle nominal and distributionally robust, generic optimization problems involving rank-dependent evaluations with possibly inverse S-shaped probability weighting functions.

1.1 Related Literature

Our work builds on the decision-theoretic literature on evaluating risk and ambiguity. Specifically, we consider in (1)–(2) the rank-dependent evaluations of Quiggin (1982) and adopt a robust approach as in Gilboa and Schmeidler (1989), Maccheroni et al. (2006), Hansen and Sargent (2001, 2007) and Laeven and Stadje (2023), which can be viewed as (generalized) decision-theoretic foundations of the classical decision rule of Wald (1950); see also Huber (1981). We contribute to this literature by developing corresponding optimization techniques.

An initial connection between robust risk measures and robust optimization has been studied in an early paper by Bertsimas and Brown (2009). They show that a decision-maker's preference, as represented by a coherent risk measure, can provide a device for constructing an uncertainty set for robust optimization purposes. The authors were able to obtain tractability results for distortion risk measures under a specific parameterization and the (strict) assumption that the underlying discrete probability space is uniform. Our paper relaxes these assumptions, while also providing a blueprint for constructing uncertainty sets tailored to general risk and ambiguity preferences. Postek et al. (2016) studied distributionally robust optimization (DRO) problems for a broad collection of uncertainty sets and risk measures. However, for classes of risk measures that are non-linear in probabilities, such as distortion risk measures, the tractability remained unknown (see their Table I); they are covered as special cases in this paper. Although different choices of uncertainty sets are possible, our paper focuses on uncertainty sets defined by ϕ-divergences, which constitute a family of divergences including the Kullback-Leibler divergence, Burg entropy and Hellinger distance. The study of ϕ-divergences in DRO is motivated by several earlier studies; see e.g., Ben-Tal and Teboulle (1986, 1987) and Ben-Tal et al. (2013).

[2]Under CEU, the non-additive measure need not be obtained by distorting an additive probability measure.

Recently, Cai et al. (2025) and Pesenti et al. (2020) studied DRO problems involving distortion risk measures with general non-concave distortion functions. Their interesting results show that the DRO problem associated with a non-concave distortion function is equivalent to that with its concave envelope approximation provided that the underlying uncertainty set obeys certain moment conditions. In this paper, we show that this equivalence is typically not satisfied in our general setting, thus requiring novel techniques. Optimization of non-expected utility models with possibly non-concave distortion functions has also been studied in innovative work of He and Zhou (2011), in particular in the context of portfolio choice, introducing the so-called quantile method; see also Carlier and Dana (2006). The effectiveness of this method relies on the ability to parametrize the distributions of all random objective functions $\{f(\mathbf{a}, \mathbf{X}), \mathbf{a} \in \mathcal{A}\}$ by a set of quantile functions that satisfy finitely many constraints. For a general convex function f and feasibility set $\mathcal{A}$ that are considered in (1), there is, however, no systematic approach to obtain such a quantile formulation. Our work provides a method to optimize (both robust and nominal) rank-dependent models for a broad class of decision problems and can directly be implemented in standard optimization software. Finally, Delage et al. (2022) and Wang and Xu (2023) analyze 'preference robust' optimization, which considers uncertainty in the distortion function rather than in the underlying probabilistic model.

The remainder of the paper is organized as follows. Section 2 introduces the setting and notation. Sections 3 to 5 study the optimization problems (1) and (2) for concave distortion functions. The techniques we develop are extended in Section 6 to non-concave, inverse S-shaped distortion functions. Section 7 presents our numerical experiments. Conclusions are in Section 8. In Electronic Companions EC.1–EC.9, we provide all the proofs, additional examples, and technical details.

2 Setup and Notation

2.1 Rank-Dependent Evaluation

Let $(\Omega, \mathcal{F})$ be a measurable space. We define the rank-dependent evaluation $\rho_{u,h,\mathbb{Q}}$ of the utility loss associated with a random variable $X : \Omega \to \mathbb{R}$ under a given probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ as the integral with respect to the non-additive measure $h \circ \mathbb{Q}$ (Denneberg, 1994), or equivalently, as a Choquet integral (Choquet, 1954):

$$\rho_{u,h,\mathbb{Q}}(X) \triangleq \int -u(X)\,\mathrm{d}(h \circ \mathbb{Q}) \tag{3}$$

$$= \int_0^\infty h\left(\mathbb{Q}[-u(X) > t]\right)\mathrm{d}t + \int_{-\infty}^0 \left(h\left(\mathbb{Q}[-u(X) > t]\right) - 1\right)\mathrm{d}t. \tag{4}$$

Here, u is a non-decreasing utility function defined on a suitable domain containing the support of X. Furthermore, the function $h : [0,1] \to [0,1]$ is a distortion, or probability weighting, function that is non-decreasing and satisfies $h(0) = 0$ and $h(1) = 1$. We note that $\rho_{u,h,\mathbb{Q}}$ is also known as a *distortion risk measure* when u is the identity function. In this paper, we consider distortion functions that may be concave, convex or inverse S-shaped.

Definition 1. We say that h is inverse S-shaped if, for some $p^0 \in (0,1)$, we have that h is concave for $p \leq p^0$ and h is convex for $p \geq p^0$.

If Ω is discrete with $|\Omega| = m$, then the integral (4) reduces to a rank-dependent sum. Let $x_{(1)} \geq x_{(2)} \geq \ldots \geq x_{(m)}$ denote the ranked realizations of X, with $\{(i)\}_{i=1}^m$ denoting the indices of the ranked realizations. The monotonicity of u preserves the ranking of $(x_i)_{i=1}^m$. Therefore, we have

$$\rho_{u,h,\mathbf{q}}(X) = \sum_{i=1}^{m} -\left(h\left(\sum_{j=i}^{m} q_{(j)}\right) - h\left(\sum_{j=i+1}^{m} q_{(j)}\right)\right) u\left(x_{(i)}\right), \tag{5}$$

with $\mathbf{q} \in [0,1]^m$ the probability vector associated to Ω and $\sum_{j=m+1}^{m} q_{(j)} \triangleq 0$, by convention.

A well-known example of a rank-dependent evaluation is the Conditional-Value-at-Risk (a.k.a. Expected Shortfall; Föllmer and Schied, 2016, Section 4.6), $\mathrm{CVaR}_{1-\alpha}(X)$, which is defined as

$$\mathrm{CVaR}_{1-\alpha}(X) \triangleq \frac{1}{1-\alpha}\int_0^{1-\alpha} \mathrm{VaR}_\gamma(X)\,\mathrm{d}\gamma, \quad 0 \leq \alpha < 1, \tag{6}$$

where $\mathrm{VaR}_\gamma(X) \triangleq \inf\{x : \mathbb{Q}(-X \leq x) \geq 1-\gamma\}$, $0 < \gamma < 1$, is the Value-at-Risk. For a certain level $\alpha \in [0,1)$, $\mathrm{CVaR}_{1-\alpha}(X)$ can be interpreted as the (sign-changed) average of the left $(1-\alpha)\cdot 100\%$ tail of the risk X. It is easily verified that $\mathrm{CVaR}_{1-\alpha}$ is a rank-dependent evaluation with linear utility function and distortion function $h(p) = \min\left\{\frac{p}{1-\alpha}, 1\right\}$. Other canonical examples of distortion functions are given in Table EC.1 of Electronic Companion EC.1. The distortion function captures attitude toward risk whereas the utility function describes attitude toward wealth (see e.g., Quiggin, 1982, Yaari, 1987, Chew et al., 1987, and Eeckhoudt and Laeven, 2021).

2.2 ϕ-Divergence Ambiguity Sets

We construct ambiguity (or uncertainty) sets using ϕ-divergences. For discrete outcome spaces Ω with $|\Omega| = m$, the ϕ-divergence $I_\phi(\mathbf{q},\mathbf{p})$ between two probability vectors $\mathbf{q},\mathbf{p} \in [0,1]^m$ is defined as

$$I_\phi(\mathbf{q},\mathbf{p}) \triangleq \sum_{i=1}^{m} p_i \phi\left(\frac{q_i}{p_i}\right). \tag{7}$$

Here, $\phi : [0,\infty) \to \mathbb{R}$ is a convex function that satisfies the following conventions: $\phi(1) = 0$, $0\phi(0/0) \triangleq 0$ and $0\phi(x/0) \triangleq x \lim_{t\to\infty} \phi(t)/t$ (see Pardo, 2006, Definition 1.1). For a nominal probability vector $\mathbf{p}$ and $r > 0$, the ϕ-divergence ambiguity set is defined as

$$\mathcal{D}_\phi(\mathbf{p}, r) \triangleq \left\{\mathbf{q} \in \mathbb{R}^m \;\middle|\; \mathbf{q} \geq \mathbf{0},\ \mathbf{q}^T\mathbf{1} = 1, I_\phi(\mathbf{q},\mathbf{p}) \leq r\right\}. \tag{8}$$

Here, $\mathbf{0}, \mathbf{1} \in \mathbb{R}^m$ are the vectors with all entries equal to 0 and 1, respectively. As outlined in Ben-Tal et al. (2013) (Section 3.2), one may construct $\mathcal{D}_\phi(\mathbf{p}, r)$ as a statistical confidence set by replacing $\mathbf{p}$ with an empirical estimator $\hat{\mathbf{p}}_n$. Indeed, as shown in Pardo (2006) (Corollary 3.1), under the null-hypothesis $H_0 : \mathbf{p} = \mathbf{p}_0$, and provided ϕ is twice continuously differentiable in a neighborhood of 1 with $\phi''(1) > 1$, the following object converges to a chi-squared distribution:[3]

$$\frac{2n}{\phi''(1)} I_\phi(\mathbf{p}_0, \hat{\mathbf{p}}_n) \rightsquigarrow \chi^2_{m-1,1-\alpha}. \tag{9}$$

Here, n is the sample size used for constructing the empirical estimator, and $\chi^2_{m-1,1-\alpha}$ is the $(1-\alpha)$-quantile of the chi-square distribution with $m-1$ degrees of freedom.[4] Using (9), one can construct an asymptotic confidence set $\mathcal{D}_\phi(\hat{\mathbf{p}}_n, r)$ by choosing

$$r = \frac{\phi''(1)}{2n} \chi^2_{m-1,1-\alpha}. \tag{10}$$

2.3 Problem Formulations, Terminology and Assumptions

We study the following nominal and robust minimization problems:

$$\inf_{\mathbf{a} \in \mathcal{A}} \rho_{u,h,\mathbf{p}}(f(\mathbf{a}, \mathbf{X})), \tag{P-Nom}$$

$$\inf_{\mathbf{a} \in \mathcal{A}} \sup_{\mathbf{q} \in \mathcal{D}_\phi(\mathbf{p}, r)} \rho_{u,h,\mathbf{q}}(f(\mathbf{a}, \mathbf{X})), \tag{P}$$

where $\mathbf{a} \in \mathbb{R}^{n_a}$ is the decision vector contained in a compact set of constraints $\mathcal{A}$ consisting of convex inequalities, $\mathbf{X} \in \mathbb{R}^l$ is a random vector, $f : \mathbb{R}^{n_a} \times \mathbb{R}^l \to \mathbb{R}$ is a jointly convex function in $(\mathbf{a}, \mathbf{x})$, $\mathcal{D}_\phi(\mathbf{p}, r)$ is a ϕ-divergence ambiguity set defined in (8) with respect to a nominal probability vector $\mathbf{p}$, and $\rho_{u,h,\mathbf{q}}(.)$ is the rank-dependent evaluation defined in (5), with respect to some probability vector $\mathbf{q} \in \mathbb{R}^m$.

Henceforth, we often consider the nominal and robust rank-dependent evaluations in a constraint form induced by the following epigraph formulation:

$$\rho_{u,h,\mathbf{p}}(f(\mathbf{a}, \mathbf{X})) \leq c, \tag{P-Nom-EG}$$

$$\sup_{\mathbf{q} \in \mathcal{D}_\phi(\mathbf{p}, r)} \rho_{u,h,\mathbf{q}}(f(\mathbf{a}, \mathbf{X})) \leq c. \tag{P-EG}$$

That is, we emphasize that we are able to deal with a robust rank-dependent evaluation both in the

[3] From a statistical perspective, it may be more natural to consider $I_\phi(\hat{\mathbf{p}}_n, \mathbf{p}_0)$, where the true model $\mathbf{p}_0$ appears as the reference model. However, in the robust optimization literature, $I_\phi(\mathbf{p}_0, \hat{\mathbf{p}}_n)$ appears more commonly. According to Theorem 3.1 and Corollary 3.1 of Pardo (2006), both objects have the same limiting distribution under the null.

[4] That is, $\mathbb{P}(Z \leq \chi^2_{m-1,1-\alpha}) = 1 - \alpha$, for $Z \sim \chi^2_{m-1}$.

objective and in the constraint, where in the latter case we consider the following type of problem:

$$\begin{aligned} &\min_{\mathbf{a}\in\mathcal{A}} \; g(\mathbf{a}) \\ &\text{s.t.} \sup_{\mathbf{q}\in\mathcal{D}_\phi(\mathbf{p},r)} \rho_{u,h,\mathbf{q}}(f(\mathbf{a},\mathbf{X})) \leq c, \end{aligned} \tag{P-constraint}$$

where g is convex in $\mathbf{a}$, and $c \in \mathbb{R}$ is a fixed parameter (in contrast to being an epigraph variable). The nominal version of (P-constraint) is defined analogously, with the only difference that the robust constraint is replaced by the nominal constraint (P-Nom-EG).

Additionally, we would like to mention the alternative, but highly related, approach of defining a robust rank-dependent evaluation, which appears in the decision theory literature (see e.g., Laeven and Stadje, 2023):

$$\tilde{\rho}_{\text{rob}}(X) \triangleq \sup_{\mathbf{q}\geq\mathbf{0},\mathbf{q}^T\mathbf{1}=1} \rho_{u,h,\mathbf{q}}(X) - \theta I_\phi(\mathbf{q},\mathbf{p}), \quad \theta > 0. \tag{11}$$

In Electronic Companion EC.4, we provide additional details on how (11) can be reformulated similar to (P-EG), and that the minimization problem (P) is equivalent to minimizing (11) for a specific θ.

Let $[m] \triangleq \{1,\ldots,m\}$ for an integer m. The following assumptions are made throughout this paper:

Assumption 1. *The optimization problem* (P) *is finite:* $-\infty <$(P)$< \infty$.

Assumption 2. *The nominal probability vector* $\mathbf{p}$ *satisfies* $p_i > 0$ *for all* $i \in [m]$.

Assumption 3. *The functions* $\phi(t), -h(p), -u(f(\mathbf{a},\mathbf{x}))$ *are lower-semicontinuous on their respective domains.*

Assumption 4. (P-constraint) *is finite and contains a Slater point, i.e., there exists* $\mathbf{a}_0 \in \text{int}(\mathcal{A})$ *such that* $\sup_{\mathbf{q}\in\mathcal{D}_\phi(\mathbf{p},r)} \rho_{u,h,\mathbf{q}}(f(\mathbf{a}_0,\mathbf{X})) < c$. *Furthermore, we assume* (P-constraint) $> \inf_{\mathbf{a}\in\mathcal{A}} g(\mathbf{a})$.

Assumption 1 can be satisfied if e.g., the image set $\{u(f(\mathbf{a},\mathbf{x})) \mid \mathbf{a}\in\mathcal{A}, \mathbf{x}\in\text{supp}(\mathbf{X})\}$ is contained in a bounded interval. Assumption 2 constitutes a weak redundancy condition. Assumption 3 is to ensure that the optimization problem (P) has an optimal solution in $\mathcal{A}$, and Assumption 4 is to ensure that (P-constraint) satisfies the strong duality theorem.

In this paper, the term *robust solution* refers to an optimal solution of (P) or (P-constraint). Similarly, a *nominal solution* refers to an optimal solution of (P-Nom) or the nominal version of (P-constraint).

2.4 Further Notation

For a convex function $g : \mathbb{R}^d \to \mathbb{R}$, we denote by g^* its convex *conjugate* $g^*(\mathbf{y}) \triangleq \sup_{\mathbf{z}\in\text{dom}(g)}\{\mathbf{y}^T\mathbf{z} - g(\mathbf{z})\}$, where $\text{dom}(g) \triangleq \{\mathbf{z} \mid g(\mathbf{z}) < +\infty\}$ is the effective domain of g. Note that g^* is always convex, since it is the pointwise supremum of a linear function in $\mathbf{y}$. Furthermore, g^* is non-decreasing if $\mathrm{dom}(g) \subset [0, \infty)$. For $\lambda > 0$, the *perspective* of g is the function $\tilde{g} : (\mathbf{z}, \lambda) \mapsto \lambda g\left(\frac{\mathbf{z}}{\lambda}\right)$. By convention, $\tilde{g}(\mathbf{0}, 0) = 0$ and $\tilde{g}(\mathbf{z}, 0) = \infty$ for $x \neq 0$. Note that $\tilde{g}$ is convex if g is convex. The *epigraph* of a function g is the set $\mathrm{Epi}(g) \triangleq \{(\mathbf{z}, t) : g(\mathbf{z}) \leq t\}$. An epigraph may have a conic representation expressed in a conic inequality $\succeq_{\mathbf{K}}$, where $\mathbf{K}$ is a proper cone (i.e., pointed, closed, convex and with a non-empty interior) and $\mathbf{K}^*$ denotes its dual cone, defined as

$$\mathbf{K}^* \triangleq \{\boldsymbol{\lambda} \in \mathbb{R}^d : \boldsymbol{\lambda}^T \mathbf{z} \geq 0, \forall \mathbf{z} \in \mathbf{K}\};$$

see Chares (2009) and Ben-Tal and Nemirovski (2019) for further details.

3 Robust Counterpart of Rank-Dependent Models

In this section, we show how the optimization problems (P-Nom)–(P) can be reformulated into rank-independent problems with finitely many constraints. Our idea hinges on the insight that a rank-dependent evaluation with a concave distortion function admits a dual representation that is itself a robust optimization problem with linear probabilities and convex uncertainty set. Although the dual representation holds only for concave distortion functions, we show in Section 6 how the same idea can be extended to encompass convex and inverse S-shaped distortion functions. This makes it possible to conduct robust optimization for a broad class of rank-dependent models.

3.1 Reformulation into the Robust Counterpart

We start by reformulating the constraints (P-EG)–(P-Nom-EG) with concave distortion functions, since they form the basis of the more complicated convex and inverse S-shaped cases described in Section 6. In Sections 3–5, without further mentioning, we assume the utility function to be concave; Section 6 allows u to be non-concave. The reformulation relies on utilizing the following dual representation, which involves a *composite uncertainty set.*

Theorem 1. *Let $h : [0, 1] \to [0, 1]$ be a concave distortion function. Then, for all $(\mathbf{a}, c) \in \mathbb{R}^{n_a+1}$, we have that* (P-EG) *is satisfied if and only if*

$$\sup_{(\mathbf{q}, \bar{\mathbf{q}}) \in \mathcal{U}_{\phi,h}(\mathbf{p})} \sum_{i=1}^{m} -\bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i)) \leq c, \tag{12}$$

where $\mathcal{U}_{\phi,h}(\mathbf{p})$ is a convex composite uncertainty set given by

$$\mathcal{U}_{\phi,h}(\mathbf{p}) \triangleq \left\{ (\mathbf{q}, \bar{\mathbf{q}}) \in \mathbb{R}^{2m} \;\middle|\; \begin{array}{l} \sum_{i=1}^m q_i = \sum_{i=1}^m \bar{q}_i = 1 \\ \sum_{i=1}^m p_i \phi\left(\frac{q_i}{p_i}\right) \leq r \\ \sum_{i \in J} \bar{q}_i \leq h\left(\sum_{i \in J} q_i\right), \;\; \forall J \subset [m] \\ q_i, \bar{q}_i \geq 0, \;\; \forall i \in [m] \end{array} \right\}. \tag{13}$$

As a special case of interest, (P-Nom-EG) *is satisfied if and only if*

$$\sup_{\bar{\mathbf{q}} \in M_h(\mathbf{p})} \sum_{i=1}^{m} -\bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i)) \leq c, \tag{14}$$

where $M_h(\mathbf{p})$ *is the set induced by* h*:*

$$M_h(\mathbf{p}) \triangleq \left\{ \bar{\mathbf{q}} \in \mathbb{R}^m \;\middle|\; \begin{array}{l} \sum_{i=1}^m \bar{q}_i = 1 \\ \sum_{i \in J} \bar{q}_i \leq h\left(\sum_{i \in J} p_i\right), \ \forall J \subset [m] \\ \bar{q}_i \geq 0, \ \forall i \in [m] \end{array} \right\}. \tag{15}$$

As a consequence, Theorem 1 implies that the robust problem (P) is equivalent to the following rank-independent problem that is more suitable for robust optimization techniques:

$$\min_{\mathbf{a} \in \mathcal{A}} \sup_{(\mathbf{q}, \bar{\mathbf{q}}) \in \mathcal{U}_{\phi,h}(\mathbf{p})} -\sum_{i=1}^{m} \bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i)). \tag{P-ref}$$

Similarly, the nominal problem (P-Nom) can be reformulated to

$$\min_{\mathbf{a} \in \mathcal{A}} \sup_{\bar{\mathbf{q}} \in M_h(\mathbf{p})} -\sum_{i=1}^{m} \bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i)). \tag{P-Nom-ref}$$

Moreover, the equivalence between (P-EG) and (12) features an interpretation somewhat similar to that in Bertsimas and Brown (2009), which connects the shape of an uncertainty set to decision theory. Indeed, this equivalence can be viewed as a device for constructing a myriad of preference-based uncertainty sets $\mathcal{U}_{\phi,h}(\mathbf{p})$ for robust optimization. In Figure EC.1 of Electronic Companion EC.5, we provide an illustrative visualization of the widely varying shapes of uncertainty sets $\mathcal{U}_{\phi,h}(\mathbf{p})$ that can be generated by selecting specific examples of the deterministic, univariate functions h and ϕ.

The following theorem establishes how the semi-infinite constraints (12) and (14) can be further reformulated into finitely many constraints, yielding a (more) tractable robust counterpart problem.

Theorem 2. *Let* $h : [0,1] \to [0,1]$ *be a concave distortion function. Then, we have that, for all* $(\mathbf{a}, c) \in \mathbb{R}^{n_a+1}$*, the inequality* (12) *holds if and only if there exist* α, β, γ, $(\nu_j)_{j=1}^{2^m-2}, (\lambda_j)_{j=1}^{2^m-2} \in \mathbb{R}$*, such that*

$$\begin{cases} \alpha + \beta + \gamma r + \sum_{i=1}^m p_i \gamma \phi^* \left(\frac{-\alpha + \sum_{j: i \in I_j} \nu_j}{\gamma} \right) + \sum_{j=1}^{2^m-2} \lambda_j (-h)^* \left(\frac{-\nu_j}{\lambda_j} \right) \leq c, & \\ -u(f(\mathbf{a}, \mathbf{x}_i)) - \beta - \sum_{j: i \in I_j} \lambda_j \leq 0, & \forall i \in [m] \\ \lambda_j, \gamma \geq 0, & \forall j \in [2^m - 2], \end{cases} \tag{16}$$

where $I_1, \ldots, I_{2^m-2}$ *are all subsets of* $[m]$*, except the empty set and the set* $[m]$ *itself.*

In the nominal case, we have that the inequality (14) *holds if and only if there exist* $\beta \in \mathbb{R}$, $(\lambda_j)_j \geq 0$ *such that*

$$\begin{cases} \beta + \sum_{j=1}^{2^m-2} \lambda_j h\left(\sum_{k \in I_j} p_k\right) \leq c, \\ -u(f(\mathbf{a}, \mathbf{x}_i)) - \beta - \sum_{j: i \in I_j} \lambda_j \leq 0, \ \forall i \in [m] \\ \beta \in \mathbb{R}, \lambda_j \geq 0, \ \forall j \in [2^m - 2]. \end{cases} \tag{17}$$

Remark 1. For technical reasons, the conjugate $(-h)^*$ in Theorem 2 is taken with respect to the domain $[0, \infty)$: $(-h)^*(y) = \sup_{t \geq 0}\{yt + h(t)\}$, by defining $h(x) = 1$ for all $x > 1$.

Remark 2. In Table EC.1, we provide $(-h)^*$ explicitly for a collection of canonical examples. (P-ref) may also be further reformulated using the "optimistic dual counterpart" (Gorissen et al., 2014). This approach is particularly useful if $(-h)^*$ cannot be computed analytically, as is the case e.g., for $h(p) = \Phi(\Phi^{-1}(p) + \nu)$, $\nu > 0$, with Φ the standard normal cdf; see Wang (2000) and Goovaerts and Laeven (2008). Further details are provided in Electronic Companion EC.6.

3.2 Conic Representability of the Robust Counterpart

In this subsection, we explore the conic representability of the robust counterpart reformulated in Theorem 2. Following Ben-Tal and Nemirovski (2019), a set $\mathcal{S} \subset \mathbb{R}^n$ is conic representable by a cone $\mathbf{K}$ if and only if

$$\mathbf{x} \in \mathcal{S} \Leftrightarrow \exists \ \mathbf{w}, \mathbf{A}, \mathbf{b} : \mathbf{A}\begin{pmatrix} \mathbf{x} \\ \mathbf{w} \end{pmatrix} - \mathbf{b} \succeq_{\mathbf{K}} \mathbf{0}.$$

A function is said to be conic representable if its epigraph is. The reformulated robust counterpart (16) in Theorem 2 contains constraints that are expressed in the perspective of the univariate conjugate functions $(-h)^*$ and ϕ^*. We focus on the conic representability of these constraints:

$$\lambda(-h)^*\left(\frac{-\nu}{\lambda}\right) \leq z, \ \text{and} \ \gamma\phi^*\left(\frac{s}{\gamma}\right) \leq t.$$

In practice, the derivation of the conjugate functions might be difficult. However, the epigraph representation of a conjugate function does not always require an explicit form of the conjugate function itself. The following two lemmas provide a generic approach for determining the conic representation of the epigraph of the perspective function through that of the function itself. Interestingly, this procedure does not require any derivation of the conjugate function. In Tables EC.1 and EC.2, we provide the explicit conic representation for many canonical examples of the conjugate functions $(-h)^*$ and ϕ^*, where the representations are composed of a combination of standard cones such as the quadratic, power, and exponential. These explicit conic representations are useful for the implementation of these constraints in standard optimization software such as CVXPY. Details on the derivations of the epigraph representations can be found in Electronic Companion EC.3.

The following two lemmas originate from Ben-Tal and Nemirovski (2019) (Propositions 2.3.2

and 2.3.4), where they were stated only for the quadratic cones. However, they can be extended to general cones.

Lemma 1. *If f is conic representable with a cone $\mathbf{K}$, i.e., there exist $(\mathbf{A}, \mathbf{v}, \mathbf{B}, \mathbf{b})$ such that*

$$\mathrm{Epi}(f) = \{(\mathbf{x}, t) : \exists\ \mathbf{w} : \mathbf{A}\mathbf{x} + t\mathbf{v} + \mathbf{B}\mathbf{w} + \mathbf{b} \succeq_{\mathbf{K}} \mathbf{0}\},$$

and $\mathbf{A}\mathbf{x} + t\mathbf{v} + \mathbf{B}\mathbf{w} + \mathbf{b} \succ_{\mathbf{K}} \mathbf{0}$ for some $(\mathbf{x}, t, \mathbf{w})$, then f^* is conic representable with the dual cone $\mathbf{K}^*$:

$$\mathrm{Epi}(f^*) = \{(\mathbf{y}, s) : \exists\ \boldsymbol{\xi} \in \mathbf{K}^* : \mathbf{A}^T\boldsymbol{\xi} = -\mathbf{y}, \mathbf{B}^T\boldsymbol{\xi} = \mathbf{0}, \mathbf{v}^T\boldsymbol{\xi} = 1, s \geq \mathbf{b}^T\boldsymbol{\xi}\}.$$

In particular, f and f^* are representable by the same cone if $\mathbf{K}$ is self-adjoint.

Lemma 2. *If f is conic representable with a cone $\mathbf{K}$, then so is its perspective.*

We provide an example illustrating how the ideas in the proofs of Lemmas 1 and 2 can be systematically applied to determine the conic representation of the epigraph of the perspective transformation of a distortion function.

Example 1. Consider $h(p) = p^r(1 - \log(p^r)),\ \ 0 < r < 1$. Instead of deriving a closed-form expression of $(-h)^*$, which can be challenging in specific cases, we determine a conic representation of the epigraph $\lambda(-h)^*\left(\frac{-\nu}{\lambda}\right) \leq t$ by first determining a conic representation of $\mathrm{Epi}(-h)$. We have

$$(p, t) \in \mathrm{Epi}(-h) \Leftrightarrow \exists\ w \geq 0 : -w + w\log(w) \leq t, w \leq p^r, w \leq 1,$$

due to the fact that $x \mapsto -x + x\log(x)$ is decreasing on $[0, 1]$. Following the proof of Lemma 1, we have that $(y, s) \in \mathrm{Epi}((-h)^*)$ if and only if

$$-s \leq \min_{p, w \geq 0, t \in \mathbb{R}} \{-yp + t \mid -w + w\log(w) \leq t, w \leq p^r, w \leq 1\}.$$

Since this is a bounded convex problem satisfying Slater's condition, the duality theorem implies that the minimization problem is equal to

$$\begin{aligned}
&\max_{\xi_1, \xi_2, \xi_3 \geq 0} -\xi_3 + \inf_{p, w \geq 0, t \in \mathbb{R}} \{-yp - \xi_2 p^r + (\xi_2 + \xi_3 - \xi_1)w + \xi_1 w\log(w) + (1 - \xi_1)t\} \\
&= \max_{\xi_2, \xi_3 \geq 0} -\xi_3 + \inf_{p, w \geq 0} \{-yp - \xi_2 p^r + (\xi_2 + \xi_3 - 1)w + w\log(w)\} \\
&= \max_{\xi_2, \xi_3 \geq 0} \{-\xi_3 - e^{-\xi_2 - \xi_3} + \left(r^{\frac{1}{1-r}} - r^{\frac{r}{1-r}}\right) \xi_2^{\frac{1}{1-r}} |y|^{1 - \frac{1}{1-r}} \mid y \leq 0\},
\end{aligned}$$

where in the last equality we computed $\inf_{p \geq 0}\{-yp - \xi_2 p^r\}$ explicitly.[5] This gives that $(y, s) \in$

[5] $\inf_{p \geq 0} -yp - \xi_2 p^r = \begin{cases} \left(r^{\frac{1}{1-r}} - r^{\frac{r}{1-r}}\right) \xi_2^{\frac{1}{1-r}} |y|^{1 - \frac{1}{1-r}} & y < 0, \quad \xi_2 > 0 \\ 0 & y \leq 0, \quad \xi_2 = 0 \\ -\infty & y \geq 0, \quad \xi_2 > 0. \end{cases}$

$\text{Epi}((-h)^*)$ if and only if there exist $\xi_2, \xi_3, \xi_4 \geq 0$ such that

$$\xi_3 + e^{-\xi_2 - \xi_3} + \left(r^{\frac{r}{1-r}} - r^{\frac{1}{1-r}}\right)\xi_4 \leq s, \quad \xi_2 \leq |y|^r \xi_4^{1-r}, \quad y \leq 0.$$

Lemma 2 then yields that $\lambda(-h)^*\left(\frac{-\nu}{\lambda}\right) \leq z$ if and only if there exist $\xi_2, \xi_3, \xi_4 \geq 0$ such that

$$\xi_3 + \lambda e^{-(\xi_2+\xi_3)/\lambda} + \left(r^{\frac{r}{1-r}} - r^{\frac{1}{1-r}}\right)\xi_4 \leq z, \quad \xi_2 \leq |\nu|^r \xi_4^{1-r}, \quad \nu \geq 0,$$

which is a combination of the power cone and the exponential cone.

In some cases, one can also calculate a conjugate function by writing it as an inf-convolution. For example, if f is a sum of individual f_i's: $f(\mathbf{x}) = \sum_{i=1}^n f_i(\mathbf{x})$, then the conjugate of a sum is the inf-convolution of the sum of the conjugates (see Theorem 2.1, Bertsimas and den Hertog, 2022):

$$\left(\sum_{i=1}^n f_i\right)^*(\mathbf{s}) = \inf_{\{\mathbf{v}^i\}_{i\in[n]}} \left\{\sum_{i=1}^n f_i^*(\mathbf{v}^i) \;\middle|\; \sum_{i=1}^n \mathbf{v}^i = \mathbf{s}\right\},$$

where the infimum can eventually be omitted when considering the epigraph formulation.

4 Solving the Robust Problem I: A Cutting-Plane Method

As shown in Theorem 2, the number of reformulated constraints grows exponentially, as 2^m, where m is the dimension of the probability vector. If m is of small or moderate size, then the reformulations in (16) and (17) can be applied to solve the minimization problems (P) and (P-Nom) exactly. However, for larger m, this becomes computationally intractable. In this section, we show how the reformulation of (P) to (P-ref) (and similarly for (P-Nom) to (P-Nom-ref)) by Theorem 1 enables us to devise a suitable cutting-plane method, which circumvents this curse of dimensionality.

4.1 The Cutting-Plane Algorithm

The standard cutting-plane algorithm is one of the most applied algorithms to solve robust optimization problems and has shown great efficiency in many applications (e.g., Mutapcic and Boyd, 2009, Bertsimas et al., 2016). The general idea is simple: Approximate the uncertainty set $\mathcal{U}$ with a suitably chosen subset $\mathcal{U}_j \subset \mathcal{U}$ and solve the corresponding robust problem, which is often a simpler problem, similar to the nominal problem. If the solution is feasible according to the robust evaluation with respect to the original set $\mathcal{U}$, then the process is terminated. Otherwise, the worst-case parameter in $\mathcal{U}$ associated with the current solution is added to $\mathcal{U}_j$, and the process is repeated.

In our case, we apply the cutting-plane method to the reformulated problem (P-ref), where the uncertainty set is given by the *composite uncertainty set* $\mathcal{U}_{\phi,h}(\mathbf{p})$. After obtaining a candidate solution using a suitable subset $\mathcal{U}_j$, we verify its feasibility with respect to the original uncertainty set $\mathcal{U}_{\phi,h}(\mathbf{p})$. This involves solving the following optimization problem, for a given solution $\mathbf{a}_*$:

$$\sup_{(\mathbf{q},\bar{\mathbf{q}})\in\mathcal{U}_{\phi,h}(\mathbf{p})} \sum_{i=1}^{m} -\bar{q}_i u(f(\mathbf{a}_*, \mathbf{x}_i)).$$

At first sight, this seems problematic due to the 2^m number of constraints in $\mathcal{U}_{\phi,h}(\mathbf{p})$. Fortunately, this can be avoided, since the equivalent robust rank-dependent evaluation is the optimization problem:

$$\sup_{\mathbf{q}\in\mathcal{D}_\phi(\mathbf{p},r)} \rho_{u,h,\mathbf{q}}(f(\mathbf{a}_*, \mathbf{X})). \tag{18}$$

Problem (18) can be solved efficiently. Indeed, for a given solution $\mathbf{a}_*$, we can assess the ranking of the outcomes: $u(f(\mathbf{a}_j, \mathbf{x}_{(1)})) \geq \ldots \geq u(f(\mathbf{a}_j, \mathbf{x}_{(m)}))$. Then, by rewriting the alternating sum in (5), problem (18) becomes equal to the following convex optimization problem:

$$\sup_{\mathbf{q}\in\mathcal{D}_\phi(\mathbf{p},r)} \sum_{k=1}^{m} h\left(\sum_{j=k}^{m} q_{(j)}\right)\left(u(f(\mathbf{a}_*, \mathbf{x}_{(j-1)})) - u(f(\mathbf{a}_*, \mathbf{x}_{(j)}))\right), \tag{19}$$

where $u(f(\mathbf{a}_*, \mathbf{x}_{(0)})) := 0$. The optimal solution of (19) is a probability vector $\mathbf{q}^* \in \mathcal{D}_\phi(\mathbf{p}, r)$. This probability vector further yields a $\bar{\mathbf{q}}^*$ such that $\bar{q}^*_{(i)} = h\left(\sum_{k=i}^{m} q^*_{(k)}\right) - h\left(\sum_{k=i+1}^{m} q^*_{(k)}\right)$ and $\bar{q}^*_{(m)} = h(q^*_{(m)})$, which is precisely the probability vector that constitutes the rank-dependent sum (5). The following lemma justifies that we may add the probability vectors $(\mathbf{q}^*, \bar{\mathbf{q}}^*)$ as the worst-case probabilities at each iteration of the cutting-plane procedure.

Lemma 3. *We have that* $(\mathbf{q}^*, \bar{\mathbf{q}}^*) \in \mathcal{U}_{\phi,h}(\mathbf{p})$.

We can now describe the cutting-plane procedure in full detail, which we do more precisely in Algorithm 1. We note that the cutting-plane algorithm always computes a *lower* bound c_j on the optimal objective value of (P-ref), since solving (20) at each iteration j with respect to a subset $\mathcal{U}_j \subset \mathcal{U}_{\phi,h}(\mathbf{p})$ is a less conservative problem. Moreover, the lower bound is improved iteratively, since $\mathcal{U}_j \subset \mathcal{U}_{j+1}$. Furthermore, by solving (19), which is a robust rank-dependent evaluation of a particular feasible solution, we obtain an *upper* bound v_j on (P-ref) at each iteration. Hence, the final step of the cutting-plane algorithm yields an upper and a lower bound on the exact optimal objective value (P-ref), with a gap value $v_j - c_j \leq \epsilon_{\text{tol}} > 0$ that can be chosen arbitrarily small.

The natural, formal question that arises is whether the cutting-plane algorithm actually terminates after finitely many iterations, thus yielding convergent upper and lower bounds as $\epsilon_{\text{tol}} \to 0$. The following theorem, which has its roots in Mutapcic and Boyd (2009), states that, for any $\epsilon_{\text{tol}} > 0$, the termination is indeed guaranteed. We note that Mutapcic and Boyd (2009) imposes Lipschitz continuity conditions, which we avoid in our setting.

Algorithm 1 Cutting-Plane Method

1: Start with $\mathcal{U}_1 = \{(\mathbf{p}, \mathbf{p})\}$. Fix a tolerance parameter $\epsilon_{\text{tol}} > 0$.

2: At the j-th iteration, solve the following problem with the uncertainty set $\mathcal{U}_j$:

$$\min_{\mathbf{a}\in\mathcal{A}} \sup_{(\mathbf{q},\bar{\mathbf{q}})\in\mathcal{U}_j} -\sum_{i=1}^{m} \bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i)). \tag{20}$$

3: Let $(\mathbf{a}_j, c_j)$ be the optimal solution and objective value of (20). Determine a ranking of the realizations:

$$-u(f(\mathbf{a}_j, \mathbf{x}_{(1)})) \leq \ldots \leq -u(f(\mathbf{a}_j, \mathbf{x}_{(m)})).$$

Then, solve the optimization problem (19), which gives an optimal objective value v_j and a solution $(\mathbf{q}_j^*, \bar{\mathbf{q}}_j^*)$.

4: If $v_j - c_j \leq \epsilon_{\text{tol}}$, then the solution is accepted and the process is terminated.

5: If not, set $\mathcal{U}_{j+1} = \mathcal{U}_j \cup \{(\mathbf{q}_j^*, \bar{\mathbf{q}}_j^*)\}$ and repeat steps 2–5.

Theorem 3. *Let h be a concave distortion function. Suppose $\sup_{\mathbf{a}\in\mathcal{A},i\in[m]} u(f(\mathbf{a}, \mathbf{x}_i)) < \infty$. Then, for all $\epsilon_{\text{tol}} > 0$, Algorithm 1 terminates after finitely many iterations.*

Remark 3. Algorithm 1 also applies to the nominal case (P-Nom), where $r \equiv 0$ in the set $\mathcal{U}_{\phi,h}(\mathbf{p})$. Furthermore, Theorem 3 still holds in this case, since its proof does not depend on the choice of r.

4.2 Robust Optimization with a Rank-Dependent Evaluation in the Constraint

In the previous subsection, we have shown that our cutting-plane algorithm yields convergent upper and lower bounds to problems (P-Nom)–(P), where the rank-dependent evaluation appears in the objective function. In this subsection, we discuss how the cutting-plane algorithm can be adapted to the robust problem (P-constraint), where the RDU model appears in the constraint, to provide a convergent lower bound. The adaptation of Algorithm 1 to problem (P-constraint) is presented in full detail in Algorithm EC.1 in Electronic Companion EC.7.[6] The following theorem establishes the convergence of the cutting-plane algorithm.

Theorem 4. *Let h be concave. Suppose $\sup_{\mathbf{a}\in\mathcal{A},i\in[m]} u(f(\mathbf{a}, \mathbf{x}_i)) < \infty$. Then, for all $\epsilon_{\text{tol}} > 0$, Algorithm EC.1 terminates after finitely many iterations. Moreover, if the function $\mathbf{a} \mapsto u(f(\mathbf{a}, \mathbf{x}))$ is concave for all $\mathbf{x} \in \mathbb{R}^l$, then the optimal objective value of the final solution obtained from Algorithm EC.1 converges to that of* (P-constraint)*, as $\epsilon_{\text{tol}} \to 0$.*

Thus, using the cutting-plane algorithm, we can obtain lower bounds that converge to the exact optimal objective value of (P-constraint), as $\epsilon_{\text{tol}} \to 0$. Naturally, one would also like to obtain an upper bound on (P-constraint), as well as a feasible solution of (P-constraint). We note that this is not guaranteed by the cutting-plane algorithm as described in Algorithm EC.1, since it only

[6] In a nutshell, we replace (20) in Algorithm 1 by (P-constraint) adapted to the set $\mathcal{U}_j$ and replace c_j in step 4 by the constraint parameter c.

provides a solution that is feasible within a tolerance $\epsilon_{\text{tol}} > 0$. With this objective in mind, we explore the rank-dependent nature of $\rho_{u,h,\mathbf{q}}(.)$ and propose a method to obtain an upper bound and a feasible solution of (P-constraint), which requires solving an optimization problem with only $3m+3$ number of constraints. We first state the following definition.

Definition 2. Given an $\mathbf{a} \in \mathcal{A}$, we let $\mathcal{I}(\mathbf{a})$ denote the set of all permutations $(i_1, \ldots, i_m)$ of the index vector $(1, \ldots, m)$ such that the ranking $-u(f(\mathbf{a}, \mathbf{x}_{i_1})) \le \ldots \le -u(f(\mathbf{a}, \mathbf{x}_{i_m}))$ holds.

The idea is that for any given solution $\mathbf{a}_0 \in \mathcal{A}$, we can determine a ranking of the realizations $\{-u(f(\mathbf{a}_0, \mathbf{x}_i))\}_{i=1}^m$ and obtain a vector of indices $(i_1, \ldots, i_m) \in \mathcal{I}(\mathbf{a}_0)$. A natural upper bound on (P-constraint) can then be computed by solving the following optimization problem induced by the ranking $(i_1, \ldots, i_m)$ of $\mathbf{a}_0$:

$$U^*(\mathbf{a}_0) \triangleq \min_{\mathbf{a} \in \mathcal{A}} \left\{ g(\mathbf{a}) \;\middle|\; \sup_{(\mathbf{q}, \bar{\mathbf{q}}) \in \mathcal{U}_{\phi,h}^{(i_1,\ldots,i_m)}(\mathbf{p})} -\sum_{i=1}^m \bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i)) \le c \right\}, \tag{21}$$

where the rank-dependent uncertainty set is

$$\mathcal{U}_{\phi,h}^{(i_1,\ldots,i_m)}(\mathbf{p}) \triangleq \left\{ (\mathbf{q}, \bar{\mathbf{q}}) \in \mathbb{R}^{2m} \;\middle|\; \begin{array}{l} \sum_{i=1}^m q_i = \sum_{i=1}^m \bar{q}_i = 1 \\ \sum_{i=1}^m p_i \phi\left(\frac{q_i}{p_i}\right) \le r \\ \sum_{j=k}^m \bar{q}_{i_j} \le h\left(\sum_{j=k}^m q_{i_j}\right), \ \forall k \in [m] \\ q_i, \bar{q}_i \ge 0, \ \forall i \in [m] \end{array} \right\}.$$

Indeed, since $\mathcal{U}_{\phi,h}(\mathbf{p}) \subset \mathcal{U}_{\phi,h}^{(i_1,\ldots,i_m)}(\mathbf{p})$, solving (21) (which can be done using Theorem 2) yields an upper bound and a feasible solution of (P-constraint). In particular, if $\mathbf{a}_0$ is a good approximation of the optimal solution of (P-constraint), such that the ranking coincides, then $U^*(\mathbf{a}_0)$ is exactly equal to the optimal objective value of (P-constraint). This observation is made precise in the following lemma and is pivotal for developing a convergent upper bound for (P-constraint).

Lemma 4. *Let $\mathbf{a}^*$ be a minimizer of* (P-constraint). *Then, for any $\mathbf{a}_0 \in \mathcal{A}$ such that $\mathcal{I}(\mathbf{a}_0) \subset \mathcal{I}(\mathbf{a}^*)$, we have that the value $U^*(\mathbf{a}_0)$ as defined in* (21) *is equal to the optimal objective value of* (P-constraint).

Lemma 4 suggests that if one approximates the optimal solution of (P-constraint) with a sequence of cutting-plane solutions as $\epsilon_{\text{tol}} \to 0$, then under certain continuity conditions, the upper bounds computed in (21) also converge to the exact value. This is established in the following theorem.

Theorem 5. *Let h be concave. Assume $\sup_{\mathbf{a} \in \mathcal{A}, i \in [m]} u(f(\mathbf{a}, \mathbf{x}_i)) < \infty$. If the functions $g(\mathbf{a})$ and $u(f(\mathbf{a}, \mathbf{x}_i))$ are continuous on $\mathcal{A}$, for all $i \in [m]$, and $\mathbf{a}_n$ is a sequence of solutions obtained from Algorithm EC.1 with $\epsilon_{\text{tol},n} \to 0$, then $(U^*(\mathbf{a}_n))_{n \ge 1}$ is a sequence of upper bounds that converges to the optimal objective value of* (P-constraint).

4.3 Choquet Expected Utility

In this subsection, we briefly discuss how our cutting-plane method can also be extended to solve optimization problems for a more general rank-dependent model, namely the Choquet expected utility model (Schmeidler, 1986, 1989). Let $c : \mathcal{F} \to [0,1]$ be a monotone set function[7] such that $c(\emptyset) = 0,\ c(\Omega) = 1$. The Choquet expected utility model evaluates a random variable X by the Choquet integral:

$$\int -X \,\mathrm{d}c \triangleq \int_0^\infty c(-X > t)\,\mathrm{d}t + \int_{-\infty}^0 (c(-X > t) - 1)\,\mathrm{d}t. \tag{22}$$

The rank-dependent evaluation $\rho_{u,h,\mathbb{Q}}$ as defined in (4) is a special case of the Choquet expected utility, where $c(A) = h(\mathbb{Q}(A))$. If c is a submodular set function,[8] then the Choquet expected utility can also be written as a worst-case expectation (see Denneberg, 1994), similar to Theorem 1:

$$\int -X\,\mathrm{d}c = \sup_{\bar{\mathbf{q}} \in \mathcal{U}_c} -\sum_{i=1}^m \bar{q}_i x_i, \quad \mathcal{U}_c \triangleq \left\{ \bar{\mathbf{q}} \in \mathbb{R}^m \;\middle|\; \begin{array}{l} \sum_{i=1}^m \bar{q}_i = 1 \\ \sum_{i\in J} \bar{q}_i \le c(J),\ \forall J \subset [m] \\ \bar{q}_i \ge 0,\ \forall i \in [m] \end{array} \right\}. \tag{23}$$

The uncertainty set $\mathcal{U}_c$ has essentially the same structure as the set $\mathcal{U}_{\phi,h}(\mathbf{p})$ defined in (13) and consists of only linear constraints. Therefore, the same procedure of the cutting-plane algorithm can be extended naturally to solve the more general rank-dependent minimization problem $\min_{\mathbf{a}\in\mathcal{A}} \int -f(\mathbf{a},\mathbf{X})\,\mathrm{d}c$. By contrast, the piecewise-linear approximation methods, which we will discuss in Section 5, constitute a more restrictive, but as we will explicate in some cases more efficient, approach.

5 Solving the Robust Problem II: Piecewise-Linear Approximation

In this section, we develop a different method for computing lower and upper bounds on the optimal objective value of (P-ref), which relies on a suitable approximation of the distortion functions, and is leveraged in Section 6.

5.1 Piecewise-Linear Distortion Functions

In this subsection, we show that if the distortion function h is concave and piecewise-linear, then the exponential complexity of the constraints in $\mathcal{U}_{\phi,h}(\mathbf{p})$ can be reduced to the order of $m \cdot K$, where m is the dimension of the probability vector, and K is the number of linear pieces of h.

[7] $c(A) \le c(B)$, if $A \subset B$.
[8] $c(A\cap B) + c(A \cup B) \le c(A) + c(B),\ \forall A, B \in \mathcal{F}$.

More precisely, we consider a function $h = \min_{j\in[K]} h_j$, with affine functions $h_j(p) = l_j \cdot p + b_j$, such that the slopes $l_1 > \cdots > l_K$ are decreasing and the intercepts $b_1 < \cdots < b_K$ are increasing. The affine functions are assumed to be defined on a set of support points $0 = x_0 < x_1 < \cdots < x_K = 1$, such that $h(p) = h_j(p)$, if $p \in [x_{j-1}, x_j]$, for all $j \in [K]$. Furthermore, we impose $b_1 \equiv 0$ and $l_K + b_K \equiv 1$, so that $h(0) = 0$ and $h(1) = 1$. We refer to this type of function as a *K-piecewise-linear distortion function.* Clearly, this function is non-decreasing and concave.

We prove the following lemma.

Lemma 5. *Let h be a K-piecewise-linear distortion function. Then, $(\mathbf{q}, \bar{\mathbf{q}}) \in \mathcal{U}_{\phi,h}(\mathbf{p})$ if and only if there exists $\mathbf{t} \triangleq (\{t_{i,j}\}_{i=1}^m)_{j=1}^K \in \mathbb{R}^{mK}$ such that the variables $(\mathbf{q}, \bar{\mathbf{q}}, \mathbf{t})$ satisfy the constraints*

$$\begin{cases} q_i, \bar{q}_i, t_{i,j} \geq 0, \ \forall i \in [m], \ \forall j \in [K] \\ \bar{q}_i \leq l_j \cdot q_i + t_{i,j}, \ \forall i \in [m], \ \forall j \in [K] \\ \sum_{i=1}^m t_{i,j} \leq b_j, \ \forall j \in [K] \\ \sum_{i=1}^m q_i = \sum_{i=1}^m \bar{q}_i = 1 \\ \sum_{i=1}^m p_i \phi\left(\frac{q_i}{p_i}\right) \leq r. \end{cases} \tag{24}$$

Hence, for K-piecewise-linear distortion functions, we obtain the following reformulation of the robust counterpart (12).

Theorem 6. *Let h be a K-piecewise-linear distortion function. Then, we have that $(\mathbf{a}, c)$ satisfies the constraint* (12) *if and only if there exist variables $\alpha, \beta, \gamma, \{\lambda_{ij}\}_{i\in[m],j\in[K]}, \{\nu_j\}_{j=1}^K \in \mathbb{R}$ such that*

$$\begin{cases} \alpha + \beta + \gamma r + \sum_{j=1}^K \nu_j b_j + \sum_{i=1}^m p_i \gamma \phi^* \left(\frac{-\alpha + \sum_{j=1}^K \lambda_{ij} l_j}{\gamma} \right) \leq c \\ -u(f(\mathbf{a}, \mathbf{x}_i)) - \beta - \sum_{j=1}^K \lambda_{ij} \leq 0, \ \forall i \in [m] \\ \lambda_{ij} \leq \nu_j, \ \forall i \in [m], \ \forall j \in [K] \\ \gamma, \lambda_{ij}, \nu_j \geq 0, \ \forall i \in [m], \ \forall j \in [K]. \end{cases} \tag{25}$$

In the nominal case, we have that (14) *holds if and only if there exist variables $\beta, \lambda_{ij}, \nu_j \in \mathbb{R}$ such that* (25) *holds with $\alpha = \gamma \equiv 0$, whence the first line becomes $\beta + \sum_{j=1}^K \nu_j b_j + \sum_{i=1}^m \sum_{j=1}^K \lambda_{ij} l_j p_i \leq c$.*

Thus, for K-piecewise-linear distortion functions, the exponential complexity in the uncertainty set $\mathcal{U}_{\phi,h}(\mathbf{p})$ can be reduced. This result also provides a tractable approximation for general non-piecewise-linear concave distortion functions, as we outline in the following subsection, exploiting that each concave function can be approximated uniformly by a piecewise-linear concave function.

5.2 Piecewise-Linear Approximation of General Concave Distortion Functions

For a general concave h, we can reduce the complexity in $\mathcal{U}_{\phi,h}(\mathbf{p})$ by approximating h with a piecewise-linear function h_ϵ, such that $\max_{x\in[0,1]} |h(x) - h_\epsilon(x)| \leq \epsilon$ for any given $\epsilon > 0$. An upper approximation $h_\epsilon \geq h$ and a lower approximation $h_\epsilon \leq h$ yield an upper and lower bound on the optimal objective value of (P-ref), respectively. Uniform approximation of general concave functions with piecewise-linear functions has been studied in the literature (see, e.g., Imamoto and Tang, 2008, Cox, 1971).

More specifically, using the concavity of h, one can approximate h from below by choosing a set of support points $\{x_i\}_{i=0}^K$ and considering the concave piecewise-linear function that connects the values $\{h(x_i)\}_i$. Given an $\epsilon > 0$, the required number of support points K can be minimized. We briefly describe our approach of determining the minimal K support points $\{x_i\}_{i=0}^K$ such that h is lower approximated by a piecewise-linear function with a prescribed error ϵ. Set $x_0 = 0$. At the i-th iteration, we choose the next support point x_{i+1} such that the maximal error is equal to ϵ:

$$e_i(x_{i+1}) \triangleq \sup_{x \in [x_i, x_{i+1}]} \left\{ h(x) - \frac{h(x_{i+1}) - h(x_i)}{x_{i+1} - x_i}(x - x_i) - h(x_i) \right\} = \epsilon. \tag{26}$$

If $e_i(1) \leq \epsilon$, then we simply choose $x_{i+1} = 1$. Otherwise, we choose x_{i+1} such that (26) holds. By construction, the piecewise-linear approximation h_ϵ induced by this set of support points has a maximum approximation error of ϵ and satisfies $h_\epsilon \leq h$, due to concavity. The existence of such x_{i+1} at each iteration is verified in the following lemma. Moreover, it implies that we can solve (26) for x_{i+1} using the bisection method.

Lemma 6. *Let h be a continuous, increasing, strictly concave function on $[0, 1]$. Then, $e_i(x_{i+1})$ is an increasing and continuous function in x_{i+1}. In particular, for any $\epsilon > 0$ and iteration i, if $e_i(1) > \epsilon$, then there exists a $x_{i+1} \in (x_i, 1)$ such that $e_i(x_{i+1}) = \epsilon$.*

5.3 The Piecewise-Linear Approximation Method and its Convergence

Given a piecewise-linear lower approximation $h_\epsilon \leq h$ of h, we can approximate the uncertainty set $\mathcal{U}_{\phi,h}(\mathbf{p})$ with $\mathcal{U}_{\phi,h_\epsilon}$ and apply Theorem 6. This yields a lower bound on (P-ref). Similarly, the concave distortion function $\tilde{h}_\epsilon$, which we define as

$$\tilde{h}_\epsilon(p) \triangleq \begin{cases} 0, & p = 0 \\ \min\{h_\epsilon(p) + \epsilon, 1\}, & 0 < p \leq 1, \end{cases} \tag{27}$$

yields an upper approximation of h with uniform error ϵ. Hence, we can also approximate $\mathcal{U}_{\phi,h}(\mathbf{p})$ by $\mathcal{U}_{\phi,\tilde{h}_\epsilon}(\mathbf{p})$ and apply Theorem 6 (note that the constraints (25) for $\tilde{h}_\epsilon$ are the same as those for h_ϵ, but with b_j replaced by $b_j + \epsilon$). This yields an upper bound for (P-ref) since $\mathcal{U}_{\phi,h}(\mathbf{p}) \subset \mathcal{U}_{\phi,\tilde{h}_\epsilon}(\mathbf{p})$. Therefore, we obtain an upper and lower bound using the piecewise-linear approximation method that we summarize in Algorithm 2. We also show that both bounds converge to the optimal objective value of the exact problem (P-ref), if the approximation error ϵ approaches zero. Thus, Algorithm 2 terminates for any parameter $\delta > 0$. This also applies, *mutatis mutandis*, to problem (P-constraint).

Algorithm 2 **Piecewise-Linear Approximation Method**

1: Set an $\epsilon > 0$. Approximate h from below by a K-piecewise-linear distortion function h_ϵ such that: (1) The number of pieces, K, is minimized; (2) $h_\epsilon(x) \leq h(x)$, $\forall x \in [0,1]$; (3) $\sup_{x\in[0,1]} h(x) - h_\epsilon(x) \leq \epsilon$.
2: Solve problem (P-ref) with $\mathcal{U}_{\phi,h}(\mathbf{p})$ replaced by $\mathcal{U}_{\phi,h_\epsilon}(\mathbf{p})$, using Theorem 6. This gives a lower bound L^* and an optimal solution $\mathbf{a}^*$. Do the same with $\tilde{h}_\epsilon$ to obtain an upper bound U^*.
3: If $U^* - L^* < \delta$ for some prescribed $\delta > 0$, then we take $\mathbf{a}^*$ as the final solution.
4: Otherwise, set $\epsilon \to \epsilon/2$ and perform all previous steps to obtain new bounds and solutions.

Theorem 7. *Let h be concave. Suppose $\sup_{\mathbf{a}\in\mathcal{A},i\in[m]} u(f(\mathbf{a},\mathbf{x}_i)) < \infty$. Then, Algorithm 2 terminates after finitely many iterations, for any $\epsilon, \delta > 0$. If the function $\mathbf{a} \mapsto u(f(\mathbf{a},\mathbf{x}))$ is concave for all $\mathbf{x} \in \mathbb{R}^l$, then this also holds when Algorithm 2 is applied to problem* (P-constraint).

6 Non-Concave Distortion Functions

In the previous sections, we have discussed how to optimize a rank-dependent model when the distortion function is concave. We now extend these ideas to convex and inverse S-shaped distortion functions, which are often found in empirical work (see e.g., Wakker, 2010, Prelec, 1998). Optimization problems with non-concave distortion functions are challenging due to their non-convex nature. Moreover, rank-dependent models with non-concave distortion functions lack a dual representation, which was a pivotal tool that enabled us to reformulate the rank-dependent problem. In the recent literature (e.g., Cai et al., 2025 and Pesenti et al., 2020), uncertainty sets with a specific structure have been identified for which optimizing the robust distortion risk measure with non-concave distortion function is equivalent to optimizing the same model with h replaced by its concave envelope $\hat{h}$.[9] However, this equivalence is typically not satisfied in our setting, as we state in the following proposition.

Proposition 1. *Let $\mathcal{U}$ be a compact set of probability vectors. Enumerate the realizations of X as $x_1 > \ldots > x_m$ and suppose that $\mathcal{U}$ does not contain the probability vector for which $q_1 = 1$, which is concentrated on x_1. Then, there exists a continuous distortion function h such that*

$$\sup_{\mathbf{q}\in\mathcal{U}} \rho_{h,\mathbf{q}}(X) < \sup_{\mathbf{q}\in\mathcal{U}} \rho_{\hat{h},\mathbf{q}}(X).$$

Therefore, it is necessary to study how to reformulate a robust or nominal rank-dependent evaluation (constraint), in the case of a convex or inverse S-shaped distortion function. We will first study how to reformulate a nominal constraint of the form (P-Nom-EG).

[9]The concave envelope of a function h is the smallest concave function that dominates h.

6.1 The Nominal Problem

Although rank-dependent models do not admit a dual representation for general non-concave distortion functions, we can still express them as inner robust optimization problems if the distortion function is convex or inverse S-shaped. The idea is to treat the concave and convex parts of the inverse S-shaped function separately, where the convex part is transformed into a concave part via the dual function $\bar{h}(p) \triangleq 1 - h(1-p)$. This result is stated in the following theorem. Henceforth, the utility function is allowed to be non-concave.

Theorem 8. *Let $h : [0,1] \to [0,1]$ be an inverse S-shaped distortion function with $p^0 \in [0,1]$ as in Definition 1. Define the dual function $\bar{h}(p) \triangleq 1 - h(1-p)$. Then, we have that, for all $(\mathbf{a}, c) \in \mathbb{R}^{n_a+1}$,* (P-Nom-EG) *is satisfied if and only if there exist $z \in \mathbb{R}, \bar{\mathbf{q}} \in N_{\bar{h}}^{cv}(\mathbf{p})$ such that*

$$\begin{cases} z + \sum_{i=1}^{m} -\bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i)) \leq c \\ \sup_{\mathbf{q} \in M_h^{ca}(\mathbf{p})} \sum_{i=1}^{m} -q_i u(f(\mathbf{a}, \mathbf{x}_i)) \leq z, \end{cases} \tag{28}$$

where

$$M_h^{ca}(\mathbf{p}) \triangleq \left\{ \mathbf{q} \in \mathbb{R}^m \;\middle|\; \begin{array}{l} q_i \geq 0 \\ \sum_{i \in J} q_i \leq h\left(\sum_{i \in J} p_i\right), \ \forall J \subset [m] : \sum_{i \in J} p_i \leq p^0 \\ \sum_{i=1}^{m} q_i = h(p^0) \end{array} \right\}, \tag{29}$$

and

$$N_{\bar{h}}^{cv}(\mathbf{p}) \triangleq \left\{ \bar{\mathbf{q}} \in \mathbb{R}^m \;\middle|\; \begin{array}{l} \bar{q}_i \geq 0 \\ \sum_{i \in J} \bar{q}_i \leq \bar{h}\left(\sum_{i \in J} p_j\right), \ \forall J \subset [m] : \sum_{i \in J} p_i \leq 1 - p^0 \\ \sum_{i=1}^{m} \bar{q}_i = \bar{h}(1-p^0) \end{array} \right\}. \tag{30}$$

Remark 4. In particular, if h is convex, we have that

$$\rho_{u,h,\mathbf{p}}(f(\mathbf{a}, \mathbf{X})) = - \sup_{\bar{\mathbf{q}} \in N_{\bar{h}}^{\text{cv}}(\mathbf{p})} \sum_{i=1}^{m} \bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i)), \tag{31}$$

with $p^0 \equiv 0$.

Theorem 8 shows that (P-Nom-EG) can be reformulated into rank-independent constraints (28), where the supremum constraint can be reformulated further as in Theorem 2. However, both (29) and (30) still feature an exponential number of constraints. As outlined in the previous section, we can circumvent this using piecewise-linear approximation. Let h be a piecewise-linear, inverse S-shaped distortion function with its concave and convex parts specified by the concave functions $h_l, \bar{h}_l$ (note that $\bar{h}_l$ is the dual of the convex part of h), respectively, on the domains $[0, p^0]$ and

$[0, 1-p^0]$. Due to concavity and monotonicity, they can be expressed as minima of affine functions:

$$h_l(p) = \min\{l_1^{(1)}p + b_1^{(1)}, \ldots, l_{K_1}^{(1)}p + b_{K_1}^{(1)}\}, \quad \bar{h}_l(p) = \min\{l_1^{(2)}p + b_1^{(2)}, \ldots, l_{K_2}^{(2)}p + b_{K_2}^{(2)}\}, \tag{32}$$

defined on the support points

$$0 = s_0^{(1)} \leq s_1^{(1)}, \ldots, s_{K_1-1}^{(1)} \leq s_{K_1}^{(1)} = p^0, \quad 0 = s_0^{(2)} \leq s_1^{(2)}, \ldots, s_{K_2-1}^{(2)} \leq s_{K_2}^{(2)} = 1 - p^0,$$

such that, for all $k \in \{1, \ldots, K_1\}, v \in \{1, \ldots, K_2\}$,

$$h_l(p) = l_k^{(1)}p + b_k^{(1)}, \forall p \in [s_{k-1}^{(1)}, s_k^{(1)}],\ \bar{h}_l(p) = l_v^{(2)}p + b_v^{(2)}, \forall p \in [s_{v-1}^{(2)}, s_v^{(2)}]. \tag{33}$$

The following theorem establishes a reformulation of the constraints in (28), when h is piecewise-linear.

Theorem 9. *Let h be a piecewise-linear, inverse S-shaped distortion function with its concave and convex part specified by $h_l, \bar{h}_l$ as in* (32)*. Then, for any $(\mathbf{a}, c) \in \mathbb{R}^{n_a+1}$,* (P-Nom-EG) *is satisfied if and only if there exist variables $\lambda_{ik}, \nu_k, t_{ik}, \bar{q}_i \geq 0$, $\beta \in \mathbb{R}$ such that*

$$\begin{cases} \beta \cdot h(p^0) + \sum_{k=1}^{K_1} \nu_k b_k^{(1)} + \sum_{i=1}^m \sum_{k=1}^{K_1} \lambda_{ik} l_k^{(1)} p_i - \sum_{i=1}^m \bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i)) \leq c \\ -u(f(\mathbf{a}, \mathbf{x}_i)) - \beta - \sum_{k=1}^{K_1} \lambda_{ik} \leq 0,\ \forall i \in [m] \\ \lambda_{ik} \leq \nu_k,\ \forall i \in [m],\ \forall k \in [K_1] \\ \bar{q}_i \leq l_k^{(2)} p_i + t_{ik},\ \forall i \in [m],\ \forall k \in [K_2] \\ \sum_{i=1}^m t_{ik} \leq b_k^{(2)},\ \forall k \in [K_2] \\ \sum_{i=1}^m \bar{q}_i = \bar{h}(1 - p^0). \end{cases} \tag{34}$$

Importantly, this yields a problem with only $O(m \cdot K_1 \vee K_2)$ constraints. We note that, not surprisingly, (34) contains a non-convex constraint due to the product term $\bar{q}_i u(f(\mathbf{a}, \mathbf{x}_i))$. Hence, the computation of (P-Nom-EG) with the reformulated constraints in (34) requires a mixed-integer nonlinear programming (MINLP) solver, such as BARON that uses the branch and reduce search method to obtain a global optimum (see e.g., Ryoo and Sahinidis, 1996). In particular, if the set $\mathcal{A}$ is polyhedral (i.e., $\mathcal{A} = \{\mathbf{a} \in \mathbb{R}^{n_a} : \mathbf{Da} \leq \mathbf{d}\}$), $f(\mathbf{a}, \mathbf{x}) = \mathbf{a}^T\mathbf{x}$ and u is piecewise-linear, then the nomimal problem (P-Nom-EG) with constraints (34) can also be solved efficiently by Gurobi (Gurobi Optimization, LLC, 2023), using bilinear and special ordered sets of type 2 (SOS2) constraints; see the precise formulation in (EC.30) in Electronic Companion EC.8, further elaboration on the SOS2 constraints in (36), and a numerical study on the performance of Gurobi in this setting in Section 7.4.

6.2 Robust Rank-Dependent Models with Inverse S-Shaped Distortion

In this subsection, we study the robust constraint (P-EG) when h is inverse-S-shaped. We note that then (P-EG) cannot be reformulated using a composite uncertainty set as in (12), since the convex part of the distortion function gives rise to a sup-inf term when combining the set $\mathcal{D}_\phi(\mathbf{p}, r)$ with (30). However, we can still circumvent this using a suitable cutting-plane approach, where we iteratively solve a nominal problem using (34), and compute the robust rank-dependent evaluation for each nominal solution.

Indeed, for any given feasible solution $\mathbf{a}_* \in \mathcal{A}$, we can assess the ranking of the outcomes $u(f(\mathbf{a}^T, \mathbf{x}_{(1)})) \geq \ldots \geq u(f(\mathbf{a}^T, \mathbf{x}_{(m)}))$. Then, we calculate the robust rank-dependent evaluation:

$$\sup_{\mathbf{q} \in \mathcal{D}_\phi(\mathbf{p}, r)} \sum_{i=1}^{m} h\left(\sum_{k=i}^{m} q_{(k)}\right) \left(u(f(\mathbf{a}_*), \mathbf{x}_{(i-1)})) - u(f(\mathbf{a}_*, \mathbf{x}_{(i)}))\right). \tag{35}$$

Problem (35) is, quite naturally, non-convex and can again be computed using a solver such as BARON. If h is piecewise-linear and the divergence function ϕ is linear or quadratic, then one can also solve (35) using SOS2 constraints in Gurobi. Specifically, given a set of support points $(t_j, h(t_j))_{j=0}^K$, where $0 = t_0 < t_1 < \cdots < t_{K-1} < t_K = 1$, the SOS2 constraints can be formulated as follows:

$$\begin{aligned}
\max_{\substack{\mathbf{q} \in \mathcal{D}_\phi(\mathbf{p}, r) \\ \boldsymbol{\lambda} \in \mathbb{R}^{K \times m} \\ \mathbf{z} \in \mathbb{R}^m}} \quad & - u(f(\mathbf{a}_*, \mathbf{x}_{(1)})) + \sum_{i=2}^{m} z_i (u(f(\mathbf{a}_*, \mathbf{x}_{(i-1)})) - u(f(\mathbf{a}_*, \mathbf{x}_{(i)}))) \\
\text{subject to} \quad & \sum_{k=i}^{m} q_{(k)} = \sum_{j=1}^{K} \lambda_{ij} t_j, \ \forall i \in \{2, \ldots, m\} \\
& z_i = \sum_{j=1}^{K} \lambda_{ij} h(t_j), \ \forall i \in \{2, \ldots, m\} \\
& \sum_{j=1}^{K} \lambda_{ij} = 1, \ \forall i \in [m] \\
& \lambda_{ij} \geq 0, \text{SOS2}, \ \forall i \in [m], \ \forall j \in [K].
\end{aligned} \tag{36}$$

The SOS2 constraints control, for each $i \in [m]$, the variables $\{\lambda_{ij}\}_{j=1}^K$ in a way such that only two adjacent $\lambda_{ij}, \lambda_{i,j+1}$ can be non-zero. Since the support points $\{t_j\}_{j=0}^K$ are ordered, the two non-zero adjacent variables $\lambda_{ij}, \lambda_{i,j+1}$ ensure that we are only optimizing within each interval $[t_j, t_{j+1}]$.

Thus, we can now introduce Algorithm 3, the cutting-plane approach that enables us to solve the robust problem (P) for piecewise-linear, inverse S-shaped distortion functions.

The following theorem establishes that this cutting-plane method terminates after finitely many iterations.

Theorem 10. *Let h be an inverse S-shaped, piecewise-linear distortion function. Suppose that*

Algorithm 3 **Cutting-Plane Method with Inverse S-Shaped Distortion Function**

1: Start with $\mathcal{U}_1 = \{\mathbf{p}\}$. Fix a tolerance parameter $\epsilon_{\text{tol}} > 0$.

2: At the j-th iteration, solve the following problem with the uncertainty set $\mathcal{U}_j$, to obtain a solution $\mathbf{a}_j$ and optimal objective value c_j:

$$\min_{\mathbf{a}\in\mathcal{A}} \sup_{\mathbf{q}\in\mathcal{U}_j} \rho_{u,h,\mathbf{q}}(f(\mathbf{a},\mathbf{X})), \tag{37}$$

by applying the reformulation in Theorem 9 to each element of $\mathcal{U}_j$.

3: Determine the robust rank-dependent evaluation v_j of $\mathbf{a}_j$ by solving (36), which gives an optimal solution $\mathbf{q}_j^*$.

4: If $v_j - c_j \leq \epsilon_{\text{tol}}$, then the solution $\mathbf{a}_j$ is accepted and the process terminated.

5: If not, set $\mathcal{U}_{j+1} = \mathcal{U}_j \cup \{\mathbf{q}_j^*\}$, and repeat steps 2–5.

$\sup_{\mathbf{a}\in\mathcal{A},i\in[m]} |u(f(\mathbf{a},\mathbf{x}_i))| < \infty$. *Then, for all* $\epsilon_{\text{tol}} > 0$, *Algorithm 3 terminates after finitely many iterations.*

Next, we examine the convergence of the piecewise-linear approximation to problems (P-Nom) and (P). In conjunction with Theorem 10, it constitutes, in a sense, the master theorem of the paper.

Theorem 11. *Let h be an inverse S-shaped distortion function and let h_ϵ be such that* $\sup_{p\in[0,1]} |h(p) - h_\epsilon(p)| \leq \epsilon$. *Suppose that* $\sup_{\mathbf{a}\in\mathcal{A},i\in[m]} |u(f(\mathbf{a},\mathbf{x}_i))| < \infty$. *Then, the optimal objective value of* (P-Nom) *for h_ϵ converges to that of h, as $\epsilon \to 0$.*

Similarly, for the robust problem (P)*, Algorithm 3, if applied to h_ϵ, yields an optimal objective value that converges to the exact optimal objective value of* (P) *for h, as both $\epsilon_{\text{tol}}, \epsilon \to 0$.*

We note that when h is an inverse-S-shaped distortion function, the piecewise-linear approximation specified by $\{h_l, \bar{h}_l\}$ in (32) provides neither a lower bound nor an upper bound on $h(p)$, for all $p \in [0,1]$. Indeed, h_l is a lower approximation of h on $[0, p^0]$, whereas $1 - \bar{h}_l(1-p)$ is an upper approximation of h on $[p_0, 1]$. Nonetheless, if one aims to bound h on the entire $[0,1]$, then we can apply the same device as in (27): either translate h_l with an ϵ (the maximal error) to obtain an upper piecewise-linear approximation of h on $[0, p^0]$, or do the same with the dual function $\bar{h}_l$ for a lower approximation on $[p^0, 1]$. Theorem 11 then implies that the corresponding lower and upper bounds on the optimal objective value of (P-Nom) (or (P)), computed from these piecewise-linear approximations of h, will converge to the exact optimal value, as $\epsilon \to 0$.

Finally, we note that Algorithm 3 can also be adapted (similar to Algorithm EC.1 in the Electronic Companion) to solve problem (P-constraint), where the rank-dependent evaluation with inverse-S-shaped distortion function is in the constraint. Hence, we also present the following theorems with similar statements as in Theorems 10 and 11.

Theorem 12. *Let h be an inverse S-shaped, piecewise-linear distortion function. Suppose that $u(f(\mathbf{a},\mathbf{x}_i))$ is continuous in $\mathbf{a} \in \mathcal{A}$ for all $i \in [m]$. Then, for any $\epsilon_{\text{tol}} > 0$, Algorithm 3 terminates* *after finitely many iterations when applied to* (P-constraint)*. Furthermore, if h is also continuous, then the final solution $\mathbf{a}_{\epsilon_{\text{tol}}}$ of Algorithm 3 gives a lower bound $g(\mathbf{a}_{\epsilon_{\text{tol}}})$ that converges to the optimal objective value of* (P-constraint) *as $\epsilon_{\text{tol}} \to 0$.*

Theorem 13. *Let h be an inverse S-shaped distortion function. Suppose that $u(f(\mathbf{a}, \mathbf{x}_i))$ is continuous in $\mathbf{a} \in \mathcal{A}$ for all $i \in [m]$. If g, appearing in the objective, and h are continuous, then for any distortion function h_ϵ such that $h_\epsilon(p) \leq h(p)$ for all $p \in [0, 1]$ and $\sup_{p\in[0,1]} |h(p) - h_\epsilon(p)| \leq \epsilon$, the optimal objective value of* (P-constraint) *for h_ϵ converges to that of h, as $\epsilon \to 0$.*

7 Numerical Examples

In this section, we illustrate the various methods developed in this paper by applying them to two canonical examples of optimization problems: the *newsvendor* problem and the *portfolio choice* problem. In both examples, we compare the solution of the robust problem (P) to the solution of the nominal problem (P-Nom). We use the phrase "robust/nominal solution" to refer to the solution of the robust/nominal problem.

7.1 Robust Single-Item Newsvendor

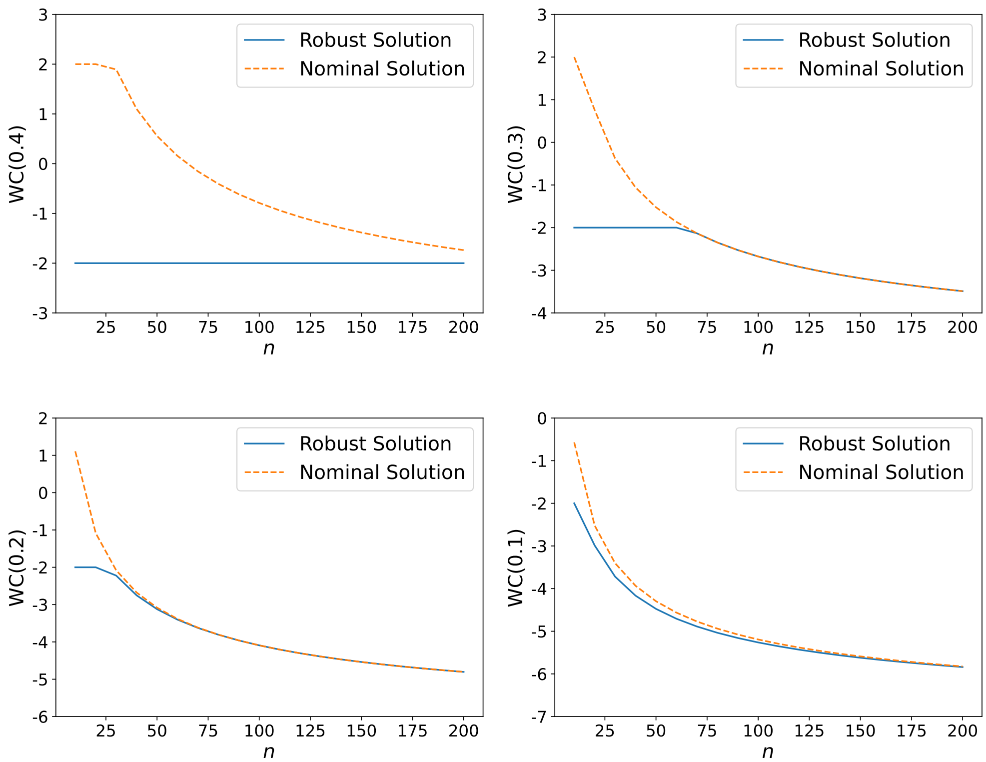
In the single-item newsvendor problem, a seller is uncertain about the demand for a certain product, and has to decide in advance how many units of the product have to be ordered and stocked in the inventory. Let $d_i \geq 0$ denote the realization of the demand in state i. Furthermore, let: c be the ordering cost of one unit, $v > c$ be the selling price, $s < c$ be the salvage value per unsold item returned to the factory, and l be the loss per unit of unmet demand, which may include both the cost of a lost sale and a penalty for the lost customer goodwill. Finally, let y be the number of items ordered, i.e., the decision variable. The profit function $\pi(d_i, y)$ is defined as

$$\begin{aligned} \pi(d_i, y) &\triangleq v \min\{d_i, y\} + s(y - d_i)_+ - l(d_i - y)_+ - cy \\ &= (s - v)(y - d_i)_+ - l(d_i - y)_+ + (v - c)y. \end{aligned} \tag{38}$$

Note that $\pi(d_i, y)$ is concave in y since it is a sum of concave piecewise-linear functions in y for all i. Assume that the demand $d_i \in \{4, 8, 10\}$, which corresponds to low, medium, and high demand, with nominal probabilities $\mathbf{p} = \{0.375, 0.375, 0.25\}$. Furthermore, we assume that the number of items ordered will not exceed the maximum demand, i.e., $y \leq 10$. Finally, we set the parameter values $c = 4, v = 6, s = 2, l = 4$.

We solve both the robust problem (P) and the nominal problem (P-Nom), where the decision variable y is subject to the constraint $0 \leq y \leq 10$, $\mathbf{d}$ is the uncertain parameter, and $\pi(\mathbf{d}, y)$ is the random objective function. The nominal and robust newsvendor problem is studied in Eeckhoudt et al. (1995) and Ben-Tal et al. (2013), respectively, assuming expected utility preferences. We choose $\text{CVaR}_{1-\alpha_0}$ (see (6)) to be our rank-dependent evaluation $\rho_{u,h,\mathbf{q}}$. This corresponds to a

Figure 1: Single-item newsvendor problem. This figure displays the worst-case evaluation $\text{WC}(\alpha_0) \triangleq \sup_{\mathbf{q}\in\mathcal{D}_\phi(\mathbf{p},r(n))} \text{CVaR}_{1-\alpha_0}(.)$ under the robust and nominal solutions, for a range of values of $r(n) = \chi^2_{2,0.95}/(2n)$ and $\alpha_0 = 0.4, 0.3, 0.2, 0.1$.

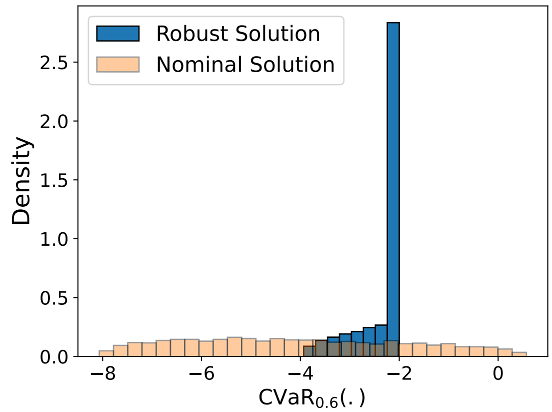


piecewise-linear distortion function $h(p) = \min\{p/(1-\alpha_0), 1\}$ and a linear utility function $u(x) = x$. For the robust problem (P), we choose the KL-divergence function $\phi(t) = t\log(t) - t + 1$, and set the radius of the divergence set $\mathcal{D}_\phi(\mathbf{p}, r)$ to be $r = \chi^2_{2,0.95}/(2n)$ as in (10), where n is the sample size that we assume that $\mathbf{p}$ is estimated from. A larger value of the objective function indicates a larger (i.e., worse) evaluation of the utility loss; a value of $\alpha_0 = 0.4$ means that the worst 60% of the utility loss distribution is considered.

As the cardinality of realizations of the uncertain parameter d is merely 3, we can readily invoke Theorem 2 to solve both the reformulated nominal and robust problems (P-ref) and (P-Nom-ref) exactly, without reducing the number of constraints. We then perform the following experiment: For each α_0, we first obtain a nominal solution to (P-Nom). Then, for each radius $r(n) = \chi^2_{2,0.95}/(2n)$, we obtain a robust solution to (P). To compare the robust solution with the nominal solution, we calculate their worst-case evaluation $\sup_{\mathbf{q}\in\mathcal{D}_\phi(\mathbf{p},r(n))} \text{CVaR}_{1-\alpha_0}(.)$ under the radius $r(n)$. This is repeated for a range of $n = 10, 20, \ldots, 200$, and $\alpha_0 = 0.4, 0.3, 0.2, 0.1$. As n increases, the ϕ-divergence radius r decreases, and thus there is less ambiguity.

The results are displayed in Figure 1. Quite naturally, we observe that the differences between the worst-case evaluations of both solutions decrease as n increases. We also observe a decrease in

Figure 2: Single-item newsvendor problem (*continued*): This figure displays, for each sampled probability vector $\mathbf{q}$ from the KL-divergence uncertainty set for $n = 50$, the corresponding $\text{CVaR}_{1-\alpha_0}(.)$ evaluation with $\alpha_0 = 0.4$ of the robust and nominal solutions.



the value of the worst-case evaluation as α_0 decreases, reflecting that the $\text{CVaR}_{1-\alpha_0}$ risk measure itself is monotonically decreasing in α_0. Furthermore, we see that for lower values of α_0, the differences between the worst-case evaluations of both solutions are already small for small sample size n. Moreover, we observe a flat, constant worst-case evaluation of the robust solution at level -2. This is the largest (i.e., worst) value that the robust rank-dependent evaluation attains for a robust solution $y^* = 7$. This conservativeness occurs when n is relatively small.

To further illustrate the differences between the robust and nominal solutions, we use the Hit-and-Run algorithm (see Electronic Companion EC.9 for further details) to sample 5,000 probability vectors $\mathbf{q}$ from the KL-divergence uncertainty set for $n = 50$. For each sampled vector $\mathbf{q}$, we calculate the $\text{CVaR}_{1-\alpha_0}$ evaluation, with $\alpha_0 = 0.4$. As shown in Figure 2, the evaluation of the nominal solution exhibits a large variance. It also exceeds the largest (i.e., worst) evaluation of the robust solution. On the other hand, the robust solution shows less variance and is concentrated at the value -2, which as mentioned above, is the most conservative evaluation that the robust optimal solution attains.

7.2 Robust Multi-Item Newsvendor

We next examine the multi-item newsvendor problem, where each item in a set of items features its own uncertain demand. Let $d_i^{(j)}$ be the i-th realization of the j-th item's demand, $i = 1, 2, 3$. We take $j = 1, 2, 3$ and consider the sum of the individual profit functions

$$\pi_{\text{tot}}(\mathbf{d}, \mathbf{y}) = \pi_1(d^{(1)}, y_1) + \pi_2(d^{(2)}, y_2) + \pi_3(d^{(3)}, y_3), \tag{39}$$

where

$$\pi_j(d^{(j)}, y_j) = v_j \min\{d^{(j)}, y_j\} + s_j(y_j - d^{(j)})_+ - l_j(d^{(j)} - y_j)_+ - c_j y_j, \tag{40}$$

Table 1: Multi-item newsvendor problem. This table displays the parameters used in the multi-item newsvendor problem.

Item(j)	c	v	s	l	$p_1^{(j)}$	$p_2^{(j)}$	$p_3^{(j)}$
1	4	6	2	4	0.375	0.375	0.25
2	5	8	2.5	3	0.25	0.25	0.5
3	4	5	1.5	4	0.127	0.786	0.087

with v_j, s_j, l_j, c_j the parameters corresponding to item j. We assume that the demand takes on the same possible values for all items, i.e., $d_i^{(j)} \in \{4, 8, 10\}$ for all j. Since each realization of $d^{(j)}$ contributes to a possible realization of $\pi_{\text{tot}}(\mathbf{y}, \mathbf{d})$, there are in total $m = 3^3 = 27$ possible realizations. We solve again problems (P) and (P-Nom) with the same preference specification as in the single-item problem. Since the $\text{CVaR}_{1-\alpha_0}$ risk measure has a piecewise-linear distortion function, we can apply Theorem 6 to solve (P-ref) and (P-Nom-ref) in this higher-dimensional setting. We take the nominal probability $\mathbf{p} \in \mathbb{R}^{27}$ to be the probability of each combination of realizations of $(d^{(1)}, d^{(2)}, d^{(3)})$, which is the product of the probabilities $\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \mathbf{p}^{(3)}$ of each individual realization. The parameters we use are given in Table 1.

Similar to the single-item problem, we investigate the differences between the worst-case evaluations of the robust and nominal solutions, for a range of n and α_0. We choose $\alpha_0 = 0.4, 0.3$ to compare the results with the single-item problem, and $\alpha_0 = 0.9, 0.8$ to explore higher values of α_0. From Figure 3, we observe that in the multi-item problem, the worst-case evaluation is much lower (i.e., better) than in the single-item problem, which speaks to the diversification benefits of a multi-item inventory. The overall pattern is similar to the single-item case: for relatively low α_0, the difference in worst-case evaluation between the robust solution and the nominal solution is smaller than for larger α_0. The same holds as the sample size n increases.

7.3 Robust Portfolio Choice with Concave Distortion Function

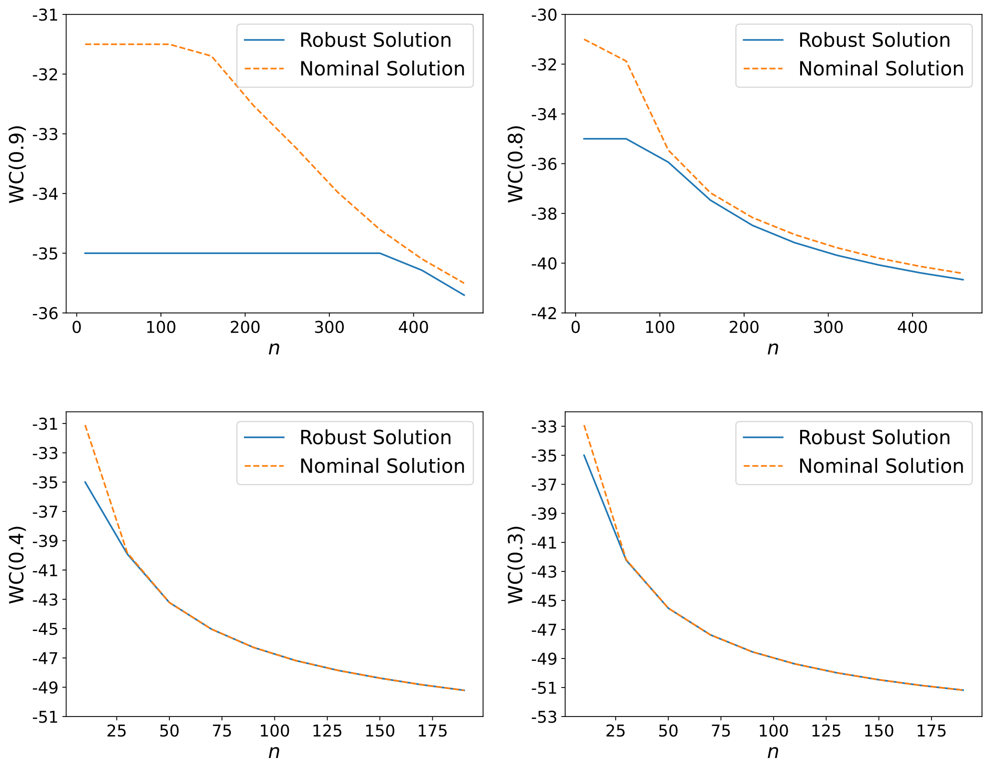
In this subsection, we investigate the performance of the cutting-plane algorithm and the piecewise-linear approximation method described in Algorithms 1 and 2, by studying robust portfolio optimization problems. We consider the returns of six portfolios formed on size and book-to-market ratio (2×3) obtained from Kenneth French's data library.[10] We use monthly returns from January 1984 to January 2014. This gives us a total of $m = 360$ return realizations for each of the six portfolios. We denote by $\mathbf{r} \in \mathbb{R}^6$ the uncertain return of the six portfolios, with empirical realizations $\mathbf{r}_1, \ldots, \mathbf{r}_{360}$. We solve the following nominal and robust portfolio optimization problems:

$$\min_{\mathbf{a} \in \mathcal{A}} \rho_{u,h,\mathbf{p}}(1 + \mathbf{a}^T \mathbf{r}) \tag{41}$$

$$\min_{\mathbf{a} \in \mathcal{A}} \sup_{\mathbf{q} \in \mathcal{D}_\phi(\mathbf{p}, r)} \rho_{u,h,\mathbf{q}}(1 + \mathbf{a}^T \mathbf{r}), \tag{42}$$

[10] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Figure 3: Multi-item newsvendor problem. This figure displays the worst-case evaluation $\mathrm{WC}(\alpha_0) \triangleq \sup_{\mathbf{q}\in\mathcal{D}_\phi(\mathbf{p},r(n))} \mathrm{CVaR}_{1-\alpha_0}(.)$ under the robust and nominal solutions, for a range of values of $r(n) = \chi^2_{2,0.95}/(2n)$ and $\alpha_0 = 0.9, 0.8, 0.4, 0.3$.



where $1 + \mathbf{a}^T\mathbf{r}$ is the total wealth after one period with initial wealth normalized to unity, which is evaluated under the rank-dependent utility model. The decision variable $\mathbf{a} \in \mathbb{R}^6$ is subject to the constraints $\mathcal{A} = \{\mathbf{a} \in \mathbb{R}^6 \mid \sum_{j=1}^6 a_j = 1, a_j \geq 0\}$. The nominal probability $\mathbf{p}$ is set to be the empirical distribution with $p_i = \frac{1}{360}$. We use the modified chi-squared divergence function $\phi(x) = (x-1)^2$, with radius $r = \frac{1}{n}\chi^2_{359,0.95}$. We choose the concave distortion function $h(p) = 1-(1-p)^2$, motivated by two decision-theoretic papers: Eeckhoudt et al. (2020) and Eeckhoudt and Laeven (2021). Furthermore, we choose the exponential utility function $u(x) = 1 - e^{-x/\lambda}$ with $\lambda = 10$.

We solve problems (41)–(42) with the cutting-plane method described in Algorithm 1, and compare its performance to the piecewise-linear method described in Algorithm 2, formally relying on Theorems 3 and 7. As shown in Table 2, both algorithms yield a very small gap between the upper and lower bounds they provide. The piecewise-linear approximation method is typically more efficient, but this is not surprising, since the cutting-plane method is a more general method and utilizes less structure of the problems at hand.

Table 2: Robust portfolio choice. This table displays the results of Algorithms 1 and 2 when applied to the robust problem (42) and the nominal problem (41). The respective tolerance parameters are set to $\epsilon_{\text{tol}} = 0.0001$ and $\epsilon = 0.001$.

Panel A: The cutting-plane method

Problem	Lower Bound	Upper Bound	# Cuts	Run Time
Robust	-0.08953	-0.08948	5	54 sec
Nominal	-0.09403	-0.09397	5	22 sec

Panel B: Piecewise-linear approximation

Problem	Lower Bound	Upper Bound	Run Time
Robust	-0.08951	-0.08948	14 sec
Nominal	-0.09401	-0.09398	7 sec

Additionally, for the cutting-plane algorithm, we calculate the worst-case rank-dependent evaluation of the nominal solution, which is equal to -0.08938. We also calculate the evaluation of the robust solution under the nominal distribution $\mathbf{p}$, which is equal to -0.09395. As we can see from the table, the worst-case evaluation of the nominal solution is not much larger than that of the robust solution (cf. the first rows of the Panels A and B), suggesting that the nominal solution is already "near-optimal" for the robust problem (42). Similarly, the robust solution is also "near-optimal" for the nominal problem (41). It appears that when the dimension m of the state space is large, the robust and nominal problems do not yield highly different solutions in this example.

7.4 Robust Portfolio Choice with Inverse S-shaped Distortion Function

In this subsection, we investigate portfolio optimization problems (41)–(42) when h is an inverse S-shaped distortion function. In particular, we examine the Prelec (1998) distortion function:

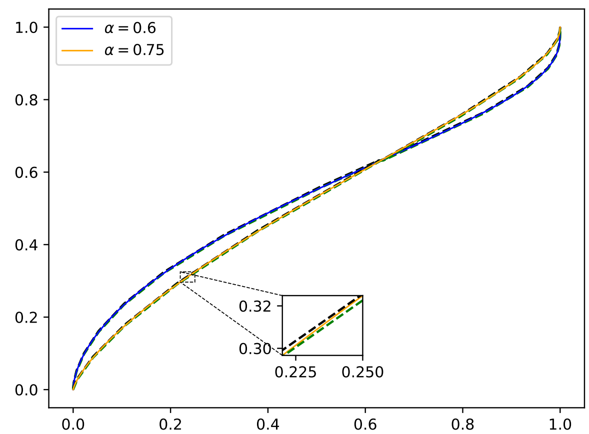
$$h_\alpha(p) \triangleq 1 - \exp(-(-\log(1-p))^\alpha), \quad 0 < \alpha < 1. \tag{43}$$

We focus primarily on $\alpha = 0.6$, 0.75, which are broadly consistent with common empirical findings (Wakker, 2010, p. 260). We consider a linear utility function, $u(x) = x$, to isolate the effect of inverse S-shaped probability weighting. To obtain an upper and lower bound on problems (41)–(42), we approximate Prelec's function from above and below, using piecewise-linear functions. The approximation procedure is carried out by applying the method described in Section 5.2 separately to the concave and convex parts of Prelec's function, where we minimize the number of linear pieces under a pre-specified approximation error.[11] The shape of Prelec's function and its respective upper and lower piecewise-linear approximations are displayed in Figure 4.

We generate 100 return data for 5 assets using the same numerical scheme as in Esfahani and Kuhn (2018), where the returns for each asset $j \in \{1, \ldots, 5\}$ are composed of two fac-

[11]The approximation error is chosen to be $\epsilon = 0.003$ leading to 19, 13, and 6 linear pieces for $\alpha = 0.6$, 0.75, 0.95.

Figure 4: Prelec's distortion. This figure displays Prelec's distortion function (43) and its upper and lower piecewise-linear approximations (dashed) for $\alpha = 0.6$, 0.75. The approximation error is set to $\epsilon = 0.003$.



tors: $r_j = \psi + \gamma_j$, with a systematic risk factor $\psi \sim N(0, 0.02)$ and an idiosyncratic risk factor $\gamma_j \sim N(0.03j, 0.25j)$, for the j-th asset. Here, $N(\mu, \sigma)$ denotes the Gaussian distribution. By construction, the asset with a higher index j has a higher expected return and standard deviation.

We obtain lower and upper bounds on the nominal problem (41) by implementing the constraints (34) of Theorem 9 for the lower and upper piecewise-linear approximation of h, respectively, in the Gurobi solver (Version 11.0.3). Next, we invoke Theorems 10–11. Hence, a lower bound for the robust problem (42) is obtained by calculating the lower bound implied by the cutting-plane method described in Algorithm 3 (where we set $\epsilon_{\text{tol}} = 0.001$), while using the lower approximation of h. Then, with the solution obtained from the cutting-plane method, we determine an upper bound on (42) by calculating its worst-case evaluation using the SOS2-constraints as in (36), where we use the upper approximation of h. For the robust problem, we consider the total variation divergence $\phi(t) = |t - 1|$. We choose the radius for the robust problem to be $r = \frac{1}{m}\chi^2_{m-1,0.95}$, where $m = 100$.

The results are displayed in Table 3. We observe very tight upper and lower bounds. Furthermore, Gurobi solves both the nominal and robust problems (41)–(42) efficiently. As we can observe from Panel A, the evaluation of the nominal solution decreases with α. This is not surprising. Indeed, as we can see in Figure 4, Prelec's function has the property that if $\alpha_0 < \alpha_1$, then $h_{\alpha_0}(p) > h_{\alpha_1}(p)$ for the concave segment $p \in [0, 1 - 1/e]$, which covers more than half of the interval $[0, 1]$. Interestingly, Panel B of Table 3 reveals that the robust evaluation of the robust solution can exhibit an opposite, mildly increasing relationship with respect to α. This is because in the robust case, if the radius r is sufficiently large, then the worst-case probability $q^*_{(m)}$ of the worst-case realization can be such that $q^*_{(m)} \geq 1 - 1/e$ (recall that (m) denotes the index of the highest ranked realization, as in (35)). When this happens, all distorted decumulative probabilities in the objective function of (35) are increasing in α, since then all decumulative probabilities are on the interval $[1 - 1/e, 1]$. If the radius r would be set sufficiently small, then the same decreasing

Table 3: Robust portfolio choice with inverse S-shaped distortion function. Upper and lower bounds (UB and LB) obtained for the nominal problem (41) and the robust problem (42) when h is Prelec's distortion function with $\alpha \in \{0.6, 0.75, 0.95\}$ (see Figure 4) and the divergence function is given by $\phi(t) = |t-1|$. The cardinality of return realizations is $m = 100$.

Panel A: Solutions Nominal Problem

α	LB	UB	Run Time (LB)	Run Time (UB)
0.6	-1.144	-1.142	1.34 sec	1.21 sec
0.75	-1.153	-1.152	0.44 sec	0.90 sec
0.95	-1.162	-1.160	0.43 sec	0.30 sec

Panel B: Solutions Robust Problem

α	LB	UB	# Cuts	Run Time (LB)	Run Time (UB)
0.6	-1.041	-1.040	9	314 sec	2.77 sec
0.75	-1.039	-1.038	9	142 sec	1.45 sec
0.95	-1.036	-1.035	9	98 sec	0.37 sec

pattern with respect to α as in Panel A would also emerge in the robust case.

8 Concluding Remarks

In this paper, we have shown that nominal and robust optimization problems involving rank-dependent models can be reformulated into rank-independent, tractable optimization problems. When the distortion function is concave, we have demonstrated that this reformulation admits a conic representation, which we have explicitly derived for canonical distortion and divergence functions. Whereas the multiplicity of constraints in the reformulation increases exponentially with the dimension of the underlying probability space, we have developed two types of algorithms, and combinations thereof, to circumvent this curse of dimensionality. We have formally established that the upper and lower bounds the algorithms generate converge to the optimal objective value. Finally, we have illustrated the good performance of our methods in two examples involving concave as well as inverse S-shaped distortion functions, efficiently yielding nominal and robust solutions that generate very tight upper and lower bounds on the optimal objective values.

As a direction for future research, one can investigate whether the approach developed in this paper can be extended to encompass robust optimization problems with general law-invariant convex functionals (Föllmer and Schied (2016), S. 4.5), which also admit a dual representation.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. *Econometrica*, 21:503–546.

Ben-Tal, A., den Hertog, D., de Waegenaere, A., Melenberg, B., and Rennen, G. (2013). Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357.

Ben-Tal, A. and Nemirovski, A. (2019). *Lectures on Modern Convex Optimization*. SIAM.

Ben-Tal, A. and Teboulle, M. (1986). Expected utility, penalty functions, and duality in stochastic nonlinear programming. *Management Science*, 32(11):1445–1466.

Ben-Tal, A. and Teboulle, M. (1987). Penalty functions and duality in stochastic programming via ϕ-divergence functionals. *Mathematics of Operations Research*, 12(2):224–240.

Bertsimas, D. and Brown, D. B. (2009). Constructing uncertainty sets for robust linear optimization. *Operations Research*, 57(6):1483–1495.

Bertsimas, D. and den Hertog, D. (2022). *Robust and Adaptive Optimization*. Dynamic Ideas LLC Belmont, Massachusetts.

Bertsimas, D., Dunning, I., and Lubin, M. (2016). Reformulation versus cutting-planes for robust optimization. *Computational Management Science*, 13(2):195–217.

Cai, J., Li, J. Y.-M., and Mao, T. (2025). Distributionally robust optimization under distorted expectations. *Operations Research*, 73(2):969–985.

Carlier, G. and Dana, R.-A. (2006). Law invariant concave utility functions and optimization problems with monotonicity and comonotonicity constraints. *Statistics & Risk Modeling*, 24(1):127–152.

Chares, P. R. (2009). *Cones and Interior-Point Algorithms for Structured Convex Optimization Involving Powers and Exponentials*. Ecole polytechnique de Louvain, Université Catholique de Louvain.

Chew, S., Karni, E., and Safra, Z. (1987). Risk aversion in the theory of expected utility with rank dependent probabilities. *Journal of Economic Theory*, 42:370–381.

Choquet, G. (1954). Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295.

Cox, M. (1971). An algorithm for approximating convex functions by means by first degree splines. *The Computer Journal*, 14(3):272–275.

Csiszár, I. (1975). i-divergence geometry of probability distributions and minimization problems. *Annals of Probability*, 3:146–158.

Delage, E., Guo, S., and Xu, H. (2022). Shortfall risk models when information of loss function is incomplete. *Operations Research*, 70(6):3511–3518.

Denneberg, D. (1994). *Non-Additive Measure and Integral*. Springer.

Eeckhoudt, L., Gollier, C., and Schlesinger, H. (1995). The risk-averse (and prudent) newsboy. *Management Science*, 41(5):786–794.

Eeckhoudt, L. R. and Laeven, R. J. (2021). Dual moments and risk attitudes. *Operations Research*, 70(3):1330–1341.

Eeckhoudt, L. R., Laeven, R. J., and Schlesinger, H. (2020). Risk apportionment: The dual story. *Journal of Economic Theory*, 185:104971.

Esfahani, P. M. and Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171:115–166.

Föllmer, H. and Schied, A. (2016). *Stochastic Finance*. Berlin: De Gruyter, 4th edition.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Goovaerts, M. J. and Laeven, R. J. (2008). Actuarial risk measures for financial derivative pricing. *Insurance: Mathematics and Economics*, 42(2):540–547.

Gorissen, B. L., Blanc, H., Ben-Tal, A., and den Hertog, D. (2014). Technical Note–Deriving robust and globalized robust solutions of uncertain linear programs with general convex uncertainty sets. *Operations Research*, 62(3):672–679.

Gurobi Optimization, LLC (2023). Gurobi Optimizer Reference Manual.

Hansen, L. P. and Sargent, T. J. (2001). Robust control and model uncertainty. *American Economic Review*, 91:60–66.

Hansen, L. P. and Sargent, T. J. (2007). *Robustness*. Princeton University Press.

He, X. D. and Zhou, X. Y. (2011). Portfolio choice via quantiles. *Mathematical Finance*, 21(2):203–231.

Huber, P. J. (1981). *Robust Statistics*. Wiley, New York.

Imamoto, A. and Tang, B. (2008). Optimal piecewise linear approximation of convex functions. *Proceedings of the World Congress on Engineering and Computer Science*.

Laeven, R. J. and Stadje, M. (2023). A rank-dependent theory for decision under risk and ambiguity. *Technical Report*.

Maccheroni, F., Marinacci, M., and Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498.

Mutapcic, A. and Boyd, S. (2009). Cutting-set methods for robust convex optimization with pessimizing oracles. *Optimization Methods & Software*, 24(3):381–406.

Pardo, L. (2006). *Statistical Inference Based on Divergence Measures*. Chapman & Hall/ CRC Boca Raton.

Pesenti, S., Wang, Q., and Wang, R. (2020). Optimizing distortion riskmetrics with distributional uncertainty. *Paper*. Available on ArXiv.

Postek, K., den Hertog, D., and Melenberg, B. (2016). Computationally tractable counterparts of distributionally robust constraints on risk measures. *SIAM Review*, 58(4):603–650.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3):497–527.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization*, 3(4):323–343.

Ryoo, H. S. and Sahinidis, N. V. (1996). A branch-and-reduce approach to global optimization. *Journal of Global Optimization*, 8(2):107–138.

Schmeidler, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society*, 97:255–261.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–323.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 3rd edition.

Wakker, P. P. (2010). *Prospect Theory for Risk and Rationality*. Cambridge University Press.

Wald, A. (1950). *Statistical Decision Functions*. Wiley, New York.

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

Wang, W. and Xu, H. (2023). Preference robust distortion risk measure and its application. *Mathematical Finance*, 33(2).

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

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