Risk-averse Regret Minimization in Multi-stage Stochastic Programs

Mehran Poursoltani[a,*], Erick Delage[a], and Angelos Georghiou[b]

[a]GERAD & Department of Decision Sciences, HEC Montréal, Montréal, Canada
[b]Department of Business and Public Administration, University of Cyprus, Nicosia, Cyprus

December 10, 2021

Abstract

Within the context of optimization under uncertainty, a well-known alternative to minimizing expected value or the worst-case scenario consists in minimizing regret. In a multi-stage stochastic programming setting with a discrete probability distribution, we explore the idea of risk-averse regret minimization, where the benchmark policy can only benefit from foreseeing Δ steps into the future. The Δ-regret model naturally interpolates between the popular ex-ante and ex-post regret models. We provide theoretical and numerical insights about this family of models under popular coherent risk measures and shed new light on the conservatism of the Δ-regret minimizing solutions.

1 Introduction

The regret minimization paradigm, introduced by Savage (1951), is claimed to provide less conservative solutions compared to the ones returned by optimizing with respect to the the worst-case scenario (Perakis and Roels 2008, Aissi et al. 2009, Natarajan et al. 2014, Caldentey et al. 2017). Given a profit function $h(\boldsymbol{x}, \boldsymbol{\zeta})$, which depends on the decision $\boldsymbol{x}$ and an uncertain vector of parameters $\boldsymbol{\zeta}$, the regret minimization approach aims at minimizing the difference between the achieved profit and the best profit that would have been made if the realization of $\boldsymbol{\zeta}$ was known before making the decision. Namely, the so-called ex-post worst-case regret minimization problem takes the form of:

$$\text{(EP-WCR)} \quad \underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \;\; \max_{\omega\in\Omega} \;\; \left\{ \max_{\boldsymbol{x}'\in\mathcal{X}} h(\boldsymbol{x}', \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}, \boldsymbol{\zeta}(\omega)) \right\},$$

where $\mathcal{X}$ is the set of admissible actions, Ω denotes the outcome space, and $\boldsymbol{x}'$ captures the decision made with full information about ω, which we will refer to as the benchmark policy.

While most of the regret minimization literature focuses on worst-case scenario analysis, there has recently been a scarce but growing interest for formulations that account for more information about the underlying potential of realization of the different outcomes. A first common approach can be referred as the ex-post risk averse regret minimization problem:

$$\text{(EP-RAR)} \quad \underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \;\; \rho\left(\max_{\boldsymbol{x}'\in\mathcal{X}} h(\boldsymbol{x}', \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}, \boldsymbol{\zeta}(\omega)) \right),$$

[*]Corresponding author
Email addresses:
mehran.poursoltani@hec.ca (Mehran Poursoltani)
erick.delage@hec.ca (Erick Delage)
georghiou.angelos@ucy.ac.cy (Angelos Georghiou)

where ρ can either be a law-invariant risk measure (see Kusuoka (2001)), e.g. expected value or Conditional Value-at-Risk (CVaR), or a worst-case risk measure (c.f. Postek et al. (2018)), e.g. a worst-case expected value that accounts for incomplete distribution information. For example, Natarajan et al. (2014) proposed an ex-post regret minimization model equipped with a worst-case CVaR risk measure that accounted for information about the marginal distribution of the different terms of $\boldsymbol{\zeta}$. Indeed, having access to distributional information enables one to employ a variety of popular risk measures, which can help further control conservatism by trading off between the expected value and tail risks of the regret with respect to a fully informed decision.

A second approach (see Perakis and Roels (2008)) departs from the traditional ex-post regret form as it instead measures regret with respect to an action $\boldsymbol{x}'$ that does not have knowledge of the realized scenario. This rather gives rise to what can be referred as the ex-ante risk averse regret minimization problem:

$$\text{(EA-RAR)} \quad \underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \;\; \max_{\boldsymbol{x}'\in\mathcal{X}} \;\; \rho\left(h(\boldsymbol{x}',\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x},\boldsymbol{\zeta}(\omega))\right) .$$

To clarify, we illustrate the distinction between the two approaches using a simple project selection problem with partial distirbution information as an example.

Example 1. *A manager must choose one of the three available projects for investment (i.e. $\mathcal{X} := \{\boldsymbol{x}_A, \boldsymbol{x}_B, \boldsymbol{x}_C\}$) and considers two possible scenarios (i.e. $\Omega := \{\omega_1, \omega_2\}$) for the projects' payoff. Although the true probability of each scenario is not known to the manager, she considers two different possibilities (i.e. $\mathcal{P} := \{\mathbb{P}^I, \mathbb{P}^{II}\}$) and employs worst-case expected value as the risk measure.*[1] *Table 1 provides the numerical details while Table 2 presents the optimal project selected under four different regret minimization formulations: {Ex-ante/Ex-post} {Worst-case/Risk averse} regret minimization. The reader is referred to Appendix A for further numerical details.*

Table 1: Numerical details of Example 1

	Project payoffs			Probabilities	
	$\boldsymbol{x}_A$	$\boldsymbol{x}_B$	$\boldsymbol{x}_C$	$\mathbb{P}^I$	$\mathbb{P}^{II}$
ω_1	1\$	5\$	4\$	80%	0%
ω_2	6\$	2\$	3\$	20%	100%

Table 2: Optimal project selected under four variants of regret minimization in Example 1

	Ex-ante	Ex-post
Worst-case	$\boldsymbol{x}_C$	$\boldsymbol{x}_C$
Risk averse	$\boldsymbol{x}_A$	$\boldsymbol{x}_C$

Specifically, both ex-post models measure the regret under each outcome by comparing to the best action in hindsight: i.e., $\boldsymbol{x}_B$ and $\boldsymbol{x}_A$ under scenarios ω_1 and ω_2 respectively. However, ex-ante model needs to consider the same action $\boldsymbol{x}'$ to compare to under all the scenarios. Namely, in the case of the risk-averse model, we have $\boldsymbol{x}'^* = \boldsymbol{x}_B$. Looking at Table 2 one can remark that while under a worst-case regret formulation, the optimal decision is unaffected by the use of ex-post or ex-ante regret, this is not the case anymore when using a risk averse setting.

Example 1 raises questions such as what are conditions under which EP-RAR and EA-RAR are equivalent, whether other formulations exists between ex-ante and ex-post that could fill in the gap between the two solutions (especially in a multi-stage setting), and finally what are the implications of these formulation in terms of level of conservatism. To the best of our knowledge, this paper investigates these questions for the first time and by presenting a new multi-stage regret minimization formulation that measures regret with respect to decisions that can exploit information revealed up to Δ stages into the future. This model effectively interpolates very naturally between the ex-ante (with $\Delta = 0$) and ex-post (with $\Delta = \infty$) models and effectively allows to study them under the same lens.

Overall, the contribution can be summarized as follows:

- Theoretically, we show that EP-RAR and EA-RAR are equivalent in terms of optimal solution in a risk neutral setting, and equivalent both in optimal solution and value when a worst-case risk measure is used if a "relatively complete recourse property" is satisfied.

[1] Note that the selected values of $\mathbb{P}^I$ and $\mathbb{P}^{II}$ are such that $\max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[X]$ can be reinterpreted as the 75%-conditional value-at-risk of X when using the probability measure $\mathbb{P}(\omega_1) = 1 - \mathbb{P}(\omega_2) = 20\%$.

- Methodologically, we introduce the Δ-regret model for multi-stage stochastic programming under a discrete probability space. We show how this model can be evaluated over a continuum of Δ values and can be reformulated as a special class of two-stage robust linear program that is amenable to a rich range of solution schemes when the stochastic program is linear.
- Numerically, we investigate the effect of Δ and risk aversion on the conservatism of solutions proposed by the Δ-regret model in a simple newsvendor problem. We further illustrate the effect of enforcing different information look-ahead levels Δ on the regret experienced in a multi-period inventory management problem.

The rest of the paper is composed as follows. Section 2 reviews the relevant literature. Section 3 presents the Δ-regret model, an interpretation that gives rise to fractional regret model, and an illustrative example involving an inventory management problem. Section 4 presents our theoretical contributions and proposed solution scheme while Section 5 presents our numerical experiments.

2 Literature review

Since the first introduction by Savage (1951), regret minimization has been used in a wide range of applications including single-period portfolio selection (Lim et al. 2012), shortest path, subset selection (Natarajan et al. 2014), spanning tree, ranking problems (Audibert et al. 2014), and in pricing and mechanism design (Caldentey et al. 2017 and Koçyiğit et al. 2021) to name a few. Broadly speaking, the regret minimization models that are found in the literature can be classified based on three elements, e.g., the type of risk measure employed for regret evaluation, the length of the planning horizon, and the type of nonanticipativity constraint imposed on the benchmark policy.

In a single-stage setting, the majority of studies focus on the ex-post worst-case regret minimization problem (see for e.g. Feizollahi and Averbakh 2014, Furini et al. 2015 and Park et al. 2021), perhaps because of the ease of requiring only support information about the unknown parameters. Under the assumption of partial distribution information, Natarajan et al. (2014) study ex-post regret using a worst-case Conditional Value-at-Risk measure. Chen et al. (2006) exploit a similar regret model but in the context of a p-median problem. In these formulations, the benchmark policy can be seen as exploiting both the information about the distribution and about the realization itself. This is in sharp contrast with the ex-ante formulation that employs a worst-case risk measure (see Yue et al. (2006) and Perakis and Roels (2008) for an application to the newsvendor problem). Indeed, these works employ a worst-case expected value to measure regret and rather interpret it as the expected value of distribution information (EVDI), due to the following equivalence:

$$\max_{\boldsymbol{x}'\in\mathcal{X}} \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}\left[h(\boldsymbol{x}',\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x},\boldsymbol{\zeta}(\omega))\right] = \sup_{\mathbb{P}\in\mathcal{P}} \left(\left(\max_{\boldsymbol{x}'\in\mathcal{X}} \;\; \mathbb{E}_{\mathbb{P}}\left[h(\boldsymbol{x}',\boldsymbol{\zeta}(\omega))\right]\right) - \mathbb{E}_{\mathbb{P}}[h(\boldsymbol{x},\boldsymbol{\zeta}(\omega))]\right).$$

In the literature studying EVDI of the newsvendor problem, one can mention that Chen and Xie (2021) assume concurrent demand and supply randomness and Zhu et al. (2013) provide closed-form solutions for the relative EVDI. Other applications of EVDI can be found in blood classification (see El-Amine et al. (2018)) and portfolio optimization (see Lim et al. (2012) and Benati and Conde (2022)).

In the multi-stage setting, most studies focus on a two-stage setting under an ex-post worst-case regret minimization (see Bertsimas and Dunning (2020), and Poursoltani and Delage (2021) and references therein). Additionally, Xu et al. (2015) study a two-stage bidding problem in an electricity market, where perfect distribution information is assumed and different risk measures (namely Value-at-Risk, conditional Value-at-Risk, and expected value) are applied on the realized ex-post regret. Similar approaches were used in Zhang et al. (2020). Lim et al. (2006) investigate ex-ante and ex-post worst-case expected regret model in a fully multi-stage framework involving either an inventory management or a portfolio optimization problem. The authors derive analytical expressions for both the benchmark policies and regret minimizing policies and draw connections between regret minimization and Bayesian learning.

There has also been an interest in the economic literature to study the role that regret can play in a dynamic environment. For instance, Hayashi (2011) and Halpern and Leung (2016) study different forms of ex-post regret models and identify conditions under which regret minimizing policies are dynamically consistent. Alternatively, one can refer to Krähmer and Stone (2008) that considers a two-stage setting where the decision maker optimizes a trade-off between expected payoff and an weighted sum of the regret experienced at different point of time. The regret in each period is measured using an unconventional method that involves computing the effect of a one step deviation from an equilibrium policy that prescribes both how to adapt to revealed information and under any history of previous actions. Finally, Hayashi (2009) explore dynamic consistency and the role of ex-post regret in optimal stopping problems, while Strack and Viefers (2021) explores in the same application the effect of using stopping time to control the horizon over which the ex-post regret is measured.

This paper can be viewed to contribute to multi-stage regret theory from an optimization point of view. Indeed, we propose for the first time an intuitive risk-averse multi-stage regret minimization problem where the pessimism of the benchmark policy set is controlled using a bound on the maximum amount of look-ahead. This Δ-regret model naturally interpolates between the ex-ante and ex-post regret models. Furthermore, we explore the properties of Δ-regret model under popular risk measures and provide a promising direction for numerical resolution of these models, which is based on the recent advances in two-stage robust optimization. This later results can be seen as an interesting extension of Poursoltani and Delage (2021) to the multi-stage setting.

3 Δ-regret minimization in multi-stage stochastic programs

We consider a multi-stage decision making environment in which at each stage $t \in \{1, \dots, T\}$ a decision maker needs to make a decision $\boldsymbol{x}_t \in \mathbb{R}^n$ based on the available historical information captured by $[\boldsymbol{\zeta}_1 \ \boldsymbol{\zeta}_2 \ \cdots \ \boldsymbol{\zeta}_{t-1}]$. Focusing on a discrete probability space $(\Omega, \Sigma, \mathbb{Q})$, where $\mathbb{Q}$ is assumed strictly positive without loss of generality, one classical decision-making approach formulates the following multi-stage stochastic program:

$$(\mathcal{MSP}) \quad \min_{\boldsymbol{x} \in \mathcal{X} \cap \mathcal{X}_{na}} \quad \rho\big(-h(\boldsymbol{x}, \boldsymbol{\zeta})\big) \tag{1}$$

where $\boldsymbol{x} : \Omega \to \mathbb{R}^{n \times T}$ is the multi-stage policy, $\boldsymbol{\zeta} : \Omega \to \mathbb{R}^{m \times T-1}$ is the concatenated matrix of the random vectors observed over the whole horizon, $h(\boldsymbol{x}, \boldsymbol{\zeta})$ is the cumulative profit of implementing policy $\boldsymbol{x}$ when $\boldsymbol{\zeta}$ is realized, ρ is a convex risk measure that maps a random cost to a risk level, $\mathcal{X} := \{\boldsymbol{x} : \Omega \to \mathbb{R}^{n \times T} \,|\, \boldsymbol{x}(\omega) \in \mathcal{X}_\omega \,,\, \omega \in \Omega\}$, with bounded $\mathcal{X}_\omega \subseteq \mathbb{R}^{n \times T}$, imposes "physical" constraints that must be satisfied by the policy under each outcome in Ω, while $\mathcal{X}_{na}$ ensures that the policy is nonanticipative with respect to the information revealed by $\boldsymbol{\zeta}$. We formalize below some of these elements.

Definition 1. *The **set of nonanticipative policies** takes the form:*

$$\mathcal{X}_{na} := \left\{ \boldsymbol{x} : \Omega \to \mathbb{R}^{n \times T} \,\middle|\, \boldsymbol{x}_t(\omega) = \boldsymbol{x}_t(\omega'), \ \forall \omega, \omega' \in \Omega : \boldsymbol{\zeta}^{[t-1]}(\omega) = \boldsymbol{\zeta}^{[t-1]}(\omega'), \ \forall t \in \{1, 2, ..., T\} \right\},$$

where $[t] := \{1, \dots, t\}$, and where $\boldsymbol{\zeta}^{[0]}(\omega) = \boldsymbol{\zeta}^{[0]}(\omega')$ is interpreted as always true.

Definition 2. *According to Föllmer and Schied (2002), letting $\mathcal{L} : \{\xi : \Omega \to \mathbb{R}\}$ be the space of all possible finite random liabilities, ρ is a **convex risk measure** if and only if it satisfies:*

- *Monotonicity: $\forall \xi^1, \xi^2 \in \mathcal{L}$, $\xi^1 \geq \xi^2$ a.s. $\Rightarrow \rho(\xi^1) \geq \rho(\xi^2)$;*
- *Translation invariance: $\forall \xi \in \mathcal{L}$, $t \in \mathbb{R}$, $\rho(\xi + t) = \rho(\xi) + t$;*
- *Convexity: $\forall \xi^1, \xi^2 \in \mathcal{L}$, and $\theta \in [0,\ 1]$, $\rho(\theta\xi^1 + (1-\theta)\xi^2) \leq \theta\rho(\xi^1) + (1-\theta)\rho(\xi^2)$.*

*Moreover, ρ is considered a **coherent risk measure** if it further satisfies:*

- *Scale invariance: $\forall \xi \in \mathcal{L}$, $\alpha \geq 0$, $\rho(\alpha\xi) = \alpha\rho(\xi)$.*

In particular, it is well known that $\rho(-h(\boldsymbol{x}, \boldsymbol{\zeta})) = \mathbb{E}_{\mathbb{Q}}[-h(\boldsymbol{x}, \boldsymbol{\zeta})]$ and Conditional Value-at-Risk (see Example 3 for a definition) fall in the class of coherent risk measure. Unless specified otherwise, in what follows we will assume that ρ is a convex risk measure.

To improve computational tractability, we will later (when indicated) focus on the class of problems where constraints and objective function are affine with respect to $\boldsymbol{x}$.

Assumption 1. *[Stochastic Linear Programming] The profit function is an affine function of* $\boldsymbol{x}$ *defined as*

$$h(\boldsymbol{x},\boldsymbol{\zeta}) := \sum_{t=1}^{T} \boldsymbol{c}_t^\top(\boldsymbol{\zeta})\boldsymbol{x}_t + d(\boldsymbol{\zeta}), \tag{2}$$

for some arbitrary $\boldsymbol{c}_t : \mathbb{R}^{m\times(T-1)} \to \mathbb{R}^n$ *and* $d : \mathbb{R}^{m\times(T-1)} \to \mathbb{R}$*. Furthermore, for each* $\omega \in \Omega$, $\mathcal{X}_\omega$ *is a bounded polyhedron formulated as:*

$$\mathcal{X}_\omega := \left\{ \boldsymbol{x} \in \mathbb{R}^{n\times T} \,\middle|\, \sum_{t=1}^{T} \boldsymbol{a}_{jt}(\boldsymbol{\zeta}(\omega))^\top \boldsymbol{x}_t(\omega) \le b_j(\boldsymbol{\zeta}(\omega))\,,\, j = 1,2,...,\mathcal{J} \right\},$$

with arbitrary $\boldsymbol{a}_{jt} : \mathbb{R}^{m\times(T-1)} \to \mathbb{R}^n$ *and* $b_j : \mathbb{R}^{m\times(T-1)} \to \mathbb{R}$*, for all* j *and* t*.*

Recall that the regret models discussed in the introduction addressed a static decision model (namely with $T = 2$). Hence, the main difference between the ex-post and ex-ante models hinged on whether the benchmark action $\boldsymbol{x}'$ could fully anticipate or not realization $\boldsymbol{\zeta}$. A natural question to pose is therefore how the concept of regret extends in the multi-stage problems where we have $T > 2$ and where the values of $\boldsymbol{\zeta}$ are progressively revealed in time. In what follows, we propose a multi-stage regret minimization formulation that measures regret with respect to a benchmark policy that can exploit information revealed up to Δ stages into the future, which we term Δ-regret. This model effectively interpolates very naturally between the ex-ante (with $\Delta = 0$) and ex-post (with $\Delta = \infty$) models and effectively allows to study them under the same lens. Section 3.1 will present the Δ-regret model. Section 3.2 will describe an environment where Δ can be interpreted as a rational number (e.g. 1/2-regret model). Finally, Section 3.3 will present an illustrative example involving a multi-stage inventory management problem.

3.1 The Δ-regret model

In a multi-stage decision making problem, a regret-averse policy maker might be interested to compare his decisions to benchmark policies that exploit shorter foresight than the total planning horizon. This gives rise to the idea of Δ-regret model where the benchmark policies are capable of predicting the future realizations up to Δ steps ahead in the future. As an immediate result of such setting, the benchmark policies can adapt to the information released till time step $t + \Delta$. Assuming ρ is a convex risk measure, the Δ-regret model in the multi-stage setting is formulated as

$$(\Delta\text{-}\textit{regret}) \quad \underset{\boldsymbol{x}\in\mathcal{X}\cap\mathcal{X}_{na}}{\text{minimize}} \quad \mathcal{R}_\Delta(\boldsymbol{x}), \tag{3a}$$

where

$$\mathcal{R}_\Delta(\boldsymbol{x}) := \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \quad \rho(h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta})), \tag{3b}$$

and where $\mathcal{X}_\Delta$ is the space of policies that violate the nonanticipativity constraints by up to a margin of Δ steps. More specifically,

$$\mathcal{X}_\Delta := \left\{ \boldsymbol{x} : \Omega \to \mathbb{R}^{n\times T} \,\middle|\, \boldsymbol{x}_t(\omega) = \boldsymbol{x}_t(\omega'),\ \forall \omega,\omega' \in \Omega : \boldsymbol{\zeta}^{[t+\Delta-1]}(\omega) = \boldsymbol{\zeta}^{[t+\Delta-1]}(\omega'),\ \forall t \in \{1,2,...,T\} \right\},$$

where we interpret $\boldsymbol{\zeta}^{[t]} := \boldsymbol{\zeta}$ when $t \ge T-1$. For any $\Delta \in \{0,1,2,3,..,T-1\}$, the Δ-regret model will evaluate the regret of the prescribed decisions as contrasted with the ones that could have been made if the uncertain parameters were revealed up to Δ steps ahead of time. Clearly, when $\Delta = 0$, $\mathcal{X}_\Delta$ reduces to $\mathcal{X}_{na}$, implying that the benchmark policy has no access to any realization beforehand. On the contrary, $\Delta = T-1$ gives the benchmark policy full access to all the realizations of $\boldsymbol{\zeta}$ at any point of time. The Δ-regret model therefore naturally interpolates between the ex-ante and ex-post regret models. In addition, regret is a non-decreasing function of Δ. These concepts are formalized in the following lemma.

Lemma 1. *The* Δ*-regret model, i.e. problem* (3)*, reduces to ex-ante and ex-post regret minimization when* $\Delta = 0$ *and* $\Delta \ge T-1$ *respectively. Moreover, its optimal value is an increasing function of* Δ*.*

Proof. Proof. Clearly when $\Delta = 0$, we have that $\mathcal{X}_\Delta = \mathcal{X}_0 = \mathcal{X}_{na}$. So that the 0-regret model reduces to the *ex-ante* form:

$$\underset{\boldsymbol{x}\in\mathcal{X}\cap\mathcal{X}_{na}}{\text{minimize}} \quad \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_{na}} \rho(h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta}))\,,$$

and in particular to

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \max_{\boldsymbol{x}'\in\mathcal{X}} \rho(h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta}))\,,$$

when dealing with a static linear problem, i.e. $T = 2$ and $\boldsymbol{c}_2 = \boldsymbol{a}_{j2} = 0$ for all j.

Alternatively, when $\Delta \geq T-1$, by definition we have that $\mathcal{X}_\Delta = \{\boldsymbol{x} : \Omega \to \mathbb{R}^{n\times T}\}$, implying that:

$$\begin{aligned}\mathcal{R}_\Delta(\boldsymbol{x}) &= \max_{\boldsymbol{x}'\in\mathcal{X}} \rho\big(h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta})\big)\\ &= \rho\big(\max_{\boldsymbol{x}'\in\mathcal{X}_\omega} h(\boldsymbol{x}',\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega),\boldsymbol{\zeta}(\omega))\big),\end{aligned}$$

which follows from monotonicity of ρ. Specifically, we first have that for all $\boldsymbol{x}' \in \mathcal{X}\cap\mathcal{X}_T$:

$$h(\boldsymbol{x}'(\omega),\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega),\boldsymbol{\zeta}(\omega)) \leq \max_{\boldsymbol{x}''\in\mathcal{X}_\omega} h(\boldsymbol{x}'',\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega),\boldsymbol{\zeta}(\omega)) \quad \forall\omega\in\Omega.$$

Hence,

$$\rho(h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta})) \leq \rho(\max_{\boldsymbol{x}''\in\mathcal{X}_\omega} h(\boldsymbol{x}'',\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega),\boldsymbol{\zeta}(\omega)))\,.$$

On the other hand, we can define $\bar{\boldsymbol{x}}'(\omega) \in \arg\max_{\boldsymbol{x}'\in\mathcal{X}_\omega} h(\boldsymbol{x}',\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega),\boldsymbol{\zeta}(\omega))$, with $\bar{\boldsymbol{x}}' \in \mathcal{X}\cap\mathcal{X}_T$ to conclude that:

$$\begin{aligned}\rho\big(\max_{\boldsymbol{x}'\in\mathcal{X}_\omega} h(\boldsymbol{x}',\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega),\boldsymbol{\zeta}(\omega))\big) &= \rho\big(h(\bar{\boldsymbol{x}}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta})\big)\\ &\leq \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_T} \rho\big(h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta})\big)\\ &\leq \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_T} \rho\big(\max_{\boldsymbol{x}''\in\mathcal{X}_\omega} h(\boldsymbol{x}'',\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega),\boldsymbol{\zeta}(\omega))\big)\\ &= \rho\big(\max_{\boldsymbol{x}''\in\mathcal{X}_\omega} h(\boldsymbol{x}'',\boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega),\boldsymbol{\zeta}(\omega))\big),\end{aligned}$$

where the first inequality follows since $\bar{\boldsymbol{x}}' \in \mathcal{X}_T$. We can therefore conclude that the T-regret model reduces to the *ex-post* model:

$$\underset{\boldsymbol{x}\in\mathcal{X}\cap\mathcal{X}_{na}}{\text{minimize}} \quad \rho(\max_{\boldsymbol{x}'\in\mathcal{X}} h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta}))\,,$$

which takes the following form when the problem is static:

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{minimize}} \quad \rho(\max_{\boldsymbol{x}'\in\mathcal{X}} h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta}))\,.$$

Finally, we turn to establishing the monotonicity of the optimal value of problem (3). Let $\Delta \leq \Delta'$, then $\mathcal{X}_\Delta \subseteq \mathcal{X}_{\Delta'}$. This implies that:

$$\mathcal{R}_\Delta(\boldsymbol{x}) = \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \rho(h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta})) \leq \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_{\Delta'}} \rho(h(\boldsymbol{x}',\boldsymbol{\zeta}) - h(\boldsymbol{x},\boldsymbol{\zeta})) = \mathcal{R}_{\Delta'}(\boldsymbol{x})\,.$$

□

3.2 Extension to fractional Δ-regret model

In this section, we explore the possibility of formulating a Δ-regret minimization model where Δ is a rational number. This opportunity arises in a special family of multi-stage stochastic programs, which we refer to as an "overdiscretized" multi-stage program.

Definition 3. *Let $\mathcal{T}_d := \{1,\, H+1,\, 2H+1, \ldots, (D-1)H+1\}$, with $(D-1)H+1 = T$ be a set of evenly distributed decision moments over the horizon T. An $\mathcal{MSP}$ is called overdiscretized if $c_t(\boldsymbol{\zeta}) := 0$ and $d_t(\boldsymbol{\zeta}) := 0$ for all $t \notin \mathcal{T}_d$ and $\mathcal{X} \subseteq \mathcal{X}_d$, where*

$$\mathcal{X}_d := \{x \in \mathbb{R}^{n\times T} | x_t = 0\,, \forall t \notin \mathcal{T}_d\}\,.$$

Note that in an overdiscretized $\mathcal{MSP}$, the set $\mathcal{T}_d$ describes the only time points at which a decision is actually implemented. Intermediate time points $t \notin \mathcal{T}_d$ play the role of capturing how the information about the random process is being revealed to the decision maker between any two decisions. This type of model can for instance be relevant in an inventory management problem where production decisions can only be made once a day, yet orders are received continuously throughout the day. While this overdiscretized property is irrelevant from the point of view of classical risk minimization, we will see that it gives rise to the concept of fractional Δ-regret.

Indeed, from the point of view of identifying the optimal strategy for this $\mathcal{MSP}$, one can simplify the representation of the overdiscretized model by adapting the discretization to focus on pure decision moments. Specifically, the overdiscretized $\mathcal{MSP}$ model can be shown equivalent to:

$$(\widehat{\mathcal{MSP}}) \quad \min_{\hat{\boldsymbol{x}} \in \hat{\mathcal{X}} \cap \hat{\mathcal{X}}_{na}} \quad \rho\big(\hat{h}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{\zeta}})\big) \tag{4}$$

where $\hat{\boldsymbol{x}} : \Omega \to \mathbb{R}^{n \times D}$, $\hat{\boldsymbol{\zeta}} : \Omega \to \mathbb{R}^{(mH) \times D}$, and where

$$\hat{h}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{\zeta}}) := \sum_{d=1}^{D} \hat{\boldsymbol{c}}_d(\hat{\boldsymbol{\zeta}})^\top \hat{\boldsymbol{x}}_d + \hat{d}(\hat{\boldsymbol{\zeta}}), \tag{5}$$

with $\hat{\boldsymbol{c}}_d(\hat{\boldsymbol{\zeta}}) := \boldsymbol{c}_{(d-1)H+1}^\top(\boldsymbol{\zeta})$ and $\hat{d}(\hat{\boldsymbol{\zeta}}) := d(\boldsymbol{\zeta})$, while $\hat{\mathcal{X}}$ and $\hat{\mathcal{X}}_{na}$ are defined such that $\hat{\boldsymbol{x}} \in \hat{\mathcal{X}} \cap \hat{\mathcal{X}}_{na}$ if and only if

$$\boldsymbol{x}_t(\omega) := \begin{cases} \hat{\boldsymbol{x}}_d(\omega) & \text{if } t = (d-1)H + 1 \\ 0 & \text{otherwise.} \end{cases} \in \mathcal{X} \cap \mathcal{X}_{na}\,.$$

From the perspective of Δ-regret minimization however, one can consider that the Δ-regret model associated to $\mathcal{MSP}$ provides a Δ/H-regret model for $\widehat{\mathcal{MSP}}$. This is interesting given that it provides the means of constructing a continuum of Δ-regret model for all $\Delta \in \mathcal{Q}_+$. Namely, for any $\Delta := q/r$ with $q, r \in \mathcal{N}$, one should assemble an overdiscretized description of the $\mathcal{MSP}$ with $H := r$, and solve the q-regret model of this $\mathcal{MSP}$. We close with a remark that a common tool for assembling such an overdiscretized model is to assume an underlying continuous stochastic process for the vector of uncertainties: e.g. in an inventory management problem, one might assume that the demand for each product is independent and follows a Poisson process.

3.3 Illustrative example of Δ-regret model

We consider the multi-stage inventory management problem previously studied in Ben-Tal et al. (2004) and Kuhn et al. (2011). We assume that each period t consists of a day. The inventory system consists of I factories which produce a single item and store it at a shared warehouse. The production cost of a single unit of the item on day t at factory i is c_{it} and the objective is to determine the optimal production level of each factory for each (x_{it}) to satisfy the uncertain demand and minimize the total production cost over a planning horizon of T days. While $\bar{x}_{it}$ indicates the production capacity of factory i on day t, the maximum production potential over the whole planning horizon is determined by $\bar{x}_{i,tot}$. The minimum and the maximum inventory levels that should be maintained at the end of each day are denoted by $\underline{x}_{wh}$ and $\bar{x}_{wh}$, respectively, and x_{wh}^0 represents the initial inventory level. If $d_t(\omega) \in \mathbb{R}$ denotes the demand of day t under scenario ω, then we let $\zeta_{t-1}(\omega) := d_t(\omega) \in \mathbb{R}$ to model the fact that the demand for day t is known when deciding of the production levels at the beginning of the day: this occurs for instance when orders for pick up need to be made at the latest one day before pickup. The $\mathcal{MSP}$ for this inventory problem takes the form where:

$$h(\boldsymbol{x}, \boldsymbol{\zeta}) := -\sum_{t=1}^{T} \sum_{i=1}^{I} \boldsymbol{c}_{it}^\top \boldsymbol{x}_{it}(\omega),$$

and

$$\mathcal{X}_\omega := \left\{ \boldsymbol{x} \in \mathbb{R}^{n \times T} \;\middle|\; \begin{array}{l} 0 \leq x_{it} \leq \bar{x}_{it},\ \forall i \in \mathcal{I},\ \forall t \in \mathcal{T} \\ \sum_{t=1}^{T} x_{it} \leq \bar{x}_{i,tot},\ \forall i \in \mathcal{I} \\ \underline{x}_{wh} \leq x_{wh}^0 + \sum_{s=1}^{t} \sum_{i=1}^{I} x_{is} - \sum_{s=1}^{t-1} \zeta_s(\omega) - d_1 \leq \bar{x}_{wh},\ \forall t \in \mathcal{T} \end{array} \right\}.$$

We consider a simple instance of this problem with 2 factories ($I = 2$), 3 days planning horizon ($T = 3$) and 5 demand pattern scenarios ($\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$). In this setting, considering the scenario tree structure depicted in Figure 1, the nonanticipativity constraints are expressed as:

$$\mathcal{X}_{na} := \left\{ \boldsymbol{x} : \Omega \to \mathbb{R}^{2\times 3} \;\middle|\; \begin{array}{l} \boldsymbol{x}_1(\omega_1) = \boldsymbol{x}_1(\omega_2) = \boldsymbol{x}_1(\omega_3) = \boldsymbol{x}_1(\omega_4) = \boldsymbol{x}_1(\omega_5) \\ \boldsymbol{x}_2(\omega_1) = \boldsymbol{x}_2(\omega_2),\ \boldsymbol{x}_2(\omega_4) = \boldsymbol{x}_2(\omega_5) \end{array} \right\}.$$

When measuring regret, the policy maker might be interested in comparing her policy to one that benefits from the same information. This is an immediate implication of $\Delta = 0$ in the Δ-regret model. Setting Δ to 1 allows her to measure her regret with respect to the policy made under one stage look-ahead information. Eventually, $\Delta = 2$ compares to policies that exploit the full information. Specifically, we have the following reductions:

$$\mathcal{X}_0 = \mathcal{X}_{na}, \quad \mathcal{X}_1 = \left\{ \boldsymbol{x} : \Omega \to \mathbb{R}^{2\times 3} \;\middle|\; \boldsymbol{x}_1(\omega_1) = \boldsymbol{x}_1(\omega_2),\ \boldsymbol{x}_1(\omega_4) = \boldsymbol{x}_1(\omega_5) \right\} \quad \mathcal{X}_2 = \left\{ \boldsymbol{x} : \Omega \to \mathbb{R}^{2\times 3} \right\}.$$

Figure 1 illustrates each policy sets in a scenario tree. This example shows how increasing Δ lifts the constraints imposed on $\boldsymbol{x}'$ gradually and the full access of the realized scenario is bestowed upon $\boldsymbol{x}'$ when Δ is at its maximum value.

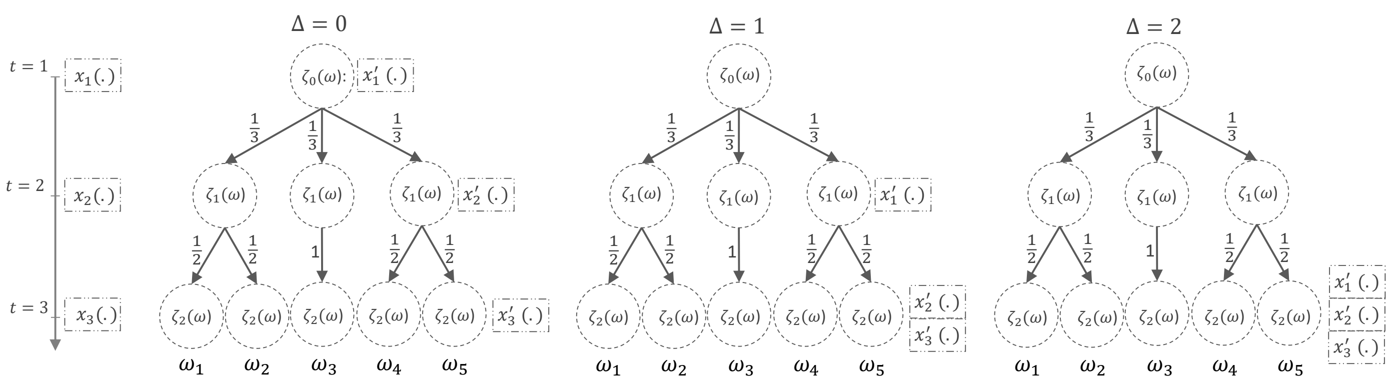

Figure 1: Comparison of adaptation power between $\boldsymbol{x}$ (beside the timeline) and $\boldsymbol{x}'$ (on the right of each tree) as a function of Δ. The nodes of the tree present what information is available at each point of time.

Now, let's instead consider that orders are received continuously throughout the day. In this context, the manager might consider that additional information about tomorrow's demand will be available in the middle of the day and could resent some regret for not having access to this information earlier. Optimizing this form of regret is possible by using the fractional regret model presented in Section 3.2. Figure 2 illustrates the corresponding policy sets associated to the $\Delta = 0.5$ and $\Delta = 1.5$ regret models. One should note that in this example $H = 0.5$, $D = 5$ such that $\mathcal{T}_d := \{1,\ 1.5,\ 2,\ 2.5,\ 3\}$.

4 Properties of Δ-regret model under risk measures

In this section, we explore interesting properties that arise under different choice of risk measures for the Δ-regret model. In particular, the first two subsections initially study the properties that emerge under specific coherent risk measures, namely the worst-case $\rho(\xi) = \operatorname{ess\,sup}(\xi)$ and expected value $\rho(\xi) = \mathbb{E}_{\mathbb{Q}}[\xi]$. We then consider the general class of coherent risk measures using their worst-case expectation representation, i.e. $\rho(\xi) := \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[\xi]$ (see Artzner et al. (1999)) We will show that under a worst-case risk measure, all Δ-regret models are equivalent if (and only if) a *relatively complete recourse* property is satisfied. This will also occur, yet only in terms of optimal solution set for models that employ an expected value. Finally, we will derive a reformulation for all coherent risk measures that take the form of a two-stage robust linear program when the stochastic program is linear and the risk measure linear programming representable.

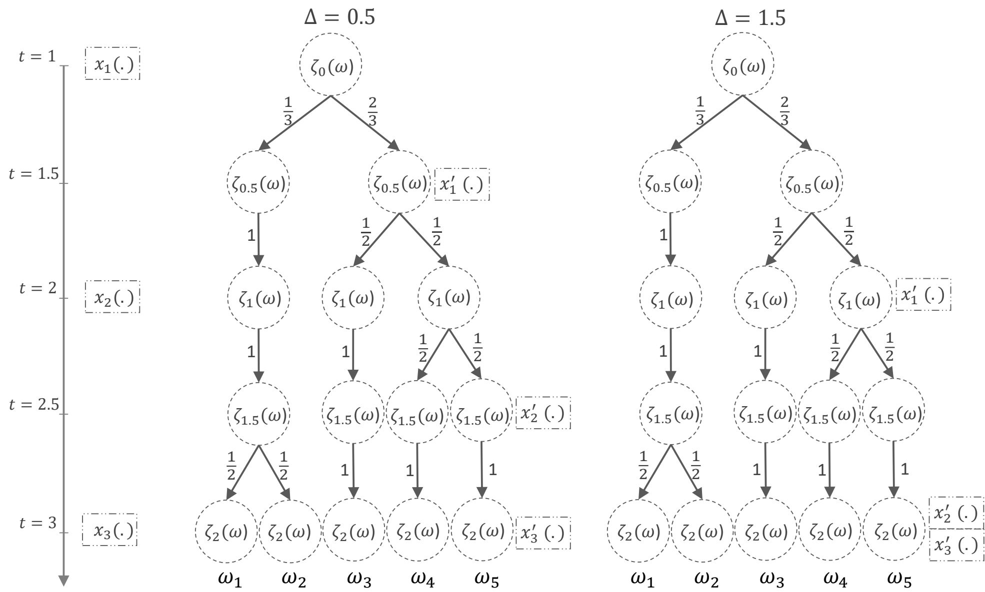

Figure 2: Comparison of adaptation power between $\boldsymbol{x}$ (beside the timeline) and $\boldsymbol{x}'$ (on the right of each tree) as a function of Δ in an overdiscretized $\mathcal{MSP}$. Note that the notation presents the indexing of the $\widehat{\mathcal{MSP}}$.

4.1 The case of $\rho(\xi) = \operatorname{ess\,sup}(\xi)$

In this section, we consider measuring the Δ-regret using the essential supremum as the risk measure:

$$\rho(\xi) := \operatorname{ess\,sup}(\xi) = \inf\left\{a \,|\, \mathbb{P}(\xi > a) = 0\right\}.$$

In particular, we will confirm conditions under which, the invariability of Δ-regret to Δ, observed in Example 1, holds. In order to present our main result, we first introduce an assumption about the $\mathcal{MSP}$.

Assumption 2. *The multi-stage stochastic program satisfies the* ***relatively complete recourse property****, i.e.,*

$$\mathcal{X}_{\omega}^{[t]} = \mathcal{X}_{\omega'}^{[t]}, \quad \forall(\omega, \omega') : \zeta^{[t-1]}(\omega) = \zeta^{[t-1]}(\omega'), \forall t,$$

where

$$\mathcal{X}_{\omega}^{[t]} := \{\boldsymbol{x} \in \mathbb{R}^{n \times t} | \exists \bar{\boldsymbol{x}} \in \mathbb{R}^{n \times T-t}, [\boldsymbol{x} \;\; \bar{\boldsymbol{x}}] \in \mathcal{X}_{\omega}\}$$

is a projection of $\mathcal{X}_{\omega}$ on the space spanned by the decision vectors $\boldsymbol{x}_1$, $\boldsymbol{x}_2$,... $\boldsymbol{x}_t$.

In simpler words, this assumption imposes that when looking at the set of feasible decisions $\boldsymbol{x}$ in hindsight, this set only includes candidates that had a probability one guarantee of being feasible at the time that they were implemented. While the decision to satisfy this assumption is an important modeling choice in designing the Δ-regret model and might affect the measured regret (see Example 2 below), it is in fact always possible to modify a multi-stage stochastic program so that the property is satisfied.

Lemma 2. *Given an $\mathcal{MSP}$, one can construct $\overline{\mathcal{MSP}}$ that produces the same optimal value and optimal solution set as $\mathcal{MSP}$ while satisfying the relatively complete recourse assumption, i.e. Assumption 2.*

Proof. Proof. Let $\overline{\mathcal{MSP}}$ be exactly the same as $\mathcal{MSP}$ except for $\mathcal{X}_{\omega}$ which is replaced with:

$$\bar{\mathcal{X}}_{\omega} := \cap_{t=1}^{T} \cap_{\omega' : \zeta^{[t-1]}(\omega) = \zeta^{[t-1]}(\omega')} \{\boldsymbol{x} \in \mathbb{R}^{n \times T} | \boldsymbol{x}_{1:t} \in \mathcal{X}_{\omega'}^{[t]}\}.$$

First, we can start by demonstrating that $\overline{\mathcal{MSP}}$ satisfies the relatively complete recourse property. Namely, for all t, if (ω, ω') is such that $\zeta^{[t-1]}(\omega) = \zeta^{[t-1]}(\omega')$, then:

$$\begin{aligned}
\bar{\mathcal{X}}_\omega^{[t]} &= \{\boldsymbol{x} \in \mathbb{R}^{n\times t} | \exists \bar{\boldsymbol{x}} \in \mathbb{R}^{n\times T-t}, [\boldsymbol{x}\ \ \bar{\boldsymbol{x}}] \in \bar{\mathcal{X}}_\omega\} \\
&= \{\boldsymbol{x} \in \mathbb{R}^{n\times t} | \exists \bar{\boldsymbol{x}} \in \mathbb{R}^{n\times T-t}, [\boldsymbol{x}\ \ \bar{\boldsymbol{x}}] \in \cap_{t=1}^T \cap_{\omega'':\zeta^{[t-1]}(\omega)=\zeta^{[t-1]}(\omega'')} \{\boldsymbol{x} \in \mathbb{R}^{n\times T} | \boldsymbol{x}_{1:t} \in \mathcal{X}_{\omega''}^{[t]}\}\} \\
&= \{\boldsymbol{x} \in \mathbb{R}^{n\times t} | \exists \bar{\boldsymbol{x}} \in \mathbb{R}^{n\times T-t}, [\boldsymbol{x}\ \ \bar{\boldsymbol{x}}] \in \cap_{t=1}^T \cap_{\omega'':\zeta^{[t-1]}(\omega')=\zeta^{[t-1]}(\omega'')} \{\boldsymbol{x} \in \mathbb{R}^{n\times T} | \boldsymbol{x}_{1:t} \in \mathcal{X}_{\omega''}^{[t]}\}\} \\
&= \{\boldsymbol{x} \in \mathbb{R}^{n\times t} | \exists \bar{\boldsymbol{x}} \in \mathbb{R}^{n\times T-t}, [\boldsymbol{x}\ \ \bar{\boldsymbol{x}}] \in \bar{\mathcal{X}}_{\omega'}\} = \bar{\mathcal{X}}_{\omega'}^{[t]}\,.
\end{aligned}$$

Next, we show that $\overline{\mathcal{MSP}}$ produces the same set of optimal solutions and optimal value as $\mathcal{MSP}$. In particular, one can show that $\mathcal{X} \cap \mathcal{X}_{na} = \bar{\mathcal{X}} \cap \mathcal{X}_{na}$ where $\bar{\mathcal{X}} := \{\boldsymbol{x} : \Omega \to \mathbb{R}^{n\times T} \,|\, \boldsymbol{x}(\omega) \in \bar{\mathcal{X}}_\omega\,,\, \omega \in \Omega\}$. First, since we have that:

$$\bar{\mathcal{X}}_\omega \subseteq \cap_{\omega':\zeta^{[T-1]}(\omega)=\zeta^{[T-1]}(\omega')} \{\boldsymbol{x} \in \mathbb{R}^{n\times T} | \boldsymbol{x}_{1:T} \in \mathcal{X}_{\omega'}^{[T]}\} \subseteq \mathcal{X}_\omega\,,$$

we can conclude that $\mathcal{X} \cap \mathcal{X}_{na} \supseteq \bar{\mathcal{X}} \cap \mathcal{X}_{na}$. Alternatively, we have that for all $\boldsymbol{x} \in \mathcal{X} \cap \mathcal{X}_{na}$, one can confirm that $\boldsymbol{x} \in \bar{\mathcal{X}}$, i.e. $\boldsymbol{x}(\omega) \in \bar{\mathcal{X}}_\omega$ for all ω. Specifically, fixing any ω, any t, and any ω' that satisfies $\zeta^{[T-1]}(\omega) = \zeta^{[T-1]}(\omega')$, we can check that $\boldsymbol{x}_{1:t}(\omega) \in \mathcal{X}_{\omega'}^{[t]}$ since

$$[\boldsymbol{x}_{1:t}(\omega)\ \ \boldsymbol{x}_{t+1:T}(\omega')] = [\boldsymbol{x}_{1:t}(\omega')\ \ \boldsymbol{x}_{t+1:T}(\omega')] \in \mathcal{X}_{\omega'}\,,$$

where we used the fact that $\boldsymbol{x} \in \mathcal{X}_{na}$ which implies that $\boldsymbol{x}_{1:t}(\omega) = \boldsymbol{x}_{1:t}(\omega')$.

□

We can now turn to the main result of this section which indicates that relatively complete recourse is a necessary and sufficient condition for the Δ-regret model to be invariant to Δ under the essential supremum risk measure.

Theorem 1. *Given that Assumption 2 holds and $\rho(\xi) = \operatorname{ess\,sup}(\xi)$, for any arbitrary $\Delta \geq 0$, the objective function of problem* (3) *reduces to*

$$\mathcal{R}_\Delta(\boldsymbol{x}) = \mathcal{R}_{T-1}(\boldsymbol{x}) = \max_{\omega\in\Omega} \max_{\boldsymbol{x}'\in\mathcal{X}_\omega} h(\boldsymbol{x}', \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega)). \tag{6}$$

Proof. Proof. The argument goes as follows:

$$\begin{aligned}
\mathcal{R}_\Delta(\boldsymbol{x}) &= \sup_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \operatorname{ess\,sup}(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})) && \text{(7a)} \\
&\leq \sup_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_{T-1}} \operatorname{ess\,sup}(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})) && \text{(7b)} \\
&\leq \sup_{\boldsymbol{x}'\in\mathcal{X}} \operatorname{ess\,sup}(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})) && \text{(7c)} \\
&\leq \operatorname{ess\,sup}(\sup_{\boldsymbol{x}'\in\mathcal{X}_\omega} h(\boldsymbol{x}', \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega))) && \text{(7d)} \\
&= \max_{\omega\in\Omega} \sup_{\boldsymbol{x}'\in\mathcal{X}_\omega} h(\boldsymbol{x}', \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega)) && \text{(7e)} \\
&= \sup_{\boldsymbol{x}'\in\mathcal{X}_{\omega^*}} h(\boldsymbol{x}', \boldsymbol{\zeta}(\omega^*)) - h(\boldsymbol{x}(\omega^*), \boldsymbol{\zeta}(\omega^*)) && \text{(7f)} \\
&\leq \sup_{\boldsymbol{x}'\in\mathcal{X}_{\omega^*}} \operatorname{ess\,sup}(h(\pi(\omega; \boldsymbol{x}', \omega^*), \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega))) && \text{(7g)} \\
&\leq \sup_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_{na}} \operatorname{ess\,sup}(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})) && \text{(7h)} \\
&\leq \sup_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \operatorname{ess\,sup}(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})) && \text{(7i)} \\
&= \mathcal{R}_\Delta(\boldsymbol{x})\,, && \text{(7j)}
\end{aligned}$$

where (7b) follows from Lemma 1 and (7c) results from relaxing the constraint that $\boldsymbol{x}' \in \mathcal{X}_{T-1}$. (7d) follows from the monotonicity of ess sup and the fact that $h(\boldsymbol{x}'(\omega), \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega)) \leq \sup_{\boldsymbol{x}'\in\mathcal{X}_\omega} h(\boldsymbol{x}', \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega))$, almost surely. (7e) follows from the fact that $\mathbb{Q}(\omega) \geq 0$ for all $\omega \in \Omega$ and Ω is finite. In

(7f) we define ω^* has any maximizer of (7e). In (7g), we let $\pi(\cdot;\boldsymbol{x}',\omega^*)$ be a nonanticipative policy which implements $\boldsymbol{x}'$ under outcome ω^* while implementing an arbitrarily chosen feasible action at each time point for all other outcomes, e.g.:

$$\pi_t(\omega;\boldsymbol{x}',\omega^*) := \begin{cases} \boldsymbol{x}'_t & \text{if } \zeta(\omega)^{[t-1]} = \zeta(\omega^*)^{[t-1]} \\ \arg\min_{\bar{\boldsymbol{x}}_t:[\pi_{[t-1]}(\omega;\boldsymbol{x}',\omega^*)^\top \;\; \bar{\boldsymbol{x}}_t]^\top \in \mathcal{X}_\omega^{[t]}} \|\bar{\boldsymbol{x}}_t\|_2 & \text{otherwise} \end{cases} \quad \forall t. \tag{8}$$

The fact that this policy exists and is in $\mathcal{X}\cap\mathcal{X}_{na}$ is due to Assumption 2 (proof below) and motivates (7h). Finally, (7i) follows from the fact that $\mathcal{X}_{na}\subseteq\mathcal{X}_\Delta$.

We finalize this proof by providing more details about the three facts regarding $\pi_t(\omega;\boldsymbol{x}',\omega^*)$. First, this policy exists since we can construct it from $t=1,\ldots,T$ with the guarantee that the arg min in (8) is non-empty given that for all t and all $\omega\in\Omega$:

$$\boldsymbol{\zeta}^{[t-1]}(\omega) = \boldsymbol{\zeta}^{[t-1]}(\omega^*) \Rightarrow \boldsymbol{\zeta}^{[t'-1]}(\omega) = \boldsymbol{\zeta}^{[t'-1]}(\omega^*),\ \forall 1\le t'\le t \Rightarrow \pi_{[t]}(\omega;\boldsymbol{x}',\omega^*)\in\mathcal{X}_\omega^{[t]},$$

while iteratively, from $t=2$ to $t=T$, and for all $\omega\in\Omega$:

$$\begin{aligned}
&\pi_{[t-1]}(\omega;\boldsymbol{x}',\omega^*)\in\mathcal{X}_\omega^{[t-1]} \;\;\&\;\; \boldsymbol{\zeta}^{[t-1]}(\omega)\neq\boldsymbol{\zeta}^{[t-1]}(\omega^*)\\
&\Rightarrow \exists\bar{\boldsymbol{x}}\in\mathbb{R}^{n\times T-t+1}, [\pi_{[t-1]}(\omega;\boldsymbol{x}',\omega^*)\;\;\bar{\boldsymbol{x}}]\in\mathcal{X}_\omega\\
&\Rightarrow \exists\bar{\boldsymbol{x}}_t\in\mathbb{R}^{n\times 1},\ [\pi_{[t-1]}(\omega;\boldsymbol{x}',\omega^*)\;\;\bar{\boldsymbol{x}}_t]\in\mathcal{X}_\omega^{[t]}\\
&\Rightarrow \arg\min_{\bar{\boldsymbol{x}}_t:[\pi_{[t-1]}(\omega;\boldsymbol{x}',\omega^*)^\top\;\;\bar{\boldsymbol{x}}_t]^\top\in\mathcal{X}_\omega^{[t]}} \|\bar{\boldsymbol{x}}_t\|_2\in\mathcal{X}_\omega^{[t]}\\
&\Rightarrow \pi_{[t]}(\omega;\boldsymbol{x}',\omega^*)\in\mathcal{X}_\omega^{[t]}\,,
\end{aligned}$$

where we first employ the definition of $\pi_{[t-1]}(\omega;\boldsymbol{x}',\omega^*)\in\mathcal{X}_\omega^{[t-1]}$, and then confirmed that the first vector of the $\bar{x}$ matrix could be used to create a member of $\mathcal{X}_\omega^{[t]}$.

Now regarding $\pi(\cdot;\boldsymbol{x}',\omega^*)\in\mathcal{X}$, this is necessary the case as we just showed that $\pi(\cdot;\boldsymbol{x}',\omega^*)\in\mathcal{X}_\omega^{[t]}$ for all t and $\omega\in\Omega$. Hence, $\pi(\cdot;\boldsymbol{x}',\omega^*)\in\mathcal{X}_\omega^{[T]}=\mathcal{X}_\omega$ for all ω. Furthermore, $\pi(\cdot;\boldsymbol{x}',\omega^*)\in\mathcal{X}_{na}$ by construction. Namely, for any t, if $\boldsymbol{\zeta}(\omega)^{[t-1]}=\boldsymbol{\zeta}(\omega')^{[t-1]}=\boldsymbol{\zeta}(\omega^*)^{[t-1]}$, then $\pi_t(\cdot;\boldsymbol{x}',\omega^*)=\boldsymbol{x}'_t$. Alternatively, for any (ω,ω') such that $\boldsymbol{\zeta}(\omega)^{[t-1]}=\boldsymbol{\zeta}(\omega')^{[t-1]}\neq\boldsymbol{\zeta}(\omega^*)^{[t-1]}$, we can exploit the fact that:

$$\boldsymbol{\zeta}^{[t-1]}(\omega)=\boldsymbol{\zeta}^{[t-1]}(\omega')\Rightarrow\boldsymbol{\zeta}^{[t'-1]}(\omega)=\boldsymbol{\zeta}^{[t'-1]}(\omega'),\ \forall 1\le t'\le t\,,$$

so that iteratively from $t'=2$ to $t'=t$, given that:

$$\pi_{t'-1}(\omega;\boldsymbol{x}',\omega^*)=\pi_{t'-1}(\omega';\boldsymbol{x}',\omega^*)$$

and that Assumption 2 implies that $\mathcal{X}_\omega^{[t']}=\mathcal{X}_\omega^{[t]}$, then necessarily

$$\begin{aligned}
\pi_{t'}(\omega;\boldsymbol{x}',\omega^*) &= \arg\min_{\bar{\boldsymbol{x}}_{t'}:[\pi_{[t'-1]}(\omega;\boldsymbol{x}',\omega^*)^T\;\;\bar{\boldsymbol{x}}_t]^T\in\mathcal{X}_\omega^{[t']}}\|\bar{\boldsymbol{x}}_{t'}\|_2\\
&= \arg\min_{\bar{\boldsymbol{x}}_{t'}:[\pi_{[t'-1]}(\omega';\boldsymbol{x}',\omega^*)^T\;\;\bar{\boldsymbol{x}}_t]^T\in\mathcal{X}_{\omega'}^{[t']}}\|\bar{\boldsymbol{x}}_{t'}\|_2=\pi_{t'}(\omega';\boldsymbol{x}',\omega^*)\,.
\end{aligned}$$

□

At first glance, the result of Theorem 1 looks intuitive since essential supremum hedges against a single worst-case scenario. Thus imposing a nonanticipative structure on $\boldsymbol{x}'\in\mathcal{X}_\Delta$ will have no effect for any value of Δ. What is less intuitive is the role of Assumption 2. In this regard, the following example supports and illustrates our claims that the relatively complete recourse property is necessary to obtain this invariance, and that $\mathcal{MSP}$ can always be reformulated to satisfy this assumption.

Example 2. *Consider a simple two-stage (i.e. $T = 2$) problem, where the set of first and second stage actions are defined as $A = \{a_1, a_2, a_3\}$ and $B = \{b_1, b_2\}$, respectively. After implementing the first stage decision, the decision maker is faced with two scenarios ω_1 or ω_2 with 10% and 90% chances respectively. We consider the following definition for $\mathcal{X}$:*

$$\mathcal{X} := \{\boldsymbol{x} := \{\omega_1, \omega_2\} \to A \times B | \boldsymbol{x}(\omega_1) \in \mathcal{X}_{\omega_1},\ \boldsymbol{x}(\omega_2) \in \mathcal{X}_{\omega_2}\},$$

with:

$$\mathcal{X}_{\omega_1} := \{(a_1, b_1), (a_3, b_1), (a_3, b_1)\}$$
$$\mathcal{X}_{\omega_2} := \{(a_2, b_2), (a_3, b_1), (a_3, b_2)\}$$

In words, if the decision maker chooses a_1, he can react to ω_1 with b_1 but has no feasible recourse against ω_2. The reverse is true for a_2, while a_3 enables both the b_1 and b_2 actions under ω_1 and ω_2 respectively. The profit function, defined only over feasible pairs, takes the form described in Table 3

Actions	$h(\boldsymbol{x}, \boldsymbol{\zeta}(\omega_1))$	$h(\boldsymbol{x}, \boldsymbol{\zeta}(\omega_2))$
(a_1, b_1)	4	-
(a_2, b_2)	-	1
(a_3, b_1)	3	-
(a_3, b_2)	-	0

Table 3: Profit function in Example 2

In this example, $\mathcal{X}_{\omega_1}^{[1]} = \{a_1, a_3\} \neq \mathcal{X}_{\omega_2}^{[1]} = \{a_2, a_3\}$, which indicates that Assumption 2 is violated. Furthermore there is only one feasible policy for the decision maker, i.e. $\{\bar{\boldsymbol{x}}\} = \{(a_3,\ b_1\mathbf{1}\{\omega = \omega_1\} + b_2\mathbf{1}\{\omega = \omega_2\})\} = \mathcal{X} \cap \mathcal{X}_{na}$. Focusing on this policy, one can compute the Δ-regret under the essential supremum measure as follows. In the case of $\Delta = 1$, the feasible space for the benchmark policy becomes $\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_1 = \mathcal{X} = \{\bar{\boldsymbol{x}},\ \bar{\boldsymbol{x}}'\}$ with

$$\bar{\boldsymbol{x}}' := (a_1\mathbf{1}\{\omega = \omega_1\} + a_2\mathbf{1}\{\omega = \omega_2\},\ b_1\mathbf{1}\{\omega = \omega_1\} + b_2\mathbf{1}\{\omega = \omega_2\}).$$

We can thus conclude that:

$$\begin{aligned}\mathcal{R}_1(\bar{\boldsymbol{x}}) &= \max_{\boldsymbol{x}' \in \{\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}'\}} \operatorname{ess\,sup}(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\bar{\boldsymbol{x}}, \boldsymbol{\zeta}))\\ &= \max(\operatorname{ess\,sup}(h(\bar{\boldsymbol{x}}, \boldsymbol{\zeta}) - h(\bar{\boldsymbol{x}}, \boldsymbol{\zeta})),\ \operatorname{ess\,sup}(h(\bar{\boldsymbol{x}}', \boldsymbol{\zeta}) - h(\bar{\boldsymbol{x}}, \boldsymbol{\zeta})))\\ &= \max(\operatorname{ess\,sup}(0),\ \operatorname{ess\,sup}(1)) = 1\,.\end{aligned}$$

On the other hand, if $\Delta = 0$, we have that $\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_0 = \mathcal{X} \cap \mathcal{X}_{na} = \{\bar{\boldsymbol{x}}\}$, i.e. the benchmark decision must be chosen among the same sets of decision as for the decision maker. This naturally leads to $\mathcal{R}_0(\bar{\boldsymbol{x}}) = 0$. We therefore showed that when Assumption 2 is violated, it is possible that $0 = \mathcal{R}_0(\boldsymbol{x}) \neq \mathcal{R}_1(\boldsymbol{x}) = 1$.

We close this example with the observation that if the $\mathcal{MSP}$ was modified as proposed in Lemma 2, then we would have:

$$\begin{aligned}\bar{\mathcal{X}}_{\omega_1} &:= \{\boldsymbol{x} \in A \times B | x_1 \in \mathcal{X}_{\omega_1}^{[1]}\} \cap \{\boldsymbol{x} \in A \times B | x_1 \in \mathcal{X}_{\omega_2}^{[1]}\} \cap \mathcal{X}_{\omega_1}\\ &= \{\boldsymbol{x} \in A \times B | x_1 \in \mathcal{X}_{\omega_1}^{[1]} \cap \mathcal{X}_{\omega_2}^{[1]}\} \cap \mathcal{X}_{\omega_1} \ = \ \{\boldsymbol{x} \in A \times B | x_1 \in \{a_3\}\} \cap \mathcal{X}_{\omega_1} = \{(a_3, b_1)\}\end{aligned}$$

while

$$\bar{\mathcal{X}}_{\omega_2} := \{\boldsymbol{x} \in A \times B | x_1 \in \mathcal{X}_{\omega_1}^{[1]}\} \cap \{\boldsymbol{x} \in A \times B | x_1 \in \mathcal{X}_{\omega_2}^{[1]}\} \cap \mathcal{X}_{\omega_2} = \{(a_3, b_2)\}\,.$$

Using $\bar{\mathcal{X}} := \{\boldsymbol{x} : \Omega \to A \times B \,|\, \boldsymbol{x}(\omega) \in \bar{\mathcal{X}}_\omega\,,\ \omega \in \Omega\}$ instead of $\mathcal{X}$ does not affect the solution of the $\mathcal{MSP}$ since in any case $\mathcal{X} \cap \mathcal{X}_{na} = \{\bar{\boldsymbol{x}}\} = \bar{\mathcal{X}} = \bar{\mathcal{X}} \cap \mathcal{X}_{na}$. Yet, using $\bar{\mathcal{X}}$ instead of $\mathcal{X}$ does affect the Δ-regret model given that under $\bar{\mathcal{X}}$ now we have that $\mathcal{R}_1(\bar{\boldsymbol{x}}) = \mathcal{R}_0(\bar{\boldsymbol{x}}) = 0$. Intuitively, from the regret perspective the difference between the two models reduces to whether the benchmark policy is allowed to implement decisions that don't have a probability one guarantee to lead to long term feasibility.

4.2 The case of $\rho(\xi) = \mathbb{E}_{\mathbb{Q}}[\xi]$

For a given probability measure $\mathbb{Q}$, the expected value can be considered as another option among the popular risk measures. However, for any value of Δ problem (3) produces the same optimal solution as $\mathcal{MSP}$. This is formalized in the following proposition.

Proposition 1. *Given that $\rho(\xi) := \mathbb{E}_{\mathbb{Q}}[\xi]$, for any arbitrary Δ, problems* (1) *and* (3) *have the same set of optimal solutions as the $\mathcal{MSP}$.*

Proof. Proof. The argument goes as follows:

$$\mathcal{R}_\Delta(\boldsymbol{x}) = \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_\Delta} \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})] \tag{9a}$$

$$= \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_\Delta} \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}', \boldsymbol{\zeta})] - \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}, \boldsymbol{\zeta})] \tag{9b}$$

$$= g(\Delta) - \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}, \boldsymbol{\zeta})] \tag{9c}$$

with $g(\Delta) := \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_\Delta} \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}', \boldsymbol{\zeta})]$, and where (9b) follows from linearity of the risk measure. The fact that Δ only affects the constant $g(\Delta)$ in (9c) allows us to conclude that the optimal solution sets of problem (3) is unaffected by Δ. We also observe that:

$$\underset{\boldsymbol{x} \in \mathcal{X} \cap \mathcal{X}_{na}}{\text{minimize}}\, \mathcal{R}_\Delta(\boldsymbol{x}) \equiv \underset{\boldsymbol{x} \in \mathcal{X} \cap \mathcal{X}_{na}}{\text{maximize}}\, \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}, \boldsymbol{\zeta})] - g(\Delta) \tag{10}$$

hence the Δ-regret model has the same optimal solution set as $\mathcal{MSP}$ when $\rho(\xi) = \mathbb{E}_{\mathbb{Q}}[\xi]$. $\square$

While Proposition 1 establishes that all Δ-regret models produce the same optimal solution under a risk neutral setting, we will see in our numerical experiments that optimal values do change for different values of Δ. Interestingly, in the case of $\Delta = 0$, one can confirm that a risk neutral decision maker never experiences regret if she acts optimally.

Corollary 1. *The optimal value of problem* (3) *with $\Delta = 0$ and $\rho(\xi) := \mathbb{E}_{\mathbb{Q}}[\xi]$ is equal to zero and achieved by the $\mathcal{MSP}$ solution.*

Proof. Proof. This follows from the fact that $\mathcal{X}_0 = \mathcal{X}_{na}$, hence replacing $g(\Delta) := \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_{na}} \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}', \boldsymbol{\zeta})]$ in equation (10) leads to:

$$\begin{aligned}\underset{\boldsymbol{x} \in \mathcal{X} \cap \mathcal{X}_{na}}{\text{minimize}}\, \mathcal{R}_\Delta(\boldsymbol{x}) &= \max_{\boldsymbol{x} \in \mathcal{X} \cap \mathcal{X}_{na}} \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}, \boldsymbol{\zeta})] - g(\Delta)\\ &= \max_{\boldsymbol{x} \in \mathcal{X} \cap \mathcal{X}_{na}} \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}, \boldsymbol{\zeta})] - \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_{na}} \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}', \boldsymbol{\zeta})] = 0\end{aligned}$$

$\square$

In particular, Proposition 1 and Corollary 1 suggest that the regret experienced by a decision maker can be decomposed into three positive components

$$\begin{aligned}\mathcal{R}_\Delta(\boldsymbol{x}) &= \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_\Delta} \rho\left(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right)\\ &= \left(\max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_\Delta} \rho\left(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right) - \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_0} \rho\left(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right)\right)\\ &\quad + \left(\max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_0} \rho\left(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right) - \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_0} \mathbb{E}_{\mathbb{Q}}\left[h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right]\right)\\ &\quad + \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_0} \mathbb{E}_{\mathbb{Q}}\left[h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right]\end{aligned}$$

$$= \left(\max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_\Delta} \rho\left(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right) - \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_0} \rho\left(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right)\right) \tag{11a}$$

$$+ \left(\max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_0} \rho\left(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right) - \max_{\boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_0} \mathbb{E}_{\mathbb{Q}}\left[h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\right]\right) \tag{11b}$$

$$+ \mathbb{E}_{\mathbb{Q}}[h(\boldsymbol{x}^*, \boldsymbol{\zeta})] - \mathbb{E}_{\mathbb{Q}}\left[h(\boldsymbol{x}, \boldsymbol{\zeta})\right] . \tag{11c}$$

The first component (11a) captures the part of the regret which comes from the information that is out of the decision maker reach. The second component (11b) captures a part of the regret that comes from risk aversion of the decision maker. Finally, the last component (11c) comes from not being optimal with respect to the risk neutral version of $\mathcal{MSP}$.

4.3 The case of coherent risk measures

In a more general context, it is well-known (see Artzner et al. (1999)) that any coherent risk measure can be represented using a worst-case expectation formulation :

$$\rho(\xi) := \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[\xi],$$

where $\mathcal{P}$ is a non-empty convex set of probability measures that, in the distributionally robust optimization literature, is also referred as an ambiguity set known to contain the true underlying measure $\mathbb{Q}$.

Definition 4. *The ambiguity set $\mathcal{P}$ is a bounded convex set which implies that*

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[\xi] = \max_{\boldsymbol{p}\in\mathcal{D}\cap\mathcal{M}} \sum_{\omega\in\Omega} p_\omega \xi(\omega)$$

where $\xi : \Omega \to \mathbb{R}$, $\mathcal{M} \subset \mathbb{R}^{|\Omega|}$ is the simplex set, and $\mathcal{D} \subset \mathbb{R}^{|\Omega|}$ denotes a general convex and compact set.

Taking advantage of Definition 4, problem (3) can be cast as a two-stage robust optimization problem. This is formalized in the following proposition.

Proposition 2. *Given $\rho(\xi) := \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[\xi]$, problem* (3) *reduces to*

$$\begin{aligned}
&\underset{\boldsymbol{x}\in\mathcal{X}\cap\mathcal{X}_{na}}{\text{minimize}} \quad \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \min_{r,\boldsymbol{v}} \quad r && \text{(12a)}\\
&\qquad\qquad s.t. \quad \delta^*(\boldsymbol{v}|\mathcal{D}) - v_\omega + h(\boldsymbol{x}'(\omega), \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega)) \leq r, \ \forall\omega \in \Omega, && \text{(12b)}
\end{aligned}$$

where $r \in \mathbb{R}$, $\boldsymbol{v} \in \mathbb{R}^{|\Omega|}$ and $\delta^*(\boldsymbol{v}|\mathcal{D}) := \sup_{\boldsymbol{p}\in\mathcal{D}} \boldsymbol{p}^T\boldsymbol{v}$ represents the support function of $\mathcal{D}$.

Proof. Proof. Taking advantage of worst-case expectation risk measure, problem (3) leads to

$$\begin{aligned}
\mathcal{R}_\Delta(\boldsymbol{x}) &= \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \ \rho\big(h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})\big)\\
&= \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \max_{\boldsymbol{p}\in\mathcal{D}\cap\mathcal{M}} \sum_{\omega\in\Omega} p_\omega h(\boldsymbol{x}'(\omega), \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega))\,.
\end{aligned}$$

Using epigraph variable r, this can be alternatively rewritten as

$$\begin{aligned}
\mathcal{R}_\Delta(\boldsymbol{x}) = \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \min_{r} \quad & r\\
\text{s.t.} \quad & \sum_{\omega\in\Omega} p_\omega h(\boldsymbol{x}'(\omega), \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega)) \leq r, \ \forall\boldsymbol{p} \in \mathcal{D}\cap\mathcal{M}\\
& r \in \mathbb{R}.
\end{aligned} \tag{13}$$

Letting $g(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{p}) := \sum_{\omega\in\Omega} p_\omega h(\boldsymbol{x}'(\omega), \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega))$ and applying Theorem 2 in Ben-Tal et al. (2015), we get:

$$\begin{aligned}
\max_{\boldsymbol{p}\in\mathcal{D}\cap\mathcal{M}} g(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{p}) &= \inf_{\boldsymbol{v}} \delta^*(\boldsymbol{v}|\mathcal{D}) - \inf_{\boldsymbol{p}\in\mathcal{M}} \ \boldsymbol{p}^\top\boldsymbol{v} - g(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{p}) && \text{(14a)}\\
&= \inf_{\boldsymbol{v}} \delta^*(\boldsymbol{v}|\mathcal{D}) - \inf_{\boldsymbol{p}:\,\boldsymbol{p}\geq 0,\, \sum_{\omega\in\Omega}\boldsymbol{p}_\omega = 1} \ \sum_{\omega\in\Omega} \boldsymbol{p}_\omega\big(v_\omega - h(\boldsymbol{x}'(\omega), \boldsymbol{\zeta}(\omega)) + h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega))\big) && \text{(14b)}\\
&= \inf_{\boldsymbol{v}} \delta^*(\boldsymbol{v}|\mathcal{D}) - \min_{\omega\in\Omega} \ v_\omega - h(\boldsymbol{x}'(\omega), \boldsymbol{\zeta}(\omega)) + h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega)) && \text{(14c)}\\
&= \inf_{\boldsymbol{v}} \max_{\omega\in\Omega} \ \delta^*(\boldsymbol{v}|\mathcal{D}) - v_\omega + h(\boldsymbol{x}'(\omega), \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}(\omega), \boldsymbol{\zeta}(\omega)), && \text{(14d)}
\end{aligned}$$

where $\boldsymbol{v} \in \mathbb{R}^{|\Omega|}$ and (14c) follows from the fact that searching over worst-case distribution is indeed searching over the worst-case outcome. Plugging this result back into equation (13) leads to the two-stage optimization model presented in (12). $\square$

In general, problem (12) is a non-linear two-stage robust optimization problem. However, under a number of popular ambiguity sets, support function of $\delta^*(v|\mathbb{P})$ renders a linear programming representation which in turn makes (12) a robust linear two-stage program. Such choices include the sets associated to Conditional Value-at-Risk or expectiles (see Bellini and Bernardino (2017)), and, in the DRO literature, some type-1 Wasserstein ambiguity sets (see Mohajerin Esfahani and Kuhn (2018)) or some sets based on hypothesis testing (see Bertsimas et al. (2018)).

Corollary 2. *Given that Assumption 1 is satisfied and that* $\mathcal{D} := \{\boldsymbol{p} \in \mathbb{R}^{|\Omega|} \,|\, \exists\, \boldsymbol{q} \in \mathbb{R}^{n_q}, B_p\boldsymbol{p} + B_q\boldsymbol{q} \leq \boldsymbol{b}\}$*, where* $B_p \in \mathbb{R}^{m\times|\Omega|}$*,* $B_q \in \mathbb{R}^{m\times n_q}$*,* $\boldsymbol{b} \in \mathbb{R}^m$*, then problem* (12) *reduces to the following robust two-stage linear optimization problem:*

$$\begin{aligned}
\min_{\boldsymbol{x}\in\mathcal{X}\cap\mathcal{X}_{na}} \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \min_{r,\,\boldsymbol{v},\,\boldsymbol{\lambda}} \quad & r \\
s.t. \quad & \boldsymbol{\lambda}^\top \boldsymbol{b} - v_\omega + \sum_{t=1}^{T} \boldsymbol{c}_t^\top(\boldsymbol{\zeta}(\omega))(\boldsymbol{x}'_t(\omega) - \boldsymbol{x}_t(\omega)) \leq r, \ \forall \omega \in \Omega \\
& B_p^\top \boldsymbol{\lambda} = \boldsymbol{v} \\
& B_q^\top \boldsymbol{\lambda} = 0 \\
& r \in \mathbb{R},\, v \in \mathbb{R}^{|\Omega|},\, \boldsymbol{\lambda} \in \mathbb{R}^m_+ .
\end{aligned} \tag{15}$$

Proof. Proof. We define

$$\mathcal{D}' := \{\boldsymbol{p}' \in \mathbb{R}^{|\Omega|+n_q} \,|\, \exists\, \boldsymbol{p} \in \mathbb{R}^{|\Omega|},\, \boldsymbol{q} \in \mathbb{R}^{n_q}, \boldsymbol{p}' = [\boldsymbol{p}^\top\ \boldsymbol{q}^\top]^\top,\, B\boldsymbol{p}' \leq \boldsymbol{b}\}$$

where $B := [B_p \ \ B_q]$ so that $\mathcal{D} := \{\boldsymbol{p} \in \mathbb{R}^{n_p} \,|\, \exists\, \boldsymbol{p}' \in \mathcal{D}', \boldsymbol{p} = A\boldsymbol{p}'\}$, where $A := [I \ \ 0]$. Since $\mathcal{D}$ is an affine projection $\mathcal{D}'$, we have that

$$\delta^*(\boldsymbol{v}|\mathcal{D}) = \sup_{\boldsymbol{p}':B\boldsymbol{p}'\leq\boldsymbol{b}} \boldsymbol{v}^\top A\boldsymbol{p}' = \inf_{\boldsymbol{\lambda}\geq 0:A^\top\boldsymbol{v}=B^\top\boldsymbol{\lambda}} \boldsymbol{b}^\top\boldsymbol{\lambda}\,,$$

where we exploited strong LP duality theory, given that $\mathcal{D}$, and implicitly $\mathcal{D}'$, is non-empty. After replacing B by using its definition and reintegrating the infimum operation in constraint (12b) we get problem (15). $\square$

Example 3. *Conditional Value-at-Risk (CVaR) evaluates the conditional expectation of the random variable* ξ *under* $\alpha\%$ *worst scenarios and mathematically takes the form of*

$$CVaR_\alpha(\xi) := \inf_t \ t + \frac{1}{1-\alpha}\mathbb{E}_{\bar{\boldsymbol{p}}}[\max(0, -\xi - t)], \tag{16}$$

where $\bar{\boldsymbol{p}}$ *denotes the reference probability distribution. It has the following worst-case expectation representation (see Rockafellar et al. (2006)):*

$$CVaR_\alpha(\xi) := \sup_{\boldsymbol{p}\in\mathcal{D}\cap\mathcal{M}} \sum_{\omega\in\Omega} p_\omega X(\omega), \tag{17}$$

where $\mathcal{D} := \{\boldsymbol{p} \in \mathbb{R}^{|\Omega|} \,|\, \boldsymbol{p} \leq \bar{\boldsymbol{p}}/(1-\alpha)\}$*, where* $\bar{p}_\omega := \mathbb{Q}(\omega)$*. Based on Corollary 2, we get the following two-stage linear program:*

$$\min_{\boldsymbol{x}\in\mathcal{X}\cap\mathcal{X}_{na}} \max_{\boldsymbol{x}'\in\mathcal{X}\cap\mathcal{X}_\Delta} \min_{r,\,\boldsymbol{v}} \quad r \tag{18a}$$

$$s.t. \quad \frac{\bar{\boldsymbol{p}}^\top\boldsymbol{v}}{1-\alpha} - v_\omega + \sum_{t=1}^{T} \boldsymbol{c}_t^\top(\boldsymbol{\zeta}(\omega))(\boldsymbol{x}'_t(\omega) - \boldsymbol{x}_t(\omega)) \leq r, \ \forall \omega \in \Omega \tag{18b}$$

$$r \in \mathbb{R},\, \boldsymbol{v} \in \mathbb{R}^{|\Omega|}_+ . \tag{18c}$$

Solving problem (15) in polynomial time is likely to be in general impossible as the class of two-stage robust linear optimization problems is known to be NP-hard, see Ben-Tal et al. (2004). However, there is a variety of exact and approximate solution schemes that can practically solve problem (15), see (Poursoltani and Delage 2021, Section 3.1). One such exact approach is the column-and-constraint generation algorithm proposed in Zeng and Zhao (2013). The procedure expresses the inner minimization problem through its KKT conditions which in turn can be expressed as a mixed-integer linear program (MILP), solving the resulting min-max problem using an iterative scheme. Such a procedure can be directly applied to problem (15) which we briefly discuss in Appendix B. We remark that the effectiveness of the column-and-constraint generation algorithm heavily relies on the the ability to efficiently solve the inner MILP. However, the McCormick inequalities which are typically used to linearize the KKT conditions lead to weak linear relaxations for the inner MILP. When using the Conditional Value-at-Risk, one can strengthen the mixed-integer formulation by taking advantage both of structure of the ambiguity set, which leads to improved computational efficiency. This strengthened formulation is also discussed in Appendix B.

We finish this section by remarking that alternative two-stage robust formulation exists for problems (12) and (15) as suggested by Bertsimas and de Ruiter (2016) for the two-stage linear programming case and de Ruiter et al. (2018) for the two-stage non-linear case. This could be of practical interest for problems (12). Under Assumption 1, $\mathcal{X} \cap \mathcal{X}_\Delta$ is a polyhedral set while in general constraint (12b) is a non-linear convex constraint due to the support function $\delta^*(\boldsymbol{v}|\mathcal{D})$. As suggested by de Ruiter et al. (2018), one can dualize the inner minimization problem in problems (12), interchanging the order of the maximization over $x' \in \mathcal{X} \cap \mathcal{X}_\Delta$ and the newly introduced dual variables λ, and finally dualizing the new inner maximization over $x' \in \mathcal{X} \cap \mathcal{X}_\Delta$. The resulting two-stage problem will have linear first and second stage constraints, while all non-linear constraint will now appear in the maximization problem. Such reformulation could open up new avenues for exact and approximate solution approaches for regret minimization problems.

5 Numerical experiments

In this section, we conduct two numerical studies to provide numerical insights on the relationship between the amount of look-ahead (Δ) allowed for the benchmark policy, the level of risk aversion on the solution quality, and the level of regret. To this end, Section 5.1 studies a single-stage single-item newsvendor problem, which reduces to either the ex-post (EP-RAR) or ex-ante (EA-RAR) problems discussed in the introduction. By comparing the two regret models as well as the solution resulting from $\mathcal{MSP}$ we aim to shed light on the conditions in which each method produces the least conservative solution. In Section 5.2, we consider the multi-period inventory problem discussed in Section 3.3. The multi-stage nature of the problem allows us to benchmark multiple values of Δ including fractional values and measure the level of regret. In addition, we examine the breakdown of regret as this is expressed in (11). In both studies, we measure risk using Conditional Value-at-Risk as it allows to easily control the level of risk aversion with a single parameter, namely α. Solutions to the regret models that appear in each study can be obtained using the column-and-constraint generation algorithm discussed in Appendix B.

5.1 Newsvendor problem

In this section, we consider a single-item newsvendor problem in which $h(x,\zeta) = p\min(x,\zeta) - cx$ denotes the profit function, $p > 0$ is the sales price, $c < p$ is the ordering cost, and ζ is the random demand following a discrete uniform distribution with 10000 scenarios taking values in $[8, 12]$. The ordering decision x is constrained in $\mathcal{X} = [8, 12]$. Using $\rho(\xi) = \mathrm{CVaR}_\alpha(\xi)$, we solve problem (3) for $\Delta = 0$ and $\Delta = 1$, as well as the corresponding CVaR minimization problem (1).

To compare the ordering decisions from the three utility functions we consider instances where $(p, c) \in \{(10, 1), (10, 5), (10, 9)\}$. Figure 3 plots the optimal regret and Figure 4 plots the optimal ordering levels as a function of α. We observe the following: (i) For $\alpha = 0$, all models reduce to minimizing expected loss, thus all models produce the same optimal solution, as suggested by Proposition 1, despite that the optimal regret is different. (ii) For $\alpha = 1$, both $\Delta = 0$ and $\Delta = 1$ produce the same optimal ordering decision. This is not surprising as the risk measure reduces to $\rho(\xi) = \mathrm{CVaR}_1(\xi) = \operatorname{ess\,sup}(\xi)$ and from Theorem 1 we know that both regret models will achieve the same optimal regret, see Figure 3. Note that in general it is possible that the two regret models produce different optimal solutions, however, this is not the case in this one-dimensional example. (iii) Although all models result in the same ordering levels for $\alpha = 0$, as the value of α increases both Δ-regret models result in significantly larger ordering levels compared to the CVaR minimization problem. If we interpret the behavior of the decision maker to be "less conservative" if her order is large, then the results provide further evidence that regret minimization models provide less conservative solutions than CVaR minimization. Similar behavior was observed when other demand distribution were considered such as symmetric and skewed Beta distributions (not presented in this paper). (iv) The ordering decisions for both Δ models is the same for $p = 2c$ due to the symmetric nature of the distributions. For $p > 2c$, $\Delta = 1$ is less conservative while the reverse is true for the $p < 2c$ case. This can be intuitively interpreted as follows: when the profit margin $p - c$ is large, see $(p, c) = (10, 1)$, then for $\Delta = 1$ under-ordering will induce a large regret as the benchmark decision is chosen after observing the demand and due to the large profit margin over producing will induce lower regret. $\Delta = 0$ on the other hand, produces slightly more conservative ordering decisions as the benchmark decision is chosen before the demand realizes. In contrast, the reverse behavior is true if profit margin $p - c$ is small, see $(p, c) = (10, 9)$. The ordering decisions from both Δ models is significantly lower as the profit margin is small with $\Delta = 0$ producing less conservative decisions than with $\Delta = 1$. (v) It is interesting to observe that for $(p, c) = (10, 1)$ and $(p, c) = (10, 9)$ both Δ-regret models have the same optimal regret. This is due to the fact that the demand follows a uniform distribution which is symmetric.

To better understand the optimal ordering levels from the Δ-regret models, we now take a look at the behavior of the benchmark ordering levels x'. For $\Delta = 1$, as the benchmark policy can adapt to the demand, the optimal ordering decisions is $x'^* = \zeta$. Thus its decision is not affected by the α level and as consequence the ordering level x^* is also constant across all α values. For $\Delta = 0$, for any $x \in \mathcal{X}$ problem (3b) is in general non-convex. For the optimal x^* and $\alpha > 0$ problem (3b) has two global maximizers depicted in red in Figure 5. Therefore, the benchmark policy can be thought as a randomized policy. For the optimal x^* we can thus equivalently express (3b) as

$$\mathcal{R}_\Delta(x^*) = \sup_{F \in \mathcal{F}(x_1'^*, x_2'^*)} \mathbb{E}_{x' \sim F} \Big[\rho(h(x', \zeta) - h(x^*, \zeta) | x') \Big] \tag{19}$$

where $\mathcal{F}(x_1'^*, x_2'^*)$ is the family of all Bernoulli distributions supported on the two global maximizers of problem (3b). Using the optimizer F^* of (19), we can now construct $\mathbb{E}_{F^*}(x'^*)$ which is depicted in blue in Figure 5. The downwards sloping behavior of $\mathbb{E}_{F^*}(x'^*)$ as α increases in $(p, c) = (10, 1)$ partly explains the downward sloping behavior of the ordering decision x^* in the range $\alpha \in (0, 0.5)$ as the decision maker is trying to minimize her regret compared to the benchmark policy thus orders less. However, in the range $\alpha \in (0.5, 1]$ the decision maker is more concerned about the worst-case behavior of the regret thus her ordering decision is less affected by the probability the adversary assigns to each of its actions. The reverse is true for $(p, c) = (10, 9)$.

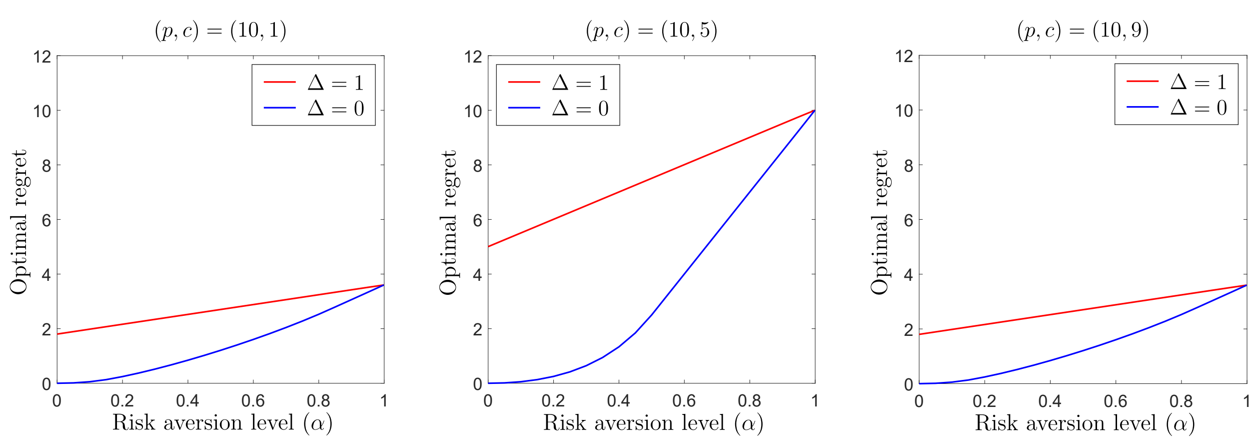

Figure 3: Optimal Regret for different value of Δ, α and (p, c).

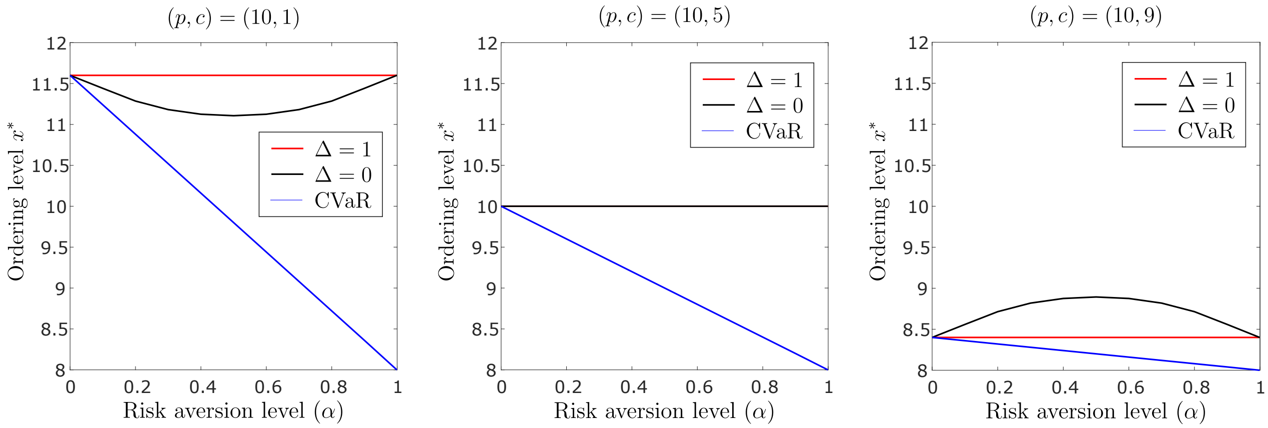

Figure 4: Optimal ordering levels for Δ-regret and CVaR minimization using a discrete Uniform demand distribution.

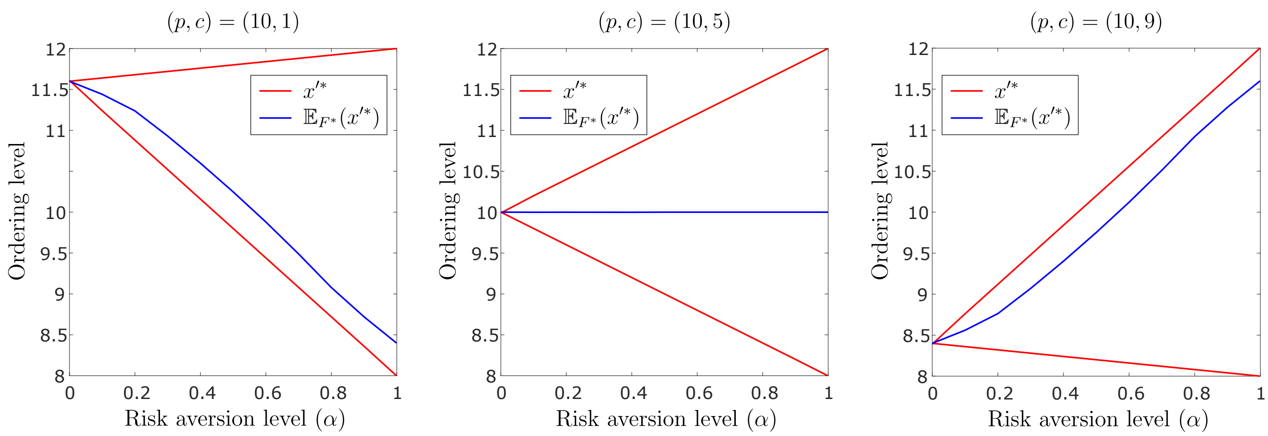

Figure 5: Optimal adversarial ordering levels. For each α the two global optimizers of (3b) are given in red, while blue depicts the mean $\mathbb{E}_{F^*}(x'^*)$ of the adversarial randomized policy.

5.2 Multi-period inventory management problem

In this section, we revisit the multi-stage inventory management problem described in Section 3.3 and consider an instance with 3 production facilities and a three day horizon. We assume that the demand information for the next day can be progressively collected in the middle and on the next morning; however, the production can only be planned at the beginning of each day at which point the full demand of the day is known. Overall, the outcome space includes 16 scenarios ($|\Omega| = 16$). The scenario tree that describes the evolution of the random demand, demand realizations and other parameters of the inventory model are presented in Appendix C. The experiments are conducted over the risk aversion levels of $\alpha = n/16$, where $n \in \{0, 1, ..., 16\}$. We remark that although we present results for a single instance, when tested on other randomly generated instances the insights from the results presented were qualitatively the same.

The first experiment compares the optimal Δ-regret values for $\Delta \in \{0, 0.5, 1, 1.5, 2\}$ at different levels of risk aversion α. The results are presented in Figure 6. Looking at the figure, one can remark that for any fixed look-ahead level, increasing the risk aversion level leads to an increased minimal Δ-regret. This originates from the fact that the Conditional Value-at-Risk only considers the worst-case $(1 - \alpha) \cdot 100\%$ of scenarios, e.g. while at $\alpha = 0\%$ it incorporates all the scenarios, at $\alpha = 100\%$ it measures the regret with respect to the worst-case scenario. On the other hand, when fixing the risk aversion level, the results demonstrate an increase in the minimal regret as the look-ahead level Δ increases from 0 to 2. This is in line with fact that as Δ increases, we are gradually relaxing the nonanticipativity constraints imposed on

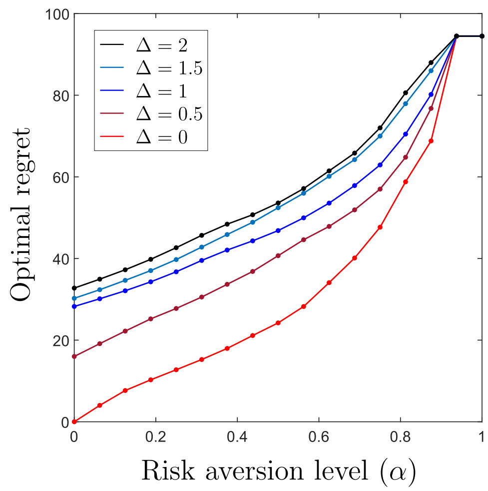

Figure 6: Optimal Δ-regret using Conditional Value-at-Risk at different risk aversion level α

the set of benchmark policies, as suggested in Lemma 1. Moreover, we can observe that as the risk aversion level reaches 100%, i.e., $\rho(\xi) = \mathrm{CVaR}_1(\xi) = \operatorname{ess\,sup}(\xi)$, all regret models achieve the same regret, verifying Theorem 1 which stated that all models are equivalent under the essential supremum risk measure. At the opposite side of the graph, one sees that the minimal risk for the regret model with $\Delta = 0$ converges to zero as predicted by Corollary 1.

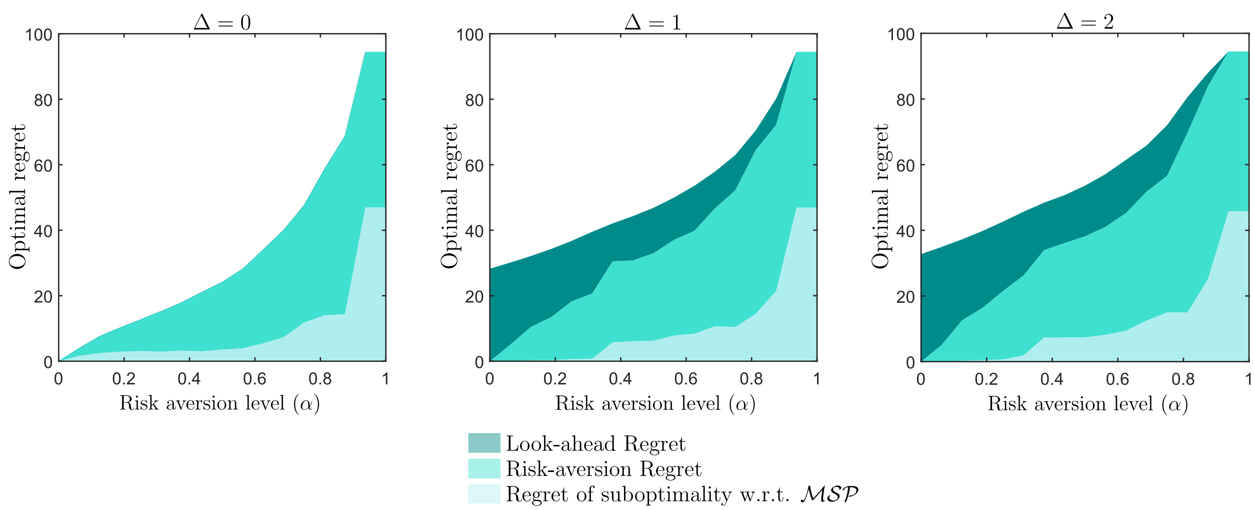

Figure 7: Regret breakdown for different Δ-regret optimal solutions.

The second set of results presents a breakdown of regret as this is expressed in equation (11). To this end, using the optimal solutions of $\Delta \in \{0, 1, 2\}$ we evaluate the three expression in (11) which can be interepretted as the "look ahead regret" (11a), "risk-aversion regret" (11b) and "regret of being suboptimal w.r.t. $\mathcal{MSP}$" (11c). Figure 7 presents the cumulative breakdown of the regret. As expected, for $\alpha = 0\%$, the "risk-aversion regret" is zero for all Δ by definition, while for $\alpha = 100\%$ the "look ahead regret" is zero for all Δ as suggested by Theorem 1. In fact, the "look ahead regret" decreases as α increases. For $\Delta = 0$, by definition the "look ahead regret" is zero indicating that regret is a combination of the "risk-aversion regret" and "regret of being suboptimal w.r.t. $\mathcal{MSP}$". It is interesting to observe that the "regret of being suboptimal w.r.t. $\mathcal{MSP}$" is relatively low in the range of $\alpha \in [0, 0.6]$ but constitutes roughly half of the total regret when $\alpha = 1$.

References

Aissi, H., Bazgan, C., and Vanderpooten, D. (2009). Min-max and min-max regret versions of combinatorial optimization problems: A survey. *European Journal of Operational Research*, 197(2):427–438.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, 9(3):203–228.

Audibert, J., Bubeck, S., and Lugosi, G. (2014). Regret in online combinatorial optimization. *Mathematics of Operations Research*, 39(1):31–45.

Bellini, F. and Bernardino, E. D. (2017). Risk management with expectiles. *The European Journal of Finance*, 23(6):487–506.

Ben-Tal, A., den Hertog, D., and Vial, J.-P. (2015). Deriving robust counterparts of nonlinear uncertain inequalities. *Mathematical Programming*, 149(1-2):265–299.

Ben-Tal, A., Goryashko, A., Guslitzer, E., and Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376.

Benati, S. and Conde, E. (2022). A relative robust approach on expected returns with bounded cvar for portfolio selection. *European Journal of Operational Research*, 296(1):332–352.

Bertsimas, D. and de Ruiter, F. (2016). Duality in two-stage adaptive linear optimization: Faster computation and stronger bounds. *INFORMS Journal on Computing*, 28(3):500–511.

Bertsimas, D. and Dunning, I. (2020). Relative robust and adaptive optimization. *INFORMS Journal on Computing*, 32(2):408–427.

Bertsimas, D., Gupta, V., and Kallus, N. (2018). Robust sample average approximation. *Mathematical Programming*, 171(1):217–282.

Caldentey, R., Liu, Y., and Lobel, I. (2017). Intertemporal pricing under minimax regret. *Operations Research*, 65(1):104–129.

Chen, G., Daskin, M. S., Shen, Z.-J. M., and Uryasev, S. (2006). The α-reliable mean-excess regret model for stochastic facility location modeling. *Naval Research Logistics*, 53(7):617–626.

Chen, Z. and Xie, W. (2021). Regret in the newsvendor model with demand and yield randomness. *Production and Operations Management*, 30(11):4176–4197.

de Ruiter, F., Zhen, J., and den Hertog, D. (2018). Dual approach for two-stage robust nonlinear optimization. *Available on Optimization-Online*.

El-Amine, H., Bish, E. K., and Bish, D. R. (2018). Robust postdonation blood screening under prevalence rate uncertainty. *Operations Research*, 66(1):1–17.

Feizollahi, M. J. and Averbakh, I. (2014). The robust (minmax regret) quadratic assignment problem with interval flows. *INFORMS Journal on Computing*, 26(2):321–335.

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and stochastics*, 6(4):429–447.

Furini, F., Iori, M., Martello, S., and Yagiura, M. (2015). Heuristic and exact algorithms for the interval min-max regret knapsack problem. *INFORMS Journal on Computing*, 27(2):392–405.

Halpern, J. Y. and Leung, S. (2016). Minimizing regret in dynamic decision problems. *Theory and Decision*, 81(1):123–151.

Hayashi, T. (2009). Stopping with anticipated regret. *Journal of Mathematical Economics*, 45(7-8):479–490.

Hayashi, T. (2011). Context dependence and consistency in dynamic choice under uncertainty: the case of anticipated regret. *Theory and decision*, 70(4):399–430.

Koçyiğit, Ç., Rujeerapaiboon, N., and Kuhn, D. (2021). Robust multidimensional pricing: Separation without regret. *Mathematical Programming*, Advance online publication.

Krähmer, D. and Stone, R. (2008). Regret in dynamic decision problems. Technical report.

Kuhn, D., Wiesemann, W., and Georghiou, A. (2011). Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, 130(1):177–209.

Kusuoka, S. (2001). On law invariant coherent risk measures. In *Advances in Mathematical Economics*, pages 83–95. Springer Japan, Tokyo.

Lim, A. E., Shanthikumar, J. G., and Vahn, G.-Y. (2012). Robust portfolio choice with learning in the framework of regret: Single-period case. *Management Science*, 58(9):1732–1746.

Lim, A. E. B., Shanthikumar, J. G., and Shen, Z. J. M. (2006). Model uncertainty, robust optimization, and learning. In *Models, Methods, and Applications for Innovative Decision Making*, pages 66–94. INFORMS TutORials in Operations Research (INFORMS, Cantonsville, MD).

McCormick, G. P. (1983). Nonlinear programming; theory, algorithms, and applications. Technical report.

Mohajerin Esfahani, P. and Kuhn, D. (2018). Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1):115–166.

Natarajan, K., Shi, D., and Toh, K.-C. (2014). A probabilistic model for minmax regret in combinatorial optimization. *Operations Research*, 62(1):160–181.

Park, J., El-Amine, H., and Mutlu, N. (2021). An exact algorithm for large-scale continuous nonlinear resource allocation problems with minimax regret objectives. *INFORMS Journal on Computing*, 33(3):1213–1228.

Perakis, G. and Roels, G. (2008). Regret in the newsvendor model with partial information. *Operations Research*, 56(1):188–203.

Postek, K., Ben-Tal, A., den Hertog, D., and Melenberg, B. (2018). Robust optimization with ambiguous stochastic constraints under mean and dispersion information. *Operations Research*, 66(3):814–833.

Poursoltani, M. and Delage, E. (2021). Adjustable robust optimization reformulations of two-stage worst-case regret minimization problems. *Operations Research*, Forthcoming.

Rockafellar, R. T., Uryasev, S., and Zabarankin, M. (2006). Generalized deviations in risk analysis. *Finance and Stochastics*, 10(1):51–74.

Savage, L. (1951). The theory of statistical decision. *Journal of the American Statistical association*, 46(253):55–67.

Strack, P. and Viefers, P. (2021). Too proud to stop: Regret in dynamic decisions. *Journal of the European Economic Association*, 19(1):165–199.

Xu, Z., Hu, Z., Song, Y., and Wang, J. (2015). Risk-averse optimal bidding strategy for demand-side resource aggregators in day-ahead electricity markets under uncertainty. *IEEE Transactions on Smart Grid*, 8(1):96–105.

Yue, J., Chen, B., and Wang, M.-C. (2006). Expected value of distribution information for the newsvendor problem. *Operations research*, 54(6):1128–1136.

Zeng, B. and Zhao, L. (2013). Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters*, 41(5):457–461.

Zhang, R., Li, G., Jiang, T., Chen, H., Li, X., Pei, W., and Xiao, H. (2020). Incorporating production task scheduling in energy management of an industrial microgrid: A regret-based stochastic programming approach. *IEEE Transactions on Power Systems*, 36(3):2663–2673.

Zhu, Z., Zhang, J., and Ye, Y. (2013). Newsvendor optimization with limited distribution information. *Optimization Methods & Software*, 28(3):640–667.

A Illustrative Example

Table 4 provides detailed calculations of obtaining the optimal decision for the project selection problem, described in Example 1, when the ex-post worst-case regret measure is exploited. Once the decision maker chooses project A ($x = x_A$), she will face either scenario ω_1 or ω_2, resulting in a payoff of 1\$ or 6\$, respectively. Having full access to the realized scenario, the optimal benchmark decision consists of choosing project B ($x' = x_B$) under ω_1 with 5\$ payoff and picking project A ($x' = x_A$) under ω_2 with 6\$ payoff, leading to a 4\$ regret for the decision maker in the first case and zero regret in the second one; consequently, the worst-case regret of choosing project A will be 4\$. Performing the same analysis for $x = x_B$ and $x = x_C$ brings about the worst-case regrets of 4\$ and 3\$, respectively. As a conclusion, aiming at minimizing the worst-case regret, the decision maker finds project C with 3\$ worst-case regret as her best option.

Table 4: EP-WCR

$\boldsymbol{x}$	ω	$\boldsymbol{x}'$	$\{.\}^\dagger$	$\sup_{\omega\in\Omega}\{.\}$
Project A	ω_1	Project B	4	4
	ω_2	Project A	0	
Project B	ω_1	Project B	0	4
	ω_2	Project A	4	
$\boldsymbol{x^*} =$ Project C	ω_1	Project B	1	**3**
	ω_2	Project A	3	
$\min_{\boldsymbol{x}\in\mathcal{X}}\sup_{\omega\in\Omega}\{.\}$				**3**

† $\{.\} := \left\{\max_{\boldsymbol{x}'\in\mathcal{X}} h(\boldsymbol{x}', \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}, \boldsymbol{\zeta}(\omega))\right\}$

Putting emphasis on risk-aversion, Table 5 clarifies the details of getting the optimal decisions for the ex-post worst-case expected regret minimization problem. In this setting, similar to the EP-WCR case, the benchmark policy has full access to the future scenario realizations, and as an immediate result, always selects project B under ω_1 and project A under ω_2. If the manager picks project A for investment, withdrawing the distributional information, the felt regret consists of 4\$ and 0\$ for ω_1 and ω_2 realizations, respectively. However, in contrast to EP-WCR, in EP-RAR with $\rho(\xi) := \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[\xi]$ the expected regret is measured with respect to the worst $\mathbb{P}$ from $\mathcal{P}$. Since the $\mathbb{P}^I$ and $\mathbb{P}^{II}$ lead to expected regrets of 3.2\$ and 0\$, the worst-case expected regret of choosing project A equals 3.2\$. Replicating the same analysis for $x = x_B$ and $x = x_C$ gives rise to worst-case expected regrets of 4\$ and 3\$. As a consequence, investment on project C with 3\$ worst-case expected regret will be the optimal choice.

An alternative to EP-RAR consists in EA-RAR, where the benchmark has no longer access to the realized scenario and only knows the true distribution. Table 6 summarizes the analysis for this problem. To be consistent with previous analysis, once again, we elaborate the details of getting the worst-case

Table 5: EP-RAR

$\boldsymbol{x}$	$\mathcal{P}$	$\boldsymbol{x}'^*$		$\mathbb{E}_{\mathbb{P}}[.]^\dagger$	$\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}_{\mathbb{P}}[.]$
		ω_1	ω_2		
Project A	$\mathbb{P}^I$	Project B	Project A	3.2	3.2
	$\mathbb{P}^{II}$	Project B	Project A	0	
Project B	$\mathbb{P}^I$	Project B	Project A	0.8	4
	$\mathbb{P}^{II}$	Project B	Project A	4	
$\boldsymbol{x^*}$ = Project C	$\mathbb{P}^I$	Project B	Project A	1.4	**3**
	$\mathbb{P}^{II}$	Project B	Project A	3	
$\min_{\boldsymbol{x}\in\mathcal{X}}\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}_{\mathbb{P}}[.]$					**3**

† $\mathbb{E}_{\mathbb{P}}[.] := \mathbb{E}_{\mathbb{P}}\left[\max_{\boldsymbol{x}'\in\mathcal{X}} h(\boldsymbol{x}', \boldsymbol{\zeta}(\omega)) - h(\boldsymbol{x}, \boldsymbol{\zeta}(\omega))\right]$

expected regret of choosing project A. In this case, the immediate payoff under ω_1 and ω_2 will be 1\$ and 6\$, respectively. Subsequently, the benchmark decision can be made after evaluating the expected regret of each of the three possible options ($x' = x_A$, x_B or x_C) under $\mathbb{P}^I$ and $\mathbb{P}^{II}$ and picking the one which maximizes the expected regret of decision maker's choice ($x = x_A$). Looking at Table 6, one remarks six expected values for $x = x_A$, representing these six settings. For instance, if $x' = x_C$ and $\mathbb{P} = \mathbb{P}^I$, the corresponding expected regret of $x = x_A$ can be derived as $\mathbb{E}_{\mathbb{P}}[h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})] = 0.8(4-1) + 0.2(3-6) = 1.8$\$. The maximum expected regret among thes six values is 2.4\$ which is associated with $x' = x_B$ under $\mathbb{P} = \mathbb{P}^I$.

Table 6: EA-RAR

$\boldsymbol{x}$	$\boldsymbol{x}'$	$\mathbb{P}$	$\mathbb{E}_{\mathbb{P}}[.]^\dagger$	$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[.]$	$\max_{\boldsymbol{x}'\in\mathcal{X}} \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[.]$
$\boldsymbol{x}^*$ = Project A	Project A	$\mathbb{P}^I$	0	0	**2.4**
		$\mathbb{P}^{II}$	0		
	Project B	$\mathbb{P}^I$	2.4	2.4	
		$\mathbb{P}^{II}$	-4		
	Project C	$\mathbb{P}^I$	1.8	1.8	
		$\mathbb{P}^{II}$	-3		
Project B	Project A	$\mathbb{P}^I$	-2.4	4	4
		$\mathbb{P}^{II}$	4		
	Project B	$\mathbb{P}^I$	0	0	
		$\mathbb{P}^{II}$	0		
	Project C	$\mathbb{P}^I$	-0.6	1	
		$\mathbb{P}^{II}$	1		
Project C	Project A	$\mathbb{P}^I$	-1.8	3	3
		$\mathbb{P}^{II}$	3		
	Project B	$\mathbb{P}^I$	0.6	0.6	
		$\mathbb{P}^{II}$	-1		
	Project C	$\mathbb{P}^I$	0	0	
		$\mathbb{P}^{II}$	0		
	$\min_{\boldsymbol{x}\in\mathcal{X}} \max_{\boldsymbol{x}'\in\mathcal{X}} \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[.]$				**2.4**

† $\mathbb{E}_{\mathbb{P}}[.] := \mathbb{E}_{\mathbb{P}}[h(\boldsymbol{x}', \boldsymbol{\zeta}) - h(\boldsymbol{x}, \boldsymbol{\zeta})]$

Performing the same analysis for $x = x_B$ and $x = x_C$ guides the manager towards investing on the project with minimum worst-case expected regret; more specifically, the minimum value in the last column of this table is 2.4\$, indicating that the best choice for the manager is to choose project A for investment with worst-case expected regret of 2.4\$. This is in contrast with the recommended option of project C coming from EP-WCR and EP-RAR problems.

B Column and constraint generation algorithm

The column and constraint generation algorithm, proposed by Zeng and Zhao (2013), is an iterative scheme which optimally solves two-stage linear robust optimization problems with right-hand-side uncertainty. Hence, we can employ it for solving problem (15). Assume that $\mathcal{X} \cap \mathcal{X}_{na}$ is non-empty, and let $\{\boldsymbol{x}'_1, \ldots, \boldsymbol{x}'_K\}$ denote the set of policies that comprise the vertices of $\mathcal{X} \cap \mathcal{X}_\Delta$, i.e., $\boldsymbol{x}'_k : \Omega \to \mathbb{R}^{n\times T}$ for all $k \in \{1, \ldots, K\}$. Let $\mathcal{K}' \subseteq \{1, \ldots, K\}$. The column and constraint generation algorithm can be viewed as a reduction to the vertex enumeration method, where at each iteration, a vertex is added to the following master problem

$$
\begin{aligned}
\mathcal{M}(\mathcal{K}') = \underset{\boldsymbol{x}, s, \{\boldsymbol{v}_k, \boldsymbol{\lambda}_k\}_{\forall k \in \mathcal{K}'}}{\text{minimize}} \quad & s \\
\text{s.t.} \quad & \boldsymbol{\lambda}_k^\top \boldsymbol{b} - v_{\omega,k} + \sum_{t=1}^{T} \boldsymbol{c}_t^\top (\boldsymbol{\zeta}(\omega))(\boldsymbol{x}'_{t,k}(\omega) - \boldsymbol{x}_t(\omega))) \leq s, \quad \forall \omega \in \Omega\,, \forall k \in \mathcal{K}' \\
& B_p^\top \boldsymbol{\lambda}_k = \boldsymbol{v}_k, \quad \forall k \in \mathcal{K}' \\
& B_q^\top \boldsymbol{\lambda}_k = 0, \quad \forall k \in \mathcal{K}' \\
& s \in \mathbb{R},\, \boldsymbol{\lambda}_k \in \mathbb{R}^m_+,\, \boldsymbol{v}_k \in \mathbb{R}^{|\Omega|}, \quad \forall k \in \mathcal{K}' \\
& \boldsymbol{x} \in \mathcal{X} \cap \mathcal{X}_{na}.
\end{aligned}
\tag{20}
$$

For any $\mathcal{K}' \subseteq \{1, \dots, K\}$, $\mathcal{M}(\mathcal{K}')$ constitutes a lower bound on the optimal value of problem (15). For a given $\boldsymbol{x} \in \mathcal{X} \cap \mathcal{X}_{na}$, we can evaluate $\mathcal{R}_\Delta(\boldsymbol{x})$ through solving the inner maximization of problem (15). Expressing the inner minimization of problem (15) thought its KKT conditions and merging it into the outer maximization problem, yields the following bilinear optimization program

$$
\begin{aligned}
\mathcal{R}_\Delta(\boldsymbol{x}) = \max_{\boldsymbol{x}',\, r,\, \boldsymbol{v},\, \boldsymbol{\lambda},\, \boldsymbol{p},\, \boldsymbol{q}} \quad & r + \boldsymbol{\lambda}^\top \boldsymbol{b} && \text{(21a)} \\
\text{s.t.} \quad & r + v_\omega \geq \sum_{t=1}^{T} \boldsymbol{c}_t^\top (\boldsymbol{\zeta}(\omega))(\boldsymbol{x}'_t(\omega) - \boldsymbol{x}_t(\omega)),\ \forall \omega \in \Omega && \text{(21b)} \\
& B_p^\top \boldsymbol{\lambda} = \boldsymbol{v} && \text{(21c)} \\
& B_q^\top \boldsymbol{\lambda} = 0 && \text{(21d)} \\
& \sum_{\omega \in \Omega} p_\omega = 1 && \text{(21e)} \\
& B_p \boldsymbol{p} + B_q \boldsymbol{q} \leq \boldsymbol{b} && \text{(21f)} \\
& p_\omega (r + v_\omega - \sum_{t=1}^{T} \boldsymbol{c}_t^\top (\boldsymbol{\zeta}(\omega))(\boldsymbol{x}'_t(\omega) - \boldsymbol{x}_t(\omega))) = 0,\ \forall \omega \in \Omega && \text{(21g)} \\
& \lambda_i \boldsymbol{e}_i^\top (\boldsymbol{b} - B_p \boldsymbol{p} - B_q \boldsymbol{q}) = 0,\ \forall i = 1, 2, ..., m && \text{(21h)} \\
& r \in \mathbb{R},\, \boldsymbol{v} \in \mathbb{R}^{|\Omega|},\, \boldsymbol{\lambda} \in \mathbb{R}^m_+,\, \boldsymbol{p} \in \mathbb{R}^{|\Omega|}_+,\, \boldsymbol{q} \in \mathbb{R}^{n_q} && \text{(21i)} \\
& \boldsymbol{x}' \in \mathcal{X} \cap \mathcal{X}_\Delta, && \text{(21j)}
\end{aligned}
$$

where $\boldsymbol{e}_i \in \mathbb{R}^m$ is the ith column of the identity matrix. Constraints (21b)-(21d) ensure primal feasibility, (21e)-(21f) ensure dual feasibility, while the bilinear constraints (21g) and (21h) ensure complementary slackness. To make the problem amenable to efficient optimization solvers, the bilinear constraints can be linearized using McCormick inequalities, see McCormick (1983). To this end, let M denotes a sufficiently large constant, typically referred to as the big-M constant in the integer programming literature. By introducing binary variables $\text{Bin}^{\boldsymbol{p}} \in \{0,1\}^{|\Omega|}$ and $\text{Bin}^{\boldsymbol{\lambda}} \in \{0,1\}^m$, constraints (21g) and (21h) can be reformulated as

$$
\begin{aligned}
& r + v_\omega - \sum_{t=1}^{T} \boldsymbol{c}_t^\top (\boldsymbol{\zeta}(\omega))(\boldsymbol{x}'_t(\omega) - \boldsymbol{x}_t(\omega)) \leq M \text{Bin}^{\boldsymbol{p}}_\omega,\ \forall \omega \in \Omega && \text{(22a)} \\
& \boldsymbol{p} \leq 1 - \text{Bin}^{\boldsymbol{p}} && \text{(22b)} \\
& \boldsymbol{b} - B_p \boldsymbol{p} - B_q \boldsymbol{q} \leq M \text{Bin}^{\boldsymbol{\lambda}} && \text{(22c)} \\
& \boldsymbol{\lambda} \leq M(1 - \text{Bin}^{\boldsymbol{\lambda}}). && \text{(22d)}
\end{aligned}
$$

Solving the resulting mixed integer linear program provides an upper bound on the optimal value of problem (15). The optimal worst-case benchmark policy $\boldsymbol{x}'$ of problem (21), can be added to the master problem to further strengthen the lower bound. Algorithm 1 describes the iterative process. The computational efficiency of Algorithm 1 heavily relies on the ability to evaluate efficiently $\mathcal{R}_\Delta(\boldsymbol{x})$ in Step 3. The choice of the big-M constant heavily influences the solution speed, i.e., choosing it too big will result to weak

Algorithm 1 Column and constraint generation algorithm, Zeng and Zhao (2013)

1: **Initialize**: $lb = -\infty$, $ub = \infty$, $\mathcal{K}' = \emptyset$.
2: Solve problem (20) and let $\boldsymbol{x}^*$ be the optimal solution. Set $lb = \mathcal{M}(\mathcal{K}')$;
3: Evaluate $\mathcal{R}_\Delta(\boldsymbol{x}^*)$ by solving problem (21) and let $\boldsymbol{x}'^*$ be the optimal solution. Set $ub = \mathcal{R}_\Delta(\boldsymbol{x}^*)$;
4: **if** $ub - lb > 0$ **then**
5: $\quad K' = K' \cup \{i\}$ where i is the index of $\boldsymbol{x}'^*$ in the set of vertices $\{\boldsymbol{x}'_1, \ldots, \boldsymbol{x}'_K\}$, and go to Step 2;
6: **else**
7: $\quad$ **Return:** $\boldsymbol{x}^*$ and $\mathcal{R}_\Delta(\boldsymbol{x}^*)$.
8: **end if**

linear relaxation leading in longer computational times. For the special case where the risk measure is the Conditional Value-at-Risk, the matrices in $\mathcal{D}$ reduce to $B_p = I \in \mathbb{R}^{|\Omega|\times|\Omega|}$, $B_q = 0$ and $\boldsymbol{b} = \bar{\boldsymbol{p}}/(1-\alpha)$. Since by construction $\boldsymbol{p} \in [0,1]^{|\Omega|}$, then constraint (22c) reduces to

$$\frac{\bar{\boldsymbol{p}}}{(1-\alpha)} - \boldsymbol{p} \leq \frac{1}{(1-\alpha)} \mathrm{diag}(\bar{\boldsymbol{p}}) \mathrm{Bin}^{\boldsymbol{\lambda}},$$

where $\mathrm{diag}(\bar{\boldsymbol{p}})$ is a diagonal matrix with $\bar{\boldsymbol{p}}$ appearing in the diagonal entries. In other words, the big-M constant can be set to $\bar{\boldsymbol{p}}_\omega/(1-\alpha)$ for the ω^{th} constraint.

C Multi-period inventory management problem

Table 7: Demand Realizations - $\boldsymbol{\zeta}_t(\omega)$

		ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}	ω_{11}	ω_{12}	ω_{13}	ω_{14}	ω_{15}	ω_{16}
	0	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
	0.5	0	0	0	0	0	0	0	0	4	4	4	4	4	4	4	4
t	1	6	6	6	6	3	3	3	3	11	11	11	11	12	12	12	12
	1.5	10	10	6	6	13	13	4	4	0	0	15	15	9	9	14	14
	2	4	1	7	14	2	4	9	12	15	2	12	8	13	4	11	3

Table 8: Instance Parameters

$\boldsymbol{c}_{it}$	$i=1$	$i=2$	$i=3$
$t=1$	14	12	20
$t=2$	18	16	22
$t=3$	24	18	28
$\bar{x}_{it} = 15, \forall i \in \mathcal{I}, \forall t \in \mathcal{T}$			
$\bar{x}_{i,tot} = 25, \forall i \in \mathcal{I}$			
$x^0_{wh} = 0$, $\underline{x}_{wh} = 0$, $\bar{x}_{wh} = 50$			

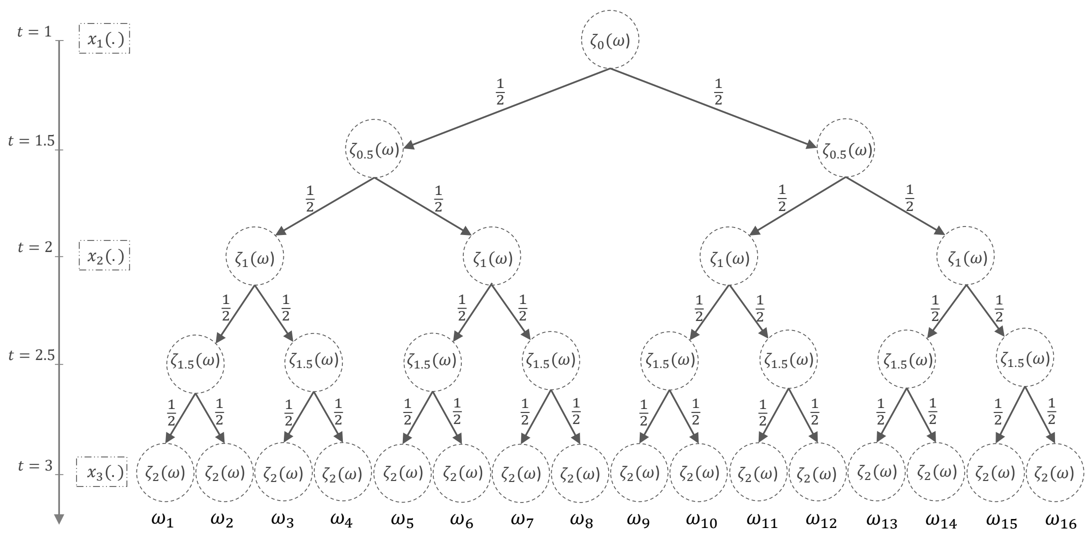

Figure 8: Scenario tree of a multi-period inventory management problem with $T = 3$, $|\Omega| = 16$ and $\mathcal{T}_d := \{1,\, 1.5,\, 2,\, 2.5,\, 3\}$.