Pareto Efficiency in Robust Optimization

Dan A. Iancu*   Nikolaos Trichakis†

August 1, 2012

Abstract

This paper formalizes and adapts the well known concept of Pareto efficiency in the context of the popular robust optimization (RO) methodology. We argue that the classical RO paradigm need not produce solutions that possess the associated property of Pareto optimality, and illustrate via examples how this could lead to inefficiencies and sub-optimal performance in practice. We provide a basic theoretical characterization of Pareto robustly optimal (PRO) solutions, and extend the RO framework by proposing practical methods that verify Pareto optimality, and generate solutions that are PRO. Critically important, our methodology involves solving optimization problems that are of the same complexity as the underlying robust problems, hence the potential improvements from our framework come at essentially no computational cost. We perform numerical experiments drawn from three different application areas (portfolio optimization, inventory management, and project management), which demonstrate that PRO solutions have a significant upside compared with solutions obtained via classical RO methods, at no extra cost or downside.

1 Introduction

Robust optimization (RO) is a relatively young methodology, developed mainly in the course of the last 15 years to analyze and optimize the performance of complex systems (refer to the survey papers Ben-Tal and Nemirovski (2007), Bertsimas et al. (2011a) and the book Ben-Tal et al. (2009) for a thorough overview of the framework). As a result of its versatility and tractability, recent years have seen an explosion of applications of RO in management science, ranging from inventory management to dynamic pricing and revenue management, portfolio optimization, and healthcare

---

*Graduate School of Business, Stanford University, daniancu@stanford.edu

†Harvard Business School, ntrichakis@hbs.edu

applications (for a comprehensive list, see Gabrel et al. (2012)[1]). The goal of this paper is not to motivate the use of robust (and, more generally, distribution free) techniques, but rather to introduce a methodological enhancement to this popular modeling approach.

RO deals with decision making problems where some of the parameters are *a priori* unknown and/or subject to uncertainty. The standard approach is to assume that all such parameters belong to particular uncertainty sets, and to take decisions so as to optimize the worst-case performance among all possible uncertainty realizations. As such, RO is indeed indifferent towards non worst-case scenarios, and the performance of decisions made under them.

In this paper, we discuss and demonstrate via numerous examples how the inherent focus of RO on optimizing performance only under worst-case outcomes might leave decisions "un-optimized" in case a non worst-case scenario materialized. Clearly, this is undesirable, particularly in circumstances where these decisions are implemented in practice, as they lead to sub-optimal performance. To the best of our knowledge, this paper is the first to reveal this extra dimension for optimization in the classical RO framework, and the first to propose a scalable, tractable method for exploiting it, in a way that *strictly* enhances the framework: indeed, our approach comes with *no downside* in either performance or computational complexity!

To formalize our findings and the fact that RO might lead to decisions that are "un-optimized" for non worst-case outcomes, we introduce the concept of Pareto efficiency in RO. The concept mimics the corresponding one in economics, engineering and multiobjective optimization: a decision that is Pareto robustly optimal is guaranteed to deliver optimized performance across all possible scenario realizations, in the same way, for instance, that a Pareto optimal solution in multiobjective optimization delivers optimized performance across all different objectives. We demonstrate via examples that decisions made using RO need not have this property. To alleviate this, we propose methods for verifying whether a decision is Pareto optimal or not, and methods for obtaining robustly optimal decisions that are provably Pareto optimal as well. Put differently, in this paper, we introduce an *essential* property that RO decisions that are to be implemented in practice need to possess (in a way that parallels the importance of Pareto efficiency in economics), and develop theory and computational tools pertaining to it. Our methodology enables a decision maker to

---

[1]This survey reports 100 papers authored in *Operations Research* and *Management Science* between 2007 and (June) 2012 containing the words "robust optimization" in their title, and a further 762 articles containing either "robust" or "robustness" in the title.

compute robustly optimal solutions that are compatible with the classical RO framework, incur no extra computational cost, and can perform strictly better in practice.

Specifically, we make the following contributions:

(a) We formalize and adapt the well accepted concept of Pareto efficiency to the classical RO framework. We demonstrate that the framework need not produce solutions that possess the associated property of Pareto optimality, and illustrate via examples how this could lead to inefficiencies and sub-optimal performance in practice.

(b) We provide a basic theoretical characterization of Pareto robustly optimal solutions.

(c) We extend the RO framework by proposing practical methods that verify Pareto optimality, and generate solutions that are also (provably) Pareto. Critically important, all our proposed methodology involves optimization problems that are of *the same computational complexity* as the underlying robust problems, hence the potential improvements from our framework come at essentially no computational cost.

(d) We perform numerical experiments drawn from three different application areas studied in the management science literature: portfolio optimization, inventory management, and project management. The studies demonstrate that Pareto robustly optimal solutions obtained via our methodology have a significant upside compared with solutions obtained via classical methods, at no extra cost or downside.

We conclude by noting that our treatment in this paper is restricted to the case of uncertain linear optimization problems. We make this choice primarily due to the overwhelming preponderance of linear models in practice, as well as for reasons of simplicity and ease of exposition. However, we see this as a first step in treating more general optimization problems appearing in the classical RO framework (Ben-Tal et al. 2009).

1.1   Literature review

Originally introduced in operations research and management science by Soyster (1973), the methodology of robust optimization (RO) has been revitalized in the late 1990s and early 2000s through the seminal work of Ben-Tal and Nemirovski, Bertsimas and Sim, and El-Ghaoui et al. (see the

review papers Ben-Tal and Nemirovski (2007), Bertsimas et al. (2011a) and the book Ben-Tal et al. (2009) for a thorough overview of the methodology and contributions).

With a strong emphasis on computational tractability and ability to accomodate a diverse range of relevant optimization models, the RO methodology has been adopted in many applications of interest in management science. Such examples include inventory management (Ben-Tal et al. 2005, Bertsimas and Thiele 2006, Bienstock and Özbay 2008, See and Sim 2010, Bertsimas et al. 2011b), dynamic pricing and revenue management (Adida and Perakis 2006, Perakis and Roels 2010, Lim and Shanthikumar 2007), assortment planning (Rusmevichientong and Topaloglu 2011), portfolio optimization (Goldfarb and Iyengar 2003, Bertsimas and Pachamanova 2008, Fabozzi et al. 2007), project management (Cohen et al. 2007, Goh et al. 2010, Adida and Joshi 2009), healthcare (Chan et al. 2006), auction design (Bandi and Bertsimas 2012), and others. The list is by no means exclusive - the interested reder can refer to the recent review papers Bertsimas et al. (2011a) and Gabrel et al. (2012), and the book Ben-Tal et al. (2009) for comprehensive references.

Despite its empirical success, however, the robust framework has been known to suffer from several potential shortcomings. One criticism is that, by focusing exclusively on the worst-case outcomes, it may result in conservative decisions, with limited potential upside. This has lead to several alternative proposals of robustness measures, such as absolute or relative regret (Savage 1972), "soft-robustness" (Ben-Tal et al. 2010), "light-robustness" (Fischetti and Monaci 2009), bw-robustness (Roy 2010, Gabrel et al. 2011), α-robustness (Kalai et al. 2012), and others. Depending on the exact assumptions and setup, such approaches typically result in the same (or slightly decreased) generality and the same (or slightly increased) computational complexity as the standard RO framework. Critically, however, all such approaches trade off some of the robustness (i.e., performance in the worst-case) in exchange for potential upside. By contrast, our approach, which parallels the notion of Pareto efficiency, guarantees the same worst-case outcome, while at the same time allowing potentially improved upside, at no increase in computational complexity.

A second (and more subtle) criticism is that, by solely focusing on worst-case outcomes, the minimax/maximin criterion may result in multiple optimal solutions, and hence generate Pareto inefficiencies in the decision process. This idea has emerged in several research streams, from fairness in resource allocation (Young 1995, Bertsimas et al. 2012), to multiobjective optimization (Ogryczak 1997, Suh and Lee 2001), but has been, to the best of our knowledge, absent from

the mainstream RO literature.[2] The unifying characteristic in the above settings is that they are concerned with a *finite* set of alternative realizations / scenarios (e.g., the multiple parties in a problem of equitable resource allocation, or the objectives in a multiobjective problem). This case, however, is typically of limited interest in RO, where the standard setup involves a continuous uncertainty set, i.e., an infinite number of realizations (Ben-Tal et al. 2009). In this sense, the proposals for correcting Pareto inefficiencies of the maximin rule in resource allocation and multiobjective optimization (e.g., the lexicographic max-min fairness scheme (Young 1995), Ogryczak (1997)) cannot be readily extended to the classical RO framework. By contrast, our definition of Pareto robustly optimal solutions is applicable to cases with both finite and infinite uncertainty sets, and hence fully compatible with the RO methodology. Furthermore, our approach still allows potential multiplicity in the Pareto robustly optimal solutions, hence not eliminating the benefits of multiplicity in RO (Iancu et al. 2012, Bertsimas et al. 2011c).

Finally, we note that the concepts of "Pareto optimality/efficiency" and "Robust Optimization" have appeared together before, typically in the area of multiobjective optimization, and in a very different spirit than that addressed in the present paper. In particular, Gorissen and den Hertog (2011) discuss the use of RO (and, more broadly, convex optimization) to approximate the Pareto frontier in multiobjective optimization problems. Several papers have also attempted introducing robust formulations of multiobjective problems (e.g., Deb and Gupta (2005), Chen et al. (2011), Luo and Zheng (2008), Cristobal et al. (2006), Ono and Nakayama (2009), Suh and Lee (2001)). The notions of robustness used are typically different than those in the classical RO framework, resulting in models that require solving very difficult optimization problems; this usually leads to the use of various heuristics, such as evolutionary algorithms, polynomial chaos or particle swarm optimization.

1.2 Notation

We use **1** to denote the vector with all components equal to one. We use e_i to denote the unit vector with 1 in the ith component. The inequality sign for vectors is used for componentwise inequality.

---

[2]The multiplicity of optimal solutions has been typically seen as a positive feature in RO, since it allowed deriving simple dynamic policies in multi-period problems (Bertsimas et al. 2010, Iancu et al. 2012), or optimizing secondary objectives (Bertsimas et al. 2011c).

We use several basic notions of convex analysis (Rockafellar 1970) that we denote as follows. For a set $S \subset \mathbb{R}^n$: we use $\text{ext}(S)$ to denote the set of its extreme points; we use $\text{conv}(S)$ to denote its convex hull; we use $\text{ri}(S)$ to denote its relative interior; we use S^* to denote its dual cone.

2   Pareto Robustly Optimal Solutions

In this section, we introduce and formally define the notion of Pareto efficiency in robust optimization (RO). For illustration purposes and to ease exposition, we consider a specific form of RO problems, for which we present our definitions and the results in Section 3. In Section 4, we discuss how our findings extend to more general forms of RO problems.

The type of RO problems we consider are linear optimization problems where only the objective is subject to uncertainty. Specifically, we consider the problem of selecting x from a polyhedral feasible set $X \subset \mathbb{R}^n$ so as to maximize $p^T x$. The objective vector p is *a priori* unknown, and belongs to a polyhedral uncertainty set $U \subset \mathbb{R}^n$. We assume that both sets X and U are bounded and that their inequality representations are

$$
X \quad = \quad \{x \in \mathbb{R}^n : Ax \leq b\} \text{, and} \tag{1a}
$$

$$
U \quad = \quad \{p \in \mathbb{R}^n : Dp \geq d\} \text{,} \tag{1b}
$$

where $A \in \mathbb{R}^{m_X \times n}$, $b \in \mathbb{R}^{m_X}$, $D \in \mathbb{R}^{m_U \times n}$ and $d \in \mathbb{R}^{m_U}$ are given. We note that our findings readily extend to the case where we only have access to a representation of U through its extreme points.

Although linear models with polyhedral uncertainty sets are a strict subset of the RO methodology, we choose them as the focus of our treatment due to the widespread use in practice, and since they are a natural first step before examining more general convex optimization models.

According to the classical RO paradigm (Ben-Tal et al. 2009), one selects x in the above setting

6

by solving the following optimization problem[3]

$$\underset{x \in X}{\text{maximize}} \quad \underset{p \in U}{\min} \, p^T x. \tag{2}$$

Let z^{RO} be the optimal value of this problem. It is easy to check that the optimal set, denoted by X^{RO}, is given by

$$X^{\text{RO}} = \left\{ x \in X : \exists \, y \in \mathbb{R}_+^{m_U} \text{ such that } D^T y = x, \quad y^T d \geq z^{\text{RO}} \right\}. \tag{3}$$

An optimal solution $x \in X^{\text{RO}}$ of problem (2) is typically referred to as a *robustly optimal solution* in this setting. It corresponds to a solution that maximizes the worst-case objective value $p^T x$, under all possible realizations of the uncertainty $p \in U$. In other words, a robustly optimal solution is selected under the sole requirement of protecting us against worst-case scenarios. We are guaranteed that no other solution does better along that requirement. If such a solution is to be used in practice, however, this raises the following questions:

- How would x perform (in terms of the objective value) in case the uncertainty scenario that actually materialized did not correspond to a worst-case one?

- Are there any guarantees that no other solution exists that, apart from protecting us from worst-case scenarios, also performs better overall, under all possible uncertainty realizations?

To better understand the meaning of these questions, we consider a stylized network flow example arising in a communication network. The example illustrates that, in general, a robustly optimal solution $x \in X^{\text{RO}}$ might perform poorly in case a worst-case scenario did not realize: in fact, we show that it is sometimes possible to find another robustly optimal solution that outperforms x under all possible uncertainty realizations.

Example 1. *Consider a communication network that consists of multiple links. Any given link is used to transmit information between two points in the network, at a rate that is determined by the*

---

[3]Note that this model can also correspond to a distributionally robust setting, where U is a subset of the probability simplex, denoting ambiguity about the true probability measure, and the goal is to maximize the worst-case expected outcome (where nature is choosing the distribution adversarially). In this sense, our model can speak to the vast literature on risk measures and choice under ambiguity (see, e.g., Ben-Tal et al. (2010) or Bertsimas et al. (2011a) for more references).
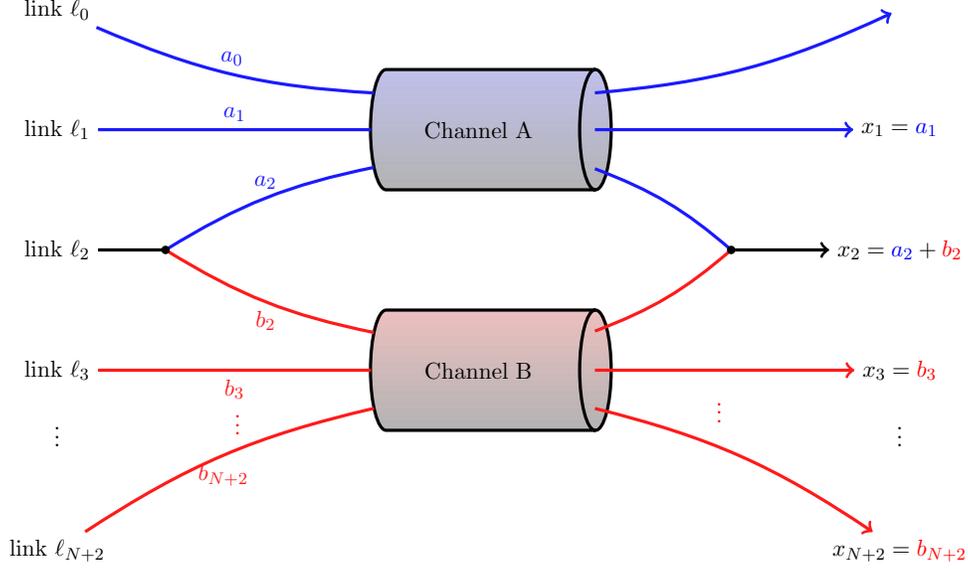
Figure 1: An illustration of the network structure in Example 1.

network manager. The links share different capacitated communication channels for transmission; capacity limitations then affect the information transmission rates of the individual links.

Specifically, consider the following network structure: there are two channels, denoted by A and B. The channels are of unit capacity and are utilized by $N + 3$ links, denoted by $\ell_0, \ell_1, \ldots, \ell_{N+2}$, with $N \geq 3$. Links ℓ_0 and ℓ_1 utilize only channel A. Link ℓ_2 utilizes both channels A and B. Links $\ell_3, \ldots, \ell_{N+2}$ utilize only channel B. Let a_i denote the transmission rate of link ℓ_i over channel A, $i = 0, 1, 2$. Accordingly, b_i denotes the transmission rate of link ℓ_i over channel B, $i = 2, 3, \ldots, N+2$. Vectors $a \in \mathbb{R}^3$ and $b \in \mathbb{R}^{N+1}$ contain the associated values. The structure is depicted in Figure 1.

Links $\ell_1, \ldots, \ell_{N+2}$ are dedicated for emergency purposes, whereas link ℓ_0 is used for general purposes. Let x_i be the transmission rate of the emergency link ℓ_i, $i = 1, 2, \ldots, N + 2$. We have

$$x_1 = a_1, \quad x_2 = a_2 + b_2, \quad and \quad x_i = b_i, \quad i = 3, 4, \ldots, N + 2.$$

In case an emergency transmission needs to be established, this is achieved by utilizing one of the dedicated links, or a combination thereof. More specifically, let f_i be the fraction of the emergency

transmission routed via link ℓ_i, $i = 1, 2, \ldots, N + 2$. The net emergency transmission rate is then

$$f^T x = \sum_{i=1}^{N+2} f_i x_i.$$

Fractions f depend on the emergency situation and are uncertain. In particular, f is assumed to belong to a probability simplex uncertainty set U, i.e.,

$$U = \left\{ f \in \mathbb{R}_+^{N+2} : \mathbf{1}^T f = 1 \right\}.$$

The problem for the network manager is then to select rates x, a and b (in case of an emergency) so as to maximize the net emergency transmission rate. If the manager uses robust optimization, the rates are selected by solving

$$
\begin{aligned}
\text{maximize} \quad & \min_{f \in U} f^T x \\
\text{subject to} \quad & x_1 = a_1 \\
& x_2 = a_2 + b_2 \\
& x_i = b_i, \quad i = 3, \ldots, N + 2 \\
& a_0 + a_1 + a_2 = 1 \\
& b_2 + b_3 + \ldots + b_{N+2} = 1 \\
& a, b \geq 0,
\end{aligned}
\tag{4}
$$

with variables x, a and b. Let X be the feasible set. It is easy to check that the optimal value is $z^{\mathrm{RO}} = 1/N$ and the optimal set is

$$X^{\mathrm{RO}} = \left\{ (x, a, b) \in X : x \geq \frac{1}{N} \mathbf{1} \right\}.$$

We solved problem (4) using the SeDuMi software package (Sturm et al. 2006), for $N = 10$, and obtained the following robustly optimal solution

$$a^{\mathrm{SeDuMi}} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}^T, \quad b^{\mathrm{SeDuMi}} = \begin{bmatrix} 0 & \frac{1}{10} & \cdots & \frac{1}{10} \end{bmatrix}^T, \quad x^{\mathrm{SeDuMi}} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{10} & \cdots & \frac{1}{10} \end{bmatrix}^T.$$

Note that, for any realization of f drawn from U, we have

$$f^T x^{\text{SeDuMi}} \geq 1/N = 1/10.$$

An associated worst-case realization of f for which the obtained performance is equal to $1/10$ is, for instance, e_3. Consider now a non-worst-case scenario for x^{SeDuMi}, e.g., $f = e_1$. Then, the obtained objective value is $e_1^T x^{\text{SeDuMi}} = 1/3$.

We now compare the robustly optimal solution we obtained from SeDuMi with the following solution

$$a^\star = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}^T, \quad b^\star = \begin{bmatrix} 0 & \frac{1}{10} & \cdots & \frac{1}{10} \end{bmatrix}^T, \quad x^\star = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{10} & \cdots & \frac{1}{10} \end{bmatrix}^T.$$

It is easy to check that this solution is also robustly optimal, i.e., $(x^\star, a^\star, b^\star) \in X^{\text{RO}}$. Hence, it has the same qualities as the solution obtained from SeDuMi in protecting us from worst-case realizations of f. However, if $f = e_1$ realizes, this solution yields strictly better performance,

$$e_1^T x^\star = 1/2 > e_1^T x^{\text{SeDuMi}} = 1/3.$$

In fact, it is easy to see that $x^\star$ performs better, or at least equally well, compared to x^{SeDuMi} for *any* realization of f drawn from U! In particular, we have

$$f^T x^\star \geq f^T x^{\text{SeDuMi}}, \quad \forall f \in U, \quad \text{and} \tag{5a}$$

$$f^T x^\star > f^T x^{\text{SeDuMi}}, \quad \forall f \in U \cap \{f : f_1 + f_2 > 0\}. \tag{5b}$$

The above discussion demonstrates that by focusing only on worst-case scenarios, the classical RO paradigm allows solutions that are *sub-optimal* for other uncertainty realizations. Clearly, this inefficiency is not important in case robust optimization is used only for purposes of assessing worst-case performance. However, it might negatively affect managers who utilize RO to actually compute solutions they implement in practice.

To rectify the aforementioned weakness of the classical RO framework, we need to additionally require from a robustly optimal solution to "perform as well as possible" across all uncertainty

scenarios that can realize. In particular, for any robustly optimal solution we select, we need to guarantee that there does not exist another solution that performs at least as well across all uncertainty realizations and strictly better at some realizations. If such a solution existed, it would be strictly preferred for all practical considerations.

We call solutions with the above property as *Pareto Robustly Optimal (PRO)* solutions. Formally, we have the following definition:

Definition 1. *A solution x is called a Pareto Robustly Optimal (PRO) solution for Problem (2) if it is robustly optimal, i.e., $x \in X^{\mathrm{RO}}$, and $\nexists \bar{x} \in X$ such that*

$$p^T \bar{x} \geq p^T x, \quad \forall p \in U, \quad and$$

$$\bar{p}^T \bar{x} > \bar{p}^T x, \; for \; some \; \bar{p} \in U.$$

In the definition above, we say that $\bar{x}$ *Pareto dominates* x. As discussed in the Introduction, the terminology we use borrows from economics and multiobjective optimization: RO can be viewed as a multiobjective optimization problem with an infinite number of objectives (corresponding to uncertainty scenarios).

Returning to Example 1, by (5), we have that the solution $(x^\star, a^\star, b^\star)$ Pareto dominates the SeDuMi solution. Moreover, one can show that the solution $(x^\star, a^\star, b^\star)$ is a PRO solution. In fact, if we denote the set of all PRO solutions with X^{PRO}, then, for Problem (2) we have

$$X^{\mathrm{PRO}} = \left\{ (x, a, b) \in X : x \geq \frac{1}{N} \mathbf{1}, \quad x_1 + x_2 = 1 \right\}.$$

Below is another toy example that illustrates the notion of PRO solutions.

Example 2. *Consider the following problem, which is of the same form as Problem (2):*

$$
\begin{aligned}
\text{maximize} \quad & \min_{p \in U} \{p_1 x_1 + p_2 x_2 + p_3 x_3\} \\
\text{subject to} \quad & x_1 - x_2 = 0 \\
& x_1 + x_3 = 0 \\
& x_1 \geq 0 \\
& x_1 \leq 1,
\end{aligned}
\tag{6}
$$

where $U = \left\{ p \in \mathbb{R}^3 : \mathbf{1} \leq p \leq \mathbf{2} \right\}$ is a hypercube uncertainty set.

For the above problem, it is easy to check that for any $x \in X$, $\begin{bmatrix} 1 & 1 & 2 \end{bmatrix}^T$ is a worst-case uncertainty realization, for which $\begin{bmatrix} 1 & 1 & 2 \end{bmatrix} x = 0$. Hence, we have that $X = X^{\mathrm{RO}}$.

We solved problem (6) using IBM CPLEX and obtained the robustly optimal solution $x^{\mathrm{cplex}} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$. Consider the solution $x^\star = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}^T$. For any uncertainty realization different from the worst-case one we identified above, i.e., $p \neq \begin{bmatrix} 1 & 1 & 2 \end{bmatrix}^T$, we have that $p^T x^\star > p^T x^{\mathrm{cplex}}$. For the worst-case realization, both solutions yield an objective value of zero. Hence, solution $x^\star$ Pareto dominates solution x^{cplex}. In fact, $x^\star$ is the only PRO solution for problem (6), so that $X^{\mathrm{PRO}} = \{x^\star\}$.

There are many interesting questions to be addressed that are theoretically and practically relevant in view of the notion of PRO solutions we have introduced. Do PRO solutions always exist and how can we find them efficiently? When is every robustly optimal solution also a PRO solution? Can we characterize X^{PRO}? Is it a convex set?

Apart from shedding light on the questions raised above, the rest of the paper is devoted to answering the following two central questions:

1. Given a robustly optimal solution $x \in X^{\mathrm{RO}}$, how do we check if x is also PRO? If it is not, how do we find a PRO solution $\bar{x} \in X^{\mathrm{PRO}}$ that Pareto dominates x?

2. How do we optimize over the set of PRO solutions X^{PRO}? From our discussion so far, it should be obvious that for practical decision making, a manager should always prefer PRO solutions. In case a manager also has a secondary objective, how can she select a PRO solution that is optimal with respect to this secondary objective?

3   Finding and Optimizing over PRO Solutions

As suggested at the end of Section 2, several questions of interest can be posed concerning the set of PRO solutions. The goal of the present section is to provide detailed answers to these questions. We focus our discussion here on the class of RO problems described by (2), when the feasible set X and the uncertainty set U are given by (1a) and (1b), respectively. In Section 4, we revisit and extend our results to several other models of interest.

3.1 Finding PRO Solutions

In current practice, a decision maker would first formulate a RO model for the particular application of interest, and then seek to solve the resulting problem, hence determining a robustly optimal solution $x \in X^{\mathrm{RO}}$. However, as suggested by the simple examples in Section 2, the solution x does not necessarily have to be a PRO solution.

In this context, the first question of interest is how to check whether a given $x \in X^{\mathrm{RO}}$ is also PRO, and, if not, how to find an $\bar{x}$ that is PRO and Pareto dominates x. The following theorem argues that both of these questions can be answered in a straightforward fashion, by solving a single linear program (LP) of compact size.

Theorem 1. *Given a point $x \in X^{\mathrm{RO}}$, consider an arbitrary point $\bar{p} \in \mathrm{ri}(U)$, and the following optimization problem:*

$$
\begin{aligned}
\underset{y}{\text{maximize}} \quad & \bar{p}^T y \\
\text{subject to} \quad & y \in U^* \\
& x + y \in X.
\end{aligned}
\tag{7}
$$

Then, either the optimal value in the problem is zero, in which case $x \in X^{\mathrm{PRO}}$, or the optimal value is strictly positive, in which case $\bar{x} = x + y^\star$ Pareto dominates x and $\bar{x} \in X^{\mathrm{PRO}}$, for any optimal solution $y^\star$.

Proof of Theorem 1. Note first that $y = 0$ is always feasible in Problem (7). Hence, the optimal value is nonnegative. The discussion separates in two disjoint cases.

1. *The optimal value is zero.* Assume that $x \notin X^{\mathrm{PRO}}$, i.e., there exists $\tilde{x} \in X$ that Pareto dominates x: (i) $p^T \tilde{x} \geq p^T x, \, \forall \, p \in U$, and (ii) $\exists \, \hat{p} \in U$ such that $\hat{p}^T \tilde{x} > \hat{p}^T x$. Without loss of generality, we can additionally take $\hat{p} \in \mathrm{ext}(U)$, e.g., as a vertex solution to $\max_{p \in U}(\tilde{x} - x)^T p$.

Note that $\tilde{x} - x$ is readily feasible in (7). We claim that $\bar{p}^T(\tilde{x} - x) > 0$. To prove this, recall that any $\bar{p} \in \mathrm{ri}(U)$ has a representation as a strict convex combination of all extreme points of U (Rockafellar 1970), i.e., $\exists \, \lambda \in \mathbb{R}^{|\mathrm{ext}\,U|}$ such that $\lambda > 0$, $\mathbf{1}^T \lambda = 1$ and $\bar{p} = \sum_{i \in \mathcal{I}} \lambda_i p_i$, where $\mathrm{ext}(U) \overset{\mathrm{def}}{=} \{p_i \, : \, i \in \mathcal{I}\}$. But then,

$$
\bar{p}^T(\tilde{x} - x) = \underbrace{\hat{p}^T(\tilde{x} - x)}_{>0} + \sum_{i \in \mathcal{I} \, : \, p_i \neq \hat{p}} \underbrace{p_i^T(\tilde{x} - x)}_{\geq 0} > 0.
$$

13

This immediately leads to the desired contradiction.

2. *The optimal value is positive.* The conditions in Definition 1 can be directly verified for $\bar{x}$, showing that $\bar{x}$ Pareto dominates x. To see that $\bar{x} \in X^{\mathrm{PRO}}$, consider testing this by solving problem (7), i.e., with x replaced by $\bar{x} = x + y^\star$, and the same $\bar{p}$. We claim that the optimal solution to this problem must be zero (which, in turn, implies that $\bar{x} \in X^{\mathrm{PRO}}$). Otherwise, if an optimal solution $\tilde{y}$ existed, with $\bar{p}^T \tilde{y} > 0$, then $y^\star + \tilde{y}$ would be feasible and provide a higher objective value than $y^\star$ when solving problem (7) to test whether x was PRO, contradicting the fact that $y^\star$ was an optimal solution. □

Let us comment on the result and its relevance. We claimed that problem (7) is an LP. This follows since the dual cone U^* of a polyhedral set U always has a polyhedral representation (Rockafellar 1970). For instance, if U is the polytope given by (1b), then, by strong LP duality,

$$U^* \overset{\text{def}}{=} \{y \in \mathbb{R}^n : y^T p \geq 0, \, \forall \, p \in U\} = \{y \in \mathbb{R}^n : \exists \lambda \in \mathbb{R}_+^{m_U} \text{ such that } D^T \lambda = y, \, d^T \lambda \geq 0\}. \quad (8)$$

This implies that problem (7) can be solved efficiently using standard software (e.g., CPLEX, Gurobi, SeDuMi, etc.), for sizes that are typical in real-world applications.

Note also that our result is stated for an arbitrary point $\bar{p}$ in the relative interior of U, so that problem (7) must only be solved *once*. Since finding a point in the relative interior of a polyhedron can be done efficiently by LP techniques (Schrijver 2000), this readily leads to an efficient procedure for testing whether $x \in X^{\mathrm{PRO}}$ and (if not) for producing points that Pareto dominate x.

In a different sense, Theorem 1 also confirms that PRO solutions to any robust LP problem always exist, and suggests the following procedure for finding them.

Corollary 1. *Consider an arbitrary point $\bar{p} \in \mathrm{ri}(U)$. Then, all the optimal solutions to the problem* $\mathrm{maximize}_{x \in X^{\mathrm{RO}}} \, \bar{p}^T x$ *are PRO.*

The proof follows analogously to that of Theorem 1, and is omitted. It is interesting to note that finding an $x \in X^{\mathrm{PRO}}$ is not substantially more difficult than finding an $x \in X^{\mathrm{RO}}$ – one only needs to solve an additional LP. Since the potential gains from working with PRO solutions can be substantial (as illustrated in our examples in Section 5), this suggests that the framework can have considerable promise in practice.

The procedure introduced in Corollary 1 also suggests a simple way to generate (potentially different) solutions in X^{PRO}, by (i) sampling different values $\bar{p}$ from $\mathrm{ri}(U)$, and (ii) solving the corresponding LP over X^{RO}. The following result confirms that *all* the points of X^{PRO} can, in fact, be generated by such a procedure.

Proposition 1. *For any $x \in X^{\mathrm{PRO}}$, there exists $\bar{p} \in \mathrm{ri}(U)$ such that $x \in \arg\max_{y \in X^{\mathrm{RO}}} \bar{p}^T y$.*

Proof of Proposition 1. Let $P \in \mathbb{R}^{n \times |\mathrm{ext}(U)|}$ denote the matrix with columns given by the extreme points of U, and, without loss of generality, assume that the last $|\mathrm{ext}(U)| - k$ columns (where $k \leq |\mathrm{ext}(U)|$) correspond to *all* the points $p \in \mathrm{ext}(U)$ such that $p^T x = z^{\mathrm{RO}}$, $\forall x \in X^{\mathrm{RO}}$. Also let $\bar{P} \in \mathbb{R}^{n \times k}$ denote the matrix obtained by keeping the first k columns of P.

Consider any point $x \in X^{\mathrm{PRO}}$. To construct the desired $\bar{p} \in \mathrm{ri}(U)$, we note that it is enough to show that $\exists \lambda \in \mathbb{R}^k$ such that $\lambda \geq \mathbf{1}$ and $x \in \arg\max_{y \in X^{\mathrm{RO}}} y^T \bar{P} \lambda$. The reason is that any such λ can be extended into $\tilde{\lambda} \stackrel{\mathrm{def}}{=} \frac{1}{\mathbf{1}^T \lambda + |\mathrm{ext}(U)| - k} \begin{bmatrix} \lambda \\ \mathbf{1} \end{bmatrix}$, which satisfies $\tilde{\lambda} > 0$, $\mathbf{1}^T \tilde{\lambda} = 1$, and $x \in \arg\max_{y \in X^{\mathrm{RO}}} y^T P \tilde{\lambda}$ if and only if $x \in \arg\max_{y \in X^{\mathrm{RO}}} y^T P \begin{bmatrix} \lambda \\ 0 \end{bmatrix}$.

To this end, assume (for the purposes of deriving a contradiction) that $\forall \lambda \geq \mathbf{1}$, we have $x \notin \arg\max_{y \in X^{\mathrm{RO}}} y^T \bar{P} \lambda$. This implies that

$$\forall \lambda \geq \mathbf{1}, \ \exists y(\lambda) \in X^{\mathrm{RO}} \text{ such that } \big(y(\lambda) - x\big)^T \bar{P} \lambda > 0 \quad \Rightarrow$$

$$\min_{\lambda \geq \mathbf{1}} \max_{y \in X^{\mathrm{RO}}} (y - x)^T \bar{P} \lambda > 0 \quad \overset{(*)}{\Leftrightarrow}$$

$$\max_{y \in X^{\mathrm{RO}}} \min_{\lambda \geq \mathbf{1}} (y - x)^T \bar{P} \lambda > 0 \quad \overset{(**)}{\Rightarrow}$$

$$\exists y \in X^{\mathrm{RO}} \ : \ \max\{ \mathbf{1}^T \mu \ : \ \mu^T \leq (y - x)^T \bar{P}, \ \mu \geq 0 \} > 0.$$

Step $(*)$ follows from the minimax theorem in convex analysis, which is applicable here since the function $(y - x)^T \bar{P} \lambda$ is bilinear in y and λ, and the set X^{RO} is compact (Rockafellar 1970). Step $(**)$ follows from strong LP duality, which holds for the given y, since the minimization over λ is feasible and bounded below. The last step implies that $x \notin X^{\mathrm{PRO}}$, since y Pareto dominates x. $\square$

The previous results also allow formulating conditions under which all the points in X^{RO} are actually PRO solutions. This is relevant since it would allow a decision maker to not worry about the issue of Pareto domination, and simply utilize any solution obtained by solving a standard

robust problem. The next result, whose proof also follows directly from Theorem 1, argues that this question can also be answered by solving a compact LP.

Corollary 2. *Consider any $\bar{p} \in \mathrm{ri}(U)$, and the following optimization problem:*

$$
\begin{aligned}
\underset{x,y}{\text{maximize}} \quad & \bar{p}^T y \\
\text{subject to} \quad & x \in X^{\mathrm{RO}} \\
& y \in U^* \\
& x + y \in X.
\end{aligned}
\tag{9}
$$

Then, $X^{\mathrm{PRO}} = X^{\mathrm{RO}}$ if and only the optimal value is zero.

Proof of Corollary 2. Let $x^\star$, $y^\star$ be optimal solutions of Problem (9). It can be checked that if the optimal value is positive, $x^\star + y^\star$ is PRO and Pareto dominates $x^\star \in X^{\mathrm{RO}}$, implying that $X^{\mathrm{PRO}} \neq X^{\mathrm{RO}}$. The reverse direction follows since an optimal value of zero implies that any $x \in X^{\mathrm{RO}}$ is PRO, by Theorem 1. $\qquad \square$

In practice, it may be relevant to look for simpler conditions guaranteeing $X^{\mathrm{PRO}} = X^{\mathrm{RO}}$. By Corollary 2, one such example is when $0 \in \mathrm{ri}(U)$ – in this case, any $x \in X^{\mathrm{RO}}$ yields the same objective value of zero, so that $X^{\mathrm{PRO}} = X^{\mathrm{RO}}$. It is important to note that this condition applies to the particular objective used in the robust problem (2), i.e., $p^T x$, where p is uncertain. Clearly, by suitably translating any uncertainty set U with nonempty relative interior, one may obtain a new model with objective $r^T x + p^T x$, where r is known exactly, $p \in \tilde{U}$ is uncertain, and $0 \in \mathrm{ri}(\tilde{U})$. For such a model, $0 \in \mathrm{ri}(\tilde{U})$ no longer guarantees that $X^{\mathrm{PRO}} = X^{\mathrm{RO}}$. While having zero in the relative interior may occur in some cases, many of the RO models considered in applications do not typically satisfy the condition (see our numerical studies in Section 5 and Ben-Tal et al. (2009) for details). In fact, the condition should never be satisfied if the physical nature of an uncertain parameter prevented it from switching signs, e.g., if it corresponded to an uncertain yield in a production process, a probability or a customer demand.

3.2 Optimizing Over the Set of PRO Solutions

The previous section discussed procedures for testing whether a given robustly optimal solution $x \in X^{\mathrm{RO}}$ is a PRO, and (more generally) how PRO solutions can be generated. In many settings,

such results could suffice – if a decision maker was ready to settle for solution x, then using a PRO solution that Pareto dominates x has essentially no downside. On the contrary, there is considerable upside, since PRO solutions can perform strictly better when non worst-case scenarios occur.

In practice, however, a manager may have a secondary objective in mind. For instance, choosing decisions out of X^{PRO} may protect against risk, but the manager may also want to maximize the performance in a representative scenario $r \in U$ (which could, for instance, correspond to the sample mean, median or mode, obtained from historical data). See, for instance, Bertsimas et al. (2011c), who use a secondary objective based on fairness.

In this context, a natural question to ask is whether linear objectives can be efficiently optimized over the set X^{PRO}. More broadly, understanding the structure and properties of the set becomes relevant. The following example confirms that X^{PRO} is, unfortunately, not a convex set in general.

Example 3 (Non-convexity of X^{PRO}). *Consider the following feasible set X and uncertainty set U:*

$$X = \{x \in \mathbb{R}^4 : x_1 \leq 1,\ x_2 + x_3 \leq 6,\ x_3 + x_4 \leq 5,\ x_2 + x_4 \leq 5\},$$
$$U = \mathrm{conv}\Big(\{e_i,\ i \in \{1,\ldots,4\}\}\Big).$$

It can be checked that $z^{\mathrm{RO}} = 1$, and $X^{\mathrm{RO}} = \{x \in X : x \geq \mathbf{1}\}$. Also, $x^1 = \begin{bmatrix} 1 & 2 & 4 & 1 \end{bmatrix}^T$ and $x^2 = \begin{bmatrix} 1 & 4 & 2 & 1 \end{bmatrix}^T$ are both PRO solutions (they are the optimal solutions to the problems of maximizing $\begin{bmatrix} \epsilon & \epsilon & 1 - 3\epsilon & \epsilon \end{bmatrix}^T$ and $\begin{bmatrix} \epsilon & 1 - 3\epsilon & \epsilon & \epsilon \end{bmatrix}^T$ over X^{RO}, respectively, for some small $\epsilon > 0$). However, the point $0.5\,x^1 + 0.5\,x^2 = \begin{bmatrix} 1 & 3 & 3 & 1 \end{bmatrix}^T \notin X^{\mathrm{PRO}}$, since it is Pareto dominated by $\begin{bmatrix} 1 & 3 & 3 & 2 \end{bmatrix}^T \in X^{\mathrm{RO}}$.

The non-convex structure of X^{PRO} suggests that solving optimization problems over the set may be computationally challenging in general. One particular case when this is simple is when $X^{\mathrm{RO}} = X^{\mathrm{PRO}}$, which can be tested using the LP in Corollary 2. When $X^{\mathrm{RO}} \neq X^{\mathrm{PRO}}$, the following result provides some intuition about the structure of the latter set.

Proposition 2. *If $X^{\mathrm{RO}} \neq X^{\mathrm{PRO}}$, then $X^{\mathrm{PRO}} \cap \mathrm{ri}(X^{\mathrm{RO}}) = \emptyset$.*

Proof of Proposition 2. We first argue that, if $X^{\mathrm{RO}} \neq X^{\mathrm{PRO}}$, then $\nexists \bar{p} \in \mathrm{ri}(U)$ such that $\bar{p}^T x = $ constant, $\forall x \in X^{\mathrm{RO}}$. To see this, note that if such a $\bar{p}$ existed, then the optimal objective function in problem (7), for any point $x \in X^{\mathrm{RO}}$, would have to be zero, implying that $X^{\mathrm{RO}} = X^{\mathrm{PRO}}$.

Assume now that $X^{\mathrm{PRO}} \cap \mathrm{ri}(X^{\mathrm{RO}}) \neq \emptyset$, and consider any x in the intersection. By Proposition 1, there must exist $\bar{p} \in \mathrm{ri}(U)$ such that $x \in \arg\max_{x \in X^{\mathrm{RO}}} \bar{p}^T x$. Furthermore, since $\bar{p}^T x$ is not constant over X^{RO}, there must exist a $y \in X^{\mathrm{RO}}$ such that $\bar{p}^T y < \bar{p}^T x$. But then, since $x \in \mathrm{ri}(X^{\mathrm{RO}})$, there exists a small enough $\epsilon > 0$ such that $\bar{x} \overset{\text{def}}{=} x + \epsilon \cdot (x - y) \in X^{\mathrm{RO}}$, for which $\bar{p}^T \bar{x} > \bar{p}^T x$, a contradiction. $\qquad\square$

The previous result shows that the set X^{PRO} is either identical to X^{RO} or is contained in the boundary of the latter. This result is somewhat encouraging, since it may allow a characterization of the convex hull of X^{PRO} in particular cases. We do not pursue this further in the present paper.

In a different sense, the results in Corollary 2 also suggest that whether a given $x \in X^{\mathrm{RO}}$ (obtained by solving a nominal robust problem) is actually PRO may critically depend on the solver used. In particular, whenever interior point methods are used[4] to obtain x, the resulting solution may very well not be PRO.

We now return to the main question of interest, i.e., developing a tractable procedure for optimizing a linear objective over the set X^{PRO}, and present two different approaches for addressing it. The first involves solving a Mixed-Integer Linear Program (MILP), and is summarized in the following proposition.

Proposition 3. *For any $r \in \mathbb{R}^n$ and any $\bar{p} \in \mathrm{ri}(U)$, let $(x^\star, \mu^\star, \eta^\star, z^\star) \in \mathbb{R}^n \times \mathbb{R}^{m_X} \times \mathbb{R} \times \{0,1\}^{m_X}$ be an optimal solution of the following MILP*

$$\underset{x,\mu,\eta,z}{\text{maximize}} \quad r^T x \tag{10a}$$

$$\text{subject to} \quad x \in X^{\mathrm{RO}} \tag{10b}$$

$$\mu \leq M(1-z) \tag{10c}$$

$$b - Ax \leq Mz \tag{10d}$$

$$DA^T \mu - d\eta \geq D\bar{p} \tag{10e}$$

$$\mu \geq 0,\ \eta \geq 0,\ z \in \{0,1\}^{m_X}, \tag{10f}$$

---

[4]This is the case if the solver SeDuMi is used, or if CPLEX is run with the interior point solver selected (as opposed to the Simplex method). However, even when the Simplex method is used, there is still no reason to *a priori* suspect that the optimal solution is a PRO (see Example 2 of Section 2).

where M is a sufficiently large value. Then, $x^\star$ is an optimal solution of the problem

$$\underset{x \in X^{\mathrm{PRO}}}{\text{maximize}} \quad r^T x.$$

Proof of Proposition 3. We need to show that $x \in X^{\mathrm{RO}}$ is PRO if and only if there exist $\mu \in \mathbb{R}^{m_X}$, $\eta \in \mathbb{R}$, and $z \in \{0,1\}^{m_X}$ that satisfy (10c)-(10f).

Fix a solution $x \in X^{\mathrm{RO}}$. By Theorem 1, x is PRO if and only if, for an arbitrarily chosen $\bar{p} \in \mathrm{ri}(U)$, the optimal value in the problem $\text{maximize}_y \{\bar{p}^T y \; : \; y \in U^*, x + y \in X\}$ is zero. Equivalently, x is PRO if and only if the optimal value in the following primal-dual LP pair is zero:

$$
\begin{array}{lll}
\underset{\lambda}{\text{maximize}} \quad \bar{p}^T D^T \lambda & = & \underset{\mu, \eta}{\text{minimize}} \quad \mu^T (b - Ax) \\[2mm]
\text{subject to} \quad d^T \lambda \geq 0 & & \text{subject to} \quad D A^T \mu - d\,\eta \geq D\bar{p} \\[2mm]
\qquad\qquad A\,D^T \lambda \leq b - Ax & & \qquad\qquad \mu \geq 0 \\[2mm]
\qquad\qquad \lambda \geq 0 & & \qquad\qquad \eta \geq 0.
\end{array}
$$

To arrive at the primal (maximization), we used the description in (1a) for the set X, the expression in (8) for the dual cone U^*, and we eliminated the auxiliary variables y. Note that the primal is trivially feasible, with the choice $\lambda = 0$, resulting in an objective value of zero. As such, whenever its optimal value is (less than or) equal to zero, strong LP duality must hold.

We then have that the optimal value is zero if and only if there exist $\mu \in \mathbb{R}^{m_X}$ and $\eta \in \mathbb{R}$ satisfying the constraints of the dual, and such that $\mu^T (b - Ax) = 0$. This is a bilinear constraint in variables x and μ. In fact, since $\mu \geq 0$ and $b - Ax \geq 0, \forall x \in X^{\mathrm{RO}}$, it is equivalent to the linear complementarity constraints $\mu_i \cdot (b - Ax)_i = 0, i = 1, \ldots, m_X$ (Luo et al. 1996). The latter constraint can be modeled using variables $z \in \{0,1\}^{m_X}$ and constraints (10c)-(10d), where M is sufficiently large.[5] $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The above approach should be very relevant in practice, since large-scale MILPs can be solved in a matter of seconds using commercially available software, such as CPLEX or Gurobi. In case this is still an onerous task, one can also resort to the following simple heuristic for optimizing

---

[5]Note that we need $M \geq \max_{i=1,\ldots,m_X} \max\big(\mu_i, \max_{x \in X} e_i^T (b - Ax)\big)$. The latter term can always be bounded, since X is compact. The former term may also be bounded, depending on the dual feasible set. In practice, one may simply choose an increasing sequence of M, stopping when the constraints for $\mu \leq M$ are no longer binding.

linear functions over X^{PRO}.

Algorithm 1 Sampling procedure for solving the problem $\max_{x \in X^{\mathrm{PRO}}} r^T x$

Require: Inequality description for X and U (i.e., A, b, D and d). Number of steps N.
1: $\hat{X}^{\mathrm{PRO}} \leftarrow \emptyset$
2: **for** $i \leftarrow 1, N$ **do**
3:      Sample a point $\bar{p} \in \mathrm{ri}(U)$.
4:      $\hat{X}^{\mathrm{PRO}} \leftarrow \hat{X}^{\mathrm{PRO}} \cup \arg\max_{x \in X^{\mathrm{RO}}} \bar{p}^T x$.
5: **end for**
6: Solve $\mathrm{maximize}_{x \in \hat{X}^{\mathrm{PRO}}} r^T x$.

The central idea behind the algorithm is to use the results in Corollary 1 and Proposition 1, which suggest that all the PRO solutions can be generated by sampling points $\bar{p}$ from the relative interior of U, and solving the LP in Step 4. By keeping track of a set $\hat{X}^{\mathrm{PRO}}$ which is a subset of the true X^{PRO}, this algorithm will always produce a lower bound to the true optimal value. However, coupled with a suitable upper bound (e.g., obtained by maximizing $r^T x$ over X^{RO}) the heuristic may prove satisfactory in cases where the (exact) MILP in Proposition 3 is difficult to solve.

4 Generalizations

In this section, we discuss several directions for extending our earlier framework and results.

4.1 Uncertainty in the Constraints

Consider a linear optimization problem where the coefficients of the constraints matrix are uncertain. Specifically, consider the following RO formulation

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{subject to} \quad & Ax \geq b, \quad \forall A \in U_A,
\end{aligned}
\tag{11}
$$

where $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ are given, and $U_A \subset \mathbb{R}^{m \times n}$ is a bounded polyhedral uncertainty set of the form typically considered in the RO literature[6] (Ben-Tal et al. 2009). Problems with uncertainty in the vectors c and b can be readily reformulated so as to have the same form as Problem (11).

---

[6]The typical modeling assumption is that the matrix A depends affinely on a collection of primitive uncertain quantities, i.e., $A(\xi) = \sum_{i=1}^{m} A_i \xi_i$, where $A_i \in \mathbb{R}^{m \times n}$ are known, and $\xi \in \mathbb{R}^m$ denotes the primitive uncertain parameters, assumed to lie in an uncertainty set $\Xi \subset \mathbb{R}^m$. As such, whenever Ξ has a compact polyhedral representation, the corresponding set $U_A = \{A(\xi) : \xi \in \Xi\}$ does, as well (see Ben-Tal et al. (2009) for details).

For instance, in case c is uncertain, one can consider an equivalent epigraph formulation, where the objective is replaced by an additional variable t and the extra constraint $t - c^T x \geq 0$ is added.

Let X^{RO} be the optimal set and z^{RO} the optimal value of Problem (11). The vector of slacks for constraints in (11) for a solution x and uncertainty realization A, denoted by $s(x, A)$, is equal to

$$s(x, A) \overset{\text{def}}{=} Ax - b, \quad \forall x \in \mathbb{R}^n, \ A \in U_A.$$

In view of the above, one can express X^{RO} as follows

$$X^{\text{RO}} = \left\{ x \in \mathbb{R}^n : c^T x \leq z^{\text{RO}}, \quad s(x, A) \geq 0, \forall A \in U_A \right\}. \tag{12}$$

A compact polyhedral description of X^{RO} is readily available whenever such a description is also available for U_A, which we have assumed to be the case here.

According to classical RO, a *robustly optimal solution* $x \in X^{\text{RO}}$ in this setting protects us from worst-case realizations of $A \in U_A$, by ensuring that the slacks are nonnegative. Moreover, it does so at a minimum cost $c^T x = z^{\text{RO}}$. That is, by selecting a robustly optimal solution, we are guaranteed that no other solution exists that yields nonnegative slacks under all uncertainty scenarios, and at a lower cost than z^{RO}. However, in a similar spirit to our discussion in Section 2, such a selection criterion does not "optimize slacks" under all possible uncertain realizations – it fails to guarantee that no other solution exists yielding "larger" slacks, and at the same cost of z^{RO}!

The above observation gives rise to the notion of PRO solutions for formulation (11). In order to evaluate the slacks provided by different solutions and uncertainty realizations, we introduce a *slack value vector* $v \in \mathbb{R}^m$ that quantifies the relative value of slack in each constraint.[7] In particular, the slacks provided by solution x, under scenario A, are valued at

$$v^T s(x, A) = v^T (Ax - b), \quad \forall x \in \mathbb{R}^n, \ A \in U_A.$$

We have the following definition:

Definition 2. *A solution x is called a Pareto Robustly Optimal (PRO) solution for Problem (11)*

---

[7]We choose to compare two slack vectors by scalarization through the value vector $v \in \mathbb{R}^n$, a method that is common in multiobjective optimization (Boyd and Vandenberghe 2004). However, other multidimensional orderings may be more suitable in particular settings (for instance, the Lorenz order). Extending our framework to such cases is an interesting direction, which we do not pursue in the present paper.

if it is robustly optimal, i.e., $x \in X^{\mathrm{RO}}$, and $\nexists \bar{x} \in X^{\mathrm{RO}}$ such that

$$v^T s(\bar{x}, A) \geq v^T s(x, A), \quad \forall A \in U_A, \quad and$$

$$v^T s(\bar{x}, \bar{A}) > v^T s(x, \bar{A}), \; for \; some \; \bar{A} \in U_A.$$

Similarly to our discussion in Section 2, PRO solutions in this setting guarantee that slacks are optimized for all uncertainty realizations $A \in U_A$. Compared with robustly optimal solutions, PRO solutions have the same qualities, i.e., they ensure feasibility at the lowest cost possible, but, in addition, they also provide potentially higher slacks (as valued through vector v). This is particularly important for the following reasons:

(a) In the majority of problems of the form (11) that are derived from practical applications, one of the constraints typically involves the *actual* realized cost, e.g., due to an epigraph reformulation.[8] As such, slack in that particular constraint immediately translates into lower actual realized cost. We refer the reader to our Numerical Studies in Sections 5.2-5.3 for examples.

(b) Solutions that provide zero or low values of slack are potentially vulnerable to unforseen model mis-specifications that frequently occur in practice. For the value of slack, we refer the reader for instance to Joshi and Boyd (2009).

Hence, similarly to the notion studied in Section 2, PRO solutions in this setting have no downside and considerable potential upside.

Methodologically, the conditions of Pareto domination in Definition 2 are equivalent to

$$(A^T v)^T \bar{x} \geq (A^T v)^T x, \quad \forall A \in U_A, \quad and$$

$$(\bar{A}^T v)^T \bar{x} > (\bar{A}^T v)^T x, \; for \; some \; \bar{A} \in U_A.$$

If we let $p \stackrel{\text{def}}{=} A^T v$ and

$$U \stackrel{\text{def}}{=} \left\{ A^T v : A \in U_A \right\}, \tag{13}$$

---

[8]Note that an epigraph reformulation is necessary not only in cases where the actual cost vector is uncertain (as discussed above), but also when the cost of the problem is piece-wise affine, see for instance Bertsimas et al. (2010).

we arrive at the same formulation of Pareto efficiency as in Definition 1. Using this observation, one can extend all our findings in Section 3.1. For instance, the extended Corollary 1 would state:

Corollary 3. *Consider an arbitrary point $\bar{p} \in \mathrm{ri}(U)$, where U is given by (13). Then, all the optimal solutions to $\mathrm{maximize}_{x \in X^{\mathrm{RO}}} \; \bar{p}^T x$ are PRO for Problem (11), where X^{RO} is given by (12).*

As such, all the remarks in Section 3.1 pertaining to our results and their practical relevance hold in the present setting, as well.

4.2 Mixed-Integer Linear Optimization Models

A second direction for extending our results is the case when some of the decisions x are constrained to be integers. This is of great relevance in practical applications, since many models of real-world processes involve integrality constraints (see, e.g., Bertsimas and Sim (2003)).

To this end, the present section is concerned with the following problem

$$\underset{x \in X}{\mathrm{maximize}} \quad \underset{p \in U}{\min} \; p^T x$$

$$X = \left\{ x \in \mathbb{R}^n \; : \; Ax \leq b, \, x_i \in \mathbb{Z}, \, \forall \, i \in \mathcal{I} \right\}$$

$$U = \left\{ p \in \mathbb{R}^n \; : \; Dp \geq d \right\},$$

where $\mathcal{I} \subseteq \{1, \ldots, n\}$ is a set of indices, and A, b, D and d are known.

Our goal is to revisit the main results of Section 3, and investigate their validity in the present setting. As a general overview, all the critical results can be suitably adapted to the new model: by solving particular MILPs, one can (a) test whether a given x is a PRO solution (and, if not, obtain a PRO solution $\bar{x}$ that Pareto dominates it), (b) generate PRO solutions, and (c) optimize linear objectives over the set of PRO solutions. The sole results of Section 3 that no longer hold here are Proposition 1 and Proposition 2 – in particular, not all PRO solutions can be recovered by maximizing linear objectives of the form $\bar{p}^T x$ where $\bar{p} \in \mathrm{ri}(U)$, and the set X^{PRO} is not necessarily in the boundary of X^{RO} when the two are different. For reasons of brevity, we refrain from re-proving the (new) claims, since the arguments exactly parallel their counterparts in Section 3.

First, note that the representation for the set of robust optimal solutions X^{RO} remains identical

to expression (3), i.e.,

$$X^{\mathrm{RO}} = \left\{ x \in X : \exists\, y \in \mathbb{R}_+^{m_U} \text{ such that } D^T y = x, \quad y^T d \geq z^{\mathrm{RO}} \right\}.$$

Here, however, the optimal (worst-case) value for the problem, z^{RO}, is obtained by solving an MILP (Bertsimas and Sim 2003), and the set X^{RO} also includes integrality constraints, due to X.

In this context, we can reaffirm our first main result in Theorem 1, which holds true without any modifications. In particular, to test whether a point $x \in X^{\mathrm{RO}}$ is also a PRO, one only needs to solve the optimization problem (7) for an arbitrarily chosen $\bar{p}$ in $\mathrm{ri}(U)$. As before, if the optimal value is exactly zero, then $x \in X^{\mathrm{PRO}}$. Otherwise (i.e., strictly positive optimal value), $x + y^\star \in X^{\mathrm{PRO}}$, and it Pareto dominates x. We note that the sole change from the result in Section 3.1 is that problem (7) is now an MILP, instead of an LP. However, the complexity of finding a PRO solution compared to just finding a solution $x \in X^{\mathrm{RO}}$ is still the same, as both now require solving MILPs.

As before, this theorem readily leads to a simple procedure for finding points $x \in X^{\mathrm{PRO}}$. In particular, a result analogous to Corollary 1 holds: for any $\bar{p} \in \mathrm{ri}(U)$, all the optimal solutions to the MILP $\mathrm{maximize}_{x \in X^{\mathrm{RO}}}\ \bar{p}^T x$ are guaranteed to be PRO points. Furthermore, just as stated in Corollary 2, determining whether $X^{\mathrm{RO}} = X^{\mathrm{PRO}}$ resumes to solving the MILP in (9), and comparing the optimal value with zero.

This result also brings us to the first main point of departure from Section 3. Recall that, by choosing different points $\bar{p} \in \mathrm{ri}(U)$, the union of the sets $\arg\max_{x \in X^{\mathrm{RO}}}\ \bar{p}^T x$ was guaranteed to contain *all* PRO solutions for the case without integrality constraints (see Proposition 1 in Section 3.1). Unfortunately, that is no longer the case here. In particular, there may be points $x \in X^{\mathrm{PRO}}$ that lie in the strict interior of the set $\mathrm{conv}(X^{\mathrm{RO}})$, and hence cannot be recovered by maximizing a linear functional over X^{RO}. The following example presents such an instance.

Example 4 (Strictly Interior Point in X^{RO}). *Let $X = \{ x \in \mathbb{Z}^2 : x \geq 0, \frac{1}{2}x_1 + \frac{1}{5}x_2 \leq 1 \}$, and $U = \mathrm{conv}(\{e_1, e_2\})$. It can be checked that*

$$X^{\mathrm{RO}} = \left\{ x \in \mathbb{Z}^2 : x \geq 2 \cdot \mathbf{1} \right\}$$
$$X^{\mathrm{PRO}} = \left\{ [0 \quad 2]^T,\ [5 \quad 0]^T,\ [1 \quad 2]^T \right\}.$$

In particular, $\bar{x} = [1\ 2]^T \in \mathrm{ri}\big(\mathrm{conv}(X^{\mathrm{RO}})\big)$, and therefore $\nexists\, p \in \mathrm{ri}(U)$ such that $\bar{x} \in \arg\max_{x \in X^{\mathrm{RO}}} p^T x$.

The example above also proves that there are cases in which $X^{\mathrm{RO}} \neq X^{\mathrm{PRO}}$, and some points in X^{PRO} lie in the strict (relative) interior of X^{RO}, so that Proposition 2 also does not hold.

Due to the same reason, the heuristic suggested in Algorithm 1 for optimizing a linear functional over X^{PRO} may not be very effective: one can certainly still apply it, but since sampling points $\bar{p} \in \mathrm{ri}(U)$ is no longer guaranteed to generate all the points in X^{PRO}, one may be unable to recover the exact (true) optimal value.

However, we note that the problem of optimizing linear objectives over X^{PRO} can nonetheless be dealt with in a scalable fashion. In particular, the main result in Proposition 3 still holds, and, for any $r \in \mathbb{R}^n$, the optimal value in the problem $\mathrm{maximize}_{x \in X^{\mathrm{PRO}}}\, r^T x$ can be obtained by solving the MILP in (10). As a result, optimizing over X^{PRO} when some decisions are integral is as easy as finding a single $x \in X^{\mathrm{RO}}$, unlike the setting discussed in Section 3.2.

5 Numerical Studies

In this section, we evaluate the implications of our findings via numerical studies. We focus on three application areas in which RO has proven to be very powerful. In particular, we study three classical problems from the literature drawn from finance, inventory management and project management applications.

We generate multiple instances of the problems we consider, using random data, with the purpose of assessing

(a) the frequency at which computed robustly optimal solutions are Pareto dominated by other solutions, and

(b) the performance gain by considering PRO solutions in practice instead of dominated robustly optimal solutions.

For every generated instance, we solve an equivalent problem to Problem (9), in order to identify dominated robustly optimal solutions and associated PRO solutions.

We find that in approximately 15% of the instances we generate (across all three problems) there exist robustly optimal solutions that are Pareto dominated by other PRO solutions. Moreover, using

the associated PRO solutions in these cases yields a relative performance gain that can be as high as 74%, compared to using the dominated robustly optimal solutions.

5.1 Portfolio Optimization

RO has been widely studied and used in financial services applications, and particularly in portfolio optimization problems. We refer the reader to Ben-Tal et al. (2000), Goldfarb and Iyengar (2003), Bertsimas and Pachamanova (2008), Natarajan et al. (2008), Calafiore (2007), Ben-Tal et al. (2010), Gabrel et al. (2012) and the book Fabozzi et al. (2007) for a thorough overview.

In this section, we consider a simple portfolio selection problem that has the form we studied in Section 2, i.e., a linear program with uncertainty in the objective.

Problem description. A manager wishes to invest her wealth in $n+1$ investment opportunities or assets. We denote the return of the ith asset with r_i, $i = 1, \ldots, n+1$. The $(n+1)$th asset yields a known, deterministic return of μ_{n+1}, i.e.,

$$r_{n+1} = \mu_{n+1},$$

with no associated risk. On the contrary, all remaining assets are risky. In particular, the return of the ith asset is equal to

$$r_i = \mu_i + \sigma_i \zeta_i, \quad i = 1, \ldots, n,$$

where μ_i is the expected return of the ith asset, ζ_i is a random, uncertain shock affecting the return of the ith asset and σ_i is a volatility parameter. The values of the shocks $\zeta \in \mathbb{R}^n$ are assumed to be bounded between -1 and 1, and are assumed to sum up to zero. Uncertainty sets of that style have been introduced and used in Bandi and Bertsimas (2012).

Let $x \in \mathbb{R}^{n+1}$ be the target portfolio composition vector, i.e., x_i is the fraction of the wealth that the manager wishes to invest in the ith asset. We require that (a) no shorting is allowed, i.e., $x \geq 0$, and that (b) the net fraction of the wealth invested in any of the asset groups $\{1, \ldots, N/4\}$, $\{N/4, \ldots, N/2\}$, $\{N/2, \ldots, 3N/4\}$ and $\{3N/4, \ldots, N\}$ does not exceed 25%, for diversification purposes. The objective of the manager is to select a portfolio composition so as to maximize its worst-case return.

A formulation of the above RO problem is as follows:

$$\text{maximize} \quad \min_{\zeta \in U} \sum_{i=1}^{n+1} r_i\, x_i$$

$$\text{subject to} \quad r_i = \mu_i + \sigma_i\,\zeta_i, \quad i = 1, \ldots, n$$

$$r_{n+1} = \mu_{n+1}$$

$$\sum_{i=1+\frac{kN}{4}}^{\frac{(k+1)\,N}{4}} x_i \le 0.25, \quad k = 0, 1, 2, 3$$

$$\mathbf{1}^T x = 1$$

$$x \ge 0,$$

with variables $r \in \mathbb{R}^{n+1}$, $x \in \mathbb{R}^{n+1}$ and $U = \{\zeta \in \mathbb{R}^n : -\mathbf{1} \le \zeta \le \mathbf{1},\ \mathbf{1}^T \zeta = 0\}$.

Data. We consider 10,000 instances of problems of size $n = 8$. The expected returns of the risky assets $\{\mu_i\}_{i=1}^n$ are set to 3%, while the volatility parameters $\{\sigma_i\}_{i=1}^n$ are independently sampled from a uniformly distributed random variable between 1% and 10%. The risk-free return μ_{n+1} is also independently sampled from the same random variable.

Results. The average number of instances in which a robustly optimal solution was identified that was not a PRO solution was 31%. For these instances, we computed the maximum performance gap (in terms of the objective) one could get if the PRO solution were used instead of the dominated robustly optimal solution we obtained. A histogram of the associated (relative) performance gaps is depicted in Figure 2(a). The median performance gap recorded was 12%, while the maximum was as high as 74%.

5.2 Inventory Management

A different stream of applications in which RO has proven particularly useful has been inventory and supply chain management. For a thorough review and many references, we direct the interested reader to Ben-Tal et al. (2005), Bertsimas and Thiele (2006), Adida and Perakis (2006), Bienstock and Özbay (2008), Bertsimas et al. (2010), Bertsimas et al. (2011b), See and Sim (2010), the review papers Bertsimas et al. (2011a), Gabrel et al. (2012) and the book Ben-Tal et al. (2009).

In this section, we consider a single warehouse multiple retailer setting, where the manager
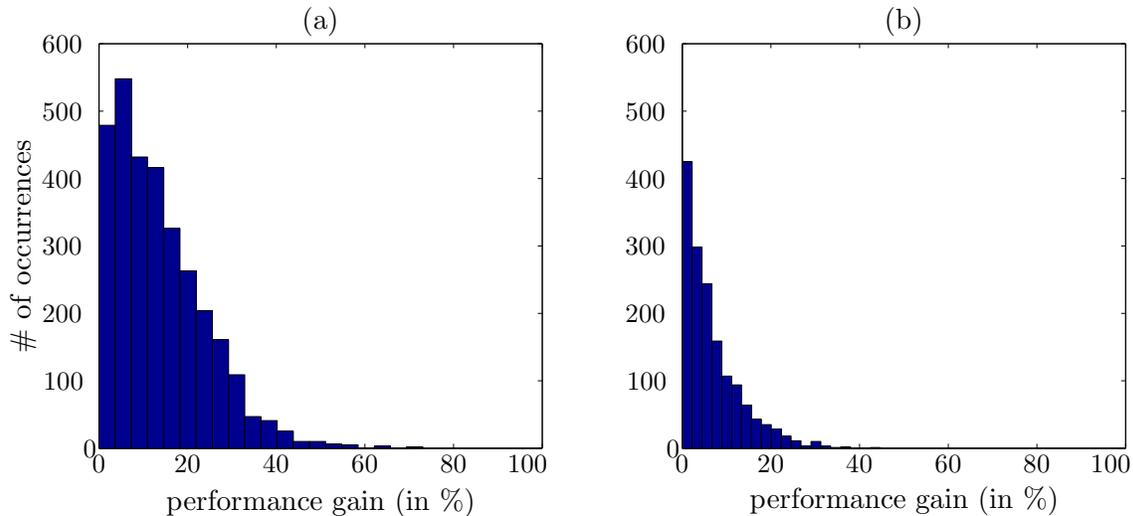
Figure 2: Histograms of maximal performance gains by considering PRO solutions in the numerical studies of (a) Section 5.1 and (b) Section 5.2.

needs to make stocking decisions in the face of uncertain demand. From a RO perspective, the resulting problem is a two-stage adjustable problem with fixed recourse.

Problem description. Consider a retail network consisting of a single warehouse and N different retail points, indexed by $i = 1, \ldots, N$, where customer demand is realized. For simplicity, we consider the case where a single item is offered for sale.

We have C units of the item available in the warehouse that need to be distributed to the retail points. The ith retail point holds zero initial inventory and is capable of stocking at most c_i units. The transportation costs for distributing, or equivalently stocking inventory at the ith retail point is t_i currency units per unit of inventory. Similarly, we assume that there is an operating cost at the ith retail point equal to h_i currency units per unit of inventory.[9] The revenues from sales at the ith retail point are r_i currency units per unit of inventory sold.

Customer demand at each point is uncertain and is driven by n_f factors that affect the market. In particular, we assume that the demand at the ith point, denoted by d_i, is equal to

$$d_i = d_i^0 + q_i^T z, \quad i = 1, \ldots, N,$$

where d_i^0 is the nominal expected demand, $z \in \mathbb{R}^{n_f}$ is a vector of the realized values of the n_f

---

[9]For simplicity, we assume no fixed transportation or operating costs, although one could include them in a straightforward manner.

28

factors, and $q_i \in \mathbb{R}^{n_f}$ is a vector of parameters that measure the exposure of demand d_i to each of the market factors. The values of the market factors are uncertain. However, we assume that they are bounded, as is their sum, as follows:

$$-b \cdot \mathbf{1} \le z \le b \cdot \mathbf{1}, \quad -B \le \mathbf{1}^T z \le B,$$

where b and B are known parameters.

Let $s \in \mathbb{R}^N$ be a vector containing the stocking decisions for all points. Similarly, $y \in \mathbb{R}^N$ is a vector with the sales at each point. Customer demand is realized after stocking decisions have been made. Thus, the sales at the ith point are equal to

$$y_i = \min \{s_i, d_i\}, \quad i = 1, \ldots, N.$$

The manager makes the stocking decisions so as to maximize worst-case profits, denoted by P. The corresponding RO formulation is as follows:

$$\text{maximize} \quad P \tag{14a}$$

$$\text{subject to} \quad P \le r^T y - (t + h)^T s \tag{14b}$$

$$y \le s \tag{14c}$$

$$y \le d^0 + Qz, \quad \forall z \in U \tag{14d}$$

$$\mathbf{1}^T s = C \tag{14e}$$

$$s \le c \tag{14f}$$

$$s \ge 0, \tag{14g}$$

with variables $P \in \mathbb{R}$, $y \in \mathbb{R}^N$ and $s \in \mathbb{R}^N$, $U = \{z \in \mathbb{R}^N : -b \cdot \mathbf{1} \le z \le b \cdot \mathbf{1}, -B \le \mathbf{1}^T z \le B\}$ and $Q = \begin{bmatrix} q_1 & \ldots & q_N \end{bmatrix}^T$. Constraint (14b) corresponds to an upper bound on the worst-case profits, equal to revenues from sales minus transportation and operations costs. According to constraints (14c)-(14d), sales are less than or equal to available stock and realized demand. Constraints (14e)-(14f) correspond to net and individual point capacity constraints, respectively.

Problem (14) is an adjustable RO problem, with fixed recourse (sales y and worst-case profits

P). To solve it, we approximate it with its Affine Adjustable Robust Counterpart (AARC), a popular and powerful heuristic (see Ben-Tal et al. (2004), Ben-Tal et al. (2005), Bertsimas et al. (2010), Iancu et al. (2012) and Ben-Tal et al. (2009)).

The AARC of Problem (14) is obtained by substituting all variables with affine functions of the uncertainty variables z and new auxiliary variables. The resulting formulation is of the same form as Problem (11). For our notion of PRO solutions, we require from robustly optimal solutions to maximize slack of constraint (14b). In other words, we employ our methodology from Section 4.1 using a slack value vector equal to a unit vector, where the component corresponding to constraint (14b) is equal to one.

Note that slack of constraint (14b) depends on the actual uncertainty realization. If a worst-case demand scenario realized, the profits on the right-hand side are equal to the worst-case profits P, and the slack is zero. For non worst-case realizations however, slack in the constraint is possible, which then translates to additional profits on top of the worst-case value of P (depending on the actual demand realization). PRO solutions ensure "optimal performance" under any demand realization by maximizing this slack, unlike robustly optimal solutions that ignore it.

Data. We consider 10,000 instances of problems of size $N = 10$. The available inventory is set to $C = 2,000$ units. All other problem data is independently sampled from uniformly distributed random variables. Inventory capacities at individual points $c \in \mathbb{R}^N$ are drawn between 300 and 500. Transportation $t \in \mathbb{R}^N$ and operations cost rates $h \in \mathbb{R}^N$ are drawn between 1 and 3. Sales revenues rates $r \in \mathbb{R}^N$ are drawn between 20 and 40.

Nominal demand values $d^0 \in \mathbb{R}^N$ are drawn between 100 and 200. We fix the number of market factors to $n_f = 2$. Exposure parameters $Q \in \mathbb{R}^{N \times n_f}$ are drawn between -2 and 2. Parameters b and B that bound the factor values are set to 5 and 25, respectively.

Results. The average number of instances in which a robustly optimal solution was identified that was not a PRO solution was 16%. For these instances, we computed the maximum performance gap (in terms of actual profits) one could get if the PRO solution were used instead of the dominated robustly optimal solution we obtained. A histogram of the associated (relative) performance gaps is depicted in Figure 2(b). The median performance gap recorded was 4.8%, while the maximum was as high as 45%.

5.3 Project Management

The third application that we consider is focused on robust models for project management. These have been studied in several papers, including Cohen et al. (2007), Goh et al. (2010), and Adida and Joshi (2009). Here, we consider a model discussed in Ben-Tal et al. (2009), where a manager needs to make resource allocation decisions in the face of uncertain processing times. Methodologically, the resulting problem is a two-stage adjustable RO problem with fixed recourse.

Problem description. Consider a project that involves multiple events and activities or tasks to be completed. Projects of that kind are typically analyzed and represented using *PERT diagrams* (Ben-Tal et al. 2009). A PERT diagram is a directed, acyclic graph consisting of a set of nodes $\mathcal{N}$, which correspond to the project events, and a set of edges $\mathcal{E}$, which correspond to the project activities or tasks.

Among the nodes there is a start node S and an end node F that correspond to the start and end of the project. The graph represents logical precedences between the project tasks and events as follows. A task, represented by an edge, starts being processed only after the event that corresponds to the node the task originates from, has occurred. The task is completed after some uncertain processing time. An event occurs only when all the tasks that correspond to all its incoming edges have been completed.

We consider the project of creating a factory studied by Ben-Tal et al. (2009). The project entails the acquirement and delivery of equipment (event A), build of facility #1 (event B) and of facility #2 (event C). The underlying tasks are as follows: (a) acquiring and delivering of equipment, (b) building facility #1 and (c) facility #2, (d) installing equipment in facility #1 and (e) in facility #2, (f) training personell at facility #1 and (g) at facility #2. The associated PERT diagram is given in Figure 3(a).

To introduce some notation, let task $e \in \mathcal{E}$ originate from node $s(e)$ and terminate at node $f(e)$. Its processing time, denoted by τ_e, is equal to

$$\tau_e = \tau_e^0 + \delta_e, \quad \forall e \in \mathcal{E},$$

where τ_e^0 is the standard processing time of task e and δ_e is an unforseen delay. We assume that

(a)   (b)
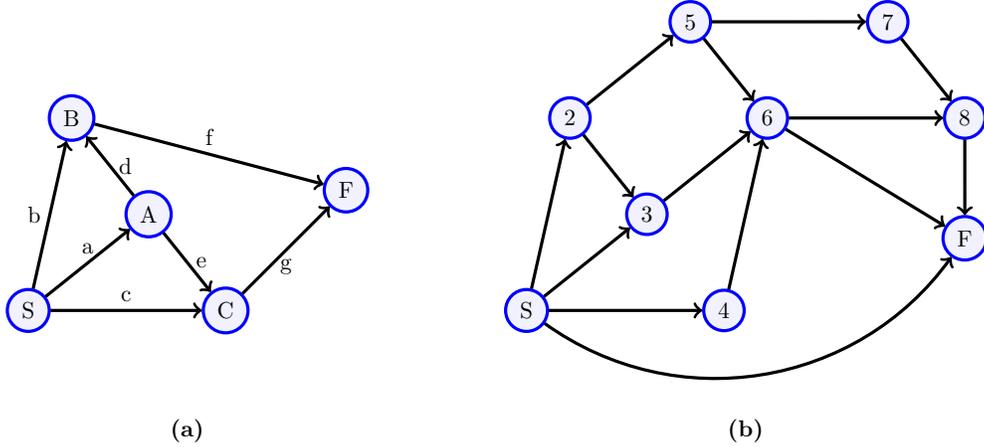
Figure 3: The PERT diagrams of the two projects considered in Section 5.3.

the unforseen delays of all tasks are bounded, as is their sum, i.e.,

$$\sum_{e\in\mathcal{E}}\delta_e \leq B, \quad \text{and} \quad 0 \leq \delta_e \leq b, \quad \forall e \in \mathcal{E},$$

where B and b are known parameters.

The manager has the ability of expediting some of the tasks by allocating a scarce resource to them, e.g., by assigning extra workforce or by spending money on purchasing/ upgrading task-related equipment. Let z_e denote the amount of the resource allocated to task $e \in \mathcal{E}$. We assume that the resource is divisible.[10] The associated reduction in processing time of task e is then $r\,z_e$, where r is the time reduction rate per unit of resource allocated. The maximum resource amount that can be allocated to any task is equal to c.

The manager has C units of the resource that need to be allocated so as to minimize the worst-case completion time T of the project. If we denote the time that event $\nu \in \mathcal{N}$ occurs with t_ν, the

---

[10]One could extend our results in case the resource is indivisible, by simply enforcing the variables z_e to be discrete.

following formulation captures the manager's problem:

$$\text{minimize} \quad T \tag{15a}$$

$$\text{subject to} \quad t_F \leq T \tag{15b}$$

$$t_S = 0 \tag{15c}$$

$$t_{f(e)} - t_{s(e)} \geq \tau_e^0 + \delta_e - r\, z_e, \quad \forall e \in \mathcal{E}, \quad \forall \delta \in U \tag{15d}$$

$$\mathbf{1}^T z = C \tag{15e}$$

$$z \leq c \tag{15f}$$

$$z \geq 0, \tag{15g}$$

with variables $T \in \mathbb{R}$, $t \in \mathbb{R}^{|\mathcal{N}|}$, $z \in \mathbb{R}^{|\mathcal{E}|}$ and $U = \{\delta \in \mathbb{R}^{|\mathcal{E}|} : 0 \leq \delta \leq b \cdot \mathbf{1}, \mathbf{1}^T \delta \leq B\}$. Constraints (15b)-(15c) determine the end and start time of the project. Constraint (15d) enforces logical precedence of tasks and events as discussed above. Constraints (15e)-(15f) limit the amount of the resource to be allocated.

Problem (15) is an adjustable RO problem, with fixed recourse (completion times of the project T and the events t). To solve it, we approximate it with its AARC, similarly to the inventory management problem in the previous section.

For our notion of PRO solutions, we require from robustly optimal solutions to maximize slack of constraint (15b), as it translates to potential reduction of the worst-case value of T (depending on the actual delay realizations). That is, we again employ our methodology from Section 4.1 using a slack value vector equal to a unit vector, where the component corresponding to constraint (15b) is equal to one.

We study separately two PERT graphs, depicted in Figure 3:

(a) the topology from Ben-Tal et al. (2009) we discussed thus far, and

(b) a more complicated topology.

Data. We consider 10,000 instances for each case (a) and (b). The standard processing times $\tau^0 \in \mathbb{R}^{|\mathcal{E}|}$ are independently sampled from a uniformly distributed random variable (a) between 1 and 10, and (b) between 5 and 20. Parameters B and b that bound the unforseen delays are set to 6 and 2 for case (a), respectively, and 25 and 5 for case (b). The processing time reduction rate
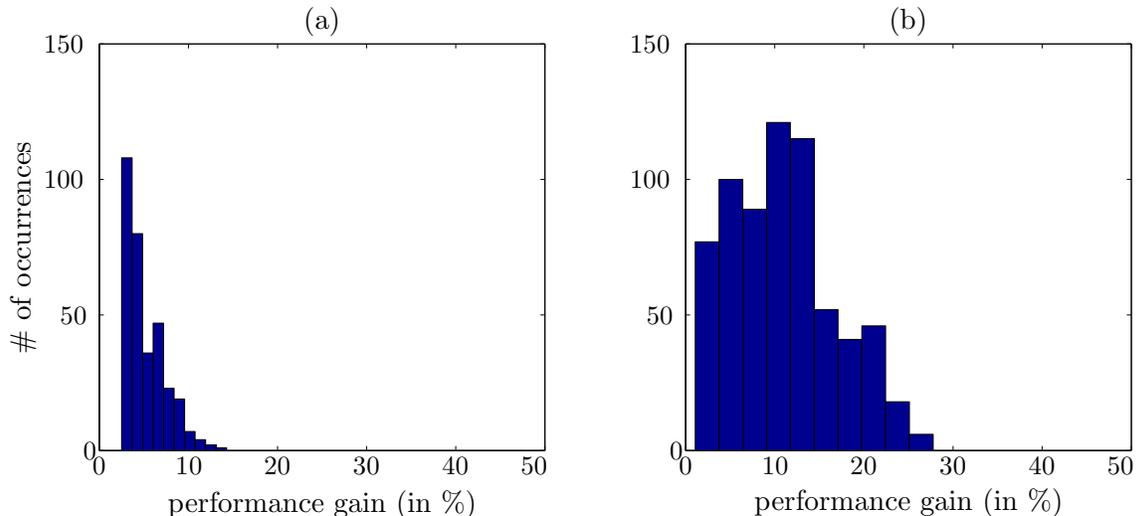
Figure 4: Histograms of maximal performance gains by considering PRO solutions in the numerical studies of Section 5.3.

r is (a) 1 and (b) 5. The maximum allowable resource at each task c is set to 1 for (a). For case (b), we replace constraint (15f) with $\tau_e^0 \geq r\, z_e$, $\forall e \in \mathcal{E}$, in order to ensure nonnegativity of task processing times. The resource amount available C is set to (a) 3 and (b) 10.

Results. The average number of instances in which a robustly optimal solution was identified that was not a PRO solution was 3.5% for case (a) and 7% for case (b). For these instances, we computed the maximum performance gap (in terms of actual completion time) one could get if the PRO solution were used instead of the dominated robustly optimal solution we obtained. A histogram of the associated (relative) performance gaps is depicted in Figure 4. The median performance gap recorded was (a) 4.5% and (b) 10.5%, while the maximum was as high as (a) 14.5% and (b) 28%.

6 Conclusions

In this paper, we adapted the well known concept of Pareto efficiency in the context of the popular robust optimization (RO) methodology. We argued that, by focusing exclusively on worst-case outcomes, the classical RO paradigm need not produce solutions that are also Pareto optimal, leading to inefficiencies and sub-optimal performance in practice. We provided a basic theoretical characterization of Pareto robustly optimal (PRO) solutions, and extended the RO framework by

proposing practical methods that verify Pareto optimality, and generate solutions that are PRO. Critically important, our approach involves solving optimization problems that are of the same complexity as the underlying robust problems, hence the potential improvements come at essentially no computational cost. Our numerical experiments, drawn from three different application areas (portfolio optimization, inventory management, and project management), demonstrated that PRO solutions have a significant upside compared with solutions obtained via classical RO methods, at no extra cost or downside.

References

Adida E, Pradnya J. 2009. A robust optimisation approach to project scheduling and resource allocation. *International J. of Services, Operations and Informatics* **4**(2) 169–193.

Adida E, Perakis G. 2006. A robust optimization approach to dynamic pricing and inventory control with no backorders. *Math. Prog.* **107** 97–129.

Bandi C, Bertsimas D. 2012. Tractable stochastic analysis in high dimensions via robust optimization. *Math. Prog.* **134** 23–70.

Ben-Tal A, Bertsimas D, Brown DB. 2010. A soft robust model for optimization under ambiguity. *Oper. Res.* **58** 1220–1234.

Ben-Tal A, El-Ghaoui L, Nemirovski A. 2009. *Robust Optimization*. Princeton Series in Applied Mathematics, Princeton University Press.

Ben-Tal A, Boaz G, Nemirovski A, Vial J-P. 2005. Retailer-supplier flexible commitments contracts: A robust optimization approach. *Manufacturing & Service Oper. Management* **7**(3) 248–271.

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

Ben-Tal A, Margalit T, Nemirovski A. 2000. Robust modeling of multi-stage portfolio problems. H. Frenk, K. Roos, T. Terlaky, S. Zhang, eds., *High Performance Optimization*. Kluwer Academic Publishers, 303–328.

Ben-Tal A, Nemirovski A. 2002. Robust optimization - methodology and applications. *Math. Prog.* **92**(3) 453–480.

Ben-Tal A, Nemirovski A. 2007. Selected topics in robust convex optimization. *Math. Prog.* **112**(1) 125–158.

Bertsimas D, Caramanis C, Brown DB. 2011a. Theory and Applications of Robust Optimization. *SIAM Review* **53**(3) 464–501.

Bertsimas D, Farias VF, Trichakis N. 2012. On the efficiency-fairness tradeoff. *Management Sci.* To appear.

Bertsimas D, Iancu DA, Parrilo PA. 2010. Optimality of Affine Policies in Multistage Robust Optimization. *Math. of Oper. Res.* **35**(2) 363–394.

Bertsimas D, Iancu DA, Parrilo PA. 2011b. A hierarchy of near-optimal policies for multistage adaptive optimization. *IEEE Transactions on Automatic Control* **56**(12) 2809 –2824.

Bertsimas D, Litvinov E, Sun X, Zhao J, Zheng T. 2011c. Adaptive robust optimization for the security constrained unit commitment problem. Working paper.

Bertsimas D, Pachamanova D. 2008. Robust multiperiod portfolio management in the presence of transaction costs. *Computers and Oper. Res.* **35**(1) 3–17.

Bertsimas D, Sim M. 2003. Robust Discrete Optimization and Network Flows. *Math. Prog.* **98**(1-3) 49–71.

Bertsimas D, Thiele A. 2006. A Robust Optimization Approach to Inventory Theory. *Oper. Res.* **54**(1) 150–168.

Bienstock D, Özbay N. 2008. Computing robust basestock levels. *Discrete Optimization* **5**(2) 389 – 414.

Boyd S, Vandenberghe L. 2004. *Convex Optimization*. Cambridge University Press.

Calafiore G. 2007. Ambiguous risk measures and optimal robust portfolios. *SIAM J. on Optimization* **18**(3) 853–877.

Chan TCY, Bortfeld T, Tsitsiklis JN. 2006. A robust approach to IMRT optimization. *Physics in Medicine and Biology* **51**(10) 2567.

Chen W, Unkelback J, Trofimov A, Madden T, Kooy H, Bortfeld T, Craft D. 2011. Including robustness in multi-criteria optimization for intensity modulated proton therapy. Working paper.

Cohen I, Golany B, Shtub A. 2007. The stochastic time-cost tradeoff problem: A robust optimization approach. *Networks* **49** 175–188.

Cristobal M A, Parks GT, Clarkson PJ. 2006. Finding robust solutions to multi-objective optimisation problems using polynomial chaos. *Proceedings of the 6th ASMO UK/ISSMO Conference on Engineering Design Optimization*.

Kalyanmoy D, Gupta H. 2005. Searching for robust pareto-optimal solutions in multi-objective optimization. Carlos Coello Coello, Arturo Hernández Aguirre, Eckart Zitzler, eds., *Evolutionary Multi-Criterion Optimization*, *Lecture Notes in Computer Science*, vol. 3410. Springer Berlin / Heidelberg, 150–164.

El-Ghaoui L, Oustry F, Lebret H. 1998. Robust Solutions to Uncertain Semidefinite Programs. *SIAM J. Optimization* **9**(1) 33–52.

Fabozzi F, Kolm P, Pachamanova D, Focardi S. 2007. *Robust Portfolio Optimization and Management*. Frank J. Fabozzi Series, Wiley.

Fischetti M, Monaci M. 2009. Light robustness. R. Ahuja, R. Möhring, C. Zaroliagis, eds., *Robust and Online Large-Scale Optimization*, *Lecture Notes in Computer Science*, vol. 5868. Springer Berlin / Heidelberg, 61–84.

Gabrel V, Murat C, Thiele A. 2012. Recent advances in robust optimization and robustness: An overview. Working paper.

Gabrel V, Murat C, Wu L. 2011. New models for the robust shortest path problem: Complexity, resolution and generalization. *Annals of Oper. Res.* 1–24.

Goh J, Hall N, Sim M. 2010. Robust optimization strategies for total cost control in project management. Working paper.

Goldfarb D, Iyengar G. 2003. Robust Portfolio Selection Problems. *Math. of Oper. Res.* **28**(1) 1–38.

Gorissen BL, den Hertog D. 2011. Approximating the pareto set of multiobjective linear programs via robust optimization. CentER Discussion Paper, 2012-031.

Iancu DA, Sharma M, Sviridenko M. 2012. Supermodularity and affine policies in dynamic robust optimization. Submitted for publication.

Joshi S, Boyd S. 2009. An efficient method for large-scale slack allocation. *Engineering Optimization* **41**(12) 1163–1176.

Kalai R, Lamboray R, Vanderpooten D. 2012. Lexicographic α-robustness: An alternative to min–max criteria. *Eur. J. of Oper. Res.* **220**(3) 722 – 728.

Lim AEB, Shanthikumar JG. 2007. Relative entropy, exponential utility, and robust dynamic pricing. *Oper. Res.* **55**(2) 198–214.

Luo B, Jinhua Z. 2008. A new methodology for searching robust pareto optimal solutions with moeas. *Proceedings of IEEE World Congress on Computational Intelligence*. 580 –586.

Luo Z-Q, Pang J-S, Ralph D. 1996. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press.

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

Ogryczak W. 1997. On the lexicographic minimax approach to location problems. *Eur. J. of Oper. Res.* **100**(3) 566 – 585.

Ono S, Nakayama S. 2009. Multi-objective particle swarm optimization for robust optimization and its hybridization with gradient search. *IEEE Congress on Evolutionary Computation*. 1629 –1636.

Perakis G, Roels G. 2010. Robust controls for network revenue management. *Manufacturing & Service Oper. Management* **12**(1) 56–76.

Rockafellar T. 1970. *Convex Analysis*. Princeton University Press.

Roy B. 2010. Robustness in operational research and decision aiding: A multi-faceted issue. *Eur. J. of Oper. Res.* **200**(3) 629 – 638.

Rusmevichientong P, Topaloglu H. 2011. Robust assortment optimization in revenue management under the multinomial logit choice model. Working paper.

Savage LJ. 1972. *The Foundations of Statistics*. 2nd ed. Dover Publications, Inc.

Schrijver A. 2000. *Theory of Linear and Integer Programming*. 2nd ed. John Wiley & Sons.

See C-T, Sim M. 2010. Robust approximation to multiperiod inventory management. *Oper. Res.* **58**(3) 583–594.

Soyster AL. 1973. Technical Note–Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming. *Oper. Res.* **21**(5) 1154–1157.

Sturm J, Imre P, Terlaky T. 2006. Sedumi 1.3. URL http://sedumi.ie.lehigh.edu/.

Suh M, Lee T. 2001. Robust optimization method for the economic term in chemical process design and planning. *Industrial & Engineering Chemistry Research* **40**(25) 5950–5959.

Young HP. 1995. *Equity: In Theory and Practice*. Princeton University Press.