Computers and Chemical Engineering 94 (2016) 75–91

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Comprehensive Pareto Efficiency in robust counterpart optimization

Ke Shang [a,*], Zuren Feng [a], Liangjun Ke [a], Felix T.S. Chan [b]

[a] State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an, China
[b] Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

ARTICLE INFO

Article history:
Received 30 March 2016
Received in revised form 27 June 2016
Accepted 26 July 2016
Available online 3 August 2016

Keywords:
Robust optimization
Pareto optimality
Linear programming
Integer programming

ABSTRACT

In this paper, an innovative concept named Comprehensive Pareto Efficiency is introduced in the context of robust counterpart optimization, which consists of three sub-concepts: Pareto Robust Optimality (PRO), Global Pareto Robust Optimality (GPRO) and Elite Pareto Robust Optimality (EPRO). Different algorithms are developed for computing robust solutions with respect to these three sub-concepts. As all sub-concepts are based on the Probability of Constraint Violation (PCV), formulations of PCV under different probability distributions are derived and an alternative way to calculate PCV is also presented. Numerical studies are drawn from two applications (production planning problem and orienteering problem), to demonstrate the Comprehensive Pareto Efficiency. The numerical results show that the Comprehensive Pareto Efficiency has important significance for practical applications in terms of the evaluation of the quality of robust solutions and the analysis of the difference between different robust counterparts, which provides a new perspective for robust counterpart optimization.

1. Introduction

Robust optimization (RO), originally introduced by Soyster (1973) and later revitalized by Ben-Tal and Nemirovski (2002), Bertsimas and Sim (2004), El Ghaoui et al. (1998) in late 1990s and early 2000s, is a technique for handling uncertainties in mathematical programming problems. In RO, an uncertainty set is firstly determined, then a robust counterpart (RC) of the original optimization problem is formulated where the solution should be feasible for any uncertain realizations in the uncertainty set. The objective of RO is to calculate a robust solution which satisfies a decision-maker's requirement (e.g., a robust solution with high quality objective value and reliability). For general review and comprehensive explanation on RO, we refer to Ben-Tal and Nemirovski (2002), Ben-Tal et al. (2009), Bertsimas et al. (2011), Gabrel et al. (2014), Gorissen et al. (2015).

The definition of the uncertainty set plays an important role in RO. It directly determines the underlying RC and then affects the whole process of RO. Many works have devoted to the construction of the uncertainty sets. The first one is considered by Soyster (1973) in which all possible uncertain realizations are included. This uncertainty set is too pessimistic and conservative which is not preferred in practice. Later El Ghaoui et al. (1998), Ben-Tal and Nemirovski (2002) consider ellipsoidal uncertainty sets and the resulting RC is a second-order cone programming (SOCP). Bertsimas and Sim (2004) define a budgeted uncertainty set which leads to a linear programming (LP). This uncertainty set is further improved by Ke et al. (2013) which is called proportion-based uncertainty set specifically suitable for 0–1 integer programming problems. Bertsimas et al. (2004) generalize the definition of the uncertainty sets by more general norms. In particular, the l_1 and l_∞ norms result in linear programming problems, and the l_2 norm results in a second-order cone programming problem. Li et al. (2011) presents a systematic study on different uncertainty sets defined by different norms and their combinations for linear and mixed integer programming problems and derived corresponding RC. Other works related to uncertainty set construction include Bertsimas and Brown (2009) which construct the uncertainty set from coherent risk measures perspective, Ben-Tal et al. (2013), Bertsimas et al. (2013) construct the uncertainty set from a data-driven and statistics perspective, etc.

With the uncertainty set defined, the robust optimal solution can be obtained by solving the corresponding RC. One important procedure is to check the quality of the robust solution, in order to make the right decision. One criterion of the solution quality is the objective value. When the uncertainty lies in the constraint, then the Probability of Constraint Violation (PCV) naturally becomes another criterion of the solution quality. For a decision maker, a solution with better objective value and lower PCV is always preferred. Many works have devoted to establishing the Probability Bounds of Constraint Violation (PBCV) when the distribution

* Corresponding author at: State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China.
E-mail address: kshang@foxmail.com (K. Shang).

http://dx.doi.org/10.1016/j.compchemeng.2016.07.022
0098-1354/

information is unknown or partially known, we refer to Bertsimas and Sim (2004), Lin et al. (2004), Paschalidis and Kang (2015), Ben-Tal et al. (2009), Li et al. (2012), Guzman et al. (2016). To the best of our knowledge, there is no systemic work on establishing PCV with known probability distributions in RO.

Iancu and Trichakis (2013) first introduced the concept of Pareto efficiency in the context of the RO methodology for linear optimization problems. The traditional RO optimizes the objective by satisfying the uncertain constraints under all possible uncertain realizations. However, as pointed out by Iancu and Trichakis (2013), the RO does not optimize the slacks of constraints, in fact, it fails to guarantee that no other solution exists yielding larger slacks and at the same objective value. Iancu and Trichakis (2013) defines the concept of Pareto Robust Optimality (PRO) based on constraint slacks. For a decision maker, the PRO solutions are always preferred to the non-PRO solutions as the non-PRO solutions can more readily generate infeasibility. One problem exists in Iancu and Trichakis (2013) is that comparing solutions feasibility by constraint slacks is very intuitive and sometimes not so accurate. A solution with less constraint slacks may has higher feasibility. Instead the PCV is the most accurate measurement of solution feasibility. Based on this observation, we can redefine the PRO by using PCV rather than constraint slacks to improve the accuracy. The premise is that the probability distribution information is known beforehand.

In this paper, we introduce an innovative concept named Comprehensive Pareto Efficiency in the context of robust counterpart optimization for linear and 0–1 integer programming problems with uncertain constraints. The main contributions are as follows:

1. Comprehensive Pareto Efficiency is initially introduced which consists of three sub-concepts: Pareto Robust Optimality (PRO), Global Pareto Robust Optimality (GPRO) and Elite Pareto Robust Optimality (EPRO).
2. Different algorithms are developed for computing robust solutions with respect to PRO, GPRO and EPRO.
3. Formulations of PCV under different probability distributions are derived, and an alternative way for calculating PCV is also presented.
4. We draw numerical studies on two applications (production planning problem and orienteering problem), to demonstrate the Comprehensive Pareto Efficiency in terms of the evaluation of the quality of robust solutions and the analysis of the difference between different robust counterparts.

The remainder of the paper is organized as follows: Section 2 reviews the robust counterpart optimization methodology, Section 3 introduces the Comprehensive Pareto Efficiency concept which consists of three sub-concepts, Section 4 describes the calculation of PCV under different probability distributions, Numerical studies are drawn in Section 5 with two applications and Section 6 concludes the whole paper.

2. Robust counterpart optimization

In this paper, we consider the following linear programming problem and 0–1 integer programming problem simultaneously:

$$\text{LP: } \max\left\{\mathbf{c}^{\mathrm{T}}\mathbf{x} : \mathbf{A}\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0}\right\} \tag{1a}$$

$$\text{0-1 IP: } \max\left\{\mathbf{c}^{\mathrm{T}}\mathbf{x} : \mathbf{A}\mathbf{x} \le \mathbf{b}, \mathbf{x} \in \{0,1\}^N\right\} \tag{1b}$$

where $\mathbf{c} \in \mathbb{R}^N$, $\mathbf{A} \in \mathbb{R}^{M\times N}$, $\mathbf{b} \in \mathbb{R}^M$.

We only consider single uncertain constraint in this paper. Suppose the ith row of $\mathbf{A}$ is affected by uncertainty, denote the transpose of ith row of $\mathbf{A}$ as uncertain vector $\mathrm{a}_i \in \mathbb{R}^N$, each element in a_i is modeled as independent and symmetric random variable. Then the ith constraint of the nominal linear programming problem and 0–1 integer programming problem turns into

$$\mathrm{a}_i^{\mathrm{T}}\mathbf{x} \le b_i \tag{2}$$

The robust optimization methodology is thus to define a so-called uncertainty set $\mathcal{U}$ for uncertain vector a_i such that the ith constraint satisfied as:

$$\mathrm{a}_i^{\mathrm{T}}\mathbf{x} \le b_i, \quad \forall \mathrm{a}_i \in \mathcal{U} \tag{3}$$

which is known as the robust counterpart of the uncertain constraint (2).

Without loss of generality, the uncertainty set $\mathcal{U}$ is defined as follows:

$$\mathcal{U} = \{\mathrm{a}_i = \mathbf{a}_i + \mathrm{A}_i\zeta \mid \zeta \in \mathcal{Z}\} \tag{4}$$

where $\mathbf{a}_i$ is the nominal value, $\mathrm{A}_i = diag(\mathrm{a}_i)$ is the perturbation set where $\mathrm{a}_i \in \mathbb{R}^{N+}$ is the perturbation vector, $\zeta \in \mathbb{R}^N$ is the vector of primitive uncertainties, and $\mathcal{Z}$ is a convex set which can be defined by a general norm of ζ as follows:

$$\mathcal{Z} = \{\zeta \in \mathbb{R}^N \mid \left\|\zeta\right\| \le \Delta\} \tag{5}$$

where $\|\cdot\|$ is any norm and Δ is the parameter controlling the size of $\mathcal{Z}$.

The key of RO is the definition of the set $\mathcal{Z}$, a particular $\mathcal{Z}$ directly determines the corresponding robust counterpart. One concern in RO is the tractability of the robust counterpart. The norm defined by ℓ_∞, ℓ_1 and ℓ_2 will lead to tractable robust counterparts (Bertsimas et al., 2004). The uncertainty set defined by ℓ_∞ is called box uncertainty set:

$$\mathcal{Z}_\infty = \{\zeta \in \mathbb{R}^N \mid \left\|\zeta\right\|_\infty \le \Theta\} \tag{6}$$

then the linear programming problem and 0–1 integer programming problem have the same robust counterpart:

$$RC_\infty = \max\left\{\mathbf{c}^{\mathrm{T}}\mathbf{x} : \mathbf{a}_i^{\mathrm{T}}\mathbf{x} + \Theta\mathrm{a}_i^{\mathrm{T}}\mathbf{x} \le b_i\right\} \tag{7}$$

In RC_∞ only the ith constraint is presented and other constraints are eliminated to keep it concise. In the following robust counterparts we only present the ith constraint as a convention, and a robust counterpart represents a linear and an 0–1 integer programming problem simultaneously.

The uncertainty set defined by ℓ_1 is called polyhedral uncertainty set:

$$\mathcal{Z}_1 = \{\zeta \in \mathbb{R}^N \mid \left\|\zeta\right\|_1 \le \Gamma\} \tag{8}$$

and the corresponding robust counterpart is:

$$RC_1 = \max\left\{\mathbf{c}^{\mathrm{T}}\mathbf{x} : \begin{array}{l}\mathbf{a}_i^{\mathrm{T}}\mathbf{x} + z\Gamma \le b_i \\ z \ge \hat{a}_{ij}x_j, \quad \forall j\end{array}\right\} \tag{9}$$

The uncertainty set defined by ℓ_2 is called ellipsoidal uncertainty set:

$$\mathcal{Z}_2 = \{\zeta \in \mathbb{R}^N \mid \left\|\zeta\right\|_2 \le \Omega\} \tag{10}$$

and the corresponding robust counterpart is:

$$RC_2 = \max\left\{\mathbf{c}^{\mathrm{T}}\mathbf{x} : \mathbf{a}_i^{\mathrm{T}}\mathbf{x} + \Omega\sqrt{\mathbf{x}^{\mathrm{T}}\mathrm{A}_i^2\mathbf{x}} \le b_i\right\} \tag{11}$$

The above three uncertainty sets are applicable when random vector ζ is unbounded, if the random vector ζ is bounded in an interval, specifically consider $\zeta \in [-1, 1]^N$, the above three uncertainty sets need to be bounded in order to limit ζ in its bound. This leads to two more uncertainty sets which are applicable when $\zeta \in [-1, 1]^N$, the first one is the intersection of the box and polyhedral sets:

$$\mathcal{Z}_{1\cap\infty} = \{\zeta \in \mathbb{R}^N \mid \left\|\zeta\right\|_1 \le \Gamma, \left\|\zeta\right\|_\infty \le 1\} \tag{12}$$

where $\Gamma \in [0, N]$ and the corresponding robust counterpart is:

$$RC_{1\cap\infty} = \max \left\{ \mathbf{c}^T\mathbf{x} : \begin{array}{l} \mathbf{a}_i^T\mathbf{x} + z\Gamma + \mathbf{e}^T\mathbf{p} \le b_i \\ z + p_j \ge \hat{a}_{ij}x_j, \quad \forall j \\ \mathbf{p} \ge \mathbf{0}, z \ge 0 \end{array} \right\} \tag{13}$$

where $\mathbf{e}$ is a N-dimensional vector with all elements equal 1.

The second one is the intersection of the box and ellipsoidal sets:

$$\mathcal{Z}_{2\cap\infty} = \{\zeta \in \mathbb{R}^N | \|\zeta\|_2 \le \Omega, \|\zeta\|_\infty \le 1\} \tag{14}$$

where $\Omega \in [0, \sqrt{N}]$ and the corresponding robust counterpart is:

$$RC_{2\cap\infty} = \max \left\{ \mathbf{c}^T\mathbf{x} : \begin{array}{l} \mathbf{a}_i^T\mathbf{x} + \mathrm{a}_i^T\mathbf{y} + \Omega\sqrt{\mathbf{z}^T\mathrm{A}_i^2\mathbf{z}} \le b_i \\ \mathbf{y} + \mathbf{z} = \mathbf{x}, \mathbf{y} \ge \mathbf{0} \end{array} \right\} \tag{15}$$

The uncertainty set $\mathcal{Z}_\infty$ is also applicable for bounded $\zeta \in [-1, 1]^N$ when parameter $\Theta \in [0, 1]$.

For a detailed summarization of all uncertainty sets defined by norms and their combinations, as well as the deduction of the robust counterparts from their corresponding uncertainty sets, we refer to Li et al. (2011).

3. Comprehensive Pareto Efficiency

3.1. Pareto Robust Optimality

Iancu and Trichakis (2013) first introduced the concept of Pareto efficiency in the context of the RO methodology for linear optimization problems. When the uncertainty lies in the constraints, the RO optimizes the objective by satisfying the uncertain constraints under all possible uncertain realizations. However the RO does not optimize the slacks of constraints, in fact, it fails to guarantee that no other solution exists yielding larger slacks and at the same objective value.

Consider the ith constraint under uncertainty set $\mathcal{U}$ as described in (3). The slack of the uncertain constraint is defined as:

$$SLK(\mathbf{x}, \mathrm{a}_i) = b_i - \mathrm{a}_i^T\mathbf{x}, \quad \forall \mathbf{x} \in \mathbb{R}^N, \mathrm{a}_i \in \mathcal{U} \tag{16}$$

Then Iancu and Trichakis (2013) defined the Pareto Robustly Optimal (PRO) solution as follows:

Definition 3.1 *(Iancu and Trichakis (2013))*. A solution $\mathbf{x}$ is called a PRO solution under constraint (3) if it is robustly optimal, and there is no other robust optimal solution $\bar{\mathbf{x}}$ such that

1. $SLK(\bar{\mathbf{x}}, \mathrm{a}_i) \ge SLK(\mathbf{x}, \mathrm{a}_i), \quad \forall \mathrm{a}_i \in \mathcal{U}$
2. $SLK(\bar{\mathbf{x}}, \bar{\mathrm{a}}_i) > SLK(\mathbf{x}, \bar{\mathrm{a}}_i), \quad \exists \bar{\mathrm{a}}_i \in \mathcal{U}$

In the above setting we say that $\bar{\mathbf{x}}$ Pareto dominates $\mathbf{x}$, the PRO solutions guarantee the slacks in the constraints are optimized for all uncertainty realizations $\mathrm{a}_i \in \mathcal{U}$. Iancu and Trichakis (2013) pointed out the reason for this setting is that solution with zero or small value of slack can more readily generate infeasibility, which implicitly means a PRO solution has the highest reliability compare with the non-PRO solutions.

The measurement of the uncertain constraint feasibility by slacks is very intuitive and sometimes not so accurate. If the probability distribution of the uncertainty is known, a more accurate measurement of the constraint feasibility is the so-called Probability of Constraint Violation (PCV), which is defined as follows:

$$PCV(\mathbf{x}) = Pr\left(\mathrm{a}_i^T\mathbf{x} > b_i\right), \quad \forall \mathbf{x} \in \mathbb{R}^N \tag{17}$$

A solution with zero or small value of PCV means it hardly generates infeasibility, which means this solution is with high reliability. This motivates us to define a more accurate PRO concept as follows:

Definition 3.2 *(PRO)*. A solution $\mathbf{x}$ is called a PRO solution under constraint (3) if it is robustly optimal, and there is no other robust optimal solution $\bar{\mathbf{x}}$ such that

$$PCV(\bar{\mathbf{x}}) < PCV(\mathbf{x})$$

In the definition above, we say that $\bar{\mathbf{x}}$ Pareto dominates $\mathbf{x}$. The new PRO in Definition 3.2 is much simpler than the old PRO in Definition 3.1, because we use PCV as the measurement of the uncertain constraint feasibility rather than constraint slacks. The old PRO definition needs to compare all slacks with all possible values inside the uncertainty set. By using PCV, we only need to compare the PCV of two solutions.

Below a toy example is presented to illustrate the notion of the new PRO concept.

Example 3.1. Consider the following nominal 0–1 integer programming:

$$\text{maximize} \quad c_1x_1 + c_2x_2 + c_3x_3 \tag{18a}$$

$$\text{subject to} \quad a_1x_1 + a_2x_2 + a_3x_3 \le b \tag{18b}$$

$$x_1 + x_2 = 1 \tag{18c}$$

$$x_2 - x_3 = 0 \tag{18d}$$

$$\{x_1, x_2, x_3\} \in \{0, 1\}^3 \tag{18e}$$

Suppose $\tilde{a}_1, \tilde{a}_2$ and $\tilde{a}_3$ obey uniform distributions and $\hat{a}_1 = \hat{a}_2 = \hat{a}_3 = 1$. Denote the primitive uncertainties as ζ_1, ζ_2 and ζ_3 which are uniformly distributed on $[-1, 1]$. There are two possible solutions in this example, $\mathbf{x}_1 = (1, 0, 0)$, $\mathbf{x}_2 = (0, 1, 1)$.

Case 1: $a_1 = 4$, $a_2 = a_3 = 2$, $c_1 = 2$, $c_2 = c_3 = 1$, $b = 5$. Here we apply $RC_{1\cap\infty}$ robust model, the robust optimal solution is

$$x^* = \begin{cases} \mathrm{x_1 \text{ or } x_2} & 0 \le \Gamma \le 1 \\ \text{no feasible solution} & \Gamma > 1 \end{cases} \tag{19}$$

We have:

1. $\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_2 = 2$.
2. $PCV(\mathbf{x}_1) = Pr(\zeta_1\hat{a}_1 > b - a_1) = Pr(\zeta_1 > 1) = 0$
   $PCV(\mathbf{x}_2) = Pr(\zeta_2\hat{a}_2 + \zeta_3\hat{a}_3 > b - a_2 - a_3) = Pr(\zeta_2 + \zeta_3 > 1) = 0.125$
   $PCV(\mathbf{x}_1) < PCV(\mathbf{x}_2)$

From Definition 3.2 we know that $\mathbf{x}_1$ is the PRO solution.

Case 2: $a_1 = 4.5$, $a_2 = a_3 = 2$, $c_1 = 2$, $c_2 = c_3 = 1$, $b = 5$. Here we apply RC_∞ robust model, the robust optimal solution is

$$x^* = \begin{cases} \mathrm{x_1 \text{ or } x_2} & 0 \le \Theta \le 0.5 \\ \text{no feasible solution} & \Theta > 0.5 \end{cases} \tag{20}$$

We have:

1. $\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_2 = 2$.
2. $PCV(\mathbf{x}_1) = Pr(\zeta_1\hat{a}_1 > b - a_1) = Pr(\zeta_1 > 0.5) = 0.25$
   $PCV(\mathbf{x}_2) = Pr(\zeta_2\hat{a}_2 + \zeta_3\hat{a}_3 > b - a_2 - a_3) = Pr(\zeta_2 + \zeta_3 > 1) = 0.125$
   $PCV(\mathbf{x}_1) > PCV(\mathbf{x}_2)$

From Definition 3.2 we know that $\mathbf{x}_2$ is the PRO solution.

Case 3: $a_1 = 5 - \frac{1}{\sqrt{2}}$, $a_2 = a_3 = 2$, $c_1 = 2$, $c_2 = c_3 = 1$, $b = 5$. Here we apply $RC_{2\cap\infty}$ robust model, the robust optimal solution is

$$x^* = \begin{cases} \mathrm{x_1 \text{ or } x_2} & 0 \le \Omega \le \dfrac{1}{\sqrt{2}} \\ \text{no feasible solution} & \Omega > \dfrac{1}{\sqrt{2}} \end{cases} \tag{21}$$

We have:

1. $\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_2 = 2$.
2. $PCV(\mathbf{x}_1) = Pr(\zeta_1 \hat{a}_1 > b - a_1) = Pr(\zeta_1 > \frac{1}{\sqrt{2}}) \approx 0.146$
   $PCV(\mathbf{x}_2) = Pr(\zeta_2 \hat{a}_2 + \zeta_3 \hat{a}_3 > b - a_2 - a_3) = Pr(\zeta_2 + \zeta_3 > 1) = 0.125$
   $PCV(\mathbf{x}_1) > PCV(\mathbf{x}_2)$

From Definition 3.2 we know that $\mathbf{x}_2$ is the PRO solution.

Remark. From the above example we have the following observations:

1. In order to calculate the PCV, the probability distributions of the uncertain variables need to be known, in the above example we assume the primitive uncertainties obey uniform distribution on $[-1, 1]$.
2. Three different robust counterpart models are applied on this example under different settings, all three models can obtain two robust optimal solutions. But all three models fail to distinguish the PRO solution and the non-PRO solution.
3. In practical applications, the PRO solution is always preferred, so it is necessary to find a method which can produce the PRO solutions rather than the non-PRO solutions.

The following proposition provides a method for calculating PRO solutions for a robust counterpart.

Proposition 1. *All the optimal solutions to* $\min_{\mathbf{x} \in X^{RO}} PCV(\mathbf{x})$ *are PRO under constraint (3), where* $X^{RO} = \{\mathbf{x} \in \mathbb{R}^N : \mathbf{c}^T\mathbf{x} \geq z^{RO}, SLK(\mathbf{x}, \mathrm{a}_i) \geq 0, \quad \forall \mathrm{a}_i \in \mathcal{U}\}$ *and* z^{RO} *is the optimal objective value under constraint (3).*

Proof. It is clear that X^{RO} is the robust optimal solution set under constraint (3). By solving $\min_{\mathbf{x} \in X^{RO}} PCV(\mathbf{x})$ over set X^{RO}, we can obtain a robust optimal solution $\mathbf{x}$ with the smallest PCV value. Based on the PRO definition 3.2, there is no other robust optimal solution $\bar{\mathbf{x}}$ such that $PCV(\bar{\mathbf{x}}) < PCV(\mathbf{x})$, which means $\mathbf{x}$ is a PRO solution. □

Proposition 1 suggests two methods for calculating PRO solutions. The first one is to solve the optimization problem $\min_{\mathbf{x} \in X^{RO}} PCV(\mathbf{x})$ directly. The objective function $PCV(\mathbf{x})$ is always a nonlinear and complex function which is difficult to solve in practice[1]. The second method is to obtain the robust optimal solution set X^{RO} first, and then filter the solution with the smallest PCV as the PRO solution. For 0–1 integer programming problem, the X^{RO} is always a finite set and can be obtained by standard software (e.g., CPLEX with Solution Pool feature), so we can get an accurate PRO solution in this situation. For linear programming problem, if there exists more than one robust optimal solution, the X^{RO} is always an infinite set. In this situation, it is difficult to get an accurate PRO solution, so we develop a heuristic algorithm to obtain an approximate PRO solution for this case. The heuristic algorithm is shown in Algorithm 1.

[1] For detailed PCV formulations see Section 4.

Algorithm 1. Calculate an approximate PRO solution for linear programming problem

```
Input: Uncertainty set U, S a set of sampled points in U, solution set X
Output: An approximate PRO solution of LP
    Set S = ∅ and X = ∅
    Randomly sample points in ri(U) where ri(U) is the relative interior
    of U and put the sampled points into set S.
    for all ā_i ∈ S do
      Solve max_{x ∈ X^RO} SLK(x, ā_i)
      Put the optimal solution into solution set X
    end for
    Filter the solution in X with the smallest PCV as an approximate
    PRO solution
```

In Algorithm 1, the idea is that to solve $\min_{\mathbf{x} \in X^{RO}} PCV(\mathbf{x})$ directly is difficult, but we can solve $\max_{\mathbf{x} \in X^{RO}} SLK(\mathbf{x}, \bar{\mathrm{a}}_i)$ easily because $SLK(\mathbf{x}, \bar{\mathrm{a}}_i)$ is a linear function. We can obtain a set of old PRO solutions in Definition 3.1 where the theoretical basis is Corollary 3 in Iancu and Trichakis (2013), then we filter the solution with the smallest PCV as an approximate PRO solution defined in this paper. As the *SLK* and *PCV* are defined upon single uncertain constraint, Algorithm 1 is designed specifically for single uncertain constraint case.

3.2. Global Pareto Robust Optimality

The PRO solution is defined with a given uncertainty set $\mathcal{U}$. As shown in Section 2, the uncertainty set $\mathcal{U}$ has a parameter Δ which controls its size. Tuning parameter Δ, different sized uncertainty sets can be generated. A PRO solution is associated with each of the uncertainty sets. In order to compare PRO solutions from different sized uncertainty sets, we introduce the concept of Global Pareto Robustly Optimal (GPRO) solution which extends the PRO concept to a globalized scope.

Without loss of generality, we consider the robust counterpart *RC* with uncertainty set $\mathcal{U}$ which is controlled by parameter Δ. Suppose we have a finite set which contains different values of parameter Δ, here denoted as $P = \{\Delta_i, 1 \leq i \leq N\}$ and $|P| = N$. For each value in P we can establish an uncertainty set, then a series of uncertainty sets with different sizes can be obtained, denoted as $\mathcal{U}(\Delta_i), 1 \leq i \leq N$. For each uncertainty set, we can get a PRO solution associated with it, then a set of PRO solutions can be obtained, denoted as $X^{PRO} = \{\mathbf{x}_i, 1 \leq i \leq N\}$.

In the following, we give the definition of GPRO solution over X^{PRO}.

Definition 3.3 *(GPRO).* Given a finite PRO solution set X^{PRO} of a robust counterpart *RC*. A PRO solution $\mathbf{x} \in X^{PRO}$ is called a GPRO solution over X^{PRO}, if there is no other PRO solution $\bar{\mathbf{x}} \in X^{PRO}$ such that:

1. $\mathbf{c}^T\bar{\mathbf{x}} = \mathbf{c}^T\mathbf{x}$ and $PCV(\bar{\mathbf{x}}) < PCV(\mathbf{x})$
2. $\mathbf{c}^T\bar{\mathbf{x}} > \mathbf{c}^T\mathbf{x}$ and $PCV(\bar{\mathbf{x}}) \leq PCV(\mathbf{x})$

In the definition above, we say that $\bar{\mathbf{x}}$ Pareto dominates $\mathbf{x}$. The PRO only considers a fixed size uncertainty set, the GPRO extends the PRO to a globalized scope which considers PRO solutions from different sized uncertainty sets. We need to consider different sized uncertainty sets and compare the objective values and PCV of the PRO solutions simultaneously.

Next a toy example is presented to illustrate the notion of the above GPRO concept.

Example 3.2. Here we consider the same 0–1 integer programming as presented in Example 3.1, see (18):

Case 1: $a_1 = 4.25$, $a_2 = a_3 = 2$, $c_1 = 2.5$, $c_2 = c_3 = 1$, $b = 5$. Here we apply $RC_{1\cap\infty}$ robust model, the robust optimal solution is

$$x^* = \begin{cases} x_1 & 0 \le \Gamma \le 0.75 \\ x_2 & 0.75 < \Gamma \le 1 \\ \text{no feasible solution} & \Gamma > 1 \end{cases} \quad (22)$$

Suppose we define parameter set $P = \{\Gamma = 0, 0.1, \ldots, 1\}$, then the PRO solution set $X^{PRO} = \{\mathbf{x}_1, \mathbf{x}_2\}$ and we have:

1. $\mathbf{c}^T\mathbf{x}_1 = 2.5 > \mathbf{c}^T\mathbf{x}_2 = 2$.
2. $PCV(\mathbf{x}_1) = Pr(\zeta_1\hat{a}_1 > b - a_1) = Pr(\zeta_1 > 0.75) = 0.125$ $PCV(\mathbf{x}_2) = Pr(\zeta_2\hat{a}_2 + \zeta_3\hat{a}_3 > b - a_2 - a_3) = Pr(\zeta_2 + \zeta_3 > 1) = 0.125$ $PCV(\mathbf{x}_1) = PCV(\mathbf{x}_2)$

From Definition 3.3 we know that $\mathbf{x}_1$ is the GPRO solution over X^{PRO}.

Case 2: $a_1 = 4.25$, $a_2 = a_3 = 2$, $c_1 = 1.5$, $c_2 = c_3 = 1$, $b = 5$. Here we apply RC_∞ robust model, the robust optimal solution is

$$x^* = \begin{cases} x_2 & 0 \le \Theta \le 0.5 \\ x_1 & 0.5 < \Theta \le 0.75 \\ \text{no feasible solution} & \Theta > 0.75 \end{cases} \quad (23)$$

Suppose we define parameter set $P = \{\Theta = 0, 0.1, \ldots, 1\}$, then the PRO solution set $X^{PRO} = \{\mathbf{x}_1, \mathbf{x}_2\}$ and we have:

1. $\mathbf{c}^T\mathbf{x}_1 = 1.5 < \mathbf{c}^T\mathbf{x}_2 = 2$.
2. $PCV(\mathbf{x}_1) = Pr(\zeta_1\hat{a}_1 > b - a_1) = Pr(\zeta_1 > 0.75) = 0.125$ $PCV(\mathbf{x}_2) = Pr(\zeta_2\hat{a}_2 + \zeta_3\hat{a}_3 > b - a_2 - a_3) = Pr(\zeta_2 + \zeta_3 > 1) = 0.125$ $PCV(\mathbf{x}_1) = PCV(\mathbf{x}_2)$

From Definition 3.3 we know that $\mathbf{x}_2$ is the GPRO solution over X^{PRO}.

Case 3: $a_1 = 4.25$, $a_2 = a_3 = 2$, $c_1 = 1.5$, $c_2 = c_3 = 1$, $b = 5$. Here we apply $RC_{2\cap\infty}$ robust model, the robust optimal solution is

$$x^* = \begin{cases} x_2 & 0 \le \Omega \le \frac{1}{\sqrt{2}} \\ x_1 & \frac{1}{\sqrt{2}} < \Omega \le 0.75 \\ \text{no feasible solution} & \Omega > 0.75 \end{cases} \quad (24)$$

Suppose we define parameter set $P = \{\Omega = 0, 0.05, 0.1, \ldots, 1\}$, then the PRO solution set $X^{PRO} = \{\mathbf{x}_1, \mathbf{x}_2\}$ and we have:

1. $\mathbf{c}^T\mathbf{x}_1 = 1.5 < \mathbf{c}^T\mathbf{x}_2 = 2$.
2. $PCV(\mathbf{x}_1) = Pr(\zeta_1\hat{a}_1 > b - a_1) = Pr(\zeta_1 > 0.75) = 0.125$ $PCV(\mathbf{x}_2) = Pr(\zeta_2\hat{a}_2 + \zeta_3\hat{a}_3 > b - a_2 - a_3) = Pr(\zeta_2 + \zeta_3 > 1) = 0.125$ $PCV(\mathbf{x}_1) = PCV(\mathbf{x}_2)$

From Definition 3.3 we know that $\mathbf{x}_2$ is the GPRO solution over X^{PRO}.

Remark. From the above example we have the following observations:

1. Three robust counterparts $RC_{1\cap\infty}$, $RC_{2\cap\infty}$ and RC_∞ are applied on this example. We can observe that PRO solutions obtained by three robust counterparts may not be GPRO solutions. This phenomenon is related to the distribution of the primitive uncertainties and the definition of the uncertainty sets. Different robust counterparts under different distributions may behave totally different.
2. The parameter set P determines PRO solution set X^{PRO}. In the above example, if we define $P = \{\Omega = 0, 0.1, \ldots, 1\}$ for case 3, then $X^{PRO} = \{\mathbf{x}_2\}$ and $\mathbf{x}_1$ cannot be obtained. In order to obtain a more complete PRO solution set X^{PRO}, the selection of P need to be denser.
3. In practical applications, the GPRO solution is always preferred. An ideal robust model will only produce GPRO solutions. From the above example we know that all three robust counterparts are not ideal under uniform distribution, the non-GPRO solutions cannot be avoided.

The following lemma describes the relations between two PRO solutions of a robust counterpart under different sized uncertainty sets.

Lemma 2. *Consider the general uncertainty set $\mathcal{U}$ controlled by parameter Δ as defined in (4) and (5). For any $\Delta_1 < \Delta_2$, $\mathbf{x}_1$ and $\mathbf{x}_2$ are PRO solutions under uncertainty sets $\mathcal{U}(\Delta_1)$ and $\mathcal{U}(\Delta_2)$ respectively, then*

1. *$\mathbf{x}_2$ is a feasible solution under uncertainty set $\mathcal{U}(\Delta_1)$*
2. $\mathbf{c}^T\mathbf{x}_1 \ge \mathbf{c}^T\mathbf{x}_2$
3. *if $\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_2$, then $PCV(\mathbf{x}_1) \le PCV(\mathbf{x}_2)$*

Proof.

(1) From the definition of uncertainty set $\mathcal{U}$, it is easy to know that $\mathcal{U}(\Delta_1) \subset \mathcal{U}(\Delta_2)$. The robust constraint (3) can be rewritten as:

$$\mathbf{a}_i^T\mathbf{x} + \max_{\zeta \in \mathcal{Z}} \left\{\zeta^T\hat{A}_i\mathbf{x}\right\} \le b_i \quad (25)$$

Denote $F(\mathbf{x}, \mathcal{Z}) = \max_{\zeta \in \mathcal{Z}} \left\{\zeta^T\hat{A}_i\mathbf{x}\right\}$. Because $\mathcal{Z}(\Delta_1) \subset \mathcal{Z}(\Delta_2)$, we have $F(\mathbf{x}_2, \mathcal{Z}(\Delta_1)) \le F(\mathbf{x}_2, \mathcal{Z}(\Delta_2))$; because $\mathbf{a}_i^T\mathbf{x}_2 + F(\mathbf{x}_2, \mathcal{Z}(\Delta_2)) \le b_i$, then we have $\mathbf{a}_i^T\mathbf{x}_2 + F(\mathbf{x}_2, \mathcal{Z}(\Delta_1)) \le b_i$, this implies $\mathbf{x}_2$ is a feasible solution under uncertainty set $\mathcal{U}(\Delta_1)$.

(2) Suppose $\mathbf{c}^T\mathbf{x}_1 < \mathbf{c}^T\mathbf{x}_2$, from above we know that $\mathbf{x}_2$ is a feasible solution under uncertainty set $\mathcal{U}(\Delta_1)$, this contradicts $\mathbf{x}_1$ is a PRO solution with uncertainty set $\mathcal{U}(\Delta_1)$, so the hypothesis is not established, and we conclude $\mathbf{c}^T\mathbf{x}_1 \ge \mathbf{c}^T\mathbf{x}_2$.

(3) If $\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_2$, suppose $PCV(\mathbf{x}_1) > PCV(\mathbf{x}_2)$, because $\mathbf{x}_2$ is a feasible solution under uncertainty set $\mathcal{U}(\Delta_1)$, this contradicts $\mathbf{x}_1$ is a PRO solution with uncertainty set $\mathcal{U}(\Delta_1)$, so the hypothesis is not established, and we conclude: if $\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_2$, then $PCV(\mathbf{x}_1) \le PCV(\mathbf{x}_2)$.

□

Next two theorems are established which show that specific robust counterpart under specific probability distribution can be an ideal robust model which will only produce GPRO solutions. The first one is robust counterpart RC_2 under Normal distribution which is described as follows:

Theorem 3.3. *If each element of the primitive uncertainties ζ_i obeys standard Normal distribution, then all PRO solutions in PRO solution set X^{PRO} generated by robust counterpart RC_2(11) are GPRO over X^{PRO}.*

Proof. If each element of the primitive uncertainties ζ_i obeys standard Normal distribution denoted as $Normal(0, 1)$, then the uncertain vector $\mathrm{a}_i = \mathbf{a}_i + \mathrm{A}_i\zeta$ obeys Normal distribution $Normal_N(\mathbf{a}_i, \mathrm{A}_i^2)$.

Consider the following chance constraint:

$$Pr\left(\mathrm{a}_i^T\mathbf{x} > b_i\right) < 1 - \varepsilon \quad (26)$$

It is well known that the above chance constraint is equivalent to the following formulation (Calafiore and El Ghaoui, 2006):

$$\mathbf{a}_i^T\mathbf{x} + \Phi^{-1}(\varepsilon)\sqrt{\mathbf{x}^T\mathrm{A}_i^2\mathbf{x}} \le b_i \quad (27)$$

where $\Phi(x)$ is the cumulative distribution function (CDF) of the standard normal distribution.

This is just the constraint in RC_2 with $\Omega = \Phi^{-1}(\varepsilon)$. Because $\Phi^{-1}(\varepsilon)$ is a monotonically increasing function of ε, for any $\varepsilon_1 < \varepsilon_2$, then $\Omega_1 < \Omega_2$. Suppose $\mathbf{x}_1$ is the PRO solution with $\mathcal{U}(\Omega_1)$ and $\mathbf{x}_2$ is the PRO solution with $\mathcal{U}(\Omega_2)$, from Lemma 2 we have $\mathbf{c}^T\mathbf{x}_1 \geq \mathbf{c}^T\mathbf{x}_2$.

(1) If $\mathbf{c}^T\mathbf{x}_1 > \mathbf{c}^T\mathbf{x}_2$, suppose $PCV(\mathbf{x}_1) \leq PCV(\mathbf{x}_2)$, because $PCV(\mathbf{x}_2) < 1-\varepsilon_2$, then we have $PCV(\mathbf{x}_1) < 1-\varepsilon_2$, this means $\mathbf{x}_1$ is a feasible solution with $\mathcal{U}(\Omega_2)$; because $\mathbf{c}^T\mathbf{x}_1 > \mathbf{c}^T\mathbf{x}_2$, this contradicts $\mathbf{x}_2$ is PRO solution under $\mathcal{U}(\Omega_2)$, so the hypothesis is not established, which means $PCV(\mathbf{x}_1) > PCV(\mathbf{x}_2)$.

(2) If $\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_2$, from Lemma 2 we have $PCV(\mathbf{x}_1) \leq PCV(\mathbf{x}_2)$. Suppose $PCV(\mathbf{x}_1) < PCV(\mathbf{x}_2)$, because $PCV(\mathbf{x}_2) < 1-\varepsilon_2$, then we have $PCV(\mathbf{x}_1) < 1-\varepsilon_2$, this means $\mathbf{x}_1$ is a feasible solution with $\mathcal{U}(\Omega_2)$; because $PCV(\mathbf{x}_1) < PCV(\mathbf{x}_2)$, this contradicts $\mathbf{x}_2$ is PRO solution under $\mathcal{U}(\Omega_2)$, so the hypothesis is not established, which means $PCV(\mathbf{x}_1) = PCV(\mathbf{x}_2)$.

We can conclude that: for any two PRO solutions $\mathbf{x}_1$ and $\mathbf{x}_2$, if $\mathbf{c}^T\mathbf{x}_1 > \mathbf{c}^T\mathbf{x}_2$, then $PCV(\mathbf{x}_1) > PCV(\mathbf{x}_2)$, if $\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_2$, then $PCV(\mathbf{x}_1) = PCV(\mathbf{x}_2)$. So for any X^{PRO}, there is no solution which can be dominated by another solution. From Definition 3.3 we know that all PRO solutions in X^{PRO} are GPRO. □

The second one is robust counterpart RC_∞ under Cauchy distribution, first we give the following lemma which describes the equivalence of chance constraint (26) and robust counterpart RC_∞:

Lemma 3.4. *If each element of the primitive uncertainties ζ_i obeys standard Cauchy distribution, then chance constraint (26) is equivalent to the following formulation:*

$$\mathbf{a}_i^T\mathbf{x} + \Psi^{-1}(\varepsilon)\mathrm{a}_i^T\mathbf{x} \leq b_i \tag{28}$$

where $\Psi(x)$ is the CDF of the standard Cauchy distribution.

Proof. If each element of the primitive uncertainties ζ_i obeys standard Cauchy distribution denoted as $Cauchy(0, 1)$, then from properties of Cauchy distribution we have: the uncertain vector $\tilde{\mathrm{a}}_i = \mathbf{a}_i + A_i\zeta$ obeys Cauchy distribution $Cauchy_N(\mathbf{a}_i, \mathrm{a}_i)$ and $\tilde{\mathrm{a}}_i^T\mathbf{x}$ obeys $Cauchy(\mathbf{a}_i^T\mathbf{x}, \mathrm{a}_i^T\mathbf{x})$.

Consider chance constraint (26), reformulate the left hand side:

$$Pr\left(\tilde{\mathrm{a}}_i^T\mathbf{x} > b_i\right) = Pr\left(\frac{\tilde{\mathrm{a}}_i^T\mathbf{x} - \mathbf{a}_i^T\mathbf{x}}{\mathrm{a}_i^T\mathbf{x}} > \frac{b_i - \mathbf{a}_i^T\mathbf{x}}{\mathrm{a}_i^T\mathbf{x}}\right) \tag{29}$$

Denote $Z = \frac{\tilde{\mathrm{a}}_i^T\mathbf{x} - \mathbf{a}_i^T\mathbf{x}}{\mathrm{a}_i^T\mathbf{x}}$, so Z obeys standard Cauchy distribution $Cauchy(0, 1)$.

Then the chance constraint (26) is equivalent to:

$$\frac{b_i - \mathbf{a}_i^T\mathbf{x}}{\mathrm{a}_i^T\mathbf{x}} \geq \Psi^{-1}(\varepsilon) \tag{30}$$

which is:

$$\mathbf{a}_i^T\mathbf{x} + \Psi^{-1}(\varepsilon)\mathrm{a}_i^T\mathbf{x} \leq b_i \tag{31}$$

where $\Psi(x)$ is the CDF of the standard Cauchy distribution. □

Based on Lemma 3.4, we have the following theorem:

Theorem 3.5. *If each element of the primitive uncertainties ζ_i obeys standard Cauchy distribution, then all PRO solutions in PRO solution set X^{PRO} generated by robust counterpart RC_∞(7) are GPRO over X^{PRO}.*

Proof. If each element of the primitive uncertainties ζ_i obeys standard Cauchy distribution, from Lemma 3.4 we known that chance constraint (26) is equivalent to the following formulation:

$$\mathbf{a}_i^T\mathbf{x} + \Psi^{-1}(\varepsilon)\mathrm{a}_i^T\mathbf{x} \leq b_i \tag{32}$$

where $\Psi(x)$ is the CDF of the standard Cauchy distribution.

This is just the constraint in RC_∞ with $\Theta = \Psi^{-1}(\varepsilon)$. Because $\Psi^{-1}(\varepsilon)$ is a monotonically increasing function of ε, for any $\varepsilon_1 < \varepsilon_2$, then $\Theta_1 < \Theta_2$. Suppose $\mathbf{x}_1$ is the PRO solution with $\mathcal{U}(\Theta_1)$ and $\mathbf{x}_2$ is the PRO solution with $\mathcal{U}(\Theta_2)$, from Lemma 2 we have $\mathbf{c}^T\mathbf{x}_1 \geq \mathbf{c}^T\mathbf{x}_2$.

Follow the proof way of Theorem 3.3 we can get: for any two PRO solutions $\mathbf{x}_1$ and $\mathbf{x}_2$, if $\mathbf{c}^T\mathbf{x}_1 > \mathbf{c}^T\mathbf{x}_2$, then $PCV(\mathbf{x}_1) > PCV(\mathbf{x}_2)$, if $\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_2$, then $PCV(\mathbf{x}_1) = PCV(\mathbf{x}_2)$. So for any X^{PRO}, there is no solution which can be dominated by another solution. From Definition 3.3 we know that all PRO solutions in X^{PRO} are GPRO. □

In Theorem 3.3 and 3.5, we proved that for any X^{PRO} produced by RC_2 under Normal distribution or RC_∞ under Cauchy distribution, all PRO solutions are GPRO over X^{PRO}. For other situations, this conclusion cannot be guaranteed, which means there may exist non-GPRO solutions in X^{PRO}. In order to obtain all GPRO solutions in X^{PRO}, we develop an efficient filtering algorithm as shown in Algorithm 2.

Algorithm 2. Filtering algorithm for obtaining GPRO solutions from X^{PRO}

Input: A set of PRO solutions $X^{PRO} = \{\mathbf{x}_i, 1 \leq i \leq N\}$ and its corresponding parameter set $P = \{\Delta_i, 1 \leq i \leq N\}$.
Output: A set of GPRO solutions X^{GPRO}.

```
Sort solutions in X^PRO according to Δ_1 < Δ_2 < ··· < Δ_N
Set X^GPRO = {x_1}, k = 1, j = 1
for j = 2 to N do
  if x_j is not dominated by x_k then
    Set X^GPRO = X^GPRO ∪ {x_j} and k = j
  end if
end for
Output X^GPRO
```

3.3. Elite Pareto Robust Optimality

The GPRO concept is defined with the general uncertainty set $\mathcal{U}$ which is controlled by parameter Δ. As shown in Section 2, different definitions of uncertainty set $\mathcal{U}$ lead to different robust counterparts. Each robust counterpart has a parameter which controls the size of the uncertainty set $\mathcal{U}$. So for each robust counterpart, we can get a corresponding GPRO solution set X^{GPRO}. Different robust counterparts may lead to different GPRO solution sets. In order to compare the GPRO solutions obtained by different robust counterparts, we further extend the GPRO concept, which is named as Elite Pareto Robust Optimality (EPRO).

Suppose we have a set which contains different robust counterparts $R = \{RC_j, 1 \leq j \leq M\}$, for each robust counterpart in R we can obtain a set of GPRO solutions X_j^{GPRO}, $1 \leq j \leq M$. The EPRO concept is defined as follows:

Definition 3.4 *(EPRO).* Given a robust counterpart set $R = \{RC_j, 1 \leq j \leq M\}$ and the corresponding GPRO solution set X_j^{GPRO}, $1 \leq j \leq M$. A GPRO solution $\mathbf{x} \in X_i^{GPRO}$ is called an EPRO solution over R, if there is no other GPRO solution $\bar{\mathrm{x}} \in \bigcup_{j=1}^{M} X_j^{GPRO}$ such that:

1. $\mathbf{c}^T\bar{\mathbf{x}} = \mathbf{c}^T\mathbf{x}$ and $PCV(\bar{\mathbf{x}}) < PCV(\mathbf{x})$
2. $\mathbf{c}^T\bar{\mathbf{x}} > \mathbf{c}^T\mathbf{x}$ and $PCV(\bar{\mathbf{x}}) \leq PCV(\mathbf{x})$

In the definition above, we say that $\bar{\mathbf{x}}$ Pareto dominates $\mathbf{x}$. The GPRO is defined within a specific robust counterpart, and the EPRO considers Pareto efficiency over a group of different robust counterparts.

Next a toy example is presented to illustrate the notion of the EPRO.

Example 3.3. Consider the following linear programming:

$$\text{maximize} \quad x_1 + x_2 \tag{33a}$$

$$\text{subject to} \quad a_1x_1 + a_2x_2 \leq 2.5 \tag{33b}$$

$$2x_1 + x_2 \leq 4 \tag{33c}$$

$$x_1 + 2x_2 \leq 4 \tag{33d}$$

$$x_1, x_2 \geq 0 \tag{33e}$$

Suppose $\tilde{a}_1$, $\tilde{a}_2$ obey uniform distributions and $a_1 = 1$, $a_2 = 1.01$, $\hat{a}_1 = \hat{a}_2 = 0.25$. Denote the primitive uncertainties as ζ_1, ζ_2 which are uniformly distributed on $[-1, 1]$.

We define robust counterpart set $R = \{RC_{1\cap\infty}, RC_\infty\}$. Each robust counterpart in R is used to solve the above linear programming problem. All models are solved by CPLEX 12.6 and the PCV of the solutions under uniform distribution is calculated according to the formulation in Section 4.

For $RC_{1\cap\infty}$, $\Gamma \in [0, 2]$, we define parameter set $P_1 = \{\Gamma = 0, 0.2, 0.4, \ldots, 2\}$. The results are shown in Fig. 1(a) where OV represents Objective Value and PCV represents Probability of Constraint Violation. We checked that all robustly optimal solutions obtained by $RC_{1\cap\infty}$ are unique under their corresponding uncertainty sets, so all RO solutions are PRO, and it is clear that all PRO solutions are GPRO.

For RC_∞, $\Theta \in [0, 1]$, we define parameter set $P_2 = \{\Theta = 0, 0.1, 0.2, \ldots, 1\}$. The results are shown in Fig. 1(b). We checked that all robustly optimal solutions obtained by RC_∞ are unique under their corresponding uncertainty sets, so all RO solutions are PRO, and it is clear that all PRO solutions are GPRO.

For two GPRO solution sets obtained by $RC_{1\cap\infty}$ and RC_∞, Fig. 1(c) shows the comparison between them, it is obvious that some GPRO solutions of RC_∞ are Pareto dominated by GPRO solutions of $RC_{1\cap\infty}$. In fact, all GPRO solutions of $RC_{1\cap\infty}$ are EPRO over R, 6 out of 11 GPRO solutions of RC_∞ are EPRO over R.

Remark. From the above example we have the following observations:

1. Robust counterpart set R which contains $RC_{1\cap\infty}$ and RC_∞ is applied on this example. The GPRO solution sets obtained by two robust counterparts are compared, which shows that some GPRO solutions may not be EPRO solutions over R. 5 GPRO solutions of RC_∞ are Pareto dominated by GPRO solutions of $RC_{1\cap\infty}$.
2. The parameter sets P_1 and P_2 determine the corresponding GPRO solution sets of two robust counterparts. In the above example the cardinality of P_1 and P_2 is 11 which is relatively sparse. In order to have a deeper comparison between two robust counterparts, the parameter sets need to be denser. We choose sparse parameter sets in this example just for illustration purpose.
3. In practical applications, the EPRO solution is always preferred. An ideal robust model will only produce EPRO solutions. In the above example we know that $RC_{1\cap\infty}$ performs better than RC_∞.

Algorithm 3. Filtering algorithm for obtaining EPRO solutions from X^{GPRO} over R

Input: Robust counterpart set $R = \{RC_j, 1 \leq j \leq M\}$, GPRO solution set X_j^{GPRO}, $1 \leq j \leq M$.
Output: A set of EPRO solutions X_i^{EPRO}

```
Sort each X_j^GPRO according to the increasing order of parameters of RC_j,
    1 ≤ j ≤ M
Set X_i^EPRO = ∅, index_j = 1, 1 ≤ j ≤ M
for all x ∈ X_i^GPRO do
  Set isEPRO = True
  for j = 1 to M and j ≠ i do
    isDominated = check(x, X_j^GPRO, index_j) /*See Algorithm 4*/
    if isDominated == True then
      Set isEPRO = False and break
    end if
  end for
  if isEPRO == True then
    Set X_i^EPRO = X_i^EPRO ∪ {x}
  end if
end for
Output X_i^EPRO
```

Algorithm 4. $check(\mathbf{x}, X_j^{GPRO}, index_j)$

Input: A GPRO solution $\mathbf{x}$ of RC_i, a GPRO solution set X_j^{GPRO} of RC_j
Output: Is solution $\mathbf{x}$ dominated by any solution in X_j^{GPRO}

```
for k = index_j to |X_j^GPRO| do
  Fetch the kth solution in X_j^GPRO, denote as x̄_k
  if c^T x̄_k ≥ c^T x then
    if x̄_k Pareto dominates x then
      Set index_j = k, isDominated = True and break
    end if
  else
    Set index_j = k, isDominated = False and break
  end if
end for
Return isDominated
```

Similar to the GPRO concept, we can prove that specific robust counterpart under specific probability distribution will only produce EPRO solutions. Based on Theorem 3.3 and 3.5, it is easy to get the following two corollaries:

Corollary 1. *If each element of the primitive uncertainties ζ_i obeys standard Normal distribution, then all GPRO solutions produced by RC_2(11) are EPRO over any robust counterpart set R where $RC_2 \in R$.*

Proof. Follow the proof way of Theorem 3.3, it is easy to propose a hypothesis and then get a contradiction. □

Corollary 2. *If each element of the primitive uncertainties ζ_i obeys standard Cauchy distribution, then all GPRO solutions produced by RC_∞(7) are EPRO over any robust counterpart set R where $RC_\infty \in R$.*

Proof. Follow the proof way of Theorem 3.5, it is easy to propose a hypothesis and then get a contradiction. □

Corollaries 1 and 2 show that: for any X^{GPRO} produced by RC_2 under Normal distribution or RC_∞ under Cauchy distribution, all GPRO solutions are EPRO over any robust counterpart set R. For other cases, this conclusion cannot be guaranteed, which means there may exist non-EPRO solutions in X^{GPRO}. In order to obtain all EPRO solutions in X^{GPRO}, we develop an efficient filtering algorithm which is described in Algorithms 3 and 4.

3.4. Computational consideration

For a robust counterpart RC with uncertainty set $\mathcal{U}$ which is controlled by parameter $\Delta \in [\Delta_l, \Delta_u]$ where Δ_l and Δ_u are the lower and upper bounds respectively, to find the GPRO or EPRO solutions, first we need to define a finite parameter set $P = \{\Delta_i \in [\Delta_l, \Delta_u], 1 \leq i \leq N\}$ and solve the RC N times under different sized uncertainty sets. With a large set P, this may lead to significant computational efforts for addressing realistic problems. In order to reduce the computational effort, we provide the following two strategies:

Strategy 1: Sparse discretization. This strategy means that we discretize interval $[\Delta_l, \Delta_u]$ sparsely. Then the obtained parameter set P will have a small number of elements and the computational effort can be decreased. But the drawback of this strategy is that with sparser parameter set P, we will get less accurate GPRO or EPRO solutions, and this may mislead the decision-maker to make a bad decision.

Strategy 2: Local discretization. This strategy means that we only discretize interval $[\Delta_l, \Delta_u]$ locally. For example we choose a sub-interval $[\Delta_l', \Delta_u'] \subset [\Delta_l, \Delta_u]$ and discretize interval $[\Delta_l', \Delta_u']$ to generate parameter set P. This will decrease the computational effort as we only focus on the discretization of a sub-interval. The

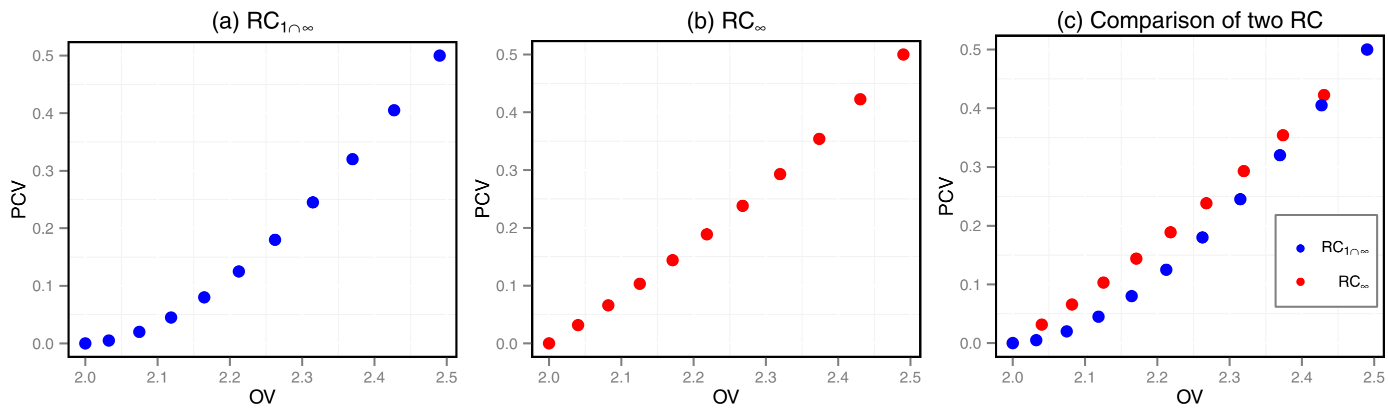

Fig. 1. Numerical results of Example 3.3.

drawback of this strategy is that we can only generate local GPRO or EPRO solutions.

For realistic problems which need large computational efforts, we can apply the above two strategies to reduce the computational efforts. For the situation when we need an overall understanding of the PRO solutions from the most conservative one to the most risky one, it is reasonable to choose Strategy 1. For the situation when we only need the PRO solutions around a specific PCV value, Strategy 2 is a better choice.

4. Probability of Constraint Violation

The Comprehensive Pareto Efficiency concepts introduced in this paper are based on PCV. The calculation of PCV needs the support of probability distributions. In this section, formulations of PCV are derived under different probability distributions. We also present an alternative way for calculating PCV when the formulation of PCV cannot be established.

Generally, we can reformulate the PCV as follows:

$$\begin{aligned} PCV(\mathbf{x}) &= Pr\left(\mathrm{a}_i^{\mathrm{T}}\mathbf{x} > b_i\right) \\ &= Pr\left(\zeta^{\mathrm{T}}\mathrm{A}_i\mathbf{x} > b_i - \mathbf{a}_i^{\mathrm{T}}\mathbf{x}\right) \end{aligned} \tag{34}$$

Denote random variable $Z = \zeta^{\mathrm{T}}\mathrm{A}_i\mathbf{x}$ and its CDF as $F_Z(z)$, then we can formulate the PCV as:

$$PCV(\mathbf{x}) = 1 - F_Z(b_i - \mathbf{a}_i^{\mathrm{T}}\mathbf{x}) \tag{35}$$

Next five concrete probability distributions are considered: Uniform, Triangular, Rademacher, Normal and Cauchy distributions. The formulations of PCV under these probability distributions are derived.

4.1. Uniform distribution

Proposition 3.6. *Suppose each element of the primitive uncertainties ζ_i obeys Uniform distribution on $[-1, 1]$, then the PCV can be calculated as:*

$$PCV(\mathbf{x}) = 1 - \frac{\sum_{\epsilon\in\{0,\pm1\}^N}\left(b_i - \mathbf{a}_i^{\mathrm{T}}\mathbf{x} + \epsilon^{\mathrm{T}}\mathrm{A}_i\mathbf{x}\right)_+^l \prod_{\sigma_j=1}\epsilon_j}{l!2^l\prod_{\sigma_j=1}\hat{a}_{ij}x_j} \tag{36}$$

where $l = \mathbf{e}^{\mathrm{T}}\sigma$, $\mathbf{e}$ is a N-dimensional vector with all elements equal 1, σ is a N-dimensional vector with each element $\sigma_j = \begin{cases} 1 & \text{if} x_j > 0 \\ 0 & \text{if} x_j = 0 \end{cases}$, $y_+ = \max(y, 0)$, ϵ is a N-dimensional vector with each element $\epsilon_j = \pm\sigma_j$.

Proof. Because each element of the primitive uncertainties ζ_i obeys Uniform distribution on $[-1, 1]$, then $Z = \zeta^{\mathrm{T}}\mathrm{A}_i\mathbf{x}$ is the sum of N independent and non-identically distributed uniform random variables.

Based on Corollary 1 in Bradley and Gupta (2002), we can formulate the probability density function (PDF) of Z as:

$$f_Z(z) = \frac{\sum_{\epsilon\in\{0,\pm1\}^N}\left(z + \epsilon^{\mathrm{T}}\mathrm{A}_i\mathbf{x}\right)_+^{l-1}\prod_{\sigma_j=1}\epsilon_j}{(l-1)!2^l\prod_{\sigma_j=1}\hat{a}_{ij}x_j} \tag{37}$$

where $l = \mathbf{e}^{\mathrm{T}}\sigma$, $\mathbf{e}$ is a N-dimensional vector with all elements equal 1, σ is a N-dimensional vector with each element $\sigma_j = \begin{cases} 1 & \text{if} x_j > 0 \\ 0 & \text{if} x_j = 0 \end{cases}$, $y_+ = \max(y,\ 0)$, ϵ is a N-dimensional vector with each element $\epsilon_j = \pm\sigma_j$.

From the PDF we can easily get the CDF of Z:

$$F_Z(z) = \frac{\sum_{\epsilon\in\{0,\pm1\}^N}\left(z + \epsilon^{\mathrm{T}}\mathrm{A}_i\mathbf{x}\right)_+^{l}\prod_{\sigma_j=1}\epsilon_j}{l!2^l\prod_{\sigma_j=1}\hat{a}_{ij}x_j} \tag{38}$$

Then the PCV can be formulated as:

$$\begin{aligned} PCV(\mathbf{x}) &= 1 - F_Z(b_i - \mathbf{a}_i^{\mathrm{T}}\mathbf{x}) \\ &= 1 - \frac{\sum_{\epsilon\in\{0,\pm1\}^N}\left(b_i - \mathbf{a}_i^{\mathrm{T}}\mathbf{x} + \epsilon^{\mathrm{T}}\mathrm{A}_i\mathbf{x}\right)_+^l\prod_{\sigma_j=1}\epsilon_j}{l!2^l\prod_{\sigma_j=1}\hat{a}_{ij}x_j} \end{aligned} \tag{39}$$

□

4.2. Triangular distribution

Proposition 3.7. *Suppose each element of the primitive uncertainties ζ_i obeys Triangular distribution on $[-1, 1]$, then the PCV can be calculated as:*

$$PCV(\mathbf{x}) = 1 - \frac{\sum_{\epsilon\in\{0,\pm1\}^{2N}}\left(b_i - \mathbf{a}_i^{\mathrm{T}}\mathbf{x} + \Sigma(\epsilon)\right)_+^l\prod_{|\epsilon_j|=1}\epsilon_j}{l!2^l\prod_{\sigma_j=1}(\hat{a}_{ij}x_j/2)^2} \tag{40}$$

where $l = 2\mathbf{e}^{\mathrm{T}}\sigma$, $\mathbf{e}$ is a N-dimensional vector with all elements equal 1, σ is a N-dimensional vector with each element $\sigma_j = \begin{cases} 1 & \text{if} x_j > 0 \\ 0 & \text{if} x_j = 0 \end{cases}$, $y_+ = \max(y, 0)$, $\Sigma(\epsilon) = \epsilon^{\mathrm{T}}\begin{bmatrix}\mathrm{A}_i/2 & 0 \\ 0 & \mathrm{A}_i/2\end{bmatrix}\begin{bmatrix}\mathbf{x} \\ \mathbf{x}\end{bmatrix}$ where $\epsilon = \begin{bmatrix}\epsilon^1 \\ \epsilon^2\end{bmatrix}$, $\epsilon_j^1 = \pm\sigma_j$ and $\epsilon_j^2 = \pm\sigma_j$.

Proof. Because each element of the primitive uncertainties ζ_i obeys Triangular distribution on $[-1, 1]$, then $Z = \zeta^{\mathrm{T}}\mathrm{A}_i\mathbf{x}$ is the sum of N independent and non-identically distributed triangular random variables.

As a triangular random variable $X \sim \text{Triangular}(-1, 1)$ can be decomposed as the sum of two uniform random variables $X_1 + X_2$ where $X_1, X_2 \sim \text{Uniform}(-1/2, 1/2)$. So Z is the sum of $2N$ independent and non-identically distributed uniform random variables.

Based on the CDF obtained in (38) we can formulate the CDF of Z as:

$$F_Z(z) = \frac{\sum_{\epsilon \in \{0,\pm 1\}^{2N}} (z + \Sigma(\epsilon))_+^l \prod_{|\epsilon_j|=1} \epsilon_j}{l! 2^l \prod_{\sigma_j=1} (\hat{a}_{ij} x_j/2)^2} \tag{41}$$

where $l = 2\mathbf{e}^T\sigma$, $\mathbf{e}$ is a N-dimensional vector with all elements equal 1, σ is a N-dimensional vector with each element $\sigma_j = \begin{cases} 1 & \text{if } x_j > 0 \\ 0 & \text{if } x_j = 0 \end{cases}$, $y_+ = \max(y, 0)$, $\Sigma(\epsilon) = \epsilon^T \begin{bmatrix} \mathrm{A}_i/2 & 0 \\ 0 & \mathrm{A}_i/2 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{x} \end{bmatrix}$ where $\epsilon = \begin{bmatrix} \epsilon^1 \\ \epsilon^2 \end{bmatrix}$, $\epsilon_j^1 = \pm\sigma_j$ and $\epsilon_j^2 = \pm\sigma_j$.

Then the PCV can be formulated as:

$$\begin{aligned} PCV(\mathbf{x}) &= 1 - F_Z(b_i - \mathbf{a}_i^T\mathbf{x}) \\ &= 1 - \frac{\sum_{\epsilon \in \{0,\pm 1\}^{2N}} \left(b_i - \mathbf{a}_i^T\mathbf{x} + \Sigma(\epsilon)\right)_+^l \prod_{|\epsilon_j|=1} \epsilon_j}{l! 2^l \prod_{\sigma_j=1} (\hat{a}_{ij} x_j/2)^2} \end{aligned} \tag{42}$$

□

4.3. Rademacher distribution

Proposition 3.8. *Suppose each element of the primitive uncertainties ζ_i obeys Rademacher distribution, then the PCV can be calculated as:*

$$PCV(\mathbf{x}) = 1 - \frac{1}{2^l} \sum_{\epsilon \in \{0,\pm 1\}^N} \Lambda\left(b_i - \mathbf{a}_i^T\mathbf{x}, \epsilon\right) \tag{43}$$

where $l = \mathbf{e}^T\sigma$, $\mathbf{e}$ is a N-dimensional vector with all elements equal 1, σ is a N-dimensional vector with each element $\sigma_j = \begin{cases} 1 & \text{if } x_j > 0 \\ 0 & \text{if } x_j = 0 \end{cases}$, $\Lambda(z, \epsilon) = \begin{cases} 1 & \text{if } z \geq \epsilon^T \mathrm{A}_i \mathbf{x} \\ 0 & \text{if } z < \epsilon^T \mathrm{A}_i \mathbf{x} \end{cases}$, $\epsilon_j = \pm\sigma_j$.

Proof. Because each element of the primitive uncertainties ζ_i obeys Rademacher distribution, the probability mass function (PMF) of ζ_i is:

$$f(\zeta_i) = \begin{cases} 1/2 & \text{if } \zeta_i = +1 \\ 1/2 & \text{if } \zeta_i = -1 \\ 0 & \text{otherwise} \end{cases} \tag{44}$$

Then $Z = \zeta^T \mathrm{A}_i \mathbf{x}$ is the sum of N independent and non-identically distributed two-point random variables.

The CDF of Z can be expressed as:

$$F_Z(z) = \frac{1}{2^l} \sum_{\epsilon \in \{0,\pm 1\}^N} \Lambda(z, \epsilon) \tag{45}$$

where $l = \mathbf{e}^T\sigma$, $\mathbf{e}$ is a N-dimensional vector with all elements equal 1, σ is a N-dimensional vector with each element $\sigma_j = \begin{cases} 1 & \text{if } x_j > 0 \\ 0 & \text{if } x_j = 0 \end{cases}$, $\Lambda(z, \epsilon) = \begin{cases} 1 & \text{if } z \geq \epsilon^T \mathrm{A}_i \mathbf{x} \\ 0 & \text{if } z < \epsilon^T \mathrm{A}_i \mathbf{x} \end{cases}$, $\epsilon_j = \pm\sigma_j$.

Then the PCV is:

$$\begin{aligned} PCV(\mathbf{x}) &= 1 - F_Z(b_i - \mathbf{a}_i^T\mathbf{x}) \\ &= 1 - \frac{1}{2^l} \sum_{\epsilon \in \{0,\pm 1\}^N} \Lambda\left(b_i - \mathbf{a}_i^T\mathbf{x}, \epsilon\right) \end{aligned} \tag{46}$$

□

4.4. Normal distribution

Proposition 3.9. *Suppose each element of the primitive uncertainties ζ_i obeys standard Normal distribution, then the PCV can be calculated as:*

$$PCV(\mathbf{x}) = 1 - \Phi\left(\frac{b_i - \mathbf{a}_i^T\mathbf{x}}{\sqrt{\mathbf{x}^T \mathrm{A}_i^2 \mathbf{x}}}\right) \tag{47}$$

where $\Phi(x)$ is the CDF of the standard Normal distribution.

Proof. If each element of the primitive uncertainties ζ_i obeys standard Normal distribution, according to the following two properties of Normal distribution:

1. If $X \sim Normal(\mu, \sigma^2)$, then $aX + b \sim Normal(a\mu + b, a^2\sigma^2)$
2. If $X \sim Normal(\mu_1, \sigma_1^2)$ and $Y \sim Normal(\mu_2, \sigma_2^2)$ are independent, then $X + Y \sim Normal(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

We have $Z = \zeta^T \mathrm{A}_i \mathbf{x}$ obeys Normal distribution $Normal(0, \mathbf{x}^T \mathrm{A}_i^2 \mathbf{x})$.

Follow the definition of PCV and the property of Normal distribution we have:

$$\begin{aligned} PCV(\mathbf{x}) &= Pr\left(\zeta^T \mathrm{A}_i \mathbf{x} > b_i - \mathbf{a}_i^T\mathbf{x}\right) \\ &= Pr\left(\frac{\zeta^T \mathrm{A}_i \mathbf{x}}{\sqrt{\mathbf{x}^T \mathrm{A}_i^2 \mathbf{x}}} > \frac{b_i - \mathbf{a}_i^T\mathbf{x}}{\sqrt{\mathbf{x}^T \mathrm{A}_i^2 \mathbf{x}}}\right) \\ &= 1 - \Phi\left(\frac{b_i - \mathbf{a}_i^T\mathbf{x}}{\sqrt{\mathbf{x}^T \mathrm{A}_i^2 \mathbf{x}}}\right) \end{aligned} \tag{48}$$

where $\Phi(x)$ is the CDF of the standard Normal distribution. □

4.5. Cauchy distribution

Proposition 3.10. *Suppose each element of the primitive uncertainties ζ_i obeys standard Cauchy distribution, then the PCV can be calculated as:*

$$PCV(\mathbf{x}) = \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{b_i - \mathbf{a}_i^T\mathbf{x}}{\mathrm{a}_i^T\mathbf{x}}\right) \tag{49}$$

Proof. If each element of the primitive uncertainties ζ_i obeys standard Cauchy distribution, according to the following two properties of Cauchy distribution:

1. If $X \sim Cauchy(x_0, \gamma)$, then $kX + l \sim Cauchy(x_0 k + l, \gamma|k|)$
2. If $X \sim Cauchy(x_0, \gamma_0)$ and $Y \sim Cauchy(x_1, \gamma_1)$ are independent, then $X + Y \sim Cauchy(x_0 + x_1, \gamma_0 + \gamma_1)$

We have $Z = \zeta^T \mathrm{A}_i \mathbf{x}$ obeys $Cauchy(0, \mathrm{a}_i^T\mathbf{x})$.

Then the CDF of Z is:

$$F_Z(z) = \frac{1}{\pi} \arctan\left(\frac{z}{\mathrm{a}_i^T\mathbf{x}}\right) + \frac{1}{2} \tag{50}$$

Then we can get the PCV as:

$$PCV(\mathbf{x}) = 1 - F_Z(b_i - \mathbf{a}_i^T\mathbf{x}) = \frac{1}{2} - \frac{1}{\pi}\arctan\left(\frac{b_i - \mathbf{a}_i^T\mathbf{x}}{\mathrm{a}_i^T\mathbf{x}}\right) \tag{51}$$

□

4.6. Other distributions

For the probability distributions under which the PCV cannot be formulated explicitly, there is an alternative way for calculating PCV which is by means of FFT technique Ruckdeschel and Kohl (2010).

As random variable Z is the sum of N independent random variables, its PDF can be expressed as the convolution of the PDF of N random variables. Ruckdeschel and Kohl (2010) proposed an algorithm which is based on the discrete Fourier transformation (DFT) and its fast computability via the fast Fourier transformation (FFT). The R package "distr" implements the FFT algorithm, various probability distributions are supported. For details we refer to Ruckdeschel et al. (2015).

5. Numerical studies

5.1. Production planning problem

The first example is a linear programming problem called production planning problem which is introduced by Li et al. (2012) and further studied in Li and Floudas (2014), Li and Li (2015), Yuan et al. (2016), Guzman et al. (2016). In this problem, a company needs to make a production plan for the coming year which is divided into 6 periods. At the beginning, there is an initial stock of 500 tons of product in storage, and it is required to have the same amount of product in storage at the end of the year. The objective is thus to maximize the sales while some constraints need to be satisfied, including the total production and storage cost is within a budget, the inventory product balances, the final stored product requirement, the production capacity limitations and the product demand upper bounds.

Then the production planning problem can be formulated as follows:

$$\text{maximize} \quad \sum_j P_j z_j \tag{52a}$$

$$\text{subject to} \quad \sum_j C_j x_j + \sum_j V_j y_j \le 400,000 \tag{52b}$$

$$500 + x_1 - (y_1 + z_1) = 0 \tag{52c}$$

$$y_{j-1} + x_j - (y_j + z_j) = 0, \quad \forall j = 2, \ldots, 6 \tag{52d}$$

$$y_6 = 500 \tag{52e}$$

$$x_j \le U_j, \quad \forall j = 1, \ldots, 6 \tag{52f}$$

$$z_j \le D_j, \quad \forall j = 1, \ldots, 6 \tag{52g}$$

$$x_j, y_j, z_j \ge 0, \quad \forall j = 1, \ldots, 6 \tag{52h}$$

For detailed problem description and data we refer to Li et al. (2012).

5.1.1. Experiment settings

We have the following experiment settings for this problem:

1. In this problem, suppose the production cost $\tilde{C}_j$ is involved in uncertainty, $\tilde{C}_j = C_j + \hat{C}_j\zeta_j$ where C_j is the nominal value, $\hat{C}_j = 0.5C_j$ is the perturbation value and ζ_j is the primitive uncertainty variable.
2. We consider two situations of the primitive uncertainty variable ζ_j: bounded and unbounded. When ζ_j is bounded in $[-1, 1]$, the selection of the robust counterpart set is $R_1 = \{RC_\infty, RC_{1\cap\infty}, RC_{2\cap\infty}\}$. When ζ_j is unbounded, the selection of the robust counterpart set is $R_2 = \{RC_\infty, RC_1, RC_2\}$.
3. In the case when ζ_j is bounded, we consider four concrete probability distributions: Uniform, Triangular, Rademacher and Arcsine distributions. In the case when ζ_j is unbounded, two concrete probability distributions: Normal and Cauchy distributions are considered.
4. For the general robust counterpart RC with uncertainty set $\mathcal{U}$ which is controlled by parameter Δ, suppose we define $\Delta \in [\Delta_l, \Delta_u]$, the parameter set P for RC is defined as follows:

$$P = \left\{\Delta_i = \frac{(\Delta_u - \Delta_l)(i-1)}{N-1} + \Delta_l, 1 \le i \le N\right\} \tag{53}$$

and we set $N = 500$ in the experiment which means 500 runs for each robust counterpart.
5. If the primitive uncertainty variable ζ_j is bounded, we define $\Theta \in [0, 1]$, $\Gamma \in [0, 4.62]$ and $\Omega \in [0, 2.07]$ for RC_∞, $RC_{1\cap\infty}$ and $RC_{2\cap\infty}$ respectively, the reason for this setting is that RC_∞, $RC_{1\cap\infty}$ and $RC_{2\cap\infty}$ generate the most conservative solution with $\Theta = 1$, $\Gamma = 4.62$ and $\Omega = 2.07$ according to a pre-calculation. If the primitive uncertainty variable ζ_j is unbounded, we define $\Theta \in [0, 6]$, $\Gamma \in [0, 23.7]$ and $\Omega \in [0, 10.87]$ for RC_∞, RC_1 and RC_2 respectively, the reason for this setting is that RC_∞, RC_1 and RC_2 have close objective values with $\Theta = 6$, $\Gamma = 23.7$ and $\Omega = 10.87$ according to a pre-calculation.

5.1.2. Numerical analysis

First we consider the case when the primitive uncertainty variable ζ_j is bounded in $[-1, 1]$. The robust counterparts in R_1 are solved by CPLEX 12.6. PCV under Uniform, Triangular and Rademacher distributions are calculated by formulations established in Section 4 and the PCV under Arcsine distribution is calculated by R package "distr". The numerical results are shown in Fig. 2. For solutions with objective value less than 2,840,000, we checked that all solutions are unique optimal solutions under their corresponding uncertainty sets, which implies all of them are PRO solutions. In the case when the objective value of an optimal solution $\mathbf{x}$ is 2,840,000, we apply Algorithm 1 to obtain an approximate PRO solution $\bar{\mathbf{x}}$. We define a metric for performance improvement between $\mathbf{x}$ and $\bar{\mathbf{x}}$ as follows:

$$\frac{\log PCV(\bar{\mathbf{x}}) - \log PCV(\mathbf{x})}{\log PCV(\bar{\mathbf{x}})} \tag{54}$$

Table A1 gives the mean and maximum performance improvement of each RC under different distributions. From the results we can see that the approximate PRO solutions obtained by Algorithm 1 improved the performance compare with the original optimal solutions.

Next for each RC in R_1, we obtain the set of GPRO solutions X^{GPRO} from PRO solution set X^{PRO} by Algorithm 2, then calculate $|X^{GPRO}|/|X^{PRO}|$ which is the ratio of the GPRO solution number to the PRO solution number, this index can reflect the ability of a RC to generate GPRO solutions. The results are shown in Table A2. From the results we know that RC_∞ and $RC_{2\cap\infty}$ perform better than $RC_{1\cap\infty}$ on generating GPRO solutions under Uniform, Triangular and Arcsine distributions, while three RC perform nearly under Rademacher distribution.

Lastly for each RC in R_1, we obtain the set of EPRO solutions X^{EPRO} over robust counterpart set R_1 by Algorithm 3. Then calculate $|X^{EPRO}|/|X^{PRO}|$ which is the ratio of the EPRO solution number to the PRO solution number, this index reflects the ability of a RC

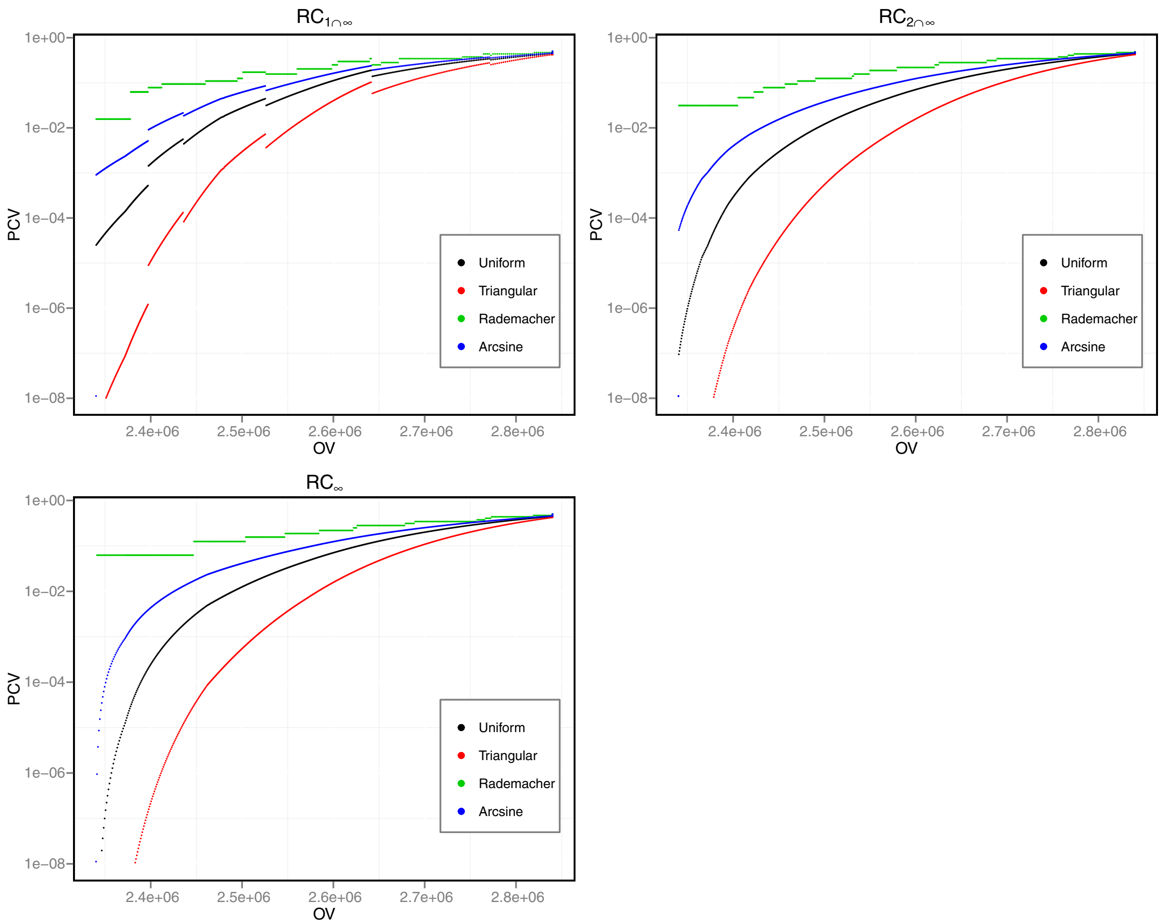

Fig. 2. Numerical results of 3 RC under different bounded distributions (production planning problem).

to generate EPRO solutions over R_1. Fig. 3 shows an intuitionistic comparison of 3 RC under different distributions and the results are shown in Table A3. From the results we know that RC_∞ and $RC_{2\cap\infty}$ perform better than $RC_{1\cap\infty}$ on generating EPRO solutions under Uniform, Triangular and Arcsine distributions, while three RC perform nearly under Rademacher distribution.

Now we consider the case when the primitive uncertainty variable ζ_j is unbounded. The numerical results are shown in Fig. 4. We checked that all optimal solutions with objective value less than 2,840,000 are PRO solutions. For each RC in R_2, we obtain the set of GPRO solutions X^{GPRO} from PRO solution set X^{PRO} and calculate $|X^{GPRO}|/|X^{PRO}|$. The results are shown in Table A4. From the results we know that RC_∞ and RC_2 perform better than RC_1 on generating GPRO solutions under two unbounded distributions.

Then for each RC, we obtain the set of EPRO solutions X^{EPRO} over robust counterpart set R_2 and calculate $|X^{EPRO}|/|X^{PRO}|$. Fig. 5 shows an intuitionistic comparison of 3 RC under two distributions and the results are shown in Table A5. From Fig. 5 and Table A5 it is clear to observe that: RC_2 overwhelms RC_∞ and RC_1 under Normal distribution and RC_∞ overwhelms RC_1 and RC_2 under Cauchy distribution, with all PRO solutions are EPRO. This further confirms the theorems and corollaries we obtained in Section 3.

5.2. Orienteering problem

The second example is a 0–1 integer programming problem called orienteering problem (OP) which is a variation of Vehicle Routing Problem (VRP). It is firstly introduced by Golden et al. (1987) and has been studied in planning, scheduling and supply chain areas. For a survey of OP we refer to Vansteenwegen et al. (2011), Gunawan et al. (2016). In this problem, a set of vertices is given, each with a score. The goal is to plan a path with limited length, that visits some vertices and maximizes the sum of the collected scores. Some constraints need to be satisfied including the path starts and ends at the depot, each vertex is visited at most one time, the length of the path has a maximum limit and the path need to have connectivity.

Then the orienteering problem can be formulated as follows:

$$\text{maximize} \quad \sum_{i\in N} s_i \sum_{j\in N^+\setminus\{i\}} x_{ij} \tag{55a}$$

$$\text{subject to} \quad \sum_{(i,j)\in A} d_{ij}x_{ij} \le L \tag{55b}$$

$$\sum_{i\in N} x_{0i} = \sum_{i\in N} x_{i0} = 1 \tag{55c}$$

$$\sum_{i\in N^+\setminus\{j\}} x_{ij} = \sum_{i\in N^+\setminus\{j\}} x_{ji} \le 1, \quad \forall j \in N \tag{55d}$$

$$u_i - u_j + 1 \le (1 - x_{ij})|N|, \quad \forall i, j \in N \tag{55e}$$

$$1 \le u_i \le |N|, \quad \forall i \in N \tag{55f}$$

$$x_{ij} \in \{0, 1\}, \quad \forall (i, j) \in A \tag{55g}$$

where N is the set of vertices, $|N|$ is the cardinality of set N, 0 is the depot location where $0 \notin N$, $N^+ = N \cup \{0\}$, A is the set of arcs connecting vertices in N^+, i, j are the index of vertex, s_i is the score of vertex i, (i, j) is the arc between i and j where $(i, j) \in A$, d_{ij} is the length of

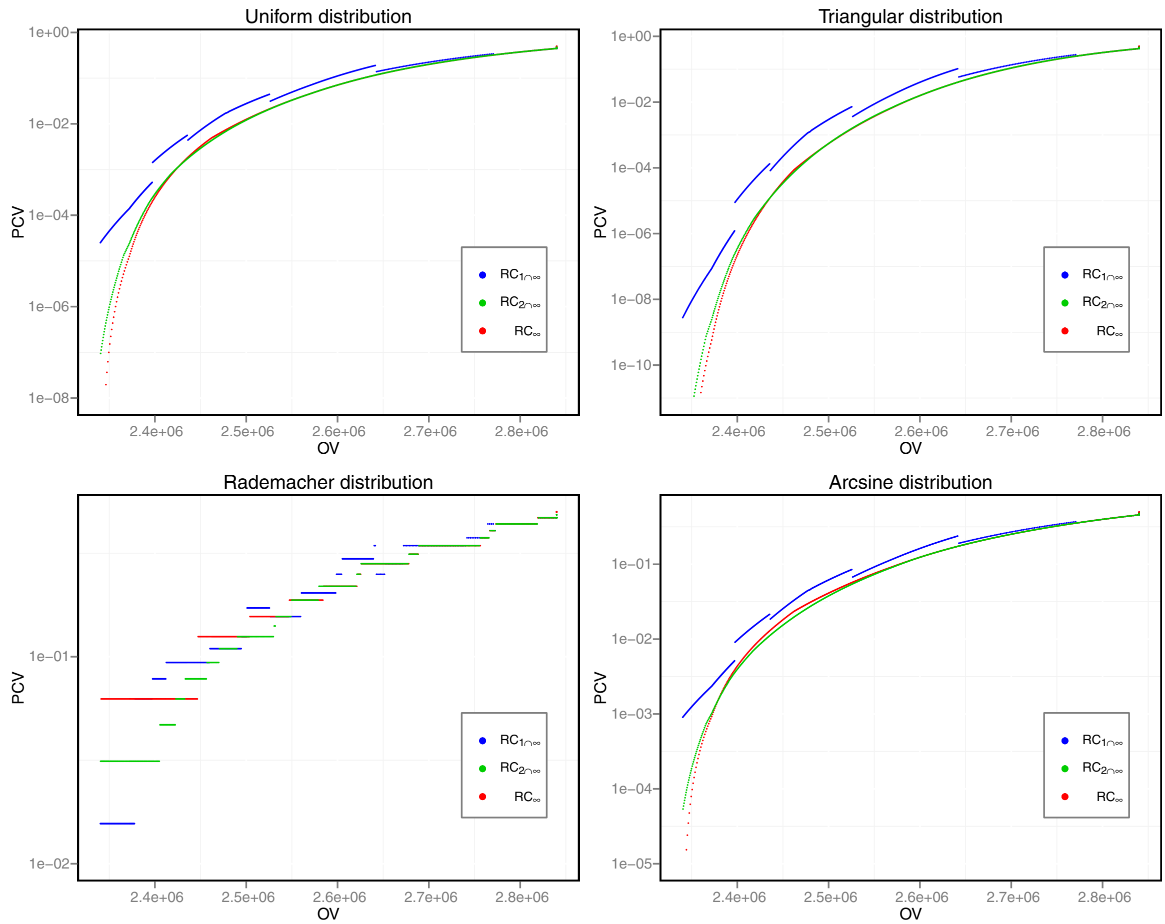

Fig. 3. Comparisons of 3 RC under different bounded distributions (production planning problem).

arc (i, j), L is the maximum length of a path, x_{ij} is the binary decision variable where $x_{ij} = 1$ if and only if arc (i, j) is visited by the path, u_i, u_j are auxiliary variables. The data is from data set "Tsiligirides problem 2" (Tsiligirides, 1984) which contains 20 nodes,[2] each with a score of either 10, 15, 20, 25, 30, 40 or 50. The maximum length L is 40.

5.2.1. Experiment settings

We have the following experiment settings for this problem:

1. In this problem, suppose the arc length $\tilde{d}_{ij}$ is involved in uncertainty, $\tilde{d}_{ij} = d_{ij} + \hat{d}_{ij}\zeta_{ij}$ where d_{ij} is the nominal value, $\hat{d}_{ij} = 0.2d_{ij}$ is the perturbation value and ζ_{ij} is the primitive uncertainty variable.
2. We consider two situations of the primitive uncertainty variable ζ_j: bounded and unbounded. When ζ_j is bounded in $[-1, 1]$, the selection of the robust counterpart set is $R_1 = \{RC_\infty, RC_{1\cap\infty}, RC_{2\cap\infty}\}$. When ζ_j is unbounded, the selection of the robust counterpart set is $R_2 = \{RC_\infty, RC_1, RC_2\}$.
3. In the case when ζ_j is bounded, we consider four concrete probability distributions: Uniform, Triangular, Rademacher and Arcsine distributions. In the case when ζ_j is unbounded, two concrete probability distributions: Normal and Cauchy distributions are considered.
4. The parameter set P is defined as in Formulation (53) and we set $N = 500$ in the experiment which means 500 runs for each robust counterpart.
5. If the primitive uncertainty variable ζ_j is bounded, we define $\Theta \in [0, 0.9]$, $\Gamma \in [0, 11.6]$ and $\Omega \in [0, 3.3]$ for RC_∞, $RC_{1\cap\infty}$ and $RC_{2\cap\infty}$ respectively, the reason for this setting is that RC_∞, $RC_{1\cap\infty}$ and $RC_{2\cap\infty}$ produce the most conservative solution with $\Theta = 0.9$, $\Gamma = 11.6$ and $\Omega = 3.3$ according to a pre-calculation. If the primitive uncertainty variable ζ_j is unbounded, we define $\Theta \in [0, 2.3]$, $\Gamma \in [0, 13.9]$ and $\Omega \in [0, 7.4]$ for RC_∞, RC_1 and RC_2 respectively, the reason for this setting is that RC_∞, RC_1 and RC_2 have the same objective values with $\Theta = 2.3$, $\Gamma = 13.9$ and $\Omega = 7.4$ according to a pre-calculation.

5.2.2. Numerical analysis

First we consider the case when the primitive uncertainty variable ζ_j is bounded in $[-1, 1]$. The robust counterparts in R_1 are solved by CPLEX 12.6. PCV under Uniform and Rademacher distributions are calculated by formulations established in Section 4 and the PCV under Triangular[3] and Arcsine distributions are calculated by R package "distr". The numerical results are shown in Fig. 6 where the small dots represent the robust optimal solutions obtained by RC and are jittered horizontally to reduce overlaps. In order to obtain the PRO solutions, the 3 RC are solved by CPLEX with the solution pool feature which generates and stores all optimal solutions, then the PCV of all optimal solutions are calculated

[2] The original data has two depots, we delete the second depot and only consider one depot in this paper. The original data set can be found with URL: http://www.mech.kuleuven.be/en/cib/op.

[3] In this problem, we cannot use the formulation of Triangular distribution to calculate PCV because the calculation is beyond the computing ability of a PC.

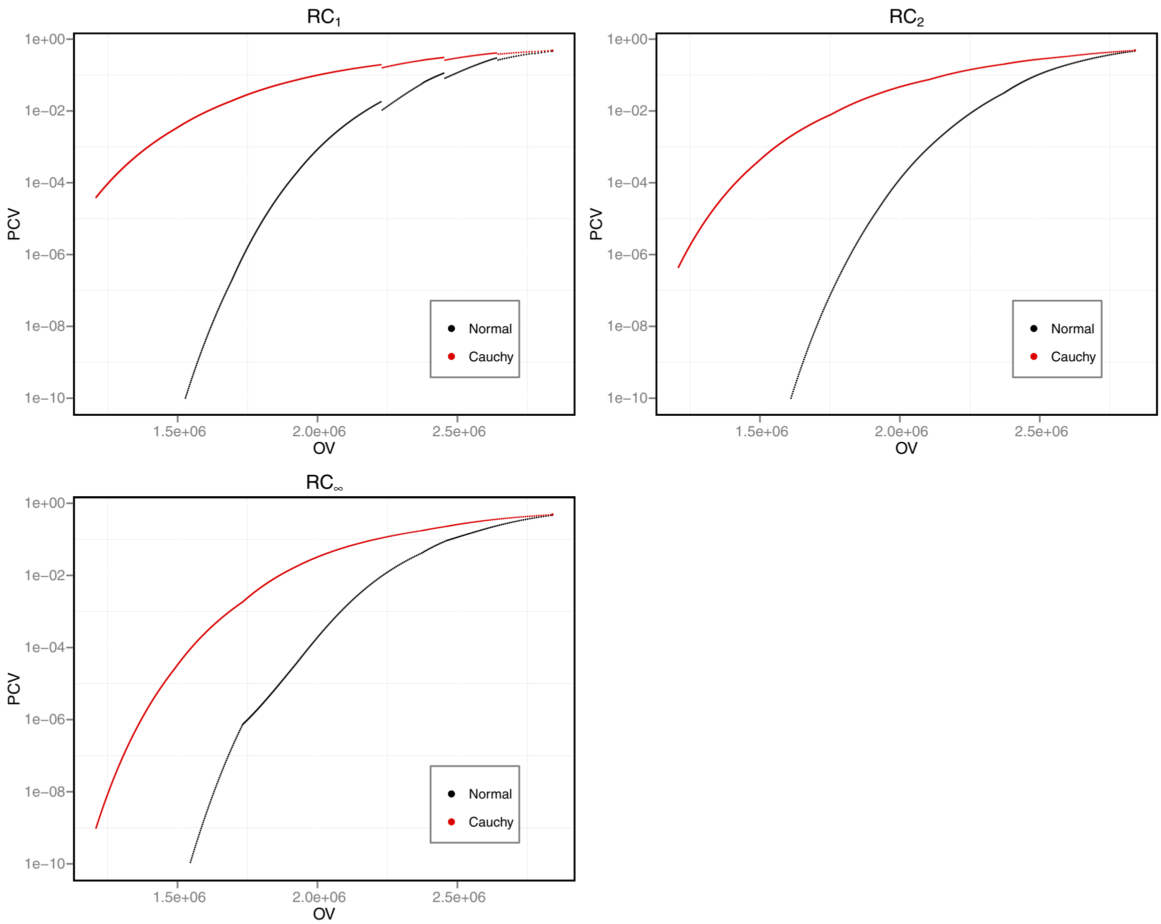

Fig. 4. Numerical results of 3 RC under Normal and Cauchy distributions (production planning problem).

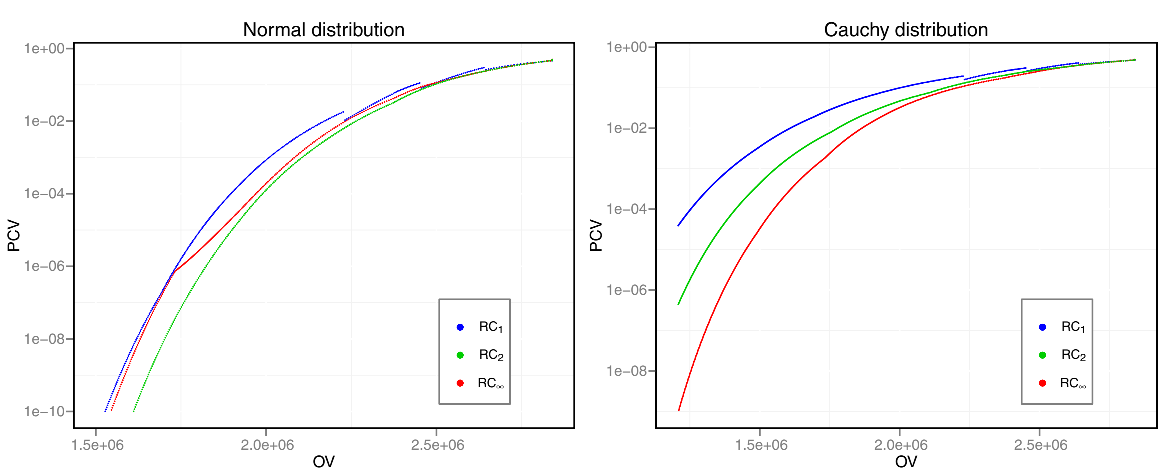

Fig. 5. Comparisons of 3 RC under Normal and Cauchy distributions (production planning problem).

and compared which identifies the PRO solutions. The big dots in Fig. 6 represent the PRO solutions found.

For all non-PRO solutions, performance improvement as defined in Formulation (54) are calculated and summarized. Table A6 gives the mean and maximum performance gain of each RC under different distributions. From the results we can see that PRO solutions have a significant performance gain compare with non-PRO solutions.

Next for each RC in R_1, we obtain the set of GPRO solutions X^{GPRO} from PRO solution set X^{PRO} and then obtain the set of EPRO solutions X^{EPRO} over R_1, and calculate $|X^{GPRO}|/|X^{PRO}|$ and $|X^{EPRO}|/|X^{PRO}|$ respectively. The results are shown in Tables A7 and A8. From the results we know that: For all RC in R_1, all PRO solutions obtained are GPRO and EPRO.

Now we consider the case when the primitive uncertainty variable ζ_j is unbounded. Fig. 7 shows the results. Table A9 gives the mean and maximum performance gain of each RC under different distributions. From the results we know that PRO solutions have a significant performance gain compare with non-PRO solutions.

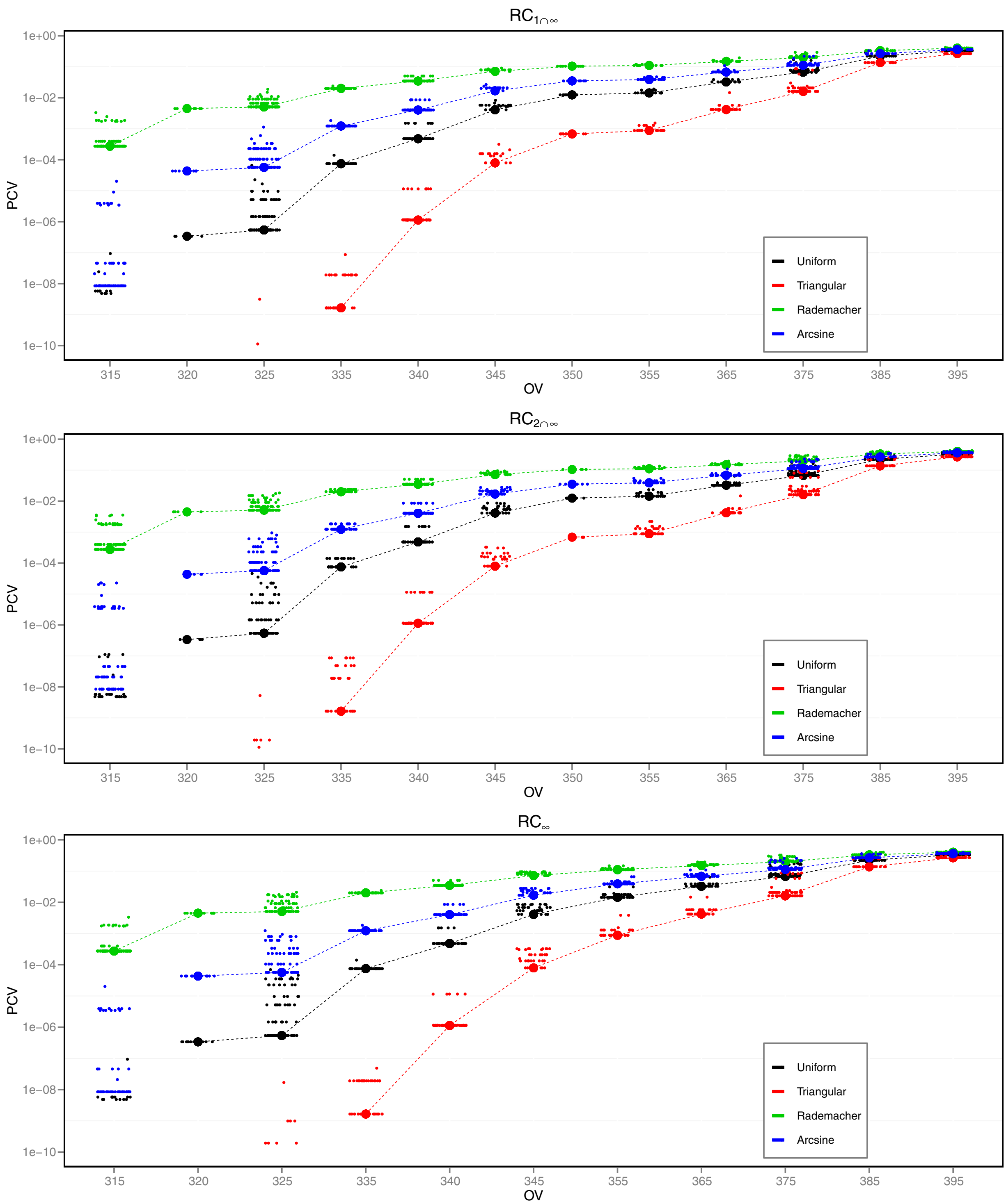

Fig. 6. Numerical results of 3 RC under different bounded distributions (orienteering problem).

Next for each RC in R_2, we obtain the set of GPRO solutions X^{GPRO} from PRO solution set X^{PRO} and then obtain the set of EPRO solutions X^{EPRO} over R_2, and calculate $|X^{GPRO}|/|X^{PRO}|$ and $|X^{EPRO}|/|X^{PRO}|$ respectively. The results are shown in Tables A10 and A11. From the results we know that RC_∞ performs the best over R_2 on generating GPRO solutions. RC_2 outperforms RC_∞ and RC_1 under Normal distribution and RC_∞ outperforms RC_1 and RC_2 under Cauchy distribution, with all PRO solutions are EPRO. This is consistent with the Theorems and Corollaries in Section 3.

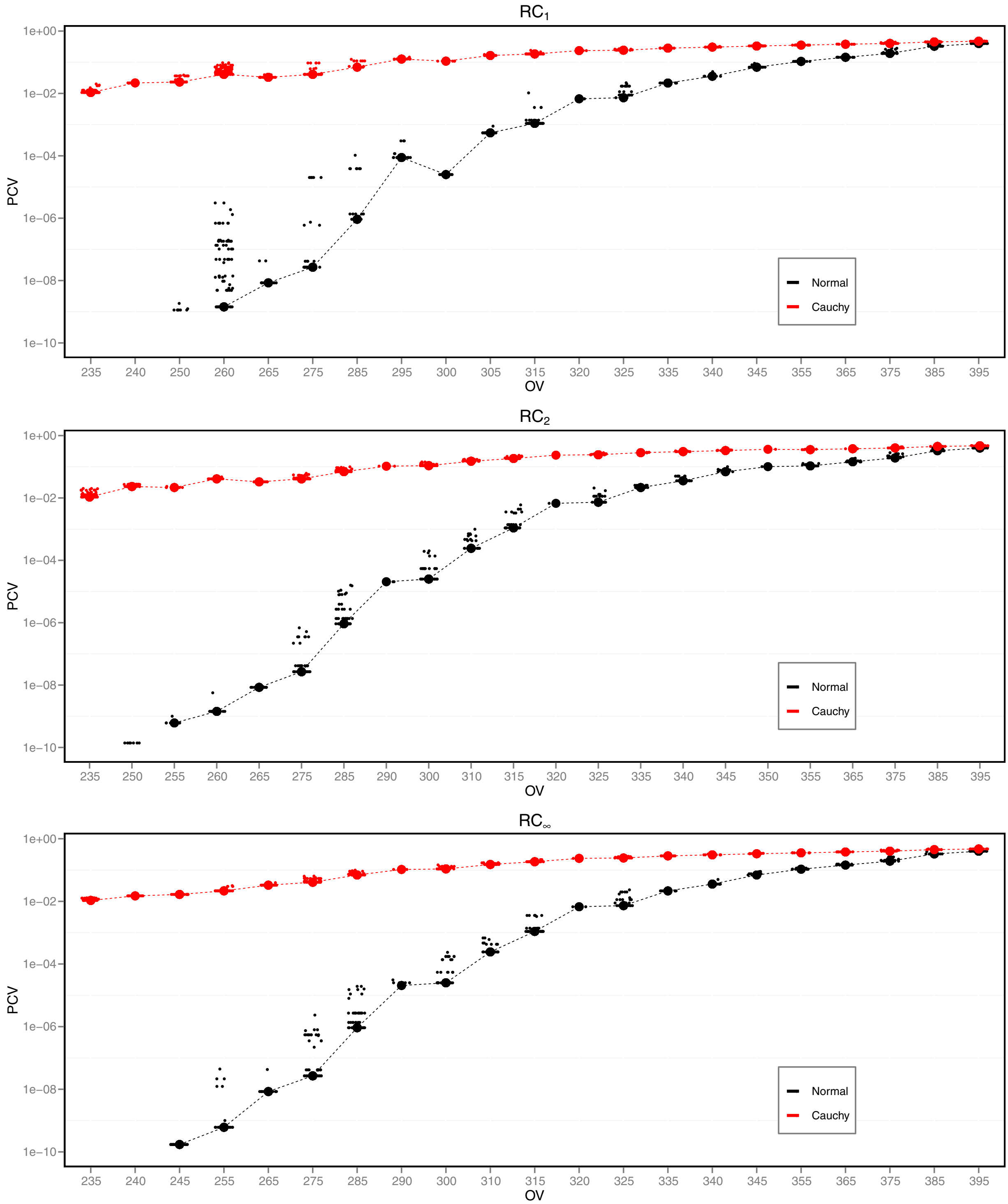

Fig. 7. Numerical results of 3 RC under Normal and Cauchy distributions (orienteering problem).

6. Conclusion

In this paper, we introduced the concept of Comprehensive Pareto Efficiency in robust counterpart optimization which includes Pareto Robust Optimality (PRO), Global Pareto Robust Optimality (GPRO) and Elite Pareto Robust Optimality (EPRO). We theoretically proved that RC_2 under Normal distribution and RC_∞ under Cauchy distribution achieve the best performance, with all PRO solutions generated are GPRO and EPRO. The numerical results of two applications show that PRO solutions can improve the performance compare with non-PRO solutions. By considering GPRO and EPRO solutions, we can have a deep understanding of the quality of the robust solutions and the difference between different robust counterparts. In general, the Comprehensive Pareto Efficiency provides a new perspective for robust counterpart optimization and has important significance in practice which can help us to find high quality solutions and make better decisions.

We only consider single uncertain constraint in this paper, it is interesting to consider multiple uncertain constraints in robust counterpart optimization, many researchers have worked on this which is known as joint chance constraint. Future work will be conducted to extend the current concept to the multiple uncertain constraints case.

Acknowledgements

This work was supported by National Natural Science Foundation of China (Nos. 61573277, 71471158), the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 15201414), the Fundamental Research Funds for the Central Universities, the Open Research Fund of the State Key Laboratory of Astronautic Dynamics under Grant 2015ADL-DW403, and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, Natural Science Basic Research Plan in Shaanxi Province of China (No. 2015JM6316). The authors also would like to thank The Hong Kong Polytechnic University Research Committee for financial and technical support. We are also thankful to the anonymous referees.

Appendix A. Detailed numerical results

Tables A1–A11

Table A1
Performance improvement obtained by the approximate PRO solutions (production planning problem).

RC	Uniform		Triangular		Rademacher		Arcsine	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max
$RC_{1\cap\infty}$	0.01	0.1397	0.0139	0.1942	0.0061	0.0852	0.0083	0.1156
$RC_{2\cap\infty}$	0.0348	0.0698	0.0492	0.0982	0.013	0.0433	0.0286	0.0574
RC_{∞}	0.0747	0.1397	0.1048	0.1942	0.0596	0.0852	0.0612	0.1156

Table A2
$|X^{GPRO}|/|X^{PRO}|$ of each RC under different bounded distributions (production planning problem).

RC	Uniform	Triangular	Rademacher	Arcsine
$RC_{1\cap\infty}$	0.87	0.87	0.03	0.87
$RC_{2\cap\infty}$	1.00	1.00	0.04	1.00
RC_{∞}	1.00	1.00	0.03	1.00

Table A3
$|X^{EPRO}|/|X^{PRO}|$ of each RC under different bounded distributions (production planning problem).

RC	Uniform	Triangular	Rademacher	Arcsine
$RC_{1\cap\infty}$	0.06	0.06	0.03	0.06
$RC_{2\cap\infty}$	0.76	0.62	0.04	0.89
RC_{∞}	0.75	0.89	0.03	0.52

Table A4
$|X^{GPRO}|/|X^{PRO}|$ of each RC under Normal and Cauchy distributions (production planning problem).

RC	Normal	Cauchy
RC_1	0.95	0.91
RC_2	1.00	1.00
RC_{∞}	1.00	1.00

Table A5
$|X^{EPRO}|/|X^{PRO}|$ of each RC under Normal and Cauchy distributions (production planning problem).

RC	Normal	Cauchy
RC_1	0.01	0.01
RC_2	1.00	0.08
RC_{∞}	0.08	1.00

Table A6
Performance improvement obtained by the PRO solutions under different bounded distributions (orienteering problem).

RC	Uniform		Triangular		Rademacher		Arcsine	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max
$RC_{1\cap\infty}$	0.0762	0.4575	0.0633	0.4075	0.0913	0.3042	0.0456	0.3066
$RC_{2\cap\infty}$	0.1065	0.4633	0.0807	0.4291	0.1051	0.3118	0.061	0.325
RC_{∞}	0.098	0.4575	0.0806	0.4977	0.1087	0.3042	0.0691	0.3783

Table A7
$|X^{GPRO}|/|X^{PRO}|$ of each RC under different bounded distributions (orienteering problem).

RC	Uniform	Triangular	Rademacher	Arcsine
$RC_{1\cap\infty}$	1.00	1.00	1.00	1.00
$RC_{2\cap\infty}$	1.00	1.00	1.00	1.00
RC_{∞}	1.00	1.00	1.00	1.00

Table A8
$|X^{EPRO}|/|X^{PRO}|$ of each RC under different bounded distributions (orienteering problem).

RC	Uniform	Triangular	Rademacher	Arcsine
$RC_{1\cap\infty}$	1.00	1.00	1.00	1.00
$RC_{2\cap\infty}$	1.00	1.00	1.00	1.00
RC_{∞}	1.00	1.00	1.00	1.00

Table A9
Performance improvement obtained by the PRO solutions under Normal and Cauchy distributions (orienteering problem).

RC	Normal		Cauchy	
	Mean	Max	Mean	Max
RC_1	0.0917	0.3799	0.0614	0.2672
RC_2	0.0643	0.2492	0.0395	0.1385
RC_{∞}	0.0664	0.2563	0.0384	0.1411

Table A10
$|X^{GPRO}|/|X^{PRO}|$ of each RC under Normal and Cauchy distributions (orienteering problem).

RC	Normal	Cauchy
RC_1	0.95	0.78
RC_2	1.00	0.89
RC_{∞}	1.00	1.00

Table A11
$|X^{EPRO}|/|X^{PRO}|$ of each RC under Normal and Cauchy distributions (orienteering problem).

RC	Normal	Cauchy
RC_1	0.92	0.66
RC_2	1.00	0.89
RC_{∞}	0.94	1.00

References

Ben-Tal, A., Nemirovski, A., 2002. Robust optimization-methodology and applications. Math. Program. 92, 453–480.
Ben-Tal, A., El Ghaoui, L., Nemirovski, A., 2009. Robust Optimization. Princeton University Press.
Ben-Tal, A., Den Hertog, D., De Waegenaere, A., Melenberg, B., Rennen, G., 2013. Robust solutions of optimization problems affected by uncertain probabilities. Manag. Sci. 59, 341–357.
Bertsimas, D., Brown, D.B., 2009. Constructing uncertainty sets for robust linear optimization. Oper. Res. 57, 1483–1495.
Bertsimas, D., Sim, M., 2004. The price of robustness. Oper. Res. 52, 35–53.
Bertsimas, D., Pachamanova, D., Sim, M., 2004. Robust linear optimization under general norms. Oper. Res. Lett. 32, 510–516.
Bertsimas, D., Brown, D.B., Caramanis, C., 2011. Theory and applications of robust optimization. SIAM Rev. 53, 464–501.

Bertsimas, D., Gupta, V., Kallus, N., 2013. Data-Driven Robust Optimization, arXiv:1401.0212.

Bradley, D.M., Gupta, R.C., 2002. On the distribution of the sum of n non-identically distributed uniform random variables. Ann. Inst. Stat. Math. 54, 689–700.

Calafiore, G.C., El Ghaoui, L., 2006. On distributionally robust chance-constrained linear programs. J. Optim. Theory Appl. 130, 1–22.

El Ghaoui, L., Oustry, F., Lebret, H., 1998. Robust solutions to uncertain semidefinite programs. SIAM J. Optim. 9, 33–52.

Gabrel, V., Murat, C., Thiele, A., 2014. Recent advances in robust optimization: an overview. Eur. J. Oper. Res. 235, 471–483.

Golden, B.L., Levy, L., Vohra, R., 1987. The orienteering problem. Naval Res. Logist. 34, 307–318.

Gorissen, B.L., Yanıkoğlu, İ., den Hertog, D., 2015. A practical guide to robust optimization. Omega 53, 124–137.

Gunawan, A., Lau, H.C., Vansteenwegen, P., 2016. Orienteering problem: a survey of recent variants, solution approaches and applications. Eur. J. Oper. Res.

Guzman, Y.A., Matthews, L.R., Floudas, C.A., 2016. New a priori and a posteriori probabilistic bounds for robust counterpart optimization: I. Unknown probability distributions. Comput. Chem. Eng. 84, 568–598.

Iancu, D.A., Trichakis, N., 2013. Pareto efficiency in robust optimization. Manag. Sci. 60, 130–147.

Ke, L., Xu, Z., Feng, Z., Shang, K., Qian, X., 2013. Proportion-based robust optimization and team orienteering problem with interval data. Eur. J. Oper. Res. 226, 19–31.

Li, Z., Floudas, C.A., 2014. A comparative theoretical and computational study on robust counterpart optimization: III. Improving the quality of robust solutions. Ind. Eng. Chem. Res. 53, 13112–13124.

Li, Z., Li, Z., 2015. Optimal robust optimization approximation for chance constrained optimization problem. Comput. Chem. Eng. 74, 89–99.

Li, Z., Ding, R., Floudas, C.A., 2011. A comparative theoretical and computational study on robust counterpart optimization: I. Robust linear optimization and robust mixed integer linear optimization. Ind. Eng. Chem. Res. 50, 10567–10603.

Li, Z., Tang, Q., Floudas, C.A., 2012. A comparative theoretical and computational study on robust counterpart optimization: II. Probabilistic guarantees on constraint satisfaction. Ind. Eng. Chem. Res. 51, 6769–6788.

Lin, X., Janak, S.L., Floudas, C.A., 2004. A new robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty. Comput. Chem. Eng. 28, 1069–1085.

Paschalidis, I.C., Kang, S.-C.,2015. Robust linear optimization: on the benefits of distributional information and applications in inventory control. In: 44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. IEEE, pp. 4416–4421.

Ruckdeschel, P., Kohl, M., 2010. General Purpose Convolution Algorithm in S4-Classes by Means of FFT, arXiv:1006.0764.

Ruckdeschel, P., Kohl, M., Stabla, T., Camphausen, F., 2015. S4 classes for distributions a manual for packages "distr", "distrex", "distrellipse", "distrmod", "distrsim", "distrtest", "distrteach", Version 2.5.

Soyster, A.L., 1973. Technical note convex programming with set-inclusive constraints and applications to inexact linear programming. Oper. Res. 21, 1154–1157.

Tsiligirides, T., 1984. Heuristic methods applied to orienteering. J. Oper. Res. Soc., 797–809.

Vansteenwegen, P., Souffriau, W., Van Oudheusden, D., 2011. The orienteering problem: a survey. Eur. J. Oper. Res. 209, 1–10.

Yuan, Y., Li, Z., Huang, B., 2016. Robust optimization under correlated uncertainty: formulations and computational study. Comput. Chem. Eng. 85, 58–71.