



On a conservative partition refinement (CPR) method for a class of two-stage stochastic programming problems

Carlos Andrés Gamboa¹ · Davi Michel Valladão¹ · Alexandre Street¹

Received: 14 January 2021 / Accepted: 5 December 2021 / Published online: 20 January 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

Two-stage stochastic programming is a mathematical framework widely used in real-life applications such as power system operation planning, supply chains, logistics, inventory management, and financial planning. Since most of these problems cannot be solved analytically, decision makers make use of numerical methods to obtain a near-optimal solution. Some applications rely on the implementation of non-converged and therefore sub-optimal solutions because of computational time or power limitations. In this context, the existing partition-refinement methods provide an optimistic solution whenever convergence is not attained. Optimistic solutions often generate high disappointment levels because they consistently underestimate the actual costs in the approximate objective function. To address this issue, we developed a conservative convergent partition-refinement method for two-stage stochastic linear programming problems with a convex recourse function of the uncertainty. Given a partition of the uncertainty support, a conservative decision can be obtained by means of a distributionally robust problem whose complexity grows exponentially with the uncertainty dimensionality. We prove the convergence of the method given a refining partition sequence and propose algorithmic schemes to address the problem of dimensionality. For problems with low-dimensional uncertainty, we developed a deterministic equivalent linear programming model; whereas, for medium-sized uncertainty dimensionality, we propose a column and constraint generation algorithm. To handle high dimensional uncertainty, we propose a simplex-based heuristic method whose complexity grows linearly with the uncertainty dimension—size of the random vector. In the presence of monotone recourse functions with regard to an uncertain parameter, we prove convergence of the proposed simplex-based heuristic method. Computational

✉ Carlos Andrés Gamboa
caagamboaro@aluno.puc-rio.br

Davi Michel Valladão
davimv@puc-rio.br

Alexandre Street
street@ele.puc-rio.br

¹ R. Marquês de São Vicente, 225, Rio de Janeiro, Brazil

experiments are presented for a farmer's problem, an aircraft allocation problem, and a Unit Commitment problem.

Keywords Two-stage stochastic programming · Exact bound partition methods · Conservative solution

Mathematics Subject Classification 90C15

1 Introduction

Two-stage stochastic programming is a mathematical framework to model decision making under uncertainty. In this context, first-stage decisions are made under uncertainty, while the second-stage (or recourse) decisions are intended to improve the impact of the first-stage decisions after observing the uncertainty realization. This framework is widely used in real-world applications covering the planning of power system operations and supply chains, logistics, inventory, and financial planning, to mention a few. Since there is no analytical solution to most two-stage stochastic optimization problems, efficient numerical methods are of paramount importance.

One of the first numerical methods explored in the stochastic programming literature utilized bounds for the expected recourse function. Initiated by [15], this approach considers minimization problems of which the recourse is a convex function of the uncertainty. It relies on a partition of the uncertainty support, the law of total probability, and Jensen's and Edmundson–Madanski inequalities in order to obtain lower and upper bounds. Partition refinement methods ensure improvements to the approximation error with a monotone sequence of limits (see [6,9–12]). It is worth mentioning that this numerical approach appeared as first solution method to stochastic optimization problems, since it is based on numerical integration techniques that are easy to implement, that is why the referenced works are fairly old. With the advancement of computers, this approach was replaced by sampling-based techniques, but in this work we want to return to the partition-based method to highlight the importance of obtaining a solution given by deterministic limits.

Starting from the law of total probability, previous works have provided an optimistic solution (Solution that underestimates the actual cost) by applying Jensen's inequality to each partition cell to obtain the lower approximation problem in a computationally tractable fashion. The optimal value of this proxy provides a lower bound to the original problem because it replaces the expected recourse function by a weighted average recourse evaluation. A deterministic optimality gap is obtained by fixing the current solution and computing an upper-bound that applies Edmundson–Madanski inequality to each cell of the partition and averages its results with the associated probability mass. Given a refining partition sequence, this method converges to the true optimal in an optimistic manner, i.e., at each iteration, and the solution underestimates the actual recourse costs.

The available partition refinement methods in literature only provide optimistic solutions when the method is not converged. However, some practical applications depend on the implementation of a non-converged sub-optimal solution because of

computational time limitations. In these cases, an optimistic solution might generate significantly higher expected costs when compared to the obtained objective function. In this context, a solution methodology that provides pre-convergence conservative solutions is of significant importance to practical applications with high computational burden and time limitations.

The objective of this study is to obtain a conservative convergent solution method for two-stage stochastic linear optimization problems of which the recourse is a convex function of the uncertainty. We formulated a distributionally robust optimization model based on a generalization of Edmundson-Madanski inequality (see [6,13,14]), and solved it to obtain a conservative solution and a tighter upper bound to the original problem. The distributionally robust model minimizes the worst expected cost over every extreme probability measure with known partition-adapted conditional expectations. A deterministic optimality gap was obtained by solving the lower-bound problem that applied Jensen's inequality to each partition cell and computing the distance between both limits. This technique was motivated by the idea that distributionally robust solutions avoid out-of-sample disappointment [19].

Nevertheless, in the presence of high dimensional uncertainty vectors, the proposed method is challenged because of the exponential growth of the number of linear constraints and variables. To handle this, we propose different solution schemes depending on the uncertainty dimensionality: (i) for problems with low-dimensional uncertainty, we developed a deterministic equivalent linear programming model, (ii) for medium-sized uncertainty dimensionality, we propose a column and constraint generation algorithm [24], and (iii) to handle high dimensional uncertainty, we propose a simplex-based heuristic method whose complexity grows linearly with the uncertainty dimension. For the latter, we prove convergence when the recourse function is monotone over the uncertainty. Therefore, the main contributions of the study are the following:

- Raise awareness on the negative impact of optimistic solution methods to solve two-stage stochastic optimization problems.
- Reformulation of the upper-bound, conservative, distributionally robust problem into a linear programming model.
- Development of two acceleration procedures, i.e., (1) an exact decomposition approach based on the column and constraint generation algorithm to solve medium-scaled problems and (2) an approximative simplex-based partitioning scheme to find robust solutions for large-scale instances.
- A new conservative solution framework for two-stage stochastic linear optimization problems based on a deterministic partition refinement algorithm.

The remainder of this paper is organized as follows. Section 2 describes the theoretical and methodological background on the existing partition-based method and related bounding optimization problems. In Sect. 3, we develop the deterministic equivalent reformulation of the upper-bound problem and propose accelerating methods to handle medium- and large-sized problems. Section 4 presents the proposed conservative partition refining (CPR) method for two-stage stochastic optimization and proves the convergence of the sequential algorithm. Section 5 presents the numerical results of

the computational experiments showing the superior pre-convergence performance of the proposed method. Finally, in Sect. 6, relevant conclusions are provided.

2 Existing methods

In this section, we introduce the nomenclature and provide a short review of the existing partition methods from which our contributions are derived.

2.1 Notation

We study two-stage stochastic linear programming problems of the form:

$$z^* := \text{Min}_{\mathbf{x} \in X} \{f(\mathbf{x}) := \mathbf{c}^\top \mathbf{x} + \mathbb{E}[Q(\mathbf{x}, \boldsymbol{\xi})]\} \tag{1}$$

where

$$\begin{aligned} Q(\mathbf{x}, \boldsymbol{\xi}) &= \text{Min}_{\mathbf{y} \geq 0} \mathbf{q}^\top \mathbf{y} \\ \text{s.t. } & \mathbf{W}\mathbf{y} \geq \mathbf{h}(\boldsymbol{\xi}) - \mathbf{T}(\boldsymbol{\xi})\mathbf{x}. \end{aligned} \tag{2}$$

is known as the recourse function or second-stage problem. This model corresponds to a linear optimization problem that minimizes the cost where \mathbf{x} denotes the first-stage decisions, \mathbf{y} denotes the second-stage decisions and the expectation $\mathbb{E}[Q(\mathbf{x}, \boldsymbol{\xi})]$ represents the expected cost of the recourse. We assume that the random vector $\boldsymbol{\xi}$ has known probability distribution with support in $\Xi \subseteq \mathbb{R}^{d_\xi}$. We denote the set of feasible first-stage decisions by $X \subseteq \mathbb{R}^{d_x}$. Here, $\mathbf{c} \in \mathbb{R}^{d_x}$, $Q : X \times \text{co}(\Xi) \rightarrow \mathbb{R}$, $\mathbf{q} \in \mathbb{R}^{d_y}$, $\mathbf{W} \in \mathbb{R}^{m_y \times d_y}$, $\mathbf{h}(\boldsymbol{\xi}) \in \mathbb{R}^{m_y}$, and $\mathbf{T}(\boldsymbol{\xi}) \in \mathbb{R}^{m_y \times d_x}$, where $\text{co}(\Xi)$ denotes the convex hull of the uncertainty support.

We assume that problem (1) has relatively complete recourse, i.e., problem (2) is feasible for every $\mathbf{x} \in X$, and every realization of the unknown data $\boldsymbol{\xi}$; the uncertainty support Ξ is compact and the expectation $\mathbb{E}[Q(\mathbf{x}, \boldsymbol{\xi})]$ exists. We further assume that $Q(\mathbf{x}, \cdot)$ is a convex function on $\text{co}(\Xi)$ for all $\mathbf{x} \in X$. Particularly, this last assumption holds when the unknown data is only in the components of $\mathbf{h}(\boldsymbol{\xi})$ and $\mathbf{T}(\boldsymbol{\xi})$, respectively.

The existing partition-based method [9,11,12,14,15,17] makes use of a partition of the support of the uncertainty to generate a monotonic sequence of limits. These limits are given by the optimal objective values of the upper- and lower-bound problems obtained by the classical inequalities of Edmundson-Mandanski and Jensen, respectively.

According to previous reported partition methods, we start from the partition of the uncertainty support Ξ . The set $\mathcal{P}_n = \{\Xi^k : k = 1, \dots, n\}$ of cells $\Xi^k \subseteq \Xi$, is a partition of the support $\Xi \subset \mathbb{R}^{d_\xi}$ if:

1. $\mathbb{P} \left(\bigcap_{k \in K} \Xi^k \right) = 0, \quad \forall K \subseteq \{1, \dots, n\},$
2. $\bigcup_{k=1}^n \Xi^k = \Xi.$

Since most real-life applications are based on bounded probability distributions for the uncertainty, in this work, we assume rectangular support for the data-generating probability distribution. Moreover, we assume rectangular partitions with cells of the form $\Xi^k = \times_{i=1}^{d_\xi} [a_i^k, b_i^k]$, because the partition refinement procedure for this partition regards the method computational tractable.

We denote by $p^k = \mathbb{P}(\xi \in \Xi^k)$ the probability mass, and $\bar{\xi}^k = \mathbb{E}[\xi | \xi \in \Xi^k]$ the conditional mean of the cell $k, k = 1, \dots, n$. It is worth mentioning that, for a general probability distribution, the computation of the probability mass p^k and the conditional mean $\bar{\xi}^k$ is not easy. In Sect. 5 we present a different methods to assess this computation.

Finally, even though the disappointment concept has been widely studied by the risk-averse theory and there exist several definitions under this concept, in this work, we assume the following definition:

Definition 1 The disappointment is the difference between the out-of-sample cost and the in-sample estimate (i.e., the optimal objective value).

2.2 Existing lower bound

According to [15], if we consider a partition of the uncertainty support Ξ , by the law of total probability, we have that the expectation in (1) can be expressed by $\mathbb{E}[Q(\mathbf{x}, \xi)] = \sum_{k=1}^n p^k \mathbb{E}[Q(\mathbf{x}, \xi) | \xi \in \Xi^k]$. Then (1) is equivalent to the following linear programming problem:

$$\text{Min}_{\mathbf{x} \in X} \left\{ f(\mathbf{x}) := \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^n p^k \mathbb{E}[Q(\mathbf{x}, \xi) | \xi \in \Xi^k] \right\} \tag{3}$$

By Jensen’s inequality we have that

$$Q(\mathbf{x}, \bar{\xi}^k) \leq \mathbb{E}[Q(\mathbf{x}, \xi) | \xi \in \Xi^k], \quad k = 1, \dots, n,$$

which gives the following lower-bound for the optimal objective value of (3):

$$\begin{aligned} z_n^L &:= \text{Min}_{\mathbf{x} \in X} \left\{ f_n^L(\mathbf{x}) := \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^n p^k Q(\mathbf{x}, \bar{\xi}^k) \right\} \\ &\leq \text{Min}_{\mathbf{x} \in X} \left\{ \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^n p^k \mathbb{E}[Q(\mathbf{x}, \xi) | \xi \in \Xi^k] \right\}. \end{aligned} \tag{4}$$

Note that the lower-bound problem on the left side of Eq. (4) underestimates the expected cost, because it only considers the finite set of conditional mean scenarios $\{\bar{\xi}^k : k = 1, \dots, n\}$ to approximate $\mathbb{E}[Q(\mathbf{x}, \xi)]$.

We can explicitly write the lower-bound problem as the following deterministic-equivalent linear program:

$$\begin{aligned}
 & \text{Min}_{\mathbf{x}, \mathbf{y}^k} \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^n p^k \cdot \mathbf{q}^\top \mathbf{y}^k \\
 & \text{s.t. } \mathbf{W}\mathbf{y}^k \geq \mathbf{h}(\bar{\boldsymbol{\xi}}^k) - \mathbf{T}(\bar{\boldsymbol{\xi}}^k)\mathbf{x}, \quad k = 1, \dots, n, \\
 & \quad \mathbf{y}^k \geq 0, \quad k = 1, \dots, n, \\
 & \quad \mathbf{x} \in X.
 \end{aligned}
 \tag{5}$$

Note that the number of blocks of linear constraints for (5) is the same as the partition size, meaning that the lower-bound problem does not demand a high computational effort. When n is not sufficiently large, the optimal solution \mathbf{x}_n^L of (5) represents an optimistic decision because of the lower approximation from below to the conditional expectation $\mathbb{E}[Q(\mathbf{x}, \boldsymbol{\xi}) | \boldsymbol{\xi} \in \Xi^k]$ by $Q(\mathbf{x}, \bar{\boldsymbol{\xi}}^k)$, $k = 1, \dots, n$, underestimating the actual conditional expected cost.

2.3 Existing upper bound

To derive an upper bound for the expectation $\mathbb{E}[Q(\mathbf{x}, \boldsymbol{\xi})]$, Edmundson-Madanski [18] initially proposed an inequality based on a convex combination of the value of $Q(\mathbf{x}, \cdot)$ at the extreme points of the convex hull of the support Ξ . Therefore, if $\{\mathbf{e}_j : j = 1, \dots, 2^{d_\xi}\}$ is the set of extreme points of $\Xi = \times_{i=1}^{d_\xi} [a_i, b_i]$, since every point $\boldsymbol{\xi} \in \Xi$ can be expressed by a convex combination of the vertices of Ξ , it is always possible to find $p_j(\boldsymbol{\xi}) \geq 0, j = 1, \dots, 2^{d_\xi}$, such that

$$\sum_{j=1}^{2^{d_\xi}} p_j(\boldsymbol{\xi}) \cdot \mathbf{e}_j = \boldsymbol{\xi} \quad \forall \boldsymbol{\xi} \in \Xi
 \tag{6}$$

and

$$\sum_{j=1}^{2^{d_\xi}} p_j(\boldsymbol{\xi}) = 1 \quad \forall \boldsymbol{\xi} \in \Xi.
 \tag{7}$$

The weights $p_j(\boldsymbol{\xi}), j = 1, \dots, 2^{d_\xi}$, can be interpreted as conditional probabilities, i.e., $p_j(\boldsymbol{\xi}) = \mathbb{P}(\mathbf{e} = \mathbf{e}_j | \boldsymbol{\xi})$, considering a random vector \mathbf{e} with support in the set of extreme points $\{\mathbf{e}_j : j = 1, \dots, 2^{d_\xi}\}$. Given that (6) and (7) are true and $Q(\mathbf{x}, \cdot)$ is a convex function for any given \mathbf{x} , the following inequality holds for any set of conditional probabilities $\{(p_1(\boldsymbol{\xi}), \dots, p_{2^{d_\xi}}(\boldsymbol{\xi}))\}_{\boldsymbol{\xi} \in \Xi}$:

$$\mathbb{E}[Q(\mathbf{x}, \boldsymbol{\xi})] = \int_{\Xi} Q\left(\mathbf{x}, \sum_{j=1}^{2^{d_\xi}} p_j(\boldsymbol{\xi}) \cdot \mathbf{e}_j\right) d\mathbb{P}(\boldsymbol{\xi}) \leq \sum_{j=1}^{2^{d_\xi}} Q(\mathbf{x}, \mathbf{e}_j) \int_{\Xi} p_j(\boldsymbol{\xi}) d\mathbb{P}(\boldsymbol{\xi}).
 \tag{8}$$

According to the above, we will consider the worst-case of the right-hand-side of the inequality (8) because of through an infinite maximization problem over the set of conditional probabilities that satisfies (6), we are able to derive an upper bound for (3) by a finite tractable optimization problem, as it can be seen in the following.

Since (8) holds for any set of conditional probabilities $\{(p_1(\xi), \dots, p_{2^{d_\xi}}(\xi))\}_{\xi \in \Xi}$ that satisfies (6) then,

$$\begin{aligned} \mathbb{E}[Q(\mathbf{x}, \xi)] &\leq \text{Max}_{p_j(\xi)} \sum_{j=1}^{2^{d_\xi}} Q(\mathbf{x}, \mathbf{e}_j) \int_{\Xi} p_j(\xi) d\mathbb{P}(\xi) \\ \text{s.t.} \quad &\sum_{j=1}^{2^{d_\xi}} \mathbf{e}_j p_j(\xi) = \xi \quad \forall \xi \in \Xi \\ &\sum_{j=1}^{2^{d_\xi}} p_j(\xi) = 1 \quad \forall \xi \in \Xi. \end{aligned} \tag{9}$$

Note that (9) is a semi-infinite optimization problem that allows us to derive an upper bound through a tractable finite relaxation. If we replace the two constraints in (9) with their expected value, i.e.,

$$\begin{aligned} \sum_{j=1}^{2^{d_\xi}} \mathbf{e}_j \int_{\Xi} p_j(\xi) d\mathbb{P}(\xi) &= \bar{\xi} \\ \sum_{j=1}^{2^{d_\xi}} \int_{\Xi} p_j(\xi) d\mathbb{P}(\xi) &= 1, \end{aligned}$$

and denote $\delta_j = \int_{\Xi} p_j(\xi) d\mathbb{P}(\xi)$, for $j = 1, \dots, 2^{d_\xi}$, we obtain an upper bound based on the following tractable finite linear optimization problem:

$$\begin{aligned} \text{Max}_{\delta \in \mathcal{D}(\bar{\xi})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e})] &:= \text{Max}_{\delta \geq 0} \sum_{j=1}^{2^{d_\xi}} Q(\mathbf{x}, \mathbf{e}_j) \delta_j \\ \text{s.t.} \quad &\sum_{j=1}^{2^{d_\xi}} \mathbf{e}_j \delta_j = \bar{\xi}, \\ &\sum_{j=1}^{2^{d_\xi}} \delta_j = 1, \end{aligned} \tag{10}$$

where $\mathcal{D}(\bar{\xi}) = \left\{ \delta \in \mathbb{R}_+^{2^{d_\xi}} : \sum_{j=1}^{2^{d_\xi}} \mathbf{e}_j \delta_j = \bar{\xi}, \sum_{j=1}^{2^{d_\xi}} \delta_j = 1 \right\}$. Thus, as per the above developments, the following inequality holds:

$$\mathbb{E}[Q(\mathbf{x}, \xi)] \leq \text{Max}_{\delta \in \mathcal{D}(\bar{\xi})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e})]. \tag{11}$$

Existing partition-based methods used the derived the upper bound (11) for a first-stage decision found by the lower bound problem (5). So, supposing that \mathbf{x}_n^L is the solution of (5), in [9,12,13] the following upper bound is proposed:

$$z^* \leq f_n^U(\mathbf{x}_n^L) := \mathbf{c}^\top \mathbf{x}_n^L + \sum_{k=1}^n p^k \text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \left(\sum_{j=1}^{2^{d_\xi}} \delta_j^k Q(\mathbf{x}_n^L, \mathbf{e}_j^k) \right), \tag{12}$$

where,

$$\begin{aligned} Q(\mathbf{x}_n^L, \mathbf{e}_j^k) &:= \text{Min}_{\mathbf{y} \geq 0} \mathbf{q}^\top \mathbf{y} \\ \text{s.t. } \mathbf{W}\mathbf{y} &\geq \mathbf{h}(\mathbf{e}_j^k) - \mathbf{T}(\mathbf{e}_j^k)\mathbf{x}_n^L, \end{aligned} \tag{13}$$

for all $j = 1, \dots, 2^{d_\xi}$ and $k = 1, \dots, n$.

The existing partition-based method is summarized in the following pseudo-algorithm:

Algorithm 1 Existing partition-based methods

Input: $\mathcal{P}^1 = \{\Xi\}$, $\epsilon > 0$ stopping criteria.

Output: \mathbf{x}_n^L the optimistic solution of the lower bound problem.

- 1: Solve (5) to determine the optimal solution \mathbf{x}_n^L .
 - 2: Evaluate \mathbf{x}_n^L in the recourse problem (13) for $j = 1, \dots, 2^{d_\xi}$, and $k = 1, \dots, n$, to get an upper bound for the optimal objective value of (3), according to Eq. (12).
 - 3: **if** $\frac{f_n^U(\mathbf{x}_n^L) - z_n^L}{f_n^L(\mathbf{x}_n^L)} \leq \epsilon$ **then**
 - 4: stop and return \mathbf{x}_n^L as the optimal solution.
 - 5: **else**
 - 6: refine the partition \mathcal{P}_n to \mathcal{P}_{n+1} and return to 1
 - 7: **end if**
-

The existing partition-based methods are based on the lower bound problem, as it underestimates the recourse cost by relying only on the partition-adapted conditional means within the recourse function assessment. Therefore, pre-convergence solutions of the lower-bound problem are optimistic and generate high disappointment levels when they are evaluated for adverse scenarios of the uncertainty realization as illustrated in Sect. 5 of the computational experiments.

3 Proposed upper-bound problem reformulation and solution methods

Motivated by the concept explored in [19], i.e., conservative solutions obtained by solving distributionally robust optimization problems have performance guarantees

against out-of-sample disappointment. Our thrust is to obtain pre-convergence conservative solutions in a computationally efficient manner. For that, we develop in this section a deterministic equivalent reformulation of the upper-bound (distributionally robust) problem that is computationally tractable for low-dimensional uncertainty. We also propose a column-constraint generation algorithm for medium-sized uncertainty dimensionality and, to handle high dimensional uncertainty, we propose a simplex-based heuristic method whose complexity grows linearly with the uncertainty dimension. In the presence of monotone recourse functions with regard to an uncertain parameter, we prove convergence of the proposed simplex-based heuristic method. The assumption of monotonicity of the resource function is reasonable in light of the fact that most real-world applications belong to this kind of problem.

3.1 Deterministic equivalent model for upper-bound problem

Following the upper approximation of the expected recourse function (11), we define the upper-bound problem

$$z_n^U := \text{Min}_{\mathbf{x} \in X} \left\{ f_n^U(\mathbf{x}) := \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^n p^k \text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^k)] \right\}, \tag{14}$$

where \mathbf{e}^k is the random vector whose support is the set of vertex of the hypercube Ξ^k and $\mathcal{D}(\bar{\xi}^k)$ is the ambiguity set generated by the conditional mean $\bar{\xi}^k$ of the cell Ξ^k .

The solutions obtained from the upper-bound problem can be seen as robust or conservative solutions. Note that problem (14) is a distributionally robust optimization problem, where the conditional-probability distribution within each cell is selected to represent the worst-case distribution preserving the conditional-average information of the cell [5]. Since the upper-bound problem (14) overestimates the actual cost (11), the upper-bound solution has a mathematical certificate against disappointment. This is specially useful when non-converged solutions are actually implemented owing to time or computational-power limitations.

However, solving the upper-bound problem to obtain the conservative solution \mathbf{x}_n^U requires significant computational effort. The number of variables of the inner problem grows exponentially with the uncertainty dimension. In this work, we propose a new framework to obtain a conservative solution by solving the upper-bound problem. To the best of our knowledge, no existing partition-based method solves the upper bound problem.

For a given cell Ξ^k of the rectangular partition, the upper-bound for the conditional-expected recourse cost (10) can be recast according to its dual formulation as follows:

$$\begin{aligned} \text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^k)] &= \text{Min}_{\pi, \theta} \pi + \theta^\top \bar{\xi}^k \\ \text{s.t. } \pi + \theta^\top \mathbf{e}_j^k &\geq Q(\mathbf{x}, \mathbf{e}_j^k), \quad j = 1, \dots, 2^{d_\xi}. \end{aligned} \tag{15}$$

where $\pi^k \in \mathbb{R}$ and $\theta^k \in \mathbb{R}^{d_\xi}$ be the dual variables associated with the first and second linear constraints of (10), respectively.

Replacing the worst-case conditional-expected recourse cost of (14) with its dual formulation (15), the upper-bound problem (14) can be recast as the following linear optimization problem:

$$\begin{aligned}
 \text{Min}_{\pi^k, \theta^k, \mathbf{x}} \quad & \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^n p^k (\pi^k + [\theta^k]^\top \bar{\xi}^k) \\
 \text{s.t.} \quad & \pi^k + [\theta^k]^\top \mathbf{e}_j^k \geq Q(\mathbf{x}, \mathbf{e}_j^k), \quad j = 1, \dots, 2^{d_\xi}, \quad k = 1, \dots, n \\
 & \mathbf{x} \in X.
 \end{aligned} \tag{16}$$

Since the recourse function is a minimization problem, we are able to obtain the deterministic equivalent

$$\text{Min}_{\pi^k, \theta^k, \mathbf{x}, \mathbf{y}} \quad \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^n p^k (\pi^k + [\theta^k]^\top \bar{\xi}^k) \tag{17a}$$

$$\text{s.t.} \quad \pi^k + [\theta^k]^\top \bar{\xi}^k \geq \mathbf{q}^\top \mathbf{y}_{\bar{\xi}^k}, \quad \forall k = 1, \dots, n \tag{17b}$$

$$\mathbf{W} \mathbf{y}_{\bar{\xi}^k} \geq \mathbf{h}(\bar{\xi}^k) - \mathbf{T}(\bar{\xi}^k) \mathbf{x}, \quad \forall k = 1, \dots, n \tag{17c}$$

$$\pi^k + [\theta^k]^\top \mathbf{e}_j^k \geq \mathbf{q}^\top \mathbf{y}_j^k, \quad \forall j \in J^k, \quad \forall k = 1, \dots, n \tag{17d}$$

$$\mathbf{W} \mathbf{y}_j^k \geq \mathbf{h}(\mathbf{e}_j^k) - \mathbf{T}(\mathbf{e}_j^k) \mathbf{x}, \quad \forall j \in J^k, \quad \forall k = 1, \dots, n \tag{17e}$$

$$\mathbf{x} \in X. \tag{17f}$$

which is an equivalent formulation of the original problem (16) whenever J^k comprises all vertices of cell Ξ^k . This equivalent formulation consists of replacing the recourse function by the second-stage objective function, including all second stage decisions as variables and adding all second stage feasibility constraints. Indeed, if the left-hand-side (LHS) of the first block of constraints is greater than or equal to the second stage cost of a feasible second-stage solution, then the LHS is greater than or equal to the minimum second stage cost given by the recourse function.

Note for instance that (17) is a linear programming problem whenever X is a polyhedral set and can be efficiently solved whenever for problems with low dimension uncertainty vector. However, problem (17) is in general an intractable problem for medium- and large-scale instances as it relies on an exponential set of constraints. To handle this challenge, we present two acceleration procedures: (1) an exact decomposition approach based on the column and constraint generation algorithm [24] to solve medium-dimensional problems and (2) an heuristic procedure for high-dimensional uncertainty based on a simplex-based heuristic method using a circumscribed simplex for each partition cell. For the latter, we prove convergence whenever the recourse function over the uncertainty dimension is monotone.

3.2 Column-constraint generation algorithm to handle medium-sized uncertainty dimensionality

For problems with medium-sized uncertainty dimension, we adapted the column and constraint generation framework to iteratively identify a subset of constraints that ensures feasibility to the original problem. The thrust of this algorithm is to obtain an equivalent formulation of (16) with significantly fewer constraints and variables.

According to [24], the column and constraint generation procedure is implemented in a master-oracle scheme. In our context, the master problem is a relaxation of an equivalent formulation of the original problem (16). Given the solution of the master (relaxed) problem, the oracle finds the vertex that generates the worst infeasibility and adds the associated block of linear constraints and decision variables to the master problem. This iterative procedure stops whenever the oracle asserts that the master solution is feasible for the original problem.

For the initial iteration $\ell = 0$ of the column-constraint algorithm, we start with an empty set $J^k = \emptyset$ of constraints, i.e., the initial master problem is a relaxation of (17). For initialization purposes, we also add a block of constraints and variables (17b)–(17c) associated with the conditional expectation $\bar{\xi}^k$ adapted to cell Ξ^k . This additional constraint does not affect the original feasible set since it can be represented as an weighted average of all constraints associated to each vertex.

For a given iteration ℓ , with an updated set J^k , we solve (17) and obtain a candidate solution $(\pi_\ell^*, \theta_\ell^*, \mathbf{x}_\ell^*)$. Then, we solve the oracle problem

$$\text{ORACLE : } \vartheta^* = \text{Max}_j \{Q(\mathbf{x}_\ell^*, \mathbf{e}_j^k) - \pi_\ell^* - [\theta_\ell^*]^\top \mathbf{e}_j^k\}, \tag{18}$$

to find the highest constraint violation given the current solution. To efficiently solve the oracle problem, we replace the recourse function by its dual representation and combine them in a single maximization problem

$$\begin{aligned} \text{Max}_{j, \lambda \geq 0} \quad & \lambda^\top (\mathbf{h}(\mathbf{e}_j^k) - \mathbf{T}(\mathbf{e}_j^k) \mathbf{x}^*) - \pi^* - [\theta^*]^\top \mathbf{e}_j^k \\ \text{s.t.} \quad & \mathbf{W}^\top \lambda \leq \mathbf{q}. \end{aligned} \tag{19}$$

Following [23], we avoid solving (19) by enumeration of vertex $\{\mathbf{e}_j^k : j = 1, \dots, d_\xi\}$ of the cell Ξ^k by introducing binary variables to represent each vertex and transform (19) into a MILP equivalent formulation presented in details in Appendix A. The algorithm stops whenever the oracle optimal value is non-positive, i.e., the current solution $(\pi_{\ell+1}^*, \theta_{\ell+1}^*, \mathbf{x}_{\ell+1}^*)$ is feasible for the original problem. If the oracle optimal value is positive, then update the set J^k including the oracle solution i^* and repeat the process solving again the master problem.

We summarize the proposed column and constraint generation algorithm in the following pseudo-algorithm:

Algorithm 2 Column and Constraint Generation framework

Input: $\hat{\xi}^k, \ell = 0$, set $J^k := \emptyset$ of the constraints generated by the vertex of the cell Ξ^k .
Output: \mathbf{x}_n^U the conservative solution of the upper bound problem.
 1: Initialization. Solve the problem (17) to derive an optimal solution. $(\pi_0^*, \theta_0^*, \mathbf{x}_0^*)$ with $J^k := \emptyset$.
 2: Maximize infeasibility. Find j^* as the optimal solution of (18).
 3: **if** $\vartheta^* \leq 0$ **then**
 4: **return** $\mathbf{x}_\ell^* = \mathbf{x}_n^U$ as the optimal solution.
 5: **else**
 6: Update $\ell = \ell + 1$. Do $J^k \cup \{j^*\}$ and go to the next step
 7: **end if**
 8: Update $(\pi_{\ell+1}^*, \theta_{\ell+1}^*, \mathbf{x}_{\ell+1}^*)$ as the optimal solution of the master problem (17) and go to 2.

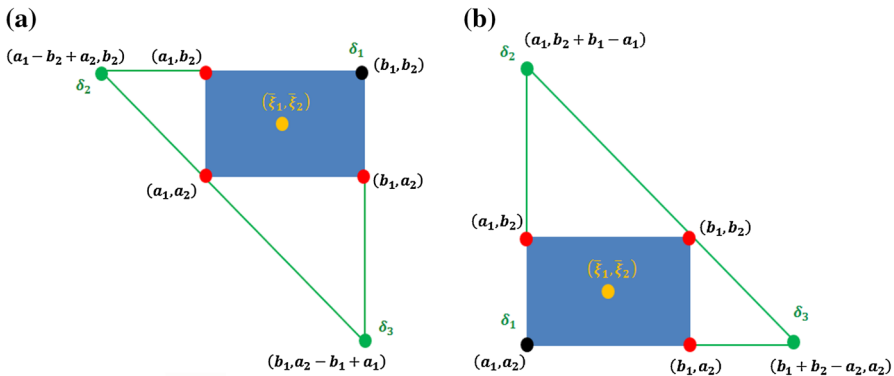


Fig. 1 The blue region in the figure corresponds to the original support $[a_1, b_1] \times [a_2, b_2]$, and the simplex, represented in green, is its extension. The vertex in black denotes the selected $\hat{\xi}$ of the realization of the uncertainty and the point highlighted at the middle indicates the mean. **a** $\hat{\xi} = (b_1, b_2)$; **b** $\hat{\xi} = (a_1, a_2)$

3.3 Simplex-based heuristic method to handle high-dimensional uncertainty

As the master problem of the column-constraint generation algorithm has exponential size with the uncertainty dimensionality, there are some instances of high dimensional uncertainty that reaches the tractability limit of this accelerating method. For these cases, we developed an heuristic method to obtain an upper bound and corresponding conservative solution.

To handle high dimensional uncertainty, we propose a heuristic solution method that extends the original box uncertainty support to a circumscribed simplex polyhedral where the number of vertexes depends linearly on the uncertainty dimension. Under the extended support, we reformulate (15) to obtain an upper-bound problem whose complexity grows linearly with the uncertainty dimension.

The proposed extension is the minimum volume simplex that contains the original cell (see Fig. 1) and one selected vertex $\hat{\xi}^k$ coinciding with the cell’s vertexes. This simplex has the property that the length of the edges that contain the original vertex $\hat{\xi}$ is equal to the sum of the length of the projection of the original support along to each dimension, respectively.

For example in Fig. 1, it is represented the extension of the support $[a_1, b_1] \times [a_2, b_2]$ by the simplex represented in green. We considered two possibilities $\{(a_1, a_2), (b_1, b_2)\}$ for the original vertex $\hat{\xi}^k$, indicated by the black point, that corresponds to the extreme events of the realization of the uncertainty. In both cases, the length of the edges of the simplex that contain the vertex $\hat{\xi} \in \{(a_1, a_2), (b_1, b_2)\}$ is equal to $(b_1 - a_1) + (b_2 - a_2)$, the sum of dimensions. We use this property as a simple rule to create the simplex. Figure 2 (b1) and (b2) illustrate how the chosen cell (highlighted in blue) in the refinement procedure is extended to the green simplex for two different partition sizes to visualize the extension procedure.

The proposed simplex-based heuristic method solves the sequential partition refinement problem by considering the extension of each cell of the partition. Based on the extension of cell Ξ^k , problem (15) can be rewritten as follows:

$$\begin{aligned} \text{Max}_{\delta \in \mathcal{R}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{v}^k)] &:= \text{Min}_{\pi, \theta} \pi + \theta^\top \bar{\xi}^k \\ \text{s.t. } \pi + \theta^\top \mathbf{v}_j^k &\geq Q(\mathbf{x}, \mathbf{v}_j^k), \quad j = 1, \dots, (d_\xi + 1), \end{aligned} \tag{20}$$

where $\mathcal{R}(\bar{\xi}^k) = \left\{ \delta_j^k \in \mathbb{R}^{d_\xi+1} : \sum_{j=1}^{d_\xi+1} \delta_j^k \mathbf{v}_j^k = \bar{\xi}^k, \sum_{j=1}^{d_\xi+1} \delta_j^k = 1 \right\}$, \mathbf{v}^k is a random vector with support in the set $\{\mathbf{v}_j^k : j = 1, \dots, d_\xi + 1\}$ of vertexes of the simplex that contains the cell Ξ^k and $\mathbf{v}_1^k := \hat{\xi}^k$ is the original vertex of the hypercube Ξ^k . Note that the vertex $\mathbf{v}_j^k, j = 2, \dots, (d_\xi + 1)$, differs from $\hat{\xi}^k$ in just one component, according to the rule to create the simplex that contains the cell Ξ^k . That said, we propose the simplex-based heuristic problem

$$\begin{aligned} \tilde{z}_n^U &:= \text{Min}_{\pi^k, \theta^k, \mathbf{x}} \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^n p^k (\pi^k + [\theta^k]^\top \bar{\xi}^k) \\ \text{s.t. } \pi^k + [\theta^k]^\top \mathbf{v}_j^k &\geq Q(\mathbf{x}, \mathbf{v}_j^k), \quad j = 1, \dots, (d_\xi + 1), k = 1, \dots, n \\ \mathbf{x} &\in X. \end{aligned} \tag{21}$$

Note that the number of blocks of linear constraints of the problem (21) is $n \cdot (d_\xi + 1)$, i.e., it depends linearly on the uncertainty dimension. With this alternative upper-bound problem we can solve the sequential partition refinement problem described in Algorithm 3 solving (21) instead (16).

In particular, by assuming monotonicity of the recourse function $Q(\mathbf{x}, \cdot)$ over the uncertainty vector $\xi \in \Xi$, and that the selected original vertex $\hat{\xi}^k$ of the hypercube Ξ^k is the worst-case of the recourse, i.e., $\hat{\xi}^k \in \arg \max_{\xi \in \Xi^k} Q(\mathbf{x}, \xi)$ for any $\mathbf{x} \in X$, the convergence of the partition refinement problem by using the simplex-based heuristic method is guaranteed.

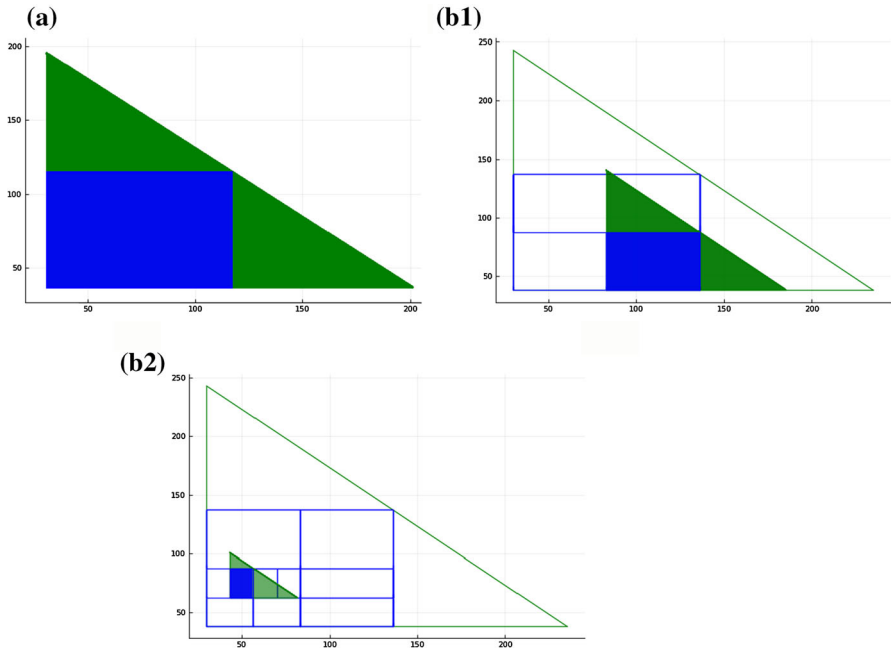


Fig. 2 (a) Extension of original support $[a_1, b_1] \times [a_2, b_2]$ for newsvendor problem for two items; (b1) and (b2) The extension of the chosen cell (highlighted in blue) in the refinement procedure is the same as that of original support for different partition sizes

Proposition 1 *Let $Q(\mathbf{x}, \xi_i)$ be a monotonic function of ξ_i , for all $i = 1, \dots, d_\xi$. Assuming that, for any \mathbf{x} , $\hat{\xi}^k \in \arg \max_{\xi \in \Xi^k} Q(\mathbf{x}, \xi)$ generates the worst-case recourse cost given the cell Ξ^k . Then, the sequences $\{\tilde{z}_n^U\}_{n=1}^\infty$ and $\{\tilde{\mathbf{x}}_n^U\}_{n=1}^\infty$ converge to the optimal objective value and optimal solution of (1), respectively, where $\tilde{\mathbf{x}}_n^U$ is the optimal solution of (21).*

The proof of Proposition 1 is presented in the Appendix B.

Note that a reformulation similar to (17) can be obtained by just considering J^k as the number of vertexes of the circumscribed simplex. For a polyhedral set X , this reformulation is a linear programming problem whose complexity grows linearly with the uncertainty vector dimension.

4 Proposed conservative partition refining (CPR) method for two-stage stochastic programming

In this section, we present a new conservative solution framework for two-stage stochastic linear optimization problems that solves the upper-bound problem. The aim is to obtain a conservative solution that avoids disappointment, i.e., the objective function cost estimate is the upper limit for the actual expected cost. We propose a tractable reformulation, namely the deterministic equivalent model, for the upper-

bound problem and proof convergence of the proposed methodology. We developed a simple and efficient partition refinement algorithm based on the structure of the optimality gap of each iteration.

We start from a sequential procedure to split the uncertainty support to obtain a refined partition. For iteration n and partition \mathcal{P}_n , we solve the upper-bound problem (16) and obtain the conservative solution \mathbf{x}_n^U . Then, we solve the lower-bound problem for the same partition \mathcal{P}_n and compute the optimality GAP—the difference of the optimal values of the upper- and lower-bound problems. The proposed CPR method is outlined in the following pseudo-algorithm:

Algorithm 3 The CPR method

Input: $\mathcal{P}^1 = \{\Xi\}$, $\epsilon > 0$ stopping criteria.

Output: \mathbf{x}_n^U the conservative solution of the upper bound problem.

- 1: Solve (16) to determine the optimal solution \mathbf{x}_n^U and the optimal objective value z_n^U of the upper bound problem for the partition \mathcal{P}_n .
 - 2: Solve (5) to determine the optimal objective value z_n^L of the lower bound problem for the partition \mathcal{P}_n .
 - 3: **if** $\frac{z_n^U - z_n^L}{z_n^U} \leq \epsilon$ **then**
 - 4: stop and return \mathbf{x}_n^U as the optimal solution.
 - 5: **else**
 - 6: refine the partition \mathcal{P}_n to \mathcal{P}_{n+1} and return to 1
 - 7: **end if**
-

The convergence of the CPR method is ensured by the Theorem 1, whose proof is presented in the Appendix C. It states that the sequence of optimal solutions and objective values of the lower- and upper-bound problems converge to the optimal solution \mathbf{x}^* and the optimal objective value z^* of the two-stage stochastic optimization problem (1), respectively. Note that the proof of the convergence does not depend on the refinement-partition procedure.

Theorem 1

- 1.1 *The optimal objective value sequence $\{z_n^U\}_{n=1}^\infty$ corresponds to conservative solutions given by the upper bound problem, for a family of partitions $\{\mathcal{P}_n\}_{n=1}^\infty$ such that \mathcal{P}_{n+1} refines \mathcal{P}_n , is non-increasing.*
- 1.2 *The optimal objective value sequence $\{z_n^L\}_{n=1}^\infty$ corresponds to optimistic solutions given by the lower bound problem, for a family of partitions $\{\mathcal{P}_n\}_{n=1}^\infty$ such that \mathcal{P}_{n+1} refines \mathcal{P}_n , is non-decreasing.*
- 1.3 *We have that the sequences $\{z_n^U\}_{n=1}^\infty$ and $\{z_n^L\}_{n=1}^\infty$ are convergent, i.e., $z_n^L \rightarrow z^* \leftarrow z_n^U$, as $n \rightarrow \infty$. Also the sequence $\{x_n^U\}_{n=1}^\infty$ and $\{x_n^L\}_{n=1}^\infty$ converge to $\mathbf{x}^* \in \text{Argmin}_{\mathbf{x} \in X} f(\mathbf{x})$.*

It is worth mentioning that the upper bound sequence obtained by computing the objective function of the upper-bound problem at the optimistic solution defines a non-increasing sequence. This fact ensures the convergence of the existing partition-based method. However, the optimistic solution is sub-optimal for the upper-bound problem, so the existing partition-based method derives an upper bound less tight.

Next, we propose a simple and efficient partition refinement algorithm that defines the sequential upper- and lower-bounds and, consequently, generates the sequence of conservative solutions $\{x_n^U\}_{n=1}^\infty$.

4.1 Solution algorithm with worst-case partition refinement (SAWPR)

Any partition-based method is supported on the refinement procedure of the partition to improve the upper and lower approximations. Next, we detail the three basic steps of the proposed partition refining procedure as (i) selection of the cell Ξ_{k^*} to be split, (ii) selection of the uncertainty direction $i^* \in \{1, \dots, d_\xi\}$ to refine the partition, and (iii) selection of the cutting point.

In the first step of the partition refinement procedure, we aim to select the cell with the highest contribution to the current optimality gap, which is composed of the difference of upper and lower approximations of the expected recourse function. Given that the partition influences the gap through the expected recourse function, we select the cell with the maximum contribution to the difference of upper and lower approximations of the expected recourse function. Let us define the selected cell as

$$k^* \in \arg \max_k \left\{ p^k \left(\text{Max}_{\xi \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}_n^U, \mathbf{e}^k)] - Q(\mathbf{x}_n^L, \bar{\xi}^k) \right) \right\}. \tag{22}$$

The second step of the partition refinement procedure is the selection of the uncertainty direction. Most existing partition-based methods consider the direction with the highest metric of non-linearity of $Q(\mathbf{x}, \cdot)$ as the optimal direction to split the optimal cell to refine the partition. This is motivated by the fact the lower and upper approximations coincide for an affine function. Initially, [5] proposed to use dual (subgradient) information at the endpoints of the cell, while [12] compares the difference between the upper and lower approximations. However, the use of dual information increases the computational burden of the sequential partition refinement problem.

In this work, we selected the uncertainty direction by solving an optimization problem that resembles the robust optimization model with an uncertainty budget proposed by [4]. For the selected cell k^* , we formulated an adversary problem in (23) that aimed to find the uncertainty realization with the highest cost for the optimistic solution given by the lower-bound problem. The intent was to select the dimension where the conditional mean was less representative of the entire conditional distribution adapted to that cell. We imposed a unitary budget constraint, i.e., only one component of the random vector is allowed to change around its nominal value (conditional mean). Hence, we selected the uncertainty direction i^* such that $\hat{\xi}_{i^*} \neq \bar{\xi}_{i^*}$, where

$$\begin{aligned} \hat{\xi} &\in \arg \max_{\xi, \mathbf{z}} Q(\mathbf{x}_n^L, \xi) \\ \text{s.t.} \quad &\bar{\xi}_i^k - z_i(\bar{\xi}_i^k - a_i^k) \leq \xi_i \quad \forall i, \\ &\xi_i \leq \bar{\xi}_i^k + z_i(b_i^k - \bar{\xi}_i^k) \quad \forall i, \end{aligned}$$

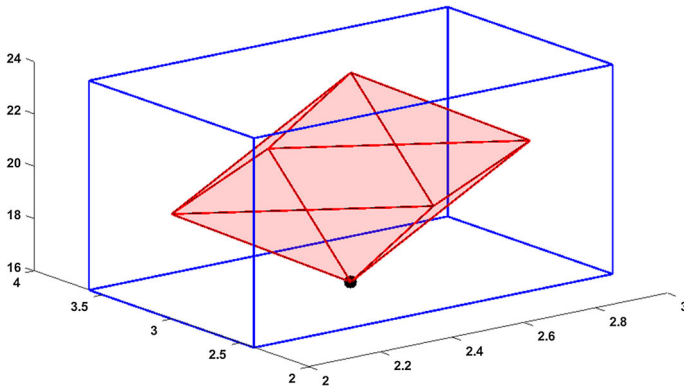


Fig. 3 Ambiguity set representing the feasible region of the problem (23) for the case $d_\xi = 3$

$$\sum_{i=1}^{d_\xi} z_i \leq 1, \quad 0 \leq z_i \leq 1 \quad \forall i. \tag{23}$$

Since (23) is a maximization problem with a convex objective function and the feasible set is a polyhedron, there is a vertex that is optimal. By construction, the number of vertices of the feasible polyhedron is $2 d_\xi$, i.e., grows linearly on the uncertainty dimension. Thus, to efficiently solve (23) it suffices to enumerate the vertices and cast the one with the highest recourse cost.

Note also that the vertices are intuitive and easily identifiable since they are defined as the conditional mean $\bar{\xi}^k$ projected on the faces of the hypercube Ξ^k . To illustrate this concept, let us consider $\Xi = \times_{i=1}^3 [a_i^k, b_i^k]$ (see Fig. 3) and a given optimistic solution \mathbf{x}_n^L . A unitary uncertainty budget leads to a “diamond” shaped polyhedral inscribed in the hypercube Ξ^k . In other words, the vertices of the feasible set differ from the center (conditional mean $\bar{\xi}^k$) in just one component i , which can assume any extreme value $\{a_i^k, b_i^k\}$. Therefore, the number of vertexes is $2 d_\xi$.

The third and last step of the partition refinement procedure is the selection of the cutting point. For simplicity and computational efficiency, we assumed that the cutting point is the component of the conditional mean $\bar{\xi}_{i^*}$ along to the uncertainty direction i^* the same as [5,16].

Finally, we summarize the proposed partition refinement procedure, namely, solution algorithm with worst-case partition refinement (SAWPR) as follows:

In Fig. 4 we present the refinement procedure for the farmer’s problem instance for three types of crops as an illustrative example of the SAWPR algorithm. Note that the cells of the partition are clustered around the region of the uncertainty support corresponding to scenarios of less land productivity which represents in particular, the worst-case for this instance.

Algorithm 4 SAWPR

Input: \mathbf{x}^* optimal solution of (14)

Output: Ξ^{k^*} and Ξ^{n+1}

1: Determine k^* the optimal cell to be split as the cell with the largest error contribution given by (22).

2: Find i^* as the component $\hat{\xi}_{i^*}$ of the optimal solution $\hat{\xi}$ of (23) such that $\hat{\xi}_{i^*} \neq \bar{\xi}_{i^*}^{k^*}$.

$$3: \Xi^{k^*} \leftarrow [a_{i^*}^{k^*}, \bar{\xi}_{i^*}^{k^*}] \times \prod_{\{1,2,\dots,d_\xi\} \setminus \{i^*\}} [a_i^{k^*}, b_i^{k^*}]$$

$$4: \Xi^{n+1} \leftarrow [\bar{\xi}_{i^*}^{k^*}, b_{i^*}^{k^*}] \times \prod_{\{1,2,\dots,d_\xi\} \setminus \{i^*\}} [a_i^{k^*}, b_i^{k^*}]$$

5 Empirical study

In this section, we present an empirical study to test the proposed CPR method. As we mentioned above (see Sect. 2.1), the partition refinement method depends on the computation of the probability mass p^k and the conditional mean $\bar{\xi}^k$ of the cell k , $k = 1, \dots, n$. For a general probability distribution, this computation is not easy. We develop an empirical study method based on the uniform probability distribution since in this case, we have an analytic solution for the conditional probability and expectation. However, to cover a more general class of distributions we provide a high-level guideline on how to deal with distributions without analytic solutions for the conditional probability and expectations.

5.1 Implementation remark

We argue that it is possible to efficiently compute the probability mass p^k and the conditional expectation $\bar{\xi}^k$ of the cell $k = 1, \dots, N$, according to the following cases:

- **Independent random variables:** Suppose that the coordinates of the random vector ξ are continuous independent random variables. Let φ the probability density function of the random vector ξ . Then, by independence, φ can be written as the product of the marginals φ_i of the coordinates ξ_i , $i = 1, \dots, d_\xi$. Thereby,

$$p^k = P(\Xi^k) = \int_{\Xi^k} \varphi(\xi) d\xi = \int_{a_1}^{b_1} \varphi(\xi_1) d\xi_1 \cdots \int_{a_{d_\xi}}^{b_{d_\xi}} \varphi_{d_\xi}(\xi_{d_\xi}) d\xi_{d_\xi},$$

and

$$\begin{aligned} \bar{\xi}^k &= \mathbb{E}[\xi \mid \xi \in \Xi^k] = \int_{\Xi^k} \xi \varphi(\xi \mid \xi \in \Xi^k) d\xi \\ &= \int_{\Xi^k} \xi \left[\frac{\varphi(\xi)}{p^k} \right] d\xi = \frac{1}{p^k} \int_{a_1}^{b_1} \xi_1 \varphi(\xi_1) d\xi_1 \cdots \int_{a_{d_\xi}}^{b_{d_\xi}} \xi_{d_\xi} \varphi_{d_\xi}(\xi_{d_\xi}) d\xi_{d_\xi}, \end{aligned}$$

SAWPR for the farmer's Problem

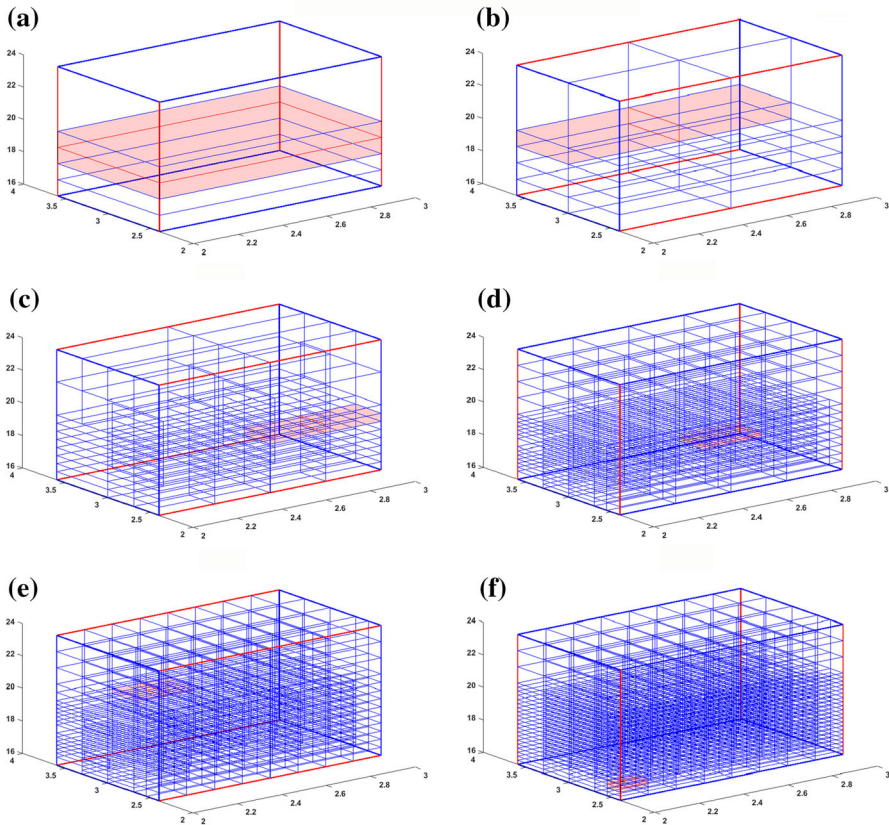


Fig. 4 Partition of the support $\Xi = \times_{i=1}^3 [a_i^k, b_i^k]$ with: **a** 4 cells, **b** 20 cells, **c** 100 cells, **d** 300 cells, **e** 500 cells, and **f** 900 cells. The shaded region shows the chosen cell and the red edges show the direction chosen to refine the partition

i.e., the calculation of the probability mass p^k and the conditional expectation $\bar{\xi}^k$ derives in the computation of sequence univariate numerical integrals, which is computationally tractable (see [20,22]).

- **Linearly dependent random variables:** Suppose that ξ can be expressed as $\xi = \mathbf{L}\zeta$, where $\mathbf{L} \in \mathbb{R}^{d_\xi \times d_\xi}$ is a non-singular matrix and $\zeta \in \mathbb{R}^{d_\xi}$ is an independent random vector. Then, we can calculate the probability mass p^k and the conditional mean $\bar{\xi}^k$ as in the previous case. To see that, let us denote by $\Xi^k = [\mathbf{a}^k, \mathbf{b}^k]$ the hypercube for the cell $k, k = 1, \dots, n$, where $\mathbf{a}^k = [a_1^k, \dots, a_{d_\xi}^k]$ and $\mathbf{b}^k = [b_1^k, \dots, b_{d_\xi}^k]$. Then,

$$\begin{aligned}
 p^k &= P(\Xi^k) = P(\mathbf{a}^k \leq \xi \leq \mathbf{b}^k) = P(\mathbf{a}^k \leq \mathbf{L}\zeta \leq \mathbf{b}^k) = P(\mathbf{L}^{-1}\mathbf{a}^k \leq \zeta \leq \mathbf{L}^{-1}\mathbf{b}^k) \\
 &= P(\zeta \in \tilde{\Xi}^k),
 \end{aligned}$$

where $\tilde{\Xi}^k$ is the hypercube defined by $\tilde{\Xi}^k = [\mathbf{L}^{-1}\mathbf{a}^k, \mathbf{L}^{-1}\mathbf{b}^k]$, and

$$\bar{\xi}^k = \mathbb{E}[\xi \mid \xi \in \Xi^k] = \mathbb{E}[\mathbf{L}\zeta \mid \mathbf{L}\zeta \in \Xi^k] = \mathbf{L}\mathbb{E}[\zeta \mid \zeta \in \tilde{\Xi}^k].$$

Therefore, the computation of the conditional probability $P(\tilde{\Xi}^k)$ and the conditional expectation $\mathbb{E}[\zeta \mid \zeta \in \tilde{\Xi}^k]$ reduces to the independent case. As example, suppose that ξ is a multivariate normal variable. Let $\Sigma \in \mathbb{R}^{d_\xi \times d_\xi}$ the covariance matrix of ξ , and $\Sigma = \mathbf{L}\mathbf{L}^\top$ its Cholesky decomposition. Then, there exists an independent random vector ζ such that $\xi = \mathbf{L}\zeta$.

- **Generically dependent random variables case:** In this case, we can assess the computation of the probability mass p^k and the conditional mean $\bar{\xi}^k$ by the importance sampling method (see [2]) via numerical simulation. Let us denote by

$$U(\xi; [\mathbf{a}^k, \mathbf{b}^k]) = \begin{cases} \prod_{i=1}^{d_\xi} \frac{1}{b_i^k - a_i^k}, & \xi \in [\mathbf{a}^k, \mathbf{b}^k], \\ 0, & \text{otherwise,} \end{cases}$$

the probability density function of the multivariate uniform distribution with support $\Xi^k = [\mathbf{a}^k, \mathbf{b}^k]$. Thereby, we can consider the change of probability measure

$$p^k = \int_{\Xi^k} \varphi(\xi) d\xi = \int_{\Xi^k} \left[\frac{\varphi(\xi)}{U(\xi; [\mathbf{a}^k, \mathbf{b}^k])} \right] U(\xi; [\mathbf{a}^k, \mathbf{b}^k]) d\xi,$$

and

$$\begin{aligned} \bar{\xi}^k &= \int_{\Xi} \xi \varphi(\xi \mid \xi \in \Xi^k) d\xi = \int_{\Xi^k} \xi \left[\frac{\varphi(\xi)}{p^k} \right] d\xi \\ &= \int_{\Xi^k} \xi \left[\frac{\varphi(\xi)}{p^k U(\xi; [\mathbf{a}^k, \mathbf{b}^k])} \right] U(\xi; [\mathbf{a}^k, \mathbf{b}^k]) d\xi. \end{aligned}$$

Therefore, according to the central limit theorem,

$$\frac{1}{S} \sum_{s=1}^S \left[\frac{\varphi(\hat{\xi}_s)}{U(\hat{\xi}_s; [\mathbf{a}^k, \mathbf{b}^k])} \right] \xrightarrow{P} p_k,$$

and

$$\frac{1}{S} \sum_{s=1}^S \hat{\xi}_s \left[\frac{\varphi(\hat{\xi}_s)}{p^k U(\hat{\xi}_s; [\mathbf{a}^k, \mathbf{b}^k])} \right] \xrightarrow{P} \bar{\xi}^k,$$

where $\hat{\xi}_s, \forall s = 1, \dots, S$ are independent random samples of the uniform distribution with density defined by $U(\xi; [\mathbf{a}^k, \mathbf{b}^k])$. Based on standard statistical results, for a given error precision and confidence level, one can define a sample size S and compute the above estimates to be used in our algorithms as proxies for the conditional probability p_k and expectation $\bar{\xi}^k$.

5.2 Computational experiments

We test the CPR method based on a study case for each uncertainty dimensionality category: low, medium, and high. We assessed the aircraft allocation problem [7] (see Appendix D) by the deterministic equivalent reformulation of the upper-bound problem. For medium-sized uncertainty dimensionality, we assessed an instance of the farmer's problem [5] (see Appendix E) with eight cultivated crops. For high uncertainty dimensionality, we assessed an instance of the farmer's problem with twenty cultivated crops and an instance of the Stochastic Unit Commitment (SUC) problem for a small single-bus system with five generators over a 24-hour time span (see Appendix F). The medium-sized uncertainty instance (8-crops) was assessed by the column and constraint generation method whereas the high uncertainty dimensionality instances (20-crops and SUC) were assessed by the proposed simplex-based heuristic method of the extension of the uncertainty support. This empirical study was made to evidence the disappointment generated by both conservative and optimistic solutions to raise awareness on the negative impact of optimistic solution methods to solve two-stage stochastic optimization problems, it does not intend to compare the proposed approach with other solution methodologies.

For the first application, different types of aircraft must be allocated to a certain route for the transportation of passengers. The number of allocated aircraft is the first-stage decision, and the recourse of the problem is defined by the number of bumped passengers when demand for seats outstrips capacity. The right-hand side uncertainty corresponds to the unknown demand of passengers modeled by a uniform probability distribution. For this problem, the dimension of the uncertainty is 5, the number of first-stage variables is 21, and the number of second stage-variables is 10 for each uncertainty realization.

Regarding the second application, the farmer problem is an example of a production model under uncertainty where the first-stage decisions correspond to the land allocation destined to rise different types of crops, and the recourse consists in trading the cultivated products in the local market to satisfy a given demand. In this problem, we assume a uniform probability distribution to model the uncertainty in the land productivity for growing each crop. The number of first-stage variables is equal to the uncertainty dimension, since they correspond to the land allocation to cultivate each type of crop while the number of second-stage variables for each uncertainty realization is twice the quantity of cultivated products because they correspond to the sold and buy quantities of each crop in the local market.

To generate different instances of the farmer problem, we varied the uncertainty dimension considering eight and 20 types of crops to create a computational experiment for each of the following situations:

- It is impossible to solve in practical time the deterministic equivalent linear model for the enumerative case considering the 2^{d_ξ} vertexes of the hypercubes of the rectangular partition, but it is possible to handle the medium-sized uncertainty dimensionality with the proposed column and constraint generation algorithm.

- It is impossible to obtain a conservative solution solving the re-formulated upper-bound problem using the column and constraint generation algorithm, and it is necessary to appeal to a heuristic solution method.

For both applications, the refinement algorithm and the algorithm for the sequential partition-based method is implemented in JuMP [8], a modeling language for mathematical optimization embedded in the Julia programming language and Gurobi was used as the linear optimization solver to run the computational experiments on a Intel Core i7, 4.0-GHz processor with 32 GB of RAM. Moreover, we have considered a uniform probability distribution to alleviate the computational burden of the CPR method in the computation of the probability mass p^k and the conditional mean $\bar{\xi}^k$ of the cell $k, k = 1, \dots, N$. For a general probability distribution, this computation can be assessed via numerical simulation (see 2.1).

5.3 Results

In general, the optimal solutions of the upper-bound problem of both the CPR method solved by the column and constraint generation algorithm and the simplex-based heuristic method of the extension of the uncertainty support represents a conservative decision policy. Indeed, these solutions are obtained from the approximation of the conditional expectation $\mathbb{E}[Q(\mathbf{x}, \xi) | \xi \in \xi^k]$ by the worst distributionally expectation

$$\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^k)]$$

and

$$\text{Max}_{\delta \in \mathcal{R}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{v}^k)],$$

considering the marginal distribution of the random vectors \mathbf{e}^k and \mathbf{v}^k with support in the set $\{\mathbf{e}_j^k : j = 1, \dots, 2^{d_\xi}\}$ and $\{\mathbf{v}_j^k : j = 1, \dots, d_\xi + 1\}$, respectively, for all $k = 1, \dots, n$.

As mentioned above, the existing partition-based methods do not solve the upper-bound problem, instead, it determines the optimistic solution \mathbf{x}_n^L given by the lower-bound problem. Nevertheless, depending on the recourse cost, sometimes it is necessary to obtain a conservative solution because the optimistic solution generates a significant disappointment. To study the disappointment of a given conservative and optimistic solution in an out-of-sample analysis, we performed a Monte Carlo simulation with a number N of scenarios to estimate the actual cost by the sample-average cost [3].

For the aircraft allocation problem, we consider two instances: (a) low recourse cost; (b) high recourse cost. For the low recourse cost—low cost for the second-stage variables, we consider a negligible cost for unattended demand. For the high recourse cost, we consider a significantly high deficit cost associated with unattended demand. Given these two instances, we depict in Fig. 5 the disappointment by confronting the out-of-

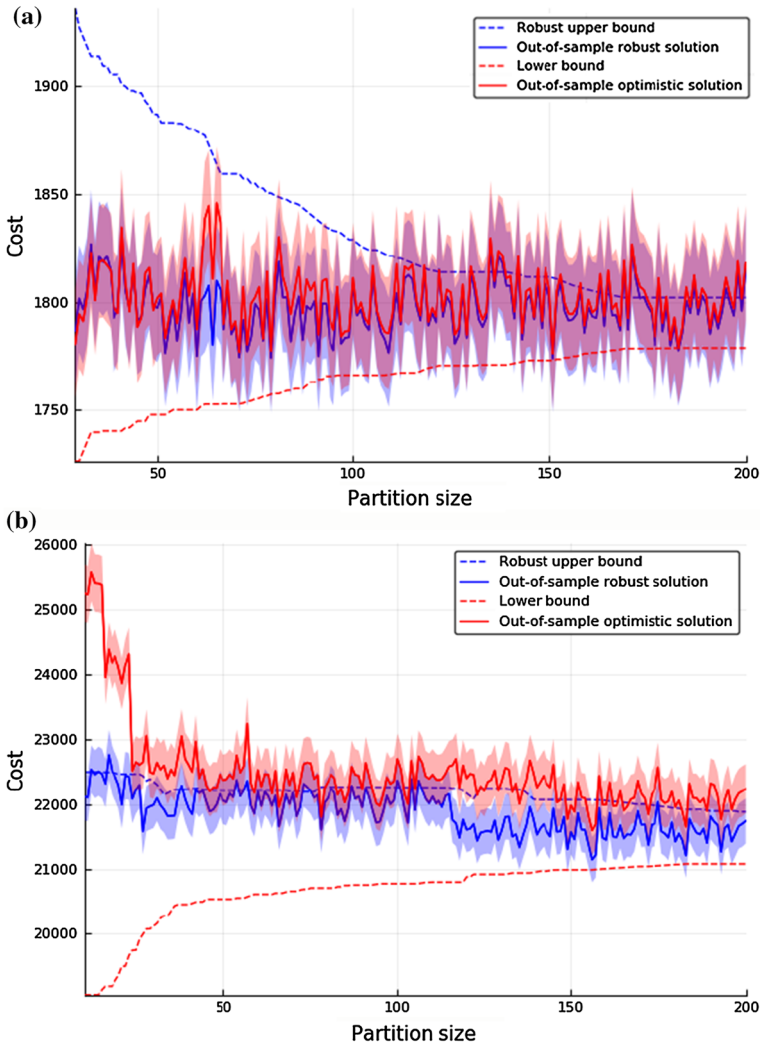


Fig. 5 Disappointment in an out-of-sample analysis for the aircraft allocation problem: **a** low recourse cost; **b** high recourse cost

sample cost against the cost estimate given by objective value of the upper (conservative) and lower (optimistic) bound problems. For didactic purposes, the results associated with the optimistic solution are presented in red while the results for the conservative solution are presented in blue. The horizontal axis of the figure reports the computational time to obtain the conservative and the optimistic solution, given a partition size. The cost estimated by the optimal value of the upper- and lower-bound problem (optimal objective value) is represented by the dashed line. The solid line corresponds to the out-of-sample cost evaluation and the shaded area corresponds to the associated 95% confidence interval. It is worth clarifying that the computational time reported in the horizontal axis does not include the time to compute the out-of-sample cost.

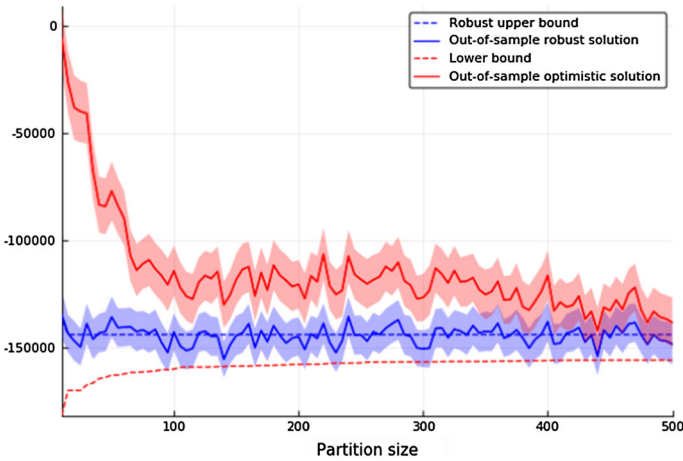


Fig. 6 Disappointment in an out-of-sample analysis of the conservative solution and optimistic solution of the farmer's problem with eight crops

For the low recourse cost instance, Fig. 5a, we observe that both (conservative and optimistic) solutions have similar out-of-sample performances. However, for the high recourse cost instance, Fig. 5b, the conservative solution out-performs the optimistic one for any partition size. Note also that, for both instances, the optimistic solution leads to significant disappointment—the out-of-sample cost evaluation is significantly higher than the cost estimated by the optimal value of the lower bound problem. On the other hand, the conservative solution has a mathematical guarantee against disappointment, which is corroborated by this empirical study—there is no statistically significant disappointment. This means that optimistic solution methods might not be suitable for applications with high recourse cost due to poor out-of-sample performance and potentially high disappointment levels.

In Fig. 6, we present the numerical results for the farmer's problem with eight crops considering a high recourse cost instance. The out-of-sample evaluation of the conservative solution shows a very similar behavior when compared to the optimal objective value of the upper-bound problem. Conversely, the out-of-sample cost evaluation for the optimistic solution is significantly higher than the optimal objective value, i.e., it shows a high disappointment level. Moreover, the conservative solution significantly out-performs the optimistic one in an out-of-sample analysis. It is important to note that the above mentioned effects are amplified whenever the algorithm is far from converging.

This result alerts us to use of the optimistic solution for two-stage stochastic linear programming problems when the recourse cost is significantly high. If the method stops before convergence—due to time or computational-power limitations—the optimistic solution is not reliable given its poor out-of-sample performance and high disappointment level. On the other hand, the conservative decision is robust even for non-converged solutions.

Finally, Fig. 7a presents the numerical results for the farmer's problem with 20-crops, and Fig. 7b reports the result for the instance of the SUC problem, both obtained by the simplex-based heuristic method for highly asymmetric recourse cost,

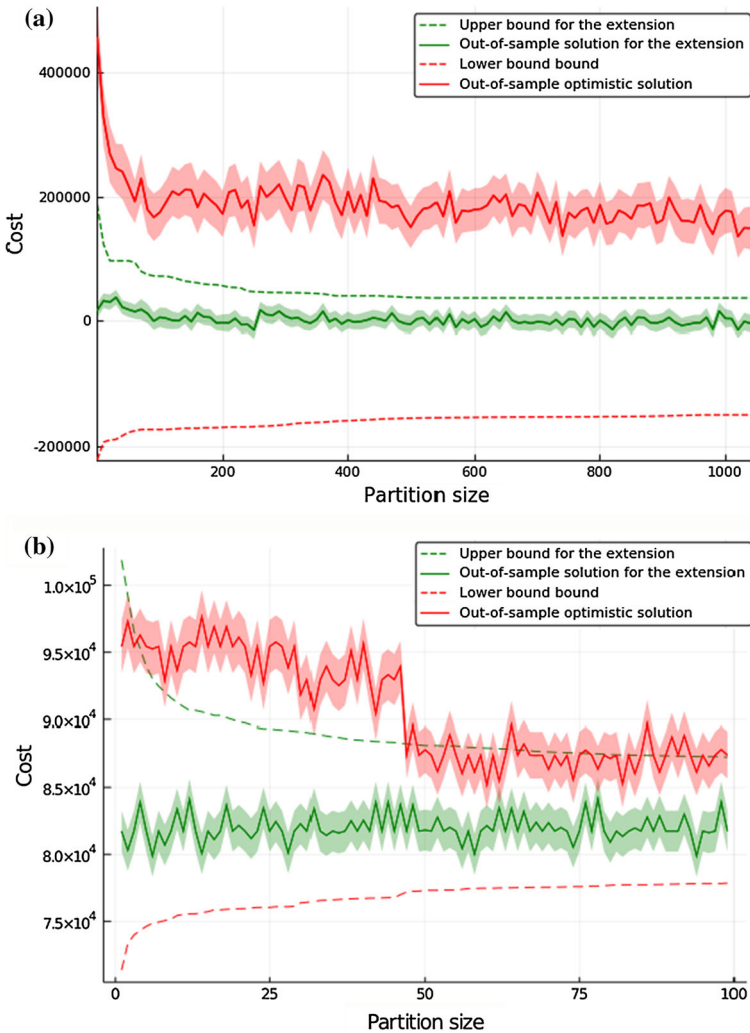


Fig. 7 Upper bound and lower bound obtained by the simplex-based heuristic method of the extension of the uncertainty support for **a** a farmer’s problem with 20-crops; **b** the stochastic unit commitment problem for a single-bus system over a 24-hour time span

which means that the corrective actions are much more expensive in one direction—unattended demand is more expensive than over-production (farmer’s problem) or over-generation (SUC problem). As in the upper-bound problem (16), solved by the column and constraint generation algorithm, the solution given by the simplex-based heuristic method (21) also avoids disappointments and significantly outperforms the optimistic solution. As before, the non-converged optimistic solution presented a significant disappointment and poor out-of-sample performance.

Table 1 displays the uncertainty dimension d_ξ , the size (# variables, # constraints) of the upper-bound problem and the computational time to solve the CPR method by applying the corresponding numerical scheme to solve each instance.

Table 1 Computational time of the proposed algorithmic schemes for each instance

Instance	$d\xi$	Size of the UB problem	Method	Time (CPUs)
Aircraft allocation problem (low cost)	5	(4064, 26496)	Equivalent deterministic	1380
Aircraft allocation problem (high cost)	5	(4064, 26496)	Equivalent deterministic	1363
Farmer's problem 8-crops	8	(6144, 2304)	Column–Constraint algorithm	32,541
Farmer's problem 20-crops	20	(62914560, 22020096)	Simplex-based method	22,278
SUC	24	$(8.05 \times 10^9, 1.77 \times 10^{10})$	Simplex-based method	32,796

6 Conclusions

We developed a conservative convergent solution method for a two-stage stochastic linear optimization problem with a convex recourse function of uncertainty based on the distributionally robust optimization model that generalizes the Edmundson–Madanski inequality. Considering that the complexity grows exponentially over an uncertainty dimension, for computational tractability, we reformulated the upper-bound problem and proposed algorithmic schemes: (i) for problems with low-dimensional uncertainty, we developed a deterministic equivalent linear programming model, (ii) for medium-sized uncertainty dimensionality, we proposed a column and constraint generation algorithm, and (iii) to handle high dimensional uncertainty, we proposed a simplex-based heuristic method whose complexity grew linearly with the uncertainty dimension.

An out-of-sample computational experiments showed that our method avoids disappointment in comparison to the non-converged sub-optimal optimistic solution given by the lower-bound problem based on Jensen's inequality when the cost of recourse was high. This raises awareness regarding the use of the optimistic solution provided by the partition-based method to solve two-stage stochastic optimization problems when the recourse cost is significantly high. Many practical applications that exhibit this type of recourse cost and depend on the implementation of a non-converged sub-optimal solution because of computational time limitations will benefit from the developed framework to obtain a conservative solution.

7 Appendix

A Oracle MIP problem

Let $\tau \in \{0, 1\}^{m_y}$ and $\kappa \in \{0, 1\}^{m_y \times d_x}$ be a binary vector and a binary matrix, respectively. We define $\mathbf{h}(\mathbf{e}_j^k) := [(1 - \tau_r)\hat{a}_r^k + \tau_r \hat{b}_r^k]$, $\mathbf{T}(\mathbf{e}_j^k) := [(1 - \kappa_{r,s})\tilde{a}_{r,s}^k + \kappa_{r,s} \tilde{b}_{r,s}^k]$ and $\mathbf{W} := [W_{r,t}]$ with $r = 1, \dots, m_y$, $s = 1, \dots, d_x$ and $t = 1, \dots, d_y$.

We write explicitly (19) as follows:

$$\begin{aligned}
 & \text{Max}_{\lambda, \alpha \in \mathbb{R}^{m_y}, \beta \in \mathbb{R}^{m_y \times m_x}, \tau, \kappa} \left\{ \sum_{r=1}^{m_y} \underbrace{[\lambda_r \tau_r \hat{b}_r^k + (\lambda_r - \lambda_r \tau_r) \hat{a}_r^k]}_{\alpha_r} - \left(\sum_{r=1}^{m_y} \sum_{s=1}^{d_x} \underbrace{\lambda_r \kappa_{r,s} \tilde{b}_{r,s}^k x_s^*}_{\beta_{r,s}} + \sum_{r=1}^{m_y} \sum_{s=1}^{d_x} \underbrace{(\lambda_r - \lambda_r \kappa_{r,s}) \tilde{a}_{r,s}^k x_s^*}_{\beta_{r,s}} \right) \right\} \\
 & \text{s.t. } \sum_{r=1}^{m_y} \lambda_r W_{r,t} \leq q_t, \quad \forall t = 1, \dots, d_y \\
 & \lambda_r - (1 - \tau_r)M \leq \alpha_r \leq \lambda_r + (1 - \tau_r)M, \quad \forall r = 1, \dots, m_y \\
 & -\tau_r M \leq \alpha_r \leq \tau_r M, \quad \forall r = 1, \dots, m_y \\
 & \lambda_r - (1 - \kappa_{r,s})M \leq \beta_{r,s} \leq \lambda_r + (1 - \kappa_{r,s})M, \quad \forall r = 1, \dots, m_y, \forall s = 1, \dots, d_x \\
 & -\kappa_{r,s} M \leq \beta_{r,s} \leq \kappa_{r,s} M, \quad \forall r = 1, \dots, m_y, \forall s = 1, \dots, d_x \\
 & \tau_r, \kappa_{r,s} \in \{0, 1\}, \quad \forall r = 1, \dots, m_y, \forall s = 1, \dots, d_x.
 \end{aligned} \tag{24}$$

where $M \in \mathbb{R}$ is sufficiently large and the products $\alpha_r = \lambda_r \tau_r$ and $\beta_{r,s} = \lambda_r \kappa_{r,s}$ are linearized.

B Proof of the Proposition 1

Proof of the Proposition 1 By duality we have that

$$\text{Max}_{\delta \in \mathcal{R}(\hat{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{v}^k)] = \text{Max}_{\delta^k \geq 0} \left\{ \sum_{j=1}^{d_\xi+1} \delta_j^k Q(\mathbf{x}, \mathbf{v}_j^k) : \sum_{j=1}^{d_\xi+1} \delta_j^k \mathbf{v}_j^k = \hat{\xi}^k, \sum_{j=1}^{d_\xi+1} \delta_j^k = 1 \right\}.$$

Since the recourse function $Q(\mathbf{x}, \xi_i)$ is monotonic for all $i = 1, \dots, d_\xi$ we have that

$$Q(\mathbf{x}, \mathbf{v}_j^k) \leq Q(\mathbf{x}, \hat{\xi}^k), \quad \forall j = 1, \dots, (d_\xi + 1),$$

since \mathbf{v}_j^k only differs from $\hat{\xi}^k$ in just one component. So, $\sum_{j=1}^{d_\xi+1} \delta_j^k Q(\mathbf{x}, \mathbf{v}_j^k) \leq Q(\mathbf{x}, \hat{\xi}^k)$ for all $\mathbf{x} \in X$. By other hand, by the existence of $\mathbb{E}[Q(\mathbf{x}, \xi)]$, it holds that

$$\sum_{k=1}^n p^k Q(\mathbf{x}, \hat{\xi}^k) \longrightarrow \mathbb{E}[Q(\mathbf{x}, \xi)] \quad \text{as } n \longrightarrow \infty,$$

therefore

$$\sum_{k=1}^n p^k \sum_{j=1}^{d_\xi+1} \delta_j^k Q(\mathbf{x}, \mathbf{v}_j^k) \longrightarrow \mathbb{E}[Q(\mathbf{x}, \xi)] \quad \text{as } n \longrightarrow \infty, \quad \forall \mathbf{x} \in X.$$

Since $\tilde{f}_n(\mathbf{x}) := \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^n p^k \sum_{j=1}^{d_\xi+1} \delta_j^k Q(\mathbf{x}, \mathbf{v}_j^k)$ is a sequence of decreasing and continuous functions pointwise converging to $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x} + \mathbb{E}[Q(\mathbf{x}, \xi)]$, then $\tilde{f}_n(\mathbf{x})$ epi-converge to $f(\mathbf{x})$. Therefore, by epi-convergence, $\{\tilde{z}_n^U\}_{n=1}^\infty$ and $\{\tilde{\mathbf{x}}_n^U\}_{n=1}^\infty$ converge to the optimal value and optimal solution of (1), respectively. \square

C Proof of the Theorem 1

If $\{f_n\}_{n=1}^\infty$ is a sequence of functions pointwise converging to f , i.e., if $f(\mathbf{x}) = \lim_{n \rightarrow \infty} f_n(\mathbf{x})$ for all $\mathbf{x} \in X$, then f_n epi-converges to f , if the sequence $\{f_n\}_{n=1}^\infty$ is monotone increasing or monotone decreasing [6] and f is continuous.

Epi-convergence is a kind of convergence very useful for approximate minimization problems in the following sense:

Suppose a sequence of function $\{f_n\}_{n=1}^\infty$ epi-converges to f . Let $z^* := \text{Min}_{\mathbf{x} \in X} f(\mathbf{x})$, $z_n^* := \text{Min}_{\mathbf{x} \in X} f_n(\mathbf{x})$, and $\mathbf{x}^* \in \text{Argmin}_{\mathbf{x} \in X} f(\mathbf{x})$, $\mathbf{x}_n^* \in \text{Argmin}_{\mathbf{x} \in X} f_n(\mathbf{x})$ the corresponding minimizers, respectively. Then, under some assumptions, we can ensure that

$$\lim_{n \rightarrow \infty} z_n^* = z^*, \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbf{x}_n^* = \mathbf{x}^*.$$

Before proving the Theorem 1, we will need the following lemma to complete the proof.

Lemma 1 *let $G = \{(\xi, \eta) \in \mathbb{R}^{d_\xi+1} : \xi \in \Xi, \eta = Q(\mathbf{x}, \xi)\}$ be the graph of the function $Q(\mathbf{x}, \cdot)$ and let $\text{co}(G)$ be its convex hull. Then*

$$\sup\{\eta \in \mathbb{R} : (\bar{\xi}, \eta) \in \text{co}(G)\} = \text{Max}_{\delta \in \mathcal{D}(\bar{\xi})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e})] \tag{25}$$

Proof of Lemma 1 It is clear that

$$\text{Max}_{\delta \in \mathcal{D}(\bar{\xi})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e})] \leq \sup\{\eta \in \mathbb{R} : (\bar{\xi}, \eta) \in \text{co}(G)\}.$$

By another hand, let $(\bar{\xi}, \eta) \in \text{co}(G)$. Then there exist $S, \xi^s \in \Xi$, and probabilities $\mathbb{P}(\xi = \xi^s) \geq 0, s = 1, \dots, S$ (parameters of the convex combination), such that $\sum_{s=1}^S \mathbb{P}(\xi = \xi^s) \xi^s = \bar{\xi}, \sum_{s=1}^S \mathbb{P}(\xi = \xi^s) Q(\mathbf{x}, \xi^s) = \eta$. Now for every s , there exist conditional probabilities $\mathbb{P}(\mathbf{e} = \mathbf{e}_j | \xi = \xi^s) \geq 0, j = 1, \dots, 2^{d_\xi}$ such that $\sum_{j=1}^{2^{d_\xi}} \mathbb{P}(\mathbf{e} = \mathbf{e}_j | \xi = \xi^s) \mathbf{e}_j = \xi^s$, i.e., $\hat{\xi} = \sum_s \sum_j \mathbb{P}(\xi = \xi^s) \mathbb{P}(\mathbf{e} = \mathbf{e}_j | \xi = \xi^s) \mathbf{e}_j$. Since $Q(\mathbf{x}, \cdot)$ is convex

$$\begin{aligned} \eta &= \sum_{s=1}^S \mathbb{P}(\xi = \xi^s) Q(\mathbf{x}, \xi^s) \\ &= \sum_s \mathbb{P}(\xi = \xi^s) Q \left(\sum_j \mathbb{P}(\mathbf{e} = \mathbf{e}_j | \xi = \xi^s) \mathbf{e}_j \right) \\ &\leq \sum_s \sum_j \mathbb{P}(\xi = \xi^s) \mathbb{P}(\mathbf{e} = \mathbf{e}_j | \xi = \xi^s) Q(\mathbf{x}, \mathbf{e}_j). \end{aligned}$$

Setting $\delta_j = \sum_s \mathbb{P}(\xi = \xi^s) \mathbb{P}(\mathbf{e} = \mathbf{e}_j | \xi = \xi^s)$ yields $\bar{\xi} = \sum_j \delta_j \mathbf{e}_j$, $\sum_j \delta_j = 1$, $\delta_j \geq 0$, and

$$\eta \leq \sum_j \delta_j Q(\mathbf{x}, \mathbf{e}_j) \leq \text{Max}_{\delta \in \mathcal{D}(\bar{\xi})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e})].$$

Therefore

$$\sup\{\eta \in \mathbb{R} : (\bar{\xi}, \eta) \in \text{co}(G)\} \leq \text{Max}_{\delta \in \mathcal{D}(\bar{\xi})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e})].$$

□

By making use of the properties of epi-convergent functions and the lemma above, we will prove the Theorem 1.

Proof of Theorem 1

1.1 Let $\Xi' \subseteq \Xi$, let G' be the graph of the function $Q(\mathbf{x}, \xi)$ with $\xi \in \Xi'$ and let $\bar{\xi}' = \mathbb{E}[\xi | \xi \in \Xi']$ be the conditional mean, then

$$\sup\{\eta : (\bar{\xi}, \eta) \in \text{co}(G)\} \geq \sup\{\eta : (\bar{\xi}', \eta) \in \text{co}(G')\}$$

which implies

$$\text{Max}_{\delta \in \mathcal{D}(\bar{\xi})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e})] \geq \text{Max}_{\delta \in \mathcal{D}(\bar{\xi}')} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}')],$$

where \mathbf{e}' is the random variable with support on the vertex of the cell Ξ' . Suppose that \mathcal{P}_{n+1} refines \mathcal{P}_n , then there exist $\Xi^k \in \mathcal{P}_n$ such that $\Xi^k = \Xi^{k'} \cup \Xi^{k''}$ with $\Xi^{k'}, \Xi^{k''} \in \mathcal{P}_{n+1}$.

Let $\mathbf{e}^k, \mathbf{e}^{k'}$, and $\mathbf{e}^{k''}$ be the random variables with support in the set of vertex of the cells $\Xi^k, \Xi^{k'}$ and $\Xi^{k''}$, respectively, and let $p^{k'} = \mathbb{P}(\xi \in \Xi^{k'})$ and $p^{k''} = \mathbb{P}(\xi \in \Xi^{k''})$. By the above we have that

$$p^{k'} \left(\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^k)] \right) \geq p^{k'} \left(\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^{k'})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^{k'})] \right),$$

and

$$p^{k''} \left(\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^k)] \right) \geq p^{k''} \left(\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^{k''})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^{k''})] \right),$$

then,

$$\underbrace{(p^{k'} + p^{k''})}_{p^k} \left(\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^k)] \right) \geq p^{k'} \left(\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^{k'})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^{k'})] \right) + p^{k''} \left(\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^{k''})} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^{k''})] \right).$$

Since the other cells in \mathcal{P}_n and \mathcal{P}_{n+1} are the same, it follows that

$$\sum_{k=1}^n p^k \left(\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^k)] \right) \geq \sum_{k=1}^{n+1} p^k \left(\text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^k)] \right),$$

which implies $f_n^U(\mathbf{x}) \geq f_{n+1}^U(\mathbf{x})$ for all $\mathbf{x} \in X$, therefore

$$z_n^U = \text{Min}_{\mathbf{x} \in X} f_n^U(\mathbf{x}) \geq \text{Min}_{\mathbf{x} \in X} f_{n+1}^U(\mathbf{x}) = z_{n+1}^U.$$

1.2 It is true that

$$\bar{\xi}^k = \frac{p^{k'} \bar{\xi}^{k'} + p^{k''} \bar{\xi}^{k''}}{p^k}$$

By convexity it holds that

$$Q(\mathbf{x}, \bar{\xi}^k) \leq \frac{p^{k'}}{p^k} Q(\mathbf{x}, \bar{\xi}^{k'}) + \frac{p^{k''}}{p^k} Q(\mathbf{x}, \bar{\xi}^{k''}).$$

Since the others cells in \mathcal{P}_n and \mathcal{P}_{n+1} are the same, it follows that

$$\sum_{k=1}^n p^k Q(\mathbf{x}, \bar{\xi}^k) \leq \sum_{k=1}^{n+1} p^k Q(\mathbf{x}, \bar{\xi}^k)$$

which implies $f_n^L(\mathbf{x}) \leq f_{n+1}^L(\mathbf{x})$ for all $\mathbf{x} \in X$, therefore

$$z_n^L = \text{Min}_{\mathbf{x} \in X} f_n^L(\mathbf{x}) \leq \text{Min}_{\mathbf{x} \in X} f_{n+1}^L(\mathbf{x}) = z_{n+1}^L.$$

1.3 We have that

$$\begin{aligned} & \text{Max}_{\delta \geq 0} \left\{ \sum_j \delta_j Q(\mathbf{x}, \mathbf{e}_j^k) \mid \sum_j \delta_j = 1, \sum_j \delta_j \mathbf{e}_j = \bar{\xi}^k \right\} \\ & \leq \text{Max}_{\delta \geq 0} \left\{ \sum_j \delta_j Q(\mathbf{x}, \mathbf{e}_j^k) \mid \sum_j \delta_j = 1 \right\}. \end{aligned}$$

Let $M^k := \text{Max}_{\xi \in \Xi^k} Q(\mathbf{x}, \xi) = \text{Max}_{\delta \geq 0} \left\{ \sum_j \delta_j Q(\mathbf{x}, \mathbf{e}_j^k) \mid \sum_j \delta_j = 1 \right\}$. By convexity

$\text{Max}_{\xi \in \Xi^k} Q(\mathbf{x}, \xi) = Q(\mathbf{x}, \mathbf{e}_j^k)$ for some j .

By the existence of $\mathbb{E}[Q(\mathbf{x}, \xi)]$, it holds that (see [1])

$$\sum_{k=1}^n p^k M^k \longrightarrow \mathbb{E}[Q(\mathbf{x}, \xi)] \text{ as } n \longrightarrow \infty,$$

therefore

$$\sum_{k=1}^n p^k \text{Max}_{\delta \in \mathcal{D}(\bar{\xi}^k)} \mathbb{E}^\delta [Q(\mathbf{x}, \mathbf{e}^k)] \longrightarrow \mathbb{E}[Q(\mathbf{x}, \xi)] \text{ as } n \longrightarrow \infty,$$

which implies $f_n^U(\mathbf{x}) \longrightarrow f(\mathbf{x})$ as $n \longrightarrow \infty$ for all $\mathbf{x} \in X$. Since the sequence $\{f_n^U\}_{n=1}^\infty$ is non-increasing and the objective function f is continuous, we have that f_n^U epi-converges to f . So, if $\mathbf{x}^* = \lim_{n \rightarrow \infty} \mathbf{x}_n^U$ then $\mathbf{x}^* \in \text{Argmin}_{\mathbf{x} \in X} f(\mathbf{x})$. Therefore

$$\lim_{n \rightarrow \infty} z_n^U = \lim_{n \rightarrow \infty} f_n^U(\mathbf{x}_n^U) = f(\mathbf{x}^*) = z^*.$$

On another hand, we have that

$$m^k := \text{Min}_{\xi \in \Xi^k} Q(\mathbf{x}, \xi) \leq Q(\mathbf{x}, \bar{\xi}^k) \leq \mathbb{E}[Q(\mathbf{x}, \xi) \mid \xi \in \Xi^k].$$

The left-hand side inequality is valid by definition of m^k , and the right-hand side inequality is Jensen’s inequality. Thus, we have the following valid inequalities:

$$\sum_{k=1}^n p^k m^k \leq \sum_{k=1}^n p^k Q(\mathbf{x}, \bar{\xi}^k) \leq \sum_{k=1}^n p^k \mathbb{E}[Q(\mathbf{x}, \xi) \mid \xi \in \Xi^k] = \mathbb{E}[Q(\mathbf{x}, \xi)].$$

The right-hand side equality is the law of total probability. By the existence of $\mathbb{E}[Q(\mathbf{x}, \xi)]$, it holds that (see [1])

$$\sum_{k=1}^n p^k m^k \rightarrow \mathbb{E}[Q(\mathbf{x}, \xi)] \text{ as } n \rightarrow \infty$$

so,

$$\sum_{k=1}^n p^k Q(\mathbf{x}, \xi^k) \rightarrow \mathbb{E}[Q(\mathbf{x}, \xi)] \text{ as } n \rightarrow \infty,$$

which implies $f_n^L(\mathbf{x}) \rightarrow f(\mathbf{x})$ for all $\mathbf{x} \in X$. Since the sequence $\{f_n^L\}_{n=1}^\infty$ is no-decreasing and the objective function f is continuous, we have that f_n^L epi-converges to f . So, if $\mathbf{x}^* = \lim_{n \rightarrow \infty} \mathbf{x}_n^L$ then $\mathbf{x}^* \in \text{Argmin}_{\mathbf{x} \in X} f(\mathbf{x})$. Therefore

$$\lim_{n \rightarrow \infty} f_n^L(\mathbf{x}_n^L) = f(\mathbf{x}^*) = z^*.$$

□

D The aircraft allocation problem

The aircraft allocation problem was one of the first stochastic linear programs ever formulated by Dantzig [7]. In this problem aircraft of different types are allocated on routes in order to minimize the operating costs. Besides the operating cost, there are costs associated with bumping passengers due to insufficient capacity to meet demand.

Let

- I = set of available aircrafts,
- R = set of routes,
- $R(i)$ = subset of routes serviced by aircraft of type i ,
- b_i = number of aircraft available of type i ,
- c_{ir} = cost of operating an aircraft of type i along route r ,
- t_{ir} = passenger capacity of aircraft i on route r ,
- ξ_r = passenger demand on route r ,
- q_r = revenue lost per bumped passenger on route r ,
- x_{ir} = number of aircraft of type i assigned to route r ,
- y_r = number of bumped passengers on route r ,
- z_r = number of empty seats on route r .

we can set up the aircraft allocation problem with the following model:

$$\begin{aligned} \text{Min}_x \quad & \sum_{i \in I} \sum_{r \in R(i)} c_{ir} \cdot x_{ir} + \mathbb{E}[Q(x, \xi)] \\ \text{s.t.} \quad & \sum_{r \in R(i)} x_{ir} \leq b_i, \quad \forall i \in I, \end{aligned} \tag{26}$$

where

$$\begin{aligned}
 Q(x, \xi) = \text{Min}_{y,z} \quad & \sum_{r \in R} q_r \cdot y_r \\
 \text{s.t.} \quad & \sum_{i \in I, r \in R(i)} t_{ir} \cdot x_{ir} + y_r - z_r = \xi_r, \quad \forall r \in R, \\
 & x_{ir} \geq 0, \quad \forall i \in I, r \in R(i), \quad y_r \geq 0 \quad \forall r \in R.
 \end{aligned}
 \tag{27}$$

The input data to execute this computational experiment was taken from the following site: www.cs.wisc.edu/~swright/stochastic/sampling/

Low recourse cost:

$$\mathbf{q} = (11, 13, 0, 13, 8, 7, 0, 7, 0, 12)^\top.$$

High recourse cost:

$$\mathbf{q} = (301, 349, 239, 700, 254, 493, 348, 70, 474, 361)^\top.$$

E The farmer’s problem

One farmer specializes in raising N types of crops. He has L (km^2) of land and he must decide how much land will be allocated to devote each crop. The fixed cost to rise the i -th type of crop is c_i per ton (T), for $i = 1, \dots, N$. By another hand, he must attend some restrictions related to his plantation; he must have at least h_i (T) of the i -th type of crop, for $i = 1, \dots, N$. Those quantities can be obtained by own plantation or buying them in a local market. The purchase price for the i -th product is s_i per ton (T), for $i = 1, \dots, N$. Additionally, every excess of the i -th type of crop can be sold at the selling price of r_i per ton (T), for $i = 1, \dots, N$.

Let

- ξ_i = productivity land for rising the i -th type of crop,
- x_i = acres of land devoted to rise the i -th type of crop,
- w_i = tons of the i -th type of crop sold,
- y_i = tons of the i -th type of crop purchased

Since the farmer wants to minimize the cost, the two-stage stochastic linear optimization problem is:

$$\begin{aligned}
 \text{Min}_{x \geq 0} \quad & \sum_{i=1}^N c_i \cdot x_i + \mathbb{E}[Q(x, \xi)] \\
 \text{s.t.} \quad & \sum_{i=1}^N x_i \leq L,
 \end{aligned}
 \tag{28}$$

where

$$\begin{aligned}
 Q(x, \xi) = \text{Min}_{y,w} \quad & \sum_{i=1}^N s_i \cdot y_i - \sum_{i=1}^N r_i \cdot w_i \\
 \text{s.t.} \quad & \xi_i \cdot x_i + y_i - w_i \geq h_i, \quad \forall i = 1, \dots, N \\
 & \mathbf{y} \geq 0, \mathbf{w} \geq 0.
 \end{aligned}
 \tag{29}$$

To simplify the problem, we assume that the productivity land for raising each crop is represented by independent random variables with uniform distribution. The data of the two instances evaluated of the farmer’s problem is reported below:

Instance with 8-crops:

Crop	Plantation cost	Purchase price	Selling price
i = 1	92	667	28
i = 2	80	905	35
i = 3	92	1024	41
i = 4	88	660	42
i = 5	91	974	25
i = 6	80	1041	35
i = 7	92	978	40
i = 8	93	997	45

Uncertainty	a_i	b_i
i = 1	2.296	8.575
i = 2	1.981	6.820
i = 3	1.079	7.730
i = 4	2.470	6.232
i = 5	1.305	6.390
i = 6	0.900	6.445
i = 7	2.043	5.430
i = 8	0.409	4.214

Total land (L) : 3500

Instance with 20-crops:

Crop	Plantation cost	Purchase price	Selling price
$i = 1$	81	853	137
$i = 2$	86	570	64
$i = 3$	85	817	135
$i = 4$	83	983	58
$i = 5$	81	1001	70
$i = 6$	93	547	131
$i = 7$	91	728	106
$i = 8$	85	966	135
$i = 9$	89	875	158
$i = 10$	94	913	87
$i = 11$	82	1010	118
$i = 12$	95	844	140
$i = 13$	92	698	133
$i = 14$	94	834	109
$i = 15$	92	724	107
$i = 16$	93	702	150
$i = 17$	84	980	95
$i = 18$	94	771	66
$i = 19$	85	752	77
$i = 20$	81	907	151

Uncertainty	a_i	b_i
$i = 1$	0.963	4.05
$i = 2$	0.844	2.72
$i = 3$	2.070	4.32
$i = 4$	2.640	4.60
$i = 5$	1.730	1.48
$i = 6$	2.400	4.99
$i = 7$	2.410	5.33
$i = 8$	1.980	4.03
$i = 9$	1.180	4.79
$i = 10$	2.060	3.27
$i = 11$	2.900	4.34
$i = 12$	1.100	4.39
$i = 13$	2.690	3.43
$i = 14$	1.260	6.12
$i = 15$	2.750	3.91
$i = 16$	1.150	5.36
$i = 17$	1.920	4.41
$i = 18$	1.760	5.68
$i = 19$	0.547	4.71
$i = 20$	0.508	3.91

Total land (L) : 5200

F The Stochastic unit commitment (SUC) problem

We use a classic and simple formulation of a data-driven two-stage stochastic single-bus unit commitment problem [21]. The first stage comprises the commitment decisions while the second stage accounts for the dispatch decisions. We develop the problem by minimizing the expected total generation cost, and the uncertainty on the right-hand-side of the problem in the net electricity load parameter (load subtracted by uncertain renewable injections).

We summarize the notations by sets, parameters, first-stage variables and second-stage variables listed as follows:

Description of the sets:

- \mathcal{I} = Set of electricity generators,
- \mathcal{T} = Set of time of periods.

Description of the parameters:

- C_i^u = Fixed cost of unit $i \in \mathcal{I}$,
- C_i^{SU} = Start-up cost of unit $i \in \mathcal{I}$,
- C_i^{SD} = Shut-down cost of unit $i \in \mathcal{I}$,
- R_i^U = Ramp-up limit of unit $i \in \mathcal{I}$,
- R_i^D = Ramp-down limit of unit $i \in \mathcal{I}$,
- \bar{P}_i = Maximum power generation of unit $i \in \mathcal{I}$,
- \underline{P}_i = Minimum power generation of unit $i \in \mathcal{I}$,
- $\widehat{\xi}(t)$ = The electricity load in time $t \in \mathcal{T}$.

Description of the first-stage variables:

- $u_{i,t}$ = Binary commit variable: 1 if the thermal generator i is on in time t ; 0 otherwise,
- $v_{i,t}$ = Start-up variable for unit $i \in \mathcal{I}$ in time $t \in \mathcal{T}$,
- $w_{i,t}$ = Shut-down variable for unit $i \in \mathcal{I}$ in time $t \in \mathcal{T}$.

Description of the second-stage variables:

- $p_{i,t}$ = Power generation of unit $i \in \mathcal{I}$ in time $t \in \mathcal{T}$.

Let us denote by $\mathbf{x} = [u_{i,t}, v_{i,t}, w_{i,t}]_{i \in \mathcal{I}, t \in \mathcal{T}}$ the first-stage variable. Based on this notation, the two-stage stochastic linear optimization problem is:

$$\text{Min}_{u_{i,t}, v_{i,t}, w_{i,t}} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} \left[C_i^u u_{i,t} + C_i^{SU} v_{i,t} + C_i^{SD} w_{i,t} \right] + \mathbb{E}[Q(\mathbf{x}, \xi)] \tag{30a}$$

$$\text{s.t. } v_{i,t} - w_{i,t} = u_{i,t} - u_{i,t-1}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \tag{30b}$$

$$v_{i,t} \leq u_{i,t}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \tag{30c}$$

$$w_{i,t} \leq 1 - u_{i,t}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \tag{30d}$$

where

$$Q(\mathbf{x}, \xi) = \min_{p_{i,t}} C_i^P p_{i,t} \tag{31a}$$

$$s.t. \quad \sum_{i \in \mathcal{I}} p_{i,t} = \xi_t, \quad \forall t \in \mathcal{T}, \quad (31b)$$

$$p_{i,t} - p_{i,t-1} \leq R_i^U u_{i,t-1} + \bar{P}_i v_{i,t}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (31c)$$

$$p_{i,t-1} - p_{i,t} \leq R_i^D u_{i,t} + \underline{P}_i w_{i,t}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (31d)$$

$$\underline{P}_i u_{i,t} \leq p_{i,t} \leq \bar{P}_i u_{i,t}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (31e)$$

$$0 \leq v_{i,t} \leq 1, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (31f)$$

$$0 \leq w_{i,t} \leq 1, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (31g)$$

$$u_{i,t} \in \{0, 1\}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}. \quad (31h)$$

In the above formulation, the constraints (30b)–(30d) are the start-up and shut-down operational constraints for each thermal unit. Constraint (31b) ensures load balance. Constraints (31c)–(31d) are the the ramping up and ramping down constraints, respectively. Constraint (31e) is the minimum and maximum power generation of the unit $i \in \mathcal{I}$ in time $t \in \mathcal{T}$. Constrains (31f)–(31g) are box constraints. Finally, constraint (31h) establishes the integrability of the binary variable $[[u]_{i,t}]_{i \in \mathcal{I}, t \in \mathcal{T}}$.

References

1. Apostol, T.: *Mathematical Analysis*. Addison-Wesley series in mathematics. Addison-Wesley, Boston (1974)
2. Beichl, I., Sullivan, F.: The importance of importance sampling. *Comput. Sci. Eng.* **1**(2), 71–73 (1999)
3. Bertsimas, D., Gupta, V., Kallus, N.: Robust sample average approximation. *Math. Program.* **171**(1–2), 217–282 (2018)
4. Bertsimas, D., Sim, M.: The price of robustness. *Oper. Res.* **52**(1), 35–53 (2004)
5. Birge, J., Louveaux: *Introduction to Stochastic Programming*, 2nd edn. Springer, New York Dordrecht Heidelberg London (2011)
6. Birge, J.R., West, R.J.: Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse. *Stoch. Program.* **84**, 54–102 (1986)
7. Dantzig, G.B.: *Linear Programming and Extensions*. Princeton University Press, Princeton (1991)
8. Dunning, I., Huchette, J., Lubin, M.: Jump: a modeling language for mathematical optimization. *SIAM Rev.* **59**(2), 295–320 (2017)
9. Edirisinghe, N.C.P., You, G.-M.: Second-order scenario approximation and refinement in optimization under uncertainty. *Ann. Oper. Res.* **64**(1), 143–178 (1996)
10. Edirisinghe, N.C.P., Ziemba, W.T.: Bounds for two-stage stochastic programs with fixed recourse. *Math. Oper. Res.* **19**(2), 292–313 (1994)
11. Edirisinghe, N.C.P., Ziemba, W.T.: Implementing bounds-based approximations in convex–concave two-stage stochastic programming. *Math. Program.* **75**(2), 295–325 (1996)
12. Frauendorfer, K.: Solving slp recourse problems with arbitrary multivariate distributions: the dependent case. *Math. Oper. Res.* **13**(3), 377–394 (1988)
13. Frauendorfer, K., Kall, P.: A solution method for slp recourse problems with arbitrary multivariate distributions: the independent case. *Prob. Control Inform. Theory* **17**, 177–205 (1988)
14. Gassmann, H., Ziemba, W.T.: A tight upper bound for the expectation of a convex function of a multivariate random variable, pp. 39–53. Springer Berlin Heidelberg, Berlin, Heidelberg (1986)
15. Huang, C.C., Ziemba, W.T., Ben-Tal, A.: Bounds on the expectation of a convex function of a random variable: with applications to stochastic programming. *Oper. Res.* **25**(2), 315–325 (1977)
16. Kall, P., Ruszczyński, A., Frauendorfer, K.: Approximation techniques in stochastic programming, pp. 33–64. Springer, Berlin (1988)
17. Kall, P., Stoyan, D.: Solving stochastic programming problems with recourse including error bounds. *Math. Oper. schung und Statistik Ser. Optim.* **13**(3), 431–447 (1982)

18. Madansky, A.: Bounds on the expectation of a convex function of a multivariate random variable. *Ann. Math. Statist.* **30**(3), 743–746 (1959)
19. Mohajerin Esfahani, P., Kuhn, D.: Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. *Math. Program.* **171**(1), 115–166 (2018)
20. O’Leary, D.: Multidimensional integration: partition and conquer. *Comput. Sci. Eng.* **6**(6), 58–66 (2004)
21. Papavasiliou, A., Oren, S.S.: Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network. *Oper. Res.* **61**(3), 578–592 (2013)
22. Piessens, R., de Doncker-Kapenga, E., Überhuber, C.W., Kahaner, D.K.: *Quadpack: a subroutine package for automatic integration*, vol. 1. Springer, Berlin (2012)
23. Street, A., Oliveira, F., Arroyo, J.M.: Contingency-constrained unit commitment with $n - k$ security criterion: a robust optimization approach. *IEEE Trans. Power Syst.* **26**(3), 1581–1590 (2011)
24. Zeng, B., Zhao, L.: Solving two-stage robust optimization problems using a column-and-constraint generation method. *Oper. Res. Lett.* **41**(5), 457–461 (2013)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.