

# GENERALIZED ADAPTIVE PARTITION-BASED METHOD FOR TWO-STAGE STOCHASTIC PROGRAMS WITH FIXED RECOURSE

CRISTIAN RAMIREZ-PICO AND EDUARDO MORENO

**ABSTRACT.** We present a method to solve two-stage stochastic problems with fixed recourse when the uncertainty space can have either discrete or continuous distributions. Given a partition of the uncertainty space, the method is addressed to solve a discrete problem with one scenario for each element of the partition (sub-regions of the uncertainty space). Fixing first stage variables, we formulate a second stage subproblem for each element, and exploiting information from the dual of these problems, we provide conditions that the partition must satisfy to obtain the optimal solution. These conditions provide guidance on how to refine the partition, converging iteratively to the optimal solution. Results from computational experiments show how the method automatically refines the partition of the uncertainty space in the regions of interest for the problem. Our algorithm is a generalization of the adaptive partition-based method presented by Song & Luedtke for discrete distributions, extending its applicability to more general cases.

## 1. INTRODUCTION

We study the following two-stage stochastic program (TSSP) with fixed recourse

$$(1) \quad \min \{c^\top x + \mathbb{E}[\mathcal{Q}(x, \xi)] \mid x \in \mathcal{X}\}$$

where  $\mathcal{X} \subseteq \mathbb{R}^n$  is a set assumed to be non-empty closed,  $\xi$  is a random vector in the probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  containing the random elements  $\{h^\xi, T^\xi\}$ , and second-stage subproblem

$$(2) \quad \mathcal{Q}(x, \xi) := \min \{q^\top y \mid Wy = h^\xi - T^\xi x, y \geq 0\},$$

where fixed recourse matrix  $W \in \mathbb{R}^{m \times n}$ , deterministic costs  $q \in \mathbb{R}^n$ , random technology matrix  $T^\xi \in \mathbb{R}^{m \times n}$  and random right-hand side (RHS) vector  $h^\xi \in \mathbb{R}^m$ . Furthermore, we assume that there exists  $\bar{x}$  such that  $\mathcal{Q}(\bar{x}, \xi)$  is feasible and bounded in the whole outcome space  $\Omega$ . Note that the support of the uncertainty set  $\xi$  can be either continuous or discrete.

In this paper, we propose a method to solve TSSPs by iteratively and automatically aggregating the uncertainty set into a small number of scenarios and disaggregating them based on the information of dual subproblem variables. This approach yields a smaller version of the original stochastic problem by reducing both the number of variables and the number of constraints by an equivalent deterministic formulation of Eq. (1). For the case of discrete distributions, this idea has been called the *adaptive partition-based method* (APM) by Song & Luedtke [27], and it is based on the results of Espinoza & Moreno [6] and Bienstock & Zuckerberg [4].

2010 *Mathematics Subject Classification.* 90C15, 90-08.

Supported by CONICYT-Fondecyt Regular 1161064.

We present an alternative and more general proof that allows us to extend APM to a more general setting, in particular, to deal TSSPs with continuous distributions for  $\Omega$ .

Let  $P \subseteq \Omega$ , and let  $T^P = \mathbb{E}[T^\xi|P]$  and  $h^P = \mathbb{E}[h^\xi|P]$  be the conditional expectations of the components of  $\xi$  given  $P$ . We denote the *aggregated subproblem* as

$$(3) \quad \mathcal{Q}(x, \mathbb{E}[\xi|P]) = \min \{q^\top y \mid Wy = h^P - T^P x, y \geq 0\}$$

The contribution of this paper is to provide conditions for a partition  $\mathcal{P}$  of  $\Omega$  such that the solution of Problem [Eq. \(1\)](#) is equivalent to solving

$$(4) \quad \min_{x \in \mathcal{X}} \left\{ c^\top x + \sum_{P \in \mathcal{P}} \mathcal{Q}(x, \mathbb{E}[\xi|P]) \cdot \mathbb{P}(P) \right\}.$$

Note that this problem is equivalent to a TSSP with a discrete distribution of  $|\mathcal{P}|$  scenarios for the uncertainty space. Moreover, this approach enables us to generate algorithms to obtain exact optimal solutions for general TSSPs.

The remainder of this paper is organized as follows. [Section 2](#) reviews the literature concerning the APM for discrete TSSPs and other approaches to solve this problem. [Section 3](#) develops the generalized adaptive partition-based method (GAPM), with the main mathematical results to validate this approach. [Section 4](#) discusses the details of the algorithms that are implemented for two well-known stochastic programming problems in [Section 5](#). Finally, concluding remarks are presented in [Section 6](#).

## 2. LITERATURE REVIEW

In past decades, researchers have developed solution strategies for multiple stochastic optimization problems. However, the majority of studies start from the deterministic equivalent formulation to obtain alternative models that are more tractable in algorithmic terms, one of the most studied and utilized problems are two-stage stochastic problems. In their seminal paper, Kleywegt et al. [\[8\]](#) show that any TSSP formulation can be approximated by solving Problem [Eq. \(1\)](#) for a discrete set of samples of  $\xi$  from the original probability space  $\Omega$ : they called this result the sample average approximation method. A key fact from the paper is that good approximations require a large number of scenarios to guarantee an  $\epsilon$ -optimal solution. Since then, most of the research on this problem has been focused on solving large-scale instances of discrete TSSPs with many scenarios.

A common and widely studied approach is to decompose TSSPs via the block structure of the scenario formulation. The most classic approach is called Benders decomposition (or the L-Shaped method as its stochastic variant [\[29\]](#)).

Most of the improvements of this approach focused on reducing the algorithm instability, such as the case of regularized decomposition [\[24\]](#), level decomposition [\[11, 33\]](#) and inexact bundle methods [\[16, 32\]](#). Recent developments with respect to Benders are proposed in [\[19, 20, 21\]](#), which primarily explore how to accelerate and parallelize the technique, and [\[1, 25\]](#), which consider how to address integer problems.

Other decomposition methodologies include stochastic decomposition [\[7\]](#), progressive hedging [\[23, 31\]](#) and stochastic dual dynamic programming [\[18\]](#) for the case of multistage stochastic problems.

A different approach was developed based on the general decomposition method proposed by Bienstock & Zuckerberg [4, 15]. Espinoza & Moreno [6] introduced an algorithm based on this decomposition method to minimize risk measures in linear programs. This idea was later extended by Song & Luedtke [27] to general TSSPs with discrete distributions, where the term *adaptive partition-based method* was coined. These studies have been extended recently by combination with Benders decomposition [17], level decomposition [17, 28], and new extensions have been made to multi-stage stochastic problems [26].

As mentioned previously, most of the recent developments are oriented to the discrete case, relying on approximation by samples of continuous probability distributions for uncertain parameters. Exact methods for TSSPs with nondiscrete distributions are scarce, and they focus mostly on particular problems and distributions that can be reformulated in a more tractable manner. Other general techniques for these problems include [3], which introduces equivalent linear and nonlinear formulations for TSSPs with simple recourse according to the probability distributions of random parameters, and [5] which poses a methodology that benefits from the reduced cost of duality and sensitivity analysis to fix the correct values of some variables in the stochastic program, thereby reducing the size of the original problem.

To the best of our knowledge, this paper is one of the first exact methods based on linear programming intended to deal with TSSPs with general continuous distributions for the stochastic parameters.

### 3. GENERALIZED ADAPTIVE PARTITION-BASED METHOD

We propose a methodology which benefits from a structure shared by aggregated and atomized subproblems, which latter allows us to derive conditions such that the scenarios (either finite or infinite number of them) belonging to a certain element  $P$ , yield the same expected value of optimal solutions, as if we solve the aggregated Problem Eq. (3).

**3.1. Relations between atomized and aggregated subproblems.** As a first step, we define the relation between subproblems Eq. (2) and aggregated subproblems Eq. (3). Indeed, Lemma 3.1 shows how a feasible solution of Eq. (3) can be constructed using information from the optimal solution of Eq. (2).

**Lemma 3.1.** *Let  $\bar{x} \in \mathcal{X}$  and  $P \subseteq \Omega$ , and let  $\hat{y}^\xi$  be the set of optimal solutions of  $\mathcal{Q}(\bar{x}, \xi)$  for  $\xi \in P$ . Then,  $\hat{y}^P := \mathbb{E}[\hat{y}^\xi | P]$  is a feasible solution for  $\mathcal{Q}(\bar{x}, \mathbb{E}[\xi | P])$ .*

*Proof.* Since  $W\hat{y}^\xi = h^\xi - T^\xi \bar{x}$  for every  $\xi \in P$ ,

$$\begin{aligned} \int_{\Omega} W\hat{y}^\xi \, d\mathbb{P}(\xi | P) &= \int_{\Omega} [h^\xi - T^\xi \bar{x}] \, d\mathbb{P}(\xi | P) \\ W \int_{\Omega} \hat{y}^\xi \, d\mathbb{P}(\xi | P) &= \int_{\Omega} h^\xi \, d\mathbb{P}(\xi | P) - \left( \int_{\Omega} T^\xi \, d\mathbb{P}(\xi | P) \right) \bar{x} \\ W \mathbb{E}[\hat{y}^\xi | P] &= \mathbb{E}[h^\xi | P] - \mathbb{E}[T^\xi | P] \bar{x}. \end{aligned}$$

Hence,  $\hat{y}^P$  is a feasible solution for  $\mathcal{Q}(\bar{x}, \mathbb{E}[\xi | P])$ .  $\square$

Since these second-stage subproblems consider only continuous variables, we can introduce a dual formulations for subproblems Eq. (2) and Eq. (3), respectively,

$$(5) \quad \mathcal{Q}^D(x, \xi) := \max \{ (h^\xi - T^\xi x)^\top \lambda^\xi \mid W^\top \lambda \leq q \}$$

and

$$(6) \quad \mathcal{Q}^D(x, P) := \max \{ (h^P - T^P x)^\top \lambda^P \mid W^\top \lambda^P \leq q \}.$$

Indices  $\xi$  and  $P$  on dual variable  $\lambda$  distinguish between atomized and aggregated subproblems.

Similarly to the primal case, we can construct a feasible solution for problem Eq. (6) based on the optimal solutions of Eq. (5).

**Lemma 3.2.** *Let  $\bar{x} \in \mathcal{X}$  and  $P \subseteq \Omega$ , and let  $\hat{\lambda}^\xi$  be the optimal solution of problem  $\mathcal{Q}^D(\bar{x}, \xi)$  for  $\xi \in P$ . Then,  $\hat{\lambda}^P := \mathbb{E}[\hat{\lambda}^\xi | P]$  is a feasible solution for  $\mathcal{Q}^D(\bar{x}, P)$*

*Proof.* Since  $W^\top \hat{\lambda}^\xi \leq q$  for all  $\xi \in P$ ,

$$W^\top \hat{\lambda}^P = W^\top \int_{\Omega} \hat{\lambda}^\xi d\mathbb{P}(\xi | P) = \int_{\Omega} W^\top \hat{\lambda}^\xi d\mathbb{P}(\xi | P) \leq \int_{\Omega} q \cdot d\mathbb{P}(\xi | P) = q.$$

Hence,  $\hat{\lambda}^P$  is a feasible solution for problem  $\mathcal{Q}^D(\bar{x}, P)$  whenever set  $P$  has positive measure.  $\square$

**3.2. Construction of an optimal partition.** The previous framework provides the set of tools necessary to set  $\mathcal{Q}(\bar{x}, \mathbb{E}[\xi | P])$  as a lower bound of  $\mathbb{E}[\mathcal{Q}(\bar{x}, \xi) | P]$ ; furthermore, we identify the conditions on  $\mathcal{P}$  to make this bound tight.

**Proposition 3.3.** *Let  $\bar{x} \in \mathcal{X}$  and  $P \subseteq \Omega$ , such that  $\mathcal{Q}(\bar{x}, \xi)$  is feasible for all  $\xi \in P$ , and let  $\hat{\lambda}^\xi$  be its dual optimal solutions. If  $\hat{\lambda}^\xi$  for  $\xi \in P$  satisfies*

$$(7a) \quad \left( \mathbb{E}[h^\xi | P] \right)^\top \left( \mathbb{E}[\hat{\lambda}^\xi | P] \right) = \mathbb{E} \left[ h^\xi{}^\top \hat{\lambda}^\xi \mid P \right]$$

$$(7b) \quad \bar{x}^\top \left( \mathbb{E}[T^\xi | P]^\top \mathbb{E}[\hat{\lambda}^\xi | P] \right) = \bar{x}^\top \mathbb{E} \left[ T^\xi{}^\top \hat{\lambda}^\xi \mid P \right]$$

then,

$$\mathcal{Q}(\bar{x}, \mathbb{E}[\xi | P]) = \mathbb{E}[\mathcal{Q}(\bar{x}, \xi) | P]$$

*Proof.* We first note that  $\mathcal{Q}(\bar{x}, \xi)$  is a convex function on  $\xi$ . Let us consider the problem  $f(\xi) = \min \{ q^\top y \mid Wy = \xi \}$ . Now, we can take random values of  $b$  namely  $b_1$  and  $b_2$ , with  $y_1^*$  and  $y_2^*$  as their respective optimal solutions. If  $b_\beta$  is a convex combination of  $b_1$  and  $b_2$ , then  $\beta y_1^* + (1 - \beta) y_2^*$  is a feasible solution of  $f(b_\beta)$ . If  $y_\beta^*$  is the optimal solution of  $f(b_\beta)$ , we can build

$$\begin{aligned} f(b_\beta) &= q^\top y_\beta^* \leq q^\top (\beta y_1^* + (1 - \beta) y_2^*) \\ &= \beta q^\top y_1^* + (1 - \beta) q^\top y_2^* = \beta f(b_1) + (1 - \beta) f(b_2) \end{aligned}$$

which verifies  $\mathcal{Q}(\bar{x}, \xi)$  is convex on  $\xi$ . Hence, by applying Jensen's inequality, we obtain  $\mathcal{Q}(\bar{x}, \mathbb{E}[\xi | P]) \leq \mathbb{E}[\mathcal{Q}(\bar{x}, \xi) | P]$ .

On the other hand, according to Lemma 3.2, we know that  $\hat{\lambda}^P := \mathbb{E}[\hat{\lambda}^\xi | P]$  is a feasible solution of  $\mathcal{Q}^D(\bar{x}, P)$ ; thus,

$$\begin{aligned} \mathcal{Q}^D(\bar{x}, P) &= \mathcal{Q}(\bar{x}, \mathbb{E}[\xi | P]) \geq \left( \mathbb{E}[h^\xi | P] - \mathbb{E}[T^\xi | P] \bar{x} \right)^\top \left( \mathbb{E}[\hat{\lambda}^\xi | P] \right) \\ &= \left( \mathbb{E}[h^\xi | P] \right)^\top \left( \mathbb{E}[\hat{\lambda}^\xi | P] \right) - \bar{x}^\top \left( \mathbb{E}[T^\xi | P] \right)^\top \left( \mathbb{E}[\hat{\lambda}^\xi | P] \right) \end{aligned}$$

Since  $\hat{\lambda}^P$  satisfies conditions [Eq. \(7\)](#), by means of the linearity of the expectation, we obtain

$$\begin{aligned}\mathcal{Q}(\bar{x}, \mathbb{E}[\xi|P]) &\geq \left( \mathbb{E} \left[ (h^\xi)^\top \hat{\lambda}^\xi \middle| P \right] \right) - \bar{x}^\top \left( \mathbb{E} \left[ (T^\xi)^\top \hat{\lambda}^\xi \middle| P \right] \right) \\ &= \mathbb{E}[\mathcal{Q}(\bar{x}, \xi) | P]\end{aligned}$$

□

**Theorem 3.4.** *Let  $x^*$  be an optimal solution of problem*

$$\min_{x \in \mathcal{X}} \left\{ c^\top x + \sum_{P \in \mathcal{P}^*} \mathcal{Q}(x, \mathbb{E}[\xi|P]) \cdot \mathbb{P}(P) \right\}$$

where  $\mathcal{P}^*$  is a numerable partition of  $\Omega$  such that for each  $P \in \mathcal{P}^*$ , the optimal dual variables of  $\mathcal{Q}(x^*, \xi)$  for  $\xi \in P$  satisfy conditions [Eq. \(7\)](#). Then,  $x^*$  is also an optimal solution of problem

$$\min_{x \in \mathcal{X}} \{ c^\top x + \mathbb{E}[\mathcal{Q}(x, \xi)] \}.$$

*Proof.* By the laws of total expectation, we know that for any numerable partition  $\mathcal{P}$  of  $\Omega$ ,

$$\mathbb{E}[\mathcal{Q}(x, \xi)] = \sum_{P \in \mathcal{P}} \mathbb{E}[\mathcal{Q}(x, \xi) | P] \cdot \mathbb{P}(P)$$

In particular, for  $x^*$  and  $\mathcal{P}^*$ , according to [Proposition 3.3](#), we obtain

$$c^\top x^* + \mathbb{E}[\mathcal{Q}(x^*, \xi)] = c^\top x^* + \sum_{P \in \mathcal{P}^*} \mathcal{Q}(x^*, \mathbb{E}[\xi|P]) \cdot \mathbb{P}(P),$$

hence,

$$\min_{x \in \mathcal{X}} \{ c^\top x + \mathbb{E}[\mathcal{Q}(x, \xi)] \} \leq \min_{x \in \mathcal{X}} \left\{ c^\top x + \sum_{P \in \mathcal{P}^*} \mathcal{Q}(x, \mathbb{E}[\xi|P]) \cdot \mathbb{P}(P) \right\}.$$

On the other hand, if  $\hat{x}$  is the optimal solution of Problem [Eq. \(1\)](#), then

$$\begin{aligned}c^\top \hat{x} + \mathbb{E}[\mathcal{Q}(\hat{x}, \xi)] &= c^\top \hat{x} + \sum_{P \in \mathcal{P}^*} \mathbb{E}[\mathcal{Q}(\hat{x}, \xi) | P] \cdot \mathbb{P}(P) \\ &\geq c^\top \hat{x} + \sum_{P \in \mathcal{P}^*} \mathcal{Q}(\hat{x}, \mathbb{E}[\xi|P]) \cdot \mathbb{P}(P) \\ &\geq \min_{x \in \mathcal{X}} \left\{ c^\top x + \sum_{P \in \mathcal{P}^*} \mathcal{Q}(x, \mathbb{E}[\xi|P]) \cdot \mathbb{P}(P) \right\},\end{aligned}$$

where the first equality is true by the laws of total expectation and the second inequality is given by Jensen's inequality and the convexity of  $\mathcal{Q}(\hat{x}, \cdot)$ . □

Note that this partition always exists, as presented in the following corollary.

**Corollary 3.5.** *If  $x^*$  is an optimal solution of problem*

$$\min_{x \in \mathcal{X}} \{ c^\top x + \mathbb{E}[\mathcal{Q}(x, \xi)] \}$$

then there exists a finite partition  $\mathcal{P}^*$  of  $\Omega$  such that

$$c^\top x^* + \mathbb{E}[\mathcal{Q}(x^*, \xi)] = c^\top x^* + \sum_{P \in \mathcal{P}^*} \mathcal{Q}(x^*, \mathbb{E}[\xi|P]) \cdot \mathbb{P}(P)$$

*Proof.* Note that the dual feasible solutions  $\lambda$  of  $Q(x^*, \xi)$  must satisfy  $W^\top \lambda \leq q$ . Hence, for each  $\xi \in \Omega$ , we have an associated extreme point of  $W^\top \lambda \leq q$  that is an optimal dual solution of  $Q(x^*, \xi)$ . This result induces a finite partition  $\mathcal{P}^*$  of  $\Omega$  such that all  $\xi \in P$  have the same dual optimal solution of  $Q(x^*, \xi)$ . Hence, the values all satisfy the conditions of [Proposition 3.3](#), and

$$\begin{aligned} c^\top x^* + \mathbb{E}[\mathcal{Q}(x^*, \xi)] &= c^\top x^* + \sum_{P \in \mathcal{P}^*} \mathbb{E}[\mathcal{Q}(x^*, \xi) | P] \cdot \mathbb{P}(P) \\ &= c^\top x^* + \sum_{P \in \mathcal{P}^*} \mathcal{Q}(x^*, \mathbb{E}[\xi | P]) \cdot \mathbb{P}(P). \end{aligned}$$

□

We finish this section by noting some differences regarding the original APM proposed in [\[27\]](#). The most relevant aspect of this proof is the possibility to extend it to the case of continuous probability space. However, this proof also shows that the method can be applied to any set of primal variables  $x \in \mathcal{X}$ . Moreover, the condition proposed in the original paper to aggregate scenarios is such that all dual variables  $\lambda^\xi$  for  $\xi \in P$  on each subset  $P \subset \Omega$  must have the same value. This is a particular case which satisfies the conditions of [Proposition 3.3](#) by means of the linearity of the expected value. Finally, [\[27\]](#) establishes that this criterion is required to have the equality between the value of the aggregated problem and the expected value of the atomized subproblems. Nonetheless, the presented conditions of [Proposition 3.3](#) provides a framework where less demanding conditions might be applied to aggregate/disaggregate scenarios, e.g., degenerated subproblems with multiple optimal dual solutions.

#### 4. ALGORITHM IMPLEMENTATION

The idea of the method is to iteratively converge to a partition  $\mathcal{P}$  satisfying the conditions of [Proposition 3.3](#). Initially, we start with a trivial partition ( $\mathcal{P} = \{\Omega\}$ ) and split the partition based on the duals of the subproblems. This split procedure is problem dependent. At each iteration, the algorithm provides a lower bound (the optimal value of the aggregated problem) and, potentially, an upper bound. The upper bound can be computed by solving the expected value of the subproblem, which is easy to compute in the discrete case by solving the subproblem for each scenario independently. However, the computation can be difficult for continuous distributions. The algorithm is presented in [Algorithm 1](#).

There are two key steps in the GAPM implementation proposed in [Algorithm 1](#). First, to execute [Line 10](#) correctly, we require an additional procedure to split the uncertain region in an appropriate manner; this step is fine-tuned according to the structure of subproblems [Eq. \(2\)](#). We discuss this point in the computational experiments in [Section 5](#). On the other hand, if an upper bound  $z_U^{(t)}$  cannot be computed, on [Line 10](#) we can still check [Proposition 3.3](#) or even compare composition of consecutive partitions  $\mathcal{P}^{(t)}$  and  $\mathcal{P}^{(t+1)}$ .

#### 5. NUMERICAL EXPERIMENTS

Since the fundamental novelty of our proposal arises when stochastic parameters have continuous probability distributions, the computational experiments are designed to enlighten algorithmic behaviour on two problems from classic literature

**Algorithm 1** An iterative implementation of GAPM**Require:** A stopping threshold  $\epsilon$  and an initial partition  $\mathcal{P}^{(0)}$  of  $\Omega$ **Ensure:** An optimal partition  $\mathcal{P}^*$  to attain the optimal solution  $x^*$  of problem

Eq. (1)

---

```

1: Set  $t := 0$ ,  $z_L^{(0)} := -\infty$  and  $z_U^{(0)} = \infty$ .
2: loop
3:    $t = t + 1$ 
4:   Solve Problem Eq. (4) for partition  $\mathcal{P}^{(t)}$  and assign its optimal value to the
     lower bound  $z_L^{(t)}$  and its optimal solution to  $\bar{x}^{(t)}$ 
5:   If possible, compute the upper bound  $z_U^{(t)} = c^\top \bar{x}^{(t)} + \mathbb{E}[\mathcal{Q}(\bar{x}^{(t)}, \xi)]$ 
6:   Solve subproblems Eq. (3) for every  $P \in \mathcal{P}^{(t)}$ 
7:   if  $z_U^{(t)} - z_L^{(t)} < \epsilon$  then
8:     exit
9:   end if
10:  Run disaggregation procedure to split uncertain region and obtain  $\mathcal{P}^{(t+1)}$ 
11:  if  $\mathcal{P}^{(t+1)}$  satisfies Eq. (7) then
12:    exit
13:  end if
14: end loop
15: return optimal solution  $x^* := \bar{x}^{(t)}$ , optimal partition  $\mathcal{P}^* := \mathcal{P}^{(t)}$  and optimal
     value  $z_L^{(t)}$ 

```

---

with this type of uncertainty. For the case of a discrete distribution, we refer the reader to the papers presented in the literature review.

We have divided the computational experiment into two parts. First, we discuss the implementation and results for a classic problem from the stochastic programming literature, the LandS instance, wherein uncertainty is presented in the RHS coefficients. The second problem is the TSSP reformulation of conditional value-at-risk (CVaR) minimization, where the uncertainty appears in the technological coefficient of the first-stage variables  $x$ . Both problems have well-defined structures that are useful to define the procedure to split the uncertainty space  $\Omega$  at each iteration of the algorithm.

**5.1. Energy planning problem - LandS.** LandS, a classic problem in stochastic programming that is studied for academic purposes, was originally proposed in [13]. LandS is an energy planning investment problem, where the goal is to decide the capacities of four new plants while minimizing allocation and operational costs. The set of power plants are supposed to meet uncertain demand of three different electric modes. In the first stage, some minimum capacities and budget constraints must be satisfied; during the second stage, energy is distributed according to the realization of the uncertain demands. The mathematical formulation is as follows:

$$(8) \quad \min_{x \geq 0} \left\{ \sum_{i \in \mathbb{I}} c_i x_i + \mathbb{E}[\mathcal{Q}(x, \xi)] : \sum_{i \in \mathbb{I}} x_i \geq m, \sum_{i \in \mathbb{I}} c_i x_i \leq b \right\}$$

where

$$(9a) \quad \mathcal{Q}(x, \xi) := \min_{y \geq 0} \sum_{i \in \mathbb{I}} \sum_{j \in \mathbb{J}} f_{ij} y_{ij}$$

$$(9b) \quad \text{s.t.} \quad \sum_{j \in \mathbb{J}} y_{ij} \leq x_i, \quad \forall i \in \mathbb{I}$$

$$(9c) \quad \sum_{i \in \mathbb{I}} y_{ij} \geq d_j^\xi, \quad \forall j \in \mathbb{J}$$

The original problem sets up an uncertain demand for  $d_1^\xi$  with three scenarios: 3, 5 or 7 units. In this experiment, we assume that  $d_1^\xi$  follows a uniform distribution in the interval  $[3, 7]$ , following the ideas from [12]. Remaining demands are considered to be deterministic.

To split the uncertainty space  $\Omega$  and compute an upper bound for the optimal value of the problem, we introduce the dual of  $\mathcal{Q}(\hat{x}, \xi)$  given by

$$(10) \quad \mathcal{Q}^D(\hat{x}, \xi) := \max_{\nu, \mu \geq 0} \sum_{j \in \mathbb{J}} \mu_j d_j^\xi - \sum_{i \in \mathbb{I}} \nu_i \hat{x}_i$$

$$\mu_j - \nu_i \leq f_{ij} \quad \forall i \in \mathbb{I}, j \in \mathbb{J}$$

where  $\nu$  and  $\mu$  correspond to the dual variables of constraints Eq. (9b) and Eq. (9c), respectively. Then, given an optimal solution of the subproblem for a given value of  $d_1^\xi$ , we can use sensitivity analysis to compute a neighbourhood around  $d_1^\xi$  in which the dual optimal variables do not change. Moreover,  $\mathcal{Q}^D(x, \xi)$  is a non-decreasing piecewise linear function on  $d_1^\xi$ , so the upper bound of Line 5 is easy to compute.

In our experiment, we start with  $\mathcal{P}^{(0)} = \{[3, 7]\}$  and in each iteration, the partition is refined by dividing the corresponding elements of  $\mathcal{P}$ , utilizing the segment extremes of piecewise linear function  $\mathcal{Q}^D(\bar{x}^{(t)}, \xi)$ .

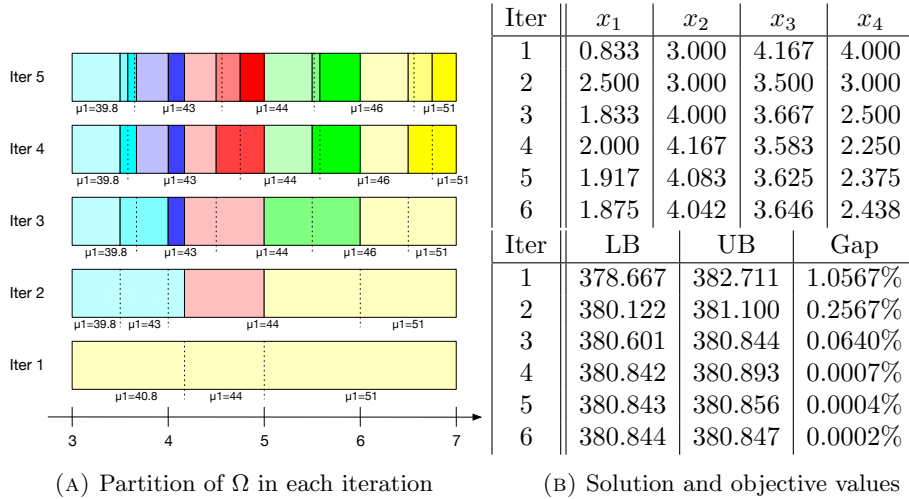


FIGURE 1. Iteration details for the LandS example

In Fig. 1, we show the resulting first six iterations of the algorithm. Columns LB and UB present the current lower bound (objective value of the aggregated problem)



and the upper bound (computed by the benefit of  $\bar{x}^{(t)}$  and subproblem optimal dual variables), respectively. The column Gap shows the relative gap between the current solution and the best upper bound obtained so far. After a few iterations, we obtain near-optimal solutions for the problem, with a gap close to the computational precision of the optimization software. In Fig. 1a, we present the partition in each iteration (highlighted by different colours), as well as the segments (dotted lines) obtained after carrying out the sensitivity analysis. The value under each segment corresponds to the dual variable of the stochastic demand constraint. Notably, these dual values do not change after iteration 3, but the extremes of the corresponding intervals change slightly in each iteration until converging to the optimal solution.

**5.2. Conditional value-at-risk linear problems.** A classic problem in risk optimization is to minimize the CVaR, which is a well-known risk measure satisfying the properties of coherency [2]. In our case, we assume a linear problem, where the objective coefficients  $\tilde{r}^\xi$  are random, and we minimize the  $\text{CVaR}(x^\top \tilde{r}^\xi)$  subject to linear constraints  $Ax \leq b$ . This problem (see [22]) can be reformulated as

$$\min_{x, \tau} \left\{ \tau + \frac{1}{\delta} \mathbb{E}[-x^\top \tilde{r}^\xi - \tau]^+ : Ax \leq b \right\}$$

In our context,  $x$  and  $\tau$  are the first-stage decisions, while the second-stage subproblem is

$$\mathcal{Q}((x, \tau), \xi) := (-x^\top \tilde{r}^\xi - \tau)^+ = \min\{z : z \geq -x^\top \tilde{r}^\xi - \tau, z \geq 0\}$$

Let us note that the dual of  $\mathcal{Q}((x, \tau), \xi)$  has a single dual variable  $\lambda$ , and it can be formulated as

$$\max_{\lambda} \{(-x^\top \tilde{r}^\xi - \tau) \cdot \lambda : \lambda \leq 1, \lambda \geq 0\}.$$

Hence, the optimal solution of this dual problem is

$$\lambda^* = \begin{cases} 1 & \text{if } -\bar{x}^\top \bar{r}^\xi - \bar{\tau} \geq 0 \\ 0 & \text{if not} \end{cases}$$

In other words, there is a hyperplane separating  $\Omega$ , where the dual variables of the subproblem  $\mathcal{Q}((\bar{x}, \bar{\tau}), \xi)$  have the same value for a given pair  $(\bar{x}, \bar{\tau})$ .

Therefore, from a partition  $\mathcal{P}^{(t)}$  of  $\Omega$ , we can compute  $r_P = \mathbb{E}[\tilde{r}^\xi | P]$  in  $p_P = \mathbb{P}(P)$  and solve the aggregated problem

$$\min_{x, \tau} \left\{ \tau + \frac{1}{\delta} \sum_{P \in \mathcal{P}^{(t)}} p_P \cdot z_P : Ax \leq b, z_P \geq -x^\top r_P - \tau, z_P \geq 0 \forall P \in \mathcal{P}^{(t)} \right\}$$

Given the optimal solution  $(\bar{x}^{(t)}, \bar{\tau}^{(t)})$  of this problem, we can split each  $P \in \mathcal{P}^{(t)}$  into subsets  $P' = P \cap \{\xi : -\bar{x}^{(t)\top} \tilde{r}^\xi \geq \bar{\tau}^{(t)}\}$  and  $P'' = P \cap \{\xi : -\bar{x}^{(t)\top} \tilde{r}^\xi \leq \bar{\tau}^{(t)}\}$  to obtain a new partition.

*Case study:* For the computational test, we solve the classic portfolio problem, where  $x$  represents the fraction of the portfolio assigned to each investment and the constraints of the first stage are  $x^\top e = 1, x \geq 0$ , ensuring to invest the whole budget in non-negative fractions. Additionally, we assume that returns  $\tilde{r}$  of each investment follow a multivariate normal distribution  $\tilde{r}^\xi \rightsquigarrow \mathcal{N}(\mu, \Sigma)$  using historical data for stocks listed on the SP500, as in [30] and [9].

Note that in each iteration, given  $(\bar{x}^{(t)}, \bar{\tau}^{(t)})$ , we can compute an upper bound for the problem expressed as

$$(11) \quad \text{CVaR}_\delta(\bar{x}^{(t)\top} \bar{r}^\xi) := \mu^\top \bar{x}^{(t)} + \frac{\sigma}{\delta} \phi(\Phi^{-1}(\delta))$$

where  $\sigma = \bar{x}^{(t)\top} \Sigma \bar{x}^{(t)}$  and  $\phi$  and  $\Phi$  are the standard normal p.d.f and standard normal quantile, respectively.

TABLE 1. Results for the CVaR portfolio example

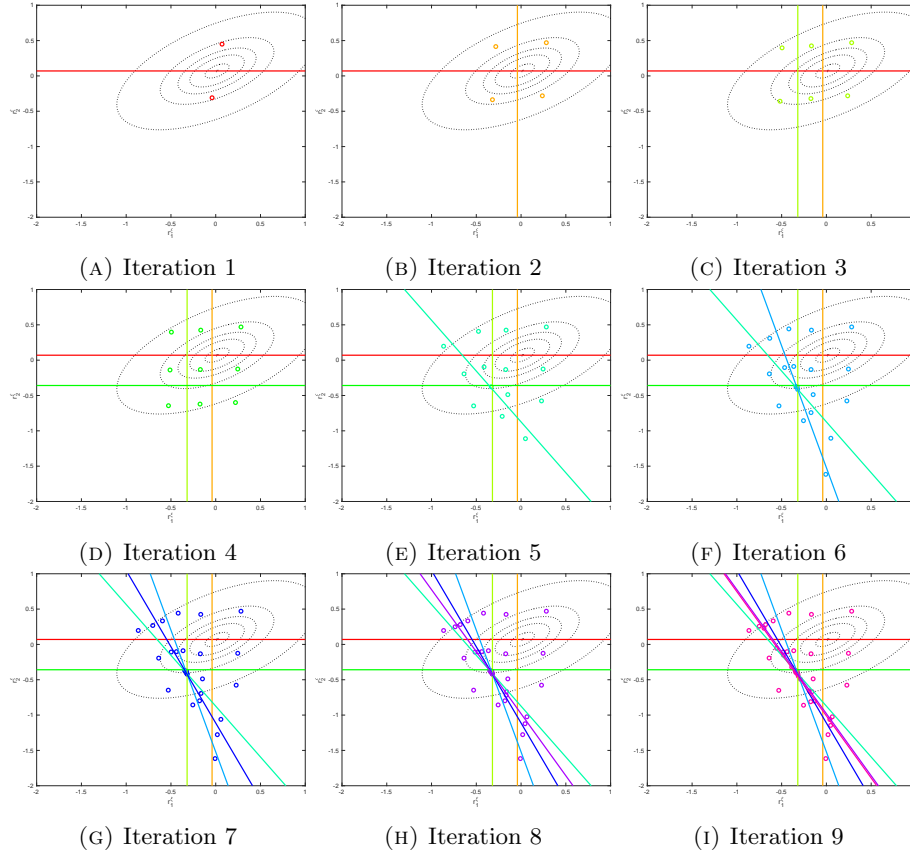
Iter	LB	UB	Gap	$ \mathcal{P}^{(t)} $	$x_1$	$x_2$
1	-0.0702	0.7641	109.184%	1	0	1
2	0.0408	0.6054	93.2602%	2	1	0
3	0.3196	0.6054	47.2124%	4	1	0
4	0.3585	0.7641	40.7887%	6	0	1
5	0.4584	0.5104	10.1866%	9	0.59	0.41
6	0.5001	0.5222	2.0277%	14	0.7752	0.2248
7	0.5043	0.5095	1.0259%	20	0.6834	0.3166
8	0.5070	0.5082	0.2305%	27	0.6371	0.3629
9	0.5082	0.5082	0.0039%	34	0.6375	0.3625

We solve the problem using two stocks and a risk level of  $\delta = 0.1$  to provide a graphical representation of the algorithm. To estimate the probabilities and expected return of each region, we use a Monte Carlo sampler of the underlying distribution. [Table 1](#) shows the results for our instance. We can notice that the problem converges quickly to the optimal solution, as well as in the previous LandS example. A more detailed analysis can be seen in [Fig. 2](#), where the region  $\Omega$  is presented, with ellipses indicating the 50%, 80%, 90%, 95% and 99% confidence intervals of the normal bidimensional distribution. In the first 4 iterations, the aggregated model considers only the riskiest scenario (bottom left dot) and invests the entire portfolio in the stock with the highest return  $r'$ . Our algorithm generates a cut that divides the uncertainty region into  $r_i^\xi \geq r'$  and  $r_i^\xi \leq r'$ , where  $i$  is the stock where the budget is invested. After Iteration 5, the portfolio starts to combine stocks, and the region of interest is divided more precisely to obtain a better estimation of the optimal problem solution.

At last, we remark that, in both computational examples, several algorithmic improvements can be implemented to solve larger and more complex problems (e.g., reaggregating regions with the same duals, considering only the last  $k$  cuts, or subdividing only the active regions; see [\[15\]](#) for more details). Nevertheless, our purpose is simply to show how the method can automatically divide a continuous random space, to iteratively define the regions of interest for the problem and converge to the optimal solution.

## 6. CONCLUSIONS

We present a generalization of the adaptive partition-based method for solving two-stage stochastic problems that contributes to extend the method to a more general setting, particularly, to consider continuous distributions of the uncertain parameters. The resulting algorithm allows to tackle this type of problems, by automatically disaggregating the uncertainty space and solving a discrete (tractable)



3. Suresh Barik, Mahendra Prasad Biswal, and Debashish Chakravarty, *Two-stage stochastic programming problems involving some continuous random variables*, Journal of Uncertain Systems **7** (2013), no. 4, 277–288.
4. Daniel Bienstock and Mark Zuckerberg, *Solving LP relaxations of large-scale precedence constrained problems*, International Conference on Integer Programming and Combinatorial Optimization, Springer, 2010, pp. 1–14.
5. Teodor G Crainic, Francesca Maggioni, Guido Perboli, and Walter Rei, *Reduced cost-based variable fixing in two-stage stochastic programming*, Annals of Operations Research (2018), 1–37.
6. Daniel Espinoza and Eduardo Moreno, *A primal-dual aggregation algorithm for minimizing conditional value-at-risk in linear programs*, Computational Optimization and Applications **59** (2014), no. 3, 617–638.
7. Julia L Hingle and Suvrajeet Sen, *Stochastic decomposition: An algorithm for two-stage linear programs with recourse*, Mathematics of operations research **16** (1991), no. 3, 650–669.
8. Anton J Kleywegt, Alexander Shapiro, and Tito Homem-de Mello, *The sample average approximation method for stochastic discrete optimization*, SIAM Journal on Optimization **12** (2002), no. 2, 479–502.
9. Guido Lagos, Daniel Espinoza, Eduardo Moreno, and Juan Pablo Vielma, *Restricted risk measures and robust optimization*, European Journal of Operational Research **241** (2015), no. 3, 771–782.
10. Pierre LECuyer, Zdravko I Botev, and Dirk P Kroese, *On a generalized splitting method for sampling from a conditional distribution*, 2018 Winter Simulation Conference (WSC), IEEE, 2018, pp. 1694–1705.
11. Claude Lemaréchal, Arkadii Nemirovskii, and Yurii Nesterov, *New variants of bundle methods*, Mathematical programming **69** (1995), no. 1-3, 111–147.
12. Jeff Linderoth, Alexander Shapiro, and Stephen Wright, *The empirical behavior of sampling methods for stochastic programming*, Annals of Operations Research **142** (2006), no. 1, 215–241.
13. Francois V. Louveaux and Yves Smeers, *Optimal investments for electricity generation: A stochastic model and a test problem*, Numerical techniques for stochastic optimization problems (Y. Ermoliev and R. J-B. Wets, eds.), Springer-Verlag, 1988.
14. Scherer Matthias and Mai Jan-frederik, *Simulating copulas: stochastic models, sampling algorithms, and applications*, vol. 4, World Scientific, 2012.
15. Gonzalo Muñoz, Daniel Espinoza, Marcos Goycoolea, Eduardo Moreno, Maurice Queyranne, and Orlando Rivera Letelier, *A study of the Bienstock–Zuckerberg algorithm: applications in mining and resource constrained project scheduling*, Computational Optimization and Applications **69** (2018), no. 2, 501–534.
16. Wellington Oliveira, Claudia Sagastizábal, and Susana Scheimberg, *Inexact bundle methods for two-stage stochastic programming*, SIAM Journal on Optimization **21** (2011), no. 2, 517–544.
17. Babak Saleck Pay and Yongjia Song, *Partition-based decomposition algorithms for two-stage stochastic integer programs with continuous recourse*, Annals of Operations Research (2017), 1–22.
18. Mario VF Pereira and Leontina MVG Pinto, *Multi-stage stochastic optimization applied to energy planning*, Mathematical programming **52** (1991), no. 1-3, 359–375.
19. Ragheb Rahmaniani, Teodor Crainic, Michel Gendreau, and Walter Rei, *A benders decomposition method for two-stage stochastic network design problems*, CIRRELT, Centre interuniversitaire de recherche sur les réseaux d’entreprise, 2017.
20. Ragheb Rahmaniani, Teodor Gabriel Crainic, Michel Gendreau, and Walter Rei, *Accelerating the benders decomposition method: Application to stochastic network design problems*, SIAM Journal on Optimization **28** (2018), no. 1, 875–903.
21. ———, *An asynchronous parallel benders decomposition method*, CIRRELT, 2019.
22. R Tyrrell Rockafellar, Stanislav Uryasev, et al., *Optimization of conditional value-at-risk*, Journal of risk **2** (2000), 21–42.
23. R Tyrrell Rockafellar and Roger J-B Wets, *Scenarios and policy aggregation in optimization under uncertainty*, Mathematics of operations research **16** (1991), no. 1, 119–147.
24. Andrzej Ruszczyński, *A regularized decomposition method for minimizing a sum of polyhedral functions*, Mathematical programming **35** (1986), no. 3, 309–333.

25. Kevin Ryan, Deepak Rajan, and Shabbir Ahmed, *Scenario decomposition for 0-1 stochastic programs: Improvements and asynchronous implementation*, Parallel and Distributed Processing Symposium Workshops, 2016 IEEE International, IEEE, 2016, pp. 722–729.
26. Murwan Siddig and Yongjia Song, *Adaptive partition-based SDDP algorithms for multistage stochastic linear programming*, Tech. report, preprint arXiv:1908.11346, 2019.
27. Yongjia Song and James Luedtke, *An adaptive partition-based approach for solving two-stage stochastic programs with fixed recourse*, SIAM Journal on Optimization **25** (2015), no. 3, 1344–1367.
28. Wim van Ackooij, Welington de Oliveira, and Yongjia Song, *Adaptive partition-based level decomposition methods for solving two-stage stochastic programs with fixed recourse*, Inform Journal on Computing **30** (2017), no. 1, 57–70.
29. Richard M Van Slyke and Roger Wets, *L-shaped linear programs with applications to optimal control and stochastic programming*, SIAM Journal on Applied Mathematics **17** (1969), no. 4, 638–663.
30. Juan Pablo Vielma, Shabbir Ahmed, and George L Nemhauser, *A lifted linear programming branch-and-bound algorithm for mixed-integer conic quadratic programs*, INFORMS Journal on Computing **20** (2008), no. 3, 438–450.
31. Jean-Paul Watson and David L Woodruff, *Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems*, Computational Management Science **8** (2011), no. 4, 355–370.
32. Christian Wolf, Csaba I Fábián, Achim Koberstein, and Leena Suhl, *Applying oracles of on-demand accuracy in two-stage stochastic programming—a computational study*, European Journal of Operational Research **239** (2014), no. 2, 437–448.
33. Victor Zverovich, Csaba I Fábián, Eldon FD Ellison, and Gautam Mitra, *A computational study of a solver system for processing two-stage stochastic lps with enhanced benders decomposition*, Mathematical Programming Computation **4** (2012), no. 3, 211–238.

FACULTY OF ENGINEERING AND SCIENCES, UNIVERSIDAD ADOLFO IBÁÑEZ, SANTIAGO, CHILE  
 E-mail address: `cristian.ramirez@edu.uai.cl`, `eduardo.moreno@uai.cl`