# Generalized adaptive partition-based method for two-stage stochastic linear programs : geometric oracle and analysis

Maël Forcier, Vincent Leclère
CERMICS, École des Ponts, Marne-la-Vallée, France

**Abstract**

Adaptive Partition-based Methods (APM) are numerical methods that solve, in particular, two-stage stochastic linear problems (2SLP). We say that a partition of the uncertainty space is *adapted* to the current first stage control $\check{x}$ if we can aggregate scenarios while conserving the true value of the expected recourse cost at $\check{x}$. The core idea of APM is to iteratively constructs an *adapted* partition to all past tentative first stage controls. Relying on the normal fan of the dual admissible set, we give a necessary and sufficient condition for a partition to be adapted even for non-finite distribution, and provide a geometric method to obtain an adapted partition. Further, by showing the connection between APM and the L-shaped algorithm, we prove convergence and complexity bounds of the APM methods. The paper presents the fixed recourse case and ends with elements to forgo this assumption.

## 1. Introduction

Stochastic programming is a powerful modeling paradigm for optimization under uncertainty that has found many applications in energy, logistics or finance (see e.g. [1]). Two-stage linear stochastic programs (2SLP) constitute an important class of stochastic programs. They have been thoroughly studied, see e.g. [2] and references therein. One reason for this interest is the availability of efficient linear solvers and the use of dedicated algorithms leveraging the special structure of linear stochastic programs ([3, 2, 4]).

### 1.1. Setting

We denote random variables as bold letters (e.g. $\boldsymbol{\xi}$) and their realization as normal scripts (e.g. $\xi$). We consider the following problem 2-stage stochastic linear problem with fixed recourse:

$$\min_{x\in\mathbb{R}^n_+} \Big\{ c^\top x + \underbrace{\mathbb{E}\big[Q(x,\boldsymbol{\xi})\big]}_{:=V(x)} \mid Ax=b \Big\}, \qquad \text{(2SLP)}$$

where the expectation is with respect to $\boldsymbol{\xi}=(\boldsymbol{T},\boldsymbol{h})$ an integrable random variable on $(\Omega,\mathcal{A},\mathbb{P})$ taking values in $\Xi\subset\mathbb{R}^{\ell\times n}\times\mathbb{R}^\ell$, and the recourse cost is

$$Q(x,\xi) := \min_{y\in\mathbb{R}^m_+} \Big\{ q^\top y \mid Tx+Wy=h \Big\}. \qquad (1)$$

The dual formulation of the recourse problem is

$$Q^D(x,\xi) := \max_{\lambda\in\mathbb{R}^\ell} \Big\{ (h-Tx)^\top\lambda \mid W^\top\lambda \leqslant q \Big\}. \qquad (2)$$

We define

$$X := \{x\in\mathbb{R}^n_+ \mid Ax=b\},$$

$$D := \{\lambda\in\mathbb{R}^\ell \mid W^\top\lambda\leqslant q\}.$$

In the rest of the paper, we assume that $D\neq\emptyset$ which implies by duality: $Q(x,\xi)=Q^D(x,\xi)$.

For the sake of simplicity, we assume throughout the paper that we are in a *relatively complete recourse* setting, that is $X\subset\mathrm{dom}(V)$. Most results can be obtained without this assumption if we add feasibility cuts (see Section 3.2).

### 1.2. Literature review

Most results for 2SLP with continuous distributions rely on discretizing the distributions. The Sample Average Approximation (SAA) method samples the costs and constraints. It relies on probabilistic results based on uniform laws of large numbers to give statistical guarantees, see [4, Chap. 5] for details. Obtaining an approximation with satisfying guarantees requires a large number of scenarios. Otherwise, when the support of the random variables are simplices, we can leverage convexity inequalities (like Jensen's and Edmundson-Madansky's) or moments inequalities to construct finite scenario trees such that the discretized problem yields upper or lower bound of the continuous one (see *e.g.* [5, 6]).

In each of these approaches, we solve an approximate version of the stochastic program, with or without guarantees. In any case, the number of scenarios increase the numerical burden of 2SLP.

In order to alleviate the computations, we can use scenario reduction methods. Some are based on heuristics, aiming at matching properties of the underlying distribution (e.g. matching moments), others are based on adequate distances on the scenario tree (see [7, 8]). Alternatively, APM methods iteratively solve an aggregated version of 2SLP over a partition of the uncertainty space by replacing each subset of scenarios by its weighted mean. We say that a partition is *adapted* to a first-stage control $x$ if the aggregated recourse problem has the same optimal objective value as the recourse problem with the original distribution. After solving an aggregated 2SLP, an APM method call a *(adapted) partition oracle* to define a new (adapted) partition at the current first-stage control. APM were first introduced by Song and Luedtke [9], who gave a partition oracle designed for fixed recourse 2SLP with finitely supported random variables. Van Ackooij, de Oliveira and Song [10] improved the performance of APM by combining it with level decomposition methods ideas. Finally, Ramirez-Pico and Moreno extended the scope of APMs, under the name GAPM, in [11] to problems with continuous distributions for the right-hand side and technology matrix (and fixed recourse cost vector and matrix). They gave a sufficient condition for a partition to be adapted. They also provided adapted partition oracles for some specific problems.

### 1.3. Contributions

The main contribution of the paper are the following: i) using polyhedral geometry tools we provide a generic adapted partition oracle, ii) we give a new necessary and sufficient condition for a partition to be adapted to $\check{x}$ even in the non-finitely supported case, iii) by casting APM methods as accelerated L-Shaped algorithms where tangeant cones are added instead of tangeant planes (affine cuts), we give convergence and complexity results for APM methods.

### 1.4. Structure of the paper

Section 2 presents the APM framework and a necessary and sufficient condition for a partition to be adapted to $\check{x}$. Section 3 uses the link between APM and L-Shaped to obtain convergence and complexity results. Finally, Section 4 presents numerical results, while Section 5 briefly extends GAPM to non-fixed recourse problem.

## 2. General framework and geometric oracle

In this section, we start by presenting a generic framework for APM algorithms, which depends on partition oracle choice.

We proceed by giving a necessary and sufficient condition for a partition oracle to be adapted, and then a geometric adapted partition oracle.

### 2.1. Partition, refinements and APM framework

A partition $\mathcal{P}$ of $\Xi$ is a collection of non-empty pairwise disjoint subsets covering $\Xi$, i.e. $\cup_{P\in\mathcal{P}} P = \Xi$, $P \cap P' = \emptyset$ and $P \neq \emptyset$ for $P \neq P' \in \mathcal{P}$. Let $P$ be a measurable subset of $\Xi$. We denote by $\mathbb{E}\big[\cdot|P\big]$ the conditional expectation $\mathbb{E}\big[\cdot|\boldsymbol{\xi} \in P\big]$ and $\mathbb{P}\big[P\big]$ the probability $\mathbb{P}\big[\boldsymbol{\xi} \in P\big]$. We say that two measurable subsets of $E, F \subset \Xi$ are $\mathbb{P}$-*equivalent*, denoted $E \sim_{\mathbb{P}} F$, if and only if they differ by a $\mathbb{P}$-negligeable set

$$E \sim_{\mathbb{P}} F \iff \mathbb{P}\big[E \cap F\big] = \mathbb{P}\big[E\big] = \mathbb{P}\big[F\big],$$

similarly we denote

$$E \subset_{\mathbb{P}} F \iff \mathbb{P}\big[E \cap F\big] = \mathbb{P}\big[E\big].$$

A $\mathbb{P}$-partition of $\Xi$ is the equivalence class of all partitions that are $\mathbb{P}$-equivalent.

Let $\mathcal{P}$ and $\mathcal{R}$ be two $\mathbb{P}$-partitions of $\Xi$. We say that $\mathcal{P}$ refines $\mathcal{R}$, denoted $\mathcal{P} \preccurlyeq_{\mathbb{P}} \mathcal{R}$, if

$$\forall P \in \mathcal{P},\ \exists R \in \mathcal{R}, \quad P \subset_{\mathbb{P}} R,$$

The *common refinement* of $\mathcal{P}$ and $\mathcal{R}$ is given by

$$\mathcal{P} \wedge \mathcal{R} = \{P \cap R \,|\, P \in \mathcal{P}, R \in \mathcal{R}\}.$$

**Definition 1** (Expected recourse cost of partition). *For $\mathcal{P}$ a $\mathbb{P}$-partition of $\Xi \subset \mathbb{R}^{\ell\times m} \times \mathbb{R}^{\ell}$ we define*

$$V_{\mathcal{P}} : x \mapsto \sum_{P\in\mathcal{P}} \mathbb{P}\big[P\big] Q\big(x, \mathbb{E}\big[\boldsymbol{\xi}|P\big]\big). \tag{4}$$

*Let $\check{x} \in \mathrm{dom}(V)$. We say that a $\mathbb{P}$-partition $\mathcal{P}$ is adapted to $\check{x}$ if $V_{\mathcal{P}}(\check{x}) = V(\check{x}) := \mathbb{E}\big[Q(\check{x}, \boldsymbol{\xi})\big]$.*

The following lemma shows that, by convexity, a finer partition yields a larger expected cost-to-go function.

**Lemma 2.** *Let $\mathcal{P}$ and $\mathcal{R}$ two $\mathbb{P}$-partitions of $\Xi$ then*

$$\mathcal{P} \preccurlyeq_{\mathbb{P}} \mathcal{R} \quad \Longrightarrow \quad V_{\mathcal{P}} \geqslant V_{\mathcal{R}}. \tag{5}$$

*Moreover,*

$$V_{\mathcal{P}\wedge\mathcal{R}} \geqslant \max(V_{\mathcal{P}}, V_{\mathcal{R}}). \tag{6}$$

*In particular,*

$$Q\big(\cdot, \mathbb{E}\big[\boldsymbol{\xi}\big]\big) \leqslant V_{\mathcal{P}} \leqslant V. \tag{7}$$

*Proof.* Since $\mathcal{P}$ is a $\mathbb{P}$-partition $\mathbb{1}_{\boldsymbol{\xi}\in R} = \sum_{P\in\mathcal{P}} \mathbb{1}_{\boldsymbol{\xi} R\cap P}$ almost surely. For any measurable set $E \subset \Xi$, $\mathbb{E}\big[\boldsymbol{\xi}\mathbb{1}_{\boldsymbol{\xi}\in E}\big] = \mathbb{P}\big[E\big]\mathbb{E}\big[\boldsymbol{\xi}|E\big]$. We then have,

$$\begin{aligned}\mathbb{P}\big[R\big]\mathbb{E}\big[\boldsymbol{\xi}|R\big] &= \mathbb{E}\big[\boldsymbol{\xi}\mathbb{1}_{\boldsymbol{\xi}\in R}\big] = \sum_{P\in\mathcal{P}} \mathbb{E}\big[\boldsymbol{\xi}\mathbb{1}_{\boldsymbol{\xi} R\cap P}\big] \\ &= \sum_{P\in\mathcal{P}} \mathbb{P}\big[R\cap P\big]\mathbb{E}\big[\boldsymbol{\xi}|R\cap P\big]\end{aligned}$$

When $\mathbb{P}\big[R\big] > 0$, by dividing this equation by $\mathbb{P}\big[R\big]$, we obtain that $\mathbb{E}\big[\boldsymbol{\xi}|R\big]$ is equal to the convex combination $\sum_{P\in\mathcal{P}} \frac{\mathbb{P}[R\cap P]}{\mathbb{P}[R]}\mathbb{E}\big[\boldsymbol{\xi}|R\cap P\big]$. Finally, consider $\check{x} \in X$, the convexity of $\xi \mapsto Q(\check{x}, \xi)$ yields

$$Q\big(\check{x}, \mathbb{E}\big[\boldsymbol{\xi}|R\big]\big) \leqslant \sum_{P\in\mathcal{P}} \frac{\mathbb{P}\big[P\cap R\big]}{\mathbb{P}\big[R\big]} Q\big(\check{x}, \mathbb{E}\big[\boldsymbol{\xi}|P\cap R\big]\big)\,.$$

Then, if $\mathcal{P} \preccurlyeq_{\mathbb{P}} \mathcal{R}$,

$$\begin{aligned}V_{\mathcal{R}}(\check{x}) &= \sum_{R\in\mathcal{R}} \mathbb{P}\big[R\big] Q\big(\check{x}, \mathbb{E}\big[\boldsymbol{\xi}|R\big]\big) \\ &\leqslant \sum_{P\in\mathcal{P}}\sum_{R\in\mathcal{R}} \mathbb{P}\big[P\cap R\big] Q\big(\check{x}, \mathbb{E}\big[\boldsymbol{\xi}|P\cap R\big]\big) \\ &= \sum_{P\in\mathcal{P}} \mathbb{P}\big[P\big] Q\big(\check{x}, \mathbb{E}\big[\boldsymbol{\xi}|P\big]\big) = V_{\mathcal{P}}(\check{x})\end{aligned}$$

The last line follows from the fact, that for $P \in \mathcal{P}$, with $\mathbb{P}\big[P > 0\big]$ and $\mathcal{P} \preccurlyeq_{\mathbb{P}} \mathcal{R}$, there exists a unique $R \in \mathcal{R}$ such that $\mathbb{P}\big[P\cap R\big] = \mathbb{P}\big[P\big]$, all other $R \in \mathcal{R}$ being such that $\mathbb{P}\big[P\cap R\big] = 0$.

Eq. (6) is a direct consequence of $V_{\mathcal{P}\wedge\mathcal{P}'} \geqslant V_{\mathcal{P}}$ and $V_{\mathcal{P}\wedge\mathcal{P}'} \geqslant V_{\mathcal{P}'}$. Thus, $V_{\mathcal{P}\wedge\mathcal{P}'} \geqslant \max(V_{\mathcal{P}}, V_{\mathcal{P}'})$. Coupling this result with $\mathcal{P} \preccurlyeq_{\mathbb{P}} \{\Xi\}$ yield the left inequality of Eq. (7) while the other can be found in [11, Prop. 1]. ☐

With those definitions we present in Algorithm 1 a generic framework for APM methods.

$k \leftarrow 0$, $z_U^0 \leftarrow +\infty$, $z_L^0 \leftarrow -\infty$, $\mathcal{P}^0 \leftarrow \{\Xi\}$ ;
**while** $z_U^k - z_L^k > \varepsilon$ **do**
  $k \leftarrow k+1$;
  Solve $z_L^k \leftarrow \min_{x\in X} c^\top x + V_{\mathcal{P}^{k-1}}(x)$ and let $x^k$ be an optimal solution ;
  Call the oracle on $x^k$ yielding $\mathcal{P}_{x^k}$ ;
  $\mathcal{P}^k \leftarrow \mathcal{P}^{k-1} \wedge \mathcal{P}_{x^k}$ ;
  $z_U^k \leftarrow \min\Big(z_U^{k-1}, c^\top x^k + V_{\mathcal{P}^k}(x^k)\Big)$ ;
**end**

**Algorithm 1:** Generic framework for APM.

### *2.2. Coarsest adapted partition*

In this section, we define $\mathcal{R}_{\check{x}}$, a particular $\mathbb{P}$-partition, and prove that it is, in a generic case, the coarsest partition adapted to $\check{x} \in X$, *i.e.*, the only partition adapted to $\check{x}$ that refines $\mathcal{R}_{\check{x}}$ is $\mathcal{R}_{\check{x}}$ itself. Indeed, we are looking for partitions that yields a precise approximation of recourse cost (exact at $\check{x}$ in the adapted case), while having the smallest possible number of elements.

When the distributions have finite support, [9] characterized the partitions adapted to $\check{x}$. Building on this result, a sufficient condition for continuous distribution can be found in [11, Prop. 2]. We now prove that, for any distribution, a partition is adapted to $\check{x}$ if and only if it refines the collection $\overline{\mathcal{R}}_{\check{x}}$ defined in (10b). Unfortunately, $\overline{\mathcal{R}}_{\check{x}}$ is not necessarily a $\mathbb{P}$-partition, thus we also provide a partition $\mathcal{R}_{\check{x}} \preccurlyeq \overline{\mathcal{R}}_{\check{x}}$ (see Figure 1 for an illustration).

Recall that $D = \{\lambda \in \mathbb{R}^\ell \mid W^\top\lambda \leqslant q\}$ and that the normal cone of $D$ at $\lambda$ is the set $N_D(\lambda) := \{\psi \in \mathbb{R}^\ell \,|\, \psi^\top(\lambda' - \lambda) \leqslant 0, \forall\lambda' \in D\}$. We denote by $\mathrm{ri}(N)$ the relative interior of a cone $N$. Let $\mathcal{N}(D) := \{N_D(\lambda) \,|\, \lambda \in D\}$ be the normal fan of $D$, *i.e.*, the (finite) collection of all normal cones of $D$ (see [12] for an introduction on normal fans and their use in stochastic programming). We denote by $\mathcal{N}(D)^{\max} := \{N \in \mathcal{N}(D) \,|\, \forall N' \in \mathcal{N}(D), N \subset N' \Rightarrow N = N'\}$ the collection of the maximal elements of $\mathcal{N}(D)$ (*i.e.*, full dimensional cones up-to lineality spaces).

**Theorem 3.** *Fix $\check{x} \in \mathrm{dom}(V)$ and $N$ a cone in $\mathbb{R}^m$. We define $E_{N,\check{x}}$ and $\overline{E}_{N,\check{x}}$, subsets of $\Xi$, as*

$$E_{N,\check{x}} := \{\xi \in \Xi \mid h - T\check{x} \in \mathrm{ri}(N)\} \tag{9a}$$

$$\overline{E}_{N,\check{x}} := \{\xi \in \Xi \mid h - T\check{x} \in N\} \tag{9b}$$

*We define $\mathcal{R}_{\check{x}}$ and $\overline{\mathcal{R}}_{\check{x}}$ as*

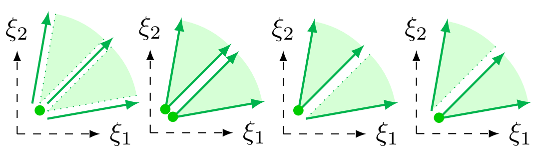
$$\mathcal{R}_{\check{x}} := \big\{E_{N,\check{x}} \mid N \in \mathcal{N}(D)\big\} \tag{10a}$$

$$\overline{\mathcal{R}}_{\check{x}} := \big\{\overline{E}_{N,\check{x}} \mid N \in \mathcal{N}(D)^{\max}\big\}. \tag{10b}$$

*Then,*

$$\mathcal{P} \preccurlyeq_{\mathbb{P}} \mathcal{R}_{\check{x}} \Longrightarrow V_{\mathcal{P}}(\check{x}) = V(\check{x})$$

$$\mathcal{P} \preccurlyeq_{\mathbb{P}} \overline{\mathcal{R}}_{\check{x}} \iff V_{\mathcal{P}}(\check{x}) = V(\check{x}).$$




(a) $\mathcal{R}_{\check{x}}$ (b) $\overline{\mathcal{R}}_{\check{x}}$ (c) $\mathcal{P}$ (d) $\mathcal{P}'$

Figure 1: $\mathcal{R}_{\check{x}}$ is a partition of $\Xi$ into 6 elements, $\overline{\mathcal{R}}_{\check{x}}$ is not a partition, $\mathcal{P}$ and $\mathcal{P}'$ are two distinct coarsest partitions (into 2 elements) with $\mathcal{R}_{\check{x}} \preccurlyeq \mathcal{P} \preccurlyeq \overline{\mathcal{R}}_{\check{x}}$ and $\mathcal{R}_{\check{x}} \preccurlyeq \mathcal{P}' \preccurlyeq \overline{\mathcal{R}}_{\check{x}}$.

**Remark 4.** *When the distribution of $\boldsymbol{\xi}$ is absolutely continuous with respect to the Lebesgue measure of $\Xi$, $\mathcal{R}_{\check{x}} \sim_{\mathbb{P}} \overline{\mathcal{R}}_{\check{x}}$, thus $\mathcal{R}_{\check{x}}$ is the coarsest partition adapted to $\check{x} \in \mathrm{dom}(V)$.*

*If $\boldsymbol{\xi}$ does not admit a density, $\mathcal{R}_{\check{x}}$ is still an adapted partition but not necessarily the coarsest, which might not exist (see Fig. 1). Nevertheless, any adapted partition should refine $\overline{\mathcal{R}}_{\check{x}}$. Unfortunately, we cannot use $\overline{\mathcal{R}}_{\check{x}}$ in Algorithm 1, as we cannot guarantee that $\overline{\mathcal{R}}_{\check{x}}$ is a $\mathbb{P}$-partition.*

**Remark 5.** *Note that Proposition 2 of [11] implies that all partition oracle returning partitions satisfying assumption (7) of [11] must be refinements of $\overline{\mathcal{R}}_{\check{x}}$ by Theorem 3. In the finite scenario case, our adaptedness condition is equivalent to Song and Luedtke's condition [9].*

We preclude the proof by a technical lemma.

**Lemma 6.** *Consider a set $P \subset \Xi$ such that $\mathbb{P}(P) > 0$, and a first-stage control $\check{x} \in \mathrm{dom}(V)$. Then,*

$$\exists R \in \mathcal{R}_{\check{x}}, \quad P \subset_{\mathbb{P}} R \Longrightarrow Q(\check{x}, \mathbb{E}\big[\boldsymbol{\xi}|P\big]) = \mathbb{E}\big[Q(\check{x},\boldsymbol{\xi})|P\big],$$

$$\exists \overline{R} \in \overline{\mathcal{R}}_{\check{x}}, \quad P \subset_{\mathbb{P}} \overline{R} \iff Q(\check{x}, \mathbb{E}\big[\boldsymbol{\xi}|P\big]) = \mathbb{E}\big[Q(\check{x},\boldsymbol{\xi})|P\big].$$

*Proof.* Since $\exists R \in \mathcal{R}_{\check{x}},\ P \subset_{\mathbb{P}} R$ implies $\exists \overline{R} \in \overline{\mathcal{R}}_{\check{x}},\ P \subset_{\mathbb{P}} \overline{R}$, we only need to prove the second equivalence.

($\Rightarrow$) Let $P$ be such that there exists $N \in \mathcal{N}(D)$ with $P \subset_{\mathbb{P}} \overline{E}_{N,\check{x}}$. By definition of $\mathcal{N}(D)$, there exists a dual point $\lambda_N \in D$ such that $N$ is the normal cone of $D$ at $\lambda_N$. By definition of a normal cone, for all $\psi \in N$ and all $\lambda \in D$, $\psi^\top(\lambda - \lambda_N) \leqslant 0$. In other words $\psi^\top\lambda_N = \max_{\lambda\in D}\psi^\top\lambda$.

As $P \subset \overline{E}_{N,\check{x}}$, for $\mathbb{P}$-almost-all $\xi \in P$, we have $h - T\check{x} \in N$. Recall that $Q(\check{x},\xi) = \sup_{\lambda\in D}(h - T\check{x})^\top\lambda$, thus, $Q(\check{x},\xi) = (h - T\check{x})^\top\lambda_N$. Hence,

$$\mathbb{E}\big[Q(\check{x},\boldsymbol{\xi})|P\big] = \mathbb{E}\big[(\boldsymbol{h} - \boldsymbol{T}\check{x})^\top\lambda_N|P\big]$$
$$= \mathbb{E}\big[\boldsymbol{h} - \boldsymbol{T}\check{x}|P\big]^\top\lambda_N = Q(\check{x},\mathbb{E}\big[\boldsymbol{\xi}|P\big])$$

as $N$ is convex and $\mathbb{E}\big[\boldsymbol{h} - \boldsymbol{T}\check{x}|P\big] \in N$.

($\Leftarrow$) For $\psi \in \mathbb{R}^\ell$, we denote the face $D^\psi := \mathrm{argmax}_{\lambda\in D}\psi^\top\lambda$. Note that, for all $\psi,\psi' \in \mathrm{ri}(N)$, with $N \in \mathcal{N}(D)$, we have $D^N := D^\psi = D^{\psi'}$.

Assume that there is no $R \in \overline{\mathcal{R}}_{\check{x}}$ such that $P \subset_{\mathbb{P}} R$. Then, for all $R \in \overline{\mathcal{R}}_{\check{x}}$, $\mathbb{P}\big[P \cap R\big] < \mathbb{P}\big[P\big]$. Since $\mathbb{P}\big[P\big] \leqslant \sum_{R\in\overline{\mathcal{R}}_{\check{x}}}\mathbb{P}\big[P\cap R\big]$, there exist $R_1$ and $R_2$ in $\overline{\mathcal{R}}_{\check{x}}$ such that $\mathbb{P}\big[P\cap R_1\big] > 0$ and $\mathbb{P}\big[P\cap R_2\big] > 0$. Let $\lambda \in D$ such that $Q(\check{x},\mathbb{E}\big[\boldsymbol{\xi}|P\big]) = \mathbb{E}\big[\boldsymbol{h} - \boldsymbol{T}\check{x}|P\big]^\top\lambda$ *i.e.*, $\lambda \in D^{\mathbb{E}[\boldsymbol{h}-\boldsymbol{T}\check{x}|P]}$. Let $N_1$ and $N_2 \in \mathcal{N}(D)^{\max}$ be such that $R_1 = \overline{E}_{N_1,\check{x}}$ and $R_2 = \overline{E}_{N_2,\check{x}}$. Since $N_1 \neq N_2$ are maximal, $D^{N_1} \cap D^{N_2} = \emptyset$. Thus, there exists at least one $i \in \{1,2\}$ such that $\lambda \notin D^{N_i}$. Then, $\mathbb{E}\big[Q(\check{x},\boldsymbol{\xi})|P\cap R_i\big] > \mathbb{E}\big[\boldsymbol{h} - \boldsymbol{T}\check{x}|P\cap R_i\big]^\top\lambda$.

Note that $Q(\check{x},\xi) = \sigma_D(h - T\check{x})$, where $\sigma_D$ is the support function of the polyhedron $D$, thus $\xi \mapsto Q(\check{x},\xi)$ is a polyhedral function. Further, its affine regions are the elements of $\overline{\mathcal{R}}_{\check{x}}$.

By convexity, for any measurable set $A$, $\mathbb{E}\big[Q(\check{x},\boldsymbol{\xi})|P\cap A\big] \geqslant Q(\check{x},\mathbb{E}\big[\boldsymbol{\xi}|P\cap A\big])$ which is equal to $\max_{\lambda'\in D}\mathbb{E}\big[\boldsymbol{h} - \boldsymbol{T}\check{x}|P\cap A\big]^\top\lambda'$. Since $\lambda \in D$, we have $\mathbb{E}\big[Q(\check{x},\boldsymbol{\xi})|P\cap A\big] \geqslant \mathbb{E}\big[\boldsymbol{h} - \boldsymbol{T}\check{x}|P\cap A\big]^\top\lambda$.

Thus, $\mathbb{E}\big[Q(\check{x},\boldsymbol{\xi})|P\big] > Q(\check{x},\mathbb{E}\big[\boldsymbol{\xi}|P\big])$. □

*Proof of Theorem 3.* By definition $\mathcal{P} \preccurlyeq_{\mathbb{P}} \overline{\mathcal{R}}_{\check{x}}$, if and only if, for all $P \in \mathcal{P}$ there exists a cell $R \in \overline{\mathcal{R}}_{\check{x}}$ such that $P \subset_{\mathbb{P}} R$. By Lemma 6 this is equivalent to, for all $P \in \mathcal{P}$, $Q(\check{x},\mathbb{E}\big[\boldsymbol{\xi}|P\big]) = \mathbb{E}\big[Q(\check{x},\boldsymbol{\xi})|P\big]$. Now, by Jensen's inequality, this equality (for all $P \in \mathcal{P}$) is equivalent to the equality of a convex sum like

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

Law of total expectation yields (11). □

**Remark 7.** *Let $x^\star$ be an optimal solution of*

$$\min_{x \in X} \quad c^\top x + V_{\mathcal{P}^\star}(x)$$

*where $\mathcal{P}^\star \preccurlyeq_{\mathbb{P}} \mathcal{R}_{x^\star}$. Then, $x^\star$ is also a solution of Problem* (2SLP)*. In other words, $\mathcal{P}^\star$ is a 0-sufficient partition according to [9, Def. 1.2].*

## 3. Comparison with other algorithms and convergence

In this section, we show that the partition-based methods can be seen as an acceleration of the cutting plane method. It then gives us a finite convergence proof with a bound on the number of steps.

### *3.1. Adapted partition and subdifferential*

We show that, for any first stage control $x \in X$, if the partition is adapted to $x$, then the subdifferential of approximate expected recourse cost coincides with the subdifferential of the true expected recourse cost.

**Lemma 8.** *Let $\check{x} \in \mathrm{dom}(V)$ and $\mathcal{P}$ be a refinement of $\mathcal{R}_{\check{x}}$, i.e. $\mathcal{P} \preccurlyeq \mathcal{R}_{\check{x}}$, then*

$$\partial V_{\mathcal{R}_{\check{x}}}(\check{x}) \subset \partial V_{\mathcal{P}}(\check{x}) \subset \partial V(\check{x})$$

*Furthermore, if $\check{x} \in \mathrm{ri}(\mathrm{dom}(V))$,*

$$\partial V_{\mathcal{R}_{\check{x}}}(\check{x}) = \partial V_{\mathcal{P}}(\check{x}) = \partial V(\check{x})$$

*Proof.* Let $g \in \partial V_{\mathcal{R}_{\check{x}}}(\check{x})$ then for all $x$, $V_{\mathcal{R}_{\check{x}}}(x) \geqslant V_{\mathcal{R}_{\check{x}}}(\check{x}) + g^\top(x - \check{x})$. By monotonicity (see (5)) $V_{\mathcal{P}}(x) \geqslant V_{\mathcal{R}_{\check{x}}}(x)$ and as $\mathcal{R}_{\check{x}}$ is adapted to $\check{x}$, we have $V_{\mathcal{R}_{\check{x}}}(\check{x}) = V(\check{x}) = V_{\mathcal{P}}(\check{x})$. Thus, $V_{\mathcal{P}}(x) \geqslant V_{\mathcal{P}}(\check{x}) + g^\top(x - \check{x})$ and $g \in \partial V_{\mathcal{P}}(\check{x})$. The proof for the second inclusion is similar.

Let $\check{x} \in \mathrm{ri}(\mathrm{dom}(V))$, we now prove that $\partial V_{\mathcal{R}_{\check{x}}}(\check{x}) = \partial V(\check{x})$. Recall that $D^N = D^\psi = \mathrm{argmax}_{\lambda \in D}\, \psi^\top \lambda$, for $\psi \in \mathrm{ri}(N)$ where $N \in \mathcal{N}(D)$. By [4, Prop 2.8 p.37], $\partial V(\check{x}) = \mathbb{E}\big[ - \boldsymbol{T}^\top D^{\boldsymbol{h} - \boldsymbol{T}\check{x}}\big] + N_{\mathrm{dom}(V)}(\check{x})$. Thus, since $\check{x} \in \mathrm{ri}(\mathrm{dom}(V))$,

$$\begin{aligned}
\partial V(\check{x}) &= \mathbb{E}\big[ - \boldsymbol{T}^\top D^{\boldsymbol{h} - \boldsymbol{T}\check{x}}\big] \\
&= \mathbb{E}\Big[ \sum_{N \in \mathcal{N}(D)} -\mathbb{1}_{\boldsymbol{h} - \boldsymbol{T}\check{x} \in \mathrm{ri}(N)} \boldsymbol{T}^\top D^N \Big] \\
&= \mathbb{E}\Big[ \sum_{N \in \mathcal{N}(D)} -\mathbb{1}_{\boldsymbol{\xi} \in E_{N,\check{x}}} \boldsymbol{T}^\top D^N \Big]
\end{aligned}$$

Further,

$$\begin{aligned}
&\mathbb{E}\big[\mathbb{1}_{\boldsymbol{\xi} \in E_{N,\check{x}}} \boldsymbol{T}^\top D^N\big] \\
&\quad = \mathbb{P}\big[E_{N,\check{x}}\big]\mathbb{E}\big[\boldsymbol{T} | E_{N,\check{x}}\big]^\top D^N \\
&\quad = \mathbb{P}\big[E_{N,\check{x}}\big]\mathbb{E}\big[\boldsymbol{T} | E_{N,\check{x}}\big]^\top D^{\mathbb{E}[\boldsymbol{h} - \boldsymbol{T}\check{x} | E_{N,\check{x}}]}
\end{aligned}$$

And by definition of $\mathcal{R}_{\check{x}}$ in (10a) , we get

$$\begin{aligned}
\partial V(\check{x}) &= \sum_{P \in \mathcal{R}_{\check{x}}} -\mathbb{P}\big[P\big]\mathbb{E}\big[\boldsymbol{T}|P\big]^\top D^{\mathbb{E}[\boldsymbol{h} - \boldsymbol{T}\check{x} | P]} \\
&= \sum_{P \in \mathcal{R}_{\check{x}}} \mathbb{P}\big[P\big] \partial_x Q(\check{x}, \mathbb{E}\big[\boldsymbol{\xi}|P\big]) \\
&= \partial V_{\mathcal{R}_{\check{x}}}(\check{x})
\end{aligned}$$

□

### *3.2. Link with L-shaped and Benders decomposition*

The classical L-shaped method (see e.g. [2, Chapter 5]) is a specification of Benders decomposition to 2SLP with finitely supported distributions. The core idea consists in representing the expected recourse cost in (2SLP), by a lift variable

$$\min_{x \in X, \theta \in \mathbb{R}} \Big\{ c^\top x + \theta \mid (x, \theta) \in \mathrm{epi}(V) \Big\}.$$

We then relax the epigraphical representation $(x, \theta) \in \mathrm{epi}(V)$, replacing it by a set of valid inequalities called *cuts*, i.e.

$$\begin{aligned}
\min_{x \in X, \theta \in \mathbb{R}} \quad & c^\top x + \theta \\
\text{s.t.} \quad & g^\top x + v \leqslant \theta, \qquad \forall (g, v) \in \mathcal{O}, \\
& f^\top x \leqslant \overline{f}, \qquad \forall (f, \overline{f}) \in \mathcal{F}.
\end{aligned}$$

More precisely, assume that we have such a relaxation of (2SLP). Let $x^k$ be an optimal first stage control of this relaxation. If it is admissible, meaning that for all scenario $\xi$ there exists an admissible recourse control $y_\xi$, we compute, through duality, a subgradient $g^k \in \partial V(x^k)$. This yields a new *optimality cut* $\theta \geqslant (g^k)^\top(x - x^k) + V(x^k)$, which is added to $\mathcal{O}$. If $x^k$ is not admissible we can add a *feasibility cut* to $\mathcal{F}$ instead by using dual optimal extreme ray (see [2, §5.1.b]). We then solve our strengthened relaxation to obtain $x^{k+1}$.

The L-Shaped method specifies that the subgradient $g^k$ can be obtained as an average over $\xi$ of subgradients $g^{k,\xi} \in \partial_x Q(x^k, \xi)$. In particular, it means that, to compute the subgradient, we can solve $|\,\mathrm{supp}(\boldsymbol{\xi})|$ smaller LP instead of a large one.

**Remark 9** (L-shaped for continuous distribution)**.** *When the distribution are non-finitely supported, we cannot apply naively this method as there is a non-finite number of scenarios. Nevertheless, we can still approximate* $\operatorname{epi}(V)$ *with cuts. We can compute* $\theta = V_{\mathcal{R}_{\check{x}}}(\check{x})$ *and a subgradient* $g \in \partial V_{\mathcal{R}_{\check{x}}}(\check{x})$ *by solving* $|\mathcal{R}_{\check{x}}|$ *linear problems of the form* (2) *through exact quantization. By Theorem 3,* $\theta = V_{\mathcal{R}_{\check{x}}}(\check{x}) = V(\check{x})$*. Further,* $g \in \partial V_{\mathcal{R}_{\check{x}}}(\check{x}) \subset \partial V(\check{x})$ *by Lemma 8. Then* $(\theta, g)$ *define an optimality cut.*

Lemma 8 shows that, at each step $k$ of Algorithm 1, we add a collection of valid cuts which are exact at $x^k$ to our collection of cuts. This means that APM methods can be seen as a Bender's decomposition method where we add more than one exact cut per iteration. In particular, when $x^k \in \operatorname{ri}(\operatorname{dom}(V))$ we add the whole tangent cone of $\operatorname{epi}(V)$ at $x$ instead of a single cut.

### *3.3. Convergence of APMs*

We start by showing that the bounds generated in Algorithm 1 are monotonic.

**Lemma 10.** *Assume that the partition oracle used is adapted. For every computed step* $k$ *we have*

$$z_L^{k-1} \leqslant z_L^k \leqslant \operatorname{val}(2\text{SLP}) \leqslant z_U^k \leqslant z_U^{k-1}$$

*Proof.* Since $\mathcal{P}^k \preccurlyeq_{\mathbb{P}} \mathcal{P}^{k-1}$, by Lemma 2, we have, for all $x \in X$,

$$\begin{aligned} c^\top x + V_{\mathcal{P}^{k-1}}(x) &\leqslant c^\top x + V_{\mathcal{P}^k}(x) \\ &\leqslant c^\top x + V(x) \end{aligned}$$

Minimizing over $x$ yields $z_L^{k-1} \leqslant z_L^k \leqslant \operatorname{val}(2\text{SLP})$. For any $k$, we have that $\mathcal{P}^k$ is adapted to $x^k \in X$, hence $V_{\mathcal{P}^k}(x^k) = V(x^k)$, thus $\operatorname{val}(2\text{SLP}) \leqslant c^\top x_k + V_{\mathcal{P}^k}(x^k)$. Further, by definition of $z_k^U$ in Algorithm 1, $z_k^U = \min_{\kappa \leqslant k} c^\top x^\kappa + V_{\mathcal{P}^\kappa}(x^\kappa)$, yielding $\operatorname{val}(2\text{SLP}) \leqslant z_U^k \leqslant z_U^{k-1}$. ☐

We now prove finite convergence of any APM.

**Theorem 11.** *Assume that the partition oracle used is adapted. If* $X \subset \mathbb{R}^n_+$ *has a finite diameter* $M \in \mathbb{R}_+$ *and* $x \mapsto c^\top x + V(x)$ *is Lipschitz with constant* $L$ *then the partition based Algorithm 1 finds an* $\varepsilon$*-solution in at most* $\left(\frac{LM}{\varepsilon} + 1\right)^n$ *iterations.*

*Proof.* We adapt the classical proof of Kelley's cutting plane algorithm to APMs. Let $k \in \mathbb{N}$ and $1 < i < k$, we have that $V(x_i) = V_{\mathcal{P}^{k-1}}(x_i) = V_{\mathcal{P}^i}(x_i)$. Let $g \in \partial V_{\mathcal{P}^{k-1}}(x_i) \subset \partial V(x_i)$ such that $\|c + g\|$ is bounded by the Lipschitz constant $L$ then

$$\begin{aligned} z_U^k - z_L^k &\leqslant c^\top x^i + V_{\mathcal{P}^i}(x^i) - \left(c^\top x^k + V_{\mathcal{P}^{k-1}}(x^k)\right) \\ &= c^\top (x^i - x^k) + V_{\mathcal{P}^{k-1}}(x^i) - V_{\mathcal{P}^{k-1}}(x^k)) \\ &\leqslant c^\top (x^i - x^k) - g^\top (x^k - x^i) \\ &\leqslant \|c + g\|_2 \|x^i - x^k\|_2 \leqslant L\|x^i - x^k\|_2. \end{aligned}$$

Then, for $k$ such that, $\varepsilon < z_k^U - z_k^L$, we have $\varepsilon < L\|x^i - x^k\|$, in particular $\|x^i - x^k\| \geqslant \varepsilon/L$. By definition of $M$ there are at most $\left(\frac{LM}{\varepsilon} + 1\right)^n$ balls of radius $\varepsilon/L$ in $X$. An $\varepsilon$-solution being obtained as soon as two points are in the same ball. ☐

## 4. Numerical examples

In this section, we detail the actual computation required by Algorithm 1 and illustrate the algorithm on numerical examples.

### *4.1. Detailing computation*

In the following two sections, we give more details on how to compute the Lines 4 to 7 of Algorithm 1.

#### *4.1.1. Master problem and subproblems*

Once $\mathbb{E}\left[\boldsymbol{\xi} \mid P\right]$ and $\mathbb{P}\left[P\right]$ have been computed for $P \in \mathcal{P}^{k-1}$, by Eq. (4) and Eq. (1), the problem of Line 4 is reduced to the following linear problem

$$\begin{aligned} \min_{x \in X, (y_P) \in (\mathbb{R}^m_+)^{\mathcal{P}^{k-1}}} \quad & c^\top x + \sum_{P \in \mathcal{P}^k} \mathbb{P}\left[P\right] q^\top y_P \\ \text{s.t.} \quad & \mathbb{E}\left[\boldsymbol{T}|P\right] x + W y_P = \mathbb{E}\left[\boldsymbol{h}|P\right] \\ & \forall P \in \mathcal{P}^k. \end{aligned}$$

Moreover, to compute the upper bound in Line 7, we need to solve at most $|\mathcal{P}^k|$ linear problems of dimension $m$

$$\begin{aligned} Q(x^k, \mathbb{E}\left[\boldsymbol{\xi}|P\right]) := \min_{y_P \in \mathbb{R}^m_+} \quad & q^\top y_P \\ \text{s.t.} \quad & \mathbb{E}\left[\boldsymbol{T}|P\right] x^k + W y_P = \mathbb{E}\left[\boldsymbol{h}|P\right] \end{aligned}$$

#### *4.1.2. Refinement, expectation and probabilities*

Recall that we can store a polyhedron $E$, either as a family of constraints $(M, \beta)$ such that $E = \{x \in \mathbb{R}^d \mid Mx \leqslant \beta\}$ ($H$-representation) or as families of vertices $(v_i)_{i \in I}$ and rays $(r_j)_{j \in J}$ such that $E = \operatorname{Conv}(v_i)_{i \in I} + \operatorname{Cone}(r_j)_{j \in J}$ ($V$-representation).

Both representation are implemented polymake, an open source software and julia library [13]. We can switch between representations through algorithms such as the *double description* [14].

We can simultaneously compute conditional expectations, probabilities and refinement as detailed in Algorithm 2.

**Data:** $\mathcal{P}^{k-1}$ and $\mathcal{R}_{x^k}$ the partition to refine, second stage distributions $\boldsymbol{T}$ and $\boldsymbol{h}$.

```
1 Set P^k := ∅;
2 for P ∈ P^{k-1} and R ∈ R_{x^k} do
3 |   Set P' := P ∩ R;
4 |   if ℙ[P'] > 0 then
5 |   |   Store ℙ[P'], 𝔼[T|P'] and 𝔼[h|P'];
6 |   |   Set P^k := P^k ∪ {P'};
7 |   end
8 end
```

**Algorithm 2:** Refinement procedure.

In this algorithm, the computation of probabilities on polyhedra in Line 5 is a $\sharp P$-complete problem in the general case, although, for a large class of distributions, formulas exists (see [12, Section 5] for a review).

*4.1.3. Explicit partition oracle*

In this section, we explain how to compute, for $\check{x} \in X$, $\mathcal{R}_{\check{x}} = \big\{E_{N,\check{x}} \,|\, N \in \mathcal{N}(D)\big\}$ where $E_{N,\check{x}} = \big\{(T,h) \,|\, h - T\check{x} \in \mathrm{ri}(N)\big\}$.

The computation of the normal fan $\mathcal{N}(D)$, already implemented in polymake, can be done thanks to a double description and active constraint sets. Note that if $N \in \mathcal{N}(D)$, then $E_{N,\check{x}}$ is a relatively open polyhedral cone of $\Xi$. In particular, if $N := \{\psi \,|\, M\psi \leqslant 0\}$ is given in a non-redundant $H$-representation where $M \in \mathbb{R}^{p\times l}$, we have $\mathrm{ri}(N) = \{\psi \,|\, M\psi \ll 0\}$. Then $E_{N,\check{x}} = \{\xi \in \Xi \,|\, H^x\xi \ll 0\}$, with $H^x = (-x_1M \cdots - x_nM \; M)$.

Unfortunately, obtaining an $H$-representation of the normal cone, from the usual $V$-representation, requires a double-description which is numerically intractable in large dimension (see McMullen bounds [15]).

The double-description can be avoided if the technology matrix $\boldsymbol{T} \equiv T$ is fixed. Indeed, in this case $E_{N,\check{x}} \sim_{\mathbb{P}} \{T\} \times \big(T\check{x} + \mathrm{ri}(N)\big)$. Thus, we can compute, at the beginning of the algorithm, a $V$-representation of all $N \in \mathcal{N}(D)$, and easily deduce a $V$-representation of $E_{N,\check{x}}$ by adding $Tx$ to each representant ray.

*4.2. Numerical examples*

We applied Algorithm 1with our geometric oracle to the problems LanDs and CV@R of [11]. We obtained the same partition, and thus the same numerical results. Finally, we treat the problem Prod-Mix for which no partition oracle were given in the literature. Our code is available at `https://github.com/maelforcier/GAPM`.

*4.2.1. Energy planing problem - LandS*

We applied numerically our method to the LandS problem and constated that our geometric oracle returned the same partition as [11].

*4.2.2. Conditional value-at-risk linear problems*

For the conditional value-at-risk problem in [11], note that our geometric oracle yields the same partition:

$$Q^D(\check{x},\xi) := \max_{\lambda\in\mathbb{R}} \quad (-\check{x}^\top r^\xi - \tau)\lambda \qquad \text{s.t.} \quad 0 \leqslant \lambda \leqslant 1$$

Here $D = [0,1]$ and $\mathcal{N}(D) = \{\mathbb{R}^-, \{0\}, \mathbb{R}^+\}$ Then, if $\check{x} \neq 0$, $\mathcal{R}_{\check{x}} = \big\{\{r|\check{x}^\top r > -\tau\}, \{r|\check{x}^\top r = -\tau\}, \{r|\check{x}^\top r < -\tau\}\big\}$.

*4.2.3. Prod-Mix*

We adapted the problem Prod-mix of `https://stoprog.org/SavedLinks/IBM_StoExt_problems/node4.php` as

$$\begin{aligned} \min_{x,\boldsymbol{y}} \quad & -c^\top x + \mathbb{E}\big[q^\top \boldsymbol{y}\big] \\ \text{s.t.} \quad & \boldsymbol{T}x - \boldsymbol{y} \leqslant \boldsymbol{h} \\ & x, \boldsymbol{y} \geqslant 0, \end{aligned}$$

where $q^\top = (5,10)$, $c^\top = (12,40)$, $\boldsymbol{T}$ follows the uniform law $\begin{pmatrix} \mathcal{U}[3.5,4.5] & \mathcal{U}[9,11] \\ \mathcal{U}[0.8,1.2] & \mathcal{U}[36,44] \end{pmatrix}$ and $\boldsymbol{h}^\top$ follows the uniform distribution $\big(\mathcal{U}[5970,6030], \; \mathcal{U}[3979,4021]\big)$. Algorithm 1 gave the results summed up in Table 1

| $k$ | $z_L^k$ | $z_U^k$ | $z_U^k - z_L^k$ | Total time | $\lvert\mathcal{P}^k\rvert$ |
|---|---|---|---|---|---|
| 1 | -18666.67 | -16939.71 | 1726.96 | 0.57 s | 4 |
| 2 | -17873.01 | -17383.73 | 489.28 | 2.1 s | 9 |
| 4 | -17744.67 | -17709.00 | 35.67 | 9.1 s | 25 |
| 6 | -17713.74 | -17711.37 | 2.37 | 23.7 s | 49 |
| 8 | -17711.71 | -17711.56 | 0.15 | 50.0 s | 81 |
| 10 | -17711.57 | -17711.56 | 0.01 | 88.0 s | 121 |

Table 1: Results of Algorithm 1 for Prod-Mix

To compare our approach with SAA, we solved the same problem 100 times, each with $10'000$ scenarios randomly drawn, yielding a 95% confidence interval centered in $-17711$, with length 4.4. This statistical confidence interval required 2058s of computation. By APM, an exact gap smaller than 4.4 is obtained after iteration 6, that is in 23s, which is here roughly the time required for solving one SAA. Thus, Algorithm 1 can be useful to find accurate values.

The most time-consuming parts of the algorithm are the computations of volumes which take 85% of the total time, because polymake only implement exact computations, which was proven to be $\sharp P$-complete [16]. To improve Algorithm 1, we could use precise rapid approximation volume algorithms, see e.g. [17].

## 5. Extensions and perspectives

We now provide an adapted partition oracle for problems with finitely supported recourse matrix $\boldsymbol{W}$ and cost $\boldsymbol{q}$. The convergence results of Section 3 can directly be applied.

We denote the dual admissible set $D_{W,q} := \{\lambda \in \mathbb{R}^\ell \,|\, W^\top \lambda \leqslant q\}$ and $\mathcal{R}_{x,W,q} := \{E_{N,\check{x}} \,|\, N \in \mathcal{N}(D_{W,q})\}$, where $\mathcal{N}(D_{W,q})$ is the normal fan of $D_{W,q}$, and $E_{N,\check{x}}$ defined in Eq. (9a). Then, by the law of total expectation and Lemma 6, straightforward computations show that $\mathcal{P}_x := \big\{\{(W,q)\} \times R \,\big|\, (W,q) \in \operatorname{supp}(\boldsymbol{W},\boldsymbol{q}),\, R \in \mathcal{R}_{x,W,q}\big\}$ is a $\mathbb{P}$-partition of $\mathbb{R}^{\ell\times m} \times \mathbb{R}^m \times \Xi$ adapted to $x$.

**Remark 12.** *Diving further into the polyhedral geometry of 2SLP we show in a forthcoming work that $q \mapsto \mathcal{R}_{x,W,q}$ is constant on $S \in \mathcal{S}_W$ where $\mathcal{S}_W$ is a partition of $\mathbb{R}^{\ell\times m}$ (the collection of relative interiors of the secondary fan of $W$ cells (see* e.g.*, [18, Chapter 5])), hence theoretically enabling GAPM methods for 2SLP with non-finitely supported $\boldsymbol{q}$.*

*More precisely, we will show that $\mathcal{P}_{\check{x}} := \big\{\{W\} \times S \times R \,|\, W \in \operatorname{supp}(\boldsymbol{W}),\, S \in \mathcal{S}_W,\, R \in \mathcal{R}_{x,W,S}\big\}$ is an adapted $\mathbb{P}$-partition to $\check{x}$, where $\mathcal{R}_{x,W,S}$ is the common value of $\mathcal{R}_{x,W,q}$ for any $q \in S$.*

*Aknowledgments.* We profusely thank the anonymous reviewer for his numerous useful suggestions. This research benefited from the support of the FMJH Program Gaspard Monge for optimization and operations research and their interactions with data science.

## References

[1] Stein W Wallace and William T Ziemba. *Applications of stochastic programming.* SIAM, 2005.

[2] John R. Birge and Francois Louveaux. *Introduction to stochastic programming.* Springer Science & Business Media, 2011.

[3] 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(4):638–663, 1969.

[4] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory.* SIAM, 2014.

[5] Daniel Kuhn. *Generalized bounds for convex multistage stochastic programs*, volume 548. Springer Science & Business Media, 2006.

[6] Nalin C.P. Edirisinghe and William T. Ziemba. Bounds for two-stage stochastic programs with fixed recourse. *Mathematics of Operations Research*, 19(2):292–313, 1994.

[7] Jitka Dupačová, Nicole Gröwe-Kuska, and Werner Römisch. Scenario reduction in stochastic programming. *Mathematical Programming*, 95(3):493–511, 2003.

[8] Holger Heitsch and Werner Römisch. A note on scenario reduction for two-stage stochastic programs. *Operations Research Letters*, 35(6):731–738, 2007.

[9] Yongjia Song and James Luedtke. An adaptive partition-based approach for solving two-stage stochastic programs with fixed recourse. *SIAM Journal on Optimization*, 25(3):1344–1367, 2015.

[10] Wim van Ackooij, Welington de Oliveira, and Yongjia Song. Adaptive partition-based level decomposition methods for solving two-stage stochastic programs with fixed recourse. *Informs Journal on Computing*, 30(1):57–70, 2018.

[11] Cristian Ramirez-Pico and Eduardo Moreno. Generalized adaptive partition-based method for two-stage stochastic linear programs with fixed recourse. *Mathematical Programming*, pages 1–20, 2021.

[12] Maël Forcier, Stéphane Gaubert, and Vincent Leclère. Exact quantization of multistage stochastic linear problems. *arXiv preprint arXiv:2107.09566*, 2021.

[13] Ewgenij Gawrilow and Michael Joswig. `polymake`: a framework for analyzing convex polytopes. In *Polytopes—combinatorics and computation (Oberwolfach, 1997)*, volume 29 of *DMV Sem.*, pages 43–73. Birkhäuser, Basel, 2000.

[14] K. Fukuda and A. Prodon. Double description method revisited. In *Selected papers from the 8th Franco-Japanese and 4th Franco-Chinese Conference on Combinatorics and Computer Science*, pages 91–111, London, UK, 1996. Springer-Verlag.

[15] Peter McMullen. The maximum numbers of faces of a convex polytope. *Mathematika*, 17(2):179–184, 1970.

[16] Martin E. Dyer and Alan M. Frieze. On the complexity of computing the volume of a polyhedron. *SIAM Journal on Computing*, 17(5):967–974, 1988.

[17] Ben Cousins and Santosh Vempala. A practical volume algorithm. *Mathematical Programming Computation*, 8(2):133–160, 2016.

[18] Jesús A. De Loera, Jörg Rambau, and Francisco Santos. *Triangulations Structures for algorithms and applications.* Springer, 2010.