STOCHASTIC OPTIMIZATION OVER A PARETO SET ASSOCIATED WITH A STOCHASTIC MULTI-OBJECTIVE OPTIMIZATION PROBLEM

HENRI BONNEL, JULIEN COLLONGE

ABSTRACT. We deal with the problem of minimizing the expectation of a real valued random function over the weakly Pareto or Pareto set associated with a Stochastic Multi-Objective Optimization Problem (SMOP) whose objectives are expectations of random functions. Assuming that the closed form of these expectations is difficult to obtain, we apply the Sample Average Approximation method (SAA-N, where N is the sample size) in order to approach this problem.

We prove that the Hausdorff-Pompeiu distance between the SAA-N weakly Pareto sets and the true weakly Pareto set converges to zero almost surely as N goes to infinity, assuming that all the objectives of our (SMOP) are strictly convex. Then we show that every cluster point of any sequence of SAA-N optimal solutions (N=1,2,. . .) is almost surely a true optimal solution.

To handle also the nonconvex case, we assume that the real objective to be minimized over the Pareto set depends on the expectations of the objectives of the (SMOP), i.e. we optimize over the outcome space of the (SMOP). Then, whithout any convexity hypothesis, we obtain the same type of results for the Pareto sets in the outcome spaces. Thus we show that the sequence of SAA-N optimal values (N=1,2 ...) converges almost surely to the true optimal value.

Keywords: Optimization over a Pareto Set, Optimization over the Pareto Outcome Set, Multiobjective Stochastic Optimization, Multiobjective Convex Optimization, Sample Average Approximation Method

AMS: 90C29, 90C25, 90C15, 90C26.

1. INTRODUCTION

Multi-Objective Optimization Problems (MOP) have become a major area of interest in Optimization and in Operation Research since Kuhn-Tucker's results (1951), even though the genesis of this theory goes back to Pareto (1906) who was inspired by Edgeworth's indifference curves.

In a (MOP) we deal with several conflicting objectives. The solution set (called Pareto or efficient set) consists of the feasible solutions which ensure some sort of equilibrium amongst the objectives. To be more precise, consider the vector function $g = (g^1, g^2, \dots, g^r)$ defined from S into $\mathbb{R}^r$ (where S is an arbitrary nonempty set). For the

$$(MOP) \qquad\qquad \min_{x\in S} g(x),$$

a point $x^* \in S$ is said to be

- *Pareto solution* iff there is no element $x \in S$ satisfying $\forall j \in \{1,\dots,r\}\ g^j(x) \le g^j(x^*)$ and $\exists j_0 \in \{1,\dots,r\}\ g^{j_0}(x) < g^{j_0}(x^*)$,
- *weakly Pareto solution* iff there is no element $x \in S$ satisfying $\forall j \in \{1,\dots,r\}\ g^j(x) < g^j(x^*)$.

That is to say, Pareto solutions are such that none of the objectives values can be improved further without deteriorating another, and weakly Pareto solutions are such that it is impossible to strictly improve simultaneously all the objectives values.

However, the Pareto set is often very large (may be infinite, and even unbounded), and technically speaking each Pareto solution is acceptable. The natural question that arises is: how to choose one solution? One possible answer is to optimize a *scalar (real valued)* function g^0 over the Pareto set associated with (MOP), i.e. to consider the problem

$$\min_{x \in E} g^0(x)$$

where E is the set of Pareto (or weakly Pareto) solutions associated with (MOP). For instance, production planning (see e.g. [5]) and portfolio management (see e.g. [41]) are practical areas where this problem arises. In general, this problem of optimizing over a Pareto set is an useful tool for a decision maker who wants to choose one solution over the embarrassing and very large Pareto set. Also, for numerical computation, solving this problem one may avoid generate all the Pareto set, and thus significantly reduce the computation time. A particular but important case of this problem is given by the situation when the scalar function to be optimized over the Pareto set depends on the objectives of the (MOP). In other word we optimize over the outcome space of the (MOP). This is the case when a decision maker wants to know the range (maximum and minimum value) of one (or more) objective over the Pareto set.

This problem of optimizing a scalar objective over the Pareto set has been intensively studied the last decades beginning with Philip's paper [44], and continued by many authors in [1, 5, 11, 12, 13, 17, 23, 24, 25, 31, 34, 35] (see ref. [48] for an extensive bibliography).

Some generalization to *semivectorial bilevel optimization problems* have been presented in [18, 26, 2, 49, 28, 19, 20].

The particular problem of optimizing a scalar function over the outcome space of a (MOP) has been studied in [6, 7, 8, 9, 40].

In all theses papers, the Pareto set is associated with a deterministic (MOP), not with a Stochastic Multi-Objective Optimization Problem (SMOP). In the deterministic case, optimizing a real valued function over the Pareto set is already very difficult due to the fact that the Pareto set is not described explicitly, and is not convex even in the linear case.

Uncertainty is inherent in most real cases, where observed phenomena are disturbed by random perturbations. Even if the presence of random vectors in optimization models complicates the mathematics governing them, it is very important to take into account this uncertainty in order to calibrate models at best.

In our paper we study the problem of optimizing the expectation of a scalar random function over a Pareto set associated with a Stochastic Multi-Objective Optimization Problem, and our study seems to be the first attempt to deal with this kind of problem.

If the expected value functions can be computed directly, the problem becomes a deterministic one. But in most cases, the closed form of the expected values is very difficult to obtain. This is the case which will be considered in this paper. In order to give approximations, we apply the well-known Sample Average Approximation (SAA-N, where N is the sample size) method. Under reasonable and suitable assumptions, we show that the (SAA-N weakly Pareto sets or SAA-N Pareto sets image) converge in the Hausdorff-Pompeiu distance sense to their true counterparts. Moreover, we show that the sequence of SAA-N optimal values converges to the true optimal value as the sample size increase.

Some results in (SMOP) using SAA-N method have been recently obtained by Fliege and Xu in [30] using a smoothing infinity norm scalarization approach to solve the SAA-N problems. Roughly speaking, the paper [30] proves that approximate Pareto solutions of the SAA-N problems tend to some approximate solution of the true problem. However this approach is not sufficient for our problem because it shows only that the deviation between the Pareto sets associated with the SAA-N problems and the true Pareto set tends to zero, hence it is possible to have true Pareto solutions which are not limits of SAA-N solutions. Optimizing a real function over the Pareto set requires that the Hausdorff-Pompeiu distance between these sets tends to zero, what is the main concern of our research.

Our paper is organized as follows.

In section 2, we introduce the problem under consideration. We consider two instances of the same problem. Firstly we consider the problem of optimizing the expectation of a real function over the Pareto set in the decision space. Secondly we consider that the real function to be optimized depends on the expectations of the objectives of (SMOP), therefore we optimize over the Pareto set in the outcome space.

In section 3, we present the basic definitions and the facts necessary for the development of our paper.

In section 4 we consider the problem of optimizing the expectation of a real function over the weakly Pareto set in the **decision space**. First, we show that the deviation of the SAA-N weakly Pareto sets from the true weakly Pareto set tends to zero almost surely as the sample size N goes to infinity. In order to show that the deviation in the other direction tends to zero, we need to assume that the (SMOP) is *strictly* convex . Thus, using some Set Valued Analysis tools and some Stability results, in Theorem 4.2 we show that the sequence of SAA-N weakly Pareto sets tends to the true weakly Pareto set in the Hausdorff-Pompeiu distance sense (which is equivalent in our framework to Painlevé-Kuratowski convergence). Moreover, we show that every cluster point of any sequence of SAA-N optimal solutions (N=1,2,...) is almost surely a true optimal solution. Hence, the sequence of SAA-N optimal values converges with probability one to the true optimal value (Theorem 4.3).

In the next section, in order to handle the nonconvex case, we need to work in the **outcome space**. This means that the real function to be optimized depends on the expectations of the objectives of (SMOP). Moreover, in this setting our real function is optimized over the image in the outcome space of the Pareto set. Using also some results of Set Valued Analysis and Stability, we show that the SAA-N images of Pareto sets tends almost surely in the Hausdorff-Pompeiu distance sense to the true Pareto set image. Thus we show that the sequence of SAA-N optimal values converges almost surely to the true optimal value (Theorem 5.2).

In section 6, using MATLAB7, we present some numerical results for an illustrative example with a (SMOP) given by a Bi-Objective Stochastic Optimization Problem.

2. Problem Statement

Let us briefly introduce the two problems under consideration. The first one, denoted by (D) will be studied in the **Decision space** (see section 4). This problem is to minimize the expectation of a scalar random function over the weakly Pareto set associated with a Stochastic Multi-Objective Optimization Problem. Note that the function to be minimized over the Pareto set may be independent of other objectives. That is to say,

$$(D) \qquad \min_{x\in E^w} \mathbb{E}\Big[F^0\Big(x,\xi(\cdot)\Big)\Big]$$

where $\xi : \Omega \to \mathbb{R}^d$ is a *random vector* defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $x \in \mathbb{R}^n$ is a *deterministic vector,* $\mathbb{E}\Big[F^0\Big(x, \xi(\cdot)\Big)\Big]$ is, for each $x \in \mathbb{R}^n$, the expectation of the scalar random variable $\omega \mapsto F^0(x, \xi(\omega))$, and E^w is the set of weakly Pareto solutions associated with the following Stochastic Multi-Objective Optimization Problem

$$(SMOP) \qquad \min_{x \in S} \mathbb{E}\Big[F\Big(x, \xi(\cdot)\Big)\Big]$$

where the feasible set $S \subset \mathbb{R}^n$. The objectives are given by

$$\mathbb{R}^n \times \Omega \ni \Big(x, \omega\Big) \mapsto F\Big(x, \xi(\omega)\Big) = \Big(F^1\Big(x, \xi(\omega)\Big), \ldots, F^r\Big(x, \xi(\omega)\Big)\Big) \in \mathbb{R}^r,$$

where $F^i : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$, $i = 1, \ldots, r$.

For problem (D) we need to assume that (SMOP) is strictly convex (see section 4 for details).

The second problem, (O), will be studied in the **Outcome space** (see section 5). This means that the scalar function to be minimized over the Pareto set associated with (SMOP) depends on the expectations of the objectives. That is to say,

$$(O) \qquad \min_{x \in E} f\Big(\mathbb{E}\Big[F\Big(x, \xi(\cdot)\Big)\Big]\Big)$$

where $f : \mathbb{R}^r \to \mathbb{R}$ is a scalar deterministic continuous function and E is the set of Pareto solutions associated with the Stochastic Multi-Objective Optimization Problem (SMOP) defined above.

But in this special case *we do not need any convexity assumption.*

In the sequel, when we talk about the **true problem**, we will refer to problem (D) or problem (O).

The purpose of the next section is to rigorously define these two problems, and to give some definitions and usefull results.

3. Preliminaries

Definition 3.1. *Let $(\Omega, \mathcal{F})$ and $(\mathbb{R}^d, \mathcal{B}_d)$ be measurable spaces, where $\mathcal{B}_d$ is the $\mathbb{R}^d$ Borel $\sigma-$field. A mapping $\xi : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}_d)$ is said to be measurable with respect to $\mathcal{F}$ and $\mathcal{B}_d$ if for any Borel set $B \in \mathcal{B}_d$, its inverse image $\xi^{-1}(B) := \{\omega \in \Omega : \xi(\omega) \in B\}$ is $\mathcal{F}$-measurable.*

A measurable mapping $\xi(\cdot)$ from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ into $\mathbb{R}^d$ is called a random vector. Note that the mapping $\xi(\cdot)$ generates the probability measure $\mathbb{P}_\xi(B) := \mathbb{P}(\xi^{-1}(B))$ on $(\mathbb{R}^d, \mathcal{B}_d)$.

The smallest closed set $\Xi \subset \mathbb{R}^d$ such that $\mathbb{P}_\xi(\Xi) = 1$ is called the support of measure $\mathbb{P}_\xi$. We can view the space $(\Xi, \mathcal{B}_\Xi)$ equipped with probability measure $\mathbb{P}_\xi$ as a probability space, where $\mathcal{B}_\Xi$ is the trace of $\mathcal{B}_d$ on Ξ. This probability space provides all relevant probabilistic information about the considered random vector.

Definition 3.2. *Let $\xi : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Xi, \mathcal{B}_\Xi, \mathbb{P}_\xi)$ be a random vector and consider a function $g : \mathbb{R}^n \times \Xi \to \mathbb{R}$. We say that g is a random function if for every fixed $x \in \mathbb{R}^n$, the function $\xi \mapsto g(x, \xi)$ is $\mathcal{B}_\Xi/\mathcal{B}_1$- measurable. For every fixed $\xi \in \Xi$ we have that*

$\mathbb{R}^n \ni x \mapsto g(x,\xi)$ is a real valued deterministic function. Note that for a random function $\mathbb{R}^n \times \Xi \ni (x,\eta) \mapsto g(x,\eta)$, we can define the corresponding expected value function $\mathbb{E}_\xi[g(x,\cdot)] = \int\limits_\Xi g(x,\eta)d\mathbb{P}_\xi(\eta)$.

Remark 3.1. *If the distribution of a random function is known, we can compute directly its expectation. Hence we consider the case where $\mathbb{E}_\xi[g(x,\cdot)]$ is very difficult to assess, and we turn to approximations such as the Sample Average Approximation method, where the expected value function is approximated by its empirical mean.*

Consider an independent identically distributed (i.i.d.) sequence $(\xi_k)_{k\in\mathbb{N}^*}$ of random vectors defined on the same probability space $(\Omega,\mathcal{F},\mathbb{P})$, and having the same distribution $\mathbb{P}_\xi$ on $(\Xi,\mathcal{B}_\Xi)$ as the random vector ξ. I.e., for each $k\geq 1,\ \xi_k : (\Omega,\mathcal{F},\mathbb{P}) \to (\Xi,\mathcal{B}_\Xi,\mathbb{P}_\xi)$ is a random vector supported by Ξ.

Let us set $\tilde{\Xi} = \prod\limits_{N=1}^{\infty} \Xi$ and let $\tilde{\mathcal{B}} = \otimes_{N=1}^{\infty}\mathcal{B}_\Xi$ denotes the smallest σ-algebra on $\tilde{\Xi}$ generated by all sets of the form $B_1\times B_2\times\cdots\times B_N\times\Xi\times\Xi\times\ldots,\ B_k\in\mathcal{B}_\Xi,\ k=1,\ldots,N,\ N=1,2,\ldots$.
The next Theorem is from General Measure Theory, and can be considered as a non trivial extension of Fubini's Theorem ([37, Theorem 10.4]) :

Theorem 3.1. *There exists a unique probability $\tilde{\mathbb{P}}_\xi$ on $(\tilde{\Xi},\tilde{\mathcal{B}})$ such that $\tilde{\mathbb{P}}_\xi(B_1\times B_2\times\cdots\times B_N\times\Xi\times\Xi\times\ldots) = \prod\limits_{k=1}^{N}\mathbb{P}_\xi(B_k)$ for all $N=1,2,\ldots$ and $B_k\in\mathcal{B}_\Xi$ for all $k=1,\ldots,N$.*

For each ranfom function $g:\mathbb{R}^n\times\Xi\to\mathbb{R}$, $x\in\mathbb{R}^n$, $N\in\mathbb{N}^*$(where $\mathbb{N}^*$ denote the set of positive integers), let $\hat{g}_N(x,\cdot)$ denote the following N-approximation

(1)
$$\hat{g}_N(x,\cdot) : (\tilde{\Xi},\tilde{\mathcal{B}},\tilde{\mathbb{P}}_\xi) \to \mathbb{R}$$
$$\tilde{\xi} = (\xi_1,\xi_2,\ldots) \mapsto \frac{1}{N}\sum_{k=1}^{N} g(x,\xi_k)$$

Definition 3.3. *For each $N\in\mathbb{N}^*$, $x\in\mathbb{R}^n$, the mapping $\tilde{\xi}\mapsto\hat{g}_N(x,\tilde{\xi})$ is called a $N-$Sample Average Approximation (SAA-N) function.*

Remark 3.2. *The random process $\tilde{\xi}$ can be viewed as a mapping from $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ and taking values in $(\tilde{\Xi},\tilde{\mathcal{B}},\tilde{\mathbb{P}}_\xi)$. Note that the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ can be constructed in a similar way as we did before for $(\tilde{\Xi},\tilde{\mathcal{B}},\tilde{\mathbb{P}}_\xi)$.*

Let $\xi : (\Omega,\mathcal{F},\mathbb{P}) \to (\Xi,\mathcal{B}_\Xi,\mathbb{P}_\xi)$ be a fixed random vector. Our $(SMOP)$ can be rewritten as follows.

$$(SMOP) \qquad \min_{x\in S}\mathbb{E}_\xi\Big[F\Big(x,\cdot\Big)\Big]$$

Recall that the feasible set $S\subset\mathbb{R}^n$ and the vector objective is given by

$$\mathbb{R}^n\times\Xi\ni\Big(x,\xi\Big)\mapsto F\Big(x,\xi\Big) = \Big(F^1\Big(x,\xi\Big),\ldots,F^r\Big(x,\xi\Big)\Big)\in\mathbb{R}^r$$

Let us reformulate the true stochastic problems under consideration.

$$(D) \qquad \min_{x\in E^w}\mathbb{E}_\xi\Big[F^0\Big(x,\cdot\Big)\Big]$$

where $F^0 : \mathbb{R}^n \times \Xi \to \mathbb{R}$ is a scalar *random* function and E^w is the weakly Pareto set associated with $(SMOP)$.

$$(O) \qquad \min_{x\in E} f\Big(\mathbb{E}_\xi\Big[F\Big(x,\cdot\Big)\Big]\Big)$$

where $f : \mathbb{R}^r \to \mathbb{R}$ is a scalar *deterministic* continuous function and E is the Pareto set associated with $(SMOP)$.

Note that all considered random functions are supposed to be $\mathcal{B}_\Xi$-measurable and $\mathbb{P}_\xi$-integrable.

In the sequel, for $y, z \in \mathbb{R}^r$, $y \leq z$ means $y_j \leq z_j$ for all $j = 1, \dots, r$ and $y < z$ means $y_j < z_j$ for all $j = 1, \dots, r$.

Let us introduce the following assumptions:

(H_0) S is a nonempty compact subset of $\mathbb{R}^n$.

(H_1) the i.i.d property holds for the random process $\tilde{\xi} \in \tilde{\Xi}$.

(H_2) $\forall j = 0, \dots, r,\ x \mapsto F^j(x,\xi)$ is finite valued and continuous on S for a.e. $\xi \in \Xi$.

(H_3) $\forall j = 0, \dots, r,\ F^j$ is dominated by an integrable function $K^j,$ i.e.

$$\mathbb{E}_\xi\left[K^j(\cdot)\right] < \infty$$

$$\left|F^j(x,\xi)\right| \leq K^j(\xi), \quad \text{for all } x \in S \text{ and for a.e. } \xi \in \Xi$$

(H_4) S is convex.

(H_5) $\forall j = 1, \dots, r,\ x \mapsto F^j(x,\xi)$ is *strictly convex* on S a.e. on Ξ.

We will specify at each time we need to use some or all of these assumptions.

The main objective of this paper is to provide solutions to the true problems ((D) and (O)) through approximations. To do so, consider the following SAA-N functions.

$$\hat{F}^0_N : \mathbb{R}^n \times (\tilde{\Xi}, \tilde{\mathcal{B}}, \tilde{\mathbb{P}}_\xi) \to \mathbb{R}$$

$$(2) \qquad (x,\tilde{\xi}) = (x,\xi_1,\xi_2,\dots) \mapsto \hat{F}^0_N(x,\tilde{\xi}) := \frac{1}{N}\sum_{k=1}^N F^0(x,\xi_k)$$

$$\hat{F}_N : \mathbb{R}^n \times (\tilde{\Xi}, \tilde{\mathcal{B}}, \tilde{\mathbb{P}}_\xi) \to \mathbb{R}^r$$

$$(3) \qquad (x,\tilde{\xi}) = (x,\xi_1,\xi_2,\dots) \mapsto \hat{F}_N(x,\tilde{\xi}) := \frac{1}{N}\sum_{k=1}^N \Big(F^1(x,\xi_k),\dots,F^r(x,\xi_k)\Big),$$

where the probability space $(\tilde{\Xi}, \tilde{\mathcal{B}}, \tilde{\mathbb{P}}_\xi)$ and the random process $\tilde{\xi} = (\xi_1, \xi_2, \dots)$ have been introduced above.

Remark 3.3. *By (H_2), for all $j = 0, \dots, r$ there exists a set $A^j \subset \Xi$ with $\mathbb{P}_\xi(A^j) = 0$ such that, $\forall \xi \in \Xi \setminus A^j,\ x \mapsto F^j(x,\xi)$ is continuous on S. Letting $A = \cup_{j=1}^r A^j$, we get $\forall \xi \in \Xi \setminus A,\ x \mapsto F(x,\xi)$ is continuous on S, and $\mathbb{P}_\xi(A) = 0$.*
Letting $\tilde{A} = \cup_{N\in\mathbb{N}^*} \underbrace{A \times \cdots \times A}_{N} \times \Xi \times \Xi \dots$, we have that $\tilde{\mathbb{P}}_\xi(\tilde{A}) = 0$, and the mapping $x \mapsto \hat{F}_N(x,\tilde{\xi})$ is continuous on S for all $N \in \mathbb{N}^*$ and for all $\tilde{\xi} \in \tilde{\Xi} \setminus \tilde{A}$.
The same rule obviously holds for $\hat{F}^0_N$ setting $\tilde{A}_0 = \cup_{N\in\mathbb{N}^*} \underbrace{A^0 \times \cdots \times A^0}_{N} \times \Xi \times \Xi \dots$.

Definition 3.4. *For each $N \in \mathbb{N}^*$ and $\tilde{\xi} \in \tilde{\Xi}$, we denote by $E_N^w(\tilde{\xi})$ (resp. $E_N(\tilde{\xi})$) the weakly-Pareto (resp. Pareto) set associated with the $N-$Sample Average Approximation Multi-Objective Optimization Problem (SAA-N MOP) $\min_{x\in S} \hat{F}_N(x,\tilde{\xi})$, where $x \mapsto \hat{F}_N(x,\tilde{\xi})$ is a vector valued SAA-N function.*

For each $N \in \mathbb{N}^*$ and $\tilde{\xi} \in \tilde{\Xi}$, consider the following problems :

$$(D_N(\tilde{\xi})) \qquad \min_{x\in E_N^w(\tilde{\xi})} \hat{F}_N^0(x,\tilde{\xi})$$

$$(O_N(\tilde{\xi})) \qquad \min_{x\in E_N(\tilde{\xi})} f(\hat{F}_N(x,\tilde{\xi}))$$

where $E_N^w(\tilde{\xi})$ (resp. $E_N(\tilde{\xi})$) is the weakly Pareto (resp. Pareto) set associated with the (SAA-N MOP)

$$\min_{x\in S} \hat{F}_N(x,\tilde{\xi}) \tag{4}$$

The scalar SAA-N function $\hat{F}_N^0$ is defined by (2), and $\hat{F}_N$ is a $\mathbb{R}^r$ valued SAA-N function defined by (3).
In the sequel we will call $(D_N(\tilde{\xi}))$ (resp. $(O_N(\tilde{\xi}))$) $(N \in \mathbb{N}^*, \tilde{\xi} \in \tilde{\Xi})$ the **SAA-N problem**. Under some reasonable assumptions, we will show that the solutions and/or optimal values of SAA-N problems for sufficiently large N are approximate solutions and/or approximate optimal values to the true problem (D) (resp. (O)).

By the Uniform Law of Large Number (ULLN) [47, Theorem 7.48], under (H_0, H_1, H_2, H_3), we obtain immediately the two following results.

Proposition 3.1. *For any $j = 0,\dots,r$, the expected value function $x \mapsto \mathbb{E}_\xi\left[F^j(x,\cdot)\right]$ is finite valued and continuous on S. Moreover,*

$$\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi}\in\tilde{\Xi}\,\Big|\,\forall\epsilon>0,\ \exists N(\epsilon,\tilde{\xi})\in\mathbb{N}^*\ :\ \forall N\geq N(\epsilon,\tilde{\xi}),\ \max_{0\leq j\leq r}\sup_{x\in S}\Big|\hat{F}_N^j(x,\tilde{\xi})-\mathbb{E}_\xi\left[F^j(x,\cdot)\right]\Big|\leq\epsilon\Big\}\Big)=1$$

Lemma 3.1. *For each convergent sequence $(x_N)_{N\in\mathbb{N}^*}$ in S, let x be its limit. Then $x \in S$ and*

$$\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi}\in\tilde{\Xi}\Big|\forall j=0,\dots,r,\ \lim_{N\to\infty}\hat{F}_N^j(x_N,\tilde{\xi})=\mathbb{E}_\xi[F^j(x,\cdot)]\Big\}\Big)=1$$

Remark 3.4. *By Proposition 3.1, there exists a set $\tilde{B} \subset \tilde{\Xi}$ with $\tilde{\mathbb{P}}_\xi(\tilde{B}) = 0$ such that for each fixed $\tilde{\xi} \in \tilde{\Xi} \setminus \tilde{B}$, $\hat{F}_N(\cdot,\tilde{\xi})$ converges to $\mathbb{E}_\xi[F(\cdot,\cdot)]$ uniformly on S. For the same reasons, there exists $\tilde{B_0} \subset \tilde{\Xi}$ with $\tilde{\mathbb{P}}_\xi(\tilde{B_0}) = 0$ such that for each fixed $\tilde{\xi} \in \tilde{\Xi}\setminus\tilde{B_0}$, $\hat{F}_N^0(\cdot,\tilde{\xi})$ converges to $\mathbb{E}_\xi[F^0(\cdot,\cdot)]$ uniformly on S.*

Definition 3.5. *Let $A, B \subset \mathbb{R}^n$ be two nonempty bounded sets.*

(1) *We denote by $d(x,B) := \inf_{x'\in B} \|x - x'\|$ the distance from $x \in \mathbb{R}^n$ to B, where $\|\cdot\|$ stands for the Euclidian norm.*

(2) *We denote $\mathbb{D}(A,B) := \sup_{x\in A} d(x,B)$ the deviation of the set A from the set B.*

(3) *Finally, we denote $\mathbb{H}(A,B) := \max(\mathbb{D}(A,B),\mathbb{D}(B,A))$ the Hausdorff-Pompeiu distance between the set A and the set B.*

Remark 3.5. *Note that in general $\mathbb{H}$ is a pseudometric. If we consider the set of all nonempty and compact subsets of $\mathbb{R}^n$, $\mathbb{H}$ becomes a metric. Also, for two nonempty bounded sets A and B, the Hausdorff-Pompeiu distance vanishes if and only if A and B have the same closure.*

Lemma 3.2. *Under (H_0, H_1, H_2, H_3), for almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the Hausdorff-Pompeiu distance between $\hat{F}_N(S, \tilde{\xi})$ and $\mathbb{E}_\xi[F(S, \cdot)]$ tends to zero as N tends to infinity, i.e.*

$$\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi} \in \tilde{\Xi}\Big| \lim_{N\to\infty} \mathbb{H}\Big(\hat{F}_N(S, \tilde{\xi}), \mathbb{E}_\xi[F(S, \cdot)]\Big) = 0\Big\}\Big) = 1$$

Proof. By Remark 3.4, there exists a set $\tilde{B} \subset \tilde{\Xi}$ with $\tilde{\mathbb{P}}_\xi(\tilde{B}) = 0$ such that for each fixed $\tilde{\xi} \in \tilde{\Xi} \setminus \tilde{B}$, $\hat{F}_N(\cdot, \tilde{\xi})$ converges to $\mathbb{E}_\xi[F(\cdot, \cdot)]$ uniformly on S. Moreover, (Remark 3.3), $\forall N \in \mathbb{N}^*, \forall \tilde{\xi} \in \tilde{\Xi} \setminus \tilde{A}, x \mapsto \hat{F}_N(x, \tilde{\xi})$ is continuous on S

Let us prove $\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi} \in \tilde{\Xi}\Big| \lim_{N\to\infty} \mathbb{D}\Big(\hat{F}_N(S, \tilde{\xi}), \mathbb{E}_\xi[F(S, \cdot)]\Big) = 0\Big\}\Big) = 1$. Arguing by contradiction, there exists a set $\tilde{D} \subset \tilde{\Xi}$ with $\tilde{\mathbb{P}}_\xi(\tilde{D}) > 0$ such that for each fixed $\tilde{\xi} \in \tilde{D}$, $\mathbb{D}\Big(\hat{F}_N(S, \tilde{\xi}), \mathbb{E}_\xi[F(S, \cdot)]\Big) \not\to 0$. Let $\tilde{\xi} \in (\tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B})) \cap \tilde{D}$ be fixed. Obviously $\tilde{\mathbb{P}}_\xi((\tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B})) \cap \tilde{D}) > 0$ and then $(\tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B})) \cap \tilde{D} \neq \emptyset$.

Since $\tilde{\xi} \in \tilde{D}$, there exist $\epsilon > 0$ and a strictly increasing mapping $\phi : \mathbb{N} \to \mathbb{N}$ such that, $\forall k \geq 1, \mathbb{D}\Big(\hat{F}_{\phi(k)}(S, \tilde{\xi}), \mathbb{E}_\xi[F(S, \cdot)]\Big) > \epsilon$. By Definition 3.5 (2), there exists $y_{\phi(k)}$ in $\hat{F}_{\phi(k)}(S, \tilde{\xi})$ such that, for all y in $\mathbb{E}_\xi[F(S, \cdot)]$, and all $k \geq$, $d(y_{\phi(k)}, y) > \epsilon$.

Moreover, there exists $(x_{\phi(k)})_{k\geq 1}$ such that $y_{\phi(k)} = \hat{F}_N(x_{\phi(k)}, \tilde{\xi})$ (all k). By the compactness of S, there exists a strictly increasing mapping $\varphi : \mathbb{N} \to \mathbb{N}$ such that $x_{\phi(\varphi(k))} \to x$ and $x \in S$.

Since $\tilde{\xi} \notin (\tilde{A} \cup \tilde{B})$, by Proposition 3.1 and Lemma 3.1, we have $y_{\phi(\varphi(k))} \to \tilde{y}$ and $\tilde{y} \in \mathbb{E}_\xi[F(S, \cdot)]$.

Then, for each fixed $\tilde{\xi} \in (\tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B})) \cap \tilde{D}$, we have a contradiction, hence $\tilde{\mathbb{P}}_\xi(\tilde{D}) = 0$.

Now we prove $\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi} \in \tilde{\Xi}\Big| \lim_{N\to\infty} \mathbb{D}\Big(\mathbb{E}_\xi[F(S, \cdot)], \hat{F}_N(S, \tilde{\xi})\Big) = 0\Big\}\Big) = 1$. Let $\tilde{\xi} \in \tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B})$ and let $y \in \mathbb{E}_\xi[F(S, \cdot)]$. There exists $x \in S$ such that $y = \mathbb{E}_\xi[F(x, \cdot)]$. The sequence $(y_N)_{N\geq 1}$ defined by $y_N = \hat{F}_N(x, \tilde{\xi})$ converges to $y = \mathbb{E}_\xi[F(x, \cdot)]$, hence $d(y, \hat{F}_N(S, \tilde{\xi})) \to 0$ as $N \to +\infty$.

Thus the sequence $(d(\cdot, \hat{F}_N(S, \tilde{\xi})))_{N\geq 1}$ is pointwise convergent on $\mathbb{E}_\xi[F(S, \cdot)]$. On the other hand, since the function $y \mapsto d(y, \hat{F}_N(S, \tilde{\xi}))$ is Lipschitz continuous with Lipschitz constant 1 (each N), the sequence of functions $(d(\cdot, \hat{F}_N(S, \tilde{\xi})))_{N\geq 1}$ is equicontinuous on $\mathbb{R}^r$, hence on the compact set $\mathbb{E}_\xi[F(S, \cdot)]$. Then, from Ascoli-Arzelà Theorem [27], we have that the sequence of functions $(d(\cdot, \hat{F}_N(S, \tilde{\xi})))_{N\geq 1}$ converges uniformly to 0 on $\mathbb{E}_\xi[F(S, \cdot)]$. Hence, $\mathbb{D}(\mathbb{E}_\xi[F(S, \cdot)], \hat{F}_N(S, \tilde{\xi})) \to 0$ as N tends to infinity for a.e. $\tilde{\xi} \in \tilde{\Xi}$.

□

We need to recall some basic facts from Set Valued Analysis (see [4, 16, 21, 32, 33, 46] for details). Let X be a separated topological space, and Y be a metric space.

Let $(A_N)_{N\in\mathbb{N}}$ be a sequence of subsets of Y. We recall that

- $\liminf_{N\to\infty} A_N$ is the set of limits of sequences $(y_N)_{N\geq 1}$ where $y_N \in A_N$ (each N)

- $\limsup_{N\to\infty} A_N$ is the set of cluster points of sequences $(y_N)_{N\geq 1}$ where $y_N \in A_N$ (each N).

Let Γ be a set valued mapping from X into Y, i.e. a function from X to the power set of Y (denoted by $\Gamma : X \rightrightarrows Y$). The *limit inferior* of Γ at $x_0 \in X$ is defined by

$$\liminf_{x\to x_0} \Gamma(x) = \{y \in Y | \forall V \text{ neighborhood of } y, \exists U \text{ neighborhood of } x_0 : \forall x \in U\backslash\{x_0\}, \Gamma(x)\cap V \neq \emptyset\}$$

while the *limit superior* of Γ at $x_0 \in X$ is defined by

$$\limsup_{x\to x_0} \Gamma(x) = \{y \in Y | \forall V \text{ neighborhood of } y, \forall U \text{ neighborhood of } x_0, \exists x \in U\backslash\{x_0\} : \Gamma(x)\cap V \neq \emptyset\}$$

Remark 3.6. [32, p. 61] *Having A, $(A_N)_{N\in\mathbb{N}^*}$ subsets of Y and taking $X = \mathbb{N}^*\cup\{+\infty\}$ endowed with the topology induced by that of $\bar{\mathbb{R}} = \mathbb{R}\cup\{-\infty,+\infty\}$, if $\Gamma : X \rightrightarrows Y$ is the set valued mapping defined by $\Gamma(N) = A_N$ (each N) and $\Gamma(+\infty) = A$ then $\liminf_{N\to\infty} A_N = \liminf_{N\to\infty} \Gamma(N)$ and $\limsup_{N\to\infty} A_N = \limsup_{N\to\infty} \Gamma(N)$.*

Definition 3.6. *Let Γ be a set valued mapping from X into Y. We say that*

- *Γ is upper continuous (u.c.) at $x_0 \in X$ if for any open set $D \subset Y$ such that $\Gamma(x_0) \subset D$, there exists a neighborhood $U \subset X$ of x_0 such that*
$$\forall x \in U,\ \Gamma(x) \subset D$$
- *Γ is lower continuous (l.c.) at $x_0 \in X$ if for any open set $D \subset Y$ such that $\Gamma(x_0) \cap D \neq \emptyset$, there exists a neighborhood $U \subset X$ of x_0 such that*
$$\forall x \in U,\ \Gamma(x) \cap D \neq \emptyset$$
- *Γ is continuous at $x_0 \in X$ if Γ is u.c. and l.c. at x_0.*
- *Γ is continuous if Γ is continuous at every $x \in X$.*

Proposition 3.2. [32, p. 55] *Let $\Gamma : X \rightrightarrows Y$ and let $x_0 \in X$.*
If Γ is l.c. at x_0 then $\Gamma(x_0) \subset \liminf_{x\to x_0} \Gamma(x)$.

Definition 3.7. *Let $\Gamma : X \rightrightarrows Y$ be a set valued mapping. We say that*

- *Γ is Hausdorff upper continuous (H-u.c.) at $x_0 \in X$ if for any $\epsilon > 0$, there exists a neighborhood $U \subset X$ of x_0 such that*
$$\forall x \in U,\ \Gamma(x) \subset \Gamma(x_0) + B_\epsilon$$
where B_ϵ denote the open ball of radius ϵ and center 0.
- *Γ is Hausdorff lower continuous (H-l.c.) at $x_0 \in X$ if for any $\epsilon > 0$, there exists a neighborhood $U \subset X$ of x_0 such that*
$$\forall x \in U,\ \Gamma(x_0) \subset \Gamma(x) + B_\epsilon$$
- *Γ is Hausdorff continuous at $x_0 \in X$ if Γ is H-u.c. and H-l.c. at x_0.*
- *Γ is H-continuous if Γ is H-continuous at every $x \in X$.*

Remark 3.7. [32, p. 59] *Γ is H-u.c. at x_0 if and only if $\lim_{x\to x_0} \mathbb{D}(\Gamma(x), \Gamma(x_0)) = 0$, and Γ is H-l.c. at x_0 if and only if $\lim_{x\to x_0} \mathbb{D}(\Gamma(x_0), \Gamma(x)) = 0$*

Now we recall some usefull property between H-u.c. and u.c. and between H-l.c. and l.c..

Proposition 3.3. *Let $\Gamma : X \rightrightarrows Y$ be a set valued mapping, and let $x_0 \in X$.*

- *If Γ is u.c. at x_0 then Γ is H-u.c. at x_0.*
- *If Γ is H-u.c. at x_0 and $\Gamma(x_0)$ is compact then Γ is u.c. at x_0.*
- *If Γ is H-l.c. at x_0 then Γ is l.c. at x_0.*
- *If Γ is l.c. at x_0 and $\Gamma(x_0)$ is compact then Γ is H-l.c. at x_0.*

Remark 3.8. *The last Proposition means if $\Gamma(\cdot)$ is compact valued, then continuity is equivalent to Hausdorff-continuity.*

Definition 3.8. *We say that $\Gamma : X \rightrightarrows Y$ is*

- *closed valued if for each $x \in X$, $\Gamma(x)$ is a closed set in Y.*
- *closed if $Graph(\Gamma) = \{(x, y)|x \in X, y \in \Gamma(X)\}$ is closed.*
- *compact at $x \in X$ if for every sequence $(x_k, y_k)_{k\geq 1}$ with $x_k \in X,\ y_k \in \Gamma(x_k)$ (each k) and $x_k \to x$, there exists a strictly increasing mapping $\phi : \mathbb{N} \to \mathbb{N}$ such that $y_{\phi(k)} \to y$ and $y \in \Gamma(x)$.*

Let $X = \mathbb{N}^* \cup \{+\infty\}$ endowed with the topology induced by that of $\bar{\mathbb{R}}$. For each fixed $\tilde{\xi} \in \tilde{\Xi}$ we define the following set valued mappings

$$\Gamma_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^r, \quad N \mapsto \hat{F}_N(S, \tilde{\xi}) \tag{5}$$

with $\Gamma_{\tilde{\xi}}(+\infty) = \mathbb{E}_\xi[F(S, \cdot)]$, where $\hat{F}_N(\cdot, \tilde{\xi})$ has been introduced in (3).

$$\Lambda_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^r, \quad N \mapsto \hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi}) \tag{6}$$

with $\Lambda_{\tilde{\xi}}(+\infty) = \mathbb{E}_\xi[F(E, \cdot)]$.

$$\Upsilon_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^n, \quad N \mapsto E_N^w(\tilde{\xi}) \tag{7}$$

with $\Upsilon_{\tilde{\xi}}(+\infty) = E^w$.

The following Lemma will be useful in the next Sections.

Lemma 3.3. *Under (H_0, H_1, H_2, H_3), for almost all $\tilde{\xi} \in \tilde{\Xi}$, $\Gamma_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^r$ defined by (5) is continuous at $+\infty$. Moreover, $\Gamma_{\tilde{\xi}}$ is compact at $+\infty$.*

Proof. Let $\tilde{\xi} \in \tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B})$. By Lemma 3.2 $\lim\limits_{N\to\infty} \mathbb{H}\Big(\hat{F}_N(S, \tilde{\xi}), \mathbb{E}_\xi[F(S, \cdot)]\Big) = 0$ which means (Remark 3.7) that $\Gamma_{\tilde{\xi}}$ is H-continuous at $+\infty$. Moreover, $\Gamma_{\tilde{\xi}}(+\infty)$ is compact. Hence (Remark 3.8) $\Gamma_{\tilde{\xi}}$ is continuous at $+\infty$.
It remains to show that $\Gamma_{\tilde{\xi}}$ is compact at $+\infty$. Let $(N_k, y_k)_{k\geq 1}$ such that $N_k \to +\infty$ and $y_k \in \Gamma_{\tilde{\xi}}(N_k)$ (each k). Then there exists a sequence $(x_k)_{k\geq 1}$ in S such that $y_k = \hat{F}_N(x_k, \tilde{\xi})$ (each k). Since S is compact, there exists $\phi : \mathbb{N} \to \mathbb{N}$ strictly increasing such that $x_{\phi(k)} \to x$ and $x \in S$. By Lemma 3.1, $y_{\phi(k)} \to y = \mathbb{E}_\xi[F(x, \cdot)]$. Hence $y \in \mathbb{E}_\xi[F(S, \cdot)] = \Gamma_{\tilde{\xi}}(+\infty)$. □

4. Results in the decision space $\mathbb{R}^n$

In this section, we work with the weakly Pareto sets.

We say that a (MOP) is convex if all its objective functions are convex and its feasible set is convex. Using the well known Scalarization Theorem for convex (MOP) (see e.g. [29, Proposition 3.7 and Proposition 3.8], or [38]), we obtain immediately the following.

Theorem 4.1. *Under (H_4, H_5), we have*

$$\bigcup_{\lambda \in \mathbb{R}^r_+ \setminus \{0\}} \underset{x \in S}{\operatorname{argmin}} \, \langle \lambda, \mathbb{E}_\xi[F(x,\cdot)] \rangle = E^w.$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^r$.
Moreover, for each $N \in \mathbb{N}^*$ and for a.e $\tilde{\xi} \in \tilde{\Xi}$, we have

$$\bigcup_{\lambda \in \mathbb{R}^r_+ \setminus \{0\}} \underset{x \in S}{\operatorname{argmin}} \left\langle \lambda, \hat{F}_N(x, \tilde{\xi}) \right\rangle = E^w_N(\tilde{\xi}).$$

Remark 4.1. *By (H_5), there exists a set $\tilde{C} \subset \tilde{\Xi}$ with $\tilde{\mathbb{P}}_\xi(\tilde{C}) = 0$ such that for each fixed $\tilde{\xi} \in \tilde{\Xi} \setminus \tilde{C}$, $x \mapsto \hat{F}_N(x, \tilde{\xi})$ is strictly convex on S.*

Proposition 4.1. *The set E^w is compact, and for each $N \geq 1$, the SAA-N weakly Pareto sets $E^w_N(\tilde{\xi})$ is compact a.e. on $\tilde{\Xi}$.*

Proof. Let $N \in \mathbb{N}^*$ and $\tilde{\xi} \in \tilde{\Xi} \setminus \tilde{A}$. Since $x \mapsto \hat{F}_N(x, \tilde{\xi})$ is continuous on the closed set S, it is easy to show that $E^w_N(\tilde{\xi})$ is closed (see e.g. [12, Theorem 3.1] or [42]). Since it is a subset of the compact set S, it is compact. The same proof applies for E^w because $x \mapsto \mathbb{E}_\xi\big[F(x,\cdot)\big]$ is continuous. □

Proposition 4.2. *For each $N \geq 1$, the SAA-N weakly Pareto set $E^w_N(\tilde{\xi}) \neq \emptyset$ a.e. on $\tilde{\Xi}$, and $E^w \neq \emptyset$ as well.*

Proof. Let $N \geq 1$. Since S is compact and $x \mapsto \hat{F}_N(x, \tilde{\xi})$ is continuous on S for almost every $\tilde{\xi} \in \tilde{\Xi}$, the first conclusion follows easily by Theorem 4.1 and Weierstrass' Theorem. In the same way we obtain $E^w \neq \emptyset$. □

Now we state the main result of this section.

Theorem 4.2. *Under $(H_0, H_1, H_2, H_3, H_4, H_5)$, for almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the Hausdorff-Pompeiu distance between the SAA-N weakly Pareto sets $E^w_N(\tilde{\xi})$ and the true weakly Pareto set E^w tends to zero as N tends to infinity, i.e.*

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, \lim_{N \to \infty} \mathbb{H}\left(E^w_N(\tilde{\xi}), E^w\right) = 0\right\}\right) = 1$$

The proof of the Theorem is an immediate consequence of the following two Lemmas.

Lemma 4.1. *Under (H_0, H_1, H_2, H_3), for almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the deviation of the SAA-N weakly Pareto sets $E^w_N(\tilde{\xi})$ from the true weakly Pareto set E^w tends to zero as N tends to infinity. In other words,*

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, \lim_{N \to \infty} \mathbb{D}\left(E^w_N(\tilde{\xi}), E^w\right) = 0\right\}\right) = 1$$

Proof. Let $\tilde{\xi} \in \tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B})$ be fixed. Then, for each fixed $N \in \mathbb{N}^*$, the set $E^w_N(\tilde{\xi})$ is nonempty by Proposition 4.2, and it is compact by Proposition 4.1.
The set-valued mapping $\Gamma_{\tilde{\xi}}$ (introduced in (5)) is continuous at $+\infty$ (Lemma 3.3) and $\Gamma_{\tilde{\xi}}(+\infty)$ is compact (Proposition 4.1). By [42, Theorem 4.3] it follows that the set-valued mapping $\Upsilon_{\tilde{\xi}}$ (introduced in (7)) is u.c. at $+\infty$. Hence (Proposition 3.3) $\Upsilon_{\tilde{\xi}}$ is H-u.c. at $+\infty$. Remark 3.7 leads to the conclusion. □

Lemma 4.2. *Under $(H_0, H_1, H_2, H_3, H_4, H_5)$, for almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the deviation of the true weakly Pareto set E^w from the SAA-N weakly Pareto sets $E^w_N(\tilde{\xi})$ tends to zero as N tends to infinity, i.e*

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi}\middle| \lim_{N\to\infty} \mathbb{D}\left(E^w, E^w_N(\tilde{\xi})\right) = 0\right\}\right) = 1$$

Proof. Let $\tilde{\xi} \in \tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B} \cup \tilde{C})$ be fixed, where $\tilde{C}$ has been introduced in Remark 4.1. Let $\hat{x} \in E^w$. By the Scalarization Theorem 4.1, there exists $\lambda \in \mathbb{R}^r_+ \setminus \{0\}$ such that $\hat{x} \in \underset{x\in S}{\operatorname{argmin}} \langle \lambda, \mathbb{E}_\xi[F(x,\cdot)]\rangle$.
Now, consider for each $N \in \mathbb{N}^*$ an element

$$x_N \in \underset{x\in S}{\operatorname{argmin}} \left\langle \lambda, \hat{F}_N(x,\tilde{\xi})\right\rangle \tag{8}$$

which is possible since the last set is nonempty according to Weierstrass' Theorem. Thus we obtain a sequence (x_N) such that $x_N \in E^w_N(\tilde{\xi})$ according to the Scalarization Theorem. Since (x_N) lies in the compact set S, there exists a strictly increasing mapping $\phi : \mathbb{N} \to \mathbb{N}$ such that $\lim_k x_{\phi(k)} = \tilde{x}$, and $\tilde{x} \in S$. By (8), $x_{\phi(k)} \in \underset{x\in S}{\operatorname{argmin}} \left\langle \lambda, \hat{F}_{\phi(k)}(x,\tilde{\xi})\right\rangle \forall k \geq 1$ and then

$$\left\langle \lambda, \hat{F}_{\phi(k)}(x_{\phi(k)},\tilde{\xi})\right\rangle \leq \left\langle \lambda, \hat{F}_{\phi(k)}(\hat{x},\tilde{\xi})\right\rangle \tag{9}$$

Since $\hat{x} \in \underset{x\in S}{\operatorname{argmin}} \langle \lambda, \mathbb{E}_\xi[F(x,\cdot)]\rangle$, taking the limit in (9) implies

$$\langle \lambda, \mathbb{E}_\xi[F(\tilde{x},\cdot)]\rangle = \langle \lambda, \mathbb{E}_\xi[F(\hat{x},\cdot)]\rangle$$

By the strict convexity hypothesis, $\tilde{x} = \hat{x}$.
Since in a compact space a sequence having a unique cluster point converges, we obtain that $\lim_N x_N = \hat{x}$.

We have shown $d(\hat{x}, E^w_N(\tilde{\xi})) \to 0$ as $N \to +\infty$ for all $\hat{x} \in E^w$.

Since E^w is compact, using Ascoli-Arzelà Theorem as in Lemma 3.2, we easily get that $\mathbb{D}(E^w, E^w_N(\tilde{\xi})) \to 0$ as N tends to infinity for a.e. $\tilde{\xi} \in \tilde{\Xi}$. □

Let $\tilde{\xi} \in \tilde{\Xi}$ be fixed and consider the set-valued mapping

$$\Sigma_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^n, \quad N \mapsto \underset{x\in\Upsilon_{\tilde{\xi}}(N)}{\operatorname{argmin}} \hat{F}^0_N(x,\tilde{\xi}) \tag{10}$$

with $\Sigma_{\tilde{\xi}}(+\infty) = \underset{x\in\Upsilon_{\tilde{\xi}}(+\infty)}{\arg\min} \mathbb{E}_\xi[F^0(x,\cdot)]$, where $\Upsilon_{\tilde{\xi}}$ was defined by (7).

Remark 4.2. *For each $N \in \mathbb{N}^*$, and for almost all $\tilde{\xi} \in \tilde{\Xi}$, by Weierstrass' Theorerm, $\Sigma_{\tilde{\xi}}(N)$ is nonempty because $x \mapsto \hat{F}^0_N(x,\tilde{\xi})$ is continuous on the compact set $E^w_N(\tilde{\xi})$. The same rule obviously holds for $\Sigma_{\tilde{\xi}}(+\infty)$*

Now, we can introduce the following optimal value function

$$V_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \to \mathbb{R}, \quad N \mapsto \min_{x\in\Upsilon_{\tilde{\xi}}(N)} \hat{F}^0_N(x,\tilde{\xi}) \tag{11}$$

with $V_{\tilde{\xi}}(+\infty) = \min_{x\in\Upsilon_{\tilde{\xi}}(+\infty)} \mathbb{E}_\xi[F^0(x,\cdot)]$.

Lemma 4.3. *Let $A \subset \mathbb{R}^n$ be a closed set, and let $g : A \to \mathbb{R}$ be a continuous function. Then the set $\underset{x\in A}{\operatorname{argmin}}\, g(x)$ is closed.*

Proof. Let $x \in \overline{\operatorname{argmin}_{x\in A} g(x)}$, where $\bar{A}$ denotes the topological closure of a set A. There exists a sequence $(x_N)_{N\geq 1}$ in $\operatorname{argmin}_{x\in A} g(x)$ such that $x_N \to x$. Thus, $g(x) = \lim g(x_N) = g(x_N)$ (for each N) hence $x \in \Sigma$. □

Theorem 4.3. *Under $(H_0, H_1, H_2, H_3, H_4, H_5)$, for almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the sequence of SAA-N optimal values $(V_{\tilde{\xi}}(N))_{N\geq 1}$ converges to the true optimal value $V_{\tilde{\xi}}(+\infty)$.*
Moreover, for almost all $\tilde{\xi}$ in $\tilde{\Xi}$ and for each sequence $(x_N^*)_{N\geq 1}$ in $\Sigma_{\tilde{\xi}}(N)$, all cluster points of $(x_N^*)_{N\geq 1}$ belong to $\Sigma_{\tilde{\xi}}(+\infty)$.

Proof. By Remark 3.3, $\forall \tilde{\xi} \in \tilde{\Xi} \setminus \tilde{A}_0, x \mapsto \hat{F}_N^0(x, \tilde{\xi})$ is continous on S. By Remark 3.4, there exists a set $\tilde{B}_0 \subset \tilde{\Xi}$ with $\tilde{\mathbb{P}}_\xi(\tilde{B}_0) = 0$ such that for each fixed $\tilde{\xi} \in \tilde{\Xi} \setminus \tilde{B}_0$, $\hat{F}_N^0(\cdot, \tilde{\xi})$ converges to $\mathbb{E}_\xi[F^0(\cdot, \cdot)]$ uniformly on S.

Let $\tilde{\xi} \in \tilde{\Xi} \setminus (\tilde{A}_0 \cup \tilde{B}_0 \cup \tilde{A} \cup \tilde{B} \cup \tilde{C})$ be fixed. Since $\Upsilon_{\tilde{\xi}}$ is Hausdorff continuous at $+\infty$ (Theorem 4.2), it is H-u.c. at $+\infty$. Moreover, $\Upsilon_{\tilde{\xi}}$ is closed valued, thus it is closed by [32, Proposition 2.5.15]. Let $\epsilon > 0$. Since $\Upsilon_{\tilde{\xi}}$ is also H-l.c. at $+\infty$, by definition, $\exists N^0(\epsilon) \in \mathbb{N}^*$ such that $\forall N \geq N^0(\epsilon)$, $\Upsilon_{\tilde{\xi}}(+\infty) \subset \Upsilon_{\tilde{\xi}}(N) + B_\epsilon$, where B_ϵ denote the open ball of radius ϵ and center 0. Let $x \in \Sigma_{\tilde{\xi}}(+\infty)$ ($\Sigma_{\tilde{\xi}}(+\infty) \neq \emptyset$ by Remark 4.2). Obviously, $x \in \Upsilon_{\tilde{\xi}}(+\infty)$ and then $x \in \Upsilon_{\tilde{\xi}}(N) + B_\epsilon\ \forall N \geq N^0(\epsilon)$. It follows that $(x + B_\epsilon) \cap \Upsilon_{\tilde{\xi}}(N) \neq \emptyset$ for $N \geq N^0(\epsilon)$.

All the assumptions of [16, Proposition 4.4] are fulfilled, hence, on one hand $\Sigma_{\tilde{\xi}}$ is u.c. at $+\infty$, and on the other hand $V_{\tilde{\xi}}$ is continuous at $+\infty$ i.e. $V_{\tilde{\xi}}(N) \to V_{\tilde{\xi}}(+\infty)$.

Since $\Sigma_{\tilde{\xi}}$ is u.c. at $+\infty$, it is H-u.c. (Proposition 3.3). Hence, for N large enough $\Sigma_{\tilde{\xi}}(N) \subset \Sigma_{\tilde{\xi}}(+\infty) + B_\epsilon$. Moreover (Lemma 4.3) $\Sigma_{\tilde{\xi}}(+\infty)$ is a closed set. By [4, Theorem 5.2.4] and Remark 3.6, we have $\limsup_{N\to+\infty} \Sigma_{\tilde{\xi}}(N) \subset \Sigma_{\tilde{\xi}}(+\infty)$ which concludes the last sentence of the Theorem. □

5. Results in the outcome space $\mathbb{R}^r$

In this section, we work with the Pareto sets image, and we assume only (H_0, H_1, H_2, H_3).

Proposition 5.1. *For each $N \geq 1$, the SAA-N Pareto set $E_N(\tilde{\xi})$ is a nonempty bounded set a.e. on $\tilde{\Xi}$. The true Pareto set E is also nonempty and bounded.*

Proof. Let $N \in \mathbb{N}^*$, $\tilde{\xi} \in \tilde{\Xi} \setminus \tilde{A}$, and let $\lambda \in \mathbb{R}^r$ such that $\lambda_i > 0\ \forall i = 1, \ldots, r$. Since $x \mapsto \hat{F}_N(x, \tilde{\xi})$ is continuous, by Weierstrass' Theorem there exists $\tilde{x} \in \operatorname{argmin}_{x\in S} \left\langle \lambda, \hat{F}_N(x, \tilde{\xi}) \right\rangle$. If $\tilde{x} \notin E_N(\tilde{\xi})$ there exists $\hat{x} \in S$ such that $\hat{F}_N(\hat{x}, \tilde{\xi}) \leq \hat{F}_N(\tilde{x}, \tilde{\xi})$ and $\hat{F}_N(\hat{x}, \tilde{\xi}) \neq \hat{F}_N(\tilde{x}, \tilde{\xi})$. Hence $\left\langle \lambda, \hat{F}_N(\hat{x}, \tilde{\xi}) \right\rangle < \left\langle \lambda, \hat{F}_N(\tilde{x}, \tilde{\xi}) \right\rangle$, a contradiction. Finally $x \in E_N(\tilde{\xi})$. Since S is compact, the boundedness follows. The same rule holds for the true Pareto set E since $x \mapsto \mathbb{E}_\xi[F(x, \cdot)]$ is continuous. □

To prove the main results of this section, we need the following Lemmas.

Lemma 5.1. *Let $(A_N)_{N\geq 1}$ be a sequence of non-empty subsets of $\mathbb{R}^r$, and let A be a subset of $\mathbb{R}^r$. If $d(x, A_N) \to 0$ as $N \to \infty$ for all x in A, then $d(x, A_N) \to 0$ as $N \to \infty$ for all x in $\bar{A}$.*

Proof. Let $\bar{x} \in \bar{A}$. Then there exists a sequence $(x_k)_{k\geq 1}$ in A such that $d(x_k, \bar{x}) \to 0$ as $k \to \infty$. Let $\epsilon > 0$ be fixed. $\exists k$ such that $d(x_k, \bar{x}) < \frac{\epsilon}{2}$. Since $x_k \in A$, there exists N^0 such that : $\forall N \geq N^0,\ d(x_k, A_N) < \frac{\epsilon}{2}$. Since $d(\bar{x}, A_N) \leq d(\bar{x}, x_k) + d(x_k, A_N)$, we obtain $d(\bar{x}, A_N) < \epsilon$ for all $N \geq N^0$. □

Lemma 5.2. *For almost all $\tilde{\xi} \in \tilde{\Xi}$, for each $N \geq 1$ and for all $y \in \hat{F}_N(S, \tilde{\xi})$, there exists $\hat{y} \in \hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi})$ such that $\hat{y} \leq y$.*

Proof. Let $\tilde{\xi} \in \tilde{\Xi} \setminus \tilde{A}$ be fixed and $N \in \mathbb{N}^*$. Let $y \in \hat{F}_N(S, \tilde{\xi})$ and let $\lambda \in int(\mathbb{R}^r_+)$. Since $x \mapsto \hat{F}_N(x, \tilde{\xi})$ is continuous, the set $Z_y = \left\{ y' \in \mathbb{R}^r \middle| y' \in (y - \mathbb{R}^r_+) \cap \hat{F}_N(S, \tilde{\xi}) \right\}$ is nonempty and compact. Thus there exists $\hat{y} \in \underset{y' \in Z_y}{\operatorname{argmin}} \langle \lambda, y' \rangle$. Obviously, we have $\hat{y} \leq y$. If $\hat{y} \notin \hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi})$, there exists a z in $\hat{F}_N(S, \tilde{\xi})$ such that $z \leq \hat{y}$ and $z \neq \hat{y}$. Hence $z \in Z_y$ and $\langle \lambda, z \rangle < \langle \lambda, \hat{y} \rangle$ which is a contradiction. Therefore, $\hat{y} \in \hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi})$. □

Proposition 5.2. *For almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the deviation of the true Pareto set image $\mathbb{E}_\xi[F(E, \cdot)]$ from the SAA-N Pareto sets image $\hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi})$ tends to zero as N tends to infinity, i.e*

$$\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi} \in \tilde{\Xi} \Big| \lim_{N\to\infty} \mathbb{D}\Big(\mathbb{E}_\xi[F(E, \cdot)], \hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi})\Big) = 0\Big\}\Big) = 1$$

Proof. Let $x \in E$ and let $\tilde{\xi} \in \tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B})$ be fixed. By Lemma 5.2, there exists a sequence $(x_N)_{N\geq 1}$ verifying $x_N \in E_N(\tilde{\xi})$ and

$$\hat{F}_N(x_N, \tilde{\xi}) \leq \hat{F}_N(x, \tilde{\xi}) \text{ for each } N \geq 1 \tag{12}$$

On one hand, since $\hat{F}_N(S, \tilde{\xi})$ is compact for each $N \geq 1$, the sequence $(\hat{F}_N(x_N, \tilde{\xi}))_{N\geq 1}$ admits at least one cluster point. Let $\hat{y}$ be such a cluster point. On the other hand, $(x_N)_{N\geq 1}$ lies in the compact S. Hence there exists a strictly increasing mapping $\phi : \mathbb{N} \to \mathbb{N}$ such that $\lim_{k\to\infty} x_{\phi(k)} = \hat{x}$ and $\lim_{k\to\infty} \hat{F}_{\phi(k)}(x_{\phi(k)}, \tilde{\xi}) = \hat{y}$. Since $\tilde{\xi} \notin (\tilde{A} \cup \tilde{B})$ $\hat{y} = \mathbb{E}_\xi[F(\hat{x}, \cdot)]$.

By (12), for each $k \geq 1$, $\hat{F}_{\phi(k)}(x_{\phi(k)}, \tilde{\xi}) \leq \hat{F}_{\phi(k)}(x, \tilde{\xi})$. Passing to the limit implies $\mathbb{E}_\xi[F(\hat{x}, \cdot)] \leq \mathbb{E}_\xi[F(x, \cdot)]$, and since $x \in E$ we have $\mathbb{E}_\xi[F(\hat{x}, \cdot)] = \mathbb{E}_\xi[F(x, \cdot)] = \hat{y}$. Thus all the cluster points of $(\hat{F}_N(x_N, \tilde{\xi}))_{N\geq 1}$ coincide, hence $\lim_{N\to\infty} \hat{F}_N(x_N, \tilde{\xi}) = \mathbb{E}_\xi[F(x, \cdot)]$.

We have shown that for all x in E and for almost every $\tilde{\xi} \in \tilde{\Xi}$, $\lim_{N\to\infty} d(\mathbb{E}_\xi[F(x, \cdot)], \hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi})) = 0$. By Lemma 5.1 and using Ascoli-Arzelà Theorem as in Lemma 3.2, we can easily show that $\lim_{N\to\infty} \mathbb{D}(\overline{\mathbb{E}_\xi[F(E, \cdot)]}, \hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi})) = 0$ a.e. on $\tilde{\Xi}$.
Since $\sup_{y \in \mathbb{E}_\xi[F(E,\cdot)]} d(y, \hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi})) \leq \sup_{y \in \overline{\mathbb{E}_\xi[F(E,\cdot)]}} d(y, \hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi}))$, the rest of the proof is straightforward. □

Proposition 5.3. *For almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the deviation of the SAA-N Pareto sets image $\hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi})$ from the true Pareto set image $\mathbb{E}_\xi[F(E, \cdot)]$ tends to zero as N tends to infinity, i.e*

$$\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi} \in \tilde{\Xi} \Big| \lim_{N\to\infty} \mathbb{D}\Big(\hat{F}_N(E_N(\tilde{\xi}), \tilde{\xi}), \mathbb{E}_\xi[F(E, \cdot)]\Big) = 0\Big\}\Big) = 1$$

Proof. Let $\tilde{\xi} \in \tilde{\Xi} \setminus (\tilde{A} \cup \tilde{B})$ be fixed. Since $\Gamma_{\tilde{\xi}}$ is continuous at $+\infty$ (Lemma 3.3), $\Gamma_{\tilde{\xi}}$ is l.c. at $+\infty$. Hence by Proposition 3.2 $\Gamma_{\tilde{\xi}}(+\infty) \subset \liminf_{N\to+\infty} \Gamma_{\tilde{\xi}}(N)$. Moreover $\Gamma_{\tilde{\xi}}$ is compact at

$+\infty$ (Lemma 3.3). All the assumptions of [32, Theorem 3.5.29] are satisfied hence $\Lambda_{\tilde{\xi}}$ is u.c. at $+\infty$. By Proposition 3.3 $\Lambda_{\tilde{\xi}}$ is H-u.c. at $+\infty$. The conclusion follows by Remark 3.7. □

The proof of the following is straightforward :

Theorem 5.1. *For almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the Hausdorff-Pompeiu distance between the SAA-N Pareto sets image $\hat{F}_N(E_N(\tilde{\xi}),\tilde{\xi})$ and the true Pareto set image $\mathbb{E}_\xi[F(E,\cdot)]$ tends to zero as N tends to infinity, i.e*

$$\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi}\in\tilde{\Xi}\Big|\ \lim_{N\to\infty}\mathbb{H}\Big(\hat{F}_N(E_N(\tilde{\xi}),\tilde{\xi}),\mathbb{E}_\xi[F(E,\cdot)]\Big)=0\Big\}\Big)=1$$

For a.a. $\tilde{\xi}\in\tilde{\Xi}$ and for all $N\in\mathbb{N}^*$, let us denote

$$U_N(\tilde{\xi})=\inf_{x\in E_N(\tilde{\xi})} f(\hat{F}_N(x,\tilde{\xi})),\ U=\inf_{x\in E} f(\mathbb{E}_\xi[F(x,\cdot)])$$

Theorem 5.2. *For almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the sequence of SAA-N optimal values $(U_N(\tilde{\xi}))_{N\geq 1}$ converges to the true optimal value U.*

Proof. Let $\tilde{\xi}\in\tilde{\Xi}\setminus(\tilde{A}\cup\tilde{B})$. Let $(a_N)_{N\geq 1}$ be a sequence of positive numbers such that $a_N\to 0$. There exists a sequence $(y_N)_{N\geq 1}$ such that for all $N\geq 1,\ y_N\in\hat{F}_N(E_N(\tilde{\xi}),\tilde{\xi})$ and

$$f(y_N)<U_N(\tilde{\xi})+a_N\leq f(y_N)+a_N \tag{13}$$

By Lemma 3.2 there exists a compact set $K\subset\mathbb{R}^r$ such that $\hat{F}_N(S,\tilde{\xi})\subset K$ and $\mathbb{E}_\xi[F(S,\cdot)]\subset K$. Since f is continuous, the sequence $(U_N(\tilde{\xi}))_{N\geq 1}$ lies in the compact set $f(K)$ hence admits at least one cluster point. Let W be such a cluster point. There exists $\phi:\mathbb{N}\to\mathbb{N}$ strictly increasing such that $U_{\phi(N)}(\tilde{\xi})\to W$. Since $(y_{\phi(N)})_{N\geq 1}$ is in the compact K, there exists $\varphi:\mathbb{N}\to\mathbb{N}$ strictly increasing such that $y_{\phi(\varphi(N))}\to y$ and $y\in K$. By Proposition 5.3, $y\in\overline{\mathbb{E}_\xi[F(E,\cdot)]}$.

By (13), $W=\lim_{N\to\infty} f(y_{\phi(\varphi(N))})$. Since f is continuous, $W=f(y)$ and $W\in f(\overline{\mathbb{E}_\xi[F(E,\cdot)]})$. If $W<U$ then $y\in\overline{\mathbb{E}_\xi[F(E,\cdot)]}\setminus\mathbb{E}_\xi[F(E,\cdot)]$. Hence there exists a sequence $(z_k)_{k\geq 1}$ in $\mathbb{E}_\xi[F(E,\cdot)]$ such that $z_k\to y$. By continuity, for k large enough $f(z_k)<U$ which is a contradiction. Thus $W\geq U$. Now we suppose that $U\neq W$, i.e. $U<W$. Hence there exists $\hat{y}\in\mathbb{E}_\xi[F(E,\cdot)]$ such that

$$f(\hat{y})<W \tag{14}$$

By Proposition 5.2, there exists a sequence $(\hat{y}_{\phi(\varphi(N))})_{N\geq 1}$ such that $\hat{y}_{\phi(\varphi(N))}$ in $\hat{F}_N(E_{\phi(\varphi(N))}(\tilde{\xi}),\tilde{\xi})$ for each N and $\hat{y}_{\phi(\varphi(N))}\to\hat{y}$. Since f is continuous,

$$\lim_{N\to\infty} f(\hat{y}_{\phi(\varphi(N))})=f(\hat{y})$$

For N large enough, the last equality and (14) imply $f(\hat{y}_{\phi(\varphi(N))})<U_{\phi(\varphi(N))}(\tilde{\xi})$, a contradiction. Hence $U=W$ and all the cluster points of $(U_N(\tilde{\xi}))_{N\geq 1}$ coincide.
Finally, $\lim_{N\to\infty} U_N(\tilde{\xi})=U$. □

6. AN ILLUSTRATIVE EXAMPLE

Consider the following Stochastic Bi-Objective Convex Optimization Problem

$$\min_{x\in S} \mathbb{E}_\xi\Big[|\xi|(x_1+x_2), |\xi|(x_1-x_2)\Big]$$

where the feasible set $S=\{x=(x_1,x_2)\in\mathbb{R}^2 | x_1^2+x_2^2\leq 1\}$. To illustrate our approach, we consider that ξ follows a standard normal distribution. Our goal is to minimize the first objective over the Pareto set, i.e

$$\min_{x\in E} \mathbb{E}_\xi[|\xi|](x_1+x_2)$$

We deduce that

$$E=\{(cos\theta, sin\theta)\in\mathbb{R}^2 : \theta\in[\tfrac{3\pi}{4},\tfrac{5\pi}{4}]\}$$

and

$$\min_{x\in E} \mathbb{E}_\xi[|\xi|](x_1+x_2) = -\tfrac{2}{\sqrt{\pi}}$$

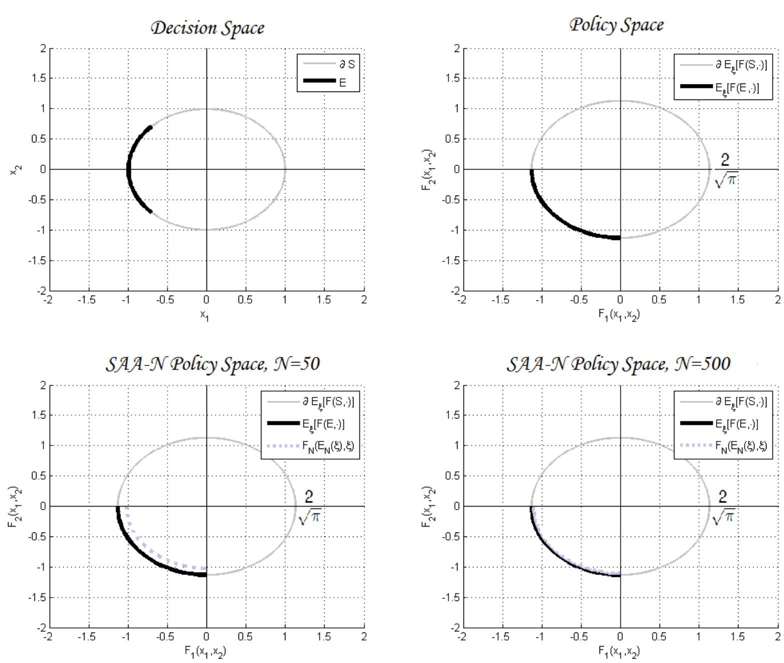
Numerically, let us solve the following SAA-N problem

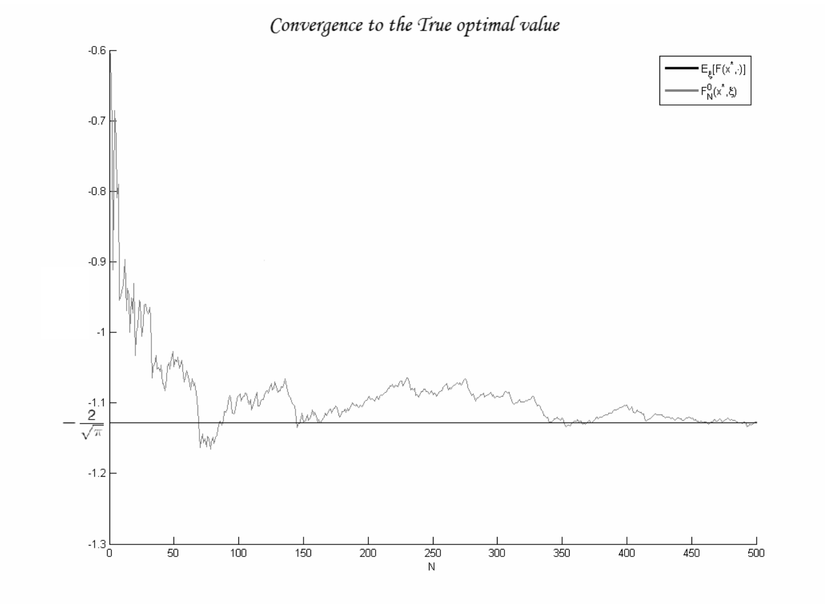
$$\min_{x\in E_N(\tilde{\xi})} \tfrac{1}{N}\textstyle\sum_{k=1}^N |\xi_k|(x_1+x_2)$$

where $E_N(\tilde{\xi})$ is the Pareto set associated with the following (SAA-N MOP)

$$\min_{x\in S} \tfrac{1}{N}\textstyle\sum_{k=1}^N |\xi_k|(x_1+x_2, x_1-x_2)$$

and the random process $\tilde{\xi}$ verifies hypothesis (H_1). Using MATLAB 7, we obtain the following numerical results





REFERENCES

[1] L. T. H. AN, L.D. MUU AND P.D. TAO, *Numerical solution for optimization over the efficient set by d.c. optimization algorithms,* Oper. Res. Lett., **19**, 117-128, 1996.

[2] Z. ANKHILI, A. MANSOURI,*An exact penalty on bilevel programs with linear vector optimization lower level,* Eur. J. Oper. Res. 197, 36-41 (2009)

[3] J.-P. AUBIN AND I. EKELAND, *Applied Nonlinear Analysis*, John Wiley & Sons, New York, 1984.

[4] J.-P. AUBIN AND H. FRANKOWSKA, *Set Valued Analysis*, Birkhäuser, Basel, 1990.

[5] H.P. BENSON, *Optimization over the Efficient Set,* J. Math. Anal. Appl. Volume **98**, pp. 562-580, 1984.

[6] H.P. BENSON AND D. LEE, *Outcome-Based Algorithm for Optimizing over the Efficient Set of a Bicriteria Linear Programming Problem,* J. Optim. Theory Appl. **88** 77-105, 1996.

[7] H.P. BENSON, *Generating the Efficient Outcome Set in Multiple Objective Linear Programs: The Bicriteria Case,* Acta Math. Vietnam. **22** 29-51, 1997.

[8] H.P. BENSON, *Further Analysis of an Outcome Set-Based Algorithm for Multiple Objective Linear Programming,* J. Optim. Theory Appl. **97** 1-10, 1998.

[9] H.P. BENSON, *Hybrid Approach for Solving Multiple Objective Linear Programs in Outcome Space,* J. Optim. Theory Appl. **98** 17-35, 1998.

[10] H. P. BENSON, *A finite, non-adjacent extreme point search algorithm for optimization over the efficient set,* J. Optim. Theory Appl., **73**, 47-64, 1992.

[11] S. BOLINTINÉANU, *Optimality conditions for minimization over the (weakly or properly) efficient set,* J. Math. Anal. Appl., **173**(2), 523-541, 1993.

[12] S. BOLINTINÉANU, *Necessary Conditions for Nonlinear Suboptimization over the Weakly-Efficient Set,* J. Optim. Theory Appl. **78**, 1993.

[13] S. BOLINTINÉANU, *Minimization of a quasi-concave function over an efficient set,* Math. Program., **61**, 89-110, 1993.

[14] S. BOLINTINÉANU, *Necessary conditions for nonlinear suboptimization over the weakly-efficient set,* J. Optim. Theory Appl., **78**, 579-598, 1993.

[15] S. BOLINTINÉANU (H. BONNEL), *Vector Variational Principles; ϵ-Efficiency and Scalar Stationarity,* J. Convex Anal. **8**, 71-85, 2001.

[16] J.F. BONNANS AND A. SHAPIRO, *Perturbation Analysis of Optimization Problem,* Springer Series in Operations Research, 2000.

[17] H. BONNEL AND C.Y. KAYA, *Optimization Over the Efficient Set of Multi-objective Control Problems,* J. Optim. Theory Appl. **147**(1), 93-112, 2010.

[18] H. BONNEL, J. MORGAN, *Semivectorial Bilevel Convex Optimal Control Problems: An Existence Result* (to appear in SIAM Journal on Control and Optimization)
[19] H. BONNEL, *Optimality Conditions for the Semivectorial Bilevel Optimization Problem*, Pacific Journal of Optimization, Vol. 2, No.3, pp. 447-468, 2006.
[20] H. BONNEL, J. MORGAN, *Semivectorial Bilevel Optimization Problem: Penalty Approach,* Journal of Optimization Theory and Applications, Vol. 131, No. 3, pp. 365-382, 2006.
[21] R.S. BURACHIK AND A.N. IUSEM, *Set-Valued Mappings and Enlargements of Monotone Operators,* Springer Optimization and Its Applications, 2008.
[22] F. H. CLARKE, *Optimization and Nonsmooth Analysis,* Wiley Interscience, New York, 1983.
[23] B. D. CRAVEN, *Aspects of multicriteria optimization,* Recent Developments in Mathematical Programming, 93-100, 1991.
[24] J. P. DAUER, *Optimization over the efficient set using an active constraint approach,* J. Oper. Res., **35**, 185-195, 1991.
[25] J. P. DAUER AND T. A. FOSNAUGH, *Optimization over the efficient set,* J. Global Optim., **7**, 261-277, 1995.
[26] S. DEMPE, N. GADHI, A.B. ZEMKOHO, *New optimality conditions for the semivectorial bilevel optimization problem,* www.optimization-online.org/DB_HTML/2011/11/3228.html
[27] J. DIEUDONNÉ, *Foundations of modern analysis,* Academic Press, New York and London, 1960.
[28] G. EICHFELDER, *Multiobjective bilevel optimization,* Math. Program., Ser. A (2010) 123:419-449
[29] M. EHRGOTT, *Multicriteria Optimization,* Springer-Verlag Berlin Heidelberg, 2000.
[30] J. FLIEGE AND H. XU, *Stochastic Multiobjective Optimization : Sample Average Approximation and Applications* J. Optim. Theory Appl. **151** 135-162, 2011.
[31] J. FÜLÖP, *A cutting plane algorithm for linear optimization over the efficient set,* Generalized Convexity, Lecture notes in Economics and Mathematical System 405, Springer-Verlag, Berlin, 374-385, 1994.
[32] A. GÖPFERT, H. RIAHI, C. TAMMER AND C. ZĂLINESCU, *Variational Methods in Partially Ordered Spaces,* Springer, Berlin, 2003.
[33] GUANG-YA CHEN, XUEXIANG HUANG AND XIAOQI YANG, *Vector Optimization : Set Valued and Variational Analysis,* Springer, Berlin, 2005.
[34] R. HORST AND N. V. THOAI, *Maximizing a concave function over the efficient or weakly-efficient set,* European J. Oper. Res., **117**, 239-252, 1999.
[35] R. HORST, N. V. THOAI, Y. YAMAMOTO, AND D. ZENKE, *On Optimization over the Efficient Set in Linear Multicriteria Programming,* J. Optim. Theory Appl., **134**, 433-443, 2007.
[36] H. ISERMANN, *Proper Efficiency and the Linear Vector Maximization Problem,* Oper. Res., **22**, 189-191, 1974.
[37] J. JACOD, P. PROTTER, *Probability Essentials,* Springer-Verlag Berlin Heidelberg, 2004.
[38] J. JAHN, *Vector Optimization,* Springer-Verlag, Berlin, Heidelberg, New York, 2004.
[39] J. JAHN, *Introduction to the Theory of Nonlinear Optimization,* Springer-Verlag, Berlin, Heidelberg, New York, 2007.
[40] N.T.B. KIM AND T.N. THANG, *Optimization over the Efficient Set of a Bicriteria Convex Programming Problem,* (to appear in Pac. J. Optim.)
[41] H. KONNO, P.T. THACH AND D. YOKOTA, *Dual Approach to Minimization on the Set of Pareto-Optimal Solutions,* J. Optim. Theory Appl. **88** 689-707, 1996.
[42] D. T. LUC, *Theory of Vector Optimization,* Lecture Notes in Econom. and Math. Systems 319, Springer-Verlag, Berlin, Heidelberg, 1989.
[43] K. M. MIETTINEN, *Nonlinear Multiobjective Optimization,* Kluwer Academic Publishers, 1998.
[44] J. PHILIP, *Algorithms for the vector maximization problem,* Math. Program., 207-229, 1972.
[45] R.T. ROCKAFELLAR, *Convex Analysis,* Princeton, New Jersey, Princeton University Press, 1972.
[46] R.T. ROCKAFELLAR, ROGER J-B WETS, *Variational Analysis,* Springer, Heidelberg, London, New York, **317**, 1998.
[47] A. SHAPIRO, D. DENTCHEVA, A. RUSZCZYNSKI, *Lectures on stochastic programming : modeling and theory,* MPS/SIAM Ser. Optim., 2009.
[48] Y. YAMAMOTO, *Optimization over the efficient set : an overview,* J. Global Optim., **22**, 285-317, 2002.
[49] Y.ZHENG AND Z. WAN, *A solution method for semivectorial bilevel programming problem via penalty method,* J. Appl. Math. Comput. 37 (2011), 207-219.

ERIM, UNIVERSITY OF NEW CALEDONIA, B.P. R4, F98851 NOUMEA CEDEX, NEW CALEDONIA
E-mail address: henri.bonnel@univ-nc.nc, julien.collonge@univ-nc.nc