Analyse post-Pareto en optimisation vectorielle stochastique et déterministe : étude théorique et algorithmes.

Julien Collonge

École Doctorale du Pacifique

Thèse de doctorat

Discipline : Mathématiques

présentée par

Julien Collonge

Analyse post-Pareto en Optimisation Vectorielle Stochastique et Déterministe : Étude Théorique et Algorithmes

dirigée par **Henri Bonnel**

Soutenue le 12 novembre 2014 devant le Jury composé de :

Rapporteurs

Régina S Burachik, Associate Professor	University of South Australia
Jörg Fliege, Professor	University of Southampton

Examinateurs

Yalcin Kaya, Senior Lecturer	University of South Australia
Ioane Muni Toke, Maître de Conférences	Université de la Nouvelle-Calédonie
Michel Théra, Professeur (Président du Jury)	Université de Limoges

Directeur

Henri Bonnel, Professeur	Université de la Nouvelle-Calédonie

Université de la Nouvelle-Calédonie
BP R4, Avenue James Cook, Nouméa 98851, Nouvelle-Calédonie

Équipe de Recherche en Informatique et Mathématiques (ERIM)
École Doctorale du Pacifique
BP R4, Avenue James Cook, Nouméa 98851, Nouvelle-Calédonie

If you can't explain it simply, you
don't understand it well enough.
(Albert Einstein)

Remerciements

Tout d'abord, je tiens à remercier Henri Bonnel, mon directeur de thèse. Pendant ces années, j'ai eu la chance de profiter de ses immenses connaissances Mathématiques et de ses remarques très enrichissantes et toujours très pertinentes. Je voudrais lui exprimer toute ma gratitude pour sa disponibilité et son soutien.

Cette thèse a été effectuée au sein de l'Université de Nouvelle-Calédonie que je souhaiterai remercier pour avoir financé ces recherches, sans oublier l'école Doctorale qui m'a donné l'opportunité d'assister au vingt-et-unième International Symposium on Mathematical Programming in Berlin (2012).

Je suis reconnaissant à Michel Théra pour avoir appuyé ma candidature au Franco-Thai Symposium (Bangkok 2012).

Je suis également reconnaissant à Regina Burachik et à Jörg Fliege d'avoir accepté la charge d'être rapporteur de cette thèse.

Merci à Ioane Muni Toke, à Yalcin Kaya et à Michel Théra d'être examinateurs de cette thèse ; j'éprouve un profond respect pour leur travail et leur parcours.

Mes remerciements vont également à mes collègues de l'Équipe de Recherche en Informatique et Mathématiques (ERIM) pour leur bonne humeur et avec qui j'ai pu passer d'agréables moments.

Un merci particulier à Michel Théra, Regina Burachik, Yalcin Kaya, et Jörg Fliege qui ont accepté d'être membre du jury malgré le déplacement.

Résumé - Abstract

Mots-clefs

Optimisation sur l'ensemble de Pareto - Optimisation sur le front de Pareto
Optimisation Vectorielle Déterministe - Optimisation Vectorielle Stochastique
Optimisation Vectorielle Convexe - Méthode de Monte-Carlo

Keywords

Optimization over a Pareto set - Optimization over the Pareto image set
Deterministic Vector Optimization - Stochastic Vector Optimization
Convex Vector Optimization - Sample Average Approximation Method

Résumé

Cette thèse relate certains aspects liés à l'analyse post-Pareto issue de Problèmes d'Optimisation Vectorielle.

Un Problème d'Optimisation Vectorielle (POV) consiste à optimiser une fonction vectorielle F définie sur un ensemble arbitraire G, et à valeurs dans un espace vectoriel ordonné H. Lorsque $H = \mathbb{R}^r$ est muni de l'ordre produit, nous parlons de Problème d'Optimisation Multicritères (POM). Plus précisément, si $\emptyset \neq X \subset G$ est l'ensemble admissible du (POM), et si le (POM) est un problème de minimisation déterministe, nous notons

$$(POM) \qquad \mathrm{MIN}_{x\in X}\Big(F(x) = (F^1(x), F^2(x), \ldots, F^r(x))\Big).$$

L'ensemble des solutions de ce problème (appelé ensemble de Pareto) est composé des solutions admissibles qui assurent un certain équilibre entre les objectifs : il est impossible d'améliorer la valeur d'un objectif sans détériorer celle d'un autre. Autrement dit, $x^* \in X$ est une solution de Pareto pour (POM) si, et seulement si, il n'existe pas de $x \in X$ tel que

$$F^j(x) \leq F^j(x^*) \ \forall j = \{1, \ldots, r\} \text{ et } F^i(x) < F^i(x^*) \text{ pour un } i \in \{1, \ldots, r\}$$

L'ensemble de Pareto est souvent très grand (infini) et son image par la fonction vectorielle F est une partie de la frontière de l'image $F(X)$ (la partie sud-ouest lorsque $r = 2$).

D'un point de vue technique, chaque solution de Pareto est acceptable.

Nous nous posons alors le problème de la sélection de l'une d'entre elles : en supposant l'existence d'un décideur qui aurait son propre critère de décision $F^0 : G \to \mathbb{R}$, nous considérons le problème post-Pareto [1] (D) de minimiser cette fonctionnelle sur l'ensemble de Pareto du (POM). Autrement dit, nous considérons le Problème d'Optimisation

$$(D) \qquad \min_{x\in P} F^0(x),$$

où P représente l'ensemble des solutions de Pareto du (POM).

Dans le cas Déterministe, le problème d'optimiser sur un ensemble de Pareto a fait l'objet de nombreuses études ces dernières décennies. Ces problèmes sont considérés comme difficiles car l'ensemble de Pareto est généralement non-convexe, il peut-être non-borné, et est rarement explicitement connu.

1. Johan Philip fut le premier en 1972 ([102]) à considérer un Problème d'Optimisation post-Pareto, et a inspiré de nombreux papiers [1, 7, 16, 17, 18, 13, 25, 42, 46, 47, 70, 77, 78] (voir [113] pour une bibliographie significative)

À l'inverse, le cas d'un Problème d'Optimisation post-Pareto Stochastique n'a été abordé que très récemment en raison de la difficulté supplémentaire pour prendre en compte l'incertitude.

Dans cette thèse, nous considérons le Problème d'Optimisation post-Pareto Stochastique (S) qui vise à minimiser l'espérance Mathématique d'une fonction aléatoire réelle sur l'ensemble de Pareto d'un (POM) Stochastique.

Autrement dit, pour un vecteur aléatoire ξ donné, nous nous posons le Problème d'Optimisation suivant

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

où E est l'ensemble des solutions de Pareto du $(POMS)$

$$\text{MIN}_{x \in X}\Big(\mathbb{E}[F(x, \xi(\cdot))] = (\mathbb{E}[F^1(x, \xi(\cdot))], \mathbb{E}[F^2(x, \xi(\cdot))], \ldots, \mathbb{E}[F^r(x, \xi(\cdot))])\Big),$$

où pour chaque $x \in X$ fixé, $\mathbb{E}[F^j(x, \xi(\cdot))]$ $(j \in \{0, \ldots, r\})]$ est l'espérance Mathématique de la variable aléatoire réelle $\omega \mapsto F^j(x, \xi(\omega))$.

Comme il est rarement possible de pouvoir définir explicitement l'espérance Mathématique d'une fonction aléatoire, ces Problèmes d'Optimisation post-Pareto Stochastiques sont très difficiles à résoudre.

Notre stratégie de résolution est la suivante :

- introduire une suite de problèmes déterministes équivalents obtenus à l'aide de la méthode de Monte-Carlo : dans le Chapitre 2 nous montrons que toute valeur d'adhérence de toute suite de solutions optimales (resp. la suite de valeurs optimales) issue des problèmes obtenus par Monte-Carlo converge presque sûrement vers une solution optimale (resp. la valeur optimale) du problème (S).

- proposer un algorithme pour encadrer la valeur optimale d'un Problème d'Optimisation post-Pareto Déterministe : dans le Chapitre 3 nous prouvons que notre algorithme finit toujours en un nombre fini d'étapes, puis nous montrons qu'il est aisément adaptable au cas post-Pareto Stochastique.

Post-Pareto Analysis in Stochastic Multi-Objective Optimization : Theoretical Results and Algorithms

Abstract

This thesis explore related aspects to post-Pareto analysis arising from Vector Optimization Problem.

A Vector Optimization Problem (POV) is to optimize a vector objective function F defined on an arbitrary set G, and taking values in a partially ordered set H. When $H = \mathbb{R}^r$ is endowed with the product ordering, we talk about Multi-Criteria Optimization Problem (MOP). For a Deterministic Minimization Problem, we note

$$(MOP) \qquad \text{MIN}_{x\in X}\Big(F(x) = (F^1(x), F^2(x), \dots, F^r(x))\Big),$$

where $\emptyset \neq X \subset G$ is its feasible set.

Its solution set (called Pareto set) consists of the feasible solutions which ensure some sort of equilibrium amongst the objectives. That is to say, Pareto solutions are such that none of the objectives values can be improved further without deteriorating another. In other words, $x^* \in X$ is said to be Pareto solution for (MOP) if, and only if there is no element $x \in X$ satisfying

$$F^j(x) \leq F^j(x^*)\ \forall j = \{1, \dots, r\} \text{ and } \exists\ i \in \{1, \dots, r\} F^i(x) < F^i(x^*).$$

The Pareto set is often very large (may be infinite, and even unbounded) and its image by the vector objective function F is a part of the boundary of the image $F(X)$ (in the case $r = 2$ the southwestern part).

Technically speaking, each Pareto solution is acceptable.

The natural question that arises is: how to choose one solution? One possible answer is to optimize an other objective over the Pareto set [2]. Considering the

2. Johan Philip was the first in 1972 [102] to consider post-Pareto Optimization Problem, and has been followed by many authors in many papers [1, 7, 16, 17, 18, 13, 25, 42, 46, 47, 70, 77, 78] (see [113] for an extensive bibliography)

existence of a decision-maker with its own criteria $F^0 : G \to \mathbb{R}$, we deal with the post-Pareto Optimization Problem (D) of minimizing the real-valued function F^0 over the Pareto set. That is to say, we deal with the following Optimization Problem:

$$(D) \qquad \min_{x \in P} F^0(x),$$

where P is the set of Pareto solutions associated with (MOP).

In the Deterministic case, the problem of optimizing a scalar objective over the Pareto set has been intensively studied the last decades. This problem is difficult due to the fact that the Pareto set is not described explicitly and is generally not convex (even in the linear case).

Contrast to the Deterministic case, post-Pareto Stochastic Optimization Problem has just been studied due to the further difficulty to take into account the uncertainty.

In this thesis, we consider the post-Pareto Stochastic Optimization Problem (S) of minimizing the expectation of a scalar random function over the Pareto set associated with a Stochastic (MOP), where each objective is given by the expectation of a scalar random function. In other words, for a given random vector ξ, we deal with the following Optimization Problem:

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

where E is the set of Pareto solutions associated with the following $(SMOP)$

$$\text{MIN}_{x \in X} \Big(\mathbb{E}[F(x, \xi(\cdot))] = (\mathbb{E}[F^1(x, \xi(\cdot))], \mathbb{E}[F^2(x, \xi(\cdot))], \ldots, \mathbb{E}[F^r(x, \xi(\cdot))]) \Big),$$

where for each fixed $x \in X$, $\mathbb{E}[F^j(x, \xi(\cdot))]$ $(j \in \{0, \ldots, r\})$ is the expectation of the scalar random variable $\omega \mapsto F^j(x, \xi(\omega))$.

Since the closed form of an expected value function is rarely possible to compute, post-Pareto Stochastic Optimization Problems are very difficult to solve.

Our resolution strategy is as follows:

– to introduce a sequence of equivalents Deterministic Problems obtained with the Sample Average Approximation method: in Chapter 2 we show that every cluster point of any sequence of optimal solutions (resp. the sequence of optimal values) of the Sample Average Approximation problems converges almost surely to an optimal solution (resp. the optimal value) of problem (S).

- to look for the optimal value of a post-Pareto Optimization Problem: in Chapter 3 we first propose an outcome deterministic algorithm, and we prove that it always terminates in a finite number of steps. Finally we show that it can be easily suited to the post-Pareto Stochastic case.

Table des matières

Introduction

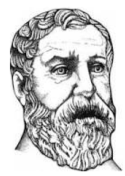
Un trajet, la forme d'un objet, le contrôle aérien, la consommation d'un moteur, la régulation des feux tricolores, le choix d'un investissement économique, ainsi que tous les systèmes susceptibles d'être modélisés sont optimisés. Pourquoi se priver d'une configuration idéale permettant un gain de temps, d'argent et d'énergie ? La meilleure configuration existe-t-elle ? Comment la déterminer ?

Les réponses à ces questions se trouvent dans le vaste chapitre de l'Optimisation à un ou plusieurs critères, qui inclut parmi d'autres l'optimisation en dimension finie ou infinie, l'optimisation continue ou combinatoire, l'optimisation linéaire, quadratique, convexe, algébrique, à plusieurs niveaux...

Cette branche des Mathématiques découpée en une multitude de sous disciplines qui s'entrecroisent a connu un essor important ces dernières années avec l'avènement de l'ordinateur, et connaît une abondance d'applications dans bien des domaines (Chimie, Physique, Aéronautique, Informatique, Militaire...). Cela vient du principe même de l'Optimisation : choisir parmi l'ensemble des alternatives réalisables celle qui permettra d'atteindre la *meilleure configuration possible*, ou tout au moins trouver le *meilleur compromis* entre plusieurs exigences souvent contradictoires. Parler d'Optimisation signifie donc implicitement parler de choix et de critères de comparaison qui permettront de les comparer.

Les premiers Problèmes d'Optimisation auraient été formulés par le mathématicien Grec **Euclide**[3] au IIIe siècle avant notre ère. L'histoire retient également le nom d'**Héron d'Alexandrie** (année 100 après J.C.) qui énonça le *principe du plus court chemin* dans le contexte de l'optique : "Le chemin le plus court qui lie un point P à un point Q et qui contient un point d'une droite d donnée, est tel qu'au point de réflexion sur la droite d, l'angle incident égal l'angle réfléchi."

Héron d'Alexandrie



L'application qui intervient dans ce principe associe à chaque point M de d la somme des distances de P à M et de M à Q. **H. d'Alexandrie** a explicitement déterminé quel point de d la rend minimale.

3. Euclide a rédigé l'un des plus célèbres textes de l'histoire des Mathématiques : les Éléments d'Euclide, œuvre constituée de 13 livres organisés thématiquement.

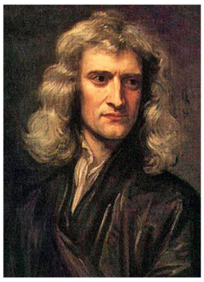
Il semblerait que ce soit les premières traces écrites de l'Optimisation, ou tout du moins le premier problème a avoir été considéré scientifiquement.

Pierre de Fermat

Au début du XVIIe siècle, le Prince des amateurs **Pierre de Fermat** (1605 ?-1665) s'est intéressé au *principe du plus court chemin*. Partant du fait que la nature choisit toujours la voie la plus courte et la plus simple, il modifia son énoncé par un principe Physique qui porte aujourd'hui son nom : "La lumière se propage d'un point à un autre sur des trajectoires telles que la durée du parcours soit localement minimale".

P. de Fermat fut l'un des premiers à avoir étudié des courbes par le biais d'équations et à se poser le problème de tangence aux courbes (on parlait à l'époque de touchante plutôt que de tangente). Ce fut le point de départ de la notion de nombre dérivé et du calcul d'extremums d'une fonction. Sa méthode des extremums (1638) (Condition Nécessaire d'optimalité pour des polynômes réels d'une variable réelle) deviendra par la suite le fondement du calcul différentiel.

Isaac Newton



Gottfried Wilhelm Leibniz

A la fin du XVIIe siècle, les premiers outils de résolution de certains Problèmes d'Optimisation relatifs à la Géométrie et à la Physique furent donnés par l'allemand **Gottfried Wilhelm Leibniz** (1646-1716) qui énonça une Condition Nécessaire du premier ordre sans contrainte : "la dérivée d'une application est nécessairement nulle en un point de minimum local", ainsi que par l'anglais **Isaac Newton** (1643-1727) qui mit au point une méthode itérative pour déterminer les zéros locaux d'une application. Quarante ans après les travaux de **P. de Fermat**, ces deux figures emblématiques des Mathématiques ont "inventé", simultanément mais néanmoins indépendamment, le Calcul Différentiel. Le procès de 1713 entre les deux hommes pour en déterminer la paternité est resté célèbre.

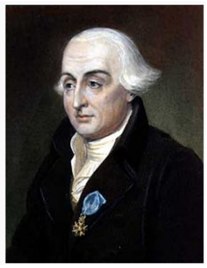
Johann Bernoulli

En 1696, **Johann Bernoulli** (1667-1748), mathématicien et physicien Suisse, se pose le problème de déterminer parmi toutes les courbes joignant deux points donnés dans un plan vertical, celle que devra nécessairement suivre un point matériel placé dans un champ de pesanteur uniforme, glissant sans frottement et sans vitesse initiale, pour qu'il présente un temps de parcours minimal entre ces deux points.

Durant le XVIIIe siècle, les travaux des mathématiciens **Leonhard Euler** (1707-1783) et **Joseph-Louis Lagrange** (1736-1813) mènent au Calcul Variationnel, une branche de l'Analyse Fonctionnelle regroupant certaines techniques d'Optimisation.

Leonhard Euler

Joseph-Louis Lagrange



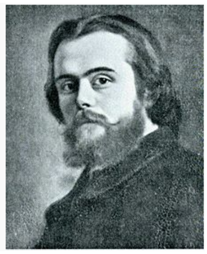
En 1744, **L. Euler** apporte une contribution particulièrement fondatrice dans le calcul des variations, qui inclut l'un de ses résultats les plus célèbres : l'équation d'Euler-Lagrange.

J.L. Lagrange s'appuie sur ces résultats préliminaires pour étendre en 1797 la Condition Nécessaire d'optimalité de **G.W. Leibniz** à certains Problèmes d'Optimisation avec contraintes d'égalité.

Tout le XIXe siècle fut marqué par l'intérêt croissant des économistes pour les Mathématiques. Ceux-ci mettent en place des modèles qu'il convient d'optimiser. Jusqu'alors, le développement des méthodes Mathématiques n'avait jamais été si important.

Antoine-Augustin Cournot

Léon Walras



Antoine-Augustin Cournot (1801-1877), mathématicien et économiste Français, a influencé les réflexions de **Léon Walras** (1834-1910), désigné plus tard comme le plus grand de tous les économistes. Dans son livre "Recherches sur les principes mathématiques de la théorie des richesses (1838)", il définit pour la première fois les mécanismes du marché (l'offre et la demande) comme des relations entre des fonctions Mathématiques. Ses théories sur l'offre et la demande sont aujourd'hui devenues de grands classiques.

En dépit des développements technologiques incessants, l'Optimisation Mono-Objectif a montré ses limites dans des situations très fréquentes où plus d'un objectif est à considérer : la qualité des résultats et des prédictions dépend de la pertinence du modèle, et certains systèmes sont si complexes qu'il est impossible de les modéliser correctement sans considérer simultanément plusieurs objectifs.

Il faudra cependant attendre la fin du XIXe siècle pour que des méthodes d'aide à la décision soient mises en place pour des problèmes faisant intervenir plusieurs objectifs. Mais les objectifs sont souvent en conflit les uns avec les autres, et une *"meilleure"* solution pour un objectif peut s'avérer plutôt mauvaise pour les autres.

C'est le cœur de la difficulté dans l'approche Multi-Objectifs.

Francis Ysidro Edgeworth

Vilfredo Pareto

La notion d'optimalité perd de ce fait tout son sens, laissant place à celle d'efficience. Les racines de cette nouvelle notion se trouvent dans les travaux des économistes **Francis Ysidro Edgeworth** (1845-1926) et **Vilfredo Pareto**[4].

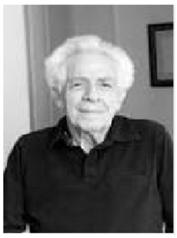
V. Pareto a apporté de nombreuses contributions importantes, notamment en économie, et plus particulièrement dans l'analyse des choix individuels. En

4. économiste italien né à Paris le 15 juillet 1848 et mort à Céligny (Suisse) le 19 août 1923.

1906, il a introduit le concept d'équilibre qui porte aujourd'hui son nom : état de la société dans lequel on ne peut pas améliorer le bien-être d'un individu sans détériorer celui d'un autre.

C'est en 1939, dans son mémoire de Master, que **William Karush** (1917-1997) généralise la méthode des multiplicateurs de Lagrange qui ne permettait jusque là que des contraintes d'égalité. **W. Karush** est également connu pour sa contribution dans les conditions de Karush-Kuhn-Tucker.

Harold William Kuhn



Albert William Tucker

Il faudra attendre 1951 et les travaux de **Harold W. Kuhn** (1925-2014) et **Albert W. Tucker** (1905-1995) pour que ces Problèmes Multi-Objectifs soient mathématiquement développés. Depuis cette période, l'Optimisation est devenue un pilier des Mathématiques Appliquées et a fait l'objet de nombreuses études, jouant aujourd'hui un rôle fondamental en économie et microéconomie, dans l'industrie et l'ingénierie, en statistiques, en théorie des jeux, ou encore en théorie du contrôle.

En théorie comme en pratique, les résultats concernant l'Optimisation Mono-Objectif sont nombreux, et souvent le *"meilleur"* élément existe et peut même être déterminé explicitement.

Cependant, les modèles considérés n'ont rarement qu'un seul objectif : **plusieurs critères souvent contradictoires doivent être optimisés simultanément.** C'est le principe même de l'Optimisation Multi-Objectifs.

Le nombre de solutions efficientes d'un Problème d'Optimisation Multi-Objectifs est rarement fini : toute solution efficiente (ou de Pareto) est optimale au sens où aucune amélioration ne peut être effectuée sur un objectif sans en dégrader au moins un autre. De ce fait, les solutions de Pareto ne sont pas comparables entre elles, et d'un point de vue technique, chacune constitue une *"meilleure"* solution. Nous nous posons naturellement la question suivante :

Comment prendre la meilleure décision quand elle n'existe pas ?

En pratique, **il est souvent nécessaire d'en choisir UNE**. Si chaque objectif est la représentation du critère d'un agent humain, il est très sage pour le décideur de faire son choix sur l'ensemble de Pareto car personne ne pourra le lui reprocher. En effet, si un agent demandait l'amélioration de son critère, le décideur lui rétorquerait que le cas échéant, il dégraderait forcément celui d'un autre agent.

Il est alors nécessaire de mettre en place une méthode pour prendre la décision la plus juste ou la plus souhaitable.

Pour cela, **Johan Philip** propose en 1972 [102] de faire appel à un autre critère d'évaluation [5] qu'il convient d'optimiser sur l'ensemble de Pareto. Depuis lors, ces Problèmes d'Optimisation post-Pareto connaissent un essor fulgurant dans tous secteurs concernés par ces Problèmes Multi-Objectifs de grande envergure, et pour lesquels les décisions doivent être optimales. Ainsi, le travail Mathématiques a un effet direct sur la mise au point d'algorithmes répondant spécifiquement aux besoins de l'industrie.

En Optimisation Déterministe, les fonctions et les variables sont explicitement connues, mais l'incertitude est inhérente à la réalité où les phénomènes observés sont perturbés aléatoirement. C'est justement ce qui a donné naissance à l'Optimisation Stochastique, dont le but est précisément d'étudier des Problèmes impliquant des Objectifs aléatoires.

Prendre une décision optimale dans un environnement évolutif et aléatoire est un problème soulevé aussi bien par les agents que par le décideur. Trouver une réponse à cette problématique est un véritable défi pour les chercheurs.

Numériquement très populaire, la Méthode de Monte-Carlo est reconnue comme l'une des méthodes les plus efficaces pour résoudre des Problèmes d'Optimisation Stochastique à un seul Objectif. Cependant, cette Méthode n'a reçu que très peu d'attention pour des Problèmes Multi-Objectifs Stochastique.

C'est un des aspects étudié dans cette thèse : nous considérerons un Problème d'Analyse post-Pareto Stochastique. Plus précisément, nous nous poserons le problème de minimiser l'espérance Mathématique d'une fonction aléatoire réelle sur l'ensemble de Pareto d'un Problème d'Optimisation Multi-Objectifs Stochastique, dont chaque objectif est l'espérance Mathématique d'une fonction aléatoire réelle.

Jusqu'à présent, seul le cas Déterministe a été considéré. Résoudre un Problème d'Optimisation post-Pareto Déterministe est déjà très difficile car l'ensemble de Pareto n'est généralement pas convexe, et peut même être non borné. Le cas Stochastique l'est encore plus car les fonctions considérées dépendent non seulement des variables de décisions, mais aussi de variables aléatoires. Comme l'espérance Mathématique d'une fonction aléatoire est très souvent abstraite, ces Problèmes d'Optimisations post-Pareto Stochastique sont très souvent impossibles à résoudre directement.

5. celui du décideur par exemple

Ce mémoire regroupe les travaux de recherche qui consistent d'une part à proposer une suite de problèmes déterministes obtenus via la méthode de Monte-Carlo, et d'autre part à mettre en place un modèle numérique pour déterminer leurs valeurs optimales.

Notre étude sera échelonnée comme suit :

1. Le type d'approche que nous développerons vise à minimiser l'espérance d'une fonction aléatoire réelle sur l'ensemble de Pareto associé à un Problème d'Optimisation Multi-Objectifs Stochastique (POMS), dont les objectifs sont définis par des espérances de fonctions aléatoires. Autrement dit, nous nous intéresserons au Problème d'Optimisation post-Pareto Stochastique :

$$(S) \qquad \min\left\{\mathbb{E}\Big[F^0\Big(x,\xi(\cdot)\Big)\Big]\Big|x\in \mathrm{ARGMIN}_{\mathbb{R}^r_+}\mathbb{E}\Big[F\Big(X,\xi(\cdot)\Big)\Big]\right\},$$

   où $\xi:\Omega\to\mathbb{R}^d$ est un *vecteur aléatoire* défini sur un espace de probabilité $(\Omega,\mathcal{F},\mathbb{P})$, $x\in\mathbb{R}^n$ est un *vecteur déterministe* qui représente les variables de décisions, $\mathbb{E}\Big[F^0\Big(x,\xi(\cdot)\Big)\Big]$ est pour chaque $x\in\mathbb{R}^n$ fixé, l'espérance de la *variable aléatoire réelle* $\omega\mapsto F^0(x,\xi(\omega))$, et $\mathrm{ARGMIN}_{\mathbb{R}^r_+}\mathbb{E}\Big[F\Big(X,\xi(\cdot)\Big)\Big]$ est l'ensemble des solutions efficientes (au sens de Pareto) associées au

$$(POMOS) \qquad \mathrm{MIN}_{x\in X}\mathbb{E}\Big[F\Big(x,\xi(\cdot)\Big)\Big],$$

   où l'ensemble admissible X est un sous-ensemble de $\mathbb{R}^n$, et les objectifs sont donnés par

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

   où $F^i:\mathbb{R}^n\times\mathbb{R}^d\to\mathbb{R}$ est une fonction aléatoire réelle pour tout $i=1,\dots,r$.

   Une des missions de notre première contribution sera de proposer une suite de problèmes telle que toute suite de solutions optimales (resp. la suite des valeurs optimales) converge vers une solution optimale (resp. la valeur optimale) du problème (S).

Pour cela, nous définirons rigoureusement les fonctions suivantes en utilisant la Méthode de Monte-Carlo :

$$\begin{aligned}
&\hat{F}_N^0 : \mathbb{R}^n \times \tilde{\Xi} \to \mathbb{R} \\
&\qquad (x, \tilde{\xi}) \mapsto \hat{F}_N^0(x, \tilde{\xi}) := \frac{1}{N} \sum_{k=1}^{N} F^0(x, \xi_k), \\
&\hat{F}_N : \mathbb{R}^n \times \tilde{\Xi} \to \mathbb{R}^r \\
&\qquad (x, \tilde{\xi}) \mapsto \hat{F}_N(x, \tilde{\xi}) := \frac{1}{N} \sum_{k=1}^{N} \Big(F^1(x, \xi_k), \ldots, F^r(x, \xi_k) \Big),
\end{aligned}$$

où $\tilde{\Xi} \ni \tilde{\xi} = (\xi_1, \xi_2, \ldots)$ est un processus aléatoire dont chaque composante est distribuée indépendamment et identiquement au vecteur aléatoire ξ.

Pour chaque $N \geq 1$ et $\tilde{\xi} \in \tilde{\Xi}$ fixés, nous étudierons le Problème d'Optimisation post-Pareto Déterministe

$$(S_N(\tilde{\xi})) \qquad \min \left\{ \hat{F}_N^0(x, \tilde{\xi}) \middle| x \in \mathrm{ARGMIN}_{\mathbb{R}_+^r} \hat{F}_N(X, \tilde{\xi}) \right\},$$

où $\mathrm{ARGMIN}_{\mathbb{R}_+^r} \hat{F}_N(X, \tilde{\xi})$ est l'ensemble des solutions de Pareto du Problème d'Optimisation Multi-Objectifs

$$(SAA - NPOMO) \qquad \mathrm{MIN}_{x \in X} \hat{F}_N(x, \tilde{\xi})$$

obtenu par la Méthode de Monte-Carlo.

Dans le Chapitre 2 nous démontrerons que les problèmes déterministes $(S_N(\tilde{\xi}))$ permettent d'approcher le problème (S), au sens où toute valeur d'adhérence de toute suite de solutions optimales issue des problèmes $(S_N(\tilde{\xi}))$ est presque sûrement une solution optimale pour le problème (S). Ainsi, la suite des valeurs optimales des problèmes $(S_N(\tilde{\xi}))$ converge presque sûrement vers la valeur optimale du problème (S) quand N tend vers l'infini.

Nous aurons dans un premier temps à montrer que la distance de Hausdorff-Pompeiu entre les ensembles de Pareto

$$\mathrm{ARGMIN}_{\mathbb{R}_+^r} \mathbb{E}\Big[F\Big(X, \xi(\cdot) \Big) \Big] \text{ et } \mathrm{ARGMIN}_{\mathbb{R}_+^r} \hat{F}_N(X, \tilde{\xi})$$

tend vers zéro presque sûrement lorsque N tend vers l'infini. Autrement dit, nous montrerons que

(a) la déviation entre les ensembles de Pareto

$$\mathrm{ARGMIN}_{\mathbb{R}_+^r} \mathbb{E}\Big[F\Big(X, \xi(\cdot) \Big) \Big] \text{ et } \mathrm{ARGMIN}_{\mathbb{R}_+^r} \hat{F}_N(X, \tilde{\xi})$$

tend vers 0 presque sûrement quand N tend vers l'infini, et

(b) la déviation entre les ensembles de Pareto

$$\text{ARGMIN}_{\mathbb{R}^r_+} \hat{F}_N(X, \tilde{\xi}) \text{ et } \text{ARGMIN}_{\mathbb{R}^r_+} \mathbb{E}\Big[F\Big(X, \xi(\cdot)\Big)\Big]$$

tend vers 0 presque sûrement quand N tend vers l'infini.

Pour démontrer le premier point, nous utiliserons des résultats provenant de l'Analyse Multivoque, et nous supposerons que toutes les fonctions considérées sont continues sur l'ensemble admissible X supposé compact.

Dans le second point traité, nous supposerons de plus que le $(POMOS)$ est strictement convexe, c'est à dire que l'ensemble admissible X est un ensemble convexe, et que chaque composante de la fonction vectorielle $F = (F^1, F^2, \ldots, F^r)$ est strictement convexe. Les Théorèmes de scalarisation pour les Problèmes Vectoriels convexes nous permettront de conclure.

Notons que pour notre Problème d'Optimisation sur l'ensemble de Pareto, il est **nécessaire** que la distance de Hausdorff-Pompeiu entre les ensembles $\text{ARGMIN}_{\mathbb{R}^r_+} \mathbb{E}\Big[F\Big(X, \xi(\cdot)\Big)\Big]$ et $\text{ARGMIN}_{\mathbb{R}^r_+} \hat{F}_N(X, \tilde{\xi})$ tende vers zéro pour que toute solution efficiente (au sens de Pareto) du $(POMOS)$ soit une limite d'une suite de solutions efficientes du $(SAA - NPOMO)$, et pour que toute valeur d'adhérence de suite de solutions efficientes du $(SAA - NPOMO)$ soit une solution efficiente du $(POMOS)$.

Ainsi, nous pourrons conclure que toute valeur d'adhérence de toute suite de l'ensemble $\text{argmin}(\hat{F}^0_N(x, \tilde{\xi}) | x \in \text{ARGMIN}_{\mathbb{R}^r_+}(\hat{F}_N(X, \tilde{\xi})))$ appartient à l'ensemble $\text{argmin}(\mathbb{E}[F^0(X, \xi(\cdot))] | x \in \text{ARGMIN}_{\mathbb{R}^r_+}(\mathbb{E}[F(X, \xi(\cdot))]))$.

Enfin, nous prouverons que la suite des valeurs optimales des problèmes $(S_N(\tilde{\xi}))$ converge presque sûrement vers la valeur optimale du problème (S) lorsque N tend vers l'infini.

Dans la deuxième partie du Chapitre 2, nous traiterons le cas particulier où la fonction à optimiser sur l'ensemble de Pareto dépend des objectifs du $(POMOS)$, et ceci afin de lever l'hypothèse restrictive de stricte convexité. Autrement dit, nous considérerons le problème d'optimiser sur l'image (par la fonction vectorielle F) de l'ensemble de Pareto, souvent appelé dans la littérature front de Pareto.

Très important en pratique, ces Problèmes d'Optimisation post-Pareto dans l'espace des Objectifs rendent possible l'évaluation des bornes de chaque objectif sur l'ensemble de Pareto, et leur résolution ne nécessite pas toujours de générer l'ensemble du front de Pareto.

Nous nous intéresserons donc au Problème d'Optimisation post-Pareto Stochastique :

$$(O) \qquad \min\left\{ f\Big(\mathbb{E}\Big[F\Big(x,\xi(\cdot)\Big)\Big]\Big) \Big| x \in \mathrm{ARGMIN}_{\mathbb{R}^r_+}\mathbb{E}\Big[F\Big(X,\xi(\cdot)\Big)\Big]\right\},$$

où $f : \mathbb{R}^r \to \mathbb{R}$ est une fonction déterministe, et $\mathrm{ARGMIN}_{\mathbb{R}^r_+}\mathbb{E}\Big[F\Big(X,\xi(\cdot)\Big)\Big]$ est l'ensemble de Pareto du

$$(POMOS) \qquad \mathrm{MIN}_{x\in X}\mathbb{E}\Big[F\Big(x,\xi(\cdot)\Big)\Big]$$

défini précédemment.

Pour chaque $N \geq 1$ et $\tilde{\xi} \in \tilde{\Xi}$ fixés, nous étudierons le Problème d'Optimisation post-Pareto Déterministe

$$(O_N(\tilde{\xi})) \qquad \min\left\{ f\Big(\hat{F}_N\Big(X,\tilde{\xi}\Big)\Big) \Big| x \in \mathrm{ARGMIN}_{\mathbb{R}^r_+}\hat{F}_N(X,\tilde{\xi})\right\},$$

où $\mathrm{ARGMIN}_{\mathbb{R}^r_+}\hat{F}_N(X,\tilde{\xi})$ est l'ensemble de Pareto du Problème d'Optimisation Multi-Objectifs :

$$(SAA - NPOMO) \qquad \mathrm{MIN}_{x\in X}\hat{F}_N(x,\tilde{\xi}).$$

Pour ces Problèmes d'Optimisation dans l'espace des Objectifs, nous n'aurons recours à aucune hypothèse de convexité. Pour démontrer que la suite des valeurs optimales issues des problèmes $(O_N(\tilde{\xi}))$ converge presque sûrement vers la valeur optimale du problème (O), nous aurons naturellement à montrer que

(a) la déviation entre les fronts de Pareto

$$\mathrm{MIN}_{\mathbb{R}^r_+}\mathbb{E}\Big[F\Big(X,\xi(\cdot)\Big)\Big] \text{ et } \mathrm{MIN}_{\mathbb{R}^r_+}\hat{F}_N(X,\tilde{\xi})$$

tend vers 0 presque sûrement quand N tend vers l'infini, puis que

(b) la déviation entre les fronts de Pareto

$$\mathrm{MIN}_{\mathbb{R}^r_+}\hat{F}_N(X,\tilde{\xi}) \text{ et } \mathrm{MIN}_{\mathbb{R}^r_+}\mathbb{E}\Big[F\Big(X,\xi(\cdot)\Big)\Big]$$

tend vers 0 presque sûrement quand N tend vers l'infini.

Autrement dit, nous prouverons que la distance de Hausdorff-Pompeiu entre ces fronts de Pareto tend vers zéro presque sûrement lorsque N tend vers l'infini. Pour cela, nous nous servirons de la propriété dite de domination, ainsi que de certains résultats issus de l'Analyse Multivoque.

En résumé, nous proposerons une suite de problèmes déterministes $(O_N(\tilde{\xi}))$, et la résolution de l'un d'eux donnera une approximation de la valeur optimale du problème (O), dont la qualité dépendra du rang N.

Cependant, à ce stade de nos recherches, la vitesse de convergence de l'erreur d'approximation vers zéro nous est encore inconnue. Cette étude est un des objectifs de notre seconde contribution.

2. Dans le Chapitre 3, nous introduirons premièrement notre algorithme. Son principe est de retourner à chaque étape un intervalle contenant la valeur optimale d'un Problème d'Optimisation post-Pareto Déterministe, dans le cas particulier où le front de Pareto est un sous ensemble de $\mathbb{R}^2$.

   Nous considérerons donc le problème de minimiser une fonction réelle sur le front de Pareto associé à un Problème d'Optimisation bi-Objectifs (POBO) :

$$(DO) \qquad \min\left\{f(z)\middle|z \in \mathrm{MIN}_{\mathbb{R}^2_+} F(X)\right\},$$

   où $f : \mathbb{R}^2 \to \mathbb{R}$ est une fonction déterministe, et $\mathrm{MIN}_{\mathbb{R}^2_+} F(X)$ est le front de Pareto du

$$(POBO) \qquad \mathrm{MIN}_{x\in X} F(x) = \left(F^1(x), F^2(x)\right),$$

   où la fonction $F : \mathbb{R}^n \to \mathbb{R}^2$, et où l'ensemble admissible X est explicitement donné par

$$X = \left\{x \in \mathbb{R}^n \middle| g(x) = \left(g^1(x), g^2(x), \ldots, g^p(x)\right) \leq 0\right\},$$

   où la fonction $g : \mathbb{R}^n \to \mathbb{R}^p$.

   Le principe de l'algorithme est le suivant : déterminer à chaque étape $k \geq 1$ deux sous-ensembles compacts $S^{(k)}$ et $D^{(k)}$ qui vérifient

$$D^{(k-1)} \subset D^{(k)} \subset \mathrm{MIN}_{\mathbb{R}^2_+} F(X) \subset S^{(k)} \subset S^{(k-1)}.$$

   Ainsi, les suites réelles définies par

$$d^{(k)} := \min\left(f(z)\middle|z \in D^{(k)}\right)$$

   et

$$s^{(k)} := \min\left(f(z)\middle|z \in S^{(k)}\right)$$

sont telles que

$$s^{(k)} \leq \min\left\{f(z)\middle|z \in \mathrm{MIN}_{\mathbb{R}^2_+} F(X)\right\} \leq d^{(k)}$$

pour tout $k \geq 0$.

La résolution du problème revient donc à construire les suites d'ensembles $D^{(k)}$ et $S^{(k)}$ telles que $\lim_{k\to\infty} d^{(k)} - s^{(k)} = 0$. Par la suite, nous montrerons que notre algorithme termine toujours en un nombre fini d'étapes, quelle que soit la qualité d'approximation choisie par l'utilisateur.

Notre but principal étant de résoudre un Problème d'Optimisation post-Pareto Stochastique, nous adapterons par la suite notre algorithme au problème suivant :

$$(O) \qquad \min\left\{f(z)\middle|z \in \mathrm{MIN}_{\mathbb{R}^2_+} \mathbb{E}\Big[F\Big(X, \xi(\cdot)\Big)\Big]\right\},$$

où $f : \mathbb{R}^2 \to \mathbb{R}$ est une fonction déterministe, et $\mathrm{MIN}_{\mathbb{R}^2_+} \mathbb{E}\Big[F\Big(X, \xi(\cdot)\Big)\Big]$ est le front de Pareto du

$$(POBOS) \qquad \mathrm{MIN}_{x\in X} \mathbb{E}\Big[F\Big(x, \xi(\cdot)\Big)\Big] = \mathbb{E}\Big[F^1\Big(x, \xi(\cdot)\Big), F^2\Big(x, \xi(\cdot)\Big)\Big],$$

dont l'ensemble admissible X est un sous ensemble déterministe de $\mathbb{R}^n$ explicitement donné par

$$X = \left\{x \in \mathbb{R}^n \middle| g(x) = \Big(g^1(x), g^2(x), \ldots, g^p(x)\Big) \leq 0\right\},$$

où la fonction $g : \mathbb{R}^n \to \mathbb{R}^p$.

Nous utiliserons alors certains résultats de notre première contribution : nous considérerons les Problèmes d'Optimisation post-Pareto

$$(O_N(\tilde{\xi})) \qquad \min\left\{f(z)\middle|z \in \mathrm{MIN}_{\mathbb{R}^2_+} \hat{F}_N(X, \tilde{\xi})\right\},$$

où pour chaque N fixé, $\mathrm{MIN}_{\mathbb{R}^2_+} \hat{F}_N(X, \tilde{\xi})$ est le front de Pareto du Problème d'Optimisation Bi-Objectifs

$$(SAA - NPOBO) \qquad \mathrm{MIN}_{x\in X} \hat{F}_N(x, \tilde{\xi}) = \frac{1}{N}\sum_{i=1}^{N} \Big(F^1(x, \xi_i), F^2(x, \xi_i)\Big).$$

Nous montrerons alors que la vitesse de convergence de la suite des valeurs optimales des problèmes $(O_N(\tilde{\xi}))$ vers la valeur optimale du problème (O)

est exponentielle.

Ensuite, quelle que soit l'erreur d'approximation $\epsilon > 0$ et le niveau de confiance $p_0 \in]0,1[$ choisis par l'utilisateur, nous proposerons un rang $N^0 = N^0(\epsilon, p_0)$ qui dépendra explicitement des données, et pour lequel la valeur optimale du problème $(O_{N^0}(\tilde{\xi}))$ est une ϵ-approximation de la valeur optimale du problème (O) avec une probabilité supérieure à p_0.

Finalement, nous utiliserons notre algorithme déterministe pour déterminer un intervalle contenant la valeur optimale du problème $(O_{N^0}(\tilde{\xi}))$, puis nous proposerons un intervalle de confiance qui contient celle du problème (O).

Pour arriver à ces résultats, nous supposerons en plus des hypothèses classiques de l'Optimisation Stochastique que la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ est non-décroissante et continue au sens de Lipschitz. De plus, nous admettrons que pour tout $x \in X$ fixé, la variance des variables aléatoires $\omega \mapsto F^j\Big(x, \xi(\omega)\Big)$ est bornée $(j = 1, 2)$, et qu'il est possible d'en estimer un majorant indépendant de x.

Ce mémoire est organisé comme suit :

Que ce soit la caractérisation de la relation d'ordre (partielle) par un cône convexe ou les différents types de solutions efficientes d'un problème d'Optimisation Vectorielle, sans oublier les Théorèmes de scalarisation, le Chapitre 1 sera consacré aux notions fondamentales de l'Optimisation Vectorielle qui ont apportées les bases nécessaires à l'élaboration de cette thèse. Certaines d'entre elles seront illustrées.

Le Chapitre 2 constituera une base théorique pour ces recherches. Pour la première fois, le problème de minimiser l'espérance d'une fonction aléatoire réelle sur l'ensemble de Pareto associé à un problème de minimisation Stochastique Multi-Objectifs sera considéré. Ce problème sera explicité en détails, puis nous construirons méthodiquement une suite de problèmes déterministes via la méthode de Monte-Carlo. Un exemple numérique sera également présenté.

Dans notre seconde contribution (Chapitre 3), nous aborderons un Problème d'Optimisation post-Pareto Stochastique, dans le cas particulier où le front de Pareto est un sous-ensemble de $\mathbb{R}^2$, et où les deux Objectifs seront supposés convexes. Nous énoncerons un algorithme déterministe que nous adapterons pour suggérer un intervalle de confiance contenant la valeur optimale de notre problème post-Pareto Stochastique. Ces travaux seront illustrés par la résolution numérique de plusieurs exemples.

Pour conclure, ce mémoire aboutira sur des Perspectives de Recherche.

Notations

$\mathbb{N} := \{0, 1, 2, \ldots\}$ l'ensemble des nombres entiers positifs ou nuls.

$\mathbb{N}^* := \{1, 2, \ldots\}$ l'ensemble des nombres entiers strictement positifs.

$\mathbb{R}$ l'ensemble des nombres réels.

$\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$ l'ensemble des nombres réels positifs ou nuls.

$\mathbb{R}^r := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{r \text{ fois}},\ r > 1.$

$\mathbb{R}^r_+ := \left\{x = (x_1, x_2, \ldots, x_r) \in \mathbb{R}^r : x_i \geq 0,\ i = 1, 2, \ldots, r\right\}$ le cône de Pareto.

Pour toute partie A d'un espace topologique, son intérieur sera noté $\mathrm{int}(A)$, son adhérence $\bar{A}$, et sa frontière $\partial A := \bar{A} \setminus \mathrm{int}(A)$.

Chapitre 1

Notions d'Optimisation Vectorielle

1.1 Introduction

Remarquons que tout Problème d'Optimisation Mono-Objectif repose sur la relation d'ordre totale dans $\mathbb{R}$.
Ainsi, chaque solution optimale représente, parmi toutes les solutions admissibles, la (une) *meilleure* solution.
Le cas Vectoriel est beaucoup plus difficile à traiter car la *meilleure* solution n'existe généralement pas.

Un problème d'Optimisation Vectorielle (POV) est caractérisé par

- L'espace des décisions G, c'est à dire l'espace des variables décrivant le modèle.
- L'ensemble admissible $X \subset G$, formé par toutes les alternatives acceptables (décisions admissibles) pour le modèle considéré.
- L'espace des objectifs H, muni d'une relation d'ordre partiel $\preceq$, appelée relation de préférence. La notation $h^1 \preceq h^2$ sera utilisée pour indiquer que $h^1 \in H$ est préféré à $h^2 \in H$.
- Une application $F : G \to H$ représentant le critère de performance (pour chaque alternative admissible $x \in X$, sa performance est donnée par $F(x)$).

En changeant éventuellement le signe de l'application vectorielle F, il est toujours possible de se ramener à un problème de minimisation, noté

$$(POV) \qquad \mathrm{MIN}_{x \in X} F(x).$$

Dans le cas particulier où l'espace des Objectifs est donné par $H = \mathbb{R}^r$, et où l'ordre partiel $\preceq$ est défini naturellement, c'est à dire que pour tout $y, z \in \mathbb{R}^r$,

$$y \preceq z \Leftrightarrow y^i \leq z^i \ \forall i = 1, \ldots, r,$$

nous parlerons de Problème d'Optimisation Multi-Objectifs (POM).
Le critère de performance est alors de la forme $F = (F^1, F^2, \ldots F^r) : X \to \mathbb{R}^r$, et chacune de ses composantes F^i $(i = 1, \ldots, r)$ est un Objectif scalaire (à valeurs réelles). La performance de l'alternative $x \in X$ est donc préférée à celle de $x' \in X$ si, et seulement si,

$$F^i(x) \leq F^i(x'),\ \forall i \in \{1, \ldots, r\}.$$

La grande difficulté d'un Problème d'Optimisation Multi-Objectifs découle du fait qu'il n'existe généralement pas d'alternative $x^* \in X$ tel que

$$\forall x \in X,\ F(x^*) \preceq F(x).$$

En effet, une telle alternative appartiendrait à l'ensemble $\bigcap_{i=1}^{r} \operatorname{argmin}_{x \in X} F^i(x)$, qui est généralement vide. Pour cette raison, x^* est dite **alternative idéale ou utopique**.

Ainsi, au lieu de rechercher des alternatives d'existence incertaine, **Vilfredo Pareto** a proposé de rechercher les alternatives de meilleur compromis. Pour être plus précis, une alternative $x^* \in X$ est efficiente au sens de Pareto si, et seulement si, il n'existe pas d'autre alternative $x \in X$ tel que

$$F(x) \preceq F(x^*) \text{ et } F(x) \neq F(x^*).$$

Autrement dit, pour un Problème d'Optimisation Multi-Objectifs, une alternative efficiente au sens de Pareto réalise la *meilleure* performance parmi les alternatives admissibles dont les performances lui sont comparables.
Cependant, il n'existe généralement pas d'alternative utopique, dont la performance serait la *meilleure de toutes*.

1.2 Cône Convexe et Relation de Préférence

Supposons maintenant que H est un espace de Hilbert. La relation de préférence $\preceq$ est une relation binaire sur H, c'est à dire une partie (non vide) $\mathcal{R}$ du produit cartésien $H \times H$. Pour indiquer que l'élément $h_1 \in H$ est préféré à $h_2 \in H$, nous pouvons indifféremment écrire $h_1 \preceq h_2$ ou $(h_1, h_2) \in \mathcal{R}$.

Commençons par rappeler quelques définitions.

Définition 1.2.1 (Espace ordonné). *On dit que le couple $(H, \preceq)$ est un espace Hilbertien ordonné si, et seulement si, la relation $\preceq$ est une relation d'ordre partielle (réflexive, transitive et antisymétrique) compatible avec les structures algébrique et topologique, c'est à dire vérifiant*

- $\forall h_1, h_2, h_3 \in H,\ h_1 \preceq h_2 \Rightarrow h_1 + h_3 \preceq h_2 + h_3.$
- $\forall h_1, h_2 \in H, \forall \lambda \in \mathbb{R}_+,\ h_1 \preceq h_2 \Rightarrow \lambda h_1 \preceq \lambda h_2.$

– *pour toute suite* $(h_n)_{n\in\mathbb{N}}$ *de* H *telle que* $\lim h_n \to h$*, si* $h_n \preceq 0$ *pour tout* $n \in \mathbb{N}$*, alors* $h \preceq 0$.

Définition 1.2.2 (Cône et cône convexe)**.** *Soit* $C \subset H$ *un ensemble non vide. On dit que* C *est un cône si, et seulement si,* $\mathbb{R}_+ C \subset C$*. Si de plus* $tx + (1-t)y \in C$ *pour tout* $x, y \in C$*,* $t \in [0,1]$*, le cône* C *est dit convexe.*

Proposition 1.2.1. *Un cône* $C \subset H$ *est convexe si, et seulement si,*

$$C + C \subset C.$$

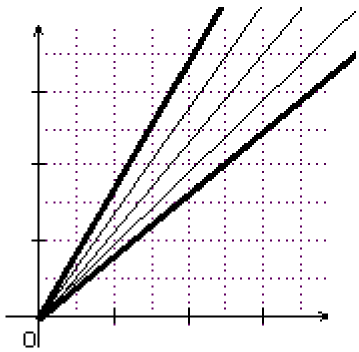


Figure 1.1 – Un exemple de cône convexe de $\mathbb{R}^2$

Définition 1.2.3 (Cône saillant)**.** *Un cône convexe* $C \subset H$ *est dit saillant si, et seulement si,* $C \cap -C = \{0\}$.

Le Théorème suivant caractérise toute relation d'ordre partielle via un cône convexe fermé saillant.

Théorème 1.2.1. *Soit* $\preceq$ *une relation d'ordre partielle sur l'espace de Hilbert* H.
L'ensemble

$$H_+ := \left\{ h \in H \middle| 0 \preceq h \right\}$$

est un cône convexe fermé saillant.

Réciproquement, soit $C \subset H$ *un cône convexe fermé saillant. Alors la relation binaire définie par*

$$\preceq := \left\{ (h_1, h_2) \in H \times H \middle| h_2 - h_1 \in C \right\}$$

est une relation d'ordre partielle sur H.
De plus, l'espace Hilbertien $(H, \preceq)$ *est ordonné et* $H_+ = C$.

Définition 1.2.4 (Cône de Pareto)**.** *Lorsque l'espace des Objectifs est donné par* $H = \mathbb{R}^r$*, le cône* $\mathbb{R}^r_+$ *est appelé cône de Pareto.*

1.3 Les différentes alternatives efficientes d'un Problème d'Optimisation Vectorielle

Dans cette partie, nous allons nous intéresser aux différents type de solutions du (POV)

$$\mathrm{MIN}_{x \in X} F(x),$$

où l'application $F : X \subset G \to (H, C)$, $C = H_+$ est le cône induit par la relation d'ordre partielle $\preceq$ de l'espace des Objectifs H (supposé être un espace de Hilbert), et l'ensemble admissible X est une partie non vide de l'espace des Décisions G (supposé être un espace vectoriel).

Rappelons que pour tout $h_1, h_2 \in H$, nous notons

$$h_1 \preceq h_2 \Leftrightarrow h_2 - h_1 \in C,$$

et si de plus $\mathrm{int}(C) \neq \emptyset$, nous notons

$$h_1 \prec h_2 \Leftrightarrow h_2 - h_1 \in \mathrm{int}(C).$$

Définition 1.3.1 (Alternatives efficientes)**.** *Une alternative $x^* \in X$ est dite **efficiente** pour (POV) si, et seulement si,*

$$(F(x^*) - C) \cap F(X) = \{F(x^*)\}.$$

Ce qui est équivalent à dire

$$\forall x \in X, F(x) \preceq F(x^*) \Rightarrow F(x) = F(x^*),$$

ou encore

$$\nexists x \in X : \ F(x) \preceq F(x^*) \ \text{et} \ F(x) \neq F(x^*).$$

L'ensemble des alternatives efficientes du (POV) est noté

$$ARGMIN_C F(X).$$

Considérons maintenant le cône convexe saillant

$$C_w := \mathrm{int}(C) \cup \{0\}.$$

Définition 1.3.2 (Alternatives faiblement efficientes)**.** *Une alternative $x^* \in X$ est dite **faiblement efficiente** pour (POV) si, et seulement si,*

$$x^* \in ARGMIN_{C_w} F(X).$$

D'une manière équivalente,

$$(F(x^*) - \mathrm{int}(C)) \cap F(X) = \emptyset,$$

ou encore

$$\nexists x \in X : \ F(x) \prec F(x^*).$$

L'ensemble des alternatives faiblement efficientes du (POV) *est noté*

$$w\text{-}ARGMIN_C F(X).$$

Définition 1.3.3 (Alternatives proprement efficientes, Henig[76]). *Une alternative* $x^* \in X$ *est dite* ***proprement efficiente*** *pour* (POV) *si, et seulement si, il existe un cône convexe* K $(K \neq H)$ *d'intérieur non vide tel que* $C \setminus \{0\} \subset \mathrm{int}(K)$ *et*

$$x^* \in ARGMIN_K F(X).$$

L'ensemble des alternatives proprement efficientes du (POV) *est noté*

$$p\text{-}ARGMIN_C F(X).$$

Lorsque $(H, C) = (\mathbb{R}^r, \mathbb{R}^r_+)$ et $F(X) + \mathbb{R}^r_+$ est un ensemble convexe, une définition équivalente est donnée par la

Définition 1.3.4 (Geoffrion,[71]). *Une alternative* $x^* \in X$ *est dite* ***proprement efficiente*** *pour* (POM) *si, et seulement si, elle est efficiente et il existe un réel* $M > 0$ *tel que pour tout* $i \in \{1, \ldots, r\}$ *et* $x \in X$ *satisfaisants* $F^i(x) < F^i(x^*)$, *il existe* $j \in \{1, \ldots, r\}$ *telle que* $F^j(x^*) < F^j(x)$ *et*

$$\frac{F^i(x^*) - F^i(x)}{F^j(x) - F^j(x^*)} \leq M$$

Remarque 1.3.1. *Dans le cas d'un* (POM), *les alternatives efficientes (resp. faiblement efficientes, resps. proprement efficientes) sont dites efficientes au sens de Pareto, où plus simplement de Pareto (resp. de Pareto faible, resp. de Pareto propre), et l'ensemble* $ARGMIN_{\mathbb{R}^r_+} F(X)$ *(resp.* $w\text{-}ARGMIN_{\mathbb{R}^r_+} F(X)$, *resp.* $p\text{-}ARGMIN_{\mathbb{R}^r_+} F(X)$*) est appelé* ***ensemble de Pareto*** *(resp. ensemble de Pareto faible, resp. ensemble de Pareto propre).*

L'ensemble $MIN_{\mathbb{R}^r_+}\Big(F(X)\Big) = F\Big(ARGMIN_{\mathbb{R}^r_+} F(X)\Big)$ *(resp.* $w\text{-}MIN_{\mathbb{R}^r_+}\Big(F(X)\Big) = F\Big(w\text{-}ARGMIN_{\mathbb{R}^r_+} F(X)\Big)$, *resp.* $p\text{-}MIN_{\mathbb{R}^r_+}\Big(F(X)\Big) = F\Big(p\text{-}ARGMIN_{\mathbb{R}^r_+} F(X)\Big)$*) est alors appelé* ***front de Pareto*** *(resp. front de Pareto faible, resp. front de Pareto propre).*

Les images des ensembles efficients sont telles que

Proposition 1.3.1. *[97, Proposition 2.2 p.41]*

$$p\text{-}MIN_C\Big(F(X)\Big) \subset MIN_C\Big(F(X)\Big) \subset w\text{-}MIN_C\Big(F(X)\Big).$$

Remarque 1.3.2. *Notons que lorsque* $H = \mathbb{R}$ *(c'est à dire la performance* F *est un objectif scalaire), si l'ensemble* $\operatorname*{argmin}_{x \in X} F(x) := \{x^* \in X | F(x^*) = \min_{x \in X} F(x)\}$ *est non vide, alors*

$$p\text{-}ARGMIN_C F(X) = ARGMIN_C F(X) = w\text{-}ARGMIN_C F(X) = \operatorname*{argmin}_{x \in X} F(x).$$

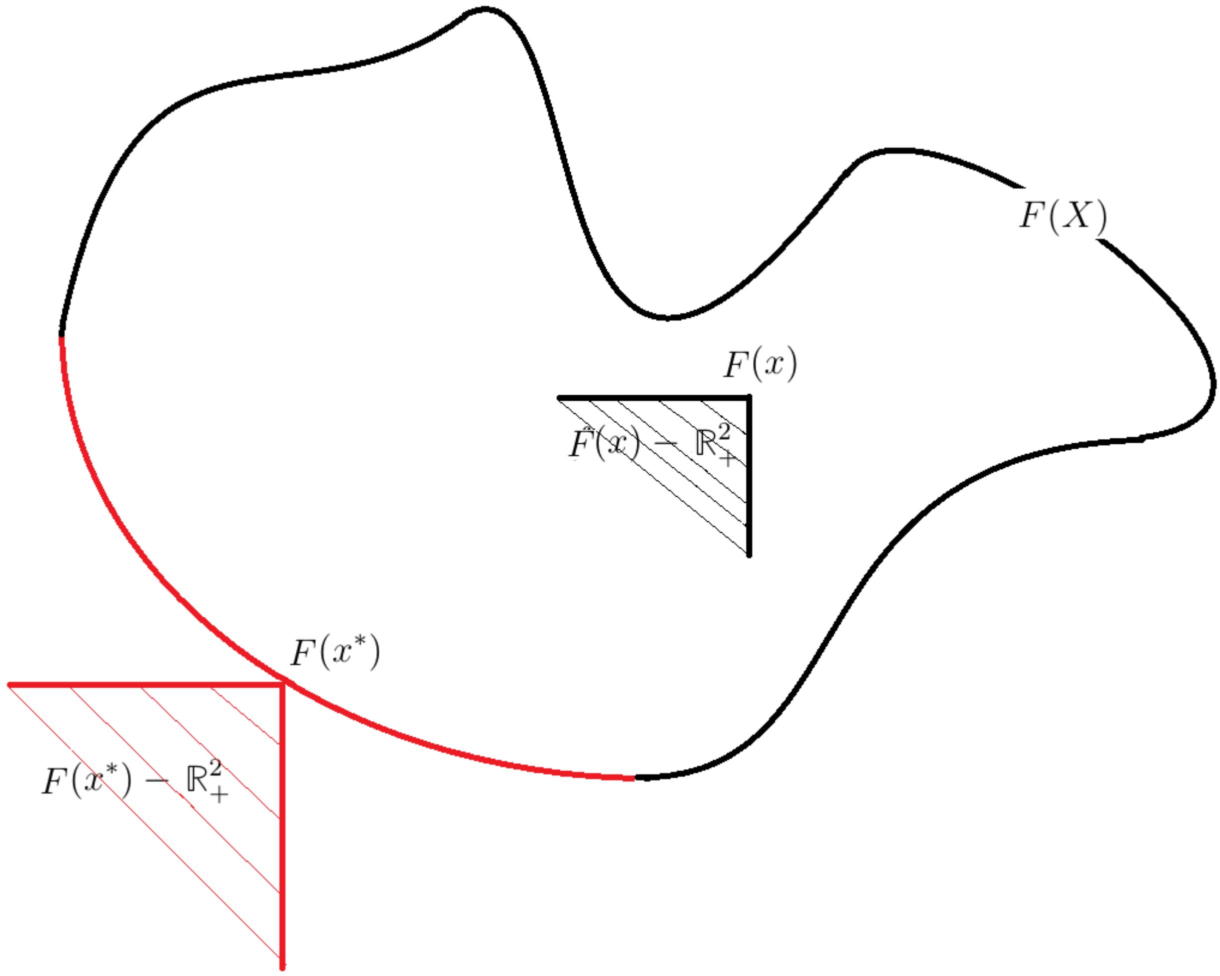


FIGURE 1.2 – En rouge, le front de Pareto $\mathrm{MIN}_{\mathbb{R}^2_+}\Big(F(X)\Big)$ de l'ensemble $F(X) \subset \mathbb{R}^2$

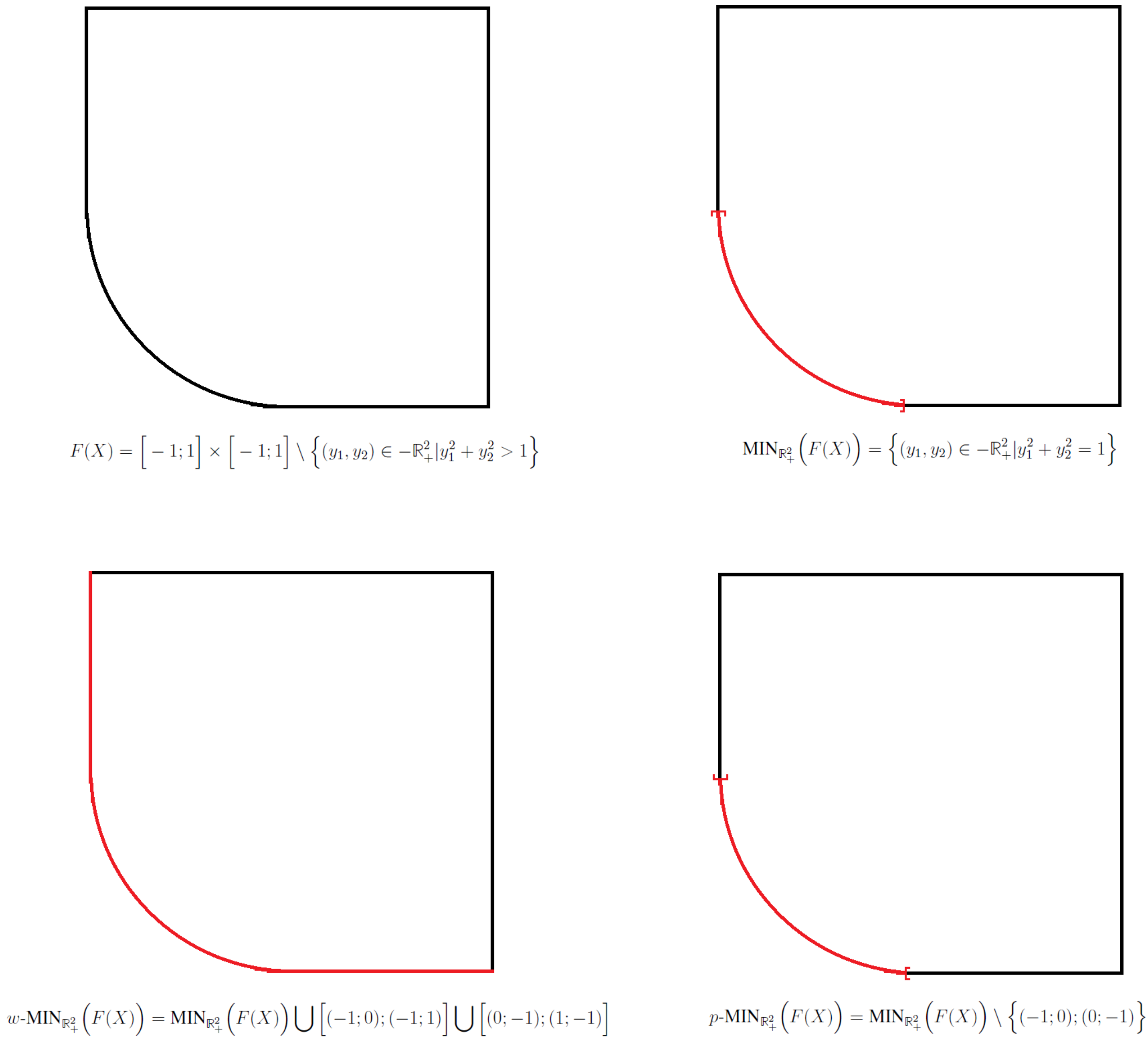


Figure 1.3 – Illustration des différents fronts de Pareto (en rouge)

1.4 Une Approche par Scalarisation

Le principe de la scalarisation consiste à remplacer un Problème d'Optimisation Vectorielle par une famille de Problèmes Scalaires.

Notons H' le dual topologique de l'espace de Hilbert H (i.e. l'ensemble des formes linéaires et continues sur H) , et $\langle \cdot, \cdot \rangle$ le produit scalaire de dualité entre H et H'.
Rappelons que $C = H_+$ est le cône induit par la relation d'ordre partielle sur H.

Définition 1.4.1 (Cônes polaires)**.** *L'ensemble*

$$C^* := \left\{ \lambda \in H' \middle| \langle \lambda, c \rangle \geq 0 \ \forall c \in C \right\}$$

est appelé ***cône polaire*** *du cône C, et l'ensemble*

$$C_s^* := \left\{ \lambda \in H' \middle| \langle \lambda, c \rangle > 0 \ \forall c \in C \setminus \{0\} \right\}$$

est le ***pseudo-intérieur*** *de C^*.*

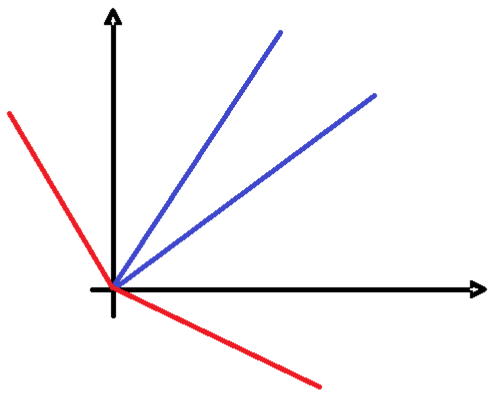


FIGURE 1.4 – En bleu un cône de $\mathbb{R}^2$, et en rouge son cône polaire

Remarque 1.4.1. *C^* et $C_s^* \cup \{0\}$ sont des cônes convexes, et C^* est un ensemble fermé.*

La définition suivante généralise l'idée de fonction convexe dans le cas vectoriel.

Définition 1.4.2 ((POV) convexe)**.** *La fonction vectorielle $F : X \subset G \to (H, C)$ est dite C**-convexe** si, et seulement si, X est un sous-ensemble convexe de l'espace vectoriel G et $\forall x_1, x_2 \in X,\ \ \forall t \in [0, 1]$,*

$$F(tx_1 + (1-t)x_2) \preceq tF(x_1) + (1-t)F(x_2).$$

Dans ce cas, le (POV) $MIN_{x \in X} F(x)$ sera dit convexe.

Une caractérisation des fonctions C-convexe est donnée par la

Proposition 1.4.1. *La fonction $F : X \to (H, C)$ est C-convexe si, et seulement si, la fonction scalaire $X \ni x \mapsto \langle \lambda, F(x) \rangle$ est convexe pour tout $\lambda \in C^*$.*

Remarque 1.4.2. *Dans le cas particulier où $H = \mathbb{R}^r$, la fonction $F = (F^1, F^2, \dots, F^r)$ est $\mathbb{R}^r_+$-convexe si, et seulement si, chaque composante scalaire F^i $(i = 1, \dots, r)$ est une fonction convexe.*

Il est important de noter que même si l'application F est C-convexe, l'ensemble $F(X)$ peut ne pas être convexe. En revanche,

Proposition 1.4.2. *[114, Theorem 3.7 p.26] Soit $F : X \to (\mathbb{R}^r, \mathbb{R}^r_+)$ une fonction $\mathbb{R}^r_+$-convexe. Alors l'ensemble*

$$F(X) + \mathbb{R}^r_+$$

est convexe.

Théorème 1.4.1. *[56, Proposition 2.1 p.22] Les ensembles $F(X)$ et $F(X) + \mathbb{R}^r_+$ admettent le même front de Pareto. Autrement dit,*

$$MIN_{\mathbb{R}^r_+}\Big(F(X)\Big) = MIN_{\mathbb{R}^r_+}\Big(F(X) + \mathbb{R}^r_+\Big).$$

Théorème 1.4.2. *Le front de Pareto faible est donné par*

$$w\text{-}MIN_{\mathbb{R}^r_+}\Big(F(X)\Big) = \partial\Big(F(X) + \mathbb{R}^r_+\Big) \cap F(X).$$

Les Théorèmes suivants, dit de Scalarisation, découlent du Théorème de Hahn-Banach (forme géométrique).
Théoriquement très populaires, ils permettent de caractériser les ensembles faiblement efficients et proprement efficients.

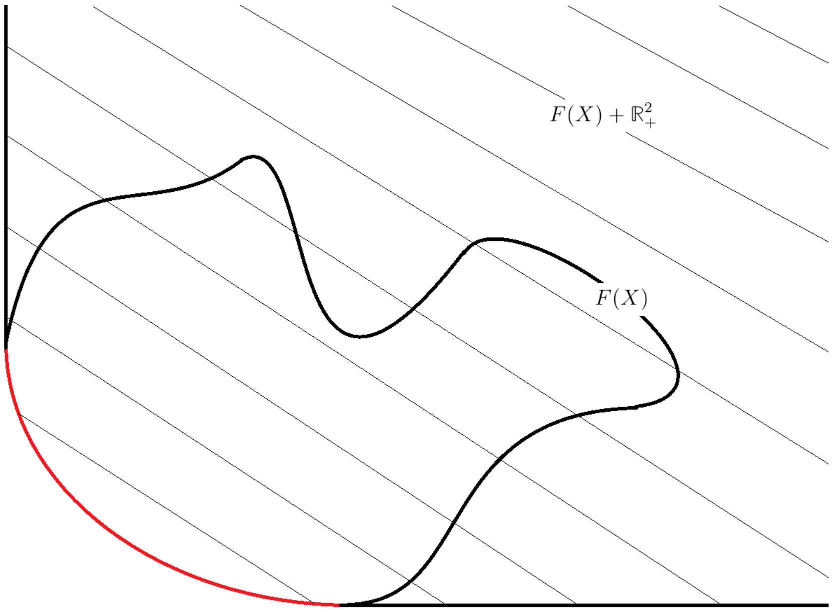


FIGURE 1.5 – Illustration de la Proposition 1.4.2 et des Théorèmes 1.4.1,1.4.2 dans un cas particulier où $\text{MIN}_{\mathbb{R}^2_+}\Big(F(X)\Big) = \text{w-MIN}_{\mathbb{R}^2_+}\Big(F(X)\Big)$ (en rouge)

Théorème 1.4.3. *Pour le Problème d'Optimisation Vectorielle $MIN_{x\in X}F(x)$,*

$$\bigcup_{\lambda\in C^*_s} \underset{x\in X}{\operatorname{argmin}}\,\langle\lambda, F(x)\rangle \subset p\text{-}ARGMIN_C F(X),$$

et

$$\bigcup_{\lambda\in C^*\setminus\{0\}} \underset{x\in X}{\operatorname{argmin}}\,\langle\lambda, F(x)\rangle \subset w\text{-}ARGMIN_C F(X).$$

Dans le cas d'un (POV) convexe, nous nous référons au Théorème suivant :

Théorème 1.4.4.

$$\bigcup_{\lambda\in C^*_s} \underset{x\in X}{\operatorname{argmin}}\,\langle\lambda, F(x)\rangle = p\text{-}ARGMIN_C F(X),$$

et

$$\bigcup_{\lambda\in C^*\setminus\{0\}} \underset{x\in X}{\operatorname{argmin}}\,\langle\lambda, F(x)\rangle = w\text{-}ARGMIN_C F(X).$$

Remarque 1.4.3. *Pour un Problème d'Optimisation Multi-Objectifs (POM), c'est à dire lorsque $(H, C) = (\mathbb{R}^r, \mathbb{R}^r_+)$, en identifiant l'espace Euclidien $\mathbb{R}^r$ à son dual topologique, nous obtenons*

$$C^* = \mathbb{R}^r_+$$

et

$$C_s^* = \text{int}(\mathbb{R}^r_+).$$

Le Théorème de Scalarisation devient alors plus applicable :

$$\bigcup_{\lambda \in \text{int}(\mathbb{R}^r_+)} \underset{x \in X}{\text{argmin}} \left(\sum_{i=1}^{r} \lambda^i F^i(x) \right) \subset p\text{-}ARGMIN_{\mathbb{R}^r_+} F(X),$$

et

$$\bigcup_{\lambda \in \mathbb{R}^r_+ \setminus \{0\}} \underset{x \in X}{\text{argmin}} \left(\sum_{i=1}^{r} \lambda^i F^i(x) \right) \subset w\text{-}ARGMIN_{\mathbb{R}^r_+} F(X),$$

avec égalité dans le cas d'un (POM) *convexe.*

Remarque 1.4.4. *Il est important de noter que pour tout* $\lambda \in \mathbb{R}^r_+ \setminus \{0\}$*, les Problèmes d'Optimisation Scalaires* $\min_{x \in X} \langle \lambda, F(x) \rangle$ *et* $\min_{x \in X} \left\langle \frac{\lambda}{||\lambda||}, F(x) \right\rangle$ *sont équivalents, au sens où ils admettent les mêmes solutions optimales. Ainsi,*

$$\bigcup_{\lambda \in \text{int}(\mathbb{R}^r_+), ||\lambda||=1} \underset{x \in X}{\text{argmin}} \left(\sum_{i=1}^{r} \lambda^i F^i(x) \right) \subset p\text{-}ARGMIN_{\mathbb{R}^r_+} F(X),$$

et

$$\bigcup_{\lambda \in \mathbb{R}^r_+, ||\lambda||=1} \underset{x \in X}{\text{argmin}} \left(\sum_{i=1}^{r} \lambda^i F^i(x) \right) \subset w\text{-}ARGMIN_{\mathbb{R}^r_+} F(X),$$

avec égalité dans le cas d'un (POM) *convexe.*

1.5 Section Compacte et Propriété de Domination

Dans cette sous-section, Y dénote un sous-ensemble non-vide de l'espace Hilbertien ordonné (H, C).

Définition 1.5.1 (Section Compacte)**.** *L'ensemble* Y *contient une section compacte si, et seulement si, il existe* $\bar{y} \in Y$ *tel que la section* $(\bar{y} - C) \cap Y$ *soit un ensemble compact.*

Le Théorème suivant donne une condition suffisante pour que l'ensemble efficient d'un (POV) soit non vide.

Théorème 1.5.1. *[56][Theorem 2.7 p.24] Pour* (POV)*, si l'ensemble* $F(X)$ *contient une section compacte, alors*

$$ARGMIN_C F(X) \neq \emptyset.$$

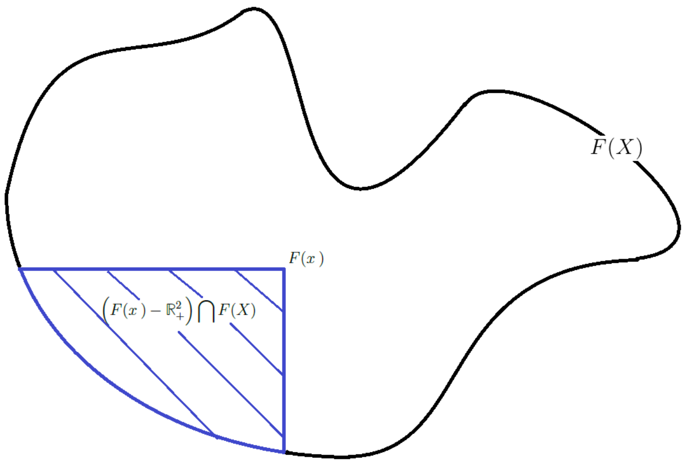


FIGURE 1.6 – En bleu, une section compacte avec $Y = F(X) \subset \mathbb{R}^2$ et $\bar{y} = F(x)$ pour un $x \in X$

Un point très important en Optimisation Vectorielle est l'identification d'alternatives dont les performances seraient meilleures que celle d'une alternative donnée. Cette propriété, dite de domination, est populaire grâce à son utilité numérique pour générer l'ensemble efficient.

Définition 1.5.2 (Propriété de domination)**.** *L'ensemble Y a la propriété de domination si, et seulement si, pour tout $y \in Y$, il existe $y^* \in MIN_C(Y)$ tel que $y - y^* \in C$.*

Proposition 1.5.1. *[97][Proposition 4.2 p.54] L'ensemble Y a la propriété de domination si, et seulement si,*

$$Y \subset MIN_C(Y) + C.$$

Définition 1.5.3 (Propriété de domination faible)**.** *L'ensemble Y a la propriété de domination faible si, et seulement si, pour tout $y \in Y$, soit $y \in w\text{-}MIN_C(Y)$, soit il existe $y^* \in w\text{-}MIN_C(Y)$ tel que $y - y^* \in \operatorname{int}(C)$.*

Proposition 1.5.2. *[97][Proposition 4.10 p.56] Tout ensemble compact a la propriété de domination faible.*

1.6 Conclusion

Très souvent infini et non-convexe, parfois non borné, l'ensemble efficient d'un Problème d'Optimisation Vectorielle est constitué de toutes les alternatives admissibles qui assurent un certain équilibre entre les différents Objectifs.

Incomparables entre elles, les performances efficientes d'un (POV) sont techniquement équivalentes, et dans de nombreuses situations pratiques, il est nécessaire de choisir une solution. Mais **il n'existe pas de solution unique qui s'impose d'elle même**.

Quelle méthode utiliser pour préférer une alternative efficiente à une autre ? C'est la question que se pose celui ou celle à qui revient la décision finale.

Johan Philip [102] fut le premier en 1972 à optimiser un nouvel Objectif (scalaire) sur l'ensemble efficient. Depuis lors ces Problèmes d'Optimisation post-Pareto ont faits l'objet de nombreuses études (voir [1, 7, 16, 17, 18, 13, 25, 42, 46, 47, 70, 77, 78] et [113] pour une bibliographie plus significative).

Ceci mène à la recherche d'une solution optimale pour ce nouvel Objectif, tel un outil pour mesurer les préférences du décideur et distinguer parmi toutes ces solutions efficientes, la plus juste ou la plus souhaitable.

Chapitre 2

Première Contribution

Stochastic Optimization over a Pareto Set Associated with a Stochastic Multi-Objective Optimization Problem

Henri Bonnel & Julien Collonge

Journal of Optimization Theory and Applications

ISSN 0022-3239
Volume 162
Number 2

J Optim Theory Appl (2014) 162:405-427
DOI 10.1007/s10957-013-0367-8

Springer

J Optim Theory Appl (2014) 162:405–427
DOI 10.1007/s10957-013-0367-8

Stochastic Optimization over a Pareto Set Associated with a Stochastic Multi-Objective Optimization Problem

Henri Bonnel · Julien Collonge

Received: 24 October 2012 / Accepted: 29 June 2013 / Published online: 18 July 2013

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, 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 in order to approach this problem.

We prove that the Hausdorff–Pompeiu distance between the weakly Pareto sets associated with the Sample Average Approximation problem and the true weakly Pareto set converges to zero almost surely as the sample size goes to infinity, assuming that our Stochastic Multi-objective Optimization Problem is strictly convex. Then we show that every cluster point of any sequence of optimal solutions of the Sample Average Approximation problems is almost surely a true optimal solution.

To handle also the non-convex case, we assume that the real objective to be minimized over the Pareto set depends on the expectations of the objectives of the Stochastic Optimization Problem, i.e. we optimize over the image space of the Stochastic Optimization Problem. Then, without any convexity hypothesis, we obtain the same type of results for the Pareto sets in the image spaces. Thus we show that the sequence of optimal values of the Sample Average Approximation problems converges almost surely to the true optimal value as the sample size goes to infinity.

Keywords Optimization over a Pareto set · Optimization over the Pareto image set · Multi-objective stochastic optimization · Multi-objective convex optimization · Sample average approximation method

H. Bonnel ( ) · J. Collonge
ERIM, University of New Caledonia, B.P. R4, 98851 Noumea Cedex, New Caledonia
e-mail: henri.bonnel@univ-nc.nc

J. Collonge
e-mail: julien.collonge@univ-nc.nc

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, 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 over the Pareto set associated with MOP. For instance, production planning (see, e.g., [1]) and portfolio management (see, e.g., [2]) 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 embarrassingly large Pareto set. Furthermore, for numerical computation, solving this problem one may avoid generate all the Pareto set (see [3, 4]), 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 words, we optimize over the image space (sometimes called 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 [5], and continued by many authors [1, 3, 4, 6–15] (see [16] for an extensive bibliography).

Some generalization to *semivectorial bilevel optimization problems* has been presented in [17–23].

The particular problem of optimizing a scalar function over the image space of a MOP has been studied in [24–28].

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. However, 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 with probability one as the sample size increase.

Some results in SMOP using SAA-N method have been recently obtained by Fliege and Xu [29] using a smoothing infinity norm scalarization approach to solve the SAA-N problems. Roughly speaking, the paper [29] 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, which is the main concern of our research.

Our paper is organized as follows.

In Sect. 2, we introduce the problem under consideration. We consider two instances of the same problem. First, we consider the problem of optimizing the expectation of a real function over the Pareto set in the decision space. Second, 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 image space.

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

In Sect. 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 goes to infinity. In order to show that the deviation in the other direction tends to zero, we need to assume that 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 almost surely 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 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 non-convex case, we need to work in the **Image 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 Pareto set image. Using again some results from Set Valued Analysis and Stability, we show that the SAA-N images of Pareto sets tend 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 Sect. 6, we present an illustrative example with a SMOP given by a Bi-objective Stochastic Optimization Problem in order to show graphically the convergence of the Pareto set images associated with the SAA-N problem to the true Pareto set image, as well as the convergence of the related optimal values to the true optimal value. We use MATLAB7 for randomly generate samples of size N and to graphically represent the different sets.

Concluding remarks are given in Sect. 7.

2 Problem Statement

We need to recall that for a (MOP) $\min_{x\in S} g(x)$, where $g=(g^1,g^2,\dots,g^r)$ is defined from S into $\mathbb{R}^r$ (S is an arbitrary nonempty set), a point $x^*\in S$ is said to be

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

Let us briefly introduce the two problems under consideration. The first one, denoted by (D), will be studied in the **Decision space** (see Sect. 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

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

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*, $\mathbb{E}[F^0(x,\xi(\cdot))]$ 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

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

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

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

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

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

The second problem (O), will be studied in the **Image space** (see Sect. 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

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

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 SMOP defined above.

However, 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 useful 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 σ-algebra. A mapping $\xi : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}_d)$ is said to be measurable with respect to $\mathcal{F}$ and $\mathcal{B}_d$ iff 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 and only 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 $\mathbb{E}_\xi[g(x, \cdot)] = \int_\Xi g(x, \eta) d\mathbb{P}_\xi(\eta)$.

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 $(\xi_k)_{k\geq 1}$ 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 $\tilde{\Xi} = \prod_{N=1}^{\infty} \Xi$ and let $\tilde{\mathcal{B}} = \bigotimes_{N=1}^{\infty} \mathcal{B}_\Xi$ denote the smallest σ-algebra on $\tilde{\Xi}$ generated by all sets of the form $B_1 \times B_2 \times \cdots \times B_N \times \Xi \times \Xi \times \cdots,\ B_k \in \mathcal{B}_\Xi$, $k = 1, \ldots, N,\ N = 1, 2, \ldots$.

The next theorem is from General Measure Theory, and can be considered as a nontrivial extension of Fubini's theorem [30, Theorem 10.4]:

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

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

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

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 an N-Sample Average Approximation (SAA-N) function.

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

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:

$$\text{(SMOP)} \quad \min_{x \in S} \mathbb{E}_\xi\big[F(x, \cdot)\big].$$

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

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

Let us reformulate the true stochastic problems under consideration.

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

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).

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

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 $y, z \in \mathbb{R}^r$, $y \le z$ means $y_j \le z_j$ 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 and 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.

$$\mathbb{E}_\xi\big[K^j(\cdot)\big] < \infty,$$
$$\big|F^j(x, \xi)\big| \le K^j(\xi) \quad \text{for all } x \in S \text{ and for a.e. } \xi \in \Xi.$$

(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:

$$\hat{F}_N^0 : \mathbb{R}^n \times (\tilde{\Xi}, \tilde{\mathcal{B}}, \tilde{\mathbb{P}}_\xi) \to \mathbb{R}$$
$$(x, \tilde{\xi}) \mapsto \hat{F}_N^0(x, \tilde{\xi}) := \frac{1}{N}\sum_{k=1}^N F^0(x, \xi_k), \tag{2}$$
$$\hat{F}_N : \mathbb{R}^n \times (\tilde{\Xi}, \tilde{\mathcal{B}}, \tilde{\mathbb{P}}_\xi) \to \mathbb{R}^r$$
$$(x, \tilde{\xi}) \mapsto \hat{F}_N(x, \tilde{\xi}) := \frac{1}{N}\sum_{k=1}^N \big(F^1(x, \xi_k), \dots, F^r(x, \xi_k)\big), \tag{3}$$

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 $A = \bigcup_{j=1}^r A^j$, we get $\forall \xi \in \Xi \setminus A,\ x \mapsto F(x, \xi)$ is continuous on S, and $\mathbb{P}_\xi(A) = 0$.

Letting $\tilde{A} = \bigcup_{N \in \mathbb{N}^*} \underbrace{A \times \cdots \times A}_{N} \times \Xi \times \Xi \cdots$, 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 $\hat{F}_N^0$ setting $\tilde{A}_0 = \bigcup_{N \in \mathbb{N}^*} \underbrace{A^0 \times \cdots \times A^0}_{N} \times \Xi \times \Xi \cdots$.

Definition 3.4 For each $N \in \mathbb{N}^*$ and $\tilde{\xi} \in \tilde{\Xi}$, we consider the following scalar deterministic optimization problems:

$$\big(\mathrm{D}_N(\tilde{\xi})\big) \qquad \min_{x \in E_N^w(\tilde{\xi})} \hat{F}_N^0(x, \tilde{\xi}),$$
$$\big(\mathrm{O}_N(\tilde{\xi})\big) \qquad \min_{x \in E_N(\tilde{\xi})} f\big(\hat{F}_N(x, \tilde{\xi})\big),$$

where $E^w_N(\tilde{\xi})$ (resp. $E_N(\tilde{\xi})$) is the weakly Pareto (resp. Pareto) set associated with the following SAA-N Multi-objective Optimization Problem:

$$\text{(SAA-N MOP)} \quad \min_{x \in S} \hat{F}_N(x, \tilde{\xi}).$$

The scalar SAA-N function $\hat{F}^0_N$ 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 ($\mathrm{D}_N(\tilde{\xi})$) (resp. ($\mathrm{O}_N(\tilde{\xi})$)) the **SAA-N problem** ($N \in \mathbb{N}^*, \tilde{\xi} \in \tilde{\Xi}$). Under some reasonable assumptions, we will show that the solutions and/or optimal values of SAA-N problems for sufficiently large N are approximations of the solutions and/or optimal values of the true problem (D) (resp. (O)).

By the Uniform Law of Large Number (ULLN) [31, Theorem 7.48], under ($\mathrm{H}_0, \mathrm{H}_1, \mathrm{H}_2, \mathrm{H}_3$), we obtain immediately the two following results.

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

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

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*

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

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 reason, there exists $\tilde{B}_0 \subset \tilde{\Xi}$ with $\tilde{\mathbb{P}}_\xi(\tilde{B}_0) = 0$ such that for each fixed $\tilde{\xi} \in \tilde{\Xi} \setminus \tilde{B}_0$, $\hat{F}^0_N(\cdot, \tilde{\xi})$ 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 and bounded sets.

- 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.
- We denote $\mathbb{D}(A, B) := \sup_{x \in A} d(x, B)$ the deviation of the set A from the set B.
- 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 pseudo-metric. If we consider the set of all nonempty and compact subsets of $\mathbb{R}^n$, $\mathbb{H}$ becomes a metric. Furthermore, for two nonempty and 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* $(\mathrm{H}_0, \mathrm{H}_1, \mathrm{H}_2, \mathrm{H}_3)$, *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.*

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

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 $\tilde{\mathbb{P}}_\xi(\{\tilde{\xi} \in \tilde{\Xi} \mid \lim_{N\to\infty} \mathbb{D}(\hat{F}_N(S, \tilde{\xi}), \mathbb{E}_\xi[F(S, \cdot)]) = 0\}) = 1$. 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 $\tilde{\xi} \in \tilde{D}$, $\mathbb{D}(\hat{F}_N(S, \tilde{\xi}), \mathbb{E}_\xi[F(S, \cdot)]) \nrightarrow 0$. Obviously $\tilde{\mathbb{P}}_\xi(\tilde{D} \setminus (\tilde{A} \cup \tilde{B})) > 0$, hence $\tilde{D} \setminus (\tilde{A} \cup \tilde{B}) \neq \emptyset$. Let then $\tilde{\xi} \in \tilde{D} \setminus (\tilde{A} \cup \tilde{B})$ be fixed.

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

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

Now we prove $\tilde{\mathbb{P}}_\xi(\{\tilde{\xi} \in \tilde{\Xi} \mid \lim_{N\to\infty} \mathbb{D}(\mathbb{E}_\xi[F(S, \cdot)], \hat{F}_N(S, \tilde{\xi})) = 0\}) = 1$.

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 of functions $(d(\cdot, \hat{F}_N(S, \tilde{\xi})))_{N\geq 1}$ 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 [32] 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 [33–38] for details). Let X be a separated topological space, and Y be a normed vector 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

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

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

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

Remark 3.6 [36, 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 $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, if $\Gamma : X \rightrightarrows Y$ is the set-valued mapping defined by $\Gamma(N) := A_N$ ($N \in \mathbb{N}^*$) 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 $x_0 \in X$ iff 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 for all $x \in U$,
$$\Gamma(x) \subset D.$$
- Γ is lower continuous (l.c.) at $x_0 \in X$ iff 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 for all $x \in U$,
$$\Gamma(x) \cap D \neq \emptyset.$$
- Γ is continuous at $x_0 \in X$ iff Γ is u.c. and l.c. at x_0.
- Γ is continuous iff Γ is continuous at every $x \in X$.

Proposition 3.2 [36, p. 55] *Let $\Gamma : X \rightrightarrows Y$ and let $x_0 \in X$. Γ is l.c. at x_0 if and only if $\Gamma(x_0) \subset \liminf_{x \to x_0} \Gamma(x)$.*

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

- Γ is Hausdorff continuous at $x_0 \in X$ iff Γ is H-u.c. and H-l.c. at x_0.
- Γ is H-continuous iff Γ is H-continuous at every $x \in X$.

Remark 3.7 [36, 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 $\lim_{x\to x_0} \mathbb{D}(\Gamma(x_0), \Gamma(x)) = 0$.

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

- closed valued iff for each $x \in X$, $\Gamma(x)$ is a closed set in Y;
- closed iff $\text{Graph}(\Gamma) := \{(x, y) | x \in X, y \in \Gamma(X)\}$ is closed;
- compact at $x \in X$ iff 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 $y \in \Gamma(x)$ (see [36]).

Note that if Γ is compact at x then the set $\Gamma(x)$ is compact. The converse is not true.

Now we recall some useful 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 that, if the set $\Gamma(x_0)$ is compact, then Γ is continuous at x_0 if and only if Γ is H-continuous at x_0.

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

$$\Gamma_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^r, \qquad \Gamma_{\tilde{\xi}}(N) := \begin{cases} \hat{F}_N(S, \tilde{\xi}) & \text{if } N \in \mathbb{N}^* \\ \mathbb{E}_\xi\big[F(S, \cdot)\big] & \text{if } N = +\infty, \end{cases} \tag{4}$$

where $\hat{F}_N(\cdot, \tilde{\xi})$ has been introduced in (3).

$$\Lambda_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^r, \qquad \Lambda_{\tilde{\xi}}(N) := \begin{cases} \hat{F}_N\big(E_N(\tilde{\xi}), \tilde{\xi}\big) & \text{if } N \in \mathbb{N}^* \\ \mathbb{E}_\xi\big[F(E, \cdot)\big] & \text{if } N = +\infty, \end{cases} \tag{5}$$

$$\Upsilon_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^n, \qquad \Upsilon_{\tilde{\xi}}(N) := \begin{cases} E_N^w(\tilde{\xi}) & \text{if } N \in \mathbb{N}^* \\ E^w & \text{if } N = +\infty. \end{cases} \tag{6}$$

The following lemma will be useful in the next sections.

Lemma 3.3 *Under* $(\text{H}_0, \text{H}_1, \text{H}_2, \text{H}_3)$, *for almost all $\tilde{\xi} \in \tilde{\Xi}$, $\Gamma_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^r$ defined by* (4) *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, $\lim_{N\to\infty} \mathbb{H}(\hat{F}_N(S, \tilde{\xi}), \mathbb{E}_\xi[F(S, \cdot)]) = 0$, which means (Remark 3.7) that $\Gamma_{\tilde{\xi}}$ is H-continuous at $+\infty$. Moreover, $\Gamma_{\tilde{\xi}}(+\infty)$ is a compact set. 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 iff all its objective functions are convex and its feasible set is convex. Using the well known Scalarization Theorem for convex MOP (see, e.g., [39, Propositions 3.7 and 3.8], [40] or [41]), we obtain immediately the following.

Theorem 4.1 *Under* $(\mathrm{H}_4, \mathrm{H}_5)$, *we have*

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

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*

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

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 *Under* $(\mathrm{H}_0, \mathrm{H}_1, \mathrm{H}_2)$, *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., [8, 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 $x \mapsto \mathbb{E}_\xi[F(x, \cdot)]$ is continuous. □

Proposition 4.2 *Under* $(\mathrm{H}_0, \mathrm{H}_1, \mathrm{H}_2)$, *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 $N \geq 1$ and let $\lambda \in \mathbb{R}^r_+ \setminus \{0\}$. Since S is compact and $x \mapsto \langle \lambda, \hat{F}_N(x, \tilde{\xi}) \rangle$ is continuous on S for almost every $\tilde{\xi} \in \tilde{\Xi}$, the first conclusion follows easily by

Weierstrass' theorem since $\arg\min_{x\in S}\langle\lambda, \hat{F}_N(x,\tilde{\xi})\rangle \subset E_N^w(\tilde{\xi})$ (see Theorem 3.5 in [39] or [43]). In the same way, we obtain $E^w \neq \emptyset$. □

Now we state the main result of this section.

Theorem 4.2 *Under* $(\mathrm{H}_0, \mathrm{H}_1, \mathrm{H}_2, \mathrm{H}_3, \mathrm{H}_4, \mathrm{H}_5)$, *for almost all* $\tilde{\xi}$ *in* $\tilde{\Xi}$, *the Hausdorff–Pompeiu distance between the SAA-N weakly Pareto sets* $E_N^w(\tilde{\xi})$ *and the true weakly Pareto set* E^w *tends to zero as* N *tends to infinity, i.e.*

$$\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi} \in \tilde{\Xi} \mid \lim_{N\to\infty} \mathbb{H}\big(E_N^w(\tilde{\xi}), E^w\big) = 0\Big\}\Big) = 1.$$

The proof of the theorem is an immediate consequence of the following two lemmas.

Lemma 4.1 *Under* $(\mathrm{H}_0, \mathrm{H}_1, \mathrm{H}_2, \mathrm{H}_3)$, *for almost all* $\tilde{\xi}$ *in* $\tilde{\Xi}$, *the deviation of the SAA-N weakly Pareto sets* $E_N^w(\tilde{\xi})$ *from the true weakly Pareto set* E^w *tends to zero as* N *tends to infinity. In other words*,

$$\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi} \in \tilde{\Xi} \mid \lim_{N\to\infty} \mathbb{D}\big(E_N^w(\tilde{\xi}), E^w\big) = 0\Big\}\Big) = 1.$$

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 $E_N^w(\tilde{\xi})$ is nonempty by Proposition 4.2, and it is compact by Proposition 4.1.

The set-valued mapping $\Gamma_{\tilde{\xi}}$ (introduced in (4)) is continuous at $+\infty$ (Lemma 3.3) and $\Gamma_{\tilde{\xi}}(+\infty)$ is a compact set (Proposition 4.1). By [42, Theorem 4.3, p.104] it follows that the set-valued mapping $\Upsilon_{\tilde{\xi}}$ (introduced in (6)) 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* $(\mathrm{H}_0, \mathrm{H}_1, \mathrm{H}_2, \mathrm{H}_3, \mathrm{H}_4, \mathrm{H}_5)$, *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* $E_N^w(\tilde{\xi})$ *tends to zero as* N *tends to infinity, i.e*

$$\tilde{\mathbb{P}}_\xi\Big(\Big\{\tilde{\xi} \in \tilde{\Xi} \mid \lim_{N\to\infty} \mathbb{D}\big(E^w, E_N^w(\tilde{\xi})\big) = 0\Big\}\Big) = 1.$$

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 $\hat{x} \in \arg\min_{x\in S}\langle\lambda, \mathbb{E}_\xi[F(x,\cdot)]\rangle$.

Now, consider for each $N \in \mathbb{N}^*$ an element

$$x_N \in \arg\min_{x\in S}\big\langle\lambda, \hat{F}_N(x,\tilde{\xi})\big\rangle, \tag{7}$$

which is possible since the last set is nonempty according to Weierstrass' theorem. Thus, according to the Scalarization theorem, we obtain a sequence (x_N) such that $x_N \in E_N^w(\tilde{\xi})$ (each N). 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 (7), $x_{\phi(k)} \in \arg\min_{x\in S}\langle\lambda, \hat{F}_{\phi(k)}(x,\tilde{\xi})\rangle$ $\forall k \geq 1$ and then

$$\langle\lambda, \hat{F}_{\phi(k)}(x_{\phi(k)},\tilde{\xi})\rangle \leq \langle\lambda, \hat{F}_{\phi(k)}(\hat{x},\tilde{\xi})\rangle. \tag{8}$$

Since $\hat{x} \in \arg\min_{x\in S}\langle\lambda, \mathbb{E}_\xi[F(x,\cdot)]\rangle$, taking the limit in (8) implies

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

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 $d(\hat{x}, E_N^w(\tilde{\xi})) \to 0$ 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 $\mathbb{D}(E^w, E_N^w(\tilde{\xi})) \to 0$ as N tends to infinity for a.e. $\tilde{\xi} \in \tilde{\Xi}$. □

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* $\arg\min_{x\in A} g(x)$ *is closed.*

Proof Let $x \in \overline{\arg\min_{x\in A} g(x)}$, where $\overline{A}$ denotes the topological closure of a set A. There exists a sequence $(x_N)_{N\geq 1}$ in $\arg\min_{x\in A} g(x)$ such that $x_N \to x$.

Thus, $g(x) = \lim g(x_N) = g(x_N)$ (each N), hence $x \in \arg\min_{x\in A} g(x)$. □

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

$$\Sigma_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \rightrightarrows \mathbb{R}^n, \quad \Sigma_{\tilde{\xi}}(N) := \begin{cases} \arg\min_{x\in\Upsilon_{\tilde{\xi}}(N)} \hat{F}_N^0(x,\tilde{\xi}) & \text{if } N \in \mathbb{N}^* \\ \arg\min_{x\in\Upsilon_{\tilde{\xi}}(+\infty)} \mathbb{E}_\xi\big[F^0(x,\cdot)\big] & \text{if } N = +\infty, \end{cases} \tag{9}$$

where $\Upsilon_{\tilde{\xi}}$ was defined in (6).

Remark 4.2 For each $N \in \mathbb{N}^*$, and for almost all $\tilde{\xi} \in \tilde{\Xi}$, by Weierstrass' theorem, the set $\Sigma_{\tilde{\xi}}(N)$ is nonempty because $x \mapsto \hat{F}_N^0(x,\tilde{\xi})$ is continuous on the compact set $E_N^w(\tilde{\xi})$. Moreover, it is compact by Lemma 4.3. The same rule obviously holds for $\Sigma_{\tilde{\xi}}(+\infty)$.

Now, we can introduce the following optimal value function:

$$V_{\tilde{\xi}} : \mathbb{N}^* \cup \{+\infty\} \to \mathbb{R}, \quad V_{\tilde{\xi}}(N) := \begin{cases} \min_{x\in\Upsilon_{\tilde{\xi}}(N)} \hat{F}_N^0(x,\tilde{\xi}) & \text{if } N \in \mathbb{N}^* \\ \min_{x\in\Upsilon_{\tilde{\xi}}(+\infty)} \mathbb{E}_\xi\big[F^0(x,\cdot)\big] & \text{if } N = +\infty. \end{cases} \tag{10}$$

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 continuous 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 [36, 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 denotes 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 [34, 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 $\Sigma_{\tilde{\xi}}(+\infty)$ is a closed set by Lemma 4.3.

Then, [33, Theorem 5.2.4] and Remark 3.6 imply that $\limsup_{N\to+\infty} \Sigma_{\tilde{\xi}}(N) \subset \Sigma_{\tilde{\xi}}(+\infty)$, which concludes the last sentence of the theorem. □

5 Results in the Image 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 and 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 $\tilde{x} \in \arg\min_{x\in S}\langle\lambda, \hat{F}_N(x, \tilde{\xi})\rangle$. 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 $\langle\lambda, \hat{F}_N(\hat{x}, \tilde{\xi})\rangle < \langle\lambda, \hat{F}_N(\tilde{x}, \tilde{\xi})\rangle$, a contradiction. Finally, $\tilde{x} \in E_N(\tilde{\xi})$. 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 nonempty 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 $d(\overline{x}, A_N) \to 0$ as $N \to \infty$ for all $\overline{x} \in \overline{A}$.*

Proof Let $\overline{x} \in \overline{A}$. Then there exists a sequence $(x_k)_{k\geq 1}$ in A such that $d(x_k, \overline{x}) \to 0$ as $k \to \infty$. Let $\epsilon > 0$ be fixed. $\exists k$ such that $d(x_k, \overline{x}) < \frac{\epsilon}{2}$. 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 $d(\overline{x}, A_N) \leq d(\overline{x}, x_k) + d(x_k, A_N)$, we obtain $d(\overline{x}, A_N) < \epsilon$ 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 $\lambda \in \operatorname{int}(\mathbb{R}^r_+)$. Since $x \mapsto \hat{F}_N(x, \tilde{\xi})$ is continuous, the set $Z_y = \{y' \in \mathbb{R}^r \,|\, y' \in (y - \mathbb{R}^r_+) \cap \hat{F}_N(S, \tilde{\xi})\}$ is nonempty and compact. Thus there exists $\hat{y} \in \arg\min_{y' \in Z_y} \langle \lambda, y' \rangle$. 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 $\langle \lambda, z \rangle < \langle \lambda, \hat{y} \rangle$, 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*

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

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 $(x_N)_{N \geq 1}$ with $x_N \in E_N(\tilde{\xi})$ and

$$\hat{F}_N(x_N, \tilde{\xi}) \leq \hat{F}_N(x, \tilde{\xi}) \quad \text{for each } N \geq 1. \tag{11}$$

On one hand, since $\hat{F}_N(S, \tilde{\xi})$ is compact (each N), 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 (11), 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 $\tilde{\xi} \in \tilde{\Xi}$

$$\lim_{N \to \infty} d\big(\mathbb{E}_\xi\big[F(x, \cdot)\big], \hat{F}_N\big(E_N(\tilde{\xi}), \tilde{\xi}\big)\big) = 0.$$

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*

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

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 [36, 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*

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

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

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

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

$$f(y_N)<U_N(\tilde{\xi})+a_N\leq f(y_N)+a_N. \tag{12}$$

By Lemma 3.2, there exists a compact set $K\subset\mathbb{R}^r$ such that $\mathbb{E}_\xi[F(S,\cdot)]\subset K$, and, for all $N\geq 1$, $\hat{F}_N(S,\tilde{\xi})\subset K$. 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 (12), $W=\lim_{N\to\infty} f(y_{\phi(\varphi(N))})$. 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

$$f(\hat{y})<W. \tag{13}$$

By Proposition 5.2, there exists a sequence $(\hat{y}_{\phi(\varphi(N))})_{N\geq 1}$ such that, for each N, $\hat{y}_{\phi(\varphi(N))}\in\hat{F}_N(E_{\phi(\varphi(N))}(\tilde{\xi}),\tilde{\xi})$ 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 (13) 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

In this section, we give a toy example in order to illustrate the Hausdorff–Pompeiu convergence of the Pareto sets in the image space and the convergence of the SAA-N optimal values.

We consider the following Stochastic Bi-objective Optimization Problem:

$$\min_{x\in S} \mathbb{E}_\xi\big[F^1(x,\cdot), F^2(x,\cdot)\big],$$

where the decision variable $x=(x_1,x_2)\in\mathbb{R}^2$, the random variable ξ follows a uniform distribution with mean $\frac{3}{2}$ and variance $\frac{1}{12}$, the feasible set $S=[0,1]\times[0,1]$, and $\mathbb{E}_\xi[F^1(x,\cdot), F^2(x,\cdot)] = \mathbb{E}[x_1+x_2+\xi, (x_1-x_2+\xi)^2]$.

For this simple example, it is possible to compute in closed form the expectation of the objectives: $\forall x=(x_1,x_2)\in\mathbb{R}^2$,

$$\mathbb{E}_\xi\big[F^1(x,\cdot)\big] = x_1+x_2+\frac{3}{2},$$

$$\mathbb{E}_\xi\big[F^2(x,\cdot)\big] = \frac{1}{3}\big((x_1-x_2+2)^3-(x_1-x_2+1)^3\big) = (x_1-x_2)^2+3(x_1-x_2)+\frac{7}{3}.$$

The set $\mathbb{E}_\xi[F^1(S,\cdot), F^2(S,\cdot)]$ is the curvilinear quadrilateral domain $ABCD$ with $A(\frac{3}{2},\frac{7}{3})$, $B(\frac{5}{2},\frac{1}{3})$, $C(\frac{7}{2},\frac{7}{3})$, $D(\frac{5}{2},\frac{19}{3})$ (see Fig. 1), and $\widehat{AB}$, $\widehat{BC}$, $\widehat{CD}$, $\widehat{DA}$ are parabolic arcs having the following parametric representations:

$$\widehat{AB}: [0,1]\ni t\mapsto \left(t+\frac{3}{2}, t^2-3t+\frac{7}{3}\right),$$

$$\widehat{BC}: [0,1]\ni t\mapsto \left(t+\frac{5}{2}, t^2+t+\frac{1}{3}\right),$$

$$\widehat{DC}: [0,1]\ni t\mapsto \left(t+\frac{5}{2}, t^2-5t+\frac{19}{3}\right),$$

$$\widehat{AD}: [0,1]\ni t\mapsto \left(t+\frac{3}{2}, t^2+3t+\frac{7}{3}\right).$$

Then the true Pareto set image can be identified graphically and is given by the arc $\widehat{AB}$ (in bold in Fig. 1). Hence the true Pareto set is $E=\{0\}\times[0,1]$.

Consider now the problem (O) (see Sect. 2) with $f(y_1,y_2):=2y_1+y_2$. Thus, in this particular case, problem (O) consists in the minimization of the function

$$x\mapsto \phi(x) := f\big(\mathbb{E}\big[F^1\big(x,\xi(\cdot)\big), F^2\big(x,\xi(\cdot)\big)\big]\big) = 2\mathbb{E}_\xi\big[F^1(x,\cdot)\big]+\mathbb{E}_\xi\big[F^2(x,\cdot)\big]$$

over the Pareto set E associated with (SBOP), i.e. we want to solve

$$\min_{(x_1,x_2)\in\{0\}\times[0,1]}\left[(x_1-x_2)^2+5x_1-x_2+\frac{16}{3}\right].$$

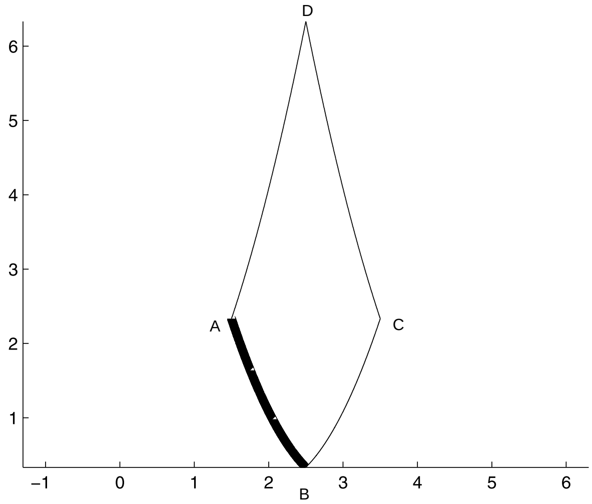


Fig. 1 $\mathbb{E}_\xi[F(E,\cdot)]$ (**bold**) and the boundary of $\mathbb{E}_\xi[F(S,\cdot)]$

Since $\phi(0,\cdot)$ is strictly convex over [0, 1], it is easy to see that its unique minimizer is $\frac{1}{2}$. Hence the true optimal value is given by

$$U = \phi\left(0, \frac{1}{2}\right) = \frac{61}{12}.$$

Now, let $\tilde{\xi} = (\xi_1, \xi_2, \ldots,) \in \tilde{\Xi}$ be given, and consider the corresponding SAA-N problem:

$$\min_{(x_1,x_2)\in E_N(\tilde{\xi})} \phi_N(x_1, x_2, \tilde{\xi}),$$

where

$$\phi_N(x_1, x_2, \tilde{\xi}) := 2\frac{1}{N}\sum_{k=1}^{N}(x_1 + x_2 + \xi_k) + \frac{1}{N}\sum_{k=1}^{N}(x_1 - x_2 + \xi_k)^2,$$

and $E_N(\tilde{\xi})$ is the SAA-N Pareto set associated with the bi-objective minimization problem

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

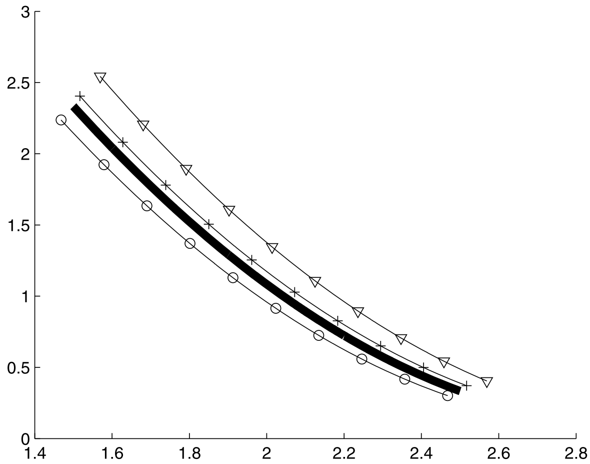
where

$$\hat{F}_N(x, \tilde{\xi}) := \frac{1}{N}\sum_{k=1}^{N}\big(x_1 + x_2 + \xi_k, (x_1 - x_2 + \xi_k)^2\big).$$

Denote

$$\mu_N := \frac{1}{N}\sum_{k=1}^{N}\xi_k, \qquad \sigma_N^2 := \frac{1}{N}\sum_{k=1}^{N}\xi_k^2 - \mu_N^2.$$

Fig. 2 $\hat{F}_5(E_5(\tilde{\xi}), \tilde{\xi})$ (△–△), $\hat{F}_{15}(E_{15}(\tilde{\xi}), \tilde{\xi})$ (○–○), $\hat{F}_{25}(E_{25}(\tilde{\xi}), \tilde{\xi})$ (×–×), $\mathbb{E}_\xi[F(E, \cdot)]$ **(bold)**



Thus,

$$\phi_N(x_1, x_2, \tilde{\xi}) = (x_1 - x_2)^2 + 2(1 + \mu_N)x_1 + 2(1 - \mu_N)x_2 + \sigma_N^2 + 2\mu_N + \mu_N^2,$$
$$\hat{F}_N(x, \tilde{\xi}) = \big(x_1 + x_2 + \mu_N, (x_1 - x_2)^2 + 2\mu_N(x_1 - x_2) + \sigma_N^2 + \mu_N^2\big).$$

The set $\hat{F}_N(S, \tilde{\xi})$ is the curvilinear quadrilateral domain $A_N B_N C_N D_N$ with $A_N(\mu_N, \sigma_N^2 + \mu_N^2)$, $B_N(1 + \mu_N, (1 - \mu_N)^2 + \sigma_N^2)$, $C_N(2 + \mu_N, \sigma_N^2 + \mu_N^2)$, $D_N(1 + \mu_N, (1 + \mu_N)^2 + \sigma_N^2)$, and $\widehat{A_N B_N}$, $\widehat{B_N C_N}$, $\widehat{C_N D_N}$, $\widehat{D_N A_N}$ are parabolic arcs having the following parametric representations:

$$\widehat{A_N B_N}: \ [0,1] \ni t \mapsto \big(t + \mu_N, t^2 - 2\mu_N t + \sigma_N^2 + \mu_N^2\big),$$
$$\widehat{B_N C_N}: \ [0,1] \ni t \mapsto \big(t + 1 + \mu_N, t^2 + 2(\mu_N - 1)t + \sigma_N^2 + (\mu_N - 1)^2\big),$$
$$\widehat{D_N C_N}: \ [0,1] \ni t \mapsto \big(t + 1 + \mu_N, t^2 - 2(\mu_N + 1)t + \sigma_N^2 + (\mu_N + 1)^2\big),$$
$$\widehat{A_N D_N}: \ [0,1] \ni t \mapsto \big(t + \mu_N, t^2 + 2\mu_N t + \sigma_N^2 + \mu_N^2\big).$$

Since $\mu_N \in [1, 2]$, the set $\hat{F}_N(S, \tilde{\xi})$ has the same shape as $\mathbb{E}_\xi[F^1(S, \cdot), F^2(S, \cdot)]$. Thus, for each N, the SAA-N Pareto set image is the arc $\widehat{A_N B_N}$, and so the set $E_N(\tilde{\xi}) = E = \{0\} \times [0, 1]$. Then, it is easy to see that $\forall N \geq 1$, the SAA-N optimal value is given by

$$U_N(\tilde{\xi}) = \phi_N(0, \mu_N - 1, \tilde{\xi}) = \sigma_N^2 + 4\mu_N - 1.$$

Using MATLAB 7, we generate a random vector $(\xi_1, \xi_2, \ldots, \xi_{250})$ of realizations of ξ which satisfy (H_1).

Figure 2 focuses on the Hausdorff convergence of the SAA-N Pareto sets image to the true Pareto set image.

Figure 3 illustrates the convergence of the SAA-N optimal values to the true optimal value for $N = 1, \ldots, 250$.

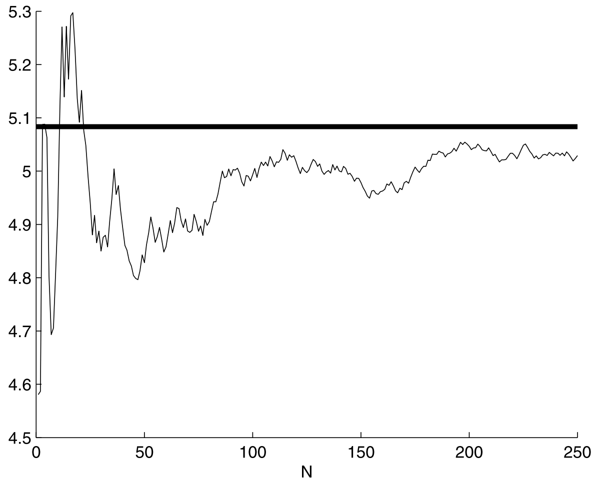


Fig. 3 Convergence of the SAA-N Optimal Values $U_N(\tilde{\xi})$ to the true optimal value U (**bold**)

7 Conclusions

In this paper, we have considered the novel problem of optimizing a scalar random function over the Pareto set associated with a Stochastic Multi-objective Optimization Problem. Since in most real cases it is impossible to solve directly this problem, we used the well known Sample Average Approximation method to obtain an approximated problem. To show that the sequence of SAA-N optimal values tends almost surely to the true optimal value, we have shown that the Hausdorff–Pompeiu distance between the SAA-N Pareto set images and the true Pareto set image vanishes at infinity.

Moreover, assuming strict convexity, we have shown that the Hausdorff–Pompeiu distance between the SAA-N Pareto sets and the true Pareto set vanishes at infinity, and then we obtained that every cluster point of any sequence of SAA-N optimal solutions is an optimal solution for the true problem.

Finally, we have illustrated graphically the Hausdorff–Pompeiu convergence of the Pareto set images to the true Pareto set image, and the convergence of SAA-N optimal values to the true optimal value using MATLAB7 with an example of a stochastic bi-objective optimization problem, the scalar random function to be optimized over the Pareto set being a conic combination of the two random objectives. For this illustrative example it is possible to find the closed form of the expectations and to determine the true Pareto and Pareto image sets as well as the Pareto and Pareto image sets associated with SAA-N problems.

Further research avenues may include the study of the numerical aspects related to the problem of stochastic optimization over a stochastic Pareto set using SAA-N approach. It is important to find an estimation of the size N of the sample in order to obtain an a priori error bound for the optimal value and/or optimal solution. Also, it is necessary to propose appropriate algorithms to solve the SAA-N problem, especially for the case when the problem is given in the image space.

Acknowledgements The authors are grateful to the anonymous referees for their useful comments which have improved the quality of the paper.

References

1. Benson, H.P.: Optimization over the efficient set. J. Math. Anal. Appl. **98**, 562–580 (1984)
2. Konno, H., Thach, P.T., Yokota, D.: Dual approach to minimization on the set of Pareto-optimal solutions. J. Optim. Theory Appl. **88**, 689–707 (1996)
3. Bolintinéanu, S.: Minimization of a quasi-concave function over an efficient set. Math. Program. **61**, 89–110 (1993)
4. Benson, H.P.: A finite, non-adjacent extreme point search algorithm for optimization over the efficient set. J. Optim. Theory Appl. **73**, 47–64 (1992)
5. Philip, J.: Algorithms for the vector maximization problem. Math. Program. **2**, 207–229 (1972)
6. An, L.T.H., Muu, L.D., Tao, P.D.: Numerical solution for optimization over the efficient set by d.c. optimization algorithms. Oper. Res. Lett. **19**, 117–128 (1996)
7. Bolintinéanu, S.: Optimality conditions for minimization over the (weakly or properly) efficient set. J. Math. Anal. Appl. **173**(2), 523–541 (1993)
8. Bolintinéanu, S.: Necessary conditions for nonlinear suboptimization over the weakly-efficient set. J. Optim. Theory Appl. **78** (1993)
9. Bonnel, H., Kaya, C.Y.: Optimization over the efficient set of multi-objective control problems. J. Optim. Theory Appl. **147**(1), 93–112 (2010)
10. Craven, B.D.: Aspects of multicriteria optimization. In: Recent Developments in Mathematical Programming, pp. 93–100 (1991)
11. Dauer, J.P.: Optimization over the efficient set using an active constraint approach. J. Oper. Res. **35**, 185–195 (1991)
12. Dauer, J.P., Fosnaugh, T.A.: Optimization over the efficient set. J. Glob. Optim. **7**, 261–277 (1995)
13. Fülöp, J.: A cutting plane algorithm for linear optimization over the efficient set. In: Generalized Convexity. Lecture notes in Economics and Mathematical System, vol. 405, pp. 374–385. Springer, Berlin (1994)
14. Horst, R., Thoai, N.V.: Maximizing a concave function over the efficient or weakly-efficient set. Eur. J. Oper. Res. **117**, 239–252 (1999)
15. Horst, R., Thoai, N.V., Yamamoto, Y., Zenke, D.: On Optimization over the Efficient Set in Linear Multicriteria Programming. J. Optim. Theory Appl. **134**, 433–443 (2007)
16. Yamamoto, Y.: Optimization over the efficient set: an overview. J. Glob. Optim. **22**, 285–317 (2002)
17. Bonnel, H., Morgan, J.: Semivectorial bilevel convex optimal control problems: an existence result. SIAM J. Control Optim. **50**(6), 3224–3241 (2012)
18. Dempe, S., Gadhi, N., Zemkoho, A.B.: New optimality conditions for the semivectorial bilevel optimization problem. J. Optim. Theory Appl. **157**(1), 54–74 (2013)
19. Ankhili, Z., Mansouri, A.: An exact penalty on bilevel programs with linear vector optimization lower level. Eur. J. Oper. Res. **197**, 36–41 (2009)
20. Zheng, Y., Wan, Z.: A solution method for semivectorial bilevel programming problem via penalty method. J. Appl. Math. Comput. **37**, 207–219 (2011)
21. Eichfelder, G.: Multiobjective bilevel optimization. Math. Program., Ser. A **123**, 419–449 (2010)
22. Bonnel, H.: Optimality conditions for the semivectorial bilevel optimization problem. Pac. J. Optim. **2**(3), 447–468 (2006)
23. Bonnel, H., Morgan, J.: Semivectorial bilevel optimization problem: penalty approach. J. Optim. Theory Appl. **131**(3), 365–382 (2006)
24. Benson, H.P., Lee, D.: Outcome-based algorithm for optimizing over the efficient set of a bicriteria linear programming problem. J. Optim. Theory Appl. **88**, 77–105 (1996)
25. Benson, H.P.: Generating the efficient outcome set in multiple objective linear programs: the bicriteria case. Acta Math. Vietnam. **22**, 29–51 (1997)
26. Benson, H.P.: Further analysis of an outcome set-based algorithm for multiple objective linear programming. J. Optim. Theory Appl. **97**, 1–10 (1998)
27. Benson, H.P.: Hybrid approach for solving multiple objective linear programs in outcome space. J. Optim. Theory Appl. **98**, 17–35 (1998)
28. Kim, N.T.B., Thang, T.N.: Optimization over the efficient set of a bicriteria convex programming problem. Pac. J. Optim. **9**, 103–115 (2013)
29. Fliege, J., Xu, H.: Stochastic multiobjective optimization: sample average approximation and applications. J. Optim. Theory Appl. **151**, 135–162 (2011)
30. Jacod, J., Protter, P.: Probability Essentials. Springer, Berlin (2004)
31. Shapiro, A., Dentcheva, D., Ruszczynski, A.: Lectures on stochastic programming: modeling and theory. MPS/SIAM Ser. Optim. (2009)

32. Dieudonné, J.J.: Foundations of Modern Analysis. Academic Press, New York (1960)
33. Aubin, J.P., Frankowska, H.: Set Valued Analysis. Birkhäuser, Basel (1990)
34. Bonnans, J.F., Shapiro, A.: In: Perturbation Analysis of Optimization Problem. Springer Series in Operations Research (2000)
35. Burachik, R.S., Iusem, A.N.: Set-Valued Mappings and Enlargements of Monotone Operators. Springer, Berlin (2008)
36. Göpfert, A., Riahi, H., Tammer, C., Zălinescu, C.: Variational Methods in Partially Ordered Spaces. Springer, Berlin (2003)
37. Chen, G.Y., Huang, X., Yang, X.: Vector Optimization: Set Valued and Variational Analysis. Springer, Berlin (2005)
38. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis. Springer, Heidelberg (1998)
39. Ehrgott, M.: Multicriteria Optimization. Springer, Berlin (2000)
40. Jahn, J.: Vector Optimization. Springer, Berlin (2004)
41. Miettinen, K.M.: Nonlinear Multiobjective Optimization. Kluwer Academic, Amsterdam (1998)
42. Luc, D.T.: Theory of Vector Optimization. Lecture Notes in Econom. and Math. Systems, vol. 319. Springer, Berlin (1989)
43. Yu, P.L.: Multiple-Criteria Decision Making: Concepts, Techniques, and Extensions. Mathematical Concepts and Methods in Science and Engineering, vol. 30. Plenum Press, New York (1985)

Chapitre 3

Seconde Contribution

A paraître dans Journal of Global Optimization

Noname manuscript No.
(will be inserted by the editor)

Optimization over the Pareto Outcome Set Associated with a Convex Bi-objective Optimization Problem: Theoretical Results, Deterministic Algorithm and Application to the Stochastic Case

Henri Bonnel · **Julien Collonge**

Received: date / Accepted: date

Abstract Our paper consists of two main parts. In the first one, we deal with the deterministic problem of minimizing a real valued function f over the Pareto outcome set associated with a deterministic convex bi-objective optimization problem (BOP), in the particular case where f depends on the objectives of (BOP), i.e. we optimize over the Pareto set in the Outcome space. In general, the optimal value U of such a kind of problem cannot be computed directly, so we propose a deterministic outcome space algorithm whose principle is to give at every step a range (lower bound, upper bound) that contains U. Then we show that for any given error bound, the algorithm terminates in a finite number of steps.

In the second part of our paper, in order to handle also the stochastic case, we consider the situation where the two objectives of (BOP) are given by expectations of random functions, and we deal with the stochastic problem (S) of minimizing a real valued function f over the Pareto outcome set associated with this Stochastic bi-objective Optimization Problem (SBOP). Because of the presence of random functions, the Pareto set associated with this type of problem cannot be explicitly given, and thus it is not possible to compute the optimal value V of problem (S). That is why we consider a sequence of Sample Average Approximation problems (SAA-N, where N is the sample size) whose optimal values converge almost surely to V as the sample size N goes to infinity.

Assuming f nondecreasing, we show that the convergence rate is exponential, and we propose a confidence interval for V.

Finally, some computational results are given to illustrate the paper.

Keywords Optimization over the Pareto image set · Multi-objective deterministic optimization · Multi-objective stochastic optimization · Multi-objective convex optimization · Sample average approximation method

Mathematics Subject Classification (2000) 90C29 · 90C25 · 90C15 · 90C26

H. Bonnel
ERIM, University of New Caledonia, B.P. R4, 98851 Noumea Cedex, New Caledonia
E-mail: henri.bonnel@univ-nc.nc

J. Collonge
ERIM, University of New Caledonia, B.P. R4, 98851 Noumea Cedex, New Caledonia E-mail: julien.collonge@univ-nc.nc

1 Introduction

For a bi-objective Optimization Problem (BOP), the solution set (also called Pareto set) consists of the feasible solutions which ensure some sort of equilibrium amongst the objectives. That is to say, Pareto solutions are such that neither of the objective values can be improved further without deteriorating the other.

The Pareto set associated with a (BOP) may be infinite and even unbounded, and then arises the decision problem of choosing one solution over this set, which is generally not explicitly given and not convex (even in the linear case). A useful and efficient way to answer this problem is to consider the problem of optimizing a real valued function over the Pareto set.

This problem has been intensively studied during the last decades beginning with Philip's paper [33], and continued by many authors in [1–17, 21, 24, 25] (see [36] for an extensive bibliography).

Since in mathematical models the number of variables is often greater than the number of objectives, Benson argues [4–7] that generating the Pareto outcome set require less computation than generating the Pareto set itself. Moreover, it is reasonable to take a decision based on the objective values rather than on variable values. So it can be useful to consider the Problem (P) of minimizing a real valued function f over the Pareto outcome set. This particular problem has been studied in [3, 13, 28].

However, the optimal value of such a kind of problem is rarely possible to compute. That is the reason why we propose (in the first part of our paper) a deterministic outcome space algorithm inspired by the one proposed in [28]. The algorithm in [28] is quite restrictive since it deals with the problem of maximizing a nondecreasing function over the Pareto outcome set associated with a convex (BOP). Our algorithm applies to any continuous function, and is as follows: at the beginning of Step $k (k \geq 1)$, we know from Step $k-1$:

- k right triangles whose union contains the Pareto outcome set.
- $k+1$ efficient points.

If the minimum value of f over the right triangles (lower bound) is close enough to the minimum value of f over the efficient points (upper bound), then the algorithm stops. Else, a new efficient point is generated and one triangle is reduced by two new right triangles with smaller hypotenuse. This process is repeated until the chosen error bound is not reached.

Due to the difficulty of optimizing over a Pareto set, authors generally deal with deterministic optimization problems, although in real life observed phenomena are randomly disturbed.

That is why, in the second part of our paper, we consider the stochastic problem (S) of minimizing a real valued function f over the Pareto outcome set associated with a Stochastic Convex Bi-Objective Optimization Problem ($SBOP$), whose objectives are expectations of random functions. This kind of problem may have many applications, including financial models like mean-variance portfolio optimization [29, 30].

Some results taking into account uncertainty have been recently obtained by Bonnel and Collonge in [13], where the problem of optimizing over a Pareto set associated with a Stochastic Multi-Objective Optimization Problem has been studied. Some approximations using the well known Sample Average Approximation (SAA-N, where N is the sample size) method have been given. In particular, it has been shown that the sequence of optimal values V_N of problems given by the SAA method converges almost surely to the optimal value V

of problem (S) as the sample size N goes to infinity. One of our goals is to find the rate of this convergence.

Moreover, for a given error bound $\varepsilon > 0$ and a given confidence level $p_0 \in]0,1[$, another goal is to find a rank $N^0(\varepsilon, p_0)$ which explicitly depends on data, such that the optimal value $V_{N^0(\varepsilon,p_0)}$ of problem SAA-$N^0(\varepsilon, p_0)$ is an ε-approximation of V with probability greater than p_0. As soon as this rank is given, it is possible to compute a range (upper bound, lower bound) of V with our algorithm, and thus to propose a confidence interval of length ε which contains V with probability greater than p_0.

Our paper is organized as follows.

In Section 2, we present the deterministic problem under consideration: optimizing a continuous real valued function over the Pareto outcome set associated with a deterministic convex bi-objective optimization problem. Under some reasonable assumptions, we show that this problem admits at least one solution in the decision space. Then we propose a deterministic outcome space algorithm to determine the optimal value of the considered problem, and finally we prove that it terminates in a finite number of steps whatever the error bound is.

Section 3 deals with the stochastic problem (S) of optimizing a real valued function over the Pareto outcome set associated with a stochastic convex bi-objective optimization problem. We begin with some basic definitions and facts necessary for the development of the stochastic case. Then we introduce problem (S) together with the sequence of SAA-N problems proposed in [13]. We recall the basic result that the sequence of SAA-N optimal values converges to the true optimal value V with probability one (see [13] for more details). Unlike the theoretical results on convergence presented in [13], we are now concerned with convergence rate and error estimates.

Thus we show that, assuming f Lipschitz continuous on a suitable compact set, non-decreasing on $\mathbb{R}^2$ and the support of the random perturbations bounded, the convergence with probability one of the sequence of SAA-N optimal values to the true optimal value V has an exponential rate. Moreover, for a given error bound $\varepsilon > 0$ and a given confidence level $p_0 \in]0,1[$, we propose a rank $N^0 = N^0(\varepsilon, p_0)$ such that the SAA-N^0 optimal value is an ε-approximation of V with probability greater than p_0. Note that N^0 explicitly depends on data and on a uniform upper bound of the standard deviation of the random bi-objective function. Thus we compute (with our algorithm) a confidence interval that contains V.

In the last Section, we present two numerical examples obtained with MATLAB7. The first example is a simple one with a two dimensional feasible set and a linear vector objective which allows us to graphically illustrate our algorithm. The second one use a four dimensional feasible set and a nonlinear vector objective in order to test the performances of the algorithm.

2 A Deterministic Outcome Space Algorithm

Let $y,z \in \mathbb{R}^2$. In this paper we use the notation $y \leq z$ to indicate $y^i \leq z^i$ for all $i = 1,2$, $y \lneqq z$ to indicate $y \leq z$ and $y \neq z$, and $y < z$ to indicate $y^i < z^i$ for all $i = 1,2$.

We note $\mathbb{R}^2_+ := \{x \in \mathbb{R}^2 | 0 \leq x\}$ the Pareto cone, and $\text{int}(\mathbb{R}^2_+) := \{x \in \mathbb{R}^2 | 0 < x\}$ its topological interior. For a subset T of $\mathbb{R}^2$, we denote by $w\text{-MIN}_{\mathbb{R}^2_+}(T) := \{y \in T | \nexists \bar{y} \in T : y - \bar{y} \in \text{int}(\mathbb{R}^2_+)\}$ the weakly minimal set of T with respect to the Pareto cone, and

$\mathrm{MIN}_{\mathbb{R}^2_+}(T) := \{y \in T |\ \nexists \bar{y} \in T \setminus \{y\} : y - \bar{y} \in \mathbb{R}^2_+\}$ the minimal set of T with respect to the cone $\mathbb{R}^2_+$.

Moreover, for a (BOP) $\mathrm{MIN}_{x\in X} F(x)$, where the function F is defined from $\mathbb{R}^n$ to $\mathbb{R}^2$, and the feasible set $X \subset \mathbb{R}^n$, we call $\mathrm{MIN}_{\mathbb{R}^2_+}(F(X))$ the Pareto outcome set (sometimes called Pareto set in the outcome space, outcome Pareto set or Pareto image set), and we call $\mathrm{ARGMIN}_{\mathbb{R}^2_+}(F(X)) := F^{-1}(\mathrm{MIN}_{\mathbb{R}^2_+}(F(X))) \cap X$ the Pareto set. Also, we denote by $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^2$. More details about solution set in vector optimization problems can be found in the monographs [19, 22, 23, 27, 31, 32].

2.1 Problem Statement

We deal with the problem of minimizing a real valued function f over the Pareto outcome set associated with a Bi-objective Optimization Problem (BOP). That is to say

$$(D) \qquad \min\Big(f(F(x))\Big| x \in \mathrm{ARGMIN}_{\mathbb{R}^2_+}(F(X))\Big),$$

where $f : \mathbb{R}^2 \to \mathbb{R}$, and $F : \mathbb{R}^n \to \mathbb{R}^2$ is the vector objective of the following

$$(BOP) \qquad \mathrm{MIN}_{x\in X} F(x) = \Big(F^1(x), F^2(x)\Big),$$

whose feasible set $X \subset \mathbb{R}^n$ is supposed nonempty and explicitly given by

$$X := \Big\{x \in \mathbb{R}^n \Big| g(x) = (g^1(x), g^2(x), \ldots, g^p(x)) \leq 0\Big\},$$

where for each $i = 1, \ldots, p$, $g^i : \mathbb{R}^n \to \mathbb{R}$.

In this section, we propose a deterministic outcome space algorithm in order to give an approximation of the optimal value of problem (D) under the following assumptions:

$$(H_d) \begin{cases} X \text{ is a nonempty and bounded subset of } \mathbb{R}^n. \\ \forall i = 1, \ldots, p,\ x \mapsto g^i(x) \text{ is convex on } \mathbb{R}^n. \\ \forall j = 1, 2,\ x \mapsto F^j(x) \text{ is continuous and convex on } X. \end{cases}$$

(H_0) f is continuous on $\mathbb{R}^2$.

Note that under (H_d), (BOP) is convex, i.e. its feasible set X is a convex set and its objectives are convex on X. Moreover, since the functions g^i are continuous on $\mathbb{R}^n$ (each i), X is a closed set and then it is compact because it is bounded.

2.2 Algorithm

Let us rewrite problem (D) in the Outcome space:

$$(P) \qquad \min\Big(f(z)\Big| z \in \mathrm{MIN}_{\mathbb{R}^2_+}(F(X))\Big).$$

To show that this outcome problem admits at least a solution, we need the following Lemma.

Lemma 1 *Let T be a compact set in $\mathbb{R}^2$ such that $T+\mathbb{R}^2_+$ is convex. Then the set $MIN_{\mathbb{R}^2_+}(T)$ is compact.*

Proof We consider a sequence a_k in $\text{MIN}_{\mathbb{R}^2_+}(T)$ converging to $a \in T$. It is easy to see that the set $w\text{-MIN}_{\mathbb{R}^2_+}(T)$ is closed and contains $\text{MIN}_{\mathbb{R}^2_+}(T)$. Hence $a \in w\text{-MIN}_{\mathbb{R}^2_+}(T)$. Suppose by contradiction that $a \notin \text{MIN}_{\mathbb{R}^2_+}(T)$. Then there exist $b \in T$ and $j \in \{1,2\}$ such that $b^j < a^j$ and $b^i = a^i$ for $i \neq j$, $i \in \{1,2\}$. For k large enough, we have $b^j < a_k^j$ and since $a_k \in \text{MIN}_{\mathbb{R}^2_+}(T)$, we have $a_k^i < b^i$. If $a_k^j < a^j$, then $a_k < a$ which contradicts $a \in w\text{-MIN}_{\mathbb{R}^2_+}(T)$. Thus $a^j \leq a_k^j$.

We have proven that for sufficiently large k, $b^j < a^j \leq a_k^j$ and $a_k^i < a^i = b^i$. Then there exists $\bar{t} \in]0,1]$ such that $a^j = \bar{t}a_k^j + (1-\bar{t})b^j$. Let $t = \frac{\bar{t}}{2}$. We have $a^j = ta_k^j + (1-t)b^j + c^j$, where $c^j = \frac{\bar{t}}{2}(a_k^j - b^j) > 0$, and $a^i = ta_k^i + (1-t)b^i + c^i$, where $c^i = \frac{\bar{t}}{2}(b^i - a_k^i) > 0$. Hence $a = ta_k + (1-t)b + c$, with $c \in \text{int}(\mathbb{R}^2_+)$.

Since $T+\mathbb{R}^2_+$ is convex, $ta_k + (1-t)b \in T + \mathbb{R}^2_+$, and then $a - c \in T + \mathbb{R}^2_+$. So there exist $a' \in T$ and $c' \in \mathbb{R}^2_+$ such that $a - c = a' + c'$. Hence $a - (c + c') \in T$, which contradicts the fact that $a \in w\text{-MIN}_{\mathbb{R}^2_+}(T)$ since $(c+c') \in \text{int}(\mathbb{R}^2_+)$. Finally, $\text{MIN}_{\mathbb{R}^2_+}(T)$ is a closed subset of the compact T, hence it is compact.

□

Proposition 1 *Problem (P) admits at least one solution.*

Proof (BOP) is convex, hence the set $F(X)+\mathbb{R}^2_+$ is convex ([37, Theorem 3.7]). Moreover, F is continuous on the compact set X, therefore the Pareto outcome set $\text{MIN}_{\mathbb{R}^2_+}(F(X))$ is a compact set by Lemma 1. Since f is continuous, Weierstrass' Theorem leads to the conclusion.

□

We denote by $U := \min\left(f(z)\middle| z \in \text{MIN}_{\mathbb{R}^2_+}(F(X))\right)$ the optimal value of problem (P), and we consider the following two sequences of problems whose optimal values converge to U:

$$(PS^k) \qquad \min_{z \in S^{(k)}} f(z),$$

$$(PD^k) \qquad \min_{z \in D^{(k)}} f(z),$$

where for all $k \geq 0$, the sets $D^{(k)}$ and $S^{(k)}$ are compact in $\mathbb{R}^2$ and satisfy:

$$D^{(k)} \subset D^{(k+1)} \subset \text{MIN}_{\mathbb{R}^2_+}(F(X)) \subset S^{(k+1)} \subset S^{(k)}.$$

For each $k \geq 0$, let $s^{(k)}$ and $d^{(k)}$ be the optimal value of problem (PS^k) and (PD^k) respectively, i.e.,

$$s^{(k)} := \min_{z \in S^{(k)}} f(z), \quad d^{(k)} := \min_{z \in D^{(k)}} f(z).$$

Hence, for each $k \geq 0$, we have $U \in [s^{(k)}, d^{(k)}]$ and $[s^{(k+1)}, d^{(k+1)}] \subset [s^{(k)}, d^{(k)}]$.

Since X is compact, according to Weierstrass' Theorem, the reals $x_0 := \min_{x \in X} F^1(x)$ and $y_1 := \min_{x \in X} F^2(x)$ are well defined. Therefore the sets $\text{argmin}\left(F^1(x)\middle| x \in X\right) := \{x \in X \,|\, F^1(x) = x_0\}$ and $\text{argmin}\left(F^2(x)\middle| x \in X\right) := \{x \in X \,|\, F^2(x) = y_1\}$ are nonempty, convex and compact, hence the reals $y_0 := \min_{x\in X}\left(F^2(x)\middle|F^1(x)=x_0\right)$ and $x_1 := \min_{x\in X}\left(F^1(x)\middle|F^2(x)=y_1\right)$ are well defined and represent the minimal values of *convex optimization problems.* We put

$$z_0 := (x_0, y_0) \in \mathbb{R}^2, \tag{1}$$
$$z_1 := (x_1, y_1) \in \mathbb{R}^2. \tag{2}$$

Remark 1 It is easy to see that $z_0, z_1 \in \mathrm{MIN}_{\mathbb{R}^2_+}(F(X))$ and $x_0 \le x_1,\ y_0 \ge y_1$. Therefore, if $z_0 \neq z_1$, then obviously $x_0 < x_1$ and $y_0 > y_1$. On the other hand, if $z_0 = z_1$, then $\mathrm{MIN}_{\mathbb{R}^2_+}(F(X)) = \{z_0\}$ and z_0 is the unique outcome optimal solution of problem (P), i.e. z_0 is such that $\min\{f(z)|z \in \mathrm{MIN}_{\mathbb{R}^2_+}(F(X))\} = f(z_0)$.

So, we will suppose in the sequel that $z_0 \neq z_1$, hence

$$x_0 < x_1 \ \text{ and } \ y_0 > y_1. \tag{3}$$

The following proposition is useful in the presentation of the algorithm, and it is inspired by the Pascoletti-Serafini Scalarization (see, e.g., [18]). Its proof is a direct consequence of Proposition 3.

Proposition 2 *Let $t \in]0,1[$, and let $z_t \in F(X)+\mathbb{R}^2_+$ be given by*

$$z_t := (1-t)z_0 + tz_1. \tag{4}$$

Now let us consider a fixed vector $r \in -int(\mathbb{R}^2_+)$. Then the problem

$$\sup\left(\lambda \ge 0 \middle| z_t + \lambda r \in F(X) + \mathbb{R}^2_+\right)$$

has a unique optimal solution $\lambda^* \in \mathbb{R}$, and z_t^* defined by

$$z_t^* := z_t + \lambda^* r \tag{5}$$

belongs to $MIN_{\mathbb{R}^2_+}(F(X))$.

Since z_0 and z_1 belongs to $\mathrm{MIN}_{\mathbb{R}^2_+}(F(X))$, we set $z_0^* := z_0$ and $z_1^* := z_1$.

Definition 1 Let $t \in]0,1[$. z_t^* defined by (5) will be called the (unique) outcome solution associated with t.

Remark 2 To find the outcome solution z_t^* associated with a given $t \in]0,1[$, we have to solve the following convex optimization problem

$$(CP_t) \qquad \sup_{(\lambda,x)}\left(\lambda \ge 0 \middle| F(x) - z_t - \lambda r \le 0,\ g(x) \le 0\right),$$

and its optimal value λ^* is such that the set $(z_t + \lambda^* r - \mathbb{R}^2_+) \cap F(X)$ is reduced to a singleton, namely $\{z_t^*\} = (z_t + \lambda^* r - \mathbb{R}^2_+) \cap F(X)$.

Definition 2 Let $t_1, t_2 \in [0,1]$ be such that $t_1 < t_2$. We denote by $Simp(t_1,t_2)$ the triangle given by the convex hull of $(z_{t_1}^*, z_{t_2}^*, z^I)$, where $z_{t_1}^* = (x_{t_1}^*, y_{t_1}^*)$ and $z_{t_2}^* = (x_{t_2}^*, y_{t_2}^*)$ are defined by (5), and $z^I := (x_{t_1}^*, y_{t_2}^*)$.

Remark 3 By (3), (4), (5), and by the fact that $z^*_{t_1}, z^*_{t_2} \in \mathrm{MIN}_{\mathbb{R}^2_+}(F(X))$ and $r \in -\mathrm{int}(\mathbb{R}^2_+)$, we obtain easily that $x^*_{t_1} < x^*_{t_2}$ and $y^*_{t_1} > y^*_{t_2}$. Thus, the convex hull of $(z^*_{t_1}, z^*_{t_2}, z^I)$ is a right triangle with nonempty interior.

ALGORITHM :
Step $k = 0$ (initializing)
Solve the following convex optimization problems:

$$x_0 = \min\left(F^1(x)\middle| g(x) \le 0\right), \qquad y_0 = \min\left(F^2(x)\middle| F^1(x) = x_0,\ g(x) \le 0\right),$$

$$y_1 = \min\left(F^2(x)\middle| g(x) \le 0\right), \qquad x_1 = \min\left(F^1(x)\middle| F^2(x) = y_1,\ g(x) \le 0\right),$$

and put $z_0 = (x_0, y_0)$ and $z_1 = (x_1, y_1)$.
Choose $\varepsilon > 0$ (the error bound) and put:

- $t^{(0)}_0 = 0, t^{(0)}_1 = 1.$
- $D^{(0)} = \{z^*_0, z^*_1\} = \{z_0, z_1\}.$
- $S^{(0)} = S^{(0)}_1 = Simp(0,1).$

Solve $d^{(0)} = \min(f(z_0), f(z_1))$ the best current upper bound.
Solve $s^{(0)} = \min_{z \in S^{(0)}} f(z)$ the best current lower bound.
IF $d^{(0)} - s^{(0)} < \varepsilon$,
THEN Stop ($U \in [s^{(0)}, d^{(0)}]$),
ELSE choose $r \in -\mathrm{int}(\mathbb{R}^2_+)$ and $k = 1$.

Step k ($k \ge 1$)

- Find $j^* \in \{1, \ldots, k\}$ such that $s^{(k-1)} = \min\left(f(z)\middle| z \in Simp(t^{(k-1)}_{j^*-1}, t^{(k-1)}_{j^*})\right)$, and put

$$\begin{cases} t^{(k)}_j = t^{(k-1)}_j,\ j = 0, \ldots, j^* - 1 \\ t^{(k)}_{j^*} = \frac{1}{2}(t^{(k-1)}_{j^*-1} + t^{(k-1)}_{j^*}) \\ t^{(k)}_j = t^{(k-1)}_{j-1},\ j = j^* + 1, \ldots, k+1 \end{cases}. \tag{6}$$

- Determine $z^*_{t^{(k)}_{j^*}}$ the outcome solution associated with $t^{(k)}_{j^*}$, i.e. solve the convex optimization problem

$$(CP_{t^{(k)}_{j^*}}) \qquad \sup_{(\lambda,x)} \left(\lambda \ge 0 \middle| g(x) \le 0, F(x) - \lambda r - z_{t^{(k)}_{j^*}} \le 0\right).$$

Denote by λ^* its optimal value, and set $z^*_{t^{(k)}_{j^*}} = z_{t^{(k)}_{j^*}} + \lambda^* r$ and

$$D^{(k)} = D^{(k-1)} \cup \{z^*_{t^{(k)}_{j^*}}\}. \tag{7}$$

Put $d^{(k)} = \min(d^{(k-1)}, f(z^*_{t^{(k)}_{j^*}}))$ the best current upper bound.

– For all $j = 1,\dots,k+1$, denote $S_j^{(k)} = Simp(t_{j-1}^{(k)}, t_j^{(k)})$ and set

$$S^{(k)} = \bigcup_{j=1}^{k+1} S_j^{(k)}. \tag{8}$$

Solve $s_{j^*}^{(k)} = \min_{z\in S_{j^*}^{(k)}} f(z)$ and $s_{j^*+1}^{(k)} = \min_{z\in S_{j^*+1}^{(k)}} f(z)$.
If $k = 1$, then put $s^{(1)} = \min_{j=1,2} s_j^{(1)}$ the best current lower bound.

If $k \geq 2$, put $s^{(k)} = \min_{j=1,\dots,k+1} s_j^{(k)}$ the best current lower bound, where
$s_j^{(k)} = s_{j-1}^{(k-1)} \quad \forall j = 3,\dots,k+1 \quad$ if $j^* = 1$,
$s_j^{(k)} = s_j^{(k-1)} \quad \forall j = 1,\dots,k-1 \quad$ if $j^* = k$,
$s_j^{(k)} = s_j^{(k-1)} \quad \forall j = 1,\dots,j^*-1$ and $s_j^{(k)} = s_{j-1}^{(k-1)} \forall j = j^*+2,\dots,k+1$ if $j^* \in \{2,\dots,k-1\}$.

IF $d^{(k)} - s^{(k)} < \varepsilon$,

THEN Stop ($U \in [s^{(k)}, d^{(k)}]$),

ELSE $k = k+1$ and go to step k.

Remark 4 Note that during step k ($k \geq 1$), the j^{*th} triangle $S_{j^*}^{(k-1)} = Simp(t_{j^*-1}^{(k-1)}, t_{j^*}^{(k-1)})$ is reduced by two triangles $S_{j^*}^{(k)} = Simp(t_{j^*-1}^{(k)}, t_{j^*}^{(k)})$ and $S_{j^*+1}^{(k)} = Simp(t_{j^*}^{(k)}, t_{j^*+1}^{(k)})$, and for all $j = 1,\dots k, j \neq j^*$, the triangle $S_j^{(k-1)}$ is unchanged. Thus, it is not difficult to see that

$$S^{(k)} = (\bigcup_{j\neq j^*} S_j^{(k-1)}) \cup (S_{j^*}^{(k)} \cup S_{j^*+1}^{(k)}). \tag{9}$$

Remark 5 If we suppose in addition f quasi-concave, the algorithm will perform better because its minimum value on a triangle is reached on a vertex. So it may happen that our algorithm returns at a step k the exact optimal value of problem (P). For example, if we want to optimize one objective of the (BOP) over its outcome Pareto set, the algorithm returns the exact optimal value at the initializing step, which can be useful for a decision maker who wants to know the range (maximum and minimum value) of one or both objectives over the Pareto set.

Now we are going to prove that for a given error bound, our algorithm terminates in a finite number of steps. We begin with some Lemmas.

Lemma 2 *Let Z be a real normed space, $C \subset Z$ a convex pointed cone with nonempty interior, and T a nonempty subset of Z. Then the set $w\text{-}MIN_C(T) := \{z \in T \mid \nexists \bar{z} \in T : \bar{z} \in z - int(C)\}$ satisfies*

$$w\text{-}MIN_C(T) = \partial(T+C) \cap T. \tag{10}$$

Proof Let $z \in w\text{-MIN}_C(T)$. If $z \notin \partial(T+C) \cap T$, since $z \in T$, and $T \subset T+C$, we must have $z + \alpha\mathbb{B} \subset T+C$ for some $\alpha > 0$ (where $\mathbb{B}$ stands for the open unit ball). Let $c \in \text{int}(C)$

with $\|c\| < \alpha$. Thus $z-c \in T+C$, i.e. $z-c = t+c'$ for some $t \in T$ and $c' \in C$. Finally, $t = z-c-c' \in z-\text{int}(C)$ which contradicts the choice of z.

Conversely, let $z \in \partial(T+C) \cap T$. If $z \notin w\text{-MIN}_C(T)$, then there is some $t \in T$ such that $z \in t+\text{int}(C) \subset \text{int}(T+C)$. Therefore $z \notin \partial(T+C)$ which is a contradiction, so (10) is proven. □

Lemma 3 *Let in the previous Lemma $Z = \mathbb{R}^2$, $C = \mathbb{R}^2_+$ and T be a compact nonempty subset of Z such that $T+C$ is convex. Let $z_1 = (x_1,y_1)$, $z_2 = (x_2,y_2) \in MIN_C(T)$, such that $x_1 < x_2$ (hence $y_1 > y_2$). Then for any $z = (x,y) \in \partial(T+C)$ with $x_1 < x < x_2$, we have*

$$(x,y) \in MIN_C(T), \tag{11}$$

and

$$y \leq y_1 + \frac{y_2-y_1}{x_2-x_1}(x-x_1). \tag{12}$$

Proof Since T is compact and C is closed, it is well known that $T+C$ is closed. Thus, using (10) with $T+C$ instead of T, since $C+C=C$, we obtain that

$$w\text{-MIN}_C(T+C) = \partial(T+C).$$

Therefore $(x,y) \in w\text{-MIN}_C(T+C)$.

Let us prove that $(x,y) \in \text{MIN}_C(T+C)$.

Arguing by contradiction, since $(x,y) \in w\text{-MIN}_C(T+C)$, we can find $(x',y') \in T+C$ such that

(i) $x' < x$, $y' = y$

or

(ii) $x' = x$, $y' < y$.

Since $T+C$ is convex, for all $\varepsilon \in [0,1]$ we have $((1-\varepsilon)(x',y') + \varepsilon(x_2,y_2)) \in T+C$ and $((1-\varepsilon)(x',y') + \varepsilon(x_1,y_1)) \in T+C$.

Assume that (i) holds. For sufficiently small $\varepsilon \in [0,1[$ we have $(1-\varepsilon)x' + \varepsilon x_2 < x$, and since $y > y_2$, $(1-\varepsilon)y' + \varepsilon y_2 = (1-\varepsilon)y + \varepsilon y_2 < y$ for all $\varepsilon \in]0,1]$. This contradicts the fact that $(x,y) \in w\text{-MIN}_C(T+C)$.

Suppose that (ii) holds. We have $(1-\varepsilon)x' + \varepsilon x_1 = (1-\varepsilon)x + \varepsilon x_1 < x$ for all $\varepsilon \in]0,1]$, and $(1-\varepsilon)y' + \varepsilon y_1 < y$ for sufficiently small $\varepsilon \in [0,1[$. This contradicts the fact that $(x,y) \in w\text{-MIN}_C(T)$. Therefore, we have shown that $(x,y) \in \text{MIN}_C(T+C)$.

By [19, Proposition 2.1], $\text{MIN}_C(T) = \text{MIN}_C(T+C)$, so (11) is proven.

To prove (12) we reason again by contradiction. So, if (12) is not satisfied, then we can find $h > 0$ such that $x-h > x_1$ and $y > y_1 + \frac{y_2-y_1}{x_2-x_1}(x-h-x_1)$. But $(x-h, y_1 + \frac{y_2-y_1}{x_2-x_1}(x-h-x_1)) \in [(x_1,y_1),(x_2,y_2)] \subset (T+C)$, and $(x-h, y_1 + \frac{y_2-y_1}{x_2-x_1}(x-h-x_1)) \in (x,y) - \text{int}(C)$, which contradicts the fact that $(x,y) \in \text{MIN}_C(T)$. □

Proposition 3 *Let the hypotheses of Lemma 3 hold. Let $(\hat{x},\hat{y}) = (1-\hat{t})(x_1,y_1) + \hat{t}(x_2,y_2)$, where $\hat{t} \in]0,1[$ is given, and let $r = (r_1,r_2) \in -int(\mathbb{R}^2_+)$. Consider*

$$\lambda^* = \sup\{\lambda \geq 0 |\ (\hat{x},\hat{y}) + \lambda r \in T+C\}.$$

Then $\lambda^* \in \mathbb{R}$,

$$(x^*, y^*) := (\hat{x}, \hat{y}) + \lambda^* r \in MIN_C(T), \tag{13}$$

and

$$x_1 < x^* < x_2, \quad y^* \le y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x^* - x_1), \tag{14}$$

in particular $y^* \in]y_2, y_1[$.

Proof Let us show first that the set $[(\hat{x}, \hat{y}) - C] \cap (T + C)$ is nonempty and compact. Since $T + C$ is convex we have that $(\hat{x}, \hat{y}) \in T + C$, and obviously $(\hat{x}, \hat{y}) \in (\hat{x}, \hat{y}) - C$.

Note that the Euclidean norm $\|\cdot\|$ is nondecreasing on $C = \mathbb{R}^2_+$.

Let (u_n) be a sequence in $[(\hat{x}, \hat{y}) - C] \cap (T + C)$. Thus, for each $n \in \mathbb{N}$, $u_n = (\hat{x}, \hat{y}) - c_n = t_n + c'_n$ for some $c_n, c'_n \in C$, $t_n \in T$. Since T is compact, there exists a convergent subsequence $(t_{n'})$, $t_{n'} \to t \in T$. Thus the sequence $(c_{n'} + c'_{n'})$ converges to $(\hat{x}, \hat{y}) - t$, hence it is bounded. Since $\|c_{n'}\| \le \|c_{n'} + c'_{n'}\|$, we obtain that the sequence $(c_{n'})$ is bounded, hence there is a subsequence $(c_{n''})$ which converges to an element $c \in C$ (because C is closed). Then, the sequence $(c'_{n''})$ converges to an element $c' \in C$. Finally, the subsequence $(u_{n''})$ converges to $(\hat{x}, \hat{y}) - c = t + c'$, so $[(\hat{x}, \hat{y}) - C] \cap (T + C)$ is compact.

Since for any $\lambda \ge 0$ we have $(\hat{x}, \hat{y}) + \lambda r \in [(\hat{x}, \hat{y}) - C]$, then $[(\hat{x}, \hat{y}) + \mathbb{R}_+ r] \cap (T + C)] \subset [(\hat{x}, \hat{y}) - C] \cap (T + C)$. Hence there exists a positive real M such that for any $\lambda \ge 0$ satisfying $(\hat{x}, \hat{y}) + \lambda r \in T + C$, we have

$$\|(\hat{x}, \hat{y}) + \lambda r\| \le M.$$

But, for such a λ we have

$$\lambda = \frac{1}{\|r\|}\|\lambda r\| \le \frac{1}{\|r\|}\Big(M + \|(\hat{x}, \hat{y})\|\Big).$$

In other words, the set $A := \{\lambda \ge 0 |\ (\hat{x}, \hat{y}) + \lambda r \in T + C\}$ (which is nonempty because $0 \in A$) is bounded from above, so $\lambda^* \in \mathbb{R}$. Then there exists a sequence (λ_n) in A such that $\lambda_n \to \lambda^*$. Therefore the sequence $\big((\hat{x}, \hat{y}) + \lambda_n r\big)$ converges to $(x^*, y^*) := (\hat{x}, \hat{y}) + \lambda^* r \in T + C$ because $T + C$ is closed (T is compact and C is closed). If $(x^*, y^*) \in \text{int}(T + C)$, then $(x^*, y^*) + \varepsilon\mathbb{B} \subset T + C$ for some $\varepsilon > 0$. So $(x^*, y^*) + \frac{\varepsilon}{2\|r\|} r = (\hat{x}, \hat{y}) + (\lambda^* + \frac{\varepsilon}{2\|r\|}) r \in T + C$ which contradicts the definition of λ^*. Thus $(x^*, y^*) \in \partial(T + C)$. Hence, according to Lemma 3, in order to finish the proof it is sufficient to show that $x^* \in]x_1, x_2[$.

We have $x^* = (1 - \hat{t})x_1 + \hat{t}x_2 + \lambda^* r_1 \le (1 - \hat{t})x_1 + \hat{t}x_2 < x_2$ (since $x_1 < x_2$).

Also, since $y_2 < y_1$, we have $y^* = (1 - \hat{t})y_1 + \hat{t}y_2 + \lambda^* r_2 < y_1$. Since $(x_1, y_1) \in \text{MIN}_C(T) = \text{MIN}_C(T + C)$, we cannot have $x^* \le x_1$, hence $x^* \in]x_1, x_2[$. □

Remark 6 As a direct consequence of the previous Proposition, for any t_1, t, t_2 with $0 \le t_1 < t < t_2 \le 1$, since $Simp(t_1, t_2) = conv((x^*_{t_1}, y^*_{t_1}), (x^*_{t_2}, y^*_{t_2}), (x^*_{t_1}, y^*_{t_2}))$ where $(x^*_{t_i}, y^*_{t_i}) = z^*_{t_i}$ is the outcome solution associated with t_i ($i = 1, 2$) , (14) immediately implies that the outcome solution $z^*_t = (x^*_t, y^*_t)$ associated with t satisfies

$$z^*_t \in Simp(t_1, t_2). \tag{15}$$

The following is obvious.

Lemma 4 *Let* $z_1, z_2, z_3 \in \mathbb{R}^2$ *and let* $A = conv(z_1, z_2, z_3)$ *the convex hull of* z_1, z_2, z_3. *Then, for all* $a \in A$, $conv(z_1, z_2, a) \subset A$.

Recall that during the k^{th} step of our algorithm, the triangle $S_{j^*}^{(k-1)} = Simp(t_{j^*-1}^{(k-1)}, t_{j^*}^{(k-1)})$ is reduced by two triangles $S_{j^*}^{(k)} = Simp(t_{j^*-1}^{(k)}, t_{j^*}^{(k)})$ and $S_{j^*+1}^{(k)} = Simp(t_{j^*}^{(k)}, t_{j^*+1}^{(k)})$. By (6), we know that $t_{j^*-1}^{(k-1)} = t_{j^*-1}^{(k)}$ and $t_{j^*}^{(k-1)} = t_{j^*+1}^{(k)}$. Thus, Lemma 4 and (15) imply

$$S_{j^*}^{(k)} \cup S_{j^*+1}^{(k)} \subset S_{j^*}^{(k-1)}.$$

Hence, by (7) and (9), the proof of the following is straightforward.

Theorem 1 *For all $k \geq 0$, the compact sets $S^{(k)}$ and $D^{(k)}$ are such that*

$$D^{(k)} \subset D^{(k+1)} \subset MIN_{\mathbb{R}^2_+}(F(X)) \subset S^{(k+1)} \subset S^{(k)}.$$

Lemma 5 *Each time a triangle S is reduced by two triangles S_1 and S_2, we have*

$$\max(\mathrm{diam}(S_1), \mathrm{diam}(S_2)) \leq c(r) \cdot \mathrm{diam}(S),$$

where for any subset $A \subset \mathbb{R}^2$, $\mathrm{diam}(A) := \sup_{x,y \in A} \|x - y\| \in [0, +\infty]$ stands for the diameter of the set A, and the real constant $c(r)$ depends only on $r \in -int(\mathbb{R}^2_+)$ and satisfies $c(r) < 1$.

Proof Without loss of generality, consider a triangle S given by $\mathrm{conv}(O, A, B)$, where the coordinates of O (resp. A, B) are $O(0,0)$ (resp. $A(0,y)$, $B(x,0)$ with $x \geq y > 0$). Let $\delta = diam(S)$, i.e. δ is the distance between A and B. Hence $x = \delta\cos(\theta)$ and $y = \delta\sin(\theta)$. Since $x \geq y, \theta \in]0, \frac{\pi}{4}]$.

Let $\alpha \in]0, \frac{\pi}{2}[$ be such that $r = (-\cos\alpha, -\sin\alpha)$. We put $T = \min(\frac{\delta\cos(\theta)}{2\cos(\alpha)}, \frac{\delta\sin(\theta)}{2\sin(\alpha)})$. Hence the point $P(\frac{x}{2} - t\cos(\alpha), \frac{y}{2} - t\sin(\alpha))$ belongs to S iff $t \in [0, T]$. Let $\gamma(t)$ be the distance between P and A, and $\beta(t)$ the distance between P and B, and let us show that $\max\limits_{t \in [0,T]}(\gamma(t)^2, \beta(t)^2) \leq c(r)^2\delta^2$, where $c(r) < 1$.

If $0 < \alpha < \theta \leq \frac{\pi}{4}$, $T = \frac{\delta\cos(\theta)}{2\cos(\alpha)}$. Then $\max\limits_{t \in [0,T]}(\gamma(t)^2, \beta(t)^2) = \beta(T)^2 \leq \frac{\delta^2}{2}(1 + \cos(2\alpha))$.

If $0 < \theta \leq \alpha < \frac{\pi}{2}$, $T = \frac{\delta\sin(\theta)}{2\sin(\alpha)}$. Then $\max\limits_{t \in [0,T]}(\gamma(t)^2, \beta(t)^2) = \beta(T)^2 \leq \frac{\delta^2}{2}(1 + \cos(\alpha))$ if $\alpha + \theta \leq \frac{\pi}{2}$, and $\max\limits_{t \in [0,T]}(\gamma(t)^2, \beta(t)^2) = \gamma(T)^2 \leq \frac{\delta^2}{2}(1 - \cos(\alpha + \frac{\pi}{4}))$ if $\frac{\pi}{2} < \alpha + \theta < \frac{3\pi}{4}$.

Note that for $\alpha > \frac{3\pi}{8}$ we have $\alpha \geq \theta$ and $\frac{\pi}{2} < \alpha + \theta < \frac{3\pi}{4}$, and for $\alpha \leq \frac{3\pi}{8}$ we have $\cos(2\alpha) < \cos(\alpha)$.

Finally,

$$c(r) = \begin{cases} \frac{1}{2}(1 + \cos(\alpha)) \text{ if } 0 < \alpha \leq \frac{3\pi}{8} \\ \frac{1}{2}(1 - \cos(\alpha + \frac{\pi}{4})) \text{ if } \frac{3\pi}{8} < \alpha < \frac{\pi}{2} \end{cases}$$

□

is such that $\max(\gamma(T), \beta(T)) \leq c(r)\delta$ and $c(r) < 1$.

Theorem 2 *The algorithm terminates in a finite number of steps.*

Proof Suppose by contradiction that the algorithm does not terminate. Let us show that the sequence $(S^{(k)})_{k \geq 0}$ has a subsequence $(S^{(\phi(k))})_{k \geq 0}$ such that one can find a sequence of index $(j_k)_{k \geq 0}$, which satisfies $\bigcap_{k \geq 0} S_{j_k}^{(\phi(k))} = \{z^*\}$, where z^* is an outcome optimal solution for problem (P), i.e. $z^* \in \mathrm{MIN}_{\mathbb{R}^2_+}(F(X))$ and $U = \min\Big(f(z)\Big| z \in \mathrm{MIN}_{\mathbb{R}^2_+}(F(X))\Big) = f(z^*)$.

Since the algorithm does not terminate, then at each step $k \geq 0$ there exists an index $j \in \{1,\ldots,k+1\}$ such that the set

$$\left\{\bar{k} > k \,\middle|\, \exists \bar{j} \in \{1,\ldots,\bar{k}\} : S_j^{(k)} = S_{\bar{j}}^{(\bar{k})} \cup S_{\bar{j}+1}^{(\bar{k})},\ \min_{z \in S^{(k)}} f(z) = \min_{z \in S_j^{(k)}} f(z)\right\} \tag{16}$$

is infinite.

Thus there exists a strictly increasing mapping $\phi : \mathbb{N} \to \mathbb{N}$ and a sequence of index $(j_k)_{k\geq 0}$ ($j_k \in \{1,\ldots,\phi(k)+1\}$) such that $s^{(\phi(k))} = \min_{z \in S^{(\phi(k))}} f(z) = \min_{z \in S_{j_k}^{(\phi(k))}} f(z)$,

$$\cdots \subset S_{j_{k+1}}^{(\phi(k+1))} \subset S_{j_k}^{(\phi(k))} \subset \cdots \subset S_{j_0}^{(\phi(0))} = S_1^{(0)}, \tag{17}$$

and, according to Lemma 5,

$$\mathrm{diam}(S_{j_{k+1}}^{(\phi(k+1))}) \leq c(r)\mathrm{diam}(S_{j_k}^{(\phi(k))}). \tag{18}$$

Therefore

$$\lim_{k\to+\infty} \mathrm{diam}(S_{j_k}^{(\phi(k))}) = 0. \tag{19}$$

Since the sets $S_{j_k}^{(\phi(k))}$ are closed for all k, we have

$$\bigcap_{k\geq 0} S_{j_k}^{(\phi(k))} = \{\bar{z}\} \tag{20}$$

for some $\bar{z} \in \mathbb{R}^2$.

Moreover, by (6), for each $k \geq 0$ we have $t_{j_{k+1}}^{(\phi(k+1))} - t_{j_{k+1}-1}^{(\phi(k+1))} = \frac{1}{2}(t_{j_k}^{(\phi(k))} - t_{j_k-1}^{(\phi(k))})$, and then

$$\bigcap_{k\geq 0} [t_{j_k-1}^{(\phi(k))}, t_{j_k}^{(\phi(k))}] = \{t^*\},\ t^* \in [0,1].$$

Let z^* be the outcome solution associated with t^*, i.e. $z^* = z_{t^*} + \lambda^* r \in \mathrm{MIN}_{\mathbb{R}^2_+}(F(X))$ for a $\lambda^* \geq 0$.
Since $\forall k \geq 0$, $t^* \in [t_{j_k-1}^{(\phi(k))}, t_{j_k}^{(\phi(k))}]$, (15) implies that $z^* \in S_{j_k}^{(\phi(k))} = Simp(t_{j_k-1}^{(\phi(k))}, t_{j_k}^{(\phi(k))})$. Hence

$$z^* \in \bigcap_{k\geq 0} S_{j_k}^{(\phi(k))}.$$

Thus, by (20),

$$z^* = \bar{z},$$

hence

$$\bigcap_{k\geq 0} S_{j_k}^{(\phi(k))} = \{z^*\}. \tag{21}$$

Now, let us show that $U = \min\left(f(z) \,\middle|\, z \in \mathrm{MIN}_{\mathbb{R}^2_+}(F(X))\right) = f(z^*)$.

On one hand, $\mathrm{MIN}_{\mathbb{R}^2_+}(F(X)) \subset S^{(\phi(k))}$ (all k), so $s^{(\phi(k))} = \min_{z \in S_{j_k}^{(\phi(k))}} f(z) \leq U$. Since the sequence of real numbers $(s^{(\phi(k))})_k$ is increasing and bounded by U, it admits a limit s, and $s \leq U$.

Now, for each $k \geq 0$, according to Weierstrass' Theorem we consider $z_k \in S_{j_k}^{(\phi(k))}$ such that $f(z_k) = s^{(\phi(k))}$.

Since $z^* \in S_{j_k}^{(\phi(k))}$, we have $\|z^* - z_k\| \leq \operatorname{diam}(S_{j_k}^{(\phi(k))}) \to 0$ as $k \to +\infty$.

But f is continuous, hence $f(z^*) = \lim_{k\to\infty} f(z_k) = \lim_{k\to\infty} s^{(\phi(k))} = s$. Therefore $U \geq f(z^*)$.

On the other hand, $z^* \in \mathrm{MIN}_{\mathbb{R}^2_+}(F(X))$, so $U \leq f(z^*)$.
Thus $U = f(z^*)$.

For all k we have $d^{(\phi(k))} \leq f(z^*_{t_{j_k}^{(\phi(k))}})$, hence

$$0 \leq d^{(\phi(k))} - s^{(\phi(k))} \leq f(z^*_{t_{j_k}^{(\phi(k))}}) - f(z_k).$$

Moreover,

$$\|z^*_{t_{j_k}^{(\phi(k))}} - z_k\| \leq \operatorname{diam}(S_{j_k}^{(\phi(k))}) \to 0 \quad (\text{as } k \to +\infty).$$

Since f is uniformly continuous on the compact $F(X)$, we obtain that $d^{(\phi(k))} - s^{(\phi(k))} \to 0$ as $k \to +\infty$.

Therefore, for any given $\varepsilon > 0$, the stopping criterion is met in a finite number of steps contradicting our ab absurdo hypothesis.

□

3 Application to the Stochastic Case

3.1 Problem Statement

The basic facts about stochastic optimization and sample average approximation method can be found in [34].

Let $\xi : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Xi, \mathcal{B}_\Xi, \mathbb{P}_\xi)$ be a fixed $\mathbb{R}^d$ valued random vector, where $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\Xi, \mathcal{B}_\Xi, \mathbb{P}_\xi)$ are probability spaces, $\Xi \subset \mathbb{R}^d$ is the support of measure $\mathbb{P}_\xi(\cdot) := \mathbb{P}(\xi^{-1}(\cdot))$ (i.e. Ξ is the smallest closed set such that $\xi^{-1}(\Xi) = \Omega$, hence $\mathbb{P}_\xi(\Xi) = 1$), and $\mathcal{B}_\Xi$ is the trace on Ξ of the $\mathbb{R}^d$ Borel σ-field.

Note that a function $h : \mathbb{R}^n \times \Xi \to \mathbb{R}$ is said to be a random function if for every fixed $x \in \mathbb{R}^n$, the function $\xi \mapsto h(x, \xi)$ is measurable. For a random function, we can define the corresponding expected value function $x \mapsto \mathbb{E}[h(x, \xi(\cdot))] = \mathbb{E}_\xi[h(x, \cdot)] := \int_\Xi h(x, \delta) d\mathbb{P}_\xi(\delta)$.

In this section, we deal with the **stochastic problem** (S) of minimizing a real valued function f over the Pareto outcome set associated with a **Stochastic** bi-objective Optimization Problem $(SBOP)$, whose objectives are expectations of random functions. That is, to say

$$(S) \qquad \min\Big(f(\mathbb{E}_\xi[F(x,\cdot)])\Big| x \in \mathrm{ARGMIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])\Big),$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ is a deterministic mapping, $F = (F^1, F^2) : \mathbb{R}^n \times \Xi \to \mathbb{R}^2$ is the vector random objective of the following

$$(SBOP) \qquad \mathrm{MIN}_{x\in X}\mathbb{E}_\xi\Big[F(x,\cdot)\Big] = \mathbb{E}_\xi\Big[\Big(F^1(x,\cdot), F^2(x,\cdot)\Big)\Big],$$

whose feasible set $X \subset \mathbb{R}^n$ is deterministic and explicitly given by

$$X := \Big\{ x \in \mathbb{R}^n \Big| g(x) \leq 0 \Big\}, \tag{22}$$

with $g : \mathbb{R}^n \to \mathbb{R}^p$.

In most real cases, we cannot compute directly the expectation of a random function. Hence problem (S), which will be called the **true problem**, becomes impossible to solve. However, under some hypothesis, we can apply the results of paper [13] to give a sequence of Sample Average Approximation (SAA-N) problems that converges to the true problem (S), in the sense where the sequence of SAA-N optimal values converges almost surely to the **true** optimal value as the sample size N goes to infinity.

Then we will be concerned with the convergence rate and error estimates.

To be more precise, we consider an independent identically distributed (i.i.d.) sequence $(\xi_k)_{k\in\mathbb{N}^*}$ ($\mathbb{N}^*$ denotes the set of positive integers) 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.

We set $\tilde{\Xi} = \prod_{N=1}^{\infty} \Xi$ and $\tilde{\mathcal{B}} = \otimes_{N=1}^{\infty} \mathcal{B}_\Xi$ the smallest σ-algebra on $\tilde{\Xi}$ generated by all sets of the form $B_1 \times B_2 \times \cdots \times B_N \times \Xi \times \Xi \times \ldots$, $B_k \in \mathcal{B}_\Xi$, $k = 1, \ldots, N$, $N = 1, 2, \ldots$.

Then ([26, Theorem 10.4]), there exists a unique probability $\tilde{\mathbb{P}}_\xi$ on $(\tilde{\Xi}, \tilde{\mathcal{B}})$ such that $\tilde{\mathbb{P}}_\xi(B_1 \times B_2 \times \cdots \times B_N \times \Xi \times \Xi \times \ldots) = \prod_{k=1}^{N} \mathbb{P}_\xi(B_k)$ for all $N = 1, 2, \ldots$, and $B_k \in \mathcal{B}_\Xi$ for all $k = 1, \ldots, N$.

For each $N \in \mathbb{N}^*$, $x \in \mathbb{R}^n$, let $\hat{F}_N^j(x, \cdot)$ $(j = 1, 2)$ be the $N-$Sample Average Approximation (SAA-N) function defined by

$$\begin{aligned} \hat{F}_N^j(x, \cdot) : (\tilde{\Xi}, \tilde{\mathcal{B}}, \tilde{\mathbb{P}}_\xi) &\to \mathbb{R} \\ \tilde{\xi} = (\xi_1, \xi_2, \ldots) &\mapsto \frac{1}{N} \sum_{k=1}^{N} F^j(x, \xi_k), \end{aligned} \tag{23}$$

where $F^1, F^2 : \mathbb{R}^n \times \Xi \to \mathbb{R}$ are the objectives of (SBOP).

Now, for each $N \in \mathbb{N}^*$ and $\tilde{\xi} \in \tilde{\Xi}$, we can consider the following **SAA-N problem**:

$$(S_N(\tilde{\xi})) \qquad \min \Big(f(\hat{F}_N(x, \tilde{\xi})) \Big| x \in \mathrm{ARGMIN}_{\mathbb{R}^2_+}(\hat{F}_N(X, \tilde{\xi})) \Big),$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ is the same as in the true problem (S), and the SAA-N function $x \mapsto \hat{F}_N(x, \tilde{\xi}) = \Big(\hat{F}_N^1(x, \tilde{\xi}), \hat{F}_N^2(x, \tilde{\xi}) \Big)$ come from the following SAA-N bi-objective Optimization Problem

$$(SAA-NBOP) \qquad \mathrm{MIN}_{x\in X} \hat{F}_N(x, \tilde{\xi}) = \frac{1}{N} \sum_{k=1}^{N} \Big(F^1(x, \xi_k), F^2(x, \xi_k) \Big),$$

whose feasible set X has been given in (22).

In this Section we need to assume the following assumptions:

$$(H_s)\begin{cases} X \text{ is a nonempty and bounded subset of } \mathbb{R}^n.\\ \forall i=1,\dots,p,\ x\mapsto g^i(x) \text{ is convex on } \mathbb{R}^n.\\ \text{the random process } \tilde{\xi}\in\tilde{\Xi} \text{ is i.i.d..}\\ \forall j=1,2,\ x\mapsto F^j(x,\xi) \text{ is finite valued and continuous on } X \text{ for almost every (a.e.) } \xi\in\Xi.\\ \forall j=1,2,\ x\mapsto F^j(x,\xi) \text{ is convex on } X \text{ for a.e. } \xi\in\Xi.\\ \forall j=1,2,\ F^j \text{ is dominated by an integrable and square integrable function } K^j\text{, i.e.,}\\ \forall (x,\xi)\in\mathbb{R}^n\times\Xi \quad |F^j(x,\xi)|\le K^j(\xi).\end{cases}$$

(H_1) f is continuous on $\mathbb{R}^2$, and there exists a real $\varepsilon_0>0$ such that f is L - Lipschitz continuous on the bounded subset $\mathbb{E}_\xi[F(X,\cdot)]+\varepsilon_0\mathbb{B}$, i.e., for all $y,z\in\mathbb{E}_\xi[F(X,\cdot)]+\varepsilon_0\mathbb{B}$ we have $|f(y)-f(z)|\le L\|y-z\|$, where $\mathbb{B}$ stands for the open unit ball.

(H_2) f is a nondecreasing function on $\mathbb{R}^2$. In other words for all $z,y\in\mathbb{R}^2$, if $y-z\in\mathbb{R}^2_+$, we have $f(z)\le f(y)$.

Note that under (H_s), $(SBOP)$ as well as the related $(SAA-NBOP)$ are convex.

Remark 7 Let $h:\mathbb{R}^2\to\mathbb{R}$ be a nondecreasing function defined on a right triangle S given by the convex hull of the points A,B,C, whose coordinates are respectively $A(u,v)$, $B(u,\bar{v})$, $C(\bar{u},v)$ with $\bar{u}>u$ and $\bar{v}>v$. Then

$$\min_{x\in S} h(x)=h(A).$$

3.2 Preliminaries

Proposition 4 *Under (H_s,H_1), problem (S) admits at least one solution.*

Proof Since $(SBOP)$ is convex, and $\mathbb{E}_\xi$ is linear, the set $\mathbb{E}_\xi[F(X,\cdot)]+\mathbb{R}^2_+$ is convex. Moreover, X is compact, hence $\mathbb{E}_\xi[F(X,\cdot)]$ is compact, therefore the Pareto outcome set $\mathrm{MIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])$ is compact by Lemma 1. Since f is continuous on $\mathbb{R}^2$, by Weierstrass' Theorem there exists a $z^*\in\mathrm{MIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])$ such that $f(z^*)=\min(f(z)|z\in\mathrm{MIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)]))$. Thus, there exists $x^*\in\mathrm{ARGMIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])$ such that $\mathbb{E}_\xi[F(x^*,\cdot)]=z^*$. □

Let $j=1,2$. In paper [13] it is shown there exists a set $\tilde{A}\subset\tilde{\Xi}$ such that $\tilde{\mathbb{P}}_\xi(\tilde{A})=0$, and for each fixed $\tilde{\xi}\in\tilde{\Xi}\setminus\tilde{A}$, $N\in\mathbb{N}^*$,

- $x\mapsto\hat{F}^j_N(x,\tilde{\xi})$ is continuous on X and finite valued,
- $x\mapsto\hat{F}^j_N(x,\tilde{\xi})$ is convex on X,
- $\hat{F}^j_N(\cdot,\tilde{\xi})$ converges to $\mathbb{E}_\xi\left[F^j(\cdot,\cdot)\right]$ uniformly on X.

This immediately implies the following Proposition.

Proposition 5 *Under (H_s,H_1), for each $N\ge 1$ and almost every $\tilde{\xi}\in\tilde{\Xi}$, problem $(S_N(\tilde{\xi}))$ admits at least one solution.*

Remark 8 Note that the two previous Propositions hold if we assume only (H_s) and f continuous on $\mathbb{R}^2$.

Definition 3 Let $A,B \subset \mathbb{R}^2$ be two nonempty and bounded sets.

- We denote by $d(x,B) := \inf_{x' \in B} \|x - x'\|$ the distance from $x \in \mathbb{R}^2$ to B (recall that $\|\cdot\|$ stands for the Euclidean norm).
- We denote $\mathbb{D}(A,B) := \sup_{x \in A} d(x,B)$ the deviation of the set A from the set B.
- 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.

Lemma 6 (see Lemma 3.2 in [13]) *Under (H_s), for almost all $\tilde{\xi}$ in $\tilde{\Xi}$, the Hausdorff-Pompeiu distance between $\hat{F}_N(X,\tilde{\xi})$ and $\mathbb{E}_\xi[F(X,\cdot)]$ tends to zero as N tends to infinity, i.e.,*

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

Let us denote by

$$V := \min\left(f(\mathbb{E}_\xi[F(x,\cdot)]) \,\middle|\, x \in \mathrm{ARGMIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])\right)$$

the optimal value of problem (S), and for each $N \in \mathbb{N}^*, \tilde{\xi} \in \tilde{\Xi}$,

$$V_N(\tilde{\xi}) := \min\left(f(\hat{F}_N(x,\tilde{\xi})) \,\middle|\, x \in \mathrm{ARGMIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi}))\right)$$

the optimal value of problem $(S_N(\tilde{\xi}))$.

The following Theorem is given in [13, Theorem 5.2].

Theorem 3 *Assume f continuous and (H_s) holds. Then the sequence of SAA-N optimal values $(V_N(\tilde{\xi}))_{N\geq 1}$ converges almost surely to the true optimal value V.*

3.3 Exponential convergence rate

In this subsection we will show that the convergence in Theorem 3 has exponential rate under (H_s, H_1, H_2) and the following hypotheses :

(E_1) There exist an integrable function $\kappa : \Xi \to \mathbb{R}_+$ and a constant $\gamma > 0$, such that

$$\|F(x,\xi) - F(y,\xi)\| \leq \kappa(\xi)\|x - y\|^\gamma$$

for all $x, y \in X$ and $\xi \in \Xi$.

(E_2) For every $x \in X$, the random variable $\xi \mapsto \|F(x,\xi)\|$ has a distribution supported on a bounded subset of $\mathbb{R}$, as well as the random variable κ.

Now, using Lemma 4.2 from [20] (see also [35, Theorem 5.1]) we obtain immediately the following.

Lemma 7 *Let assumptions (H_s, E_1, E_2) hold. Then, for any $\varepsilon > 0$, there exist positive reals $\alpha(\varepsilon)$ and $\beta(\varepsilon)$ independent of N, such that, for sufficiently large N, we have*

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, \sup_{x \in X} \|\hat{F}_N(x,\tilde{\xi}) - \mathbb{E}_\xi[F(x,\cdot)]\| \geq \varepsilon\right\}\right) \leq \alpha(\varepsilon)e^{-\beta(\varepsilon)N}. \tag{24}$$

As a consequence we have the following.

Lemma 8 *Let assumptions (H_s, E_1, E_2) hold. Then, for any $\varepsilon > 0$, there exist positive reals $\alpha(\varepsilon)$ and $\beta(\varepsilon)$ independent of N, such that, for sufficiently large N, we have*

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, \mathbb{H}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)]\right) \le \varepsilon\right\}\right) \ge 1 - \alpha(\varepsilon)e^{-\beta(\varepsilon)N}. \tag{25}$$

Proof Let N such that (24) holds. Using the definition of the Hausdorff-Pompeiu distance it is easy to see that

$$\left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, \sup_{x\in X} \|\hat{F}_N(x,\tilde{\xi}) - \mathbb{E}_\xi[F(x,\cdot)]\| < \varepsilon\right\} \subset \left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, \mathbb{H}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)]\right) \le \varepsilon\right\},$$

and the conclusion follows immediately. □

Definition 4 We say that a subset Y of $\mathbb{R}^2$ has the domination property if, and only if, for each $y \in Y$, there exists $y^* \in \mathrm{MIN}_{\mathbb{R}^2_+} Y$ such that $y^* \le y$.

Theorem 4 *Let assumptions $(H_s, H_1, H_2, E_1, E_2)$ hold. Then, for any $\varepsilon > 0$, there exist positive reals $\hat{\alpha}(\varepsilon)$ and $\hat{\beta}(\varepsilon)$ independent of N, such that*

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, |V_N(\tilde{\xi}) - V| \le \varepsilon\right\}\right) \ge 1 - \hat{\alpha}(\varepsilon)e^{-\hat{\beta}(\varepsilon)N} \tag{26}$$

for sufficiently large N.

Proof Without loss of generality we can suppose that $\varepsilon < \varepsilon_0$ and $L \ge 1$, where the Lipschitz constant L and the real ε_0 have been introduced in (H_1). Then, for any N we obviously have

$$\left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, \mathbb{H}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)]\right) \le \frac{\varepsilon}{L}\right\} \subset \left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, \hat{F}_N(X,\tilde{\xi}) \subset \mathbb{E}_\xi[F(X,\cdot)] + \varepsilon_0\mathbb{B}\right\}. \tag{27}$$

Let N such that (25) holds with $\frac{\varepsilon}{L}$ instead of ε, and let $\hat{\alpha}(\varepsilon) := \alpha(\frac{\varepsilon}{L})$, and $\hat{\beta}(\varepsilon) := \beta(\frac{\varepsilon}{L})$. Thus

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi} \,\middle|\, \mathbb{H}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)]\right) \le \frac{\varepsilon}{L}\right\}\right) \ge 1 - \hat{\alpha}(\varepsilon)e^{-\hat{\beta}(\varepsilon)N}. \tag{28}$$

Let $\tilde{\xi} \in \tilde{H}$, where $\tilde{H} := \{\tilde{\xi} \in \tilde{\Xi} | \mathbb{H}(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)]) \le \frac{\varepsilon}{L}\}$.

Now let $y_N^* \in \operatorname{argmin}\left(f(y) \,\middle|\, y \in \mathrm{MIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi}))\right)$ and $y^* \in \operatorname{argmin}\left(f(y) \,\middle|\, y \in \mathrm{MIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])\right)$ (Propositions 4 and 5 ensure the existence of y_N^* and y^*).

Suppose first that $f(y_N^*) \le f(y^*)$. Since $\tilde{\xi} \in \tilde{H}$ and $\mathrm{MIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi})) \subset \hat{F}_N(X,\tilde{\xi})$, by (28) there exist $y \in \mathbb{E}_\xi[F(X,\cdot)]$ and $h \in \mathbb{B}$ such that $y = y_N^* + \frac{\varepsilon}{L}h$.

The domination property holds for the set $\mathbb{E}_\xi[F(X,\cdot)]$ (see [31][Corollary 4.4 p.54]), hence there exists $\hat{y} \in \mathrm{MIN}_{\mathbb{R}^2_+} \mathbb{E}_\xi[F(X,\cdot)]$ such that $\hat{y} \le y$. Thus, by (27) and (H_1, H_2) we obtain easily

$$f(\hat{y}) \le f(y) = f(y_N^* + \frac{\varepsilon}{L}h) \le f(y_N^*) + \varepsilon \le f(y^*) + \varepsilon \le f(\hat{y}) + \varepsilon,$$

and therefore $|f(y_N^*) - f(y^*)| \le \varepsilon$.

The case $f(y^*) < f(y_N^*)$ can be handled in a similar way, i.e. there exist $y_N \in \hat{F}_N(X,\tilde{\xi})$ and $h \in \mathbb{B}$ such that $y_N = y^* + \frac{\varepsilon}{L}h$.

Moreover, there exits $\tilde{y} \in \mathrm{MIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi}))$ such that $\tilde{y} \leq y_N$ by the domination property. Finally, by (27) and (H_1, H_2) we obtain

$$f(\tilde{y}) \leq f(y_N) = f(y^* + \frac{\varepsilon}{L}h) \leq f(y^*) + \varepsilon < f(y_N^*) + \varepsilon \leq f(\tilde{y}) + \varepsilon,$$

hence $|f(y_N^*) - f(y^*)| \leq \varepsilon$.

In other words we have proven that $|V_N(\tilde{\xi}) - V| \leq \varepsilon$. Therefore $\tilde{H} \subset \left\{ \tilde{\xi} \in \tilde{\Xi} \middle| |V_N(\tilde{\xi}) - V| \leq \varepsilon \right\}$.

Hence, by (28) we obviously have

$$\tilde{\mathbb{P}}_\xi \left(\left\{ \tilde{\xi} \in \tilde{\Xi} \middle| |V_N(\tilde{\xi}) - V| \leq \varepsilon \right\} \right) \geq 1 - \hat{\alpha}(\varepsilon) e^{-\hat{\beta}(\varepsilon)N}.$$

□

3.4 A confidence interval

In this subsection we assume that it is possible to estimate a uniform upper bound of the standard deviation of the random bi-objective function $F(x,\xi)$. In other words, we assume that it is possible to estimate an upper bound of $\max_{j=1,2} \max_{x\in X} \sigma_j^2(x)$, where for each fixed $x \in X$, $\sigma_j(x)$ stands for the standard deviation of the real random function $F^j(x,\xi)$.

Then, for any $\varepsilon > 0$ and $p_0 \in]0,1[$ we propose a rank $N_0 = N_0(\varepsilon, p_0)$ which explicitly depends on data and the above upper bound such that the SAA-N^0 optimal value is an ε-approximation of V with probability greater than p_0. In other words, we find a confidence interval.

Proposition 6 *Let $\varepsilon > 0$ and $p_0 \in]0,1[$ be given. Under (H_s), for each $N \geq N^0$, we have*

$$\tilde{\mathbb{P}}_\xi \left(\tilde{\xi} \in \tilde{\Xi} \middle| \mathbb{H}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)] \right) \leq \varepsilon \right) \geq p_0, \tag{29}$$

where

$$N^0 = \max_{j=1,2} 2 \frac{\max_{x\in X} \sigma_j^2(x)}{\varepsilon^2(1-p_0)}. \tag{30}$$

Proof By the Uniform Law of Large Number (ULLN) [34, Theorem 7.48], we obtain immediately that

$$\tilde{\mathbb{P}}_\xi \left(\left\{ \tilde{\xi} \in \tilde{\Xi} \middle| \forall \varepsilon > 0, \exists N(\varepsilon,\tilde{\xi}) \in \mathbb{N}^* : \forall N \geq N(\varepsilon,\tilde{\xi}), \max_{1\leq j\leq 2} \sup_{x\in X} \left| \hat{F}_N^j(x,\tilde{\xi}) - \mathbb{E}_\xi \left[F^j(x,\cdot) \right] \right| \leq \frac{\varepsilon}{\sqrt{2}} \right\} \right) = 1.$$

So let $\tilde{\xi} \in \tilde{\Xi}$, $\varepsilon > 0$, and $N \geq 1$ be fixed such that

$$\max_{1\leq j\leq 2} \sup_{x\in S} \left| \hat{F}_N^j(x,\tilde{\xi}) - \mathbb{E}_\xi \left[F^j(x,\cdot) \right] \right| \leq \frac{\varepsilon}{\sqrt{2}}, \tag{31}$$

and let us show that $\mathbb{H}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)] \right) \leq \varepsilon$.

By contradiction, if $\mathbb{H}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)] \right) > \varepsilon$, we have $\mathbb{D}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)] \right) > \varepsilon$ or $\mathbb{D}\left(\mathbb{E}_\xi[F(X,\cdot)], \hat{F}_N(X,\tilde{\xi}) \right) > \varepsilon$.

Suppose first that $\mathbb{D}\Big(\hat{F}_N(X,\tilde{\xi}),\mathbb{E}_\xi[F(X,\cdot)]\Big) > \varepsilon$. By the definition of the deviation, since the sets $\hat{F}_N(X,\tilde{\xi})$ and $\mathbb{E}_\xi[F(X,\cdot)]$ are compact, one can find $x\in X$ such that $\|\hat{F}_N(x,\tilde{\xi})-\mathbb{E}_\xi[F(x,\cdot)]\| > \varepsilon$. Hence there exists at least a $j\in\{1,2\}$ such that $|\hat{F}^j_N(x,\tilde{\xi})-\mathbb{E}_\xi[F^j(x,\cdot)]| > \frac{\varepsilon}{\sqrt{2}}$, a contradiction to (31).

In the same way, if we assume $\mathbb{D}\Big(\mathbb{E}_\xi[F(X,\cdot)],\hat{F}_N(X,\tilde{\xi})\Big) > \varepsilon$, we obtain a contradiction.

Therefore, we have proven that $\mathbb{H}\Big(\hat{F}_N(X,\tilde{\xi}),\mathbb{E}_\xi[F(X,\cdot)]\Big) \leq \varepsilon$ whenever (31) holds.

Let $p_0 \in]0,1[$ be given. Now, we are going to show that (31) holds with probability greater than p_0 for every $N \geq N^0 = \max_{j=1,2} 2\frac{\max_{x\in X}\sigma_j^2(x)}{\varepsilon^2(1-p_0)}$.

Let $j=1,2$. Since the random process $\tilde{\xi}$ is i.i.d., it is well known that the sequence of real random variable $(F^j(x,\xi_k))_{k\geq 1}$ is i.i.d.. Moreover, the variance $\sigma_j^2(x)$ of $F^j(x,\xi)$ is finite valued for all $x\in X$ and almost every $\xi\in\Xi$, then the SAA-N real random function $\frac{1}{N}\sum_{k=1}^{N} F^j(x,\xi_k)$ has mean $\mathbb{E}_\xi\Big[F^j(x,\cdot)\Big]$ and variance $\frac{\sigma_j^2(x)}{N}$. Moreover, the mapping $x\mapsto \sigma_j^2(x)$ is continuous on the compact set X, hence there exists $x^j\in X$ such that $\sigma_j^2(x^j) = \max_{x\in X}\sigma_j^2(x)$.

By Bienaymé-Tchebychev Inequality ([26][Corollary 5.2 p.29]), for all fixed $x\in X$ we have

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi}\in\tilde{\Xi}\,\Big|\, |\hat{F}^j_N(x,\tilde{\xi})-\mathbb{E}_\xi[F^j(x,\cdot)]| \leq \frac{\varepsilon}{\sqrt{2}}\right\}\right) \geq 1-2\frac{\sigma_j^2(x)}{N\varepsilon^2},$$

and therefore, for all $N \geq N^0 = \max_{j=1,2} 2\frac{\sigma_j^2(x^j)}{\varepsilon^2(1-p_0)}$, we have

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi}\in\tilde{\Xi}\,\Big|\, \max_{1\leq j\leq 2}\sup_{x\in S}|\hat{F}^j_N(x,\tilde{\xi})-\mathbb{E}_\xi[F^j(x,\cdot)]| \leq \frac{\varepsilon}{\sqrt{2}}\right\}\right) \geq p_0.$$

□

Theorem 5 *Let $\varepsilon > 0$ and $p_0 \in]0,1[$ be given. Under (H_s,H_1,H_2), for each $N \geq N^0$, we have*

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi}\in\tilde{\Xi}\,\Big|\, |V_N(\tilde{\xi})-V| \leq \varepsilon\right\}\right) \geq p_0,$$

where $N^0 = \max_{j=1,2} 2L^2\frac{\max_{x\in X}\sigma_j^2(x)}{\varepsilon^2(1-p_0)}$ is as in Proposition 6 with ε replaced by $\frac{\varepsilon}{L}$, and L has been introduced in assumption (H_1).

Proof Let $N \geq N^0$ be fixed.

The rest of the proof is identical to the proof of Theorem 4 replacing relation (28) by relation (29), and the expression $1-\hat{\alpha}(\varepsilon)e^{-\hat{\beta}(\varepsilon)N}$ by p_0. □

Remark 9 Notice that it is possible to get rid of assumption (H_2) in Theorem 5 (and therefore in this subsection) if we succeed to estimate the Hausdorff-Pompeiu distance between the SAA-N Pareto set image and the true Pareto set image.

To be more precise, let $\varepsilon > 0$ and $p_0 \in]0,1[$ be given, and suppose that we can compute the rank $\tilde{N} = \tilde{N}(\varepsilon, p_0) \geq 1$ ($\tilde{N}$ exists by [13, Theorem 5.1] combined with (29)) such that for all $N \geq \tilde{N}$

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi} \middle| \mathbb{H}\left(\mathrm{MIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi})), \mathrm{MIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])\right) \leq \frac{\varepsilon}{L}\right\}\right) \geq \frac{p_0+1}{2} \tag{32}$$

and

$$\tilde{\mathbb{P}}_\xi\left(\tilde{\xi} \in \tilde{\Xi} \middle| \mathbb{H}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)]\right) \leq \frac{\varepsilon}{L}\right) \geq \frac{p_0+1}{2}. \tag{33}$$

We can suppose without loss of generality that $\varepsilon \leq \varepsilon_0$ and $L \geq 1$. Let us prove that for all $N \geq \tilde{N}$

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi} \middle| |V_N(\tilde{\xi}) - V| \leq \varepsilon\right\}\right) \geq p_0.$$

Indeed, let $N \geq \tilde{N}$ be fixed, and let

$$\tilde{\xi} \in \tilde{H}' := \left\{\tilde{\xi} \in \tilde{\Xi} \middle| \mathbb{H}\left(\mathrm{MIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi})), \mathrm{MIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])\right) \leq \frac{\varepsilon}{L}\right\} \cap \left\{\tilde{\xi} \in \tilde{\Xi} \middle| \mathbb{H}\left(\hat{F}_N(X,\tilde{\xi}), \mathbb{E}_\xi[F(X,\cdot)]\right) \leq \frac{\varepsilon}{L}\right\}.$$

We have

$$\tilde{\mathbb{P}}_\xi(\tilde{H}') \geq p_0.$$

According to Weierstrass' Theorem, let $y_N^* \in \operatorname{argmin}\left(f(y) \middle| y \in \mathrm{MIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi}))\right)$ and $y^* \in \operatorname{argmin}\left(f(y) \middle| y \in \mathrm{MIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])\right)$.

Since $\tilde{\xi} \in \tilde{H}'$, by the definition of the Hausdorff-Pompeiu distance, there exists $\tilde{y} \in \mathrm{MIN}_{\mathbb{R}^2_+}(\mathbb{E}_\xi[F(X,\cdot)])$ and $\tilde{y}_N \in \mathrm{MIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi}))$ such that

$$\|y_N^* - \tilde{y}\| \leq \frac{\varepsilon}{L} \text{ and } \|\tilde{y}_N - y^*\| \leq \frac{\varepsilon}{L}.$$

Thus, by (H_1), since $\tilde{H}' \subset \left\{\tilde{\xi} \in \tilde{\Xi} \middle| \hat{F}_N(X,\tilde{\xi}) \subset \mathbb{E}_\xi[F(X,\cdot)] + \varepsilon_0 \mathbb{B}\right\}$ we have

$$f(y_N^*) \leq f(\tilde{y}_N) \leq f(y^*) + \varepsilon \leq f(\tilde{y}) + \varepsilon \leq f(y_N^*) + 2\varepsilon.$$

Hence, $|f(y_N^*) - f(y^*)| \leq \varepsilon$, and therefore

$$\tilde{\mathbb{P}}_\xi\left(\left\{\tilde{\xi} \in \tilde{\Xi} \middle| |V_N(\tilde{\xi}) - V| \leq \varepsilon\right\}\right) \geq p_0 \quad (N \geq \tilde{N}).$$

STOCHASTIC ALGORITHM :

Step 1:
Choose $\eta > 0$ the error bound, and $p_0 \in]0,1[$ the confidence level. Determine N^0 for $\varepsilon = \frac{\eta}{4}$, i.e.,

$$N^0 = \max_{j=1,2} 32L^2 \frac{\max_{x \in X} \sigma_j^2(x)}{\eta^2(1-p_0)}. \tag{34}$$

Randomly generate $(\xi_1, \xi_2, \ldots, \xi_{N^0})$.

Step 2:
Use the deterministic algorithm with $\varepsilon = \frac{\eta}{2}$ to solve problem $(S_{N^0}(\tilde{\xi}))$.

Remark 10 By Theorem 5, we know that

$$\tilde{\mathbb{P}}_\xi\left(\tilde{\xi} \in \tilde{\Xi} \middle| V \in \left[V_{N^0}(\tilde{\xi}) - \frac{\eta}{4}, V_{N^0}(\tilde{\xi}) + \frac{\eta}{4}\right]\right) \geq p_0.$$

Moreover, the algorithm stop at a step k^0 when $d^{(k^0)} - s^{(k^0)} < \frac{\eta}{2}$, and we have

$$V_{N^0}(\tilde{\xi}) \in [s^{(k^0)}, d^{(k^0)}].$$

Hence, when the algorithm terminates, we have

$$\tilde{\mathbb{P}}_\xi\left(\tilde{\xi} \in \tilde{\Xi} \middle| V \in \left[s^{(k^0)} - \frac{\eta}{4}, d^{(k^0)} + \frac{\eta}{4}\right]\right) \geq p_0,$$

and $(d^{(k^0)} + \frac{\eta}{4}) - (s^{(k^0)} - \frac{\eta}{4}) < \eta$.

4 Numerical Examples

Our goal in this section is to solve two stochastic problems with our stochastic algorithm.

The first example is a simple one since we consider that the feasible set is a subset of $\mathbb{R}^2$ and the vector objective function linear. However it allows us to give an explicit description of the feasible set image, and then to illustrate graphically how our algorithm runs. Moreover, for such an example, we are able to compute the closed form of the optimal value λ^* of problem (CP_t) (each $t \in]0,1[$).

In the second example, we consider the case where the the feasible set is a subset of $\mathbb{R}^4$ and the vector objective function non-linear. In this situation we are not able to give an explicit description of the feasible set in the objective space, but we will give some details step by step.

In both example we note p_0 the confidence level and η the tolerated error bound (recall that the algorithm run with an error bound of $\frac{\eta}{2}$). Moreover, k^0 indicates the number of steps so that the algorithm terminates, $s^{(k^0)}$ is the lower bound at step k^0 and $d^{(k^0)}$ the upper bound. Therefore the optimal value of the SAA-N^0 problem is such that $V_{N^0} \in [s^{(k^0)}, d^{(k^0)}]$, and the true optimal value V satisfies $\tilde{\mathbb{P}}_\xi\left(\tilde{\xi} \in \tilde{\Xi} \middle| V \in \left[s^{(k^0)} - \frac{\eta}{4}, d^{(k^0)} + \frac{\eta}{4}\right]\right) \geq p_0$.

All the numerical results have been obtained with MATLAB7.

4.1 Example 1

Consider the following $(SBOP)$

$$\text{MIN}_{w \in X} \mathbb{E}_\xi\Big(F(w,\cdot)\Big),$$

where the decision variable $w \in \mathbb{R}^2$, the random variable $\xi = (\xi^1, \xi^2) \in \mathbb{R}^2$ is such that ξ^1 and ξ^2 follows both a standard normal distribution with a correlation coefficient of -0.5, the feasible set X is the closed unit ball in $\mathbb{R}^2$ and the objective function of $(SBOP)$ is given by

$$F(w,\xi) = (F^1(w,\xi),F^2(w,\xi)) = (2(w_1+w_2)+\xi^1, w_1-w_2+\xi^2).$$

Now we consider the stochastic problem (S) of minimizing the functional $f:\mathbb{R}^2 \to \mathbb{R},\ z=(x,y)\mapsto x^3+y$ (which is quasi-concave and nondecreasing) over the Pareto outcome set associated with the previous $(SBOP)$, i.e.,

$$(S) \qquad \min\Big(\Big[\mathbb{E}_\xi\Big(F^1(w,\cdot)\Big)\Big]^3 + \Big[\mathbb{E}_\xi\Big(F^2(w,\cdot)\Big)\Big]\Big| w\in \mathrm{ARGMIN}_{\mathbb{R}^2_+}\mathbb{E}_\xi[F(X,\cdot)]\Big).$$

Obviously, we have that $\mathbb{E}_\xi\Big(F^1(w,\cdot),F^2(w,\cdot)\Big) = \Big(2(w_1+w_2), w_1-w_2\Big)$, hence problem (S) can be rewritten in the Outcome space as

$$\min\Big(x^3+y\Big|(x,y)\in \mathrm{MIN}_{\mathbb{R}^2_+}(Z)\Big),$$

where

$$Z=\mathbb{E}_\xi[F(X,\cdot)] = \Big\{(x,y)\in\mathbb{R}^2\Big|\Big(\frac{x}{2\sqrt{2}}\Big)^2+\Big(\frac{y}{\sqrt{2}}\Big)^2\le 1\Big\}.$$

Therefore, problem (S) can be solved directly with our deterministic algorithm.

Step 0 **(see Fig1)**

Let the error bound be $\varepsilon = 5*10^{-6}$. We compute $z_0 = (-2\sqrt{2},0)$ and $z_1=(0,-\sqrt{2})$.

Moreover, $d^{(0)} = f(z_0) = -16\sqrt{2}$, and since f is nondecreasing on $\mathbb{R}^2$, we have $s^{(0)} = f(-2\sqrt{2},-\sqrt{2}) = -17\sqrt{2}$.

Since $d^{(0)}-s^{(0)} = \sqrt{2} > \varepsilon$, we set $r = (-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2})$.

Step 1 **(see Fig2)**

Note that for each $t\in]0,1[$, the convex problem

$$(CP_t) \qquad \sup_{(\lambda,z)}\Big(\lambda\ge 0\Big| z\in z_t+\lambda r-\mathbb{R}^2_+,\ z\in Z\Big)$$

can be rewritten equivalently as

$$\sup\Big(\lambda\ge 0\Big|\Big(\frac{x_t+\lambda r}{2\sqrt{2}}\Big)^2+\Big(\frac{y_t+\lambda r}{\sqrt{2}}\Big)^2=1\Big),$$

where $(x_t,y_t) = z_t = ((t-1)2\sqrt{2},-t\sqrt{2})$.

Hence, the optimal value of problem (CP_t) is given by

$$\lambda^* = \frac{4}{5}\Big(-1-t+\sqrt{-9t^2+12t+1}\Big). \tag{35}$$

At this step we obviously have $j^*=1$ and $t^{(1)}_{j^*}=\frac{1}{2}$, and according to (35), the optimal value of problem $(CP_{\frac{1}{2}})$ is $\lambda^* = 0.5436$.

We compute $z^*_{\frac{1}{2}} = (-1.7986,-1.0915)$, $d^{(1)} = d^{(0)} = -16\sqrt{2}$, and $s^{(1)} = f(-2\sqrt{2},-1.0915) = -16\sqrt{2}-1.0915$. Therefore $d^{(1)}-s^{(1)} = 1.0915 > \varepsilon$.

Step 2 **(see Fig3)**

We compute $j^* = 1, t_{j^*}^{(2)} = \frac{1}{4}, z_{\frac{1}{4}}^* = (-2.4630, -0.6953), d^{(2)} = d^{(0)}, s^{(2)} = f(-2\sqrt{2}, -0.6953)$. Therefore $d^{(2)} - s^{(2)} = 0.6953 > \varepsilon$.

After 698 steps, the algorithm stopped and returned $s^{(698)} = -22.627422$ and $d^{(698)} = -22.627417$. Hence $V \in [-22.627422, -22.627417]$.

Now, consider the corresponding sequence of $SAA-N$ problems ($N \in \mathbb{N}^*, \tilde{\xi} \in \tilde{\Xi}$)

$$(S_N(\tilde{\xi})) \qquad \min \left(\hat{F}_N^1(w, \tilde{\xi})^3 + \hat{F}_N^2(w, \tilde{\xi}) \middle| w = (w_1, w_2) \in \text{ARGMIN}_{\mathbb{R}^2_+} (\hat{F}_N(X, \tilde{\xi})) \right),$$

which can obviously be rewritten in the Outcome space as

$$\min \left(x^3 + y \middle| (x, y) \in \text{MIN}_{\mathbb{R}^2_+} (Z_N(\tilde{\xi})) \right),$$

where

$$Z_N(\tilde{\xi}) = \hat{F}_N(X, \tilde{\xi}) = \left\{ (x,y) \in \mathbb{R}^2 | \left(\frac{x - \frac{1}{N} \sum_{i=1}^N \xi_i^1}{2\sqrt{2}} \right)^2 + \left(\frac{y - \frac{1}{N} \sum_{i=1}^N \xi_i^2}{\sqrt{2}} \right)^2 \leq 1 \right\}.$$

Then, for a given N, we are able to solve problem $(S_N(\tilde{\xi}))$ with our algorithm.

Using Bienaymé-Tchebychev inequality, it is easy to see that $\tilde{\mathbb{P}}_\xi \left(\tilde{\xi} \in \tilde{\Xi} \middle| Z_N(\tilde{\xi}) \subset 3\mathbb{B} \right) \geq p_0$ if and only if $N \geq \frac{1}{(1-p_0)(17-12\sqrt{2})}$. Moreover, $f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto x^3 + y$ is $\sqrt{730}$-Lipschitz continuous on $3\mathbb{B}$, hence assumption (H_1) holds with $\varepsilon_0 = 3 - 2\sqrt{2}$ and $L = \sqrt{730}$.

p_0	η	N^0	k^0	$V_{N^0}(\tilde{\xi}) \in [s^{(k^0)}, d^{(k^0)}]$	$\tilde{\mathbb{P}}_\xi(\tilde{\xi} \in \tilde{\Xi} \lvert V \in [s^{(k^0)} - \frac{\eta}{4}, d^{(k^0)} + \frac{\eta}{4}]) \geq p_0$	elapsed time
0.95	0.2	$1.2 * 10^7$	7	$V_{N^0}(\tilde{\xi}) \in [-22.6909, -22.6209]$	$\tilde{\mathbb{P}}_\xi(\tilde{\xi} \in \tilde{\Xi} \lvert V \in [-22.7409, -22.5709]) \geq 0.95$	2.4 seconds
0.95	0.1	$4.7 * 10^7$	8	$V_{N^0}(\tilde{\xi}) \in [-22.6665, -22.6234]$	$\tilde{\mathbb{P}}_\xi(\tilde{\xi} \in \tilde{\Xi} \lvert V \in [-22.6915, -22.5984]) \geq 0.95$	4.2 seconds
0.95	0.05	$1.8 * 10^8$	10	$V_{N^0}(\tilde{\xi}) \in [-22.6452, -22.6250]$	$\tilde{\mathbb{P}}_\xi(\tilde{\xi} \in \tilde{\Xi} \lvert V \in [-22.6577, -22.6125]) \geq 0.95$	8.7 seconds

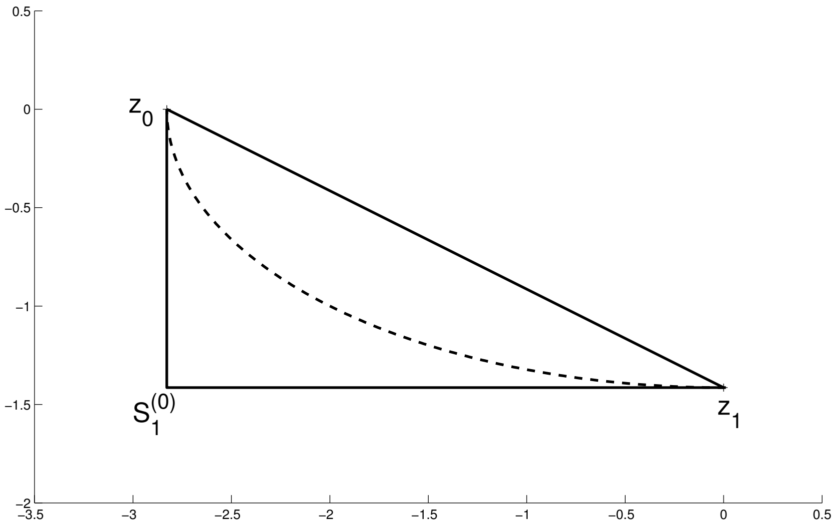


Fig. 1 Step 0 The Pareto outcome set (dashed line), the set $D^{(0)} = \{z_0, z_1\}$ and the set $S^{(0)} = S_1^{(0)}$

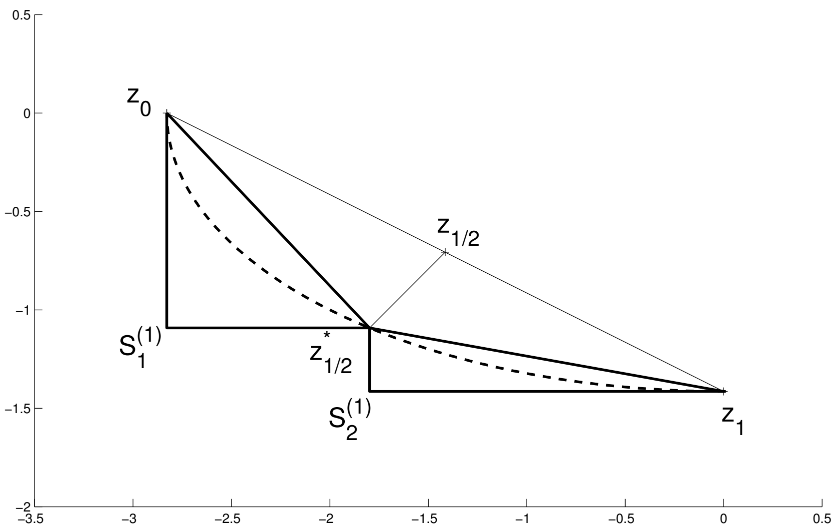


Fig. 2 Step 1 The Pareto outcome set (dashed line), the set $D^{(1)} = \{z_0, z^*_{\frac{1}{2}}, z_1\}$ and the set $S^{(1)} = S_1^{(1)} \cup S_2^{(1)}$

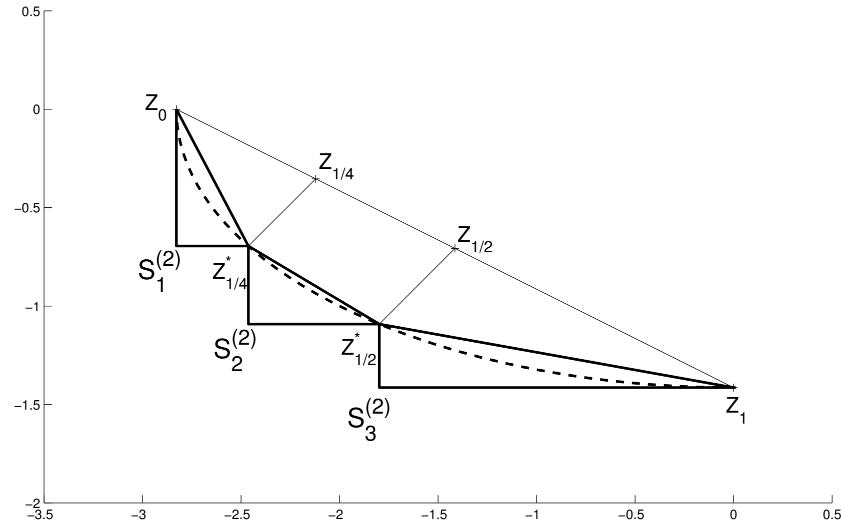


Fig. 3 Step 2 The Pareto outcome set (dashed line), the set $D^{(2)} = \{z_0, z^*_{\frac{1}{4}}, z^*_{\frac{1}{2}}, z_1\}$ and the set $S^{(2)} = S_1^{(2)} \cup S_2^{(2)} \cup S_3^{(2)}$

4.2 Example 2

Now we consider the following (*SBOP*)

$$\mathrm{MIN}_{w\in X}\mathbb{E}_\xi\Big(F(w,\cdot)\Big),$$

where the decision variable $w \in \mathbb{R}^4$, the random variable $\xi = (\xi^1, \xi^2) \in \mathbb{R}^2$ is such that ξ^1 and ξ^2 follows both a uniform distribution with mean 1 and variance $\frac{1}{3}$, and the feasible set X is defined by

$$X := \Big\{w \in \mathbb{R}^4 \Big| \|w\| \leq 1, \sum_{i=1}^4 w_i \geq 0\Big\}.$$

The objective function of (*SBOP*) is given by

$$F(w,\xi) = (F^1(w,\xi), F^2(w,\xi)) = \Big(\xi^1(2\exp(w_1+w_2)+(w_3+w_4)^2), \xi^2(4(w_1+w_2)^2+\frac{1}{2}\exp(w_3+w_4))\Big).$$

We are interested by the stochastic problem (S) of minimizing the mapping $f : \mathbb{R}^2 \to \mathbb{R}$, $z = (x,y) \mapsto (x+y)^3$ over the Pareto outcome set associated with the previous (*SBOP*), i.e.,

$$(S) \qquad \min\Big(\Big(\mathbb{E}_\xi\Big[F^1(w,\cdot)\Big]+\mathbb{E}_\xi\Big[F^2(w,\cdot)\Big]\Big)^3\Big| w = (w_1,w_2,w_3,w_4) \in \mathrm{ARGMIN}_{\mathbb{R}^2_+}\mathbb{E}_\xi[F(X,\cdot)]\Big).$$

We obviously have

$$\mathbb{E}_\xi[F(w,\cdot)] = \Big(2\exp(w_1+w_2)+(w_3+w_4)^2, 4(w_1+w_2)^2+\frac{1}{2}\exp(w_3+w_4)\Big).$$

Therefore, we are able to directly solve the stochastic problem (S) with our deterministic algorithm :

Step 0

Let the error bound be $\varepsilon = 0.05$. We compute $z_0 = (1.4559, 2.1684)$, $z_1 = (2.1249, 0.4853)$ and $d^{(0)} = f(z_1) = 17.7827$.

Since the mapping $(x,y) \mapsto (x+y)^3$ is nondecreasing on $\mathbb{R}^2$, by Remark7 we have $s^{(0)} = f(1.4559, 0.4853) = 7.3151$.

Since $d^{(0)} - s^{(0)} = 10.4676 > \varepsilon$, we set $r = (-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$.

Remark 11 Note that for each $t \in]0,1[$, the optimal value λ^* of the convex problem

$$(CP_t) \qquad \sup_{(\lambda,z)} \Big(\lambda \geq 0 \Big| z \in z_t + \lambda r - \mathbb{R}^2_+,\ z \in Z\Big)$$

is computed with the function "fmincon" of Matlab.

Step 1

At this step we have $j^* = 1$, $t^{(1)}_{j^*} = \frac{1}{2}$, and we compute $\lambda^* = 0.2343$ (the optimal value of problem $(CP_{\frac{1}{2}})$).

Then we have $z^*_{\frac{1}{2}} = (1.5561, 1.0925)$ and $d^{(1)} = f(z_1) = 17.7827$. According to Remark7, we have $s^{(1)} = f(1.5561, 0.4853) = 8.5071$. Finally $d^{(1)} - s^{(1)} = 9.2756 > \varepsilon$.

Step 2

We have $j^* = 2$, $t^{(2)}_{j^*} = \frac{3}{4}$, $\lambda^* = 0.2231$. Therefore $z^*_{\frac{3}{4}} = (1.7315, 0.6830)$. We compute $d^{(2)} = f(z^*_{\frac{3}{4}}) = 14.1285$ and $s^{(2)} = f(1.7345, 0.4853) = 10.9384$. Therefore $d^{(2)} - s^{(2)} = 3.1901 > \varepsilon$.

Step 3

We have $j^* = 3$, $t^{(2)}_{j^*} = \frac{7}{8}$, $\lambda^* = 0.1517$. Therefore $z^*_{\frac{7}{8}} = (1.8895, 0.5439)$. We compute $d^{(2)} = f(z^*_{\frac{3}{4}}) = 14.1285$ and $s^{(2)} = f(1.5561, 0.6830) = 11.2254$. Therefore $d^{(2)} - s^{(2)} = 2.9031 > \varepsilon$.

After 234 seconds and 151 steps, our deterministic algorithm returns $s^{(151)} = 13.9853$, $d^{(151)} = 13.9903$, hence the true optimal value V of problem S is such that $V \in [13.9853, 13.9903]$.

Now we consider the sequence of $SAA-N$ problems ($N \in \mathbb{N}^*, \tilde{\xi} \in \tilde{\Xi}$) given by

$$(S_N(\tilde{\xi})) \qquad \min\Big(\Big(\hat{F}^1_N(w,\tilde{\xi}) + \hat{F}^2_N(w,\tilde{\xi})\Big)^3 \Big| w = (w_1,w_2,w_3,w_4) \in \mathrm{ARGMIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi}))\Big),$$

which can be rewritten in the Outcome space as

$$\min\Big(\Big(x+y\Big)^3 \Big| (x,y) \in \mathrm{MIN}_{\mathbb{R}^2_+}(\hat{F}_N(X,\tilde{\xi}))\Big).$$

Note that the mapping, $f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto (x+y)^3$ is $600\sqrt{2}$-Lipschitz continuous on $10\mathbb{B}$.

The results obtained with our stochastic algorithm are given in the following array:

p_0	η	N^0	k^0	$V_{N^0}(\tilde{\xi}) \in [s^{(k^0)}, d^{(k^0)}]$	$\tilde{\mathbb{P}}_\xi(\tilde{\xi} \in \tilde{\Xi} \mid V \in [s^{(k^0)} - \frac{\eta}{4}, d^{(k^0)} + \frac{\eta}{4}]) \geq p_0$	elapsed time
0.95	0.2	$4.5 * 10^6$	35	$V_{N^0}(\tilde{\xi}) \in [13.9264, 14.0004]$	$\tilde{\mathbb{P}}_\xi(\tilde{\xi} \in \tilde{\Xi} \mid V \in [13.8764, 14.0540]) \geq 0.95$	12 seconds
0.95	0.1	$1.8 * 10^7$	48	$V_{N^0}(\tilde{\xi}) \in [13.9473, 13.9953]$	$\tilde{\mathbb{P}}_\xi(\tilde{\xi} \in \tilde{\Xi} \mid V \in [13.9223, 14.0203]) \geq 0.95$	24 seconds
0.95	0.05	$7.2 * 10^7$	69	$V_{N^0}(\tilde{\xi}) \in [13.9658, 13.9903]$	$\tilde{\mathbb{P}}_\xi(\tilde{\xi} \in \tilde{\Xi} \mid V \in [13.9553, 14.0028]) \geq 0.95$	52 seconds

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. H.P. BENSON, *Optimization over the Efficient Set,* J. Math. Anal. Appl. Volume **98**, pp. 562-580, (1984).
3. 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).
4. H.P. BENSON, *Generating the Efficient Outcome Set in Multiple Objective Linear Programs: The Bicriteria Case,* Acta Math. Vietnam. **22** 29-51, (1997).
5. H.P. BENSON, *Further Analysis of an Outcome Set-Based Algorithm for Multiple Objective Linear Programming,* J. Optim. Theory Appl. **97** 1-10, (1998).
6. H.P. BENSON, *Hybrid Approach for Solving Multiple Objective Linear Programs in Outcome Space,* J. Optim. Theory Appl. **98** 17-35, (1998).
7. H.P. BENSON, *An outer approximation algorithm for generating all efficient extreme points in the outcome set of a multiple objective linear programming problem,* J. Global Optim. **13**, 1-24, (1998).
8. 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).
9. S. BOLINTINÉANU, *Optimality conditions for minimization over the (weakly or properly) efficient set,* J. Math. Anal. Appl., **173**(2), 523-541, (1993).
10. S. BOLINTINÉANU, *Necessary Conditions for Nonlinear Suboptimization over the Weakly-Efficient Set,* J. Optim. Theory Appl. **78**, (1993).
11. S. BOLINTINÉANU, *Minimization of a quasi-concave function over an efficient set,* Math. Program., **61**, 89-110, (1993).
12. S. BOLINTINÉANU, *Necessary conditions for nonlinear suboptimization over the weakly-efficient set,* J. Optim. Theory Appl., **78**, 579-598, (1993).
13. H. BONNEL AND J. COLLONGE, *Stochastic Optimization over a Pareto Set Associated with a Stochastic Multi-Objective Optimization Problem, J. Optim. Theory Appl., (2013), (DOI) 10.1007/s10957-013-0367-8*
14. 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).
15. B. D. CRAVEN, *Aspects of multicriteria optimization,* Recent Developments in Mathematical Programming, 93-100, (1991).
16. J. P. DAUER, *Optimization over the efficient set using an active constraint approach,* J. Oper. Res., **35**, 185-195, (1991).
17. J. P. DAUER AND T. A. FOSNAUGH, *Optimization over the efficient set,* J. Global Optim., **7**, 261-277, (1995).
18. G. EICHFELDER, *Adaptative Scalarization Methods in Multiobjective Optimization,* Springer-Verlag Berlin Heidelberg, (2008.)
19. M. EHRGOTT, *Multicriteria Optimization,* Springer-Verlag Berlin Heidelberg, (2000).
20. J. FLIEGE AND H. XU, *Stochastic Multiobjective Optimization : Sample Average Approximation and Applications* J. Optim. Theory Appl. **151** 135-162, (2011).
21. 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).
22. A. GÖPFERT, H. RIAHI, C. TAMMER AND C. ZĂLINESCU, *Variational Methods in Partially Ordered Spaces,* Springer, Berlin, (2003).
23. GUANG-YA CHEN, XUEXIANG HUANG AND XIAOQI YANG, *Vector Optimization : Set Valued and Variational Analysis,* Springer, Berlin, (2005).

24. 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).
25. 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).
26. J. JACOD, P. PROTTER, *Probability Essentials,* Springer-Verlag Berlin Heidelberg, (2004).
27. J. JAHN, *Vector Optimization,* Springer-Verlag, Berlin, Heidelberg, New York, (2004).
28. N.T.B. KIM AND T.N. THANG, *Optimization over the Efficient Set of a Bicriteria Convex Programming Problem,* Pac. J. Optim., **9** 103-115, (2013).
29. 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).
30. H. KONNO, M. INORI, *Bond Portfolio Optimization by Bilinear Fractional Programming,* Journal of the Operations Research Society of Japan, **32** 143-158, (1989).
31. D. T. LUC, *Theory of Vector Optimization,* Lecture Notes in Econom. and Math. Systems 319, Springer-Verlag, Berlin, Heidelberg, (1989).
32. K. M. MIETTINEN, *Nonlinear Multiobjective Optimization,* Kluwer Academic Publishers, (1998).
33. J. PHILIP, *Algorithms for the vector maximization problem,* Math. Program., 207-229, (1972).
34. A. SHAPIRO, D. DENTCHEVA, A. RUSZCZYNSKI, *Lectures on stochastic programming : modeling and theory,* MPS/SIAM Ser. Optim., (2009).
35. A. SHAPIRO, H.XU, *Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation,* Taylor and Francis Group, Optimization **57**, No.3 395-418, (2008).
36. Y. YAMAMOTO, *Optimization over the efficient set : an overview,* J. Global Optim., **22**, 285-317, (2002).
37. P.L. YU, *Multiple-Criteria Decision Making,* Plenum Press, New York and London (1985).

Perspectives de Recherche

– Notre principale perspective de recherche vise à généraliser notre approche par l'étude d'un Problème d'Optimisation à deux niveaux : à savoir le **niveau supérieur** qui est un Problème d'Optimisation **scalaire** et le **niveau inférieur** qui est **vectoriel**.

  Plus précisément, nous nous intéresserons au Problème d'Optimisation Semi-Vectoriel suivant :

$$(POSV) \qquad \min \Big(f(y,x) \Big| (y,x) \in Y \times X, x \in P(y) \Big),$$

  où :
  $Y \subset \mathbb{R}^m$ représente l'ensemble des variables de décision du leader (le niveau supérieur),
  $X \subset \mathbb{R}^n$ représente l'ensemble des variables de décision des agents (le niveau inférieur),
  $f : Y \times X \to \mathbb{R}$ représente l'objectif du leader, et
  pour tout $y \in Y$, $P(y) \subset X$ est l'ensemble de Pareto faible du Problème d'Optimisation Multi-Objectifs

$$(POMO) \qquad \text{MIN}_{x \in X} F(y,x),$$

  où $F = (F^1, F^2, \ldots, F^r) : Y \times X \to \mathbb{R}^r$ représente les objectifs des différents agents. Autrement dit,

$$P : Y \rightrightarrows X, P(y) = \text{w-ARGMIN}_{\mathbb{R}^r_+} F(y,X).$$

  En pratique, le leader prend une décision ($y \in Y$), puis les agents prennent la leur ($x \in X$). Bien entendu, la coopération entre le leader et les agents a un impact direct sur la décision prise par ces derniers. Par conséquent, plusieurs cas de figure se présentent :

  1. les agents coopèrent entièrement avec le leader : quel que soit la décision $y \in Y$ prise par le leader, les agents prennent une des décisions $x_y \in P(y)$ qui soit la plus favorable pour le leader. Autrement dit, la

décision x_y prise par les agents est telle que

$$f(y, x_y) = \min\left(f(y, x)\middle| x \in P(y)\right).$$

Dans ce cas, dit ***Optimiste***, le Problème d'Optimisation Semi-Vectoriel $(POSV)$ se réécrit :

$$(POSVO) \qquad \min_{y \in Y} \min_{x \in P(y)} f(y, x).$$

2. les agents vont contre la volonté du leader : quel que soit la décision $y \in Y$ prise par le leader, les agents prennent une des $x_y \in P(y)$ qui soit la plus défavorable pour le leader. Autrement dit, la décision x_y prise par les agents est telle que

$$f(y, x_y) = \max\left(f(y, x)\middle| x \in P(y)\right).$$

Dans ce cas, dit ***Pessimiste***, le Problème d'Optimisation Semi-Vectoriel $(POSV)$ se réécrit :

$$(POSVP) \qquad \min_{y \in Y} \max_{x \in P(y)} f(y, x).$$

3. la coopération entre les agents et le leader est partielle : en fonction de la décision $y \in Y$ prise par le leader, les agents prennent une des décisions $x_y \in P(y)$ qui vérifie

$$f(y, x_y) = \beta(y) \min\left(f(y, x)\middle| x \in P(y)\right) + (1 - \beta(y)) \max\left(f(y, x)\middle| x \in P(y)\right),$$

où $\beta(y) \in]0; 1[$ représente le niveau de coopération.

Remarquons que lorsque l'ensemble des variables de décision du leader Y est réduit à un singleton, le Problème d'Optimisation Semi-Vectoriel considéré est un Problème d'Optimisation post-Pareto.

Dans l'avenir, plusieurs pistes de réflexion sont à envisager pour poursuivre ces travaux :

1. **Montrer l'existence des solutions :**

$$\text{Supposons } (H) \begin{cases} X \text{ est un sous-ensemble convexe et compact de } \mathbb{R}^n, \\ \forall y \in Y,\ \forall i = 1, \ldots, r,\ x \mapsto F^i(y, x) \text{ est continue et convexe sur } X. \end{cases}$$

Sous ces hypothèses, pour tout $y \in Y$ et pour tout $\lambda \in \Lambda_w := \{c \in \mathbb{R}^r_+, ||c|| = 1\}$, la fonction réelle $X \ni x \mapsto \sum_{i=1}^{r} \lambda^i F^i(y,x)$ est convexe, et l'ensemble $X'(y,\lambda) := \underset{x \in X}{\operatorname{argmin}} \left(\sum_{i=1}^{r} \lambda^i F^i(y,x) \right)$ est non vide.

De plus, pour tout $y \in Y$, le Théorème de Scalarisation nous permet d'affirmer que

$$\bigcup_{\lambda \in \Lambda_w} X'(y,\lambda) = \text{w-ARGMIN}_{\mathbb{R}^r_+} F(y,X).$$

Ainsi, le Problème d'Optimisation Semi-Vectoriel $(POSVO)$ est équivalent au problème

$$\min_{y \in Y} \min_{\lambda \in \Lambda_w} \min_{x' \in X'(y,\lambda)} f(y,x'),$$

et le Problème d'Optimisation Semi-Vectoriel $(POSVP)$ est équivalent au problème

$$\min_{y \in Y} \max_{\lambda \in \Lambda_w} \min_{x' \in X'(y,\lambda)} f(y,x').$$

En supposant de plus

(H_s) $\forall y \in Y$, $\exists i = 1, \ldots, r$, $x \mapsto F^i(y,x)$ est strictement convexe sur X,

le Problème d'Optimisation scalaire $\min_{x \in X} \left(\sum_{i=1}^{r} \lambda^i F^i(y,x) \right)$ $(y \in Y, \lambda \in \Lambda_w)$ admet une unique solution qui sera notée $x(y,\lambda)$.

Ainsi, le Problème d'Optimisation Semi-Vectoriel $(POSVO)$ est équivalent au problème

$$\min_{y \in Y} \min_{\lambda \in \Lambda_w} f(y, x(y,\lambda)),$$

et le Problème d'Optimisation Semi-Vectoriel $(POSVP)$ est équivalent au problème

$$\min_{y \in Y} \max_{\lambda \in \Lambda_w} f(y, x(y,\lambda)).$$

Comme Λ_w est un ensemble compact, en admettant que les applications f et $\lambda \mapsto x(y,\lambda)$ sont continues, les réels $\min_{\lambda \in \Lambda_w} f(y, x(y,\lambda))$ et $\max_{\lambda \in \Lambda_w} f(y, x(y,\lambda))$ sont bien définis pour tout $y \in Y$.

En supposant maintenant que Y est un sous-ensemble compact de $\mathbb{R}^m$, il est facile de vérifier que les Problèmes d'Optimisation Semi-Vectoriel $(POSVO)$ et $(POSVP)$ admettent tous deux au moins une solution optimale.

2. **Donner des conditions explicites d'optimalité dans le cas où le niveau inférieur est quadratique :**

   Supposons que pour tout $y \in Y = \mathbb{R}^m$ et pour tout $i \in \{1, \ldots, r\}$, les objectifs des agents sont de la forme

   $$F^i(y, x) = \frac{1}{2} x^T A_i(y) x - b_i(y)^T x,$$

   où $A_i(y)$ est une matrice réelle $n \times n$ symétrique définie positive et $b_i(y)$ est un vecteur de $\mathbb{R}^n$. Notons que l'hypothèse (H_s) introduite précédemment est vérifiée.

   De plus, supposons que pour chaque $\lambda \in \Lambda_w$, la matrice $\sum\limits_{i=1}^{r} \lambda^i A_i(y)$ est définie positive.

   Ainsi, pour tout $y \in Y$ et $\lambda \in \Lambda_w$, l'unique solution optimale $x(y, \lambda)$ du problème d'Optimisation scalaire $\min\limits_{x \in X} \left(\sum\limits_{i=1}^{r} \lambda^i F^i(y, x) \right)$ est explicitement donnée par

   $$x(y, \lambda) = \left(\sum_{i=1}^{r} \lambda^i A_i(y) \right)^{-1} \left(\sum_{i=1}^{r} \lambda^i b_i(y) \right).$$

3. **Étendre ces résultats au cas où** $\forall y \in Y,\ P(y) = \textbf{p-ARGMIN}_{\mathbb{R}^r_+} F(y, X)$ **est l'ensemble de Pareto propre du (POMO) :**

   Sous (H), en vertu du Théorème de Scalarisation,

   $$\bigcup_{\lambda \in \Lambda_p} \underset{x \in X}{\operatorname{argmin}} \left(\sum_{i=1}^{r} \lambda^i F^i(y, x) \right) = \text{p-ARGMIN}_{\mathbb{R}^r_+} F(y, X),$$

   où $\Lambda_p := \{c \in \operatorname{int}(\mathbb{R}^r_+), ||c|| = 1\}$.

   Comme Λ_p est un ensemble ouvert, en supposant les fonctions différentiables et en appliquant la règle de **Pierre de Fermat**, nous espérons pouvoir établir des conditions explicites d'optimalité.

4. **Considérer et traiter le cas Stochastique :**

   Nous étendrons l'étude réalisée au cours de notre première contribution au cas semi-Vectoriel. Les résultats de convergence au sens de Hausdorff-Pompeiu obtenus dans l'espace des décisions sont suffisamment fertiles pour que nous supposions la démarche féconde.

Nous envisageons également de traiter le cas où le niveau inférieur est quadratique et stochastique.

- **Directions de Recherches :**

  - Proposer des applications en finance : l'adaptation de notre algorithme pour développer un modèle espérance-variance semble possible.

  - Concevoir un algorithme opérationnel pour un nombre quelconque d'objectifs et envisager la recherche des solutions optimales.

  - Lever les hypothèses de convexité.

  - Traiter des cas dynamiques en considérant des Équations Différentielles Stochastiques.

  - Travailler avec des cônes quelconques.

  - ...

Annexe A

Applications Multivoques

A.1 La distance de Hausdorff-Pompeiu

Soient (X, d) un espace métrique, et A et B deux sous-ensembles non-vides et bornés de X.

La distance d'un point $x \in X$ à l'ensemble A est donnée par

$$\text{dist}(x, A) := \inf \Big(d(x, a) : a \in A \Big).$$

Définition A.1.1. *La **distance de Hausdorff-Pompeiu** entre les ensembles A et B est donnée par*

$$\mathbb{H}(A, B) := \max \Big(\sup_{a \in A} \text{dist}(a, B); \sup_{b \in B} \text{dist}(b, A) \Big).$$

Le réel $\mathbb{D}(A, B) := \sup_{a \in A} \text{dist}(a, B)$ (resp. $\mathbb{D}(B, A) := \sup_{b \in B} \text{dist}(b, A)$) est appelé **Déviation** entre les ensembles A et B (resp. **Déviation** entre les ensembles B et A).

La distance de Hausdorff-Pompeiu est une pseudo-distance sur $\mathcal{P}(X) := \{A | A \subset X\}$:

Théorème A.1.1. *Soient $A, B, C \in \mathcal{P}(X)$.Les assertions suivantes sont vraies :*

1. $\mathbb{H}(A, B) \geq 0$.
2. $\mathbb{H}(A, A) = 0$.
3. $\mathbb{H}(A, B) = \mathbb{H}(B, A)$.
4. $\mathbb{H}(A, B) \leq \mathbb{H}(A, C) + \mathbb{H}(C, B)$.

Remarque A.1.1. *En général, $\mathbb{H}$ n'est pas une distance car $\mathbb{H}(A, B) = 0$ n'implique pas toujours $A = B$. En fait, si $\mathbb{H}(A, B) = 0$, alors $A \subset \bar{B}$ et $B \subset \bar{A}$.*

Toutefois, la distance de Hausdorff-Pompeiu est une distance sur $\mathcal{P}_c(X) := \{A | A \subset X, A \text{ compact}\}$.

Soit $\epsilon > 0$. Pour tout $A \in \mathcal{P}(X)$, posons $\tilde{A}_\epsilon := \Big\{x \in X | \text{dist}(x, A) < \epsilon\Big\}$.

Théorème A.1.2. *Soient $A, B \in \mathcal{P}(X)$. Les déviations entre les ensembles A et B sont*

- $\sup\limits_{a \in A} \text{dist}(a, B) = \inf(\epsilon > 0 | A \subset \tilde{B}_\epsilon)$, *et*
- $\sup\limits_{b \in B} \text{dist}(b, A) = \inf(\epsilon > 0 | B \subset \tilde{A}_\epsilon)$,

d'où

$$\mathbb{H}(A, B) = \inf\Big(\epsilon > 0 | A \subset \tilde{B}_\epsilon \text{ et } B \subset \tilde{A}_\epsilon\Big).$$

A.2 Convergence de suite d'ensembles

Soient (X, d) un espace métrique, et $(A_n)_{n \in \mathbb{N}}$ une suite de sous-ensembles non-vides de X.

Définition A.2.1 (convergence au sens de Kuratowski)**.**

- *La limite supérieure de la suite $(A_n)_{n \in \mathbb{N}}$ est un sous ensemble de X donnée par*
$$\limsup_{n \to \infty} A_n := \Big\{x \in X \Big| \liminf_{n \to \infty} \text{dist}(x, A_n) = 0\Big\}.$$
- *La limite inférieure de la suite $(A_n)_{n \in \mathbb{N}}$ est un sous ensemble de X donnée par*
$$\liminf_{n \to \infty} A_n := \Big\{x \in X \Big| \lim_{n \to \infty} \text{dist}(x, A_n) = 0\Big\}.$$

Si $\limsup\limits_{n \to \infty} A_n = \liminf\limits_{n \to \infty} A_n$, on dit que la limite de la suite $(A_n)_{n \in \mathbb{N}}$ existe et on pose

$$\lim_{n \to \infty} A_n = \limsup_{n \to \infty} A_n = \liminf_{n \to \infty} A_n = A,$$

La suite $(A_n)_{n \in \mathbb{N}}$ est alors dite convergente vers A au sens de Kuratowski.

Proposition A.2.1.

- *La limite supérieure de la suite $(A_n)_{n \in \mathbb{N}}$ est l'ensemble des valeurs d'adhérences des suites $(u_n)_{n \in \mathbb{N}}$ telles que $u_n \in A_n$ $\forall n \in \mathbb{N}$.*
- *La limite inférieure de la suite $(A_n)_{n \in \mathbb{N}}$ est l'ensemble des limites des suites $(u_n)_{n \in \mathbb{N}}$ telles que $u_n \in A_n$ $\forall n \in \mathbb{N}$.*

Définition A.2.2 (convergence au sens de Hausdorff)**.**
Soient $(A_n)_{n \in \mathbb{N}}$ une suite de sous-ensembles fermés non-vides de X, et $A \subset X$ un sous-ensemble fermé non-vide. La suite $(A_n)_{n \in \mathbb{N}}$ est dite convergente vers A au sens de Hausdorff si, et seulement si,

$$\lim_{n \to \infty} \mathbb{H}(A_n, A) = 0.$$

Remarque A.2.1. *Si* $\lim_{n\to\infty} \mathbb{H}(A_n, A) = 0$*, alors pour tout* $\epsilon > 0$*, il existe* $N \in \mathbb{N}$ *tel que* $A_n \subset \tilde{A}_\epsilon$ *pour tout* $n \geq N$*.*

Théorème A.2.1. *Soient* $(A_n)_{n\in\mathbb{N}}$ *une suite de sous-ensembles fermés non-vides de* X*, et* $A \subset X$ *un sous-ensemble fermé non-vide. Si la suite* $(A_n)_{n\in\mathbb{N}}$ *converge vers* A *au sens de Hausdorff, alors elle converge au sens de Kuratowski.*

La réciproque du dernier Théorème n'est généralement pas vraie, excepté dans le cas où les ensembles sont des compacts :

Théorème A.2.2. *Soient* $(A_n)_{n\in\mathbb{N}}$ *une suite de sous-ensembles compacts non-vides de* X*, et* $A \subset X$ *un un sous-ensemble compact non-vide. S'il existe un compact de* X *qui contient tous les ensembles* A_n*, alors la convergence de la suite* $(A_n)_{n\in\mathbb{N}}$ *vers* A *au sens de Kuratowski implique la convergence au sens de Hausdorff.*

A.3 Applications Multivoques

Soient (X, d_X) et (Y, d_Y) deux espaces métriques.

Définition A.3.1. *Une application* Γ *définie sur* X *et à valeurs dans* Y *est dite multivoque si, et seulement si, à chaque point* $x \in X$*, il existe un ensemble* $\Gamma(x) \subset Y$*. Une telle application est notée*

$$\Gamma : X \rightrightarrows Y.$$

Définition A.3.2. *Soit* $\Gamma : X \rightrightarrows Y$ *une application multivoque. On définit*

- *le domaine de* Γ *par*

$$Dom(\Gamma) := \{x \in X | \Gamma(x) \neq \emptyset\},$$

- *l'image de* Γ *par*

$$Im(\Gamma) := \bigcup_{x\in Dom(\Gamma)} \Gamma(x),$$

- *le graphe de* Γ *par*

$$Graph(\Gamma) := \{(x, y) | y \in \Gamma(x), x \in Dom(\Gamma)\}.$$

Définition A.3.3. *Une application multivoque* $\Gamma : X \rightrightarrows Y$ *est dite*

- *à valeurs fermées (resp. valeurs ouvertes, resp. valeurs compactes) si, et seulement si, pour chaque* $x \in X$*, l'ensemble* $\Gamma(x)$ *est un ensemble fermé (resp. ouvert, resp. compact).*
- *fermée (resp. ouverte, resp. compacte) si* $Graph(\Gamma)$ *est un ensemble fermé (resp. ouvert, resp. compact).*

Définition A.3.4. *La limite inférieure de* $\Gamma : X \rightrightarrows Y$ *au point* $x_0 \in X$ *est donnée par*

$$\begin{aligned}\liminf_{x \to x_0} \Gamma(x) \; := \; & \{y \in Y | \forall V \text{ voisinage de } y, \exists U \text{ voisinage de } x_0 : \\ & \forall x \in U \setminus \{x_0\},\ \Gamma(x) \cap V \neq \emptyset\}.\end{aligned}$$

La limite supérieure de Γ *au point* $x_0 \in X$ *est définie par*

$$\begin{aligned}\limsup_{x \to x_0} \Gamma(x) \; := \; & \{y \in Y | \forall V \text{ voisinage de } y, \forall U \text{ voisinage de } x_0, \\ & \exists x \in U \setminus \{x_0\} :\ \Gamma(x) \cap V \neq \emptyset\}.\end{aligned}$$

Remarque A.3.1. *Soient* A, $(A_n)_{n\in\mathbb{N}}$ *des sous-ensembles non-vides de* Y, *et* $X = \mathbb{N} \cup \{+\infty\}$ *muni de la topologie induite par celle de* $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$.

Si $\Gamma : X \rightrightarrows Y$ *est définie pour tout* $n \in \mathbb{N}$ *par* $\Gamma(n) := A_n$ *et* $\Gamma(+\infty) := A$, *alors* $\liminf_{n\to\infty} A_n = \liminf_{n\to\infty} \Gamma(n)$ *et* $\limsup_{n\to\infty} A_n = \limsup_{n\to\infty} \Gamma(n)$.

Définition A.3.5. *L'application multivoque* $\Gamma : X \rightrightarrows Y$ *est dite :*

- *continue supérieurement au point* $x_0 \in X$ *si, et seulement si, pour tout ensemble ouvert* $D \subset Y$ *tel que* $\Gamma(x_0) \subset D$, *il existe un voisinage* $U \subset X$ *de* x_0 *tel que pour tout* $x \in U$, $\Gamma(x) \subset D$.
- *continue inférieurement au point* $x_0 \in X$ *si, et seulement si, pour tout ensemble ouvert* $D \subset Y$ *tel que* $\Gamma(x_0) \cap D \neq \emptyset$, *il existe un voisinage* $U \subset X$ *de* x_0 *tel que pour tout* $x \in U$, $\Gamma(x) \cap D \neq \emptyset$.
- *continue au point* $x_0 \in X$ *si, et seulement si,* Γ *est continue supérieurement et continue inférieurement en* x_0.
- *continue si, et seulement si,* Γ *est continue en chaque point de* X.

Proposition A.3.1. *L'application multivoque* $\Gamma : X \rightrightarrows Y$ *est continue inférieurement au point* $x_0 \in X$ *si, et seulement si,* $\Gamma(x_0) \subset \liminf_{x\to x_0} \Gamma(x)$.

Définition A.3.6. *On dit que l'application multivoque* $\Gamma : X \rightrightarrows Y$ *est*

- *continue supérieurement au sens de Hausdorff au point* $x_0 \in X$ *si, et seulement si, pour tout* $\epsilon > 0$, *il existe un voisinage* $U \subset X$ *de* x_0 *tel que pour tout* $x \in U$,
$$\Gamma(x) \subset \Gamma(x_0) + B_\epsilon,$$
où B_ϵ *désigne la boule ouverte de rayon* ϵ.
- *continue inférieurement au sens de Hausdorff au point* $x_0 \in X$ *si, et seulement si, pour tout* $\epsilon > 0$, *il existe un voisinage* $U \subset X$ *de* x_0 *tel que pour tout* $x \in U$,
$$\Gamma(x_0) \subset \Gamma(x) + B_\epsilon.$$
- *continue au sens de Hausdorff en* $x_0 \in X$ *si, et seulement si, elle est continue supérieurement au sens de Hausdorff et continue inférieurement au sens de Hausdorff au point* x_0.

– *continue au sens de Hausdorff si, et seulement si, elle est continue au sens de Hausdorff en chaque point de* X.

Remarque A.3.2. *L'application multivoque* $\Gamma : X \rightrightarrows Y$ *est continue supérieurement au sens de Hausdorff au point* $x_0 \in X$ *si, et seulement si,*

$$\lim_{x \to x_0} \mathbb{D}(\Gamma(x), \Gamma(x_0)) = 0,$$

et Γ *est continue inférieurement au sens de Hausdorff au point* $x_0 \in X$ *si, et seulement si,*

$$\lim_{x \to x_0} \mathbb{D}(\Gamma(x_0), \Gamma(x)) = 0.$$

Finalement, Γ *est continue au sens de Hausdorff au point* $x_0 \in X$ *si, et seulement si,*

$$\lim_{x \to x_0} \mathbb{H}(\Gamma(x_0), \Gamma(x)) = 0.$$

Définition A.3.7. *Une application multivoque* $\Gamma : X \rightrightarrows Y$ *est dite compacte au point* $x \in X$ *si, et seulement si, pour chaque suite* $(x_k, y_k)_{k \in \mathbb{N}}$ *telle que* $x_k \in X$, $y_k \in \Gamma(x_k)$ *(pour chaque* k*) et* $x_k \to x$*, il existe une application strictement croissante* $\phi : \mathbb{N} \to \mathbb{N}$ *telle que* $y_{\phi(k)} \to y$ *et* $y \in \Gamma(x)$.

Remarque A.3.3. *Si* $\Gamma : X \rightrightarrows Y$ *est compacte au point* $x \in X$*, alors l'ensemble* $\Gamma(x)$ *est un ensemble compact.*

La proposition suivante regroupe plusieurs résultats liant continuité et continuité au sens de Hausdorff.

Proposition A.3.2. *Soit* $\Gamma : X \rightrightarrows Y$ *une application multivoque, et soit un point* $x_0 \in X$.

– *Si* Γ *est continue supérieurement en* x_0*, alors* Γ *est continue supérieurement en* x_0 *au sens de Hausdorff.*
– *Si* Γ *est continue supérieurement au sens de Hausdorff en* x_0 *et* $\Gamma(x_0)$ *est un ensemble compact, alors* Γ *est continue supérieurement en* x_0.
– *Si* Γ *est continue inférieurement en* x_0 *et* $\Gamma(x_0)$ *est un ensemble compact, alors* Γ *est continue inférieurement au sens de Hausdorff en* x_0.
– *Si* Γ *est continue inférieurement au sens de Hausdorff en* x_0*, alors* Γ *est continue inférieurement en* x_0.

Remarque A.3.4. *Soit* $x_0 \in X$*. Si l'ensemble* $\Gamma(x_0)$ *est compact, alors la continuité de l'application multivoque* $\Gamma : X \rightrightarrows Y$ *au point* x_0 *est équivalente à la continuité au sens de Hausdorff en* x_0.

Annexe B

Code Matlab

Commençons par rappeler l'un des Problèmes d'Optimisation post-Pareto Stochastique étudié numériquement au cours de notre seconde contribution (Chapitre 3).

Nous avons considéré le Problème d'Optimisation Stochastique bi-Objectifs :

$$(SBOP) \qquad \mathrm{MIN}_{w\in X}\mathbb{E}_\xi\Big(F(w,\cdot)\Big),$$

où $w \in \mathbb{R}^4$, $\xi = (\xi^1, \xi^2) \in \mathbb{R}^2$ est tel que ξ^1 et ξ^2 suivent toutes deux une loi uniforme de moyenne 1 et de variance $\frac{1}{3}$, l'ensemble admissible est

$$X := \Big\{ w = (w_1, w_2, w_3, w_4) \in \mathbb{R}^4 \Big| \|w\| \leq 1, \sum_{i=1}^{4} w_i \geq 0 \Big\},$$

et la fonction objectif du $(SBOP)$ est donnée par

$$F(w,\xi) = \Big(\xi^1(2\exp(w_1+w_2) + (w_3+w_4)^2), \xi^2(4(w_1+w_2)^2 + \frac{1}{2}\exp(w_3+w_4))\Big).$$

Nous nous sommes intéressés au Problème d'Optimisation post-Pareto Stochastique (O) qui vise à minimiser la fonction $f : \mathbb{R}^2 \to \mathbb{R},\ z = (x,y) \mapsto (x+y)^3$ sur l'ensemble de Pareto associé au $(SBOP)$ précédent. Autrement dit, nous nous sommes posés le problème

$$(O) \quad \min \bigg(\Big(\mathbb{E}_\xi\Big[F^1(w,\cdot)\Big] + \mathbb{E}_\xi\Big[F^2(w,\cdot)\Big]\Big)^3 \bigg| w \in \mathrm{ARGMIN}_{\mathbb{R}^2_+}\mathbb{E}_\xi[F(X,\cdot)]\bigg).$$

Dans cette annexe, nous fournissons le code Matlab utilisé pour donner une approximation de la valeur optimale du problème (O).

```matlab
%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% ENSEMBLE ADMISSIBLE %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [c,ceq]=constraint(x)
c1=(x(1)-0)^2+(x(2)-0)^2+(x(3)-0)^2+(x(4)-0)^2-1;%inégalité
c2=-x(1)-x(2)-x(3)-x(4);%inégalité
c=[c1;c2];
ceq=[];%égalité
end

%%%%%%%%%%%%%%%%%%
%%% OBJECTIF 1 %%%
%%%%%%%%%%%%%%%%%%
function f=objfun1(x)
f=2*exp(x(1)+x(2))+(x(3)+x(4))^2;
end

%%%%%%%%%%%%%%%%%%
%%% OBJECTIF 2 %%%
%%%%%%%%%%%%%%%%%%
function f=objfun2(x)
f=1/2*exp(x(3)+x(4))+4*(x(1)+x(2))^2;
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% FONCTION DU DECIDEUR %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [out]=f(y)
out=(y(1)+y(2))^3;
end


%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% FONCTION PRINCIPALE %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [ outD outS approx Sareduire] = main(t, epsilon, N1, N2)

%CALCUL DES COORDONNES DE y0
x0=[0;0;0;0];
[x,fval]=fmincon('objfun1',x0,[],[],[],[],[],[],'constraint');
y01=fval*N1;
y02=objfun2(x)*N2;
y0=[y01;y02];

%CALCUL DES COORDONNES DE y1
[x,fval]=fmincon('objfun2',x0,[],[],[],[],[],[],'constraint');
y12=fval*N2;
y11=objfun1(x)*N1;
y1=[y11;y12];

%CALCUL DES COORDONNES DES yt
yt1=(1-t)*y01+t*y11;
yt2=(1-t)*y02+t*y12;
yt=[yt1 ; yt2];
```

```
%CALCUL DES COORDONNES DES yt*
ytetoile=y0;
A=[0 0 0 0 0 0 0;0 0 0 0 0 0 0;
    0 0 0 0 0 0 0;0 0 0 0 0 0 0;
    0 0 0 0 0 0 0;0 0 0 0 0 1 0;
    0 0 0 0 0 0 1];%pour fixer yt
Aeq=[]
beq=[]
lb=[]
ub=[]
for k=2:(length(t)-1)
    wt=yt(:,k);
    x0=[0;-0.1;0.1;-0.1;0.1;wt(1);wt(2)];
    b=[0 0 0 0 0  wt(1) wt(2)]';
    [x,fval]=fmincon('lambda',x0,A,b,Aeq,beq,lb,ub,'lambdacon');
    ytetoile=[ytetoile yt(:,k)+fval];
end
ytetoile=[ytetoile y1];

%CALCUL DES COORDONNES DES SIMPLEX
S=[];
for k=1:length(t)-1
  S=[S  [ytetoile(:,k)';ytetoile(:,k+1)';[ytetoile(1,k)  ytetoile(2,k+1)]] ];
end

%CALCUL DU MINIMUM DE f SUR LES POINTS DE DISCRETISATION
D=ytetoile;
fmintempD=f(D(:,1));
for k=2:length(t)
    if f(D(:,k))<fmintempD
        fmintempD=f(D(:,k));
    end
end
fmintempD;

%CALCUL DU MINIMUM DE f SUR LES SIMPLEX
fmintempS=zeros(length(t)-1);
fmintempS=fmintempS(1,:);
for k=1:length(t)-1
    fmintempS(k)=min(min(f(D(:,k)),f(D(:,k+1))),f([D(1,k) D(2,k+1)]));
end
fmintempS;

%CALCUL DU SIMPLEX A REDUIRE
Sareduire=1;
k=1;
while fmintempS(k)~=min(fmintempS)
    k=k+1;
end
Sareduire=k;

%CALCUL DE L'ERREUR D'APPROXIMATION
outS=min(fmintempS);
outD=fmintempD;
approx=outD-outS;
```

```
end


%%%%%%%%%%%%%%%%%%%%%%%
%%% FONCTION LAMBDA %%%
%%%%%%%%%%%%%%%%%%%%%%%
%pour résoudre les problèmes (CPt)
function [f,g] = lambda(w)
t = w(1);
f = t;%f=lambda
if nargout>1
     g=1;%sa dérivée
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% CONTRAINTES SUR LAMBDA %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [c,ceq,gc,gceq] = lambdacon(w)
c=[w(1) ;%LAMBDA<=0
    w(2)^2+w(3)^2+w(4)^2+w(5)^2-1;%ENSEMBLE ADMISSIBLE
    -w(2)-w(3)-w(4)-w(5);%ENSEMBLE ADMISSIBLE
    2*exp(w(2)+w(3))+(w(4)+w(5))^2-w(1)-w(6);%RESTER DANS LES IMAGES OBJECTIF1
    1/2*exp(w(4)+w(5))+4*(w(2)+w(3))^2-w(1)-w(7)];%RESTER DANS LES IMAGES 
OBJECTIF2
ceq=[];
if nargout>2%DERIVEES
     gc=[1    0         0       -1                -1;
         0    2*w(2)    -1       2*exp(w(2)+w(3))  8*(w(2)+w(3));
         0    2*w(3)    -1       2*exp(w(2)+w(3))  8*(w(2)+w(3));
         0    2*w(4)     1       2*(w(4)+w(5))     1/2*exp(w(4)+w(5));
         0    2*w(5)    -1       2*(w(4)+w(5))     1/2*exp(w(4)+w(5));
         0    0         0        -1                0;
         0    0         0         0                -1];
     gceq=[];
end


%%%%%%%%%%%%%%%%%%%%%%
%%% INITIALISATION %%%
%%%%%%%%%%%%%%%%%%%%%%
clear all
tic
eta=0.05;%ERREUR TOLEREE
epsilon=eta/2;%ERREUR DE L'ALGORITHME
p0=0.99;%NIVEAU DE CONFIANCE
N=32*600*sqrt(2)*1/3*1/(eta^2*(1-p0));%RANG N0
ro=0;%CORRELATION ENTRE LES VARIABLES ALEATOIRES
X=2*rand(2,N);%LOI DES VARIABLES ALEATOIRES [0;2]
cov=[1 ro;ro 1];%MATRICE DE COVARIANCE
A=chol(cov);
N=(A'*X);%V.A. CORRELEES
N1=mean(N(1,:));%MOYENNE DE LA PREMIERE V.A.
N2=mean(N(2,:));%MOYENNE DE LA SECONDE V.A.
t=[0 1/2 1];%INITIALISATION DE LA DISCRETISATION
[outD outS approx Sareduire]=main(t, epsilon, N1, N2);%INITIALISATION DE LA 
BOUCLE
```

```
%RESOLUTION DU PROBLEME
while approx>eta/2
    t=sort([t  (t(Sareduire)+t(Sareduire+1))/2]);
    [outD outS approx Sareduire]=main(t,epsilon,N1,N2);
end
outD%VALEUR DE f SUR LA DISCRETISATION
outS%VALEUR DE f SUR LES SIMPLEX
approx%VERIFICATION DE L'APPROXIMATION CHOISIE
nb_etape=length(t)-1%NOMBRE D'ITERATIONS

toc

%%%%%%%%%%%%%%%%%
%%% GRAPHIQUE %%%
%%%%%%%%%%%%%%%%%

z=boule_d(4);%SIMULATION DE LA BOULE UNITE POUR LE GRAPHIQUE
x=z;
%CALCUL DES COORDONNES DES POINTS DE L'IMAGE F(X)
f1p=[];
f2p=[];
for k=1:length(x)
    if sum(x(k,:))>=0
        f1p=[f1p objfun1(x(k,:))];
        f2p=[f2p objfun2(x(k,:))];
    else
        f1p=[f1p 0];
        f2p=[f2p 0];
    end
end
f1p=f1p*N1;
f2p=f2p*N2;


%CALCUL DES COORDONNES DE y0
x0=[0;0;0;0];
[x,fval]=fmincon('objfun1',x0,[],[],[],[],[],[],'constraint');
y01=fval*N1;
y02=objfun2(x)*N2;
y0=[y01;y02];

%CALCUL DES COORDONNES DE y1
[x,fval]=fmincon('objfun2',x0,[],[],[],[],[],[],'constraint');
y12=fval*N2;
y11=objfun1(x)*N1;
y1=[y11;y12];

%CALCUL DES COORDONNES DES yt
yt1=(1-t)*y01+t*y11;
yt2=(1-t)*y02+t*y12;
yt=[yt1 ; yt2];

%CALCUL DES COORDONNES DES yt*
ytetoile=y0;
```

```
A=[0 0 0 0 0 0 0;0 0 0 0 0 0 0;
    0 0 0 0 0 0 0;0 0 0 0 0 0 0;
    0 0 0 0 0 0 0;0 0 0 0 0 1 0;
    0 0 0 0 0 0 1];
Aeq=[]
beq=[]
lb=[]
ub=[]
for k=2:(length(t)-1)
    wt=yt(:,k);
    x0=[0;-0.1;0.1;-0.1;0.1;wt(1);wt(2)];
    b=[0 0 0 0 0  wt(1) wt(2)]';
    [x,fval]=fmincon('lambda',x0,A,b,Aeq,beq,lb,ub,'lambdacon');
    ytetoile=[ytetoile yt(:,k)+fval];
end
ytetoile=[ytetoile y1];

%CALCUL DES COORDONNES DES SIMPLEX
S=[];
for k=1:length(t)-1
  S=[S  [ytetoile(:,k)';ytetoile(:,k+1)';[ytetoile(1,k)  ytetoile(2,k+1)]] ];
end

%CALCUL DU MINIMUM DE f SUR LES POINTS DE DISCRETISATION
D=ytetoile;
fmintempD=f(D(:,1));
for k=2:length(t)
    if f(D(:,k))<fmintempD
        fmintempD=f(D(:,k));
    end
end
fmintempD;

%CALCUL DU MINIMUM DE f SUR LES SIMPLEX
fmintempS=zeros(length(t)-1);
fmintempS=fmintempS(1,:);
for k=1:length(t)-1
    fmintempS(k)=min(min(f(D(:,k)),f(D(:,k+1))),f([D(1,k) D(2,k+1)]));
end
fmintempS;

%AFFICHAGE DES RESULTATS
outS=min(fmintempS)
outD=fmintempD
approx=outD-outS

%GRAPHIQUE
figure
T=[1,2,3];
hold on
axis equal
grid off

for k=1:2:length(S)
    patch('Vertices', S(:,k:k+1), 'Faces', T, 'EdgeColor', 'k', 'FaceColor',
'w')
```

```
end
plot(f1p,f2p,'.red')

plot([y0(1) y1(1)]',[y0(2) y1(2)]','k')
for k=2:length(t)-1
   plot([yt(1,k) ytetoile(1,k)]', [yt(2,k) ytetoile(2,k)]','k')
end
for k=1:length(t)
   plot( yt(1,k),yt(2,k),'k+')
end

hold off
```

Table des figures

Bibliographie

[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] E.M. Bednarczuk and J. Przybyla, *The vector-valued variational principle in Banach spaces ordered by cones with nonempty interiors,* SIAM J. Optim., **18(3)**, 907-913, 2007.

[6] H.P. Benson, *Existence of efficient solutions for vector maximization problems,* J. Optim. Theory Appl., **26(4)**, 569-580, 1978.

[7] H.P. Benson, *Optimization over the Efficient Set,* J. Math. Anal. Appl., **98**, 562-580, 1984.

[8] 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.

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

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

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

[12] H.P. Benson, *An outer approximation algorithm for generating all efficient extreme points in the outcome set of a multiple objective linear programming problem,* J. Global Optim., **13(1)**, 1-24, 1998.

[13] 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.

[14] H. P. Benson, *Simplicial Branch and Reduce Algorithm for Convex Programs with a Multiplicative Constraint,* J. Optim. Theory Appl., **145**, 213-233, 2010.

[15] H. P. Benson, *Existence of efficient solutions for vector maximisation problems,* J. Optim. Theory Appl., **26(4)**, 569-580, 1978.

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

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

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

[19] S. Bolintinéanu and M. El Maghri, *Pénalisation dans l'optimisation sur l'ensemble faiblement efficient,* Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle, tome 31 (3), 295-310, 1997.

[20] S. Bolintinéanu, *Vector Variational Principles ; ϵ-Efficiency and Scalar Stationarity,* J. Convex Anal. **8**, 71-85, 2001.

[21] J.F. Bonnans, R. Cominetti and A. Shapiro, *Sensitivity analysis of optimization problems under second order regular constraints,* Math. Oper. Res., **23(4)**, 806-831, 1998.

[22] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problem,* Springer Series in Operations Research, 2000.

[23] H. Bonnel and J. Collonge, *Stochastic Optimization over a Pareto set associated with a Stochastic Multiobjective Optimization Problem,* J. Optim. Theory Appl., **162(2)**, 405-427, 2014.

[24] H. Bonnel and J. Collonge, *Optimization over the Pareto Outcome set associated with a Convex Bi-Objective Optimization Problem : Theoretical Results, Deterministic Algorithm and Application to the Stochastic case,* to appear in J. Global Optim.

[25] H. Bonnel and C.Y. Kaya, *Optimization Over the Efficient Set of Multi-objective Convex Optimal Control Problems,* J. Optim. Theory Appl. **147(1)**, 93-112, 2010.

[26] H. Bonnel and N.S. Pham, *Nonsmooth Optimization over the (weakly or properly) Pareto set of a linear-quadratic multiobjective Control problem : Explicit optimality Conditions,* J. Industrial Management Optim., **7(4)**, 789-809, 2011.

[27] H. Bonnel, J. Morgan, *Semivectorial Bilevel Convex Optimal Control Problems : Existence Results,* SIAM J. Control Optim., **50(6)**, 3224-3241, 2012.

[28] H. Bonnel, J. Morgan, *Optimality Conditions for Semivectorial Bilevel Convex Optimal Control Problems,* Computational and Analytical Mathematics In Honor of Jonathan Borwein's 60th Birthday, Chapter : 4, Publisher : Springer, Editors : David H. Bailey, Heinz H. Bauschke, Peter Borwein, Frank Garvan, Michel Théra, Jon D. Vanderwerff, Henry Wolkowicz, pp.43-74, ISBN : 978-1-4614-7620-7, 2013.

[29] H. Bonnel, *Optimality Conditions for Semivectorial Bilevel Optimization Problem*, Pacific Journal of Optimization, **2(3)**, 447-468, 2006.

[30] H. Bonnel, *Remarks about approximate solutions in vector optimization*, Pacific Journal of Optimization, **5(1)**, 53-73, 2009.

[31] H. Bonnel, J. Morgan, *Semivectorial Bilevel Optimization Problem : Penalty Approach,* J. Optim. Theory Appl., **131(3)**, 365-382,2006.

[32] H. Bonnel *Post-Pareto analysis for multiobjective parabolic control systems,* www.optimization-online.

[33] J.M. Borwein, *Proper efficient points for maximizations with respect to cones,* SIAM J. Control Optim., **15**, 57-63, 1977.

[34] J.M. Borwein, *On the existence of Pareto efficient points,* Math. Oper. Res., **8(1)**, 64-73, 1983.

[35] R.S. Burachik and A.N. Iusem, *Set-Valued Mappings and Enlargements of Monotone Operators,* Springer Optimization and Its Applications, 2008.

[36] R.S. Burachik, and M.M. Rizvi, *On weak and strong Kuhn-Tucker optimality conditions for smooth multiobjective optimization,* J. Optim. Theory Appl., **155(2)**, 477-491, 2012.

[37] R.S. Burachik, C.Y. Kaya and M.M. Rizvi, *A new Scalarization Technique to Approximate Pareto Fronts of Problems with Disconnected Feasible Sets,* J. Optim. Theory Appl., **162**, 428-446, 2014.

[38] V. Chankong and Y. Y. Haimes, *Multiobjective decision making,* Volume 8 of North Holland Series in system science and engineering. Elsevier, Amsterdam, 1983.

[39] F. H. Clarke, *Optimization and Nonsmooth Analysis,* Wiley Interscience, New York, 1983.

[40] B. Colson, P. Marcotte and G. Savard, *Bilevel programming : a survey,* 4OR, **3(2)**, 87-107, 2005.

[41] B. Colson, P. Marcotte and G. Savard, *An overview of bilevel optimization,* Ann. Oper. Res., **153**, 235-256, 2007.

[42] B. D. Craven, *Aspects of multicriteria optimization,* Recent Developments in Mathematical Programming, 93-100, 1991.

[43] H.W. Corley, *An existence result for maximization with respect to cones,* J. Optim. Theory Appl., **31(2)**, 277-281, 1980.

[44] H.W. Corley, *Optimality conditions for maximization of set-valued functions,* J. Optim. Theory Appl., **58**, 1-10, 1988.

[45] G. B. Dantzig, *Linear programming and extensions,* Princeton University Press, Princeton, NJ, 1998.

[46] J. P. Dauer, *Optimization over the efficient set using an active constraint approach,* J. Oper. Res., **35**, 185-195, 1991.

[47] J. P. Dauer and T. A. Fosnaugh, *Optimization over the efficient set,* J. Global Optim., **7**, 261-277, 1995.

[48] S. Dempe, *Foundations of bilevel programming of Non-convex Optimization and Its Applications,* Kluwer, Dordrecht, **61**, 2002.

[49] S. Dempe, J. Dutta and B. S. Mordukhovich, *New necessary optimality conditions in optimistic bilevel programming,* Optimization, **56(5-6)**, 577-604, 2007.

[50] 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*

[51] S. Deng, *On efficient solutions in vector optimization,* J. Optim. Theory Appl., **96(1)**, 201-209,1998.

[52] J. Dieudonné, *Foundations of modern analysis,* Academic Press, New York and London, 1960.

[53] J. Dutta and C.Y. Kaya, *A new scalarization and numerical method for constructing the weak Pareto front of multi-objective optimization problems,* Optimization **60**, 1091-1104, 2011.

[54] F. Y. Edgeworth, *Mathematical Psychics,* University Microfilms International, 1987.

[55] M. Ehrgott, *Multicriteria Optimization, Classification and Methodology,* Shaker Verlag, Aachen, 1997.

[56] M. Ehrgott, *Multicriteria Optimization,* Lect. Notes Econ. Maths. Syst. Springer Berlin, 2000.

[57] M. Ehrgott, *Multicriteria Optimization,* 2nd edition, 2005.

[58] M. Ehrgott, *An approximation algorithm for convex multi-objective programming problems,* J. Global Optim., **50**, 397-416, 2011.

[59] G. EICHFELDER, *Scalarization for adaptively solving multiobjective optimization problems* , Comput. Optim. Appl., 2007.

[60] G. EICHFELDER, *Adaptative Scalarization Methods in Multiobjective Optimization,* Springer-Verlag Berlin Heidelberg, 2008.

[61] G. EICHFELDER , *Multiobjective bilevel optimization,* Math. Program. Ser., **213**, 419-449, 2010.

[62] J. ENGWERDA, *Necessary and sufficient conditions for Pareto optimal solutions of cooperative differential games,* SIAM J. Control Optim., **48(6)**, 3859-3881, 2010.

[63] A. V. FIACCO, *Sensitivity analysis for nonlinear programming using penalty methods,* Math. Program., **10**, 287-311, 1976.

[64] A. V. FIACCO, *Introduction to Sensitivity and Stability analysis in nonlinear programming,* Mathematics in Science and Engineering, Academic Press, **165**, London 1983.

[65] J. FLIEGE, *An efficient interior point method for convex multicriteria optimization problems,* Math. Oper. Res., **31(4)**, 825-845, 2006.

[66] J. FLIEGE AND R. WERNER, *Robust multiobjective optimization and applications in portfolio optimization,* European Journal of Operational Research **234(2)**, 422-433, 2014.

[67] J. FLIEGE AND L. N. VICENTE, *Multicriteria approach to bilevel optimization* J. Optim. Theory Appl. **131(2)**, 209-225, 2006.

[68] J. FLIEGE AND H. XU, *Stochastic Multiobjective Optimization : Sample Average Approximation and Applications,* J. Optim. Theory Appl. **151**, 135-162, 2011.

[69] F. FLORES-BAZAN, S. LAENGLE AND G. LOYOLA, *Characterizing the efficient points without closedness or free-disposability,* Central European Journal of Operations Research, **21(2)**, 401-410,2013.

[70] 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.

[71] A. M. GEOFFRION, *Proper efficiency and the theory of vector maximization* J. Math. Anal. Appl., **22**, 618-630, 1968.

[72] D. GOURION AND D. T. LUC, *Generating the weakly efficient set of non-convex multi-objective problems* J. Global Optim., 2007.

[73] A. GÖPFERT, H. RIAHI, C. TAMMER AND C. ZĂLINESCU, *Variational Methods in Partially Ordered Spaces,* Springer, Berlin, 2003.

[74] GUANG-YA CHEN, XUEXIANG HUANG AND XIAOQI YANG, *Vector Optimization : Set Valued and Variational Analysis,* Springer, Berlin, 2005.

[75] C.Y. Kaya and H. Maurer, *A numerical method for non convex multi-objective optimal control problems,* Comput. Optim. Appl. **57**, 685-702, 2014.

[76] M. I. Henig, *Proper efficiency with respect to cones,* J. Optim. Theory Appl. **36**, 387-407, 1982.

C. Hillermeier and J. Jahn, *Multiobjective optimization : survey of methods and industrial applications,* Surv. Math. Ind., **11**, 1-42, 2005.

[77] 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.

[78] 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.

[79] R. Horst and H. Tuy, *Global Optimization : Deterministic Approaches,* Springer, Berlin, 1990.

[80] C. L. Hwang and A. S. M. Masud, *Multiple objective decision making - Methods and applications,* Lect. Notes Econ. Maths. Syst., **164**, Springer Berlin, 1979.

[81] H. Isermann, *Proper Efficiency and the Linear Vector Maximization Problem,* Oper. Res., **22**, 189-191, 1974.

[82] H. Isermann, *The enumeration of the set of all efficient solutions for a linear multiple objective program,* Oper. Res. Quaterly, **28(3)**, 711-725, 1977.

[83] J. Jacod, P. Protter, *Probability Essentials,* Springer-Verlag Berlin Heidelberg, 2004.

[84] J. Jahn, *Scalarization in vector Optimization,* Math. Program., **29**, 203-218, 1984.

[85] J. Jahn, *A characterization of properly minimal elements of a set,* SIAM J. Cont. Opti., **23**, 649-656, 1985.

[86] J. Jahn, *Mathematical Vector Optimization in Partially Ordered Spaces,* Methoden und Verfarhen der mathematischen Physik, Peter Lang, Frankfurt and Main **31**, 1986.

[87] J. Jahn, *Introduction to the theory of nonlinear optimization,* Springer, Berlin, 1994.

[88] J. Jahn, *Vector Optimization : Theory, Applications and Extensions,* Springer, Berlin, 2004.

[89] J. Jahn, *Introduction to the Theory of Nonlinear Optimization,* Springer-Verlag, Berlin, Heidelberg, New York, 2007.

[90] N. T. B. KIM AND T. N. THANG, *Optimization over the Efficient Set of a Bicriteria Convex Programming Problem,* P. J. Optim., **9**, 103-115, 2013.

[91] N. T. B. KIM AND O. DE WECK, *Adaptative weight sum method for bi-objective optimization,* Structural and Multidisciplinary Optimization, **29** 149-158, 2005.

[92] 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.

[93] G.-H. LIN AND M. FUKUSHIMA, *Stochastic equilibrium problems and stochastic mathematical, Pacific Journal of Optimization,* **6(3)**, *455-482, 2010.*

[94] P. LORIDAN, ϵ-solutions in vector minimization problems, *J. Optim. Theory Appl.,* **43** *265-276, 1984.*

[95] P. LORIDAN AND J. MORGAN, Convergence of Approximate Solutions and Values in Parametric Vector Optimization, *Non-convex Optimization and Its Applications,* **38**, *335-350, 2000.*

[96] D. T. LUC, Connectedness of the efficient sets in quasi-concave vector optimization, *J. Math. Anal. Appl.,* **122** *346-354, 1987.*

[97] D. T. LUC, Theory of Vector Optimization, *Lecture Notes in Econom. and Math. Systems 319, Springer-Verlag, Berlin, Heidelberg, 1989.*

[98] M. E. MAGHRI AND M. LAGHDIR, Pareto differential calculus for convex vector mappings and applications to vector optimization, *SIAM J. Optim.,* **19(4)**, *1970-1994, 2009.*

[99] M. AIT MANSOUR, C. MALIVERT AND M. THÉRA, Semicontinuity of vector-valued mappings, *Optimization,***56 (1-2)**, *241-252, 2007.*

[100] K. M. MIETTINEN, Nonlinear Multiobjective Optimization, *International series in Operations Research and Management Science, Kluwer Academic Publishers, Dordrecht, 1999.*

[101] A. PASCOLETTI AND P. SERAFINI, Scalarizing vector optimization problems, *J. Optim. Theory Appl.,* **42(4)** , *499-524, 1984.*

[102] J. PHILIP, Algorithms for the vector maximization problem, *Math. Program., 207-229, 1972.*

[103] R.T. ROCKAFELLAR, Convex Analysis, *Princeton, New Jersey, Princeton University Press, 1970.*

[104] R.T. ROCKAFELLAR, ROGER J-B WETS, Variational Analysis, *Springer, Heidelberg, London, New York,* **317**, *1998.*

[105] Y. Sawaragi, H. Nakayama and T. Tanino*,* Theory of Multiobjective Optimization, *Number 176 in Mathematics in Science and Engineering. Academic Press, London, 1985.*

[106] A. Shapiro, D. Dentcheva, A. Ruszczynski*,* Lectures on stochastic programming : modeling and theory, *MPS/SIAM Ser. Optim., 2009.*

[107] R. M. Soland*,* Multicriteria optimization : A general characterization of efficient solutions, *Decision science,* **10***, 26-38, 1979.*

[108] R. E. Steuer*,* Multiple criteria optimization : Theory, computation and application, *John Wiley Sons, New York, 1985.*

[109] J. Stoer and C. Witzgall*,* Convexity and Optimization in Finite dimensional Spaces, Springer, Berlin, *1970.*

[110] T. Tanaka*,* Approximately Efficient solutions in Vector Optimization, *J. Multi-Criteria Decision Analysis,* **5***, 271-278, 1996.*

[111] S. Vogel*,* A stochastic approach to stability in stochastic programming, *J. of Comp. and Applied Math.,* **56***, 65-96, 1994.*

[112] R. E. Wendell and D. N. Lee*,* Efficiency in multiple objective optimization problems, *Math. Program.,* **312***, 406-414, 1977.*

[113] Y. Yamamoto*,* Optimization over the efficient set : an overview, *J. Global Optim.,* **22***, 285-317, 2002.*

[114] P. L. Yu*,* Multiple-Criteria Decision Making : Concepts, Techniques, and Extensions, *Plenum Press, New York and London, 1985.*

[115] X. Zengkun*,* Necessary conditions for suboptimization over the weakly efficient set associated to generalized invex multiobjective programming, *J. Math. Analysis and Appl. Comput.,* **201***, 502-515, 1996.*

[116] X.Y. Zheng and K.F. Ng*,* The Lagrange multiplier rule for multifunctions in Banach spaces, *SIAM J. Optim.,* **17(4)***, 1154-1175, 2006.*

[117] X.Y. Zheng and Z. Wan*,* A solution method for semivectorial bilevel programming problem via penalty method, *J. Appl. Math. Comput.*, **37**, 207-219, 2011.

[118] X.Y. Zheng and Z. Wan, *Hamiltonian necessary conditions for a multiobjctive optimal control problem with endpoint constraints,* SIAM J. Control Optim., **39(1)**, 97-112, 2000.