

Stochastic Approximation in Convex Multiobjective Optimization

Carlo Alberto De Bernardi, Enrico Miglierina

*Dipartimento di Matematica per le Scienze Economiche, Finanziarie ed Attuariali,
Università Cattolica del Sacro Cuore, Milano, Italy
carloalberto.debernardi@unicatt.it, enrico.miglierina@unicatt.it*

Elena Molho

*Dipartimento di Scienze Economiche e Aziendali, Università degli Studi di Pavia, Italy
elena.molho@unipv.it*

Jacopo Somaglia

*Politecnico di Milano, Dipartimento di Matematica, Milano, Italy
jacopo.somaglia@polimi.it*

Received: March 6, 2023

Accepted: August 21, 2023

Given a strictly convex multiobjective optimization problem with objective functions $f_1, \dots, f_N$, let us denote by x_0 its solution, obtained as minimum point of the linear scalarized problem, where the objective function is the convex combination of $f_1, \dots, f_N$ with weights $t_1, \dots, t_N$. The main result of this paper gives an estimation of the averaged error that we make if we approximate x_0 with the minimum point of the convex combinations of n functions, chosen among $f_1, \dots, f_N$, with probabilities $t_1, \dots, t_N$, respectively, and weighted with the same coefficient $1/n$. In particular, we prove that the averaged error considered above converges to 0 as n goes to ∞ , uniformly w.r.t. the weights $t_1, \dots, t_N$. The key tool in the proof of our stochastic approximation theorem is a geometrical property, called by us small diameter property, ensuring that the minimum point of a convex combination of the functions $f_1, \dots, f_N$ continuously depends on the coefficients of the convex combination.

Keywords: Multiobjective optimization, continuity of solution map, convex combinations of convex functions, small diameter property.

2020 Mathematics Subject Classification: Primary 90C29, 46N10; secondary 90C25.

1. Introduction

The main aim of the present paper is to develop a result about the approximation of solutions of a convex multiobjective optimization problem in the spirit of [2, Theorem 1, p.28]. In that paper the authors provided a result whose meaning was explained by the authors themselves (the quotation is translated in English since the original paper is written in French, see [2, Remark at p. 30]):

The research of the authors has been partially supported by the GNAMPA (INdAM – Istituto Nazionale di Alta Matematica). C. A. De Bernardi, E. Miglierina and E. Molho have been partially supported by the Ministry for Science and Innovation, Spanish State Research Agency (Spain), under project PID2020-112491GB-I00.

“This theorem means that to calculate a Pareto Optimum for the functions $f_1, \dots, f_N$, it suffices, approximately, to calculate it for n of them, and, better, most of the choices of n elements among the N will give a good result.”

In order to make clear the comment above, we briefly describe what is a multiobjective optimization problem. For a complete overview about this field we refer the reader to the, now classical, monographs [10] and [9]. Let $f_1, \dots, f_N$ be N functions from a normed space X to $\mathbb{R}$. A point $x_0 \in X$ is a Pareto Optimum point (or efficient point) for $f_1, \dots, f_N$ when there is no $x \in X$ such that $f_k(x) \leq f_k(x_0)$ for $k = 1, \dots, N$ and $f_i(x) < f_i(x_0)$ for some $i \in \{1, \dots, N\}$. One of the most common techniques to find the Pareto Optimum points is the so called linear scalarization, i.e., solving a family of scalar optimization problems where the objective function is given by a convex combination of the functions $f_1, \dots, f_N$ (see, e.g., [10, Ch.4] and [9, Ch.5]). In our paper we deal with a special case of linear scalarization for a multiobjective optimization problem where, under convexity assumptions on the functions $f_1, \dots, f_N$, each Pareto Optimum is completely characterized as minimum point of a convex combination of the functions $f_1, \dots, f_N$. Namely, if the functions $f_1, \dots, f_N$ are strictly convex, a point x_0 is a Pareto Optimum point for $f_1, \dots, f_N$ if and only if there exists $t = (t_1, \dots, t_N) \in \Sigma_N := \{t \in [0, 1]^N : \sum_{i=1}^N t_i = 1\}$ such that x_0 is a minimum point for the scalar function $u_t(x) = \sum_{k=1}^N t_k f_k(x)$ (see [2, Lemma 1, p. 27]).

The relevance of linear scalarization in the theory and practice of multiobjective optimization motivates the study of the properties of a convex combination of functions. Moreover, as we will see, in order to obtain a result in the spirit of [2], a key tool will be a geometrical property ensuring that the minimum point of a convex combination of continuous strictly convex functions $f_1, \dots, f_N$ continuously depends on the coefficients of the convex combination.

The approach developed by Enflo and Beauzamy in [2] is deeply original with respect to the field of multiobjective optimization and gives an unusual viewpoint about the approximation of solution set. Despite these interesting features, the paper [2] seems not to have been widely considered in the literature about multiobjective optimization. For this reason, we are interested in studying this result and try to weaken its assumptions. The key point of the approach of [2] is to study the Pareto Optimum points of a convex multiobjective optimization problem with objective functions $f_1, \dots, f_N$, by identifying these points with the solutions of minimum points of convex combinations of $f_1, \dots, f_N$. In particular, we need to ensure the continuity of the map $\phi(t) = \operatorname{argmin}(u_t)$ where $t \in \Sigma_N$ and u_t is defined as above. Moreover, in view of linear scalarization procedure, this type of result is interesting in itself.

In the present paper we restrict our attention to the case where the functions $f_1, \dots, f_N$ are such that, for every $k = 1, \dots, N$:

- (a) f_k is a continuous strictly convex function that is bounded on bounded sets;
- (b) f_k is coercive (i.e. for each $C \in \mathbb{R}$ the set $\{x \in X : f(x) \leq C\}$ is bounded).

Under these assumptions, it is quite straightforward to show that the function ϕ is continuous whenever X is a finite-dimensional normed space. On the other hand, when X is infinite-dimensional, it may happen that ϕ is not continuous.

Indeed, we provide an example of two functions defined on the Hilbert space ℓ_2 and satisfying assumptions (a)–(b) above, such that the corresponding function ϕ is not continuous (see Example 3.7). In the infinite-dimensional setting, the continuity of ϕ can be obtained by adding a property, already introduced in [3] and [12] under different terminologies, called by us *small diameter property*. Section 2 is mainly devoted to the study of the small diameter property. It is worth pointing out that in Proposition 2.4 we prove a result that allows to lift strongly exposed point from the norm of X to a function $f: X \rightarrow \mathbb{R}$, hence providing a good tool to find examples of functions satisfying the small diameter property. In Section 3, it is shown that ϕ is continuous if the functions $f_1, \dots, f_N$ satisfy the small diameter property and additionally properties (a)–(b) mentioned above. Finally, Section 4 provides a “stochastic” approximation of Pareto Optimum points for $f_1, \dots, f_N$ not depending on the choice of the value t associated to each Pareto Optimum (Theorem 4.4). This theorem can be applied to a larger setting than the result of [2], as proved in Remark 4.8, where, by using special norms built up in [4], we prove that there exists a set of functions $f_1, \dots, f_N$ not satisfying the assumptions of [2, Theorem 1, p.28] but enjoying the conditions required by our result. Moreover, an example shows that the small diameter property assumption cannot be dropped.

2. Notation and preliminaries

In this section we introduce some notions and basic results that we shall need in the sequel of the paper. The section is divided in two parts: the former is devoted to the study of geometrical properties of convex sets and functions, the latter to a brief introduction to multiobjective optimization. We consider only nontrivial real normed linear spaces. If X is a normed space with topological dual X^* , then B_X , U_X , and S_X are the closed unit ball, the open unit ball, and the unit sphere of X , respectively.

2.1. Geometrical properties of convex sets and functions

We are going to introduce several notions that are widely used in convex analysis and in the study of geometry of Banach spaces (see, e.g., [1, 5, 8, 11] and the references therein). Let D be a convex subset of a normed space X . A point $x \in D$ is *supported* by a linear functional $x^* \in X^* \setminus \{0\}$ if $x^*(x) = \sup x^*(D)$.

A point $x \in D$ is called an *exposed point* of D if x is supported by $x^* \in X^*$ and $\{y \in D: x^*(y) = x^*(x)\} = \{x\}$. A point $x \in D$ is *strongly exposed* by $x^* \in X^*$ if it is supported by x^* and $x_n \rightarrow x$ for all sequences $\{x_n\} \subset D$ such that we have $\lim_{n \rightarrow \infty} x^*(x_n) = \sup_D x^*$. In this case, x is also strongly exposed by λx^* , for any $\lambda > 0$. Moreover, a point $x \in D$ is strongly exposed by $x^* \in X^* \setminus \{0\}$ if and only if x^* is bounded over D and $\text{diam } S(x^*, \delta, D) \rightarrow 0$ as $\delta \rightarrow 0$, where

$$S(x^*, \delta, D) := \{y \in D: x^*(y) > \sup_D x^* - \delta\}.$$

Let $f: D \rightarrow \mathbb{R}$ be a convex continuous function. By

$$\text{epi}(f) := \{(x, t) \in D \times \mathbb{R}: t \geq f(x)\}$$

we denote the *epigraph* of the function f . Moreover, for $x \in D$, $\partial f(x)$ denotes the *subdifferential* of f at x . A point $(x, f(x)) \in \text{epi}(f)$ is supported by $(x^*, -1)$ if and only if $x^* \in \partial f(x)$.

Definition 2.1. Let D be a convex subset of X , $x \in D$, and $f: D \rightarrow \mathbb{R}$.

- Let $x^* \in \partial f(x)$, we say that x is an x^* -small diameter point for f if each sequence $\{x_n\} \subset D$, satisfying $x^*(x_n) - f(x_n) \rightarrow x^*(x) - f(x)$, is norm convergent.
- If, for each $x^* \in \partial f(x)$, the point x is an x^* -small diameter point for f , then we say that x is a *small diameter point* for f .
- f has the *small diameter property* (SDP) if each $y \in D$ is a small diameter point for f .

We recall the following result which relates small diameter points for f in X with strongly exposed points of $\text{epi}(f)$ in $X \times \mathbb{R}$.

Lemma 2.2. [3] *Let D be a convex subset of X , $f: D \rightarrow \mathbb{R}$ a continuous convex function, $x \in X$, and $x^* \in \partial f(x)$. Then x is an x^* -small diameter point for f if and only if $(x, f(x))$ is a strongly exposed point of $\text{epi}(f)$ by $(x^*, -1)$.*

We deduce that if $f: X \rightarrow \mathbb{R}$ has the SDP, then for each $x \in X$, $(x, f(x))$ is supported by $(x^*, -1)$ if and only if $(x, f(x)) \in \text{epi}(f)$ is strongly exposed by $(x^*, -1)$. We are going to show an elementary fact about the SDP; we shall need this result in the next section (cf. the proof of [12, Theorem 4.2]).

Proposition 2.3. *Let $f, g: X \rightarrow \mathbb{R}$ be convex continuous functions. If f has the SDP then $f + g$ has the SDP.*

Proof. Let $x_0 \in X$ and $x^* \in \partial(f + g)(x_0)$. We are going to show that x_0 is a x^* -small diameter point for $f + g$. Indeed, since $f + g$ is a convex function and $x^* \in \partial(f + g)(x_0)$, we have

$$x^*(y) - (f + g)(y) \leq x^*(x_0) - (f + g)(x_0),$$

for any $y \in X$. Now, let $\{x_n\} \subset X$ be a sequence such that $x^*(x_n) - (f + g)(x_n)$ converges to $x^*(x_0) - (f + g)(x_0)$. By [11, Theorem 3.23] there exists $x_1^* \in \partial f(x_0)$ and $x_2^* \in \partial g(x_0)$ such that $x^* = x_1^* + x_2^*$. Fix $\epsilon > 0$, then eventually we have

$$\begin{aligned} \epsilon &> x^*(x_0) - (f + g)(x_0) - x^*(x_n) + (f + g)(x_n) \\ &= x_1^*(x_0) - f(x_0) - x_1^*(x_n) + f(x_n) + x_2^*(x_0) - g(x_0) - x_2^*(x_n) + g(x_n) \\ &\geq x_1^*(x_0) - f(x_0) - x_1^*(x_n) + f(x_n) \geq 0. \end{aligned}$$

Since f has the SDP, the previous chain of inequalities and the arbitrariness of $\epsilon > 0$ show that $\{x_n\}$ converges to x_0 . Hence, $f + g$ has the SDP. $\square$

The next result shows, under suitable hypotheses, how to lift strongly exposed points from the norm of a normed space X to the graph of a function $f: X \rightarrow \mathbb{R}$. It provides a good tool for constructing examples of functions having the SDP.

Proposition 2.4. *Let $(X, \|\cdot\|)$ be a normed space, $x_0 \in S_X$, and suppose that $x_0 \in B_X$ is strongly exposed by $x^* \in S_{X^*}$. Let $h: [0, \infty) \rightarrow \mathbb{R}$ be a continuous convex strictly increasing function and let $\lambda \in (0, \infty)$ and $\lambda^* \in \mathbb{R}$, suppose that λ is a λ^* -small diameter point for h . Then:*

- (i) $h(\|\cdot\|)$ is a continuous convex function on X ;
- (ii) λx_0 is a $\lambda^* x^*$ -small diameter point for $h(\|\cdot\|)$;
- (iii) the origin of X is a 0-small diameter point for $h(\|\cdot\|)$, where $0 \in X^*$ denotes the null functional.

Proof. The first point is well-known and easy to prove. Let us prove (ii). Clearly, since h is strictly increasing, λ^* is positive. Hence, for each $x \in X$, we have

$$\lambda^* x^*(x) - h(\|x\|) \leq \lambda^*(x^*(x) - \|x\|) + \lambda^* \lambda - h(\lambda) \leq \lambda^* x^*(\lambda x_0) - h(\|\lambda x_0\|), \quad (1)$$

where the first inequality holds since $\lambda^* \in \partial h(\lambda)$. Now, let us consider a sequence $\{x_n\} \subset X$ and suppose that

$$\lambda^* x^*(x_n) - h(\|x_n\|) \rightarrow \lambda^* x^*(\lambda x_0) - h(\|\lambda x_0\|).$$

Since, for $n \in \mathbb{N}$,

$$\lambda^* x^*(x_n) - h(\|x_n\|) \leq \lambda^* \|x_n\| - h(\|x_n\|) \leq \lambda^* \lambda - h(\lambda),$$

we have that $\|x_n\| \rightarrow \lambda$, in particular $\|x_n\| > 0$ eventually. Moreover, by (1), we have: $\lambda^*(x^*(x_n) - \|x_n\|) \rightarrow 0$ and hence $x^*\left(\frac{x_n}{\|x_n\|}\right) \rightarrow 1$. Since $x_0 \in B_X$ is strongly exposed by $x^* \in S_{X^*}$, we have $\frac{x_n}{\|x_n\|} \rightarrow x_0$, and hence $x_n \rightarrow \lambda x_0$. The proof of (ii) is complete.

In order to prove (iii), we can suppose without any loss of generality that $h(0) = 0$. Then, it is sufficient to observe that, by the continuity of the inverse function of h , if $h(\|x_n\|) \rightarrow h(\|0\|) = 0$ then $\|x_n\| \rightarrow 0$. $\square$

A function $f: X \rightarrow \mathbb{R}$ is said *strictly convex* on a convex set $D \subset X$ if, for every $x, y \in D$, the following relation holds: $f\left(\frac{x+y}{2}\right) < \frac{1}{2}(f(x) + f(y))$. We now recall the stronger notion of uniformly convex function.

Definition 2.5. Let D be a convex subset of a normed space X , $x_0 \in D$ and $f: D \rightarrow \mathbb{R}$. Then we say that

- f is *uniformly convex at x_0* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every $y \in D$ such that $\|x_0 - y\| \geq \varepsilon$ it holds

$$f\left(\frac{x_0 + y}{2}\right) \leq \frac{1}{2}f(x_0) + \frac{1}{2}f(y) - \delta.$$

- f is *uniformly convex on D* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every $x, y \in D$ such that $\|x - y\| \geq \varepsilon$ it holds

$$f\left(\frac{x + y}{2}\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y) - \delta.$$

Remark 2.6. The following implications hold.

- (i) If a function $f: X \rightarrow \mathbb{R}$ is uniformly convex on a convex subset $D \subset X$, then it is uniformly convex at x , for each $x \in D$.
- (ii) If $f: X \rightarrow \mathbb{R}$ is uniformly convex at x , for each $x \in X$, then f has the SDP.
- (iii) If $f: X \rightarrow \mathbb{R}$ has the SDP then it is strictly convex.

Proof. The implication contained in (i) is trivial, let us prove (ii). Suppose on the contrary that f has not the SDP, then there exist $x \in X$, $x^* \in \partial f(x)$, $\varepsilon > 0$, and a sequence $\{x_n\} \subset X$ such that $x^*(x_n) - f(x_n)$ goes to $x^*(x) - f(x)$, when $n \rightarrow \infty$, but $\|x_n - x\| \geq \varepsilon$, whenever $n \in \mathbb{N}$. Since f is uniformly convex in x , there exists $\delta > 0$ such that

$$f\left(\frac{x+x_n}{2}\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(x_n) - \delta,$$

for each $n \in \mathbb{N}$. From which we get

$$2\delta + x^*(x) - f(x) + x^*(x_n) - f(x_n) \leq 2x^*\left(\frac{x+x_n}{2}\right) - 2f\left(\frac{x+x_n}{2}\right).$$

Hence, for sufficiently large n we get

$$\delta + 2[x^*(x) - f(x)] \leq 2\left[x^*\left(\frac{x+x_n}{2}\right) - f\left(\frac{x+x_n}{2}\right)\right].$$

The last inequality contradicts the fact that the point $(x, f(x)) \in \text{epi}(f)$ is supported by the functional $(x^*, -1)$.

Finally, let us prove (iii). If $f: X \rightarrow \mathbb{R}$ is not strictly convex, then its graph contains a segment of length $\ell > 0$, therefore one can find a point $x \in X$ and $x^* \in X^*$ for which $\text{diam}(S(y^*, \delta, \text{epi}(f))) \geq \ell$, where $y^* = (x^*, -1)$. Therefore f has not the SDP. $\square$

In general, none of the implications contained in the previous remark can be reversed (this is easy to see for (i) and (ii), and an example of strictly convex function without the SDP is contained in Example 3.7 below). On the other hand, the next proposition, whose proof is left to the reader, shows that strict convexity and uniform convexity coincide on compact sets.

Proposition 2.7. *Let K be a compact and convex subset of a normed space X . Let $f: K \rightarrow \mathbb{R}$ be a continuous and strictly convex function. Then f is uniformly convex on K .*

2.2. Convex multiobjective optimization problem

Multiobjective optimization is a topic of optimization theory that finds many applications in Economics, Operations Research and Engineering. Here we just give a brief account on the notions of solution and on the linear scalarization procedure. For a complete overview about this topic, we refer the readers to the, now classical, monographs [10] and [9]. Moreover, we quote also [7], even if this book is devoted only to the finite-dimensional setting. Even if the results we mention in this subsection can be proved in an easy way, their proofs can be found in the above mentioned books, often in a more general framework.

Let $f_1, \dots, f_N$ be N real valued functions defined on a normed linear space X .

Definition 2.8. A point $x_0 \in X$ is a *Pareto Optimum point* for $f_1, \dots, f_N$ when there is no $x \in X$ such that $f_k(x) \leq f_k(x_0)$ for each $k = 1, \dots, N$ and $f_i(x) < f_i(x_0)$ for some $i \in \{1, \dots, N\}$.

A weaker notion of solution for a multiobjective optimization problem is the notion of weak Pareto Optimum point.

Definition 2.9. A point $x_0 \in X$ is a *weak Pareto Optimum point* for $f_1, \dots, f_N$ when there is no $x \in X$ such that $f_k(x) < f_k(x_0)$ for each $k = 1, \dots, N$.

Of course, each Pareto Optimum point is also a weak Pareto Optimum point, but in general the reverse implication does not hold. Nevertheless, the two notions coincide under suitable convexity assumptions. Namely, the following result holds.

Proposition 2.10. *Let $f_1, \dots, f_N$ be strictly convex functions. Then each weak Pareto Optimum Points for $f_1, \dots, f_N$ is also Pareto Optimum for $f_1, \dots, f_N$.*

A multiobjective optimization problem consists in finding the (weak) Pareto Optimum points for objective functions $f_1, \dots, f_N$. It is easy to see that, in general, Pareto Optimum points are not unique. A common approach to solve a multiobjective optimization problem is to reduce it to a parametrized family of scalar optimization problems (this procedure is called scalarization of a multiobjective optimization problem).

Many different scalarization methods have been developed in the literature (see the monographs quoted above and the references therein). Among them, one of the most used is the linear scalarization or weighted sum scalarization. Let us denote by Σ_N the $(N - 1)$ -dimensional simplex, i.e., $\Sigma_N = \{t \in [0, 1]^N : \sum_{i=1}^N t_i = 1\}$ and, for $t = (t_1, \dots, t_N) \in \Sigma_N$, let $u_t(x) = \sum_{i=1}^N t_i f_i(x)$. We have the following sufficient optimality condition.

Proposition 2.11. *Let $x_0 \in X$. If there exists $\tilde{t} \in \Sigma_N$ such that $u_{\tilde{t}}(x_0) \leq u_{\tilde{t}}(x)$ for every $x \in X$, then x_0 is a weak Pareto Optimum point for $f_1, \dots, f_N$.*

The previous sufficient condition becomes also necessary if we add convexity assumptions for the objective functions.

Proposition 2.12. *Let $x_0 \in X$ and $f_1, \dots, f_N$ be convex functions. If x_0 is a weak Pareto Optimum point for $f_1, \dots, f_N$, then there exists $\tilde{t} \in \Sigma_N$ such that $u_{\tilde{t}}(x_0) \leq u_{\tilde{t}}(x)$ for every $x \in X$.*

In the sequel of the paper we always consider objective functions $f_1, \dots, f_N$ that are strictly convex. Therefore, by combining Propositions 2.10, 2.11 and 2.12, we obtain the following remark, that plays a key role in our setting.

Remark 2.13. Let us consider a multiobjective optimization problem with strictly convex objective functions $f_1, \dots, f_N$. Then a point $x_0 \in X$ is a Pareto Optimum point for $f_1, \dots, f_N$ if and only if x_0 is a minimum point for the function $u_t(x)$ for some $t \in \Sigma_N$.

3. Continuity of argmin map for a convex combinations of functions

This section is devoted to find some conditions ensuring that the minimum point of a convex combination of continuous strictly convex functions $f_1, \dots, f_N$ continuously depends on the coefficients of the convex combination. In order to provide existence of minima we restrict our attention to reflexive spaces.

Definition 3.1. Let X be a reflexive Banach space. A function $f: X \rightarrow \mathbb{R}$ is called a *U-function* if it satisfies the following conditions:

- (a) f is a continuous strictly convex function that is bounded on bounded sets.
- (b) f is coercive (i.e. for each $C \in \mathbb{R}$ the set $\{x \in X: f(x) \leq C\}$ is bounded).

Let us notice that *U-functions* have a unique minimum point. Indeed, let f be a *U-function* and take $C \in \mathbb{R}$ in such a way that the set $S_C = \{x \in X: f(x) \leq C\}$ is non-empty. Since f is continuous and convex, S_C is closed and convex, therefore it is a weakly closed subset of X , hence S_C is weakly compact. Since f is a convex function, it is weakly lower semicontinuous, therefore it admits a minimum point in S_C . Uniqueness follows by strict convexity of f .

Let $f_1, \dots, f_N$ be a finite family of *U-functions*. For

$$t = (t_1, \dots, t_N) \in \Sigma_N,$$

we recall that $u_t: X \rightarrow \mathbb{R}$ is the function defined by $u_t = \sum_{i=1}^N t_i f_i$. Since, for each $t \in \Sigma_N$, u_t is a *U-function*, we can consider the map $\phi: \Sigma_N \rightarrow X$ defined by

$$\phi(t) := \operatorname{argmin}(u_t).$$

Let us stress the fact that the map ϕ implicitly depends on the choice of the functions $f_1, \dots, f_N$.

Remark 3.2. For a family of *U-functions* $f_1, \dots, f_N$ as above, consider the set $\mathcal{M} = \phi(\Sigma_N)$ containing all minimum points of the functions u_t ($t \in \Sigma_N$). Proceeding as in [2, p. 28], it is not difficult to see that the set $\mathcal{M}$ is bounded.

Remark 3.3. (Multiobjective optimization with *U-functions*) When we consider a convex multiobjective optimization problem where the objective functions are *U-functions*, by Remark 2.13 it follows that the set $\mathcal{M} = \phi(\Sigma_N)$ is the set of all Pareto Optimum points for the functions $f_1, \dots, f_N$.

If X is finite-dimensional, for any choice of *U-functions* $f_1, \dots, f_N$, it is possible to show, by a standard compactness argument, that the function $\phi: \Sigma_N \rightarrow X$ is continuous (here, Σ_N is endowed with the standard product topology). Let us point out that, in the infinite-dimensional case, this is not true in general: Example 3.7 below contains two *U-functions* f_1, f_2 , for which the corresponding function ϕ is not continuous. The main aim of the present section is to provide suitable geometric assumptions, that ensure the continuity of the function ϕ in the infinite-dimensional setting (see Theorem 3.5, below).

We start by proving the continuity of the map ϕ at a given point of Σ_N . In the sequel, for a positive real number α and a *U-function* f , we denote by $S_\alpha(f)$ the strict sublevel set of f at level $(\min f + \alpha)$, i.e.,

$$S_\alpha(f) = \{x \in X: f(x) < \min(f) + \alpha\}.$$

Lemma 3.4. Let $f_1, \dots, f_N$ be a finite family of *U-functions* and $\hat{t} \in \Sigma_N$. Suppose that $\operatorname{diam}(S_\alpha(u_{\hat{t}}))$ goes to zero for α going to zero. Then the map $\phi: \Sigma_N \rightarrow X$ is continuous at $\hat{t}$.

Proof. We write $x_0 = \phi(\hat{t}) = \operatorname{argmin}(u_{\hat{t}})$. We claim that if a sequence $\{v^n\} \subset \Sigma_N$ converges to $\hat{t}$, then the sequence of functions $\{u_{v^n}\}$ converges uniformly to $u_{\hat{t}}$ on bounded sets. Indeed, let $D \subset X$ be a bounded set and assume $x \in D$. Define $\hat{t} = (\hat{t}_1, \dots, \hat{t}_N)$ and $v^n = (v_1^n, \dots, v_N^n)$, then

$$\begin{aligned} |u_{\hat{t}}(x) - u_{v^n}(x)| &= \left| \sum_{i=1}^N \hat{t}_i f_i(x) - \sum_{i=1}^N v_i^n f_i(x) \right| \\ &\leq \max_i \sup_{y \in D} |f_i(y)| \sum_{j=1}^N |\hat{t}_j - v_j^n|. \end{aligned}$$

Taking the supremum over $x \in D$ and letting n to infinity we get the claim.

Now suppose that the function $\varphi: \Sigma_N \rightarrow X$ is not continuous at $\hat{t}$. Then there exist a sequence $\{t_n\} \subset \Sigma_N$, converging to $\hat{t}$, and a positive real number $\xi > 0$ such that $\|x_n - x_0\| > \xi$, where $x_n := \phi(t_n)$ for every $n \in \mathbb{N}$. Since $\operatorname{diam}(S_\alpha(u_{\hat{t}}))$ goes to zero, there exists $\alpha > 0$ such that $\operatorname{diam}(S_\alpha(u_{\hat{t}})) < \xi/2$.

By Remark 3.2, the set $\mathcal{M}$ is bounded. Since $\{u_{t_n}\}$ converges uniformly to $u_{\hat{t}}$ on $\mathcal{M}$, there exists $n \in \mathbb{N}$ such that $|u_{t_n}(x_0) - u_{\hat{t}}(x_0)| < \alpha/3$ and $|u_{t_n}(x_n) - u_{\hat{t}}(x_n)| < \alpha/3$. Since $x_0 \in S_\alpha(u_{\hat{t}})$ and $\|x_n - x_0\| > \xi$, we have $u_{\hat{t}}(x_n) \geq \alpha + u_{\hat{t}}(x_0)$. Hence, we obtain

$$\begin{aligned} u_{t_n}(x_n) &= u_{t_n}(x_n) - u_{\hat{t}}(x_n) + u_{\hat{t}}(x_n) > 2\alpha/3 + u_{\hat{t}}(x_0) \\ &> \alpha/3 + u_{\hat{t}}(x_0) > u_{t_n}(x_0). \end{aligned}$$

A contradiction to the fact that $x_n = \phi(t_n) = \operatorname{argmin}(u_{t_n})$. $\square$

Theorem 3.5. *Let $f_1, \dots, f_N$ be a finite family of U -functions. The following assertions hold*

- (i) *If f_i has the SDP for some $i \in \{1, \dots, N\}$, then the map $\varphi: \Sigma_N \rightarrow X$ is continuous on $\{t = (t_1, \dots, t_N) \in \Sigma_N: t_i \neq 0\}$.*
- (ii) *If f_i has the SDP for every $i = 1, \dots, N$, then the map $\varphi: \Sigma_N \rightarrow X$ is continuous on Σ_N .*

Proof. Let us prove (i), (ii) follows immediately. Suppose that $t = (t_1, \dots, t_N) \in \Sigma_N$ is such that $t_i \neq 0$, and let us prove that ϕ is continuous at t . By Proposition 2.3, the function u_t is a U -function which satisfies the SDP. If $x_t = \phi(t)$, by Lemma 2.2, we have that $(x_t, u_t(x_t))$ is strongly exposed by $(0, -1) \in X^* \times \mathbb{R}$. Hence, $\operatorname{diam}(S_\alpha(u_t))$ goes to zero for α going to zero. The assertion follows by Lemma 3.4. $\square$

Remark 3.6. Let us point out that if a continuous convex function $f: X \rightarrow \mathbb{R}$ has the SDP then, by [12, Proposition 3.13], it is coercive. Therefore, under this additional assumption, condition (b) in Definition 3.1 is redundant.

The following example shows that the map $\phi: \Sigma_N \rightarrow X$ is not always continuous, even if $f_1, \dots, f_N$ are U -functions. Therefore, this example points out the key role played in Theorem 3.5 by the SDP.

Example 3.7. Let $X = \ell_2$, $h_1, h_2: X \rightarrow \mathbb{R}$, and $g: [0, \infty) \rightarrow \mathbb{R}$ be defined by

$$\begin{aligned} h_1(x) &= \sum_{n=2}^{\infty} \frac{(x_n - x_1)^2}{2^n}, & x &= (x_1, x_2, x_3, \dots) \in \ell_2, \\ h_2(x) &= \sum_{n=2}^{\infty} \frac{(x_n + x_1 - 1)^2}{2^n}, & x &= (x_1, x_2, x_3, \dots) \in \ell_2, \\ g(y) &= \begin{cases} (y - 2)^2 & \text{if } y \geq 2, \\ 0 & \text{if } 0 \leq y < 2. \end{cases} \end{aligned}$$

Then define

$$f_1(x) = h_1(x) + g(\|x\|), \quad f_2(x) = h_2(x) + g(\|x\|) \quad (x \in \ell_2).$$

Then f_1 and f_2 are real-valued U -functions on ℓ_2 such that:

- (1) f_1 attains its minimum at $x = 0$;
- (2) for each $t \in (0, 1)$, the function $u_t := (1 - t)f_1 + tf_2$ attains its minimum outside $2U_X$ (i.e. $\|\phi(t)\| \geq 2$).

In particular, the map $\phi: \Sigma_2 \rightarrow X$ is not continuous at $(t_1, t_2) = (1, 0) \in \Sigma_2$.

Proof. Let us first prove that f_1 and f_2 are U -functions. Because of the fact that, for each $x = (x_1, x_2, x_3, \dots) \in \ell_2$, we have

$$h_2((x_1, x_2, x_3, \dots)) = h_1((1 - x_1, x_2, x_3, \dots)),$$

it is sufficient to prove it just for f_1 .

Continuity of f_1 is trivial, while condition (b), in Definition 3.1, follows by the definition of the function g . We claim that h_1 is bounded on bounded sets. Indeed, if D is a bounded set and $x \in D$, we get

$$|h_1(x)| = \left| \sum_{n=2}^{\infty} \frac{(x_n - x_1)^2}{2^n} \right| \leq \sup_{n \in \mathbb{N}} |x_n - x_1|^2 \sum_{n=2}^{\infty} \frac{1}{2^n},$$

and the claim is proved. By our claim and by our construction, f_1 is bounded on bounded sets. It remains to show that f_1 is strictly convex. Since $g(\|\cdot\|)$ is a convex function, it is enough to show that h_1 is strictly convex. Let $x, y \in \ell_2$ be distinct, $\lambda \in (0, 1)$, and let us prove that

$$\theta_1 := h_1(\lambda x + (1 - \lambda)y) - [\lambda h_1(x) + (1 - \lambda)h_1(y)] < 0.$$

To do this, let us consider the function $k_1: \ell_2 \times \ell_2 \rightarrow \mathbb{R}$ defined by

$$k_1(x, y) = \sum_{n=2}^{\infty} \frac{(x_n - x_1)(y_n - y_1)}{2^n}, \quad x = (x_1, x_2, \dots), y = (y_1, y_2, \dots) \in \ell_2,$$

and observe that

- (i) $h_1(x) > 0$ if and only if $x \neq 0$;
- (ii) $h_1(x \pm y) = h_1(x) + h_1(y) \pm 2k_1(x, y)$.

Hence, by using (i) and (ii), we get

$$\begin{aligned}
 \theta_1 &= \sum_{n=2}^{\infty} \frac{(\lambda x_n - \lambda x_1 + (1-\lambda)y_n - (1-\lambda)y_1)^2}{2^n} - \lambda h_1(x) - (1-\lambda)h_1(y) \\
 &= \lambda^2 h_1(x) + (1-\lambda)^2 h_1(y) + 2\lambda(1-\lambda)k_1(x, y) - \lambda h_1(x) - (1-\lambda)h_1(y) \\
 &= \lambda(\lambda-1)h_1(x) - (1-\lambda)(\lambda)h_1(y) + 2\lambda(1-\lambda)k_1(x, y) \\
 &= -\lambda(1-\lambda)h_1(x-y) < 0.
 \end{aligned}$$

This concludes the proof of the fact that f_1, f_2 are U -functions.

Now, condition (1) trivially holds, let us prove (2). Fix $t \in (0, 1)$, we claim that function $\tilde{u}_t = (1-t)h_1 + th_2$ admits no minimum points. It is easy to see, by standard computation, that the functional $x_t^* \in \ell_2$ defined as

$$\left(\sum_{n=2}^{\infty} \frac{(x_1 - x_n)(1-t) + (x_n + x_1 - 1)t}{2^{n-1}}, \dots, \frac{(1-t)(x_n - x_1) + t(x_n + x_1 - 1)}{2^{n-1}}, \dots \right)$$

is the derivative of the function $\tilde{u}_t$ at the point $x = (x_1, x_2, \dots)$. By [11, Proposition 1.26], $\tilde{u}_t$ has a minimum at x if and only if $x_t^* = 0$. Assume, on the contrary, that $x_t^* = 0$, then for every $n \geq 2$ the following condition must be satisfied

$$\frac{(1-t)(x_n - x_1) + t(x_n + x_1 - 1)}{2^{n-1}} = 0.$$

Which yields $x_n = x_1 - 2tx_1 + t$. Substituting it in the first coordinate we get

$$\sum_{n=2}^{\infty} \frac{(1-t)(x_1 - x_1 + 2tx_1 - t) + t(x_1 - 2tx_1 + t + x_1 - 1)}{2^{n-1}},$$

which is null if and only if each term of the series is equal to zero. It follows that

$$\begin{aligned}
 (1-t)(2tx_1 - t) + (2x_1 - 2tx_1 + t - 1) &= 0 \\
 4tx_1 - 4t^2x_1 + 2t^2 - 2t &= 0 \\
 x_1 &= \frac{1}{2}.
 \end{aligned}$$

Which implies $x_n = \frac{1}{2}$ for every $n \in \mathbb{N}$. Since $x \in \ell_2$, we get a contradiction, and the claim is proved.

Now, suppose on the contrary that $\|\phi(t)\| < 2$. Then, since $u_t|_{2U_X} = \tilde{u}_t|_{2U_X}$, we get that $\tilde{u}_t$ attains its minimum at $\phi(t)$, a contradiction by our claim. The proof is concluded. $\square$

We conclude this section by pointing out that, by means of Theorem 3.5, we can obtain some information about the topological properties of the set of Pareto Optimum points for a multiobjective optimization problem with $f_1, \dots, f_N$ as objective functions. This type of results has been widely studied in the field of multiobjective optimization (see, e.g., [10, Ch.6] and [7, Sec.3.4]). Here we just quote a corollary of our result, that immediately follows by Remark 3.3 and Theorem 3.5

Corollary 3.8. *Let $f_1, \dots, f_N$ be a finite family of U-functions. If f_i has the SDP for every $i = 1, \dots, N$, then the set of Pareto Optimum points for $f_1, \dots, f_N$ is a compact and connected subset of X .*

4. Approximation of the solutions for a convex multiobjective optimization problem

This section is devoted to develop an approximation result for the solutions of a convex multiobjective optimization problem in the spirit of the paper [2]. We start by recalling the main theorem proved in [2].

Theorem 4.1. *Let X be a uniformly convex Banach space. Let $f_1, \dots, f_N$ be a finite family of U-functions which additionally satisfy the following conditions:*

- (a) *for each $i \in \{1, \dots, N\}$, f_i is Gâteaux differentiable at each point of X .*
- (b) *for each convex bounded set $D \subset X$, there exists a constant C_D and $q \geq 2$ such that, for each $x_1, x_2 \in D$*

$$f_i\left(\frac{x_1 + x_2}{2}\right) \leq \frac{1}{2}(f_i(x_1) + f_i(x_2)) - C_D \|x_1 - x_2\|^q,$$

for every $i \in \{1, \dots, N\}$.

Then, there exist $K, \gamma > 0$ such that if $t \in \Sigma_N$, $x_t = \phi(t)$, and $\{F_k\}$ is a sequence of independent random variables defined on the same probability space (Ω, μ) which take the values $f_1, \dots, f_N$ with probability $t_1, \dots, t_N$, we have, for each $n \in \mathbb{N}$

$$\int_{\Omega} \left\| \operatorname{argmin} \left(\frac{1}{n} \sum_{k=1}^n F_k(\omega) \right) - x_t \right\| d\mu \leq \frac{K}{n^\gamma}. \quad (2)$$

Remark 4.2. From the point of view of multiobjective optimization, this result can be read as follows. First of all, x_t is a Pareto Optimum point for the functions $f_1, \dots, f_N$ associated to the weights $(t_1, \dots, t_N) = t \in \Sigma_N$, i.e, x_t is the unique minimum point for the function u_t (see Remark 3.3). Second, each realization of the random variable $\frac{1}{n} \sum_{k=1}^n F_k(\omega)$ is a convex combination of n functions chosen among $f_1, \dots, f_N$ where each term has a coefficient $1/n$. Hence, for a fixed $\omega \in \Omega$, $\operatorname{argmin}(\frac{1}{n} \sum_{k=1}^n F_k(\omega))$ is a Pareto Optimum point for the functions $F_1(\omega), \dots, F_n(\omega)$ where $F_j(\omega) \in \{f_1, \dots, f_N\}$. Taking into account these interpretations of the quantities appearing in the inequality (2), we conclude that the expected value of the error made by approximating x_t with a special Pareto Optimum point (the one relating to all equal weights) of a convex multiobjective optimization problem with n objective functions chosen among $f_1, \dots, f_N$, where each of the n chosen functions has the probability t_j of being f_j , for each $j = 1, \dots, N$, is controlled by K/n^γ .

The most interesting feature of the estimation given by Theorem 4.1 is that each solution of the multiobjective problem with $f_1, \dots, f_N$ as objective functions can be approximated in a good way by considering the minimum point of the convex combination (with equal coefficients) of n functions chosen among the N objective functions of the original multiobjective problem. Moreover, most of the choices of the n functions among the N original objective functions give a good approximation

(as pointed out in [2], see also the introduction of the present paper). Indeed, on average we provide a good estimate of the error, that depends only on the number n of the chosen functions and on their geometrical properties. Finally, it is worth pointing out that this estimate does not depend on $t \in \Sigma_N$, in particular it does not depend on the choice of the Pareto Optimum point x_t .

The aim of the present section is to present a new version of the above result under the assumption (which markedly weakens conditions (b) contained in the above theorem) that the functions $f_1, \dots, f_N$ involved are Gâteaux differentiable U -functions satisfying the SDP. On the other hand, we show that our new version of Theorem 4.1 fails if we drop the SDP assumption, see Example 4.7.

Let us notice that condition (b), in Theorem 4.1, implies that each f_i is uniformly convex on bounded sets. On the other hand, in [2] it is claimed that condition (b) can be replaced by the assumption that each f_i is uniformly convex on bounded sets. Since the proof of this last assertion is not contained in [2], we decided, for the convenience of the reader, to include in our paper the corresponding standard modifications of some preliminary results, used in [2] to prove Theorem 4.1. The next technical lemma shows that for a finite family of uniformly convex functions there exists a common modulus of convexity with some additional properties. The proof is inspired by [13].

Lemma 4.3. *Let $f_1, \dots, f_N$ be a family of real-valued continuous functions defined on a convex subset D of a Banach space X . Suppose that each f_i is uniformly convex on D . Then there exists a function $\delta: (0, \infty) \rightarrow (0, \infty)$ such that*

(i) *for every $x, y \in D$ and $i \in \{1, \dots, N\}$ we have*

$$f_i\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f_i(x) + f_i(y)) - \delta(\|x - y\|).$$

(ii) *the function $t \mapsto \delta(t)/t$ is strictly increasing on $(0, \infty)$.*

(iii) $\lim_{t \rightarrow 0} \delta(t)/t = 0$.

Proof. Let $i \in \{1, \dots, N\}$. Since f_i is uniformly convex on D , combining [13, Remark 2.6] with [14, Proposition 3.5.1] there exists a map $\delta_i: (0, \infty) \rightarrow (0, \infty)$ such that:

(a) for every $x, y \in D$ we have

$$f_i\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f_i(x) + f_i(y)) - \delta_i(\|x - y\|);$$

(b) the function $t \mapsto \delta_i(t)/t^2$ is nondecreasing on $(0, \infty)$.

Let $s > t > 0$, then we have

$$\frac{\delta_i(t)}{t^2} \leq \frac{\delta_i(s)}{s^2},$$

hence, it follows that

$$0 < \frac{\delta_i(t)}{t} \leq t \frac{\delta_i(s)}{s^2},$$

which goes to zero when t goes to zero. Similarly, let $s > t > 0$, then we have

$$\frac{\delta_i(t)}{t} \leq t \frac{\delta_i(s)}{s^2} < \frac{\delta_i(s)}{s}.$$

Therefore, the function $\delta_i(t)/t$ is strictly increasing. Finally, defining

$$\delta(t) = \min\{\delta_1(t), \dots, \delta_N(t)\},$$

we get the desired function. □

Now, we are in the position of stating the main result of this section.

Theorem 4.4. *Let X be a uniformly convex Banach space. Let $f_1, \dots, f_N$ be a finite family of U -functions which additionally satisfy the following conditions:*

- (a) *for each $i \in \{1, \dots, N\}$, f_i is Gâteaux differentiable at each point of X .*
- (b) *for each $i \in \{1, \dots, N\}$, f_i has SDP.*

Then there exists a sequence $\{\epsilon_n\}$ of positive numbers converging to 0, such that if $t = (t_1, \dots, t_N) \in \Sigma_N$ and $\{F_k\}$ is a sequence of independent random variables defined on the same probability space (Ω, μ) which take the values $f_1, \dots, f_N$ with probability $t_1, \dots, t_N$, we have, for each $n \in \mathbb{N}$

$$\int_{\Omega} \left\| \operatorname{argmin} \left(\frac{1}{n} \sum_{k=1}^n F_k(\omega) \right) - \phi(t) \right\| d\mu \leq \epsilon_n.$$

As we have already pointed out in Remark 4.2, also Theorem 4.4 can be interpreted as an approximation result for a multiobjective optimization problem. Beside the weaker assumptions with respect to those in Theorem 4.1, it is worth pointing out that also the errors average upper bound ϵ_n can be explicitly computed. Indeed, from the proof of Theorem 4.4 we will see that

$$\epsilon_n = \operatorname{diam}(\mathcal{M}_1) n^{\beta} + \eta(n^{\beta}),$$

where $\mathcal{M}_1$ is the closed convex hull of the set of Pareto Optimum points for $f_1, \dots, f_N$, β is a negative real constant and η is a function vanishing as the argument goes to 0 and depending on the geometrical properties of the objective functions.

In order to prove our main result, we recall the statements of two technical lemmata contained in [2]. Let us mention that the constant p which appears in the following result is exactly the type of the Banach space X^* (for the definition of type and cotype see, e.g., [1, p. 137]).

Lemma 4.5. *Let X be a uniformly convex Banach space and let A be a bounded subset of X^* . Then, there exist $p \in (1, 2]$ and $C > 0$ such that,*

$$\int_{\Omega} \left\| \frac{1}{n} \sum_{k=1}^n G_k - x^* \right\| d\mu \leq C n^{\frac{1}{p}-1}$$

whenever $x^* \in X^*$ and $\{G_k\}$ is a sequence of independent random variables defined on the same probability space (Ω, μ) with values in A , such that

$$\int_{\Omega} G_k d\mu = x^*, \quad k \in \mathbb{N}.$$

The next result is contained in a slightly different formulation in [2]. For the sake of completeness we include a proof.

Lemma 4.6. *Let $f_1, \dots, f_N$ be a finite family of real-valued U -functions defined on X and let $D \subset X$ be a convex subset. Let $t \in \Sigma_N$, $x_t = \phi(t)$, and suppose that there exists a function $\delta: (0, \infty) \rightarrow (0, \infty)$ such that for every $x, y \in D$ and $i \in \{1, \dots, N\}$*

$$f_i\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f_i(x) + f_i(y)) - \delta(\|x - y\|)$$

holds. Then, for each $z \in D$, we have

$$2\delta(\|x_t - z\|) < u_t(z) - u_t(x_t).$$

Proof. It is sufficient to observe that, for every $z \in D$, we have

$$u_t(x_t) < u_t\left(\frac{x_t + z}{2}\right) < \frac{1}{2}(u_t(x_t) + u_t(z)) - \delta(\|x_t - z\|),$$

from which we get the assertion. $\square$

Proof of Theorem 4.4. Let us consider the set $\mathcal{M}$ of all minimum points of convex combinations of the functions $\{f_1, \dots, f_N\}$, i.e, $\mathcal{M} = \phi(\Sigma_N)$. Since Σ_N is a compact set and, by Theorem 3.5, the function ϕ is continuous, we have that $\mathcal{M}$ is a compact set. Therefore, $\mathcal{M}_1 = \overline{\text{conv}}(\mathcal{M})$, the closure of the convex hull of $\mathcal{M}$, is compact too (see, e.g., [6, Theorem V.2.6]). Hence, by Proposition 2.7, the functions $f_1, \dots, f_N$ restricted to the set $\mathcal{M}_1$ are uniformly convex. Therefore, by Lemma 4.3, there exists a function $\delta: (0, \infty) \rightarrow (0, \infty)$, satisfying conditions (i)–(iii) in Lemma 4.3. If $k \in \mathbb{N}$, $t \in \Sigma_N$, and $x_t = \phi(t)$, we define a random variable G_k^t over (Ω, μ) with values in X^* as follows: if $\omega \in \Omega$ and $F_k(\omega) = f_i$, put $G_k^t(\omega) = f_i'(x_t)$.

Then, $\{G_k^t\}_k$ is a sequence of X^* -valued independent random variables taking values $f_1'(x_t), \dots, f_N'(x_t)$ with probabilities $t_1, \dots, t_N$. Moreover, since the functions $f_1, \dots, f_N$ are bounded on bounded sets, it is easy to see that $\sup_{x \in \mathcal{M}_1} \max_i \|f_i'(x)\|$ is finite and hence there exists a bounded set $A \subset X^*$, such that G_k^t takes values in A , whenever $k \in \mathbb{N}$ and $t \in \Sigma_N$. Since, for each $k \in \mathbb{N}$ and $t \in \Sigma_N$, we have $\int_{\Omega} G_k^t d\mu = \sum_{i=1}^N t_i f_i'(x_t) = 0$, Lemma 4.5 ensures the existence of constants $p \in (1, 2]$ and $C > 0$ such that

$$\int_{\Omega} \left\| \frac{1}{n} \sum_{k=1}^n G_k^t \right\| d\mu \leq C n^{\frac{1}{p}-1},$$

whenever $t \in \Sigma_N$. Therefore $\mu(\Omega_1) \leq n^{\frac{1-p}{2p}}$, where

$$\Omega_1 = \left\{ \omega \in \Omega: \left\| \frac{1}{n} \sum_{k=1}^n G_k^t(\omega) \right\| \geq C n^{\frac{1-p}{2p}} \right\}.$$

Let $\omega \in \Omega \setminus \Omega_1$ and $n \in \mathbb{N}$; then for each $k \leq n$ we have $G_k^t(\omega) = f_{i_k}'(x_t)$ for some $i_k \in \{1, \dots, N\}$. Let $x = \operatorname{argmin}(\frac{1}{n} \sum_{k=1}^n f_{i_k})$ and

$$\psi(r) = \frac{1}{n} \sum_{k=1}^n f_{i_k}(x + r(x_t - x)).$$

The map ψ' is increasing over the unit interval $[0, 1]$ and attains its minimum at $r = 0$. By Lemma 4.6 we have

$$\begin{aligned} 2\delta(\|x - x_t\|) &\leq \psi(1) - \psi(0) = \int_0^1 \psi'(r) dr \leq |\psi'(1)| \\ &\leq \left\| \frac{1}{n} \sum_{k=1}^n f_{i_k}'(x_t) \right\| \|x - x_t\| \leq Cn^{\frac{1-p}{2p}} \|x - x_t\|. \end{aligned}$$

Defining $h(r) = \delta(r)/r$ and $\eta(r) = h^{-1}(r)$ we get

$$\frac{\delta(\|x - x_t\|)}{\|x - x_t\|} \leq \frac{Cn^{\frac{1-p}{2p}}}{2},$$

from which we get

$$\|x - x_t\| \leq \eta\left(\frac{Cn^{\frac{1-p}{2p}}}{2}\right).$$

Therefore,

$$\begin{aligned} \int_{\Omega} \left\| \operatorname{argmin} \left(\frac{1}{n} \sum_{k=1}^n F_k(\omega) \right) - x_t \right\| d\mu &\leq \int_{\Omega_1} \left\| \operatorname{argmin} \left(\frac{1}{n} \sum_{k=1}^n F_k(\omega) \right) - x_t \right\| d\mu \\ &+ \int_{\Omega \setminus \Omega_1} \left\| \operatorname{argmin} \left(\frac{1}{n} \sum_{k=1}^n F_k(\omega) \right) - x_t \right\| d\mu \leq \operatorname{diam}(\mathcal{M}_1) n^{\frac{1-p}{2p}} + \eta\left(\frac{Cn^{\frac{1-p}{2p}}}{2}\right). \end{aligned}$$

If, for $n \in \mathbb{N}$, we define

$$\varepsilon_n = \operatorname{diam}(\mathcal{M}_1) n^{\frac{1-p}{2p}} + \eta\left(\frac{Cn^{\frac{1-p}{2p}}}{2}\right),$$

the proof is complete. $\square$

Example 4.7. Let f_1 and f_2 be defined as in Example 3.7. It is clear that both f_1 and f_2 are Gâteaux differentiable on X . If $t = 0$, then $\operatorname{argmin}(u_t) = \{0\}$, where $u_t = (1-t)f_1 + tf_2$. While, if $t \in (0, 1]$, then $\|\operatorname{argmin}(u_t)\| \geq 1$. Suppose that $t \in (0, 1)$. Let (Ω, μ) be a probability space. Let us consider random variables $F_k: \Omega \rightarrow \{f_1, f_2\}$ that satisfy $\mu(\{\omega \in \Omega: F_k(\omega) = f_1\}) = 1 - t$ and $\mu(\{\omega \in \Omega: F_k(\omega) = f_2\}) = t$. Let

$$\begin{aligned} \Omega_1 &= \left\{ \omega \in \Omega: \operatorname{argmin} \left(\frac{1}{n} \sum_{k=1}^n F_k(\omega) \right) = \{0\} \right\} \\ &= \{\omega \in \Omega: F_k(\omega) = f_1 \forall k = 1, \dots, n\}. \end{aligned}$$

Then, we observe that $\mu(\Omega_1) = (1 - t)^n$. Which implies,

$$\begin{aligned} \int_{\Omega} \left\| \operatorname{argmin} \left(\frac{1}{n} \sum_{k=1}^n F_k(\omega) \right) - x_t \right\| d\mu &\geq \int_{\Omega_1} \left\| \operatorname{argmin} \left(\frac{1}{n} \sum_{k=1}^n F_k(\omega) \right) - x_t \right\| d\mu \\ &\geq \mu(\Omega_1) = (1 - t)^n. \end{aligned}$$

Therefore, one cannot obtain an estimate which does not depend on t .

Remark 4.8. Let us point out that there exist Gâteaux differentiable U -functions with the SDP not satisfying condition (b), in Theorem 4.1. To show this, notice that in [4] an example is provided, in each infinite-dimensional separable Banach space (also in every reflexive space), of an average locally uniformly rotund, Gâteaux differentiable norm $|\cdot|$ which is not locally uniformly rotund. For reflexive Banach spaces, by [4], a norm is average locally uniformly rotund, if and only if each $x \in S_X$ is strongly exposed by each functional which supports x . Hence, Proposition 2.4 implies that the function $|\cdot|^2$ is a Gâteaux differentiable U -function with the SDP not satisfying condition (b), in Theorem 4.1.

The above construction shows that there are situations in which the hypotheses of Theorem 4.4 are satisfied but Theorem 4.1 cannot be applied.

References

- [1] F. Albiac, N. Kalton: *Topics in Banach Space Theory*, Graduate Texts in Mathematics, Springer, New York (2006).
- [2] B. Beauzamy, P. Enflo: *Théorème de point fixe et d'approximation*, Ark. Mat. 23 (1985) 19–34.
- [3] J. B. Collier: *A class of strong differentiability spaces*, Proc. Amer. Math. Soc. 53 (1975) 420–422.
- [4] C. A. De Bernardi, J. Somaglia: *Rotund Gâteaux smooth norms which are not locally uniformly rotund*, Proc. Amer. Math. Soc. 52/4 (2024) 1689–1701.
- [5] C. A. De Bernardi, J. Somaglia, L. Veselý: *Star-finite coverings of Banach spaces*, J. Math. Anal. Appl. 491 (2020), art.no. 124384, 21 pp.
- [6] N. Dunford, J. T. Schwartz: *Linear Operators. Part I: General theory*, reprint of the 1958 original, John Wiley & Sons, New York (1988).
- [7] M. Ehrgott: *Multicriteria Optimization*, 2nd ed., Springer, Berlin (2005).
- [8] M. Fabian, P. Habala, P. Hájek, V. Montesinos, V. Zizler: *Banach Space Theory. The Basis for Linear and Nonlinear Analysis*, CMS Books in Mathematics / Ouvrages de Mathématiques de la SMC, Springer, New York (2011).
- [9] J. Jahn: *Vector Optimization. Theory, Applications, and Extensions*, Springer, Berlin (2004).
- [10] D. T. Luc: *Theory of Vector Optimization*, Lecture Notes in Economics and Mathematical Systems 319, Springer, Berlin (1989).
- [11] R. R. Phelps: *Convex Functions, Monotone Operators and Differentiability*, Lecture Notes in Mathematics 1364, Springer, Berlin (1993).

- [12] M. Volle, C. Zălinescu: *Strongly adequate functions on Banach spaces*, J. Convex Analysis 20 (2013) 655–668.
- [13] C. Zălinescu: *On uniformly convex functions*, J. Math. Anal. Appl. 95 (1983) 344–374.
- [14] C. Zălinescu: *Convex Analysis in General Vector Spaces*, World Scientific Publishing, Singapore (2002).