


Review

The Averaged Hausdorff Distances in Multi-Objective Optimization: A Review

Johan M. Bogoya ¹, Andrés Vargas ¹  and Oliver Schütze ^{2,3,*}

¹ Departamento de Matemáticas, Pontificia Universidad Javeriana, Cra. 7 N. 40-62, Bogotá D.C. 111321, Colombia; jbogoya@javeriana.edu.co (J.M.B.); a.vargasd@javeriana.edu.co (A.V.)

² Computer Science Department, CINVESTAV-IPN, Av. IPN 2508, Col. San Pedro Zacatenco, Mexico City 07360, Mexico

³ Dr. Rodolfo Quintero Ramirez Chair, UAM Cuajimalpa, Mexico City 05348, Mexico

* Correspondence: schuetze@cs.cinvestav.mx

Received: 27 August 2019; Accepted: 17 September 2019; Published: 24 September 2019



Abstract: A brief but comprehensive review of the averaged Hausdorff distances that have recently been introduced as quality indicators in multi-objective optimization problems (MOPs) is presented. First, we introduce all the necessary preliminaries, definitions, and known properties of these distances in order to provide a state-of-the-art overview of their behavior from a theoretical point of view. The presentation treats separately the definitions of the (p, q) -distances $GD_{p,q}$, $IGD_{p,q}$, and $\Delta_{p,q}$ for finite sets and their generalization for arbitrary measurable sets that covers as an important example the case of continuous sets. Among the presented results, we highlight the rigorous consideration of metric properties of these definitions, including a proof of the triangle inequality for distances between disjoint subsets when $p, q \geq 1$, and the study of the behavior of associated indicators with respect to the notion of compliance to Pareto optimality. Illustration of these results in particular situations are also provided. Finally, we discuss a collection of examples and numerical results obtained for the discrete and continuous incarnations of these distances that allow for an evaluation of their usefulness in concrete situations and for some interesting conclusions at the end, justifying their use and further study.

Keywords: Averaged Hausdorff distance; evolutionary multi-objective optimization; Pareto compliance; performance indicator; power means

1. Introduction

In many real-world applications, the problem of concurrent or simultaneous optimization of several objectives is an essential task known as a multi-objective optimization problem (MOP). One important problem in multi-objective optimization is to compute a suitable finite size approximation of the solution set of a given MOP, the so-called Pareto set and its image, the Pareto front.

The Hausdorff distance d_H (e.g., Reference [1]) measures how far two subsets of a metric space are from each other. Due to its properties, it is frequently used in many research areas such as computer vision [2–4], fractal geometry [5], the numerical computation of attractors in dynamical systems [6–8], or convergence of multi-objective algorithms to the Pareto set/front of a given multi-objective optimization problem [9–15]. One possible drawback of the classical Hausdorff distance, however, is that it punishes single outliers which leads to inequitable performance evaluations in some cases. As one example, we mention here multi-objective evolutionary algorithms. On the one hand, such algorithms are known to be very effective in the (global) approximation of the Pareto set/front. On the other hand, it is also known that the final approximations (populations) may contain some outliers (e.g., Reference [16]). For such cases, the Hausdorff distance may indicate a “bad” match

of population and Pareto set/front, while the approximation quality may be indeed “good”. To avoid exactly this problem, Schütze et al. introduced the *averaged* Hausdorff distance Δ_p in Reference [16], but the initial definition only works for finite approximations of the solution set and does not behave as a proper metric in the formal mathematical sense. In Reference [17], the indicator $\Delta_{p,q}$ has been proposed by the first two authors of this paper. $\Delta_{p,q}$ is an averaged Hausdorff distance that fixes the metric behavior of Δ_p . Later, in Reference [18], a broader definition was given on metric measure spaces, suitable for the consideration of continuous approximations of the solution set. Moreover, this generalized indicator $\Delta_{p,q}$ preserves the nice metric properties of the initial finite case and reduces to it when using the standard discrete measure.

While the averaged Hausdorff distance has so far mostly been used for performance assessment of multi-objective evolutionary algorithms (using benchmark functions), it has also been used on MOPs coming from real-world problems including the multi-objective software next release problem [19], arc routing problems [20], power flow problems [21], engineering design problems [22], foreground detection [23], and contract design [24]. Several other indicators have also been proposed in the literature, like the hypervolume indicator or R indicators, each one with its own advantages and drawbacks, but their consideration is beyond the scope of this work. Information concerning other indicators can be found, e.g., in References [25–27].

The material reviewed in this work is based on recently published works [17,18,28]. The remainder of the document is organized as follows: in Section 2, we will briefly state the required background for MOPs and power means. In Section 3, we review the p -averaged Hausdorff distance Δ_p . In Section 4, we will discuss its generalization, the (p, q) -averaged Hausdorff distance $\Delta_{p,q}$, explaining individually the finite and continuous cases. In Section 5, we will consider some aspects of the metric properties of Δ_p and $\Delta_{p,q}$. In Section 6, we study the Pareto compliance of the performance indicators related to Δ_p and $\Delta_{p,q}$. In Section 7, we will present some examples and numerical experiments. Finally, in Section 8, we will draw our conclusions and will discuss possible paths for future research in this direction.

2. Preliminaries

In this review, we introduce tools from a metric perspective that deal with two related contexts: distances between finite subsets of a metric space and distances between general measurable subsets of a metric measure space. The second context actually contains the first, but we deal separately with both of them, starting with the simpler setting of finite subsets before passing to the more general situation of arbitrary measurable sets that also contains the important special case of continuous sets. To emphasize each context, we use the convention that general sets will be denoted by X, Y , and Z , but when they are finite, the labels A, B , and C will be used.

2.1. Multi-Objective Optimization

First, we briefly present some basic aspects of multi-objective optimization problems (MOPs) required for the understanding of this paper. For a more thorough discussion, we refer the interested reader, e.g., to References [12,29,30].

A continuous MOP is rigorously formalized as the minimization of an appropriate function:

$$\min_{x \in Q} F(x),$$

where F denotes a vector-valued function with components $f_i: Q \rightarrow \mathbb{R}$, for $i = 1, \dots, k$, called objective functions. Explicitly,

$$\begin{aligned} F: Q &\longrightarrow \mathbb{R}^k, \\ x &\longmapsto F(x) := (f_1(x), \dots, f_k(x)). \end{aligned}$$

The optimality of a candidate solution to a MOP depends on a *dominance* relation [31] given in terms of the partial order introduced below.

Definition 1. For $x, y \in Q$, the partial Pareto ordering \preceq associated with the MOP determined by F is defined as

$$x \preceq y \text{ if and only if } f_i(x) \leq f_i(y), \text{ for all } i = 1, \dots, k.$$

For $x, y, z \in Q$ and $X, Y \subset Q$, the following notions of dominance (\prec) and non-dominance ($\not\prec$) are standard in this context:

x is dominated by y , written $y \prec x$ if $y \preceq x$ and $F(x) \neq F(y)$.

z is dominated by X , written $X \preceq z$ if $x \preceq z$ for some $x \in X$, otherwise $X \not\prec z$.

X is dominated by Y , written $Y \preceq X$ if $\forall x \in X \exists y \in Y$ such that $y \preceq x$, otherwise $Y \not\prec X$.

In addition, $x \in Q$ is called a Pareto-optimal point if it is nondominated, i.e., $\nexists y \in Q$ with $y \prec x$. Finally, the Pareto set $P \subset Q$ consists of all Pareto-optimal points and the Pareto front is defined as its image $F(P) \subset \mathbb{R}^k$.

MOPs commonly possess the important characteristic that, when mild smoothness conditions are fulfilled, the solution (or Pareto) set P and its image the Pareto front $F(P) \subset \mathbb{R}^k$ consist of d dimensional subsets for $d = k - 1$ (or even less) when the problem involves k objective functions ([32]).

As an example, let us describe a simple unconstrained MOP [33,34] given by

$$f_1, \dots, f_k : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$f_i(\mathbf{x}) = \sum_{j=1}^n (x_j - a_j^i)^2,$$

where $\mathbf{a}^i = (a_1^i, \dots, a_n^i) \in \mathbb{R}^n$ and $i = 1, \dots, k$. The \mathbf{a}^i 's correspond to the minimizers of each quadratic objective f_i , and the Pareto set of this problem consists of a $(k - 1)$ simplex containing all the \mathbf{a}^i 's as vertices, i.e.,

$$\text{simp}_{k-1} := \text{simp}(\mathbf{a}^1, \dots, \mathbf{a}^k) = \left\{ \sum_{i=1}^k \mu_i \mathbf{a}^i : \mu_1, \dots, \mu_k \geq 0 \text{ and } \sum_{i=1}^k \mu_i = 1 \right\}.$$

In the particular case when $n = 1, k = 2, a^1 = 0$, and $a^2 = 2$, the problem becomes

$$F : \mathbb{R} \longrightarrow \mathbb{R}^2$$

$$F(x) = (x^2, (x - 2)^2). \tag{1}$$

This is the so-called Schaffers problem [35]. Figure 1 illustrates the objectives f_1 and f_2 , and the Pareto front $F(P)$ for this MOP. In this case, the Pareto set corresponds to $P = [0, 2]$ and the Pareto front is a continuous convex curve in \mathbb{R}^2 joining $(0, 4)$ with $(4, 0)$.

In many real-world applications, MOPs arise naturally. As one example, in almost all scheduling problems (e.g., References [36–41]), the total execution time (make-span) is of primary interest. However, the consideration of this objective is in many cases not enough since other quantities such as the tardiness or the energy consumption also play an important role and can consequently, according to the given problem, also add objectives to the resulting multi-objective problem.

For the numerical treatment of MOPs, there exist already many established approaches. For instance, there are mathematical programming techniques [29,42], point-wise iterative methods that are capable of detecting single local solutions of a given MOP. Via use of a clever sequence of these resulting scalar objective optimization problems, a suitable finite size approximation of the entire

Pareto front can be computed in certain cases [43–46]. Multi-objective continuation methods take advantage of the fact that the Pareto set at least locally forms a manifold [47–52]. Starting with an initial (local) solution, further candidates are computed along the Pareto set of the given MOP. All of these methods typically yield high convergence rates but are, in turn, of local nature. A possible alternative is given by set oriented methods such as subdivision and cell mapping techniques [53,54] and evolutionary algorithms [55–59] that are of global nature and are capable of computing a finite size approximation of the Pareto front in one single run.

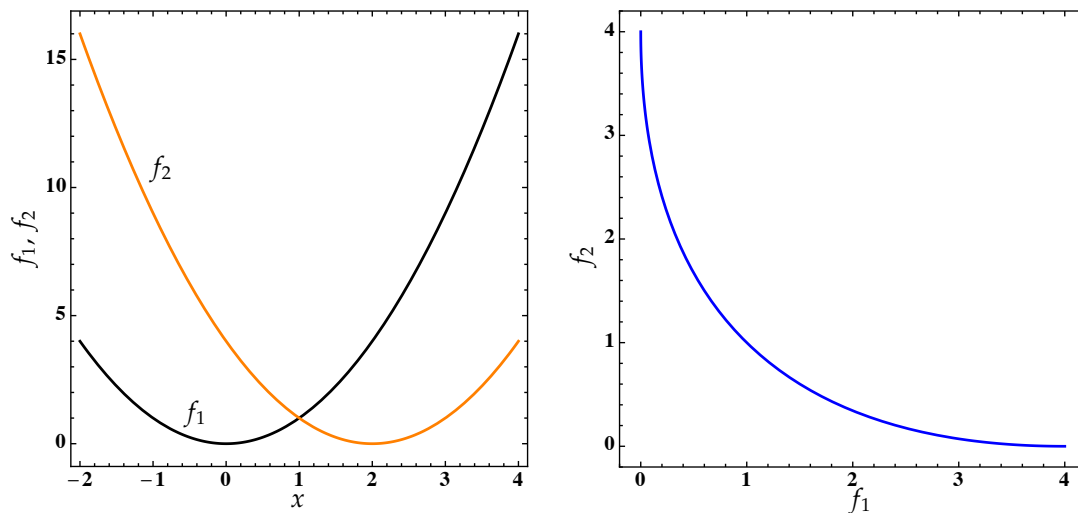


Figure 1. Left: the objectives $f_1(x) = x^2$ and $f_2(x) = (x - 2)^2$ from a multi-objective optimization problem (MOP; Equation (1)). Right: the corresponding Pareto set over the interval $[0, 2]$.

2.2. Finite Power Means

A comprehensive reference on the theory and properties of means is given in Reference [60], where proofs of the statements presented here for finite power means and for integral power means in the following subsection can be found (see also Reference [18] for integral means).

For a finite set $A \subset [0, \infty)$ and a nonzero real p , the p -average or the p power mean of A is given by

$$\mathcal{M}^p(a) := \left(\frac{1}{|A|} \sum_{a \in A} a^p \right)^{\frac{1}{p}},$$

where $|A|$ denotes the cardinality of A . The simpler notation $\mathcal{M}^p(A) := \mathcal{M}^p(a)$ will also be employed. Moreover, in order to simplify the forthcoming expressions, we introduce the abbreviation

$$\sum_{a \in A} a := \frac{1}{|A|} \sum_{a \in A} a$$

to denote the arithmetic mean of the elements of a finite set $A \subset \mathbb{R}_{\geq 0}$.

It is well known that limit cases of power means recover familiar quantities, for example,

$$\lim_{q \rightarrow 0} \mathcal{M}^q(A) = \left(\prod_{a \in A} a \right)^{\frac{1}{|A|}},$$

is the standard geometric mean of the elements of A . The special case $p = -1$ corresponds to the harmonic mean,

$$\text{harm}(A) := \mathcal{M}^{-1}(A).$$

Moreover, the p -average of any finite set can also be defined for any p in the extended real line $\overline{\mathbb{R}} := [-\infty, \infty]$ by taking appropriate limits.

$$\lim_{q \rightarrow \infty} \mathcal{M}^q(A) = \max(A), \quad \text{and} \quad \lim_{q \rightarrow -\infty} \mathcal{M}^q(A) = \min(A).$$

Proposition 1. *Let A and B be finite subsets of $[0, \infty)$ and $p, q \in \overline{\mathbb{R}}$ be arbitrary constants. Then, the following properties hold for finite power means:*

1. $\mathcal{M}^p(A) \leq \mathcal{M}^p(B)$.
2. For $p \leq q$: $\mathcal{M}^p(A) \leq \mathcal{M}^q(A)$.
3. For a matrix of nonnegative elements $D = (d_{a,b})$ with $a \in A$ and $b \in B$:

$$\mathcal{M}^p(D) := \mathcal{M}^p \left(\mathcal{M}^p(d_{a,b}) \right) = \mathcal{M}^p \left(\mathcal{M}^p(d_{a,b}) \right).$$

4. For $p \geq 1$: $\mathcal{M}^p(\{a + b \mid a \in A, b \in B\}) \leq \mathcal{M}^p(A) + \mathcal{M}^p(B)$.
5. For the harmonic mean: $\text{harm}(A) \leq |A| \min(A)$.

2.3. Integral Power Means in Measure Spaces

In order to present this part with sufficient generality, let us denote by (\mathcal{S}, μ) a measure space. Let $\mathfrak{M}(\mathcal{S})$ be the σ algebra of measurable subsets of \mathcal{S} and $\mathfrak{M}_{<\infty}(\mathcal{S})$ be the collection of those subsets with finite measure.

Now, we recall some fundamental properties of integral power means in this setting needed for the forthcoming sections. For $p \in \mathbb{R} \setminus \{0\}$ and a measurable function $f: X \subset \mathcal{S} \rightarrow [0, \infty)$ defined on a subset $X \in \mathfrak{M}_{<\infty}(\mathcal{S})$, the p power mean or p -average of f over X is given by

$$\mathcal{M}^p_{x \in X}(f(x)) := \left(\frac{1}{\mu(X)} \int_X f(x)^p \, d\mu \right)^{\frac{1}{p}}. \tag{2}$$

For convenience, RHS of Equation (2) will be denoted simply as

$$\int_X f^p \, d\mu := \frac{1}{|X|} \int_X f(x)^p \, d\mu,$$

where $|X| := \mu(X)$ refers in this context to the measure of X and not to its cardinality as in the finite case. For brevity, when the measure μ employed is clear, $d\mu$ will be abbreviated by dx to highlight the variable being integrated. The shorthand $\mathcal{M}^p(f(X)) := \mathcal{M}^p_{x \in X}(f(x))$ will also be employed.

For $p \geq 1$, the integral p mean corresponds to $\mathcal{M}^p(f(X)) = |X|^{-\frac{1}{p}} \|f\|_p$, where $\|\cdot\|_p$ is the standard p norm of the Lebesgue space $L^p(X, \mu)$. The cases $p = \pm\infty$ can also be included by taking the limits $p \rightarrow \pm\infty$. In fact, since the essential supremum of the function f on X is $\|f\|_\infty = \text{ess sup}_{x \in X} f(x)$, and when f is not identically zero its essential infimum is precisely $\|1/f\|_\infty^{-1} = \text{ess inf}_{x \in X} f(x)$; by calculating the limits, we obtain that

$$\mathcal{M}^\infty_{x \in X}(f(x)) := \lim_{p \rightarrow \infty} \left(\int_X f^p \, d\mu \right)^{\frac{1}{p}} = \|f\|_\infty,$$

and similarly,

$$\mathcal{M}^{-\infty}_{x \in X}(f(x)) := \lim_{p \rightarrow -\infty} \left(\int_X f^p \, d\mu \right)^{\frac{1}{p}} = \left\| \frac{1}{f} \right\|_\infty^{-1}.$$

Note that $\|\cdot\|_\infty$ corresponds to the norm of the space $L^\infty(X, \mu)$. For $p = 0$, it is possible to define \mathcal{M}^p as the integral generalization of the notion of geometric mean, and it is given explicitly by

$$\mathcal{M}_{x \in X}^0(f(x)) := \exp\left(\int_X \log f \, d\mu\right).$$

Proposition 2. For subsets $X, Y \in \mathfrak{M}_{<\infty}(\mathbb{R}^k)$, nonnegative measurable functions $f, g: X \rightarrow [0, \infty)$, and any product-measurable function $d: X \times Y \rightarrow [0, \infty)$, the integral power mean \mathcal{M}^p satisfies that

1. For $p \in \overline{\mathbb{R}}$, $k \in [0, \infty)$: $\mathcal{M}_{x \in X}^p(k) = k$ and $\mathcal{M}_{x \in X}^p(kf(x)) = k \mathcal{M}_{x \in X}^p(f(x))$.
2. For $p \in \overline{\mathbb{R}}$: $\mathcal{M}_{x \in X}^p\left(\mathcal{M}_{y \in Y}^p(d(x, y))\right) = \mathcal{M}_{y \in Y}^p\left(\mathcal{M}_{x \in X}^p(d(x, y))\right)$.
3. For $p \in [1, \infty]$: $\mathcal{M}_{x \in X}^p(f(x) + g(x)) \leq \mathcal{M}_{x \in X}^p(f(x)) + \mathcal{M}_{x \in X}^p(g(x))$.
4. For $p \in \overline{\mathbb{R}}$ and $f \leq g$: $\mathcal{M}_{x \in X}^p(f(x)) \leq \mathcal{M}_{x \in X}^p(g(x))$.
5. For $p, q \in [0, \infty]$ with $p \leq q$: $\mathcal{M}_{x \in X}^p(f(x)) \leq \mathcal{M}_{x \in X}^q(f(x))$.

3. The p -Averaged Hausdorff Distance

When trying to measure the distance between subsets of Euclidean space or even an arbitrary metric space, a natural choice is the well-known Hausdorff distance d_H that is extensively employed in many different contexts. However, its use is of limited practical value to measure the distance to the Pareto set/front in typical MOPs, such as stochastic search methods implemented by an evolutionary algorithm. This is due to the fact that these algorithms may produce a set of outliers that can be heavily punished by d_H . As a partial remedy, the use of an *averaged* Hausdorff distance Δ_p was first proposed in Reference [16] to replace d_H .

Let $d: \mathcal{S} \times \mathcal{S} \rightarrow [0, \infty)$ denote a distance function on a metric space \mathcal{S} for which the standard properties of the identity of indiscernibles, nonnegativity, symmetry, and subadditivity (more commonly known as the triangle inequality) are satisfied.

Definition 2. Given a point $x_0 \in \mathcal{S}$ and subsets $X, Y \subset \mathcal{S}$, we have

1. A pointwise distance to sets: $d(x_0, X) := \inf \{d(x_0, x) \mid x \in X\}$.
2. A pre-distance between sets: $d(Y, X) := \sup \{d(y, X) \mid y \in Y\}$.
3. The Hausdorff distance between sets: $d_H(X, Y) := \max \{d(X, Y), d(Y, X)\}$.

For simplicity, throughout the text, the metric d can be assumed to be the standard Euclidean distance $d(x, y) := \|x - y\|$ induced on some $\mathcal{S} \subset \mathbb{R}^k$ by the Euclidean 2 norm of \mathbb{R}^k , but the theory carries over to any general metric space (\mathcal{S}, d) .

Definition 3. Let $p \in \mathbb{N}$. For finite subsets $A, B \subset \mathcal{S}$, their (modified) p generational distance is

$$\text{GD}_p(A, B) := \left(\frac{1}{|A|} \sum_{a \in A} d(a, B)^p\right)^{\frac{1}{p}},$$

and their (modified) p inverted generational distance is

$$\text{IGD}_p(A, B) := \left(\frac{1}{|B|} \sum_{b \in B} d(b, A)^p\right)^{\frac{1}{p}}.$$

From them, the p averaged Hausdorff distance is obtained by taking the maximum

$$\Delta_p(A, B) := \max \{GD_p(A, B), IGD_p(A, B)\}.$$

The indicators GD_p and IGD_p in Definition 3 correspond to simple adjustments to the definitions of the generational distance [61] and the inverted generational distance [62].

The standard Hausdorff distance is recoverable from Δ_p by taking the limit $\lim_{p \rightarrow \infty} \Delta_p = d_H$, but for any finite value of p , the distance Δ_p is obtained from standard p power means of all the distances employed to calculate the supremum in part 2 of Definition 2, which is needed to define d_H .

The advantage of using Δ_p as an indicator is that it does not immediately disqualify a few outliers in a candidate set, contrary to what d_H does and that, among the possible configurations of (finite) candidate solutions to a MOP, it assigns lesser distances to the Pareto front to those solutions appearing evenly spread along its whole domain (see, e.g., Reference [63]). The behavior of Δ_p as a quality indicator is studied, e.g., in References [16,28], and it corresponds to the particular case $q \rightarrow -\infty$ of the results for general (p, q) -indicators presented in Section 6.

Concerning its metric properties, Δ_p has the drawback of not being a proper metric in the usual sense because for any non-unit set $A \subset S$ the distance $\Delta_p(A, A) > 0$. This problem will be fixed in the following section with a simple modification. Nevertheless, independently from that, for a positive number p , the distance Δ_p does not satisfy the triangle inequality but only a weaker version of it. Indeed, as a consequence of Corollary 3, we have that

$$\Delta_p(A, C) \leq N^\alpha (\Delta_p(A, B) + \Delta_p(B, C)),$$

where $N = \max\{|A|, |B|, |C|\} \geq 1$ and $\alpha = 1/p$.

For further details concerning Δ_p , its properties, and its relation to other indicators, the reader can consult, e.g., References [16,63].

4. The (p, q) -Averaged Hausdorff Distance

To better evaluate the optimality of a certain candidate set to approximate the Pareto set/front of a MOP, several generalizations of the averaged Hausdorff distance Δ_p have been recently introduced.

4.1. (p, q) -Distances between Finite Sets

Definition 4. For $p, q \in \mathbb{R} \setminus \{0\}$, the generational (p, q) -distance $GD_{p,q}(A, B)$ between two finite subsets $A, B \subset S$ is given by

$$GD_{p,q}(A, B) := \left(\sum_{a \in A} \left(\sum_{b \in B} d(a, b)^q \right)^{\frac{p}{q}} \right)^{\frac{1}{p}}.$$

The distance $GD_{p,q}(A, B)$ can be extended for values of $p = 0$ or $q = 0$, by taking the limits $p \rightarrow 0$ or $q \rightarrow 0$, respectively. In such cases, properties of finite power means suggest the following definitions:

$$GD_{p,0}(A, B) := \left(\sum_{a \in A} \left(\prod_{b \in B} d(a, b) \right)^{\frac{p}{|B|}} \right)^{\frac{1}{p}}, \quad \text{when } p \neq 0,$$

$$GD_{0,q}(A, B) := \left(\prod_{a \in A} \left(\sum_{b \in B} d(a, b)^q \right)^{\frac{1}{q}} \right)^{\frac{1}{|A|}}, \quad \text{when } q \neq 0, \text{ and}$$

$$GD_{0,0}(A, B) := \left(\prod_{a \in A} \left(\prod_{b \in B} d(a, b) \right)^{\frac{1}{|B|}} \right)^{\frac{1}{|A|}} \quad \text{if } p = q = 0.$$

We can also calculate $GD_{p,q}$ when $p \rightarrow \pm\infty$ or $q \rightarrow \pm\infty$ by changing the corresponding sum with a minimum or a maximum according to the case. In particular, we have the nice relation

$$\lim_{q \rightarrow -\infty} GD_{p,q}(A, B) = GD_p(A, B). \tag{3}$$

Note that the definition of $GD_{p,q}$ has two drawbacks, namely $GD_{p,q}(A, B)$ does not necessarily vanish if $A = B$ and in general $GD_{p,q}(A, B) \neq GD_{p,q}(B, A)$, hence it does not define a proper metric. In order to get one, a slight modification is needed.

Definition 5. Let $p, q \in \mathbb{R} \setminus \{0\}$. For finite subsets $A, B \subset \mathcal{S}$, their (p, q) -averaged Hausdorff distance is

$$\Delta_{p,q}(A, B) := \max\{GD_{p,q}(A, B \setminus A), GD_{p,q}(B, A \setminus B)\}.$$

Notice that $GD_p(A, B) = GD_p(A, B \setminus A)$ when $A \cap B = \emptyset$, thus using Equation (3) and Definition 5, we easily obtain

$$\lim_{q \rightarrow -\infty} \Delta_{p,q}(A, B) = \Delta_p(A, B).$$

In this way, for finite and disjoint sets, the indicator $\Delta_{p,q}$ is a generalization of Δ_p . Similarly to the relation

$$GD_p(A, B) = |A|^{-\frac{1}{p}} \|D_{AB}\|_p,$$

between the $GD_p(A, B)$ and the matrix ℓ_p norm $\|D_{AB}\|_p$ of the distance matrix $D_{AB} := [d(a, b)]_{a,b}$ for $a \in A$ and $b \in B$, we also have the following relation between the (p, q) -generational distance $GD_{p,q}(A, B)$ and the matrix $\ell_{p,q}$ norm $\|D_{AB}\|_{p,q}$, where the definition of the latter is precisely that of $GD_{p,q}$ but replacing all the normalized sums \sum by standard ones Σ (see, e.g., Reference [64]):

$$GD_{p,q}(A, B) = \mathcal{M}_{a \in A}^p \left(\mathcal{M}_{b \in B}^q (d(a, b)) \right) = |A|^{-\frac{1}{p}} |B|^{-\frac{1}{q}} \|D_{AB}\|_{p,q}.$$

A useful property of the distance $\Delta_{p,q}$ is that the parameters can be adjusted independently to achieve some desired spread of the archives by choosing an appropriate q and that they can be located with custom closeness to the Pareto front of a MOP by an adequate choice of p .

4.2. (p, q) -Distances between Measurable Sets

With the aid of Proposition 2, the results of the previous section can be generalized to subsets of a metric space (\mathcal{S}, d) endowed with an appropriate measure μ . For concreteness, \mathcal{S} can be taken to be a subset of \mathbb{R}^k carrying the metric induced from the Euclidean metric of \mathbb{R}^k and endowed with an appropriate non-null measure μ . Notice that, in our intended applications, μ will not be the restriction of the standard Lebesgue measure of \mathbb{R}^k to \mathcal{S} for the simple reason that it can easily vanish as it happens on any hypersurface or lower dimensional subsets of \mathbb{R}^k . In this case, a lower dimensional measure is needed and alternatives like the Hausdorff measure on \mathcal{S} can be used, since it gives rise to the standard notion of d dimensional volume for d submanifolds of \mathbb{R}^k . When these submanifolds are parametrized by functions from subsets of \mathbb{R}^d , the same volume will be obtained by a change of variable formulae from the standard Lebesgue measure on those subsets of \mathbb{R}^d .

A very important observation in this context is that any set-theoretic relation obtained from measure-related calculations needs to be understood to hold almost everywhere (*a.e.*). Therefore, for $X, Y \in \mathfrak{M}_{<\infty}(\mathcal{S})$, the statements $X = Y$ or $X \subset Y$ mean that the relations hold *a.e.*, i.e., $\mu\{X \neq Y\} = 0$ or $\mu\{X \not\subseteq Y\} = 0$, respectively. In other words, in this setting, we will always identify $X \in \mathfrak{M}_{<\infty}(\mathcal{S})$ with its equivalence class $[X] := \{Y \mid X = Y, \text{ a.e.}\}$. This means that those classes will be regarded as the elements of $\mathfrak{M}_{<\infty}(\mathcal{S})$, removing the need to carry the abbreviation *a.e.* all the time. Henceforth, to simplify complicated formulae, $d(x, y)$ will be shortened to $d_{x,y}$.

Definition 6. Let $p, q \in \mathbb{R} \setminus \{0\}$. For finite-measure subsets $X, Y \in \mathfrak{M}_{<\infty}(\mathcal{S})$, their generational (p, q) -distance is given by

$$GD_{p,q}(X, Y) := \mathfrak{M}^p_{x \in X} \left(\mathfrak{M}^q_{y \in Y} (d_{x,y}) \right) = \left(\int_X \left(\int_Y d_{x,y}^q dy \right)^{\frac{p}{q}} dx \right)^{\frac{1}{p}}.$$

The cases $p < 0$ or $q < 0$ are well defined only if X and Y are disjoint subsets.

Similarly to the finite case, $GD_{p,q}$ can be extended to values of $p, q \in \overline{\mathbb{R}}$, but there are two drawbacks: $GD_{p,q}(X, X) = 0$ only if X is a unit-set or singleton, and $GD_{p,q}(X, Y)$ can differ from $GD_{p,q}(Y, X)$. To fix this undesirable behavior, we repeat the strategy used in the finite case as follows.

Definition 7. Let $p, q \in \mathbb{R} \setminus \{0\}$. For finite-measure subsets $X, Y \in \mathfrak{M}_{<\infty}(\mathcal{S})$, their (p, q) -averaged Hausdorff distance is given by

$$\Delta_{p,q}(X, Y) := \max \{GD_{p,q}(X, Y \setminus X), GD_{p,q}(Y, X \setminus Y)\}.$$

Remark 1. In general, the (p, q) -distances are maps: $\mathfrak{M}_{<\infty}(\mathcal{S}) \times \mathfrak{M}_{<\infty}(\mathcal{S}) \rightarrow [0, \infty)$. On the collection of finite subsets of \mathcal{S} , the standard counting measure can be taken as the underlying one needed for these measure-theoretic notions of $GD_{p,q}$ and $\Delta_{p,q}$, and in this case, these distances become precisely the finite-case distances given in Definitions 4 and 5.

Remark 2. For disjoint subsets X and Y , Definition 5 in the finite case and Definition 7 above in the measurable case reduce to the simpler form

$$\Delta_{p,q}(X, Y) := \max \{GD_{p,q}(X, Y), GD_{p,q}(Y, X)\},$$

which is the one we will actually use in most situations. The more general definition for non-disjoint subsets is given with the purpose that the distance so-defined changes continuously as one set approaches the other until their distance vanishes. In other words, the general definition allows the distance to become a continuous function with respect to the metric topology that it determines. Nevertheless, for practical purposes dealing with applications and for most of the results presented below, the simpler definition between disjoint subsets suffices.

5. Metric Properties

To explain some of the terminology used in this section, we recall to the reader that the standard triangle inequality for a distance function $d: \mathcal{S} \times \mathcal{S} \rightarrow [0, \infty)$ is usually weakened in two different but related ways by postulating the existence of a constant $C > 0$ such that, for any points $x, y, z \in \mathcal{S}$, one of the following conditions hold:

- The C relaxed triangle inequality: $d(x, z) \leq C(d(x, y) + d(y, z))$.
- The C inframetric inequality: $d(x, z) \leq C \max\{d(x, y), d(y, z)\}$.

Since the second condition implies the first one by using the very same constant $C > 0$ and, reciprocally, the C relaxed triangle inequality implies the $2C$ inframetric one, both conditions are equivalent for an appropriate choice of constants. A semimetric satisfying any one of these conditions will be simply called an inframetric.

For arbitrary measurable sets in \mathcal{S} , the following results summarize the metric properties of $GD_{p,q}$ and $\Delta_{p,q}$. Using the counting measure, these properties also apply to finite sets. For more details, see Reference [17] in the finite case and Reference [18] in the generalized measure-theoretic context.

Theorem 1. For $p, q \in [1, \infty]$, the generational (p, q) -distance $GD_{p,q}$ is subadditive in $\mathfrak{M}_{<\infty}(\mathcal{S})$, i.e., for any $X, Y, Z \in \mathfrak{M}_{<\infty}(\mathcal{S})$, the triangle inequality holds true:

$$GD_{p,q}(X, Z) \leq GD_{p,q}(X, Y) + GD_{p,q}(Y, Z).$$

Proof. The proof follows easily by simple steps using the properties in Proposition 2. We start from the standard triangle inequality for $d(\cdot, \cdot)$:

$$d_{x,z} \leq d_{x,y} + d_{y,z} \quad (x \in X, y \in Y, z \in Z),$$

taking at both sides the q -average over Z and using 1–3 of Proposition 2 to arrive at

$$\mathcal{M}^q(d_{x,z}) \leq \mathcal{M}^q(d_{x,y} + d_{y,z}) \leq \mathcal{M}^q(d_{x,y}) + \mathcal{M}^q(d_{y,z}) = d_{x,y} + \mathcal{M}^q(d_{y,z}). \tag{4}$$

Now, there are two independent cases for the parameters $p, q \in [1, \infty)$. We explain here only the case $p \leq q$, but the case $q < p$ follows by similar arguments; see Thm. 2 in Reference [18]. Calculating the p -average over X at both sides of Equation (4) and using 1, 3, and 5 of Proposition 2, we get

$$\mathcal{M}^p\left(\mathcal{M}^q(d_{x,z})\right) \leq \mathcal{M}^p\left(d_{x,y} + \mathcal{M}^q(d_{y,z})\right) = \mathcal{M}^p(d_{x,y}) + \mathcal{M}^q(d_{y,z}). \tag{5}$$

Since the LHS of Equation (5) is $\text{GD}_{p,q}(X, Z)$, after a further p -average over Y at both sides of Equation (5) and parts 1, 3, and 5 of Proposition 2, we obtain

$$\text{GD}_{p,q}(X, Z) \leq \mathcal{M}^p\left(\mathcal{M}^p(d_{x,y}) + \mathcal{M}^q(d_{y,z})\right) = \mathcal{M}^p\left(\mathcal{M}^p(d_{x,y})\right) + \text{GD}_{p,q}(Y, Z).$$

But from 2, 4, and 5 of Proposition 2, the first term at the RHS above satisfies

$$\mathcal{M}^p\left(\mathcal{M}^p(d_{x,y})\right) = \mathcal{M}^p\left(\mathcal{M}^p(d_{x,y})\right) \leq \mathcal{M}^p\left(\mathcal{M}^q(d_{x,y})\right) = \text{GD}_{p,q}(X, Y). \quad \square$$

Corollary 1. If $p, q \in \mathbb{R} \setminus \{0\}$, the (p, q) -averaged Hausdorff distance $\Delta_{p,q}$ is a semimetric on the space $\mathfrak{M}_{<\infty}(\mathcal{S})$ of finite-measure subsets of \mathcal{S} . Furthermore, if $p, q \in [1, \infty)$ the distance $\Delta_{p,q}$ behaves as a proper metric when it is restricted to disjoint subsets of $\mathfrak{M}_{<\infty}(\mathcal{S})$.

Proof. From Definition 7, we obtain the relations $\Delta_{p,q}(\cdot, \cdot) \geq 0$ and $\Delta_{p,q}(X, Y) = \Delta_{p,q}(Y, X)$ for any $X, Y \in \mathfrak{M}_{<\infty}(\mathcal{S})$ and all $p, q \in \mathbb{R} \setminus \{0\}$. Moreover, from Definition 6, it follows that $\text{GD}_{p,q}(X, Y \setminus X) = 0$ if and only if $X = \emptyset$ or $Y \subseteq X$ (hence, $Y \setminus X = \emptyset$). Therefore, for $X, Y \neq \emptyset$,

$$\Delta_{p,q}(X, Y) = 0 \iff X = Y,$$

i.e., $\Delta_{p,q}$ is a semimetric on the collection of finite-measure subsets $\mathfrak{M}_{<\infty}(\mathcal{S})$. Finally, for disjoint X and Y , it is clear that $\text{GD}_{p,q}(X, Y \setminus X) = \text{GD}_{p,q}(X, Y)$; thus, by Theorem 1, the triangle inequality holds for both arguments inside the maximum that defines $\Delta_{p,q}$ when $p, q \in [1, \infty)$. This implies that the triangle inequality is also valid for $\Delta_{p,q}$. \square

Theorem 2. Let $X, Y, Z \in \mathfrak{M}_{<\infty}(\mathcal{S})$ be subsets admitting positive constants $r < R$ such that $r \leq d_{u,v} \leq R$ for any $u \in X \cup Y$ and $v \in Y \cup Z$. Then, for all $p, q \in \mathbb{R} \setminus \{0\}$, $|p|, |q| \geq 1$ and at least one of them negative a relaxed triangle inequality holds for $\text{GD}_{p,q}$, namely

$$\text{GD}_{p,q}(X, Z) \leq \frac{R^2}{r^2} \left(\text{GD}_{p,q}(X, Y) + \text{GD}_{p,q}(Y, Z) \right).$$

Proof. Step 1: Let $p \in \mathbb{R} \setminus \{0\}$, and suppose that $q < 0$. We will prove that

$$\text{GD}_{p,|q|}(X, Y) \leq \frac{R}{r} \text{GD}_{p,q}(X, Y). \tag{6}$$

For all $x \in X$ and $y, z \in Y$, we have $\frac{r}{R} \leq \frac{d_{x,y}}{d_{x,z}} \leq \frac{R}{r}$. Thus,

$$\frac{R}{r} \geq \left(\int_Y \int_Y \left[\frac{d_{x,y}}{d_{x,z}} \right]^{|q|} dy dz \right)^{\frac{1}{|q|}} = \left(\int_Y d_{x,y}^{|q|} dy \right)^{\frac{1}{|q|}} \left(\int_Y d_{x,z}^{-|q|} dz \right)^{\frac{1}{|q|}}.$$

Since $q = -|q|$, this means that $\mathcal{M}^{|q|}(d_{x,y}) \leq \frac{R}{r} \mathcal{M}^q(d_{x,y})$. Taking the p -average \mathcal{M}^p at both sides and from 1 and 4 of Proposition 2, we find $\mathcal{M}^p_{x \in X}(\mathcal{M}^{|q|}(d_{x,y})) \leq \frac{R}{r} \mathcal{M}^p_{x \in X}(\mathcal{M}^q(d_{x,y}))$, which is exactly Equation (6).

Step 2: Now, for $q \in \mathbb{R} \setminus \{0\}$ and $p < 0$, we will prove that

$$GD_{|p|,q}(X, Y) \leq \frac{R}{r} GD_{p,q}(X, Y). \tag{7}$$

By assumption, we have $\frac{r}{R} \leq \frac{d_{x,y}}{d_{u,y}} \leq \frac{R}{r}$ for any $y \in Y$ and all $x, u \in X$. Similarly as before and using 1 and 4 of Proposition 2, we conclude from the RHS part that $\mathcal{M}^q(d_{x,y}) \leq \frac{R}{r} \mathcal{M}^q(d_{u,y})$. However, since $p = -|p|$, after taking a p -average of the quotient of means, it follows from

$$\left(\int_X \left(\mathcal{M}^q(d_{x,y}) \right)^{|p|} dx \right)^{\frac{1}{|p|}} \left(\int_X \left(\mathcal{M}^q(d_{u,y}) \right)^p du \right)^{\frac{1}{|p|}} = \left(\int_X \int_X \left[\frac{\mathcal{M}^q(d_{x,y})}{\mathcal{M}^q(d_{u,y})} \right]^{|p|} dx du \right)^{\frac{1}{|p|}} \leq \frac{R}{r},$$

that $\mathcal{M}^{|p|}_{x \in X}(\mathcal{M}^q(d_{x,y})) \leq \frac{R}{r} \mathcal{M}^p_{x \in X}(\mathcal{M}^q(d_{x,y}))$, which is now (7).

Step 3: The previous steps can be summarized in the expression

$$GD_{|p|,|q|}(X, Y) \leq \frac{R}{r} GD_{|p|,q}(X, Y) \leq \frac{R^2}{r^2} GD_{p,q}(X, Y). \tag{8}$$

Using again 4 of Proposition 2 and Definition 6, we get $GD_{p,q}(X, Z) \leq GD_{|p|,|q|}(X, Z)$. From this, the subadditivity for $GD_{|p|,|q|}$ (Theorem 1), and Equation (8), we conclude

$$GD_{p,q}(X, Z) \leq GD_{|p|,|q|}(X, Y) + GD_{|p|,|q|}(Y, Z) \leq \frac{R^2}{r^2} (GD_{p,q}(X, Y) + GD_{p,q}(Y, Z)). \quad \square$$

Remark 3. For parameters $(p, q) \in \mathbb{R}^2$ that lie in the orange or blue sectors in Figure 2, the distance $GD_{p,q}$ fulfills a C relaxed triangle inequality for a constant $C = R^2/r^2$ only if the condition $r \leq d_{u,v} \leq R$ holds for all $u \in X \cup Y$ and $v \in Y \cup Z$. On bounded and topologically separated sets (i.e., not having common limit points), this condition always holds, and on them, $\Delta_{p,q}$ becomes an inframetric as explained below.

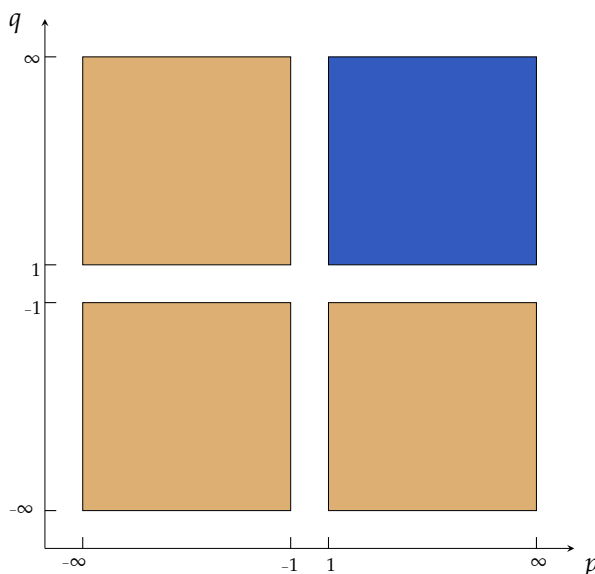


Figure 2. According to Corollary 1, when acting on disjoint subsets, $\Delta_{p,q}$ behaves as a proper metric if (p, q) lies in the blue sector, and according to Corollary 2, it behaves like an inframetric if (p, q) lies in the orange sectors.

Corollary 2. Under the same hypotheses of Theorem 2, the (p, q) -averaged Hausdorff distance $\Delta_{p,q}$ satisfies

$$\Delta_{p,q}(X, Z) \leq \frac{R^2}{r^2} \left(\Delta_{p,q}(X, Y) + \Delta_{p,q}(Y, Z) \right).$$

Proof. It is immediate using Theorem 2 and Definition 7. \square

When the involved sets are finite, a generally sharper inframetric relation holds. For emphasis, we employ in this context the notation A, B, C for those subsets of \mathcal{S} .

Theorem 3. If $p, q \in \mathbb{R}$ and $|p|, |q| > 1$, the (p, q) -distance $\text{GD}_{p,q}$ satisfies the relaxed triangle inequality

$$\text{GD}_{p,q}(A, C) \leq N^\alpha \left(\text{GD}_{p,q}(A, B) + \text{GD}_{p,q}(B, C) \right),$$

for all finite subsets $A, B, C \subset \mathcal{S}$, where $N := \max\{|A|, |B|, |C|\} \geq 1$ and $\alpha := |p|^{-1} + |q|^{-1}$.

Proof. For arbitrary $p \neq 0$, let us assume that $q < 0$, so that $|q| = -q$. We can write

$$\text{GD}_{p,|q|}(A, B) = \left(\sum_{a \in A} \left(\left[\sum_{b \in B} d(a, b)^{-q} \right]^{-1} \right)^{\frac{p}{q}} \right)^{\frac{1}{p}} = \left(\sum_{a \in A} \left(|B|^{-2} \text{harm} \left\{ d(a, b)^q \right\} \right)^{\frac{p}{q}} \right)^{\frac{1}{p}},$$

which, when combined with property 5 of Proposition 1, yields

$$\text{GD}_{p,|q|}(A, B) \leq \left(\sum_{a \in A} \left(|B|^{-1} \min_{b \in B} \left\{ d(a, b)^q \right\} \right)^{\frac{p}{q}} \right)^{\frac{1}{p}} \leq |B|^{\frac{1}{|q|}} \left(\sum_{a \in A} \left(\sum_{b \in B} d(a, b)^q \right)^{\frac{p}{q}} \right)^{\frac{1}{p}} = |B|^{\frac{1}{|q|}} \text{GD}_{p,q}(A, B).$$

A similar relation is true for any $q \neq 0$ if $p < 0$. In conclusion, $\text{GD}_{|p|,|q|}(A, B) \leq N^\alpha \text{GD}_{p,q}(A, B)$, where

$$\alpha := \begin{cases} |\min\{p, q\}|^{-1} & \text{if } pq < 0, \\ |p|^{-1} + |q|^{-1} & \text{if } p < 0, q < 0. \end{cases}$$

If N^α does not need to be sharp, α can always be chosen to take the larger value $|p|^{-1} + |q|^{-1}$.

Now, for $|p|, |q| \geq 1$, the final result follows from the triangle inequality for $GD_{|p|,|q|}$:

$$GD_{p,q}(A, C) \leq GD_{|p|,|q|}(A, C) \leq N^\alpha \left(GD_{p,q}(A, B) + GD_{p,q}(B, C) \right). \quad \square$$

Corollary 3. *If $p, q \in \mathbb{R}$ and $|p|, |q| \geq 1$, the (p, q) -distance $\Delta_{p,q}$ satisfies the relaxed triangle inequality:*

$$\Delta_{p,q}(A, C) \leq N^\alpha \left(\Delta_{p,q}(A, B) + \Delta_{p,q}(B, C) \right),$$

for all finite subsets $A, B, C \subset \mathcal{S}$, with $N := \max\{|A|, |B|, |C|\} \geq 1$, and $\alpha := |p|^{-1} + |q|^{-1}$.

Proof. The corollary follows immediately from Theorem 3 and Definition 5. \square

To conclude this section, we return to the general setting of arbitrary measurable sets to explain the behavior of $\Delta_{p,q}$ when changing the value of its parameters p and q .

Theorem 4. *Let $X, Y \in \mathfrak{M}_{<\infty}(\mathcal{S})$ and suppose that $p, p', q, q' \in \overline{\mathbb{R}}$ satisfy $p \leq p'$ and $q \leq q'$. Then,*

$$\Delta_{p,q}(X, Y) \leq \Delta_{p',q}(X, Y) \quad \text{and} \quad \Delta_{p,q}(X, Y) \leq \Delta_{p,q'}(X, Y).$$

Proof. It follows easily from part 5 of Proposition 2 and Definition 7. \square

6. The (p, q) -Distances as Quality Indicators

Let Q be a decision space Q and $F: Q \rightarrow \mathbb{R}^k$ be a multi-objective function on it, of which the associated MOP consists in the simultaneous minimization of its k component functions f_1, \dots, f_k . A candidate solution to this problem is *Pareto-optimal* if all elements of its image in $F(Q) \subset \mathbb{R}^k$ are *nondominated* in the sense of Pareto [31]; see Definition 1. For the forthcoming discussion, let us introduce the following abbreviated and useful notation. For $X, Y \subset Q$ and any $z \in Q$, we define the following

$$\begin{aligned} X_{\preceq z} &:= \{x \in X \mid x \preceq z\}, & X_{\preceq Y} &:= \{x \in X \mid \exists y \in Y : x \preceq y\}, \\ X_{\not\preceq z} &:= \{x \in X \mid x \not\preceq z\}, & X_{\not\preceq Y} &:= \{x \in X \mid \nexists y \in Y : x \preceq y\}. \end{aligned}$$

From these definitions, it follows that, for arbitrary $z \in Q$ and $X, Y \subset Q$, there are partitions:

$$X = X_{\preceq z} \sqcup X_{\not\preceq z}, \quad \text{and} \quad X = X_{\preceq Y} \sqcup X_{\not\preceq Y},$$

where \sqcup stands for the disjoint union of subsets. A similar notation with the subindices \prec, \succ , and \succeq can also be used in an analogous way. Let us recall that an archive $X \subset Q$ is, by definition, a subset of mutually non-dominated points; therefore, for any $x, x' \in X$, the condition $x \preceq x'$ implies $x = x'$. This basic property implies that $F: Q \rightarrow \mathbb{R}^k$ is a bijection when restricted to any archive $X \subset Q$ and, therefore, the points in $F(X) \subset F(Q)$ can be univocally labeled by the elements of X . Moreover, for a finite archive $A \subset Q$, both sets have the same number of elements $|A| = |F(A)|$.

Now, we introduce a couple of strengthened notions of dominance between sets (archives) that are required for the validity of most of the results in this section.

Definition 8. *An archive X is well-dominated by an archive Y if*

1. X is dominated by Y , written $Y \succeq X$, i.e., $\forall x \in X, \exists y \in Y$ s.t. $y \preceq x$, and
2. Y consists only of dominating points of X , i.e., $\forall y \in Y, \exists x \in X$ s.t. $y \preceq x$.

Moreover, X is said to be strictly well dominated by Y if

3. $\exists y \in Y \setminus X, \exists x \in X \setminus Y$ such that $y \prec x$.

For an archive $X \subset Q$, the $GD_{p,q}$, $IGD_{p,q}$, and $\Delta_{p,q}$ quality (or performance) indicators assigned to it will be defined as the distance of its image $F(X)$ to the Pareto front $F(P)$, i.e.,

$$\mathcal{I}_{p,q}^{GD}(X) := GD_{p,q}(F(X), F(P)), \quad \mathcal{I}_{p,q}^{IGD}(X) := IGD_{p,q}(F(X), F(P)),$$

$$\text{and } \mathcal{I}_{p,q}^{\Delta}(X) := \Delta_{p,q}(F(X), F(P)).$$

In this section, we study the behavior of $\mathcal{I}_{p,q}^{GD}$, $\mathcal{I}_{p,q}^{IGD}$, and $\mathcal{I}_{p,q}^{\Delta}$ as performance indicators. An example of a weakly Pareto-compliant performance indicator is the Degree of Approximation (DOA; see Reference [10]).

6.1. Pareto Compliance of (p, q) -Indicators in the Finite Case

In order to obtain general conclusions on the features of the averaged Hausdorff distance $\Delta_{p,q}$ as a quality indicator, we consider first the behavior of $GD_{p,q}$. For additional details on the material presented in this section and other related results in the context of the p -averaged Hausdorff Distance Δ_p , the reader is referred to Reference [28].

For the following statements, we will abbreviate $\delta_q(a, B) := (\sum_{b \in B} d(a, b)^q)^{\frac{1}{q}}$. Clearly, with this notation, $\mathcal{I}_{p,q}^{GD}(A) = (\sum_{a \in A} \delta_q(F(a), F(P))^p)^{\frac{1}{p}}$, where in the averaged sum \sum , we are labeling the points in $F(A)$ by the elements of the archive A , taking advantage of the fact that $|A| = |F(A)|$, as it also will be done with all the averages in this section.

Theorem 5. *Let $A, B \subset Q$ be finite archives with A strictly well dominated by B . For all $a \in A$ and $b \in B$,*

1. $b \prec a$ implies that $\delta_q(F(b), F(P)) < \delta_q(F(a), F(P))$;
2. $b \preceq a$ implies that $\frac{|B_{\preceq a}|}{|B|} \leq \frac{|A_{\succeq b}|}{|A|}$ (or equivalently an strict equality);

then, $\mathcal{I}_{p,q}^{GD}(B) < \mathcal{I}_{p,q}^{GD}(A)$.

Proof. By condition 1, for all $a \in A$ and $b \in B_{\preceq a}$, the inequality $\delta_q(F(b), F(P))^p \leq \delta_q(F(a), F(P))^p$ holds true. After averaging over all $b \in B_{\preceq a}$ at both sides, we have

$$\sum_{b \in B_{\preceq a}} \delta_q(F(b), F(P))^p \leq \delta_q(F(a), F(P))^p,$$

and averaging once again over all $a \in A$ produces

$$\sum_{a \in A} \left(\sum_{b \in B_{\preceq a}} \delta_q(F(b), F(P))^p \right) \leq \sum_{a \in A} \delta_q(F(a), F(P))^p. \tag{9}$$

From property 2 and noticing that each $b \in B$ appears $|A_{\succeq b}|$ times in the initial sum, the LHS becomes

$$\sum_{a \in A} \sum_{b \in B_{\preceq a}} \frac{1}{|B_{\preceq a}|} \delta_q(F(b), F(P))^p \geq \sum_{a \in A} \sum_{b \in B_{\preceq a}} \frac{|A|}{|B|} \frac{1}{|A_{\succeq b}|} \delta_q(F(b), F(P))^p = \sum_{b \in B} \delta_q(F(b), F(P))^p.$$

Returning to Equation (9), we conclude that $\mathcal{I}_{p,q}^{GD}(B) \leq \mathcal{I}_{p,q}^{GD}(A)$. Lastly, part 3 of Definition 8 for strictly well-dominated sets guarantees that this is an strict inequality, proving the assertion. \square

Remark 4. *Since we are dealing with finite archives, condition 2 of Theorem 5 regarding the relative size of some of their parts is equivalent to the condition $|F(B_{\preceq a})|/|F(B)| \leq |F(A_{\succeq b})|/|F(A)|$ regarding the relative size of their images. This is not necessarily the case in the context of measurable subsets, but see Remark 5.*

Figure 3 shows examples where Theorem 5 holds with $q \rightarrow -\infty$. In this case, the q -averaged distance $\delta_q(a, B)$ becomes the standard distance $d(a, B)$ between a point and a set.

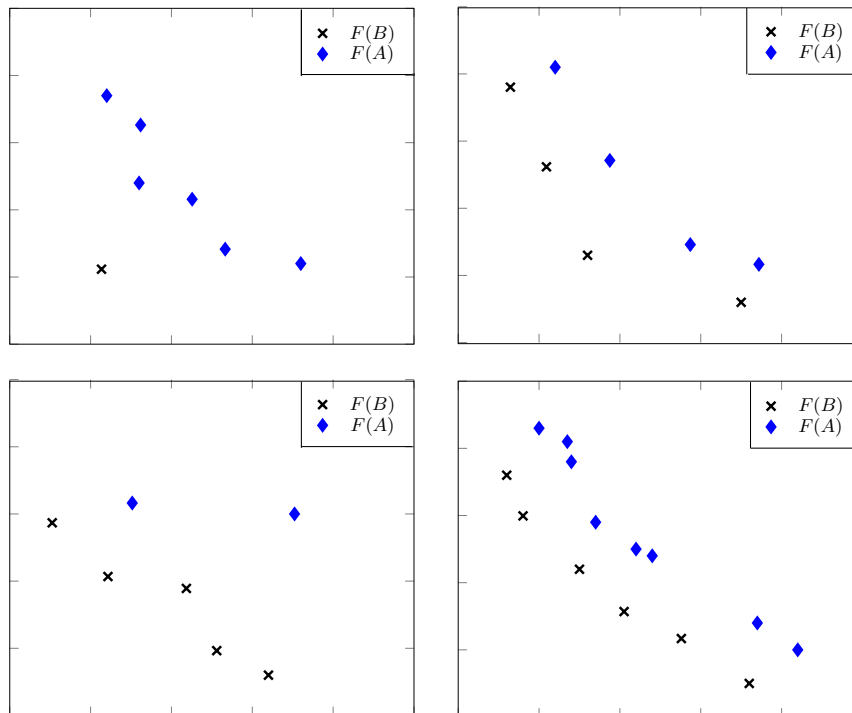


Figure 3. Different scenarios where the GD_p value of archive B is better (smaller) than the GD_p value of archive A independently of the Pareto set and where the additional assumptions made in Theorem 5 are easily verifiable.

For the inverted generational distance $IGD_{p,q}$ in the finite case, we provide here two useful results without explicit proofs. The necessary steps are similar to the arguments used to prove the analogous statements for IGD_p in Prop. 3.8 of Reference [28] and Thm. 3.9 in Reference [28]. Those statements correspond here to the limit $q \rightarrow -\infty$, and the main difference in the proofs is that the Euclidean distance $d(a, B)$ needs to be changed everywhere by the q -average $\delta_q(a, B)$, as it was done above for the proof of Theorem 5 that generalizes the proof of Thm. 3.4 in Reference [28]. The reader can also find there additional remarks on similar hypotheses to the ones needed for Theorem 6 below.

Proposition 3. Let $A, B \subset Q$ be finite and strictly well-dominated archives with $B \preceq A$ such that for all $a \in A, b \in B$, and $x \in P$:

$$b \prec a \text{ implies } \delta_q(F(b), F(x)) < \delta_q(F(a), F(x));$$

then, $\mathcal{I}_{p,q}^{IGD}(B) < \mathcal{I}_{p,q}^{IGD}(A)$.

Figure 4 illustrates two situations where the hypotheses of Proposition 3 are satisfied. Now, to state a more general result concerning the Pareto compliance of the $IGD_{p,q}$ indicator, we will further abbreviate the minimal p -average of distances $\delta_q(F(a), F(P_{\neq B}))$ by

$$\delta_A := \min_{a \in A} \left\{ \left(\sum_{x \in P_{\neq B}} \delta_q(F(x), F(a))^p \right)^{\frac{1}{p}} \right\} = \min_{a \in A} IGD_{p,q}(F(a), F(P_{\neq B})).$$

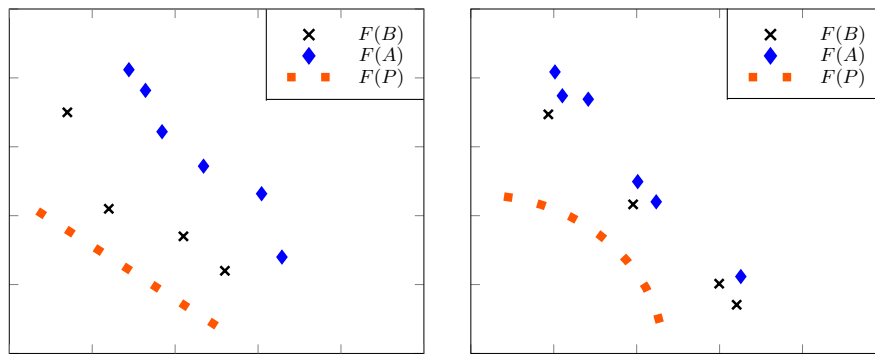


Figure 4. Two situations where $IGD_{p,q}(B)$ is better (smaller) than $IGD_{p,q}(A)$ for sufficiently negative q : Here, the hypotheses of Proposition 3 hold true.

Theorem 6. Let $A, B \subset Q$ be finite and strictly well-dominated archives such that $B \preceq A$. If at least one of the following conditions is satisfied,

1. $\forall a \in A, \forall b \in B: b \prec a$ implies $IGD_{p,q}(F(b), F(P_{\neq B})) < IGD_{p,q}(F(a), F(P_{\neq B}))$;
2. $\exists a_0 \in A$ such that $\forall x \in P_{\neq B}: a_0 \in \arg \min_{a \in A} \delta_q(F(x), F(a))$;
3. $\forall x \in P_{\neq B}: \delta_q(F(A), F(x)) = \delta_A$;

then $\mathcal{I}_{p,q}^{IGD}(B) < \mathcal{I}_{p,q}^{IGD}(A)$.

Finally, a general statement on the Pareto compliance of the finite case of the (p, q) -averaged Hausdorff distance $\Delta_{p,q}$ follows as a consequence of Theorems 5 and 6.

Theorem 7. Let $A, B \subset Q$ be finite and well-dominated archives such that $B \preceq A$. If for all $a \in A, b \in B$:

$$b \preceq a \text{ implies } \frac{|B_{\preceq a}|}{|B|} = \frac{|A_{\succeq b}|}{|A|},$$

and at least one of the following conditions is satisfied:

1. $\forall a \in A, \forall b \in B, \forall x \in P: b \prec a$ implies $\delta_q(F(b), F(x)) < \delta_q(F(a), F(x))$;
2. $\exists a_0 \in A$ such that $\forall x \in P_{\neq B}: a_0 \in \arg \min_{a \in A} \delta_q(F(x), F(a))$;
3. $\forall x \in P_{\neq B}: \delta_q(F(A), F(x)) = \delta_A$;

then, $\mathcal{I}_{p,q}^\Delta(B) < \mathcal{I}_{p,q}^\Delta(A)$.

Figure 5 illustrates four situations where Theorems 6 and 7 apply with very large q . In the first row, the left diagram is a modification of the second case in Figure 4 where condition 1 holds. In the right diagram, the diamond lying at the lower left corner of $F(A)$ represents the image $F(a_0)$ of a point a_0 satisfying condition 2. Finally, both diagrams in the second row exhibit cases where the points of $F(P)$ are equidistant to corresponding points in $F(A)$, making condition 3 valid, with δ_A being this distance.

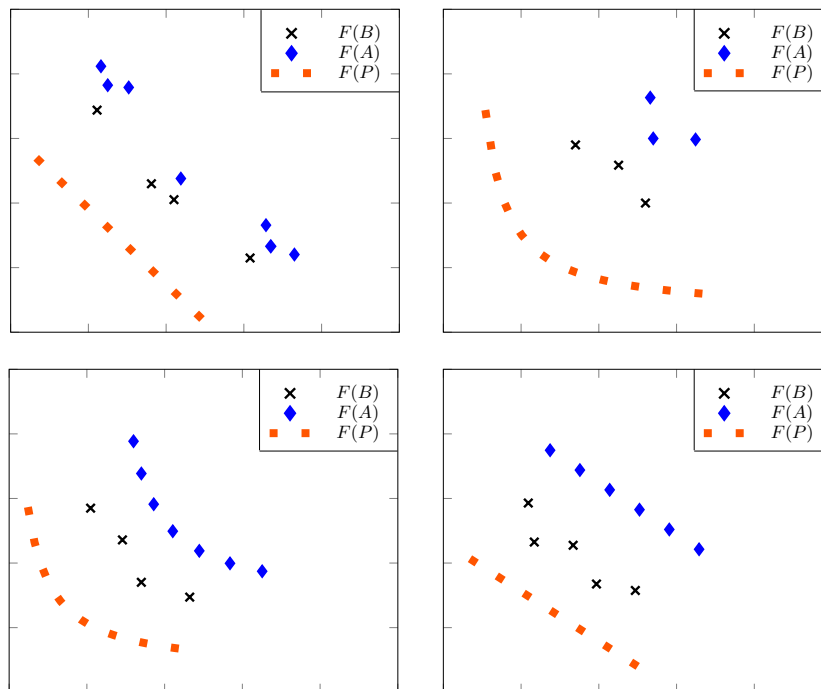


Figure 5. Four examples where $IGD_p(B)$ is smaller (better) than $IGD_{p,q}(A)$ for sufficiently negative q : In each case, at least one of the requirements of Theorem 6 is satisfied.

6.2. Pareto Compliance of (p, q) -Indicators in the General Case

We consider now the behavior of the generalized $GD_{p,q}$ distance with respect to the Pareto-compliance, concentrating on the most important aspects that describe its characteristics and using similar hypotheses to the ones needed in the previous section for $\Delta_{p,q}$ in the finite case.

Here, we will continue to assume that the decision space Q with objective function $F: Q \rightarrow \mathbb{R}^k$ defining the MOP under consideration has a Pareto set $P \subset Q$ with corresponding Pareto front $F(P) \subset F(Q)$. Also, we assume that the objective space $F(Q) \subset \mathbb{R}^k$ carries a metric d that, for simplicity, can be taken to be the one inherited from the Euclidean distance $d(\cdot, \cdot)$ in \mathbb{R}^k . In addition, to define the (p, q) -indicators on MOPs that require general non-finite sets, we need a measure space (\mathcal{S}, μ) , that here will be taken to be $\mathcal{S} := F(Q)$ endowed with a non-null measure μ according to the comments at the beginning of Section 4.2. In this context $X, Y \subset Q$ will denote arbitrary subsets such that $F(X), F(Y) \subset F(Q)$ are measurable with non-null and finite measures.

Remark 5. Recall that, here, $|F(X)| = \mu(F(X))$ denotes the measure of $F(X) \subset \mathcal{S}$. In this context, Q will not be asked to carry a measure and the notation $|X|$ will have no a priori meaning for $X \subset Q$. Nevertheless, it is possible to induce a measure on those subsets of Q where F is bijective by taking the pullback μ^* of μ to them, making the identity $|X| = \mu^*(X) := \mu(F(X)) = |F(X)|$ trivially true. This can be done for all archives but not for subsets where F is not bijective. When it is Q that carries a measure, a push-forward measure can be always defined on its image $F(Q)$, making this identity true for all sets. This was implicitly assumed in the presentation provided in Reference [18] (Section 3.4). For clarity, we avoid here this identification and state everything from the assumption that the measure μ is defined only on $S := F(Q)$.

Before stating the complete result, let us recall that a partition of a set X is a collection of disjoint and non-empty subsets of X whose union is the whole of X and a partition of an archive $X \subset Q$ induces a partition of $F(X) \subset F(Q)$ by the bijectivity of F restricted to X . For convenience, we abbreviate the measure-theoretic q -averaged distance from a point $F(x) \in F(Q)$ to a set $F(Z) \subset F(Q)$ by $\delta_q(F(x), F(Z)) := (\int_{v \in F(Z)} d(F(x), v)^q)^{\frac{1}{q}}$.

Theorem 8. For $p, q \in \overline{\mathbb{R}}$, let $X, Y \subset Q$ denote archives of which the images $F(X)$ and $F(Y)$ are of non-null finite measures in $F(Q)$. Moreover, assume that

1. there exist finite partitions $X = \bigsqcup_{i=1}^m X_i$ and $Y = \bigsqcup_{i=1}^m Y_i$ such that $\forall i \in \{1, \dots, m\}$:
 - (a) $F(X_i) \subset F(X)$ and $F(Y_i) \subset F(Y)$ are subsets of non-null finite measure in $F(Q)$;
 - (b) $\forall x \in X_i, \forall y \in Y_i: x \preceq y$;
2. $\forall x \in X, \forall y \in Y: x \preceq y$ implies $\delta_q(F(x), F(P)) \leq \delta_q(F(y), F(P))$;

then, $\mathcal{I}_{p,q}^{GD}(X) \leq \mathcal{I}_{p,q}^{GD}(Y)$.

Proof. By 1(a) of Theorem 8, the sets X and Y can be subdivided into the same number m of subsets, and by 1(b), if $x \in X_i$ and $y \in Y_i$ for any $i \in \{1, \dots, m\}$, then $\delta_q(F(x), F(P)) \leq \delta_q(F(y), F(P))$. Therefore, we can take successive integral p -averages over $F(X_i)$ and, afterwards, over $F(Y_i)$ at both sides of this inequality to find that, for each i , we have

$$a_i^p := \frac{1}{|F(X_i)|} \int_{F(X_i)} \delta_q(v, F(P))^p \, dv \leq \frac{1}{|F(Y_i)|} \int_{F(Y_i)} \delta_q(v, F(P))^p \, dv =: b_i^p. \tag{10}$$

For those $i \in \{1, \dots, m\}$ violating the inequality $\frac{|X_i|}{|X|} \leq \frac{|Y_i|}{|Y|}$, we subdivide X_i into a sufficiently large partition of m_i subsets $X_{i,1}, X_{i,2}, \dots, X_{i,m_i}$, with images by F of non-null finite measure, so as to guarantee that, for all $j \in \{1, \dots, m\}$, we get

$$w_{i,j} := \frac{|F(X_{i,j})|}{|F(X)|} \leq \frac{|F(Y_i)|}{|F(Y)|} =: \tilde{w}_i. \tag{11}$$

Notice that this is indeed possible because each $F(X_i)$ is of non-null finite measure. Since $\forall x \in X_i, \forall y \in Y_i$, we have $x \preceq y$, an inequality similar to Equation (10) also holds for them, i.e.,

$$a_{i,j}^p := \frac{1}{|F(X_{i,j})|} \int_{F(X_{i,j})} \delta_q(v, F(P))^p \, dv \leq \frac{1}{|F(Y_i)|} \int_{F(Y_i)} \delta_q(v, F(P))^p \, dv =: b_i^p,$$

for all $i \in \{1, \dots, n\}, j \in \{1, \dots, m_i\}$. However, $|F(X)| = \sum_{i=1}^m |F(X_i)|$, where $|F(X_i)| = \sum_{j=1}^{m_i} |F(X_{i,j})|$ and $|F(Y)| = \sum_{i=1}^m |F(Y_i)|$. Therefore, with the notation of Equation (11), a simple calculation shows that $\sum_{i=1}^m \sum_{j=1}^{m_i} w_{i,j} = \sum_{i=1}^m \tilde{w}_i = 1$, implying that $w_{i,j}$ and \tilde{w}_i are normalized weights useful for weighted averages. Since $0 \leq a_{i,j} \leq b_i$ and $0 \leq w_{i,j} \leq \tilde{w}_i \leq 1$, simple properties of discrete weighted power mean imply the inequality $\sum_{i=1}^m \sum_{j=1}^{m_i} w_{i,j} a_{i,j}^p \leq \sum_{i=1}^m \tilde{w}_i b_i^p$. Thus, we can finally write

$$\begin{aligned} \mathcal{I}_{p,q}^{GD}(F(Y))^p &= \frac{1}{|F(X)|} \sum_{i=1}^m \sum_{j=1}^{m_i} \int_{F(X_{i,j})} \delta_q(v, F(P))^p \, dv = \sum_{i=1}^m \sum_{j=1}^{m_i} \frac{|F(X_{i,j})|}{|F(X)|} a_{i,j}^p = \sum_{i=1}^m \sum_{j=1}^{m_i} w_{i,j} a_{i,j}^p \\ &\leq \sum_{i=1}^m \tilde{w}_i b_i^p = \sum_{i=1}^m \frac{|F(Y_i)|}{|F(Y)|} b_i^p = \frac{1}{|F(Y)|} \sum_{i=1}^m \int_{F(Y_i)} \delta_q(v, F(P))^p \, dv = \mathcal{I}_{p,q}^{GD}(Y)^p. \quad \square \end{aligned}$$

Remark 6. From condition 1 of Theorem 8, it follows the simpler (and somewhat weaker) dominance conditions:

- (a') $X \preceq Y$ (i.e., $\forall y \in Y, \exists x \in X$ such that $x \preceq y$), and
- (b') $\forall x \in X, \exists y \in Y$ such that $x \preceq y$.

For simple situations where (a') and (b') are valid, the partitions needed for part 1 of Theorem 8 are not difficult to find; however, this is not always possible as the right side of Figure 6 indicates. Indeed, Figure 6 presents some examples where (a') and (b') hold true, but $\mathcal{I}_{p,q}^{GD}(X) \leq \mathcal{I}_{p,q}^{GD}(Y)$ can be both true (left side) and false (right side). Furthermore, it is possible to show that X and Y comply (left side) and do not comply (right side) with condition 1 of Theorem 8, respectively.

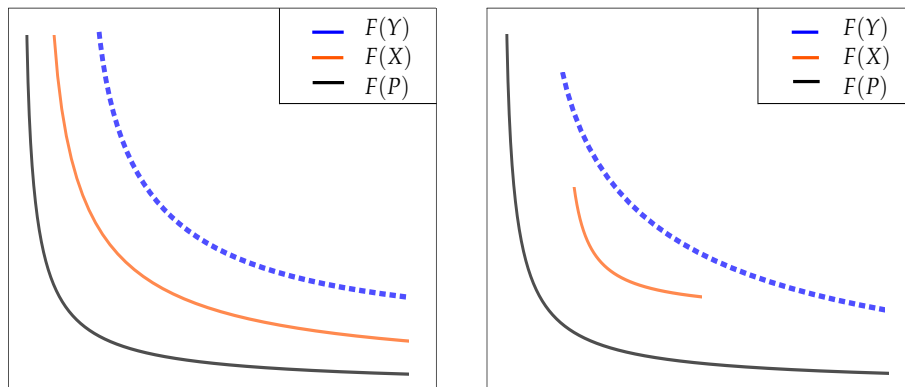


Figure 6. (Left) A situation where the Pareto front $F(P)$ and the images $F(X)$ and $F(Y)$ of continuous archives satisfy $\mathcal{I}_{p,q}^{GD}(X) \leq \mathcal{I}_{p,q}^{GD}(Y)$ and condition 1(a) of Theorem 8 holds true. (Right) A modification of the previous situation where conditions (a') and (b') of Remark 6 are satisfied but $\mathcal{I}_{p,q}^{GD}(X) \not\leq \mathcal{I}_{p,q}^{GD}(Y)$. Here, there are no possible partitions of the archives satisfying part 1(a) of Theorem 8.

Remark 7. An important advantage of using $GD_{p,q}$ over GD_p is that condition 2 of Theorem 8 provides the possibility of choosing an appropriate $q \in \mathbb{R}$ for which the condition $\delta_q(F(x), F(P)) \leq \delta_q(F(y), F(P))$ holds when $x \preceq y$, ensuring in this way the compliance to Pareto optimality for $GD_{p,q}$. This freedom is lacking for GD_p because, in the limit $q \rightarrow -\infty$, the distance $\delta_q(F(x), F(P))$ becomes the standard distance $d(F(x), F(P))$, which does not allow for any choice.

7. Examples and Numerical Experiments

In this section, we present some numerical experiments involving finite sets first, and afterwards, we study the case of continuous sets.

7.1. Working with $\Delta_{p,q}$ over Finite Sets

Let us take a hypothetical Pareto front P given by the line segment from $(0, 1)$ to $(1, 0)$ in \mathbb{R}^2 , i.e., the set of all points

$$(t, 1 - t) \in \mathbb{R}^2, \quad \text{for } 0 \leq t \leq 1.$$

This is the same example considered in Reference [16] p. 506 and enables us to make a comparison with values of Δ_p . In order to use the finite version of $\Delta_{p,q}$, we discretize P by taking 11 uniformly distributed points; we call this set P' . We assume two archives: X_1 is obtained from P' by changing $(0, 1)$ for $(0, 10)$, including an outlier, and by adding $1/10$ to the remaining ordinates. X_2 is obtained from P' by adding 5 to each ordinate. See Figure 7.

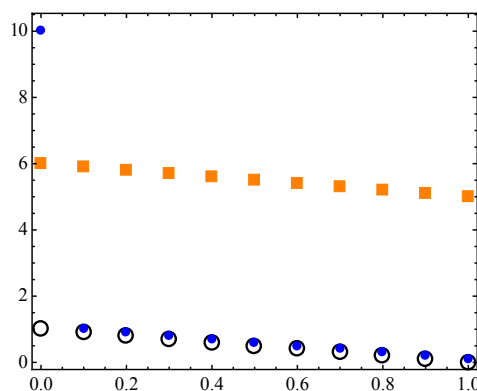


Figure 7. A hypothetical Pareto front discretization P' (black circles) and two different archives: X_1 (blue dots) and X_2 (orange squares).

As explained in Section 3, we know that

$$\Delta_\infty(A, B) := \lim_{p \rightarrow \infty} \Delta_p(A, B)$$

coincides with the standard Hausdorff distance d_H . In this case, we obtained

$$\begin{aligned} \Delta_1(P', X_1) &= 0.9091, & \Delta_1(P', X_2) &= 4.5412, \\ d_H(P', X_1) &= 9, & d_H(P', X_2) &= 5; \end{aligned}$$

and according to Theorem 4 and Reference [16] p.512, these values must increase as p increases.

Tables 1 and 2 show that we can find values of p and q such that the (p, q) -averaged distance does not punish heavily the outliers, for example, $p = q = 1$ or $p = 1$ and $q = -1$. We remark that the values of $\Delta_{p,q}(P', X_1)$ do not present a significative change under variations of $q \leq 1$ for a fixed p . Thus, it is possible to work with $q = 1$, in which case $\Delta_{p,q}$ is a metric according to Corollary 1, and to still obtain values close to the ones given by the inframetric Δ_p , with the same $p \geq 1$.

Table 1. $\Delta_{p,q}(P', X_1)$ for several values of p and q .

$q \backslash p$	1	2	5	10	20
$-\infty$	0.9091	2.7153	5.5714	7.0811	7.9831
-100	0.9272	2.7701	5.6839	7.2241	8.1443
-20	0.9537	2.8367	5.8202	7.3974	8.3396
-5	0.9895	2.8624	5.8705	7.4613	8.4117
-1	1.1131	2.8782	5.8848	7.4795	8.4322
1	1.3243	2.9112	5.8920	7.4886	8.4425
2	2.9277	2.9295	5.8956	7.4932	8.4476
5	5.8920	5.8956	5.9063	7.5068	8.4630
10	7.4886	7.4932	7.5068	7.5292	8.4882

Table 2. $\Delta_{p,q}(P', X_2)$ for several values of p and q .

$q \backslash p$	1	2	5	10	20
$-\infty$	4.5412	4.5497	4.5751	4.6160	4.6867
-100	4.6442	4.6529	4.6790	4.7209	4.7933
-20	4.8425	4.8518	4.8795	4.9239	5.0003
-5	4.9624	4.9720	5.0007	5.0465	5.1250
-1	5.0008	5.0105	5.0394	5.0856	5.1646
1	5.0203	5.0301	5.0591	5.1055	5.1848
2	5.0301	5.0398	5.0690	5.1154	5.1949
5	5.0591	5.0690	5.0983	5.1450	5.2248
10	5.1055	5.1154	5.1450	5.1921	5.2725

For large values of p , the behavior of $\Delta_{p,q}$ presents the same disadvantages of Δ_p or of the standard Hausdorff distance. For example, in Tables 1 and 2, it can be observed that all distances for $p \geq 5$ are useless because they imply that the distance from the discrete Pareto front P' to the archive X_1 is larger than its distance to the archive X_2 . Figure 7 suggests that this is an undesirable outcome.

Table 3 shows that $\Delta_{p,q}$ is close to a metric when $q \leq -1$ and $p \geq 1$. The percentage of triangle inequality violations decreases as p increases or q decreases.

Table 3. Triangle inequality violations, in percentage, for several values of p and q : Here, we randomly chose 80 sets, each one containing 2 points in $[0, 10]^2$, and verified the triangle inequality for all possible set permutations (that is, 492,960).

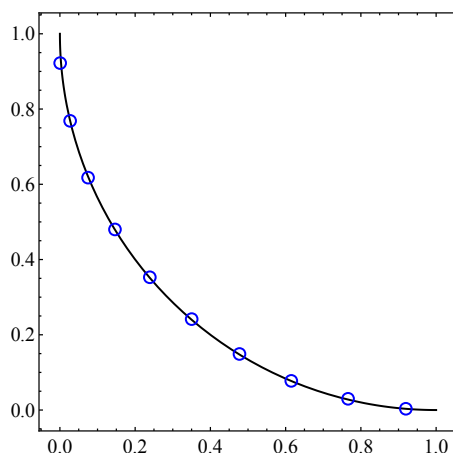
$q \backslash p$	1	2	5	10
-1	0.05396	0	0	0
-2	0.10265	0.00041	0	0
-5	0.28815	0.01217	0	0
-10	0.35622	0.05031	0.00041	0
-20	0.43046	0.08439	0.00446	0.00041

7.2. Optimal Archives for Discretized Spherical Pareto Fronts

We now consider two standard Pareto fronts: The spheric convex and spheric concave quarter-circles, see Figures 8 and 9.

$$P_1 = \left\{ (\cos(\theta) + 1, \sin(\theta) + 1) : -\pi \leq \theta \leq -\frac{\pi}{2} \right\},$$

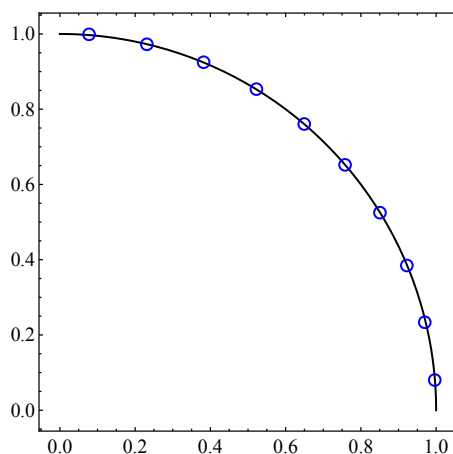
$$P_2 = \left\{ (\cos(\theta), \sin(\theta)) : 0 \leq \theta \leq \frac{\pi}{2} \right\}. \tag{12}$$



0.003083	0.921541
0.027631	0.766555
0.076121	0.617317
0.147360	0.477502
0.239594	0.350552
0.350552	0.239595
0.477502	0.147360
0.617317	0.076121
0.766555	0.027631
0.921541	0.003083

$$\Delta_{1,-1}(A, P_1) = 0.179112$$

Figure 8. Optimal $\Delta_{1,-1}$ archive A for the connected Pareto front P_1 given by Equation (12) with 10 elements (blue circles) and at the right is the respective archive coordinates and the $\Delta_{1,-1}$ distance.



0.996918	0.078459
0.972370	0.233445
0.923880	0.382684
0.852639	0.522498
0.760405	0.649448
0.649448	0.760405
0.522498	0.852639
0.382683	0.923879
0.233445	0.972370
0.078459	0.996917

$$\Delta_{1,-1}(A, P_1) = 0.174482$$

Figure 9. Optimal $\Delta_{1,-1}$ archive A for the connected Pareto front P_2 given by Equation (12) with 10 elements (blue circles) and at the right is the respective archive coordinates and the $\Delta_{1,-1}$ distance.

To numerically find the optimal $\Delta_{p,q}$ archive of size M , we discretized the Pareto front with 1000 equidistant points (which is an acceptable discretization according to Reference [63] p. 603) and randomly chose an initial M sized archive. Then, we used a random-walk (or step climber) evolutionary algorithm, moving one point at a time. Finally, we refined the optimal archive with the “evenly spaced” construction suggested by Reference [63] p. 607.

When finding optimal $\Delta_{p,q}$ archives, our numerical experiments suggest a clear geometrical influence of the parameters p and q . For values of p in $(-\infty, -1)$, the optimal archive sets are basically the same. When $q \in [-1, 1]$ increases, the optimal archive tends to lose dispersion, converging to one point. When $q \geq 1$, the optimal archive collapses to one point, and when $q \in (-\infty, -1]$, the corresponding optimal archives are basically the same (see Figure 10). When $p \geq -1$ increases, the optimal archive moves away from the Pareto set (see Figure 11).

The following Figures 10 and 11 show certain “optimal” archives A for the Pareto front P_1 in Equation (12), where the optimality means that the distance $\Delta_{p,q}(X, P_1)$ is minimum when $X = A$. Because of the choice of the parameters p and q , this solution is clearly different from the one shown in the Figure 8.

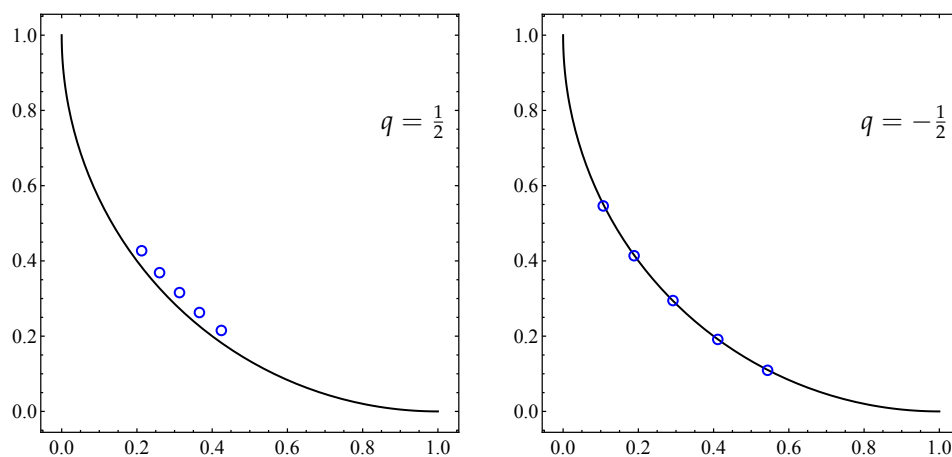


Figure 10. Optimal $\Delta_{1,q}$ five-point set archives A for the connected Pareto front P_1 given by Equation (12) with $p = 1$ and $q = \pm 1/2$.

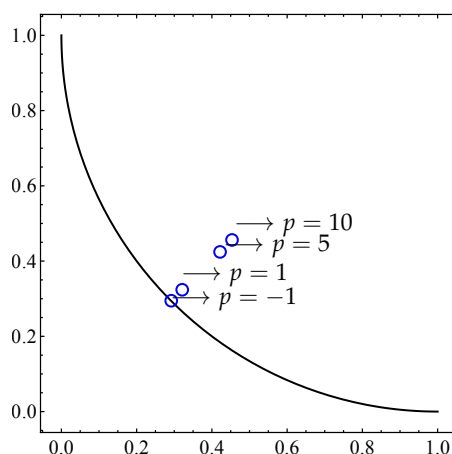


Figure 11. Optimal $\Delta_{p,-1}$ one-point archives A for the connected Pareto front P_1 given by Equation (12) with $q = -1$ and different values of p : In all cases, the archives are located in the line $x = y$.

7.3. Optimal Archives for Disconnected and Discretized Pareto Sets

In this section, we present the optimal $\Delta_{p,q}$ archives for a disconnected step Pareto front:

$$P_3^{(s,\gamma)} = \left\{ \left(t, 1 - \gamma t + (\gamma - 1) \frac{\lfloor st \rfloor}{s} \right) : 0 \leq t \leq 1 \right\}, \tag{13}$$

where s is the number of steps, $\gamma > 0$ is a small constant responsible for the step’s twist, and $\lfloor \cdot \rfloor$ stands for the integer part function.

Figure 12 shows numerical optimal $\Delta_{1,-1}$ archives of sizes 20. The archive coordinates reveal that

$$A \cap P_3^{(5, \frac{1}{10})} = \emptyset,$$

i.e., the optimal archive points do not lie over the Pareto front but they are so close to it that this is hardly noticeable. It is also evident that the archives are evenly distributed along the Pareto front.

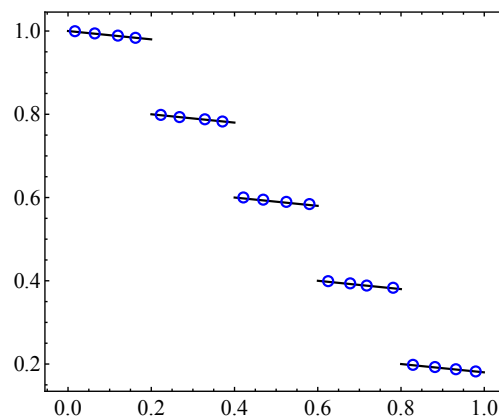


Figure 12. Numerical optimal $\Delta_{1,-1}$ archive A for the disconnect step Pareto front $P_3^{(5)}$ given by Equation (13) with 20 elements: here, we obtain $\Delta_{1,-1} \left(A, P_3^{(5, \frac{1}{10})} \right) = 0.111132$.

7.4. General Example for Continuous Sets

In this first example, we are going to construct simple and illustrative continuous sets A and B . Let A be the straight segment in \mathbb{R}^2 from $a = (-1, 0)$ to $b = (1, 0)$, that is

$$A = \overline{ab}. \tag{14}$$

For a small positive $\varepsilon > 0$ and a variable $\delta > 0$, let $B_\delta \subset \mathbb{R}^2$ be the set given by the following union of straight segments

$$B_\delta = \overline{cd_\delta} \cup \overline{e_\delta f_\delta} \cup \overline{g_\delta h}, \tag{15}$$

where $c = (-1, \varepsilon)$, $d_\delta = (-\delta, \varepsilon)$, $e_\delta = (-\delta, 1)$, $f_\delta = (\delta, 1)$, $g_\delta = (\delta, \varepsilon)$, and $h = (1, \varepsilon)$. We can regard the set B_δ as a continuous approximation of A , where the central segment $\overline{e_\delta f_\delta}$ can be seen as the outlier.

In the following Figure 13, we can see the sets A and B_δ for $\varepsilon = 0.10$ and $\delta = 0.10, 0.20$. According to Table 4, as δ decreases, the $\Delta_{p,q}$ distance between the approximation B_δ and the set A also decreases.

We remark that the classical Hausdorff distance between A and B_δ produces the value 1 for any $\delta > 0$. Thus, by working with the (p, q) -distance instead of d_H , we can detect “better” approximations.

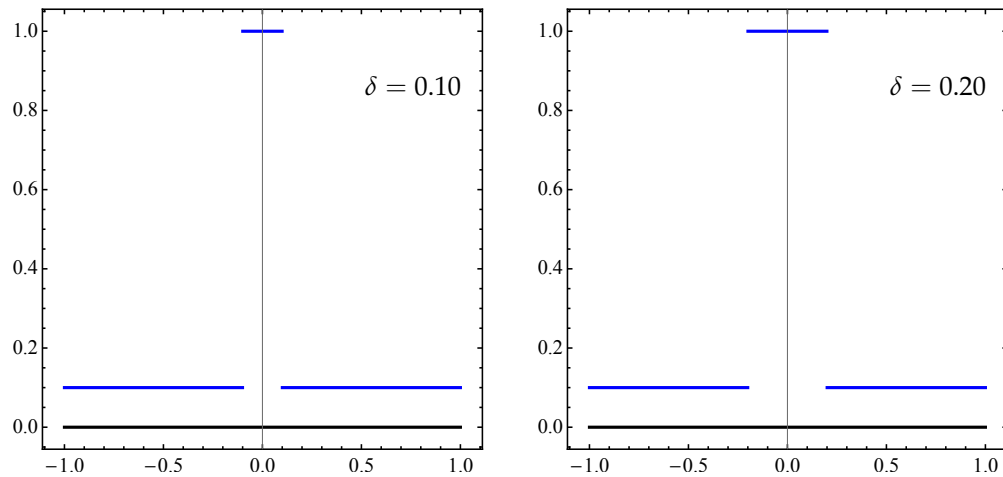


Figure 13. The black horizontal segment is the set A from Equation (14), and the blue piecewise map is the respective approximation given by the set B_δ from Equation (15) for two values of δ and $\varepsilon = 0.10$.

Table 4. $\Delta_{p,q}$ results between the sets A and B_δ in Equations (14) and (15) for $\varepsilon = 0.10$ and some parameter values of p, q , and δ .

p	q	$\Delta_{pq}(A, B_{0.05})$	$\Delta_{pq}(A, B_{0.10})$	$\Delta_{pq}(A, B_{0.20})$	$\Delta_{pq}(A, B_{0.40})$
1	1	0.7149	0.7464	0.8091	0.9324
1	1	0.4105	0.4506	0.5311	0.6945
1	100	0.1503	0.1961	0.2878	0.4711
1	200	0.1479	0.1934	0.2844	0.4663
1	10,000	0.1451	0.1901	0.2802	0.4602

7.5. Approximating Pareto Set and Front of a MOP

Finally, we address the problem to approximate the Pareto sets and fronts of given MOPs. As an example, we will consider the bi-objective *Lamé super-sphere problem* [32] which is defined as follows:

$$\min_x F: \mathbb{R}^n \longrightarrow \mathbb{R}^2, \tag{16}$$

where $F(x) = (f_1(x), f_2(x))$ is defined as

$$f_1(x) = \left(\frac{1}{n} \sum_{i=1}^n x_i^2 \right)^{\frac{\gamma}{2}} \quad \text{and} \quad f_2(x) = \left(\frac{1}{n} \sum_{i=1}^n (x_i - 1)^2 \right)^{\frac{\gamma}{2}}$$

where $x \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$. For $n = 2$, the Pareto sets and fronts of this problem are shown in Figures 14 and 15 for $\gamma = 2$ and $\gamma = 1/2$, respectively.

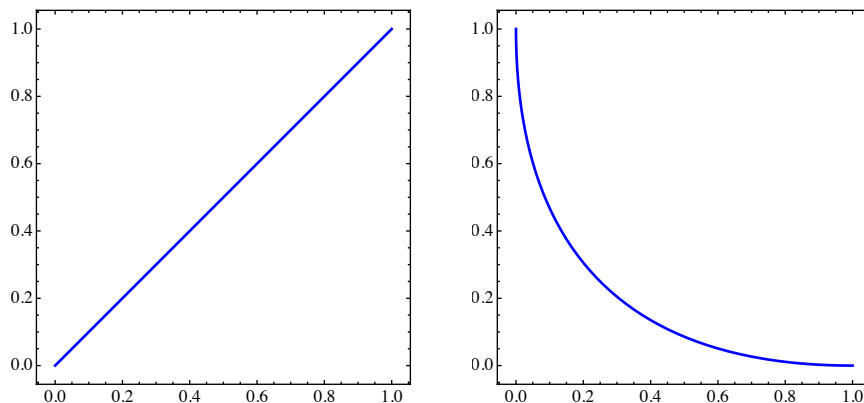


Figure 14. (Left) Pareto set. (Right) Pareto front of MOP (Equation (16)) for $n = 2$ and $\gamma = 2$.

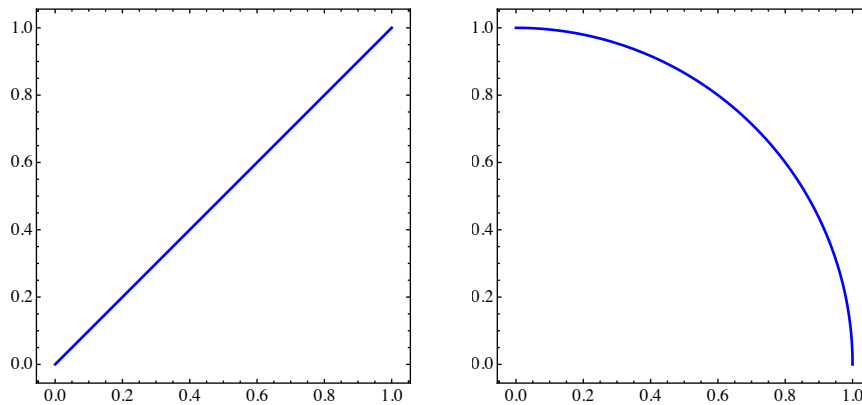


Figure 15. The same as in Figure 14 but for $\gamma = 1/2$.

In a first step, we discuss the principle difference of discrete and continuous archives when approximating the Pareto set/front on a hypothetical example. For this, we assume that we are given the 5-element archive $A = \{x_1, \dots, x_5\} \subset \mathbb{R}^2$; those elements are given by

$$x_1 = (-0.02, 0.03), \quad x_2 = (0.29, 0.20), \quad x_3 = (0.41, 0.49), \quad x_4 = (0.70, 0.62), \quad x_5 = (1.02, 0.98).$$

Hence, we can see A as a 5-element approximation of the Pareto set and its image $F(A)$ as a 5-element approximation of the Pareto front. Now, instead of A , we may use a polygonal curve that is defined by A :

$$B := \overline{x_1x_2} \cup \dots \cup \overline{x_4x_5}.$$

In the following, we will call A a discrete archive while we call the polygon approximation B the continuous archive. Figures 16 and 17 show the approximations A and B as well as their images $F(A)$ and $F(B)$. Apparently, the approximation qualities are much better for the linear interpolates. This impression gets confirmed by the values of $\Delta_{p,q}$ for this problem that are shown in Table 5. We can observe the following two behaviors: (i) the distances are much better for the continuous archives and the differences are even larger in objective space, and (ii) the distances decrease with decreasing q (which is in accordance to the result of Theorem 4).

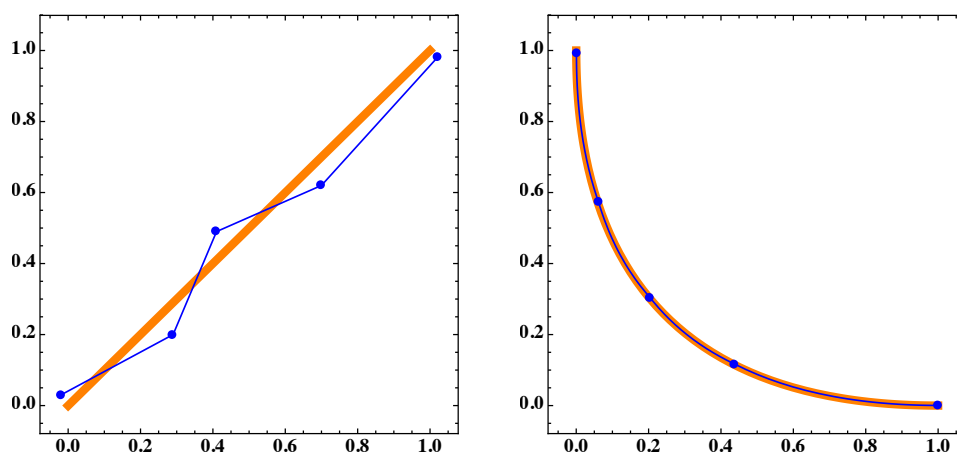


Figure 16. (Left) The blue dots A and the blue polygonal line B are the discrete and continuous approximations, respectively, for the Pareto set which corresponds to the orange thick segment, of MOP (Equation (16)) for $n = 2$. (Right) respective sets $F(A)$ and $F(B)$ of the Pareto front for $\gamma = 2$.

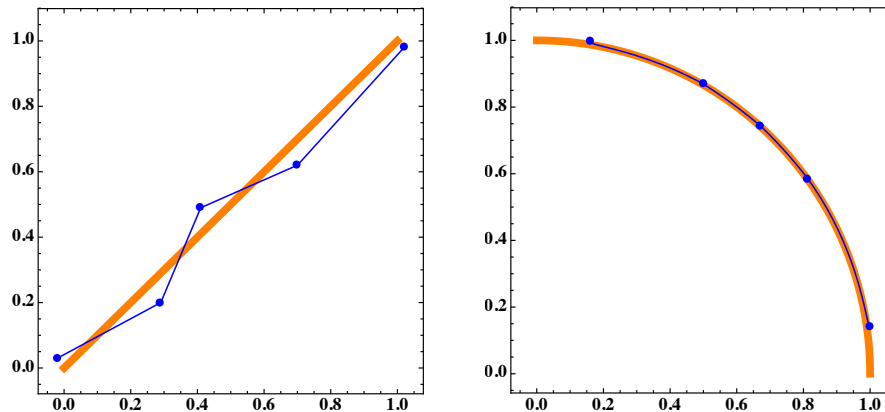


Figure 17. The same as in Figure 16 but for $\gamma = 1/2$.

Table 5. $\Delta_{p,q}$ results for the approximations of the Pareto set and front for MOP (Equation (16)).

	p	q	Decision Space		Objective Space	
			Finite Arch.	Cont. Arch.	Finite Arch.	Cont. Arch.
$\gamma = 2$	1	1	0.5262	0.4775	0.4377	0.3851
	1	1	0.2710	0.2017	0.2051	0.1070
	1	100	0.1121	0.0341	0.0862	0.0040
	1	200	0.1112	0.0333	0.0855	0.0039
	1	10,000	0.1103	0.0324	0.0848	0.0038
$\gamma = \frac{1}{2}$	1	1	0.5262	0.4775	0.5520	0.4965
	1	1	0.2710	0.2017	0.2587	0.1120
	1	100	0.1121	0.0341	0.1079	0.0012
	1	200	0.1112	0.0333	0.1071	0.0012
	1	10,000	0.1103	0.0324	0.1062	0.0011

In a next step, we consider discrete archives that have been generated from multi-objective evolutionary algorithms together with their resulting continuous archives. For multi-objective evolutionary algorithms, we have chosen the widely used methods NSGA-II [65] and MOEA/D [66]. We stress, however, that any other MOEA could be chosen and that the conclusions we draw out by our results apply in principle to any other such algorithm. Table 6 shows the parameter setting we have used for our studies.

Table 6. Parameter setting for NSGA-II and MOEA/D: Here, n denotes the dimension of the decision variable space.

Algorithm	Parameter	Value
NSGA-II	Population size	12
	Number of generations	500
	Crossover probability	0.8
	Mutation probability	$1/n$
	Distribution index for crossover	20
	Distribution index for mutation	20
MOEA/D	Population size	12
	# weight vectors	12
	Number of generations	500
	Crossover probability	1
	Mutation probability	$1/n$
	Distribution index for crossover	30
	Distribution index for mutation	20
	Aggregation function	Tchebycheff
Neighborhood size	3	

Figures 18 and 19 and Table 7 show the results of NSGA-II where we have used 500 generations and a population size of 12. Figures 20 and 21 and Table 8 show the respective results for MOEA/D where we have also used 500 generations and population size 12. We can see that, for both algorithms, the $\Delta_{p,q}$ values are significantly better for the continuous archives. We can also make another observation: the $\Delta_{p,q}$ values oscillate for the results of the dominance-based algorithm NSGA-II which is indeed typical. For the continuous archives, these oscillations are less notorious, which indicates that the use of continuous archives may have a smoothing effect on the approximations, which is highly desired.

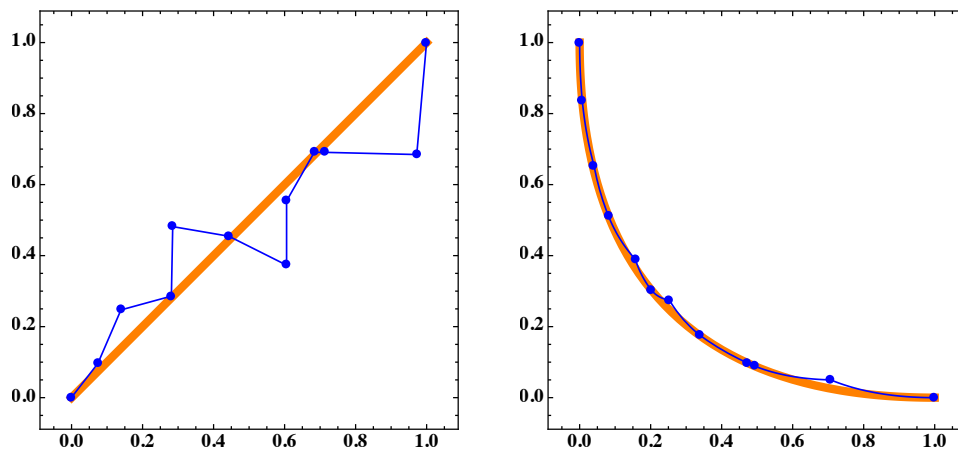


Figure 18. (Left) the blue dots A and the blue polygonal line are the discrete and continuous approximations, respectively, for the Pareto set and the orange thick segment is for the 410th generation of the NSGA-II algorithm of MOP (Equation (16)) for $n = 2$. (Right) corresponding sets $F(A)$ and $F(B)$ of the Pareto front for $\gamma = 2$.

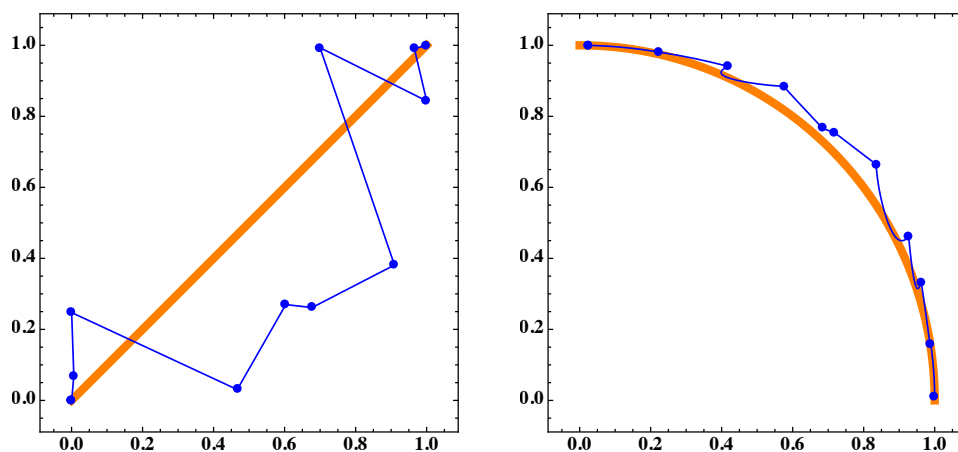


Figure 19. The same as in Figure 18 but for $\gamma = 1/2$.

Table 7. For MOP (Equation (16)), the Table shows the $\Delta_{p,q}$ results for the finite and continuous Pareto front approximations. We used the NSGA-II generated archives for $p = 1$ and $q = -10$.

Generation	$\gamma = 1/2$		$\gamma = 2$	
	Finite Arch.	Cont. Arch.	Finite Arch.	Cont. Arch.
50	0.0439	0.0147	0.0696	0.0160
100	0.0498	0.0109	0.0540	0.0102
200	0.0613	0.0118	0.0716	0.0207
250	0.0651	0.0265	0.0572	0.0061
400	0.0602	0.0102	0.0723	0.0276
450	0.0630	0.0154	0.0584	0.0088
460	0.0612	0.0154	0.0658	0.0098
470	0.0523	0.0102	0.0566	0.0083
480	0.0754	0.0269	0.0684	0.0241
490	0.0510	0.0091	0.0584	0.0118
500	0.0722	0.0097	0.0560	0.0103

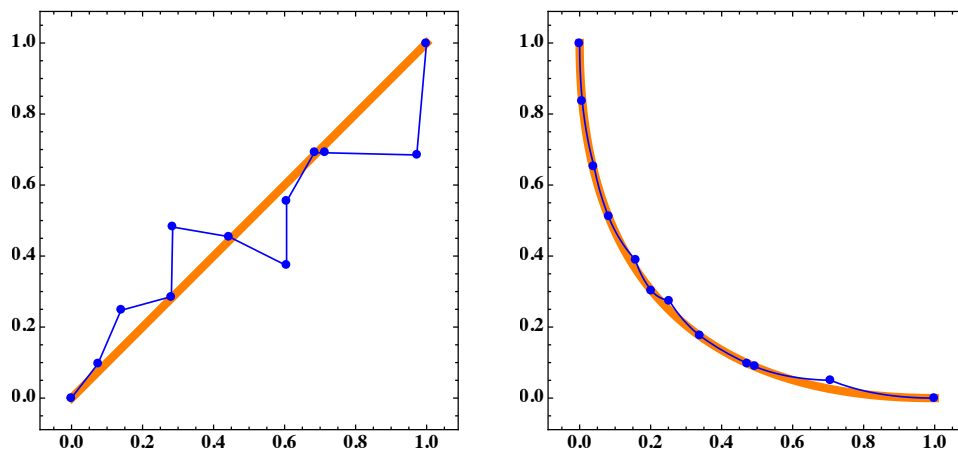


Figure 20. (Left) the blue dots A and the blue polygonal line are the discrete and continuous approximations, respectively, for the Pareto set and the orange thick segment is for the 410th generation of the MOEA/D algorithm of MOP (Equation (16)) for $n = 2$. (Right) corresponding sets $F(A)$ and $F(B)$ of the Pareto front for $\gamma = 2$.

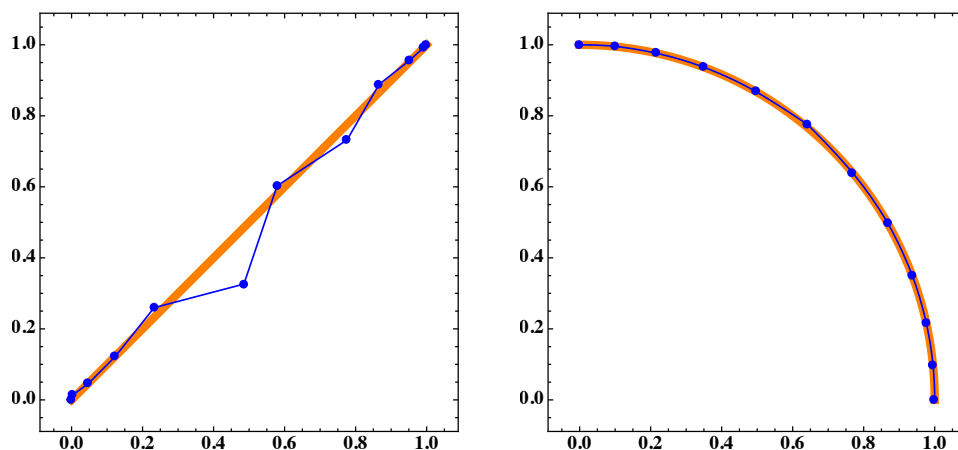


Figure 21. The same as in Figure 20 but for $\gamma = 1/2$.

Table 8. For MOP (Equation (16)), the Table shows the $\Delta_{p,q}$ results for the finite and continuous Pareto front approximations. We used the MOEA/D generated archives for $p = 1$ and $q = -10$.

Generation	$\gamma = 1/2$		$\gamma = 2$	
	Finite Arch.	Cont. Arch.	Finite Arch.	Cont. Arch.
50	0.0610	0.0171	0.0648	0.0119
100	0.0519	0.0051	0.1093	0.0016
200	0.0536	0.0037	0.0781	0.0009
250	0.0522	0.0037	0.0790	0.0008
400	0.0511	0.0017	0.0784	0.0009
450	0.0511	0.0017	0.0784	0.0009
460	0.0509	0.0012	0.0784	0.0009
470	0.0509	0.0012	0.0784	0.0009
480	0.0509	0.0010	0.0783	0.0009
490	0.0509	0.0010	0.0783	0.0009
500	0.0509	0.0010	0.0783	0.0009

We want to investigate the last statement further on. To this end, we consider the following convex bi-objective problem: the objectives are given by $f_1, f_2: \mathbb{R}^3 \rightarrow \mathbb{R}$, where

$$\begin{aligned} f_1(x) &= (x_1 + 1)^2 + x_2^2 + x_3^2 \\ f_2(x) &= (x_1 - 1)^2 + x_2^2 + x_3^2. \end{aligned} \tag{17}$$

Figure 22 shows the Pareto set and front of MOP (Equation (17)). The Pareto set is given by the straight segment joining $(0, 0, 0)$ and $(1, 0, 0)$.

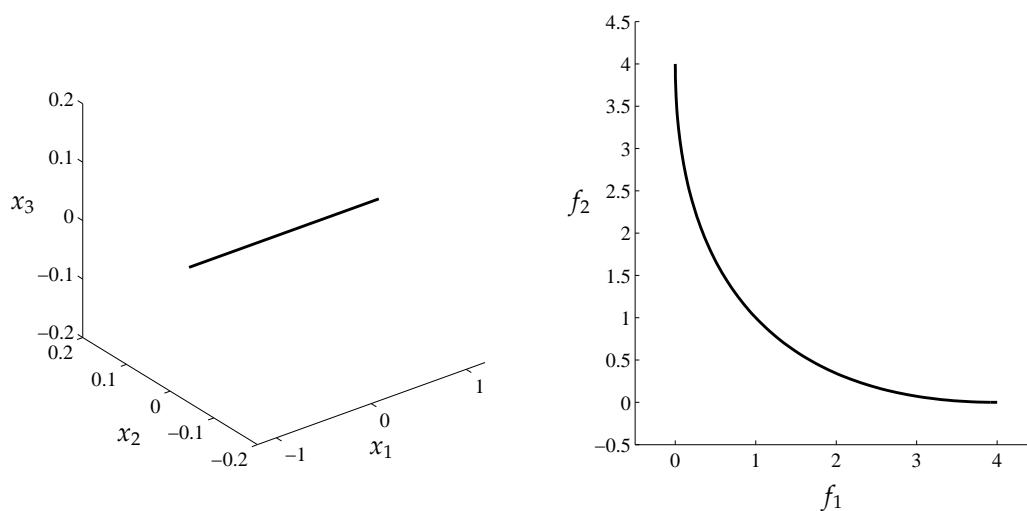


Figure 22. (Left) Pareto set. (Right) Pareto front of MOP (Equation (17)).

The values of $\Delta_{p,q}$ obtained by NSGA-II for the discrete archives (using population size 20) as well as for the respective continuous archives can be seen in Figure 23 and Table 9. Also for this example, the values for the continuous archives are much better and the oscillations are significantly reduced compared to the discrete archives.

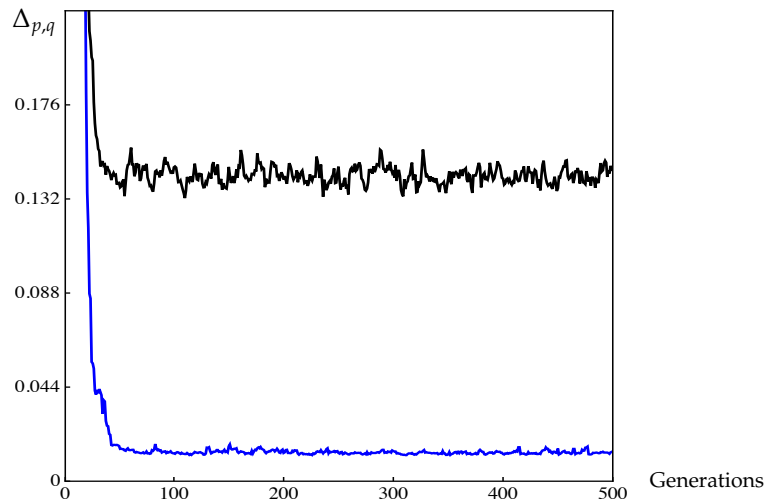


Figure 23. The black curve is the $\Delta_{p,q}$ value for the discrete approximation, and the blue one is the respective curve for the continuous approximation of NSGA-II for MOP (Equation (17)).

Figures 24–26 show the results of both kinds of archives after 300, 400, and 500 generations which confirm this observation. The results show that NSGA-II is indeed capable of computing points near the Pareto front while the distribution of the points vary. This is a known fact since there exists no “limit archive” for this algorithm (as it is, e.g., not based on the averaged Hausdorff distance or any other performance indicator). When considering the respective results of the continuous archives, however, NSGA-II computes (at least visually) nearly perfect approximations of the Pareto front. The $\Delta_{p,q}$ values reflect this.

Table 9. $\Delta_{p,q}$ results between the Pareto Front and its respective discrete and continuous approximations of NSGA-II for MOP (Equation (17)): The data shown is the averaged over the 20 independent runs above.

Generation	Continuous Archive	Finite Archive
20	0.1333	0.2401
40	0.0176	0.1451
60	0.0090	0.1561
80	0.0088	0.1355
100	0.0065	0.1472
120	0.0074	0.1412
140	0.0081	0.1395
160	0.0075	0.1549
180	0.0092	0.1468
200	0.0074	0.1429
220	0.0066	0.1408
240	0.0075	0.1397
260	0.0066	0.1460
280	0.0074	0.1439
300	0.0084	0.1421
320	0.0070	0.1352
340	0.0070	0.1373
360	0.0081	0.1454
380	0.0079	0.1413
400	0.0066	0.1388
420	0.0063	0.1400
440	0.0097	0.1384
460	0.0067	0.1418
480	0.0067	0.1421
500	0.0076	0.1426

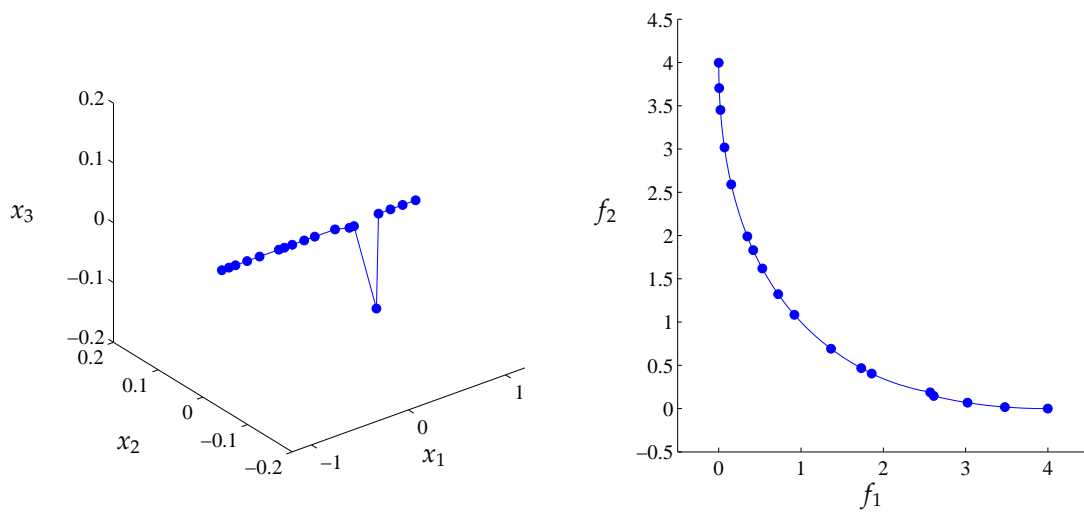


Figure 24. (Left) The blue dots and the blue polygon line are the discrete and continuous approximation, respectively, for the Pareto set of MOP (Equation (17)) in the 300th generation. (Right) respective sets $F(A)$ and $F(B)$ of the Pareto front.

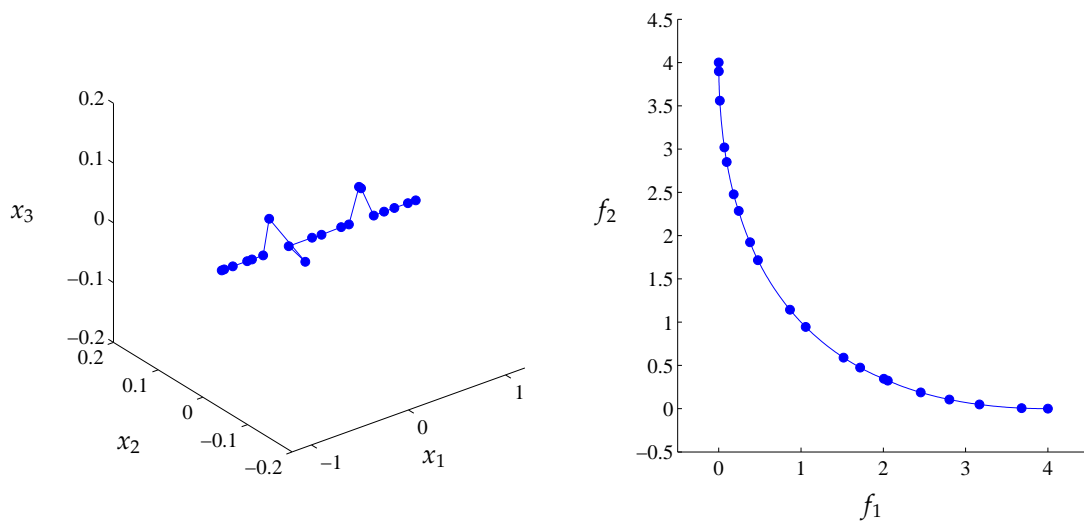


Figure 25. The same as in Figure 24 but for the 400th generation.

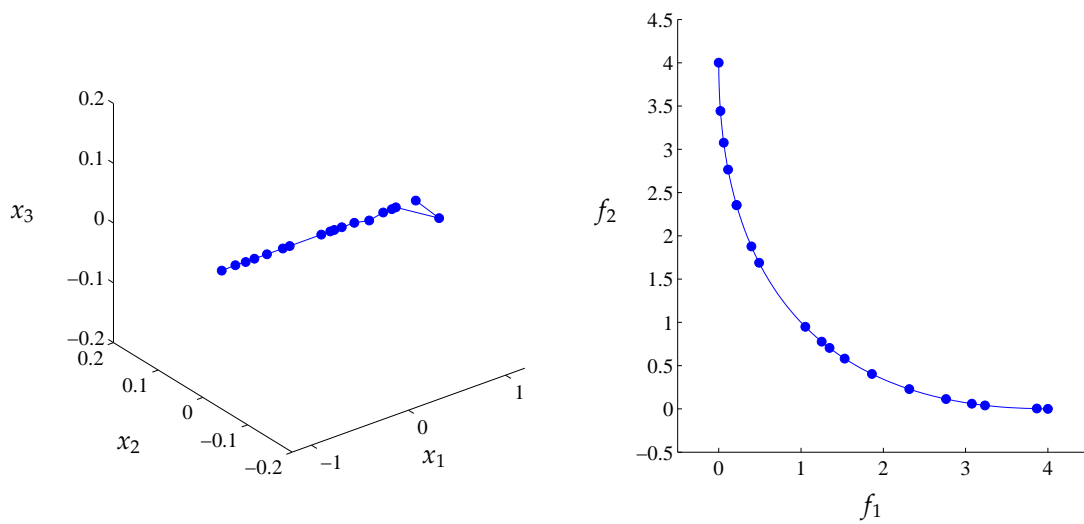


Figure 26. The same as in Figure 24 but for the 500th generation.

8. Conclusions and Perspectives

In this paper, we have presented a comprehensive overview of the averaged Hausdorff distances that have recently appeared in connection with the study of MOPs.

Among the averaged Hausdorff distances studied here, the generalized $\Delta_{p,q}$ as defined for arbitrary measurable sets was shown to provide a general and robust definition for applications that carries good metric properties, is adequate for use with continuous approximations of the Pareto set of a MOP, and even reduces to the previously introduced definition for discrete approximations.

Concerning the appearance of the additional parameter q in the definition of $\Delta_{p,q}$ which could give the impression of an overly complicated expression, it is important to highlight, as it was observed in Remark 7, that it can provide the possibility to choose a suitable value of q in order to make $GD_{p,q}$ as Pareto compliant as possible for the MOP under consideration. This is an argument in favor of the flexibility provided by the generalized version $GD_{p,q}$, which is not available for the GD_p distance, and this particular aspect worths further investigation.

Nevertheless, since the freedom provided by the two parameters p and q may appear as excessive and perhaps undesirable in many applications, there remains to find a practical recipe to determine and fix these parameters according to the characteristics of the problem under consideration. Certainly, the desired spreads of the optimal archives, the distance of an approximation to the Pareto front, and the convexity of these fronts need to be taken into account in order to determine an appropriate set of preferred values for these parameters depending on the situation.

To achieve these aims, more theoretical as well as numerical studies of optimal solutions associated with Pareto fronts with different convexities must be carried out and experiments evaluating how the Pareto compliance can be enhanced in each situation by the choice of parameters need to be performed.

Finally, we stress that the results we have shown in Section 7 show the advantage of a new performance indicator that is able to compute the performance of a continuous approximation of the solution set. Continuous approximations, e.g., of the Pareto set/front of multi-objective optimization problems have not been considered so far, though both Pareto set and front typically form continuous sets in case the objectives are continuous. The examples have indicated that the consideration of continuous archives (via use of interpolation on the populations generated by the evolutionary algorithms) could allow a reduction in population sizes and, hence, a significant reduction of the computational effort of the evolutionary algorithms. This is because the time complexity for all existing multi-objective evolutionary algorithms is quadratic in the population size and in each generation of the algorithm. To verify this statement, more computations are needed, which is left for future work.

Author Contributions: J.M.B. and A.V. obtained the theoretical results concerning the (p, q) -averaged Hausdorff distance, and O.S. conceived and designed the experiments; J.M.B. and A.V. performed the experiments and provided the related figures and tables; O.S. analyzed the data and contributed with the text. J.M.B. and A.V. wrote the paper.

Funding: The first two authors were partially supported by Vicerrectoría de Investigación, Pontificia Universidad Javeriana, Bogotá D.C., Colombia. The third author was supported by Conacyt Basic Science project No. 285599 and SEP Cinvestav project no. 231.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Heinonen, J. *Lectures on Analysis on Metric Spaces*; Springer: New York, NY, USA, 2001.
2. de Carvalho, F.; de Souza, R.; Chavent, M.; Lechevallier, Y. Adaptive Hausdorff distances and dynamic clustering of symbolic interval data. *Pattern Recogn. Lett.* **2006**, *27*, 167–179. [[CrossRef](#)]
3. Huttenlocher, D.P.; Klanderman, G.A.; Rucklidge, W.A. Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **1993**, *15*, 850–863. [[CrossRef](#)]
4. Yi, X.; Camps, O.I. Line-based recognition using a multidimensional Hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 901–916.

5. Falconer, K. *Fractal Geometry: Mathematical Foundations and Applications*, 2nd ed.; Mathematical foundations and applications; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2003.
6. Aulbach, B.; Rasmussen, M.; Siegmund, S. Approximation of attractors of nonautonomous dynamical systems. *Discrete Contin. Dyn. Syst. Ser. B* **2005**, *5*, 215–238.
7. Dellnitz, M.; Hohmann, A. A subdivision algorithm for the computation of unstable manifolds and global attractors. *Numerische Mathematik* **1997**, *75*, 293–317. [[CrossRef](#)]
8. Emmerich, M.; Deutz, A.H. Test problems based on Lamé superspheres. In Proceedings of the 4th International Conference on Evolutionary Multi-criterion Optimization EMO'07, Matsushima, Japan, 5–8 March 2007; Springer: Berlin/Heidelberg, Germany, 2007; pp. 922–936.
9. Dellnitz, M.; Schütze, O.; Hestermeyer, T. Covering Pareto sets by multilevel subdivision techniques. *J. Optim. Theory Appl.* **2005**, *124*, 113–155. [[CrossRef](#)]
10. Dilettoso, E.; Rizzo, S.A.; Salerno, N. A weakly Pareto compliant quality indicator. *Math. Comput. Appl.* **2017**, *22*, 25. [[CrossRef](#)]
11. Padberg, K. Numerical Analysis of Transport in Dynamical Systems. Ph.D. Thesis, University of Paderborn, Paderborn, Germany, 2005.
12. Peitz, S.; Dellnitz, M. A survey of recent trends in multiobjective optimal control—Surrogate models, feedback control and objective reduction. *Math. Comput. Appl.* **2018**, *23*, 30. [[CrossRef](#)]
13. Schütze, O. Set Oriented Methods for Global Optimization. Ph.D. Thesis, University of Paderborn, Paderborn, Germany, 2004.
14. Schütze, O.; Coello Coello, C.A.; Mostaghim, S.; Talbi, E.G.; Dellnitz, M. Hybridizing evolutionary strategies with continuation methods for solving multi-objective problems. *Eng. Optim.* **2008**, *40*, 383–402. [[CrossRef](#)]
15. Schütze, O.; Laumanns, M.; Coello Coello, C.A.; Dellnitz, M.; Talbi, E.G. Convergence of stochastic search algorithms to finite size Pareto set approximations. *J. Glob. Optim.* **2008**, *41*, 559–577. [[CrossRef](#)]
16. Schütze, O.; Esquivel, X.; Lara, A.; Coello Coello, C.A. Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization. *IEEE Trans. Evol. Comput.* **2012**, *16*, 504–522. [[CrossRef](#)]
17. Vargas, A.; Bogoya, J.M. A generalization of the averaged Hausdorff distance. *Computación y Sistemas* **2018**, *22*, 331–345. [[CrossRef](#)]
18. Bogoya, J.M.; Vargas, A.; Cuate, O.; Schütze, O. A (p, q) -averaged Hausdorff distance for arbitrary measurable sets. *Math. Comput. Appl.* **2018**, *23*, 51. [[CrossRef](#)]
19. Cai, X.; Li, Y.; Fan, Z.; Zhang, Q. An external archive guided multiobjective evolutionary algorithm based on decomposition for combinatorial optimization. *IEEE Trans. Evolut. Comput.* **2015**, *19*, 508–523.
20. Shang, R.; Wang, Y.; Wang, J.; Jiao, L.; Wang, S.; Qi, L. A multi-population cooperative coevolutionary algorithm for multi-objective capacitated arc routing problem. *Inf. Sci.* **2014**, *277*, 609–642. [[CrossRef](#)]
21. Zhang, J.; Tang, Q.; Li, P.; Deng, D.; Chen, Y. A modified MOEA/D approach to the solution of multi-objective optimal power flow problem. *Appl. Soft Comput.* **2016**, *47*, 494–514. [[CrossRef](#)]
22. Dhiman, G.; Kumar, V. Multi-objective spotted hyena optimizer: A Multi-objective optimization algorithm for engineering problems. *Knowl. Based Syst.* **2018**, *150*, 175–197. [[CrossRef](#)]
23. López-Rubio, F.J.; López-Rubio, E. Features for stochastic approximation based foreground detection. *Comput. Vision Image Underst.* **2015**, *133*, 30–50. [[CrossRef](#)]
24. Kerkhove, L.P.; Vanhoucke, M. Incentive contract design for projects: The owner's perspective. *Omega* **2016**, *62*, 93–114. [[CrossRef](#)]
25. Hansen, M.P.; Jaszkiwicz, A. *Evaluating the Quality of Approximations to the Non-Dominated Set*; IMM, Department of Mathematical Modelling, Technical University of Denmark: Kongens Lyngby, Denmark, 1998.
26. Zitzler, E.; Thiele, L. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Trans. Evol. Comput.* **1999**, *3*, 257–271. [[CrossRef](#)]
27. Siwel, J.; Yew-Soon, O.; Jie, Z.; Liang, F. Consistencies and contradictions of performance metrics in multiobjective optimization. *IEEE Trans. Evol. Comput.* **2014**, *44*, 2329–2404.
28. Vargas, A. On the Pareto compliance of the averaged Hausdorff distance as a performance indicator. *Universitas Scientiarum* **2018**, *23*, 333–354. [[CrossRef](#)]
29. Miettinen, K. *Nonlinear Multiobjective Optimization*; Kluwer Academic Publishers: Tranbjerg, Denmark, 1999.

30. Ehrgott, M.; Wiecek, M.M. Multiobjective programming. In *Multiple Criteria Decision Analysis: State of the Art Surveys*; Springer: New York, NY, USA, 2005; pp. 667–722.
31. Pareto, V. *Manual of Political Economy*; The Macmillan Press: London, UK, 1971.
32. Hillermeier, C. *Nonlinear Multiobjective Optimization: A Generalized Homotopy Approach*; Springer Science & Business Media: Berlin, Germany, 2001; Volume 135.
33. Köppen, M.; Yoshida, K. Many-objective particle swarm optimization by gradual leader selection. In Proceedings of the 8th international conference on adaptive and natural computing algorithms (ICANNGA 2007), Warsaw, Poland, 11–14 April 2007; Springer: Berlin/Heidelberg, Germany, 2007; pp. 323–331.
34. Schütze, O.; Lara, A.; Coello Coello, C.A. On the influence of the number of objectives on the hardness of a multiobjective optimization problem. *IEEE Trans. Evolut. Comput.* **2011**, *15*, 444–455. [[CrossRef](#)]
35. Schaffer, J.D. Multiple Objective Optimization with Vector Evaluated Genetic Algorithms. Ph.D. Thesis, Vanderbilt University, Nashville, TN, USA, 1984.
36. Amini, A.; Tavakkoli-Moghaddam, R. A bi-objective truck scheduling problem in a cross-docking center with probability of breakdown for trucks. *Comput. Ind. Eng.* **2016**, *96*, 180–191. [[CrossRef](#)]
37. Li, M.W.; Hong, W.C.; Geng, J.; Wang, J. Berth and quay crane coordinated scheduling using multi-objective chaos cloud particle swarm optimization algorithm. *Neural Comput. Appl.* **2017**, *28*, 3163–3182. [[CrossRef](#)]
38. Dulebenets, M.A. A comprehensive multi-objective optimization model for the vessel scheduling problem in liner shipping. *Int. J. Prod. Econ.* **2018**, *196*, 293–318. [[CrossRef](#)]
39. Goodarzi, A.H.; Nahavandi, N.; Hessameddin, S. A multi-objective imperialist competitive algorithm for vehicle routing problem in cross-docking networks with time windows. *J. Ind. Syst. Eng.* **2018**, *11*, 1–23.
40. Venturini, G.; Iris, C.; Kontovas, C.A.; Larsen, A. The multi-port berth allocation problem with speed optimization and emission considerations. *Transp. Res. Part D Transp. Environ.* **2017**, *54*, 142–159. [[CrossRef](#)]
41. Chargui, T.; Bekrar, A.; Reghioui, M.; Trentesaux, D. Multi-objective sustainable truck scheduling in a rail-road physical internet cross-docking hub considering energy consumption. *Sustainability* **2019**, *11*, 3127. [[CrossRef](#)]
42. Fliege, J.; Graña, L.M.; Svaiter, B.F. Newton’s method for multiobjective optimization. *SIAM J. Opt.* **2009**, *20*, 602–626. [[CrossRef](#)]
43. Das, I.; Dennis, J. Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. *SIAM J. Opt.* **1998**, *8*, 631–657. [[CrossRef](#)]
44. Eichfelder, G. *Adaptive Scalarization Methods in Multiobjective Optimization*; Springer: Berlin Heidelberg, Germany, 2008.
45. Fliege, J. Gap-free computation of Pareto-points by quadratic scalarizations. *Math. Methods Operat. Res.* **2004**, *59*, 69–89. [[CrossRef](#)]
46. Pereyra, V. Fast computation of equispaced Pareto manifolds and Pareto fronts for multiobjective optimization problems. *Math. Comput. Simul.* **2009**, *79*, 1935–1947. [[CrossRef](#)]
47. Wang, H. Zigzag search for continuous multiobjective optimization. *INFORMS J. Comp.* **2013**, *25*, 654–665. [[CrossRef](#)]
48. Martin, B.; Goldsztejn, A.; Granvilliers, L.; Jermann, C. Certified parallelotope continuation for one-manifolds. *SIAM J. Numer. Anal.* **2013**, *51*, 3373–3401. [[CrossRef](#)]
49. Pereyra, V.; Saunders, M.; Castillo, J. Equispaced Pareto front construction for constrained bi-objective optimization. *Math. Comput. Model.* **2013**, *57*, 2122–2131. [[CrossRef](#)]
50. Martin, B.; Goldsztejn, A.; Granvilliers, L.; Jermann, C. On continuation methods for non-linear bi-objective optimization: Towards a certified interval-based approach. *J. Glob. Optim.* **2014**, *64*, 1–14. [[CrossRef](#)]
51. Schütze, O.; Martín, A.; Lara, A.; Alvarado, S.; Salinas, E.; Coello Coello, C.A. The directed search method for multiobjective memetic algorithms. *J. Comput. Optim. Appl.* **2016**, *63*, 305–332. [[CrossRef](#)]
52. Martín, A.; Schütze, O. Pareto Tracer: A predictor-corrector method for multi-objective optimization problems. *Eng. Optim.* **2018**, *50*, 516–536. [[CrossRef](#)]
53. Jahn, J. Multiobjective search algorithm with subdivision technique. *Comput. Optim. Appl.* **2006**, *35*, 161–175. [[CrossRef](#)]
54. Sun, J.Q.; Xiong, F.R.; Schütze, O.; Hernández, C. *Cell Mapping Methods-Algorithmic Approaches and Applications*; Springer: Singapore, 2019.
55. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; John Wiley & Sons: Chichester, UK, 2001.

56. Coello Coello, C.A.; Lamont, G.B.; Van Veldhuizen, D.A. *Evolutionary Algorithms for Solving Multi-Objective Problems*, 2nd ed.; Springer: New York, NY, USA, 2007.
57. Sun, Y.; Gao, Y.; Shi, X. Chaotic multi-objective particle swarm optimization algorithm incorporating clone immunity. *Mathematics* **2019**, *7*, 146. [[CrossRef](#)]
58. Wang, P.; Xue, F.; Li, H.; Cui, Z.; Xie, L.; Chen, J. A multi-objective DV-hop localization algorithm based on NSGA-II in internet of things. *Mathematics* **2019**, *7*, 184. [[CrossRef](#)]
59. Pei, Y.; Yu, J.; Takagi, H. Search acceleration of evolutionary multi-objective optimization using an estimated convergence point. *Mathematics* **2019**, *7*, 129. [[CrossRef](#)]
60. Bullen, P.S. *Handbook of Means and Their Inequalities*; Vol. 560, Mathematics and its Applications; Kluwer Academic Publishers Group: Dordrecht, The Netherlands, 2003; p. xxviii+537.
61. Van Veldhuizen, D.A.; Lamont, G.B. Multiobjective evolutionary algorithm test suites. In Proceedings of the 1999 ACM symposium on Applied Computing, San Antonio, TX, USA, 28 February–2 March 1999; ACM: New York, NY, USA, 1999; pp. 351–357.
62. Coello Coello, C.A.; Cruz Cortés, N. Solving multiobjective optimization problems using an artificial immune system. *Genet. Program. Evol. Mach.* **2005**, *6*, 163–190. [[CrossRef](#)]
63. Rudolph, G.; Schütze, O.; Grimme, C.; Domínguez-Medina, C.; Trautmann, H. Optimal averaged Hausdorff archives for bi-objective problems: Theoretical and numerical results. *Comput. Optim. Appl.* **2016**, *64*, 589–618. [[CrossRef](#)]
64. Goldberg, M. Equivalence constants for ℓ_p norms of matrices. *Linear Multilinear Algebra* **1987**, *21*, 173–179. [[CrossRef](#)]
65. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [[CrossRef](#)]
66. Zhang, Q.; Li, H. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.* **2007**, *11*, 712–731. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).