

Article

# A Bounded Archiver for Hausdorff Approximations of the Pareto Front for Multi-Objective Evolutionary Algorithms

Carlos Ignacio Hernández Castellanos <sup>1,\*</sup>  and Oliver Schütze <sup>2,\*</sup> 

<sup>1</sup> Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas (IIMAS), Universidad Nacional Autónoma de México (UNAM), Mexico City 04510, Mexico

<sup>2</sup> Computer Science Department, CINVESTAV-IPN, Mexico City 07360, Mexico

\* Correspondence: carlos.hernandez@iimas.unam.mx (C.I.H.C.); schuetze@cs.cinvestav.mx (O.S.)

† These authors contributed equally to this work.

**Abstract:** Multi-objective evolutionary algorithms (MOEAs) have been successfully applied for the numerical treatment of multi-objective optimization problems (MOP) during the last three decades. One important task within MOEAs is the archiving (or selection) of the computed candidate solutions, since one can expect that an MOP has infinitely many solutions. We present and analyze in this work ArchiveUpdateHD, which is a bounded archiver that aims for Hausdorff approximations of the Pareto front. We show that the sequence of archives generated by ArchiveUpdateHD yields under certain (mild) assumptions with a probability of one after finitely many steps a  $\Delta^+$ -approximation of the Pareto front, where the value  $\Delta^+$  is computed by the archiver within the run of the algorithm without any prior knowledge of the Pareto front. The knowledge of this value is of great importance for the decision maker, since it is a measure for the “completeness” of the Pareto front approximation. Numerical results on several well-known academic test problems as well as the usage of ArchiveUpdateHD as an external archiver within three state-of-the-art MOEAs indicate the benefit of the novel strategy.



**Citation:** Hernández Castellanos, C.I.; Schütze, O. A Bounded Archiver for Hausdorff Approximations of the Pareto Front for Multi-Objective Evolutionary Algorithms. *Math. Comput. Appl.* **2022**, *27*, 48. <https://doi.org/10.3390/mca27030048>

Academic Editors: Sebastian Peitz and Gianluigi Rozza

Received: 30 April 2022

Accepted: 28 May 2022

Published: 1 June 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** evolutionary multi-objective optimization; archiving; convergence

## 1. Introduction

This work is dedicated to the 60th birthday of Professor Kalyanmoy Deb, a pioneer and highly impactful and influential proponent of Evolutionary Multi-Objective Optimization (EMO) for the last three decades. In particular, the seminal work *Combining Convergence and Diversity in Evolutionary Multiobjective Optimization* by Marco Laumanns, Lothar Thiele, Kalyanmoy Deb, and Eckart Zitzler [1] has been a motivation of the second author to consider the challenging and fruitful field of archiving in EMO.

Multi-objective optimization problems (MOPs), i.e., problems where several conflicting objectives have to be optimized concurrently, naturally arise in many real-world applications (e.g., [2–8]). While one can expect *one* optimal solution if “only” one objective is being considered, the solution set of an MOP (the so-called Pareto set, respectively, its image, the Pareto front) typically forms at least locally a manifold of a certain dimension [9]. One important task in multi-objective optimization (MOO) is hence to identify a “suitable” finite size approximation of these solution sets. Multi-objective evolutionary algorithms (MOEAs) represent an important class of algorithms for the numerical treatment of such problems. MOEAs have caught the interest of researchers and practitioners due to their global nature and robustness and since they require only minimal assumptions on the model (e.g., [2,10]). The process to elect a subset of the candidate solutions generated by the MOEA is called *selection* or *archiving*. Existing archiving/selection strategies can be roughly divided into two classes (see subsequent section for more details): (i) mechanisms that maintain sets whose cardinalities are equal or do not exceed a certain pre-defined cardinality—which we will call bounded archivers in the following—and (ii) archivers that

are based on the concept of  $\epsilon$ -dominance. Such archivers generate sequences of archives  $A_i$  with monotonic behavior, i.e., no deterioration of cyclic behavior can be observed during the run of an algorithm. Furthermore, for  $i \rightarrow \infty$ , these archives yield certain limit approximation qualities that can be adjusted a priori, mainly by choosing the values of  $\epsilon \in \mathbb{R}^k$  which comes rather naturally at least if the MOP arises from a real-world application. On the other hand, the magnitudes of the final archives are entirely determined by  $\epsilon$  and some other design parameters, which are set a priori and are supposed to remain fixed during the computation, and the size of the Pareto front, which is a priori of course unknown. It has turned out that most EMO researchers prefer to have a fixed number of elements in the archives, e.g., for the sake of a better comparison to other methods but also to avoid the necessity of storing an unexpected large amount of candidate solutions. The latter problem is apparently by construction not given by bounded archives. For most strategies from class (i), however, no theoretical analyses such as convergence properties are known. For many distance-based methods, it is further known that cyclic behavior and deterioration can occur during the run of the algorithm. It is hence fair to say that these methods do not tap the full potential, since any MOEA using such a strategy will not converge regardless of the regions they explore during the run of the algorithm. An exception is the bounded archive proposed in [11], which yields under certain (mild) assumptions and with the probability of one  $\epsilon$ -Pareto set in the limit, where  $\epsilon \in \mathbb{R}$  is the smallest possible value with respect to to the bound of the archives.

In this paper, we propose a bounded archiver that is based on distance, dominance and  $\epsilon$ -dominance that offers quasi-monotonic behavior and yields approximation qualities in the limit. More precisely, ArchiveUpdateHD aims for Hausdorff approximations of the Pareto front (i.e., evenly spread solutions along the Pareto front). Under certain (mild) assumptions on the generation process, it will be shown that the Hausdorff distances of the images of the archives  $F(A_i)$  and the Pareto front  $F(P_Q)$  are bounded by a value  $\Delta^+$ , which is computed by the archiver during the run of the algorithm. Numerical experiments show that this value indeed represents a good approximation of the actual Hausdorff distance (while a better strategy is proposed for bi-objective problems). During the run of the algorithm, two design parameters are adjusted adaptively during the run of the algorithm (one being the value of  $\epsilon$  for the  $\epsilon$ -dominance). Since these values will become stationary during the search process, one can expect monotonic behavior from a certain stage of the search process. The knowledge of the Hausdorff distance of  $F(A_i)$  and  $F(P_Q)$  is important information for the decision maker (DM), since it represents the maximal error in the approximation. If not needed (i.e., depending on the chosen initial design parameters), the magnitudes of the archivers will not reach the pre-defined size  $N$ . Else, the value  $\Delta^+$  computed by the archiver is an important piece of information, since it tells the DM if the approximation is “complete enough” or not. In the latter case, the computation may have to be repeated using an increased value of  $N$ . A preliminary version of this work can be found in [12], which is restricted to bi-objective problems and contains fewer empirical results.

The rest of this document is structured as follows: Section 2 briefly summarizes the background that is required for the understanding of the sequel and presents the related work. In Section 3, the new archiver ArchiveUpdateHD is discussed and analyzed, first for bi-objective problems and after that for the general number of objectives. In Section 4, some numerical results and comparisons are presented and discussed. Finally, in Section 5, conclusions are drawn, and possible paths for future research are mentioned.

## 2. Background and Related Work

Here, we consider continuous multi-objective optimization problems (MOPs) that can be expressed as follows:

$$\min_{x \in Q} F(x). \quad (\text{MOP})$$

The map  $F$  is defined by the individual objective functions  $f_i$ , i.e.,

$$F : Q \rightarrow \mathbb{R}^k, \quad F(x) = (f_1(x), \dots, f_k(x))^T, \tag{1}$$

where each  $f_i : Q \rightarrow \mathbb{R}, i = 1, \dots, k$ , is assumed to be continuous. We stress, however, that the archiver presented below can also be applied to discrete problems.  $Q$  is the domain or feasible set of the problem, which is typically expressed by equality and inequality constraints. We assume  $Q$  to be compact (i.e., closed and bounded). If  $k = 2$  objectives are considered, the problem is also termed a bi-objective problem (BOP).

In order to define optimality in multi-objective optimization, the concept of dominance can be used: for two vectors  $x, y \in \mathbb{R}^k$  we say that  $x$  is less than  $y$  ( $x <_p y$ ) if  $x_i < y_i, i \in \{1, \dots, k\}$ , analogously for the relation  $\leq_p$ . We say that  $y \in Q$  is *dominated* by  $x \in Q$  ( $x \prec y$ ) with respect to (MOP) if  $F(x) \leq_p F(y)$  and  $F(x) \neq F(y)$ , else we say that  $y$  is *non-dominated* by  $x$ . Finally,  $x^* \in Q$  is called Pareto optimal or simply optimal with respect to (MOP) if there exists no  $y \in Q$  that dominates  $x$ . The Pareto set  $P_Q$  is the set of all optimal solutions with respect to (MOP), and its image  $F(P_Q)$  is called the Pareto front. One can expect that both the Pareto set and Pareto front are from at least locally objects of dimension  $k - 1$ , where  $k$  is the number of objectives considered in the MOP [9].

For the convergence analysis, we will consider a very general class of algorithms, called Generic Stochastic Search Algorithm (GSSA), first considered by Laumanns et al. [1]. An algorithm of this class consists of a process to generate new candidate solutions together with an update strategy. Algorithm 1 shows the pseudocode of GSSA.

---

**Algorithm 1** Generic Stochastic Search Algorithm

---

- 1:  $P_0 \subset Q$  drawn at random
  - 2:  $A_0 = \text{ArchiveUpdate}(P_0, \emptyset)$
  - 3: **for**  $j = 0, 1, 2, \dots$  **do**
  - 4:      $P_{j+1} = \text{Generate}(P_j)$
  - 5:      $A_{j+1} = \text{ArchiveUpdate}(P_{j+1}, A_j)$
  - 6: **end for**
- 

In the following, we define the Hausdorff distance  $d_H$  and the averaged Hausdorff distance  $\Delta_p$ , which we will use to assess the approximation qualities of the obtained Pareto front approximations (toward the actual Pareto fronts).

**Definition 1.** Let  $u \in \mathbb{R}^n$  and  $A, B \subset \mathbb{R}^n$ . The semi-distance  $\text{dist}(\cdot, \cdot)$  and the Hausdorff distance  $d_H(\cdot, \cdot)$  are defined as follows:

- (a)  $\text{dist}(u, A) := \inf_{v \in A} \|u - v\|_\infty$
- (b)  $\text{dist}(B, A) := \sup_{u \in B} \text{dist}(u, A)$
- (c)  $d_H(A, B) := \max\{\text{dist}(A, B), \text{dist}(B, A)\}$

**Definition 2 ([13]).** Let  $A, B \subset \mathbb{R}^n$  be finite sets. The value

$$\Delta_p(A, B) = \max(\text{GD}_p(A, B), \text{IGD}_p(A, B)), \tag{2}$$

where

$$\begin{aligned} \text{GD}_p(A, B) &= \left( \frac{1}{|A|} \sum_{a \in A} \text{dist}(a, B)^p \right)^{1/p} \\ \text{IGD}_p(A, B) &= \left( \frac{1}{|B|} \sum_{b \in B} \text{dist}(b, A)^p \right)^{1/p}, \end{aligned} \tag{3}$$

and  $p \in \mathbb{N}$ , is called the averaged Hausdorff distance between  $A$  and  $B$ .

We further define some objects that specify certain approximation qualities of Pareto front approximations. All of these objects are based on the concept of  $\epsilon$ -dominance, which we will define first.

**Definition 3** ( $\epsilon$ -dominance). Let  $\epsilon = (\epsilon_1, \dots, \epsilon_k)^T \in \mathbb{R}_+^k$  and  $x, y \in \mathbb{R}^n$ .  $x$  is said to  $\epsilon$ -dominate  $y$  (in short:  $x \prec_\epsilon y$ ) with respect to (MOP) if

$$F(x) - \epsilon \leq_p F(y) \quad \text{and} \quad F(x) - \epsilon \neq F(y). \tag{4}$$

**Definition 4** ( $\epsilon$ -(approximate) Pareto front, [1]). Let  $\epsilon \in \mathbb{R}_+^k$  and  $A \subset \mathbb{R}^n$ .

(a)  $F(A)$  is called an  $\epsilon$ -approximate Pareto front of (MOP) if every point  $x \in Q$  is  $\epsilon$ -dominated by at least one  $a \in A$ , i.e.,

$$\forall x \in Q : \exists a \in A : a \prec_\epsilon x. \tag{5}$$

(b)  $F(A)$  is called an  $\epsilon$ -Pareto front if  $F(A)$  is an  $\epsilon$ -approximate Pareto front and if every point  $a \in A$  is a Pareto point of (MOP).

**Definition 5** ( $\Delta$ -tight  $\epsilon$ -(approximate) Pareto front, [14]). Let  $\epsilon \in \mathbb{R}_+^k$  and  $A \subset \mathbb{R}^n$ .

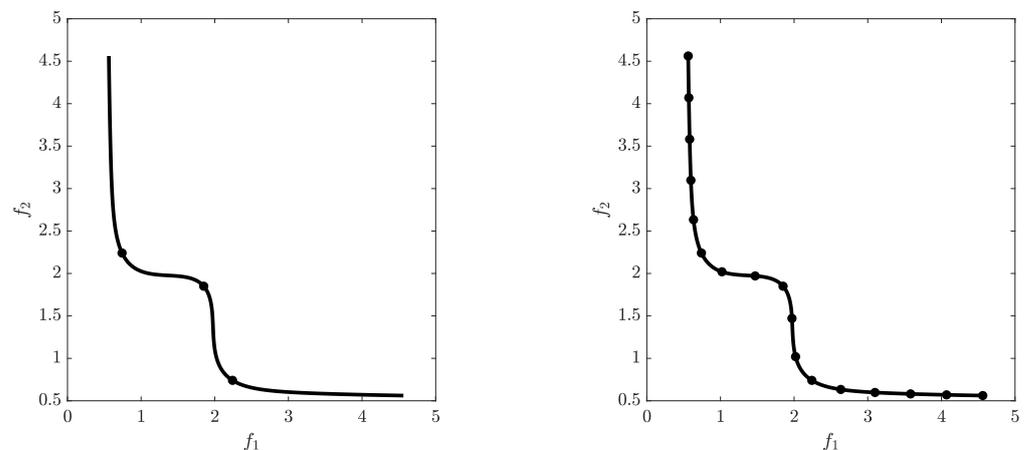
(a)  $F(A)$  is called a  $\Delta$ -tight  $\epsilon$ -approximate Pareto front of (MOP) if  $A$  is an  $\epsilon$ -approximate Pareto front of (MOP) and if in addition

$$\text{dist}(F(P_Q), F(A)) \leq \Delta. \tag{6}$$

(b)  $F(A)$  is called a  $\Delta$ -tight  $\epsilon$ -Pareto front if  $A$  is an  $\epsilon$ -Pareto front of (MOP) and if in addition

$$d_H(F(P_Q), F(A)) \leq \Delta. \tag{7}$$

The archiver we propose in this work, ArchiveUpdateHD, aims for  $\Delta$ -tight  $\epsilon$ -Pareto fronts for particular values of  $\Delta$  and  $\epsilon$ . The sole usage of  $\epsilon$ -dominance for the Pareto front approximations may lead to gaps in particular when parts of the front are flat. The  $\Delta$ -tight  $\epsilon$ -(approximate) Pareto fronts also take into account the distance of the Pareto front toward the candidate set, leading to better approximations in the Hausdorff sense (see Figure 1).



**Figure 1.** Gaps in the approximation can occur when  $\epsilon$ -dominance is used exclusively in the selection/archiving of the candidate solutions (left).  $\Delta$ -tight  $\epsilon$ -(approximate) Pareto fronts also consider the distance of the Pareto front toward the archive (right).

Since one can expect infinitely many solutions for a continuous MOP, it is inevitable in (continuous) evolutionary multi-objective optimization (EMO) that not all promising

solutions can be kept during the run of an algorithm. Instead, one has to elect a subset of candidate solutions in each iteration so that this sequence eventually leads to a “suitable” representation of the Pareto set/front of the given problem. This process is typically called “selection” within MOEAs and “archiving” if an external set of candidate solutions (archive) is maintained during the run of a MOEA (though of course both terms can be used interchangeably).

Three main classes of MOEAs exist: (a) dominance-based [15–18], (b) decomposition-based [19–25], and (c) indicator-based [26–30] algorithms. The selection strategies for MOEAs of class (b) or (c) are rather straightforward: the selection in a decomposition-based MOEA is done implicitly by the chosen scalarizing functions, and the selection in an indicator-based MOEA is typically handled via considering the indicator contributions. These two approaches come on the one hand with a monotonic behavior of the sequence of approximations (i.e., no deterioration can occur). On the other hand, these selection strategies do not guarantee convergence toward the best approximation (e.g., [29,31]). The selection of the first dominance-based MOEAs is based on non-dominated sorting in combination with niching techniques (e.g., [32–34]). Due to missing elite preservation, none of these methods converge in the mathematical sense. Later MOEAs such as SPEA [35], PAES [18], SPEA-II [16], and NSGA-II [15] include such elite preservation leading to much better overall performance. However, also for these algorithms, no convergence properties (again, in the mathematical sense) are known. Rudolph [36–39] and Hanne [40–43] have studied convergence properties of MOEA frameworks. These studies are mainly concerned with the convergence of individuals of the populations toward the Pareto set/front, while the magnitudes and the distributions of the resulting populations are not considered.

Archiving strategies with bounded archive size based on adaptive grid selection have been considered in [44–46]. Bounded archivers in particular for hypervolume approximations have been proposed in [47,48]. Both archivers yield monotonic behavior in the approximation qualities of the obtained sequence of archives.

Laumanns et al. considered the class of algorithms GSSA as described above [1,49] which allows to focus on the archiver under certain (mild) assumptions on the generator. In both studies, archivers were considered, aiming for several  $\epsilon$  approximations of the Pareto front, where finitely many iterations were considered. Later, further archivers have been proposed based on  $\epsilon$ -dominance using the framework of GSSA to perform convergence analysis [14,50–54].

In [11], Laumanns and Zenklusen propose two bounded archivers that use adaptive schemes to obtain  $\epsilon$  approximations of the Pareto front. Another adaptive archiving strategy is proposed in [55] that utilizes a particular discretization of the objective space of the given problem. A strategy that is based on the convex hull of individual minima in order to increase diversity of the solutions is proposed in [56].

Recently, the use of external archives has become more popular [57–62] in particular for the treatment of real-world applications where function evaluations are expensive, and where it is hence advisable to maintain all promising candidate solutions. Consequently, most of these archivers are unbounded [63–67]. For the treatment of in particular MOPs with many objectives—also called many objective problems—MOEAs have been proposed that utilize *two* archives, one aiming for convergence and one aiming for diversity [68,69].

### 3. ArchiveUpdateHD

We will in this section propose and discuss the novel archiver ArchiveUpdateHD. Since the considerations of the distances as well as the Hausdorff approximations can be done more accurately for  $k = 2$ —where we can assume the Pareto front to locally form a curve, and hence, the elements of the approximations can be arranged via a sorting in objective space—we first address the bi-objective case and will afterwards consider the archiver for problems with  $k > 2$ .

### 3.1. The Bi-Objective Case

The pseudocode of ArchiveUpdateHD for bi-objective problems is shown in Algorithm 2. This archiver aims for approximations of the Pareto front of a given BOP in the Hausdorff sense (i.e., for solutions that are evenly spread along the Pareto front). The archiver is based (i) on the distances among the candidate solutions (lines 18–36 of Algorithm 2), (ii) “classical” dominance or elite preservation (lines 5 and 9) as well as (iii) the concept of  $\epsilon$ -dominance (line 5). The archiver can roughly be divided into two parts: an acceptance strategy to decide if an incoming candidate solution  $p$  should be considered (line 5), and a pruning technique (mainly lines 18–36, but also lines 11–14) which is applied if the size of the archive has exceeded a predefined budget  $N$  of archive entries.

In the following, we will describe ArchiveUpdateHD as in Algorithm 2 in more detail. This algorithm contains several elements that have to be incorporated in order to guarantee convergence. After the convergence analysis (Theorem 1), we will discuss more practical realizations of the algorithm.

In line 5, it is decided if a candidate solution  $p$  should be (at least temporarily) added to the existing archive  $A$ . This is the case if (a) none of the entries  $a \in A$   $\Theta\epsilon$ -dominates  $p$  ( $\Theta \in (0, 1)$  being a safety factor needed to guarantee convergence, see below for practical realizations), or if (b) none of the entries  $a \in A$  dominates  $p$  and for none of the entries  $a \in A$  the distance  $\|F(a) - F(p)\|$  is less or equal than  $\Theta\Delta$ . Throughout this work,  $\|\cdot\|$  denotes the Euclidean norm. We stress that this acceptance strategy is identical to the one of the archiver ArchiveUpdateTight2 [14], which we will need for the upcoming convergence analysis.

If the candidate solution  $p$  is accepted, it will be added to  $A$ . Next, all other entries  $a \in A$  dominated by  $p$  will be discarded (lines 8–10). Hence, all archives generated by ArchiveUpdateHD only contain mutually non-dominated elements (elite preservation). If the distance  $\|F(p) - F(a)\|$  is larger than  $\epsilon$  for any of these dominated archive entries  $a$ , a “reset” is executed for  $\Delta$  and  $\epsilon$ :  $\Delta_{min}$  is set to  $\kappa\Delta_{min}$  (where  $\kappa > 1$  another safety factor). Next,  $\Delta$  and  $\epsilon$  are updated using this new minimal value. The idea behind this reset is as follows: if  $p \prec a$  and the distance of  $F(a)$  and  $F(p)$  is larger than  $\Delta$ , then  $p$  and  $a$  could be located in different connected components of the set of (local) solutions of the bi-objective problem. Since the values both of  $\Delta$  and  $\epsilon$  are determined by the length of the (known) Pareto front, their values have to be set back, since a “jump” to a new connected component may lead to a new length. See Figure 2 for a hypothetical scenario. The value of  $\Delta_{min}$  has to be (slightly) increased in each reset in order to avoid the possible of a cyclic behavior in the sequence of archives (which, in fact, has not been observed in our computations).

If  $|A|$  exceeds the predefined magnitude  $N$ , it is decided in lines 18–36 which of the elements of  $A$  has to be discarded (pruning). For  $k = 2$  objectives, we can order all the entries of the archives (e.g., as done here: in ascending order wrt objective  $f_1$ ). Then, the vector  $d \in \mathbb{R}^N$  of distances can be simply computed via:

$$d_i := \|F(a_{i+1}) - F(a_i)\|, \quad i = 1, \dots, N. \quad (8)$$

For an index  $m$  chosen from  $\arg \min d$ , either  $a_m$  or  $a_{m+1}$  is then removed from  $A$ , which is done in lines 23–33. The aim of ArchiveUpdateHD is to maintain good approximations of the end points of the Pareto front. Accordingly,  $a_2$ , respectively,  $a_N$ , are always discarded instead of  $a_1$  and  $a_{N+1}$ , respectively (lines 23–26). The rationale behind the selection in lines 28–33 is to keep the archive of size  $N$  with the most evenly distributed elements.

**Algorithm 2** ArchiveUpdateHD

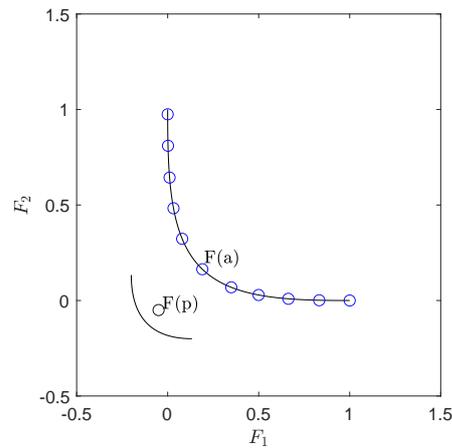
**Require:** Problem (MOP), where  $k = 2$ ,  $P$ : current population,  $A_0$ : current archive,  $\Delta_0 > 0$ : current value of  $\Delta$ ,  $\Delta_{min}$ : minimal value of  $\Delta$ ,  $\Theta \in (0, 1)$ ,  $\kappa > 1$ : safety factors,  $N$ : upper bound for archive size

**Ensure:** updated archive  $A$ , updated values for  $\Delta$ ,  $\Delta_{min}$ , and  $\epsilon$

```

1:  $A := A_0$ 
2:  $\Delta := \Delta_0$ 
3:  $\epsilon := (\Delta, \dots, \Delta)^T$ 
4: for all  $p \in P$  do
5:   if  $\nexists a \in A : a \prec_{\Theta\epsilon} p$ , or  $\exists a \in A : a \prec p$  and  $\forall a \in A : \|F(a) - F(p)\| > \Theta\Delta$  then
6:      $A := A \cup \{p\}$ 
7:   end if
8:   for all  $a \in A$  do
9:     if  $p \prec a$  then
10:       $A := A \cup \{p\} \setminus \{a\}$ 
11:      if  $\|F(p) - F(a)\|_\infty > \Delta$  then ▷ reset  $\Delta$  and  $\epsilon$ 
12:         $\Delta_{min} := \kappa\Delta_{min}$ 
13:         $\Delta := \Delta_{min}$ 
14:         $\epsilon := (\Delta, \dots, \Delta)^T$ 
15:      end if
16:    end if
17:  end for
18: if  $|A| = N + 1$  then ▷ apply pruning
19:    $\Delta := \frac{N+1}{N}\Delta$ 
20:    $\epsilon := \frac{N+1}{N}\epsilon$ 
21:   sort  $A$  (e.g., according to  $f_1$ )
22:   compute  $d \in \mathbb{R}^N$  as in (8)
23:   choose  $m \in \arg \min d$ 
24:   if  $m = 1$  then
25:      $A := A \setminus \{a_2\}$  ▷ remove 2nd entry
26:   else if  $m = N$  then
27:      $A := A \setminus \{a_N\}$  ▷ remove 2nd but last entry
28:   else
29:      $dl := \|F(a_{m+1}) - F(a_{m-1})\|$ 
30:      $dr := \|F(a_{m+2}) - F(a_m)\|$ 
31:     if  $dl < dr$  then
32:        $A := A \setminus \{a_m\}$ 
33:     else
34:        $A := A \setminus \{a_{m+1}\}$ 
35:     end if
36:   end if
37: end if
38: end for
39: return  $\{A, \epsilon, \epsilon_{min}, \Delta\}$ 

```



**Figure 2.** A hypothetical scenario that can happen for multi-modal problems: first, a front that is only locally optimal is detected by the search process and approximated by the archiver. If later, a candidate  $p$  is computed such that  $F(p)$  lies on a “better” front, the current values of  $\Delta$  and  $\epsilon$  may not be adequate any more to suitably approximate this front.

In the following, we investigate the limit behavior of ArchiveUpdateHD.

**Theorem 1.** Let (MOP) be given and  $Q \subset \mathbb{R}^n$  be compact, and let there be no weak Pareto points in  $Q \setminus P_Q$ . Furthermore, let  $F$  be continuous and injective, and

$$\forall x \in Q \text{ and } \forall \delta > 0 : P(\exists l \in \mathbb{N} : P_l \cap B_\delta(x) \cap Q \neq \emptyset) = 1. \tag{9}$$

Then, an application of Algorithm 1, where ArchiveUpdateHD (Algorithm 2) is used to update the archive, leads to a sequence of archives  $A_l, l \in \mathbb{N}$ , where the following holds:

(a) There exists a  $l_1 \in \mathbb{N}$  and  $\Delta^+ > 0$  such that

$$\Delta_l = \Delta^+, \quad \forall l \geq l_1, \quad \text{with probability one.}$$

(b) There exists with probability one a  $l_2 \in \mathbb{N}$  such that  $A_l$  is a  $\Delta^+$ -tight  $\epsilon$ -approximate Pareto front with respect to (MOP) for all  $l \geq l_2$ , where  $\epsilon = (\Delta^+, \dots, \Delta^+)^T$ .

(c)

$$\lim_{l \rightarrow \infty} \text{dist}(A_l, P_Q) = 0, \quad \text{with probability one.}$$

(d) There exists a  $l_3 \in \mathbb{N}$  such that

$$d_H(F(A_l), F(P_Q)) \leq \Delta^+, \quad \forall l \geq l_3, \quad \text{with probability one.}$$

**Proof.** We first show that during the run of the algorithm, only finitely many changes of the value of  $\Delta$  (and hence also of  $\epsilon$ ) can occur. Since  $F$  is continuous and the domain  $Q \subset \mathbb{R}^n$  is compact, also the image  $F(Q)$  is compact, and hence, in particular bounded. ArchiveUpdateHD changes the value of  $\Delta$  in two cases: if (i) a reset of  $\Delta$  and  $\epsilon$  is executed (line 12) or if (ii) the pruning technique is applied (line 19). In case of (i), the value of  $\Delta_{min}$  is increased by a constant factor  $\kappa > 1$ . The value of  $\Delta$  after the  $i$ -th reset is hence equal to or larger than  $\kappa^i \Delta_{min}^0$ , where  $\Delta_{min}^0 > 0$  denotes the value of  $\Delta_{min}$  at the start of the algorithm. A reset is applied if the distance of the image of the candidate solution  $p$  to the image of an archive element  $a$  is larger than the current value of  $\Delta$  (line 11). Since  $F(P_Q)$  is bounded, only a finite number of such resets can be applied during the run of the algorithm.

Case (ii) happens if the magnitude of the current archive is  $N + 1$ . New candidate solutions  $p$  are added to the archive in lines 5 and 6 and lines 9 and 10. Lines 9 and 10 describe a dominance replacement which does not increase the magnitude of the archive.

Hence, such replacements do not lead to an application of the pruning. A candidate  $p$  can be further added to the current archive  $A$  if one of the following statements is true (line 5):

$$\begin{aligned} \mathcal{E}_1 \quad & \exists a \in A : a \prec_{\Theta\epsilon} p, \text{ or} \\ \mathcal{E}_2 \quad & \exists a \in A : a \prec p \text{ and } \forall a \in A : \|F(a) - F(p)\| > \Theta\Delta \end{aligned} \tag{10}$$

Since  $F(Q)$  is bounded, there exists for every  $a \in Q$  a (large enough)  $\Delta_a > 0$  so that  $a \prec_{\Theta\epsilon} p$ , where  $\epsilon = (\Delta_a, \dots, \Delta_a)^T$ . Similarly,  $\|F(a) - F(p)\| < \Theta\Delta_a$  if  $\Delta_a$  is large enough. Since in each pruning step, the value of  $\Delta$  is increased by the factor of  $(N + 1)/N$  and since only finitely many resets are executed, also only finitely many prunings can be applied during the run of the algorithm.

Note that ArchiveUpdateHD differs from ArchiveUpdateTight2 in two parts: the reset strategy (lines 11–15) and the pruning technique (lines 18–37), and that both these parts come with a change of the values of  $\Delta$  and  $\epsilon$ . In other words, ArchiveUpdateHD is identical to ArchiveUpdateTight2 as long as no change in  $\Delta$  and  $\epsilon$  occurs. For this case, we can hence apply the theoretical results on ArchiveUpdateTight2 for ArchiveUpdateHD. Now, consider a fixed value of  $\Delta$  (and hence also  $\epsilon$ ). During the run, it can either be the case that (i) all magnitudes of  $A_l$  are less than or equal to  $N$  (i.e., no pruning is applied), or that (ii) this magnitude is  $N + 1$  at one point, leading to an application of the pruning technique. In case (i), we can use Theorem 7.4 of [54] on ArchiveUpdateTight2: there exists with probability of one a  $\bar{l} \in \mathbb{N}$  such that the sets  $F(A_l)$  form a  $\Delta$ -tight  $\epsilon$ -approximate Pareto front for all  $l \geq \bar{l}$ . Note that once  $F(A_l)$  forms such an object, no more resets can occur: assume there exists a candidate solution  $p$  that dominates an element  $a \in A_l$ , and where  $\|F(p) - F(a)\|_\infty > \Delta$ . The latter means that

$$\max_{i=1, \dots, k} f_i(a) - f_i(p) > \Delta \tag{11}$$

which in turn means that  $a$  does not  $\epsilon$ -approximate  $p$ , which is a contradiction to the assumption on  $A_l$ . In case of (ii), the value of  $\Delta$  is simply not large enough for the  $N$ -element archive to form a  $\Delta$ -tight  $\epsilon$ -approximate Pareto front. Again, by Theorem 7.4 of [54], there exists in this case with a probability of one a finite iteration number where the magnitude will exceed  $N$ . As discussed above, the pruning can only be applied finitely many times during the run of the algorithm. Hence, the value of  $\Delta$  will, with a probability of one, stay fixed from one iteration onwards, which proves part (a).

Parts (b) and (c) follow from Theorem 7.4 of [54] and part (a), and finally, part (d) follows from parts (b) and (c) and the definition of the Hausdorff distance.  $\square$

**Remark 1.** (a) Equation (9) is an assumption that has to be made on the generation process. It means that every neighborhood of every feasible point  $x \in Q$  will be “visited” with probability one by `Generate()` after finitely many steps. For MOEAs, this, e.g., ensured if Polynomial Mutation [70,71] is used or another mutation operator for which the support of the probability density functions equal to  $Q$  (at least for box-constrained problems). We hence think that this assumption is rather mild.

- (b) The complexity of the consideration of one candidate solution  $p$  is  $O(N \log(N))$ , which is determined by the sorting of the current archive  $A$  in line 20.
- (c)  $\Theta \in (0, 1)$  and  $\kappa > 1$  are safety factors needed to guarantee the convergence properties. In our computations, however, we have not observed any impact of these values if both are chosen near to one. We hence suggest to use  $\Theta = \kappa = 1$  (i.e., practically not to use these safety factors).
- (d) The above consideration is done for  $\epsilon = (\Delta, \dots, \Delta)^T \in \mathbb{R}^k$ , i.e., using the same value for all entries of  $\epsilon$ . If the values for the objectives along the Pareto front differ significantly, one

can of course instead use  $\Delta = \epsilon = (\Delta_1, \dots, \Delta_k)^T$  using different values  $\Delta_i$ . In that case, the following modifications have to be done: (i) the last condition in line 5 has to be replaced by

$$\nexists a \in A : |f_i(a) - f_i(p)| \leq \Delta_i, \quad i = 1, \dots, k.$$

Furthermore, (ii) the condition for the reset in line 11 has to be replaced by

$$\exists i \in \{1, \dots, k\} : |f_i(p) - f_i(a)| > \Delta_i.$$

- (e) The value of  $\Delta$  computed throughout the algorithm yields an approximation quality of the archivers in the Hausdorff sense. The theoretical upper bound of the final value  $\Delta^+$  is twice the value of the actual Hausdorff approximation as the following discussion shows (refer to Figure 3): assume we are given a linear front with slope  $-1$ , and we are given a budget of  $N = 2$  elements (the discussion is analog for general  $N$ ). The ideal archive as computed by ArchiveUpdateHD is in this case  $A = \{a_1, a_2\}$ , where the  $a_i$ 's are the end points of the Pareto set. Assume we have  $F(a_1) = (0, 1)^T$  and  $F(a_2) = (1, 0)^T$ ; then, the Hausdorff distance of the Pareto front and  $A$  is  $1/2$  determined by the point  $y_m = (1/2, 1/2)^T$ . Given this archive, for any value  $\Delta < 1$  and assuming that  $F(Q)$  is large enough, there exists a candidate  $p$  such that  $p$  is not dominated by  $a_1$  or  $a_2$  and that  $\|F(a_i) - F(p)\| > \Delta, i = 1, 2$ . Hence,  $p$  will be added to the archiver—and later on discarded (lines 23–26). The latter leads to an increase of  $\Delta$ .

On the one hand, one suggesting strategy would be to take  $\frac{1}{2}\Delta^+$  as a Hausdorff approximation of the Pareto front in particular, since most Pareto fronts have at least one element where the slope of the tangent space is  $-1$ . On the other hand, the use of  $\epsilon$ -dominance prevents that the images  $F(a), a \in A$ , are perfectly evenly distributed along the Pareto front so that  $\frac{1}{2}\Delta^+$  is not that accurate for some problems. In fact, this factor of two can only be observed for linear fronts, while  $\Delta^+$  already yields a good approximation in general (see, e.g., the subsequent results for MOPs with more than two objectives). However, we have observed that the following estimation gives even better approximations of the Hausdorff distances: given  $A = \{a_1, \dots, a_N\}$ , which is sorted (e.g., according to objective  $f_1$ ), the current Hausdorff approximation  $h$  is computed as follows:

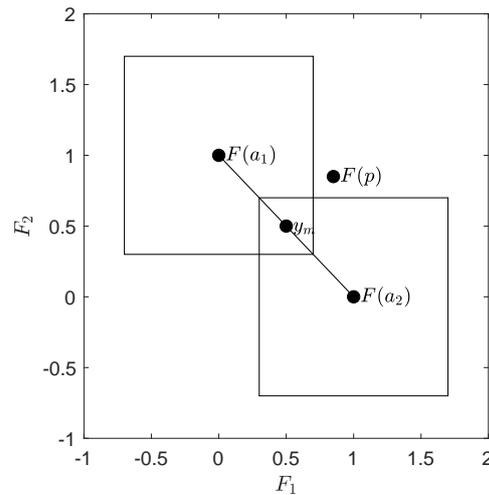
$$\tilde{d}_i := \begin{cases} \|F(a_{i+1}) - F(a_i)\|, & \text{if } \|F(a_{i+1}) - F(a_i)\| \leq 2\Delta \\ 0, & \text{else} \end{cases}, \quad i = 1, \dots, N - 1 \tag{12}$$

$$h := \frac{1}{2} \max_{i=1, \dots, N-1} \tilde{d}_i.$$

Note that the distance is set to 0 if the distance between two neighboring candidate solutions is larger or equal to  $2\Delta$ , which has been done to take into account approximations of Pareto fronts that fall into several connected components.

- (f) Several norms are used within the algorithm. While one is—except in line 11, see the above proof—in principle free for the choice of the norms, we suggest taking the infinity norm in line 5 in order to reduce the issue mentioned in the previous part, and the 2 norm in lines 28 and 29 in order to obtain a (slightly) better distribution of the entries along the Pareto front.

Algorithm 3 shows the modifications of ArchiveUpdateHD discussed above, which have been used for the calculations presented in this work. Hereby,  $\Delta_{min} \in \mathbb{R}_+^k$  denotes the vector of minimal elements for each entry  $\Delta_i$ .



**Figure 3.** Linear Pareto front with slope  $-1$ . If for  $N = 2$ , the archive is given by  $A = \{a_1, a_2\}$  such that  $F(a_1)$  and  $F(a_2)$  are the end points of the Pareto front, then the Hausdorff distance of  $A$  and the Pareto front is given by  $h = \|F(a_1) - y_m\|_\infty$ , where  $y_m$  is the arithmetic mean of  $F(a_1)$  and  $F(a_2)$ . For  $\Delta < 2h$ , there may exist candidate solutions  $p$  that will be considered by the archive (line 5 of Algorithm 1) but discarded in the same step (lines 23 to 26 of Algorithm 1), leading to an increase of  $\Delta$ .

**Remark 2.** For the performance assessment of MOEAs, it is typically advisable to take instead of the Hausdorff distance  $d_H$  the averaged Hausdorff distance  $\Delta_p$ . The main reason for this is that MOEAs may compute a few outliers in particular if the MOP contains weakly dominated solutions that are not optimal (also called dominance resistance solutions [72]). On the other hand, we stress that  $\Delta_p$ , opposed to  $d_H$ , is not a metric in the mathematical sense, since the triangle inequality does not hold. We refer, e.g., to [13,38,73,74] for more discussion on this matter.

In the following, we discuss one possibility to obtain an approximation of the value of  $\Delta_p$  from a given archive  $A$ . To this end, we first investigate the value of  $\Delta_p$  if the elements of  $A$  are perfectly located around a linear connected Pareto front (if  $N$  is large enough, we can expect that this approximation works fine for any connected Pareto front). That is, all  $a_i$  values are optimal. Furthermore, if  $A$  is sorted,  $F(a_1)$  and  $F(a_N)$  are the end points of the Pareto front, and the distance of two consecutive elements  $F(a_i)$  and  $F(a_{i+1})$  is given by  $2h$  (leading to  $d_H = h$ ). Since all the  $a_i$  values are optimal, the  $\Delta_p$  value is hence given by the value of  $IGD_p$ , which can be computed as follows:

$$\begin{aligned}
 \Delta_p(F(A), F(P_Q)) &= IGD_p(F(A), F(P_Q)) = \left( \frac{1}{F(a_N) - F(a_1)} \int_{F(a_1)}^{F(a_N)} \text{dist}(t, F(A))^p dt \right)^{\frac{1}{p}} \\
 &= \left( \frac{1}{(N-1)2h} 2(N-1) \int_0^h t^p dt \right)^{\frac{1}{p}} \\
 &= \left( \frac{1}{h} \left[ \frac{1}{p+1} t^{p+1} \right]_0^h \right)^{\frac{1}{p}} = \left( \frac{1}{h} \frac{1}{p+1} h^{p+1} \right)^{\frac{1}{p}} \\
 &= \sqrt[p]{\frac{1}{p+1}} \cdot h
 \end{aligned}
 \tag{13}$$

Hereby, we have used the formulation of  $IGD_p$  for continuous Pareto fronts as discussed in [73]. It remains to compute  $h$ . Since the assumption that all the images of the  $a_i$  values are evenly spread is ideal, we cannot simply take  $\frac{1}{2} \|F(a_{i+1}) - F(a_i)\|$  for an arbitrarily index  $i \in \{1, \dots, N-1\}$ . Instead, it makes sense to use the average of these distances:

$$h \approx \frac{\sum_{i=1}^{N-1} \tilde{d}_i}{2m}, \quad (14)$$

where  $\tilde{d}_i$  is as in (12) and  $m$  denotes the number of elements of  $\tilde{d}_i$  that are not equal to zero. This leads to the approximation  $d_p$  of the averaged Hausdorff distance  $\Delta_p$  of the Pareto front by a given archive  $A$ :

$$d_p := \sqrt[p]{\frac{1}{p+1} \cdot \frac{\sum_{i=1}^{N-1} \tilde{d}_i}{2m}}. \quad (15)$$

In order to obtain a first impression on the effect of the archiver, we apply it to several test problems. More precisely, we use ArchiveUpdateHD together with the generator, which is simply choosing candidate solutions uniformly at random from the domain of the problem. As test problems, we use CONV (convex front), DENT ([75], convex-concave front), RUD1 and RUD2 (disconnected fronts), LINEAR (linear front) and RUD3 (convex front). The first five test problems are uni-modal, while RUD3 has next to the Pareto front eight local fronts. RUD3 is taken from [76], and RUD1 and RUD2 are straightforward modifications of RUD3 to obtain the given Pareto front shapes.

Figure 4 shows the final approximations of the fronts using  $N = 30$  for the archive size and initial values of  $\Delta$  small enough so that this threshold is reached for all problems. As it can be seen, in all cases, evenly distributed solutions along the Pareto fronts have been obtained. Figure 5 shows the actual Hausdorff and averaged Hausdorff values of the computed archives in each step for one run of the algorithm ( $d_H$  and  $\Delta_2$ , i.e.,  $p = 2$  has been used for the averaged Hausdorff distance), together with their approximations  $h$  and  $d_2$ . For all problems, the archiver is capable of quickly determining a good approximation of both  $d_H$  and  $\Delta_2$  during the run of the algorithm. Tables A1 and A2 show the approximation qualities averaged over 30 independent runs, which support the observations from Figure 5. Figure 6 shows the evolution of the value of  $\Delta$  during one run of the algorithm for DENT and RUD3. For the uni-modal problem DENT, the value of  $\Delta$  is essentially increasing monotonically (i.e., not counting the first few iteration steps), while for the multi-modal problem RUD3, more than 10 restarts occur. Nevertheless, in both cases, a final value  $\Delta^+$  is reached, which is in accord with Theorem 1.

Figure A1 shows the box collections

$$C(A_f) := \bigcup_{a \in A_f} B_{\Delta_f}(F(a)) \quad (16)$$

of the final archives  $A_f$  and the final value  $\Delta_f$  for the test problems, where  $B_\Delta(x)$  denotes the  $\Delta$ -ball around  $x$  using the maximum norm. The figure indicates that the Hausdorff distance of  $F(A_f)$  and the respective Pareto fronts is indeed less or equal to  $\Delta_f$  for all problems.

---

**Algorithm 3**  $\{A, \Delta, h\} := \text{ArchiveUpdateHD}(P, A_0, \Delta_0, N)$

---

**Require:** Problem (MOP), where  $k = 2$ ,  $P$ : current population,  $A_0$ : current archive,  $\Delta_0 \in \mathbb{R}_+^k$ : current values of  $\Delta$ ,  $N$ : upper bound for archive size

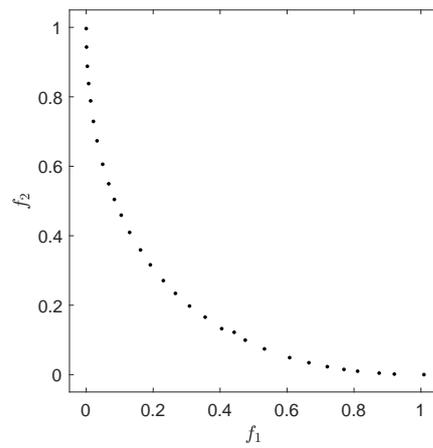
**Ensure:** updated archive  $A$ , updated values for  $\Delta$ , Hausdorff approximation  $h$

```

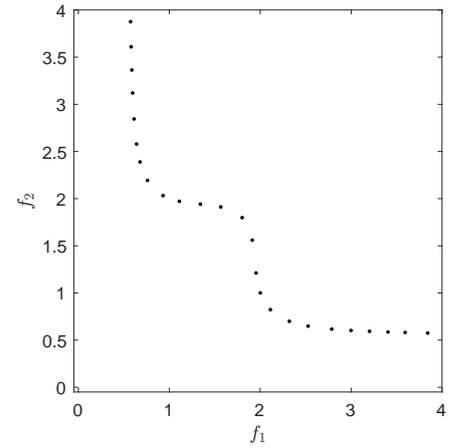
1:  $A := A_0$ 
2:  $\Delta := \Delta_0$ 
3:  $\epsilon := \Delta$ 
4: for all  $p \in P$  do
5:   if  $\nexists a \in A : a \prec_\epsilon p$ , or  $\nexists a \in A : a \prec p$  and  $\exists a \in A : |f_i(a) - f_i(p)| \leq \Delta_i$ ,
       $i = 1, \dots, k$  then
6:      $A := A \cup \{p\}$ 
7:   end if
8:   for all  $a \in A$  do
9:     if  $p \prec a$  then
10:       $A := A \cup \{p\} \setminus \{a\}$ 
11:      if  $\exists i \in \{1, \dots, k\} : f_i(a) - f_i(p) > \Delta_i$  then ▷ reset  $\Delta$  and  $\epsilon$ 
12:         $\Delta := \Delta_{min}$ 
13:         $\epsilon := \Delta$ 
14:      end if
15:    end if
16:  end for
17: if  $|A| = N + 1$  then ▷ apply pruning
18:    $\Delta := \frac{N+1}{N} \Delta$ 
19:    $\epsilon := \frac{N+1}{N} \epsilon$ 
20:   sort  $A$  (e.g., according to  $f_1$ )
21:   compute  $d \in \mathbb{R}^N$  as in (8)
22:   choose  $m \in \arg \min d$ 
23:   if  $m = 1$  then
24:      $A := A \setminus \{a_2\}$  ▷ remove 2nd entry
25:   else if  $m = N$  then
26:      $A := A \setminus \{a_N\}$  ▷ remove 2nd but last entry
27:   else
28:      $dl := \|F(a_{m+1}) - F(a_{m-1})\|_2$ 
29:      $dr := \|F(a_{m+2}) - F(a_m)\|_2$ 
30:     if  $dl < dr$  then
31:        $A := A \setminus \{a_m\}$ 
32:     else
33:        $A := A \setminus \{a_{m+1}\}$ 
34:     end if
35:   end if
36: end if
37: end for
38: sort  $A$  (e.g., according to  $f_1$ ) ▷ compute Hausdorff approximation
39: compute  $\tilde{d}_i, i = 1, \dots, |A| - 1$  as in (12)
40:  $h := \frac{1}{2} \max_{i=1, \dots, |A|-1} \tilde{d}_i$ 
41: return  $\{A, \Delta, h\}$ 

```

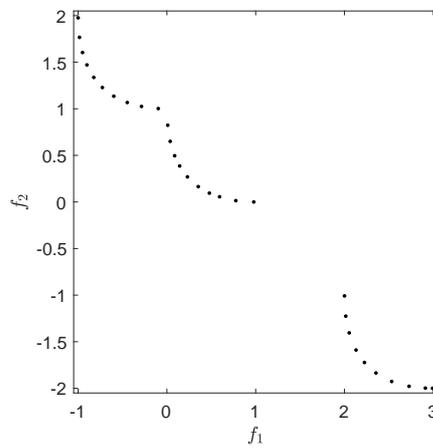
---



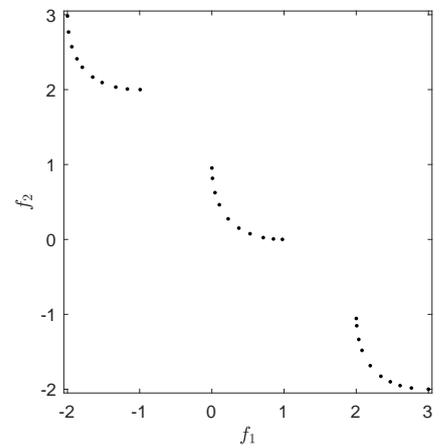
(a) CONV



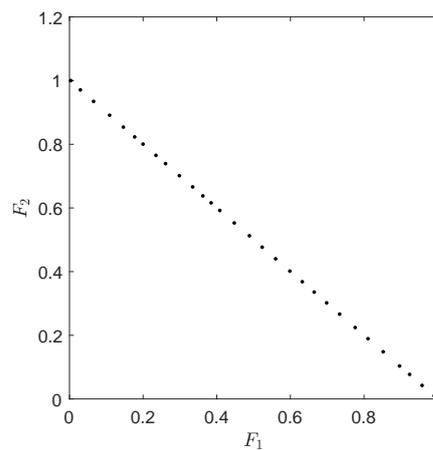
(b) DENT



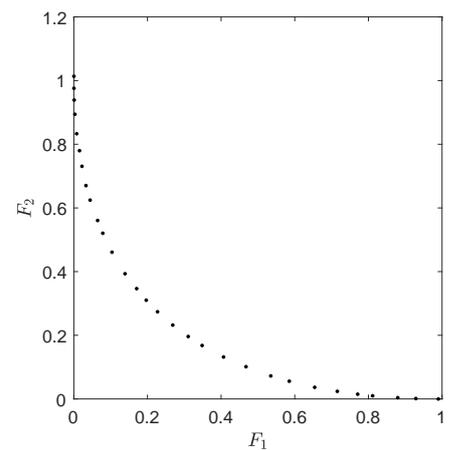
(c) RUD1



(d) RUD2

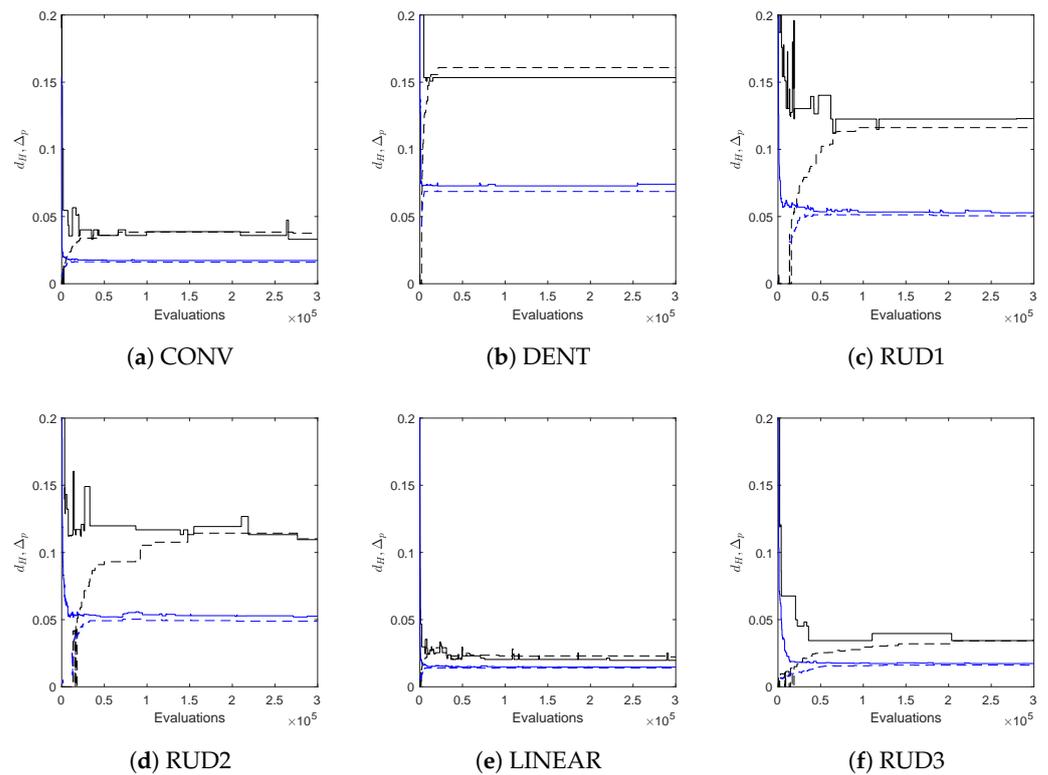


(e) LINEAR

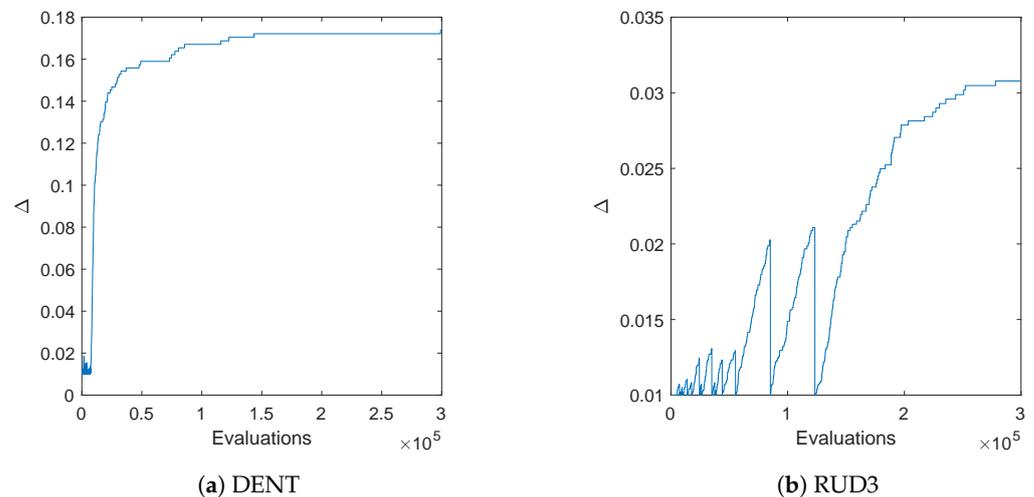


(f) RUD3

**Figure 4.** Numerical results of ArchiveUpdateHD on six BOPs with different shapes of the Pareto fronts. For the sake of clarity, we omitted the fronts that already become apparent by the approximations.



**Figure 5.** Hausdorff and averaged Hausdorff approximations ( $d_H$  and  $\Delta_2$ , respectively) obtained by ArchiveUpdateHD for one single run for six bi-objective problems (see Figure 4) together with their approximations  $h$  and  $d_2$ .  $d_H$  is plotted black solid,  $h$  is black dashed,  $\Delta_2$  is blue solid, and  $d_2$  is blue dashed.



**Figure 6.** Evolution of the value of  $\Delta$  for one run of the algorithm on DENT and RUD3.

### 3.2. The General Case

Next, we consider the archiver for MOPs with more than two objectives. Algorithm 4 shows the pseudocode of ArchiveUpdateHD for such problems. The archiver is essentially identical to the one for BOPs; however, it comes with two modifications, since one cannot expect the Pareto front to form a one-dimensional object any more and another one prevents too many unnecessary resets during the run of the algorithm.

1. The distances cannot be sorted any more as in (8). Instead, one has to consider the distances

$$d_{i,j} = \|F(a_i) - F(a_j)\|, \quad i, j = 1, \dots, |A|, \quad j > i, \tag{17}$$

for a given archive  $A$ . Furthermore, more sophisticated considerations of the distances as, e.g., in lines 27 and 28 of Algorithm 3 cannot be considered any more. Instead, we have chosen to first compute

$$d_{i_m, j_m} \in \arg \min_{\substack{i, j=1, \dots, N+1 \\ j > i}} d_{i,j}, \tag{18}$$

and then to remove  $a_l$  from the archiver, where  $l$  is chosen randomly from  $\{i_m, j_m\}$ . Similar as for the bi-objective case, an exception can of course be made for the best found solutions for each objective value.

2. The approximation of the Hausdorff distance cannot be done as in (12) any more. Instead, we choose the value of  $\Delta$  as an approximation for  $d_h(F(A), F(P_Q))$ , which is motivated by Theorem 1.
3. The reset is completed if there exists an entry  $a$  of the current archive  $A$  and a candidate solution  $p$  that dominates  $a$  and

$$f_i(a) - f_i(p) > \Delta_i \quad i = 1, \dots, k.$$

That is, the improvement is larger than  $\Delta_i$  for all objectives. It has been observed that if one only asks for an improvement in one objective (as done for the bi-objective case), too many resets are performed in particular for MOPs that contain a “flat” region of the Pareto front.

Note that none of these changes affects the statements made in Theorem 1. Hence, the statements of Theorem 1 also hold if Algorithm 4 is used for MOPs with  $k > 2$  objectives. We stress that this algorithm can of course also be used for the treatment of BOPs; however, in that case, Algorithm 3 seems to be better suited, since both distance considerations and Hausdorff approximation are more sophisticated.

Figure 7 shows an application of Algorithm 4 on the test function DTLZ2 with three objectives (concave and connected Pareto front) for  $N = 300$  and  $N = 500$ . The evolution of the approximated value  $\Delta$  of the Hausdorff distance  $d_h(F(A), F(P_Q))$  together with the real value can be found in Figure 8. Hereby, we have used ArchiveUpdateHD as the external archiver of NSGA-II. The same result could have been obtained using randomly chosen test points within the domain  $Q$ , however, for a much higher amount of test points. Figures 9 and 10 show the respective results for DTLZ7, whose Pareto front is disconnected and convex-concave. In all cases, the archiver is capable of finding evenly spread solutions along the Pareto front, and the value of  $\Delta$  is already after some iterations quite close to the actual Hausdorff distance. In order to suitably handle weakly optimal solutions, we have used the approach we describe in the following remark.

**Remark 3.** *It is known that distance-based archiving/selection for MOPs that contains weakly optimal solutions that are not optimal (dominance-resistant solutions) may lead to unsatisfactory results, since candidates may be included in the archive that are far away from the Pareto front. In [77], it has been suggested to consider the modified objectives*

$$\tilde{f}_i(x) = (1 - \alpha)f_i(x) + \frac{\alpha}{m} \sum_{i=1}^k f_i(x), \quad i = 1, \dots, k, \tag{19}$$

where  $\alpha > 0$  is “small”, instead of the original objectives  $f_i$ ,  $i = 1, \dots, k$ . We have adopted this approach for the treatment of the ZDT and DTLZ functions in this work, using  $\alpha = 0.02$ .

---

**Algorithm 4**  $\{A, \Delta\} := \text{ArchiveUpdateHD}(P, A_0, \Delta_0, N)$

---

**Require:** Problem (MOP),  $P$ : current population,  $A_0$ : current archive,  $\Delta_0 \in \mathbb{R}_+^k$ : current value of  $\Delta$ ,  $N$ : upper bound for archive size

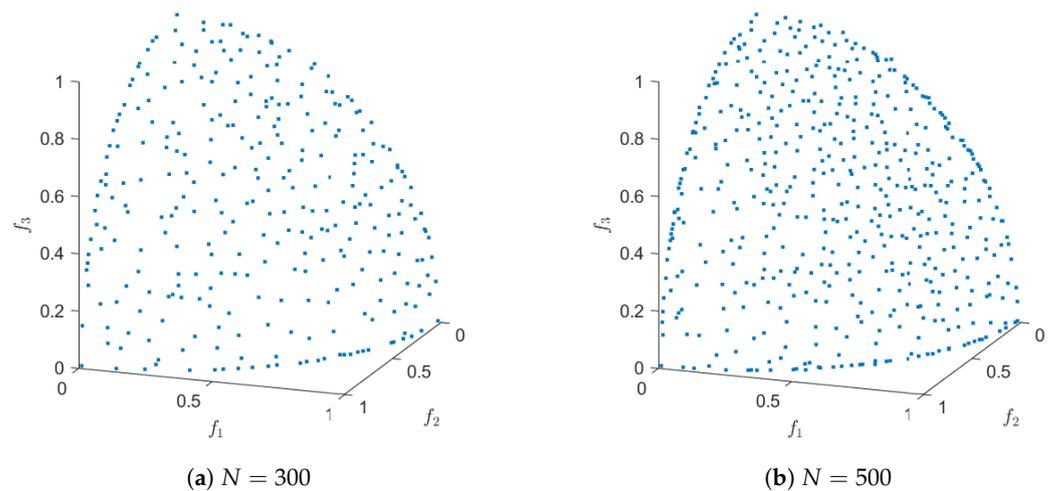
**Ensure:** updated archive  $A$ , updated value of  $\Delta$

```

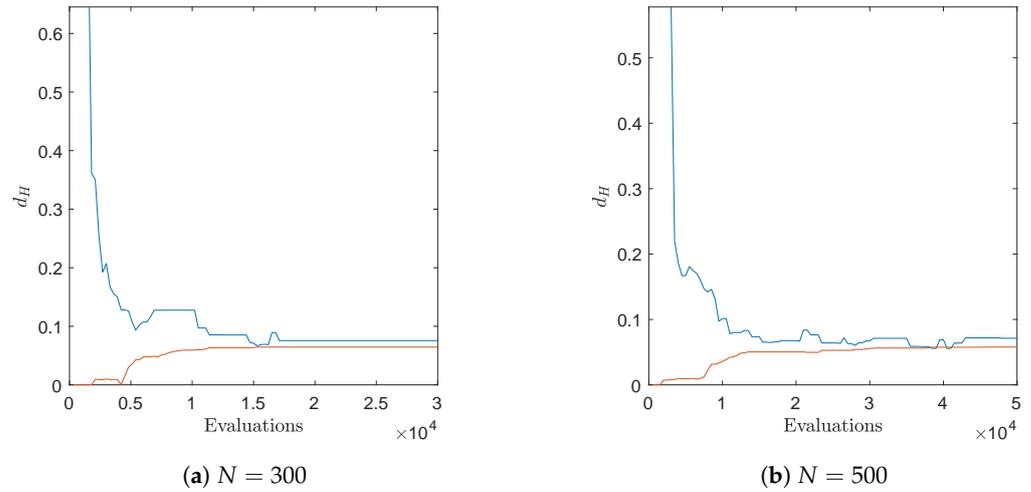
1:  $A := A_0$ 
2:  $\Delta := \Delta_0$ 
3:  $\epsilon := \Delta$ 
4: for all  $p \in P$  do
5:   if  $\nexists a \in A : a \prec_\epsilon p$ , or  $\nexists a \in A : a \prec p$  and  $\nexists a \in A : |f_i(a) - f_i(p)| \leq \Delta_i, i = 1, \dots, k$  then
6:      $A := A \cup \{p\}$ 
7:   end if
8:   for all  $a \in A$  do
9:     if  $p \prec a$  then
10:       $A := A \cup \{p\} \setminus \{a\}$ 
11:      if  $f_i(a) - f_i(p) > \Delta_i, i = 1, \dots, k$ , then ▷ reset  $\Delta$  and  $\epsilon$ 
12:         $\Delta := \Delta_{min}$ 
13:         $\epsilon := \Delta$ 
14:      end if
15:    end if
16:  end for
17:  if  $|A| = N + 1$  then ▷ apply pruning
18:     $\Delta := \frac{N+1}{N} \Delta$ 
19:     $\epsilon := \frac{N+1}{N} \epsilon$ 
20:    compute  $d_{i,j}$  as in (17)
21:    choose  $d_{i_m, j_m} \in \arg \min_{\substack{i,j=1,\dots,N+1 \\ j>i}} d_{i,j}$ 
22:    choose  $l$  randomly from  $\{i_m, j_m\}$ 
23:     $A := A \setminus \{a_l\}$ 
24:  end if
25: end for
26: return  $\{A, \Delta\}$ 

```

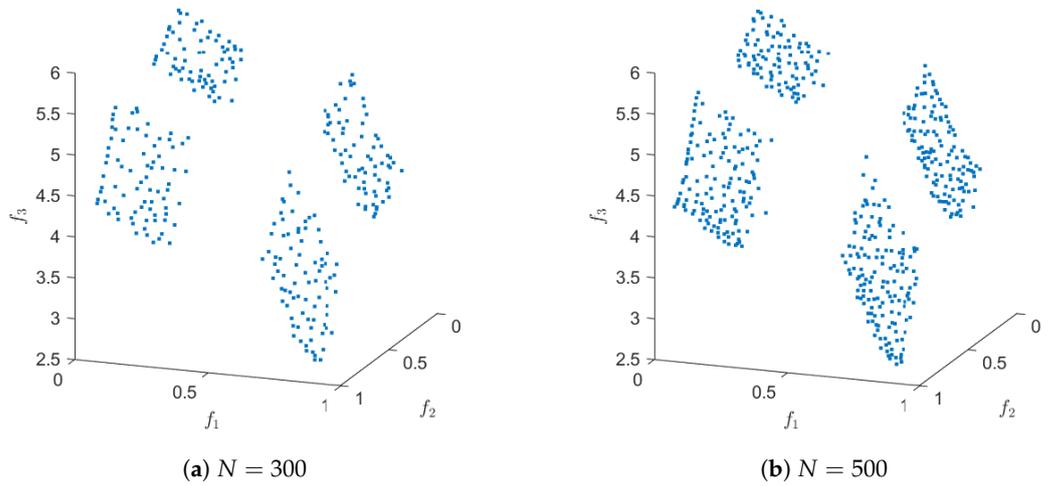
---



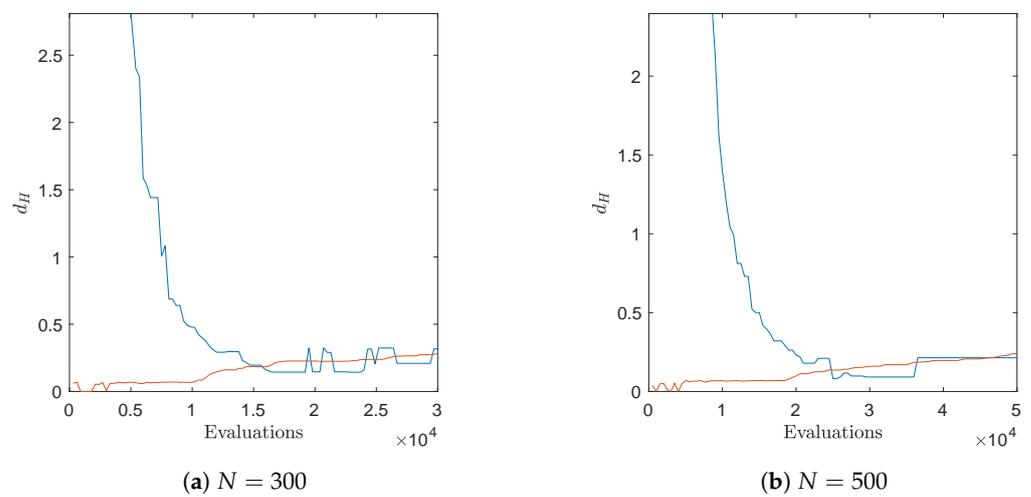
**Figure 7.** Results of ArchiveUpdateHD on DTLZ2 for different values of  $N$ .



**Figure 8.** Real (blue) and approximated (red) Hausdorff distances during the run of one algorithm for DTLZ2.

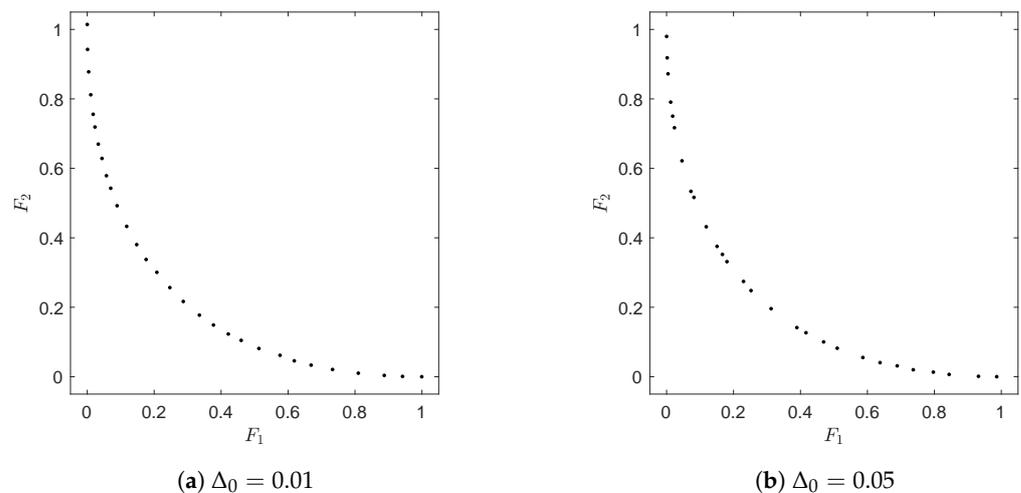


**Figure 9.** Results of ArchiveUpdateHD on DTLZ7 for different values of  $N$ .



**Figure 10.** Real (blue) and approximated (red) Hausdorff distances during the run of one algorithm for DTLZ7.

**Remark 4.** We finally stress that the archiver  $A$  only reaches the magnitude  $N$  if  $\Delta$  (and hence  $\epsilon$ ) is chosen “small enough”, which does not represent a drawback in our opinion. In real-world applications, the values of  $\Delta$  have a physical meaning. As a hypothetical example, consider that one objective in the design of the car is its maximal speed (e.g.,  $f_1 = s_{max}$ ), and the decision maker considers two cars to have different maximal speeds if  $s_{max}$  differs by at least 10 km/h. In this case  $\Delta_1 = 10$  is a suitable choice for ArchiveUpdateHD. Hence, depending on these values and the size of the Pareto front, it may happen that less than  $N$  elements are needed to suitably represent the solution set. In turn, if  $\Delta^+$  is (significantly) larger than the target values, this gives a hint to the decision maker that  $N$  has to be increased and that the computation has to be repeated in order to obtain a “complete” approximation. Figure 11 shows two results of ArchiveUpdateHD on CONV for two different starting values of  $\Delta$ .



**Figure 11.** Results of ArchiveUpdateHD on CONV for two different initial values of  $\Delta$  using  $N = 30$ . For  $\Delta_0 = 0.01$ , the final archive contains 30 elements, while there are only 28 elements for  $\Delta_0 = 0.05$ . The solutions on the left are more evenly spread along the Pareto front due to distance considerations in the pruning technique. For the solution on the right, no pruning technique has been applied during the run of the algorithm.

#### 4. Numerical Results

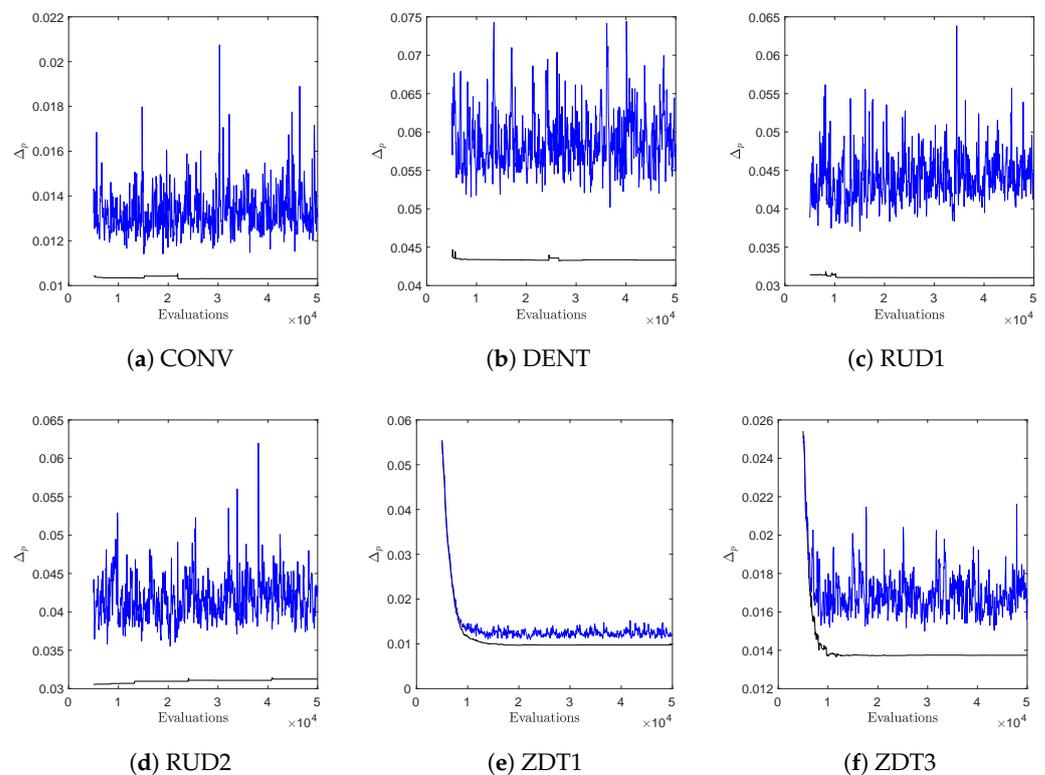
In this section, we show some more numerical results to further demonstrate the advantage of the proposed archiver. As base MOEAs, we have chosen to take the state-of-the-art algorithms NSGA-II (dominance based), MOEA/D (decomposition based), and SMS-EMOEA (indicator based). We have used the implementations of the algorithms as well as the reference fronts provided by PlatEMO [78]. For sake of a fair comparison, we will in the following equip these MOEAs with ArchiveUpdateHD as an external archiver, where the upper bound  $N$  is chosen equal to the population sizes. For each run of an algorithm, we have fed the archiver with exactly the same candidate solution as for the respective base MOEA.

Motivated by Theorem 1 and by the discussion made in Remark 2, we will primarily use  $\Delta_p$  ( $p = 2$ ) for the performance assessment of the MOEA results. However, we will also use the Hpyervolume indicator [79], leading to some surprising results.

We first make a comparison with NSGA-II to investigate possible cyclic behavior during the run of an algorithm. It is known that distance-based selection/archiving strategies may lead to such cyclic behavior, which means that from a certain stage of the search, no more improvements can be expected (and in particular no convergence). The selection strategy of NSGA-II is mainly distance based (since from a certain point on, all individuals of the population are mutually non-dominated). For this, we have set both population size and  $N$  to 50 and have run NSGA-II for 1000 generations, using  $P_c = 1$  and  $\eta_c = 20$  for SBX, and  $P_m = 1/N$  and  $\eta_m = 20$  for polynomial mutation.

For ArchiveUpdateHD, we have chosen the first  $\epsilon_0$  small enough so that archive size  $|A| = N$  was reached for all problems. In Figure 12, typical evolutions of the values of the approximation qualities  $\Delta_p$  are shown over time for six selected BOPs (similar plots are obtained for all problems considered in this study). Hereby, “NSGA-II” stands for the population of the MOEA, and “NSGA-II-A” stands for the respective archive that was fed with the same candidate solutions as NSGA-II. While NSGA-II reveals clear cyclic behavior in all cases, this is not the case for NSGA-II-A. The latter is due to the acceptance strategy of ArchiveUpdateHD that is based on  $\epsilon$ -dominance. As discussed above, the value of  $\Delta$  (and hence also of  $\epsilon$ ) will become large enough during the run of the algorithm so that only dominance replacements will occur, which, however, cannot lead to cyclic behavior.

Apart from the “quasi-monotonic” behavior, one can also observe that the  $\Delta_2$  values of NSGA-II-A are for all test problems significantly lower than the ones of NSGA-II.



**Figure 12.** Approximation qualities of the Pareto fronts (measured by  $\Delta_2$ ) during one run of the algorithm for NSGA-II (blue) and the archives NSGA-II-A (black) for six selected BOPs.

Next, we investigate the performance of ArchiveUpdateHD as an external archive for the three MOEAs using an extended set of test functions. We first consider bi-objective problems. For this, we have chosen the ZDT problems [80], where we have used the modified objectives as expressed in (19) using  $\alpha = 0.02$  to handle weak Pareto optimal solutions that are not optimal. Next to these six test problems, we have taken another four BOPs, which were selected due to the shapes of their fronts: LIN ([81], linear front), CONV (convex front), as well as DENT and SSW [82], which have both convex–concave Pareto fronts. The boxplots of the results are shown in Figures 13 and 14, based on 30 independent runs, for each using 1000 generations and a population size of 50. The Wilcoxon rank-sum is shown in Table 1. In the following, we will compare the results of the base algorithms against the respective solutions that use ArchiveUpdateHD.

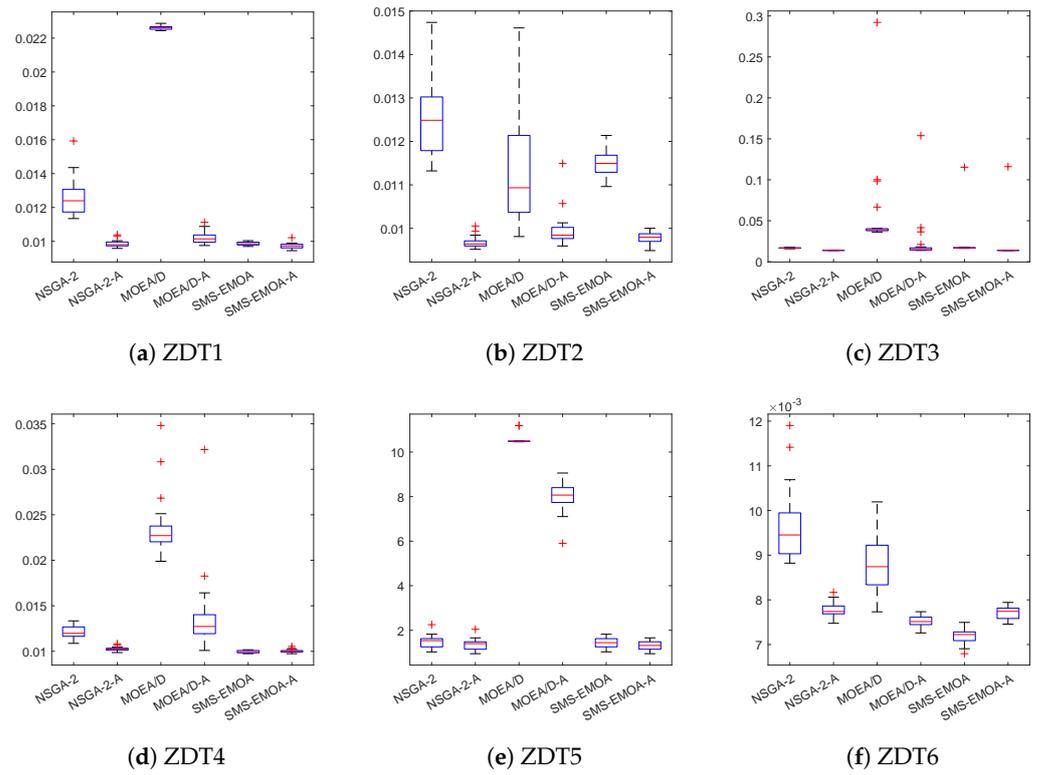


Figure 13. Boxplots for the obtained results for the ZDT test functions.

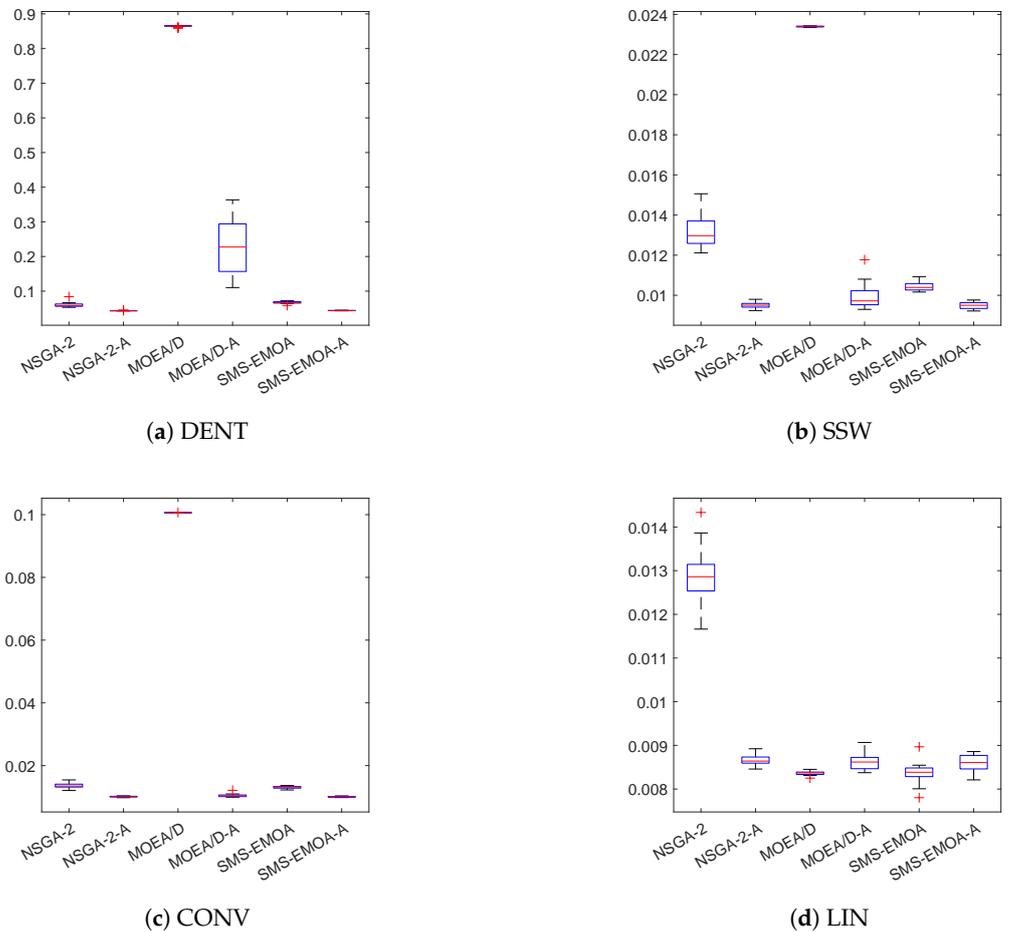


Figure 14. Boxplots for the obtained results for DENT, SSW, CONV and LIN.

The performance of the external archives is better in 10 out of 10 cases for NSGA-II, in 9 out of 10 cases for MOEA/D, and in 8 out of 10 cases for SMS-EMOA. ArchiveUpdateHD loses against the MOEA/D and SMS-EMOA on test problem LIN, which has a linear and thus most possible regular Pareto front. Both MOEA/D and SMS-EMOA are able to compute perfect solutions for this problem. Such perfect approximations cannot be expected from ArchiveUpdateHD due to its acceptance strategy. While this strategy is responsible for suppressing any cyclic behavior, it also prevents that all the solutions even of the limit archive are perfectly evenly spread along the Pareto front. For more complex Pareto fronts, the situation, however, changes. Figures A2 and A3 show the average results obtained by the different methods on problems DENT and ZDT, respectively. For DENT, the use of ArchiveUpdateHD leads in all three cases to significantly better Pareto front approximations. This is similar to ZDT3, while the improvements are less, since all three base MOEAs can already detect very good approximations.

**Table 1.** Comparison (wins/ties/losses) of the results of the base MOEAs against their archive equipped variants on the bi-objective test problems. The Wilcoxon rank-sum test has been used for statistical significance, where  $p$ -value  $< 0.05$ .

Method 1	Method 2	Wins	Ties	Losses
NSGA-2	NSGA-2-A	0	0	10
MOEA/D	MOEA/D-A	1	0	9
SMS-EMOA	SMS-EMOA-A	2	0	8

In a next step, we investigate the effect of ArchiveUpdateHD on several test problems with  $k = 3$  objectives. For this, we have chosen the seven DTLZ test problems, the test functions IDTLZ1 and ITDLZ2 [21] with “inverted” fronts, and MaF1 to 5 [83]. Table 2 shows the Wilcoxon rank-sum for these 14 test problems using the indicators  $\Delta_p$ , HV, as well as the classical Hausdorff distance  $d_H$ . Figure A4 shows the boxplots for all algorithms and test problems for  $\Delta_p$ , Figures A5 and A6 show the results of the algorithms on IDTLZ1 and MaF2, respectively, and Figure A7 shows the selected behaviors of the Hausdorff approximations. For the latter, we have taken the median runs with respect to  $\Delta_p$ . As it can be seen from Table 2, the use of ArchiveUpdateHD as the external archiver is highly beneficial in almost all cases. More precisely, starting with  $\Delta_p$ , NSGA-II-A is better than NSGA-II in 12 out of 14 cases, and it only becomes (slightly) beaten on DTLZ5 and 6, which is likely owed to the degeneration the Pareto fronts of these two test problems (which has to be investigated in more detail in the future). MOEA/A is superior to MOEA/D in 10 out of 14 cases with one tie (DTLZ4) and 3 losses (DTLZ1-3), which is due to the regular structure of these Pareto fronts where MOEA/D can hardly be beaten. Finally, SMS-EMOA-A yields better results in all of the 14 cases. The situation is quite similar when considering the other two performance indicators. While this was expected for  $d_H$ , the results are surprising for HV: note that SMS-EMOA-A also outperforms SMS-EMOA on all of the 14 test functions when considering the hypervolume indicator.

Finally, Figures A8 and A9 show evolutions of the obtained Hausdorff distances of NSGA-II-A for the same test problems but now using  $k = 4$  and  $k = 5$  objectives. The results already give evidence that the values of  $\Delta$  obtained by ArchiveUpdateHD yield satisfying approximations of the actual values of  $d_H$  also for problems with more objectives. The only exceptions are DTLZ5 and DTLZ6 (both for four and five objectives) as well as MaF4 for  $k = 5$ . By Theorem 1, we know that the runs of the algorithms have simply not been long enough, while it is in turn unclear how long these runs should have been. While these results are satisfying, more investigation has to be done in particular for the treatment of many objective problems, which we leave for future study.

**Table 2.** Comparison (wins (1) / ties (0) / losses (−1)) of the results of the base MOEAs against their archive equipped variants on the 14 three-objective test problems. The Wilcoxon rank-sum test has been used for statistical significance, where  $p$ -value < 0.05.

Indicator	Method 1	Method 2	Result	GroupCount
$\Delta_2$	NSGA-2	NSGA-2-A	−1	12
$\Delta_2$	NSGA-2	NSGA-2-A	0	0
$\Delta_2$	NSGA-2	NSGA-2-A	1	2
$\Delta_2$	MOEAD	MOEAD-A	−1	10
$\Delta_2$	MOEAD	MOEAD-A	0	1
$\Delta_2$	MOEAD	MOEAD-A	1	3
$\Delta_2$	SMS-EMOA	SMS-EMOA-A	−1	14
$\Delta_2$	SMS-EMOA	SMS-EMOA-A	0	0
$\Delta_2$	SMS-EMOA	SMS-EMOA-A	1	0
HV	NSGA-2	NSGA-2-A	−1	11
HV	NSGA-2	NSGA-2-A	0	1
HV	NSGA-2	NSGA-2-A	1	2
HV	MOEAD	MOEAD-A	−1	9
HV	MOEAD	MOEAD-A	0	1
HV	MOEAD	MOEAD-A	1	4
HV	SMS-EMOA	SMS-EMOA-A	−1	14
HV	SMS-EMOA	SMS-EMOA-A	0	0
HV	SMS-EMOA	SMS-EMOA-A	1	0
$d_H$	NSGA-2	NSGA-2-A	−1	13
$d_H$	NSGA-2	NSGA-2-A	0	1
$d_H$	NSGA-2	NSGA-2-A	1	0
$d_H$	MOEAD	MOEAD-A	−1	10
$d_H$	MOEAD	MOEAD-A	0	1
$d_H$	MOEAD	MOEAD-A	1	3
$d_H$	SMS-EMOA	SMS-EMOA-A	−1	14
$d_H$	SMS-EMOA	SMS-EMOA-A	0	0
$d_H$	SMS-EMOA	SMS-EMOA-A	1	0

### 5. Conclusions and Future Work

In this paper, we have presented and analyzed the archiving strategy ArchiveUpdateHD for use within set-based stochastic search algorithms such as multi-objective evolutionary algorithms (MOEAs) for the treatment of multi-objective optimization problems (MOPs). ArchiveUpdateHD is a bounded archiver that is based on distance dominance,  $\epsilon$ -dominance and the distances among the candidate solutions and that aims for evenly spread solutions along the Pareto front of a given MOP. We have shown that the images  $F(A_i)$  of the sequence of archives  $A_i$  generated by this archiver form under certain (mild) conditions of the process to generate candidate solutions with a probability of one of a  $\Delta^+$ -approximation of the Pareto front in the Hausdorff sense, and all entries of  $A_i$  converge to Pareto optimal solutions with a probability of one and for  $i \rightarrow \infty$ . Furthermore, the value  $\Delta^+$  is computed by ArchiveUpdateHD during the run of the algorithm (without any prior knowledge of the Pareto front). Since this value represents the maximal error in the representation, it is of important value for the decision maker (DM). In particular, if the magnitude of the archives reaches the pre-defined value  $N$ , the value of  $\Delta^+$  gives a feedback if the approximation is “complete enough” or not. Empirical studies on several benchmark test problems have shown the benefit of the novel strategy, among others, that the obtained value  $\Delta^+$  gives a good approximation of the actual Hausdorff approximation. For bi-objective problems, we have presented an alternative way to compute this value,

which can even be considered to be tight from the practical point of view. Finally, we have used ArchiveUpdateHD as the external archiver for three state-of-the-art MOEAs (NSGA-II, MOEA/D, and SMS-EMOA), indicating that it is capable of significantly improving the overall performance of these algorithms.

One important next step which we will leave for future work is to use the mechanisms behind ArchiveUpdateHD as the selection strategy within an MOEA. If an external archiver is used, two archives (instead of only one) have to be maintained, leading to an additional overhead, which could be avoided. It is hence intended to utilize ArchiveUpdateHD to design a new class of MOEAs that aims for (averaged) Hausdorff approximations of the Pareto fronts (as, e.g., done in [84–86]).

**Author Contributions:** C.I.H.C. and O.S. have contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The codes will be made publicly available after acceptance.

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

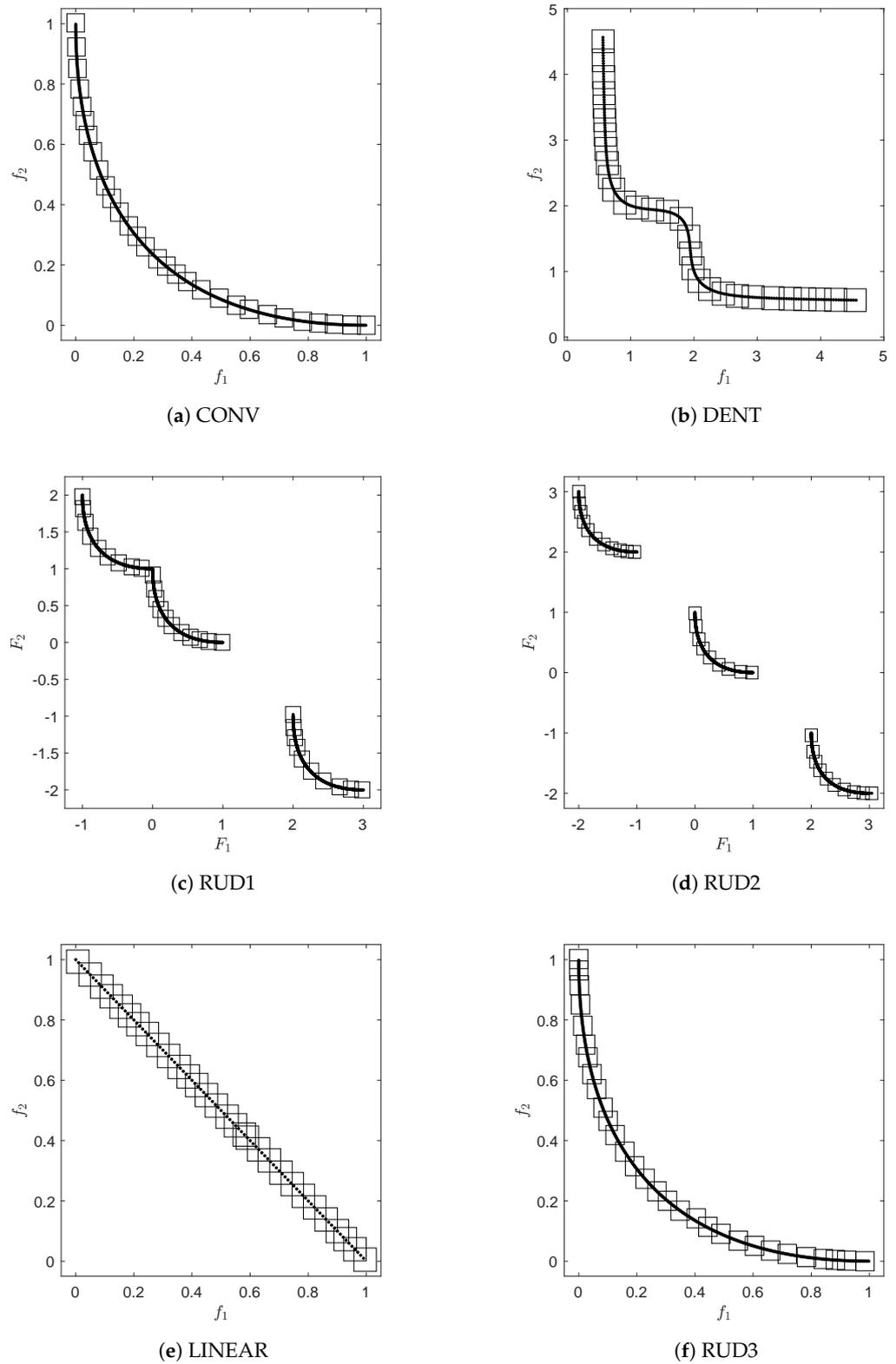
### Appendix A

**Table A1.** Hausdorff distances  $d_H$  and approximations  $h$  computed by ArchiveUpdateHD for the six bi-objective problems.

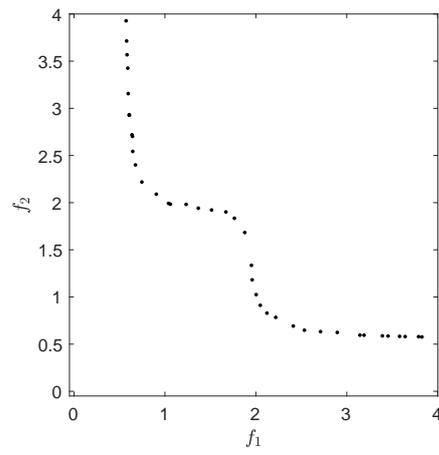
Problem	Mean $d_H$	std $d_H$	Mean $h$	std $h$
CONV	0.0345	0.0014861	0.036105	0.0016771
DENT	0.15428	0.012059	0.15872	0.0091187
RUD1	0.11337	0.0075653	0.10027	0.024311
RUD2	0.11295	0.0083812	0.085445	0.038149
LINEAR	0.02094	0.0019671	0.01977	0.0068899
RUD3	0.035354	0.0023404	0.025694	0.011484

**Table A2.** Averaged Hausdorff distances  $\Delta_2$  and approximations  $d_2$  computed by ArchiveUpdateHD for the six bi-objective problems.

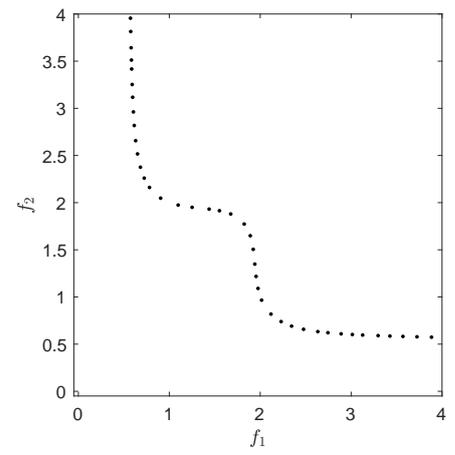
Problem	Mean $\Delta_2$	std $\Delta_2$	Mean $d_2$	std $d_2$
CONV	0.016626	0.00021014	0.01602	$4.9072 \times 10^{-5}$
DENT	0.070802	0.00075385	0.069309	0.00036096
RUD1	0.052171	0.00072633	0.050631	0.0022879
RUD2	0.050967	0.00068389	0.049108	0.00034928
LINEAR	0.014494	0.00033184	0.013892	0.00018454
RUD2	0.016653	0.00028272	0.01595	0.00034355



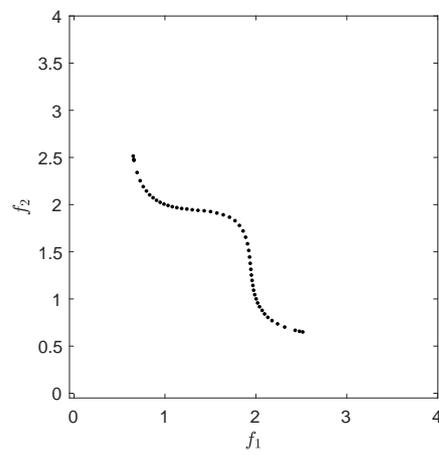
**Figure A1.** The box coverings  $C(A_f)$  for the final archives indicate that the Hausdorff distance between  $F(A_f)$  and the Pareto fronts is less than the final value  $\Delta_f$  computed by ArchiveUpdateHD for all test problems.



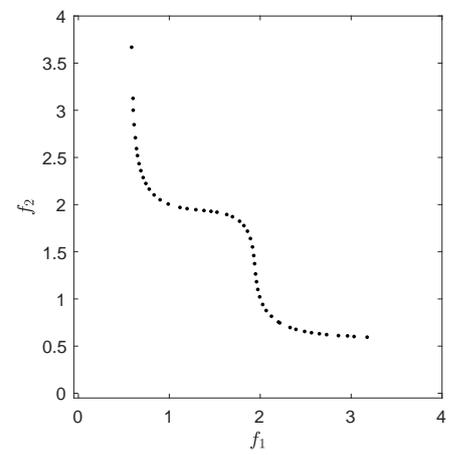
(a) NSGA-II



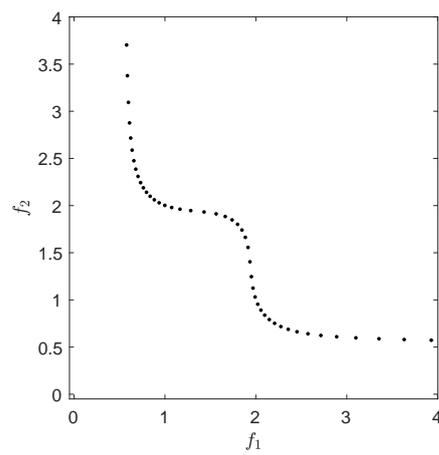
(b) NSGA-II-A



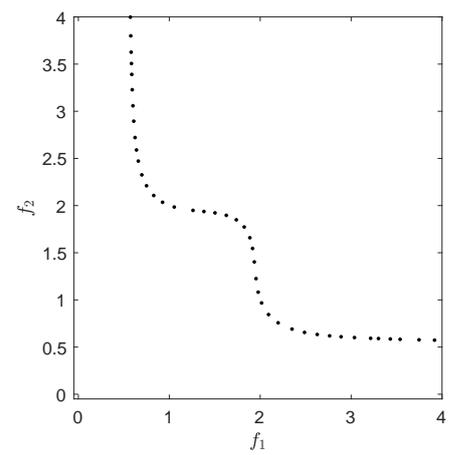
(c) MOEA/D



(d) MOEA/D-A

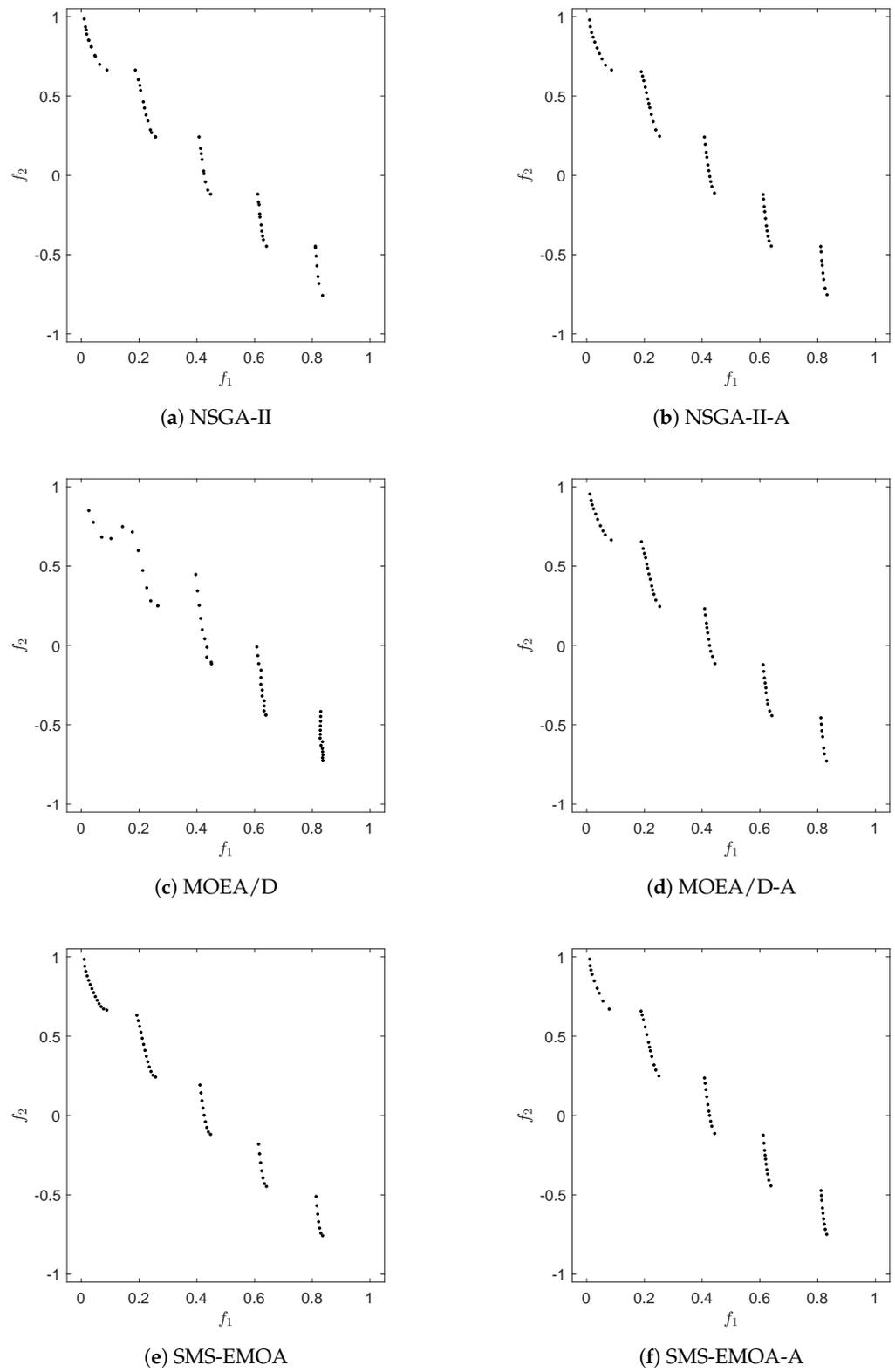


(e) SMS-EMOA



(f) SMS-EMOA-A

**Figure A2.** Numerical results of the different algorithms and archiving/selection strategies on DENT.



**Figure A3.** Numerical results of the different algorithms and archiving/selection strategies on ZDT3.

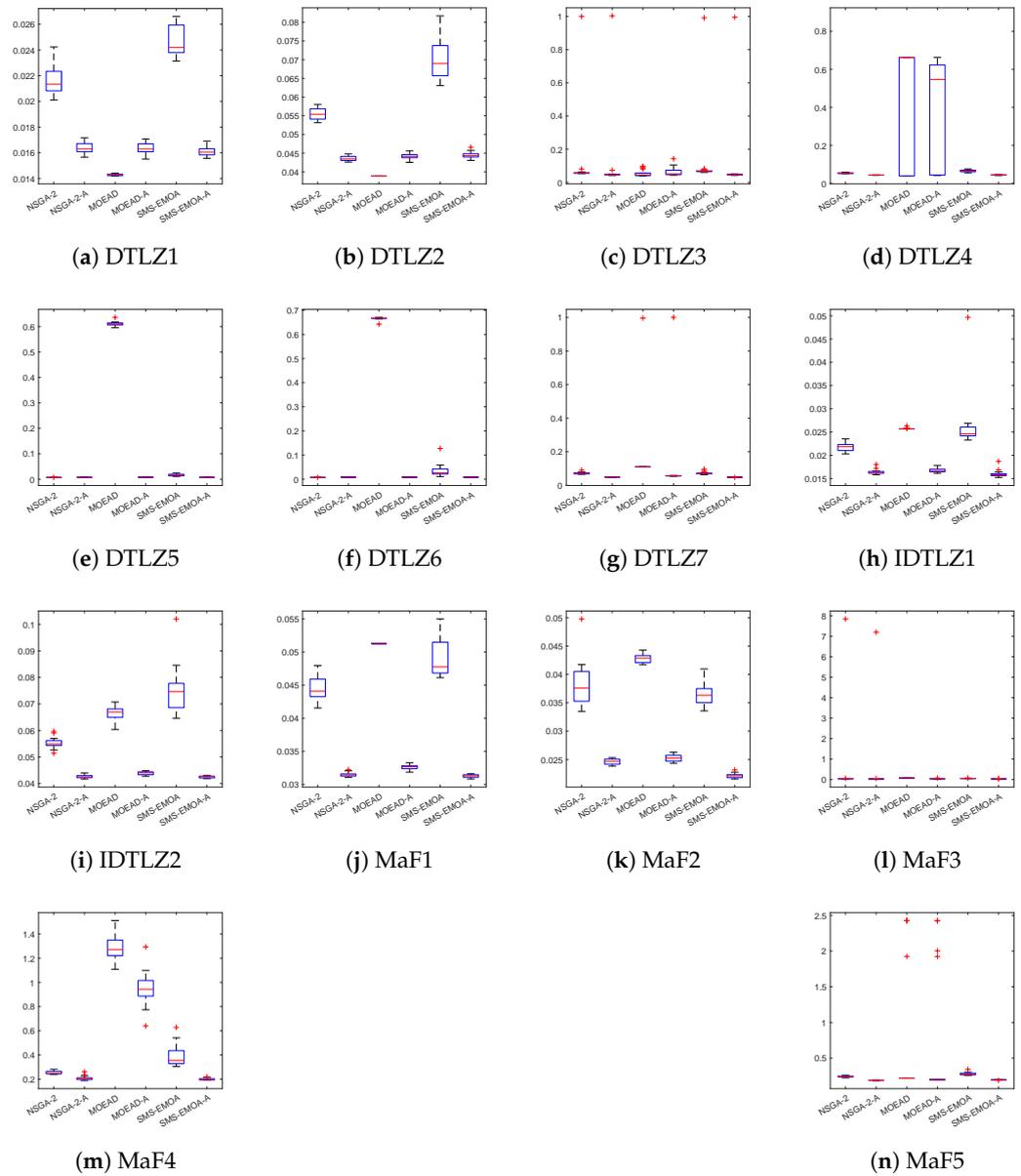
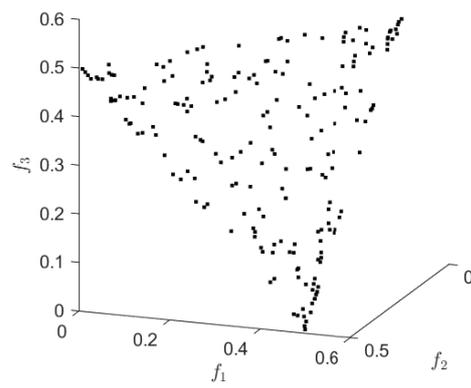
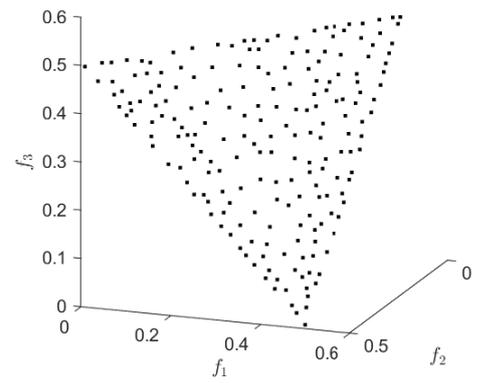


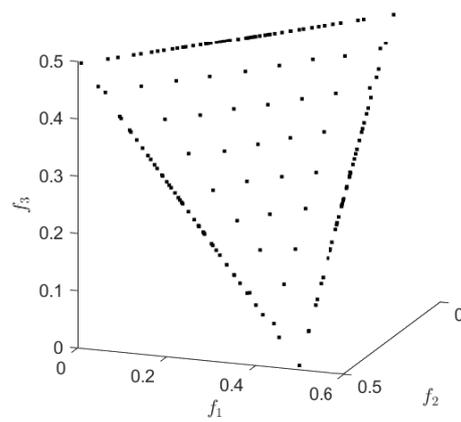
Figure A4. Boxplots for the considered three-objective test functions.



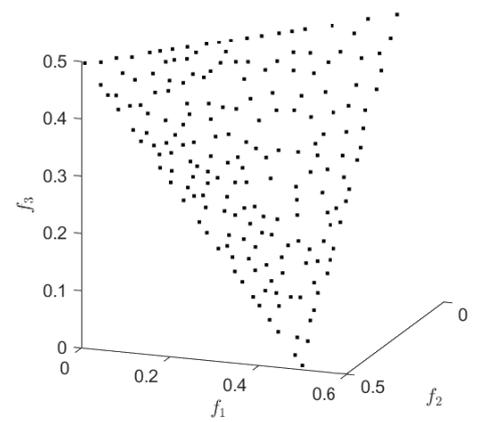
(a) NSGA-II



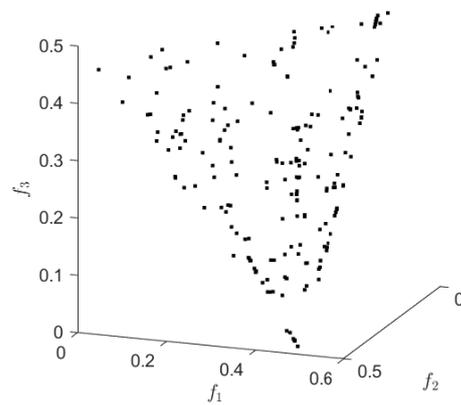
(b) NSGA-II-A



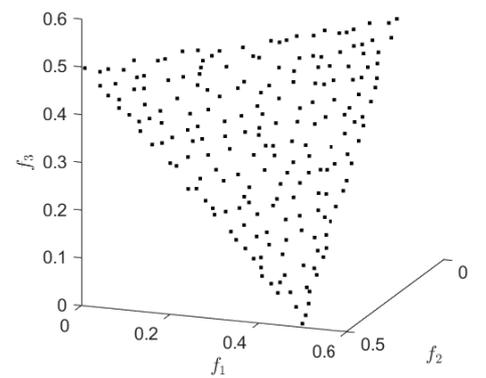
(c) MOEA/D



(d) MOEA/D-A

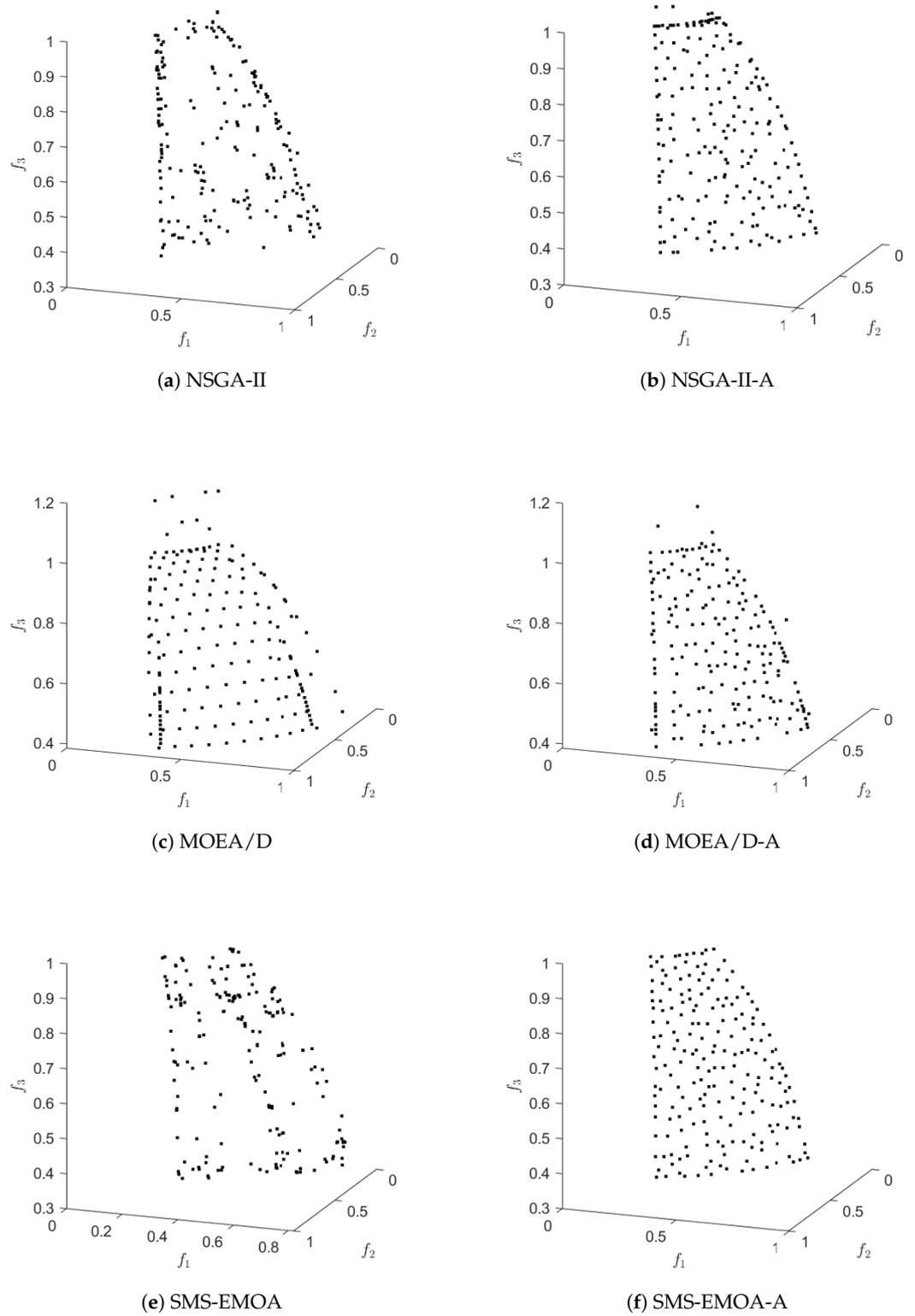


(e) SMS-EMOA

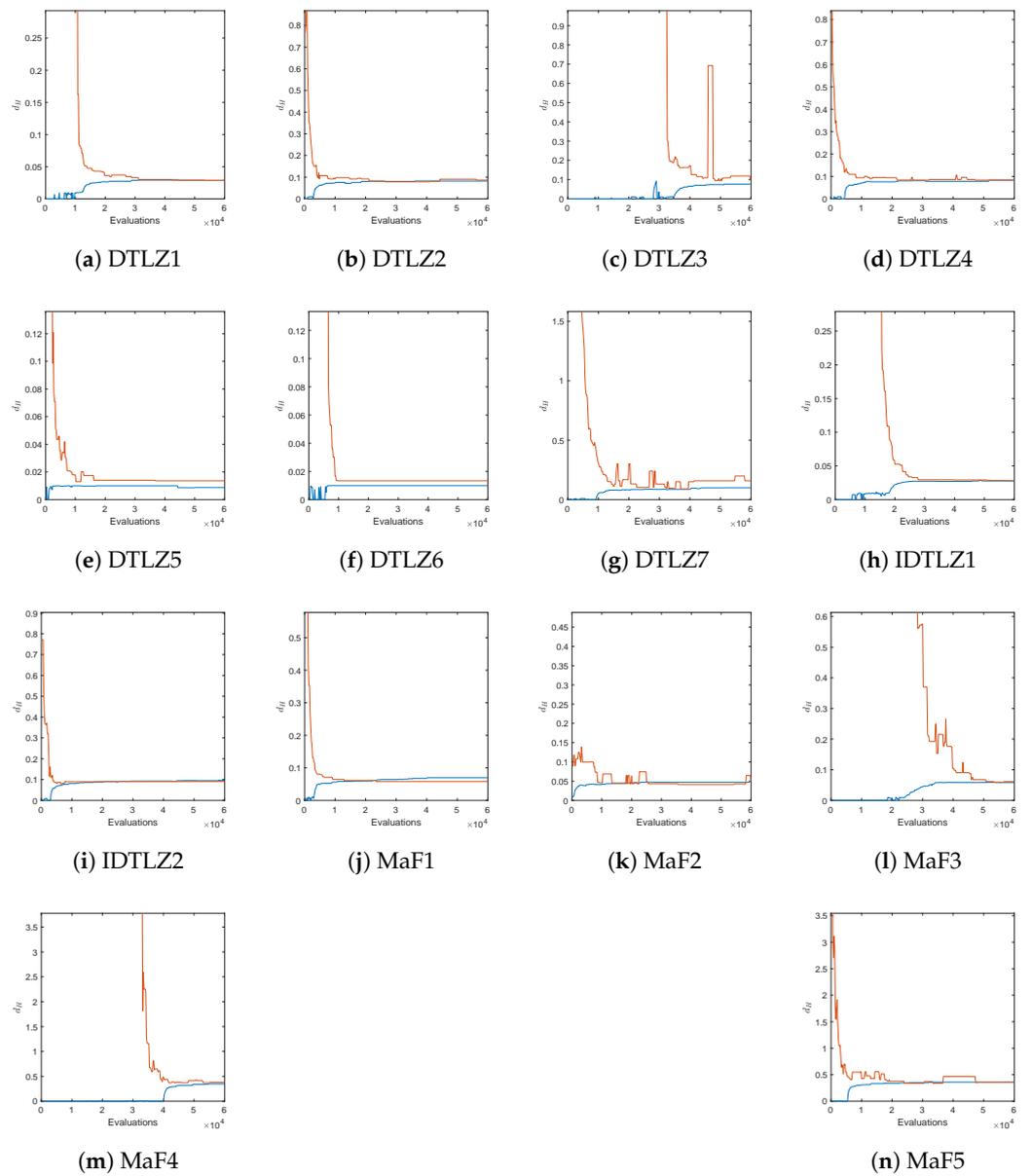


(f) SMS-EMOA-A

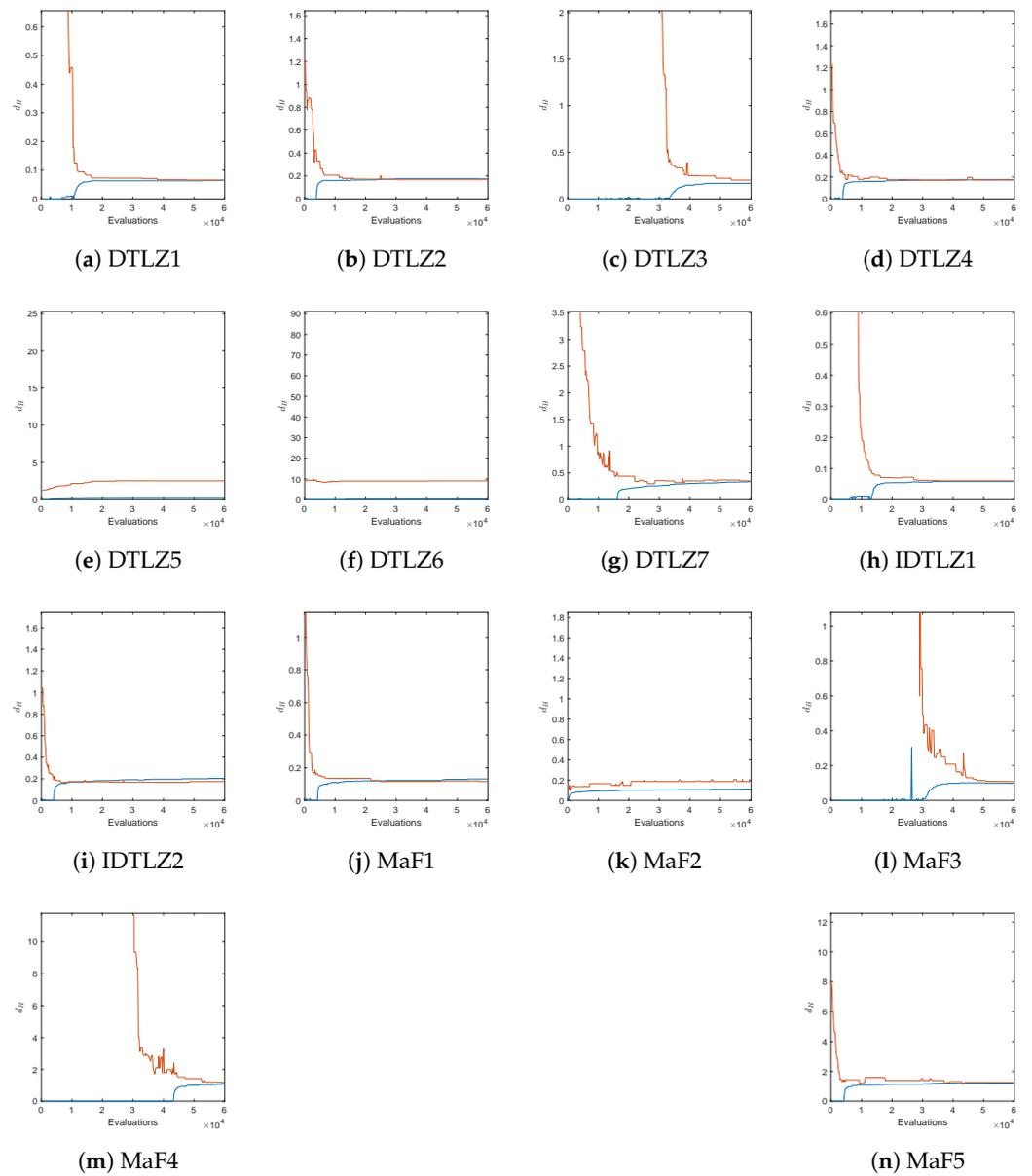
**Figure A5.** Numerical results of the different algorithms and archiving/selection strategies on IDTLZ1 for  $k = 3$ .



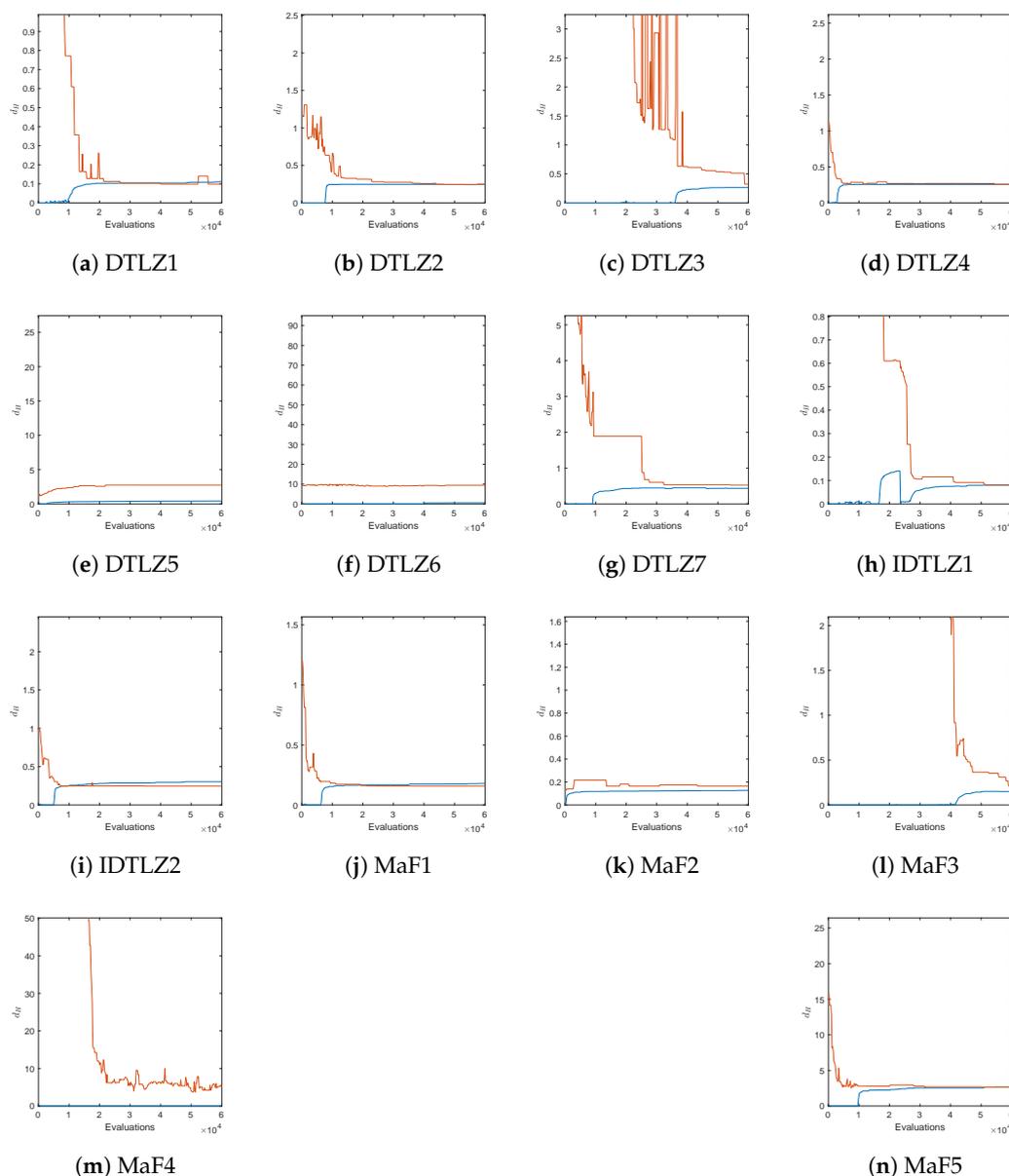
**Figure A6.** Numerical results of the different algorithms and archiving/selection strategies on MaF2 for  $k = 3$ .



**Figure A7.** Evolution of the Hausdorff distances  $d_H(F(A), F(P_Q))$  of NSGA-II-A and the computed approximations  $\Delta$  for several test functions, using  $k = 3$  objectives.



**Figure A8.** Evolution of the Hausdorff distances  $d_H(F(A), F(P_Q))$  of NSGA-II-A and the computed approximations  $\Delta$  for several test functions, using  $k = 4$  objectives.



**Figure A9.** Evolution of the Hausdorff distances  $d_H(F(A), F(P_Q))$  of NSGA-II-A and the computed approximations  $\Delta$  for several test functions, using  $k = 5$  objectives.

**References**

1. Laumanns, M.; Thiele, L.; Deb, K.; Zitzler, E. Combining convergence and diversity in evolutionary multiobjective optimization. *Evol. Comput.* **2002**, *10*, 263–282. [[CrossRef](#)] [[PubMed](#)]
2. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; John Wiley & Sons: Chichester, UK, 2001; ISBN 0-471-87339-X.
3. Stewart, T.; Bandte, O.; Braun, H.; Chakraborti, N.; Ehrgott, M.; Göbel, M.; Jin, Y.; Nakayama, H. Real-World Applications of Multiobjective Optimization. In *Proceedings of the Multiobjective Optimization*; Lecture Notes in Computer Science; Slowinski, R., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5252, pp. 285–327.
4. Sun, J.Q.; Xiong, F.R.; Schütze, O.; Hernández, C. *Cell Mapping Methods—Algorithmic Approaches and Applications*; Springer: Berlin/Heidelberg, Germany, 2019.
5. Aguilera-Rueda, V.J.; Cruz-Ramírez, N.; Mezura-Montes, E. Data-Driven Bayesian Network Learning: A Bi-Objective Approach to Address the Bias-Variance Decomposition. *Math. Comput. Appl.* **2020**, *25*, 37. [[CrossRef](#)]
6. Estrada-Padilla, A.; Lopez-Garcia, D.; Gómez-Santillán, C.; Fraire-Huacuja, H.J.; Cruz-Reyes, L.; Rangel-Valdez, N.; Morales-Rodríguez, M.L. Modeling and Optimizing the Multi-Objective Portfolio Optimization Problem with Trapezoidal Fuzzy Parameters. *Math. Comput. Appl.* **2021**, *26*, 36. [[CrossRef](#)]

7. Frausto-Solis, J.; Hernández-Ramírez, L.; Castilla-Valdez, G.; González-Barbosa, J.J.; Sánchez-Hernández, J.P. Chaotic Multi-Objective Simulated Annealing and Threshold Accepting for Job Shop Scheduling Problem. *Math. Comput. Appl.* **2021**, *26*, 8. [[CrossRef](#)]
8. Castellanos-Alvarez, A.; Cruz-Reyes, L.; Fernandez, E.; Rangel-Valdez, N.; Gómez-Santillán, C.; Fraire, H.; Brambila-Hernández, J.A. A Method for Integration of Preferences to a Multi-Objective Evolutionary Algorithm Using Ordinal Multi-Criteria Classification. *Math. Comput. Appl.* **2021**, *26*, 27. [[CrossRef](#)]
9. Hillermeier, C. *Nonlinear Multiobjective Optimization: A Generalized Homotopy Approach*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2001; Volume 135.
10. 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.
11. Laumanns, M.; Zenklusen, R. Stochastic convergence of random search methods to fixed size Pareto front approximations. *Eur. J. Oper. Res.* **2011**, *213*, 414–421. [[CrossRef](#)]
12. Hernández, C.; Schütze, O. A Bounded Archive Based for Bi-objective Problems based on Distance and epsilon-dominance to avoid Cyclic Behavior. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2022), Boston, MA, USA, 9–13 July 2022.
13. Schütze, O.; Esquivel, X.; Lara, A.; Coello, C.A.C. Using the averaged Hausdorff distance as a performance measure in evolutionary multi-objective optimization. *IEEE Trans. Evol. Comput.* **2012**, *16*, 504–522. [[CrossRef](#)]
14. Schütze, O.; Laumanns, M.; Tantar, E.; Coello, C.A.C.; Talbi, E.G. Computing gap free Pareto front approximations with stochastic search algorithms. *Evol. Comput.* **2010**, *18*, 65–96. [[CrossRef](#)]
15. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evol. Comput. IEEE Trans.* **2002**, *6*, 182–197. [[CrossRef](#)]
16. Zitzler, E.; Laumanns, M.; Thiele, L. SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization. In Proceedings of the Evolutionary Methods for Design, Optimisation and Control with Application to Industrial Problems (EUROGEN 2001), Athens, Greece, 19–21 September 2001; Giannakoglou, K., Tsahalis, D., Périaux, J., Papailiou, K., Fogarty, T., Eds.; International Center for Numerical Methods in Engineering (CIMNE): Barcelona, Spain, 2002; pp. 95–100.
17. Fonseca, C.M.; Fleming, P.J. An overview of evolutionary algorithms in multiobjective optimization. *Evol. Comput.* **1995**, *3*, 1–16. [[CrossRef](#)]
18. Knowles, J.D.; Corne, D.W. Approximating the nondominated front using the Pareto Archived Evolution Strategy. *Evol. Comput.* **2000**, *8*, 149–172. [[CrossRef](#)] [[PubMed](#)]
19. Zhang, Q.; Li, H. MOEA/D: A Multi-objective Evolutionary Algorithm Based on Decomposition. *IEEE Trans. Evol. Comput.* **2007**, *11*, 712–731. [[CrossRef](#)]
20. Deb, K.; Jain, H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints. *Trans. Evol. Comput.* **2014**, *18*, 577–601. [[CrossRef](#)]
21. Jain, H.; Deb, K. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point Based Nondominated Sorting Approach, Part II: Handling Constraints and Extending to an Adaptive Approach. *IEEE Trans. Evol. Comput.* **2014**, *18*, 602–622. [[CrossRef](#)]
22. Martínez, S.Z.; Coello, C.A.C. A multi-objective particle swarm optimizer based on decomposition. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2011), Dublin, Ireland, 12–16 July 2011; pp. 69–76.
23. Zuiani, F.; Vassile, M. Multi Agent Collaborative Search based on Tchebycheff decomposition. *Comput. Optim. Appl.* **2013**, *56*, 189–208. [[CrossRef](#)]
24. Moubayed, N.A.; Petrovski, A.; McCall, J. (DMOPSO)-M-2: MOPSO Based on Decomposition and Dominance with Archiving Using Crowding Distance in Objective and Solution Spaces. *Evol. Comput.* **2014**, *22*, 47–77. [[CrossRef](#)] [[PubMed](#)]
25. Zapotecas-Martínez, S.; López-Jaimes, A.; García-Nájera, A. Libea: A Lebesgue indicator-based evolutionary algorithm for multi-objective optimization. *Swarm Evol. Comput.* **2019**, *44*, 404–419. [[CrossRef](#)]
26. Beume, N.; Naujoks, B.; Emmerich, M. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *Eur. J. Oper. Res.* **2007**, *181*, 1653–1669. [[CrossRef](#)]
27. Zitzler, E.; Thiele, L.; Bader, J. SPAM: Set Preference Algorithm for multiobjective optimization. In Proceedings of the Parallel Problem Solving From Nature PPSN X, Dortmund, Germany, 13–17 September 2008; pp. 847–858.
28. Wagner, T.; Trautmann, H. Integration of Preferences in Hypervolume-based multiobjective evolutionary algorithms by means of desirability functions. *IEEE Trans. Evol. Comput.* **2010**, *14*, 688–701. [[CrossRef](#)]
29. Schütze, O.; Domínguez-Medina, C.; Cruz-Cortés, N.; de la Fraga, L.G.; Sun, J.Q.; Toscano, G.; Landa, R. A scalar optimization approach for averaged Hausdorff approximations of the Pareto front. *Eng. Optim.* **2016**, *48*, 1593–1617. [[CrossRef](#)]
30. Sosa-Hernández, V.A.; Schütze, O.; Wang, H.; Deutz, A.; Emmerich, M. The Set-Based Hypervolume Newton Method for Bi-Objective Optimization. *IEEE Trans. Cybern.* **2020**, *50*, 2186–2196. [[CrossRef](#)] [[PubMed](#)]
31. Bringmann, K.; Friedrich, T. Convergence of Hypervolume-Based Archiving Algorithms. *IEEE Trans. Evol. Comput.* **2014**, *18*, 643–657. [[CrossRef](#)]
32. Fonseca, C.M.; Fleming, P.J. Genetic algorithms for multiobjective optimization: Formulation, discussion, and generalization. In Proceedings of the 5th International Conference on Genetic Algorithms, Urbana-Champaign, IL, USA, 17–21 June 1993; pp. 416–423.

33. Srinivas, N.; Deb, K. Multiobjective optimization using nondominated sorting in genetic algorithms. *Evol. Comput.* **1994**, *2*, 221–248. [[CrossRef](#)]
34. Horn, J.; Nafpliotis, N.; Goldberg, D.E. A niched Pareto genetic algorithm for multiobjective optimization. In Proceedings of the First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Computation, Orlando, FL, USA, 27–29 June 1994; IEEE Press: Piscataway, NJ, USA, 1994; pp. 82–87.
35. 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](#)]
36. Rudolph, G. Finite Markov Chain results in evolutionary computation: A Tour d’Horizon. *Fundam. Inform.* **1998**, *35*, 67–89. [[CrossRef](#)]
37. Rudolph, G. On a multi-objective evolutionary algorithm and its convergence to the Pareto set. In Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC 1998), Anchorage, AK, USA, 4–9 May 1998; IEEE Press: Piscataway, NJ, USA, 1998; pp. 511–516.
38. Rudolph, G.; Agapie, A. Convergence Properties of Some Multi-Objective Evolutionary Algorithms. In Proceedings of the 2000 IEEE Congress on Evolutionary Computation (CEC), La Jolla, CA, USA, 16–19 July 2000; IEEE Press: Piscataway, NJ, USA, 2000.
39. Rudolph, G. Evolutionary Search under Partially Ordered Fitness Sets. In Proceedings of the International NAISO Congress on Information Science Innovations (ISI 2001), Dubai, United Arab Emirates, 17–21 March 2001; ICSC Academic Press: Sliedrecht, The Netherlands, 2001; pp. 818–822.
40. Hanne, T. On the convergence of multiobjective evolutionary algorithms. *Eur. J. Oper. Res.* **1999**, *117*, 553–564. [[CrossRef](#)]
41. Hanne, T. Global multiobjective optimization with evolutionary algorithms: Selection mechanisms and mutation control. In Proceedings of the Evolutionary Multi-Criterion Optimization, First International Conference, EMO 2001, Zurich, Switzerland, 7–9 March 2001; Springer: Berlin/Heidelberg, Germany, 2001; pp. 197–212.
42. Hanne, T. A multiobjective evolutionary algorithm for approximating the efficient set. *Eur. J. Oper. Res.* **2007**, *176*, 1723–1734. [[CrossRef](#)]
43. Hanne, T. A Primal-Dual Multiobjective Evolutionary Algorithm for Approximating the Efficient Set. In Proceedings of the 2007 IEEE Congress on Evolutionary Computation (CEC), Singapore, 25–28 September 2007; IEEE Press: Piscataway, NJ, USA, 2007; pp. 3127–3134.
44. Knowles, J.D.; Corne, D.W. Properties of an adaptive archiving algorithm for storing nondominated vectors. *IEEE Trans. Evol. Comput.* **2003**, *7*, 100–116. [[CrossRef](#)]
45. Corne, D.W.; Knowles, J.D. Some multiobjective optimizers are better than others. In Proceedings of the IEEE Congress on Evolutionary Computation, Canberra, Australia, 8–12 December 2003; IEEE Press: Piscataway, NJ, USA, 2003; pp. 2506–2512.
46. Knowles, J.D.; Corne, D.W. Bounded Pareto archiving: Theory and practice. In *Proceedings of the Metaheuristics for Multiobjective Optimisation*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 39–64.
47. Knowles, J.D.; Corne, D.W.; Fleischer, M. Bounded archiving using the Lebesgue measure. In Proceedings of the IEEE Congress on Evolutionary Computation, Canberra, Australia, 8–12 December 2003; IEEE Press: Piscataway, NJ, USA, 2003; pp. 2490–2497.
48. López-Ibáñez, M.; Knowles, J.D.; Laumanns, M. On Sequential Online Archiving of Objective Vectors. In Proceedings of the Evolutionary Multi-Criterion Optimization (EMO 2011), Ouro Preto, Brazil, 5–8 April 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 46–60.
49. Laumanns, M.; Zitzler, E.; Thiele, L. On the effects of archiving, elitism, and density based selection in evolutionary multi-objective optimization. In Proceedings of the International Conference on Evolutionary Multi-Criterion Optimization, Zurich, Switzerland, 7–9 March 2001; Springer: Berlin/Heidelberg, Germany, 2001; pp. 181–196.
50. Schütze, O.; Laumanns, M.; Coello, C.A.C.; Dellnitz, M.; Talbi, E.G. Convergence of Stochastic Search Algorithms to Finite Size Pareto Set Approximations. *J. Glob. Optim.* **2008**, *41*, 559–577. [[CrossRef](#)]
51. Schütze, O.; Lara, A.; Coello, C.A.C.; Vasile, M. On the Detection of Nearly Optimal Solutions in the Context of Single-Objective Space Mission Design Problems. *J. Aerosp. Eng.* **2011**, *225*, 1229–1242. [[CrossRef](#)]
52. Schütze, O.; Vasile, M.; Coello, C.A.C. Computing the Set of Epsilon-Efficient Solutions in Multiobjective Space Mission Design. *J. Aerosp. Comput. Inf. Commun.* **2011**, *8*, 53–70. [[CrossRef](#)]
53. Schütze, O.; Hernandez, C.; Talbi, E.G.; Sun, J.Q.; Naranjani, Y.; Xiong, F.R. Archivers for the Representation of the Set of Approximate Solutions for MOPs. *J. Heuristics* **2019**, *5*, 71–105. [[CrossRef](#)]
54. Schütze, O.; Hernández, C. *Archiving Strategies for Evolutionary Multi-Objective Optimization Algorithms*; Springer: Berlin/Heidelberg, Germany, 2021.
55. Luong, H.N.; Bosman, P.A.N. Elitist Archiving for Multi-Objective Evolutionary Algorithms: To Adapt or Not to Adapt. In *Proceedings of the Parallel Problem Solving from Nature—PPSN XII*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 72–81.
56. Zapotecas Martínez, S.; Coello Coello, C.A. An archiving strategy based on the Convex Hull of Individual Minima for MOEAs. In Proceedings of the 2010 IEEE Congress on Evolutionary Computation (CEC), Barcelona, Spain, 18–23 July 2010; IEEE Press: Piscataway, NJ, USA, 2010; pp. 1–8.
57. Cai, X.; Li, Y.; Fan, Z.; Zhang, Q. An External Archive Guided Multiobjective Evolutionary Algorithm Based on Decomposition. *IEEE Trans. Evol. Comput.* **2015**, *19*, 508–523.
58. Wang, F.; Zhang, H.; Li, Y.; Zhao, Y.; Rao, Q. External archive matching strategy for MOEA/D. *Soft Comput.* **2018**, *22*, 7833–7846. [[CrossRef](#)]

59. Tanabe, R.; Ishibuchi, H. An analysis of control parameters of MOEA/D under two different optimization scenarios. *Appl. Soft Comput.* **2018**, *70*, 22–40. [[CrossRef](#)]
60. Bezerra, L.C.T.; López-Ibáñez, M.; Stützle, T. Archiver effects on the performance of state-of-the-art multi- and many-objective evolutionary algorithms. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '19), Prague, Czech Republic, 13–17 July 2019; pp. 620–628.
61. Hernández, C.I.; Schütze, O.; Sun, J.Q.; Ober-Blöbaum, S. Non-Epsilon Dominated Evolutionary Algorithm for the Set of Approximate Solutions. *Math. Comput. Appl.* **2020**, *25*, 3. [[CrossRef](#)]
62. Patil, M.B. Improved performance in multi-objective optimization using external archive. *Sādhanā* **2020**, *45*, 1–10. [[CrossRef](#)]
63. Brockhoff, D.; Tran, T.D.; Hansen, N. Benchmarking numerical multiobjective optimizers revisited. In Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation, Madrid, Spain, 11–15 July 2015; pp. 639–646.
64. Wang, R.; Zhou, Z.; Ishibuchi, H.; Liao, T.; Zhang, T. Localized weighted sum method for many-objective optimization. *IEEE Trans. Evol. Comput.* **2016**, *22*, 3–18. [[CrossRef](#)]
65. Pang, L.M.; Ishibuchi, H.; Shang, K. Algorithm Configurations of MOEA/D with an Unbounded External Archive. *arXiv* **2020**, arXiv:2007.13352.
66. Ishibuchi, H.; Pang, L.M.; Shang, K. Solution Subset Selection for Final Decision Making in Evolutionary Multi-Objective Optimization. *arXiv* **2020**, arXiv:2006.08156.
67. Ishibuchi, H.; Pang, L.M.; Shang, K. A New Framework of Evolutionary Multi-Objective Algorithms with an Unbounded External Archive. In *ECAI 2020*; IOS Press: Amsterdam, The Netherlands, 2020; pp. 283–290.
68. Praditwong, K.; Yao, X. A New Multi-objective Evolutionary Optimisation Algorithm: The Two-Archive Algorithm. In Proceedings of the Computational Intelligence and Security, Harbin, China, 15–19 December 2007; Wang, Y., Cheung, Y.M., Liu, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 95–104.
69. Wang, H.; Jiao, L.; Yao, X. Two\_Arch2: An Improved Two-Archive Algorithm for Many-Objective Optimization. *IEEE Trans. Evol. Comput.* **2015**, *19*, 524–541. [[CrossRef](#)]
70. Deb, K.; Agrawal, S. A niched-penalty approach for constraint handling in genetic algorithms. In Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms (ICANNGA-99), Portoroz, Slovenia, 6–9 April 1999; Springer: Berlin/Heidelberg, Germany, 1999; pp. 235–243.
71. Deb, K.; Deb, D. Analysing mutation schemes for real-parameter genetic algorithms. *Int. J. Artif. Intell. Soft Comput.* **2014**, *4*, 1–28. [[CrossRef](#)]
72. Ikeda, K.; Kita, H.; Kobayashi, S. Failure of Pareto-based MOEAs: Does non-dominated really mean near to optimal? In Proceedings of the 2001 IEEE Congress on Evolutionary Computation (CEC), Seoul, Korea, 27–30 May 2001; pp. 957–962.
73. 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](#)]
74. Vargas, A.; Bogoya, J. A Generalization of the Averaged Hausdorff Distance. *Computación y Sistemas* **2018**, *22*, 331–345. [[CrossRef](#)]
75. Witting, K. Numerical Algorithms for the Treatment of Parametric Multiobjective Optimization Problems and Applications. Ph.D. Thesis, Department of Mathematics, University of Paderborn, Paderborn, Germany, 2012.
76. Rudolph, G.; Naujoks, B.; Preuss, M. Capabilities of EMOA to Detect and Preserve Equivalent Pareto Subsets. In *Proceedings of the Evolutionary Multi-Criterion Optimization: 4th International Conference, EMO 2007, Matsushima, Japan, 5–8 March 2007*; Obayashi, S., Deb, K., Poloni, C., Hiroyasu, T., Murata, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 36–50.
77. Ishibuchi, H.; Matsumoto, T.; Masuyama, N.; Nojima, Y. Effects of dominance resistant solutions on the performance of evolutionary multi-objective and many-objective algorithms. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '20), Cancún, Mexico, 8–12 July 2020; Coello, C.A.C., Ed.; ACM: New York, NY, USA, 2020; pp. 507–515. [[CrossRef](#)]
78. Tian, Y.; Cheng, R.; Zhang, X.; Jin, Y. PlatEMO: A MATLAB Platform for Evolutionary Multi-Objective Optimization. *IEEE Comput. Intell. Mag.* **2017**, *3*, 73–87. [[CrossRef](#)]
79. Zitzler, E.; Thiele, L. Multiobjective optimization using evolutionary algorithms—A comparative case study. In *Proceedings of the Parallel Problem Solving from Nature—PPSN V*; Eiben, A.E., Bäck, T., Schoenauer, M., Schwefel, H.P., Eds.; Springer: Berlin/Heidelberg, Germany, 1998; pp. 292–301.
80. Zitzler, E.; Deb, K.; Thiele, L. Comparison of multiobjective evolutionary algorithms: Empirical results. *Evol. Comput.* **2000**, *8*, 173–195. [[CrossRef](#)]
81. Emmerich, M.; Deutz, A. Test problems based on Lamé superspheres. In *Proceedings of the Evolutionary Multi-Criterion Optimization: 4th International Conference, EMO 2007, Matsushima, Japan, 5–8 March 2007*; Obayashi, S., Deb, K., Poloni, C., Hiroyasu, T., Murata, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 922–936.
82. Schaeffler, S.; Schultz, R.; Weinzierl, K. Stochastic Method for the Solution of Unconstrained Vector Optimization Problems. *J. Optim. Theory Appl.* **2002**, *114*, 209–222. [[CrossRef](#)]
83. Cheng, M.; Tian, Y.; Zhang, X.; Yang, S.; Jin, Y.; Yao, X. A benchmark test suite for evolutionary many-objective optimization. *Complex Intell. Syst.* **2017**, *3*, 67–81. [[CrossRef](#)]
84. Rudolph, G.; Trautmann, H.; Sengupta, S.; Schütze, O. Evenly Spaced Pareto Front Approximations for Tricriteria Problems Based on Triangulation. In Proceedings of the Evolutionary Multi-Criterion Optimization Conference (EMO 2013), Sheffield, UK, 19–22 March 2013; pp. 443–458.

- 
85. Rudolph, G.; Grimme, C.; Schütze, O.; Trautmann, H. An Aspiration Set EMOA based on Averaged Hausdorff Distances. In Proceedings of the Learning and Intelligent Optimization Conference (LION 2014), Gainesville, FL, USA, 16–21 February 2014; pp. 153–156.
  86. 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](#)]