Comput Optim Appl (2016) 64:589–618
DOI 10.1007/s10589-015-9815-8

Optimal averaged Hausdorff archives for bi-objective problems: theoretical and numerical results

Günter Rudolph[1] · Oliver Schütze[2] · Christian Grimme[3] · Christian Domínguez-Medina[4] · Heike Trautmann[3]

Received: 4 November 2014 / Published online: 23 December 2015

Abstract One main task in evolutionary multiobjective optimization (EMO) is to obtain a suitable finite size approximation of the Pareto front which is the image of the solution set, termed the Pareto set, of a given multiobjective optimization problem. In the technical literature, the characteristic of the desired approximation is commonly expressed by closeness to the Pareto front and a sufficient spread of the solutions obtained. In this paper, we first make an effort to show by theoretical and empirical findings that the recently proposed *Averaged Hausdorff* (or Δ_p-) indicator indeed aims at fulfilling both performance criteria for bi-objective optimization problems. In the second part of this paper, standard EMO algorithms combined with a specialized archiver and a postprocessing step based on the Δ_p indicator are introduced which sufficiently approximate the Δ_p-optimal archives and generate solutions evenly spread along the Pareto front.

✉ Heike Trautmann
trautmann@uni-muenster.de; trautmann@wi.uni-muenster.de

Günter Rudolph
guenter.rudolph@tu-dortmund.de

Oliver Schütze
schuetze@cs.cinvestav.mx

Christian Grimme
christian.grimme@uni-muenster.de

Christian Domínguez-Medina
hdomingueza09@sagitario.cic.ipn

1 Department of Computer Science, TU Dortmund University, Dortmund, Germany

2 Department of Computer Science, CINVESTAV IPN, Mexico City, Mexico

3 Department of Information Systems, University of Münster, Münster, Germany

4 Computer Research Center, National Polytechnic Institute, Mexico City, Mexico

Keywords Evolutionary computation · Δ_p indicator · Hausdorff distance · Evolutionary multiobjective optimization

1 Introduction

In many real-world applications, it is required to consider several conflicting objectives concurrently leading to *multiobjective optimization problems* (MOPs). One important characteristic of MOPs is that the solution set, the so-called Pareto set, is typically not given by a singleton as for scalar optimization problems. Instead, the Pareto set as well as its image, the Pareto front (PF), typically form a $(k-1)$-dimensional object, where k is the number of objectives involved in the problem [13]. Since these sets can apart from trivial examples not be computed analytically, numerical methods are required that compute suitable finite size approximations. Evolutionary multiobjective algorithms (EMOAs) have caught the interest of many researchers in the recent past since they accomplish this task outstandingly (e.g., [2,4–6,32]). Due to their population based and global approach, EMOAs are capable of computing an entire set of candidate solutions A within one run of the algorithm such that the image of A (denoted by $f(A)$) is 'well-distributed' and sufficiently close to the Pareto front $f(X^*)$, where X^* denotes the Pareto set.

One problem that remains is to measure the performance of these algorithms, i.e., the relation of $f(A)$ to $f(X^*)$. For this, several performance indicators have been proposed so far such as the hypervolume indicator [34] and the R indicators [12]. For an overview and discussion of such indicators we refer e.g. to [1,17,35].

One desire on $f(A)$ that is expressed in the specialized literature is that its elements should ideally be evenly spread around the PF [6,32] in order to give the decision maker an unbiased view of his/her optimal possibilities. A particular example where evenly spread solutions along the PF are advantageous is the online-optimization of mechatronical systems [22,31]. The problem is that the above mentioned indicators do not always prefer such distributions. For instance, the preferences of the R2 indicator heavily rely on the specified distribution of the weight vectors, and it is known that the hypervolume indicator prefers solutions in knee regions of the PF [9]. Though the related attainment surface is indeed close to the Pareto front, this can e.g. not be utilized for the above online optimization problem. An attempt to compute Pareto front approximations with uniform gap can be found in [9], where the evenly spread in terms of Manhattan distance is achieved only in the limit of acute dominance cones with opening angle 0 degree. In this case the size of the dominance cones collapses to zero and finding improvements gets intractable. Therefore this approach is equipped with severe numerical problems.

The indicator used within this study is Δ_p [27] which can be viewed as an averaged Hausdorff distance between $f(A)$ and $f(X^*)$. The aim of this paper is two-fold: first, we make an attempt to show that Δ_p indeed prefers evenly spread solutions around the PF. Second, we present how to compute such Δ_p approximations (also called averaged Hausdorff approximations) by means of EMOAs combined with a special archiving and postprocessing strategy. We explicitly do not aim at proposing the Δ_p indicator as superior to other indicators such as e.g. the hypervolume indicator w.r.t.

to performance evaluation of an EMOA, but we consider it as a viable choice for cases where an evenly spread is desired. The introduced methodology can be applied after *any* EMOA has been run on a particular problem, and the Δ_p indicator is used to extract a most uniformly distributed subset of all nondominated solutions generated within the course of the used EMOA.

To accomplish the first part we investigate optimal archives, i.e., archives of a certain size m that minimize the value of a given indicator, for the indicators Δ_p and "Least IGD under Zero GD" (short: L-IGD) where (I)GD denotes "(Inverted) Generational Distance". The indicator L-IGD measures the IGD value of points whose images lie on the Pareto front. We will show that both indicators are closely related: optimal L-IGD archives converge to optimal Δ_p archives with increasing m. Further, if the PF of a bi-objective problem is connected and concave, both optimal archives coincide. We will further on show that optimal Δ_p archives for problems with linear and spherical concave PFs are indeed evenly distributed on the front. For problems with spherical convex PFs we will show that the optimal Δ_p archive is not on but close to the PF and that at least the optimal L-IGD archives are evenly distributed along these fronts. Finally, we will investigate archives of problems with more complex fronts empirically indicating that in general the Δ_p indicator prefers evenly distributed points along the PFs. In the second part specialized archivers combined with standard EMOAs are introduced which are used as an offline postprocessing step after the optimization. It is experimentally shown that the Δ_p indicator can be efficiently used to improve the original EMOA solution w.r.t. the desired generation of equally spaced PF approximations. Furthermore we illustrate that these solutions sufficiently match the derived optimal archives.

The first EMOA designed for the computation of Hausdorff approximations of the PF was proposed in [11] for bi-objective problems. The challenge for all such algorithms is that the PF is apparently not known a priori. In [11] this was solved by creating a piecewise linear front out of the nondominated solutions of the current archive. This together with an application of the PL metric [20] was the basis of the selection mechanism. Later studies dealt with tri-objective problems. In [28], k-dimensional objective vectors were mapped to a two-dimensional space via *Multi-Dimensional Scaling* (*MDS*) allowing to utilize the archiving strategy used in [11]. In [25], specialized triangulations and boundary detection concepts to approximate the 3D surface of the considered PF approximations were used. Finally, MOPs with four objectives were addressed in [7]. Instead of an approximation of the reference front, the *Part and Selection Algorithm* (PSA, [26]) has been used on the candidate population set. The PSA is capable of quickly selecting an evenly spread subset out of a given data set and serves thus as an alternative to determine the reference solution set, in particular for problems with more than 3 objectives. Numerical results indicated that the PSA based method indeed delivers quite evenly spread solutions along the PF which matches the intention of the Δ_p indicator.

The remainder of this paper is organized as follows: Sect. 2 provides the methodological background. Optimal archives for the Δ_p and related indicators for specific PF shapes are addressed from a theoretical perspective in Sect. 3. In Sect. 4 specialized evolutionary algorithm concepts are introduced which aim at generating equally spaced PF approximations that are optimal w.r.t. the Δ_p indicator while Sect. 5 shows

the results of systematic experimental studies regarding the latter techniques. Comparisons with the optimal archives are provided for special cases in Sect. 6. Finally, conclusions are drawn in Sect. 7.

2 Background

Let $f : X \to \mathbb{R}^k$ be a vector-valued function with domain $X \subseteq \mathbb{R}^n$. In the following we consider MOPs of the form

$$\min\{f(x) : x \in X\}, \tag{MOP}$$

where $f(x) = (f_1(x), \ldots, f_k(x))'$ gathers $k \geq 2$ real-valued objective functions that are to be minimized simultaneously. The optimality of a MOP is defined by the concept of *dominance* [21].

Definition 1 Let $u, v \in F \subseteq \mathbb{R}^k$ where F is equipped with the partial order $\preceq$ defined by $u \preceq v \Leftrightarrow \forall i = 1, \ldots k : u_i \leq v_i$. If $u \prec v \Leftrightarrow (u \preceq v \wedge u \neq v)$ then v is said to be *dominated by* u. An element u is termed *nondominated* relative to $V \subseteq F$ if there is no $v \in V$ that dominates u. The set $\mathsf{ND}(V, \preceq) = \{u \in V \,|\, \nexists\, v \in V : v \prec u\}$ is called the *nondominated set* relative to V.

If $F = f(X)$ is the objective space of some MOP with decision space $X \subseteq \mathbb{R}^n$ and objective function $f(\cdot)$ then the set $F^* = \mathsf{ND}(f(X), \preceq)$ is called the *Pareto front*. Elements $x \in X$ with $f(x) \in F^*$ are termed *Pareto-optimal* and the set X^* of all Pareto-optimal points is called the *Pareto set*.

Moreover, for some $Q \subseteq X \subseteq \mathbb{R}^n$ and $f : X \to \mathbb{R}^k$ the set $\mathsf{ND}_f(Q, \preceq) = \{x \in Q : f(x) \in \mathsf{ND}(f(Q), \preceq)\}$ contains those elements from Q whose images are nondominated in image space $f(Q) = \{f(x) : x \in Q\} \subseteq \mathbb{R}^k$.

Since neither the Pareto set nor the Pareto front can typically be computed analytically, one task in multiobjective optimization is to numerically detect a finite size approximation of $F^* = f(X^*)$. In this work we are particularly interested in a small set distance between the latter and F^* accompanied by a sufficiently good spread. A natural candidate for measuring distances between sets is the Hausdorff distance.

Definition 2 The value $d_H(A, B) := \max(\mathrm{d}(A, B), \mathrm{d}(B, A))$ is termed the *Hausdorff distance* between two sets $A, B \subset \mathbb{R}^k$, where

$$\mathrm{d}(B, A) := \sup\{\mathrm{d}(u, A) : u \in B\} \quad \text{and} \quad \mathrm{d}(u, A) := \inf\{\|u - v\| : v \in A\}$$

for a vector norm $\|\cdot\|$.

The Hausdorff distance is widely used in many fields. It has, however, certain limitations when measuring the distance of the outcome of an EMOA to the PF since outliers generated by EMOAs, i.e. points far away from the remaining ones and the Pareto front, are punished too strongly by d_H. Moreover, as pointed out in [10,27], the distance gets smaller only if the extreme points get closer whereas d_H is indifferent w.r.t. closer proximity for other points. As a remedy, we follow the suggestion of [27] and use the *averaged* Hausdorff distance.

Definition 3 The value $\Delta_p(A, B) = \max(\mathsf{GD}_p(A, B), \mathsf{IGD}_p(A, B))$ with

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

for $p > 0$ is termed the *averaged Hausdorff distance*[1] between sets A and B as given in Definition 2.

The indicator Δ_p can be viewed as a composition of slight variations of the Generational Distance (GD, see [29]) and the Inverted Generational Distance (IGD, see [3]). It is $\Delta_\infty = d_H$, but for finite values of p the indicator Δ_p averages the distances considered in d_H. Hence, as opposed to d_H, Δ_p does in particular punish single (or few) outliers in a candidate set to a lesser extent the smaller is p.

Note that $\mathsf{GD}_p$and $\mathsf{IGD}_p$are defined on discrete (or discretized) sets. This works well in practice, however, for theoretical purposes it is advantageous in certain cases to consider the complete set. While archives can considered to be discrete, this does not hold for Pareto fronts. Therefore, we introduce the indicators also for the continuous case in the next section in order to assess the numerical effects caused by discretizations.

3 Optimal Δ_p archives

An optimal archive A^* of m elements and an indicator $I : X^m \subset (\mathbb{R}^n)^m \to \mathbb{R}$, where w.l.o.g. lower values are preferred, is characterized by

$$A^* \in \arg \min_{\substack{A \subset X \\ |A|=m}} I(f(A)) \tag{1}$$

where A is a multiset in general and operator $\subset$ acts accordingly. An optimal archive with respect to (MOP), archive size m and indicator I can be viewed as a minimizer of a $m \cdot n$-dimensional scalar optimization problem.

In this section, we consider optimal archives for Δ_p and the indicator $\mathsf{L}\text{-}\mathsf{IGD}_p$ ("Least IGD_p under Zero GD_p") which is related to Δ_p and will be in particular used for convex PFs. In the following we consider linear and spherical fronts analytically and some other fronts with more complicated structure numerically. The results indicate that Δ_p indeed aims for uniformly distributed points along the Pareto front.

3.1 Indicators

We assume that the bi-objective PF can be represented by a continuous mapping that is parameterized by a scalar from a compact set.

[1] We note that $\Delta_p(\cdot, \cdot)$ is not a metric but an inframetric (a metric with relaxed triangle inequality) [27]. Keeping this fact in mind, we take the liberty to use the term *distance* in this context.

Definition 4 A continuous mapping $\varphi : [a, b] \to \mathbb{R}^2$ with $[a, b] \subset \mathbb{R}$ is said to be a *path* from $\varphi(a) \in \mathbb{R}^2$ to $\varphi(b) \in \mathbb{R}^2$. If the mapping is injective then the image of the path is termed a *Jordan arc*.

Thus, connected and closed Pareto fronts in two dimensions are Jordan arcs whereas disconnected Pareto fronts are assumed to be a collection of such arcs, each with its own parameter interval and continuous mapping.

3.1.1 Averaged Hausdorff distance Δ_p

Following (1), an optimal Δ_p archive with m elements solves

$$\min_{\substack{A \subset X \\ |A|=m}} \Delta_p(f(A), F^*). \tag{2}$$

Since here we deal with continuous Pareto fronts F^*, we have to adapt Δ_p to this context: let $Y = \{y_1, \ldots, y_m\} \subset f(X)$ where $y_1, \ldots, y_m$ are arranged in lexicographic order. First,

$$\mathsf{GD}_p(Y, F^*) = \left[\frac{1}{|Y|} \sum_{y \in Y} d(y, F^*)^p \right]^{\frac{1}{p}} = \left[\frac{1}{|Y|} \sum_{y \in Y} d(\varphi(s^*(y)), y)^p \right]^{\frac{1}{p}}, \tag{3}$$

where $s^*(y) = \arg\min\{d(\varphi(s), y) : s \in [a, b]\}$ is that scalar s representing the point $\varphi(s)$ on the front with smallest distance to given $y \in Y$. Evidently, we must be able to solve the optimization problem for $s^*(y)$ analytically to derive an elementary expression for GD_p. Second,

$$\mathsf{IGD}_p(Y, F^*) = \left[\frac{1}{L} \int_a^b d(\varphi(s), Y)^p \, ds \right]^{\frac{1}{p}} = \left[\frac{1}{L} \sum_{i=1}^{m} \int_{s_{i-1}}^{s_i} d(\varphi(s), y_i)^p \, ds \right]^{\frac{1}{p}}, \tag{4}$$

where $L := L(a, b)$ is the length of the Jordan arc and $s_i = \min\{s \in [a, b] : d(\varphi(s), y_{i+1}) < d(\varphi(s), y_i)\}$, for $i = 1, \ldots, m-1$, are 'switch points' with boundary points $s_0 = a$ and $s_m = b$. The switch points determine when the subsequent point in the candidate set Y becomes closer to the Pareto front than its predecessor. Clearly, unless the switch points can be derived analytically the development of an elementary expression for IGD_p becomes unlikely.

3.1.2 Least IGD_p *under zero* GD_p *(*L-IGD_p*)*

One characteristic of optimal Δ_p archives A^* is that the GD_p value does not have to be zero, i.e., images of elements of A^* do not have to lie on the PF. To see this, consider a strictly convex PF of a bi-objective problem with small and constant curvature. Then, the optimal IGD_∞ archive for $m = 1$ is given by the middle point of the line segment

that connects the two end points of the PF (if feasible) which is by strict convexity of the PF not Pareto optimal. For $p < \infty$, this point will move towards the PF, but will also not lie on it. Though the GD_p values for already relatively small values of m are typically reasonably small and the Δ_p approximation is determined by the IGD_p value (see Sect. 3.5 for such examples) a suggesting alternative to Δ_p could be to minimize the IGD_p value under the constraint that the elements are Pareto optimal. Using (1) the related optimization problem thus reads

$$\min_{\substack{A\subset X^*\\ |A|=m}} \mathsf{IGD}_p(f(A), F^*). \tag{5}$$

The following result provides some insights into the relation of the optimal archives of Δ_p and the indicator *Least* IGD_p *under Zero* $\mathsf{GD}_p(\mathsf{L\text{-}IGD}_p)$.

Theorem 1 (a) $\min\limits_{\substack{A\subset F\\ |A|=m}} \Delta_p(A, F^*) \le \min\limits_{\substack{A\subset F^*\\ |A|=m}} \mathsf{IGD}_p(A, F^*)$

(b) *Let the closure of* F^* *be bounded and* A_m *be a minimizer of (5) then*

$$\lim_{m\to\infty} \mathsf{IGD}_p(A_m, F^*) = 0.$$

Proof (a) Follows since $F^* \subset F$ and since $\mathsf{GD}_p(F^*, A) = 0$ for $A \subset F^*$.

(b) For $\epsilon > 0$ define $\delta := \epsilon(L(a,b)/(b-a))^{1/p}$, where a, b, and $L(a,b)$ are as in (4). Denote by $\bar{F^*}$ the closure of F^*, then

$$C_\delta := \bigcup_{y\in\bar{F^*}} B_\delta(y),$$

where $B_\delta(y) := \{x \in \mathbb{R}^k : d(x,y) < \delta\}$, is an open covering of $\bar{F^*}$. Since $\bar{F^*}$ is compact it follows by the Theorem of Heine-Borel that there exists a finite subcover S_δ of C_δ. That is, for every $y \in F^*$ there exists an element $s \in S_\delta$ such that $d(y,s) \le \delta$. Now let A_m be a minimizer of (5) where m is the finite number of open hyperballs of subcover S_δ, then

$$\begin{aligned}\mathsf{IGD}_p(A_m, F^*) \le \mathsf{IGD}_p(S_\delta, F^*) &= \left[\frac{1}{L(a,b)}\int_a^b d(\varphi(s), S_\delta)^p ds\right]^{1/p}\\ &\le \left[\frac{1}{L(a,b)}\int_a^b \delta^p ds\right]^{1/p} = \left[\frac{b-a}{L(a,b)}\frac{L(a,b)}{b-a}\cdot\epsilon^p\right]^{1/p} = \epsilon.\end{aligned}$$

□

By the above result it follows that (a) solutions of (5) can be viewed as upper bounds for optimal Δ_p archives, and (b) that these bounds are getting tight for increasing number of m. We will use this result in the following in particular for the investigation of optimal Δ_p archives for spherical convex PFs.

In case the Pareto front is connected and concave, it follows that an optimal archive of (5) is also an optimal Δ_p archive as the following result shows.

In [10], a similar result can be found in the context of level set approximations.

Theorem 2 *Let $k = 2$ and the PF F^* be connected and concave. Then for the optimal Δ_p archive A^* it holds that $\mathsf{GD}_p(A^*) = 0$.*

Proof Let A^* be an optimal Δ_p archive. Assume there exists $a \in A^*$ such that $y := F(a) \notin F^*$. Let s_a and s_b be the switch points in (4) such that the IGD value of A^* in $\varphi([s_a, s_b])$ is determined by y, and let $v_1 := \varphi(s_a) - \varphi(s_b)$. Since $\varphi(s_a)$ and $\varphi(s_b)$ are nondominated we can choose $v_2 \in \mathbb{R}^2$ s.t. it is orthogonal to v_1 and both of its entries are negative. The pair $\{v_1, v_2\}$ forms an orthogonal basis of the $\mathbb{R}^2$. Let l_{v_2} be the line through y along v_2 and p_f be the orthogonal projection of some $f \in F^*$ onto l_{v_2}. Since F^* is concave and connected there exists an $\epsilon > 0$ s.t. $y_\epsilon := y + \epsilon v_2 \in F^*$. It is

$$\|f - y\|_2^2 = \|f - p_f\|_2^2 + \|p_f - y\|_2^2.$$

By concavity of F^* it is either $p_f \in F^*$ or p_f is not feasible, and thus

$$\|f - y_\epsilon\|_2^2 = \|f - p_f\|_2^2 + \|p_f - y_\epsilon\|_2^2,$$

where $\|p_f - y_\epsilon\|_2^2 < \|p_f - y\|_2^2$ (since $\epsilon > 0$) from which it follows that $\|f - y_\epsilon\|_2 < \|f - y\|_2$. Now let $b \in X^*$ s.t. $F(b) = y_\epsilon$. Then, for $\tilde{A} := A^* \cup \{b\} \setminus \{a\}$ we have constructed an archive that is better than A^* w.r.t. both GD_p and IGD_p, a contradiction to the optimality of A^*. □

3.2 Optimal archives for linear Pareto fronts

In the following we show that the elements of the optimal Δ_p archives are evenly spread around the PF if it is linear and connected.

Theorem 3 *Suppose the PF of a bi-objective problem is linear and connected with end points $(m_1, M_2)^T$ and $(M_1, m_2)^T$, where $m_i < M_i$, $i = 1, 2$. Then the optimal Δ_p archive with m elements is given by $Y = \{y_1, \ldots, y_m\}$, where*

$$y_i = \begin{pmatrix} m_1 \\ M_2 \end{pmatrix} + \frac{t_i - m_1}{M_1 - m_1} \begin{pmatrix} M_1 - m_1 \\ m_2 - M_2 \end{pmatrix}, i = 1, \ldots, m,$$

and

$$t_i = m_1 + \left(\frac{1}{2n} + \frac{i-1}{n} \right) (M_1 - m_1), i = 1, \ldots, m. \tag{6}$$

Proof We present here the proof for $p = 1$. The analysis for general value of p is analog but more lengthy.

The Pareto front can be expressed by the curve $\gamma : [m_1, M_1] \to \mathbb{R}^2$,

$$\gamma(t) = \begin{pmatrix} m_1 \\ M_1 \end{pmatrix} + \frac{t - m_1}{M_1 - m_1} D_F,$$

where $D_F = (M_1 - m_1, m_2 - M_2)^T$. By Theorem 2 we can assume that $Y \subset F^*$, i.e., there exist values $t_i \in [m_1, M_1]$ s.t. $y_i = \gamma(t_i), i = 1, \ldots, m$. Further, assume that the y_i's are ordered according to the first objective value. Then the values of $t \in [m_1, M_1]$ s.t. a point $y = \gamma(t)$ on the Pareto front is closest to y_i is given by

$$\begin{aligned} t &\in [m_1, \xi_1] \quad \text{for } y_1 \\ t &\in [\xi_{i-1}, \xi_i] \quad \text{for } y_i,\ i = 2, \ldots, m-1 \\ t &\in [\xi_{n-1}, M_1] \quad \text{for } y_m, \end{aligned}$$

where $\xi_i = (t_i + t_{i+1})/2$, and we obtain

$$\begin{aligned} \mathsf{IGD}_1(Y, F^*) &= \frac{1}{L(m_1, M_1)} \int_{m_1}^{M_1} d(\gamma(t), Y) dt \\ &= \left[\int_{m_1}^{\xi_1} \|\gamma(t) - y_1\|_2 dt + \sum_{i=2}^{m-1} \int_{\xi_{i-1}}^{\xi_i} \|\gamma(t) - y_i\|_2 dt \right. \\ &\quad \left. + \int_{\xi_{m-1}}^{M_1} \|\gamma(t) - y_m\|_2 dt \right] \cdot \frac{1}{L(m_1, M_1)} \\ &= \frac{\|D_F\|_2}{2\, L(m_1, M_1)} \underbrace{\left[(t_1 - m_1)^2 + \frac{1}{2} \sum_{i=1}^{m-1} (t_{i+1} - t_i)^2 + (M_1 - t_m)^2 \right]}_{=: g(t_1, \ldots, t_m)} \end{aligned}$$

The value of IGD_1 is thus minimal if the value of g is minimal. Setting the gradient of g to zero leads to the linear system of equations

$$\begin{pmatrix} 3 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 3 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_{m-1} \\ t_m \end{pmatrix} = \begin{pmatrix} 2m_1 \\ 0 \\ \vdots \\ 0 \\ 2M_1 \end{pmatrix} \tag{7}$$

where the square matrix is just the Hessian matrix of the quadratic function g. Notice that the tridiagonal Hessian matrix is regular since it is irreducibly diagonally dominant[2] [14]. Moreover, weak diagonal dominance and positive diagonal entries imply positive semi-definiteness (p.s.d.) of the Hessian matrix. But a regular p.s.d. matrix

[2] A tridiagonal matrix is irreducibly diagonally dominant if it is weakly diagonal dominant, both off-diagonals have nonzero entries, and there exists at least one row that is strictly diagonal dominant.

must be positive definite which implies strict convexity of g. Therefore g has a unique local minimum which is the global minimum.

It is straightforward to see that the choice of t_i as in (6) is the only way to solve (7). Since this solution is in the interior of $[m_1, M_1]^m$ it is also the global solution of g restricted to $[m_1, M_1]^m$ which leads to the claim. □

We note that in [10] a similar result for a variant of Δ_p has been shown in the context of level set approximations.

The problem to find the optimal archives apparently gets more complicated when the Pareto front is not linear and/or not connected. Note, however, that each part of the locally differentiable Pareto front gets linear with increasing zoom into its neighborhood. Thus, for instance, the 'non-optimality' of the Δ_p archives as discussed in Sect. 3.1.2 vanishes with increasing m.

3.3 Optimal archives for spherical concave Pareto fronts

Suppose the PF is given by $F^* = \{(\sin\omega, \cos\omega)' : \omega \in [0, \pi/2]\}$, i.e., the quarter of the unit circle in the positive quadrant. By Theorem 2 we can restrict ourselves to the consideration of the L-IGD$_p$ value which we adapt now for the current context: the archive of size $m \in \mathbb{N}$ is given by $A = \{\varphi(\alpha_i) : i = 1, \ldots, m\}$, where $\varphi(\omega) = (\sin(\omega), \cos(\omega))'$ for $\omega \in [0, \pi/2]$. All points $\varphi(\omega)$ with angle $\omega \in [0, \alpha_1]$ have closest distance to $\varphi(\alpha_1)$. In addition, all points $\varphi(\omega)$ with angle $\omega > \alpha_1$ up to angle $(\alpha_1 + \alpha_2)/2$, which exactly divides the angle between α_1 and α_2, have closest distance to $\varphi(\alpha_1)$. In an analogous manner all points with angles between $(\alpha_{i-1} + \alpha_i)/2$ and $(\alpha_i + \alpha_{i+1})/2$ for $i = 2, \ldots, m-1$ are closest to $\varphi(\alpha_i)$. Thus, we can express the IGD$_p$ by

$$\mathsf{IGD}_p(\alpha) = \left[\frac{1}{L(0,\pi/2)} \sum_{i=1}^{m} \int_{\frac{\alpha_{i-1}+\alpha_i}{2}}^{\frac{\alpha_i+\alpha_{i+1}}{2}} d(\varphi(\omega), \varphi(\alpha_i))^p \, d\omega \right]^{\frac{1}{p}} \tag{8}$$

with $\alpha_0 := -\alpha_1$, $\alpha_{m+1} := \pi - \alpha_m$ and $0 \le \alpha_1 \le \ldots \le \alpha_m \le \pi/2$.

Theorem 4 *Let $F^* = \{(\sin\omega, \cos\omega)' : \omega \in [0, \pi/2]\}$ and $m \in \mathbb{N}$ be the number of points on the Pareto front. If $p = 2$ and $d(x, y) = \|x - y\|_2$ then*

$$\alpha_i^* = (2i - 1)\frac{\pi}{4m} \quad \text{for } i = 1, \ldots, m$$

are the optimal angles such that $y_i = (\sin\alpha_i^*, \cos\alpha_i^*)'$ for $i = 1, \ldots, m$ minimize Δ_p. In this case,

$$\Delta_2 = \mathsf{IGD}_2 = \sqrt{2 - 2 \cdot \frac{4m}{\pi} \sin\left(\frac{\pi}{4m}\right)} \le \frac{\pi}{4\sqrt{3}} \cdot \frac{1}{m}.$$

Proof If $p = 2$ Eq. (8) can be expressed via $\mathsf{IGD}_2(\alpha) = \sqrt{2\,g(\alpha)/\pi}$ where

$$
\begin{aligned}
g(\alpha) &= \sum_{i=1}^{m} \int_{\frac{\alpha_{i-1}+\alpha_i}{2}}^{\frac{\alpha_i+\alpha_{i+1}}{2}} d\left(\begin{pmatrix}\sin\omega\\ \cos\omega\end{pmatrix}, \begin{pmatrix}\sin\alpha_i\\ \cos\alpha_i\end{pmatrix}\right)^2 d\omega \\
&= \sum_{i=1}^{m} \int_{\frac{\alpha_{i-1}+\alpha_i}{2}}^{\frac{\alpha_i+\alpha_{i+1}}{2}} \left[(\sin\omega - \sin\alpha_i)^2 + (\cos\omega - \cos\alpha_i)^2\right] d\omega \\
&= \sum_{i=1}^{m} \int_{\frac{\alpha_{i-1}+\alpha_i}{2}}^{\frac{\alpha_i+\alpha_{i+1}}{2}} \left[2 - 2\,\underbrace{(\cos\omega\cdot\cos\alpha_i + \sin\omega\cdot\sin\alpha_i)}_{\cos(\omega-\alpha_i)}\right] d\omega \\
&= 2\sum_{i=1}^{m} \left[\omega - \sin(\omega-\alpha_i)\right]_{\frac{\alpha_{i-1}+\alpha_i}{2}}^{\frac{\alpha_i+\alpha_{i+1}}{2}} \\
&= \pi - 2\sum_{i=1}^{m} \left[\sin\left(\frac{\alpha_i-\alpha_{i-1}}{2}\right) + \sin\left(\frac{\alpha_{i+1}-\alpha_i}{2}\right)\right] \\
&= \pi - 2\left[\sin\alpha_1 + 2\sum_{i=1}^{m-1} \sin\left(\frac{\alpha_{i+1}-\alpha_i}{2}\right) + \sin\left(\frac{\pi}{2}-\alpha_m\right)\right] \qquad (9)
\end{aligned}
$$

Partial differentiation of (9) yields the necessary conditions

$$\frac{\partial g(\alpha)}{\partial \alpha_1} = 2\left[\cos\left(\frac{\alpha_2-\alpha_1}{2}\right) - \cos\alpha_1\right] \stackrel{!}{=} 0 \qquad (10)$$

$$\frac{\partial g(\alpha)}{\partial \alpha_i} = 2\left[\cos\left(\frac{\alpha_{i+1}-\alpha_i}{2}\right) - \cos\left(\frac{\alpha_i-\alpha_{i-1}}{2}\right)\right] \stackrel{!}{=} 0 \qquad (11)$$

$$\frac{\partial g(\alpha)}{\partial \alpha_m} = 2\left[\cos\left(\frac{\pi}{2}-\alpha_m\right) - \cos\left(\frac{\alpha_m-\alpha_{m-1}}{2}\right)\right] \stackrel{!}{=} 0 \qquad (12)$$

with $i = 2, \ldots, m-1$ in (11). Condition (10) leads to $\alpha_2^* = 3\,\alpha_1^*$ whereas condition (11) immediately reveals that consecutive angles must have the same distance. Moreover, we may derive the recurrence equation $\alpha_{i+1}^* = 2\,\alpha_i^* - \alpha_{i-1}^*$ for $i = 2, \ldots, m-1$. This recurrence equation together with the initial condition $\alpha_2^* = 3\,\alpha_1^*$ has the solution

$$\alpha_i^* = (2\,i-1)\,\alpha_1^* \qquad (13)$$

for $i = 1, \ldots, m$. It remains to identify α_1^*. Condition (12) may be expressed via $\alpha_m^* = (\pi + \alpha_{m-1}^*)/3$. Replacement of α_m^* and α_{m-1}^* with the expressions from (13) finally delivers $\alpha_1^* = \pi/(4\,m)$.

Since $\alpha_{i+1}^* - \alpha_i^* = 2\alpha_1^*$ for $i = 1, \ldots, m-1$ and $\alpha_m^* = \pi/2 - \alpha_1^*$ we obtain $g(\alpha^*) = \pi - 4m \cdot \sin\left(\frac{\pi}{4m}\right)$ after insertion of the optimal angles in (9).

We still have to show that $g(\alpha^*)$ is indeed a local/global minimum. Partial derivation of (10) to (12) and insertion of α^* leads to the tridiagonal Hessian matrix

$$\begin{pmatrix} 3\sin\frac{\pi}{4m} & -\sin\frac{\pi}{4m} & & & 0 \\ -\sin\frac{\pi}{4m} & 2\sin\frac{\pi}{4m} & -\sin\frac{\pi}{4m} & & \\ & \ddots & \ddots & \ddots & \\ & & -\sin\frac{\pi}{4m} & 2\sin\frac{\pi}{4m} & -\sin\frac{\pi}{4m} \\ 0 & & & -\sin\frac{\pi}{4m} & 3\sin\frac{\pi}{4m} \end{pmatrix}$$

which is irreducibly diagonally dominant and since the diagonal entries are positive it follows that the Hessian matrix is positive definite implying that α^* is a local minimum.

Finally, for assuring that the local minimum is the global minimum we have to show that solutions on the border have a larger value. Recall that we insisted on the equalities $0 \le \alpha_1 \le \alpha_2 \le \ldots \le \alpha_m \le \pi/2$. As a consequence, solutions on the border must have at least $\alpha_1 = 0$ or $\alpha_m = \pi/2$. If $\alpha_1 = 0$ then successive angles (with increasing index) may be zero and if $\alpha_m = \pi/2$ then preceding angles (with decreasing index) may have the value $\pi/2$. A closer look at (9) reveals that only three cases are of interest: 1. $\alpha_1 = 0$, 2. $\alpha_m = \pi/2$, and 3. $\alpha_1 = 0 \wedge \alpha_m = \pi/2$. If further angles are placed on the border (as described previously) they do not contribute any value to the sum in (9), only the dimensionality of the problem decreases (leading to larger IGD_2 resp. Δ_2 values [27]).

Notice that the first case is analogous to the second case for symmetry reasons. Therefore, we begin with the second case by setting $\alpha_m = \pi/2$ and insisting that all other angles are in the open interval $(0, \pi/2)$. Insertion in (9) and subsequent partial differentiation leads to the necessary conditions (10), (11) for $i = 1, \ldots, m-2$ and

$$\frac{\partial g(\alpha)}{\partial \alpha_{m-1}} = 2\left[\cos\left(\frac{\pi}{4} - \frac{\alpha_{m-1}}{2}\right) - \cos\left(\frac{\alpha_{m-1} - \alpha_{m-2}}{2}\right)\right] \stackrel{!}{=} 0. \tag{14}$$

With the same argumentation leading to the solution of (10)–(12) we find optimal $\alpha_1^+ = \pi/(4m-2)$ and $\alpha_i^+ = (2i-1)\,\alpha_1^+$ for $i = 1, \ldots, m-1$ with fixed $\alpha_m^+ = \pi/2$. Insertion leads to

$$\mathsf{IGD}_2(\alpha^+) = \sqrt{2 - 2 \cdot \frac{4m-2}{\pi} \sin\left(\frac{\pi}{4m-2}\right)} > \mathsf{IGD}_2(\alpha^*)$$

since $\sin(x)/x$ is positive and strictly decreasing on $(0, 1)$. As for the third case, set $\alpha_1 = 0 \wedge \alpha_m = \pi/2$ and proceed as before. We obtain optimal $\tilde{\alpha}_2 = \pi/(2m-2)$ and

$\tilde{\alpha}_i = (i-1)\,\tilde{\alpha}_2$ for $i = 2, \ldots, m-1$ with fixed $\tilde{\alpha}_1 = 0$ resp. $\tilde{\alpha}_m = \pi/2$ leading to

$$\mathsf{IGD}_2(\tilde{\alpha}) = \sqrt{2 - 2 \cdot \frac{4m-4}{\pi} \sin\left(\frac{\pi}{4m-4}\right)} \; > \; \mathsf{IGD}_2(\alpha^*).$$

As a consequence, we have shown that the local minimum is indeed the global minimum of IGD_2 and it is located in the interior of $[0, \pi/2]^m$. The upper bound for the IGD_2 easily follows from the inequality $\sin(x) \geq x - x^3/6$ being valid for $0 \leq x \leq 1$. □

Theorem 5 *Let $F^* = \{(\sin\omega, \cos\omega)' : \omega \in [0, \pi/2]\}$ and $m \in \mathbb{N}$ be the number of points on the Pareto front. If $p = 1$ and $d(x, y) = \|x - y\|_2$ then*

$$\alpha_i^* = (2i-1)\,\frac{\pi}{4m} \quad \textit{for } i = 1, \ldots, m$$

are the optimal angles such that $y_i = (\sin\alpha_i^*, \cos\alpha_i^*)'$ for $i = 1, \ldots, m$ minimize Δ_p. In this case,

$$\Delta_1 = \mathsf{IGD}_1 = 2 \cdot \frac{8m}{\pi}\left[1 - \cos\left(\frac{\pi}{8m}\right)\right] \leq \frac{\pi}{8} \cdot \frac{1}{m}.$$

Proof If $p = 1$ Eq. (8) can be expressed via $\mathsf{IGD}_1(\alpha) = 2\,g(\alpha)/\pi$ where

$$\begin{aligned}
g(\alpha) &= \sum_{i=1}^{m} \int_{\frac{\alpha_{i-1}+\alpha_i}{2}}^{\frac{\alpha_i+\alpha_{i+1}}{2}} d\left(\begin{pmatrix}\sin\omega\\ \cos\omega\end{pmatrix}, \begin{pmatrix}\sin\alpha_i\\ \cos\alpha_i\end{pmatrix}\right) d\omega \\
&= \sum_{i=1}^{m} \int_{\frac{\alpha_{i-1}+\alpha_i}{2}}^{\frac{\alpha_i+\alpha_{i+1}}{2}} \sqrt{(\sin\omega - \sin\alpha_i)^2 + (\cos\omega - \cos\alpha_i)^2}\, d\omega \\
&= 2\sum_{i=1}^{m} \int_{\frac{\alpha_{i-1}+\alpha_i}{2}}^{\frac{\alpha_i+\alpha_{i+1}}{2}} \sqrt{\frac{1 - \cos(\omega - \alpha_i)}{2}}\, d\omega \\
&= 2\sum_{i=1}^{m} \int_{\frac{\alpha_{i-1}+\alpha_i}{2}}^{\alpha_i} \sin\left(\frac{\alpha_i - \omega}{2}\right) d\omega + 2\sum_{i=1}^{m} \int_{\alpha_i}^{\frac{\alpha_i+\alpha_{i+1}}{2}} \sin\left(\frac{\omega - \alpha_i}{2}\right) d\omega \\
&= 4\sum_{i=1}^{m}\left[1 - \cos\left(\frac{\alpha_i - \alpha_{i-1}}{4}\right)\right] + 4\sum_{i=1}^{m}\left[1 - \cos\left(\frac{\alpha_{i+1} - \alpha_i}{4}\right)\right] \\
&= 8m - 4\left[\cos\left(\frac{\alpha_1}{2}\right) + 2\sum_{i=1}^{m-1}\cos\left(\frac{\alpha_{i+1} - \alpha_i}{4}\right) + \cos\left(\frac{\pi}{4} - \frac{\alpha_m}{2}\right)\right] \qquad (15)
\end{aligned}$$

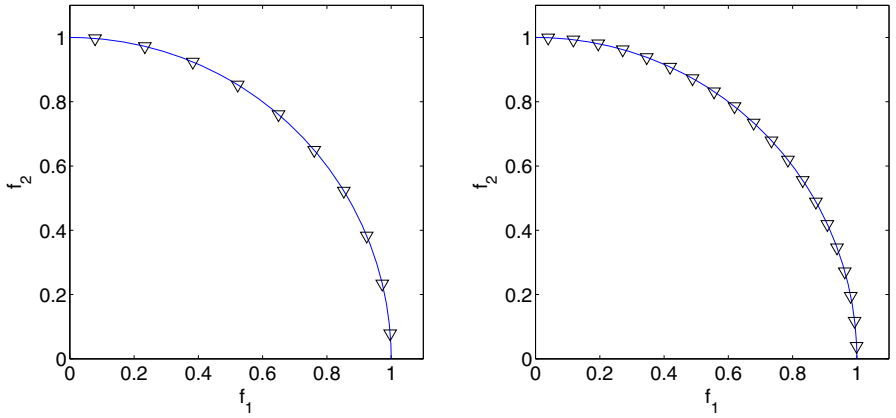

Fig. 1 Optimal Δ_p archives for $p \in \{1, 2\}$, $m = 10$ (*left*) and $m = 20$ (*right*) for a concave spherical front

Partial differentiation of (15) yields the necessary conditions

$$\frac{\partial g(\alpha)}{\partial \alpha_1} = 2\left[\sin\left(\frac{\alpha_1}{2}\right) - \sin\left(\frac{\alpha_2 - \alpha_1}{4}\right)\right] \stackrel{!}{=} 0 \tag{16}$$

$$\frac{\partial g(\alpha)}{\partial \alpha_i} = 2\left[\sin\left(\frac{\alpha_i - \alpha_{i-1}}{4}\right) - \sin\left(\frac{\alpha_{i+1} - \alpha_i}{4}\right)\right] \stackrel{!}{=} 0 \tag{17}$$

$$\frac{\partial g(\alpha)}{\partial \alpha_m} = 2\left[\sin\left(\frac{\alpha_m - \alpha_{m-1}}{4}\right) - \sin\left(\frac{\pi}{4} - \frac{\alpha_m}{2}\right)\right] \stackrel{!}{=} 0 \tag{18}$$

with $i = 2, \ldots, m-1$ in (17). Condition (16) leads to $\alpha_2^* = 3\,\alpha_1^*$ and it is easy to see that conditions (17) and (18) lead to the same optimal angles as for the case with $p = 2$. Insertion of the optimal angles in (15) delivers the result stated in the theorem. With the same argumentation as in the previous proof it can be shown that these angles lead to a unique local and therefore global minimum. The upper bound for the IGD_1 easily follows from the inequality $1 - \cos(x) \le x^2/2$ being valid for $0 \le x \le 1$. □

Figure 1 shows the images of the optimal Δ_p archives for the front considered in Theorems 4 and 5 for $p \in \{1, 2\}$ and for archive size $m \in \{10, 20\}$. Evidently, the images are evenly spread along the front.

The theory developed so far acts on the assumption that IGD_p and GD_p are determined for continuous Pareto fronts with innumerable many points. In practice such a calculation is typically not possible and we have to switch to finite discretizations of the innumerable front. Therefore we are interested in the *discretization error* defined by the difference of the indicator value $\Delta_p = \max\{\mathsf{IGD}_p, \mathsf{GD}_p\}$ for the innumerable and the finite Pareto front.

For this purpose we placed $N \in \mathbb{N}$ points on the concave Pareto front considered in this section by using N equidistant angles $\omega \in [0, \pi/2]$ using the parametric representation $\varphi(\omega) = (\sin\omega, \cos\omega)'$. Table 1 displays Δ_p values using a discretized Pareto front for optimal archives of size $m \in \{10, 100\}$ and sample sizes

Table 1 Comparison of Δ_p values with $p = 1$ and $p = 2$ for varying archive size m and increasing number of samples N in the discretized Pareto front versus the optimal value for the continuous Pareto front ($N = \infty$)

N	$m = 10$		$m = 100$	
	$p = 1$	$p = 2$	$p = 1$	$p = 2$
∞	0.0392649	0.0453380	0.00392699	0.00453449
10^6	0.0392649	0.0453380	0.00392699	0.00453450
10^5	0.0392653	0.0453384	0.00392703	0.00453454
10^4	0.0392688	0.0453425	0.00392738	0.00453494
10^3	0.0393042	0.0453833	0.00393092	0.00453903

$N \in \{10^3, 10^4, 10^5, 10^6\}$ for $p \in \{1, 2\}$ in comparison to the true innumerable Pareto front (indicated by sample size $N = \infty$). Apparently, even a moderate granularity of the discretization delivers a close approximation to the true Δ_p value.

3.4 Optimal archives for spherical convex Pareto fronts

Now consider that $F^* = \{(1-\cos\omega, 1-\sin\omega)' : \omega \in [0, \pi/2]\}$ is a convex spherical front. Due to the discussion in Sect. 3.1.2 elements of the optimal Δ_p archive will not be Pareto optimal. Instead of deriving the optimal archives for Δ_p we can at least present the respective ones for $\mathsf{L\text{-}IGD}_p$ which follows by the above subsection and which can be seen as good approximations of the Δ_p archives (cf. Theorem 1).

Corollary 1 *Let $F^* = \{(1-\cos\omega, 1-\sin\omega)' : \omega \in [0, \pi/2]\}$ and $m \in \mathbb{N}$ be the number of points on the Pareto front. If $p \in \{1, 2\}$ and $d(x, y) = \|x - y\|_2$ then*

$$\alpha_i^* = (2\,i - 1)\,\frac{\pi}{4\,m} \quad \textit{for } i = 1, \ldots, m$$

are the optimal angles such that $y_i = (1 - \cos\alpha_i^*, 1 - \sin\alpha_i^*)'$ for $i = 1, \ldots, m$ minimize IGD_p under zero GD_p. In this case,

$$\Delta_1 \;\le\; \mathsf{L\text{-}IGD}_1 = 2 \cdot \frac{8\,m}{\pi}\left[1 - \cos\left(\frac{\pi}{8\,m}\right)\right] \;\le\; \frac{\pi}{8} \cdot \frac{1}{m},$$

$$\Delta_2 \;\le\; \mathsf{L\text{-}IGD}_2 = \sqrt{2 - 2 \cdot \frac{4\,m}{\pi}\,\sin\left(\frac{\pi}{4\,m}\right)} \;\le\; \frac{\pi}{4\sqrt{3}} \cdot \frac{1}{m}.$$

Proof Follows from the inequality provided by Theorem 1(a) and by the proofs of Theorems 4 and 5 since the concave PF is a reflection of the convex one, which does not affect the distances. □

The fact that the inequalities may be strict can be illustrated as follows: let $\tilde{A}^*_{\Delta_1}$ denote the archive obtained by numerically solving optimization problem (2) with respect to Δ_1 (i.e., the numerically optimal Δ_1 archive).

Δ_1 archive
(0.0033483782, 0.9215673964)’
(0.0278878575, 0.7666225241)’
(0.0763640789, 0.6174241462)’
(0.1475834114, 0.4776455857)’
(0.2397960730, 0.3507282266)’
(0.3507288039, 0.2397954934)’
(0.4776455702, 0.1475834074)’
(0.6174241262, 0.0763640741)’
(0.7666225340, 0.0278878672)’
(0.9215674026, 0.0033483728)’

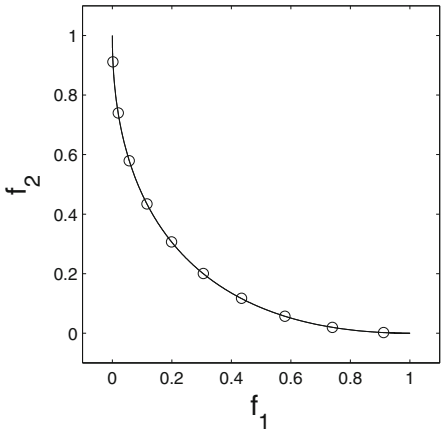

Fig. 2 Numerically optimal Δ_1 archive $f(\tilde{A}^*_{\Delta_1})$ of size $m = 10$ for the spherically convex Pareto front with nonzero GD_1 (*left* values of archive points; *right* graphical representation)

The Δ_1 distance between this archive for $m = 10$ archive points (see Fig. 2) and the true Pareto front $f(X^*)$ is

$$\Delta_1(f(\tilde{A}^*_{\Delta_1}),\, f(X^*)) = 0.03926277,$$

which is equal to $\mathsf{IGD}_1(f(\tilde{A}^*_{\Delta_1}),\, f(X^*))$, with corresponding

$$\mathsf{GD}_1(f(\tilde{A}^*_{\Delta_1}),\, f(X^*)) = 0.000266746 > 0,$$

whereas $\mathsf{L\text{-}IGD}_1(f(A^*_{\mathsf{L\text{-}IGD}_1},\, f(X^*)) = 0.03926486$ for the optimal $\mathsf{L\text{-}IGD}_1$ archive of Corollary 1. Moreover, although the optimal Δ_1 archive is not on the front, this deviation cannot be recognized visually from Fig. 2 since the GD_1 value is that small.

3.5 Optimal archives for nonlinear and disconnected Pareto fronts

In the preceding section we observed that the GD_1 value is about two orders of magnitude smaller than the IGD_1 value for the numerically optimal Δ_1 archive. This behavior can be noticed also for general nonlinear and disconnected Pareto fronts. For illustration we consider test problems DENT [30] with connected convex/concave PF and ZDT3 [33] with nonlinear and disconnected PF. Figure 3 shows the numerically optimal Δ_1 archives of size $m = 10$ obtained by numerically solving problem (2). Again, it cannot be detected visually from Fig. 3 that the archive points do not lie on the PF but the solutions appear to be evenly spread along the Pareto fronts.

Table 2 reveals that the GD_1 value is nonzero but about three orders of magnitude smaller than the IGD_1 value indicating that the optimal Δ_1 archive is close to the true Pareto front and that the auxiliary problem of minimizing $\mathsf{L\text{-}IGD}_1$ would deliver a good approximation.

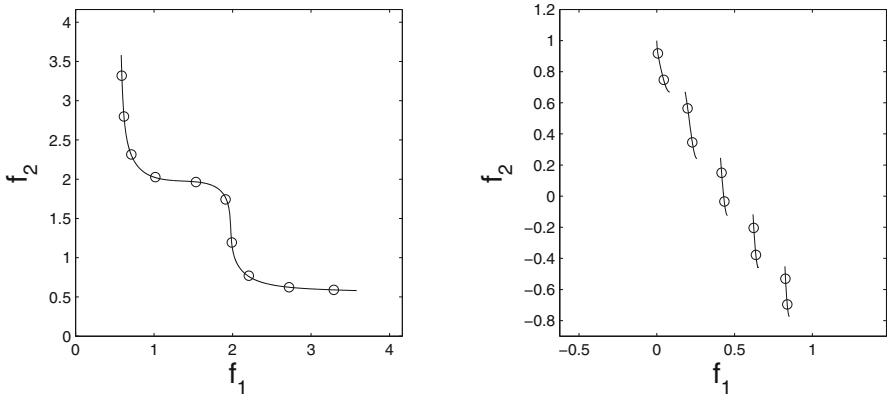

Fig. 3 Numerically optimal Δ_1 archives for $m = 10$ for the Pareto fronts of test problems DENT (*left*) and ZDT3 (*right*)

Table 2 Indicator values for numerically optimal Δ_1 archives for test problems DENT and ZDT3

Problem	$\Delta_1(f(\tilde{A}^*_{\Delta_1}), F^*)$	$\mathsf{IGD}_1(f(\tilde{A}^*_{\Delta_1}), F^*)$	$\mathsf{GD}_1(f(\tilde{A}^*_{\Delta_1}), F^*)$
DENT	0.290388	0.290388	0.000440
ZDT3	0.109733	0.109733	0.000701

It remains to answer the question about the difference between an optimal Δ_1 archive and an optimal $\mathsf{L\text{-}IGD}_1$ archive. Even for the relatively small archive size of $m = 10$ the Hausdorff distance between the archive numerically optimal with respect to Δ_1 and the archive numerically optimal with respect to $\mathsf{L\text{-}IGD}_1$ is $d_H(f(\tilde{A}^*_{\Delta_1}), f(\tilde{A}^*_{\mathsf{L\text{-}IGD}_1})) = 0.015208$ for DENT and 0.006062 for ZDT3. For larger archive sizes the distance gets smaller leading to the conclusion that for practically reasonable archive sizes $m \geq 100$ the differences between optimal Δ_1 and optimal $\mathsf{L\text{-}IGD}_1$ archive become negligible.

4 Specific algorithms

In this section, algorithmic approaches are presented which aim at approximating the Δ_p optimal archives. The main idea is to combine evolutionary multiobjective optimization techniques with specific postprocessing strategies applied to the archive of all solutions generated during the optimization run.

4.1 Concepts of PS-EMOA

PSEMOA is an evolutionary algorithm proposed in [7] that uses PSA as the tool to select the best individuals during the evolution of the search.

4.1.1 Part and selection algorithm (PSA)

The PSA has been recently introduced in [26] as an algorithm for selecting m 'well-spread' points out of a set of $n > m$ points. PSA consists of two steps: (1) the set of interest is partitioned into m subsets in order to group similar members into the same subset, and (2) selection of one representative member from each generated subset in order to get a diversed subset of m points of the set of interest. A more detailed description is as follows:

Partitioning a set This is based on the concept of the *dissimilarity* of a set:

Let $A = \{\mathbf{f}_1 = (f_{11}, \ldots, f_{1k})', \ldots, \mathbf{f}_n = (f_{n1}, \ldots, f_{nk})'\}$ (i.e., n objective vectors $\mathbf{f}_i \in \mathbb{R}^k$) and $a_j = \min\{f_{ij} : i = 1, \ldots, n\}$, $b_j = \max\{f_{ij} : i = 1, \ldots, n\}$ for $j = 1, \ldots, k$, then $diss(A) = \max\{b_j - a_j : j = 1, \ldots, k\}$ measures the *dissimilarity* of set A.

In each step, the set with the greatest dissimilarity among its members is the one that is divided. This is repeated m times to obtain the desired number of subsets. Figure 4 demonstrates the steps of partitioning the set of interest into $m = 5$ subsets.

Selection of points Once the set A has been partitioned into the m subsets $A_1, \ldots, A_m$, the 'most suitable' element from each subset must then be chosen in order to obtain a subset $A_{(r)}$ of A that contains m elements. This is of course problem dependent. Figure 4 (bottom right) illustrates the rule of choosing the nearest points to the virtual centers. The centers of the gray rectangles are marked with a cross. In each subset the member closest to the center is circled (a random member is circled in the subset with only two members). The representative set $A_{(r)} = \{a_1, a_2, a_3, a_4, a_5\}$ is the set of all circled points.

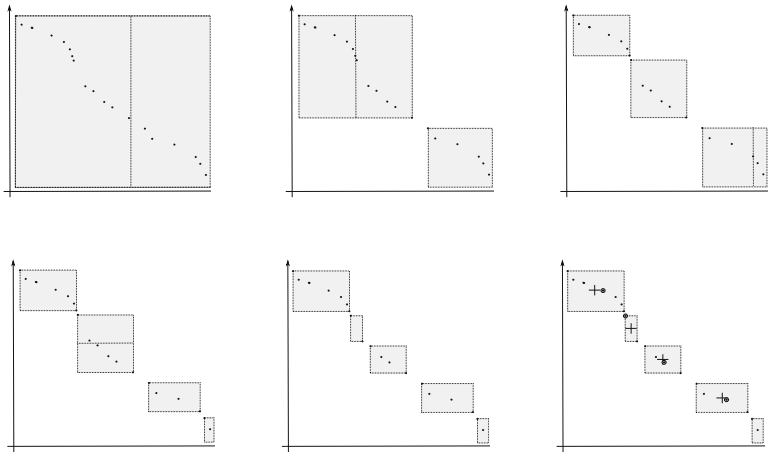

Fig. 4 Partitioning of 24 elements in bi-objective space into 5 subsets (*gray boxes*) and selection of the subset $A_{(r)}$ using the center points (*bottom right*) (borrowed from [26])

4.1.2 PSEMOA

The steps of PSEMOA are as follows: First, a population is chosen at random. When a new offspring out of the population is created via crossover and mutation during the evolution, the set A is defined, which is the set of nondominated points of the union of the current population and the offspring. Since PSA preserves any outliers encountered in the search, a strategy to delete these points is needed (the maximum values per objective are deleted from A). If the size of the set of nondominated points free of outliers is less than the maximum population size, then a nondominated sorting of A is performed [6], and the next generation is selected according to this sorting. If the size of A is bigger than the maximum population size, then PSA selects μ points out of A as the next generation. In order to increase the quality of the approximations found by PSEMOA, the Δ_p external archive strategy and the offline version defined in [7] are used. The offline version allows PSA to have much more information to generate a good reference set needed by the Δ_p external archive, which will guide the external archive to better results. The pseudo code of PSEMOA can be found in Algorithm 1.

4.2 Concepts of Δ_p-EMOA

Suppose that some EMOA has generated an approximation of the Pareto front for some MOP. Typically, this approximation does not yield a finite point set in objective space that is evenly distributed. Therefore, we propose the following *offline* approach:

1. Run your favorite EMOA that is equipped with a tiny add-on:
   - as soon as an offspring is generated and evaluated store a copy in a file
2. After termination of your favorite EMOA:
   - construct an evenly spaced reference front from a given approximation of the Pareto front (e.g. from last population of favorite EMOA)
   - feed each stored offspring into the Δ_p archive updater (see Algorithm 2) sequentially
   - output: archive A

Here, the reference front is constructed as follows: calculate a linear interpolation from a given approximation of the Pareto front. Since the length L of the resulting polygonal line can be derived easily, division by the number of desired archive points m yields the size of the equal spacing $\delta = L/m$ which is used to place m points along the polygonal line with distance δ, where the extreme points of the discretization are moved half the length inwards the polygonal line. After the reference front R has been constructed it is used in the Δ_p archive updater to decide which point should be added to or deleted from the archive. An update operation can be realized as sketched in Algorithm 2.

This naive approach takes $\Theta(|A| \cdot (|A| \cdot |R| \cdot d))$ time units, whereas the quick update version below (Algorithm 3) only needs $\Theta(|A| \cdot |R| \cdot d)$ time units: calculating $d(a, R)$ takes $\Theta(|R| \cdot d)$ whereas calculating $d(r, A)$ takes $\Theta(|A| \cdot d)$ time. Therefore the first loop needs $\Theta(|A| \cdot |R| \cdot d)$ time, the second loop $\Theta(|R| \cdot |A| \cdot d)$ time, and

Algorithm 1 PSEMOA [7] with Δ_p external archive and *offline* approach.

choose $maxEval$ = maximum number of evaluations and $\mu \in \mathbb{N}$
set $t = 0$, $cntEval = 0$
draw multiset P_0 with μ elements $\in \mathbb{R}^n$ at random.
while $cntEval < maxEval$ **do**
  generate multiset Q with μ offspring $\in \mathbb{R}^n$ from P_t by variation
  save/append Q to an external file
  $cntEval = cntEval + \mu$
  $A = \mathsf{ND}_f(P_t \cup Q, \preceq)$
  if t is even **then**
    $A = deleteOutliers(A)$.
  end if
  if $|A| < \mu$ **then**
    build ranking $R_1, \ldots, R_h$ from $P_t \cup Q$ by nondominated sorting
    set $j = 1$, $P_{t+1} = \emptyset$
    while $|P_{t+1} \cup R_j| \leq \mu$ **do**
      $P_{t+1} = P_{t+1} \cup R_j$
      $j = j + 1$
    end while
    if $|P_{t+1}| < \mu$ **then**
      add $\mu - |P_{t+1}|$ individuals from R_j at random
    end if
  else
    $P_{t+1} = PSA(A, \mu)$
  end if
  increment t
end while
$E = \mathsf{ND}_f(\{\text{all points of the external file}\}, \preceq)$
$E' = eliminateAllNondominatedOutliers(E \cup P_t)$
$R = PSA(E', \mu)$, $A = \emptyset$
for all $x \in E$ **do**
  $A = \Delta_p\text{-}update(x, A, R)$.
end for
output A

Algorithm 2 Δ_1-update [25]

Require: archive set A, reference set R, new element x
1: $A = \mathsf{ND}_f(A \cup \{x\}, \preceq)$
2: **if** $|A| > N_R := |R|$ **then**
3: **for all** $a \in A$ **do**
4: $h(a) = \Delta_1(A \setminus \{a\}, R)$
5: **end for**
6: $A^* = \{a^* \in A : a^* = \mathsf{argmin}\{h(a) : a \in A\}\}$
7: **if** $|A^*| > 1$ **then**
8: $a^* = \mathsf{argmin}\{\mathsf{GD}_p(A \setminus \{a\}, R) : a \in A^*\}$ {ties broken at random}
9: **end if**
10: $A = A \setminus \{a^*\}$
11: **end if**

the third loop $\Theta(|A|)$ time. Hence, in total $\Theta(|A| \cdot |R| \cdot d)$ time units are required for an update.

We stress that the archiver Δ_1-update considers only contributions of the indicator value of single points as e.g. done in [2,18] using the hypervolume indicator. By

this, convergence toward the optimal archive cannot be guaranteed as the resulting strategy can be viewed as a cyclic search procedure [23]. More sophisticated selection strategies are subject of ongoing research.

Algorithm 3 Quick Δ_1-update [25]

Note: $x+= y \Leftrightarrow x = x + y$
Require: archive set A, reference set R, new element x
$A = \mathsf{ND}_f(A \cup \{x\}, \preceq)$
if $|A| > N_R := |R|$ **then**
  $GD_p = IGD_p = 0$
  for all $a \in A$ **do**
    $GD_p(a) = d(a, R)$ // GD_p contribution of archive point a
    $GD_p += GD_p(a)$ // add GD_p contribution of a
    $IGD_p^1(a) = IGD_p^2(a) = 0$ // initialize for later use
  end for
  for all $r \in R$ **do**
    let $a^* \in A$ such that $d(r, A) = d(r, a^*)$ // closest archive point a^*
    $d_1 = d(r, a^*)$ // distance to closest archive point
    $d_2 = d(r, A \setminus \{a^*\})$ // distance to 2nd closest archive point
    $IGD_p += d_1$ // add IGD_p contribution of r
    $IGD_p^1(a^*) += d_1$ // sum IGD_p contributions with a^* involved
    $IGD_p^2(a^*) += d_2$ // sum IGD_p contributions if a^* deleted
  end for
  $dp_{\min} = gdp_{\min} = \infty$
  for all $a \in A$ **do**
    $gdp = GD_p - GD_p(a)$ // value of GD_p if a deleted
    $igdp = IGD_p - IGD_p^1(a) + IGD_p^2(a)$ // value of IGD_p if a deleted
    $dp = \max\left\{\frac{gdp}{|A|-1}, \frac{igdp}{|R|}\right\}$ // Δ_1 if a deleted
    if $dp < dp_{\min} \vee (dp = dp_{\min} \wedge gdp < gdp_{\min})$ **then**
      $dp_{\min} = dp$ // store smallest Δ_1 seen so far
      $dp_{\min} = gdp$ // store smallest gdp since last improvement of $dp_{\min}$
      $a^* = a$ // save archive point with smallest Δ_1 seen so far
    end if
  end for
  $A = A \setminus \{a^*\}$
end if

The most obvious order of feeding the stored pairs $(x, f(x))$ into the archive updater is the order of their generation. We call this the 'forward update.' In this manner, many individuals will pass the initial dominance check, so that subsequent Δ_p calculations are necessary. Some time saving may be achieved by feeding the stored pairs into the archive update in inverted order. We call this the 'backward update.' Since points that have been generated in later iterations of the EMOA are more likely to dominate previous points, most points from the rear of the inverted sequence will probably not pass the initial dominance check, so that subsequent Δ_p calculations can be avoided. Since the order of the points presented to the archive clearly affects the final outcome of the archive, we shall compare both approaches experimentally in the next section.

5 Experiments: Δ_p approximations

5.1 Experimental setup

In order to experimentally evaluate the Δ_p archive approximation for multiple multiobjective approaches we applied forward and backward Δ_p-update as offline strategies after algorithms' execution. During forward Δ_p-update previously traced generations are successively presented to the update mechanism starting with the initial population, while the backward Δ_p-update procedure starts with the final population of an algorithm's run. Inside the Δ_p-update mechanism we implemented a linear interpolation as well as a PSA-based strategy in order to generate a set of evenly spaced reference points for deciding on a solution being added to the archive or not. Both strategies use the aggregated solutions of the whole run (i.e. from all generations) to create the interpolation or clustering, respectively. These Δ_p update strategies where implemented in JAVA™and plugged into the jMetal framework [8].

As benchmark problems we used well known test problems from multiobjective optimization like the four bi-objective problems which also have been used in [11], an instance of the sphere problem (SPHERE) with $X \subseteq \mathbb{R}^2$ having a convex and connected PF, DTLZ3 with $X \subseteq \mathbb{R}^{10}$ and a concave and connected front, the DENT problem with $X \subseteq \mathbb{R}^2$ and a convex–concave–convex and connected front, as well as ZDT3 with $X \subseteq \mathbb{R}^{20}$ and a convex/concave and disconnected front. Further, we considered WFG1 [16] with $X \subseteq \mathbb{R}^{k+\ell}, k = 2, \ell = 4$ as another problem with a convex-concave front as well as the rotated DTLZ2 problem (R-DTLZ2, $X \subseteq \mathbb{R}^{10}$) proposed in [15] for the 2007 CEC competition. Of course, the rotated problem does not expose a PF different to DTLZ2 or DTLZ3, however, algorithms may perform differently in decision space, which can result in different Δ_p approximations.

In order to compare our Δ_p approximation quality for all algorithms and test problems we generated reference fronts covered by 1000 uniformly spaced points. For all benchmark problems except WFG1 we were able to find a parametric form for exactly calculating the optimal fronts' length L (by rectification). With that we placed 1000 points uniformly on the PF using distance $\Delta L = L/1000$ between points and starting with the extreme points moved inwards by $\Delta L/2$. For WFG1 we used the PF resulting from flooding the search space with solution candidates around the known Pareto set meaning that a fine-granulated grid in each dimension of the decision space is evaluated in the region of the known Pareto set. Afterward, non-dominated solutions are preserved and 1000 well distributed solutions are extracted.[3]

The algorithmic basis for our evaluation were two state-of-the-art general purpose EMOAs (NSGA2 [6] and SMS-EMOA [2] with standard parameter settings: SBX crossover with $p_x = 0.9$ and polynomial mutation with $p_{mut} = 1/n$). In SMS-EMOA hypervolume was computed with a reference point offset of 100. Further, we used PSEMOA as special purpose EMOA with inherent internal clustering for generating evenly distributed solutions. For the same reason we also applied a modified version of NSGA2 proposed by Kukkonen and Deb [19] in which the original crowding distance

[3] This will not lead to exactly uniformly distributed Pareto-optimal solutions. However, the number of reference solutions is by far larger than the approximated set's size.

method is replaced by a sequential variant for gaining an evenly distributed PF approximation. Instead of keeping the μ individuals with largest crowding distance value for the next generation, the individuals with smallest crowding distance values are successively removed based on recomputed crowding distance values. This is done until only μ solutions are left. In the following we will call this approach *sequential crowding distance NSGA2* (SCD-NSGA2). All used algorithms were either used directly from the jMetal framework (NSGA2 and SMS-EMOA were available) or implemented and plugged into the framework.

For experimentation we optimized each test problem with all four algorithms for 50,000 function evaluations and for population sizes $\mu \in \{10, 20, 100\}$. Each optimization run was repeated 20 times. Subsequently, the data generated from the optimizers' runs were submitted to the offline Δ_p archivers. Both forward and backward strategies were parametrized with $p \in \{1, 2\}$.

5.2 Results

Figures 5 and 6 exemplary show the algorithm results for population sizes $\mu = 20$ and $\mu = 100$. A population size of $\mu = 10$ turned out to be too small to generate distinguishable algorithmic results while the general implications are revealed in this case nevertheless. Overall, results become slightly more explicit with increasing population size but overall findings coincide regardless of the population size used. As the findings for $p = 2$ are completely in line with the respective ones for $p = 1$, only the latter are presented in detail in the following.

It becomes obvious that – independent of the specific EMOA which is run as the primary algorithm (PSEMOA, NSGA2, SMS-EMOA, SCD-NSGA2) – the archives generated by Δ_p-updates improve the approximation quality of the PF regarding the Δ_p indicator w.r.t. the benchmark front generated by using optimal PL-metric archives on the known analytical PF representations of the respective test problems (see Sect. 5.1 for details). Each individual figure presents the boxplot of the respective original algorithm on the leftmost side and substantial decreases in Δ_p are achieved for the whole range of strategies in almost all settings. While the focus of the experiments clearly is not on benchmarking the primary EMOAs, note that the PSEMOA in general is not competitive with the other algorithms included into the study. This behavior most likely will change with increasing objective space dimension though, so that it is important to find that the archivers in principle work for bi-objective problems as well.

For all test problems but those consisting of disconnected PFs (ZDT3) the archive strategy using linear interpolations of the set of all nondominated points generated in the course of the algorithm run ($+(fDP)$ and $+(bDP)$) for constructing the benchmark front for the Δ_p archiver outperform the respective PSA-based strategies ($+(fPSA)$ and $+(bPSA)$), in many cases in a statistically significant manner. In each plot, it is tested whether there is a significant deviation in location between $+(fDP)$ and $+(fPSA)$ on the one hand as well as $+(bDP)$ and $+(bPSA)$ on the other hand. Significant differences were confirmed by a Wilcoxon rank sum test ($p \leq 0.05$) and annotated with a $\star$ in Figs. 5 and 6.

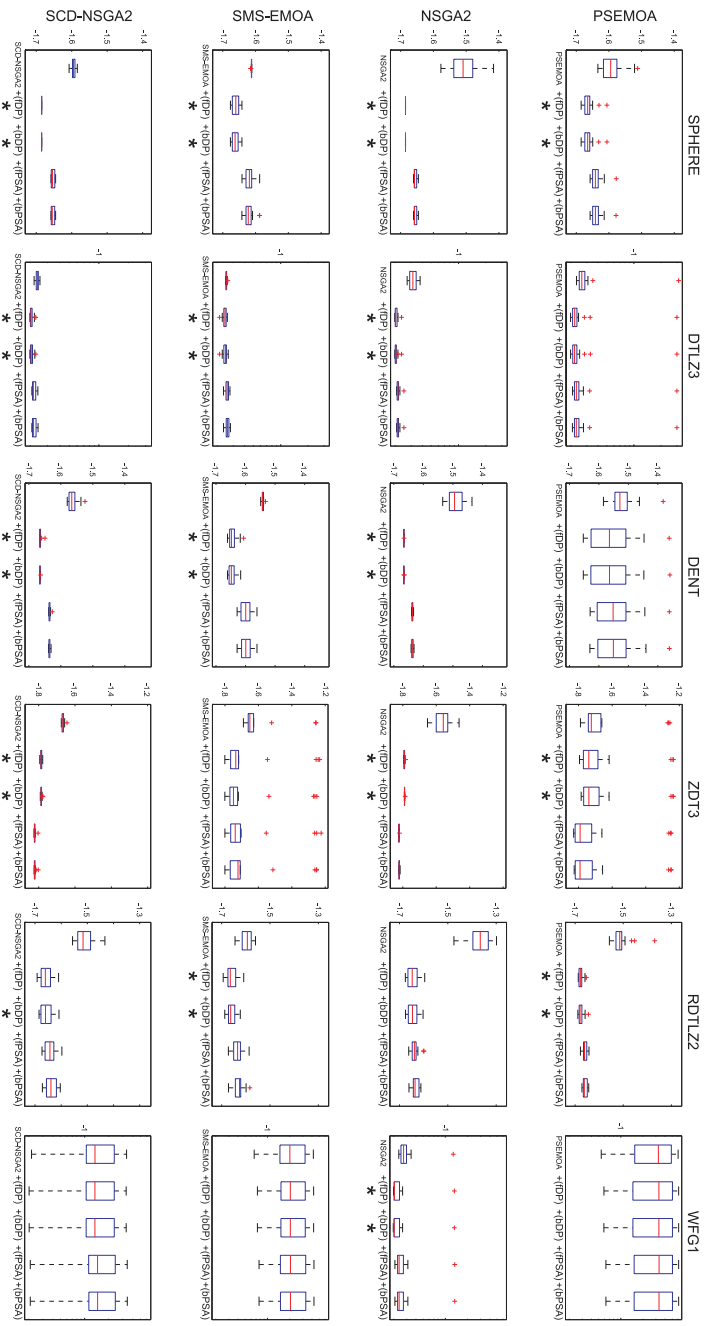

Fig. 5 Δ_p-results of algorithm comparisons for $p = 1$ and $\mu = 20$. Forward update strategies are denoted by +fDP and +fPSA, backward strategies by +bDP and +bPSA. For the scale only exponents are given

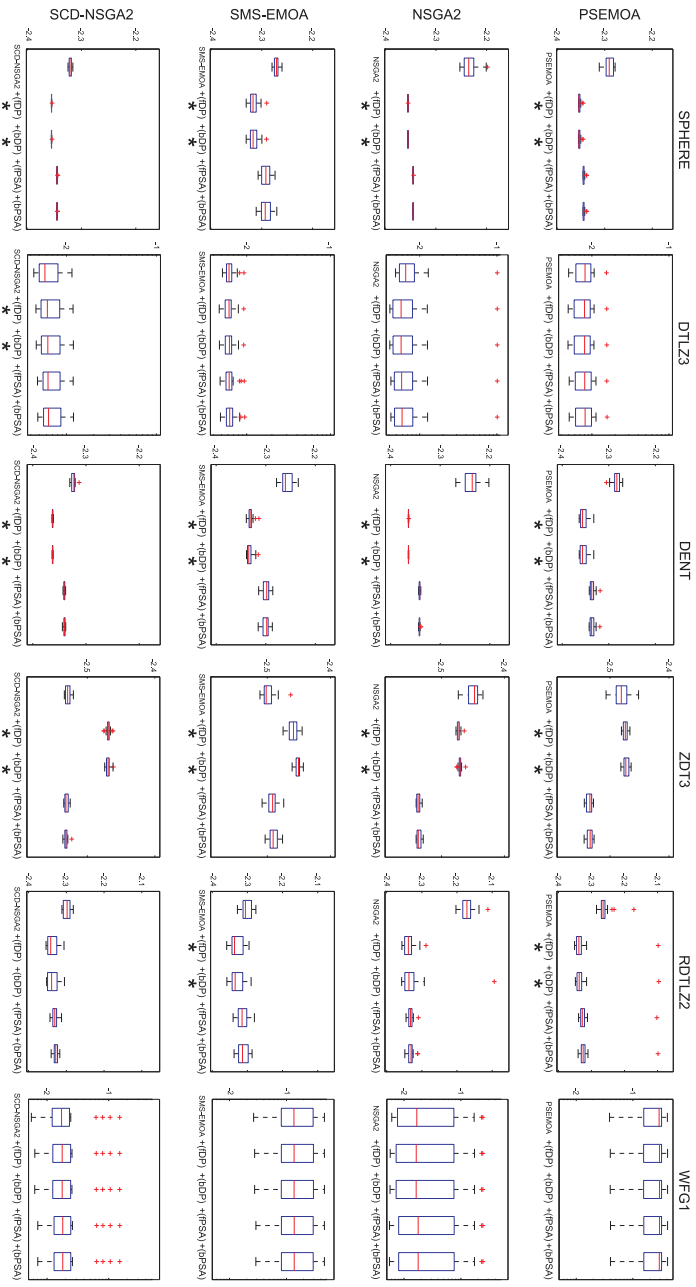

Fig. 6 Δ_p-results of algorithm comparisons for $p = 1$ and $\mu = 100$. Forward update strategies are denoted by +fDP and +fPSA, backward strategies by +bDP and +bPSA. For the scale only exponents are given

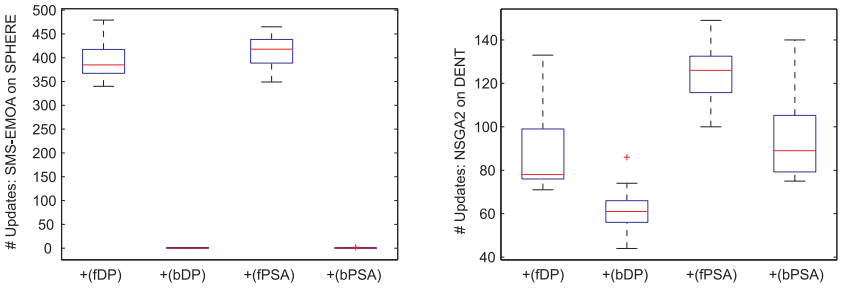

Fig. 7 Exemplary frequencies of Δ_p archive updates: SMS-EMOA with $\mu = 100$ (*left*), NSGA-II with $\mu = 20$ (*right*)

For WFG1 differences between the primary EMOAs and the sophisticated updates strategy results are very small and not statistically significant. Detailed analysis of the algorithm runs revealed that the test problem poses extreme difficulties onto the optimizers so that the PF approximations of the primary EMOAs are not good enough to allow for substantial improvements of the approximation set. In addition, the candidate solutions generated in the course of the primary algorithm run are not diverse enough to allow for an efficient selection regarding equally spaced solutions on the front. By using extremely high budgets of function evaluations (FE) in the order of 500,000 or even one million FE, however, the results show the same tendency as discussed above.

Final approximation quality turns out to be independent from the type of the direction of the update strategy. However, the backward strategy which starts the Δ_p -update loop using the final population of the primary EMOA, has the crucial advantage of being much more efficient regarding the number of necessary update steps and thus requires much less time than the forward strategy which uses the initial population as a starting point. Figure 7 exemplary visualizes the necessary archive updates for the different strategies on the sphere and the DENT problem for different population sizes and EMOA. Qualitatively, the behavior revealed here coincides for the complete set of plots. The inverse update strategy requires a substantially and statistically lower amount of archive update steps than the usual forward strategy. Thus the algorithm results can be computed in a much more efficient way.

In case of disconnected PFs, reflected by the results for the ZDT3 problem, the PSA-approach for constructing the benchmark front for the Δ_p archiver is superior to the one using linear interpolations which is perfectly plausible, since the linear interpolation strategy may lead to benchmark fronts consisting of points in the empty areas between the PF parts which cannot be reached. In [24] it was already reported that the latter strategy reaches its limits in these situations. In many cases, e.g. for all algorithms using $\mu = 20$, however, both types of archiving strategies lead to an improvement of the primary algorithm regarding Δ_p as less points are placed in between the different PF parts.

Table 3 Distance between best numerically obtained archive and optimal IGD_1 archive on the Pareto front in terms of $10^3 \times \Delta_1$, where m denotes the archive size

Archive size m	DTLZ3		RDTLZ2		DTLZ3		RDTLZ2	
	20	100	20	100	20	100	20	100
PSEMOA (+f)	4.57	4.18	17.52	5.55	16.41	4.75	19.54	4.49
PSEMOA (+b)	4.57	4.12	14.34	5.61	15.95	4.83	23.50	4.61
NSGA2 (+f)	1.98	1.79	9.41	5.18	20.64	4.64	19.29	4.49
NSGA2 (+b)	1.98	1.75	9.40	5.20	20.64	4.71	19.33	4.45
SMS-EMOA (+f)	6.76	5.36	7.79	4.63	24.39	4.78	15.91	5.02
SMS-EMOA (+b)	6.76	5.41	8.99	4.68	24.39	4.86	16.03	5.02
SCD-NSGA2 (+f)	2.86	3.57	5.89	5.06	19.62	5.10	21.29	4.44
SCD-NSGA2 (+b)	2.86	3.55	13.26	4.62	19.70	5.12	15.05	4.68

The four rightmost columns contain results from reference fronts built with PSA, whereas the left four columns contain results from reference fronts built with linear interpolation. The extension $(+f)$ indicates the forward update strategy whereas $(+b)$ the update in inverse order

6 Theoretical versus approximative results

In Sect. 3 we were able to calculate optimal archives for linear and spherical fronts. Since the Pareto fronts of test problems DTLZ3 and RDTLZ2 are spherical, we are in the position to compare experimentally obtained archives from Sect. 5 with proven optimal archives from Sect. 3. In Table 3 the archive points from theory are optimal with regard to $\mathsf{L\text{-}IGD}_1$.

As can be seen from the results, all four EMOAs come close to the optimal archive with the help of the Δ_p archive update strategy. Moreover, the differences between the EMOAs becomes smaller for increasing archive size. A similar observation can be made regarding the chosen strategy for the construction of the reference fronts: for a small archive of size 20 the PSA method is much worse than the linear interpolation. But these differences almost diminish for the archive of size 100.

7 Conclusions and outlook

Summarizing, it becomes obvious that the Δ_p indicator is an adequate means for generating equally spaced Pareto front approximations. The theoretical results revealed that the optimal Δ_p archive for connected and concave Pareto fronts is located on the front as GD_p=0. For the special cases of spherical concave as well as linear fronts the exact locations of the solutions can be provided. In case of spherical convex fronts the latter can be derived under the condition of GD_p=0.

The specialized EMOA, i.e., standard EMOA complemented by an *a posteriori* Δ_p-update strategy, proved to outperform the standard EMOA w.r.t. Δ_p, since they are able to generate Pareto front approximations which are much closer to exactly equally spaced nondominated solutions than standard EMOA. Among the analyzed different archive update strategies the updater based on the linearization of the Pareto

front approximation of the standard EMOA performed best for all Pareto front shapes but disconnected ones. In the latter situations the PSA strategy for generating the reference front for the archiver is superior. No substantial differences between forward or backward selection within the archiver can be detected. However, the backward approach is much more efficient and is thus recommended, independent from the strategy for generating the reference front. The experiments additionally showed no qualitative differences for $p = 1$ and $p = 2$. A comparison of the theoretically optimal with the approximated optimal archives generated by the specialized EMOA revealed that differences are neglectable, so that we conclude that the specialized EMOA are very well suited for the desired purpose.

Furthermore, all results offer promising perspectives for higher-dimensional problems. Successful strategies were already presented in [25,28]. Next steps include theoretical analyses regarding optimal Δ_p archives for three objectives as well as experimental comparisons using the specialized EMOA and different archivers. The crucial step in higher dimensions is the construction of an equally spaced higher-dimensional reference front. While for three-objective problems specialized strategies can be developed, starting from four dimensions the PSA approach turns out to be the most promising one [7] which is supported by the results presented in this paper.

Acknowledgments HT and CG acknowledge support by the European Center of Information Systems (ERCIS). OS acknowledges support from Conacyt Project No. 128554. CDM acknowledges support by the Consejo Nacional de Ciencia y Tecnología (CONACYT). All authors acknowledge support from CONACYT Project No. 207403, DFG Project No. TR 891/5-1 and DAAD Project No. 57065955.

References

1. Auger, A., Bader, J., Brockhoff, D., Zitzler, E.: Theory of the hypervolume indicator: optimal μ-distributions and the choice of the reference point. In: Proceedings of the Tenth ACM SIGEVO Workshop on Foundations of Genetic Algorithms (FOGA), pp. 87–102. ACM Press (2009)
2. Beume, N., Naujoks, B., Emmerich, M.: SMS-EMOA: multiobjective selection based on dominated hypervolume. Eur. J. Oper. Res. **181**(3), 1653–1669 (2007)
3. Coello Coello, C.A., Cruz Cortés, N.: Solving multiobjective optimization problems using an Artificial Immune System. Genet. Program. Evolvable Mach. **6**(2), 163–190 (2005)
4. Coello Coello, C.A., Lamont, G.B., Van Veldhuizen, D.A.: Evolutionary Algorithms for Solving Multi-objective Problems, 2nd edn. Springer, New York (2007)
5. Deb, K.: Multi-objective Optimization Using Evolutionary Algorithms. Wiley, New York (2001)
6. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. **6**(2), 182–197 (2002)
7. Dominguez-Medina, C., Rudolph, G., Schütze, O., Trautmann, H.: Evenly spaced Pareto fronts of quad-objective problems using PSA partitioning technique. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC 2013), pp. 3190–3197. IEEE Press, Piscataway, NJ (2013)
8. Durillo, J.J., Nebro, A.J.: jMetal: a Java framework for multi-objective optimization. Adv. Eng. Softw. **42**(10), 760–771 (2011)
9. Emmerich, M., Deutz, A., Kruisselbrink, J., Shukla, P.: Cone-based hypervolume indicators: construction, properties, and efficient computation. In: Purshouse, R., Fleming, P., Fonseca, C., Greco, S., Shaw, J. (eds.) Proceedings of Evolutionary Multi-Criterion Optimization (EMO 2013), pp. 111–127. Springer, Berlin (2013)
10. Emmerich, M.T., Deutz, A.H., Kruisselbrink, J.W.: On quality indicators for black-box level set approximation. In: EVOLVE-A Bridge Between Probability, Set Oriented Numerics and Evolutionary Computation, pp. 157–185. Springer, Berlin (2013)

11. Gerstl, K., Rudolph, G., Schütze, O., Trautmann, H.: Finding evenly spaced fronts for multiobjective control via averaging Hausdorff-measure. In: Proceedings of 8th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE), pp. 1–6. IEEE Press (2011). doi:10.1109/ICEEE.2011.6106656
12. Hansen, M.P., Jaszkiewicz, A.: Evaluating the quality of approximations of the non-dominated set. IMM Technical Report IMM-REP-1998-7, Institute of Mathematical Modeling, Technical University of Denmark, Lyngby (1998)
13. Hillermeier, C.: Nonlinear Multiobjective Optimization—A Generalized Homotopy Approach. Birkhäuser, Basel (2001)
14. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1985)
15. Huang, V., Qin, A., K.Deb, Zitzler, E., Suganthan, P., Liang, J., Preuss, M., Huband, S.: Problem definitions for performance assessment of multi-objective optimization algorithms. Technical Report TR-13, Nanyang Technological University, Singapore (2007). http://www3.ntu.edu.sg/home/epnsugan/index_files/CEC-07/CEC07.htm
16. Huband, S., Hingston, P., Barone, L., While, L.: A review of multiobjective test problems and a scalable test problem toolkit. IEEE Trans. Evol. Comput. **10**(5), 477–506 (2006)
17. Knowles, J., Corne, D.: On metrics for comparing nondominated sets. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC 2002), vol. 1, pp. 711–716. IEEE Press, Piscataway, NJ (2002)
18. Knowles, J.D., Corne, D.W., Fleischer, M.: Bounded archiving using the Lebesgue measure. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC 2003), vol. 4, pp. 2490–2497. IEEE Press, Piscatawa, NJ (2003)
19. Kukkonen, S., Deb, K.: Improved pruning of non-dominated solutions based on crowding distance for bi-objective optimization problems. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC 2006), pp. 1179–1186. IEEE Press, Piscataway, NJ (2006)
20. Mehnen, J., Wagner, T., Rudolph, G.: Evolutionary optimization of dynamic multi-objective test functions. In: Proceedings of the Second Italian Workshop on Evolutionary Computation (GSICE2). ACM Press (2006). CD-ROM; http://ls11-www.cs.uni-dortmund.de/people/rudolph/publications/papers/MWR06.pdf
21. Pareto, V.: Manual of Political Economy. The MacMillan Press, London (1971)
22. Pottharst, A., Baptist, K., Schütze, O., Böcker, J., Fröhlecke, N., Dellnitz, M.: Operating point assignment of a linear motor driven vehicle using multiobjective optimization methods (2004). In: Proceedings of the 11th International Conference EPE-PEMC 2004. Riga, Latvia
23. Powell, M.J.D.: On search directions for minimization algorithms. Math. Program. **4**(1), 193–201 (1973)
24. Rudolph, G., Trautmann, H., Schütze, O.: Homogene Approximation der Paretofront bei mehrkriteriellen Kontrollproblemen. Automatisierungstechnik (at) **60**(10), 612–621 (2012)
25. Rudolph, G., Trautmann, H., Sengupta, S., Schütze, O.: Evenly spaced Pareto front approximations for tricriteria problems based on triangulation. In: Proceedings of 7th International Conference on Evolutionary Multi-Criterion Optimization (EMO 2013), pp. 443–458. Springer, Berlin (2013)
26. Salomon, S., Avigad, G., Goldvard, A., Schütze, O.: PSA—a new scalable space partition based selection algorithm for MOEAs. In: Schütze, O., et al. (eds.) EVOLVE—A Bridge Between Probability, Set Oriented Numerics, and Evolutionary Computation II (Proceedings), vol. 175, pp. 137–151. Springer, Berlin (2013)
27. Schütze, O., Esquivel, X., Lara, A., Coello Coello, C.A.: Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization. IEEE Trans. Evol. Comput. **16**(4), 504–522 (2012)
28. Trautmann, H., Rudolph, G., Dominguez-Medina, C., Schütze, O.: Finding evenly spaced Pareto fronts for three-objective optimization problems. In: Schütze, O., et al. (eds.) EVOLVE—A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation II (Proceedings), pp. 89–105. Springer, Berlin (2013)
29. Veldhuizen, D.A.V.: Multiobjective evolutionary algorithms: classifications, analyses, and new innovations. Ph.D. thesis, Department of Electrical and Computer Engineering. Graduate School of Engineering. Air Force Institute of Technology, Wright-Patterson AFB, Ohio (1999)
30. Witting, K.: Numerical Algorithms for the Treatment of Parametric Optimization Problems and Applications. PhD thesis, University of Paderborn (2012)
31. Witting, K., Schulz, B., Dellnitz, M., Böcker, J., Fröhleke, N.: A new approach for online multiobjective optimization of mechatronic systems. Int. J. Softw. Tools Technol. Transf. **10**(3), 223–231 (2008)

32. Zhang, Q., Li, H.: MOEA/D: a multiobjective evolutionary algorithm based on decomposition. IEEE Trans. Evol. Comput. **11**(6), 712–731 (2007)
33. Zitzler, E., Deb, K., Thiele, L.: Comparison of multiobjective evolutionary algorithms: empirical results. Evol. Comput. **8**(2), 173–195 (2000)
34. Zitzler, E., Thiele, L.: Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. IEEE Trans. Evol. Comput. **3**(4), 257–271 (1999)
35. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., da Fonseca, V.G.: Performance assessment of multiobjective optimizers: an analysis and review. IEEE Trans. Evol. Comput. **7**(2), 117–132 (2003)