J Glob Optim (2008) 41:559–577
DOI 10.1007/s10898-007-9265-7

Convergence of stochastic search algorithms to finite size pareto set approximations

Oliver Schütze · Marco Laumanns · Carlos A. Coello Coello · Michael Dellnitz · El-Ghazali Talbi

Received: 7 November 2006 / Accepted: 3 October 2007 / Published online: 5 February 2008

Abstract In this work we investigate the convergence of stochastic search algorithms toward the Pareto set of continuous multi-objective optimization problems. The focus is on obtaining a finite approximation that should capture the entire solution set in a suitable sense, which will be defined using the concept of ϵ-dominance. Under mild assumptions about the process to generate new candidate solutions, the limit approximation set will be determined entirely by the archiving strategy. We propose and analyse two different archiving strategies which lead to a different limit behavior of the algorithms, yielding bounds on the obtained approximation quality as well as on the cardinality of the resulting Pareto set approximation.

Keywords Multi-objective optimization · Convergence · ϵ-dominance · Stochastic search algorithms

Mathematics Subject Classification (2000): 65C20 · 90B50

O. Schütze (✉) · C. A. Coello Coello
Computer Science Department, CINVESTAV-IPN, Av. IPN No. 2508, Col. San Pedro Zacatenco, Mexico City 07300, Mexico
e-mail: schuetze@cs.cinvestav.mx

C. A. Coello Coello
e-mail: ccoello@cs.cinvestav.mx

M. Laumanns
Institute for Operations Research, ETH Zurich, 8092 Zurich, Switzerland
e-mail: laumanns@ifor.math.ethz.ch

M. Dellnitz
Faculty for Computer Science, Electrical Engineering and Mathematics, Institute for Mathematics, University of Paderborn, Warburger Strasse 100, 33098 Paderborn, Germany
e-mail: dellnitz@upb.de

E.-G. Talbi
LIFL, CNRS Bât M3, Cité Scientifique, INRIA Futurs, 59655 Villeneuve d'Ascq, France
e-mail: talbi@lifl.fr

1 Introduction

A common goal in multi-objective optimization is to identify the set of Pareto-optimal solutions (the efficient set) and its image in objective space, the Pareto front (the efficient frontier). Except for special cases, where the Pareto set is finite or, e.g., representable by a finite collection of faces of a polyhedron (such as in multi-objective linear programming), it is in general not practicable to determine the entire Pareto set. Therefore, a suitable approximation concept is needed.

Various approximation concepts based on ϵ-efficiency are surveyed in [4]. As most of them deal with infinite sets, they are not practical for our purpose of producing and maintaining a representative subset of finite size. The use of discrete ϵ-approximations of the Pareto set was suggested simultaneously by Evtushenko and Potapov [1], Reuter [10], and Ruhe and Fruhwirt [13]. The general idea is that each Pareto-optimal point is approximately dominated by some point of the approximation set.

Despite the existence of suitable approximation concepts, investigations on the *convergence* of particular algorithms toward such approximation sets, that is, their ability to obtain a suitable Pareto set approximation in the limit, have remained rare. Several studies, such as [3,12], consider only the convergence to the entire Pareto set, or to a certain subset without considering the approximation quality.

Finally, the issue of stochastic convergence toward finite-size Pareto set approximations was raised in the area of evolutionary multi-objective optimization, mostly under the assumption of finite search space. One option is to use Markov chain results assuming the underlying search processes to be Markovian [11]. Another option is to define an order homomorphism of the natural dominance relation of approximation sets into a totally ordered set of quality values, thus enforcing a monotonicity of the sequence of solution sets maintained by an algorithm. As shown in [5,6], this entails convergence to a subset of the Pareto set as a local optimum of the quality indicator, but no approximation guarantee could given. Knowles and Corne [6] also analyzed the adaptive grid archiving proposed in [7] and proved that after finite time, even though the solution set itself might permanently oscillate, it will always represent an ϵ-approximation whose approximation quality depends on the granularity of the adaptive grid and on the number of allowed solutions. The results depend on the additional assumption that the grid boundaries converge after finite time, which is fulfilled in certain special cases.

In [8], two archiving algorithms were proposed that provably maintain a finite-size approximation of all points ever generated during the search process. As an immediate corollary, these archiving strategies were claimed to ensure convergence to a Pareto set approximation of given quality for any iterative search algorithm that fulfills certain mild assumptions about the process to generate new search points. While this claim holds trivially in the case of discrete (or discretized) search spaces, its extension to the continuous case is not straightforward. A restriction to *discretized* models, however, can lead to problems when, e.g., memetic strategies are used (metaheuristic search algorithms mixed with local search strategies which itself use step size control).

The goal of this article is to establish archiving strategies to obtain finite Pareto set approximations for stochastic multi-objective optimization algorithms working in continuous domains. We start by considering the first archiving strategy, which is a variant of the strategy from [8], and prove convergence with probability one to an ϵ-approximate Pareto set in the limit. Then we propose a new archiving strategy that additionally ensures that all elements of the limit set are Pareto-optimal points themselves. For both strategies, we give bounds on the approximation quality and on the cardinality of the limit solution set. Finally, computational results are given on two test problems to demonstrate the effects and benefits of both archiving strategies.

2 Background

In the following, we consider continuous unconstrained multi-objective optimization problems

$$\min_{x \in \mathbb{R}^n} \{F(x)\}, \tag{MOP}$$

where the function F is defined as the vector of the objective functions

$$F: \mathbb{R}^n \to \mathbb{R}^k, \quad F(x) = (f_1(x), \ldots, f_k(x)),$$

and where each $f_i: \mathbb{R}^n \to \mathbb{R}$ is continuous.

Definition 2.1 (a) Let $v, w \in \mathbb{R}^k$. Then the vector v is *less than* w ($v <_p w$), if $v_i < w_i$ for all $i \in \{1, \ldots, k\}$. The relation $\leq_p$ is defined analogously.

(b) A point $y \in \mathbb{R}^n$ is *dominated* by a point $x \in \mathbb{R}^n$ (in short: $x \prec y$) with respect to (MOP) if $F(x) \leq_p F(y)$ and $F(x) \neq F(y)$ (i.e., there exists a $j \in \{1, \ldots, k\}$ such that $f_j(x) < f_j(y)$), else y is called nondominated by x.

(c) A point $x \in \mathbb{R}^n$ is called *Pareto optimal* or a *Pareto point* if there is no $y \in \mathbb{R}^n$ which dominates x.

(d) A point $x \in \mathbb{R}^n$ is called weakly *Pareto optimal* if there does not exist another point $y \in \mathbb{R}^n$ such that $F(y) <_p F(x)$.

We now define a weaker concept of dominance, called (absolute) ϵ-dominance, which is used as the approximation concept in the remainder of this study.

Definition 2.2 Let $\epsilon = (\epsilon_1, \ldots, \epsilon_k) \in \mathbb{R}^k_+$ and $x, y \in \mathbb{R}^n$. x is said to ϵ-dominate y (in short: $x \prec_\epsilon y$) with respect to (MOP) if $F(x) - \epsilon \leq_p F(y)$ and $F(x) - \epsilon \neq F(y)$.

Note that the ϵ-dominance relation—unlike the ordinary dominance defined above—is not transitive, i.e., if $x \prec_\epsilon y$ and $y \prec_\epsilon z$ it does *not* follow that $x \prec_\epsilon z$, but it follows that $x \prec_{2\epsilon} z$. This fact will be used in later considerations as well as the following: if $x \prec y$ and $y \prec_\epsilon z$ it follows that $x \prec_\epsilon z$.

Definition 2.3 [8] Let $\epsilon \in \mathbb{R}^k_+$.

(a) A set $F_\epsilon \subset \mathbb{R}^n$ is called an ϵ-approximate Pareto set of (MOP) if every point $x \in \mathbb{R}^n$ is ϵ-dominated by at least one $f \in F_\epsilon$, i.e.,

$$\forall x \in \mathbb{R}^n: \exists f \in F_\epsilon: \; f \prec_\epsilon x$$

(b) A set $F^*_\epsilon \subset \mathbb{R}^n$ is called an ϵ-Pareto set if F^*_ϵ is an ϵ-approximate Pareto set and if every point $f \in F^*_\epsilon$ is a Pareto point of (MOP).

Let $B_\delta(x_0) := \{x \in \mathbb{R}^n: \|x - x_0\| < \delta\}$ be the open ball with center $x_0 \in \mathbb{R}^n$ and radius $\delta \in \mathbb{R}_+$. A k-dimensional box B can be represented by a center $c \in \mathbb{R}^k$ and a radius $r \in \mathbb{R}^k_+$:

$$B = B(c, r) = \{x \in \mathbb{R}^k: \; |x_i - c_i| \leq r_i \quad \forall i = 1, \ldots, k\}.$$

Next, we need the following distances between different sets.

Definition 2.4 Let $u \in \mathbb{R}^n$ and $A, B \subset \mathbb{R}^n$. The semi-distance dist$(\cdot,\cdot)$ and the Hausdorff distance $d_H(\cdot,\cdot)$ are defined as follows:

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

Algorithm 1 gives a framework of a generic stochastic multi-objective optimization algorithm, which is considered in this work. Here, $Q \subset \mathbb{R}^n$ denotes the domain of the MOP, P_j the candidate set (or population) of the generation process at iteration step j, and A_j the corresponding archive. Theorem 2.5 states a convergence result which is closely related to the present work, but which leads in general to unbounded archive sizes.

Algorithm 1 Generic Stochastic Search Algorithm

```
1: P_0 ⊂ Q drawn at random
2: A_0 = ArchiveUpdate(P_0, ∅)
3: for j = 0, 1, 2, ... do
4:     P_{j+1} = Generate(P_j)
5:     A_{j+1} = ArchiveUpdate(P_{j+1}, A_j)
6: end for
```

Theorem 2.5 *[14] Let an MOP $F: \mathbb{R}^n \to \mathbb{R}^k$ be given, where F is continuous, let $Q \subset \mathbb{R}^n$ be compact. Further, let there be no weak Pareto point in $Q \backslash P_Q$ (where P_Q denotes the set of Pareto points of $F\big|_Q$), and*

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

Then an application of Algorithm 1, where all nondominated points are kept, i.e., $ArchiveUpdate(P, A) := \{x \in P \cup A: y \nprec x \ \forall y \in P \cup A\}$, generates a sequence of archives $\{A_i\}_{i \in \mathbb{N}}$, such that

$$\lim_{i \to \infty} d_H(F(P_Q), F(A_i)) = 0 \quad \textit{with probability one.}$$

3 The algorithms

In the following we investigate two different strategies for the archiving of the solutions found by the algorithm leading to different limit behaviors of the sequence of archives (under certain additional conditions).

First, we assume that the entries of $\epsilon \in \mathbb{R}^k_+$ are 'small', and thus that it is sufficient to obtain an ϵ-approximate Pareto set. For this, we consider an archiving strategy very similar to the one proposed in [8], here given as Algorithm 2. It computes the subsequent archive A of a given archive A_0, a population P, an $\epsilon \in \mathbb{R}^k_+$, and a $\Theta \in (0, 1)$. Using this strategy, the sequence of archives has a limit behavior described in Theorem 3.2. To show this, we first need the following obvious but crucial property of the archiving strategy.

Algorithm 2 $A := ArchiveUpdateEps1\ (P, A_0)$

$A := A_0$
for all $p \in P$ **do**
  if $\exists a \in A :\ a \prec_{\Theta\epsilon} p$ **then**
    CONTINUE ▹ *do not execute lines* 6 – 11
  end if
  for all $a \in A$ **do**
    if $p \prec a$ **then**
      $A := A \backslash \{a\}$
    end if
  end for
  $A := A \cup \{p\}$
end for

Lemma 3.1 *Let* $A_0, P \subset \mathbb{R}^n$ *be finite sets,* $\epsilon \in \mathbb{R}^k_+$, *and* $A :=$ *ArchiveUpdateEps1* (P, A_0). *Then the following holds:*

$$\forall x \in P \cup A_0\colon\ \exists a \in A\colon a \prec_{\Theta\epsilon} x.$$

Proof Roughly speaking, the statement follows since points a are only discarded from the archive if in turn another point p with $p \prec a$ is inserted (this 'replacement' is given in lines 7, 8 and 11 in Algorithm 2). To be more precise, let $P = \{p_1, p_2, \ldots, p_l\}$, $l \in \mathbb{N}$. Without loss of generality we assume that all points p_i are considered in this ordering (i.e., in the for-loop in line 2 of Algorithm 2). There are two cases we have to distinguish.
Case A $x \in A_0$. Define $p'_0 := x$ and

$$p'_i := \begin{cases} p_i & \text{if} \quad p_i \text{`replaces'}\, p'_{i-1} \\ p'_{i-1} & \text{else} \end{cases}, \quad i = 1, \ldots, l.$$

It holds that $p'_l \in A$ and either $p'_l = x$ or $p'_l \prec x$ (due to the transitivity of $\prec$). In both cases it is $p'_l \prec_{\Theta\epsilon} x$.
Case B $x \in P$. Let $x = p_j$, $j \in \{1, \ldots, l\}$. After the j-th iteration of the outer for-loop in Algorithm 2 there exists an element $a_j \in A$ with $a_j \prec_{\Theta\epsilon} p_j$ (line 3 resp. line 11 of Algorithm 2). Define $p'_j := a_j$ and p'_i, $i = j+1, \ldots, l$, as above. It follows that $p'_l \in A$ and $p'_l \prec_{\Theta\epsilon} x$ as claimed. □

Theorem 3.2 *Let an MOP* $F : \mathbb{R}^n \to \mathbb{R}^k$ *be given, where F is continuous, let* $Q \subset \mathbb{R}^n$ *be a compact set and* $\epsilon \in \mathbb{R}^k_+$. *Further let*

$$\forall x \in Q \quad and \quad \forall \delta > 0 : \qquad P\,(\exists l \in \mathbb{N}\colon\ P_l \cap B_\delta(x) \cap Q \neq \emptyset) = 1 \tag{2}$$

Then an application of Algorithm 1, where $ArchiveUpdateEps1$ *is used to update the archive, leads to a sequence of archives* A_l, $l \in \mathbb{N}$, *such that there exists with probability one an* $l_0 \in \mathbb{N}$ *such that* A_l *is an* ϵ*-approximate Pareto set for all* $l \geq l_0$.

Proof Let $m_i = \min_{x \in Q} f_i(x)$ and $M_i = \max_{x \in Q} f_i(x)$, $1 \leq i \leq k$. Define

$$Q' = [m_1 - \Theta\epsilon_1, M_1] \times \cdots \times [m_k - \Theta\epsilon_k, M_k]$$

and

$$D(A, \epsilon) := \{y \in Q' \mid \exists a \in A : F(a) - \Theta\epsilon \leq_p y\}.$$

A point p is only inserted into the given archive A if there exists no $a \in A$ which $\Theta\epsilon$-dominates p. Assume A and p are given, and that p is inserted into A. Denote by A' the resulting archive. By Lemma 1 it follows that

$$D(A,\epsilon) \subset D(A',\epsilon)$$

whether or not p dominates points $a \in A$. Define

$$B_p := \{y \in \mathbb{R}^k \mid y \leq_p F(p) \text{ and } F(p) - \Theta\epsilon \leq_p y\}$$

It is $\overset{\circ}{B}_p \cap D(A,\epsilon) = \emptyset$, where $\overset{\circ}{B}_p$ denotes the interior of B_p (assume there exists an $y \in \overset{\circ}{B}_p \cap D(A,\epsilon)$, i.e., there exists an $a \in A$ with $F(a) - \Theta\epsilon \leq_p y <_p F(p)$, and thus $a \prec_{\Theta\epsilon} p$, a contradiction). Since $B_p = B(F(p) - \frac{\Theta\epsilon}{2}, \frac{\Theta\epsilon}{2})$, i.e., B_p is a k-dimensional box with center $c = F(p) - \frac{\Theta\epsilon}{2}$ and radius $r = \frac{\Theta\epsilon}{2}$, we have

$$Vol(D(A',\epsilon)) \geq Vol(D(A,\epsilon)) + \underbrace{Vol(B_p)}_{=\prod_{i=1}^k \Theta\epsilon_i > 0}, \tag{3}$$

where $Vol(D)$ denotes the k-dimensional volume of a set $D \subset \mathbb{R}^k$. Since $0 \leq Vol(D(A,\epsilon)) \leq Vol(Q') \leq \infty$, it follows only finitely many insertions can be done.

Using this observation we can now show the existence of a number l_0 such that the A_{l_0} is an ϵ-approximate Pareto set. For this, we assume that there exists no such number l_0. That is, for every $l \in \mathbb{N}$ there exists a point $x = x(l) \in Q$ which is not ϵ-dominated by any of the elements of the archive A_l.

Let $l \in \mathbb{N}$. Since there exists no $a \in A_l$ which ϵ-dominates $x(l)$ and since $\Theta\epsilon <_p \epsilon$ there exists also no $a \in A$ which $\Theta\epsilon$-dominates $x(l)$. Moreover, since $\Theta < 1$ and F is continuous there exists a neighborhood U_l of $x(l)$ such that

$$\nexists a \in A_l\colon a \prec_{\Theta\epsilon} u \quad \forall u \in U_l \tag{4}$$

By (2) it follows that there exists with probability one a number $i_l \in \mathbb{N}$ and a point $x_{i_l} \in P_{i_l} \cap Q$. By (4) and the construction of $ArchiveUpdateEps1$ (lines 3 resp. 11 of Algorithm 2) it follows that either this point has to be inserted, or that another point $\tilde{x}_l$ has been inserted in one of the iteration steps $l+1,\ldots,i_l$. In any case, at least one insertion must have occurred from iteration step $l+1$ to i_l.

Proceeding in this manner, a sequence of infinitely many insertions can be constructed, which contradicts the observation made above. Thus, there must exist with probability one an l_0 such that A_{l_0} forms an ϵ-approximate Pareto set as claimed. Further, all subsequent archives are ϵ-approximate Pareto sets due to Lemma 1, and the proof is complete. □

Before we proceed with the next archiving strategy we make some remarks on the assumptions made above as well as on the approximation quality of the limit archive.

Remarks 3.3

(a) The value of $\Theta < 1$ is needed to guarantee the convergence in the probabilistic sense. Consider for instance $F = (x^2, x^2)$, $\epsilon = (1,1)$, and $\Theta = 1$. Further, let the archive be given by $A = \{1\}$. This archive does not form an ϵ-approximate Pareto set, and the probability to improve this set is zero since $x^* = 0$ is the only point which is not ϵ-dominated by 1. The following consideration shows directly (unlike the proof above, which is done by contradiction) that this cannot happen when choosing $\Theta < 1$:

Let $x \in Q$. By (2) it follows that there exists with probability one for every $i \in \mathbb{N}$ an $l_i \in \mathbb{N}$ and a point x_i such that $x_i \in P_{l_i} \cap B_{1/i}(x) \cap Q$. By construction of $ArchiveUpdateEps1$ there exists for every x_i an element $a_i \in A_{l_i}$ such that $a_i \prec_{\Theta\epsilon} x_i$ (and by Lemma 1 there exist further for all $l \geq l_i$ an entry $a_i^l \in A_l$ such that $a_i^l \prec_{\Theta\epsilon} x_i$). Since $\lim_{i\to\infty} x_i = x$ and since F is continuous there exists an $l \in \mathbb{N}$ such that

$$|f_i(x_l) - f_i(x)| \leq \frac{1-\Theta}{2}\epsilon_i, \quad i = 1, \ldots, k. \tag{5}$$

Since $a_l \prec_{\Theta\epsilon} x_l$ and by (5) we have

$$\begin{aligned} f_i(a_l) - \Theta\epsilon_i \quad &\leq f_i(x_l) \leq f_i(x) - \tfrac{1-\Theta}{2}\epsilon_i \quad \forall i = 1, \ldots, k, \text{ and} \\ f_j(a_l) - \Theta\epsilon_i \; &< f_j(x_l) \leq f_j(x) - \tfrac{1-\Theta}{2}\epsilon_j \quad \text{for a } \; j \in \{1, \ldots, k\}, \end{aligned} \tag{6}$$

and hence a_l is $\frac{3\Theta-1}{2}\epsilon$-dominating x, and since $\frac{3\Theta-1}{2} < 1$ we have also $a_l \prec_\epsilon x$. That is, for $x \in Q$ there exists with probability one an entry in the archive which ϵ-dominates the given point x as desired. Despite these theoretical considerations, $\Theta = 1$ can be chosen in practise.

(b) Assumption (2) is the crucial part to obtain the convergence. For general ϵ and general F it is certainly not possible to postulate less. Given a fixed $\epsilon \in \mathbb{R}^k_+$ it would in principle be sufficient to require condition (2) only for the δ which is given in the proof above as well as for finitely many points $x \in Q$. However, this is nearly impossible to check in practice.

(c) Here we have used the absolute ϵ-dominance. If $0 \notin f_i(P_Q)$, $i = 1, \ldots, k$, alternatively the *relative* ϵ-dominance as in [8] can be used yielding similar results.

(d) We have restricted the domain to a compact subset of the $\mathbb{R}^n$. The following (academic) example shows that we can run into trouble if Q is not compact: consider the MOP

$$F\colon \mathbb{R}_+ \to \mathbb{R}^2$$
$$F(x) = \left(-x, -\frac{1}{x}\right)$$

In this case, the Pareto set is given by $\mathcal{P} = \mathbb{R}_+$. Since $F(P)$ is not bounded below it can not be represented by a finite archive using ϵ-dominance. However, this changes if $Q = [a,b]$, $a < b$, $a, b > 0$ is chosen as the domain. In this case, we have $\mathcal{P} = Q = [a,b]$ and both objectives are Lipschitz continuous on Q with Lipschitz constants

$$L_1 = \max_{x\in[a,b]} |f_1'(x)| = 1 \quad \text{and} \quad L_2 = \max_{x\in[a,b]} |f_2'(x)| = \frac{1}{a^2}.$$

By the Lipschitz continuity it follows that $|f_i(x) - f_i(y)| \leq \epsilon_i$ if $|x-y| \leq \epsilon_i/L_i$, $i = 1, 2$. Thus, the set of equally distributed points

$$x_j = a + \frac{j}{N}(b-a), \quad j = 0, \ldots, N,$$

where $N = \left\lceil \frac{b-a}{\min(\epsilon_1/L_1, \epsilon_2/L_2)} \right\rceil$, forms an ϵ-approximate Pareto set (even an ϵ-Pareto set since since all $x_j \in \mathcal{P}$).

Remarks 3.4

(a) The 'exclusion strategy' in lines 3 and 4 of Algorithm 2 makes it possible that the sequence of archives reaches a steady state solution which forms an ϵ-approximate Pareto

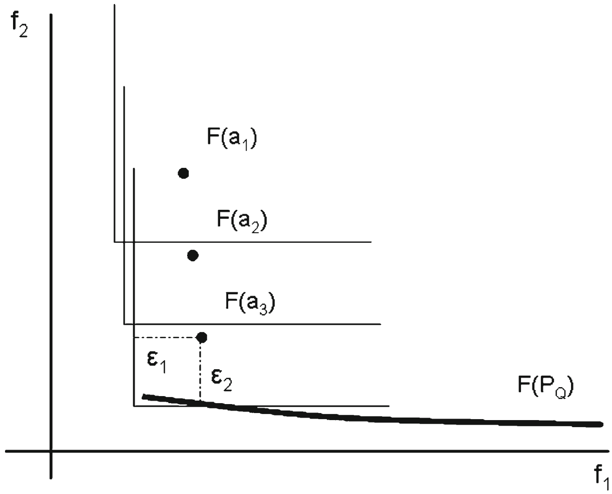

Fig. 1 Possible example of a set which was generated by $ArchiveUpdateEps1$ with $dist(F(A), F(\mathcal{P}_Q)) \gg \epsilon$

set after finitely many steps. On the other hand, exactly this feature prevents that we can guarantee convergence of the entries of an archive toward an ϵ-Pareto set. Further, it can happen that $dist(F(A), F(\mathcal{P}_Q))$ gets 'large' (that is, images $F(a), a \in A$, could be 'far away' from the Pareto front), as the following example shows (cf. Fig. 1): assume that the elements a_1, a_2, a_3 are inserted into the archive in this order (and w.l.o.g. assume that $\Theta = 1$). By construction of $ArchiveUpdateEps1$, these points will not be removed in the subsequent steps since there exists no point $p \in Q$ which dominates a_i, $i \in \{1, 2, 3\}$, but is not ϵ-dominated by any of them. In such a manner an example can be constructed with $dist(F(A), F(\mathcal{P}_Q)) = max_{i=1,..,k}(M_i - m_i)$ with $m_i = \min_{x \in Q} f_i(x)$ and $M_i = \max_{x \in Q} f_i(x)$. However, this bad theoretical value has never been observed in our computations.

(b) Note that the number of entries in an archive is (amongst others) dependent on the insertion order of the candidate solutions as the previous example shows (compare to Fig. 1): if a_1, a_2, a_3 are inserted in this order, all elements are added to the archive, but if a_3 is added first to the archive, neither a_1 nor a_2 will be added to the archive since both points are ϵ-dominated by a_3.

Next, we assume that the entries of ϵ are relatively large. This can be the case when the decision maker prefers to obtain few, widespread solutions of the MOP, or in order to be able to approximate the entire Pareto set with a limited archive, in particular when considering more than two objectives. Hence, convergence of the entries of the sequence of archives toward the Pareto set is desired. For this, we propose to use the archiving strategy that is described in Algorithm 3. In the following, we discuss the limit behavior of this approach.

Lemma 3.5 *Let* $A_0, P \subset \mathbb{R}^n$ *be finite sets,* $\epsilon \in \mathbb{R}^k_+$*, and* $A := ArchiveUpdateEps2\ (P, A_0)$*. Then the following holds:*

$$\forall x \in P \cup A_0: \quad \exists a \in A: a \prec_{\Theta\epsilon} x.$$

Proof Analogue to the proof of Lemma 3.1. □

Algorithm 3 $A := ArchiveUpdateEps2\ (P, A_0)$

$A := A_0$
for all $p \in P$ **do**
  if $\nexists a \in A :\ a \prec_{\Theta\epsilon} p$ **then**
    $A := A \cup \{p\}$
  end if
  for all $a \in A$ **do**
    if $p \prec a$ **then**
      $A := A \cup \{p\}\backslash\{a\}$
    end if
  end for
end for

Theorem 3.6 *Let (MOP) be given and $Q \subset \mathbb{R}^n$ be compact, and let there be no weak Pareto points in $Q\backslash P_Q$. Further, let F be injective and*

$$\forall x \in Q \quad and \quad \forall \delta > 0: \qquad P\,(\exists l \in \mathbb{N}:\ P_l \cap B_\delta(x) \cap Q \neq \emptyset) = 1 \tag{7}$$

Then an application of Algorithm 1, where $ArchiveUpdateEps2$ 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 with probability one an $l_0 \in \mathbb{N}$ such that A_l is an ϵ-approximate Pareto set for all $l \geq l_0$.
(b)

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

Proof (a) Analogue to the proof of Theorem 3.2 (a).
(b) A new element p is added to the archive A by $ArchiveUpdateEps2$ only in one of the following two cases: (1) if there exists no $a \in A$ which $\Theta\epsilon$-dominates p (denote by *type 1 insertion* for the remainder of this proof), and (2), if p dominates an element $a \in A$, which will in turn be discarded from the archive (denote by *type 2 insertion*). Since A is finite and P_Q is compact it follows that

$$dist(A, P_Q) = \max_{a\in A} \min_{p\in P_Q} ||a - p||.$$

We prove the claim in the following way: we show that for every $a \in A$ there exists a sequence a_i of dominating points with $a_i \to a^* \in P_Q$ which will be inserted in future instances of the archive. Since $dist(a^*, P_Q) = 0$ and since all the points a_i, $i \in \mathbb{N}$, do not have to be considered for the limit archive due to the insertion of type 2 it remains to show that there are only finitely many insertions of type 1 (which is the only manipulation of the archive by which the value of $dist(A, P_Q)$ can be increased). This will be done in the second part.
Let $l_0 \in \mathbb{N}$ and $a_0 \in A_{l_0}$. If $a_0 \in P_Q$ it follows that $a_0 \in A_{l+m}$, $\forall m \in \mathbb{N}$. In this case set $a_i := a_0$. Now assume that $a_0 \notin P_Q$. Define

$$\begin{aligned} &M\colon Q \to \mathbb{R} \\ &M(x) := \max_{p\in P_Q} \min_{i=1,\ldots,k} (f_i(x) - f_i(p)) \end{aligned} \tag{8}$$

Under the assumptions made above it holds that

$$M(x) \geq 0\ \forall x \in Q \quad \text{and} \quad M(x) = 0 \Leftrightarrow x \in P_Q.$$

Let $p_0 \in P_Q$ be the argument of the maximum of $M(a_0)$. Since $a_0 \notin P_Q$ and a_0 is not a weak Pareto point it follows that $M(a_0) > 0$ and $F(p_0) <_p F(a_0)$. Since F is continuous there exists a neighborhood U_{p_0} of p_0 such that

$$F(y) <_p F(p_0) + \frac{M(a_0)}{2} \cdot (1, \dots, 1) \quad \forall y \in U_{p_0},$$

and thus, that $F(y) <_p F(a_0),\ \forall y \in U_{p_0}$. By (7) it follows that *Generate* (see Algorithm 1) generates with probability one after finitely many steps a point $b \in U_{p_0} \cap Q$. Now there are two cases: (1) b is added to the archive (in this case set $a_1 := b$), and (2), a_0 has already been replaced by an element $\tilde{a} \in \mathbb{R}^n$ such that b and $\tilde{a}$ are mutually nondominating (in this case set $a_1 := \tilde{a}$). In both cases there exists a $j \in \{1, \dots, k\}$ such that

$$f_j(a_1) < f_j(p_0) + \frac{M(a_0)}{2}.$$

Proceeding in an analogous way we obtain a sequence $\{a_i\}_{i\in\mathbb{N}}$ of dominating points. Since the sequence $\{F(a_i)\}_{i\in\mathbb{N}}$ is bounded below and F is injective it follows that $a_i \to a^* \in Q$ for $i \to \infty$.

Next we show that $a^* \in P_Q$. For this, assume that $a^* \notin P_Q$. Define p^* as the argument of the maximum of $M(a^*)$. Since $a^* \notin P_Q$ and a^* is not a weak Pareto point it follows that $F(p^*) <_p F(a^*)$ and $M(a^*) > 0$. Proceeding further as above we obtain a point a^{**} and an element $j \in \{1, \dots, k\}$ such that

$$\begin{aligned} f_j(a^{**}) &< f_j(p^*) + \frac{M(a^*)}{2} \le f_j(p^*) + \frac{f_j(a^*) - f_j(p^*)}{2} \\ &= \frac{f_j(p^*) + f_j(a^*)}{2} < f_j(a^*) \end{aligned}$$

This is a contradiction to the assumption of the convergence of the sequence, and thus it must be that $a^* \in P_Q$.

Further, it can be shown analogously to $ArchiveUpdateEps1$ (see proof of Theorem 3.2) that only finitely many insertions of type 1 can be done during the run of the algorithm. Thus, only finitely many times the value of $dist(A, P_Q)$ can be increased (an insertion of type 2 decreases the value), and the claim for the limit archive follows. □

4 Bounds on the archive sizes

In the following, we give bounds on the magnitude of the limit archives with respect to $\epsilon \in \mathbb{R}^k_+$ and the chosen archiving strategy. We assume that $|P_0| = 1$, and thus also $|A_0| = 1$. The lower bound of the limit archive for both archiving strategies is obviously given by 1. For this, consider e.g., $f_1 = f_2 = \dots = f_k$ to be a convex function which takes its (unique) minimum inside Q. The upper bounds for the different archiving strategies are derived separately.

Theorem 4.1 *Let* $m_i = \min_{x\in Q} f_i(x)$ *and* $M_i = \max_{x\in Q} f_i(x)$, $1 \le i \le k$, *and* $|A_0| = 1$. *Then, when using* $ArchiveUpdateEps1$, *the archive size maintained in Algorithm 1 for all* $l \in \mathbb{N}$ *is bounded as*

$$|A_l| \le \sum_{\substack{i_1,\dots,i_{k-1}=1 \\ i_1 > \dots > k_{k-1}}}^{K} \prod_{j=1}^{k-1} \left\lceil \frac{M_{i_j} - m_{i_j}}{\Theta \epsilon_{i_j}} \right\rceil, \tag{9}$$

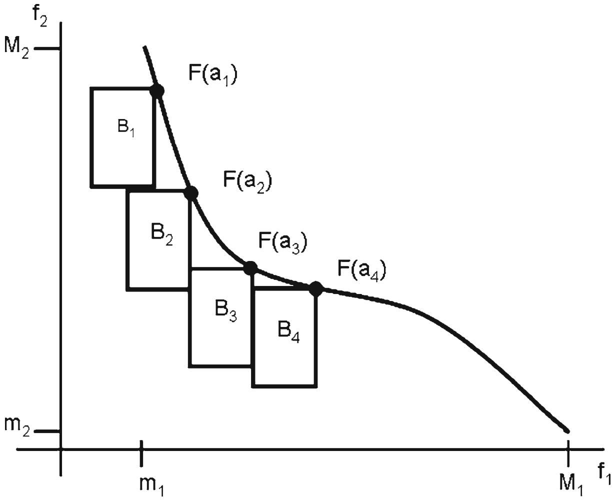

Fig. 2 The entries a_i of each archive lie on a (virtual) curve c. Since interiors of the boxes B_i (with upper right corners $F(a_i)$) are mutually non-intersecting, the minimal component-wise distance between two entries is determined by $\Theta\epsilon$

Proof Consider a sequence $p_1, p_2, \ldots$ of points which are all accepted by $ArchiveUpdateEps1$ in this order (i.e., starting with $A_0 = \{p_1\}$). Consider the i-th step and let $A_i = \{a_1, \ldots, a_l\}$ with $l \leq i$. Define $B_j := B(F(a_j) - \frac{\Theta\epsilon}{2}, \frac{\Theta\epsilon}{2}),\ \ j = 1, \ldots, l$ (compare to proof of Theorem 3.2). Using inductive arguments we see that (a) all elements in A_i are mutually nondominating, and that (b) the interiors of all the boxes $B_j,\ \ j = 1, \ldots, l$, are mutually non-intersecting. Since the points a_j are the upper right corners of the boxes B_j and since the interiors of these boxes are mutually non-intersecting the minimal distance between two points a_{j_1} and a_{j_2}, $j_1 \neq j_2$, is determined by $\Theta\epsilon$ (see Fig. 2). Thus we are able to bound the number of entries in the archives if we can bound the number of such boxes which can be placed in the image space such that their lower left corners are mutually nondominating and the boxes are mutually non-intersecting.

Let us first consider a bi-objective model (i.e., $k = 2$), since in this case the proof is geometrically descriptive and already captures the basic idea. Since all points a_j are mutually nondominating, the images of these points are all located on a (virtual) continuously differentiable curve

$$\begin{aligned} c\colon [m_1, M_1] &\to \mathbb{R}^2 \\ u &\mapsto (u, f(u)) \end{aligned} \tag{10}$$

where $f\colon [m_1, M_1] \to [m_2, M_2]$ is a strictly monotonically decreasing (but not necessarily surjective) function. The length of this curve can be bounded as follows:

$$\begin{aligned} L(c) &= \int_{m_1}^{M_1} \|c'(u)\| \mathrm{d}u = \int_{m_1}^{M_1} \sqrt{|1|^2 + |f'(u)|^2} \mathrm{d}u \\ &\leq \int_{m_1}^{M_1} 1 \mathrm{d}u + \int_{m_1}^{M_1} |f'(u)| \mathrm{d}u = \int_{m_1}^{M_1} 1 \mathrm{d}u - \int_{m_1}^{M_1} f'(u) \mathrm{d}u \\ &\leq (M_1 - m_1) + (M_2 - m_2) \end{aligned} \tag{11}$$

Thus, for $k = 2$ we see that $|A_i| \leq \left\lceil \frac{(M_1 - m_1)}{\Theta\epsilon_1} \right\rceil + \left\lceil \frac{(M_2 - m_2)}{\Theta\epsilon_2} \right\rceil$, $i \in \mathbb{N}$, as claimed above. Now we turn our attention to the general case, i.e., let $k \geq 2$ be given. Define

$$\begin{aligned} K &:= [m_1, M_1] \times \cdots \times [m_{k-1}, M_{k-1}], \\ K_{(i)} &:= [m_1, M_1] \times \cdots \times [m_{i-1}, M_{i-1}] \times [m_{i+1}, M_{i+1}] \times \cdots \times [m_{k-1}, M_{k-1}], \\ u_{(i)} &:= (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_{k-1}), \; i = 1, \ldots, k-1. \end{aligned} \tag{12}$$

In analogy to the bi-objective case, the images of the elements of the archives are located in the graph of a map Φ which is characterized as follows:

$$\begin{aligned} &\Phi\colon K \to \mathbb{R}^k \\ &\Phi(u_i, \ldots, u_{k-1}) = (u_1, \ldots, u_{k-1}, f(u_1, \ldots, u_{k-1})), \end{aligned} \tag{13}$$

where $f\colon K \to [m_k, M_k]$ is a sufficiently smooth function satisfying the monotonicity conditions $\frac{\partial f}{\partial u_i} u < 0$, $\forall u \in K$ and $\forall i = 1, \ldots, k-1$. Then, the $(k-1)$-dimensional volume of Φ with parameter range K can be bounded as follows:

$$\begin{aligned} Vol(\Phi) &= \int_K \sqrt{||\nabla f||^2 + 1} \mathrm{d}u = \int_K \sqrt{\left(\frac{\partial f}{\partial u_1}\right)^2 + \ldots + \left(\frac{\partial f}{\partial u_{k-1}}\right)^2 + 1} \mathrm{d}u \\ &\leq \int_K \left|\frac{\partial f}{\partial u_1}\right| \mathrm{d}u + \ldots + \int_K \left|\frac{\partial f}{\partial u_{k-1}}\right| \mathrm{d}u + \int_K 1 \mathrm{d}u \\ &= \sum_{i=1}^{k-1} \left(\int_{K_{(i)}} \left(\int_{m_i}^{M_i} \left|\frac{\partial f}{\partial u_i}\right| \mathrm{d}u_i \right) \mathrm{d}u_{(i)} \right) + \int_K 1 \mathrm{d}u \\ &= \sum_{i=1}^{k-1} \left(\int_{K_{(i)}} \left(- \int_{m_i}^{M_i} \frac{\partial f}{\partial u_i} \mathrm{d}u_i \right) \mathrm{d}u_{(i)} \right) + \int_K 1 \mathrm{d}u \\ &\leq \sum_{\substack{i_1, \ldots, i_{k-1} = 1 \\ i_1 > \ldots > k_{k-1}}}^{k} \prod_{j=1}^{k-1} (M_{i_j} - m_{i_j}) \end{aligned} \tag{14}$$

This bound of the volume leads directly to the bound of the cardinality of the archives as stated above, which completes the proof. □

Figure 1 indicates how to construct an example for the bi-objective case where the size of the archive reaches this maximal bound. However, this bound is typically much too pessimistic in practice since (a) the shape of the Pareto front can be arbitrary (e.g., disconnected), and thus its $(k-1)$-dimensional volume much smaller than estimated above, (b) in case the Pareto front contains 'flat' regions these are typically covered by few solutions (see example in Sect. 5.1), and (c) the number of entries in the archive depends on the ordering of the candidates (see Remark 3.4(c)).

However, the next example motivates that the usage of this archiving strategy and its broad bound can be of advantage, e.g., for discrete problems, where the Pareto set is a finite—but possibly huge—set.

Example 4.2 As an example consider bi-criteria {0, 1}-knapsack problems of the following form (e.g., [2]):

$$f_1, f_2\colon \{0,1\}^n \to \mathbb{R}$$
$$f_i(x) = \sum_{j=1}^{n} c_j^i x_j, \quad i = 1, 2 \tag{15}$$

such that

$$\sum_{j=1}^{n} w_j x_j \le W,$$

where $c_j^i \in [0, 1]$ represents the (normalized) value of item j on criterion i, $i = 1, 2. w_j$ is the weight of item j, and W the overall knapsack capacity. The cardinalities of the Pareto sets of such capacity constrained models are typically relatively large. For instance, for the following special case the number of nondominated solutions can be stated explicitly: for $c_j^1 + c_j^2 = 1$ and $w_j = k,\ j = 1, \dots, n$ (i.e., constant sum of the criteria coefficients and equal weighted items) and $l := \lfloor \frac{W}{k} \rfloor \le n$ the cardinality of the Pareto set is given by $|\mathcal{P}| = \binom{n}{l} = \frac{n!}{(n-l)!l!}$. Hence, e.g., for $n = 20$ items and $l = 10$ the Pareto set consists of $|P_{(20,10)}| = 184,756$ efficient solutions, and for $n = 50,\ l = 20$ we have $|P_{(50,20)}| = 1.26 \times 10^{13}$.

As an example for the maximal magnitude of the archive when using *ArchiveUpdateEps*1 we assume for simplicity that ϵ has equal entries, $\epsilon = (\bar{\epsilon}, \dots, \bar{\epsilon})$. Since for all efficient solutions it holds that $|x| := x^T x = l$ ([2]) and since $c_j^i \in [0, 1]$ we can assume that $f_i(a) \in [0, l],\ \forall a \in A$, and thus $m_i = 0$ and $M_i = l$. Hence we obtain (for $\Theta = 1$)

$$|A| \le \left(\frac{l}{\bar{\epsilon}} + \frac{l}{\bar{\epsilon}}\right) = 2\frac{l}{\bar{\epsilon}}$$

Choosing $\bar{\epsilon} = 0.1$ – which corresponds to 10% of the value of one item – we obtain the upper bounds $|A_{(20,10)}| \le 200$ for $(n, l) = (20, 10)$ and $|A_{(50,20)}| \le 500$ for $(n, l) = (50, 20)$. Although pessimistic, these estimates are considerably lower than the magnitudes of the entire Pareto sets themselves.

Theorem 4.3 *Let* $m_i = \min_{x \in Q} f_i(x)$ *and* $M_i = \max_{x \in Q} f_i(x)$, $1 \le i \le k$, *and* $|A_0| = 1$. *Then, when using ArchiveUpdateEps*2*, the archive size maintained in Algorithm 1 is bounded for all* $l \in \mathbb{N}$ *as*

$$|A_l| \le \prod_{i=1}^{k} \left\lceil \frac{M_i - m_i}{\Theta \epsilon_i} \right\rceil. \tag{16}$$

Proof We can consider the process of including solutions into the archive over time as a process for constructing a directed graph G. Starting with an empty graph, we add a new node for each new solution p that is added to the archive A in line 4 or line 8 of the algorithm. If p is added in line 8 (meaning the condition in line 7 is true), we add arcs (p, a) from p to each solution a that is discarded in line 8 due to $p \prec a$. Let $V_t := \bigcup_{1 \le j \le t} A_j$ be the union of all archives up to iteration t and $V_t' \subseteq V_t$ the subset of those archive members that have been added in line 4. Thus, the node set of G_t after iteration t is V_t, and G_t itself is a forest whose roots are the current archive members A_t and whose leafs are the elements of V_t'. Obviously, the number of roots must be smaller than the number of leafs, so $|A_t| \le |V_t'|$.

To bound $|V_t'|$, the number of elements that ever entered the archive in line 4, we again consider the boxes $B_v := B(F(v) - \frac{\Theta\epsilon}{2}, \frac{\Theta\epsilon}{2})$ for all $v \in V_t'$. Due to line 3, a solution p generated in iteration $t' \le t$ cannot be accepted in line 4 if $F(p)$ lies inside the box B_v of any previously accepted element of $v \in V_t'$, otherwise $a \prec_{\Theta\epsilon} p$ for some current archive member $a \in A_t$ as there exists $a \in A_t$ with $F(a) \leq F(v)$ and $v \prec_{\Theta\epsilon} p$. If p was accepted in line 4, then $F(p)$ cannot lie inside the box B_v of any subsequently accepted element of $v \in V_t'$ neither, as this would entail $p \prec v$. Hence, the interiors of the boxes B_v must be mutually non-intersecting. The maximum number of non-intersecting boxes with side length $\Theta\epsilon$ and centers c with $m_i \leq c_i \leq M_i$ is $\prod_{i=1}^{k} \lceil (M_i - m_i)/(\Theta\epsilon_i) \rceil$, thus the claimed bound on the archive size follows. □

In the following example we construct a sequence of objective vectors to show that the bound of $ArchiveUpdateEps2$ is tight for $n = 2$.

Example 4.4 Let an MOP be given with $n = 2$ and $F(Q) = [0, M_1] \times [0, M_2]$, where $M_i = r_i \cdot t_i, t_i = \Theta\epsilon_i$, and $r_i \in \mathbb{N}$ a given integer, for $i = 1, 2$. We are going to construct an archive consisting of r_2 groups of r_1 elements each, where the elements can become arbitrarily close within a group, and the groups are separated by a distance of t_2.

For each group i, $1 \leq i \leq r_2$ repeat the following: first, generate the r_1 vectors

$$v_{i,j}^1 = ((i-1)r_1\delta_1 + (j-1)(t_1+\delta_1), M_2 - (i-1)(r_2\delta_2 + t_2) - (j-1)\delta_2), \quad j = 1, \ldots, r_1,$$

where $\delta_1 \leq t_1/(r_1 r_2))$ and $\delta_2 \leq t_2/(r_2^2 + r_1)$, which ensures that all these vectors are inside $F(Q)$. Now generate the $r_1 - 1$ further vectors

$$v_{i,j}^2 = ((i-1)r_1\delta_1 + (j-1)\delta_1, M_2 - (i-1)(r_2\delta_2 + t_2) - (j-1)\delta_2), \quad j = 2, \ldots, r_1,$$

each of which replacing its dominated counterparts with identical second coordinate in the archive.

After all r_2 groups have been constructed in this way, the final archive consists of $r_1 \cdot r_2$ elements, which is exactly the value of the upper bound given for $ArchiveUpdateEps2$ since $r_i = \frac{M_i - m_i}{\Theta\epsilon_i}$, $i = 1, 2$, by construction of this example.

5 Numerical results

In this section we make a comparative study on two test problems in order to illustrate the effect of the different archiving strategies. In both examples the overall computational time is significantly lower when using the novel archiving strategies compared to the 'standard' one, which stores all nondominated solutions obtained during the search process.
The computations have been done on an Intel Xeon 3.2 GHz processor.

5.1 Example 1

First we consider the following bi-objective example (K. Witting and A. Hessel Von Molo, Private Communication, 2004):

$$\begin{aligned} & f_1, f_2 \colon \mathbb{R}^2 \to \mathbb{R}^2 \\ f_1(x, y, \lambda) &= \tfrac{1}{2}(\sqrt{1 + (x+y)^2} + \sqrt{1 + (x-y)^2} + x - y) + \lambda \cdot e^{-(x-y)^2} \\ f_2(x, y, \lambda) &= \tfrac{1}{2}(\sqrt{1 + (x+y)^2} + \sqrt{1 + (x-y)^2} - x + y) + \lambda \cdot e^{-(x-y)^2} \end{aligned} \tag{17}$$

To obtain a fair comparison of the different archivers we have decided to take a random search operator for the generation process. Moreover, we have taken the same sequence of points for all settings. For the subsequent comparisons we have used the following archiving strategies:

(ND) *ArchiveUpdateND*, i.e., all nondominated points are kept,
(Eps1) *ArchiveUpdateEps1*, and
(Eps2) *ArchiveUpdateEps2*,

Figure 3 shows one example for resulting limit sets. Hereby we have taken $N = 500{,}000$ randomly chosen points in $Q = [-0.5, 0.5]^2$, and have chosen $\lambda = 0, \epsilon = (0.15, 0.15)$, and $\Theta = 1$. The set obtained by $ArchiveUpdateND$ is already very close to the real Pareto front. However, in that case the time which had to be spent to update the archive[1] is tremendous compared to both other strategies, see Table 1 for the average running times.

It has to be noted that the value of N is relatively high, and that such a value is higher than needed to obtain a 'reasonable' approximation for this low-dimensional problem. However, very often stochastic search procedures are run 'over night', and in this case typically a huge number of maximal function calls is adjusted. And in these cases—and when the execution time for a function call is not too large—the storage of the 'good' solutions can be the bottleneck of the search procedure.

Next we consider the same MOP but use the value $\lambda = 0.85$. In this case the Pareto front contains a dent when choosing the domain as $Q = [-1.5, 1.5]^2$. Figure 4 displays a result similar as for the previous one for $N = 200{,}000$ randomly chosen points, and Table 2 shows the corresponding averaged running times. Note that the approximation obtained by using $ArchiveUpdateEps1$ reveals some gaps, which might be unwanted in certain applications (e.g., [9]). These gaps can occur in regions where the Pareto front is 'flat', which is due to the nature of ϵ-dominance. For possible ways of overcoming this problem, we refer to subsequent work [15].

5.2 Example 2

Next we consider the following three-objective MOP ([14]):

$$f_1, f_2, f_3 \colon \mathbb{R}^{10} \to \mathbb{R}$$
$$f_i(x) = \sum_{\substack{j=1 \\ j \neq i}}^{10} (x_j - a_j^i)^2 + (x_i - a_i^i)^4, \tag{18}$$

where

$$\begin{aligned} a^1 &= (1, 1, 1, 1, \ldots) && \in \mathbb{R}^{10} \\ a^2 &= (-1, -1, -1, -1, \ldots) && \in \mathbb{R}^{10} \\ a^3 &= (1, -1, 1, -1, \ldots) && \in \mathbb{R}^{10} \end{aligned}$$

For the comparative study we have taken $N = 5{,}000{,}000$ randomly chosen points within the domain $Q = [-1, 1]^{10}$. Figure 5 shows one numerical result using the three archiving strategies, where we have chosen $\epsilon = (1/3, 1/3, 1/3)$ and $\Theta = 1$. Apparently, the approximation qualities of the three sets are similar, that is, no set seems to be significantly 'better' than any of the others. To measure the real approximation quality the *epsilon indicator* [16] is used, where $I_\epsilon(A, B)$ gives the smallest value of $\bar{\epsilon} \in \mathbb{R}$ such that A is an ϵ-approximate Pareto set of B where $\epsilon = (\bar{\epsilon}, \ldots, \bar{\epsilon})$, i.e.,

$$I_\epsilon(A, B) := \min\{\bar{\epsilon} \in \mathbb{R} \mid \forall b \in B\ \exists a \in A : \ a \prec_\epsilon b\}.$$

For the data obtained in this test the values can be seen in Table 3. This shows that the differences in the different approximation qualities are only very marginal. Such a similarity

[1] The elements of all archives were stored using a linear list.

Fig. 3 Three limit achives obtained by different archiving strategies for MOP (17) for $\lambda = 0$

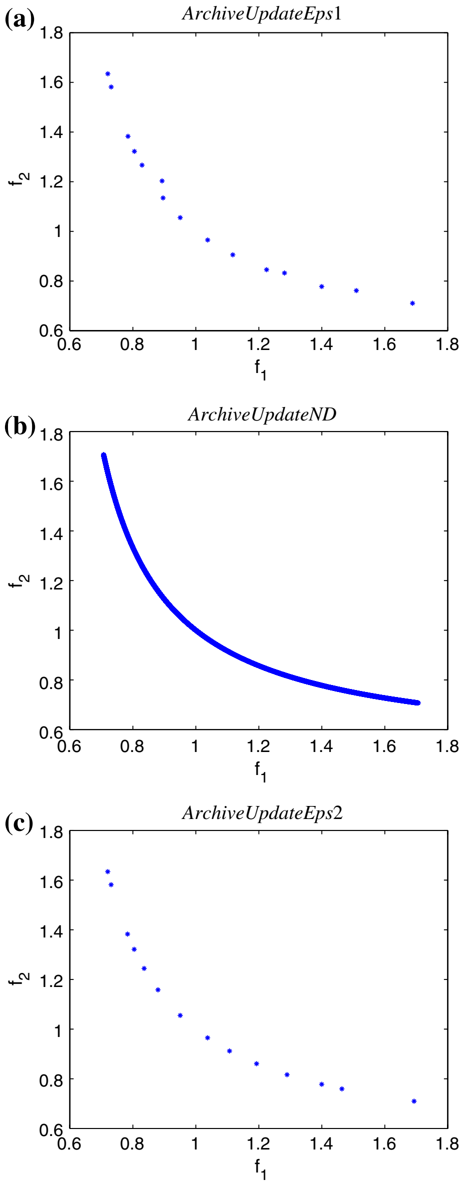

Table 1 Comparison of the magnitudes of the final archive ($|A_N|$, rounded) and the corresponding update times (T, in seconds) for different archiving strategies for MOP (17) and for $\lambda = 0$

	ND	Eps_1	Eps_2
$\|A_N\|$	11,874	15	14
T	3730	0.56	2.6

We have taken the average result of 100 test runs

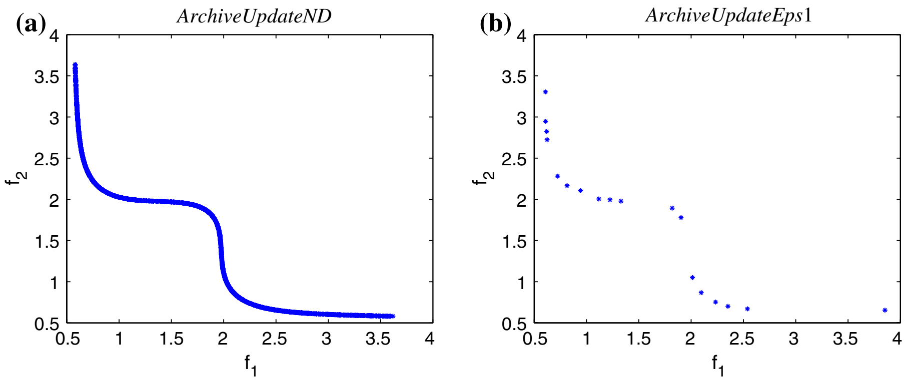

Fig. 4 Numerical result on MOP (17) for $\lambda = 0.85$ and $\epsilon = (0.3, 0.3)$. The set obtained by $ArchiveUpdateEps1$ reveals gaps in the approximation

Table 2 Comparison of the magnitudes of the final archive ($|A_N|$, rounded) and the corresponding update times (T, in seconds) for different archiving strategies for MOP (17) and for $\lambda = 0.85$

	ND	Eps_1	Eps_2
$\lvert A_N \rvert$	1481	15	14
T	134.03	0.28	1.37

We have taken the average result of 100 test runs

does not hold, however, for the different running times (see Table 4). In fact, these times offer a huge variety. First, it is obvious that the archiver $ArchiveUpdateEps1$ is much faster in this case than $ArchiveUpdateEps2$. This is most probably due to the fact that $ArchiveUpdateEps1$ rejects more solutions, and that this rejection process is done faster than in the second approach—with the sacrifice of dropping the ability to converge toward an ϵ-Pareto set. However, these two archivers are both much faster than $ArchiveUpdateND$. Compared to $ArchiveUpdadateEps2$ the difference is of one order of magnitude, and compared to $ArchiveUpdadateEps1$ the difference is even of two orders of magnitude.

Thus, regarding both the similarity of the approximation quality and the significant difference in the running times we can draw the conclusion that the archiving strategies considered in this work, $ArchiveUpdateEpsi, i = 1, 2$, are advantageous compared to the 'standard' one for this application—and probably for others as well.

6 Conclusion and future work

We have proposed generic stochastic search algorithms for obtaining ϵ-approximate Pareto sets as well as ϵ-Pareto sets of a continuous multi-objective optimization problem in the limit. We have presented a convergence result for these algorithms, have given bounds on the cardinality of the corresponding archives, and have finally presented some numerical results.

For future work, there are a lot of interesting topics which can be addressed to advance the present work. One could, for instance, consider the speed of convergence, in particular when the methods presented above are hybridized with local search strategies. Further, we

Fig. 5 Numerical result on MOP (18) using different archiving strategies

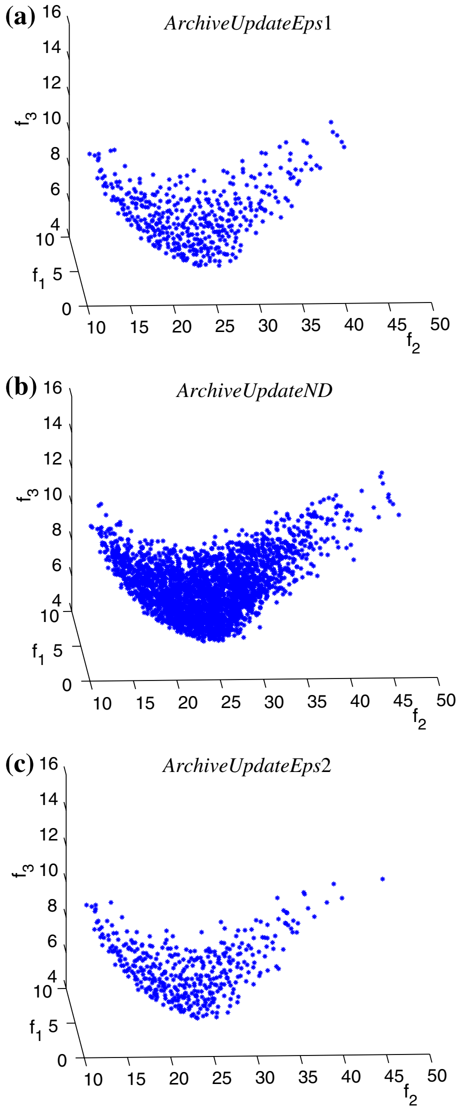

Table 3 Values of the epsilon indicator $I_\epsilon(A, B)$ on the sets shown in Fig. 5: ND denotes the set obtained by $ArchiveUpdateND$ and $Epsi$ the sets obtained by $ArchiveUpdateEpsi$, $i = 1, 2$

	$A = ND$ $B = Eps1$	$A = ND$ $B = Eps2$	$A = Eps1$ $B = Eps2$
$I_\epsilon(A, B)$	0	0	0.1545
$I_\epsilon(B, A)$	0.1599	0.1588	0.1586

Table 4 Comparison of the magnitudes of the final archive ($|A_N|$, rounded) and the corresponding update times (T, in seconds) for different archiving strategies for MOP (18)

	ND	Eps_1	Eps_2
$\lvert A_N \rvert$	2533	419	417
T	7326	10	660

We have taken the average result of 100 test runs.

intend to apply this theoretical framework in search for the development of fast and reliable multi-objective optimization algorithms.

Acknowledgements The authors would like Emilia Tantar for help and fruitful discussions on the content of this article. They also thank the anonymous reviewers for their valuable comments which greatly helped them to improve the contents of this paper. The third author gratefully acknowledges support from CONACyT Project no. 45683-Y.

References

1. Evtushenko, Y.G., Potapov, M.A.: Methods of numerical solution of multicriterion problem. Soviet Math. Doklady **34**, 420–423 (1987)
2. Gomes da Silva, C., Climaco, J., Figueira, J.: Geometrical configuration of the Pareto frontier of bi-criteria {0,1}-knapsack problems. Research Report 16, INESC-COIMBRA (2004)
3. Hanne, T.: On the convergence of multiobjective evolutionary algorithms. Euro. J. Operat. Res. **117**(3), 553–564 (1999)
4. Helbig, S., Pateva, D.: On several concepts for ϵ-efficiency. OR Spektrum **16**(3), 179–186 (1994)
5. Knowles, J.D.: Local-search and hybrid evolutionary algorithms for Pareto optimization. PhD thesis, The University of Reading, Department of Computer Science, Reading, UK (2002)
6. Knowles, J.D., Corne, D.: Properties of an adaptive archiving algorithm for storing nondominated vectors. IEEE Trans. Evol. Comput. **7**(2), 100–116 (2003)
7. Knowles, J.D., Corne, D.W.: Approximating the non-dominated front using the Pareto Archived Evolution Strategy. Evol. Comput. **8**(2), 149–172 (2000)
8. Laumanns, M., Thiele, L., Deb, K., Zitzler, E.: Combining convergence and diversity in evolutionary multiobjective optimization. Evol. Comput. **10**(3), 263–282 (2002)
9. 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. In: Proceedings of the 11th International Conference EPE-PEMC 2004, Riga, Latvia (2004)
10. Reuter, H.: An approximation method for the efficiency set of multiobjective programming problems. Optimization **21**, 905–911 (1990)
11. Rudolph, G.: Finite Markov chain results in evolutionary computation: A tour d'horizon. Fundamenta Informaticae **35**, 67–89 (1998)
12. Rudolph G., Agapie A.: On a multi-objective evolutionary algorithm and its convergence to the Pareto set. In: Congress on Evolutionary Computation (CEC2000), pp. 1010–1016 (2000)
13. Ruhe, G., Fruhwirt, B.: ϵ-optimality for bicriteria programs and its application to minimum cost flows. Computing **44**, 21–34 (1990)
14. Schütze, O.: Set oriented methods for global optimization. PhD thesis, University of Paderborn (2004). <http://ubdata.uni-paderborn.de/ediss/17/2004/schuetze/>
15. Schütze, O., Laumanns, M., Tantar, E., Coello Coello, C.A., Talbi, E.-G.: Convergence of stochastic search algorithms to gap-free Pareto front approximations. In: GECCO '07: Proceedings of the 9th annual Conference on Genetic and Evolutionary Computation, pp. 892–901. ACM Press, New York, NY, USA (2007)
16. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C., Fonseca, V.: Performance assessment of multiobjective optimizers: an analysis and review. IEEE Trans. Evol. Comput. **7**(2), 117–132 (2003)