Convergence of Stochastic Search Algorithms to Finite Size Pareto Set Approximations

Oliver Schütze[1], Marco Laumanns[2], Carlos A. Coello Coello[3], Michael Dellnitz[4], and El-ghazali Talbi[1].

[1] INRIA Futurs, LIFL, CNRS Bât M3, Cité Scientifique
59655 Villeneuve d'Ascq, France
e-mail: {schuetze,talbi}@lifl.fr

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

[3] CINVESTAV-IPN, Electrical Engineering Department
Av. IPN No. 2508, Col. San Pedro Zacatenco
México D.F. 07300, MEXICO.
e-mail: ccoello@cs.cinvestav.mx

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

Abstract

In this work we study the convergence of generic 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 investigate 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.

Key words: multi-objective optimization, convergence, ϵ-dominance, stochastic search algorithms.

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 representable by a finite collection of line segments (such as in multi-objective linear programming), it is in general not practicable to determine the entire Pareto set. Instead, a suitable approximation concept is needed.

Various approximation concepts based on ϵ-efficiency are given in [3]. 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. Using discrete ϵ-approximations of the Pareto set was suggested simultaneously by [1], [5], and [7]. 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 towards such approximation sets, that is, their ability to obtain a suitable Pareto set approximation in the limit, have remained rare. Several studies, such as [2, 6], consider only the convergence to the entire Pareto set, or to a certain subset without considering the approximation quality.

In [4] the issue of convergence towards a finite-size Pareto set approximation was finally addressed for a general class of iterative search algorithms. Two archiving algorithms were proposed that provably maintain a finite-size approximation of all points ever generated during the search process. This led to the claim that these archiving strategies will 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. Consideration of *discretized* models, however, can lead to problems when, e.g., using memetic strategies (metaheuristic search algorithms mixed with local search strategies which itself use step size control).

The goal of this paper is to establish convergence results with respect to finite Pareto set approximations for stochastic multi-objective optimization algorithms working in continuous domains. We start by considering the first archiving strategy from [4] 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 itself. For both strategies we give bounds on the approximation quality and on the cardinality of the limit solution set.

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

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

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 vector $y \in \mathbb{R}^n$ is dominated *by a vector $x \in \mathbb{R}^n$ (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 non-dominated 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 weakly Pareto optimal *if there does not exist another point $y \in \mathbb{R}^n$ such that $F(y) <_p F(x)$.*

In the following we will define a weaker concept of dominance, so-called (absolute) ϵ-dominance, which will be used for our further studies.

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*

(i) $f_i(x) - \epsilon_i \leq f_i(y) \quad \forall i = 1, \ldots, k$, and

(ii) $f_j(x) - \epsilon_j < f_j(y) \quad$ for at least one $j \in \{1, \ldots, k\}$.

We have to emphasize that the ϵ-dominance relation – unlike the 'classical' one 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 *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.*

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

*(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).*

Further, 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 $\delta \in \mathbb{R}_+$.

Algorithm 1 gives a framework of a generic stochastic multi-objective optimization algorithm, which will be considered in this work. Theorem 2.4 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: $P_0 \subset Q$ drawn at random
2: $A_0 = ArchiveUpdate(P_0, \emptyset)$
3: **for** $j = 0, 1, 2, \ldots$ **do**
4: $\quad P_{j+1} = Generate(P_j)$
5: $\quad A_{j+1} = ArchiveUpdate(P_{j+1}, A_j)$
6: **end for**

Theorem 2.4 *[8] 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*

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

Then an application of Algorithm 1, where all non-dominated points are kept, i.e., $ArchiveUpdate(P, A) := \{x \in P \cup A : y \not\prec x \ \forall y \in P \cup A\}$, generates a sequence of archives $\{A_i\}_{i \in \mathbb{N}}$, such that

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

where $d(\cdot, \cdot)$ denotes the Hausdorff distance.

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 the archiving strategy proposed in [4], here given as Algorithm 2. It computes the subsequent archive A of a given archive A_0, a population P, and an $\epsilon \in \mathbb{R}^k_+$. Using this strategy, the sequence of archives has a limit behavior described in Theorem 3.2. To show this, we need first the following obvious but crucial property of the archiving strategy.

Algorithm 2 $A := ArchiveUpdate1_\epsilon\ (P, A_0)$

1: $A := A_0$
2: **for all** $p \in P$ **do**
3:     **if** $\exists a \in A:\ a \prec_{\epsilon/3} p$ **then**
4:         CONTINUE $\triangleright$ *do not execute lines* 6 – 11
5:     **end if**
6:     **for all** $a \in A$ **do**
7:         **if** $p \prec a$ **then**
8:             $A := A \backslash \{a\}$
9:         **end if**
10:     **end for**
11:     $A := A \cup \{p\}$
12: **end for**

Lemma 3.1 *Let $A_0, P \subset \mathbb{R}^n$ be finite sets, $\epsilon \in \mathbb{R}^k_+$, and $A := ArchiveUpdate1_\epsilon\ (P, A_0)$. Then the following holds:*

$$\forall x \in P \cup A_0 : \quad \exists a \in A : a \prec_{\epsilon/3} x.$$

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

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

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 $p'_l \prec_{\epsilon/3} x$.
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 $a_j \prec_{\epsilon/3} p_j$ (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 $p'_l \prec_{\epsilon/3} x$ 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*

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

Then an application of Algorithm 1, where $ArchiveUpdate1_\epsilon$ 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 a $l_0 \in \mathbb{N}$ such that A_l is an ϵ-approximate Pareto set for all $l \geq l_0$.

(b) Assume there exists a $l_0 \in \mathbb{N}$ such that A_{l_0} is an ϵ-approximate Pareto set. Then

$$A_l = A_{l_0}, \quad \forall l \geq l_0.$$

Proof:

(a) Since Q is compact and F is continuous it follows that $F\big|_Q$ is uniformly continuous. Hence for $\epsilon/3 \in \mathbb{R}^k_+$ there exists a $\delta > 0$ such that

$$x \prec_{\epsilon/3} y \quad \forall x, y \in Q \text{ with } \|x - y\| < \delta. \tag{3}$$

Define

$$G := \bigcup_{p \in P_Q} B_\delta(p)$$

G is an open cover of P_Q. Since P_Q is compact it follows – due to the theorem of Heine-Borel – that there exists a finite subcover

$$\mathcal{S} := \bigcup_{i=1}^{s} B_\delta(p_i) \supset P_Q, \quad p_i \in P_Q, i = 1, \ldots, s.$$

By (2) it follows that there exist with probability one s numbers $l_1, \ldots, l_s \in \mathbb{N}$ such that each $B_\delta(p_i) \cap Q, i = 1, \ldots, s$, gets 'visited' by *Generate* after l_i iteration steps. That is, $P_{l_i}, i = 1, \ldots, s$, contains with probability one a point $b_i \in B_\delta(p_i) \cap Q$, and thus, A_{l_i} contains with probability one a vector d_i with $d_i \prec_{\epsilon/3} b_i$. By construction of $ArchiveUpdate1_\epsilon$ there exists for all $l \geq l_i$ with probability one a $d_i^l \in A_l$ such that $d_i^l \prec_{\epsilon/3} b_i$ (see Lemma 1). Set $l_0 := \max\{l_1, \ldots, l_s\}$.

Now let $x \in Q$. For x there exists a $p \in P_Q$ such that $F(p) \leq_p F(x)$ and since $\mathcal{S}$ is a cover of P_Q there exists an $i \in \{1, \ldots, s\}$ with $p \in B_\delta(p_i)$. Let l_0, b_i, and d_i^l be as described above and let $l \geq l_0$. Since b_i and p are inside $B_\delta(p_i)$ it follows by (3) that $b_i \prec_{\epsilon/3} p_i$ and $p_i \prec_{\epsilon/3} p$. Hence we have with probability one:

$$d_i^l \prec_{\epsilon/3} b_i \prec_{\epsilon/3} p_i \prec_{\epsilon/3} p, \quad l \geq l_i.$$

Thus, we have that $d_i^l \prec_\epsilon x$, $l \geq l_0$, with probability one as desired.

(b) This follows immediately by the construction of *ArchiveUpdate1*$_\epsilon$ (to be more precise, by lines 3 – 5 of Algorithm 2).

Remarks 3.3 *(a) 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 vectors $x \in Q$. However, this is nearly impossible to check in practise.*

(b) 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 [4] can be used yielding similar results.*

(c) 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

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

In this case, the Pareto set is given by $P = \mathbb{R}_+$. 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.

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 'capture' 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 which is described in Algorithm 3. In the following we will discuss the limit behavior of this approach.

Algorithm 3 $A := \mathit{ArchiveUpdate2}_\epsilon\ (P, A_0)$

```
A := A_0
for all p ∈ P do
    if ∄a ∈ A : a ≺_{ε/3} p then
        A := A ∪ {p}
    end if
    for all a ∈ A do
        if p ≺ a then
            A := A ∪ {p}\{a}
        end if
    end for
end for
```

Lemma 3.4 *Let $A_0, P \subset \mathbb{R}^n$ be finite sets, $\epsilon \in \mathbb{R}^k_+$, and $A := \mathit{ArchiveUpdate2}_\epsilon\ (P, A_0)$. Then the following holds:*

$$\forall x \in P \cup A_0 : \quad \exists a \in A : a \prec_{\epsilon/3} x.$$

Proof: Analogue to the proof of Lemma 3.1.

Theorem 3.5 *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*

$$\forall x \in Q \text{ and } \forall \delta > 0: \qquad P\left(\exists l \in \mathbb{N} \;:\; P_l \cap B_\delta(x) \cap Q \neq \emptyset\right) = 1 \tag{4}$$

Then an application of Algorithm 1, where $ArchiveUpdate2_\epsilon$ 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 a $l_0 \in \mathbb{N}$ such that A_l is an ϵ-approximate Pareto set for all $l \geq l_0$.

(b) There exists with probability one a $l_1 \in \mathbb{N}$ such that

$$|A_{l+1}| = |A_l|, \quad \forall l \geq l_1.$$

(c) The limit archive

$$A_\infty := \lim_{l \to \infty} A_l$$

is an ϵ-Pareto set with probability one.

Proof:

(a) Analogue to the proof of Theorem 3.2 (a).

(b) By (a) it follows that there exists with probability one a $l_0 \in \mathbb{N}$ such that A_{l_0} is an ϵ-approximate Pareto set. Assume that this number l_0 is given. $|A_{l_0}|$ is certainly finite. Further let $l \geq l_0$. By construction of $ArchiveUpdate2_\epsilon$ the archive A_l is also an ϵ-approximate Pareto set. That is, further points are only inserted to the archive if in turn at least one dominated solution is deleted (line 8 of Algorithm 3). Thus it holds that

$$|A_{l+1}| \leq |A_l| \quad \forall l \geq l_0.$$

Since on the other hand $|A_l| \geq 1 \; \forall l \in \mathbb{N}$, the sequence $\{|A_l|\}_{l \in \mathbb{N}}$ of the magnitudes of the archives is bounded below and monotonically decreasing and converges thus to an element $N_A \in \mathbb{N}$. Further, since $|A_l| \in \mathbb{N}, l \in \mathbb{N}$, there exists a $l_1 \in \mathbb{N}$ such that $|A_l| = N_A, \; \forall l \geq l_1$.

(c) By (b) it follows that there exists with probability one a $l_1 \in \mathbb{N}$ such that $|A_{l+1}| = |A_l|, \; \forall l \geq l_1$. Assume that this number l_1 is given. Consider an element $a_0 \in A_l$ with $l \geq l_1$. If $a_0 \in P_Q$ it follows that $a_0 \in A_{l+m}, \; \forall m \in \mathbb{N}$, and thus also $a_0 \in A_\infty$. Assume that $a_0 \notin P_Q$. Define

$$\begin{aligned} M &: Q \to \mathbb{R} \\ M(x) &:= \max_{p \in P_Q} \min_{i=1,\dots,k} (f_i(x) - f_i(p)) \end{aligned} \tag{5}$$

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 no 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 neigborhood U_{p_0} of p_0 such that

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

and thus, that $F(y) <_p F(a_0)$, $\forall y \in U_{p_0}$. By (4) it follows that *Generate* 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 non-dominating (in this case set $a_1 := \tilde{a}$). In both cases there exists a $j \in \{1, \ldots, k\}$ 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 below bounded and F is injective it follows that $a_i \to a^* \in Q$ for $i \to \infty$.
It remains to 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 no 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 $j \in \{1, \ldots, k\}$ such that

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

This is a contradiction to the assumption of the convergence of the sequence, and thus it must be that $a^* \in P_Q \cap A_\infty$. Since $a_0 \in A_l, l \geq l_1$, was chosen arbitrarily it follows that A_∞ is a ϵ-Pareto set and the proof is complete.

4 Bounds on the Archive Sizes

In the following we give bounds on the magnitude of the limit archives A_∞ with respect to $\epsilon \in \mathbb{R}^k_+$ and the chosen archiving strategy.
For this, we have to introduce some notations: denote by m_i and M_i the minimal resp. maximal value of objective f_i, $i = 1, \ldots, k$, inside Q (these values exist since F is continuous and Q is compact). Further, we need k-dimensional boxes, which can be represented by a center $c \in \mathbb{R}^k$ and a radius $r \in \mathbb{R}^k_+$:

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

In the following we assume that $|P_0| = 1$, and thus also $|A_0| = 1$. The lower bound of $|A_\infty|$ for both archiving strategies is obviously given by 1. For this, consider e.g. $f_1 = f_2 = \ldots = f_k$ to be a convex function which takes its (unique) minimum inside Q. The upper bounds for the different archiving strategies are derived separately in the following.

Theorem 4.1 *Let $m_i = \min_{x\in Q} f_i(x)$ and $M_i = \max_{x\in Q} f_i(x), 1 \leq i \leq k$, and $|A_0| = 1$. Then, when using ArchiveUpdate1$_\epsilon$, the archive size maintained in Algorithm 1 for all $l \in \mathbb{N}$ is bounded as*

$$|A_l| \leq \left\lceil \frac{1}{\epsilon_m} \sum_{\substack{i_1,\ldots,i_{k-1}=1\\ i_1>\ldots>k_{k-1}}}^{k} \prod_{j=1}^{k-1} (M_{i_j} - m_{i_j}) \right\rceil, \tag{6}$$

where $\epsilon_m := \min\limits_{i=1,\ldots,k} \frac{\epsilon_i}{3}$.

Proof: Consider a sequence $p_1, p_2, \ldots$ of points which are all accepted by $ArchiveUpdate1_\epsilon$ 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 $B_j := B(F(a_j) - \epsilon/6, \epsilon/6)$, $j = 1, \ldots, l$. Using inductive arguments we see that (a) all elements in A_i are mutually non-dominating, 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 given by ϵ_m (see Figure 1). 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.
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 non-dominating, the images of these points are all located on a (virtual) continuously differentiable curve

$$\begin{aligned} c : [m_1, M_1] &\to \mathbb{R}^2 \\ u &\mapsto (u, f(u)) \end{aligned} \tag{7}$$

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

$$\begin{aligned} L(c) &= \int_{m_1}^{M_1} \|c'(u)\| du = \int_{m_1}^{M_1} \sqrt{|1|^2 + |f'(u)|^2} du \\ &\leq \int_{m_1}^{M_1} 1du + \int_{m_1}^{M_1} |f'(u)| du = \int_{m_1}^{M_1} 1du - \int_{m_1}^{M_1} f'(u) du \\ &\leq (M_1 - m_1) + (M_2 - m_2) \end{aligned} \tag{8}$$

Thus, for $k = 2$ we see that $|A_i| \leq \left\lceil \frac{(M_1-m_1)+(M_2-m_2)}{\epsilon_m} \right\rceil$, $i \in \mathbb{N}$, as claimed above.

Now we turn our attention to the general case, i.e. let $k \geq 2$ be given. Define $K := [m_1, M_1] \times \ldots [m_{k-1}, M_{k-1}]$, $K_{(i)} := [m_1, M_1] \times \ldots \times [m_{i-1}, M_{i-1}] \times [m_{i+1}, M_{i+1}] \times \ldots \times [m_{k-1}, M_{k-1}]$, and $u_{(i)} := (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_{k-1})$, $i = 1, \ldots, k-1$. 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:

$$\begin{aligned} \Phi &: K \to \mathbb{R}^k \\ \Phi(u_i, \ldots, u_{k-1}) &= (u_1, \ldots, u_{k-1}, f(u_1, \ldots, u_{k-1})), \end{aligned} \tag{9}$$

where $f : K \to [m_k, M_k]$ 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:

$$\begin{aligned} Vol(\Phi) &= \int_K \sqrt{||\nabla f||^2 + 1} du = \int_K \sqrt{\left(\frac{\partial f}{\partial u_1}\right)^2 + \ldots + \left(\frac{\partial f}{\partial u_{k-1}}\right)^2 + 1} du \\ &\leq \int_K \left|\frac{\partial f}{\partial u_1}\right| du + \ldots + \int_K \left|\frac{\partial f}{\partial u_{k-1}}\right| du + \int_K 1 du \\ &= \sum_{i=1}^{k-1} \left(\int_{K_{(i)}} \left(\int_{m_i}^{M_i} \left|\frac{\partial f}{\partial u_i}\right| du_i \right) du_{(i)} \right) + \int_K 1 du \\ &= \sum_{i=1}^{k-1} \left(\int_{K_{(i)}} \left(- \int_{m_i}^{M_i} \frac{\partial f}{\partial u_i} du_i \right) du_{(i)} \right) + \int_K 1 du \\ &\leq \sum_{\substack{i_1, \ldots, i_{k-1} = 1 \\ i_1 > \ldots > k_{k-1}}}^{k} \prod_{j=1}^{k-1} (M_{i_j} - m_{i_j}) \end{aligned} \tag{10}$$

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

Theorem 4.2 *Let $m_i = \min_{x \in Q} f_i(x)$ and $M_i = \max_{x \in Q} f_i(x), 1 \leq i \leq k$, and $|A_0| = 1$. Then, when using ArchiveUpdate2$_\epsilon$, the archive size maintained in Algorithm 1 is bounded for all $l \in \mathbb{N}$ as*

$$|A_l| \leq \prod_{i=1}^{k} \left\lceil 3 \frac{M_i - m_i}{\epsilon_i} \right\rceil . \tag{11}$$

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 $V_t := \bigcup_{1 \leq j \leq t} A_j$ be

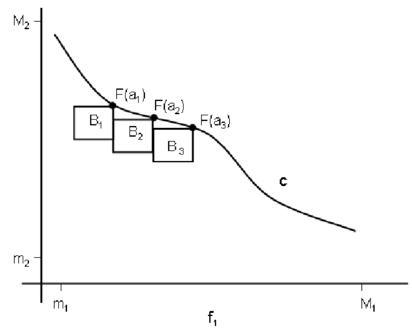

Figure 1: The entries a_i of each archive lie on a (virtual) curve c. Since the boxes B_i (with upper right corners $F(a_i)$) are mutually non-intersecting, it follows that the minimal distance of two entries is given by ϵ_m.

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 elemets of V_t'. Obviously, the number of roots must be smaller than the number of leafs, so $|A_t| \leq |V_t'|$.

To bound $|V_t'|$, the number of elements that ever entered the archive in line 4, we again consider the boxes $B_v := B(F(v) - \epsilon/6, \epsilon/6)$ for all $v \in V_t'$. Due to line 3, a solution p generated in iteration $t' \leq t$ 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 $a \prec_{\epsilon/3} p$ for some current archive member $a \in A_t$ as there exists $a \in A_t$ with $F(a) \leq F(v)$ and $v \prec_{\epsilon/3} p$. 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 $\epsilon/3$ and centers c with $m_i \leq c_i \leq M_i$ is $\prod_{i=1}^{k} \lceil 3(M_i - m_i)/\epsilon_i \rceil$, thus the claimed bound on the archive size follows.

5 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, and have given bounds on the cardinality of the corresponding archives.

For future work, there are a lot of interesting topics which can be addressed to advance the present work. One could e.g. consider the speed of the convergence, in particular when the methods presented above are hybridized with local search strategies. Further, we intend to apply this theoretical framework in search for the development of fast and reliable multi-objective optimization algorithms.

References

[1] Evtushenko, Y. G., Potapov, M. A.: Methods of numerical solution of multicriterion problem. Soviet mathematics – doklady, **34**, 420–423 (1987)

[2] Hanne, T: On the convergence of multiobjective evolutionary algorithms. European Journal Of Operational Research, **117(3)**, 553–564 (1999)

[3] Helbig, S., Pateva, D.: On several concepts for ϵ-efficiency. OR Spektrum, **16(3)**, 179–186 (1994)

[4] Laumanns, M., Thiele, L., Deb, K., Zitzler, E.: Combining convergence and diversity in evolutionary multiobjective optimization. Evolutionary Computation, **10(3)**, 263–282 (2002)

[5] Reuter, H.: An approximation method for the efficiency set of multiobjective programming problems. Optimization, **21**, 905–911 (1990)

[6] Rudolph, G., Agapie, A.: On a multi-objective evolutionary algorithm and its convergence to the Pareto set. Paper presented at the Congress on Evolutionary Computation (CEC'2000), La Jolla Marriott, San Diego, CA, USA, July 16-19, 2000.

[7] Ruhe, G., Fruhwirt, B.: ϵ-optimality for bicriteria programs and its application to minimum cost flows. Computing, **44** 21–34 (1990)

[8] Schütze, O.: Set Oriented Methods for Global Optimization. Ph.D. thesis, University of Paderborn (2004) <http://ubdata.uni-paderborn.de/ediss/17/2004/schuetze/>.