# Computing Finite Size Representations of the Set of Approximate Solutions of an MOP with Stochastic Search Algorithms

Oliver Schuetze
CINVESTAV-IPN
Computer Science
Department
Mexico City 07300,
MEXICO
schuetze@cs.cinvestav.mx

Carlos A. Coello Coello
CINVESTAV-IPN
Computer Science
Department
Mexico City 07300,
MEXICO
ccoello@cs.cinvestav.mx

Emilia Tantar
INRIA Futurs, LIFL, CNRS,
Cite Scientifique
59655 Villeneuve d'Ascq,
FRANCE
Emilia.Tantar@lifl.fr

El-Ghazali Talbi
INRIA Futurs, LIFL, CNRS,
Cite Scientifique
59655 Villeneuve d'Ascq,
FRANCE
talbi@lifl.fr

## ABSTRACT

In this work we study the convergence of generic stochastic search algorithms toward the entire set of approximate solutions of continuous multi-objective optimization problems. Since the dimension of the set of interest is typically equal to the dimension of the parameter space, we focus on obtaining a finite and tight approximation, measured by the Hausdorff distance. 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 investigate a novel archiving strategy theoretically and empirically. For this, we analyze the convergence behavior of the algorithm, yielding bounds on the obtained approximation quality as well as on the cardinality of the resulting approximation, and present some numerical results.

## Categories and Subject Descriptors

F.2.1 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms and Problems; G.1.6 [**Numerical Analysis**]: Optimization

## General Terms

Algorithms, Performance

## Keywords

multi-objective optimization, convergence, Pareto set, approximate solution, stochastic search algorithms

## 1. INTRODUCTION

Since the notion of $\epsilon$-efficiency for multi-objective optimization problems (MOPs) has been introduced more than two decades ago ([6]), this concept has been studied and used by many researchers, e.g. to allow (or tolerate) nearly optimal solutions ([6], [15]), to approximate the set of optimal solutions ([8]), or in order to discretize this set ([5], [11]). $\epsilon$-efficient solutions or approximate solutions have also been used to tackle a variety of real world problems including portfolio selection problems ([16]), a location problem ([1]), or a minimal cost flow problem ([8]).

As an illustrative example where it could make sense from the practical point of view to consider in addition to the exact solutions also approximate ones we consider a plane truss design problem, where the volume of the truss as well as the displacement of the joint to a given position have to be minimized (see also Section 6.2). Since the designs of this problem—as basically in all other engineering problems—have to obey certain physical contraints such as in this case the weight and stability of the structural element, the objective values of all feasible solutions are located within a relatively tight and a priori appreciable range. Hence, the maximal tolerable loss of a design compared to an 'optimal' one with respect to the objective values can easily be determined quantitatively and qualitatively by the decision maker (DM) before the optimization process. The resulting set of exact *and* approximate (but physically relevant) solutions obtained by the optimization algorithm[1] leads in general to a larger variety of possibilities to the DM than 'just' the

---

[1]Here we assume an idealized algorithm, since in practise every solution is an approximate one.

set of exact solutions: this is due to the fact that solutions which are 'near' in objective space can differ significantly in design space (e.g., when the model contains symmetries, or see Section 6.3 for another example).

The computation of such approximate solutions has been addressed in several studies. In most of them, scalarization methods have been empoyed (e.g., [15], [1], [3]). By their nature, such algorithms can deliver only single solutions by one single execution. The only work so far which deals with the approximation of the entire set of approximate solutions (denote by $E_\epsilon$) is [10], where an archiving strategy for stochastic search algorithms is proposed for this task. Such a sequence of archives obtained by this algorithm provably converges—under mild assumptions on the process to generate new candidate solutions—to $E_\epsilon$ in the limit and in the probabilistic sense. This result, though satisfactory for most discrete MOPs, is at least from the practical viewpoint not sufficient for continuous models (i.e., continous objectives defined on a continuous domain): in this case, the set of approximate solutions typically forms an $n$-dimensional object, where $n$ denotes the dimension of the parameter space (see Section 3). Thus, it may come to performance problems since it can easily happen that a given threshold on the magnitude of the archives is exceeded before a 'sufficient' approximation of the set of interest in terms of diversity and/or convergence is obtained. An analogue statement holds for the approximation of the Pareto front, which is 'only' $(k-1)$-dimensional for MOPs with $k$ objectives, and suitable discretizations have been subject of research since several years (e.g., [5], [4], [11]).

The scope of this paper is to extend the work in [10] and to develop a framework for *finite size* representations of the set $E_\epsilon$ with stochastic search algorithms such as evolutionary multi-objective (EMO) algorithms. This will call for the design of a novel archiving strategy to store the 'required' solutions found by the stochastic search process. We will further analyze the convergence behavior of this method, yielding bounds on the approximation quality as well as on the cardinality of the resulting approximations. Finally, we will demonstrate the practicability of the novel approach by several examples.

The remainder of this paper is organized as follows: in Section 2, we state the required background including the set of interest $P_{Q,\epsilon}$ which we analyze in Section 3. In Section 4, we propose a novel archiving strategy for the approximation of $P_{Q,\epsilon}$ and state a convergence result, and give further on an upper bound on the resulting archive sizes in Section 5. In Section 6, we present some numerical results, and make finally some conclusions in Section 7.

## 2. BACKGROUND

In the following we consider continuous multi-objective optimization problems

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

where $Q \subset \mathbb{R}^n$ and $F$ is defined as the vector of the objective functions $F : Q \to \mathbb{R}^k, \quad F(x) = (f_1(x), \dots, f_k(x))$, and where each $f_i : Q \to \mathbb{R}$ is continuously differentiable. Later we will restrict the search to a compact set $Q$, the reader may think of an $n$-dimensional box.

**Def 2.1** (a) *Let $v, w \in Q$. Then the vector $v$ is* less than

$w$ ($v <_p w$), if $v_i < w_i$ for all $i \in \{1, \dots, k\}$. The relation $\leq_p$ is defined analogously.

(b) *$y \in Q$ is* dominated *by a point $x \in Q$ ($x \prec y$) with respect to (MOP) if $F(x) \leq_p F(y)$ and $F(x) \neq F(y)$.*

(c) *$x \in Q$ is called a* Pareto optimal point *or* Pareto point *if there is no $y \in Q$ which dominates $x$.*

The set of all Pareto optimal solutions is called the *Pareto set* (denoted by $P_Q$). This set typically —i.e., under mild regularity assumptions—forms a $(k-1)$-dimensional object. The image of the Pareto set is called the *Pareto front*.

We now define another notion of dominance, which is the basis of the approximation concept used in this study.

**Def 2.2** *Let $\epsilon = (\epsilon_1, \dots, \epsilon_k) \in \mathbb{R}^k_+$ and $x, y \in Q$.*

(a) *$x$ is said to $\epsilon$-dominate $y$ ($x \prec_\epsilon y$) with respect to (MOP) if $F(x) - \epsilon \leq_p F(y)$ and $F(x) - \epsilon \neq F(y)$.*

(b) *$x$ is said to $-\epsilon$-dominate $y$ ($x \prec_{-\epsilon} y$) with respect to (MOP) if $F(x) + \epsilon \leq_p F(y)$ and $F(x) + \epsilon \neq F(y)$.*

The notion of $-\epsilon$-dominance is of course analogous to the 'classical' $\epsilon$-dominance relation but with a value $\tilde{\epsilon} \in \mathbb{R}^k_-$. However, we highlight it here since it will be used frequently in this work. While the $\epsilon$-dominance is a weaker concept of dominance, $-\epsilon$-dominance is a stronger one.

We now define the set which we want to approximate in the sequel (see Fig. 1 for two examples).

**Def 2.3** *Denote by $P_{Q,\epsilon}$ the set of points in $Q \subset \mathbb{R}^n$ which are not $-\epsilon$-dominated by any other point in $Q$, i.e.*

$$P_{Q,\epsilon} := \{x \in Q | \, \nexists y \in Q : y \prec_{-\epsilon} x\}. \qquad (1)$$

To see that $P_{Q,\epsilon}$ typically forms an $n$-dimensional set let $x_0 \in P_Q$ (such a point, for instance, always exists if $Q$ is compact). That is, there exists no $y \in Q$ such that $y \prec x_0$. Since $F$ is continuous and $\epsilon \in \mathbb{R}^k_+$ there exists a neigborhood $U$ of $x_0$ such that

$$\nexists y \in Q : \, y \prec_{-\epsilon} u \quad \forall u \in U \cap Q, \qquad (2)$$

and thus, $U \cap Q \subset P_{Q,\epsilon}$, and we are done since $U$ is $n$-dimensional.

The following result and notions are used for the upcoming proof of convergence.

**Theorem 2.4 ([9])** *Let (MOP) be given and $q : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $q(x) = \sum_{i=1}^k \hat{\alpha}_i \nabla f_i(x)$, where $\hat{\alpha}$ is a solution of*

$$\min_{\alpha \in \mathbb{R}^k} \left\{ \left\| \sum_{i=1}^k \alpha_i \nabla f_i(x) \right\|_2^2 ; \alpha_i \geq 0, i = 1, \dots, k, \sum_{i=1}^k \alpha_i = 1 \right\}.$$

*Then either $q(x) = 0$ or $-q(x)$ is a descent direction for all objective functions $f_1, \dots, f_k$ in $x$. Hence, each $x$ with $q(x) = 0$ fulfills the first-order necessary condition for Pareto optimality.*

**Def 2.5** *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:*

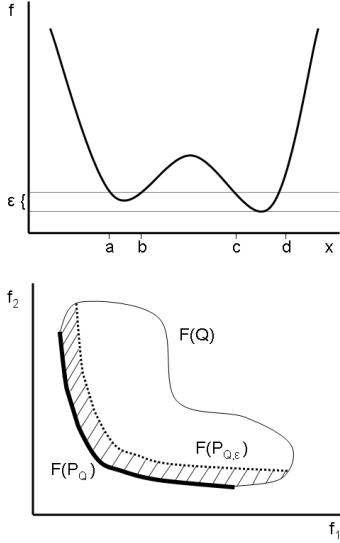**Figure 1: Two different examples for sets $P_{Q,\epsilon}$. Above for $k=1$ and in parameter space with $P_{Q,\epsilon} = [a,b] \cup [c,d]$. Below an example for $k=2$ in image space.**

(a) $dist(u, A) := \inf\limits_{v \in A} \|u - v\|$

(b) $dist(B, A) := \sup\limits_{u \in B} \; dist(u, A)$

(c) $d_H(A, B) := \max\{dist(A,B), dist(B,A)\}$

**Def 2.6** *A set $S \subset \mathbb{R}^n$ is called* not connected *if there exist open sets $O_1, O_2$ such that $S \subset O_1 \cup O_2$, $S \cap O_1 \neq \emptyset$, $S \cap O_2 \neq \emptyset$, and $S \cap O_1 \cap O_2 = \emptyset$. Otherwise, $S$ is called* connected.

Denote by $\overline{A}$ the closure of a set $A \subset \mathbb{R}^n$, by $\overset{\circ}{A}$ its interior, and by $\partial A = \overline{A} \backslash \overset{\circ}{A}$ the boundary of $A$.

**Def 2.7** (a) *A point $x \in Q$ is called a* weak Pareto point *if there exists no point $y \in Q$ such that $F(y) <_p F(x)$.*

(b) *A point $x \in Q$ is called $-\epsilon$ weak Pareto point if there exists no point $y \in Q$ such that $F(y) + \epsilon <_p F(x)$.*

Algorithm 1 gives a framework of a generic stochastic multi-objective optimization algorithm, which will be 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.

---

**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:      $P_{j+1} = Generate(P_j)$
5:      $A_{j+1} = ArchiveUpdate(P_{j+1}, A_j)$
6: **end for**

---

## 3. THE ALGORITHM

Here we present and analyze a novel archiving strategy which aims for a finite size representation of $P_{Q,\epsilon}$.

The algorithm which we propose here, $ArchiveUpdateP_{Q,\epsilon}$, is given in Algorithm 2. Hereby is $1\Delta := (\Delta, \ldots, \Delta) \in \mathbb{R}^k_+$, where $\Delta \in \mathbb{R}_+$, and $d_\infty$ the infinity norm distance. As discussed in Section 1, the scope is to prevent that all points of $P_{Q,\epsilon}$ are taken to the archive. The 'exclusion strategy' is realized in line 3 of Algorithm 2: if there already exists a point $a_2$ in the current archive such that its distance to the candidate solution $p$ is less or equal to $\Delta^* < \Delta$ (measured in image space), the candidate $p$ will be discarded from the archive. Thus, $\Delta$ can be viewed as the discretization parameter of the algorithm or the minimal spread (both in image space) of the archive.

Next we show the convergence of the sequence of archives when using this algorithm under certain assumptions.

---

**Algorithm 2** $A := ArchiveUpdateP_{Q,\epsilon} \; (P, A_0, \Delta)$

**Require:** population $P$, archive $A_0$, $\Delta \in \mathbb{R}_+$, $\Delta^* \in (0, \Delta)$
**Ensure:** updated archive $A$
1: $A := A_0$
2: **for all** $p \in P$ **do**
3:      **if** $\nexists a_1 \in A : a_2 \prec_{-\epsilon} p$ and $\nexists a_2 \in A : d_\infty(F(a_2), F(p)) \leq \Delta^*$ **then**
4:          $A := A \cup \{p\}$
5:          **for all** $a \in A$ **do**
6:              **if** $p \prec_{-(\epsilon+1\Delta)} a$ **then**
7:                  $A := A \backslash \{a\}$
8:              **end if**
9:          **end for**
10:      **end if**
11: **end for**

---

**Theorem 3.1** *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 $\epsilon \in \mathbb{R}^k_+$, $\Delta, \Delta^* \in \mathbb{R}_+$ with $\Delta^* < \Delta$. For the generation process—i.e., Generate() in Algorithm 1—we assume*

$$\forall x \in Q \; and \; \forall \delta > 0 : \quad P(\exists l \in \mathbb{N} \; : \; P_l \cap B_\delta(x) \cap Q \neq \emptyset) = 1, \tag{3}$$

*where $B_\delta(x_0) := \{x \in \mathbb{R}^n \; : \; \|x - x_0\| < \delta\}$, $P_l$ as in Alg. 1, $P(A)$ the probability for event $A$, and for the MOP*

*(A1) Let there be no weak Pareto point in $Q \backslash P_Q$.*

*(A2) Let there be no $-\epsilon$ weak Pareto point in $Q \backslash \overline{P_{Q,\epsilon}}$,*

*(A3) Define $\mathcal{B} := \{x \in Q | \exists y \in P_Q \; : \; F(y) + \epsilon = F(x)\}$. Let $\mathcal{B} \subset \overset{\circ}{Q}$ and $q(x) \neq 0$ for all $x \in \mathcal{B}$, where $q$ is as defined in Theorem 2.4.*

*Then, an application of Algorithm 1, where $ArchiveUpdateP_{Q,\epsilon}(P, A, \Delta)$ is used to update the archive, leads to a sequence of archives $A_l, l \in \mathbb{N}$, where the following holds:*

*(a) For all $l \in \mathbb{N}$ it holds*

$$\|F(a_1) - F(a_2)\|_\infty \geq \Delta^*, \; \forall a_1, a_2 \in A_l : a_1 \neq a_2, \; \forall l \in \mathbb{N}.$$

*(b) There exists with probability one an $l_0 \in \mathbb{N}$ such that for all $l \geq l_0$:*

(b1) $dist(F(P_{Q,\epsilon}), F(A_l)) < \Delta$

(b2) $dist(F(A_l), F(P_{Q,\epsilon})) \leq dist(F(P_{Q,\epsilon+2\Delta}), F(P_{Q,\epsilon}))$

(b3) $d_H(F(P_{Q,\epsilon}), F(A_l)) \leq D$, where
$\quad D = max(\Delta, dist(F(P_{Q,\epsilon+2\Delta}), F(P_{Q,\epsilon})))$

PROOF. Before we state the proof we have to make some remarks: a point $p$ is discarded from an existing archive $A$ in two cases (see line 3 of Algorithm 2):

$$\begin{aligned} (D1) \quad & \exists a_1 \in A : a_1 \prec_{-\epsilon} p, \quad or \\ (D2) \quad & \exists a_2 \in A : \|F(a_2) - F(p)\|_\infty \leq \Delta^*. \end{aligned} \quad (4)$$

Further, we define by

$$B_\delta^\infty(x) := \{y \in \mathbb{R}^k : \|x - y\|_\infty < \delta\}$$

a $k$-dimensional open box around $x \in \mathbb{R}^k$. Now we are in the position to state the proof.

*Claim (a)*: follows immediately by construction of the algorithm and by an inductive argument.

*Claim (b1)*: By *(a)* it follows that for an element $a$ from a given archive $A$ it holds

$$F(\tilde{a}) \notin B_{\Delta^*}^\infty(F(a)), \quad \forall \tilde{a} \in A\backslash\{a\}, \quad (5)$$

Since further $Q$ is compact and $F$ is continuous it follows that $F(Q)$ is bounded, and thus, there exits an upper bound for the number of entries in the archive for a given MOP, denoted by $n_0 = n_0(\Delta^*, F(Q))$ (see also next section).
Since $\overline{P_{Q,\epsilon}}$ is compact and
$dist(F(P_{Q,\epsilon}), F(A_l)) = dist(F(\overline{P_{Q,\epsilon}}), F(A_l))$, and since $A_l, l \in \mathbb{N}$, is finite it follows that

$$dist(F(P_{Q,\epsilon}), F(A_l)) = \max_{p \in P_{Q,\epsilon}} \min_{a \in A_l} \|F(p) - F(a)\|_\infty$$

That is, the claim is right for an archive $A_l, l \in \mathbb{N}$, if for every $p \in P_{Q,\epsilon}$ there exists an element $a \in A_l$ such that $\|F(p) - F(a)\|_\infty < \Delta$. Thus, $F(P_{Q,\epsilon})$ must be contained in $C_{A_l,\Delta}$, where

$$C_{A,\Delta} := \bigcup_{a \in A} B_\Delta^\infty(F(a)).$$

First we show that if there exists an $l_0 \in \mathbb{N}$ with $dist(F(P_{Q,\epsilon}), F(A_l)) < \Delta$, this property holds for all $l \geq l_0$. Assume that such an $l_0$ is given. Define

$$\tilde{A} := \{a \in A_{l_0} | \exists p \in P_{Q,\epsilon} : \|F(p) - F(a)\|_\infty < \Delta\} \quad (6)$$

Since it holds that

$$p \in P_{Q,\epsilon} \text{ and } a \in Q : \|F(p) - F(a)\| \leq \Delta \Rightarrow a \in P_{Q,\epsilon+1\Delta}$$

it follows that $\tilde{A} \subset P_{Q,\epsilon+1\Delta}$, and thus, no element $a \in \tilde{A}$ will be discarded further on due to the construction of $ArchiveUpdateP_{Q,\epsilon}$. Since $dist(F(P_{Q,\epsilon}), F(A_l)) < \Delta$ it follows that for all $p \in P_{Q,\epsilon}$ there exists an element $a \in \tilde{A}$ such that $\|F(p) - F(a)\|_\infty < \Delta$. By the above discussion this holds for all $l \geq l_0$, and since no element $a \in \tilde{A}$ is discarded during the run of the algorithm, and the claim follows.
It remains to show the existence of such an integer $l_0$, which we will do by contradiction: first we show that by using $ArchiveUpdateP_{Q,\epsilon}$ and under the assumptions made above only finitely many replacements can be done during the run of the algorithm. Then we construct a contradiction by showing that under the asssumptions made above infinitely many replacements have to be done during the run of the algorithm with the given setting.

Let a finite archive $A_0$ be given. If a point $p \in \mathbb{R}^n$ replaces a point $a \in A_0$ (see lines 4 and 7 of Algorithm 2) it follows by construction of $ArchiveUpdateP_{Q,\epsilon}$ that

$$F(p) <_p F(a) - \Delta \quad (7)$$

Since the relation '$\prec$' is transitive, there exists for every $a \in A$ a 'history' of replaced points $a_i \in A_{l_i}$ where equation (7) holds for $a_i$ and $a_{i-1}$. Since $F(Q)$ is bounded there exist $l_i, u_i \in \mathbb{R}, i = 1, \ldots, k$, such that $F(Q) \subset [l_1, u_1] \times \ldots \times [l_k, u_k]$. After $r$ replacements there exists at least one $a \in A_{l(r)}$ such that the length $h$ of the history of $a$ is at least $h \geq \lceil r/n_0 \rceil$, where $n_0$ is the maximal number of entries in the archive (see above). Denote by $a_0 \in A_0$ the root of the history of $a$. For $a, a_0$ it follows that

$$F(a) \leq F(a_0) - h\Delta$$

For $\tilde{h} > h_{max} := \Delta^{-1} \max_{i=1,\ldots,k} u_i - l_i$ we obtain a contradiction since in that case there exists $i \in \{1, \ldots, n\}$ with $f_i(a) < l_i$ and thus $F(a) \notin F(Q)$. Hence it follows that there can be done only finitely many such replacements during the run of an algorithm.
Assume that such an integer $l_0$ as claimed above does not exist, that is, that $F(P_{Q,\epsilon}) \not\subset C_{A_l,\Delta}$ for all $l \in \mathbb{N}$. Hence there exists a sequence of points

$$p_i \in P_{Q,\epsilon} : \quad y_i = F(p_i) \in F(P_{Q,\epsilon})\backslash C_{A_i,\Delta} \quad \forall i \in \mathbb{N}. \quad (8)$$

Since $P_{Q,\epsilon} \subset Q$ and $Q$ is compact there exists an accumulation point $p^* \in \overline{P_{Q,\epsilon}}$, that is, there exists a subsequence $\{i_j\}_{j \in \mathbb{N}}$ with

$$p_{i_j} \to p^* \in \overline{P_{Q,\epsilon}} \text{ for } j \to \infty. \quad (9)$$

In [10] it was shown that under the assumptions (A1)–(A3) it follows that

$$\overline{\overset{\circ}{P_{Q,\epsilon}}} = \overline{P_{Q,\epsilon}}, \quad (10)$$

i.e., that $P_{Q,\epsilon}$ is not 'flat' anywhere. Hence, the set

$$\tilde{U}_1 := B_{(\Delta-\tilde{\Delta})/2}^\infty(y^*) \cap \overset{\circ}{P_{Q,\epsilon}}, \quad (11)$$

where $y^* := F(p^*)$, is not empty. By (3) it follows that there exists with probability one an $l_1 \in \mathbb{N}$ and an element $\tilde{x}_1 \in P_{l_0+l_1}$ generated by Generate() with $\tilde{y}_1 = F(\tilde{x}_1) \in \tilde{U}_1$. There are two cases for the archive $A_{l_0+l_1}$: (a) $x_1$ can be discarded from the archive, or (b) $x_1$ is added to it. Assume first that $x_1$ is discarded. Since $x_1 \in P_{Q,\epsilon}$ there exists no $\bar{x} \in Q$ such that $\bar{x} -\epsilon$-dominates $x_1$. Hence, (D1) can not occur (see (4)), and thus, there must exist an $a_2 \in A_{l_0+l_1}$ such that $\|F(a_2) - F(x_1)\|_\infty \leq \Delta^*$ (see (D2)). Thus, whether $x_1$ is added to the archive or not there exists an $\tilde{a}_1 \in A_{l_0+l_1}$ such that $\|F(\tilde{a}_1 - y^*\|_\infty \leq \Delta$ (since in case $x_1$ is added to the archive $\tilde{a}_1 = x_1$ can be chosen), and we obtain

$$\begin{aligned} \|F(\tilde{a}_1) - \tilde{y}\|_\infty &\leq \|F(\tilde{a}_1) - F(x_1)\|_\infty + \|F(x_1) - \tilde{y}\|_\infty \\ &< \Delta \quad \forall \tilde{y} \in U_1 \end{aligned} \quad (12)$$

By (8) and (9) there exist integers $j_1, \tilde{l}_1 \in \mathbb{N}$ with

$$y_{i_{j_1}} \in \tilde{U}_1\backslash C_{l_0+l_1+\tilde{l}_1,\Delta}. \quad (13)$$

Since by (12) it holds that $\|y_{i_{j_1}} - F(a_1)\|_\infty < \Delta$ it follows that $a_1 \notin A_{l_0+l_1+\tilde{l}_1}$, which is only possible via a replacement in Algorithm 2 (lines 4 and 7).
In an analogous way a sequence $\{a_i\}_{i \in \mathbb{N}}$ of elements can

be constructed which have to be replaced by other elements. Since this leads to a sequence of infinitely many replacements. This is a contradiction to the assumption, and the proof is complete.

*Claim (b2):* Let $\tilde{A}$ and $l_0$ as above, and let $l \geq l_0$. Further, let $x \in Q \backslash P_{Q,\epsilon+2\Delta}$, that is, there exists a $p \in P_{Q,\epsilon}$ such that $p \prec_{-(\epsilon+2\Delta)} x$. Since $l \geq l_0$ there exists an $a \in \tilde{A} \subset A_l$ such that $\|F(p) - F(a)\|_\infty < \Delta$. Combining both facts we see that $a \prec_{-(\epsilon+1\Delta)} x$. Thus, no element $x \in Q \backslash P_{Q,\epsilon+2\Delta}$ is contained in $A_l, l \geq l_0$, or will ever be added to the archive further on. The claim follows since the archive can only contain elements in $P_{Q,\epsilon+2\Delta}$ (see also Examples 3.3 and 3.4).

*Claim (b3):* follows immediately by (b1) and (b2). $\square$

**Remarks 3.2** *(a) For $\Delta = \Delta^* = 0$ the archiver coincides with the one proposed in [10], which reads as*

$$UpdateP_{Q,\epsilon}(P, A) := \{x \in P \cup A : y \nprec_{-\epsilon} x \; \forall y \in P \cup A\}. \tag{14}$$

*(b) The convergence result holds for a scalar $\Delta_0 \in \mathbb{R}_+$ which is used for the discretization of the $\epsilon$-efficient front. However, analogue results can be obtained by using a vector $\Delta \in \mathbb{R}_+^k$. In this case, the exclusion strategy in line 3 of Algorithm 2 has to be replaced by*

$$\nexists a_2 \in A : \quad F(p) \in B(F(a_2), \Delta), \tag{15}$$

*where*

$$B(y, \Delta) := \{x \in \mathbb{R}^k : |x_i - y_i| \leq \Delta_i, \; i = 1, .., k\}.$$

*Further, elements a have to be discarded from the archive if they are $-(\epsilon + \Delta)$ dominated by p (lines 6-8).*

*(c) The parameter $\Delta^* \in \mathbb{R}_+$ with $\Delta^* < \Delta$ is used for theoretical purposes. In practise, $\Delta^* = \Delta$ can be chosen.*

*(d) Note that the convergence result also holds for discrete models. In that case, assumption (3) can be modified using Markov chains such that it can easier be verified (see e.g., [7]).*

The next two examples show that with using $ArchiveUpdateP_{Q,\epsilon}$ one cannot prevent to maintain points $x \in P_{Q,\epsilon+2\Delta} \backslash P_{Q,\epsilon}$ in the limit archive, and that the distance between $F(P_{Q,\epsilon+1\Delta})$ (respectively $F(P_{Q,\epsilon+2\Delta})$) and $F(P_{Q,\epsilon})$—i.e., the maximal approximation error of the image of the limit archive and the $\epsilon$-efficient front $F(P_{Q,\epsilon})$, see part (b3) of the theorem above—can get large in some (pathological) examples.

**Example 3.3** *Consider the following MOP:*

$$F : \mathbb{R} \to \mathbb{R}, \qquad F(x) = x \tag{16}$$

*Let $Q = [0, 5]$, $\epsilon = 1, \Delta = 0.1$, and let $\Delta^* = \Delta$. Thus, we have $P_{Q,\epsilon} = [0, 1]$. Assume that $A = \{a_1\}$ with $a_1 = 1.2$. If next $a_2 = 0.1$ is considered, it will be inserted into the archive since $d_\infty(F(a_1), F(a_2)) > \Delta$ and since $a_2 \in P_{Q,\epsilon}$ is not $-\epsilon$-dominated by $a_1$ nor by any other point $x \in Q$, and will thus remain in the archive further on. Since $a_2$ is not $-(\epsilon+\Delta)$-dominating $a_1$ we have for the updated archive $A = \{a_1, a_2\}$. Hence, no element $a \in [0, \Delta]$ will be taken to the archive since for these points it holds $d_\infty(F(a), F(a_2)) \leq \Delta^*$, and thus, $a_2 \in P_{Q,\epsilon+2\Delta} \backslash P_{Q,\epsilon}$ will not be discarded from the archive during the run of the algorithm.*
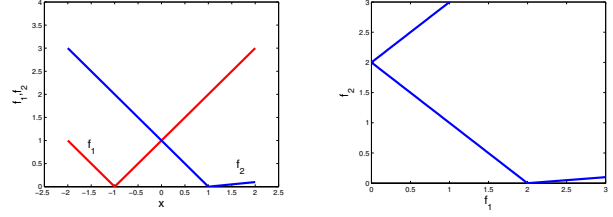


(a) Parameter space     (b) Image space

**Figure 2: Example of MOP (17) for $\alpha = 0.1$.**

**Example 3.4** *Let the MOP be given by $F : \mathbb{R} \to \mathbb{R}^2$, where*

$$f_1(x) = |x+1|, \qquad f_2(x) = \begin{cases} |x - 1| & for \; x \leq 1 \\ \alpha|x - 1| & for \; x > 1 \end{cases}, \tag{17}$$

*where $\alpha \in (0, 1)$ (see also Figure 2). For simplicity we assume that $\epsilon = (\bar{\epsilon}, \bar{\epsilon}) \in \mathbb{R}_+^2$. It is $P_Q = [-1, 1]$ with*

$$F(-1) = (0, 2), \qquad F(1) = (2, 0) \tag{18}$$

*Further, it is*

$$F(-1 - \bar{\epsilon}) = (\bar{\epsilon}, 2 + \bar{\epsilon}), \qquad F(1 + \frac{\bar{\epsilon}}{\alpha}) = (2 + \frac{\bar{\epsilon}}{\alpha}, \bar{\epsilon}) \tag{19}$$

*Using this and some monoticity arguments on $f_1$ and $f_2$ we see that*

$$P_{Q,\epsilon} = \left(-1 - \bar{\epsilon}, 1 + \frac{\bar{\epsilon}}{\alpha}\right] \tag{20}$$

*Since $F(1 + \frac{\bar{\epsilon}+\Delta}{\alpha}) = (2 + \frac{\bar{\epsilon}+\Delta}{\alpha}, \bar{\epsilon} + \Delta)$ it follows that*

$$dist(F(P_{Q,\epsilon+1\Delta}), F(P_{Q,\epsilon})) = \frac{\Delta}{\alpha}, \tag{21}$$

*which can get large for small values of $\alpha$.*

# 4. BOUNDS ON THE ARCHIVE SIZES

Here we give an upper bound $U$ on the size of the limit archive obtained by the novel strategy, and discuss that the order of $U$ is already optimal.

**Theorem 4.1** *Let $\epsilon \in \mathbb{R}_+^k$, $\Delta^*, \Delta \in \mathbb{R}_+$ with $\Delta^* < \Delta$ be given. Further 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 $l_0$ as in Theorem 3.1. Then, when using $ArchiveUpdateP_{Q,\epsilon}$, the archive size maintained in Algorithm 1 for all $l \geq l_0$ is bounded as*

$$|A_l| \leq \left(\frac{1}{\Delta^*}\right)^k \sum_{i=1}^{k}(\epsilon_i + 2\Delta + \Delta^*) \prod_{\substack{j=1 \\ j \neq i}}^{k}(M_j - m_j + \Delta^*). \tag{22}$$

PROOF. Let $l \geq l_0$ and the archive $A_l$ be given. Since $A_l \subset P_{Q,\epsilon+2\Delta}$ (see Theorem 3.1) we are interested in an upper bound on the volume of $F(P_{Q,\epsilon+2\Delta})$. For this, we consider first the $(k-1)$-dimensional volume of the Pareto front $F(P_Q)$. Due to the nature of nondominance we can assume that $F(P_Q)$ is located in the graph of a map

$$\Phi_f : K \to \mathbb{R}^k$$
$$\Phi_f(u_1, \ldots, u_{k-1}) = (u_1, \ldots, u_{k-1}, f(u_1, \ldots, u_{k-1})), \tag{23}$$

where $K := [m_1, M_1] \times \ldots \times [m_{k-1}, M_{k-1}]$ and $f : K \to [m_k, M_k]$. Analogue to [11] one can bound the $(k-1)$-dimensional volume of $\Phi_f$ with parameter range $K$ as follows:

$$Vol_{k-1}(\Phi_f) = \int_K \sqrt{||\nabla f||^2 + 1}\, du \leq \sum_{i=1}^{k} \prod_{\substack{j=1 \\ j \neq i}}^{k} (M_j - m_j), \tag{24}$$

where $\nabla f$ denotes the gradient of $f$. Considering this and the nature of $-\epsilon$-dominance we can bound the $k$-dimensional volume of $F(P_{Q,\epsilon+2\Delta})$ by:

$$Vol_k(F(P_{Q,\epsilon+2\Delta})) \leq \sum_{i=1}^{k}(\epsilon_i + 2\Delta) \prod_{\substack{j=1 \\ j \neq i}}^{k}(M_j - m_j), \tag{25}$$

Since $||F(a_1) - F(a_2)|| \geq \Delta^*$ for all $a_1, a_2 \in A_l$ it follows that the boxes

$$B_{\frac{1}{2}\Delta^*}^{\infty}(F(a)), \quad a \in A_l, \tag{26}$$

are mutually nonoverlapping. Further, if $F(a) \in F(P_{Q,\epsilon+2\Delta})$, then $B_{\frac{1}{2}\Delta^*}^{\infty}(F(a))$ is included in a $\Delta^*/2$-neighborhood $\tilde{F}$ of $F(P_{Q,\epsilon+2\Delta})$ with

$$Vol_k(\tilde{F}) \leq \sum_{i=1}^{k}(\epsilon_i + 2\Delta + \Delta^*) \prod_{\substack{j=1 \\ j \neq i}}^{k}(M_j - m_j + \Delta^*). \tag{27}$$

The maximal number of entries in $A_l$ can now be estimated by

$$|A_l| \leq \frac{Vol_k(\tilde{F})}{Vol_k(B_{\frac{1}{2}\Delta^*}^{\infty}(F(a)))}, \tag{28}$$

and the claim follows since the volume of $B_{\frac{1}{2}\Delta^*}^{\infty}(F(a))$ is obviously given by $(\Delta^*)^k$. $\square$

In particular interesting is certainly the growth of the magnitudes of the (limit) archives for vanishing discretization parameter $\Delta$. Since for every meaningful computation the value $\Delta$ will be smaller than every entry of $\epsilon$, we can assume $\epsilon_i = c_i \Delta$ with $c_i > 1$. Using (22) and for simplicity $\Delta = \Delta^*$ we see that

$$|A_l| \leq \left(\frac{1}{\Delta}\right)^{k-1} \sum_{i=1}^{k}(c_i + 3) \prod_{\substack{j=1 \\ j \neq i}}^{k}(M_j - m_j + \Delta^*). \tag{29}$$

Thus, the growth of the magnitudes is of order $\mathcal{O}\left(\left(\frac{1}{\Delta}\right)^{k-1}\right)$ for $\Delta \to 0$. Regarding the fact that $P_Q$, which is contained in $P_{Q,\epsilon}$ for all values of $\epsilon \in \mathbb{R}_+^k$, typically forms a $(k-1)$-dimensional object, we see that the order of the magnitude of the archive with respect to $\Delta$ is already optimal: the cost for the approximation of $F(P_{Q,\epsilon})$ in terms of the number of elements which have to be stored in the archive is—at least from the theoretical viewpoint—as expensive as the 'classical' problem of approximating the Pareto front. This is due to the fact that the discretization (line 3 of Algorithm 2) is realized in image space. An analogue result for a discretization in parameter space, however, can not hold since $P_{Q,\epsilon}$ is $n$-dimensional.

**Table 1: Comparison of the magnitudes of the final archive ($|A_{final}|$, rounded) and the corresponding update times ($T$, in seconds) for MOP (30) and for different values of $\Delta$. We have taken the average result of 100 test runs.**

| $\Delta$ | $|A_{final}|$ | $T$ |
|----------|---------------|-------|
| 0 | 3836 | 32.98 |
| 0.01 | 827 | 6.22 |
| 0.05 | 68 | 1.80 |

## 5. NUMERICAL RESULTS

Here we demonstrate the practicability of the novel archiver on three examples. For this, we run and compare $ArchiveUpdateP_{Q,\epsilon}$ for different values of $\Delta$ including $\Delta_0 = 0$, which is the archiver proposed in [10] which accepts all test points which are not $-\epsilon$ dominated by any other test point (see Remark 4.3 (a)). To obtain a fair comparison we have decided to take a random search operator for the generation process (the same sequence of points for all settings).

### 5.1 Example 1

First we consider the MOP suggested by Tanaka [13]:

$$F : \mathbb{R}^2 \to \mathbb{R}^2, \quad F(x_1, x_2) = (x_1, x_2) \tag{30}$$

where

$$C_1(x) = x_1^2 + x_2^2 - 1 - 0.1 \cos(16 \arctan(x_1/x_2)) \geq 0$$
$$C_2(x) = (x_1 - 0.5)^2 + (x_2 - 0.5)^2 \leq 0.5$$

Figure 3 shows a numerical result for $N = 200,000$ randomly chosen points within $Q = [0, \pi]^2$ and for three different values of the discretization parameter: $\Delta_0 = 0$, $\Delta_1 = 0.01$ and $\Delta_2 = 0.05$. As anticipated, the granularity of the resulting archive increases with the value of $\Delta$. Thus, the approximation quality decreases, but, in turn, the running time of the algorithm decreases (see Table 1).

### 5.2 Example 2

Next, we consider a real-life engineering problem, namely the design of a four-bar plane truss [12]:

$$F : \mathbb{R}^4 \to \mathbb{R}^2$$
$$f_1(x) = L(2x_1 + \sqrt{2}x_2 + \sqrt{2}x_3 + x_4)$$
$$f_2(x) = \frac{FL}{E}\left(\frac{2}{x_1} + \frac{2\sqrt{2}}{x_2} - \frac{2\sqrt{2}}{x_3} + \frac{1}{x_4}\right) \tag{31}$$

$f_1$ models the volume of the truss, and $f_2$ the displacement of the joint. The model constants are the length $L$ of each bar ($L = 200$ cm), the elasticity constants $E$ and $\sigma$ of the materials involved ($E = 2 \times 10^5$ kN/cm$^3$, $\sigma = 10$ kN/cm$^2$), and the force $F$ which causes the stress of the truss ($F = 10$ kN). The parameters $x_i$ represent the cross sectional areas of the four bars of the truss. The physical restrictions are given by

$$Q = [F/\sigma, 3F/\sigma] \times [\sqrt{2}F/\sigma, 3F/\sigma]^2 \times [F/\sigma, 3F/\sigma] \tag{32}$$
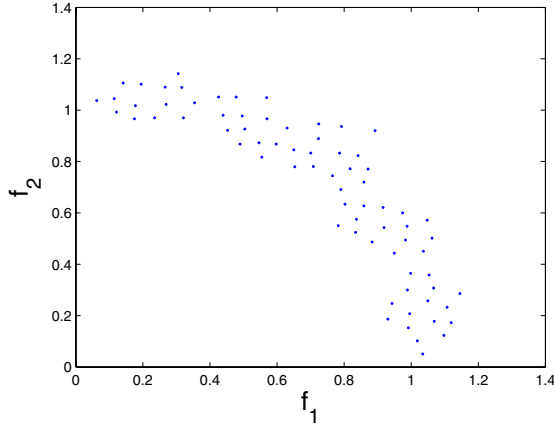
For the allowed tolerances we follow the suggestion made in [3] and set $\epsilon_1 = 50$ cm$^3$ and $\epsilon_2 = 0.0005$ cm. Figure 4 shows a result for $N = 500,000$ randomly chosen points within $Q$ and for $\Delta = (10, 0.0001)$, i.e., $\Delta_i = \epsilon_i/5$ (see Remark 4.3 (b)). The final archive consists of 78 elements, and the
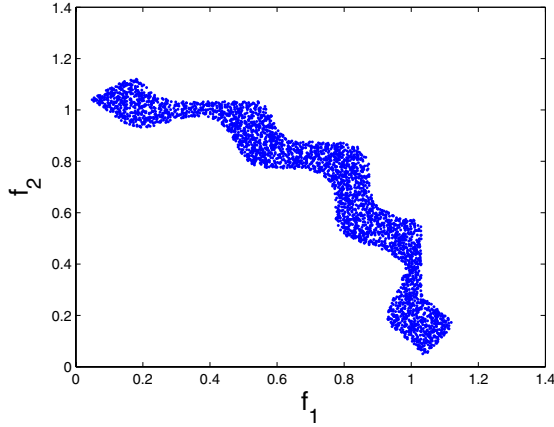
(a) $\Delta_0 = 0$,    $|A_{final}| = 3824$
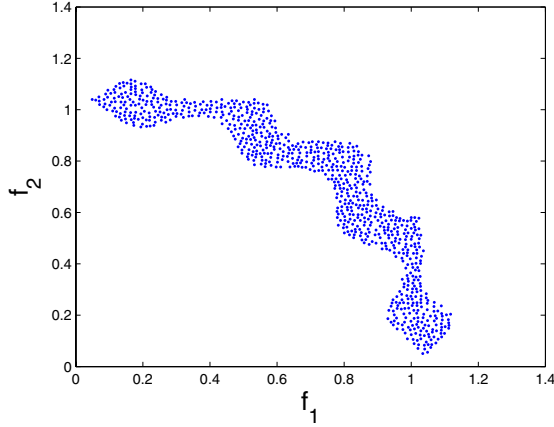


(b) $Delta_1 = 0.01$,    $|A_{final}| = 834$



(c) $Delta_2 = 0.05$,    $|A_{final}| = 73$

**Figure 3: Results for MOP (30) for different values of $\Delta$ leading to different granularities of the approximation.**
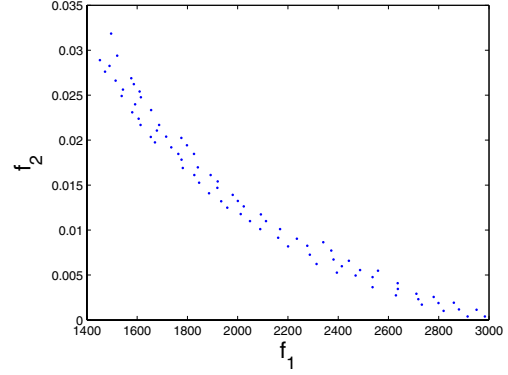


**Figure 4: Numerical result for MOP (31). Here, we have chosen $\epsilon = (50, 0.0005)$ and $\Delta = (10, 0.0001)$.**

computational time was 5.5 seconds. In contrast, a run of the same algorithm with the same setting but with $\Delta = 0$ took 4 minutes and 21 seconds leading to an archive with 8377 elements.

## 5.3 Example 3

Finally, we consider a bi-objective $\{0,1\}$-knapsack problem which should demonstrate that the additional consideration of approximate solutions can be beneficial for the decision maker.

$$f_1, f_2 : \{0,1\}^n \to \mathbb{R}, \;\; f_1(x) = \sum_{j=1}^{n} c_j^1 x_j, \;\; f_2(x) = \sum_{j=1}^{n} c_j^2 x_j \tag{33}$$

s.t.

$$\sum_{j=1}^{n} w_j x_j \leq W, \; x_j \in \{0,1\}, \; j = 1, \dots, n,$$

where $c_j^i$ represents the value of item $j$ on criterion $i, i = 1, 2$; $x_j = 1, j = 1, \dots, n$, if item $j$ is included in the knapsack, else $x_j = 0$. $w_j$ is the weight of item $j$, and $W$ the overall knapsack capacity. Both objectives have to be maximized. Figure 5 shows one numerical result obtained by an evolutionary strategy[2] for an instance with $n = 30$ items and with randomly chosen values $c_j^i \in [8, 12]$, weights $w_j = 1$, and capacity $W = 15$. For $\epsilon = (2, 2)$ and $\Delta = 0.1$ a total of 182 elements forms the final archive, and only six of them are nondominated. When taking, for instance, $x_0$ as reference (assuming, e.g., that this point has been selected by the DM out of the nondominated points) and assuming a tolerance of 1 which represents a possible loss of 0.6% compared to $x_0$ for each objective value, the resulting region of interest includes seven approximate solutions (see Figure 5). These solutions, though similar in objective space, differ significantly in parameter space: two solutions differ compared to $x_0$ in 8 items, one in 10, and 4 solutions differ even in 12 items. Thus, in this case it is obvious that by tolerating approximate solutions—where the loss of them can be determined a priori—a larger variety of possibilities is offered to the DM.

---

[2]A modification of the algorithm presented in [14] which uses the novel archiver.
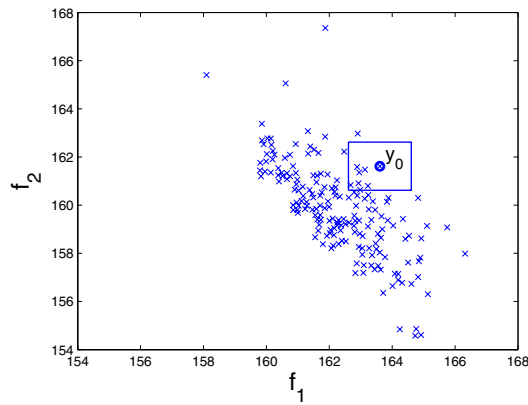
**Figure 5: Numerical result for MOP (31) with $\epsilon = (2, 2)$ and $\Delta = 0.1$. The rectangle indicates a possible region of interest around $y_0 = F(x_0)$ (see text).**

## 6. CONCLUSION AND FUTURE WORK

We have proposed and investigated a novel archiving strategy for stochastic search algorithms which allows—under mild assumptions on the generation process—for a finite size approximation of the set $P_{Q,\epsilon}$ which contains all $\epsilon$-efficient solutions of an MOP within a compact domain $Q$. We have proven convergence of the algorithm toward a finite size representation of the set of interest in the probabilistic sense, yielding bounds on the approximation quality and the cardinality of the archives. Finally, we have presented some numerical results indicating the usefulness of the approach. The consideration of approximate solutions certainly leads to a larger variety of possible options to the DM, but, in turn, also to a higher demand on the related decision making process. Thus, the support for this problem could be one focus of future research. Further, it could be interesting to integrate the archiving strategy directly into the stochastic search process (as e.g. done in [2] for an EMO algorithm) in order to obtain a fast and reliable search procedure.

## 7. REFERENCES

[1] R. Blanquero and E. Carrizosa. A. d.c. biobjective location model. *Journal of Global Optimization*, 23(2):569–580, 2002.

[2] K. Deb, M. Mohan, and S. Mishra. Evaluating the $\epsilon$-domination based multi-objective evolutionary algorithm for a quick computation of Pareto-optimal solutions. *Evolutionary Computation*, 13(4):501–525, 2005.

[3] A. Engau and M. M. Wiecek. Generating epsilon-efficient solutions in multiobjective programming. *European Journal of Operational Research*, 177(3):1566–1579, 2007.

[4] J. D. Knowles and D. W. Corne. *Metaheuristics for Multiobjective Optimisation*, volume 535 of *Lecture Notes in Economics and Mathematical Systems*, chapter Bounded Pareto Archiving: Theory and Practice, pages 39–64. Springer, 2004.

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

[6] P. Loridan. $\epsilon$-solutions in vector minimization problems. *Journal of Optimization, Theory and Application*, 42:265–276, 1984.

[7] Günter Rudolph and Alexandru Agapie. Convergence Properties of Some Multi-Objective Evolutionary Algorithms. In *Proceedings of the 2000 Conference on Evolutionary Computation*, volume 2, pages 1010–1016, Piscataway, New Jersey, July 2000. IEEE Press.

[8] G. Ruhe and B. Fruhwirt. $\epsilon$-optimality for bicriteria programs and its application to minimum cost flows. *Computing*, 44:21–34, 1990.

[9] S. Schaeffler, R. Schultz, and K. Weinzierl. Stochastic method for the solution of unconstrained vector optimization problems. *Journal of Optimization Theory and Applications*, 114(1):209–222, 2002.

[10] O. Schütze, C. A. Coello Coello, and E.-G. Talbi. Approximating the $\epsilon$-efficient set of an MOP with stochastic search algorithms. In A. Gelbukh and A. F. Kuri Morales, editors, *Mexican International Conference on Artificial Intelligence (MICAI-2007)*, pages 128–138. Springer-Verlag Berlin Heidelberg, 2007.

[11] O. Schütze, M. Laumanns, C. A. Coello Coello, M. Dellnitz, and E.-G. Talbi. Convergence of stochastic search algorithms to finite size Pareto set approximations. To appear in Journal of Global Optimization, 2007.

[12] W. Stadler and J. Dauer. Multicriteria optimization in engineering: A tutorial and survey. In M. P. Kamat, editor, *Structural Optimization: Status and Future*, pages 209–249. American Institute of Aeronautics and Astronautics, 1992.

[13] M. Tanaka. GA-based decision support system for multi-criteria optimization. In *Proceedings of International Conference on Systems, Man and Cybernetics*, pages 1556–1561, 1995.

[14] E. Tantar, O. Schütze, J. R. Figueira, C. A. Coello Coello, and E.-G. Talbi. Computing and selecting $\epsilon$-efficient solutions of $\{0, 1\}$-knapsack problems. To appear in Proceedings of the 19th International Conference on Multiple Criteria Decision Making (MCDM-2008), 2008.

[15] D. J. White. Epsilon efficiency. *Journal of Optimization Theory and Applications*, 49(2):319–337, 1986.

[16] D. J. White. Epsilon-dominating solutions in mean-variance portfolio analysis. *European Journal of Operational Research*, 105(3):457–466, 1998.