Approximating the Knee of an MOP with Stochastic Search Algorithms

Oliver Schütze[1], Marco Laumanns[2], and Carlos A. Coello Coello[1]

[1] CINVESTAV-IPN, Computer Science Department, Mexico City, Mexico
{schuetze,ccoello}@cs.cinvestav.mx,
[2] ETH Zurich, Institute for Operations Research, Zurich, Switzerland
laumanns@ifor.math.ethz.ch

Abstract. In this paper we address the problem of approximating the 'knee' of a bi-objective optimization problem with stochastic search algorithms. Knees or entire knee-regions are of particular interest since such solutions are often preferred by the decision makers in many applications. Here we propose and investigate two update strategies which can be used in combination with stochastic multi-objective search algorithm (e.g., evolutionary algorithms) and aim for the computation of the knee and the knee-region, respectively. Finally, we demonstrate the applicability of the approach on two examples.

1 Introduction

In many real world problems several objective functions have to be optimized simultaneously. One typical goal for such *multi-objective optimization problems* (MOPs) is to identify the entire set of optimal solutions (the Pareto set) and its image in objective space, the Pareto front. However, since the Pareto set typically forms a (k-1)-dimensional object, where k denotes the number of objectives, this task may become too hard, in particular for more objectives. Instead, one can e.g. integrate the decision maker (DM) into the search process (e.g., with *interactive methods* [11]) or can compute selected points out of the Pareto set, which we address here. One such particular solution is the 'knee'[3] or the 'maximal bulge' of the Pareto front which is often preferred by many DMs since it represents for them the 'optimal compromise' in multi-objective optimization. In this paper we propose and investigate two archiving strategies for stochastic search algorithms which aim for the computation of such a knee and entire knee-regions (i.e., solutions where the bulge is maximal or nearly maximal) respectively. We consider here the bi-objective case (i.e., $k = 2$), but the results may be extended for larger number of objectives.
Knees or other related user preference areas in multi-objective optimization have been addressed in many works so far ([3, 11, 2, 4, 9, 1, 6, 13, 12, 7, 10, 5]). For instance, in [1] a multi-objective evolutionary algorithm is presented which focuses

[3] There exist different characterizations of the knee in literature which, however, lead to the same or to similar solutions in many cases.

on the knee-regions of an MOP (using a different characterization of the knee). The approach which we propose here can be viewed as a possible alternative to this work. One advantage of our strategies is that they can easily be integrated into any given archiving strategy for a stochastic search procedure. In that case, the (additional) approximation of the knee comes for 'free' in the sense that no additional function call has to be spent.

2 Background

In the following we consider continuous multi-objective optimization problems

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

where $Q \subset \mathbb{R}^n$ is compact and F is defined as the vector of the objective functions $F : Q \to \mathbb{R}^k, \quad F(x) = (f_1(x), \ldots, f_k(x))$, and where each $f_i : Q \to \mathbb{R}$ is continuous.

Definition 1. *Let $v, w \in Q$. Then the vector v is* less than w *($v <_p w$), if $v_i < w_i$ for all $i \in \{1, \ldots, k\}$. The relation $\leq_p$ is defined analogously. $y \in Q$ is* dominated *by a point $x \in Q$ ($x \prec y$) with respect to (MOP) if $F(x) \leq_p F(y)$ and $F(x) \neq F(y)$. $x \in Q$ 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). The image of the Pareto set $F(P_Q)$ is called the *Pareto front*. Further, we need the following distances between different sets.

Definition 2. *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: $dist(u, A) := \inf_{v \in A} \|u - v\|$, $dist(B, A) := \sup_{u \in B} dist(u, A)$, and $d_H(A, B) := \max\{dist(A, B), dist(B, A)\}$*

Finally, we need to define some straight lines in $\mathbb{R}^2$. For $y_1, y_2 \in \mathbb{R}^2$, $y_1 \neq y_2$, we define by $\mathcal{L}(y_1, y_2) := y_1 + \mathbb{R}(y_2 - y_1)$ the line which goes through y_1 and y_2.

3 Characterization of the Knee

In this section we state one possible way to define the knee and modify it such that we can use it for our purpose.
According to [3], a knee of a Pareto curve is found by solving the following nonlinear programming problem (NLP):

$$\max_{p \in P_Q} dist(F(p), \mathcal{L}(F(p_1^*), F(p_2^*))), \tag{1}$$

where $p_i^* \in \arg\min_{x \in P_Q} f_i(x)$, $i = 1, 2$ (see also Figure 1). The knee as characterized by (1) can be interpreted as the maximal bulge of the curve with respect to the

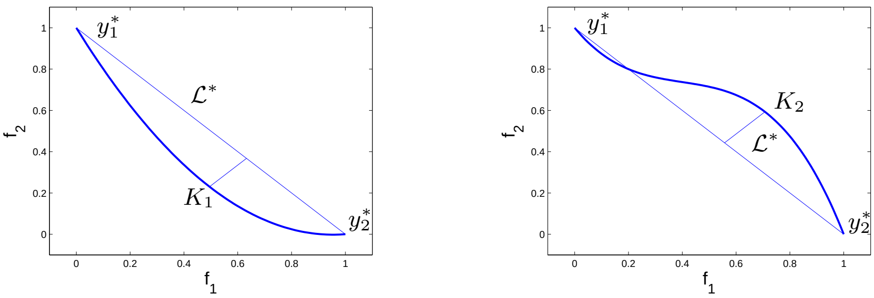
line $\mathcal{L}(F(p_1^*), F(p_2^*))$ which contains the two extreme points of the curve. We have chosen for this characterization since it requires no gradient information and is invariant to scalarization of the objectives.
Since we are interested in 'convex bulges' and not in 'concave bulges' which do intuitively not fit to the idea of minimization (see Figure 1, or [3]), we define the distance of the image $F(p)$ of a candidate solution to $\mathcal{L}(F(p_1^*), F(p_2^*))$ as follows:

$$D(p, p_1^*, p_2^*) := \begin{cases} dist(F(p), \mathcal{L}(F(p_1^*), F(p_2^*))) & \text{if } f_2(p) \leq g(f_1(p)) \\ \text{-}\ dist(F(p), \mathcal{L}(F(p_1^*), F(p_2^*))) & \text{else} \end{cases}, \quad (2)$$

where $g(x) = \mathcal{L}(F(p_1^*), F(p_2^*))$. Using this function and the fact that we are interested in convex bulges, we can modify NLP (1) by

$$\max_{x \in Q} D(x, p_1^*, p_2^*), \quad (3)$$

which will be our 'knee' in the sequel.



(a) 'convex bulge' (b) 'concave bulge'

Fig. 1. Two 'knees' K_1, K_2 for different Pareto fronts as characterized by the maximal bulge of the Pareto curve with respect to $\mathcal{L}^* := \mathcal{L}(y_1^*, y_2^*)$.

4 The Algorithms

Here we propose two different update strategies for the approximation of a single knee as well as entire knee-regions and investigate the limit behavior of these algorithms.
First we are interested in obtaining *one* maximal bulge. Since in most cases (e.g., for all convex problems) 'the' knee is indeed unique is it sufficient to store one approximation — in addition to the approximations of the extreme points of the Pareto curve, since they are also not known a priori. Algorithm 1 shows one

possible way to do this. The input parameters are the approximations m_1^0, m_2^0 of the extreme points, the current approximation K_0 of the knee as well as the new candidate solution $p \in Q$. Outputs are the new approximations of the extreme points (m_1, m_2) and of the knee (K). Theorem 1 shows that the maximal bulge (measured in objective space) is reached in the limit under certain assumptions and in the probabilistic sense.

Algorithm 1 $\{m_1, m_2, K\} := ArchiveUpdateMaxBulge1\ (p, K^0, m_1^0, m_2^0)$

if $f_1(p) < f_1(m_1^0)$ **then**
  $m_1 := p$
else
  $m_1 := m_1^0$
end if
if $f_2(p) < f_1(m_2^0)$ **then**
  $m_2 := p$
else
  $m_2 := m_2^0$
end if
if $D(p, m_1, m_2) > D(K^0, m_1, m_2)$ **then**
  $K := p$
else
  $K := K^0$
end if

Theorem 1. *Let (MOP) be given and $Q \subset \mathbb{R}^n$ be compact, let there be no weak Pareto points in $Q \backslash P_Q$, and $K^0, m_1^{(0)}, m_2^{(0)} \in Q$. Further, let $p_i^*, i = 1, 2$, as defined above with $F(p_1^*) \neq F(p_2^*)$, and*

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

where $B_\delta(x) := \{y \in \mathbb{R}^n \ : \ \|y - x\| < \delta\}$ and $P(A)$ denotes the probability for event A. Then, if Algorithm 1 is used to update the sequences $K_l, m_1^{(l)}, m_2^{(l)}, l \in \mathbb{N}$, it holds with probability one

(a)

$$m_1^{(l)} \to p_1^* \in \arg\min_{x \in P_Q} \ f_1(x) \quad \text{for } l \to \infty$$
$$m_{s_l}^{(l)} \to p_2^* \in \arg\min_{x \in P_Q} \ f_2(x) \quad \text{for } l \to \infty$$

(b)

$$D(K_l) \to \max_{x \in Q} D(x, p_1^*, p_2^*) \quad \text{for } l \to \infty.$$

Proof. (a) We prove the convergence of the sequence $(m_1^{(l)})_{l\in\mathbb{N}}$, the other statement is analogue. The claim follows, roughly speaking, by assumption (4) on the process to generate new candidate solutions and by the fact that the point with the smallest value according to f_1 which is found during the search is kept in the archive. To be more precise, let $x_1^* \in \arg\min_{x\in P_Q} f_1(x)$. By (4) it follows that there exists for every $i \in \mathbb{N}$ with probability one a number j_i and a point $p_{j_i} \in B_{1/i}(x_1^*) \cap Q$. By construction of Alg. 1 it is $f_1(m_1^{(j_i)}) \leq f_1(p_{j_i})$. Thus, the claim follows since $p_{j_i} \to x_1^*$ for $i \to \infty$.

(b) The straight lines $\mathcal{L}_l(F(a_{m_1}^{(l)}), F(a_{m_2}^{(l)}))$ can be written as $g_l(x) = m_l x + b_l$. Let $a := f_1(p_1^*)$ and $b := f_1(p_2^*)$. Denote by $S_p = (x_p, y_p) \in \mathcal{L}_l$ the vector with minimal distance to the candidate solution p_l. It is easy to verify that $x_p \in [a,b]$ (see e.g. the Appendix). Thus, it is sufficient to consider the functions g_l on the interval $[a,b]$. Since $F(p_1^*) \neq F(p_2^*)$ and by part (a) of this theorem it follows that the g_l's are converging uniformly to $g = \mathcal{L}(F(p_1^*), F(p_2^*))$ on $[a,b]$, and thus we have with probability one

$$\max_{x\in Q} D(x, m_1^{(l)}, m_2^{(l)}) \to \max_{x\in Q} D(x, p_1^*, p_2^*), \quad \text{for } l \to \infty. \tag{5}$$

Let $p^* \in \arg\max_{x\in P_Q} D(x, p_1^*, p_2^*)$. By (4) it follows that there exists with probability one a subsequence of p_{j_i} of the candidate solutions such that $p_{j_i} \in B_{1/i}(p^*) \cap Q$. By construction of Alg. 1 it follows that $D(K_{j_i}, a_{m_1}^{(j_i)}, a_{m_2}^{(j_i)}) \geq D(p_{j_i}, a_{m_1}^{(j_i)}, a_{m_2}^{(j_i)})$. Using this and (5) we obtain with probability one

$$D(K_l, a_{m_1}^{(l)}, a_{m_2}^{(l)}) \to \max_{x\in Q} D(x, p_1^*, p_2^*),\ l \to \infty \tag{6}$$

and the proof is complete.

Next, we are interested to approximate beyond one knee solution the subset of the Pareto front where the bulge is 'large' since this entire set could be interesting for the decision maker ([1]). That is, for $M := \max_{x\in Q} D(x, p_1^*, p_2^*)$ and given a threshold $\Delta \in \mathbb{R}_+$ we are interested in the following set:

$$K_\Delta := \{x \in P_Q | D(x, p_1^*, p_2^*) \geq M - \Delta\} \tag{7}$$

Note that in case the knee is not unique all these points are included in K_Δ for every value of Δ, which is another motivation to approximate this set.
In Algorithm 2 we propose one possible archiving strategy which aims for the approximation of K_Δ. The notation is as in Alg. 1 with the difference that K is a set of points. In the following we investigate the limit behavior of the strategy under the same assumptions as above (Thm. 2). Before we can do this we need the following result.

Lemma 1. *Let $m_1, m_2, z, d \in Q$ with $f_1(m_1) < f_1(x) < f_1(m_2)$, and $g(f_1(x)) \leq f_2(x)$, where $g(\cdot) = \mathcal{L}_l(F(m_1), F(m_2))$, and $d \prec z$, and let m_1 and m_2 be mutually nondominating. Then $D(d, m_1, m_2) > D(z, m_1, m_2)$.*

Algorithm 2 $\{m_1, m_2, K\} := ArchiveUpdateMaxBulge2\ (p, K^0, m_1^0, m_2^0, \Delta)$

1: **if** $f_1(p) < f_1(m_1^0)$ **then**
2: $\quad m_1 := p$
3: **else**
4: $\quad m_1 := m_1^0$
5: **end if**
6: **if** $f_2(p) < f_1(m_2^0)$ **then**
7: $\quad m_2 := p$
8: **else**
9: $\quad m_2 := m_2^0$
10: **end if**
11: $\tilde{K} := K^0 \cup \{p\}$
12: $\tilde{M} := \max_{k \in \tilde{K}} D(k, m_1, m_2)$
13: $K := nondom(\{k \in \tilde{K} : D(k, m_1, m_2) \geq \tilde{M} - \Delta\})$

Proof. Assume that $D(d, m_1, m_2) \leq D(z, m_1, m_2)$. Let $g(x_1) = ax_1 + b$. Since m_1 and m_2 are mutually nondominating it follows that a is negative. Define by g_2 the straight line which is parallel to g and which goes through z, i.e., $g_2(x_1) = ax_1 + b_2$. Since by assumption $D(d, m_1, m_2) \leq D(z, m_1, m_2)$ it follows that $f_2(d) \geq g_2(f_1(d))$. Since the slope a of g_2 is negative it follows that either $f_1(d) \geq f_1(z)$ or $f_2(d) \geq f_2(z)$ which is a contradiction to $d \prec z$, and thus, it must the that $D(d, m_1, m_2) > D(z, m_1, m_2)$.

Theorem 2. *Using the definitions above, let $M > 0$, $\Delta \in \mathbb{R}_+$ with $M - \Delta > 0$, and let*

$$\lim_{i \to \infty} K_{\Delta_i} \to K_\Delta \tag{8}$$

for every sequence $(\Delta_i)_{i \in \mathbb{N}}$ with $\Delta_i < \Delta$ and $\Delta_i \to \Delta$ for $i \to \infty$. Then, if Algorithm 2 is used to update the sequences $K_l, m_1^{(l)}, m_2^{(l)}, l \in \mathbb{N}$, and under the assumptions made in Thm. 1 it holds with probability one

(a)

$$m_1^{(l)} \to p_1^* \in \arg\min_{x \in P_Q} f_1(x) \quad \textit{for } l \to \infty$$

(b)

$$d_H(F(K_\Delta), F(K_l)) \to 0 \quad \textit{for } l \to \infty$$

Proof. (a) Analogue to proof of Thm 1 (a).

(b) First we show that $dist(F(K_\Delta), F(K_l)) \to 0$ for $l \to \infty$ with probability one. Since K_l, $l \in \mathbb{N}$, is finite and K_Δ is compact it follows that

$$dist(F(K_\Delta), F(K_l)) = \max_{p \in K_\Delta} \min_{k \in K_l} \|F(p) - F(k)\|. \tag{9}$$

By (8) it is sufficient to consider points $p \in P_Q$ with $D(p, p_1^*, p_2^*) > M - \Delta$. Let p be such a point. By Thm. 1 it follows that $\tilde{M}_l$ (see line 12 of Alg. 2) converges to M with probability one. Further, since D and F are continuous it follows that there exist with probability one a neigborhood U of p and an integer l_0 such that

$$D(u, m_1^{(l)}, m_2^{(l)}) > \tilde{M}_l - \Delta, \quad \forall u \in U,\ \forall l \geq l_0. \tag{10}$$

By (4) it follows that there exists with probability one for every $j \in \mathbb{N}$ a point $p_{l_j} \in U \cap B_{1/j}(p) \cap Q$. By construction of Alg. 2 the point p_{l_j} will either be added to the archive (in that case denote $d_j := p_{l_j}$), or there already exists a point $d_j \in K_{l_j}$ which dominates p_{l_j}. Due to (10) the point d_j will only be discarded from the archive if in turn a dominated solution is found. By this and since $p_j \to p$ and thus $F(d_j) \to F(p)$ for $j \to \infty$ it follows that

$$dist(F(p), F(K_l)) = \min_{k \in K_l} \|F(p) - F(k)\| \to 0 \quad \text{with probability one,} \tag{11}$$

and the claim follows. It remains to show that also

$$dist(F(K_l), F(K_\Delta)) = \max_{k \in K_l} \min_{p \in K_\Delta} \|F(k) - F(p)\| \tag{12}$$

vanishes for $l \to \infty$ and in the probabilistic sense. For this we have to show that every point $x \in Q \backslash K_\Delta$ will be discarded (if added before) from the archive after finitely many steps, and that this point will never be added further on, both with probability one. Let $x \in Q \backslash K_\Delta$, that is, we have either (a) $D(x, p_1^*, p_2^*) < M - \Delta$ or (b) $x \notin P_Q$. First we consider case (a). Since the sequence $\tilde{M}_l \to M$ (see above) and by continuity of D there exists with probability one an integer l_0 with

$$D(x, m_1^{(l)}, m_2^{(l)}) < \tilde{M}_l - \Delta, \quad \forall l \geq l_0, \tag{13}$$

and by this, that x is not a member of K_l for $l \geq l_0$.
Next, let $x \notin P_Q$. By case (a) we can assume that $D(x, p_1^*, p_2^*) \geq M - \Delta > 0$. Since x is not a weak Pareto point there exists a point $p \in P_Q$ with $F(p) <_p F(x)$. By continuity of D and F, by part (a) of this theorem, and by Lemma 1 it follows that there exists a neighborhood U of p and an integer l_0 such that:

$$\begin{aligned} F(u) &<_p F(x), \quad \forall u \in U, \text{ and} \\ D(u, m_1^{(l)}, m_2^{(l)}) &> D(x, m_1^{(l)}, m_2^{(l)}), \quad \forall u \in U,\ \forall l \geq l_0 \\ D(u, m_1^{(l)}, m_2^{(l)}) &> \tilde{M} - \Delta, \quad \forall u \in U,\ \forall l \geq l_0. \end{aligned} \tag{14}$$

By (4) it follows that there exists with probability one an integer $j_0 > l_0$ such that that the candidate solution p_{j_0} is in $U \cap Q$. Further, by (14) and by construction of Alg. 2 it follows that p_{j_0} will be either added to the archive or that there already exists a point d which dominates p_{j_0}. In further iterations of the algorithm, this point is only discarded if a dominated solution is found (using (14) and Lemma 1). Since $\prec$ is transitive all these points dominate x, and hence is not a member of K_l for all integers $l \geq j_0$ with probability one, and the proof is complete.

Since the archiver in Alg. 2 accepts all points in K_Δ and does not discard them further on it follows that in the course of the computation $|K_l| \to \infty$ for $l \to \infty$. In order to prevent this, one could select a subset of K_l in each step, e.g., by the techniques proposed in [8] or other pruning techniques.

5 Numerical Results

Here we present some numerical results on two MOPs: a convex problem and an MOP ([14]) which has two optimal points with maximal bulge:

$$\begin{aligned} F_1 &: [-2,2]^2 \to \mathbb{R}^2 \\ F_1(x) &= \left((x_1-1)^2 + (x_2-1)^2, (x_1+1)^2 + (x_2+1)^2\right) \end{aligned} \tag{15}$$

and

$$\begin{aligned} F_2 &= (f_1, f_2) : [-1.5, 1.5]^2 \to \mathbb{R}^2 \\ f_1(x,y) &= \frac{1}{2}(\sqrt{1+(x+y)^2} + \sqrt{1+(x-y)^2} + x - y) + \lambda \cdot e^{-(x-y)^2} \\ f_2(x,y) &= \frac{1}{2}(\sqrt{1+(x+y)^2} + \sqrt{1+(x-y)^2} - x + y) + \lambda \cdot e^{-(x-y)^2} \end{aligned} \tag{16}$$

For the generation of the sequence $(p_l)_{l\in\mathbb{N}}$ of candidate solutions we have taken a random search operator. Figures 2 and 3 show two numerical results—i.e., one result for every archiving strategy—for each of the models. MOP (16) contains two maximal bulges, and hence, the archiver $ArchiveUpdateMaxBulge1$ can only reach one of them (Fig. 3 (a)). However, this does not occur when using the second archiver (Fig. 3 (b)).

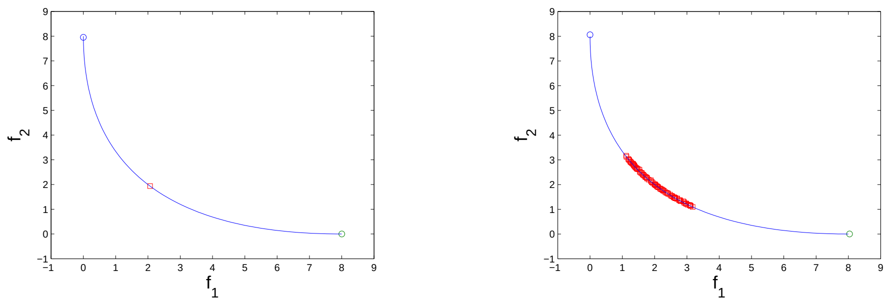


Fig. 2. Numerical results for MOP (15) with $N = 10,000$ randomly chosen points within $Q = [-2,2]^2$ for Alg. 1 (left) and for Alg. 2 for $\Delta = 0.2$. The circles represent the final extreme points, and the square(s) the approximation of the knee (region).

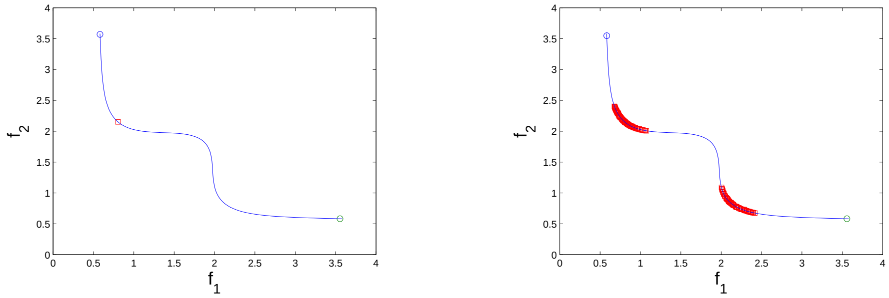


Fig. 3. Two numerical results for MOP (16) with $N = 10,000$ randomly chosen points within $Q = [-1.5, 1.5]^2$ for Alg. 1 (left) and for Alg. 2 (right) for $\Delta = 0.1$. The circles represent the final extreme points, and the square(s) the approximation of the knee (region).

6 Conclusions and Future Work

In this paper we have proposed and investigated two update strategies for the approximation of knees respectively knee-regions of multi-objective optimization problems with stochastic search algorithms. The advantage of these methods is that they can be used either as standalone-algorithms together with any stochastic search procedure or integrated into any other archiving strategy (e.g., distance based ones) without causing additional function calls. We have demonstrated on two examples where we have used a random search operator that the novel strategies are capable of approximating the desired regions with reasonable effort.

For future research, there are mainly two points which have to be addressed. First, a generalization of the obtained results for $k > 2$ would be desirable. Further, the integration of the archivers into stochastic search procedures is of particular interest: since the archivers focus on a real subset of the Pareto front, a natural demand on the resulting algorithm is that it should be more efficient in terms of function calls than algorithms which aim for the approximation of the entire Pareto front. This is, however, ad hoc not straightforward since it is well-known that the approximation of the nadir points can be a challenging task itself.

Acknowledgements The third author gratefully acknowledges support from the CONACyT project no. 45683-Y.

References

1. J. Branke, K. Deb, H. Dierolf, and M. Osswald. Finding knees in multi-objective optimization. In *Parallel Problem Solving from Nature*, volume 3242 of *LNCS*, pages 722–731. Springer, DEC 2004.

2. J. Branke, T. Kaussler, and H. Schmeck. Guidance in evolutionary multi-objective optimization. *Advances in Engineering Software*, 32:499–507, 2001.
3. I. Das. On characterizing the "knee" of the Pareto curve based on Normal Boundary Intersection. *Structural Optimization*, 18:107–115, 1999.
4. K. Deb. Multi-objective evolutionary algorithms: introducing bias among pareto-optimal solutions. pages 263–292, 2003.
5. F. di Pierro, S. F. Khu, and D. A. Savic. An investigation on preference order ranking scheme for multiobjective evolutionary optimization. *Evolutionary Computation, IEEE Transactions on*, 11(1):17–45, 2007.
6. J. Handl and J. Knowles. Exploiting the trade-off: the benefits of multiple objectives in data clustering. *Proceedings of the Third International Conference on Evolutionary Multicriterion Optimization*, pages 547–560.
7. H. Ishibuchi, Y. Nojima, K. Narukawa, and T. Doi. Incorporation of decision maker's preference into evolutionary multiobjective optimization algorithms. In *GECCO*, pages 741–742, 2006.
8. S. Kukkonen and K. Deb. A fast and effective method for pruning of non-dominated solutions in many-objective problems. In *PPSN*, pages 553–562, 2006.
9. C. A. Mattson, A. A. Mullur, and A. Messac. Smart Pareto filter: Obtaining a minimal representation of multiobjective design space. *Engineering Optimization*, 36:721–740, 2004.
10. J. Mehnen and H. Trautmann. Integration of expert's preferences in pareto optimization by desirability function techniques. In *Proceedings of the 5th CIRP International Seminar on Intelligent Computation in Manufacturing Engineering (CIRP ICME '06)*, 2006.
11. K. Miettinen. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, 1999.
12. L. Rachmawati and D. Srinivasan. A multi-objective evolutionary algorithm with weighted-sum niching for convergence on knee regions. *Proceedings of the 8th annual conference on Genetic and evolutionary computation*, pages 749–750, 2006.
13. L. Rachmawati and D. Srinivasan. A Multi-Objective Genetic Algorithm with Controllable Convergence on Knee Regions. *Evolutionary Computation, 2006. CEC 2006. IEEE Congress on*, pages 1916–1923, 2006.
14. K. Witting and M. Hessel von Molo. Private communication, 2006.

7 Appendix

Given points $p, a_1, a_N \in Q$, the distance of $F(p)$ to the straight line $\mathcal{L}(F(a_1), F(a_N))$ can be computed as follows: since $\mathcal{L}(F(a_1), F(a_N)) \subset \mathbb{R}^2$, it can be written as a function $g_1 : \mathbb{R} \to \mathbb{R}$, $g_1(x) = m_1 x + b_1$ with $m_1 = -(f_2(a_1) - f_2(a_N))/(f_1(a_N) - f_1(a_1))$, and $b_1 = f_2(a_1) - m_1 f_1(a_1)$, where the interpolation conditions $g_1(f_1(a_j)) = f_2(a_j), j \in \{1, N\}$, are used. To compute the distance of $F(p)$ and g_1 we define the auxiliary function $g_2(x) = m_2 x + b_2$ with $g_2(f_1(p)) = f_2(p)$ and which is orthogonal to g_1. Doing so, this leads to the coefficients $m_2 = -1/m_1$ and $b_2 = f_2(p) - m_2 f_1(p)$. The intersection of g_1 and g_2 is given by the point $S_p = (x_s, y_s)$ with $x_s = (b_2 - b_1)/(m_1 - m_2)$, $y_s = m_2 x_s + b_2$, and thus, we have

$$dist(F(p), \mathcal{L}(F(a_1), F(a_N))) = \|F(p) - S_p\| \tag{17}$$