Archiving Strategies for Evolutionary Multi-objective Optimization Problems

Dr. Oliver Schütze, Cinvestav

Hagenberg, Sept. 16, 2024

Outline

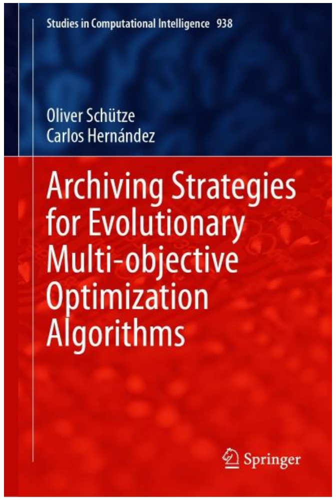
Background

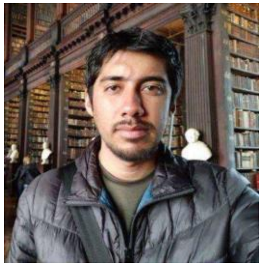
- MOO
- Motivation
- Overview of existing strategies
- Setting for convergence analysis

Archivers for PF approximations

- unbounded archiver
- impicitly bounded archivers
- a priori bounded archiver (HD approximations)

Conclusions and open issues

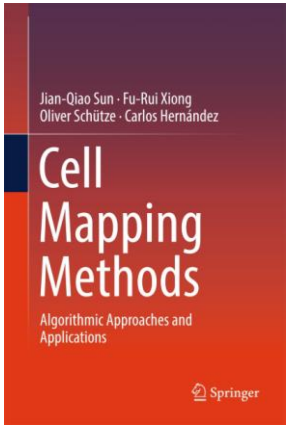


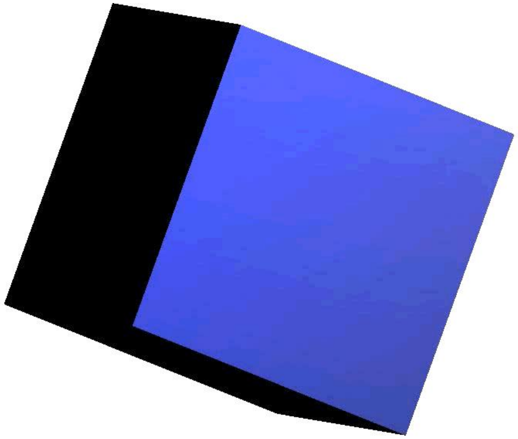


Carlos Hernandez
UNAM, Mexico

About Myself

- MSc in Bayreuth and PhD in Paderborn (Michael Dellnitz)
  Dynamical systems approach to optimization
- First EA conferences: EMO 01, PPSN VII
- Cinvestav, Mexico City, since 2007
- Numerical and Evolutionary Optimization:
  local saerch, hyrbrid methods, performance
  indicators, archiving.

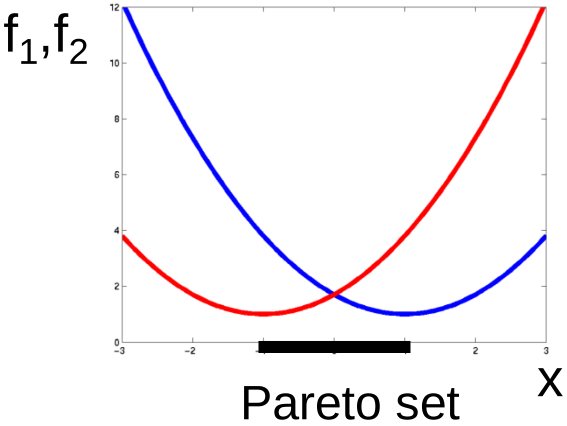




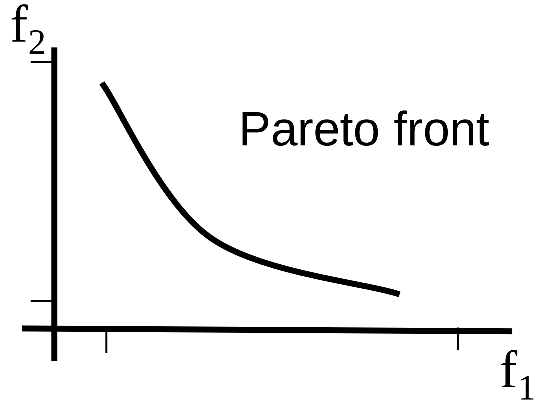
Multi-objective Optimization

Multi-objective Optimization Problem

$$\min F = \begin{cases} f_1 : Q \subset R^n \to R \\ \vdots \\ f_k : Q \subset R^n \to R \end{cases} \qquad \textbf{(MOP)}$$



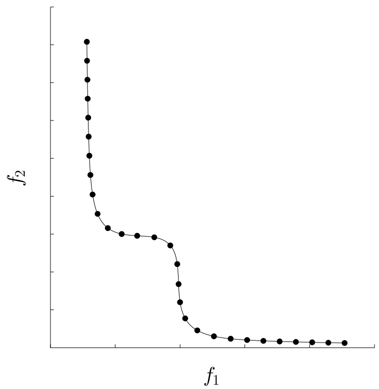
Definition: $x \in Q$ is Pareto optimal $:\Leftrightarrow$
there exists no point $y \in Q \setminus \{x\}$ s.t.
$F(y) \leq F(x)$ and $F(y) \neq F(x)$.



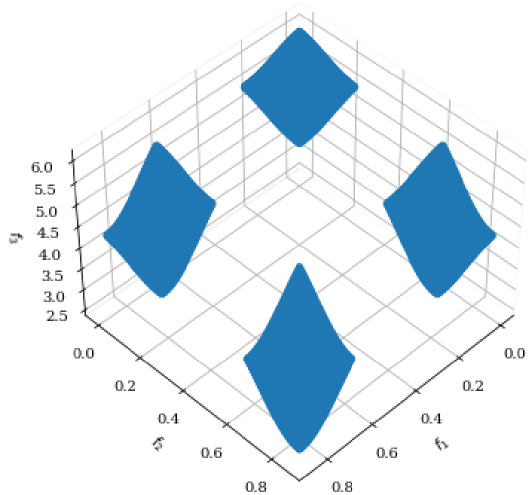
P_Q and $F(P_Q)$ typically form sets of dimension (k-1)

Pareto Front Shapes

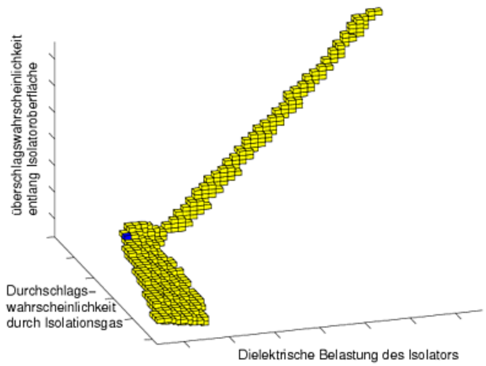
The picture can't be displayed.



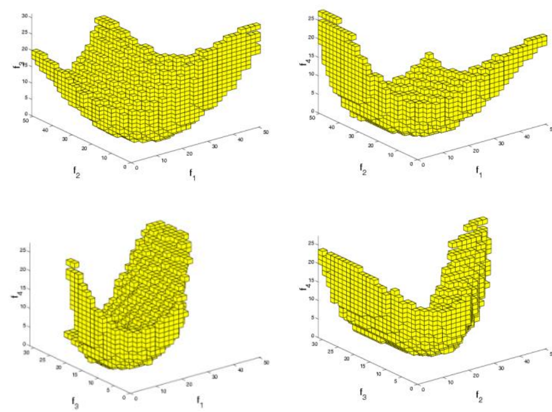
k=2



k=3



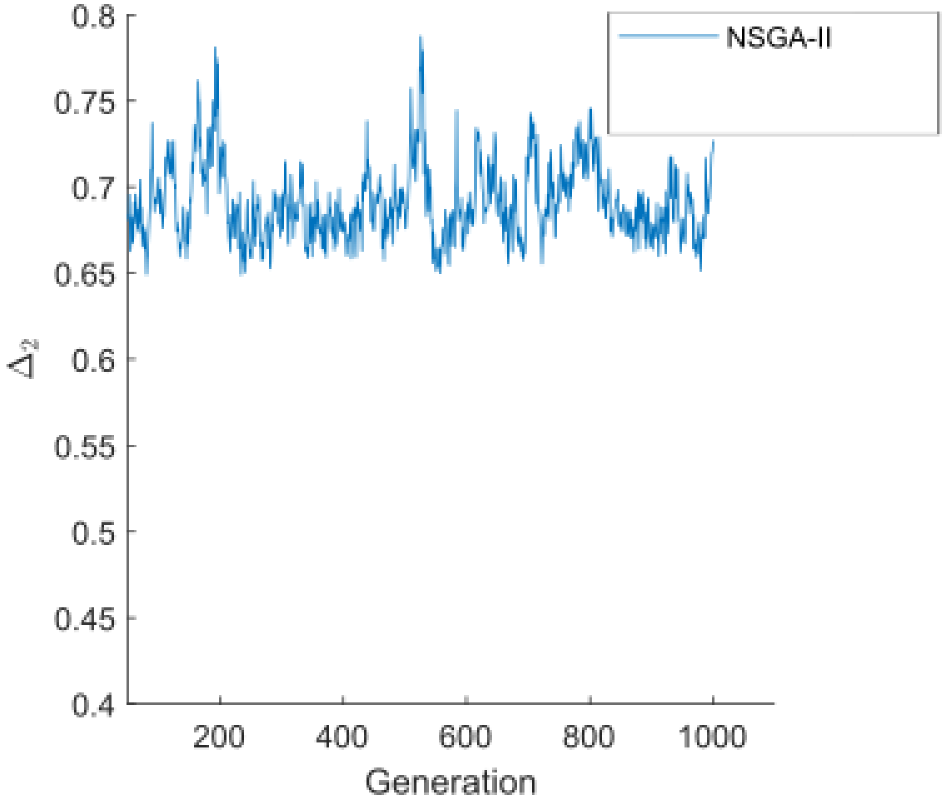
k=3



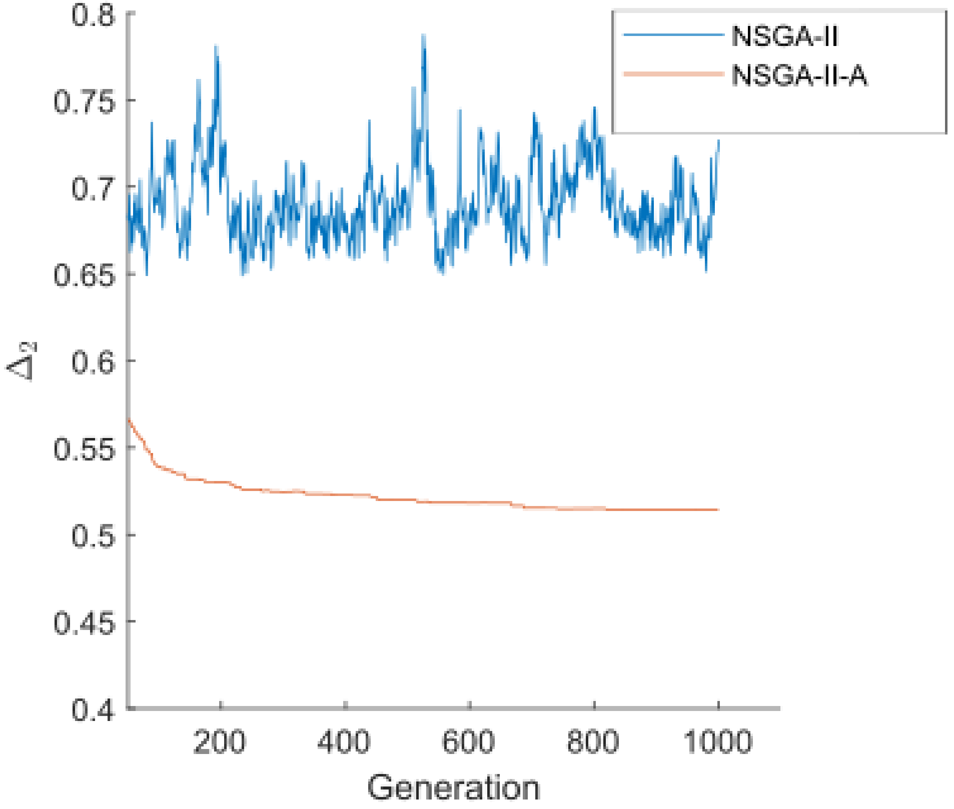
k=4

Motivation I: Behavior of NSGA-II

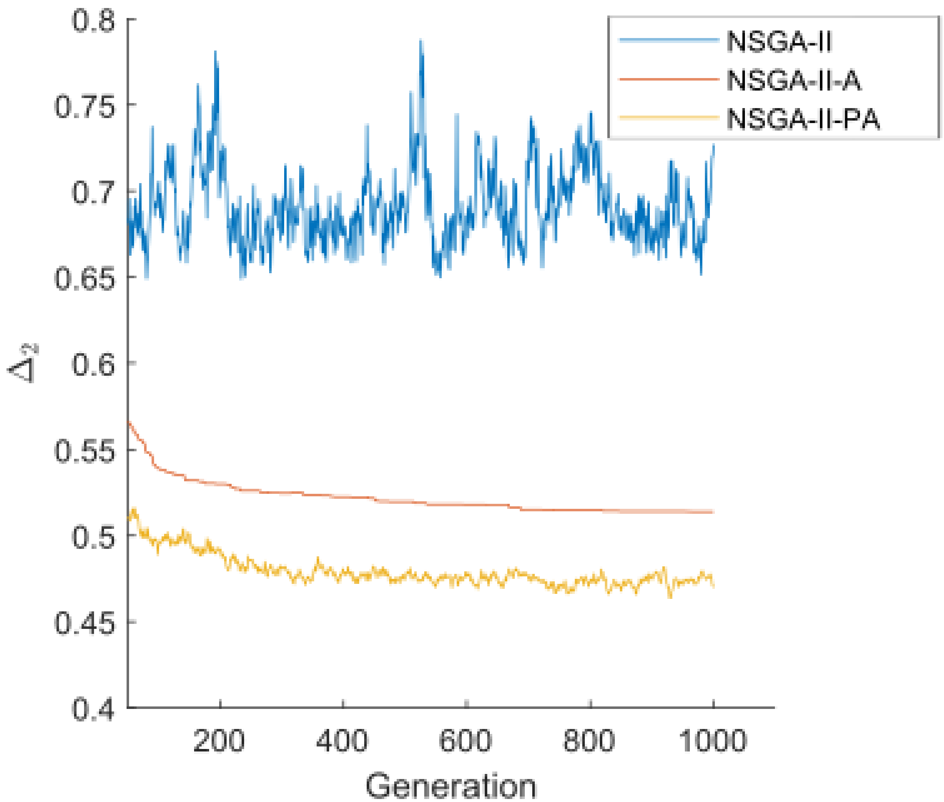
The picture can't be displayed.



Motivation I: Behavior of NSGA-II



Motivation I: Behavior of NSGA-II

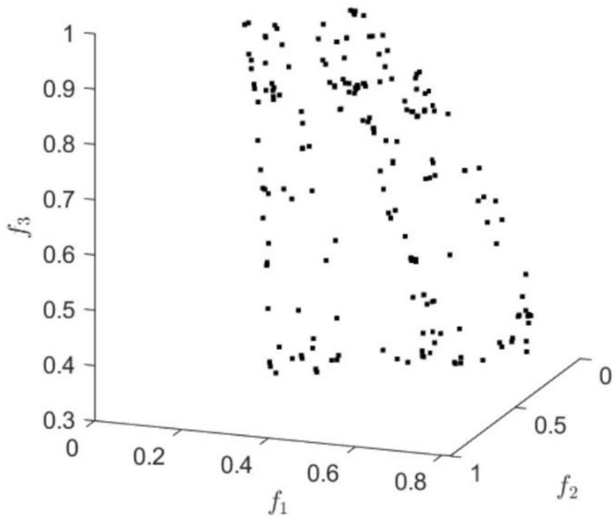


Desired:

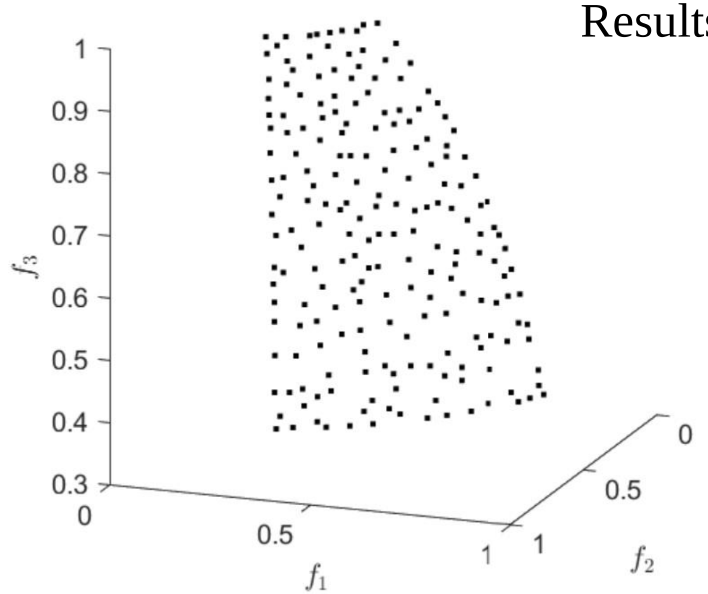
- Impact on performance
- Monotonic behavior (longer runs → better approximations)
- Convergence (limit behavior, approximation qualities)

Motivation II: SMS-EMOA

Results on MaF2



SMS-EMOA
(final population)



SMS-EMOA-A
(final external archive using
ArchiveUpdate-HD)

[Hernández, Sch., 2022]

Archiving/Selection in EMO

1.) **Unbounded archivers**: store all promising solutions
mostly PF, but e.g. also entire PS, all local solutions, all nearly optimal solutions
Theoretical aspects: Schütze et al.
Practical aspects: Ishibuchi et al.

2.) **Implicitly bounded archivers:** archive size depends on problem and design parameters (e.g., based on ε-dominance)
First considerations: Laummans et al.
Further considerations: Schütze et al.

3.) **A priori bounded archivers:** $|A| = \mu$
First theoretical investigations: Rudolph, Hanne (convergence of population toward PS)

Selection within each MOEA:
dominance based: nondominating sorting and niching (Goldberg)
decomposition based: SOP (total order)
indicator based: indicator value/contribution (SOP)

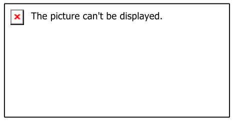
Generic Stochastic Search Algorithm

The picture can't be displayed.

Algorithm Generic Stochastic Search Algorithm

$P_0 \subset Q$ drawn at random
$A_0 := \text{ArchiveUpdate}\ (P_0, \emptyset)$
for $j = 0, 1, 2, \ldots$ **do**
  $P_{j+1} := \text{Generate}\ (P_j)$
  $A_{j+1} := \text{ArchiveUpdate}\ (P_{j+1}, A_j)$
end for

[Laumanns et al., 2002]

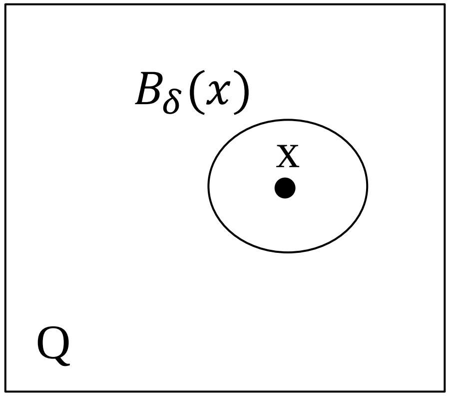


Assumption on Generator

Assumption (A1): Each neighborhood of each point *x* in the domain *Q* will be visited after finitely many steps

$$\forall x \in Q, \forall \delta > 0 : \; P(\exists l \in \mathbb{N} : Pl \cap B_\delta(x) \cap Q) = 1$$

E.g., given for Polynomial Mutation: support of probability density function is identical to Q (at least for box-constrained problems)



Discrete Problems:

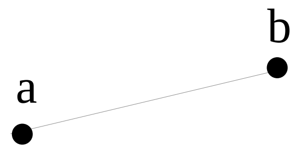
(A1): *Generate()* is a homogeneous finite Markov Chain with irreducible transition matrix

Semi-distance *dist*

Let $a, b \in R^n$ and $A, B \subset R^n$ be compact

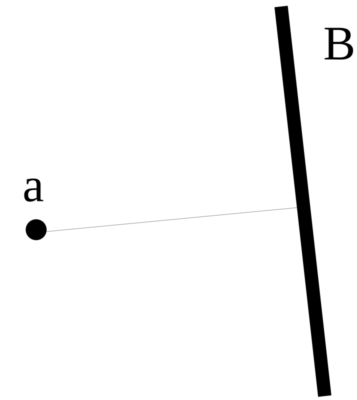
1.) Distance between two points

$$d(a,b) = \|a-b\|_2$$



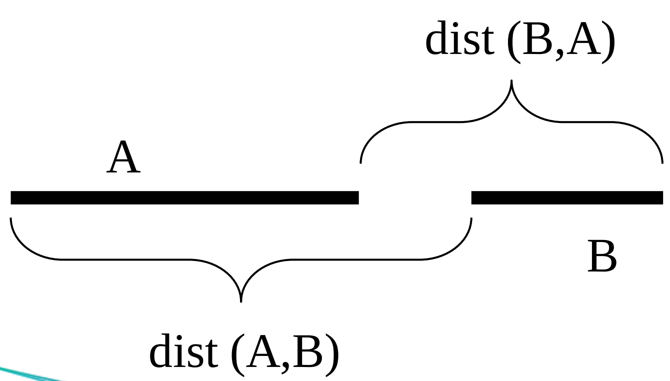
2.) Distance between point and set

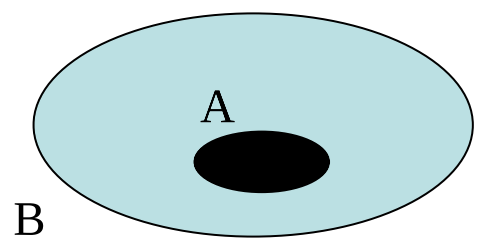
$$dist(a,B) = min_{b \in B}\|a-b\|_2$$



3.) Distance '*dist*' between two sets

$$dist(A,B) = max_{a \in A} dist(a,B)$$





$A \subset B, A \neq B$
$dist(A,B) = 0$
$dist(B,A) > 0$

Hausdorff Distance d_H

Felix Hausdorff, 1868-1942

4.) d_H: ‘symmetrize’ dist

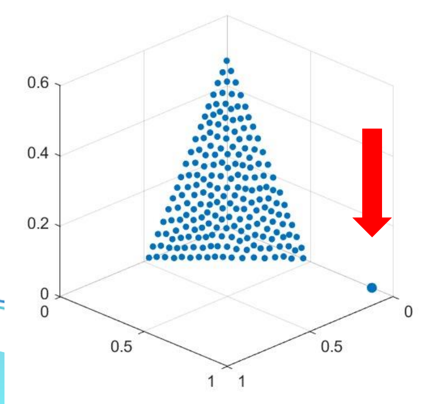
$$d_H(A,B) = \max(dist(A,B), dist(B,A))$$
$$= \max(max_{a\in A} min_{b\in B} \|a-b\|_2, max_{b\in B} min_{a\in A} \|b-a\|_2)$$

Properties

P1) $d_H(\mathbb{Q},\mathbb{R}) = 0$

$\mathbb{Q} \neq \mathbb{R}$, but $\mathbb{Q}$ ‘perfect approximation’ of $\mathcal{R}$ ($clos(\mathbb{Q}) = \mathbb{R}$)

P2) d_H punishes (single) outliers



Use averaged Hausdorff distance

$$\Delta_p(A,B) = \max(GD_p(A,B), IGD_p(A,B))$$

Task 1: Store all Nondominated Solutions

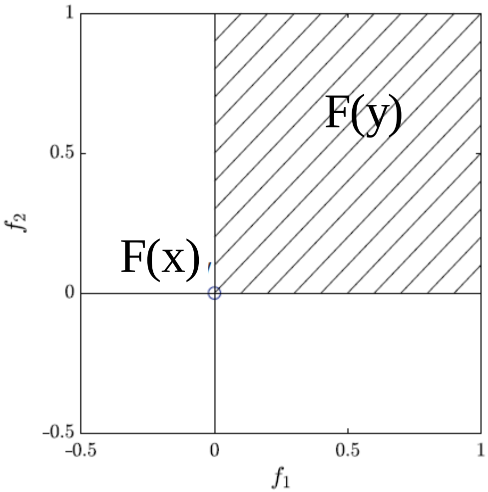
The picture can't be displayed.

ArchiveUpdateP_Q: maintain all non-dominated candidate solutions found during the run of the algorithm

→ Use concept of **dominance**

Let $x, y \in Q$. x is said to dominate y ($x \prec y$) iff

- $f_i(x) \leq f_i(y) \quad \forall\, i = 1, \ldots, k$
- $f_j(x) < f_j(y)$ $for\ a\ j \in \{1, .., k\}$



classical dominance

ArchiveUpdateP_Q

The picture can't be displayed.

Input:
current archive A_0, candidate set P,
(design parameters)

Algorithm A := ArchiveUpdateP_Q (P, A_0)

Require: population P, archive A_0
Ensure: updated archive A

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

Acceptance strategy

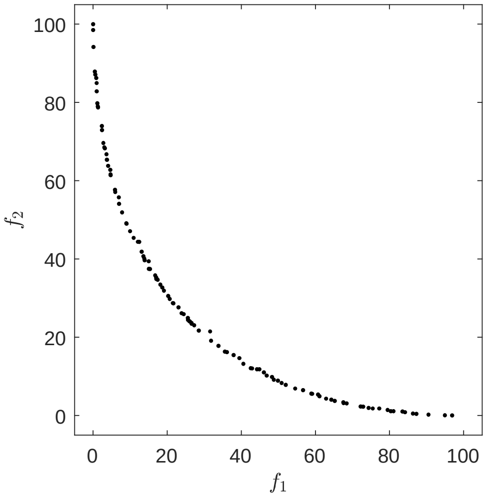
Cleaning strategy

Input:
updated archive A

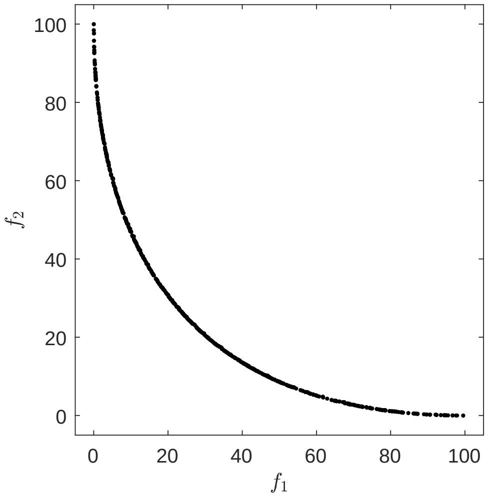
Numerical Results

The picture can't be displayed.

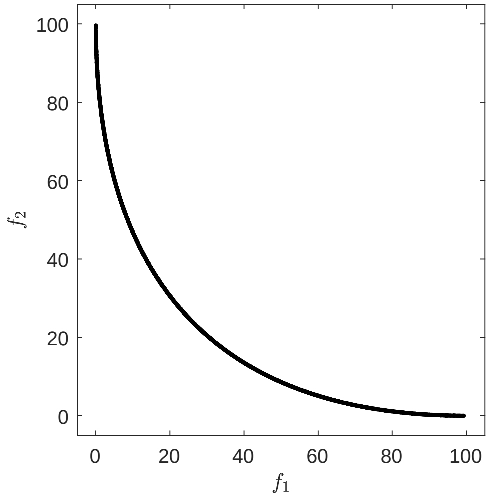
Archiver is fed with candidate solutions that are uniformly chosen at random from the domain.



N=1,000
$|A_f| = 110$



N=10,000
$|A_f| = 481$



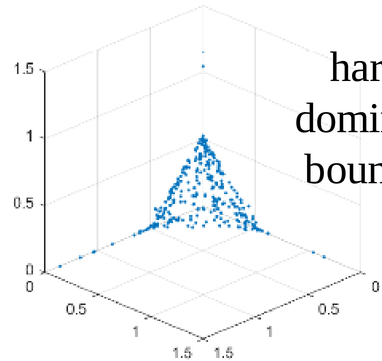
N=100,000
$|A_f| = 2349$

Issues when Using Dominance

1.) In the presence of weakly optimal solutions

[Ishibuchi et al, 2020]: use auxiliary objectives

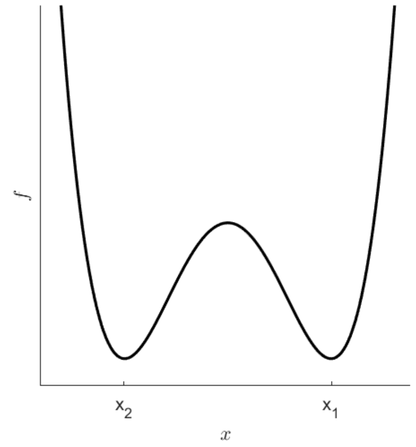
$$\tilde{f}_i(\mathrm{x}) = (1-\alpha) f_i(x) + \frac{\alpha}{k}\sum_{i=1}^{k} f_i\,(x)$$



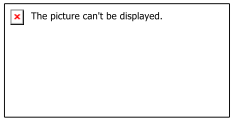
hardly dominated boundary

2.) Dominance is defined in objective space

Expect: two sub-sequences leading to x_1 and x_2 (oscillation), but no convergence toward $\{x_1,x_2\}$

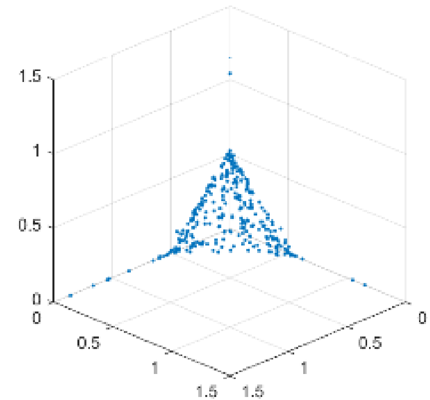


Theoretical Results



'Thm 1' (Monotonicity) Let $P_0, \dots, P_l \subset \mathbb{R}^n$ be finite, and A_i, be the archives obtained by ArchiveUpdatePQ, and $C_l = P_0 \cup \cdots \cup P_l$.

$$A_l = \{x \in C_l : \nexists y \in C_l : y \prec x\}$$



Thm 2 (Convergence I) Under the above assumptions, an applica of ArchiveUpdateP_Q leads to a sequence of archives A_i with

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

Thm 3 (Convergence II) As in Thm 2, and let there be no weak Pareto points in $Q \setminus P_Q$, then

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

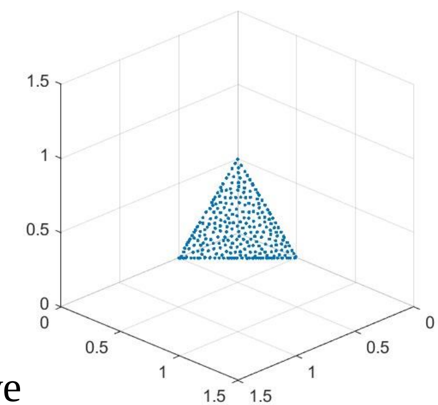
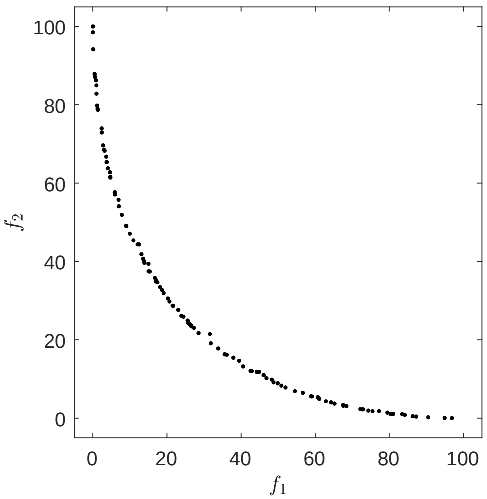


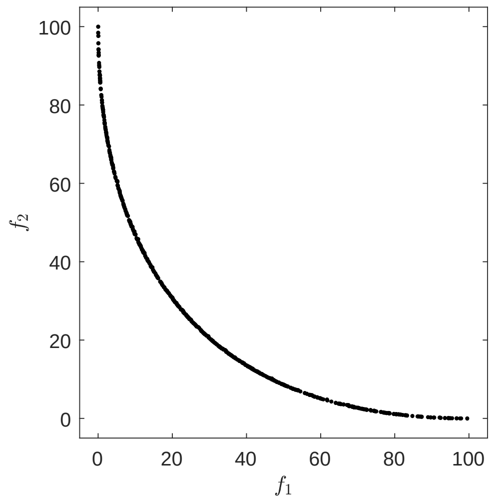
Fig. below: application of Δ_p–Newton method on result above **[Wang et al., 2024]**

Numerical Results

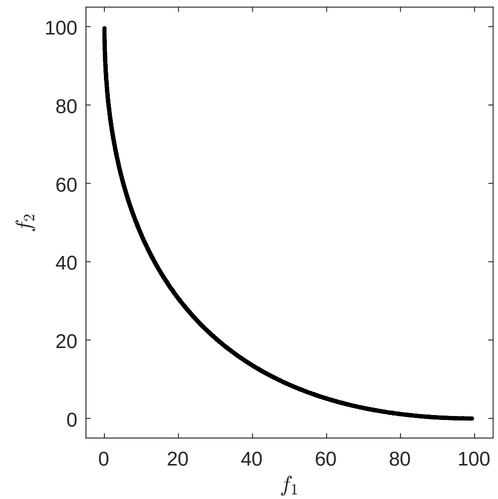
Archiver is fed with candidate solutions that are uniformly chosen at random from the domain.



N=1,000
$|A_f| = 110$



N=10,000
$|A_f| = 481$



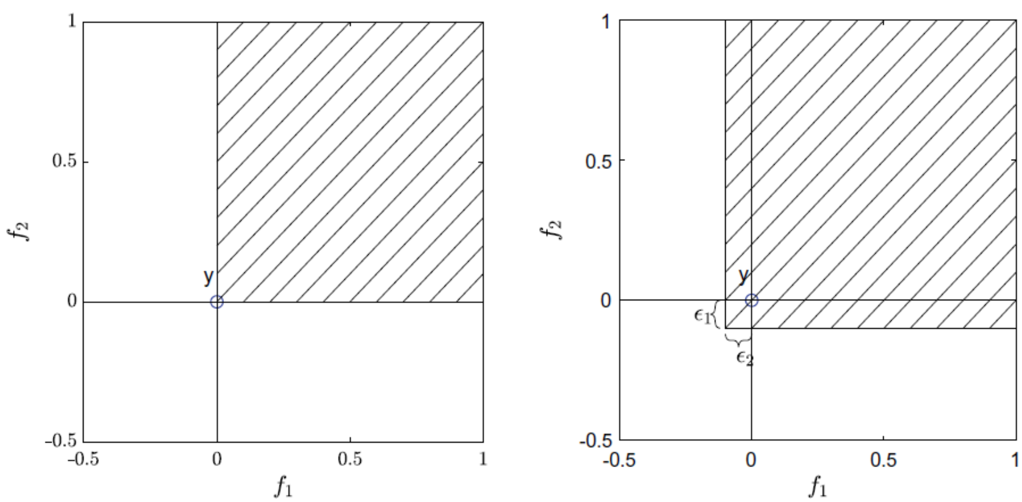
N=100,000
$|A_f| = 2349$

Task 2: Obtain Finite Size PF Approximation

ε −dominance

Let $x, y \in Q$ *and* $\varepsilon \epsilon R_+^k$. x is said to ε − dominate y $(x \prec_\varepsilon y)$ iff

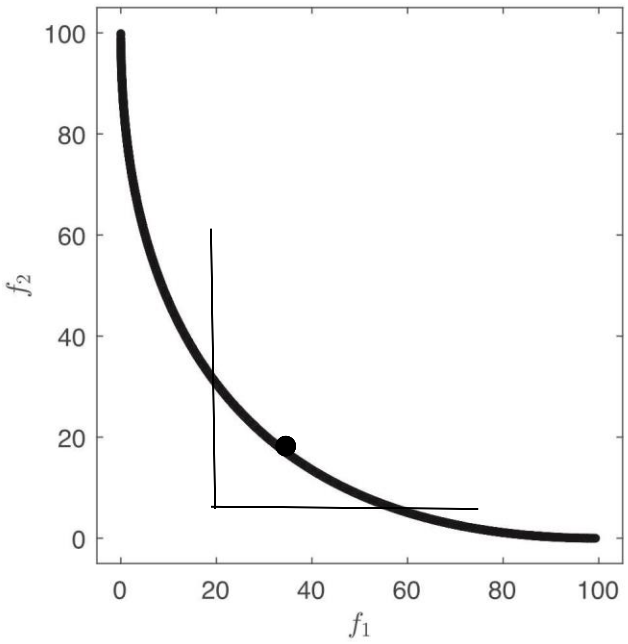
- $f_i(x) - \varepsilon_i \leq f_i(y) \quad \forall\, i = 1, \ldots, k$
- $f_j(x) - \varepsilon_j < f_j(y)$ *for a* $j \in \{1, .., k\}$



(a) classical dominance (b) ϵ-dominance

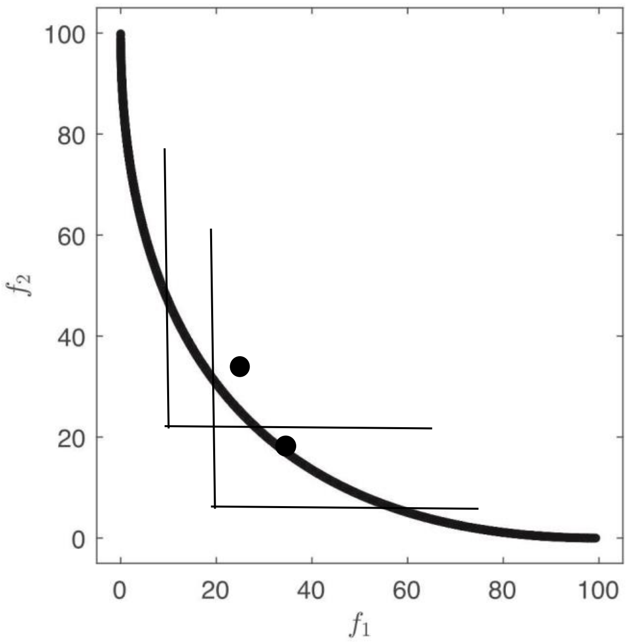
ε-approximations of the PF

The picture can't be displayed.

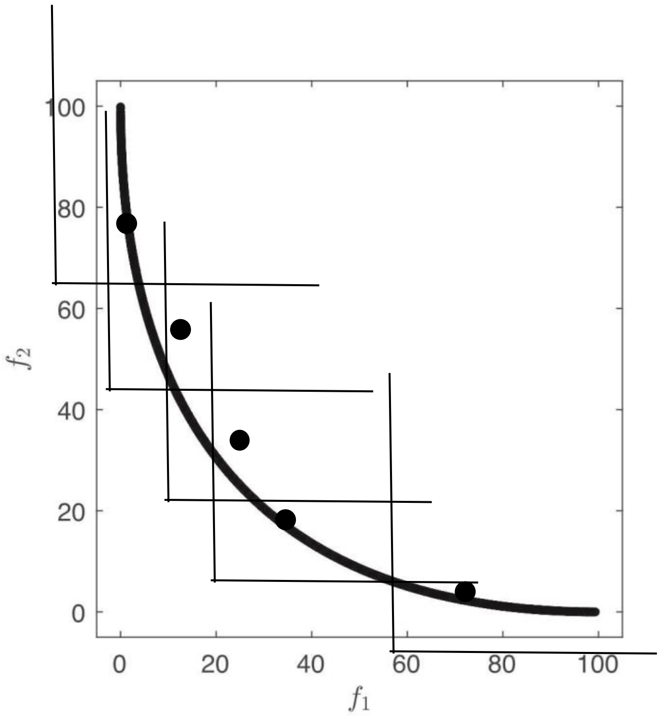


ε-approximations of the PF

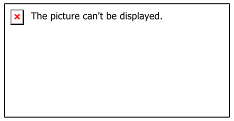
The picture can't be displayed.

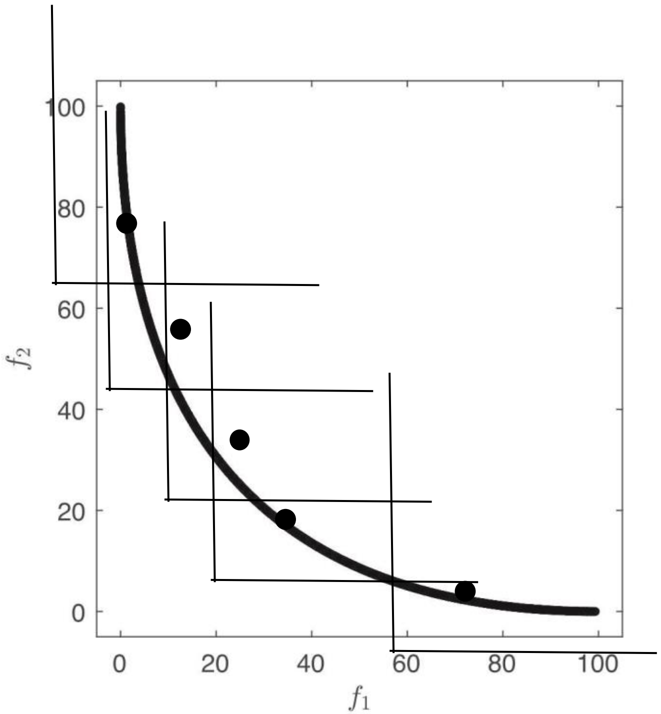


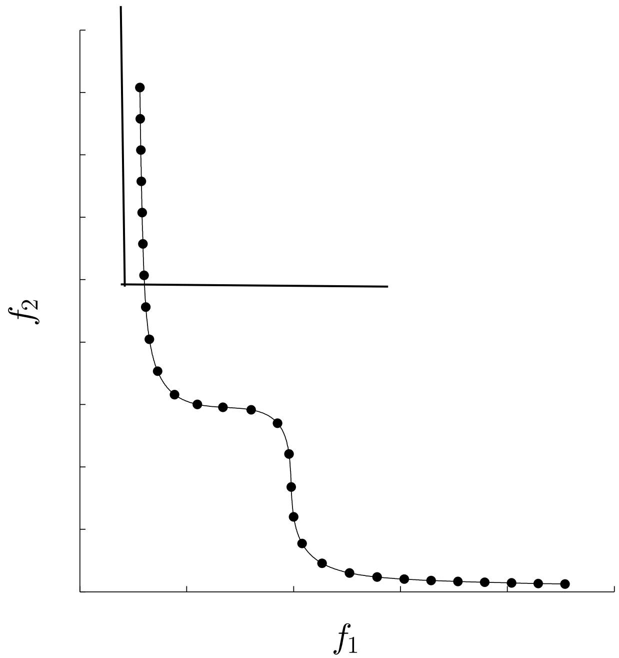
ε-approximations of the PF



ε-approximations of the PF





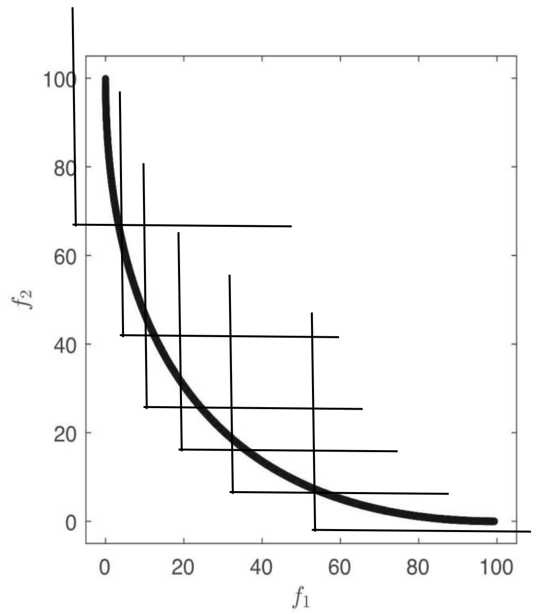


O. Schütze

Archivers for ε-approximations of the PF

ArchiveUpdateEps1 obtain ε–approx PS

‘cover’ the entire PF via ε–dominance



ArchiveUpdateEps2 obtain $\varepsilon -$ PS

as AU-Eps1, but aim for Pareto optimal archive elements

ArchiveUpdateTight1 obtain $(\tau, \varepsilon) -$tight ε–approx PS

as AU-Eps1, but target for a gap-free approximation (wrt a given value τ)

ArchiveUpdateTight2 obtain $\varepsilon -$tight ε–PS

as AU-Tight1, but aim for Pareto optimal archive elements

ArchiveUpdateTight2

The picture can't be displayed.

Theoretical results for ArchiveUpdateTight2:

- Monotonic behavior (in the sense of $\varepsilon -$ dominance)
- Yields ε-approximations of the Pareto front that are gap free (wrt τ) after finitely many steps with probabilty one , i.e.,

$$d_H\big(F\big(P_Q\big), F(A_i)\big) \leq \tau, \forall i \geq i_0$$

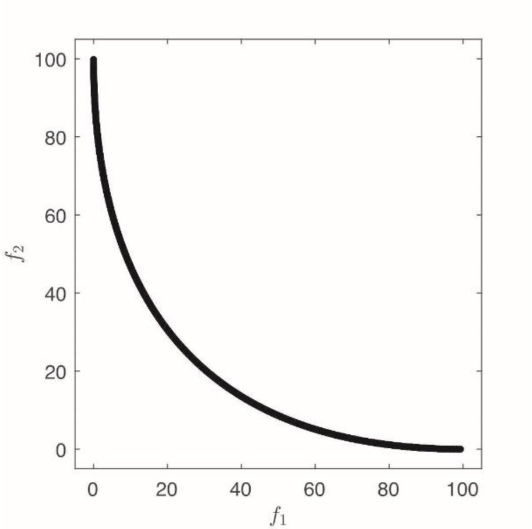
- Archive elements converge toward Pareto set, i.e.,

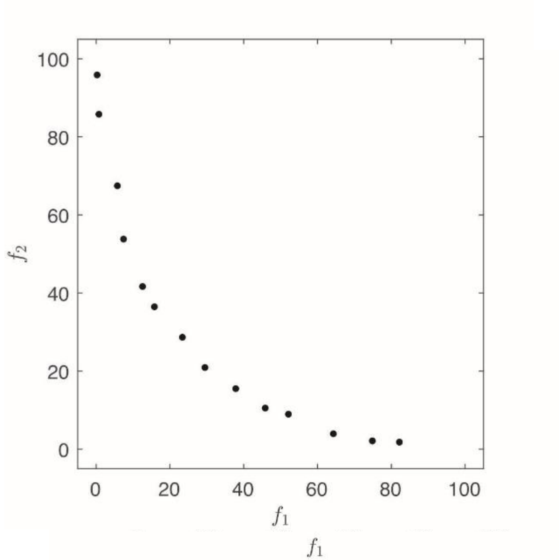
$$\lim_{i \to \infty} \mathrm{dist}\big(A_i, P_Q\big) = 0$$

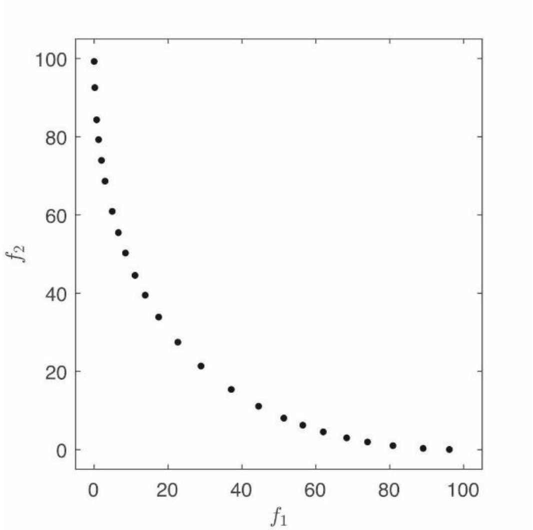
Comparison

The picture can't be displayed.

N = 100,000







AU-P_Q
$|A| = 2353$
$T = 39.1\ s$

AU-Eps1
$|A| = 14$
$T = 1.4\ s$
static ($i \geq i_0$)

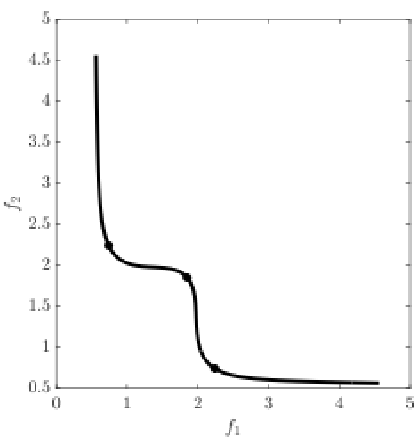
AU-Tight2
$|A| = 24$
$T = 3.5\ s$
quasi-static

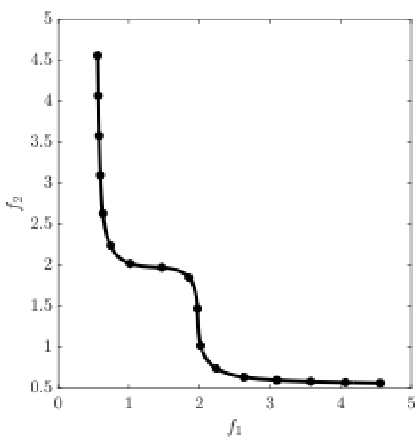
Task 3: A Priori Bounded Archiver

The picture can't be displayed.

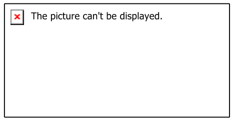
Task: A priori bounded archiver that aims for Hausdorff approximations of the Pareto front

- Use ArchiveUpdateTight archivers as basis
- Main question: how to adaptively choose discretization parameters ε and τ?





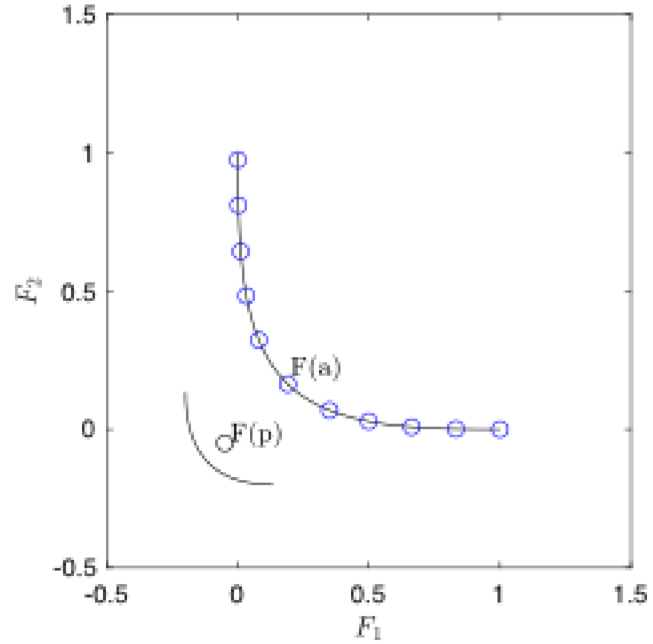
[Rudolph et al., 2016]
Optimal Hausdorff archives prefer evently spread solutions along the PF



Update of ε and τ (Basic Idea)

1.) **Archive overflow:** increase values of ε and τ

2.) **F(p) promising but too far from F(A) :**
restart ε and τ
new connected component of the PF may be found
or a new (local) front which makes current values
of the discretization parameters obsolete



Results ArchiveUpdateHD

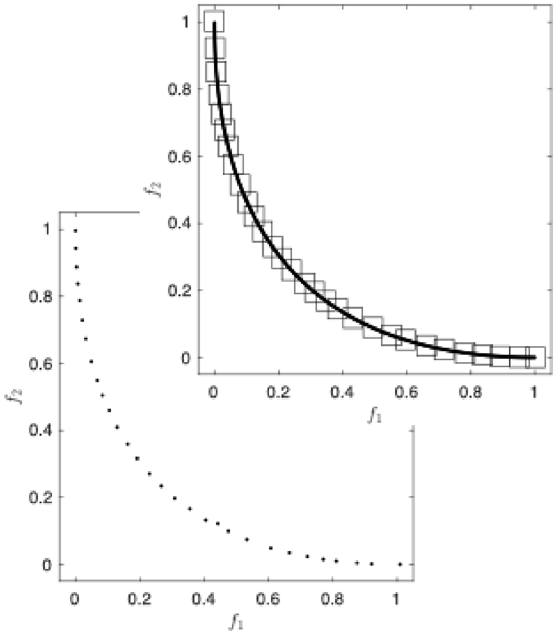
Under the assumptions of Thm 3 (e.g., no weak Pareto points in $Q\backslash P_Q$) we can expect from an application of ArchiveUpdateHD:

- Quasi-monotonic behavior: $\tau_i = \tau^+$ for all i$\geq i_1$
- Gap free approximation: $d_H(F(P_Q), F(A_i)) \leq \tau^+$for all i$\geq i_2$
- Archive elements converge toward Pareto set: $\lim_{i\to\infty} \text{dist}(A_i, P_Q) = 0$

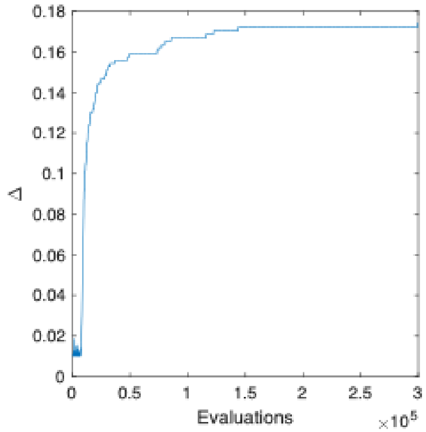
$\tau^+$is a (tight) upper bound on $d_H(F(P_Q), F(A_i))$ which is computed during the run of AU-HD with out any prior knowledge or assumption on the MOP/PF.

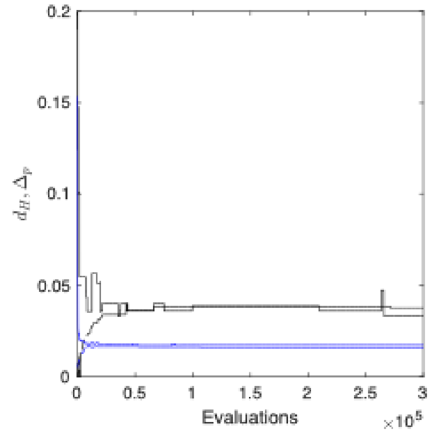
[Sch. and Hernandez., 2022]

Example 1: CONV2 (unimodal)



box covering



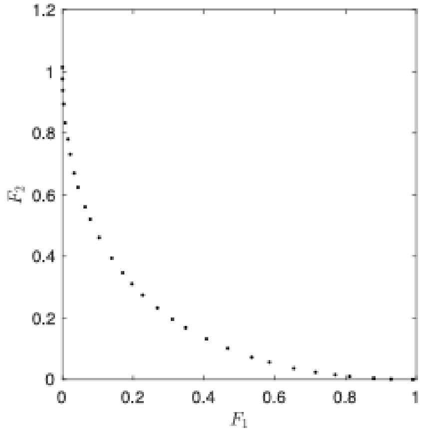


final archive

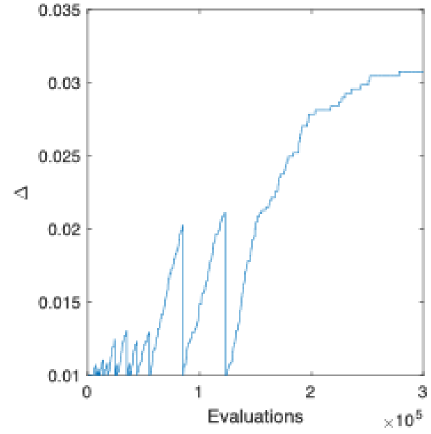
τ_i

estimated vs real approximation quality (d_H and Δ_2)

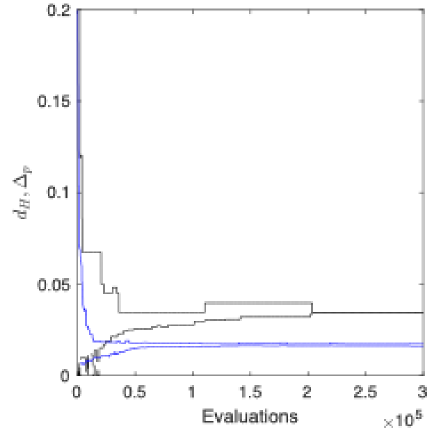
Example 2: RUD3 (multimodal)



final archive



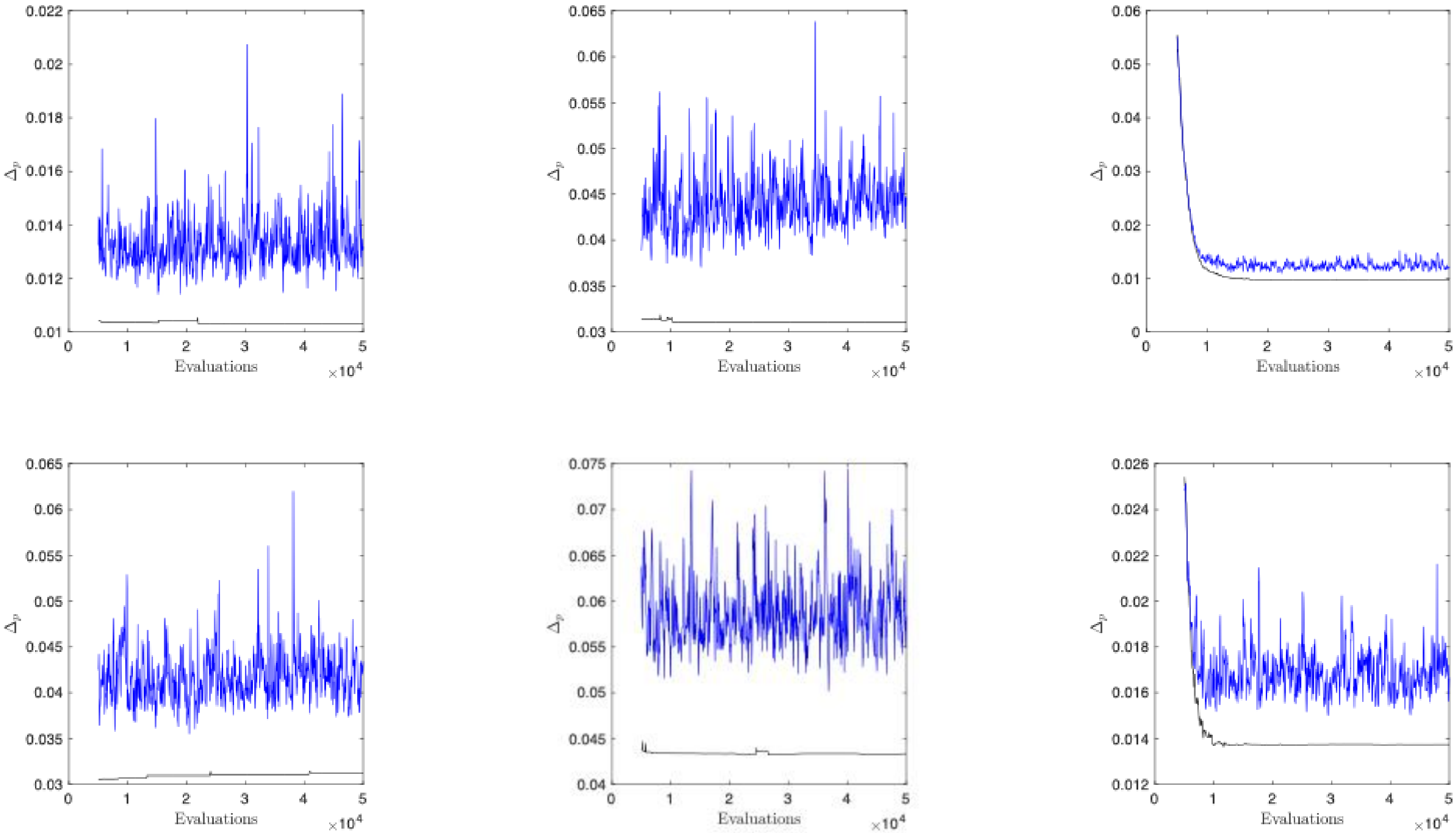
τ_i



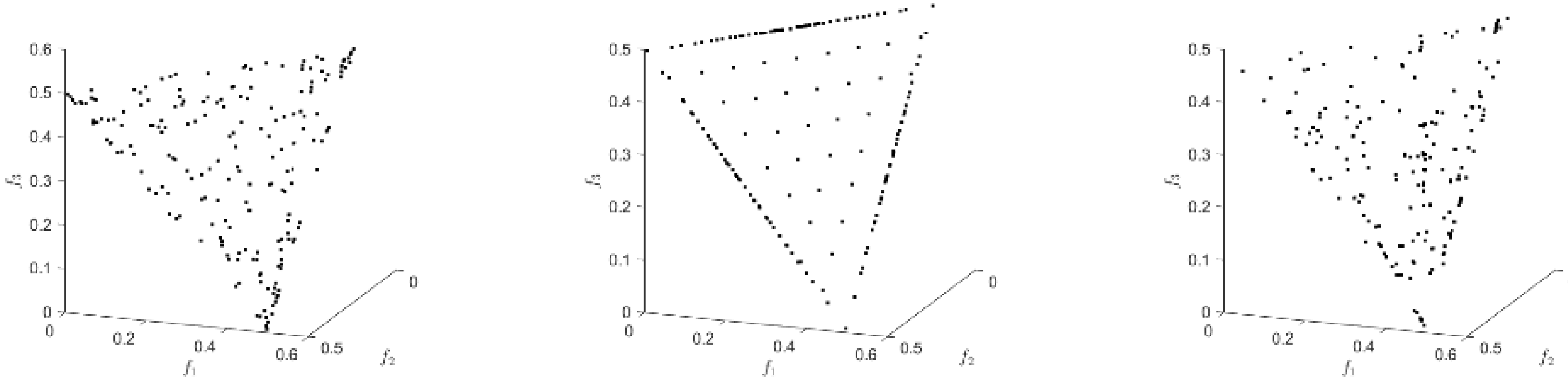
estimated vs real approximation quality (d_H and Δ_2)

The picture can't be displayed.

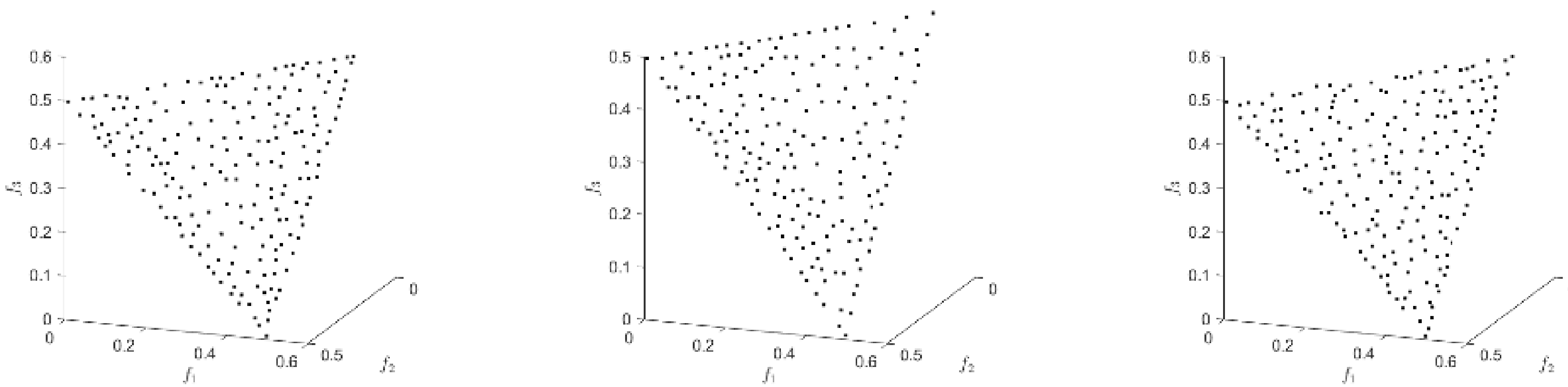
Approx. Qualities (Δ_2) NSGA-II vs NSGA-II-A



Base MOEA vs External Archive



NSGA-II MOEA/D SMS-EMOA



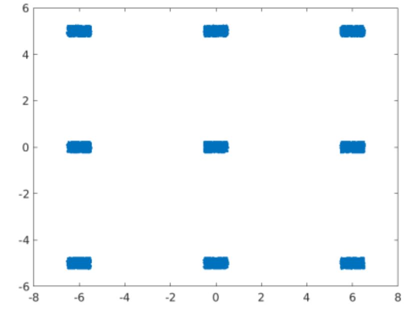
Beyond Pareto Fronts

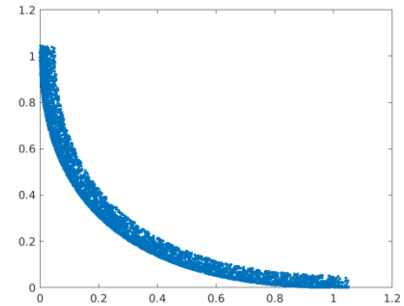
The picture can't be displayed.

Archiving is not restricted to Pareto front approximations, but is relevant e.g. for

- Detection of nearly optimal solutions
- Multi-objective Multimodal optimization (Pareto set or locally optimal solutions)
- Landscape analysis
- Particular regions of interest (e.g., knee regions, aspiration sets)
- Bilevel multi-objective optimization
- Robust multi-objective optimization
- …

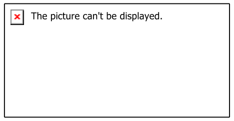
Finally, archiving is of course also not restricted to EMO.





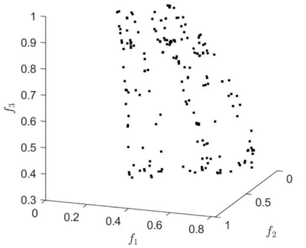
$N_{Q,\varepsilon}$ of SYMPART

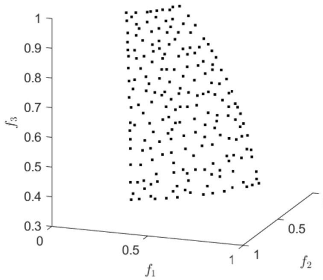
[Sch. et al., 2024]



Conclusions

- Archiving in EMO, focus on PF approximations
- PF manifold → not one single 'best' archiver to be expected
- Examples of unbounded, implicitly bounded and a priori bounded archivers
- Archiving has
  - the main influence of the convergence properties of the algorithm
  - a significant impact on the overall perofrmance of the algorithm
- Existing archivers can be integrated into any MOEA (either as alternative selection strategy or as external archivers)





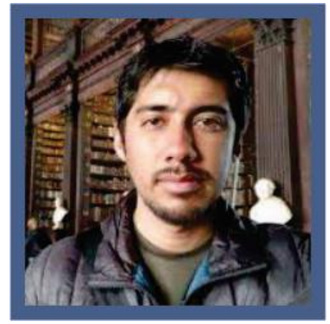
Yet, archiving is still underexplored in EMO (EO)!

Open Issues

- Other sets of interest
- Other problem classes
- Constraint handling (CDP, CV)
- Convergence rates
- Application to discrete problems
- Improvement of overall performance
  - Interplay of algorithm elements
  - External archives: 2 selection strategies used

Thank You I

The picture can't be displayed.

Thank You II

The picture can't be displayed.

Questions?

Emails:

- schuetze@cs.cinvestav.mx
- carlos.hernandez@iimas.unam.mx

Web pages:

- neo.cinvestav.mx/Group
- http://cihdezc.github.io/
- https://github.com/NumericalEvolutionaryOptimization/archivers