# DMS - Direct MultiSearch

J. F. Aguilar Madeira

IDMEC - IST and ADM - ISEL

# Presentation Outline

# Presentation Outline

# Problem Features

**MultiObjective Derivative-Free Optimization**

$$\min_{x\in\Omega\subseteq\mathbb{R}^n} F(x) \equiv \left(f_1(x), f_2(x), \ldots, f_m(x)\right)^\top$$

$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\},\ j = 1, \ldots, m > 1$$

# Problem Features

**MultiObjective Derivative-Free Optimization**

$$\min_{x \in \Omega \subseteq \mathbb{R}^n} F(x) \equiv (f_1(x), f_2(x), \ldots, f_m(x))^\top$$

$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \; j = 1, \ldots, m > 1$$

- We will make use of the extreme barrier function:

$$f_j(x) = \begin{cases} f_j(x) & \text{if } x \in \Omega, \\ +\infty & \text{otherwise.} \end{cases}$$

# Problem Features

**MultiObjective Derivative-Free Optimization**

$$\min_{x \in \Omega \subseteq \mathbb{R}^n} F(x) \equiv (f_1(x), f_2(x), \ldots, f_m(x))^\top$$

$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \; j = 1, \ldots, m > 1$$

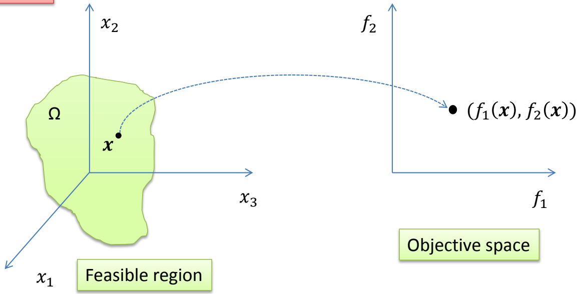
# Problem Features

**MultiObjective Derivative-Free Optimization**

$$\min_{x \in \Omega \subseteq \mathbb{R}^n} F(x) \equiv (f_1(x), f_2(x), \ldots, f_m(x))^\top$$

$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \; j = 1, \ldots, m > 1$$

Example



# Problem Features

**MultiObjective Derivative-Free Optimization**

$$\min_{x \in \Omega \subseteq \mathbb{R}^n} F(x) \equiv (f_1(x), f_2(x), \ldots, f_m(x))^\top$$

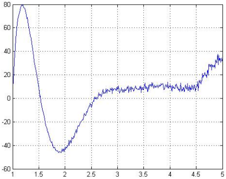
$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \; j = 1, \ldots, m > 1$$

# Problem Features

**MultiObjective Derivative-Free Optimization**

$$\min_{x \in \Omega \subseteq \mathbb{R}^n} F(x) \equiv (f_1(x), f_2(x), \ldots, f_m(x))^\top$$

$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \; j = 1, \ldots, m > 1$$



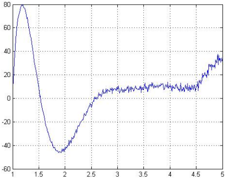

- expensive function evaluations (possibly subject to noise)
- black box functions, i.e. functions for which we do not know an algebraic expression (are evaluated through a computer code)

# Problem Features

**MultiObjective Derivative-Free Optimization**

$$\min_{x \in \Omega \subseteq \mathbb{R}^n} F(x) \equiv (f_1(x), f_2(x), \ldots, f_m(x))^\top$$

$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \; j = 1, \ldots, m > 1$$





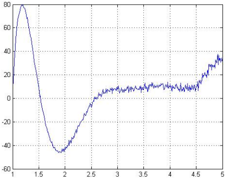
- expensive function evaluations (possibly subject to noise)
- black box functions, i.e. functions for which we do not know an algebraic expression (are evaluated through a computer code)
- functions with unknown derivatives

# Problem Features

**MultiObjective Derivative-Free Optimization**

$$\min_{x \in \Omega \subseteq \mathbb{R}^n} F(x) \equiv (f_1(x), f_2(x), \ldots, f_m(x))^\top$$

$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \ j = 1, \ldots, m > 1$$





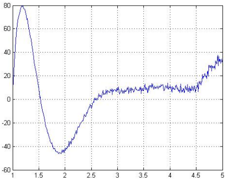
- expensive function evaluations (possibly subject to noise)
- black box functions, i.e. functions for which we do not know an algebraic expression (are evaluated through a computer code)
- functions with unknown derivatives
- source code not available for use (e.g. laboratorial results)

# Problem Features

**MultiObjective Derivative-Free Optimization**

$$\min_{x \in \Omega \subseteq \mathbb{R}^n} F(x) \equiv (f_1(x), f_2(x), \ldots, f_m(x))^\top$$

$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \; j = 1, \ldots, m > 1$$





- expensive function evaluations (possibly subject to noise)
- black box functions, i.e. functions for which we do not know an algebraic expression (are evaluated through a computer code)
- functions with unknown derivatives
- source code not available for use (e.g. laboratorial results)
- unpractical to compute approximations to derivatives

# Problem Features

## Dealing with conflicting objectives in MOO

In case of conflicting objectives we may ...

1. State some objectives as constraints, or
2. Aggregate objectives by means of an utility function $u : \mathbb{R}^m \to \mathbb{R}$

$$\text{Minimize } u(f_1(x), ..., f_m(x)), \text{ or}$$

3. Prioritize objective functions (lexicographic order), or
4. Compute all Pareto optimal solutions and analyze trade-offs, then decide

# Problem Features

**Dealing with conflicting objectives in MOO**

In case of conflicting objectives we may ...

1. State some objectives as constraints, or
2. Aggregate objectives by means of an utility function $u : \mathbb{R}^m \to \mathbb{R}$

$$\text{Minimize } u(f_1(x), ..., f_m(x)), \text{ or}$$

3. Prioritize objective functions (lexicographic order), or
4. Compute all Pareto optimal solutions and analyze trade-offs, then decide

## Classification

- $1, 2, 3$ are *a-priori methods*: decision-making before optimization.
- $4$ is a *a-posteriori method*: decision-making after optimization.
- *Progressive methods* allows user interaction during search and can combine $1, 2, 3$ or $4$.

# Problem Features

**Pareto domination**

$x \prec y$ ($x$ dominates $y$) $\iff$ for all $i \in \{1, \ldots, m\}$: $f_i(x) \leq f_i(y)$, and exists $j \in \{1, \ldots, m\}$: $f_j(x) < f_j(y)$

# Problem Features

**Pareto domination**

$x \prec y$ ($x$ dominates $y$) $\iff$ for all $i \in \{1, \ldots, m\}$:
$f_i(x) \leq f_i(y)$, and exists $j \in \{1, \ldots, m\}$: $f_j(x) < f_j(y)$

Some authors state the previous definition by considering a strict partial order in the cone $\mathbb{R}^m_+ = \{y \in \mathbb{R}^m : y \geq 0\}$. In this case, given two points $x, y$ in $\Omega$, we have the following equivalencies:

# Problem Features

**Pareto domination**

$x \prec y$ ($x$ dominates $y$) $\iff$ for all $i \in \{1, \dots, m\}$: $f_i(x) \leq f_i(y)$, and exists $j \in \{1, \dots, m\}$: $f_j(x) < f_j(y)$

Some authors state the previous definition by considering a strict partial order in the cone $\mathbb{R}^m_+ = \{y \in \mathbb{R}^m : y \geq 0\}$. In this case, given two points $x, y$ in $\Omega$, we have the following equivalencies:

**Pareto domination**

$$x \prec y \iff F(x) \prec_F F(y) \iff F(y) - F(x) \in \mathbb{R}^m_+ \setminus \{0\}$$

# Problem Features

**Pareto domination**

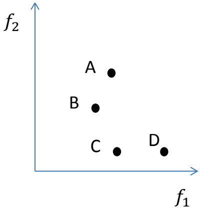
$x \prec y$ ($x$ dominates $y$) $\iff$ for all $i \in \{1, \ldots, m\}$:
$f_i(x) \leq f_i(y)$, and exists $j \in \{1, \ldots, m\}$: $f_j(x) < f_j(y)$

# Problem Features

## Pareto domination

$x \prec y$ ($x$ dominates $y$) $\iff$ for all $i \in \{1, \ldots, m\}$:
$f_i(x) \leq f_i(y)$, and exists $j \in \{1, \ldots, m\}$: $f_j(x) < f_j(y)$

Example




Objective space

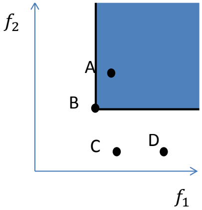
# Problem Features

## Pareto domination

$x \prec y$ ($x$ dominates $y$) $\iff$ for all $i \in \{1, \ldots, m\}$:
$f_i(x) \leq f_i(y)$, and exists $j \in \{1, \ldots, m\}$: $f_j(x) < f_j(y)$

Example

B dominates A




Objective space

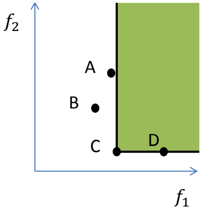
**Pareto domination**

$x \prec y$ ($x$ dominates $y$) $\iff$ for all $i \in \{1, \ldots, m\}$: $f_i(x) \leq f_i(y)$, and exists $j \in \{1, \ldots, m\}$: $f_j(x) < f_j(y)$

Example

B dominates A

C dominates D




Objective space

# Problem Features

## Pareto domination

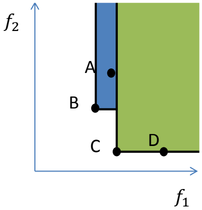
$$x \prec y \; (x \text{ dominates } y) \iff \text{for all } i \in \{1, \ldots, m\}:$$
$$f_i(x) \leq f_i(y), \text{ and exists } j \in \{1, \ldots, m\}: \; f_j(x) < f_j(y)$$

Example

B dominates A

C dominates D

B and C are nondominated




Objective space

# Problem Features

## Pareto domination

$$x \prec y \ (x \text{ dominates } y) \iff \text{for all } i \in \{1,\ldots,m\}: \\ f_i(x) \leq f_i(y), \text{ and exists } j \in \{1,\ldots,m\}: \ f_j(x) < f_j(y)$$

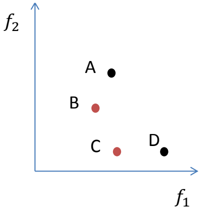
Example

B dominates A

C dominates D

B and C are nondominated

Pareto front: {B,C}




Objective space

# Presentation Outline

# Direct Search Methods (DSM) Main Lines

On single objective optimization

# Direct Search Methods (DSM) Main Lines

On single objective optimization

**Definition**

- Sample the objective function at a finite number of points at each iteration.

# Direct Search Methods (DSM) Main Lines

On single objective optimization

**Definition**

- Sample the objective function at a finite number of points at each iteration.
- Base actions on those function values.

# Direct Search Methods (DSM) Main Lines

On single objective optimization

**Definition**

- Sample the objective function at a finite number of points at each iteration.
- Base actions on those function values.
- Do not depend on derivative approximation or explicit or implicit models for the objective function.

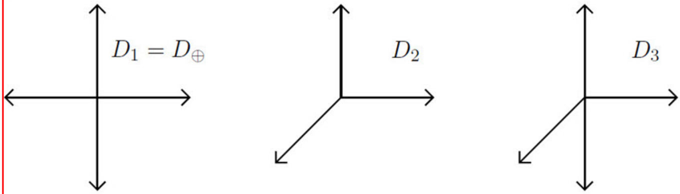
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.

# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.

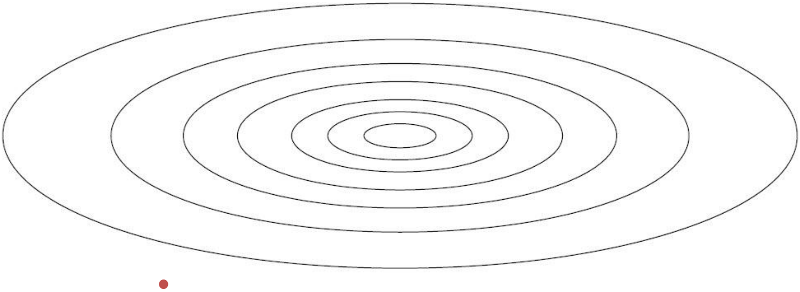
Examples of positive spanning sets / positive bases




$$D_\oplus = D_1 = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \quad \text{maximal positive basis}$$

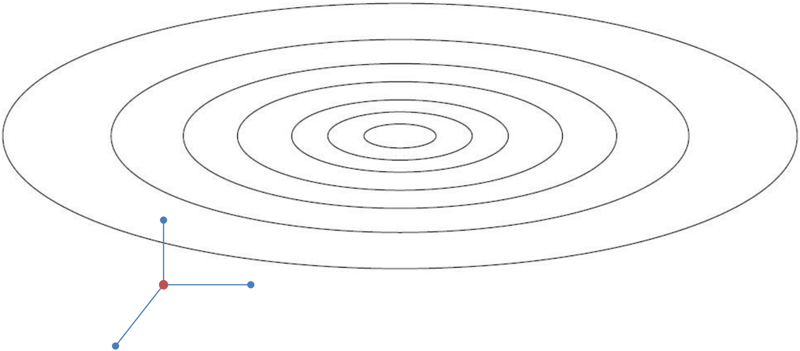
$$D_2 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \quad \text{minimal positive basis}$$

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



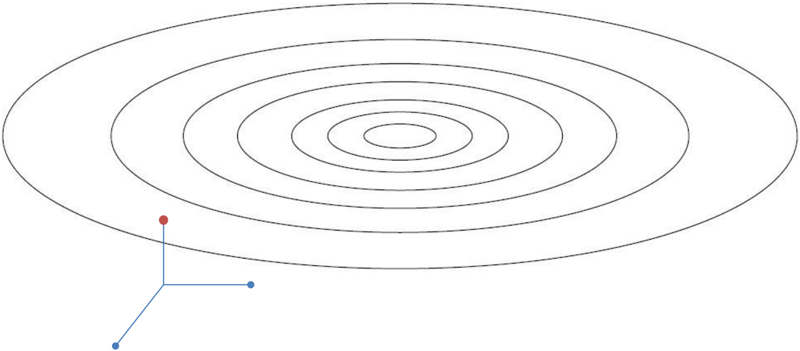
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



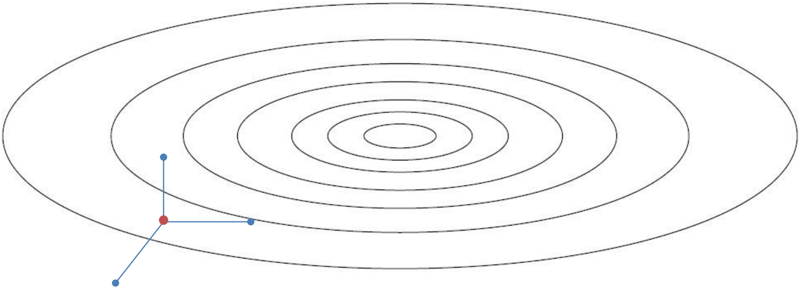
## Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



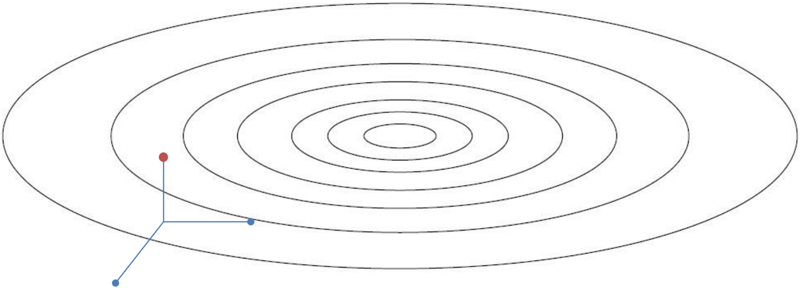
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



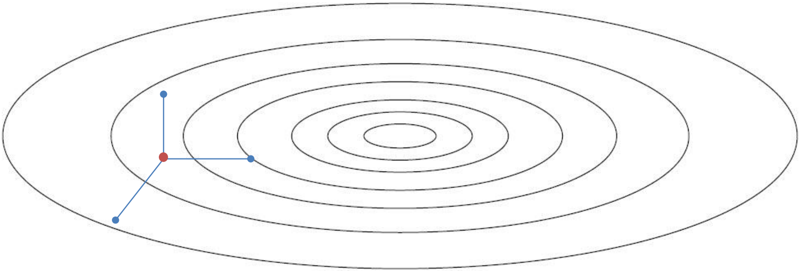
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



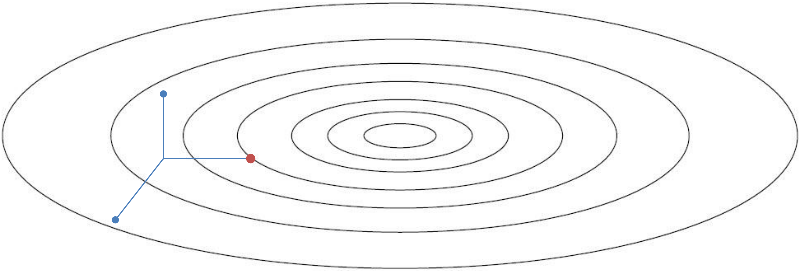
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



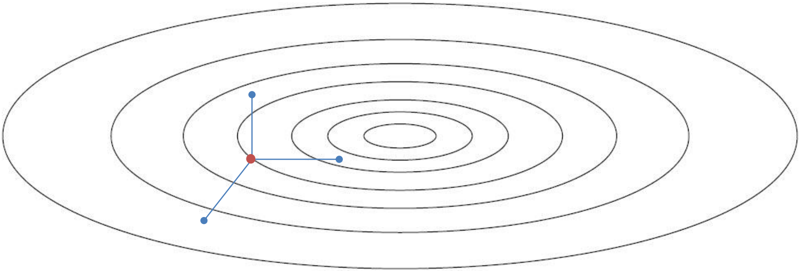
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



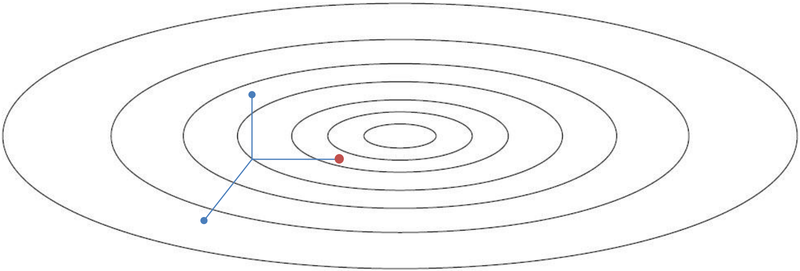
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



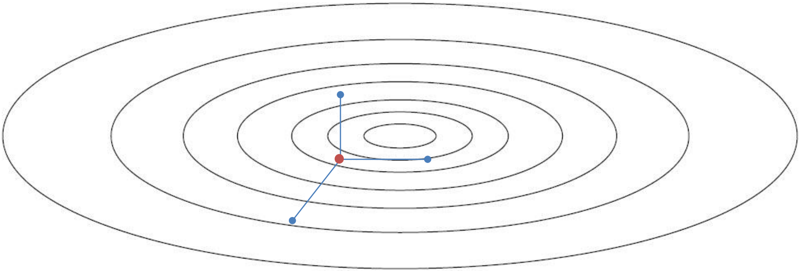
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



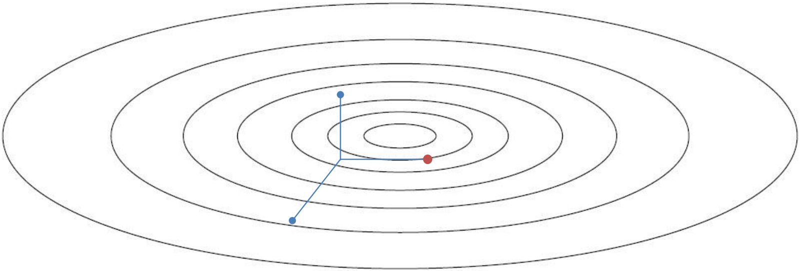
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



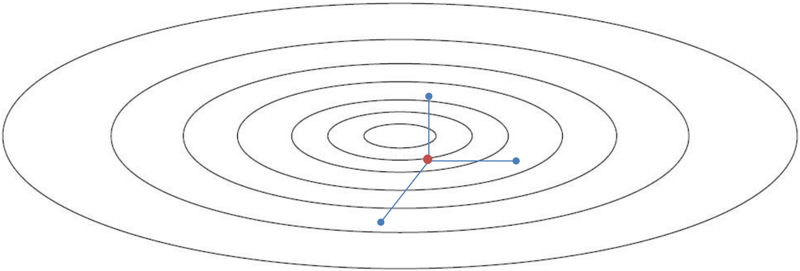
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



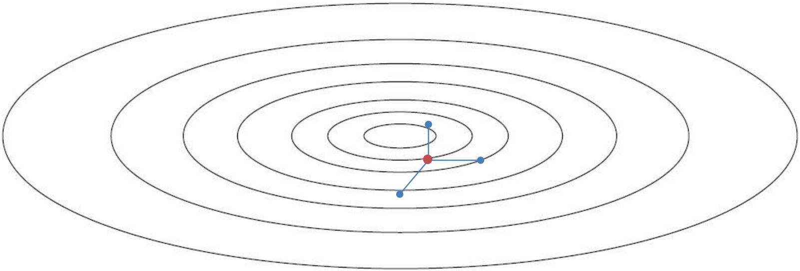
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



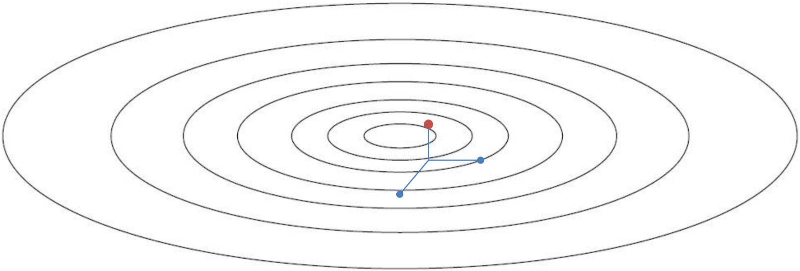
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



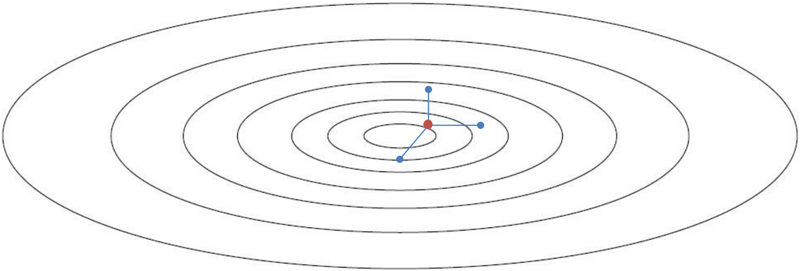
# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.

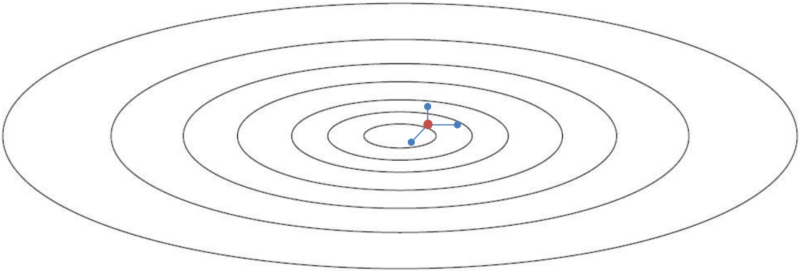


# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.

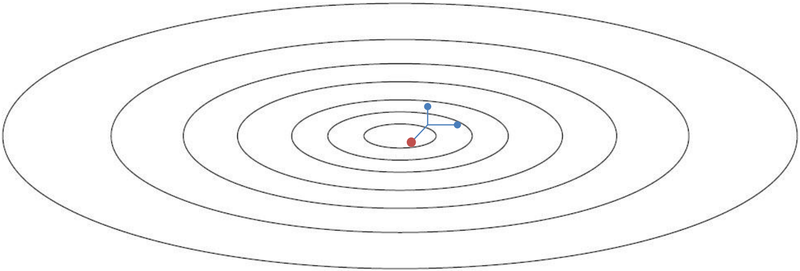


- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



# Direct Search Methods (DSM) Main Lines

- Direct search of directional type: Achieve descent by using positive spanning sets and moving in the directions of the best points.



# Direct Search Methods (DSM) Main Lines

On multiobjective optimization

On multiobjective optimization

- **DSM for Multiobjective Optimization (MOO)** are in the **beginning of their development**:

# Direct Search Methods (DSM) Main Lines

On multiobjective optimization

- **DSM for Multiobjective Optimization (MOO)** are in the **beginning of their development**:

- *Zhong et al.*(2010) proposed an empirical algorithm based on compass search, but for which no convergence analysis was provided

# Direct Search Methods (DSM) Main Lines

On multiobjective optimization

- **DSM for Multiobjective Optimization (MOO)** are in the **beginning of their development**:
- *Zhong et al.*(2010) proposed an empirical algorithm based on compass search, but for which no convergence analysis was provided
- to our knowledge, only two DSM were proposed for general Derivative-free Multiobjective Optimization:
  - MULTIMADS (2010)
  - DMS (2011)

# Presentation Outline

1. Motivation and General Concepts
2. Direct Search Methods (DSM) for MOO
3. Direct MultiSearch (DMS)
4. Conclusions and references
5. Convergence Analysis in DSM
6. Numerical Results

# Direct MultiSearch (DMS) [a] Main Lines

[a] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz and L. N. Vicente (2011): Direct Multisearch for Multiobjective Optimization. SIAM J. Optim , 21(3), pages 1109-1140.

# Direct MultiSearch (DMS) [a] Main Lines

[a] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz and L. N. Vicente (2011): Direct Multisearch for Multiobjective Optimization. SIAM J. Optim , 21(3), pages 1109-1140.

- does not aggregate any of the objective functions

# Direct MultiSearch (DMS) [a] Main Lines

[a] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz and L. N. Vicente (2011): Direct Multisearch for Multiobjective Optimization. SIAM J. Optim , 21(3), pages 1109-1140.

- does not aggregate any of the objective functions
- generalizes ALL direct-search methods of directional type to multiobjective optimization (MOO) (such as pattern search and generalized pattern search (GPS), generating set search (GSS), and mesh adaptive direct search (MADS))

# Direct MultiSearch (DMS) [a] Main Lines

[a] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz and L. N. Vicente (2011): Direct Multisearch for Multiobjective Optimization. SIAM J. Optim , 21(3), pages 1109-1140.

- does not aggregate any of the objective functions
- generalizes ALL direct-search methods of directional type to multiobjective optimization (MOO) (such as pattern search and generalized pattern search (GPS), generating set search (GSS), and mesh adaptive direct search (MADS))
- makes use of search/poll paradigm

# Direct MultiSearch (DMS) [a] Main Lines

[a] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz and L. N. Vicente (2011): Direct Multisearch for Multiobjective Optimization. SIAM J. Optim , 21(3), pages 1109-1140.

- the search step is optional:
    - does not interfere in the convergence analysis of the algorithms
    - the purpose of the search step is to improve the computational performance / disseminate the search

# Direct MultiSearch (DMS) [a] Main Lines

[a] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz and L. N. Vicente (2011): Direct Multisearch for Multiobjective Optimization. SIAM J. Optim , 21(3), pages 1109-1140.

- the search step is optional:
  - does not interfere in the convergence analysis of the algorithms
  - the purpose of the search step is to improve the computational performance / disseminate the search
- keeps a list of feasible nondominated points

# Direct MultiSearch (DMS) [a] Main Lines

[a] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz and L. N. Vicente (2011): Direct Multisearch for Multiobjective Optimization. SIAM J. Optim , 21(3), pages 1109-1140.

- the search step is optional:
  - does not interfere in the convergence analysis of the algorithms
  - the purpose of the search step is to improve the computational performance / disseminate the search
- keeps a list of feasible nondominated points
- tries to capture the whole Pareto front from the polling procedure

# Direct MultiSearch (DMS) [a] Main Lines

[a] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz and L. N. Vicente (2011): Direct Multisearch for Multiobjective Optimization. SIAM J. Optim , 21(3), pages 1109-1140.

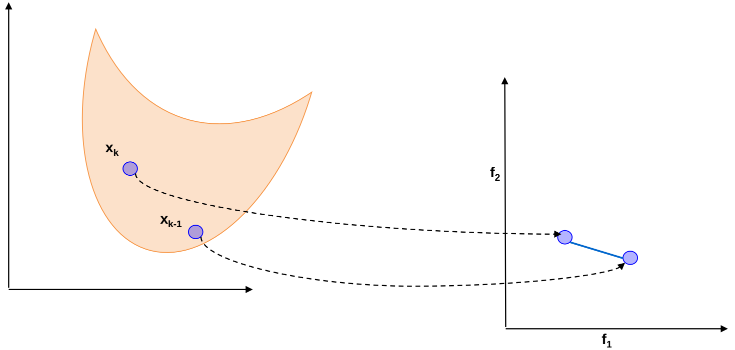
- the search step is optional:
  - does not interfere in the convergence analysis of the algorithms
  - the purpose of the search step is to improve the computational performance / disseminate the search
- keeps a list of feasible nondominated points
- tries to capture the whole Pareto front from the polling procedure
- poll centers are chosen from the list of feasible nondominated points

# Direct MultiSearch (DMS) [a] Main Lines

[a] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz and L. N. Vicente (2011): Direct Multisearch for Multiobjective Optimization. SIAM J. Optim , 21(3), pages 1109-1140.

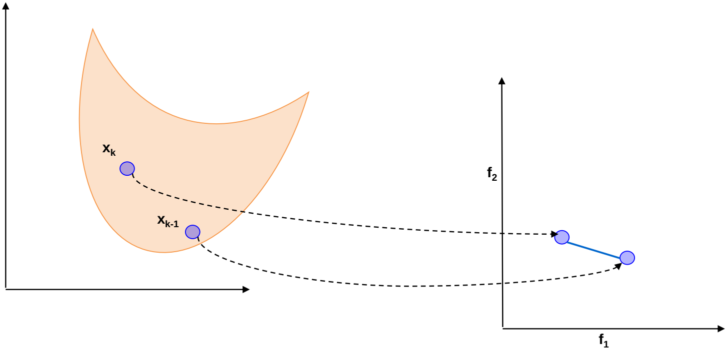
- the search step is optional:
  - does not interfere in the convergence analysis of the algorithms
  - the purpose of the search step is to improve the computational performance / disseminate the search
- keeps a list of feasible nondominated points
- tries to capture the whole Pareto front from the polling procedure
- poll centers are chosen from the list of feasible nondominated points
- successful iterations correspond to list changes: a new feasible nondominated point has been identified

# Poll Step Example (Biobjective Problem)




$L_k$

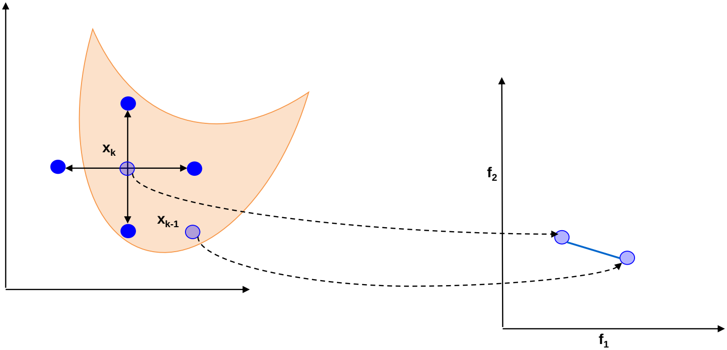
# Poll Step Example (Biobjective Problem)



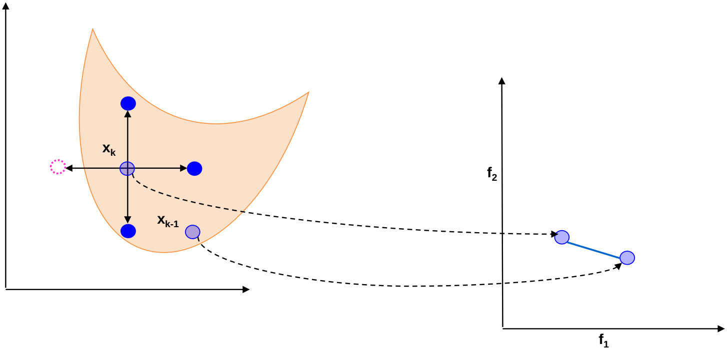

$L_k$

$L_k$ = list of nondominated points in iteration $k$

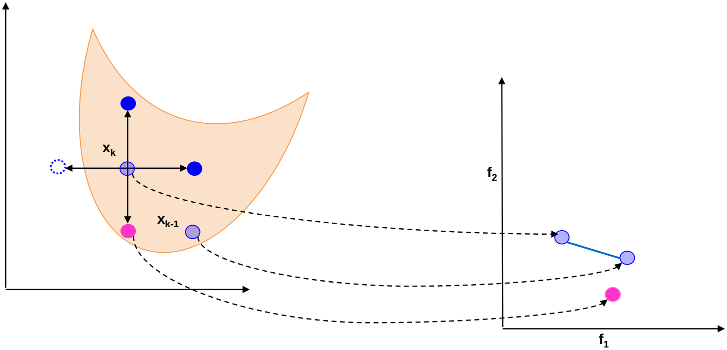
# Poll Step Example (Biobjective Problem)



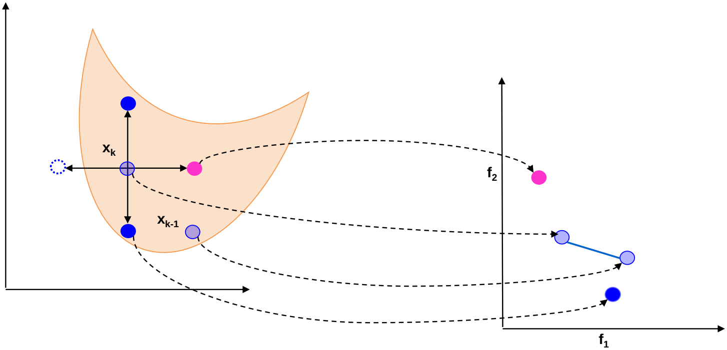
# Poll Step Example (Biobjective Problem)



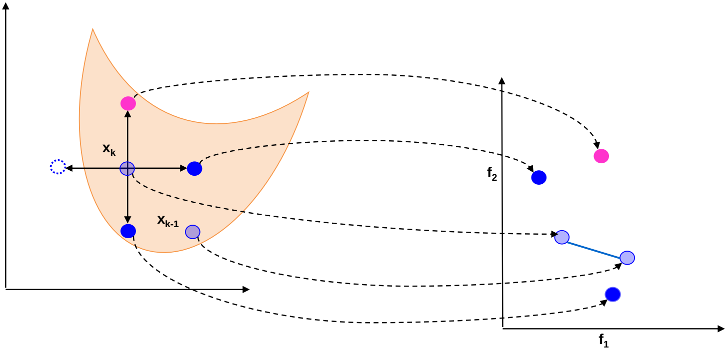
# Poll Step Example (Biobjective Problem)



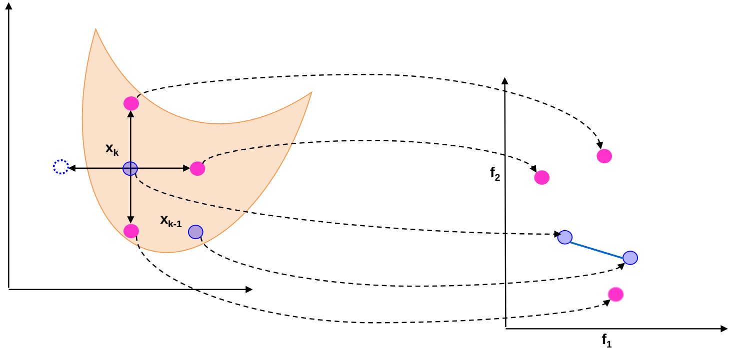
# Poll Step Example (Biobjective Problem)



# Poll Step Example (Biobjective Problem)



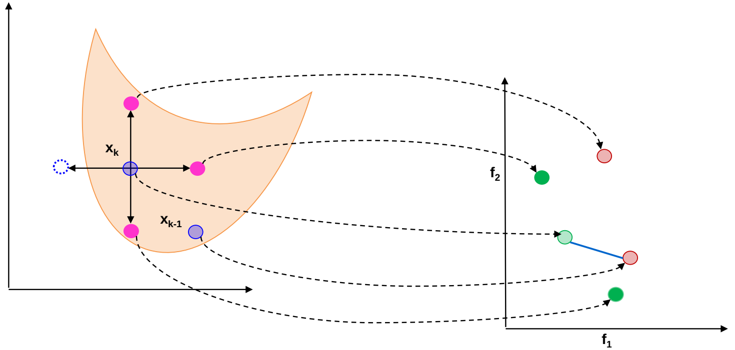
# Poll Step Example (Biobjective Problem)




$L_{add}$

$L_{add}$ = List of new feasible points

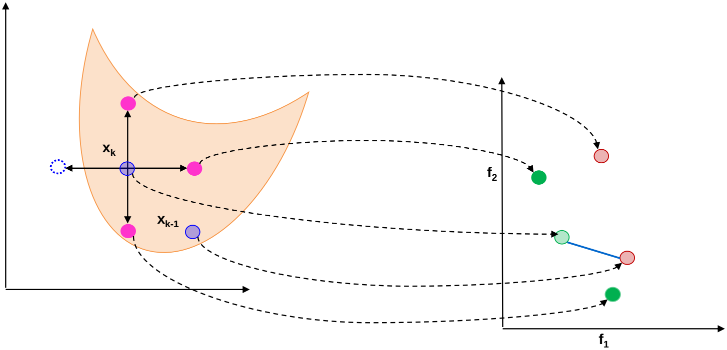
# Poll Step Example (Biobjective Problem)




$L_{filtered}$

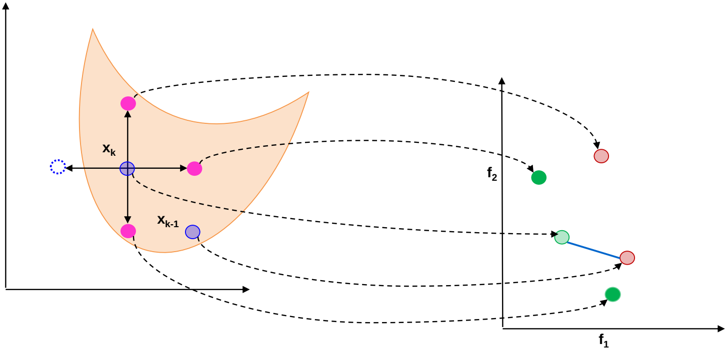
Remove dominated points from $L_k \cup L_{add}$

# Poll Step Example (Biobjective Problem)




$L_{trial} = L_{filtered}$

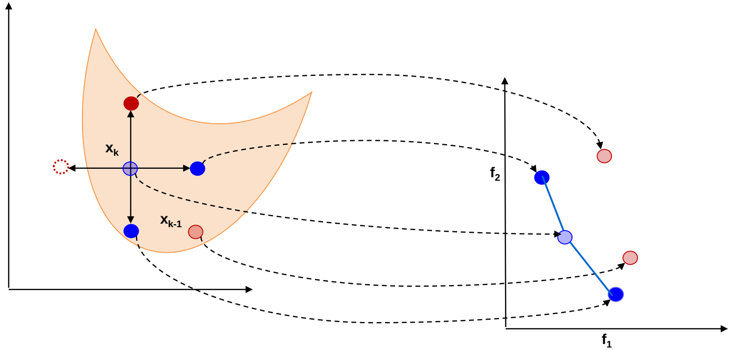
# Poll Step Example (Biobjective Problem)




$L_{trial} = L_{filtered}$

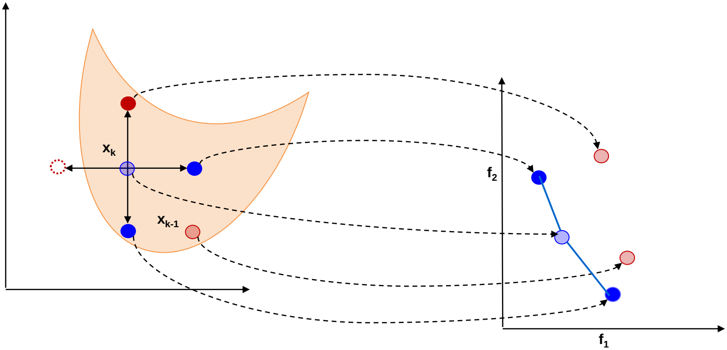
If $L_{trial} \neq L_k$ the iteration is declared successful, the step size is maintained.

# Poll Step Example (Biobjective Problem)




$L_{k+1}$

# Poll Step Example (Biobjective Problem)




$L_{k+1}$

$L_{k+1}$ = new list of nondominated points

# Direct MultiSearch for MOO

**Initialization**

Choose $x_0 \in \Omega$ with $F(x_0) < +\infty$, $\alpha_0 > 0$. Set $L_0 = \{(x_0; \alpha_0)\}$

**Selection of iterate point**

Order $L_k$ and select $(x_k; \alpha_k) \in L_k$

**Search Step (Optional)**

Evaluate a finite set of points $L_{add} = \{(z_s; \alpha_k)\}_{s \in S}$

$(L_k; L_{add}) \hookrightarrow L_{filtered} \hookrightarrow L_{trial}$

Suc → $L_{k+1} = L_{trial}$ (back to Selection of iterate point)

Unsuc ↓

**Poll Step**

Evaluate $L_{add} = \{(x_k + \alpha_k d; \alpha_k),\ d \in D_k\}$, with $D_k \subseteq \mathcal{D}$

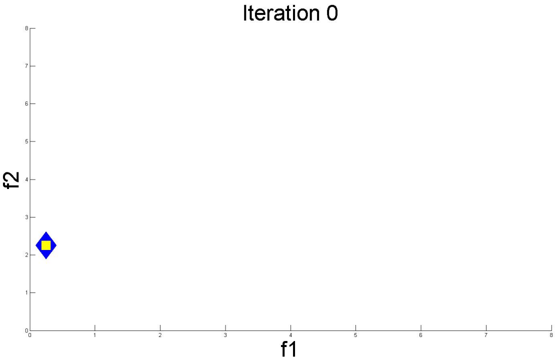
$(L_k; L_{add}) \hookrightarrow L_{filtered} \hookrightarrow L_{trial}$

Suc → $L_{k+1} = L_{trial}$ (back to Selection of iterate point)

Unsuc ↓

Decrease the step size (back to Selection of iterate point)

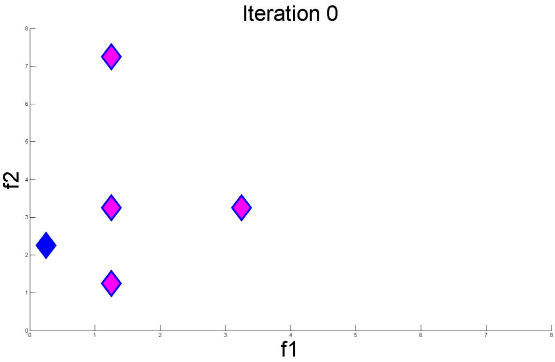
# Numerical Example - Problem SP1 [Huband et al.]




♦ Evaluated points since beginning

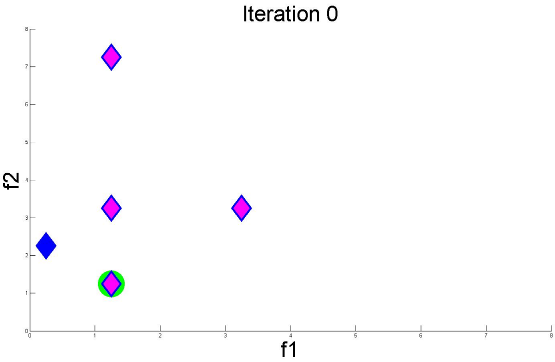
■ Current iterate list

# Numerical Example - Problem SP1 [Huband et al.]



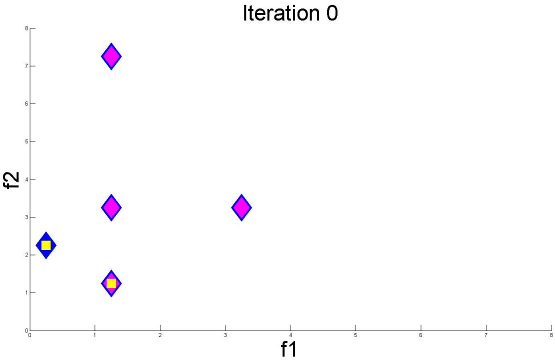

- Evaluated poll points
- Evaluated points since beginning

# Numerical Example - Problem SP1 [Huband et al.]



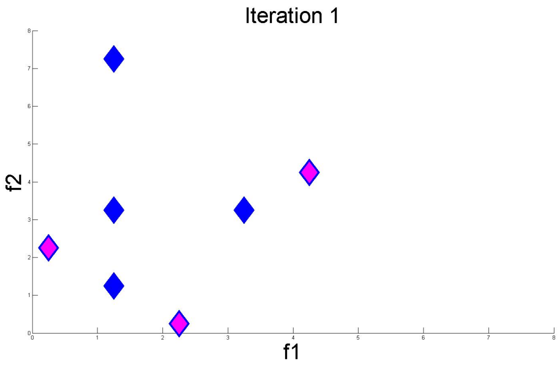

- Nondominated evaluated poll points

# Numerical Example - Problem SP1 [Huband et al.]



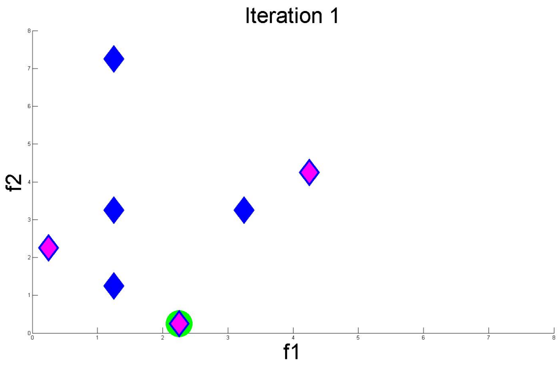

- Evaluated poll points
- Evaluated points since beginning
- Current iterate list

# Numerical Example - Problem SP1 [Huband et al.]



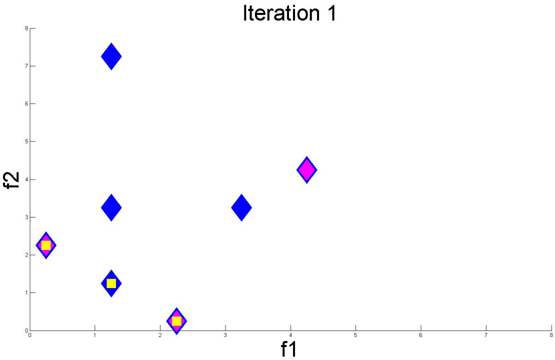

- Evaluated poll points
- Evaluated points since beginning

# Numerical Example - Problem SP1 [Huband et al.]




- Nondominated evaluated poll points

# Numerical Example - Problem SP1 [Huband et al.]

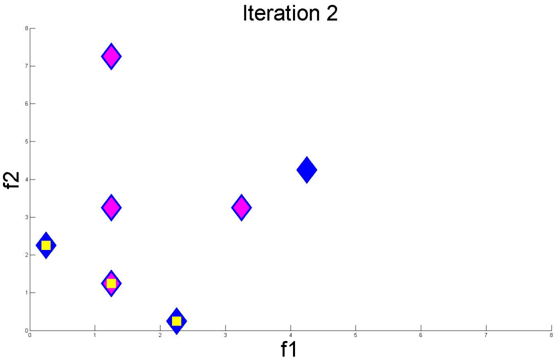



Evaluated poll points

Evaluated points since beginning

Current iterate list

# Numerical Example - Problem SP1 [Huband et al.]

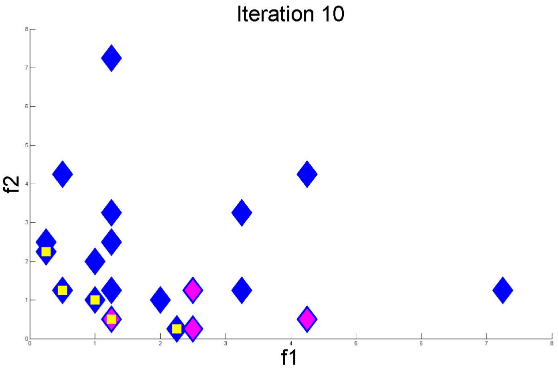



♦ Evaluated poll points

♦ Evaluated points since beginning

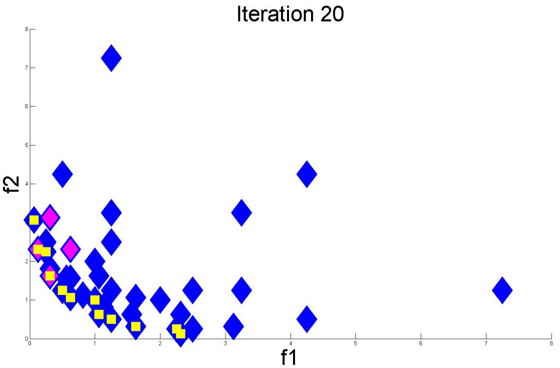
■ Current iterate list

# Numerical Example - Problem SP1 [Huband et al.]



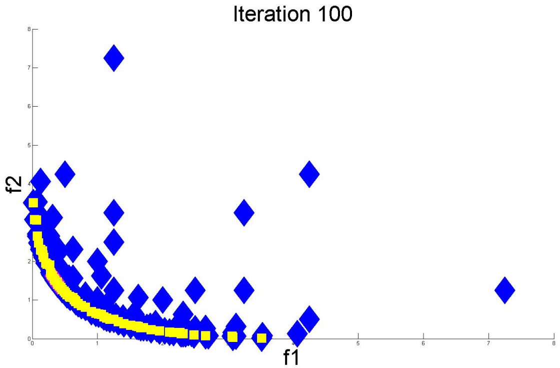

- Evaluated poll points
- Evaluated points since beginning
- Current iterate list

# Numerical Example - Problem SP1 [Huband et al.]



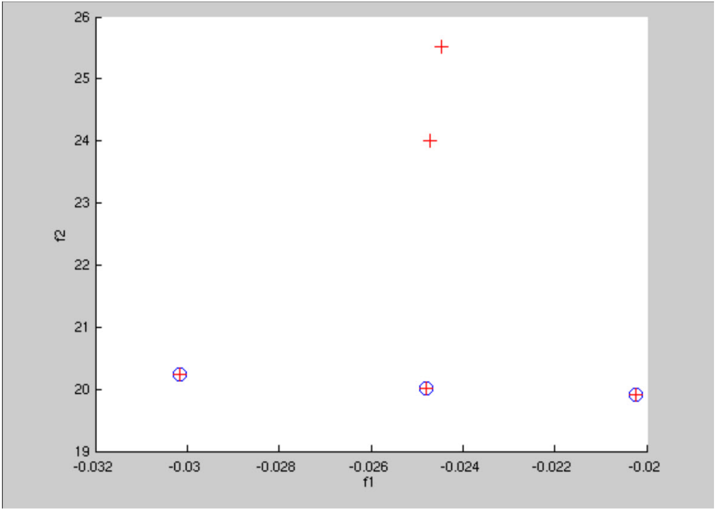

- Evaluated poll points
- Evaluated points since beginning
- Current iterate list

# Numerical Example - Problem SP1 [Huband et al.]



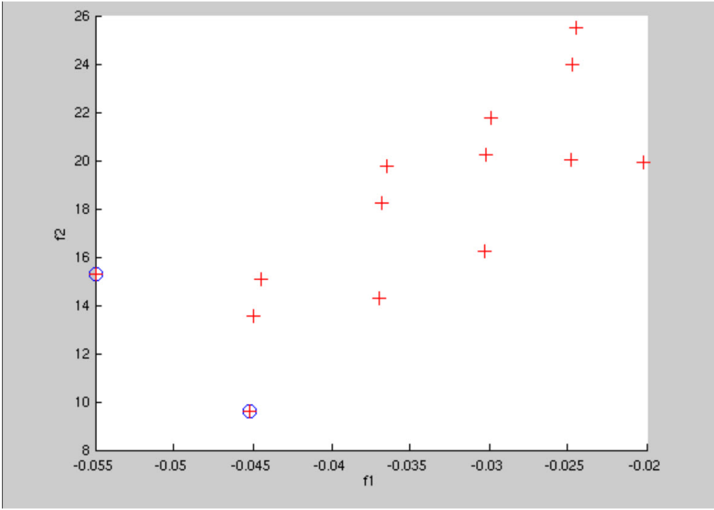

- Evaluated poll points
- Evaluated points since beginning
- Current iterate list

# Numerical Example - Laumanns [Laumanns et al.]



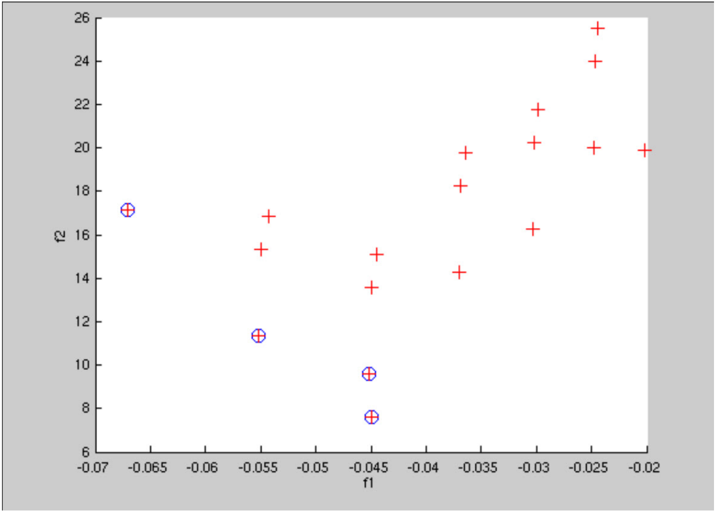

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



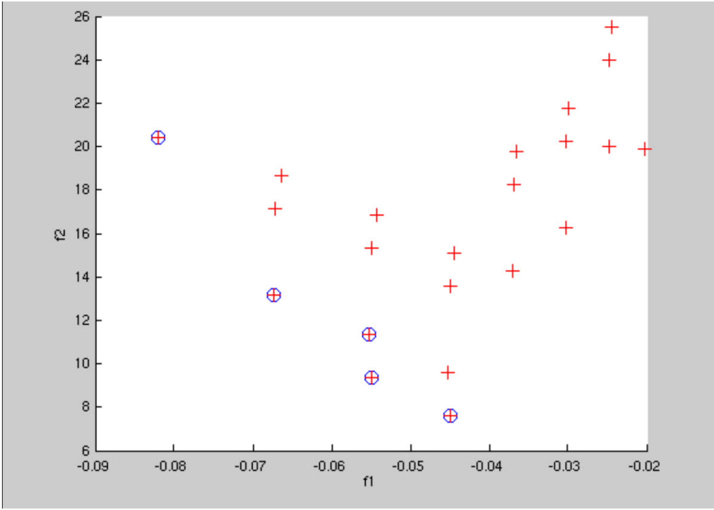

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



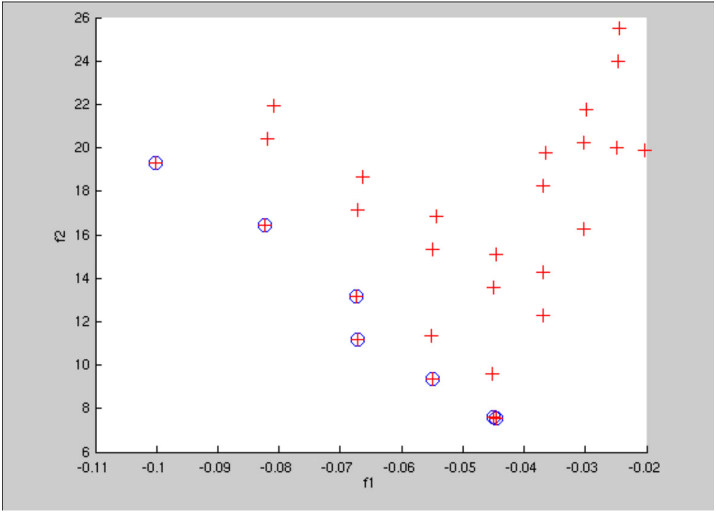

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



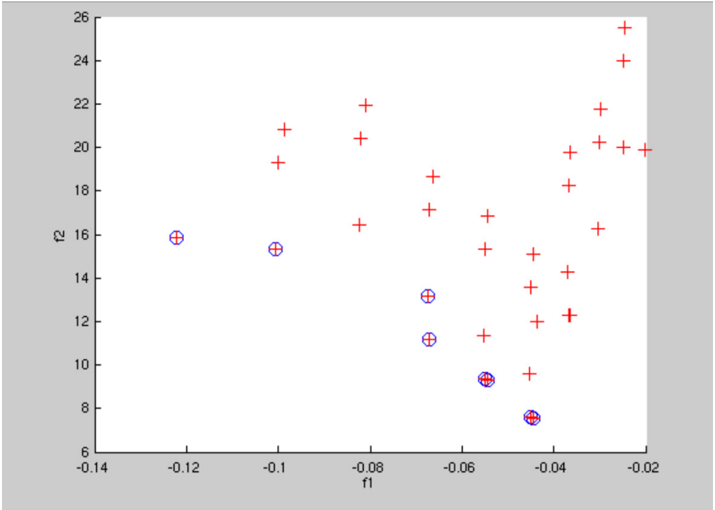

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



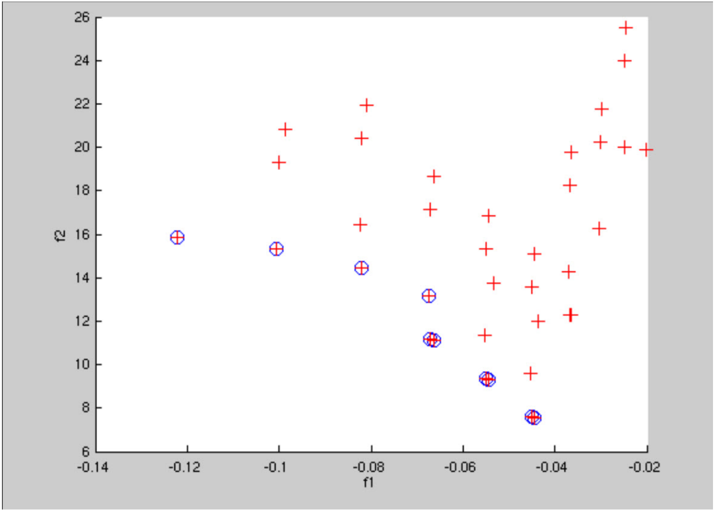

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



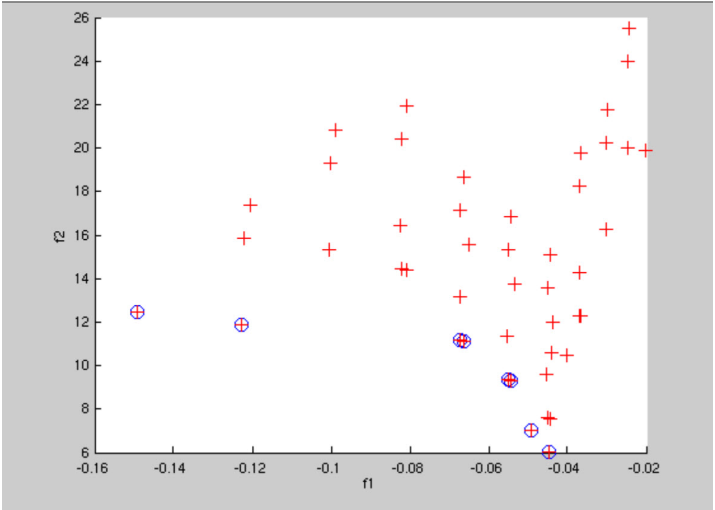

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



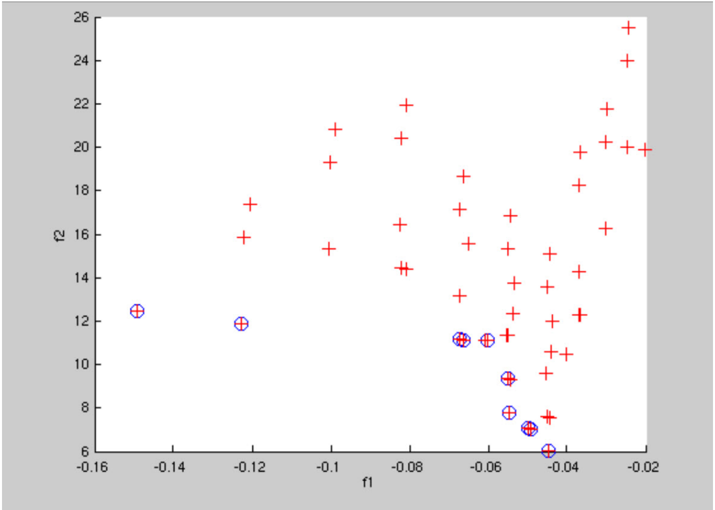

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



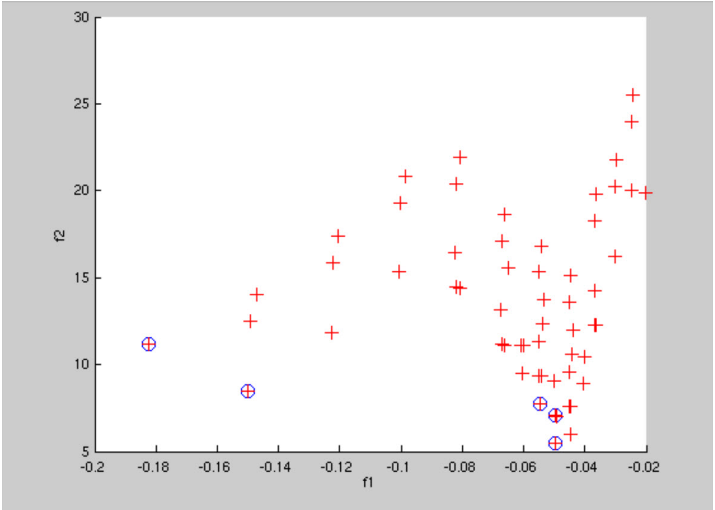

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



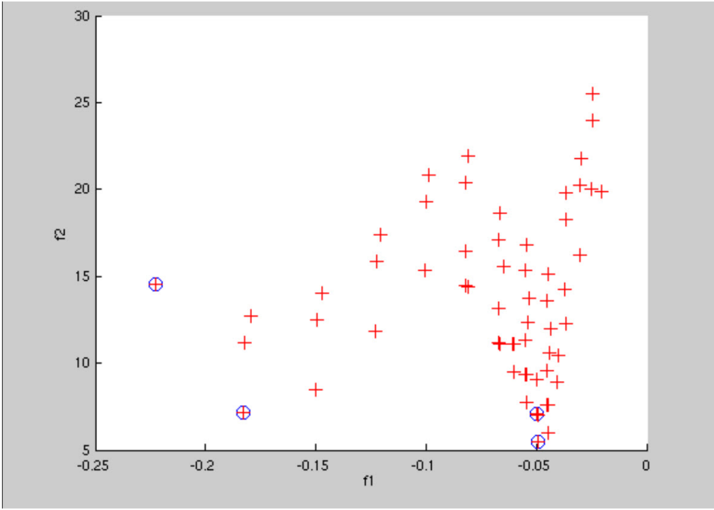

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



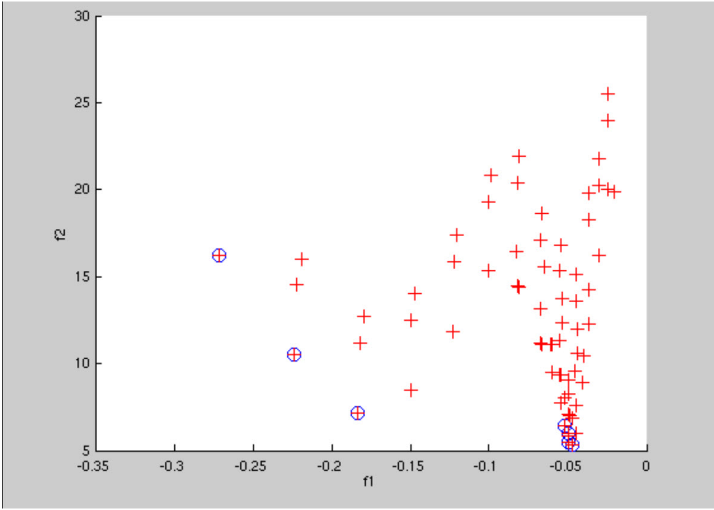

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



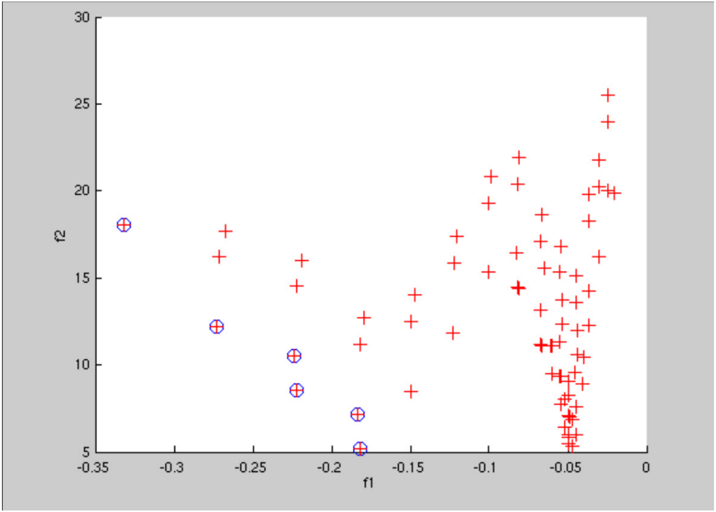

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



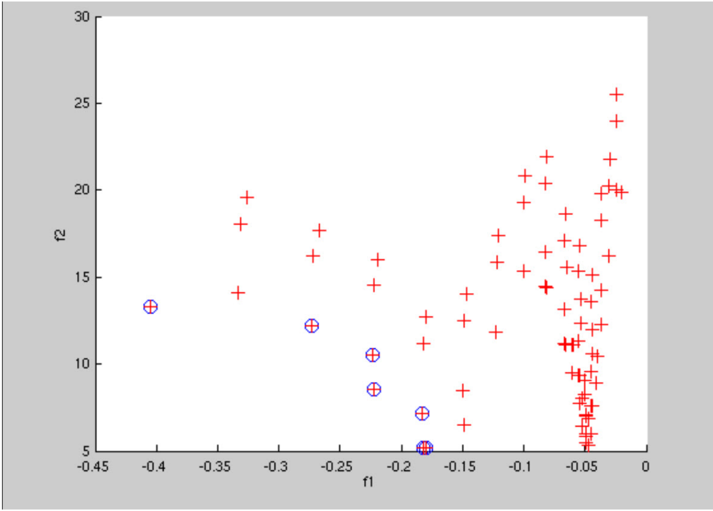

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



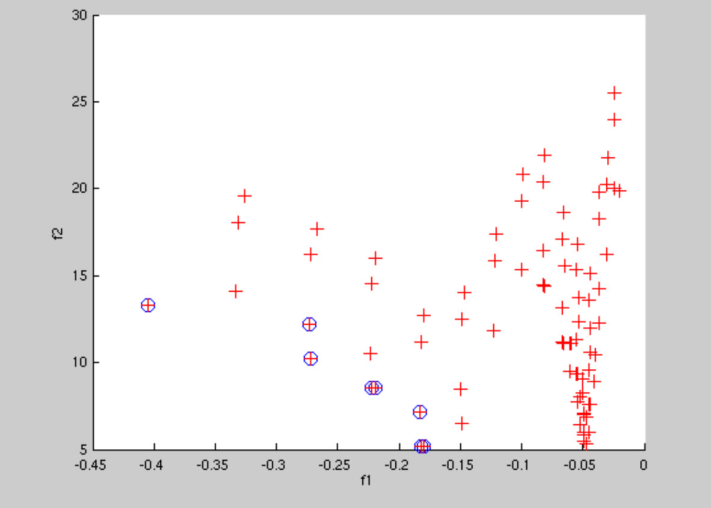

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



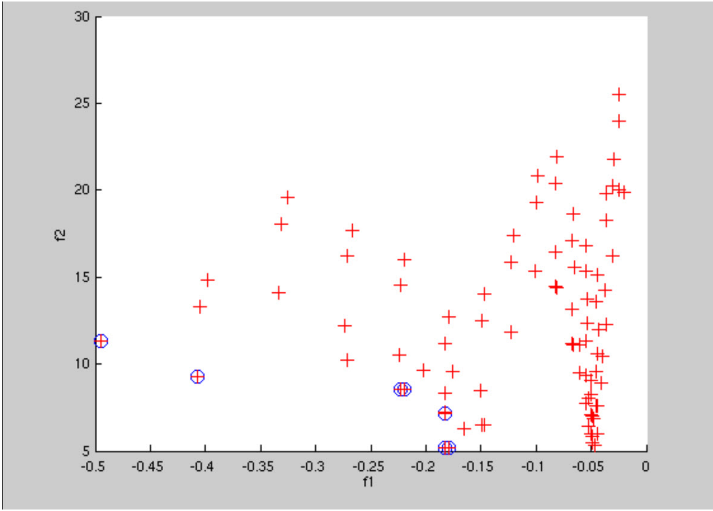

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



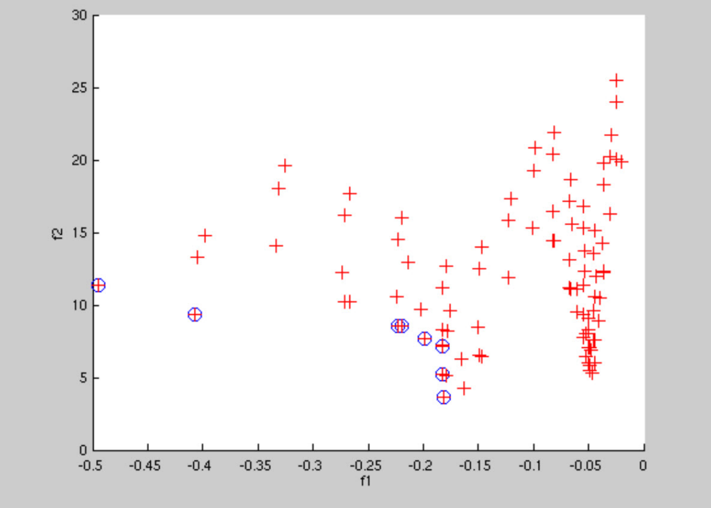

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



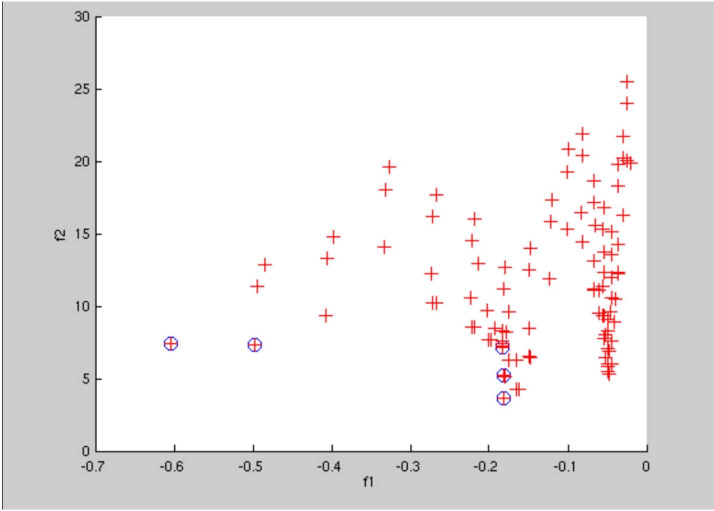

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



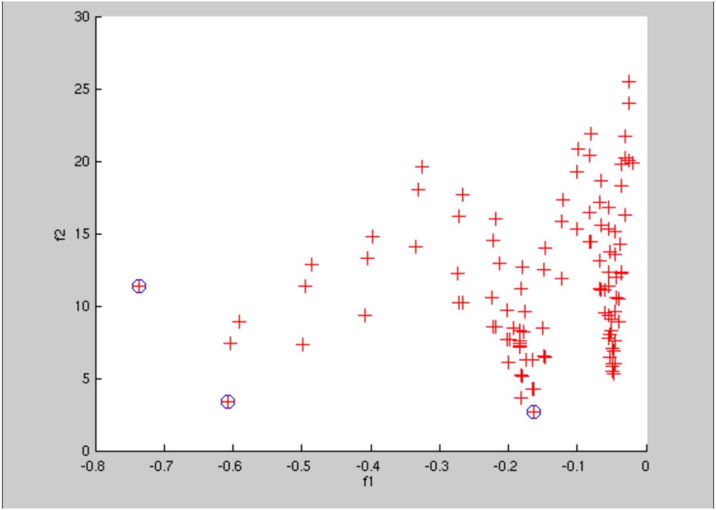

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



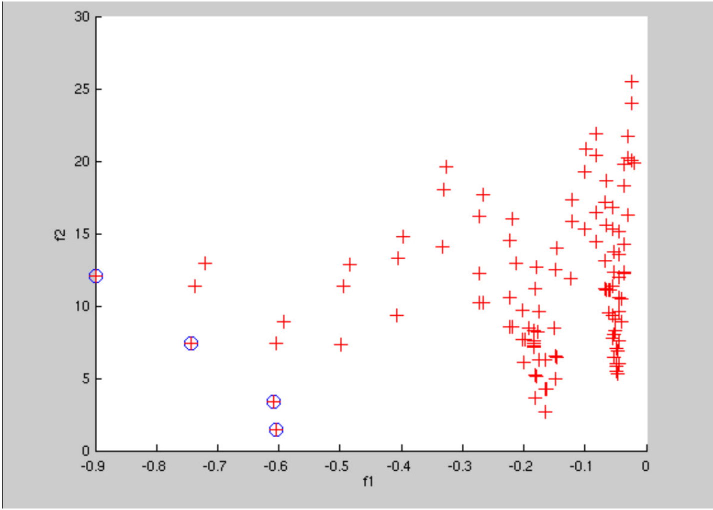

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



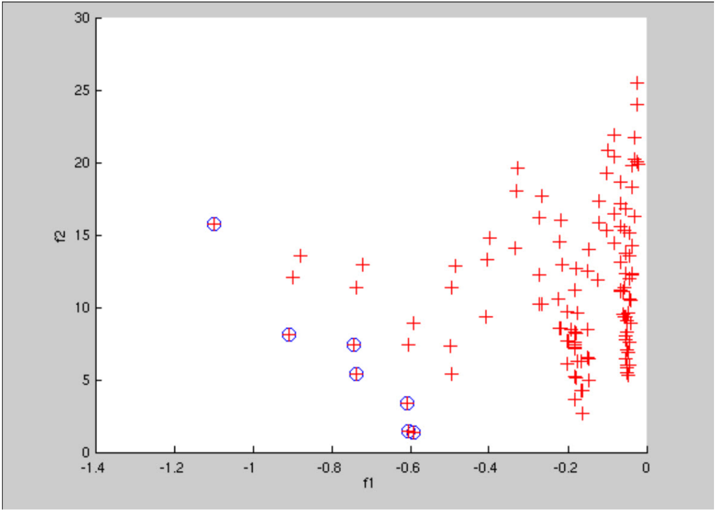

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



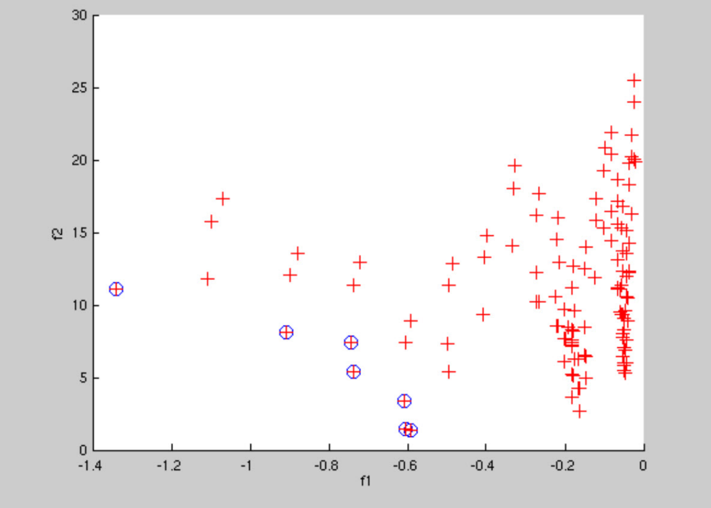

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



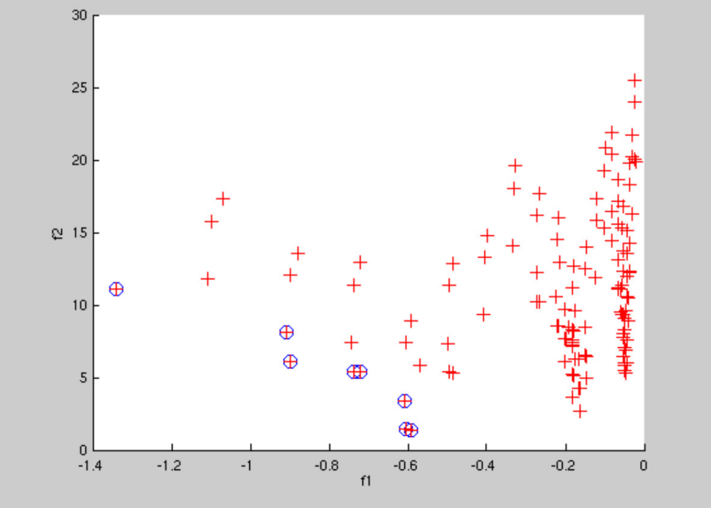

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



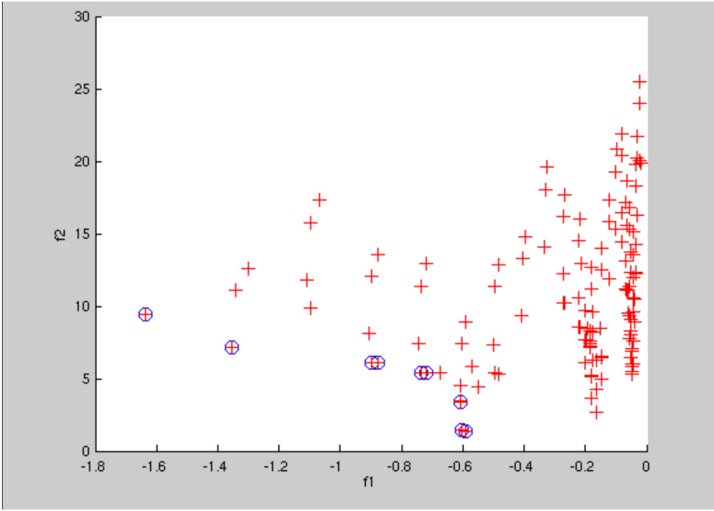

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



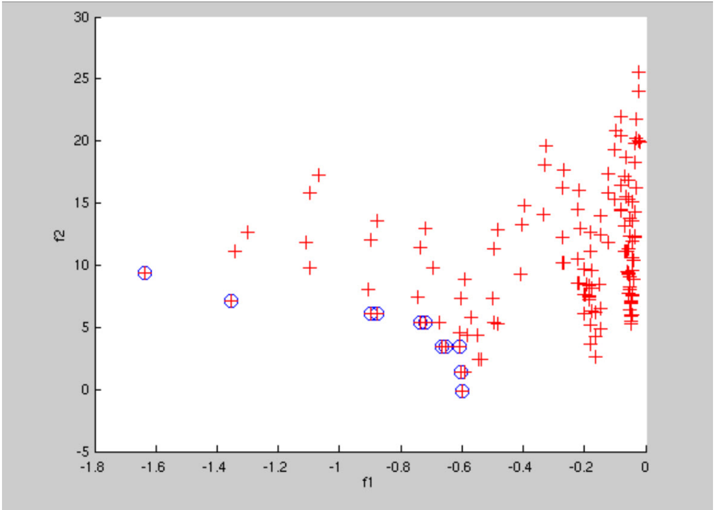

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



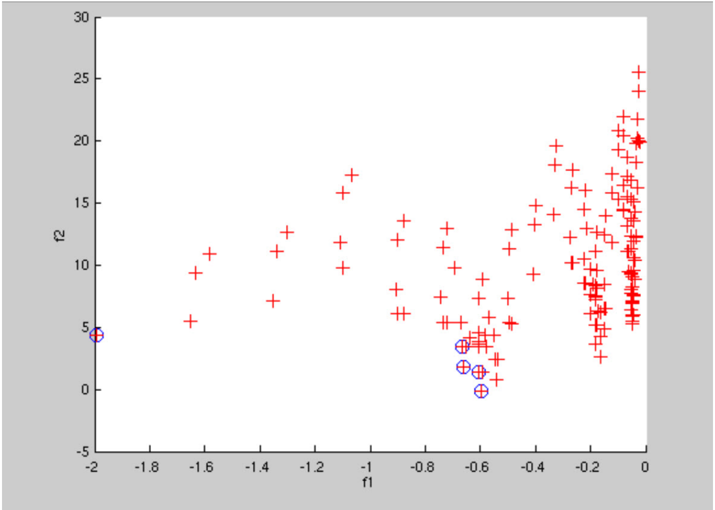

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



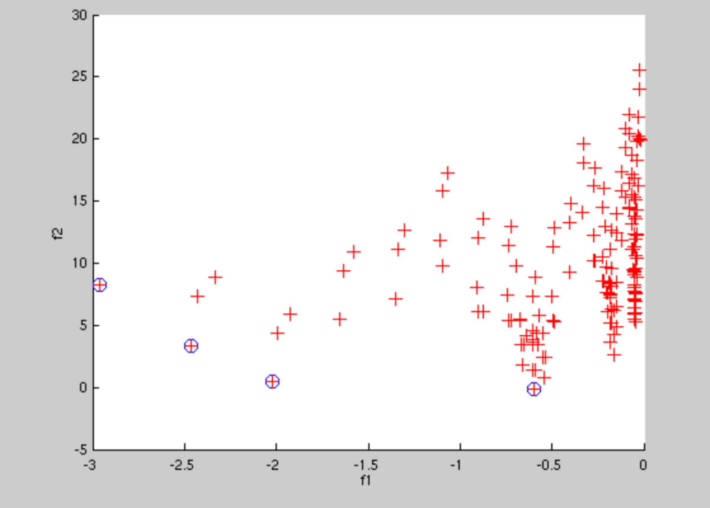

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



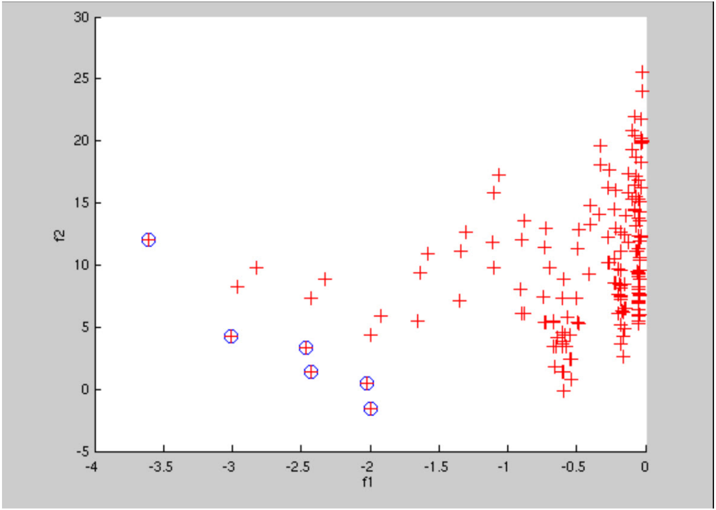

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



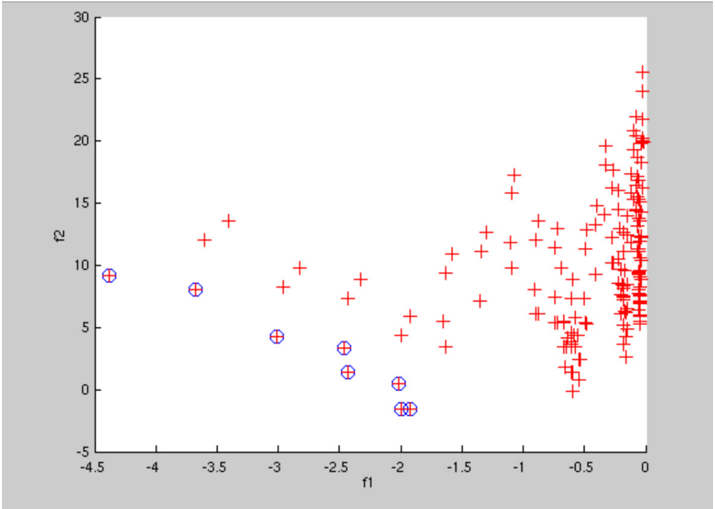

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



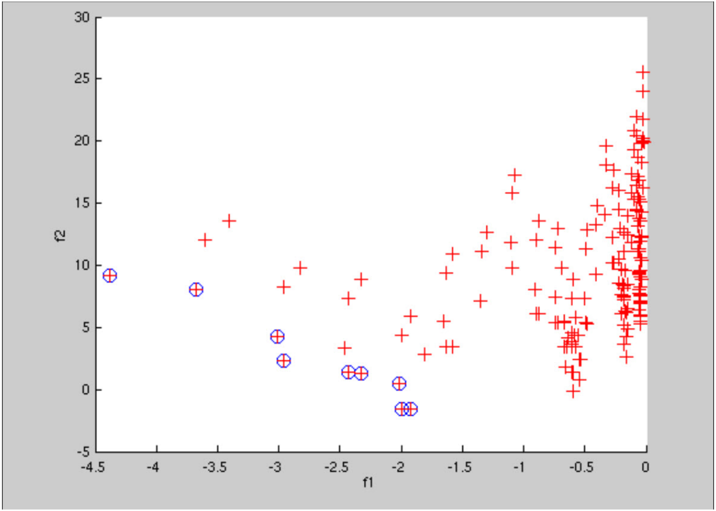

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



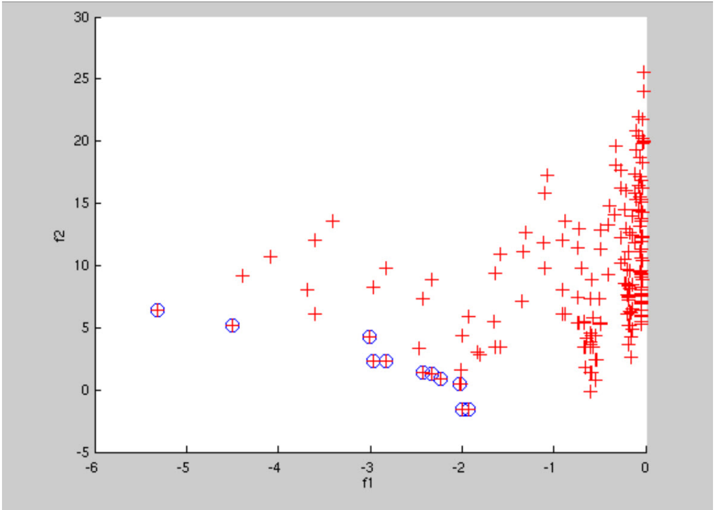

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



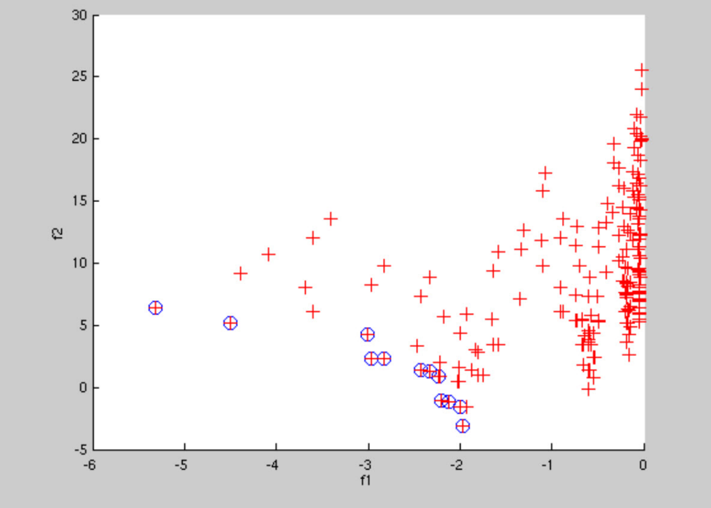

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



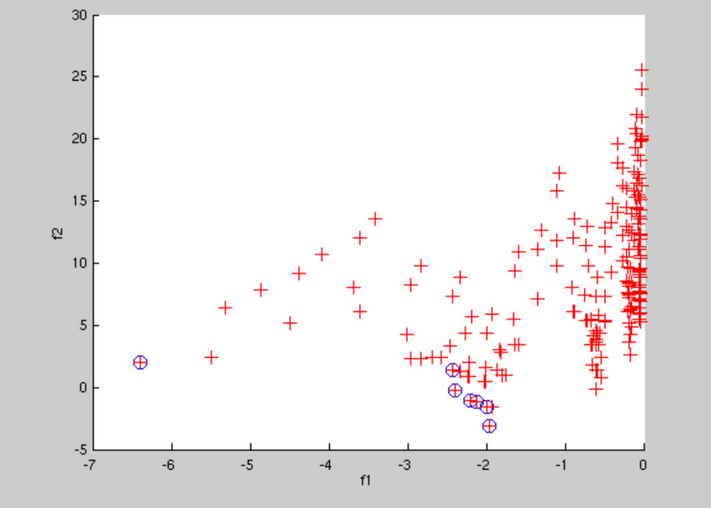

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



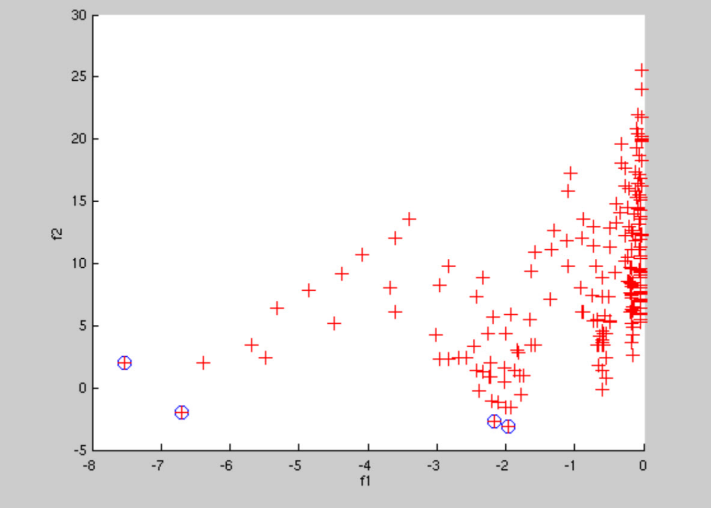

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



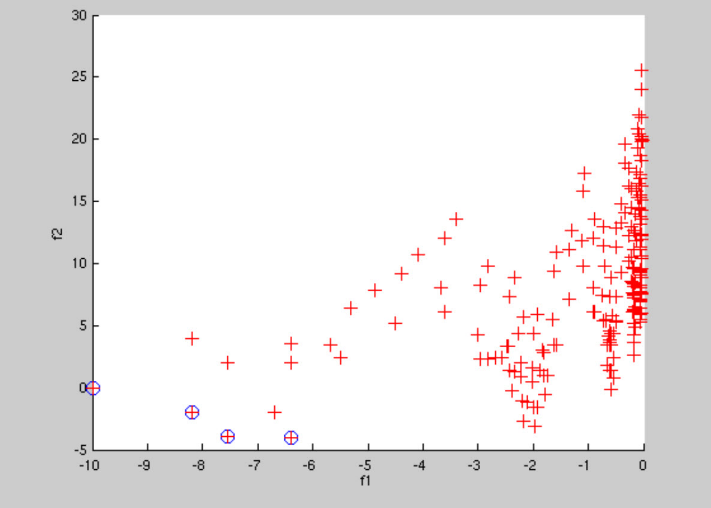

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



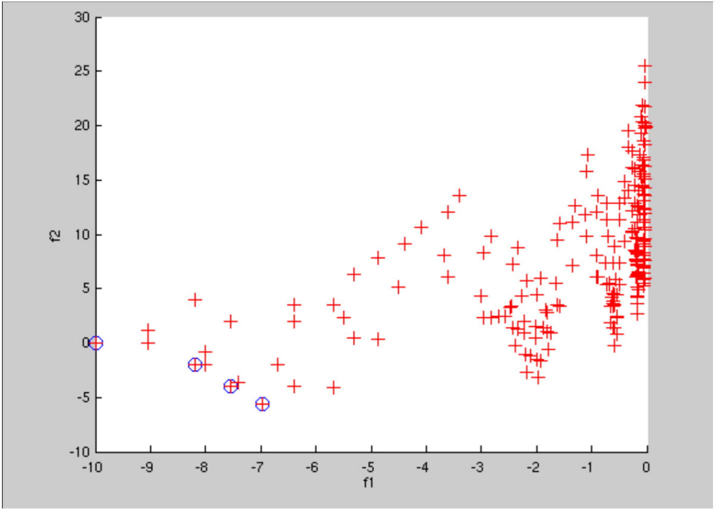

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



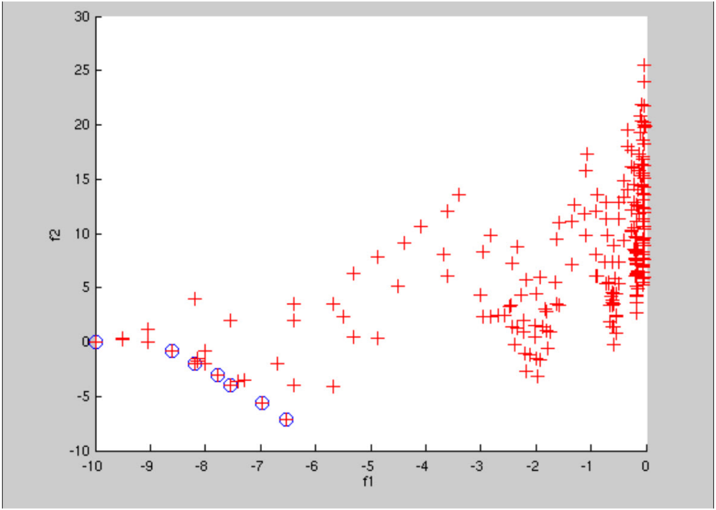

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



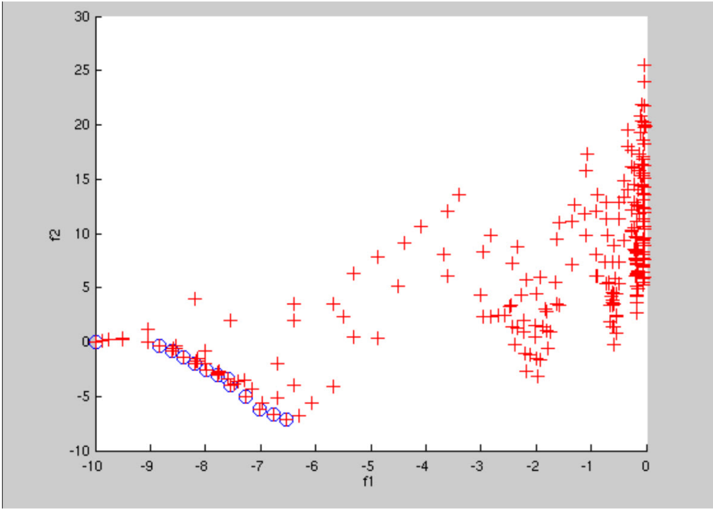

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



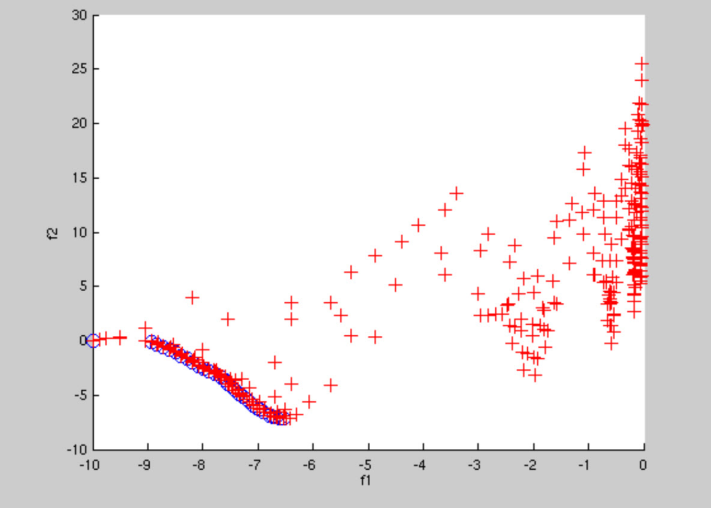

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]



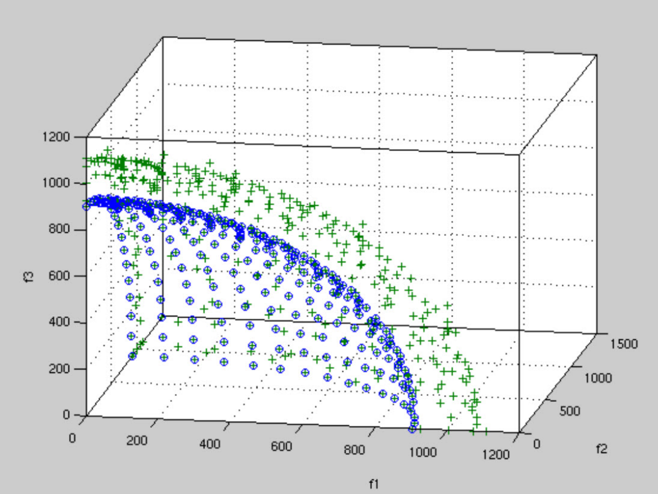

n=2,m=2

# Numerical Example - Laumanns [Laumanns et al.]




n=2,m=2

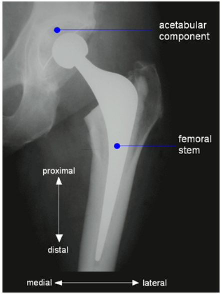
# Numerical Example - DTLZ2 [Deb et al.]

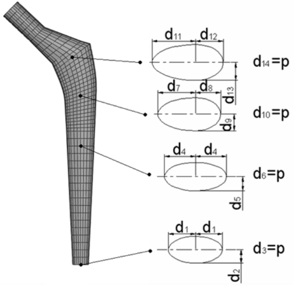



n=12,m=3

# Practical Applications

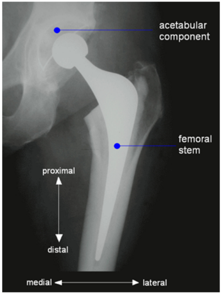
## Design of Cemented Hip Prostheses (Ruben, Fernandes, Folgado, and Madeira [2012])

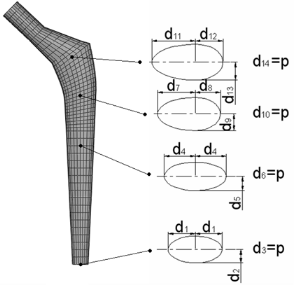



- 14 variables; 4 objectives
- 1 function evaluation $\simeq$ 2 minutes (ABAQUS)
- linear constraints

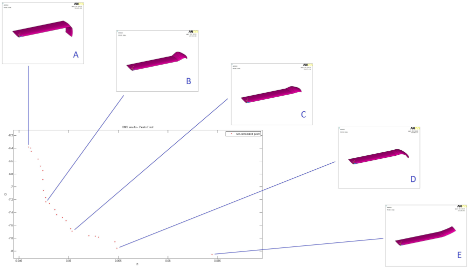
Design of Cemented Hip Prostheses (Ruben, Fernandes, Folgado, and Madeira [2012])




- 14 variables; 4 objectives
- 1 function evaluation $\simeq$ 2 minutes (ABAQUS)
- linear constraints
- more than 1000 points in the approximation to the Pareto front

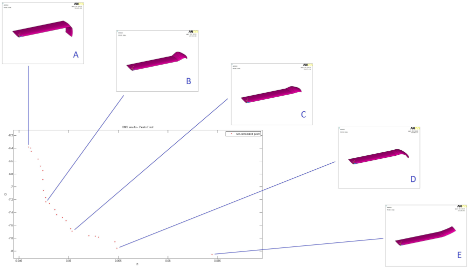
## Aerospace Design (wing shape) (Falcão, Gomes and Madeira [2012])




- 11 variables; 2 objectives
- 1 function evaluation $\simeq$ 12 minutes (ANSYS)
- bound and black-box constraints

# Practical Applications

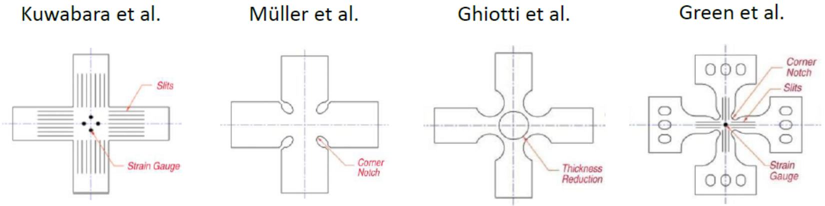
Aerospace Design (wing shape) (Falcão, Gomes and Madeira [2012])




- 11 variables; 2 objectives
- 1 function evaluation $\simeq$ 12 minutes (ANSYS)
- bound and black-box constraints
- 50 points in the approximation to the Pareto front

## An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])

- There are no standards for biaxial specimens.
- There are many geometries proposed in the literature, however none was optimized for low forces.
- Many of these geometries are not appropriated for fatigue tests (only for static tests) because stress concentrations will promote failure outside the gauge area.



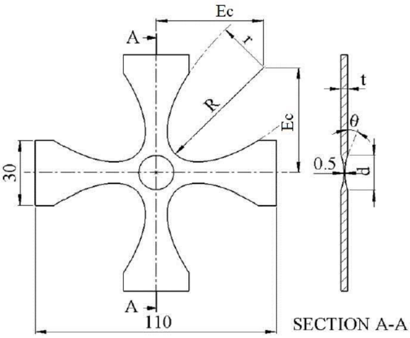
An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])

It is important to ensure that:

- Maximum stress is located in the central region of the specimen in order to avoid excessive stress that could cause failure outside of the gauge area.
- By experimental experience, the maximum allowed stress outside the gauge area should be at least $20\%$ less than in the gauge area.
- Stress at centre gauge area should be uniform in all directions to avoid fatigue crack grow deviations.

## An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])

Cruciform Studied for Crack Initiation



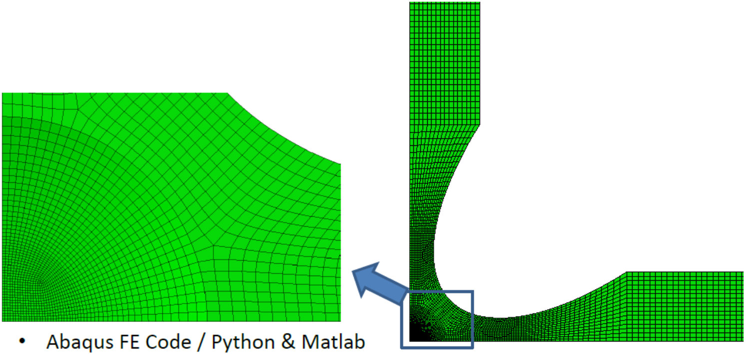

| Input variables | Min. | Max. |
|---|---|---|
| Arm Thickness, t | 2 | 5 |
| Minor Ellipse Radius, r | 15 | 30 |
| Major Ellipse Radius, R | 56 | 70 |
| Center Spline Diameter, d | 14 | 20 |
| Spline Exit Angle, θ | 30º | 90º |

| Constrains | value |
|---|---|
| Center Thickness, $t_c$ | 0,5 mm |
| Stress var. at center | 10% |
| Arms safety factor | 20% |
| Ellipse Center, $E_c$ | 49 mm |

Machining → Center thickness > =0,5 mm

## An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])

Finite Element Model



- Abaqus FE Code / Python & Matlab
- 1/8 geometry modeled (appropriated BC applied)
- Unitary force of 1 kN in each arm (worse biaxial fatigue loading case).
- 27.000 Hexahedral Elements with 20-nodes (C3D20)
  (new models 16.000 Elements (C3D8))

An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])

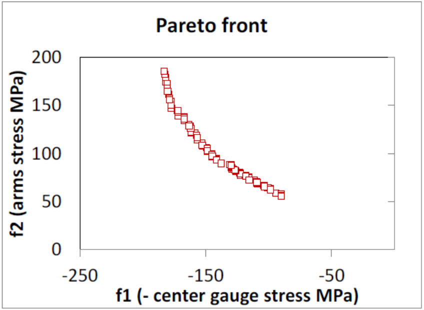
Objectives:

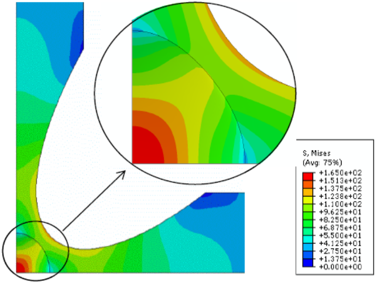
- Maximize stress in center gauge area
- Minimize stress in the arms

Constrains:

- ✓ Ensure a safety factor of 20% between center and arms stress

## An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])

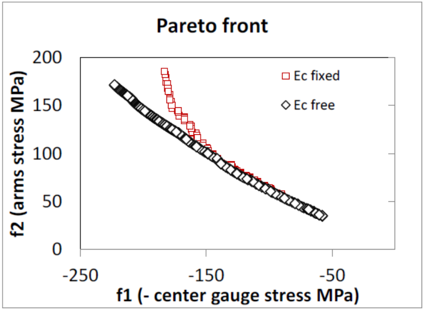


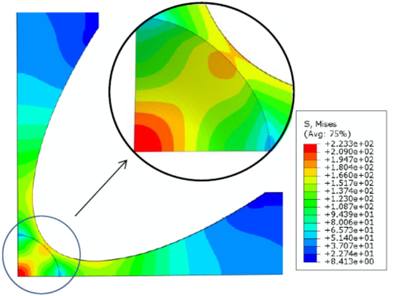

| | |
|---|---|
| Arm Thickness, t | 3 |
| Minor Ellipse Radius, r | 22 |
| Major Ellipse Radius, R | 59 |
| Center Spline Diameter, d | 15 |
| Spline Exit Angle, θ | 30º |

| Safety factor of 20% | value |
|---|---|
| Center gauge area stress | 160.3 MPa |
| Arms stress | 130.1 MPa |

## An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])




| | |
|---|---|
| Arm Thickness, t | 3.2 |
| Minor Ellipse Radius, r | 22.6 |
| Major Ellipse Radius, R | 62.4 |
| Center Spline Diameter, d | 14.4 |
| Spline Exit Angle, θ | 31.3º |
| Ellipse center | 49.8 |

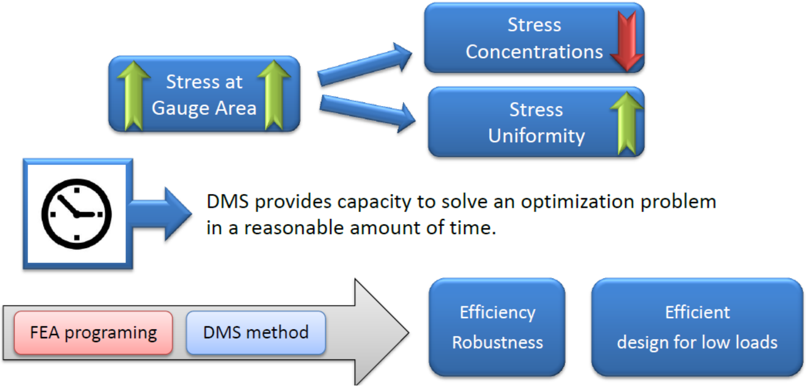



| Safety factor of 20% | value |
|---|---|
| Center gauge area stress | 220,4 MPa |
| Arms stress | 175.8 MPa |

# Practical Applications

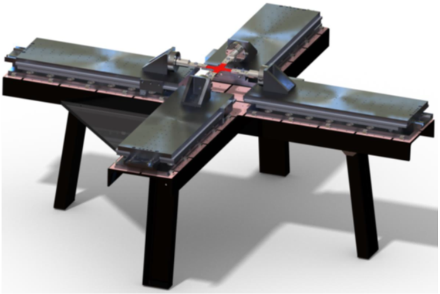
## An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])

Conclusions for this example

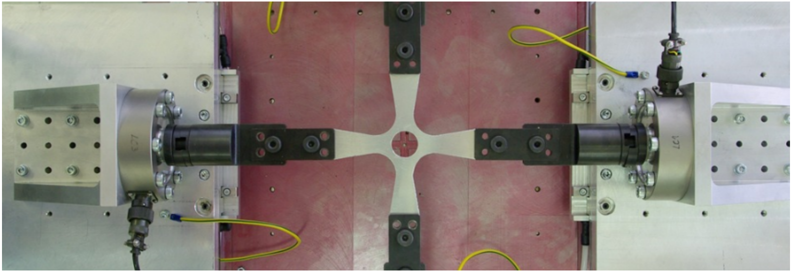


# Practical Applications

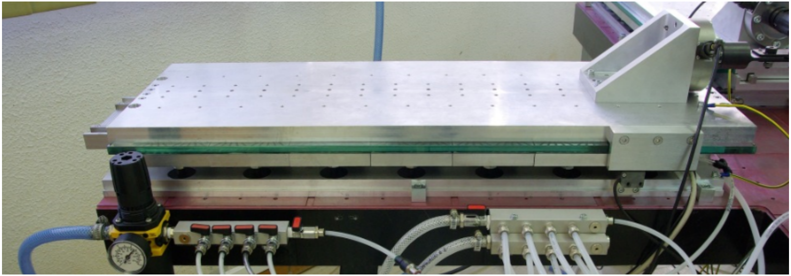
An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])



An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])



An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])



An Optimized Biaxial Cruciform Specimen for Low Capacity Testing Machines (I. Guelho, L. Reis, M. Freitas, B. Li, J. F. A. Madeira, R. A. Cláudio [2013])

# Practical Applications

Comparison of Multi-objective Algorithms Applied to Feature Selection (O. Turksen, S. Vieira, JFA Madeira, A. Apaydin, JMC Sousa ),Towards Advanced Data Analysis by Combining Soft Computing and Statistics , Springer Berlin, 2013, DOI: 10.1007/978-3-642-30278-7

- Real world data sets tend to be complex, very large, and normally contain many **irrelevant features** (variables, inputs, attributes).
- One of the most important steps in data analysis for **classification problems** is **feature selection**.
- Feature selection has been an active research area on many fields, such as data mining, pattern recognition, image understanding, machine learning and statistics.
- The main idea of feature selection is to **choose a subset of available features, by eliminating redundant features with little or no predictive information**.

# Practical Applications

Comparison of Multi-objective Algorithms Applied to Feature Selection (O. Turksen, S. Vieira, JFA Madeira, A. Apaydin, JMC Sousa ),Towards Advanced Data Analysis by Combining Soft Computing and Statistics , Springer Berlin, 2013, DOI: 10.1007/978-3-642-30278-7

- **Two key decisions** for feature subset selection:
  - The number of selected features
  - The best features to be selected
- An **effective feature selection method** can
  - minimize the classification error
  - improve the prediction accuracy
  - discover the relevant features

**Feature subset selection problem** ⟺ **Multi-objective optimization (MOO) problem**

↘ ↙

**Multi-objective feature selection problem (MOFS)**

Comparison of Multi-objective Algorithms Applied to Feature Selection (O. Turksen, S. Vieira, JFA Madeira, A. Apaydin, JMC Sousa ),Towards Advanced Data Analysis by Combining Soft Computing and Statistics , Springer Berlin, 2013, DOI: 10.1007/978-3-642-30278-7

- **Objectives:**
  - minimizing the number of features
  - minimizing the misclassification rate
- **Classifier:**
  - Takagi-Sugeno fuzzy modeling
- **MOO methods:**
  - NSGA II
  - Modified AMOSA
  - DMS
- **Data sets:**
  - Wisconsin Breast Cancer Original (WBCO)
  - Wisconsin Diagnostic Breast Cancer (WDBC)
  - Wisconsin Prognostic Breast Cancer (WPBC)
  - Sonar

# Practical Applications

Comparison of Multi-objective Algorithms Applied to Feature Selection (O. Turksen, S. Vieira, JFA Madeira, A. Apaydin, JMC Sousa ),Towards Advanced Data Analysis by Combining Soft Computing and Statistics , Springer Berlin, 2013, DOI: 10.1007/978-3-642-30278-7

**MOFS problem**

$$\min f_1(\boldsymbol{x}) = \sum_{k=1}^{N} x_k$$

$$\min f_2(\boldsymbol{x}) = (1 - accuracy(\boldsymbol{x}))$$

$$\boldsymbol{x} \in \{0,1\}^N$$

# Practical Applications

Comparison of Multi-objective Algorithms Applied to Feature Selection (O. Turksen, S. Vieira, JFA Madeira, A. Apaydin, JMC Sousa ),Towards Advanced Data Analysis by Combining Soft Computing and Statistics , Springer Berlin, 2013, DOI: 10.1007/978-3-642-30278-7

- Four benchmark data sets, selected from the UCI Machine Repository (Asuncion and Newman, 2007), are used.

**Table 1.** Description of the used data sets.

| No | datasets used | # features | # classes | # samples |
|---|---|---|---|---|
| 1 | WBCO | 9 (integer) | 2 | 699 |
| 2 | WDBC | 32 (real) | 2 | 569 |
| 3 | WPBC | 34 (real) | 2 | 198 |
| 4 | SONAR | 60 (real & integer) | 2 | 208 |

Comparison of Multi-objective Algorithms Applied to Feature Selection (O. Turksen, S. Vieira, JFA Madeira, A. Apaydin, JMC Sousa ),Towards Advanced Data Analysis by Combining Soft Computing and Statistics , Springer Berlin, 2013, DOI: 10.1007/978-3-642-30278-7

**Table 5.** Feature subsets for WBCO data set with 10 fold cross-validation

| NF | Selected features | | | Value of 1-accuracy (%) |
|---|---|---|---|---|
| | NSGAII | Modified AMOSA | DMS | |
| 1 | {2} | {2} | {2} | 7.153 |
| 2 | {2, 6} | {2, 6} | {2, 6} | 4.721 |
| 3 | {1, 2, 6} | {1, 2, 6} | - | 4.435 |
| 3 | - | - | {1, 3, 6} | 3.720 |
| 4 | - | - | {1, 3, 4, 6} | 3.577 |
| 5 | - | - | {1,2,3,5,6} | 3.434 |

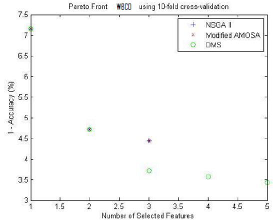



**Fig.3 Pareto front for WBCO data set**

Comparison of Multi-objective Algorithms Applied to Feature Selection (O. Turksen, S. Vieira, JFA Madeira, A. Apaydin, JMC Sousa ),Towards Advanced Data Analysis by Combining Soft Computing and Statistics , Springer Berlin, 2013, DOI: 10.1007/978-3-642-30278-7

**Table 6.** Feature subsets for WDBC data set with 10 fold cross validation

| NF | Selected features NSGAII | Modified AMOSA | DMS | Value of 1-accuracy (%) |
|---|---|---|---|---|
| 1 | {24} | - | {24} | 8.260 |
| 2 | {24,28} | - | - | 6.151 |
| 2 | - | - | {21,25} | 4.745 |
| 3 | {22,24,28} | - | - | 4.569 |
| 3 | - | - | {2,21,25} | 4.042 |
| 3 | - | - | {2,24,28} | 4.042 |
| 3 | - | - | {22,24,29} | 4.042 |
| 3 | - | - | {24,25,29} | 4.042 |
| 4 | {8,22,24, 25} | - | - | 4.394 |
| 4 | - | {21,26,29,30} | - | 6.503 |
| 4 | - | - | {2,3, 24,25} | 3.339 |
| 4 | - | - | {2,21,28,29} | 3.339 |
| 5 | {8,10,22, 24,25} | - | - | 3.866 |
| 5 | - | {2,3,8,21,29} | - | 5.800 |
| 5 | - | - | {2,14,21,25,28} | 2.812 |
| 5 | - | - | {14,21,22,25,28} | 2.812 |
| 6 | - | - | {2,14,21,25,28, 29} | 2.636 |
| 6 | - | - | {2,14,24,25,28, 29} | 2.636 |

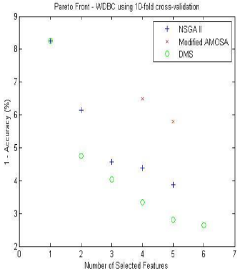



**Fig.4 Pareto front for WDBC data set**

Comparison of Multi-objective Algorithms Applied to Feature Selection (O. Turksen, S. Vieira, JFA Madeira, A. Apaydin, JMC Sousa ),Towards Advanced Data Analysis by Combining Soft Computing and Statistics , Springer Berlin, 2013, DOI: 10.1007/978-3-642-30278-7

**Table 7.** Feature subsets for WPBC data set with 10-fold cross validation.

| NF | Selected features: NSGAII | Selected features: Modified AMOSA | Selected features: DMS | Value of 1-accuracy (%) |
|---|---|---|---|---|
| 1 | {25} | - | - | 23.74 |
| 1 | - | - | {5} | 22.73 |
| 2 | {1,25} | - | {1,25} | 19.70 |
| 2 | - | - | {1,22} | 19.70 |
| 3 | - | {11,23,24} | - | 21.72 |
| 3 | - | - | {1,13,25} | 18.18 |
| 4 | - | {1,3,7,22} | - | 18.69 |
| 4 | - | - | {1,13,22,32} | 17.17 |
| 5 | {1,6,8,13,25} | - | - | 19.19 |
| 5 | - | - | {1,13,20,22,32} | 16.67 |
| 5 | - | - | {1,13,24,26,32} | 16.67 |
| 6 | {1,6,8,13,19,25} | - | | 18.18 |
| 6 | - | - | {1,6,11,13,18,32} | 16.16 |
| 6 | - | - | {1,13,22,26,27,32} | 16.16 |
| 7 | - | - | {1,13,20,22,26,27,32} | 15.66 |
| 10 | - | {1,2,5,8,13,14,15, 17, 22,24} | - | 18.18 |
| 12 | - | - | {1,2,6,11,12,13,14,17, 18,22,24,32} | 14.65 |
| 14 | - | - | {1,2,7,9,12,13,14, 17,18,20,22,24,26,29} | 13.64 |
| 20 | - | - | {1,2,5,9,12,13,14, 16,17,18,20,21,22,24, 25,26,27,28,29,31} | 13.13 |

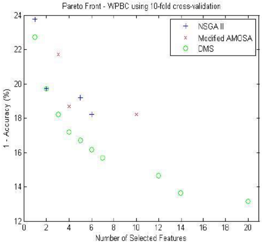



**Fig.5 Pareto front for WPBC data set**

Comparison of Multi-objective Algorithms Applied to Feature Selection (O. Turksen, S. Vieira, JFA Madeira, A. Apaydin, JMC Sousa ),Towards Advanced Data Analysis by Combining Soft Computing and Statistics , Springer Berlin, 2013, DOI: 10.1007/978-3-642-30278-7

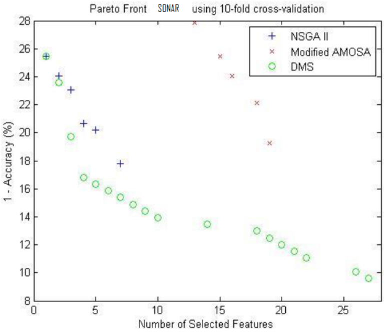



**Fig.6 Pareto front for Sonar data set**

# Conclusions and references

- Development and analysis of a novel approach (Direct MultiSearch) for MOO, generalizing ALL direct-search methods.

# Conclusions and references

- Development and analysis of a novel approach (Direct MultiSearch) for MOO, generalizing ALL direct-search methods.

- Direct MultiSearch (DMS) exhibits highly competitive numerical results for MOO.

# Conclusions and references

- Development and analysis of a novel approach (Direct MultiSearch) for MOO, generalizing ALL direct-search methods.

- Direct MultiSearch (DMS) exhibits highly competitive numerical results for MOO.

# Conclusions and references

- Development and analysis of a novel approach (Direct MultiSearch) for MOO, generalizing ALL direct-search methods.

- Direct MultiSearch (DMS) exhibits highly competitive numerical results for MOO.

A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz, and L. N. Vicente, Direct multisearch for multiobjective optimization, SIAM Journal on Optimization, 21 (2011) 1109-1140.

# Presentation Outline

# Globalization Strategies

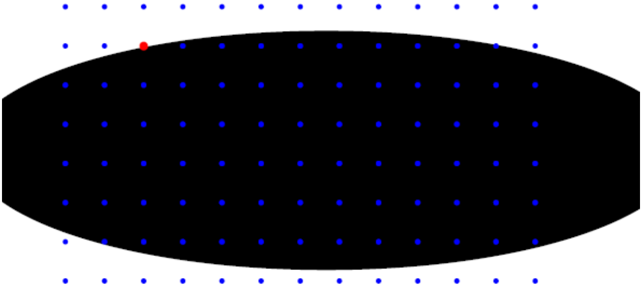
(1) Using Integer Lattices (Torczon [1997], Audet and Dennis [2002])

- requires only simple decrease ($x$ is accepted instead of the current iterate $x_k$ if $f(x) < f(x_k)$)
- poll directions and step size must satisfy integer/rational requirements
- search step is restricted to an implicit mesh

# Globalization Strategies

(1) Using Integer Lattices (Torczon [1997], Audet and Dennis [2002])

- requires only simple decrease ($x$ is accepted instead of the current iterate $x_k$ if $f(x) < f(x_k)$)
- poll directions and step size must satisfy integer/rational requirements
- search step is restricted to an implicit mesh



(2) Imposing Sufficient Decrease (Kolda, Lewis, and Torczon [2003])

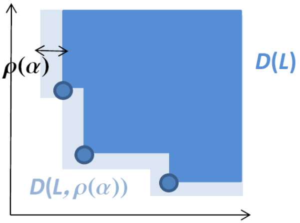
- use of a forcing function $(f(x) < f(x_k) - \rho(\alpha_k))$

  $\rho : (0, +\infty) \to (0, +\infty)$, continuous and nondecreasing, satisfying

  $$\rho(t)/t \to 0 \text{ when } t \downarrow 0$$

- directions can be randomly generated
- $x$ is nondominated $\Leftrightarrow x \notin D(L, \rho(\alpha))$

$$D(L) \subset D(L, \rho(\alpha))$$



# Refining Subsequences and Directions

For both globalization strategies:

**Theorem (Refining Subsequences)**

*There is at least a convergent subsequence of iterates* $\{x_k\}_{k \in K}$, *corresponding to unsuccessful poll steps, such that* $\lim_{k \in K} \alpha_k = 0$.

| | |
|---|---|
| DMS (integer lattices, sufficient decrease): | Custódio, Madeira, Vaz, and Vicente [2011] |
| DS (integer lattices): | Torczon [1997], Audet and Dennis [2002] |
| DS (sufficient decrease): | Kolda, Lewis, and Torczon [2003] |

# Refining Subsequences and Directions

For both globalization strategies:

**Theorem (Refining Subsequences)**

*There is at least a convergent subsequence of iterates* $\{x_k\}_{k \in K}$*, corresponding to unsuccessful poll steps, such that* $\lim_{k \in K} \alpha_k = 0$.

| | |
|---|---|
| DMS (integer lattices, sufficient decrease): | Custódio, Madeira, Vaz, and Vicente [2011] |
| DS (integer lattices): | Torczon [1997], Audet and Dennis [2002] |
| DS (sufficient decrease): | Kolda, Lewis, and Torczon [2003] |

Let $x_*$ be the limit point of a convergent refining subsequence $\{x_k\}_{k \in K}$.

**Definition (Refining Directions)**

*Refining directions for* $x_*$ *are limit points of* $\{d_k / \|d_k\|\}_{k \in K}$*, where* $d_k \in D_k$ *and* $x_k + \alpha_k d_k \in \Omega$.

Audet and Dennis [2006]

Clarke Generalized Directional Derivative

For $F$ Lipschitz continuous near $x_*$ and $d \in \mathbb{R}^n$:

$$f_j^\circ(x_*; d) \;=\; \limsup_{x' \to x_* \; t \downarrow 0} \frac{f_j(x' + td) - f_j(x')}{t}$$

# Pareto-Clarke Stationarity (Unconstrained Case $\Omega = \mathbb{R}^n$)

Clarke Generalized Directional Derivative

For $F$ Lipschitz continuous near $x_*$ and $d \in \mathbb{R}^n$:

$$f_j^\circ(x_*; d) \;=\; \limsup_{x' \to x_* \; t \downarrow 0} \frac{f_j(x' + td) - f_j(x')}{t}$$

**Definition**

$x_*$ *is a Pareto-Clarke critical point of* $F$*:*

$$\forall d \in \mathbb{R}^n, \exists j = j(d) \in \{1, \ldots, m\}, f_j^\circ(x_*; d) \geq 0$$

# Pareto-Clarke Stationarity (Constrained Case $\Omega \subset \mathbb{R}^n$)

- In the algorithm use the extreme barrier function:

$$F_\Omega(x) = \begin{cases} F(x) & \text{if } x \in \Omega, \\ (+\infty, \ldots, +\infty)^\top & \text{otherwise.} \end{cases}$$

# Pareto-Clarke Stationarity (Constrained Case $\Omega \subset \mathbb{R}^n$)

- In the algorithm use the extreme barrier function:

$$F_\Omega(x) = \begin{cases} F(x) & \text{if } x \in \Omega, \\ (+\infty, \ldots, +\infty)^\top & \text{otherwise.} \end{cases}$$

Again, assume $F$ is Lipschitz continuous near $x_*$.

**Definition**

*$x_*$ is a Pareto-Clarke critical point of $F$:*

$$\forall d \in T_\Omega^{Cl}(x_*), \exists j = j(d) \in \{1, \ldots, m\}, f_j^\circ(x_*; d) \geq 0$$

$T_\Omega^{Cl}(x_*)$ is the tangent cone to $\Omega$ at $x_*$ (redefined in the nonsmooth, Clarke way)

# Pareto-Clarke Stationarity (Constrained Case $\Omega \subset \mathbb{R}^n$)

- In the algorithm use the extreme barrier function:

$$F_\Omega(x) = \begin{cases} F(x) & \text{if } x \in \Omega, \\ (+\infty, \ldots, +\infty)^\top & \text{otherwise.} \end{cases}$$

Again, assume $F$ is Lipschitz continuous near $x_*$.

**Definition**

*$x_*$ is a Pareto-Clarke critical point of $F$:*

$$\forall d \in T_\Omega^{Cl}(x_*), \exists j = j(d) \in \{1, \ldots, m\}, f_j^\circ(x_*; d) \geq 0$$

$T_\Omega^{Cl}(x_*)$ is the tangent cone to $\Omega$ at $x_*$ (redefined in the nonsmooth, Clarke way)

Moreover, the Clarke derivatives must be appropriately redefined...

# Pareto-Clarke Stationarity (Constrained Case $\Omega \subset \mathbb{R}^n$)

Clarke-Jahn Generalized Directional Derivative

$$f_j^\circ(x_*; v) \; = \; \limsup_{\substack{x' \to x_*,\, x' \in \Omega \\ t \downarrow 0,\, x' + tv \in \Omega}} \frac{f_j(x' + tv) - f_j(x')}{t},$$

for $v \in \mathrm{int}(T_\Omega^{Cl}(x_*))$,

# Pareto-Clarke Stationarity (Constrained Case $\Omega \subset \mathbb{R}^n$)

Clarke-Jahn Generalized Directional Derivative

$$f_j^\circ(x_*; v) = \limsup_{\substack{x' \to x_*,\, x' \in \Omega \\ t \downarrow 0,\, x' + tv \in \Omega}} \frac{f_j(x' + tv) - f_j(x')}{t},$$

for $v \in \text{int}(T_\Omega^{Cl}(x_*))$,

and then (Audet and Dennis [2006]), for $d \in T_\Omega^{Cl}(x_*)$

$$f_j^\circ(x_*; d) = \lim_{v \in \text{int}(T_\Omega^{Cl}(x_*)),\, v \to d} f_j^\circ(x_*; v),$$

# Convergence Results

Consider a refining subsequence converging to $x_*$ (and assume that $F$ is Lipschitz continuous near $x_*$)

**Theorem**

*If $d \in int(T_\Omega^{Cl}(x_*))$ is a refining direction for $x_*$ then:*

$$\exists j = j(d) \in \{1, \ldots, m\} : f_j^\circ(x_*; d) \geq 0$$

DMS: Custódio, Madeira, Vaz, and Vicente [2011]

DS: Audet and Dennis [2006], Vicente and Custódio [2010]

# Proof Sketch

$$f_j^\circ(x_*;d) \quad = \quad \limsup_{\substack{x' \to x_*,\, x' \in \Omega \\ t \downarrow 0,\, x' + td \in \Omega}} \frac{f_j(x'+td) - f_j(x')}{t}$$

# Proof Sketch

$$
\begin{aligned}
f_j^\circ(x_*;d) &= \limsup_{\substack{x' \to x_*,\, x' \in \Omega \\ t \downarrow 0,\, x' + td \in \Omega}} \frac{f_j(x'+td) - f_j(x')}{t} \\
&\geq \limsup_{k \in K} \frac{f_j(x_k + \alpha_k \|d_k\| (d_k/\|d_k\|)) - f_j(x_k)}{\alpha_k \|d_k\|} - r_k
\end{aligned}
$$

# Proof Sketch

$$f_j^\circ(x_*; d) = \limsup_{\substack{x' \to x_*, x' \in \Omega \\ t \downarrow 0, x' + td \in \Omega}} \frac{f_j(x' + td) - f_j(x')}{t}$$

$$\geq \limsup_{k \in K} \frac{f_j(x_k + \alpha_k \|d_k\| (d_k / \|d_k\|)) - f_j(x_k)}{\alpha_k \|d_k\|} - r_k$$

$$= \limsup_{k \in K} \frac{f_j(x_k + \alpha_k d_k) - f_j(x_k) + \rho(\alpha_k \|d_k\|)}{\alpha_k \|d_k\|} - \frac{\rho(\alpha_k \|d_k\|)}{\alpha_k \|d_k\|} - r_k$$

# Proof Sketch

$$f_j^\circ(x_*; d) = \limsup_{\substack{x' \to x_*, x' \in \Omega \\ t \downarrow 0, x' + td \in \Omega}} \frac{f_j(x' + td) - f_j(x')}{t}$$

$$\geq \limsup_{k \in K} \frac{f_j(x_k + \alpha_k \|d_k\| (d_k / \|d_k\|)) - f_j(x_k)}{\alpha_k \|d_k\|} - r_k$$

$$= \limsup_{k \in K} \frac{f_j(x_k + \alpha_k d_k) - f_j(x_k) + \rho(\alpha_k \|d_k\|)}{\alpha_k \|d_k\|} - \frac{\rho(\alpha_k \|d_k\|)}{\alpha_k \|d_k\|} - r_k$$

Since $\{x_k\}_{k \in K}$ is a refining subsequence, for each $k \in K$, $x_k + \alpha_k d_k$ does not dominate $x_k$

# Proof Sketch

$$
\begin{aligned}
f_j^\circ(x_*; d) \quad &= \limsup_{\substack{x' \to x_*, x' \in \Omega \\ t \downarrow 0, x' + td \in \Omega}} \frac{f_j(x' + td) - f_j(x')}{t} \\
&\geq \limsup_{k \in K} \frac{f_j(x_k + \alpha_k \|d_k\| (d_k / \|d_k\|)) - f_j(x_k)}{\alpha_k \|d_k\|} - r_k
\end{aligned}
$$

$$
= \limsup_{k \in K} \frac{f_j(x_k + \alpha_k d_k) - f_j(x_k) + \rho(\alpha_k \|d_k\|)}{\alpha_k \|d_k\|} - \frac{\rho(\alpha_k \|d_k\|)}{\alpha_k \|d_k\|} - r_k
$$

Since $\{x_k\}_{k \in K}$ is a refining subsequence, for each $k \in K$, $x_k + \alpha_k d_k$ does not dominate $x_k$

Thus, for each $k \in K$ it is possible to find $j(k) \in \{1, \ldots, m\}$ such that

$$
f_{j(k)}(x_k + \alpha_k d_k) - f_{j(k)}(x_k) + \rho(\alpha_k \|d_k\|) \geq 0
$$

# Convergence Results

Consider a refining subsequence converging to $x_*$ (and assume that $F$ is Lipschitz continuous near $x_*$)

**Theorem**

*If the set of refining directions for $x_*$ is dense in $T_\Omega(x_*)$ then $x_*$ is a Pareto-Clarke critical point:*

$$\forall d \in T_\Omega^{Cl}(x_*), \exists j = j(d) \in \{1, \ldots, m\}, f_j^\circ(x_*; d) \geq 0$$

# Convergence Results

Consider a refining subsequence converging to $x_*$ (and assume that $F$ is Lipschitz continuous near $x_*$)

**Theorem**

*If the set of refining directions for $x_*$ is dense in $T_\Omega(x_*)$ then $x_*$ is a Pareto-Clarke critical point:*

$$\forall d \in T_\Omega^{Cl}(x_*), \exists j = j(d) \in \{1, \ldots, m\}, f_j^\circ(x_*; d) \geq 0$$

Note: For $m = 1$ reduces to the classical results of DS.

# Presentation Outline

1. Motivation and General Concepts
2. Direct Search Methods (DSM) for MOO
3. Direct MultiSearch (DMS)
4. Conclusions and references
5. Convergence Analysis in DSM
6. Numerical Results

# Numerical Testing Framework

Problems

- 100 bound constrained MOO problems with:
- number of variables between 1 and 30
- number of objectives between 2 and 4

# Numerical Testing Framework

Problems

- 100 bound constrained MOO problems with:
- number of variables between 1 and 30
- number of objectives between 2 and 4

Solvers

- DMS tested against 8 different MOO solvers (complete results available at `http://www.mat.uc.pt/dms`)
- results reported only for
  AMOSA – simulated annealing code
  BIMADS – based on Mesh Adaptive Direct Search
  NSGA-II (C version) – genetic algorithm code

All solvers tested with default values

# DMS Numerical Options

- No search step
- List initialization: line sampling
- List selection: all current nondominated points
- List ordering: new points added at the end of the list, poll center moved to the end of the list

# DMS Numerical Options

- No search step
- List initialization: line sampling
- List selection: all current nondominated points
- List ordering: new points added at the end of the list, poll center moved to the end of the list
- Positive basis: $[I \;\; -I]$
- Step size parameter: $\alpha_0 = 1$, halved at unsuccessful iterations
- Stopping criteria: minimum step size of $10^{-3}$ or a maximum of 20000 function evaluations

# Performance Metrics – Purity

$F_{p,s}$ (approximated Pareto front computed by solver $s$ for problem $p$)

$F_p$ (approximated Pareto front computed for problem $p$, using results for all solvers)

Purity value for solver $s$ on problem $p$:

$$\frac{|F_{p,s} \cap F_p|}{|F_{p,s}|}$$

# Performance profiles [Dolan and Moré]

$t_{p,s} =$ lower values of $t_{p,s}$ indicate better performance

$$\rho_s(\tau) = \frac{|\{p \in \mathcal{P} : r_{p,s} \leq \tau\}|}{|\mathcal{P}|}$$

with $r_{p,s} = t_{p,s} / \min\{t_{p,\bar{s}} : \bar{s} \in \mathcal{S}\}$

# Performance profiles [Dolan and Moré]

$t_{p,s}$ = lower values of $t_{p,s}$ indicate better performance

$$\rho_s(\tau) = \frac{|\{p \in \mathcal{P} : r_{p,s} \leq \tau\}|}{|\mathcal{P}|}$$

with $r_{p,s} = t_{p,s} / \min\{t_{p,\bar{s}} : \bar{s} \in \mathcal{S}\}$

- incorporates results for all problems and all solvers.

# Performance profiles [Dolan and Moré]

$t_{p,s}$ = lower values of $t_{p,s}$ indicate better performance

$$\rho_s(\tau) = \frac{|\{p \in \mathcal{P} : r_{p,s} \leq \tau\}|}{|\mathcal{P}|}$$

with $r_{p,s} = t_{p,s} / \min\{t_{p,\bar{s}} : \bar{s} \in \mathcal{S}\}$

- incorporates results for all problems and all solvers.
- Allows to access 'efficiency' and robustness.

# Performance profiles [Dolan and Moré]

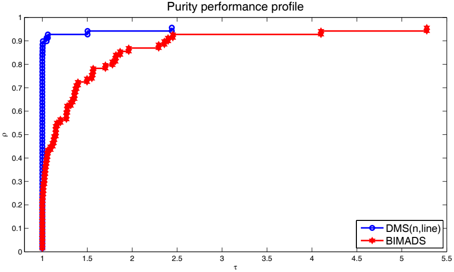
$t_{p,s}$ = lower values of $t_{p,s}$ indicate better performance

$$\rho_s(\tau) = \frac{|\{p \in \mathcal{P} : r_{p,s} \leq \tau\}|}{|\mathcal{P}|}$$

with $r_{p,s} = t_{p,s} / \min\{t_{p,\bar{s}} : \bar{s} \in \mathcal{S}\}$

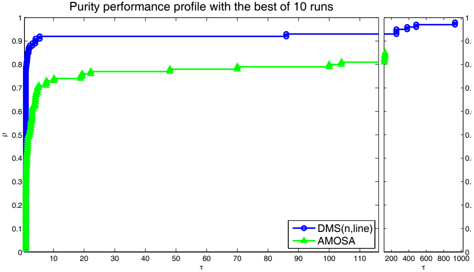
- incorporates results for all problems and all solvers.
- Allows to access 'efficiency' and robustness.
- $\rho_s(\tau)$, with $\tau = 1$, is the probability of the solver $s$ winning over the remaining ones, represents 'efficiency' of solver $s$
- $\rho_s(\tau)$, with $\tau$ large, give robustness of solver $s$

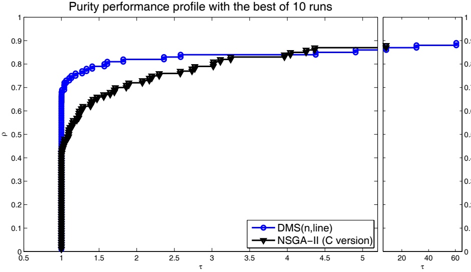
# Comparing DMS to Other Solvers (Purity)




Purity Metric
(percentage of points generated in the reference Pareto front)

$$\frac{|F_{p,s} \cap F_p|}{|F_{p,s}|}$$

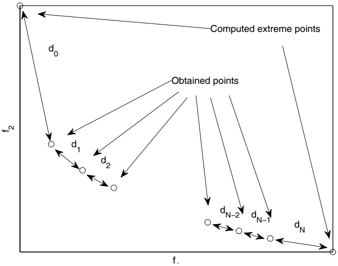


# Performance Metrics – Spread

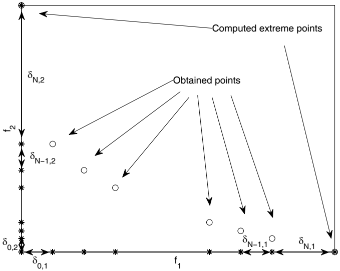
## Gamma Metric
(largest gap in the Pareto front)

$$\Gamma_{p,s} = \max_{j\in\{1,\ldots,m\}} \left( \max_{i\in\{0,\ldots,N\}} \{\delta_{i,j}\} \right)$$




## Delta Metric
(uniformity of gaps in the Pareto front)

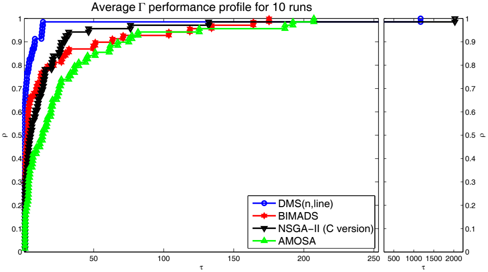
$$\Delta_{p,s} = \max_{j\in\{1,\ldots,m\}} \left( \frac{\delta_{0,j} + \delta_{N,j} + \sum_{i=1}^{N-1} |\delta_{i,j} - \bar{\delta}_j|}{\delta_{0,j} + \delta_{N,j} + (N-1)\bar{\delta}_j} \right)$$



# Comparing DMS to Other Solvers (Spread)

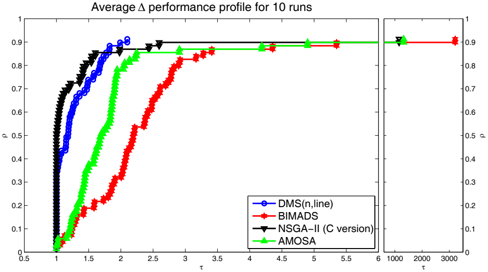
## Gamma Metric

largest gap in the Pareto front




## Delta Metric

uniformity of gaps in the Pareto front



# Data Profiles [Moré and Wild (2009)]

Indicate how likely is an algorithm to ‘solve’ a problem, given some computational budget.

Let $h_{p,s}$ be the number of function evaluations required for solver $s$ to solve problem $p$.

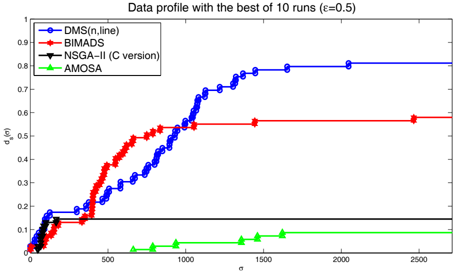
Consider

$$d_s(\sigma) \;=\; \frac{|\{p \in \mathcal{P} : h_{p,s} \leq \sigma\}|}{|\mathcal{P}|}$$

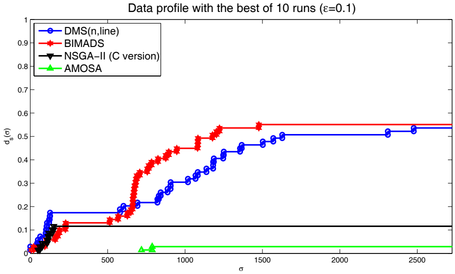
Problem solved to $\epsilon$ – accuracy (up to some level $\epsilon$ of accuracy) :

$$\frac{|F_{p,s} \cap F_p|}{|F_p|/|\mathcal{S}|} \;\geq\; 1 - \varepsilon$$

$\epsilon = 0.5$




$\epsilon = 0.1$




# maximum function evaluations = 5000

# Improving DMS Performance

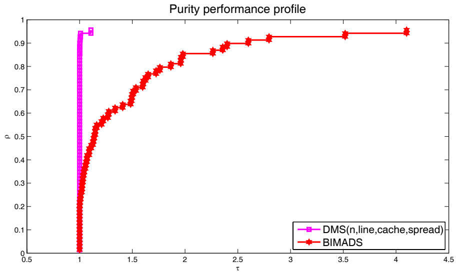
- Cache implementation: before evaluating a point, one checks, using a infinity norm, whether it has been previously evaluating. Objective function values only computed for points that dist at least $10^{-3}$ from any previously evaluated point

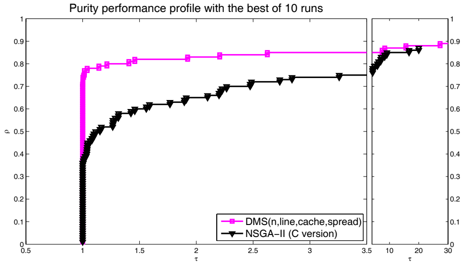
# Improving DMS Performance

- Cache implementation: before evaluating a point, one checks, using a infinity norm, whether it has been previously evaluating. Objective function values only computed for points that dist at least $10^{-3}$ from any previously evaluated point
- Ordering strategy for $L_k$ based on the $\Gamma$ metric: poll centers correspond to the largest gap in the approximated Pareto front

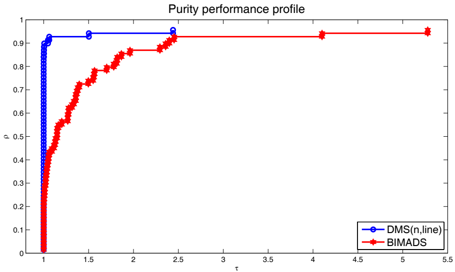
# Improving DMS Performance (Purity)

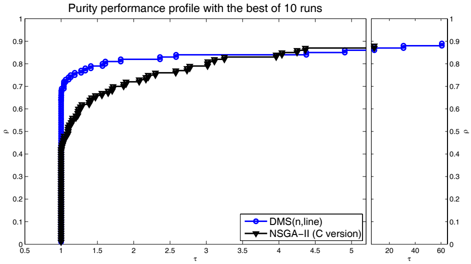
## New Version




## Old Version

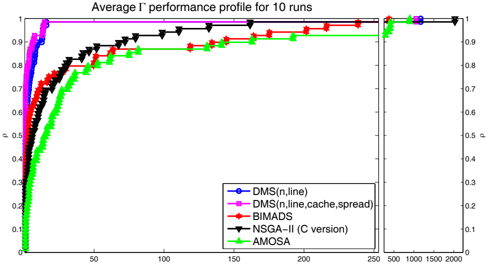


# Improving DMS Performance (Spread)

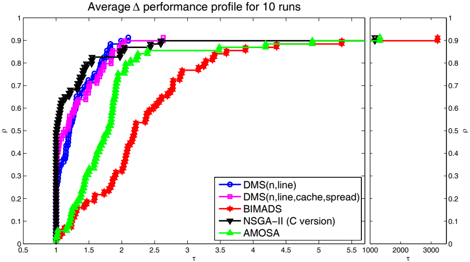
## Gamma Metric

largest gap in the Pareto front

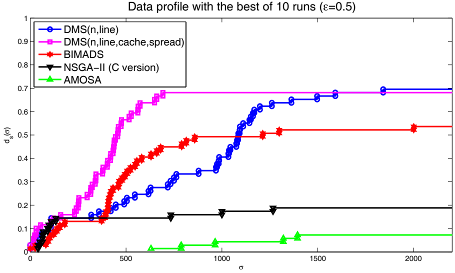



## Delta Metric

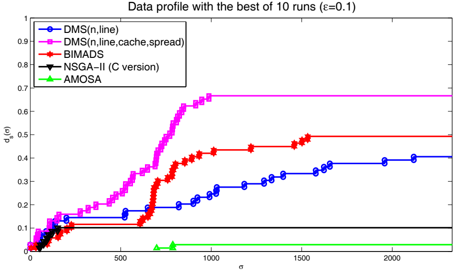
uniformity of gaps in the Pareto front



$\epsilon = 0.5$




$\epsilon = 0.1$




# maximum function evaluations = 5000