Direct Search for Multiobjective Optimization

Luís Nunes Vicente
University of Coimbra

(main co-authors: A. L. Custódio and A. I. F. Vaz)

4 mai 2011 — Journées de l'Optimisation 2011

http//www.mat.uc.pt/~lnv

Problem Features

Multiobjective Derivative-Free Optimization

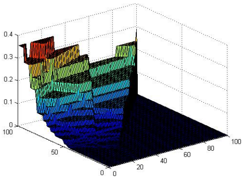
$$\min_{x \in \Omega \subseteq \mathbb{R}^n} f(x) \equiv (f_1(x), f_2(x), \ldots, f_m(x))^\top$$

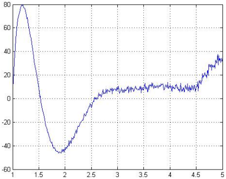
$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \; j = 1, \ldots, m \geq 1$$

Multiobjective Derivative-Free Optimization

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

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





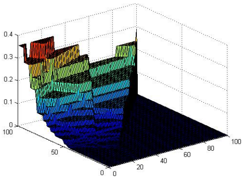
- several objectives, often conflicting

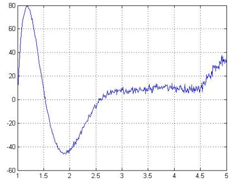
Problem Features

Multiobjective Derivative-Free Optimization

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

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



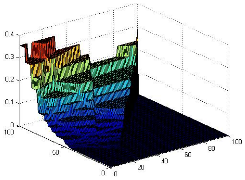


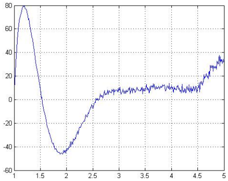
- several objectives, often conflicting
- functions with unknown derivatives

Multiobjective Derivative-Free Optimization

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

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





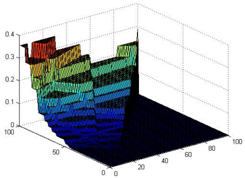
- several objectives, often conflicting
- functions with unknown derivatives
- expensive function evaluations, possibly subject to noise

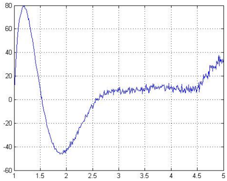
Problem Features

Multiobjective Derivative-Free Optimization

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

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



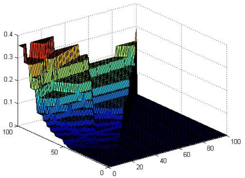


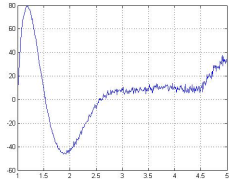
- several objectives, often conflicting
- functions with unknown derivatives
- expensive function evaluations, possibly subject to noise
- source code not available for use

Multiobjective Derivative-Free Optimization

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

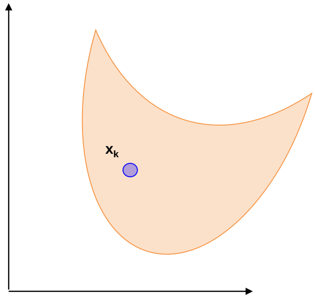
$$f_j : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \; j = 1, \ldots, m \geq 1$$



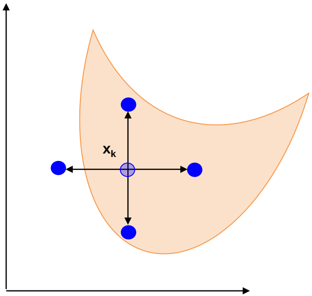


- several objectives, often conflicting
- functions with unknown derivatives
- expensive function evaluations, possibly subject to noise
- source code not available for use
- unpractical to compute approximations to derivatives

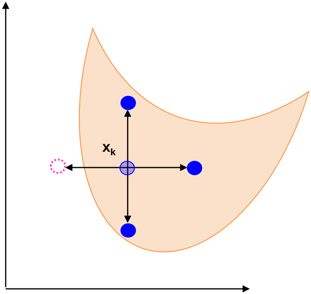
Poll Step Example — Single Objective



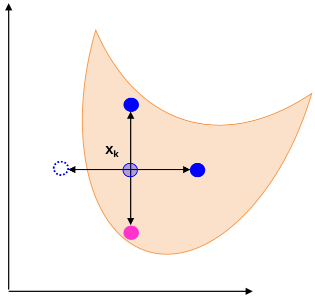
Poll Step Example — Single Objective



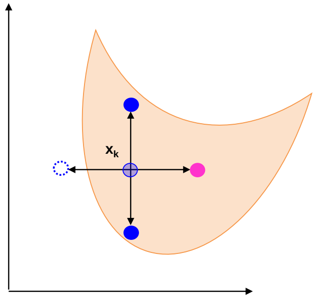
Poll Step Example — Single Objective



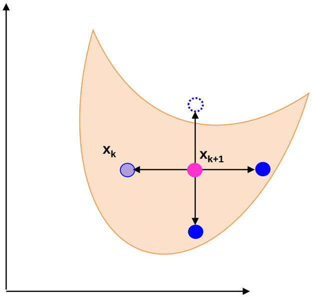
Poll Step Example — Single Objective



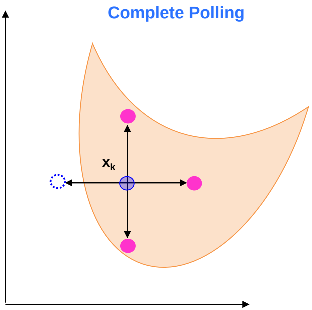
Poll Step Example — Single Objective



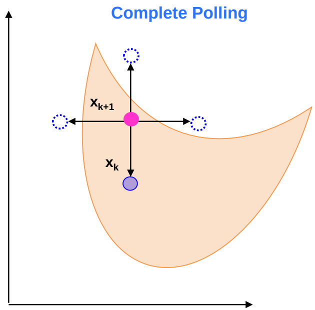
Poll Step Example — Single Objective



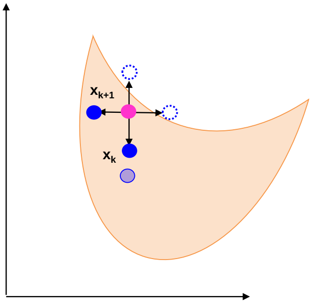
Poll Step Example — Single Objective



Poll Step Example — Single Objective



Poll Step Example — Single Objective



Pareto Dominance

$$x \prec y \text{ (}x\text{ dominates }y\text{)} \quad \Longleftrightarrow f(x) \prec_f f(y) \Longleftrightarrow$$

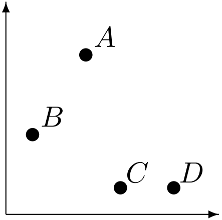
$$f(x) \leq f(y), \text{ with } f(x) \neq f(y)$$

Pareto Dominance

$$x \prec y \text{ (} x \text{ dominates } y\text{)} \iff f(x) \prec_f f(y) \iff$$

$$f(x) \leq f(y), \text{ with } f(x) \neq f(y)$$

- B dominates A
- C dominates D
- B and C are nondominated

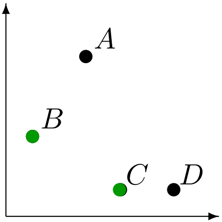


Pareto Dominance

$$x \prec y \text{ (} x \text{ dominates } y\text{)} \iff f(x) \prec_f f(y) \iff$$

$$f(x) \leq f(y), \text{ with } f(x) \neq f(y)$$

- B dominates A
- C dominates D
- B and C are nondominated



Pareto Front: $\{B, C\}$

Direct MultiSearch (DMS) Main Lines

- does not aggregate any of the objective functions

Direct MultiSearch (DMS) Main Lines

- does not aggregate any of the objective functions
- generalizes ALL direct search (DS) methods of directional type to multiobjective optimization (MOO)

Direct MultiSearch (DMS) Main Lines

- does not aggregate any of the objective functions
- generalizes ALL direct search (DS) methods of directional type to multiobjective optimization (MOO)
- tries to capture the whole Pareto front from the polling procedure

Direct MultiSearch (DMS) Main Lines

- keeps a list of feasible nondominated points

Direct MultiSearch (DMS) Main Lines

- keeps a list of feasible nondominated points
- poll centers are chosen from the list

Direct MultiSearch (DMS) Main Lines

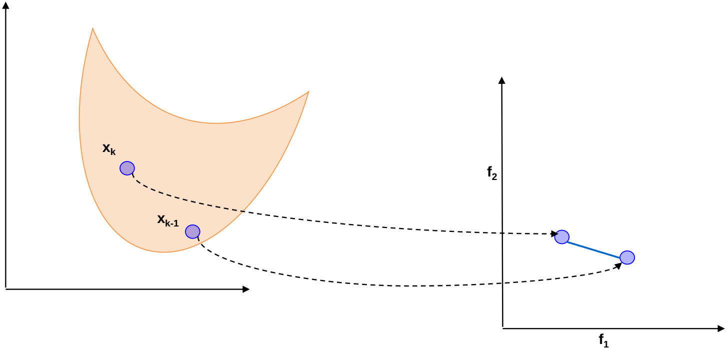
- keeps a list of feasible nondominated points
- poll centers are chosen from the list
- successful iterations correspond to list changes,

Direct MultiSearch (DMS) Main Lines

- keeps a list of feasible nondominated points
- poll centers are chosen from the list
- successful iterations correspond to list changes,

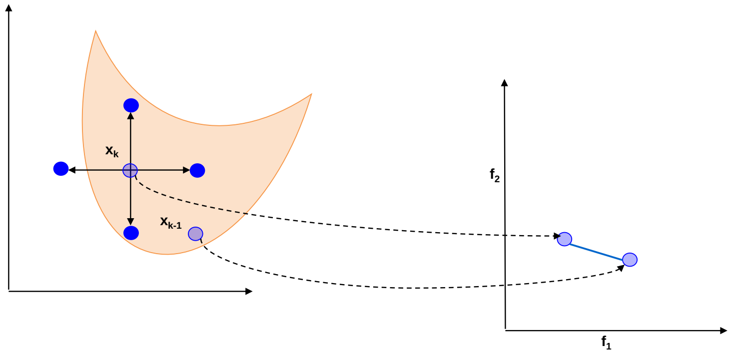
  an iteration is successful only if it produces a feasible nondominated point

Poll Step Example — Biojective

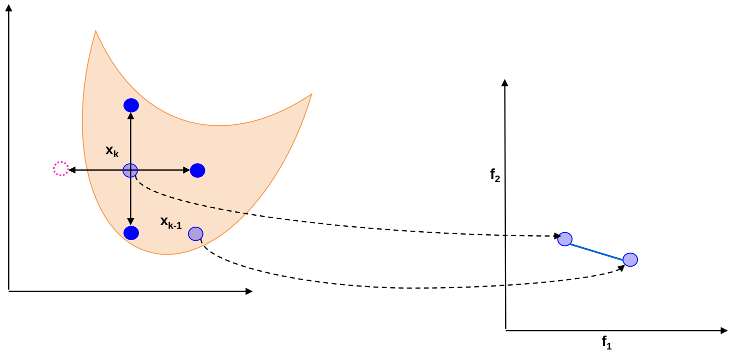


L_k

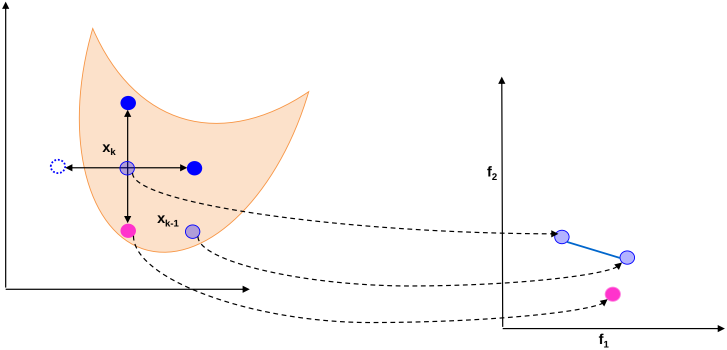
Poll Step Example — Biojective



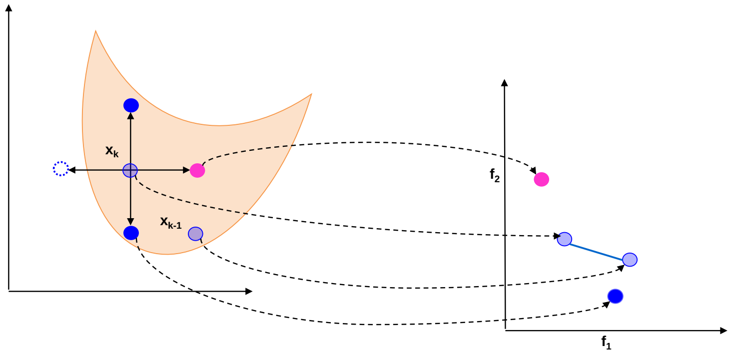
Poll Step Example — Biojective

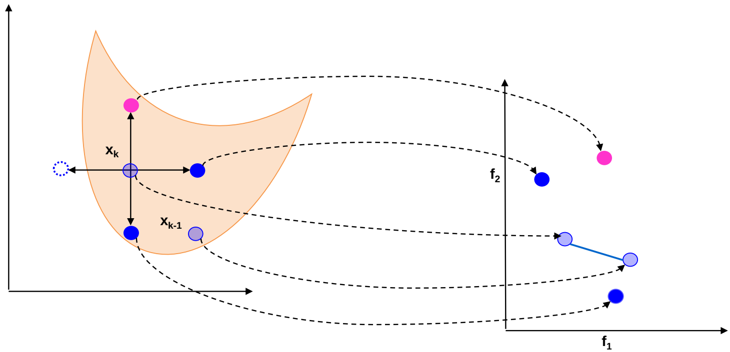


Poll Step Example — Biojective

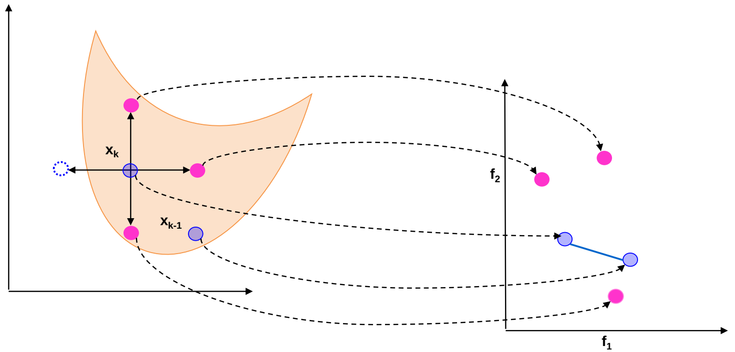


Poll Step Example — Biojective



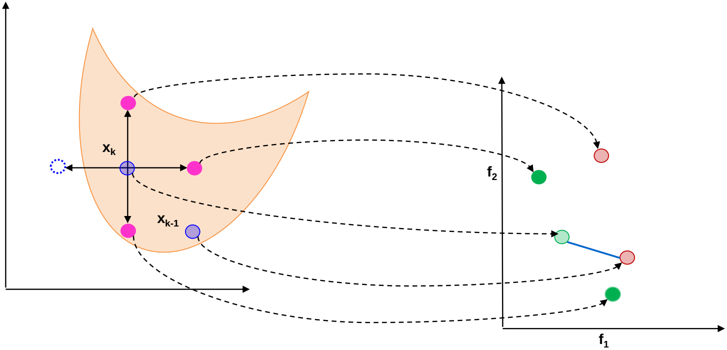


x_k
x_{k-1}
f_2
f_1

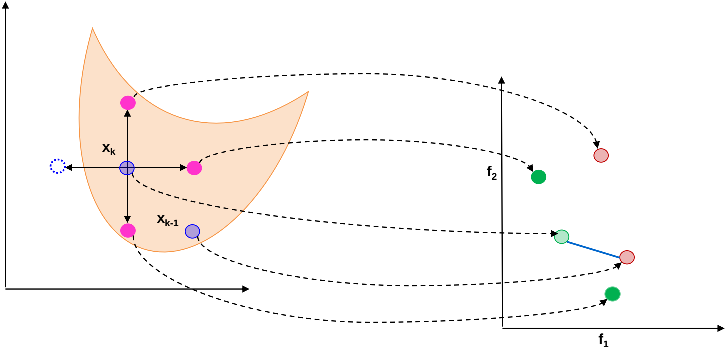


L_{add}

Poll Step Example — Biojective

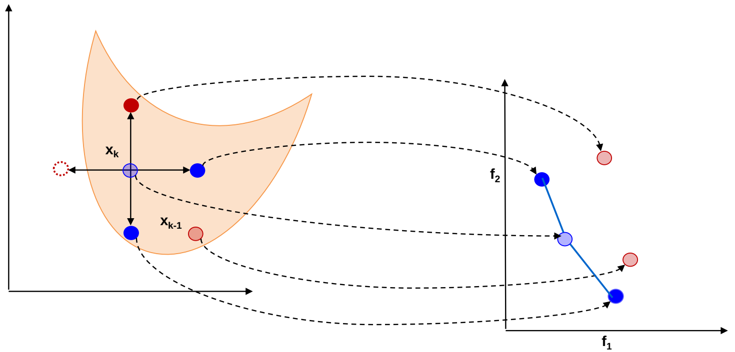


$L_{filtered}$



$$L_{trial} = L_{filtered}$$

Poll Step Example — Biojective



L_{k+1}

Direct MultiSearch (DMS) Main Lines

- makes use of search/poll paradigm

Direct MultiSearch (DMS) Main Lines

- makes use of search/poll paradigm
- incorporates an optional search step (only to disseminate the search)

Direct MultiSearch (DMS) Main Lines

- makes use of search/poll paradigm
- incorporates an optional search step (only to disseminate the search)
- but, again, tries to capture the whole Pareto front from the polling procedure

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

$L_{k+1} = L_{trial}$

Unsuc

Poll Step

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

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

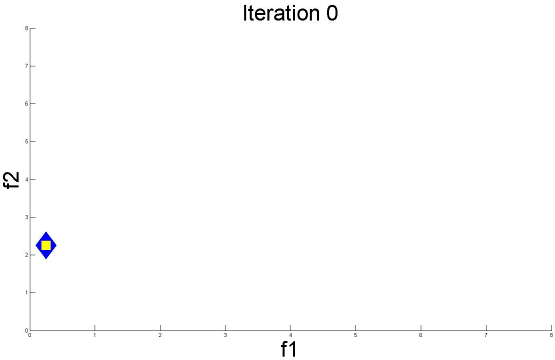
Suc

$L_{k+1} = L_{trial}$

Unsuc

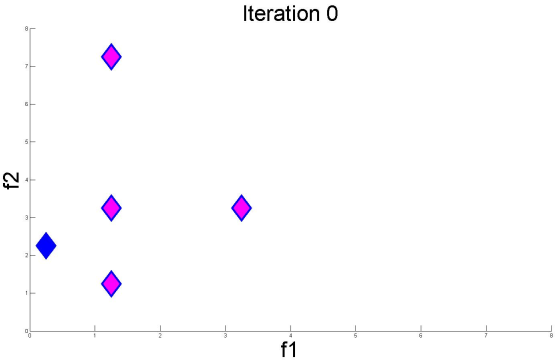
Decrease the step size

Numerical Example – Problem SP1 (Huband et al. [2005])



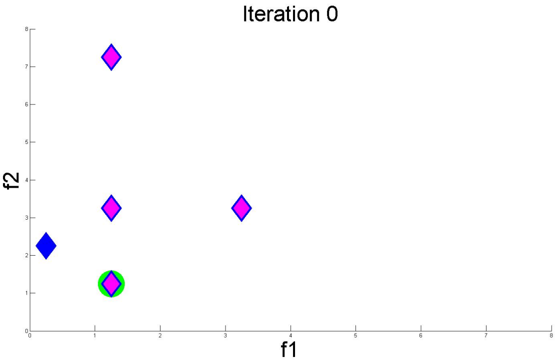
♦ Evaluated points since beginning
■ Current iterate list

Numerical Example – Problem SP1 (Huband et al. [2005])



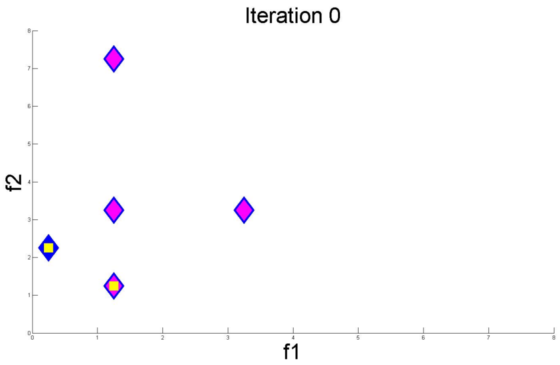
- Evaluated poll points
- Evaluated points since beginning

Numerical Example – Problem SP1 (Huband et al. [2005])



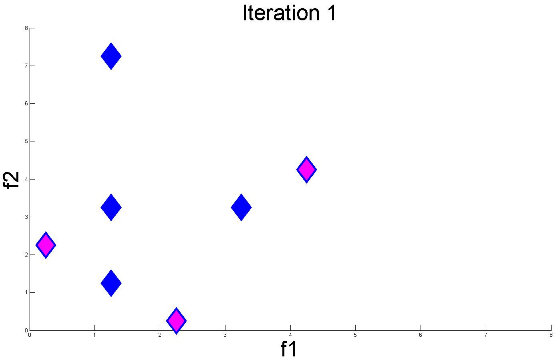
- Nondominated evaluated poll points

Numerical Example – Problem SP1 (Huband et al. [2005])



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

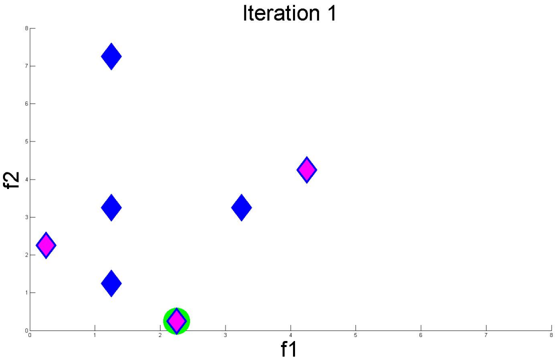
Numerical Example – Problem SP1 (Huband et al. [2005])



♦ Evaluated poll points

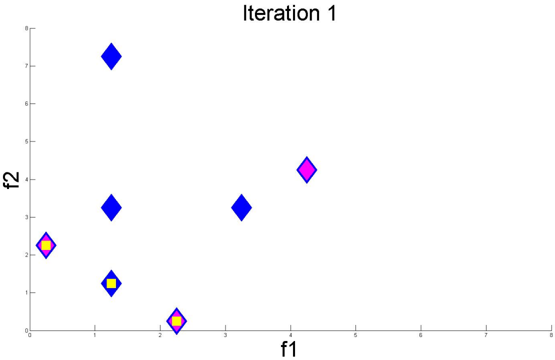
♦ Evaluated points since beginning

Numerical Example – Problem SP1 (Huband et al. [2005])

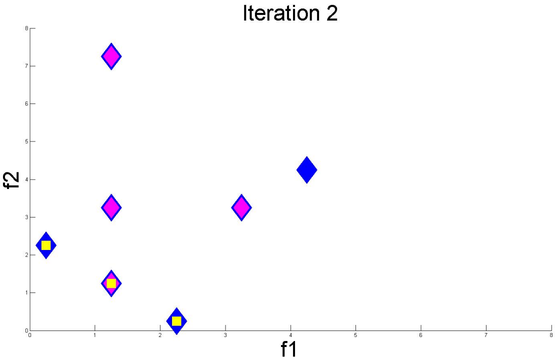


- Nondominated evaluated poll points

Numerical Example – Problem SP1 (Huband et al. [2005])

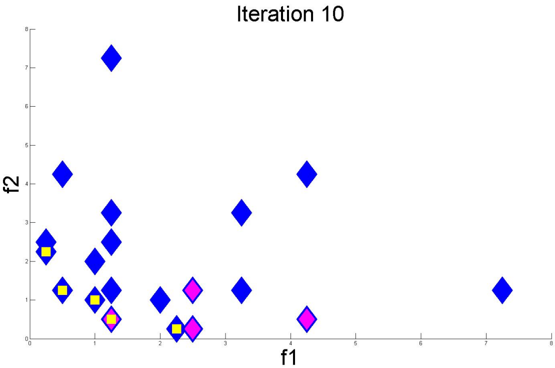


- Evaluated poll points
- Evaluated points since beginning
- Current iterate list

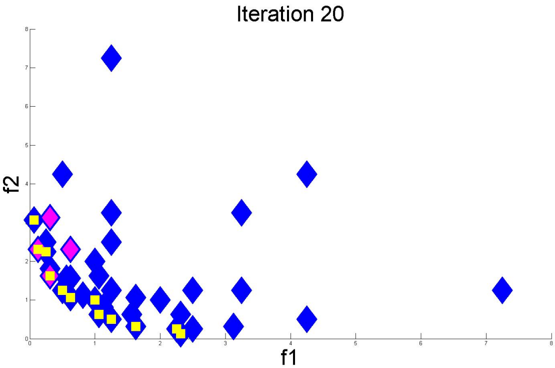


- Evaluated poll points
- Evaluated points since beginning
- Current iterate list

Numerical Example – Problem SP1 (Huband et al. [2005])

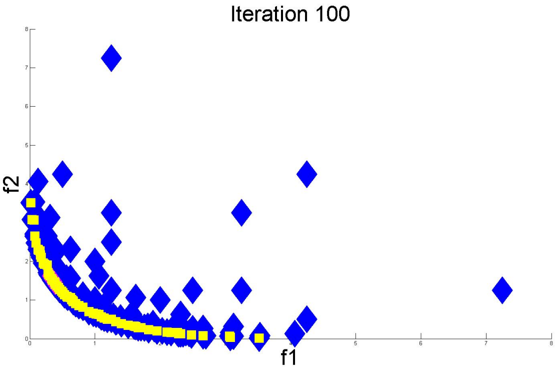


- Evaluated poll points
- Evaluated points since beginning
- Current iterate list



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

Numerical Example – Problem SP1 (Huband et al. [2005])



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

Plenty of Algorithmic Flexibility in DMS

- In the list initialization

Plenty of Algorithmic Flexibility in DMS

- In the list initialization
- In applying a search step

Plenty of Algorithmic Flexibility in DMS

- In the list initialization
- In applying a search step
- In the polling strategies: complete or opportunistic

Plenty of Algorithmic Flexibility in DMS

- In the list initialization
- In applying a search step
- In the polling strategies: complete or opportunistic
- In the selection of the trial list, as long as previous feasible nondominated points are not removed from the iterate list

Globalization

As in DS, convergence to stationarity from arbitrary starting points (global convergence) is ensured from polling

Globalization

As in DS, convergence to stationarity from arbitrary starting points (global convergence) is ensured from polling

Constraints are handled using the extreme barrier function

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

Globalization Strategies

Using Integer/Rational Lattices (Torczon [1997], Audet and Dennis [2003])

- requires only simple decrease
- poll directions and step size must satisfy integer/rational requirements
- search step is restricted to an implicit mesh

Globalization Strategies

Using Integer/Rational Lattices (Torczon [1997], Audet and Dennis [2003])

- requires only simple decrease
- poll directions and step size must satisfy integer/rational requirements
- search step is restricted to an implicit mesh

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

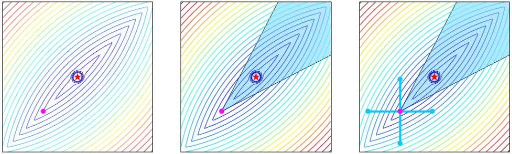
- use of a forcing function

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

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

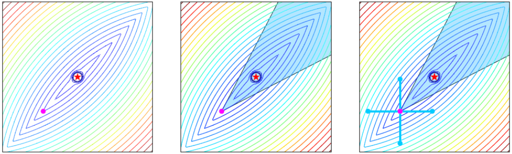
Polling Directions

If f is nonsmooth, a finite $\#$ of polling directions may not suffice



Kolda, Lewis, and Torczon [2003]

If f is nonsmooth, a finite $\#$ of polling directions may not suffice



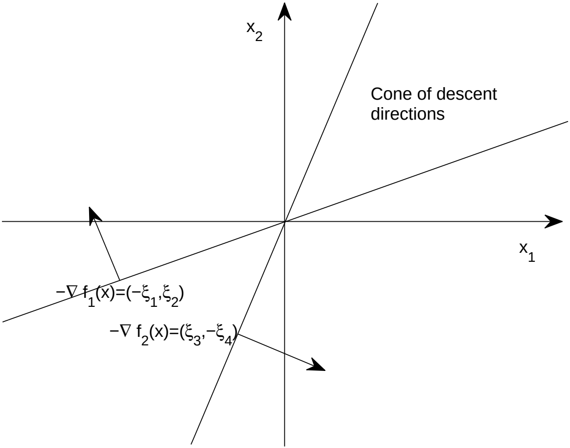
Kolda, Lewis, and Torczon [2003]

Nonsmooth Optimization

In addition to globalization requirements, the union of all normalized poll directions should be asymptotically dense in the unit sphere

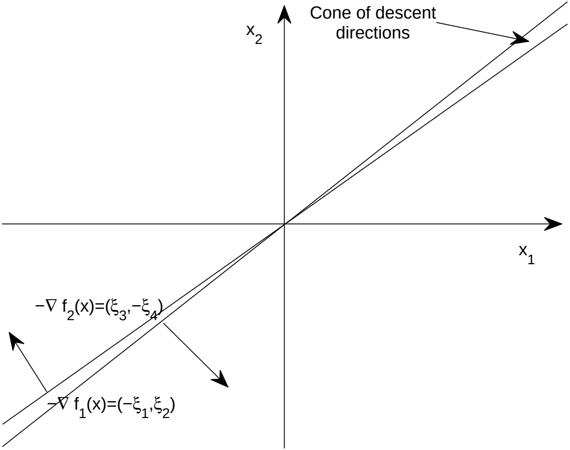
(possible strategies: randomly generated directions, LTMADS, ORTHOMADS)

Polling Directions



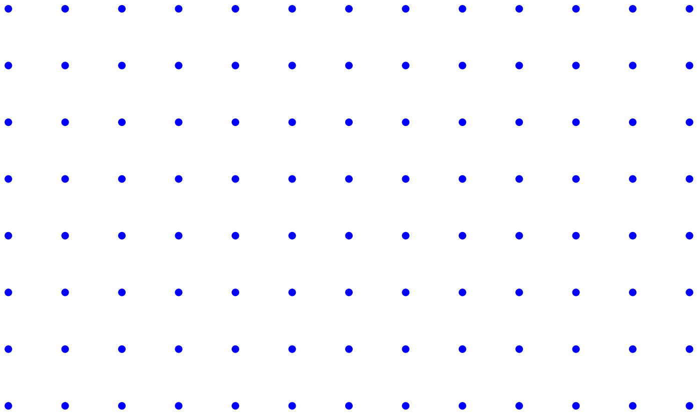
The cone of descent directions for all objective functions can be made narrower.

Polling Directions



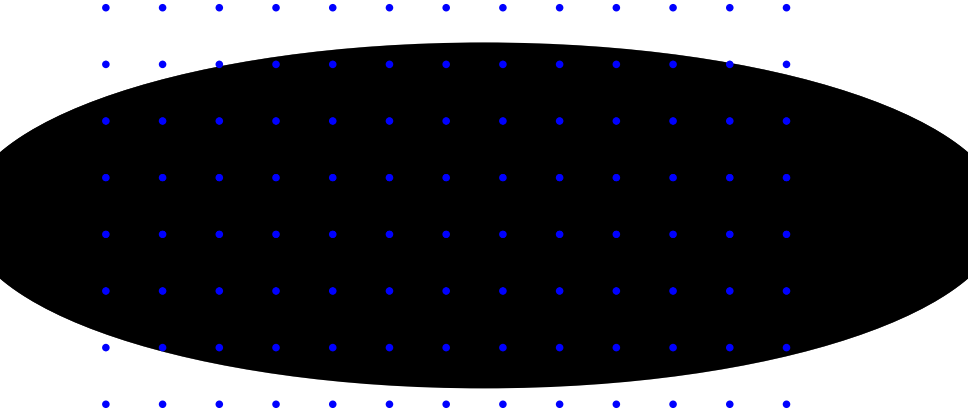
Hence, one does need also here the polling directions to be asymptotically dense in the unit sphere

All potential iterates lie on an integer lattice when the step size α_k is bounded away from zero



Intuitively, if α_k does not $\longrightarrow 0$, points in this integer lattice would be separated by a finite and positive distance

Refining Subsequences — Integer Lattices



It would therefore be impossible to fit an infinity of iterates inside a bounded level set

Refining Subsequences — Integer Lattices

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: Custódio, Madeira, Vaz, and Vicente [2010]

DS: Torczon [1997], Audet and Dennis [2003]

An iteration is successful only if it produces a nondominated point that has sufficiently decreased one of the f_j's relatively to at least a point in the list

An iteration is successful only if it produces a nondominated point that has sufficiently decreased one of the f_j's relatively to at least a point in the list

Also, $\rho(\alpha_k)$ is a monotonically increasing function of the step size α_k

Refining Subsequences — Sufficient Decrease

An iteration is successful only if it produces a nondominated point that has sufficiently decreased one of the f_j's relatively to at least a point in the list

Also, $\rho(\alpha_k)$ is a monotonically increasing function of the step size α_k

Thus, α_k cannot be bounded away from zero since otherwise one of the f_j's would tend to $-\infty$

Refining Subsequences — Sufficient Decrease

An iteration is successful only if it produces a nondominated point that has sufficiently decreased one of the f_j's relatively to at least a point in the list

Also, $\rho(\alpha_k)$ is a monotonically increasing function of the step size α_k

Thus, α_k cannot be bounded away from zero since otherwise one of the f_j's would tend to $-\infty$

Intuitively, insisting on a sufficient decrease will make it harder to have a successful step and therefore will generate more unsuccessful poll steps

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: Custódio, Madeira, Vaz, and Vicente [2010]

DS: Lewis, Tolds, and Torczon [2003]

Stationarity results for DS or DMS consist of nonnegativity of generalized directional derivatives along certain limit directions

Refining Directions

Stationarity results for DS or DMS consist of nonnegativity of generalized directional derivatives along certain limit directions

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]

First, let us focus on the unconstrained case, $\Omega = \mathbb{R}^n$

Clarke Generalized Directional Derivative

First, let us focus on the unconstrained case, $\Omega = \mathbb{R}^n$

Clarke Generalized Directional Derivative

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

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

Clarke Stationarity

Assume that f is Lipschitz continuous near x_*

Clarke Stationarity

Assume that f is Lipschitz continuous near x_*

Definition

x_* is a Clarke critical point if

$$\forall d \in \mathbb{R}^n, f^\circ(x_*; d) \geq 0$$

Clarke Stationarity — Constrained Case

Also assuming that f is Lipschitz continuous near x_*

Definition

x_* *is a Clarke critical point if*

$$\forall d \in T_\Omega(x_*), f^\circ(x_*; d) \geq 0$$

Also assuming that f is Lipschitz continuous near x_*

Definition

x_* is a Clarke critical point if

$$\forall d \in T_\Omega(x_*), f^\circ(x_*; d) \geq 0$$

$T_\Omega(x_*)$ is the tangent cone to Ω at x_*

Clarke Stationarity — Constrained Case

Also assuming that f is Lipschitz continuous near x_*

Definition

x_* is a Clarke critical point if

$$\forall d \in T_\Omega(x_*), f^\circ(x_*; d) \geq 0$$

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

Clarke Stationarity — Constrained Case

Also assuming that f is Lipschitz continuous near x_*

Definition

x_* is a Clarke critical point if

$$\forall d \in T_\Omega(x_*), f^\circ(x_*; d) \geq 0$$

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

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

Clarke-Jahn Generalized Directional Derivative

For f Lipschitz continuous near x_*

Clarke-Jahn Generalized Directional Derivative

For f Lipschitz continuous near x_*

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

Clarke Stationarity — Constrained Case

Clarke-Jahn Generalized Directional Derivative

For f Lipschitz continuous near x_*

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

for $v \in \operatorname{int}(T_\Omega(x_*))$

Clarke Stationarity — Constrained Case

Clarke-Jahn Generalized Directional Derivative

For f Lipschitz continuous near x_*

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

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

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

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

Pareto-Clarke Stationarity

Assume that f is Lipschitz continuous near x_*

Definition

x_* *is a Pareto-Clarke critical point if*

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

Convergence Results

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

Convergence Results

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

Theorem

If $d \in \mathrm{int}(T_\Omega(x_*))$ 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 [2010]

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

Proof Sketch

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

Proof Sketch

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

Proof Sketch

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

Proof Sketch

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

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

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

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_*)

Convergence Results

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

By passing from the interior to the closure of the tangent cone, as in DS $(m = 1)$

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

By passing from the interior to the closure of the tangent cone, as in DS $(m = 1)$

Theorem

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

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

Numerical Testing Framework

Problems

- 100 bound constrained MOO problems (AMPL models available at http://www.mat.uc.pt/dms)
- number of variables between 1 and 30
- number of objectives between 2 and 4

Numerical Testing Framework

Problems

- 100 bound constrained MOO problems (AMPL models available at http://www.mat.uc.pt/dms)
- number of variables between 1 and 30
- number of objectives between 2 and 4

Solvers

- DMS version 0.1 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
- Polling directions: $[I \ \ -I]$
- 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

Consider

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

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

Performance Metrics – Purity

Consider

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

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

The Purity value for solver s on problem p is

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

Performance Profiles (Dolan and Moré [2002])

Let $t_{p,s}$ be a metric for which lower values indicate better performance

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

Performance Profiles (Dolan and Moré [2002])

Let $t_{p,s}$ be a metric for which lower values indicate better performance

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

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

Performance Profiles (Dolan and Moré [2002])

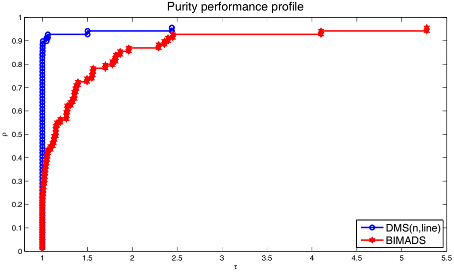
Let $t_{p,s}$ be a metric for which lower values indicate better performance

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

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

- incorporates results for all problems and all solvers
- allows to access ‘efficiency’ and robustness
- $\rho_s(1)$ represents ‘efficiency’ of solver s
- $\rho_s(\tau)$, with τ large, gives robustness of solver s

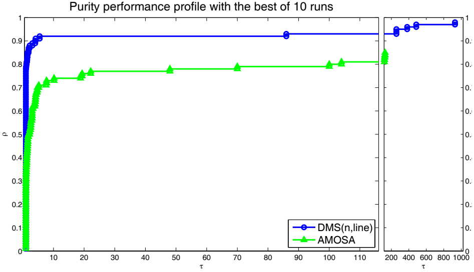
Comparing DMS to Other Solvers (Purity)

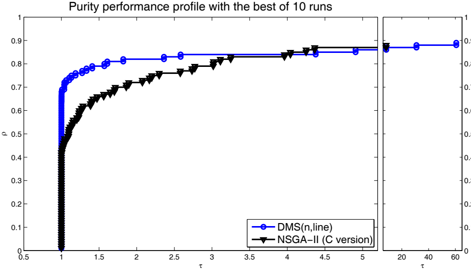


Purity Metric

(percentage of points generated in the reference Pareto front)

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

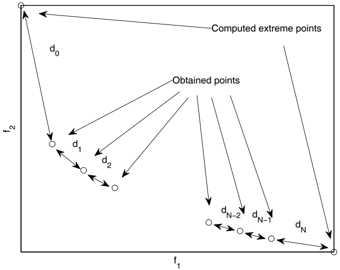




Gamma Metric

(largest gap in the Pareto front)

$$\Gamma_{p,s} \; = \; \max_{i \in \{0,\ldots,N\}} \{d_i\}$$



Delta Metric

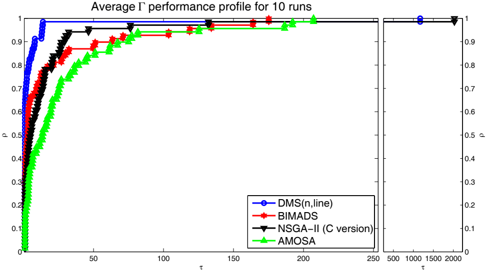
(uniformity of gaps in the Pareto front)

$$\Delta_{p,s} \; = \; \frac{d_0 + d_N + \sum_{i=1}^{N-1} |d_i - \bar{d}|}{d_0 + d_N + (N-1)\bar{d}}$$

Gamma Metric

(largest gap in the Pareto front)

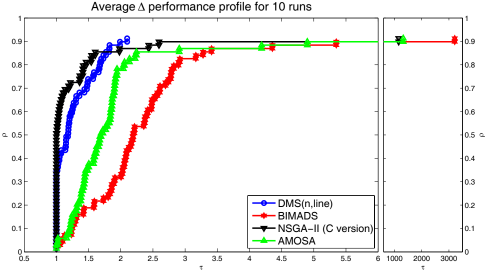
$$\Gamma_{p,s} \;=\; \max_{i\in\{0,\ldots,N\}}\{d_i\}$$



Delta Metric

(uniformity of gaps in the Pareto front)

$$\Delta_{p,s} \;=\; \frac{d_0+d_N+\sum_{i=1}^{N-1}|d_i-\bar{d}|}{d_0+d_N+(N-1)\bar{d}}$$



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

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

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

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

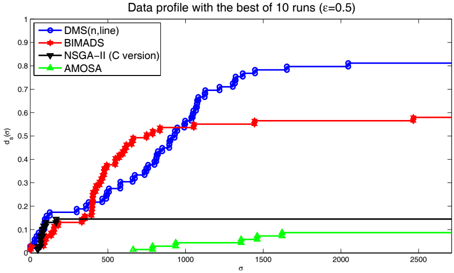
$$d_s(\sigma) \;=\; \frac{|\{p \in \mathcal{P} : h_{p,s} \leq \sigma\}|}{|\mathcal{P}|}$$

A problem is solved to ϵ–accuracy if

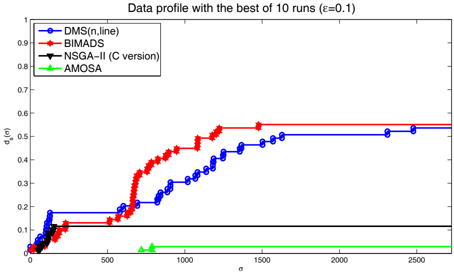
$$\frac{|F_{p,s} \cap F_p|}{|F_p|/|\mathcal{S}|} \;\geq\; 1 - \varepsilon$$

Comparing DMS to Other Solvers

$\epsilon = 0.5$



$\epsilon = 0.1$



maximum function evaluations = 5000

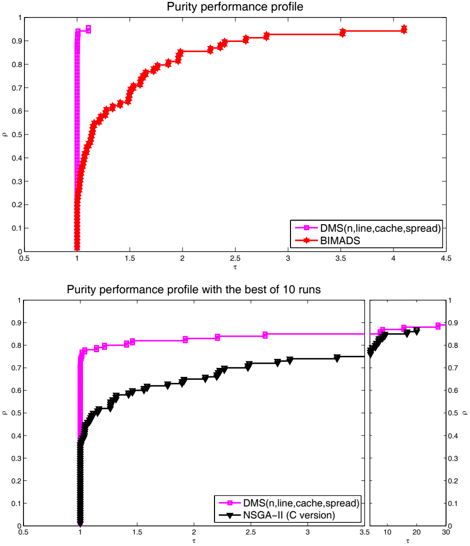
Improving DMS Performance

- Cache implementation: objective function values only computed for points that dist at least 10^{-3} from any previously evaluated point

Improving DMS Performance

- Cache implementation: 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 Γ metric: poll centers correspond to the highest Γ metric value (ties broken by the largest step size)

Improving DMS Performance (Purity)



Purity Metric

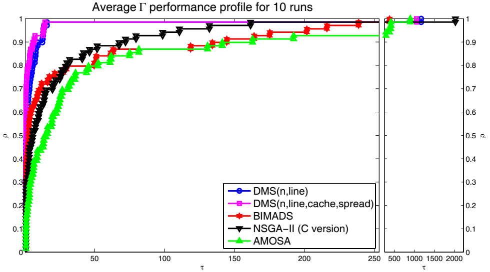
(percentage of points generated in the reference Pareto front)

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

Gamma Metric

(largest gap in the Pareto front)

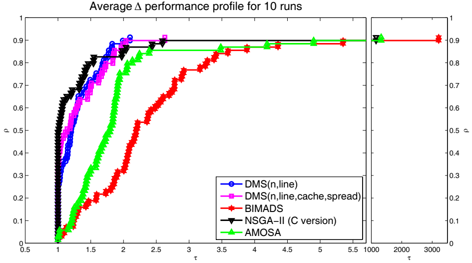
$$\Gamma_{p,s} \;=\; \max_{i\in\{0,\ldots,N\}}\{d_i\}$$



Delta Metric

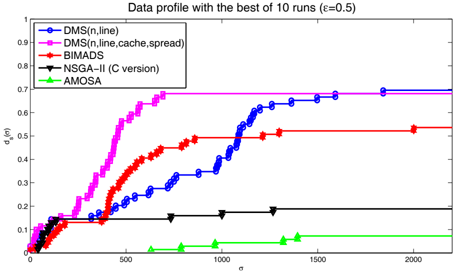
(uniformity of gaps in the Pareto front)

$$\Delta_{p,s} \;=\; \frac{d_0+d_N+\sum_{i=1}^{N-1}|d_i-\bar{d}|}{d_0+d_N+(N-1)\bar{d}}$$

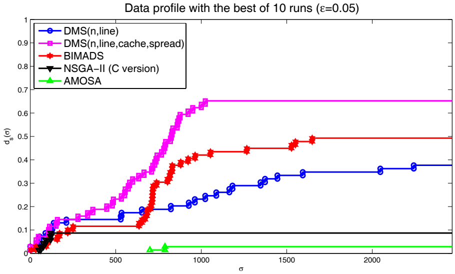


Improving DMS Performance (Data Profiles)

$\epsilon = 0.5$



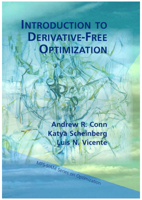
$\epsilon = 0.05$



maximum function evaluations = 5000

Our References

- A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz, and L. N. Vicente, Direct multisearch for multiobjective optimization, in review for SIAM Journal on Optimization
- L. N. Vicente and A. L. Custódio, Analysis of direct searches for discontinuous functions, to appear in Mathematical Programming
- A. R. Conn, K. Scheinberg, and L. N. Vicente, Introduction to Derivative-Free Optimization, MPS-SIAM Book Series on Optimization, SIAM, Philadelphia, 2009



Optimization 2011 (July 24–27, New University of Lisbon, Portugal)

Deadline for Abstract Submission – March 31

optimization 2011
Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa
Caparica *Campus*
July 24-27

plenary speakers

Gilbert Laporte | HEC Montréal
New trends in vehicle routing

Jean Bernard Lasserre | LAAS-CNRS, Toulouse
Moments and semidefinite relaxations for parametric optimization

José Mario Martínez | State University of Campinas
Unifying inexact restoration, SQP, and augmented Lagrangian methods

Mauricio G.C. Resende | AT&T Labs - Research
Using metaheuristics to solve real optimization problems in telecommunications

Nick Sahinidis | Carnegie Mellon University
Recent advances in nonconvex optimization

Stephen J. Wright | University of Wisconsin
Algorithms and applications in sparse optimization