A Direct Multisearch Approach (DMS) for Many-Objective Derivative-Free Optimization

Everton José da Silva and Ana Luísa Custódio

UI/BD/151246/2021 UIDB/00297/2020 UIDP/00297/2020

Presentation Outline

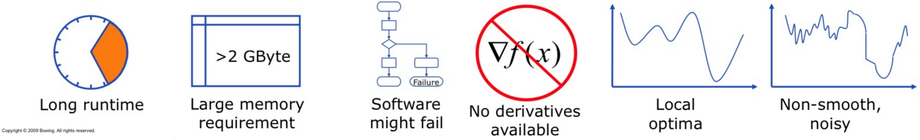
Presentation Outline

Multiobjective Derivative-free Optimization

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

$$F : \mathbb{R}^n \to \{\mathbb{R} \cup \{+\infty\}\}^m \text{ with } m \geq 2$$

- several objectives, often conflicting
- expensive function evaluation
- impossible to use or approximate derivatives

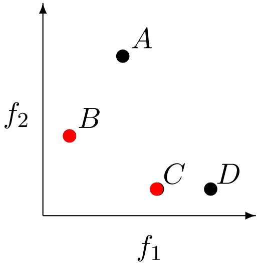


Direct MultiSearch (DMS) Main Lines

- does not aggregate any of the objective function components
- makes use of Pareto dominance

Pareto Dominance (x dominates y)

$F(x) \leq F(y)$, with $F(x) \neq F(y)$



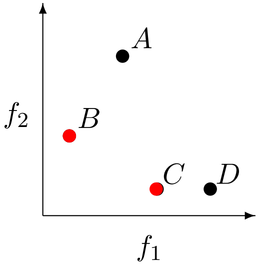
Pareto Front: $\{B, C\}$

Direct MultiSearch (DMS) Main Lines

- does not aggregate any of the objective function components
- makes use of Pareto dominance

Pareto Dominance (x dominates y)

$F(x) \leq F(y)$, with $F(x) \neq F(y)$



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

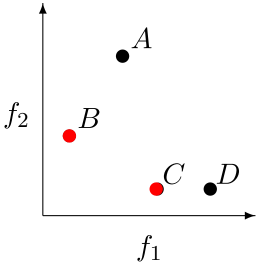
- generalizes directional direct search to MOO

Direct MultiSearch (DMS) Main Lines

- does not aggregate any of the objective function components
- makes use of Pareto dominance

Pareto Dominance (x dominates y)

$F(x) \leq F(y)$, with $F(x) \neq F(y)$

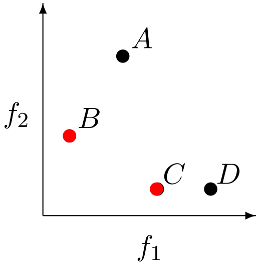


Pareto Front: $\{B, C\}$

- generalizes directional direct search to MOO
- considers the search/poll paradigm with an optional search step

Direct MultiSearch (DMS) Main Lines

- does not aggregate any of the objective function components
- makes use of Pareto dominance



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

Pareto Dominance (x dominates y)

$F(x) \leq F(y)$, with $F(x) \neq F(y)$

- generalizes directional direct search to MOO
- considers the search/poll paradigm with an optional search step
- computes approximations to the complete Pareto front

Direct MultiSearch (DMS) Main Lines

- constraints are addressed by an extreme barrier approach

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

Direct MultiSearch (DMS) Main Lines

- constraints are addressed by an extreme barrier approach

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

- keeps a list of feasible nondominated points

Direct MultiSearch (DMS) Main Lines

- constraints are addressed by an extreme barrier approach

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

- keeps a list of feasible nondominated points

- poll centers are chosen from the list

Direct MultiSearch (DMS) Main Lines

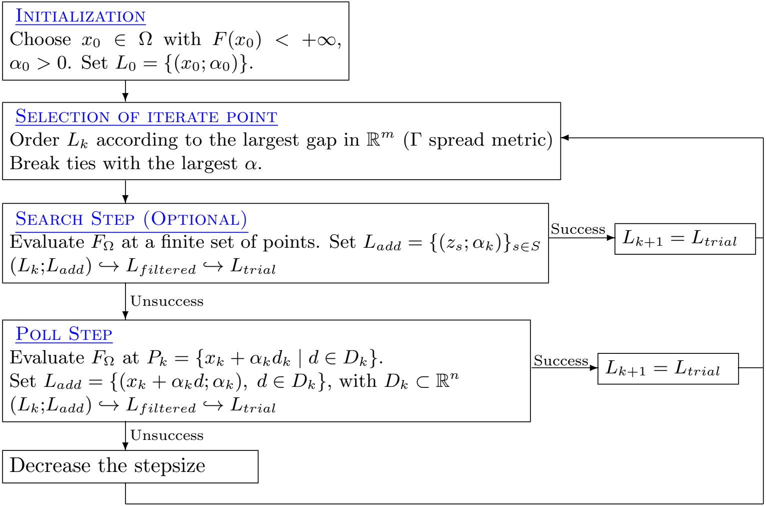
- constraints are addressed by an extreme barrier approach

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

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

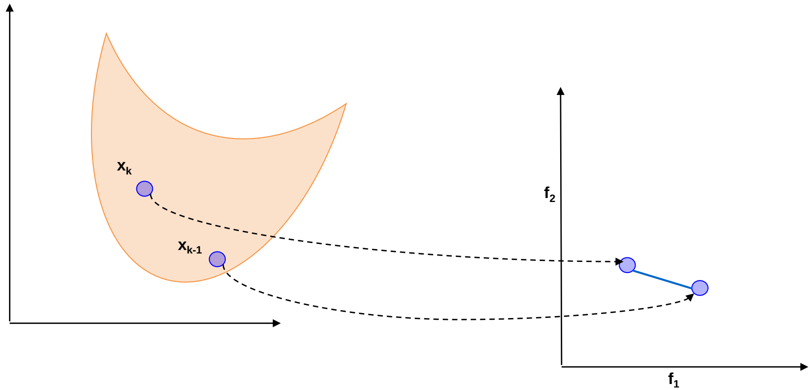
successful iteration $\Leftrightarrow$ new feasible nondominated point

Direct Multisearch - Algorithmic Structure



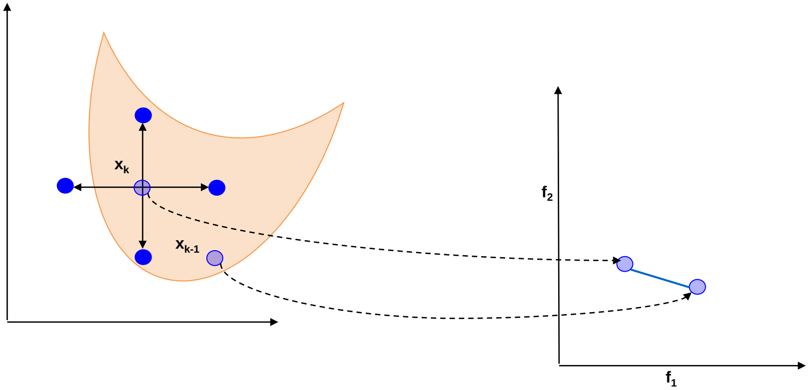
D_k **positive basis in** $\mathbb{R}^n$**;** $n + 1 \leq |D_k| \leq 2n$

Poll Step Example (Biobjective Problem)

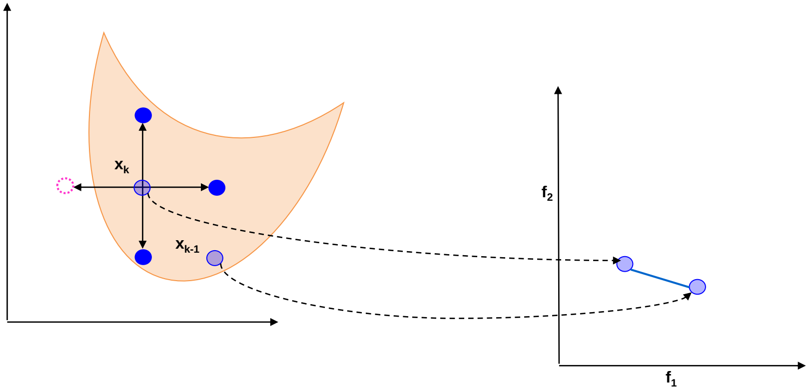


L_k

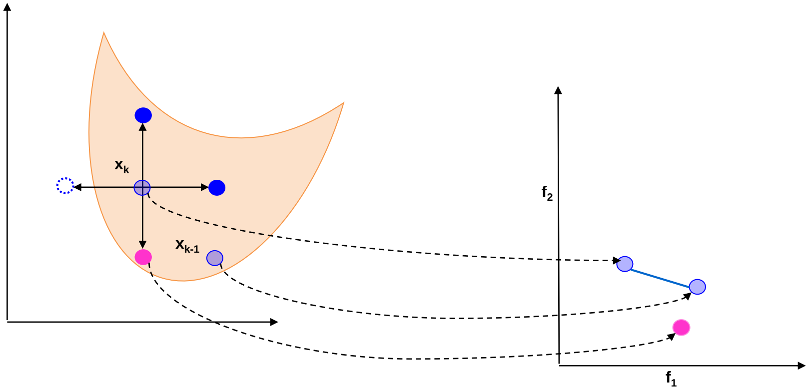
Poll Step Example (Biobjective Problem)



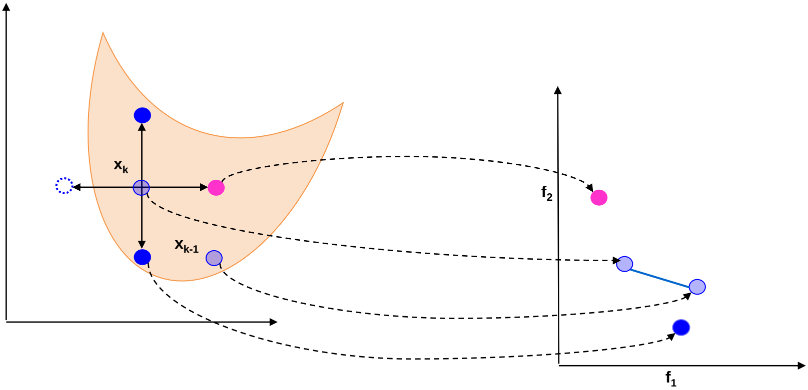
Poll Step Example (Biobjective Problem)



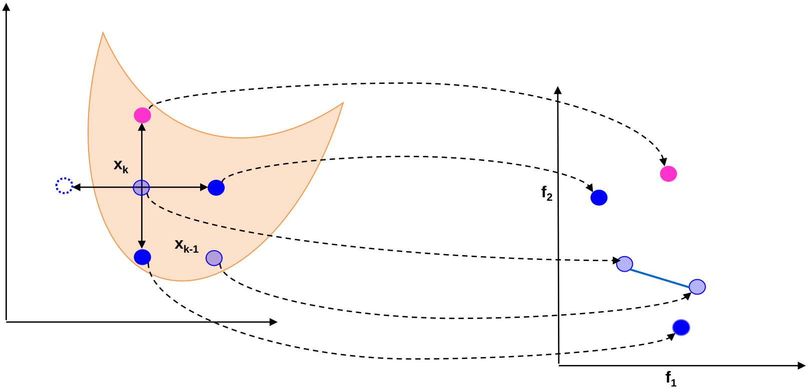
Poll Step Example (Biobjective Problem)



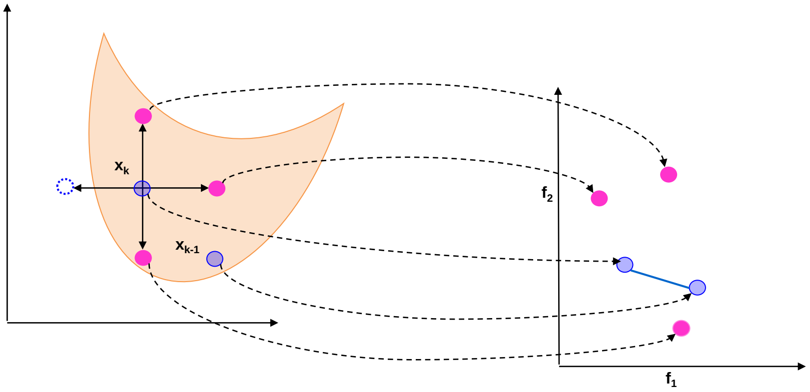
Poll Step Example (Biobjective Problem)



Poll Step Example (Biobjective Problem)

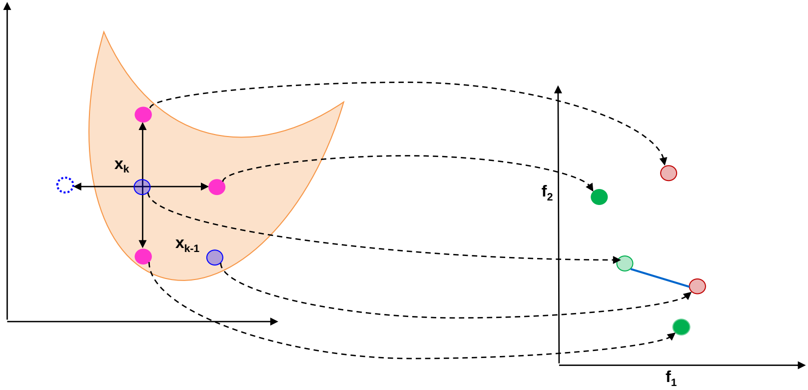


Poll Step Example (Biobjective Problem)



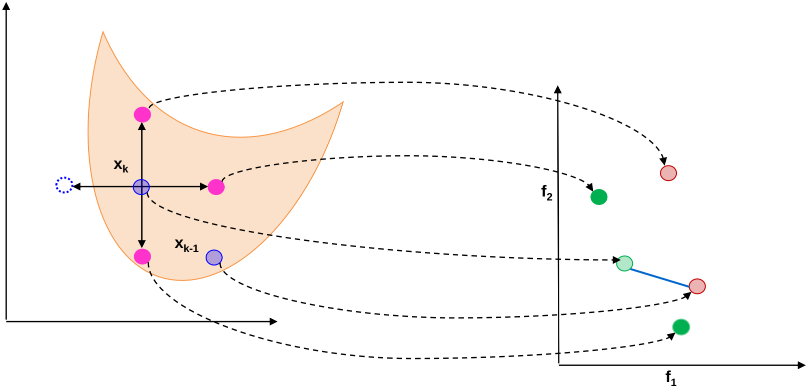
L_{add}

Poll Step Example (Biobjective Problem)



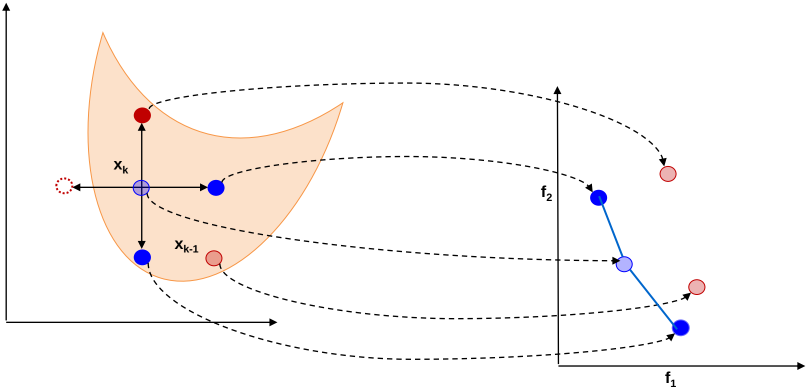
$L_{filtered}$

Poll Step Example (Biobjective Problem)



$L_{trial} = L_{filtered}$

Poll Step Example (Biobjective Problem)



L_{k+1}

What happens if the number of objectives is large? $(m \geq 4)$

- Proliferation of nondominated solutions

What happens if the number of objectives is large? ($m \geq 4$)

- Proliferation of nondominated solutions
- Challenge in keeping diversity, while progressing toward the true Pareto front

What happens if the number of objectives is large? $(m \geq 4)$

- Proliferation of nondominated solutions
- Challenge in keeping diversity, while progressing toward the true Pareto front

Many-Objective Derivative-free Optimization

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

$$F : \mathbb{R}^n \to \{\mathbb{R} \cup \{+\infty\}\}^m \text{ with } m \geq 4$$

What happens if the number of objectives is large? ($m \geq 4$)

- Proliferation of nondominated solutions
- Challenge in keeping diversity, while progressing toward the true Pareto front

Many-Objective Derivative-free Optimization

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

$$F : \mathbb{R}^n \to \{\mathbb{R} \cup \{+\infty\}\}^m \text{ with } m \geq 4$$

What happens if the problem dimension is large? (n big)

- Increase in the dimension of the positive basis

What happens if the number of objectives is large? ($m \geq 4$)

- Proliferation of nondominated solutions
- Challenge in keeping diversity, while progressing toward the true Pareto front

Many-Objective Derivative-free Optimization

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

$$F : \mathbb{R}^n \to \{\mathbb{R} \cup \{+\infty\}\}^m \text{ with } m \geq 4$$

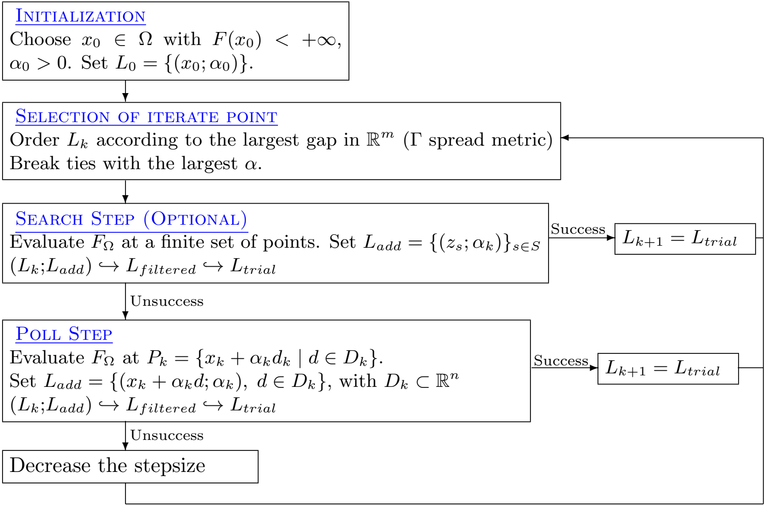
What happens if the problem dimension is large? (n big)

- Increase in the dimension of the positive basis
- Strategies to promote diversity required

Presentation Outline

1. Introduction
2. DMS-Reduction
3. Numerical Experiments
4. A Chemical Engineering Application
5. Conclusions and Future Work

DMS - Algorithmic Structure



D_k **positive basis in** $\mathbb{R}^n$**;** $n + 1 \leq |D_k| \leq 2n$

DMS-Reduction - Algorithmic Structure

Initialization

Choose $x_0 \in \Omega$, $\alpha_0 > 0$, and $r < n$.
Set $L_0 = \{(x_0; \alpha_0)\}$

Selection of iterate point

Choose $m_k \leq m$ and order L_k according to the largest gap in $\mathbb{R}^{m_k}$ (Γ spread metric in the reduced objective space). Break ties with the largest α.

Search Step (Optional)

Evaluate F_Ω at a finite set of points. Set $L_{add} = \{(z_s; \alpha_k)\}_{s \in S}$
$(L_k; L_{add}) \hookrightarrow L_{filtered} \hookrightarrow L_{trial}$

Success → $L_{k+1} = L_{trial}$

Unsuccess ↓

Poll Step

Compute an adequate matrix $P_k \in \mathbb{R}^{r \times n}$.
Evaluate F_Ω at $P_k = \{x_k + \alpha_k P_k^\top d_k \mid d \in D_k\}$.
Set $L_{add} = \{(x_k + \alpha_k P_k^\top d; \alpha_k),\ d \in D_k\}$, with $D_k \subset \mathbb{R}^r$
$(L_k; L_{add}) \hookrightarrow L_{filtered} \hookrightarrow L_{trial}$

Success → $L_{k+1} = L_{trial}$

Unsuccess ↓

Decrease the stepsize

D_k positive basis in $\mathbb{R}^r$

Objective Subspace Building

How to select the dimension of the objective subspace (m_k)?

- Use Principal Component Analysis (PCA) to dynamically select a variable number of objectives at each iteration

Objective Subspace Building

How to select the dimension of the objective subspace (m_k)?

- Use Principal Component Analysis (PCA) to dynamically select a variable number of objectives at each iteration
- Use a fixed number of objectives at each iteration $m_k = \bar{m}$

Objective Subspace Building

How to select the dimension of the objective subspace (m_k)?

- Use Principal Component Analysis (PCA) to dynamically select a variable number of objectives at each iteration
- Use a fixed number of objectives at each iteration $m_k = \bar{m}$ $(= 2)$

Objective Subspace Building

How to select the dimension of the objective subspace (m_k)?

- Use Principal Component Analysis (PCA) to dynamically select a variable number of objectives at each iteration
- Use a fixed number of objectives at each iteration $m_k = \bar{m}$ $(= 2)$

How to select the m_k objectives?

- Most conflicting
  - Determined using a correlation matrix: Spearman or Kendall

Objective Subspace Building

How to select the dimension of the objective subspace (m_k)?

- Use Principal Component Analysis (PCA) to dynamically select a variable number of objectives at each iteration
- Use a fixed number of objectives at each iteration $m_k = \bar{m}$ $(= 2)$

How to select the m_k objectives?

- Most conflicting
  - Determined using a correlation matrix: Spearman or Kendall
- Random

Objective Subspace Building

How to select the dimension of the objective subspace (m_k)?

- Use Principal Component Analysis (PCA) to dynamically select a variable number of objectives at each iteration
- Use a fixed number of objectives at each iteration $m_k = \bar{m}$ $(= 2)$

How to select the m_k objectives?

- Most conflicting
  - Determined using a correlation matrix: Spearman or Kendall
- Random
- Cyclic

Objective Subspace Building

How to select the dimension of the objective subspace (m_k)?

- Use Principal Component Analysis (PCA) to dynamically select a variable number of objectives at each iteration
- Use a fixed number of objectives at each iteration $m_k = \bar{m}$ $(= 2)$

How to select the m_k objectives?

- Most conflicting
  - Determined using a correlation matrix: Spearman or Kendall
- Random
- Cyclic

Anytime that dominance is checked, either at the search or poll steps, all the m objectives are taken into account

Variable Subspace Building

Correlation Approach

Select the subset of $r < n$ variables to keep, according to the correlation between objectives and variables. Consider $\mathcal{D}_k = [I_r \;\; -I_r]$. Compute $P_k \in \mathbb{R}^{r \times n}$ by adding $n - r$ columns of zeros to I_r, at columns corresponding to the indexes of the variables removed.

Variable Subspace Building

Correlation Approach

Select the subset of $r < n$ variables to keep, according to the correlation between objectives and variables. Consider $\mathcal{D}_k = [I_r \;\; -I_r]$. Compute $P_k \in \mathbb{R}^{r \times n}$ by adding $n - r$ columns of zeros to I_r, at columns corresponding to the indexes of the variables removed.

Example: If $n = 4$, $r = 2$, and we want to remove the first and the third variables, we get:

$$\mathcal{D}_k = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad P_k = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

then

$$P_k^\top D_k = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

Variable Subspace Building

Sketching Approach

Consider $\mathcal{D}_k = [I_r \;\; -I_r]$. Generate a sketching matrix $P_k \in \mathbb{R}^{r \times n}$, where $r < n$.

Variable Subspace Building

Sketching Approach

Consider $\mathcal{D}_k = [I_r \;\; -I_r]$. Generate a sketching matrix $P_k \in \mathbb{R}^{r \times n}$, where $r < n$.

- Hashing: P_k is a s-hashing matrix, where $s \in \mathbb{N}$ is the number of nonzero entries at randomly selected locations, each taking value $\pm 1/\sqrt{s}$
- Gaussian: P_k has entries which are independent and identically distributed in $\mathcal{N}(0, 1/r)$
- Orthogonal: $P_k = \sqrt{n/r} I_{r \times n} Q^\top$ where $I_{r \times n}$ denotes the first r rows of the $n \times n$ identity matrix I_n, and $Q \in \mathbb{R}^{n \times n}$ is the orthogonal factor in the QR decomposition $Z = QR \in \mathbb{R}^{n \times n}$ of a matrix Z with i.i.d. standard normal entries such that the diagonal entries of R are positive

L. Roberts and C. Royer
Direct Search Based on Probabilistic Descent in Reduced Spaces
SIAM J. Optim., 33 (4): 3057-3082, 2023.

Presentation Outline

Numerical Settings

- Testbed: DTLZ1-DTLZ6 and WFG1-WFG9
  - Number of components of the objective function (m) between 4 and 10
  - Number of decision variables (n) between 26 and 38
  - Final test set comprising 105 instances of bound-constrained problems
- DMS-Reduction
  - Deterministic:
    - Reduced space built using r equal to 10% of most correlated variables
    - $m = 2$ most conflicting
    - Correlation approach for variable subspace building
  - Stochastic:
    - Reduced space built using $r = 5$
    - $m = 2$ chosen randomly
    - Sketching approach for variable subspace building (1-hashing)
- Initialization DMS-Reduction*
- Stopping criterion all versions of DMS
  - $\alpha_k < 10^{-3}$ for all points in the list
  - Fixed budget of function evaluations

Initialization: Extreme Points & Middle Points

Extreme Points:

For each $\ell \in \{1, \ldots, m\}$ compute the extreme point given by

$$\overline{x}^{\ell} = \arg\min_{x \in \Omega} f_{\ell}(x)$$

Initialization: Extreme Points & Middle Points

Extreme Points:

For each $\ell \in \{1, \ldots, m\}$ compute the extreme point given by

$$\overline{x}^{\ell} = \arg\min_{x \in \Omega} f_{\ell}(x)$$

- Extreme points computed using the SID-PSM algorithm, initialized with the centroid of the box defined by bound constraints

A. L. Custódio, H. Rocha, and L. N. Vicente
Incorporating minimum Frobenius norm models in direct search
Comput. Optim. Appl., 46: 265-278, 2010.

A. L. Custódio and L. N. Vicente
Using sampling and simplex derivatives in pattern search methods
SIAM J. Optim., 18: 537-555, 2007.

Initialization: Extreme Points & Middle Points

Extreme Points:

For each $\ell \in \{1, \ldots, m\}$ compute the extreme point given by

$$\overline{x}^{\ell} = \arg\min_{x \in \Omega} f_{\ell}(x)$$

- Extreme points computed using the SID-PSM algorithm, initialized with the centroid of the box defined by bound constraints

A. L. Custódio, H. Rocha, and L. N. Vicente
Incorporating minimum Frobenius norm models in direct search
Comput. Optim. Appl., 46: 265-278, 2010.

A. L. Custódio and L. N. Vicente
Using sampling and simplex derivatives in pattern search methods
SIAM J. Optim., 18: 537-555, 2007.

- Stopping criteria: Minimum stepsize 10^{-3} or a maximum of 500 function evaluations

Initialization: Extreme Points & Middle Points

Extreme Points:

For each $\ell \in \{1, \ldots, m\}$ compute the extreme point given by

$$\overline{x}^{\ell} = \arg\min_{x \in \Omega} f_{\ell}(x)$$

- Extreme points computed using the SID-PSM algorithm, initialized with the centroid of the box defined by bound constraints

A. L. Custódio, H. Rocha, and L. N. Vicente
Incorporating minimum Frobenius norm models in direct search
Comput. Optim. Appl., 46: 265-278, 2010.

A. L. Custódio and L. N. Vicente
Using sampling and simplex derivatives in pattern search methods
SIAM J. Optim., 18: 537-555, 2007.

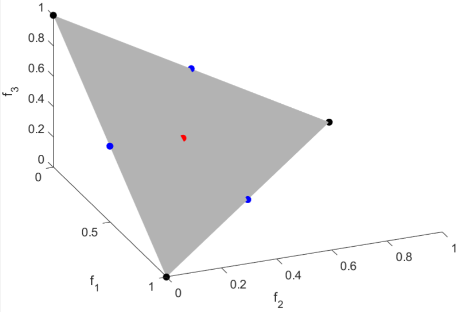
- Stopping criteria: Minimum stepsize 10^{-3} or a maximum of 500 function evaluations
- Budget Management: SID-PSM computes f_{ℓ}, not F, per iteration. Average budget for extreme points computation rounded up plus the final evaluation of F performed for each extreme point

Initialization: Extreme Points & Middle Points

Middle Points (Optional)

For each integer $\xi \in \{2, \ldots, m\}$, consider the set of combinations of ξ components of objective function (which has $\binom{m}{\xi}$ elements). Let J_ξ^ℓ be one of its elements. Compute the middle point as

$$\hat{x}^\ell = \frac{1}{\xi} \sum_{j \in J_\xi^\ell} \overline{x}^j, \quad \ell = 1, \ldots, \binom{m}{\xi}$$

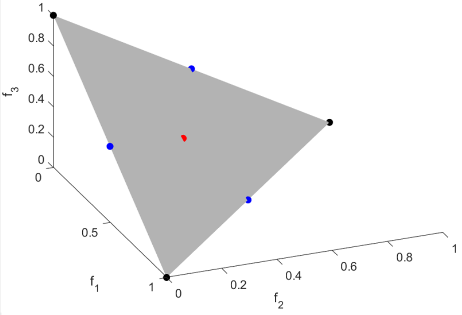


Initialization: Extreme Points & Middle Points

Middle Points (Optional)

For each integer $\xi \in \{2, \dots, m\}$, consider the set of combinations of ξ components of objective function (which has $\binom{m}{\xi}$ elements). Let J_ξ^ℓ be one of its elements. Compute the middle point as

$$\hat{x}^\ell = \frac{1}{\xi} \sum_{j \in J_\xi^\ell} \overline{x}^j, \quad \ell = 1, \dots, \binom{m}{\xi}$$



Final initialization list

$$L_0 := \left\{ (\overline{x}^\ell; \alpha_0), \ell = 1, \dots, m \right\} \cup \left\{ \bigcup_{\xi=2}^{m} \bigcup_{\ell=1}^{\binom{m}{\xi}} \{(\hat{x}^\ell; \alpha_0)\} \right\}$$

Other Solvers

- NSGA-III – Nondominated Sorting Genetic Algorithm

  K. Deb and H. Jain
  An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints
  IEEE, 18: 577-601, 2014.

- KnEA – Knee point driven Evolutionary Algorithm

  X. Zhang, Y. Tian and Y. Jin
  A knee point-driven evolutionary algorithm for many-objective optimization
  IEEE, 19: 761–776, 2015.

- MOEA/DD – Many-objective evolutionary algorithm based on dominance and decomposition

  K. Li, K. Deb, Q. Zhang and S. Kwong
  An evolutionary many-objective optimization algorithm based on dominance and decomposition
  IEEE, 19: 694–716, 2015.

- MOMBI-II – Many-objective metaheuristic based on the R2 indicator II

  R. H. Gómez and C. A. Coello Coello
  Improved metaheuristic based on the R2 indicator for many-objective optimization
  GECCO, 7:679–686, 2015.

- GrEA – Grid-Based Evolutionary Algorithm

  S. Yang, M. Li, X. Liu and J. Zheng
  A grid-based evolutionary algorithm for many-objective optimization
  IEEE, 17: 721–736, 2013.

*Implementations in the MATLAB-based platform `PlatEMO`, freely available at `https://github.com/BIMK/PlatEMO`

Metrics for Performance Profiles (Dolan and Moré [2002])

- Purity

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

- Spreads Γ and Δ

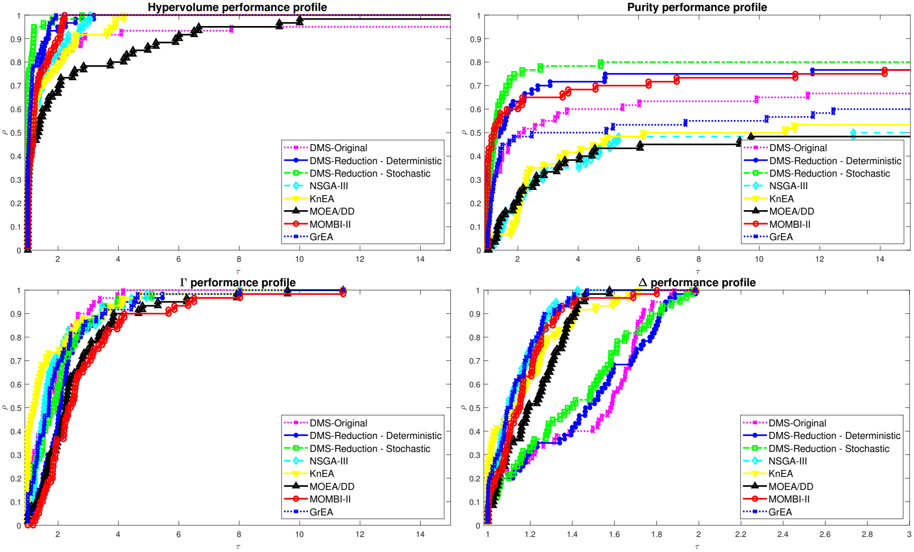
$$\Gamma_{p,s} = \max_{j \in \{1,\ldots,m\}} \left(\max_{i \in \{0,\ldots,N\}} \{d_i\} \right)$$

$$\Delta = \max_{j \in \{1,\ldots,m\}} \left(\frac{d_0 + d_N + \sum_{i=1}^{N-1} |d_i - \bar{d}|}{d_0 + d_N + (N-1)\bar{d}} \right)$$

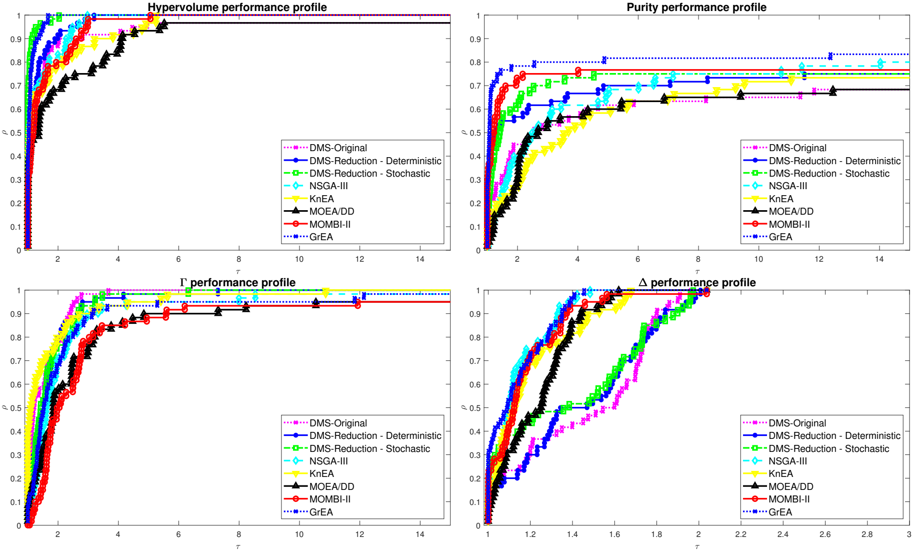
- Hypervolume

$$HI_{p,s} = Vol\{b \in \mathbb{R}^m \mid b \leq U_p \wedge \exists a \in F_{p,s} : a \leq b\}$$

5k function evaluations

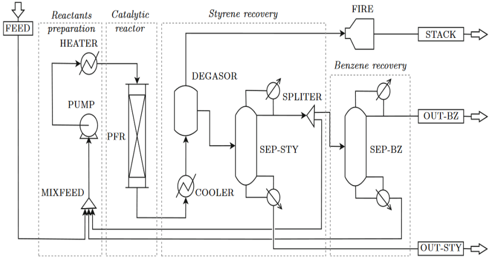


20k function evaluations



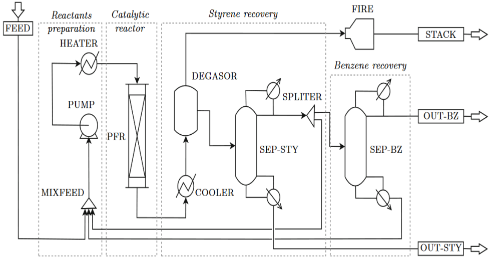
Presentation Outline

A Chemical Engineering Application - Styrene Production



- Triobjective problem
  - Net present value ($)
  - Purity of styrene
  - Overall ethylbenzene conversion into styrene

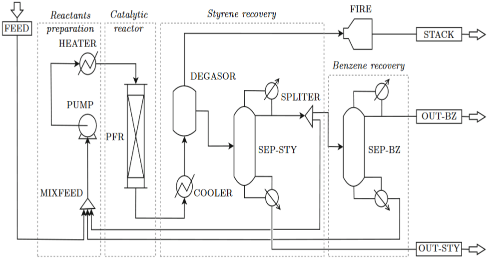
A Chemical Engineering Application - Styrene Production



- Triobjective problem
  - Net present value ($)
  - Purity of styrene
  - Overall ethylbenzene conversion into styrene

- Blackbox simulation ($\simeq$ 1s/eval)

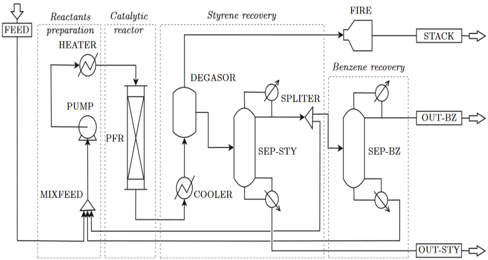
A Chemical Engineering Application - Styrene Production



- Triobjective problem
  - Net present value ($)
  - Purity of styrene
  - Overall ethylbenzene conversion into styrene

- Blackbox simulation ($\simeq$ 1s/eval)
- 8 variables subject to bounds - normalized to [0,100] intervals

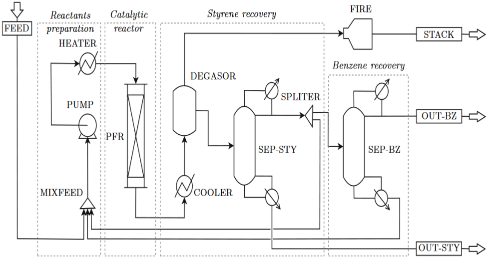
A Chemical Engineering Application - Styrene Production



- Triobjective problem
  - Net present value ($)
  - Purity of styrene
  - Overall ethylbenzene conversion into styrene

- Blackbox simulation ($\simeq$ 1s/eval)
- 8 variables subject to bounds - normalized to [0,100] intervals
- 9 constraints related with industrial and environmental regulations, treated as hidden constraints

A Chemical Engineering Application - Styrene Production

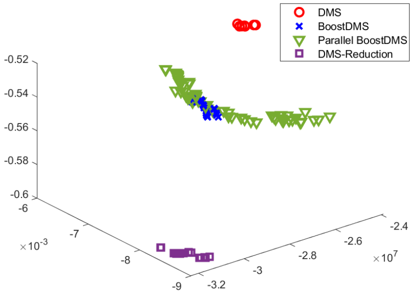


- Triobjective problem
  - Net present value ($)
  - Purity of styrene
  - Overall ethylbenzene conversion into styrene

- Blackbox simulation ($\simeq$ 1s/eval)
- 8 variables subject to bounds - normalized to [0,100] intervals
- 9 constraints related with industrial and environmental regulations, treated as hidden constraints
- Initialization at a single point provided in literature

C. Audet, V. Béchard, and S. Le Digabel
Nonsmooth optimization through mesh adaptive direct search and variable neighborhood search
J. Glob. Optim., 41: 299-318,2008.

Computed Pareto Fronts: All DMS versions



- **DMS** Pareto front is dominated by both **BoostDMS** and **Parallel BoostDMS**,which are dominated by **DMS-Reduction**

Solver	#Points
DMS	7
BoostDMS	26
Parallel BoostDMS	66
DMS-Reduction	16

Solver	Purity	Hypervolume	Gamma	Delta	# Function Evaluations
DMS	0%	0.0001	7.96e+6	1.0517	6741
BoostDMS	0%	0.0989	4.34e+6	1.1228	20000
Parallel BoostDMS	0%	0.1321	3.32e+6	1.2266	20000
DMS-Reduction	100%	0.9881	7.27e+6	1.1227	7995

Solvers Able to Handle Constraints

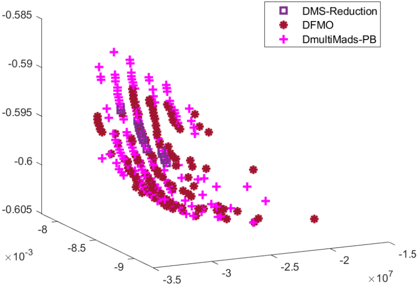
- DFMO

  G. Liuzzi, S. Lucidi and F. Rinaldi
  A derivative-free approach to constrained multiobjective nonsmooth optimization
  SIAM J. Optim., 26: 2744–2774, 2016.

- DMultiMADS-PB

  J. Bigeon, S. Le Digabel and L. Salomon
  Handling of constraints in multiobjective blackbox optimization
  Comput. Optim. Appl., 89: 69–113, 2024.

Computed Pareto Fronts



Solver	#Points
DFMO	122
DMultiMADS-PB	149
DMS-Reduction	16

Solver	Purity	Hypervolume	Gamma	Delta	# Function Evaluations
DMS-Reduction	87.5%	0.3675	13.8e+6	1.2073	7995
DFMO	46.7%	0.7033	5.11e+6	1.6886	20000
DMultiMADS-PB	58.4%	0.7149	5.20e+6	1.7989	20000

Presentation Outline

Conclusions and Future Work

Conclusions

- DMS-Reduction introduces a novel approach that leverages search directions from variable subspaces
- It is tailored to effectively address challenges in many-objective scenarios
- The approach simultaneously reduces the number of objective function components and problem variables, thereby tackling large-scale challenges
- Extensive numerical experiments demonstrate the method's competitiveness and potential

Conclusions and Future Work

Conclusions

- DMS-Reduction introduces a novel approach that leverages search directions from variable subspaces
- It is tailored to effectively address challenges in many-objective scenarios
- The approach simultaneously reduces the number of objective function components and problem variables, thereby tackling large-scale challenges
- Extensive numerical experiments demonstrate the method's competitiveness and potential

Future work

- Developing theoretical analysis for DMS-Reduction
- Implementation of a filter-inexact restoration approach to handle constraints

Thank you for your attention!