Parameter Calibration Algorithms

Xin HUANG

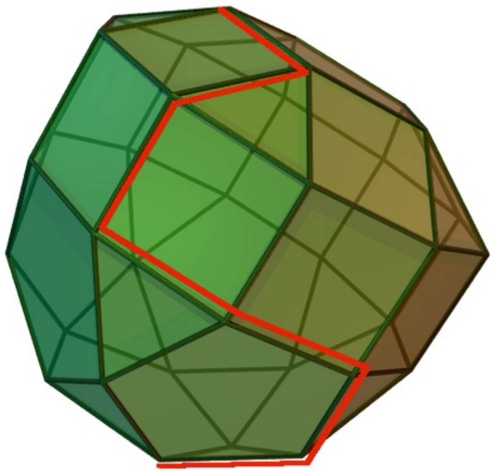
Jun 2015

Outline

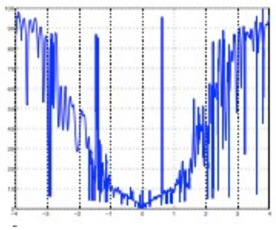
- Single-Objective Optimization
  - Simplex Downhill Method (local optimization)
  - CMA-ES
  - SCE-UA
  - PSO
  - DE
  - EGO
- Multiple-Objective Optimization
- Benchmark Functions
  - DTLZ,WGF,BBOB
- Benchmark Models
  - Licom2 : LASG/IAP Climate System Ocean Model
  - GAMIL : Grid-point Atmospheric Model
  - CLM : Community Land Model

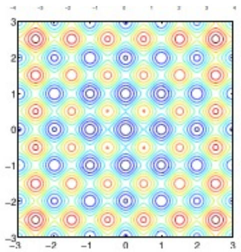
Simplex Downhill Method

- Theory
  - Reflection $\mathbf{x}_r = \mathbf{x}_o + \alpha(\mathbf{x}_o - \mathbf{x}_{n+1})$
  - Expansion $\mathbf{x}_e = \mathbf{x}_o + \gamma(\mathbf{x}_o - \mathbf{x}_{n+1})$
  - Contraction $\mathbf{x}_c = \mathbf{x}_o + \rho(\mathbf{x}_o - \mathbf{x}_{n+1})$
- Local Search
  - Sensitive to initial samples
  - Fast convergence speed
  - Multi start simplex-downhill



CMA-ES

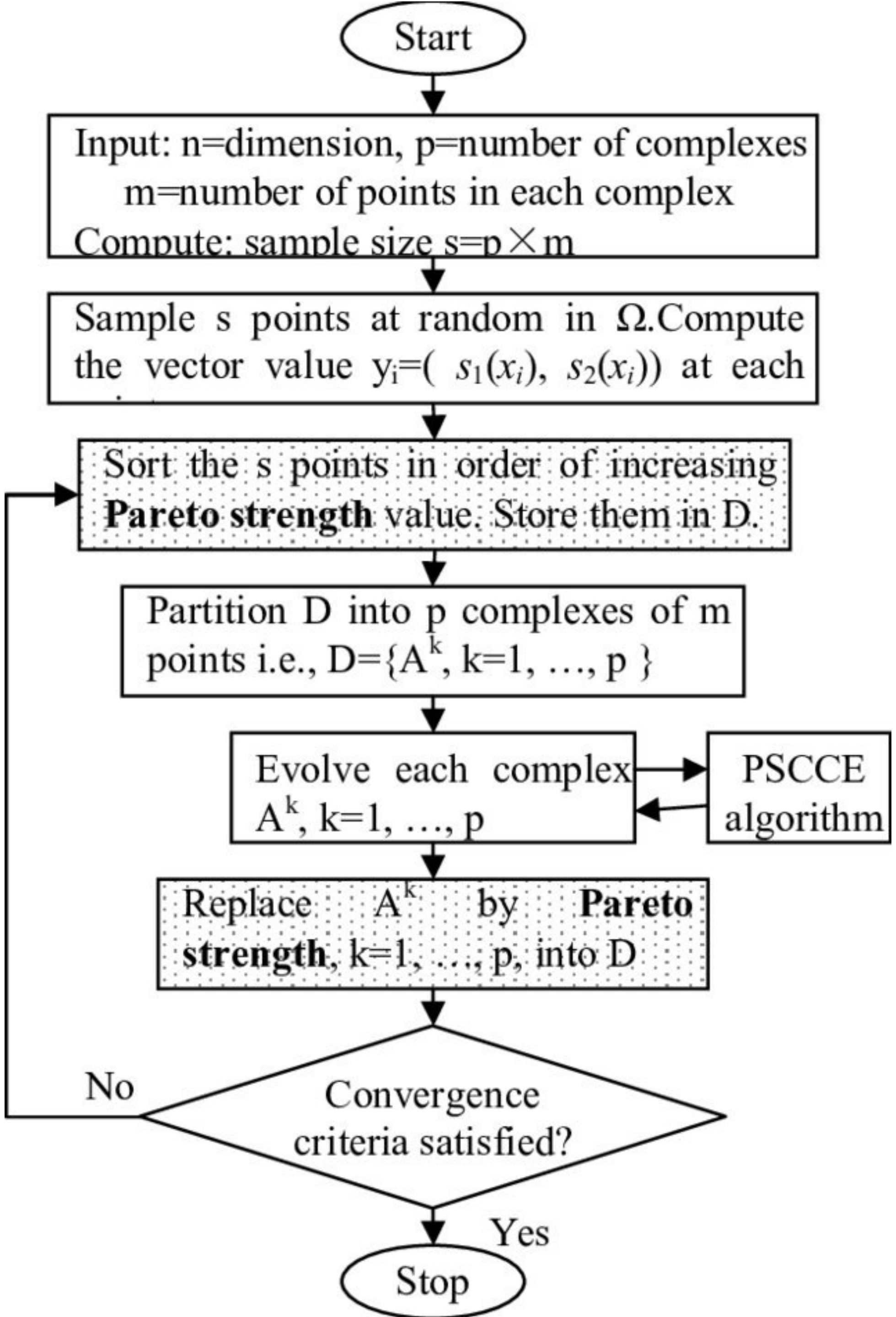




- Theory
  - Use Gaussian Distribution to generate new solutions
  - Update and improve covariance matrix
  - Strong ability on non-linear, multi-modal and rough model
- Advantages
  - Stable, not sensitive with initial points
  - Easy to Parallel
  - Relatively High Efficiency (success in 50 iterations, 50*12 samples)

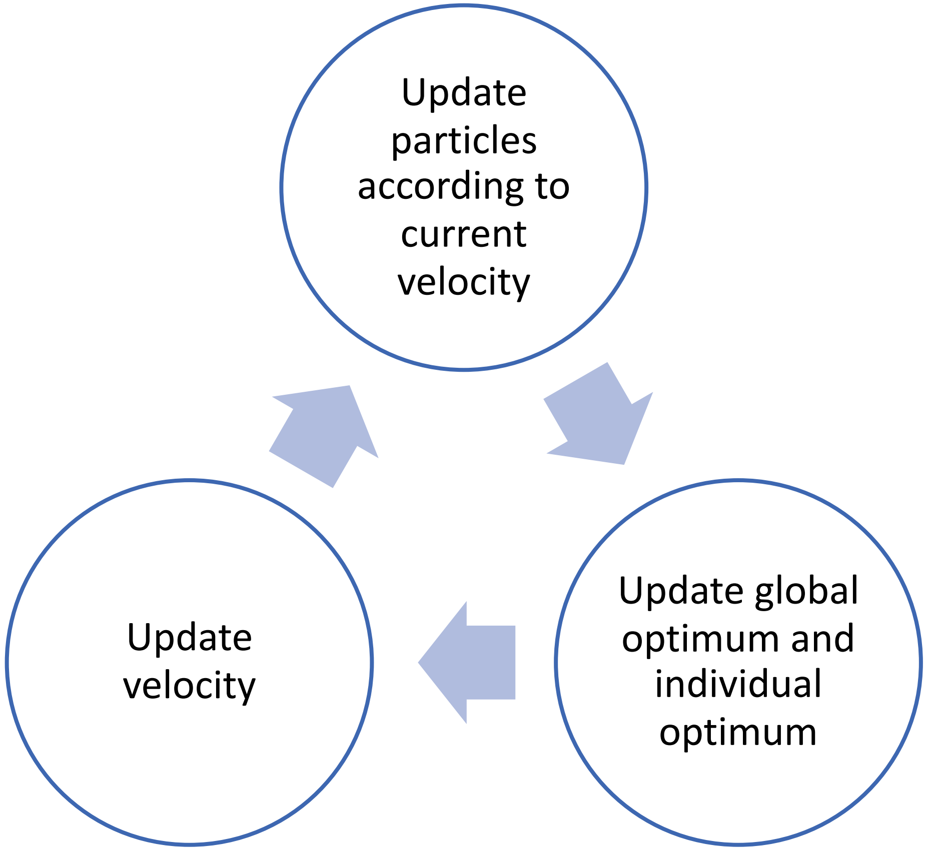
Shuffled Complex Evolution method (SCE-UA)

- Combines the advantages of stochasticity, simplicity and genetic algorithms
- This algorithm is widely applied in hydrological models



Flowchart of SCE-UA algorithm combing Pareto strength technique (ICNC '08)

Particle Swarm Optimization (PSO)



Advantages

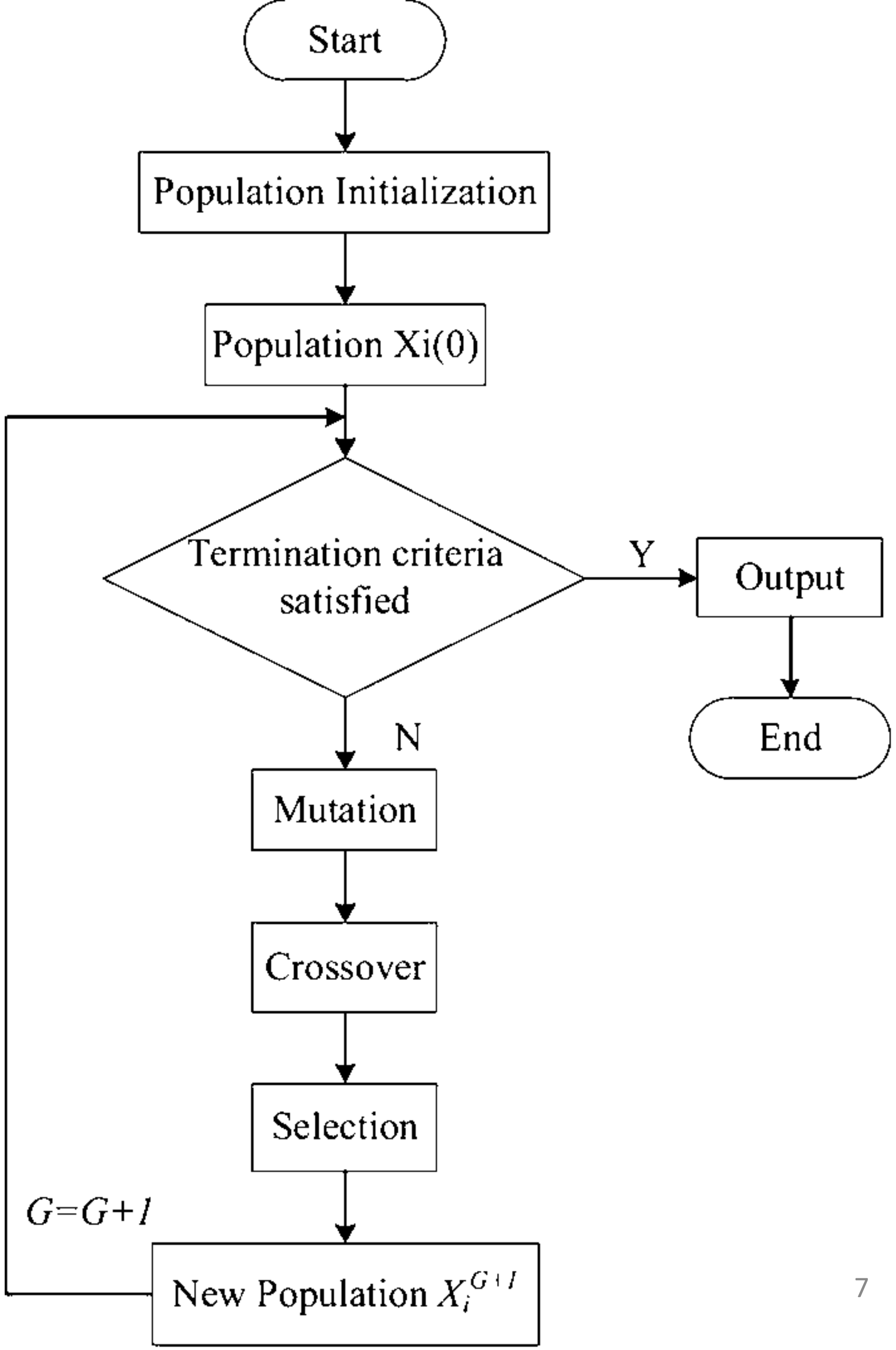
- Insensitive to scaling of design variables
- Easily parallelized for concurrent processing
- Derivative free
- Very few algorithm parameters
- A very efficient global search algorithm

Disadvantages

- PSO’s optimum local searchability is weak

Differential Evolution (DE)

- Robust, stable, make few or no assumptions about the optimization problems
- GA is more suitable for discrete optimization, PSO and DE are more natural for continuous optimization.



Flowchart of DE algorithm

Efficient Global Optimization (EGO)

- Frame
  - Initial Sample (11*D-1)
  - Build Surrogate Model and find the point with max EI

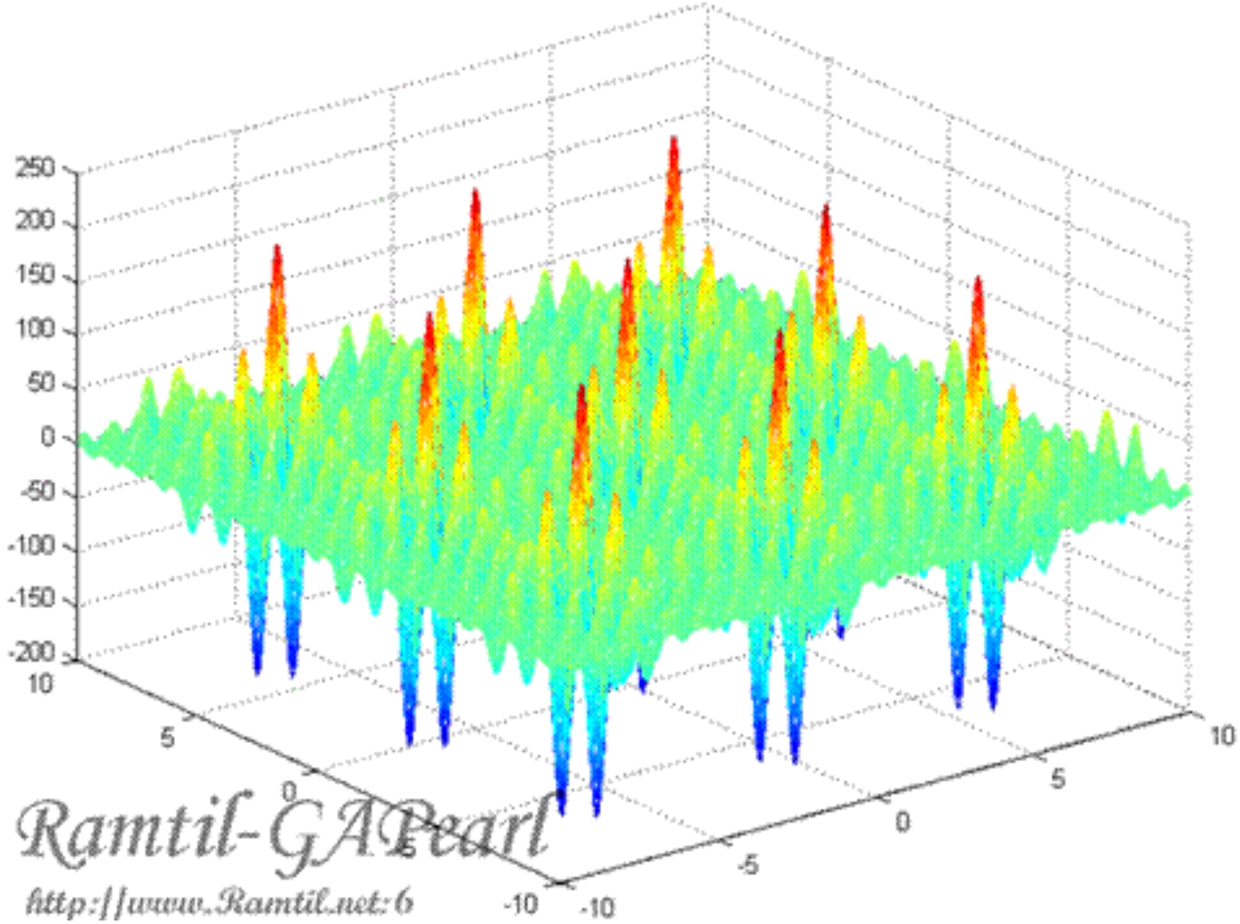
$$E[I(x)] = (f_{\min} - \hat{y})\Phi(\frac{f_{\min} - \hat{y}}{\hat{s}}) + s\Psi(\frac{f_{\min} - \hat{y}}{\hat{s}})$$

  - Update Model and Iteration
- No parallelization

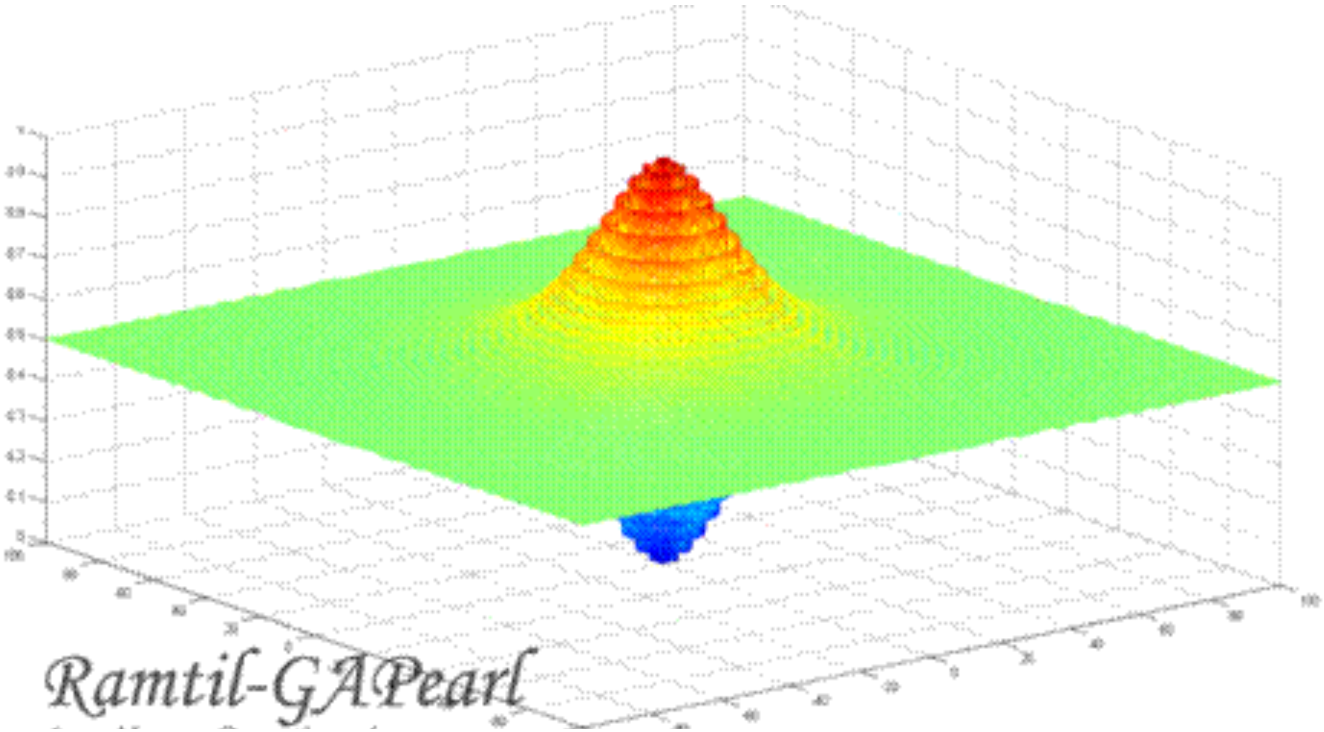
Comparison of different algorithms

- Benchmark functions
  - GoldStein_Price (many local optima, 1 global optima)
  - Schaffer1 (many local optima, 1 global optima, non-smooth)
  - Six_hump (6 local optima, 1 global optima)
  - Shubert (760 local optima, 18 global optima)
- Metrices for comparisons
  - The optima achieved
  - The minimum distance from the optimal solution and real solution
  - Percentage of good solutions (a good solution is a point whose distance to the real solution is within 1% distance)

$$\text{Shubert}\begin{cases} \min f(x_1, x_2) = \sum_{i=1}^{5} i\cos[(i+1)x_1 + i] \ * \\ \sum_{i=1}^{5} i\cos[(i+1)x_2 + i] \\ -10 \le x_i \le 10 (i = 1,2) \end{cases}$$



Schaffer1 .
$$\begin{cases} \min f(x_1, x_2) = 0.5 + \dfrac{\sin^2\sqrt{x_1^2+x_2^2} - 0.5}{[1+0.001(x_1^2+x_2^2)]^2} \\ -100 \le x_i \le 100 (i = 1,2) \end{cases}$$



Optimal solutions after 50 iterations

Better!

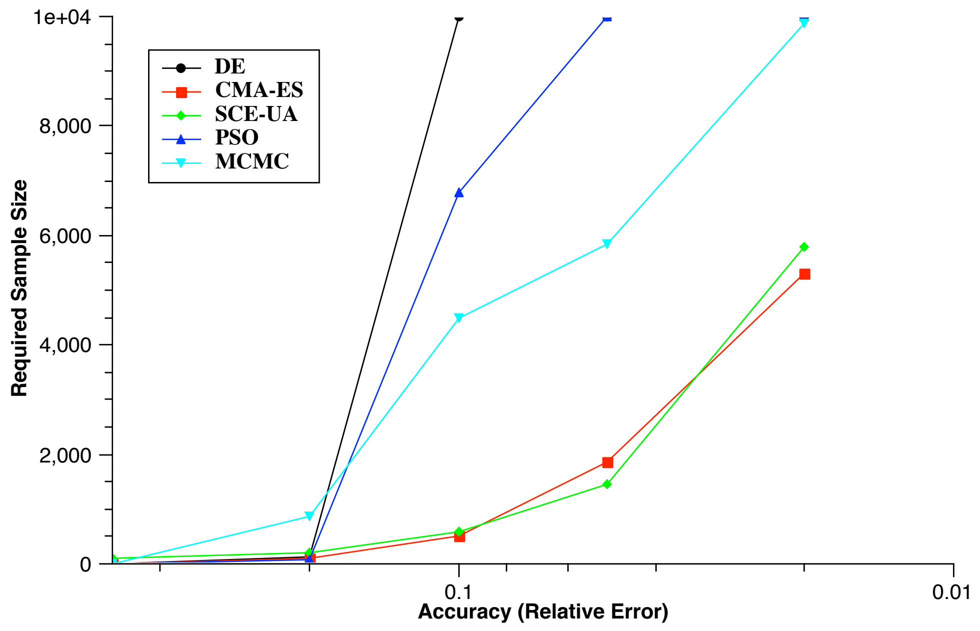
	DE	SCE-UA	PSO
Schaffer1	0.033483	0.016338	0.001333
Schaffer2	0.01584	0.02455	0.0205
Shubert	10.99433	27.99533	7.119333
Six_hump	0.000448	0.00347	2E-6

Minimum distance to the real solution

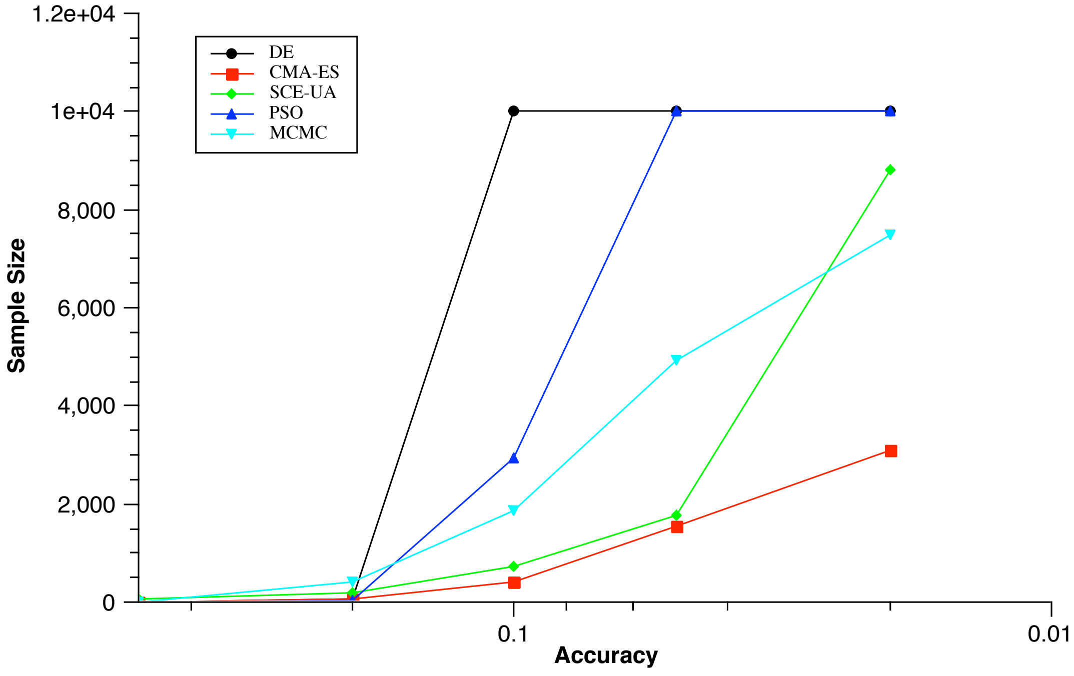
Better!

	DE	SCE-UA	PSO
Golden_Stein Price	0.0117	0.053911	0.00088
Schaffer1	2.238333	2.432518	1.845
Six_hump	0.0107	0.056155	0.0014

Comparison of optimization algorithms on CLM-CASA' model



$$J = \sum \frac{(z_{i,j} - x_{i,j})^2}{2\sigma_{i,j}^2}$$



$$J' = \frac{\sum (z_{i,j} - x_{i,j})^2}{var(X)}$$

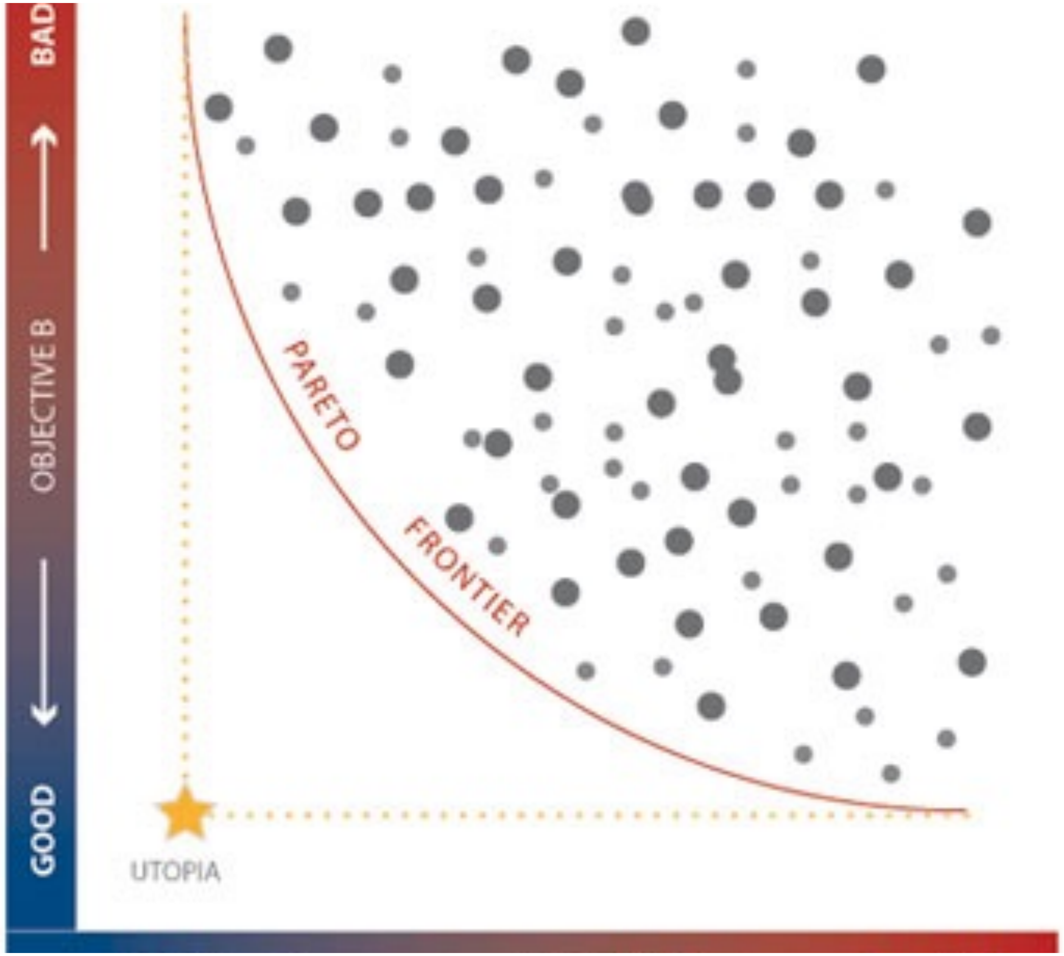
Multiple-Objective Optimization

Pareto Frontier

$$\min \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))^{\mathrm{T}}$$

$$\text{subject to } \mathbf{x} \in \Omega \subseteq \mathbb{R}^n$$

- The Target of Multi Objective Optimization is to find Pareto Frontier
- A common approach is to weight different objective and then transform it into a single-objective optimization problem
- Multiple-Objective algorithms



Convert Multi-Objective to Single-Objective Optimization problem

- In GAMIL2 Tuning, 16 objectives to 1 metrics

U200	U850
V200	V850
T200	T850
Z500	PRECT
Q400	Q850
FLUT	FLUTC
FSNTOA	FSNTOAC
LWCF	SWCF

$$(\sigma_m^F)^2 = \sum_{i=1}^{I} w(i)\left(x_m^F(i) - x_o^F(i)\right)^2 \quad (1)$$

$$(\sigma_r^F)^2 = \sum_{i=1}^{I} w(i)\left(x_r^F(i) - x_o^F(i)\right)^2 \quad (2)$$

$$\chi^2 = \frac{1}{N^F}\sum_{F=1}^{N^F}\left(\frac{\sigma_m^F}{\sigma_r^F}\right)^2 \quad (3)$$

Multiple-Objective Optimization (MOO) Algorithms

- NSGA
  - Genetic Algorithms
  - Based on un-dominated sort
  - NSGA-iii : use reference point
  - Population size : 10~1000
- MOEA/D
  - Using clustering functions to transformer problem to some single problems
  - Based on single objective optimization (e.g., PSO, SA, DE, GA)
  - Population size: 100~10000 (subproblem size)
- EFR
  - Multi Objectives means Multi rankings
  - Average Ranking

NSGA-II/iii

- 遗传算法应用+非支配排序：
  - 在选择个体时，NSGA-II使用非支配排序将2N（子代+父代）个解分为F1，F2，F3，…层，最终需要在某一层中选取一些解，此时使用拥挤度排序，拥挤度低的将优先被选择
  - NSGA-III改进了这一策略，使用一些参考点辅助选择个体
  - Advance:
    - 子种群分割保持收敛性

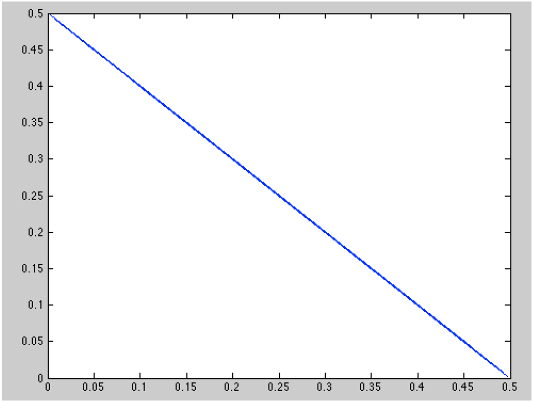
MOEA/D

- 基于分解的多目标优化
  - 讲多目标问题通过聚合函数（线性加权，切比雪夫函数）组合为N个单目标子问题，在问题优化过程中同时优化这些问题。
  - 聚合函数：$g_j^{te}(\mathbf{x}|\boldsymbol{\lambda}_j, \mathbf{z}^*) = \max_{k=1}^{m}\{\lambda_{j,k}|f_k(\mathbf{x}) - z_k^*|\}$，不同的子问题有不同的权向量λ，不同子问题的相似程度通过权向量描述。
  - MOEA/D维护一个拥有N个个体的种群，每个个体代表一个子问题的当前最优解，进化时通过在相似子问题的解中选取父母生成子代个体，并替换参考点和个体。
  - Advance:
    - 聚合函数的研究（可变权向量，新型权值函数）
    - 个体选择方法（小生境模式）

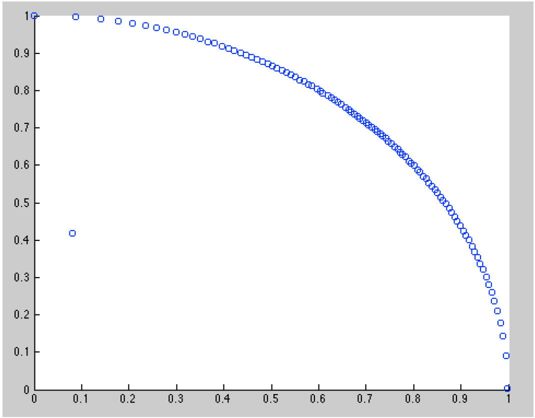
EFR

- EFR整体框架和NSGA相似，但是在选择个体时采用如下策略
  - EFR使用一系列评估函数,F1,F2,…FN，针对每个评估函数，每个解讲获得一个rank
  - N个评估函数，因此每个解有N个rank值
  - 可以使用平均排序，最大排序和层次排序得到每个解得最终rank值，选择个体

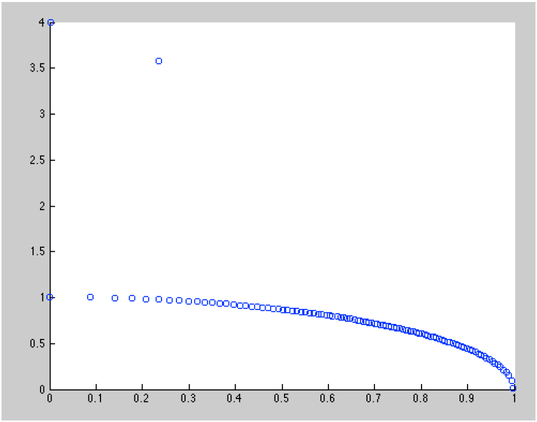
Benchmark Functions: DTLZ



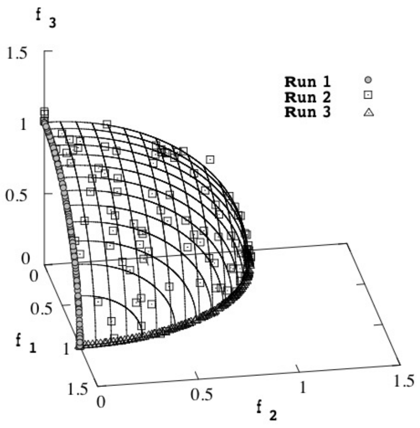
DTLZ1



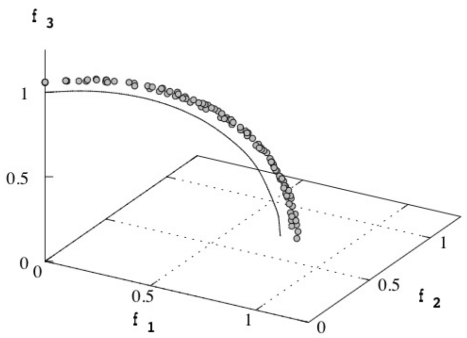
DTLZ2



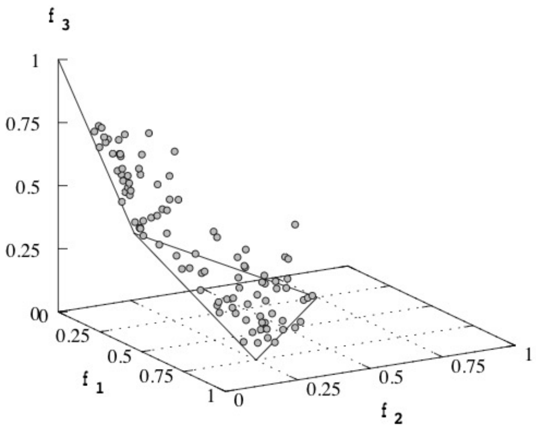
DTLZ3



DTLZ4



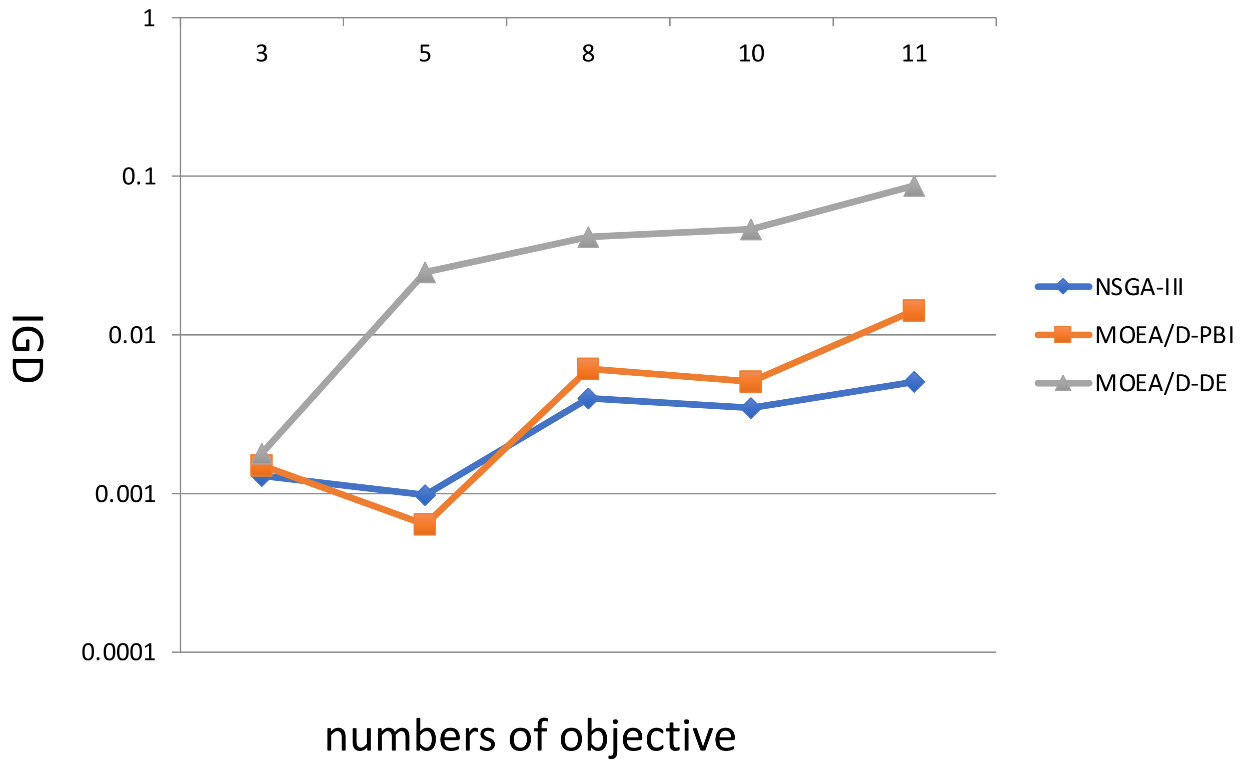
DTLZ5



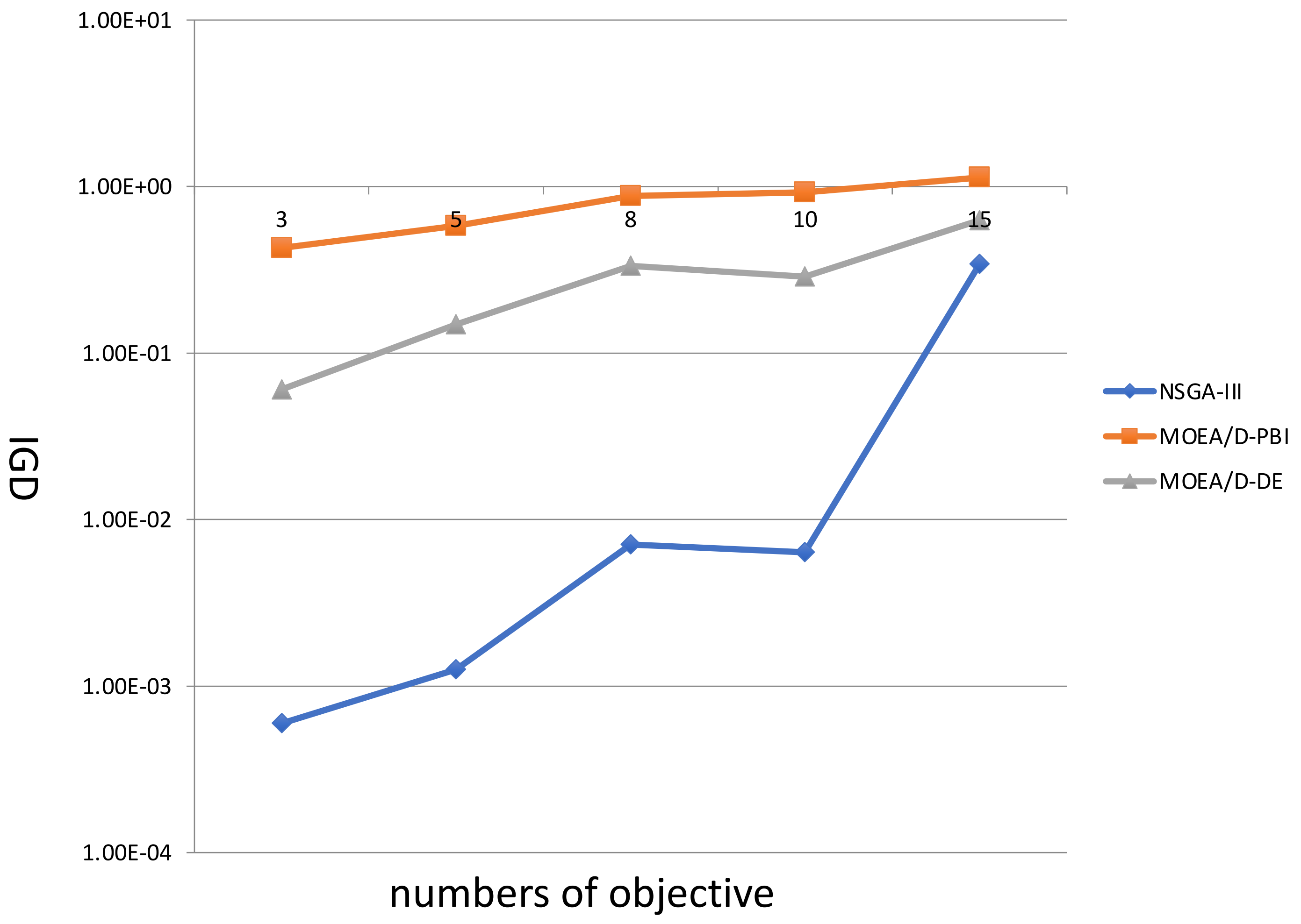
DTLZ7

Performances of different MOO algorithms on DTLZ1

- x : numbers of objective, y : Inverted Generational Distance (IGD)



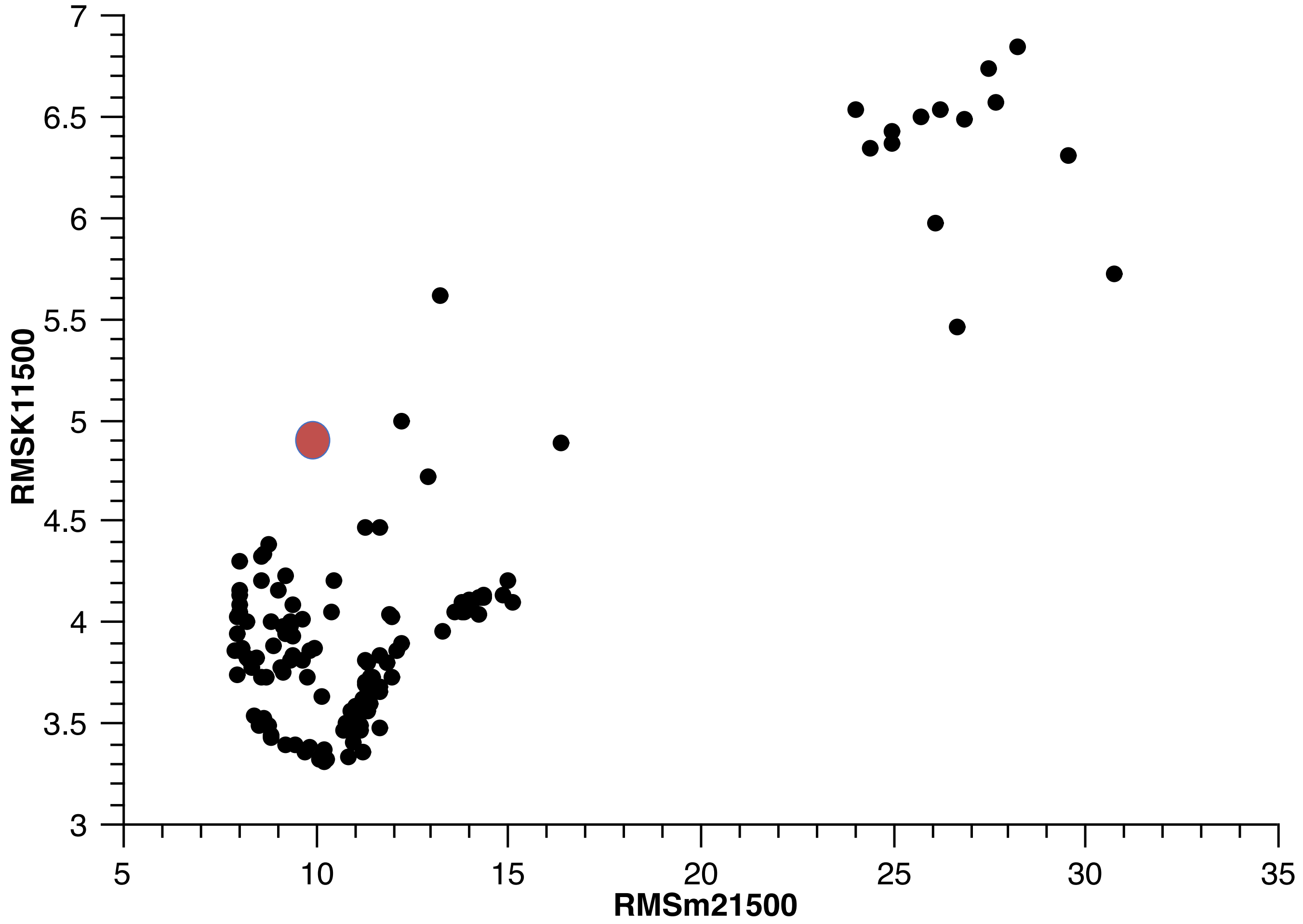
Performances of different MOO algorithms on DTLZ1



Experiments on Licom2

- Licom2 Ocean Model
- 4 input parameters
- Two Objects:
  - RMS_K1_1500 (default: 4.764616)
  - RMS_M2_1500 (default: 11.59034)

Comparison with Single-Objective Optimization



Default Res:
RmsM2 = 11.59034
RmsK1 =4.764616

Minimum M2:
RmsM2 = 7.878525
RmsK1 = 3.856853

Minimum K1:
RmsM2 = 10.22993
RmsK1 = 3.315709

Challenge

- High Dimension Output
  - Licom2 has only 2 objectives, so it’s very easy. GAMIL2 has 16 objects.
- Reduce the sample points
  - Surrogate Model?
- Multi Objective Sensitivity Analysis
- Interaction among different objects

Dimension reduction for multiple-objective optimization

$\delta - dominance$

$$\preceq^{\varepsilon}_{\mathcal{F}'} := \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in A \land \forall i \in \mathcal{F}' : f_i(\mathbf{x}) - \varepsilon \leq f_i(\mathbf{y})\}$$

Accept δ error – dominate

Measure the conflict degree of one objective subset and whole set

e.g. : (f1,f2,f4) and (f1,f2,f3,f4) conflict : 0

(f2,f4) and (f1,f2,f3,f4)

conflict :0.5

Objective Reduction of GAMIL

- Data : 907 sample points , 16 objectives
- {4,13} : 0.0347   {12,15} 0.0447
- {8,11} : 0.169   {4,10,13}:0.178 {2 3} 0.186
- {4,10,13,14}:0.232 {5,7} :0.244
- {1,6} 0.349   {2,3,5,7} 0.571
- {0,1,6} 0.616 {2,3,5,7,8,11} 0.922

Thanks