Data-Driven Evolutionary Optimization of Complex Systems: Big vs Small Data

Yaochu Jin

Head, Nature Inspired Computing and Engineering (NICE)

Department of Computer Science, University of Surrey, United Kingdom

yaochu.jin@surrey.ac.uk

Outline

- Complexity in evolutionary optimization of real-world problems
- Data driven evolutionary optimization
  - Offline (Big) data driven evolutionary optimization
  - Online small data driven evolutionary optimization
- Concluding remarks

Strengths and Weaknesses of EAs for Optimization

UNIVERSITY OF SURREY

- No need for analytical objective functions and no requirement for derivative information
- Less vulnerable to local optimums
- Less vulnerable to uncertainty (relative quality is more important)
- Well suited for multi-objective optimization
- No theoretical guarantee for global optimum
- Population based search -- computationally intensive

Complexities in Real-World Optimization

- Complexity in problem formulation and solution representation

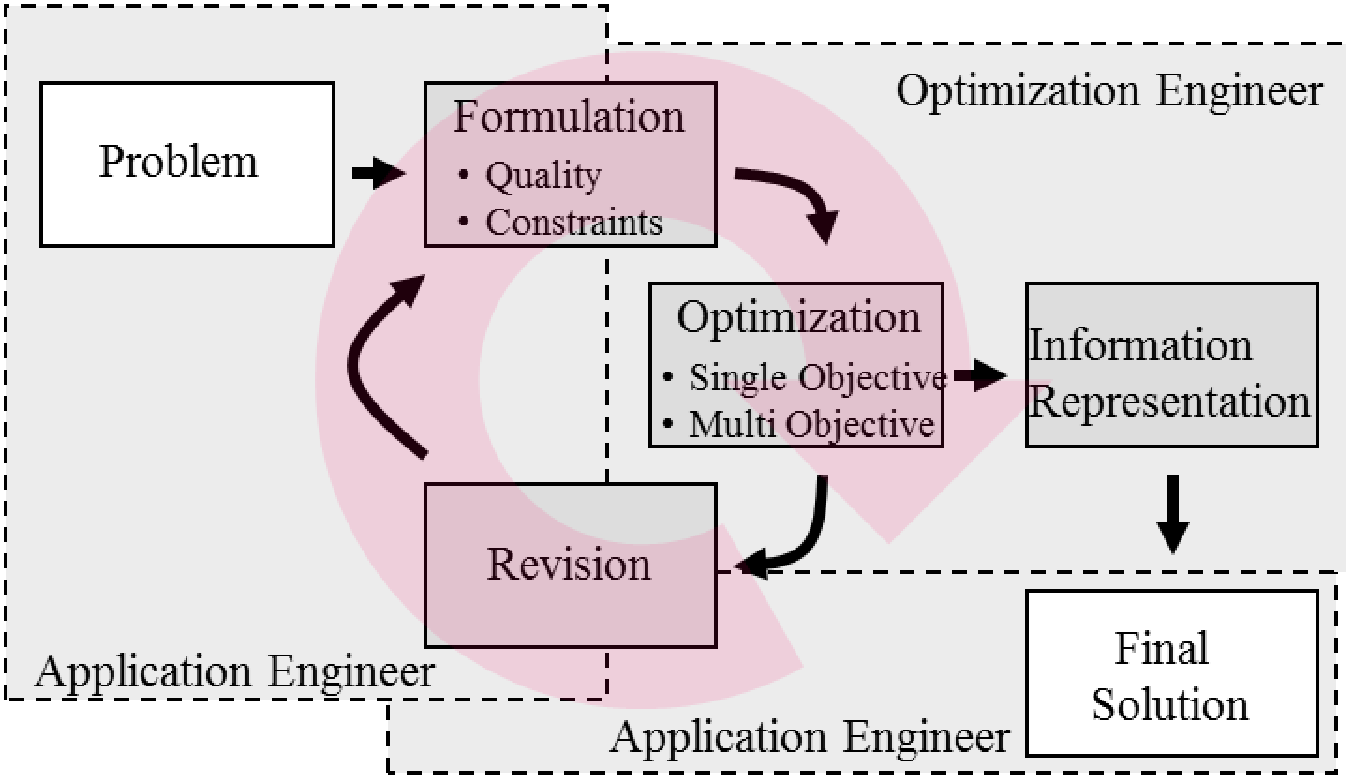
- Complexity in scale
  - Large number of decision variables
  - Large number of objectives

- Complexity in handling uncertainty

- Complexity in quality evaluation

Complexity in Problem Formulation and Solution Representation

Complexity in Problem Formulation

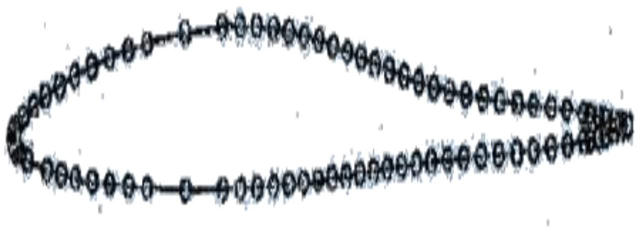


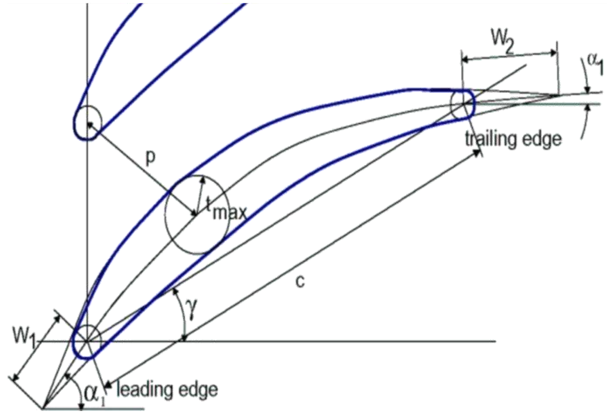
- The formulation of the objective function is an iterative process
- Representation is critical : multiple sub-systems, optimization-control coupling
- Different objectives/ constraints / decision variables may have to be considered / weighted differently at different stages
- Different resources are available at different stages

Complexity in Shape Representation

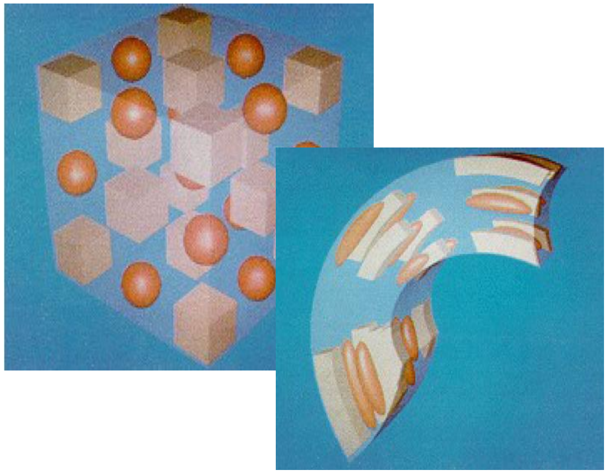
Expert Knowledge

Direct representation

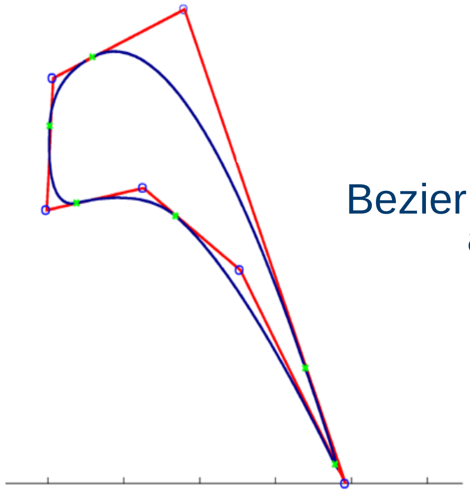




Free Form Deformation(FFD)

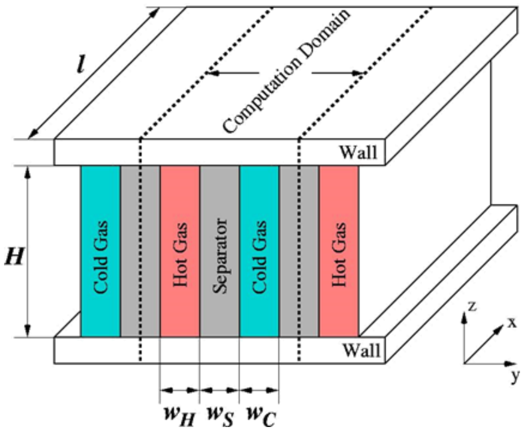


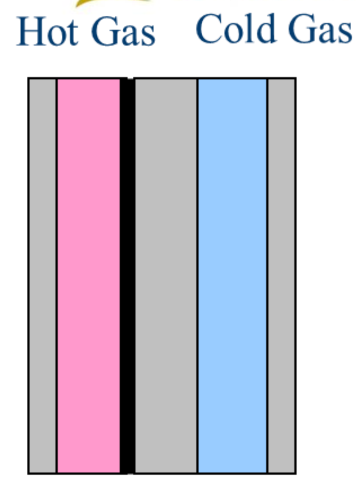
Bezier Curves, B-Splines
and NURBS

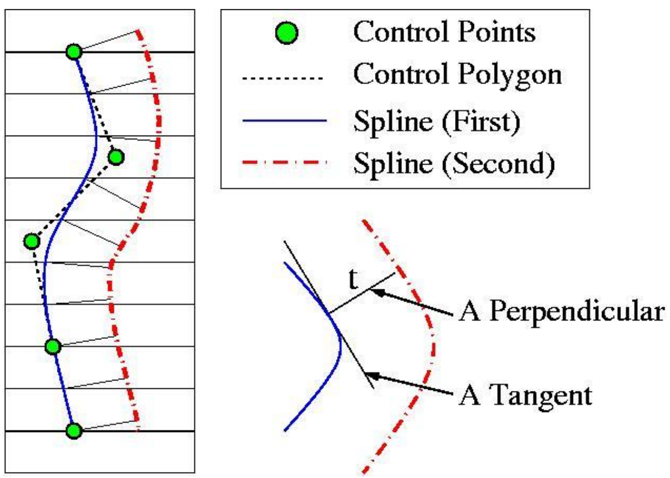


Shape Representations in Micro Heat Exchanger

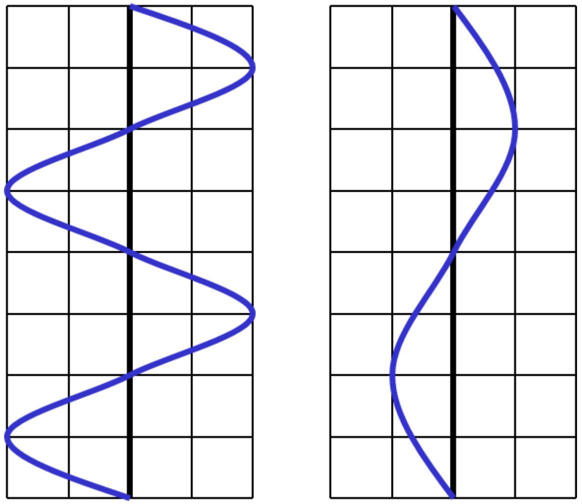
UNIVERSITY OF SURREY





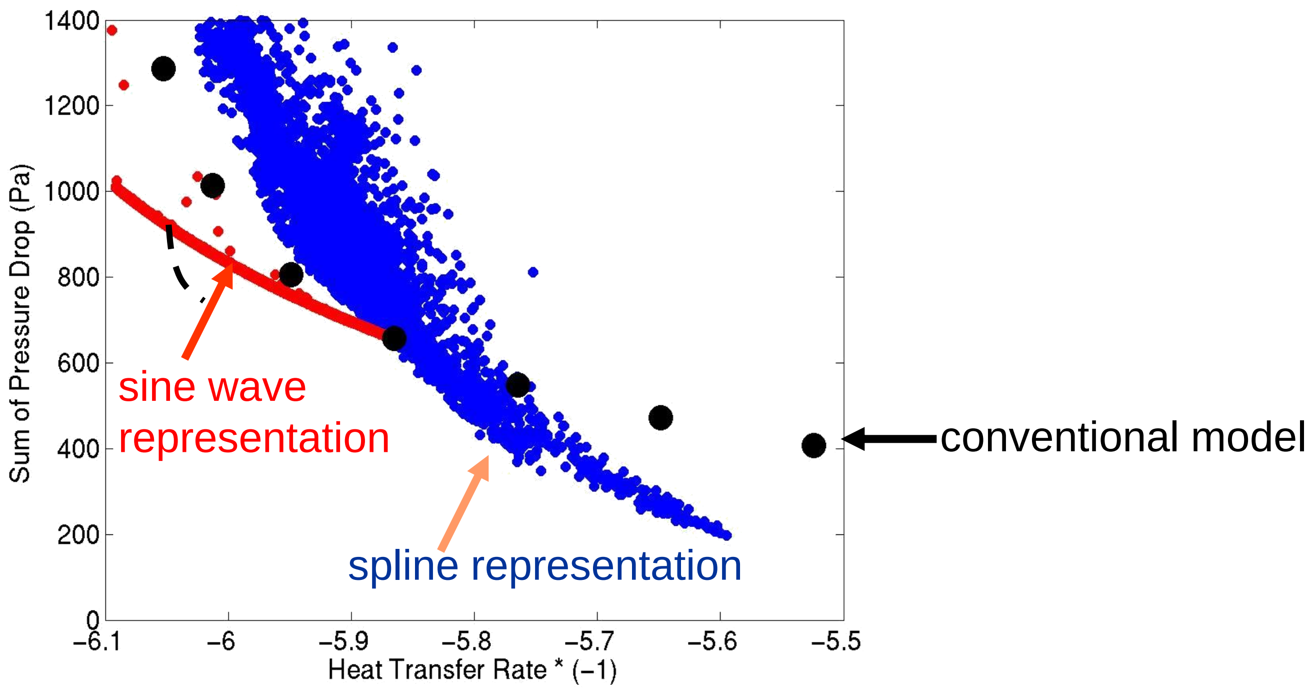


A Spline representation



A frequency-amplitude representation

Micro Heat Exchanger Optimization - Results



- Maximize the heat transfer rate (thermodynamic)
- Minimize the sum of pressure drop with a penalty (aerodynamic)

Complexity in Scale:
Large Decision and Objective Space

Swarm Intelligence for Large-Scale Optimization

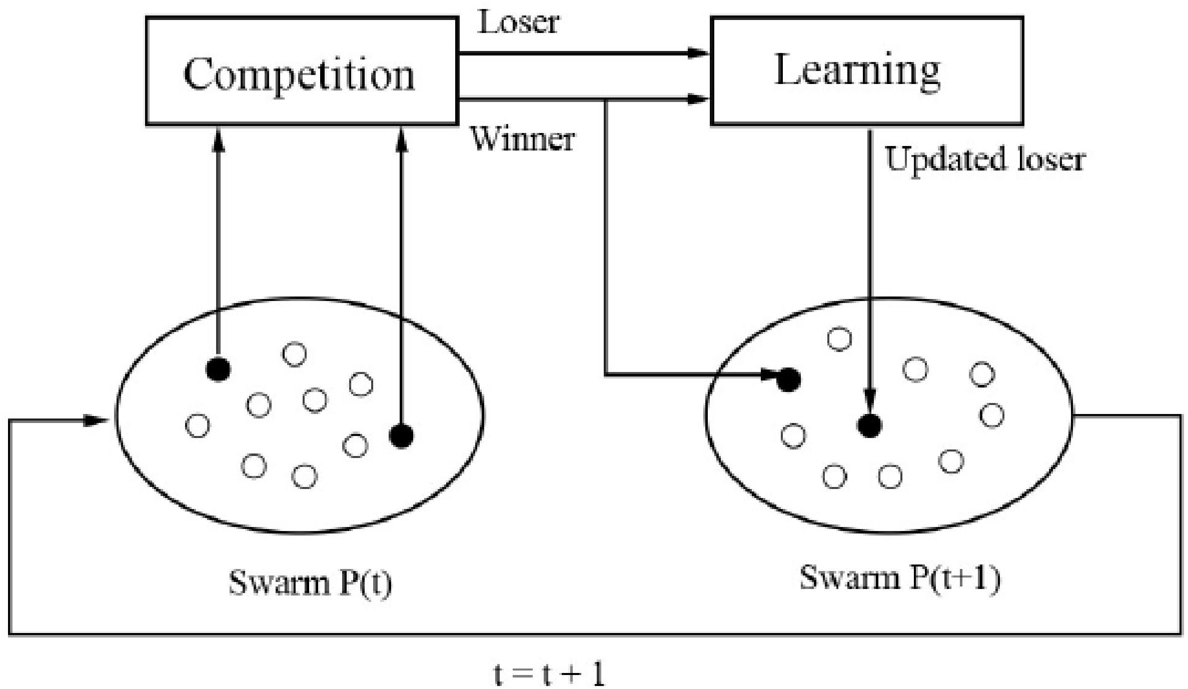
UNIVERSITY OF SURREY

- Large-scale evolutionary optimization
  - Divide and conquer by random grouping
  - Detection of correlation between decision variables
  - **A competitive swarm optimizer (CSO)**
  - **A social-learning based particle swarm optimizer (SL-PSO)**

R. Cheng and **Y. Jin**. A competitive swarm optimizer for large-scale optimization. *IEEE Transactions on Cybernetics*, 45(2):191-205, 2015

R. Cheng and **Y. Jin**. A social learning particle swarm optimization algorithm for scalable optimization. *Information Sciences*, 291:43-60, 2015

Competitive Swarm Optimization (CSO)

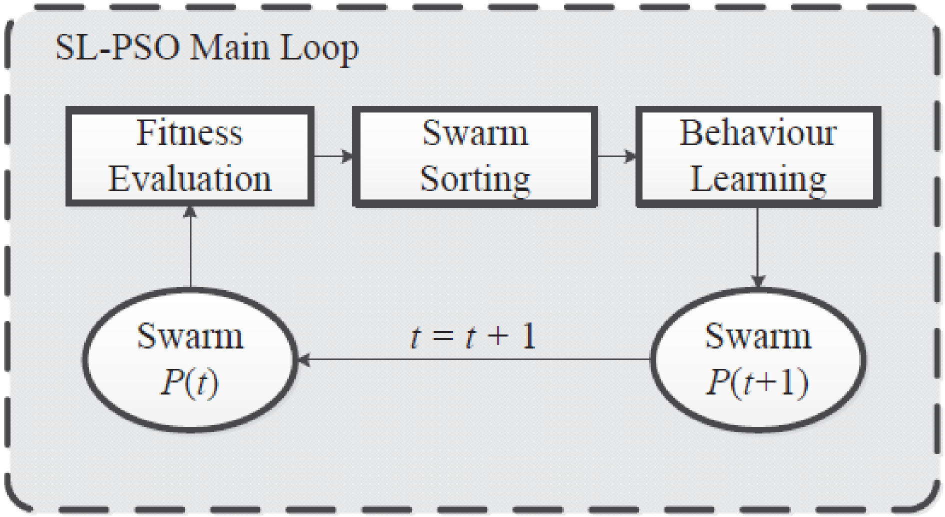


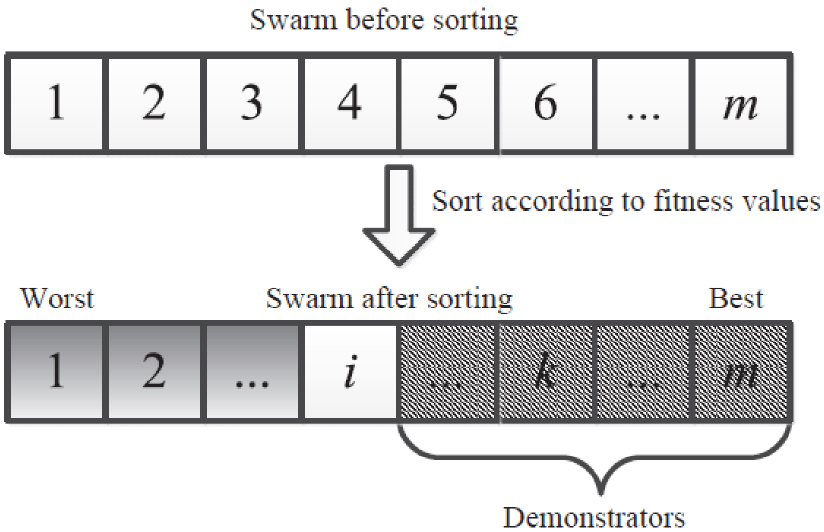
- Randomly pick two solutions
- Compare their fitness. The winner is directly passed to the next generation while the loser will be updated as follows:

$$\begin{aligned} V_{l,k}(t+1) = & R_1(k,t)V_{l,k}(t) \\ & + R_2(k,t)(X_{w,k}(t) - X_{l,k}(t)) \\ & + \varphi R_3(k,t)\left(\bar{X}_k(t) - X_{l,k}(t)\right) \end{aligned}$$

- Neither global nor personal best is used

A Social Learning PSO (SL-PSO)





$$X_{i,j}(t+1) = \begin{cases} X_{i,j}(t) + \Delta X_{i,j}(t+1), & \text{if } p_i(t) \leq P_i^L \\ X_{i,j}(t), & \text{otherwise} \end{cases} \qquad 0 \leq p_i(t) \leq P_i^L \leq 1$$

$$\Delta X_{i,j}(t+1) = r_1(t) \cdot \Delta X_{i,j}(t) + r_2(t) \cdot I_{i,j}(t) + r_3(t) \cdot \epsilon \cdot C_{i,j}(t),$$

$$\begin{cases} I_{i,j}(t) = X_{k,j}(t) - X_{i,j}(t), \\ C_{i,j}(t) = \overline{X}_j(t) - X_{i,j}(t). \end{cases}$$

Large Number of Objectives – Many-Objective Optimization

- Computational challenges
  - Calculation of performance some indicators becomes intractable
- Performance degradation
  - Loss of selection pressure in Pareto-based approaches
- Solution assessment becomes tricky
  - The performance become very sensitive and also easily biased
  - Solution sets are no loner comparable
  - Diversity becomes trickier to measure
- Can we still be able to find a “representative” subset of the Pareto front?

B. Li, J. Li, K. Tang, and X. Yao. Many-objective evolutionary algorithms: A survey. *ACM Computing Surveys*, 48:13–35, 2015
H. Ishibuchi, N. Tsukamoto, and Y. Nojima. Evolutionary manyobjective optimization: A short review. In: *Proceedings of IEEE Congress on Evolutionary Computation*, pages 2419–2426. IEEE, 2008
H. Wang, Y. Jin and X. Yao. Diversity assessment in many-objective optimization*. IEEE Transactions on Cyber*netics, 2016 (accepted)

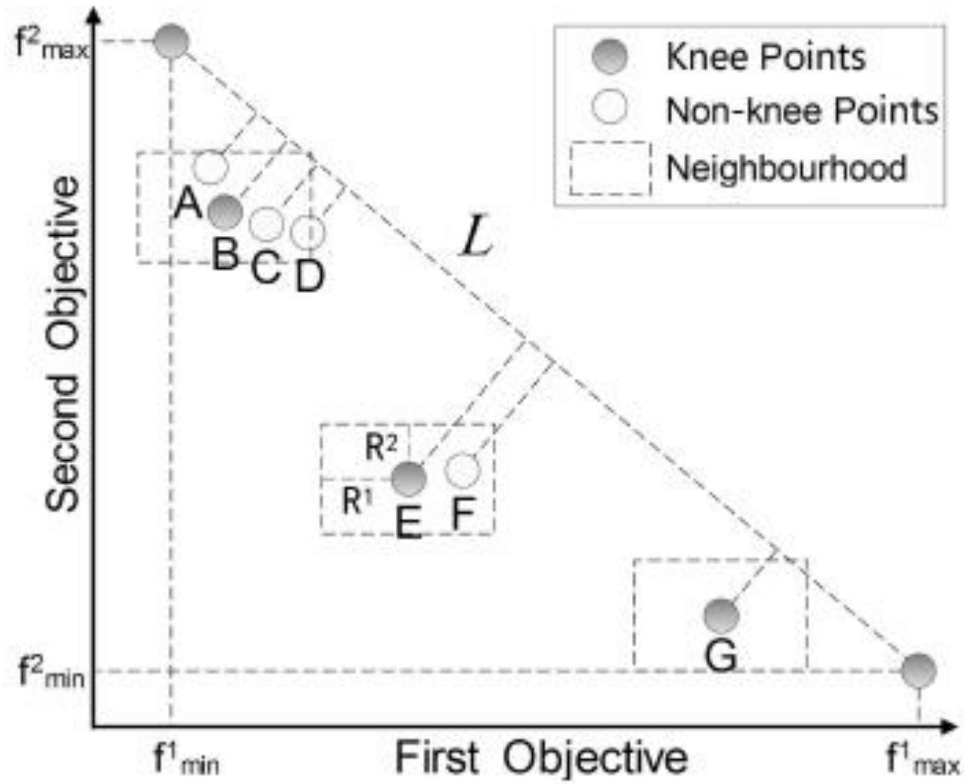
Evolutionary Many-Objective Optimization

EAs for solving MaOPs may largely be divided into the following categories:

- **Preference based**, including decomposition approaches
- **Convergence acceleration**, mainly by modifying the dominance relationship or by **including additional criteria**
- Performance indicator based

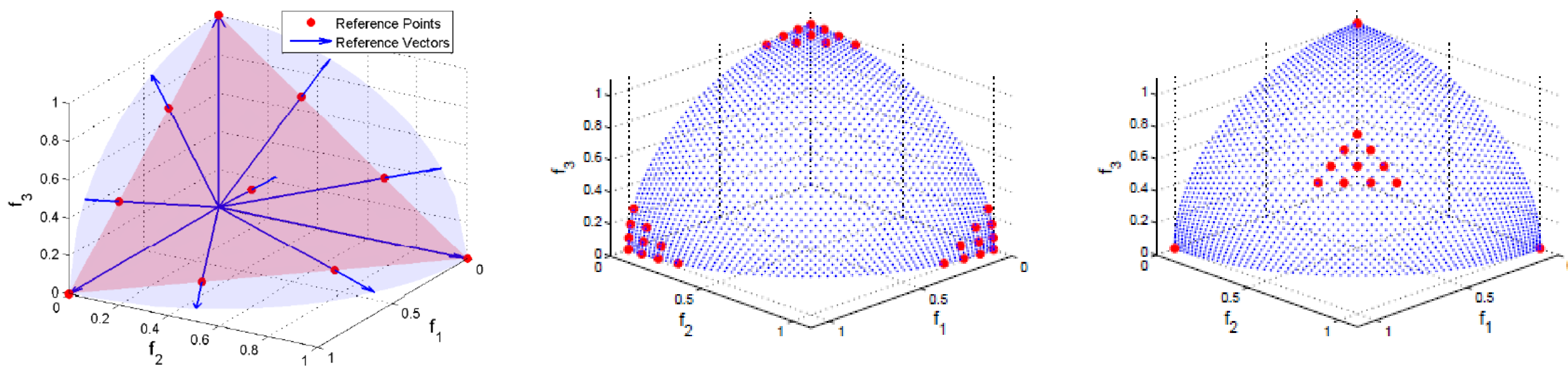
Use of "Knee-Points" to Accelerate Convergence

UNIVERSITY OF SURREY



X. Zhang, Y. Tian, and Y. Jin, A Knee Point Driven Evolutionary Algorithm for Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation*, 19(6):761-776, 2015

Specification of Preferences



Angle-penalized distance (APD):

$$d^j = (1 + P(\theta^j)) \cdot ||\bar{f}^j||,$$

$$P(\theta^j) = k \cdot (\frac{t}{t_{max}})^{\alpha} \cdot \frac{\theta^j}{\gamma_v},$$

R. Cheng, Y. Jin, M. Olhofer and B. Sendhoff. A reference vector guided evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 2016 (Accepted)

Efficient Non-Dominated Sorting

- Non-dominated sorting becomes extremely time-consuming in case of
  - A large population size
  - A large number of objectives

- Computationally efficient non-dominated sorting
  - **ENS**: An efficient non-dominated sorting algorithm for 2 or 3 objectives with a large population size
  - **A-ENS**: An approximate non-dominated sorting for many objectives
  - **T-ENS**: An accurate tree-based non-dominated sorting for large-scale many-objective optimization

X. Zhang, Y. Tian, R. Cheng, and **Y. Jin**. An efficient approach to non-dominated sorting for evolutionary multi-objective \ optimization. *IEEE Transactions on Evolutionary Computation*, 19(2):201-213, 2015
X. Zhang, Y. Tian, **Y. Jin**. Approximate non-dominated sorting for evolutionary many-objective optimization. *Information Sciences*, 2016 (accepted)
X. Zhang, Y. Tian, R. Cheng, **and Y. Jin**. A decision variable clustering-based evolutionary algorithm for large-scale many-objective optimization. 2016 (Submitted)

Code for the ENS variants available!

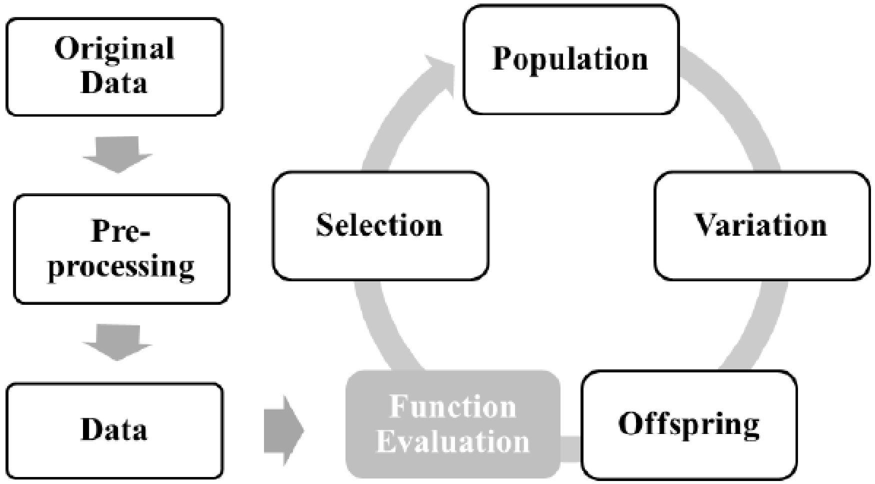
Complexity in Quality Evaluation

Complexity in Quality Evaluation

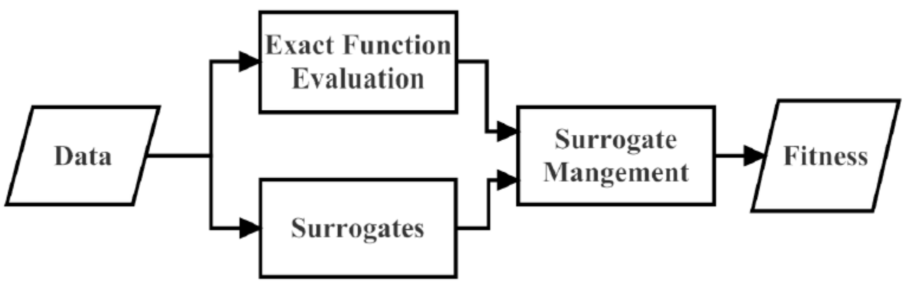
- An analytic fitness function is not available
  - Very time-consuming numerical simulations
  - Expensive experiments
  - History production data only

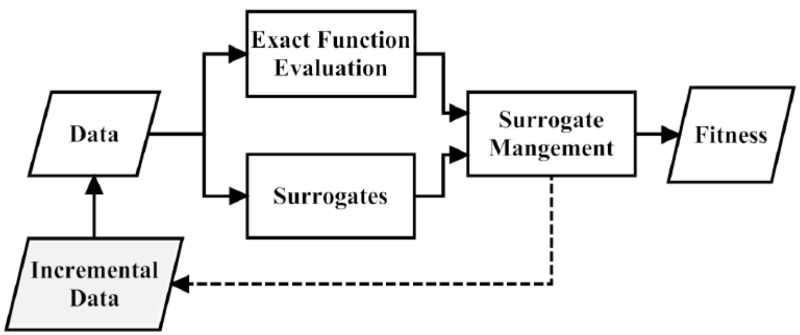
Data Driven Evolutionary Optimization

Data Driven Optimization – Offline and Online



Data-driven evolutionary optimization





Offline data-driven optimization

Online data-driven optimization

H. Wang, Y. Jin and J. O. Jansen. Data-driven surrogate-assisted multi-objective evolutionary optimization of a trauma system. *IEEE Transactions on Evolutionary Computation*, 2016 (accepted)

Offline (Big) Data Driven Evolutionary Optimization

UNIVERSITY OF SURREY

Scottish Trauma system design

Major trauma centre (MTC)

- A specialist hospital responsible for providing a regional or national service for patients with major trauma.

Trauma unit (TU)

- A trauma unit is a hospital which manages less severely injured patients.

Local emergency hospital (LEH)

- A local emergency hospital is a hospital which only deals with minor injuries.

H. Wang, Y. Jin and J. O. Jansen. Data-driven surrogate-assisted multi-objective evolutionary optimization of a trauma system. *IEEE Transactions on Evolutionary Computation*, 2016 (accepted)

Data Driven Evolutionary Trauma System Design

UNIVERSITY OF SURREY

Objective 1: Minimize total travel time

Objective 2: Minimize the exceptions number (the cases "triaged-to-MTC" diverted to a TU)

Constraint 1: Number of helicopter transfers

Constraint 2: MTC case volume

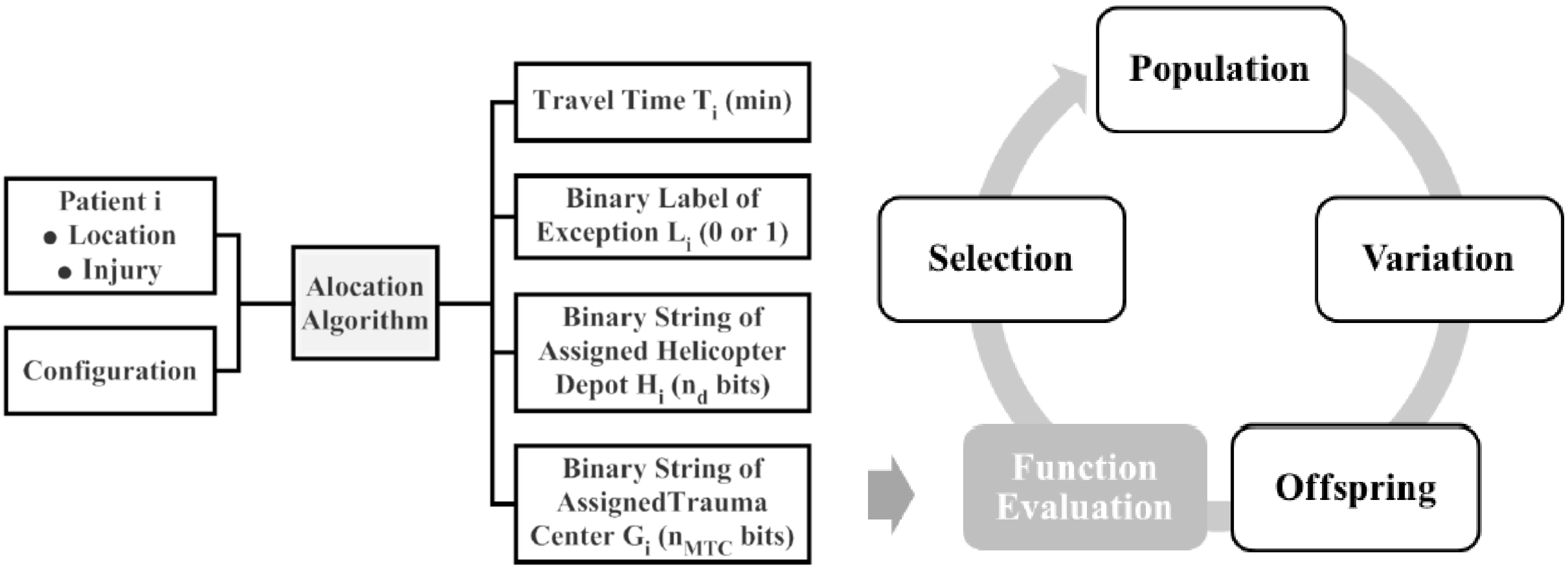
Constraint 3: TU proximity to avoid two close TUs

Main research question is: how to reduce computational time given the large amount of data (the amount of data could be much larger)?

- The EA is able to find the Pareto optimal solutions
- The computation time can be reduced as much as possible

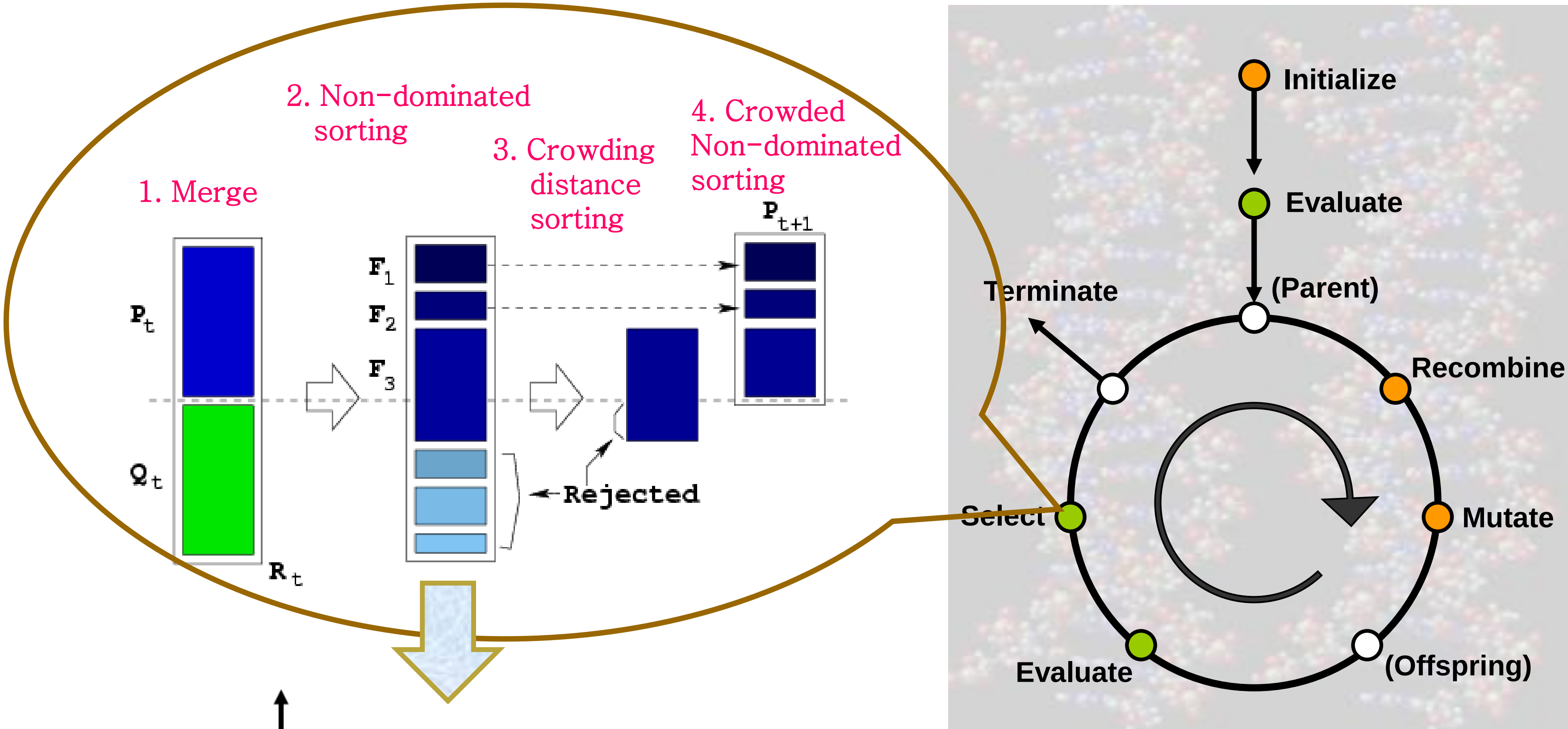
Data Driven Evolutionary Trauma System Design

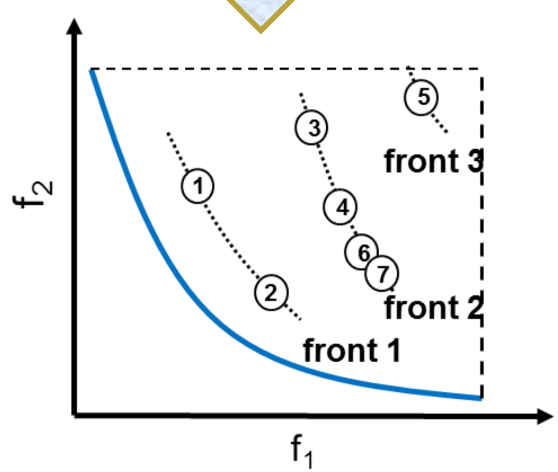
- **40,000** incident records (location, injury, patient)
- 18 trauma centers in Scotland



- **Fitness evaluations is highly time-consuming when the number of records is huge**

Elitist Non-dominated Sorting GA (NSGA-II)





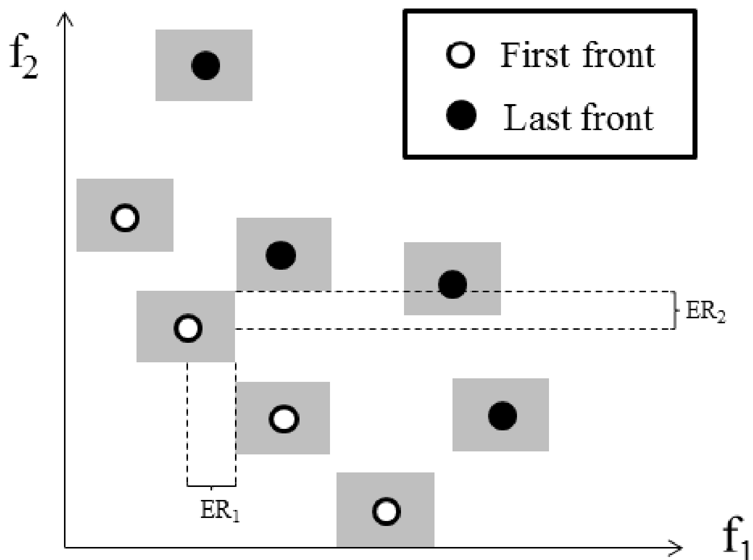
Non-dominated sorting

Data Driven Evolutionary Trauma System Design

UNIVERSITY OF SURREY

Solution

- Group the data into a number of clusters and use the cluster centers to evaluate the objective and constraint functions
- How to make sure the accuracy is good enough?



First and last front of individuals to be selected

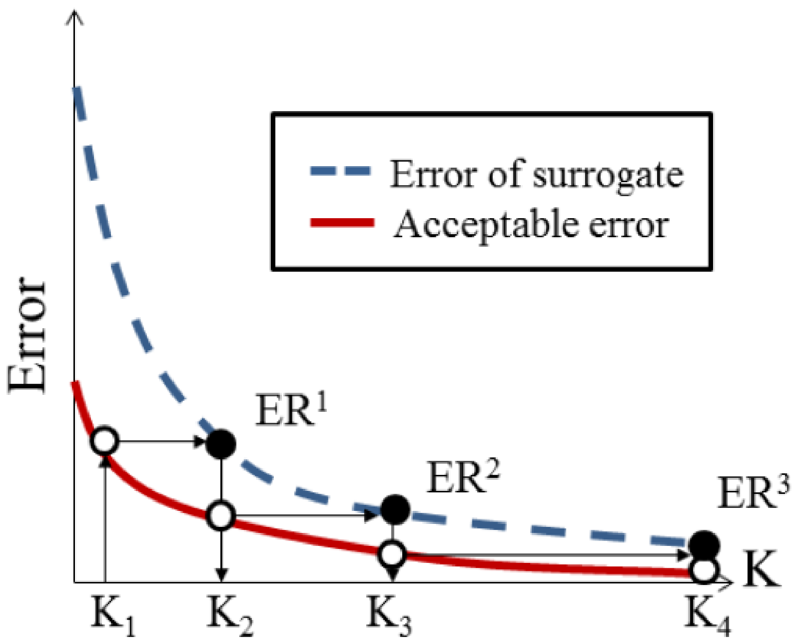
Estimated error for a given number of clusters:

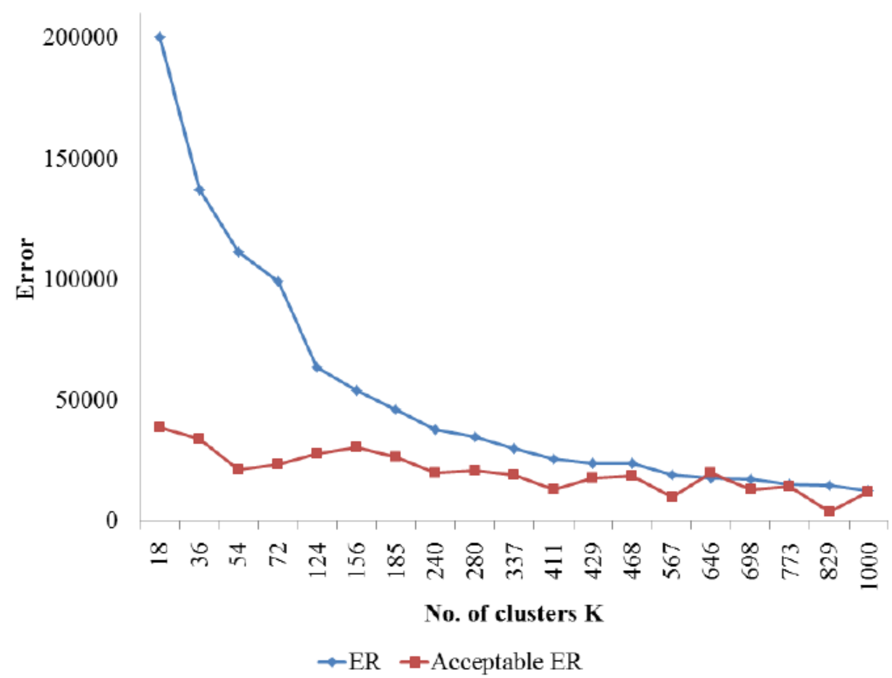
$$ER = \frac{1}{\beta_1 + \beta_2 K},$$

- The maximum error should not change the individuals to be selected

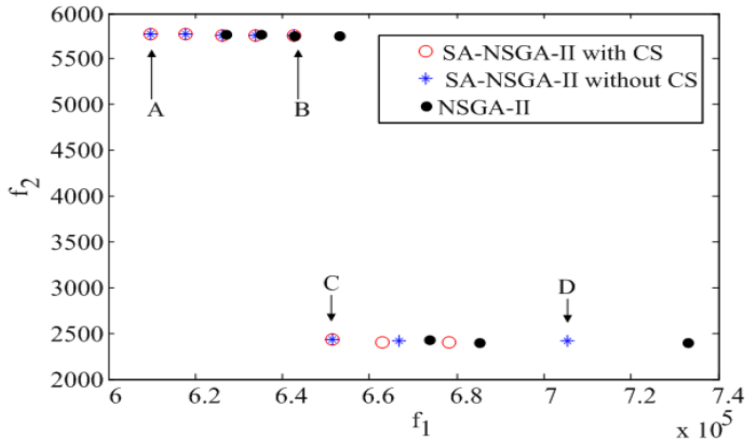
Data Driven Evolutionary Trauma System Design

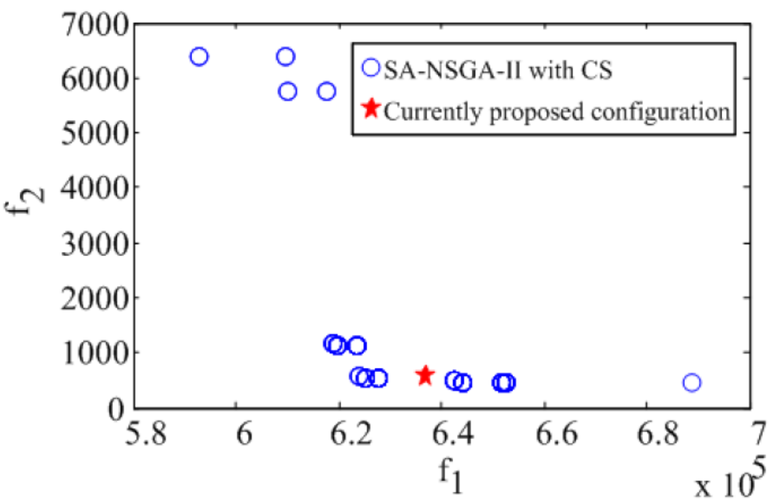
- Adaptation of K

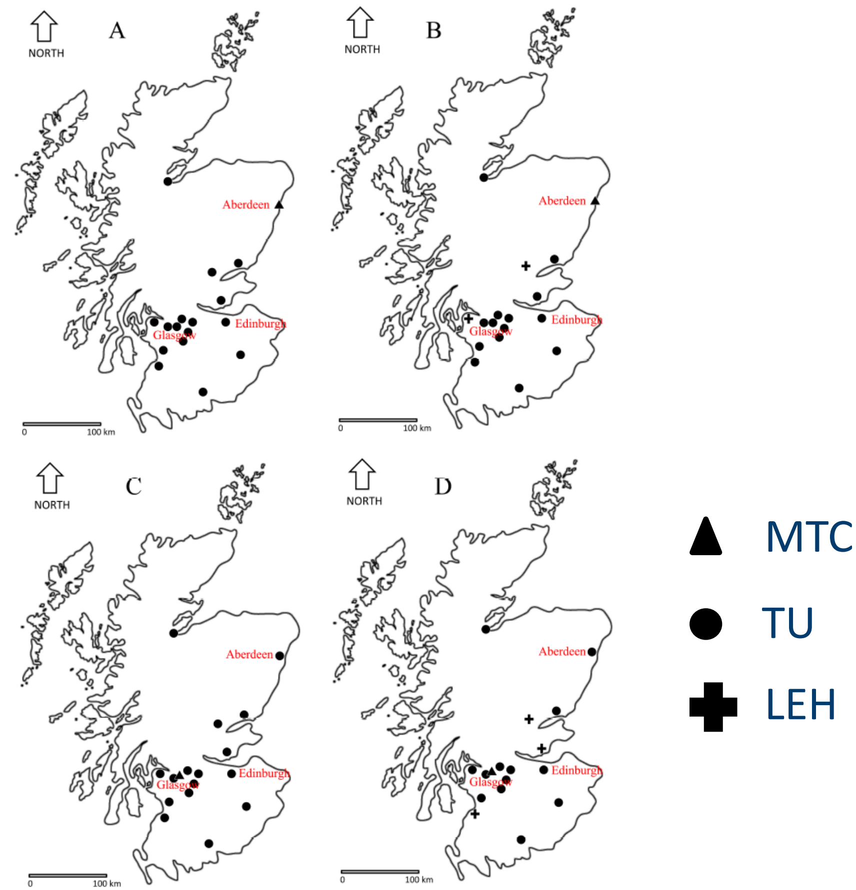




Data Driven Evolutionary Trauma System Design







	IGD	Time (s)
SA-NSGA-II with CS	2.47e-02±1.96e-02	6.54e+03±3.57e+03
SA-NSGA-II without CS	2.51e-02±2.79e-02	10.31e+03±5.00e+03
NSGA-II	0.00e+00±0.00e+00	8.14e+04±5.36e+02

Discussions

- Online big data driven evolutionary optimization, e.g., stream data
- Efficient learning of big data
- Noisy and / or heterogeneous data

Online (Small) Data Driven Evolutionary Optimization

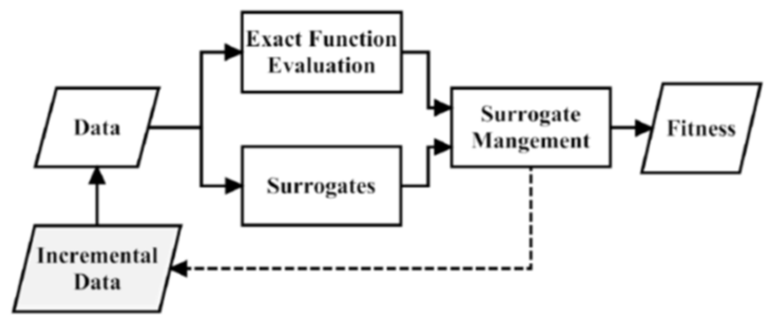
Small Data – Collecting Data is Expensive

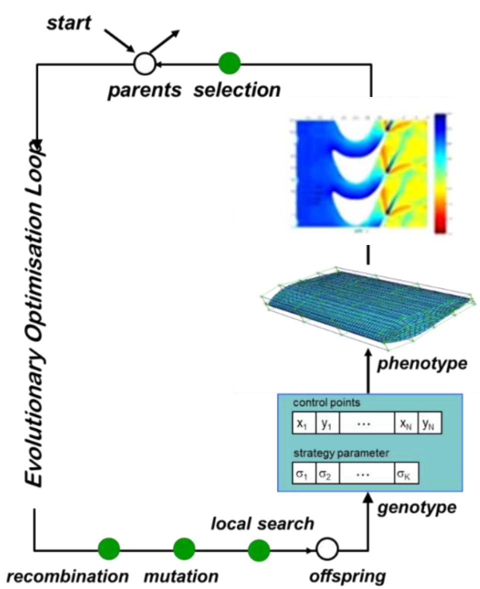
- In many cases, collecting data is very expensive
  - Highly time-consuming numerical simulation
  - Expensive physical experiments
  - Real-industrial processes only

Surrogate-Assisted Evolutionary Optimization

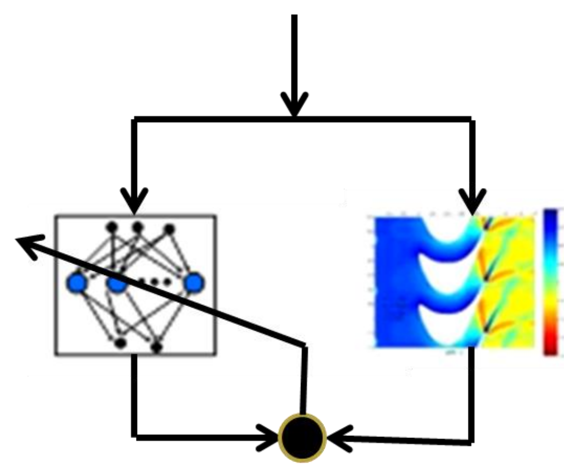
UNIVERSITY OF SURREY

- Use a meta-model /surrogate to replace the expensive fitness evaluations, e.g., CFD simulations
    - Choose a surrogate
    - Collect data
    - Train the surrogate
    - Replace the CFD





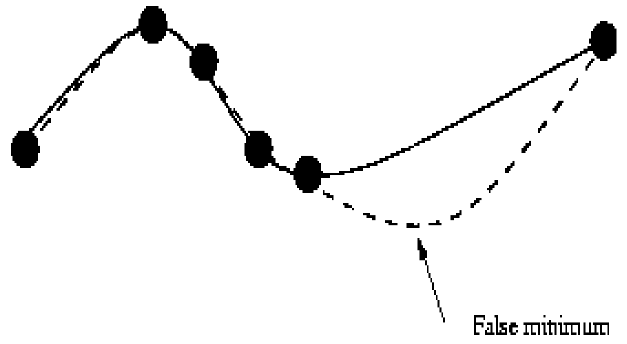
Online data-driven optimization



CFD simulations / Experiments

Model Management

- Use surrogates only: Risk of converging to a false optimum



- Surrogate management / evolution control
  - population-based
  - generation-based
  - individual-based
  - local search
  - Combination of the above

Y. Jin. A comprehensive survey of fitness approximation in evolutionary computation. *Soft Computing*, 9(1), 3-12, 2005
Y. Jin. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation*, 1(2):61-70, 2011

Model Management – Main Questions

Which individuals are to be re-evaluated using the real-fitness function?

- Solutions that are of potentially good performance
- Solutions whose estimated fitness have large amount of uncertainty
  - Less explored
  - Effective for model improvement

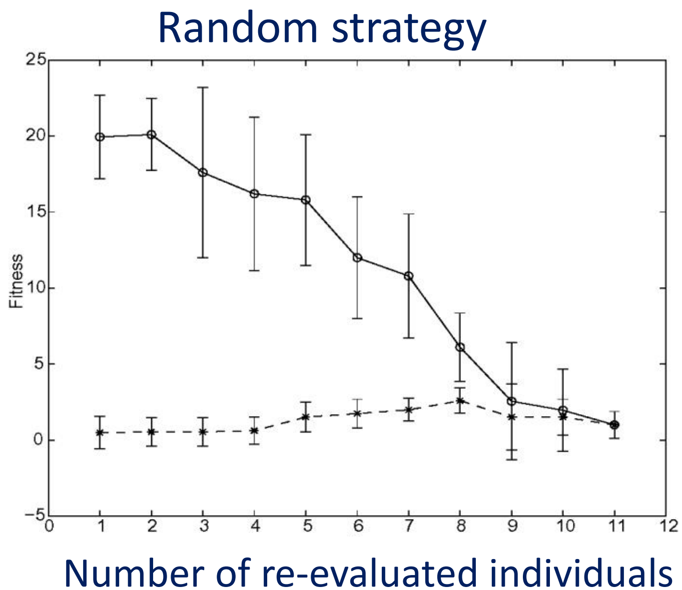
- How to measure uncertainty?

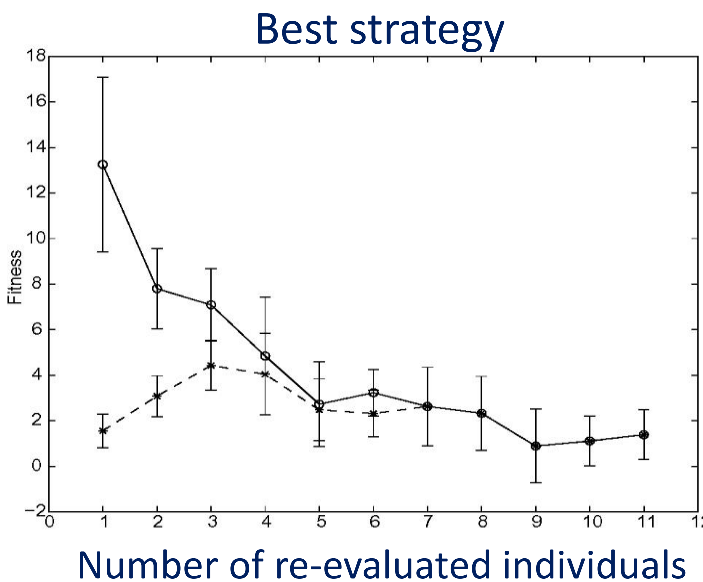
- How to measure model quality? (Jin et al 2003, Huesken et al 2005)

M. Huesken, Y. Jin and B. Sendhoff. Structure optimization of neural networks for evolutionary design optimization. *Soft Computing*, 9(1), 21-28, 2005

Y. Jin, M. Huesken and B. Sendhoff. Quality measures for approximate models in evolutionary computation. In: *Proceedings of the GECCO Workshop on "Learning, Adaptation and Approximation in Evolutionary Computation*", pp.170-174, Chicago, 2003

Model Management – Promising Ones





- 12-D Ackley function
- (3,12)-ES
- average over 10 runs

- The best strategy is more efficient than the random strategy
- In the best strategy, about half of the individual should be controlled to guarantee correct convergence

Y. Jin, M. Olhofer and B. Sendhoff. A framework for evolutionary optimization with approximate fitness functions. *IEEE Transactions on Evolutionary Computation*, 6(5): 481-494 (2002)

Model Management – Promising and Uncertain Ones

Given a stochastic model (Gaussian process),

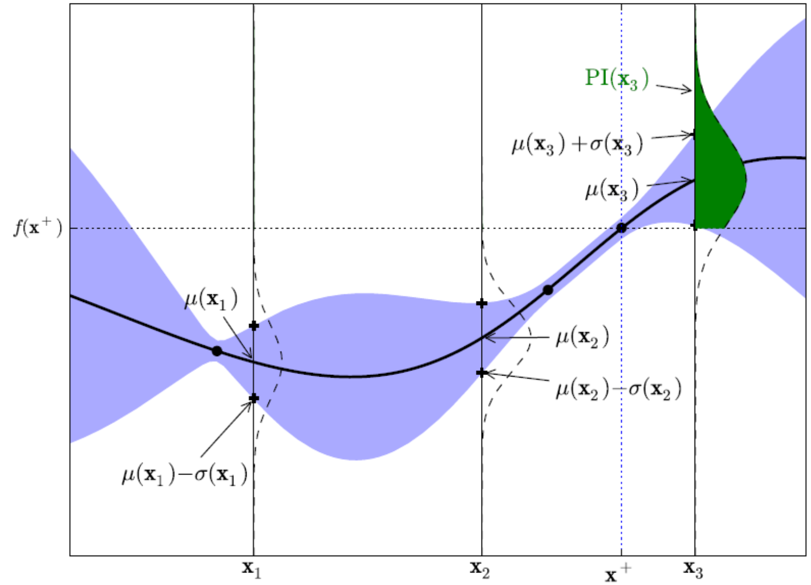
- Mean fitness value:

$$f = \mu(x);$$

- Lower confidence bound (LCB)

$$f = \mu(x) - \alpha\sigma(x)\ (\alpha > 0)$$

- Expected improvement (EI)



(Taken from Brochu et al, 2010)

$$\mathrm{EI}(\mathbf{x}) = \begin{cases} (\mu(\mathbf{x}) - f(\mathbf{x}^+))\Phi(Z) + \sigma(\mathbf{x})\phi(Z) & \text{if } \sigma(\mathbf{x}) > 0 \\ 0 & \text{if } \sigma(\mathbf{x}) = 0 \end{cases} \qquad Z = \frac{\mu(\mathbf{x}) - f(\mathbf{x}^+)}{\sigma(\mathbf{x})}$$

- Probability of Improvement (PI)

$\Phi(\cdot)$ is the normal cumulative distribution function.

$$\mathrm{PI}(\mathbf{x}) = P(f(\mathbf{x}) \geq f(\mathbf{x}^+)) = \Phi\left(\frac{\mu(\mathbf{x}) - f(\mathbf{x}^+)}{\sigma(\mathbf{x})}\right)$$

M. Emmerich, K.C. Giannakoglou, B. Naujoks, Single- and multiobjective evolutionary optimization assisted by Gaussian random field metamodels. *IEEE Transactions on Evolutionary Computation*, 10 (4) : 421–439 (2006)

Eric Brochu, Vlad M. Cora and Nando de Freitas. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning, 2010. https://arxiv.org/pdf/1012.2599

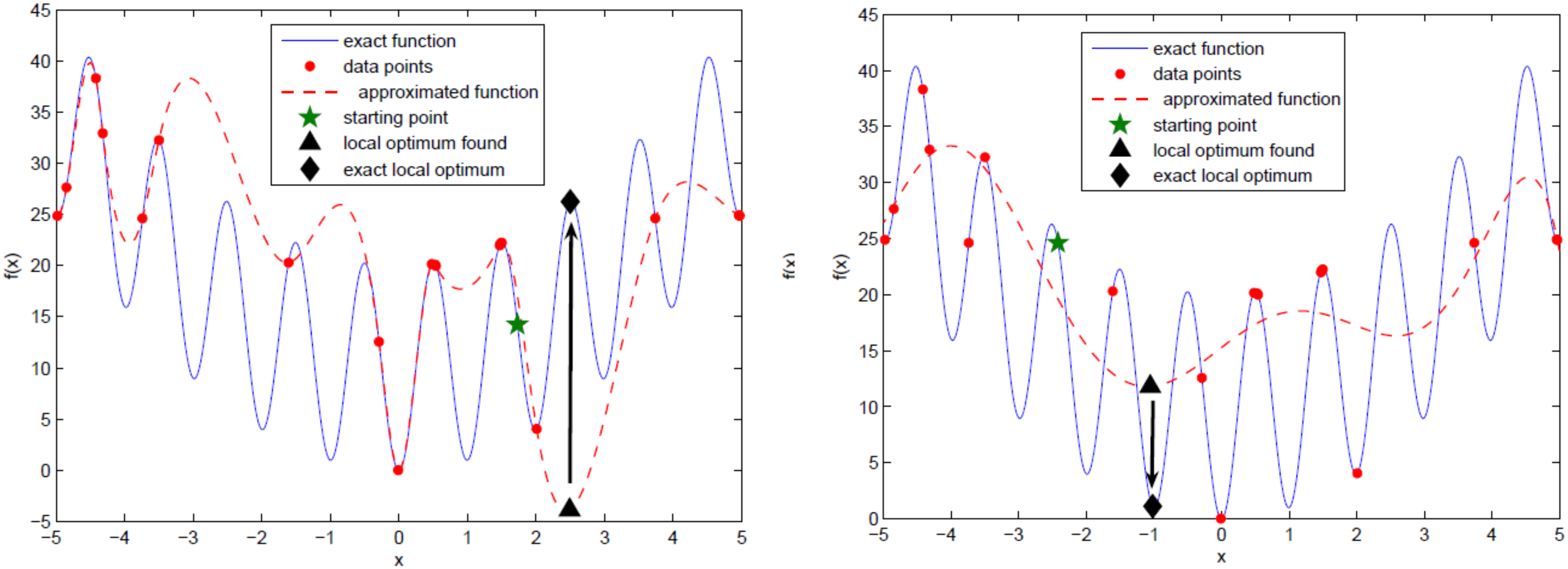
Model Management in Multi-Objective Optimization

UNIVERSITY OF SURREY

- Each objective is considered separately (similar to single objective optimization)
- Multiple objectives are converted to a scalar objective function and then use the model management criteria for single objective optimization
  - Random weights
  - Uniformly distributed weights
- Use a scalar performance indicator, e.g., hypervolume

D. Horn, T. Wagner, D. Biermann, C. Weihs, and B. Bischl. Model-Based Multi-Objective Optimization: Taxonomy, Multi-Point Proposal, Toolbox and Benchmark. In: *Evolutionary Multi-Criterion Optimization*, LNCS 9018, pages 64–78.

Potential Benefit of a Global Model

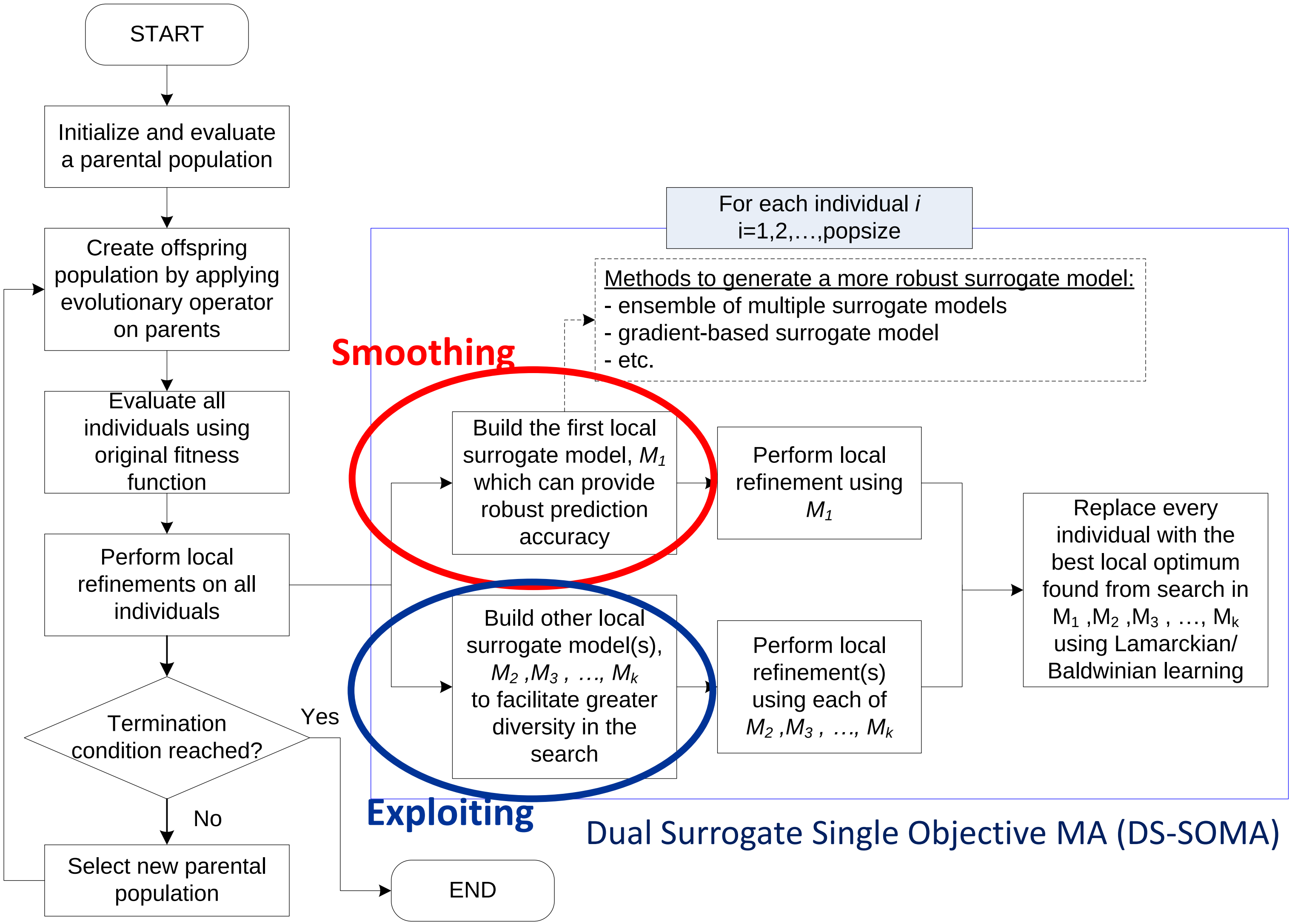


A global model might help smoothen the fitness landscape

D. Lim, Y. Jin, Y.-S. Ong, and B. Sendhoff. Generalizing surrogate-assisted evolutionary computation. *IEEE Transactions on Evolutionary Computation*, 14(3):329 - 355, 2010

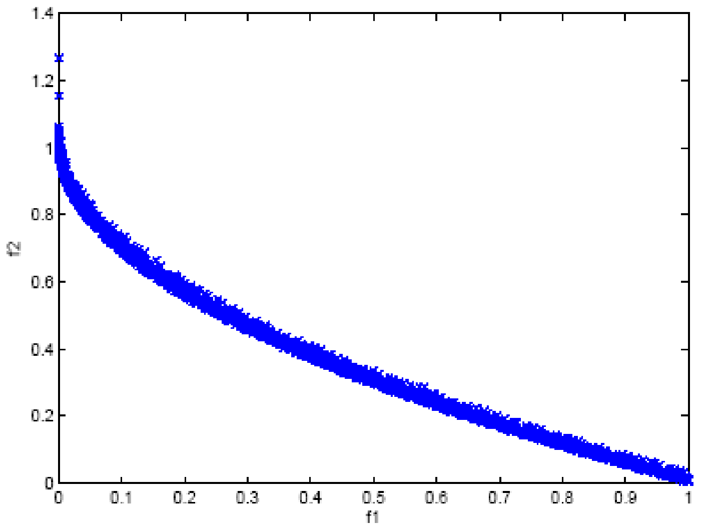
C. Sun, Y. Jin, J. Zeng and Y. Yu. A two-layer surrogate-assisted particle swarm optimization algorithm. *Soft Computing*, 19(6):1461-1475, 2015

Dual Surrogates in Memetic Algorithms

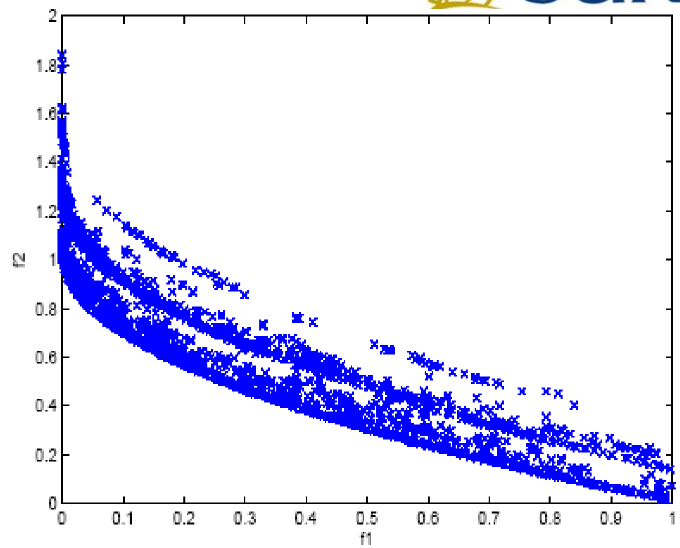


Dual Surrogate Single Objective MA (DS-SOMA)

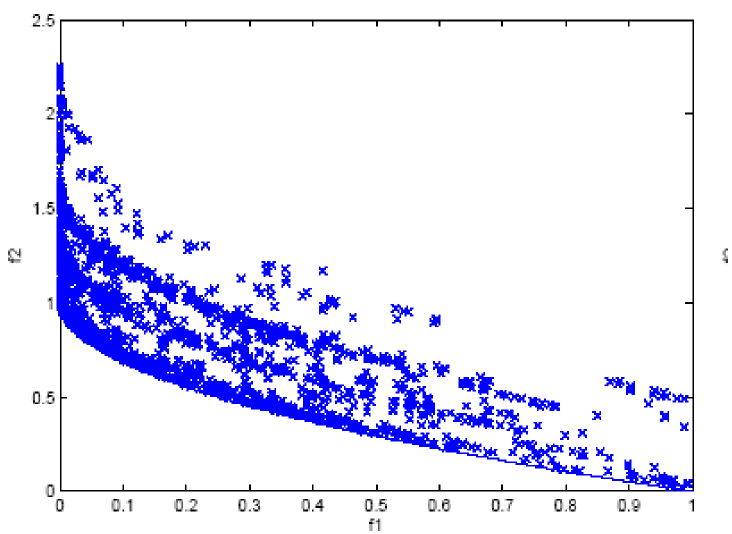
Results: Multi-Objective



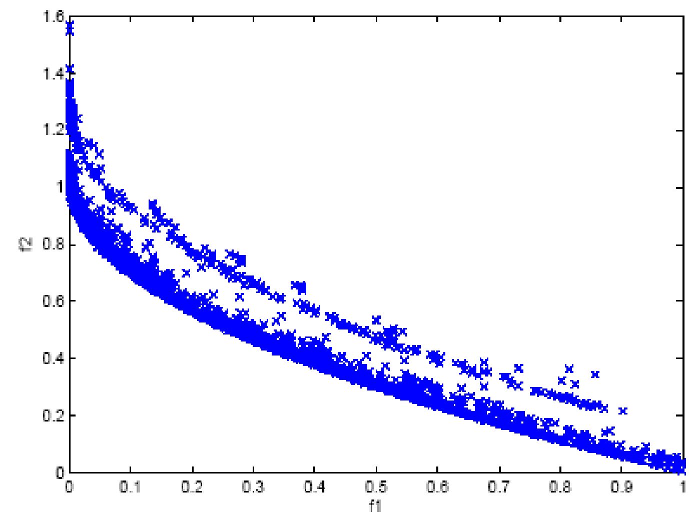
DS-MOMA



SS-MOMA- PR+RBF+GP



SS-MOMA-PR



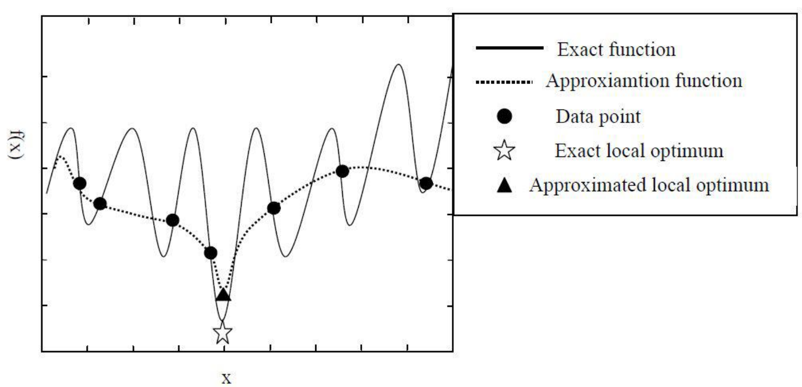
SS-MOMA-Perfect

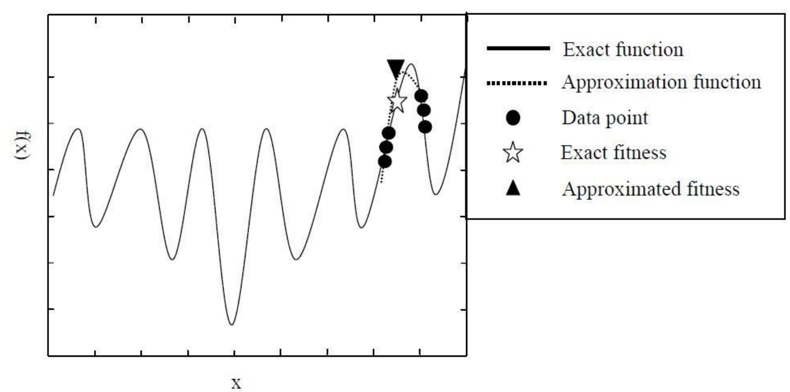
All results from 20 runs

Global and Local Surrogate Models

Two-layer (global and local) surrogate-assisted PSO

Construct a global surrogate model;
Approximate a fitness value for each individual in the swarm using the global surrogate model;
for each particle i in the swarm **do**
    Find its neighbors in the local database;
    if there are enough samples to construct a local surrogate **then**
        Construct a local surrogate model;
        Approximate the fitness of particle i using the local surrogate model;
        $\tilde{f}(\mathbf{x}_i) = \min\{\tilde{f}_g(\mathbf{x}_i), \tilde{f}_l(\mathbf{x}_i)\}$
    else
        $\tilde{f}(\mathbf{x}_i) = \tilde{f}_g(\mathbf{x}_i)$
    end if
end for





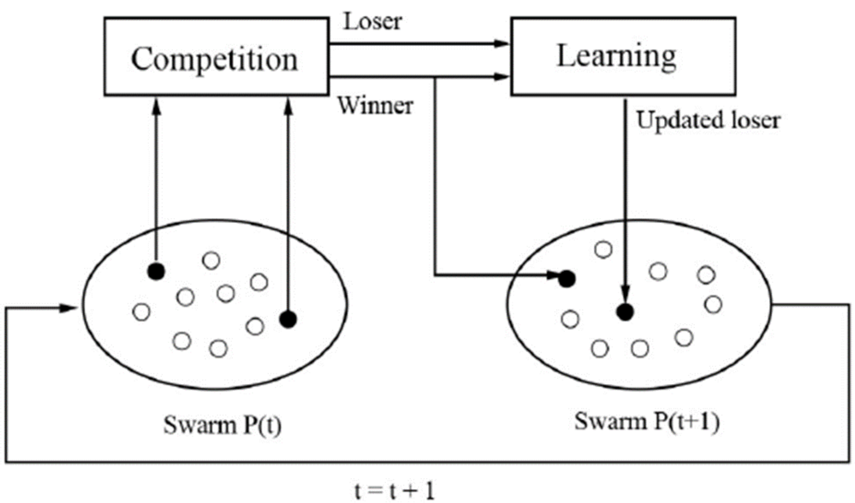
C. Sun, Y. Jin, J. Zeng and Y. Yu. A two-layer surrogate-assisted particle swarm optimization algorithm. *Soft Computing*, 19(6):1461-1475, 2015

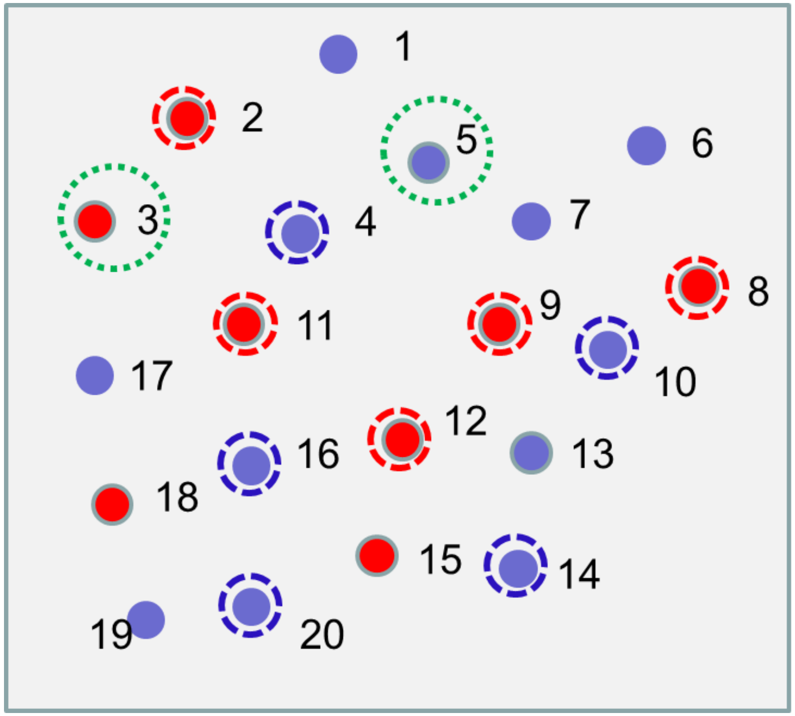
Surrogate-Assisted Large-Scale Evolutionary Optimization?

UNIVERSITY OF SURREY

- Dimension mostly limited up to 10 in GP-assisted EAs, mostly up to 30
  - Curse of dimensionality
  - Dramatic increase in computational cost for training surrogates, e.g., Gaussian processes – **it can take hours to build a GP model ☹**

Fitness Estimation Assisted CSO for Large-Scale Optimization





- Fitness estimation in competitive swarm optimization for dimensions up to 500
  - Fitness estimated based on positional relationships

Particles whose fitness is estimated (*ES*)

Particles whose fitness is calculated (*EV*)

Loser particles whose fitness is estimated (*ES*)

Winner particle whose fitness is calculated (*EV*)

C. Sun, J. Ding, J. Zeng and Y. Jin. Fitness approximation assisted competitive swarm optimizer for large scale expensive optimization problems. *Memetic Computing*, 2016 (accepted)

Concluding Remarks

- Evolutionary optimization of complex systems is promising yet challenging
- Data-driven evolutionary optimization becomes increasing important
- Surrogate-assisted evolutionary optimization will not only be essential for evolutionary optimization of complex systems but also provides a platform for integrating evolution and learning techniques
  - Which to re-evaluate (sample)
    - Active learning
  - Small data
    - Semi-supervised learning
    - Transfer learning
  - big data
    - Deep learning