Solving optimization problems with machine learning
Application to materials science

Gérard Ramstein

Laboratoire des Sciences du Numérique de Nantes

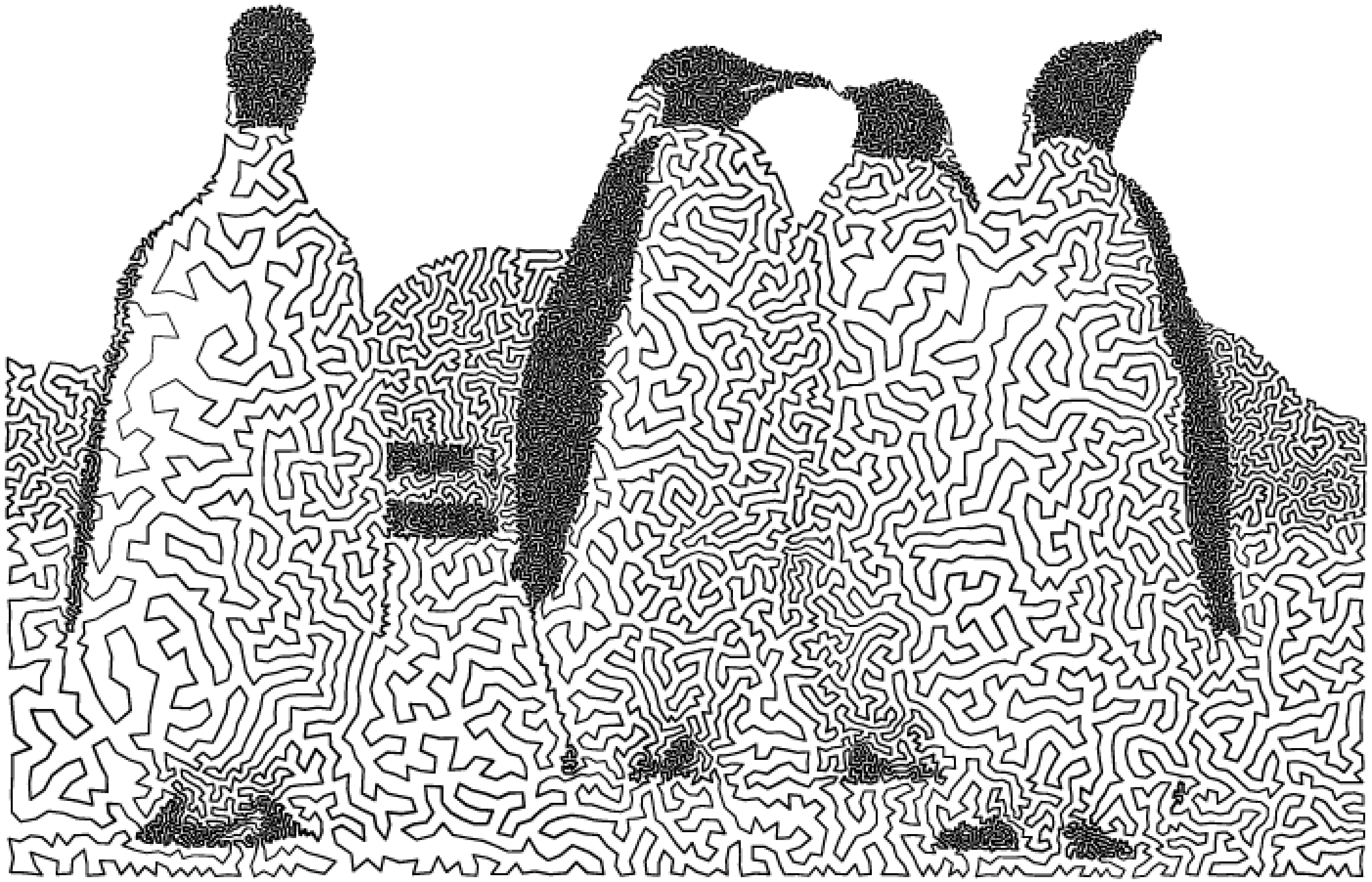
Collaboration with F. Tancret (Institut des Matériaux de Nantes)

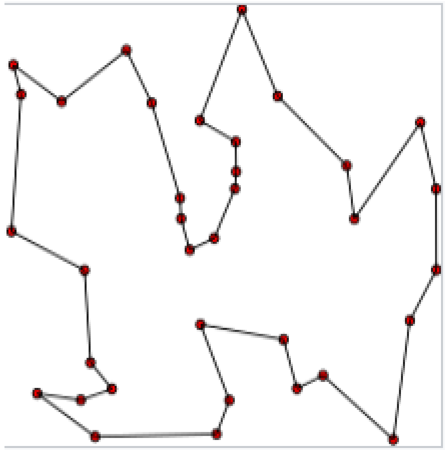
Outline

1. Optimization and ML
2. Constrained Multiobjective Optimization
3. Bayesian Optimization

1. Optimization and ML

What is Optimization ?





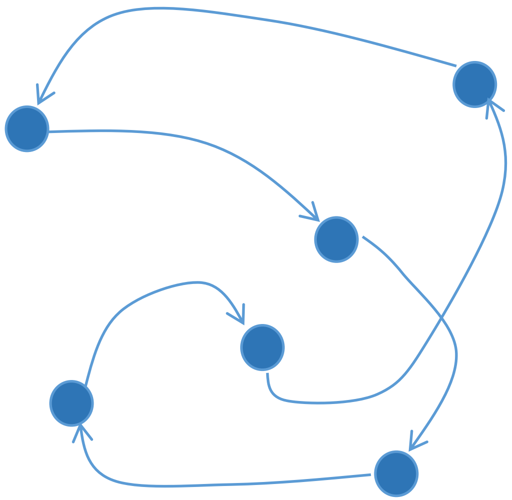
GOAL:
shortest possible path that visits each city exactly once

https://makezine.com/2010/09/04/traveling-salesman-problem-art/

Modelling

SOLUTION SPACE

DECISION SPACE



mapping

Loss function
Cost function

…

Objective function
Fitness function

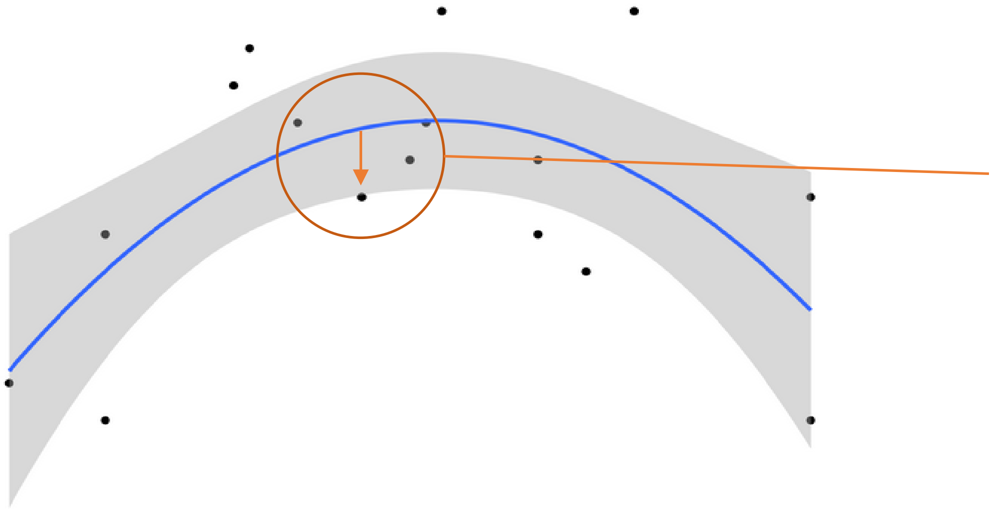
Description of a candidate solution

City1 -> City2->…->City1

distance

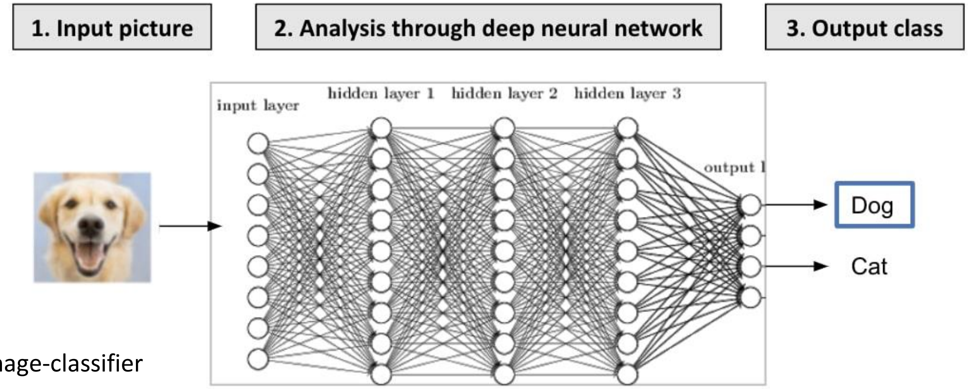
Many problems can be formulated as an optimization problem

Regression



Min dist ● → ●

Classification

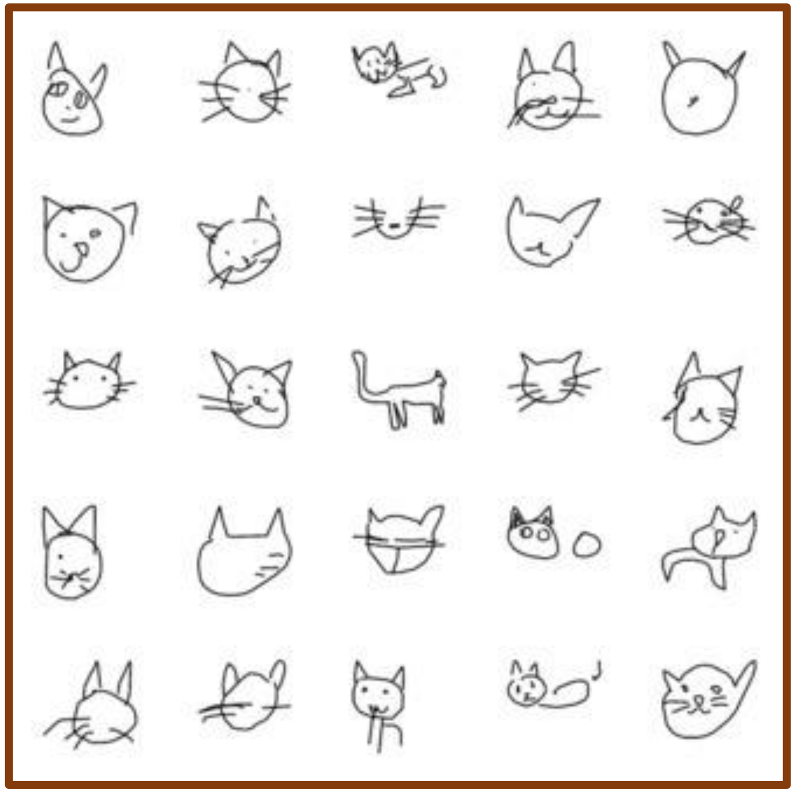


https://levity.ai/blog/what-is-an-image-classifier

Log Loss:

$$\frac{1}{N}\sum_{i=1}^{N} -\left(y_i * \log(p_i) + (1-y_i) * \log(1-p_i)\right)$$

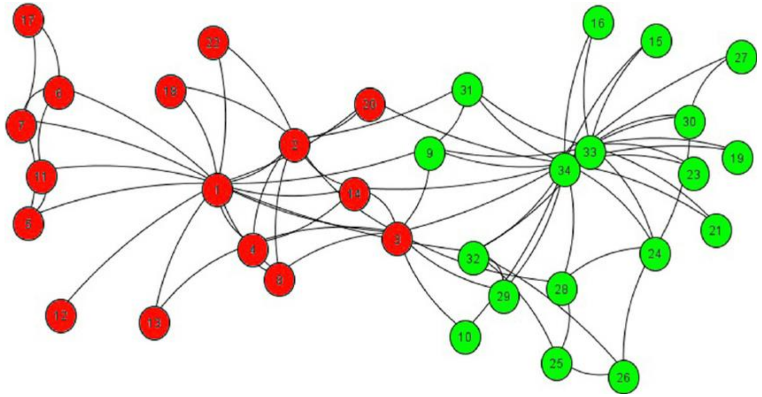
What is Machine Learning ?



15 million players have contributed to
50 millions of drawings

Learning Types

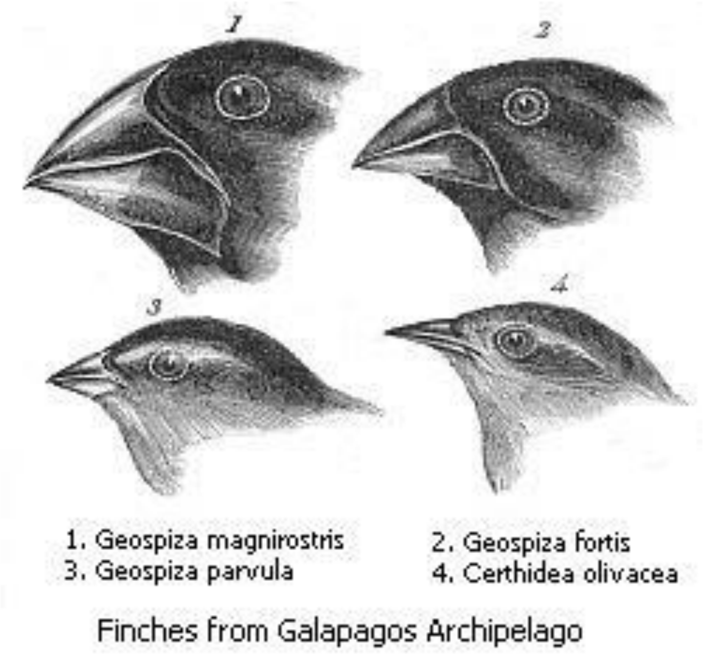
- Supervised: data are labeled with predefined classes
- Unsupervised: only the data
- By reinforcement: no labels, but a reward.



An example of optimization algorithm

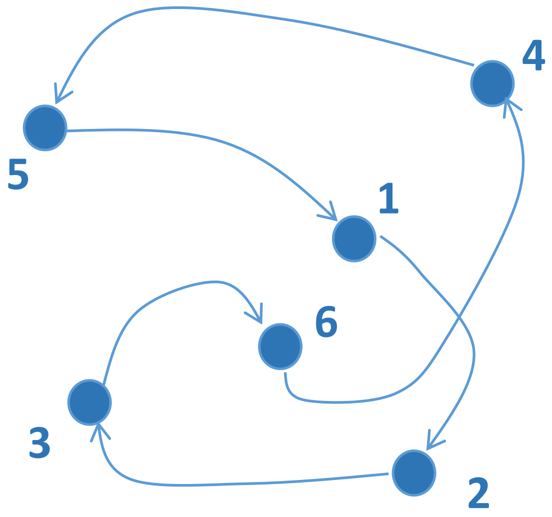
Evolutionary Algorithm

- population evolves generation after generation
- genetic heritage evolves through random transformations



https://fr.wikipedia.org/wiki/S%C3%A9lection_naturelle

A genetic algorithm for TSP

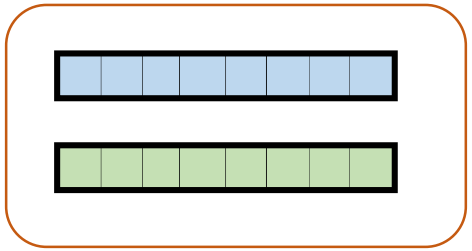


(NANTES, ANGERS, LE MANS, PARIS, NANTES)

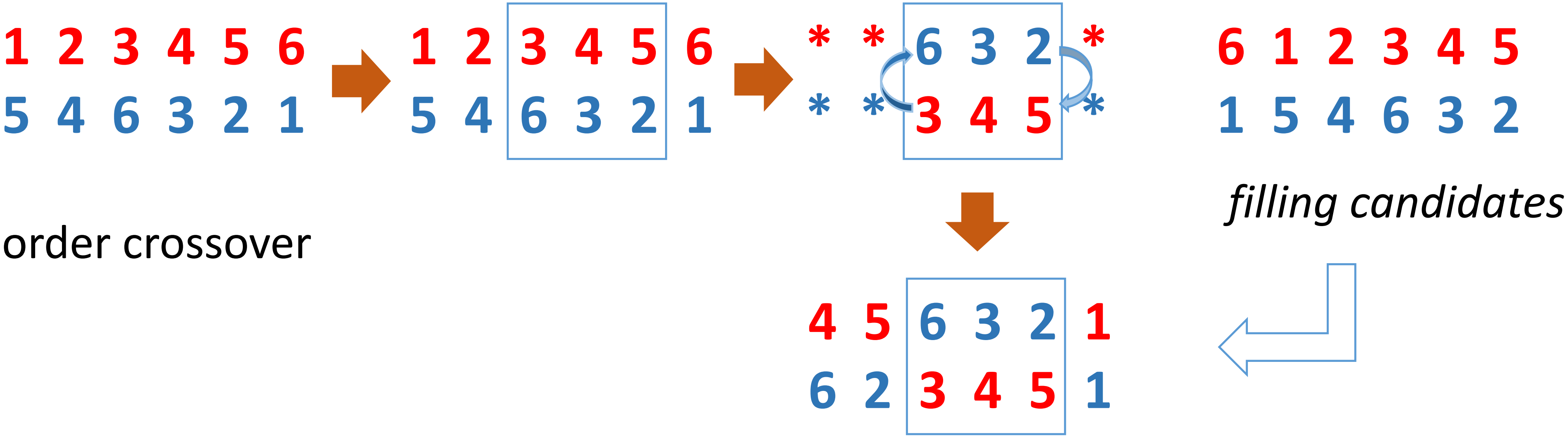
Genotype

(NANTES, ANGERS, LE MANS, PARIS, NANTES) ⇨ 1 2 3 4
(NANTES, LE MANS ANGERS, PARIS, NANTES) ⇨ 1 3 2 4
(NANTES, PARIS, ANGERS, LE MANS, NANTES) ⇨ 1 4 2 3
…

Population

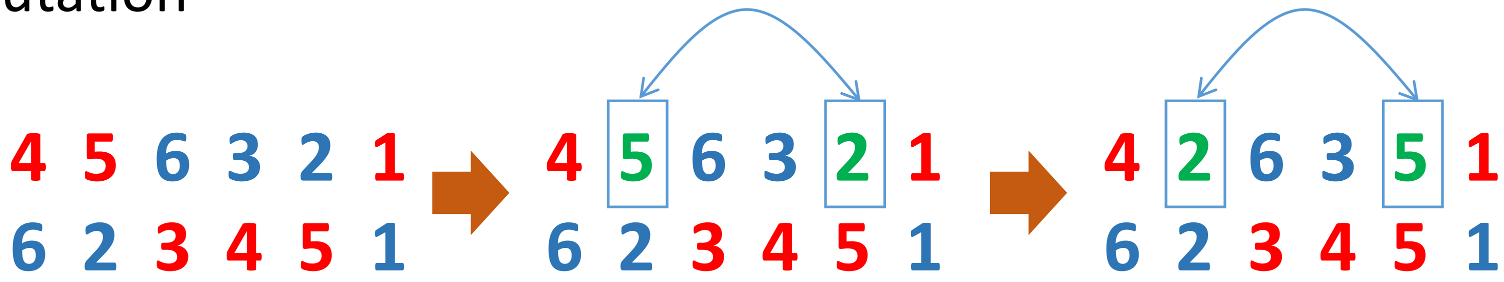


Variation: Crossover and Mutation

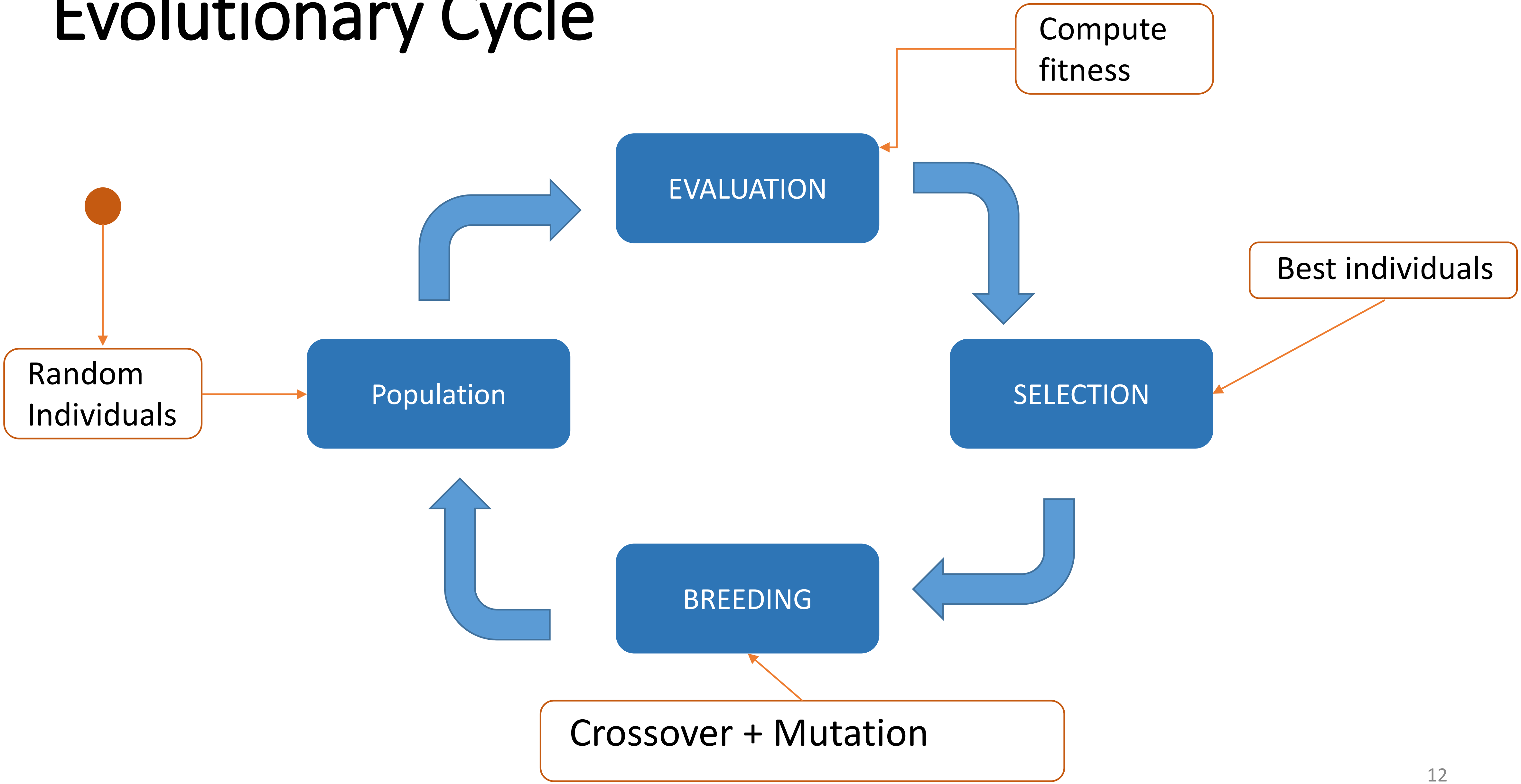


order crossover

mutation



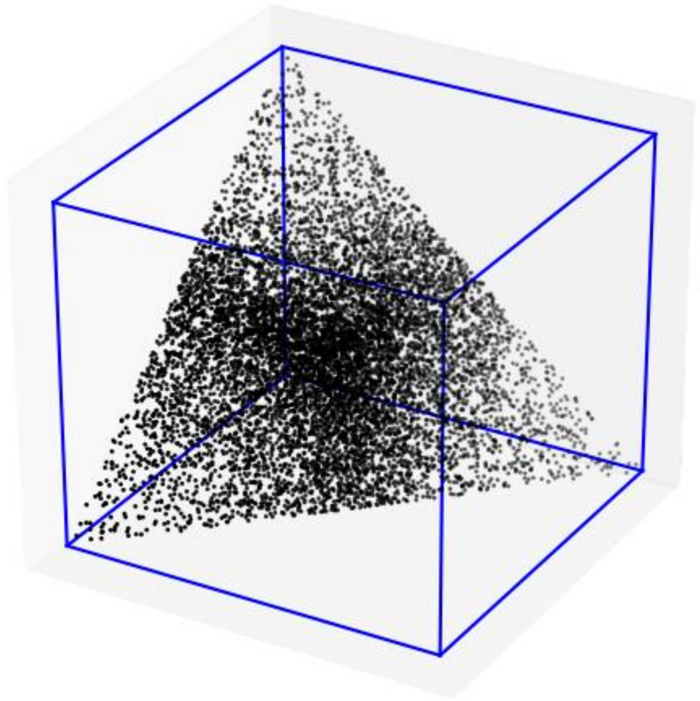
Evolutionary Cycle



Initial population and space filling design

Mixture:
X1, X2, X3

X1+X2+X3=1



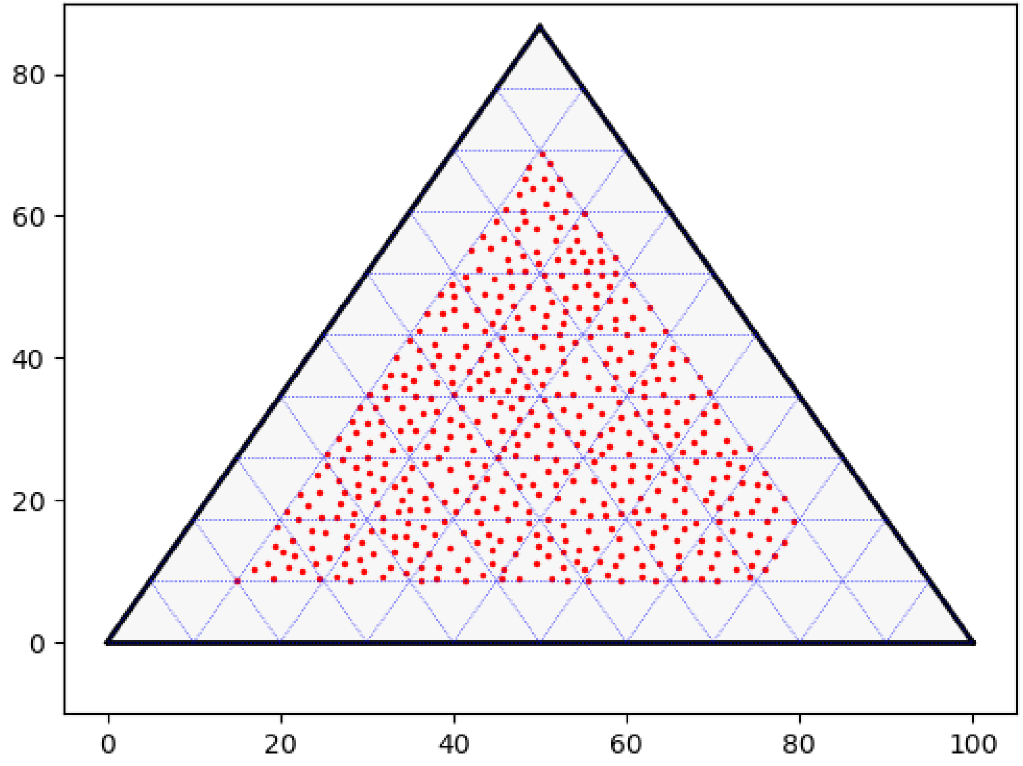
Example of a biased method for compositional data
(normalization of the proportions)

A better approach

Mixture:
X1, X2, X3

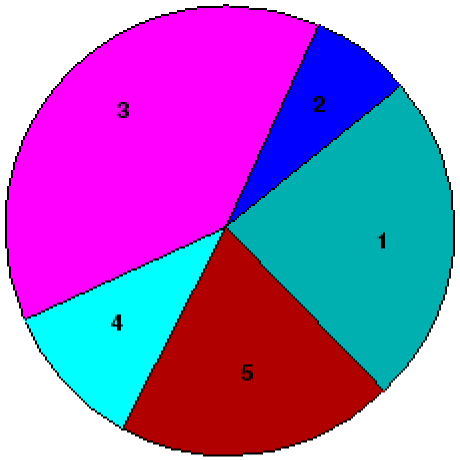
X1+X2+X3=1

size	1,000
Lower bound	[0.1,0.1,0.1]
Upper bound	[0.7,0.8,0.8]



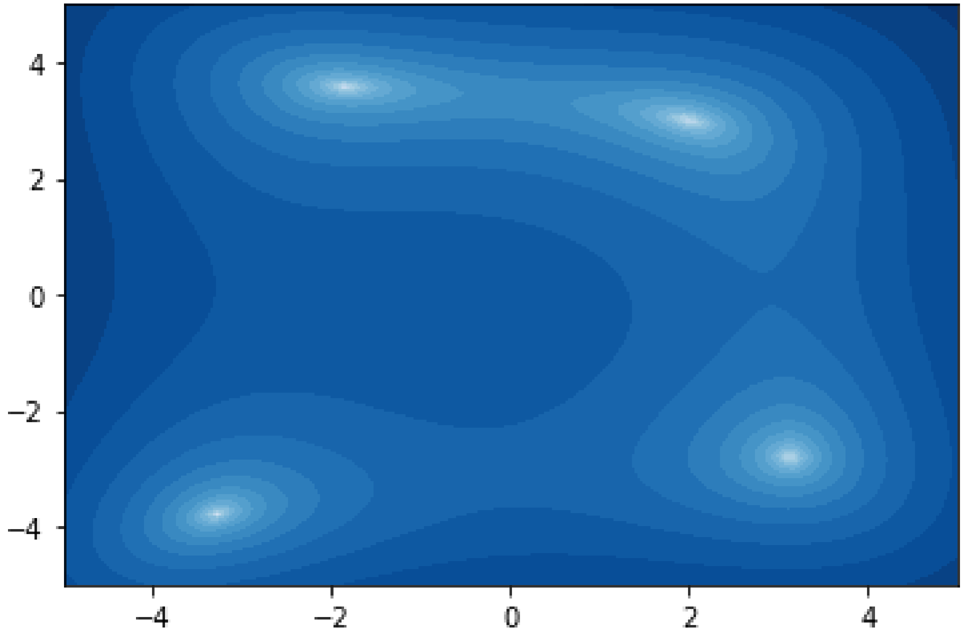
Selection

- Tournament selection
  - choose 2 individuals at random
  - select the best one
- Fitness proportionate selection
  - Selection probability is proportional to the fitness of the individual
  - Similar to the roulette wheel in a casino

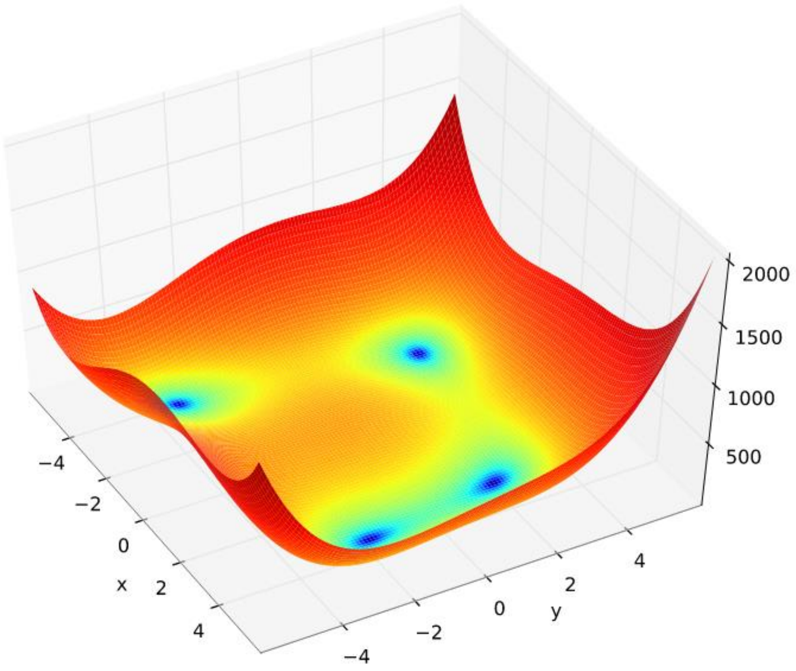


Population	Fitness
1	25.0
2	5.0
3	40.0
4	10.0
5	20.0

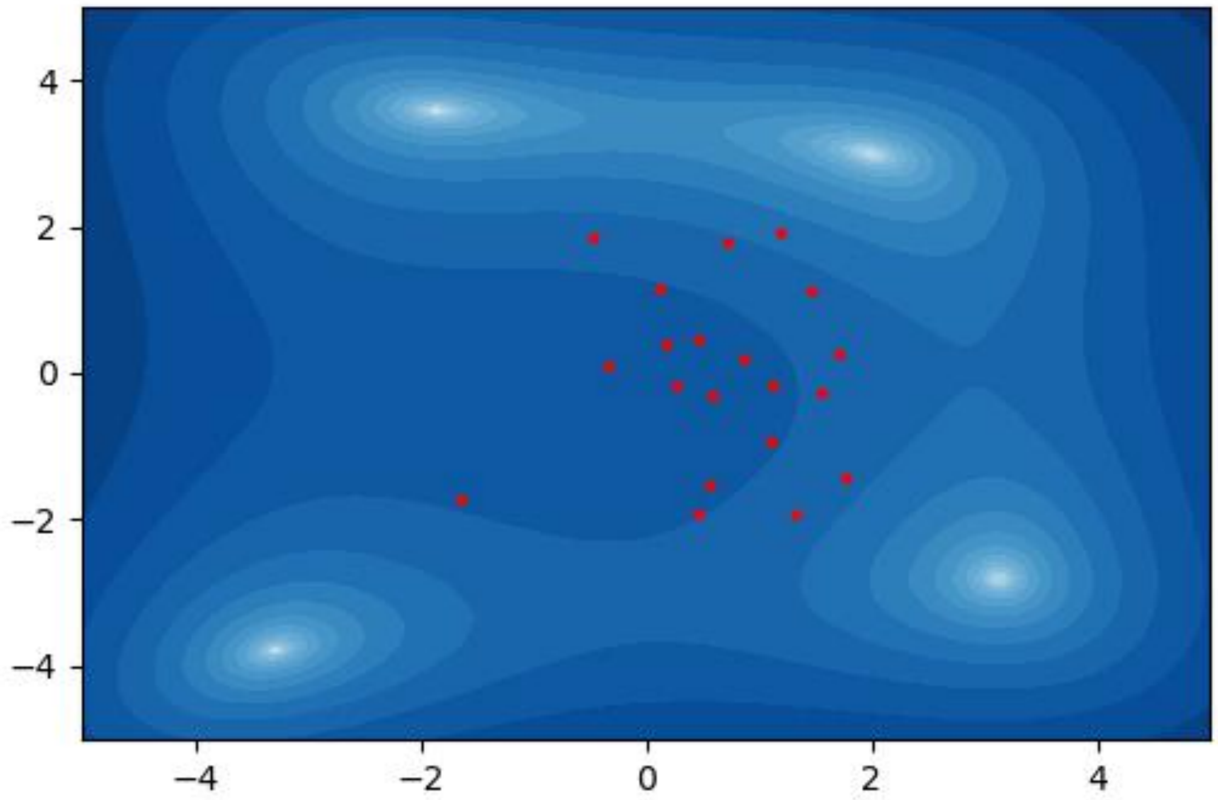
Himmelblau's multimodal function



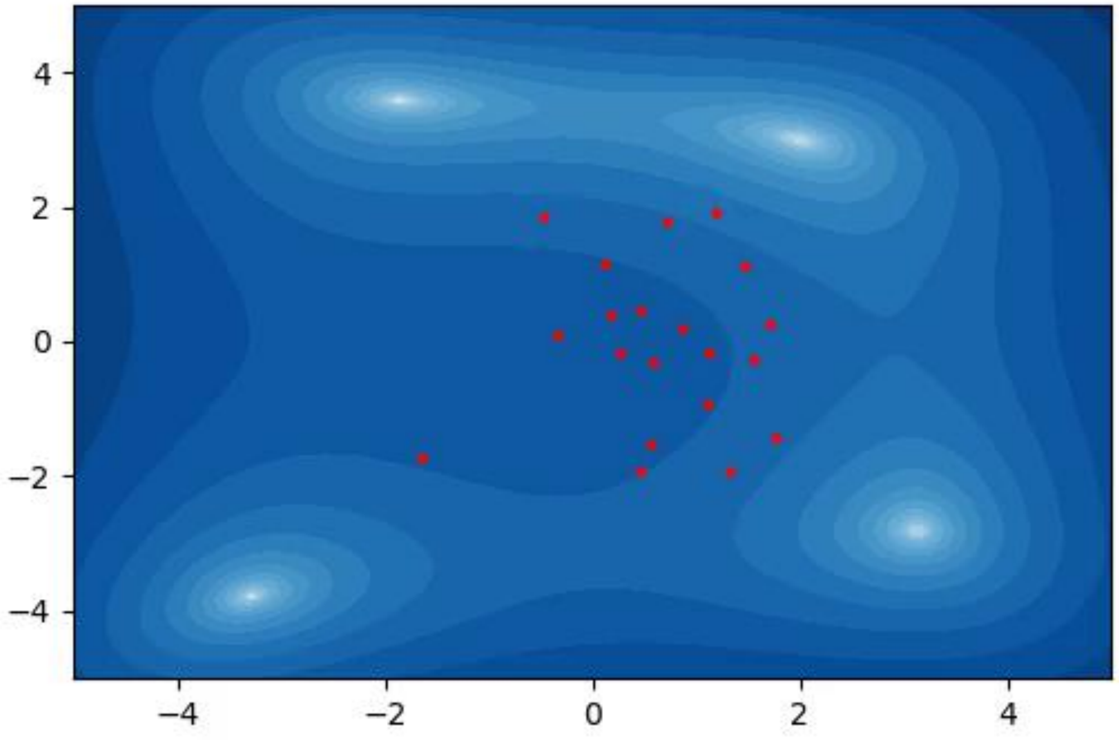
F(x,y) ≥ 0
4 points such as F(x,y)=0



Application to Himmelblau's function (part 1)

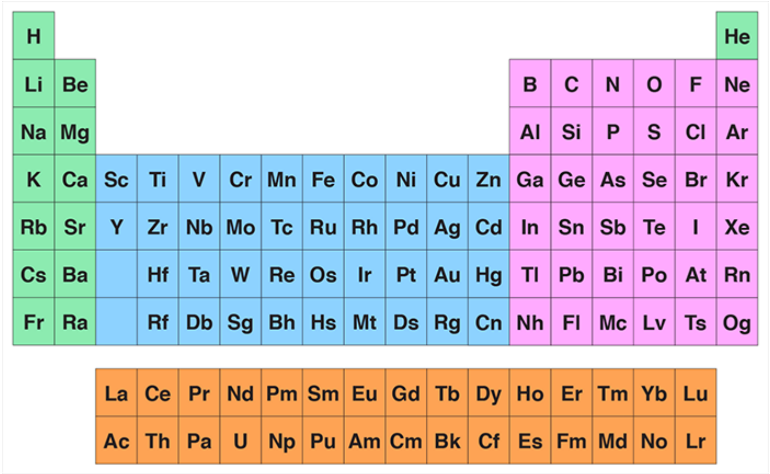


Application to Himmelblau's function (part 2)



2. Constrained Multiobjective Optimization

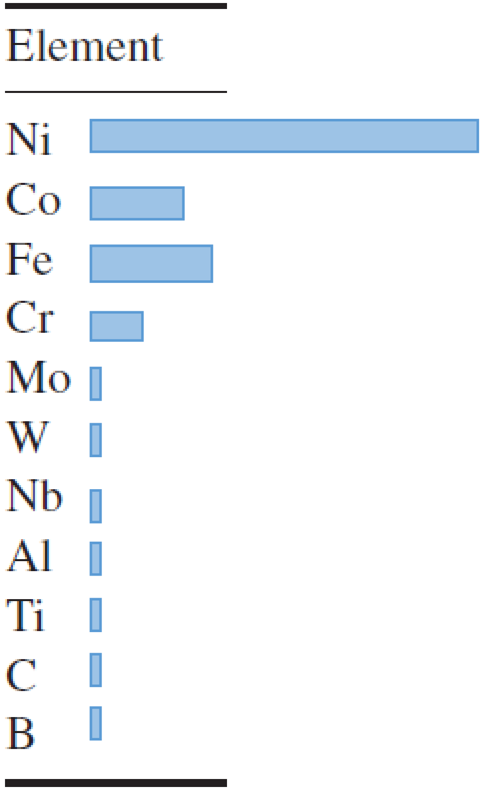
Superalloy Design



Element	Bounds	Step
Ni	Bal.	—
Co	0–21	
Fe	0–30	
Cr	15–25	
Mo	0–2	0.1
W	0–6	
Nb	0–5	
Al	0–5	
Ti	0–5	
C	0–0.15	0.01
B	0.005	

nickel-base superalloys: 10^{16} combinations

Objectives and Constraints



Equilibrium characteristics

- high-temperature stability
- Processability

Thermomechanical properties

- tensile strength
- creep resistance
- …

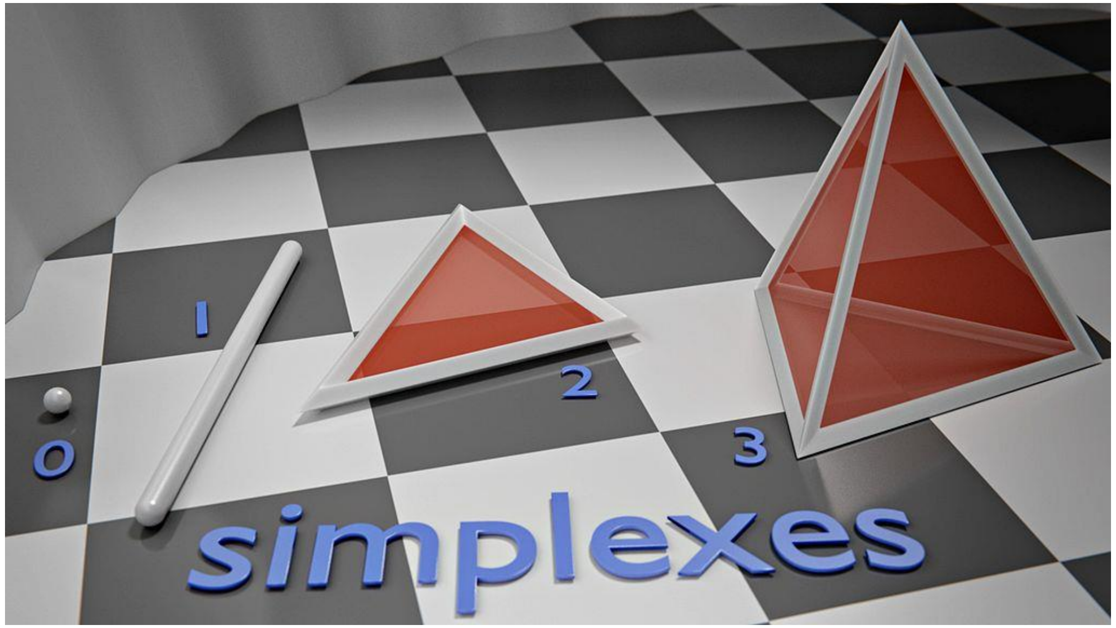
Constraint Handling in MO

- Only generate feasible solutions
- Repair unfeasible solutions
- Consider a constraint as an objective

Example of constraints: the simplex

(1) $x_1 + x_2 + x_3 + x_4 = 1$

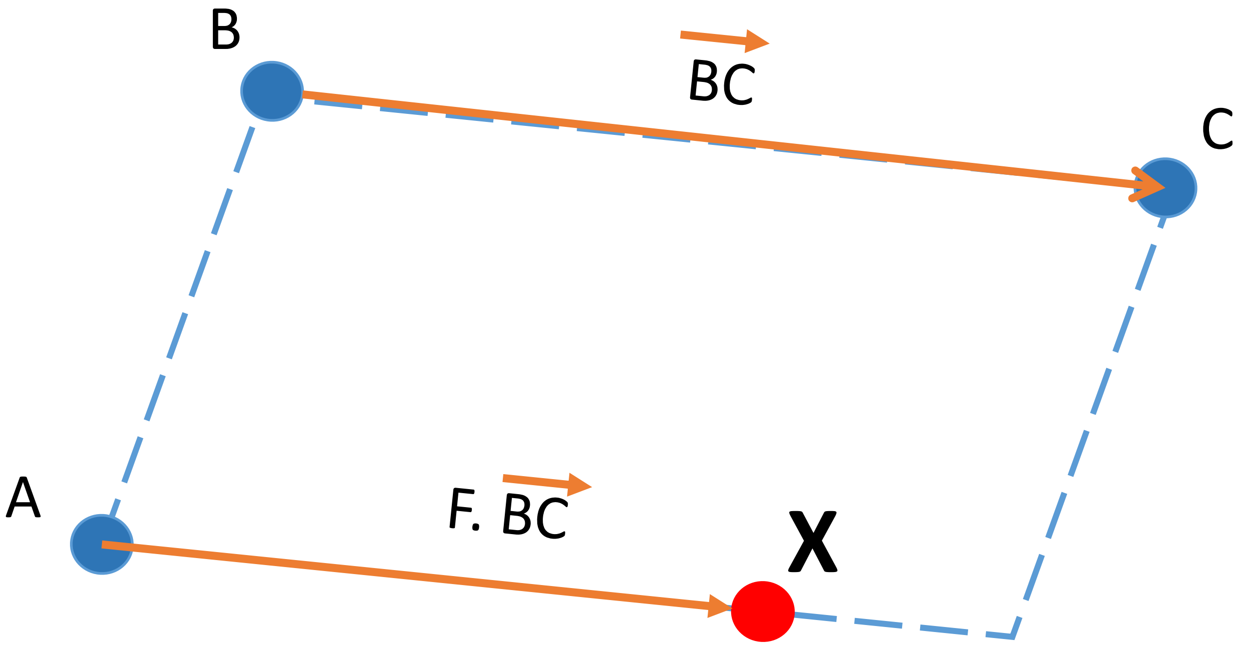
(2) $x_i \geq 0$



https://en.wikipedia.org/wiki/Simplex

A differentiel evolution operator

- Three parents
- A is the anchor
- Orientation given by B and C

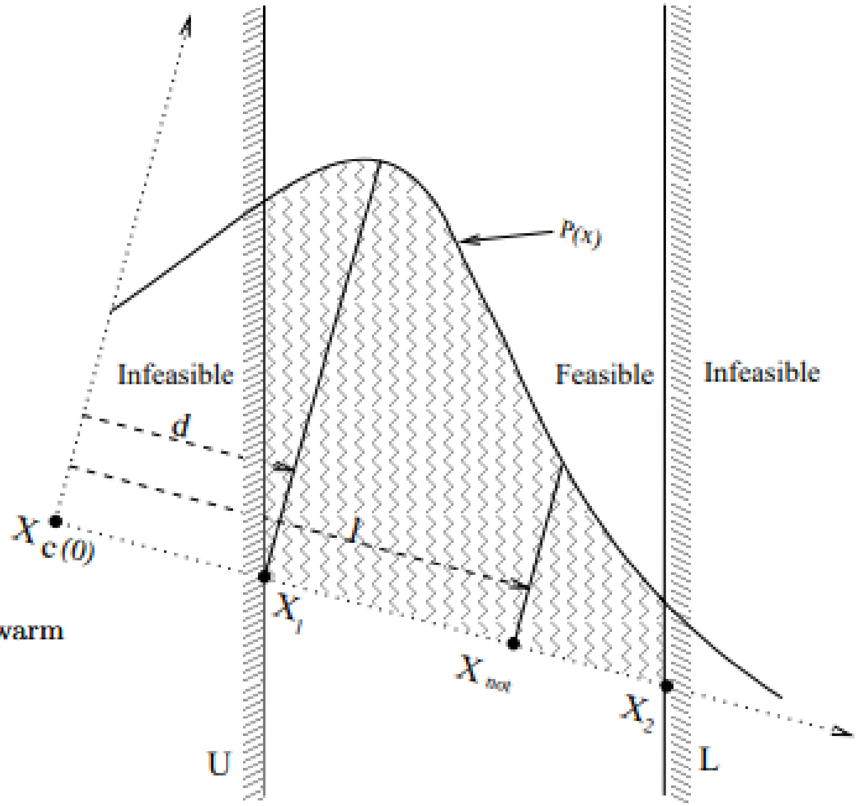


X lies in the same hyperplane as its parents

Second constraint

(1) $x_1 + x_2 + x_3 + x_4 = 1$

(2) $x_i \geq 0$



Inverse Parabolic Spread Distribution

Nikhil Padhye, Kalyan Deb, and Pulkit Mittal. Boundary handling approaches in particle swarm optimization. *Advances in Intelligent Systems and Computing*, 201, 12 2013.

G. Ramstein et al.
A multi-objective differential evolution approach for optimizing mixtures
OLA 2022

Constraint as objective

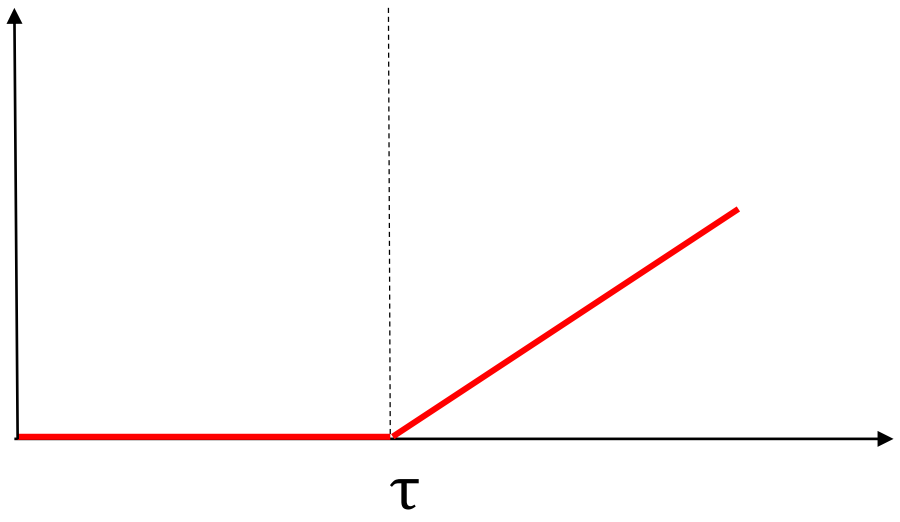
Example:

$G(X) < \tau$

Pseudo objective:

$f(X, \tau) = G(X) - \tau$ if $G(X) > \tau$

$= 0$ otherwise



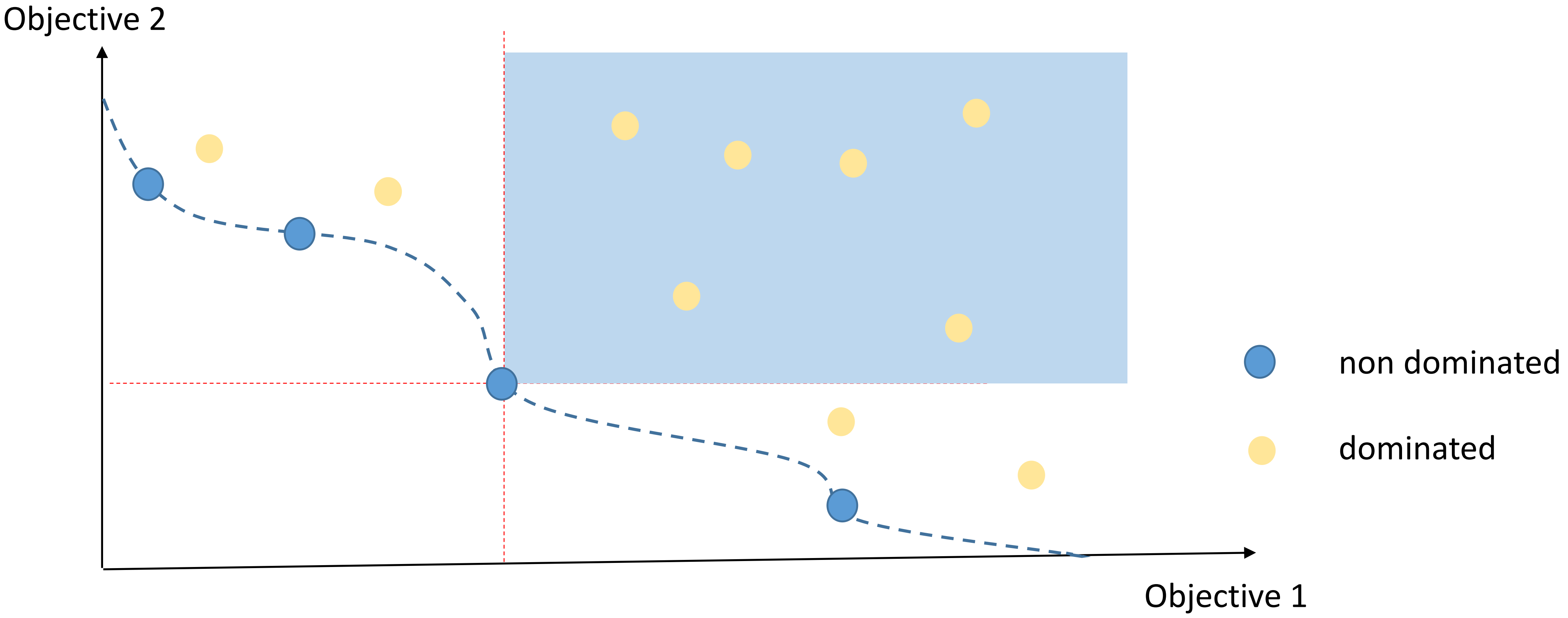
Objectives and Dominance

x1 dominates x2 if:

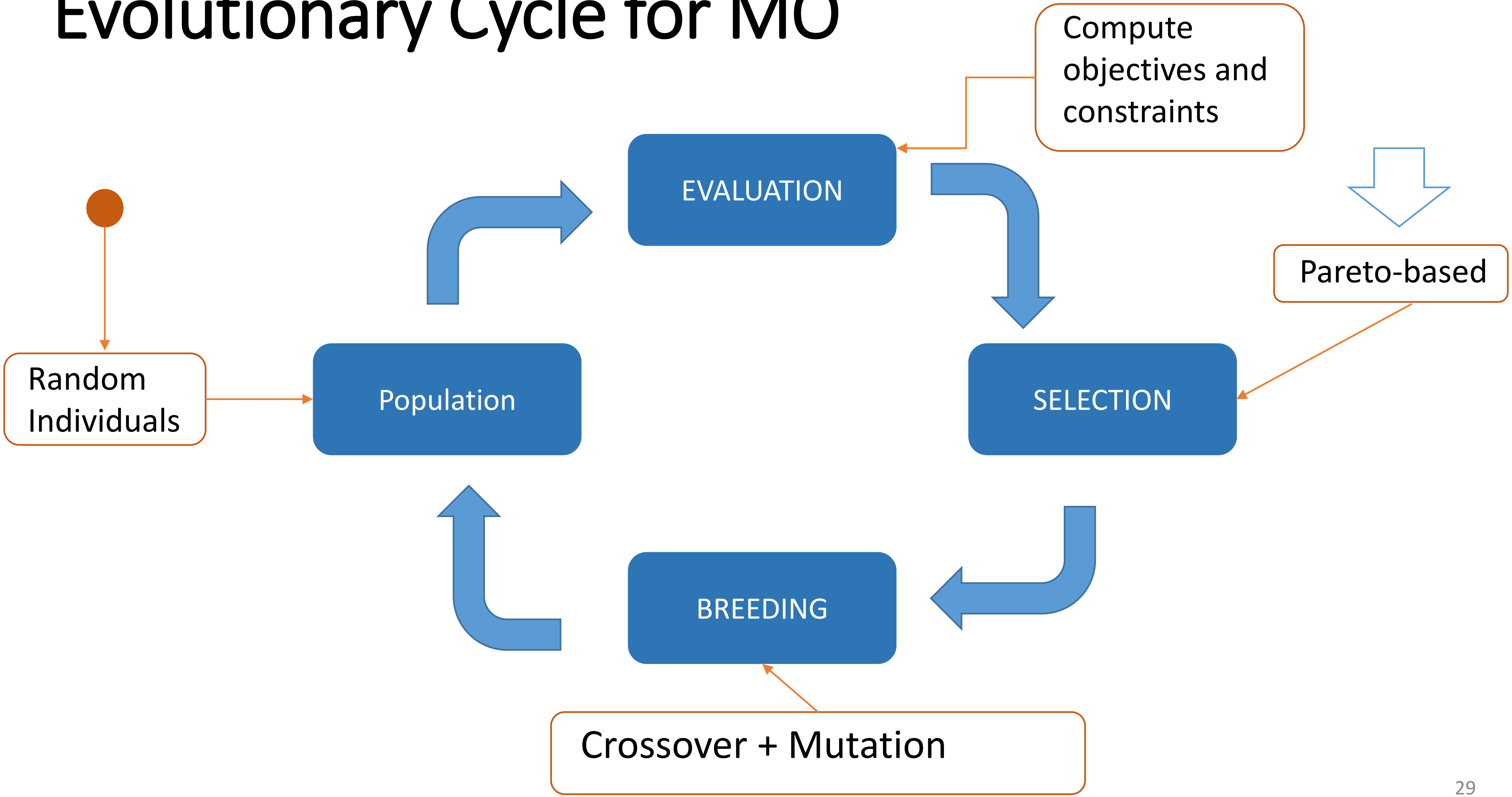
- solution x1 is not worse than x2 in all objectives
- solution x1 is strictly better than x2 in at least one objective

Objectives and Pareto Front

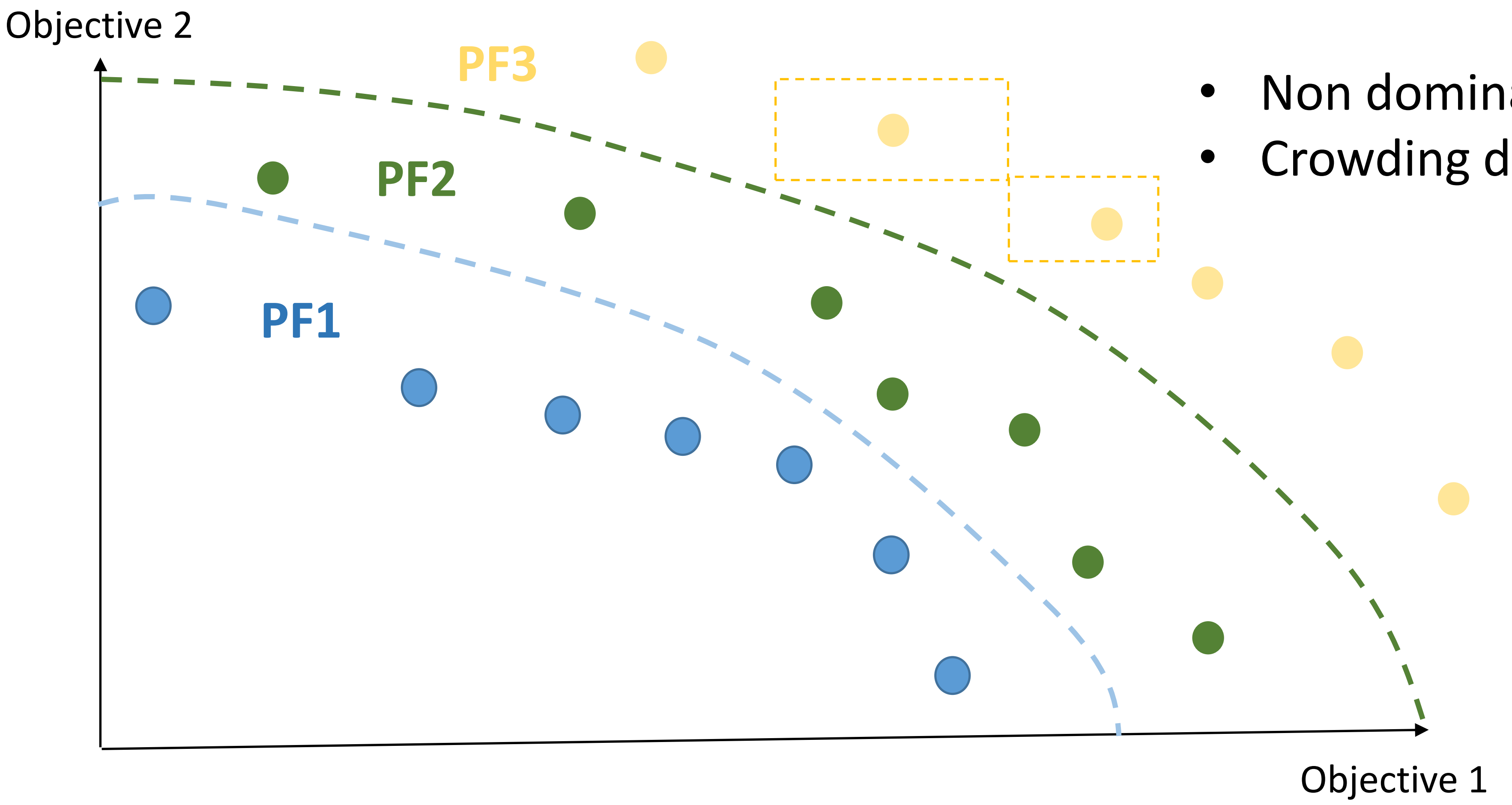
Multiobjective: a set of solutions



Evolutionary Cycle for MO



Pareto-based Selection (NSGA II)



- Non dominated Sorting
- Crowding distance Sorting

A bi-objective Optimization example

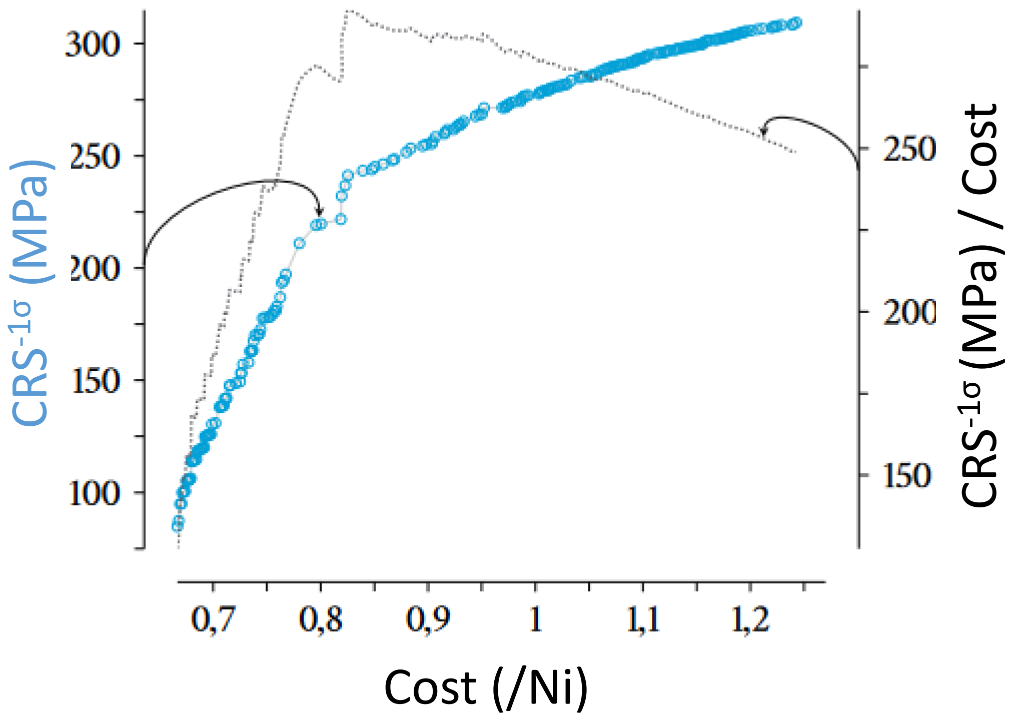
Constraints:

- alloys should only be constituted of γ, γ' and $M_{23}C_6$ phases at both 700 °C and 750 °C;
- a maximum amount of 25 mol% γ' at 750 °C is tolerated to ensure weldability;
- a minimum amount of 24 at% Cr in the (γ) matrix at 750 °C is required for corrosion resistance.

Objectives:

- the minimisation of the heat price, estimated on the basis of the cost of individual elements as found on the market on the date of study, relative to the cost of nickel;
- the maximisation of the 'lowered' creep rupture stress, $CRS^{-1\sigma}$, at 750 °C after 10^3 h.

Results

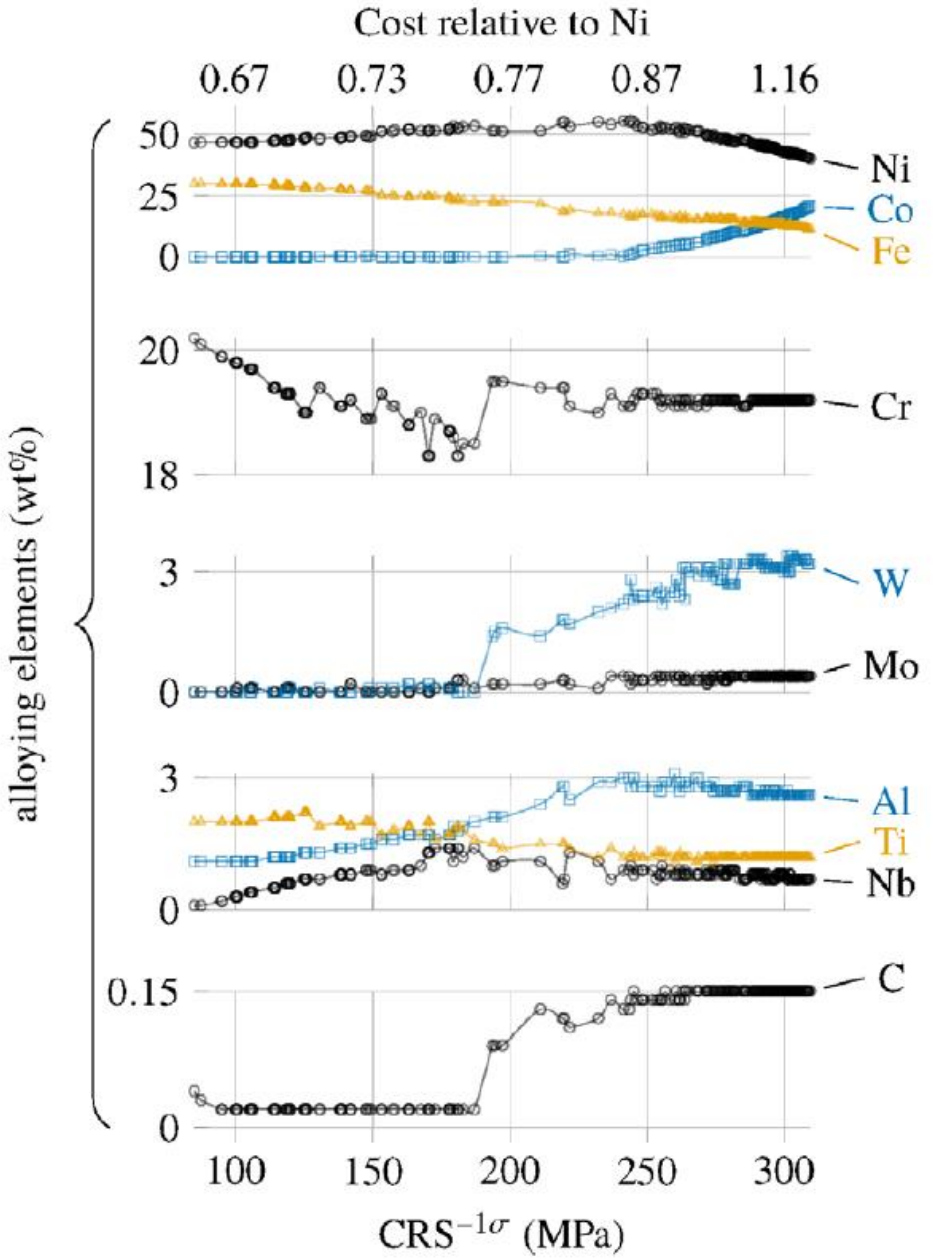


Pareto Front: cost vs CRS

- the stronger an optimized alloy, the greater its cost
- The creep resistance varies with the price range of the alloy.

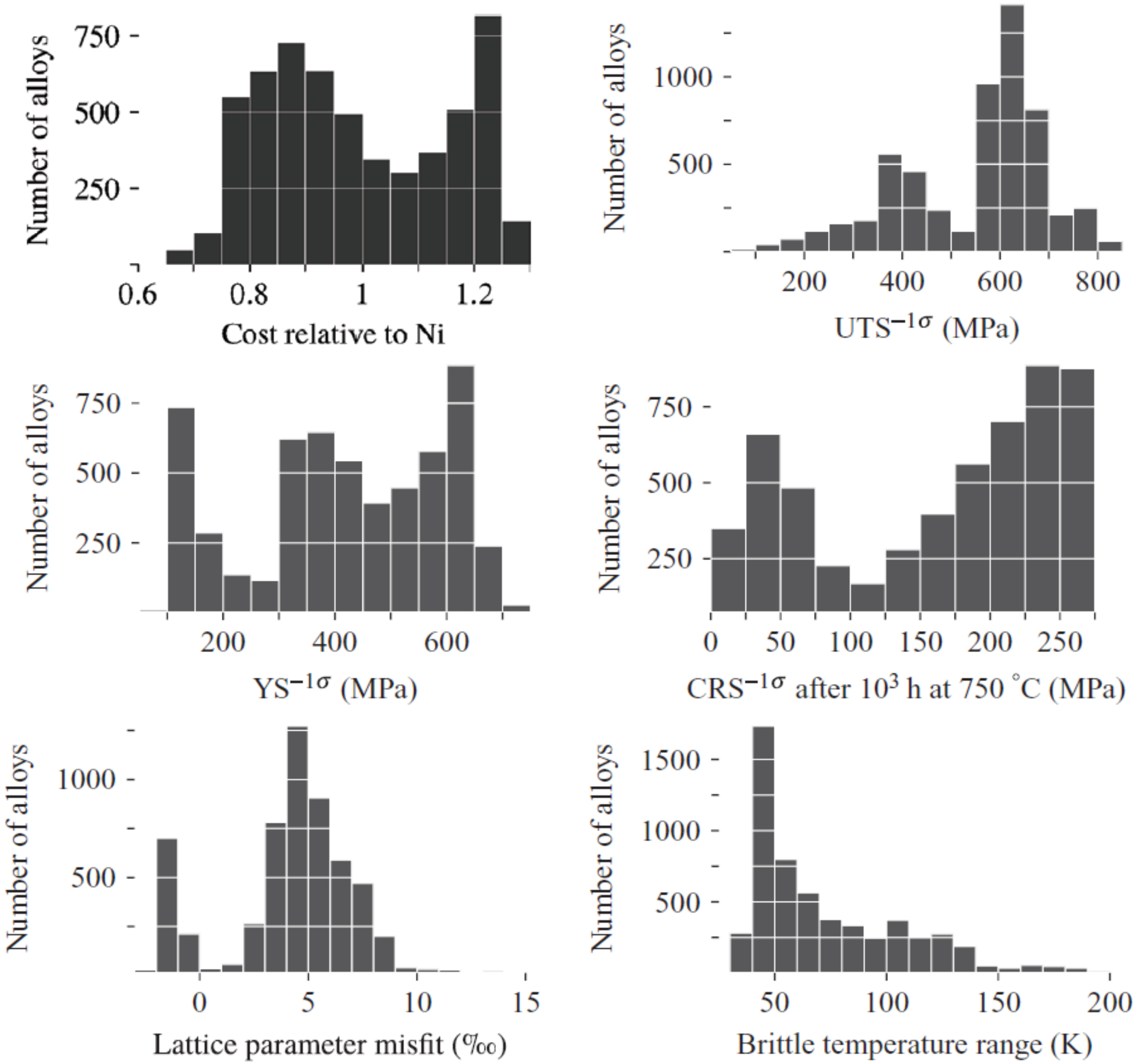
A quick analysis of the results:

gradual substitution of nickel by cobalt
for the most creep-resistant alloys, reaching
almost 21 wt% for the most expensive ones



For more details and
a 'real-world' case (6 objectives):

Edern Menou *et al* 2016
Modelling Simul. Mater. Sci. Eng.
24 055001



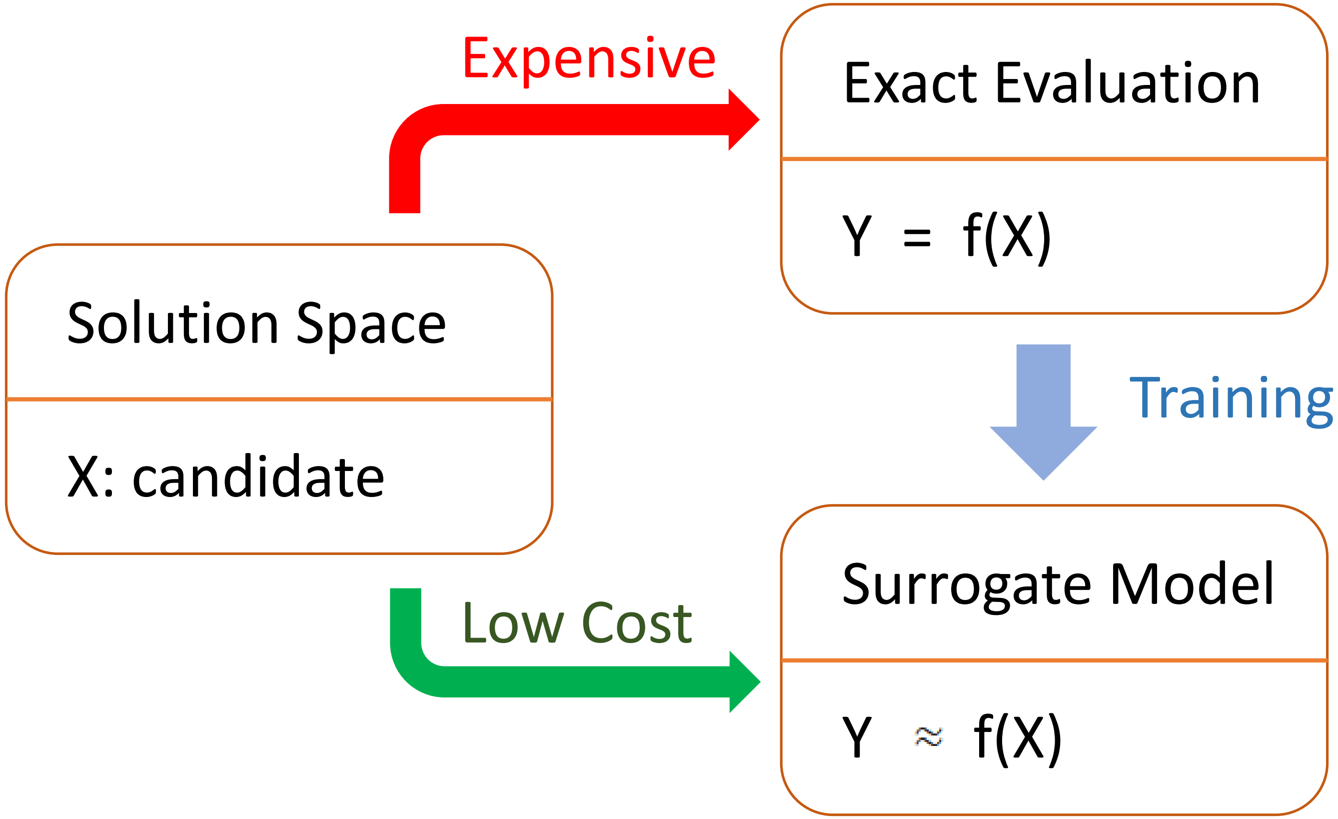
3. Bayesian Optimization

Bayesian Optimization

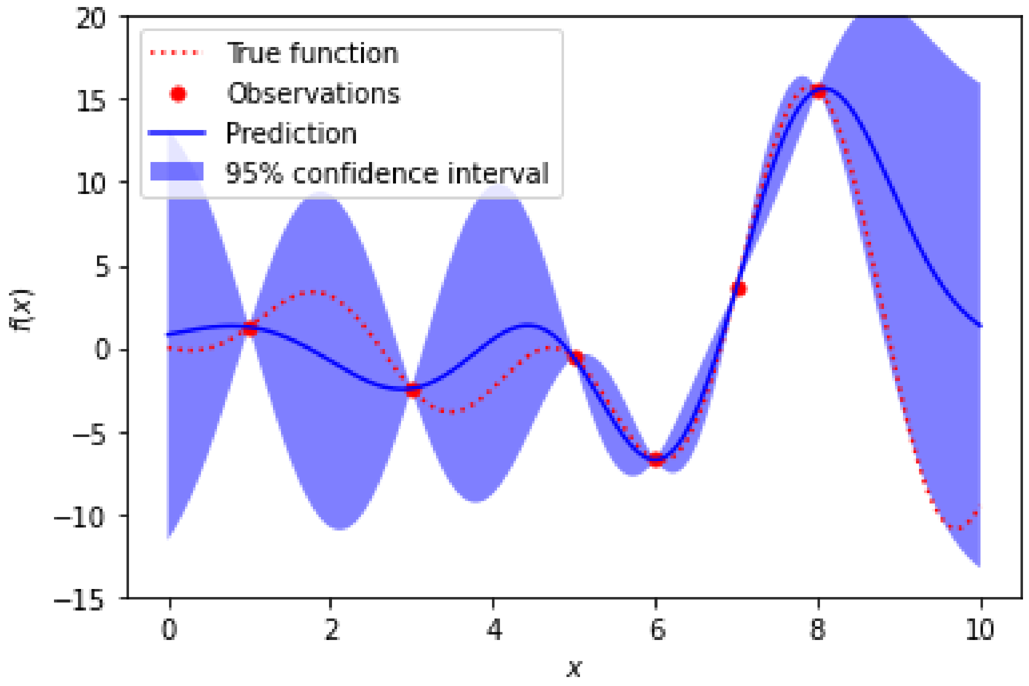
CMO works well, but the exploration of the decision space can be very time consuming
(e.g. thermodynamic properties of alloys)

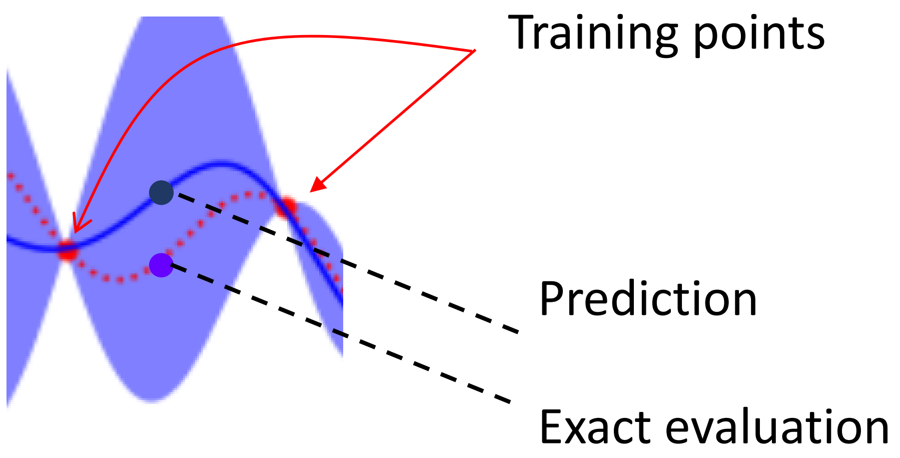
⇨ Introduction of a surrogate model

Surrogate Model

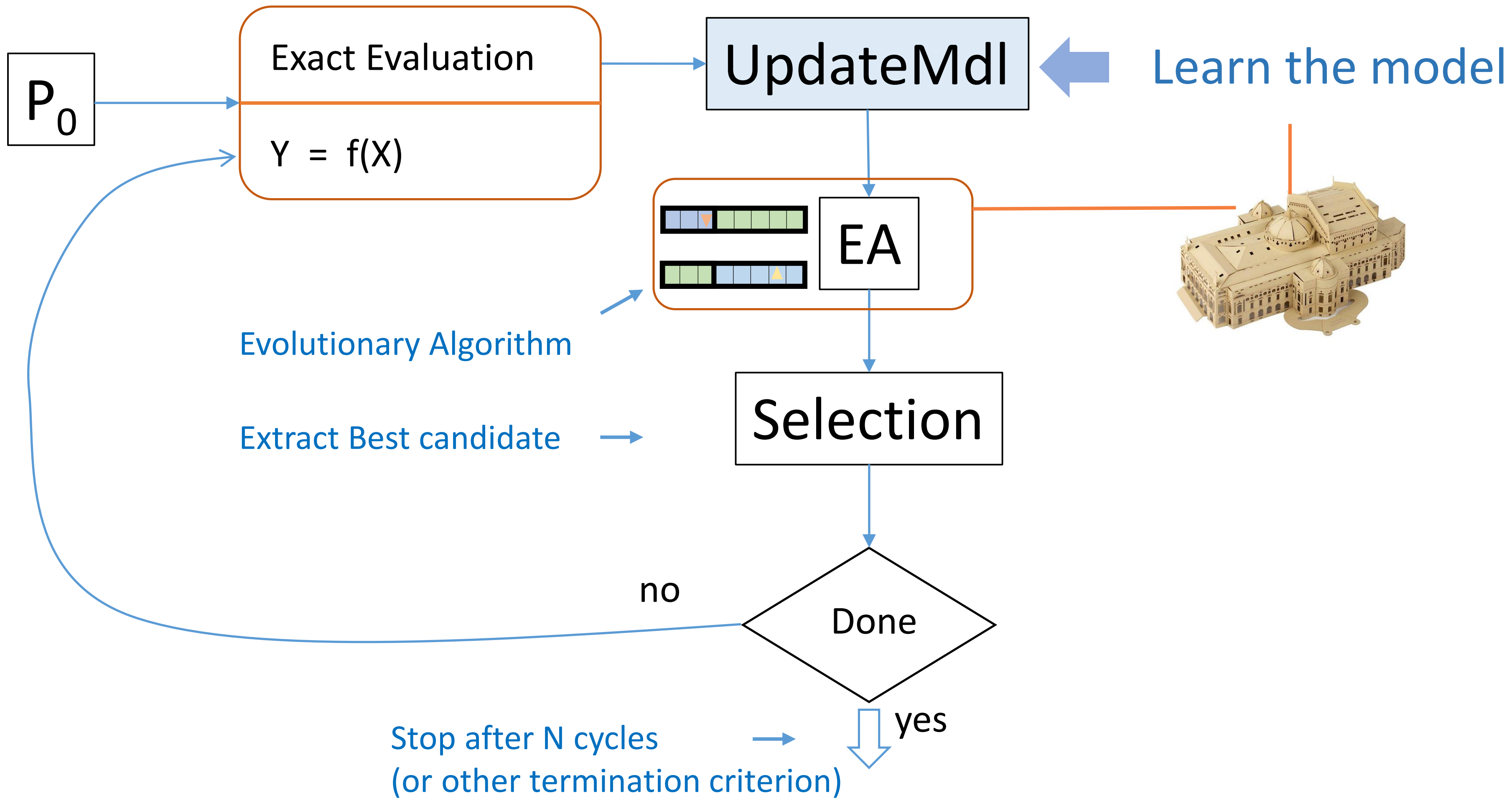


Gaussian Process Regression (Kriging)



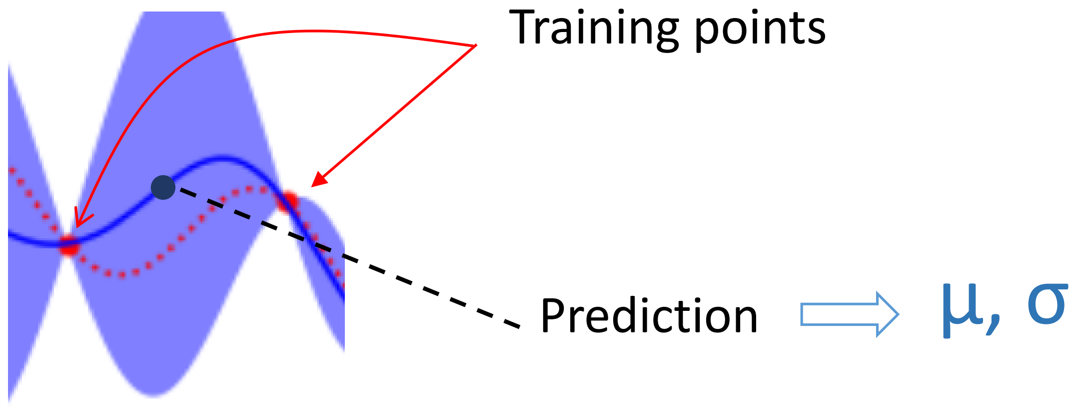


GENERAL ARCHITECTURE



Acquisition Functions

evaluate the usefulness of design points for achieving objectives/constraints



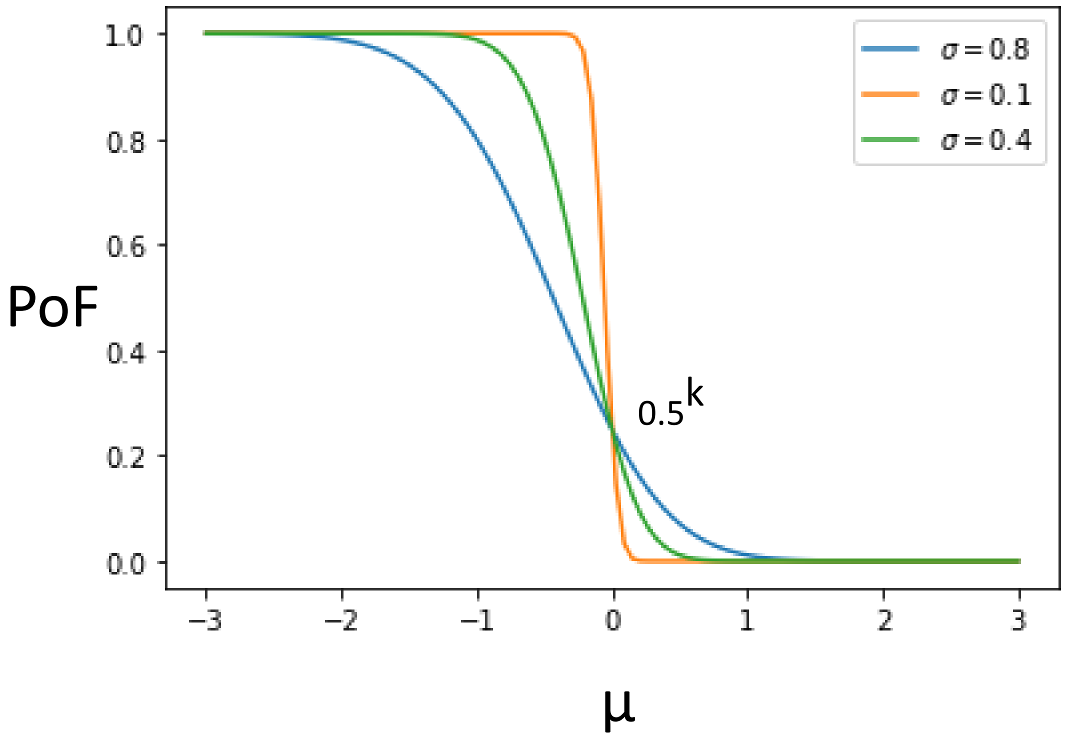
What are the best candidates in current EA population ?

Probability of Feasibility (PoF)

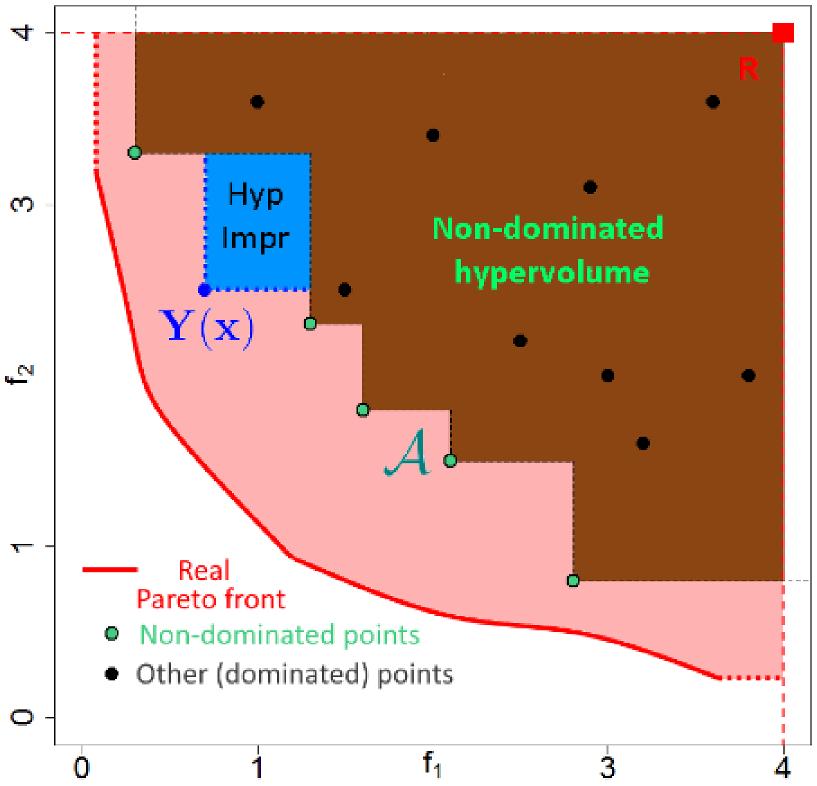
Given k constraints,

Knowing k couples (μ, σ),

What is PoF ?

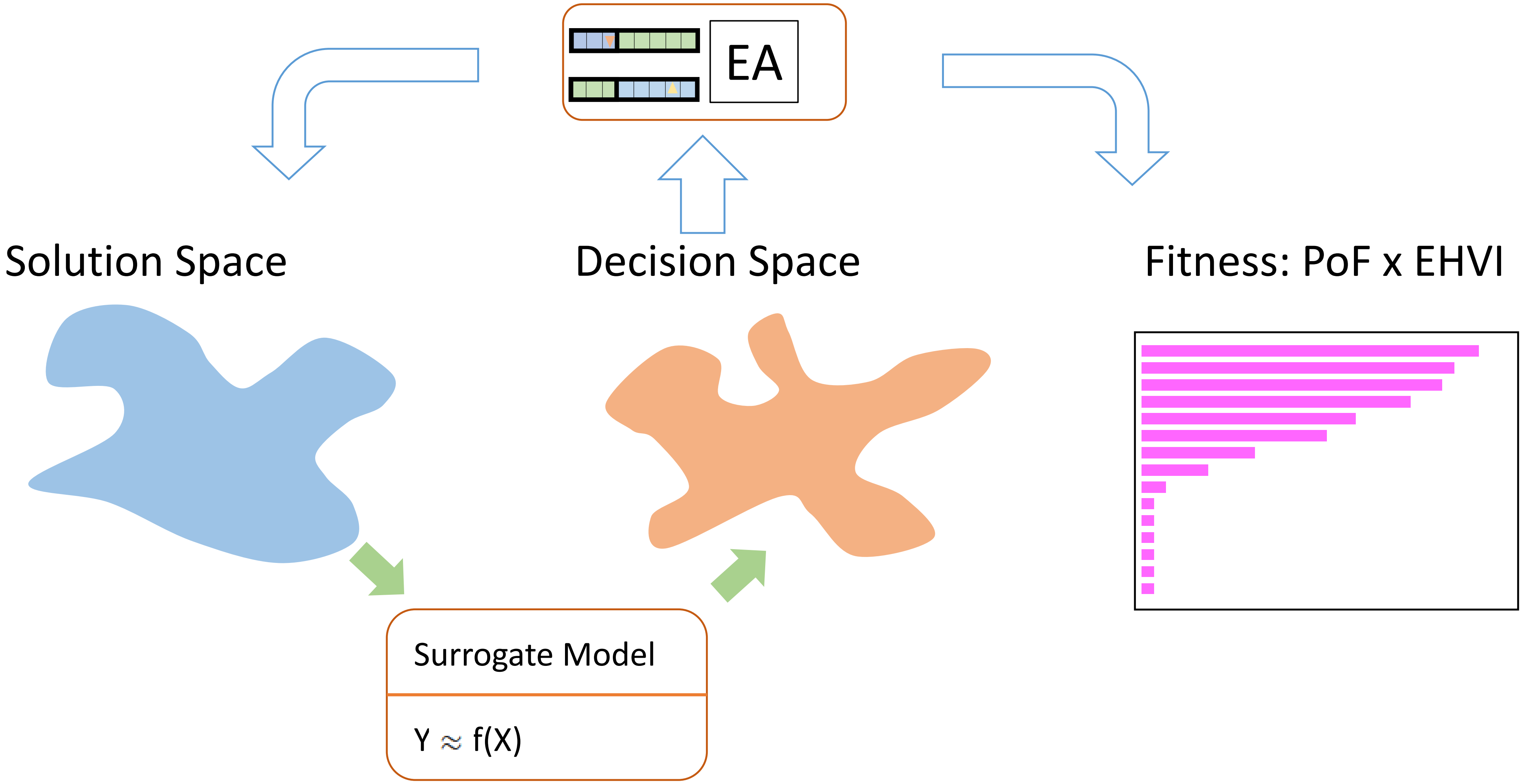


EHVI



Hyp
Impr
Non-dominated
hypervolume
R
Y(x)
A
Real
Pareto front
Non-dominated points
Other (dominated) points
f1
f2
0
1
3
4
0
1
3
4

Data flow


EA
Solution Space
Decision Space
Fitness: PoF x EHVI
Surrogate Model
Y ≈ f(X)

Illustrative Example

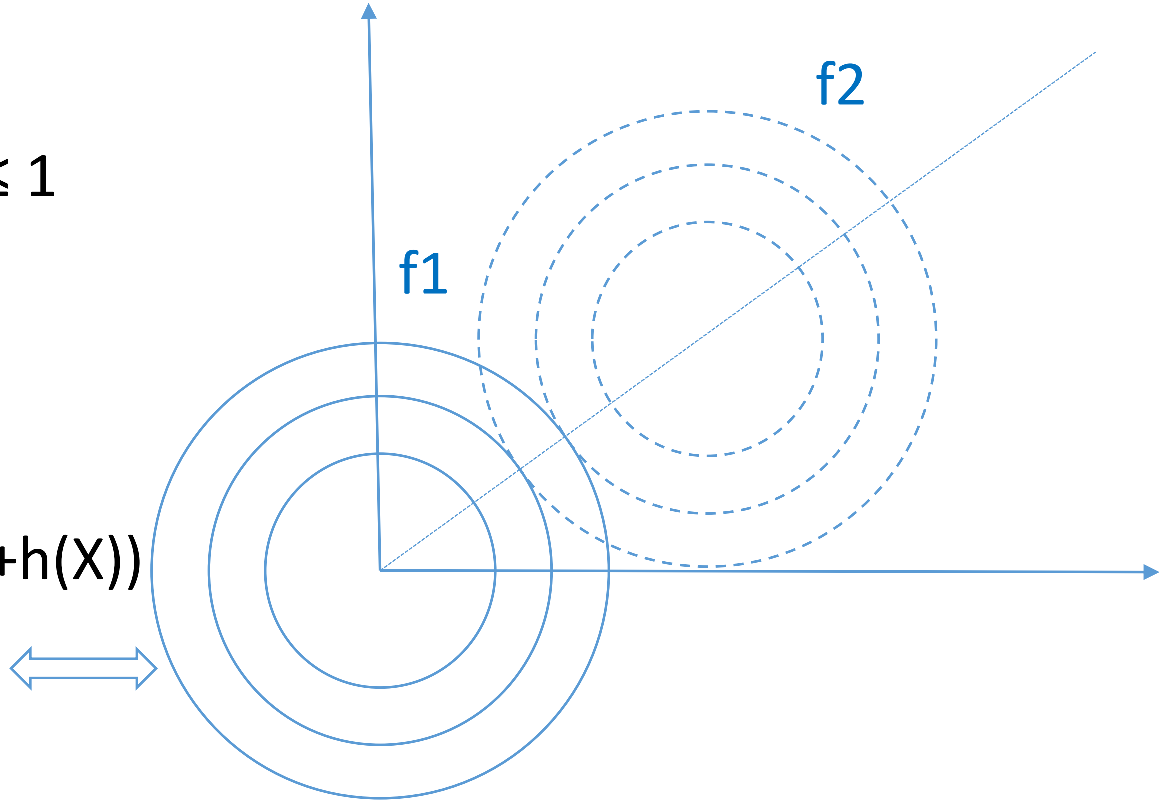
Minimize:

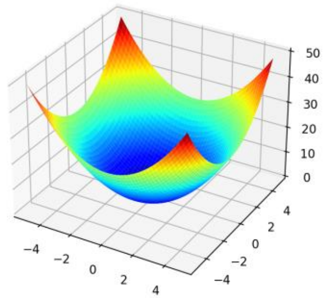
$X = (x_1, x_2, x_3, x_4, x_5, x_6), 0 \leq x_i \leq 1$

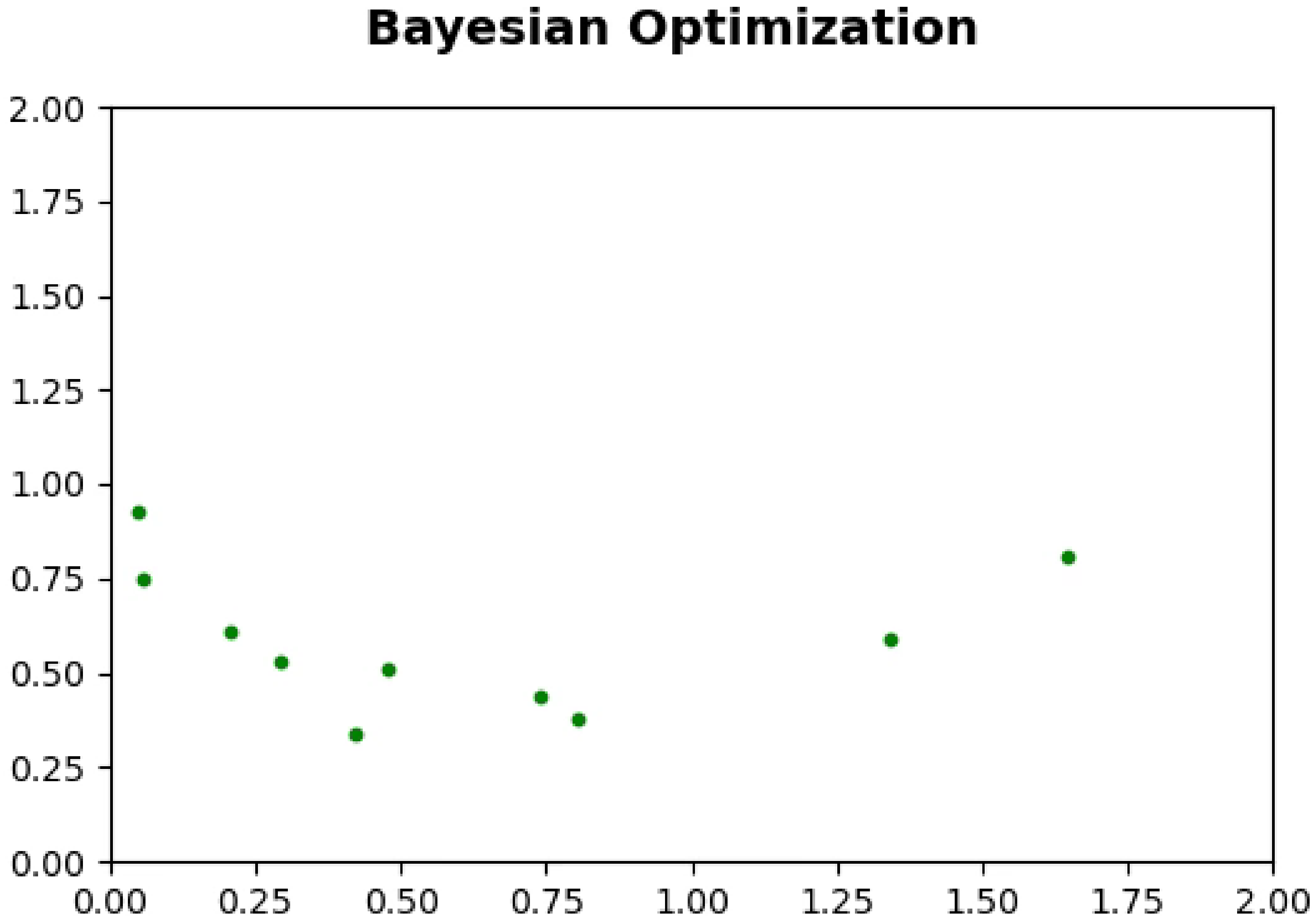
$h(X) = 10.(x_4^2 + x_5^2 + x_6^2)$

$f_1(X) = (2x_1^2 + 2x_2^2)(1+h(X))$

$f_2(X) = (2(x_1 - 0.5)^2 + 2(x_2 - 0.5)^2)(1+h(X))$

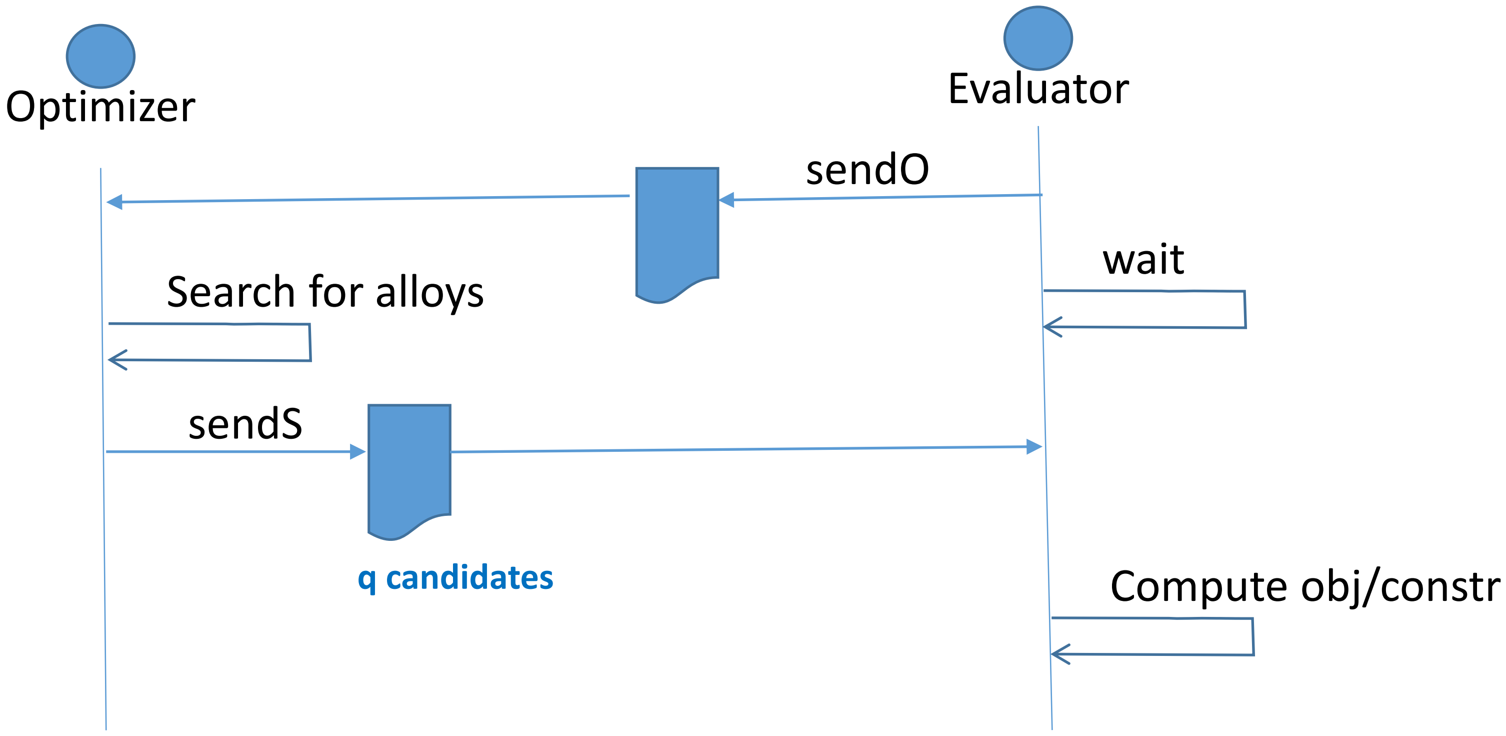




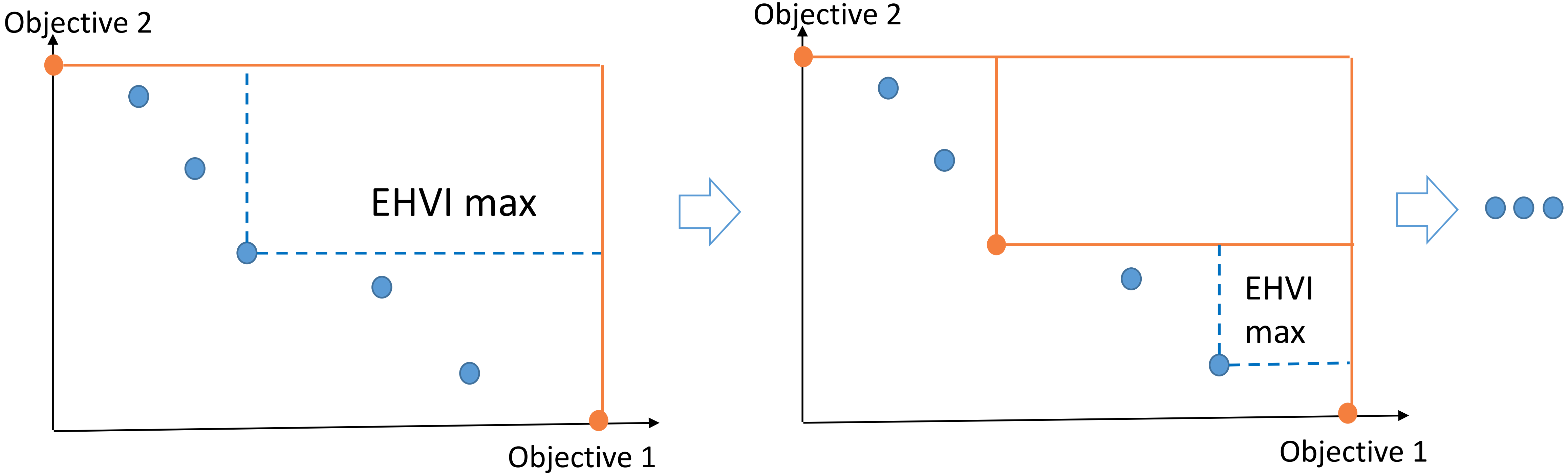

Bayesian Optimization
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

Batch Bayesian Optimization

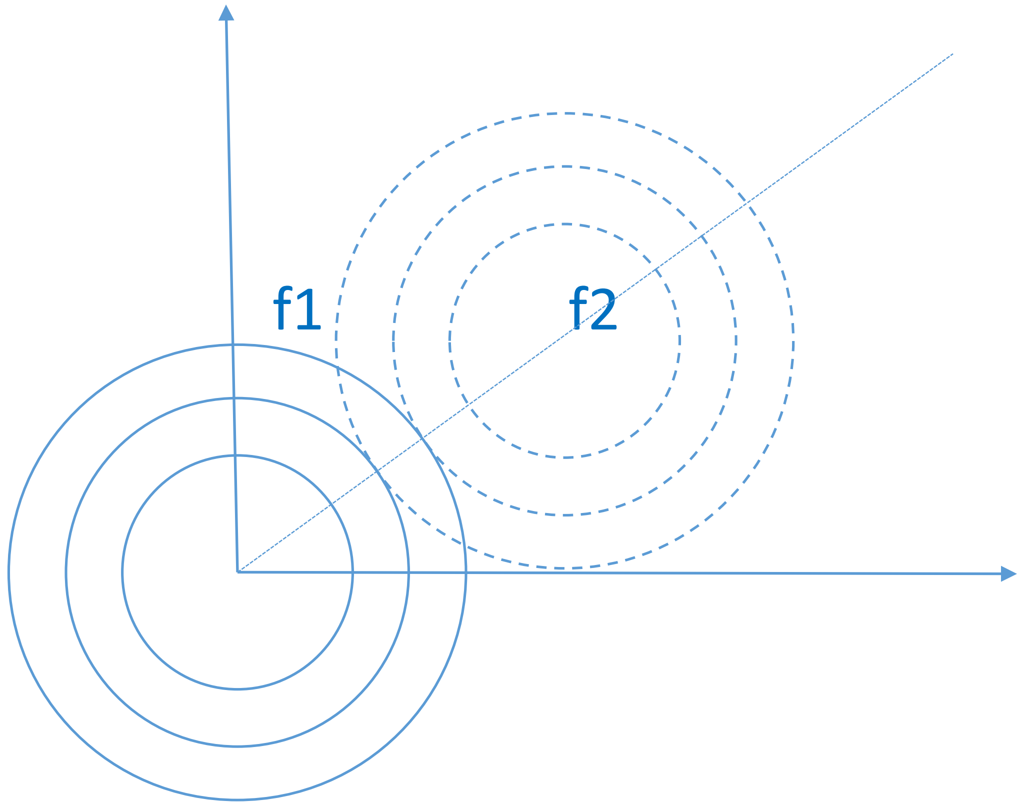
Extract q candidate solutions instead of a unique one at each iteration

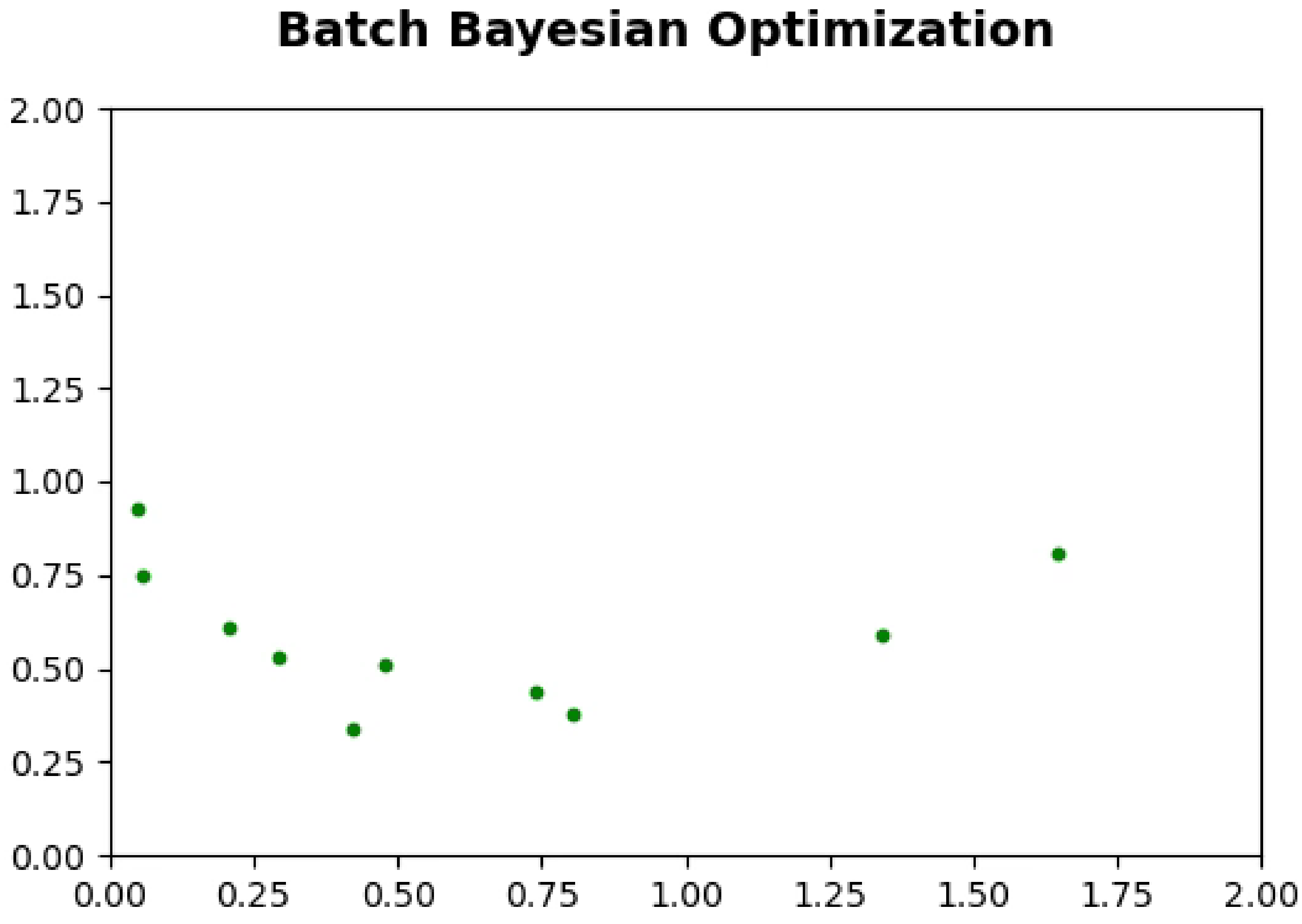


A simple technique: Kriging Believer



Illustrative Example




Batch Bayesian Optimization
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

Conclusion

- Optimization is a major issue in many fields of activity
- classical trial-and-error alloy development can be ineffective
- exploration of a search space of billions of alloys is challenging
- Handling multiple constraints and many objectives.
- Bayesian Optimization techniques reduce the computational cost.
- multi-criteria decision analysis: choosing an alloy among several thousand Pareto-optimal ones.