



# Metti<sup>5</sup>

Advanced Spring School

Thermal Measurements & Inverse Techniques  
Station Biologique de ROSCOFF  
June 13-18, 2011

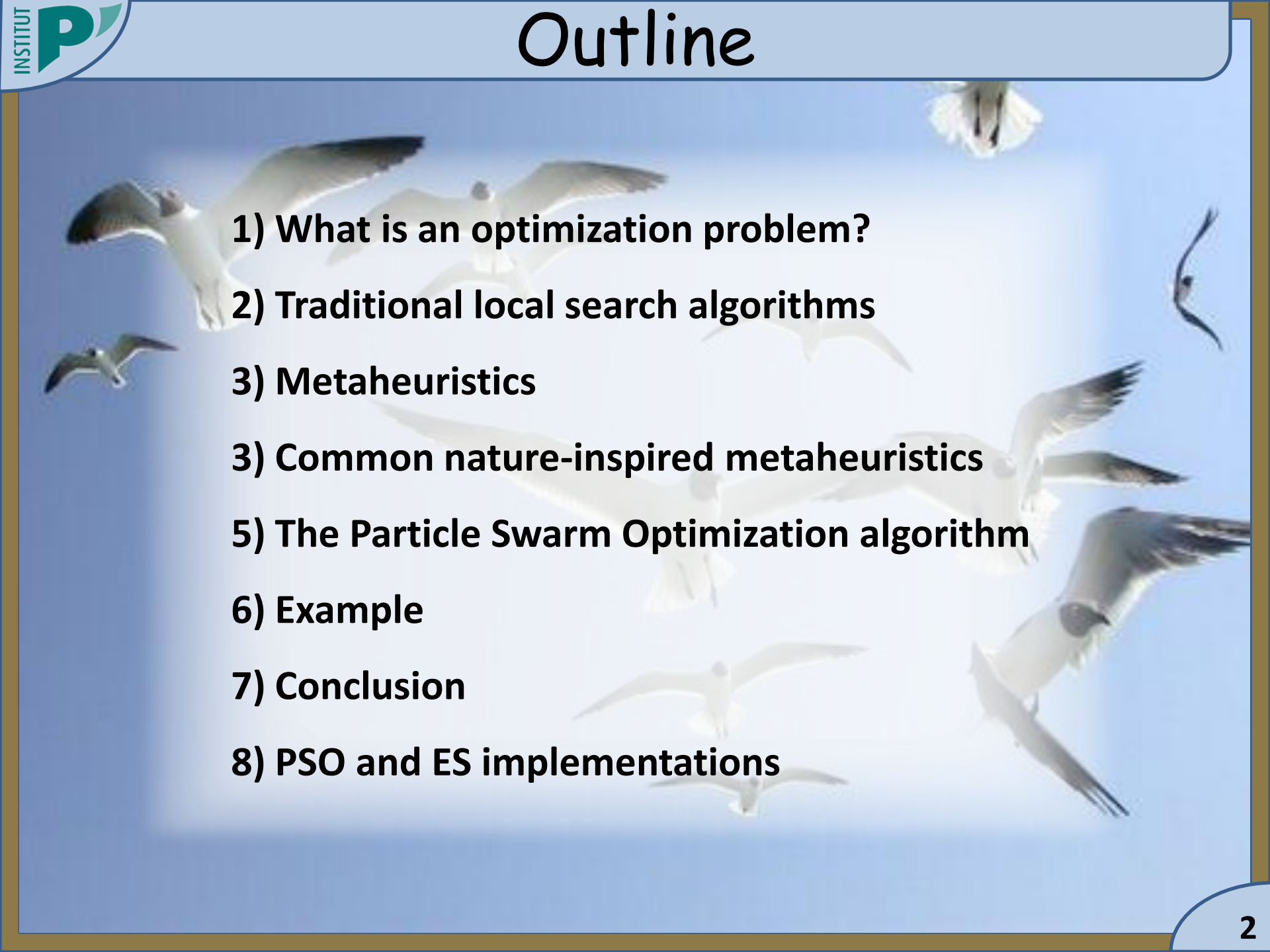


## Zero order optimization algorithms

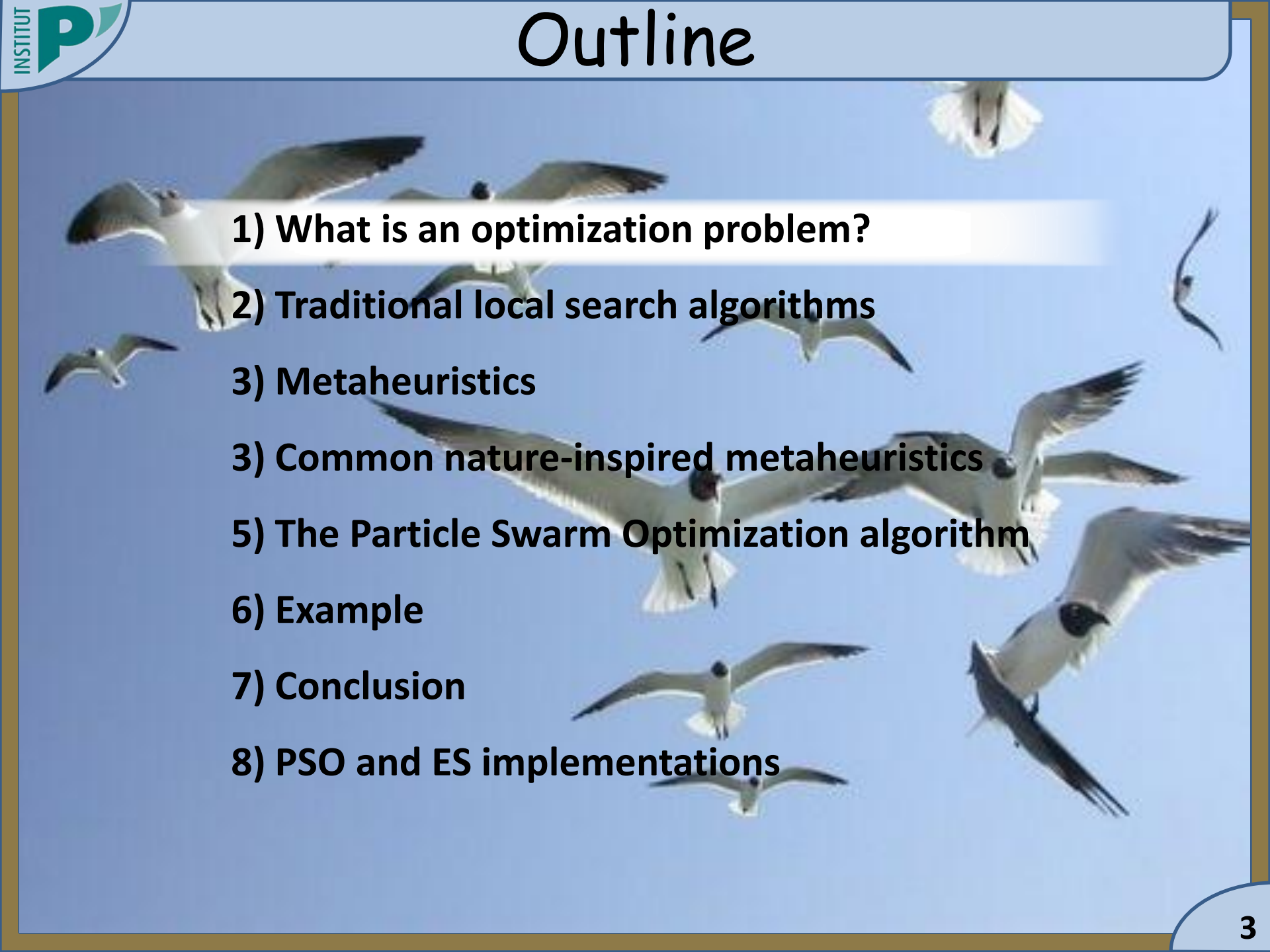
Tutorial 2 - Emmanuel RUFFIO\*, Daniel PETIT, Didier SAURY, Manuel GIRAULT\*  
emmanuel.ruffio@let.ensma.fr, manuel.girault@let.ensma.fr

Institut P' (UPR CNRS 3346)  
CNRS, ENSMA, Université de Poitiers  
Département fluides, thermique, combustion  
ENSMA - BP. 40109  
86961 Futuroscope Chasseneuil



- 
- 1) What is an optimization problem?**
  - 2) Traditional local search algorithms**
  - 3) Metaheuristics**
  - 3) Common nature-inspired metaheuristics**
  - 5) The Particle Swarm Optimization algorithm**
  - 6) Example**
  - 7) Conclusion**
  - 8) PSO and ES implementations**

# Outline

- 
- 1) What is an optimization problem?**
  - 2) Traditional local search algorithms**
  - 3) Metaheuristics**
  - 3) Common nature-inspired metaheuristics**
  - 5) The Particle Swarm Optimization algorithm**
  - 6) Example**
  - 7) Conclusion**
  - 8) PSO and ES implementations**

# An optimization problem

**Continuous optimization**

$$\beta = (\beta_1, \dots, \beta_{N_\beta})^T \in R^{N_\beta}$$

**Single objective optimization**

$$J(\beta) \in R$$

**Boundary constraints**

$$L_i \leq \beta_i \leq U_i \quad \forall i \in [1; N_\beta]$$

**Search space**

$$S = [L_1; U_1] \times \dots \times [L_{N_\beta}; U_{N_\beta}]$$

~~Multiobjective  
Multimodale  
Dynamic  
Combinatorial~~

~~Linear constraints  
 $A \cdot \beta \leq B$~~

~~Non-linear constraints  
 $f(\beta) = 0$~~

~~Soft constraints~~

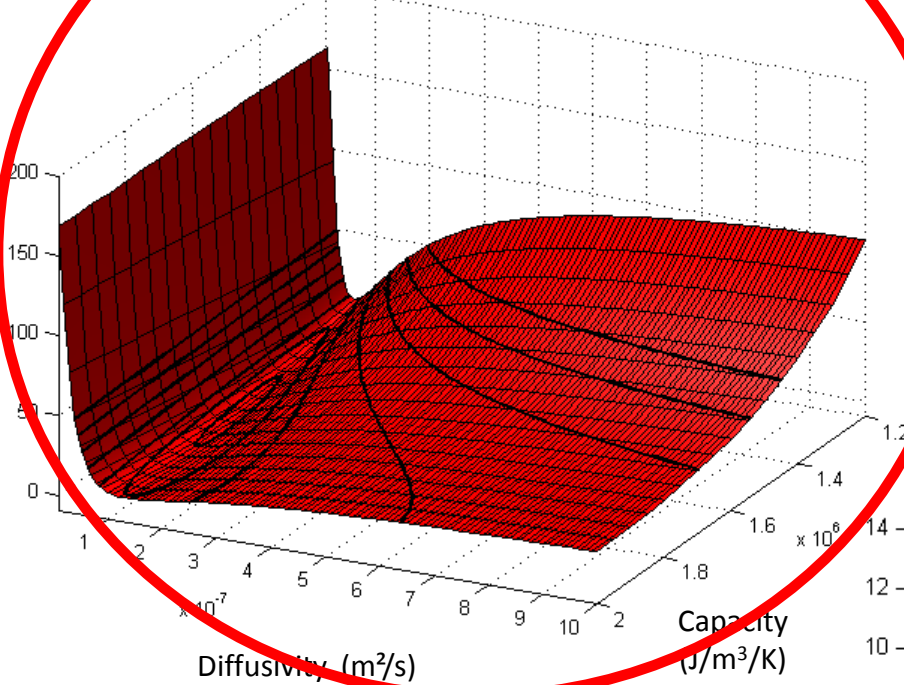
**Our optimization problem**

$$\hat{\beta} = \arg[\min_{\beta \in S} J(\beta)]$$

# What kind of objective function?

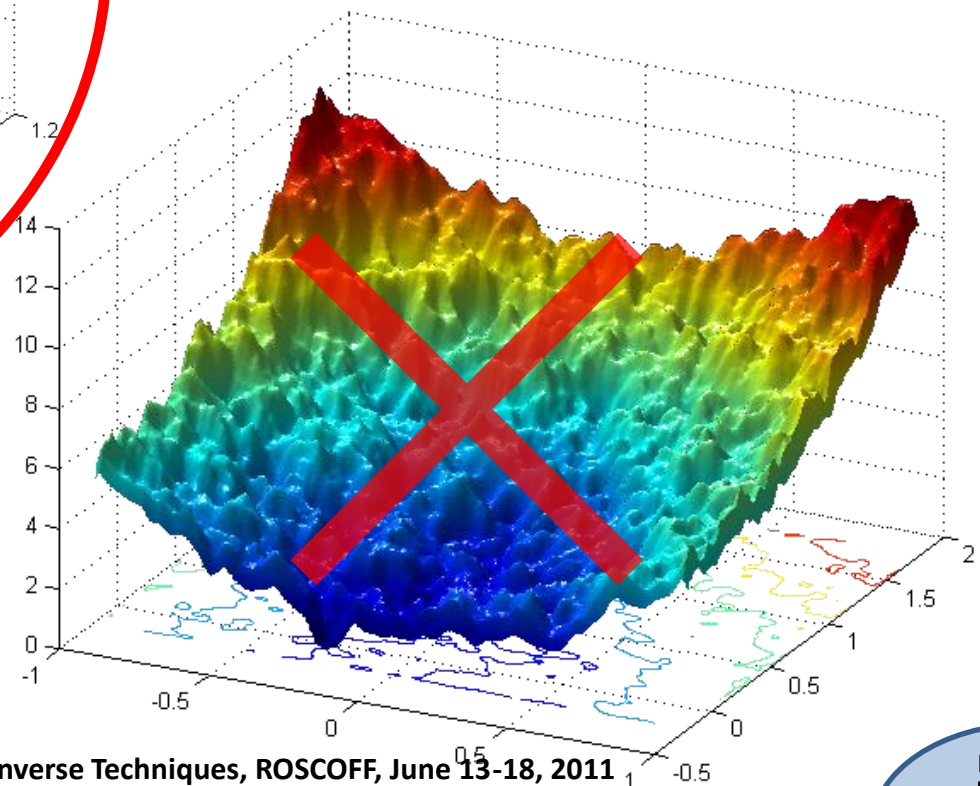
## The simple case

Ex: Least squares  
(Flash method for parameter estimation)

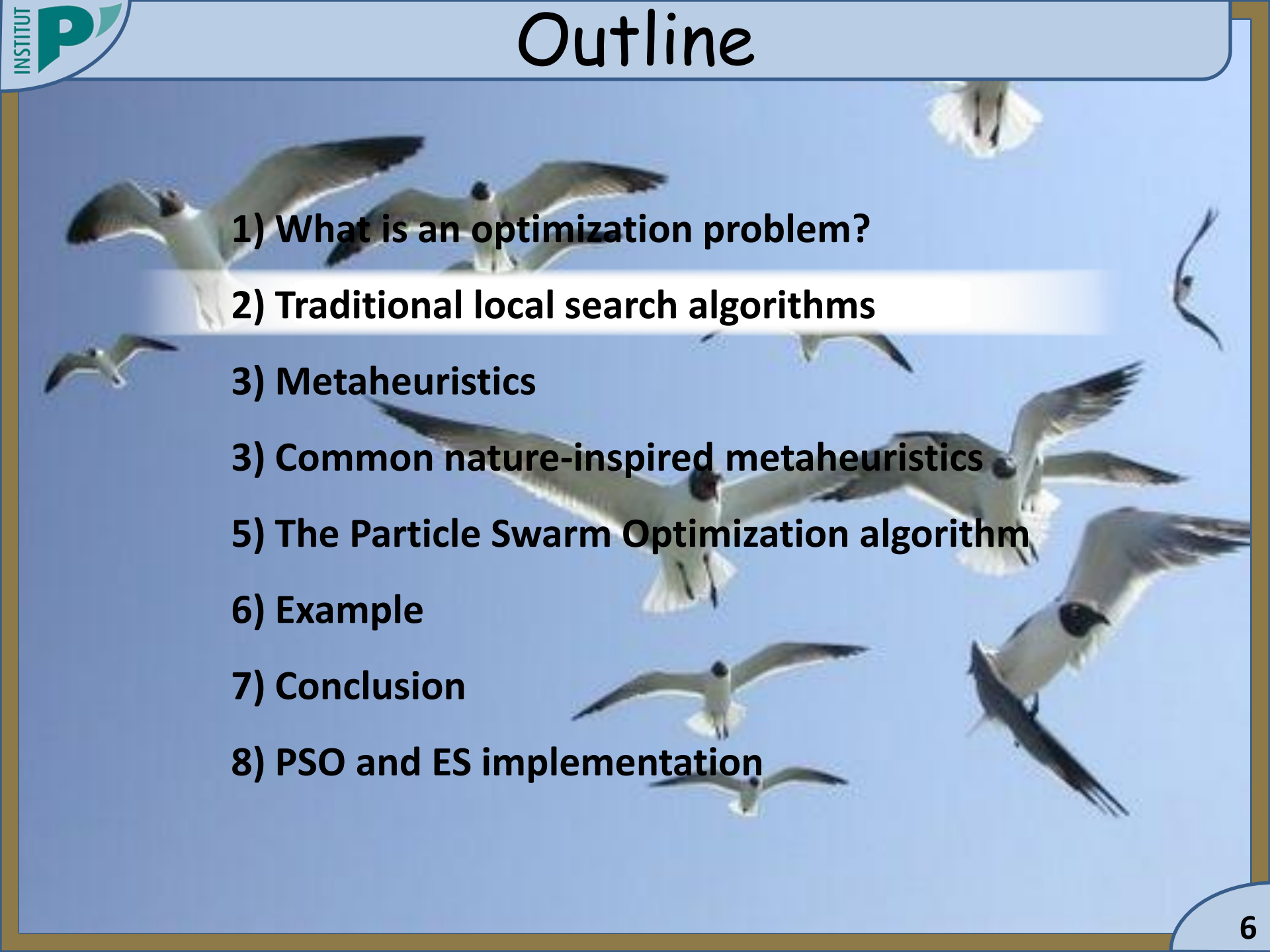


## A little bit more challenging?

Ex: Least median of squares



# Outline

- 
- 1) What is an optimization problem?
  - 2) Traditional local search algorithms
  - 3) Metaheuristics
  - 3) Common nature-inspired metaheuristics
  - 5) The Particle Swarm Optimization algorithm
  - 6) Example
  - 7) Conclusion
  - 8) PSO and ES implementation

# Gradient-based methods (1/2)

Background of gradient-based methods:  $J(\beta_k - \alpha_k \cdot D_k) < J(\beta_k)$

Objective function

Step size

Descent direction

Unknown parameters

## In which direction “ $D_k$ ” ?

The name of a gradient-based algorithm depends on the way  $D_k$  is computed.

### 1) First-order method:

- Gradient :  $D_k = \nabla J(\beta_k)$

### 2) Second-order method:

- Newton:  $D_k = [H(\beta_k)]^{-1} \cdot \nabla J(\beta_k)$  and  $\alpha_k \approx 1$

Hessian Matrix

### 3) Pseudo second-order method:

- Quasi-Newton method:  $H(\beta_k)$  is approximated (DFP, SR1, BFGS,...)

- Levenberg-Marquardt (least-squares):  $J(\beta) = \sum_i (y^{\text{exp}} - y(\beta))^2$   
 $H(\beta_k)$  is approximated using the sensitivity matrix  $X(\beta_k) = \frac{\partial y}{\partial \beta}(\beta_k)$

# Gradient-based methods (2/2)

Background of gradient-based methods:  $J(\beta_k - \alpha_k \cdot D_k) < J(\beta_k)$

Objective function

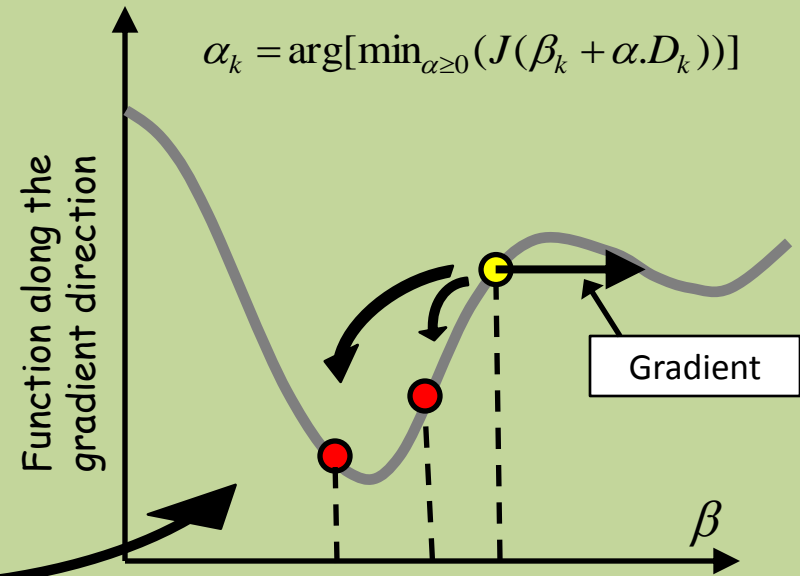
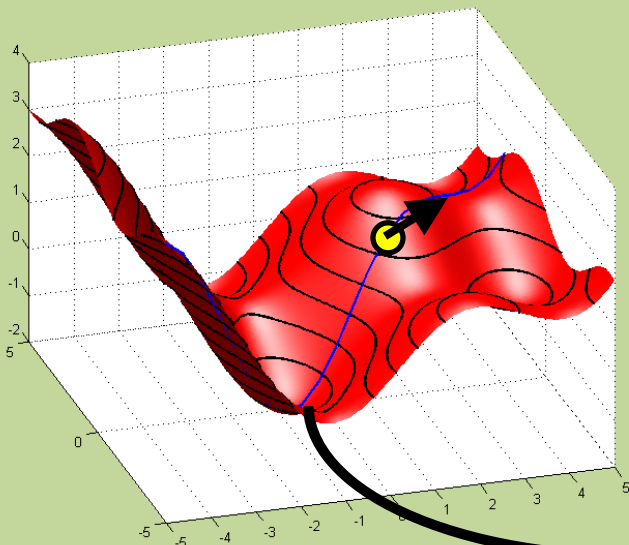
Step size

Descent direction

Unknown parameters

## How far in that direction ? What is the step size " $\alpha_k$ " ?

Different ways to compute the step size  $\alpha_k$  : (1) **Armijo**, (2) **Goldstein**, (3) **Wolfe** method or (4) the traditional **linear search**:

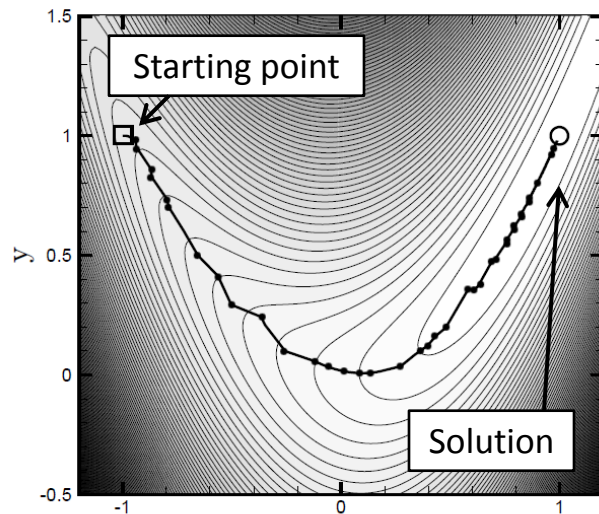
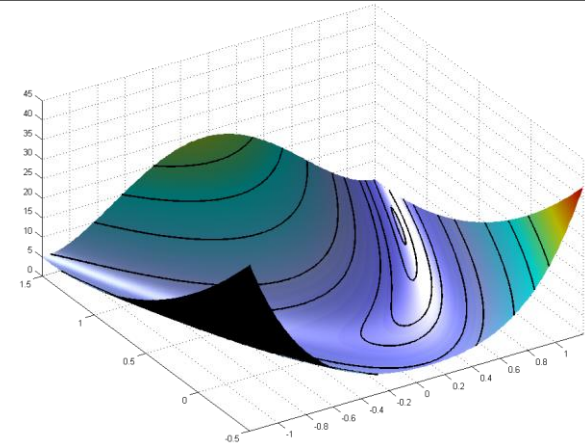


# Order and efficiency

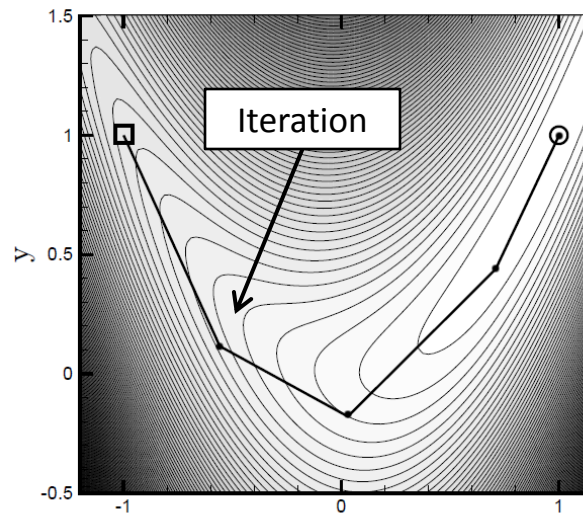
Iterative procedure:  $\beta_{k+1} = \beta_k - \alpha_k \cdot D_k$  until a convergence criterion is satisfied

The Rosenbrock function

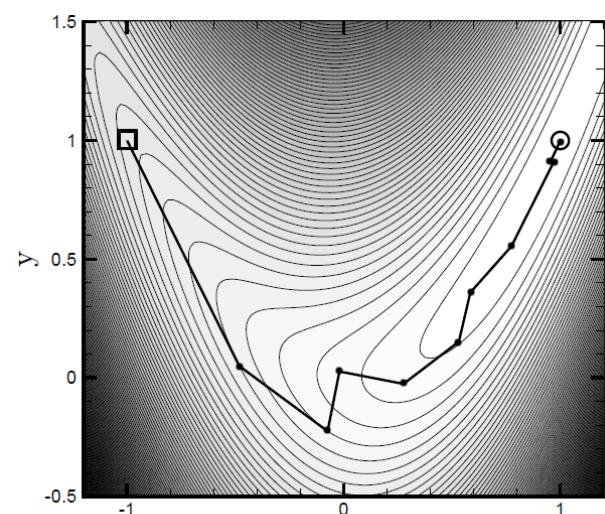
$$J(\beta) = \sum_{i=1}^{N_{\beta}-1} \left[ (1 - \beta_i)^2 + 10 \cdot (\beta_{i+1} - \beta_i^2)^2 \right]$$



**First order**  
Gradient method



**Second order**  
Newton's method



**Pseudo second order**  
Quasi-Newton's method

# Simplex methods (1/2)

**Definition :** a simplex is a generalization of a triangle to arbitrary dimension. If  $k$  is the number of dimensions, then a simplex is defined by  $k+1$  points in this space.

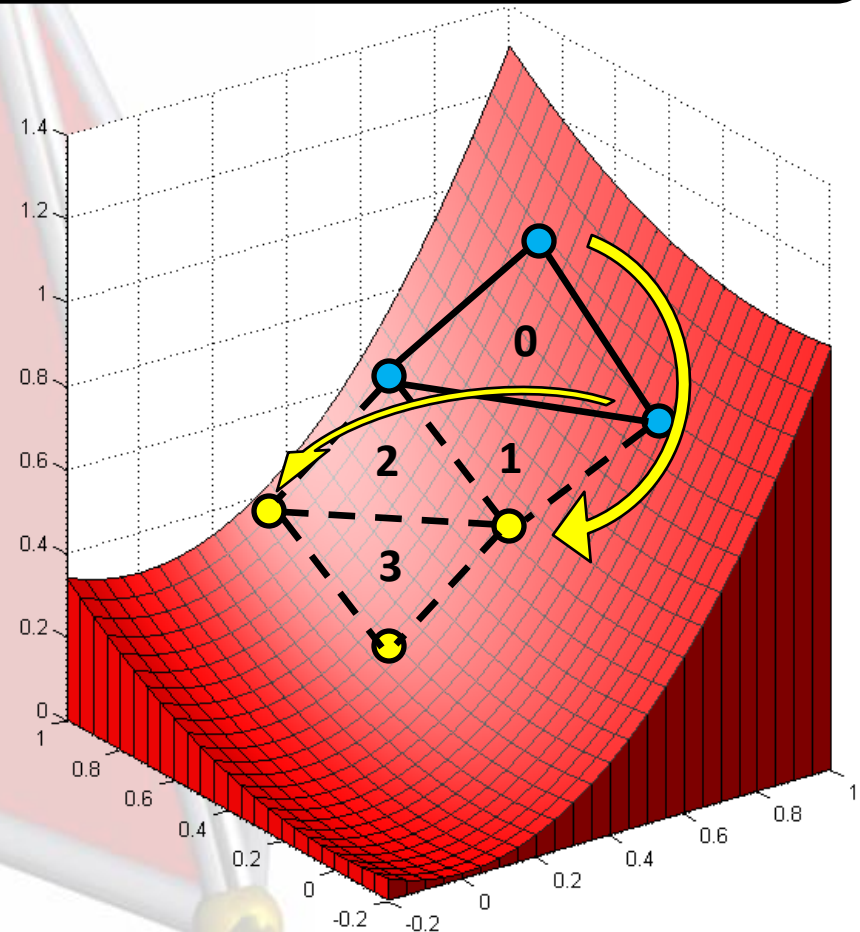
**The algorithm :** (George Dantzig, 1947)

- Deterministic
- Derivative-free (zero-order) 🧐
- Local search

## Simple but efficient method

The worst is iteratively replaced by its symmetric with respect to the center of gravity of the other points of the simplex

-> Few rules to handle tricky situations

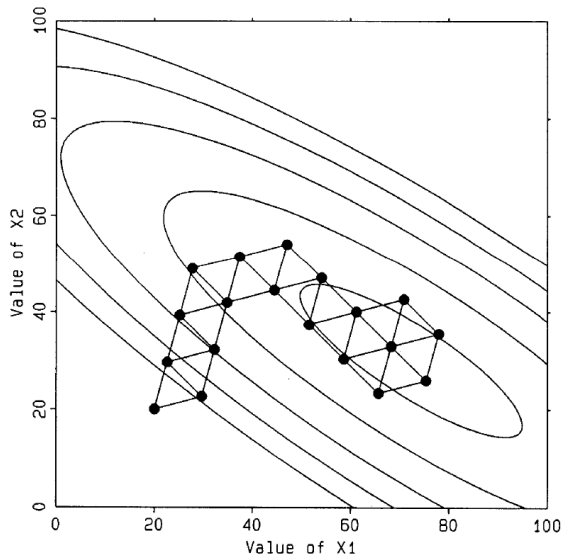


A 2D-case  $J(x, y)$

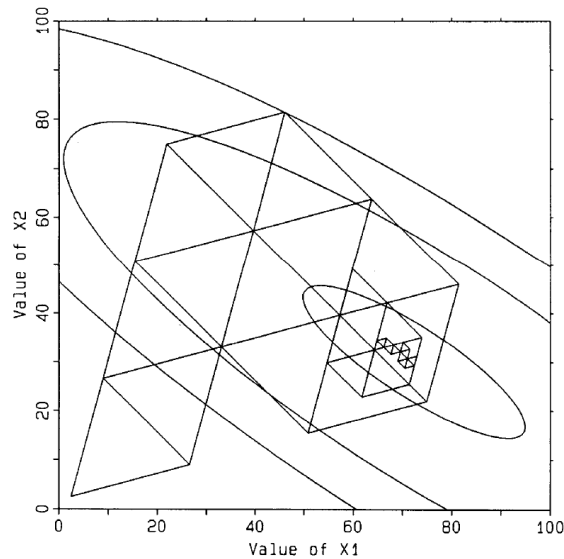
# Simplex methods (2/2)

Many variants of the algorithm !

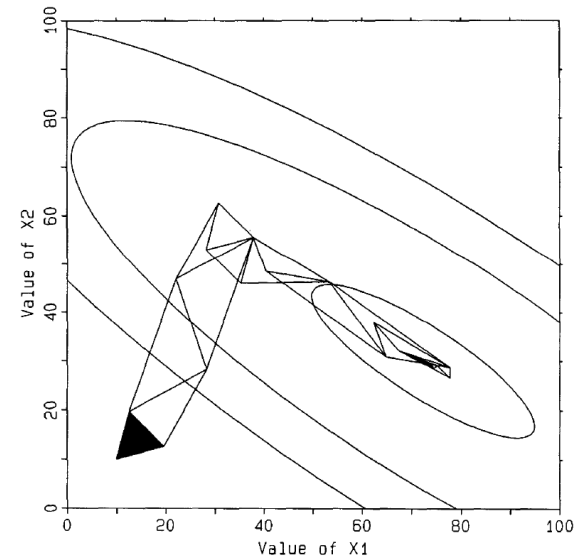
**The simplest one**  
Fixed size



**Simple but efficient**  
Sequential fixed size



**Trickier**  
Variable size

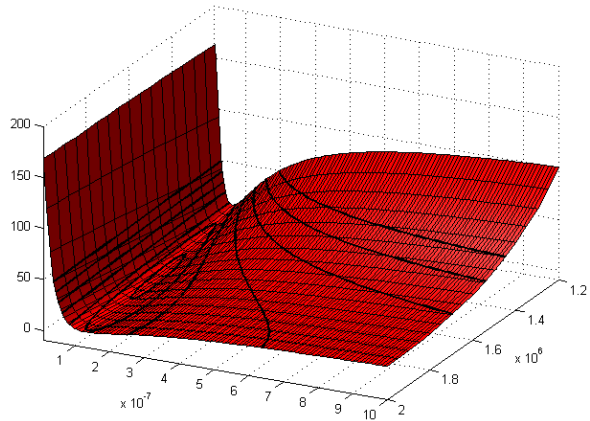


# What about this one ?

The simple case

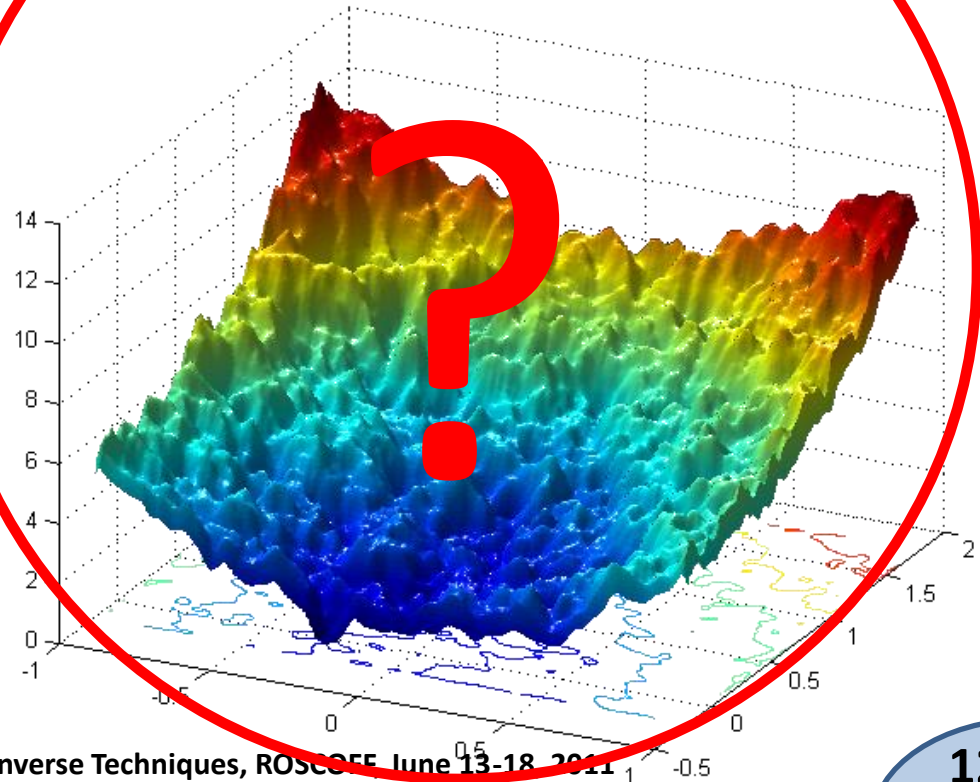
Ex: Least squares

(Flash method for parameter estimation)

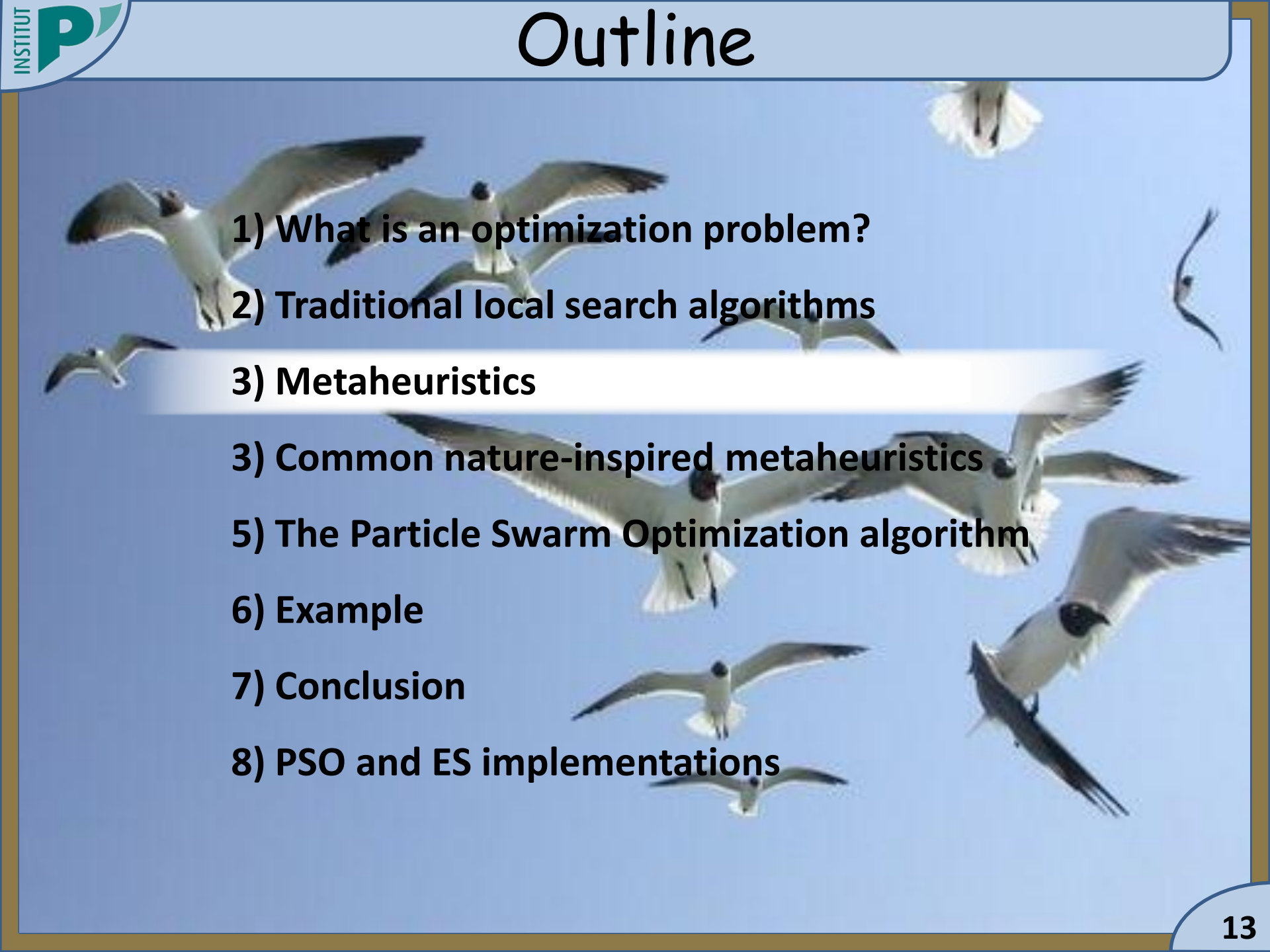


A little bit more challenging?

Ex: Least median of squares



# Outline

- 
- 1) What is an optimization problem?
  - 2) Traditional local search algorithms
  - 3) Metaheuristics
  - 3) Common nature-inspired metaheuristics
  - 5) The Particle Swarm Optimization algorithm
  - 6) Example
  - 7) Conclusion
  - 8) PSO and ES implementations

# What is a difficult problem ?

*"A metaheuristic is formally defined as an iterative generation process which guides a subordinate heuristic by*

## A hard problem:

The time required using any currently known algorithm **increases exponentially** with the size of the problem.

**Here we are...**

*moves or generating new starting solutions for the local search in a more intelligent way than just providing*

## Then, what can we do ? three possibilities:

1. Use algorithms that compute **all feasible solutions** and keep best one, but it may take **exponential time**.
2. Use "approximation algorithms" that always run in **polynomial time** but they may not **always produce the optimal solution**
3. Find a new job

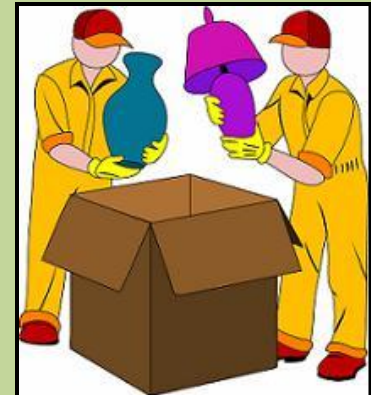
# A metaheuristic ? (1/3)

**Heuristic** (from the Greek heuriskein: "**find**" or "**discover**"):

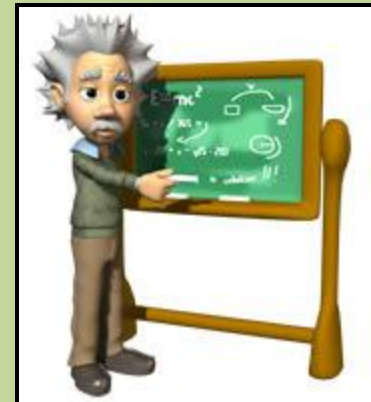
A commonsense rule (or set of rules) based on experience that aids problem solving

## Examples:

- If you are **packing odd-shaped items** into a box: start with the largest remaining items, fit the smaller items into the spaces left



- If you are having difficulty **understanding** a problem, try drawing a **picture**.



# A metaheuristic ? (2/3)

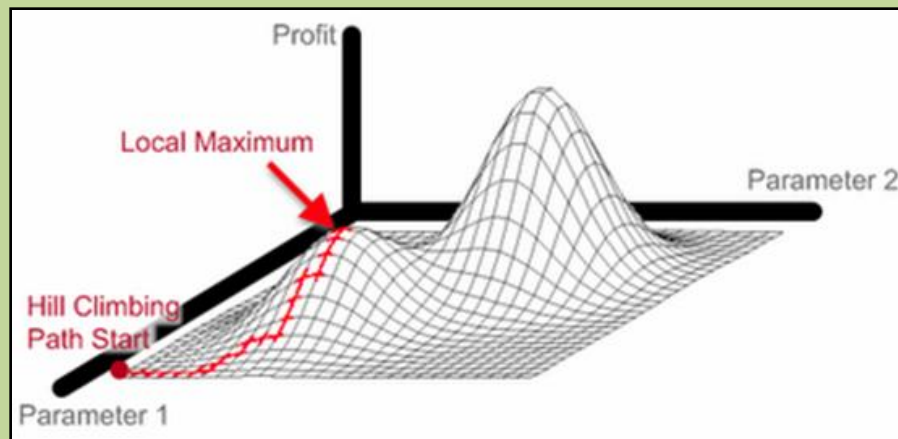
## Heuristic in computer science:

An algorithm that generally produces **acceptable solutions** with a **reasonable amount of time** in many situations, but:

- the solution can be bad
- the algorithm can be horribly slow

## Example:

**Greedy heuristic:** to maximize profit, choose the solution of the neighborhood that maximizes local profit (Hill climbing)



# A metaheuristic ? (3/3)

*"A metaheuristic is formally defined as an iterative generation process which guides a subordinate heuristic by combining intelligently different concepts for exploring*

**Meta** (from Greek for “**beyond**” or “**higher**”):

a prefix used to indicate a concept which is an abstraction from another concept.

*increase their performance. The main goal is to avoid the disadvantages of iterative improvement and, in particular, multiple descent by allowing the local search to escape*

## **Meta-heuristic: No commonly agreed definition !**

- Metaheuristics are strategies that guide the **search process**.
- Algorithms that find **acceptable solutions** of problems by using **several heuristics**.
- A **metaheuristic** is a heuristic method for solving a very general class of computational problems by combining heuristics in the hope of obtaining a more efficient or more robust procedure.

# Fundamental properties of metaheuristics

- Efficiently **explore** the search space.
- May be **approximation** and **stochastic** algorithms.
- Mechanisms to avoid getting trapped **local optima**.
- Use **search experience** to be more efficient than random search.

## Three principles

### Exploration

(diversification, global search)

Finding areas of the search space  
that are still not investigated

### Exploitation

(Intensification, local search)

Improving the current solutions found  
by performing generally small changes

### Memory

(Search experience)  
Store information



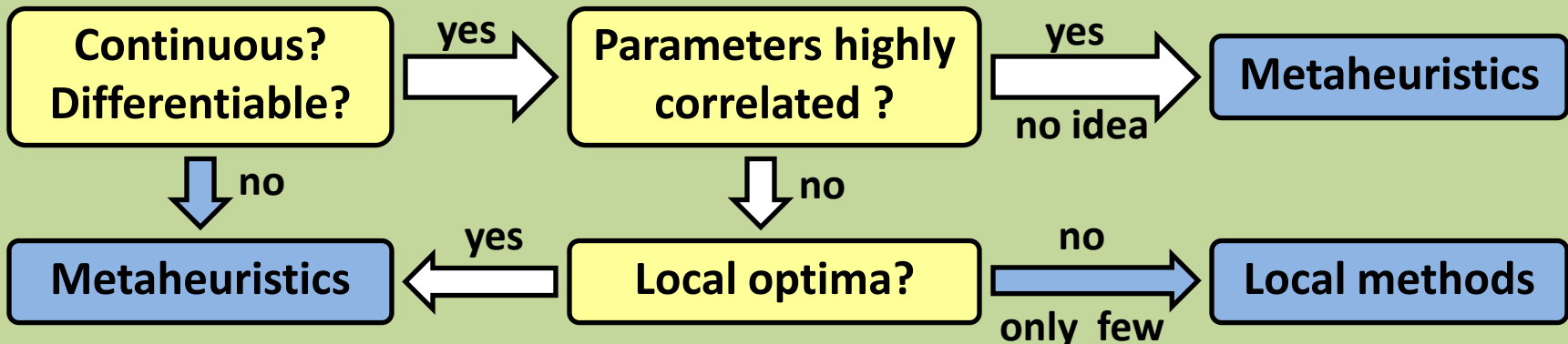
# And what...?

**Ok, metaheuristics are designed to find acceptable solutions of NP-hard problems**

But for us, humble researchers and engineers in thermal sciences who don't understand anything to complexity computation theory, what does really matter in practice ?

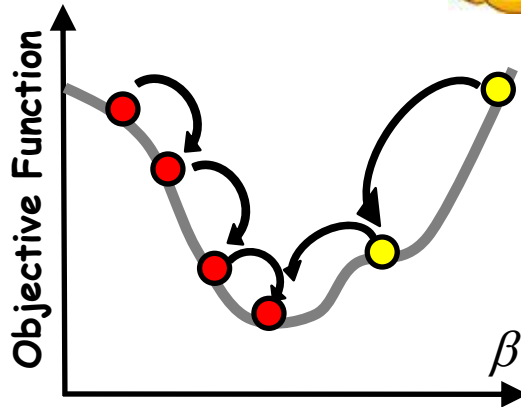
- 1) Which optimization algorithm can I use ?**
- 2) For my problem, is this algorithm suitable ?**

**It depends ... on your objective function:**

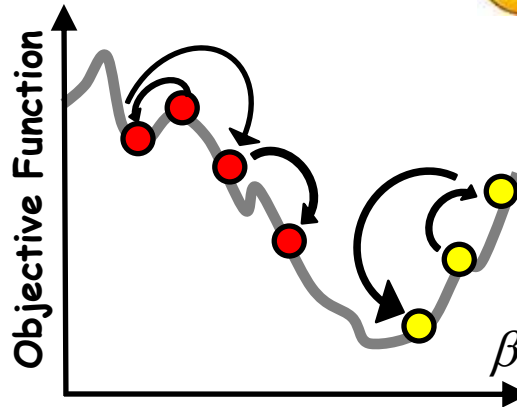


# Is it a hard objective function?

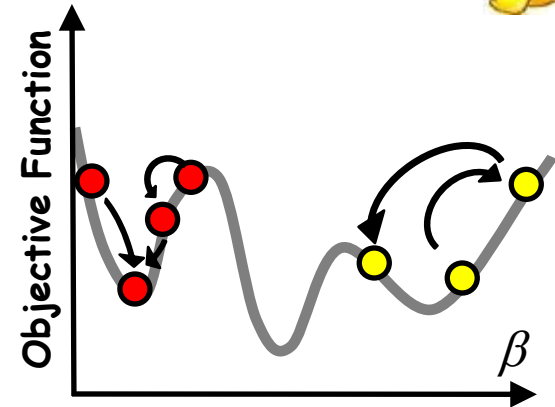
Best Case



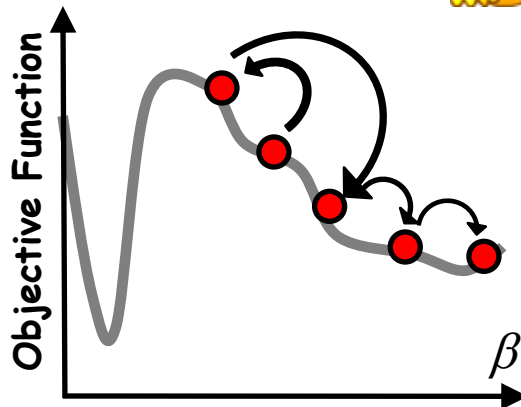
Low total variation



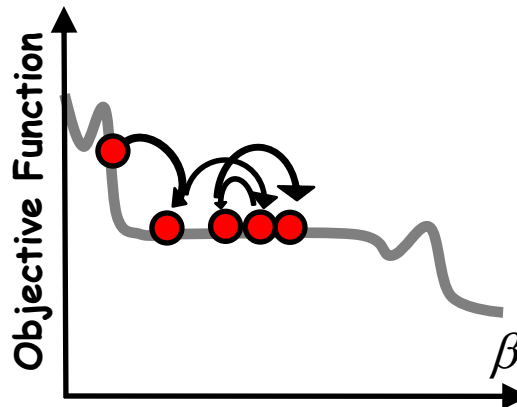
Local optima



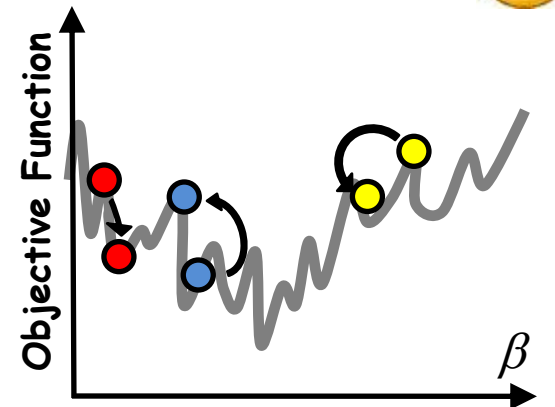
Misleading



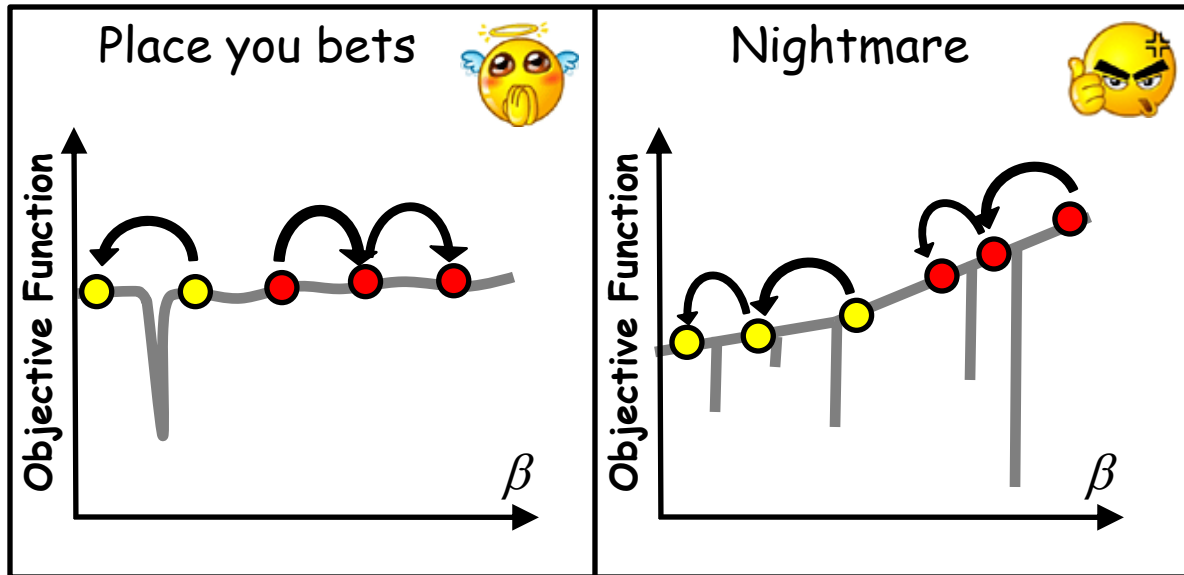
Neutral



Rugged



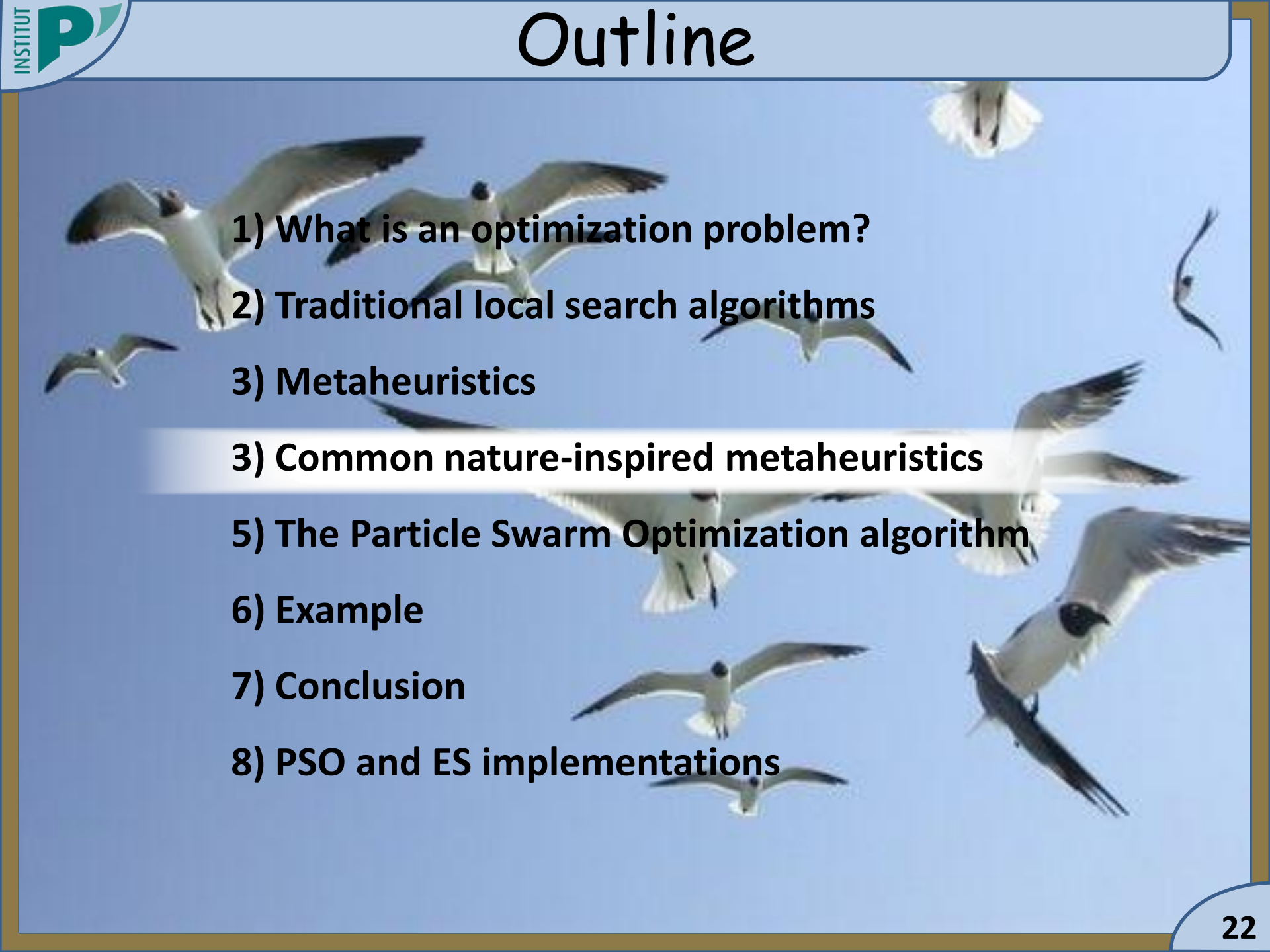
# Probably not for us



- Robust optimum are often preferred
- Kind of situation hardly possible in thermal modelling



# Outline

- 
- 1) What is an optimization problem?
  - 2) Traditional local search algorithms
  - 3) Metaheuristics
  - 3) Common nature-inspired metaheuristics
  - 5) The Particle Swarm Optimization algorithm
  - 6) Example
  - 7) Conclusion
  - 8) PSO and ES implementations

## Evolutionary computation

*“Evolution is a process that results in heritable changes in a population spread over many generations”\**

### Evolutionary algorithm

#### Evolutionary strategies

Differential  
evolution

Genetic  
Algorithms

### Swarm Intelligence

Particle Swarm

Ant Colony

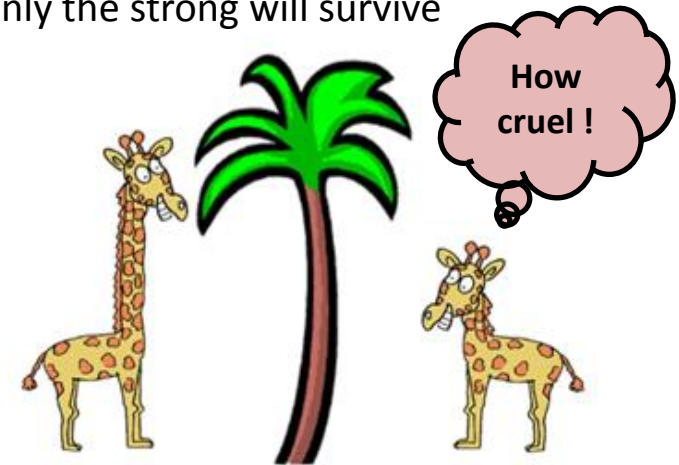
\* <http://www.talkorigins.org/faqs/evolution-definition.html>, 03/11/2009

# Evolutionary algorithms (1/3)

## An evolutionary algorithm (EA):

- generic population-based algorithm
- stochastic
- inspired by the **theory of natural selection**

Only the strong will survive



## Technical terms:

Objective  
function

Fitness  
(Rescaled function)

~~Solutions,  
Candidates~~

Individuals  
(chromosome)

~~Parameters~~

Genes

~~Set of  
solutions~~

Population

## Main principle:

An EA manipulates a **population of individuals** that **compete** with each other through **biological inspired mechanisms** called **operators**. The fittest individuals produce more **offspring**. Generations after generations, the **population improves** and converges to a solution.

## Operator:

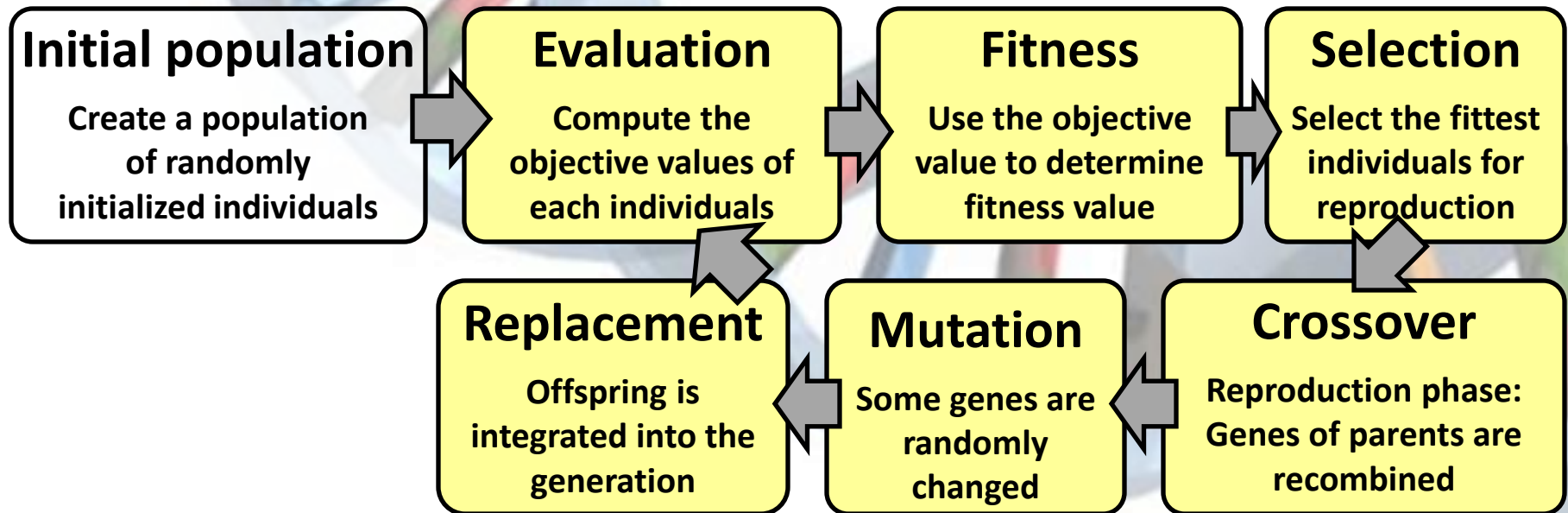
In mathematics, any symbol that indicates an operation to be performed. An operator may be regarded as a **function**, **transformation**, or map. (*Encyclopedia Britannica, 2008*).

# Evolutionary algorithms (3/3)

An algorithm is an “**evolutionary algorithm**” if it uses particular **operators** inspired by biological evolution:

**Selection, crossover, mutation and replacement**

## Generic structure



Inspired from Thomas Weise, *Global Optimization Algorithms – Theory and Application* –, e-book, 2008-05-20

## Evolutionary computation

*“Evolution is a process that results in heritable changes in a population spread over many generations”\**

### Evolutionary algorithm

#### Evolutionary strategies

Differential  
evolution

#### Genetic Algorithms

### Swarm Intelligence

Particle Swarm

Ant Colony

\* <http://www.talkorigins.org/faqs/evolution-definition.html>, 03/11/2009

# GA and ES (1/3)

Example:  $f(x, y, z) = x^2 + y^2 + z^2$

## Evolution strategies (ES) : “(4/2 + 4)-ES”

2 random parents  
for each children

Crossover

Mutations  
create children

(2)	(4)
1	0
2	6
-1	3
$\sigma=1$	$\sigma=2$



M
0
2
3
$\sigma=2$

$+ N(0, \sigma) =$   
 $+ N(0, \sigma) =$   
 $+ N(0, \sigma) =$   
 $X \begin{cases} \alpha & 50\% \\ 1/\alpha & 50\% \end{cases}$

C1
2
0
5
1.5

F=29



New parameter

New population:

The 4 best individuals are kept

(2)	(3)	C2	C1	<del>(1)</del>	C4	(4)	<del>C3</del>
6	13	15	29	38	50	53	70
1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>th</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>	7 <sup>th</sup>	8 <sup>th</sup>

In “(4/2, 4)-ES” : parents are replaced by the 4 best children

× 4

F=15



F=70



F=50



(1) F 38



x	2
y	-3
z	5

(2) F 6



x	1
y	2
z	-1

(3) F 13



x	3
y	2
z	2

(4) F 53




x	0
y	6
z	3


# GA and ES (2/3)

Example:  $f(x, y, z) = x^2 + y^2 + z^2$


## Genetic algorithm (GA)

(1) 


F	38
x	2
y	-3
z	5

(2) 

F	6
x	1
y	2
z	-1

(3) 

F	13
x	3
y	2
z	2

(4) 

F	53
x	0
y	6
z	3

Individual  
are ranked

1 <sup>st</sup>	(2)	6
2 <sup>nd</sup>	(3)	13
3 <sup>th</sup>	(1)	38
4 <sup>th</sup>	(4)	53

Selection  
probability

38%
29%
21%
12%

Individuals  
selected

(2)
(2)
(3)
(1)

4 couples are  
created

(2)+(3)
(2)+(2)
(2)+(1)
(1)+(3)

Replacement

(1)
(2)
(3)
(4)

Mutations  
(very few)

C1
C2'
C3
C4

Crossover

C1
C2
C3
C4

(2)	1	2	-1
(3)	3	2	2

=

C1	1	2	2
----	---	---	---

Crossover point

# GA and ES (3/3)

(1)



F	38
x	2
y	-3
z	5

(2)



F	6
x	1
y	2
z	-1

(3)



F	13
x	3
y	2
z	2

(4)



F	53
x	0
y	6
z	3

## Evolution strategies (ES) : “(4/2 + 4)-ES”

2 random parents for each children

(2)	(4)
1	0
2	6
-1	3
$\sigma=1$	$\sigma=2$

Crossover

M
1
2
3
$\sigma=2$

Mutations  
4 children

C1
2
0
5
1.5

Selection of the 4 best Individuals

(2)	6	(1)	38
(3)	13	C4	50
C2	15	(4)	53
C1	29	C3	70

## Genetic algorithm (GA)

Selection

1 <sup>st</sup>	(2)	6	→	38%
2 <sup>nd</sup>	(3)	13	→	29%
3 <sup>th</sup>	(1)	38	→	21%
4 <sup>th</sup>	(4)	53	→	13%

4 couples are created

(2)+(3)
(2)+(2)
(2)+(1)
(1)+(3)

Crossover

C1
C2
C3
C4

Mutations

C1
C2'
C3
C4

New population

(1)
(2)
(3)
(4)

## Many similarities between ES and GA:

- Same operators and similar structure
- Same concepts (non guided mutation, selection pressure...)

"It is selection, and only selection, that directs evolution in directions that are nonrandom with respect to advantage." (Dawkins, 1986, p.312).

## As all metaheuristics, many, many... and many variants:

Example: different operators for EA:

### Selection

Uniform  
Rank<sup>1</sup>  
Tournament  
Random<sup>2</sup>...

### Crossover

1-point<sup>1</sup>  
Multi-point  
Uniform<sup>2</sup>  
Arithmetic...

### Mutation

Permutation  
Random variable<sup>2</sup>  
Replacement<sup>1</sup>  
...

### Replacement

Total<sup>1</sup>  
Elitist<sup>2</sup>  
Steady state  
...

1 – variant used for GA in previous slides

2 – variant used for ES in previous slides

## Evolutionary computation

*“Evolution is a process that results in heritable changes in a population spread over many generations”\**

### Evolutionary algorithm

#### Evolutionary strategies

Differential  
evolution

Genetic  
Algorithms

### Swarm Intelligence

Particle Swarm

Ant Colony

\* <http://www.talkorigins.org/faqs/evolution-definition.html>, 03/11/2009

## A swarm:

A set of (mobile) agents which are liable to **communicate directly or indirectly** (by acting on their local environment) with each other, and which collectively carry out a distributed problem solving.

<http://www.molbio.ku.dk/MolBioPages/abk/PersonalPages/Jesper/Swarm.html>, 05/11/2009

## Swarm intelligence (SI)

Type of artificial intelligence based on the **collective** behavior of **decentralized, self-organized** systems. (wikipedia)

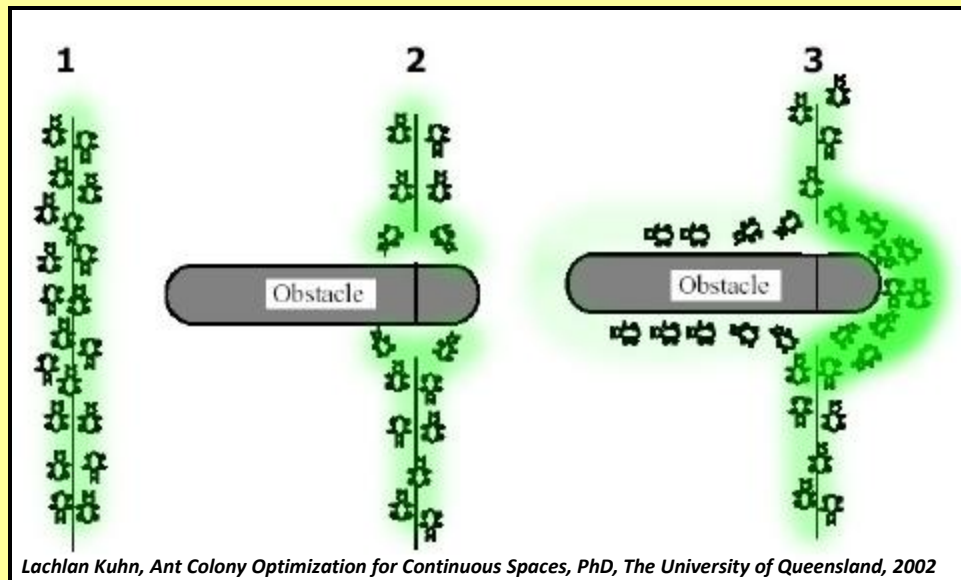
- Very simple rules
- Multiplicity, randomness, messiness
- No centralized control structure dictating how individual agents should behave
- Interactions lead to the emergence of "intelligent" global behavior

## Examples

Ant colonies, bird flocking, animal herding, bacterial growth, and fish schooling.

## Ant colony optimization (ACO) (Dorigo 1992):

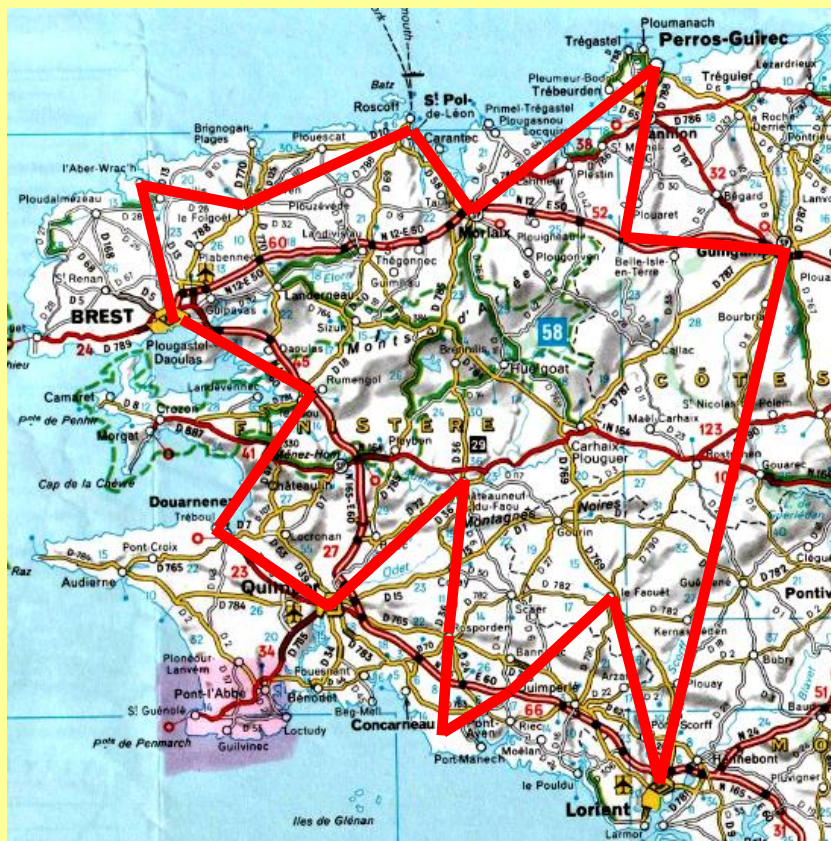
- inspired by the ants' foraging behavior (wandering in search of food)
- indirect communication using chemical pheromone trails (stigmergy)
- shortest path between food sources and their nest



First and mostly applied to discrete optimization problem

## Traveling salesman problem (TSP) (1930):

Given a list of cities and their pairwise distances, find a shortest possible tour that visits each city exactly once.



*Solution found by my swarm of neurons*

## Algorithm:

### For each iteration

**For each ant “k”**

**Choose a starting city** randomly

**For each** non visited city

## Add a random one

## Intensity Visibility

**End**

## Evaluate the tour

## Ants depose pheromone

# End

## Evaporation of pheromone

# End

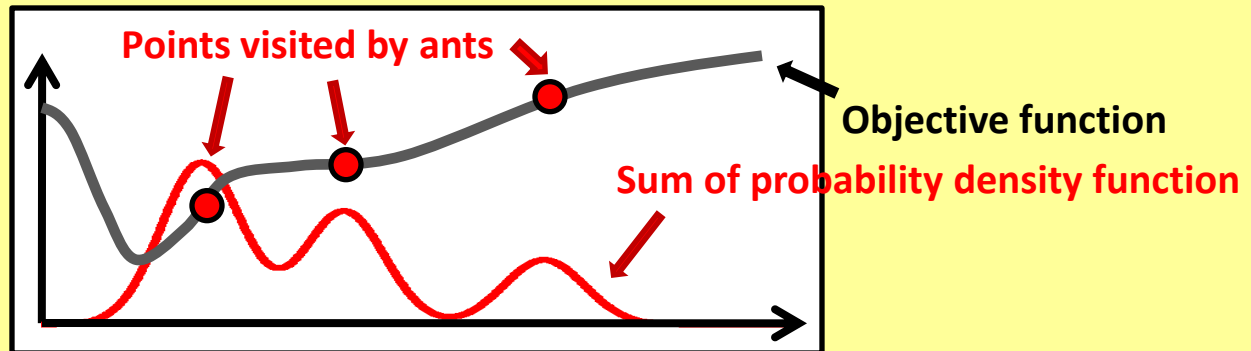
## Problem:

-infinite number of possible paths in the continuous domain

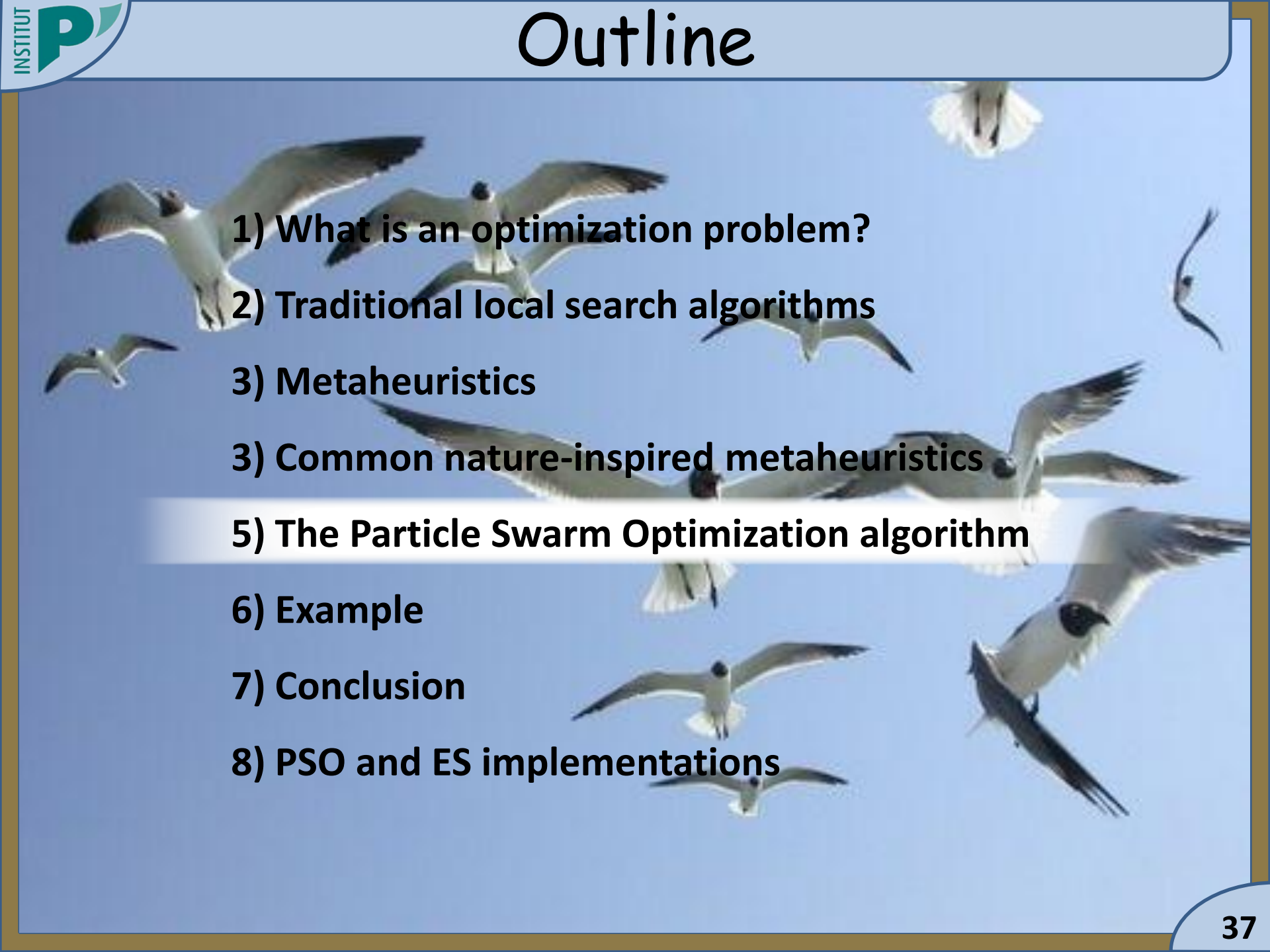
## A solution:

- pheromone trails are replaced by probability density functions
- evaporation is replaced by the forgetting the worst solutions
- ants randomly sent with respect to the sum of probability density function

## Example:

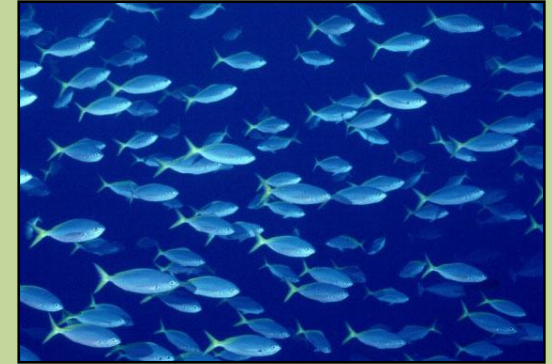


# Outline

- 
- 1) What is an optimization problem?
  - 2) Traditional local search algorithms
  - 3) Metaheuristics
  - 3) Common nature-inspired metaheuristics
  - 5) The Particle Swarm Optimization algorithm
  - 6) Example
  - 7) Conclusion
  - 8) PSO and ES implementations

## Particle Swarm Optimization (PSO) (Eberhart, Kennedy 1995):

- inspired from the flocking of birds and fish
- how individuals interact
- how individuals learn from experience



## Properties:

- stochastic
- derivative-free
- global



- population-based
- nature-inspired
- adapted to continuous problems

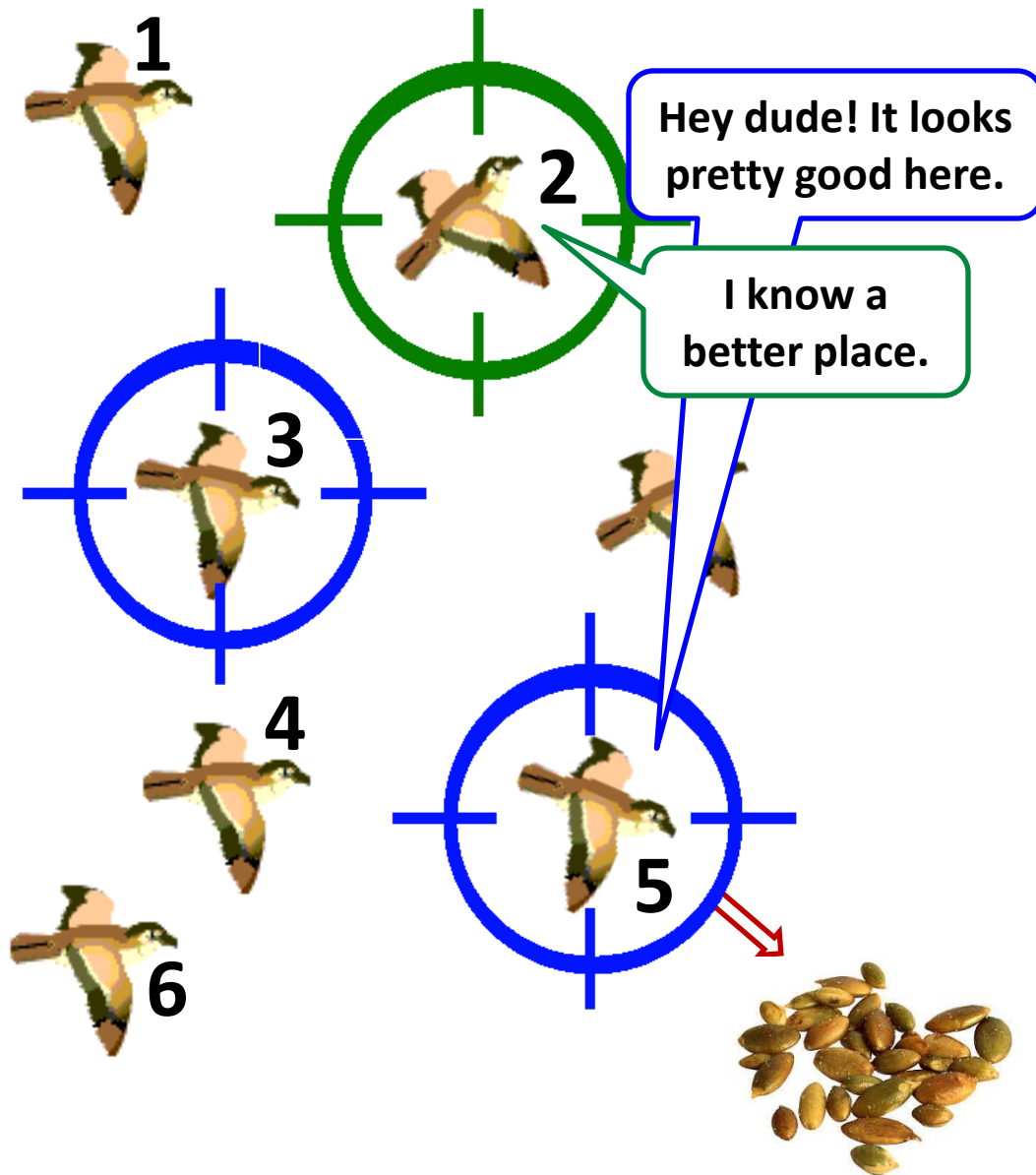
## Technical term:

~~Solutions,  
Individuals~~

Particle



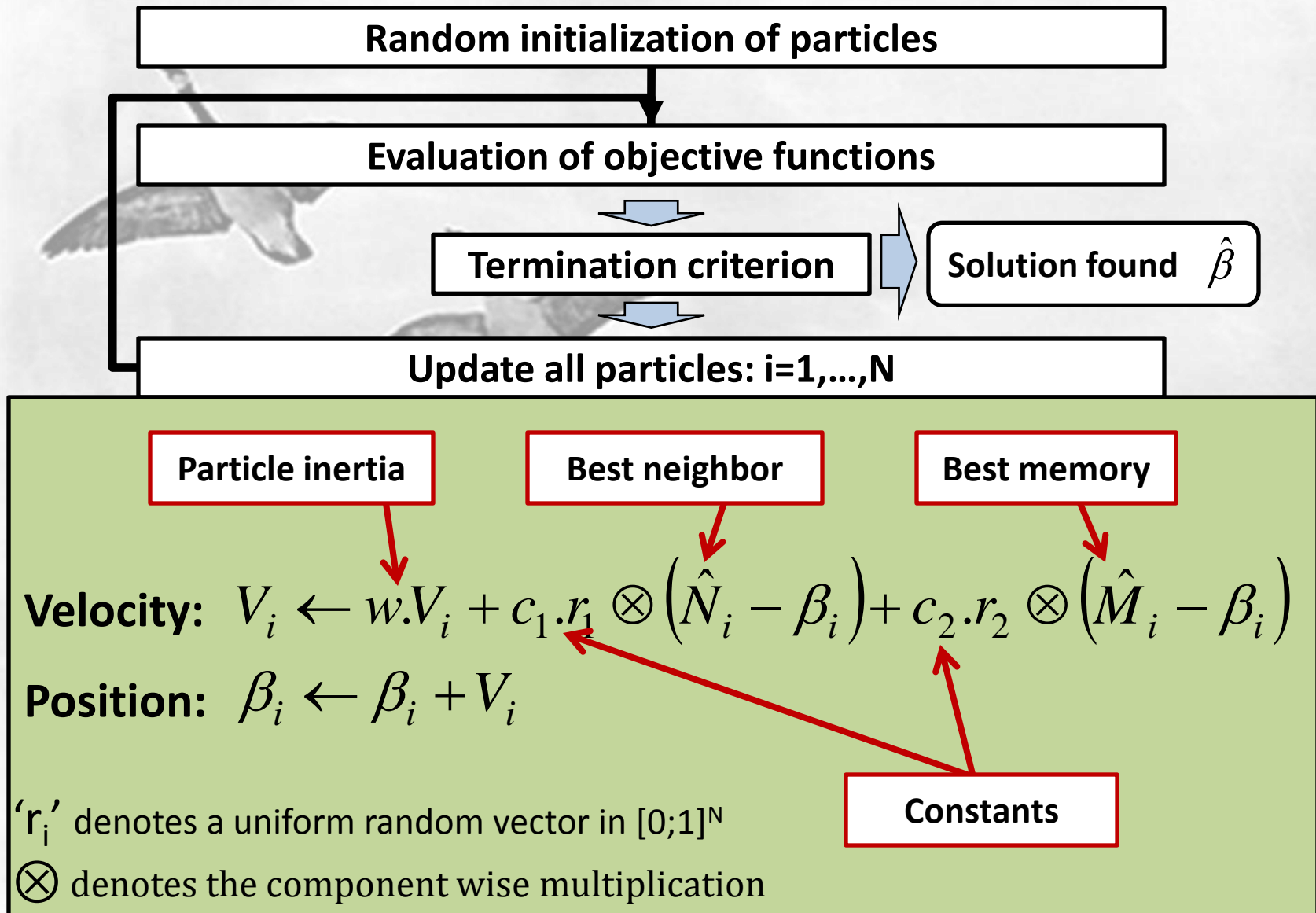
# Swarm of particles



## Particle 2

- Position  $\beta \in S \subset R^{N_\beta}$
- Velocity  $V(=\Delta\beta)$
- Performance  $J(\beta) \in R$
- Best position  $\hat{M}$ 
  - = Best experience
  - = Best memory
- Best performance  $J(\hat{M})$
- Informants **3** and **5**
  - = Neighbors

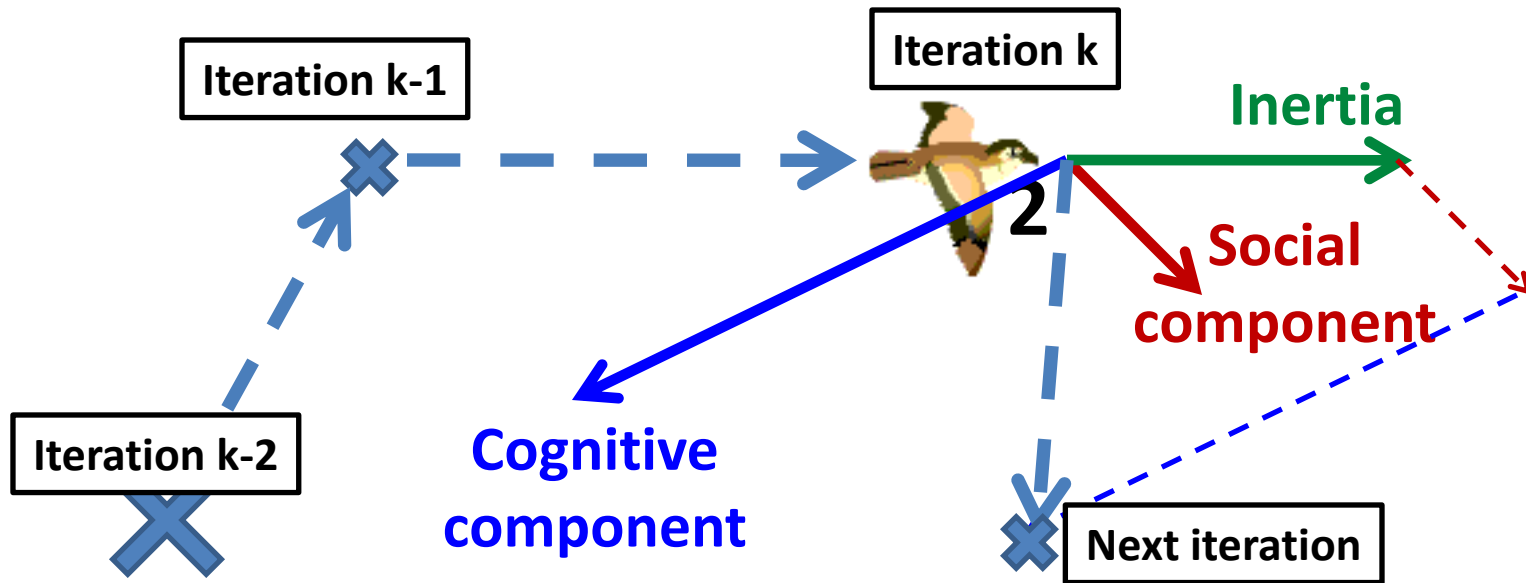
# The algorithm



# Hard to understand ? Draw a picture.

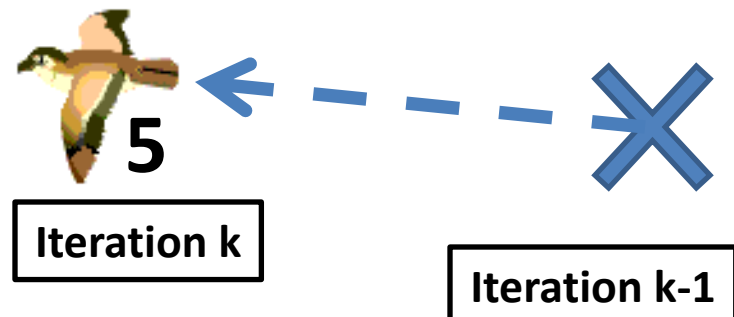
$$V_i \leftarrow w.V_i - c_1.r_1 \otimes (\hat{N}_i - \beta_i) + c_2.r_2 \otimes (\hat{M}_i - \beta_i)$$

$$\beta_i \leftarrow \beta_i + V_i$$



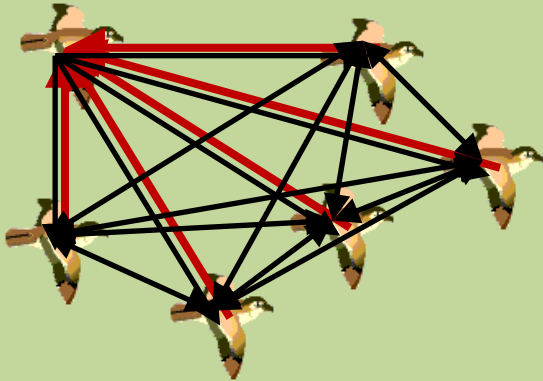
## Note: Effect of component wise multiplication

In fact, cognitive and social components are not necessarily oriented toward the best memory and the best neighbor respectively.

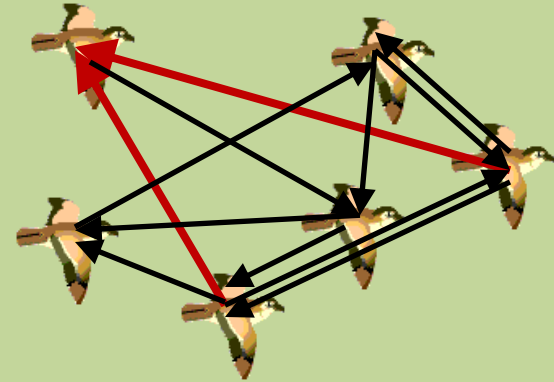


# The neighborhood

All particles are neighbors

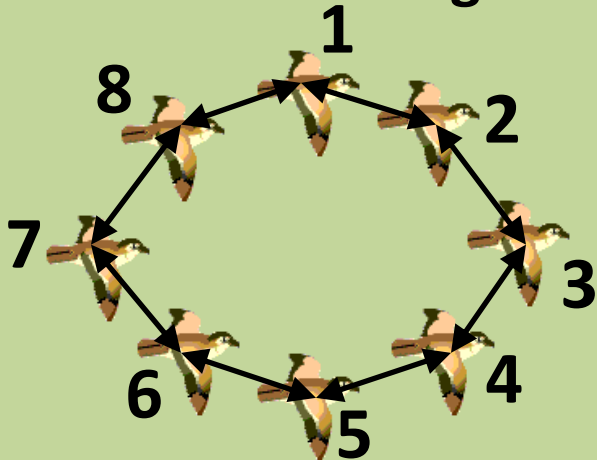


Random neighbors

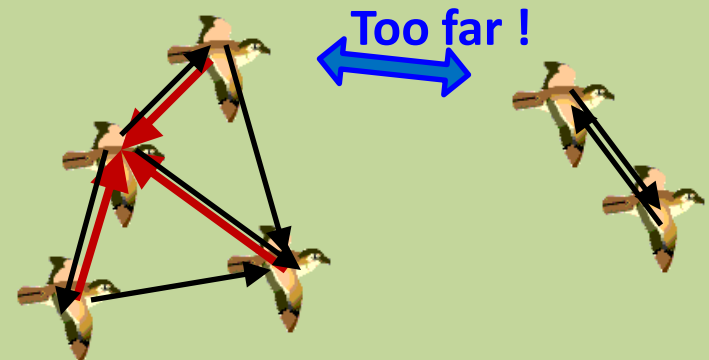


Update each iteration !

Index-based neighbors



Geographical neighborhood



Change each iteration !

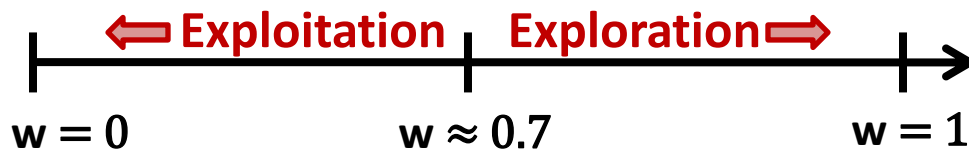
# Algorithm adjustments

Population size = 20 particles

Neighborhood type = random or ring (index-based)

Neighborhood size = 3 or 4 particles

**Particle inertia ( $\approx$ friction)**



**Velocity:**  $V_i \leftarrow w.V_i + c_1.r_1 \otimes (\hat{N}_i - \beta_i) + c_2.r_2 \otimes (\hat{M}_i - \beta_i)$

**Position:**  $\beta_i \leftarrow \beta_i + V_i$

**Velocity**

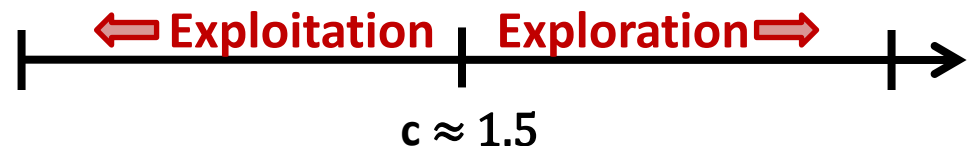
Often clamped to avoid explosion:

$$-V_{\max} \leq V_i \leq V_{\max}$$

**Acceleration coefficients ( $c_1 = c_2$ )**

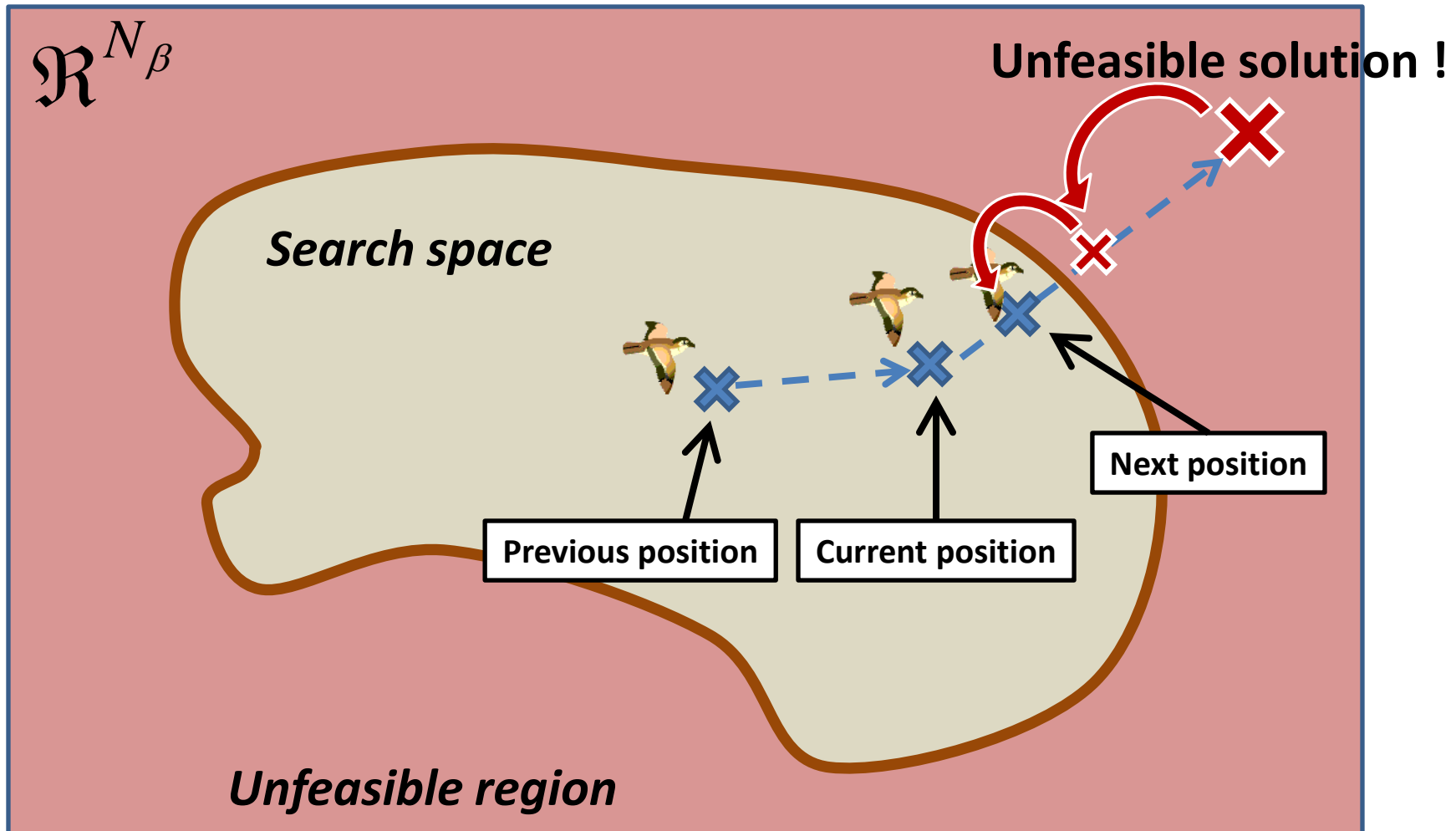
Not independent of inertia

Same compromise exploitation/exploration



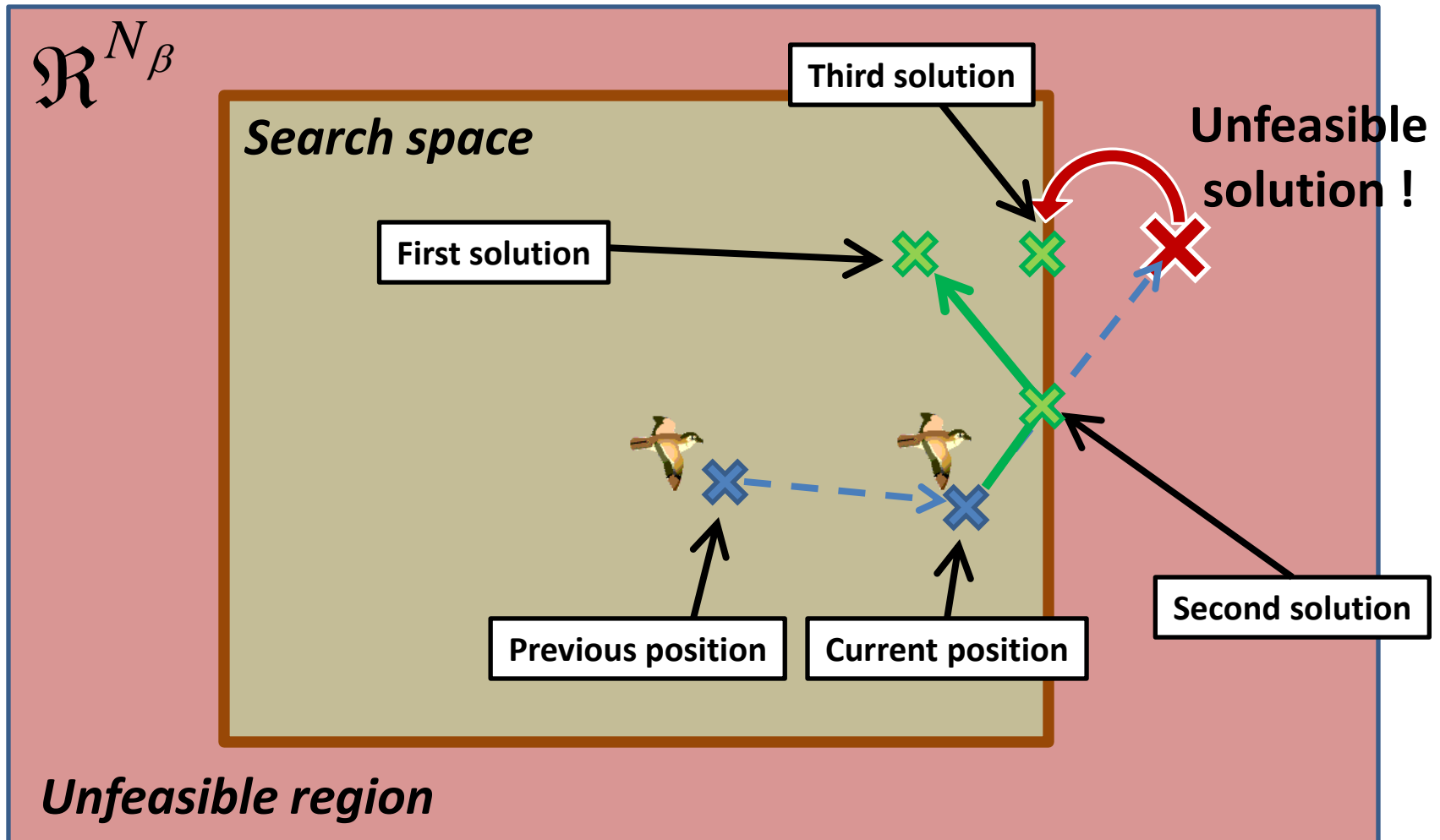
# Handling hard constraints (1/2)

## A method for the general case



# Handling hard constraints (1/2)

In our case, simple boundary constraints



## Constriction coefficient (Clerc, Eberhart, Kennedy):

$$\left. \begin{aligned} V_i &\leftarrow \chi \cdot \left[ V_i + \frac{\varphi}{2} \cdot r_1 \otimes (\hat{N}_i - \beta_i) + \frac{\varphi}{2} \cdot r_2 \otimes (\hat{M}_i - \beta_i) \right] \\ \beta_i &\leftarrow \beta_i + V_i \end{aligned} \right\} \text{ with } \chi = \frac{2}{\varphi - 2 + \sqrt{\varphi^2 - 4\varphi}}$$

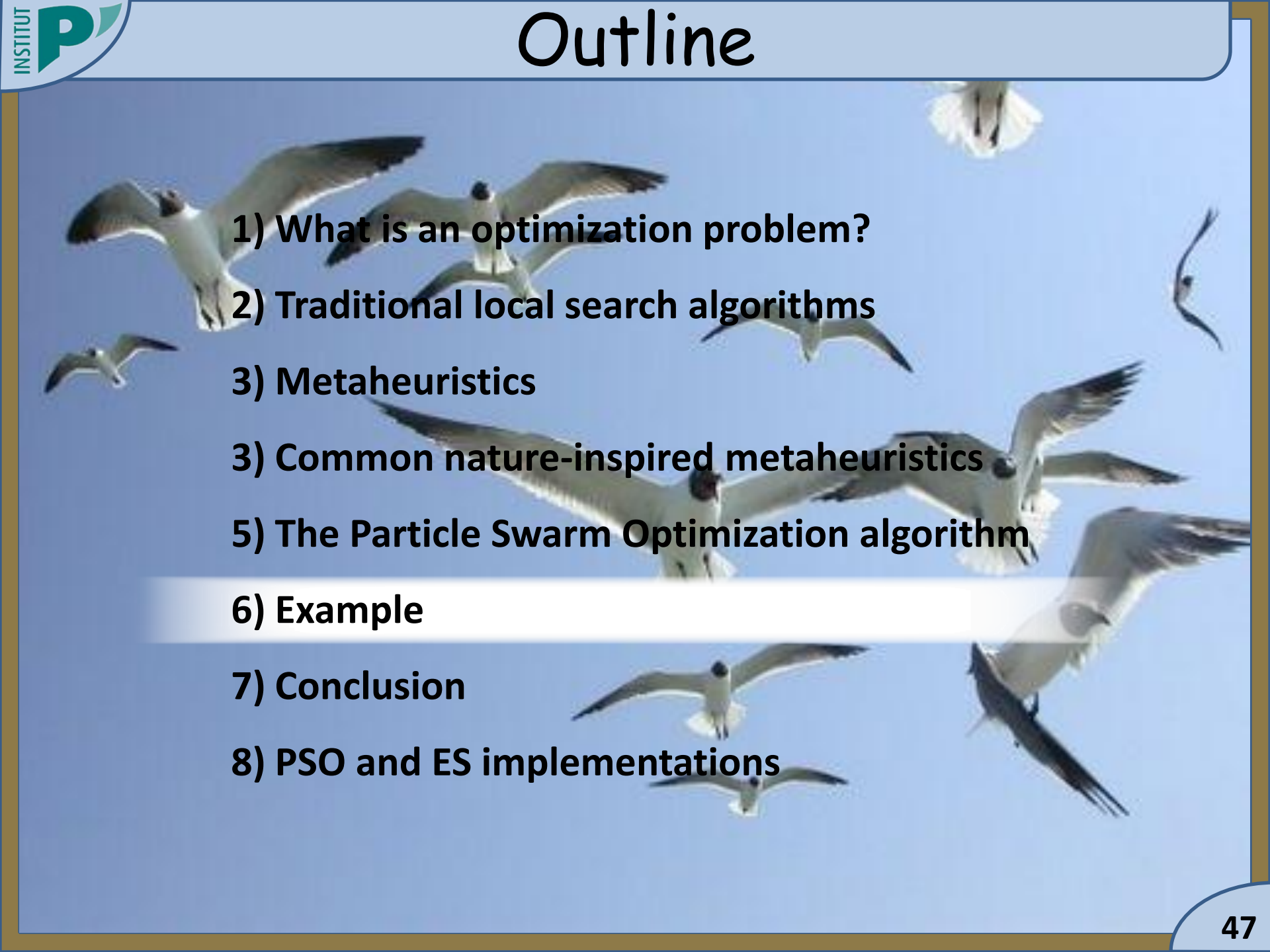
## Time dependant:

$$\begin{aligned} V_i &\leftarrow w(t) \cdot V_i + c_1 \cdot r_1 \otimes (\hat{N}_i - \beta_i) + c_2 \cdot r_2 \otimes (\hat{M}_i - \beta_i) \\ \beta_i &\leftarrow \beta_i + V_i \end{aligned}$$

## Fully informed (Mendes, Kennedy, Neves) :

$$\begin{aligned} V_i &\leftarrow \chi \cdot \left[ V_i + \frac{\varphi}{K_i} \sum_k^{K_i} r \otimes (N_k - \beta_i) \right] \\ \beta_i &\leftarrow \beta_i + V_i \end{aligned} \quad \text{With } K_i \text{ the neighborhood size}$$

# Outline

- 
- 1) What is an optimization problem?
  - 2) Traditional local search algorithms
  - 3) Metaheuristics
  - 3) Common nature-inspired metaheuristics
  - 5) The Particle Swarm Optimization algorithm
  - 6) Example
  - 7) Conclusion
  - 8) PSO and ES implementations

## Context

- Anisotropic material
- Temperature dependant parameters

$$\begin{cases} \lambda_x(T) = \lambda_{x0} + \lambda_{x1}.T \\ \lambda_y(T) = \lambda_{y0} + \lambda_{y1}.T \\ C(T) = C_0 + C_1.T \end{cases} \Rightarrow$$

Unknown parameters

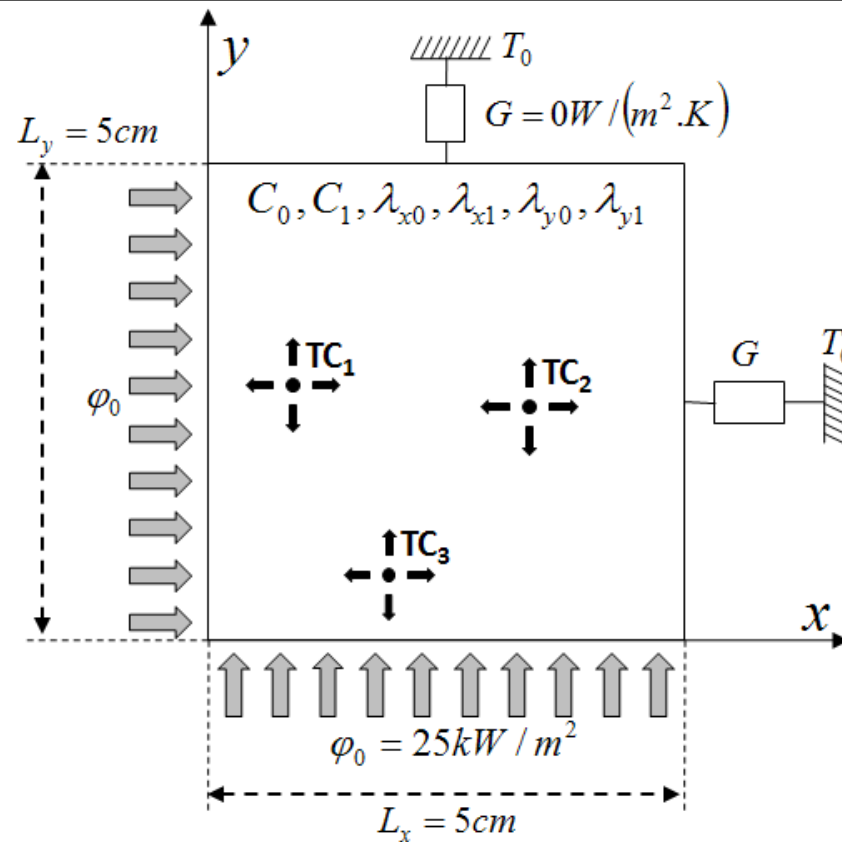
$$\beta = (\lambda_{x0}, \lambda_{x1}, \lambda_{y0}, \lambda_{y1}, C_0, C_1)$$

## Objective

Estimate  $\beta$  with the precision

# The numerical experiment

- 2D-sample of an orthotropic media (5x5cm)
- Insulated on the right and top boundaries
- Heated on the left and bottom with a constant heat flux



Jarny, An inverse analysis to estimate linearly temperature dependent, 1995

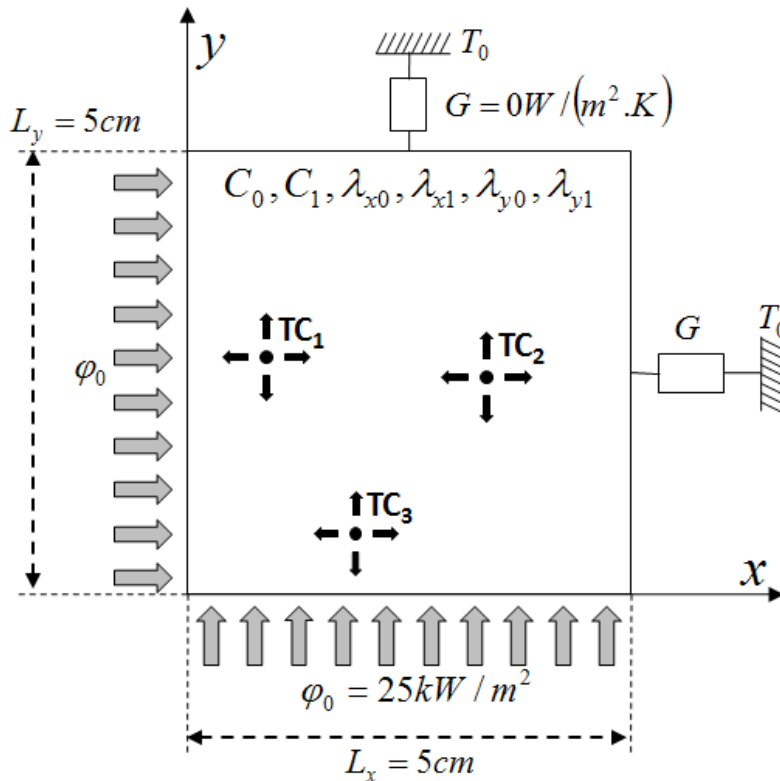
# Optimal experiment design (1/3)

## Uncertainty of estimated parameters:

-Effect of measurement noise:  $\text{cov}_\beta(\sigma_m) = \sigma_m^2 [X_\beta^T \cdot X_\beta]^{-1}$

-Effect of uncertain parameters:

$$\text{cov}_\beta(\sigma_\gamma) = [X_\beta^T \cdot X_\beta]^{-1} X_\beta^T X_\gamma \text{cov}(e_\gamma) X_\gamma^T X_\beta [X_\beta^T \cdot X_\beta]^{-1}$$



## Optimization problem

Where should I put thermocouple to estimate  $\beta$  with the best precision ?

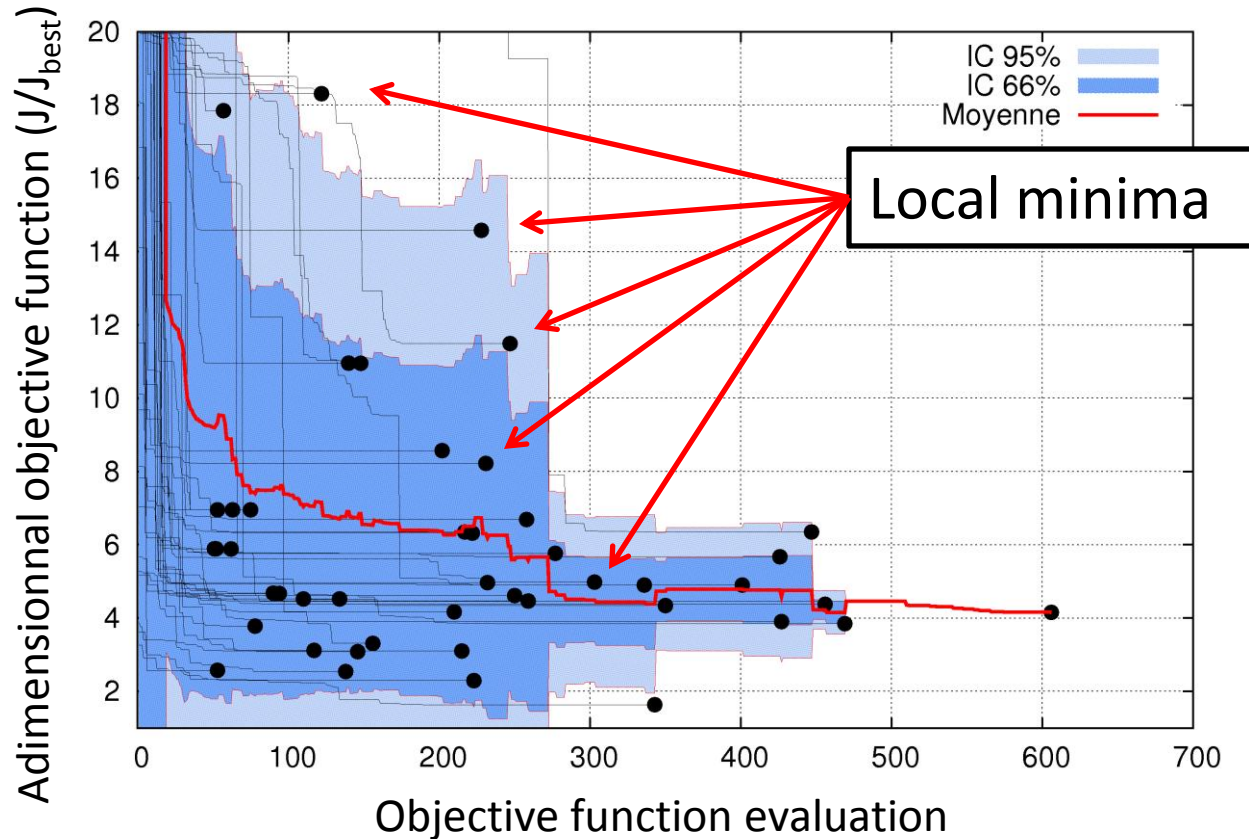
## Optimality criterion

$$J(P) = \sum_k \left( \frac{\sigma_{\beta_k}}{\beta_k} \right)^2 \quad \text{With } P \text{ the adjustable parameters}$$

**Solution**  $\hat{P} = \arg[\min_P (J(P))]$

# Optimal experiment design (2/3)

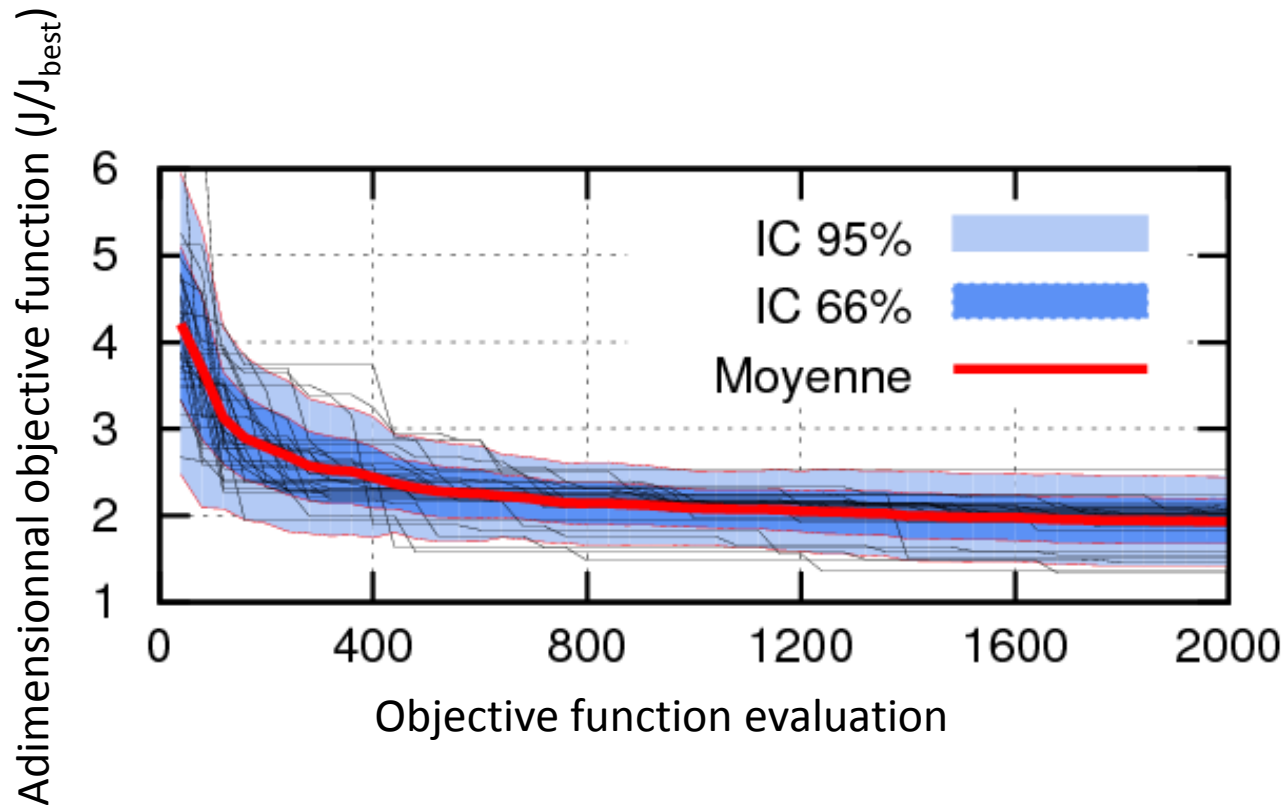
## Minimization using gradient-based methods



Several runs give several very different results

# Optimal experiment design (3/3)

## Minimization using PSO or EA

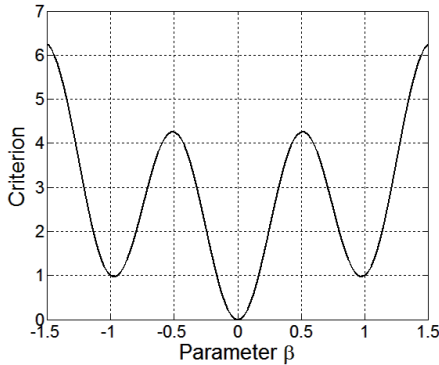


**Several runs give always a good result**

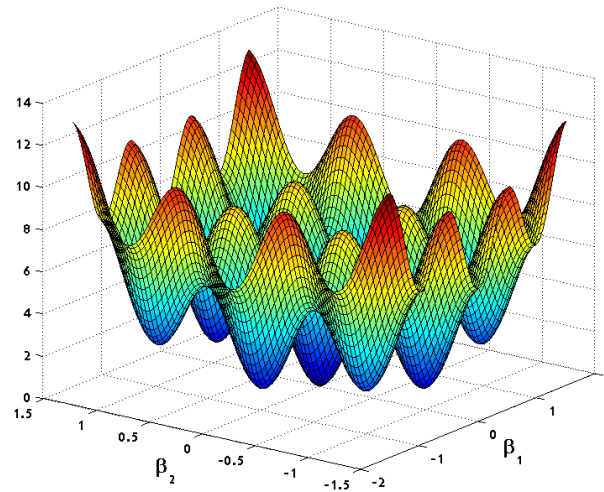
# Rastrigin function (1/2)

**Definition:** 
$$J(\beta) = A N + \sum_{i=1}^N [\beta_i^2 - A \cos(2\pi\beta_i)]$$

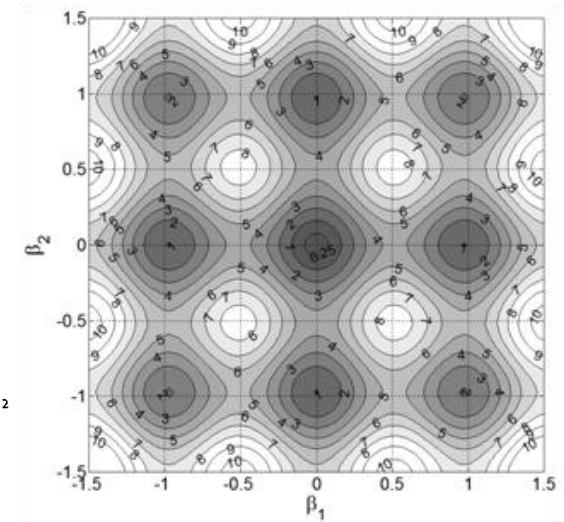
With N the dimension  
and A a constant.



**1d-Rastrigin function**

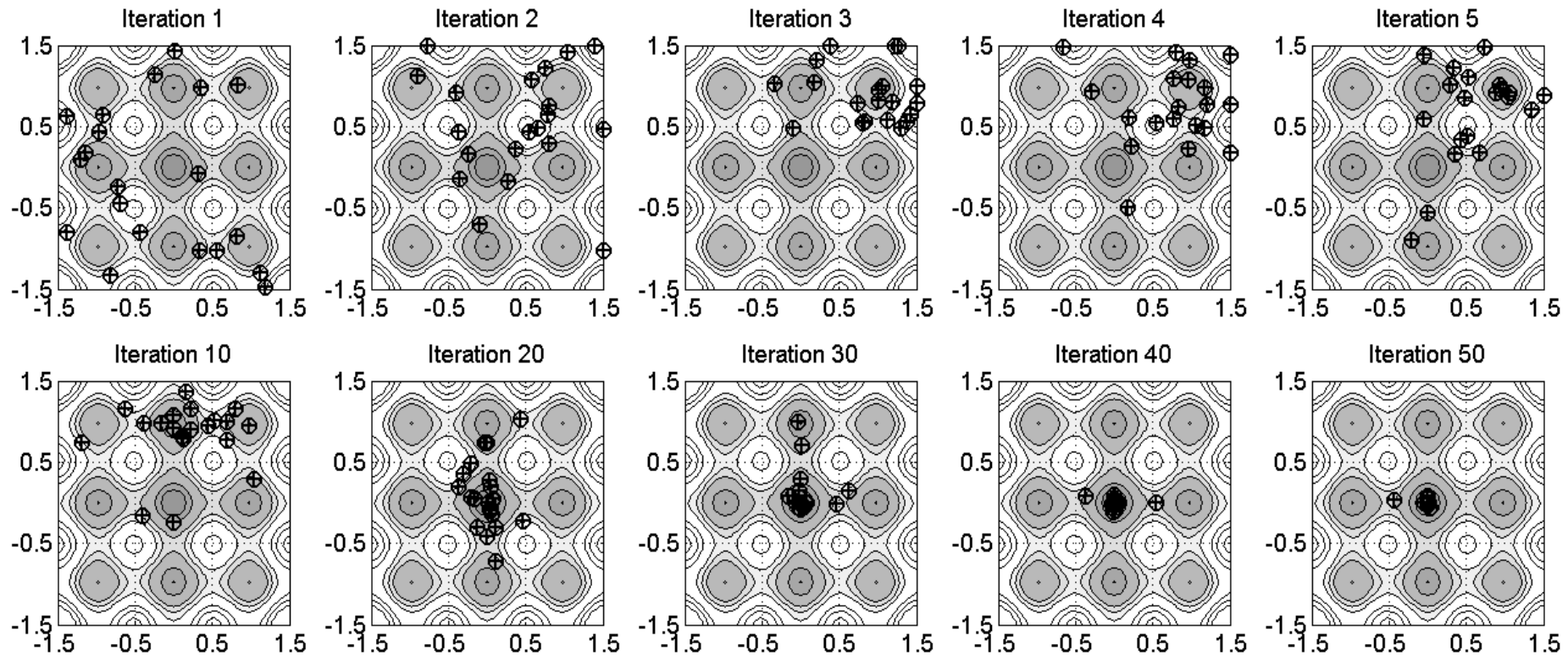


**2d-Rastrigin function**

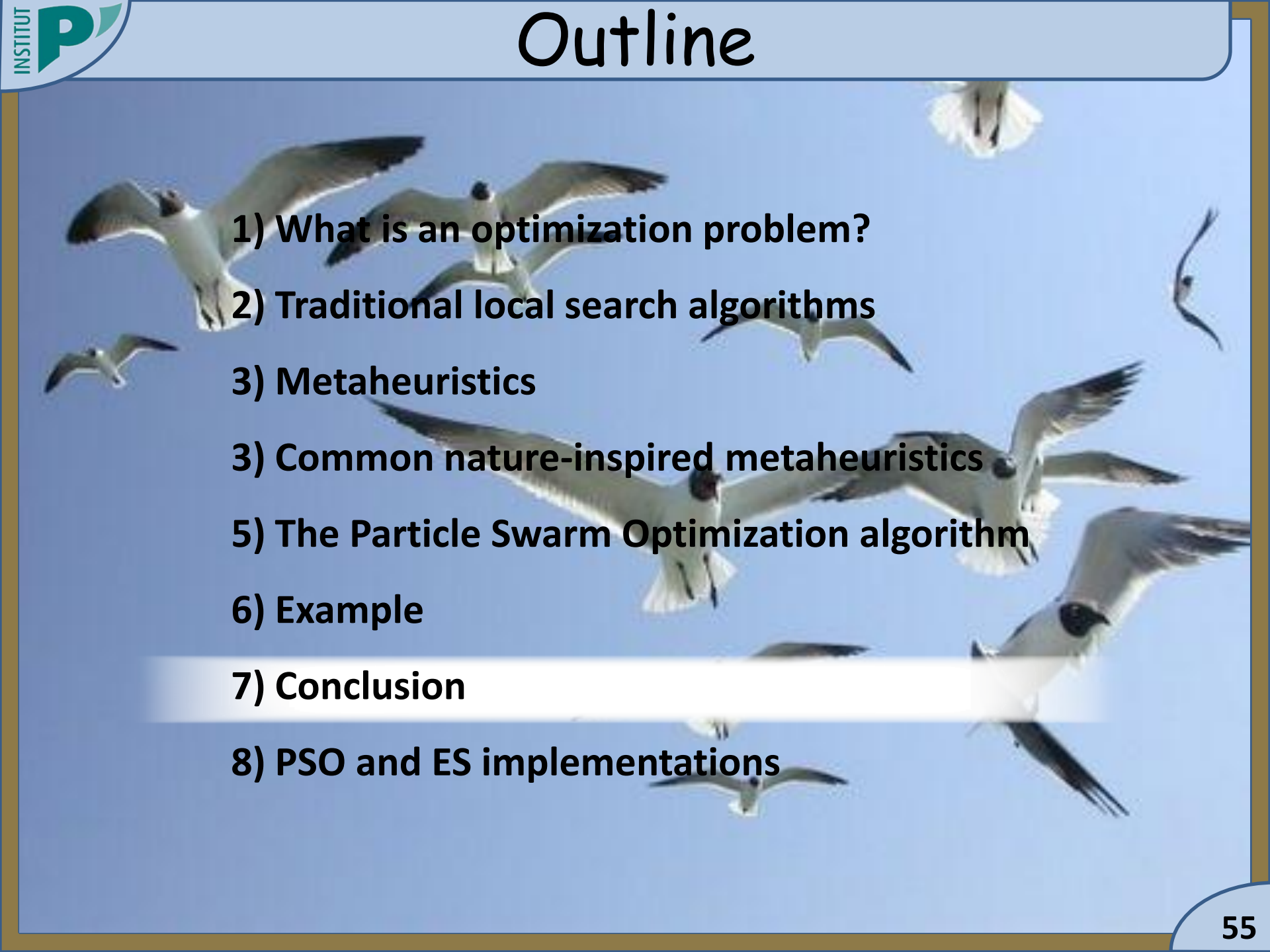


# Rastrigin function (2/2)

## Particle positions with respect to the iteration



# Outline

- 
- 1) What is an optimization problem?
  - 2) Traditional local search algorithms
  - 3) Metaheuristics
  - 3) Common nature-inspired metaheuristics
  - 5) The Particle Swarm Optimization algorithm
  - 6) Example
  - 7) Conclusion
  - 8) PSO and ES implementations

## Metaheuristics

- are designed to solve hard optimization problem
- try to find a compromise exploration/exploitation  
(somewhere between pure random and deterministic search)
- find good solutions in reasonable time (“approximation algorithms”)
- are often nature-inspired, mostly population-based, nearly always stochastic
- are often global, derivative-free

## Remarks

- Very few assumptions on the objective function shape
- The more information you provide, the more efficient the algorithm will be
- One never knows if the current best solution is the optimal one
- The choice of a termination criterion is not straightforward

## Good metaheuristic ?

- Not (too much) problem specific
- Easy to apply, easy to implement
- With only few parameters

## Best metaheuristic ?

- Such a thing doesn't exist
- Compared on all objective functions, all algorithms are equivalent ("No free lunch theorem")
- A algorithm can be more efficient on a specific set of objective functions
- Hybrid methods (Ex: PSO + Gradient)

# Conclusion (3/3)

## Metaheuristics

### Population-based

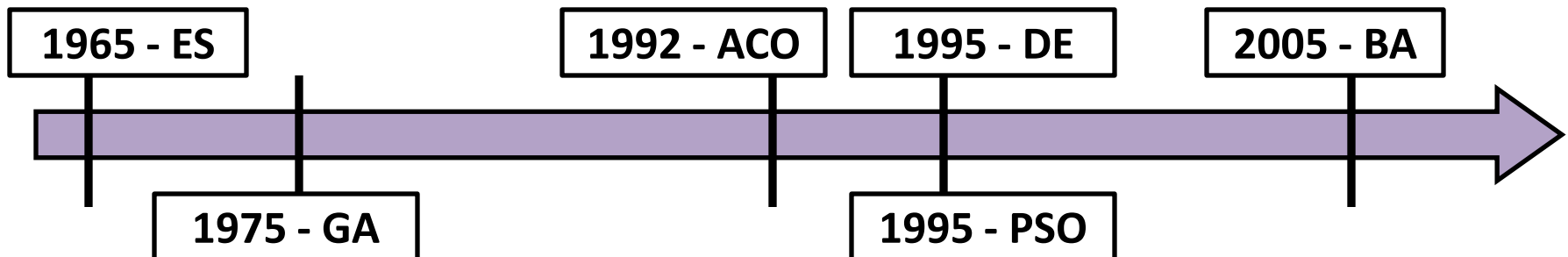
#### Nature-inspired

##### Evolutionary algorithms

- Evolution strategies (ES)*
- Genetic algorithms (GA)*
- Differential Evolution (DE)*

##### Swarm intelligence based

- Particle Swarm (PSO)*
- Ant colony (ACO)*
- Bee algorithm (BA)*



## ➤ Simplex

- C. Porte, Méthodes directes d'optimisation – Méthodes à une variable et Simplex, Techniques de l'Ingénieur

## ➤ Global optimization

- Weise, Global optimization algorithms, e-book, 2006
- J. Dréo, A. pétrowski, P. Siarry, E. Taillard. Métaheuristiques pour l'optimisation difficile. Paris : Editions Eyrolles, 2003 pp.70, 154. 2-212-11368-4

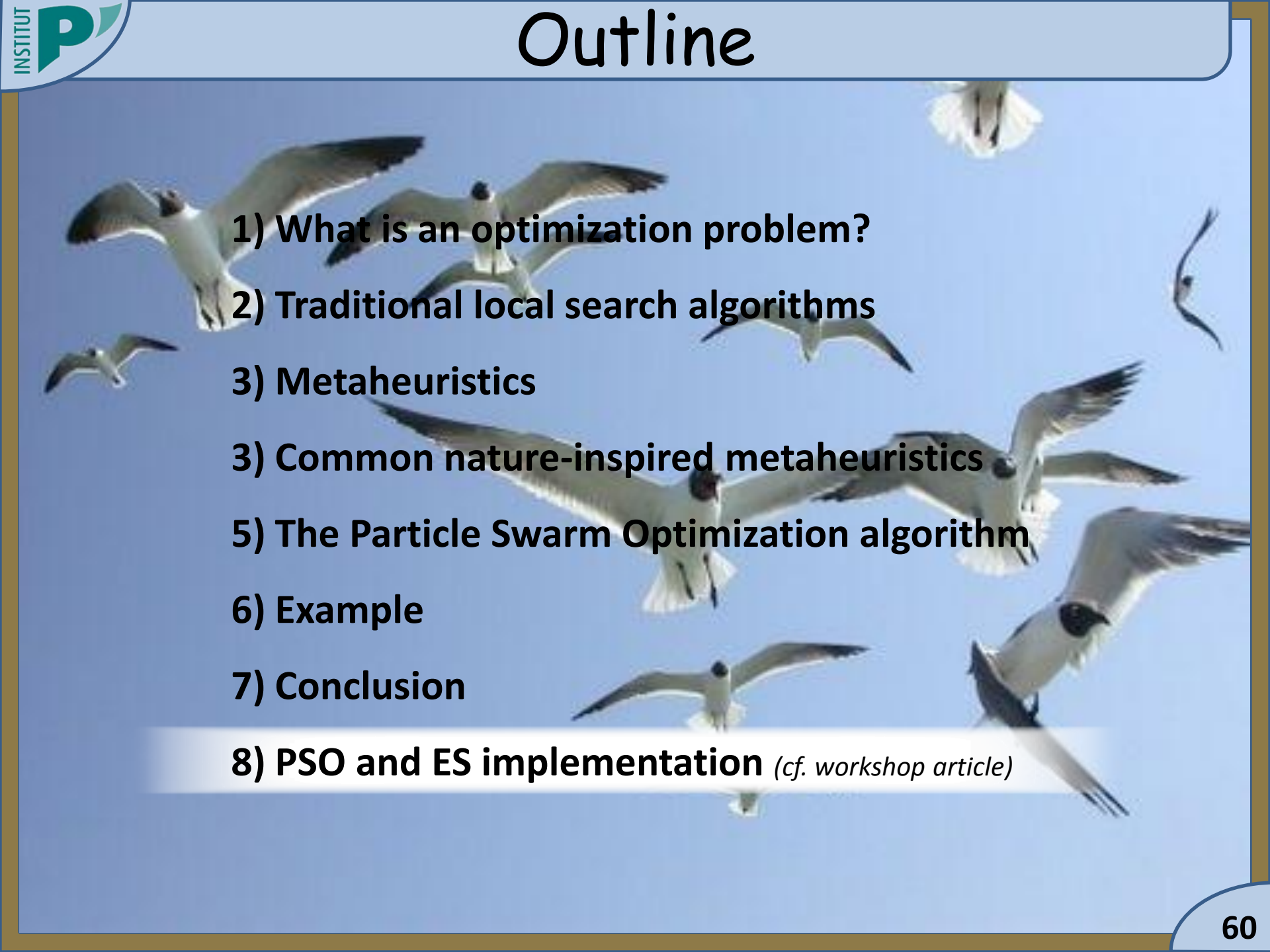
## ➤ Evolutionary algorithm

- Renders, J-M. Algorithmes génétiques et réseaux de neurones. Paris : Editions Hermès, 1995. 2-86601-467-7

## ➤ Swarm Intelligence based algorithm

- F. Van Den Bergh, An analysis of Particle Swarm Optimizers, PhD, University of Pretoria, November 2001
- J. Kennedy, Swarm Intelligence, Chapter 6, ook
- D. Merkle, M. Middendorf, Swarm Intelligence, Chapter 14, book
- F. Van Den Bergh, A.P. Engelbrecht, A study of particle swarm optimization particle trajectories, Information Sciences 176 (2006) 937-971
- C. Blum, Ant colony optimization: Introduction and recent trends, Physics of Life Reviews 2 (2005) 353-373
- G. Gilchev, I.C. Parmee, The Ant Colony Metaphor for searching Continuous Design Spaces, Proceedings of the AISB Workshop on Evolutionary Computation, April 3-4, 1995

# Outline

- 
- 1) What is an optimization problem?
  - 2) Traditional local search algorithms
  - 3) Metaheuristics
  - 3) Common nature-inspired metaheuristics
  - 5) The Particle Swarm Optimization algorithm
  - 6) Example
  - 7) Conclusion
  - 8) PSO and ES implementation *(cf. workshop article)*