Fitness Landscapes: the metaphor and beyond

Sébastien Verel

LISIC - Université du Littoral Côte d'Opale, Calais, France
http://www-lisic.univ-littoral.fr/~verel/

Tutorial LION9, January, 13th 2015

Situation map

Université du Littoral Côte d'Opale :
Boulogne, Calais, Dunkerque, Saint Omer.

Situation map

Université du Littoral Côte d'Opale :
Boulogne, Calais, Dunkerque, Saint Omer.

Outline

Fitness Landscapes, the metaphor and beyond :

1. Position and definition of fitness Landscape
2. Multimodal, rugged and neutral fitness landscapes
3. Beyond : graph point of view, Local Optima Network

Fitness landscapes : Motivations

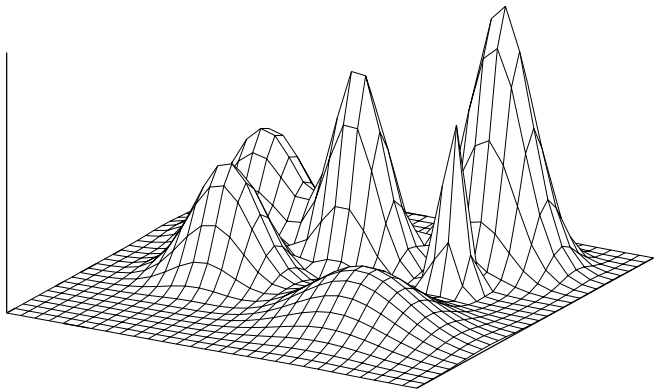
Basic idea

- Concept to study the search space
  from the point of view of local search
- Description of the search space "geometry"
  to design effective search algorithms

L. Barnett, U. Sussex, DPhil Diss. 2003

"the more we know of the **statistical properties** *of a class of fitness landscapes, the better equipped we will be for the* **design** *of effective search algorithms for such landscapes"*

Fitness landscapes in biology

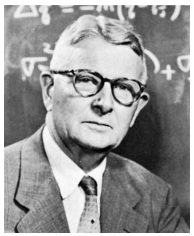


Origin in biological science :
S. Wright 1932 [45]

Biological evolution :

- Metaphorical uphill struggle across a "fitness landscape"
  - mountain **peaks** represent high "fitness" (ability to survive),
  - **valleys** represent low fitness.
- Evolution proceeds : population of organisms performs an "**adaptive walk**"

Original plots of S. Wright



S. Wright. "The roles of mutation, inbreeding, crossbreeding, and selection in evolution.", 1932.

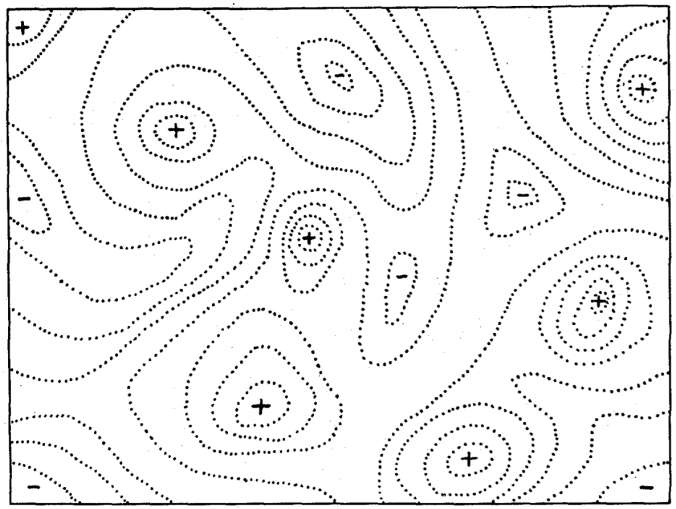


FIGURE 2.—Diagrammatic representation of the field of gene combinations in two dimensions instead of many thousands. Dotted lines represent contours with respect to adaptiveness.

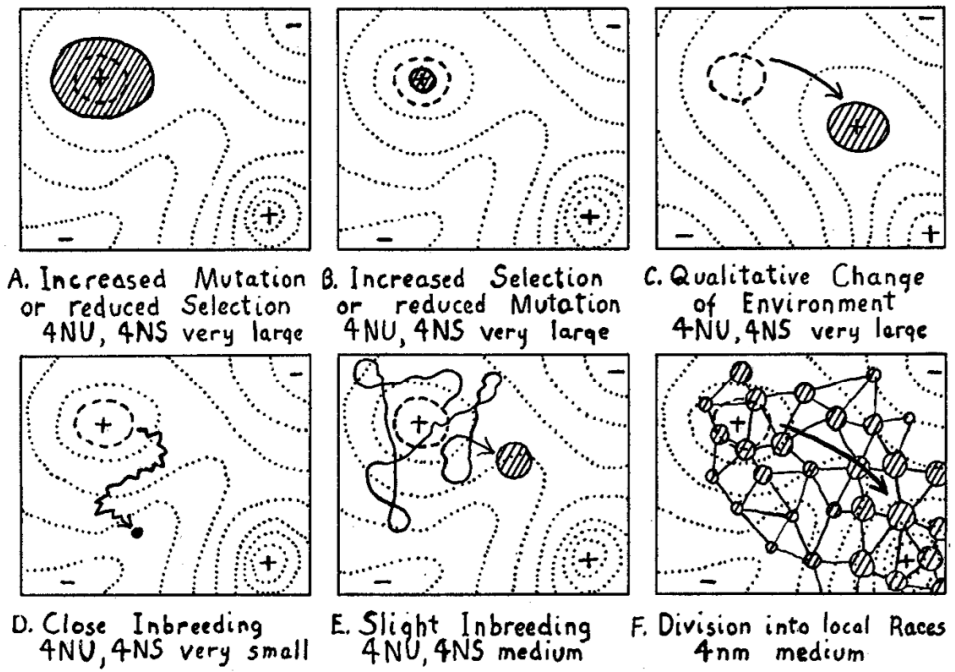
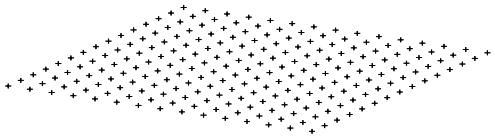
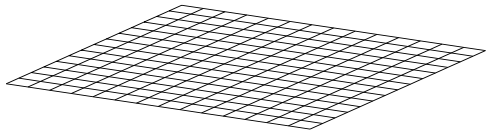


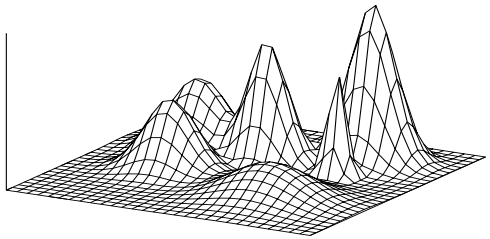
FIGURE 4.—Field of gene combinations occupied by a population within the general field of possible combinations. Type of history under specified conditions indicated by relation to initial field (heavy broken contour) and arrow.

photo : Wright, Sewall. Photograph. Encyclopædia Britannica Online. Web. 7 Jul. 2013.

Fitness landscapes in biology and others sciences







In biology :

- Model of species evolution

Extended to model dynamical systems :

- statistical physic,
- molecular evolution,
- ecology, etc.

Fitness landscapes in biology

2 sides of Fitness Landscapes

- **Metaphor** : most profound concept in evolutionary dynamics
  - give pictures of evolutionary process
  - be careful of misleading pictures :
    "smooth low-dimensional landscape without noise"
- **Quantitative** concept : predict the evolutionary paths
  - Quasispecies equation :
    mean field analysis with differential equations
  - Stochastic process :
    Markov chain
  - Network analysis

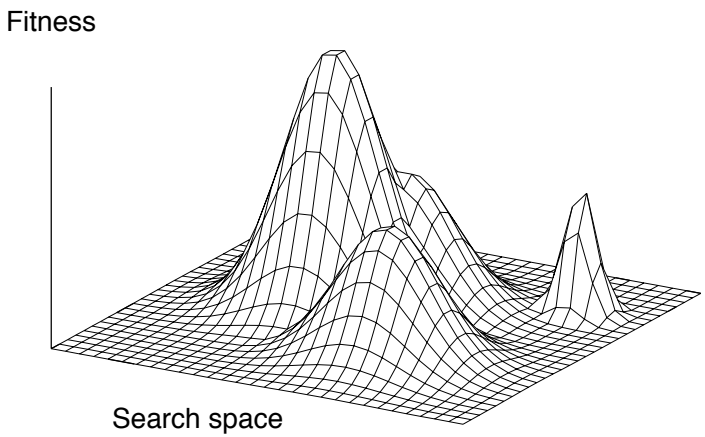
Fitness landscapes in biology

2 sides of Fitness Landscapes

- **Metaphor** : most profound concept in evolutionary dynamics
  - give pictures of evolutionary process
  - be careful of misleading pictures :
    "smooth low-dimensional landscape without noise"
- **Quantitative** concept : predict the evolutionary paths
  - Quasispecies equation :
    mean field analysis with differential equations
  - Stochastic process :
    Markov chain
  - Network analysis

Fitness landscape can be used
in the field of optimization !

Definition of fitness landscape for optimization

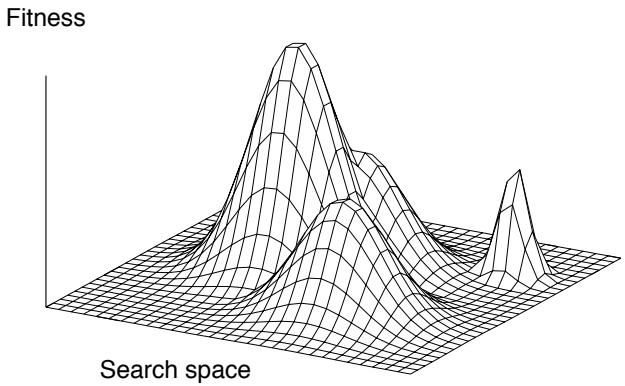


Definition

Fitness landscape $(\mathcal{S}, \mathcal{N}, f)$:

- **search space** :
$$\mathcal{S}$$
- **neighborhood relation** :
$$\mathcal{N} : \mathcal{S} \rightarrow 2^{\mathcal{S}}$$
- **objective function** :
$$f : \mathcal{S} \rightarrow \mathbb{R}$$

What is a neighborhood ?



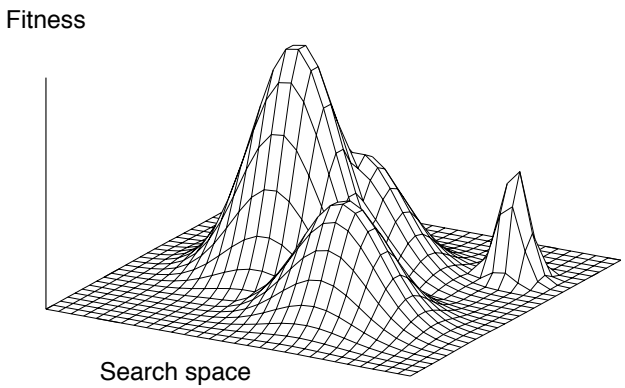
Neighborhood function :

$$\mathcal{N} : \mathcal{S} \rightarrow 2^{\mathcal{S}}$$

Set of "**neighbor**" solutions associated to each solution

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \Pr(y = op(x)) > 0\}$$

What is a neighborhood ?



Neighborhood function :

$$\mathcal{N} : \mathcal{S} \rightarrow 2^{\mathcal{S}}$$

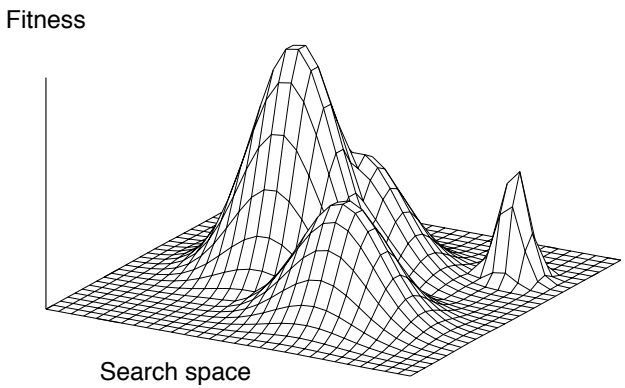
Set of "**neighbor**" solutions associated to each solution

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \Pr(y = op(x)) > 0\}$$

or

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \Pr(y = op(x)) > \varepsilon\}$$

What is a neighborhood ?



Neighborhood function :

$$\mathcal{N} : \mathcal{S} \rightarrow 2^{\mathcal{S}}$$

Set of "**neighbor**" solutions associated to each solution

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \Pr(y = op(x)) > 0\}$$

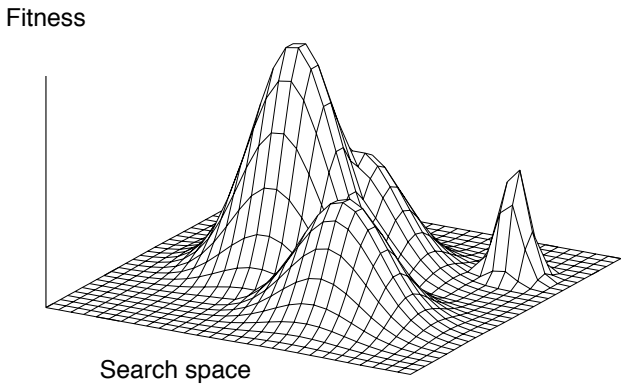
or

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \Pr(y = op(x)) > \varepsilon\}$$

or

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \text{distance}(x, y) = 1\}$$

What is a neighborhood ?



Neighborhood function :

$$\mathcal{N} : \mathcal{S} \rightarrow 2^{\mathcal{S}}$$

Set of "**neighbor**" solutions associated to each solution

Important !

Neighborhoood must be based on the operator(s) of the algorithm

Neighborhood $\Leftrightarrow$ Operator

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \Pr(y = op(x)) > 0\}$$
or
$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \Pr(y = op(x)) > \varepsilon\}$$
or
$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \text{distance}(x, y) = 1\}$$

Typical example : bit strings

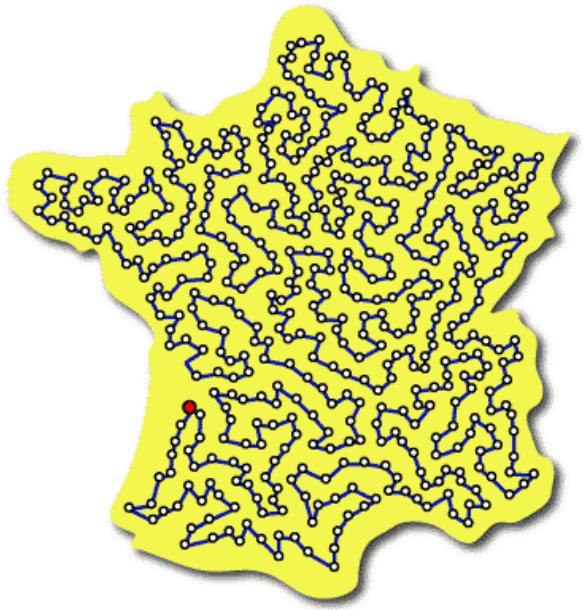
Search space : $\mathcal{S} = \{0,1\}^N$

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid d_{Hamming}(x,y) = 1\}$$

Example :
$\mathcal{N}(01101) = \{11101, 00101, 01001, 01111, 01100\}$

Typical example : permutations

Traveling Salesman Problem :
find the shortest tour which cross one time every town

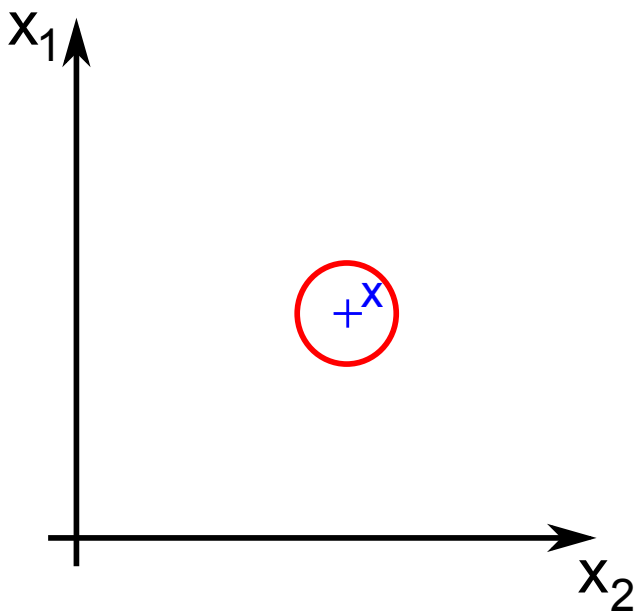


Search space : $\mathcal{S} = \{\sigma \mid \sigma \text{ permutations } \}$

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \Pr(y = op_{2opt}(x)) > 0\}$$

Not so typical example : continuous optimization

Still an open question...



Search space : $\mathcal{S} = [0,1]^d$

$$\mathcal{N}_\alpha(x) = \{y \in \mathcal{S} \mid \|y - x\| \leqslant \alpha\} \text{ with } \alpha > 0$$

More than 1 operator...

What can we do with 2 operators (ex : memetic algorithm) ?

$\mathcal{N}_1(x) = \{y \in \mathcal{S} \mid y = op_1(x)\}$ $\mathcal{N}_2(x) = \{y \in \mathcal{S} \mid y = op_2(x)\}$

More than 1 operator...

What can we do with 2 operators (ex : memetic algorithm) ?

$\mathcal{N}_1(x) = \{y \in \mathcal{S} \mid y = op_1(x)\}$ $\mathcal{N}_2(x) = \{y \in \mathcal{S} \mid y = op_2(x)\}$

Severals possibilities according to the goal :

- Study 2 landscapes : $(\mathcal{S}, \mathcal{N}_1, f)$ and $(\mathcal{S}, \mathcal{N}_2, f)$
- Study the landscape of "union" : $(\mathcal{S}, \mathcal{N}, f)$

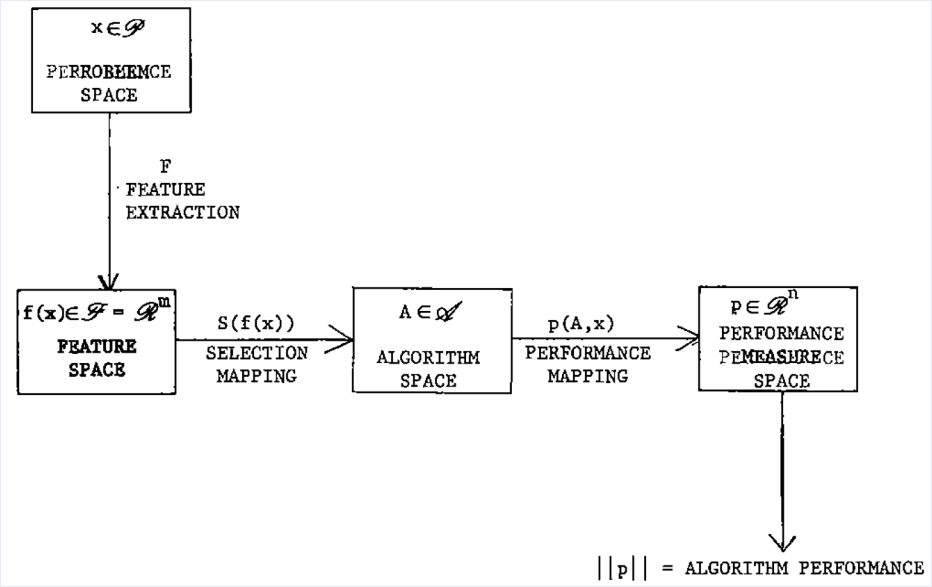
$$\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2 = \{y \in \mathcal{S} \mid y = op_1(x) \text{ or } y = op_2(x)\}$$

- Study the landscape of "composition" : $(\mathcal{S}, \mathcal{N}, f)$

$$\mathcal{N} = \{y \in \mathcal{S} \mid y = op \circ op'(x) \text{ with } op, op' \in \{id, op_1, op_2\}\}$$

Position of fitness landscape analysis

Rice framework (1976)



Malan framework (2014)

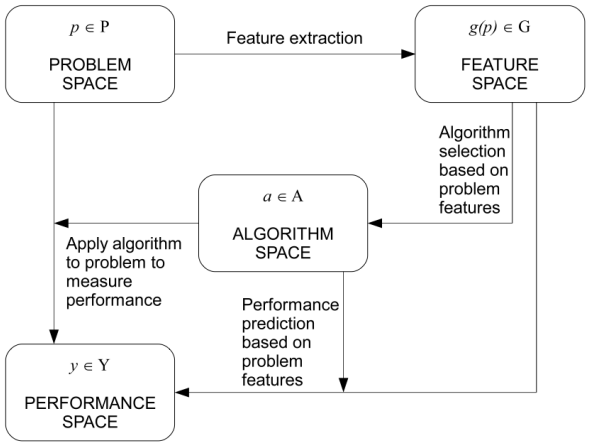
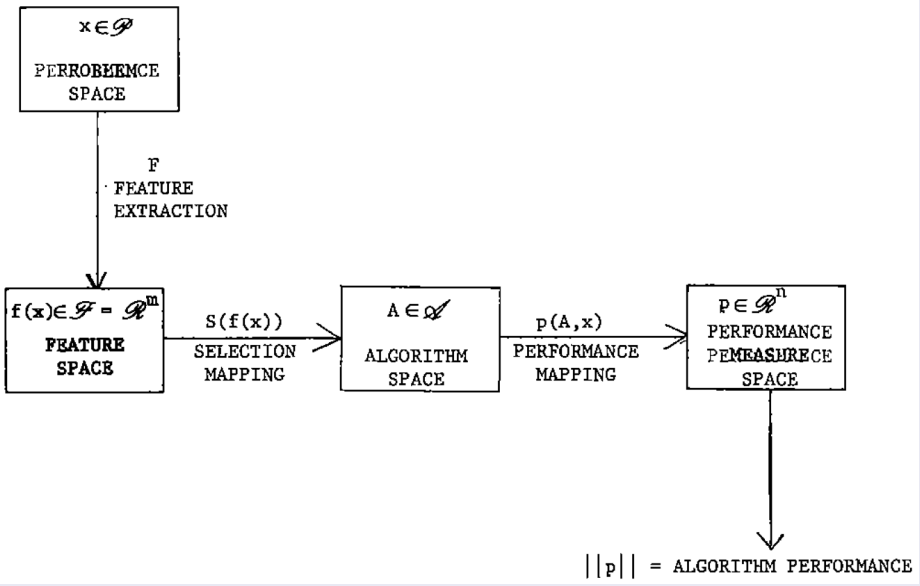


Figure 1.1: A framework for describing the general problems of algorithm selection and performance prediction based on problem features (based Rice's model [132]).

Position of fitness landscape analysis

Rice framework (1976)



Malan framework (2014)

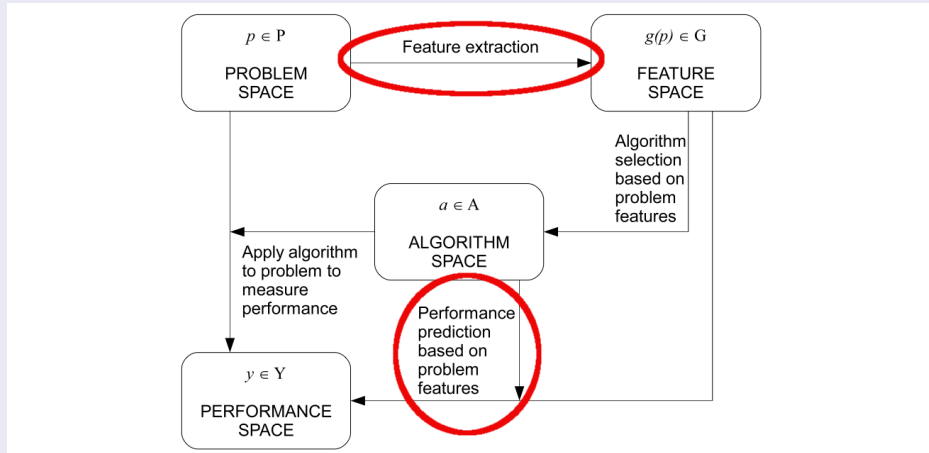


Figure 1.1: A framework for describing the general problems of algorithm selection and performance prediction based on problem features (based Rice's model [132]).

Fitness landscape analysis : features extraction *vs.* performance

Fitness landscape : Tool for black-box optimization

Black-box optimization

We have only $\{(x_0, f(x_0)), (x_1, f(x_1)), ...\}$
the computation of f is given by an "oracle"

No information is either not available or needed on the definition of objective function

- Objective function given by a computation, or a simulation
- Objective function can be irregular, non differentiable, non continuous, etc.

$\Rightarrow$ fitness landscape is "well-suited" for black-box optimization

FL and BBO share the same foundation :
no *a priori* definition or properties on objective function f are required

Goal of the fitness landscapes analysis

- Geometry/features of fitness landscape
  $\Rightarrow$ dynamics/performance of a local search algorithm
- Geometry is linked to the problem difficulty :
  - "If there are a lot of local optima, the probability to find the global optimum is lower."
  - "If the fitness landscape is flat, discovering better solutions is rare."
  - "What is the best search direction in the landscape ?"
  - ...

Study of the fitness landscape **features**
allows to study
the **performance** of search algorithms

Fitness landscapes analysis in "practice"

1. **To compare the difficulty of two search spaces** :
   - One problem, different codings : $(\mathcal{S}_1, \mathcal{N}_1, f_1)$ and $(\mathcal{S}_2, \mathcal{N}_2, f_2)$ different coding, mutation operator, objective function, etc.

     Which one is easier to solve ? [*e.g. see next papers !*]

Fitness landscapes analysis in "practice"

1. **To compare the difficulty of two search spaces** :
   - One problem, different codings : $(\mathcal{S}_1, \mathcal{N}_1, f_1)$ and $(\mathcal{S}_2, \mathcal{N}_2, f_2)$ different coding, mutation operator, objective function, etc.

     Which one is easier to solve ? [*e.g. see next papers !*]
2. **To choose the algorithm** :
   - analysis of global geometry of the landscape

     Which algorithm can I use ? [*e.g. see next paper !*]

Fitness landscapes analysis in "practice"

1. **To compare the difficulty of two search spaces** :
   - One problem, different codings : $(\mathcal{S}_1, \mathcal{N}_1, f_1)$ and $(\mathcal{S}_2, \mathcal{N}_2, f_2)$ different coding, mutation operator, objective function, etc.

     Which one is easier to solve ? [*e.g. see next papers !*]
2. **To choose the algorithm** :
   - analysis of global geometry of the landscape

     Which algorithm can I use ? [*e.g. see next paper !*]
3. **To tune the parameters** :
   - *off-line* analysis of structure of fitness landscape

     Which is the best mutation operator ? the size of the population ? number of restarts ? etc. [*see LION papers*]

Fitness landscapes analysis in "practice"

1. **To compare the difficulty of two search spaces** :
   - One problem, different codings : $(\mathcal{S}_1, \mathcal{N}_1, f_1)$ and $(\mathcal{S}_2, \mathcal{N}_2, f_2)$ different coding, mutation operator, objective function, etc.

     Which one is easier to solve ? [*e.g. see next papers !*]
2. **To choose the algorithm** :
   - analysis of global geometry of the landscape

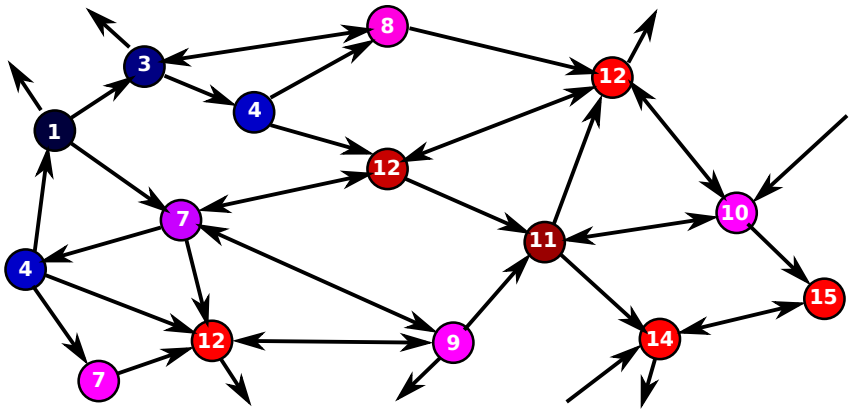
     Which algorithm can I use ? [*e.g. see next paper !*]
3. **To tune the parameters** :
   - *off-line* analysis of structure of fitness landscape

     Which is the best mutation operator ? the size of the population ? number of restarts ? etc. [*see LION papers*]
4. **To control the parameters during the run** :
   - *on-line* analysis of structure of fitness landscape

     Which is the optimal mutation operator according to the estimation of the structure ? [*see LION papers*]

Back to definition : graph point of view

Fitness landscape $(\mathcal{S}, \mathcal{N}, f)$ is :
an oriented graph $(\mathcal{S}, \mathcal{N})$ with valuated nodes given by f.

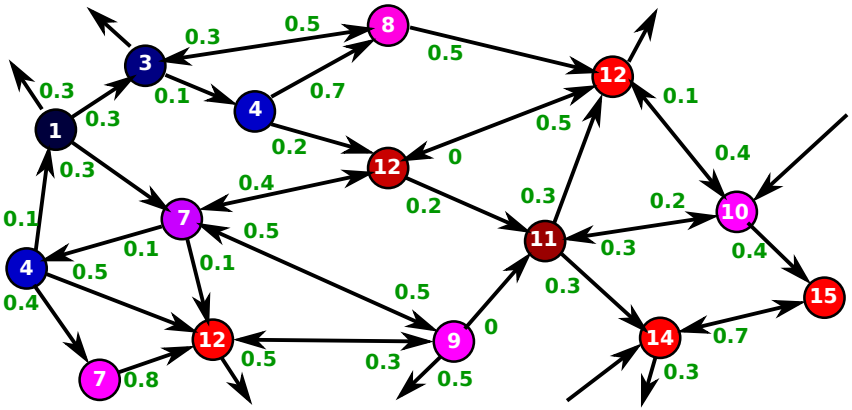


Remarks :

- Model of the search space
- Non specific to a particular local search

Back to definition : graph point of view

Fitness landscape $(\mathcal{S}, \mathcal{N}, f)$ is :
an oriented graph $(\mathcal{S}, \mathcal{N})$ with valuated nodes given by f.



Remarks :

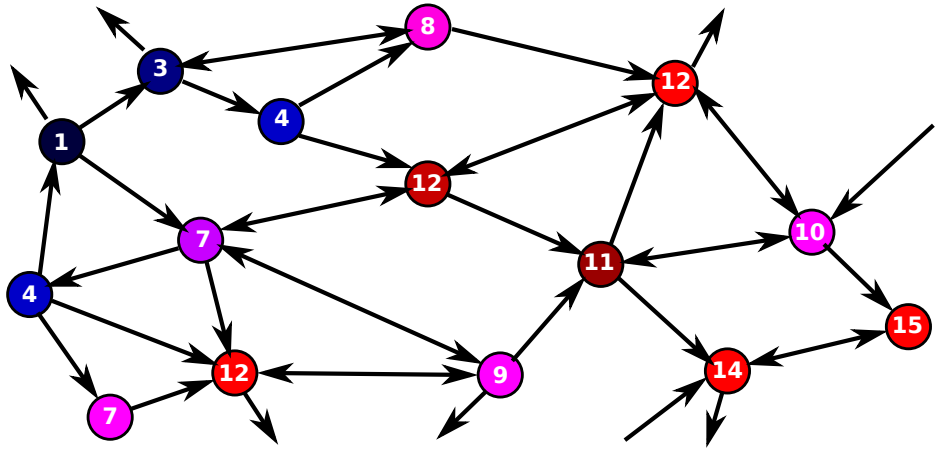
- Model of the search space
- Non specific to a particular local search
- A specific local search puts probability transitions on **edges** according to **f** and **history** of the search

Fitness landscape and complex systems

Complex system : local *vs.* global properties

- Sample the neighborhood to have information on **local features** of the search space
- From local information : deduce **global feature** such as general shape, difficulty, performance, etc.

$\Rightarrow$ Analysis using complex systems tools :



Short summary of this part

Study of the **structure** of the fitness landscape
allows to study the **difficulty**, and
allows to **design** good optimization algorithms

Fitness landscape is a **graph** $(\mathcal{S}, \mathcal{N}, f)$:

- nodes are solutions which have a value (fitness),
- edges are defined by the neighborhood relation.

pictured as a real landscape

Short summary of this part

Study of the **structure** of the fitness landscape
allows to study the **difficulty**, and
allows to **design** good optimization algorithms

Fitness landscape is a **graph** $(\mathcal{S}, \mathcal{N}, f)$:

- nodes are solutions which have a value (fitness),
- edges are defined by the neighborhood relation.

pictured as a real landscape

Next section : study of the two main geometries

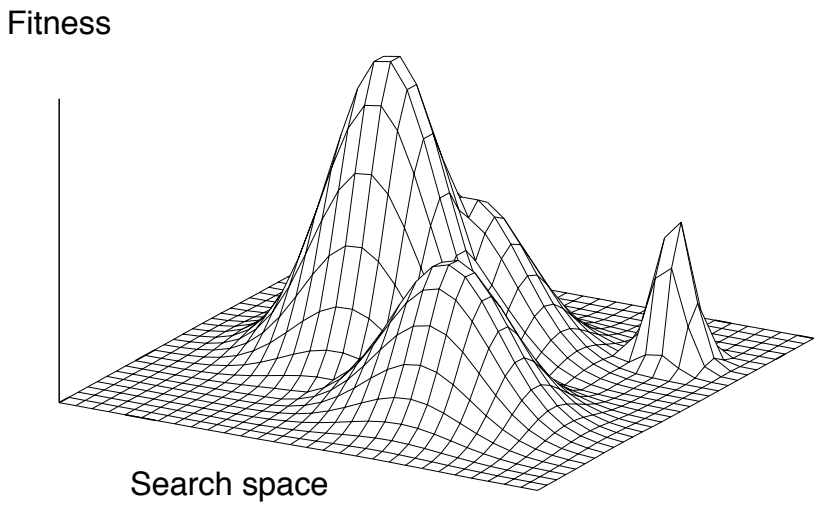
- multimodal and ruggedness
- neutral

Multimodal Fitness landscapes

Local optima s^*

no neighbor solution with higher fitness value

$$\forall s \in \mathcal{N}(s^*), \quad f(s) \leqslant f(s^*)$$



Typical example : bit strings

Search space : $\mathcal{S} = \{0,1\}^N$

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid d_{Hamming}(x,y) = 1\}$$

Example :
$x = 01101$ and $f_1(x) = f_2(x) = f_3(x) = 5$

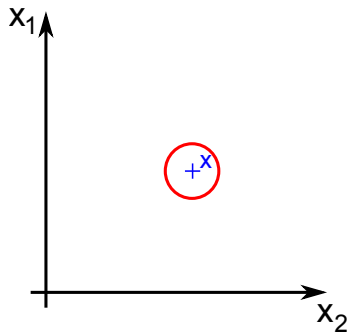
	11101	00101	01001	01111	01100
f_1	4	2	3	0	3
f_2	2	3	6	2	3
f_3	1	5	2	2	4

Question

Is x is a local maximum for f_1, f_2, and/or f_3 ?

Not so typical example : continuous optimization

Still an open question...



Search space : $\mathcal{S} = [0,1]^d$

$$\mathcal{N}_\alpha(x) = \{y \in \mathcal{S} \mid \|y - x\| \leqslant \alpha\}$$
with $\alpha > 0$

Classical definition of local optimum

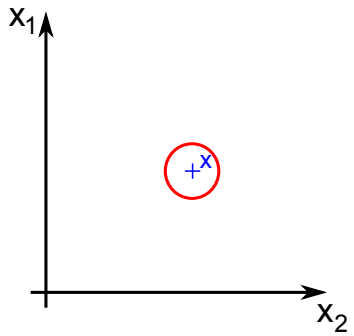
x is local maximum iif
$\exists \varepsilon > 0, \forall y$ such that $\|y - x\| \leqslant \varepsilon,\ f(y) \leqslant f(x)$

Questions

Local search definition with $\mathcal{N}_\alpha \Rightarrow$ classical definition ?
Classical definition $\Rightarrow$ local search definition with $\mathcal{N}_\alpha$?

Not so typical example : continuous optimization

Still an open question...



Search space : $\mathcal{S} = [0,1]^d$

$$\mathcal{N}_\alpha(x) = \{y \in \mathcal{S} \mid \|y - x\| \leqslant \alpha\}$$
with $\alpha > 0$

Classical definition of local optimum

x is local maximum iif
$\exists \varepsilon > 0, \forall y$ such that $\|y - x\| \leqslant \varepsilon,\ f(y) \leqslant f(x)$

Questions

Local search definition with $\mathcal{N}_\alpha \Rightarrow$ classical definition ?
Classical definition $\Rightarrow$ local search definition with $\mathcal{N}_\alpha$?

Still some works to do...

Multimodal Fitness landscapes : find local optima

Adaptive walk

$(s_0, s_1, \ldots)$ where $s_{i+1} \in \mathcal{N}(s_i)$ and $f(s_i) < f(s_{i+1})$

Multimodal Fitness landscapes : find local optima

Adaptive walk

$(s_0, s_1, \ldots)$ where $s_{i+1} \in \mathcal{N}(s_i)$ and $f(s_i) < f(s_{i+1})$

Hill-Climbing algorithm (best-improvement)

Choose initial solution $s \in \mathcal{S}$
repeat
  choose $s' \in \mathcal{N}(s)$ such that $f(s') = max_{y \in \mathcal{N}(s)} f(y)$
  if $f(s) < f(s')$ **then**
    $s \leftarrow s'$
  end if
until s is a Local Optimum

Multimodal Fitness landscapes : find local optima

Adaptive walk

$(s_0, s_1, \ldots)$ where $s_{i+1} \in \mathcal{N}(s_i)$ and $f(s_i) < f(s_{i+1})$

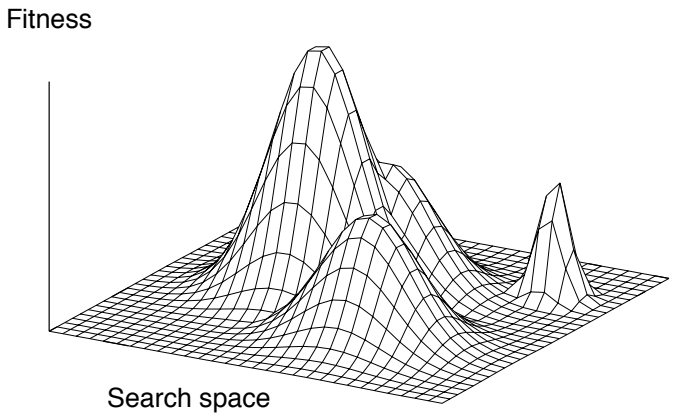
Hill-Climbing algorithm (best-improvement)

Choose initial solution $s \in \mathcal{S}$
repeat
  choose $s' \in \mathcal{N}(s)$ such that $f(s') = max_{y \in \mathcal{N}(s)} f(y)$
  if $f(s) < f(s')$ **then**
    $s \leftarrow s'$
  end if
until s is a Local Optimum

Basin of attraction of s^*

$$\{s \in \mathcal{S} \mid HillClimbing(s) = s^*\}.$$

Multimodal Fitness landscapes and difficulty



The idea :

- if the size of attractive basin of global optimum is "small",
- then, the "time" to find the global optimum is "long"

Optimisation difficulty :

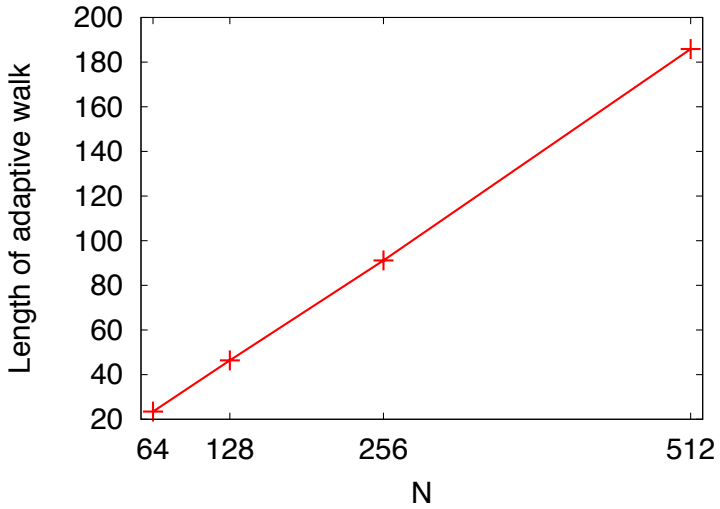
Number and size of attractive basins (Garnier *et al* [10])

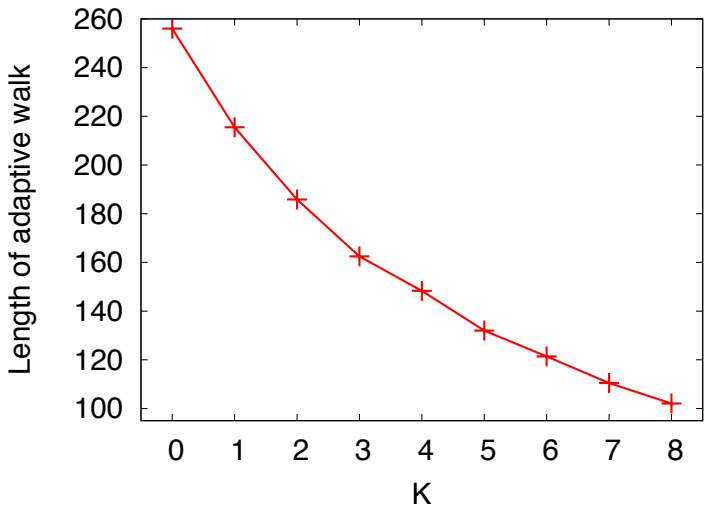
Feature to estimate basin size :

- **Length of adaptive walks**

complexity : sample size $\times$ l $\times|\mathcal{N}|$

Multimodal Fitness landscapes and difficulty





ex. nk-landscapes with $n = 512$

The idea :

- if the size of attractive basin of global optimum is "small",
- then, the "time" to find the global optimum is "long"

Optimisation difficulty :

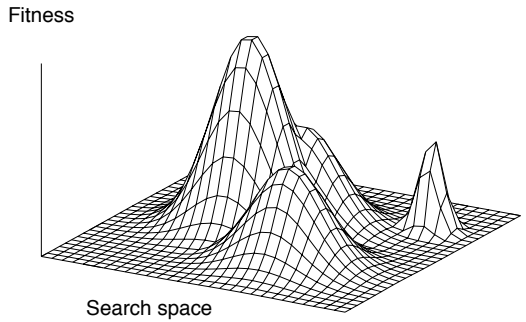
Number and size of attractive basins (Garnier *et al* [10])

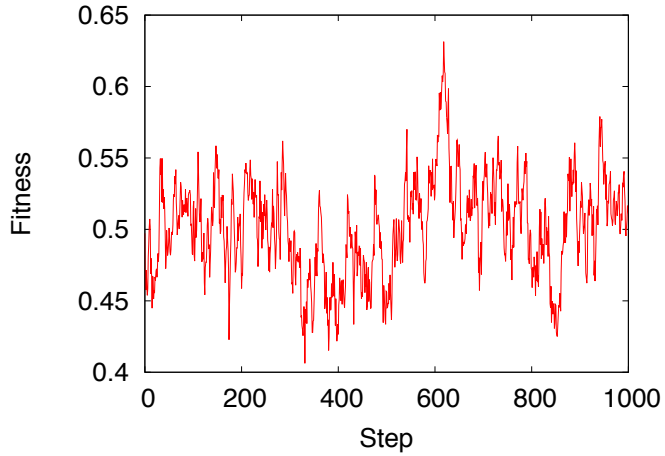
Feature to estimate basin size :

- **Length of adaptive walks**

complexity : sample size $\times$ l $\times |\mathcal{N}|$

Random Walk to measure the ruggedness





Random walk :

- $(s_1, s_2, \ldots)$ where $s_{i+1} \in \mathcal{N}(s_i)$ and equiprobability on $\mathcal{N}(s_i)$

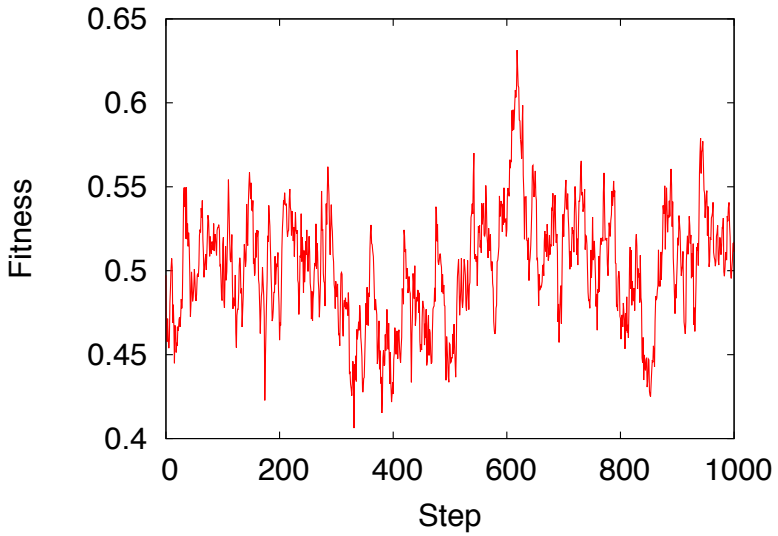
The idea :

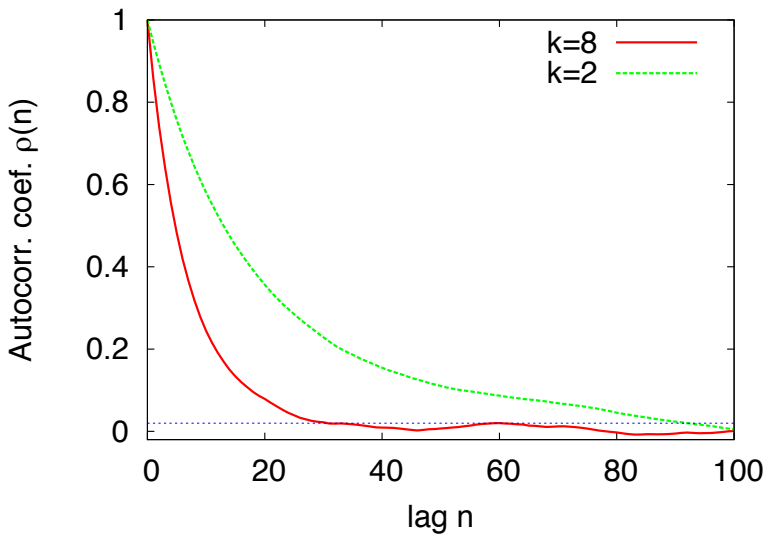
- if the profile of fitness is irregular,
- then, the "information" between neighbors is low.

Feature :

- Study the fitness profile like a signal

Rugged/smooth fitness landscapes





Autocorrelation function of time series of fitnesses along a random walk (Weinberger 90 [37]) :

$$\rho(n) = \frac{\mathrm{E}[(f(s_i) - \bar{f})(f(s_{i+n}) - \bar{f})]}{\mathrm{var}(f(s_i))}$$

Autocorrelation length $\tau = \frac{1}{\rho(1)}$

"How many random steps such that correlation becomes insignificant"

- small τ : rugged landscape
- long τ : smooth landscape

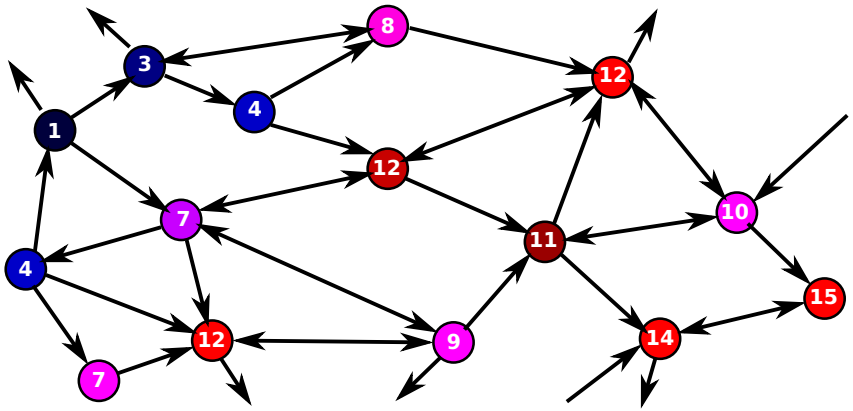
complexity : sample size $\approx 10^3$

Results on rugged fitness landscapes (Stadler 96 [28])

Ruggedness decreases with the size of thoses problems

Problem	parameter	$\rho(1)$
symmetric TSP	n number of towns	$1 - \frac{4}{n}$
anti-symmetric TSP	n number of towns	$1 - \frac{4}{n-1}$
Graph Coloring Problem	n number of nodes α number of colors	$1 - \frac{2\alpha}{(\alpha-1)n}$
NK landscapes	N number of proteins K number of epistasis links	$1 - \frac{K+1}{N}$
random max-k-SAT	n number of variables k variables per clause	$1 - \frac{k}{n(1-2^{-k})}$

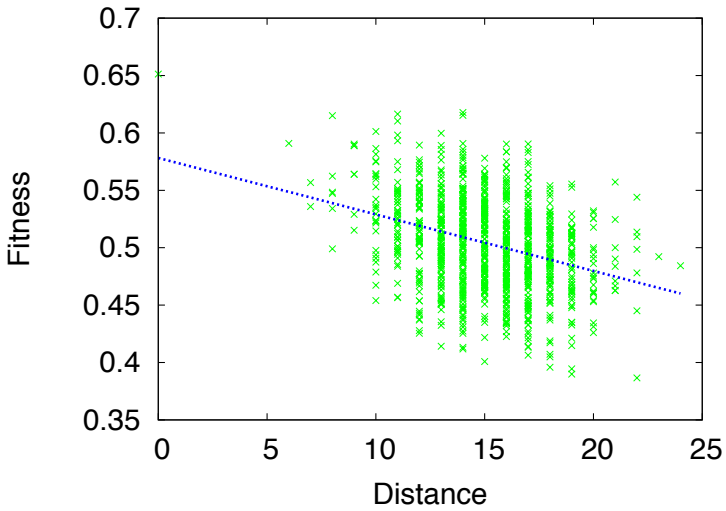
Results on rugged fitness landscapes (Stadler 96 [28])



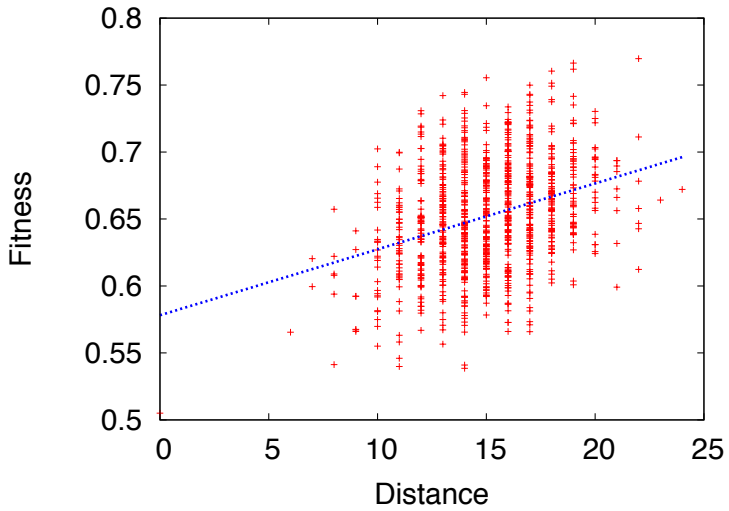
Related to Laplacian graph, elementary landscapes

Fitness Distance Correlation (FDC) (Jones 95 [15])

Correlation between fitness and distance to global optimum



easy



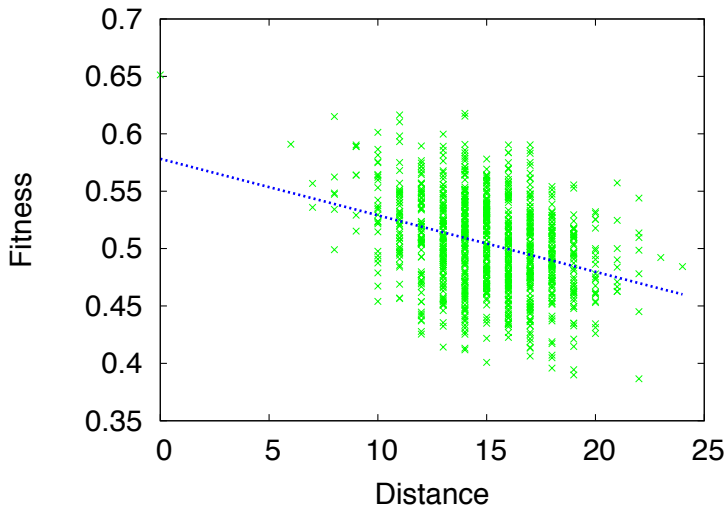
hard

Classification based on experimental studies :

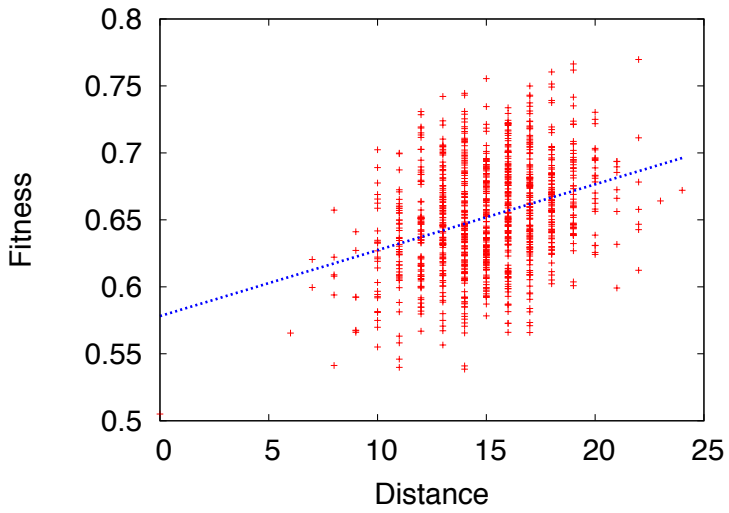
- $\rho < -0.15$, easy optimization
- $\rho > 0.15$, hard optimization
- $-0.15 < \rho < 0.15$, undecided zone

Fitness Distance Correlation (FDC) (Jones 95 [15])

Correlation between fitness and distance to global optimum



easy



hard

- Important concept to understand
- Not useful in "practice" (difficult to estimate)

Short break

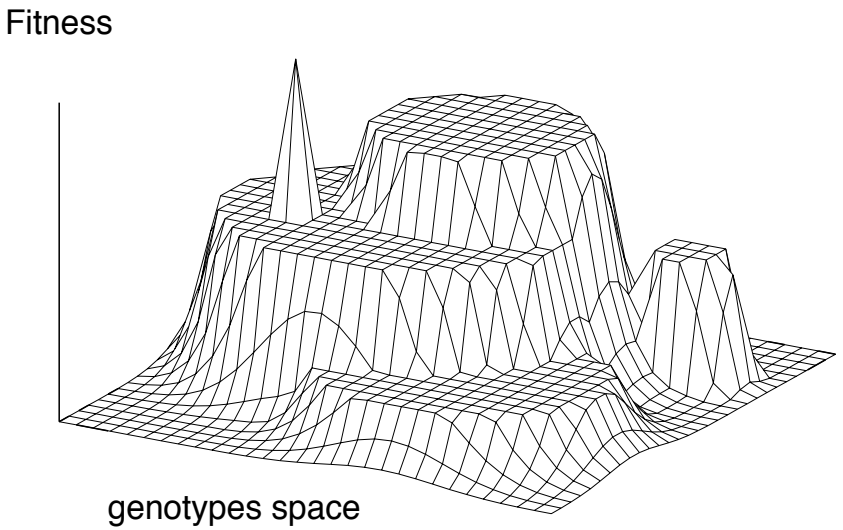
Why not playing to a not so serious game :
`http://www-lisic.univ-littoral.fr/~verel/RESEARCH/notSeriousGame/index.html`

Neutral Fitness Landscapes

Neutral theory (Kimura $\approx$ 1960 [17])

Theory of mutation and random drift

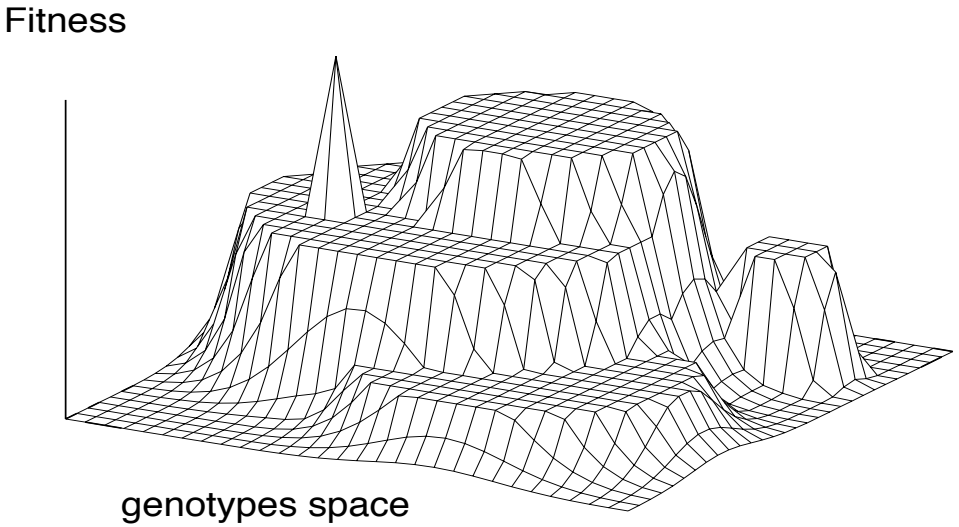
A considerable number of mutations have no effects on fitness values



- plateaus
- neutral degree
- neutral networks [Schuster 1994 [27], RNA folding]

Neutral Fitness Landscapes

- Redundant problem (symmetries, etc.) [12]
- Problem "not well" defined or dynamic environment [14]
- Unused variables, discrete values, etc.



Real-world problems :

- Robot controler
- Circuit design
- Genetic Programming
- Protein folding
- Learning problems
- Scheduling problems
- Graph problems...

Neutrality and difficulty

- In our knowledge, there is no definitive answer about the relation between neutrality and problem hardness
- Certainly, it is dependent on the "nature" of neutrality

Solving optimization problem and neutrality

3 ways to deals with neutrality :

- Decrease the neutrality : reduce the entropy barrier
- Increase the neutrality : reduce the fitness barrier
- Unchange the neutrality : use a specific algorithm

Sharp description of the geometry
of neutral fitness landscapes is needed

Neutrality and difficulty

We know for certain that :

- **No information** is better than **Bad information** :
  From a non-optimal solution, hard trap functions are more difficult than needle-in-a-haystack functions
- **Good information** is better than **No information** :
  Onemax problem is much easier than needle-in-a-haystack functions

Neutrality and difficulty

We know for certain that :

- **No information** is better than **Bad information** :
  From a non-optimal solution, hard trap functions are more difficult than needle-in-a-haystack functions
- **Good information** is better than **No information** :
  Onemax problem is much easier than needle-in-a-haystack functions

- When there is No information :
  you should have a good method to create it !

Objects of neutral fitness landscapes

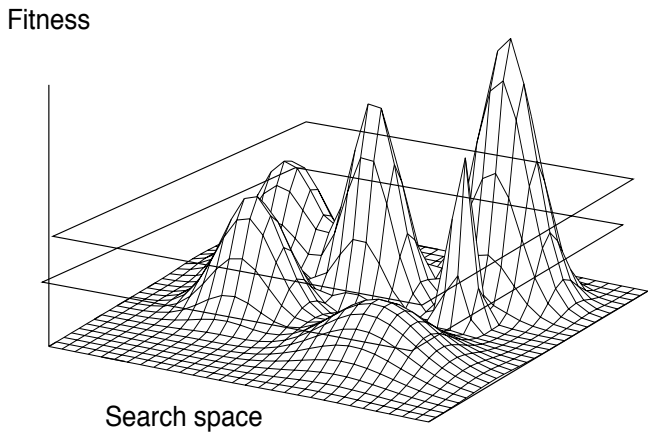
Description of multimodal fitness landscapes is based on :

- Local optima
- Basins of attraction

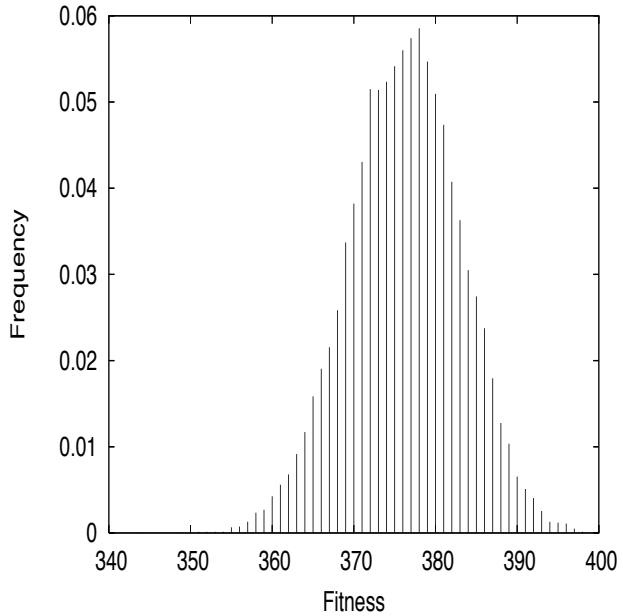
Description of neutral fitness landscapes is based on :

- **Neutral sets** :
  set of solutions with the same fitness
- **Neutral networks** :
  neutral sets with neighborhood relation

Neutral sets : Density Of States



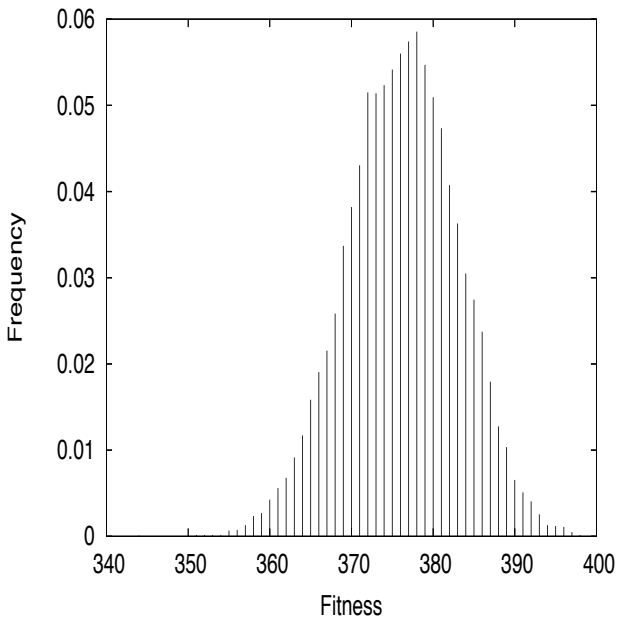
Set of solutions with fitness value



Density of states (D.O.S.)

- Introduce in physics (Rosé 1996 [26])
- Optimization (Belaidouni, Hao 00 [4])

Neutral sets : Density Of States



Informations given :

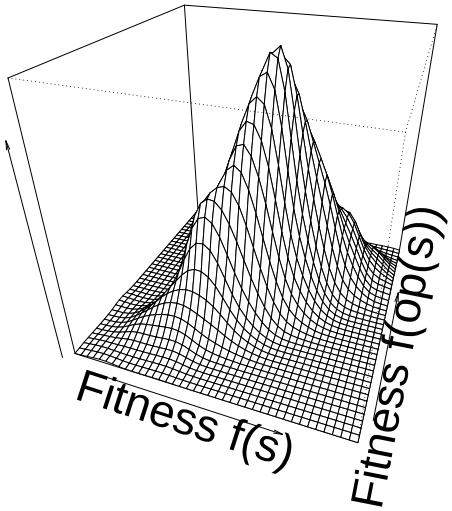
- Performance of random search
- Tail of the distribution is an indicator of difficulty :
  - the faster the decay, the harder the problem
- But do not care about the neighborhood relation

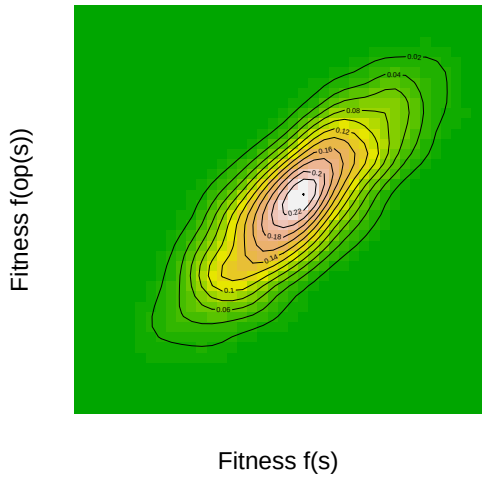
Features :

- **Average**, **sd**, **kurtosis**, etc.

complexity : sample size

Neutral sets : Fitness Cloud [Verel et al. 2003]





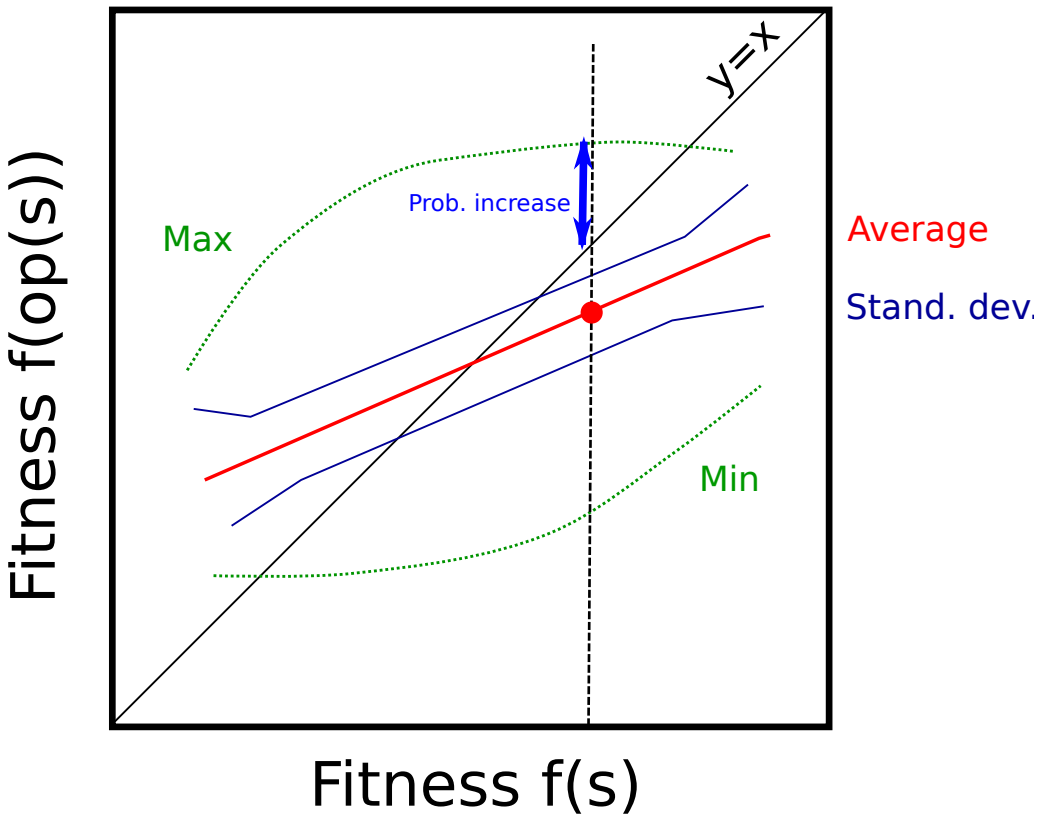
- $(\mathcal{S}, \mathcal{F}, \Pr)$: probability space
- $op : \mathcal{S} \to \mathcal{S}$ stochastic operator of the local search

- $X(s) = f(s)$
- $Y(s) = f(op(s))$

Fitness Cloud of *op*

Conditional probability density function of Y given X

Fitness cloud : Measure of evolvability



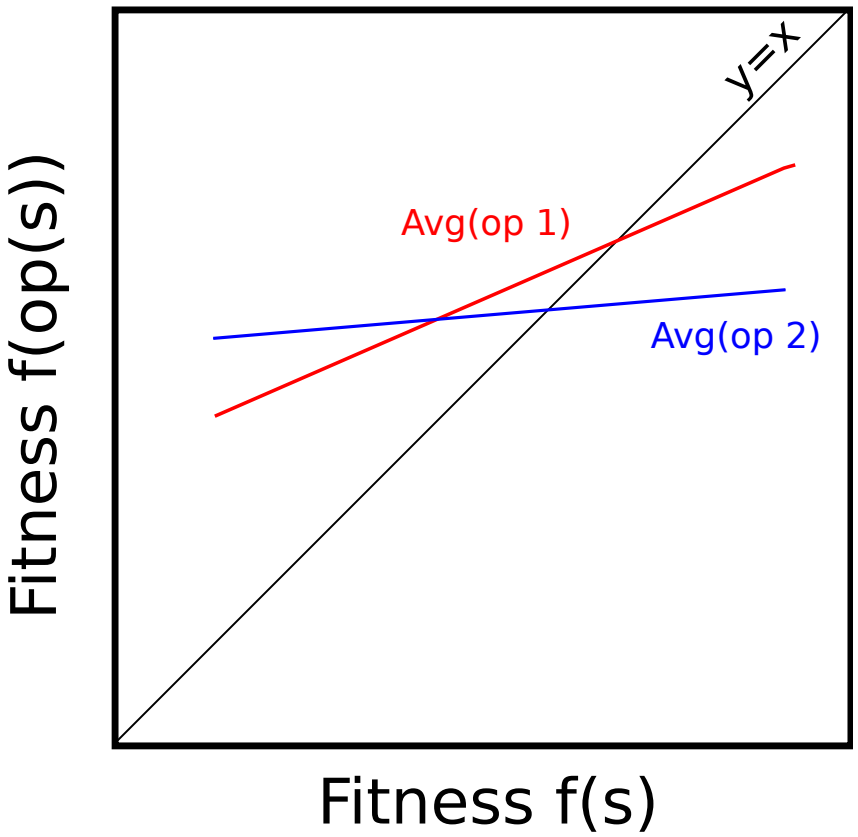
Evolvability

Ability to evolve : fitness in the neighborhood compared to the fitness of the solution

- Probability of finding better solutions
- Average fitness of better neighbor solutions
- Average and standard deviation of fitnesses

Fitness cloud : Comparison of difficulty

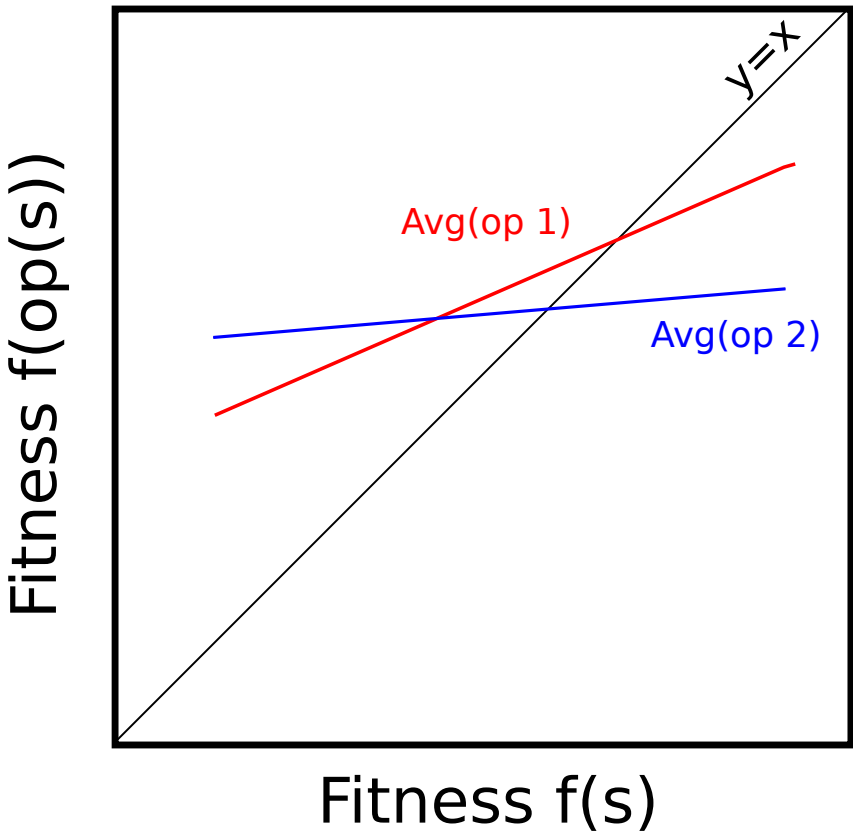
Average of evolvability



- Operator 1 ? ? Operator 2

Fitness cloud : Comparison of difficulty

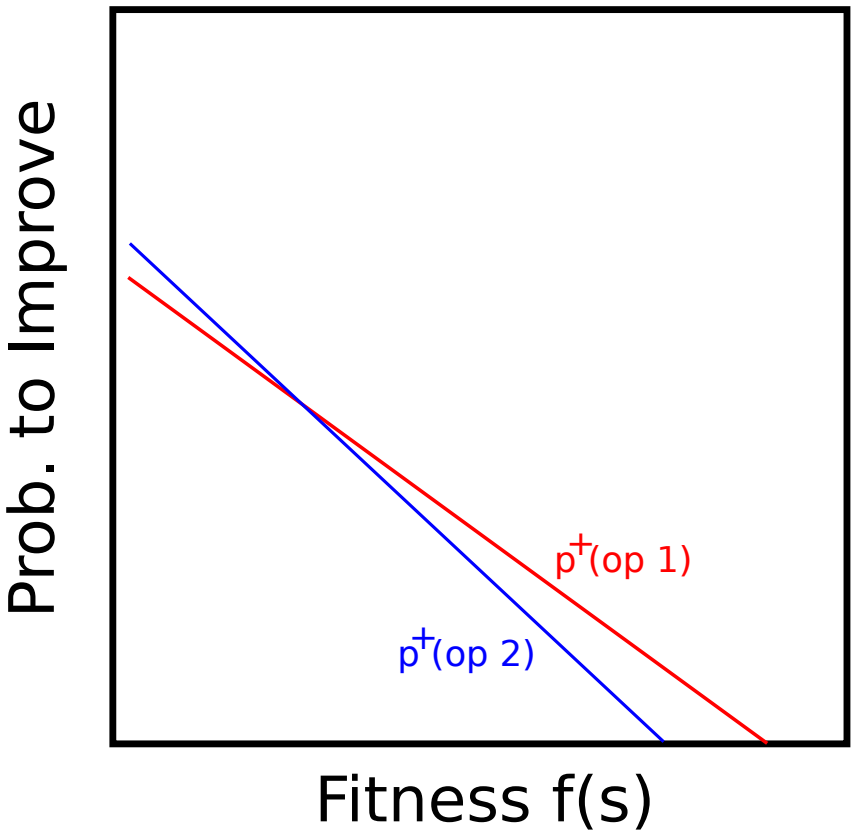
Average of evolvability



- Operator 1 > Operator 2
- Because Average 1 more correlated to fitness
- Linked to autocorrelation
- Average is often a line :
  - See works on Elementary Landscapes (Stadler, D. Wihtley, F. Chicano and others)
  - See the idea of Negative Slope Coefficient (NSC)

Fitness cloud : Comparison of difficulty

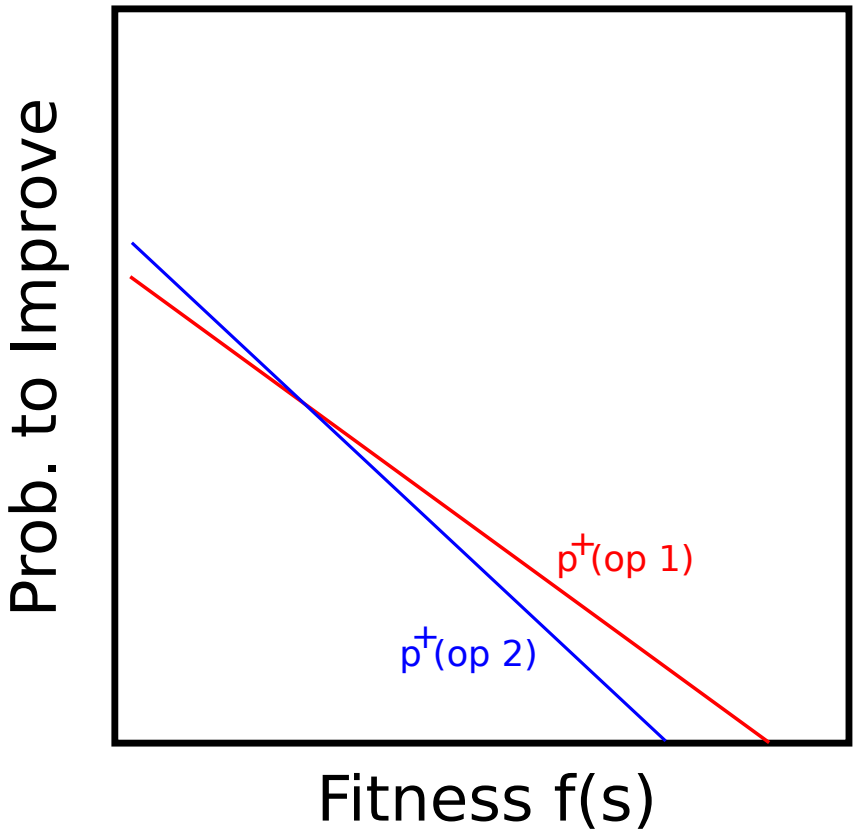
Probability to improve



- Operator 1 ? ? Operator 2

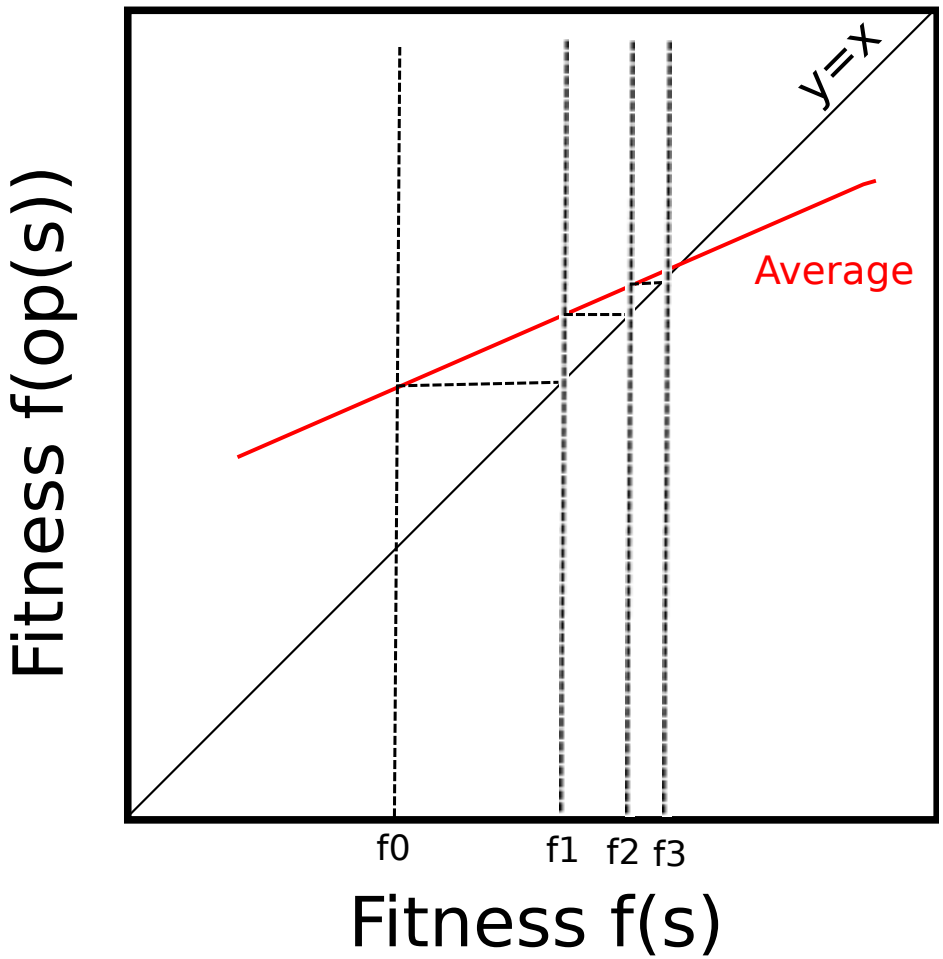
Fitness cloud : Comparison of difficulty

Probability to improve



- Operator 1 > Operator 2
- Prob. to improve of 1 is often higher than Prob. to improve of 2
- Probability to improve is often a line
- See also works on fitness-probability cloud (G. Lu, J. Li, X. Yao)
- See theory of EA and fitness level technics.

Fitness cloud : estimation of convergence point

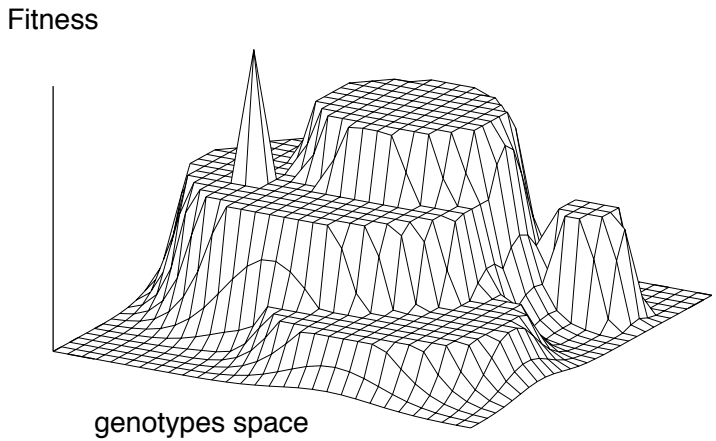


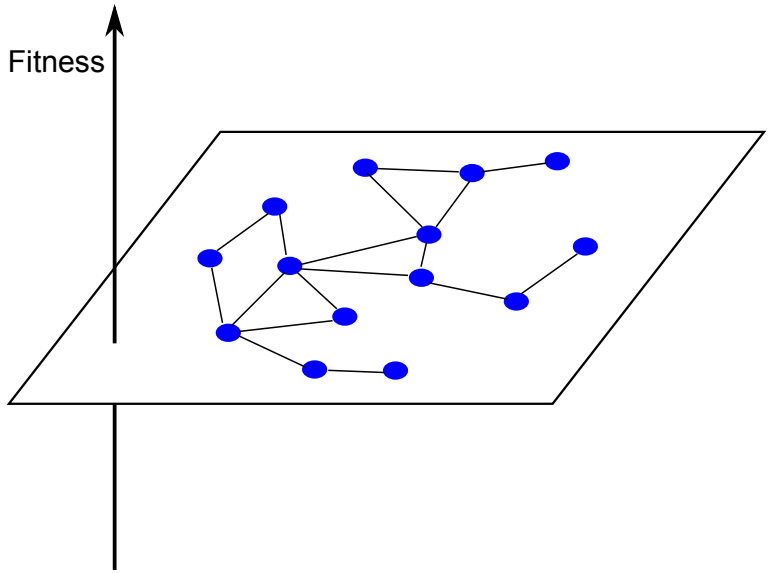
- Approximation (only approximation) of the fitness value after few steps of local operator
- Indication on the quality of the operator
- See fitness level technic

Outline

- Neutral sets (done) :
  set of solutions with the same fitness
  $\Rightarrow$ No structure
- Fitness cloud (done) :
  Bivariate density $(f(s), f(op(s)))$
  $\Rightarrow$ Neighborhood relation **between** neutral sets
- Neutral networks (to be done) :
  $\Rightarrow$ neutral sets with neighborhood relation : graph

Neutral networks (Schuster 1994 [27])





Basic definition of Neutral Network

Node = solution with same fitness value

Edge = neighborhood relation

Definitions

Test of neutrality

$$isNeutral : S \times S \rightarrow \{true, false\}$$

For example, $isNeutral(s_1, s_2)$ is *true* if :

- $f(s_1) = f(s_2)$.
- $|f(s_1) - f(s_2)| \leqslant 1/M$ with M is the search population size.
- $|f(s_1) - f(s_2)|$ is under the evaluation error.

Neutral neighborhood

of s is the set of neighbors which have the same fitness $f(s)$

$$\mathcal{N}_{neut}(s) = \{s^{'} \in \mathcal{N}(s) \mid isNeutral(s, s^{'})\}$$

Neutral degree of s

Number of neutral neighbors : $nDeg(s) = \sharp(\mathcal{N}_{neut}(s) - \{s\})$.

Definitions

Neutral walk

$W_{neut} = (s_0, s_1, \ldots, s_m)$

- for all $i \in [0, m-1]$, $s_{i+1} \in \mathcal{N}(s_i)$
- for all $(i, j) \in [0, m]^2$, $isNeutral(s_i, s_j)$ is *true*.

Neutral Network

graph $G = (N, E)$

- $N \subset \mathcal{S}$: for all s and s' from N, there is a neutral walk belonging to N from s to s',
- $(s_1, s_2) \in E$ if they are neutral neighbors : $s_2 \in \mathcal{N}_{neut}(s_1)$

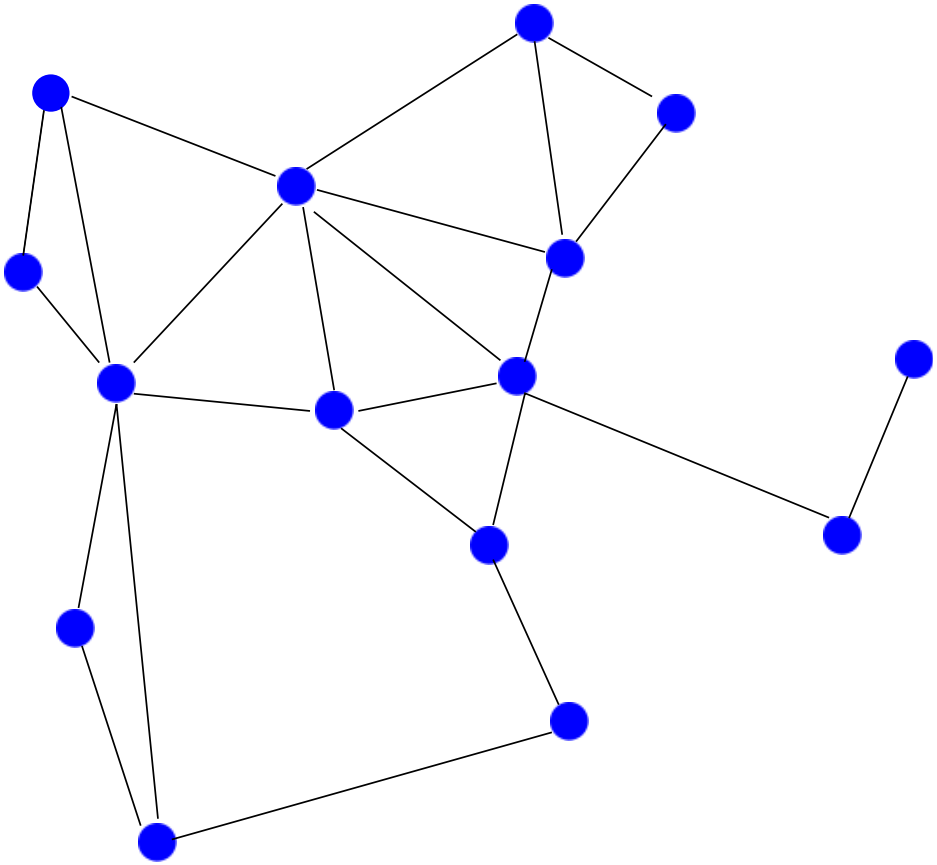
A fitness landscape is neutral
if there are many solutions with high neutral degree.

Features inside neutral network

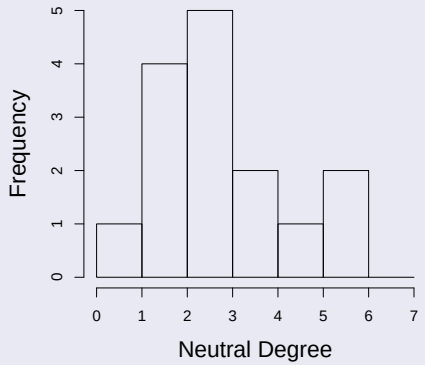
Classical graph metrics :

1. **Size of NN** :
   - number of nodes of NN,
2. **Neutral degree distribution** :
   - measure of the quantity of "neutrality"
3. **Autocorrelation of neutral degree** (Bastolla 03 [3]) during neutral random walk :
   - comparaison with random graph,
   - measure of the correlation structure of *NN*

Features inside neutral network

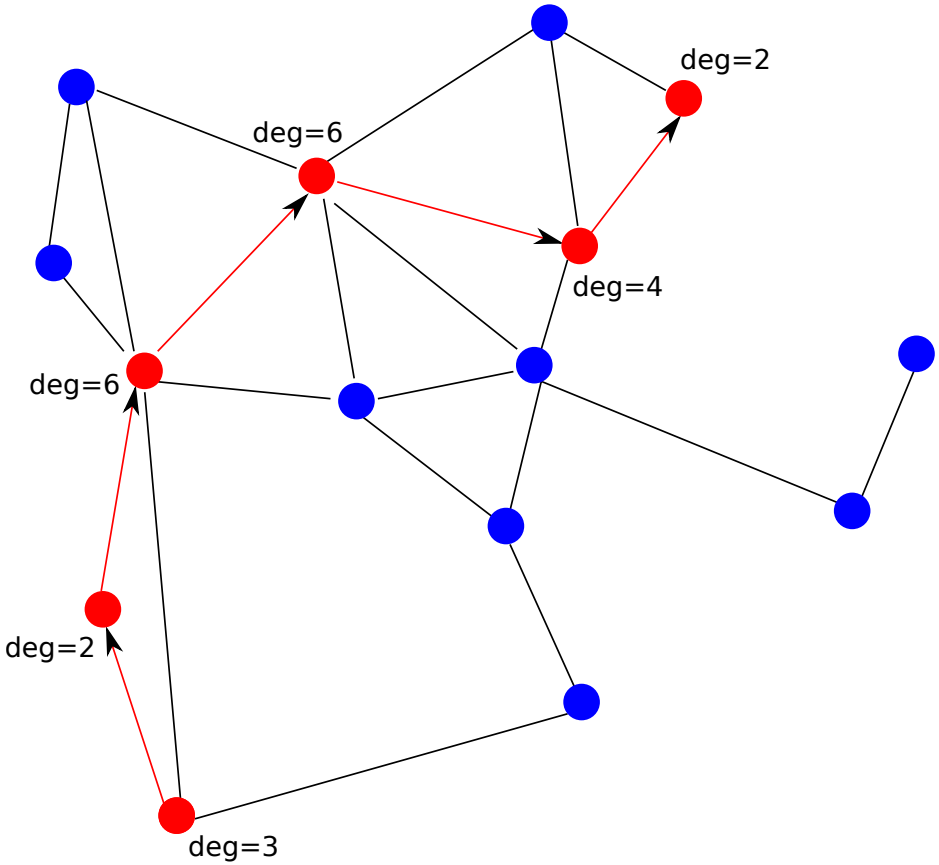


1. **Size**
   avg,distribution,etc.
2. **Neutral degree distribution**

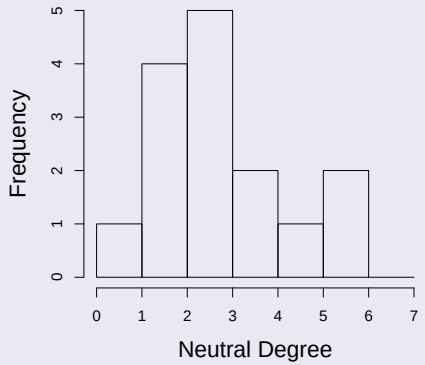


3. **Autocorrelation of neutral degree**
   - random walk on NN
   - autocorr. of degrees

Features inside neutral network

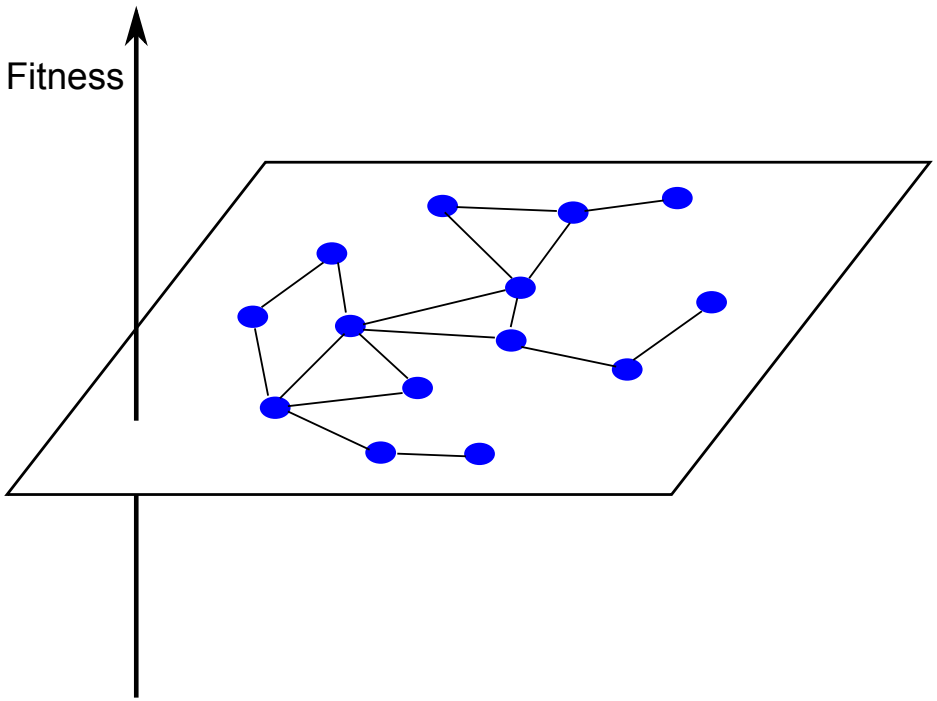


1. **Size**
   avg,distribution,etc.
2. **Neutral degree distribution**



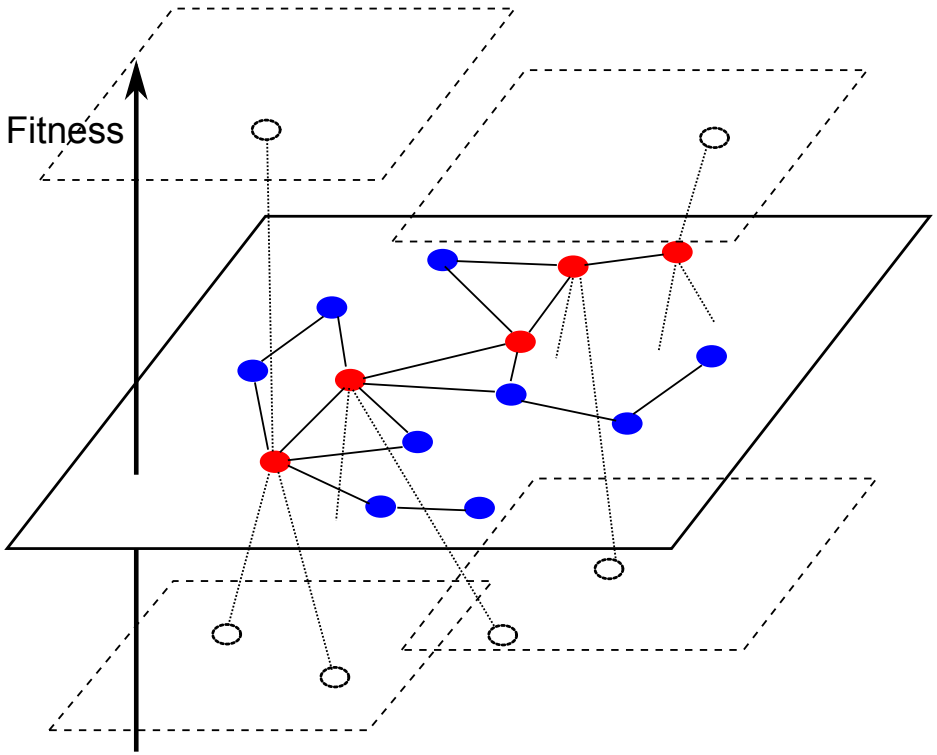
3. **Autocorrelation of neutral degree**
   - random walk on NN
   - autocorr. of degrees

Features between neutral networks



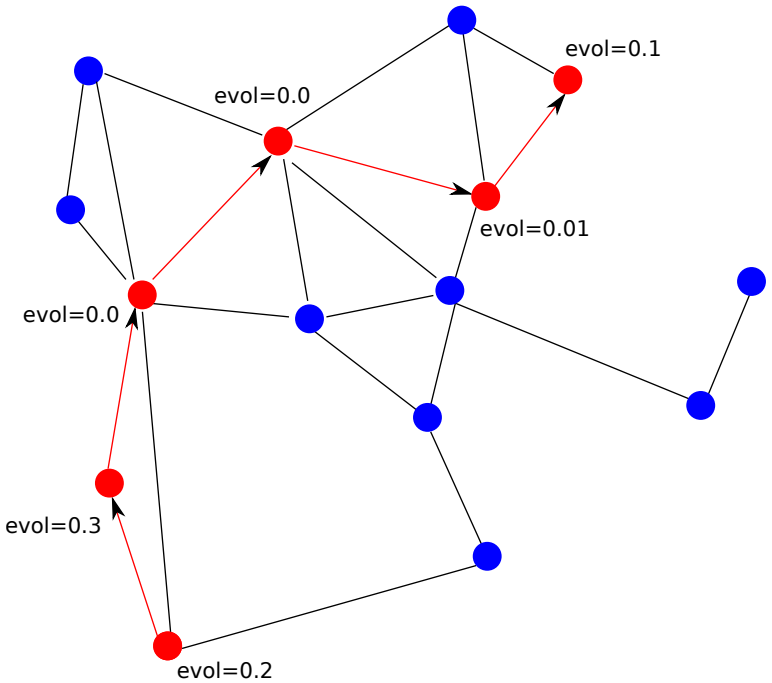
1. **Rate of innovation** (Huynen 96 [13]) : the number of new accessible structures (fitness) per mutation
2. **Autocorrelation of evolvability** [34] : autocorrelation of the sequence $(evol(s_0), evol(s_1), \ldots)$.

Features between neutral networks



1. **Rate of innovation** (Huynen 96 [13]) : the number of new accessible structures (fitness) per mutation
2. **Autocorrelation of evolvability** [34] : autocorrelation of the sequence $(evol(s_0), evol(s_1), \ldots)$.

Features between neutral networks



- Autocorrelation of evolvability :
  - Autocorrelation of $(evol(s_0), evol(s_1), \ldots)$.
  - Evolvability *evol* :
    - average fitness in the neigh.
    - prob. to improve, etc.
- Informations :
  - if high correlation $\Rightarrow$ "easy" (you can use this information)
  - if low correlation $\Rightarrow$ "difficult"

Summary of neutral fitness landscapes features

- Density of States :
  Size of neutral sets
- Fitness cloud and related statistics :
  Evolvability of solutions
- Size and Neutral degrees distribution :
  "How neutral is the fitness landscape ?"
- Autocorrelation of neutral degrees :
  Network "structure"
- Rate of innovation :
  Low information for combinatorial optimization
- Autocorrelation of evolvability :
  Evolution of evolvability on NN

Fitness landscape for Continuous optimization

Do we really need fitness landscape analysis technics for continuous optimization ?

Fitness landscape for Continuous optimization

Do we really need fitness landscape analysis technics for continuous optimization ?

Standard mathematical tools exists :

- gradient, local optima, curvature, etc.

But the same question remains :

- Relation with local search algorithms,
- Black-box context

"Exploratory Landscape Analysis" works

TU Dortmund group : Bartz-Beielstein, Mersmann, Preuss, Trautmann, etc. / Abell, Malitsky, Tierney [gecco 11,12, ppsn 11, lion 12] / [lion 13]

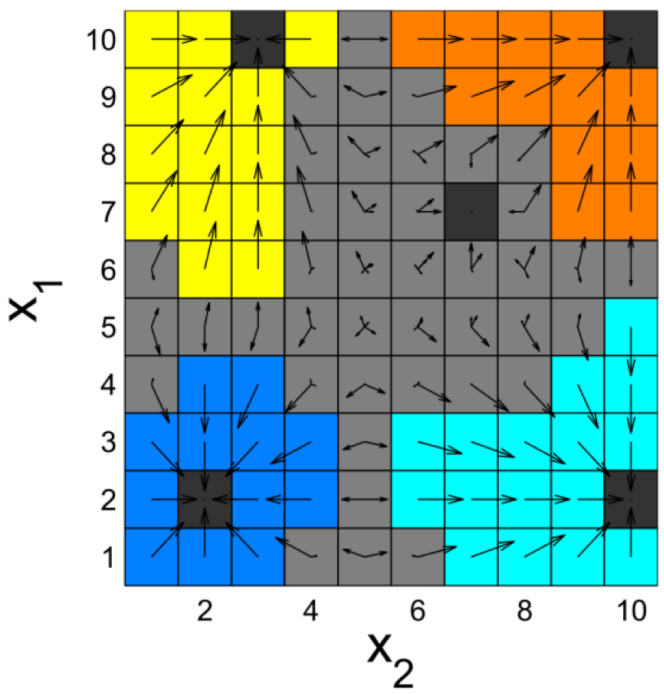
6 groups of features

- **y-distribution**
  skewness, and curtosis of the fitness distribution ($\approx$ DOS)
- **level set**
  Split the data according to fitness. Features is the errors of different classifiers.
- **metamodel**
  R2 of linear and quadratic regressions with and without interactions
- **convexity**
  $f(w\,x + (1-w)y)$
  $-[w\,f(x) + (1-w)f(y)]$
- **local search**
  Statistics based on the clustering of local optima found by Nelder-Mead
- **curvature**
  Statistics based on the gradient estimation

! search space sampling : latin hypercube DOE

Cell Mapping Techniques

for Exploratory Landscape Analysis [Kerschke et al.14]



Principle :

- Split the search space into cells.
  from nonlinear dynamical systems analysis [1987]

Metrics :

- Basins of attraction
- Inside cells : gradient homogeneity and convexity
- cell-to-cell : related to gradient/selection

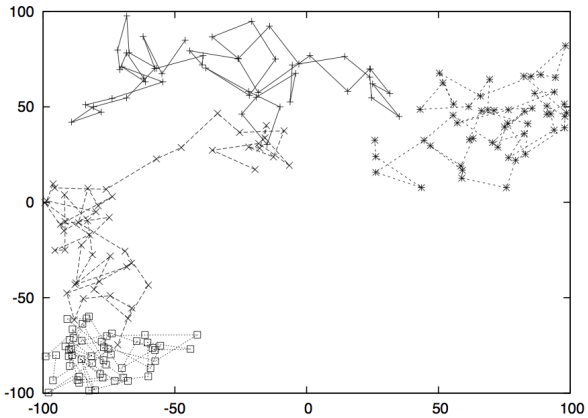
Searchability of continuous optimization problems

Malan and Engelbrecht, Swarm Intelligence, 2014.

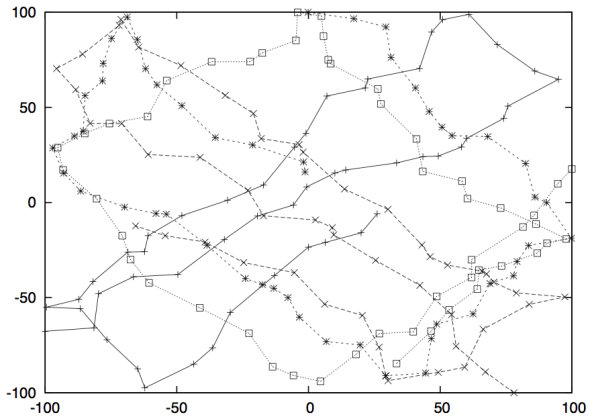
Definition of searchability

"The evolvability of a fitness landscape is defined as the ability of a given search process to move to a place in the landscape of better fitness and is henceforth referred to as searchability."

Sampling of the search space :

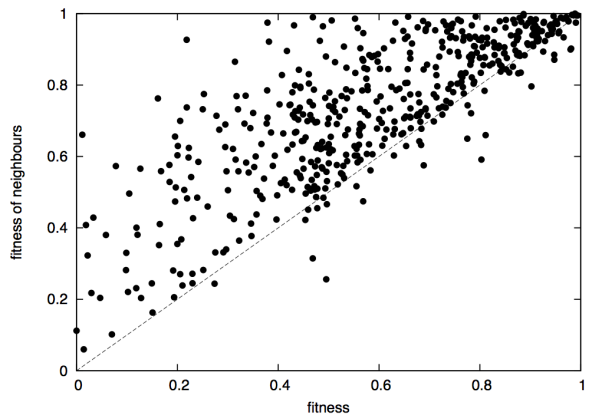


uniform random walk

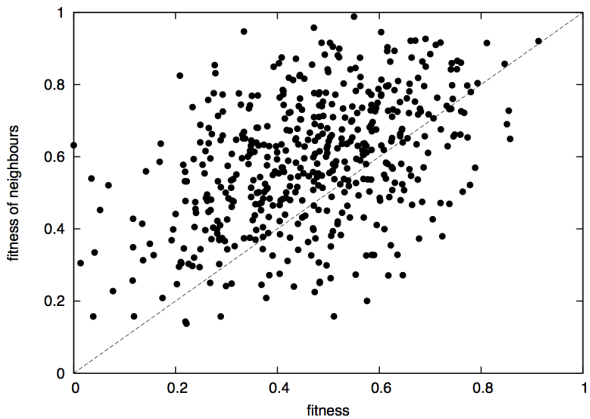


progressive random walk

Features customized for continuous optimization



(a) $2D$ Spherical



(b) $2D$ Rastrigin

- **FDC**
  best point of the sample
  and euclidian distance
- **Information landscape**
  matrix of improvement (0, 0.5, or 1)
  without neighborhood
  distance to a ref. landscape (sphere)
- **Fitness cloud**
  neighborhood related to PSO
  2 steps :
  -first step : mutation with gaussian,
  -second step : PSO "formula".

Fitness landscape for Multi-Objective optimization

Context

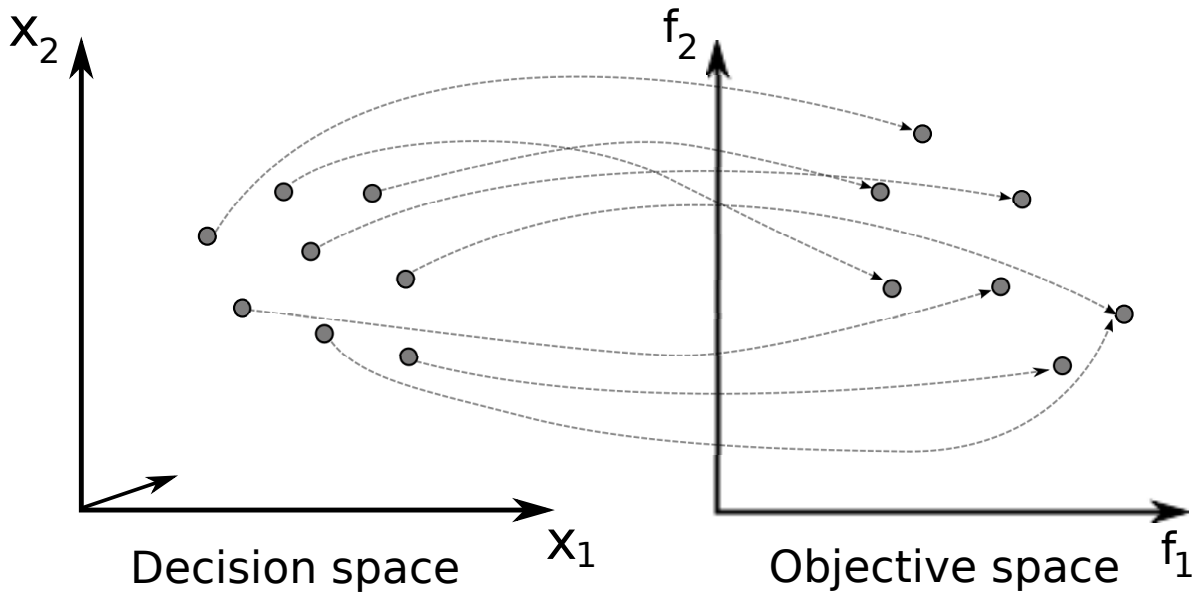
- MO problems are hard :
  underlying single-objective functions, join functions, etc.
- Learning the problem structure :
  to understand and improve algorithms
- Questions :
  - What are the relevant problem features ?
  - Replace one-dimensional by d-dimensional objective function ?

Join work with :
Arnaud Liefooghe, Clarisse Daehnens, Laetitia Jourdan, Kiyoshi Tanaka, Hernan Aguirre, Fabio Daolio, Manuel López-Ibáñez, Mathieu Basseur, Adrien Goëffon

Multiobjective optimization

Multiobjective optimization problem

- X : set of feasible solutions in the solution space
- An objective function vector $f = (f_1, f_2, \ldots, f_m)$ (max), $m \geqslant 2$
- $Z = f(X) \subseteq \mathbb{R}^m$: set of feasible vectors in the objective space



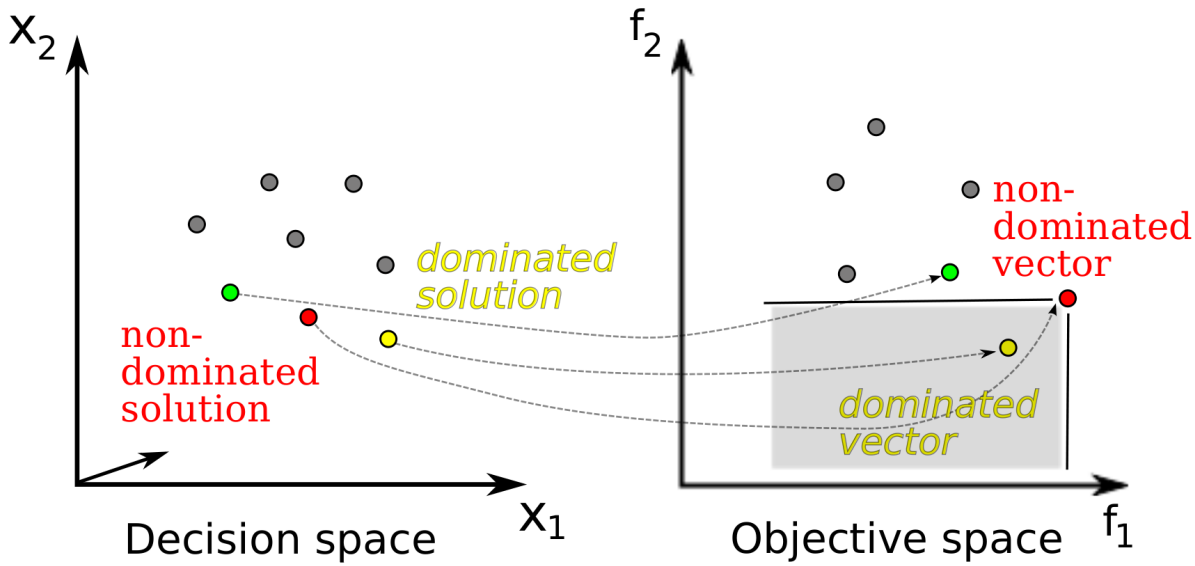
Multiobjective optimization

Pareto dominance

Pareto dominance relation

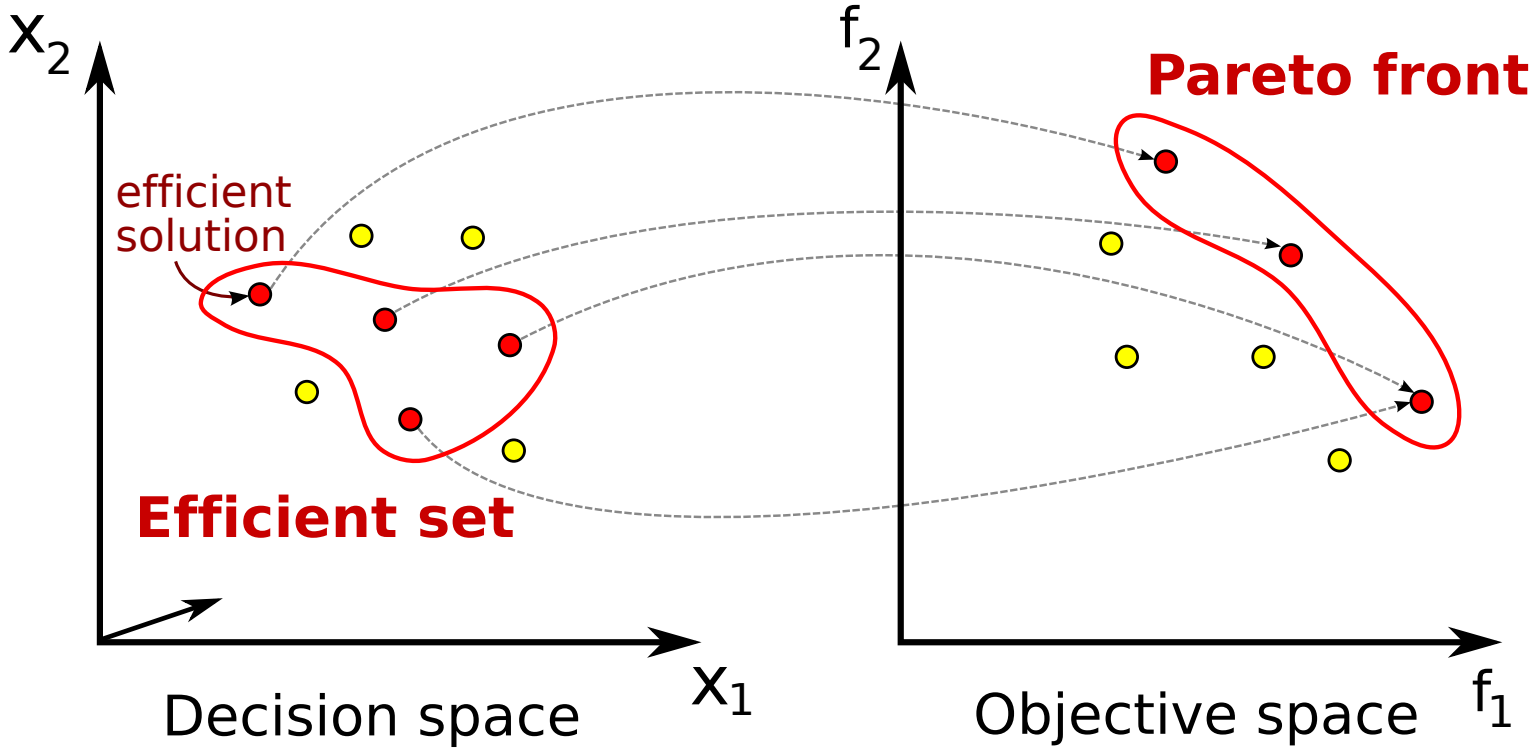
A solution $x \in X$ dominates a solution $x' \in X$ ($x' \prec x$) iff

- $\forall i \in \{1, \ldots, m\}$, $f_i(x) \geqslant f_i(x')$
- $\exists j \in \{1, \ldots, m\}$ such that $f_j(x) > f_j(x')$



Multiobjective optimization

Pareto set *vs.* Pareto front



High-level features from the Pareto set

npo **Number of Pareto optimal solutions**

Cardinality of the (exact) Pareto set ($|X^\star|$)

hv **Hypervolume indicator** [Zitzler et al., 1999]

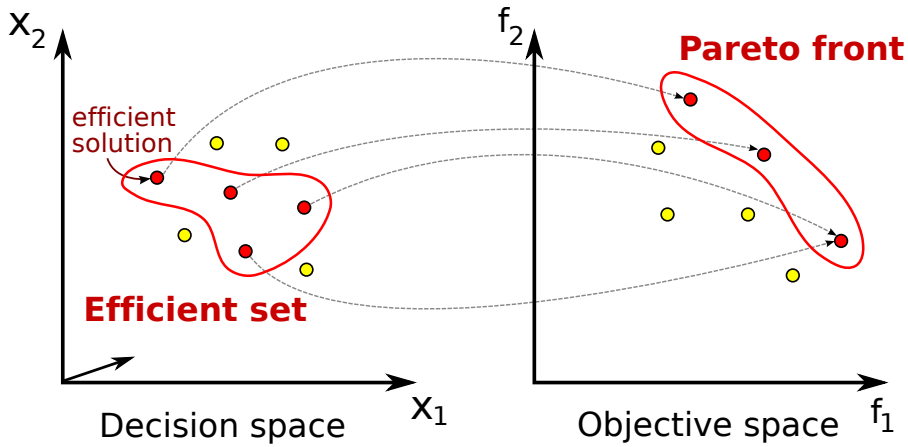
Objective space dominated by $X^\star$, ref. point $= (0.0, \ldots, 0.0)$

avgd **Average distance between Pareto solutions**

Average Hamming distance between any pair of Pareto sol.

maxd **Maximum distance between Pareto solutions**

Maximum Hamming distance between two Pareto solutions



High-level features from the Pareto graph

Connectedness [Ehrgott & Klamroth 1997], *d*-Pareto graph [Paquete et al. 2008]

`nconnec` **Number of connected components**

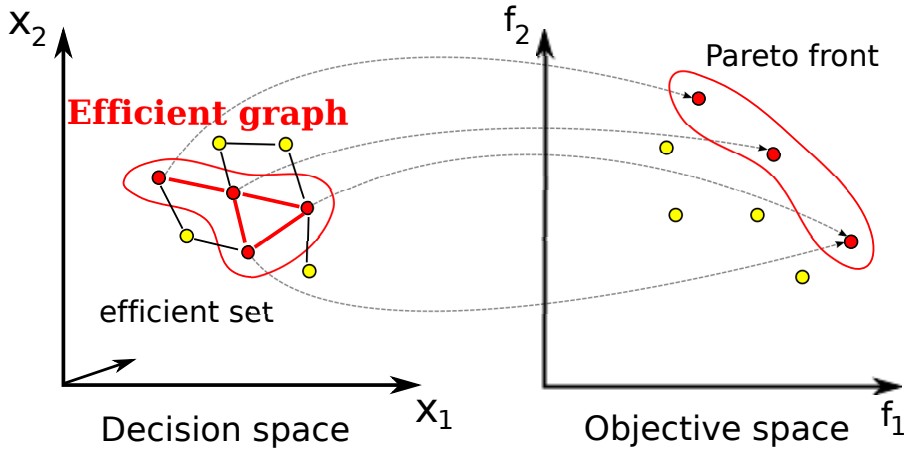
Number of connected components in the 1-Pareto graph (Hamming distance 1)

`lconnec` **Size of the largest connected component**

Size of the largest connected component in the 1-Pareto graph

`kconnec` **Minimum distance to be connected**

Smallest distance *d* such that the *d*-Pareto graph is connected



High-level features from the fitness landscape

`nplo` **Number of Pareto local optima**

Neighborhood $\mathcal{N}$ = 1-bit-flip (Hamming distance 1)

$x \in X$ is a *Pareto local optimum* iff $\nexists x' \in \mathcal{N}(x)$ s.t. $x \prec x'$

`ladapt` **Length of a Pareto-based adaptive walk**

Length of a single solution-based Pareto hill-climber

`corhv` **Correlation of solution hypervolume** (ruggedness)

Autocorrelation of the hypervolume value along a random walk

$\phi : X \rightarrow \mathbb{R} : \phi(x) = \mathtt{hv}(\{x\})$

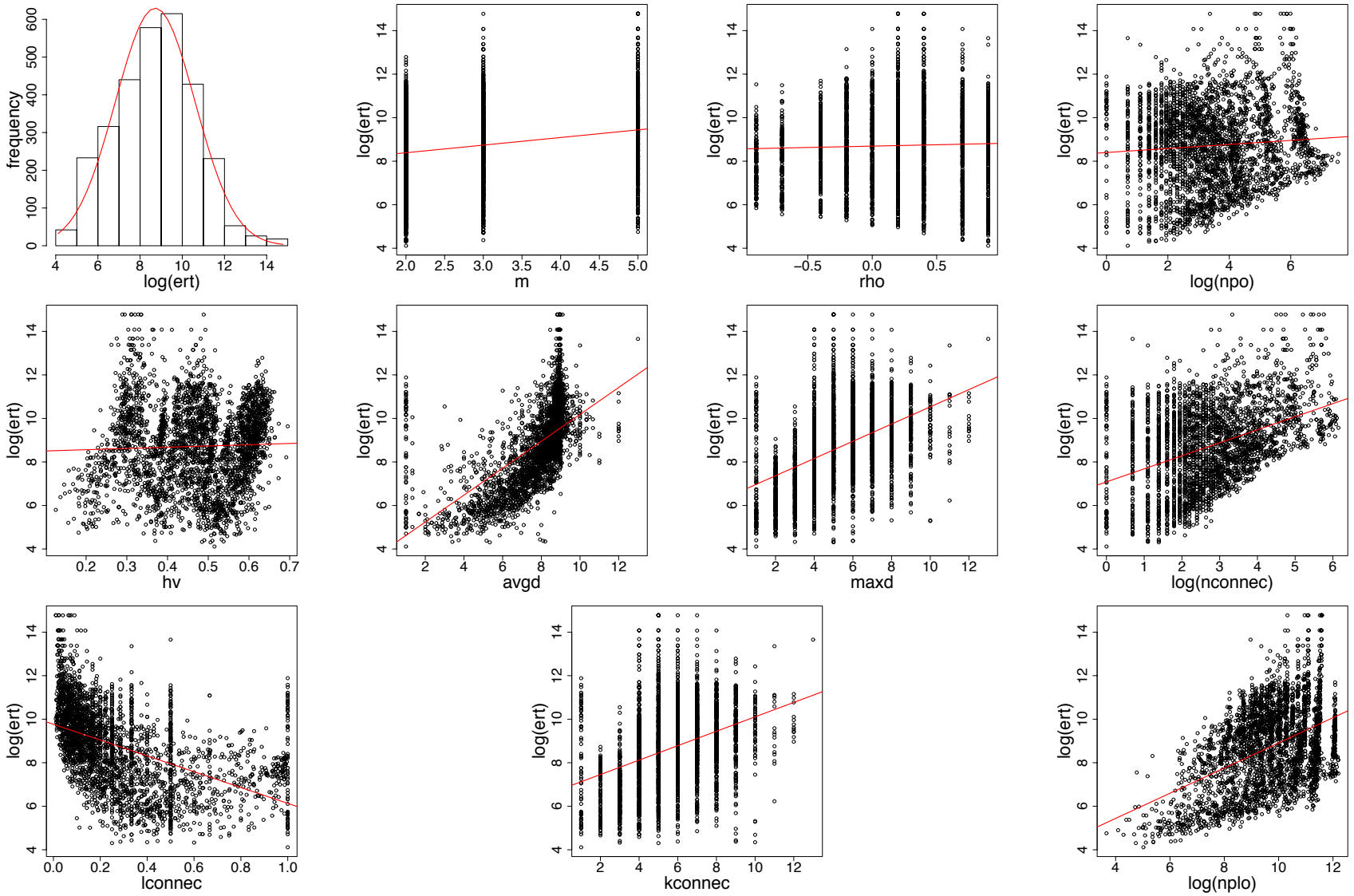
`corlhv` **Correlation of local hypervolume** (ruggedness)

$\phi : X \rightarrow \mathbb{R} : \phi(x) = \mathtt{hv}(\mathcal{N}(x) \cup \{x\})$

Summary of problem features

	low-level features	
m	number of objective functions	known
ρ	objective correlation	known (estimated)
	high-level features	
npo	number of Pareto optimal solutions	enumeration (bounds)
hv	hypervolume	enumeration
avgd	average distance between Pareto sol.	enumeration
maxd	maximum distance between Pareto sol.	enumeration
nconnec	number of connected components	enumeration
lconnec	size of the largest connected component	enumeration
kconnec	minimum distance to be connected	enumeration
nplo	number of Pareto local optima	enumeration (estim.)
ladapt	length of a Pareto-based adaptive walk	estimation
corhv	correlation length of solution hypervolume	estimation
corlhv	correlation length of local hypervolume	estimation

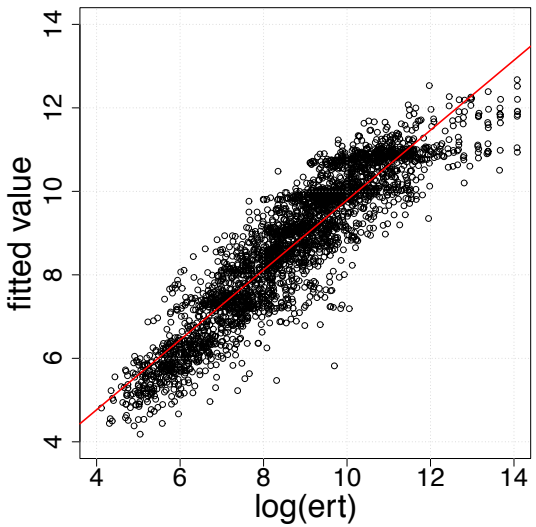
Example of MO performance prediction *vs.* single feature



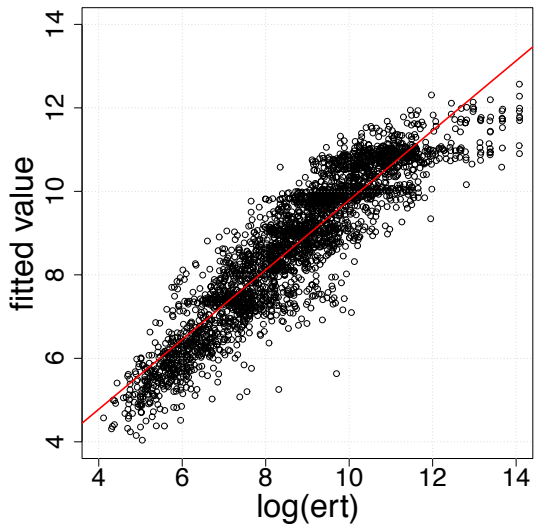
Example of MO performance prediction *vs.* features

Multiple linear regression model(s)

$$\log(\texttt{ert}) = \beta_0 + \beta_1 \cdot k + \beta_2 \cdot m + \beta_3 \cdot \rho + \beta_4 \cdot \texttt{npo} + \ldots + \beta_{14} \cdot \texttt{corlhv} + e$$



all features



$k, m, \log(\texttt{nplo}), \texttt{hv}, \rho$

Set-based multiobjective fitness landscapes

The key idea

EMO algorithms are local search algorithms performing on sets

Set-based multiobjective fitness landscape

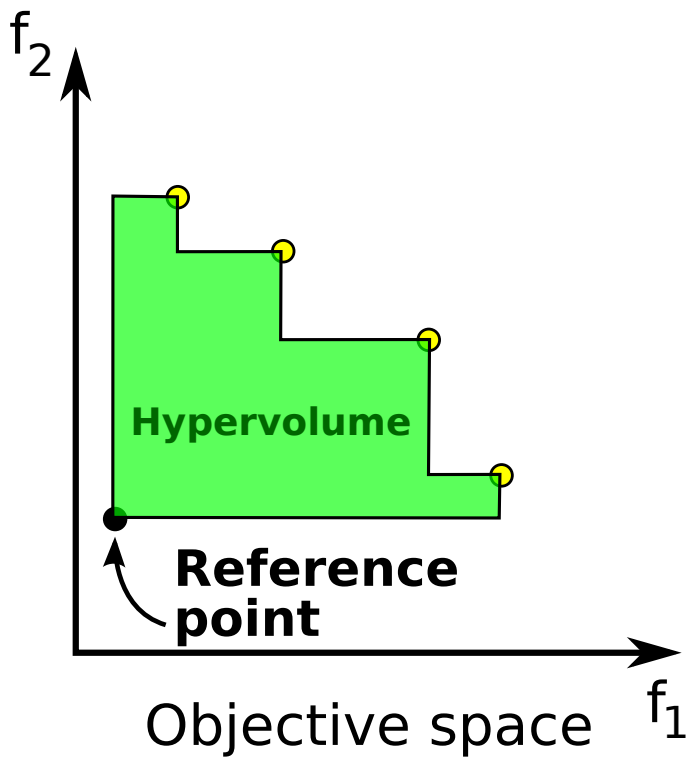
$$(\Sigma, \mathtt{N}, \mathtt{I})$$

- **Set-domain search space** (Σ)
  $\Sigma \subset 2^X$ is a set of feasible solution-sets
  (where X is the set of feasible solutions)
- **Set-domain neighborhood relation** $(\mathtt{N})$
  $\mathtt{N} : \Sigma \rightarrow 2^{\Sigma}$ is a neighborhood relation between solution-sets
- **Set-domain fitness function** $(\mathtt{I})$
  $\mathtt{I} : \Sigma \rightarrow \mathbb{R}$ is a unary quality indicator,
  i.e. a fitness function measuring the quality of solution-sets

Set-domain fitness function

Quality indicator

$\mathtt{I}(\sigma) \in \mathbb{R}$: quality of a set of solutions σ



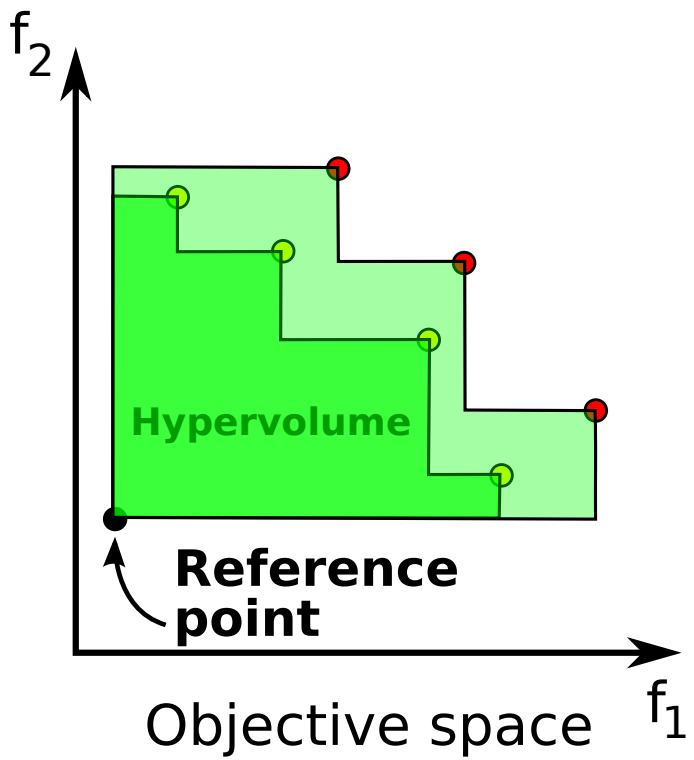
Hypervolume ($\mathtt{I}_H$)

- Compliant with the (weak) Pareto dominance relation :
  $\sigma \prec \sigma' \Rightarrow \mathtt{I}_H(\sigma) \leqslant \mathtt{I}_H(\sigma')$
- A single parameter :
  the reference point
- Minimal set maximizing $\mathtt{I}_H$:
  subset of the Pareto optimal set

Set-domain fitness function

Quality indicator

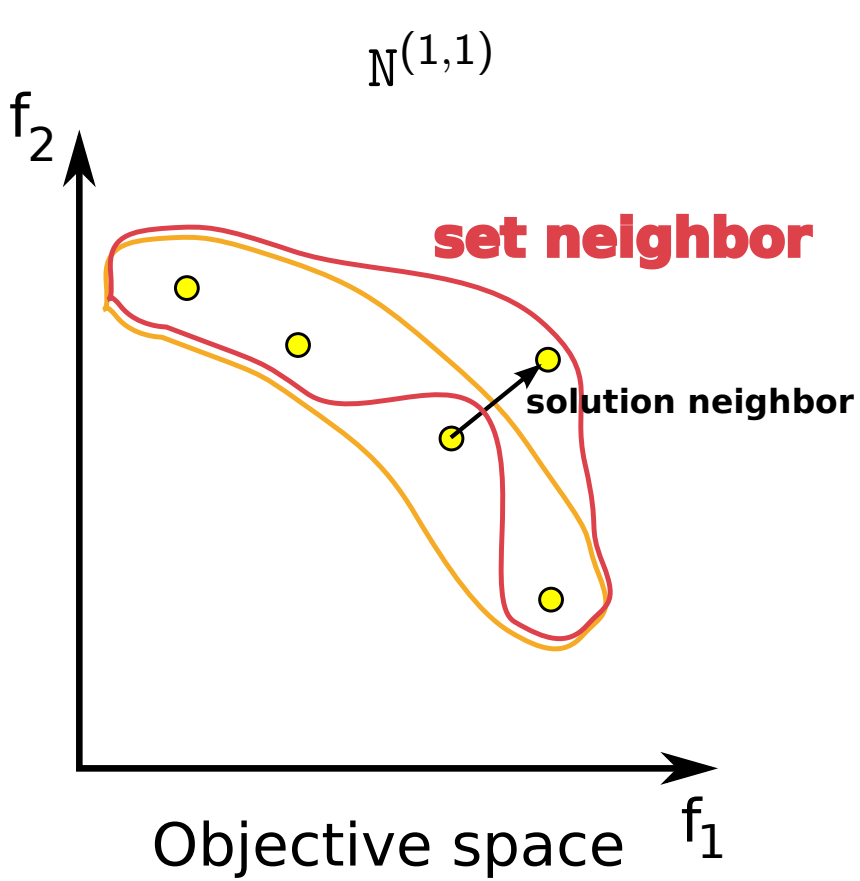
$\mathtt{I}(\sigma) \in \mathbb{R}$: quality of a set of solutions σ

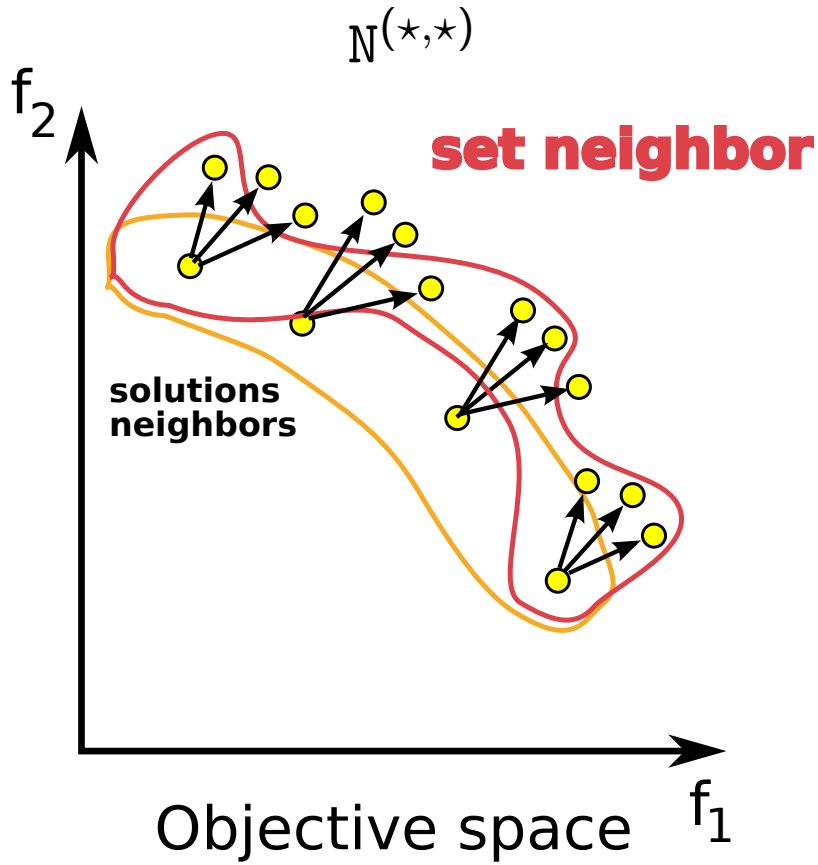


Hypervolume ($\mathtt{I}_H$)

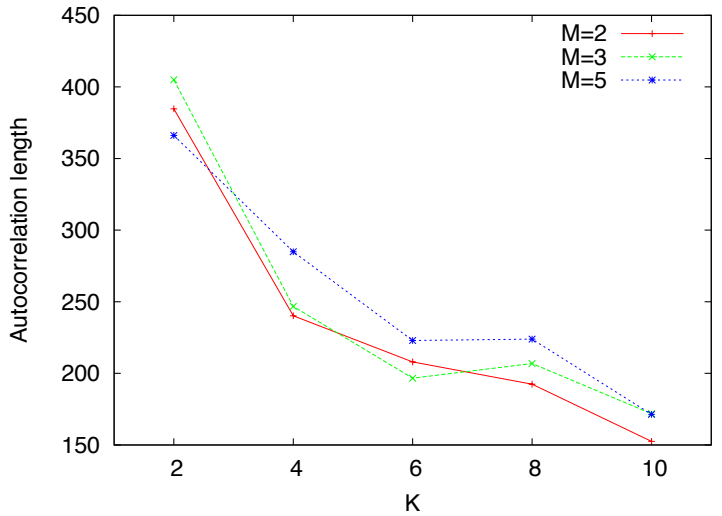
- Compliant with the (weak) Pareto dominance relation :
  $\sigma \prec \sigma' \Rightarrow \mathtt{I}_H(\sigma) \leqslant \mathtt{I}_H(\sigma')$
- A single parameter :
  the reference point
- Minimal set maximizing $\mathtt{I}_H$:
  subset of the Pareto optimal set

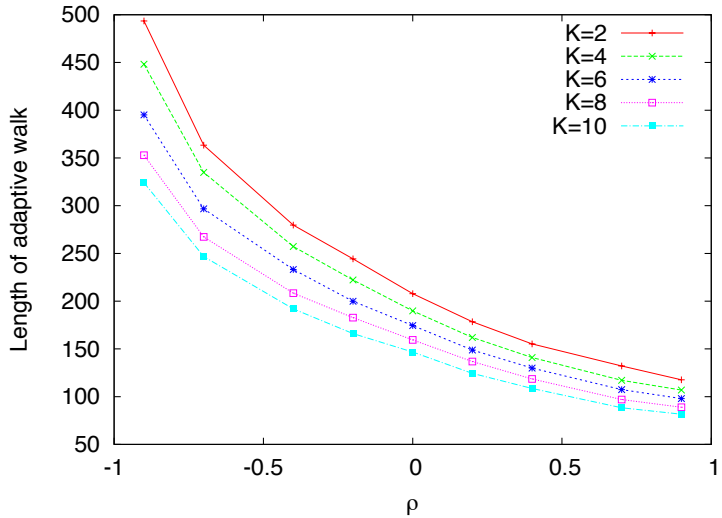
Set-domain neighborhood relations





Ruggedness and multimodality





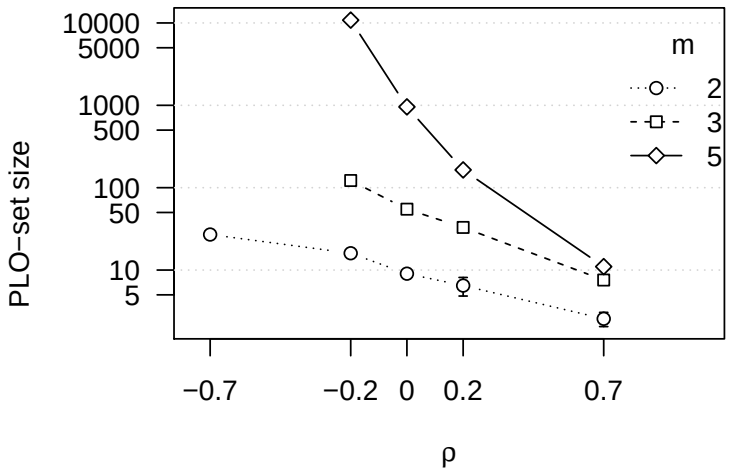
Autocorrelation length (τ)

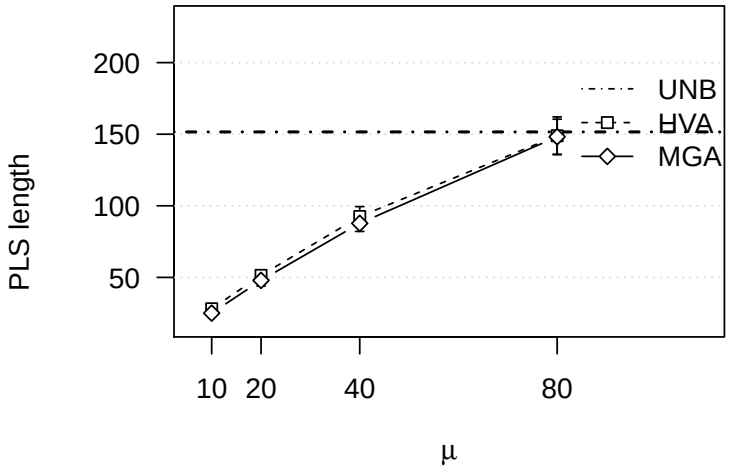
- non-linearity degree K
  $\rightarrow$ τ decreases wrt. K
- number of objectives M
  $\rightarrow$ no influence
- objective correlation ρ
  $\rightarrow$ no influence

Adaptive walk length ($\mathcal{L}$)

- non-linearity degree K
  $\rightarrow$ $\mathcal{L}$ decreases wrt. K
- number of objectives M
  $\rightarrow$ no clear trend
- objective correlation ρ
  $\rightarrow$ $\mathcal{L}$ decreases wrt. ρ!!

Pareto Local Optima set *vs.* archive size





Pareto set-size

for unbounded solution-set (archive)

- objective correlation ρ
  $\rightarrow$ decreases wrt. ρ
- number of objectives M
  $\rightarrow$ increases wrt. M

Pareto Local search Length

for bounded solution-set (archive)

- Increases with archive size

Fitness landscapes analysis for MO optimization

Summary

- A number of features are shown to be related to problem difficulty
- Better knowledge on problem difficulty for MO optimization
- Performance can be predicted using such relevant features (for some algo/problems)
- And now,
  - toward a better design ?
  - toward a algorithm portfolio ?
  - toward more theoretical works ?
  - design low complexity features ?...

Basic Methodology of fitness landscapes analysis

- Density of States : pure random search, initialization ?
- Length of adaptive walks : multimodality ?
- Autocorrelation of fitness : ruggedness ?
- Neutral Degree Distribution : neutrality ?
- Fitness Cloud : Quality of the operator, evolvability ?
- Fitness Distance Correlation from best known
- Neutral walks and evolvability : neutral information ?

Recent review : Katherine M. Malan, Information Sciences, (2013)

Basic Methodology of fitness landscapes analysis

- Density of States : pure random search, initialization ?
- Length of adaptive walks : multimodality ?
- Autocorrelation of fitness : ruggedness ?
- Neutral Degree Distribution : neutrality ?
- Fitness Cloud : Quality of the operator, evolvability ?
- Fitness Distance Correlation from best known
- Neutral walks and evolvability : neutral information ?
- ... be creative from your algorithm and problem point of view
- ... be careful on the computed measures : one measure is not enough, and must be very well understand

Recent review : Katherine M. Malan, Information Sciences, (2013)

Sofware to perform fitness landscape analysis

Framework **ParadisEO**

http://paradiseo.gforge.inria.fr

Software Framework for Metaheuristics (local search, EA, continuous, discrete, parallel, island, fitness landscape, etc.)
See documentation, tutorials

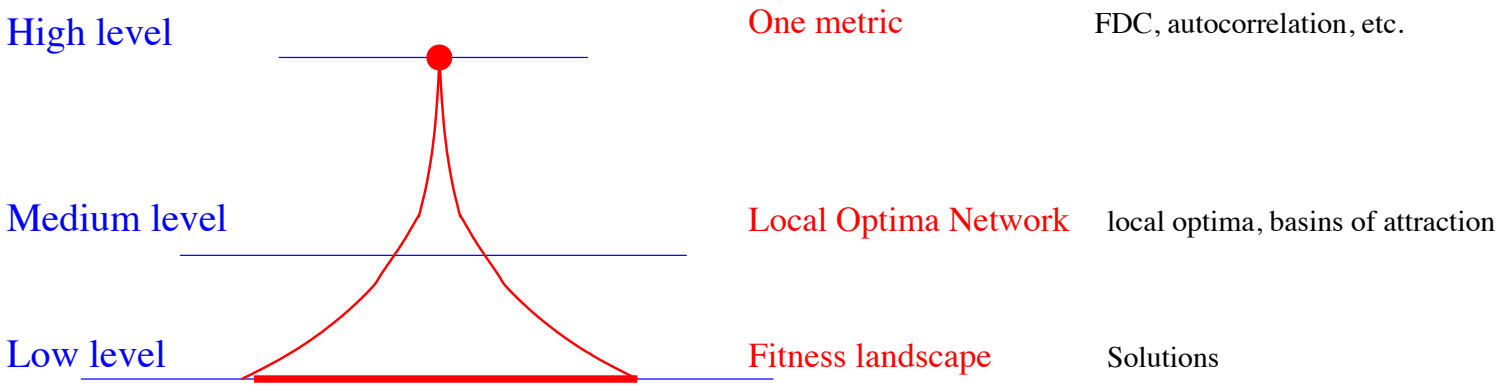
```
moAutocorrelationSampling<Neighbor> sampling(randomInitialization,
                                             neighborhood,
                                             evalFunction,
                                             neighborEvaluation,
                                             nbStep);

sampling();

sampling.fileExport(str_out);
```

Beyond : Local Optima Networks

Join work

Gabriela Ochoa, University of Stirling, Scotland,
Marco Tomassini, University of Lausanne, Switzerland,
Fabio Daolio, university of Shinshu, Japan,

Motivation and general idea : Levels of description



- **Fitness landscapes** :
  based on an huge number of solutions
- **One metric** :
  based on one real number
- **Local optima Network** :
  based on local optima

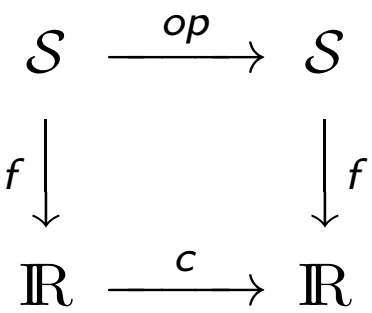
Motivations

$$\mathcal{S} \xrightarrow{op} \mathcal{S}$$

Principle of variables aggregation

- Solutions level :
  - Stochastic local search operator,
  - Exponential number of solutions,
  - Exponential size of the stochastic matrix of the process (Markov chain)

Motivations

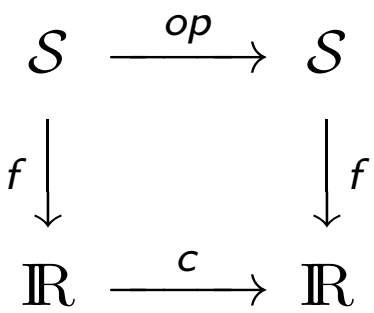


Principle of variables aggregation

- Solutions level :
  - Stochastic local search operator,
  - Exponential number of solutions,
  - Exponential size of the stochastic matrix of the process (Markov chain)
- Projection into the fitness space :
  - Reduce the complexity of Markov chain
  - Potentially loose some information
  - Relevant information remains when :

$$f(op^n(x)) \approx c^n(f(x))$$

Motivations



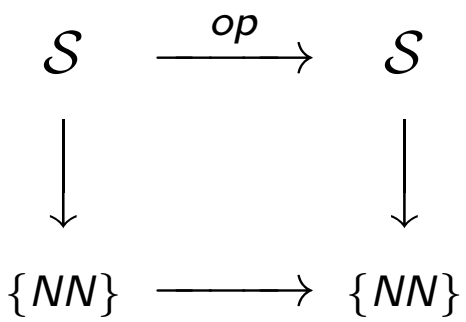
fitness cloud

Principle of variables aggregation

- Solutions level :
  - Stochastic local search operator,
  - Exponential number of solutions,
  - Exponential size of the stochastic matrix of the process (Markov chain)
- Projection into the fitness space :
  - Reduce the complexity of Markov chain
  - Potentially loose some information
  - Relevant information remains when :

$$f(op^n(x)) \approx c^n(f(x))$$

Motivations



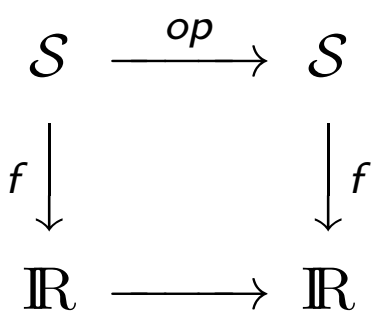
Neutral networks perspective

Principle of variables aggregation

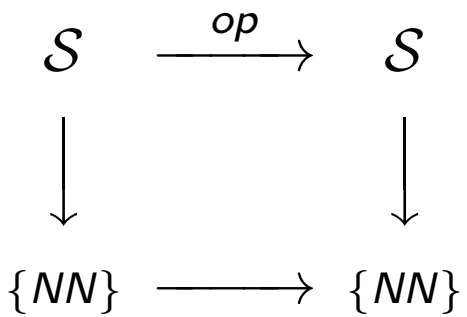
- Solutions level :
  - Stochastic local search operator,
  - Exponential number of solutions,
  - Exponential size of the stochastic matrix (Markov chain)
- Projection into the set of NN :
  - Reduce the complexity of Markov chain
  - Potentially loose some information
  - Relevant information remains

Motivations

aggregation of variables, commutative diagram



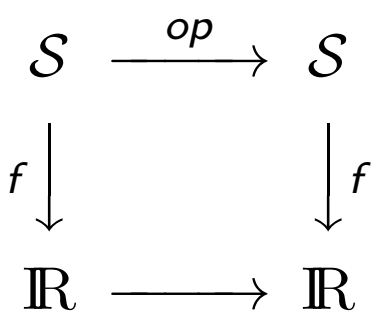
Fitness Cloud



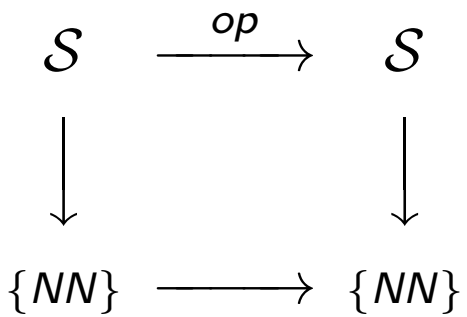
Neutral Networks

Motivations

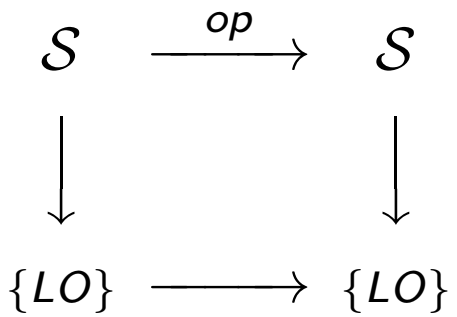
aggregation of variables, commutative diagram



Fitness Cloud



Neutral Networks



Local Optima Networks

Networks and complex systems

Key idea

Bring the tools of *complex networks* analysis to the study the structure of combinatorial fitness landscapes

Methodology

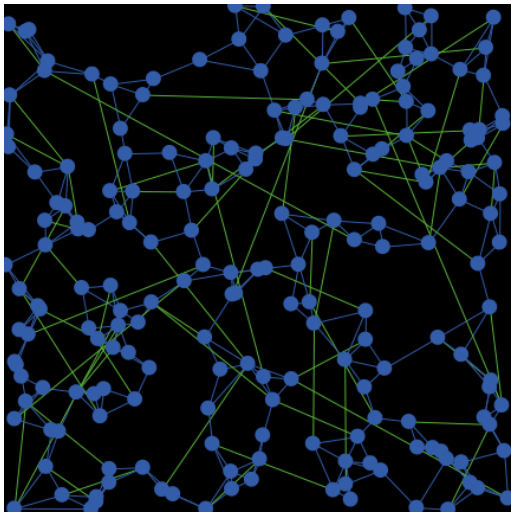
- **Design a network** that represents the landscape (Doye, 2002)
  - Vertices : local optima
  - Edges : a notion of adjacency between local optima
- **Extract features** :
  "complex" network analysis
- **Use the network features** :
  search algorithm design, difficulty, etc.

J. P. K. Doye, The network topology of a potential energy landscape : a static scale-free network., *Phys. Rev. Lett.*, 88 :238701, 2002.

Small-world networks

(Watts and Strogatz, 1998)

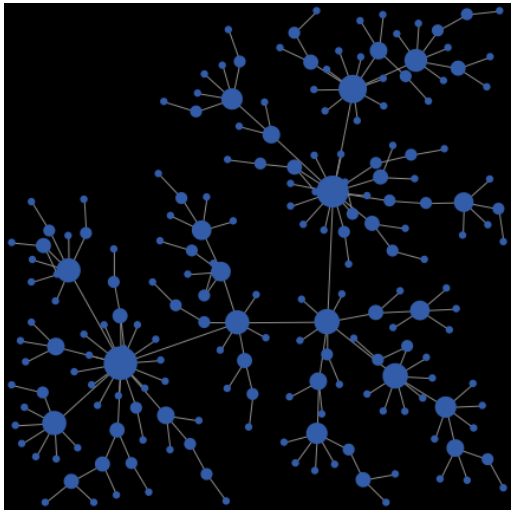
- Neither ordered nor completely random
- Clustering coefficient, local density, is high
- Shortest path length, avg distance between nodes, is short



Scale-free networks

(Barabasi and Albert, 1999)

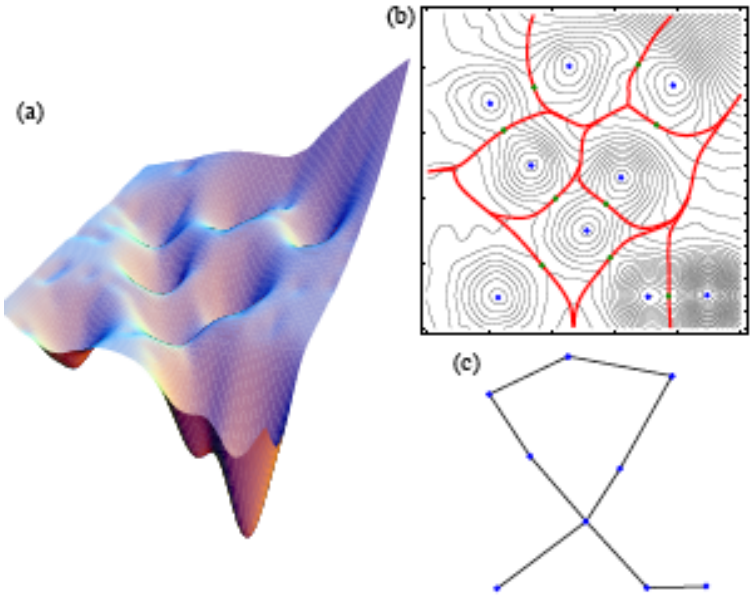
- Degree distribution is right-skewed with a heavy tail
- Most of the nodes have less-than-average degree, whilst a small fraction of hubs have a large number of connections
- Described mathematically by a power-law



Energy surface and inherent networks (Doye, 2002)

Inherent network

- **Nodes** : energy minima
- **Edges** : two nodes are connected if the energy barrier separating them is sufficiently low (transition state)



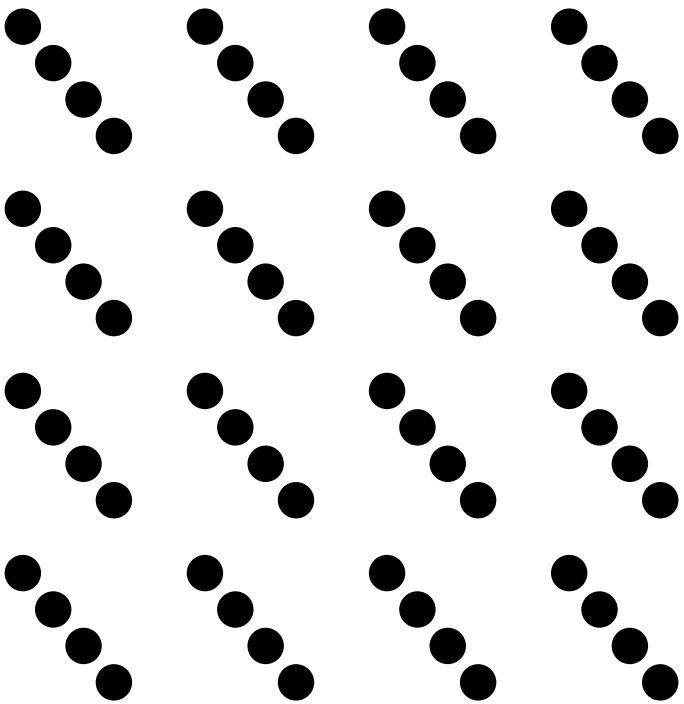
(a) Model of 2D energy surface

(b) Contour plot, partition of the configuration space into basins of attraction surrounding minima

(c) Landscape as a network

Basins of attraction in combinatorial optimization

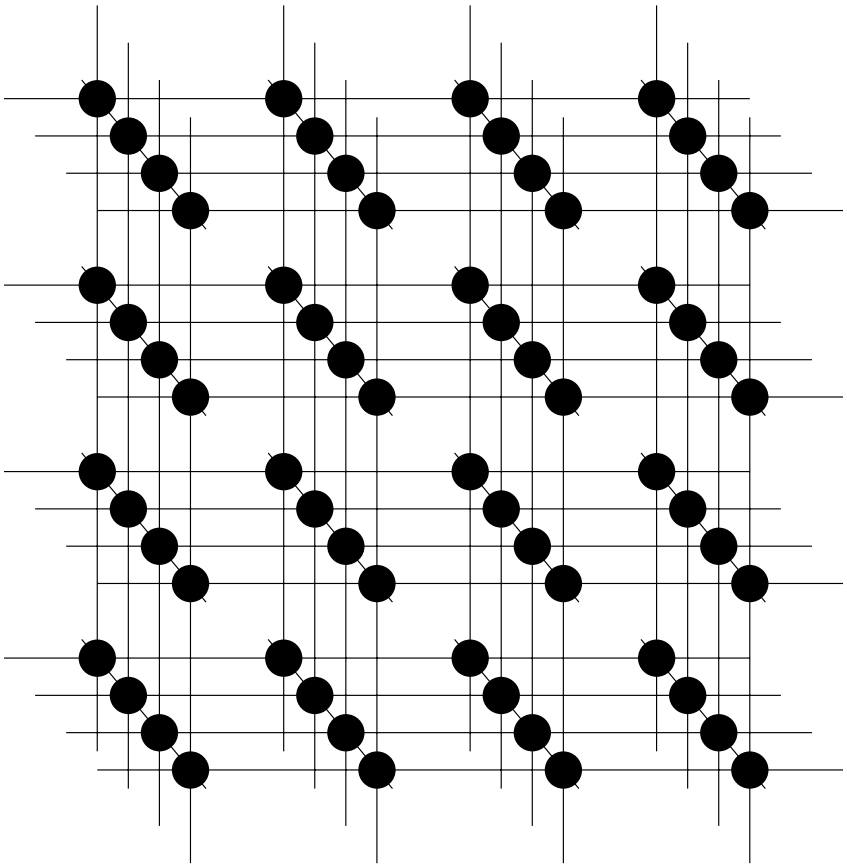
Example of small NK landscape with $N = 6$ and $K = 2$



- Bit strings of length $N = 6$
- $2^6 = 64$ solutions
- one point = one solution

Basins of attraction in combinatorial optimization

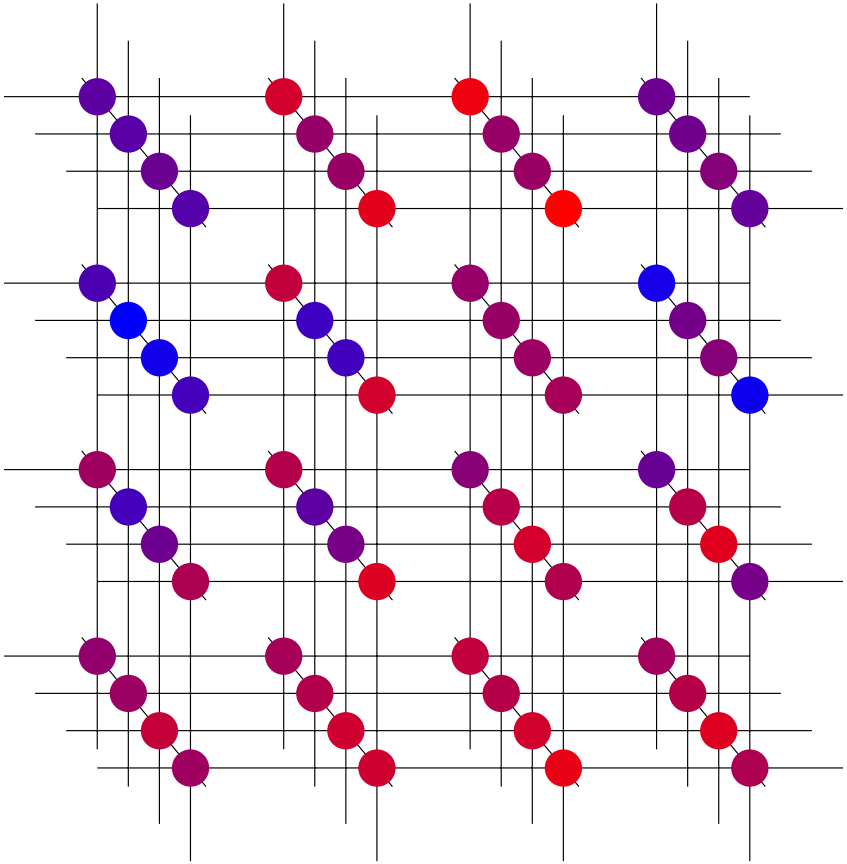
Example of small NK landscape with $N = 6$ and $K = 2$



- Bit strings of length $N = 6$
- Neighborhood size $= 6$
- Line between points = solutions are neighbors
- Hamming distances between solutions are preserved (except for at the border of the cube)

Basins of attraction in combinatorial optimization

Example of small NK landscape with $N = 6$ and $K = 2$

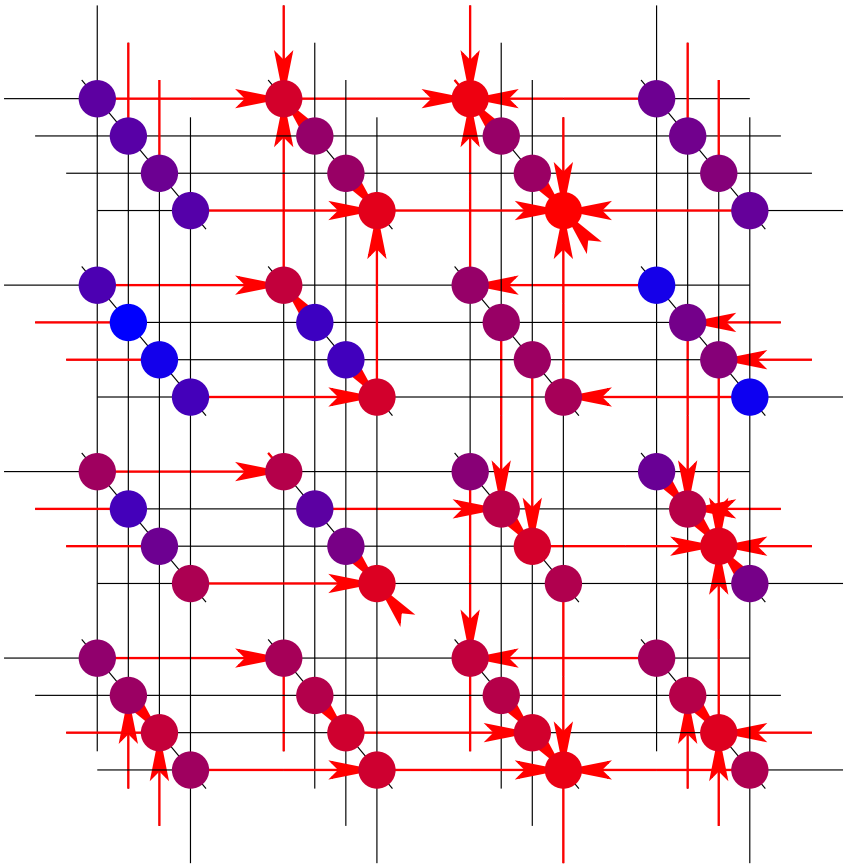


Color represent fitness value

- high fitness
- low fitness

Basins of attraction in combinatorial optimization

Example of small NK landscape with $N = 6$ and $K = 2$



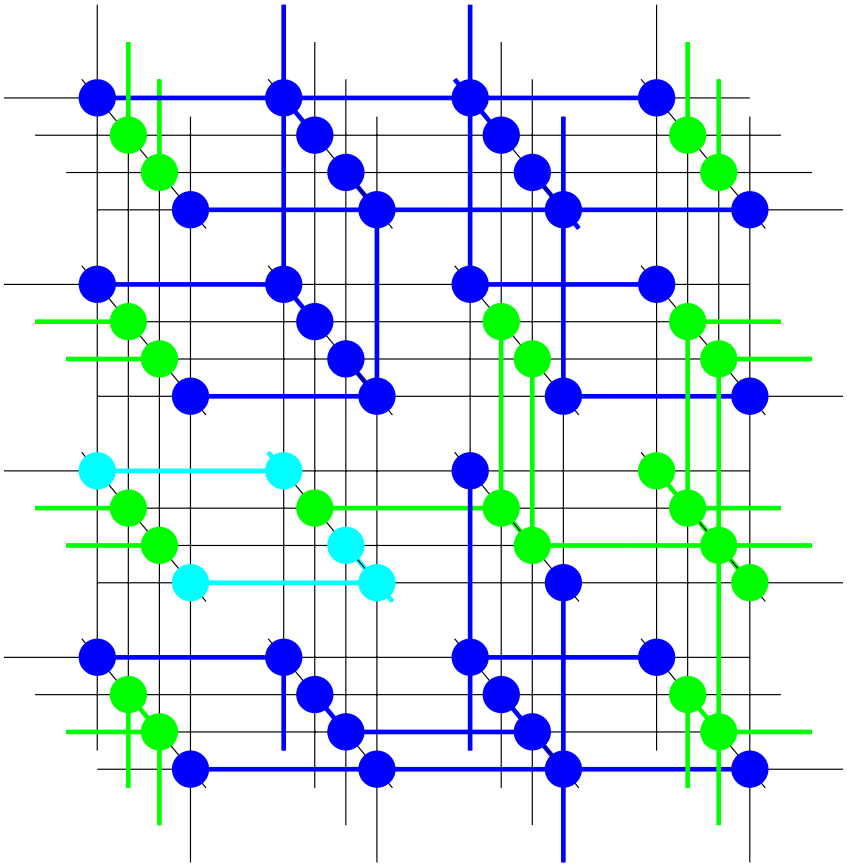
- Color represent fitness value
  - high fitness
  - low fitness
- $\rightarrow$ point towards the solution with highest fitness in the neighborhood

Exercise :

Why not make a Hill-Climbing walk on it ?

Basins of attraction in combinatorial optimization

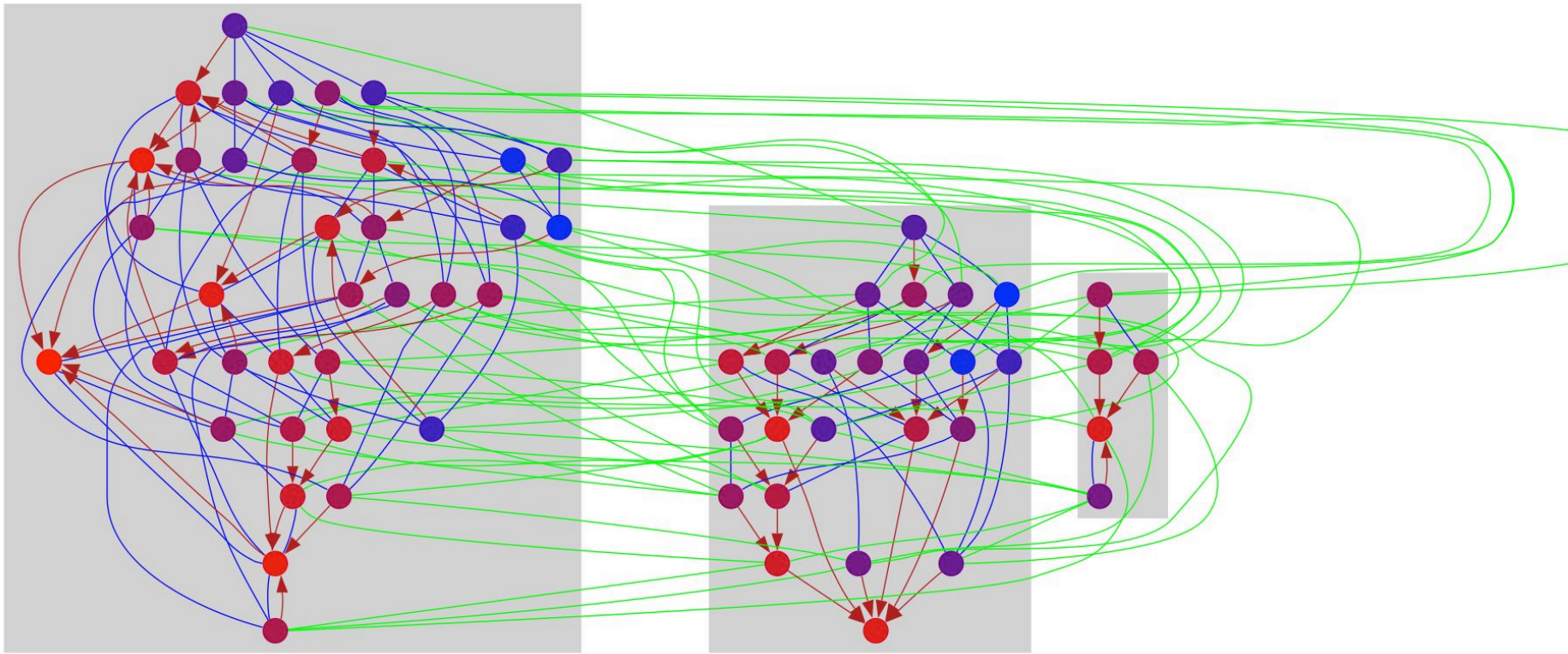
Example of small NK landscape with $N = 6$ and $K = 2$



- Each color corresponds to one basin of attraction
- Basins of attraction are interlinked and overlapped
- Basins have no "interior"

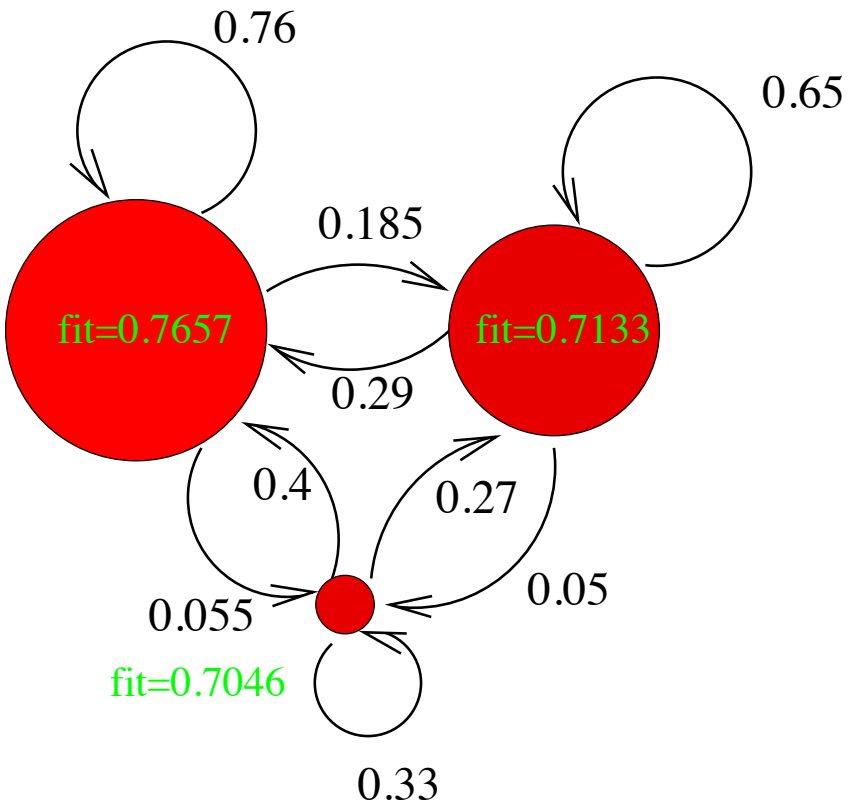
Basins of attraction in combinatorial optimization

Example of small NK landscape with $N = 6$ and $K = 2$



- Basins of attraction are interlinked and overlapped !
- Most neighbors of a given solution are outside its basin

Local optima network



- Nodes :
  local optima
- Edges :
  transition probabilities

Basin of attraction

Hill-Climbing algorithm (best-improvement)

Choose initial solution $s \in S$
repeat
  choose $s' \in \mathcal{N}(s)$ such that $f(s') = max_{x \in \mathcal{N}(s)} f(x)$
  if $f(s) < f(s')$ **then**
    $s \leftarrow s'$
  end if
until s is a Local optimum

Basin of attraction of s^* :

$$b_{s^*} = \{s \in \mathcal{S} \mid HillClimbing(s) = s^*\}.$$

local optima network

Local optima network

- Nodes : set of local optima $\mathcal{S}^*$
- Edges : notion of connectivity between basins of attraction

local optima network

Local optima network

- Nodes : set of local optima $\mathcal{S}^*$
- Edges : notion of connectivity between basins of attraction

2 possible definitions of edges

- **Basin-transition edges** :
  transition between random solutions from basin b_i to basin b_j
  (GECCO 2008 [23])
  (ALIFE 2008 [35], Phys. Rev. E 2008 [32], CEC 2010)
- **Escape edges** :
  transition from Local Optimum i to basin b_j
  (EA 2011, GECCO 2012, PPSN 2012, EA 2013)

Basin-transition edges : random transition between basins

Edges

e_{ij} between LO_i and LO_j if $\exists\ s_i \in b_i$ and $s_j \in b_j\ :\ s_j \in \mathcal{N}(s_i)$

Basin-transition edges : random transition between basins

Edges

e_{ij} between LO_i and LO_j if $\exists\ s_i \in b_i$ and $s_j \in b_j$: $s_j \in \mathcal{N}(s_i)$

Prob. from solution s to solution $s^{'}$

$$p(s \rightarrow s^{'}) = \Pr(s^{'} = op(s))$$

For example, $\mathcal{S} = \{0,1\}^N$ and bit-flip operator

- if $s^{'} \in \mathcal{N}(s)$, $p(s \rightarrow s^{'}) = \frac{1}{N}$
- if $s^{'} \notin \mathcal{N}(s)$, $p(s \rightarrow s^{'}) = 0$

Basin-transition edges : random transition between basins

Edges

e_{ij} between LO_i and LO_j if $\exists\ s_i \in b_i$ and $s_j \in b_j : s_j \in \mathcal{N}(s_i)$

Prob. from solution s to solution $s^{'}$

$$p(s \rightarrow s^{'}) = \Pr(s^{'} = op(s))$$

Prob. from solution s to basin b_j

$$p(s \rightarrow b_j) = \sum_{s^{'} \in b_j} p(s \rightarrow s^{'})$$

Basin-transition edges : random transition between basins

Edges

e_{ij} between LO_i and LO_j if $\exists\ s_i \in b_i$ and $s_j \in b_j : s_j \in \mathcal{N}(s_i)$

Prob. from solution s to solution $s^{'}$

$$p(s \rightarrow s^{'}) = \Pr(s^{'} = op(s))$$

Prob. from solution s to basin b_j

$$p(s \rightarrow b_j) = \sum_{s^{'} \in b_j} p(s \rightarrow s^{'})$$

Weights : Transition prob. from basin b_i to basin b_j

$$w_{ij} = p(b_i \rightarrow b_j) = \frac{1}{\sharp b_i} \sum_{s \in b_i} p(s \rightarrow b_j)$$

Escape edges : transition prob. from local optimum

Edges

e_{ij} between LO_i and LO_j if $\exists s : distance(LO_i, s) \leqslant D$ and $s \in b_j$.

Weights

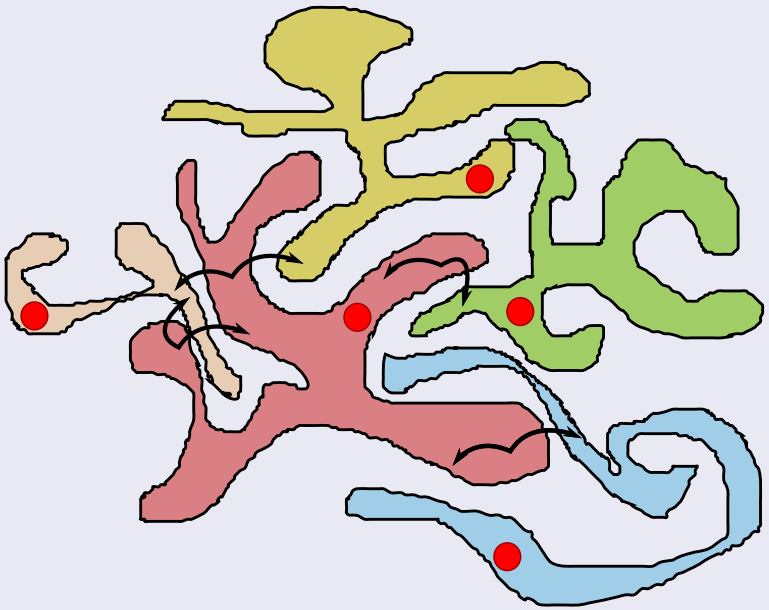
$$w_{ij} = \sharp\{s \in S \mid d(LO_i, s) \leqslant D \text{ and } s \in b_j\}$$

can be normalized by the number of solutions at distance D

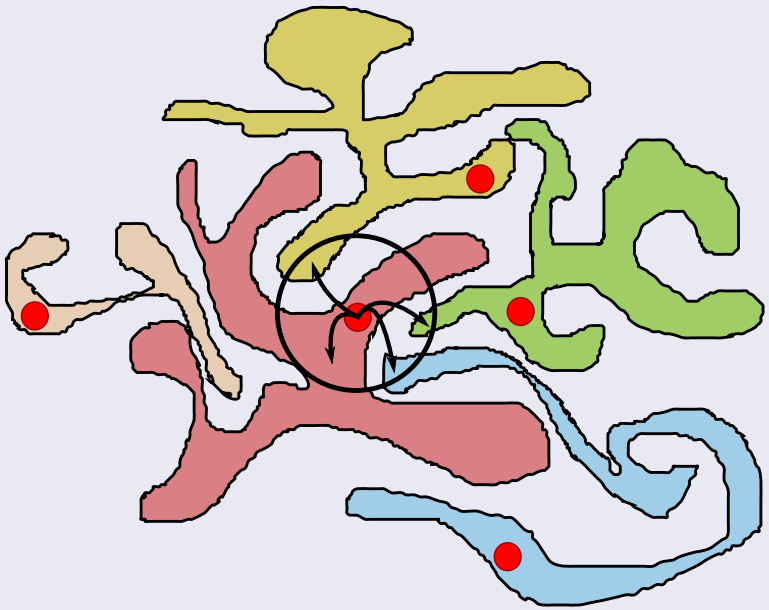
$\Rightarrow$ local optima network : weighted oriented graph

Basin *vs.* Escape edges

LON with basin edges



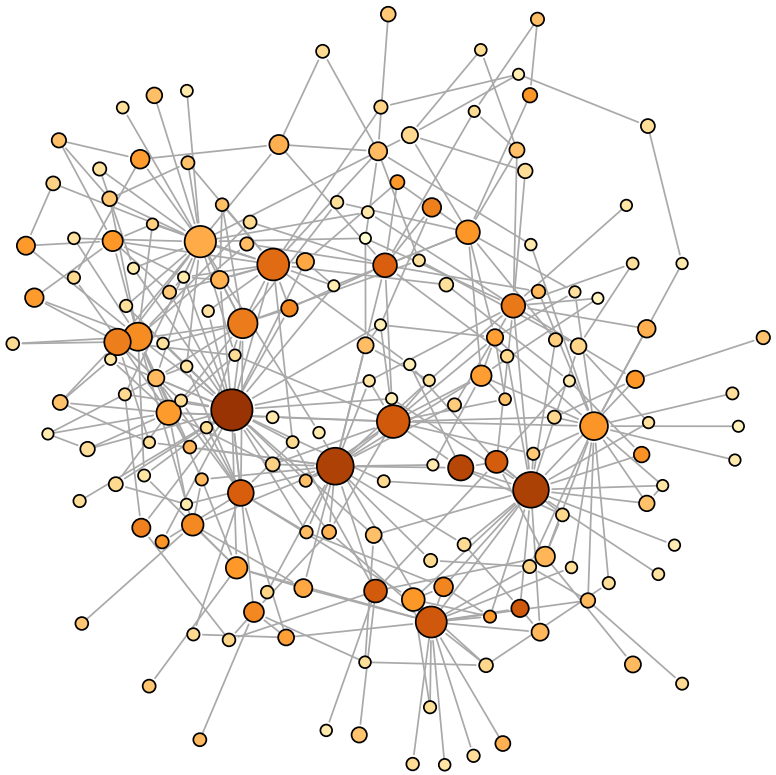
LON with escape edges



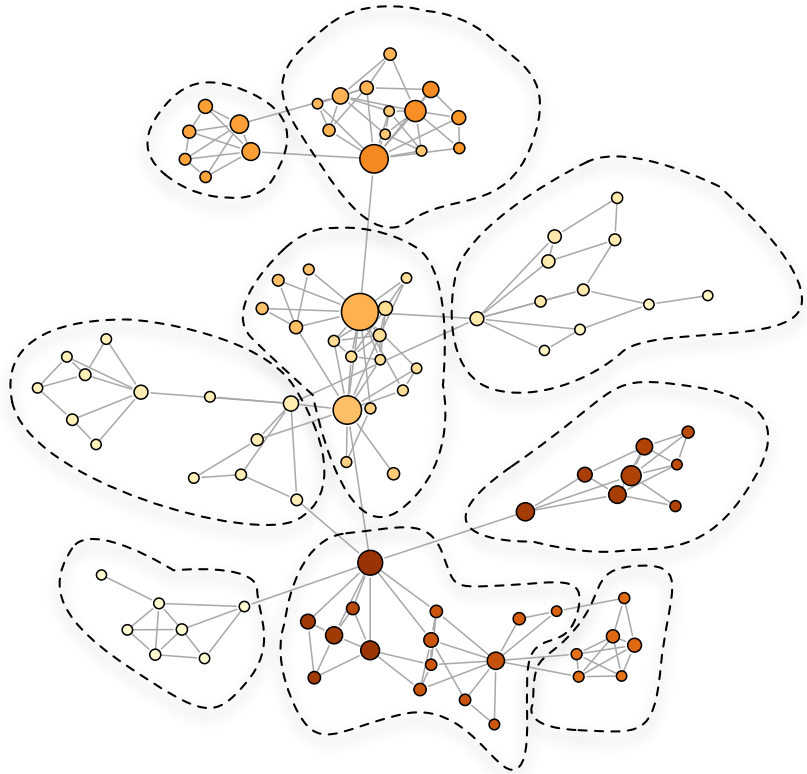
Examples of Local Optima Network with basin edges

Quadratic Assignment Problem [CEC'10 - Phys A'11]

Random instance



Real-like instance



Structure of the Local Optima Network related to search difficulty

NK-landscapes
[Kauffman 1993]

$$x \in \{0,1\}^n \qquad f(x) = \frac{1}{n}\sum_{i=1}^{n} f_i(x_j, x_{i_1}, \ldots, x_{i_k})$$

Two parameters

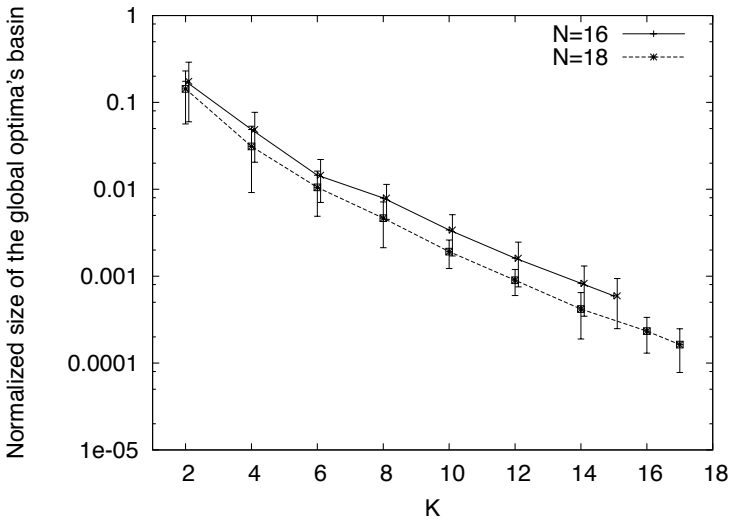
- Problem size n
- Non-linearity k $<$ n
  (multi-modality, epistatic interactions)
  - $k = 0$: linear problem, one single maxima
  - $k = n - 1$: random problem, number of local optima $\frac{2^N}{N+1}$

remarks : "same" results with QAP, flow shop.

Local Optima Network Features

- **Basin of attraction** :
  - Size :
    average, distribution, etc.
  - Fitness of local optima :
    average, distribution, correlation, etc.
- **Network features** :
  - clustering coefficient,
  - shortest path to the global optimum,
  - average path length,
  - weight distribution,
  - disparity,
  - boundary of basins

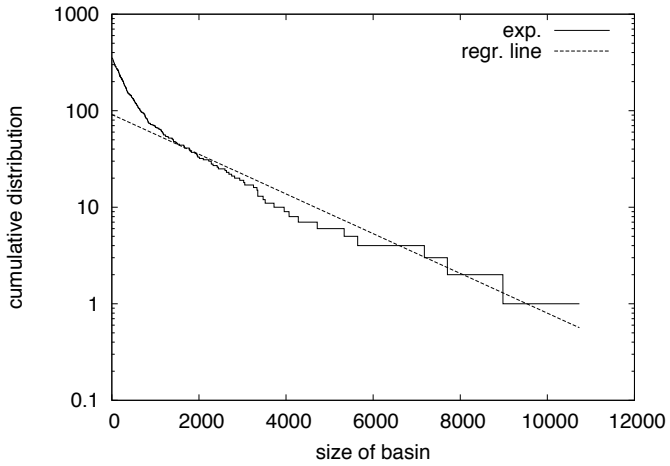
Global optimum basin size *vs.* non-linearity degree *k*



Size of the global maximum basin as a function of non-linearity degree *k*

- Basin size of maximum decreases exponentially with non-linearity degree
- $\Rightarrow$ Difficulty of (best-improvement) hill-climber from a random solution

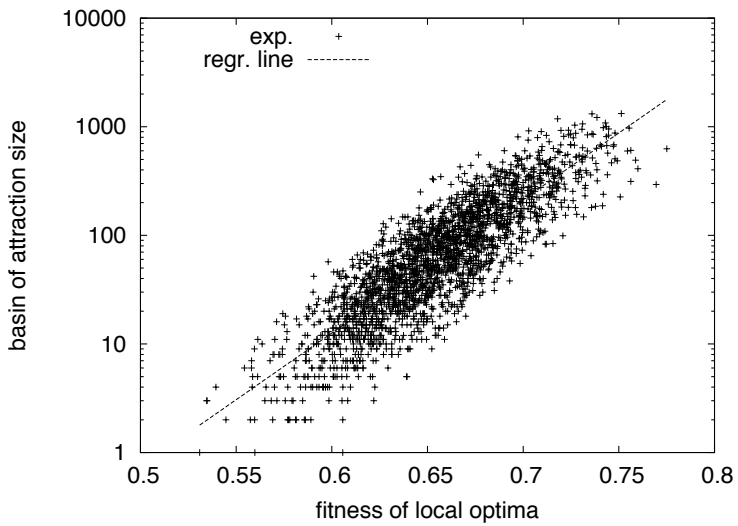
Distribution of basin sizes



Cumulative distribution of basins sizes for $n = 18$ and $k = 4$

- Log-normal cumulative distribution (**not uniform !**) :
  - large number of small basins,
  - small number of large basins.
- Effect of non-linearity :
  the distribution becomes more uniform with non-linearity degree k

Fitness of local optima *vs.* basin size

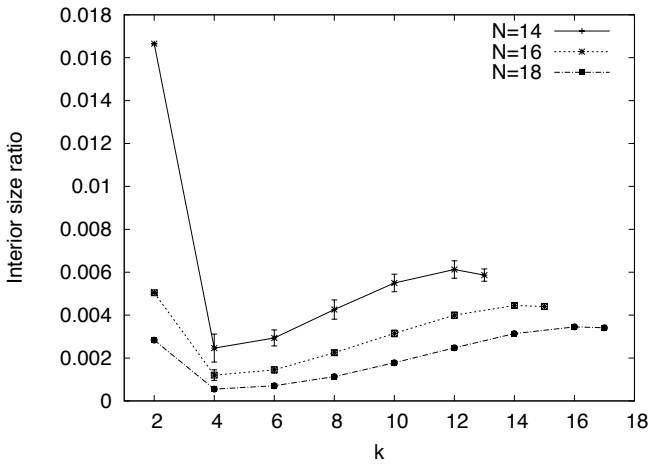


Correlation fitness of local optima *vs.* their corresponding basins sizes

The highest, the largest !

- On average, the global optimum easier to find than one given other local optimum
- But more difficult to find, as the number of local optima increases exponentially with increasing K

Basin : Interior and border sizes



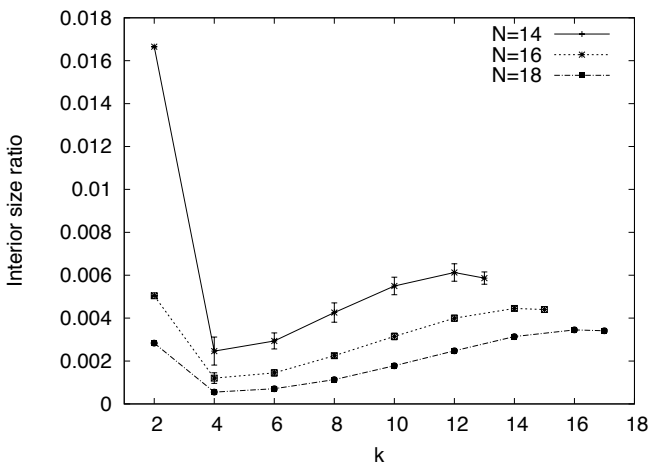
Average of basins interior size ratio

Question :

Do basins look like a "mountain" with interior and border ?

solution $\in$ interior
if all neighbors are in the same basin

Basin : Interior and border sizes



Average of basins interior size ratio

Question :

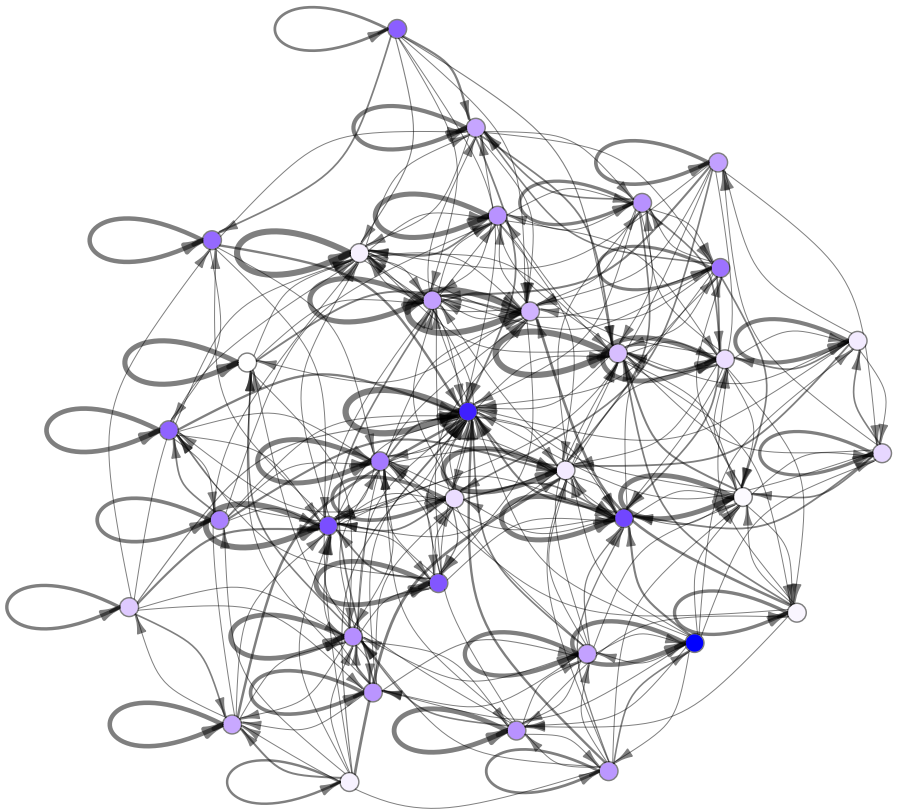
Do basins look like a "mountain" with interior and border ?

solution $\in$ interior
if all neighbors are in the same basin

Answer

- Interior is very small
- Nearly all solutions $\in$ border

Network Metrics



- nv : $\sharp$vertices
- lv : avg path length
  $d_{ij} = 1/w_{ij}$
- lo : path length to best
- fnn : fitness corr.
  $(f(x), f(y))$ with $y \in \mathcal{N}(x)$
- wii : self loops
- cc : clust. coef.
- $zout$: out degree
- $y2$: disparity
- knn : degree corr.
  $(deg(x), deg(y))$ with $y \in \mathcal{N}(x)$

Some formal definitions

Weighted clustering coefficient

local density of the network

$$c^w(i) = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{w_{ij} + w_{ih}}{2} a_{ij} a_{jh} a_{hi}$$

where $s_i = \sum_{j \neq i} w_{ij}$, $a_{nm} = 1$ if $w_{nm} > 0$, $a_{nm} = 0$ if $w_{nm} = 0$ and $k_i = \sum_{j \neq i} a_{ij}$.

Disparity

dishomogeneity of nodes with a given degree

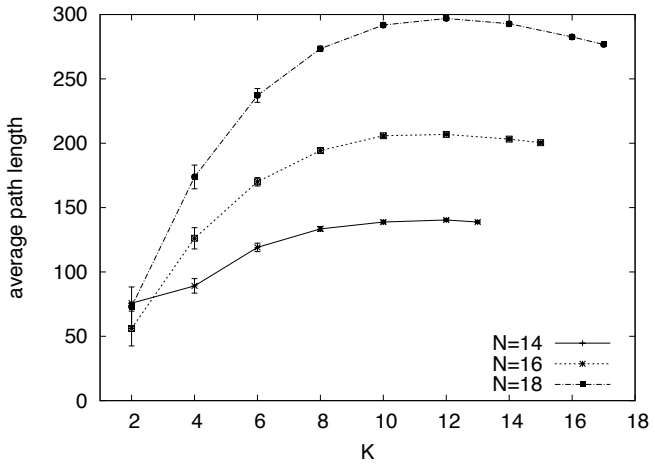
$$Y_2(i) = \sum_{j \neq i} \left(\frac{w_{ij}}{s_i} \right)^2$$

Group averages of all the observed variables

aggregate	nv	lo	lv	fnn	wii	cc	zout	y2	knn	ets ($\times 10^4$)
$K = 2$	43_{28}	33.5_{14}	187_{51}	$0.703_{0.19}$	105_{11}	$0.425_{0.086}$	$6.9_{1.8}$	$0.392_{0.075}$	$0.155_{0.4}$	$2.16_{3.3}$
$K = 4$	221_{39}	53.7_{12}	214_{15}	$0.587_{0.07}$	83.9_{3}	$0.263_{0.013}$	14.3_{1}	$0.219_{0.016}$	$-0.536_{0.13}$	$8.39_{7.74}$
$K = 6$	748_{70}	66.7_{13}	$188_{4.8}$	$0.535_{0.041}$	$67.5_{1.9}$	$0.19_{0.005}$	$24.8_{0.86}$	$0.124_{0.0059}$	$-0.778_{0.035}$	$32.6_{36.1}$
$K = 8$	1669_{73}	$76.6_{9.1}$	$171_{1.9}$	$0.431_{0.025}$	$53.3_{0.88}$	$0.159_{0.0012}$	$35.7_{0.57}$	$0.0769_{0.002}$	$-0.856_{0.022}$	$51.8_{61.1}$
$K = 10$	3148_{110}	$90.7_{8.4}$	$166_{1.2}$	$0.342_{0.016}$	$40.7_{0.78}$	$0.143_{0.00085}$	$47.2_{0.57}$	$0.0491_{0.0011}$	$-0.904_{0.011}$	$81.5_{70.6}$
$K = 12$	5270_{104}	108_{12}	$170_{0.64}$	$0.255_{0.015}$	$30.8_{0.35}$	$0.133_{0.00054}$	$57.8_{0.39}$	$0.0334_{0.00046}$	$-0.928_{0.0093}$	276_{544}
$K = 14$	8100_{121}	$125_{8.6}$	$181_{0.6}$	$0.19_{0.011}$	$23.5_{0.25}$	$0.128_{0.00032}$	$66.9_{0.33}$	$0.0245_{0.00022}$	$-0.944_{0.0063}$	300_{288}
$K = 16$	11688_{101}	146_{11}	$197_{0.42}$	$0.143_{0.0073}$	$18.2_{0.11}$	$0.125_{0.00023}$	$74.6_{0.17}$	$0.0196_{7.8e-05}$	$-0.948_{0.0055}$	414_{632}
$K = 17$	13801_{74}	156_{12}	$205_{0.42}$	$0.133_{0.01}$	$16.1_{0.06}$	$0.125_{0.00021}$	$78.2_{0.13}$	$0.0179_{6.5e-05}$	$-0.944_{0.0063}$	793_{844}

Shortest path length between local optima

Average distance (shortest path) between nodes

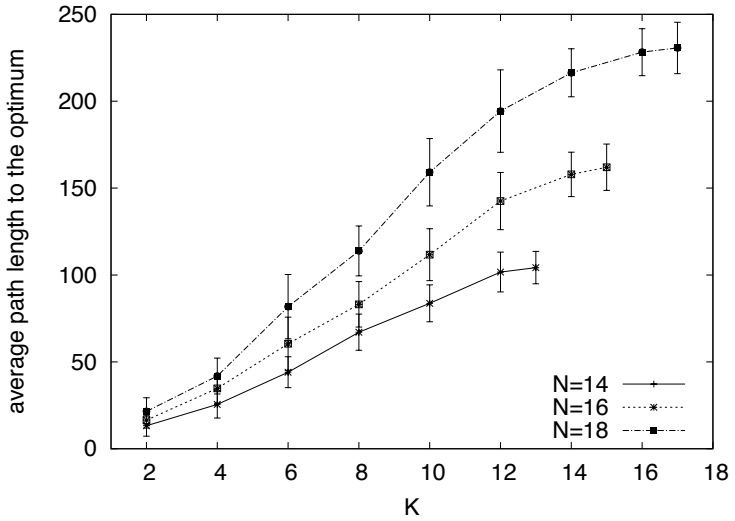


- Related question :

  Are the basins "far" from each other ?
- Increase with problem size n
  (♯ of nodes increases exponentially)
- Increase with non-linearity deg. k up to 10, then stagnates

Shortest path length to global optima

Average path length to the global optimum



- Related question :

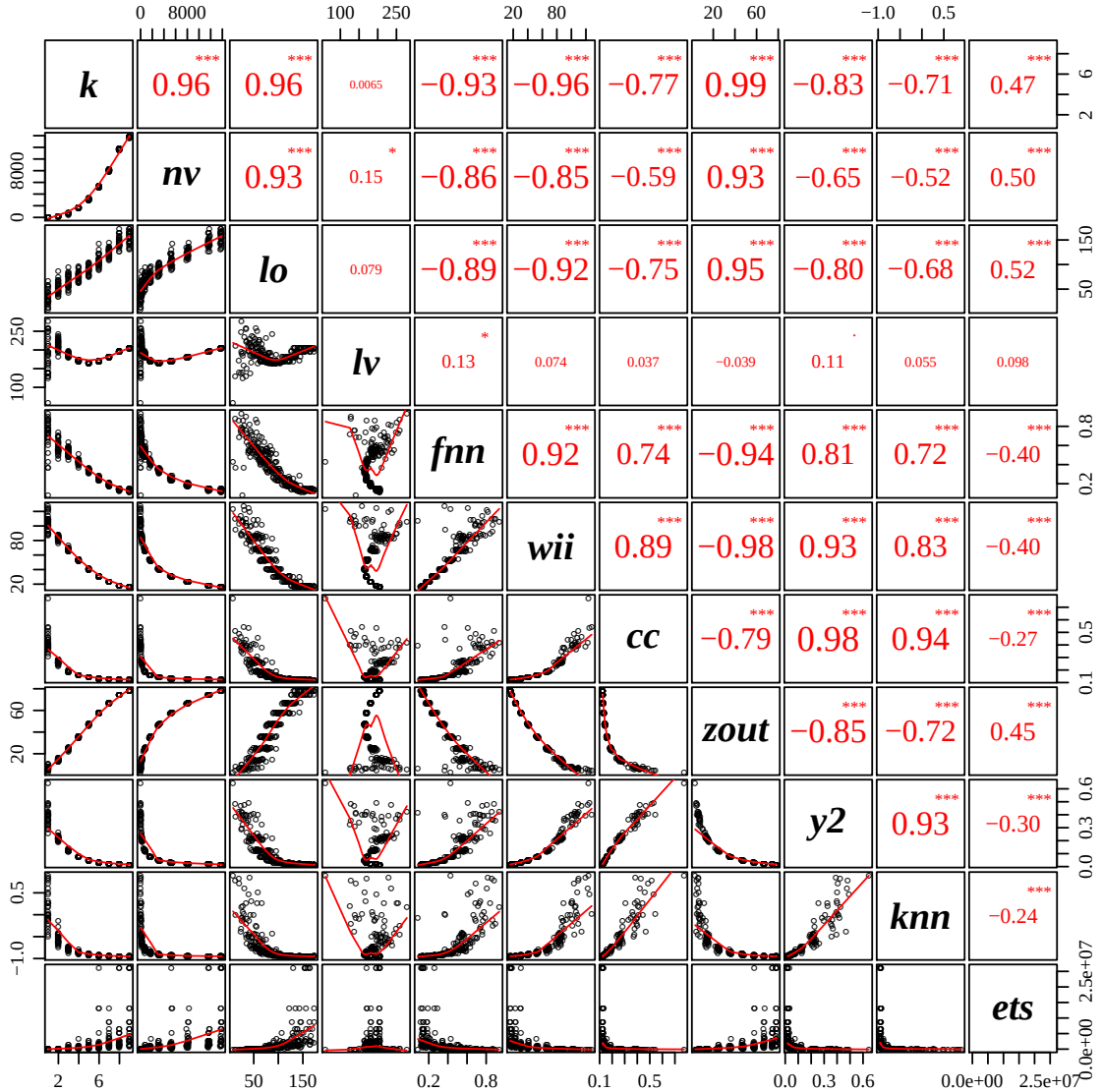
  Is the global optimum basin is far ?
- Increase steadily with non-linearity degree k

Extract information from features

Advanced questions

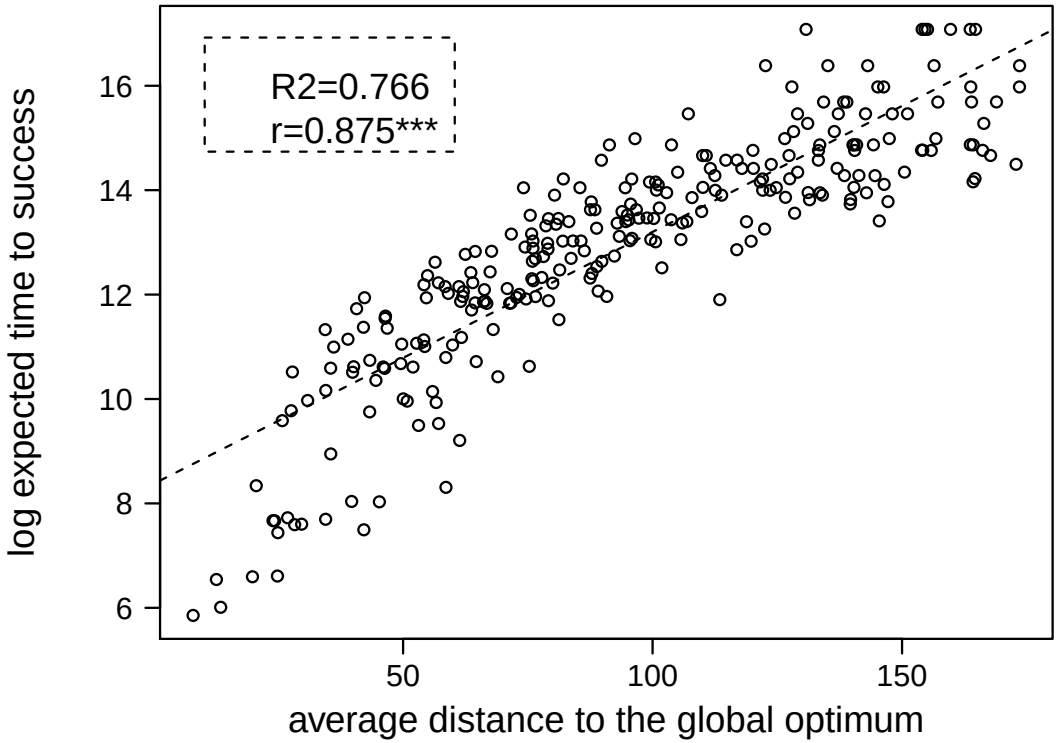
- What information gives each feature ?
- Can we predict the performance from those features ?

Correlation Matrix



ILS Performance vs LON Metrics

NK-landscapes [GECCO'12]



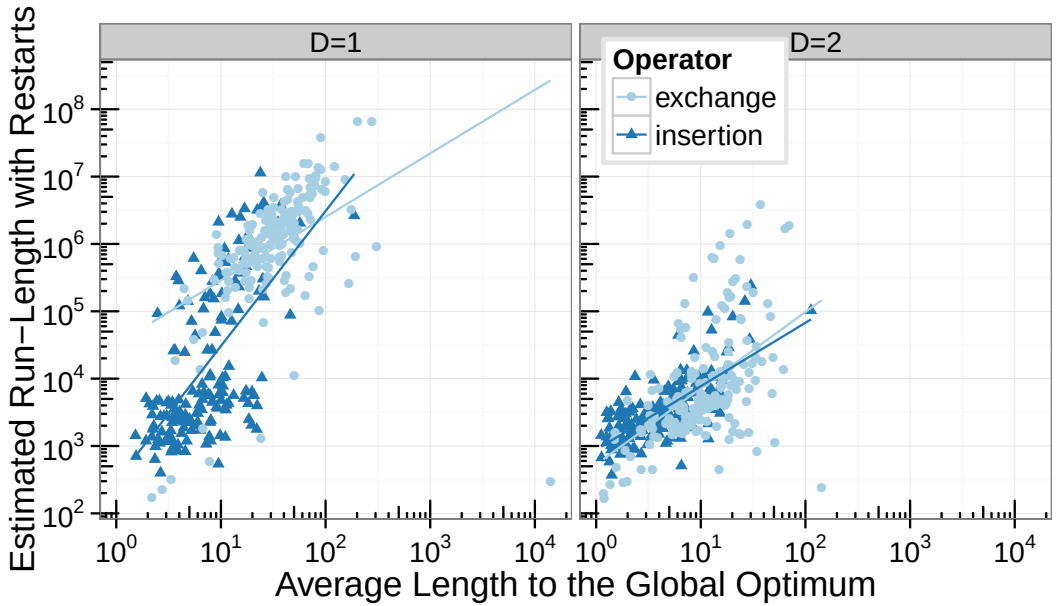
Expected running times

vs.

Average shortest path to the global optimum.

ILS Performance vs LON Metrics

Flow-Shop Scheduling Problem [EA'13]



Expected running times

vs.

Average shortest path to the global optimum.

ILS Performance vs LON Metrics
[GECCO'12]

1. Multiple linear regression on all possible predictors :

$$\log(ets) = \beta_0 + \beta_1 k + \beta_2 \log(nv) + \beta_2 lo + \cdots + \beta_{10} knn + \varepsilon$$

2. Step-wise backward elimination of each predictor in turn.

Predictor	$\widehat{\beta_i}$	Std. Error	p-value
(Intercept)	10.3838	0.58512	$9.24 \cdot 10^{-47}$
lo	0.0439	0.00434	$1.67 \cdot 10^{-20}$
zout	−0.0306	0.00831	$2.81 \cdot 10^{-04}$
y2	−7.2831	1.63038	$1.18 \cdot 10^{-05}$
knn	−0.7457	0.40501	$6.67 \cdot 10^{-02}$

Multiple R-squared : 0.8494, Adjusted R-squared : 0.8471.

Tuning of parameters and Local Optima Network

- Relevant features using "complex system" technics
- Search difficulty is related to the features of LON :
  - Classification of instances
  - Understanding the search difficulty
  - Guideline of design
- Time complexity can be estimated :
  - Tuning with **Portefolio** technics (ongoing work)

Future on local optima network

- Design a method for sampling large search space (under writing)
- Compare the properties of Loc. Opt. Network and the optimal tradeoff between exploration and exploitation
- Study the LON like a fitness landscape

Summary on fitness landscapes

Fitness landscape is a representation of

- search space
- notion of neighborhood
- fitness of solutions

Summary on fitness landscapes

Fitness landscape is a representation of

- search space
- notion of neighborhood
- fitness of solutions

Goal :

- **local description** : fitness between neighbor solutions
  Ruggedness, local optima, fitness cloud, neutral networks, local optima networks...
- and to deduce **global features** :
  - Difficulty !
  - To decide (and control) a good choice of the representation, operator and fitness function

Open issues

- How to control the **parameters** of the algorithm with the local description of fitness landscape ?
- Links between **neutrality** and time complexity (difficulty) ?
- Can fitness landscape describe the dynamics of a **population** of solutions ?
- Fitness landscape for **parallel** algorithm (island model) ?
- What about **crossover** ?

Open issues

- Which "**aggregation of variables**" shows relevant properties of the optimization problem according to the local search heuristic ?

$$\begin{array}{ccc} \mathcal{S} & \xrightarrow{op} & \mathcal{S} \\ p \downarrow & & \downarrow p \\ Z & \xrightarrow{op_Z} & Z \end{array}$$

Open issues

- Which "**aggregation of variables**" shows relevant properties of the optimization problem according to the local search heuristic ?

$$\begin{array}{ccc} \mathcal{S} & \xrightarrow{op} & \mathcal{S} \\ p \downarrow & & \downarrow p \\ Z & \xrightarrow{op_Z} & Z \end{array}$$

Thank you for your attention !

L. Barnett.
Ruggedness and neutrality - the NKp family of fitness landscapes.
In C. Adami, R. K. Belew, H. Kitano, and C. Taylor, editors, *ALIFE VI, Proceedings of the Sixth International Conference on Artificial Life*, pages 18–27. ALIFE, The MIT Press, 1998.

Lionel Barnett.
Netcrawling - optimal evolutionary search with neutral networks.
In *Proceedings of the 2001 Congress on Evolutionary Computation CEC2001*, pages 30–37, COEX, World Trade Center, 159 Samseong-dong, Gangnam-gu, Seoul, Korea, 27-30 2001. IEEE Press.

U. Bastolla, M. Porto, H. E. Roman, and M. Vendruscolo.
Statiscal properties of neutral evolution.
Journal Molecular Evolution, 57(S) :103–119, August 2003.

Meriema Belaidouni and Jin-Kao Hao.
An analysis of the configuration space of the maximal constraint satisfaction problem.
In *PPSN VI : Proceedings of the 6th International Conference on Parallel Problem Solving from Nature*, pages 49–58, London, UK, 2000. Springer-Verlag.

P. Collard, M. Clergue, and M. Defoin Platel.
Synthetic neutrality for artificial evolution.
In *Artificial Evolution : Fourth European Conference AE'99*, pages 254–265. Springer-Verlag, 2000.
Selected papers in Lecture Notes in Computer Sciences 1829.

J. C. Culberson.
Mutation-crossover isomorphisms and the construction of discrimination function.
Evolutionary Computation, 2 :279–311, 1994.

J. P. K. Doye.
The network topology of a potential energy landscape : a static scale-free network.
Phys. Rev. Lett., 88 :238701, 2002.

J. P. K. Doye and C. P. Massen.
Characterizing the network topology of the energy landscapes of atomic clusters.
J. Chem. Phys., 122 :084105, 2005.

Ricardo Garcia-Pelayo and Peter F. Stadler.
Correlation length, isotropy, and meta-stable states.
Physica D, 107 :240–254, 1997.
Santa Fe Institute Preprint 96-05-034.

Josselin Garnier and Leila Kallel.
Efficiency of local search with multiple local optima.
SIAM Journal on Discrete Mathematics, 15(1) :122–141, 2002.

P. Gitchoff and G. Wagner.
Recombination induced hypergraphs : A new approach to mutation-recombination isomorphism, 1996.

David E. Goldberg and Philip Segrest.
Finite markov chain analysis of genetic algorithms.
In *ICGA*, pages 1–8, 1987.

M. Huynen.
Exploring phenotype space through neutral evolution.
Journal Molecular Evolution, 43 :165–169, 1996.

E. Izquierdo-Torres.
The role of nearly neutral mutations in the evolution of dynamical neural networks.
In J. Pollack and al, editors, *Ninth International Conference of the Simulation and Synthesis of Living Systems (Alife 9)*, pages 322–327. MIT Press, 2004.

T. Jones.
Evolutionary Algorithms, Fitness Landscapes and Search.
PhD thesis, University of New Mexico, Albuquerque, 1995.

S. A. Kauffman.
The Origins of Order.
Oxford University Press, New York, 1993.

M. Kimura.
The Neutral Theory of Molecular Evolution.
Cambridge University Press, Cambridge, UK, 1983.

J. Lobo, J. H. Miller, and W. Fontana.
Neutrality in technology landscape, 2004.

M. Newman and R. Engelhardt.
Effect of neutral selection on the evolution of molecular species.

In *Proc. R. Soc. London B.*, volume 256, pages 1333–1338, 1998.

Erik Van Nimwegen, James P. Crutchfield, and Martijn Huynen.
Metastable evolutionary dynamics : Crossing fitness barriers or escaping via neutral paths ?
Technical Report 99-07-041, SanteFe institute, 1999.

Marmion, M.-É., Dhaenens, C., Jourdan, L., Liefooghe, A., Verel, S., 2011. NILS : a neutrality-based iterated local search and its application to flowshop scheduling. In : Proceedings of the 12th European Conference of Evolutionary Computation in Combinatorial Optimization. Vol. 6622 of EvoCOP 2011. LNCS, Springer, pp. 191– 202.

Marmion, M.-É., Dhaenens, C., Jourdan, L., Liefooghe, A., Verel, S., 2011. On the neutrality of flowshop scheduling fitness landscapes. In : Proceedings of the 5th Learning and

Intelligent OptimizatioN Conference. Vol. 6683 of LION 2011. LNCS, Springer, pp. 238–252.

Gabriela Ochoa, Marco Tomassini, Sébastien Verel, and Christian Darabos.
A Study of NK Landscapes' Basins and Local Optima Networks.
In *Proceedings of the 10th annual conference on Genetic and evolutionary computation Genetic And Evolutionary Computation Conference*, pages 555–562, Atlanta États-Unis d'Amérique, 07 2008. ACM New York, NY, USA.
best paper nomination.

M. Defoin Platel.
Homologie en Programmation Génétique - Application à la résolution d'un problème inverse.
PhD thesis, Université Nice Sophia Antipolis, France, 2004.

Eduardo Rodriguez-Tello, Jin-Kao Hao, and Jose Torres-Jimenez.
A new evaluation function for the minla problem.
In *Proceedings of the MIC 2005*, pages 796–801, Vienna Austria, 2005.

Helge Rosé, Werner Ebeling, and Torsten Asselmeyer.
The density of states - a measure of the difficulty of optimisation problems.
In *Parallel Problem Solving from Nature*, pages 208–217, 1996.

P. Schuster, W. Fontana, P. F. Stadler, and I. L. Hofacker.
From sequences to shapes and back : a case study in RNA secondary structures.
In *Proc. R. Soc. London B.*, volume 255, pages 279–284, 1994.

Peter F. Stadler.
Landscapes and their correlation functions.

J. Math. Chem., 20 :1–45, 1996.

Peter F. Stadler and W. Schnabl.
The landscape of the traveling salesmen problem.
Phys. Letters, A(161) :337–344, 1992.

Peter F. Stadler and Gunter P. Wagner.
Algebraic theory of recombination spaces.
Evolutionary Computation, 5(3) :241–275, 1997.

Terry Stewart.
Extrema selection : Accelerated evolution on neutral networks.
In *Proceedings of the 2001 Congress on Evolutionary Computation CEC2001*, pages 25–29, COEX, World Trade Center, 159 Samseong-dong, Gangnam-gu, Seoul, Korea, 27-30 May 2001. IEEE Press.

Marco Tomassini, Sébastien Verel, and Gabriela Ochoa.

Complex-network analysis of combinatorial spaces : The NK landscape case.
Physical Review E : Statistical, Nonlinear, and Soft Matter Physics, 78(6) :066114, 12 2008.
89.75.Hc ; 89.75.Fb ; 75.10.Nr.

Vesselin K. Vassilev and Julian F. Miller.
The advantages of landscape neutrality in digital circuit evolution.
In *ICES*, pages 252–263, 2000.

Sebastien Verel, Philippe Collard, and Manuel Clergue.
Measuring the evolvability landscape to study neutrality.
In M. Keijzer and et al., editors, *Poster at Genetic and Evolutionary Computation – GECCO-2006*, pages 613–614, Seatle, 8-12 July 2006. ACM Press.

Sébastien Verel, Gabriela Ochoa, and Marco Tomassini.

The Connectivity of NK Landscapes' Basins : A Network Analysis.
In *Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems Artificial Life XI*, pages 648–655, Winchester France, 08 2008. MIT Press, Cambridge, MA.
tea team.

F. Chicano, F. Daolio, G. Ochoa, S. Verel, M. Tomassini, and E. Alba.
Local optima networks, landscape autocorrelation and heuristic search performance.
In *Proceedings of Parallel Problem Solving from Nature - PPSN XII*, volume 7492 of *Lecture Notes in Computer Science*, pages 337–347. Springer, 2012.

E. D. Weinberger.

Correlated and uncorrelatated fitness landscapes and how to tell the difference.
In *Biological Cybernetics*, pages 63 :325–336, 1990.

F. Daolio, M. Tomassini, S. Verel, and G. Ochoa.
Communities of minima in local optima networks of combinatorial spaces.
Physica A : Statistical Mechanics and its Applications, 390(9) :1684–1694, 2011.

G. Ochoa, S. Verel, and M. Tomassini.
First-improvement vs. best-improvement local optima networks of nk landscapes.
In *Proceedings of Parallel Problem Solving from Nature - PPSN XI*, volume 6238 of *Lecture Notes in Computer Science*, pages 104–113. Springer, 2010.

L. Vanneschi, M. Tomassini, P. Collard, and S. Verel.

Negative slope coefficient. a measure to characterize genetic programming fitness landscapes.
In P. Collet et al., editor, *Proceedings of the 9th European Conference on Genetic Programming*, volume 3905 of *Lecture Notes in Computer Science*, pages 178–189. Springer, Berlin, Heidelberg, New York, 2006.

S. Verel, F. Daolio, G. Ochoa, and M. Tomassini.
Local optima networks with escape edges.
In *Proceedings of the International Conference on Artificial Evolution, EA-2011*, volume 7401 of *Lecture Notes in Computer Science*, pages 49–60. Springer, 2012.

S. Verel, G. Ochoa, and M. Tomassini.
Local optima networks of NK landscapes with neutrality.
IEEE Transactions on Evolutionary Computation, 15(6) :783–797, 2011.

Katherine M. Malan and Andries P. Engelbrecht.
A survey of techniques for characterising fitness landscapes and some possible ways forward.
Information Sciences, (0) :–, 2013.

Guanzhou Lu, Jinlong Li, and Xin Yao.
Fitness-probability cloud and a measure of problem hardness for evolutionary algorithms.
In *Proceedings of the 11th European conference on Evolutionary computation in combinatorial optimization*, EvoCOP'11, pages 108–117, Berlin, Heidelberg, 2011. Springer-Verlag.

S. Wright.
The roles of mutation, inbreeding, crossbreeding, and selection in evolution.
In *Proceedings of the Sixth International Congress of Genetics 1*, pages 356–366, 1932.