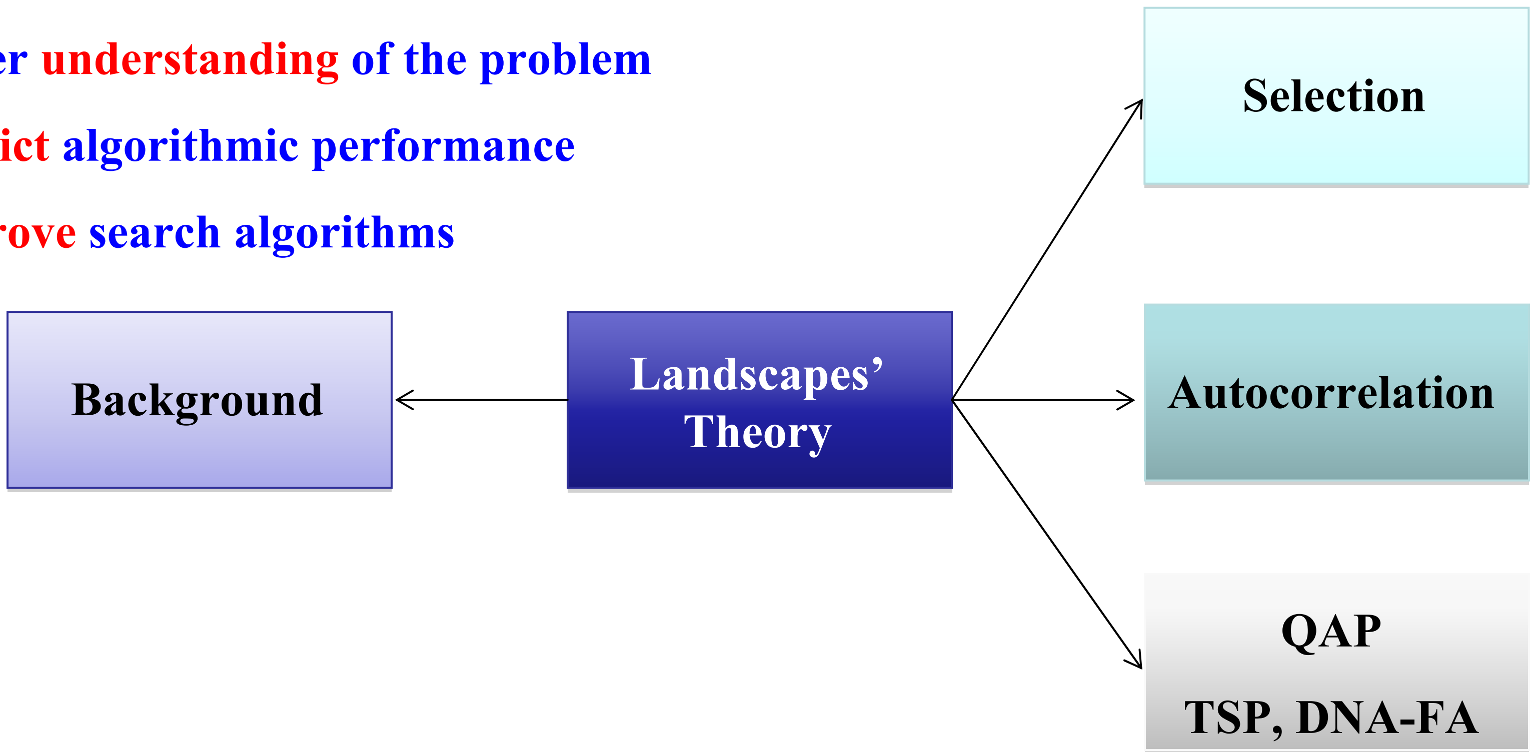
Elementary Landscape Decomposition of the Quadratic Assignment Problem

Francisco Chicano, Gabriel Luque, and Enrique Alba

Motivation

- Landscapes' theory is a tool for analyzing optimization problems
- Applications in Chemistry, Physics, Biology and Combinatorial Optimization
- Central idea: study the search space to obtain information
    - Better understanding of the problem
    - Predict algorithmic performance
    - Improve search algorithms



Landscape Definition

- A **landscape** is a triple (X,N,f) where
  - X is the solution space
  - N is the neighbourhood operator
  - f is the objective function

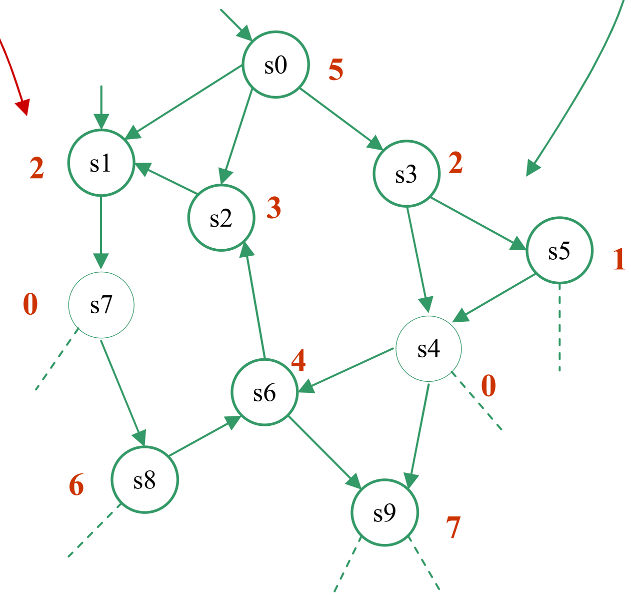
The pair (X,N) is called **configuration space**

- The **neighbourhood operator** is a function

  $N: X \rightarrow \mathcal{P}(X)$

- Solution y is **neighbour of** x if $y \in N(x)$
- **Regular** and **symmetric** neighbourhoods
  - $d=|N(x)| \quad \forall x \in X$
  - $y \in N(x) \Leftrightarrow x \in N(y)$
- **Objective function**

  $f: X \rightarrow R$ (or N, Z, Q)



Elementary Landscapes: Formal Definition

- **An elementary function is an eigenvector of the graph Laplacian (plus constant)**

Adjacency matrix

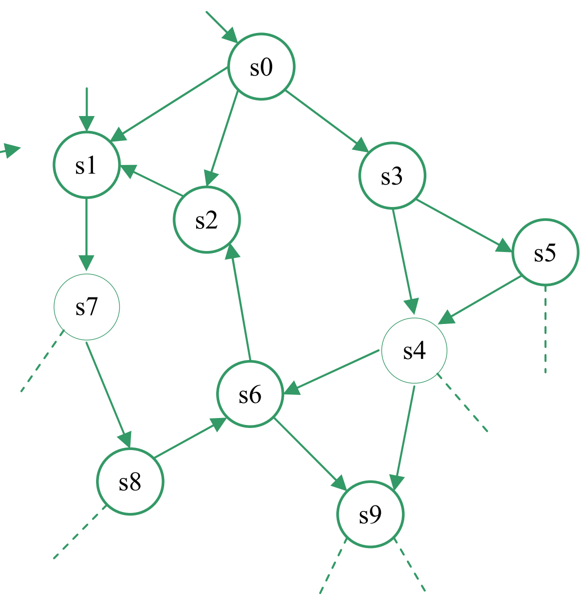
$$A_{xy} = \begin{cases} 1 & \text{if } y \in N(x) \\ 0 & \text{otherwise} \end{cases}$$

Degree matrix

$$D_{xy} = \begin{cases} |N(x)| & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

- **Graph Laplacian:** $\Delta = A - D$

Depends on the configuration space



- **Elementary function: eigenvector of Δ (plus constant)**

$$\Delta \times (\vec{f} - b) = \lambda \cdot (\vec{f} - b)$$

Eigenvalue (λ)

Elementary Landscapes: Characterizations

- **An elementary landscape is a landscape for which**

$$\operatorname*{avg}_{y \in N(x)}\{f(y)\} = \alpha f(x) + \beta \quad \forall x \in X$$

where

$$\operatorname*{avg}_{y \in N(x)}\{f(y)\} \stackrel{\text{def}}{=} \frac{1}{d} \sum_{y \in N(x)} f(y)$$

Depend on the problem/instance (α, β)

Linear relationship

- **Grover's wave equation**

$$\operatorname*{avg}_{y \in N(x)}\{f(y)\} = f(x) + \frac{k}{d}\left(\bar{f} - f(x)\right) \quad \forall x \in X$$

$$\bar{f} = \frac{1}{|X|} \sum_{y \in X} f(y)$$

$$\alpha = 1 - \frac{k}{d} \qquad \beta = \frac{k}{d}\bar{f}$$

Characteristic constant: $k = -\lambda$

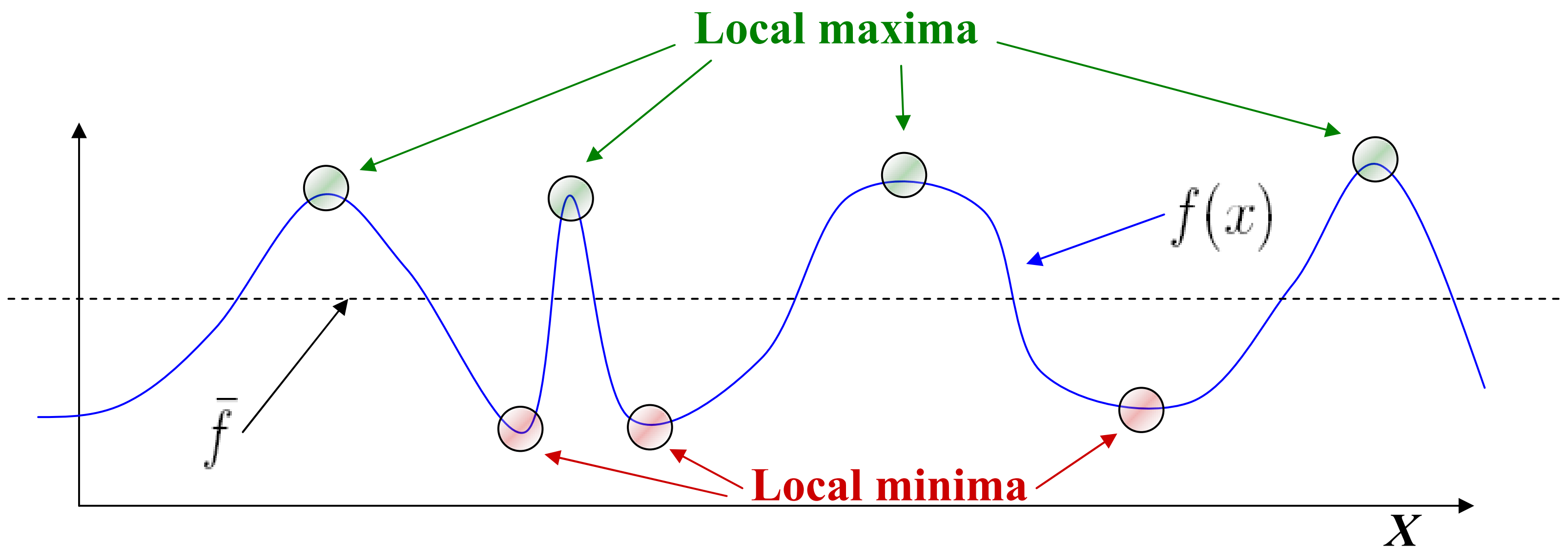
Elementary Landscapes: Properties

- **Several properties of elementary landscapes are the following**

$$f(x) < \min\left\{\operatorname*{avg}_{y\in N(x)}\{f(y)\}, \bar{f}\right\} \quad \text{or} \quad f(x) > \max\left\{\operatorname*{avg}_{y\in N(x)}\{f(y)\}, \bar{f}\right\}$$

where $f(x) \neq \bar{f}$

- **Local maxima and minima**

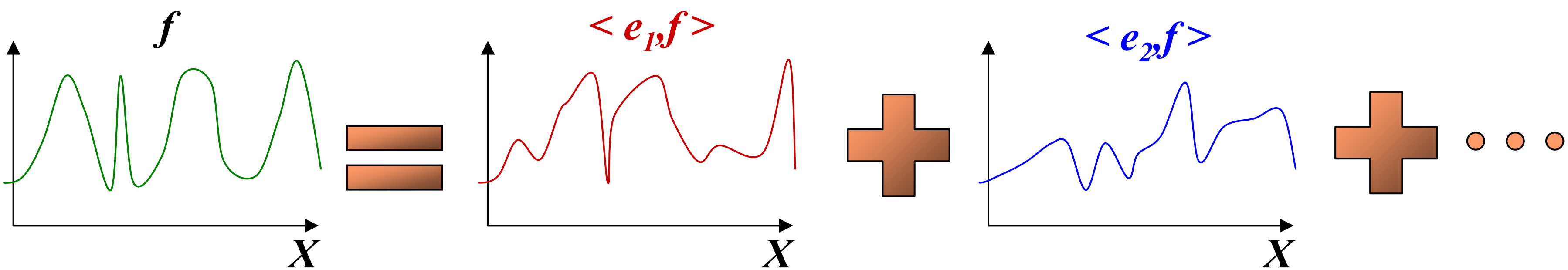


Elementary Landscapes: Examples

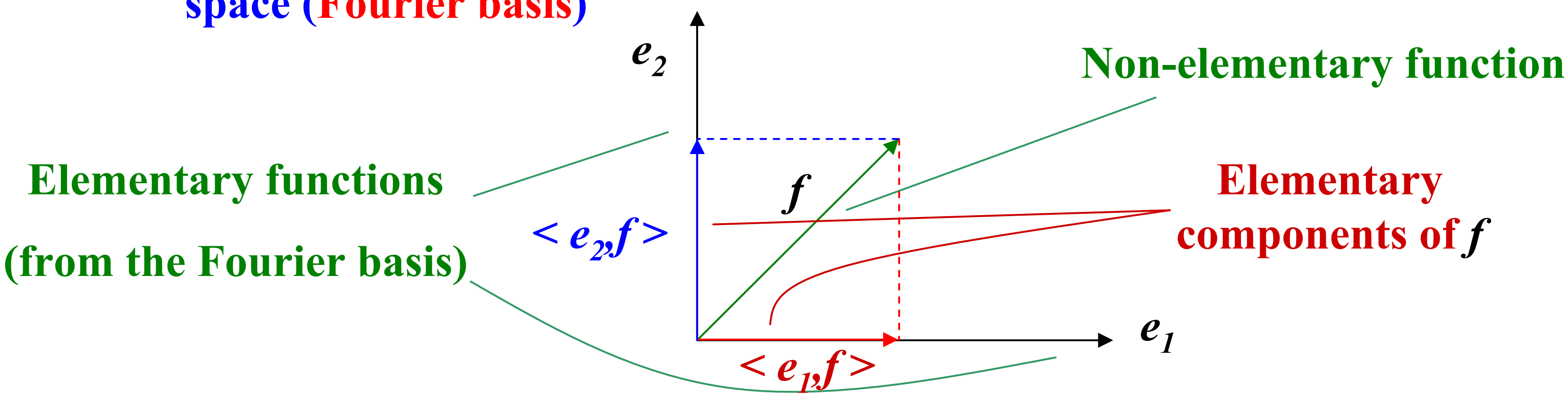
Problem	Neighbourhood	d	k
Symmetric TSP	2-opt	$n(n-3)/2$	$n-1$
	swap two cities	$n(n-1)/2$	$2(n-1)$
Antisymmetric TSP	inversions	$n(n-1)/2$	$n(n+1)/2$
	swap two cities	$n(n-1)/2$	$2n$
Graph α-Coloring	recolor 1 vertex	$(\alpha-1)n$	2α
Graph Matching	swap two elements	$n(n-1)/2$	$2(n-1)$
Graph Bipartitioning	Johnson graph	$n^2/4$	$2(n-1)$
NEAS	bit-flip	n	4
Max Cut	bit-flip	n	4
Weight Partition	bit-flip	n	4

Landscape Decomposition

- **What if the landscape is not elementary?**
- **Any landscape can be written as the sum of elementary landscapes**



- **There exists a set of eigenfunctions of Δ that form a basis of the function space (Fourier basis)**

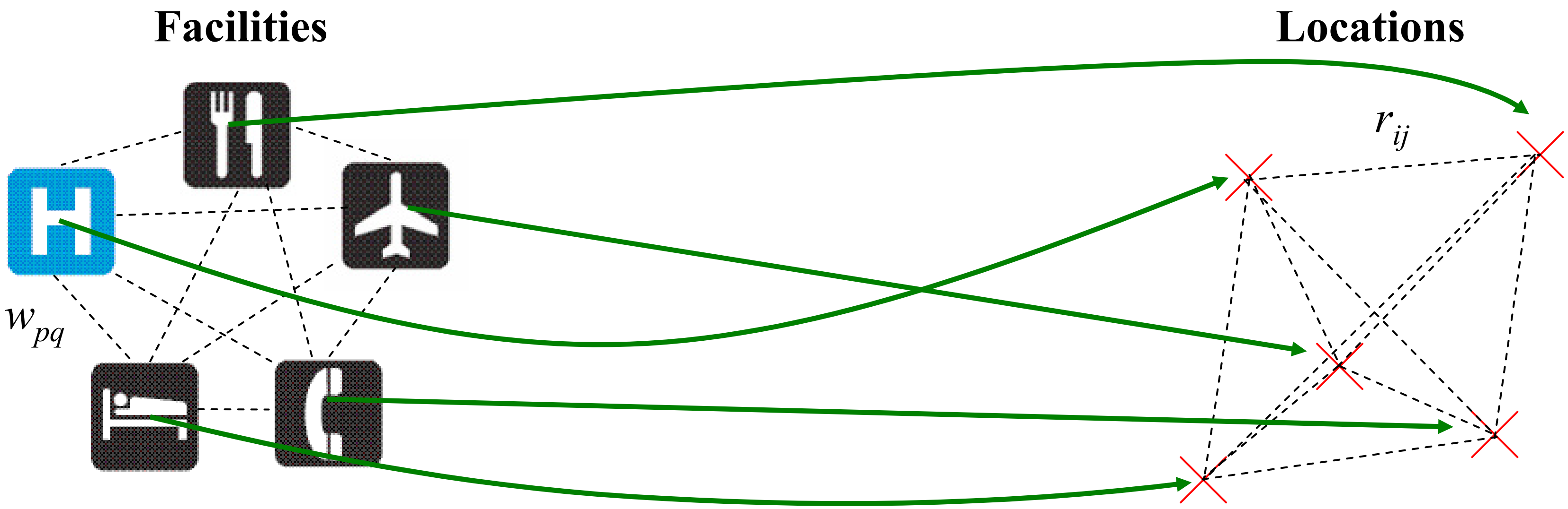


Landscape Decomposition: Examples

Problem	Neighbourhood	d	Components
General TSP	inversions	$n(n-1)/2$	2
	swap two cities	$n(n-1)/2$	2
Subset Sum Problem	bit-flip	n	2
MAX k-SAT	bit-flip	n	k
NK-landscapes	bit-flip	n	$k+1$
Radio Network Design	bit-flip	n	max. nb. of reachable antennae
Frequency Assignment	change 1 frequency	$(\alpha-1)n$	2
QAP	swap two elements	$n(n-1)/2$	3

Quadratic Assignment Problem: Definition

- A **QAP** instance is composed of n facilities and n locations

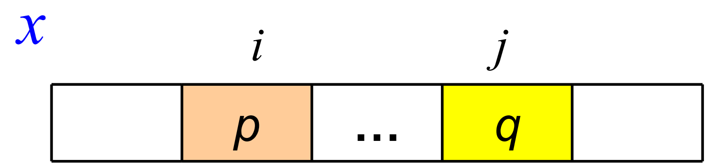


- A **distance** r_{ij} is specified between each pair of locations
- A **flow** w_{pq} is specified between each pair of facilities
- The problem consists in assigning the facilities to the locations **minimizing cost**

$$f(x) = \sum_{i,j=1}^{n} r_{ij} w_{x(i)x(j)}$$ where x is a permutation (solution)

Quadratic Assignment Problem: Decomposition

- **Auxiliary φ functions**



$$\phi^{\alpha,\beta,\gamma,\varepsilon,\zeta}_{(i,j),(p,q)}(x) = \begin{cases} \alpha & \text{if } x(i) = p \wedge x(j) = q \\ \beta & \text{if } x(i) = q \wedge x(j) = p \\ \gamma & \text{if } x(i) = p \oplus x(j) = q \\ \varepsilon & \text{if } x(i) = q \oplus x(j) = p \\ \zeta & \text{if } x(i) \neq p,q \wedge x(j) \neq p,q \end{cases}$$

- **The Ω functions are defined after the φ functions**

Characteristic constant

$$\Omega^{1}_{(i,j),(p,q)} \stackrel{\text{def}}{=} \phi^{n-3,1-n,-2,0,-1}_{(i,j),(p,q)}$$

$k_1 = 2n$

$$\Omega^{2}_{(i,j),(p,q)} \stackrel{\text{def}}{=} \phi^{n-3,n-3,0,0,1}_{(i,j),(p,q)}$$

$k_2 = 2(n-1)$

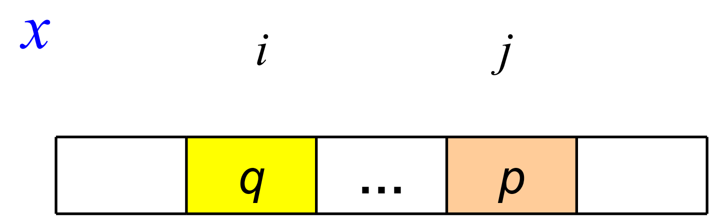
$$\Omega^{3}_{(i,j),(p,q)} \stackrel{\text{def}}{=} \phi^{2n-3,1,n-2,0,-1}_{(i,j),(p,q)}$$

$k_3 = n$

Quadratic Assignment Problem: Decomposition

- **Auxiliary φ functions**

$$\phi_{(i,j),(p,q)}^{\alpha,\beta,\gamma,\varepsilon,\zeta}(x) = \begin{cases} \alpha & \text{if } x(i) = p \wedge x(j) = q \\ \beta & \text{if } x(i) = q \wedge x(j) = p \\ \gamma & \text{if } x(i) = p \oplus x(j) = q \\ \varepsilon & \text{if } x(i) = q \oplus x(j) = p \\ \zeta & \text{if } x(i) \neq p,q \wedge x(j) \neq p,q \end{cases}$$



- **The Ω functions are defined after the φ functions**

Characteristic constant

$$\Omega^1_{(i,j),(p,q)} \stackrel{\text{def}}{=} \phi_{(i,j),(p,q)}^{n-3,1-n,-2,0,-1}$$

$k_1 = 2n$

$$\Omega^2_{(i,j),(p,q)} \stackrel{\text{def}}{=} \phi_{(i,j),(p,q)}^{n-3,n-3,0,0,1}$$

$k_2 = 2(n-1)$

$$\Omega^3_{(i,j),(p,q)} \stackrel{\text{def}}{=} \phi_{(i,j),(p,q)}^{2n-3,1,n-2,0,-1}$$

$k_3 = n$

Quadratic Assignment Problem: Decomposition

- **Using the swap neighborhood, the fitness function can be decomposed into three elementary components:**

$$f_{c1}(x) = \sum_{\substack{i,j,p,q=1 \\ i \neq j \\ p \neq q}}^{n} r_{ij} w_{pq} \frac{\Omega^1_{(i,j),(p,q)}(x)}{2n}$$

$k_1 = 2n$

$$f_{c2}(x) = \sum_{\substack{i,j,p,q=1 \\ i \neq j \\ p \neq q}}^{n} r_{ij} w_{pq} \frac{\Omega^2_{(i,j),(p,q)}(x)}{2(n-2)}$$

$k_2 = 2(n-1)$

Kronecker's delta

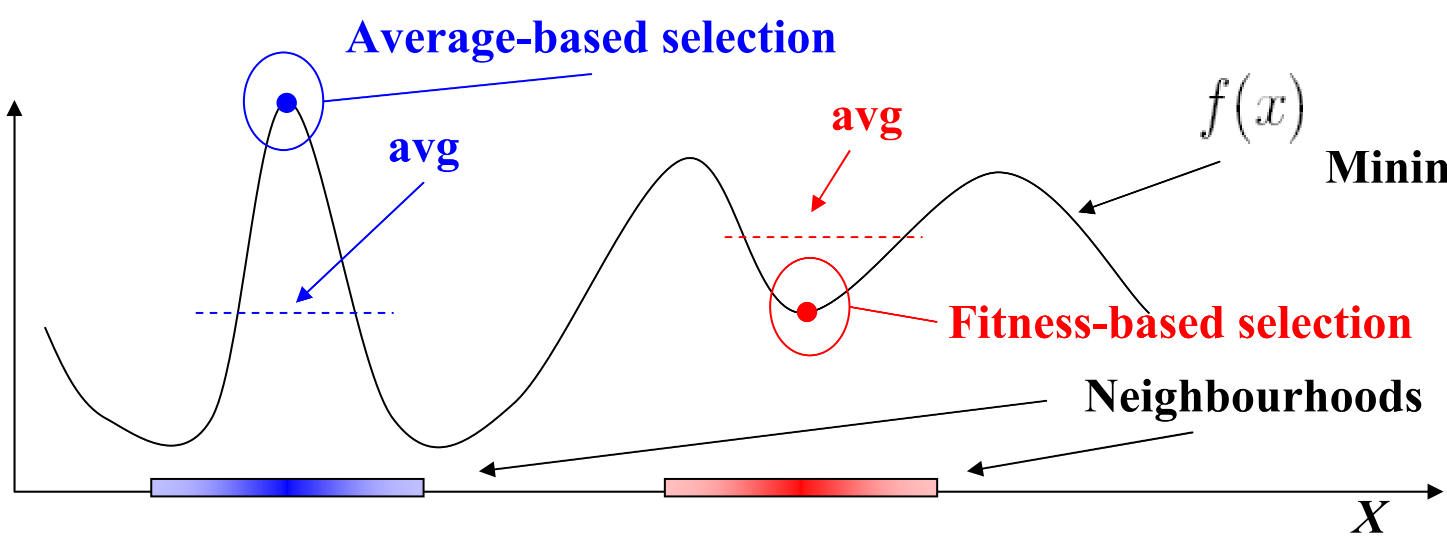
$$f_{c3}(x) = \sum_{\substack{i,j,p,q=1 \\ i \neq j \\ p \neq q}}^{n} r_{ij} w_{pq} \frac{\Omega^3_{(i,j),(p,q)}(x)}{n(n-2)} + \sum_{i,p=1}^{n} r_{ii} w_{pp} \delta^{p}_{x(i)}$$

$k_3 = n$

$$f(x) = f_{c1}(x) + f_{c2}(x) + f_{c3}(x)$$

New Selection Strategy

- **Selection operators usually take into account the fitness value of the individuals**



- **We can improve the selection operator by selecting the individuals according to the average value in their neighbourhoods**

New Selection Strategy

- In elementary landscapes the traditional and the new operator are the same!

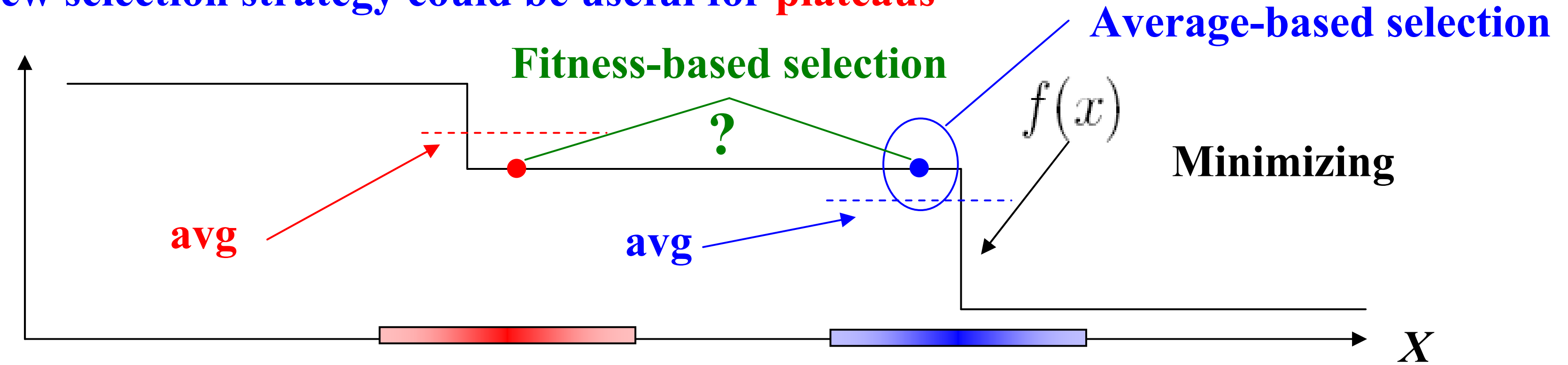
  Recall that... $$\operatorname*{avg}_{y \in N(x)}\{f(y)\} = \alpha f(x) + \beta \quad \forall x \in X$$

- However, they are not the same in non-elementary landscapes. If we have *n* elementary components, then:

$$\operatorname*{avg}_{y \in N(x)}\{f(y)\} = \alpha_0 + \alpha_1 f(x) + \sum_{i=2}^{n} \alpha_i f_i(x) \quad \forall x \in X$$

Elementary components

- The new selection strategy could be useful for plateaus



Autocorrelation

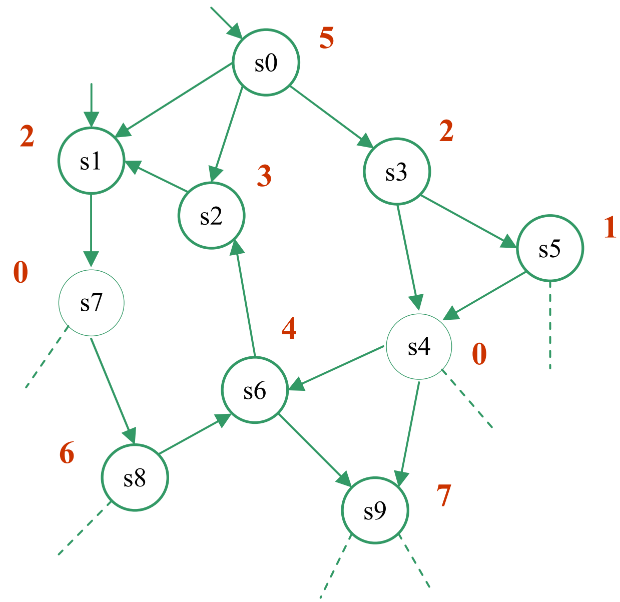
- Let $\{x_0, x_1, \ldots\}$ a simple random walk on the configuration space where $x_{i+1} \in N(x_i)$
- The random walk induces a time series $\{f(x_0), f(x_1), \ldots\}$ on a landscape.
- The autocorrelation function is defined as:

$$r(s) = \frac{\langle f(x_t) f(x_{t+s}) \rangle_t - \langle f \rangle^2}{\langle f^2 \rangle - \langle f \rangle^2}$$

- The autocorrelation length and coefficient:

$$l = \sum_{s=0}^{\infty} r(s) \qquad \xi = \frac{1}{1 - r(1)}$$

- Autocorrelation length conjecture:



> The number of local optima in a search space is roughly
> $$M \approx |X| / |X(x_0, l)|$$

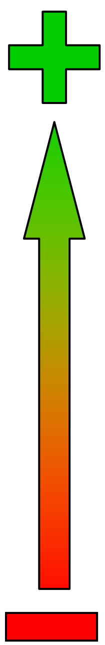
Solutions reached from x_0 after l moves

Autocorrelation Length Conjecture

- **The higher the value of *l* and ξ the smaller the number of local optima**
- ***l* and ξ is a measure of ruggedness**

Angel, Zissimopoulos. Theoretical Computer Science 263:159-172 (2001)

Length Coefficient



Ruggedness	SA (configuration 1)		SA (configuration 2)	
	% rel. error	nb. of steps	% rel. error	nb. of steps
$10 \le \zeta < 20$	0.2	50,500	0.1	101,395
$20 \le \zeta < 30$	0.3	53,300	0.2	106,890
$30 \le \zeta < 40$	0.3	58,700	0.2	118,760
$40 \le \zeta < 50$	0.5	62,700	0.3	126,395
$50 \le \zeta < 60$	0.7	66,100	0.4	133,055
$60 \le \zeta < 70$	1.0	75,300	0.6	151,870
$70 \le \zeta < 80$	1.3	76,800	1.0	155,230
$80 \le \zeta < 90$	1.9	79,700	1.4	159,840
$90 \le \zeta < 100$	2.0	82,400	1.8	165,610

Autocorrelation and Landscapes

- **If f is a sum of elementary landscapes:**

Fourier coefficients

$$r(s) = \sum_{i \neq 0} \frac{a_i^2}{\sum_{j \neq 0} a_j^2} \left(1 - \frac{k_i}{d}\right)^s \qquad l = \sum_{i \neq 0} \frac{d \; a_i^2}{k_i \sum_{j \neq 0} a_j^2}$$

- **Summing all the squared coefficients with the same k_i:**

$$r(s) = \sum_i W_i \left(1 - \frac{k_i}{d}\right)^s \qquad \xi = \frac{d}{\sum_i W_i k_i} \qquad l = d \sum_i \frac{W_i}{k_i}$$

where

$$W_i = \frac{\overline{f_{ci}^2} - \overline{f_{ci}}^2}{\overline{f^2} - \overline{f}^2}$$

Autocorrelation for QAP

- **Using the landscape decomposition we can compute *l* and *ξ***

$$\xi = \frac{n(n-1)}{(4-2W_3)n - 4W_2} \qquad l = \frac{(1+W_3)(n-1)+W_2}{4}$$

- **The bounds for these parameters are:**

$$\frac{n-1}{4} \leq \xi \leq \frac{n-1}{2} \qquad \frac{n-1}{4} \leq l \leq \frac{n-1}{2}$$

- **We designed an $O(n^2)$ algorithm for this computation (optimal complexity)**
- **We computed *l* and *ξ* for all the instances in the QAPLIB**

Autocorrelation for QAP

- **Using the landscape decomposition we can compute *l* and *ξ***

Instance	ξ	ℓ	Instance	ξ	ℓ
bur26a	11.825	12.130	esc32b	8.000	8.000
bur26b	11.727	12.073	esc32c	8.000	8.000
bur26c	12.109	12.291	esc32d	8.000	8.000
bur26d	12.050	12.258	esc32e	8.000	8.000
bur26e	12.032	12.248	esc32f	8.000	8.000
bur26f	11.962	12.208	esc32g	8.000	8.000
bur26g	12.323	12.407	esc32h	8.000	8.000
bur26h	12.296	12.392	esc64a	16.000	16.00
chr12a	3.096	3.171	had12	3.743	4
chr12b	3.201	3.346	had14	4.319	4.
chr12c	3.044	3.079	had16	4.405	
chr15a	3.917	4.049	had18	5.08	
chr15b	4.126	4.388	had20		
chr15c	3.843	3.920	kra30a		
chr18a	4.585	4.658	kra30		
chr18b	4.632	4.742			
chr20a	5.105	5.195			
chr20b	5.035	5.067	lipa2		
chr20c	5.260	[illegible]			

$$\xi = \frac{\ldots 1) + W_2}{(4 \ldots}$$

- **The bounds f...**

$$\frac{n - \ldots}{4 \ldots} \leq \ell \leq \frac{n-1}{2}$$

- **We designed ... putation (optimal complexity)**
- **We compute... in the QAPLIB**

Subproblems of QAP: TSP and DNA-FA

- **TSP is a particular case of QAP with two elementary components (f_{c1} and f_{c2})**

$$f(x) = \sum_{i=1}^{n-1} w_{x(i)x(i+1)} + w_{x(n)x(1)} \qquad r_{ij} = \begin{cases} 1 & \text{if } j = (i \mod n) + 1 \\ 0 & \text{otherwise} \end{cases}$$

- **The autocorrelation parameters are:** $\xi = \frac{n(n-1)}{4(n-W_2)}$ $\quad l = \frac{n-1+W_2}{4}$

- **DNA Fragment Assembly is a technique to reconstruct an original DNA sequence a large number of fragments**

Alba, Luque, Khuri, CEC 2005:57-65

- **Two fitness functions are possible and both define QAP subproblems**

$$F1(x) = \sum_{i=1}^{n-1} w_{x(i)x(i+1)} \qquad r_{ij} = \delta_i^{j-1}$$

$$F2(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} |i-j| \cdot w_{x(i)x(j)} \qquad r_{ij} = |i-j|$$

Conclusions & Future Work

Conclusions

- Elementary landscape decomposition is a useful tool to understand a problem
- The decomposition can be used to design new operators
- We can exactly determine the autocorrelation functions
- We present the decomposition for QAP (permutation representation)

Future Work

- Methodology for landscape decomposition
- Search for additional applications of landscapes' theory in EAs
- Design new operators and search methods based on landscapes' information
- Analyze other problems

Thanks for your attention !!!