ACAT05

May 22 - 27, 2005

DESY, Zeuthen, Germany

Search for the Higgs boson at LHC by using Genetic Algorithms

Mostafa MJAHED

Ecole Royale de l'Air, Mathematics and Systems Dept.

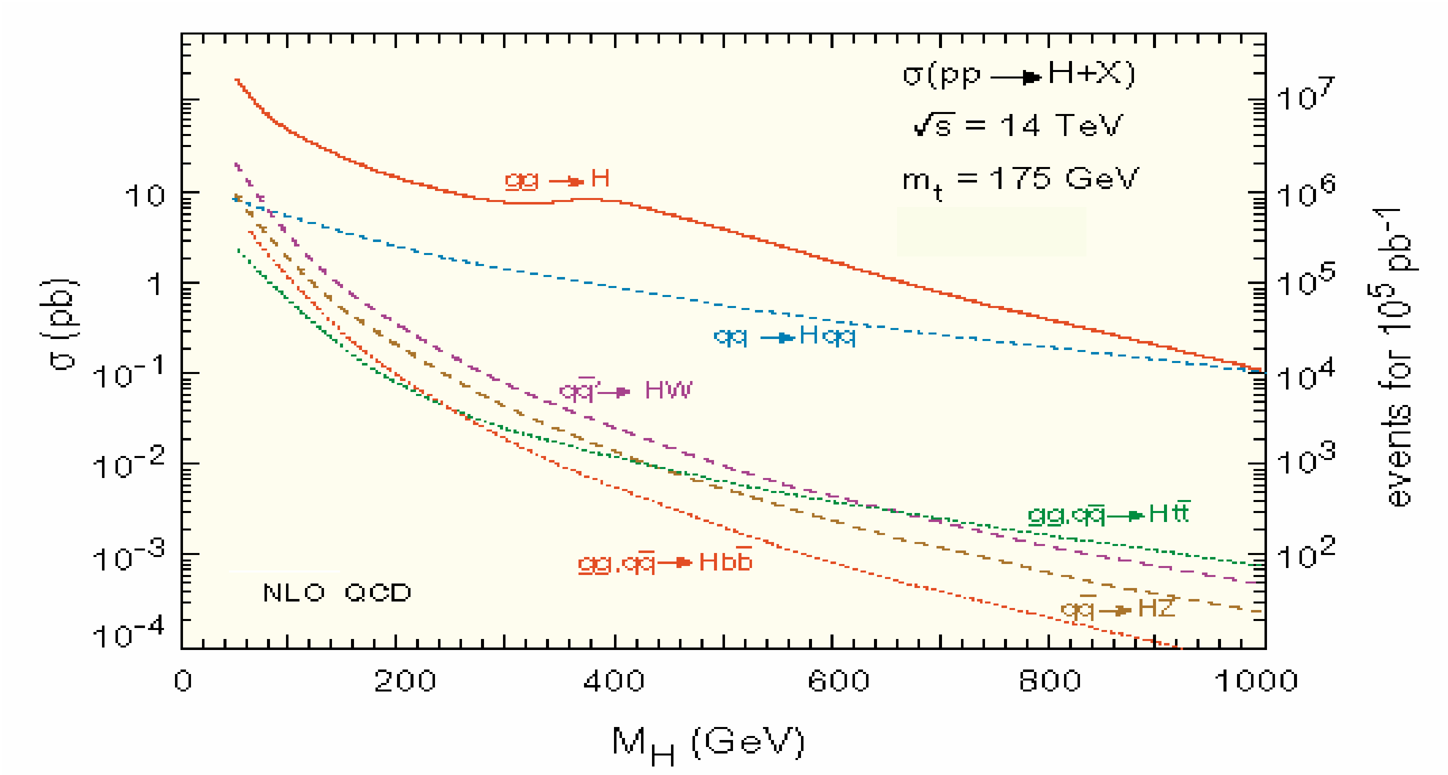
Marrakech, Morocco

Search for the Higgs boson at LHC by using Genetic Algorithms

- Introduction
- Genetic Algorithms
- Search for the Higgs boson at LHC by using Genetic Algorithms
  - Optimization of discriminant functions
  - Optimization of Neural weights
  - Hyperplans search
  - Hypersurfaces search
- Conclusion

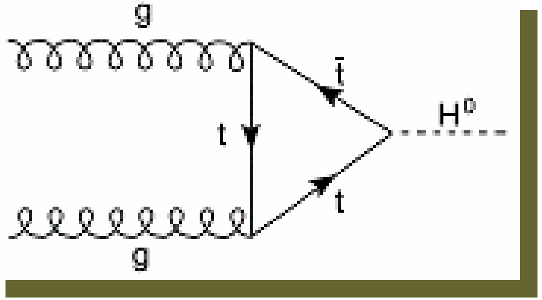
Introduction

Introduction
Higgs production at LHC

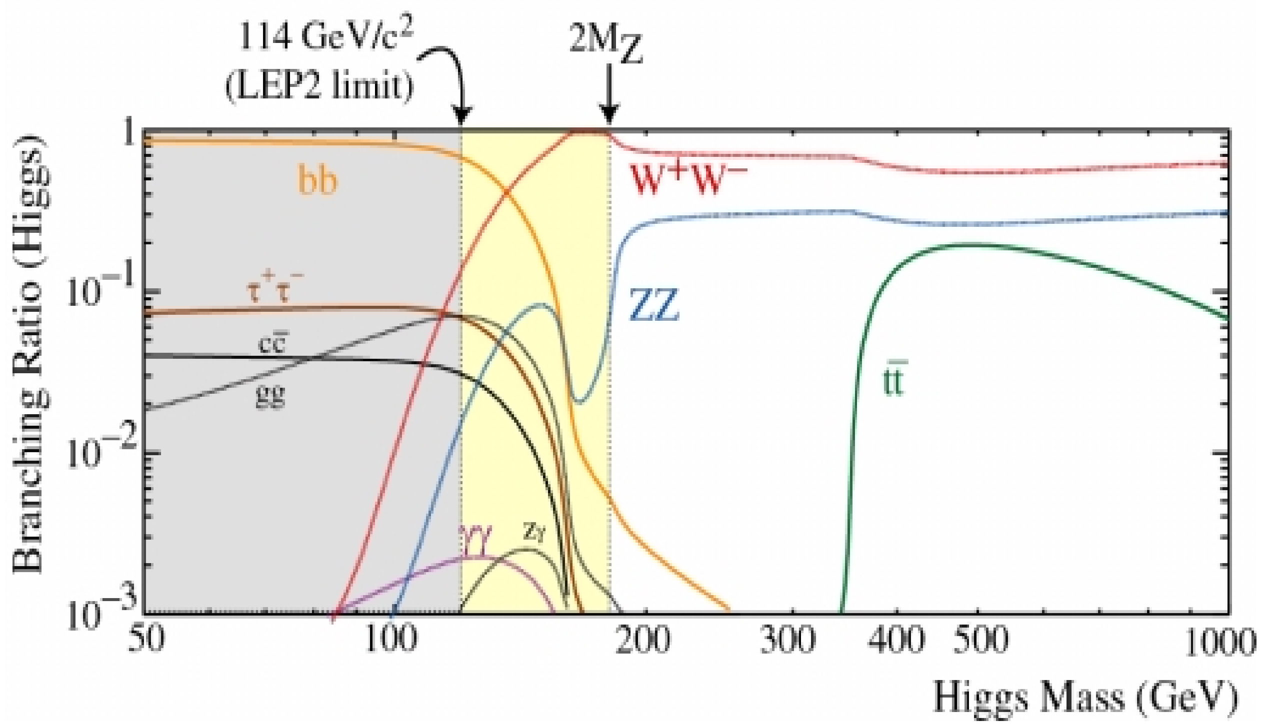


- *Several mechanisms contribute to the production of SM Higgs boson in proton collisions*

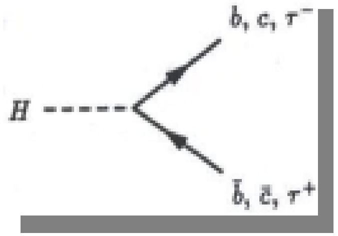
- *The dominant mechanism is the gluon fusion process,* $gg \rightarrow H$

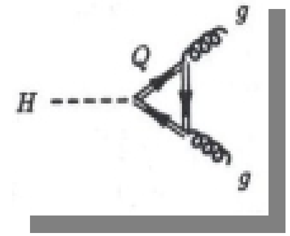


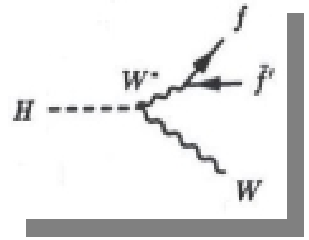
Introduction
Decay Modes



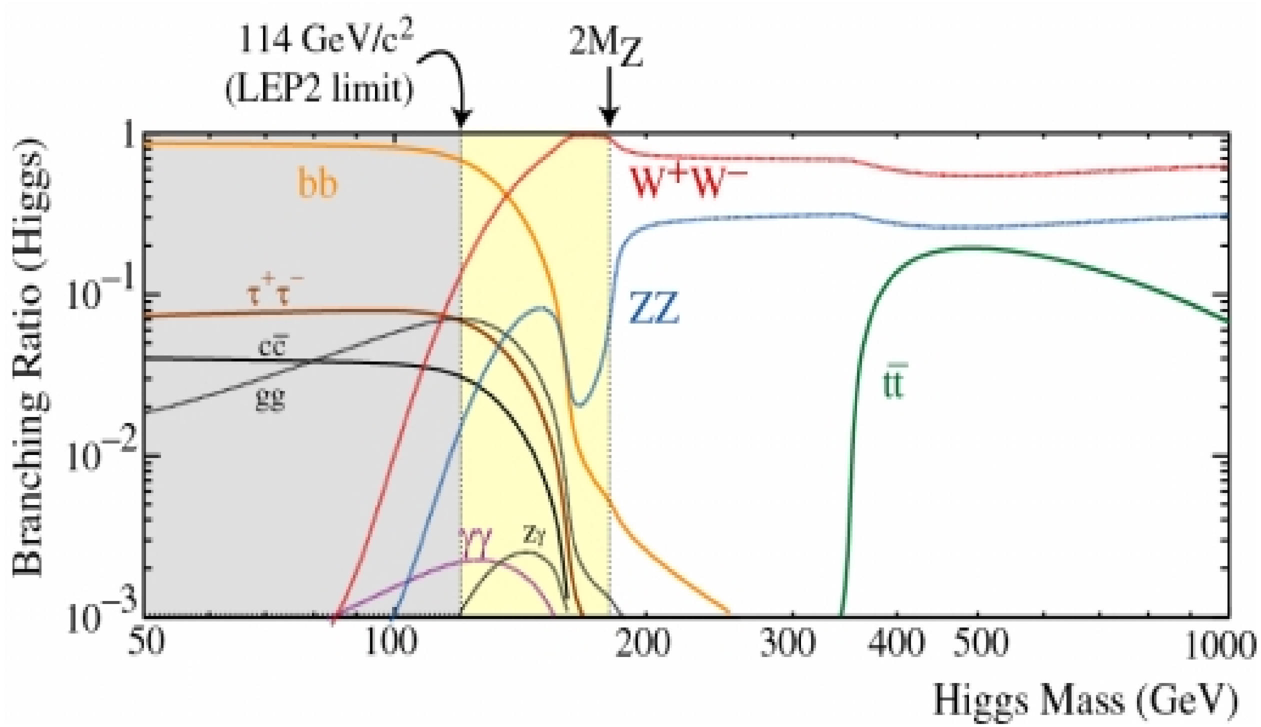
- *decay into quarks:* $H \rightarrow bb$ *and* $H \rightarrow cc$
- *leptonic decay* $H \rightarrow \tau^{+} \tau^{-}$
- *gluonic decay* $H \rightarrow g\, g$
- *decay into virtual W boson pair:* $H \rightarrow W^{+}W^{-}$







Introduction
Main discovery modes



$M_H < 2M_Z$:

- $H \to bb+X$
- $H \to W^+W^- \to l\nu l\nu$
- $H \to ZZ \to 4l$
- $H \to \gamma\gamma$

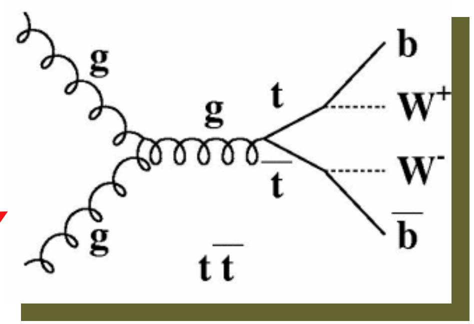
$H \rightarrow W^+W^- \rightarrow ll\nu\nu$

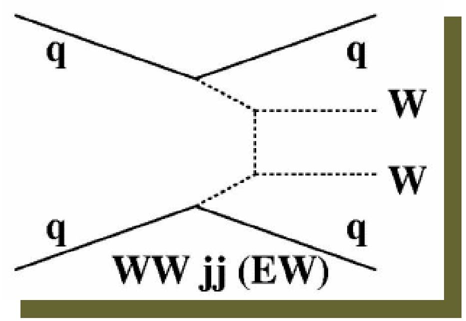
- **The decay channel chosen:**

  $H \rightarrow W^+W^- \rightarrow e^+\mu^-\nu\nu,\ \mu^+e^-\nu\nu,\ e^+e^-\nu\nu,\ \mu^+\mu^-\nu\nu.$

- **Signature:**
  - *Two charged oppositely leptons with large transverse momentum P_T.*
  - *Two energetic jets in the forward detectors.*
  - *Large missing transverse momentum P'_T*

- **Main background:**
  - *tt production: with $tt \rightarrow WbWb \rightarrow l\nu j\, l\nu j$*
  - *QCD W^+W^- +jets production*
  - *Electroweak WWjj*





$H \to W^+W^- \to ll\nu\nu$

Main Variables

- $\Delta\eta_{ll}$, $\Delta\phi_{ll}$: *the pseudo-rapidity and the azimuthal angle differences between the two leptons*

- $\Delta\eta_{jj}$, $\Delta\phi_{jj}$: *the pseudo-rapidity and the azimuthal angle differences between the two jets*

- M_{ll}, M_{jj}: *the invariant mass of the two leptons and jets,*

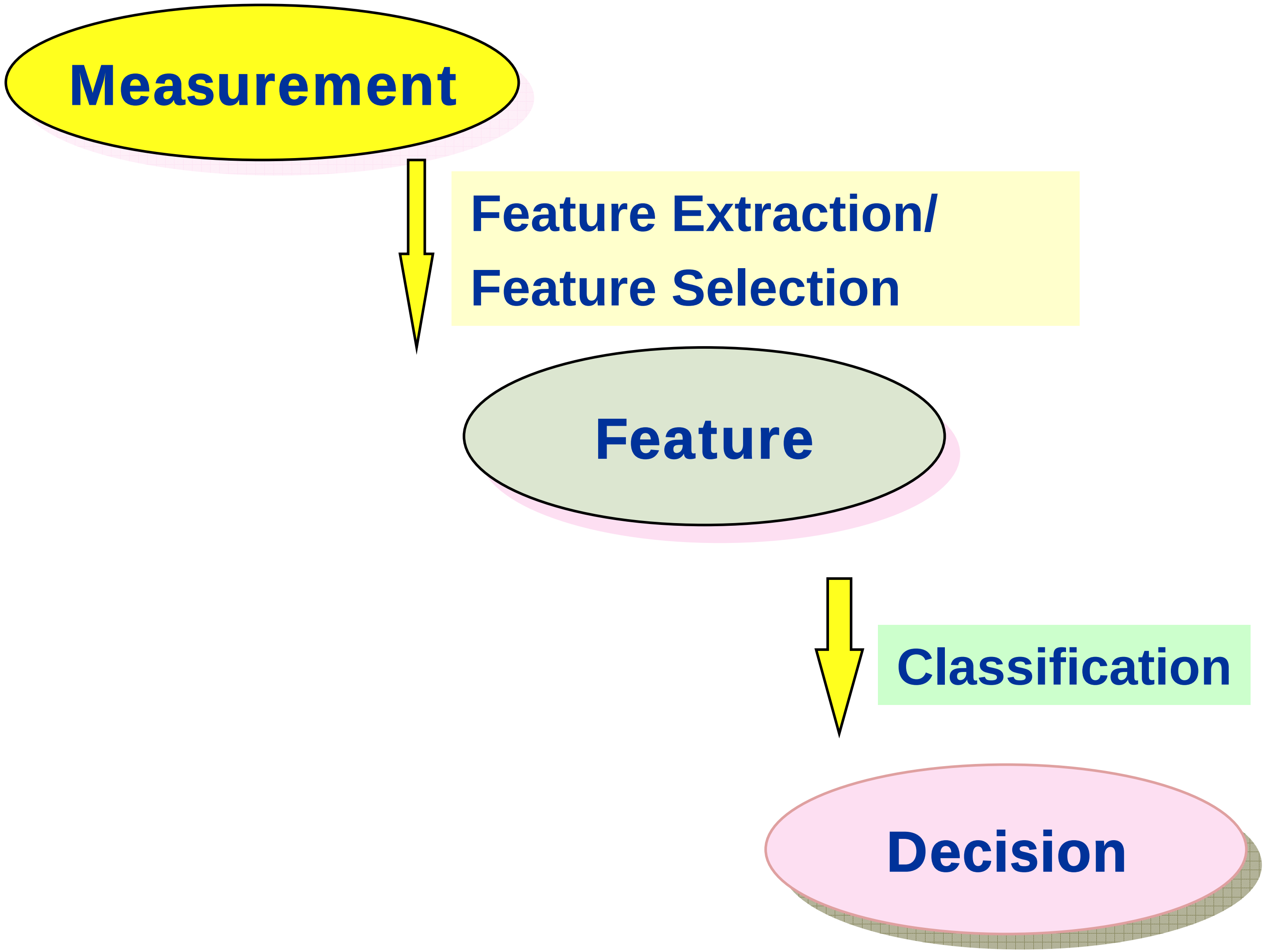
- M_{nm} *(n,m = 1,2,3) some rapidity weighted transverse momentum*

$$M_{nm} = \sum_{i \in event} \eta_i^{\,n} \cdot p_{iT}^{\,m} \qquad n, m = 1, 2,3, \ldots$$

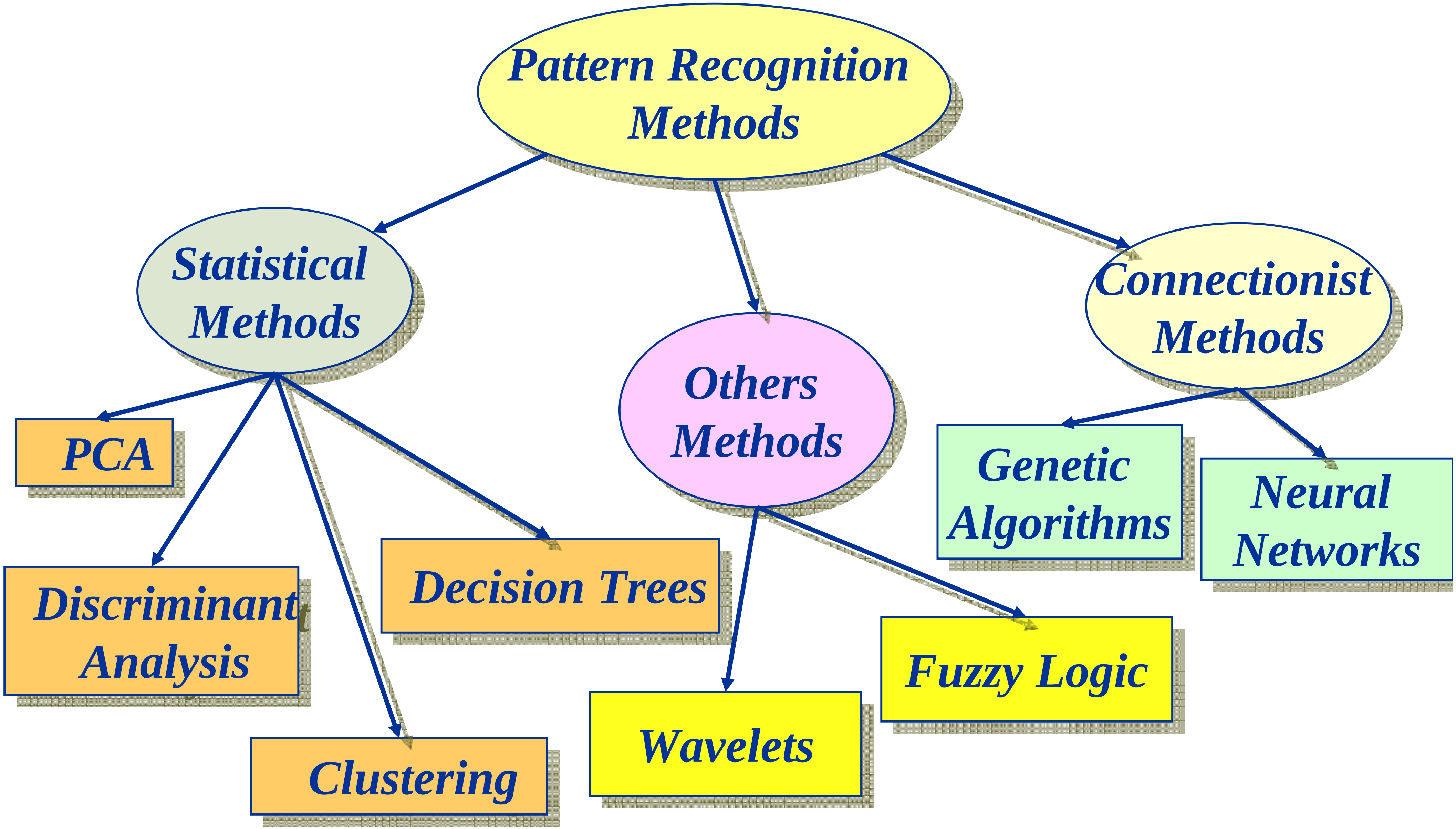
η_i *rapidity of the leptons or jets,* p_{iT} *their transverse momentums.*

Genetic Algorithms

Pattern Recognition



Pattern Recognition Methods



Genetic Algorithms

- Based on Darwin's theory of "survival of the fittest"
  Living organisms reproduce, individuals evolve / mutate, individuals survive or die based on fitness

- The output of a genetic algorithm is the set of "fittest solutions" that will survive in a particular environment

- The input is an initial set of possible solutions

- The process
  - Produce the next generation (by a cross-over function)
  - Evolve solutions (by a mutation function)
  - Discard weak solutions (based on a fitness function)

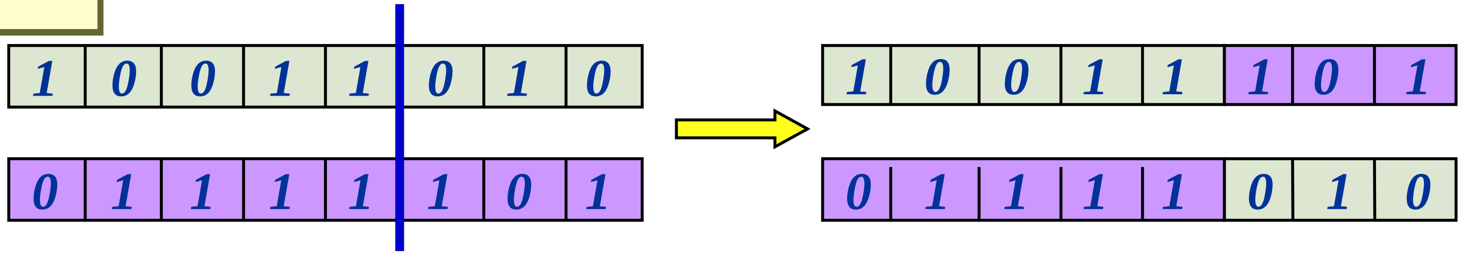
Genetic Algorithms

- **Preparation:**
  - Define an encoding to represent solutions (i. e., use a character sequence to represent a class)
  - Create possible initial solutions (and encode them as strings)
  - Perform the 3 genetic functions to operate on a cluster encoding: Crossover, Mutate, Fitness Test

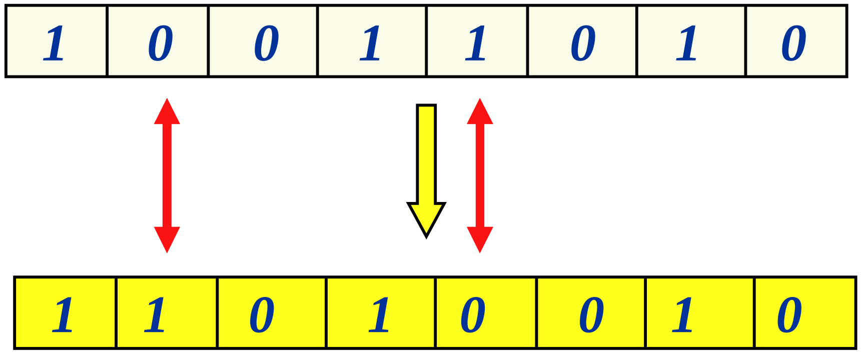
- **Why Genetic Algorithms (GAs) ?**
  - Many real life problems cannot be solved in polynomial amount of time using deterministic algorithm
  - Sometimes near optimal solutions that can be generated quickly are more desirable than optimal solutions which require huge amount of time
  - Problems can be modeled as an optimization one.

Genetic Functions

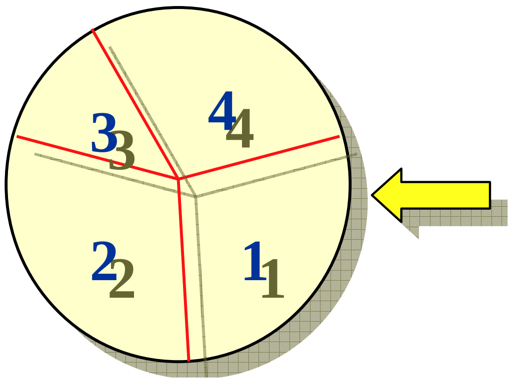
Crossover

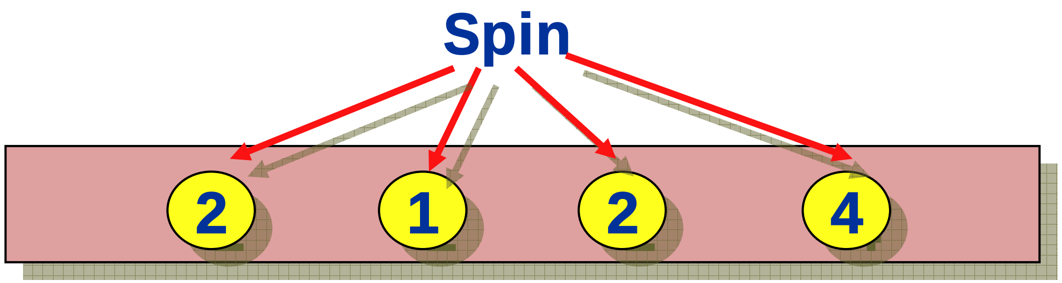


Mutation

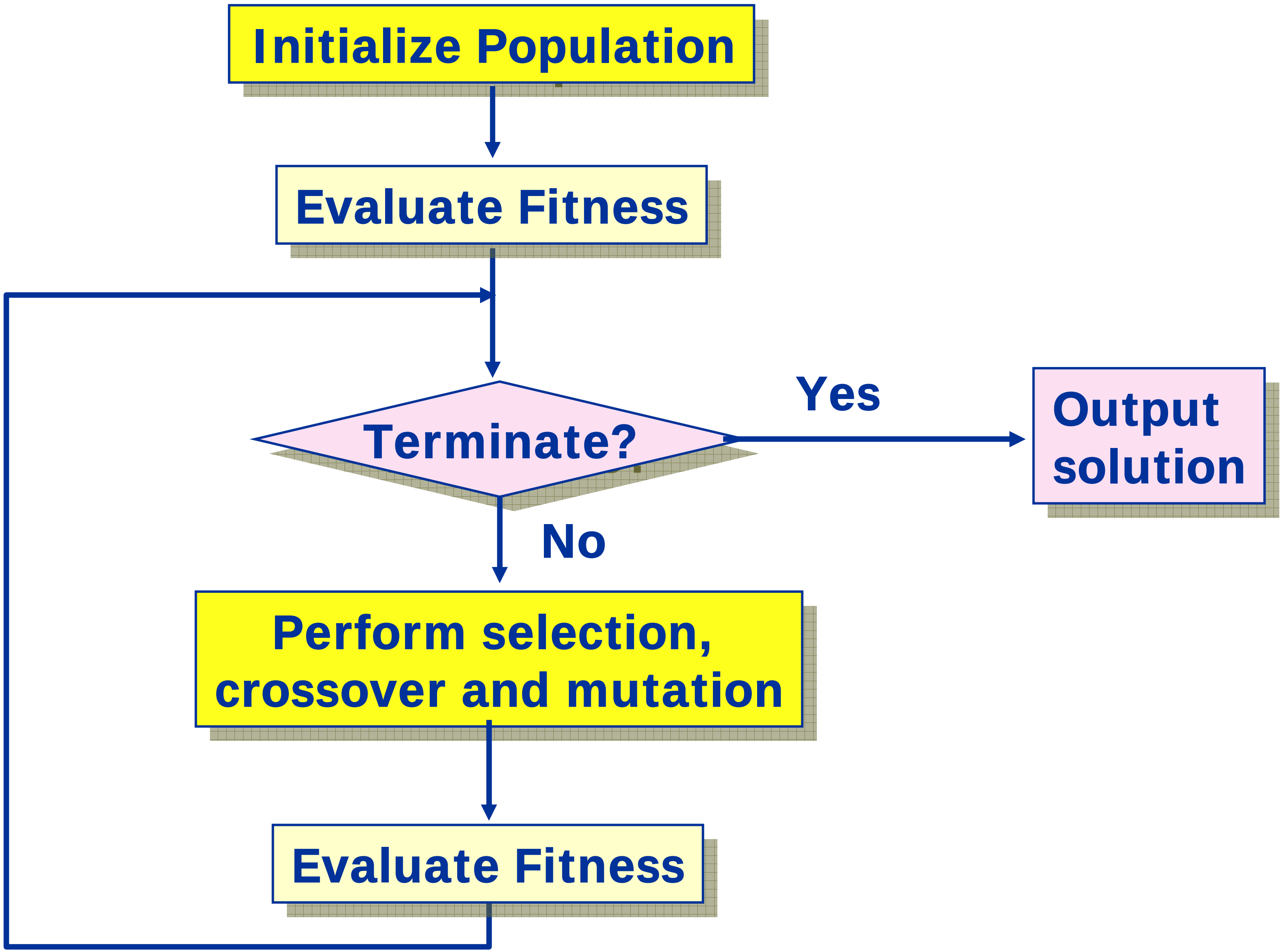


Roulette Wheel Selection





GA Flowchart



GAs for Pattern Classification

- Optimization of discriminant functions
- Optimization of Neural weights
- Hyperplans search
- Hypersurfaces search

Efficiency and Purity of Classification

- **Validation**

Test Events	Classification	
	C_1	C_2
C_1: N_1	N_{11}	N_{12}
C_2: N_2	N_{21}	N_{22}
Total	M_1	M_2

- *Efficiency for C_i classification*

$$E_i = \frac{N_{ii}}{N_i}$$

- *Purity for C_i events*

$$P_i = \frac{N_{ii}}{M_i}$$

- *Misclassification rate for C_i or Error*

$$\varepsilon_i = \frac{N_{ij}}{N_i}$$

Search for the Higgs boson at LHC by using Genetic Algorithms

Search for the Higgs boson at LHC by using Genetic Algorithms

Signal

$pp \to HX \to W^+W^-X \to l^+\nu l^-\nu X$

Background

- $pp \to ttX \to WbWbX \to l\nu j\, l\nu jX$
- $pp \to qqX \to W^+W^-\, X$

- *Events generated by the LUND MC PYTHIA 6.1 at $\sqrt{s}$ = 14 TeV*
- $M_H = 115 - 150$ *GeV/c²*
- *10000 Higgs events and 10000 Background events are used*

- *Research of discriminating variables*

Variables

- $\Delta\eta_{ll}$, $\Delta\phi_{ll}$: *the pseudo-rapidity and the azimuthal angle differences between the two leptons*

- $\Delta\eta_{jj}$, $\Delta\phi_{jj}$: *the pseudo-rapidity and the azimuthal angle differences between the two jets*

- M_{ll} : *the invariant mass of the two leptons*

- M_{jj} : *the invariant mass of the two jets*

- *Rapidity-impulsion weighted Moments* M_{nm} : $(n=1, \ldots, 6)$

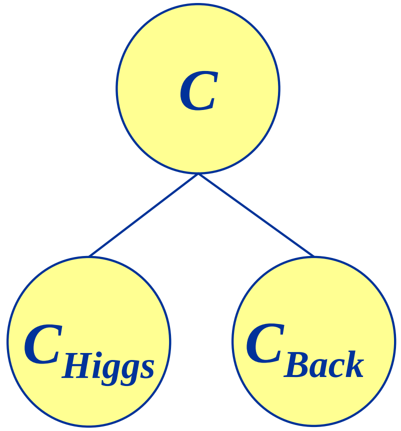
η_i *rapidity:*

$$M_{nm} = \sum_{i \in Jet} \eta_i^{\,n} \cdot p_{iT}^{\ \,m}$$

$$\eta_i = \frac{1}{2}.Log\left(\frac{E_i + p_{i//}}{E_i - p_{i//}}\right)$$

$\Delta\eta_{ll}$, $\Delta\phi_{ll}$, $\Delta\eta_{jj}$, $\Delta\phi_{jj}$, M_{ll} , M_{jj}, M_{11}, M_{21}, M_{31}, M_{41}

Optimization of Discriminant Functions (1)



- **Discriminant Analysis**

- $F(x) = (g_{signal} - g_{back})^T V^{-1} x = \Sigma \alpha_i x_i$

- ***The most separating discriminant Function $F_{Higgs/Back}$ between the classes C_{Higgs} and C_{Back} is :***

$$F_{Higgs/Back} = -0.02 + 0.12\Delta\eta_{ll} + 0.4\Delta\eta_{jj} + 0.35\,M_{ll} + 0.61\,M_{jj} + 0.74M_{11} + 1.04M_{21}$$

- ***The classification of a test event x_0 is then obtained according to the condition:***

$$\text{if } F_{Higgs/Back}(x_o) \geq 0 \text{ then } x_o \in C_{Higgs} \text{ else } x_o \in C_{Back}$$

- **Classification of test events**

Test events	Classification	
	Efficiency	Purity
C_{Higgs}	0.601	0.606
C_{Back}	0.610	0.604

Optimization of Discriminant Functions (2)

- **GA Parameters**
  - Discriminant Function
  - Fitness
  - Number of generations
  - Selection, Crossover, Mutation

-0.02	0.12	0.4	0.35	0.61	0.74	1.01	0.395
α_0^i	α_1^i	α_2^i	α_3^i	α_4^i	α_5^i	α_6^i	ε^I

Fitness Fn: ε
Misclassification rate

Test Events	Classification			
	Eff.	Purity	Eff	ε
C_{Higgs}	0.601	0.606	0.6055	0.3945
C_{Back}	0.610	0.604		

- **GA Code**
  Matlab GA Toolbox

Optimization of Discriminant Functions (3)

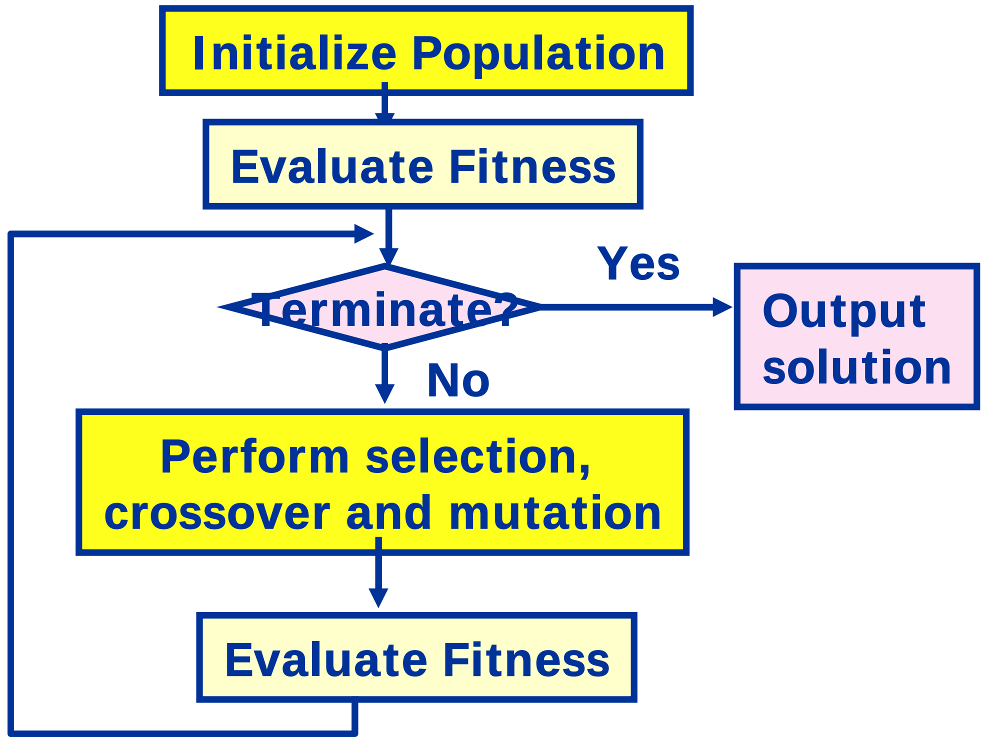
Chromosome 1	α_0^{1}	α_1^{1}	α_2^{1}	α_3^{1}	α_4^{1}	α_5^{1}	α_6^{1}	ε^{1}
....								
Chromosome N	α_0^{N}	α_1^{N}	α_2^{N}	α_3^{N}	α_4^{N}	α_5^{N}	α_6^{N}	ε^{N}

Generation of N solutions

Coefficients of F

Fitness Fn: ε

- Genetic Process



Optimization of Discriminant Functions (4)

• Optimization Results

Number of generations =10000

CPU Time: 120 s

- Optimal Disc. Fn

-0.02	0.12	0.4	0.35	0.61	0.74	1.01		0.35

Test Events	Classification			
	Eff.	Purity	Eff	ε
C_{Higgs}	0.652	0.649	0.65	0.35
C_{Back}	0.648	0.650		

Optimization of Neural Weights (1)

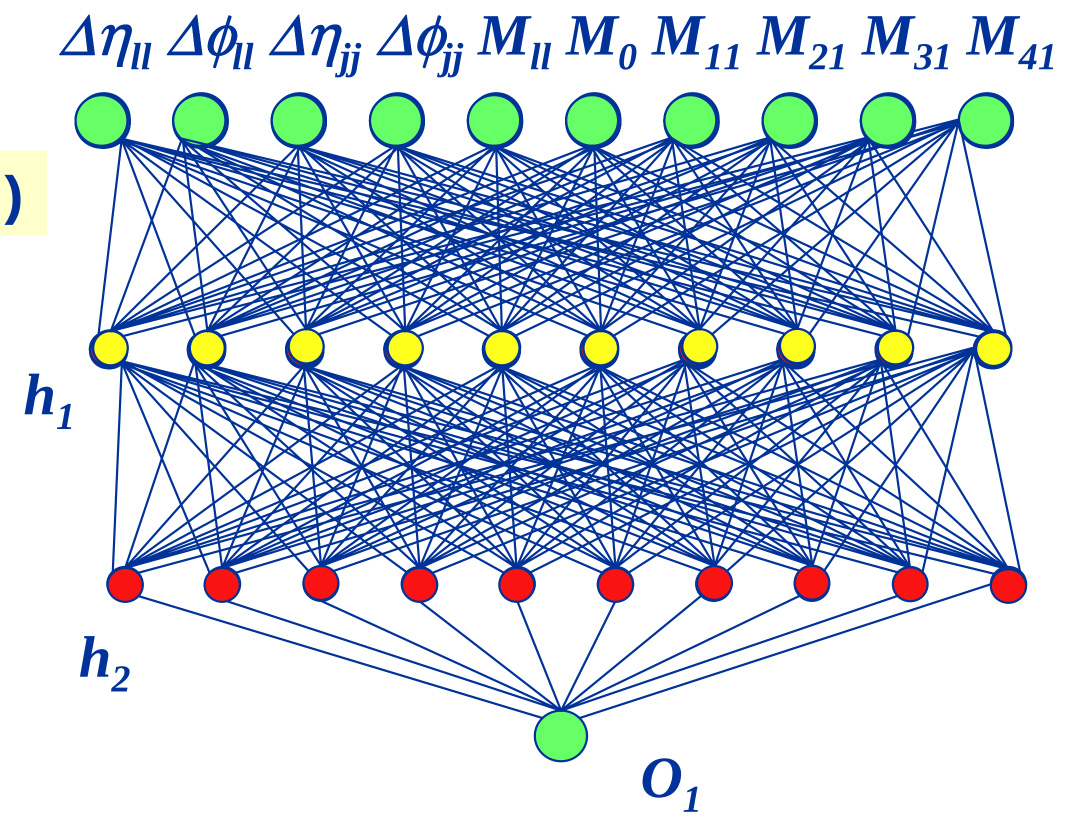
Neural Analysis

NN Architecture: (10, 10, 10, 1)

$$o_1 = \sum_i w_{i1}^{h2o} h_{2i}$$

$$h_{2i} = f\left(\sum_j w_{ji}^{h1h2} h_{1j} - \theta_i\right)$$

$$h_{1i} = f\left(\sum_j w_{ji}^{xh1} x_j - \theta_i\right)$$



Classification of test events

if $O_1(x) \geq 0.5$ then $x \in C_{Higgs}$ else $x \in C_{Back}$

Test Evts	Classification			
	Eff.	Pur.	Eff	ε
C_{Higgs}	0.654	0.663	0.661	0.338
C_{Back}	0.669	0.659		

Optimization of Neural Weights (2)

GA Parameters

- Connection Weights + thresholds

$W_{ij}^{xh1},\ \theta_i^{h1}$	$W_{ij}^{h1h2},\ \theta_i^{h2}$	W_{ij}^{h2o}	ε^I
100 +10	100 + 10	10	ε^I

- Total number of parameters to be optimized : 230
- Fitness: Misclassification rate, ε

Optimization Results

Number of generations =1000

CPU Time: 6 mn

Test Evts	Classification			
	Eff.	Pur.	Eff	ε
C_{Higgs}	0.691	0.696	0.695	0.305
C_{Back}	0.699	0.693		

Hyperplan search (1)

- $H(x) = \alpha_0 + \Sigma\, \alpha_i\, x_i = 0\,, \qquad i=1,\,10$

Hyperplan j

α_0^j	α_1^j	α_2^j	α_3^j	α_4^j	α_5^j	α_6^j	α_7^j	α_8^j	α_9^j	α_{10}^j	ε

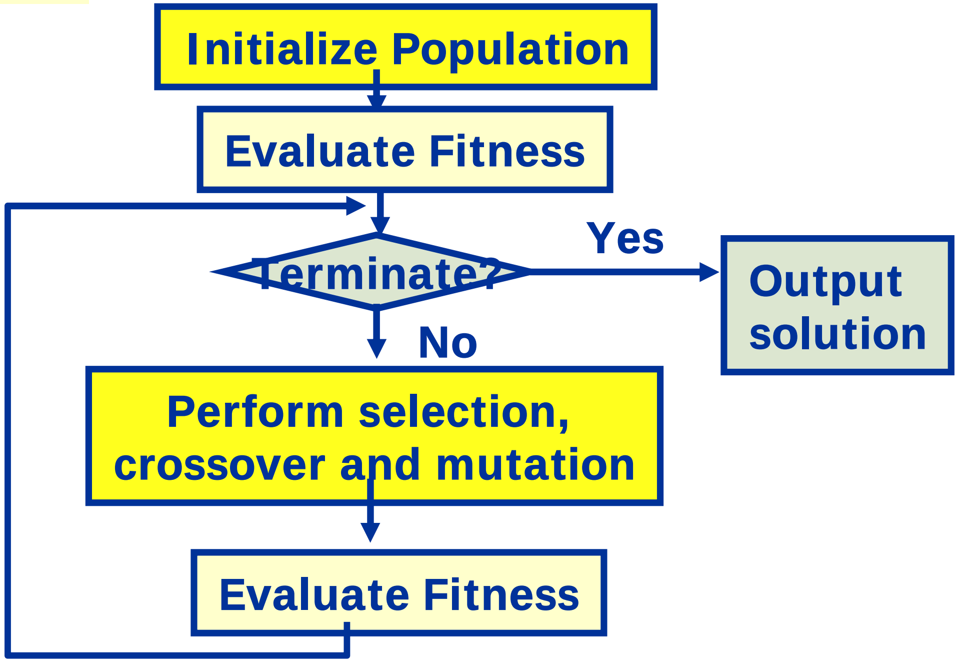
- **11 parameters to optimize**

Classification rule

$if\ H(x) \geq 0 \quad then\ \ x \in C_{Higgs} \quad else\ \ x \in C_{Back}$

Genetic Process

- Generation of N hyperplans H^j, j=1: N=20
- Number of generations =10000
- CPU Time: 4 mn



Hyperplan search (2)

- **Hyperplan search Results**

  - **Classification of test events**

Test Evts	Classification			
	Eff.	*Pur.*	*Eff*	ε
C_{Higgs}	0.661	0.655	0.656	0.344
C_{Back}	0.651	0.657		

  - **Same results than Discriminant functions optimization**

Test Events	Classification			
	Eff.	*Pur.*	*Eff*	ε
C_{Higgs}	0.652	0.649	0.65	0.35
C_{Back}	0.648	0.650		

Hypersurface search (1)

- $$S(x) = \alpha_0 + \sum_{i=1}^{10} \alpha_i\, x_i + \sum_{i=1}^{10} \beta_i\, x_i^{\,2} + \sum_{i=1}^{10} \gamma_i\, x_i^{\,3}$$

Hypersurface j

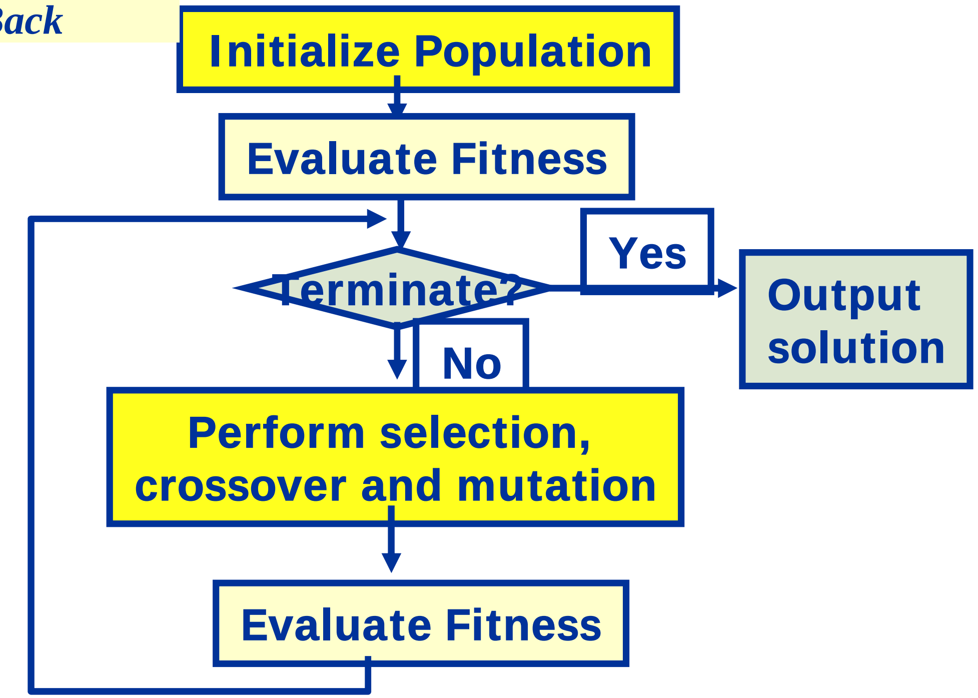
$\alpha_I^{\,j}$ $i=0{:}10$	$\beta_i^{\,j}$ $i=1{:}10$	$\gamma_i^{\,j}$ $i=1{:}10$	ε^j

- **31 parameters to optimize**

- **Classification rule**

$$if\ S(x) \geq 0\ \ then\ \ x \in C_{Higgs}\ \ else\ \ x \in C_{Back}$$

- **Genetic Process**
  - Generation of N hyperplans S^j, $j=1{:}\ N=20$
  - Number of generations =10000
  - CPU Time: 6 mn



Hypersurface search (2)

- **Hypersurface search Results**

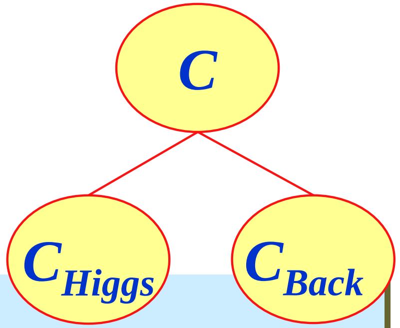
  - **Classification of test events**

Test Evts	Classification			
	Eff.	*Pur.*	*Eff*	*ε*
C_{Higgs}	0.689	0.696	0.691	0.309
C_{Back}	0.693	0.693		

  - **Same results as NN weights optimization**

Test Evts	Classification			
	Eff.	*Pur.*	*Eff*	*ε*
C_{Higgs}	0.691	0.696	0.695	0.305
C_{Back}	0.699	0.693		

Conclusion



Methods

- *Importance of Pattern Recognition Methods*
- *The improvement of an any identification is subjected to the multiplication of multidimensional effect offered by PR methods and the discriminating power of the proposed variables.*

- *Genetic Algorithms Method allows to minimize the classification error and to improve efficiencies and purities of classifications.*
- *The performances are in average 3 to 5 % higher than those obtained with other methods.*

- *Discriminant Functions Optimization : comparative to hyperplan search approach*
- *Neural Weights Optimization : comparative to hypersurface search approach*

Conclusion (continued)

- ***Variables***

Characterisation of Higgs Boson events: Other variables should be examined

- ***Physics Processes***

Other processes should be considered

Detector effects should be added to the simulated events