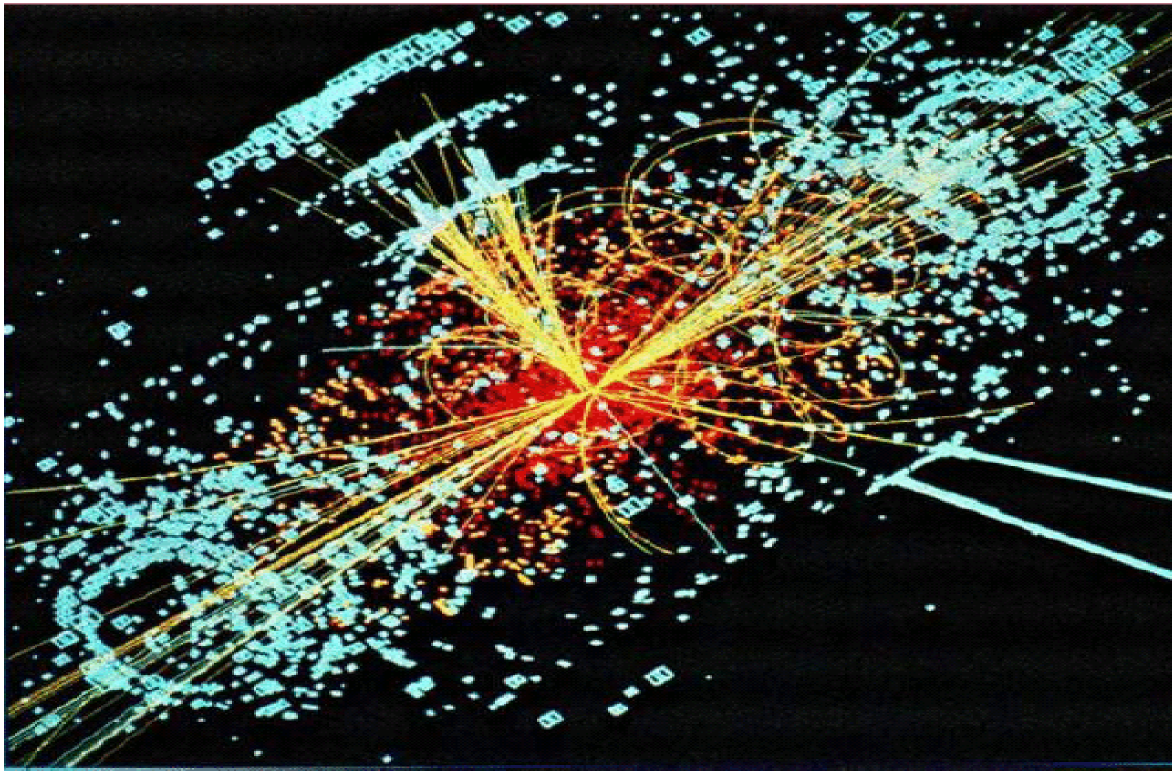
Search for the SM Higgs in the H →WW → eνμν channel in CMS*



Michele Selvaggi (UA)

Outline

- ***The experiment***
  - LHC
  - CMS
- ***Standard Model Higgs***
  - Higgs boson
  - Production and decay modes
- ***$H \rightarrow WW^* \rightarrow e\nu\mu\nu$ channel***
  - Signal topology
  - Backgrounds
  - Selections
    - Cut-based
    - Multivariate
  - Systematics

Higgs Scalar

- *Standard Model = non-Abelian SU(2) x U(1) gauge theory*

  Local gauge invariance ➔ all fields a priori massless

  Massive fields terms (W,Z,e,μ ..) BREAK invariance

- *Higgs Mechanism*

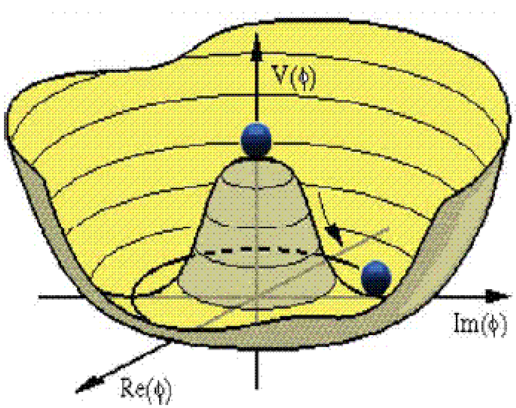
  Postulate the existence of a hidden sector, potential with SU(2) doublet

  Couples to known particles and gives them mass by acquiring v.e.v

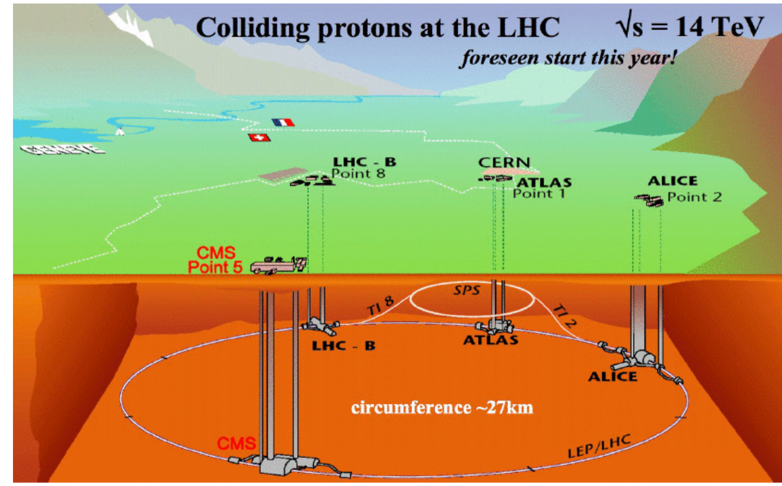
  2 complex components ➔ 4 d.o.f.

  - *3 d.o.f give masses to $W^{\pm}$, Z*
  - *1 d.o.f left scalar neutral particle H ,*

m_H free parameter

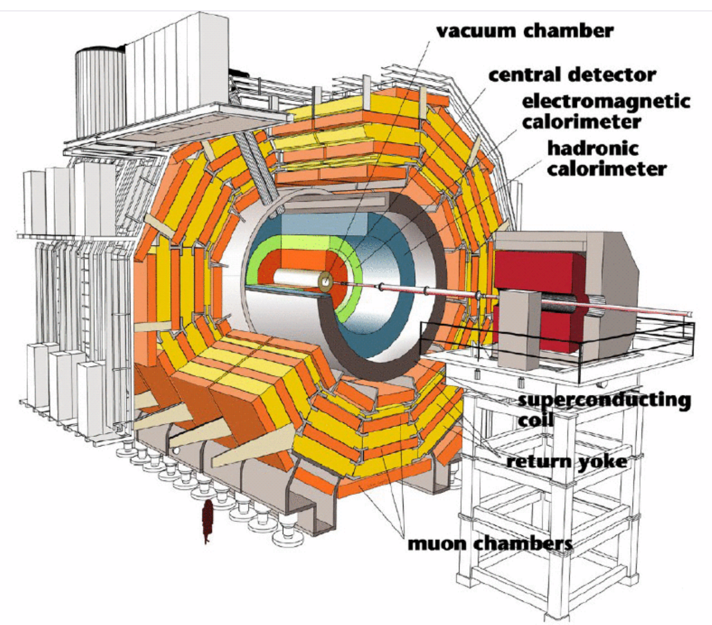


CMS and LHC



Design P-P collisions at $s=14\ (TeV)^2$

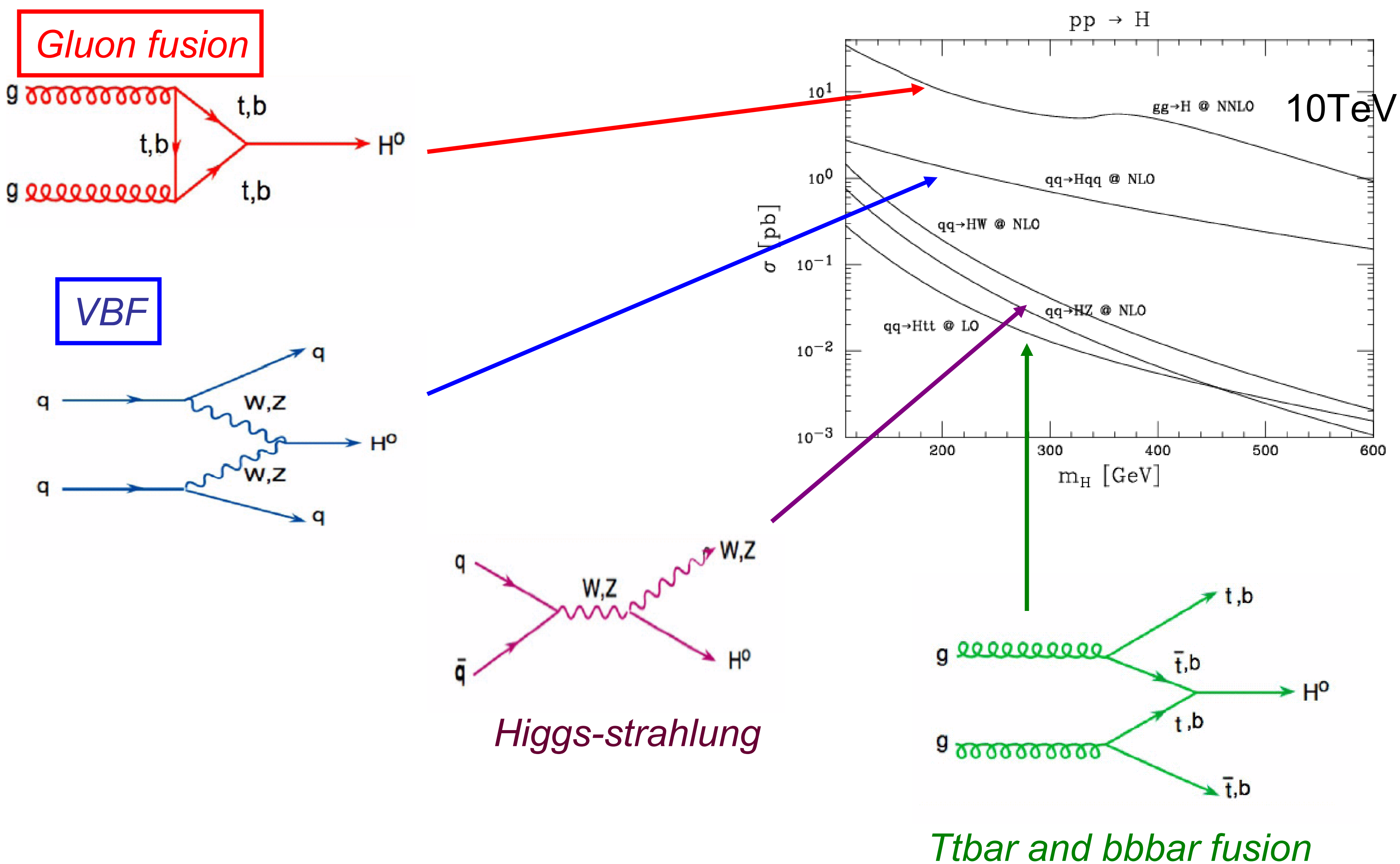
- *1st injection energy mid-november 09*
- *Progressively move to 3.5 TeV beam*
- *In 2010 → 5 TeV*
- *Later (??) → 7 TeV*



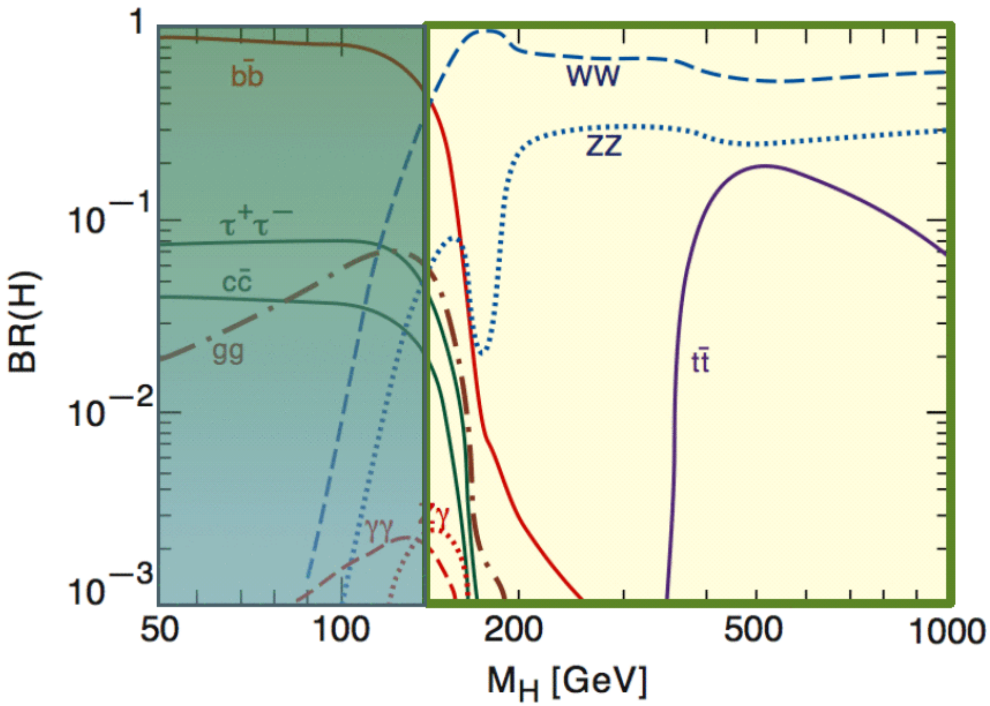
Design properties

- *Good tracking system (SiStrip + pixels)*
- *ECAL (Scintilating $PbWO_4$ crystals)*
- *HCAL (plastic scintilator)*
- *Muons (DT, RPC+CSC)*

Higgs production at LHC



Decay modes



m_H (GeV)	$\sigma_{NLO}(gg)$ (pb)	$\sigma_{NLO}(VBF)$ (pb)	$BR(H \rightarrow WW)$ (%)
120	37	5	13
140	28	4	49
160	22	3	90

Two main scenario

- m_H < 140 GeV

H ➔ bb (BR ~ 85 % but messy)

H ➔ tau tau (cleaner)

- m_H > 140 GeV

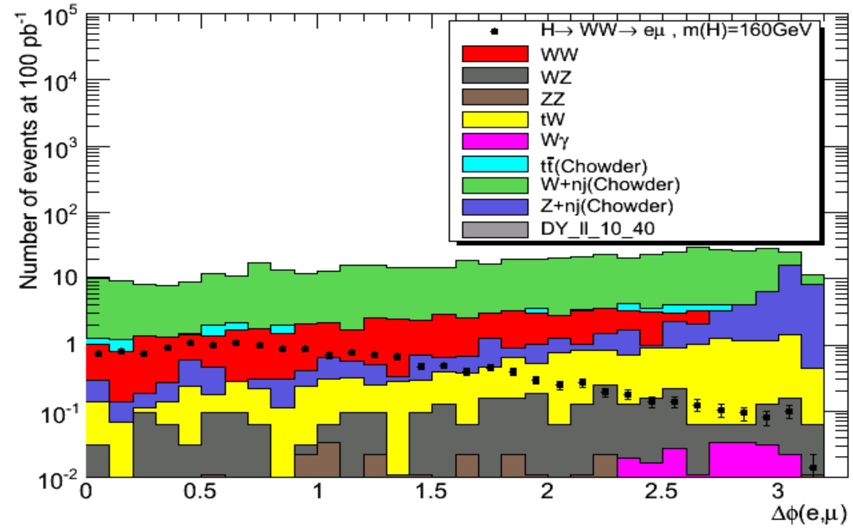
H ➔ WW ➔ 2l2v (dominant but no mass peak)

H ➔ ZZ ➔ 4l (clean mass reconstruction but lower BR)

$H \rightarrow WW \rightarrow e\,\nu\,\mu\,\nu$

Signal Topology

- 2 oppositely charged leptons at central pseudorapidities
- no central jet activity (CJV)
- high Missing E_T
- colinear emitted leptons (spin correlation)



Backgrounds

- *Dominants*: WW, ttbar, W+jets
- also tW, Z+jets, WZ ..

m_H [GeV]	σ^{NLO} [pb]	$BR(H \rightarrow WW)$	$\sigma^{NLO} \times BR(H \rightarrow WW \rightarrow \ell\nu\ell\nu)$ [pb]
120	40.31	0.13	0.56
130	35.08	0.28	1.06
140	31.02	0.48	1.58
150	27.55	0.68	1.98
160	24.63	0.89	2.34
170	22.18	0.96	2.26
180	20.17	0.93	1.99
190	18.43	0.77	1.51
200	16.76	0.73	1.30

Process	$WW \rightarrow \ell\ell'\nu\bar{\nu}$	$t\bar{t}$	$W \rightarrow \ell\nu$	$DY \rightarrow \ell\ell$	Wt	ZZ	WZ	$W+\gamma$
σ^{NLO} [pb]	12.1	840.0	19518.3 × 3	1957.9 × 3	62.0	15.3	49.9	4.7

Cut-based analysis

Strategy

- candidate events have to pass trigger and common skim with at least two reconstructed leptons with p_T > 10 GeV
- select exactly 2 leptons
    - leptons i and j should have opposite electric charge sign
    - $|\eta(\text{lepton}_i)| \leq 2.5$ and $|\eta(\text{lepton}_j)| \leq 2.5$
    - $p_\text{T}(\text{lepton}_i) \geq 10$ GeV and $p_\text{T}(\text{lepton}_j) \geq 10$ GeV
- Missing E_T > 30 GeV
- m_{ll} > 12 GeV
- Central Jet Veto

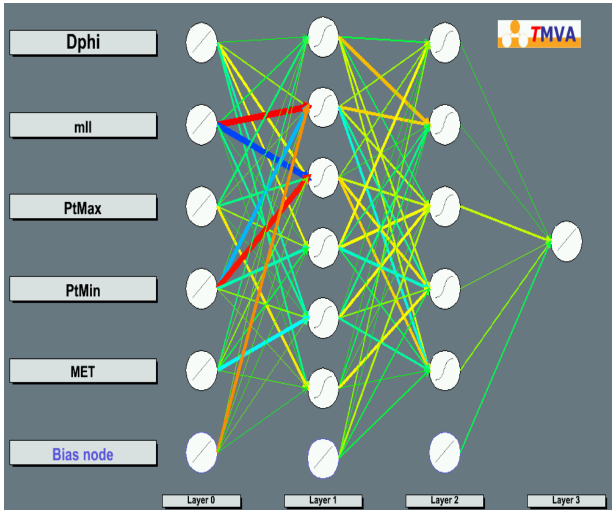
m_H	[GeV]	120	130	140	150	160	170	180	190	200
min E_T^miss	[GeV]	35	35	30	45	45	45	55	65	65
max E_T^miss	[GeV]	60	60	75	105	105	105	105	90	90
max $\Delta\phi_{\ell\ell}$	[deg.]	90	90	70	70	70	65	65	60	65
max $m_{\ell\ell}$	[GeV]	40	40	50	45	45	50	60	80	95
min $p_\text{T}^{\ell,\text{min}}$	[GeV]	15	15	15	15	25	25	20	10	15
min $p_\text{T}^{\ell,\text{max}}$	[GeV]	20	20	25	30	25	25	30	35	35
max $p_\text{T}^{\ell,\text{max}}$	[GeV]	65	65	50	50	50	55	65	70	80
ϵ_S	[$\times 10^{-2}$]	1.0	1.0	1.3	0.9	1.2	1.3	1.3	0.8	0.9
N_S		5.4	10.2	20.8	18.7	28.5	30.0	25.9	12.1	11.7
N_B		77.7	77.7	101.7	46.3	30.6	38.0	51.7	42.5	47.9

NN analysis (1)

Strategy

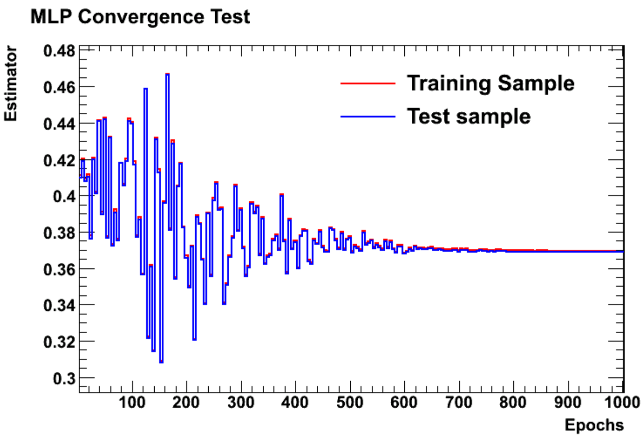
For each mass 130 GeV< m_H < 200 GeV split signal and background tree into two samples with same amount of events:
- training sample
- test sample

➔ **Independent** samples

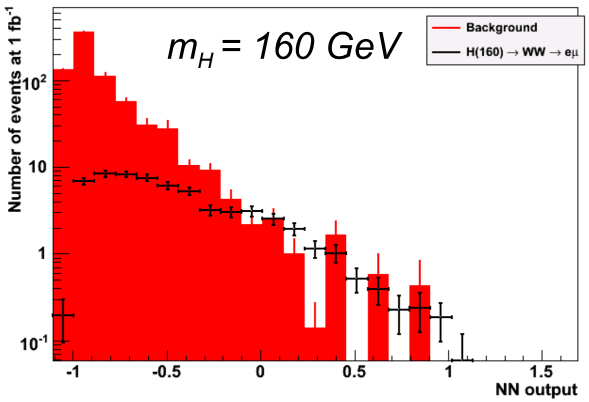


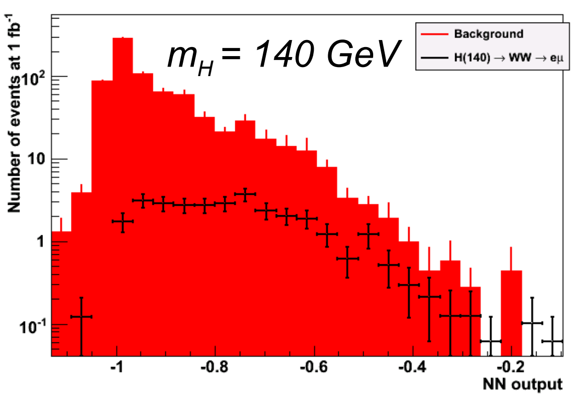
Use 5 discrimination variables ($\Delta\varphi$, Missing E_T, Pt_{Min}, Pt_{Max}, m_{ll})

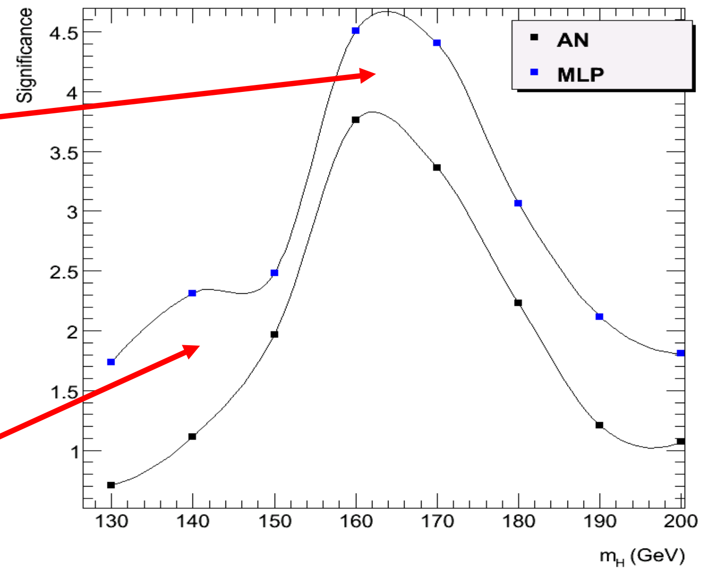
- The structure of the MLP classifier is selected.
- Train the NN with training sample ➔ fix the weights associated to nodes
- Test sample to check that no overtraining has occurred
- Let some MLP parameters vary to check if the configuration is (almost) stable (learning rate, decay rate, number of cycles, ordering input variables)



NN analysis (2)







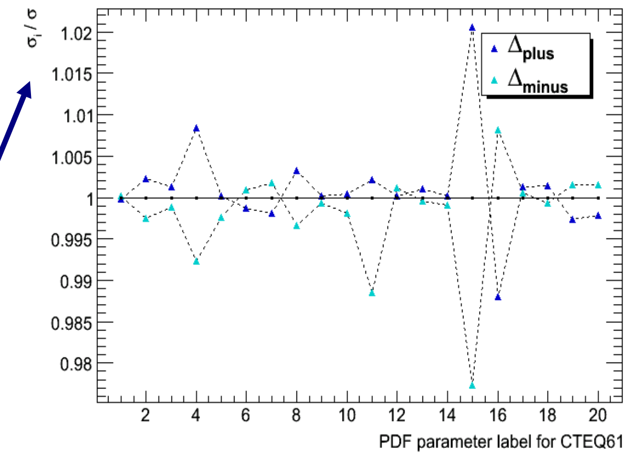
Assuming 20 % uncertainty on backgrounds

	$m_H[GeV]$							
$H \to WW \to \ell^\pm \ell^\mp$	130	140	150	160	170	180	190	200
N_S	14.7	10.5	18.2	23.6	31.8	21,6	19.9	17.4
N_B	31.6	13.4	26.2	16.5	25.7	24.9	36.1	37.1
$\frac{N_S}{\sqrt{N_B+0.04N_B^2}}$	1.74	2.31	2.48	4.51	4.40	3.06	2.12	1.81

PDF uncertainties

Crucial to study the propagation of PDF uncertainties to physical observables such as cross sections, selection efficiencies .

$$\sigma_{had} = \sum_{i,j} \int f_m(x_1, Q) f_p(x_2, Q) dx_1 dx_2 \sigma_{part}$$



Use reweighting

$$f_0(x, Q) \quad , \quad f_{k^\pm}(x, Q) \qquad 1 \leq k \leq D$$

$$w^i_{k\pm} = \frac{f_{k^\pm}(x^i_1, f^i_1, Q^i) f_{k^\pm}(x^i_2, f^i_2, Q^i)}{f_0(x^i_1, f^i_1, Q^i) f_0(x^i_2, f^i_2, Q^i)} \qquad \sigma^\pm_k = \sum_{i=1}^{N_{events}} w^i_{k\pm}$$

Calculate varied quantities and propagate

$$\Delta X^\pm = \sqrt{\sum_{k=1}^{D} \left[max(X^\pm_k - X_0, X^\mp_k - X_0, 0)\right]^2}$$

PDF Uncertainties for backgrounds					
Sample	Cross Section (pb)	CTEQ 6.1	CTEQ 6mE	CTEQ 6.5	CTEQ 6.6
$t\bar{t}$	375	+5.1 % -5.0 %	+4.9 % -4.6 %	+5.1 % -4.4 %	+4.9 % -4.6 %
WW	7.4	+3.7 % -4.7 %	+3.6 % -4.1 %	+3.0 % -3.1 %	+3.2 % -3.2 %
W+jets	40000	+4.1 % -5.0 %	+3.5 % -4.1 %	+3.4 % -3.5 %	+3.2 % -3.3 %