Statistical analysis for H -> ZZ* ->4l

By

Reham Aly – Nicola De Filippis – Ahmed Abdelaliem

30/7/2014

FP7-EENPP, Coordination meeting

For

WP1 group.

Outlines

- Aim of work
- What is Data cards
- Steps of working
- Next Step

Aim of work

- The purpose to produce statistical analysis like producing Exclusion Limit Plots & Significance of excess (P-value Test) & Spin analysis &…… for High mass higgs boson search

By

- preparing Data cards
  - Firstly for the Legency paper to validate our work (To check if we get the same plots).
  - Secondly for high mass search (Our Target).

-Statistical analysis is the last step of analysis since we collect the final number of events after full selection from data- Signal - background for different decay channels (4mu – 4e – 2e2mu).

Strategy

- Data : integrated luminosity of 5.1fb-1 at 7TeV & 19.7 fb-1 at 8 TeV.
- Signature: 4L (4mu – 4e – 2e2mu) final state.
- Signal :
- Monte Carlo samples for Fall 11(2011) & Summer 12 (2012) in the full range of SM search (we examine 187 hypothetical Higgs boson masses mH in a range between 110GeV and 1000GeV)
- Background :
  - irreducible ZZ background (ggZZ & qqZZ) .
  - reducible background Zjets.

- To exploit all the properties of the resonance -> multi-dimensional fit is implemented (3D fit).
  1. The four-lepton mass m4l
  2. The kinematical discriminant KD
  3. The production mode discriminant Di-jet tagged (untagged) category.

What is DataCards

```
imax 1
jmax 7
kmax *
------------
shapes * * hzz4l_4muS_8TeV_0.input.root w:$PROCESS
------------
bin a1_0
observation 13
------------
## mass window [105.7,140.7]
bin a1_0 a1_0 a1_0 a1_0 a1_0 a1_0 a1_0 a1_0
process ggH qqH WH ZH ttH bkg2d_qqzz bkg2d_ggzz bkg2d_zjets
process -4 -3 -2 -1 0 1 2 3
rate 1.0000 1.0000 1.0000 1.0000 1.0000 7.2035 0.1360 0.7119
------------
lumi_8TeV lnN 1.044 1.044 1.044 1.044 1.044 1.044 1.044 -
pdf_gg lnN 1.0720 - - - 1.0780 - 1.0710 -
pdf_qqbar lnN - 1.0270 1.0350 1.0349 - 1.0342 - -
pdf_hzz4l_accept lnN 1.02 1.02 1.02 1.02 1.02 - - -
QCDscale_ggH lnN 1.0750 - - - - - -
QCDscale_qqH lnN - 1.0020 - - - - -
QCDscale_VH lnN - - 1.0040 1.0155 - - - -
QCDscale_ttH lnN - - - - 1.0655 - - -
```

Number of channels

Number of processes -1

Number of sources of systematic uncertainties

Label of the channel

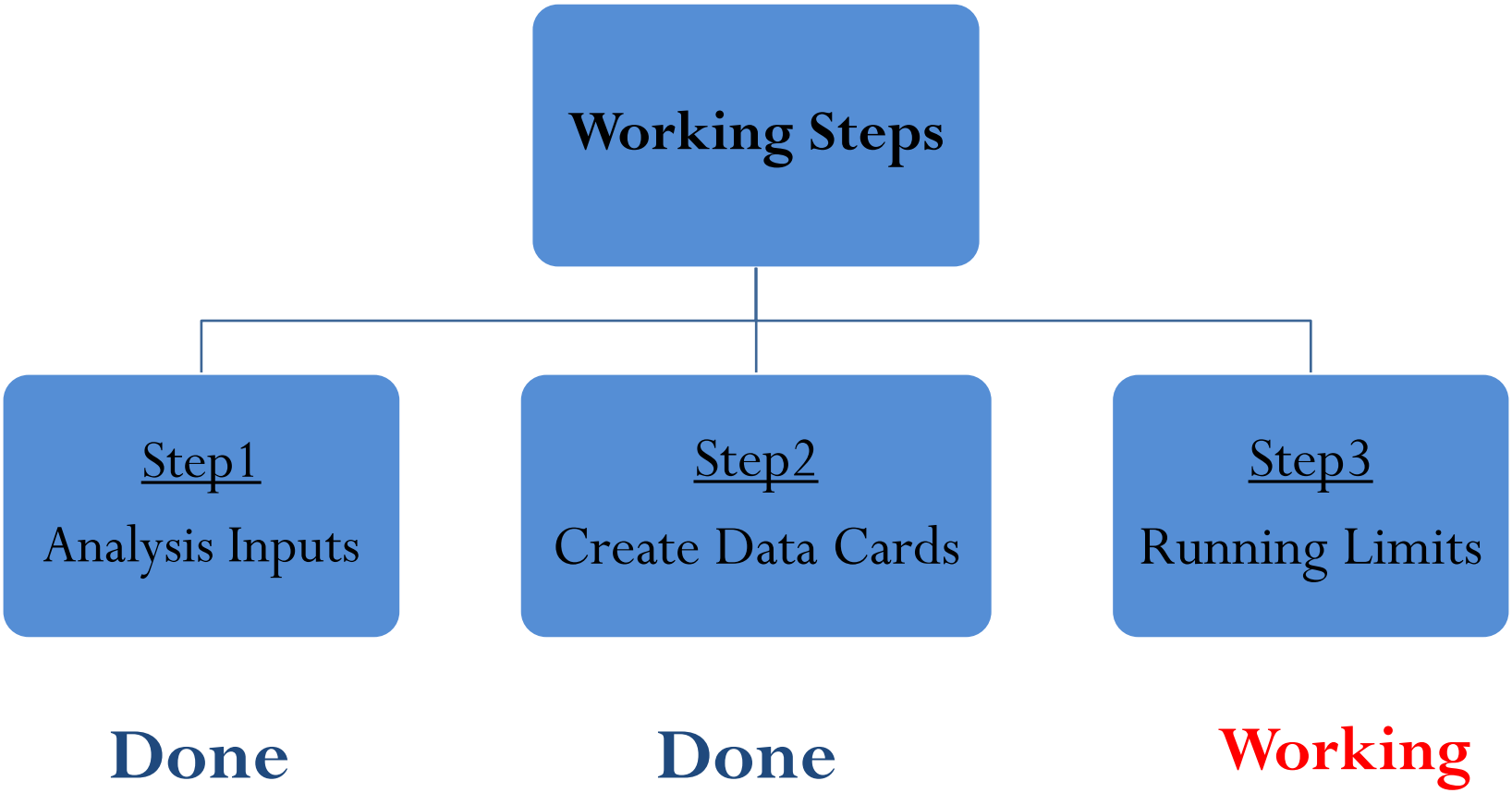
Number of observed events in these channel

Processes Signal & Bkg

Event yield in each process

Systematic uncertainty

Working steps



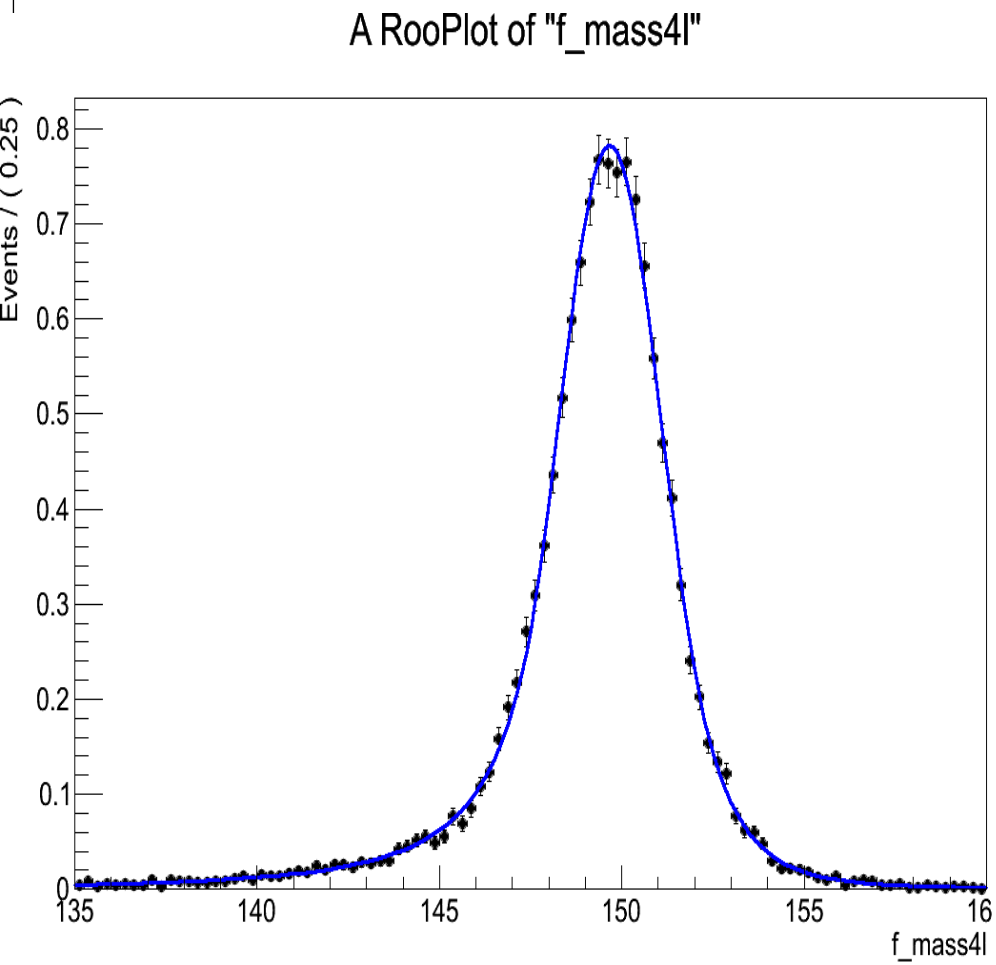
Step1 : Analysis Inputs

Take the inputs from HZZ4L analysis and convert to datacards inputs .

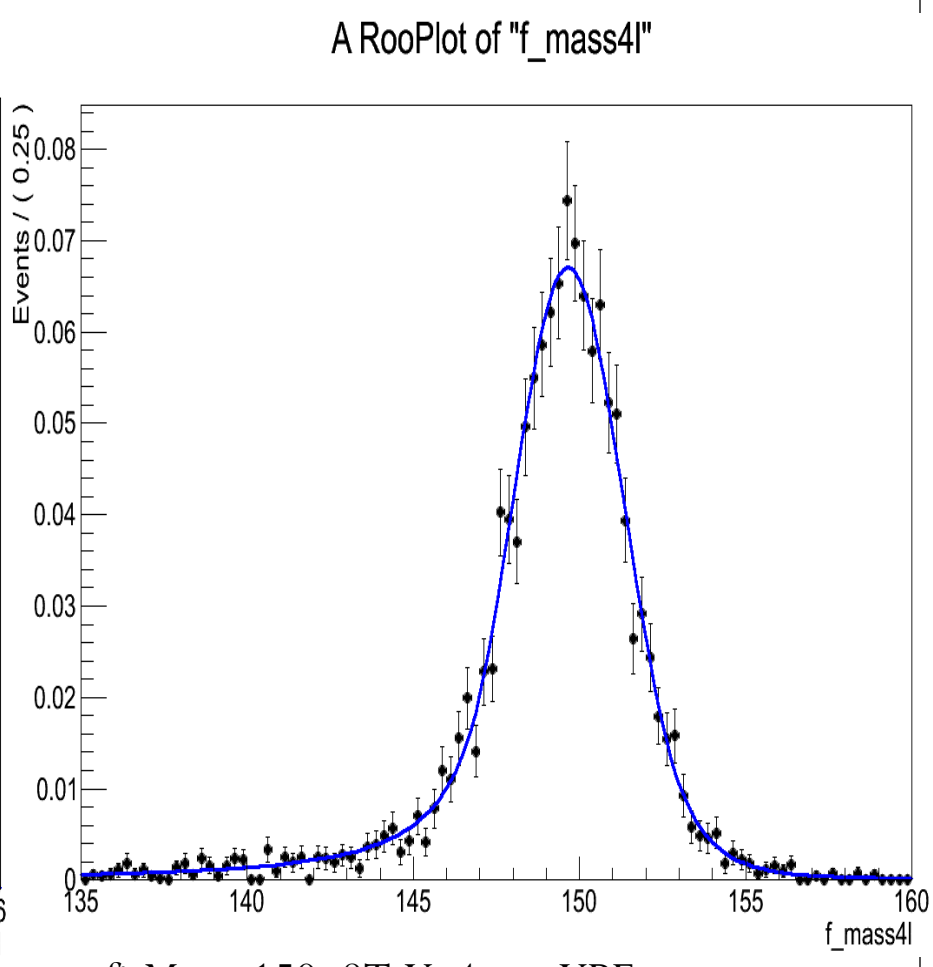
- Prepare data for 7&8 TeV (event yields after full selection)
- ZZ background rates (qqZZ&& ggZZ)- Zjets Rates
- Signal and bkg fitting (Next slide)
- Systematic uncertainties (official numbers)
- Merge these fragments in one file.

Output: Two txt File for each channel one for Dijet tagged and the other for untagged categories .

Signal Fitting

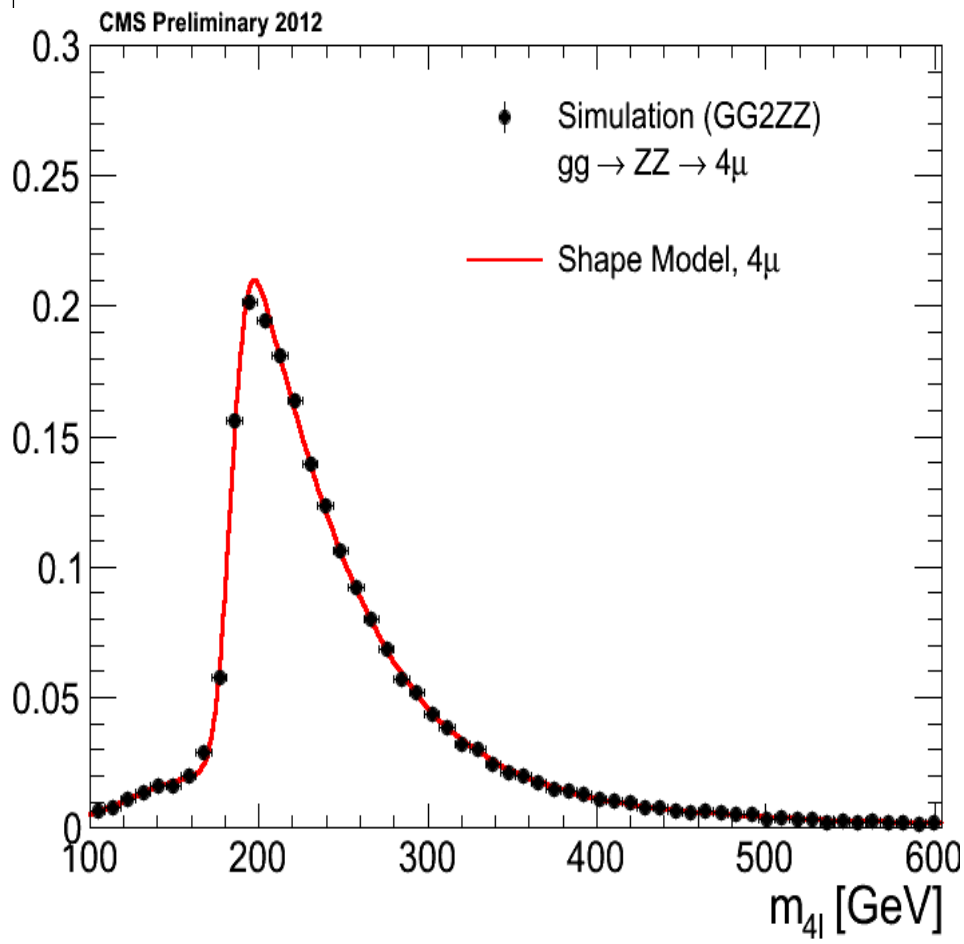


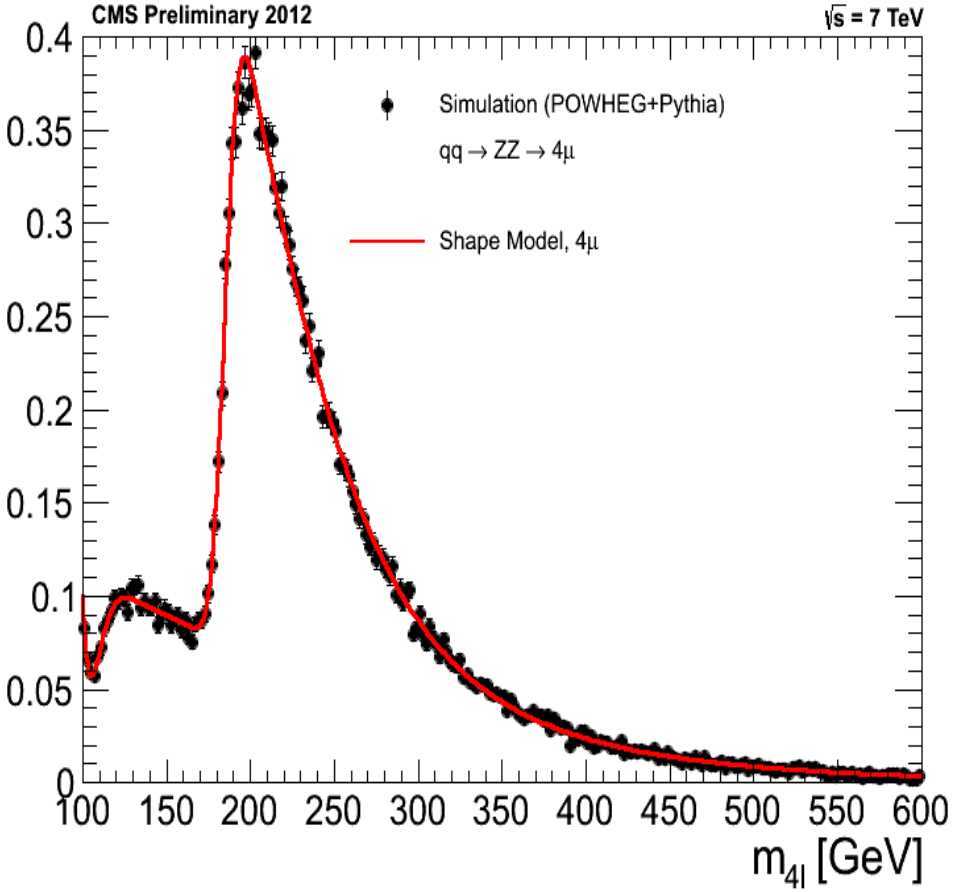
fitMass_150_8TeV_4mu_ggH



fitMass_150_8TeV_4mu_VBF

Background Fitting





Step2 : Create DataCards

- I prepared the datacards for 7 & 8 TeV for all 187 mass points (110-1000) GeV.

In each mass point I have three channels (4mu – 4e – 2e2mu) for each channel the are two file corresponding to diJet tagging (njets>=2) and untagged (njets<2) and the corresponding workspace for all these txt files (Root files).

- Combine all the txt files in one file and root files also in all the mass point to pass to the third step Running limits .

Step3 : Running Limits

- Exclusion limits "Characterizing the absence of the signal"

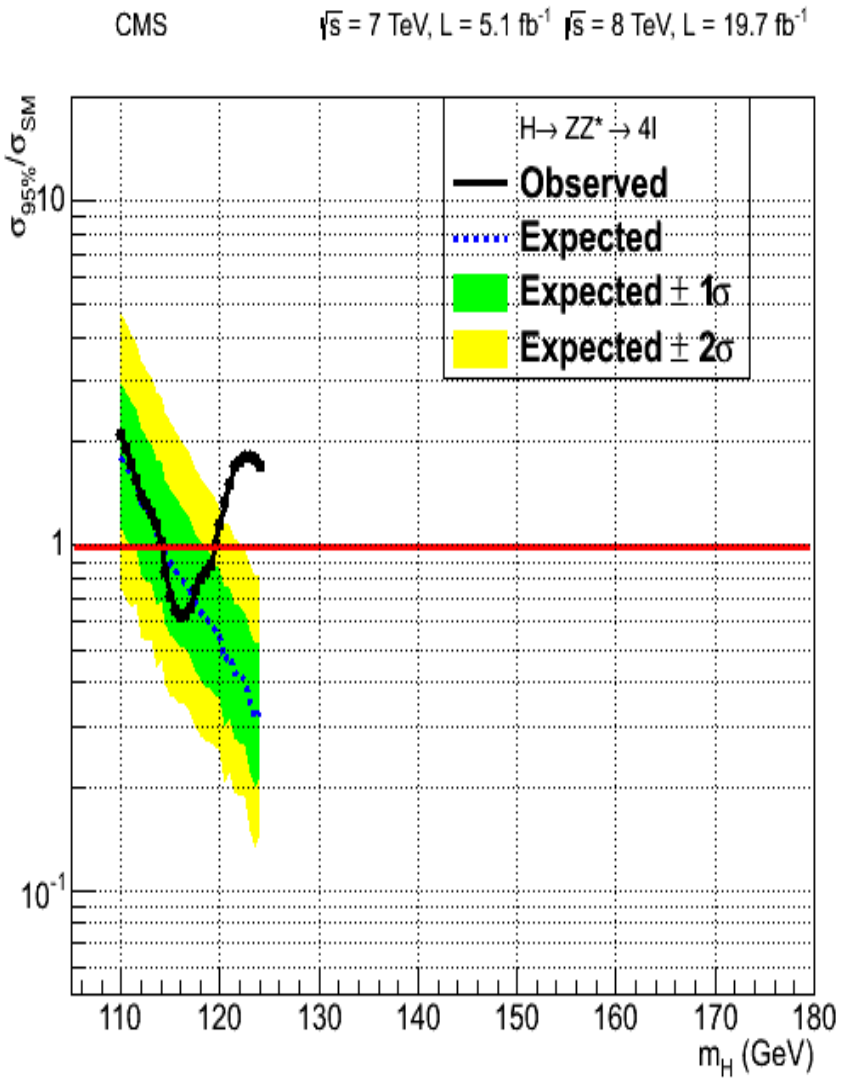
"It gives the probability that higgs exist ruled out given my data"

- we are using the modified frequentist criterion (CLs method)

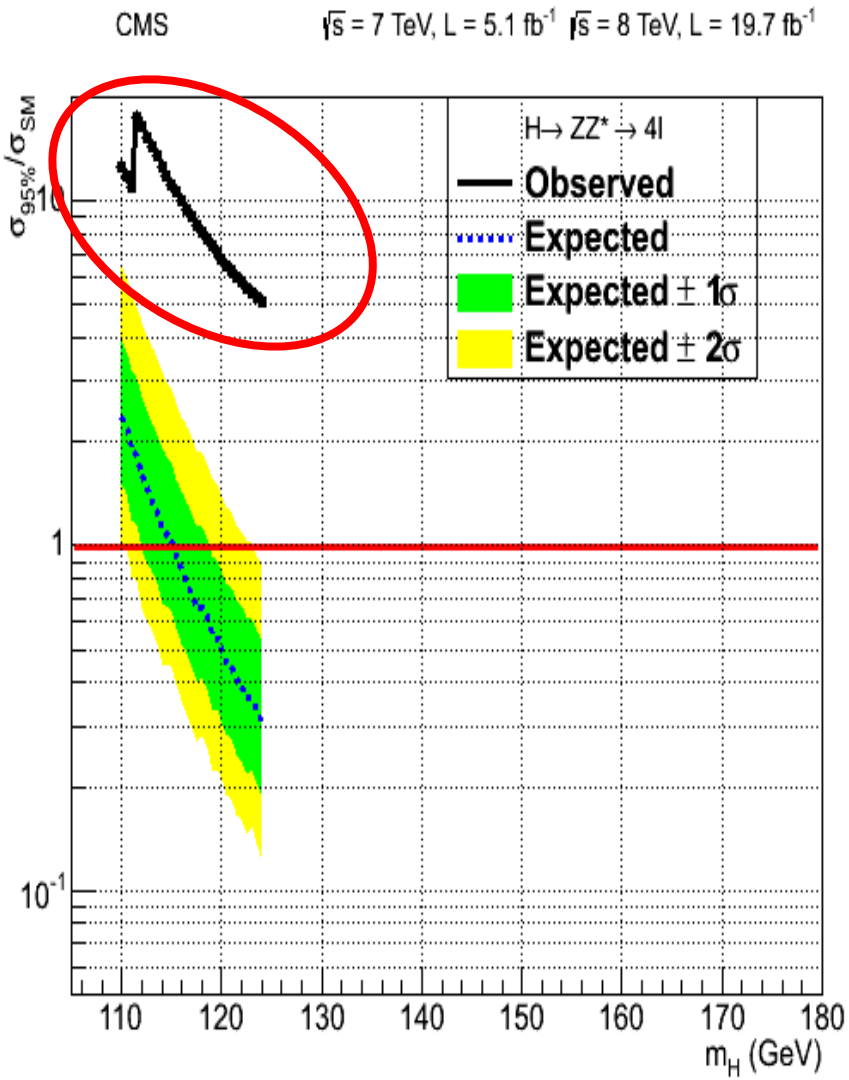
(Determine how the signal or Bkg like my data)

- I run limits for sample of mass points mH (110 - 124).
- problem: in the observed limit it is so high although Observed number of events is Same as Legency one in all data cards.

Masses 110 – 124 GeV



Legency Cards



My Cards

Next Step

- Try to understand why I have observed limits So high
- Then Run limits for the other mass points
- Making the Same work but for High mass Signal.


Thank
You!