Summary

08/01/2020

Danyi Zhang

What I did:

- Read some papers about statistical analysis
  - G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys*, J. C 71, 1554 (2011).
  - **6.2 Statistical analysis** in the note ATL-COM-PHYS-2016-1609
- Set up the HZZ Workspace and run examples
- Try to understand the analysis idea of the workspace

Workspace

- Inside workspace
- How to use:
  - collect the terms (desicribe some details)
  - Give command
  - Get the result!
- We only use part of the utilities for checks of the result:
  - plot NLL scan (test smoothness of likelihood). Tool: scan_poi

each category:

$$P = \{\underbrace{\mu \cdot n^0_{sig}}_{①} \cdot \underbrace{(1+\alpha(\theta))}_{②} \cdot \underbrace{p_{sig}(m_{4\ell}; m_H, \theta)}_{③} + \underbrace{n^0_{bkg}}_{①} \cdot \underbrace{(1+\alpha(\theta))}_{②} \cdot \underbrace{p_{bkg}(m_{4\ell}; \theta)}_{③}\} \cdot \underbrace{\prod_i G(\theta; \theta_g, 1)}_{④}$$

①: the number of expected events; *RooRealVar.*

②: normalization uncertainties; *RooStats::HistFactory::FlexibleInterpVar.*

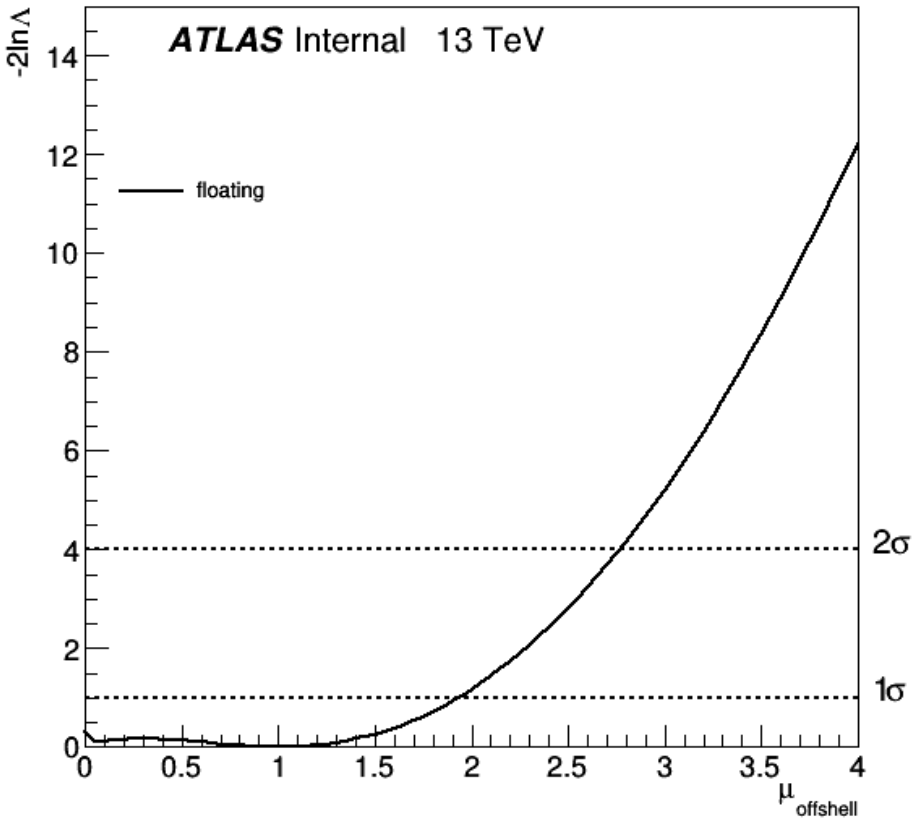
③: PDF for signal and background, including variations from systematic uncertainties; Depends on observable and modeling.

④: Gaussian constraint term, mimic the auxiliary measurement.

- **θ** is nuisance parameter, represents responding systematic uncertainties, default: 0
- $\boldsymbol{\theta_g}$ is global observable, default: 0.
- **μ** is signal strength, parameter of interest (POI).

Analysis idea of NLL scan

- NLL: negative log likelihood $\Lambda(\mu) = \frac{L\left(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu)\right)}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}$
- For the case of a single parameter of interest,
  - $-2ln\Lambda(\mu) = \frac{(\mu-\hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N})$ —— a parabola-like curve
- Scan:
  - scan_poi combined.root out.root ws_name mu_name data_name mH:100:120:130 mH:125.,mu:1.0
  - scan_poi output. root scan.root combined mu asimovData mu:100:0:5
- The example result is shown on the right-hand side



About the code

- I also read the plot code: plot_scan_multiple.python, draw1DNLL.cxx
- I understood the code is used to plot the $-2ln\Lambda(\mu)$ as a function of μ
- But I found some lines which I didn't get the idea:

```
double value = 2* (_nll - minNLL);    // What does value refer to? Why to define like this?
value = value<0?0:value;
massvalue.push_back(_mass);
nllvalue.push_back(value);
```

- This is from the function: `getGraphFromFile` defined in draw1DNLL.cxx

What we want to do

- We want to see the effect of fixing or floating the qqZZ normalisation on the signal strength $\mu_{off-shell}$
- Fixing & floating
  - Floating: profiling the normalization in the fit, mu_ZZ will be a free parameter
  - Fixing: to fix the normalization, which means to estimate the background from their MC simulation