HEP 2013
Stockholm
18-24 July 2013
(info@eps-hep2013.eu)

RooStats

A Framework for Advanced Statistical Analysis

L. Moneta (CERN/PH-SFT)
on behalf of the RooStats team
(K. Cranmer, C. Gumpert, S. Kreiss, G. Kukartsev, G. Lewis, S. Jain, W. Verkerke)

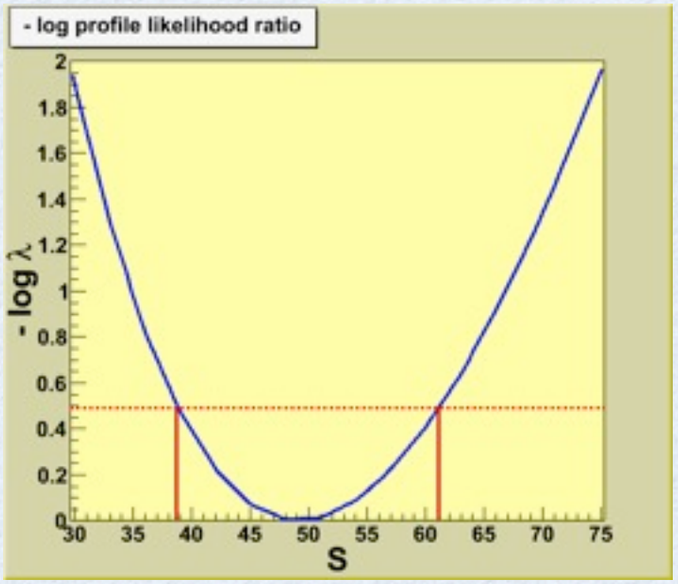
- Introduction to RooStats
- RooStats technology (RooFit)
- RooStats Calculators
  - Hypothesis tests
  - Frequentist interval/limit calculators (CLs)

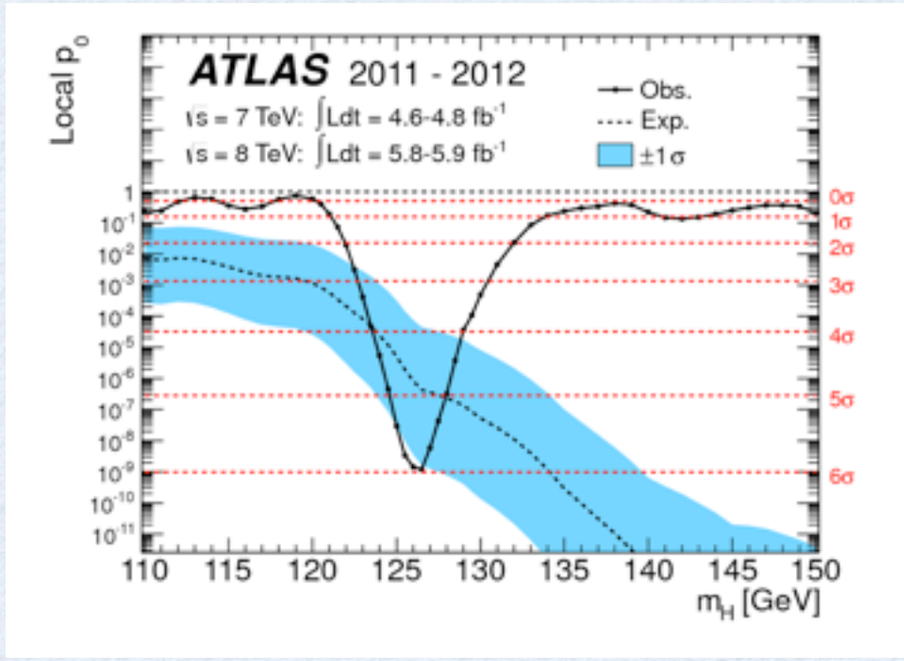
What is RooStats?

- Project to provide advanced statistical tools for LHC experiments
  - factorize model description from statistical calculations
  - implement different statistics techniques
    - Frequentist, Bayesian or Likelihood Based
  - provide utilities for combination of results
- Joint contribution between ATLAS, CMS, RooFit and ROOT

Statistical Application

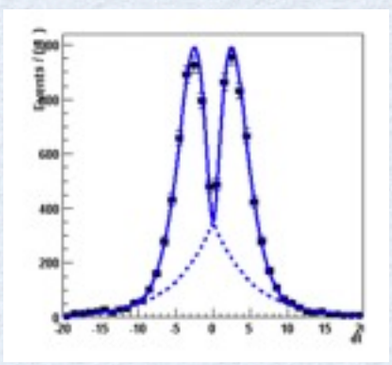
- Point estimation (covered by RooFit)
- Estimation of confidence (credible) intervals (lower/upper limits)
- Hypothesis tests (e.g discovery significance)





Underlying Technology

- Built on top of RooFit
  - provide generic model description
    - binned (histogram based)
    - unbinned (parametrized) models
  - provide tools to facilitate model creation
  - tools for combination of models

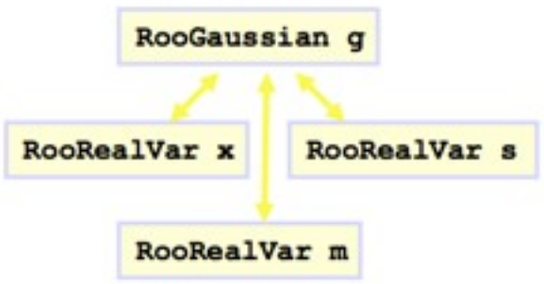


- Use core ROOT libraries
  - minimization (Minuit), numerical integration, etc.
  - additional tools provided when needed (e.g. Markov-Chain MC)

RooStats Technology

Building models with RooFit

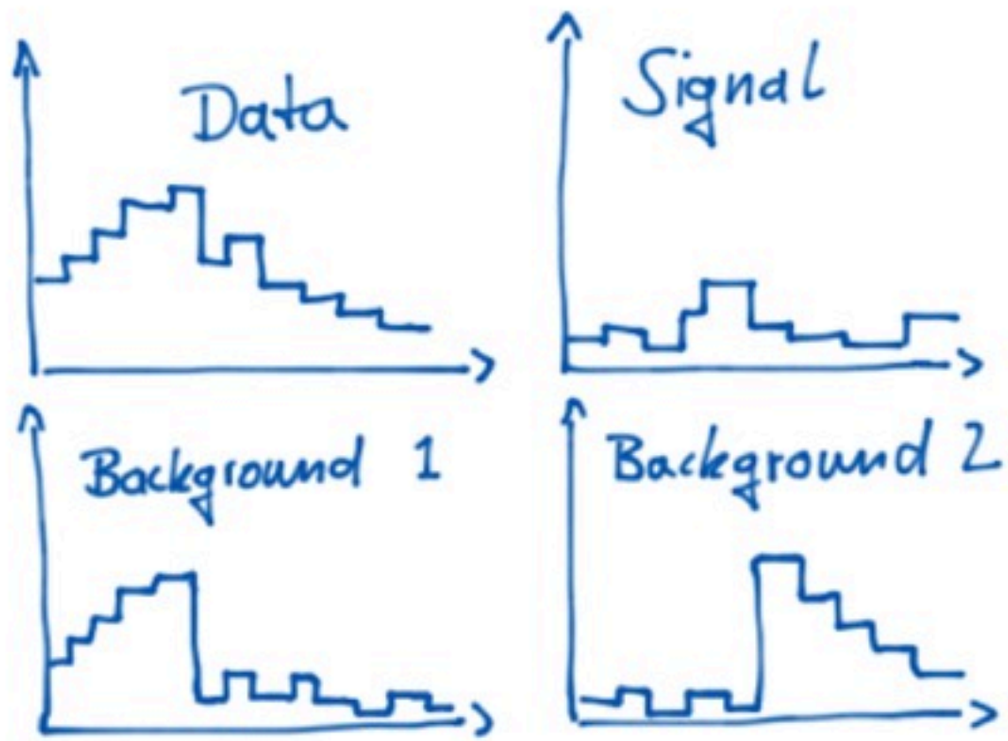
Mathematical concept		RooFit class
variable	x	`RooRealVar`
function	$f(x)$	`RooAbsReal`
PDF	$f(x)$	`RooAbsPdf`
space point	$\vec{x}$	`RooArgSet`
integral	$\int_{x_{\min}}^{x_{\max}} f(x)dx$	`RooRealIntegral`
list of space points		`RooAbsData`




```
RooRealVar x("x","x",2,-10,10)
RooRealVar s("s","s",3) ;
RooRealVar m("m","m",0) ;
RooGaussian g("g","g",x,m,s)
```

Gaussian(x,m,s)

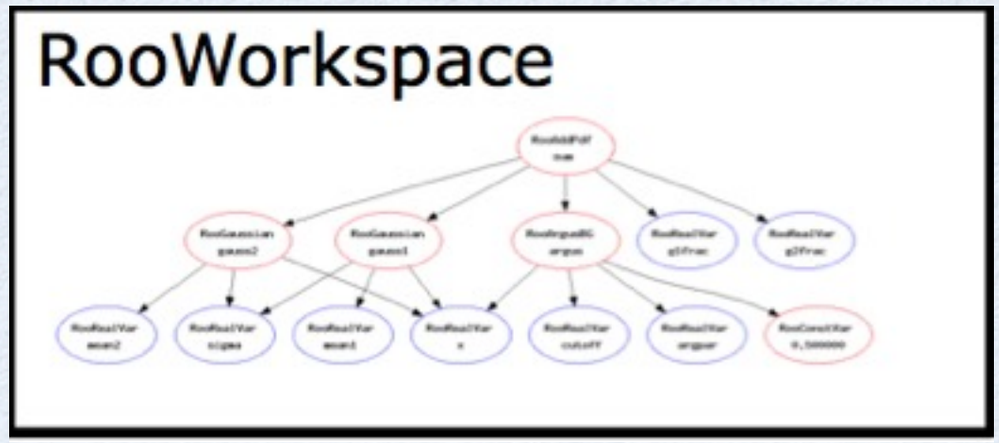
Models with HistFactory



$$\mathcal{P}(n_m, a_p|\mu, \alpha_p) = \prod_{c\in\text{channels}} \prod_{b\in\text{bins}_c} \text{Pois}(n_{cb}|\nu_{cb}) \cdot G(L_0|L, \Delta_L) \cdot \prod_{p\in\text{Syst}} P_p(a_p|\alpha_p)$$

RooFit Workspace

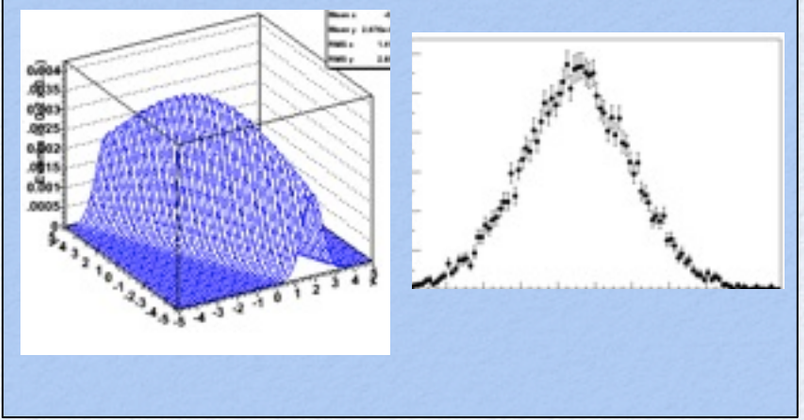
- Container for all model objects
  - PDF and their parameters
  - uncertainty and their shapes
  - (multiple) data sets
- Maintain complete description of the model
  - can be saved in a ROOT file
- All information (likelihood function) is available for further analysis



RooFit/RooStats at LHC (Higgs Analysis)

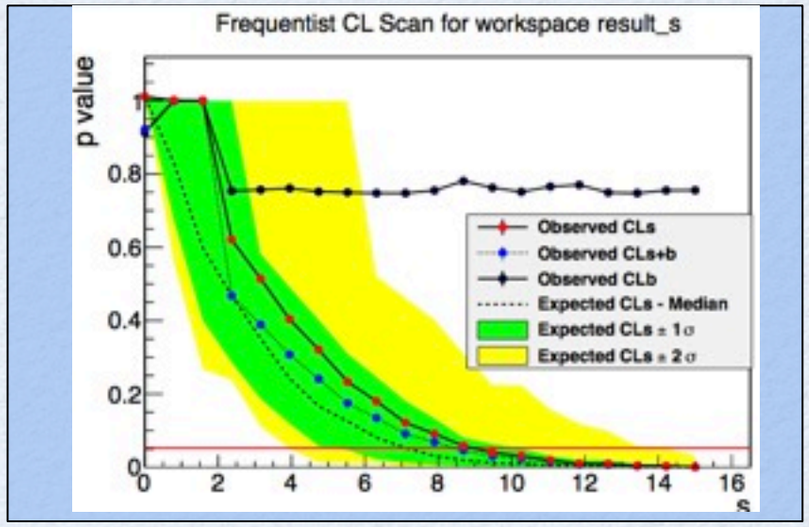
Class RooWorkspace

Simplify packaging and sharing of models



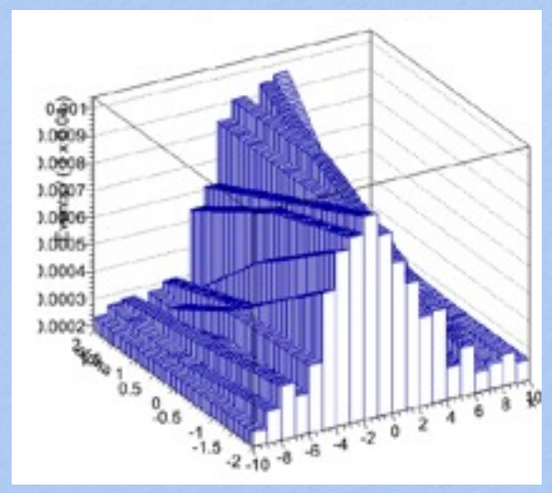
RooStats toolkit

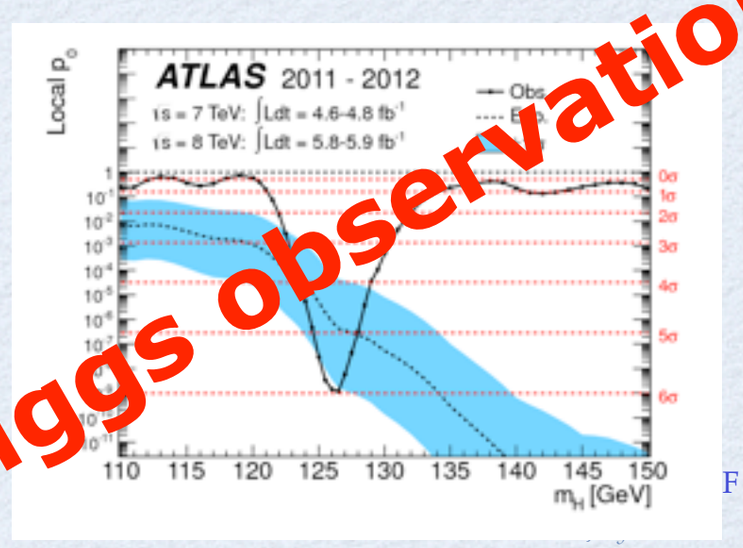
Statistical tests based on likelihoods from RooFit models



HistFactory package

Constructing models from Monte Carlo templates

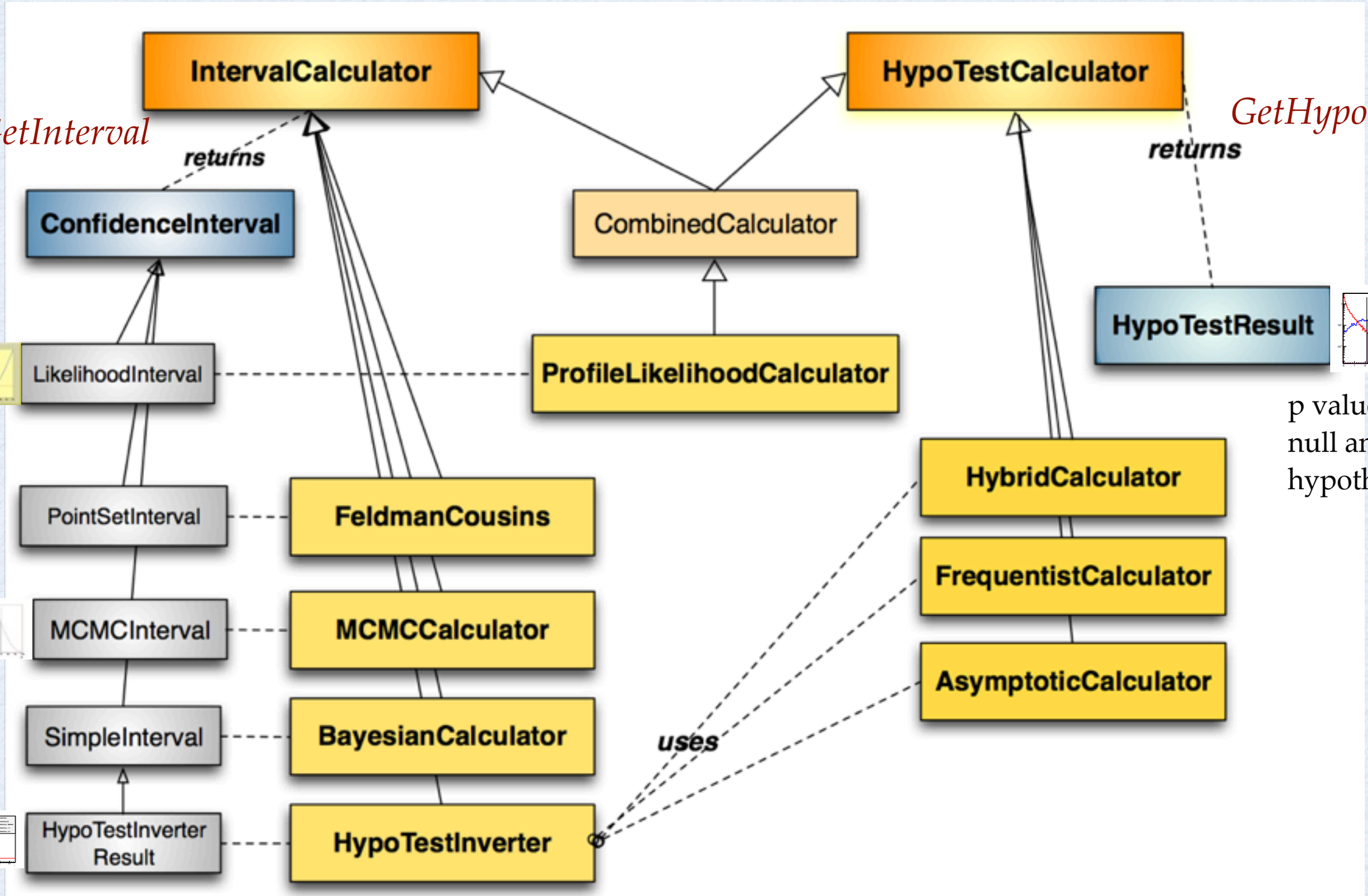




Higgs observation

RooStats Design

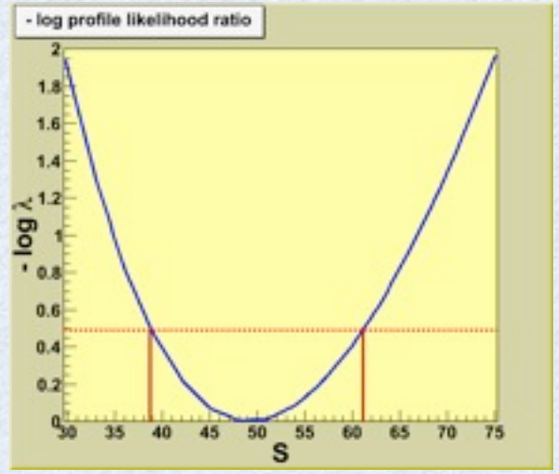
- C++ interfaces and classes mapping to real statistical concepts

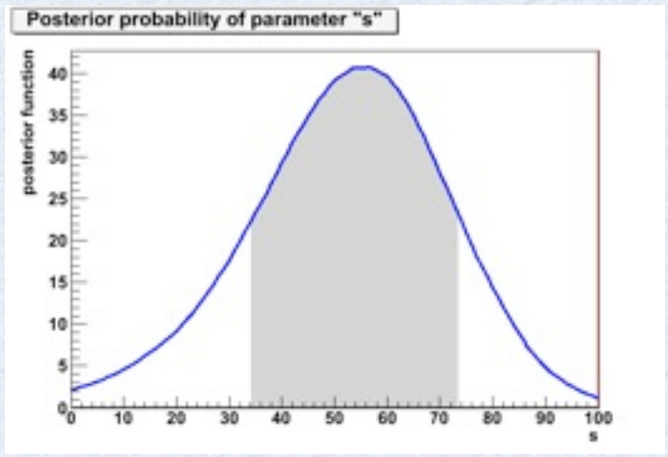


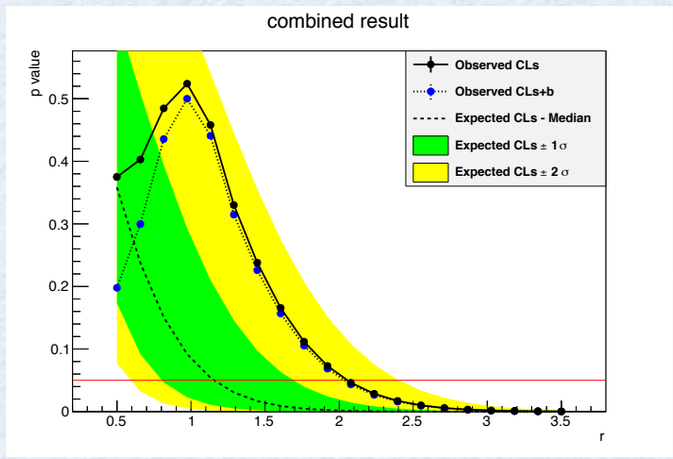
p values for null and alt. hypothesis

RooStats Calculators

- Interval Calculators
  - ProfileLikelihoodCalculator
    - interval estimation using the asymptotic properties of the likelihood function (Minos)
  - BayesianCalculator
    - interval estimation based on Bayes theorem using adaptive numerical integration
  - MCMCCalculator
    - Bayesian calculator using Markov-Chain Monte Carlo
  - HypoTestInverter
    - frequentist interval calculation using hypothesis test
    - can compute CLs limits or Feldman-Cousins interval

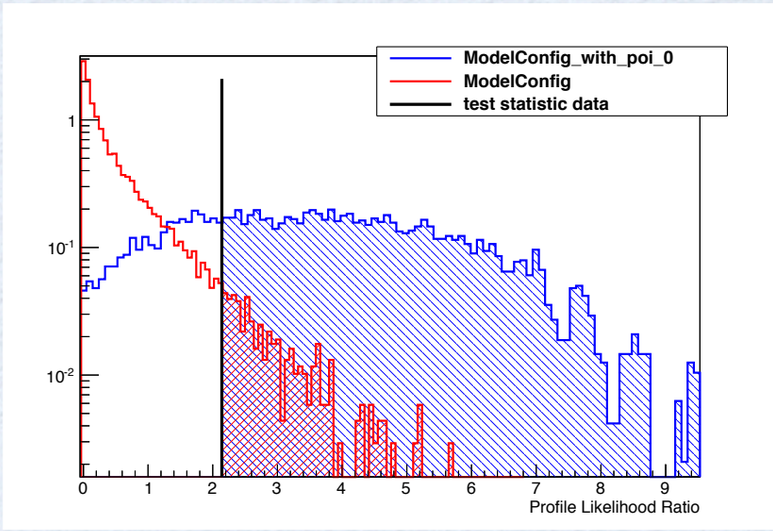


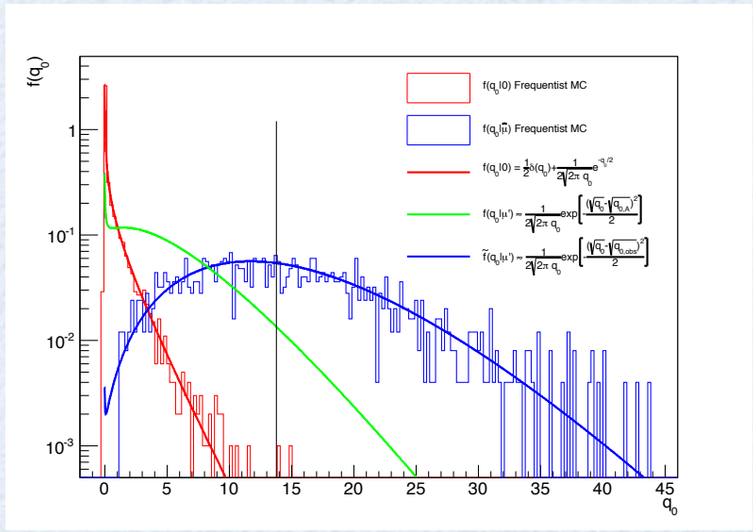




RooStats Calculators (2)

- Hypothesis Test Calculators
  - FrequentistCalculator
    - frequentist hypothesis tests using pseudo-experiments to determine the test statistics distributions (parametric bootstrap)
  - HybridCalculator
    - same as frequentist calculator by using a bayesian treatment (marginalization) of systematic uncertainties
  - AsymptoticCalculator
    - hypothesis tests using asymptotic likelihood formulae
      - ➡ Cowan, Cranmer, Gross, Vitells, arXiv:1007.1727,EPJC 71 (2011) 1-1

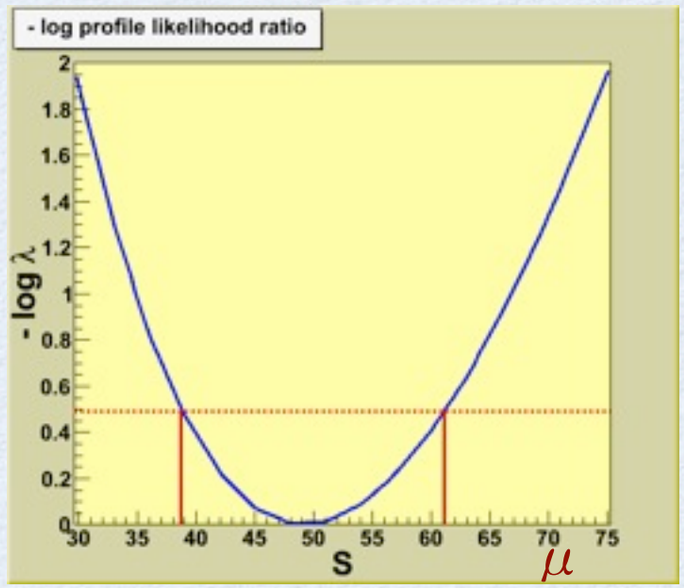




Using RooStats Calculators

- All RooStats calculators require same input:
  - model (described by the ModelConfig class which is linked to a workspace)
  - observed data
- Result is a ConfidenceInterval object or a HypoTestResult object
- Classes for plotting the result are also provided

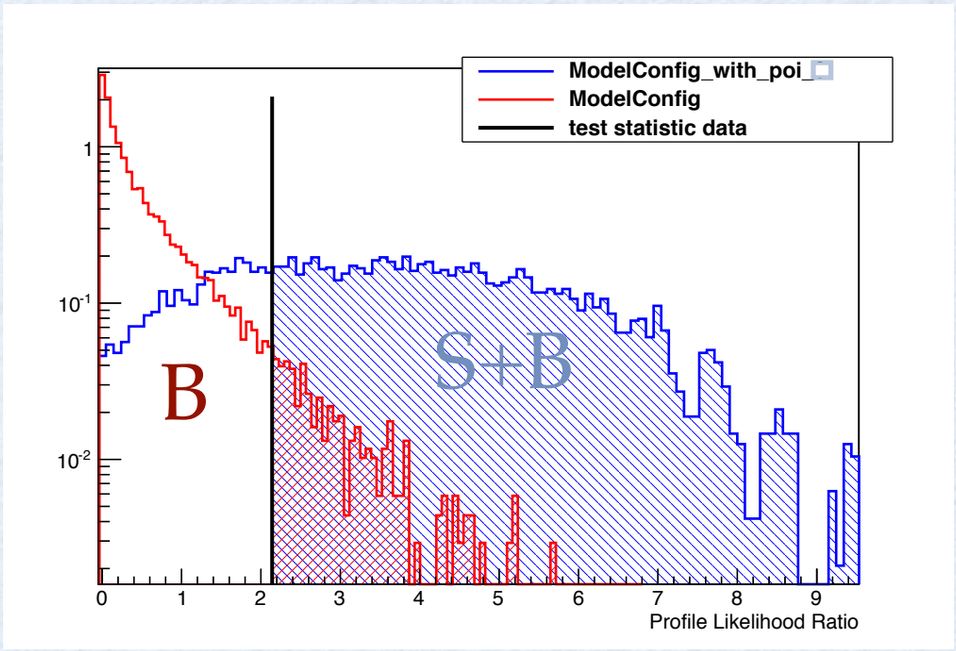
```
// create the class using data and model
ProfileLikelihoodCalculator plc(data, model);

// set the confidence level
plc.SetConfidenceLevel(0.683);

// compute the interval
LikelihoodInterval* interval = plc.GetInterval();

// plot the interval
LikelihoodIntervalPlot plot(interval);
plot.Draw();
```



RooStats Hypothesis Test

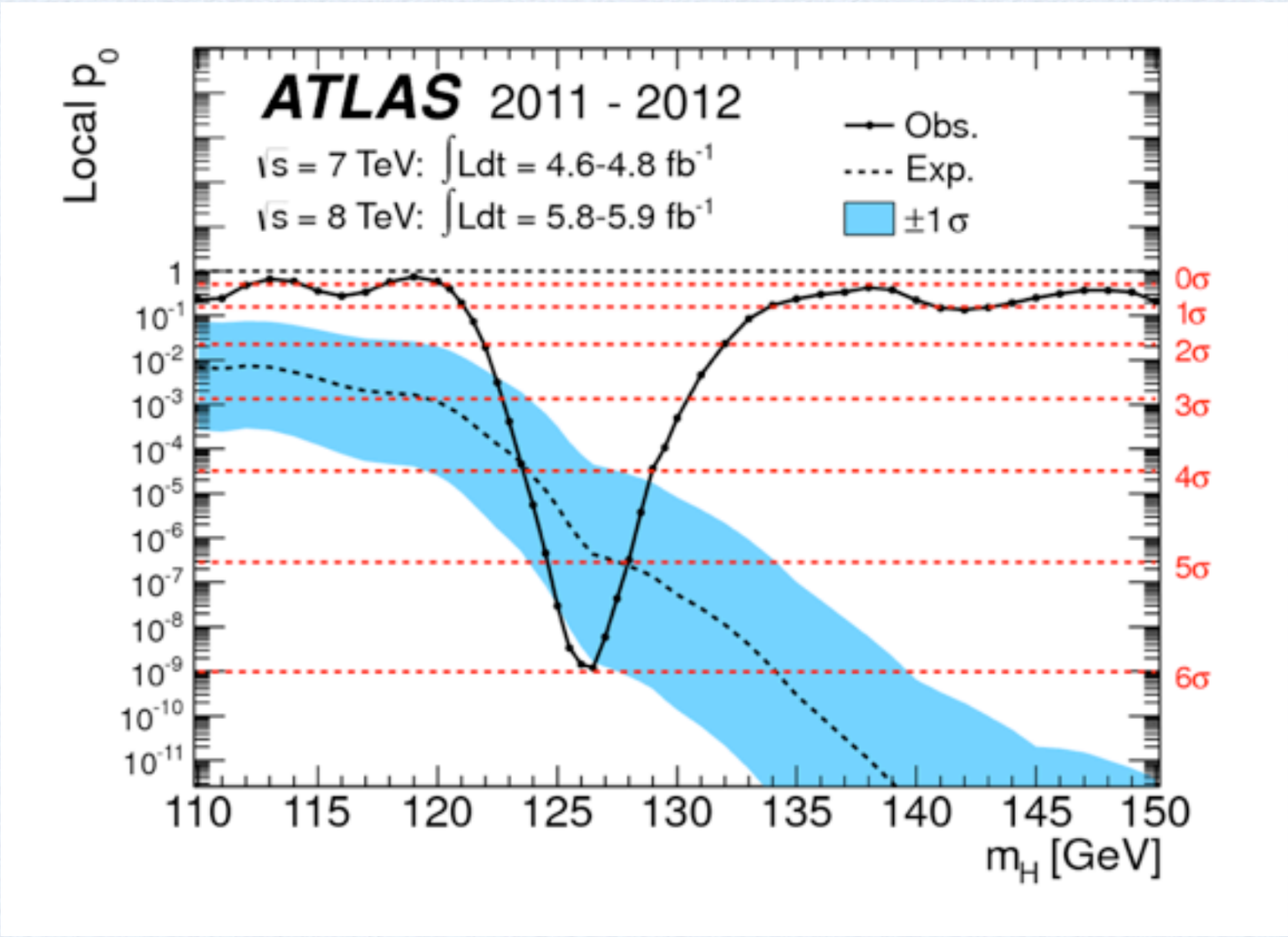
- Define null and alternate model. For discovery test
  - null: Background only model ($\mu = 0$)
  - alternate: Signal + Background model (e.g. $\mu = 1$)
- Select test statistics to use
  - e.g profile likelihood ratio (preferred due to known asymptotic formulae)
- Select type of calculator
  - asymptotic or based on toys
  - treatment of nuisance parameters
- Result is p-value for null (p_0) and alternate models (p_{s+b})

$$\lambda(\mu) = \frac{L(x|\mu, \hat{\nu})}{L(x|\hat{\mu}, \hat{\nu})}$$



Example: Discovery Significance

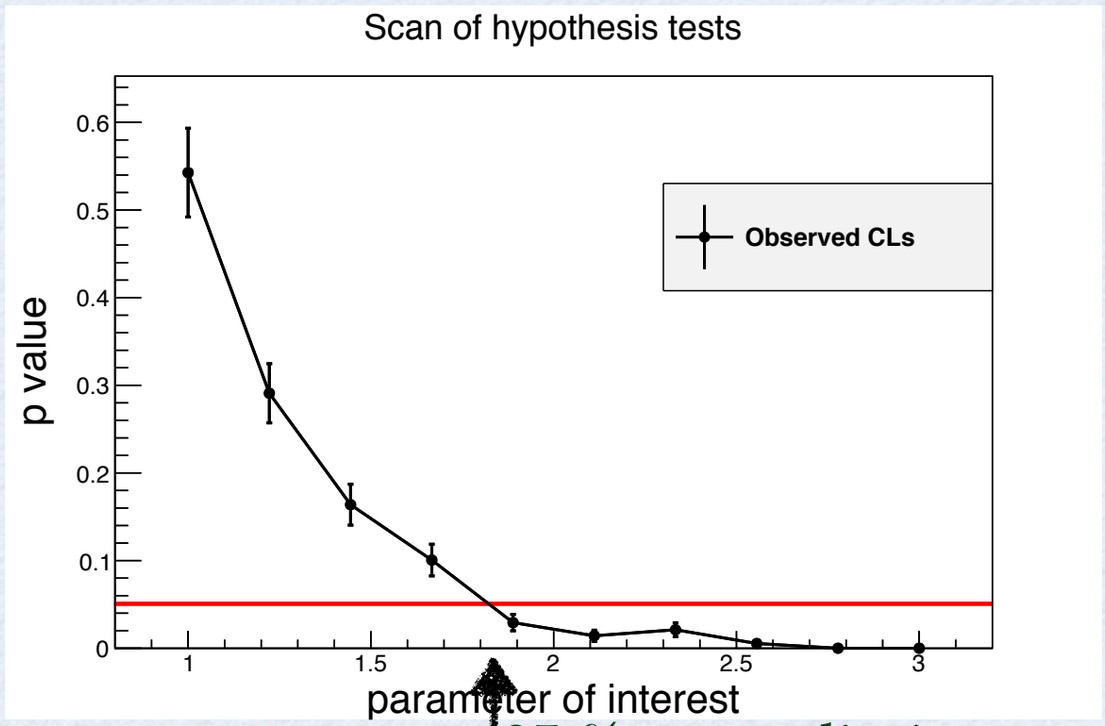
- Performing the tests for different mass hypotheses (*i.e* different signal models):

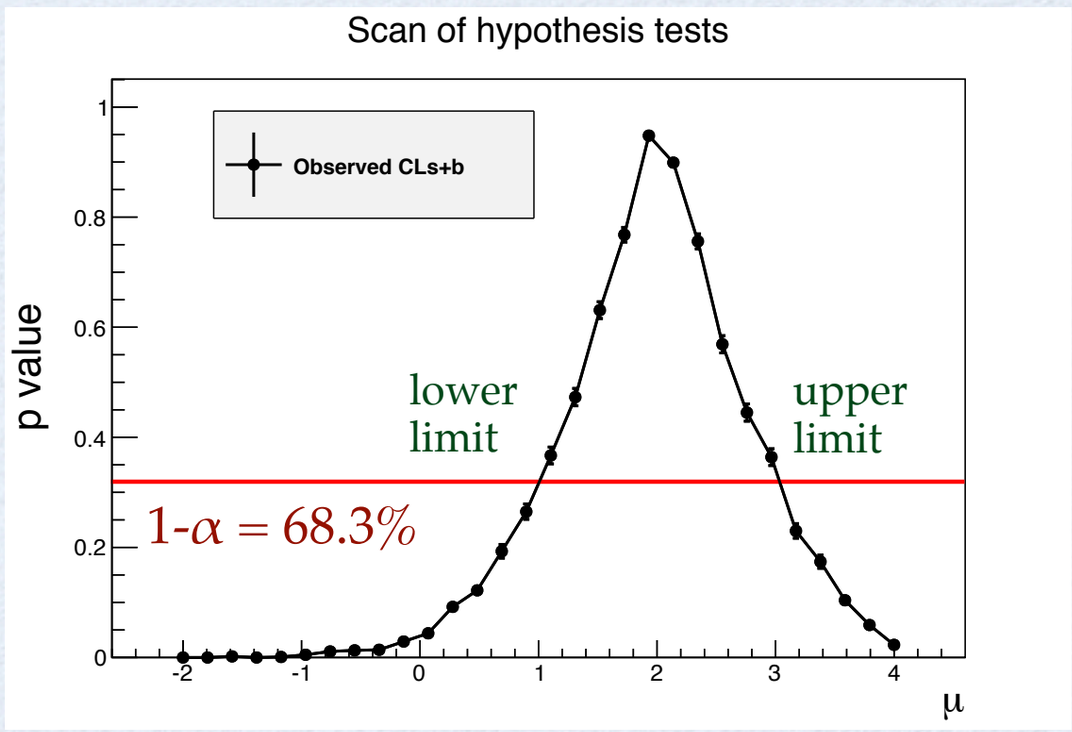


Expected significance is obtained from median of alternate (S+B) model

Hypothesis Test Inversion

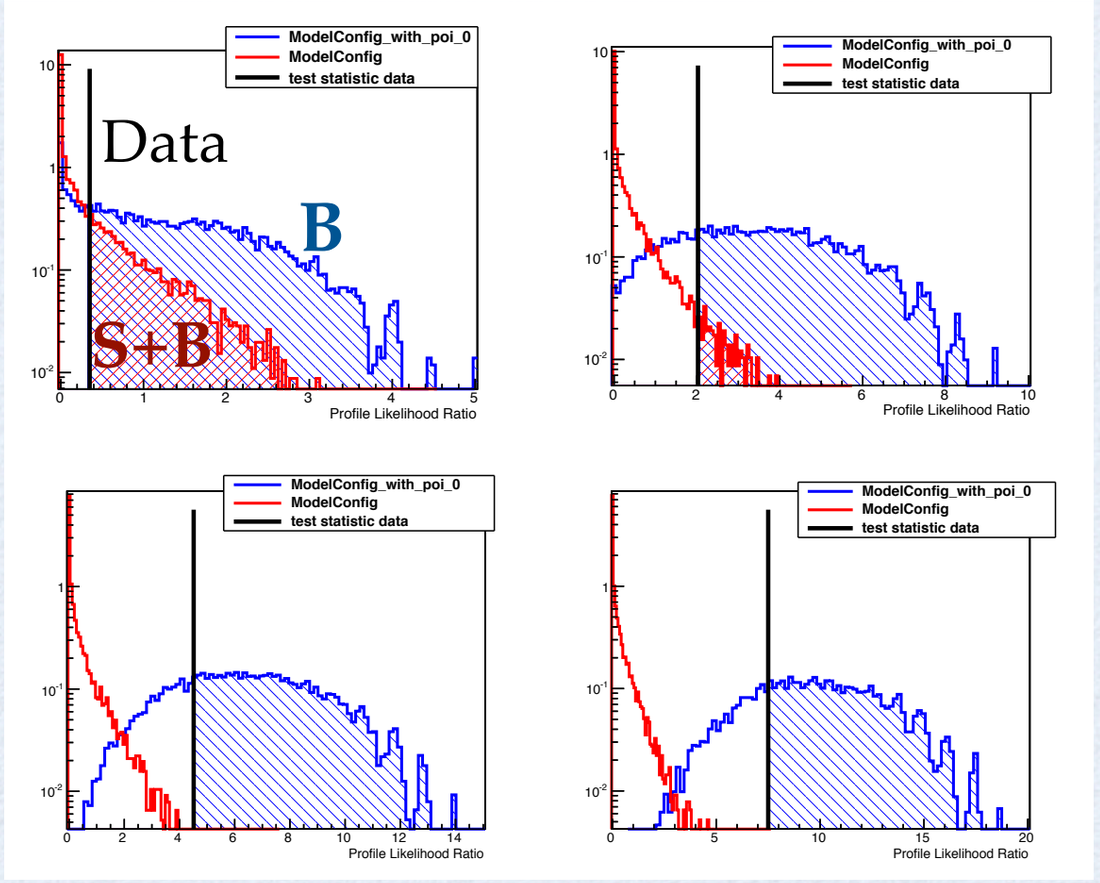
- Perform an hypothesis test at each value of the parameter
- Interval can be derived by inverting the p-value curve, function of the parameter of interest (μ)
  - value of μ which has p-value α (e.g. 0.05), is the upper limit of 1-α confidence interval (e.g. 95%)
  - for upper limits use $CL_s = CL_{s+b}/CL_b$



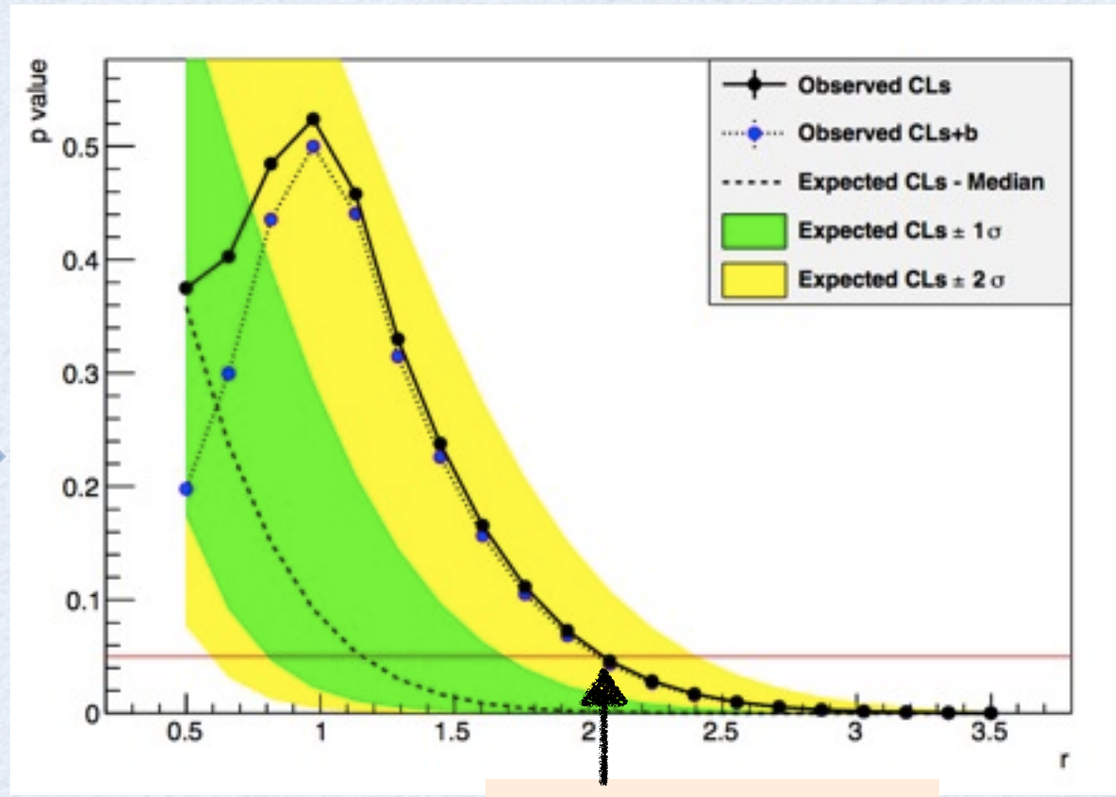


RooStats Hypo Test Inversion

- Can use Frequentist, Hybrid or Asymptotic calculator
- Compute observed, expected limits and bands



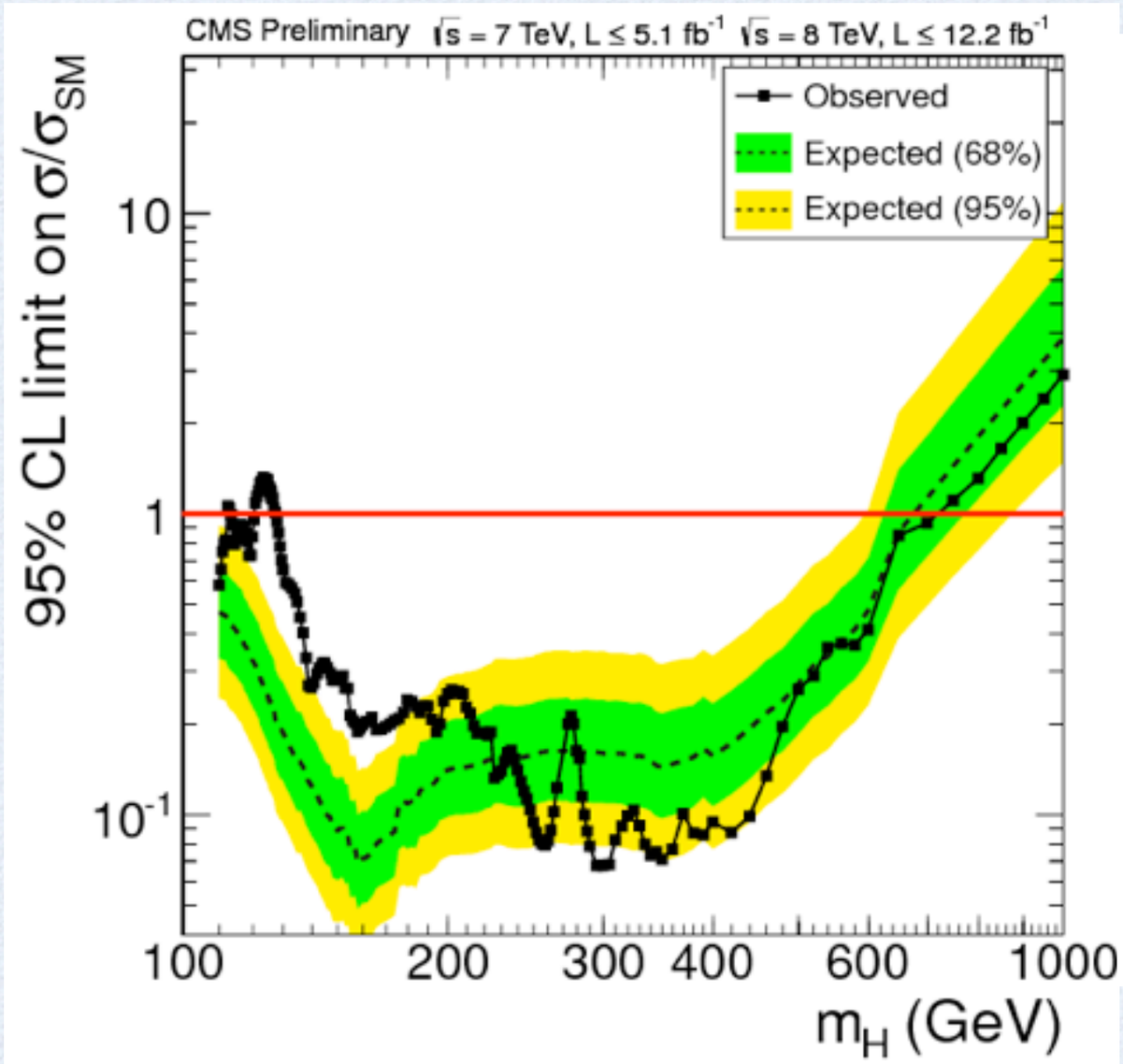
CL_s Upper Limits

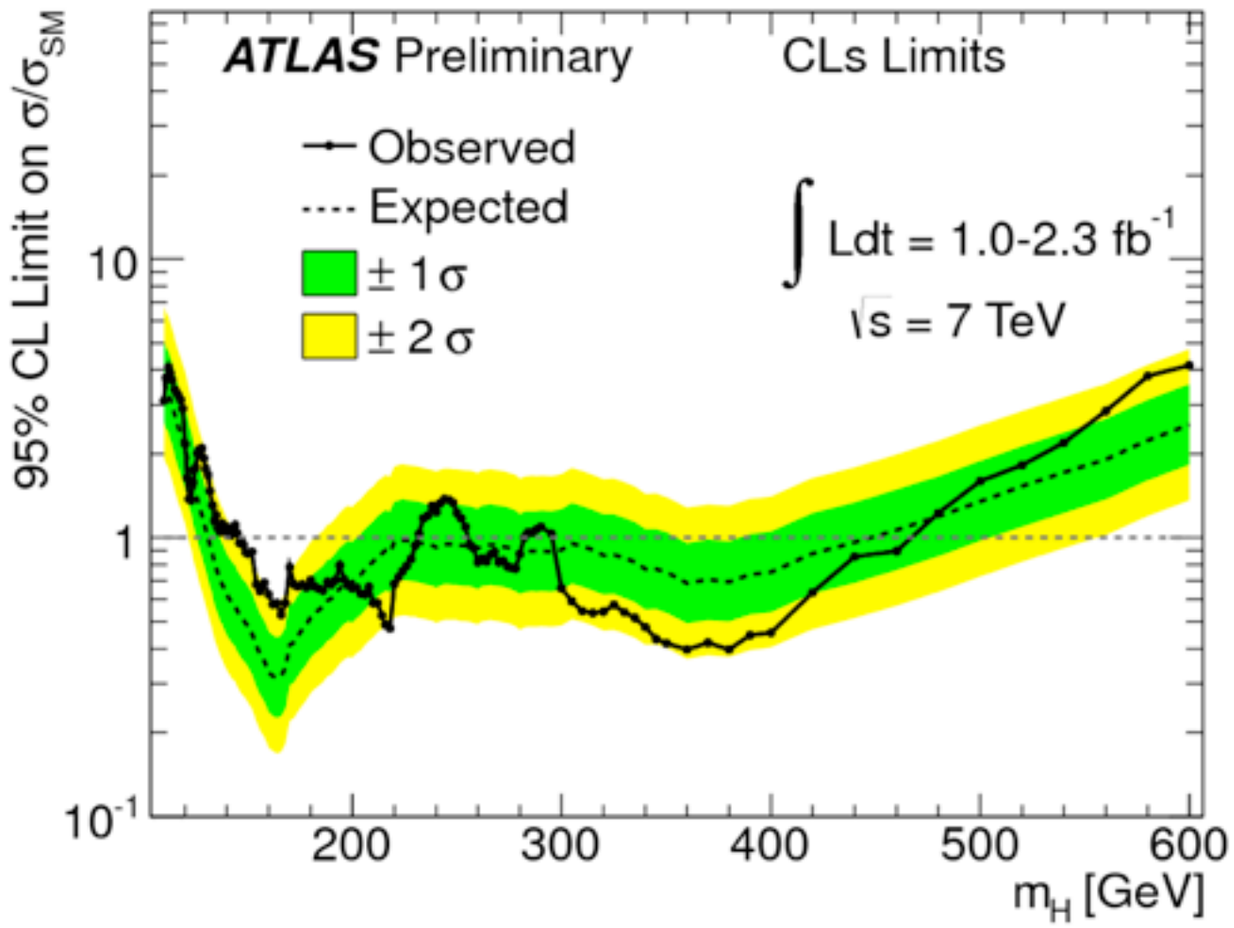


95% Limit

Example: Computing Limits

- By computing limits for different mass hypothesis:





Bayesian Analysis in RooStats

- **RooStats** provides classes for
  - marginalize posterior and estimate credible interval

posterior probability

likelihood function

prior probability

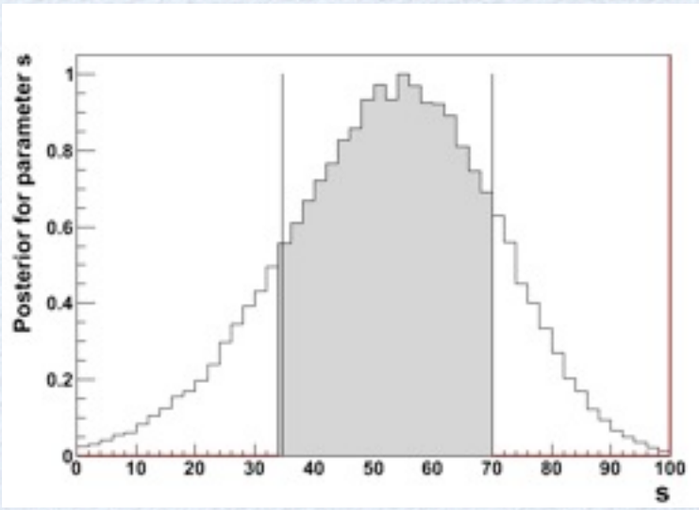
nuisance parameters marginalization

$$P(\mu|x) = \frac{\int L(x|\mu,\nu)\Pi(\mu,\nu)d\nu}{\underbrace{\int\int L(x|\mu,\nu)\Pi(\mu,\nu)d\mu d\nu}_{\text{normalisation term}}}$$

POI data

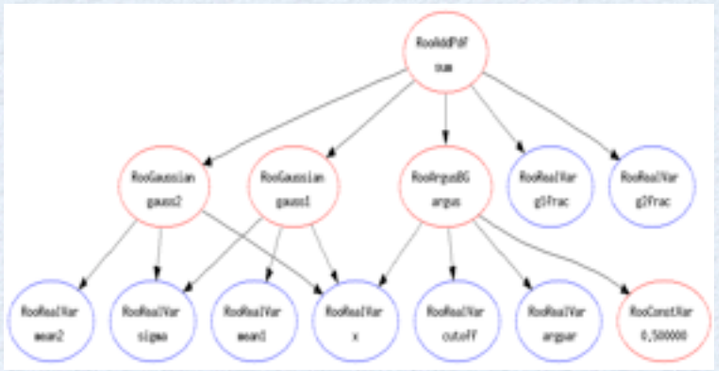
Bayesian Theorem

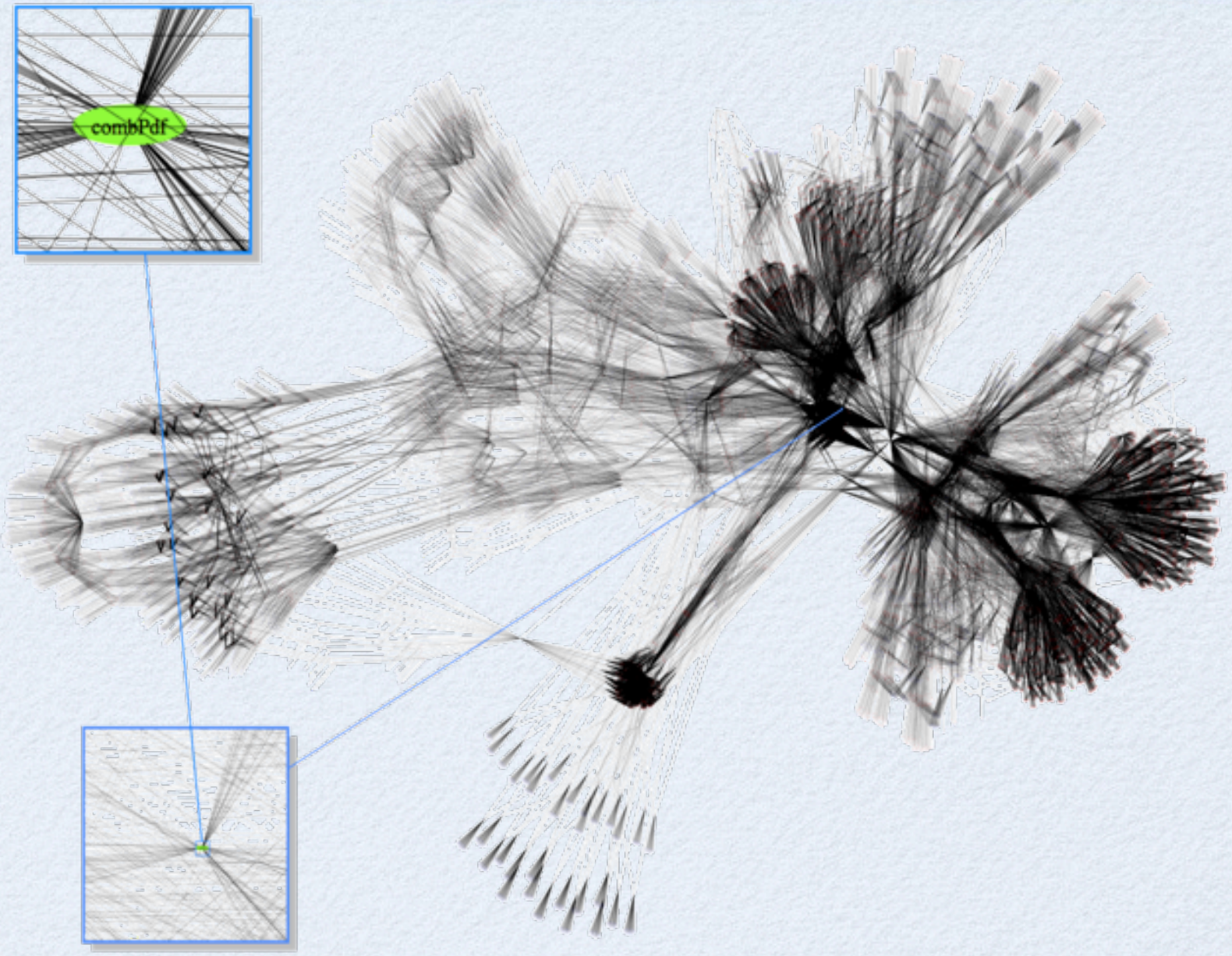
  - support for different integration algorithms:
    - adaptive (numerical)
    - MC integration
    - Markov-Chain
      - can work with models with many parameters (e.g few hundreds)



- Any prior can be given (up to know uniform prior are normally used)
- Working to include Reference priors (least informative and objective)
  - see *L. Demortier, S. Jain, H. B. Prosper, Phys. Rev. D82, 034002, 2010*

How well does it scale?





Graph of the full ATLAS Higgs combination model

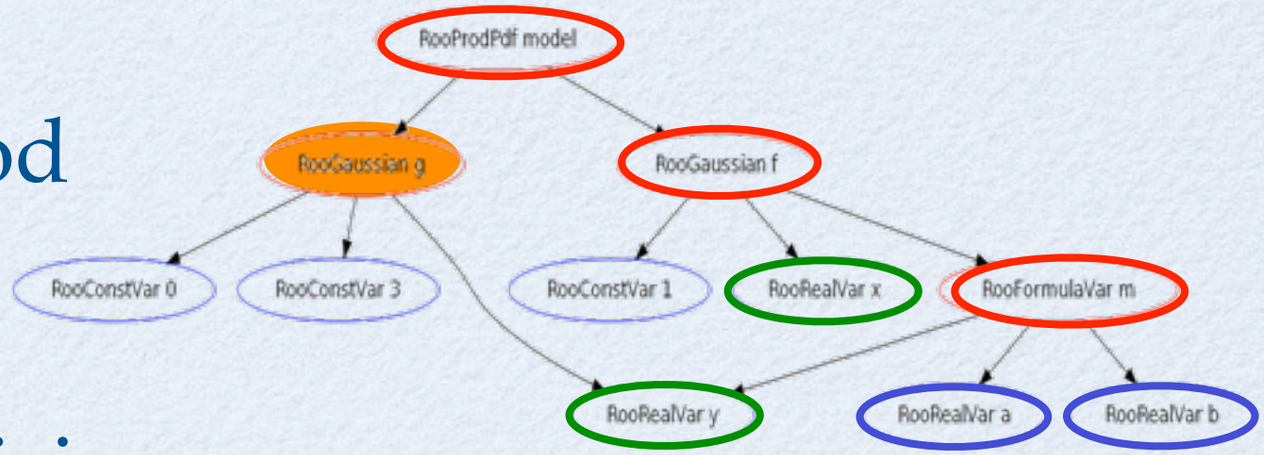
Model has ~23.000 function objects, ~1600 parameters

Reading/writing of full model takes ~4 seconds

ROOT file with workspace is ~6 Mb

Complex models Handling

- A large effort has been put recently in RooFit to handle complex models.
  - optimization of likelihood evaluations
  - improve numerical precision



- We plan improvements in
  - Minuit minimization for large number of parameters
  - make more use of vectorization and parallelizations (e.g. porting on GPU) to speed up evaluations

Summary

- RooFit/RooStats allows you to perform advanced statistical data analysis
  - LHC results (*e.g.* Higgs observation)
- Capable of using different tools and interpretations (Frequentist/Bayesian) on the same model
- Generic tools capable to deal with large variety of models
  - based on histograms or un-binned data
  - multi-dimensional observations
- Provide tools to facilitate complex model building
  - HistFactory for histogram based analysis

BackUp Slides

Profile Likelihood Calculator

- Method based on properties of the likelihood function
- Profile likelihood function:

$$\lambda(\mu) = \frac{L(x|\mu, \hat{\hat{\nu}})}{L(x|\hat{\mu}, \hat{\nu})}$$

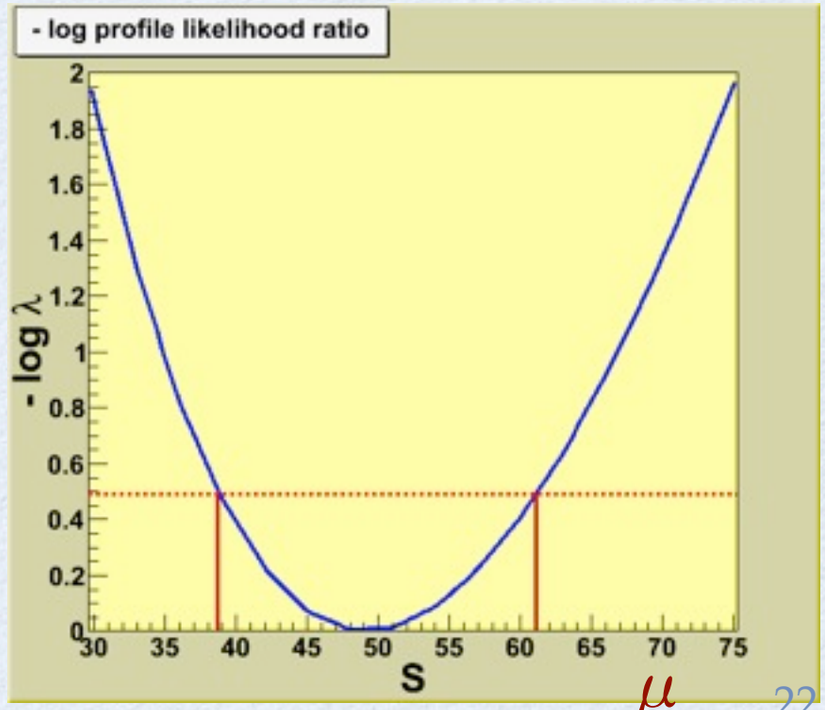
→ maximize w.r.t nuisance parameters ν and fix POI μ

→ maximize w.r.t. all parameters

λ is a function of only the parameter of interest μ

- Uses asymptotic properties of λ based on Wilks' theorem:
- from a Taylor expansion of $\log\lambda$ around the minimum:
  - ➔ $-2\log\lambda$ is a parabola (λ is a gaussian function)
  - ➔ interval on μ from $\log\lambda$ values

- Method of MINUIT/MINOS
  - lower/upper limits for 1D
  - contours for 2 parameters



Running RooStats

- ROOT macros taking all the same inputs (very easy !)
  - workspace, ModelConfig and data set names
- StandardProfileLikelihoodDemo.C

run ProfileLikelihoodCalculator - get interval and produce plot

```
root[]StandardProfileLikelihoodDemo("ws.root","w","ModelConfig","data")
```

- StandardBayesianNumericalDemo.C

run Bayesiancalculator: get a credible interval and produce plot of posterior function

```
root[]StandardBayesianNumericalDemo("ws.root","w","ModelConfig","data")
```

- StandardBayesianMCMCDemo.C

run bayesian MCMCCalculator: get a credible interval and produce plot of posterior function

```
root[]StandardBayesianMCMCDemo("ws.root","w","ModelConfig","data")
```

- StandardHypoTestInvDemo.C

run for CLs (with frequentist calculator (type = 0) and one-side PL test statistics (type = 3) scan 10 points in [0,100]

```
root[] StandardHypoTestInvDemo.C("ws.root","w","ModelConfig","","data",0,3, true, 10, 0, 100)
```

run for Asymptotic CLs (scan 20 points in [0,100])

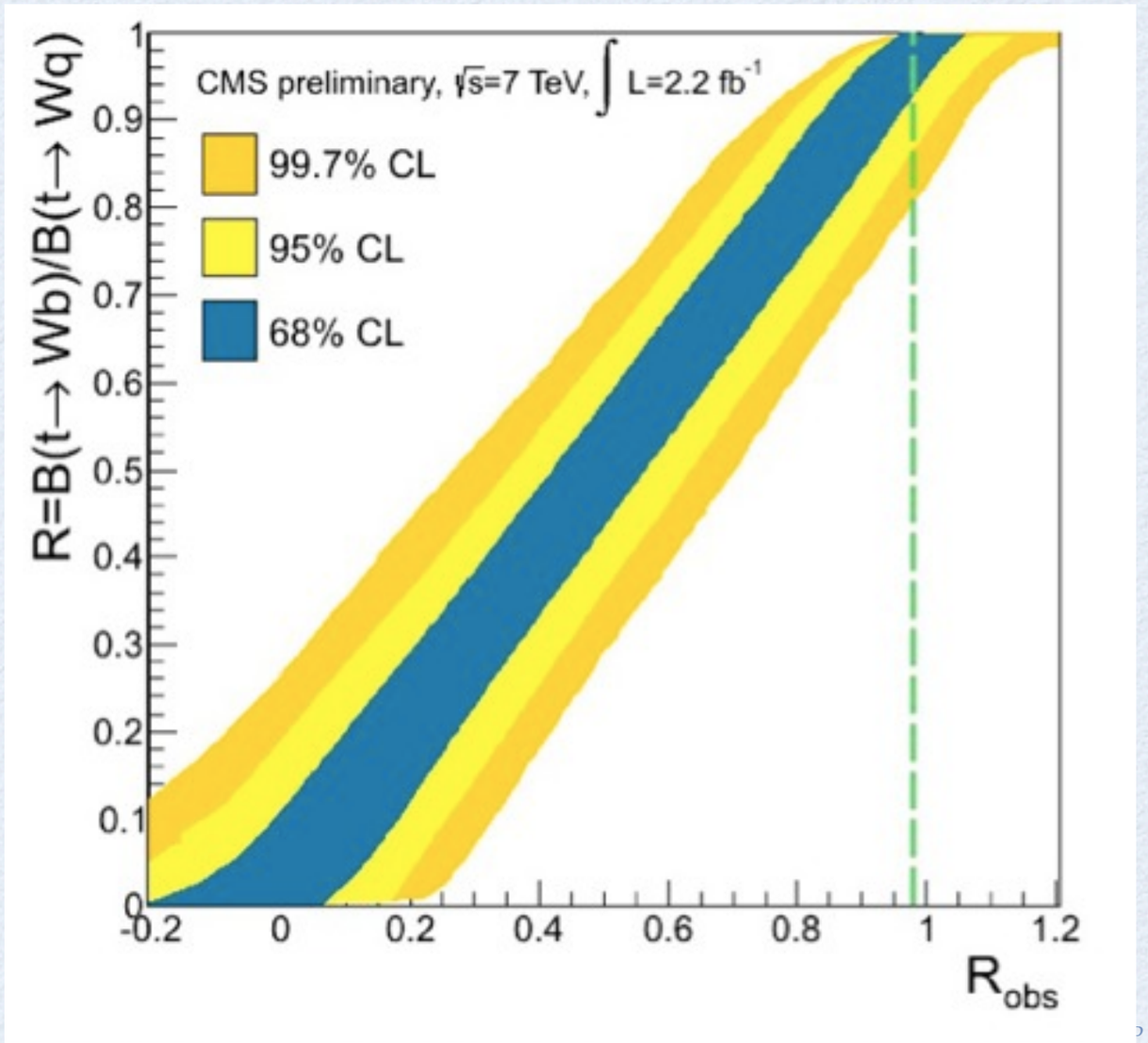
```
root[] StandardHypoTestInvDemo.C(“ws.root","w","ModelConfig","","data",2,3, true, 20, 0, 100)
```

Frequentist Hypothesis Tests

- Ingredients:
  - Null Hypothesis: the hypothesis being tested (e.g. $\theta = \theta_0$), assumed to be true and one tries to reject it
  - Alternate Hypothesis: the competitive hypothesis (e.g. $\theta \neq \theta_0$)
  - Test statistics: a function of the data, t(X) ,used for defining the critical region in multidimensional data: $X \in w \rightarrow t(X) \in w_t$
    - w is the critical region, a subspace of all possible data for which the null hypothesis is rejected
    - size of test : $\alpha = P(X \in w \mid H_0)$
    - power of test : $1- \beta = P(X \in w \mid H_1)$

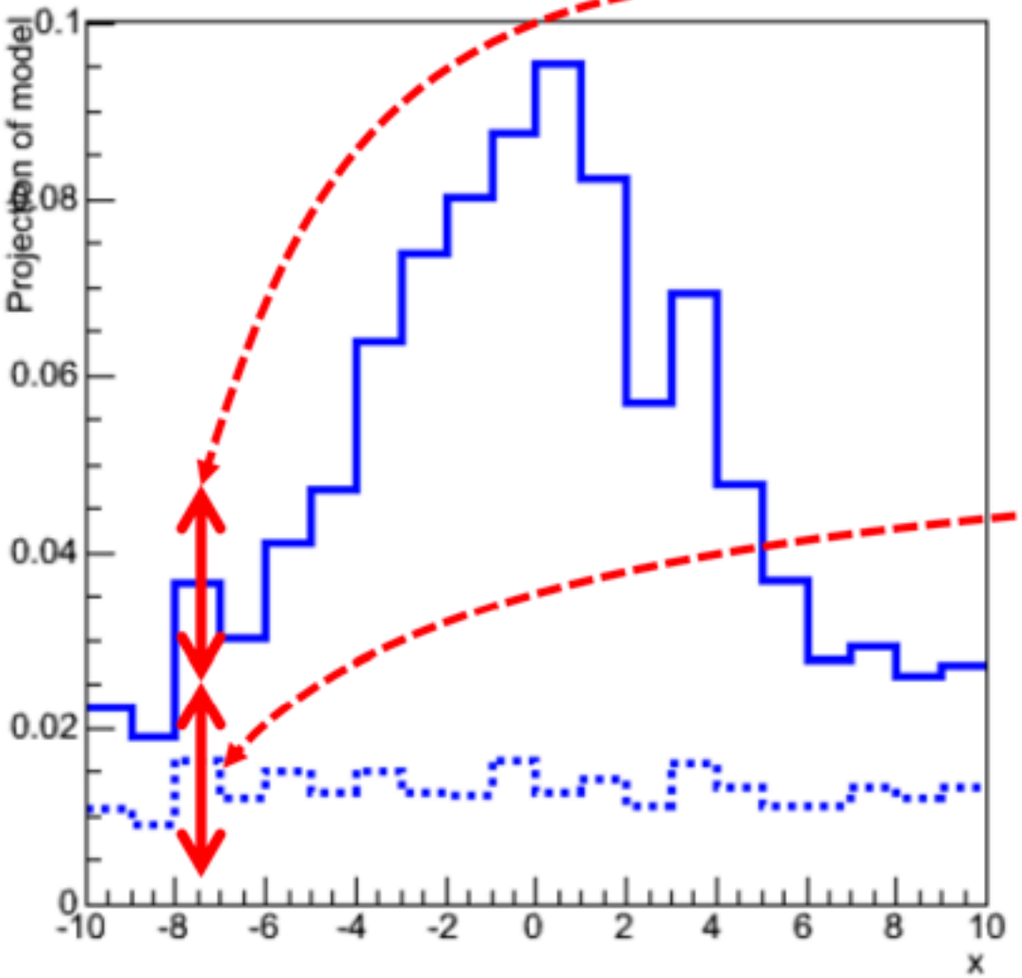
Example: Feldman-Cousins interval

- Same RooStats code but with different configuration (using a 2-sided test statistics) can compute also a Feldman-Cousins interval



HistFactory: modeling uncertainties on templates

- Histogram-shaped model not new (RooHistPdf), but key issue in physics analysis is in modeling uncertainties on this model (RooHistPdf has no degrees of freedom)
  - HistFactory: Model MC statistical uncertainties by allowing each histogram bin to float.
  - Then constrain rate in each bin with a Poisson distribution based on MC event count ("subsidiary measurements")



$$L(\tilde{y}_{0,s},\ldots,\tilde{y}_{n,s}) = \prod_{i=0}^{n} P_s(\tilde{y}_{i,s}\,|\,\theta_{i,s}) P_b(\tilde{y}_{i,b}\,|\,\theta_{i,b})$$

Signal simulation event count for bin i

Background simulation event count for bin i

HistFactory

- Tool available in ROOT (in roofit/histfactory) to build models based on histograms
  - generalization of number counting models

$$\mathcal{P}(n_b|\mu) = \mathrm{Pois}(n_{\mathrm{tot}}|\mu S + B) \left[\prod_{b\in\mathrm{bins}} \frac{\mu\nu_b^{\mathrm{sig}} + \nu_b^{\mathrm{bkg}}}{\mu S + B} \right]$$

where n_b is the data histogram

in general HistFactory produces model of this form

$$\mathcal{P}(n_{cb}, a_p \,|\, \phi_p, \alpha_p, \gamma_b) = \prod_{c\in\mathrm{channels}} \prod_{b\in\mathrm{bins}} \mathrm{Pois}(n_{cb}|\nu_{cb}) \cdot G(L_0|\lambda, \Delta_L) \cdot \prod_{p\in\mathbb{S}+\mathbf{\Gamma}} P_p(a_p|\alpha_p)$$

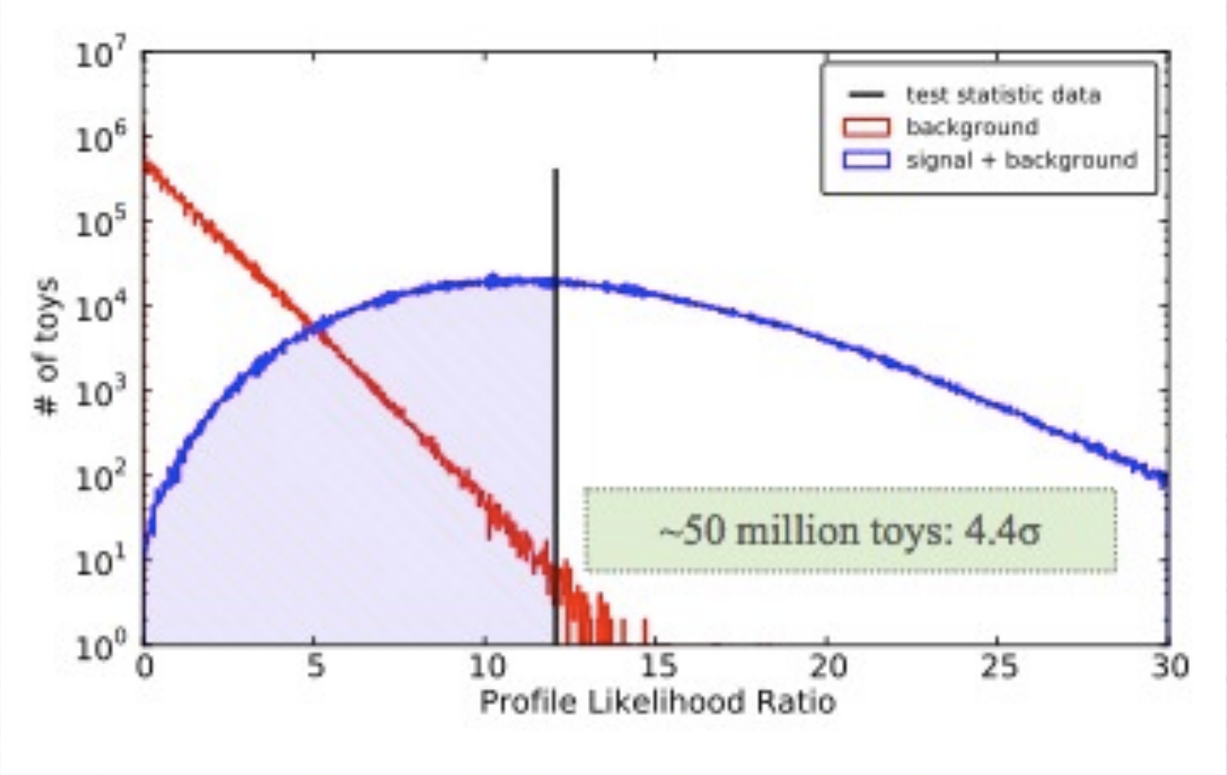
luminosity constraint

parameter constraint

HistFactory can be configured with XML files or directly in C++/Python (New in 5.34)

Parallelization

- Proof support for generating toys



- RooStats results are summarized in dedicated classes which have functionality for streaming and merging
  - used for running on grid jobs

Documentation

- RooStats TWiki: https://twiki.cern.ch/twiki/bin/view/RooStats/WebHome
- RooStats users guide (not really completed)
  - http://root.cern.ch/viewcvs/branches/dev/roostats/roofit/roostats/doc/usersguide/RooStats_UsersGuide.pdf
- For reference and citation: ACAT 2010 proceedings papers: http://arxiv.org/abs/1009.1003
- RooStats tutorial macros: http://root.cern.ch/root/html534/tutorials/roostats/index.html
- HistFactory document: https://cdsweb.cern.ch/record/1456844/files/CERN-OPEN-2012-016.pdf
- RooStats user support:
  - Request support via ROOT talk forum: http://root.cern.ch/phpBB2/viewforum.php?f=15 (questions on statistical concepts accepted)
  - contact me directly (email: Lorenzo.Moneta at cern.ch)
- Contacts for statistical questions:
  - ATLAS statistics forum:
    - TWiki: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/StatisticsTools
  - CMS statistics committee:
    - TWiki: https://twiki.cern.ch/twiki/bin/view/CMS/StatisticsCommittee