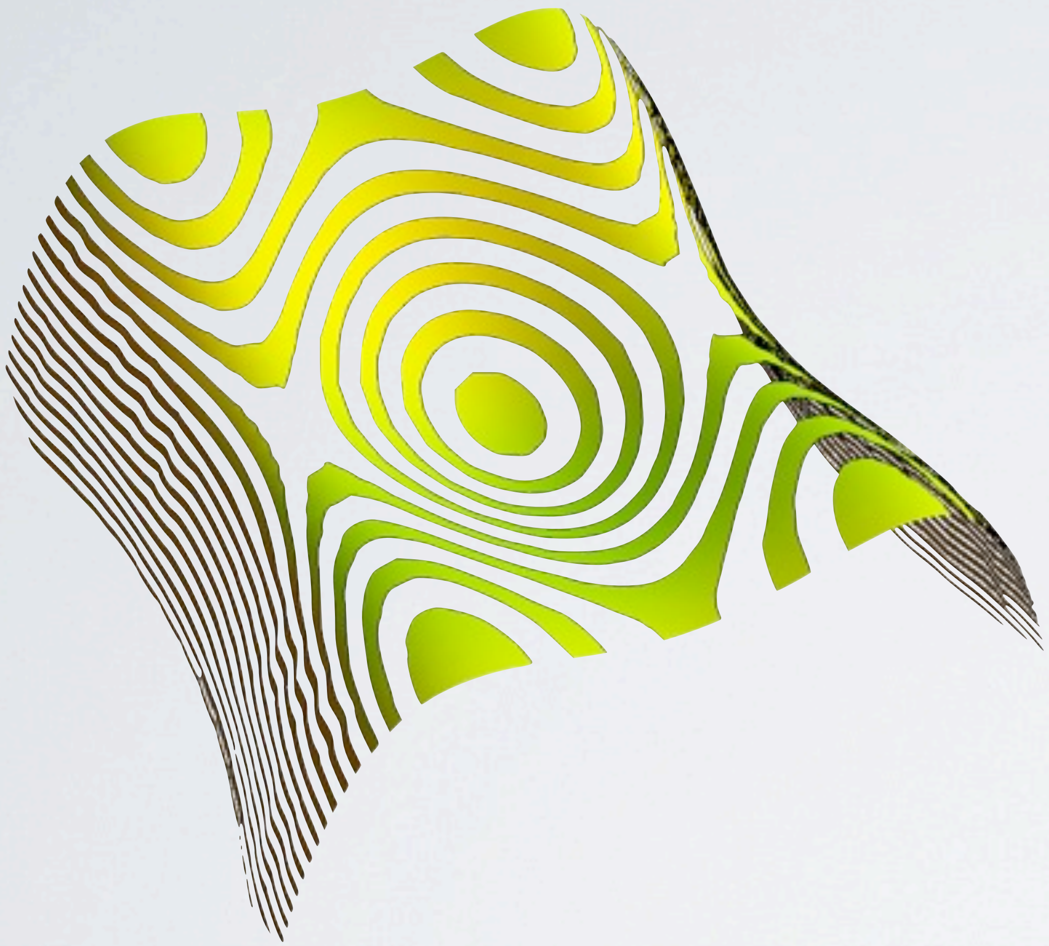
NEW YORK UNIVERSITY



RooStats

Kyle Cranmer, Sven Kreiss, Lorenzo Moneta
on behalf of the RooStats Team

CHEP 2012, New York, May 24

What is RooStats?

RooStats philosophy is to factorize the modeling from the statistical tests

- ➡ many other common statistical tools in HEP mix these two, give little flexibility in statistical tests
- ➡ statistical formalism is controversial (Bayesian and frequentist debate is not going to go away)
  - they are answering different questions, provide tools for both

Focus physicists attention on improving the modeling

- ➡ provide suite of statistical tools that can work on arbitrary models

Design:

- ➡ RooFit is used for Modeling and core functionality
- ➡ RooStats provides the higher-level statistical tests
  - interfaces mirror abstractions made by statisticians over last 80 years
- ➡ The two projects are intimately related

Outline

Building models, testing models, storing results

Problems that RooStats addresses:

- ➡ Discovery: one distribution.
- ➡ Limit: Scan and generate many distributions.

toy Monte Carlo techniques can be very slow

- ➡ PROOF
- ➡ importance sampling
- ➡ asymptotics

high dimensional parameter spaces

- ➡ MCMC, MultiNest
- ➡ pushing MINUIT to the limit

combining many complex analyses, correlating common parameters

Higgs Combination

RooStats Announcements

Collaborative project to provide and consolidate advanced statistical tools needed by LHC experiments.

Joint contribution from ATLAS, CMS, ROOT and RooFit: developments over sighted by ATLAS and CMS statistics committees.

> **"A method is more important than a discovery, since the right method will lead to new and even more important discoveries."**
> **- L.D. Landau**

Current Developers**:** K. Cranmer, G. Lewis, S. Kreiss (ATLAS), G. Schott, G. Kukartsev (CMS), Lorenzo Moneta (ROOT & CMS), Wouter Verkerke (RooFit & ATLAS), A. Lazzaro (OpenLab) Contributions from: K. Belasco, A. De Cosa, M. Pellicioni, D. Piparo, G. Petrucciani, S. Schmitz, M. Wolf, M. Baak

Included since ROOT v5.22; RooStats is developing fast and the latest stable version of ROOT is recommended: currently v5.34

Example macros in **$ROOTSYS/tutorials/roostats**

Citation: "The RooStats project", http://arxiv.org/abs/1009.1003 Proceedings of the ACAT2010 Conference

Underlying Technology

ROOT: General purpose data analysis tool.

RooFit: Modeling language which can build any binned and unbinned model. It also provides a *workspace* that includes the model and observed data that can be written to disk.

RooStats::HistFactory: Tool with a simple interface that takes ROOT histograms as input and builds a RooFit model.

Models can be written to disk and shared. Ideal for combining results.

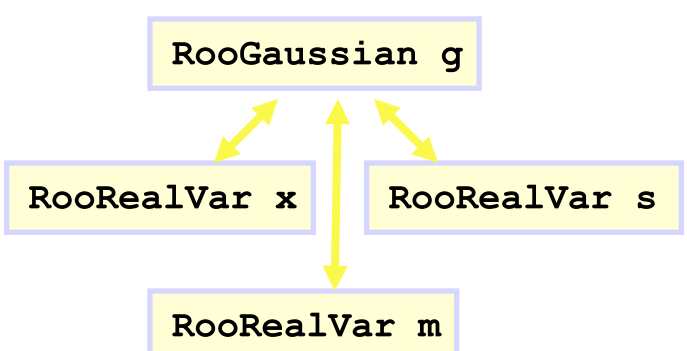
RooStats: Collection of statistical methods for any RooFit model. The methods are independent of the model and can be used with simple test cases as well as with large combined workspaces.

Underlying Technology II

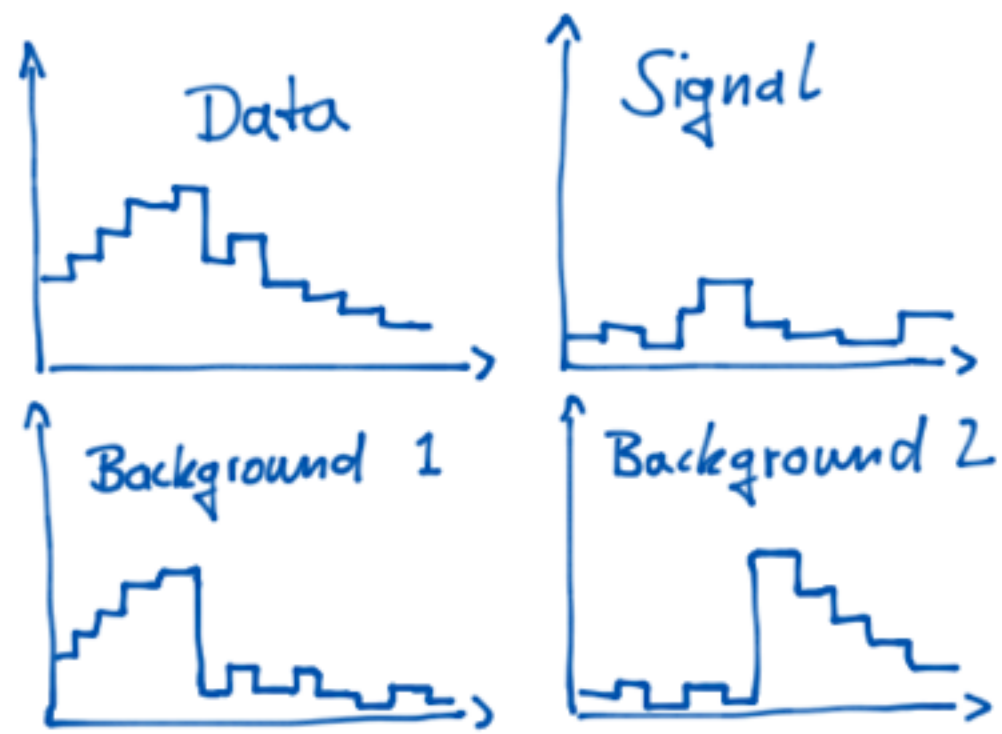
Building models with RooFit

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`

$Gauss(x|m,s)$



Building models with HistFactory



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

Many ways to build a model. All models can be used with all statistical methods.

HistFactory

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

luminosity constraint

parameter constraint

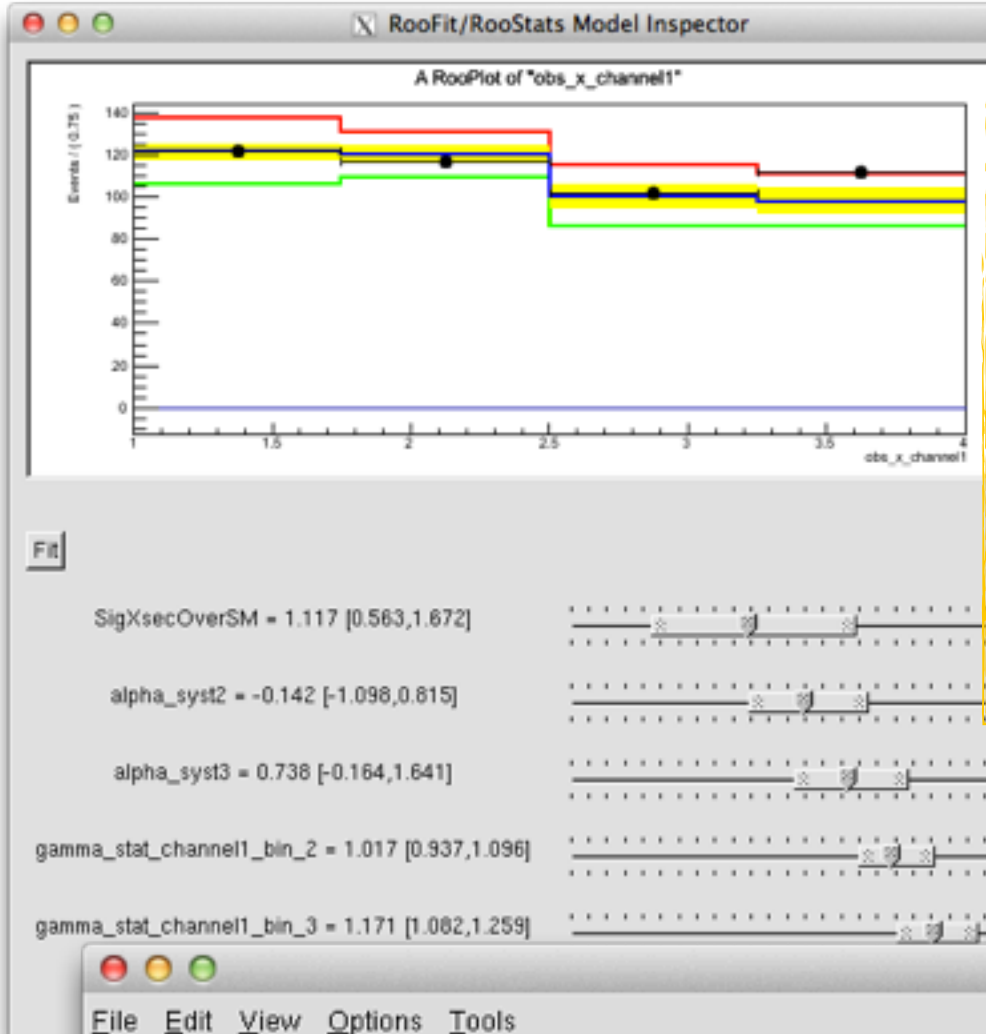
```
<!DOCTYPE Channel  SYSTEM 'HistFactorySchema.dtd'>

  <Channel Name="channel1" InputFile="./data/example.root" HistoName="" >
    <Data HistoName="data" HistoPath="" />
    <Sample Name="signal" HistoPath="" HistoName="signal">
      <OverallSys Name="syst1" High="1.05" Low="0.95"/>
      <NormFactor Name="SigXsecOverSM" Val="1" Low="0." High="3." Const="True" />
    </Sample>
    <Sample Name="background1" HistoPath="" NormalizeByTheory="True" HistoName="background1">
      <OverallSys Name="syst2" Low="0.95" High="1.05"/>
    </Sample>
    <Sample Name="background2" HistoPath="" NormalizeByTheory="True" HistoName="background2">
      <OverallSys Name="syst3" Low="0.95" High="1.05"/>
      <!-- <HistoSys Name="syst4" HistoPathHigh="" HistoPathLow="histForSyst4"/>-->
    </Sample>
  </Channel>
```

filename

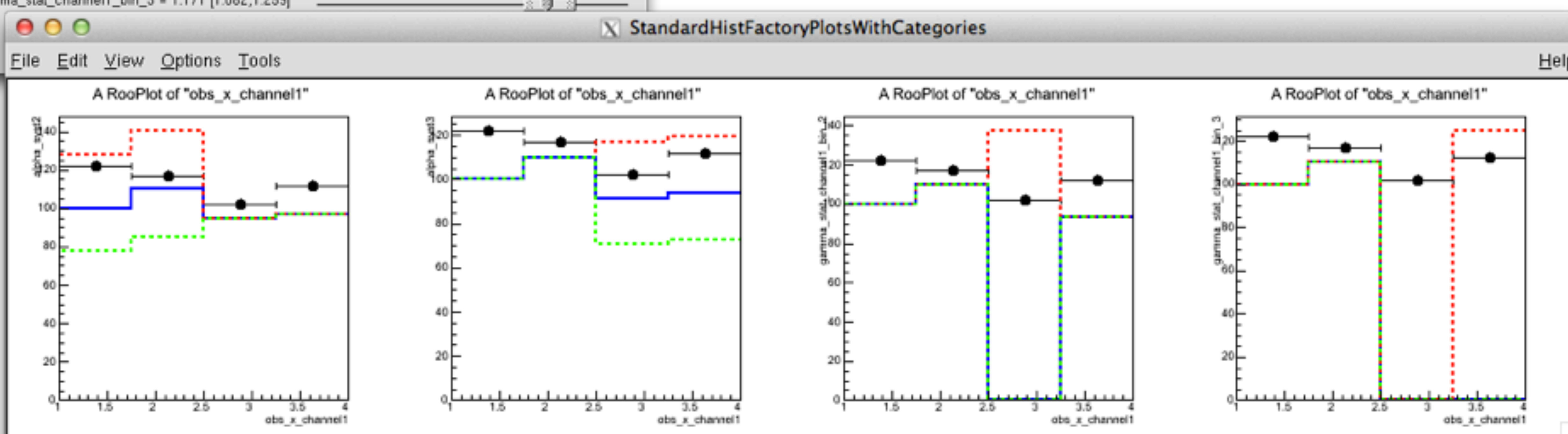
names of the histograms

Builds model from input histograms and an XML configuration like the one above. It can treat statistical uncertainties due to low Monte Carlo statistics properly. The new ROOT 5.34 also contains a **C++ and Python** interface. The configuration and all input histograms are now stored with the model and can be re-used. For example, the XML can be recreated from the model root file.

Investigating Models: ModelInspector.C, StandardHistFactoryPlotsWithCategories.C




```
shortened RooFit Print("t") of the model:
RooSimultaneous::newSimPdf = 0.23083
  RooProdPdf::model_channel1_newSimPdf = 0.23083
    RooRealSumPdf::channel1_model_newSimPdf = 220
      RooProduct::L_x_background1_channel1_StatUncert_newSimPdf = 0
        RooProduct::background1_channel1_StatUncert_newSimPdf = 0
          RooProduct::background1_channel1_Exp_newSimPdf = 0
            FlexibleInterpVar::background1_channel1_epsilon_newSimPdf = 1
              RooPolyVar::alphaOfBeta_syst2 = 0
                RooConstVar::-3.4801 = -3.4801
                RooConstVar::3.4801 = 3.4801
                RooRealVar::beta_syst2 = 1
            RooHistFunc::background1_channel1_nominal = 0
              RooRealVar::obs_x_channel1 = 1.75
```



Upper and lower limits

The statistical methods described in "Procedure for the LHC Higgs boson search combination in Summer 2011" [ATL-PHYS-PUB-2011-11, CMS NOTE-2011/005] are implemented in RooStats.

Common classes of tools:

- ➡ point estimation: determine the best estimate of a parameter
- ➡ estimation of confidence (credible) intervals
  - multi-dimensional contours or just a lower/higher limit
- ➡ hypothesis tests: evaluation of p-value for one or multiple hypotheses (significance)
- ➡ goodness-of-fit: how well a model describes the data

Analysis combination:

- ➡ Performed using the full information contained in the Likelihood function

RooStats: all Calculators

HypoTestCalculators

- AsymptoticCalculator
  - calculates a p-value according to an analytic expression for the asymptotic form of the test statistic distribution
- FrequentistCalculator
  - frequentist calculation (profile nuisance parameters)
- HybridCalculator
  - hybrid Bayes-Frequentist calculation (marginalize nuisance parameters)
- ProfileLikelihoodCalculator
  - the method of MINUIT/MINOS, based on Wilks' theorem

IntervalCalculators

- HypoTestInverter
  - takes a HypoTestCalculator and forms an IntervalCalculator
- ProfileLikelihoodCalculator
  - method of MINUIT/MINOS, based on Wilks' theorem
- NeymanConstruction
  - general purpose Neyman Construction class, highly configurable: choice of TestStatistic, TestStatSampler (defines ensemble/conditioning), integration boundary (upper, lower, central limits), and parameter points to scan
- FeldmanCousins
  - specific configuration of NeymanConstruction for Feldman-Cousins (generalized for nuisance parameters)
- MCMCCalculator
  - Bayesian Markov Chain Monte Carlo (Metropolis Hastings), proposal function is highly customizable
- BayesianCalculator
  - Bayesian posterior calculated via numeric integration routines, currently only supports one parameter

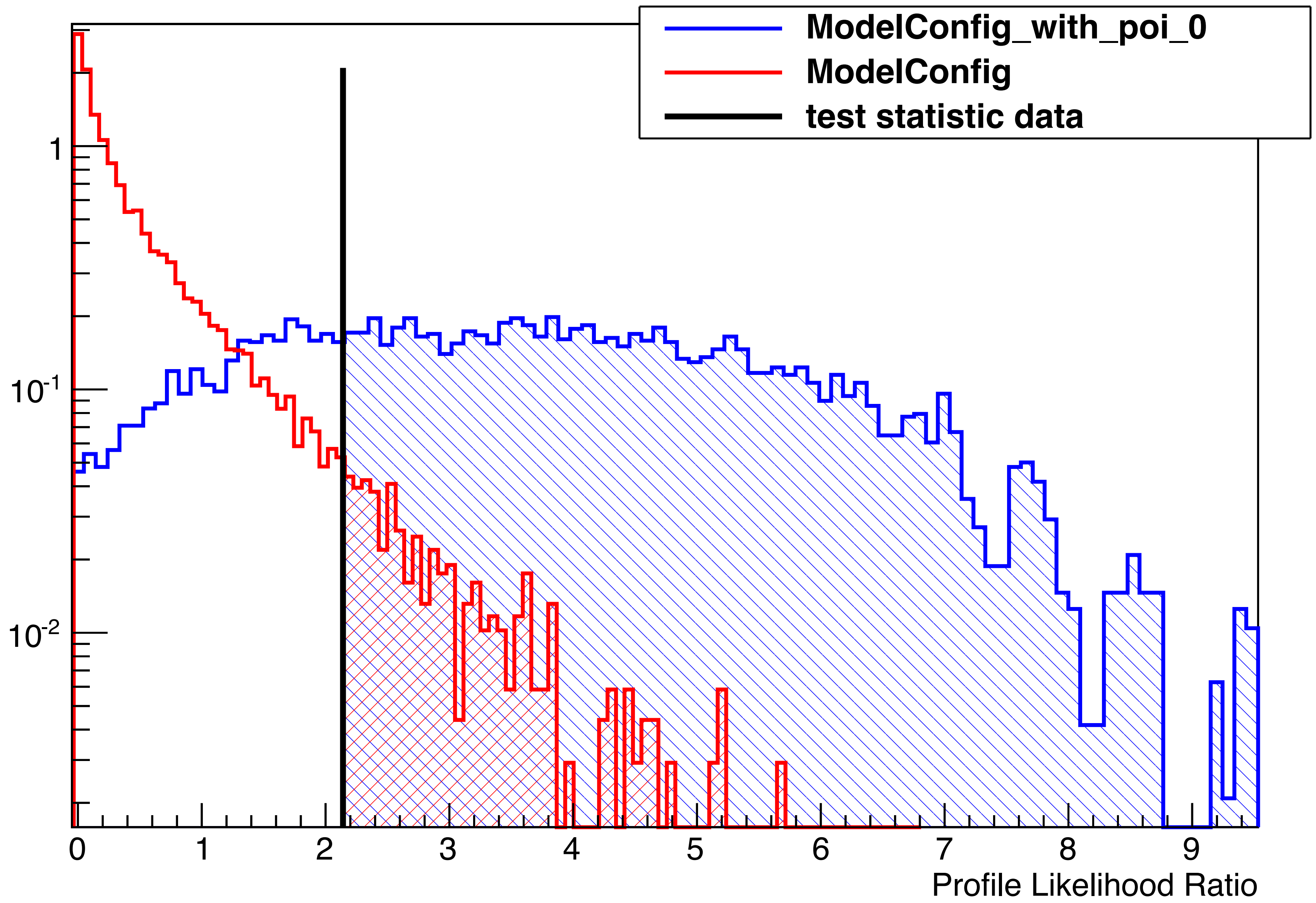
RooStats and LHC statistics

At the TeVatron, every input is binned. HistFactory handles those binned inputs and creates statistical models in a simple way.
RooFit and RooStats can handle binned, unbinned and mixed models.

For most statistical tools like MCLimit, Collie, TLimit, ... the model is tied to the method.

RooStats interfaces follow abstractions made by statisticians over decades.

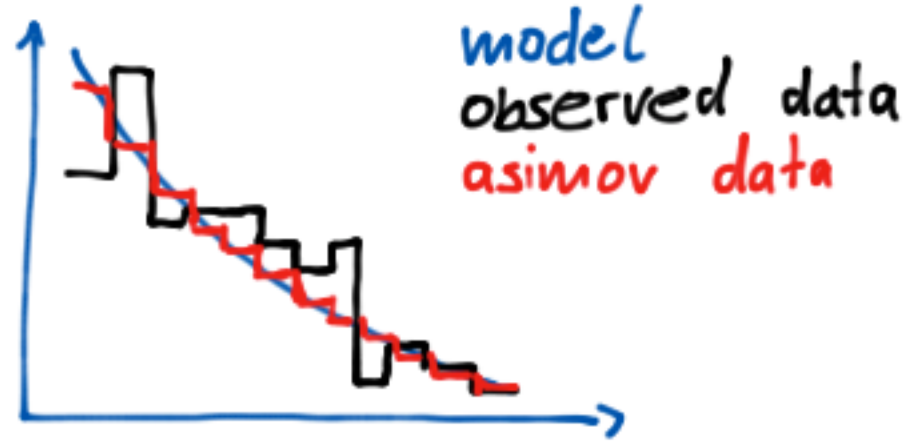
FrequentistCalculator, ToyMCSampler



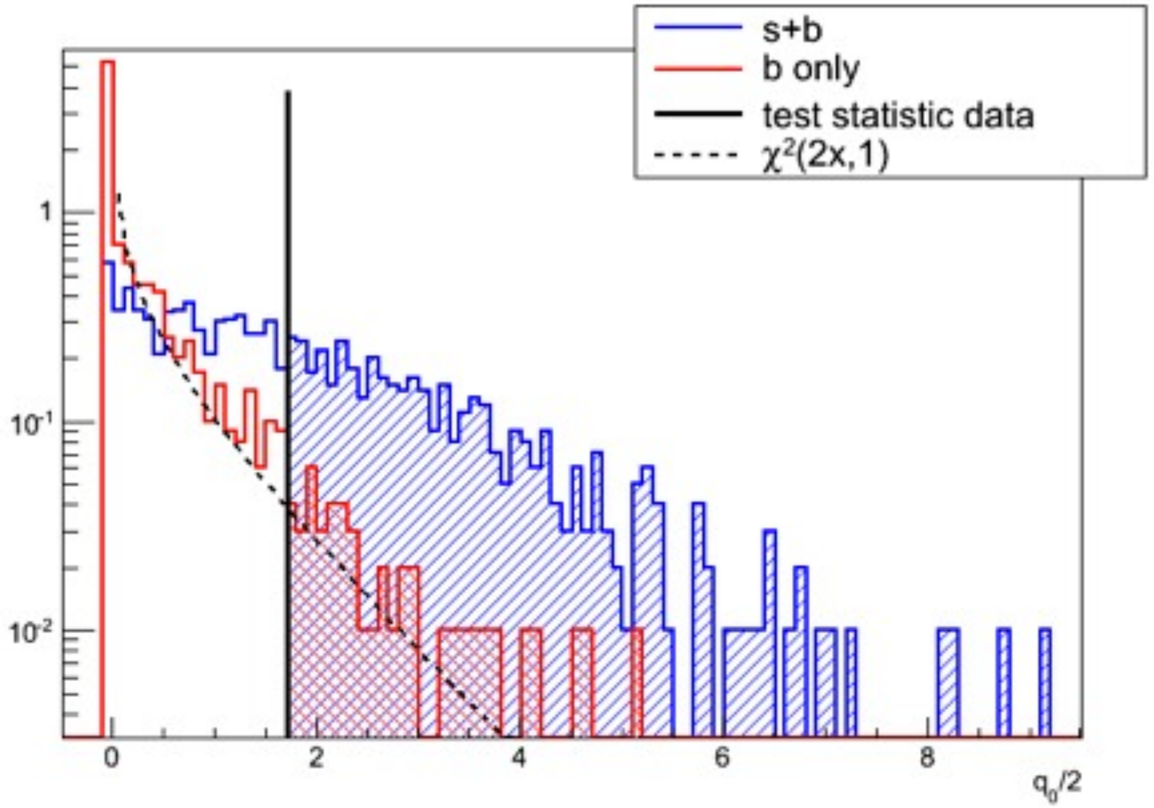
`FrequentistCalculator` calculates LHC-style p-values and follows the recommendations for the treatment of nuisance parameters. The `ToyMCSampler` is generating the distributions which also supports PROOF for parallelization.

Asymptotic Calculator [based on "Asimov Paper": Cowan, Cranmer, Gross, Vitels, arXiv:1007.1727]

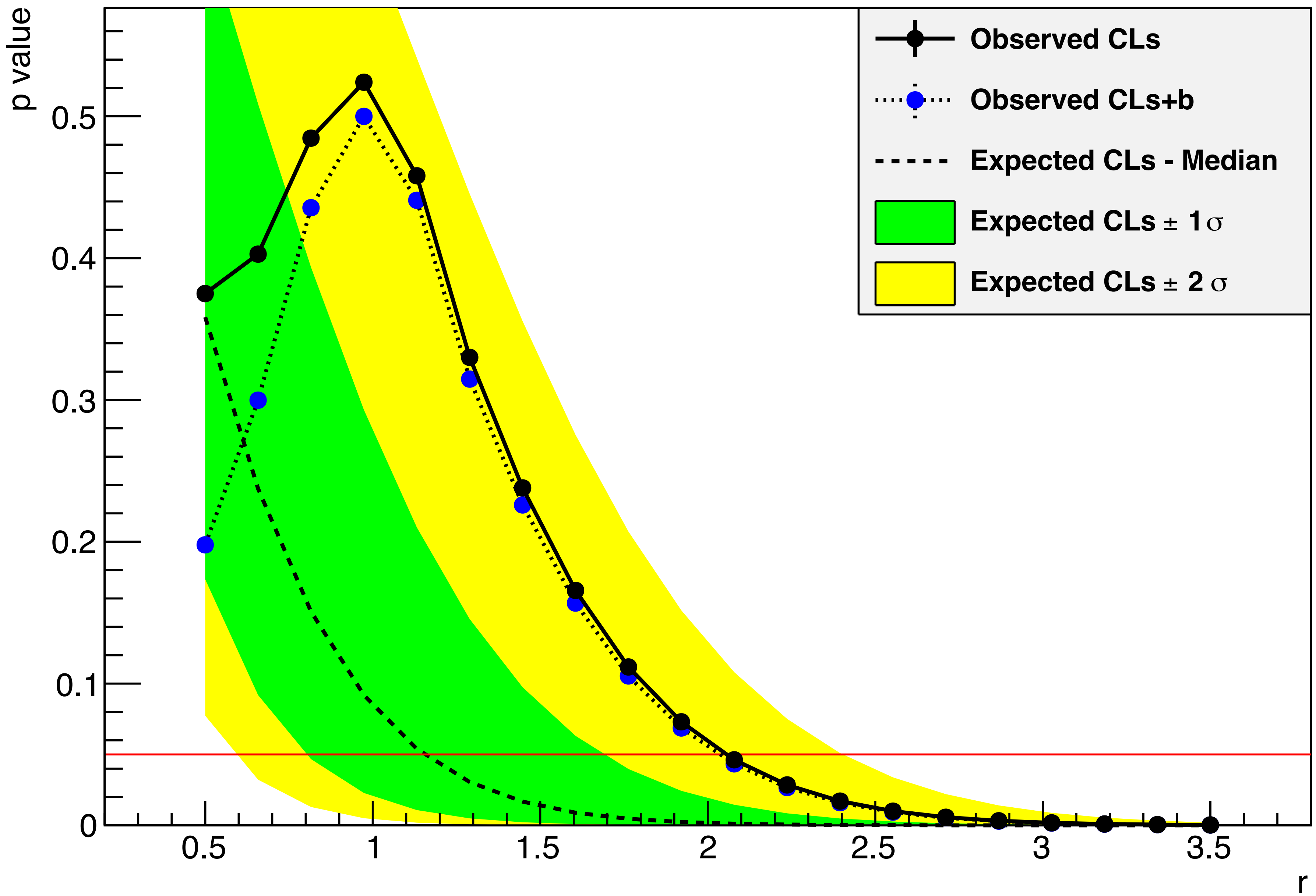
Can generate expected data (Asimov data).



Very fast, but the current asymptotic formulas do not describe all aspects of the model (e.g. energy scale uncertainties).

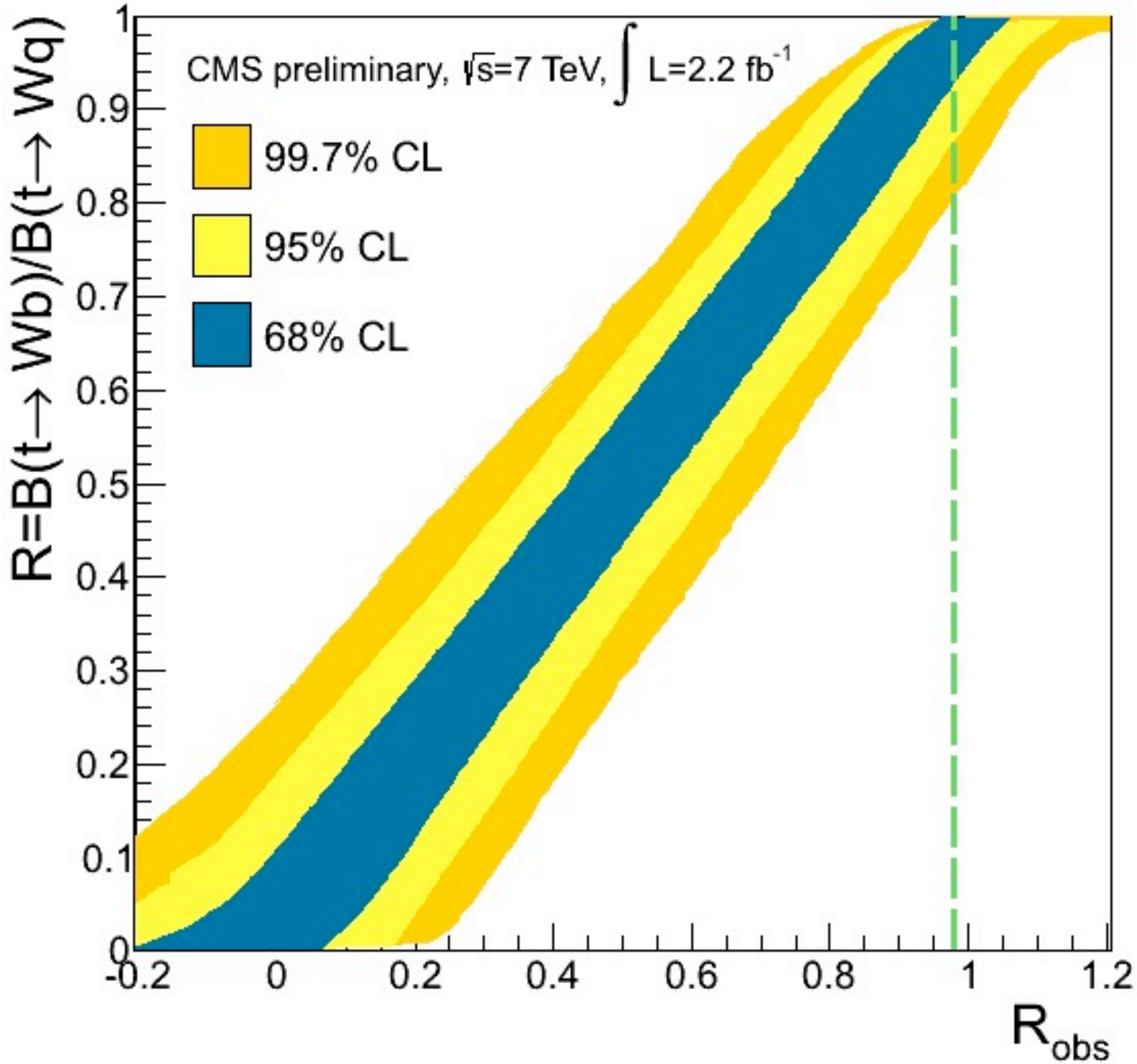


HypoTestInverter: CLs and PCL limits using StandardHypoTestInvDemo.C

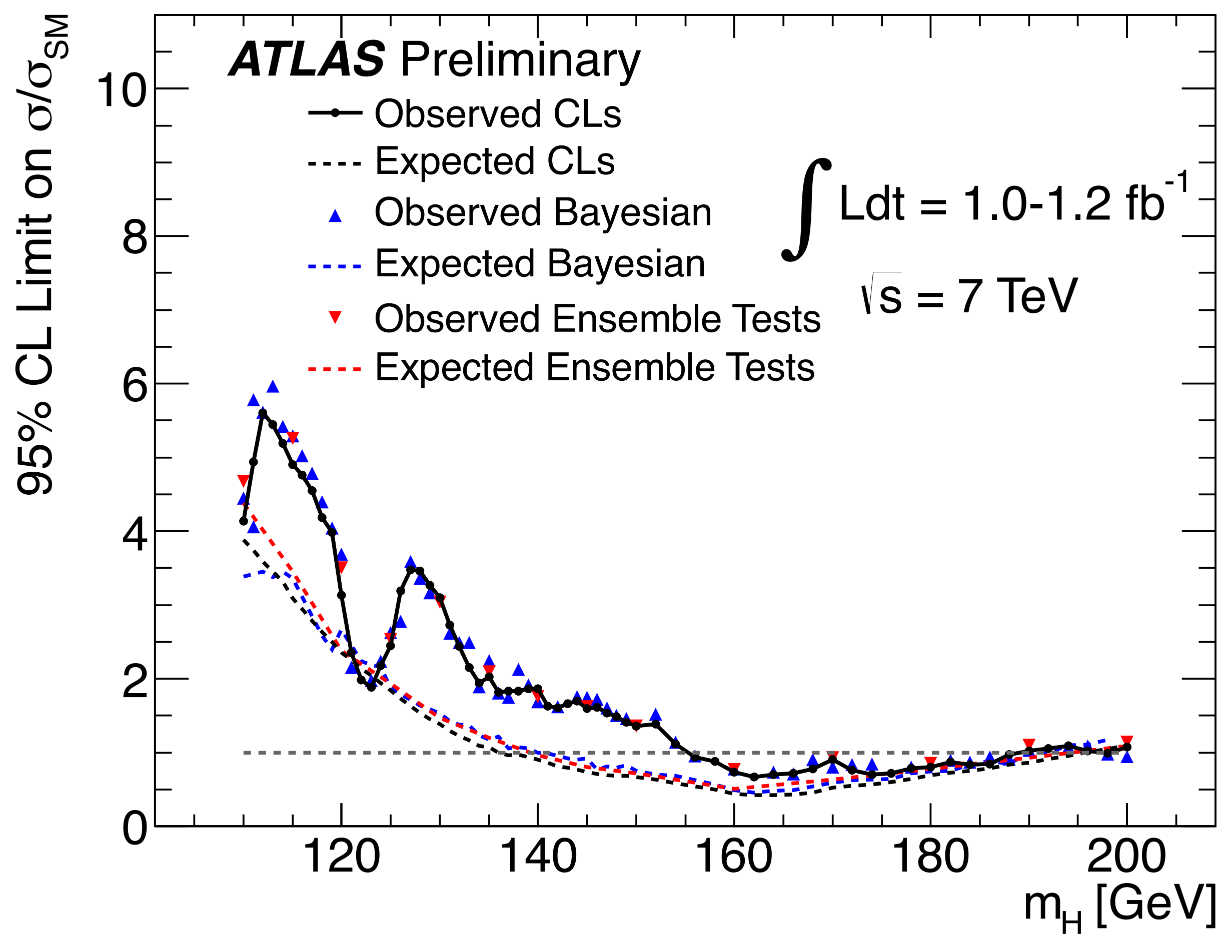


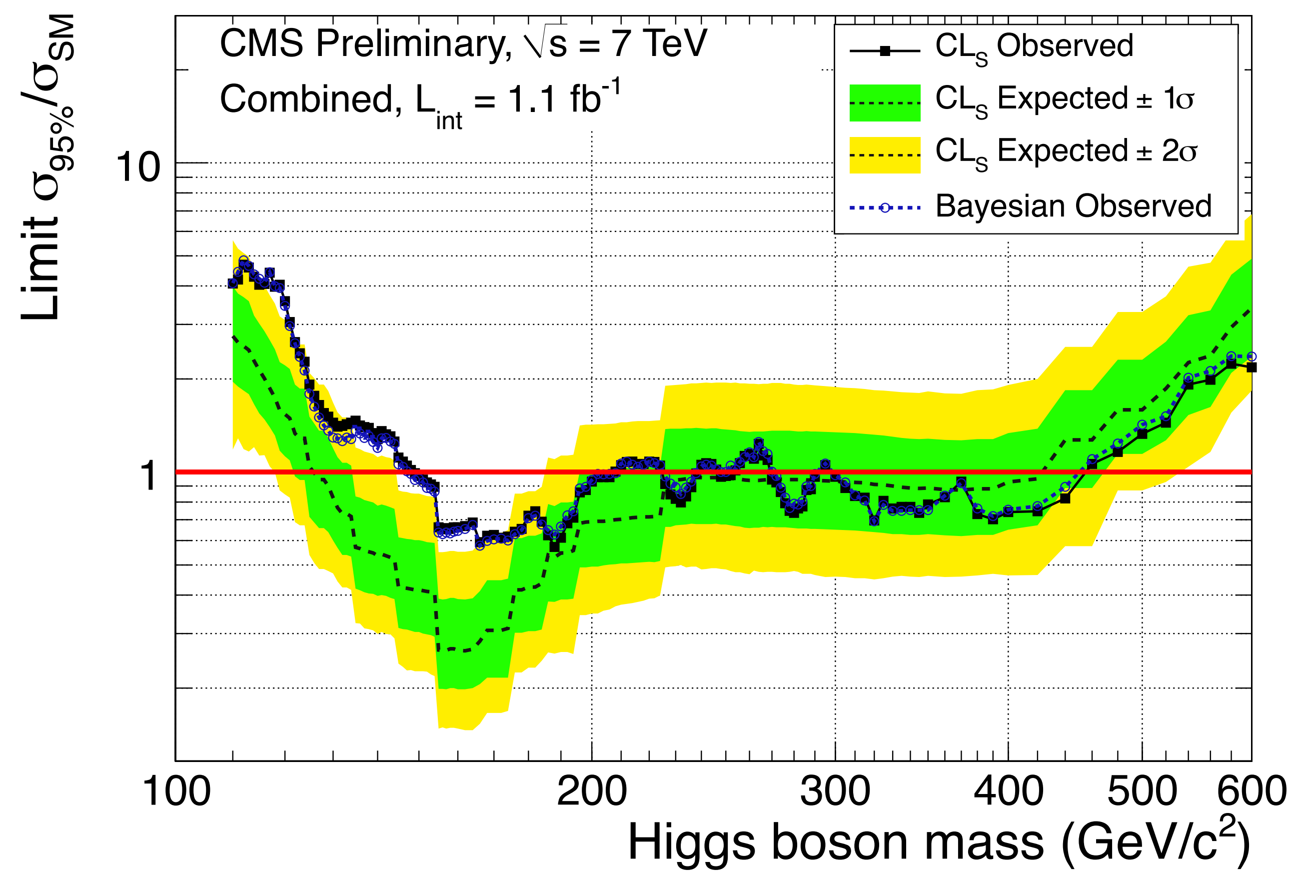
This example is done using the AsymptoticCalculator to generate the individual distributions. Can also be done with toys using the Frequentist- and HybridCalculator.

HypoTestInverter: Feldman-Cousins limits using StandardHypoTestInvDemo.C



One Model, Many Methods

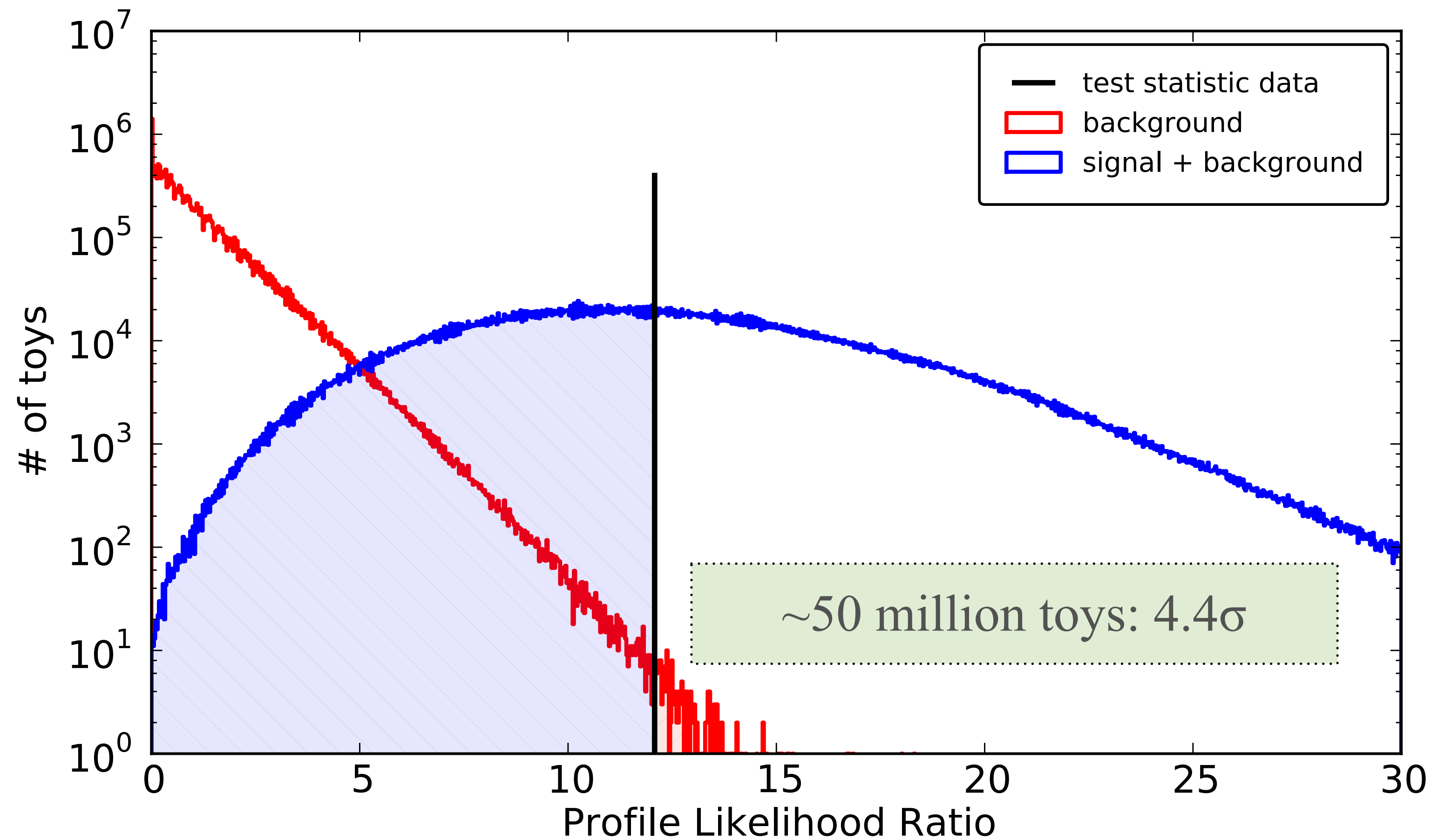




An example how well three different statistical tools agree. The results are produced from the same input model.

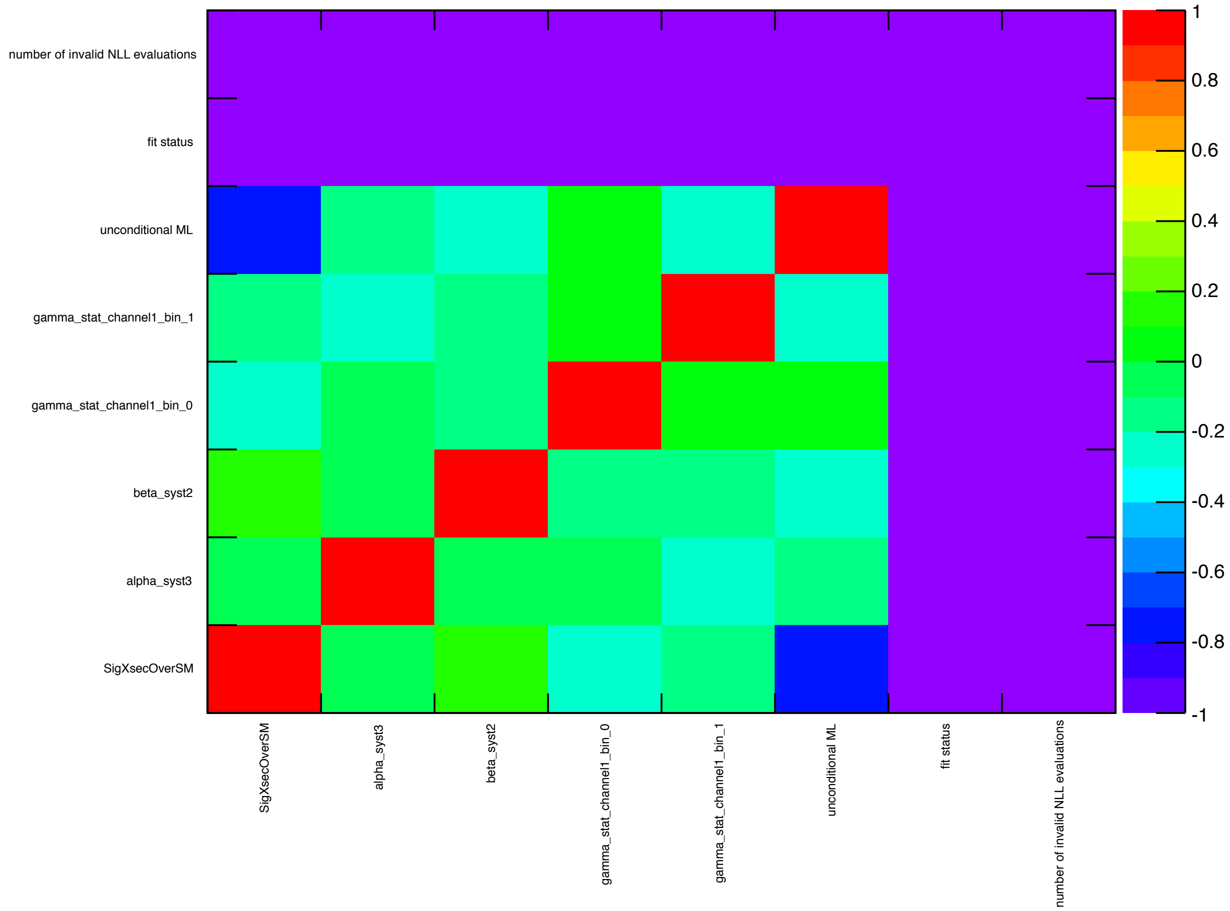
Parallelization and Managing Results

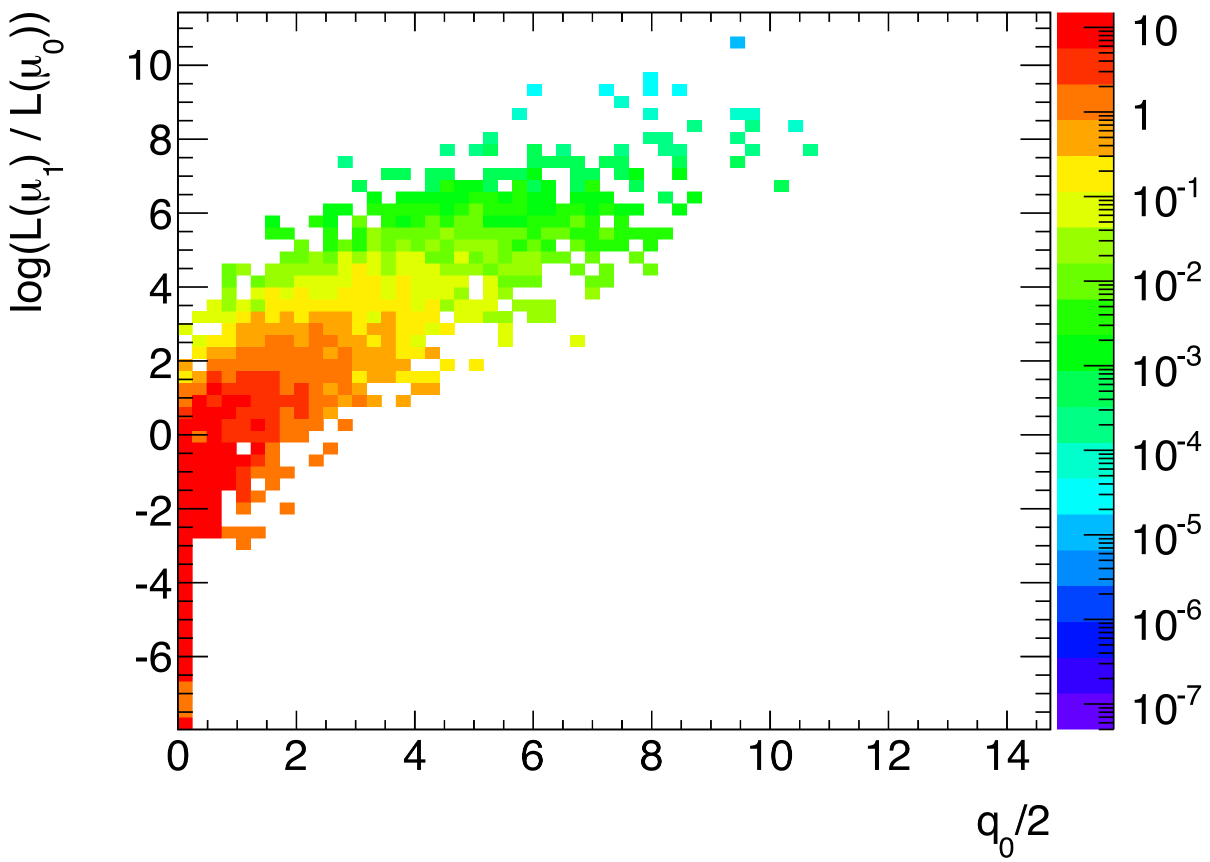
PROOF: all toy studies can be parallelized to run on many cores in a single machine and to clusters (for example when working with a Tier3).



Full results are stored in dedicated result classes. They can be streamed and stored and have built-in functionality to merge themselves (for example for Condor, PBS or Grid jobs).

New Developments for Toys: Detailed Output and Multiple Test Statistics





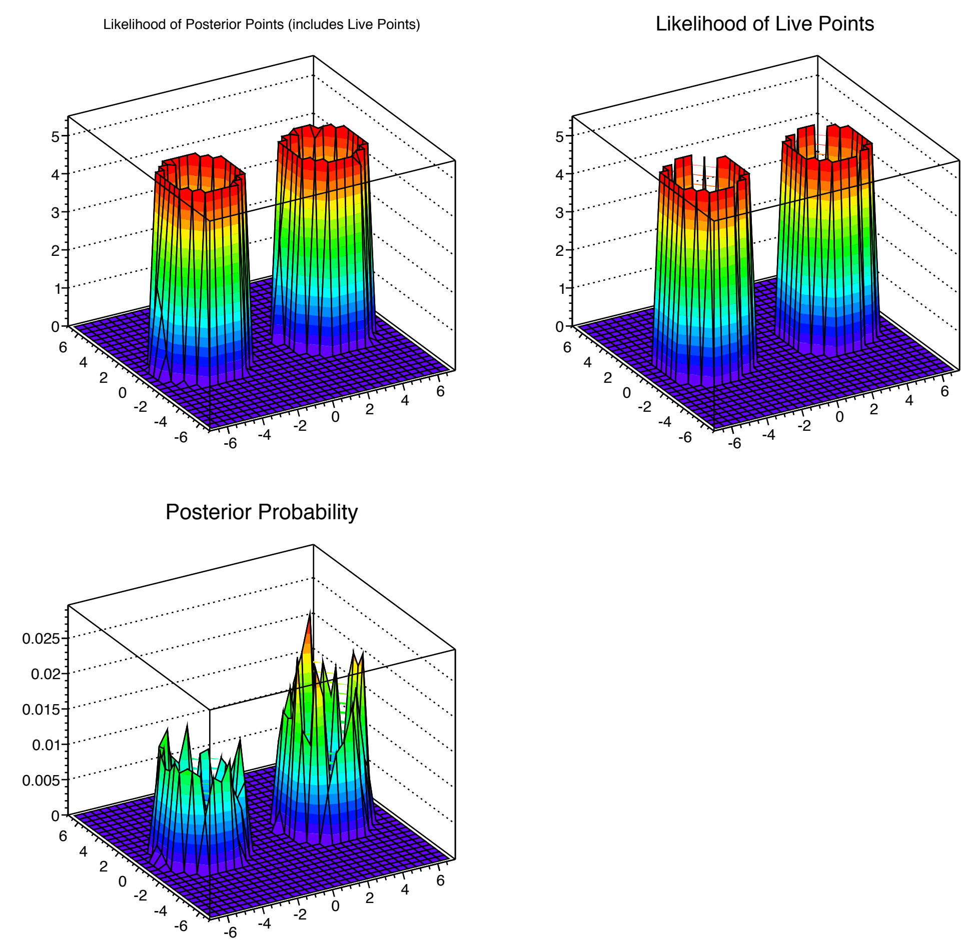
New Developments for Bayesian Methods: MCMC and MultiNest

The result class for MCMC, `MarkovChain`, has better support for merging now. The Bayesian tools support arbitrary priors. MCMC works with high dimensional models (used with ~400 parameters in the ATLAS Higgs combination and the CMS Higgs combination).

A completely new tool to RooStats is a connector to MultiNest which is a multimodal nested sampler [Feroz, Hobson, Bridges, http://arxiv.org/abs/0809.3437].

Also here, the model is separate from the method and can be created with the Workspace factory:

```
RooWorkspace w("rings");
   w.factory("x[0,-10,10]");
   w.factory("y[0,-10,10]");
   w.factory("z[0,-10,10]");
   RooArgList axes(*w.var("x"), *w.var("y"),
*w.var("z"));
   w.factory("expr::v1('sqrt((@0-@1)*(@0-@1)+(@2-
@3)*(@2-@3))', x, c11[-3.5], y, c12[0])");
   w.factory("RooGaussian::g1(v1,r1[2],w1[.1])");
   w.factory("expr::v2('sqrt((@0-@1)*(@0-@1)+(@2-
@3)*(@2-@3))', x, c21[3.5], y, c22[0])");
   w.factory("RooGaussian::g2(v2,r2[2],w2[.1])");
   w.factory("ASUM::model(5*g1,5*g2)");
```



Importance Sampling: Intro

Based on ideas from Alex Read and Michael Woodroofe

- ➡ Michael Woodroofe's talk at Banff 2010: http://people.stat.sfu.ca/~lockhart/richard/banff2010/woodroofe.pdf where he covered importance sampling and a method to create a suitable importance density.

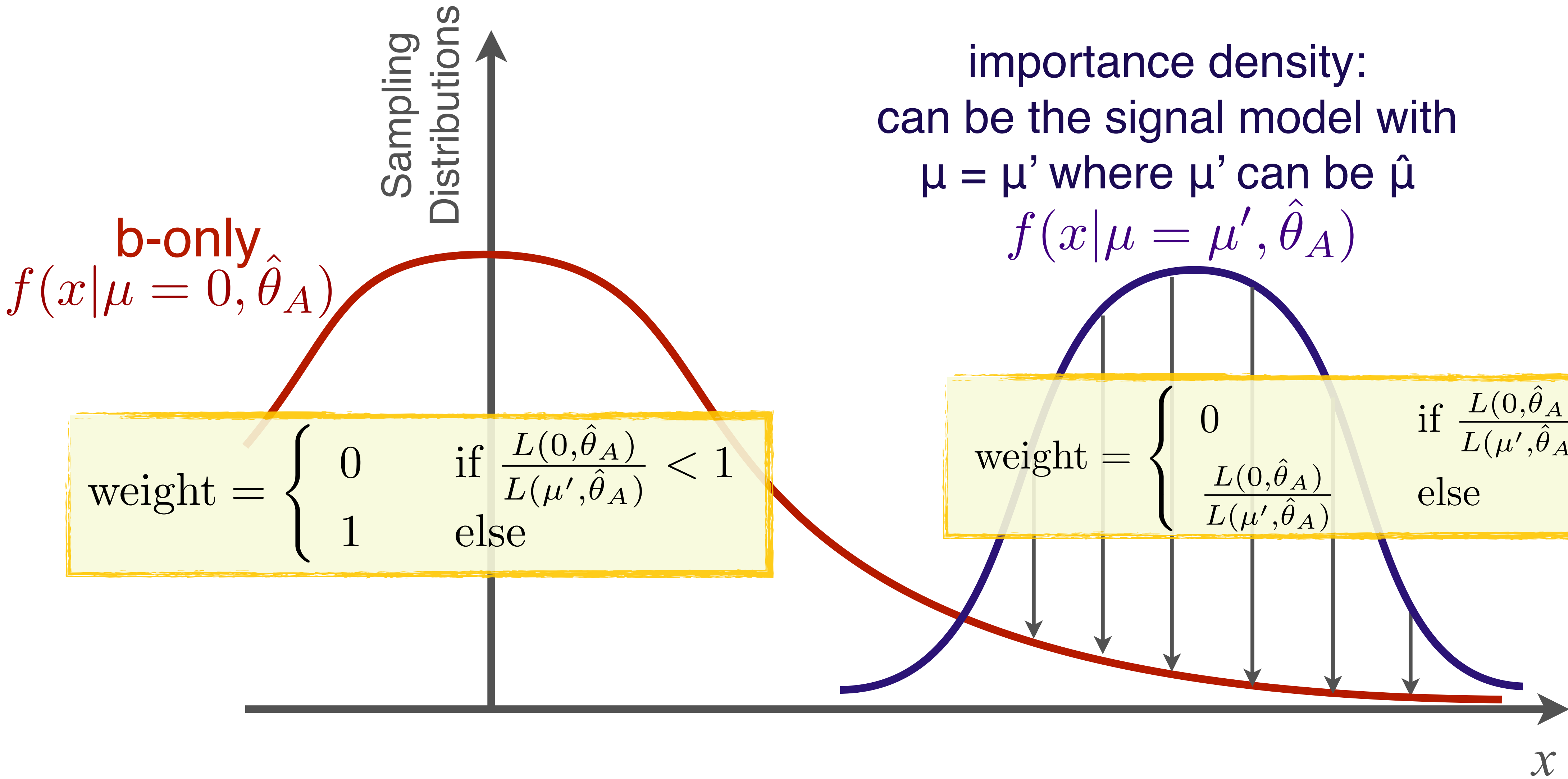
Technical: Importance Sampling is implemented in RooStats::ToyMCImportanceSampler. The standard RooStats::ToyMCSampler is extended to handle this and "detailed output."

Now in ROOT 5.33

Idea: use a different density (importance density) to generate toys, but re-weight the result according to the ratio of their Likelihoods.

- ➡ Can populate a small tail with few toys.

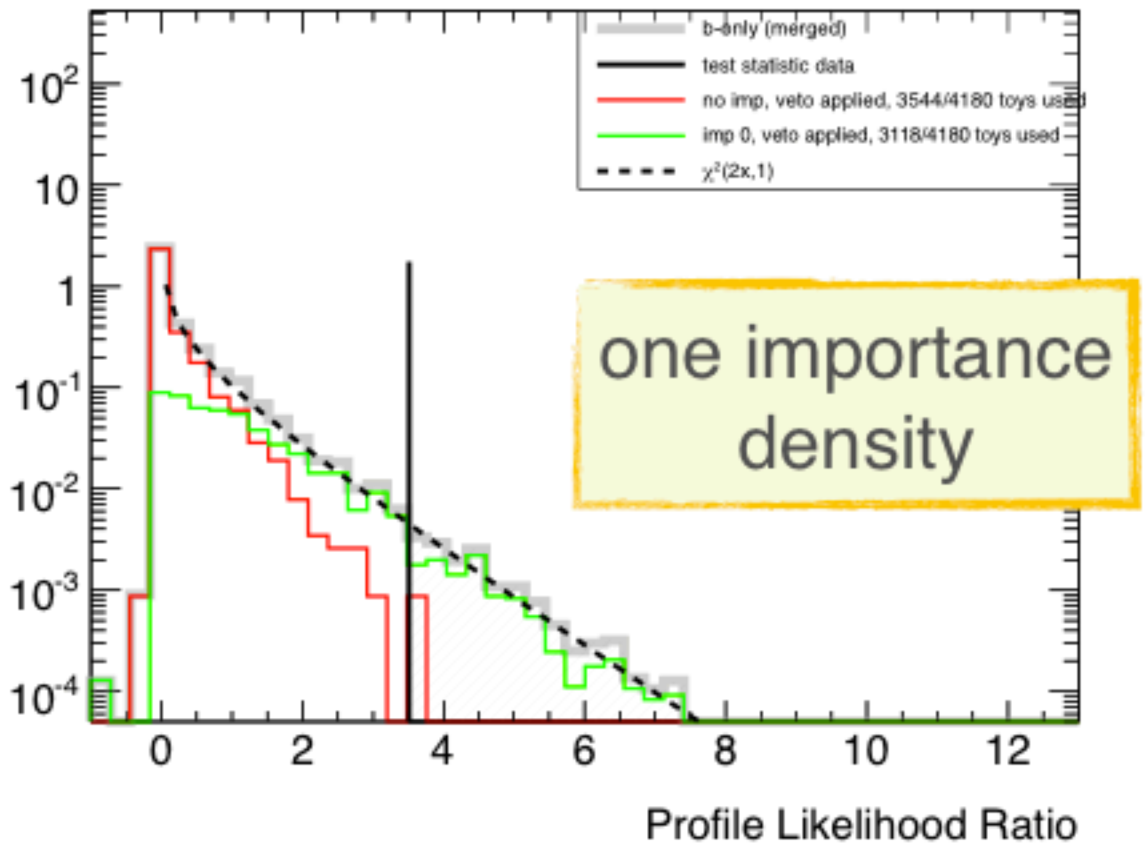
Importance Sampling with One Importance Density

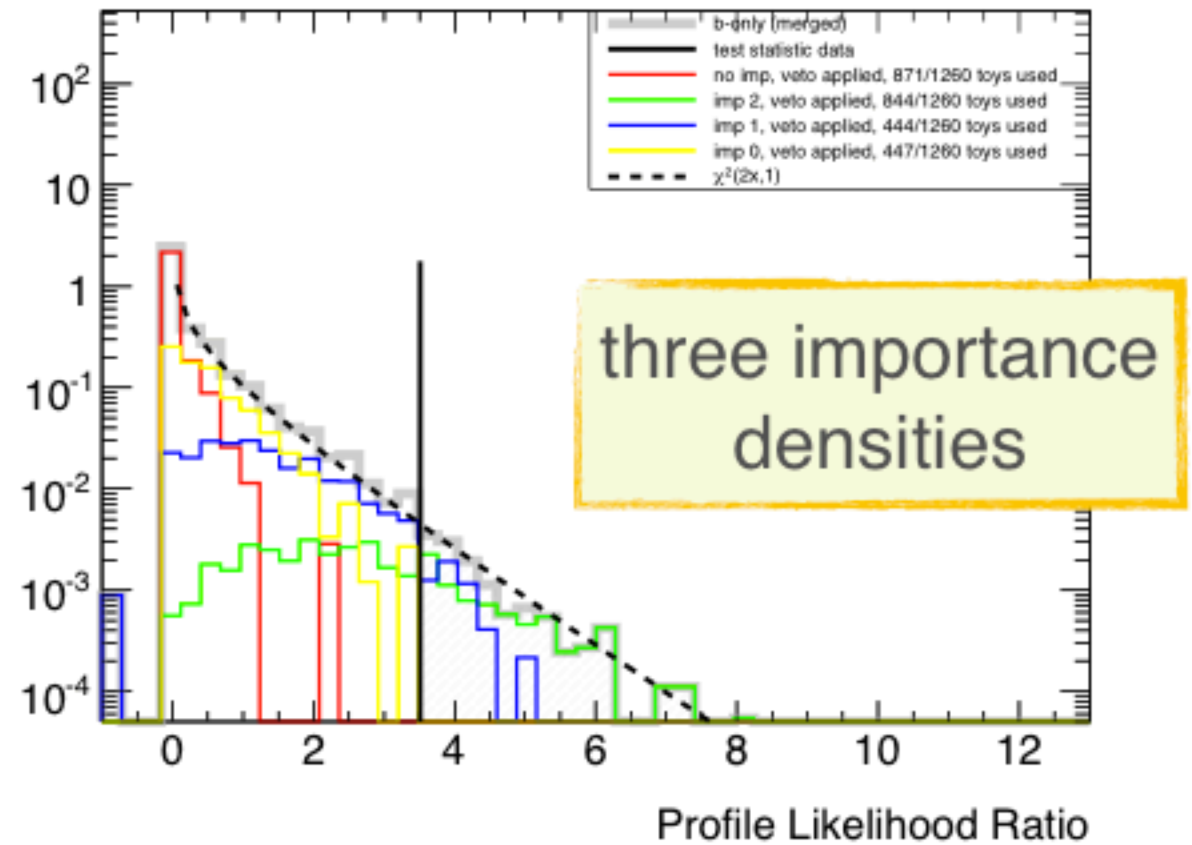


merged distribution = toys from b-only + toys from imp dens

A point is only used in the merged sampling distribution when its Likelihood is the largest of the tested densities.

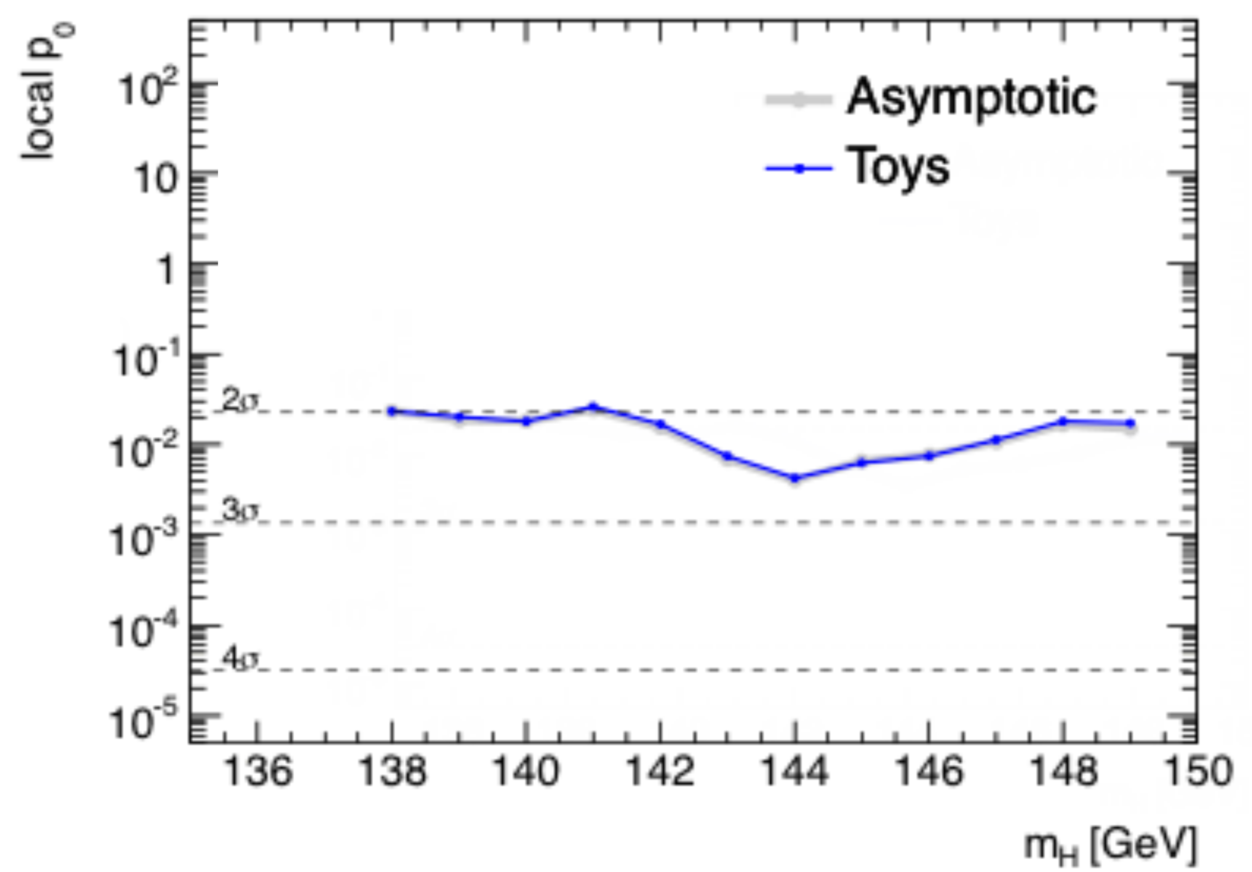
Some Tests





The agreement is good between toys from Importance Sampling and Asymptotics.

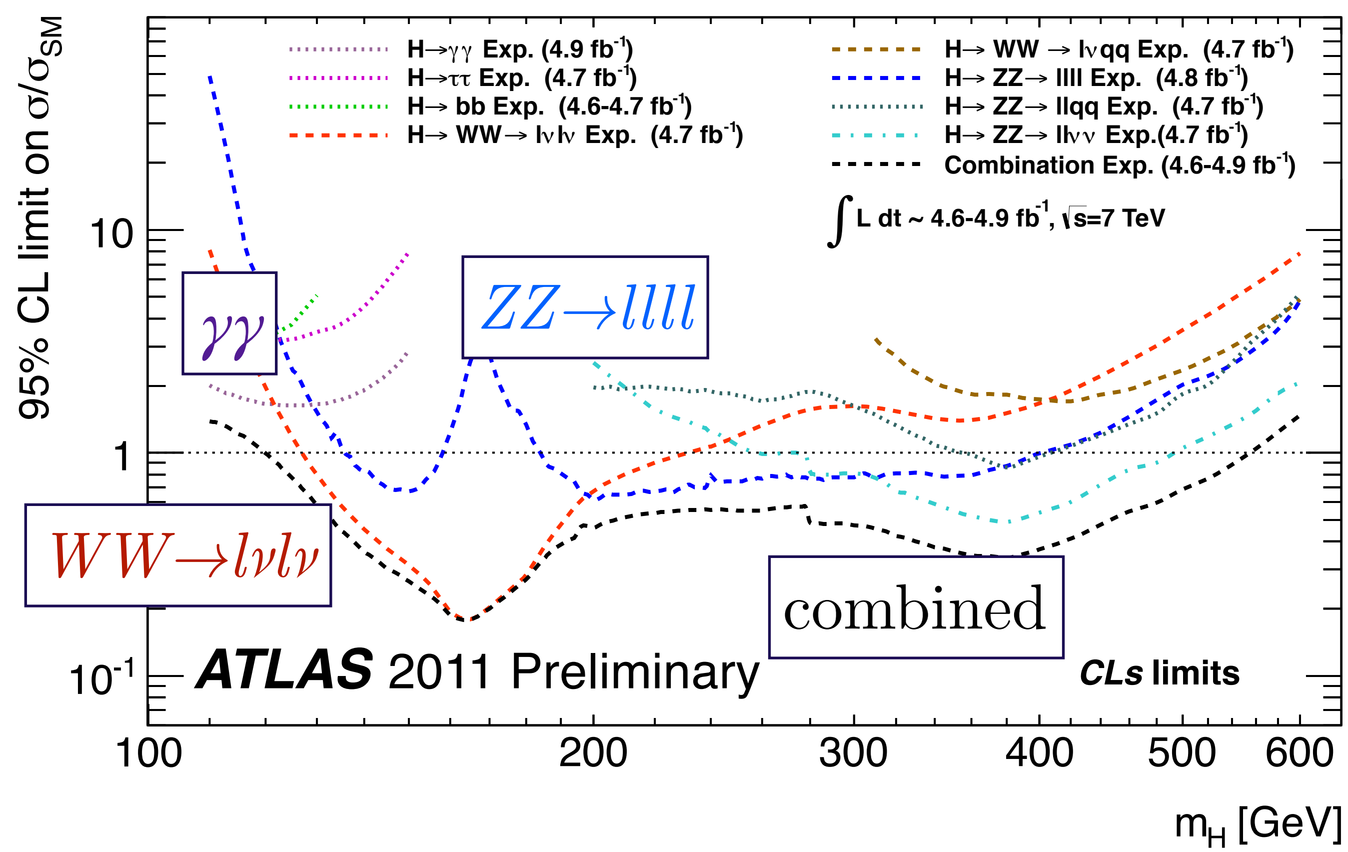
This is also validated with ATLAS Council results at the ~3.5σ level.



Higgs Boson: The only undiscovered particle in the Standard Model. It's mass m_H is the only fundamental parameter of the Standard Model that is unknown. Understanding electroweak symmetry breaking is one of the primary objectives of the LHC. Standard Model and non-Standard Model scenarios are considered.

There are various options for production and decay of the Higgs. Their importance depends on m_H. For the currently interesting region, the production through gluon-gluon-Fusion and Vector-Boson-Fusion and the decays to $\gamma\gamma$, $ZZ\rightarrow llll$ and $WW\rightarrow l\nu l\nu$ contribute the most.

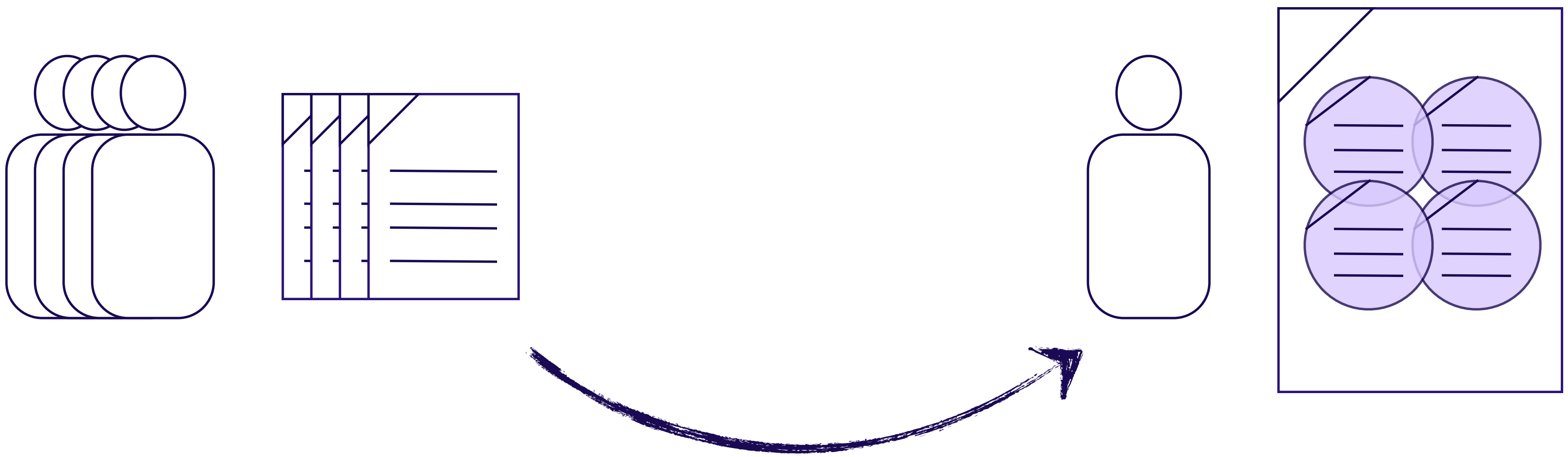
Expected upper limits are related to the "weight" of each channel in the combination:
[ATLAS-CONF-2012-019, 5 March 2012]



Combination: a Collaborative Effort

one subgroup per decay mode:
$\gamma\gamma$, ZZ, WW, ...

combined model includes
common parameters



subgroups provide digital form of their
model to combination group

Combination group creates combined models from 12 groups with 68 sub channels containing a total of over 400 parameters.

The combination between 2011 and 2012 data will be an extra challenge because some parts of the models are independent and this will essentially double the number of nuisance parameters.

Combined results: upper limits, discovery significances and best fit signal strengths.

The combination group is now also working on methods for property determinations for a possible low mass Higgs.

ATLAS+CMS Combination Result Summer 2011

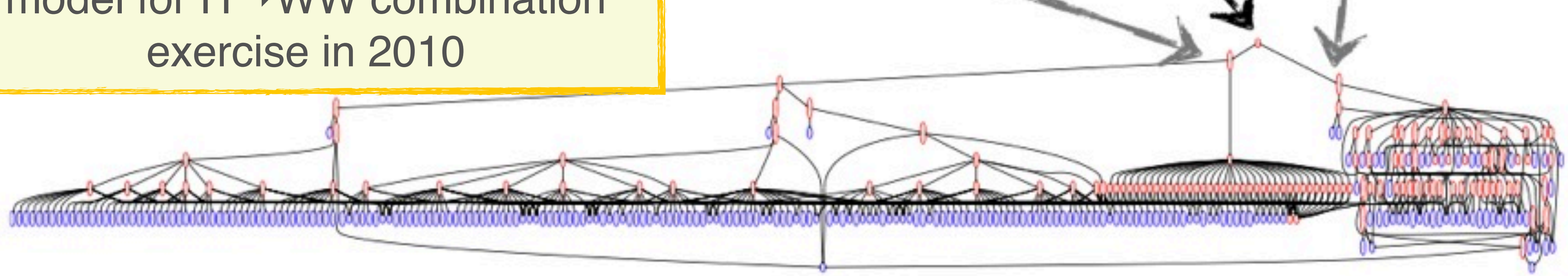
The full model has 12 observables and ~50 parameters

top level model

ATLAS part

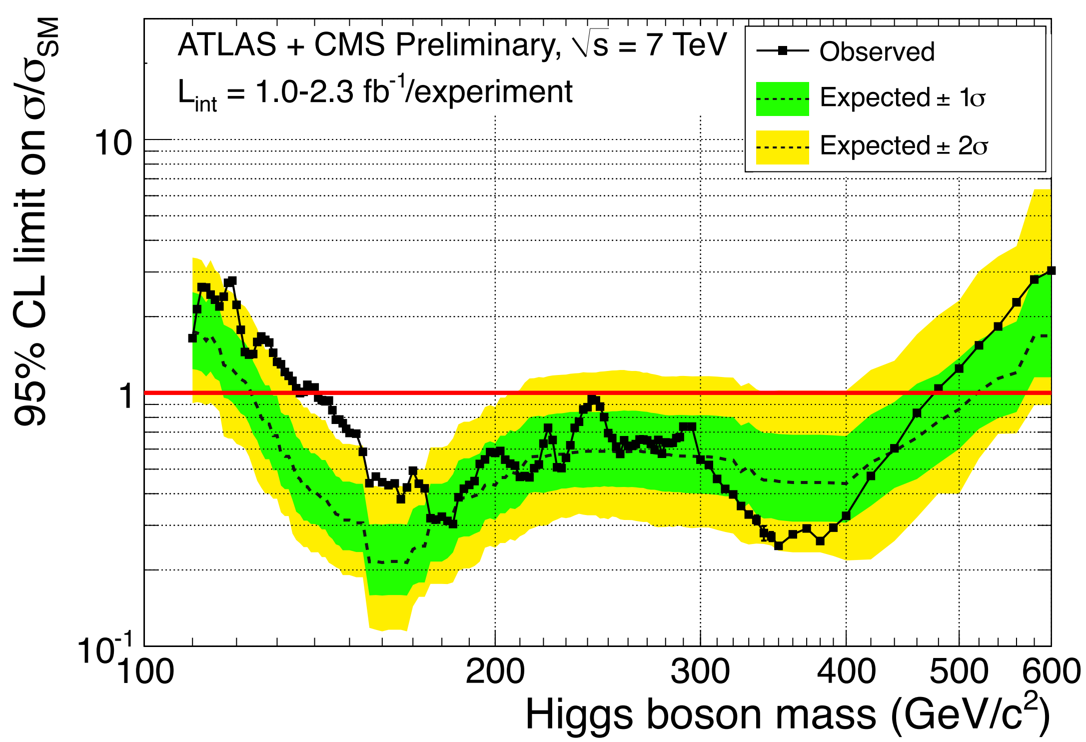
CMS part

model for H→WW combination exercise in 2010



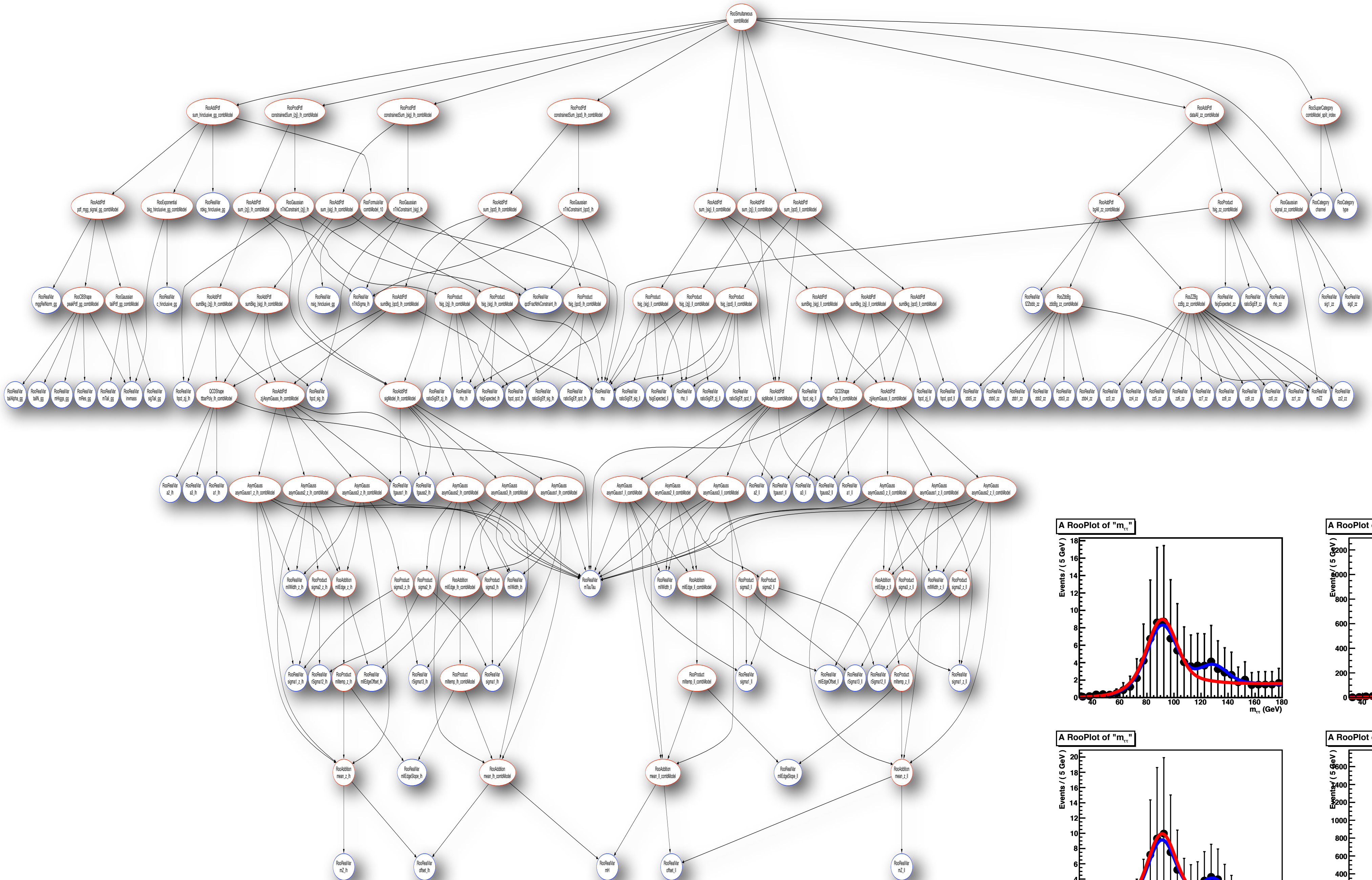
parameter of interest

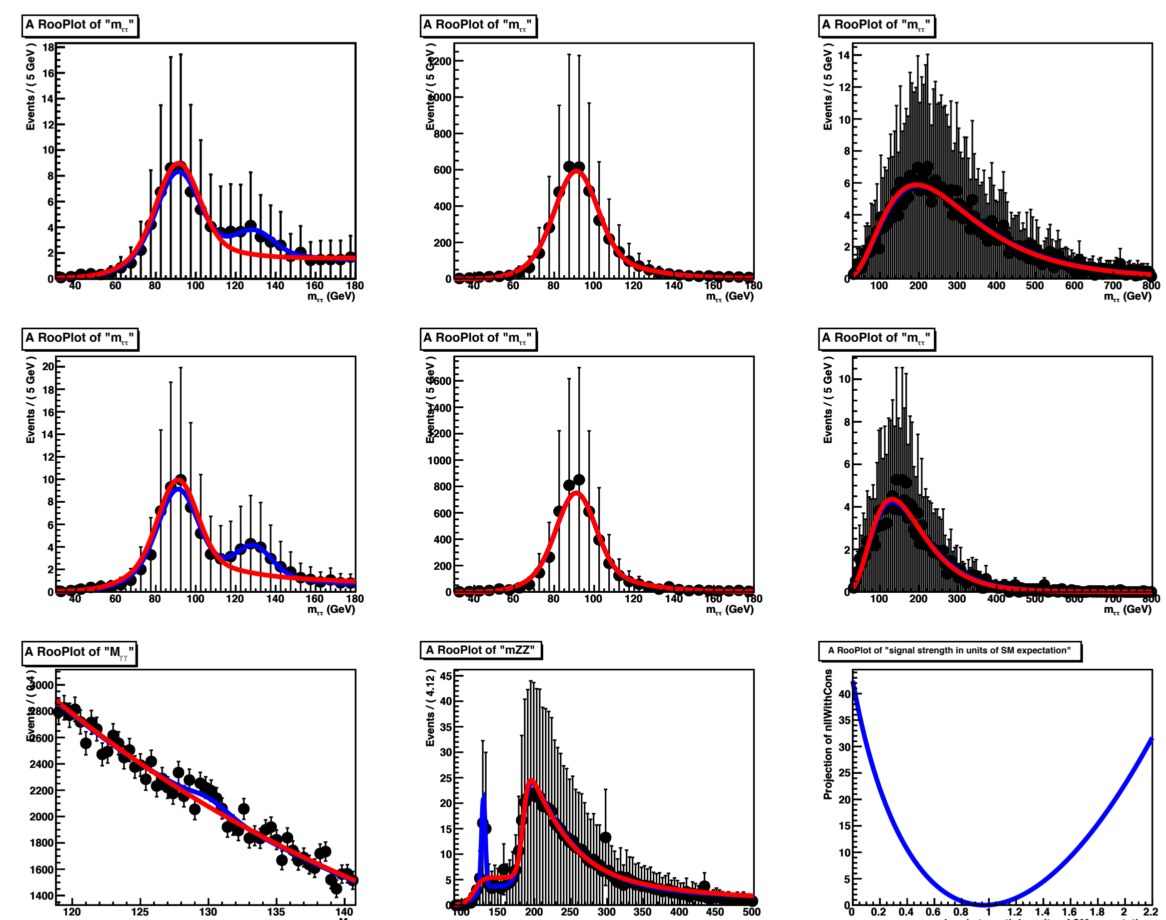
$$\mu = \frac{\sigma BR}{\sigma_{SM} BR_{SM}}$$



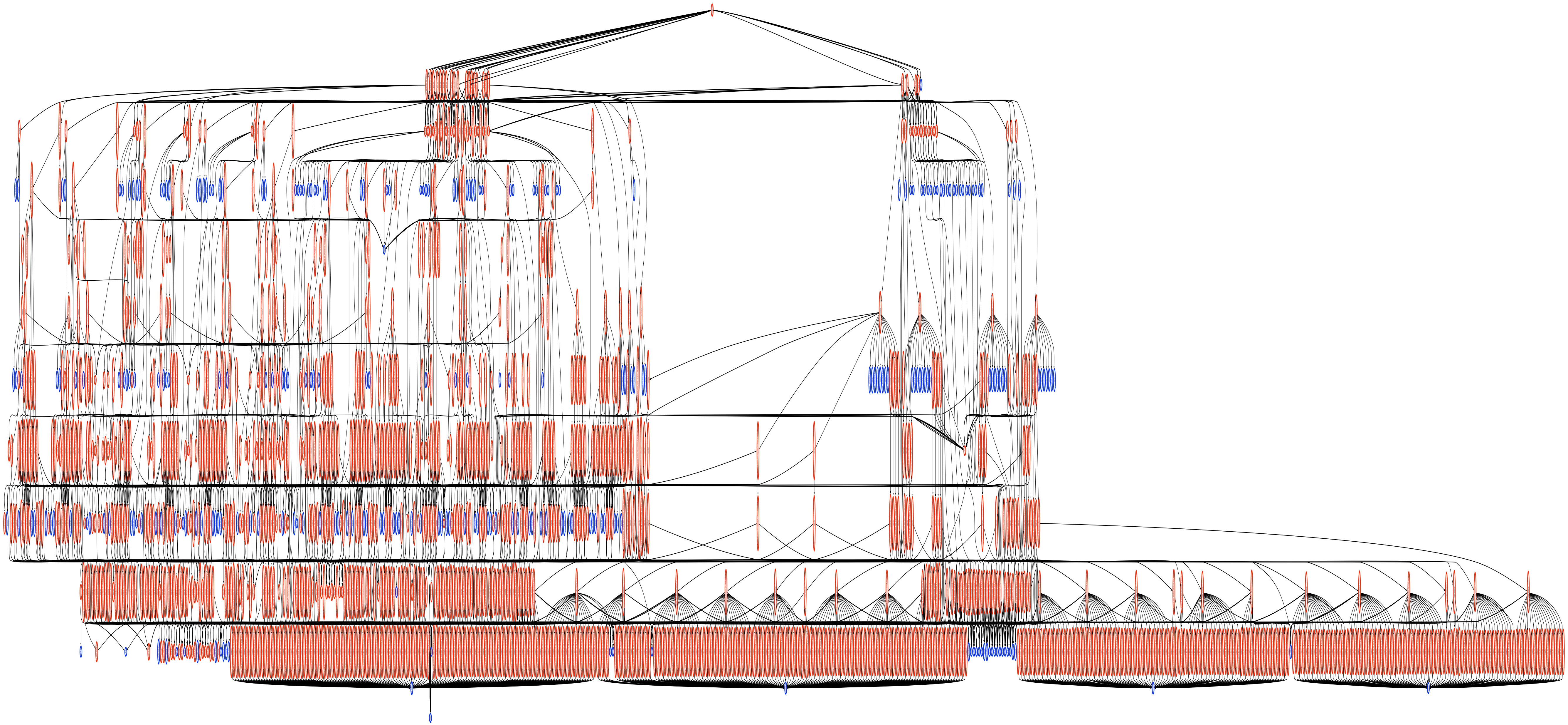
The number of observables and parameters has grown by a factor of ~10 since this exercise.

Visualizing Models I

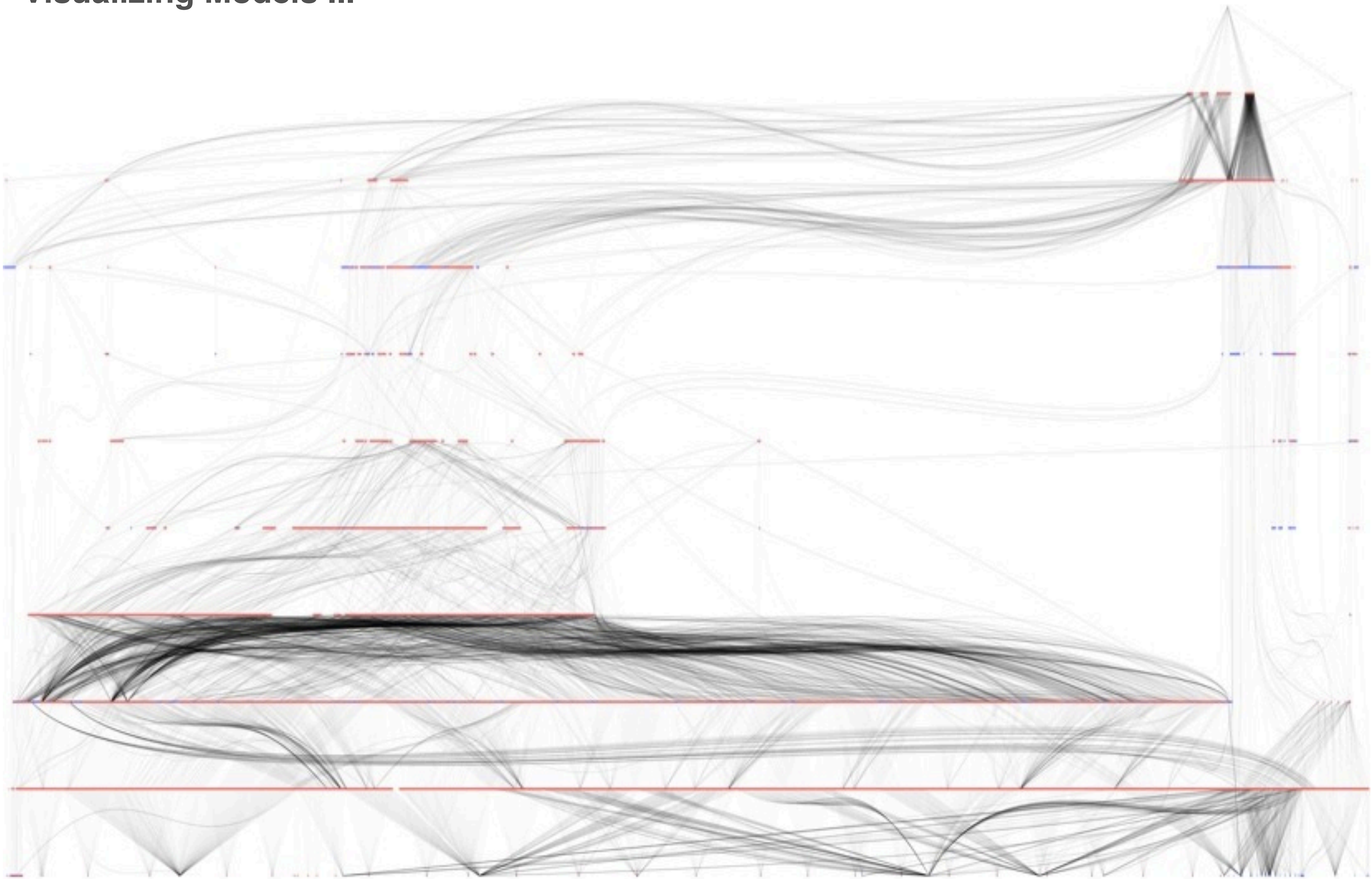




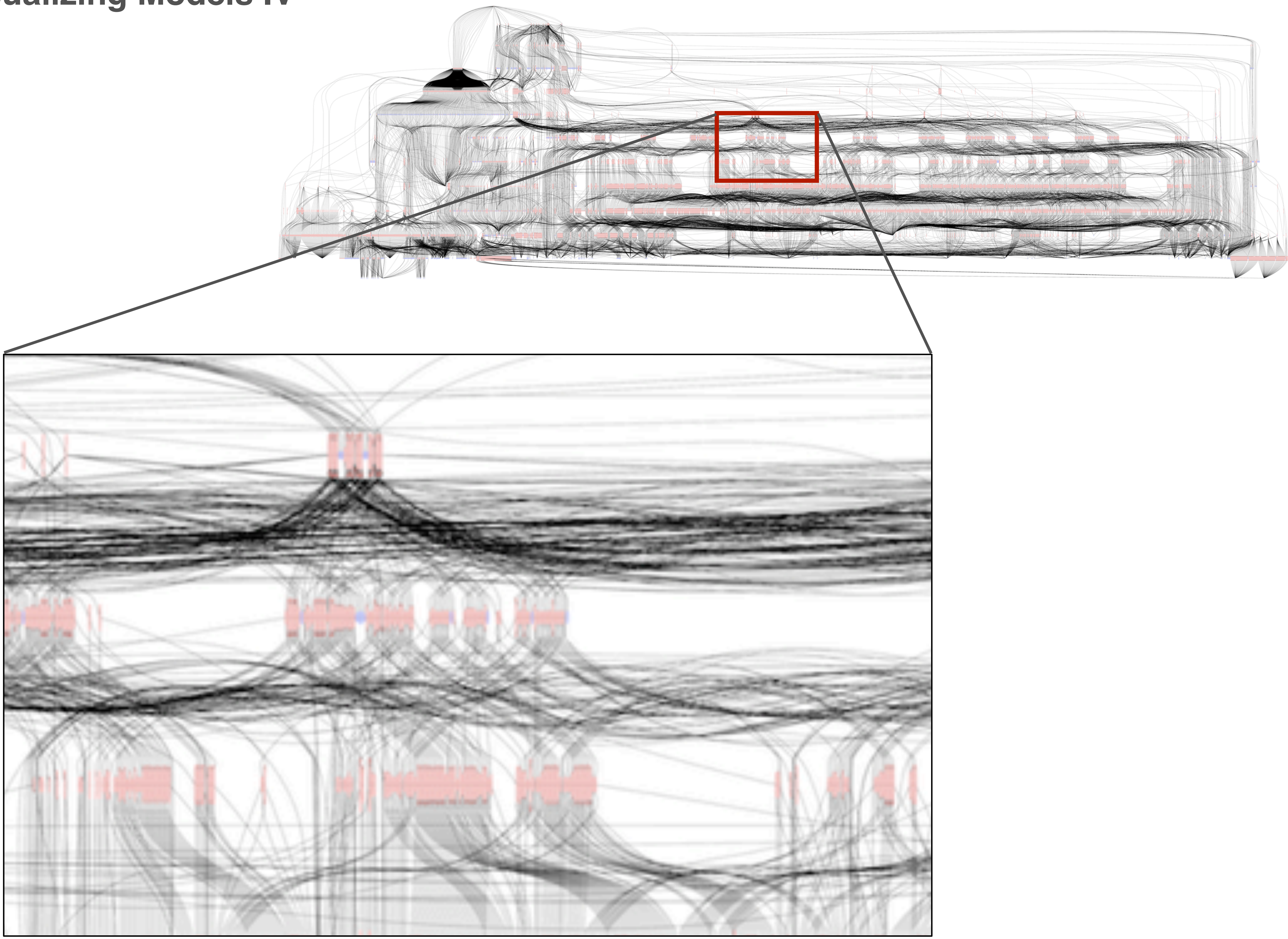
Visualizing Models II



Visualizing Models III



Visualizing Models IV



Summary

RooStats supports many methods for discovery and limits (asymptotics, toys, Bayesian)

- ➡ is powerful (used for real-world complicated analyses and combinations, see plot on top right)
- ➡ proven in individual search channels (e.g. various Higgs and SUSY channels), top physics, exotic analysis and Higgs combinations

Separates model, method and storage: each is improving all the time.

Preparations for Higgs combinations with 7 and 8 TeV and Higgs property measurements are underway.