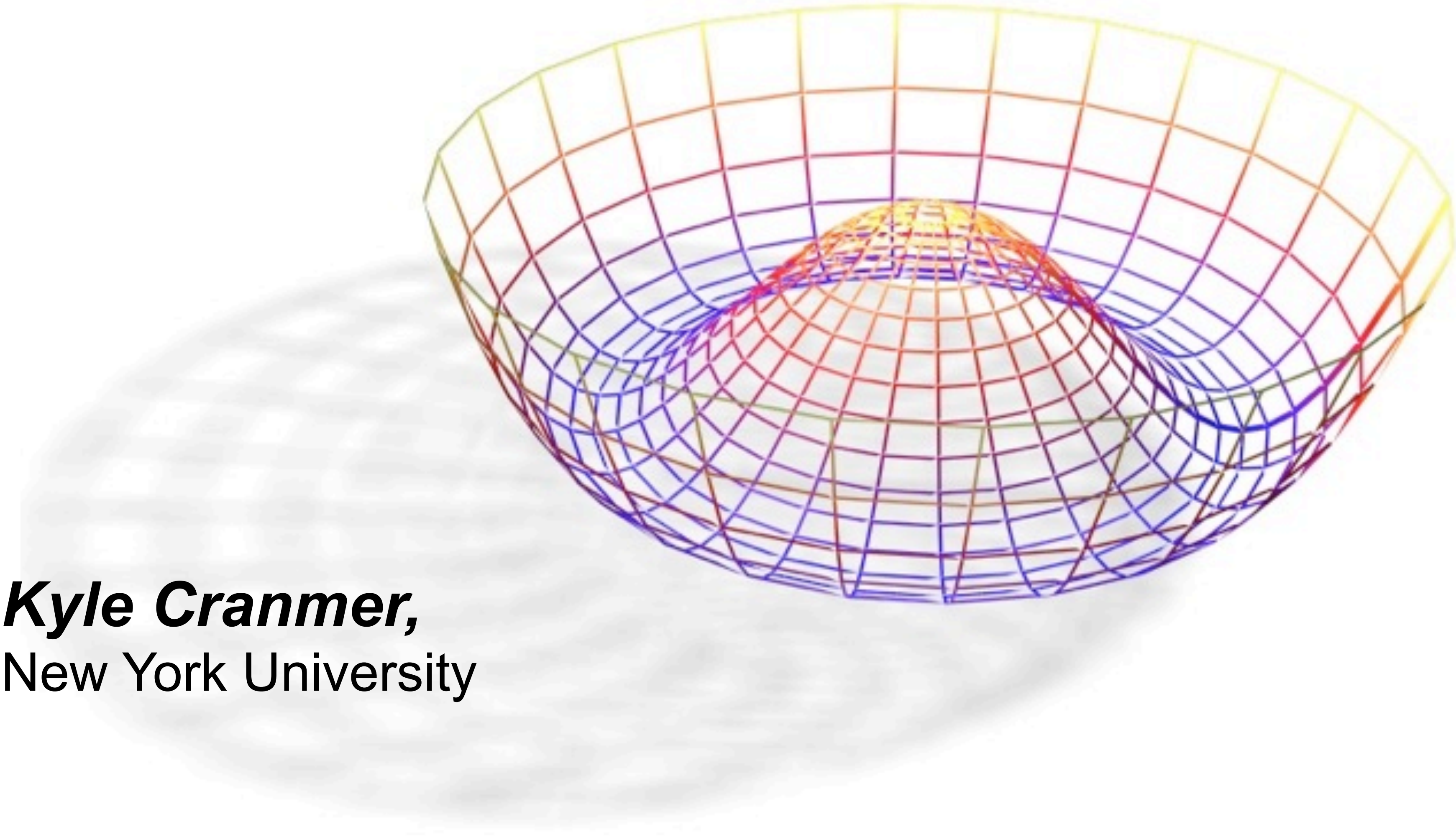
Center for
Cosmology and
Particle Physics

Bringing it all together: statistical interpretation of new physics



Kyle Cranmer,
New York University

Introduction

Michelangelo's Likelihood Mandate (MLM):

A general assessment of the status and needs of the tools for setting limits on (or fitting) parameters of BSM models, using the multitude of data from searches at the LHC

Two related communities and ongoing discussions

- Characterization & Simplified Models
- Fitting Model Parameters

Introduction

Michelangelo's Likelihood Mandate (MLM):

A general assessment of the status and needs of the tools for setting limits on (or fitting) parameters of BSM models, using the multitude of data from searches at the LHC

Two related communities and ongoing discussions

- Characterization & Simplified Models → parametrization
- Fitting Model Parameters → interpretation

Introduction

Michelangelo's Likelihood Mandate (MLM):

A general assessment of the status and needs of the tools for setting limits on (or fitting) parameters of BSM models, using the multitude of data from searches at the LHC

Two related communities and ongoing discussions

- Characterization & Simplified Models → parametrization
- Fitting Model Parameters → interpretation

Potential new tasks

- **Input for the Strategy Group**
  - LPCC and experiments required to produce combined assessment of the 2010-11(-12) findings in Higgs and BSM searches
  - TH community, and other expl communities (e.g. LinCol, SuperB, ...), will use this to assess the implications of LHC data for BSM and future exptl projects

➡ We need to prepare the framework/tools to enable:

  - combination of limits/evidence from ATLAS/CMS(/LHCb)
  - use of the results by the rest of the community (e.g. SUSY-models' fitters)
- This will require coordination with
  - ATLAS-CMS statistics forum
  - Fitters' groups
  - all LHC "search " efforts (Higgs, B decays, exotica of all sorts ....)
  - ...

Introduction

Michelangelo's Likelihood Mandate (MLM):

A general assessment of the status and needs of the tools for setting limits on (or fitting) parameters of BSM models, using the multitude of data from searches at the LHC

Two related communities and ongoing discussions

- Characterization & Simplified Models → parametrization
- Fitting Model Parameters → interpretation

Potential new tasks

- **Input for the Strategy Group**
  - LPCC and experiments required to produce combined assessment of the 2010-11(-12) findings in Higgs and BSM searches
  - TH community, and other expl communities (e.g. LinCol, SuperB, ...), will use this to assess the implications of LHC data for BSM and future exptl projects

➡ We need to prepare the framework/tools to enable:

- combination of limits/evidence from ATLAS/CMS(/LHCb)
- use of the results by the rest of the community (e.g. SUSY-models' fitters)

- This will require coordination with
  - ATLAS-CMS statistics forum
  - Fitters' groups
  - all LHC "search " efforts (Higgs, B decays, exotica of all sorts ....)
  - ...

Goals for this meeting

- Review the progress made by the experiments
- Status report on the SLAC WG
- Collect further input from all fields (TH + exps)
- In the context of simplified models, start outlining the roadmap and the workflow to go from analysis, to publication, to combination of the results of different experiments, to conclude with the exploitation of the published results by a random theorist.

analysis

format of the published result

combination among experiments

use of the results by a theorist, in the context of a new model

Scope and goals

My goal is to make this talk complementary to other contributions

- I respect and have participated in the simplified model approach
  - for the purpose of this talk, like any other parametrized model for the signal
- I also think that most models have more than 2 parameters
  - thus are hard to publish on a piece of paper
- And that most parametrized models have a limited range of validity
  - beyond which we may need detector simulation (eg. something like RECAST)

Goals for this meeting

- Review the progress made by the experiments
- Status report on the SLAC WG
- Collect further input from all fields (TH + exps)
- In the context of simplified models, start outlining the roadmap and the workflow to go from analysis, to publication, to combination of the results of different experiments, to conclude with the exploitation of the published results by a random theorist.

analysis

format of the published result

combination among experiments

use of the results by a theorist, in the context of a new model

Activities in the *fitting Groups

DESY SUSY/BSM Workshop

CENTER FOR COSMOLOGY AND PARTICLE PHYSICS

Analysis Centre: SUSY/BSM Fit Working Group

chaired by Philip Bechtle (DESY), Johannes Haller (University Hamburg), Tilman Plehn (University of Edinburgh & MPI Muenchen), Thomas Schoerner-Sadenius (DESY), Georg Weiglein (IPPP Durham), Peter Wienemann (University of Bonn)

from Monday 26 July 2010 at **14:00** to Wednesday 28 July 2010 at **14:00** (Europe/Berlin)
at DESY

Link to Agenda

14:00 **Welcome** *05'*

14:05 **Overview of Analysis Centre Activities** *10'*
Speakers: Thomas Schoerner-Sadenius (DESY)
Material: Slides

14:15 **Lessons from the past** *20'*
Speakers: Klaus Moenig (DESY)
Material: Slides

14:35 **Introduction to Gfitter** *20'*
Speakers: Max Baak (CERN)
Material: Slides

14:55 **Introduction to Fittino** *20'*
Speakers: Peter Wienemann (University of Bonn)
Material: Slides

15:50 **break** *30'*

16:20 **Introduction to HiggsBounds** *20'*
Speakers: Philip Bechtle (DESY)
Material: Slides

09:00 **Introduction to Mastercode (theory + fitting code)** *20'*
Speakers: Sven Heinemeyer (IFCA (CSIC, Santander))
Material: Slides

09:20 **Introduction to SPheno** *20'*
Speakers: Werner Porod (Universität Würzburg)

11:00 **Theory issues** *20'*
Comparison of codes. Reference SM values for low energy ot theory predictions? How to include them in the fit?
Speakers: Georg Weiglein (DESY)
Material: Slides

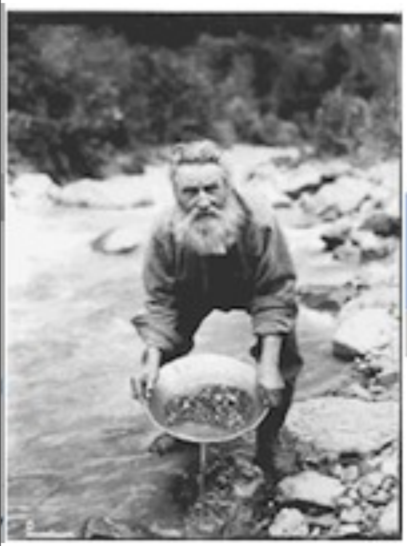
11:20 **Calculator comparisons** *20'*
Speakers: Matthias Hamer (Uni Goettingen)
Material: Slides

12:00 **Lunch break** *2h00'*

14:00 **Database of measurements** *2h00'*
Low energy measurements, LEP, Tevatron, astrophysics
Speakers: Xavier Prudent (IKTP Dresden)
Material: Slides

16:00 **break** *30'*

16:30 **Common interface for inputs** *1h30'*
Make use of RooStats workspaces?
Speakers: Kyle Cranmer (New York University)
Material: Slides



Link to Agenda

OKC PROSPECTS Workshop

15-17 September 2010 AlbaN

Physics Codes - Chair: Klaus Desch (14:00-17:40) Room: FD5 (Svedbergsalen)

14:00-14:50 (00h50')	[256]	DarkSUSY - J. Edsjö
14:50-15:40 (00h50')	[258]	SOFTSUSY: an update on MSSM spectrum calculation - B. Allanach
15:40-16:10 (00h30')		*Coffee Break*
16:10-17:10 (01h00')	[257]	micrOMEGAs, latest developments - G. Belanger
17:10-17:40 (00h30')	[277]	Incorporating Event Rates Into MSSM Parameter Determinations - B. O'LEARY

11:00-11:50 (00h50')	[260]	Publishing likelihoods with RooStats - K. Cranmer

Scanning Packages - Chair: Pat Scott (14:40-18:40) Room: FB42

14:40-15:30 (00h50')	[263]	SuperBayeS - R. Ruiz de Austri
15:30-16:20 (00h50')	[264]	Sfitting - T. Plehn
16:20-16:50 (00h30')		*Coffee Break*
16:50-17:40 (00h50')	[265]	Determination of SUSY parameters with Fittino - K. Desch
17:40-18:40 (01h00')	[279]	Panel Discussion on Scanning Packages - led by K. Desch, T. Plehn and R. Ruiz de Austri

Issues being discussed in these groups

Statistical

- Bayesian vs. Frequentist
- Prior dependence of results (2007)
- Coverage of results (2010)

Technical

- spectrum calculators: RGE accuracy and speed
- consistency of results from different codes
- scanning high dimensional models
  - random scans, minuit, MCMC, MultiNest, genetic algorithms, etc.
- cool logos

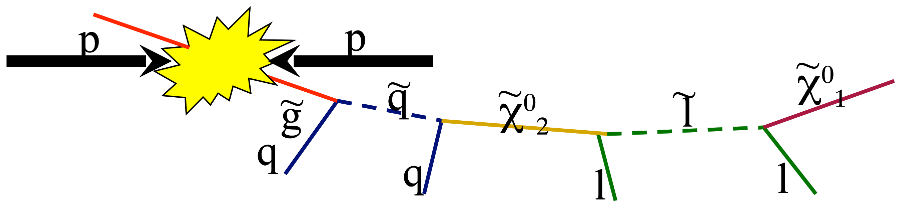
Physics

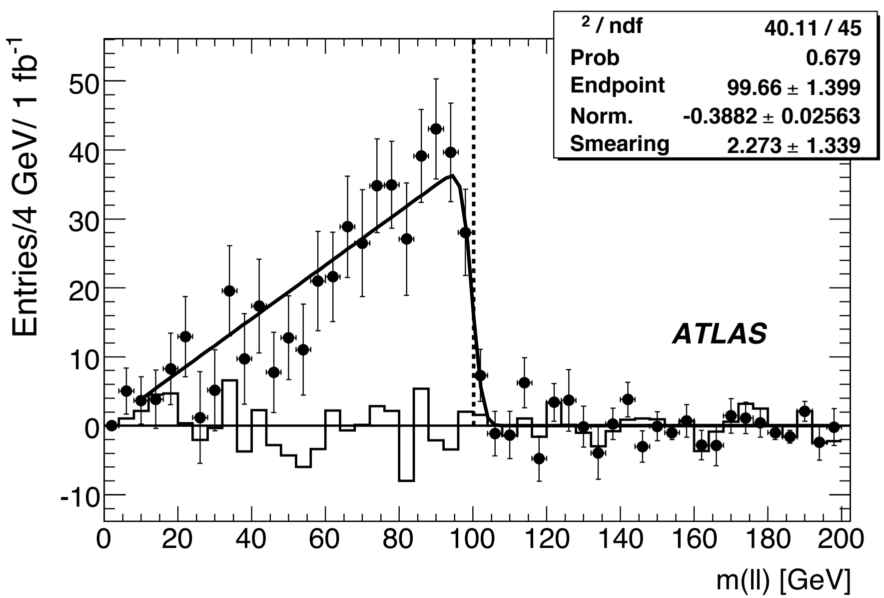
- global fits to existing precision data
  - "weather forecasts" for LHC and direct dark matter searches
- Expected results and preparation for global fits with LHC data

Recently, I've been trying to draw attention to the inputs to these programs

- eg. the likelihood functions

Inference on which parameters?





$$m_{ll}^{\mathrm{edge}} = m_{\tilde{\chi}^0_2}\sqrt{1-\left(\frac{m_{\tilde{l}}}{m_{\tilde{\chi}^0_2}}\right)^2}\sqrt{1-\left(\frac{m_{\tilde{\chi}^0_1}}{m_{\tilde{l}}}\right)^2}$$

$$(m_{\tilde{l}}, m_{\tilde{\chi}^0_2}, m_{\tilde{\chi}^0_1}, \ldots)$$

$$(m_0, m_{1/2}, \tan\beta, A_0, \mathrm{sign}(\mu))$$

Most of the effort of the fitting groups has been on inferring parameters of more fundamental theories

- often needed more fundamental theory to relate observations in different experiments

However, most of the technical and statistical tools can be applied to inference on the physical parameters (sparticle masses, cross-sections, BRs)

- and for similar experiments it is clear how to relate and combine measurements

Supersymmetric Parameter Estimation

Several groups have been using MCMC to infer supersymmetric parameters based on WMAP and various particle physics results

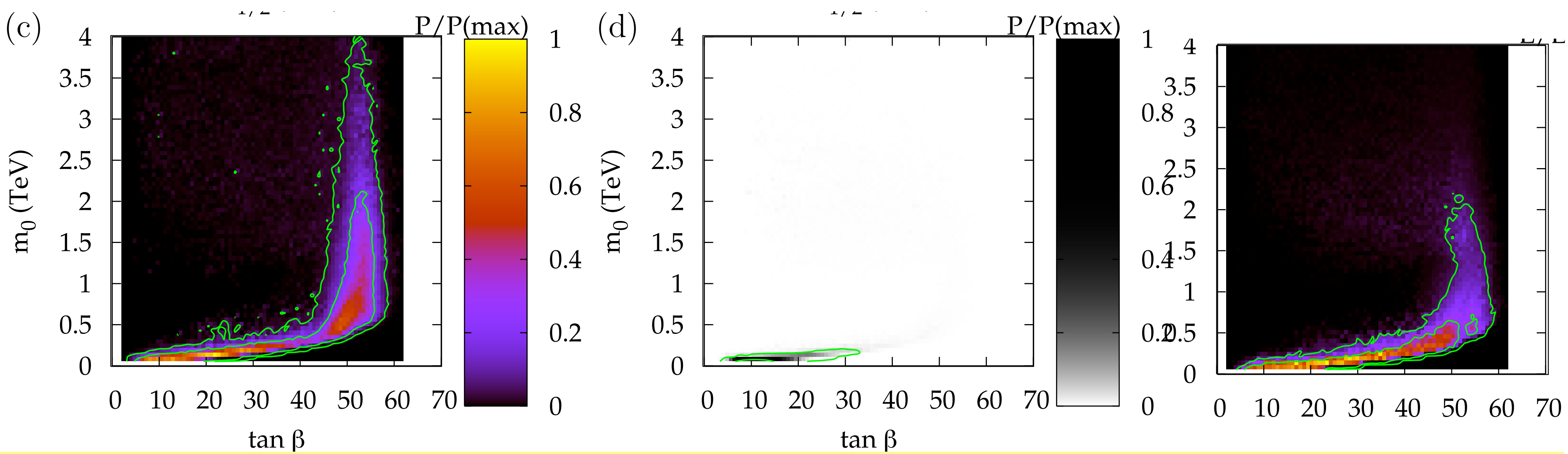
Initially these groups were all using flat priors

$$p(m_0, M_{1/2}, A_0, \tan\beta, s|\text{data}) = p(\text{data}|m_0, M_{1/2}, A_0, \tan\beta, s)\frac{p(m_0, M_{1/2}, A_0, \tan\beta, s)}{p(\text{data})}$$

I joined one group that used a hierarchical Bayesian technique relating the prior of several parameters to one energy scale with physical significance

$$p(m_0, M_{1/2}, A_0, \mu, B) = \int_0^\infty dM_S\; p(m_0, M_{1/2}, A_0, \mu, B|M_S)\; p(M_S)$$

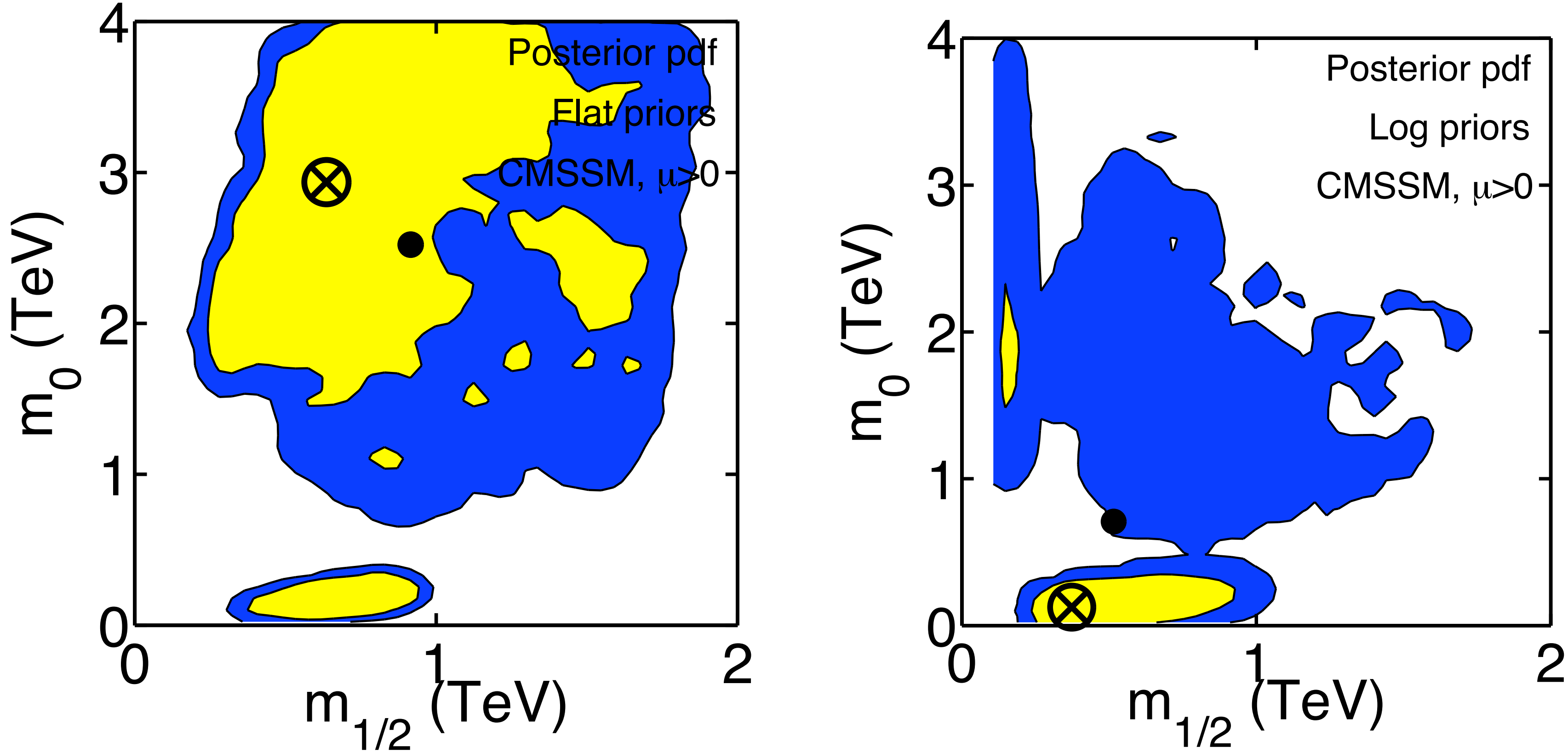
We also repeated the analysis using the profile likelihood ratio (no priors)



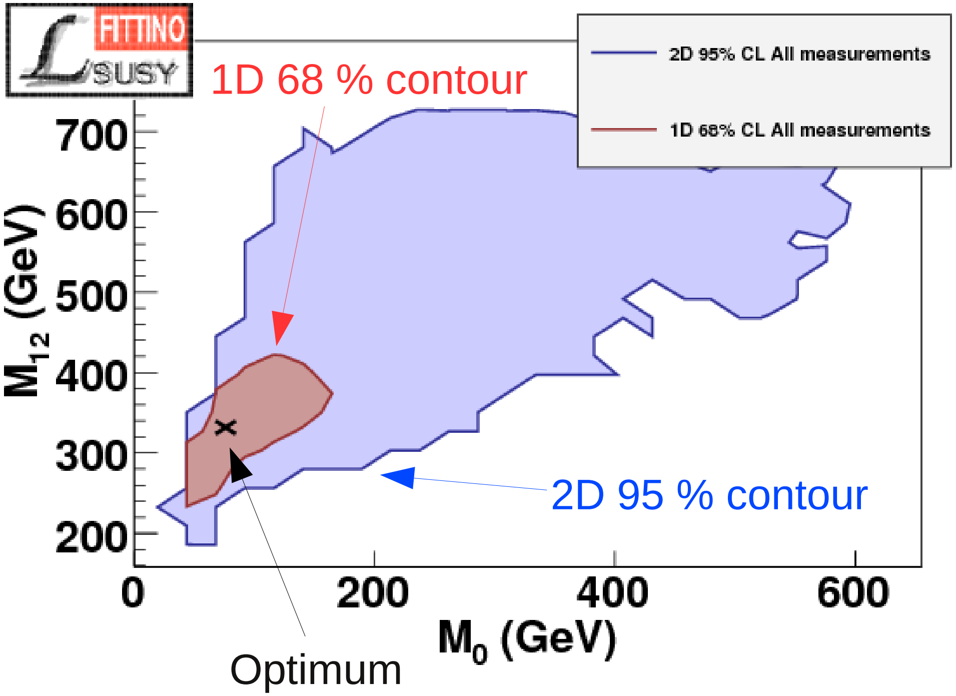
[hep-ph/0507283, hep-ph/0601089, arXiv:0705.0487] , B. C. Allanach, K. Cranmer, C. G. Lester, and A. M. Weber, J. High Energy Phys. 08 (2007) 023.

Prior dependence

Center for Cosmology and Particle Physics



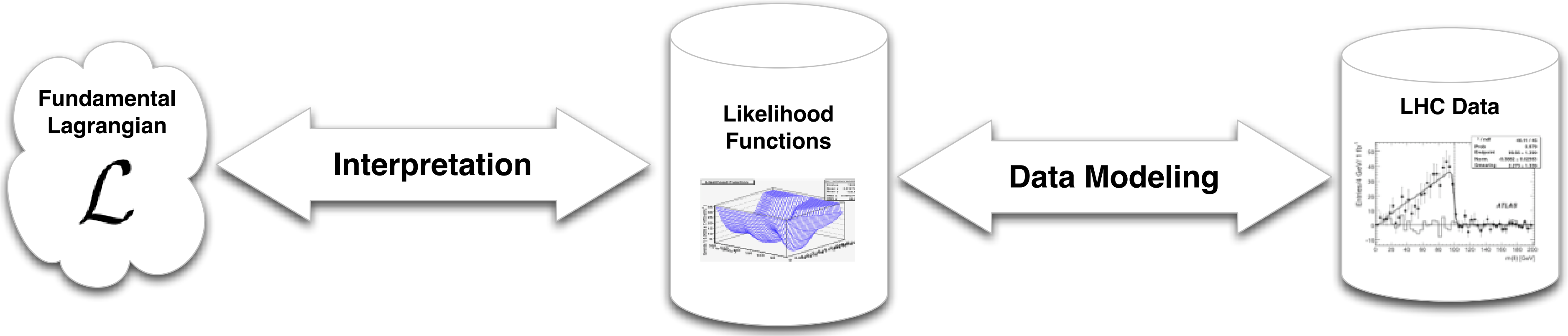
Observable	Experimental Value	Uncertainty stat	Uncertainty syst	Exp. Reference
$\mathcal{B}(B \to s\gamma)/\mathcal{B}(B \to s\gamma)_{\rm SM}$	1.117	0.076	0.096	[47]
$\mathcal{B}(B_s \to \mu\mu)$	$< 4.7\times10^{-8}$			[47]
$\mathcal{B}(B_d \to \ell\ell)$	$< 2.3\times10^{-8}$			[47]
$\mathcal{B}(B \to \tau\nu)/\mathcal{B}(B \to \tau\nu)_{\rm SM}$	1.15	0.40		[48]
$\mathcal{B}(B_s \to X_s\ell\ell)/\mathcal{B}(B_s \to X_s\ell\ell)_{\rm SM}$	0.99	0.32		[47]
$\Delta m_{B_s}/\Delta m_{B_s}^{\rm SM}$	1.11	0.01	0.32	[49]
$\frac{\Delta m_{B_s}/\Delta m_{B_s}^{\rm SM}}{\Delta m_{B_d}/\Delta m_{B_d}^{\rm SM}}$	1.09	0.01	0.16	[47,49]
$\Delta\epsilon_K/\Delta\epsilon_K^{\rm SM}$	0.92	0.14		[49]
$\mathcal{B}(K \to \mu\nu)/\mathcal{B}(K \to \mu\nu)_{\rm SM}$	1.008	0.014		[50]
$\mathcal{B}(K \to \pi\nu\bar{\nu})/\mathcal{B}(K \to \pi\nu\bar{\nu})_{\rm SM}$	< 4.5			[51]
$a_\mu^{\rm exp} - a_\mu^{\rm SM}$	30.2×10^{-10}	8.8×10^{-10}	2.0×10^{-10}	[52,53]
$\sin^2\theta_{\rm eff}$	0.2324	0.0012		[46]
Γ_Z	2.4952 GeV	0.0023 GeV	0.001 GeV	[46]
R_l	20.767	0.025		[46]
R_b	0.21629	0.00066		[46]
R_c	0.1721	0.003		[46]
$A_{\rm fb}(b)$	0.0992	0.0016		[46]
$A_{\rm fb}(c)$	0.0707	0.0035		[46]
A_b	0.923	0.020		[46]
A_c	0.670	0.027		[46]
A_l	0.1513	0.0021		[46]
A_τ	0.1465	0.0032		[46]
$A_{\rm fb}(l)$	0.01714	0.00095		[46]
$\sigma_{\rm had}$	41.540 nb	0.037 nb		[46]
m_h	> 114.4 GeV		3.0 GeV	[54,55,56]
$\Omega_{\rm CDM}h^2$	0.1099	0.0062	0.012	[57]
$1/\alpha_{em}$	127.925	0.016		[58]
G_F	$1.16637\times10^{-5}\,{\rm GeV}^{-2}$	$0.00001\times10^{-5}\,{\rm GeV}^{-2}$		[58]
α_s	0.1176	0.0020		[58]
m_Z	91.1875 GeV	0.0021 GeV		[46]
m_W	80.399 GeV	0.025 GeV	0.010 GeV	[58]
m_b	4.20 GeV	0.17 GeV		[58]
m_t	172.4 GeV	1.2 GeV		[59]
m_τ	1.77684 GeV	0.00017 GeV		[58]
m_c	1.27 GeV	0.11 GeV		[46]



Ideal scenario for fitting packages

Center for
Cosmology and
Particle Physics

The ideal interface between the data and the inference to the effective or fundamental Lagrangian parameters is through a likelihood function that accurately incorporates all the **experimental systematics** and is **parametrized** in a way relevant to the theory in question



Is this feasible?

- It is the basic model on which Zfitter, GFitter, SFitter, Fittino, MasterCode, Kismet, SuperBayes, etc. are based
- unfortunately, likelihood functions are usually simplistic and based on a few 1-d measurements

A few points

Objective part of Bayesian inference is encoded in Likelihood

$$P(\text{theory}|\text{data}) = \frac{L(\text{data}|\text{theory})\pi(\text{theory})}{P(\text{data})}$$

- improvements in Likelihood is not Bayesian vs. Frequentist

Prior may be based on data $\pi(\text{theory}) \propto L'(\text{data'}|\text{theory})\eta(\text{theory})$

- but it also depends on the initial prior $\eta(\text{theory})$

In the same way that the "Bayesian calculus" allows for propagation of belief, the measurements can be combined with the likelihood function

$$L_{\text{tot}}(\text{data'}|\text{theory}) = L(\text{data'}|\text{theory})L'(\text{data'}|\text{theory})$$

Current scenario

Center for
Cosmology and
Particle Physics

Taken from the GFitter paper

[23]This procedure only uses the M_H value under consideration, where Higgs-mass hypothesis and measurement are compared. It thus neglects that in the SM a given signal hypothesis entails background hypotheses for all M_H values other than the one considered. An analysis accounting for this should provide a statistical comparison of a given hypothesis with all available measurements. This however would require to know the correlations among all the measurement points (or better: the full experimental likelihood as a function of the Higgs-mass hypothesis), which are not provided by the experiments to date. The difference to the hypothesis-only test employed here is expected to be small at present, but may become important once an experimental Higgs signal appears, which however has insufficient significance yet

Origins I: The First "Statistics in HEP" conference

WORKSHOP ON CONFIDENCE LIMITS

CERN, Geneva, Switzerland
17–18 January 2000

CERN 2000–005

Massimo Corradi

Does everybody agree on this statement, to publish likelihoods?

Louis Lyons

Any disagreement ? Carried unanimously. That's actually quite an achievement for this Workshop.

...[Fred James wants to be able to calculate coverage, Don Groom wants to able to calculate goodness of fit]...

Cousins

I thought the point of unanimity was that publishing the likelihood function was a *necessary* condition, not a sufficient condition.

But a practical problem remained: *How* to communicate multi-D likelihood?

http://indico.cern.ch/conferenceDisplay.py?confId=100458

RooStats: Project info

Started in 2005, when René Brun asked me to help organize statistical tools in ROOT

- Main goals are to provide a common framework for various statistical techniques (Frequentist, Bayesian, Likelihood based,...)

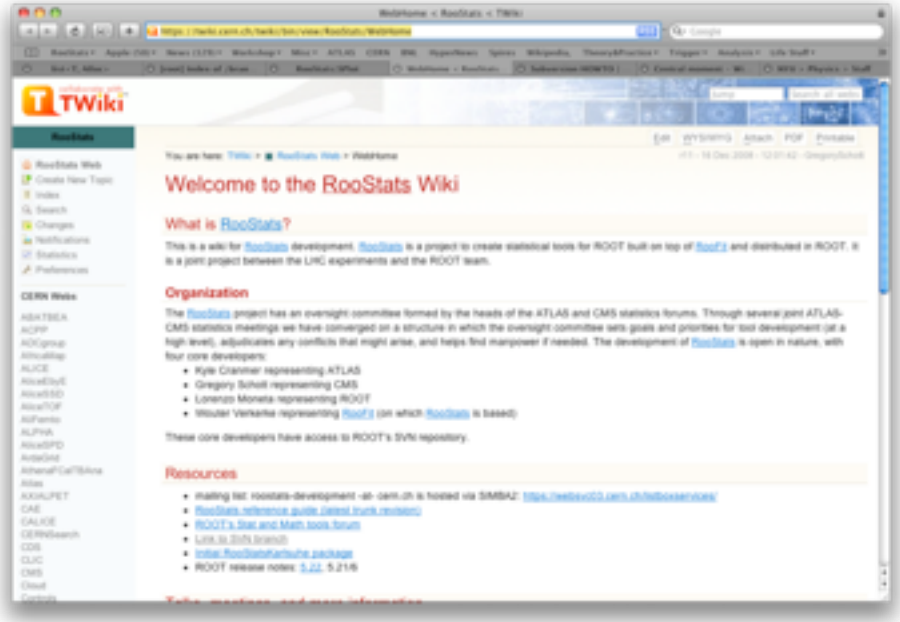
We want tools to work with probability models of arbitrary complexity (which implies interfaces, etc.)

- Decided to base tools on RooFit's data modeling language and core interfaces

Initially an ATLAS/CMS project, but other experiments are interested (LHCb, Fermi, ...)

- core developers
  - K. Cranmer (ATLAS), Lorenzo Moneta (ROOT), Gregory Schott (CMS), Wouter Verkerke (RooFit)
- open project, you are welcome to contribute
  - ~10 others contributing now, growing fast

Included since ROOT v5.22 (we are now on 5.27)



https://twiki.cern.ch/twiki/bin/view/RooStats/WebHome

RooStats has been a topic of conversation in every combined ATLAS/CMS statistics forum meeting

- In July, we showed the first toy ATLAS/CMS Higgs combination using the tools of RooFit/RooStats.
  - see agenda:

http://indico.cern.ch/conferenceDisplay.py?confId=100458

Major Goals and Status

Center for Cosmology and Particle Physics

Goal: Standardize interface for major statistical procedures so that they can work on an arbitrary RooFit model & dataset and handle many parameters of interest and nuisance parameters (systematics).

- **Status**: Done
  - **ConfIntervalCalculator** & **HypoTestCalculator** interface for tools
  - they return **ConfidenceInterval** and **HypoTestResult**

Goal: Implement most accepted techniques from Frequentist, Bayesian, and Likelihood-based approaches

- **Status**: Done / Ongoing
  - **ProfileLikelihoodCalculator**: (Likelihood) the method of MINUIT/MINOS
  - **FeldmanCousins**: (Frequentist) a generalization of F-C that can incorporate systematics
  - **MCMCCalculator**: (Bayesian) uses Metropolis-Hastings algorithm (native or BAT)
  - **HybridCalculator**: (Bayesian/Frequentist Hybrid) used at LEP and Tevatron

Goal: Provide utilities to perform combined measurements

- **Status:** Partially done / Ongoing
  - **RooWorkspace** allows one to save arbitrary RooFit model (even with custom code) into a .root file. PDFs and DataSets have been extended to facilitate combinations.
  - Same technology can aid in digital publishing

today's focus

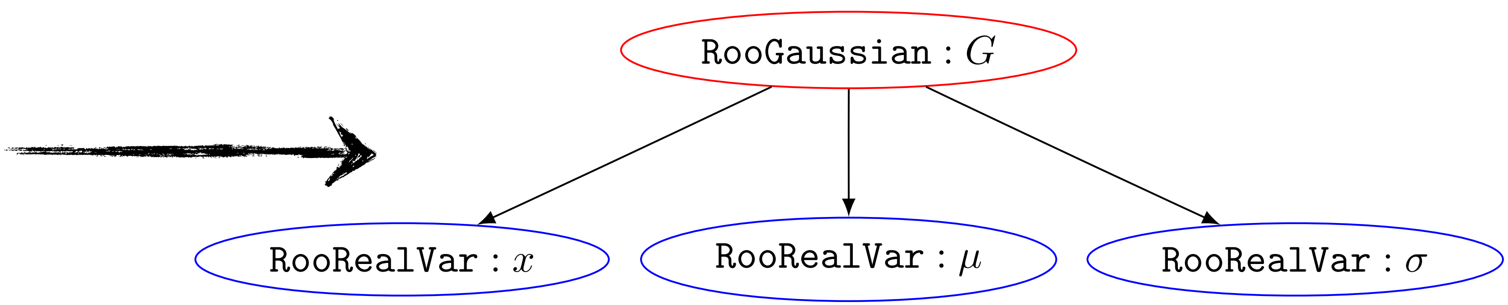
What goes in a Workspace

The workspace stores the full probability model and any data necessary to evaluate the likelihood function

- it is the code necessary to evaluate the likelihood function at an arbitrary point in the parameter space. It is not a big table of likelihood values!
- we are using the same ROOT technology that the LHC experiments are using to save their data
  - well supported, and supports "schema evolution" / backwards compatibility
- the probability model also allows you to generate toy data for any given parameter point
  - necessary for frequentist methods, goodness of fit, coverage
- PDFs and functions can be extended by the user (source stored in workspace)

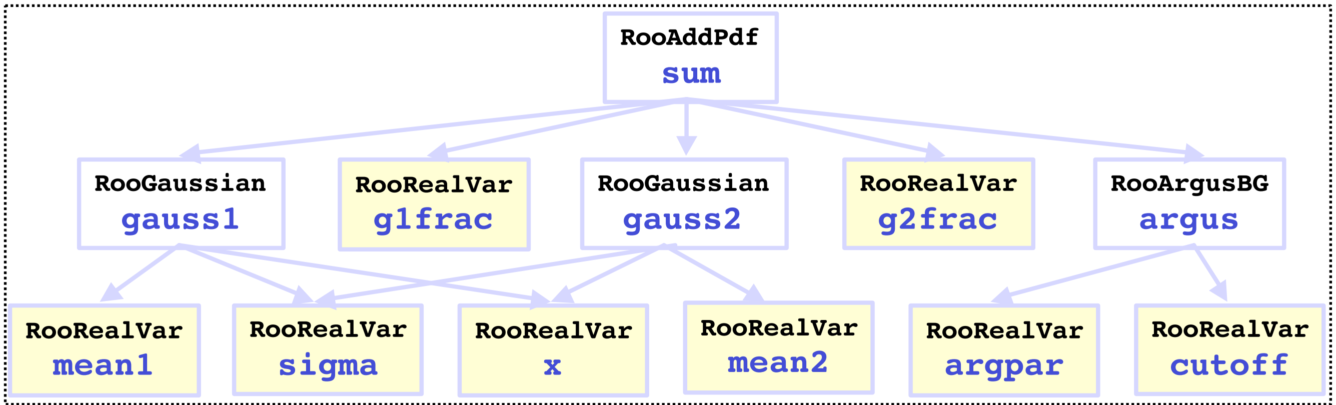
I will show some visualization of real-life LHC probability models. Let's start with a simple example:

$$G(x|\mu,\sigma)$$

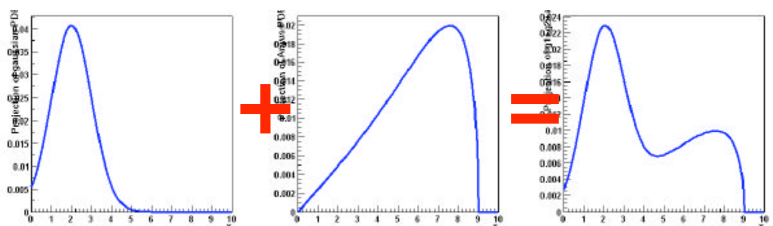


RooFit: A data modeling toolkit

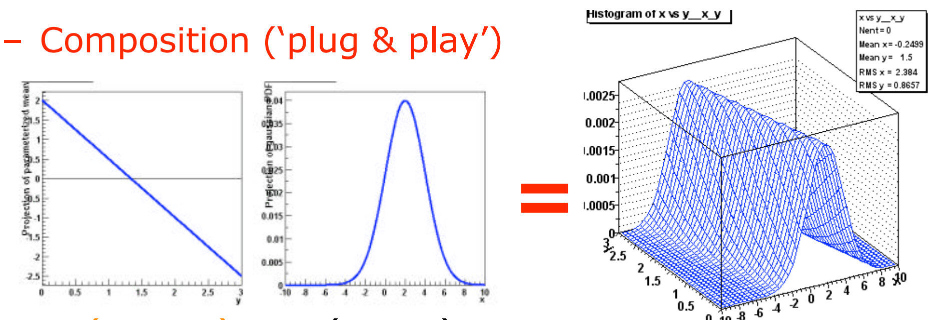
A major tool at BaBar. Fit complicated models with >100 parameters!



– Addition



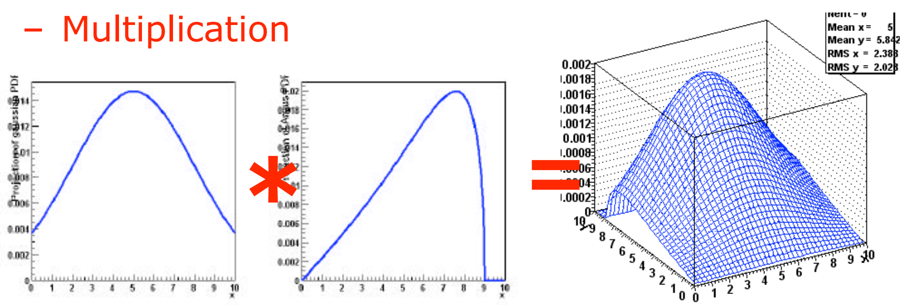
– Composition ('plug & play')



$m(y;a_0,a_1)$ $g(x;m,s)$ $g(x,y;a0,a1,s)$

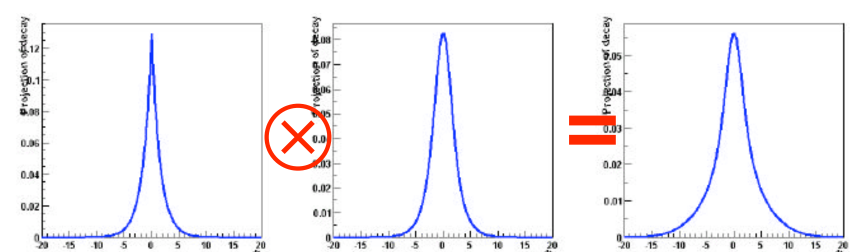
Possible in *any* PDF
No explicit support in PDF code needed

– Multiplication



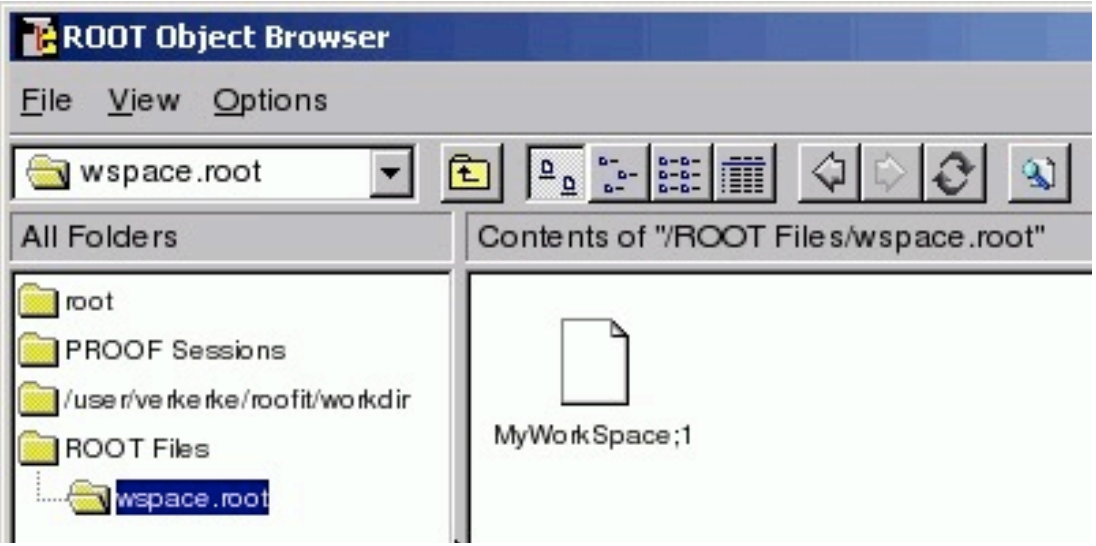
Wouter Verkerke,

– Convolution



Wouter Verkerke, UCSB

The RooFit/RooStats workspace

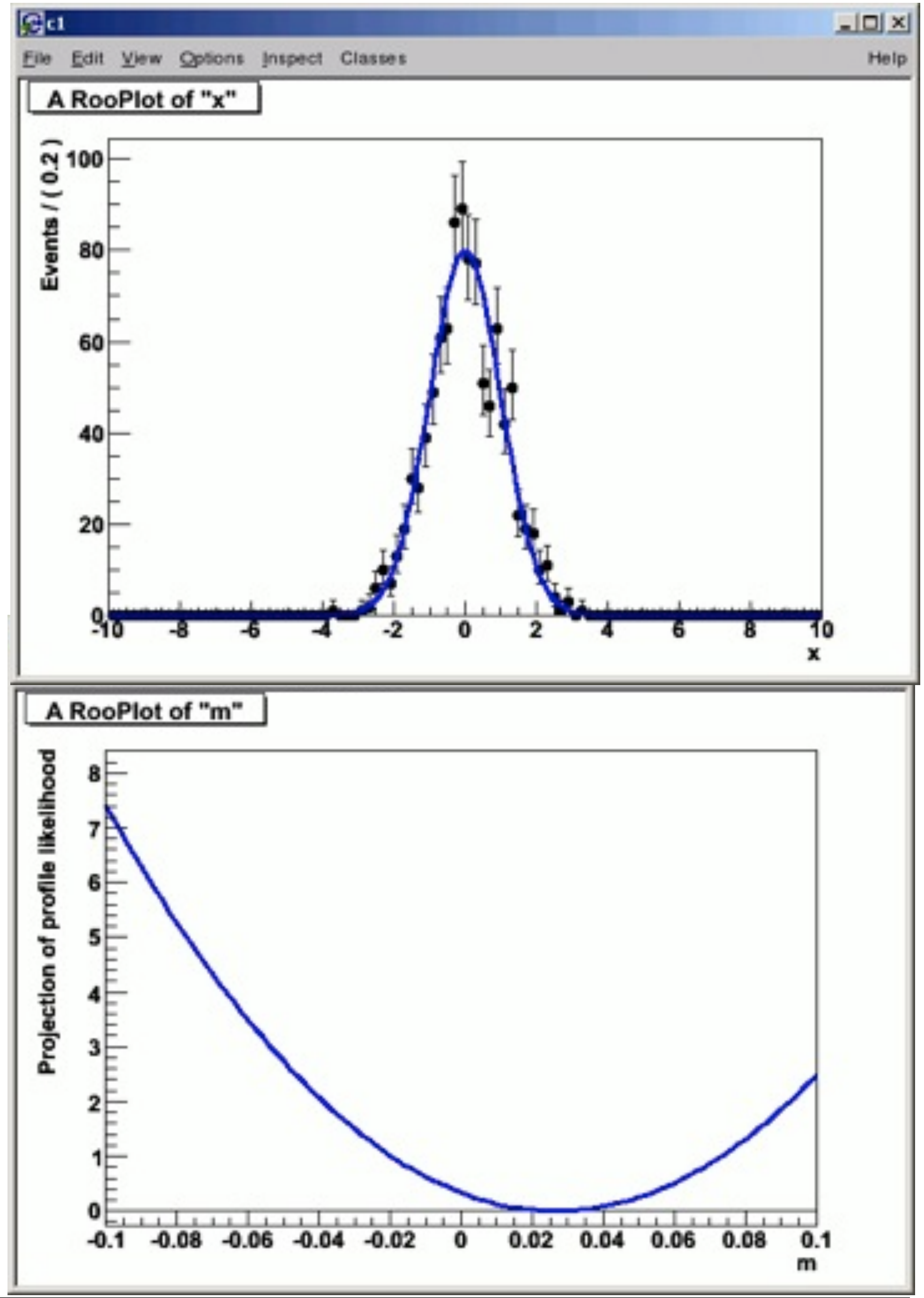


RooStat's Workspace can save in a file the full likelihood model and the minimal data necessary to reproduce likelihood function.

The technology is generic, we decide how to parametrize the model.

Being used by ATLAS/CMS for very complicated models

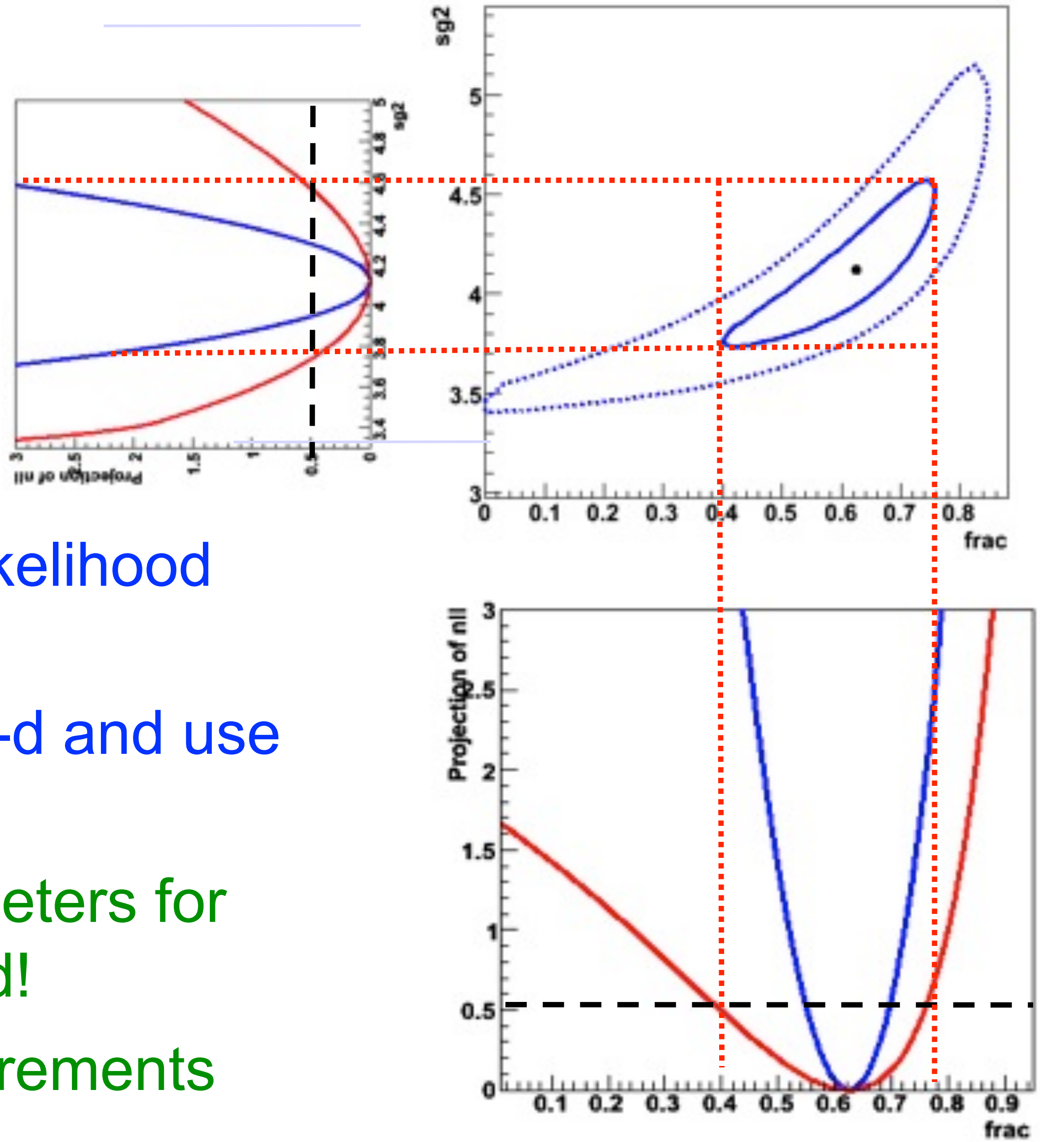
Need this for combinations, exciting potential for publishing results.



Extracting Contours from these results

The workspace can represent arbitrary models with many parameters of interest and many nuisance parameters

This contour is NOT an ellipse!

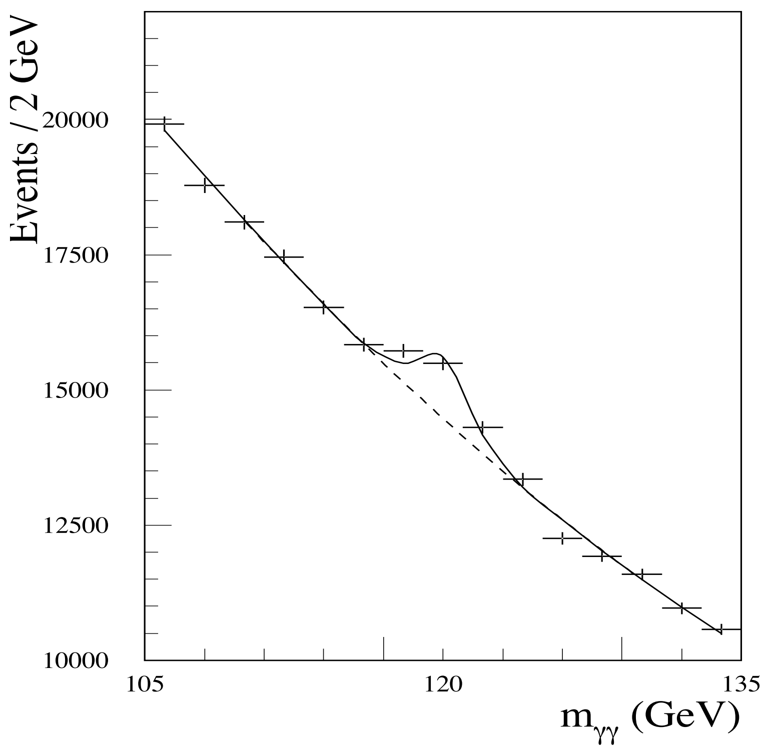


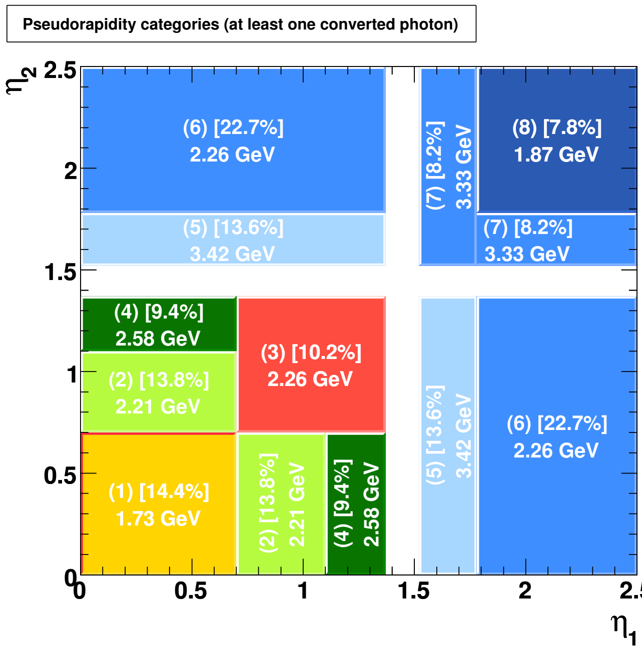
- One can plot 2-d contours, 1-d likelihood functions.
- One can evaluate likelihood in N-d and use to evaluate a theoretical model
- If the model has nuisance parameters for systematics, they will be included!
- Easy to combine multiple measurements

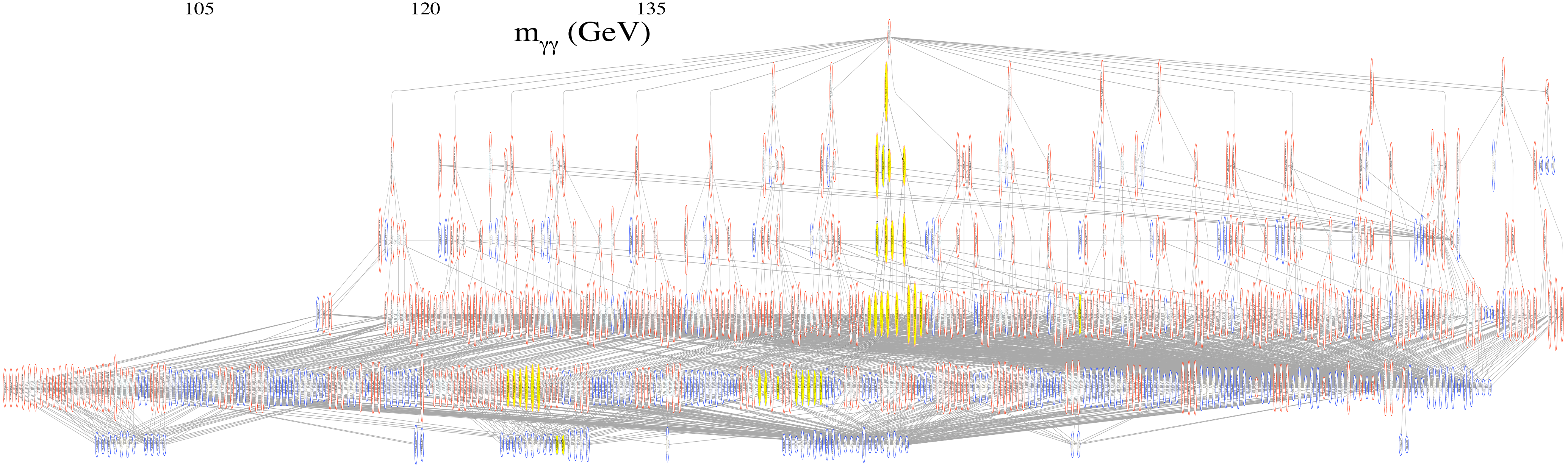
Taken from Wouter Verkerke, NIKHEF

Examples of Real-Life LHC Models

ATLAS H->γγ


Events / 2 GeV
20000
17500
15000
12500
10000
105
120
135
$m_{\gamma\gamma}$ (GeV)


Pseudorapidity categories (at least one converted photon)
η_2
η_1
(1) [14.4%] 1.73 GeV
(2) [13.8%] 2.21 GeV
(3) [10.2%] 2.26 GeV
(4) [9.4%] 2.58 GeV
(5) [13.6%] 3.42 GeV
(6) [22.7%] 2.26 GeV
(7) [8.2%] 3.33 GeV
(8) [7.8%] 1.87 GeV



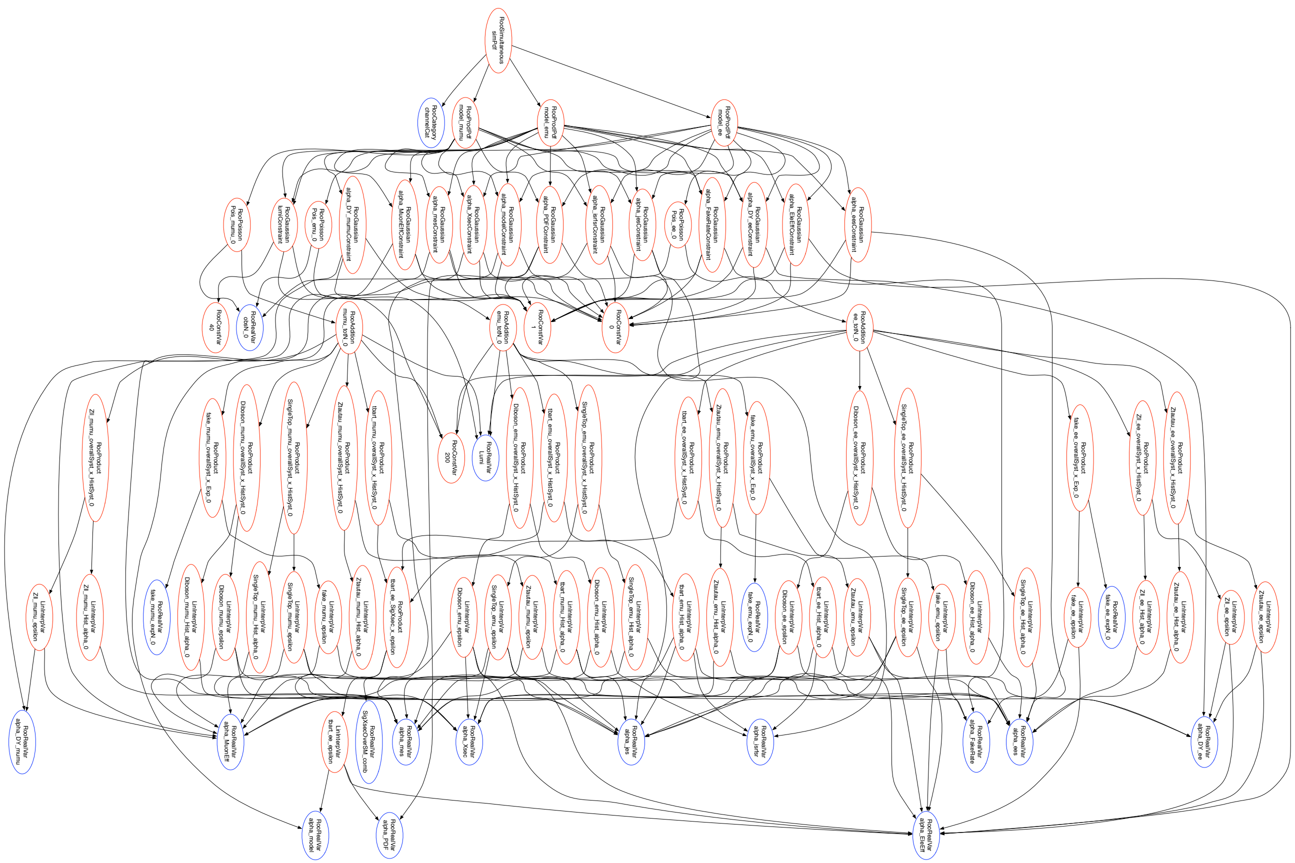
3-channel top combination

The graph below represents this PDF

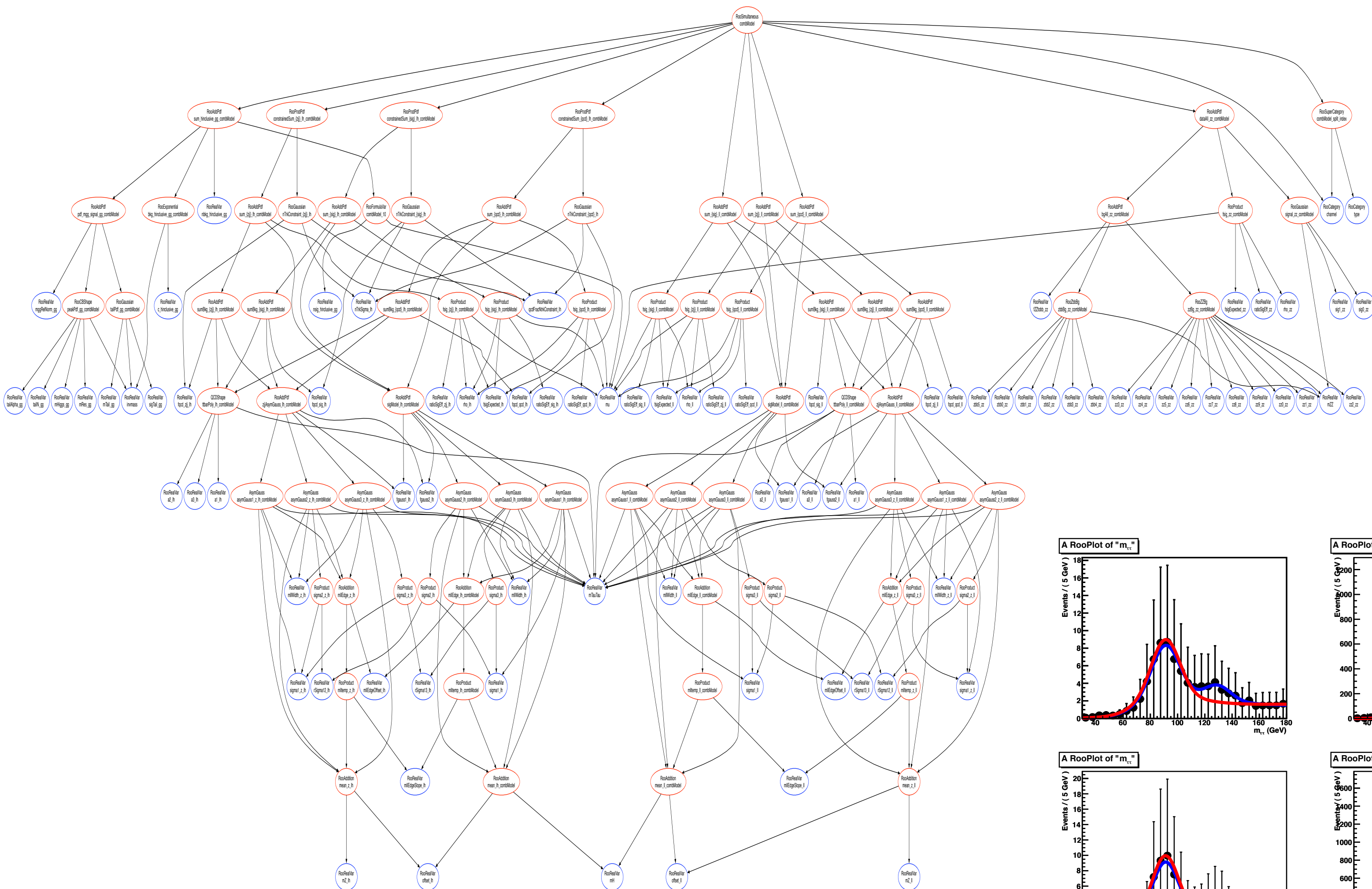
$$L(\sigma_{sig}, \mathscr{L}, \alpha_j) = \prod_{l \in \{ee, \mu\mu, e\mu\}} \left\{ \prod_{i \in bins} \left[Pois(N_i^{obs} | N_{i,tot}^{exp}) \, Gaus(\tilde{\mathscr{L}} | \mathscr{L}, \sigma_{\mathscr{L}}) \prod_{j \in syst} Gaus(0 | \alpha_j, 1) \right] \right\}$$

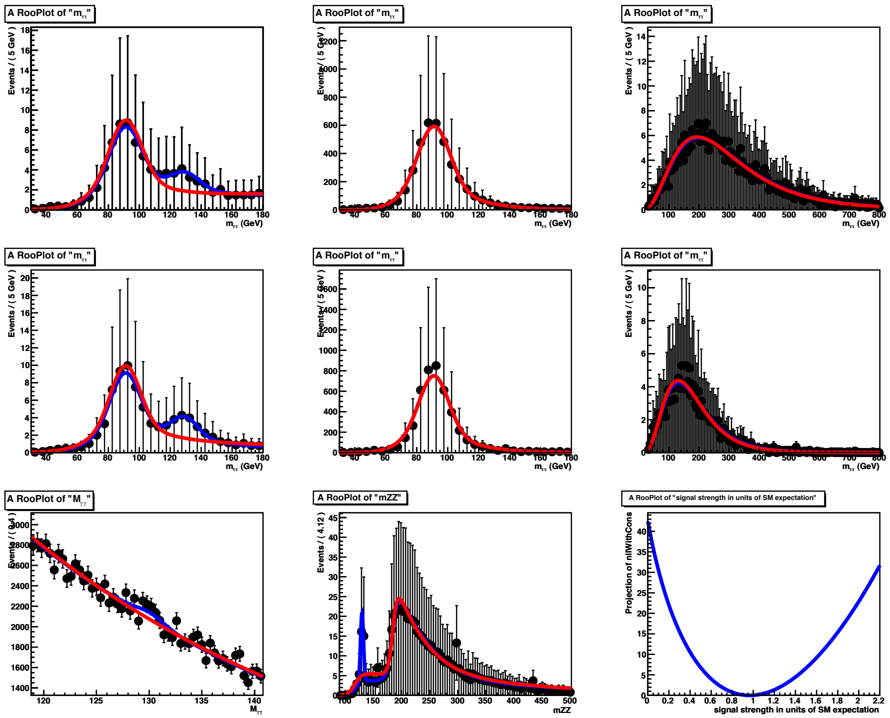
- where there are several relations between the expected means in the different channels

3 observations from data
13 control samples
1 parameter of interest
13 nuisance parameters



4-channel ATLAS Higgs combination



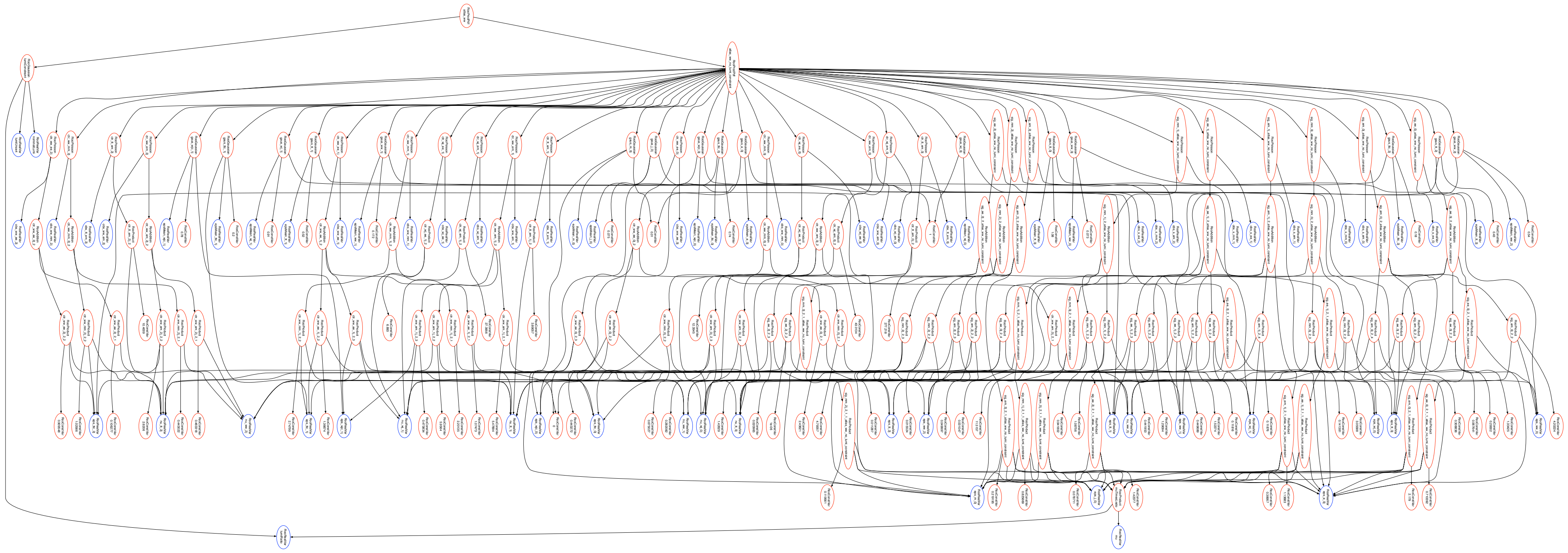


9 observations of continuous variables
1 parameter of interest
27 nuisance parameters

9 Channel ATLAS H->WW combination

CENTER FOR
COSMOLOGY AND
PARTICLE PHYSICS

top level model



parameter of interest

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

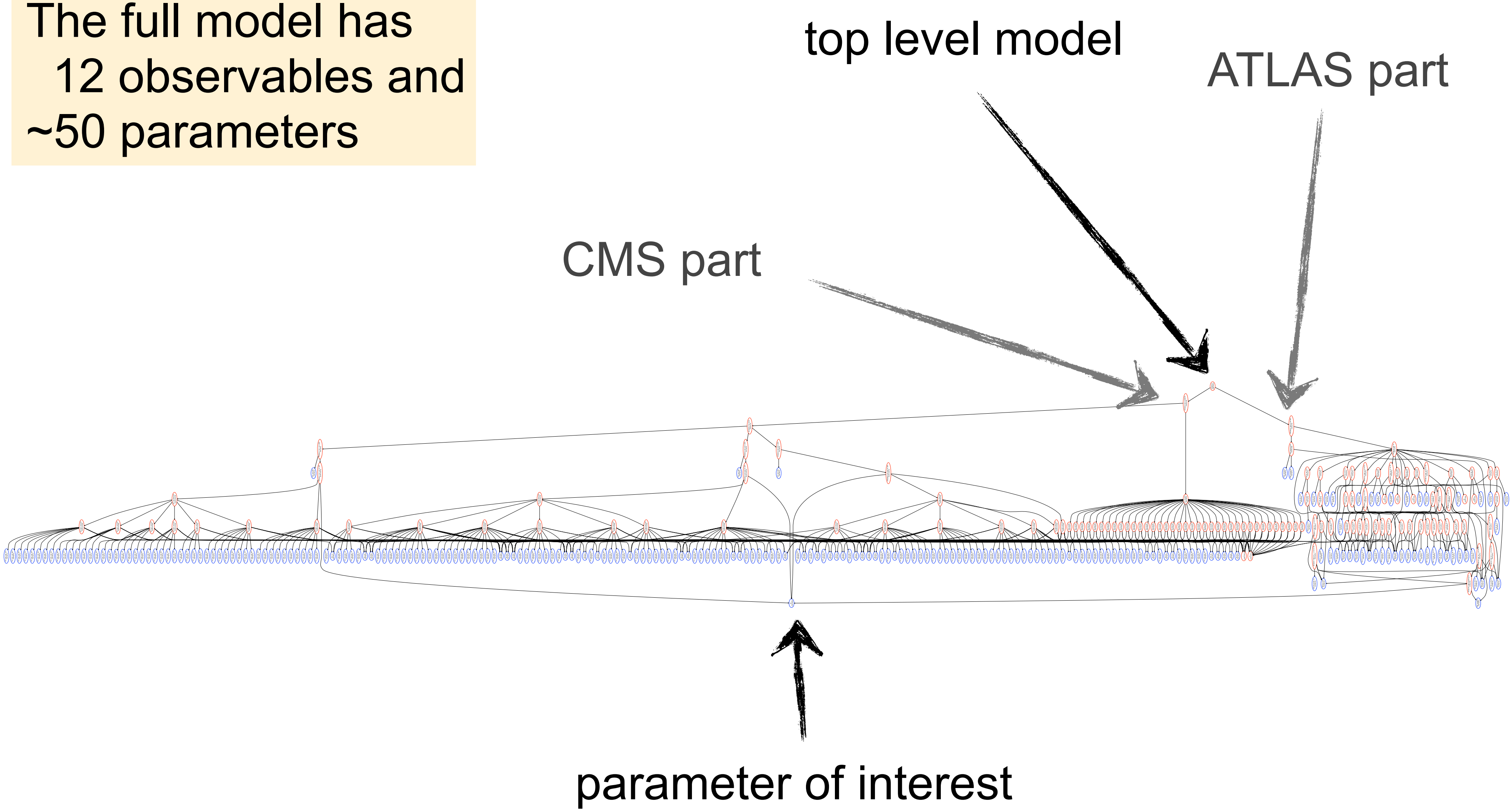
25 measurements from data

1 parameter of interest and 24 nuisance parameters

Visualization of the ATLAS+CMS Workspace

Center for
Cosmology and
Particle Physics

The full model has
12 observables and
~50 parameters



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

How to use the workspace

CENTER FOR
COSMOLOGY AND
PARTICLE PHYSICS

Using the likelihood function stored in a workspace does not mean:

- becoming an expert in RooFit/RooStats
- ever using a ROOT prompt.

To get started, I have a simple example:

- a C++ program with a main()
- with a real Makefile that links to the RooFit and RooStats libraries in ROOT

```
int main(int argc, char** argv) {

  evaluateLikelihoodFunction();

  return 0;
}
```

The program:

- opens the ROOT file and gets the workspace
- gets the log-likelihood function
- evaluates it at some random parameter points

```
void evaluateLikelihoodFunction(){

  // open root file with workspace
  // here you have to know name of workspace
  // you can get that by opening the file in root
  // and typing ".ls"
  // here we know it's called "w"
  TFile f("AnExampleWorkspace.root");
  RooWorkspace* w = (RooWorkspace*) f.Get("w");

  // now we need to get stuff out of the workspace
  // we can see what's inside with the Print() method
  w->Print();

  // the top-level PDF is called "model" in this case
  RooAbsPdf* model = w->pdf("model");
  // the only dataset is called "modelData"
  RooAbsData* data = w->data("modelData");

  // from the PDF and the data we can create the likelihood function
  RooNLLVar* nll = (RooNLLVar*) model->createNLL(*data);

  // now we need to know the parameters it depends on
  cout << " this likelihood depends on the following parameters"<<endl;
  RooArgSet* parameters = nll->getParameters(*data);
  parameters->Print();
  // get a pointer to the parameters
  RooRealVar* m1 = w->var("m1");
  RooRealVar* m2 = w->var("m2");

  cout << "\n\n--------------------------" << endl;
  cout << "Here are some random evaluations " << endl;

  for(int i=0; i<100; ++i){
    m1->randomize();
    m2->randomize();
    cout << "at (m1,m2) = ("<<m1->getVal()
         << ", " << m2->getVal()
         << ") -log likleihood = "
         << nll->getVal() << endl;
  }

}
```

A Critical Look at What We Are Doing Now

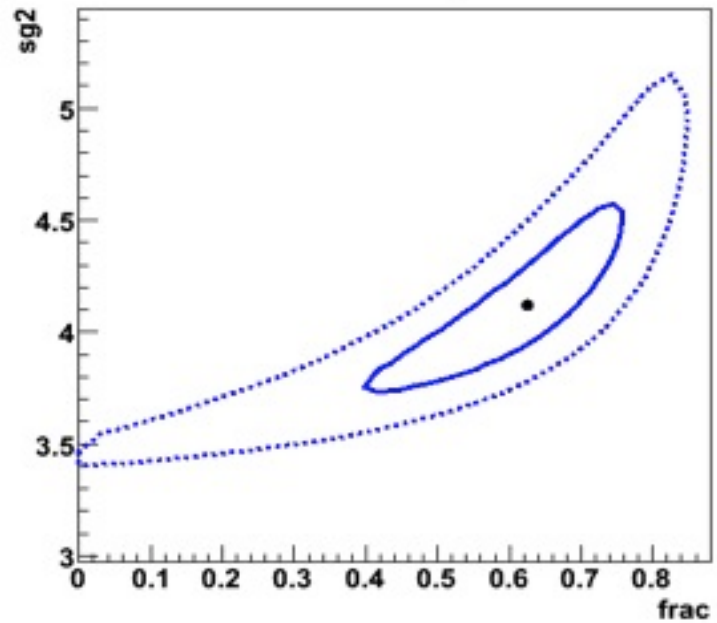
Correlated systematics

Clearly, several systematic effects will be correlated between the different measurements, and this must be taken into account

- That means the likelihood function needs to be a function of nuisance parameters.

Table 4: Endpoint positions for SU3 and SU4, in GeV. The first error is statistical, the second and third are the systematic and the jet energy scale uncertainty, respectively. The theoretical values are also given for ease of comparison to the left of the fitted values. The integrated luminosity assumed is 1 fb^{-1} for SU3 and 0.5 fb^{-1} for SU4.

Endpoint	SU3 truth	SU3 measured	SU4 truth	SU4 measured
$m_{\ell\ell q}^{\mathrm{edge}}$	501	$517 \pm 30 \pm 10 \pm 13$	340	$343 \pm 12 \pm 3 \pm 9$
$m_{\ell\ell q}^{\mathrm{thr}}$	249	$265 \pm 17 \pm 15 \pm 7$	168	$161 \pm 36 \pm 20 \pm 4$
$m_{lq(\mathrm{low})}^{\mathrm{max}}$	325	$333 \pm 6 \pm 6 \pm 8$	240	$201 \pm 9 \pm 3 \pm 5$
$m_{lq(\mathrm{high})}^{\mathrm{max}}$	418	$445 \pm 11 \pm 11 \pm 11$	340	$320 \pm 8 \pm 3 \pm 8$



Current approach:

- Not clear from a table like this if errors are anti-correlated
- The relationship can be non-trivial, not able to be represented by a simple covariance matrix

Wouldn’t it be nice to be able to inflate the uncertainties and see what happens? Especially, when results don’t agree.

Asymmetric Errors

CENTER FOR COSMOLOGY AND PARTICLE PHYSICS

I see that several groups interpret asymmetric errors reported by the experiments into a likelihood function:

- Of course, this is an ill-posed problem
- Rober Barlow considered 8 ways of doing this for cases in HEP, and the results can vary significantly

1: Traditional Method

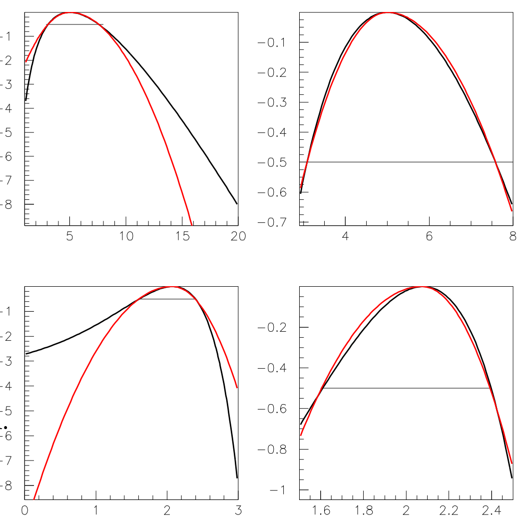
Split Gaussian with σ_+ above, σ_- below.

Top curves show Poisson $n = 5$.
Bottom curves show log Gaussian.

Horizontal line shows $[-\sigma_-, \sigma_+]$

Black is true likelihood. Red is model.

Right shows central region in detail.



Agreement fair in central region, not good outside.

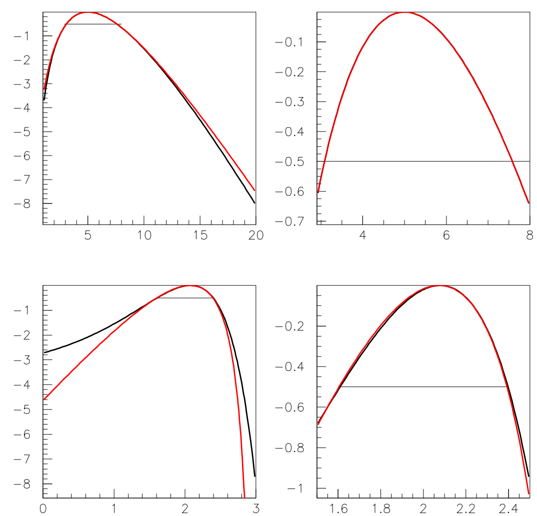
Roger Barlow:PHYSTAT05 *Asymmetric Statistical Errors* *Slide 4*

8: Gaussian - linear V

Same idea but change variance rather than σ

$f(a) = -\frac{1}{2}\frac{a^2}{V}$ with $V = V_0 + V'a$

$V_0 = \sigma_-\sigma_+ \qquad V' = \sigma_+ - \sigma_-$



Acceptable on log Gauss, Excellent on Poisson

Roger Barlow:PHYSTAT05 *Asymmetric Statistical Errors* *Slide 11*

http://www.physics.ox.ac.uk/phystat05/Talks/slides.pdf

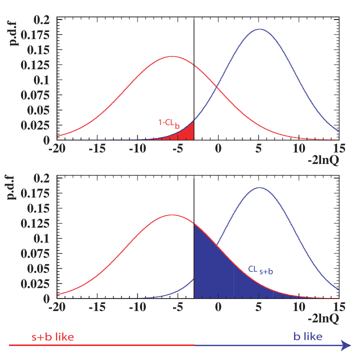
Likelihoods from CL_{s+b}

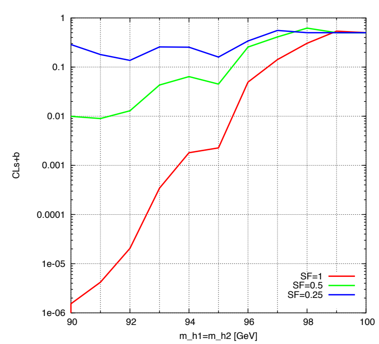
Center for Cosmology and Particle Physics

Because the experiments do not publish likelihoods, groups are often forced to try to infer the likelihood from other information, like CL_{s+b}

- The basis of this is that the CL_{s+b} was based on a LLR test statistic, which is assumed to be distributed as a chi-square distribution but let's look more closely
  - what is the "LLR" exactly?
  - Is it distributed as a chi-square distribution?

From Exclusions to Fits





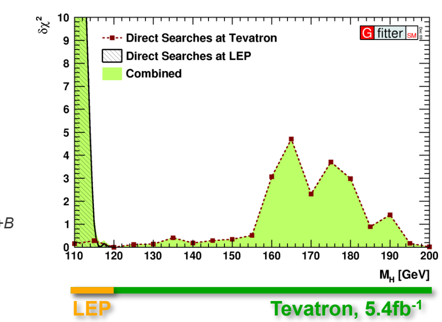
- Transfer (single-sided) CL_{s+b} into (double-sided) χ^2:
  $\chi^2_H = 2\,\mathrm{InvErf}^2(1 - 2CL_{s+b})$
- In the future (see below): Using HiggsBounds we can have χ^2 for each model with (pseudo)scalar particles
- Same for SUSY would be nice, but even more complicated…

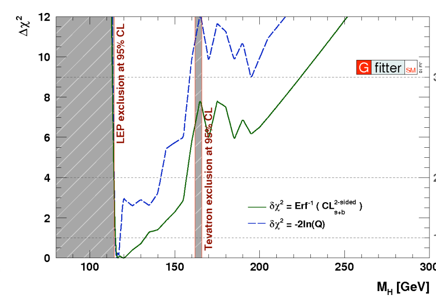
New: e.g. CL_{s+b} for MSSM/2HDM/…

P. Bechtle: HiggsBounds SUSYFit 2010 Workshop 26.07.2010 5

Statistical interpretation Higgs limits

- Stat. interpretation in global fit: 2-sided CL_{S+B}
  - Experiments measure test statistics: $LLR = -2lnQ$, where $Q = L_{S+B}/L_B$
  - LLR is transformed by experiments into one-sided CL_{S+B} using toy-MC experiments
    - Sensitive to too few Higgs-like events
  - We transform 1-sided CL_{S+B} into 2-sided CL^{2s}_{S+B}
    - We measure deviations from the SM.
    - Eg. also interested in too many Higgs-like events.
  - χ^2 contribution calculated via inverse error function, assuming symmetric pdf : $\delta\chi^2 = Erf^{-1}(1 - CL^{2s}_{S+B})$
- Alternative treatments (thanks to fruitful discussion with Tevatron people):
  - Use one-sided CL_{S+B}: however, different interpretation – want SM Higgs (not any Higgs)
  - Directly use $\delta\chi^2 \approx$ LLR: Bayesian interpretation, lacks pseudo-MC information





Max Baak Global Fit of electroweak SM and beyond

Three common "LLR" test statistics

We express cross-section as $\mu = \sigma/\sigma_{SM}$ for convenience.

Effect of systematics is parametrized by one or more "nuisance parameters" denoted ν .

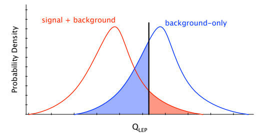
- best fit point is: $\hat{\mu}, \hat{\nu}$
- best fit of nuisance parameters with μ fixed is $\hat{\hat{\nu}}$ (aka "profiled")

In principle, s+b and b-only models can have different parametrizations

Three common test statistics used in the field are:

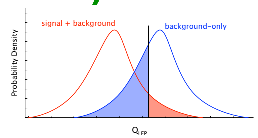
- simple likelihood ratio (used at LEP, nuisance parameters fixed)

$$Q_{LEP} = L_{s+b}(\mu = 1)/L_b(\mu = 0)$$



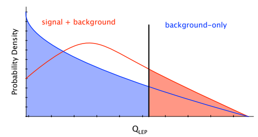
- ratio of profiled likelihoods (used commonly at Tevatron)

$$Q_{TEV} = L_{s+b}(\mu = 1, \hat{\hat{\nu}})/L_b(\mu = 0, \hat{\hat{\nu}}')$$



- profile likelihood ratio (related to Wilks's theorem)

$$\lambda(\mu) = L_{s+b}(\mu, \hat{\hat{\nu}})/L_{s+b}(\hat{\mu}, \hat{\nu})$$



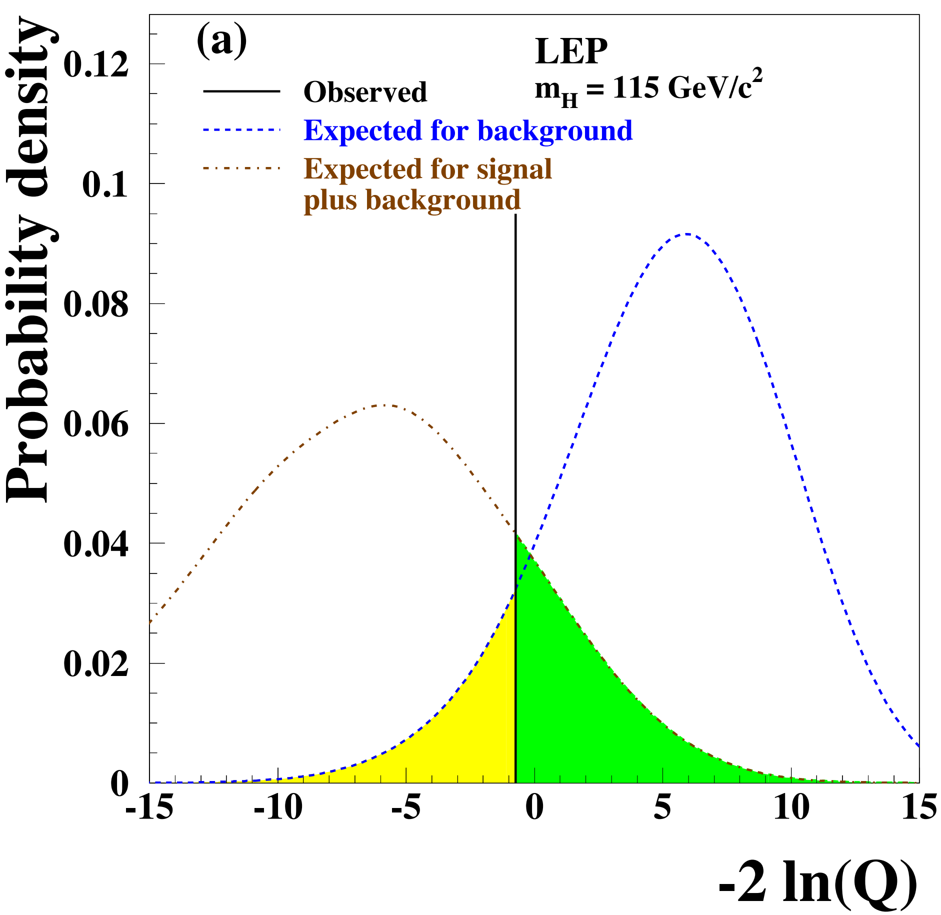
The LEP Test Statistic

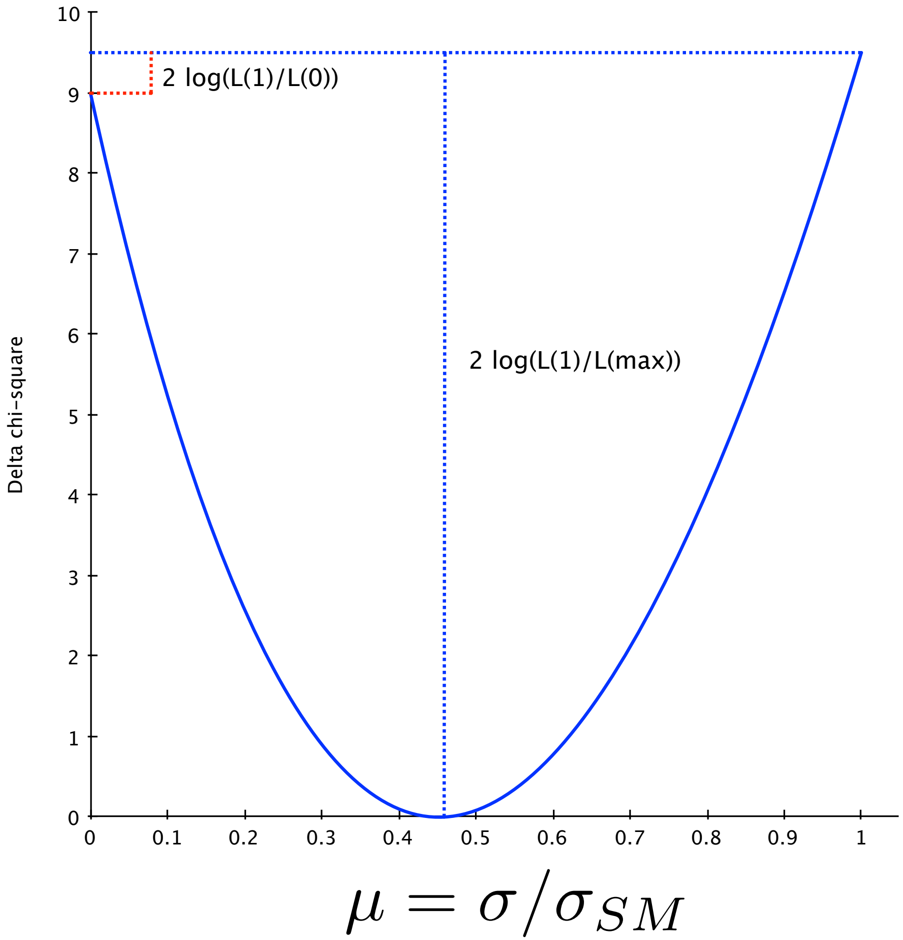
Center for Cosmology and Particle Physics

The "LLR" used by LEP was $Q_{LEP} = L_{s+b}(\mu=1)/L_b(\mu=0)$

- This is definitely not distributed as a chi-square,

To get a chi-square distribution, the denominator needs to be the best-fit point and the variable is non-negative





Moving Forward

Looking back, to look forward

CENTER FOR COSMOLOGY AND PARTICLE PHYSICS

Most of the LEP Higgs results were presented as exclusions in the space of some theoretical model

- SM Higgs mass assuming SM branching ratios and production rates
- exclusions in the $(m_A, \tan\beta)$ plane assuming some SUSY scenario

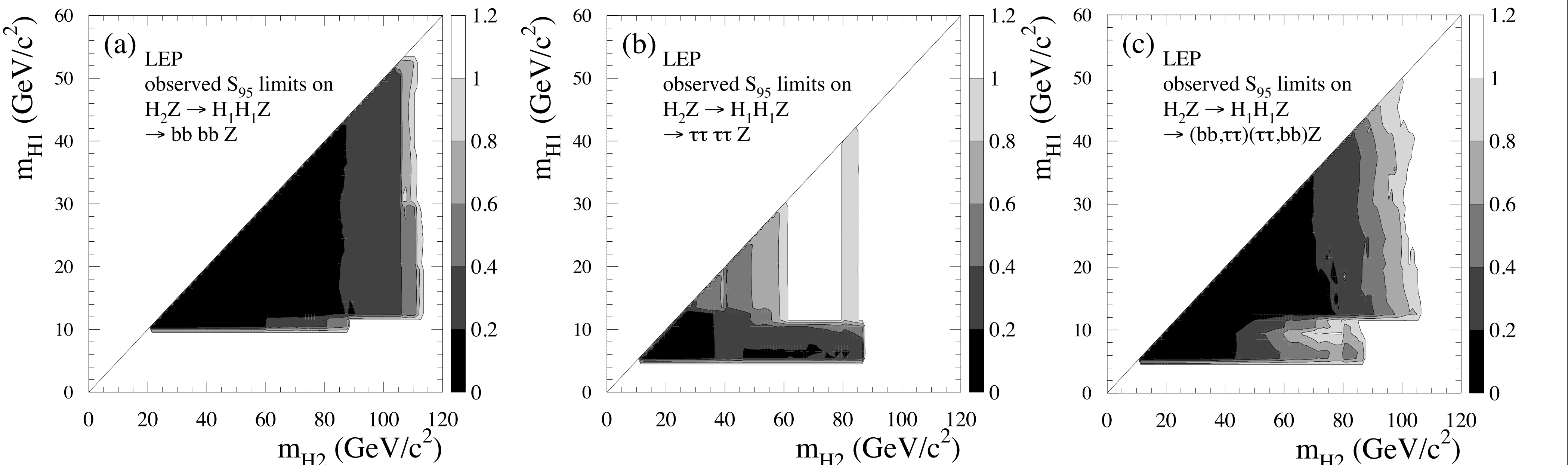
Recently, I've been considering non-standard models, and the most useful results are presented in terms of limits on cross section X branching ratio

- in terms of experimentally observable masses (see HiggsBounds)

But... when combining different decays or production modes, assumptions are made on the relative branching ratios and cross-sections -- not ideal.

(factor x SM cross section that corresponds to 95% exclusion)

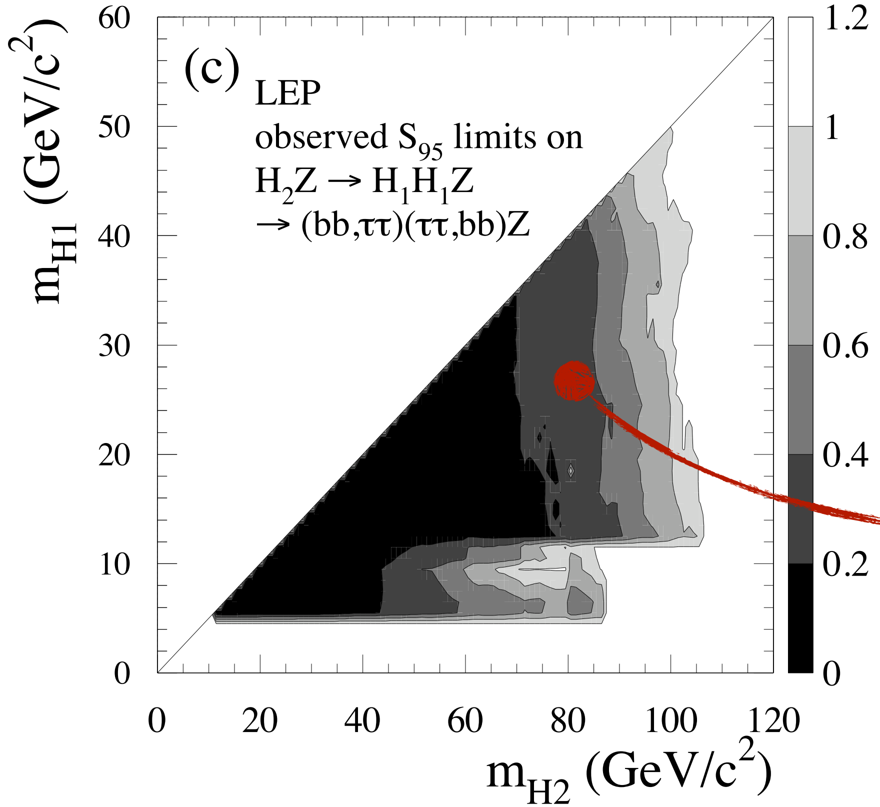
Eur.Phys.J. C33 (2004) [hep-ex/0602042]

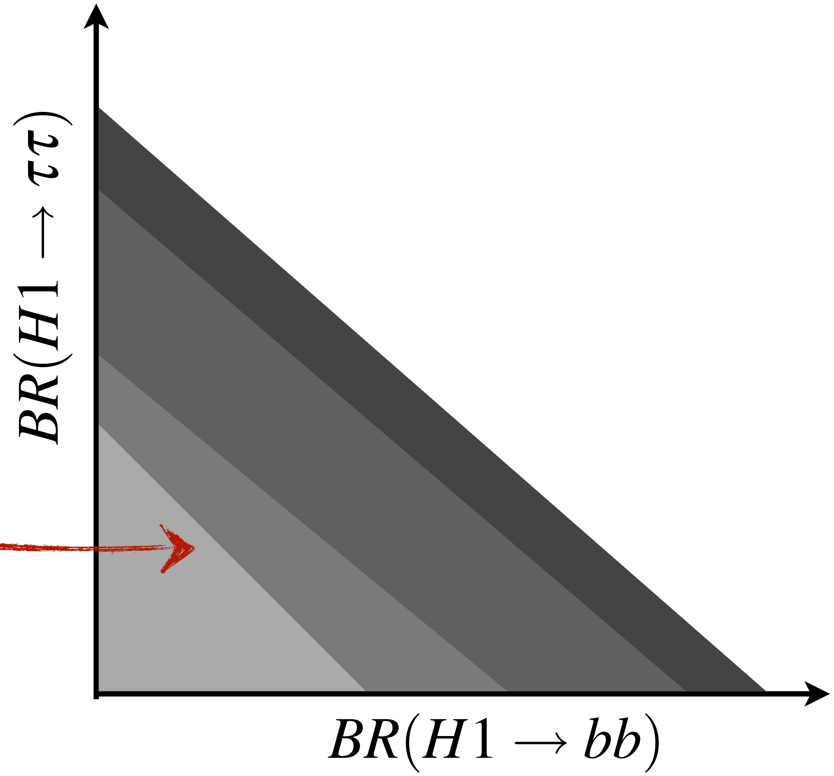


For mixed decays

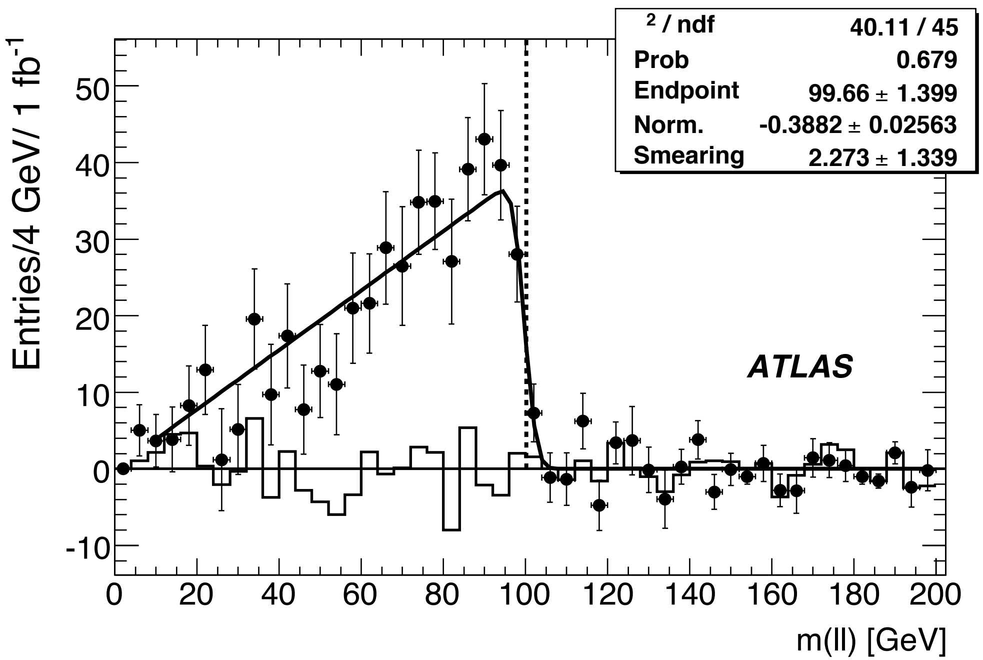
In the case of mixed decays, one would really like to be able know the limit for any point in the branching ratio space

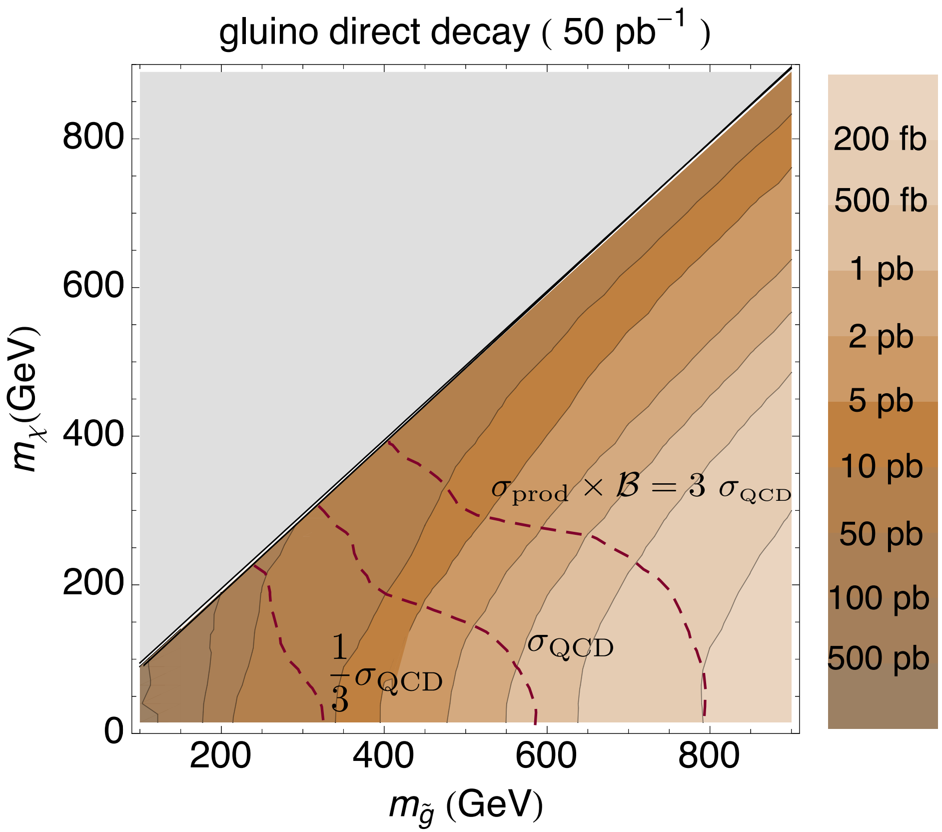
- the LEP analysis only considered 50/50 branching ratios
- difficult to publish on paper
- effect of branching ratios are easy to parametrize, discussed several times
- digital publishing or some type of online service is easy to imagine





Let's try to connect the two pictures



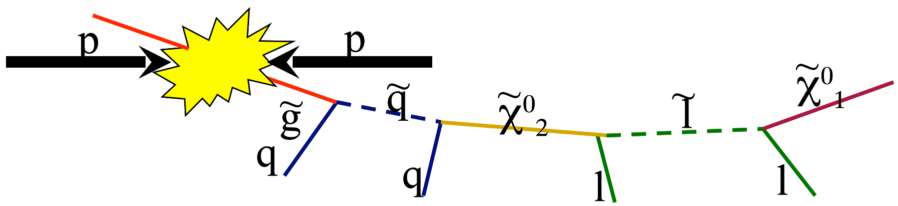


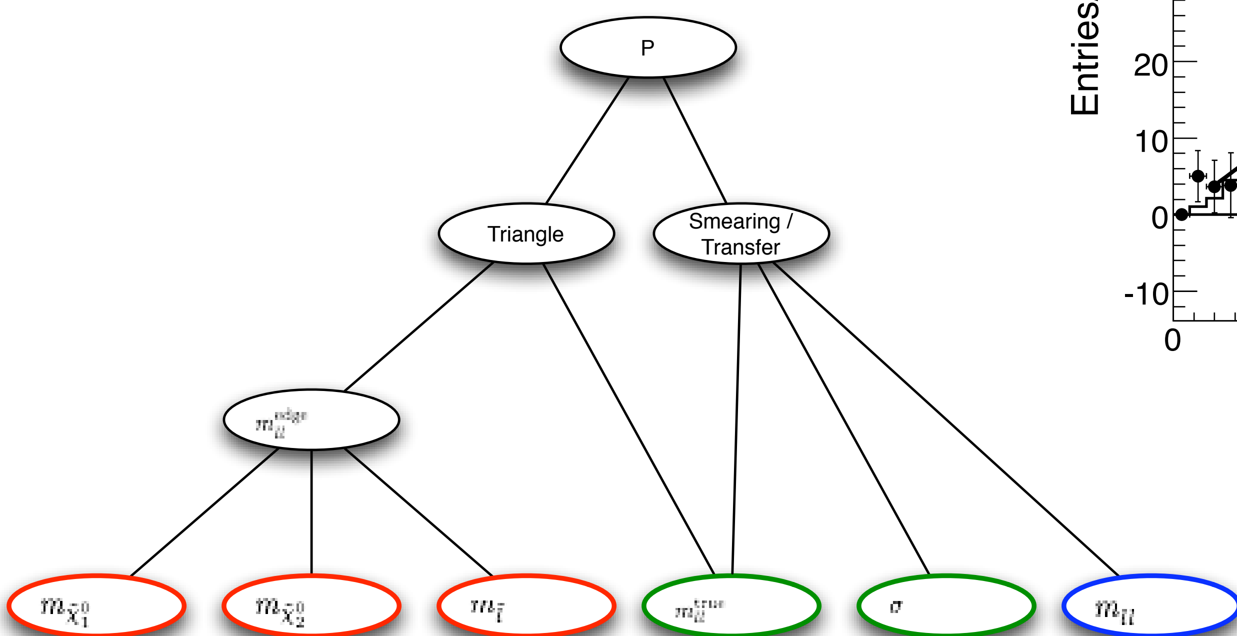
How will we model the data

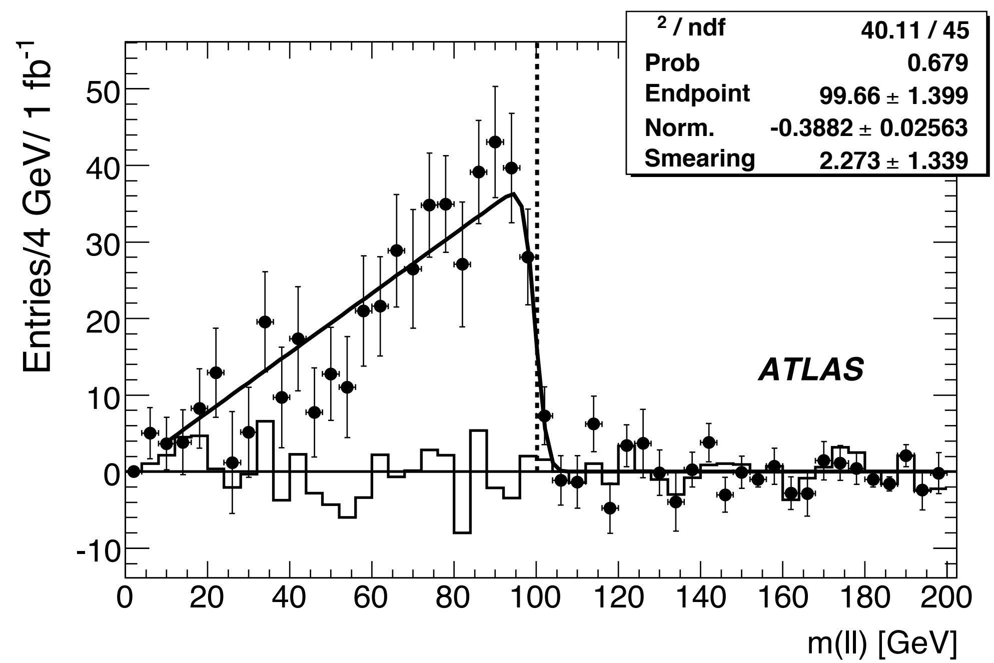
Center for Cosmology and Particle Physics

The technology is general, but it is still up to the experiments to decide how they will model the data.

- Thoughtful parametrization requires planning and coordination between theorists and experimentalists (eg. groups like this)







$$P(m_{ll}|m_{\tilde{\chi}_1^0}, m_{\tilde{\chi}_2^0}, m_{\tilde{l}}, \sigma) = \text{Triangle}(m_{ll}^{true}, m_{ll}^{\text{edge}}(m_{\tilde{\chi}_1^0}, m_{\tilde{\chi}_2^0}, m)) \oplus \text{Smearing}(m_{ll}^{true}, m_{ll})$$

$$m_{ll}^{\text{edge}} = m_{\tilde{\chi}_2^0}\sqrt{1-\left(\frac{m_{\tilde{l}}}{m_{\tilde{\chi}_2^0}}\right)^2}\sqrt{1-\left(\frac{m_{\tilde{\chi}_1^0}}{m_{\tilde{l}}}\right)^2}$$

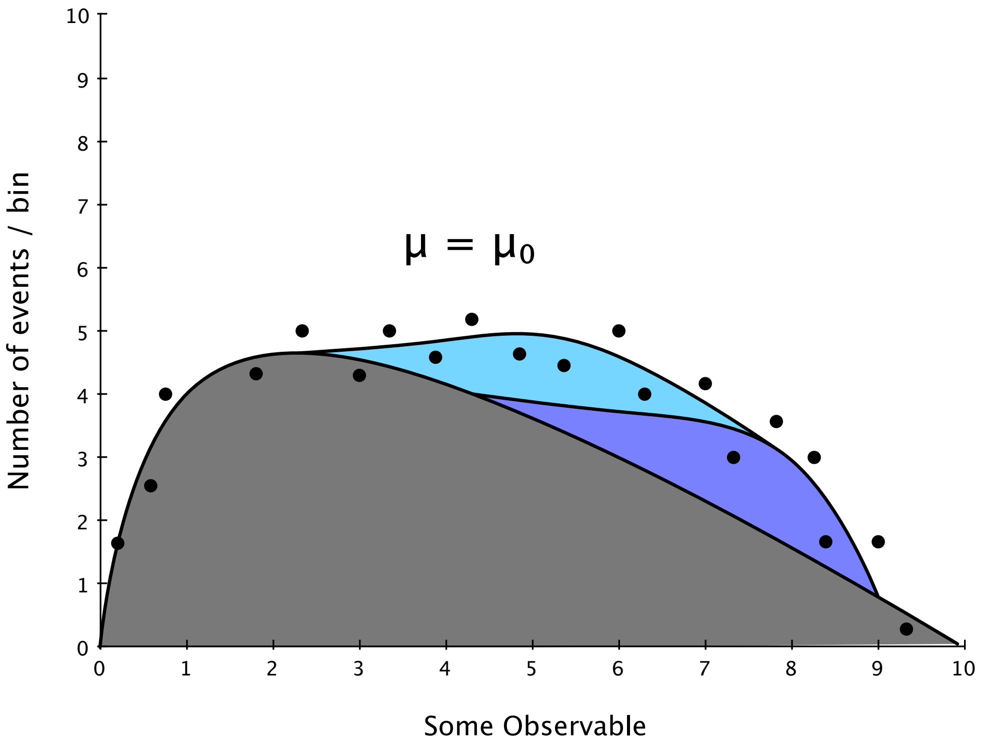
Acceptance

As we move around the model's parameter space the distributions change, thus changing the signal efficiency and acceptance.

- This is harder to parametrize than branching ratios fro different topologies

Remember that even at this fixed point in the model's parameter space, the efficiency and acceptance can change as you vary the nuisance parameters associated with systematic effects.

- at first, maybe we can neglect this effect and it's an adequate approximation



Acceptance

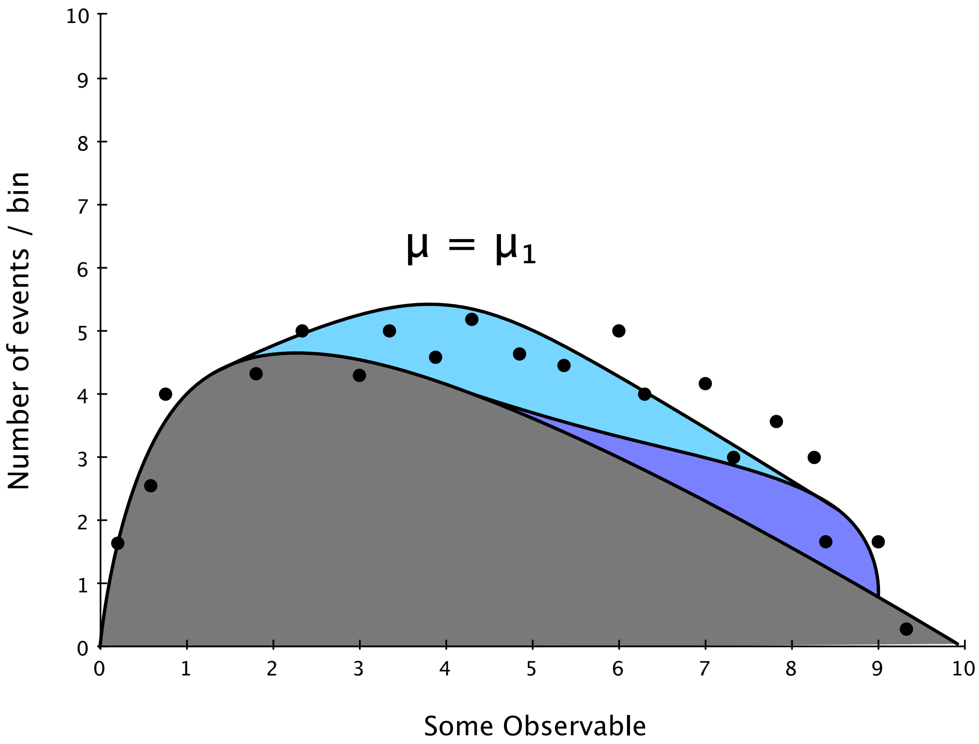
Center for Cosmology and Particle Physics

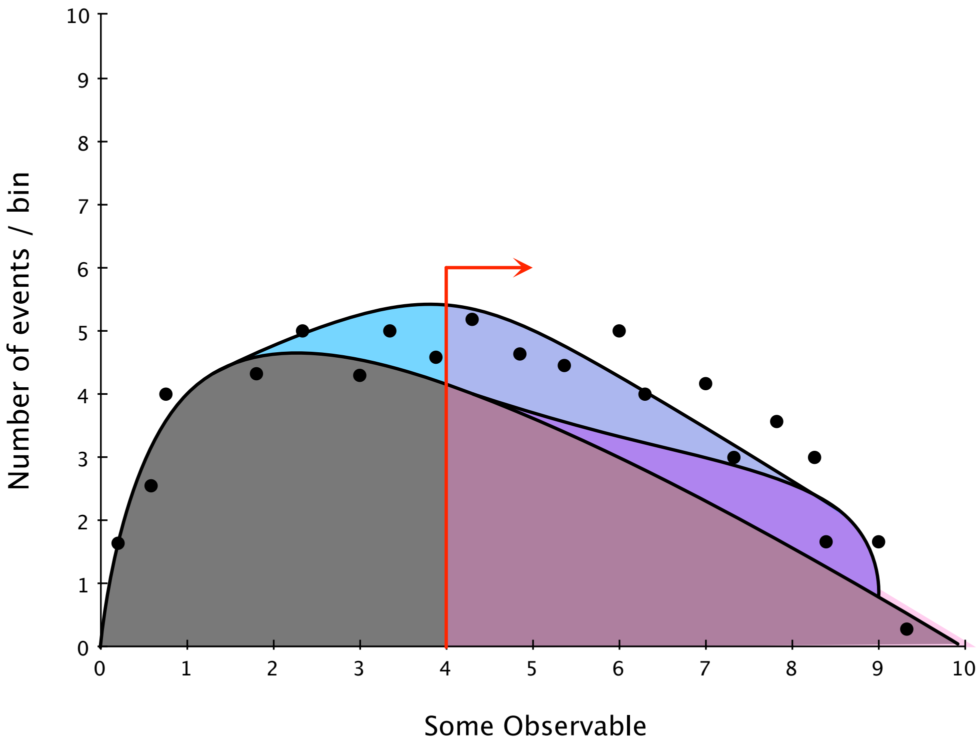
As we move around the model's parameter space the distributions change, thus changing the signal efficiency and acceptance.

- This is harder to parametrize than branching ratios fro different topologies

Remember that even at this fixed point in the model's parameter space, the efficiency and acceptance can change as you vary the nuisance parameters associated with systematic effects.

- at first, maybe we can neglect this effect and it's an adequate approximation





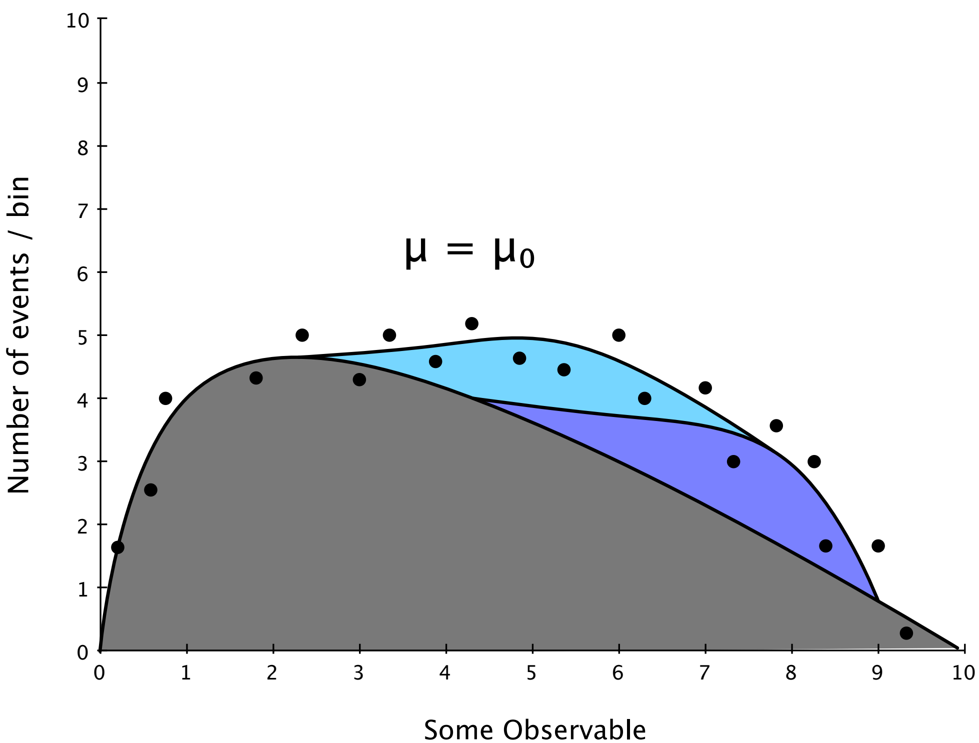
Acceptance

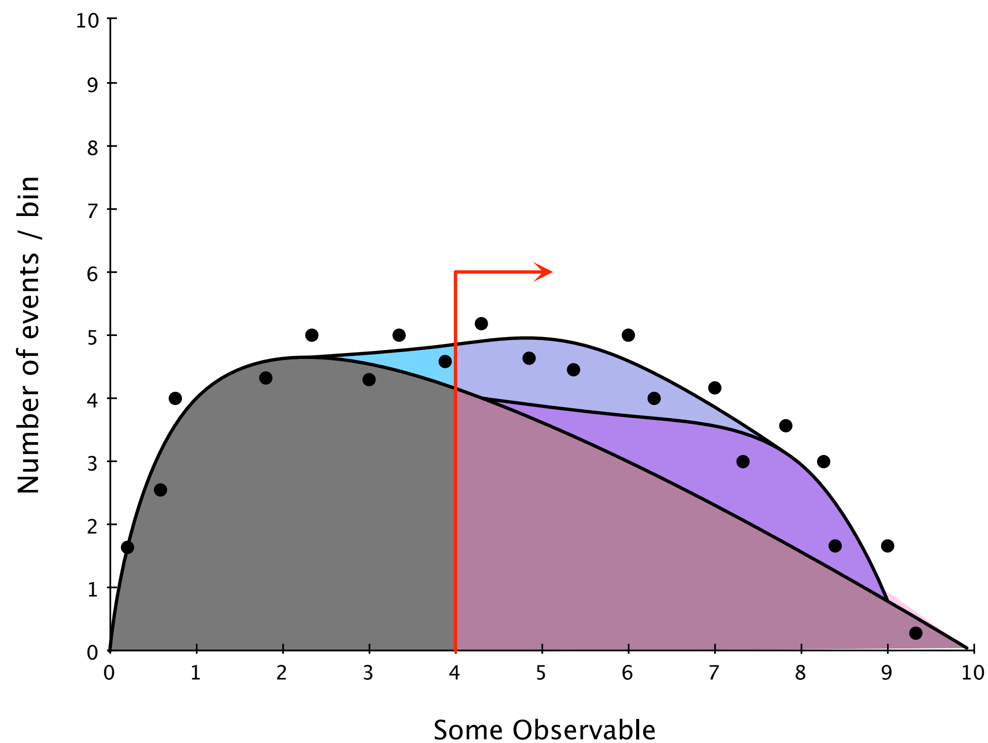
As we move around the model's parameter space the distributions change, thus changing the signal efficiency and acceptance.

- This is harder to parametrize than branching ratios fro different topologies

Remember that even at this fixed point in the model's parameter space, the efficiency and acceptance can change as you vary the nuisance parameters associated with systematic effects.

- at first, maybe we can neglect this effect and it's an adequate approximation



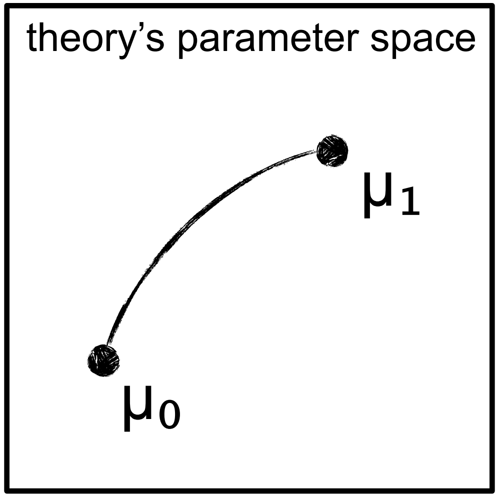


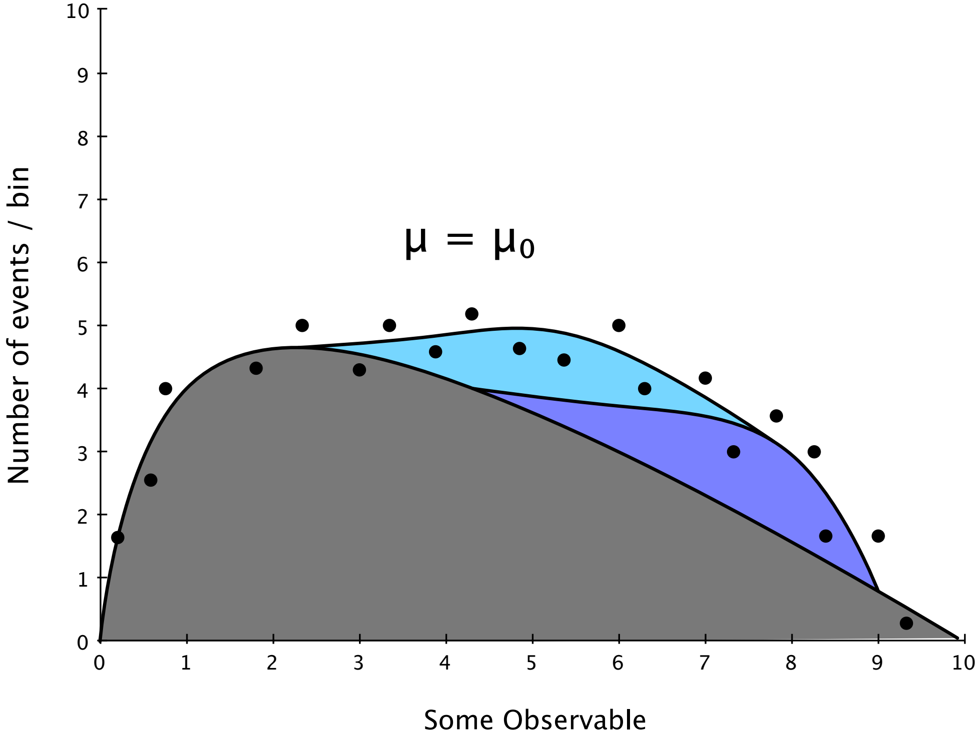
Interpolating

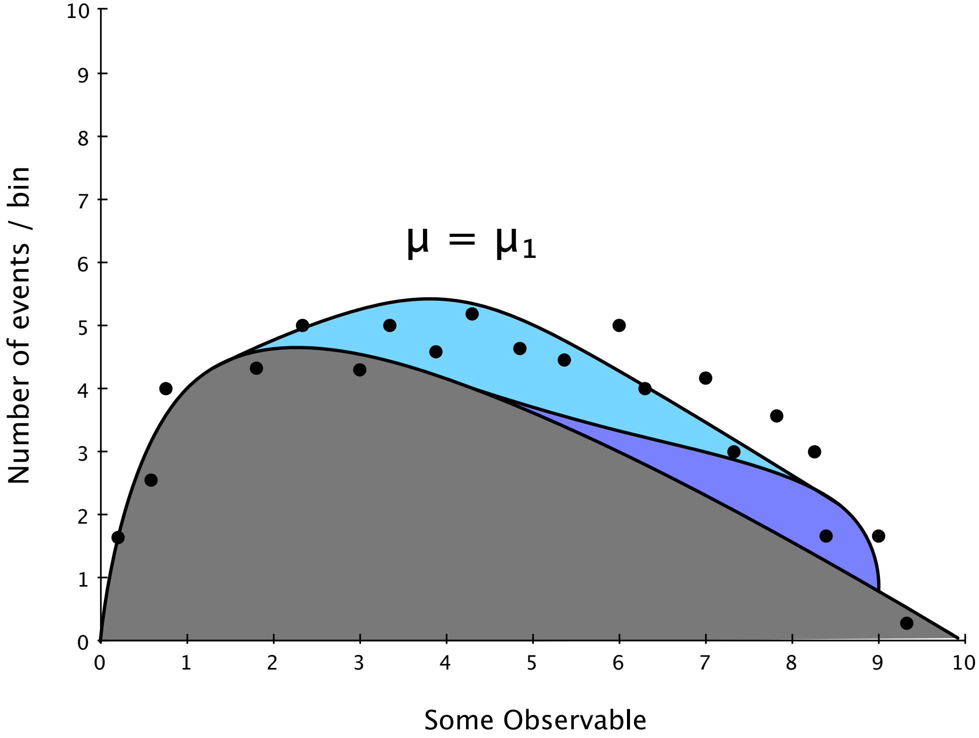
Center for Cosmology and Particle Physics

Most statistical techniques require ability to evaluate likelihood at arbitrary points in the parameter space ⇒

- Either need to have evaluated model at sufficiently many discrete parameter points
- Or have a way of interpolating expected signal distribution (including efficiency & acceptance)
  - Often by interpolating between template histograms







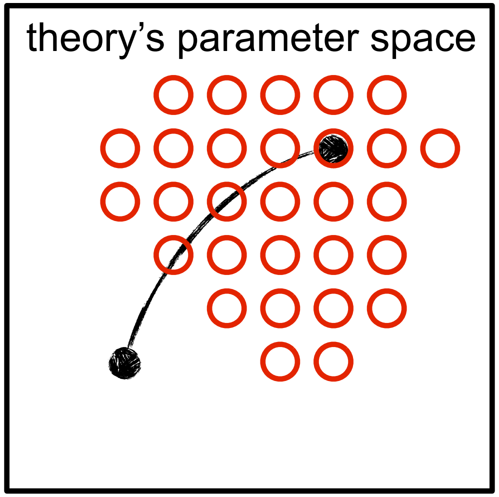
Extrapolating

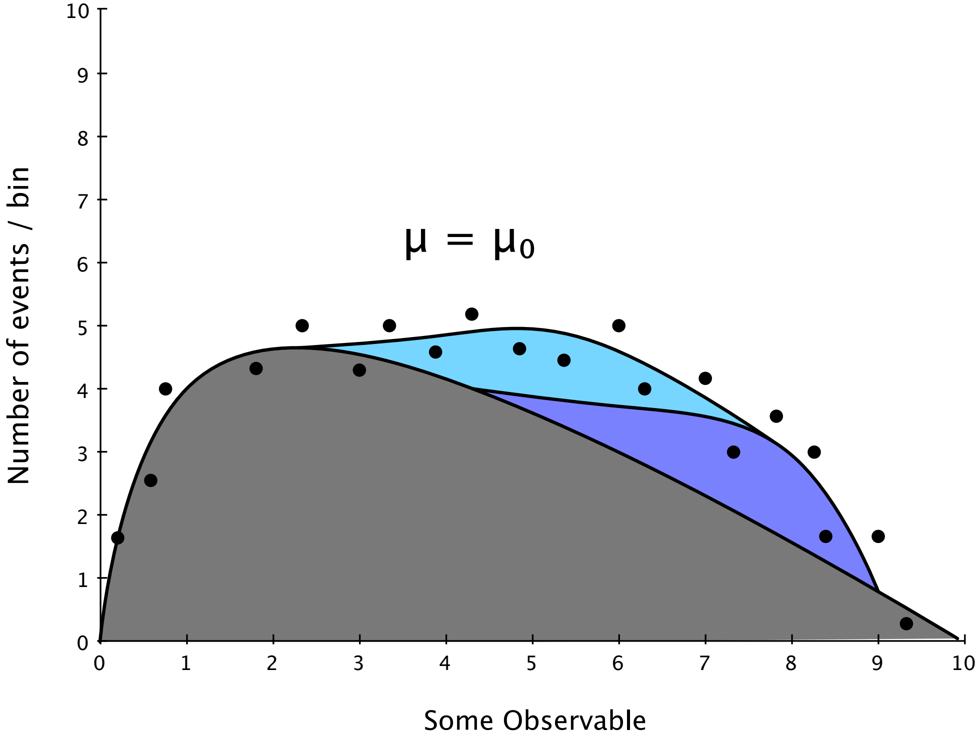
Initially the experiments may scan some initial set of model points.

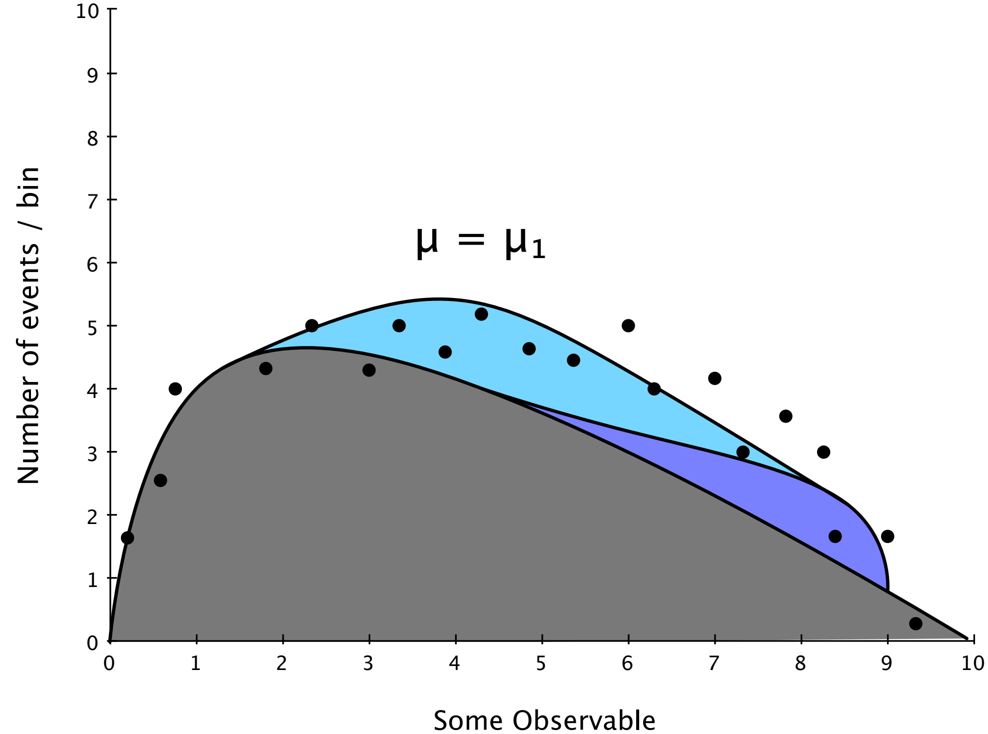
- These define a domain of validity for the model

How does one go to model points outside this set?

- if the experiments had a service to provide signal templates for new model points, then one can interpolate between these new anchor points.







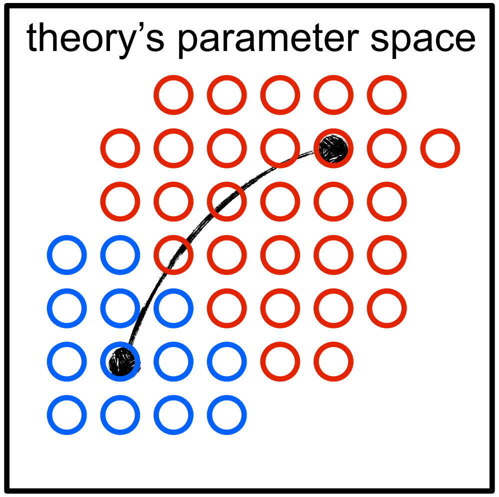
Extrapolating

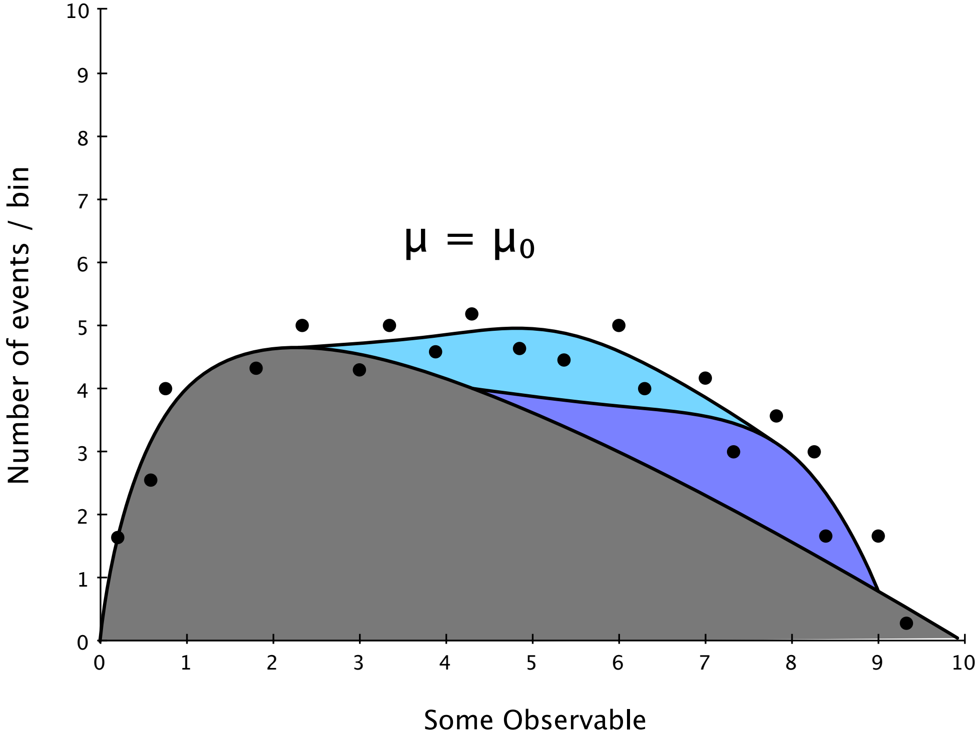
Initially the experiments may scan some initial set of model points.

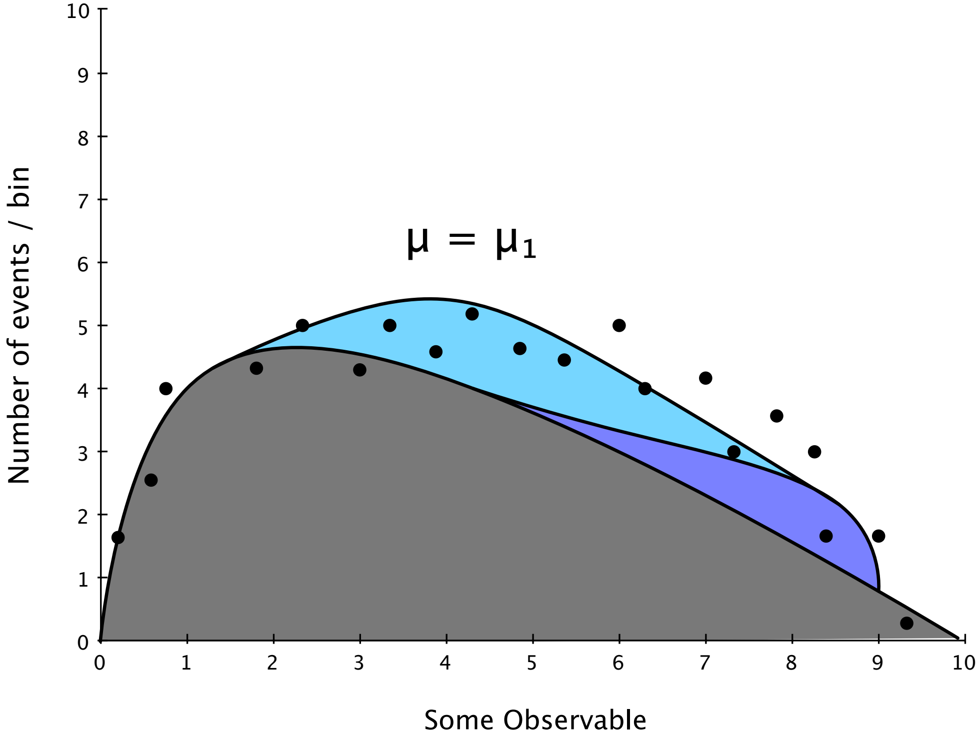
- These define a domain of validity for the model

How does one go to model points outside this set?

- if the experiments had a service to provide signal templates for new model points, then one can interpolate between these new anchor points.



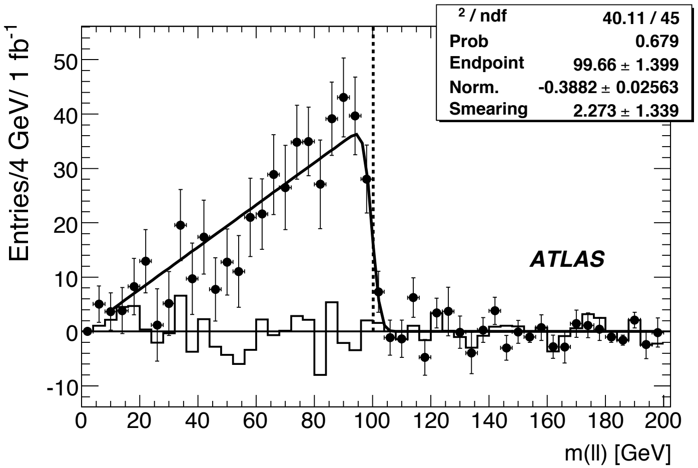




Moving Forward

It seems the best way to move forward is to prepare some workspaces corresponding to benchmark tests.

- This will be very helpful for comparison of different fitters
- a very simple multivariate Gaussian for degugging
- eg. a SUSY ll_{max}, ql_{low}, ql_{high}, qll_{thresh}, qll_{edge} example
  - start with only the measured edges
  - could extend to the full shape
- A Higgs example:
  - h/H -> tau tau with both taus visible?
  - the toy ATLAS+CMS combination?



Do the Intervals Cover?

Michael Bridges, KC, Farhan Feroz, Mike Hobson, Roberto Ruiz de Austri, Roberto Trotta

Similar work by Yashar Akrami

Coverage

Coverage is the probability that interval contains (covers) the true value

- Property of the method used to produce confidence/credible interval
- For any given data, the interval either covers or it doesn't
- Requires repeating the procedure on pseudo-data several times

Hard-core frequentist intervals "cover by construction"

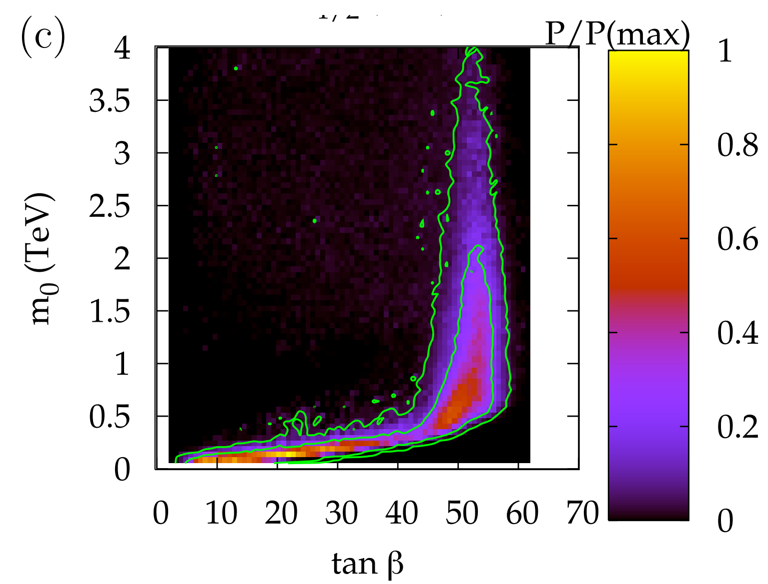
- Neyman Construction, aka "Feldman-Cousins" in HEP
- Likelihood-based intervals **are not** guaranteed to cover
- Bayesian methods **are not** guaranteed to cover (not their goal)

I think of coverage as a **calibration of our statistical apparatus**

- How often do the intervals cover the true value?
- Is it really 95%?

What should be the view today: Objective Bayesian analysis is the best frequentist tool around.

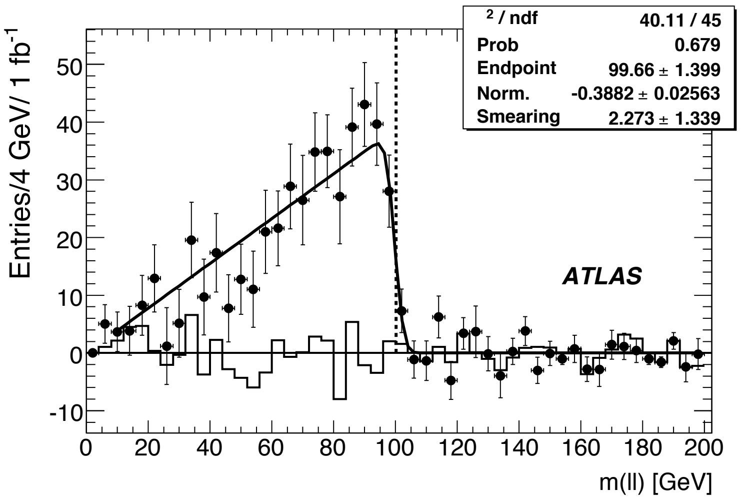
-Jim Berger



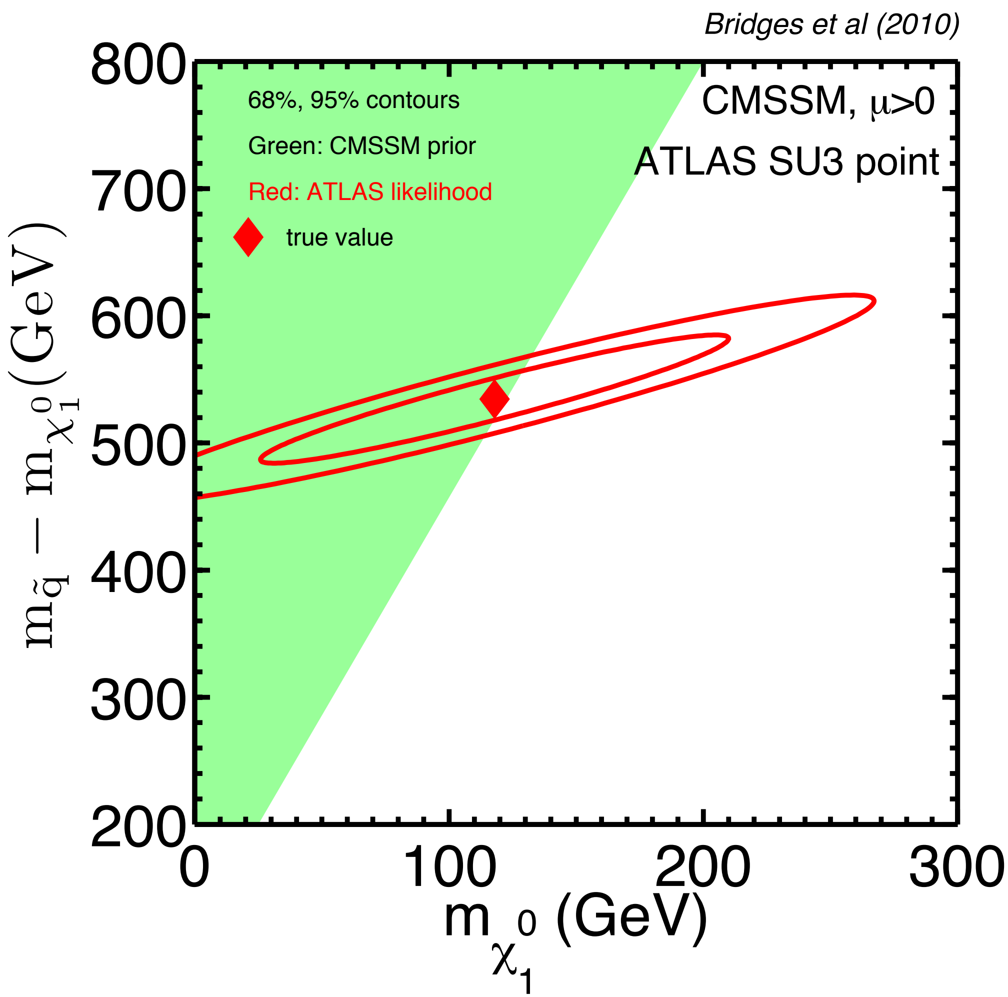
Do our current intervals cover?

To study this we considered a simplified model based on the ATLAS analysis of the “SU3” benchmark point.

- **Model**: A multivariate Gaussian likelihood function based on the published ATLAS covariance matrix SM. Is likelihood really a multi-variate Gaussian?



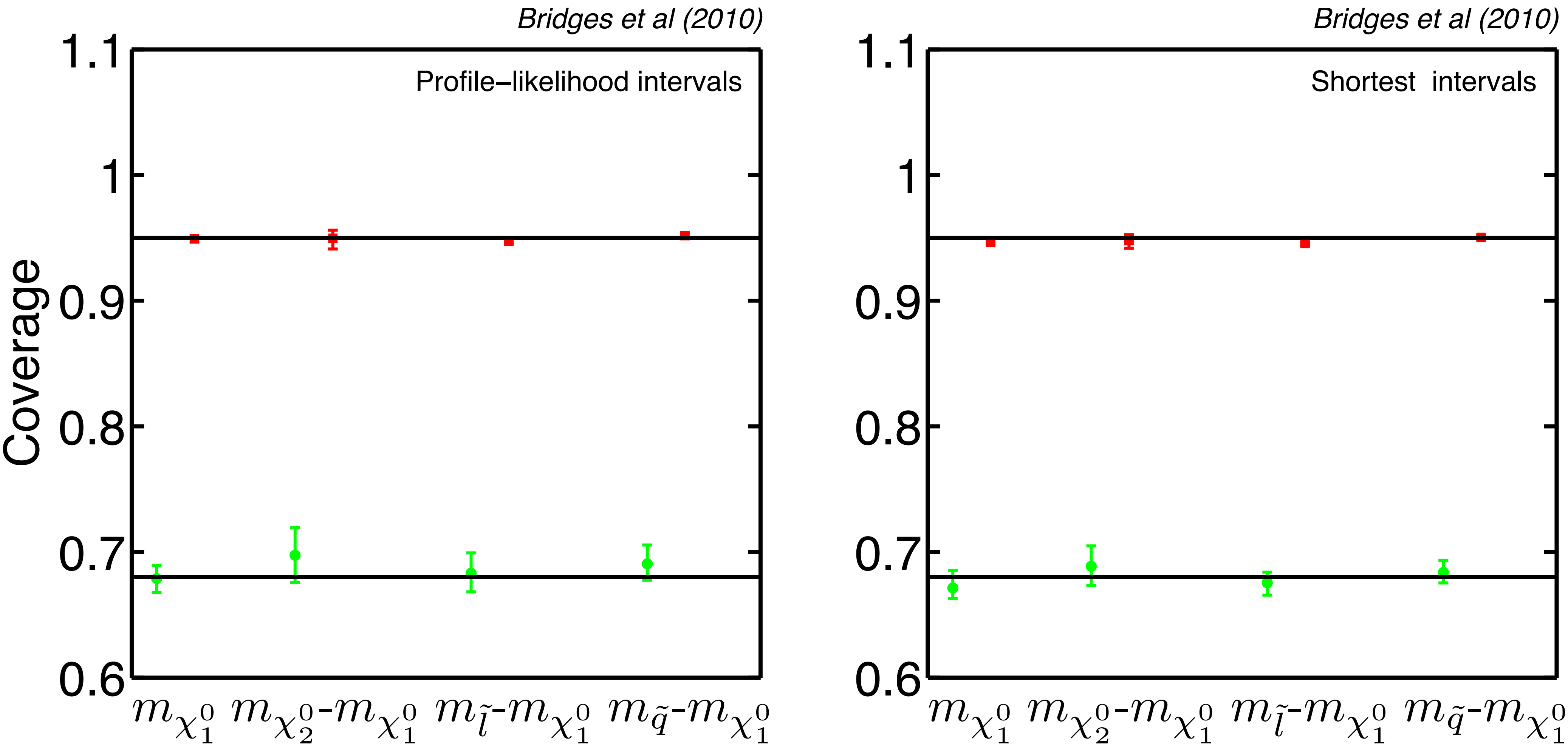
Observable	SU3 m_{meas} [GeV]
$m_{\tilde{\chi}_1^0}$	$88 \pm 60 \mp 2$
$m_{\tilde{\chi}_2^0}$	$189 \pm 60 \mp 2$
$m_{\tilde{q}}$	$614 \pm 91 \pm 11$
$m_{\tilde{\ell}}$	$122 \pm 61 \mp 2$
Observable	**SU3 Δm_{meas} [GeV]**
$m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0}$	$100.6 \pm 1.9 \mp 0.0$
$m_{\tilde{q}} - m_{\tilde{\chi}_1^0}$	$526 \pm 34 \pm 13$
$m_{\tilde{\ell}} - m_{\tilde{\chi}_1^0}$	$34.2 \pm 3.8 \mp 0.1$



Initial Coverage Result

Start by checking coverage of the weak-scale model:

- no pull-back to CMMS, parameters are mean of multivariate Gaussian



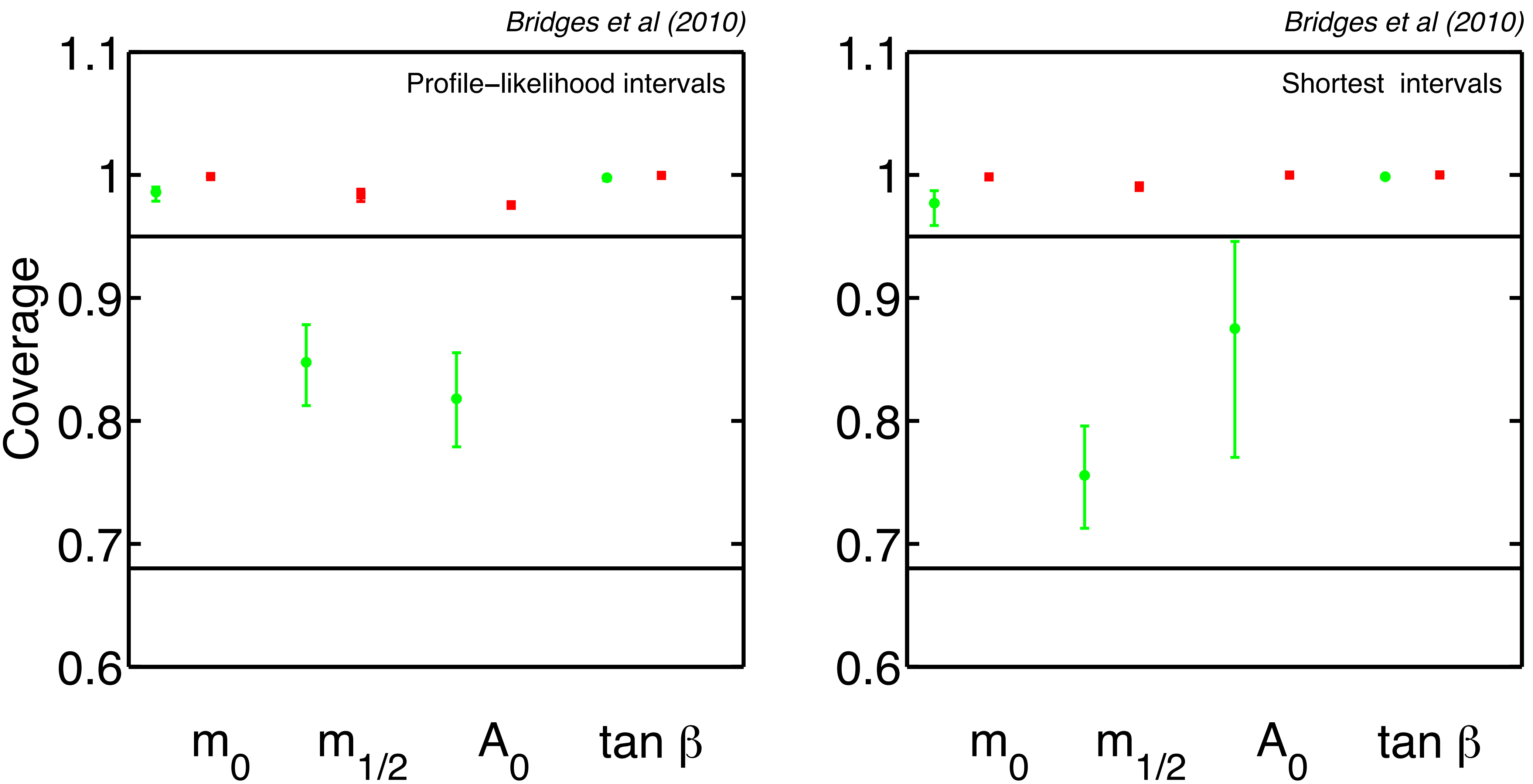
This "has to work". If it didn't, would be an algorithmic problem.

Initial Coverage Result

CENTER FOR COSMOLOGY AND PARTICLE PHYSICS

When we pull back to the CMSSM params., we see significant over-coverage

- consistently with profiling, MCMC, Multinest, etc.

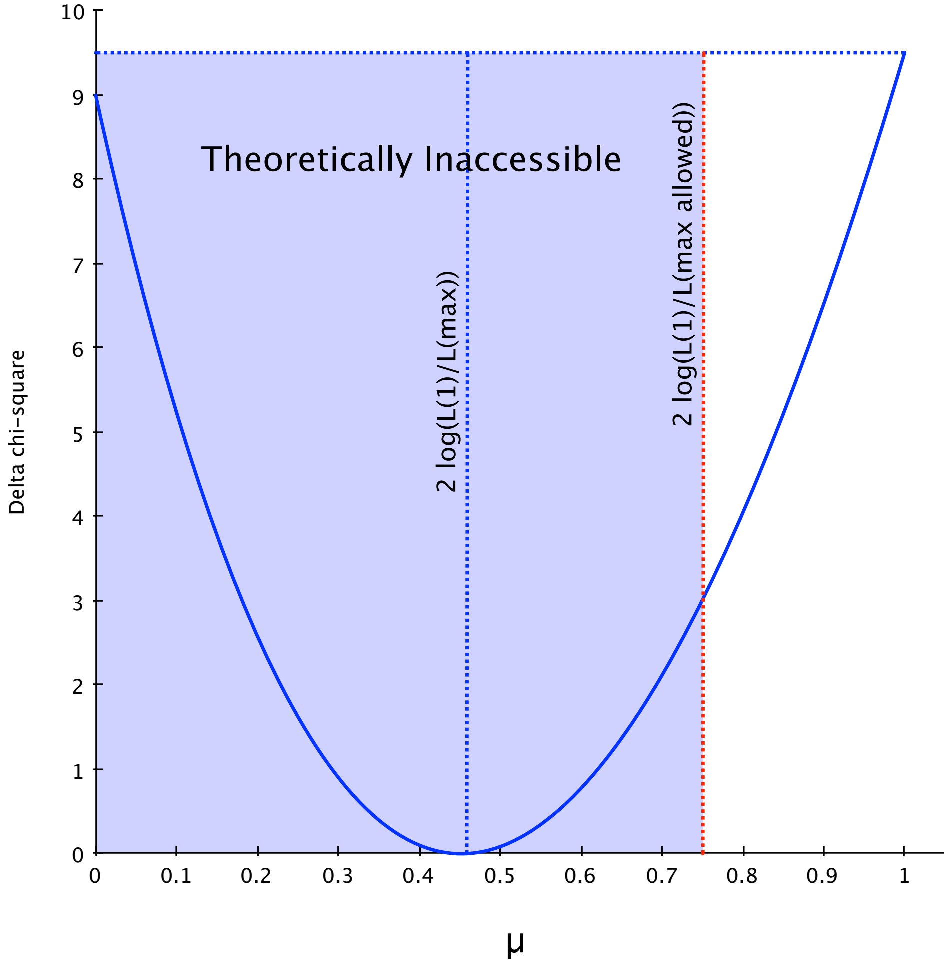


What is the source of the over-coverage?

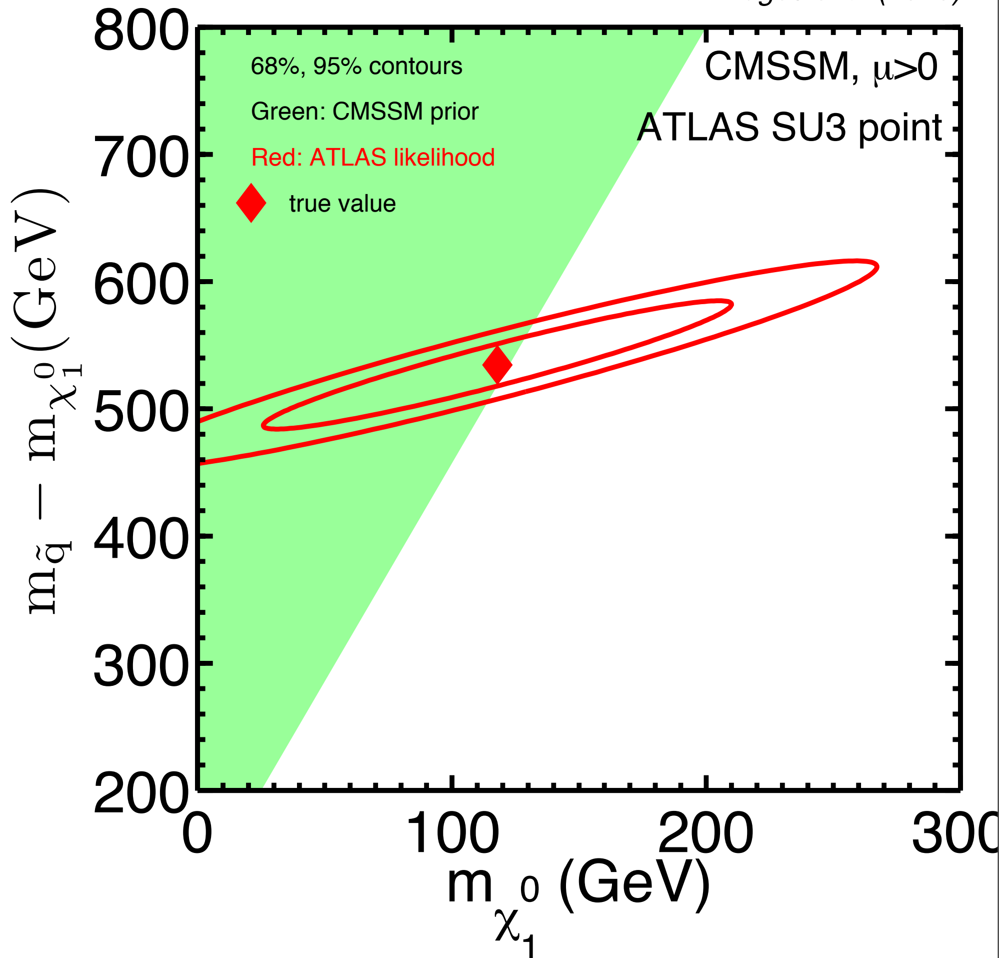
Effect of boundaries

The requirements that a CMSSM point is physical (LSP, EWSB, Tachions) introduce boundaries in the parameter space.

- These boundaries mean convergence to a χ^2 distribution (Wilks) is slow
- leads to a higher cut-off on -2 ln L ⇒ larger interval ⇒ over-coverage



Bridges et al (2010)

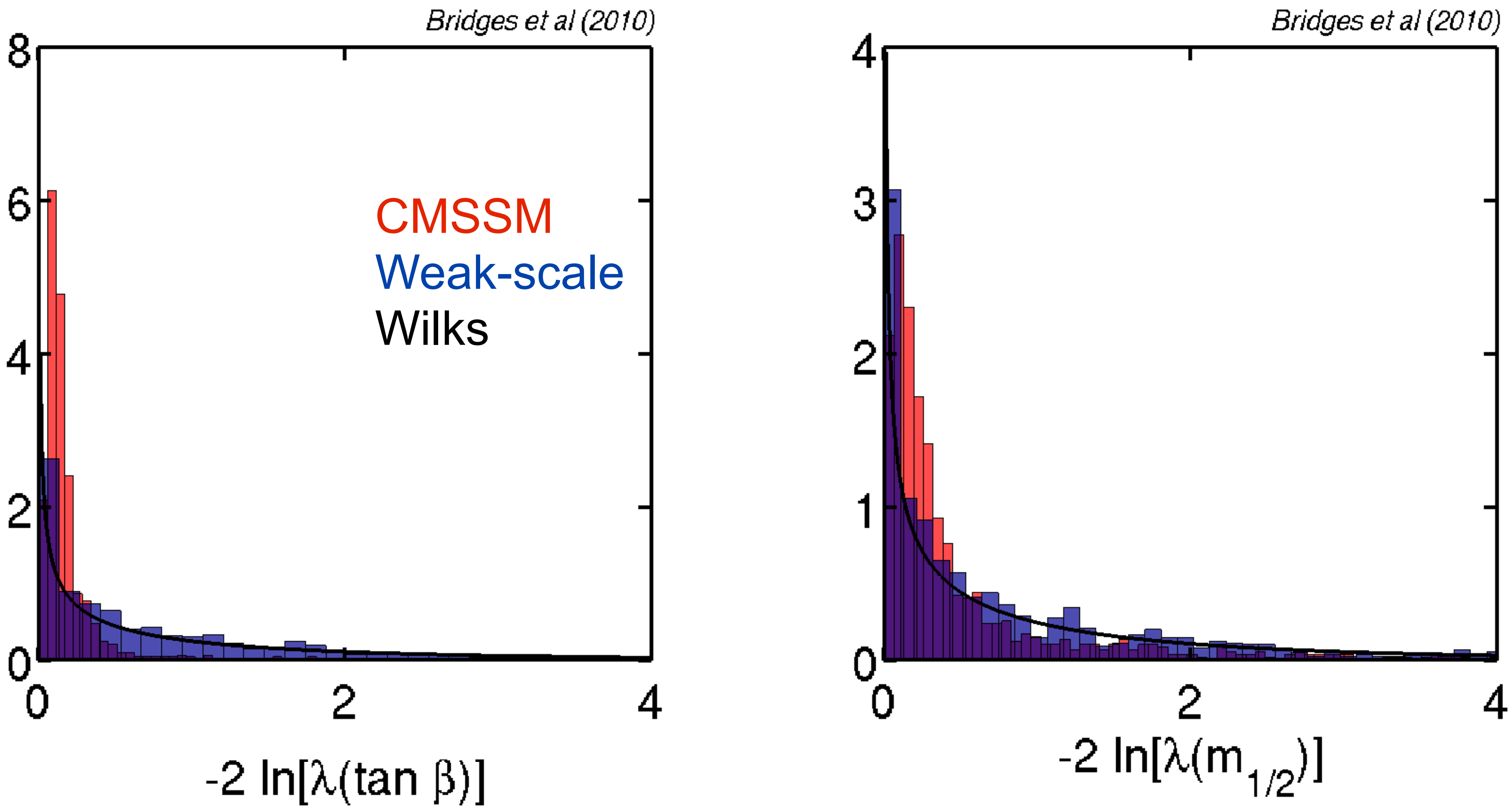


Checking "Validity" of Wilks

Center for
Cosmology and
Particle Physics

To check that this was the effect, we plot $-2\ln\lambda$ evaluated at true point

- Confirms expectation: distribution is **not** χ^2 for CMSSM (is for weak-scale)



Conclusions

Fitting groups are very active, work mainly in context of specific models

- basis for their inference is the likelihood function, which is typically written by the fitting groups from inferring what is written in the experimental papers
- may be a poor approximation for the true likelihood
- typically does include a reasonable treatment of the systematics

The RooStats project has reached a certain level of maturity and is rapidly being adopted by the LHC experiments

- The toy ATLAS/CMS Higgs combination was a milestone for the project

The workspace technology that is so important for combinations (of different channels within an experiment or between experiments) also provides enormous opportunity for communicating experimental results to the fundamental lagrangian fitters.

- Given the effort that is going into making the RGE's more precise and the fitting techniques, we should make sure the inputs (likelihood functions) are sensible

In order to move forward, we should

- agree on some useful benchmark examples and prepare the workspace for them so different tools can start working on their interfaces to the workspace
- realize that a thoughtful parametrization of the model requires planning and clear requests to the experimental community
- make a clear request to and/or mandate for the experiments and provide them