LHC Higgs Combination

Kétévi A. Assamagan
BNL

Please check this out:

Elusive Higgs slips from sight again
http://www.newscientist.com/article/dn20442

Introduction

- **LHC Higgs combination – LHC-HCG – initiated by ATLAS and CMS Spokespersons and Physics coordinators**

Mandate

- Define the strategy for statistical combination of Higgs results from LHC experiments
  - for EPS & Lepton-Photon conferences (summer 2011) and beyond
- Prepare the inputs needed for the combination
- Produce and help disseminate the combined results for SM Higgs
- Consult with the Statistics committees of ATLAS and CMS regarding the statistical tools and procedures employed
- Consult with the Higgs cross section working group for input on various cross sections and branching ratios
- Several follow up meetings of ATLAS & CMS Higgs coordinators
  - Developed draft memo of understanding & timeline of work

Timeline

Date	Goal / Achievement
Dec 6	Kick-off open meeting (OM) to discuss combination strategy
January	Compare Roostats & other commonly used tools
February	Stat methods, $H\rightarrow WW$ inputs
March	Precisions $H\rightarrow WW$+0 jet comparisons, 1% agreement on limits reached
April 7	Discuss $H\rightarrow\gamma\gamma$, $H\rightarrow ZZ$ inputs. 1-2% agreement on $H\rightarrow WW$+0/1/2 jets limits
April	Combine multiple channels, finalise stat. methods
May 12	OM to report on combination exercise on $H\rightarrow$ WW, ZZ, $\gamma\gamma$
May 19	WM to share summer analysis strategy, analysis definition. Prepare for combinations with data, start documentation on combined results
June30	WM to share prelim results with data, start combination process; exact process of sharing to be determined by SP & PC
July 14	WM on Combined results for EPS, review documentation
July 15	Start approval process within ATLAS & CMS; exact procedure to be determined by SP & PC
July 21	**EPS starts in Grenoble ; option to update for LP'11**

Initial Composition

Role	ATLAS	CMS
Higgs WG convener	Bill Murray	Vivek Sharma
Overall Contact	Kétévi A. Assamagan	Andrey Korytov
Statistical Com. Rep.	Eilam Gross	Gregory Schott
Higgs XS Rep.	Rei Tanaka	Chiara Mariotti

In addition:

ATLAS and CMS Spokespersons and Physics Coordinators

With the participation of relevant experts as needed

Objectives

- **SM Higgs combination first**
  - **Summer 2011 EPS & EP meetings**
  - **Perhaps 0.5 to 1/fb per experiment**
- **Later, extend to the beyond the SM Higgs**

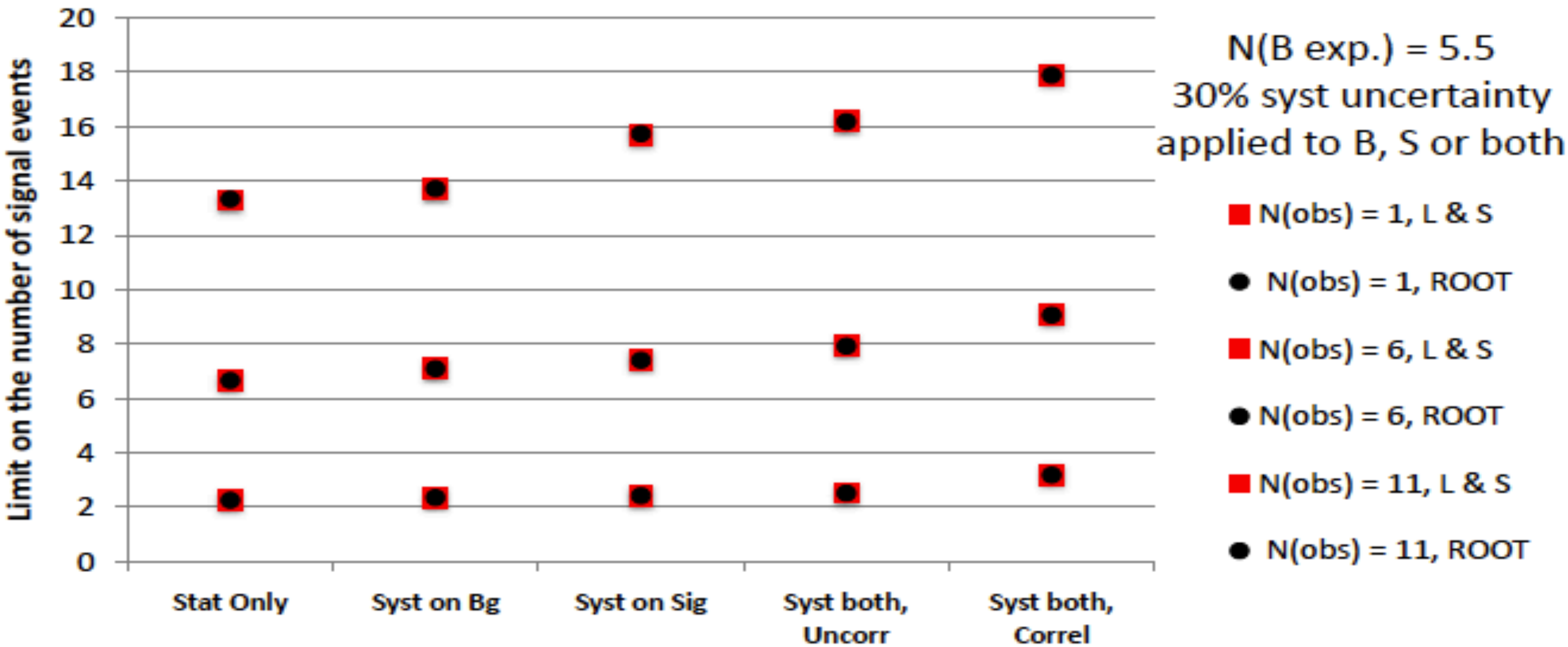
The issues to be addressed

- The common framework to do the combination
- The ATLAS+CMS combined Likelihood
- Systematic Errors
  - Correlated and un-correlated between experiments
- Nuisance parameters
- Test statistics
- Method(s) to be used for combination
- Treatment of nuisance parameters and auxiliary measurements in toy Exp.
- Discovery protocols
- ATLAS+CMS handshakes
- Documentation
- Time

The common framework

- **In the RooStat framework**
- **Validation of RooStat tools done with independent codes. Also independent checks done in MCLimit**

One Example. Profile Likelihood: Comparison of simple counting experiment with different n(obs) and systematic uncertainties



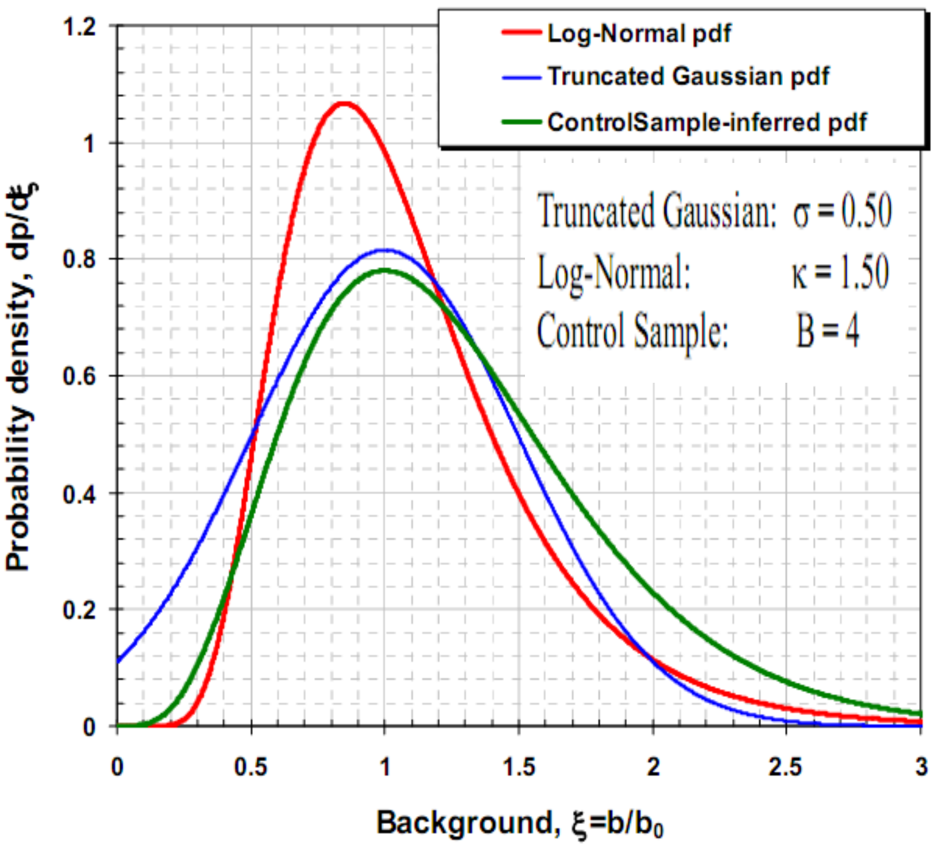
The ATLAS+CMS combined Likelihood

- From the RooStat framework
  - Build WorkSpaces with the data, a model and PDFs
  - Provide ATLAS WorkSpace, CMS WorkSpace and their combinations on both sides
  - Run various limit calculators on both and compare
  - Toy WorkSpaces have been built and exercised for H → WW + 0/1/2jets, H → gg, H → ZZ and their combinations
  - Converged on a uniform name convention

Nuisance Parameters

An uncertainty on a nuisance parameter x (e.g. background, efficiency, cross section, luminosity, etc.) can be in general described in a form of some probability density function $pdf(x)$:

- **Gamma PDF for the nuisance parameters of un-correlated systematics**
- **Log-normal PDF for the nuisance parameters of correlated systematics**



Systematic Errors

- Un-correlated systematic errors
  - e.g., MC statistics
  - Control sample measurements, …
- Correlated systematics uncertainties
  - Luminosity
  - Theoretical uncertainties on cross-sections and cross-section x Acceptance
- Uncertainty on the total cross-section σ_{tot}. This is the starting point
- To set limit on σ x BR, we are interested in the uncertainties on Acceptance (A)
- Uncertainty on the cross-section within a limited Acceptance: σA. Needed when setting limit in combining channels.

a. Systematic errors associated with PDF+α_s uncertainties

- **First, We group all processes in 4 categories based on the prevailing production source**
- **Second, we assume PDF+α_s systematic errors between all processes in one group are 100% positively correlated and not correlated between processes from different groups.**

Group	Examples of processes	Name convention
gg	gg→H, ttH, Zbb, ttbar (incl. single top), gg→VV, …	pdf_gg
qqbar	VH, V, VV, γγ	pdf_qqbar
qq	VBF H	pdf_qq
qg	γ+jets	pdf_qg

b. Systematic errors associated with QCD scale uncertainties

We assume that all physics processes have uncorrelated QCD scale uncertainties, except for very closely related processes (e.g., W and Z production; WW, WZ, and ZZ production, etc.) that we take as 100% correlated. The naming convention to be used by CMS and ATLAS in the corresponding workspaces is suggested to be as follows:

Processes	Name convention
gg→H	**QCDscale_ggH**
VH	**QCDscale_VH**
VBF H	**QCDscale_qqH**
ttH	**QCDscale_ttH**
V	**QCDscale_V**
V + heavy flavor QQ	**QCDscale_VQQ**
VV up to NLO	**QCDscale_VV**
gg→VV	**QCDscale_ggVV**
tt (incl. single top)	**QCDscale_ttbar**

4. Acceptance and extrapolation factor uncertainties

Given that the cuts are ever evolving, calculations of the acceptance and extrapolation factor uncertainties are to be performed within the ATLAS and CMS Higgs groups according to the prescriptions from the LHC Higgs cross-section group

We currently assume that the acceptance and extrapolation factor uncertainties are independent from the total cross section uncertainties discussed in section 3, except for the acceptance associated with jet counting in $H \rightarrow WW$ + 0/1/2-jets analyses. This exception is discussed in the next section.

Two data-driven techniques used by ATLAS and CMS to estimate WW and ttbar backgrounds in $H \rightarrow WW \rightarrow 2l2\nu$ + 0jet. Error dominated by QCD scale. Associated nuisance parameters:

Brief description of the extrapolation	Name convention
To be done	QCDscale_WW_EXTRAP
To be done	QCDscale_ttbar_EXTRAP

5. (Cross Section) x (Acceptance) uncertainties

- **Uncertainties on acceptance of all cuts except jet counting are treated as independent from the total cross-section.**
- **However, for gg → H → WW+0/1/2jets, the fractions of 0-, 1- and 2-jet bins are sensitive to the choice of the QCD scales. The level of sensitivity is very similar to the total cross-section uncertainties**
- **We need a recipe to evaluate the uncertainties in the jet-bin fraction (acceptance)**
  - **Some guidance has emerged from this workshop: see the proposal in Rei Tanaka's summary talk**

Uncertainty in the jet bin fraction

The overall prescription for getting numbers is as follows. First, the $gg \to H$ cross section uncertainties are taken from the LHC Higgs CS group YR [Ref]. For example, for m_H=160, the up/down uncertainties can be characterized by κ_{up}=1.109 and κ_{down}=1/1.072=0.93.

Superseded by new Proposal? – See the summary talk of Rei Tanaka

To evaluate uncertainties in the jet bin fractions (acceptance), we use HNNLO ME program with parton-level cuts taken close to the actual values used in an analysis[1]. By varying QCD scales up/down by a factor of two one obtains results summarized in the table below. The product of the total cross section κ and acceptance κ gives the overall $\sigma_{gg} \times \mathcal{A}$ for each jet bin.

Table X. Lognormal κ-factors describing uncertainties in the full gg cross section (red, taken from YR), acceptance (green, calculated using HNNLO), and CSxAcceptance (blue). The grayed out column shows variation in the total CS as obtained with HNNLO as a sanity check. The table shows numbers for the SM Higgs mass of 160 GeV.

	YR	HNNLO						
	σ_{gg}	σ_{gg}	$\mathcal{A}_0$	$\mathcal{A}_1$	$\mathcal{A}_2$	$\sigma_{gg} \times \mathcal{A}_0$	$\sigma_{gg} \times \mathcal{A}_1$	$\sigma_{gg} \times \mathcal{A}_2$
QCD up	0.93	0.90	1.05	0.94	0.70	12%	13%	27%
QCD down	1.11	1.11	0.95	1.02	1.50	16%	13%	61%

Test Statistics

- **The test statistic is the profile likelihood ratio:**

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

μ Signal strength

$\hat{\mu}$ Preferred μ in the first fit

$\hat{\nu}$ Preferred nuisance parameters in the first fit

$\hat{\hat{\nu}}$ Preferred nuisance parameters in the 2nd fit

The MINOS Technique

Method to be used for combination

- **Tentatively agreed on CL_s as the method to be used for the combination this summer**

$$1 - CL_S = 1 - \frac{(1 - CL_{SB})}{1 - CL_B}$$

where the p-value calculated on the BackToys distribution, $1 - CL_B$, is used as normalization factor of the p-value $1 - CL_{SB}$ calculated on the SigBackToys distribution. This definition yields more conservative results, but has the advantage of never excluding zero times the Standard Model (i.e., the background only hypothesis) at 2σ. Moreover, as this is the approach used at the Tevatron experiments (see for instance Reference [37]), it can be used for comparison purposes. This approach has been used already

Treatment of nuisance parameters and auxiliary measurements in toy experiments

- **Sampling the test statistic by:**
  - Toy MC randomizing nuisance parameters according PDF
  - Toy MC with nuisance parameters fixed at their nominal values
  - Using asymptotic approximation

These options and a proposal on how to proceed are being discussed...

Higgs mass m_H grid

(motivation for the choice of steps driven $\gamma\gamma$, ZZ mass resolutions)

mass range	step	number of points
110-140	0.5	61
140-160	1	20
160-260	2	50
260-290	2	15
290-350	5	12
350-400	10	5
400-500	20	5
550, 600	20	5
Total number of points		**173**

Full tables for Higgs XS and BR (with interpolation for missing points with spline fit) are now being prepared.

ATLAS+CMS Handshakes

- - ATLAS "pseudo-data" Workspaces for different channels:
  - H → WW (→llνν) + 0/1/2j
  - H → γγ
  - H → ZZ → 4l
- - CMS "pseudo-data" Workspaces for different channels:
  - H → WW (→llνν) + 0/1/2j
  - H → γγ
  - H → ZZ → 4l
- These two WorkSpaces are allowed to (and do) give different results.
- The combined ATLAS+CMS "pseudo-data" WorkSpaces
  - Combined independent on both sides and cross-checked
- For this exercise: Channels: H->WW->2l2v, H->γγ,
  H->ZZ->4l (CMS only)
  - 5800 un-binned events in 37 channels
  - 98 nuisance parameters: 12 common ones, 31 ATLAS-specific ones, 55 CMS-specific ones. 10 nuisances were correlated: luminosity, 3 x pdf (gg,qq, qqbar), 6 x scales (ggH, qqH, ggVV, qqVV, V, ttbar).

ATLAS+CMS Handshakes

PL = Profile Likelihood

LEP-style CLs ($Q = -2\log\frac{L(\mu=1)}{L(\mu=0)}$, no profiling)

Task	Done by ATLAS	Done by CMS
CLs @μ=1 for ATLAS Workspaces (WW)	**0.096±0.002**	**0.104±0.002**
CLs @μ=1 for CMS Workspaces (WW)	**0.0012±0.0005**	**0.0009±0.0003**
Combine 2 workspaces and calculate PL limit, **H → WW / H→$\gamma\gamma$**	**0.5501 (WW)** **4.7651 ($\gamma\gamma$)**	**0.5501 (WW)** **4.7653 ($\gamma\gamma$)**
Combine 2 Workspaces and calculate PL asymptotic limit (H → WW+$\gamma\gamma$)	**0.2724**	**0.2724**
Combine 2 Workspaces and calculate LEP-like CLs limit (H → WW)	**0.519±0.003**	**0.508±0.003**
Combine 2 Workspaces and calculate LEP-like CLs limit (H → WW+$\gamma\gamma$+ZZ)		**0.626±0.004**

Based on 1/fb pseudo data

Agreement within 1-2%

Documentation

- **To summarize what we have done and agreed on in preparation for the ATLAS+CMS combination this summer**
- **When appropriate, some pieces of this document may be recycled in one or another form in the ultimate publication with the actual combination of data results**

LHC Higgs Combination Group report

A. Armbruster[1], K. A. Assamagan[2], M. Chen[3], K. Cranmer[4], E. Gross[5], A. Korytov[3], C. Mariotti[6], W. J. Murray[7], G. Petrucciani[8], J. Qian[1], G. Schott[8], V. Sharma[8], R. Tanaka[10], F. Tarrade[2], and H. Wang[11]

[1]University of Michigan, Ann Arbor, MI, USA
[2]BNL, Upton, NY, USA
[3]University of Florida, Gainesville, FL, USA
[4]New York University, New York, NY, USA
[5]Weiseman Institute, Rehovot, Israel
[6]University of Torino, Torino, Italy
[7]Rutherford Appleton Laboratory, Didcot, UK
[8]Universtiy of California, San Diego, CA, USA
[9]Universitaet Karlsruhe, Karlsruhe, Germany
[10]LAL, Orsay, France
[11]University of Wisconsin, Madison, WI, USA

Conclusions

- **Efforts have started on LHC Higgs combination**
  - **For limit setting and discovery**
- **A lot of progress so far**
  - **Combined likelihood**
  - **Test statistic**
  - **Method for limit setting**
  - **PDF for nuisance parameter**
  - **Higgs mass grid**
- **Toy combination show excellent agreement between ATLAS and CMS on their combined models**
- **But a lot to do still …**

➔

To be addressed still

- Recipe for correlated theoretical systematic uncertainties
  - Expect some guidance/convergence at this workshop
- Dealing with mass points for which we have no simulation
- Treatment of nuisance parameters and auxiliary measurements in toy experiments
- CL_s with profiling of systematic errors and study the CPU consumption
- Look-elsewhere effect
  - the combination will be probing O(200) Higgs mass points with some non-trivial correlations between them
- Format of presenting results
  - Limits on the overall signal strength modifier μ (both observed and expected), including bands. mass range excluded (both observed and expected) …
  - Excess: local p-value; p-value taking into account the look-elsewhere effect
- Documentation
  - Complete report to ATLAS and CMS managements

Jianming's proposal for jet-bin uncertainties

1) calculate cross sections in exclusive bins as
sigma_0 = f0 * sigma_tot
sigma_1 = f1 * sigma_tot
sigma_2 = f2 * sigma_tot

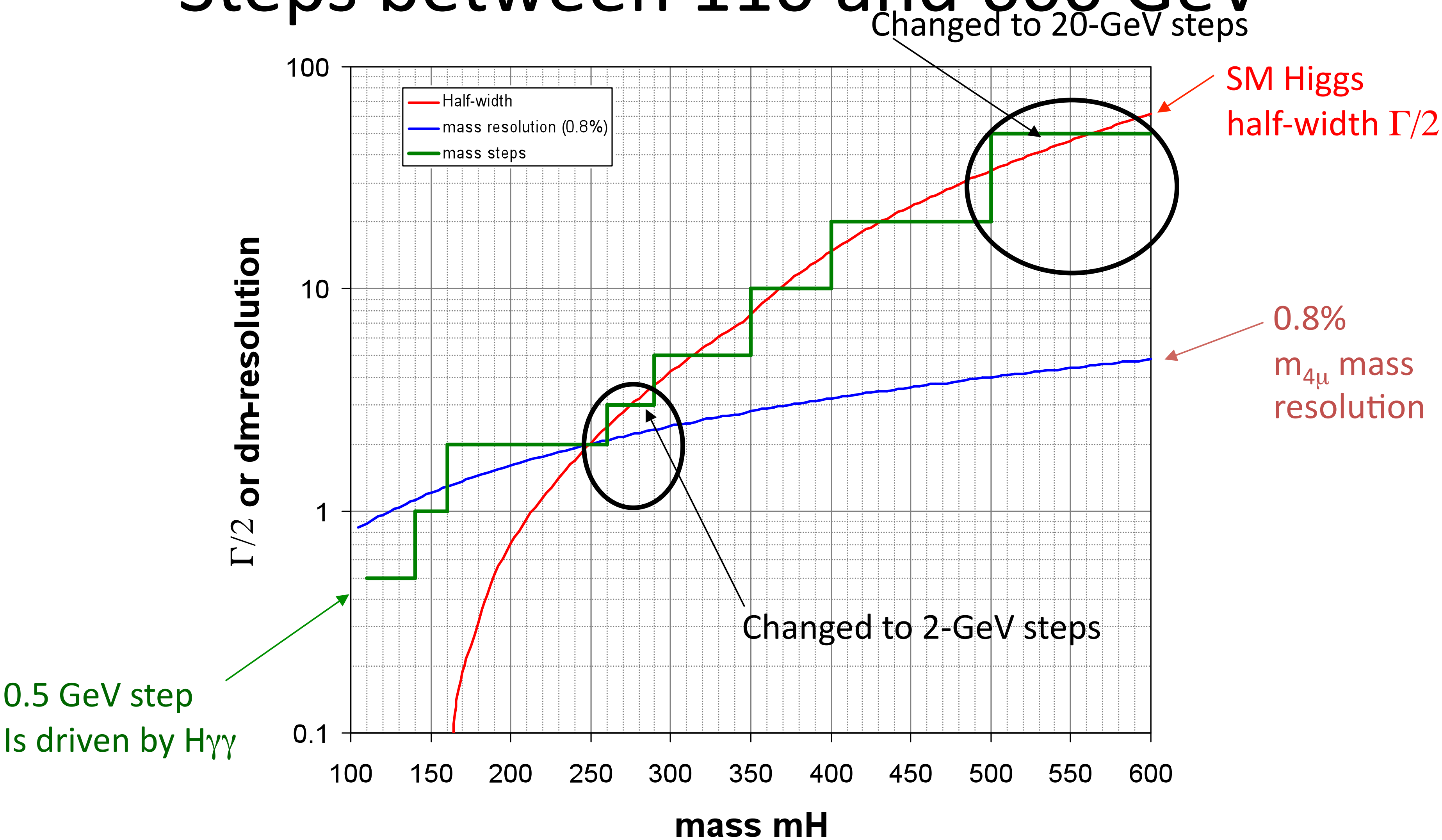
2) calculate error using standard error propagation procedure assuming fi and sigma_tot are independent:
d(sigma_0) = sqrt[d(f0)^2*sigma_tot^2 + f0^2*d(sigma_tot)^2]
d(sigma_1) = sqrt[d(f1)^2*sigma_tot^2 + f1^2*d(sigma_tot)^2]
d(sigma_2) = sqrt[d(f2)^2*sigma_tot^2 + f2^2*d(sigma_tot)^2]

This will guarantee that
d(sigma_i)/sigma_i > d(sigma_tot)/sigma_tot
a point emphasized by many theorists.

3) In combining Higgs search results from these three bins, the full correlation matrix among fi is taken into account. For mH=160 GeV:
1.00 -0.95 -0.98
-0.95 1.00 0.88
-0.98 0.88 1.00

BACKUP

Steps between 110 and 600 GeV



On modeling of uncertainties

- An uncertainty on a nuisance parameter x (e.g. background, efficiency, cross section, luminosity, etc.) can be in general described in a form of some probability density function $pdf\ (x)$:

- Truncated Gaussian
  - in general, not recommended within CMS (here shown for completeness)

- Log-normal
  - commonly used alternative

- Gamma distribution
  - recently added, intended for describing uncertainties associated with limited statistics in control samples, Monte Carlo

- …

Truncated Gaussian

- The Gaussian distribution with the unphysical tail below x=0 simply chopped off by hand :

$$\frac{dp}{dx} = \begin{cases} \frac{C}{\sqrt{2\pi}\sigma} \exp(-\frac{(x-\tilde{x})^2}{2\sigma^2}), & \text{for } x > 0 \\ 0, & \text{for } x \leq 0 \end{cases} \text{, where } C = \frac{2}{1 + \mathrm{erf}(\frac{1}{\sqrt{2}\sigma})}$$

For correlated errors with different sigma's, the largest sigma should be used to define truncation. The narrower correlated Gaussians get truncated before reaching x=0. This looks very unphysical.

Finite value of $pdf(\varepsilon)$ at ε=0 for signal acceptance is pathological in calculation of Bayesian limits. Ad hoc truncation above zero is too arbitrary…

Log-normal

- The normal distribution for $\ln(x)$:

$$\frac{dp}{dx} = \frac{1}{\sqrt{2\pi}\ln\kappa}\,\frac{1}{x}\,\exp\left(-\frac{(\ln(x/\tilde{x}))^2}{2(\ln(\kappa))^2}\right)$$

- where x-*tilde* is the best estimate on the nuisance parameter, κ is the factor error on x

- No tail toward negative values of x and the probability density at x=0 is always zero

- When the overall uncertainty in x arises from uncertainties in multiple multiplicative factors (like various efficiencies), the central limit theorem implies that it is the $\ln x$ distribution that would tend to become Gaussian.

Log-normal (cont'd)

•When uncertainties become very large (e.g., when we say "a factor of two uncertainty"), they map very naturally onto the log-normal *pdf*, while the Gaussian distribution obviously becomes completely inappropriate.

•For small errors $\kappa \sim 1 = 1+\sigma$, the log-normal distribution is basically a Gaussian with the mean *x-tilde* and relative error σ

•Larger tail than a Gaussian. Consequently, a significance of an event excess will be more conservative when one uses a log-normal *pdf* for the background uncertainties

Gamma-function

•When estimating background in the signal region as $x = \rho B$, where B is an event count taken from a control sample and small (or when run into a problem of limited MC statistics with very few (B) MC events passing cuts), then the natural choice is the Gamma-function :

$$\frac{dp}{dx} = \frac{1}{\rho} \frac{(x/\rho)^B \exp(-x/\rho)}{\Gamma(B+1)}$$

•Most probable value: $\rho \cdot B$

•Mean value: $\rho \cdot (B+1)$

•Dispersion: $\rho \cdot \mathrm{sqrt}(B+1)$

Gamma-function (cont'd)

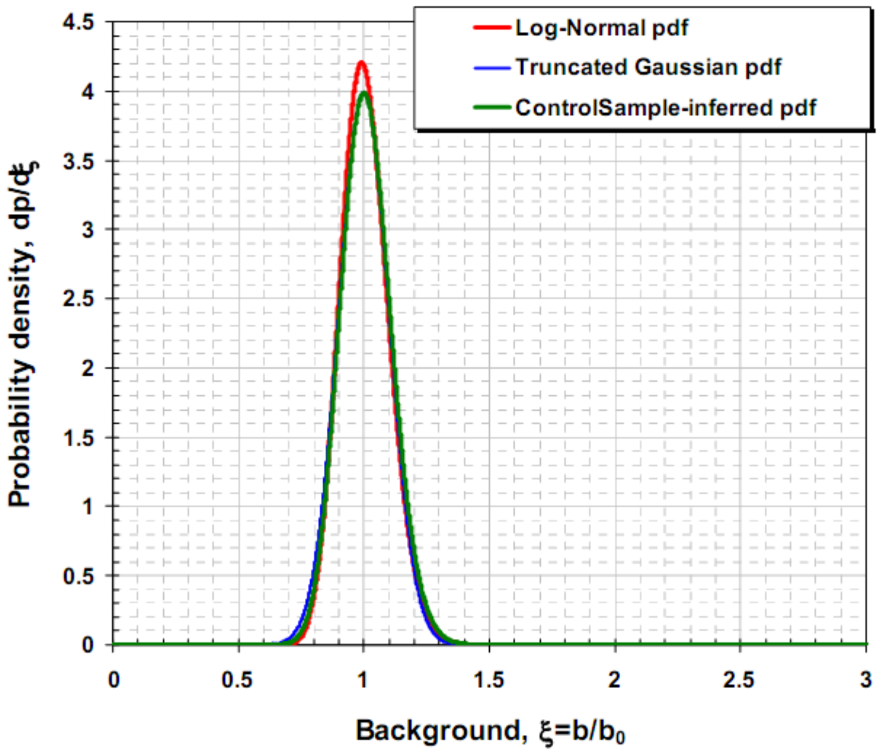
- Similarly to log-normal, $pdf(x)=0$ for $x=0$

- Similarly to log-normal, it has a longer tail toward larger values in comparison to the Gaussian

- For large B,
- it becomes similar to Gaussian with $\sigma = 1/\sqrt{B}$
-
- $B = 0$ is a perfectly allowable situation,
- giving $pdf(x) = (1/\rho)\exp(-x/\rho)$

Comparison of *pdf* ’s

Truncated Gaussian: $\sigma = 0.10$
Log-Normal: $\kappa = 1.10$
Control Sample: $B = 100$

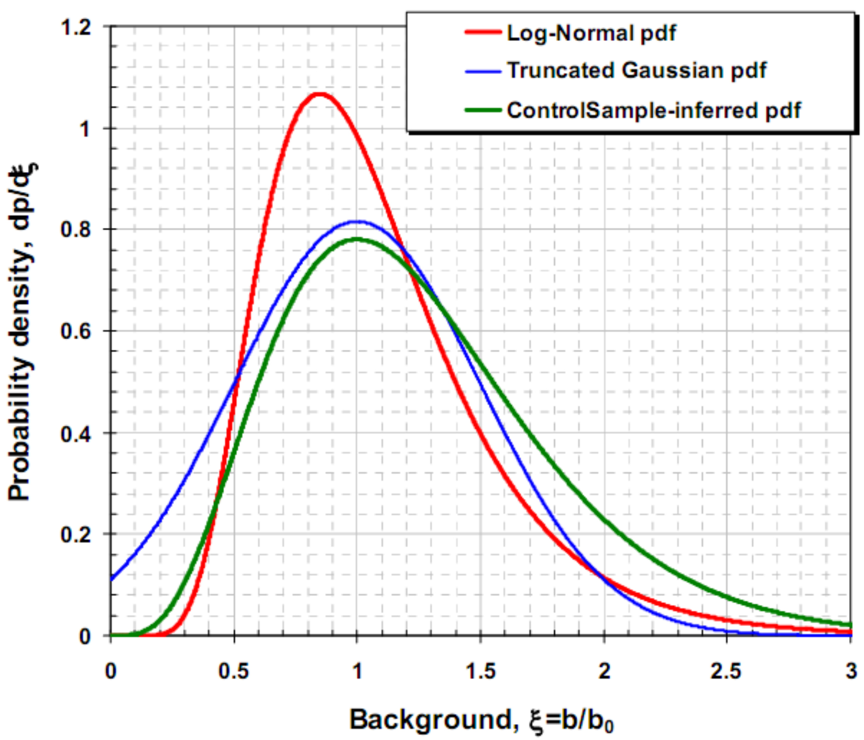


- With expected bkg = 1, observe 5 event, significance calculated from p-value :

Truncated Gaussian	Log-normal	Gamma
2.658± 0.005	**2.652± 0.005**	**2.647± 0.005**

Comparison of *pdf*'s

Truncated Gaussian: $\sigma = 0.50$
Log-Normal: $\kappa = 1.50$
Control Sample: $B = 4$



With expected bkg = 4, observe 4 event, 95% CL upper limit on signal yield (Bayesian with flat prior) :

Truncated Gaussian	Log-normal	Gamma
6.78	6.35	6.33

- With expected bkg = 1, observe 7 event, significance calculated from p-value :

Truncated Gaussian	Log-normal	Gamma
3.25 ± 0.01	3.10 ± 0.01	2.89 ± 0.01

What about B=0 ?

- If we have only 0 event in control sample, the only choice is Gamma *pdf*

- For example, for ρ = 0.1:

  - 95% C.L. upper limit on signal yield, when we observe 1 event (Bayesian with flat prior): 4.66

  - Significance when observe 5 events (from p-value): 4.4

100%-Correlated in log-normal

- The $pdf(b|\tilde{b},\kappa)$ can be effectively emulated in pseudo-experiments by generating a random x according to the normal (Gaussian) *pdf*, $g(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$, and then taking

$$b = \tilde{b} \cdot e^{x \cdot \ln \kappa}$$

- For 100%-correlated uncertainties: one should use one random number $g(x)$ for modifying all correlated uncertainties in a given pseudo-experiment

100%-Correlated in Gamma

- e.g. HZZ analysis, one would use Z→$\mu\mu$ yield (B_0) in data to normalize ZZ→4μ ($b_1=\rho_1 B_0$) and ZZ→2e2μ ($b_2=\rho_2 B_0$)

- In a toy experiment, one generates one B according to its gamma distribution , and uses the B for both ZZ→4μ and ZZ→4e

Dealing with mass points for which we have no simulation

Three classes of analyses:

1. Cut-and-count
2. 1d-shape analysis after cuts
3. MVA-based analyses

Cut-and-count

Cuts are required to change smoothly with m_H: cut(m_H)

Then:

- One can easily get **expected bkgd** event yields for any given m_H mass
- Similarly, we get the **observed event counts** for any m_H mass points
- Since cuts change smoothly, one can expect that the **signal efficiency** for any given m_H mass point can be simply interpolated between nearby simulated mass points
- Signal **CS x BR** are interpolated linearly between tabulated mass points from the Higgs CS Yellow Report

1d shape analysis ($m_{\gamma\gamma}$, m_T, m_{4l}, etc.)

Cuts are required to change smoothly with m_H: cut(m_H)

Then:

- One can easily get **expected bkgd(m)**
- Similarly, we get the **observed events** (binned or unbinned in m)
- Signal **$pdf_S(m|m_H)$** can be obtained by horizontal morphing of signal pdf's obtained for simulated points (see next slide)
- Signal efficiency **$\varepsilon_S(m_H)$** can be obtained by interpolating between efficiencies obtained for simulated points
- Signal **CS x BR** are interpolated linearly between tabulated mass points from the Higgs CS Yellow Report

NOTE:

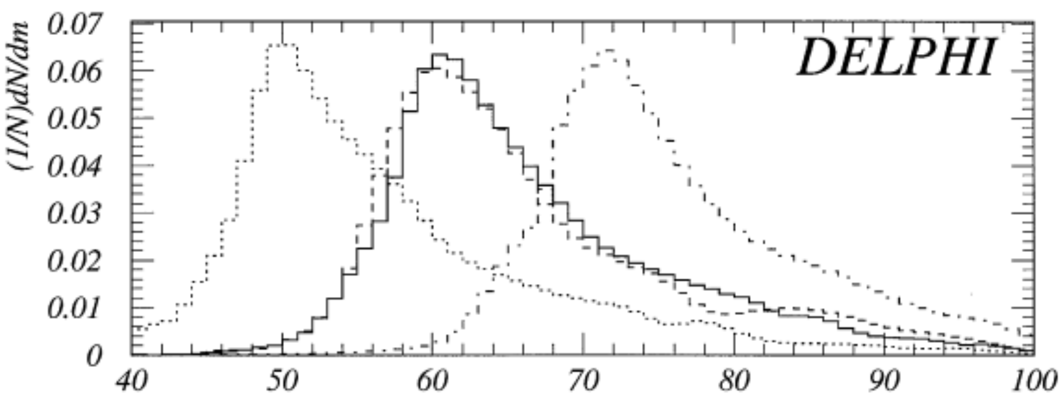
The last three steps can be done in one go, if one has a parametric form for **CS x BR x $\varepsilon_S(m_H)$ x $pdf_S(m|m_H)$.** *One can simply interpolate between parameters obtained for simulated/tabulated mass points.*

Horizontal morphing

- Used at LEP for Higgs templates.

  Nucl. Instrum. Meth. A425 (1999) 357-360

  http://inspirebeta.net/record/501018/



- Main limitations:
  - Doesn't work well for histograms with few bins
  - Works only for one morphing parameter

MVA-based analyses

- MVA trained for Higgs mass m_A:
  - MVA function $y=f_A(x)$
  - signal output distribution $h_A(y)$

- MVA trained for Higgs mass m_B:
  - MVA function $y=f_B(x)$
  - signal output distribution $h_B(y)$

- for any mass in between: $m = \alpha m_A + \beta m_B$
  - MVA function: $f(x) = \alpha f_A(x) + \beta f_B(x)$
  - expected signal MVA output: $h(y) = \alpha h_A(y) + \beta h_B(y)$
  - background output is derived using f(x) and can be checked that it matches the linear interpolation

NOTE:

we do not expect that interpolated MVA must be identical to what one would get by actually training MVA on the intermediate mass point. However, their sensitivities are expected to be nearly identical, if steps are not too crude.