ATLAS Discovery Experience

Aaron Armbruster

University of Michigan

On behalf of ATLAS

February 13, 2013

Outline

- Combination Procedure
- Asymptotic Limit Bands
- Low count $\hat{\mu}$ / Physical PDF
- P-Value Uncapping
- Asymmetric Uncertainty Handling

Combination Procedure

- Individual models (likelihoods, datasets, etc) provided by subgroups electronically via RooFit workspaces
- Start with the individual likelihoods $\mathcal{L}_i(\mu, \theta_i) = \mathcal{L}_i^0(\mu, \theta_i) \times \prod_j^{M_i} \mathcal{A}(\theta_i^j)$
  - μ is parameter of interest
  - θ_i are the set of nuisance parameters used in channel i
  - $\mathcal{L}_i^0$ is the main body of the likelihood (eg observable distribution)
  - $\mathcal{A}(\theta_i^j)$ are auxiliary constraints for each θ_i^j (eg unit gaussian)
- Build a combined likelihood $\mathcal{L}(\mu, \theta) = (\prod_i^N \mathcal{L}_i^0(\mu, \theta_i)) \times (\prod_j^M \mathcal{A}(\theta^j))$
  - θ is now the set of all $unique$ nuisance parameters
  - Some θ_i^j are shared between channels. This must be recognized to ensure proper correlation.

Combination Details

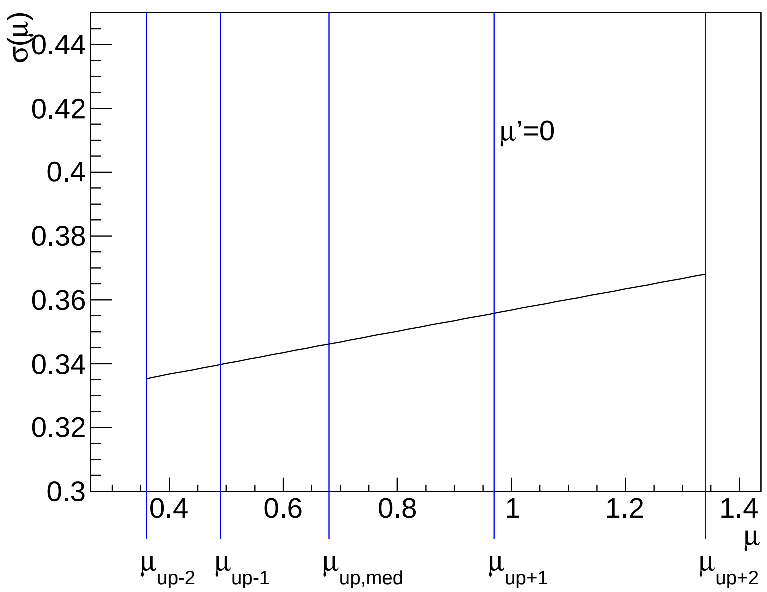
- At first (one or two combinations), ATLAS results were fully based on toys
- As model grew, these became impractical
  - $\sim$570 nuisance parameters at time of discovery
  - $\sim$310 of these are due to MC stats, treated Barlow-Beeston style
- $\sim$10-30 minutes per fit$\rightarrow$20-60 minutes per toy
  - $\mathcal{O}$(millions) CPU hours to produce full result
- Large model gives many fit failures, leads to false toys in tails of distribution
  - Diagnostic tools used to ensure tail events were truly failures
  - p-values become falsely enhanced, gives conservative result
- This led us to use asymptotics for all results, validated with toy and Bayesian tests

Asymptotic Bands

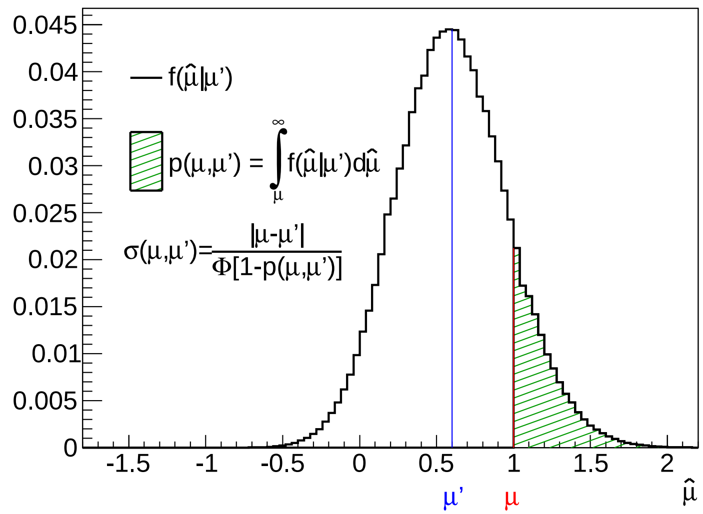
Asymptotic Limit Bands

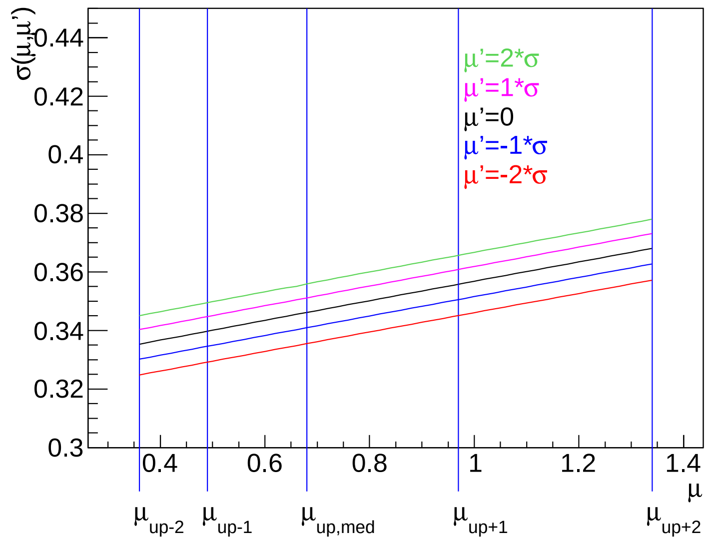
- When σ is μ-dependent, simple equation for limit bands fails

$$
\begin{aligned}
\text{CLs}(\mu_{\text{up+N}}) \equiv \alpha &\approx (1-\Phi(\frac{\mu_{up+N}-N\sigma_1}{\sigma_2}))/(\Phi(\frac{\mu_{up+N}}{\sigma_3}-\frac{\mu_{up+N}-N\sigma_1}{\sigma_2})) \\
&\neq (1-\Phi(\frac{\mu_{up+N}-N\sigma_1}{\sigma_2}))/(\Phi(N)) \\
\Rightarrow \mu_{up+N} &\neq \sigma\{\Phi^{-1}[1-\alpha\Phi(N)]+N\}
\end{aligned}
\quad (1)
$$



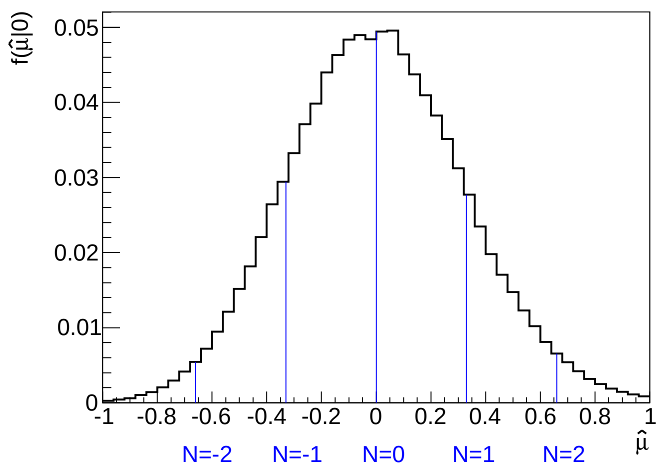
Estimating σ

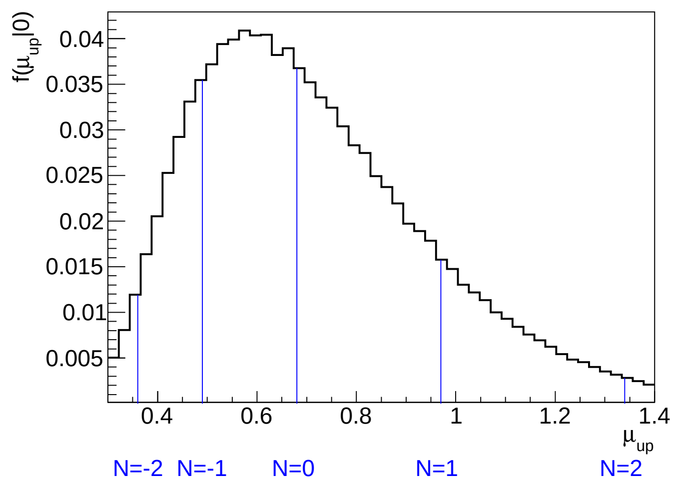




- $\sigma \approx \frac{|\mu - \mu'|}{\sqrt{t_{\mu,A_{\mu'}}}}$ really involves two μ's: μ and μ'
- To conceptualize, imagine the toy distribution $f(\hat{\mu}|\mu')$ from which σ can be extracted
  - μ' is the hypothesized value $\sim$ median of the distribution
  - μ is the tested value from which you integrate to extract σ

Mapping





- In result using toys, μ_{up} is found for each b-only toy by scanning test statistic until is crosses some calibrated threshold
  - Band is derived from the quantiles of the distribution of μ_{up}
- If asymptotic properties hold, there should be a one-to-one mapping between the quantiles of $f(\hat{\mu}|0)$ and $f(\mu_{up}|0)$
  - The N'th quantile of $f(\hat{\mu}|0)$, call it μ'_N, therefore characterizes the Asimov dataset for the N'th quantile expected limit

Using the N'th Quantile Asimov

- Construct five Asimov datasets, one for each N=-2,...,2
  - μ'_N found by scanning $\mu' = 0$ Asimov NLL to find $-2\log\lambda(\mu'_N) = N^2$
  - Nuisance parameters $\hat{\theta}(\mu'_N)$ also taken from $\mu' = 0$ Asimov fit when constructing μ'_N Asimov dataset
  - These Asimov datasets characterize quantiles of $f(\hat{\mu}|0)$ that we wish to map to $f(\mu_{up+N}|0)$
- Use each Asimov dataset to find the crossing $\mathrm{CLs}(\mu_{\mathrm{up+N}}) = \alpha$
  - Implicit in the $\mathrm{CLs}(\mu_{\mathrm{up+N}}) = \alpha$ is the mapping function to go from $f(\hat{\mu}|0)$ to $f(\mu_{up}|0)$

Comparison

Comparison of expected upper limit band

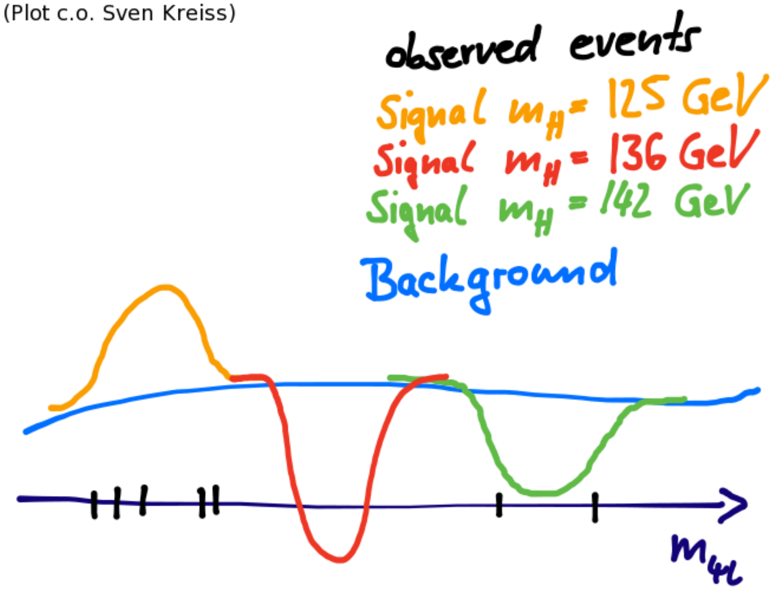
	Low systematics			High systematics		
Quantile	Old μ_{up}	Improved μ_{up}	Toy μ_{up}	Old μ_{up}	New μ_{up}	Toy μ_{up}
+2	1.27	1.34	1.32	1.27	2.78	2.82
+1	0.95	0.97	0.96	0.94	1.25	1.25
0	0.68	0.68	0.68	0.68	0.68	0.68
-1	0.49	0.48	0.48	0.49	0.42	0.42
-2	0.37	0.36	0.36	0.36	0.29	0.28

- New procedure reproduces toy results well
- Differences are especially striking when systematics are large
  - μ-dependence of σ is large in these cases, which is exactly the scenario the new method was designed to address

Low count $\hat{\mu}$ / Physical PDF

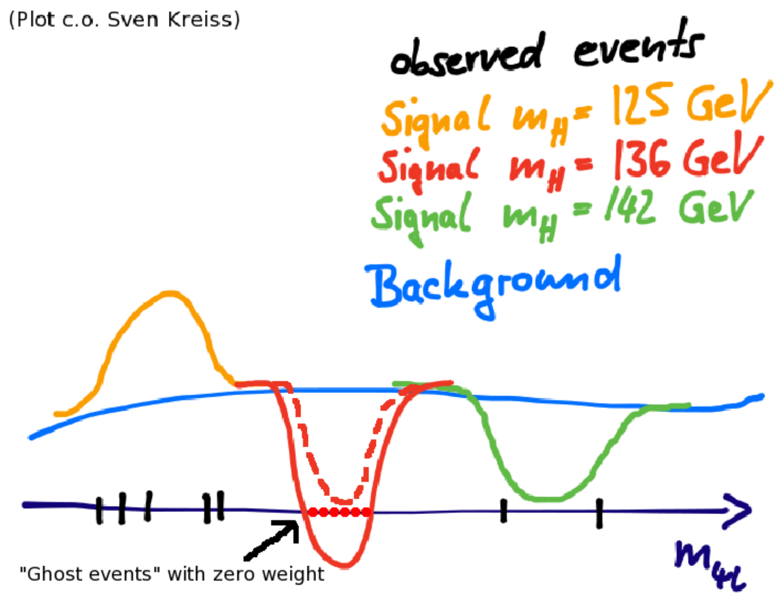
Low count $\hat{\mu}$ / Physical PDF

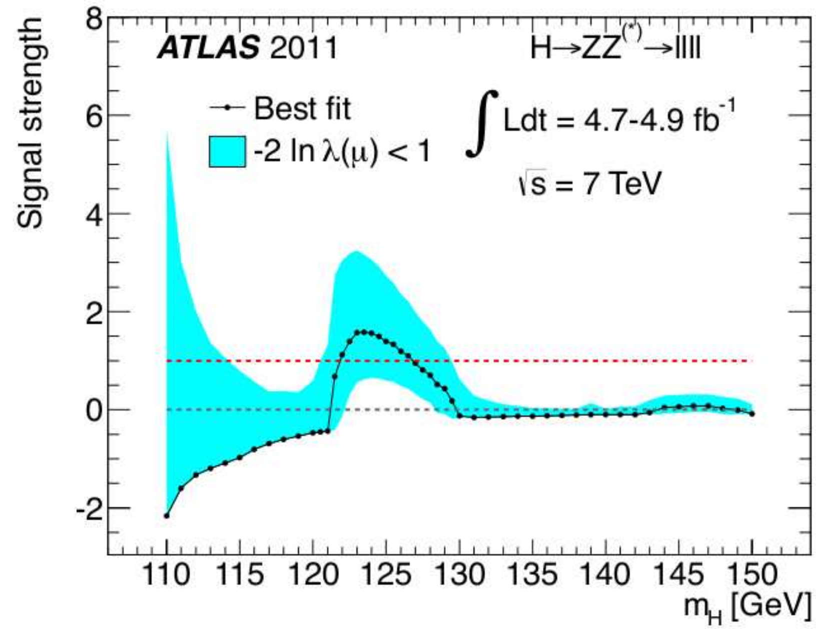
(Plot c.o. Sven Kreiss)



- In resonance models with low event counts like $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$, PDF can become negative when fitting signal with no observed events
- Issue is mostly technical
  - Likelihood only evaluated on data points
  - Difficult to check PDF is physical everywhere in a general way

Low count $\hat{\mu}$ / Physical PDF

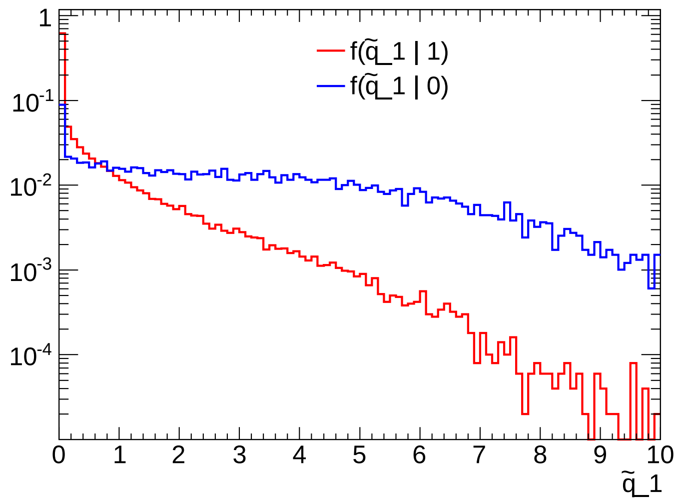


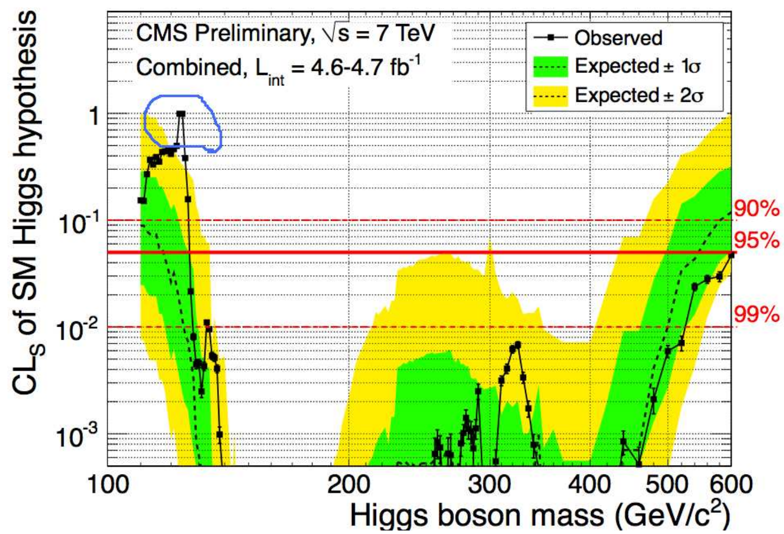


- "Ghost events" with zero weight can be added around signal peak to force RooFit to check if PDF is negative
- This causes the familiar wall in the $\hat{\mu}$ plot on the right
  - $\hat{\mu}_{\text{min}} \approx -\frac{B(x_{\text{peak}})}{S(x_{\text{peak}})}$

P-Value Uncapping

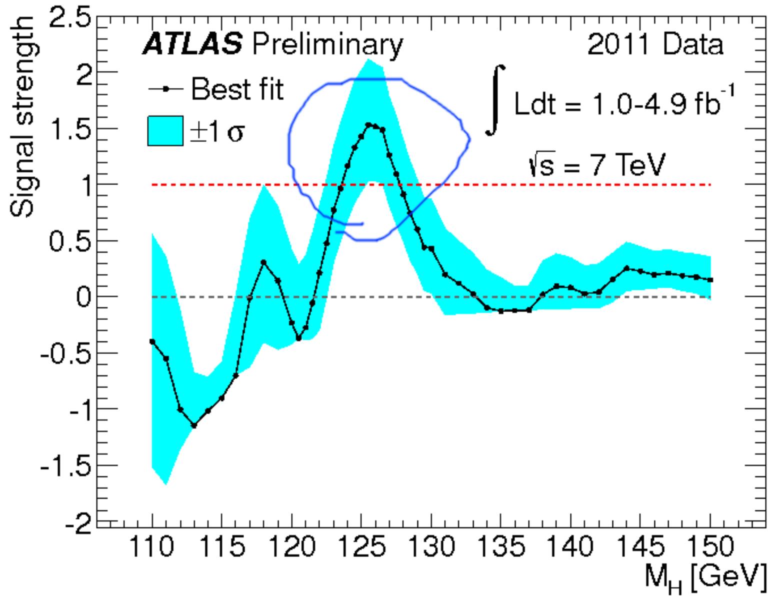
P-Value Uncapping

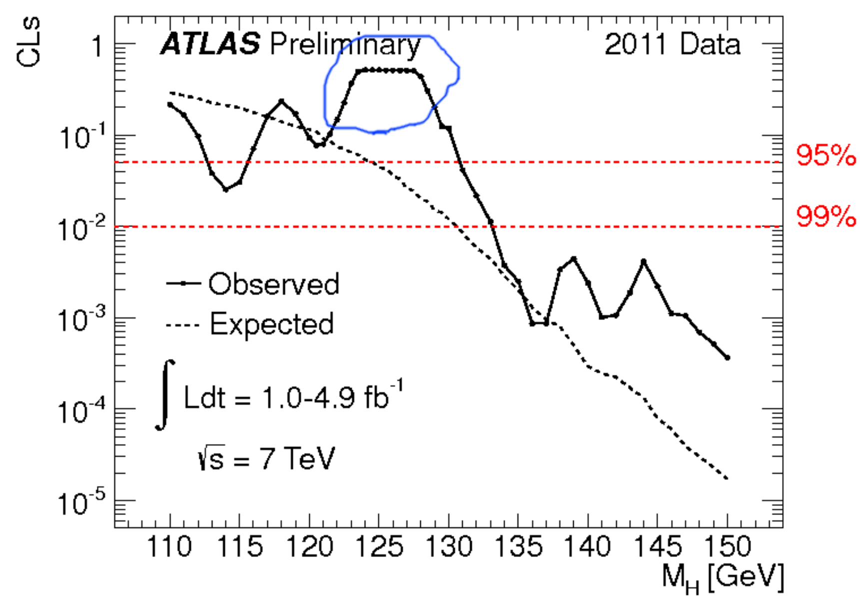




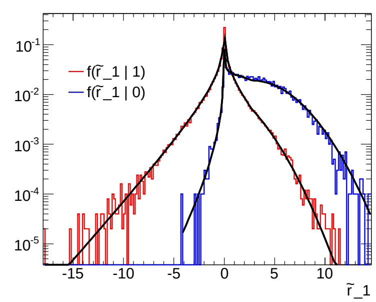
- Classically, CLs→1 continuously as Q$\to -\infty$
- Condition $\tilde{q}_\mu = 0$ for $\hat{\mu} > \mu$ leads to $\delta(\tilde{q}_\mu)$ in $f(\tilde{q}_\mu|\mu')$ that breaks this
  - Similar issue for q_0 when $\hat{\mu} < 0$
- $\tilde{q}_\mu$, q_0 can be redefined to reveal structure of excesses and deficits

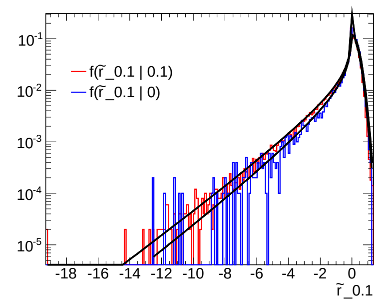
Example


Signal strength
ATLAS Preliminary
2011 Data
Best fit
±1 σ
∫ Ldt = 1.0-4.9 fb-1
√s = 7 TeV
M_H [GeV]


CLs
ATLAS Preliminary
2011 Data
95%
99%
Observed
Expected
∫ Ldt = 1.0-4.9 fb-1
√s = 7 TeV
M_H [GeV]

Redefining $\tilde{q}_\mu$





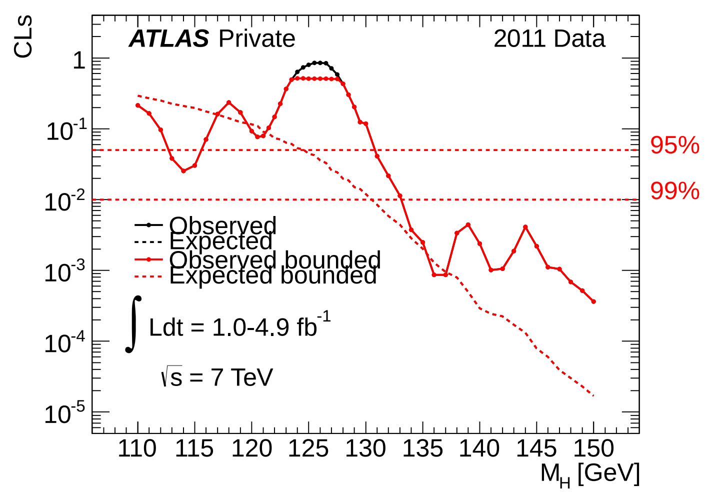
- We can gain back information lost in $\delta(\tilde{q}_\mu)$ by defining a new test stat $\tilde{r}_\mu$

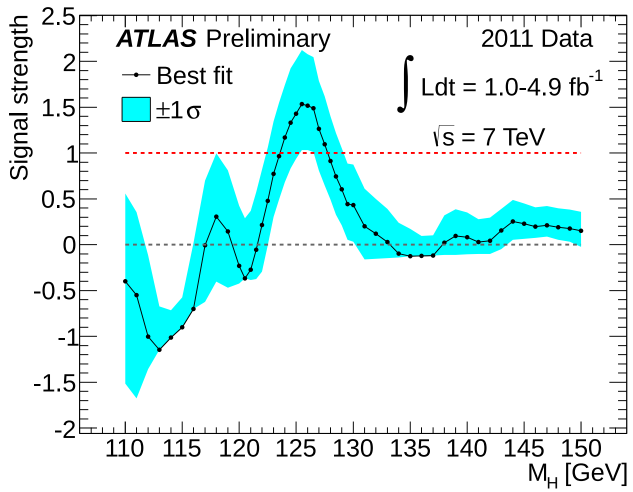
$$\tilde{r}_\mu = \begin{cases} +(-2\ln\lambda(0)) & \hat{\mu} \leq 0 & (same) \\ +(-2\ln\lambda(\mu)) & 0 < \hat{\mu} < \mu & (same) \\ -(-2\ln\lambda(\mu)) & \hat{\mu} \geq \mu & (new!) \end{cases}$$

$$F(\tilde{r}_\mu|\mu') = \begin{cases} \Phi(\frac{\tilde{r}_\mu - (\mu^2 - 2\mu\mu')/\sigma^2}{2\mu/\sigma}) & \frac{\mu^2}{\sigma^2} \leq \tilde{r}_\mu & (same) \\ \Phi(\sqrt{\tilde{r}_\mu} + \frac{\mu-\mu'}{\sigma}) & 0 < \tilde{r}_\mu < \frac{\mu^2}{\sigma^2} & (same) \\ \Phi(-\sqrt{-\tilde{r}_\mu} + \frac{\mu-\mu'}{\sigma}) & \tilde{r}_\mu \leq 0 & (new!) \end{cases}$$

- Identical for expected p-values and observed cases of $\hat{\mu} < \mu$

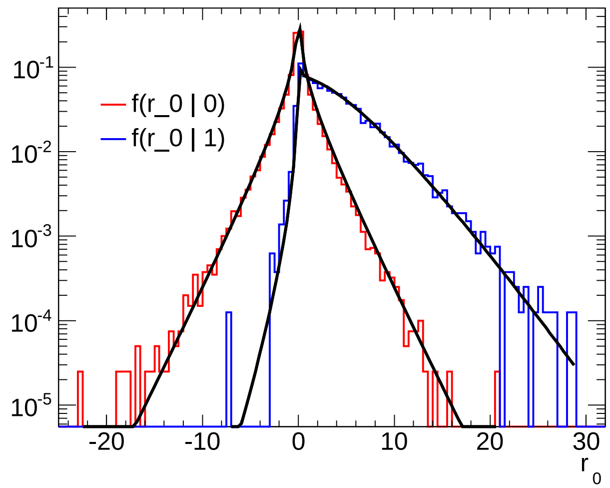
Comparison

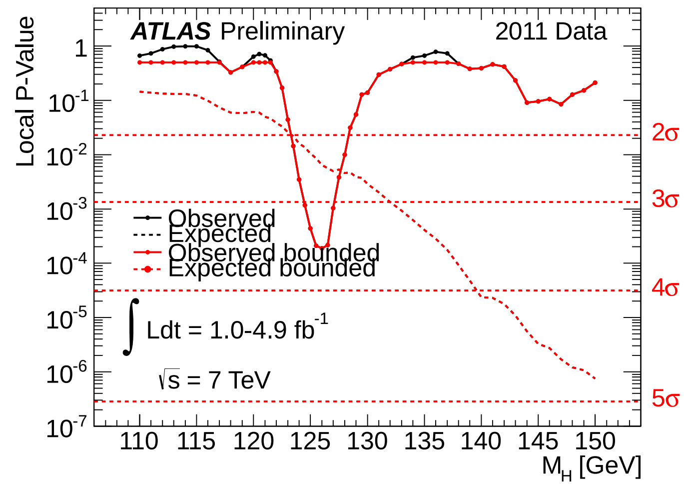




- As expected, behavior is identical for $\hat{\mu} < 1$
- For $\hat{\mu} > 1$, the new CDF kicks in and the CLs=0.5 cap is broken

Redefining q_0





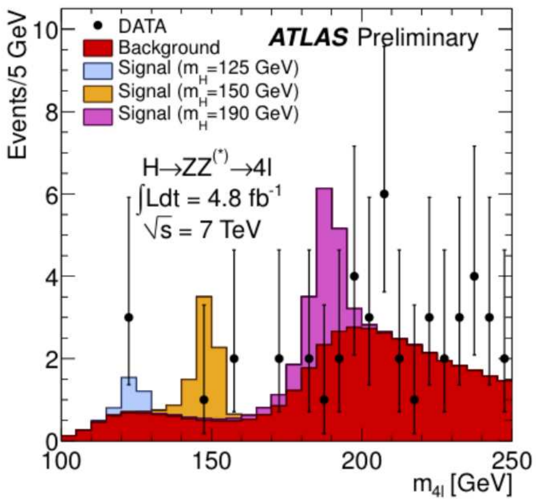
- r_0 can be defined similar to $\tilde{r}_\mu$

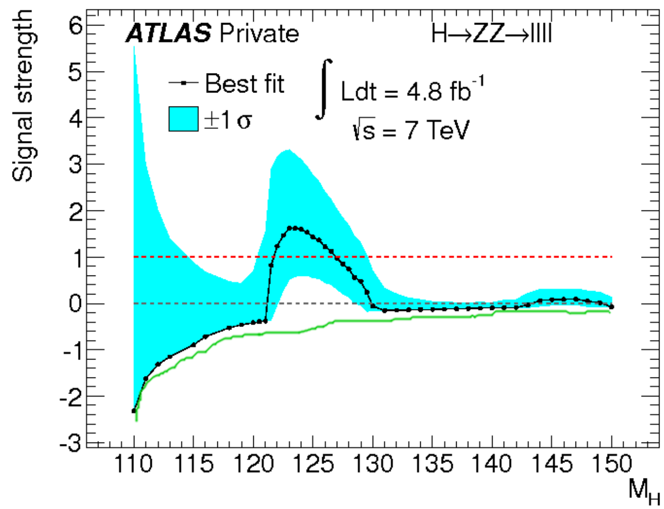
$$r_0 = \begin{cases} +(-2\ln\lambda(\mu)) & \hat{\mu} > 0 \quad (same) \\ -(-2\ln\lambda(\mu)) & \hat{\mu} \leq 0 \quad (new!) \end{cases}$$

$$F(r_0|\mu') = \begin{cases} \Phi(\sqrt{r_0} - \frac{\mu'}{\sigma}) & r_0 > 0 \quad (same) \\ \Phi(-\sqrt{-r_0} - \frac{\mu'}{\sigma}) & r_0 \leq 0 \quad (new!) \end{cases}$$

- Identical for expected p-values and observed cases of $\hat{\mu} > 0$

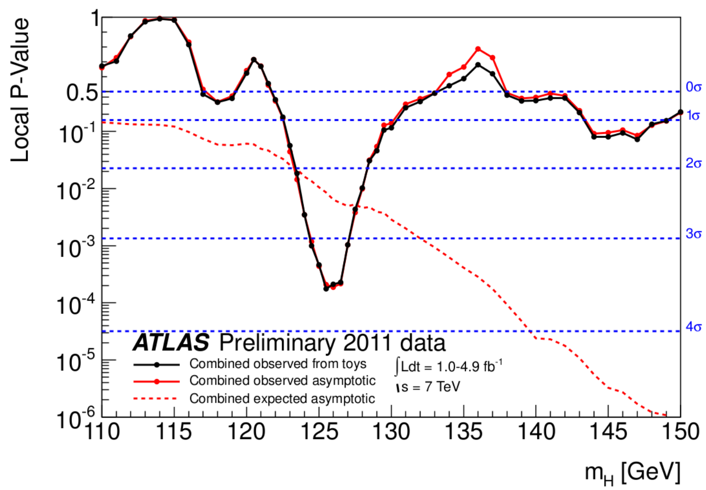
Physical PDF issues

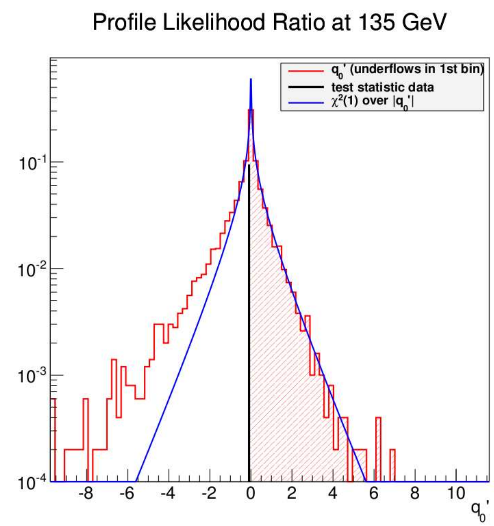




- Ghost events come back to haunt us!
- To deal with the physical PDF issue, ghost events would have to be added to each toy in the ensemble
  - This will break the asymptoticity of r_0 toys if not dealt with

Practical example





- p_0 with toys for ATLAS combination courtesy of Tim Adye
  - Negative r_0 toy distribution deviates from asymptotics
  - $\hat{\mu}$ goes more negative than it should, leading to larger r_0 values
- Physical pdf isn't explicitely required within RooFit toys
- NB: Asymptotics can be taught about physical pdf
  - Difficult to do in a model independent way for toys

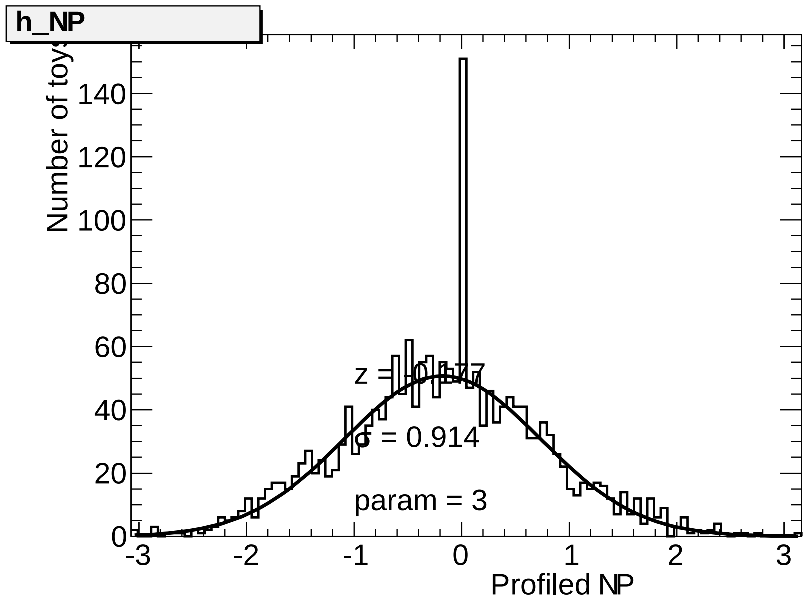
Asymmetric Uncertainties / NP Distributions

Asymmetric Uncertainties

- Systematics are treated as lognormally distributed
  - For a normally distributed nuisance parameter θ, $\nu(\theta) \equiv \kappa^\theta$ is lognormally distributed
  - $\nu(\theta)$ multiplies our expected rates: $B(\theta) = B_0\nu(\theta)$
- κ determined by calculating $\nu(\theta)$ at $\theta = \pm 1$
  - Eg, vary JES by $\pm 1\sigma$ and look at effect on background
  - Leads to two values of κ: $\kappa_\pm = \nu(\pm 1)$
  - If $\kappa_+ \neq \kappa_-$, uncertainty is asymmetric
- Asymmetric uncertainties were at first handled by bifurcating $\nu(\theta)$

$$\nu(\theta) = \begin{cases} \kappa_+^\theta, & \theta \geq 0 \\ \kappa_-^\theta, & \theta < 0 \end{cases} \tag{2}$$

Asymmetric Uncertainties



- Look at posterior distribution of maximum likelihood estimators $\hat{\theta}$
- Bifurcation leads to unphysical delta function in distribution
  - This causes kink in likelihood vs θ and a discontinuous first derivative
  - Instabilities in Minuit cause delta at $\theta = 0$

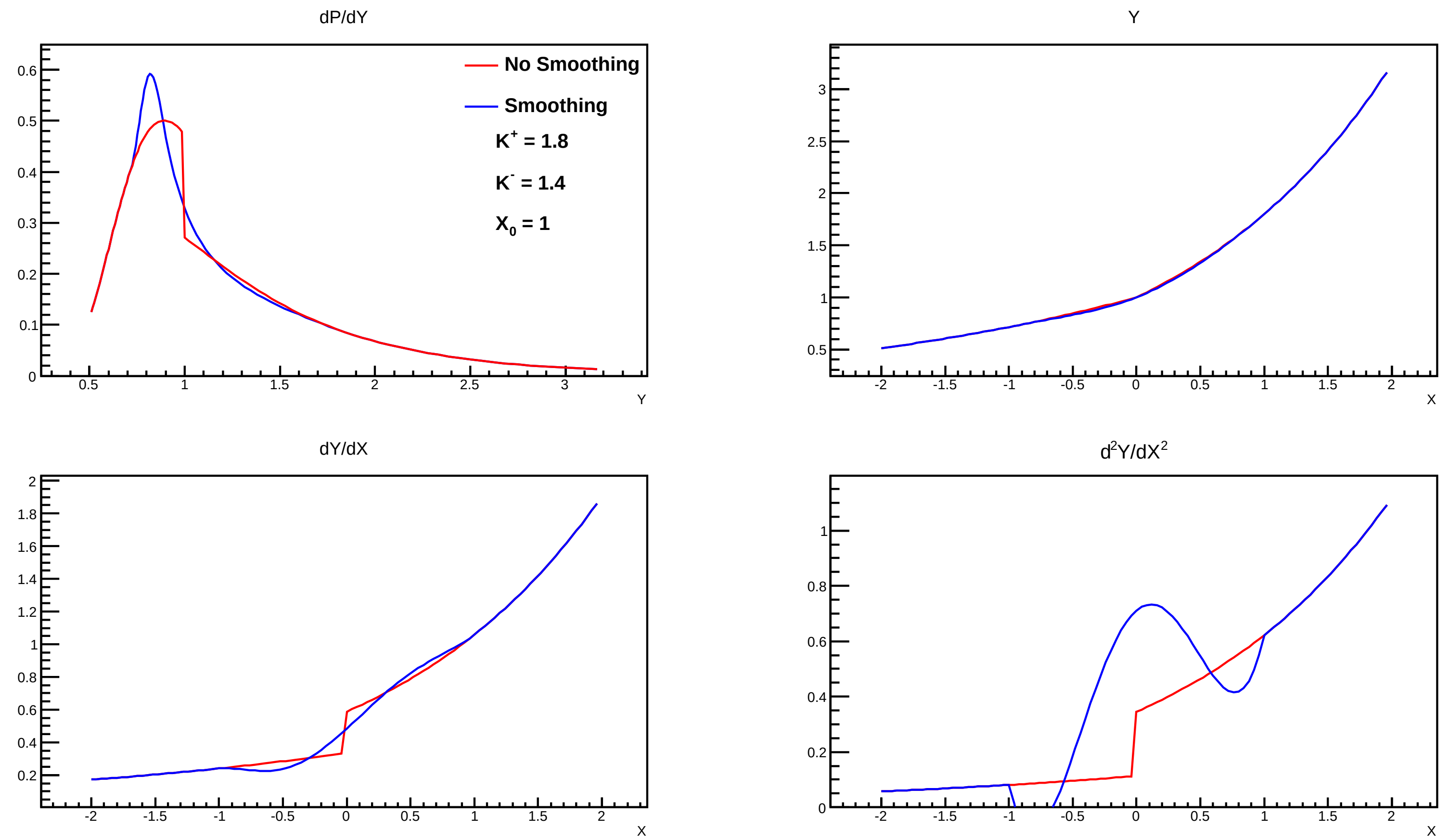
Response Function Interpolation

- Introduce a polynomial function in the transition region
  - A 6D polynomial is required to satisfy ν, $\frac{d\nu}{d\theta}$, $\frac{d^2\nu}{d\sigma^2}$ continuity
- Embed the transition function in some range $(-X_0, +X_0)$ (nominally (-1,1))

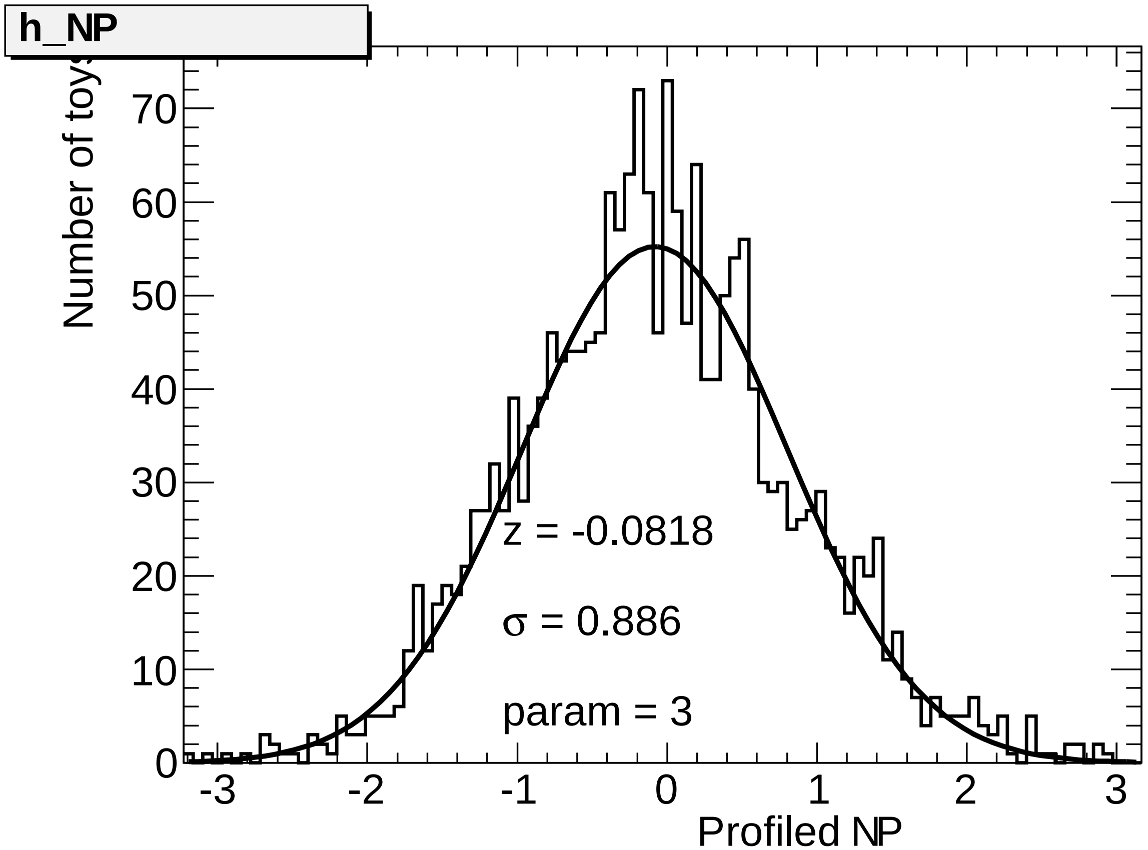
$$\nu(\theta) = \begin{cases} \kappa_+^\theta, & \theta \geq X_0 \\ 1 + \sum_{i=1}^{6} a_i \theta^i, & -X_0 < \theta < X_0 \\ \kappa_-^\theta, & X \leq -X_0 \end{cases}$$

- a_i determined by requiring continuity up to the second derivative at the boundaries $-X_0, +X_0$

Response Function Comparison



Result



- Delta function is removed and θ returns to its proper gaussian distribution

Summary

- Several statistics and modelling issues addressed in last two years during Higgs search
  - These became apparent as model matured and as we learned to develop tools to understand our results
  - Typically developed out of necessesity during conference rush
  - PDF physicality, asymmetric systematic handling
- Asymptotic methods have grown to better approximate toy results
  - Band formulation, uncapping