Setting limits
for the Higgs boson search
(2/2)

Luca Lista

INFN Napoli

Examples from real measurements

A rare process limit using event counting and combination of multiple channels

Search for B $\rightarrow \tau\nu$ at BaBar

Upper limits to $B \to \tau\nu$ at BaBar

- Reconstruct one $B^{\pm}$ with a complete hadronic decay ($e^+e^- \to \Upsilon(4S) \to B^+B^-$)
- Look for a tau decay on other side with missing energy (neutrinos)
  - Five decay channels used: $\mu^-\nu\nu$, $e^-\nu\nu$, $\pi^-\nu$, $\pi^-\pi^0\nu$, $\pi^-\pi^+\pi^-\nu$
- Likelihood function: product of Poissonian likelihoods for the five channels
- Background is known with finite uncertainties from side-band applying scaling factors (taken from simulation)

BABAR Collaboration, Phys.Rev.Lett.95:041804,2005, *Search for the Rare Leptonic Decay $B^- \to \tau^- \nu_\tau$*

Combined likelihood

- Combine the five channels with likelihood ($n_{\mathrm{ch}} = 5$):

$$L(s+b) = \prod_{i=1}^{n_{\mathrm{ch}}} \frac{e^{-(s_i+b_i)}(s_i+b_i)^{n_i}}{n_i!} \qquad L(b) = \prod_{i=1}^{n_{\mathrm{ch}}} \frac{e^{-b_i} b_i^{n_i}}{n_i!}$$

- Define likelihood ratio estimator, as for combined LEP Higgs search:

$$Q = \frac{L(s+b)}{L(b)} = e^{-\sum_{i=0}^{n_{\mathrm{ch}}} s_i} \prod_{i=0}^{n_{\mathrm{ch}}} \left(1 + \frac{s_i}{b_i}\right)^{n_i}$$

- In case the scan of $-2\ln Q$ vs s shows a significant minimum, a non-null measurement of s can be determined
- More discriminating variables may be incorporated in the likelihood definition

Upper limit evaluation

- Use toy Monte Carlo to generate a large number of counting experiments
- Evaluate the C.L. for a signal hypothesis defined as the fraction of C.L. for the $s+b$ and b hypotheses:

$$\mathrm{CL}_s = \frac{\mathrm{CL}_{s+b}}{\mathrm{CL}_b} = \frac{N_{Q_{s+b} \le Q}}{N_{Q_b \le Q}}$$

- Modified frequentist approach

Including (Gaussian) uncertainties

- Nuisance parameters are the background yields b_i known with some uncertainty from side-band extrapolation
- Convolve likelihood with a Gaussian PDF (assuming negligible the tails at negative yield values!)

$$L(s+b) = \prod_{i=1}^{n_{\mathrm{ch}}} \int_{-\infty}^{+\infty} \mathrm{d}b' \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-(b'_i-b_i)^2/2\sigma_i^2} \frac{e^{-(s_i+b'_i)}(s_i+b'_i)^{n_i}}{n_i!}$$

$$L(b) = \prod_{i=1}^{n_{\mathrm{ch}}} \int_{-\infty}^{+\infty} \mathrm{d}b' \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-(b'_i-b_i)^2/2\sigma_i^2} \frac{e^{-b'_i}b_i'^{n_i}}{n_i!}$$

  - Note: b_i is the estimated background, not the “true” one!
- … but C.L. evaluated anyway with a frequentist approach (Toy Monte Carlo)!
- Analytical integrability leads to huge CPU saving!
(L.L., A 517 (2004) 360–363)

Analytical expression

- Simplified analytic Q derivation:

$$Q = \frac{L(s+b)}{L(b)} = e^{-\sum_{i=0}^{n_{\mathrm{ch}}} s_i} \prod_{i=0}^{n_{\mathrm{ch}}} \frac{p_{n_i}(s_i + b_i - \sigma_i^2, \sigma_i)}{p_{n_i}(b_i - \sigma_i^2, \sigma_i)}$$

- Where $p_n(\alpha, \beta)$ are polynomials defined with a recursive relation:

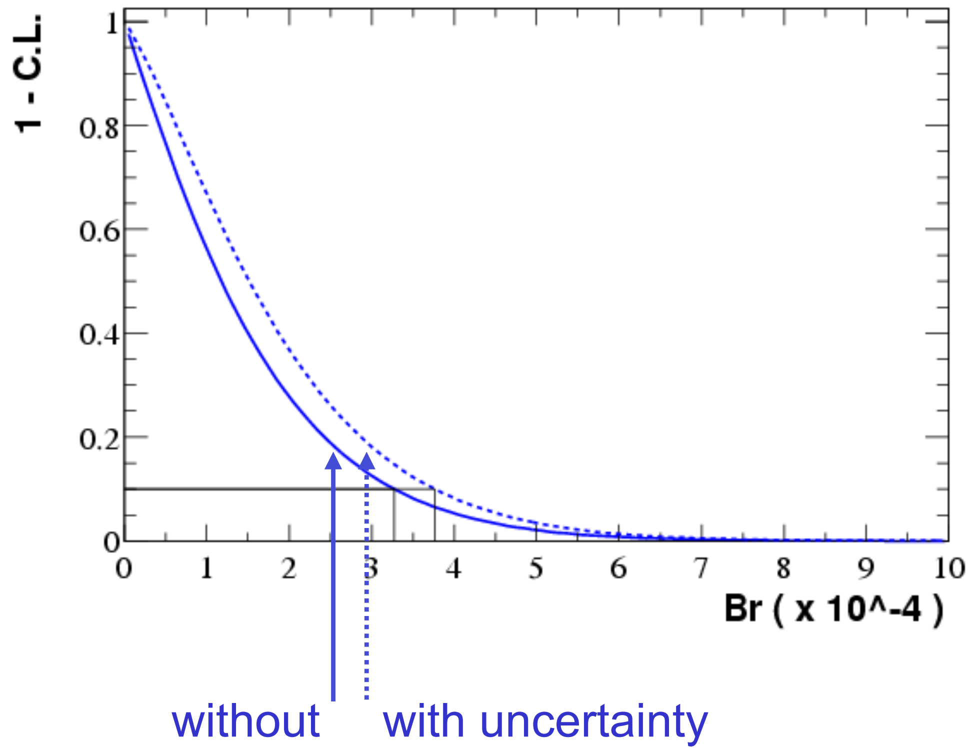
$$\begin{cases} p_0(\alpha,\beta) = 1, \\ p_1(\alpha,\beta) = \alpha, \\ p_2(\alpha,\beta) = \alpha^2 + \beta^2, \\ p_3(\alpha,\beta) = \alpha^3 + 3\alpha\beta^2, \\ p_4(\alpha,\beta) = \alpha^4 + 6\alpha^2\beta^2 + 3\beta^4, \\ p_5(\alpha,\beta) = \alpha^5 + 10\alpha^3\beta^2 + 15\alpha\beta^4, \\ \vdots \end{cases}$$

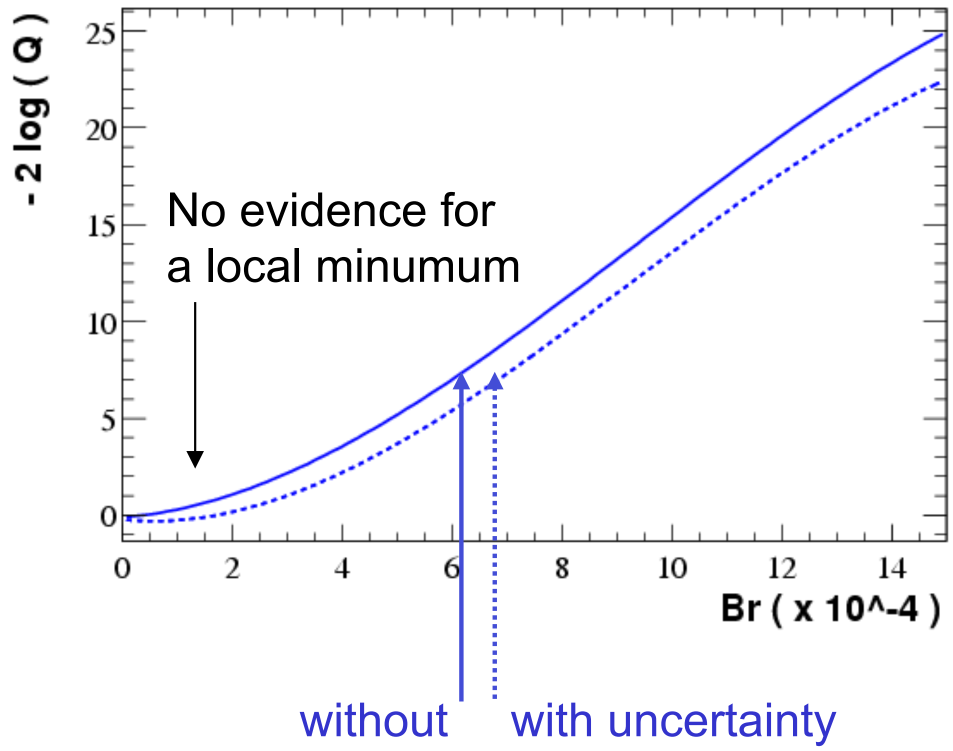
$$p_n(\alpha,\beta) = \alpha p_{n-1}(\alpha,\beta) + (n+1)\beta^2 p_{n-2}(\alpha,\beta)$$

… but in many cases it's hard to be so lucky!

Branching ratio: $\int L dt = 82$ fb^{-1}

- Low statistics scenario






```
RooStats::HypoTestInverter
```

$B \rightarrow \tau\nu$ was eventually measured

- … and is now part of the PDF

2011 Review of Particle Physics.
Please use this **CITATION**: K. Nakamura *et al.* (Particle Data Group), Journal of Physics G**37**, 075021 (2010) and 2011 partial update for the 2012 edition.

B^+ BRANCHING RATIOS

back to $B^\pm$

$\Gamma(\tau^+ \nu_\tau)/\Gamma_{total}$ | References | History since 1996

Γ_{26}/Γ

See the note on "Decay Constants of Charged Pseudoscalar Mesons" in the D_s^+ Listings.

VALUE (10^{-4})	*CL%*	*DOCUMENT ID*	*TECN*	*COMMENT*
1.65 ± 0.34	**OUR AVERAGE**			
$1.7 \pm 0.8 \pm 0.2$		[1,2] AUBERT	10E BABR	$e^+ e^- \rightarrow \Upsilon(4S)$
$1.54^{+0.38}_{-0.37}{}^{+0.29}_{-0.31}$		[1,3] HARA	10 BELL	$e^+ e^- \rightarrow \Upsilon(4S)$
$1.8^{+0.9}_{-0.8} \pm 0.45$		[4,1] AUBERT	08D BABR	$e^+ e^- \rightarrow \Upsilon(4S)$
$1.79^{+0.56}_{-0.49}{}^{+0.46}_{-0.51}$		[4,1] IKADO	06 BELL	$e^+ e^- \rightarrow \Upsilon(4S)$
* * * We do not use the following data for averages, fits, limits, *etc.* * * *				
$0.9 \pm 0.6 \pm 0.1$		[1,2] AUBERT	07AL BABR	Repl. by AUBERT 2010E
<2.6	90	[1] AUBERT	06K BABR	$e^+ e^- \rightarrow \Upsilon(4S)$
<4.2	90	[1] AUBERT,B	05B BABR	Repl. by AUBERT 2006K
<8.3	90	[5] BARATE	01E ALEP	$e^+ e^- \rightarrow Z$
<8.4	90	[1] BROWDER	01 CLE2	$e^+ e^- \rightarrow \Upsilon(4S)$
<5.7	90	[6] ACCIARRI	97F L3	$e^+ e^- \rightarrow Z$
<104	90	[7] ALBRECHT	95D ARG	$e^+ e^- \rightarrow \Upsilon(4S)$
<22	90	ARTUSO	95 CLE2	$e^+ e^- \rightarrow \Upsilon(4S)$
<18	90	[8] BUSKULIC	95 ALEP	$e^+ e^- \rightarrow Z$

[1] Assumes equal production of B^+ and B^0 at the $\Upsilon(4S)$.
[2] Requires one reconstructed semileptonic B decay $B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell X$ in the recoil.
[3] Requires one reconstructed semileptonic B decay $B^- \rightarrow D^{(*)0} \ell^- \bar{\nu}_\ell X$ in the recoil.
[4] The analysis is based on a sample of events with one fully reconstructed tag B in a hadronic decay mode $B^- \rightarrow D^{(*)0} X^-$.
[5] The energy-flow and b-tagging algorithms were used.
[6] ACCIARRI 1997F uses missing-energy technique and $f(b \rightarrow B^-) = (38.2 \pm 2.5)\%$.
[7] ALBRECHT 1995D uses full reconstruction of one B decay as tag.
[8] BUSKULIC 1995 uses same missing-energy technique as in $\bar{b} \rightarrow \tau^+ \nu_\tau X$, but analysis is restricted to endpoint region of missing-energy distribution.

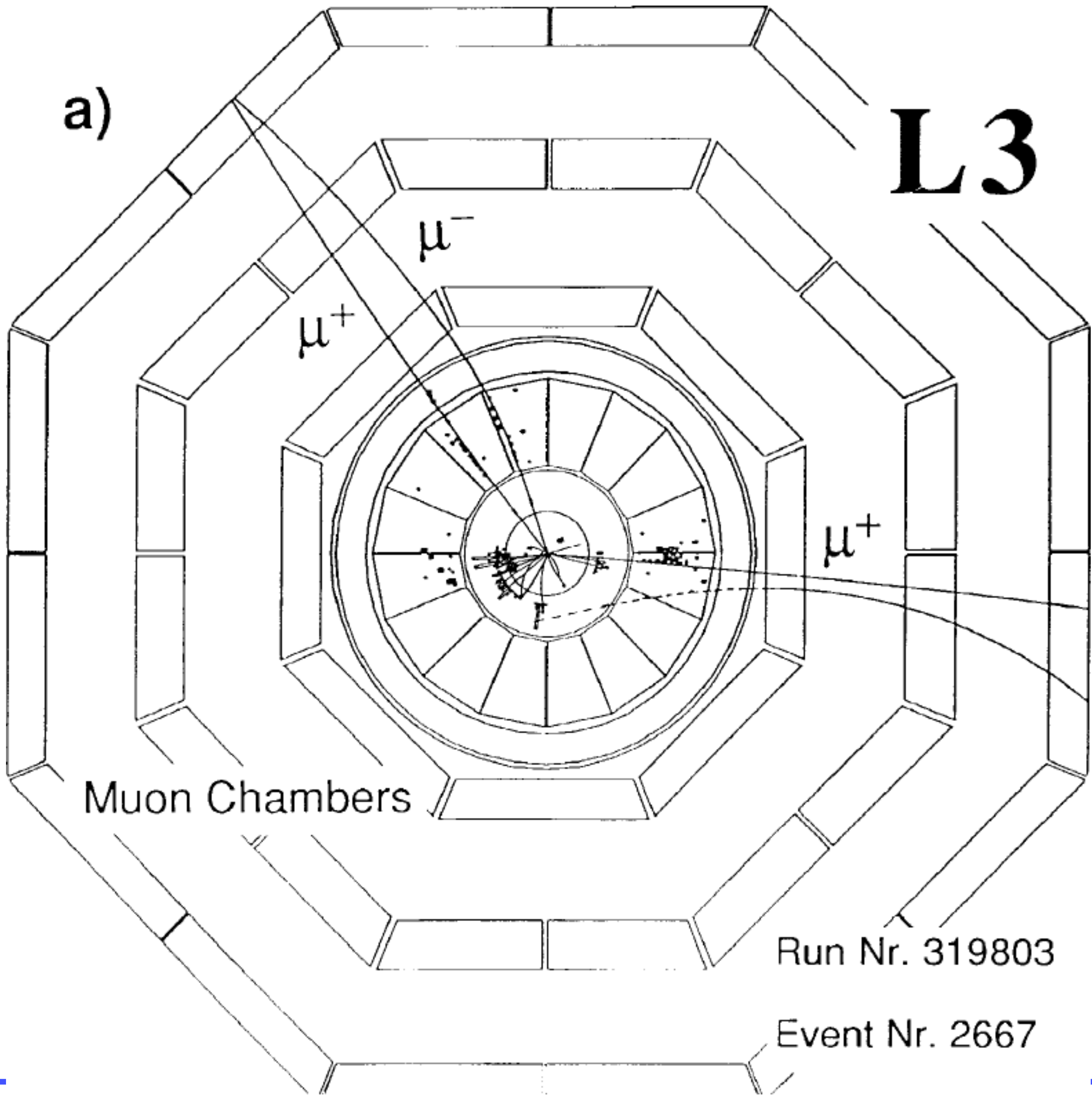
A Bayesian approach to Higgs search with small background

Higgs search at LEP-I (L3)

Higgs search at LEP

- Production via $e^+e^- \to HZ^* \to b\bar{b}l^+l^-$
- Higgs candidate mass measured via missing mass to lepton pair
- Most of the background rejected via kinematic cuts and isolation requirements for the lepton pair
- Search mainly dominated by statistics
- A few background events survived selection (first observed in L3 at LEP-I)

First Higgs candidate ($m_H \cong 70$ GeV)



Extended likelihood approach

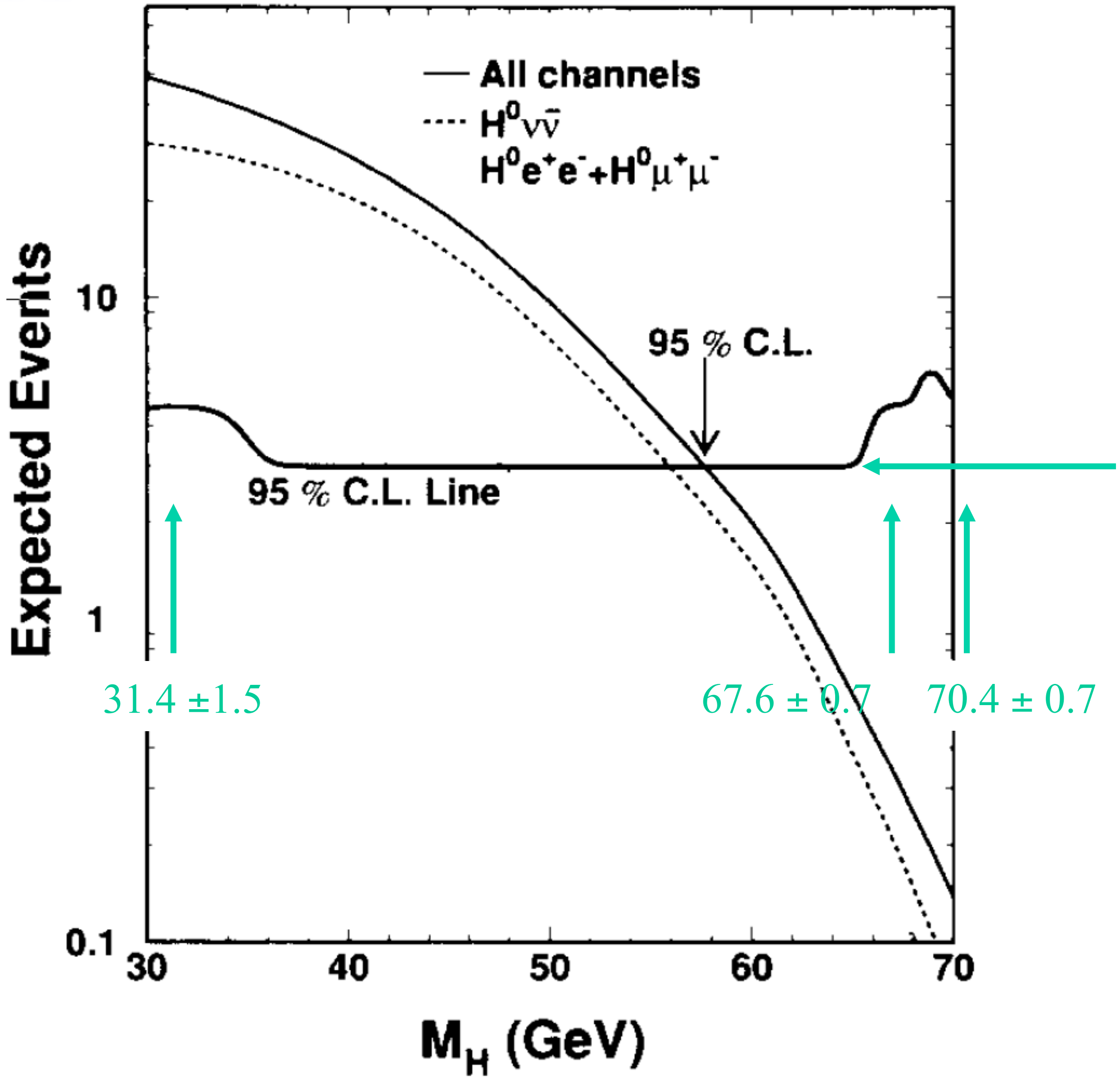
- Assume both signal and background are present, with different PDF for mass distribution: Gaussian peak for signal, flat for background (from Monte Carlo samples):

$$L = e^{-(s+b)} \prod_{i=1}^{n} (s P_s(m_i) + b P_b(m_i))$$

- Bayesian approach can be used to extract the upper limit, with uniform prior, $\pi(s) = 1$:

$$1 - \mathrm{CL} = \frac{\int_{s^{\mathrm{up}}}^{\infty} e^{-s} \prod_{i=1}^{n} (s P_s(m_i) + b P_b(m_i)) \mathrm{d}s}{\int_{0}^{\infty} e^{-s} \prod_{i=1}^{n} (s P_s(m_i) + b P_b(m_i)) \mathrm{d}s}$$

Application to Higgs search at L3

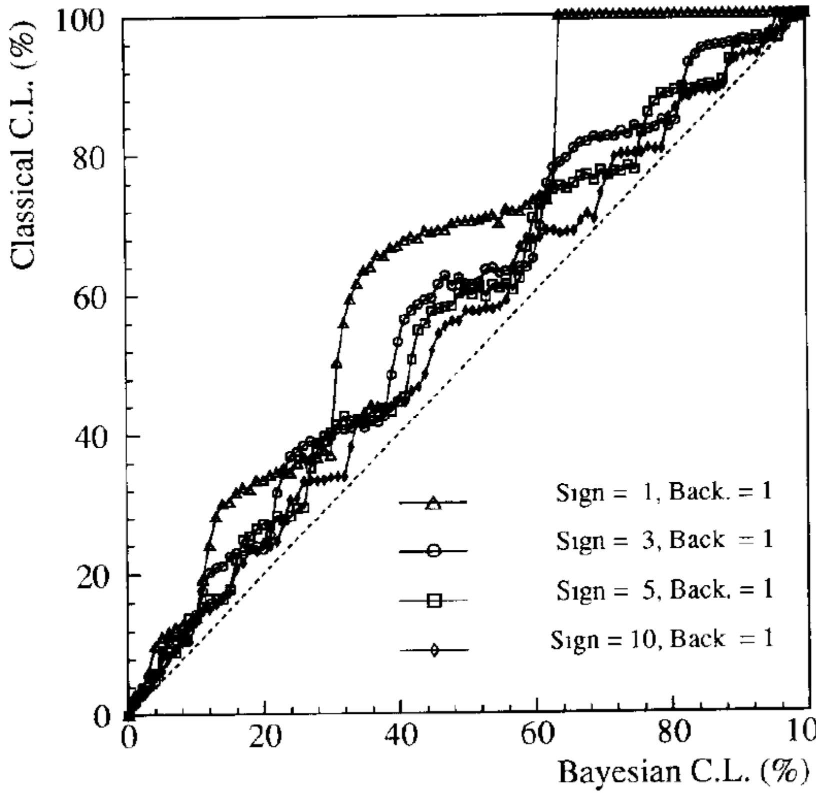


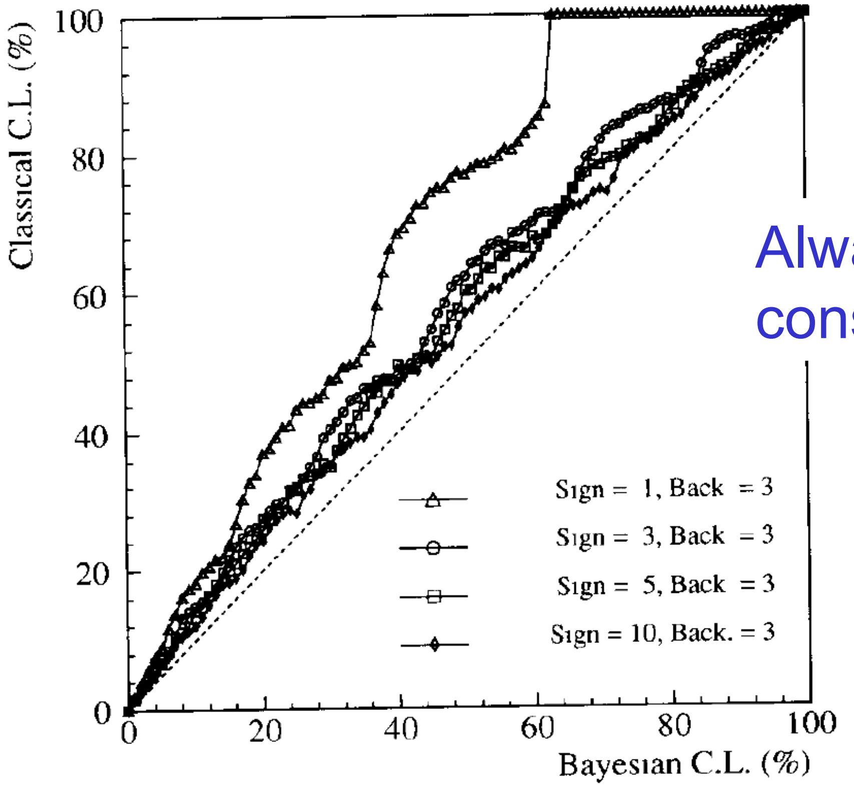
LEP-I

3 events
“standard” limit

Comparison with frequentist C.L.

- Toy MC can be generated for different signal and background scenarios
- frequentist coverage ("classical" CL) can be computed counting the fraction of toy experiments above/below the Bayesian limit





Always conservative!

Higgs search at LEP-II

Combined search using CLs

Combined Higgs search at LEP-II

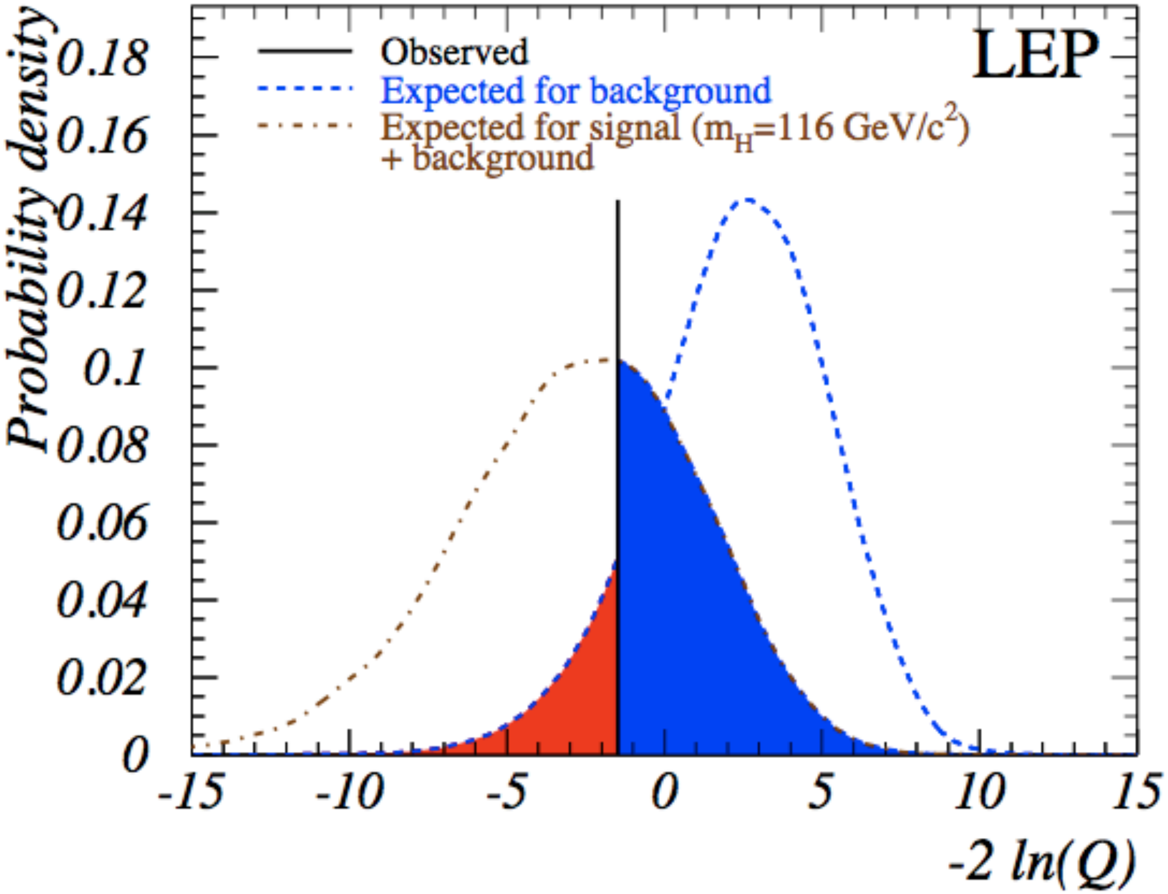
- Extended likelihood definition:

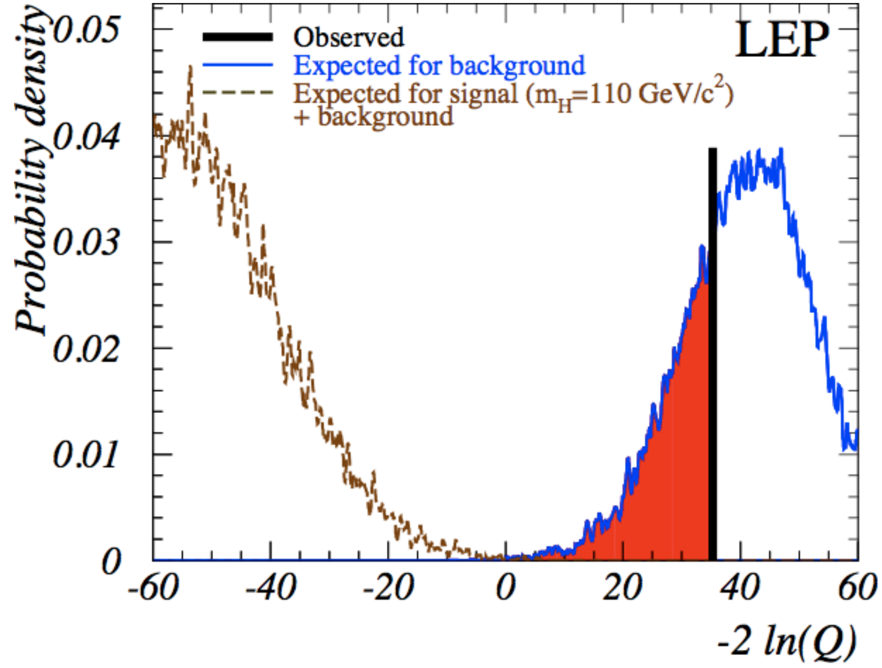
$$L(\eta) = \prod_{k=1}^{n_{\mathrm{ch}}} \frac{e^{-(\eta s_k(m_H)+b_k)} \cancel{(\eta s_k(m_H) + b_k)^{n_k}}}{n_k!} \times \prod_{j=1}^{n_k} \frac{\eta s_k(m_H) S_k(\vec{x_{jk}}; m_H) + b_k B_k(\vec{x_{jk}})}{\cancel{\eta s_k(m_H) + b_k}}$$

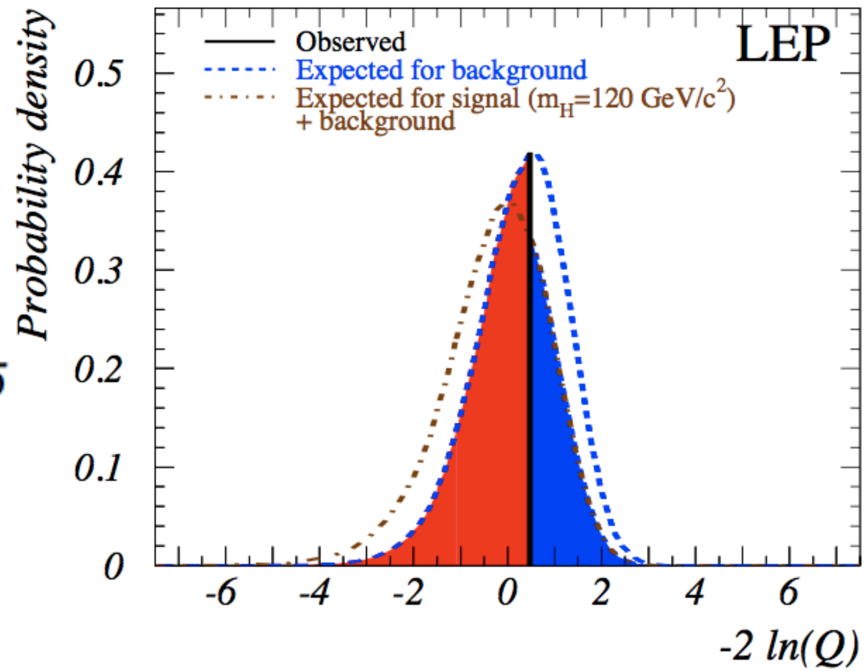
- η = 0 for b only, 1 for $s + b$ hypotheses
- Likelihood ratio:

$$-2 \ln Q(m_H) = 2 \sum_{k=1}^{n_{\mathrm{ch}}} \left[s_k(m_H) - \sum_{j=1}^{n_k} \ln \left(1 + \frac{s_k(m_H) S_k(\vec{x_{jk}}; m_k)}{b_k B_k(\vec{x_{jk}})} \right) \right]$$

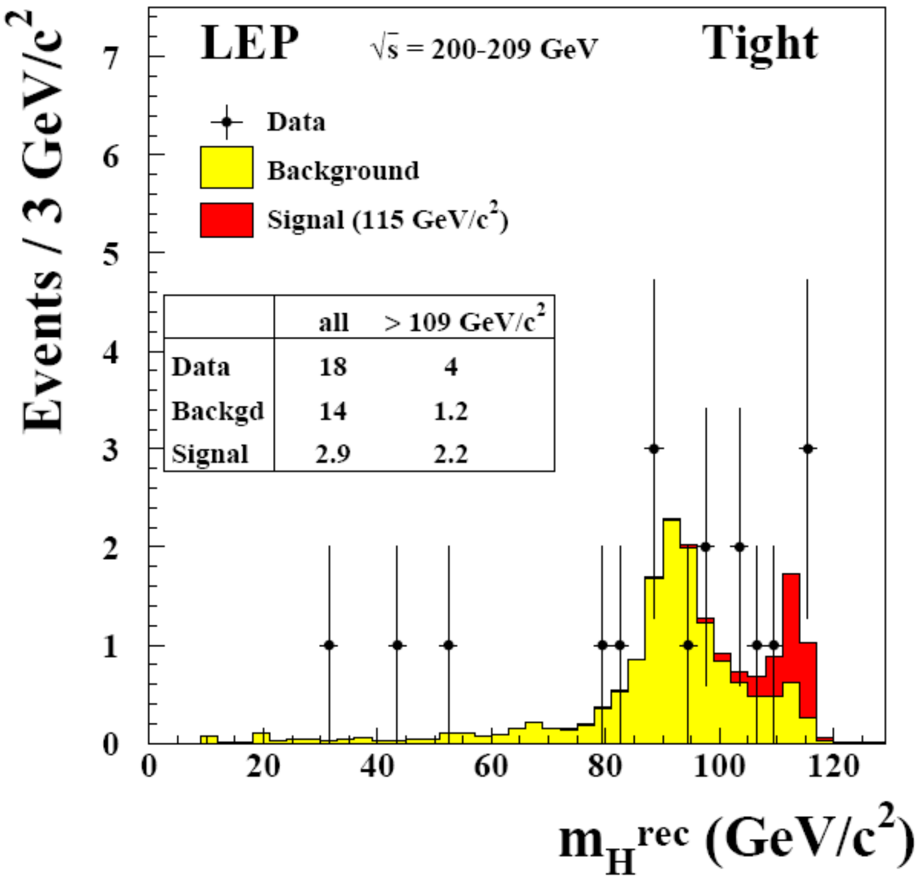
CLs PDF plot

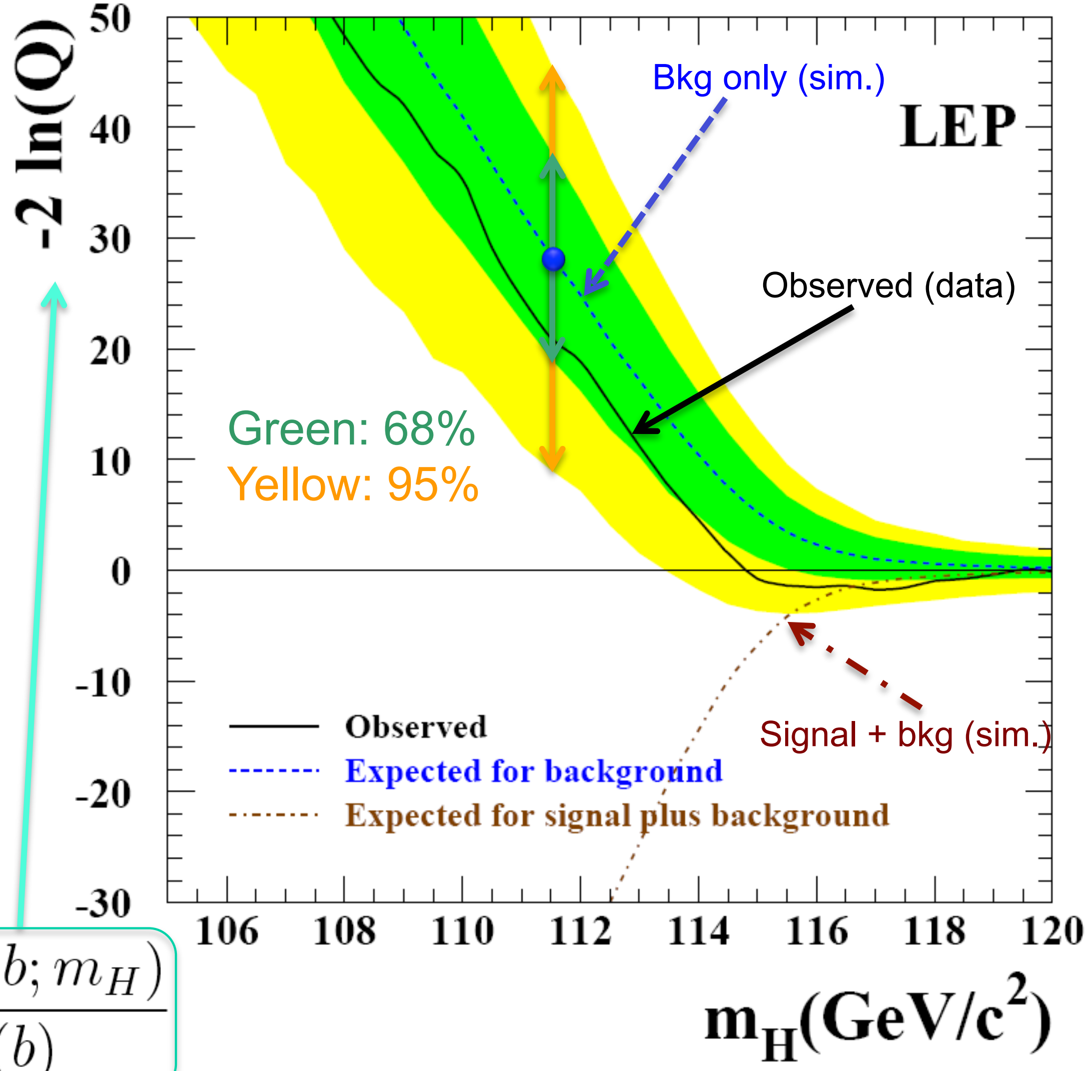

Probability density
Observed
Expected for background
Expected for signal (m_H=116 GeV/c²)
+ background
LEP
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-15
-10
-5
0
5
10
15
-2 ln(Q)


Probability density
Observed
Expected for background
Expected for signal (m_H=110 GeV/c²)
+ background
LEP
0.05
0.04
0.03
0.02
0.01
0
-60
-40
-20
0
20
40
60
-2 ln(Q)


Probability density
Observed
Expected for background
Expected for signal (m_H=120 GeV/c²)
+ background
LEP
0.5
0.4
0.3
0.2
0.1
0
-6
-4
-2
0
2
4
6
-2 ln(Q)

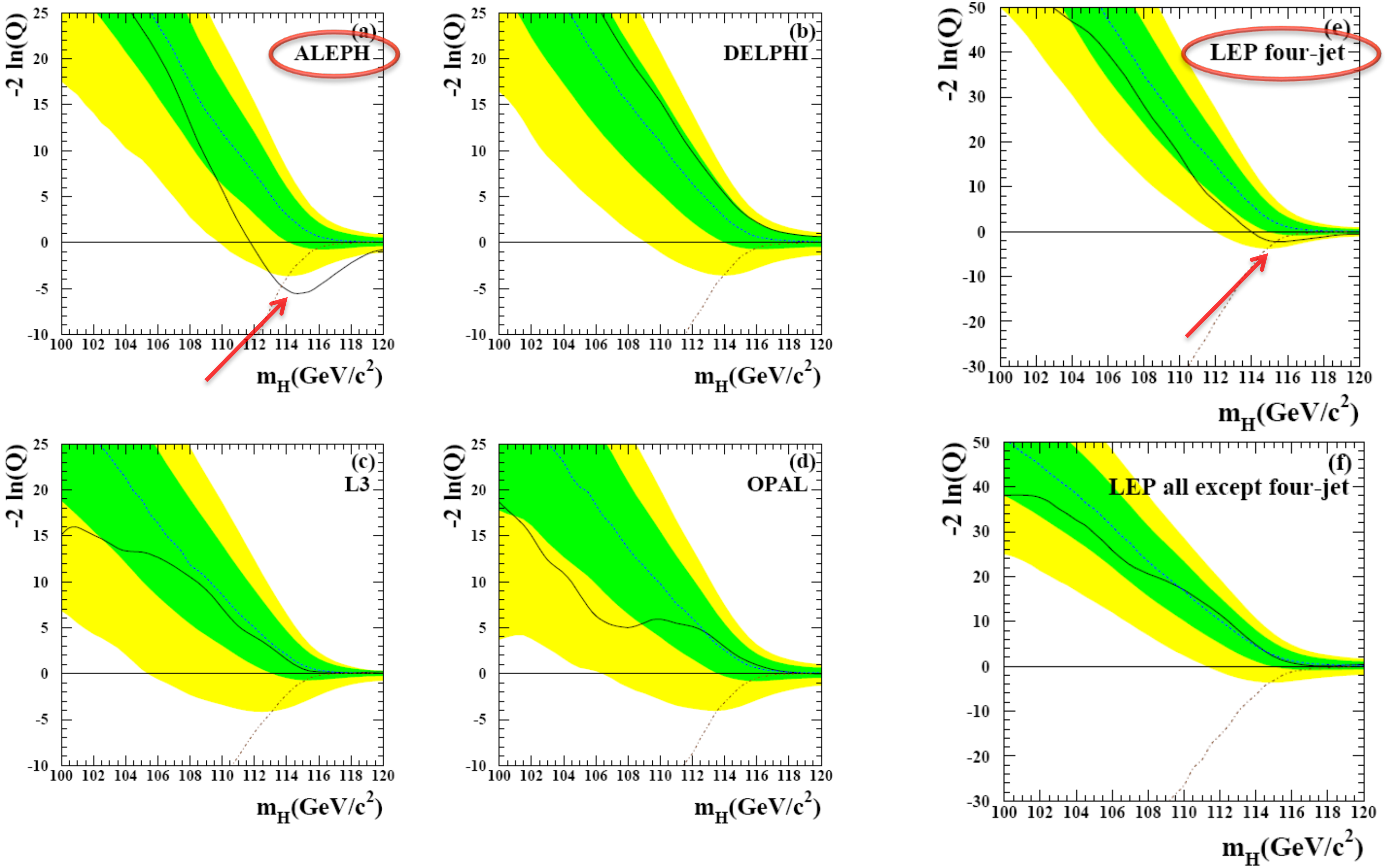
Mass scan plot



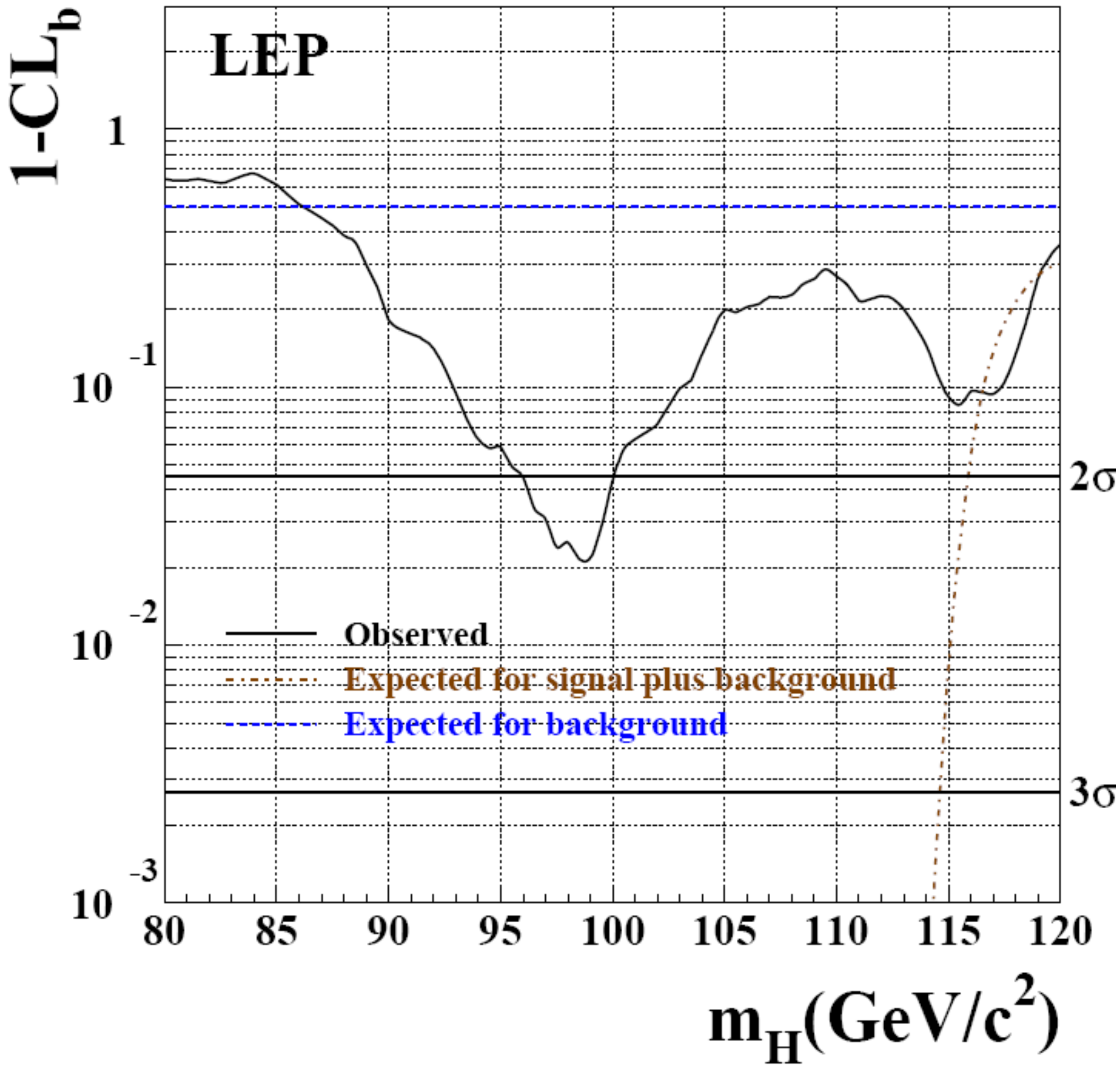


$$-2\ln Q = -2\ln\frac{L(s+b;m_H)}{L(b)}$$

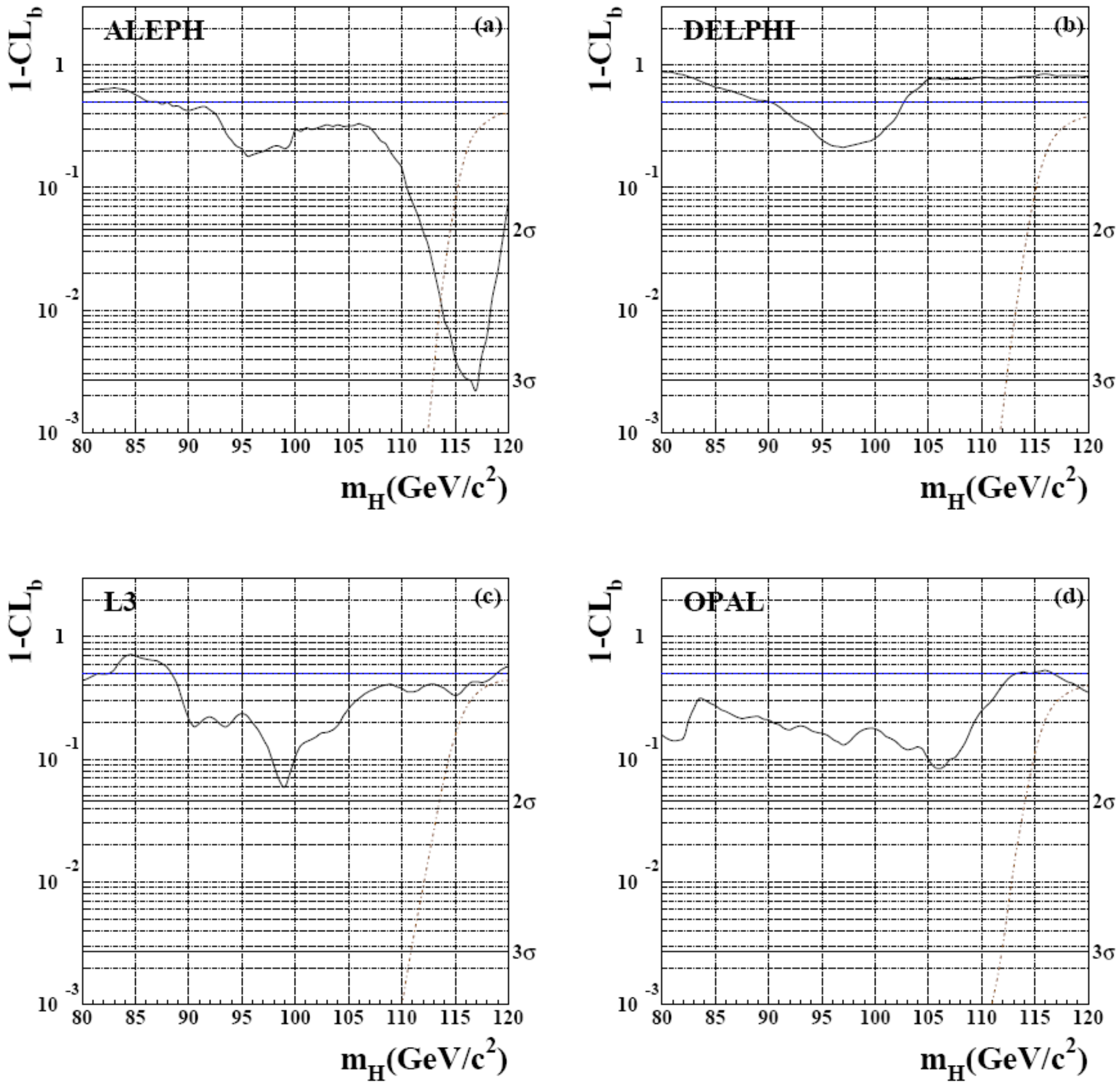
By experiment & channel



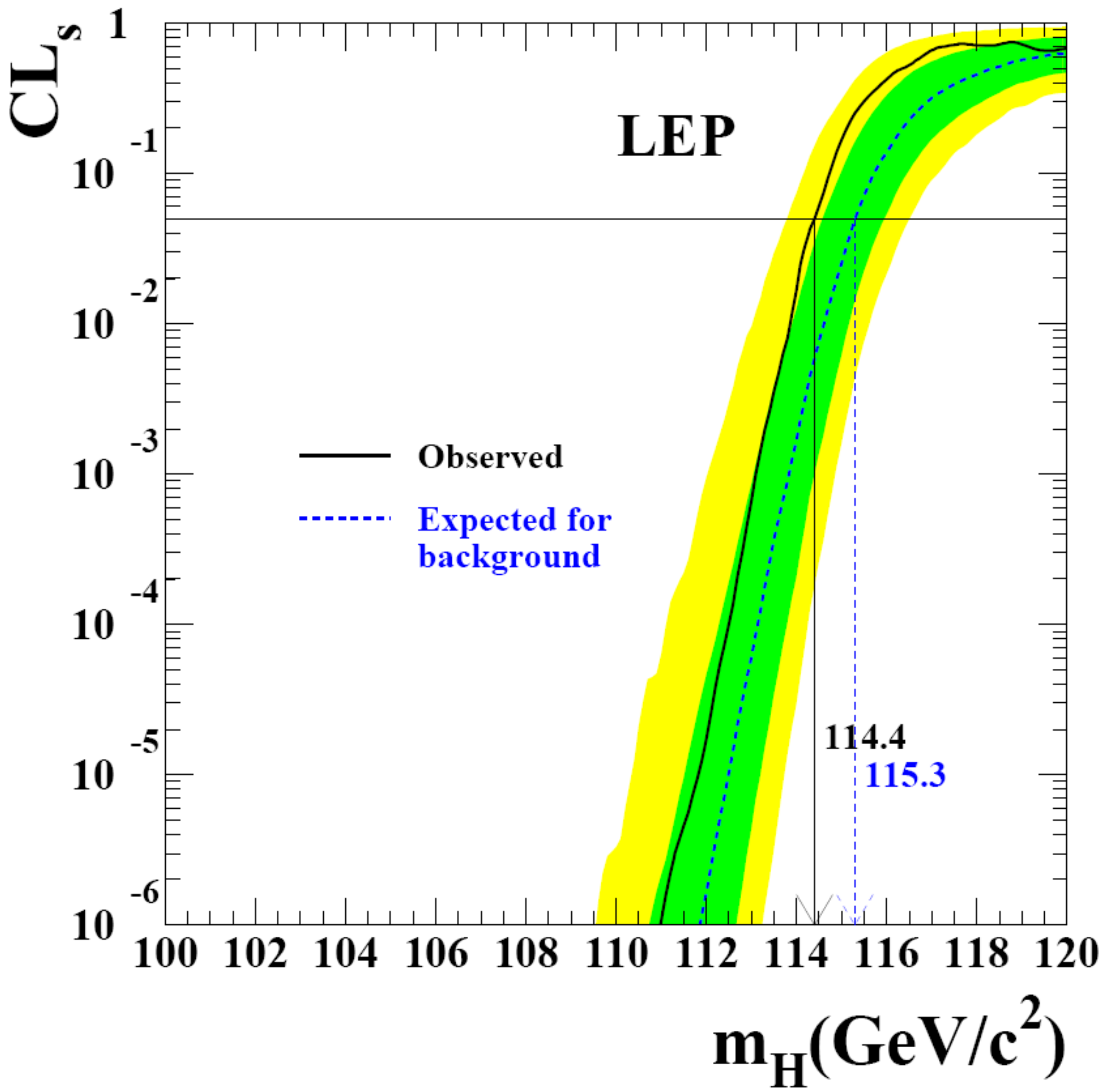
Background hypothesis C.L.



Background C.L. by experiment



Signal hypothesis C.L.



Higgs search at LHC

Combined search using CLs

Higgs search at LHC method

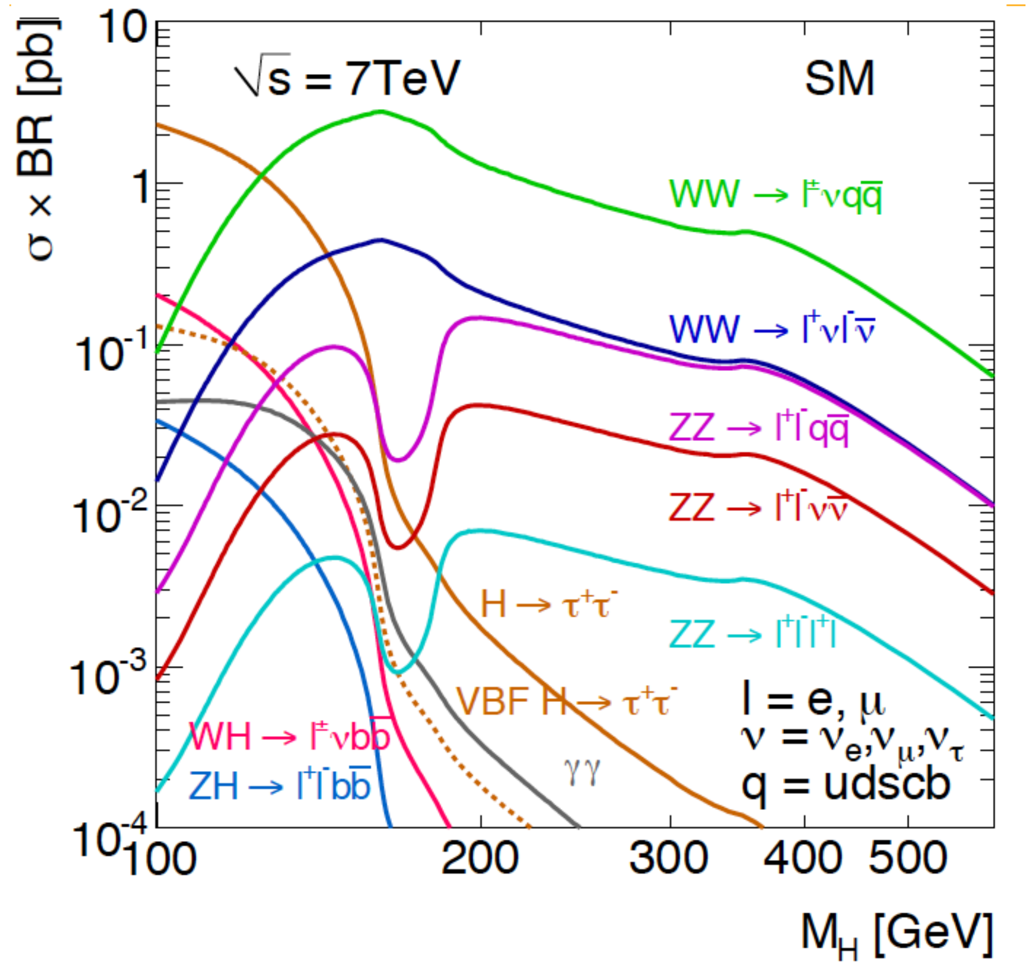
- Agreed method between ATLAS and CLS
- Test statistics:

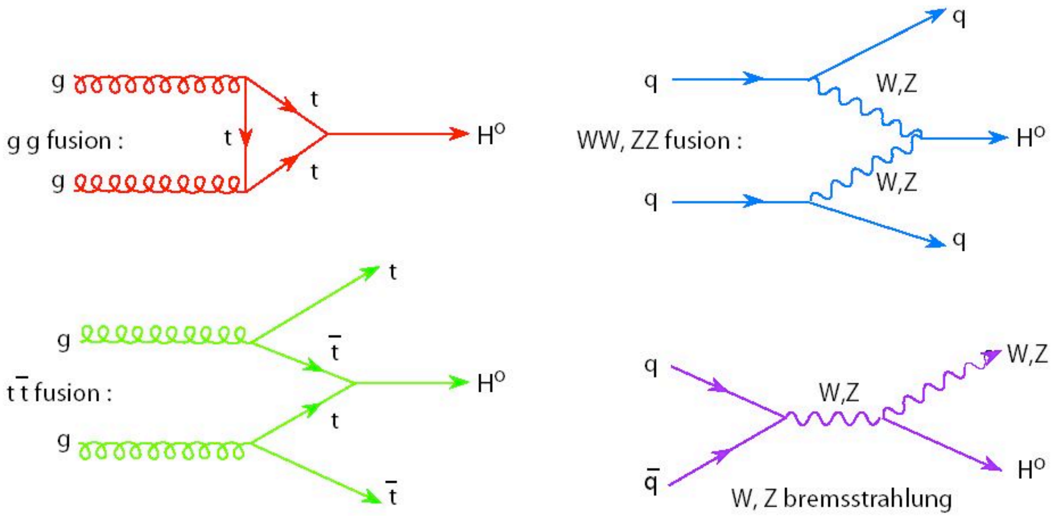
$$q_\mu = -2\ln \frac{L(\text{data}; \mu\hat{\hat{\theta}}_\mu)}{L(\text{data}; \hat{\mu}, \hat{\theta}_\mu)}$$

- Has good asymptotic behavior
- Nuisance parameters are profiled
- Uncertainties are modeled with log-normal PDFs
- CLs protects against unphysical limits in cases of large downward background fluctuations
- Observed and median expected values of CLs limits presented as 68% and 95% belts

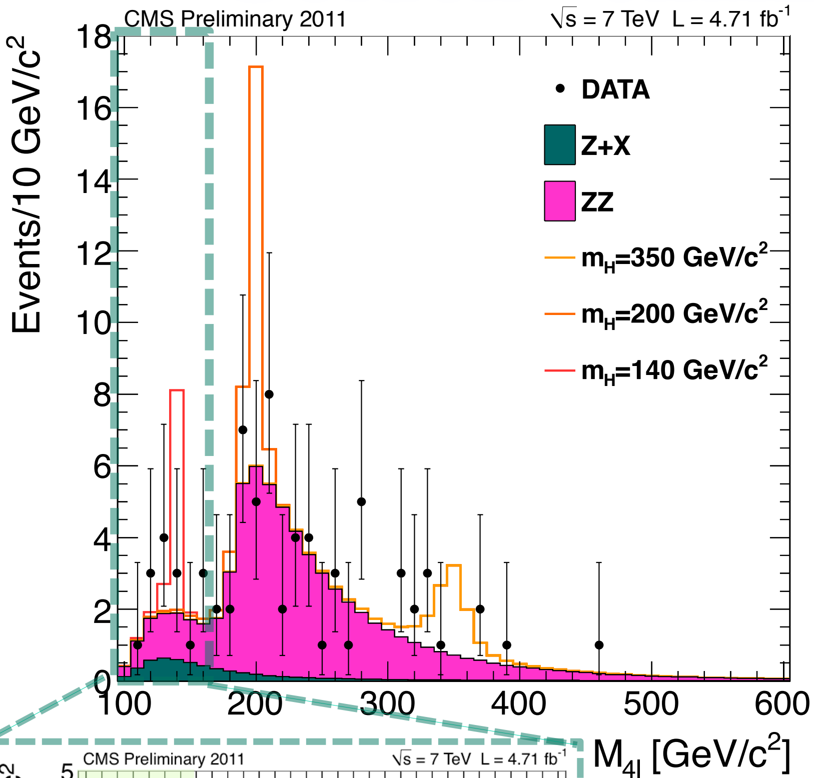
Higgs boson production at LHC

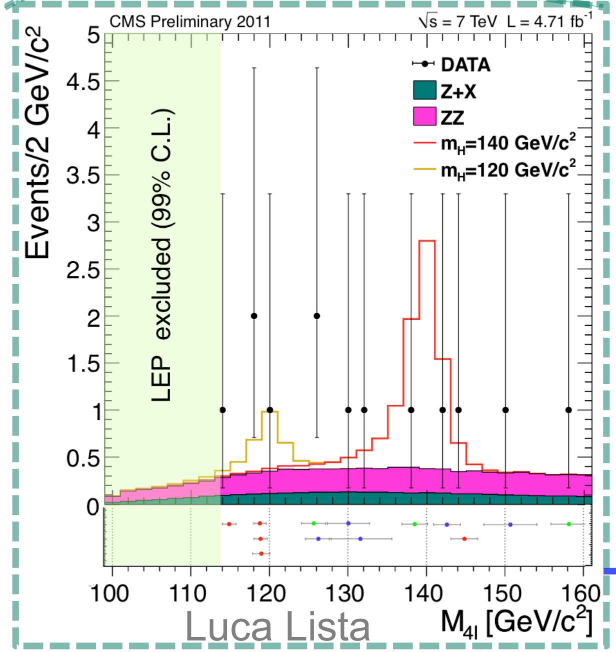
- Decays are favored into heavy particles (top, Z, W, b, …)
- Most abundant production via "gluon fusion" and "vector-boson fusion"

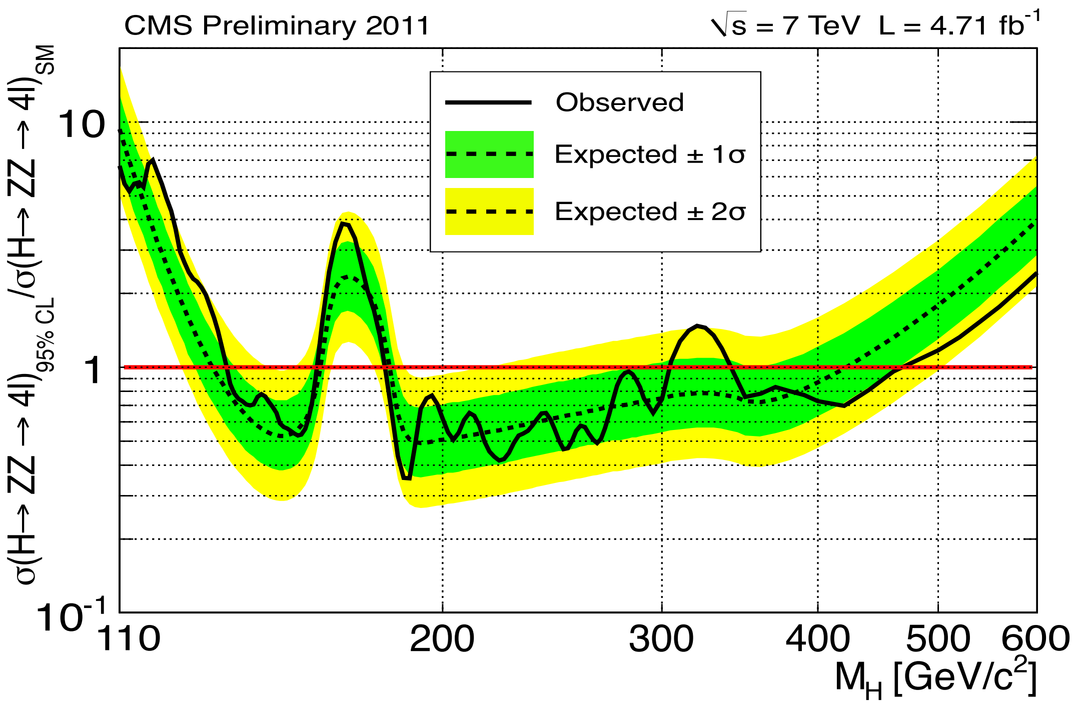




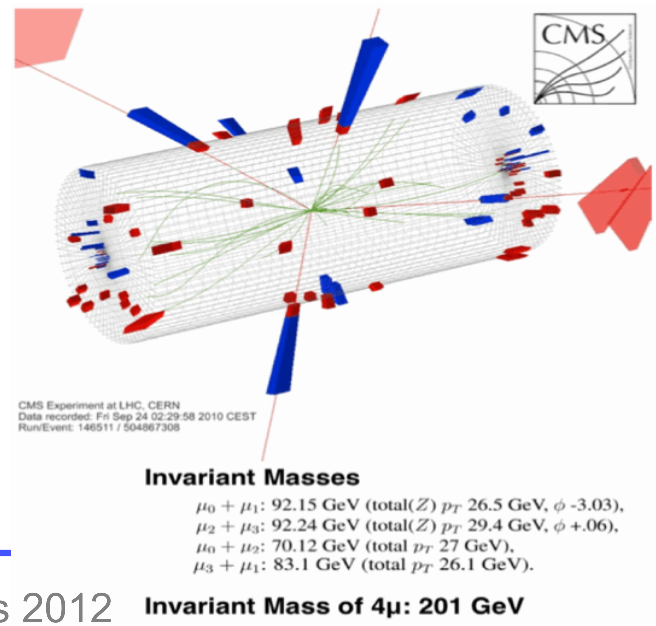
"Golden" channel: H→ZZ →4l (l=e,μ)



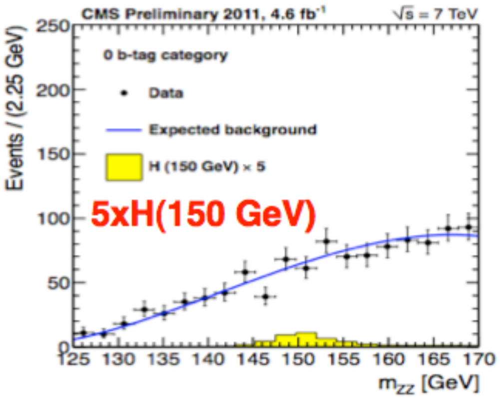


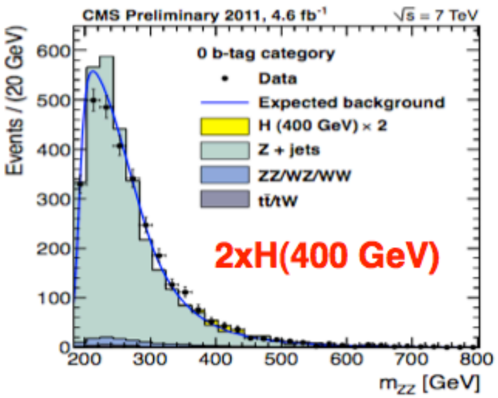


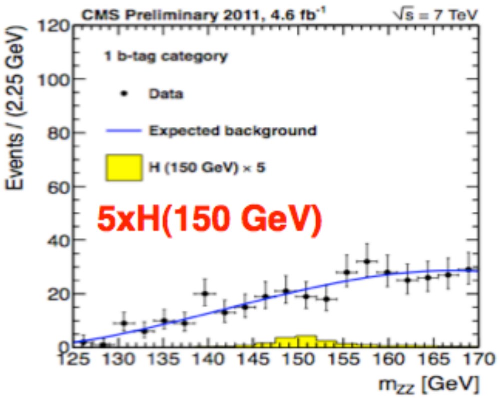
Mass resolution is ~2-3 GeV

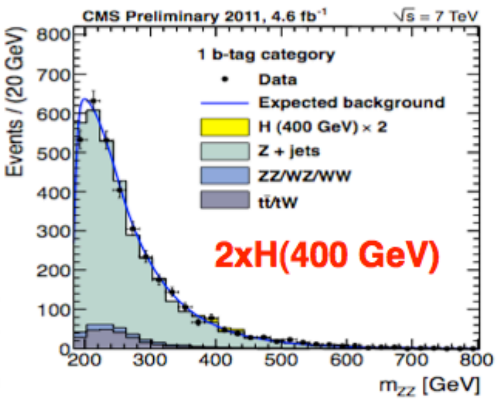


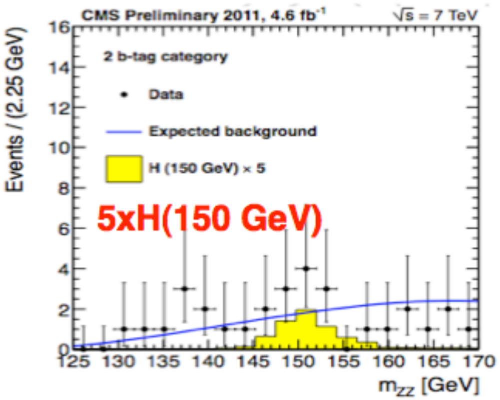
Jets: H→ZZ→2l2q (l=e,μ)

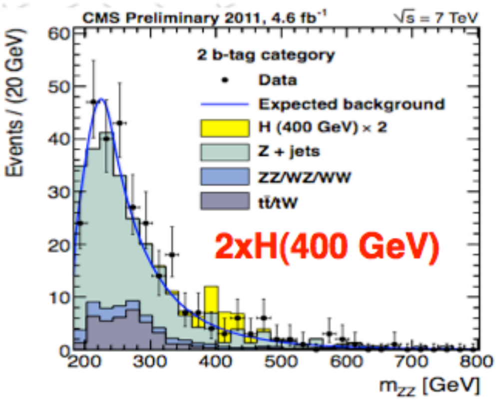


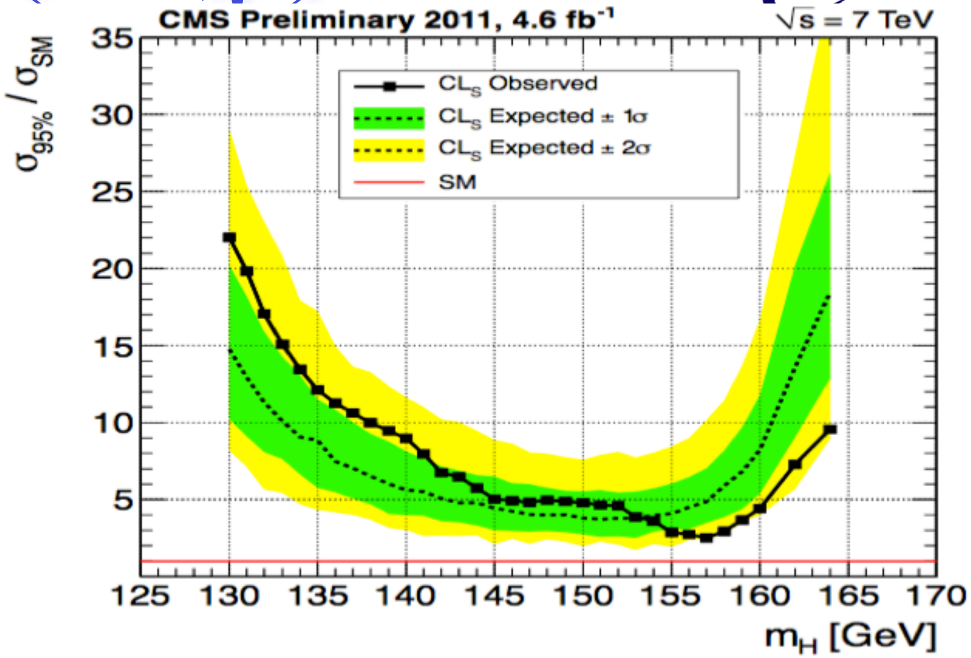


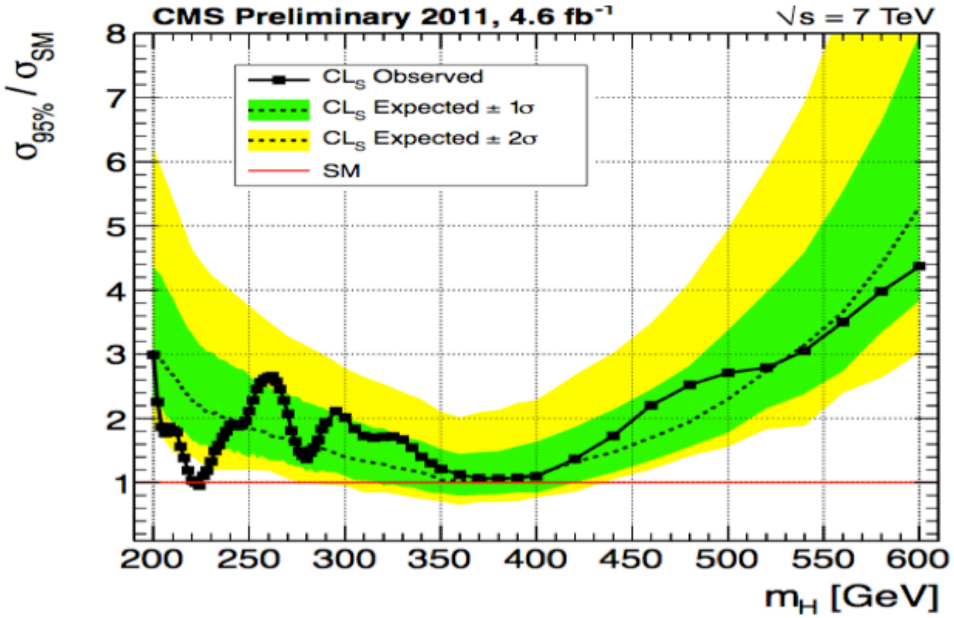






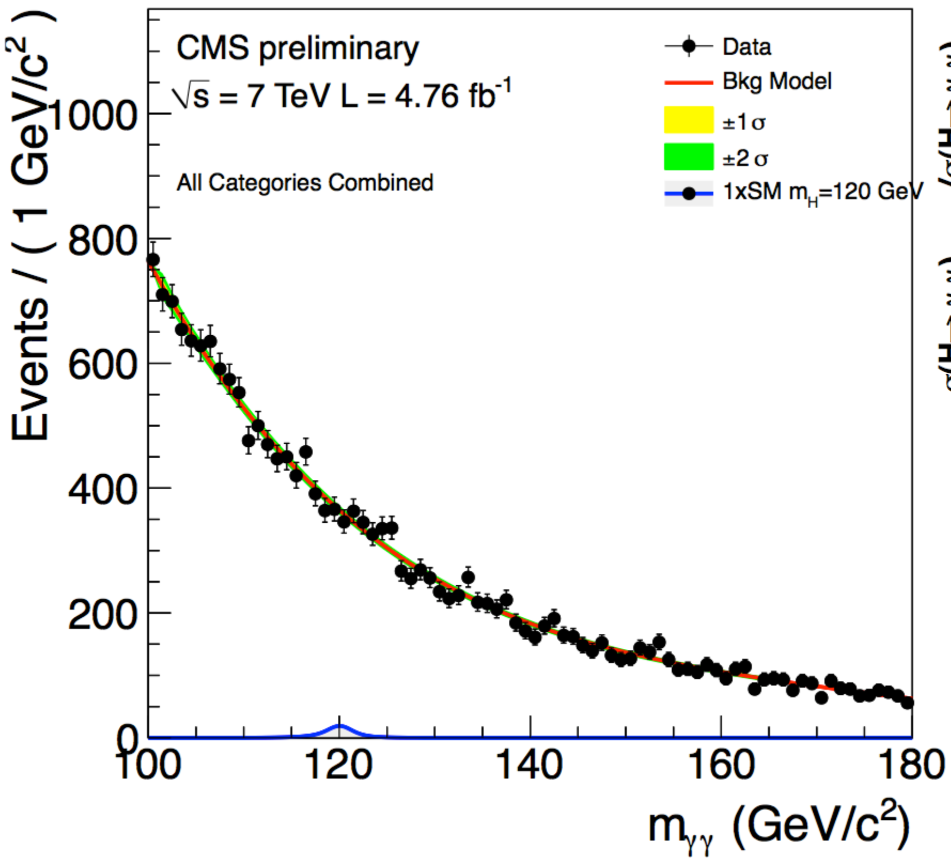


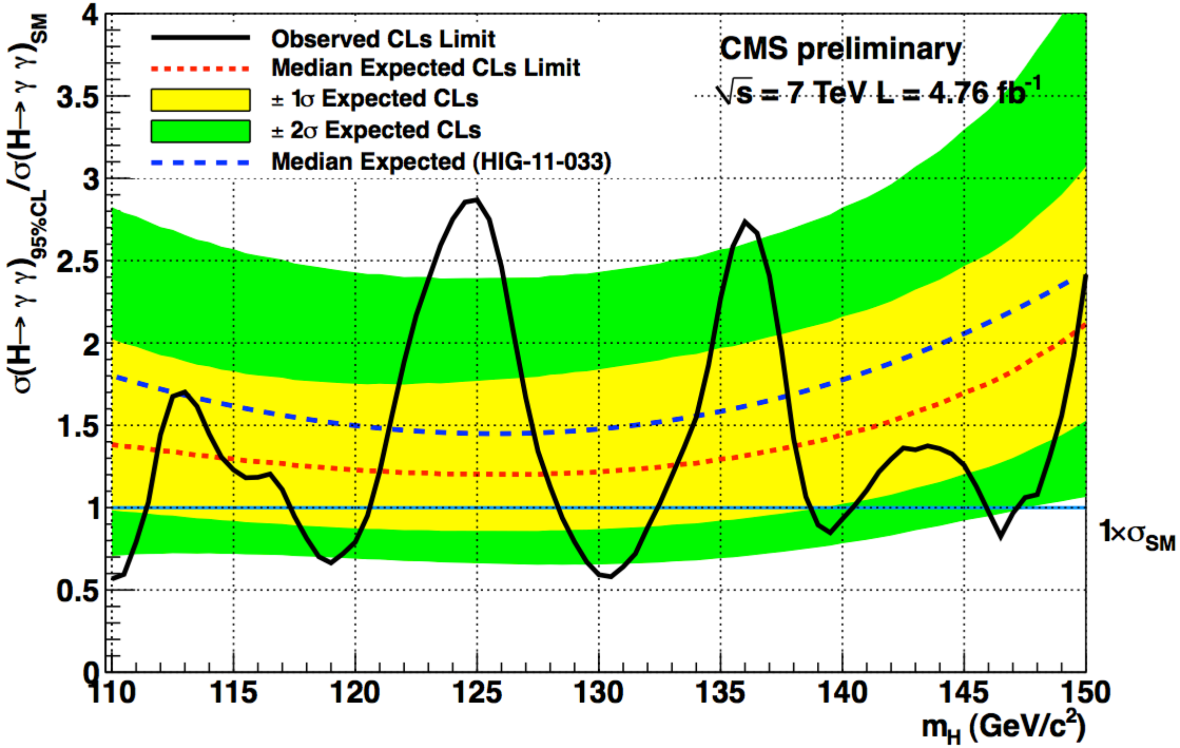




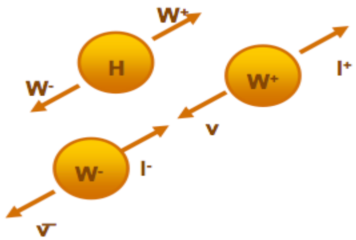
H→γγ

- Large background, good resolution

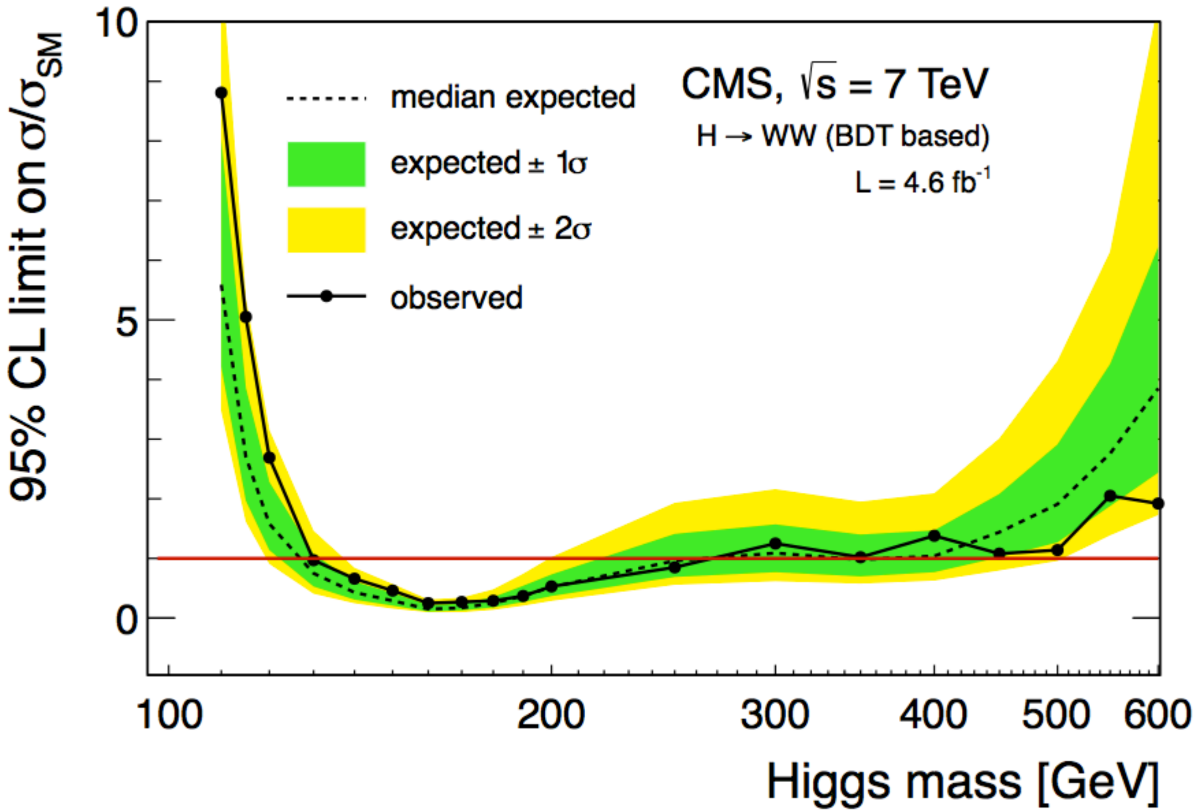


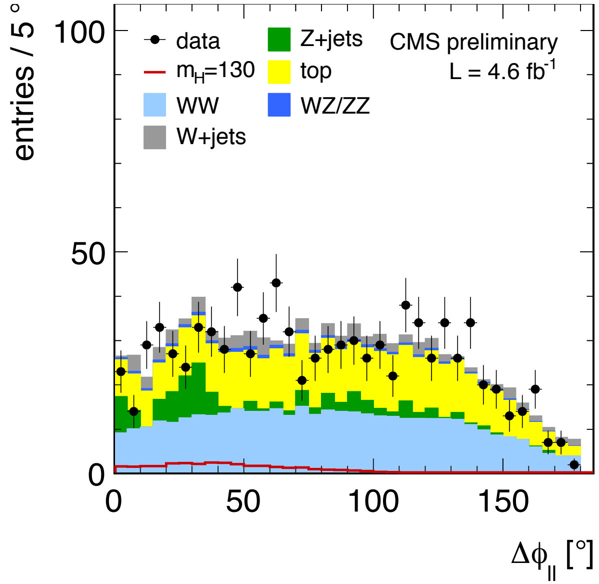


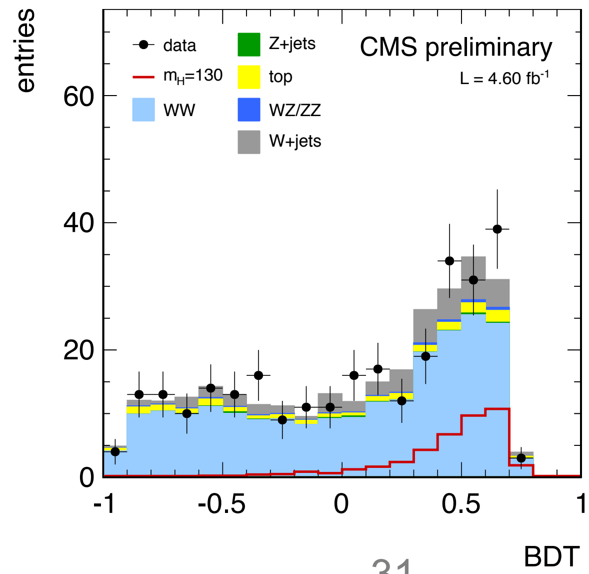
H→WW→2l2ν



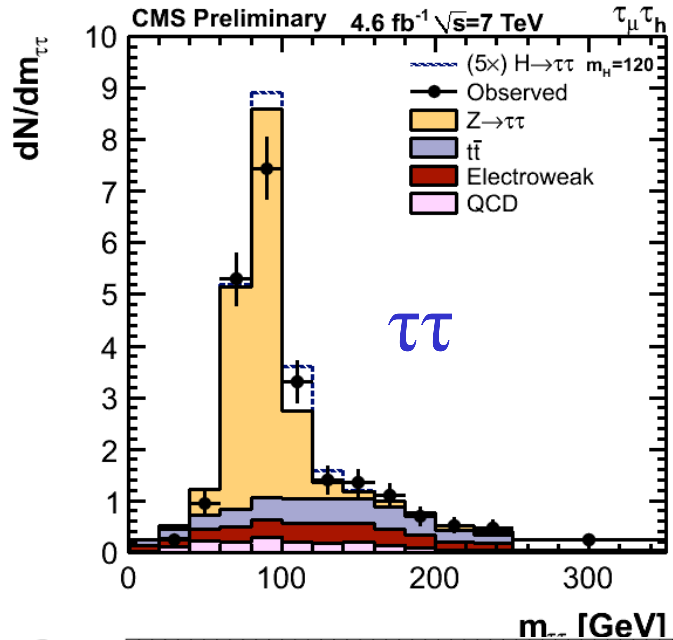
- Can't reconstruct Higgs mass due to neutrinos
- Signal can be discriminated vs background using angular distribution (Higgs boson has spin zero)
  - Two leptons tend to be alligned in Higgs event
- A multivariate analysis maximizes sig/bkg separation

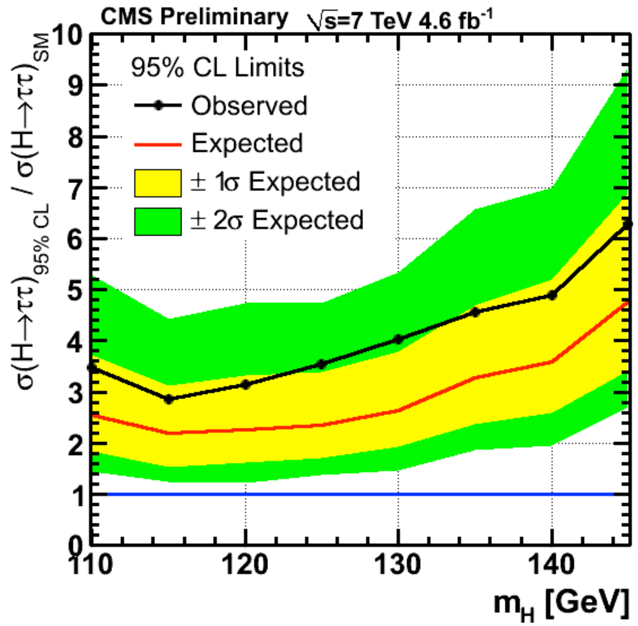


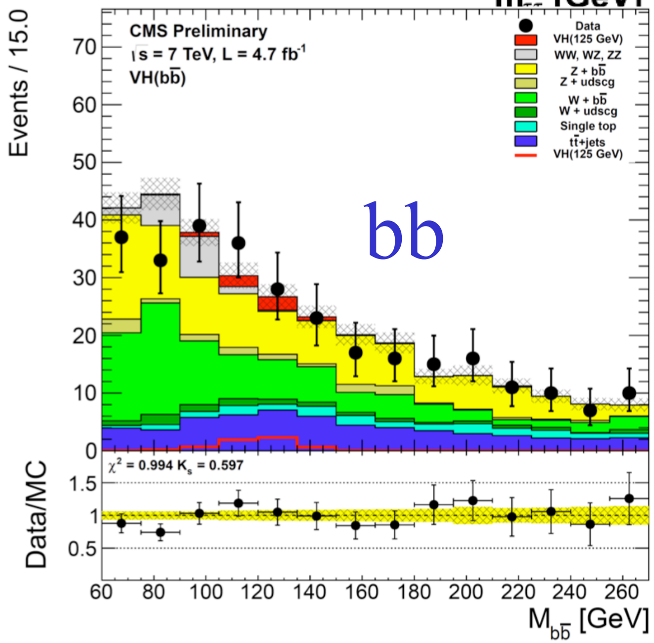


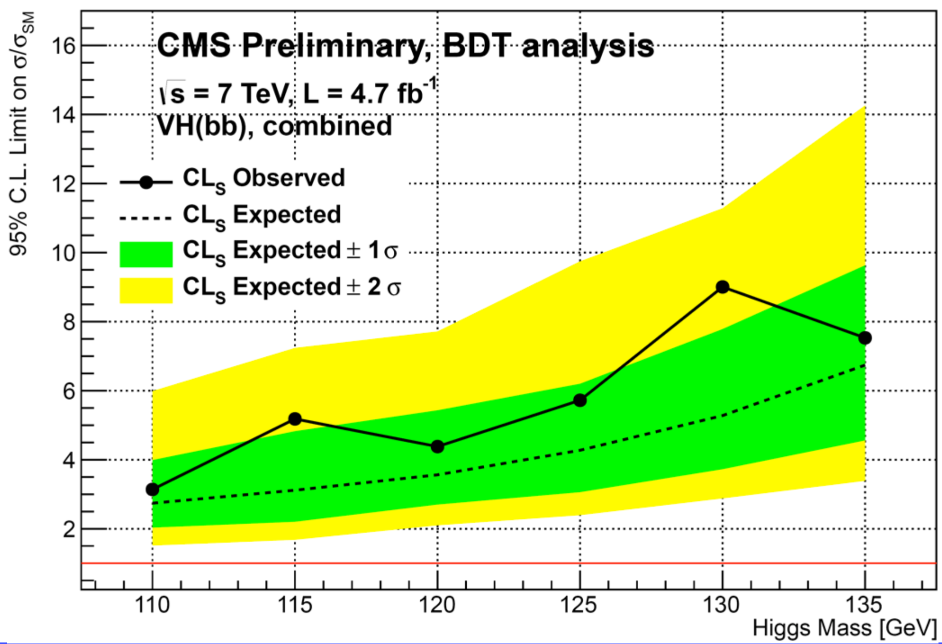


Low mass sensitive channels



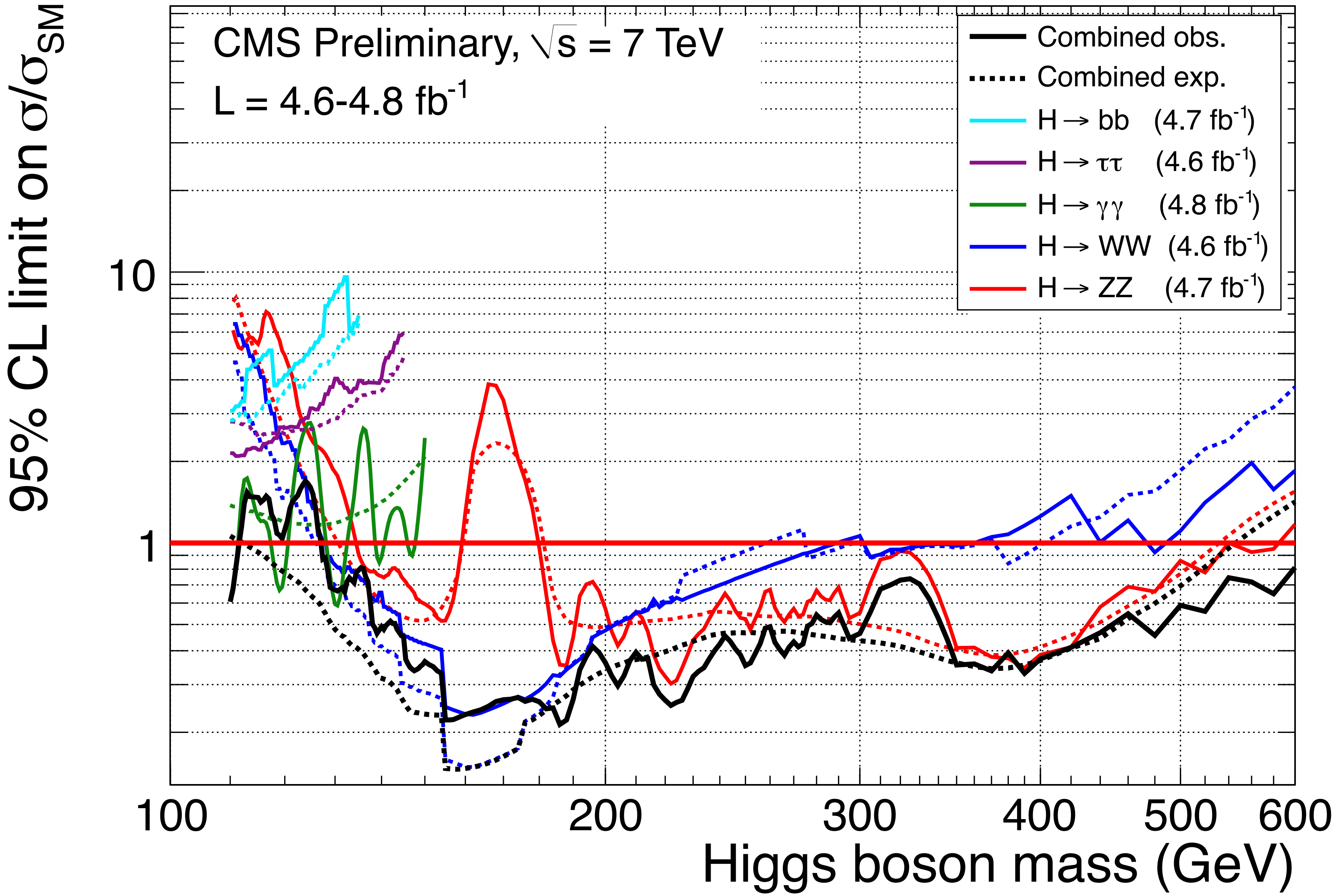






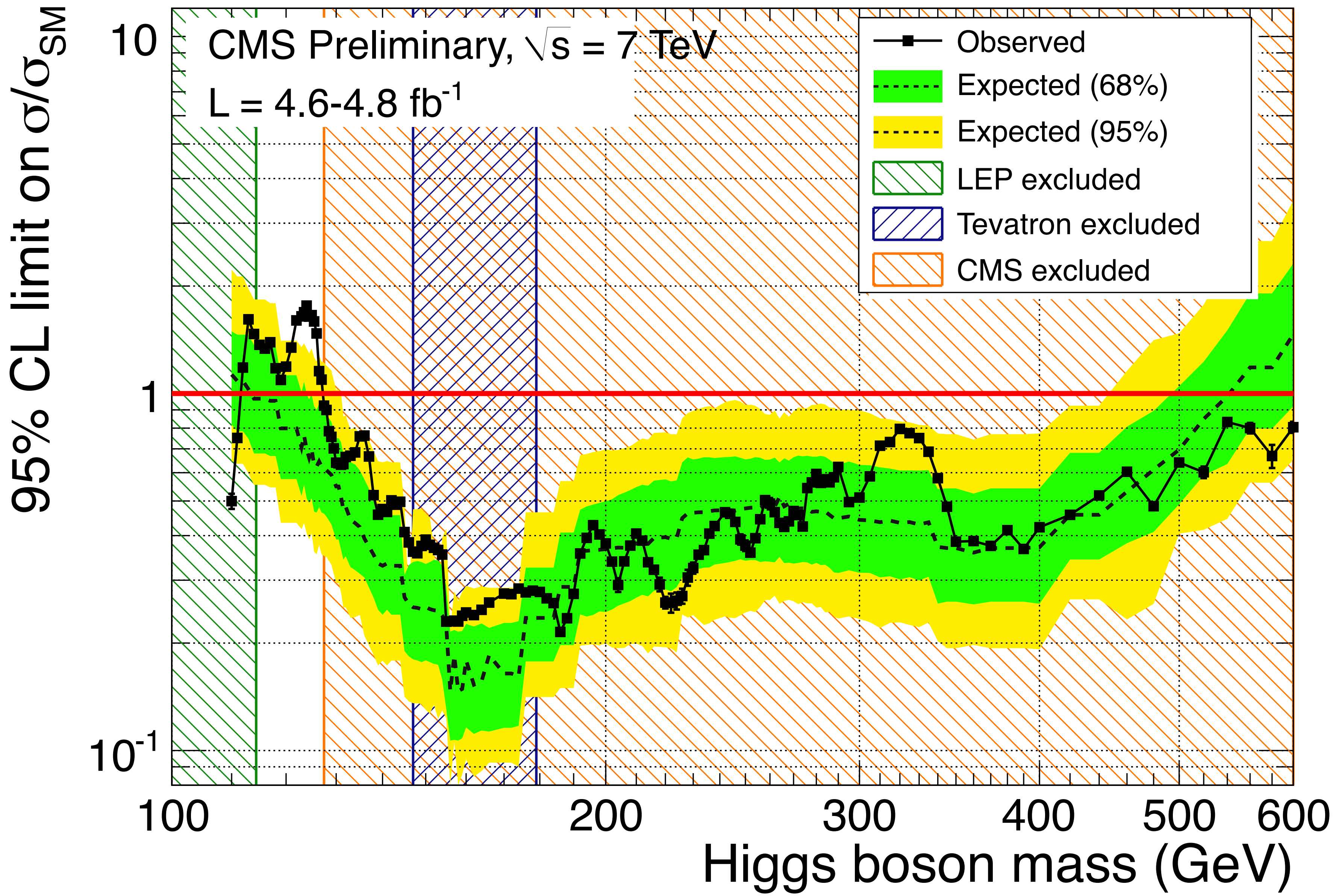
Combining limits to σ/σSM

Phys. Lett. B 710 (2012) 26-48, arXiv:1202.1488

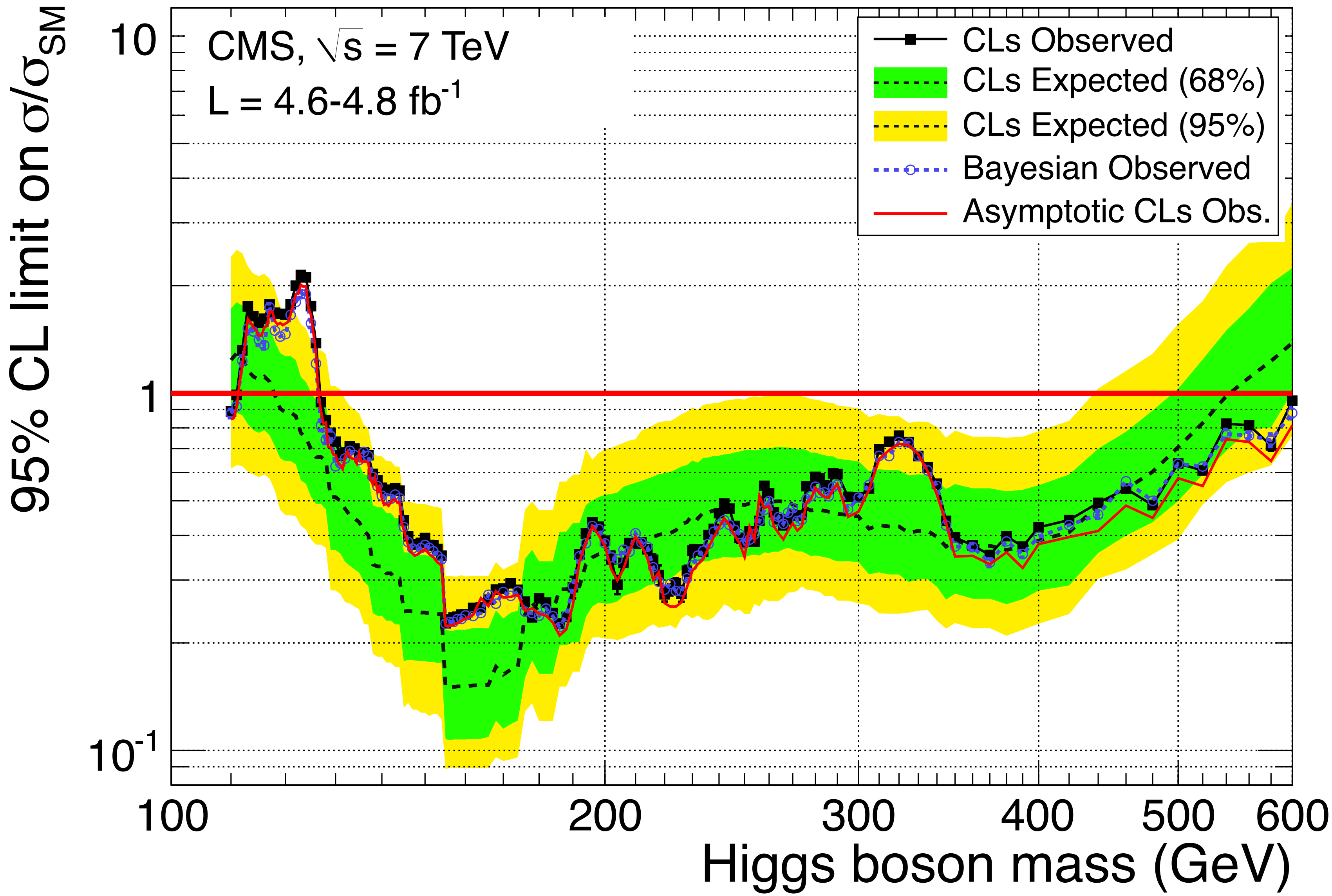


Excluded range: 127.5–600 GeV al 95% CL (expected: 114.5-543 GeV)

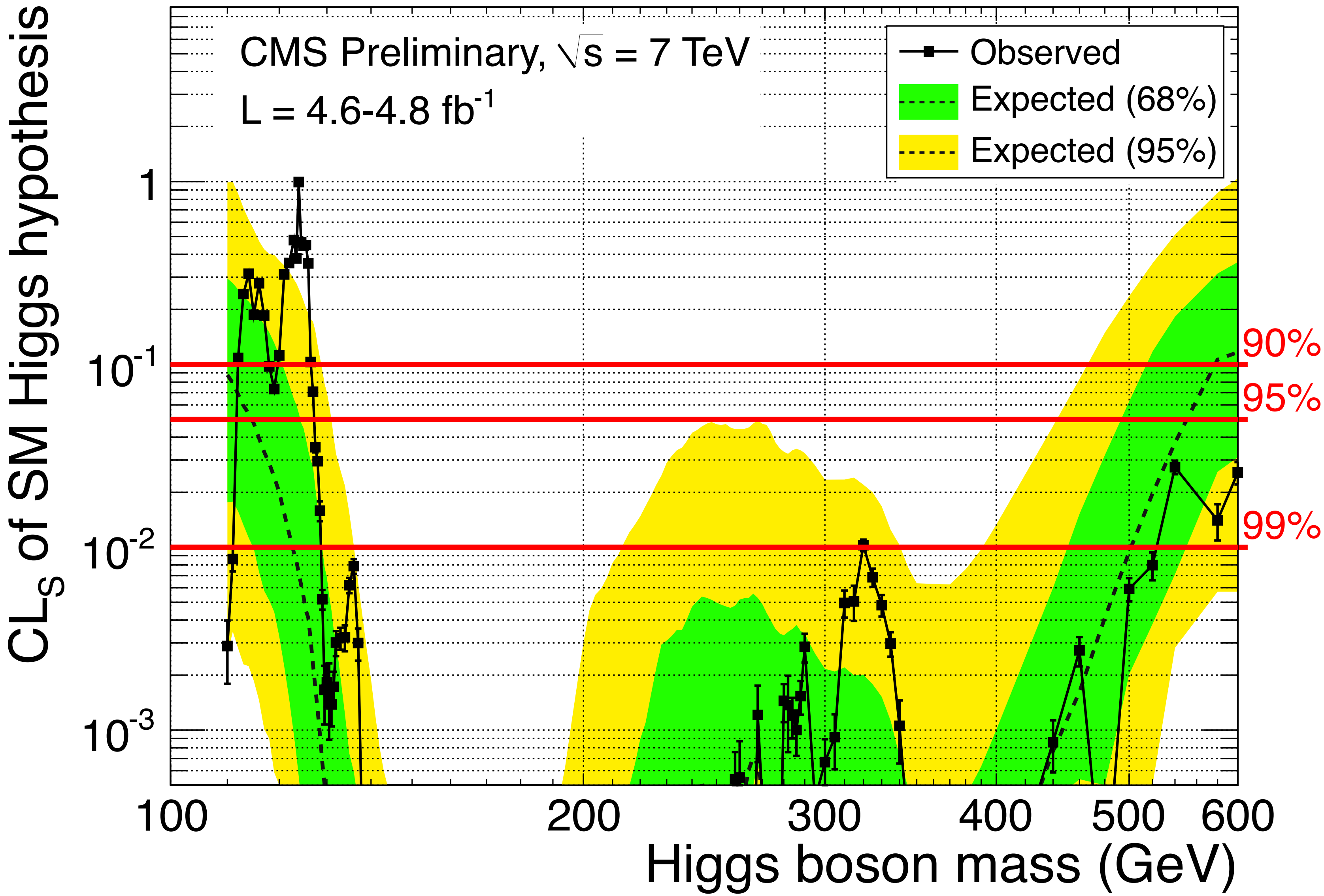
Exclusion plot at 95% CL



CL_s vs Bayesian and asymptotic



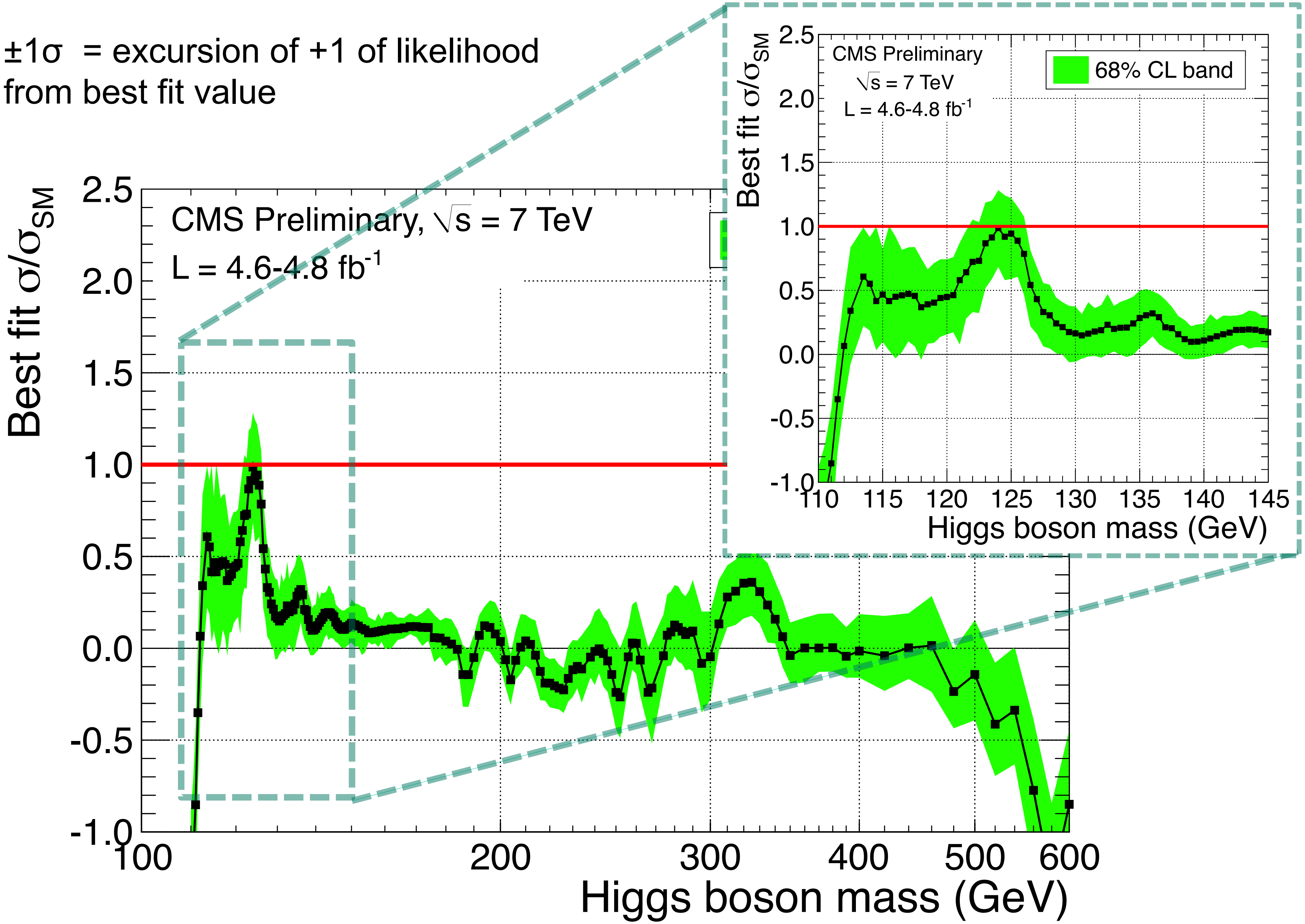
What if we use 99%?



Excluded range: 127.5–600 GeV at 95% CL, 129–525 GeV at 99% CL

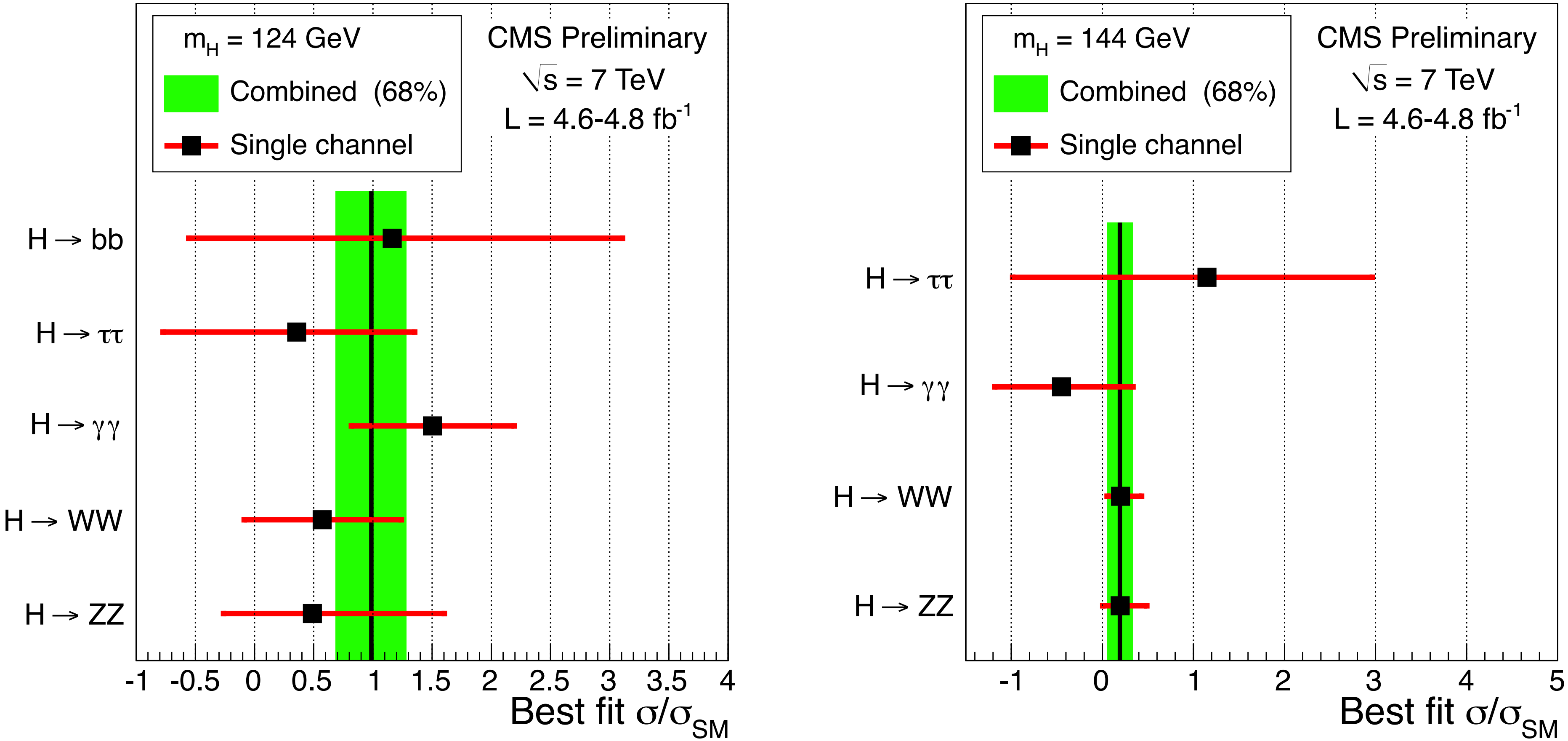
Cross section “measurement”

±1σ = excursion of +1 of likelihood from best fit value



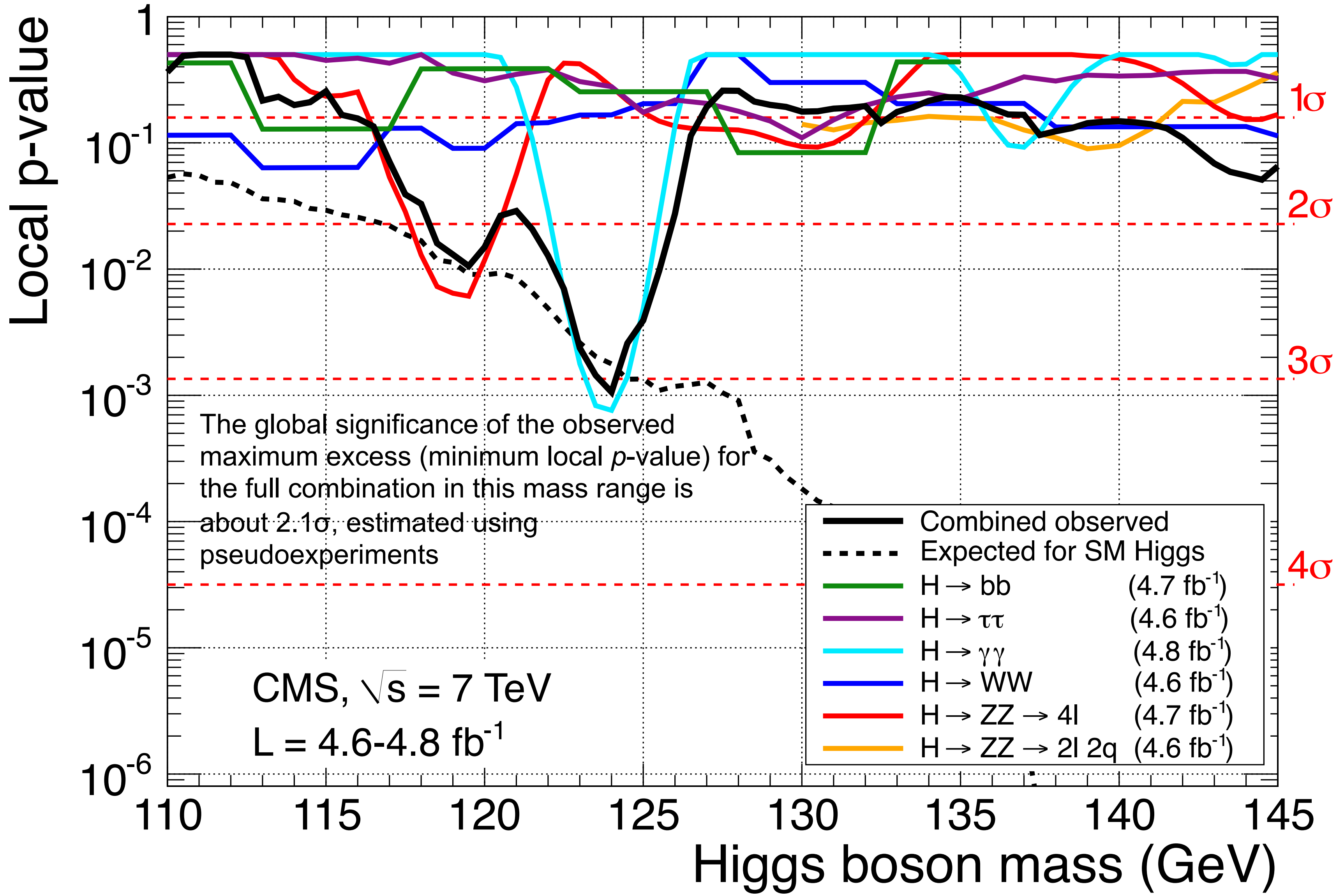
Comparing different channels

- Best fit to σ/σ_{SM} separately in various canals
- A modest excess is present consistently in all low-mass sensitive channels



"Hint" or fluctuation?

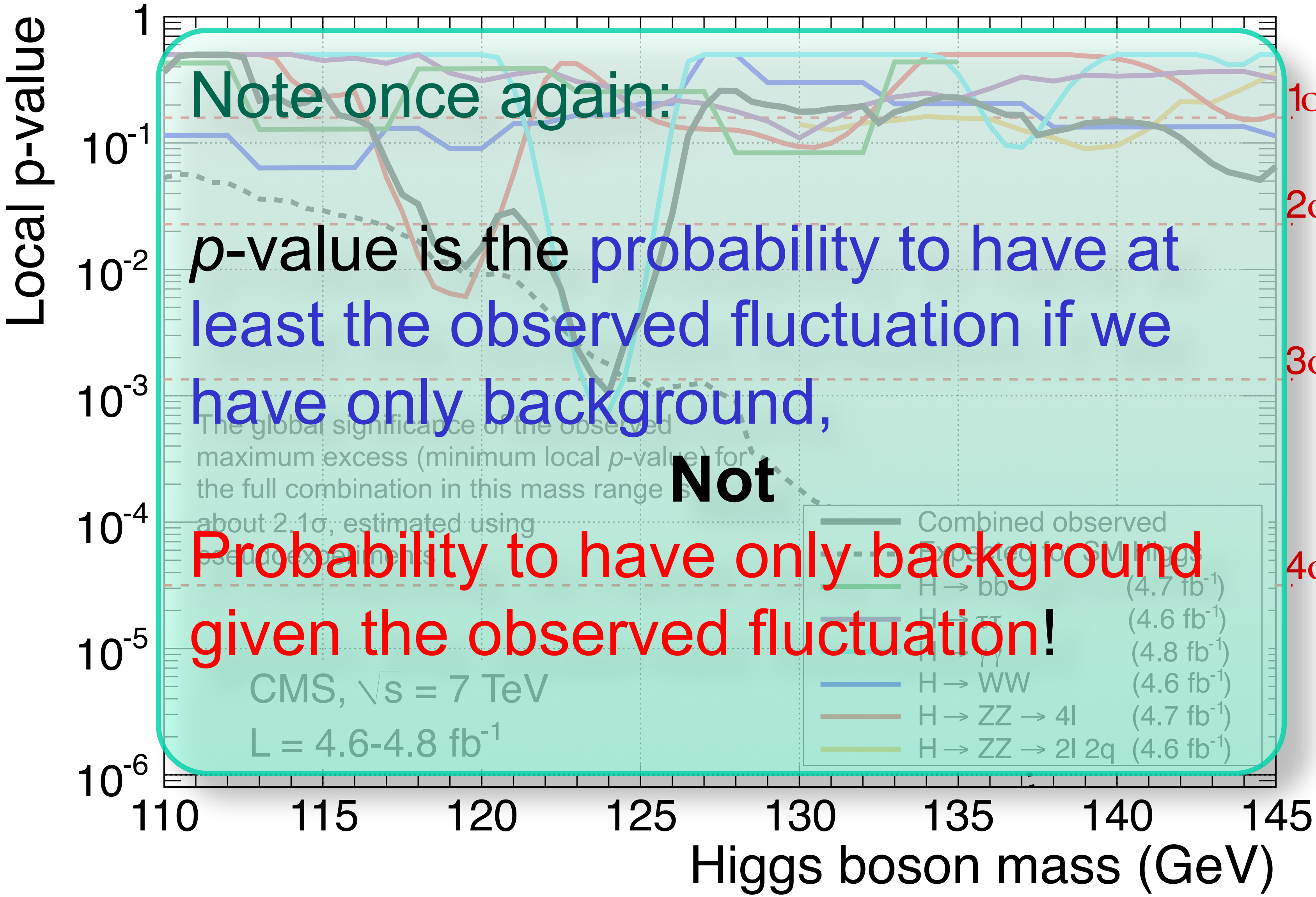
Probability of a bkg fluctuation ≥ than the observed one



$$p = \int_{n\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathrm{d}x = 1 - \frac{1}{2}\mathrm{erf}\left(\frac{n}{\sqrt{2}}\right)$$

"Hint" or fluctuation?

Probability of a bkg fluctuation ≥ than the observed one



Note once again:

p-value is the probability to have at least the observed fluctuation if we have only background,

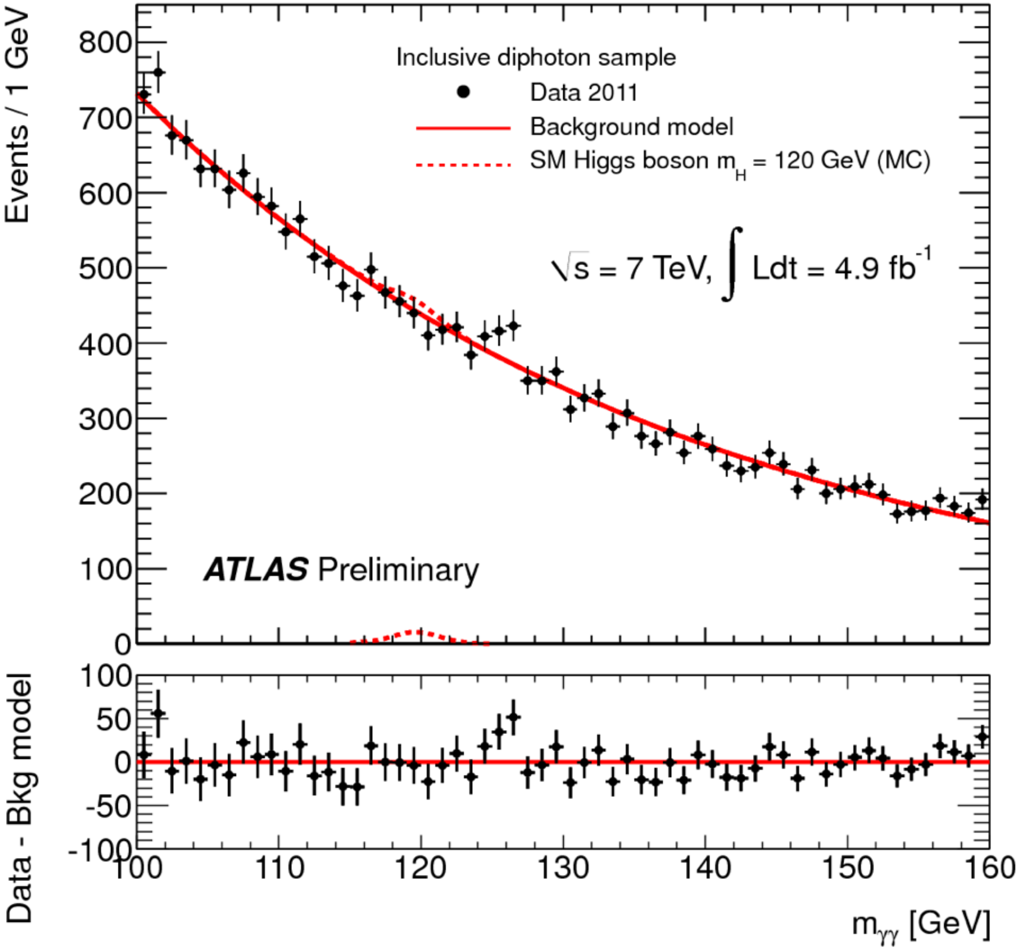
Not

Probability to have only background given the observed fluctuation!

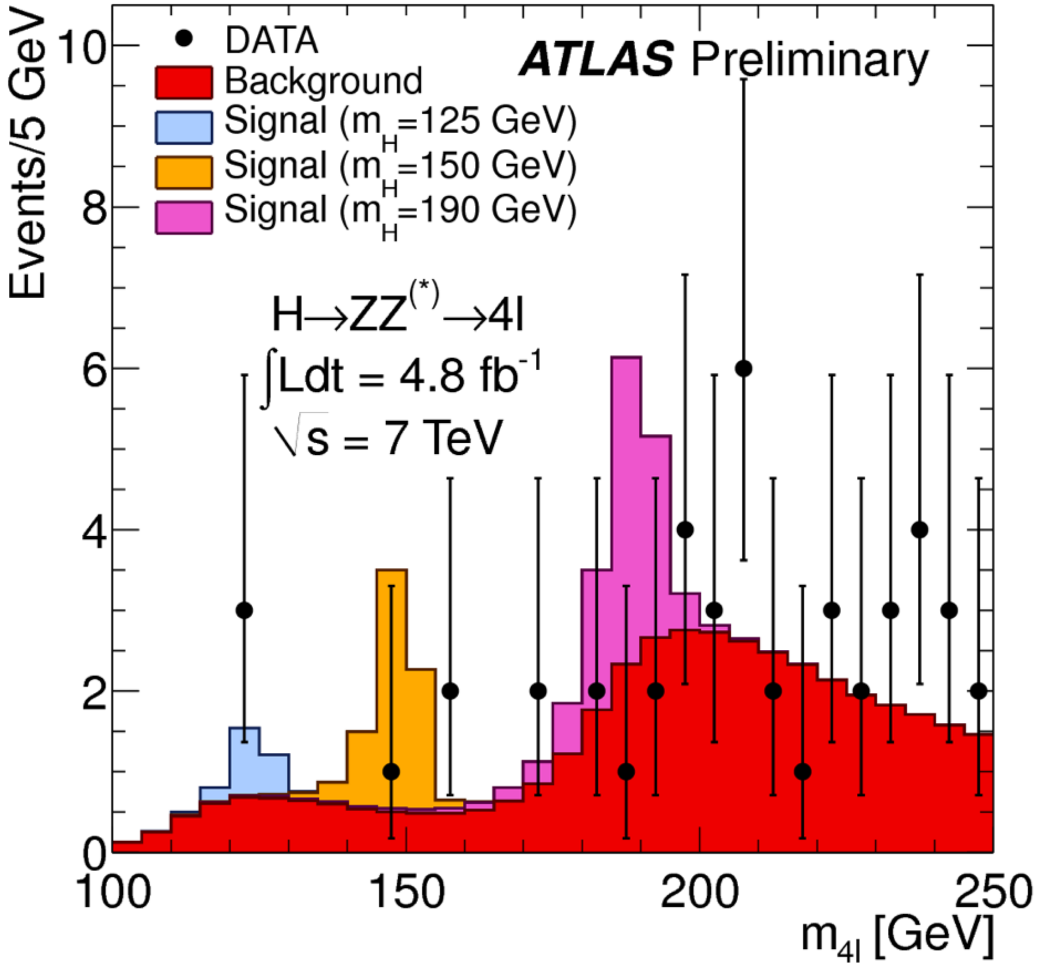
$$p = \int_{n\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathrm{d}x = 1 - \frac{1}{2}\mathrm{erf}\left(\frac{n}{\sqrt{2}}\right)$$

ATLAS: γγ, 4l

arXiv:1202.1414

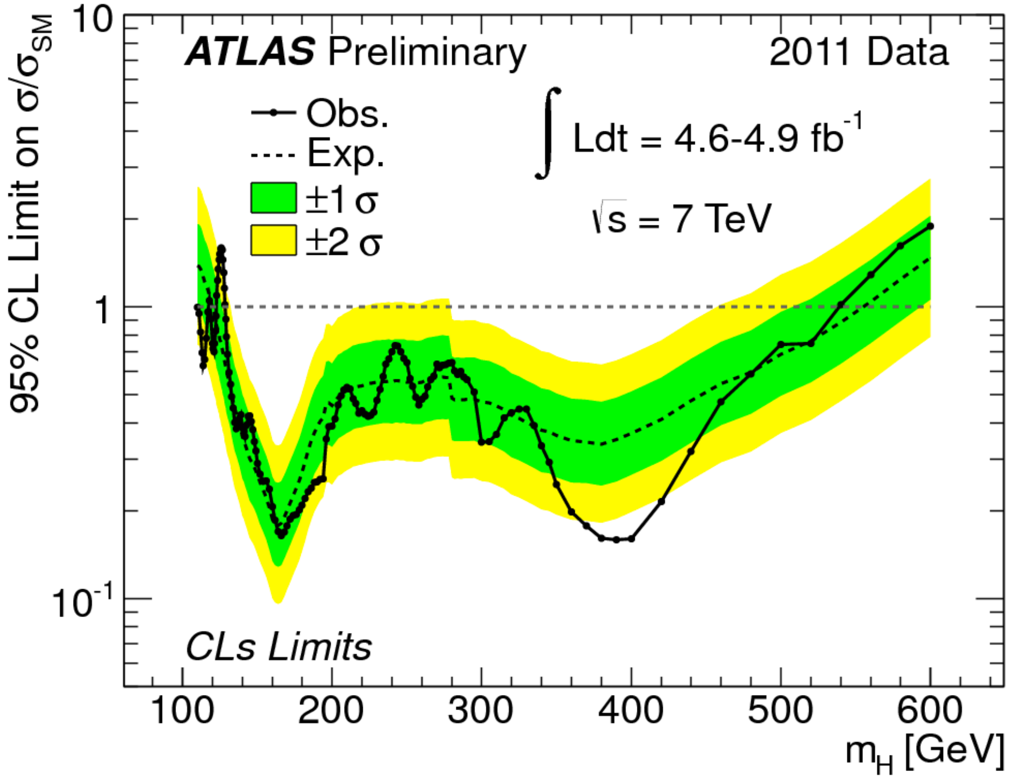


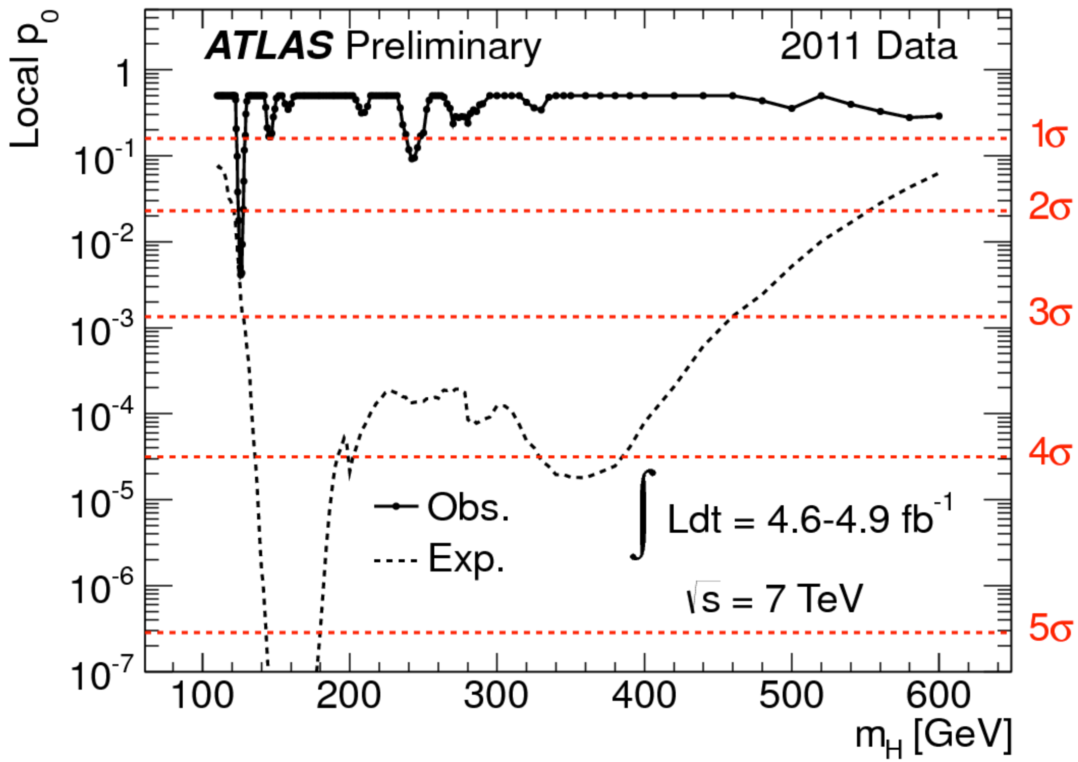
arXiv:1202.1415



ATLAS: combined limit

arXiv:1202.1408
ATLAS-CONF-2012-019

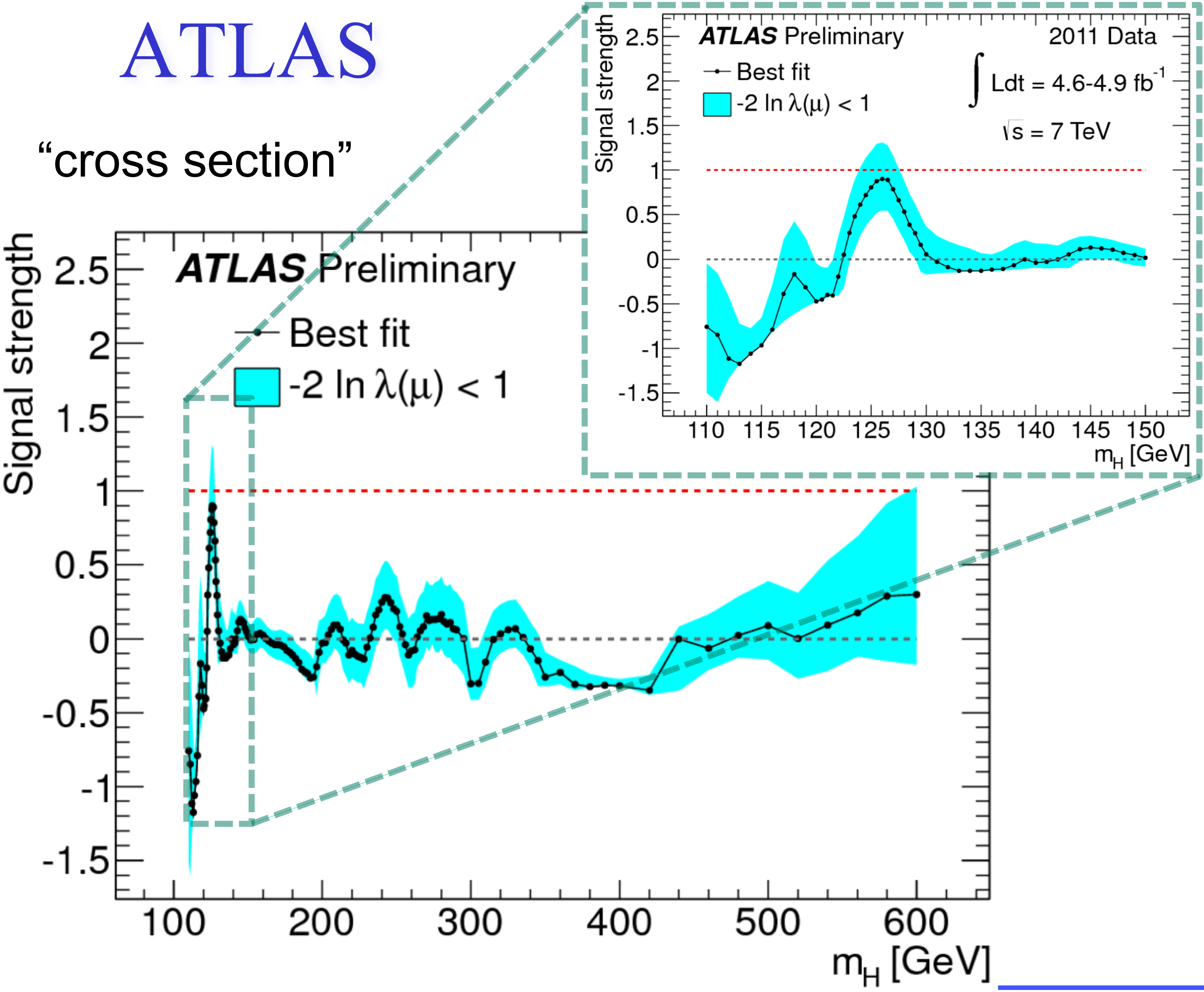




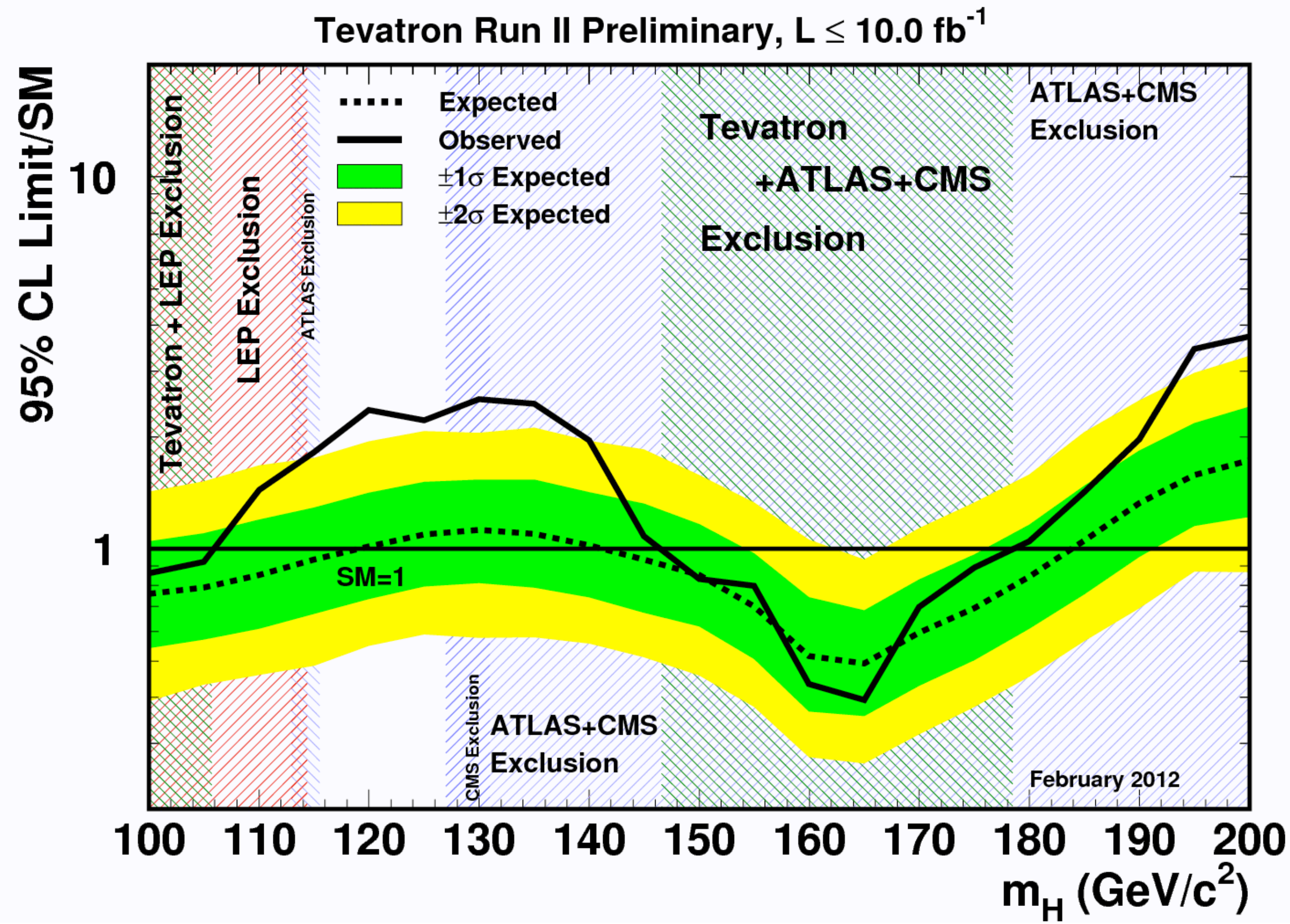
- Local sifnificance: 2.8σ (γγ), 2.1σ (ZZ*→4l), 1.4σ (WW*→lνlν)
- Global significance (LEE) 2.2σ (110-600 GeV)
- Excluded ranges: (95% CL):110–115.5 GeV, 118.5-122.5 GeV, 129–539GeV (expected: 120–550 GeV)

ATLAS

“cross section”



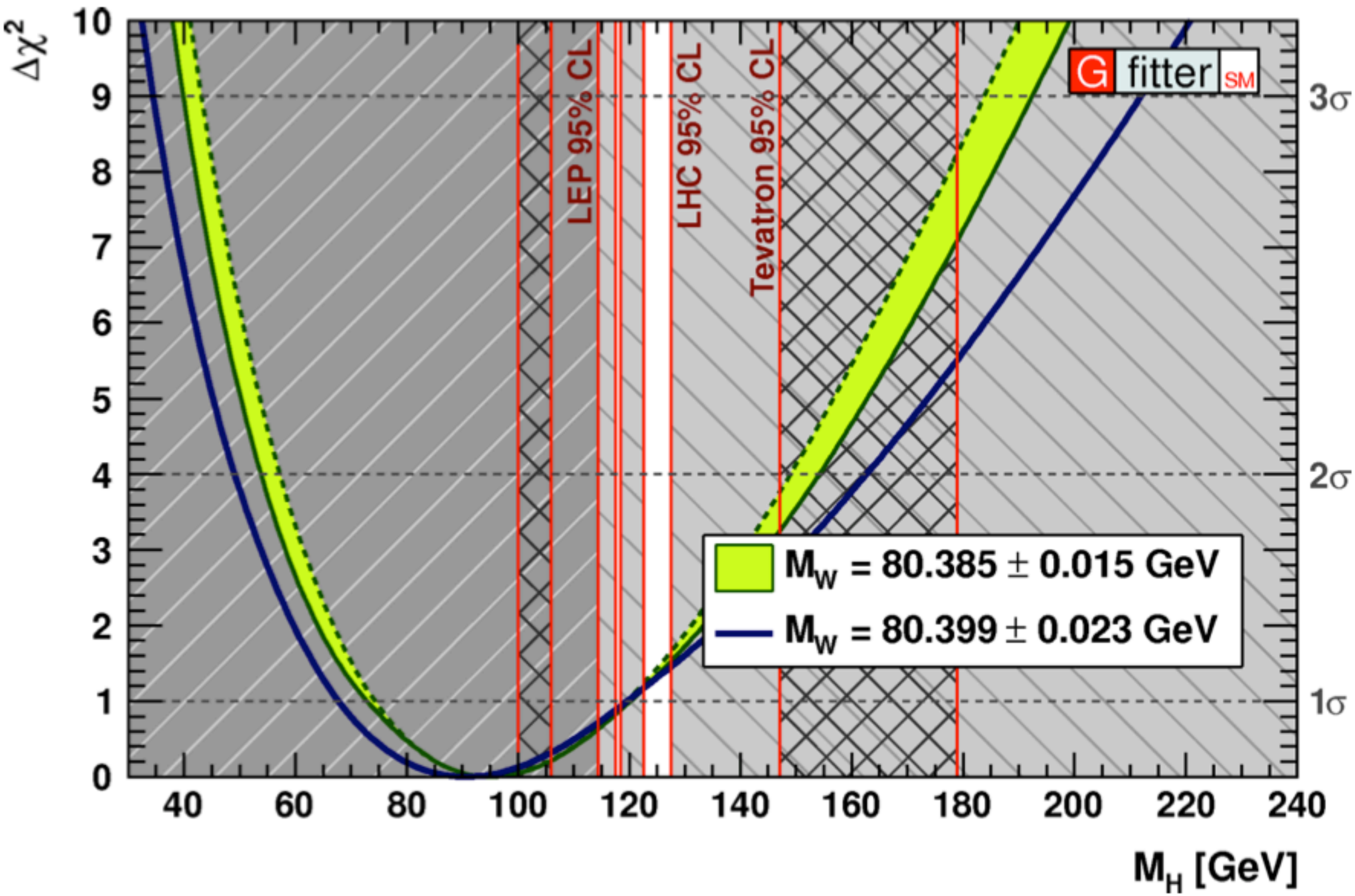
Latest from Tevatron



arXiv:1203.3774

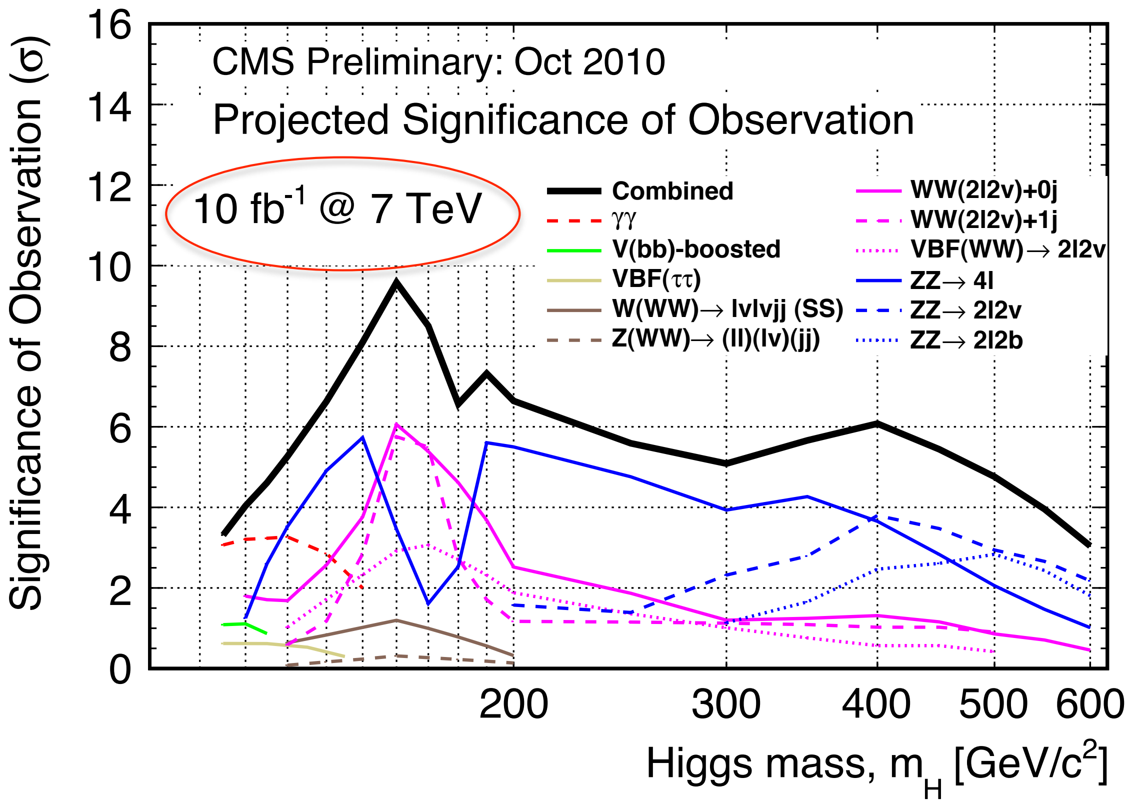
- Excluded ranges: 100-107 GeV, 147-179, expected: 100-119 GeV, 141-184 GeV
- Local significance (120 GeV): 2.7σ, global significance (LEE); 2.2σ

Latest SM fit



Perspectives for 2012

- LHC energy increased from 7 a 8 TeV. ATLAS and CMS are taking data now (+10% in cross section)
- In 2012 LHC should deliver about 4 times the 2011 integrated luminosity (~20fb^{-1})



- Higgs boson discovery or exclusion is very likely by 2012

In conclusion

- Many recipes and approaches available
- Bayesian and Frequentist approaches lead to similar results in the easiest cases, but may diverge in frontier cases
- Be ready to master both approaches!
- … and remember that Bayesian and Frequentist limits have very different meanings

- If you want your paper to be approved soon:
  - Be consistent with your assumptions
  - Understand the meaning of what you are computing
  - Try to adopt a popular and consolidated approach (even better, software tools, like RooStats), wherever possible
  - Debate your preferred statistical technique in a statistics forum, not a physics result publication!