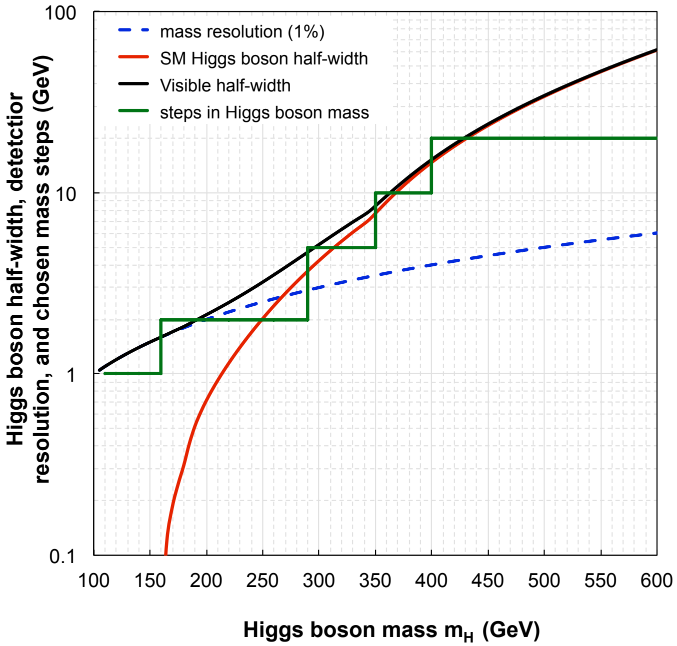
Combined results on SM Higgs Search with the CMS Detector

CMS Collaboration

Combination scope

channel	mass range (GeV/c^2)	luminosity (fb^{-1})	number of sub-channels	type of analysis
$H \to \gamma\gamma$	110-140	1.1	8	mass shape (unbinned)
$H \to \tau\tau$	110-140	1.1	6	mass shape (binned)
$H \to WW \to 2\ell 2\nu$	110-600	1.1	5	MVA (binned); cut&count
$H \to ZZ \to 4\ell$	110-600	1.1	3	mass shape (unbinned)
$H \to ZZ \to 2\ell 2\nu$	250-600	1.1	2	cut&count
$H \to ZZ \to 2\ell 2q$	226-600	1.0	6	mass shape (unbinned)
TOTAL (6)	**110-600**	**1.0-1.1**	**30**	

- **143 m_H-points for combination in the 110 – 600 GeV mass range**
- **6 analyses with a wide range of strategies**
- **30 exclusive sub-channels in combination**
- **142 nuisance parameters, majority of which affect more than one sub-channel**

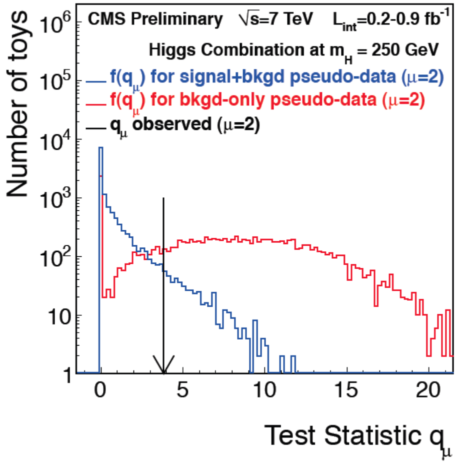


Two slides on statistics: limits

Pursue both frequentist and Bayesian paradigms, which allows us to validate robustness of results…

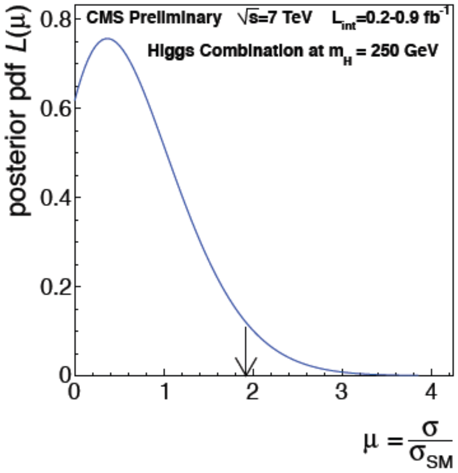
- **$\mathrm{CL_s}$ (exact formulation as agreed with ATLAS, details are in backup)**

$$\mathrm{CL_s} = \frac{P\left(q_\mu \geq q_\mu^{obs} \,|\, \mu s(\hat{\theta}_\mu^{obs}) + b(\hat{\theta}_\mu^{obs})\right)}{P\left(q_\mu \geq q_\mu^{obs} \,|\, b(\hat{\theta}_0^{obs})\right)} = 0.05$$



- **Bayesian (with flat prior on signal strength)**

$$\int_0^{\mu_{95\%CL}} p(\mu\,|\,\mathrm{data})\, d\mu \;=\; 0.95$$



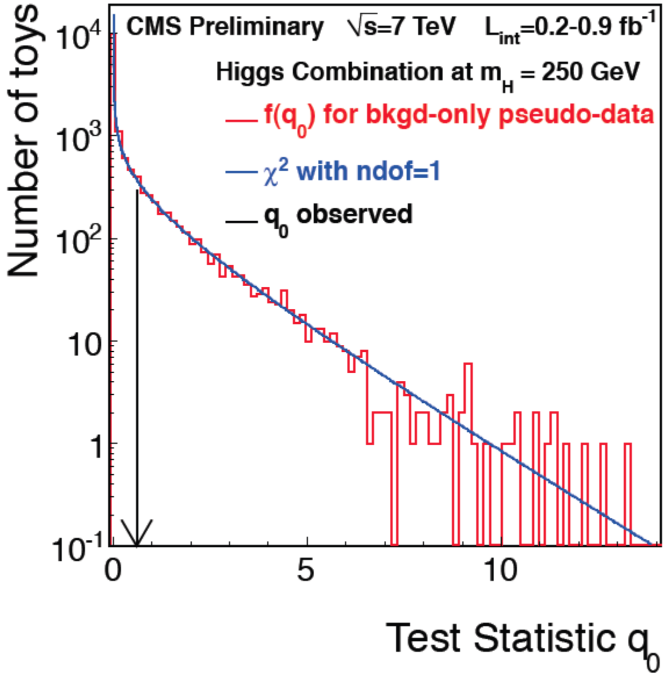
Two slides on statistics: excess

test statistic – profile likelihood

$$q_0 \;=\; -2\ln\frac{\mathcal{L}(\mathrm{data}|0,\hat{\theta}_0)}{\mathcal{L}(\mathrm{data}|\hat{\mu},\hat{\theta})} \qquad \text{and } \hat{\mu} \geq 0$$

"local" p-values – from asymptotic approximation

$$\tilde{p} = \frac{1}{2}\left[1 - \mathrm{erf}\left(\sqrt{q_0^{\mathrm{obs}}/2}\right)\right]$$



IMPORTANT:

- **small "local" p-value means one has a local excess w.r.t. expectations**
- **it does not tell us whether the excess is due to a signal or not**
- **nor does it tell us whether the excess is consistent with THE expected signal**
- **moreover, one must be ware of a potentially large look-elsewhere effect (LEE)** that can considerably de-rate "significance" of the minimal p-value found in a search involving scans over a broad phase space with a good "local" resolution

Preamble for combination

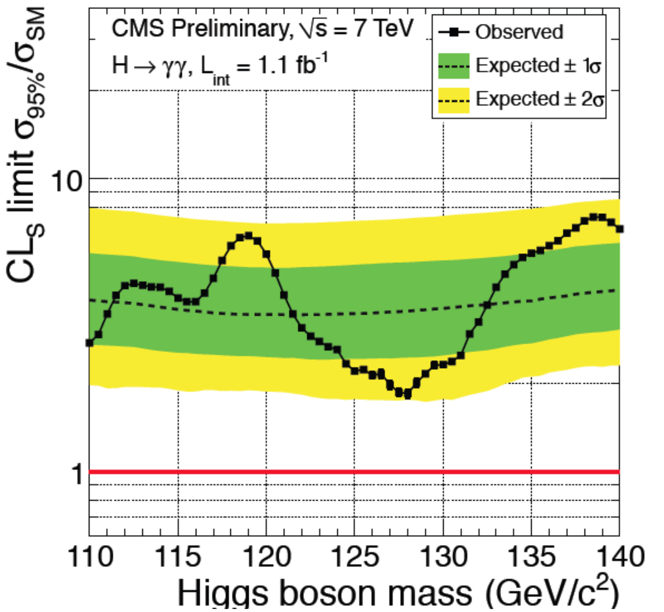
- Summary of results obtained by six analyses entering the combination
- Features seen in the individual analyses will manifest themselves in the combination

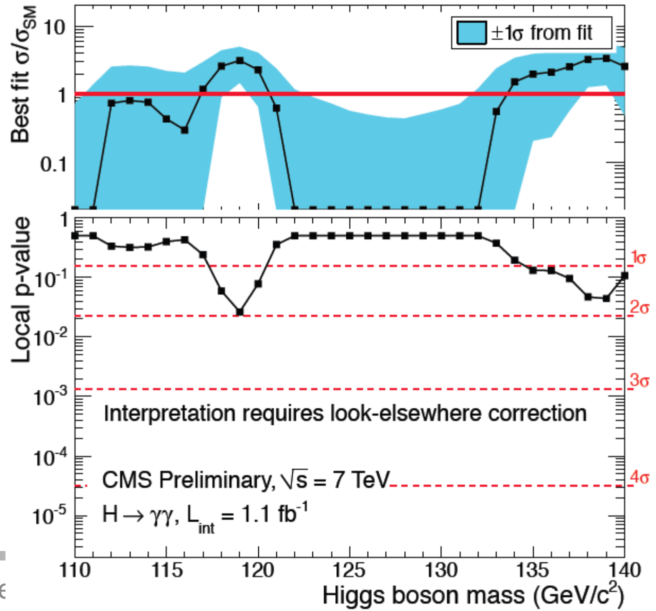
Low mass range: H → $\gamma\gamma$

- **unbinned $m_{\gamma\gamma}$ distributions in 8 event categories**

- **observed exclusion: 2-7 × σ_{SM}**
  - **variations are within ±2σ statistical bands**
  - **correlation "length" agrees with the instrumental mass resolution**

- **no significant excess of events:**
  - two bumps with local p-values 3-4% (<2σ)
  - LEE: probability to observe a 2σ-excess for background-only hypothesis is ~60%
  - two bumps would require ~ $3\times\sigma_{SM}$ cross section



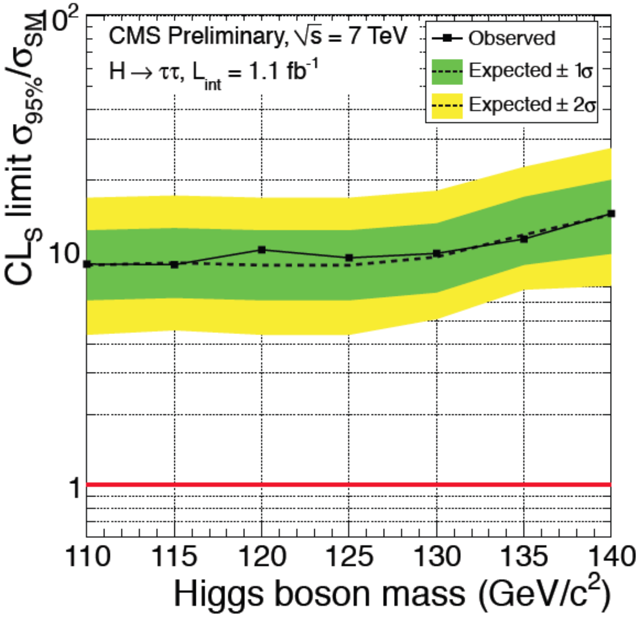


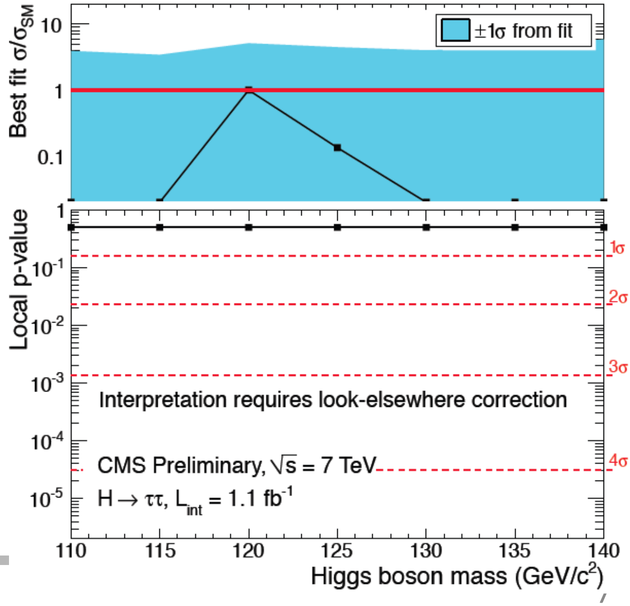
Low mass range: H → ττ

- binned m_{vis} distributions in 6 exclusive final states

- observed exclusion: ~10 × σ_{SM}
  - observed ≈ expected
  - shape is rather featureless, due to the broad $m_{\tau\tau}$ resolution

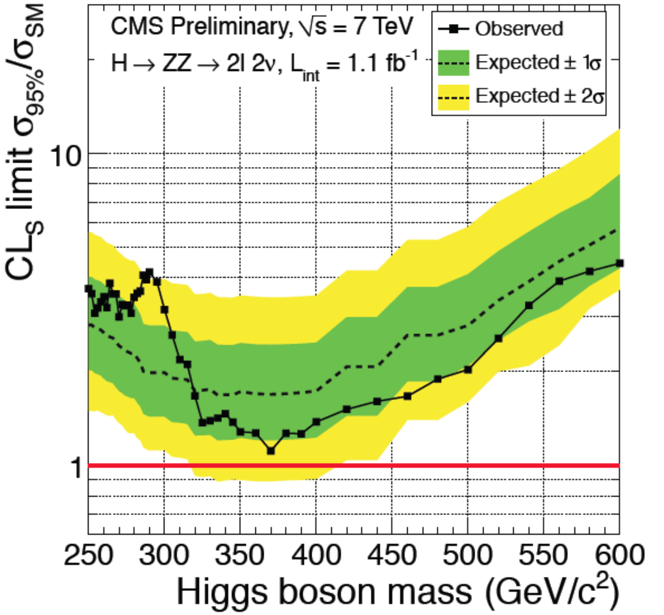
- no significant excess:
  - LEE trial factor ~ 1

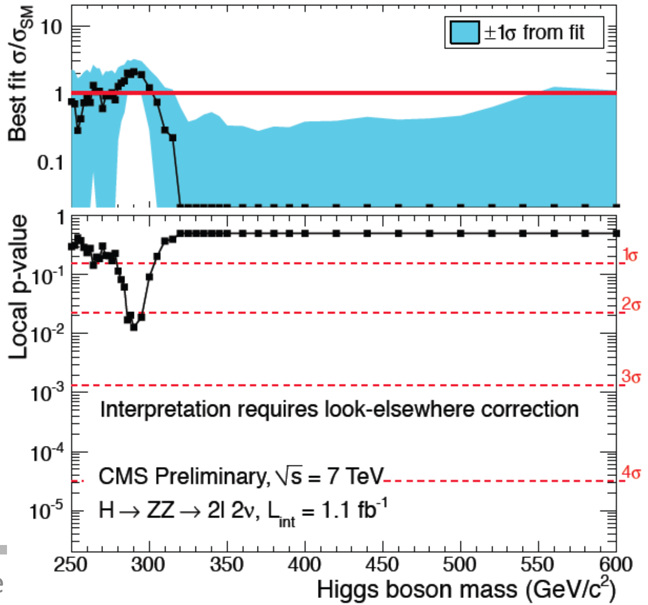




High mass range: H → ZZ → 2l2ν

- **cut-and-count with sliding cuts(m_H) in two event categories**
- **observed exclusion: 1-4 × σ_{SM}**
  - **variations are within ±2σ statistical bands**
  - **correlation "length" agrees with the "effective" m_H mass window size of 50 (200) GeV at low (high) Higgs mass**
- **no significant excess of events:**
  - one bump at m_H~290 with local p-values ~1%
  - the bump would require ~ 2×σ_{SM} cross section



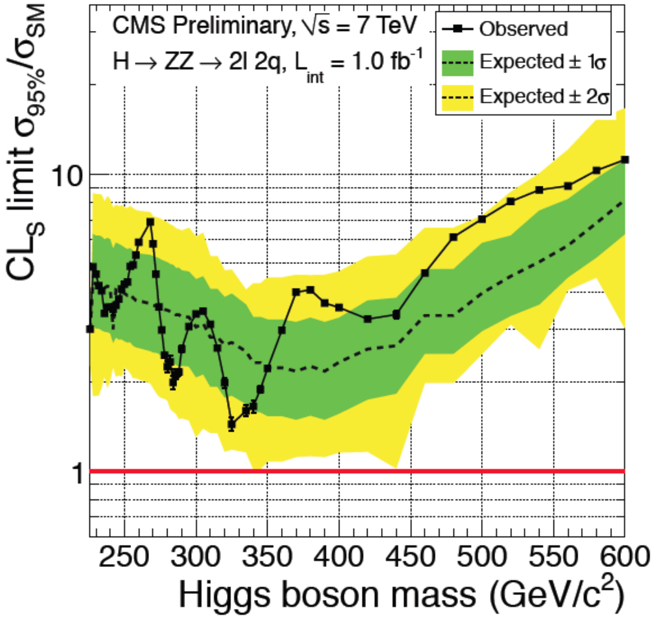


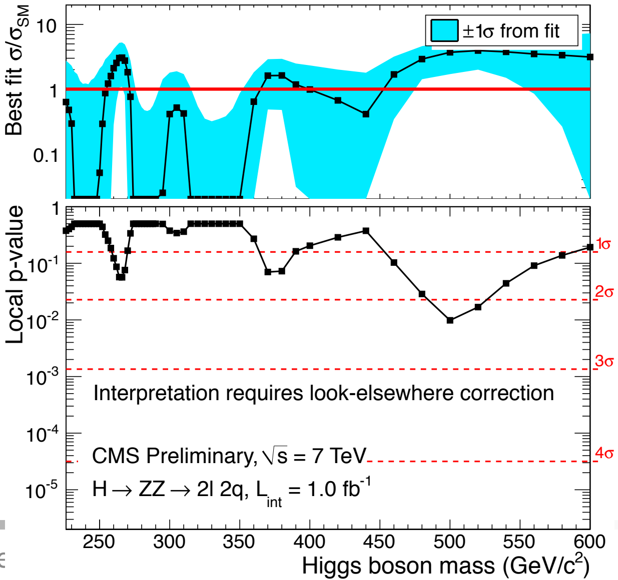
High mass range: H → ZZ →2l2q

- **unbinned m_{2l2j} analysis in six event categories**

- **observed exclusion: $2\text{-}10 \times \sigma_{SM}$**
  - **variations are within ±2σ statistical bands**
  - **correlation "length" agrees with the "effective" m_{ZZ} peak width of 3% (6%) at 250 (500) GeV Higgs mass**

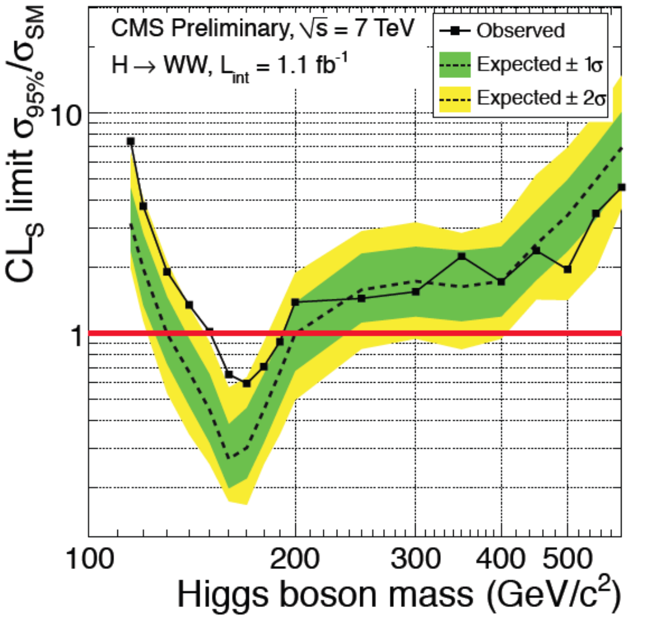
- **no significant excess of events:**
  - three bumps with local p-values ~1-5%
  - the two smallest p-value bumps would require ~ $4\times\sigma_{SM}$ cross section

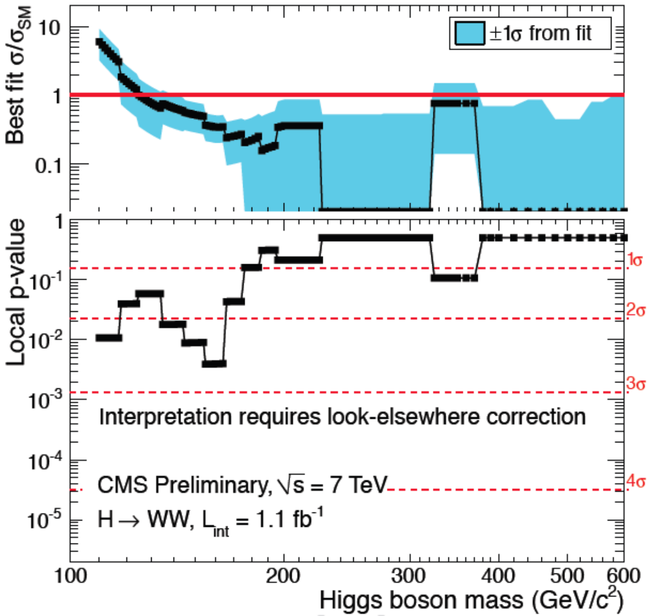




Full mass range: $H \rightarrow WW \rightarrow 2l2\nu$

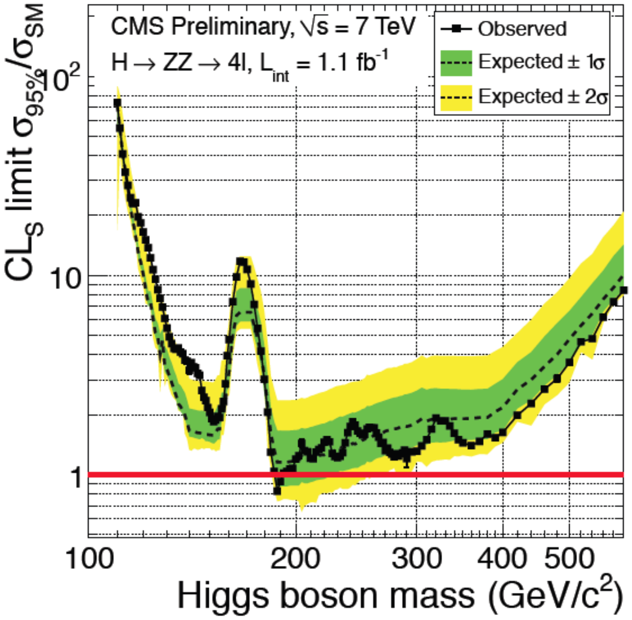
- binned MVA-output shape analysis in 4 event categories + 1 cut-and-count analysis
- very poor mass resolution
- observed exclusion: m_H=150-193 GeV for SM H
  - at low masses <200 GeV, the limits are not as strong as expected for bkgd-only and show a broad +2σ deviation
  - at high masses >200 GeV, observation ≈ expectation
- correlation "length" at low m_H is ±30 GeV
  - deviations, whether due to bkgd fluctuations or signal, will always appear as flattish shifts up/down
- observed some excess of events:
  - broad ~2σ excess at low masses
  - LEE: ~O(3); hence, the excess is approximately as unlikely as it appears (a few %, can happen)
  - **The bets-fit values of σ/σ_{SM} disfavor a SM Higgs signal explanation in the < 120 and > 150 GeV/c^2**

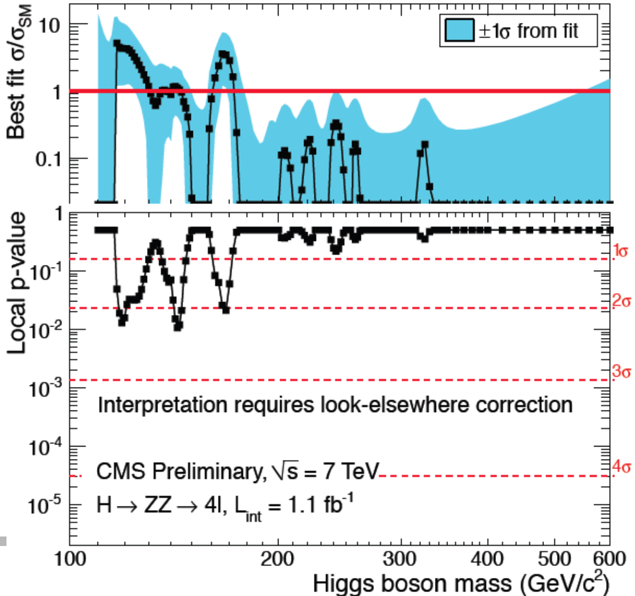




Full mass range: H → ZZ → 4l

- **unbinned m_{4l} analysis in three event categories**
- **observed exclusion: 1-100 × σ_{SM}**
  - **variations are within ±2σ statistical bands**
  - **correlation "length" agrees with the "effective" m_{ZZ} peak width**
- **observed some excess of events**
  - Wiggles in p-value follow individual events; p-value ~ 0.01
  - LEE~O(100) and washes out significance of excesses
  - Two pairs of events at m_H~120 and ~160 GeV would imply too large signal CS, one pair around m_H~140 GeV would not be inconsistent with a signal—but the statistical precision of these assessments is very poor.





Combination

- Six analyses:
  - only one reached exclusion sensitivity [130-200]
  - all analyses have up and down wiggles vs m_H
- Next step is combination…

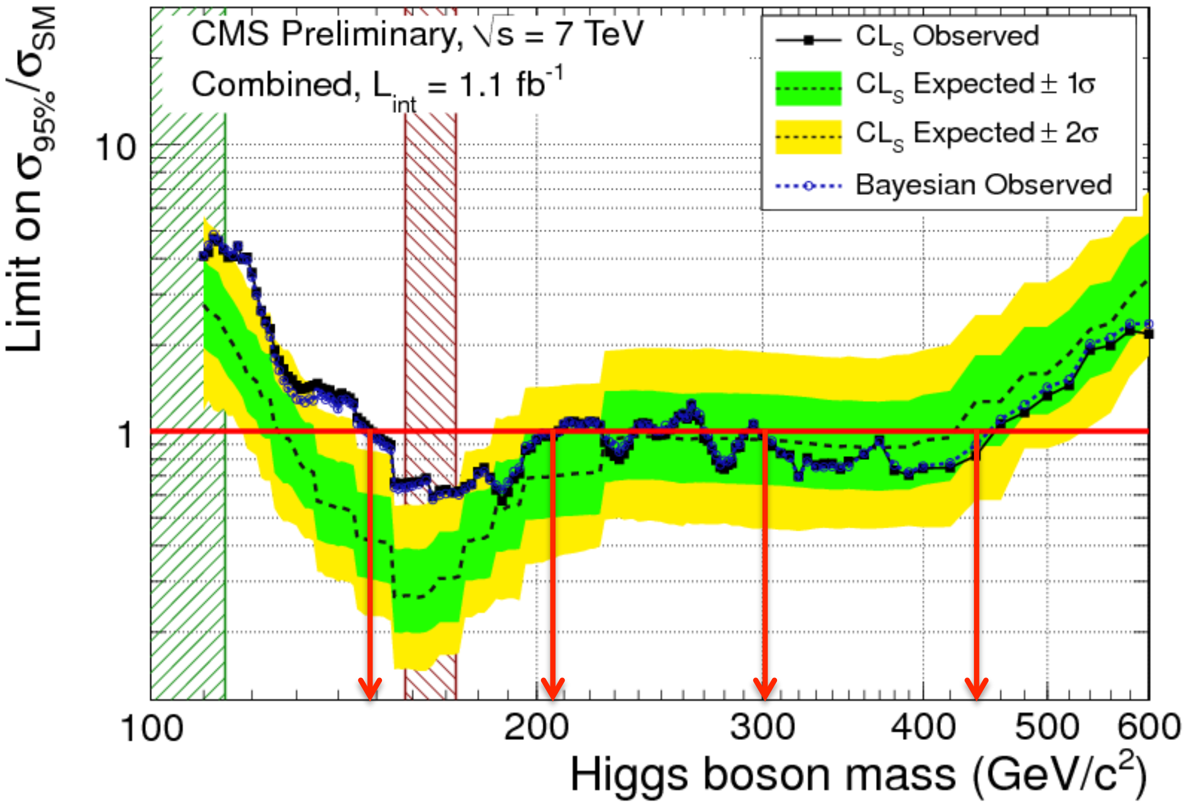
Combination: 95% CL limit on σ/σ_{SM}

Expected (GeV)

[127-420]

Excluded (GeV)

[149-206] … [300-440]
and 3 short segments
in between



- At low mass, the excess is driven by H→WW, with a little boost from H→ZZ→4l
- Remarkable agreement between CL$_s$ and Bayesian approaches: 0.3±4.6%
- Step-like structure is due to m_H-discreteness in the H→WW analysis
- The range [200-300] GeV is just below the line of expected exclusion at 95% C.L.

Combination: CL_s value

Excluded at 95% CL (GeV)

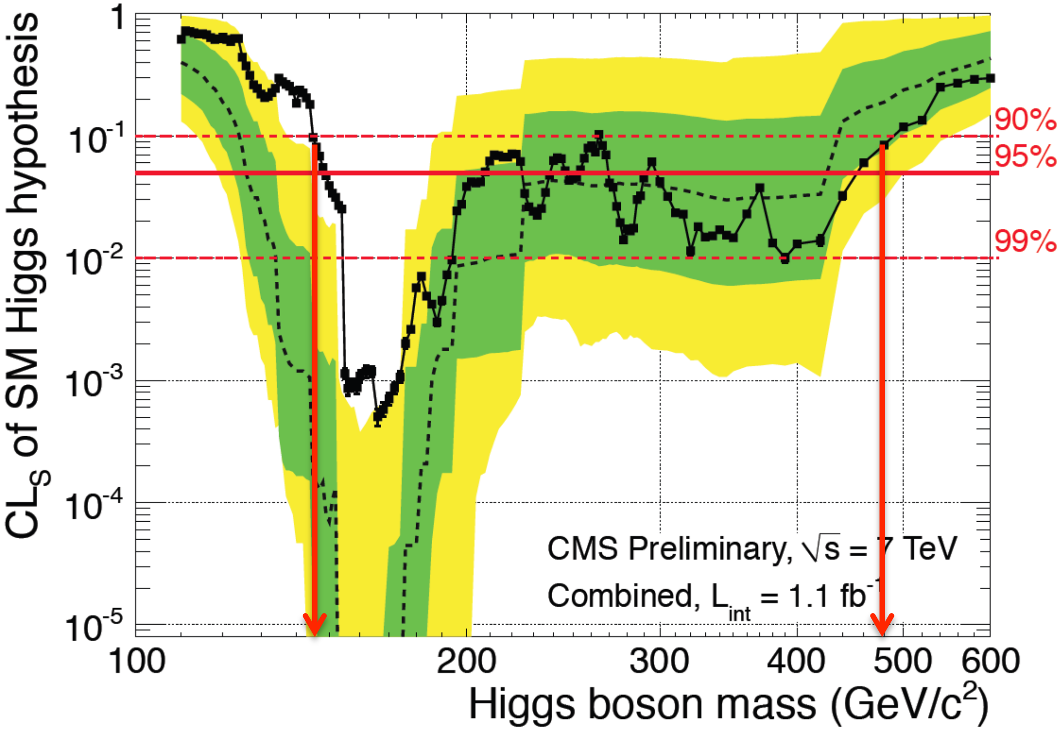
[149-206] … [300-440]

and 3 short segments

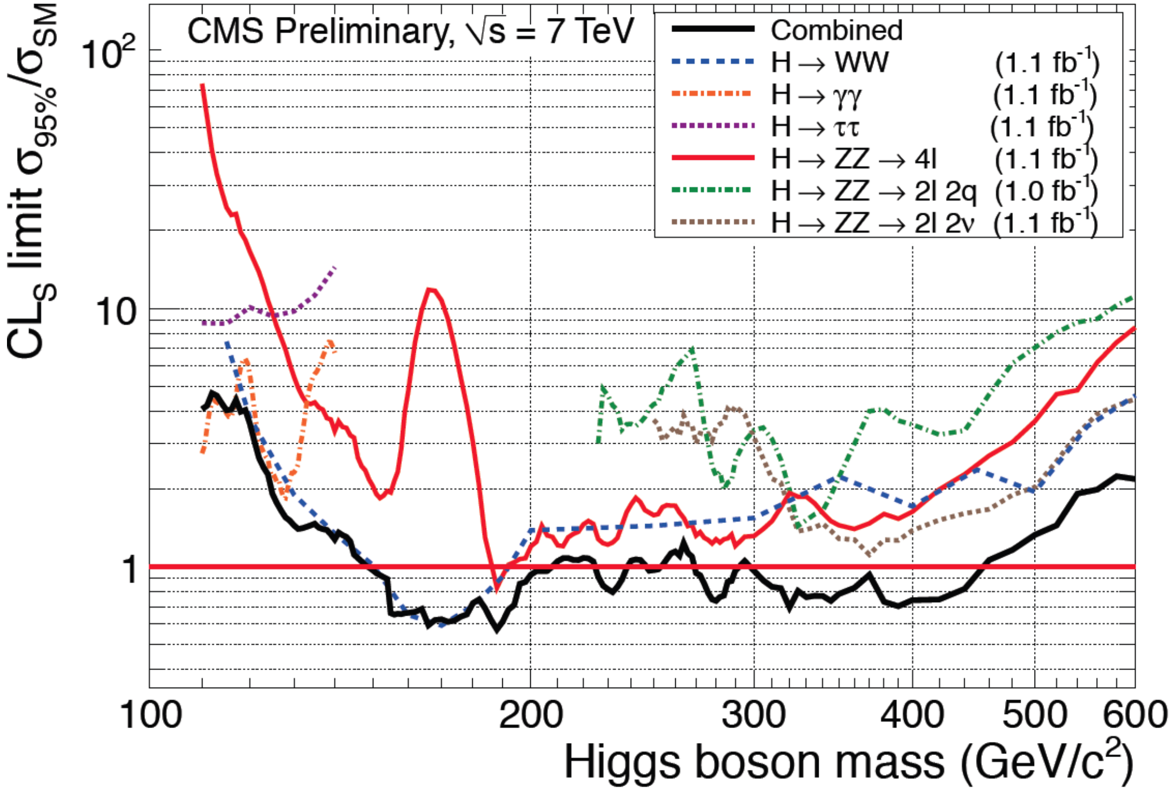
in between

Excluded at 90% CL

[145-480] GeV

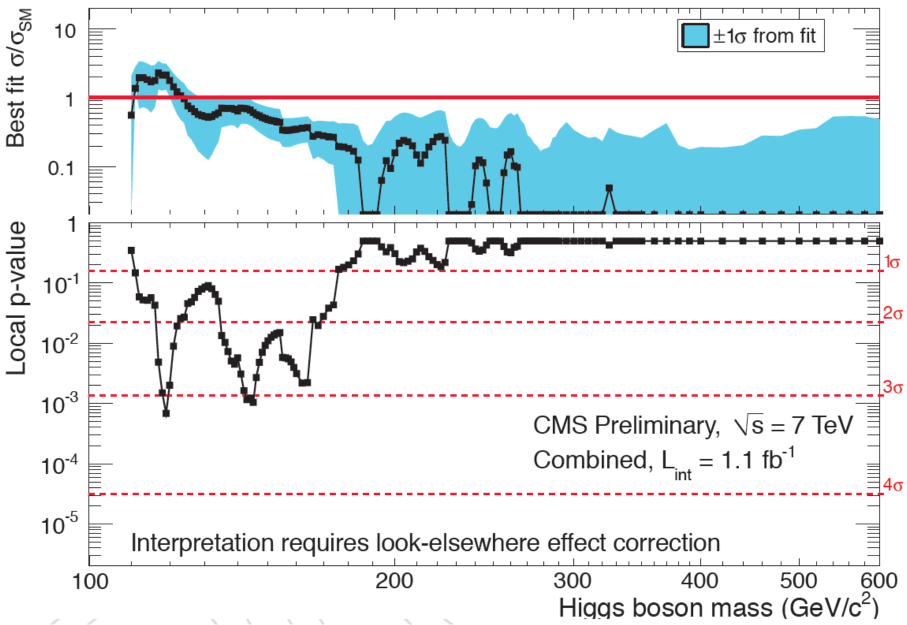


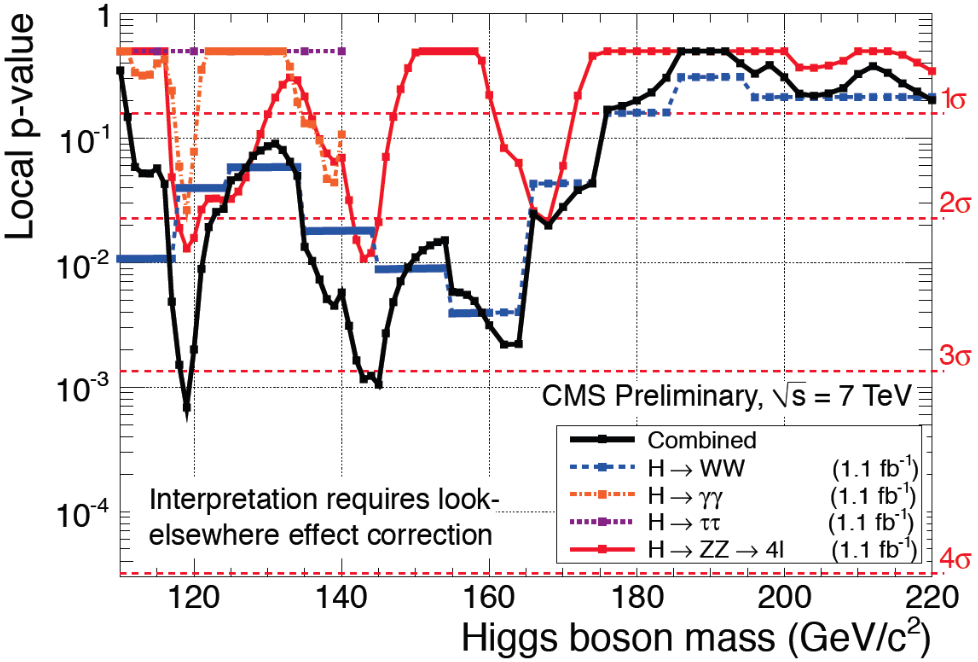
Combination: interplay of six analyses



- **At high masses, the combination gives a large gain over all individual analyses**
- **At very low mass, excess in the H→WW analysis makes combination equal or even more conservative than the H→γγ search would imply on its own**

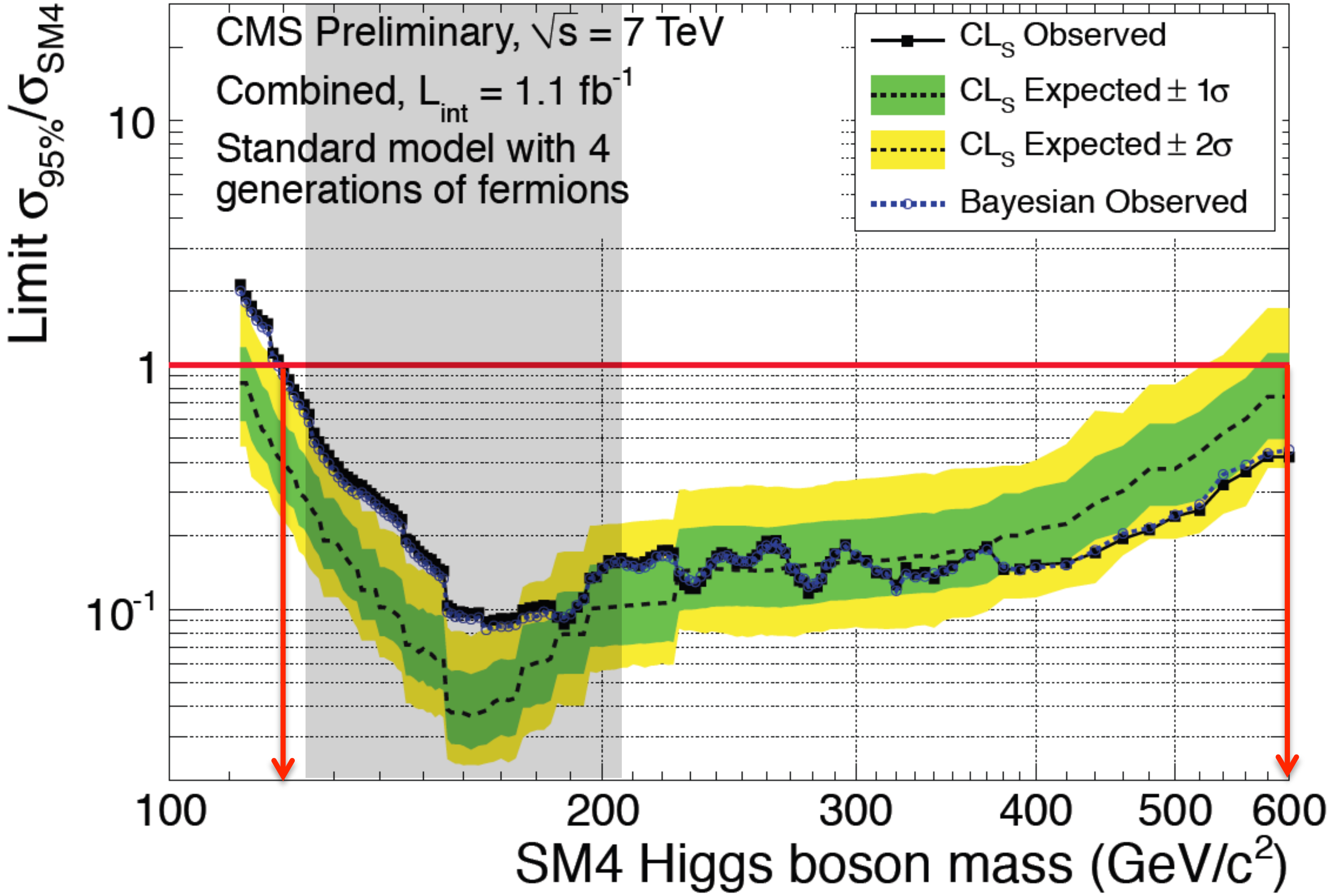
Local p-value and best-fit σ/σ_{SM}





- **RECALL: Small p-value means an excess. It does NOT tell us whether this is a signal or not, NOR does it say if the excess is consistent with the expected signal**
- **The origin of the overall structure:**
  - **the broad ~2σ excess comes mostly from H→WW**
  - **H→ZZ→4l and H→γγ , subject to large LEE, add some structure on top of the H**
- **LEE effect for the combination is not yet determined; individual channels entering the combination have LEE from O(1) to O(100)**

Reinterpretation for SM4 Higgs



If there are 4 fermion generations, the SM4 Higgs boson is excluded in the 120-600 GeV mass range at 95% CL

Combination summary

- **SM Higgs boson**
  - **Excluded at 95% C.L. in two regions 149-206 and 300-440 GeV, and a few segments in between**
  - **Excluded at 90% C.L. from 145-480 GeV**

- **SM4 Higgs boson**
  - **Excluded at 95% C.L. from 120-600 GeV**

- Disentangling the source(s) of some event excesses in low mass analyses will require more data, **which are rapidly coming**

Backup

Model for systematic errors (1)

uncertainty --> nuisance parameter θ, whose best estimate is $\tilde{\theta}$

Bayesian declares his degree of believe on the true value of nuisance θ: $\rho(\theta|\tilde{\theta})$

frequentist wants to know pdf of "measuring" $\tilde{\theta}$, should nuisance be θ: $p(\tilde{\theta}\,|\,\theta)$

Two paradigms can be connected via Bayes' theorem:

$$\rho(\theta|\tilde{\theta}) \sim p(\tilde{\theta}\,|\,\theta) \cdot \pi_\theta(\theta)$$

posterior as a prior for Bayesian analysis

"measurement" PDF for frequentist analysis

primordial flat prior

Type of syst. error	posterior $\rho(\theta\|\tilde{\theta})$	frequentist $p(\tilde{\theta}\,\|\,\theta)$	prior $\pi_\theta(\theta)$
Unconstrained	flat	flat	flat
Gaussian, Lognormal	$\rho(\theta\,\|\,\tilde{\theta}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\theta-\tilde{\theta})^2}{2}\right)$	$p(\tilde{\theta}\,\|\,\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\tilde{\theta}-\theta)^2}{2}\right)$	flat
Statistical	$\rho(\theta\,\|\,N) = \frac{\theta^N}{N!}\ \exp(-\theta)$	$p(N\,\|\,\theta) = \frac{\theta^N}{N!}\ \exp(-\theta)$	flat

Typical examples

Gaussian: $b = b_0 (1 + \varepsilon\theta)$

Lognormal: $b = b_0\, \kappa^\theta$

Statistical: $b = \alpha \times \theta$

Model for systematic errors (2)

All uncertainties are broken into independent sources

- **each independent source gets assigned its own independent nuisance**
- **one source can affect more than one observable: 100% correlated**
- **effect strengths are not necessarily equal: $\varepsilon_1 \neq \varepsilon_2$ or $\kappa_1 \neq \kappa_2$**
- **correlations are either positive or negative: $(\varepsilon_1 > 0, \varepsilon_2 < 0)$ or $(\kappa_1 > 1, \kappa_2 < 1)$**

Limits: Bayesian paradigm

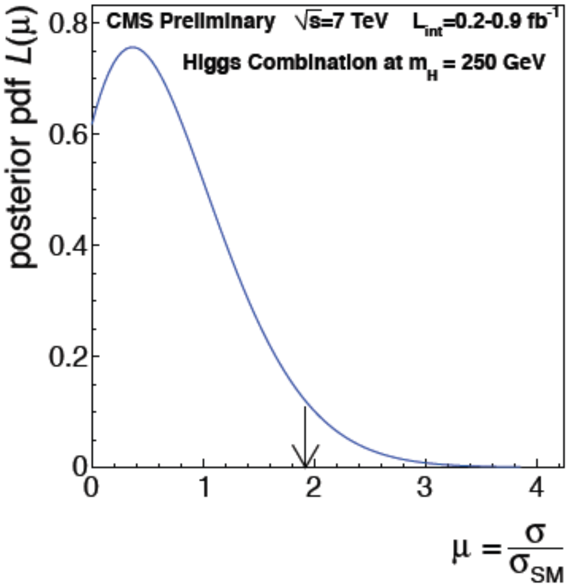
Posterior on signal strength μ

- using flat prior
- marginalization of nuisance parameters

$$p(\mu\,|\,\text{data}) = \frac{1}{C}\int_\theta p\left(\text{data}\,|\,\underline{\mu} s(\theta)+b(\theta)\right)\ \rho_\theta(\theta)\ \underline{\pi_\mu(\mu)}\ \underline{d\theta}.$$

Deriving limit on μ

$$\int_0^{\mu_{95\%CL}} p(\mu\,|\,\text{data})\,d\mu \;=\; 0.95$$



Limits: modified frequentist CL$_s$

CL$_s$ definitions are agreed on with ATLAS → Feel free to compare results

Pure frequentist. *Aided by Bayes' theorem,* all systematic errors are "measurements"

Likelihood

$$\mathcal{L}(\mathrm{data}\,|\,\mu,\theta) = \mathrm{Poisson}\,(\,\mathrm{data}\,|\,\mu\cdot s(\theta)+b(\theta)\,)\cdot p(\tilde{\theta}|\theta)$$

Test statistic

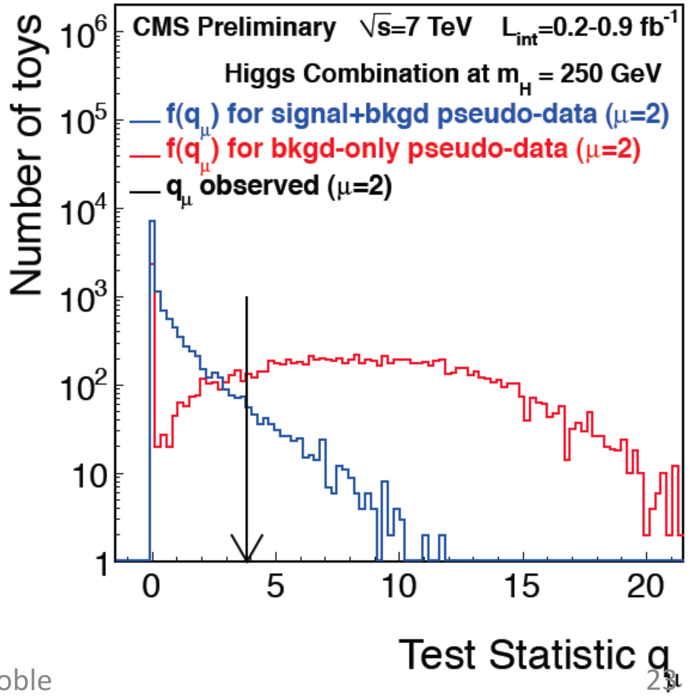
$$\tilde{q}_\mu = -2\ln\frac{\mathcal{L}(\mathrm{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\mathrm{data}|\hat{\mu},\hat{\theta})}, \quad \text{with a constraint } 0\le\hat{\mu}\le\mu$$

Pseudo-data (toys)

- fit data to find two best sets of nuisances $\hat{\theta}_0^{obs}$ and $\hat{\theta}_\mu^{obs}$
- prepare sampling distributions of test statistics
  - bkgd-only pseudo-data: $(\mathrm{data}\,,\tilde{\theta}\,)$ for $b(\hat{\theta}_0^{obs})$
  - signal+bkgd pseudo-data: $(\mathrm{data},\ \tilde{\theta}\ )$ for $\mu s(\hat{\theta}_\mu^{obs})+b(\hat{\theta}_\mu^{obs})$

Define

$$\mathrm{CL_s} = \frac{P\left(q_\mu \ge q_\mu^{obs}\,|\,\mu s(\hat{\theta}_\mu^{obs})+b(\hat{\theta}_\mu^{obs})\right)}{P\left(q_\mu \ge q_\mu^{obs}\,|\,b(\hat{\theta}_0^{obs})\right)}$$



Limits: modified frequentist CL$_s$

- **Likelihood**

$$\mathcal{L}(\text{data}\,|\,\mu,\theta) = \text{Poisson}\,(\,\text{data}\,|\,\mu\cdot s(\theta)+b(\theta)\,)\cdot p(\tilde{\theta}|\theta)$$

- **Test statistics**

$$\tilde{q}_\mu \;=\; -2\ln\frac{\mathcal{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})}, \qquad \text{with a constraint } 0\le\hat{\mu}\le\mu$$

- **Pseudo-data (toys)**
  - fit data to find two best sets of nuisances $\hat{\theta}_0^{obs}$ and $\hat{\theta}_\mu^{obs}$
  - prepare sampling distributions of test statistics
    - bkgd-only pseudo-data: $(\text{data}\,,\tilde{\theta})$ for $b(\hat{\theta}_0^{obs})$
    - signal+bkgd pseudo-data: $(\text{data},\,\tilde{\theta}\,)$ for $\mu s(\hat{\theta}_\mu^{obs})+b(\hat{\theta}_\mu^{obs})$

- **Define** CL$_s$ =

$$\text{CL}_s = \frac{P\left(q_\mu \ge q_\mu^{obs}\,|\,\mu s(\hat{\theta}_\mu^{obs})+b(\hat{\theta}_\mu^{obs})\right)}{P\left(q_\mu \ge q_\mu^{obs}\,|\,b(\hat{\theta}_0^{obs})\right)}$$

LEP did not use syst. error pdf's in Likelihood

TEV puts in $\rho(\theta|\tilde{\theta})$, = $p(\tilde{\theta}\,|\,\theta)$ for pdf's we use

LEP and TEV, fix μ=0 in the denominator

LEP does not profile nuisance (there aren't any)

TEV does profile for nuisances

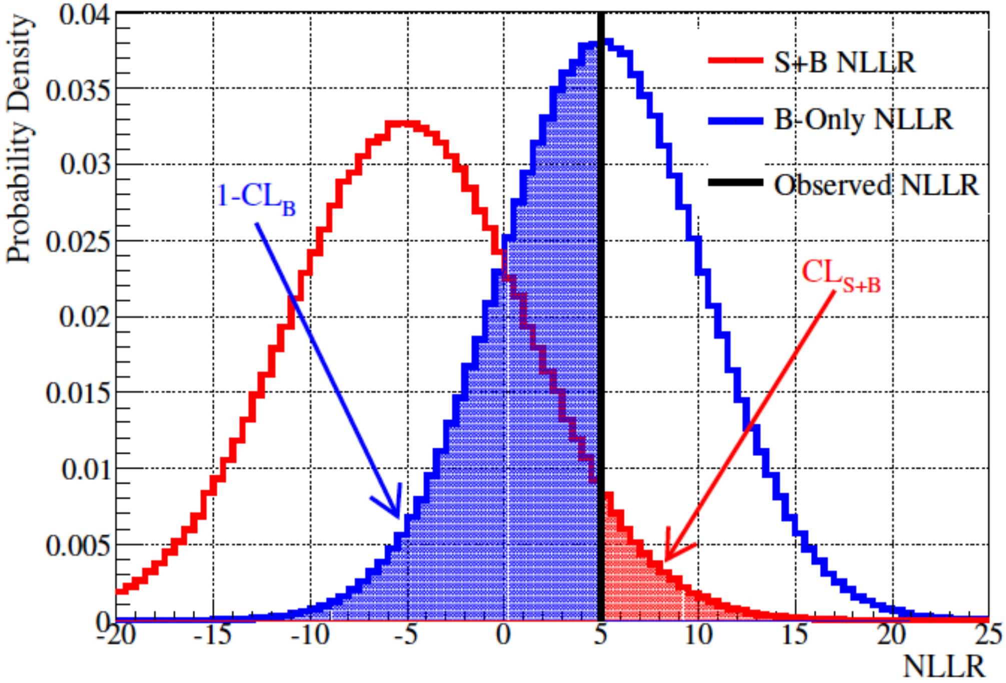
TEV/LEP use $\tilde{\theta}$ to generate θ and then generate pseudo-data using new s(θ) and b(θ), which is explicitly Bayesian

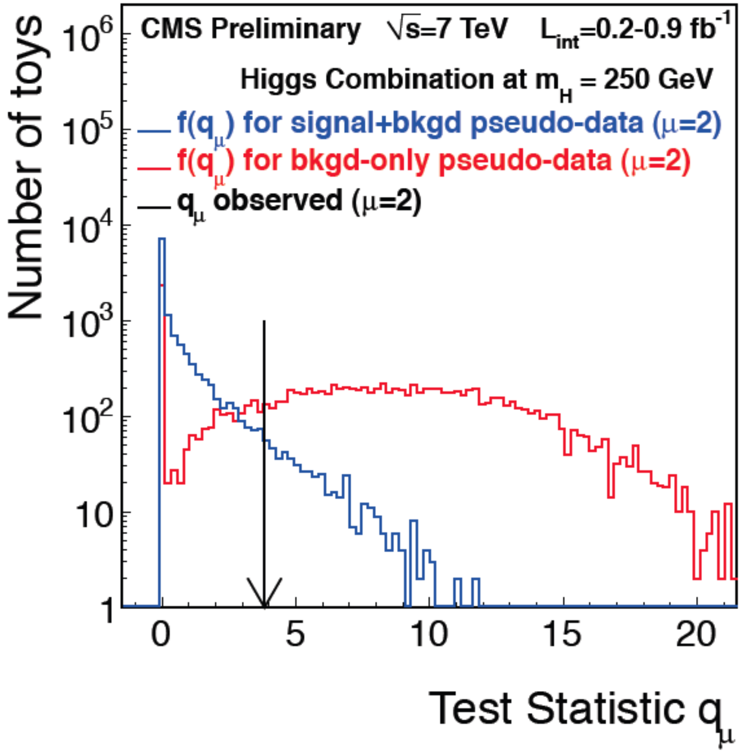
LHC-type CL$_s$: what is different?

Visually, sampling distributions are very different from LEP/TEV

Numerically, results obtained by tossing pseudo-data are very similar

Example from Collie documentation (D0 note 5595)





Quantifying an excess

test statistic – profile likelihood

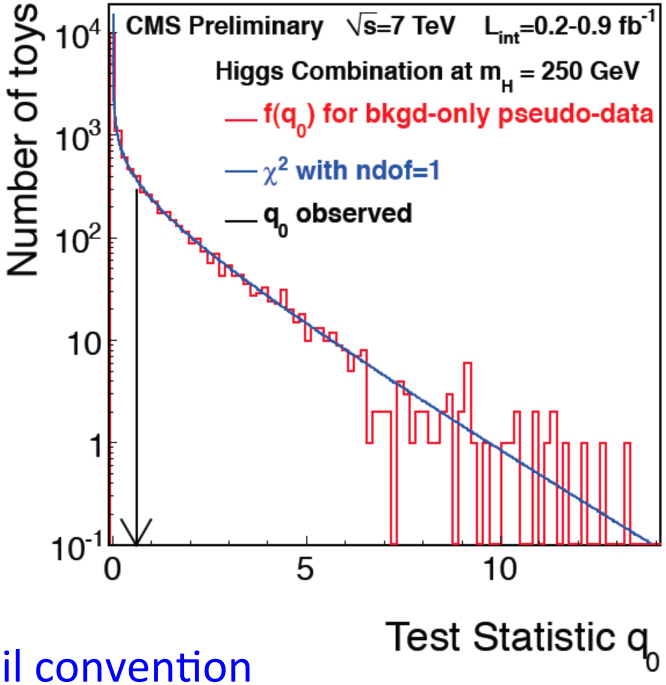
$$q_0 = -2\ln\frac{\mathcal{L}(\text{data}|0,\hat{\theta}_0)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})} \quad \text{and } \hat{\mu} \geq 0$$

best fit value, not necessary consistent with SM Higgs

"local" p-values

- either from tossing pseudo data
- or from from asymptotic approximation

$$\tilde{p} = \frac{1}{2}\left[1 - \operatorname{erf}\left(\sqrt{q_0^{\text{obs}}/2}\right)\right]$$

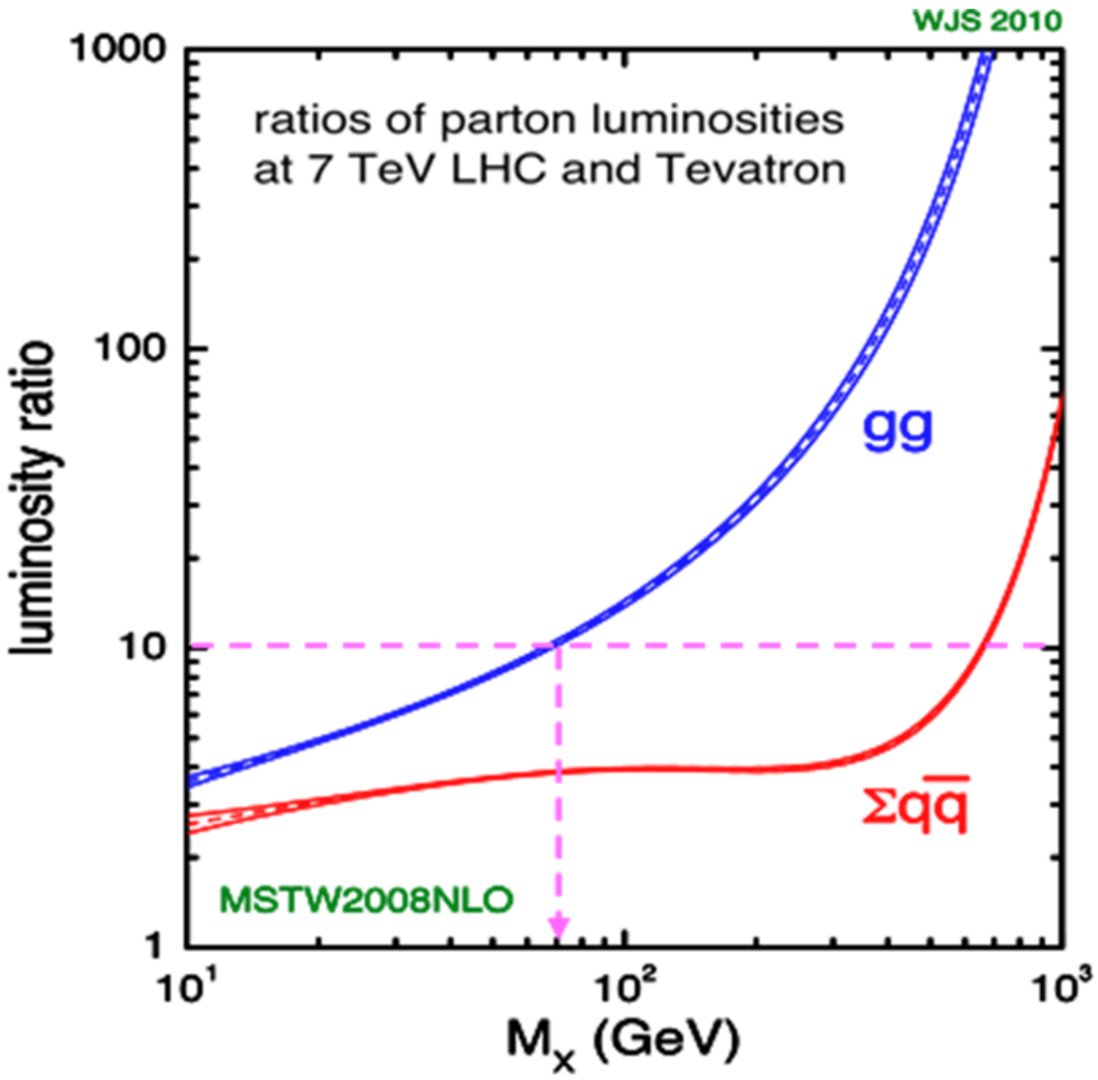


"local" significance Z: one-sided normal distribution tail convention

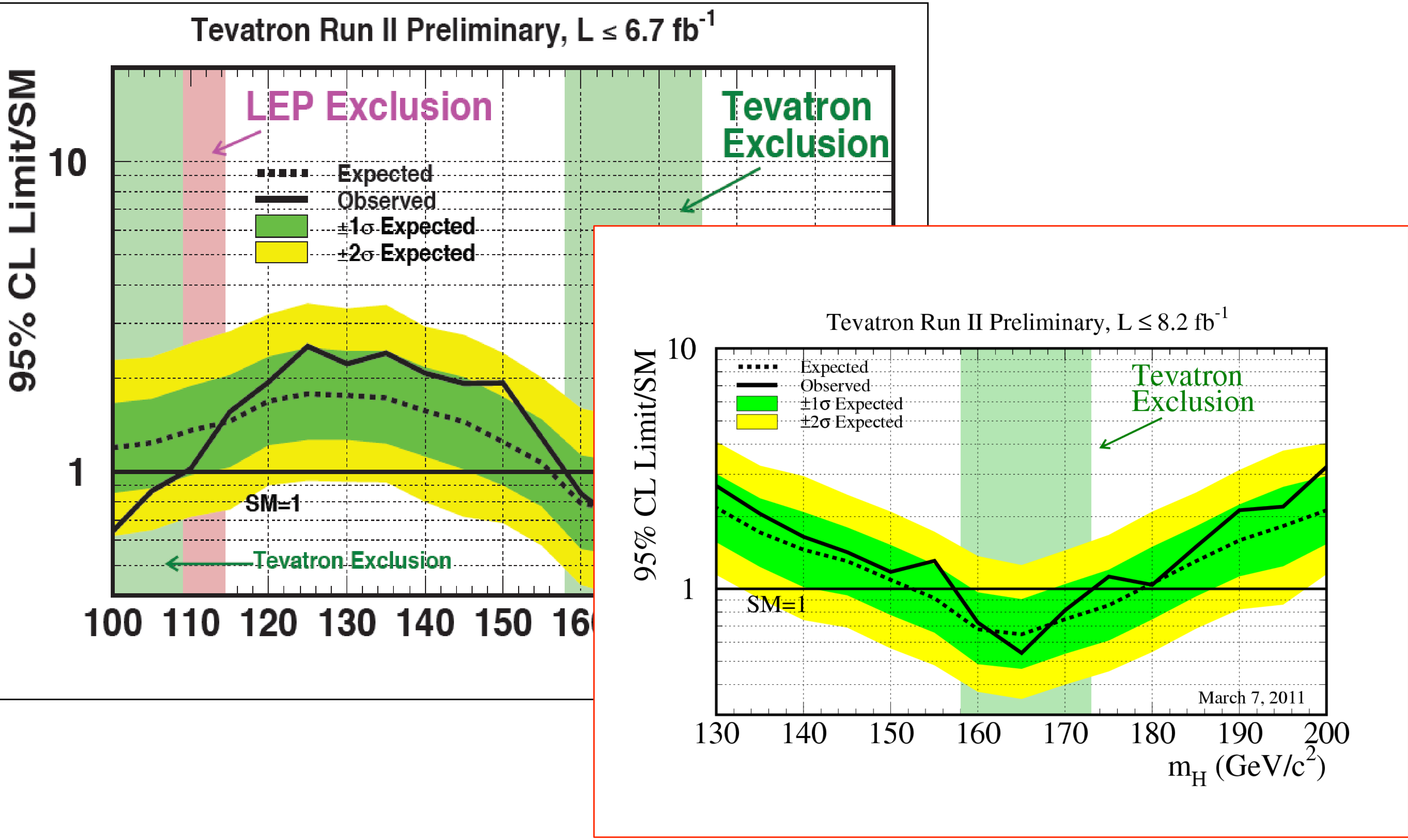
$$p = \int_Z^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)\,dx$$

look-elsewhere effect is straightforward for simple background models, but not for their combination

LHC/TEV partonic luminosities



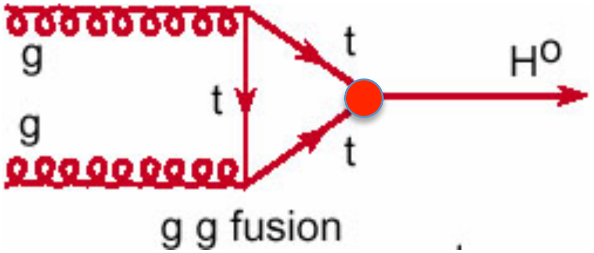
SM Higgs at Tevatron



What is SM4 Higgs?

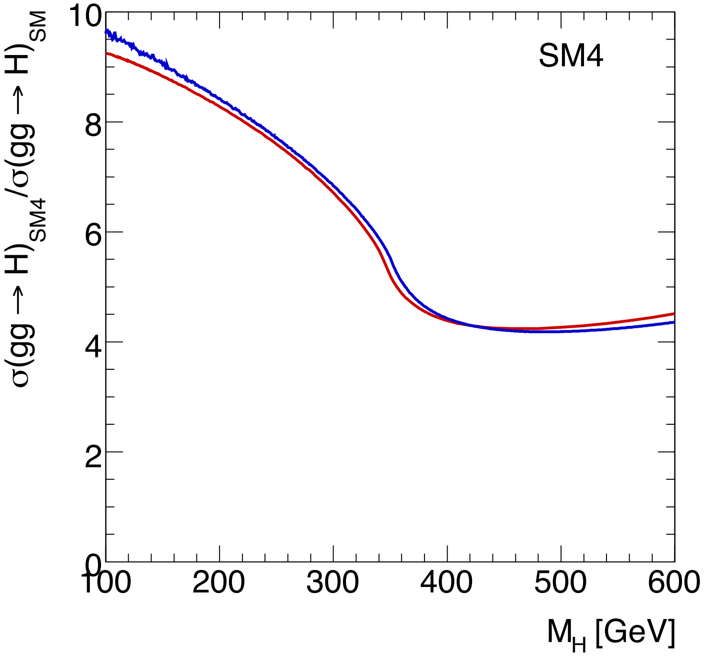
ggH: top-quark loop is a dominant process and is almost quark-mass independent

- quark-Higgs coupling is ~m_q
- "kinematic" penalty for a heavy quark in the loop is ~$1/m_q$
- around $m_H \sim 2m_{top}$, CS has "resonant" enhancement



One more generation implies 2 additional quarks

- **ggH gets a factor of 3 enhancement in ME (t + t' + b') and, hence, a factor of 9 in the cross-section**
- around $m_H \sim 2m_{top}$, "the top-quark resonant" piece obviously does not get enhanced, and the overall ratio of $\sigma_{SM4} / \sigma_{SM}$ is smaller than a factor of 9



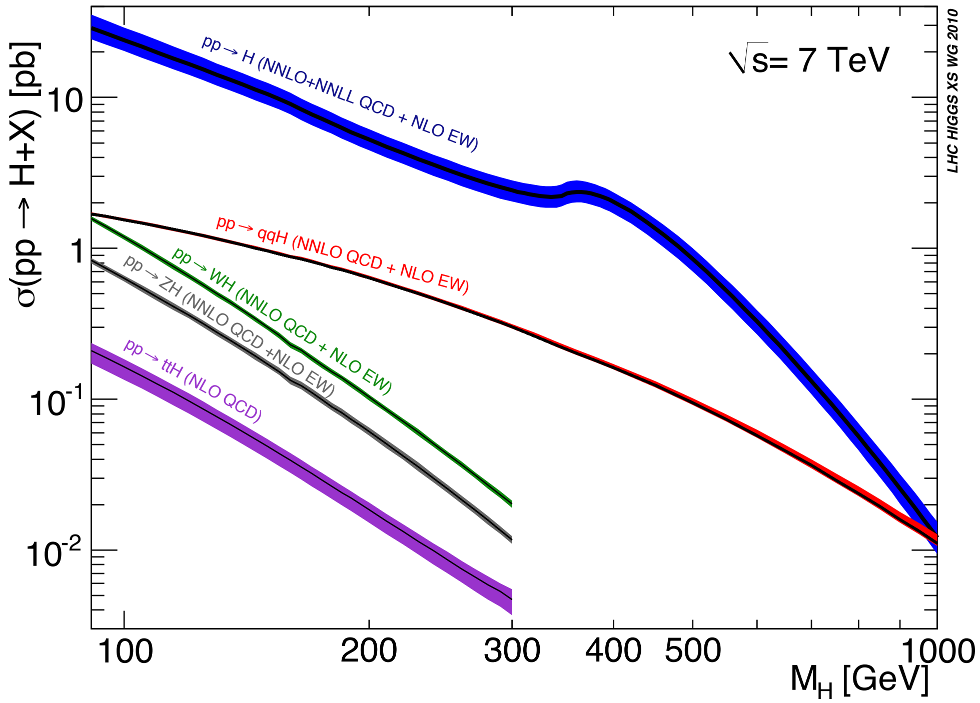
NOTE1: VBF, VH, ttH production do not get any enhancements

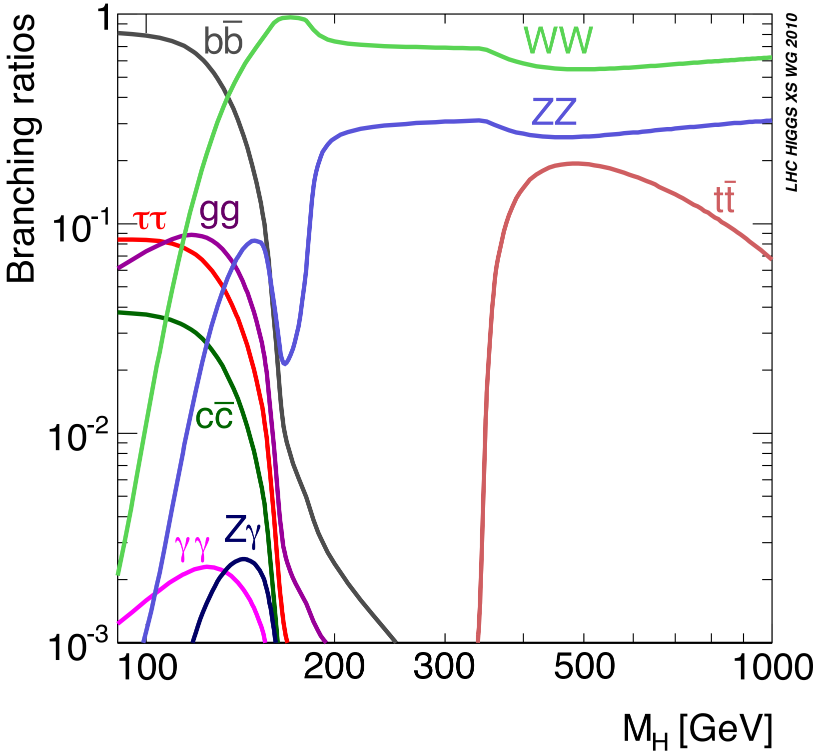
NOTE2: Decays going via loop diagrams are also affected

- Γ(gg) becomes larger (additional quarks in the loop)
- Γ(γγ) becomes smaller (cancellation between q and W loops)

NOTE3: SM4 with heavy Higgs does not contradict precision EWK measurements

SM Higgs: production & decays





Cross sections, branching ratios, and their errors
are prepared by the LHC Higgs Cross Section group

Inter-analysis correlated nuisances

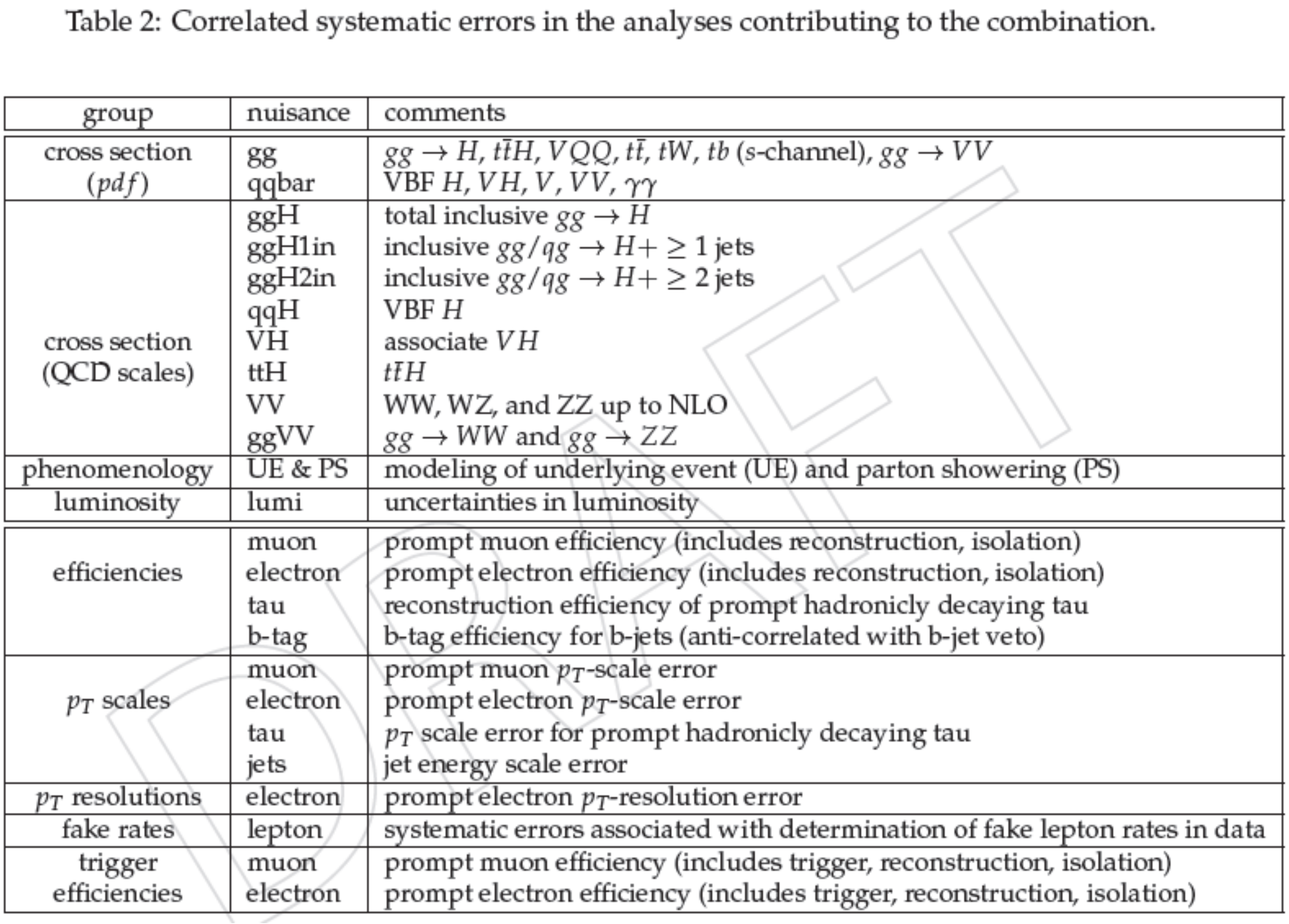
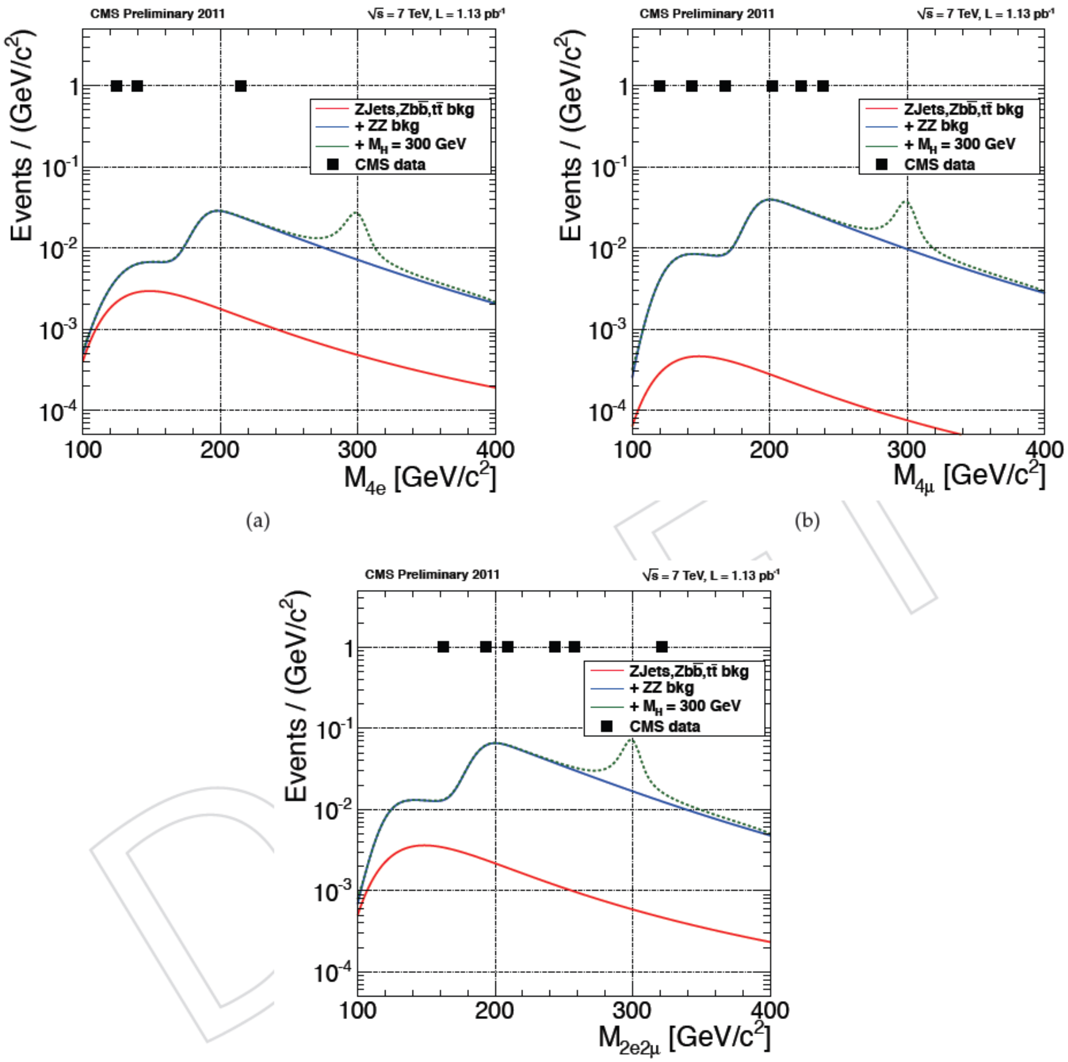


Table 2: Correlated systematic errors in the analyses contributing to the combination.

group	nuisance	comments
cross section (pdf)	gg	$gg \to H$, $t\bar{t}H$, VQQ, $t\bar{t}$, tW, tb (s-channel), $gg \to VV$
	qqbar	VBF H, VH, V, VV, $\gamma\gamma$
cross section (QCD scales)	ggH	total inclusive $gg \to H$
	ggH1in	inclusive $gg/qg \to H + \geq 1$ jets
	ggH2in	inclusive $gg/qg \to H + \geq 2$ jets
	qqH	VBF H
	VH	associate VH
	ttH	$t\bar{t}H$
	VV	WW, WZ, and ZZ up to NLO
	ggVV	$gg \to WW$ and $gg \to ZZ$
phenomenology	UE & PS	modeling of underlying event (UE) and parton showering (PS)
luminosity	lumi	uncertainties in luminosity
efficiencies	muon	prompt muon efficiency (includes reconstruction, isolation)
	electron	prompt electron efficiency (includes reconstruction, isolation)
	tau	reconstruction efficiency of prompt hadronicly decaying tau
	b-tag	b-tag efficiency for b-jets (anti-correlated with b-jet veto)
p_T scales	muon	prompt muon p_T-scale error
	electron	prompt electron p_T-scale error
	tau	p_T scale error for prompt hadronicly decaying tau
	jets	jet energy scale error
p_T resolutions	electron	prompt electron p_T-resolution error
fake rates	lepton	systematic errors associated with determination of fake lepton rates in data
trigger efficiencies	muon	prompt muon efficiency (includes trigger, reconstruction, isolation)
	electron	prompt electron efficiency (includes trigger, reconstruction, isolation)

4l events

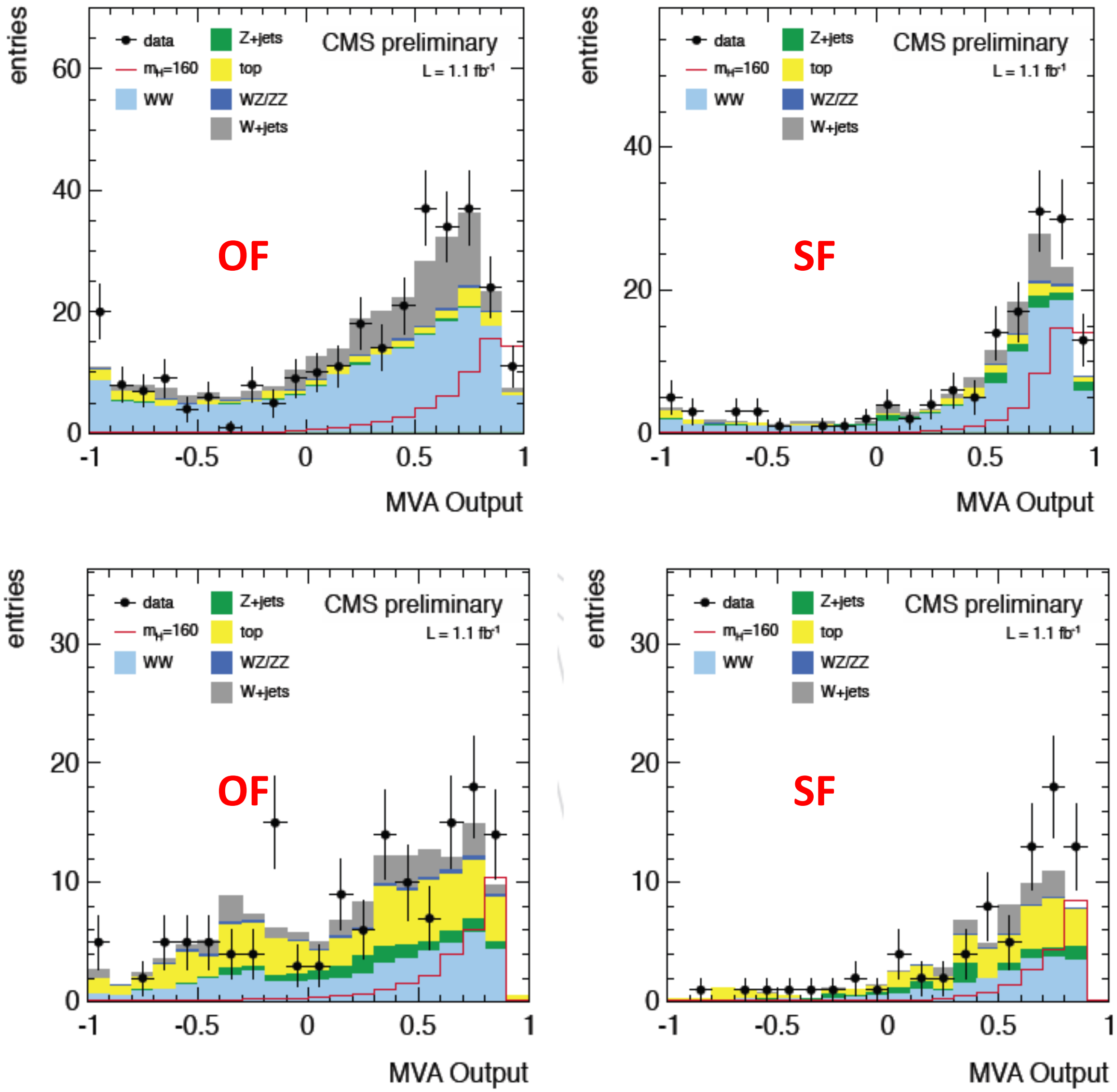


CMS Preliminary 2011
$\sqrt{s}$ = 7 TeV, L = 1.13 pb^{-1}
Events / (GeV/c^2)
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
100
200
300
400
M_{4e} [GeV/c^2]
$M_{4\mu}$ [GeV/c^2]
$M_{2e2\mu}$ [GeV/c^2]
ZJets,Zb$\bar{b}$,t$\bar{t}$ bkg
+ ZZ bkg
+ M_H = 300 GeV
CMS data

(a)

(b)

H→WW MVA outputs



0 jets

1 jet