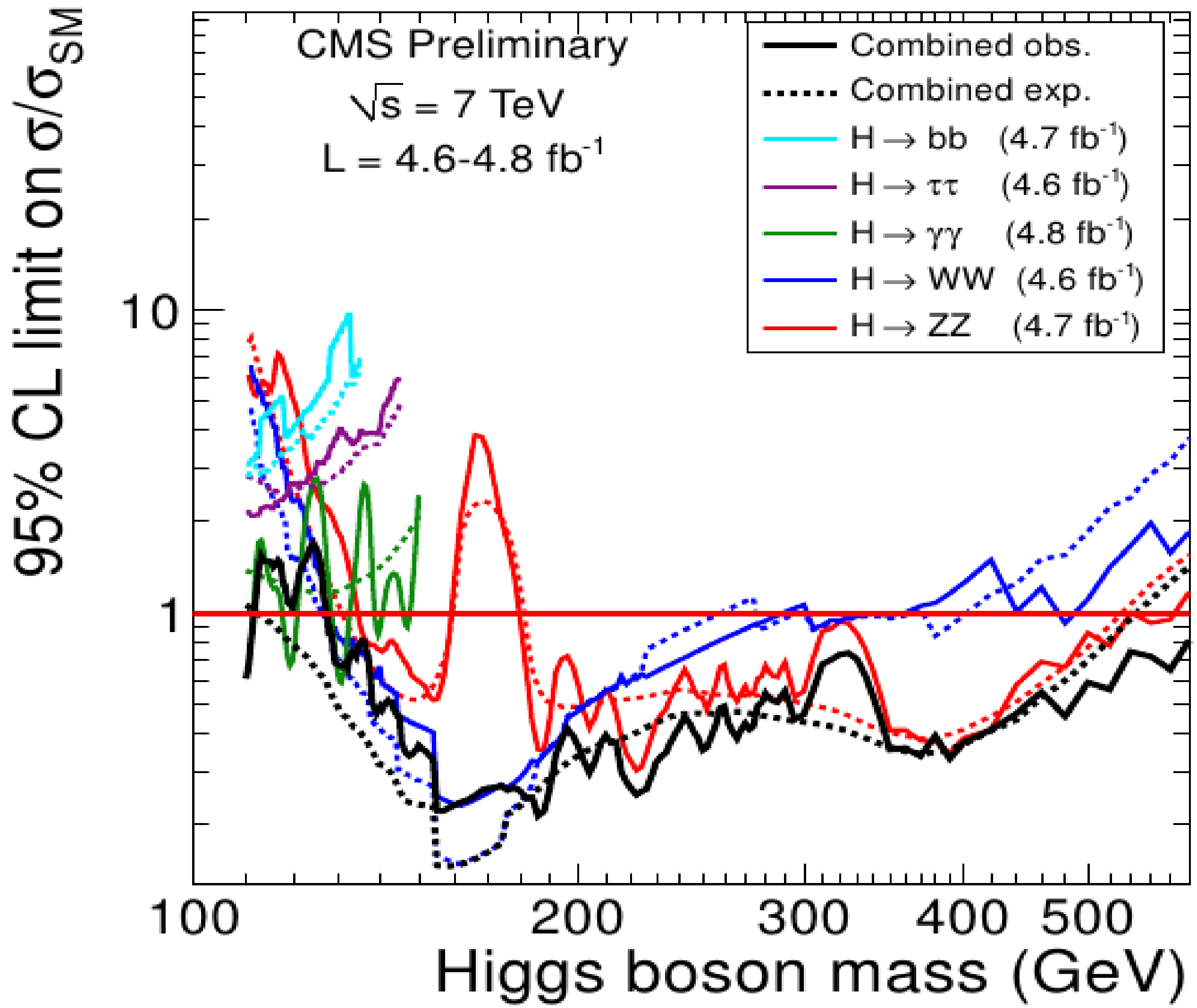
Standard Model Higgs Boson Searches at CMS

Matteo Sani, UCSD
on behalf of the CMS Collaboration

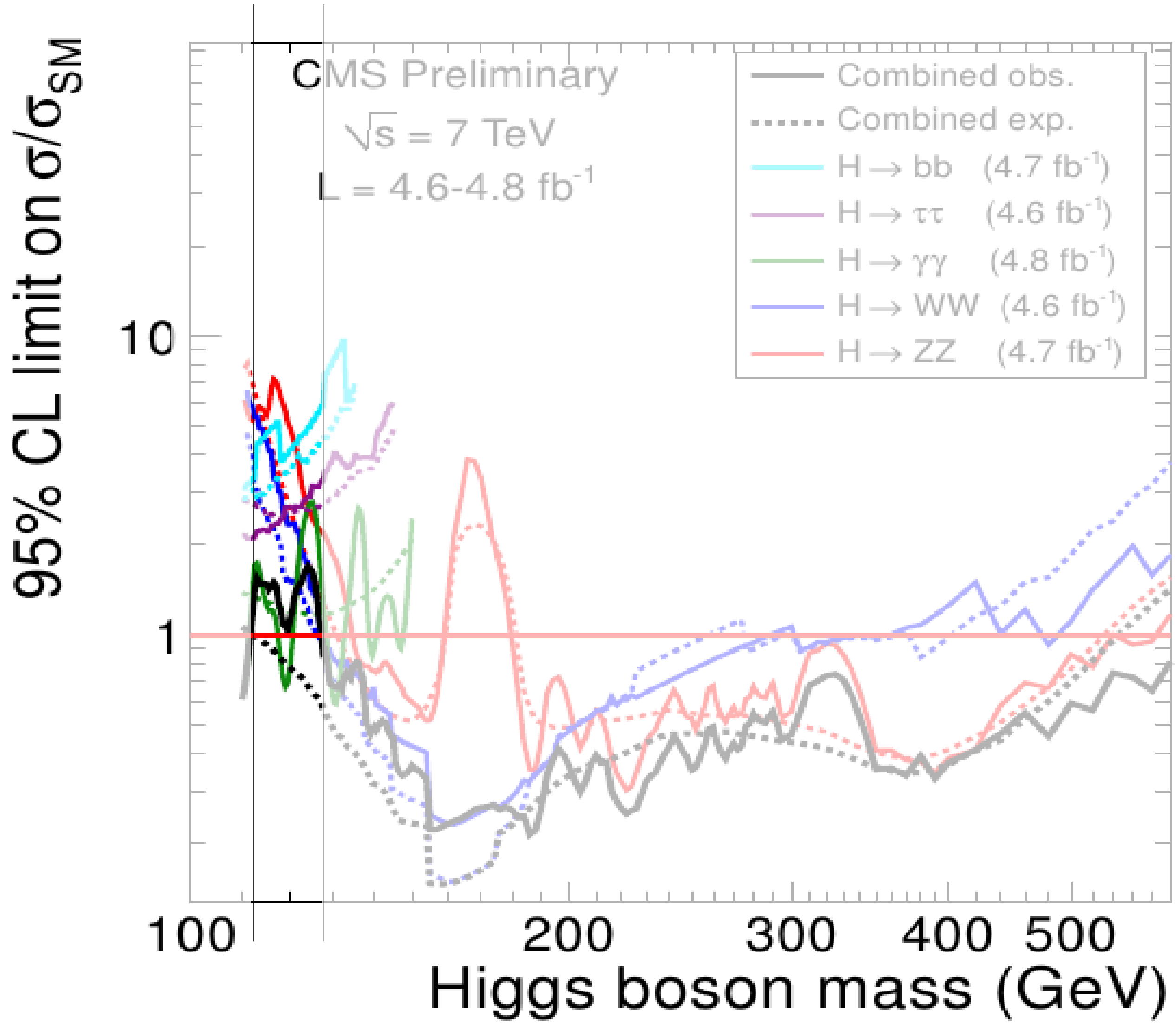
Where we were (end 2011)

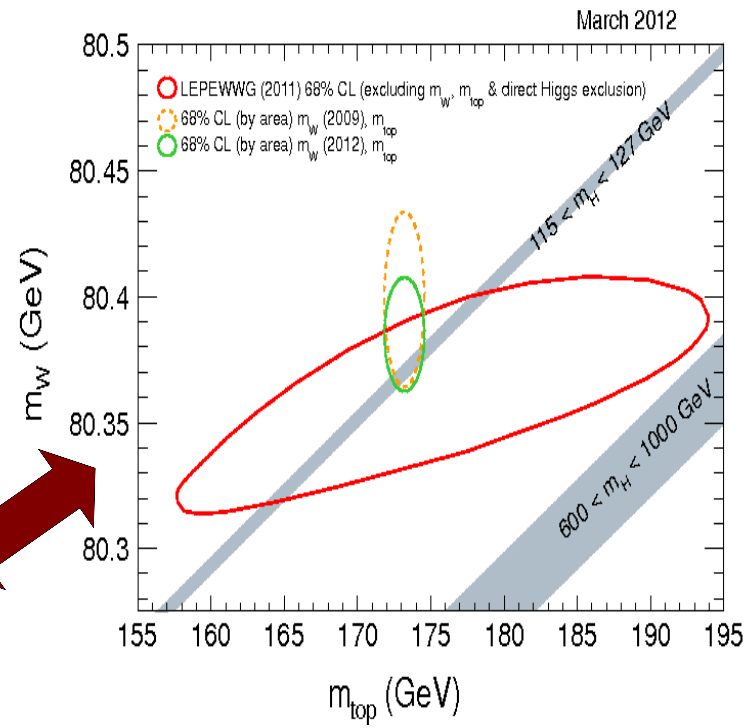
The SM Higgs phenomenology is precisely predicted
except for its mass (MH).



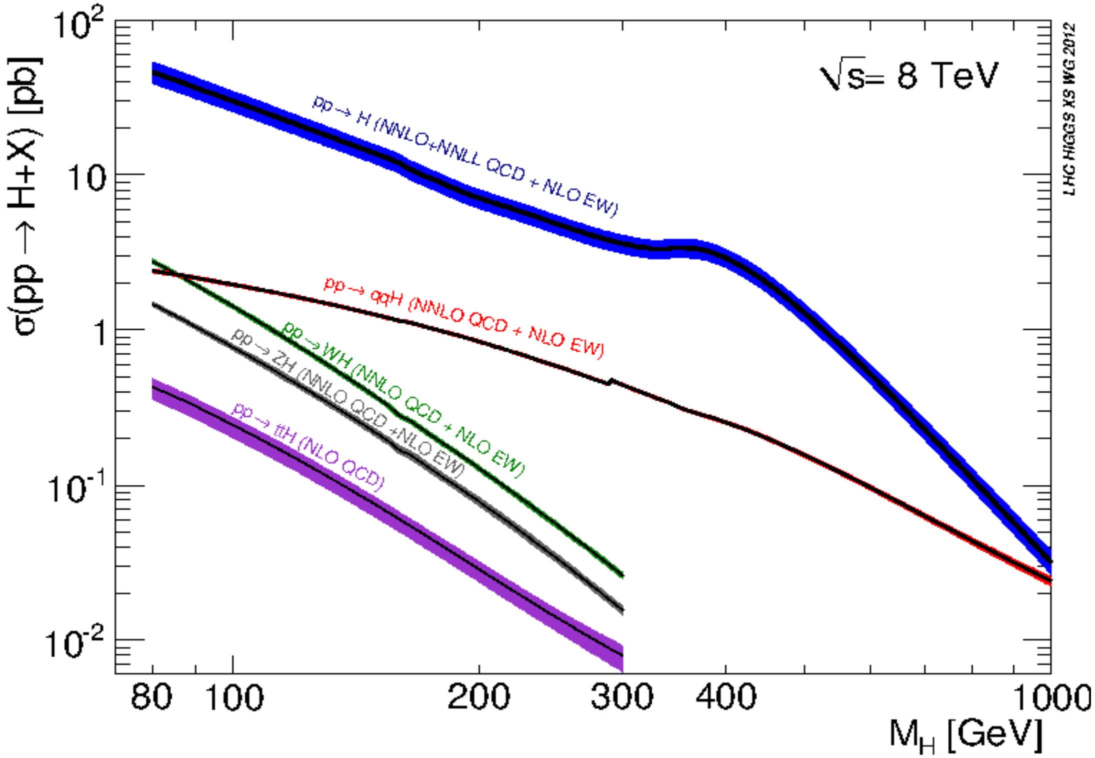
Where we were (end 2011)

The SM Higgs phenomenology is precisely predicted except for its mass (MH).

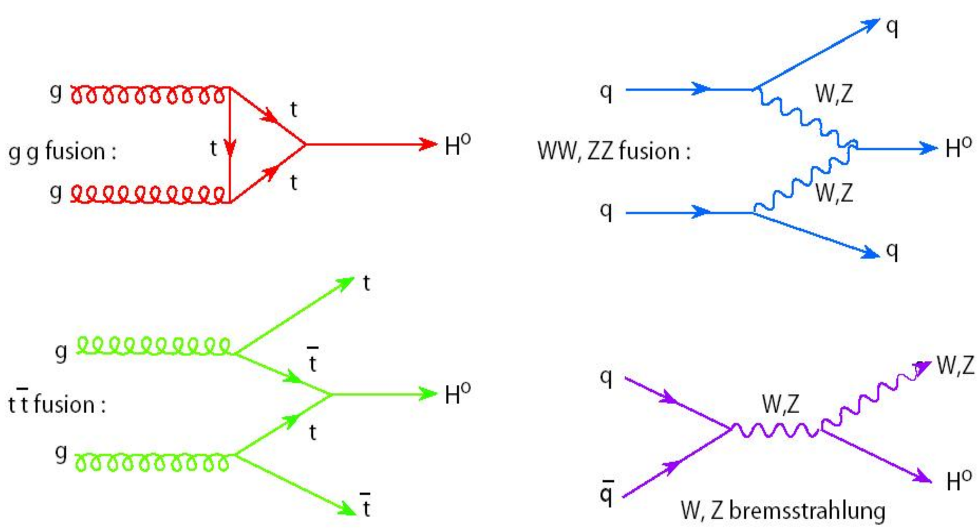


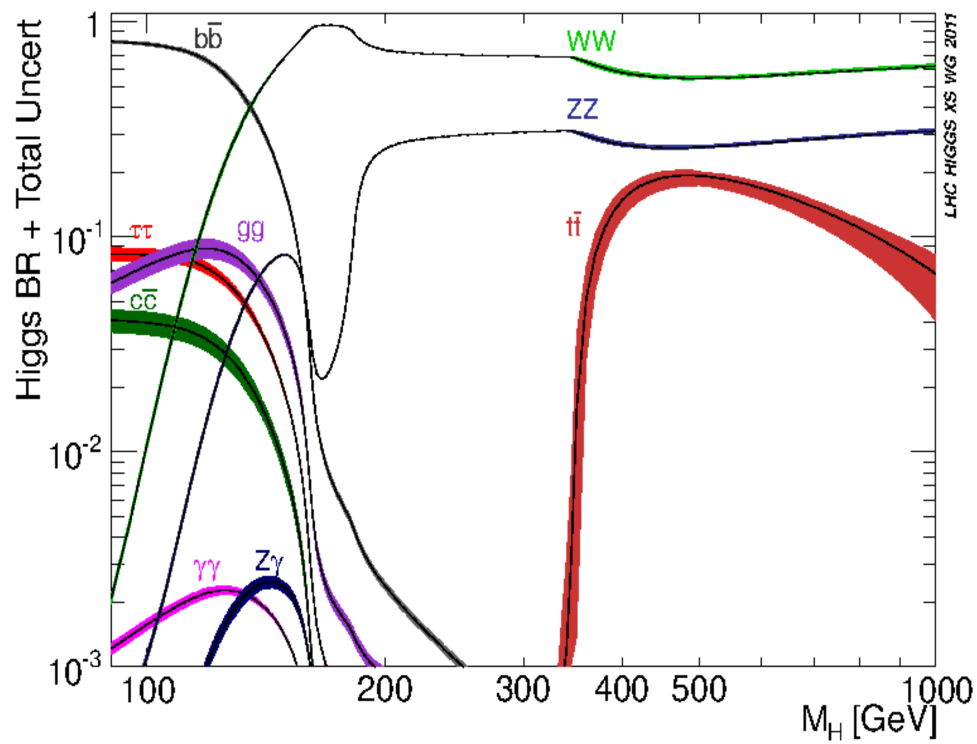


Higgs Production and BR at LHC

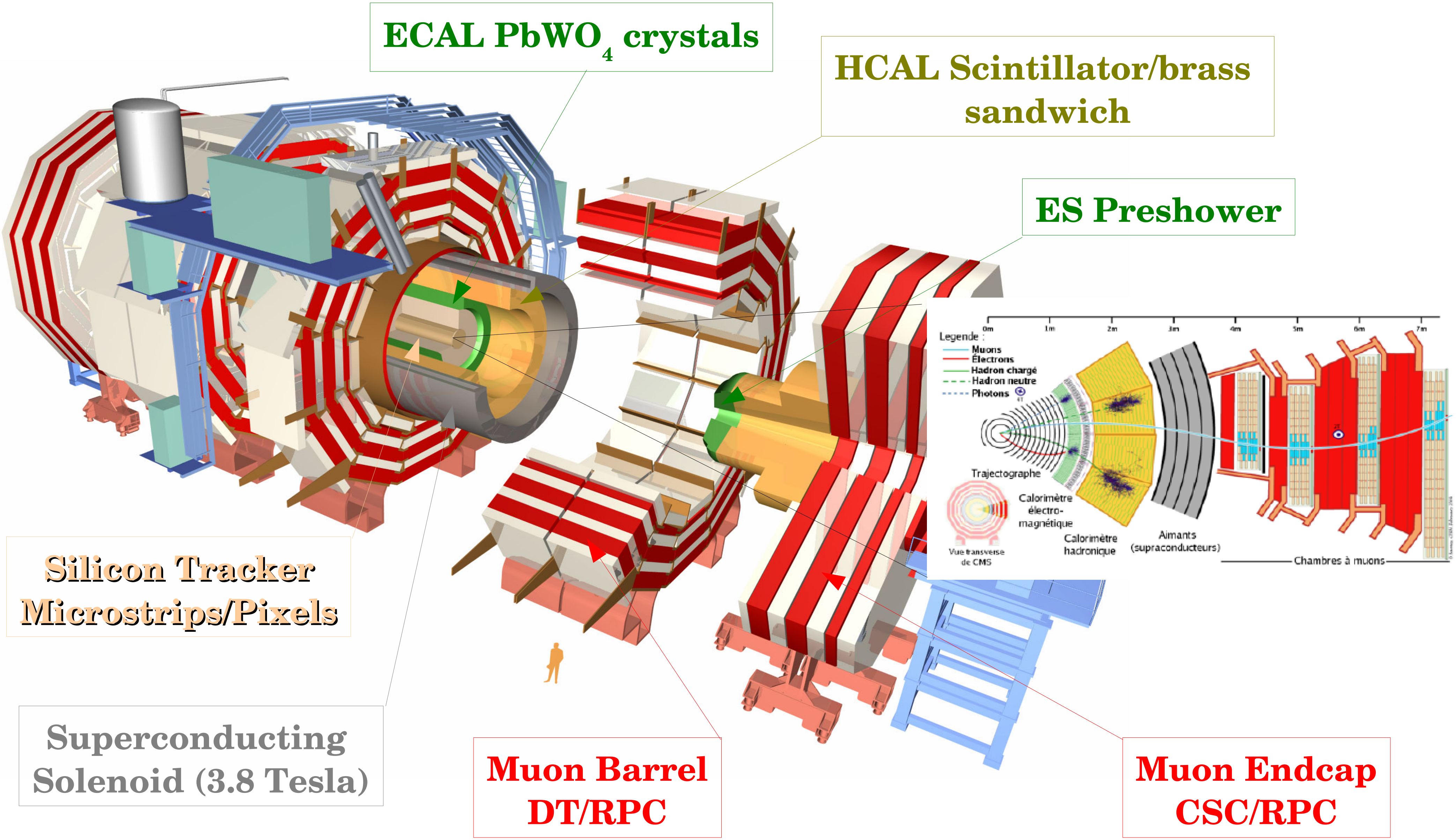


Dominant production mechanism: gluon fusion (gg→H)



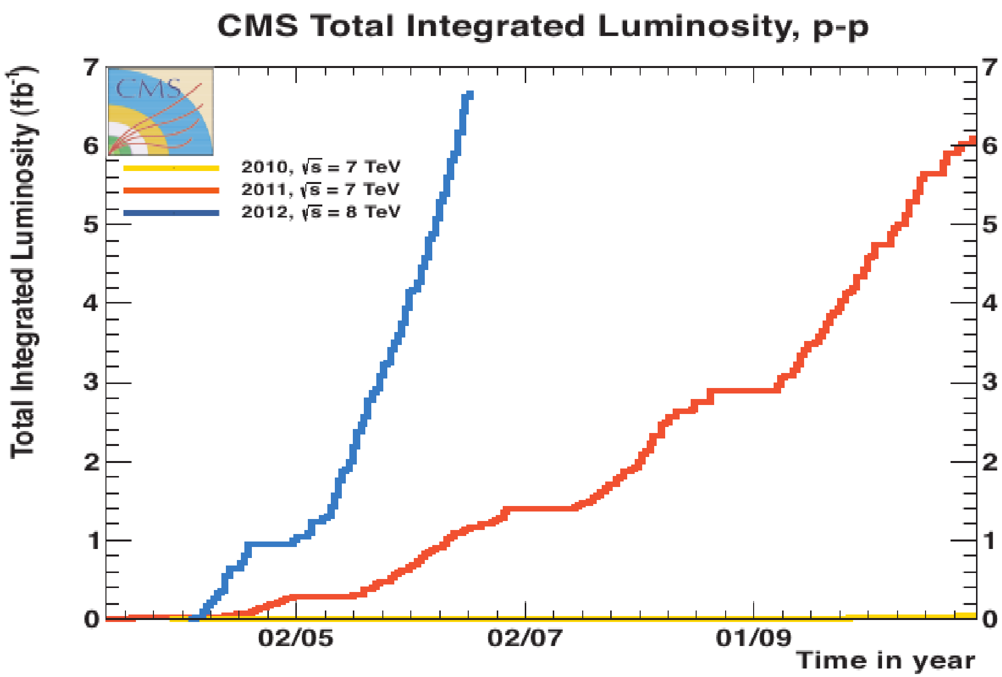


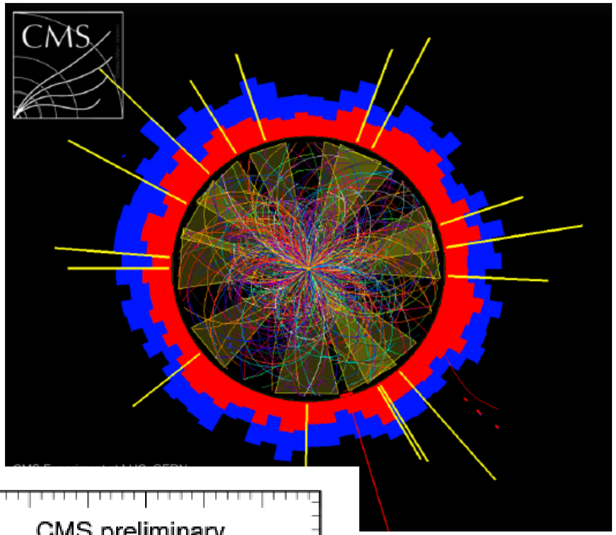
Compact Muon Solenoid

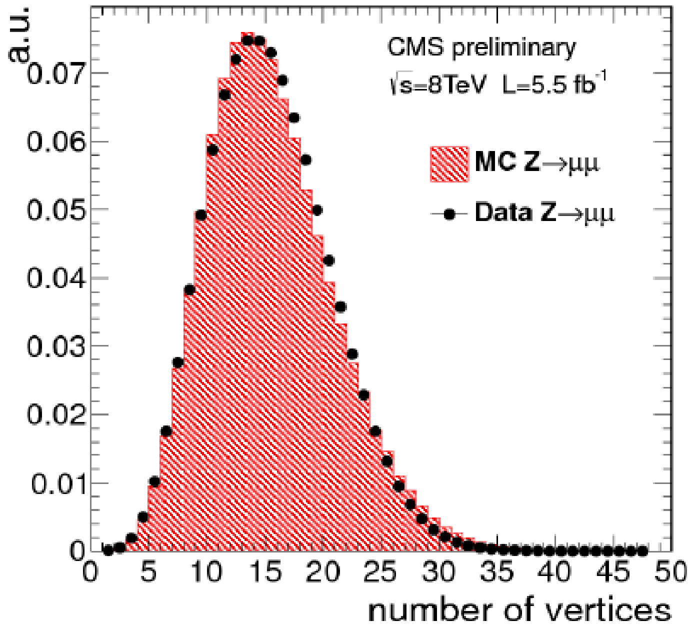


Data Collection and Reconstruction

CMS has collected 5.1 fb^{-1} at $\sqrt{s}$ = 7 TeV + 5.3 fb^{-1} at $\sqrt{s}$ = 8 TeV high quality data at LHC pp collider.







High pileup disguises possible signal events.
All analysis have been improved to cope with the new data taking conditions.

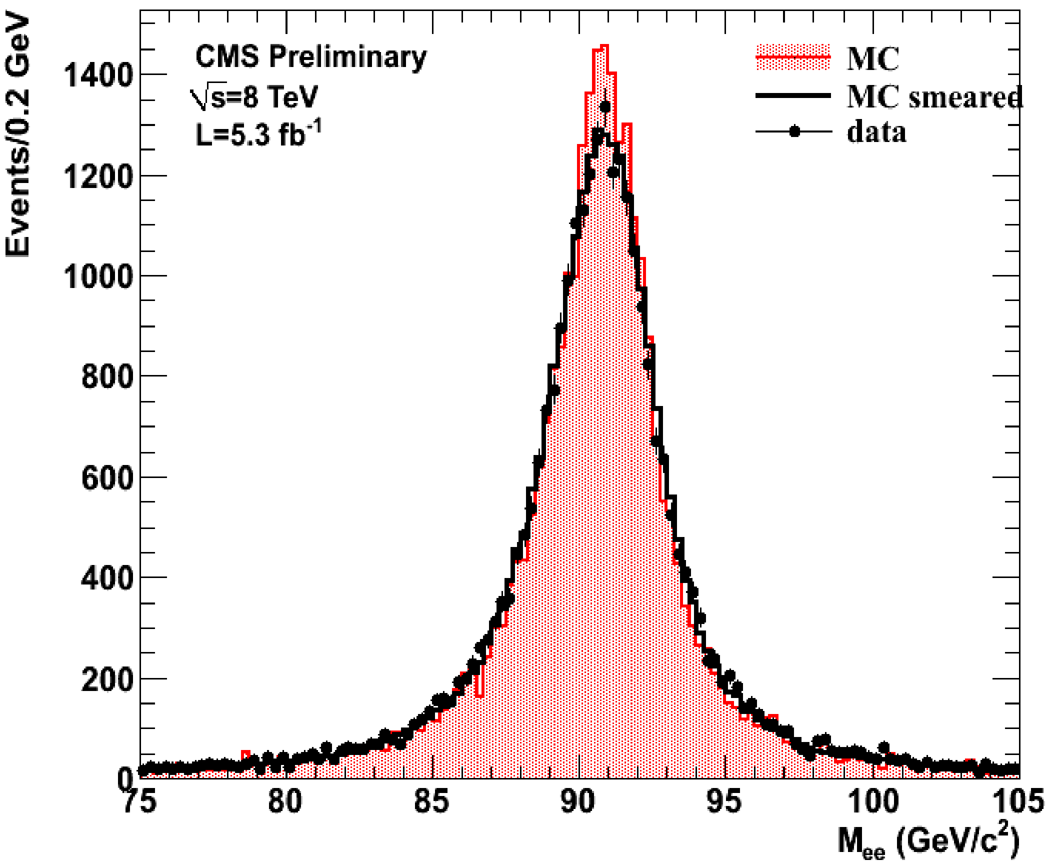
CMS Search Channels

Decay	Production Topology	Physics Analysis Summary
$H \to \gamma\gamma$	Inclusive, VBF	HIG-12-015
$H \to ZZ$	Inclusive	HIG-12-016
$H \to WW$	0/1 jet, VBF	HIG-12-017
$H \to bb$	WH, ZH, ttH	HIG-12-019 HIG-12-025
$H \to \tau\tau$	Inclusive, VBF	HIG-12-018
Combination		HIG-12-020

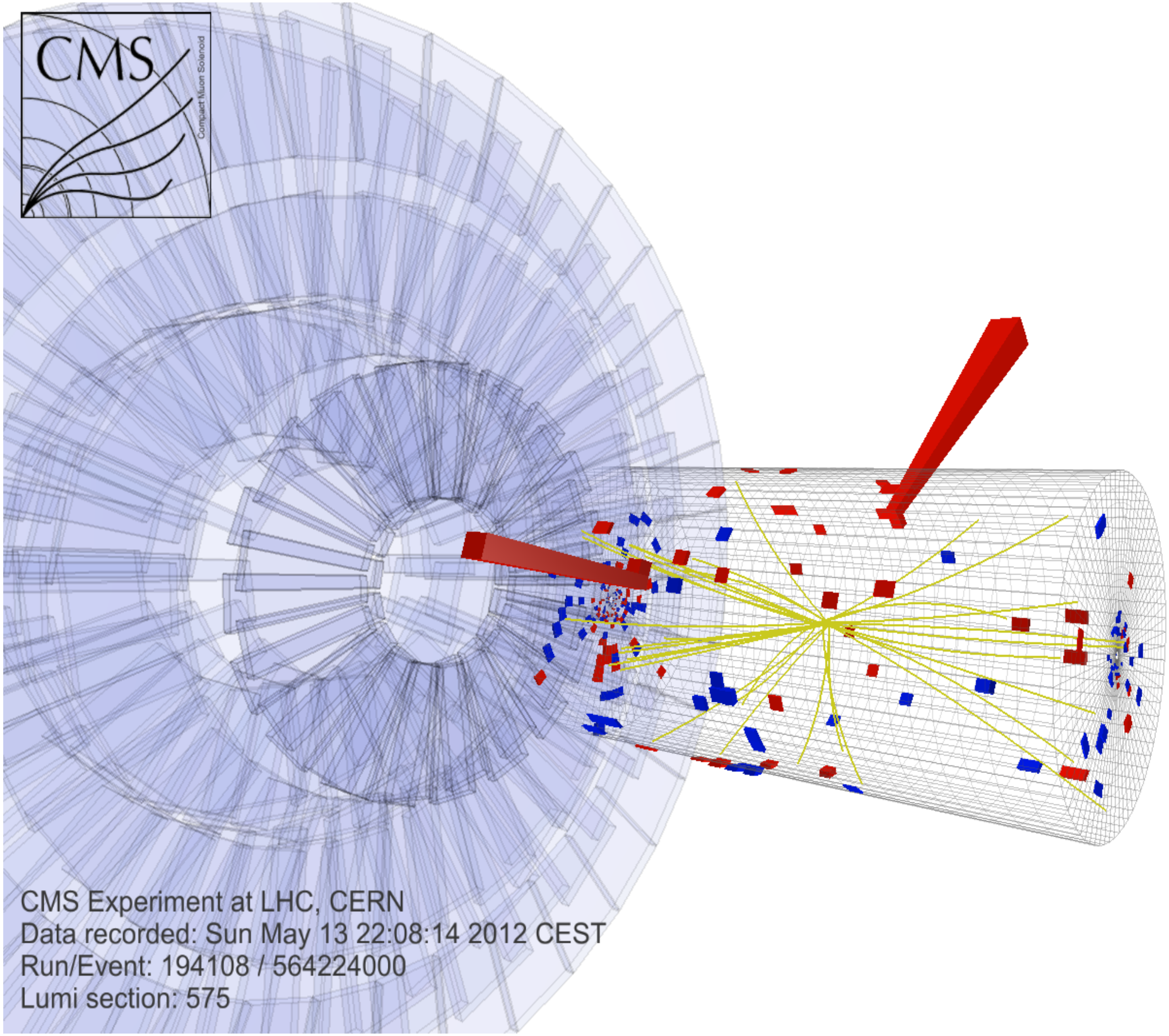
In all decay modes blinded analysis have been performed.

H→γγ

- Discovery channel at low mass:
  - **Low signal rate** (BR ~ 2.10^{-3}) but **very clear signature** with a narrow peak (**mass resolution 1-2%**) on top of continuous background;
  - Background: di-photon prompt production (irreducible), γ+jet and di-jet QCD, DY (reducible).
- Good performance of ECAL and photon reconstruction are crucial for the analysis.
- Reconstructed cluster energies corrected using a MC trained multivariate regression for:
  - Containment of the shower
  - Conversion recovery
  - Pile-up resilience



Event Display



Event parameters:

$M_{\gamma\gamma}$ = 125.9 GeV

$p_T^{\gamma 1}$ = 89.8 GeV

$p_T^{\gamma 2}$ = 46.5 GeV

$\eta_{\gamma 1}$ = 0.06

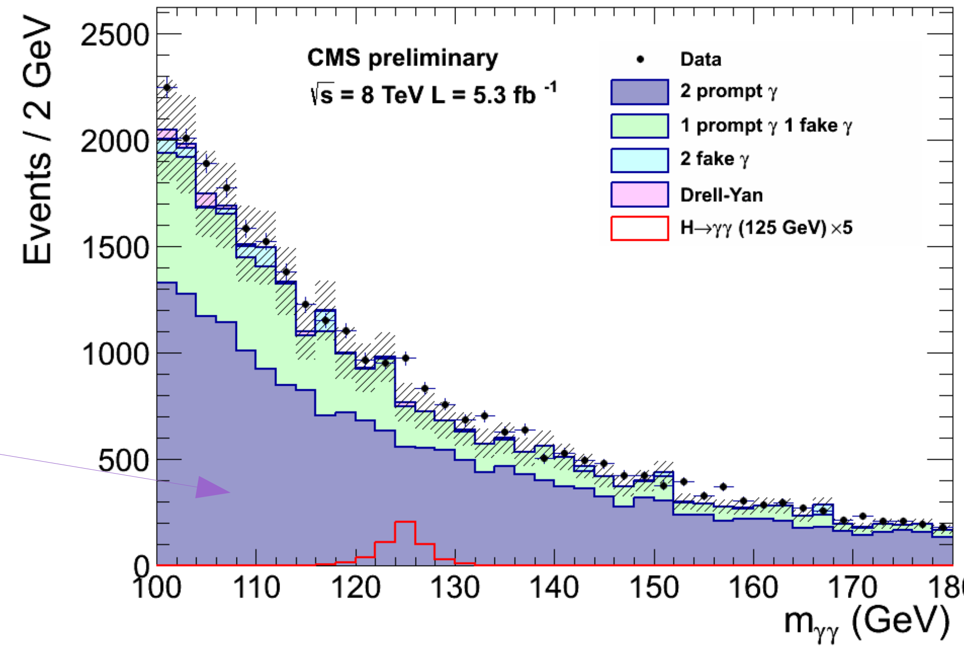
$\eta_{\gamma 2}$ = -0.81

σ_M/M =0.89%

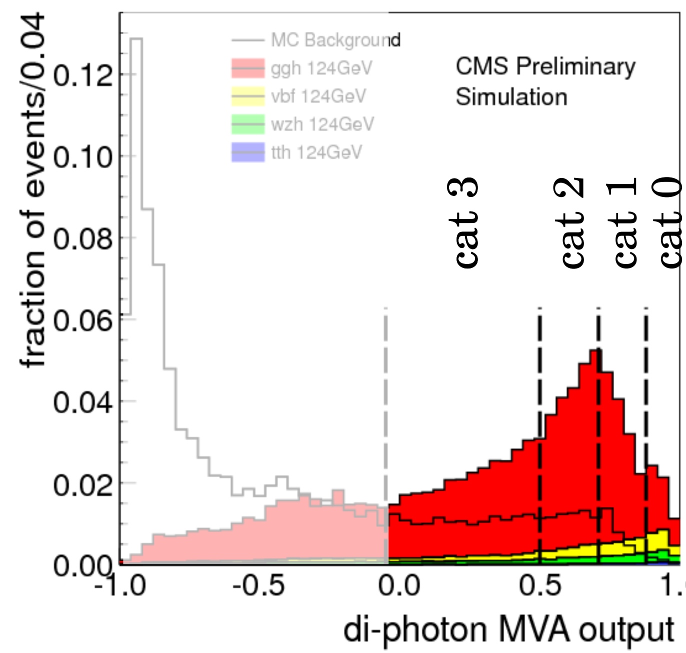
$p_T^{\gamma\gamma}$ =78.4 GeV

Event Selection and Classification

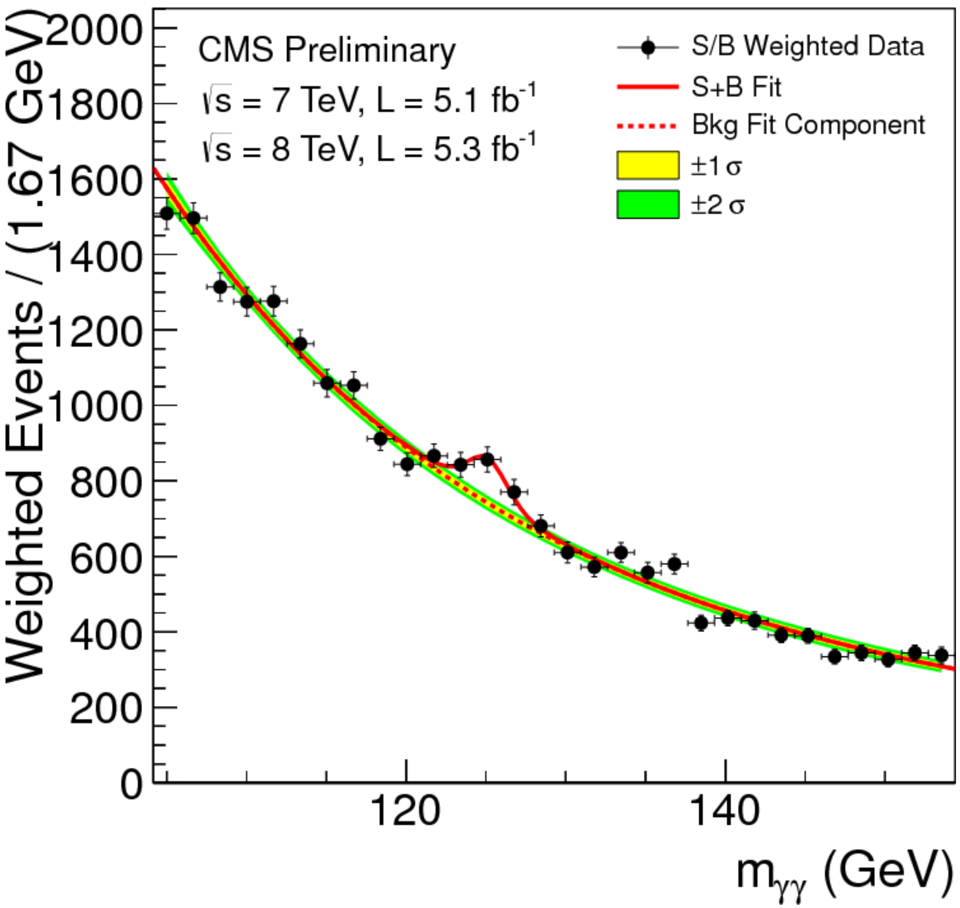
- **Selection:** two isolated photons selected with photon ID based on Boosted Decision Tree (BDT) (separate prompt γ from π^0):
  - Correct γ momentum for the vertex (PV); its choice done using BDT algorithm.
  - Irreducible background is dominant ~70%

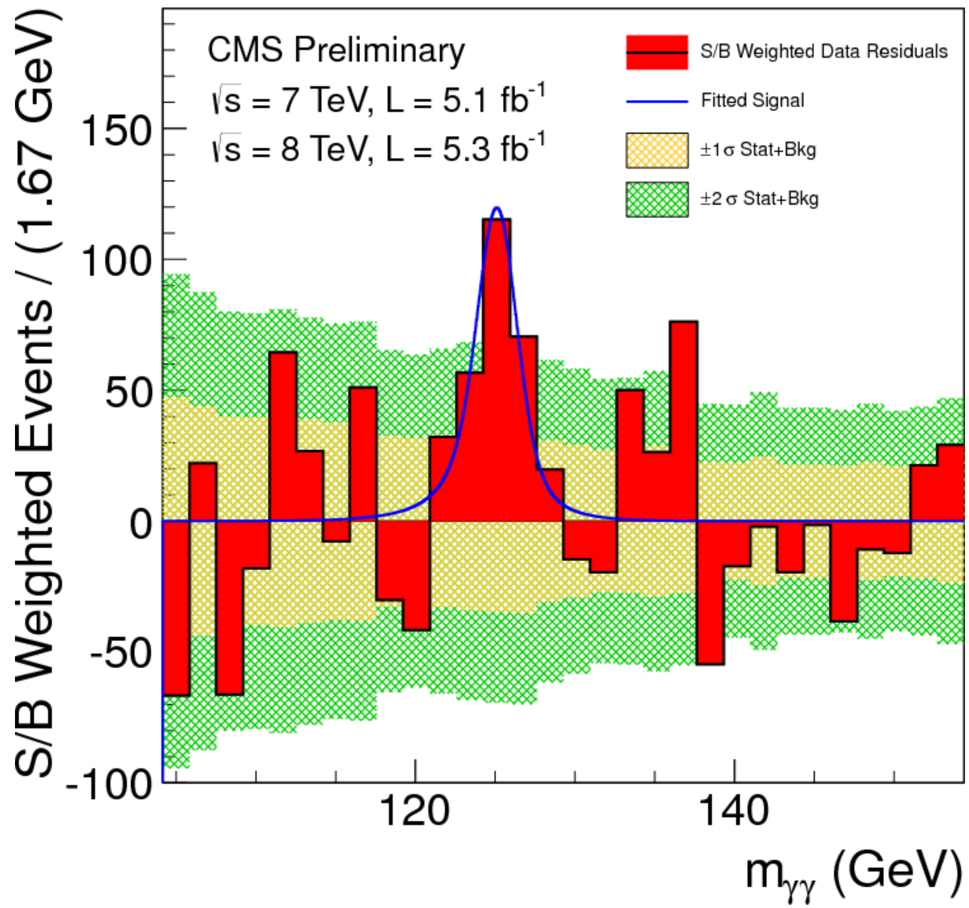


- **Classification:** Multivariate (MVA) classifier with high score events with:
  - Signal-like kinematic (predominantly high $p_T^{\gamma\gamma}$)
  - Good di-photon mass resolution
- Six classes: 4 di-photon MVA and 2 di-jet tagged categories (boundaries optimized for best expected limit on MC):
  - Background estimated from data.
- Statistical analysis done by fitting the invariant mass distribution for each category.



Invariant Mass Spectrum



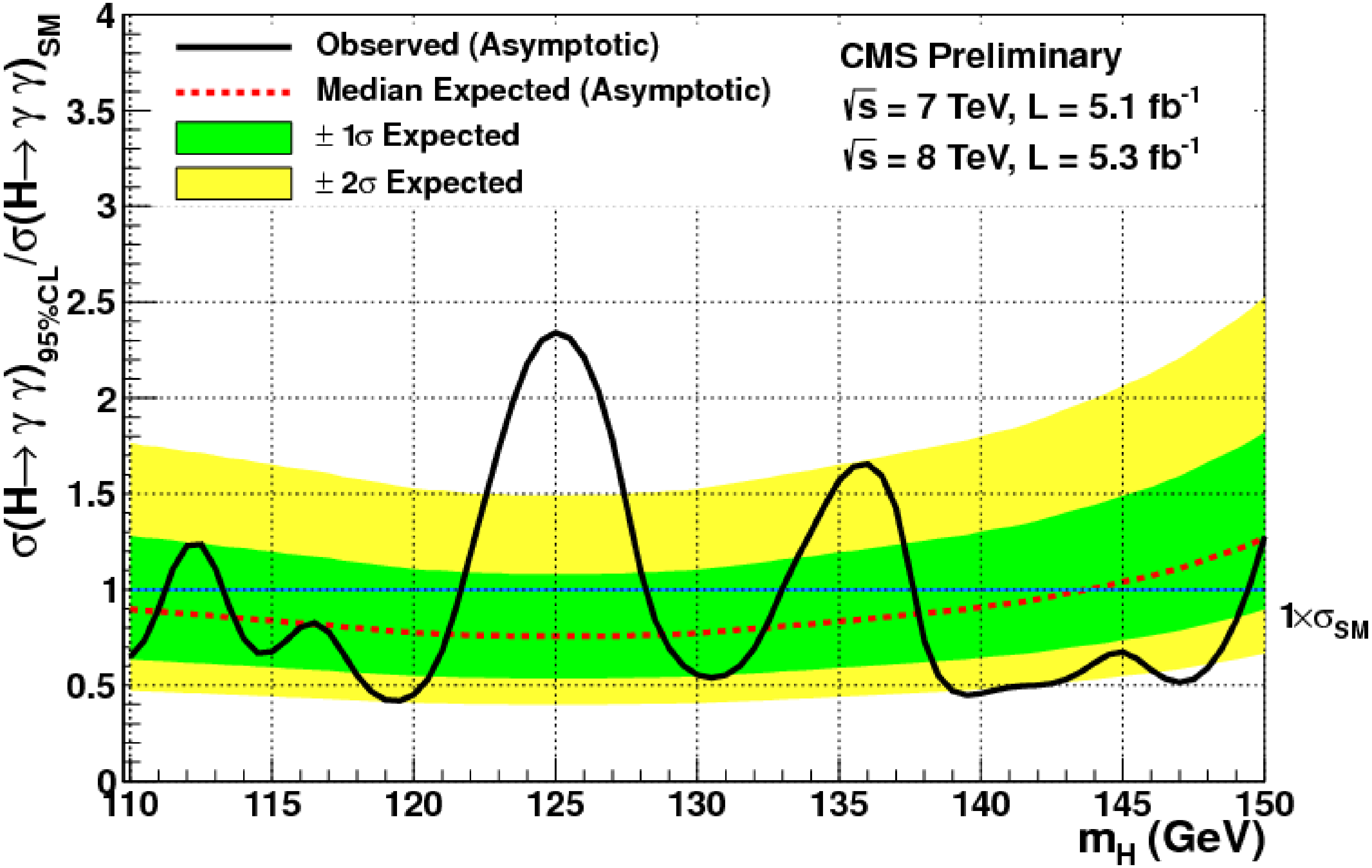


Signal over background in the combined mass spectrum is driven by worse categories.

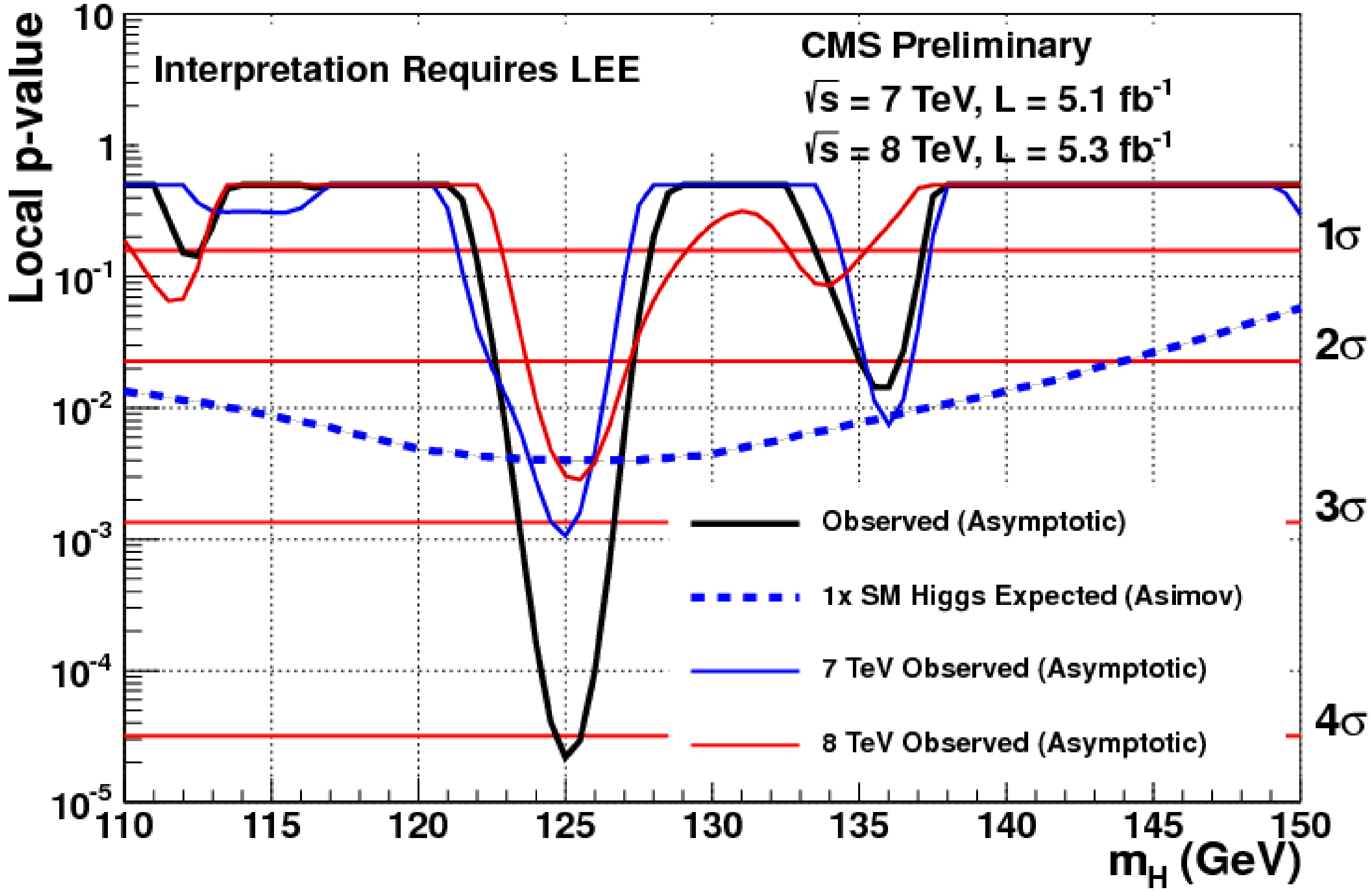
<u>Event weighted mass spectrum with expected S/B makes the peak more visible.</u>

Exclusion Limit

- Reached expected sensitivity 0.76xσ_{SM} at 125 GeV.
- **SM Higgs excluded at 95% CL: [114-121], [129-132], [138-149] GeV.**
- Cross-checked with (independent) cut-based and sideband background model analysis: similar results within experimental uncertainties.



p-Value and Significance

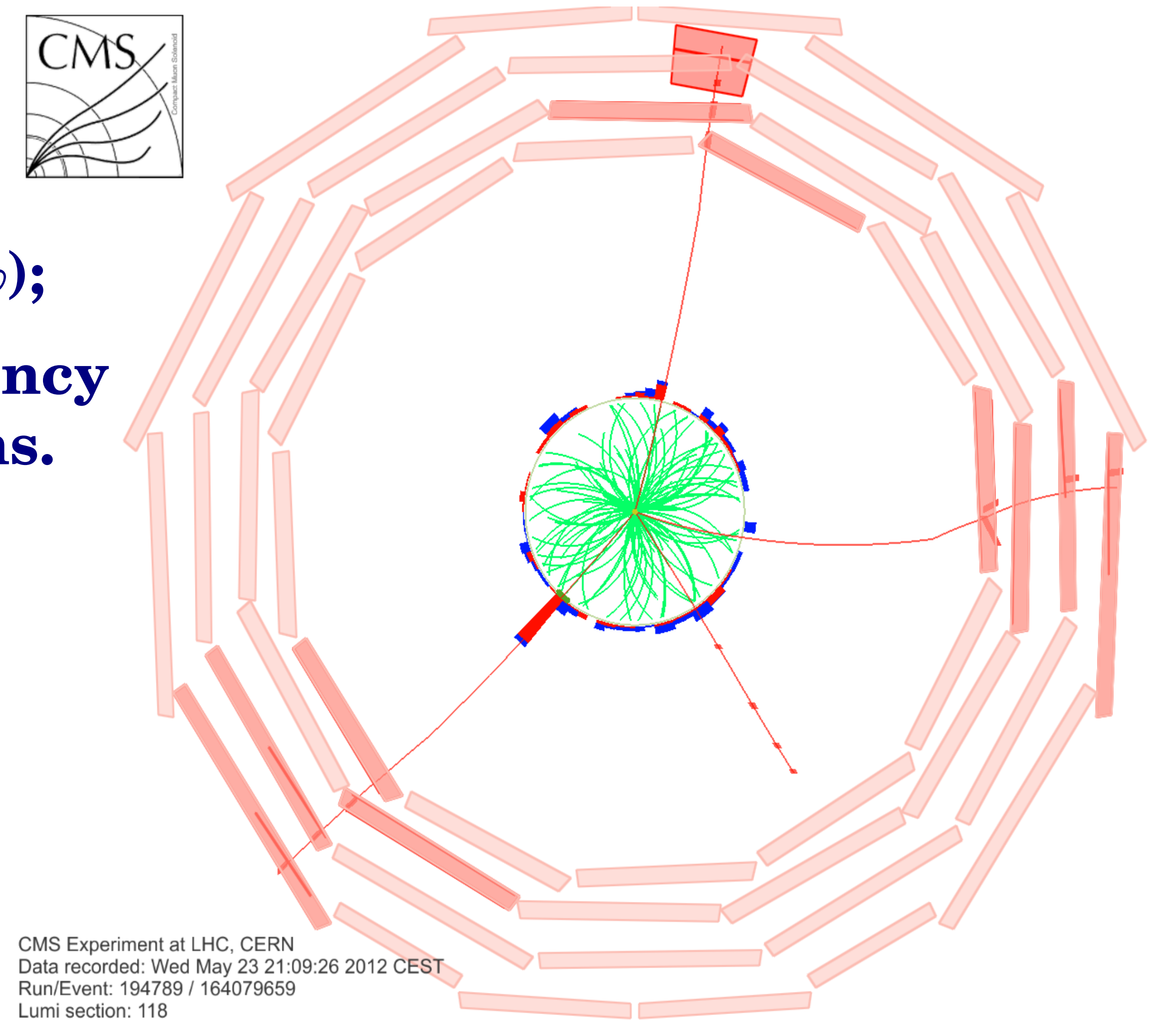


The excess around 125 GeV is both in 7 TeV and 8 TeV data: there is evidence of a narrow resonance around 125 GeV mass at 4.1σ.

Look-elsewhere effect (LEE) correction in the full $H \rightarrow \gamma\gamma$ search range 110-150 GeV gives 3.2σ global significance

$H \rightarrow ZZ^{(*)} \rightarrow 4l$

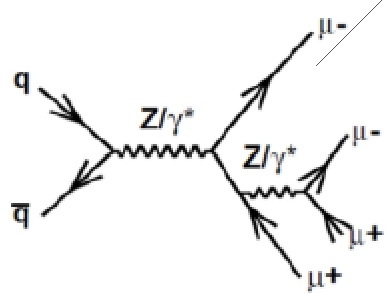
- *"Golden channel"* with a clean experimental signature, four isolated leptons ($2l2\tau$ decay included):
  - Benefits from excellent e and μ resolutions;
  - **Narrow resonance in four lepton mass spectrum (mass resolution 1-2%);**
  - Challenge: **high efficiency for very low p_T leptons.**

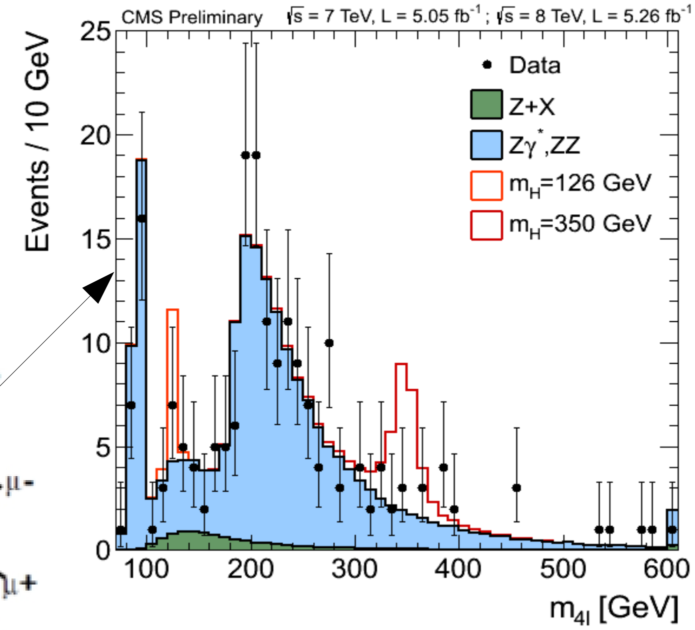


- Background:
  - $ZZ^{(*)}$ (irreducible);
  - Z+jets, $t\bar{t}$ and WZ (reducible).

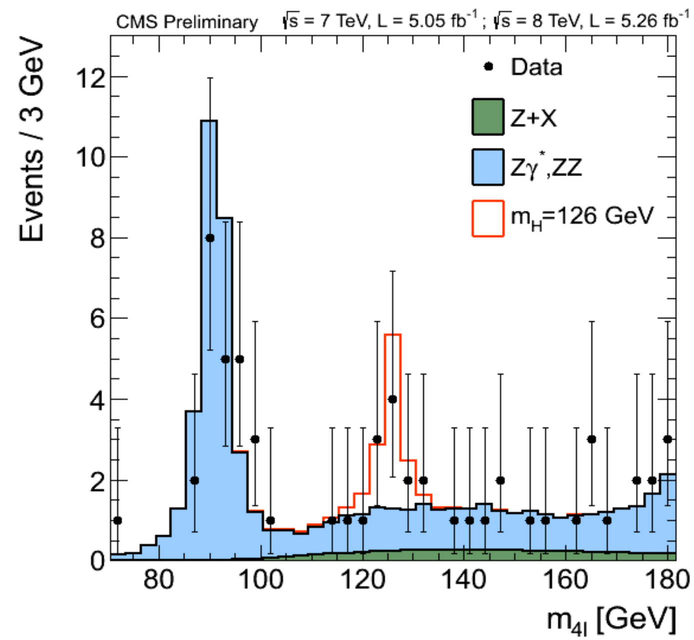
Event Selection and Background Estimation

- **Selection:** four identified and isolated leptons with $p_T^{\mu} > 5$ GeV ($p_T^{e} > 7$ GeV) compatible with PV and FSR photon recovery:
    - Z1 (opposite-sign di-lepton pair closest to MZ), and Z2 (off-shell);
    - Final kinematic cuts: M(4l) > 100 GeV; M(Z2) > 12 GeV





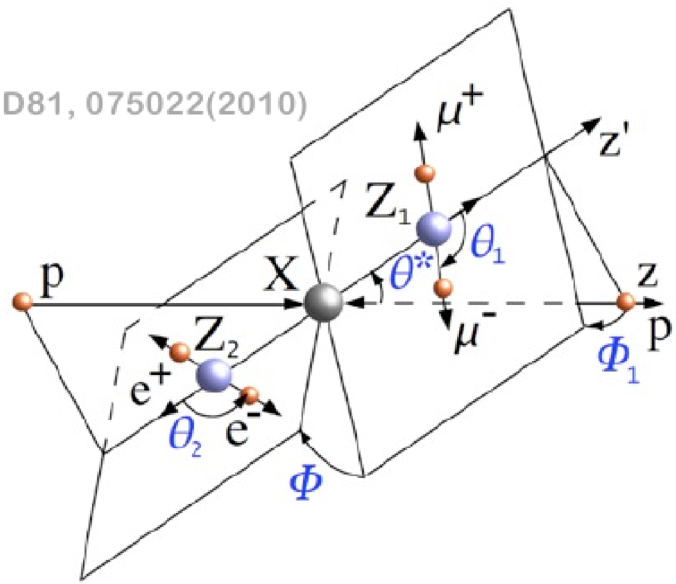
- **Background estimation:**
    - Irreducible $ZZ^{(*)}$: estimated with MC (normalized to theoretical σ_{xsec});
    - Reducible backgrounds: estimated from control samples requiring 2(3) tight leptons and 2(1) "loose" (complementary approach requires same-sign/same-flavour pairs);
        - "fake rate" measured with "loose" leptons.



Statistical Analysis

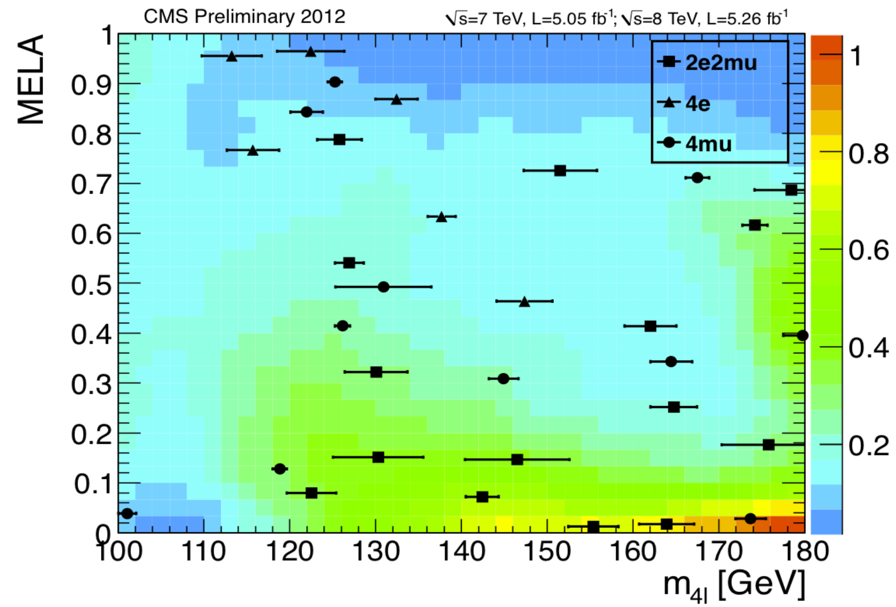
- **MELA** (Matrix Element Likelihood Analysis) allows to discriminate spin 0 particle from background:
  - Decay kinematic fully described by five angles and two masses.

PRD81, 075022(2010)

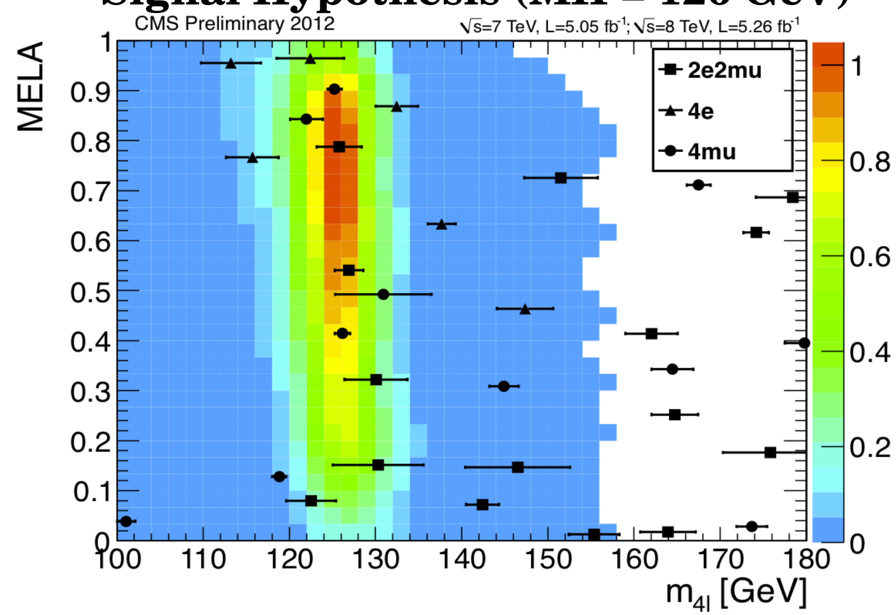


2D Fit performed on (MELA; M4l) distributions

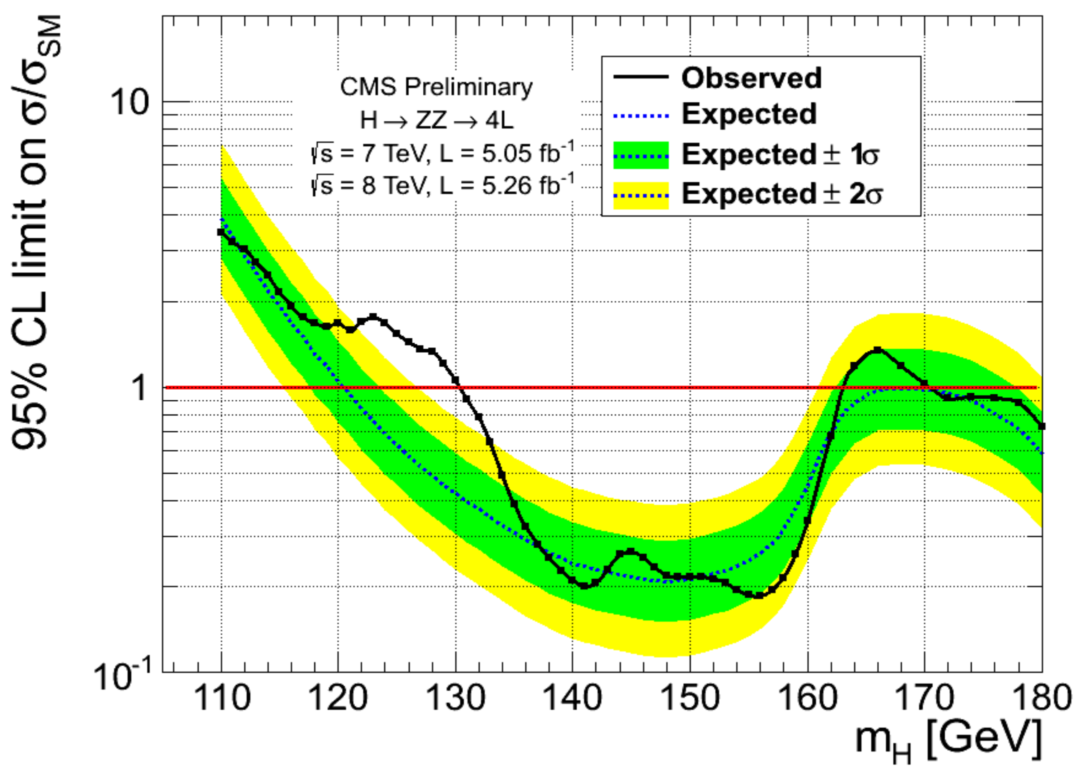
Background Hypothesis

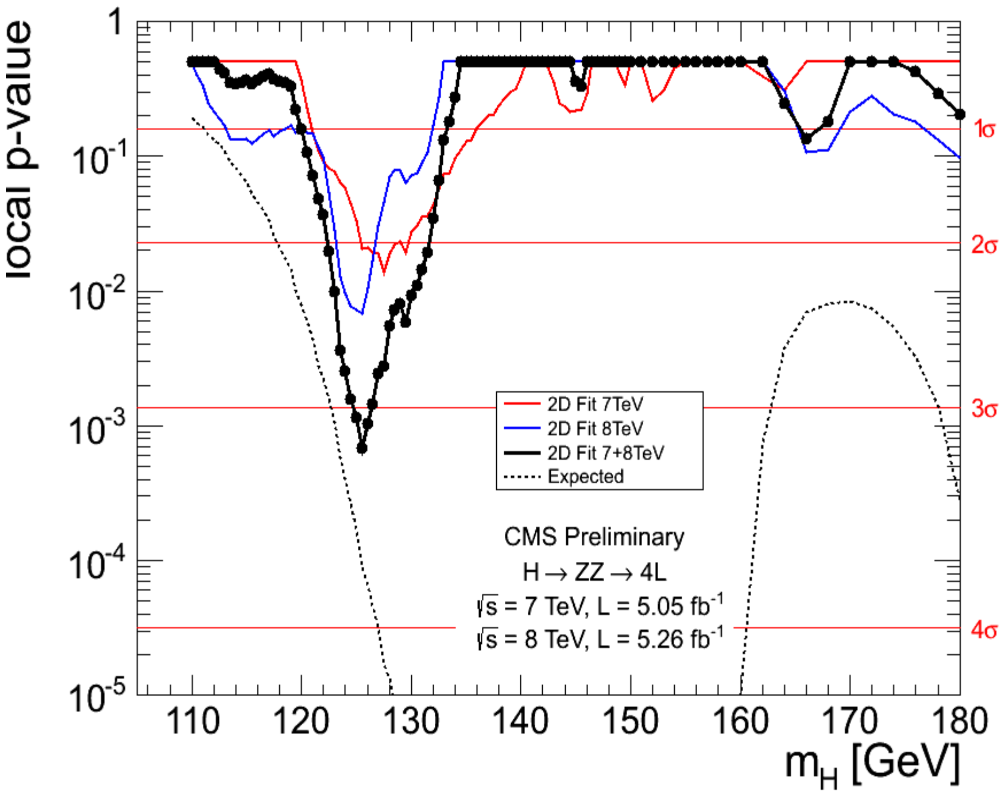


Signal Hypothesis (MH = 126 GeV)



Exclusion Limit and p-Value





Exclusion at 95% CL

Expected: [121-550] GeV

Observed: [131-162] and [172-530] GeV

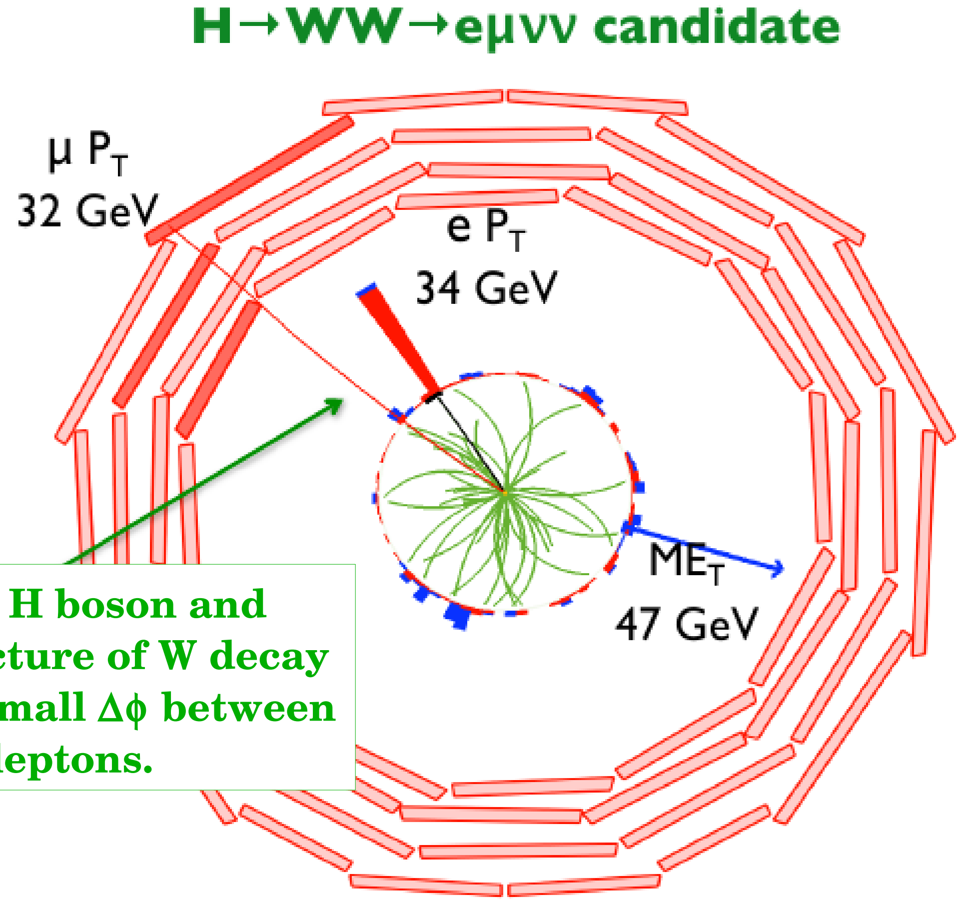
Local Significance at 125.5 GeV

Expected: 3.8σ

Observed: 3.2σ

$H \to WW \to 2l2\nu$

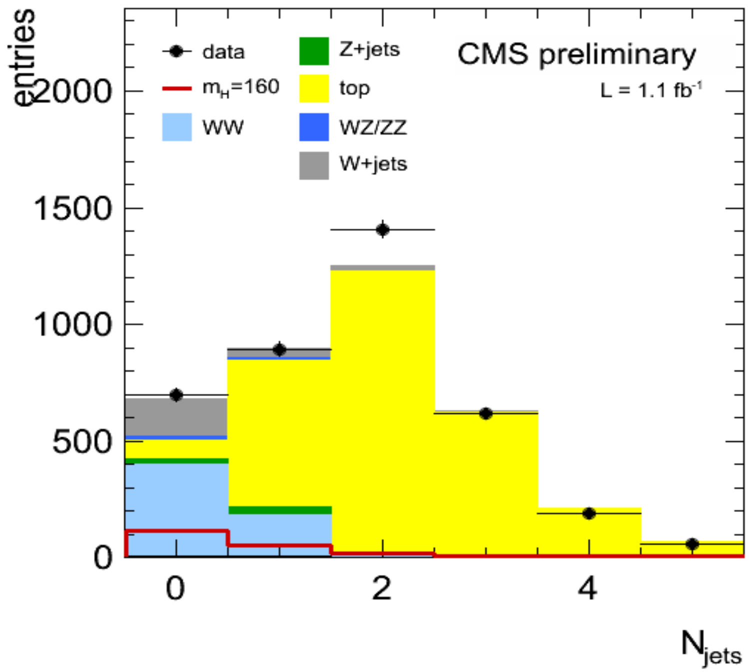
- Final state with the highest rate in most of the Higgs mass search range.
- Signature: two high p_T isolated leptons and MET.
- **No narrow peak:**
  - **Mass resolution ~20%;**
  - **Background estimation is crucial for the analysis (main contributions estimated from data).**
- Background: WW (irreducible), Z+jets, WZ, ZZ, $t\bar{t}$ and W+jets (reducible).



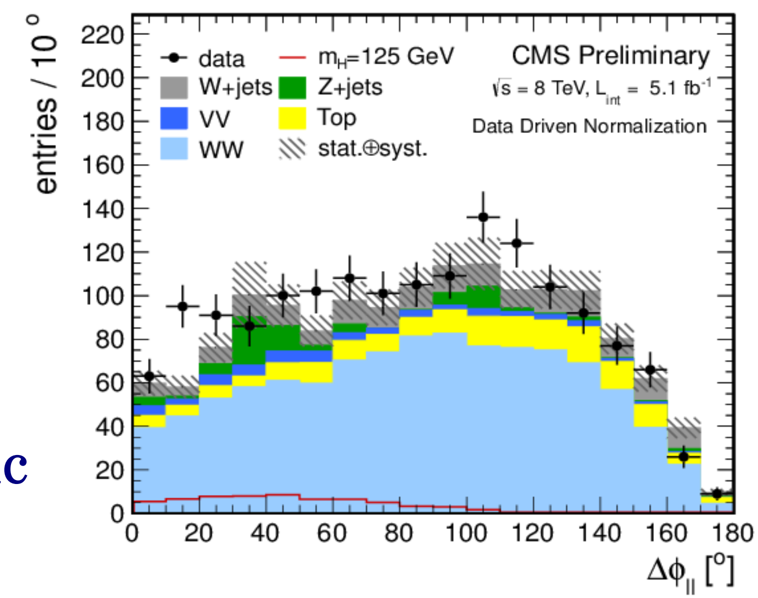
Scalar H boson and V-A structure of W decay favours small Δφ between leptons.

Event Selection

- **Selection:**
  - Two identified oppositely charged isolated leptons compatible with PV.
  - Jet-ID to remove pileup jets, p_T^{ll} (non-prompt leptons), no third lepton (WZ/ZZ), m_{ll} (low mass resonances), Z peak veto and MH hypothesis dependent cuts (DY).

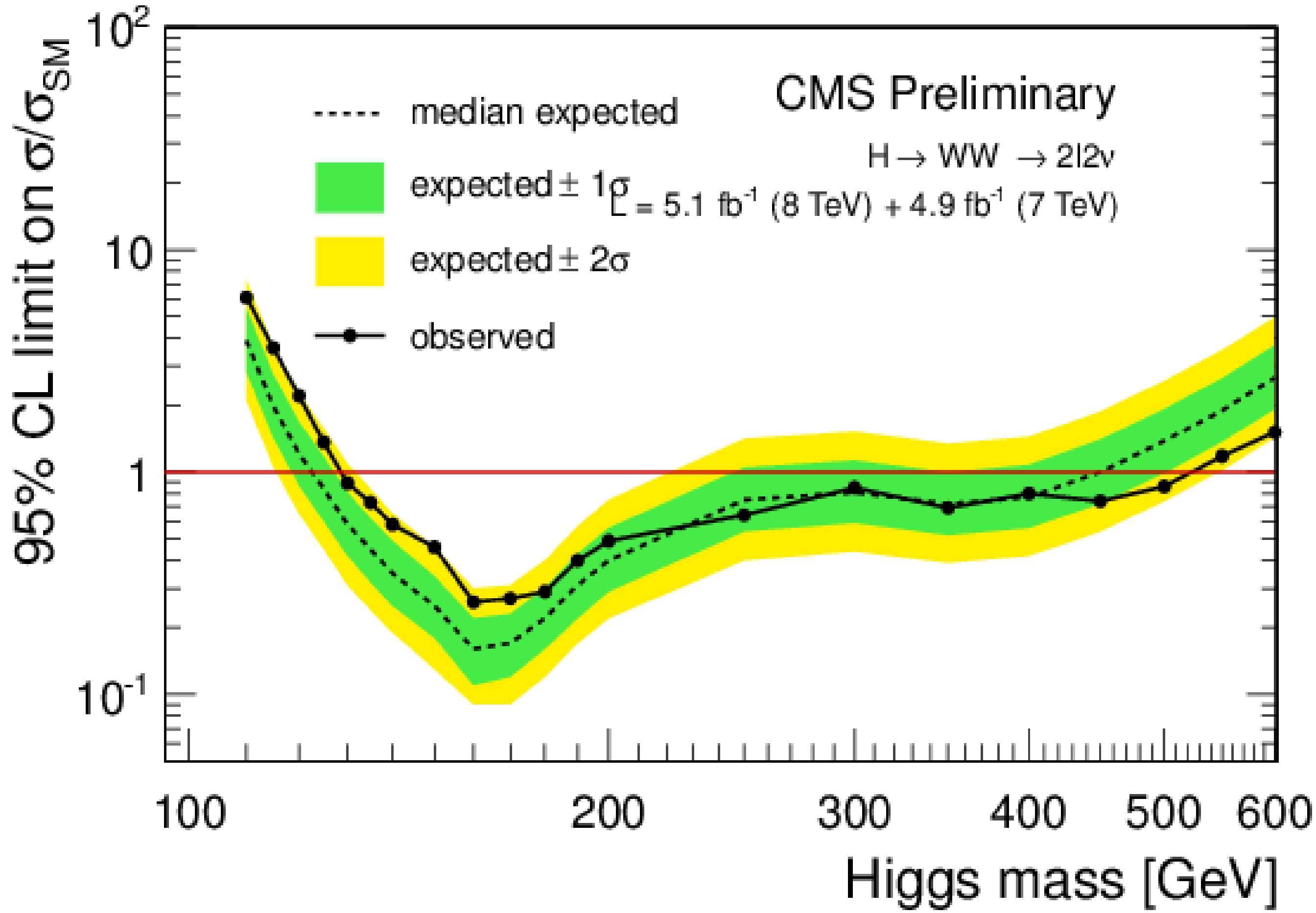


- **Categorize events by jet multiplicity:**
  - **0-jet:** dominated by WW, most sensitive to gg→H;
  - **1-jet:** dominated by $t\bar{t}$ and tW, apply anti b-tag;
  - **2-jet:** specific selection for VBF ($\Delta\eta$(jj), m_{jj} cuts)



- Further discrimination provided by kinematic variables: p_T, m_{ll}, $\Delta\phi_{ll}$, M_T.

Exclusion Limit



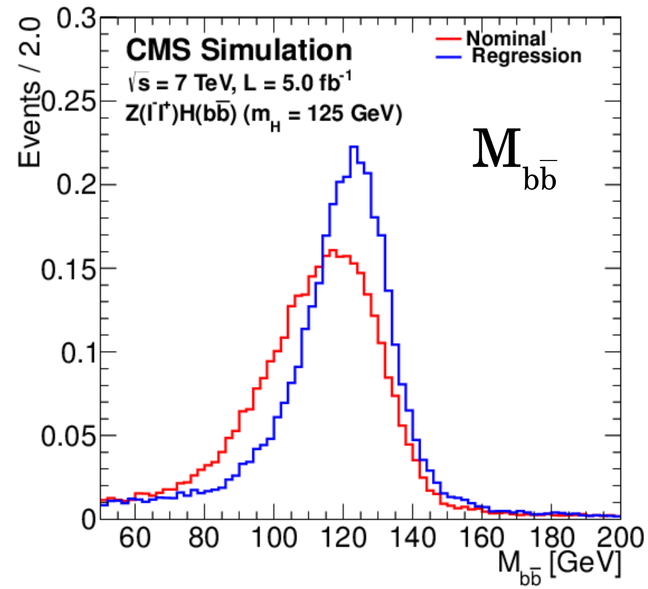
Expected limit: [122-450] GeV

Observed limit: [129-520] GeV

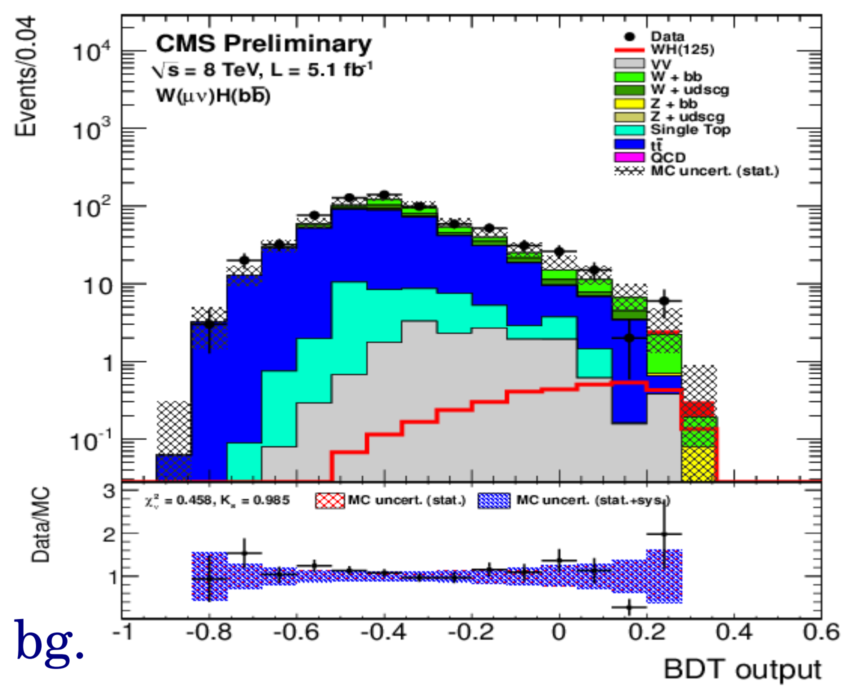
The excess at low mass is still both compatible with SM Higgs Boson with MH=125 GeV and with a fluctuation of the background.

H→b$\bar{b}$

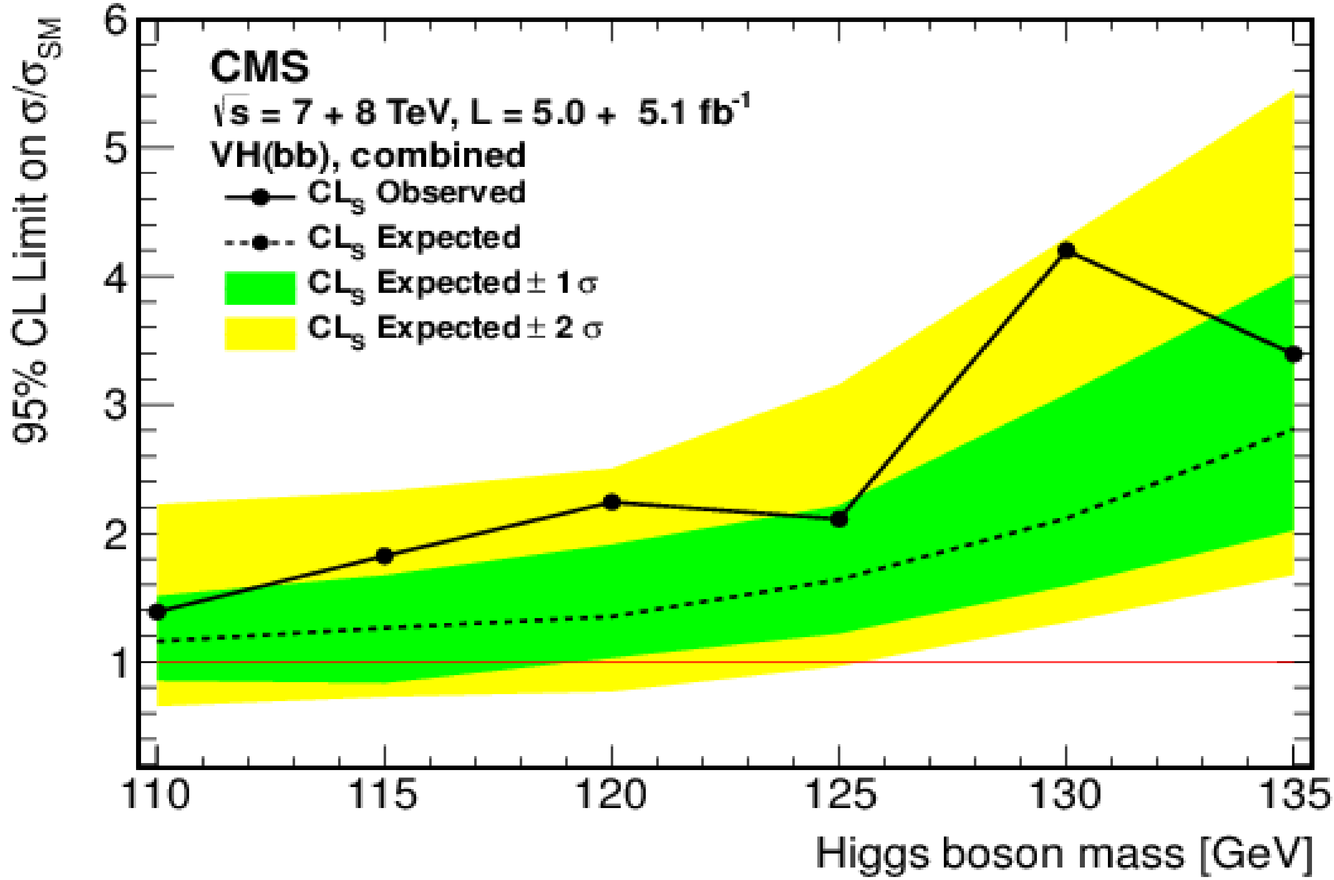
- **Largest BR below 130 GeV** but **very large background** (not possible to use gg→H, additional tag needed):
  - Two exclusive channels: VH and (recently added) t$\bar{t}$H.



- **VH analysis (5 modes Z(ll)H, W(lν)H, Z(νν)H, l = e, μ):**
  - Require high momentum vector boson and 2-b tagged jets back-to-back;
  - Use data control regions to constrain most important backgrounds (V+jets, t$\bar{t}$);
  - b-jet energy regression to improve mass resolution (~15%);
  - Statistical analysis with fit to shape of BDT output used to discriminate signal to bg.



Exclusion Limit (VH)

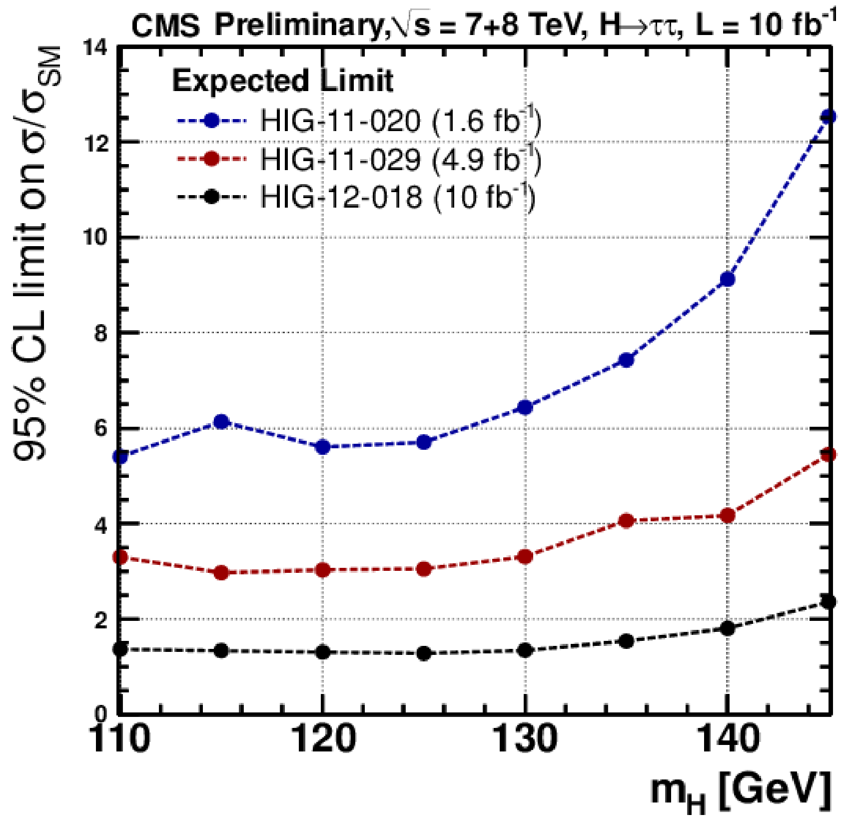


Sensitivity enhanced by 50% due to improvements in the analysis (almost reached SM sensitivity below 115 GeV).

Mild excess between 115 and 135 GeV, expected 1.6xσ_{SM} at 125 GeV, observed 2xσ_{SM}

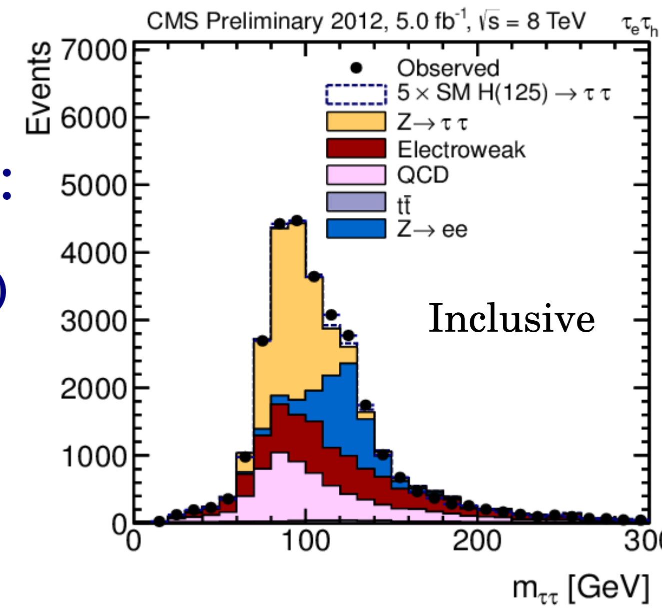
$H \rightarrow \tau\tau$

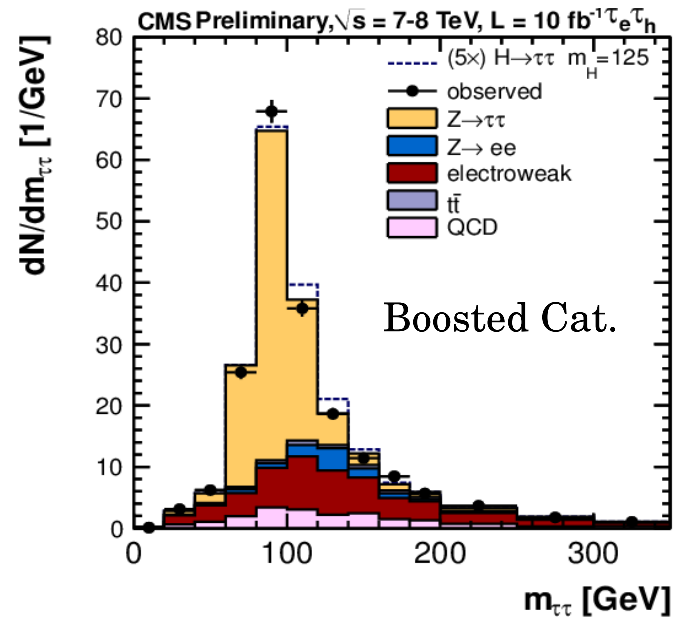
- **Complicated analysis**, combination of many different sub-channels:
  - Four independent final states according to number of τ leptonic decays ($e\tau_h$, $\mu\tau_h$, $e\mu$, $\mu\mu$).
  - **Sensitive to all Higgs production channels;**
  - **Probe of direct coupling to lepton;**
- Largest backgrounds: DY, W+jets, QCD, estimated with data-driven techniques.
- Analysis re-optimized since 2011:
  - Improved lepton and τ ID criteria;
  - New di-τ mass reconstruction gives 20% better resolution.



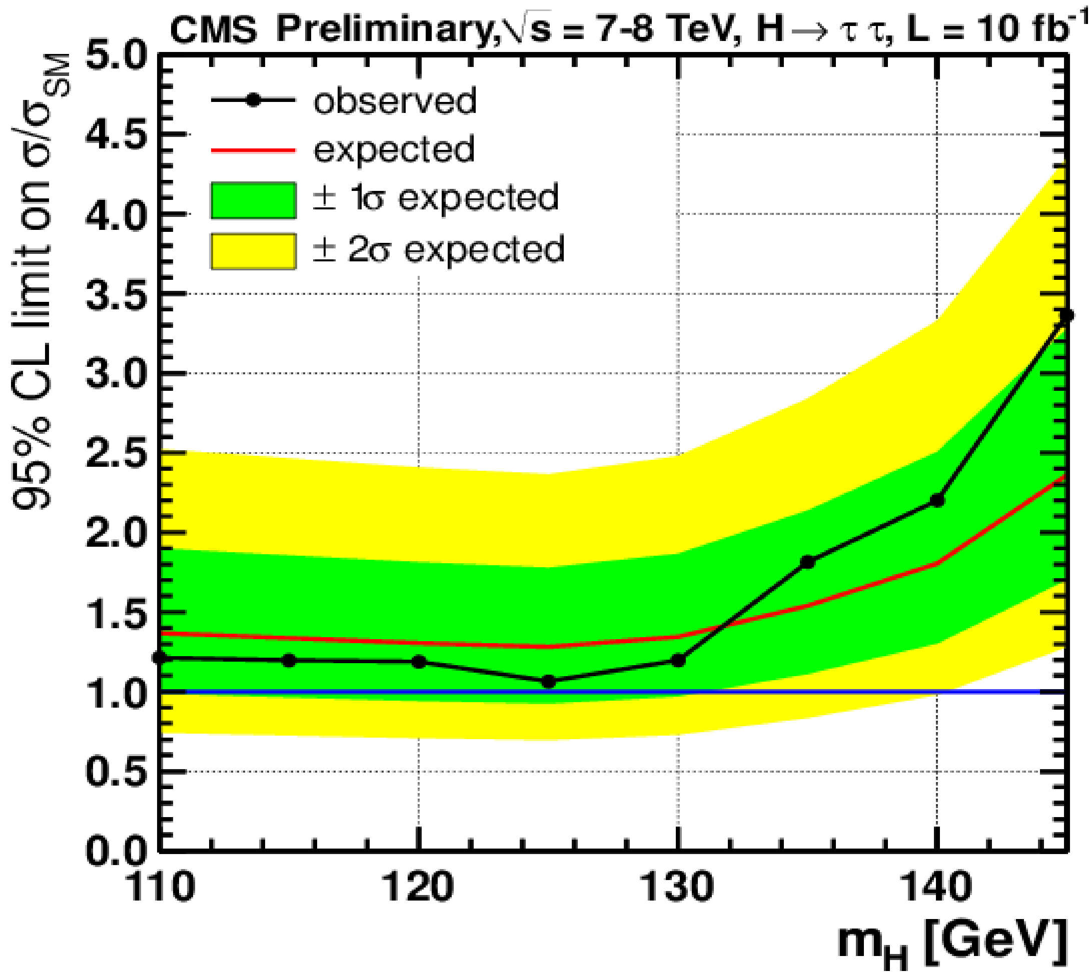
Event Categorization

- To enhance sensitivity, selected events are split in 5 categories based on jet multiplicity and transverse momentum of visible τ decay:
  - **VBF**: require two jets, BDT discriminator (new)
  - **Boosted**: high p_T Higgs Boson recoiling against high-p_T jet. Higher and more precise MET improves di-τ mass resolution (further split in high/low p_T).
  - **0-jet**: all selected events not belonging to previous categories (also split in high/low p_T sub-categories)
- Simultaneous binned maximum likelihood fit to all di-τ invariant mass distribution in each categories.





Exclusion Limit



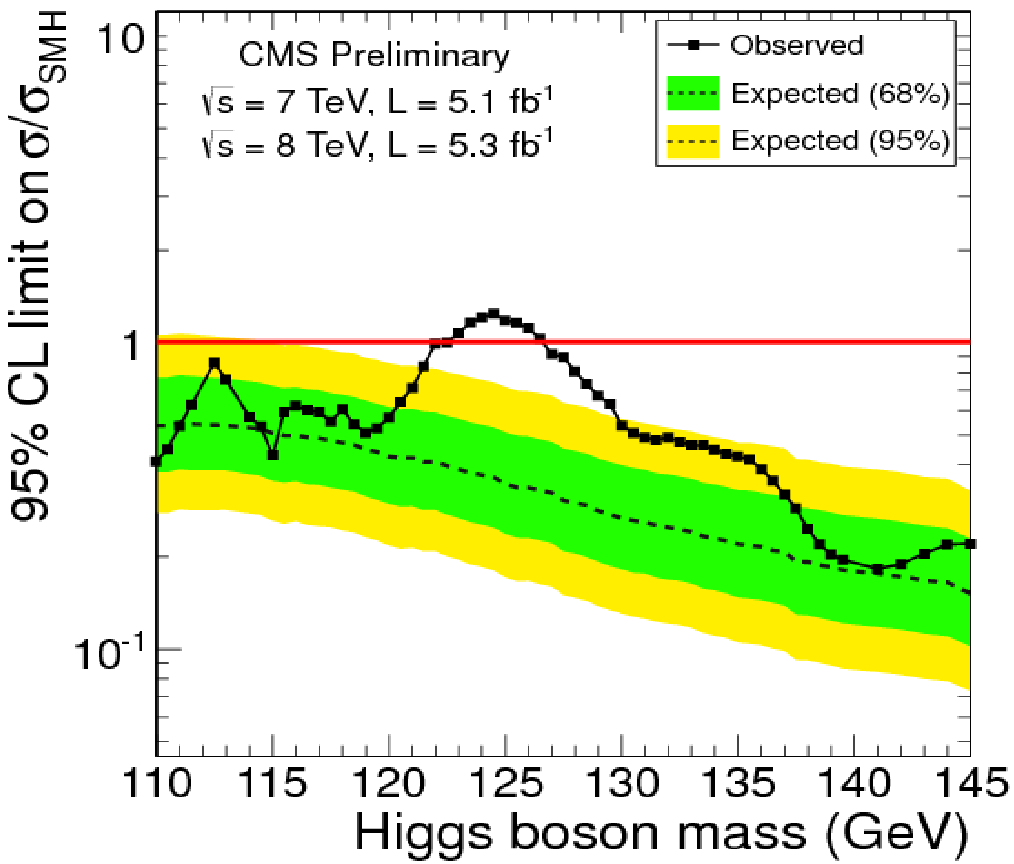
No significant deviation from the background only hypothesis is observed.

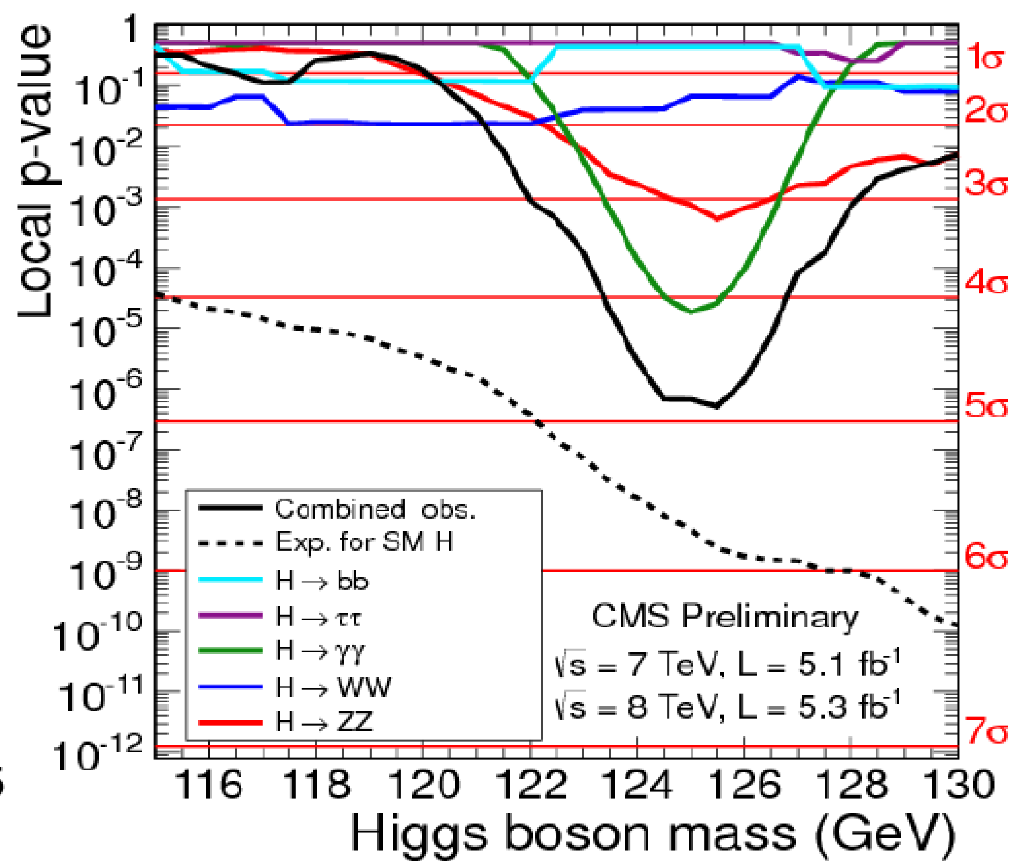
Observed exclusion limit 1.06xσ_{SM} at MH= 125 GeV

Combined Results

Combined Significance

Most of the analysis have used **10 fb^{-1}** and have been improved with respect to 2011.

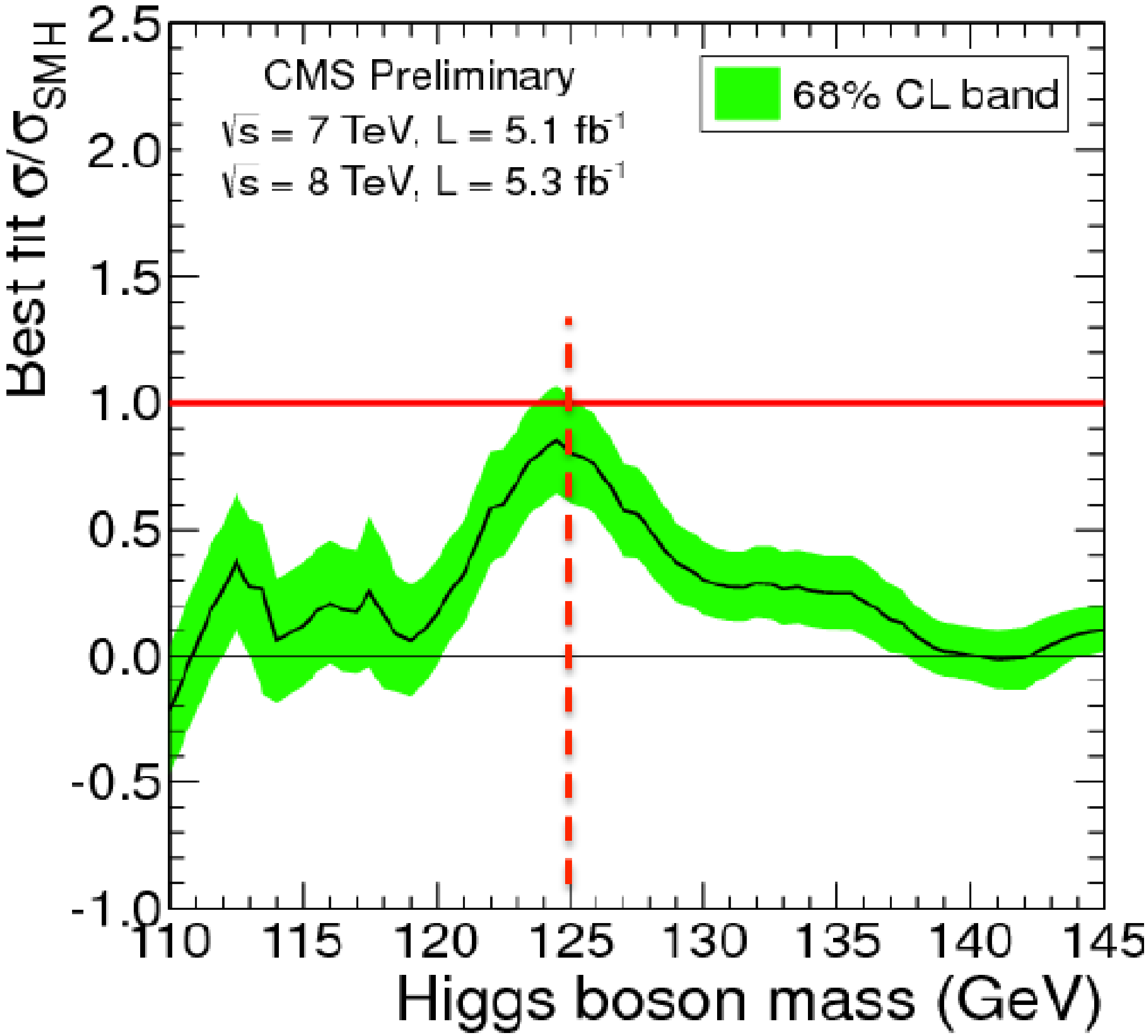




Local significance of the excess 4.9σ (expected significance for SM Higgs signal is 5.9σ).

This excess can be interpreted as the observation of a new Boson with mass around 125 GeV.

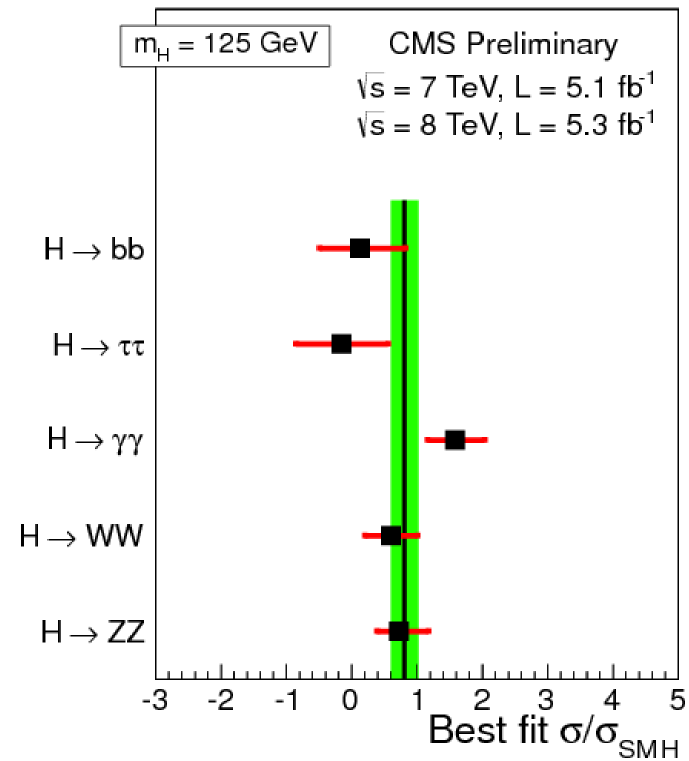
Signal Strength



Best fit signal strength:

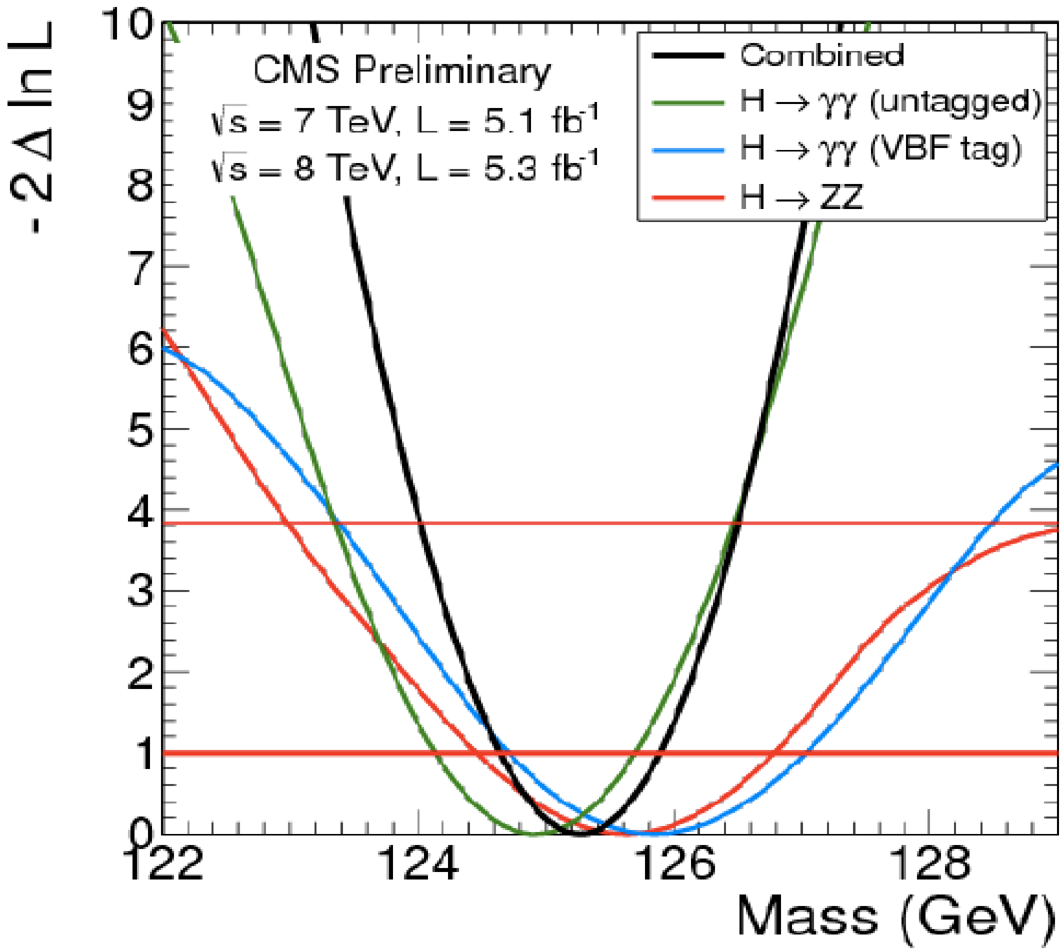
$(0.80 \pm 0.22) \times \sigma_{SMH}$

Compatible with the expectations from a SM Higgs Boson signal.



Mass Measurement

- Perform a fit of the mass with freely floating signal strength to minimize model dependence.



Systematical uncertainty on the mass driven by energy scale uncertainty in γγ (for the moment conservative 0.5% will improve in the future).

M = 125.3 ± 0.4 (stat.) ± 0.5 (syst.) = 125.3 ± 0.6 GeV

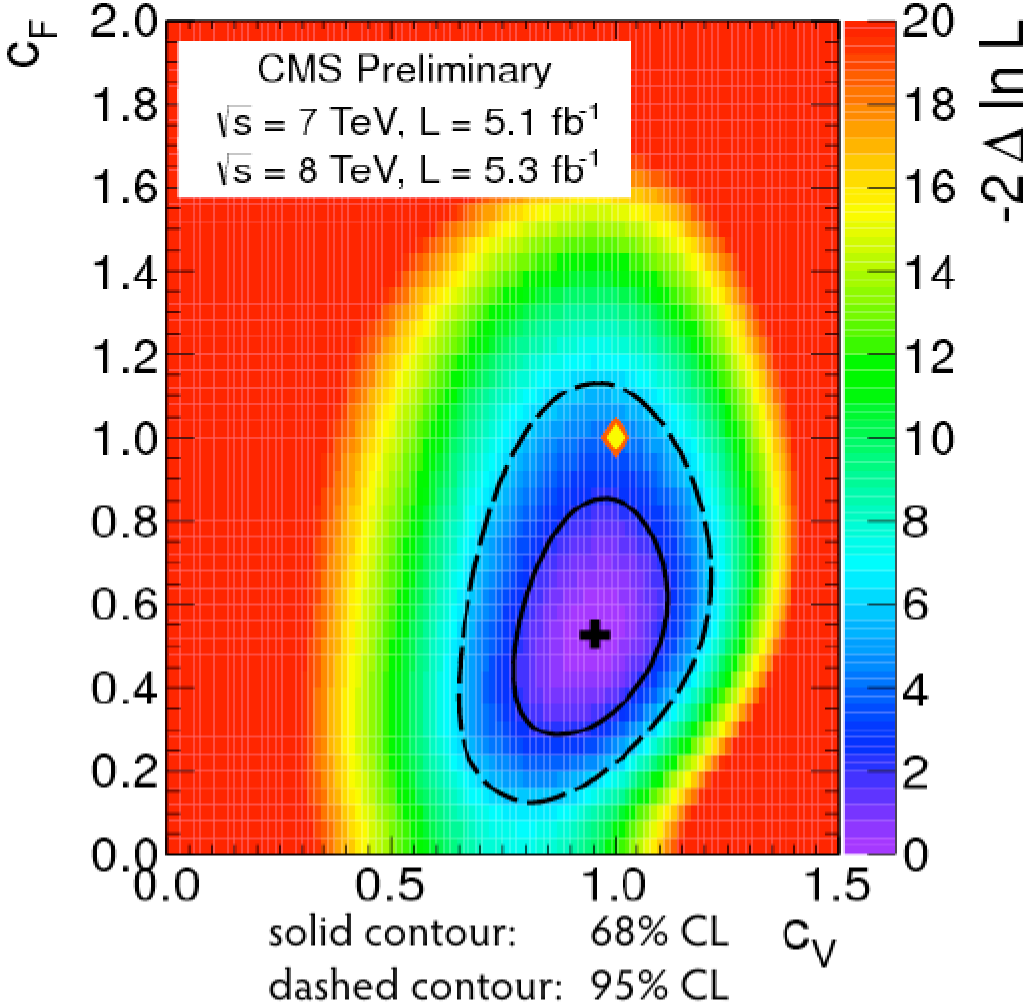
Couplings

- Test compatibility w.r.t. SM predictions by introducing two parameters (c_V, c_F) modifying the expected signal yields in each mode through LO expressions.

CMS data compatible with SM prediction at 95% CL.

Best fit c_F driven to low values by VBF $\gamma\gamma$ excess and $\tau\tau$ deficit.

More data is needed to draw any definite conclusion.

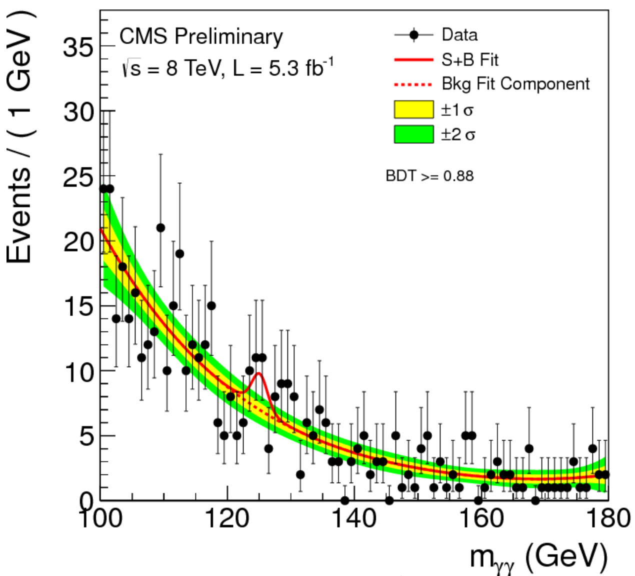


Summary

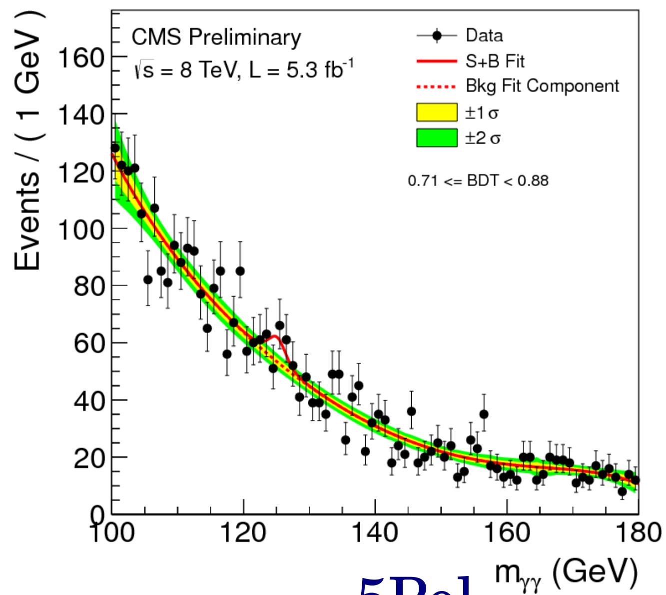
- In the searches for a SM Higgs Boson at CMS, **a new state with mass 125.3 ± 0.6 GeV has been observed,** dominantly in the γγ and 4l modes.
- Within the limited precision of the current data the observation is compatible with the predictions for a SM Higgs Boson signal, despite the larger excess in γγ and the deficit in ττ, $b\bar{b}$ modes.
- **More data is needed to draw any conclusions on the nature of this new particle** and in particular to:
  - Improve mass measurement;
  - Measure spin and parity using angular distributions in ZZ, WW and γγ;
  - Search for deviations from the SM in the couplings.

Backup

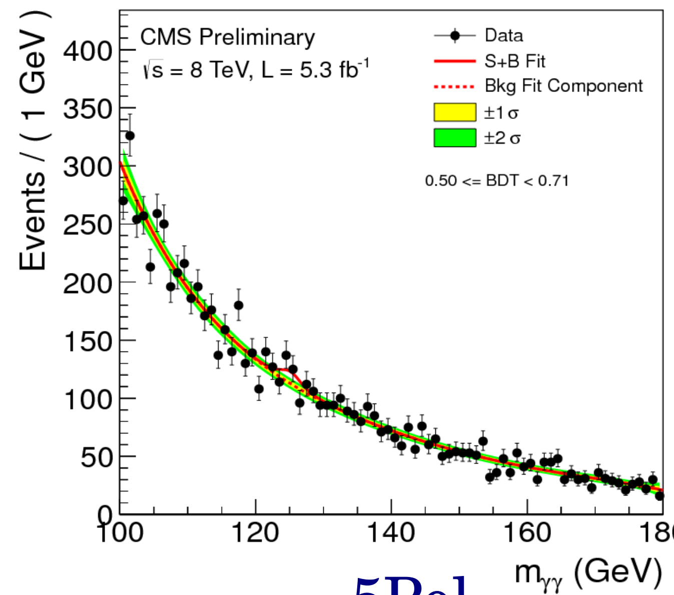
Invariant Mass (H→γγ)



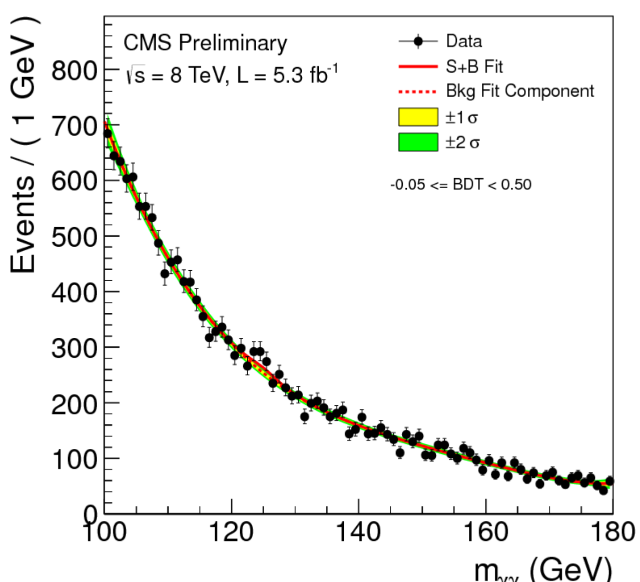
4Pol



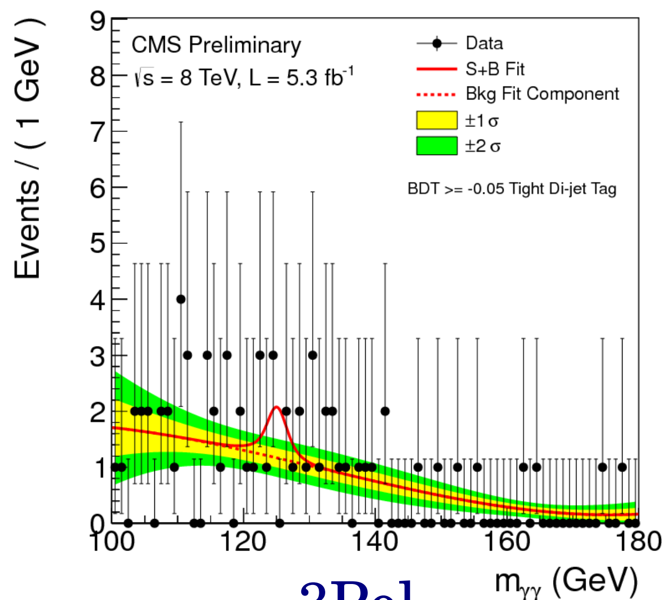
5Pol



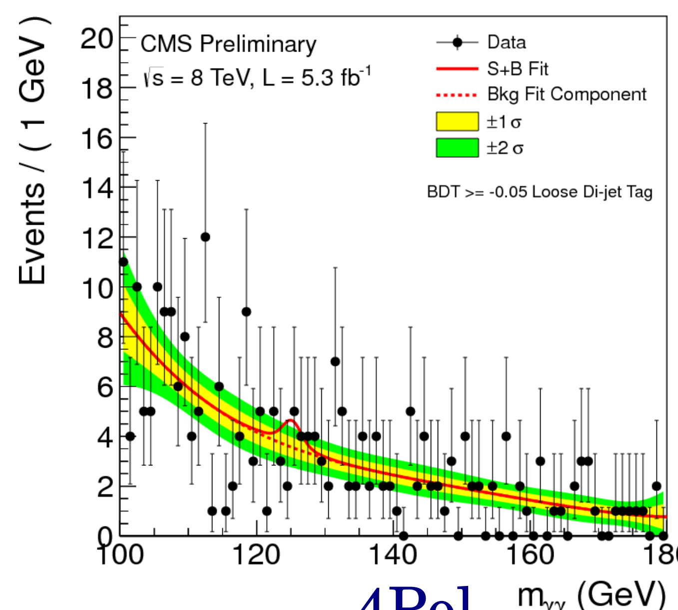
5Pol



5Pol

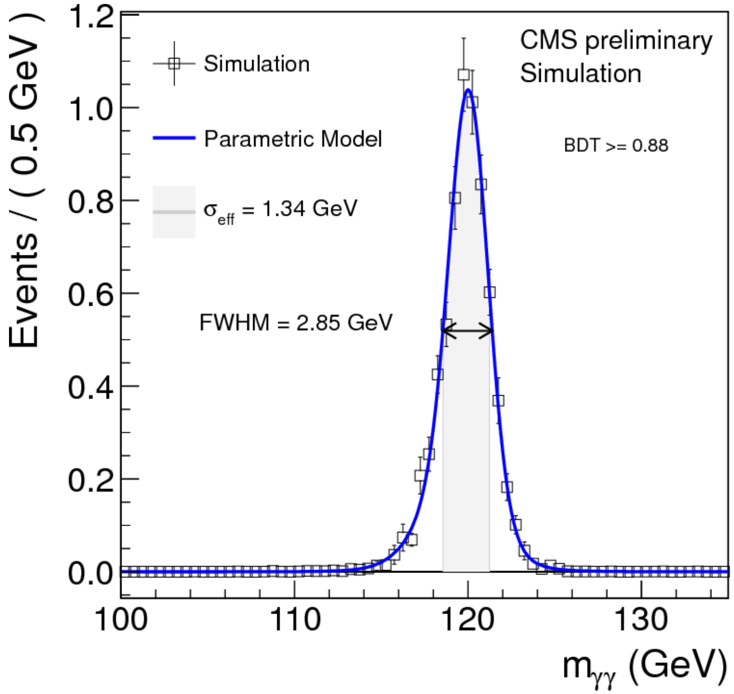


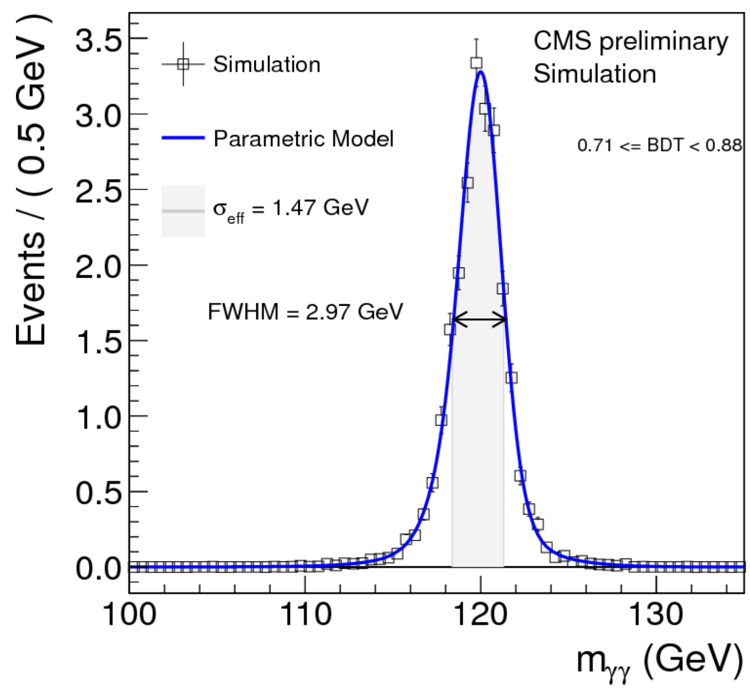
3Pol



4Pol

Mass Resolution (H→γγ)





Parametric model:
sum of Gaussians

Expected signal and estimated background

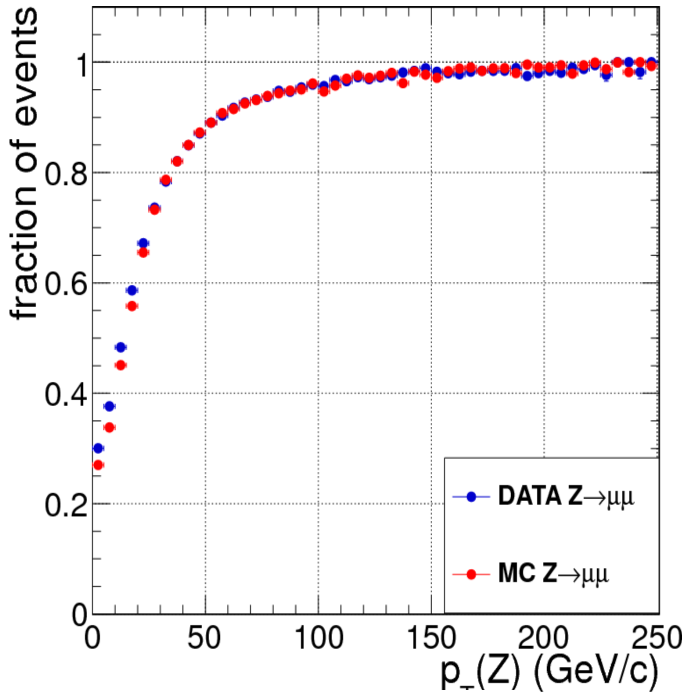
Event classes		SM Higgs boson expected signal (m_H=125 GeV)							Background $m_{\gamma\gamma}$ = 125 GeV (ev./GeV)
		Total	ggH	VBF	VH	ttH	σ_{eff} (GeV)	FWHM/2.35 (GeV)	
7 TeV 5.1 fb^{-1}	Untagged 0	3.2	61%	17%	19%	3%	1.21	1.14	3.3 ± 0.4
	Untagged 1	16.3	88%	6%	6%	1%	1.26	1.08	37.5 ± 1.3
	Untagged 2	21.5	91%	4%	4%	–	1.59	1.32	74.8 ± 1.9
	Untagged 3	32.8	91%	4%	4%	–	2.47	2.07	193.6 ± 3.0
	Dijet tag	2.9	27%	73%	1%	–	1.73	1.37	1.7 ± 0.2
8 TeV 5.3 fb^{-1}	Untagged 0	6.1	68%	12%	16%	4%	1.38	1.23	7.4 ± 0.6
	Untagged 1	21.0	88%	6%	6%	1%	1.53	1.31	54.7 ± 1.5
	Untagged 2	30.2	92%	4%	3%	–	1.94	1.55	115.2 ± 2.3
	Untagged 3	40.0	92%	4%	4%	–	2.86	2.35	256.5 ± 3.4
	Dijet tight	2.6	23%	77%	–	–	2.06	1.57	1.3 ± 0.2
	Dijet loose	3.0	53%	45%	2%	–	1.95	1.48	3.7 ± 0.4

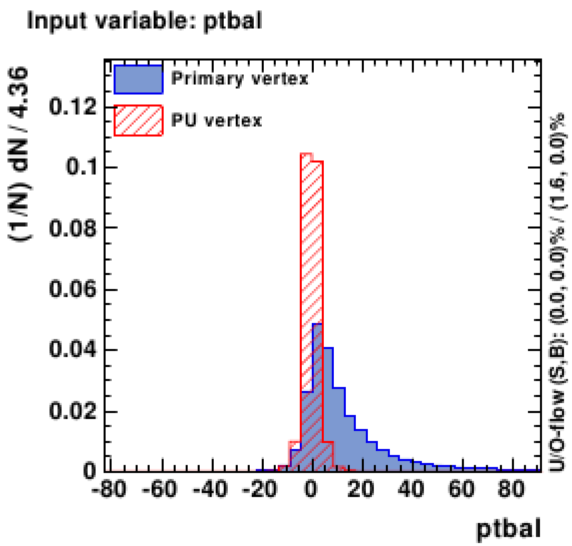
Vertex Selection (H→γγ)

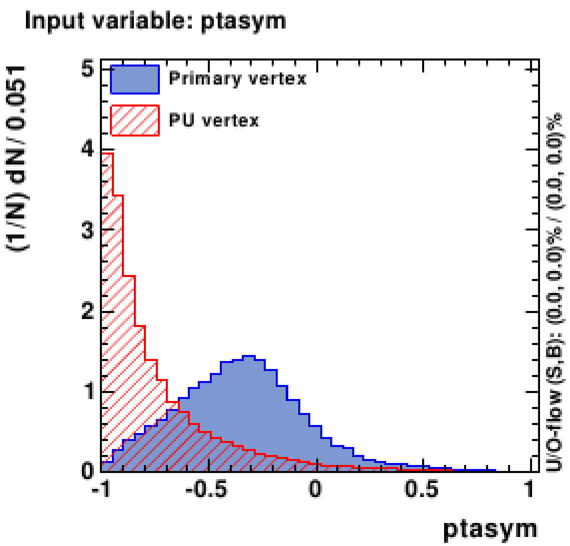
sumpt2: $\sum_i |\vec{p}_T^{\,i}|^2$.

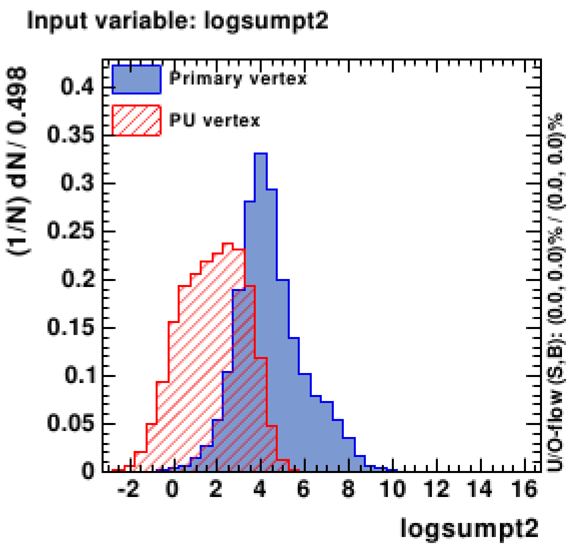
ptbal: $-\sum_i (\vec{p}_T^{\,i} \cdot \frac{\vec{p}_T^{\,\gamma\gamma}}{|\vec{p}_T^{\,\gamma\gamma}|})$.

ptasym: $(|\sum_i \vec{p}_T^{\,i}| - p_T^{\gamma\gamma}) \,/\, (|\sum_i \vec{p}_T^{\,i}| + p_T^{\gamma\gamma})$.





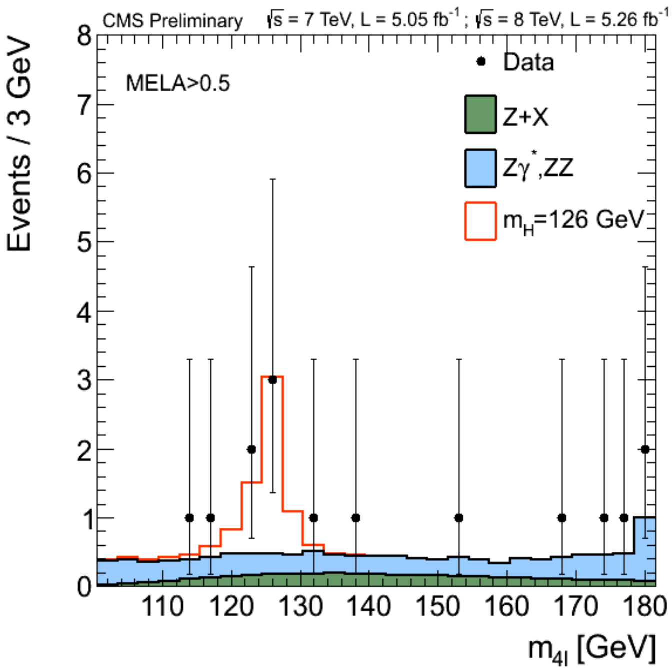


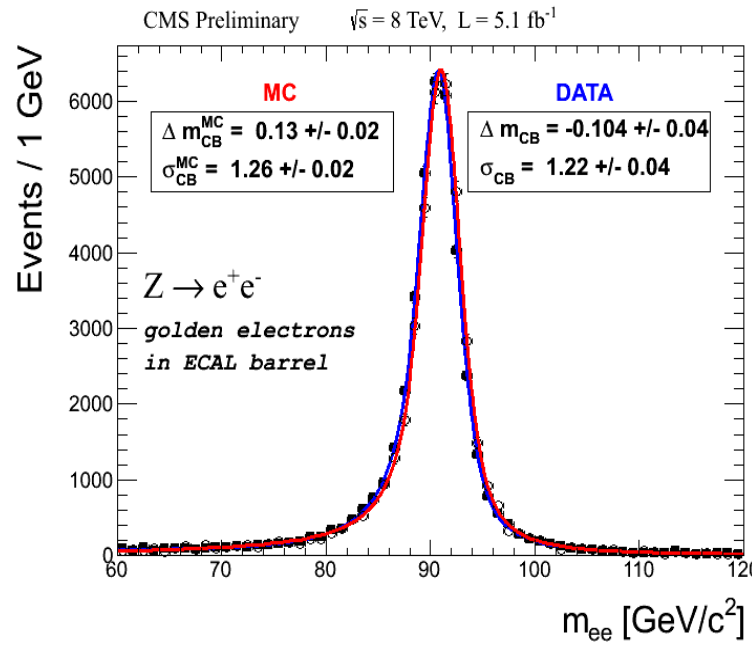


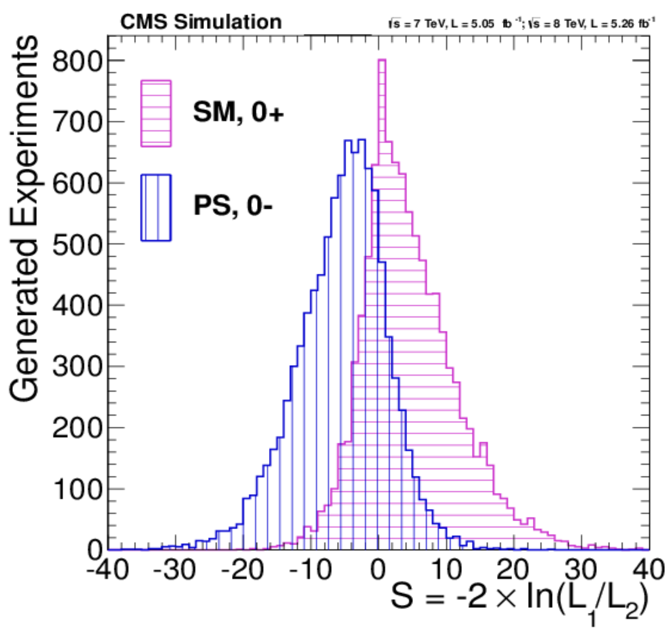
Systematics (H→γγ)

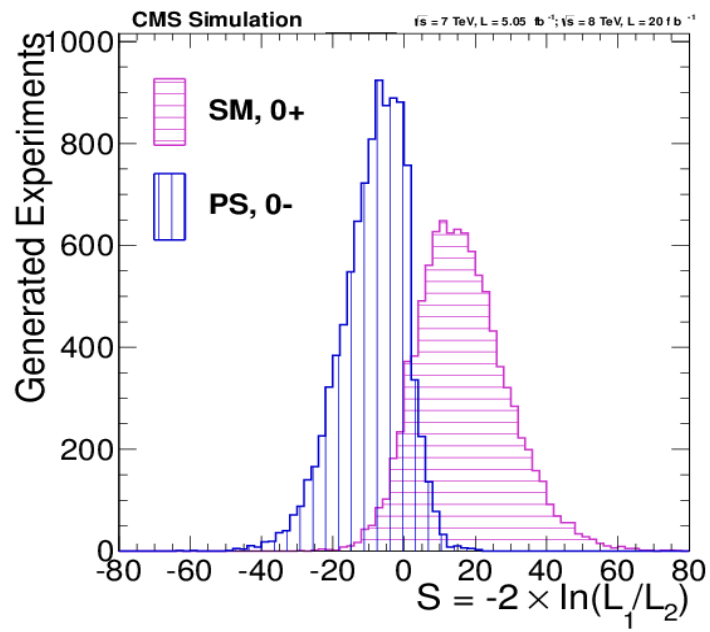
Sources of systematic uncertainty		Uncertainty	
Per photon		Barrel	Endcap
Photon selection efficiency		0.8%	2.2%
Energy resolution ($\Delta\sigma/E_{MC}$)	$R_9 > 0.94$ (low η, high η)	0.22%, 0.60%	0.90%, 0.34%
	$R_9 < 0.94$ (low η, high η)	0.24%, 0.59%	0.30%, 0.52%
Energy scale ($(E_{data} - E_{MC})/E_{MC}$)	$R_9 > 0.94$ (low η, high η)	0.19%, 0.71%	0.88%, 0.19%
	$R_9 < 0.94$ (low η, high η)	0.13%, 0.51%	0.18%, 0.28%
Photon identification BDT	(Effect of up to 4.3% event class migration.)	$\pm$0.01 (shape shift)	
Photon energy resolution BDT	(Effect of up to 8.1% event class migration.)	$\pm$10% (shape scaling)	
Per event			
Integrated luminosity		4.4%	
Vertex finding efficiency		0.2%	
Trigger efficiency	One or both photons $R_9 < 0.94$ in endcap	0.4%	
	Other events	0.1%	
Dijet selection			
Dijet-tagging efficiency	VBF process	10%	
	Gluon-gluon fusion process	50%	
	(Effect of up to 15% event migration among dijet classes.)		
Production cross sections		Scale	PDF
Gluon-gluon fusion		+12.5% -8.2%	+7.9% -7.7%
Vector boson fusion		+0.5% -0.3%	+2.7% -2.1%
Associated production with W/Z		1.8%	4.2%
Associated production with $t\bar{t}$		+3.6% -9.5%	8.5%
Scale and PDF uncertainties	(Effect of up to 12.5% event class migration.)	(y, p_T)-differential	

H→ZZ(*)→4l


CMS Preliminary
Events / 3 GeV
MELA>0.5
Data
Z+X
Zγ*,ZZ
m_H=126 GeV
m_4l [GeV]


CMS Preliminary
√s = 8 TeV, L = 5.1 fb^-1
Events / 1 GeV
MC
Δ m_CB^MC = 0.13 +/- 0.02
σ_CB^MC = 1.26 +/- 0.02
DATA
Δ m_CB = -0.104 +/- 0.04
σ_CB = 1.22 +/- 0.04
Z → e+e-
golden electrons
in ECAL barrel
m_ee [GeV/c^2]


CMS Simulation
Generated Experiments
SM, 0+
PS, 0-
S = -2 × ln(L1/L2)

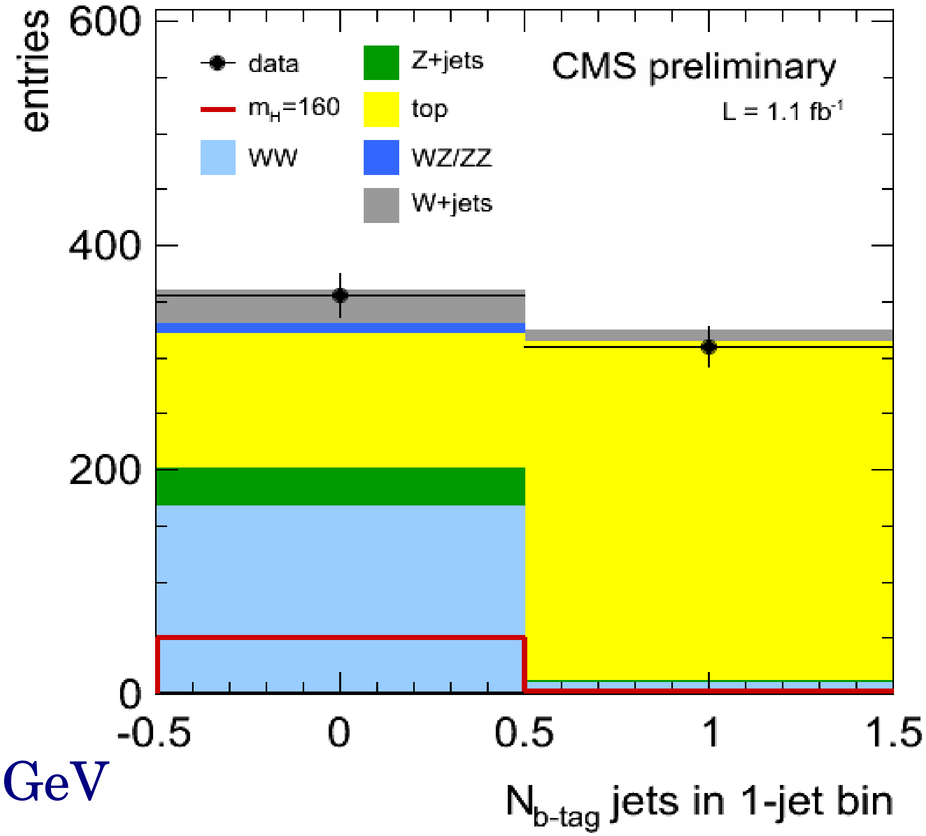

CMS Simulation
Generated Experiments
SM, 0+
PS, 0-
S = -2 × ln(L1/L2)

Systematics (H→ZZ)

Trigger	1.5%
Lepton ID	2-3%
Th ID and Iso	6%
Th energy scale	3%
E-p calibration	0.5%
Luminosity	4.5%
H ssec	15-20%
BR	2%

Background Estimation (H→WW)

- A combination of techniques is used to determine background contamination after the selection (main background estimated from data).
  - W+jets: from fake-enriched sample ("tight-fail") weighted for fake→lepton probability
  - $t\bar{t}$: from b-tagged events
  - DY: from events under Z-peak
  - WW: from signal free region (m_{ll} > 100 GeV for MH<200 GeV. For high MH simulation is used).



- Cut and count analysis whose systematic uncertainty is dominated by background normalization:
  - W+jets (~40% fake→lepton eff.), DY (~60% statistical uncertainty of control sample), $t\bar{t}$ (~25% b-tag), WW (~15%).

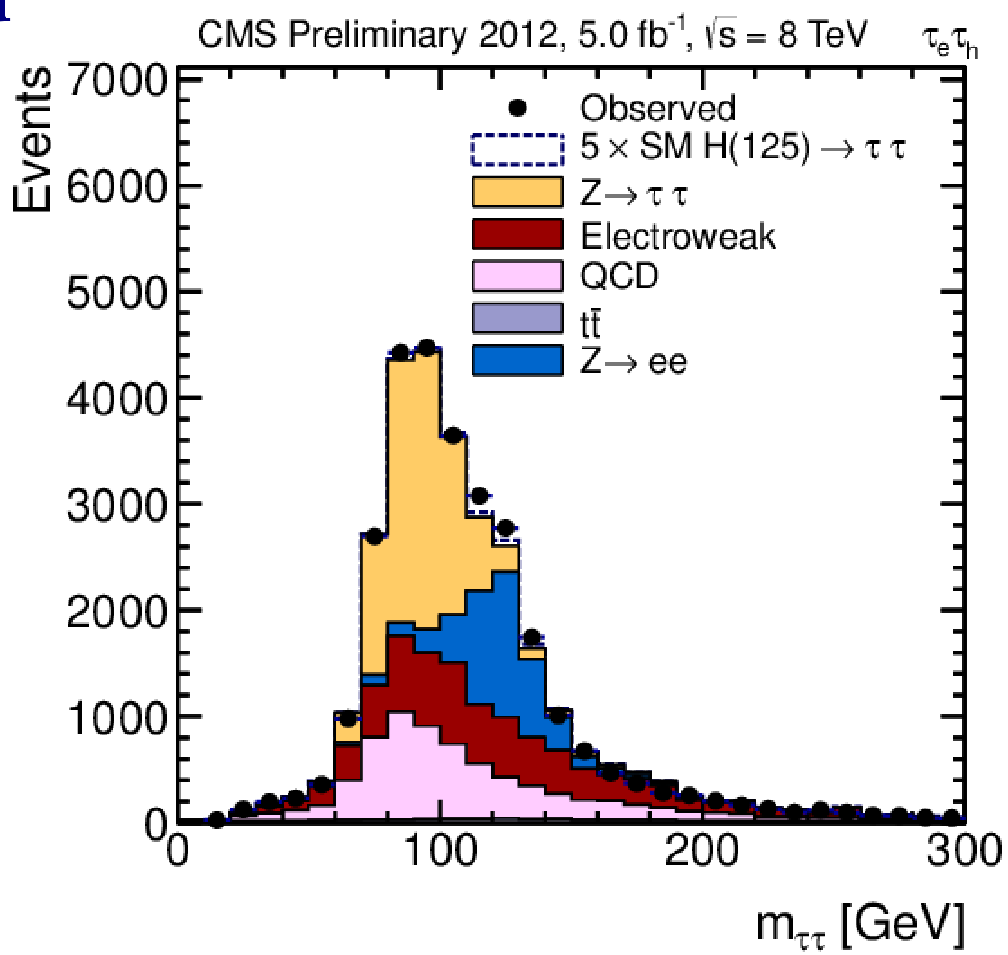
Background Estimation

- Z→ττ: to avoid MC mis-modelling estimated from observed sample of Z→μμ using *embedding technique*. Normalization determined from Z→μμ yield in data.

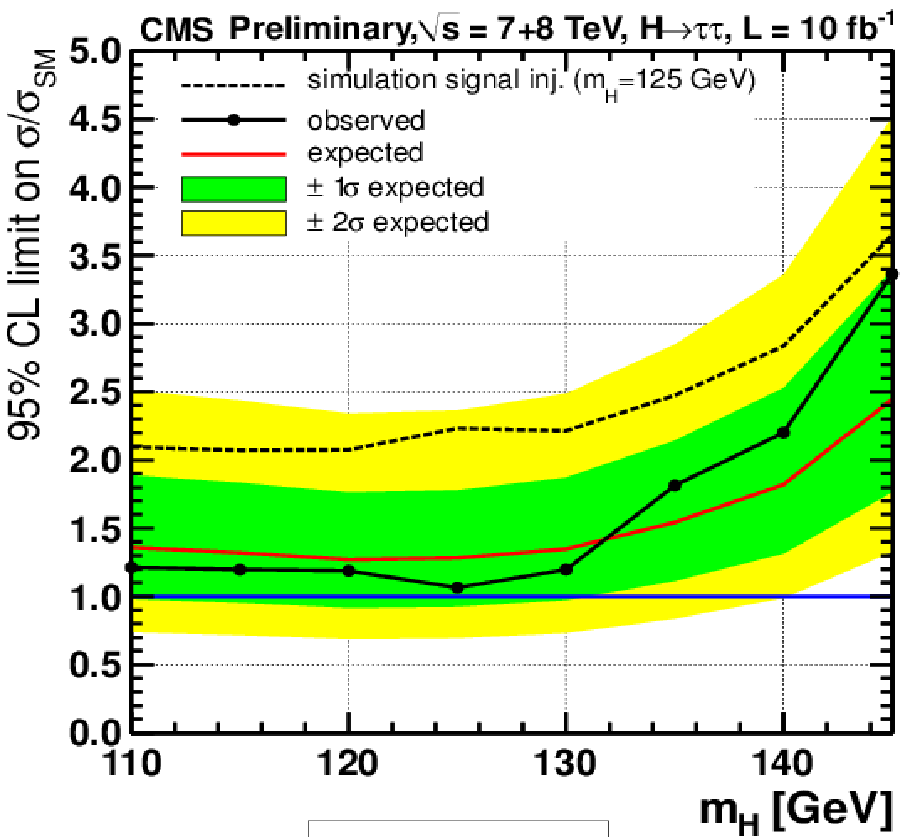
- Multijets: one jet is misidentified as e or μ and a second jet as τ_h; estimated from the number of observed same-charged τ pair events.

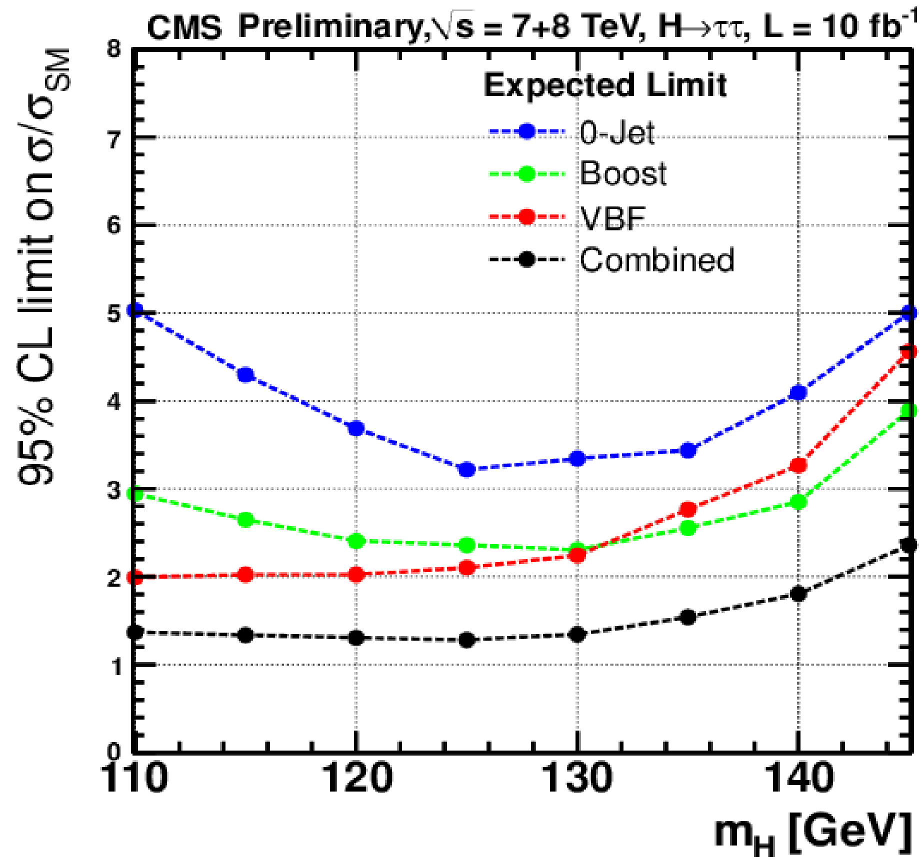
- W+jets: one jet misidentified as τ_h; estimated using high MT sidebands.

- DY→ ll: estimated from MC corrected for data/MC discrepancies.



H→ττ





BDT inputs

- m_{jj}
- $\Delta\eta_{jj}$
- $\Delta\phi_{jj}$
- $\Delta\phi$
- di-jet p_T
- di-τ p_T + MET
- visible di-τ p_T
- Δη (di-τ, jet)

Systematics ($H \to \tau\tau$)

Experimental Uncertainties		Propagation into Limit Calculation		
Uncertainty	Uncert.	*0-Jet*	*Boost*	*VBF*
Electron ID & Trigger (*)	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$
Muon ID & Trigger (*)	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$
Tau ID & Trigger (*)	$\pm 7\%$	$\pm 7\%$	$\pm 7\%$	$\pm 7\%$
JES (Norm.) (*)	$\pm 2.5 - 5\%$	$\mp 1\%$	$\pm 5\%$	$\pm 10\%$
b-Tag Efficiency (*)	$\pm 10\%$	$\mp 1\%$	$\mp 2\%$	$\mp 2\%$
Mis-Tagging (*)	$\pm 30\%$	$\mp 1\%$	$\mp 1\%$	$\mp 1\%$
Norm. $Z \to \tau\tau$	$\pm 3\%$	$\pm 3\%$	$\pm 5\%$	$\pm 13\%$
Norm. $t\bar{t}$ (*)	$\pm 10 - 30\%$	$\pm 10\%$	$\pm 12\%$	$\pm 30\%$
Norm *EWK*	$\pm 30\%$	$\pm 30\%$	$\pm 15 - 30\%$	$\pm 30 - 100\%$
Norm Fakes	$\pm 10 - 30\%$	$\pm 10\%$	$\pm 10\%$	$\pm 30\%$
Lumi (Signal & EWK)	$\pm 2.2(5)\%$	$\pm 2.2(5)\%$	$\pm 2.2(5)\%$	$\pm 2.2(5)\%$
Norm. $W + jets$	$\pm 10 - 30\%$	$\pm 10\%$	$\pm 10 - 30\%$	$\pm 30\%$
Norm. Z: l fakes τ_h	$\pm 20 - 100\%$	$\pm 20 - 30\%$	$\pm 20 - 100\%$	$\pm 30\%$
Norm. Z: jet fakes τ_h	$\pm 20\%$	$\pm 20\%$	$\pm 20\%$	$\pm 30\%$

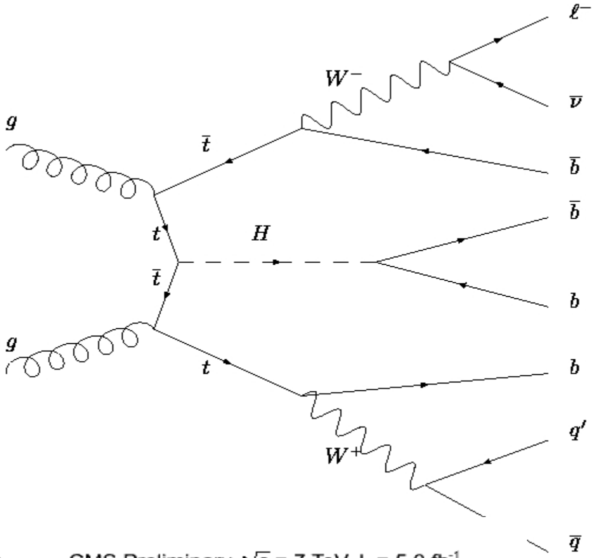
Theory Uncertainties (SM)		Propagation into Limit Calculation		
Uncertainty	Uncert.	*0-Jet*	*Boost*	*VBF*
PDF (*)	-	$\pm 2 - 8\%$	$\pm 2 - 8\%$	$\pm 2 - 8\%$
$\mu_r/\mu_f(gg \to H)$ (*)	-	$\pm 8\%$	$\pm 10\%$	$\pm 30\%$
$\mu_r/\mu_f(qq \to H)$ (*)	-	$\pm 3.5\%$	$\pm 4\%$	$\pm 10\%$
$\mu_r/\mu_f(qq \to VH)$ (*)	-	$\pm 4\%$	$\pm 4\%$	$\pm 4\%$
UE & PS (*)	-	$\mp 4\%$	$\pm 4\%$	$\pm 4\%$

ANN (H→b$\bar{b}$)

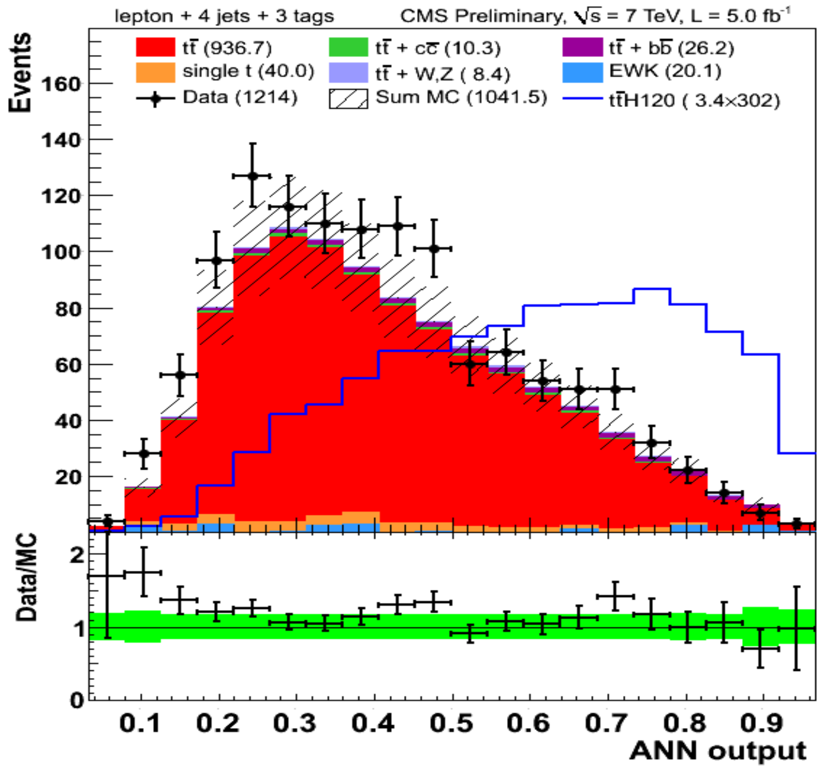
Variable
p_{Tj}: transverse momentum of each Higgs daughter
$m(\mathrm{jj})$: dijet invariant mass
$p_T(\mathrm{jj})$: dijet transverse momentum
$p_T(\mathrm{V})$: vector boson transverse momentum (or pfMET)
$\mathrm{CSV_{max}}$: value of CSV for the b-tagged jet with largest CSV value
$\mathrm{CSV_{min}}$: value of CSV for the b-tagged jet with second largest CSV value
$\Delta\phi(\mathrm{V,H})$: azimuthal angle between V (or E_T^{miss}) and dijet
$\lvert\Delta\eta(\mathrm{jj})\rvert$; difference in η between Higgs daughters
$\Delta R(\mathrm{j1,j2})$; distance in η–ϕ between Higgs daughters (not for $\mathrm{Z}(\ell\ell)\mathrm{H}$)
N_{aj}: number of additional jets ($p_T > 30\,\mathrm{GeV}$, $\lvert\eta\rvert < 4.5$)
$\Delta\phi(\mathrm{E_T^{miss}, jet})$: azimuthal angle between E_T^{miss} and the closest jet (only for $\mathrm{Z}(\nu\nu)\mathrm{H}$)
$\Delta\theta_{pull}$: color pull angle [62] (not for $\mathrm{Z}(\ell\ell)\mathrm{H}$)

$H \to b\bar{b}$ ($t\bar{t}H$)

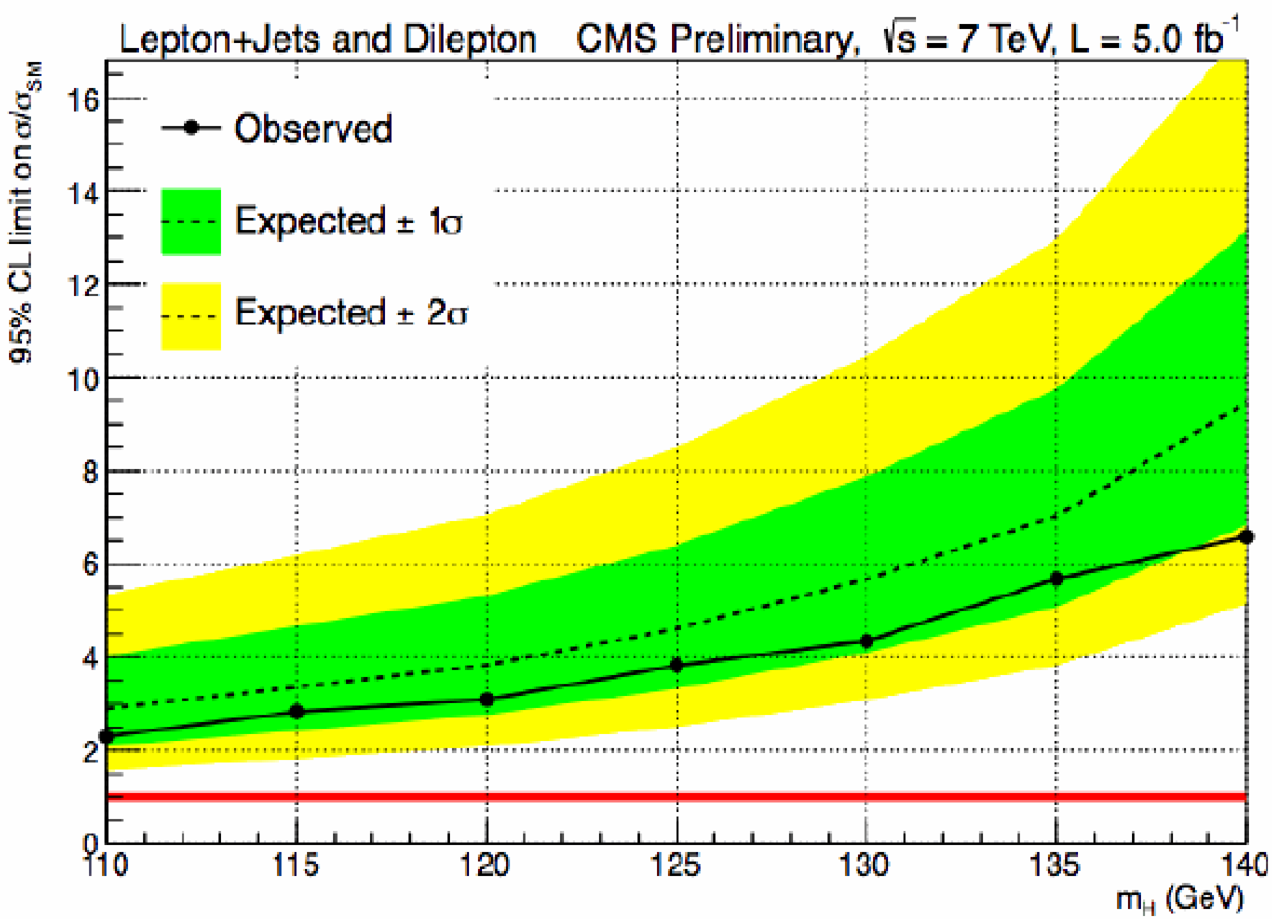
- New analysis performed only on 2011 data.
- Studied lepton+jet and di-lepton top decays
  - Major background from $t\bar{t}$ (+jet) events
- Split events by top decay and by number of jets and b-tags.



- Artificial Neural Network (ANN) to separate $t\bar{t}$ and $t\bar{t}H$:
  - b-tag, kinematic and angular correlation (e.g. min ΔR between all pairs of b-tagged jets)
- Simultaneous maximum likelihood fit of ANN shape in each jet/tag category



Exclusion Limit ($t\bar{t}H$)



Sensitivity dominated by lepton+jet mode
(only 5-10% improvement from di-lepton)

No excess seen, expect 4.6xσ_{SM} at 125 GeV, observed 3.8xσ_{SM}