Search for the SM Higgs Boson at CMS

Francesco Fabozzi

INFN-Napoli & Università della Basilicata

On behalf of the CMS collaboration

LNF-2013-1 Winter Institute on Hot Topics at LHC

LNF, 24 January 2013

Outline

- This talk will be focused on the recent results from Standard Model (SM) Higgs boson searches at CMS

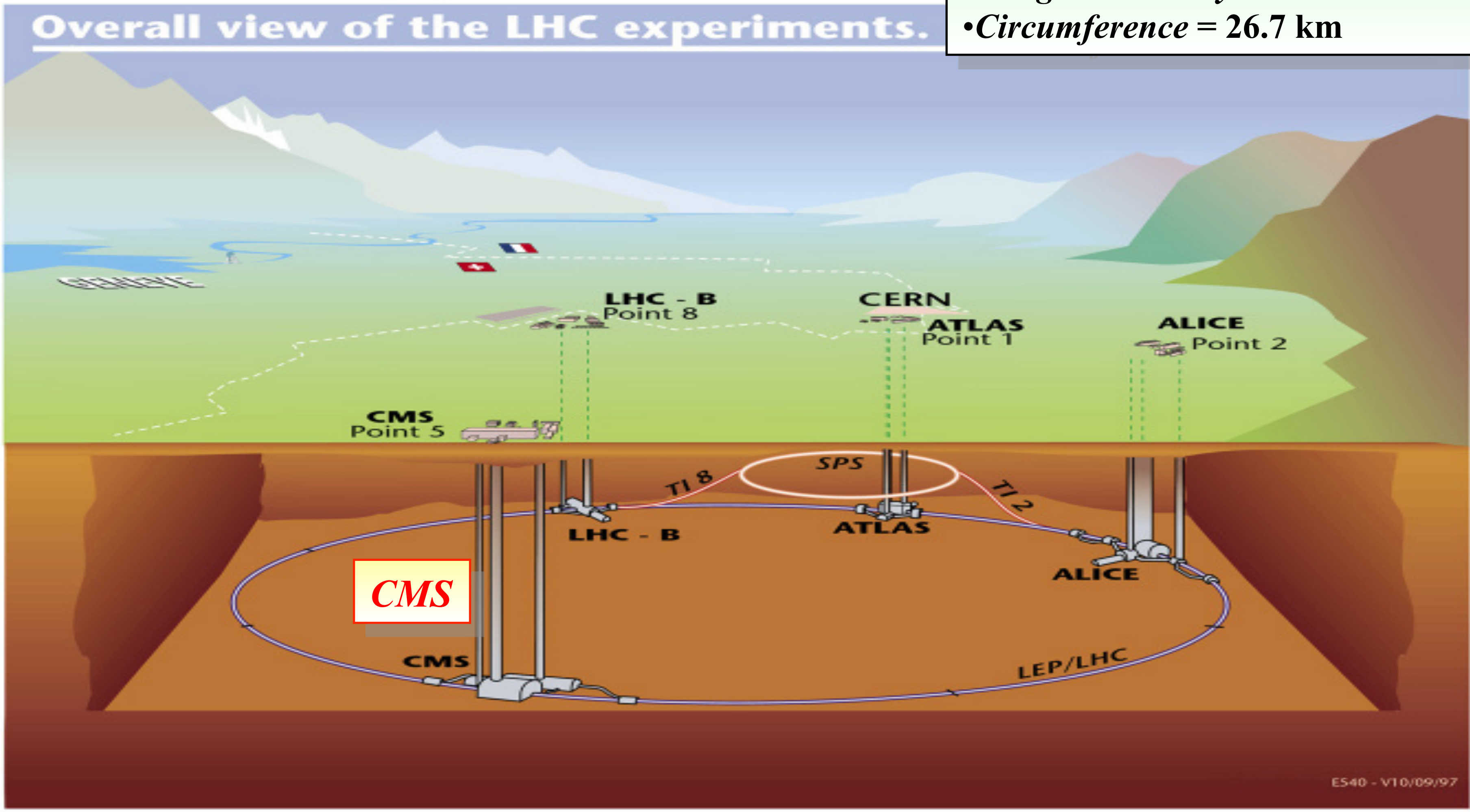
- Introduction
- Search channels of SM Higgs boson
  - Focus on the low mass range
- Analysis strategies and results
- Combination of results, first measurements of boson properties

The CMS experiment @ LHC

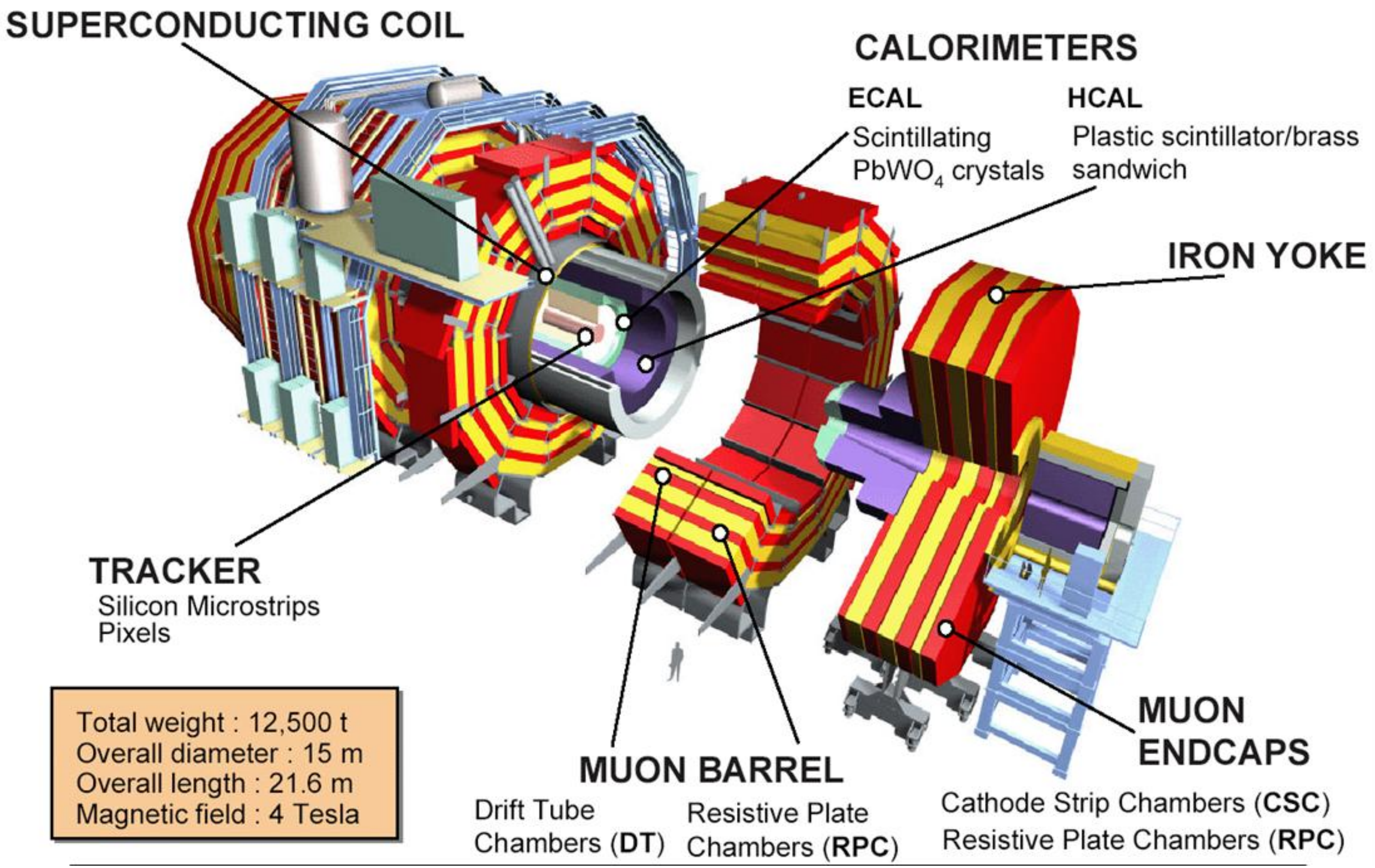
2010 – 2011: pp collisions at sqrt(s) = 7 TeV
2012: pp collisions at sqrt(s) = 8 TeV
Goal: 14 TeV after 2013-14 shutdown

The Large Hadron Collider

- ***p-p collider***
- ***Design luminosity* = 10^{34} cm^{-2} s^{-1}**
- ***Circumference* = 26.7 km**



The CMS detector

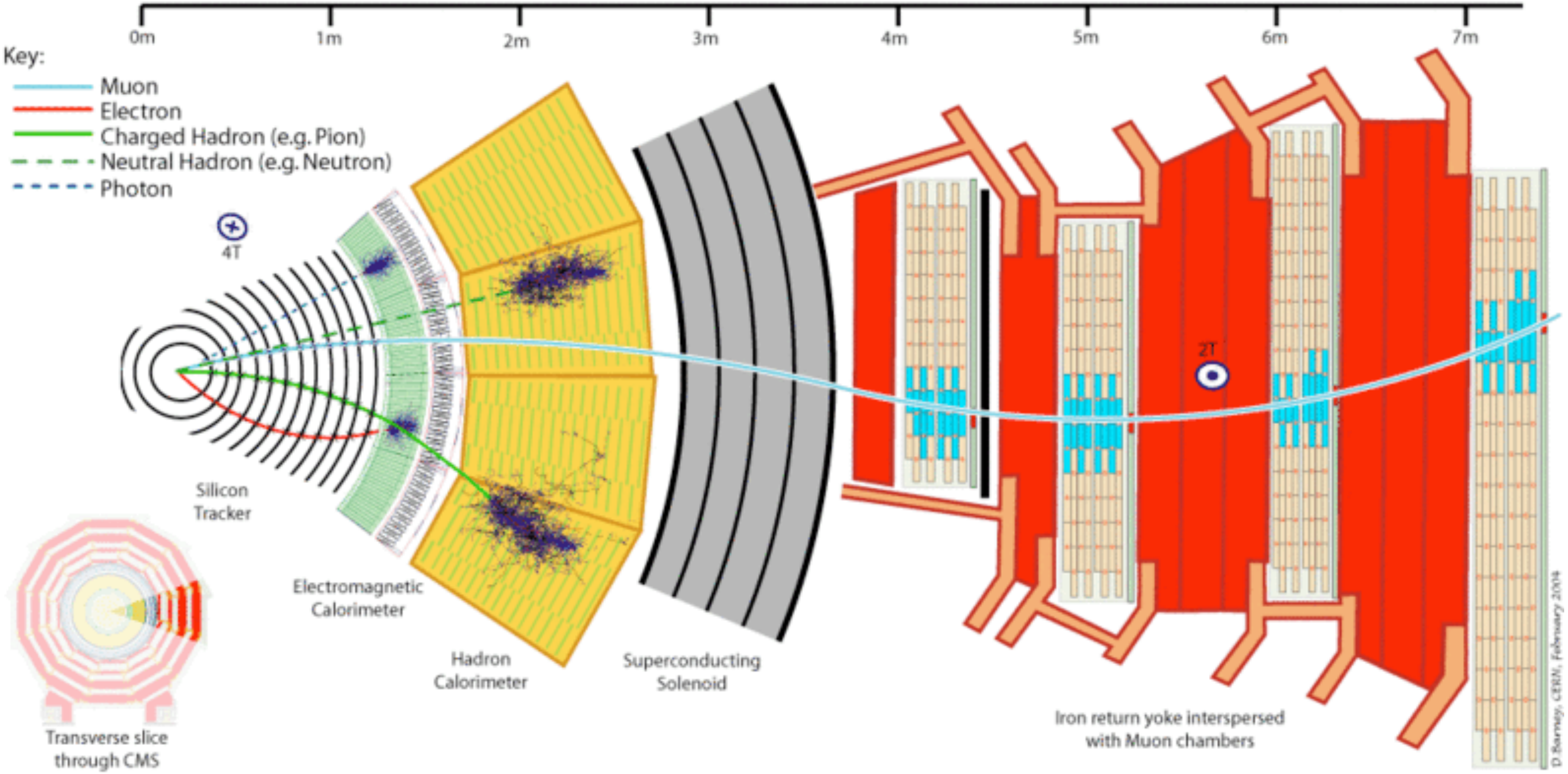


The CMS collaboration

About 3000 collaborators from 179 Institutes and 41 countries

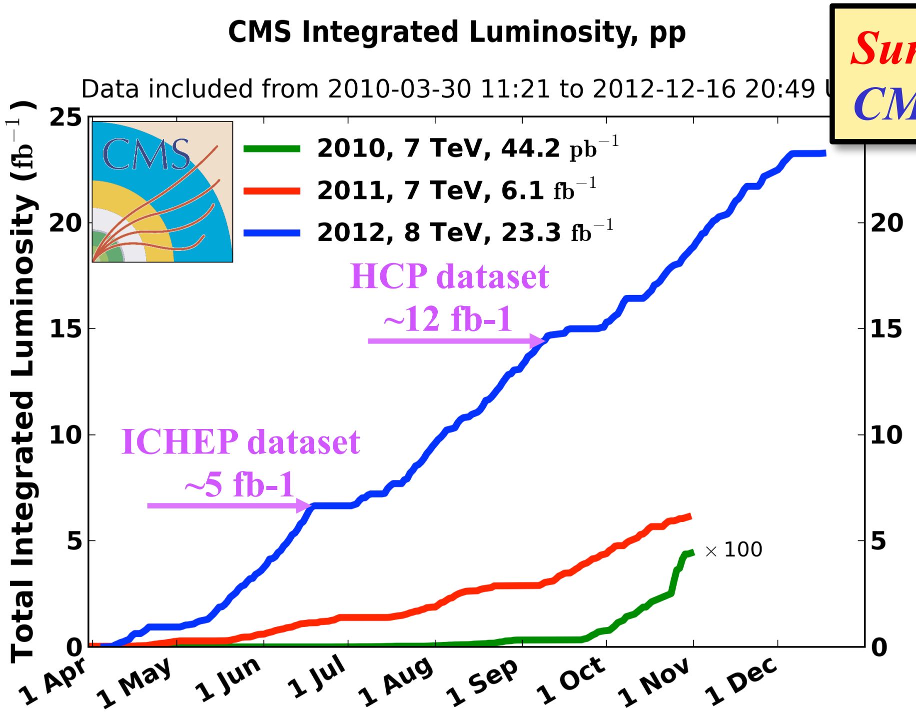
Note: detector image in full scale

Particle reconstruction at CMS



A particle-flow algorithm combines the information from all the sub-detectors in order to provide mutually exclusive lists of particles

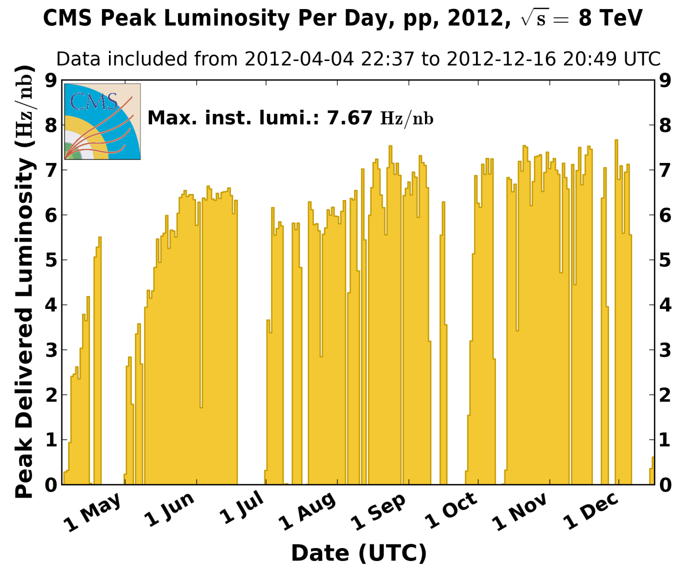
LHC impressive performances!

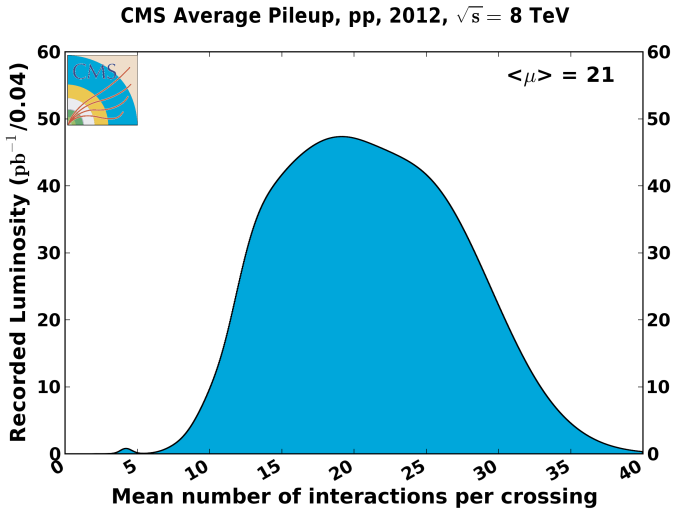


Summary of p-p runs in 2010-2012

CMS Data taking efficiency > 90%

•*Peak luminosity record = 7.67 10^{33} cm^{-2} s^{-1}* (was 3.5 10^{33} cm^{-2} s^{-1} in 2011)

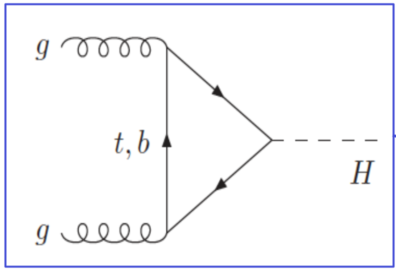




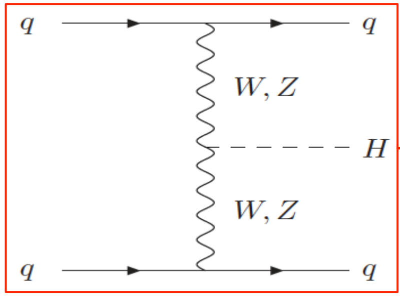
•*Multiple p-p interactions in the same bunch crossing ("pile-up")*

•*Average "pile up" interactions ~ 21 (was 10 in 2011)*

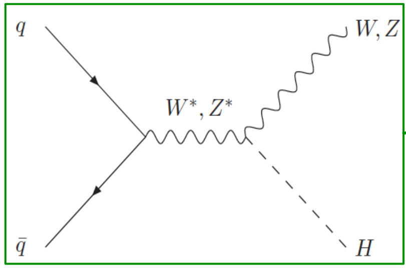
SM Higgs production and decay at LHC



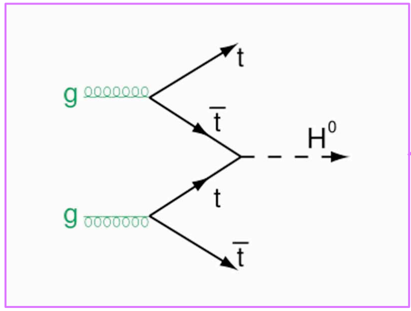
gg fusion



Vector Boson Fusion

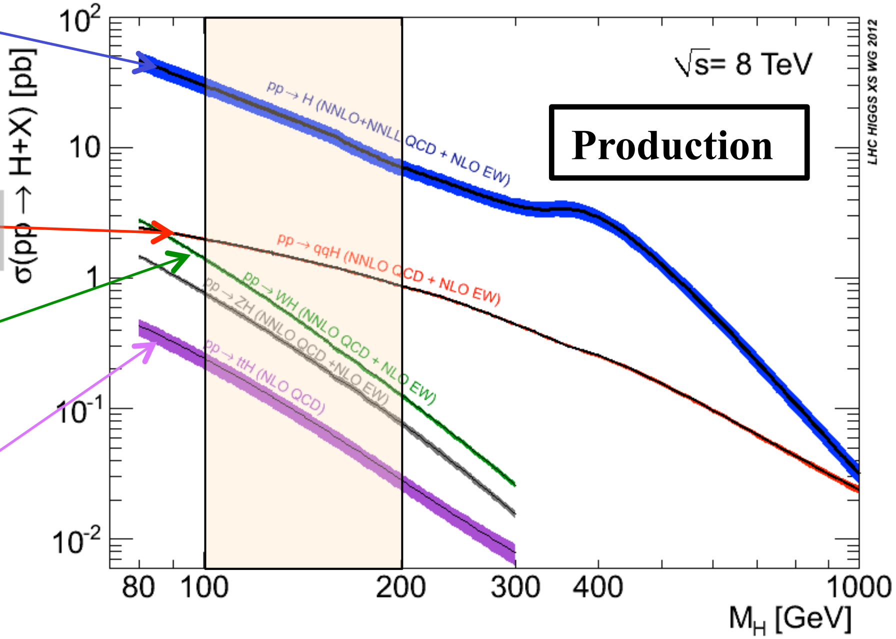


Associated VH production



Associated ttH production

High focus in the low-mass region



σ =23 (29) – 10 (14) pb at 7 (8) TeV for m_H in 110 – 160 GeV interval

New boson observation @ 125 GeV

Physics Letters B 716 (2012) 30–61

Contents lists available at SciVerse ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC ☆

CMS Collaboration *

CERN, Switzerland

This paper is dedicated to the memory of our colleagues who worked on CMS but have since passed away. In recognition of their many contributions to the achievement of this observation.

A R T I C L E   I N F O

Article history:
Received 31 July 2012
Received in revised form 9 August 2012
Accepted 11 August 2012
Available online 18 August 2012
Editor: W.-D. Schlatter

Keywords:
CMS
Physics
Higgs

A B S T R A C T

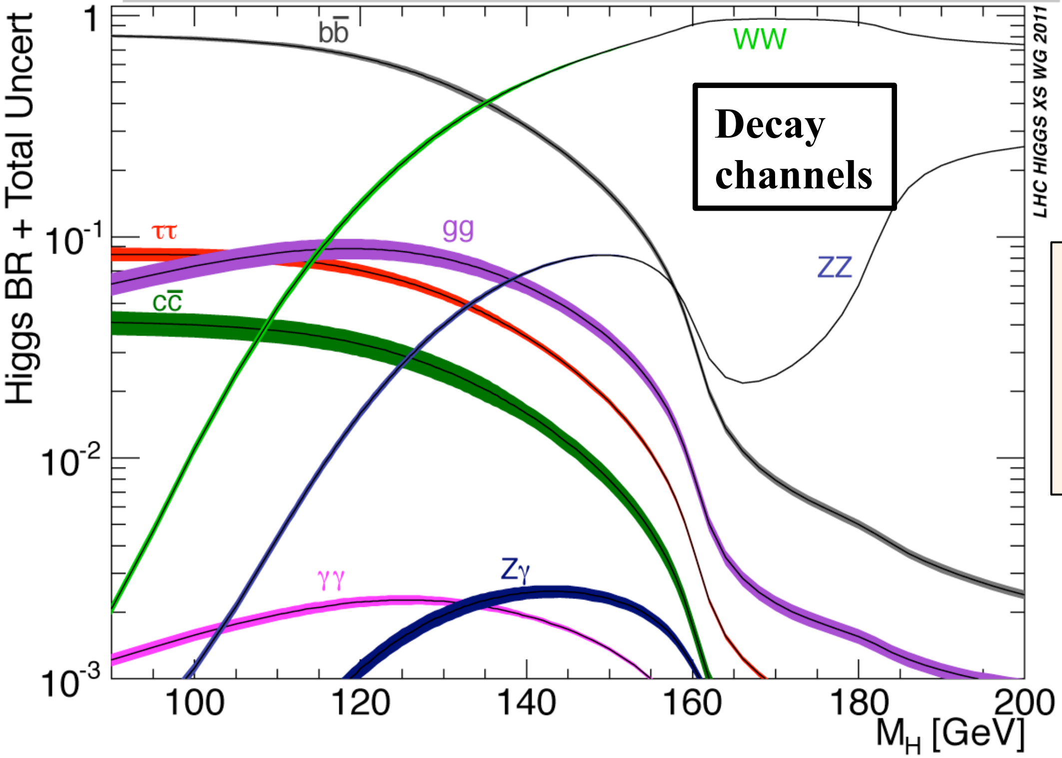
Results are presented from searches for the standard model Higgs boson in proton–proton collisions at $\sqrt{s} = 7$ and 8 TeV in the Compact Muon Solenoid experiment at the LHC, using data samples corresponding to integrated luminosities of up to 5.1 fb^{-1} at 7 TeV and 5.3 fb^{-1} at 8 TeV. The search is performed in five decay modes: $\gamma\gamma$, ZZ, W^+W^-, $\tau^+\tau^-$, and $b\bar{b}$. An excess of events is observed above the expected background, with a local significance of 5.0 standard deviations, at a mass near 125 GeV, signalling the production of a new particle. The expected significance for a standard model Higgs boson of that mass is 5.8 standard deviations. The excess is most significant in the two decay modes with the best mass resolution, $\gamma\gamma$ and ZZ; a fit to these signals gives a mass of $125.3 \pm 0.4(\text{stat.}) \pm 0.5(\text{syst.})$ GeV. The decay to two photons indicates that the new particle is a boson with spin different from one.

Also SM Higgs exluded @95%CL in 110-122.5 GeV and 127-600 GeV m_H intervals

Question: is it the SM Higgs?
Goal: Study properties of resonance and fully explore its decay channels

SM Higgs production and decay at LHC

A multiplicity of decay channels at low mass



Decay channels

Different mass resolution and sensitivity
Note: sensitivity it's not only related to the BR

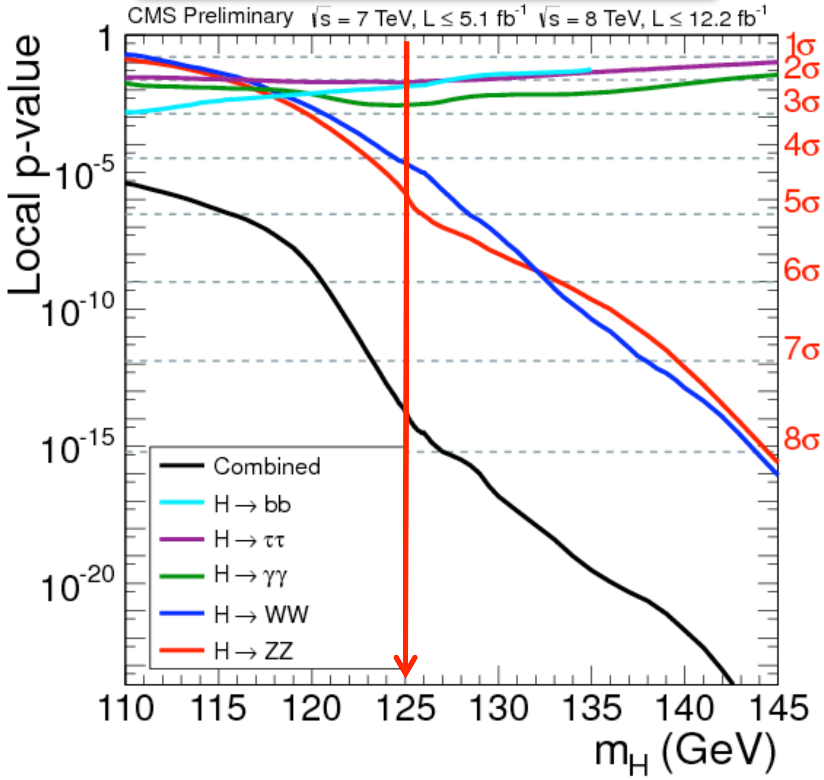
SM Higgs search channels at CMS

5 channels explored

Channel	m_H range (GeV)	Luminosity (fb-1)		m_H resolution
		7 TeV	8 TeV	
H-> γγ	110-150	5	5	1-2 %
H->bb	110-135	5	12	10%
H->ττ	110-145	5	12	20%
H->WW	110-600	5	12	20%
H->ZZ	110-1000	5	12	1-2%

Sensitive also to high m_H

Expected sensitivities for a SM Higgs boson



Expected significance @ m_H = 125 GeV: 7.8 σ

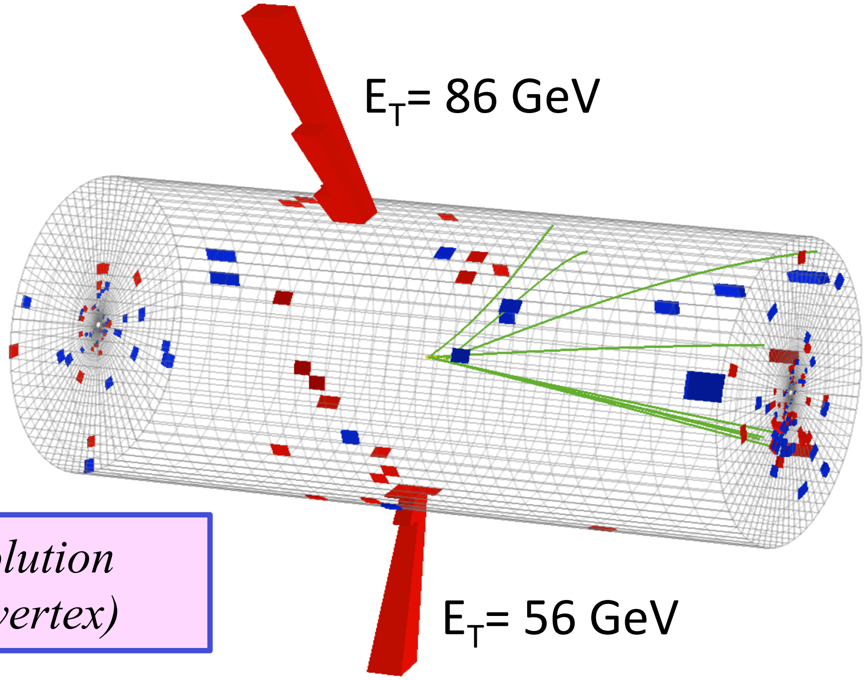
$H \rightarrow \gamma\gamma$

$H \rightarrow \gamma\gamma$ channel

PLB 716 (2012) 30-61

Br ≈ 0.14 -0.23% in search range

Clean signature: two isolated, high E_T gamma; narrow peak (1-2% resolution) in $m_{\gamma\gamma}$ over decreasing background



Mass resolution dominated by ECAL energy resolution (crucial is correct assignement of the di-photon vertex)

Photon energy and uncertainty estimated event-by-event with multivariate regression technique

Main backgrounds:

- prompt γγ production (irreducible)
- γ+jet and jet-jet (< 30% contribution)

$H \rightarrow \gamma\gamma$ channel: event categories

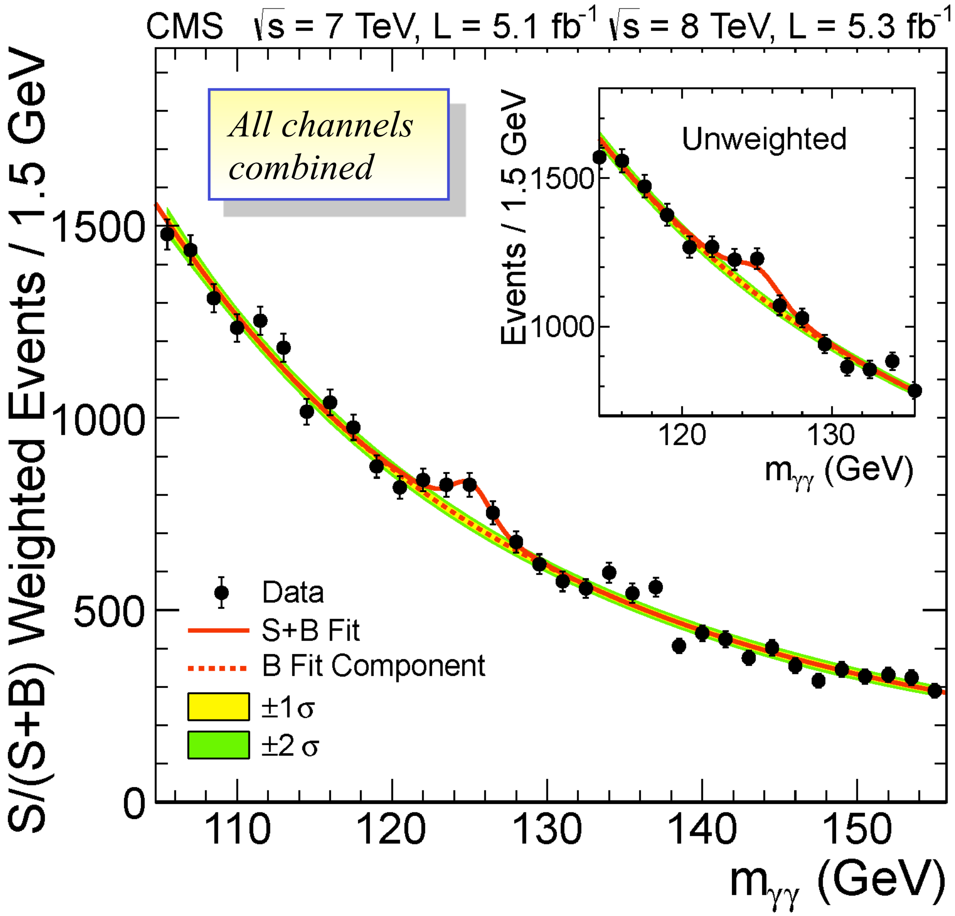
Events classification in independent categories of different S/B, selection optimized for each category

- *VBF-category: highest S/B, 2 high-p_T jets, $\Delta\eta > 3.5$; sub-splitting in tight ($m_{jj} > 500$ GeV) and loose ($250 < m_{jj} < 500$ GeV) category; cuts exploiting correlation between di-photon and di-jet system applied*
- *Remaining events split in 4 categories: based on output of di-photon BDT discriminant exploiting signal-like kinematics, di-photon mass resolution, photon BDT identification*
- *Cut-based analysis performed as cross-check*

	Event categories	SM Higgs boson expected signal ($m_H = 125$ GeV) Events	ggH	VBF	VH	ttH	σ_{eff} (GeV)	FWHM/2.35 (GeV)	Background $m_{\gamma\gamma} = 125$ GeV (events/GeV)
8 TeV, 5.3 fb^{-1}	BDT 0	6.1	68%	12%	16%	4%	1.38	1.23	7.4 ± 0.6
	BDT 1	21.0	87%	6%	6%	1%	1.53	1.31	54.7 ± 1.5
	BDT 2	30.2	92%	4%	4%	–	1.94	1.55	115.2 ± 2.3
	BDT 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

H→γγ channel: signal extraction

Fit to $m_{\gamma\gamma}$ data distribution in each category

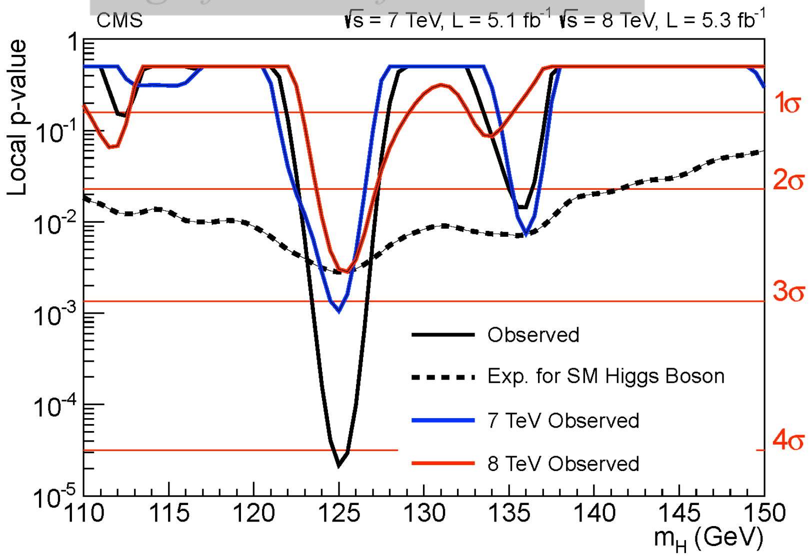


- *Bkg parametrized with polynomials*
- *Parametrization tested for robustness w.r.t. possible bias on the signal extraction*

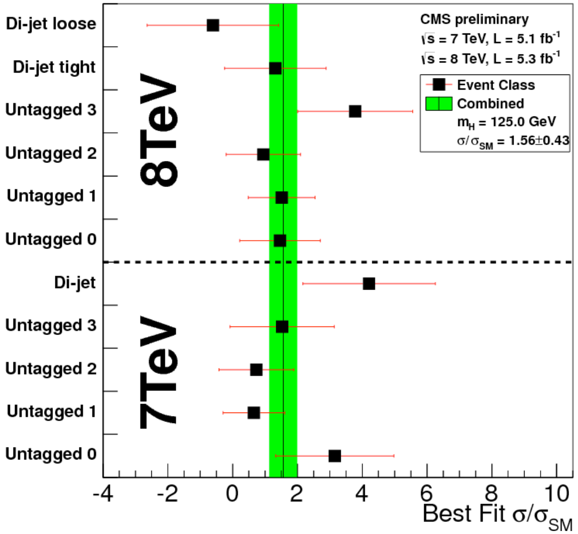
Events in the plot weighted by S/(S+B) of each category

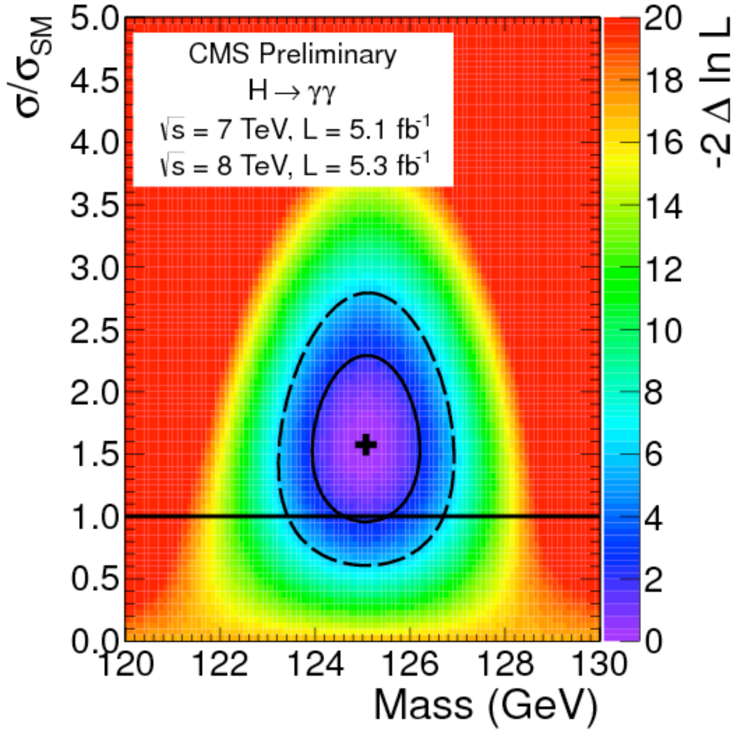
H→γγ channel: results

Significance of the excess



4σ level observation @ 125 GeV





$m_X = 125.1 \pm 0.4(stat.) \pm 0.6(syst.)$ GeV

$\sigma/\sigma_{SM} = 1.56 \pm 0.43$ @ 125 GeV

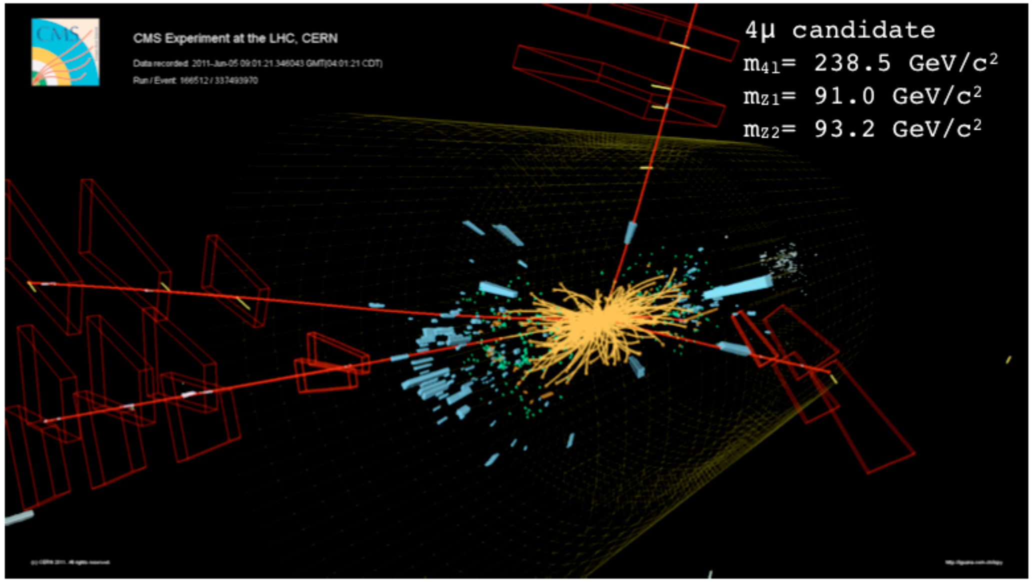
$H \rightarrow ZZ$

H→ZZ channel

4l channel (=4μ, 4e, 2e2μ)

PAS HIG-12-041

Signature: two pairs of isolated, high p_T leptons of opposite sign and originating from the primary vtx;



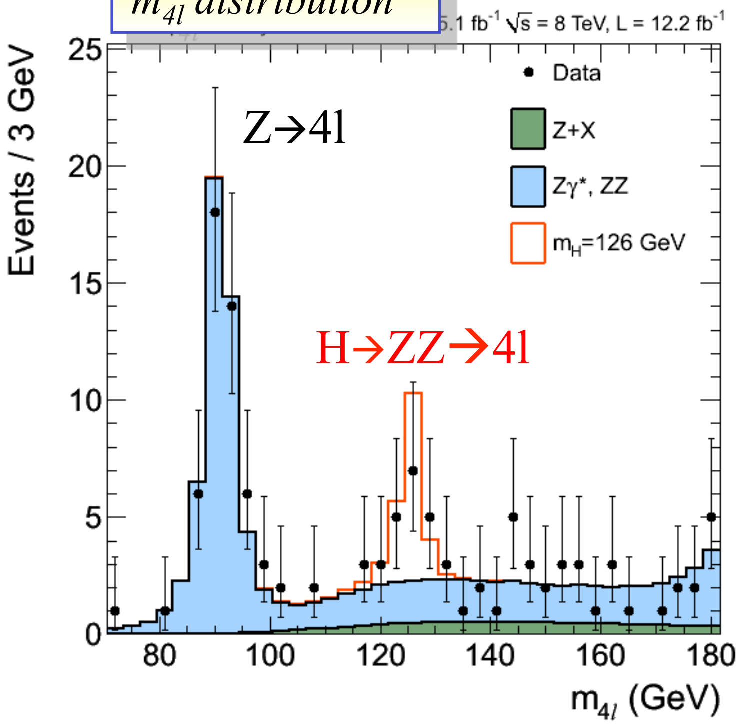
Z bosons can be off-mass shell:
$40<M_{Z1}<120$ GeV
$12<M_{Z2}<120$ GeV

High lepton efficiency through a broad p_T range is crucial

Dominant backgrounds:
- non resonant ZZ (irreducible)
- Z+jets (mainly Zbb) and ttbar

H→ZZ: selected events

m_{4l} distribution



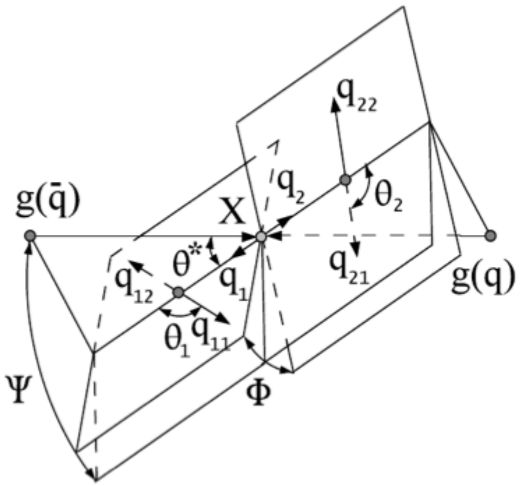
Candidates in the range 110-160 GeV

Channel	4e	4μ	$2e2\mu$	4ℓ
ZZ background	4.7 ± 0.6	9.6 ± 1.0	12.5 ± 1.4	26.8 ± 1.8
Z+ X	$3.4^{+3.0}_{-2.3}$	$1.6^{+1.2}_{-0.9}$	$5.6^{+5.4}_{-3.6}$	$10.6^{+5.3}_{-4.4}$
All backgrounds	$8.0^{+3.1}_{-2.3}$	$11.2^{+1.6}_{-1.4}$	$18.1^{+5.6}_{-3.8}$	$37.3^{+6.6}_{-4.7}$
$m_H =$ 125 GeV	2.4 ± 0.4	4.6 ± 0.5	5.9 ± 0.7	12.9 ± 0.9
$m_H =$ 126 GeV	2.7 ± 0.4	5.1 ± 0.6	6.6 ± 0.8	14.4 ± 1.1
Observed	12	16	19	47

- *ZZ bkg determined from simulation*
- *Reducible bkg extrapolated from control regions (e.g. same charge lepton pair with loosened lepton isolation/ID)*

H→ZZ: signal extraction

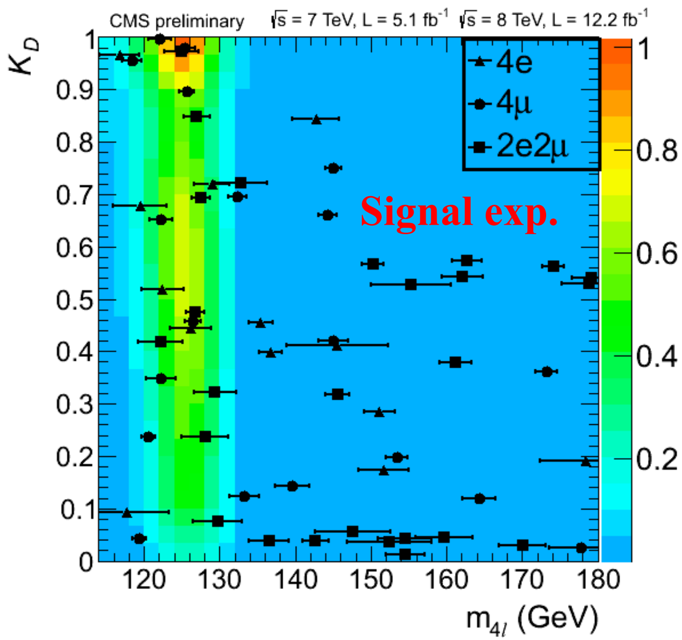
Simultaneous likelihood fit of (m_{4l}, K_D) distribution for the three channels

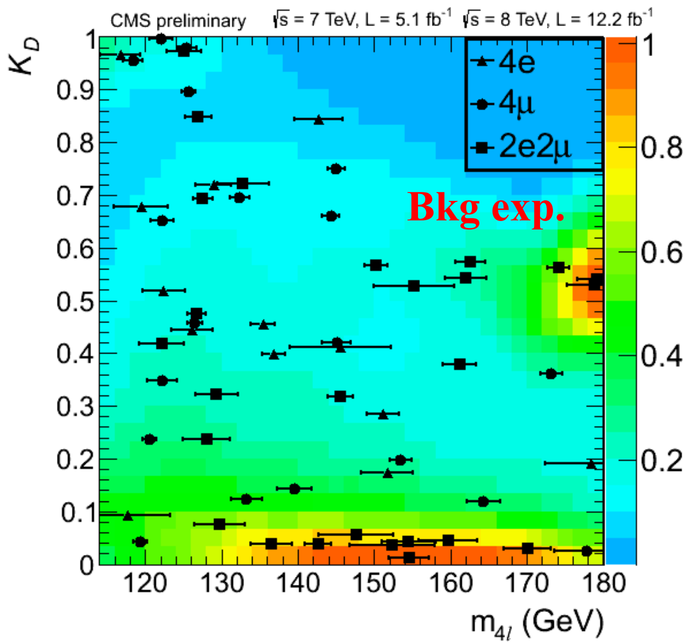


Kinematics fully described by 5 angles and m_{Z1}, m_{Z2} for a given m4l
Analytical parametrization derived for signal and bkg

Kinematic discriminant K_D

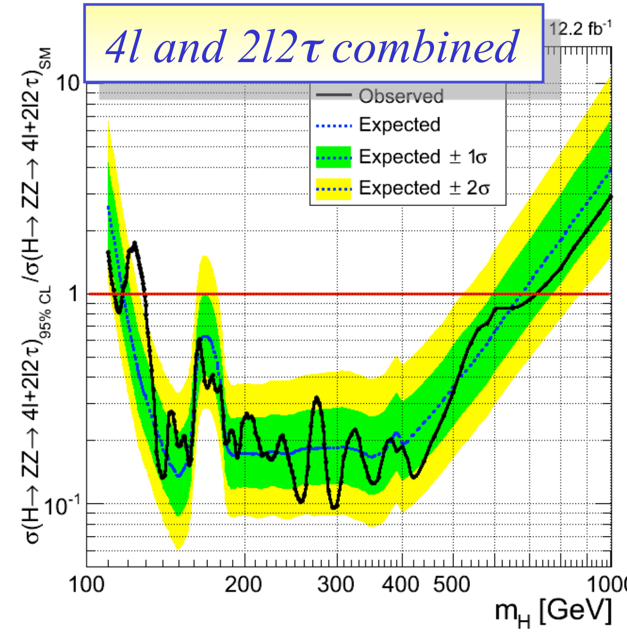
$$1/K_D = 1 + \frac{P_{background}(m_1, m_2, \theta_1, \theta_2, \Psi, \Phi, \theta^*)}{P_{signal}(m_1, m_2, \theta_1, \theta_2, \Psi, \Phi, \theta^*)}$$





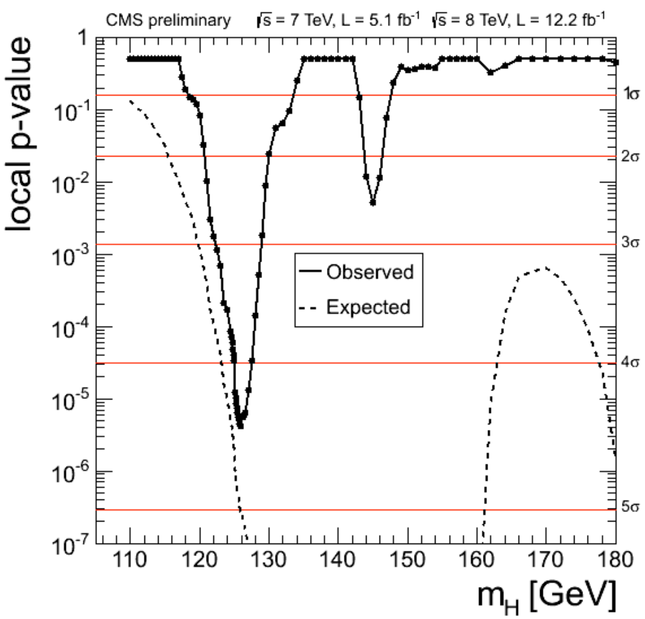
H→ZZ: results

4l and 2l2τ combined

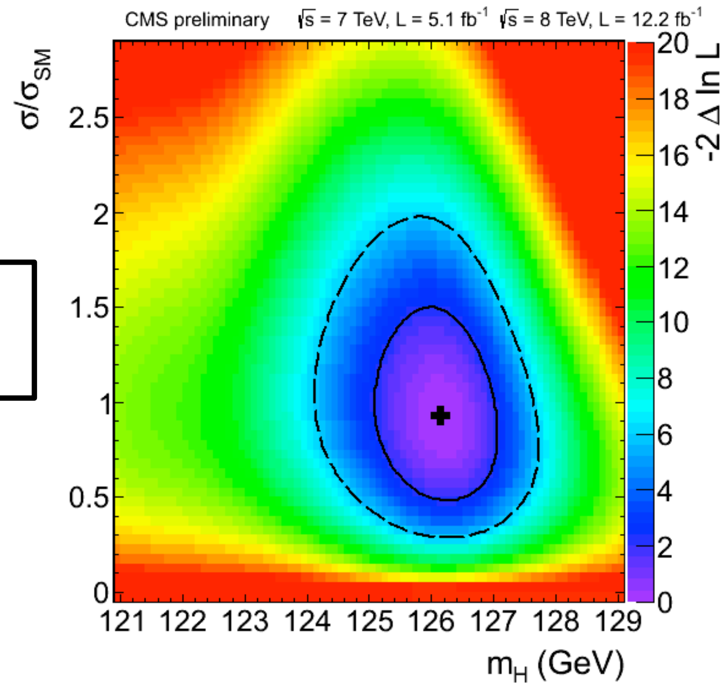


Search complemented for $m_H > 180$ GeV by $2l2\tau$ channel ($l = e,\mu$; $2\tau = \tau_h\tau_h$, $\tau_\mu\tau_h$, $\tau_e\tau_h$, $\tau_e\tau_\mu$),

Observed limit →2 exclusion intervals @95%CL:
113-116, 129-700 GeV



Max significance:
4.5 σ @ 126 GeV



$m_X = 126.2 \pm 0.6(stat.) \pm 0.2(syst.)$ GeV

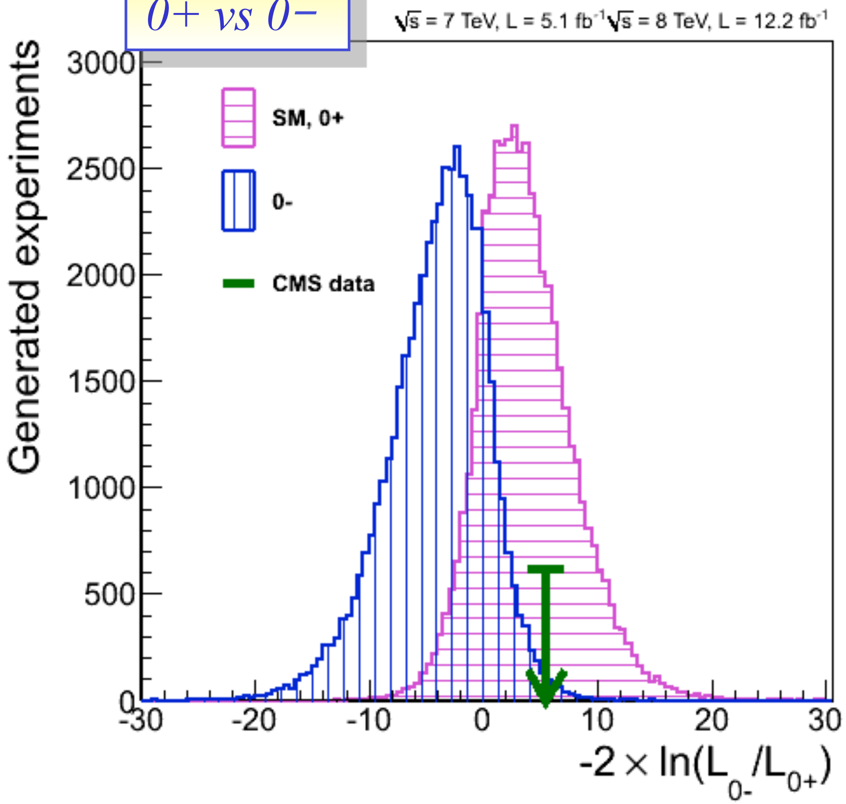
$\sigma/\sigma_{SM} = 0.80\ ^{+0.35}_{-0.28}$ @ 126 GeV

H→ZZ: spin-parity properties

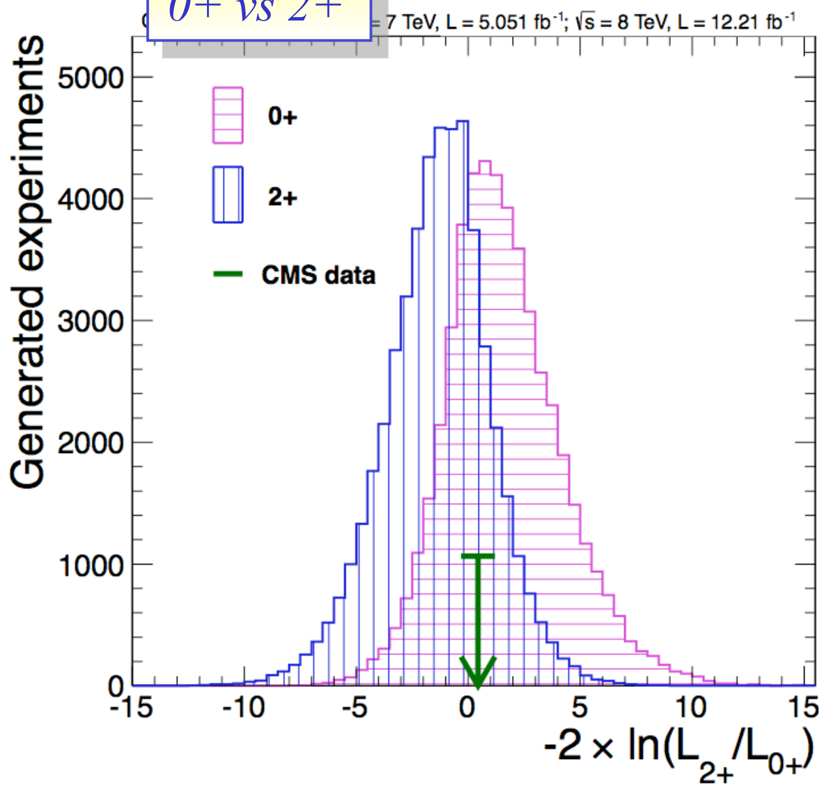
Using similar K_D approach to discriminate between different J^P states

$$\mathcal{D}_{J^P} = \frac{\mathcal{P}_{SM}}{\mathcal{P}_{SM} + \mathcal{P}_{J^P}} = \left[1 + \frac{\mathcal{P}_{J^P}(m_1, m_2, \vec{\Omega}|m_{4\ell})}{\mathcal{P}_{SM}(m_1, m_2, \vec{\Omega}|m_{4\ell})}\right]^{-1}$$

0+ vs 0−



0+ vs 2+



Observation disfavours 0- at the 2.4 σ-level

Only 1σ separation expected

$H \rightarrow WW$

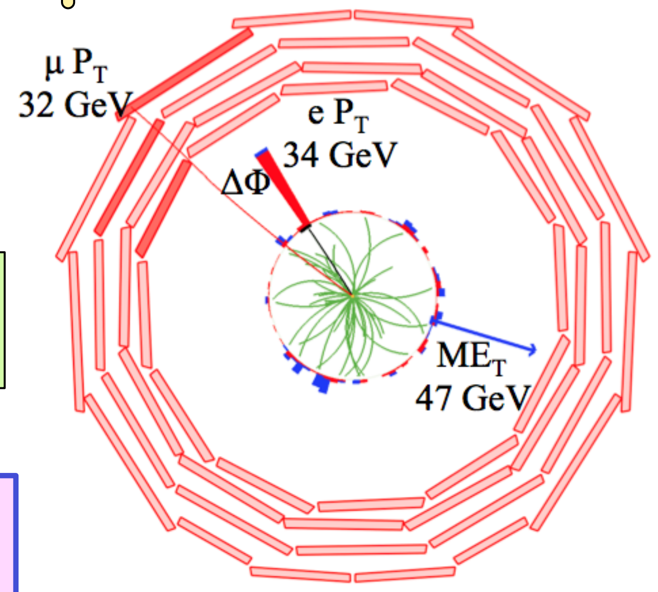
H→WW→2l2ν channel

PAS HIG-12-042

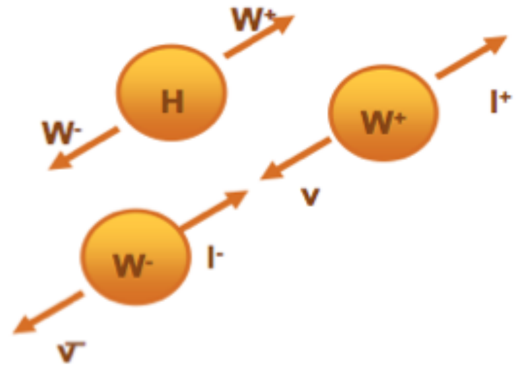
Three channels:
- *μμ, ee (same-flavour events – SF)*
- *eμ (different-flavour events - DF)*

Signature: two isolated, high p_T leptons of opposite sign; large missing transverse energy (MET) in the event

MET is crucial for DY rejection
- *Use of projected MET to nearest lepton direction improves DY→ττ rejection*



- *Best sensitivity for $m_H \approx 2m_W$*
- *Robust lepton ID and MET reconstruction in high PU environment for being sensitive down to 120 GeV*



Scalar boson + V-A decay
=> Small leptons opening angle

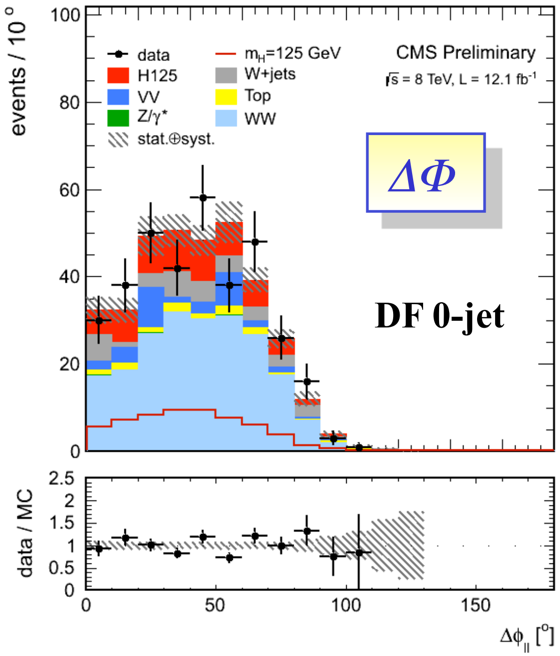
H→WW→2l2ν: event categories

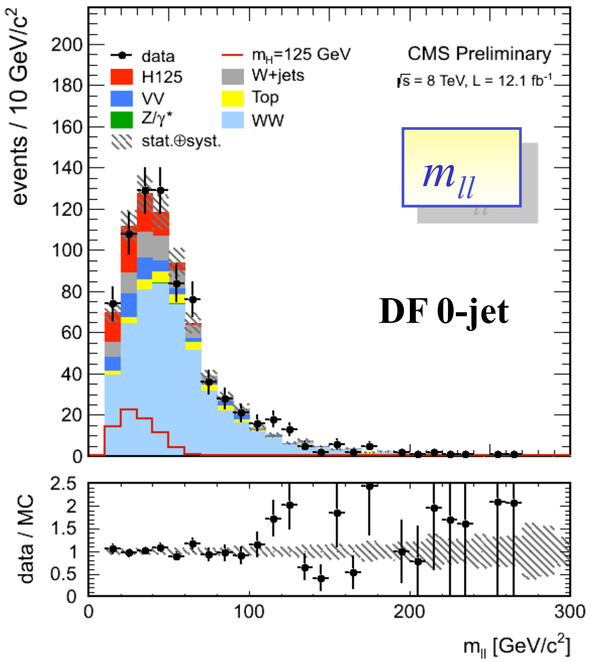
- *Events classification in independent 0-, 1-, 2-jet (VBF) exclusive categories (by counting jets with $p_T > 30$ GeV)*
- *Selection optimized for the different categories*

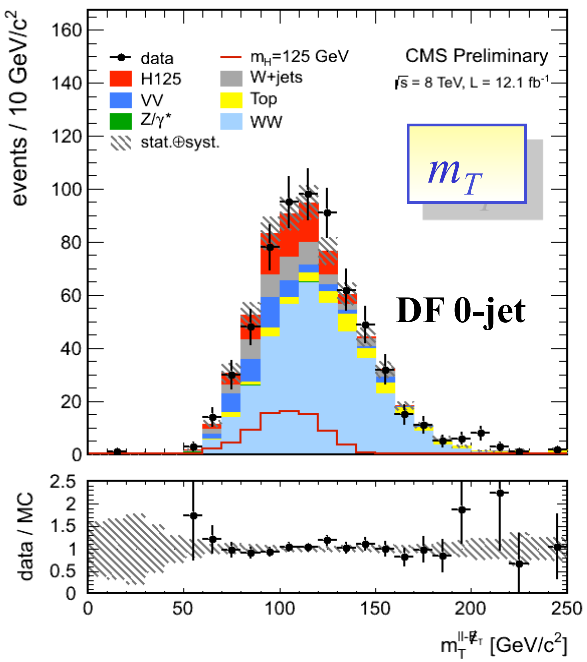
DF 0-jets category provides the best sensitivity
DF 1-jets + SF 0-jets contribute for another 10% to the sensitivity
VBF -> highest S/B but low yield

Dominant backgrounds:

- DF + 0-jets -> non resonant WW, W+jets,
- DF+n-jets -> also ttbar
- SF -> also DY



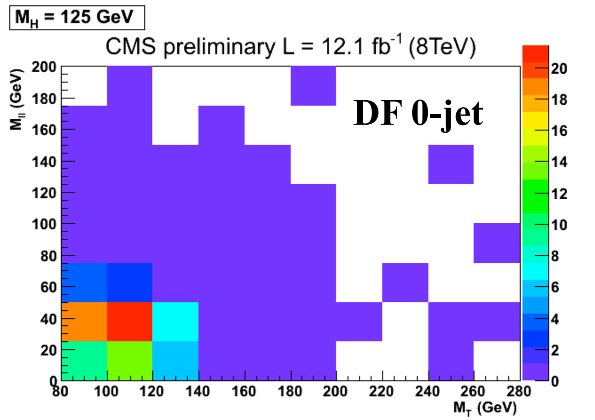


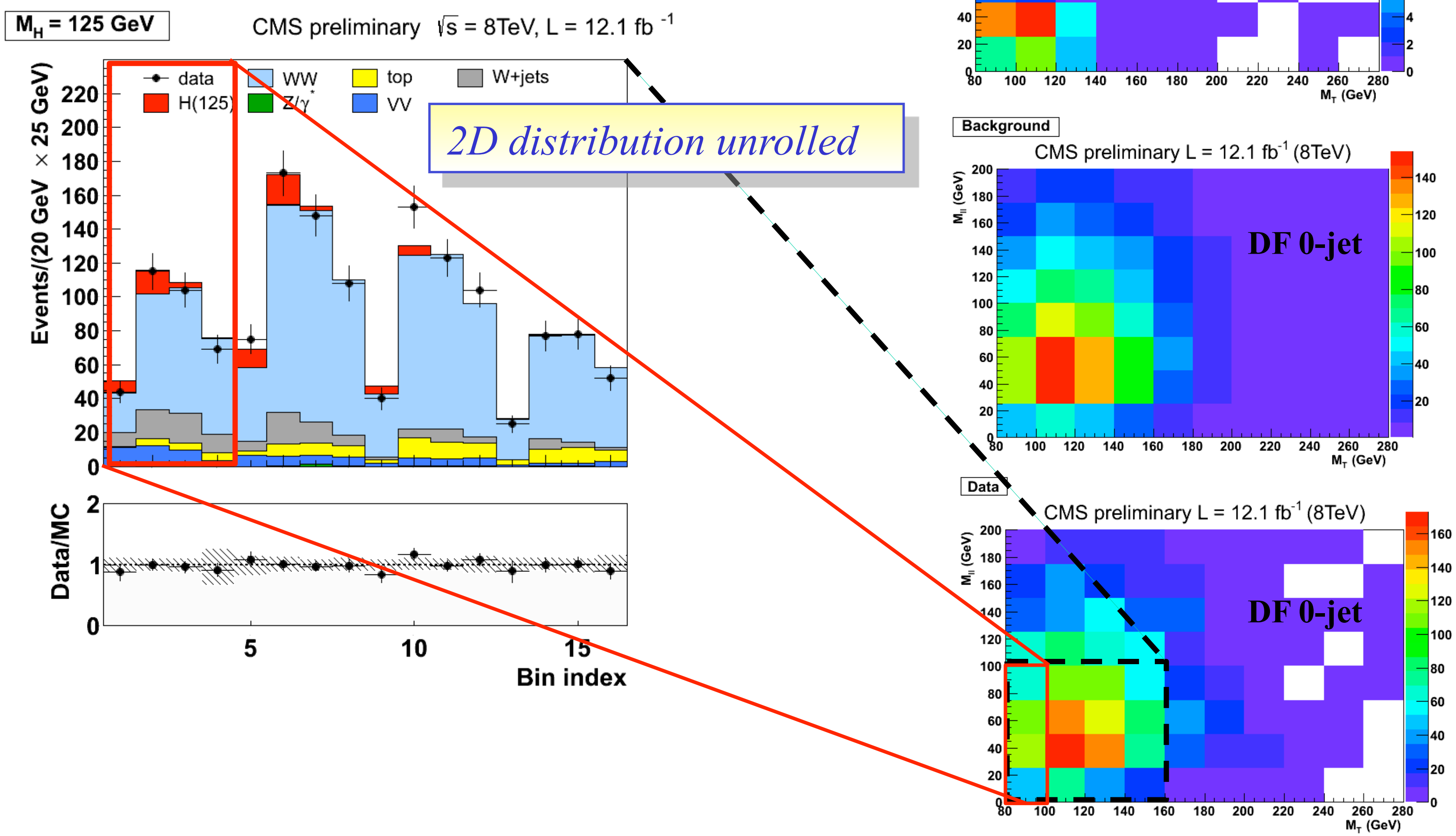


H→WW→2l2ν: signal extraction

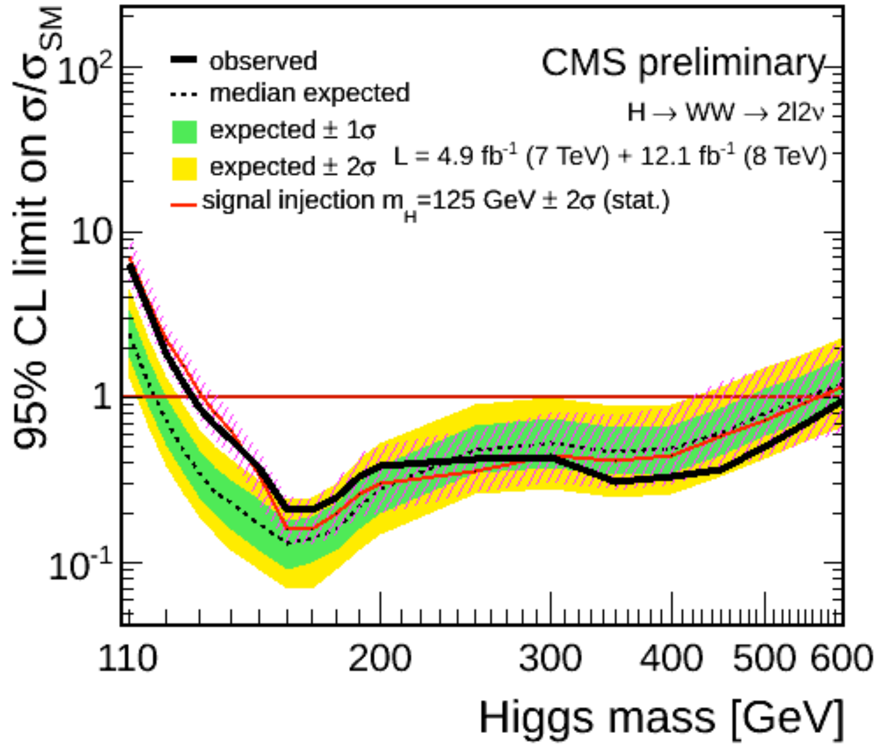
- *DF-0jet, -1jet: 2D shape analysis in* (m_{ll}, m_T)
- *Cut & count for the other categories*

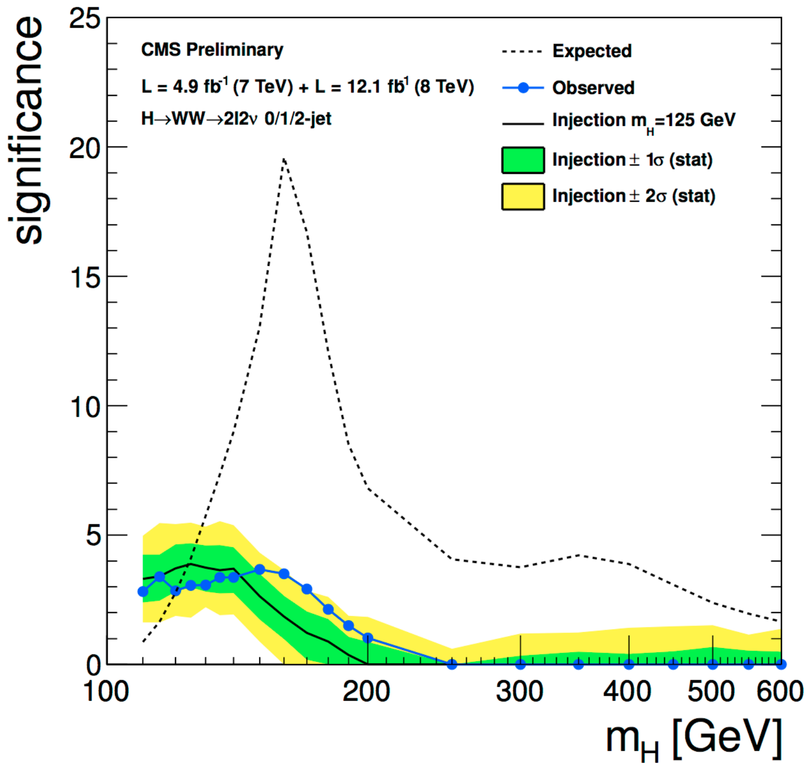
$$m_T = \sqrt{2p^{ll}_T ET^{miss}(1 - \cos\Delta\phi_{ETmiss,\,ll})}$$





H→WW→2l2ν: results





Excluded:128-600 GeV @ 95% CL

Observation compatible with an expected signal @ 125 GeV

Significance of the excess: 3.1 σ @ 125 GeV (expected significance 4.1 σ)

Best fit σ/σ_{SM}= 0.74 ± 0.25 @ 125 GeV

$H \rightarrow \tau\tau$

H→ττ channel

Channels considered: $\tau_h\tau_h$, $e\tau_h$, $\mu\tau_h$, $e\mu$, $\mu\mu$

PAS HIG-12-043

Signature: two isolated, high p_T leptons/tau with opposite charge; large missing transverse energy (MET) in the event

Main backgrounds: DY→ττ (irreducible), W+jets

W+jets rejection: in signal MET ~ collinear with direction of visible decay products

Additional criteria for bkg rejection

- *Veto events with additional leptons*
- *Veto events with b-tagged jets*

Improved MET measurement in high-PU environment with a multivariate regression technique

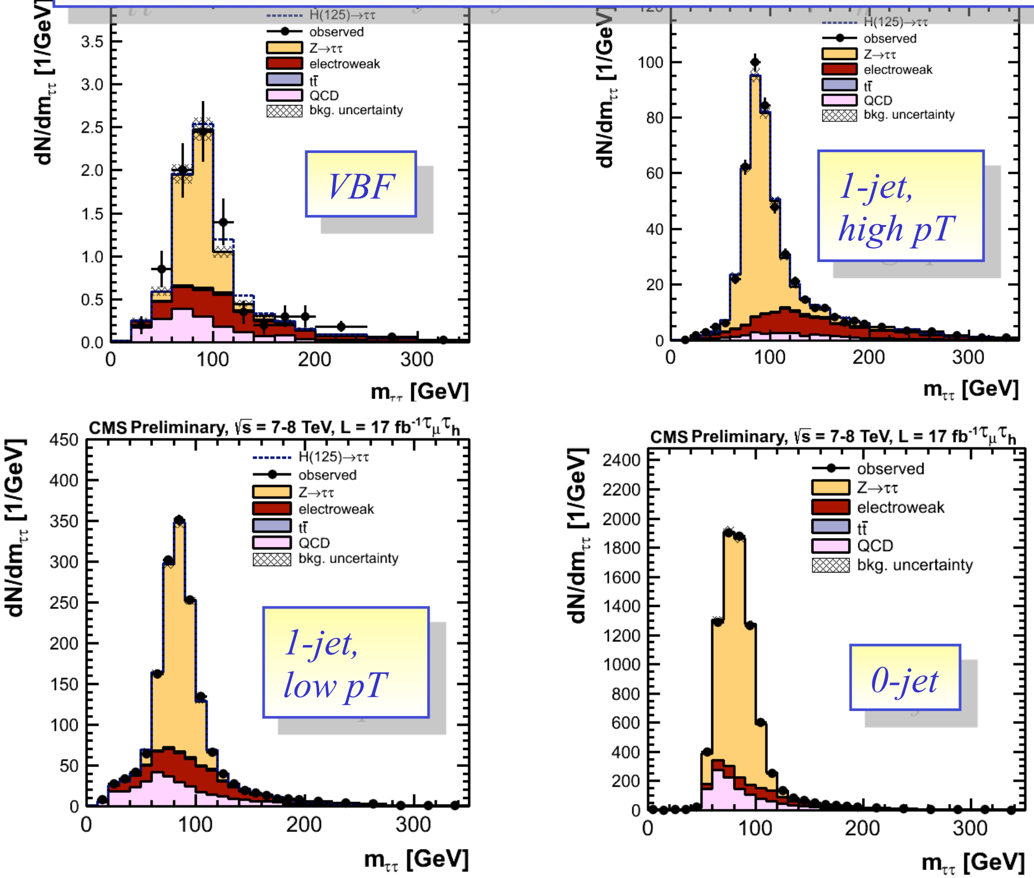
H→ττ: event categories

For each channel, further splitting of events in categories by counting jets with p_T >30 GeV

- 0-jet used only to constrain bkg
- 1-jet, high τ p_T
- 1-jet, low τ p_T
- 2 jet (p_T >30) – VBF

Final selection criteria optimized in each channel/ category

$m_{\tau\tau}$ distribution after full selection in $\mu\tau_h$ channel



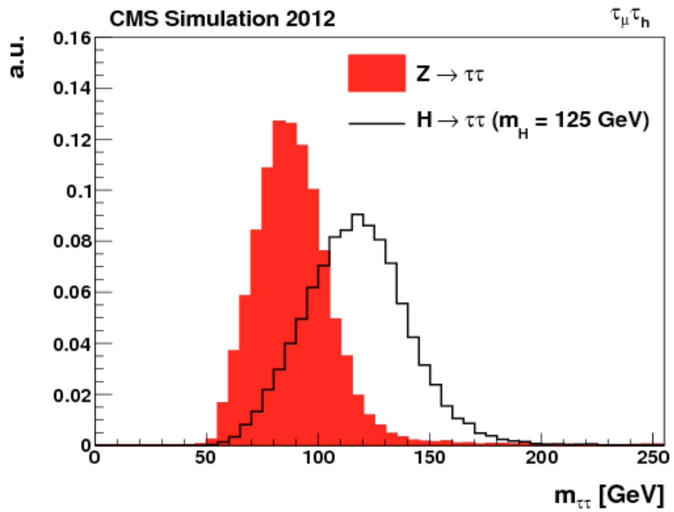
1-jet high p_T provides the highest sensitivity

- *Exploits boosted kinematics (and also larger MET, better resolution)*

H→ττ: signal extraction

Simultaneous binned likelihood fit to the $m_{\tau\tau}$ invariant mass distributions

ML technique for reconstruction of ττ invariant mass provides mass resolution ~ 20%

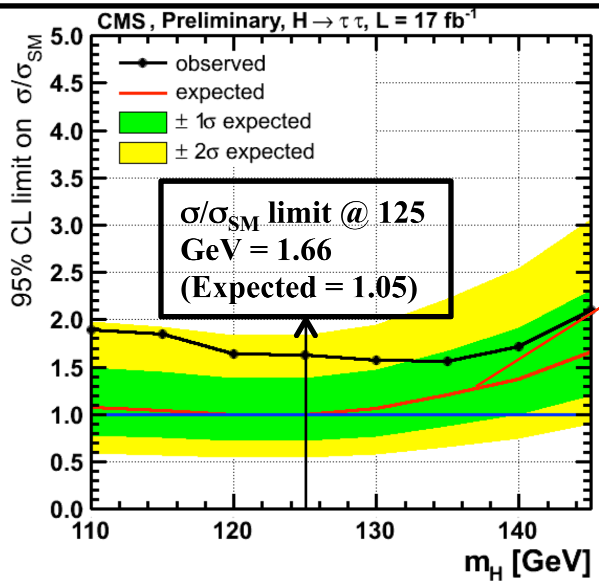


Data-driven methods for bkg estimation, e.g.:

- DY→$\tau\tau$: "embedding" of simulated tau decay in Z→μμ control sample (Z→μμ constrains also normalization)
- W+jets: sideband of m_T
- QCD: same sign events, fake rate method

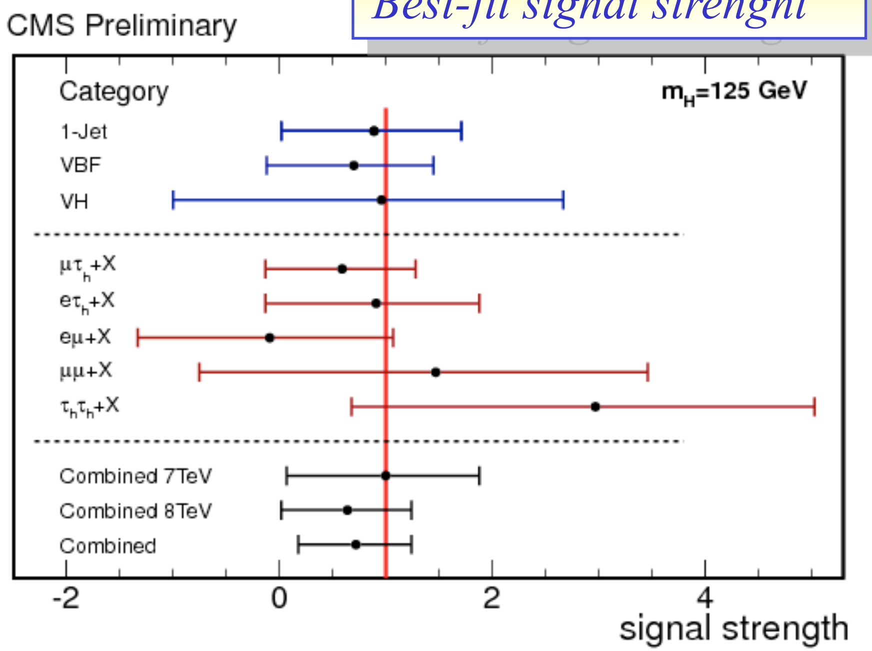
H→ττ channel: results

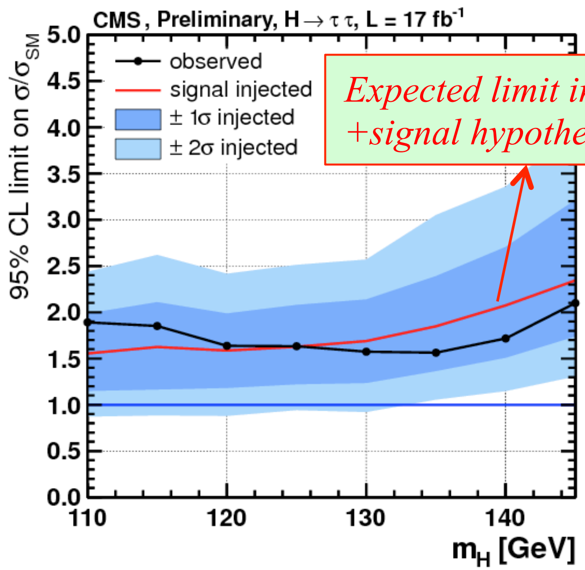
Plots include results from a dedicated VH analysis



Expected limit in bkg-only hypothesis

Best-fit signal strenght





Combined: $\sigma/\sigma_{SM} = 0.7 \pm 0.5$ @ 125 GeV

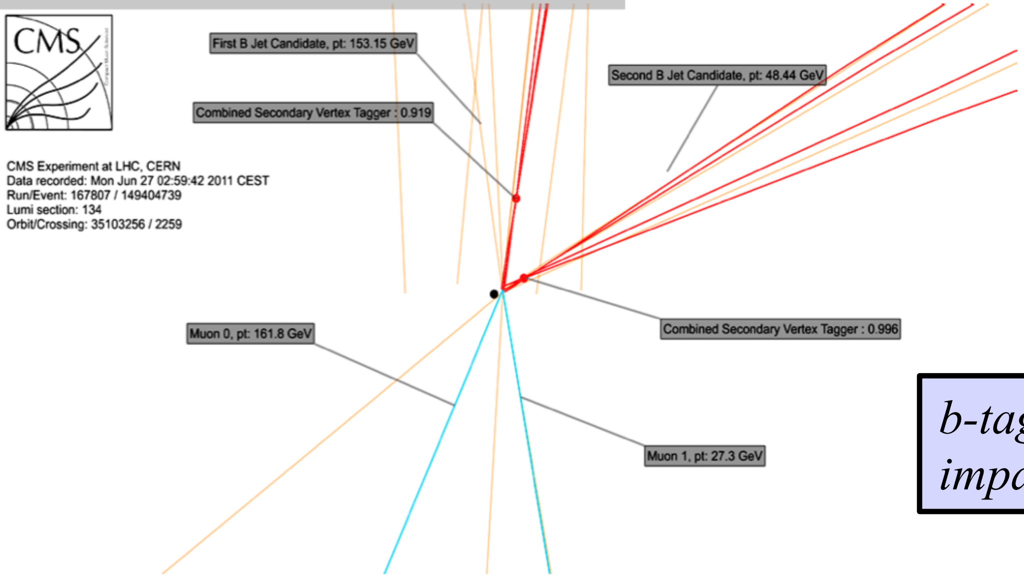
$H \rightarrow bb$

H→bb channel

- *Favoured BR at low mass, but huge bkg in inclusive channel (di-jet QCD production)*
- *Associated VH production channel must be exploited*
- *Channels considered: VH, with V=Z, W; Z→μμ, ee, νν; W→μν, eν*

PAS HIG-12-044

ZH→μμbb candidate event



Signature: two high p_T jets passing b-tagging requirements + leptons/MET from Z/W bosons

b-tagging from displaced vertex and track impact parameters within the jet

Background: mainly W/Z+jets and top; also diboson and multijet events

H→bb channel: event categories

Exploit boosted kinematics of the jj system and of V boson for bkg rejection

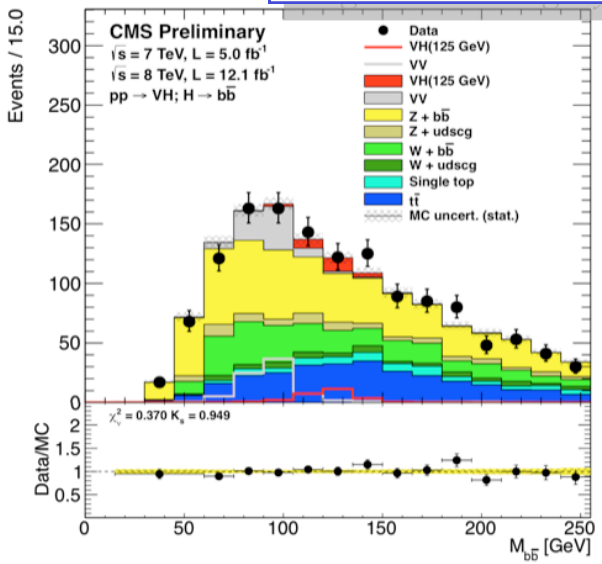
- Large p_T of di-jet and V
- Large $\Delta\Phi(V, H)$

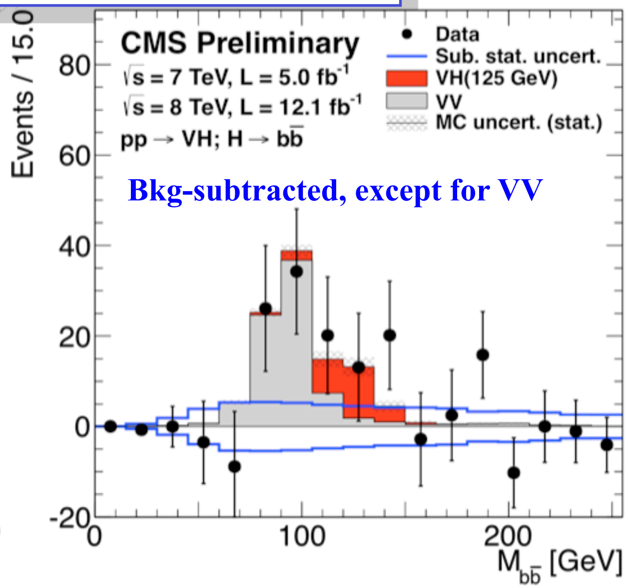
Splitting of events in low- and high-boost categories; analysis optimized for each category

Variable	$W(\ell\nu)H$	$Z(\ell\ell)H$	$Z(\nu\nu)H$
$p_T(V)$	[120 – 170] (> 170)	[50 – 100] (> 100)	–
E_T^{miss}		–	[130 – 170] (> 170)

Recalibration of jet energy using a MVA regression technique improves di-jet resolution by 10%

High-boost category, all channels combined



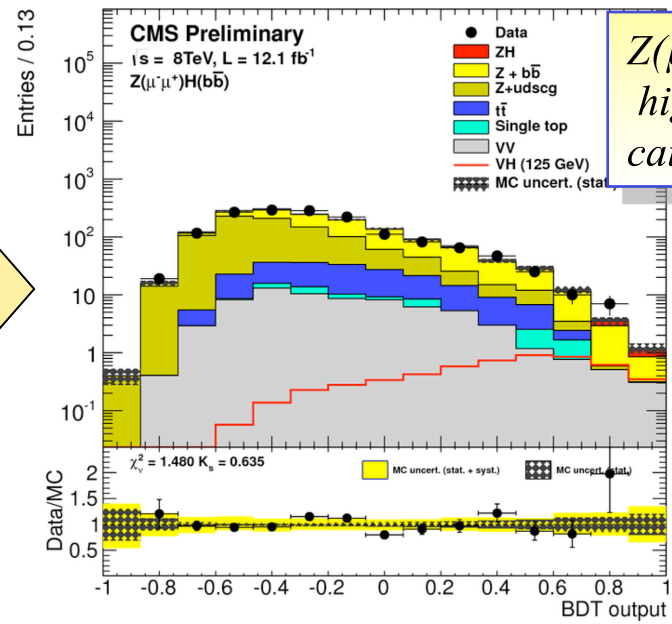


- Control regions defined for dominant bkg (V+jets, top)
- data/MC correction factors derived for bkg normalization

H→bb: signal extraction

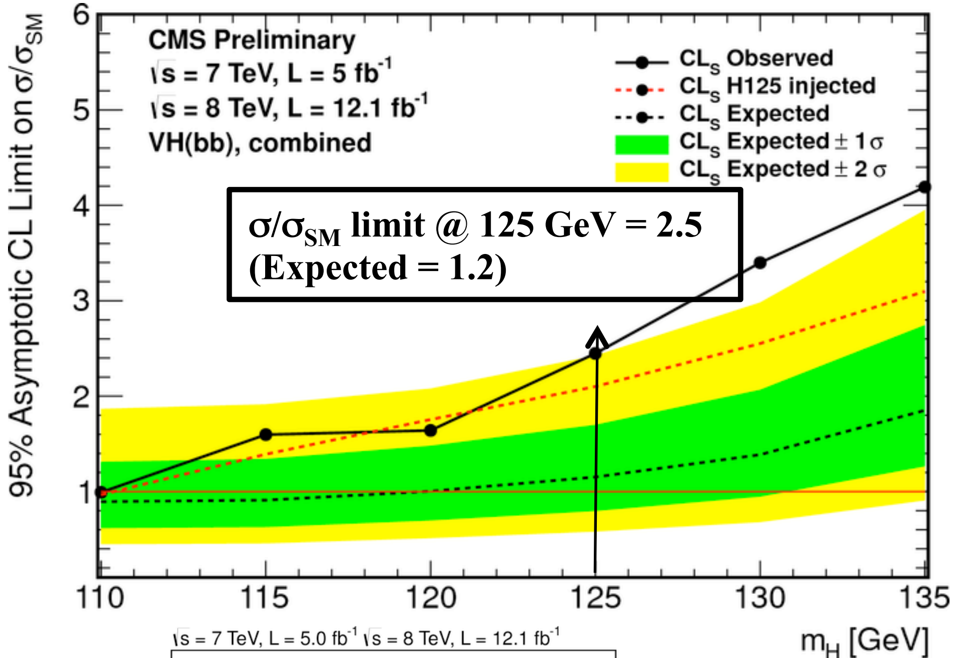
Fit distribution of a BDT discriminant using in input m_{jj}+additional kinematic variables

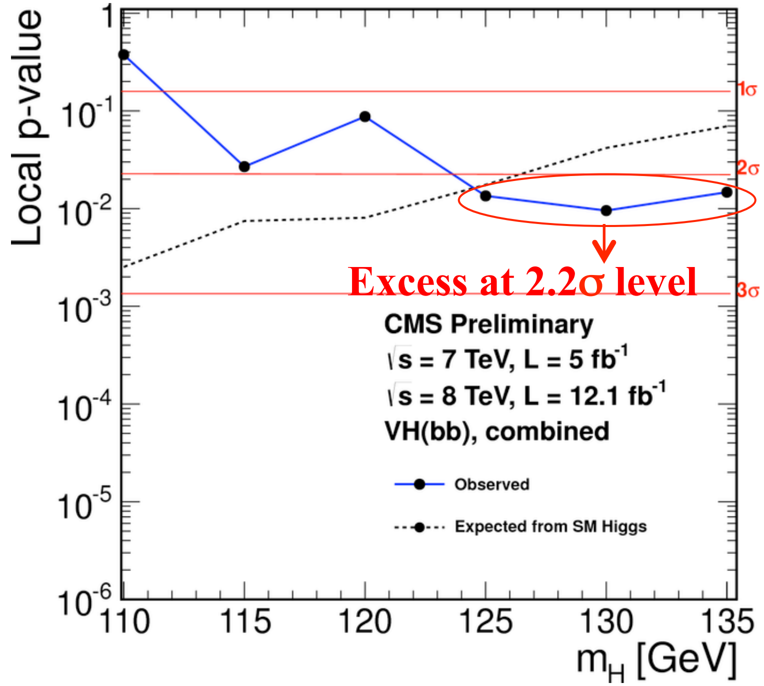
p_{Tj}: transverse momentum of each Higgs daughter
$m(\mathrm{jj})$: dijet invariant mass
$p_T(\mathrm{jj})$: dijet transverse momentum
$p_T(\mathrm{V})$: vector boson transverse momentum (or E_T^{miss})
$\mathrm{CSV_{max}}$: value of CSV for the Higgs daughter with largest CSV value
$\mathrm{CSV_{min}}$: value of CSV for the Higgs daughter with second largest CSV value
$\Delta\phi(\mathrm{V,H})$: azimuthal angle between V (or E_T^{miss}) and dijet
$|\Delta\eta(\mathrm{jj})|$: difference in η between Higgs daughters
$\Delta R(\mathrm{jj})$: distance in η–ϕ between Higgs daughters
N_{aj}: number of additional jets
$\Delta\phi(\mathrm{E_T^{miss}, jet})$: azimuthal angle between E_T^{miss} and the closest jet (only for $Z(\nu\nu)H$)
$\Delta\theta_{\mathrm{pull}}$: color pull angle [35] →Useful to determine if jets come from a color-singlet object

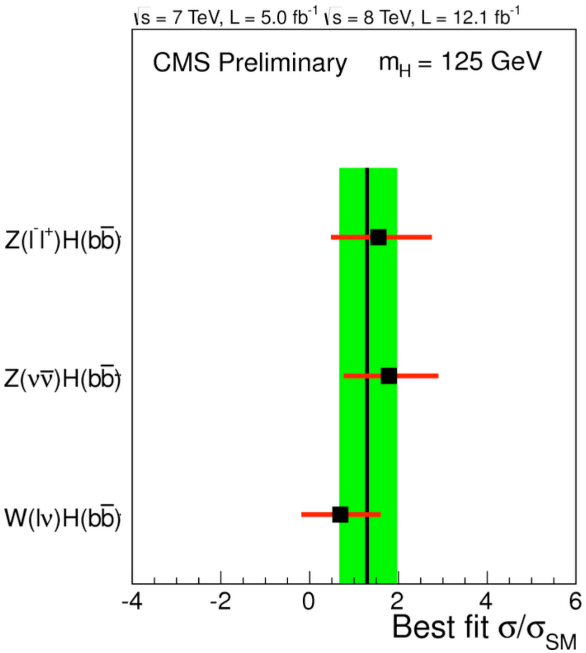


Z(μμ)H, high-boost category

H→bb channel: results





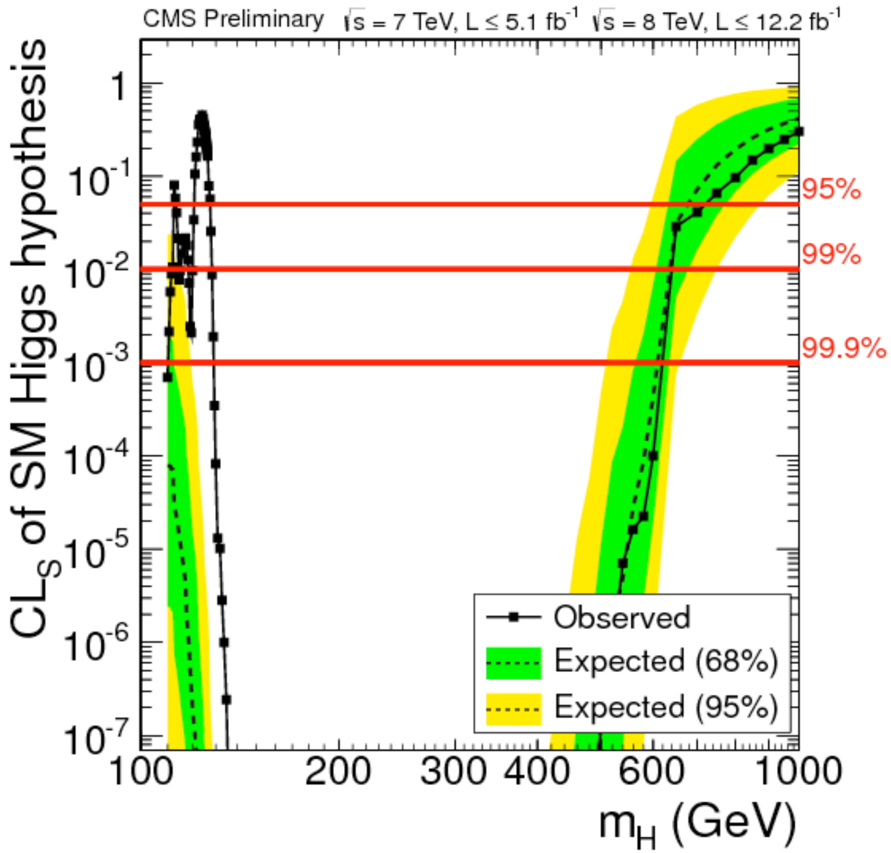


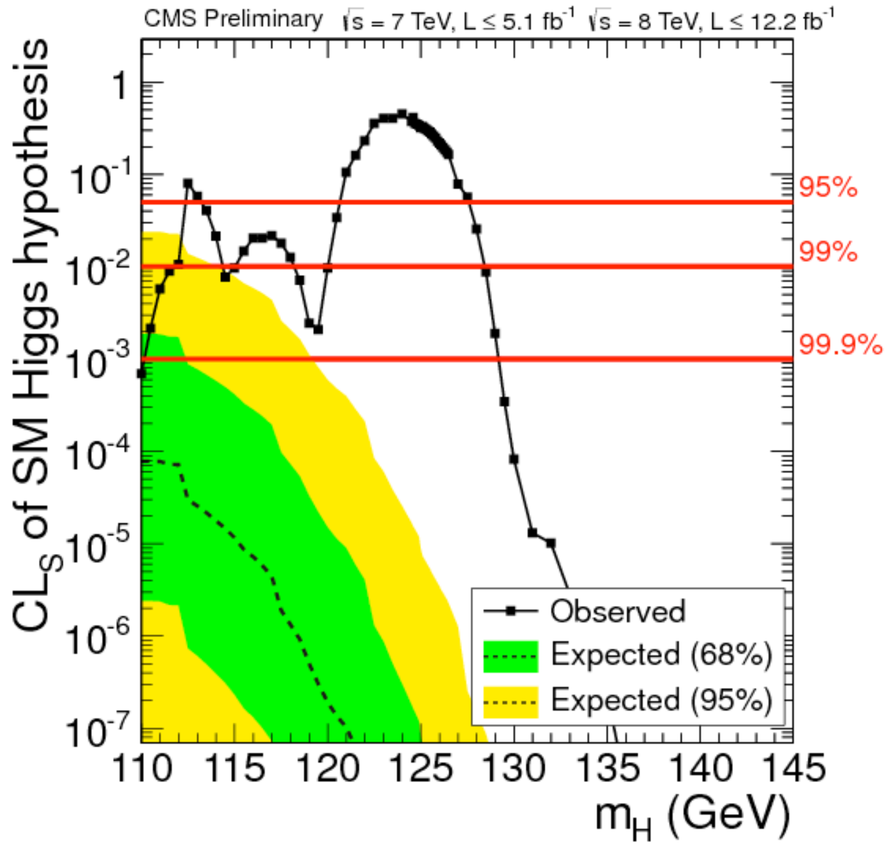
Combined: $\sigma/\sigma_{SM} = 1.3\ ^{+0.7}_{-0.6}$ *@ 125 GeV*

Combination

Exclusion limits

PAS HIG-12-045



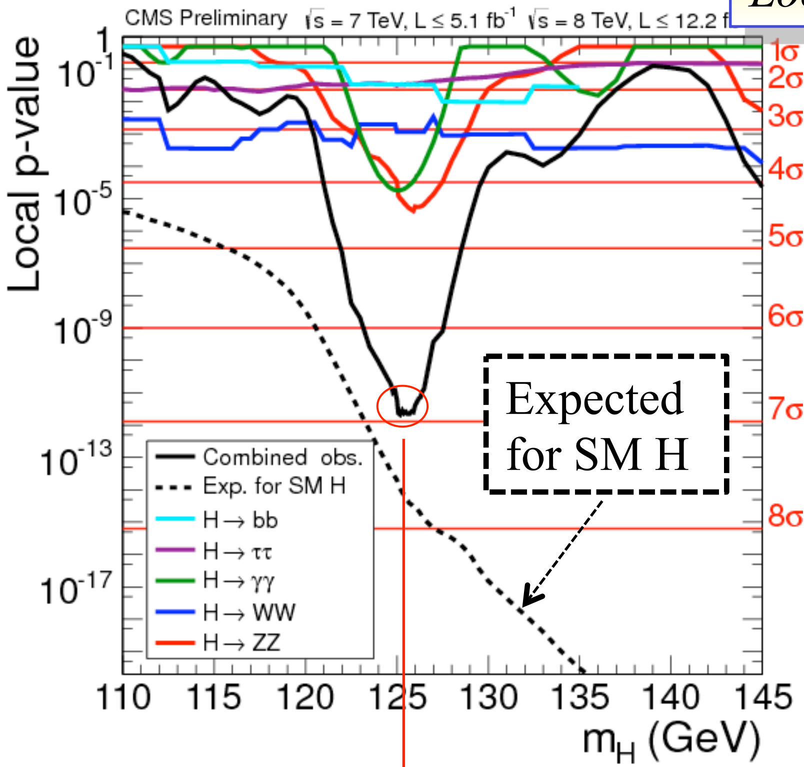


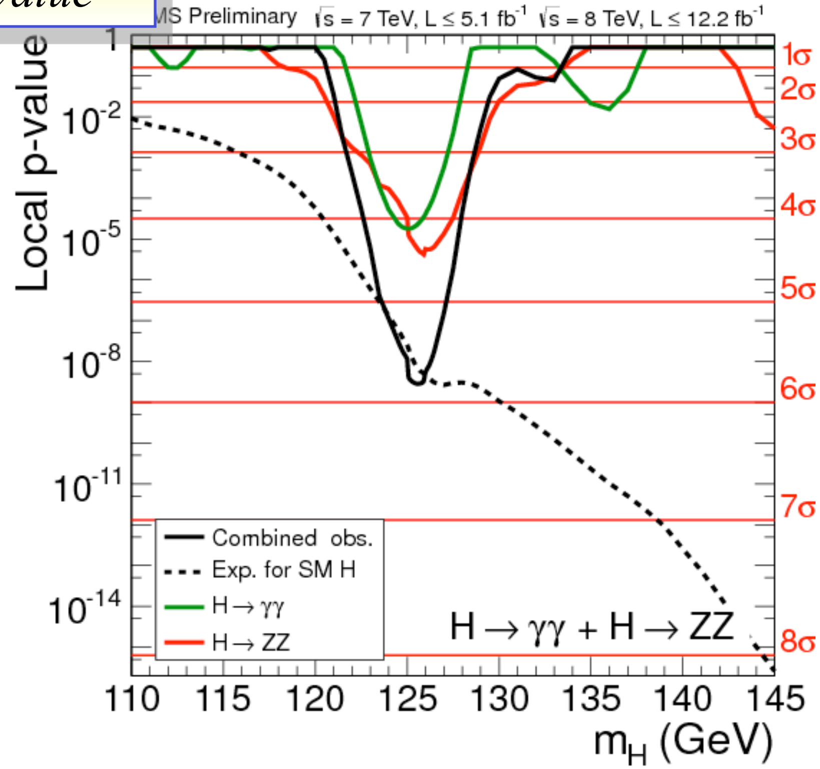
High mass:
m_H excluded @ 95%CL till to 700 GeV

Low mass:
interval 121-128 GeV not excluded @ 95%CL (excess in data due to the new boson)

Significance of the excess

Local p-value





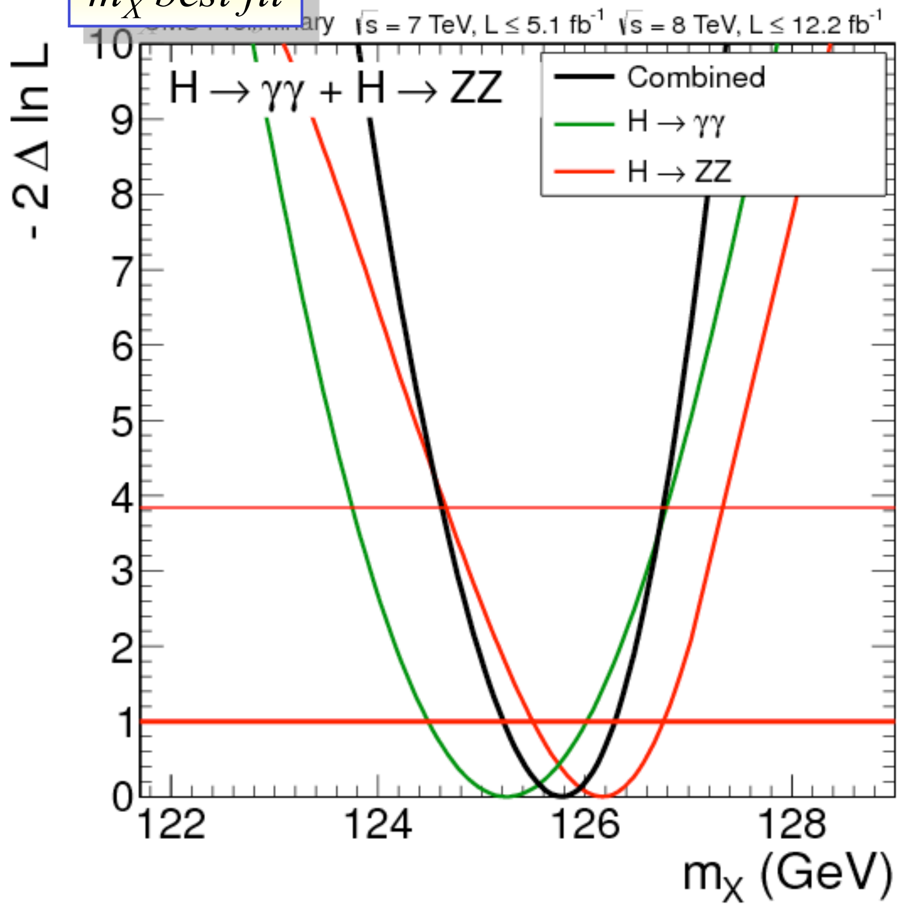
Observed significance: 6.9σ @ 125.8 GeV
Expected significance: 7.8σ @ 125.8 GeV

5.8*σ* significance from γγ and ZZ alone

Mass measurement

Combination of γγ and ZZ (high resolution channels)

m_X best fit

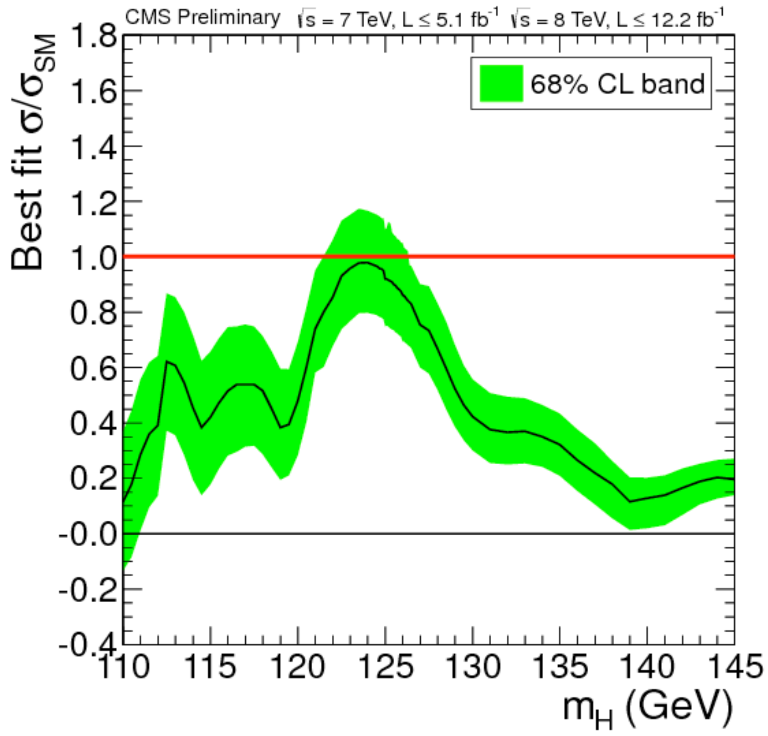


For m_X measurement, no SM constraints on the yield of γγ and ZZ

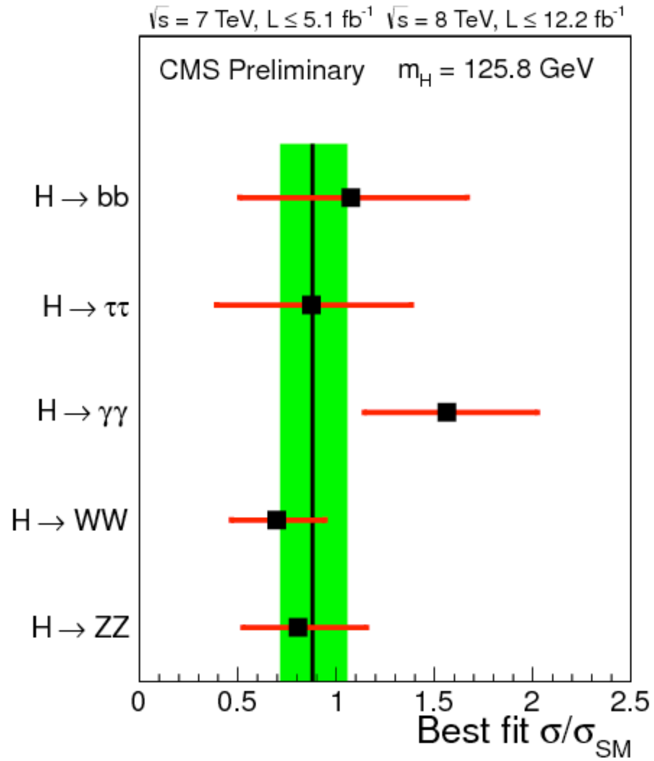
$m_X = 125.8 \pm 0.4(stat.) \pm 0.4(syst.)$ GeV

Signal strenght compatibility with SM

Signal strenght best fit



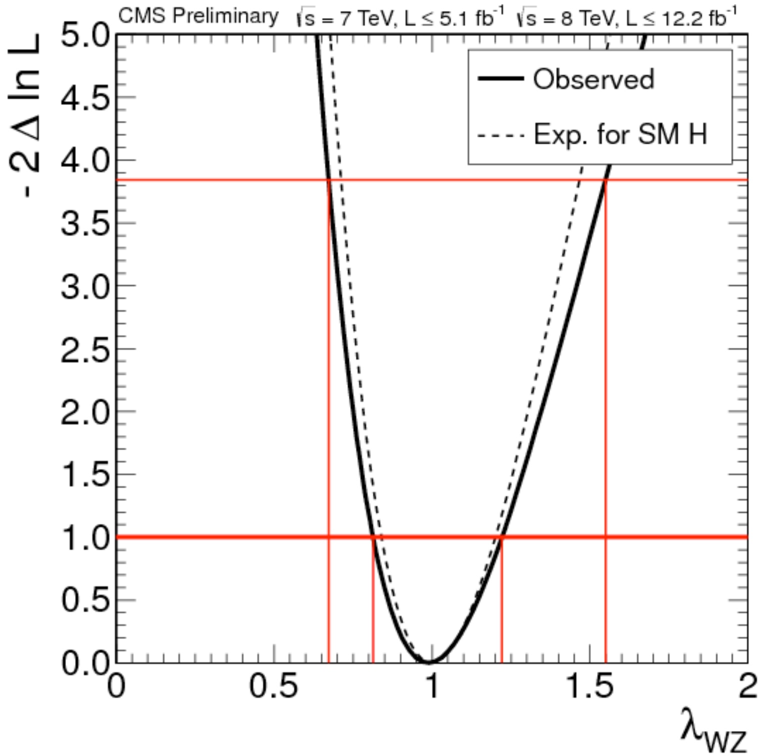
Compatibility of measurements in different channels



σ/σ_{SM} @ 125.8 GeV = 0.88 ± 0.21

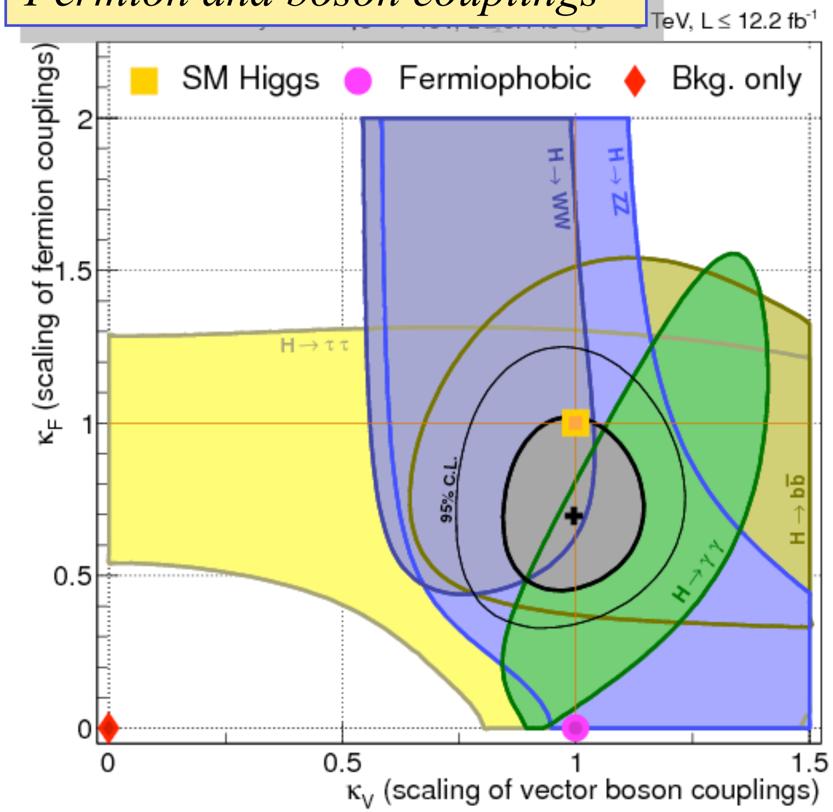
Compatibility with SM couplings

Introducing scale factors κ_i modifying the SM couplings (κ_i=1 for SM)



$\lambda_{WZ}=\kappa_W/\kappa_Z$ =1 in the SM
("custodial symmetry)

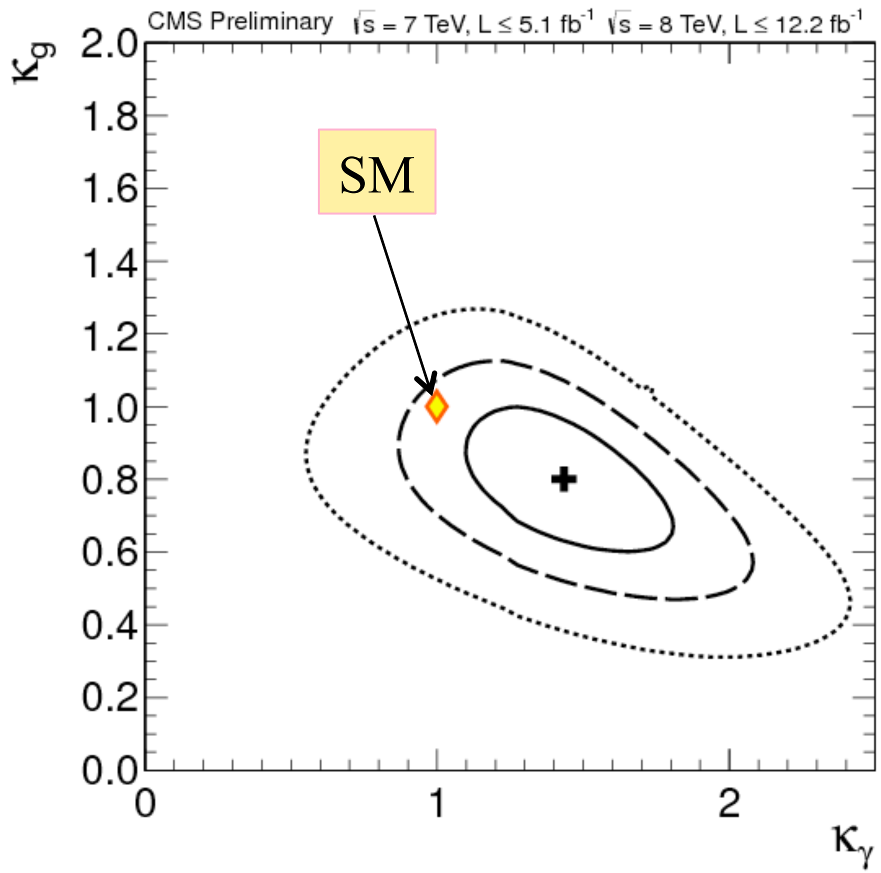
Fermion and boson couplings



Fermiophobic scenario excluded at CL > 95%

Compatibility with SM couplings

Introducing scale factors κ_i modifying the SM couplings (κ_i=1 for SM)



Couplings with gg and γγ proceeds via loop ➔ sensitive to new physics

Conclusions

- New boson observation @ 125 GeV reported last July
  - Question: is it the SM Higgs?
  - Main goal is to understand its properties and fully investigate all possible decay channels
- First answers based on 2011(~5 fb^{-1}) + 2012 (~12 fb^{-1}, still partial) data sample
  - 5 main channels investigated (a large multiplicity of sub-channels and categories)
- Excess of events is confirmed with significance of 6.9σ
  - Significance now > 3 σ in WW,
  - broad excess of events also in fermionic channels
- Mass best fit 125.8 ± 0.4(stat.) ± 0.4(syst.) GeV
  - signal strenght compatible with SM
- Events observed in ZZ channel favour the 0+ hypothesis vs. 0-
- Still more data from the full 2012 needs to be analyzed!

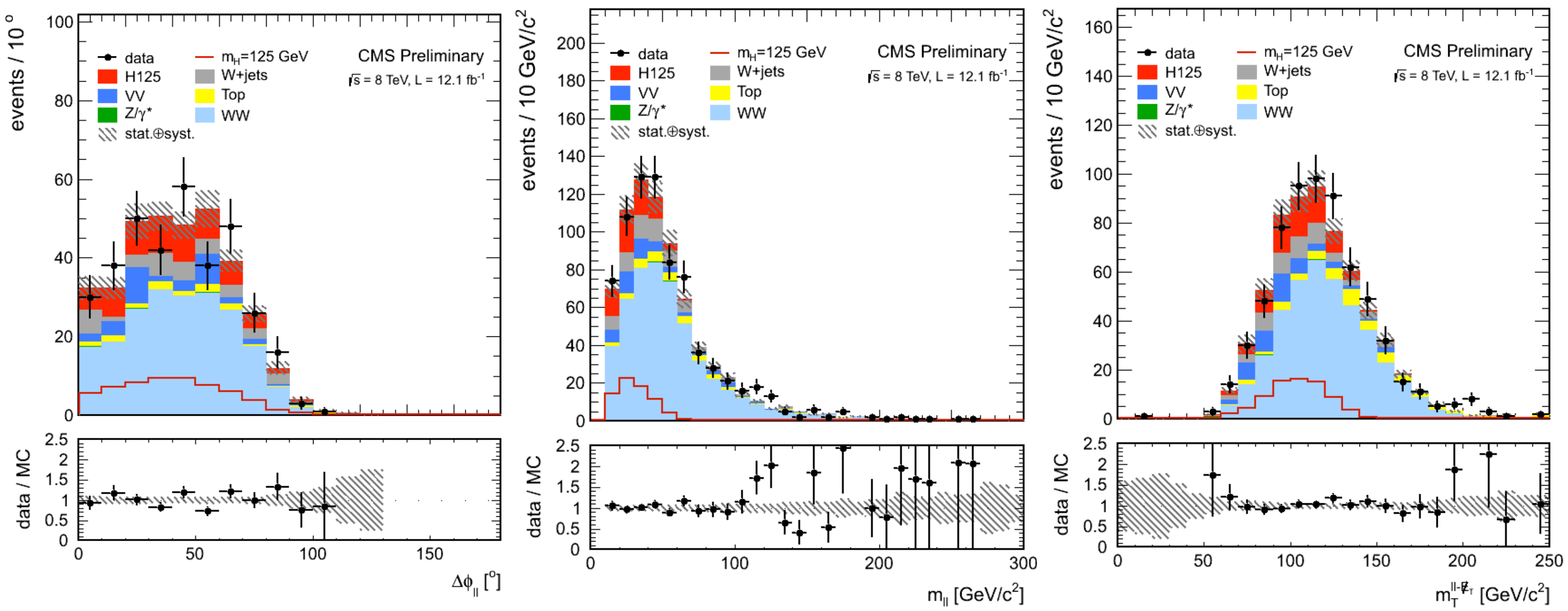
Backup slides

H→WW→2l2ν: bkg determination

Data driven bkg estimates:

- *W+jets -> using jet fake-rates applied to control sample (one lepton failing ID)*
- *Top -> "top-tagging" of events*
- *WW -> use control region m_{ll}>100 GeV, extrapolation to signal region*
- *DY bkg normalization from number of events within 7.5 GeV around Z-boson mass*

Uncertainty on the bkg determination is the largest source of systematics (along with theoretical uncertainty on Higgs cross-section)



Systematic uncertainties

Systematics are, of course, channel and analysis dependent

The following table summarizes sources and uncertainties in a particular analysis as an example

Table from H→bb analysis

Source	Range
Luminosity	2.2-4.4%
Lepton efficiency and trigger (per lepton)	3%
$Z(\nu\nu)H$ triggers	3%
Jet energy scale	2–3%
Jet energy resolution	3–6%
Missing transverse energy	3%
b-tagging	3–15%
Signal cross section (scale and PDF)	4%
Signal cross section (p_T boost, EWK/QCD)	5–10% / 10%
Signal Monte Carlo statistics	1-5%
Backgrounds (data estimate)	$\approx$ 10%
Single-top (simulation estimate)	15–30%
Dibosons (simulation estimate)	30%

Limited statistics of the control samples

$H\to\tau\tau$: event selection

For each channel, further splitting of events in categories by counting jets with p_T >30 GeV

- 0-jet used only to constrain bkg
- 1-jet, high τ p_T
- 1-jet, low τ p_T
- 2 jet (p_T >30) – VBF

Expected and observed events in $\mu\tau_h$ channel for each category

Process	*0-Jet*	*1-Jet*	*VBF*
$Z\to\tau\tau$	64948 ± 3628	13292 ± 745	80 ± 9
QCD	12515 ± 622	3156 ± 222	32 ± 5
W+jets	5160 ± 615	3610 ± 233	36 ± 3
Z+jets (l/jet faking τ)	1200 ± 193	452 ± 52	1 ± 0.4
$t\bar{t}$	8 ± 0.8	350 ± 34	4 ± 1
Dibosons	95 ± 10	216 ± 25	1 ± 0.4
Total Background	83926 ± 3736	21076 ± 815	154 ± 10
$H\to\tau\tau$	-	112 ± 9	8 ± 0.9
Data	81297	21107	174

Signal Eff.

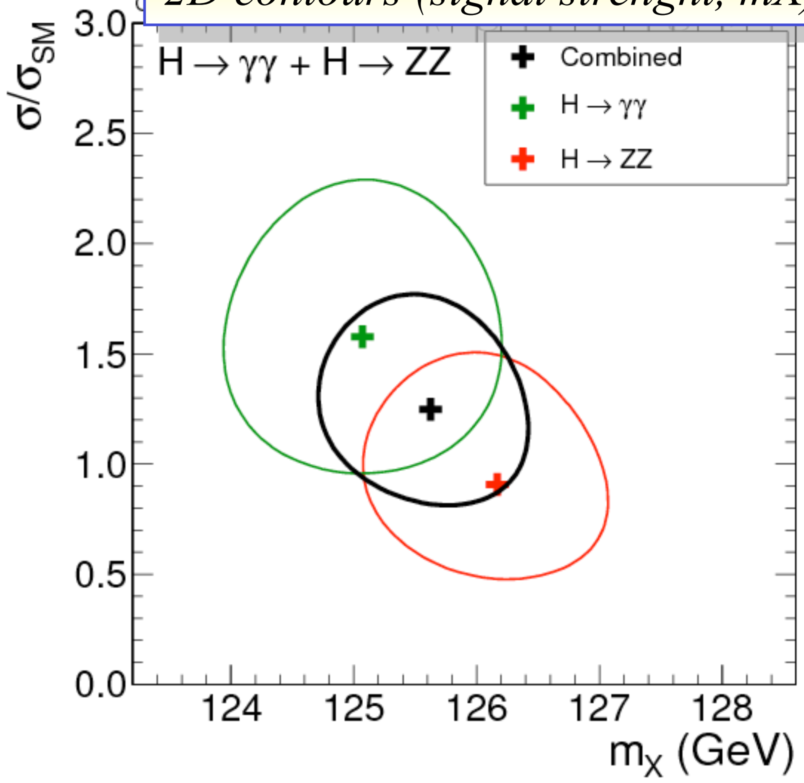
$gg\to H$	-	$4.50 \cdot 10^{-3}$	$8.41 \cdot 10^{-5}$
$qq\to H$	-	$9.75 \cdot 10^{-3}$	$3.90 \cdot 10^{-3}$
$qq\to Ht\bar{t}$ or VH	-	$6.19 \cdot 10^{-3}$	$1.52 \cdot 10^{-5}$

Final selection criteria optimized in each channel/ category

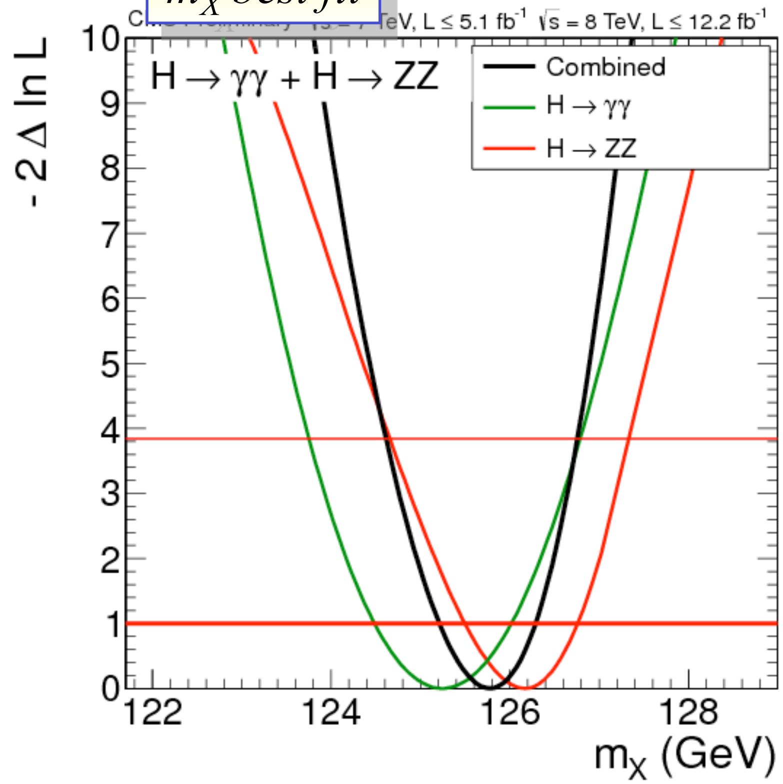
Mass measurement

Combination of $\gamma\gamma$ and ZZ (high resolution channels)

2D contours (signal strenght, mX)



m_X best fit



In the combination the relative yield of $\gamma\gamma$ and ZZ is fixed to SM

For m_X measurement, no SM constraints (the relative yield of $\gamma\gamma$ and ZZ is not fixed)

m_X = 125.8±0.4(stat.)±0.4(syst.) GeV