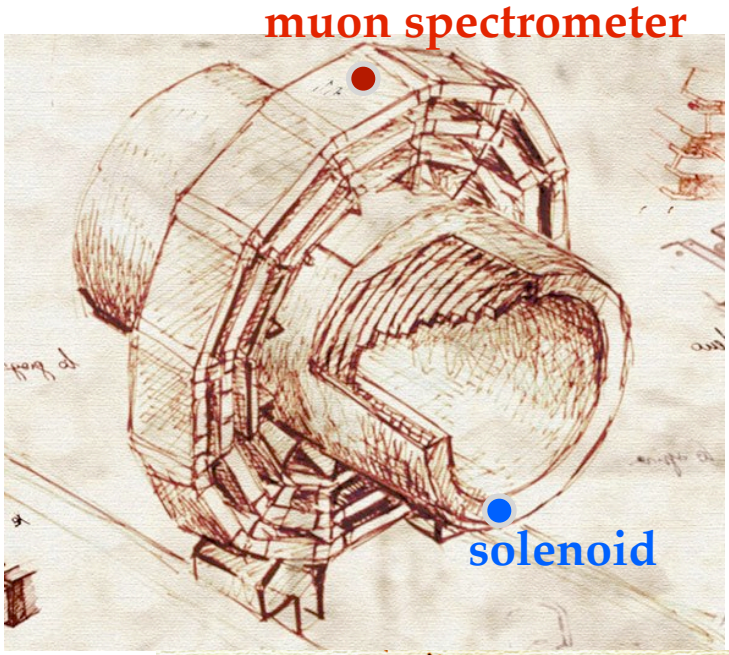
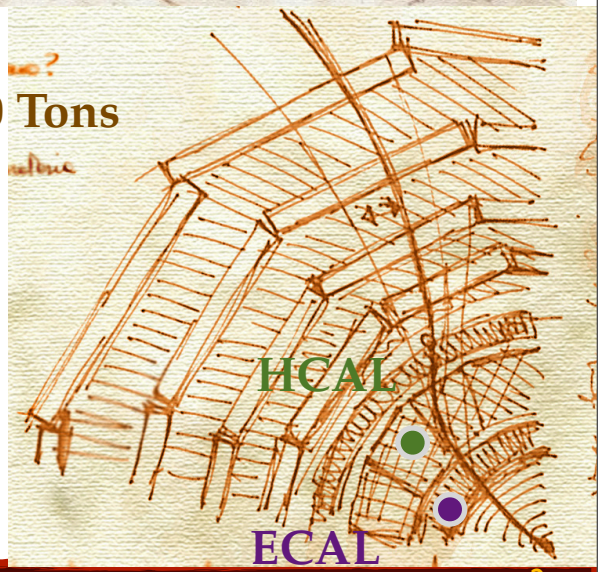
Standard Model Higgs search with CMS

Emanuele Di Marco
IFAE, Ferrara 11-13 April 2012

The CMS detector

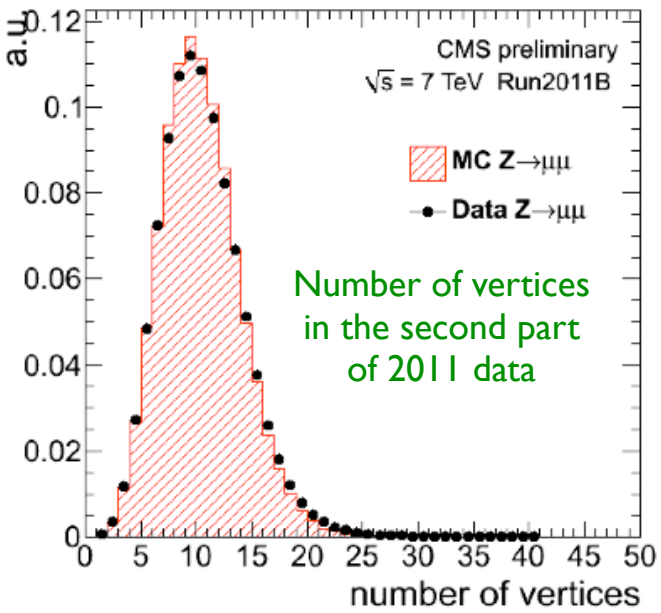
- Silicon tracker $|\eta| < 2.5$, B = 3.8T solenoid
  - $\sigma/p_T \approx 1.5 \cdot 10^{-4}\, p_T \oplus 0.005$
- Calorimetry $|\eta|_{ECAL} < 3$, $|\eta|_{HCAL} < 5$
  - ECAL: $PbWO_4$ crystals, high resolution $M(\gamma\gamma)$
    - $\sigma/E \approx 2.8\%/\sqrt{E} \oplus 12\%/E \oplus 0.3\%$
  - HCAL: brass+Scintillator
    - $\sigma/E \approx 100\%/\sqrt{E} \oplus 5\,\%$
- Muon Spectrometer: B = 2T (solenoid return yoke)
  - $\sigma/p_T \approx 1\%$ @ 40 GeV, $|\eta| < 2.4$

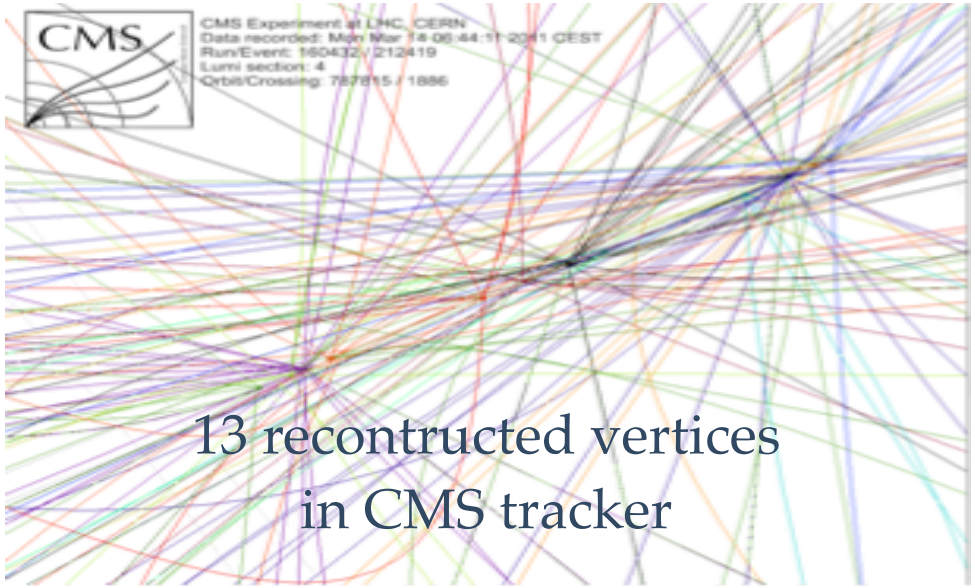




2011 dataset and reconstruction

- Excellent performance of LHC and CMS in 2011
  - More than 5 fb^{-1} of pp collisions collected at 7 TeV CM energy
  - Peak luminosity 3.5 x 10^{33} $cm^{-2}s^{-1}$
  - Data taking efficiency 90%
  - 90% of collected data good for all analyses
  - Mean pileup 10 events

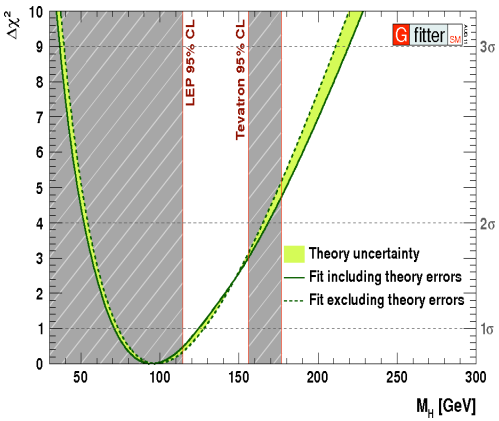


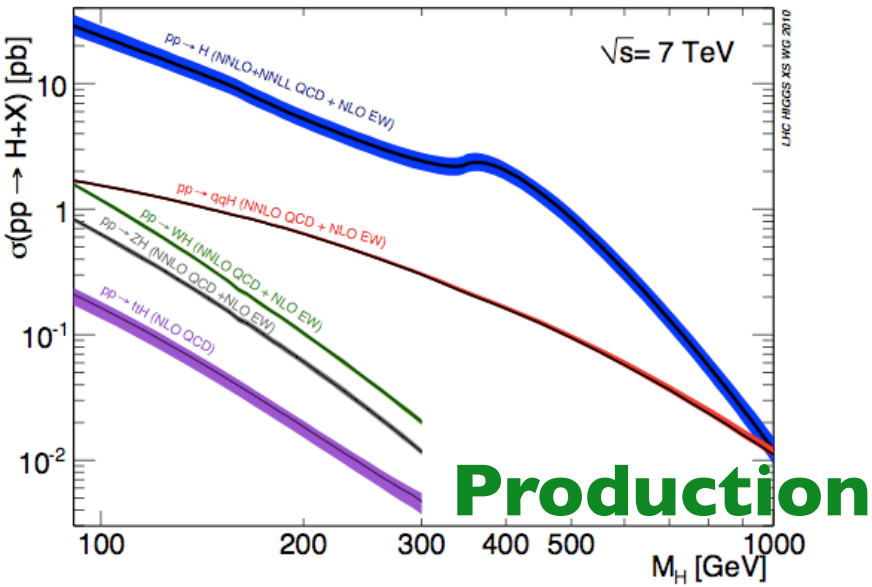


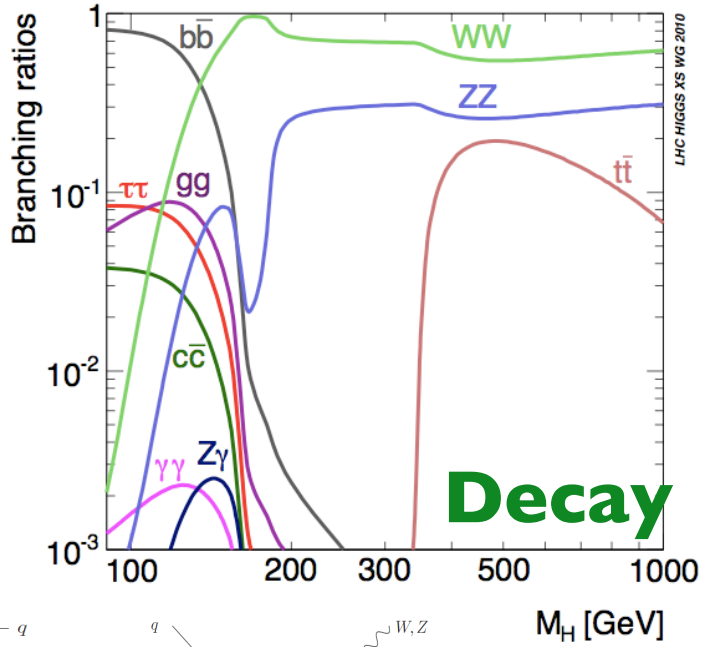
- A Particle Flow algorithm (PF) has been developed
  - Provides a global event description in form of a list of particles and avoids double-counting
- Large improvements in τ, jet and missing transverse energy (MET) measurement

Higgs boson at LHC

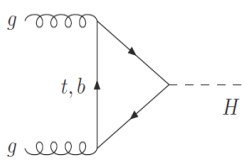
- Indirect constraints from precision EW measurements
  - M_H < 169 GeV at 95% CL (standard fit)
  - M_H < 143 GeV at 95% CL (including direct searches, before LHC)
- Direct searches
  - LEP: M_H < 114.4 GeV excluded at 95% CL
  - Tevatron: latest results in previous talks
- SM scalar boson favored at low mass, above the LEP limit



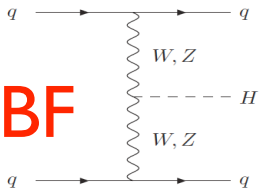




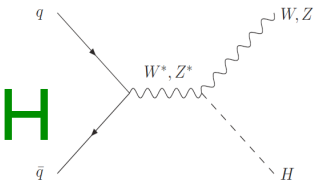
gg→H



VBF



VH

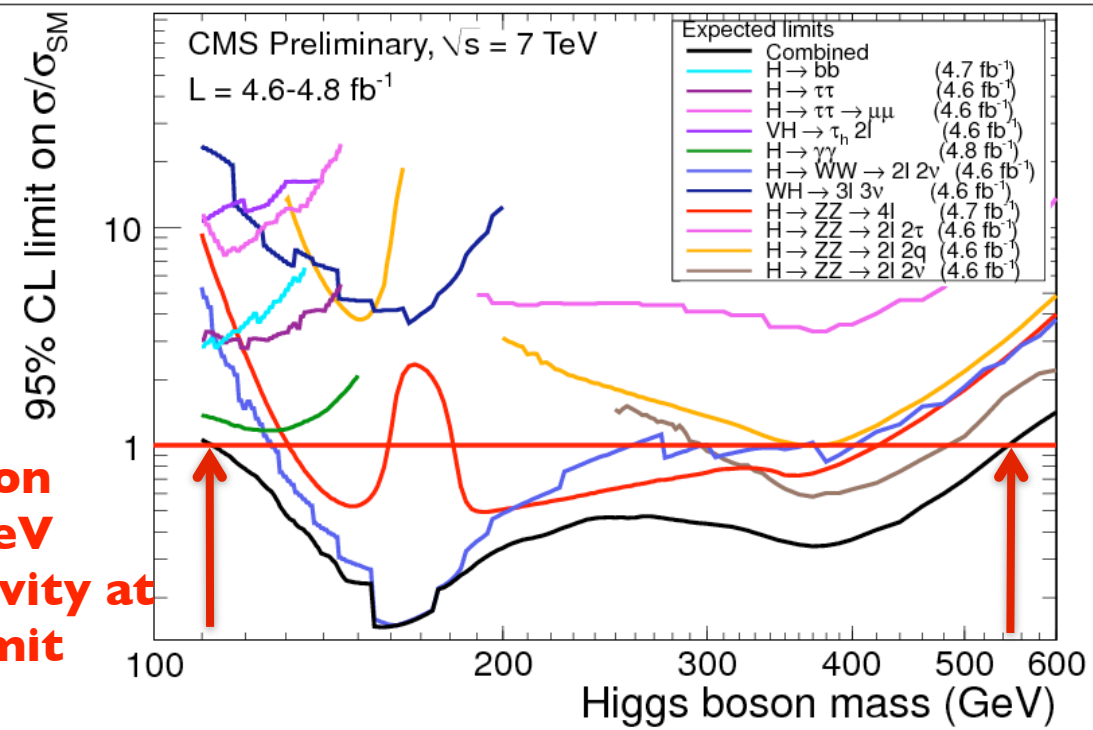


Higgs boson channel at CMS

Low mass

High mass

	Channel	m_H range (GeV)	Luminosity (fb^{-1})	Sub-channels	m_H resolution
new	$H \to \gamma\gamma$	110–150	4.8	2	1–2%
	$H \to \tau\tau \to e\tau_h/\mu\tau_h/e\mu + X$	110–145	4.6	9	20%
new	$H \to \tau\tau \to \mu\mu + X$	110–140	4.5	3	20%
new	$WH \to e\mu\tau_h/\mu\mu\tau_h + \nu$'s	100–140	4.7	2	20%
	$(W/Z)H \to (e\nu/\mu\nu/ee/\mu\mu/\nu\nu)(bb)$	110–135	4.7	5	10%
	$H \to WW^* \to 2\ell 2\nu$	110–600	4.6	5	20%
new	$WH \to W(WW^*) \to 3\ell 3\nu$	110–200	4.6	1	20%
	$H \to ZZ^{(*)} \to 4\ell$	110–600	4.7	3	1–2%
	$H \to ZZ^{(*)} \to 2\ell 2q$	130–164 200–600	4.6	6	3% 3%
	$H \to ZZ \to 2\ell 2\tau$	190–600	4.7	8	10–15%
	$H \to ZZ \to 2\ell 2\nu$	250–600	4.6	2	7%



Expected combined 95% exclusion 114.5-543 GeV

Exclusion sensitivity at LEP lower limit

Search mass range 110-600 GeV

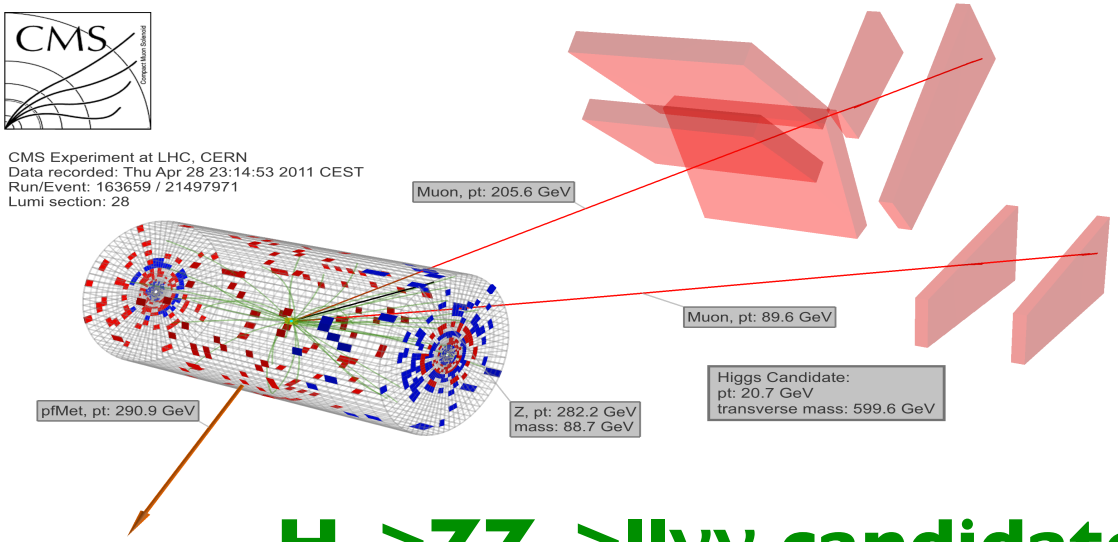
High mass channels

$H \rightarrow ZZ \rightarrow 2l2\nu$

$H \rightarrow ZZ \rightarrow 2l2q$

H→ZZ→2l2ν

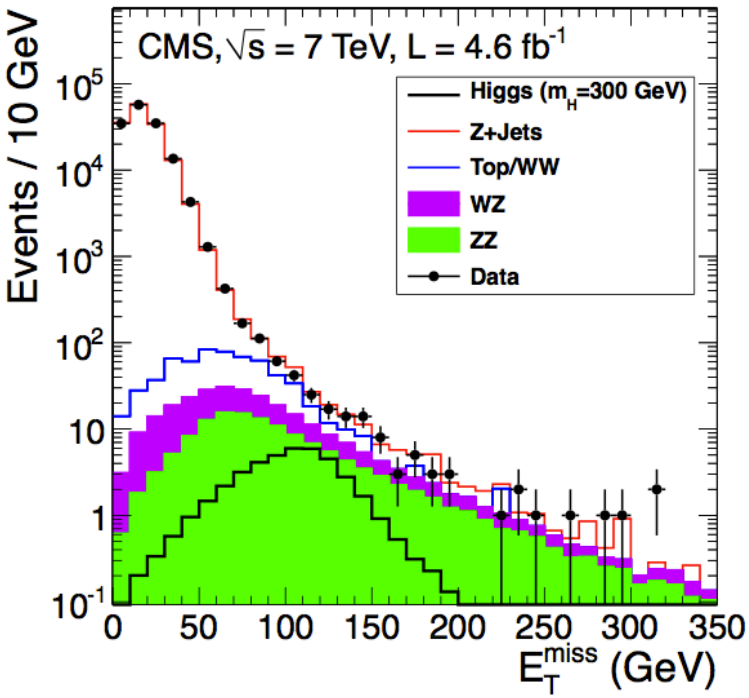
- Most sensitive channel for high mass search
  - BR 6 times larger than ZZ→4l
- **Neutrinos implies mass resolution 7%**
- Only accessible for high mass (M_H >~ 250 GeV):
  - the two Z bosons are boosted
  - large MET due to invisible decay
- Main backgrounds
  - ZZ (irreducible)
  - Z+jets, tt, WZ



H->ZZ->llνν candidate

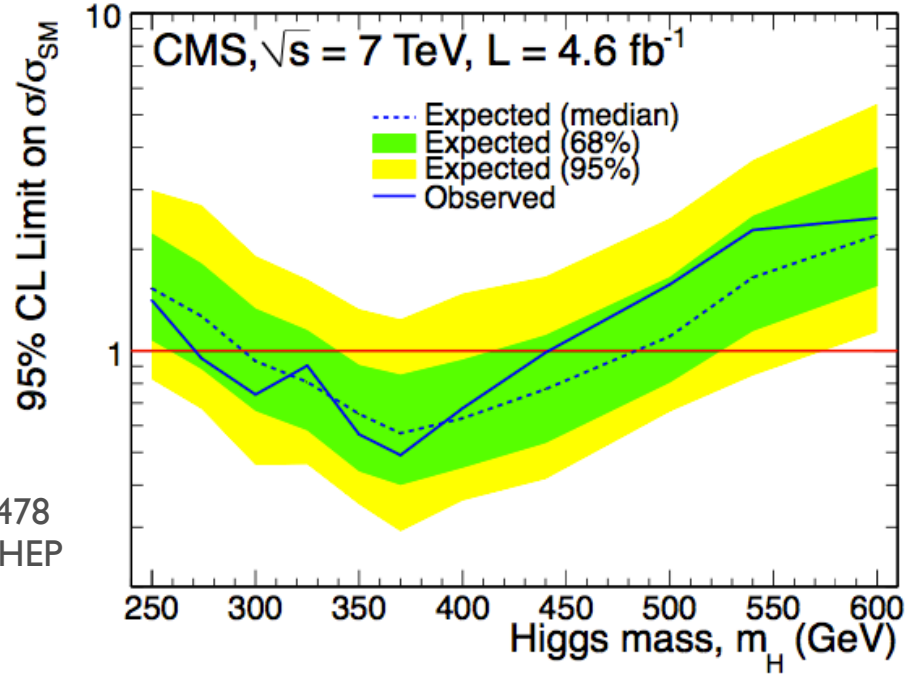
H→ZZ→2l2ν: results

MET after preselection



- Two analyses: cut and count and M_T fit
- Transverse mass definition:

$$M_T^2 = \left(\sqrt{p_T(\ell\ell)^2 + M(\ell\ell)^2} + \sqrt{E_T^{miss\,2} + M(\ell\ell)^2}\right)^2 - (\vec{p}_T(\ell\ell) + \vec{E}_T^{miss})^2$$



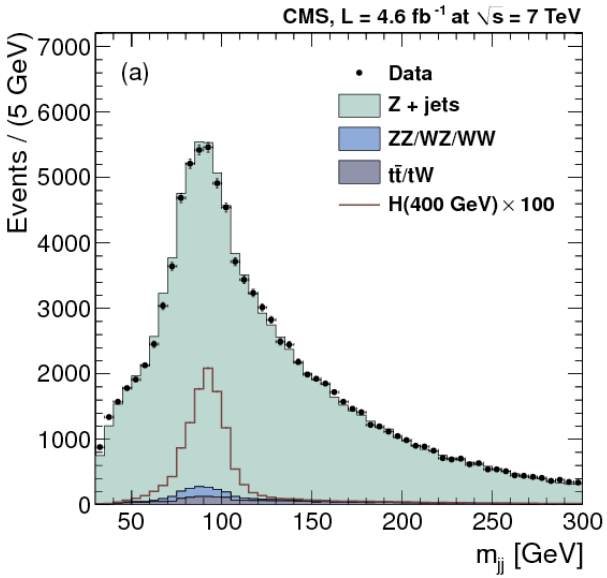
arXiv:1202.3478
Published by JHEP

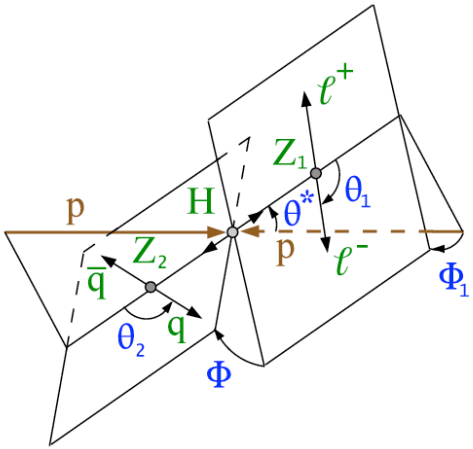
- BG estimation
  - Z+jets estimated using γ+jet to model the MET distribution
  - Non resonant BG normalization from eμ events
  - ZZ and WZ from MC

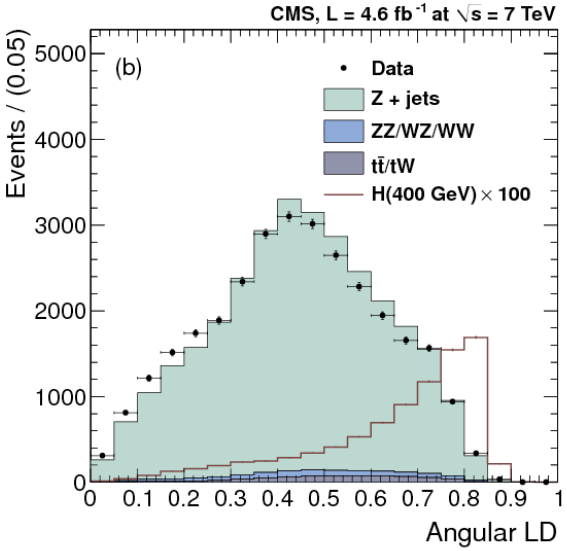
- No excess observed in data
- Expected 95% exclusion for M_H in [290-480] GeV
- Observed 95% CL exclusion for M_H [270-440] GeV

H→ZZ→2l2q

- Two opposite-charge, same-flavour leptons consistent with Z mass
- Two jets in Z peak with p_T > 30 GeV, within tracking acceptance
  - Analysis divided in 0/1/2 b-tag event categories
  - Use of quark-gluon discriminant based on jet shape/constituents
  - Additional use of angular likelihood discriminant at high mass



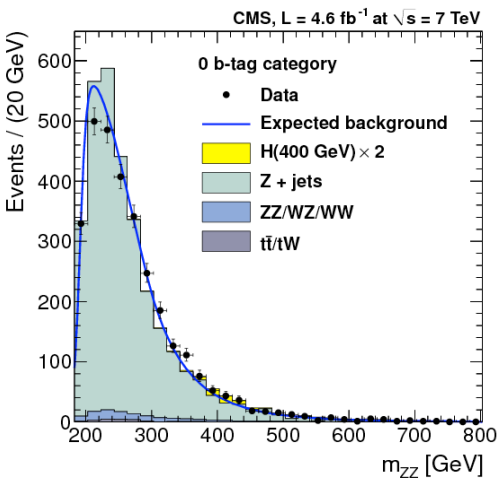


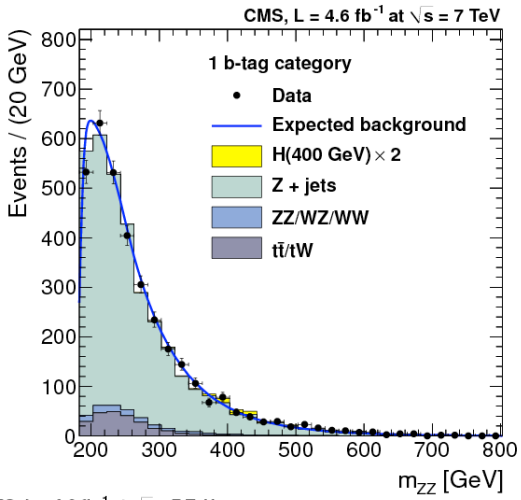


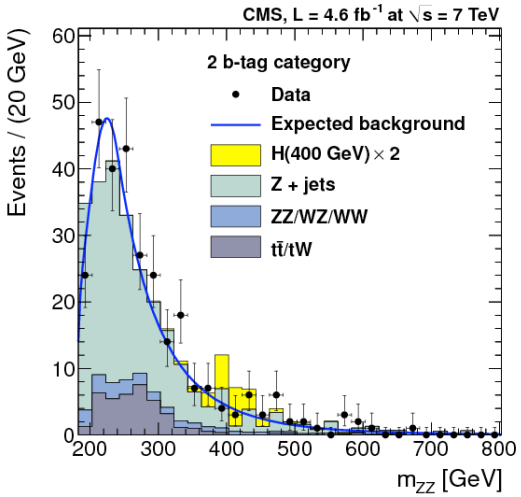
- Dominant Z+jets background estimated from m_{jj} sidebands

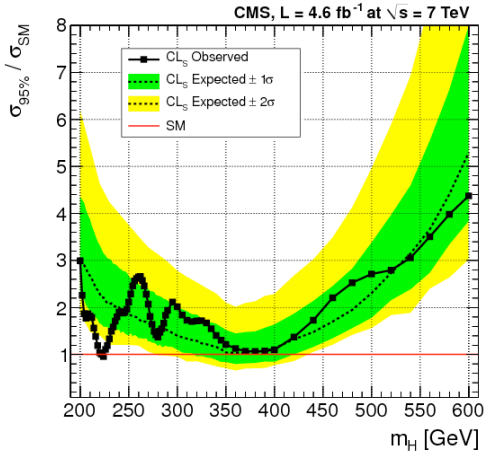
H→ZZ→2l2q: results

- Full distribution of mZZ used to extract the signal (default) OR cut-and count
- Background modeled from m_{jj} sidebands: ($60<m_{jj}<75$ GeV) ∪ ($105<m_{jj}<130$ GeV)
- eμjj used also as control sample (Z+jets free)









Sensitivity approaching SM cross-section in high mass region

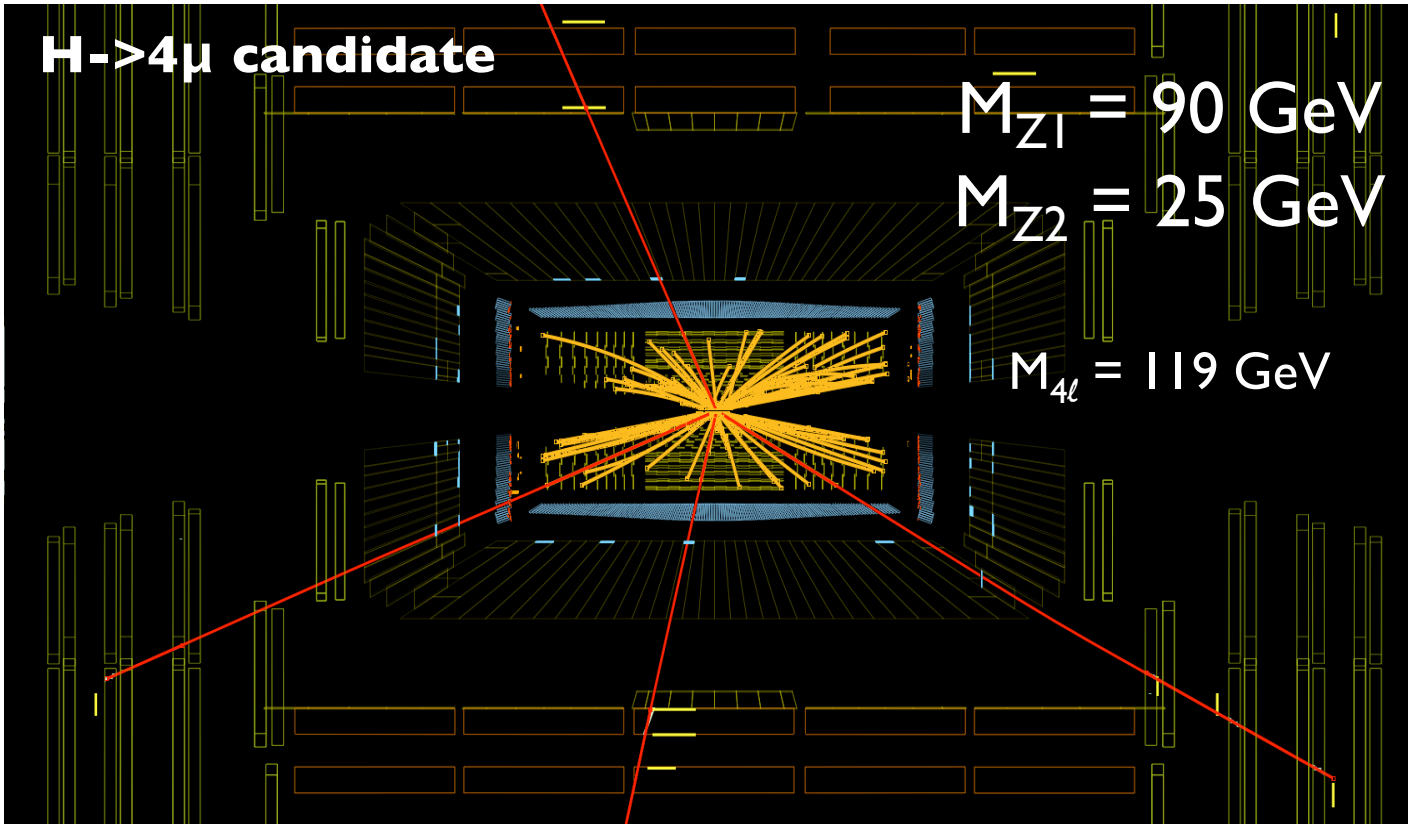
Higgs→ZZ*: the golden mode

search-range: [110 - 600] GeV
max sensitivity: 180 GeV
ZZ→4l: clean signature, narrow peak

H→ZZ→ 4l (4μ, 4e, 2e2μ)

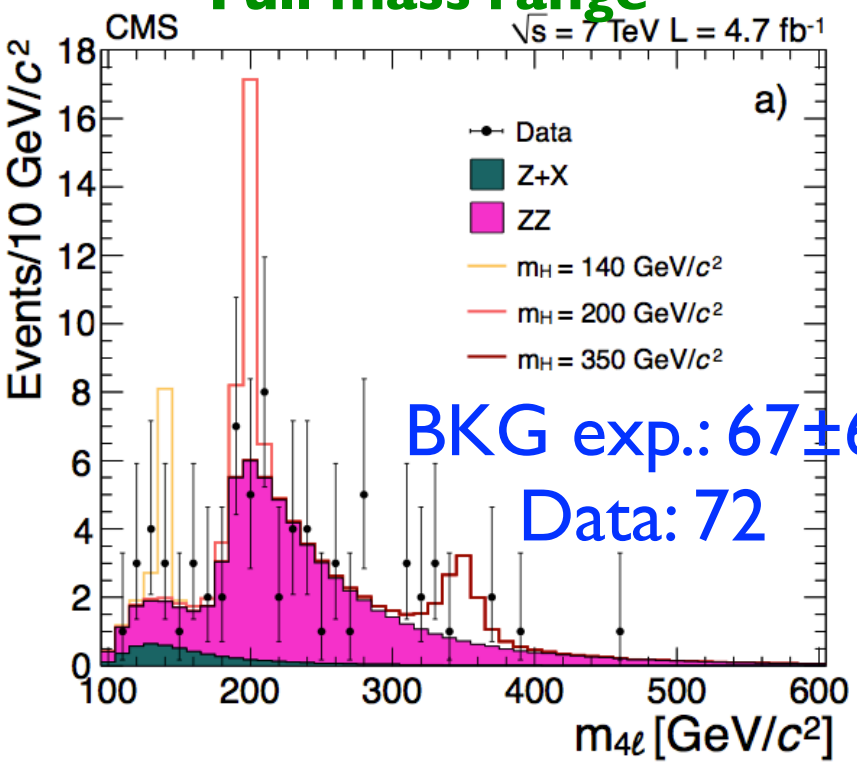
- Clean channel: 2 high mass pairs of isolated electrons or muons
- **Narrow mass peak**
  - Very good mass resolution 1-2 %
- Small BR ~10^{-3} at high mass

- Background
  - irreducible: ZZ, Reducible: Z+jets, Zbb, tt, WZ
- Most important aspect:
  - highest possible lepton reconstruction and id efficiency down to very low P_T



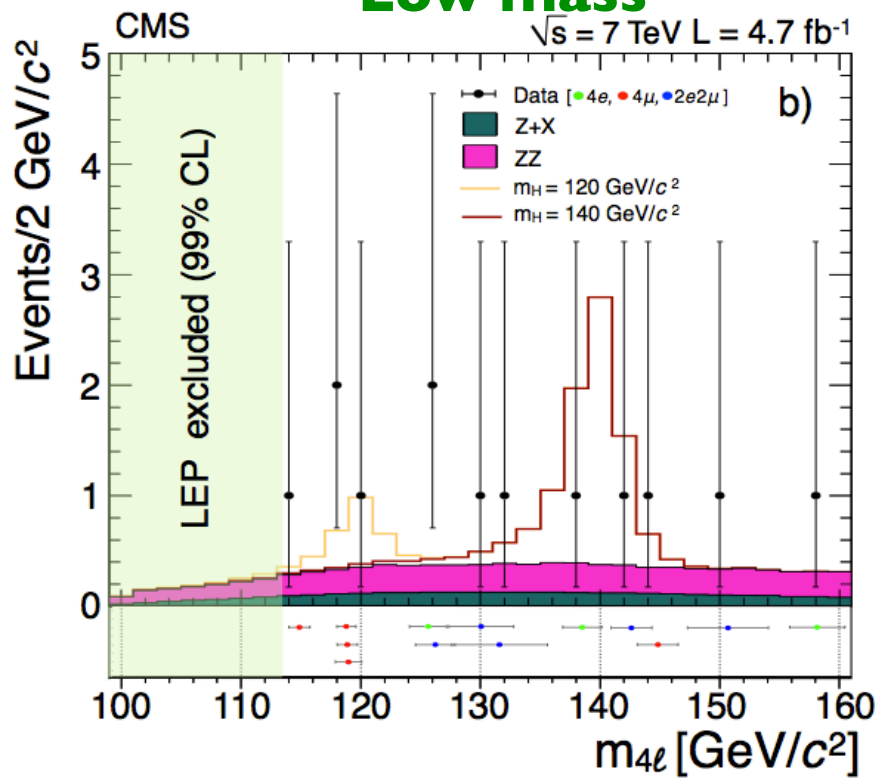
H→ZZ→ 4l: mass spectrum

Full mass range



BKG exp.: 67±6
Data: 72

Low mass



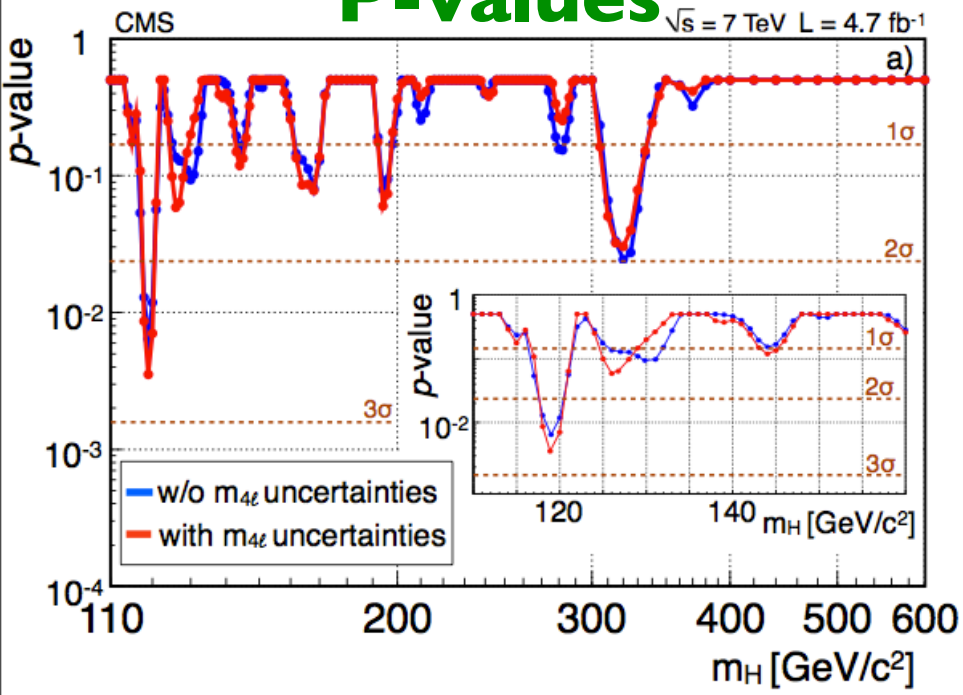
Channel	$4e$	4μ	$2e2\mu$
ZZ continuum	12.27 ± 1.16	19.11 ± 1.75	30.25 ± 2.78
Z+X	1.67 ± 0.55	1.13 ± 0.55	2.71 ± 0.96
All background	13.94 ± 1.28	20.24 ± 1.83	32.96 ± 2.94
$m_H = 120\,\mathrm{GeV}/c^2$	0.25	0.62	0.68
$m_H = 140\,\mathrm{GeV}/c^2$	1.32	2.48	3.37
$m_H = 350\,\mathrm{GeV}/c^2$	1.95	2.61	4.64
Observed	12	23	37

BKG exp.: 9.5±1.3
Data: 13
In 100-160 GeV

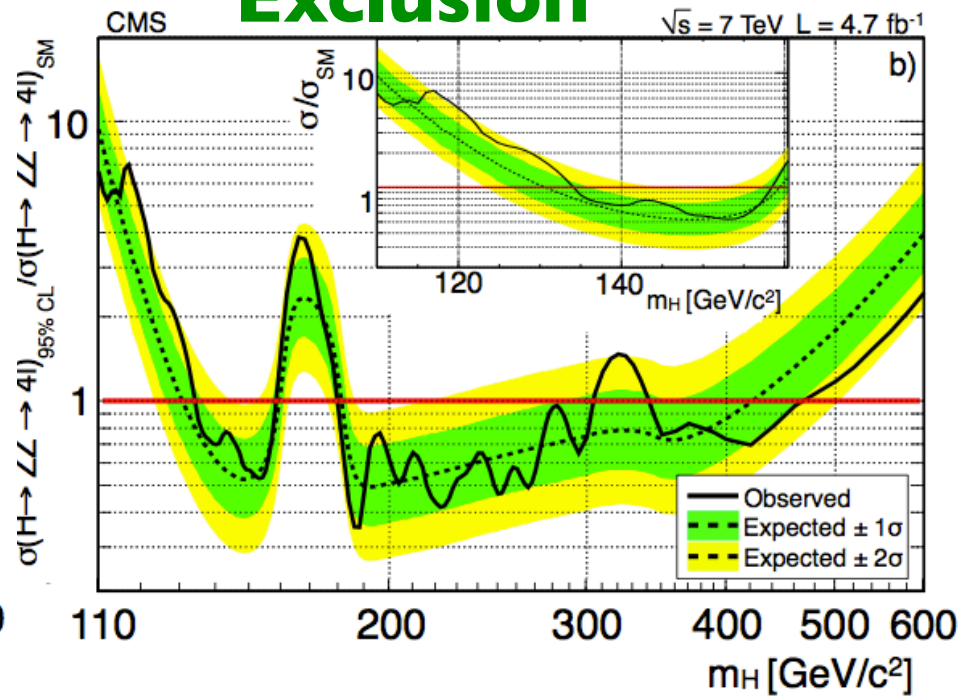
H→ZZ→ 4l: results

arXiv:1202.1997
Published by PRL

P-values



Exclusion



- Largest excess observed at 119.5 GeV
  - local significance 2.5 σ
  - global significance 1.0 σ in the full mass range, 1.6 in the mass range 100-160 GeV

SM scalar boson excluded at 95% CL for M_H in [134-158], [180-305] and [340-465] GeV

Higgs→WW*: the draught-horse

search-range: [110 - 600] GeV
max sensitivity: 160 GeV
WW→2l2ν: high BR, clean signature, no mass peak

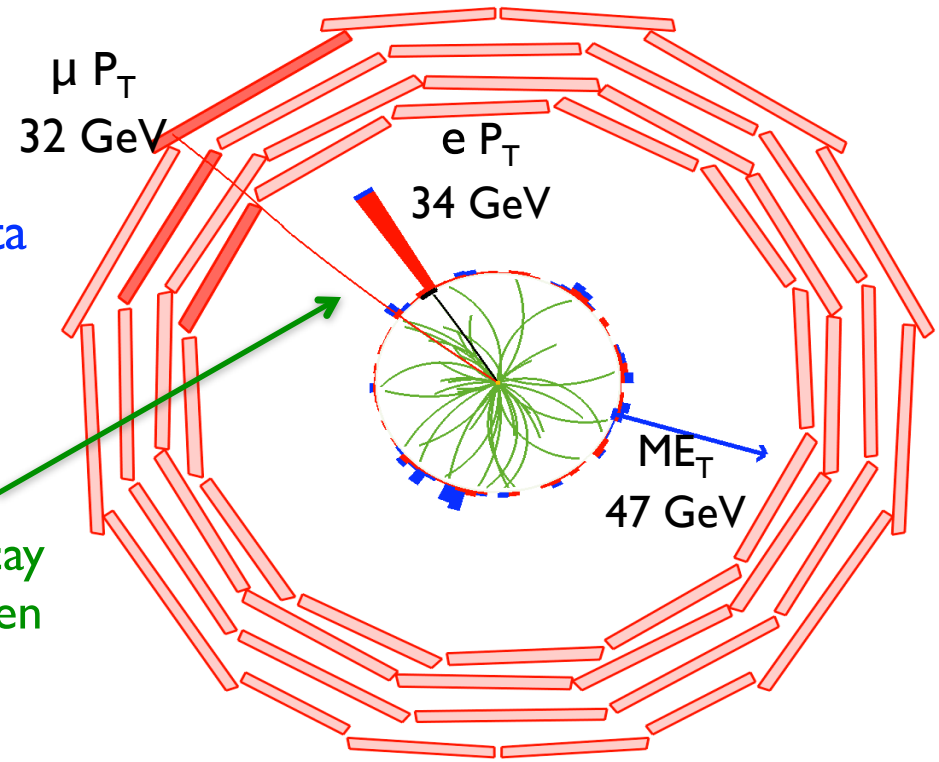
H→WW→ 2l2ν: strategy

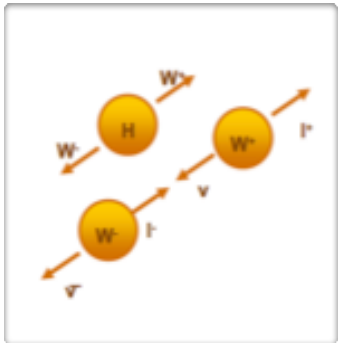
- Most sensitive channel around $2xM_W$

(125 <~ M_H <~ 200 GeV)

- **No narrow mass peak (mass resolution ~20%)**
- Two high p_T isolated leptons + MET
- Main backgrounds
  - WW (irreducible)
  - Z+jets, WZ, ZZ, tt, W + jets
- BKG estimation crucial
  - Main BG estimated from data

H->WW->eμνν candidate

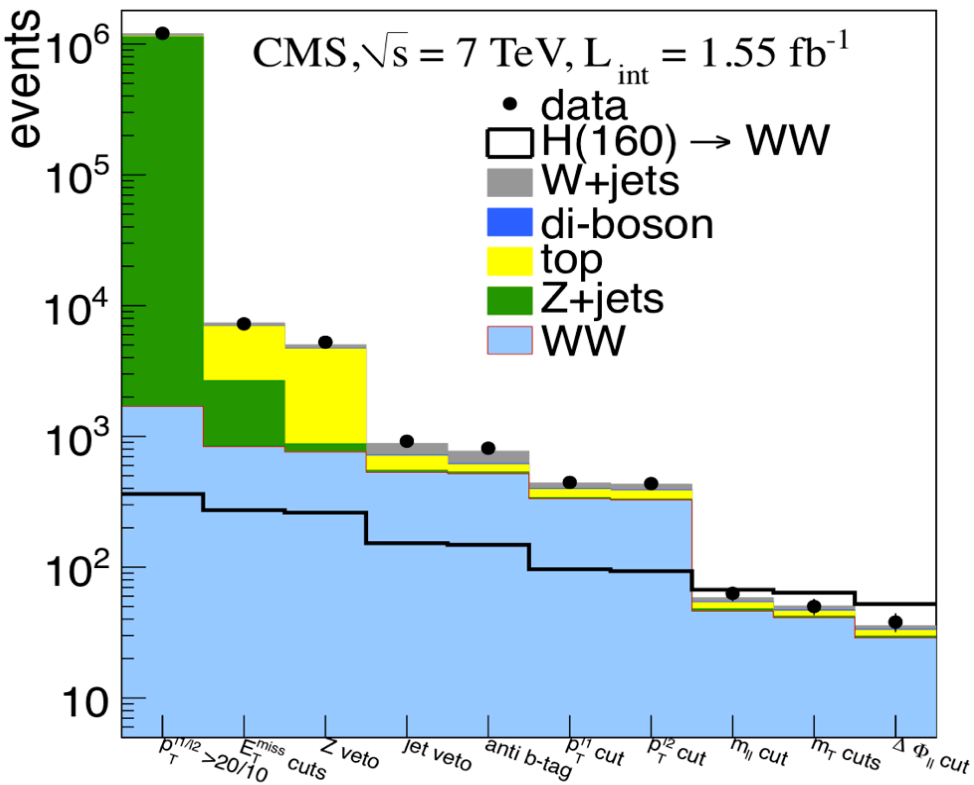




Scalar H boson +
V-A structure of W decay
favors small Δφ between
leptons

H→WW→ 2l2ν: selection

- **start with DY-dominated sample**, reduced by MET cut and Z-veto
- **top dominated sample** rejected with jet veto (=0, 1 or peculiar VBF 2 jets)
- b-tagging veto further reduces it
- this is **WW dominated sample**
- apply m_H dependent cuts: kinematics strongly depends on the Higgs mass
- $\Delta\phi_{ll}$, p_T^1, p_T^2, m_{ll}, m_T cuts
- ⇒ background composition depends on m_H



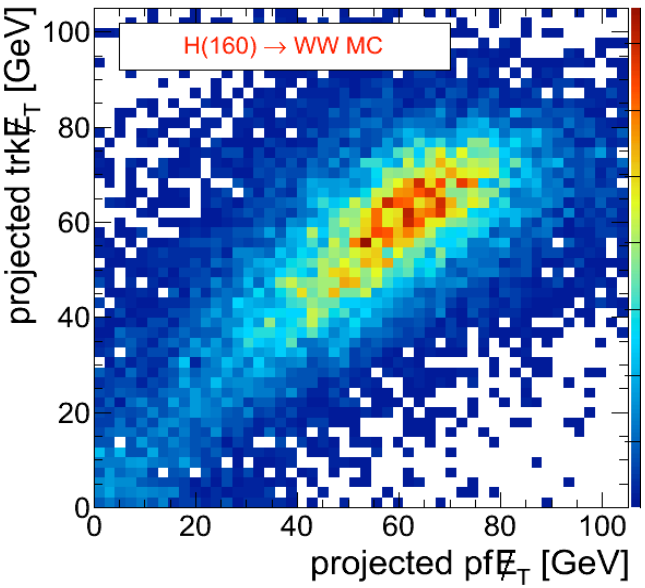
expected (data-driven) vs observed (eg. 0-jet, m_H=160 GeV)

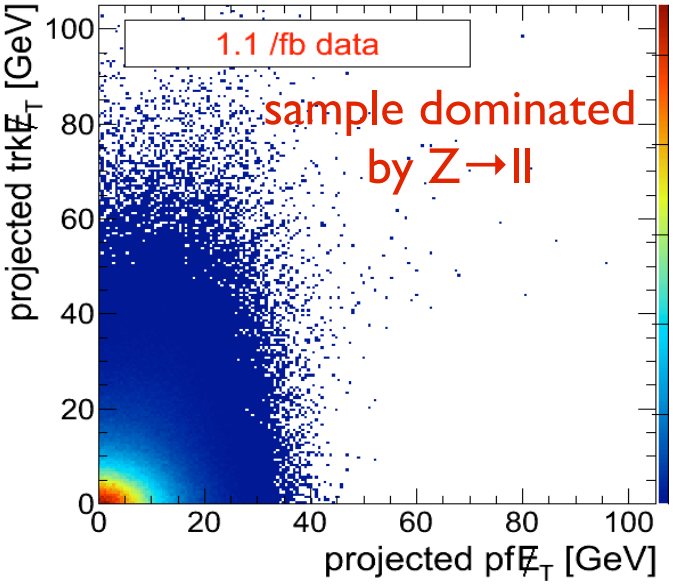
The pileup challenge

- Pileup worsen E_T^{miss} resolution
- Two estimators considered:
  - particle-flow ETmiss: charged + neutral component
  - track ETmiss: only tracks consistent with nominal primary vertex ⇒ reduced PU dependency
  - take the min of the two
- Project along ll-system (Z→ττ rejection):

$$\mathrm{proj}E_T^{\mathrm{miss}} = \begin{cases} E_T^{\mathrm{miss}} & \text{if } \Delta\phi_{min} \geq \frac{\pi}{2}, \\ E_T^{\mathrm{miss}} \sin(\Delta\phi_{min}) & \text{if } \Delta\phi_{min} < \frac{\pi}{2} \end{cases}$$

$$\Delta\phi_{min} = min(\Delta\phi(\ell_1, E_T^{\mathrm{miss}}), \Delta\phi(\ell_2, E_T^{\mathrm{miss}}))$$

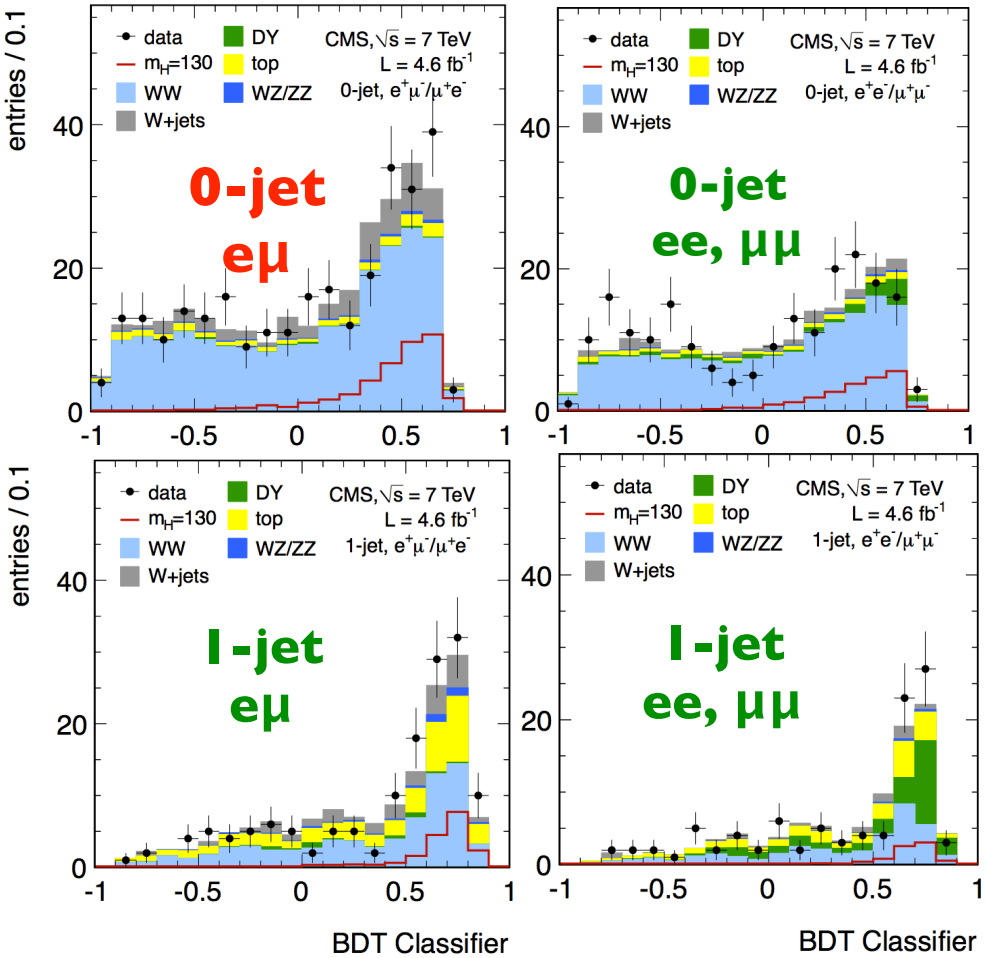




Shape analysis for 0 and 1 jet bins

- BDT trained at different masses, input variables:
  - $P_{tleptons}$, M_{ll}, $\Delta\varphi_{ll}$, ΔR_{ll}, M_t (for dilepton system and each lepton)
- Overall uncertainties:
  - signal efficiency ~20%
  - background ~15%
  - shape systematics included
- Use a fit to BDT output distribution for CLs estimation
- Many cross checks made, using Matrix Element method, using fit with only M_{ll} variable
- VBF (2jet bin): cut&count analysis

M_H=130 GeV



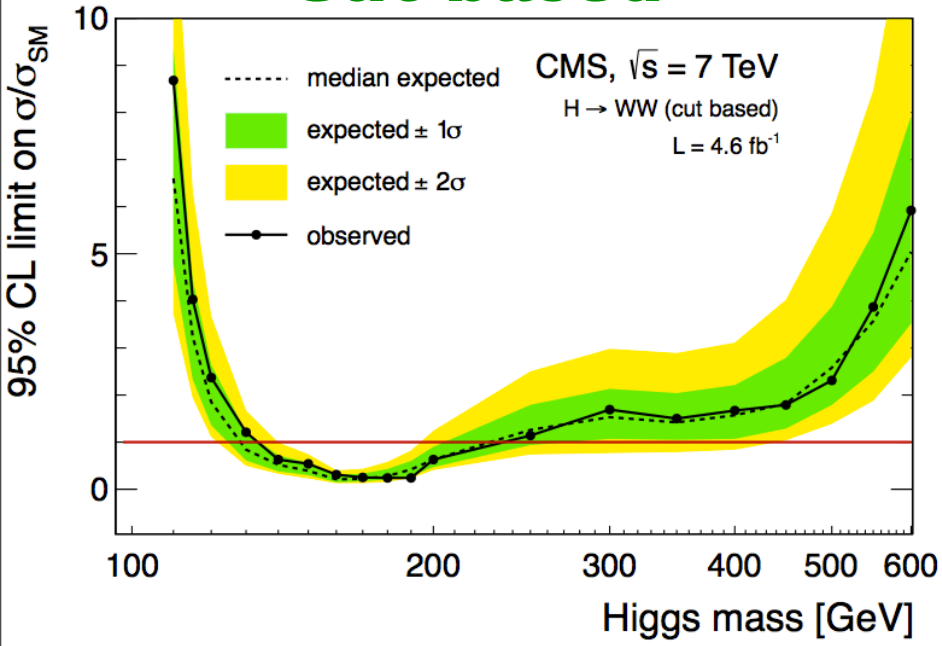
Most sensitive channel is eμ in 0-jet bin higher s/b and smaller systematic errors

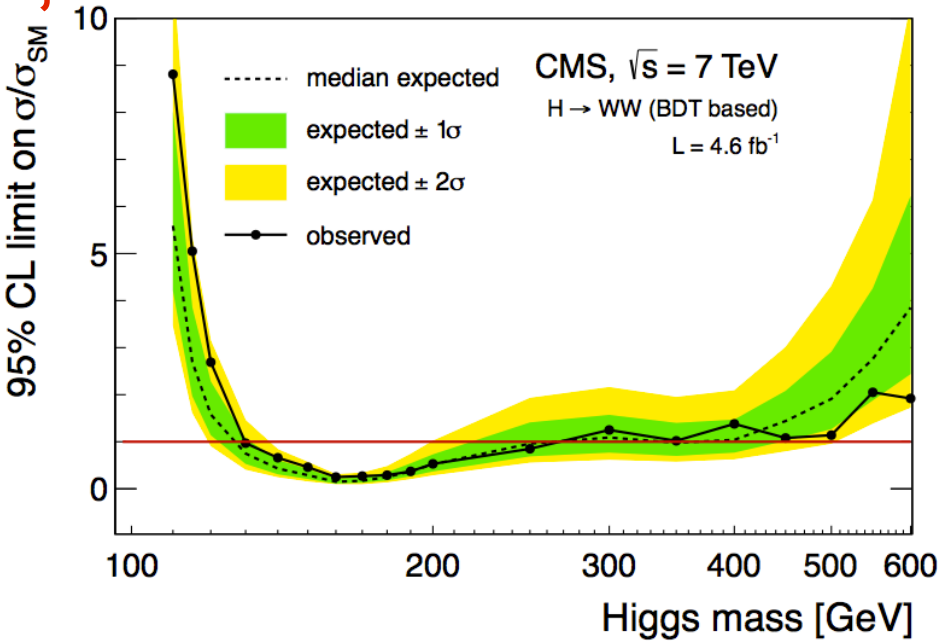
H→WW→ 2l2ν: results

- No significant excess in the full mass range

arXiv:1202.1489
Published by PLB

Cut based 0+1+2 jets **MVA based**



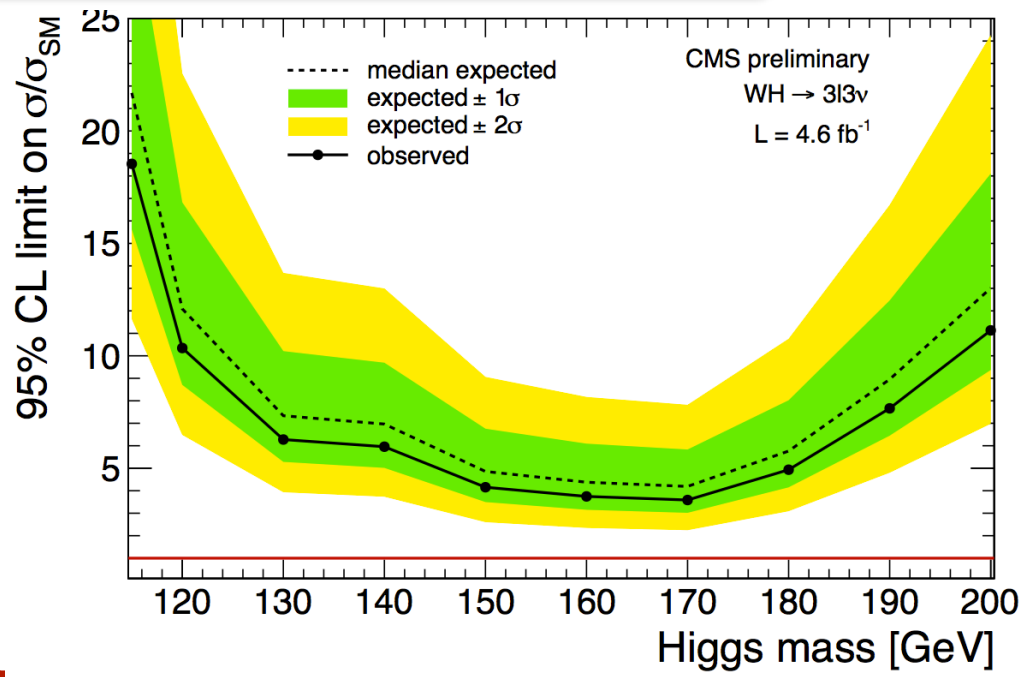


- Multivariate analysis more sensitive, especially at low mass
  - 95% C.L. expected exclusion for M_H in [127-270] GeV
  - 95% C.L. observed exclusion for M_H in [129-270] GeV
- Slight excess at low mass in shape analysis

WH→WWW→ 3l3ν

- Similar to WW analysis
- Cut and count analysis with mass independent selection
- Main backgrounds estimated from data

stage	WH (120) $H \to \tau\tau$	WH (120) $H \to WW$	data	all bkg.	WZ $\to 3\ell\nu$	ZZ $\to 4\ell$	top+Z/γ^*
3-lepton preselection	2.1 ± 0.0	3.5 ± 0.1	950	968.3 ± 11.9	482.9 ± 1.8	78.4 ± 0.9	348.0 ± 9.7
min-MET > 40 GeV	1.0 ± 0.0	1.8 ± 0.1	244	270.5 ± 4.4	208.2 ± 1.1	7.9 ± 0.3	54.5 ± 4.3
Z removal	0.4 ± 0.0	1.0 ± 0.1	40	47.9 ± 3.1	15.9 ± 0.4	0.7 ± 0.1	31.3 ± 3.1
top veto	0.1 ± 0.0	0.6 ± 0.1	12	14.2 ± 1.3	8.8 ± 0.4	0.4 ± 0.1	4.9 ± 1.3
$\Delta R_{\ell^+\ell^-}$ & $m_{\ell\ell}$	0.1 ± 0.0	0.5 ± 0.1	7	8.4 ± 0.9	5.7 ± 0.2	0.3 ± 0.1	2.6 ± 0.9

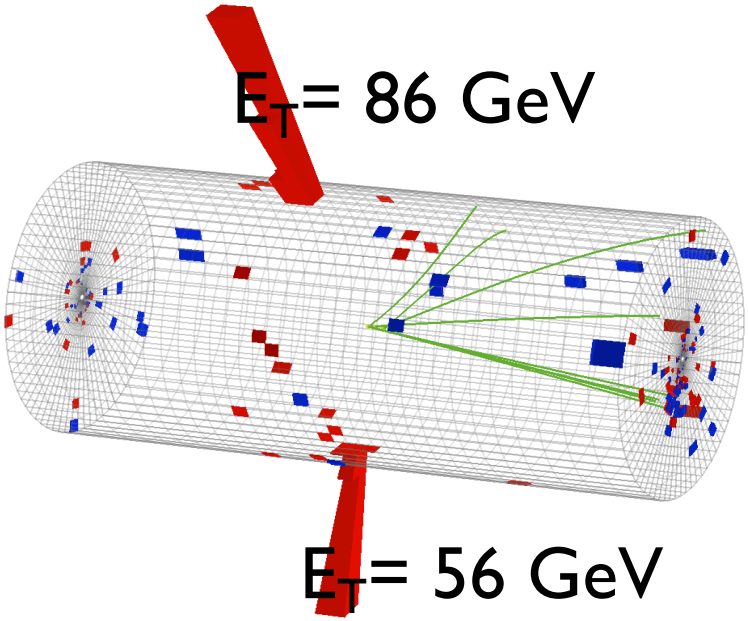


CMS document HIG-11-034

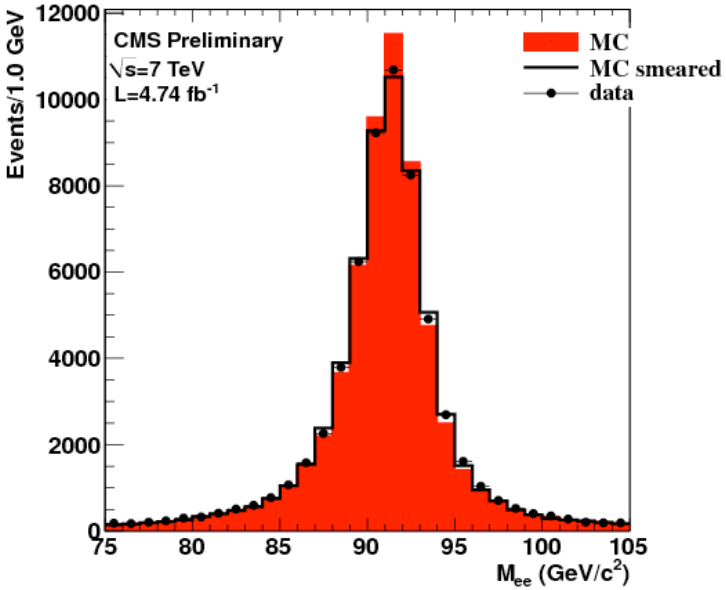
Higgs→γγ: the low mass driver

search-range: [110 - 150] GeV
max sensitivity: < 125 GeV
narrow peak, large background

H→γγ: basics



- Small BR: ~2x10^{-3}
- Two isolated high E_t photons
- VBF channel has two additional jets from outgoing quarks
- **Narrow mass peak**
  - very good mass resolution 1-2%

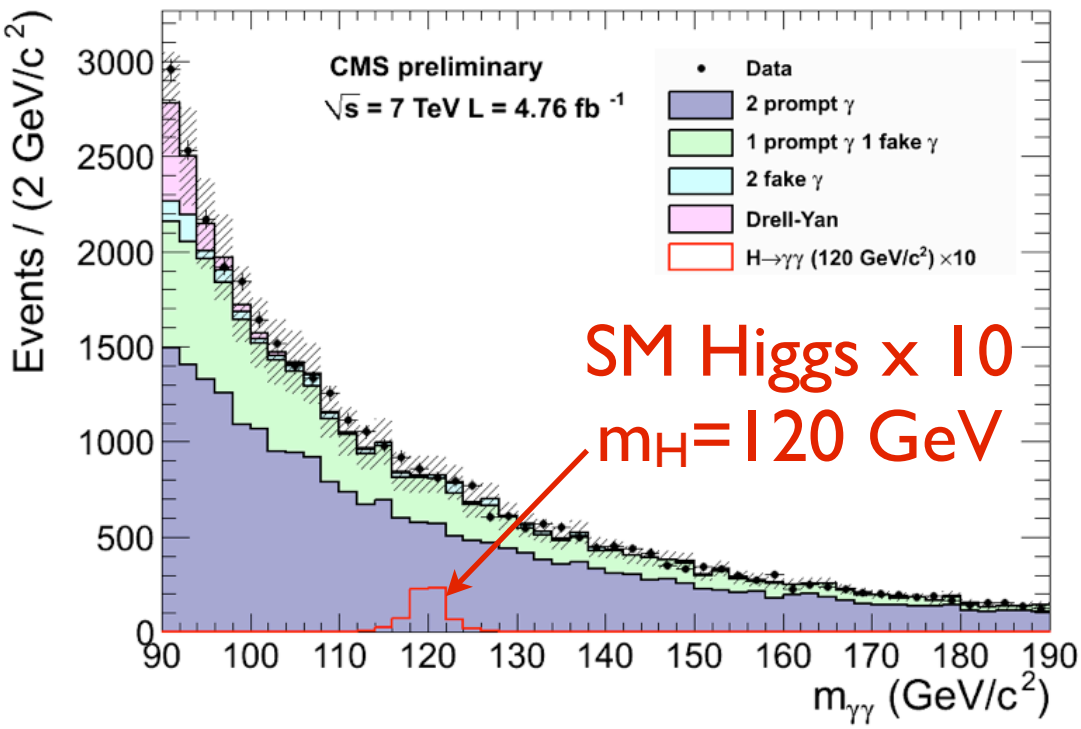


Z→ee with both e in barrel, best resolution category

- Signature: small mass peak over large smoothly decreasing background
  - Irreducible: 2γ QCD production
  - Reducible: γ+jet with 1 additional fake photon, DY with electrons faking photons
- Studied mass range: 110-150 GeV

H→γγ: analysis strategy

- Select two photons with $p_T > m_{\gamma\gamma}/3(4)$ with appropriate selection on shower shape and isolation
- Multivariate energy corrections for local and global electromagnetic cluster containment (resolution and energy scale corrections from Z → ee)
- Primary vertex selection uses track recoil against di-photon system with conversion pointing where available (correct vertex in ~ 83% of events)

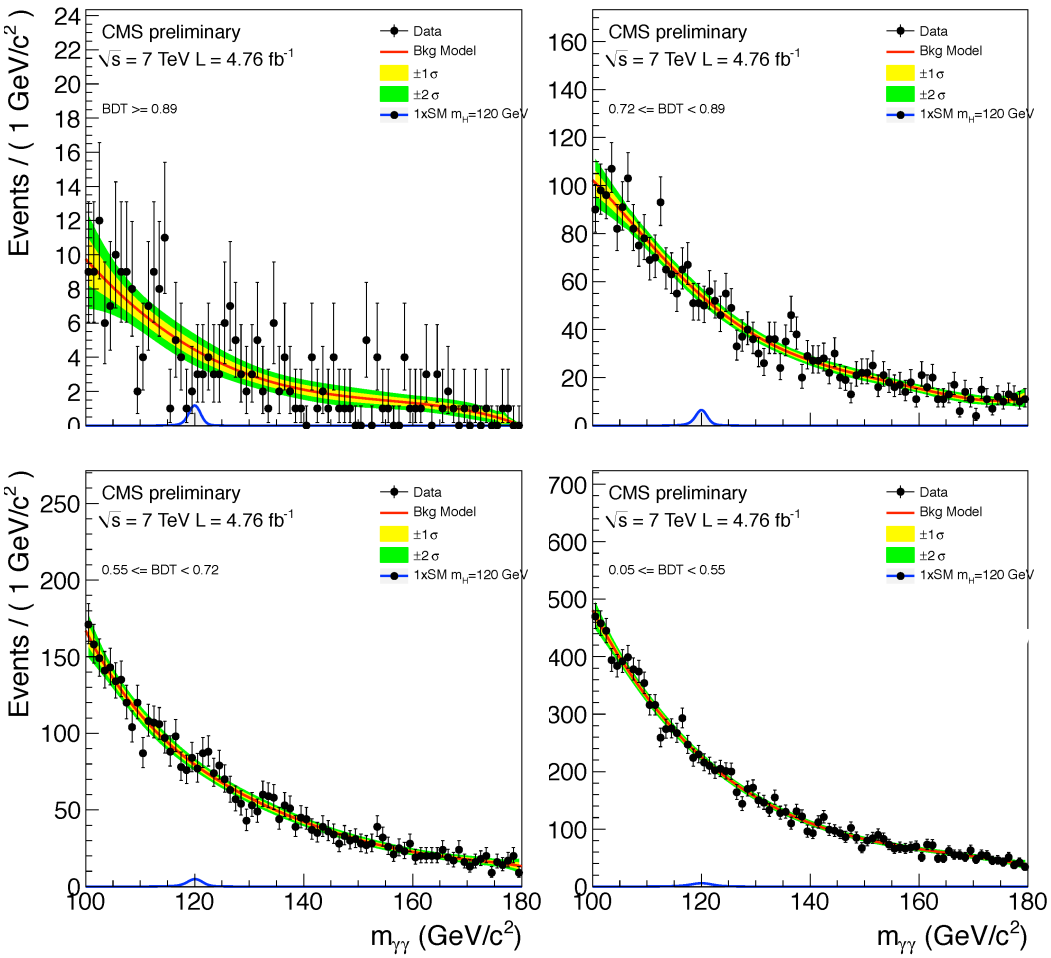


arXiv:1202.1487
Published by PLB

H→γγ: inclusive analysis / VBF

- 4 non-VBF event classes split based on the diphoton BDT classifier output
- BG is estimated by fitting to a polynomial in the full mass range (3rd to 5th order)
  - Possible BG bias is always less than 20% of the statistical error
  - Different BG estimation in cross check analysis gives consistent results
- VBF channel has lower yield but larger purity

Events passing VBF selection removed

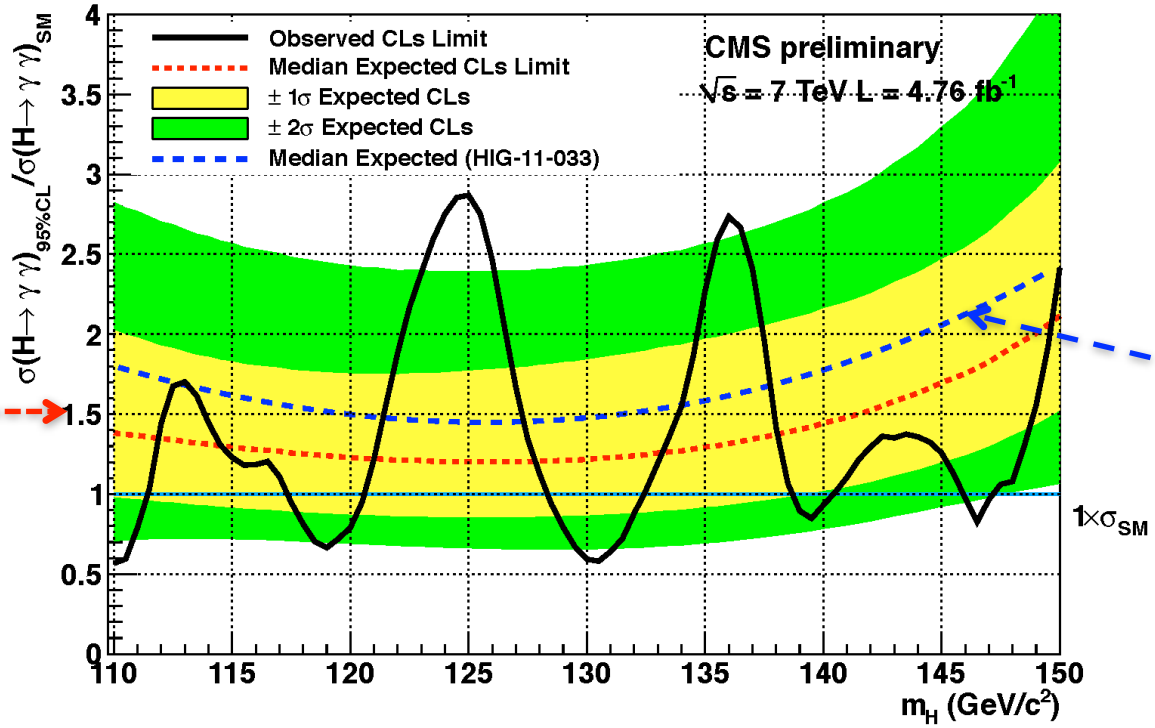


m_H=120 GeV	class 0	class 1	class 2	class 3	Dijet class
Total signal expected events	3.4	19.3	18.7	33.0	2.8
Data (events/GeV)	4.5	55.1	81.3	229.1	2.1
resolution FWHM/2.35 (%)	0.9	0.9	1.2	1.7	1.1

H→γγ: exclusion limits

CMS document
HIG-12-001

Expected from MVA analysis Improvement ~20%

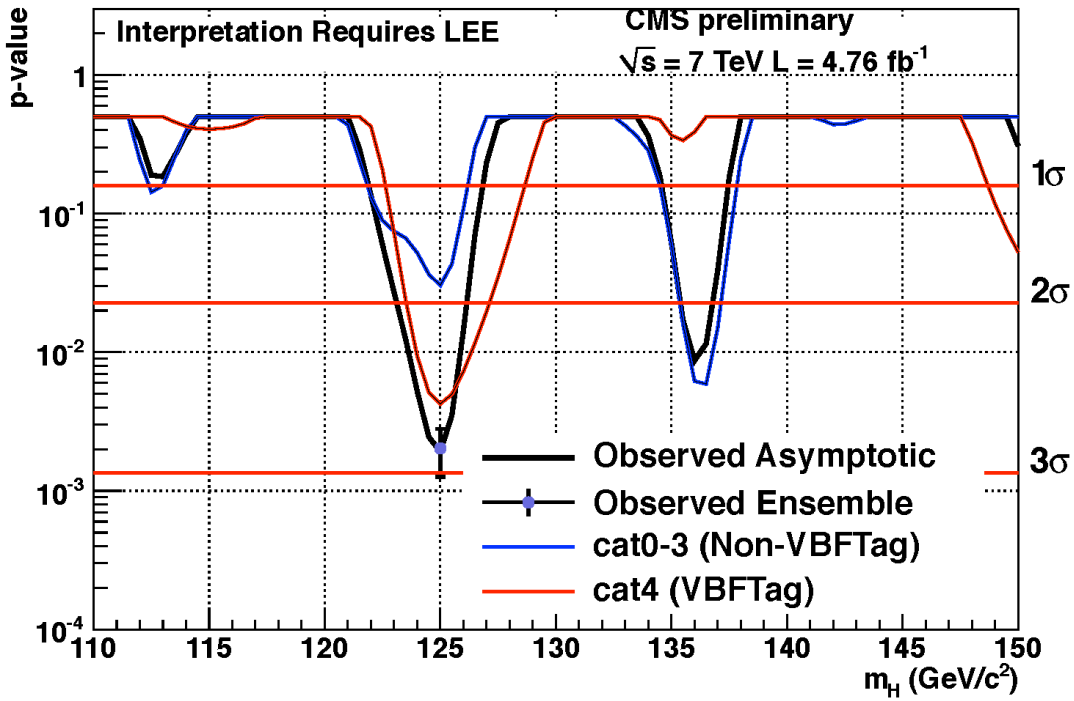


Expected from cut based analysis

- Expected 95% CL exclusion: 1.2-2 x SM
- Excluded at 95% CL:
  110.0-111.0, 117.5-120.5, 128.5-132.0, 139.0-140.0, 146.0-147.0 GeV
- Cut based analysis gives consistent results
  - arXiv:1202.1487, published by PLB
- Cross check MVA analysis also gives consistent results

H→γγ: p-values

- P-value: probability that a BG only fluctuation is more signal-like than observation



Similar results as cut based

- Largest excess around 125 GeV
  - Local significance 2.9 σ
  - Global significance 1.6 σ
- Look elsewhere effect (LEE) estimated in the mass range 110-150 GeV with toy experiments

other low mass channels

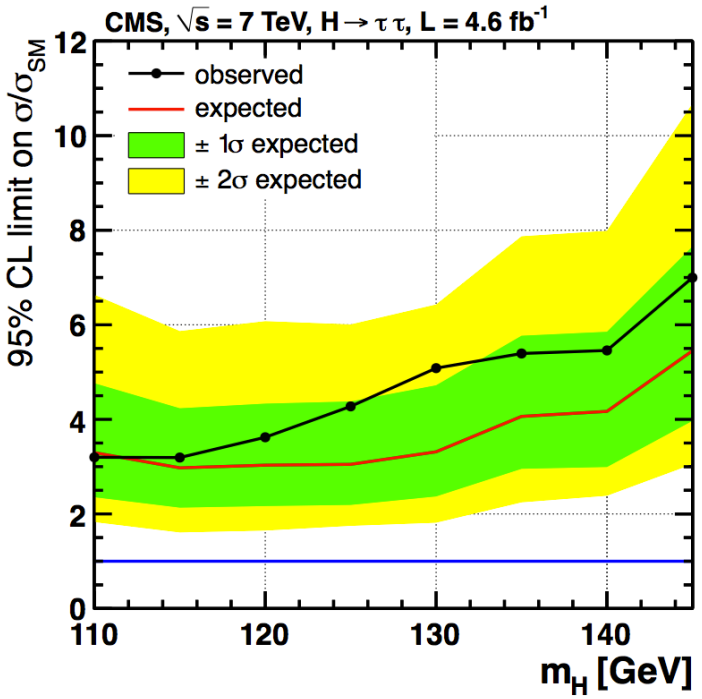
$H \rightarrow \tau\tau$

$H \rightarrow bb$

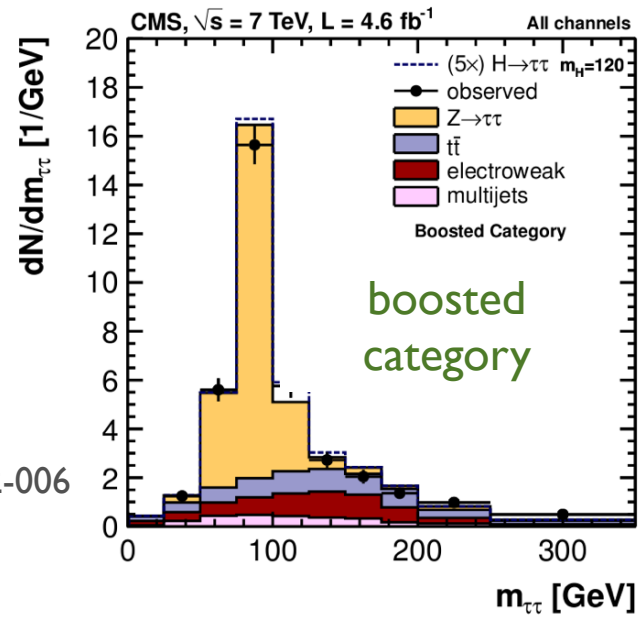
H→ττ

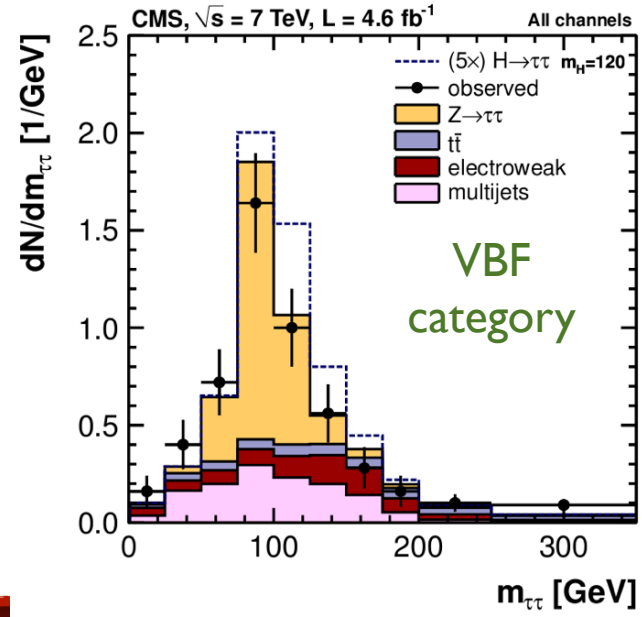
- **No narrow mass peak** (σ_M ~20%)
- Three different sub-channels for SM:
  - VBF production, two additional forward jets
  - Boosted: one jet with P_t>150 GeV
  - gg-fusion 0 or 1 additional jets

- H->ττ->μμ also studied CMS document HIG-12-007
- WH->lττ->eμτh, μμτh also studied CMS document HIG-12-006



arXiv:1202.3697
Published by PLB

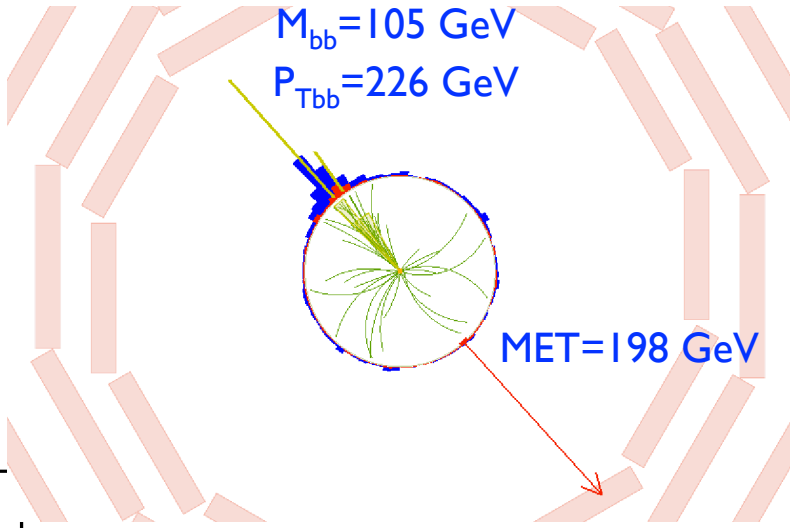


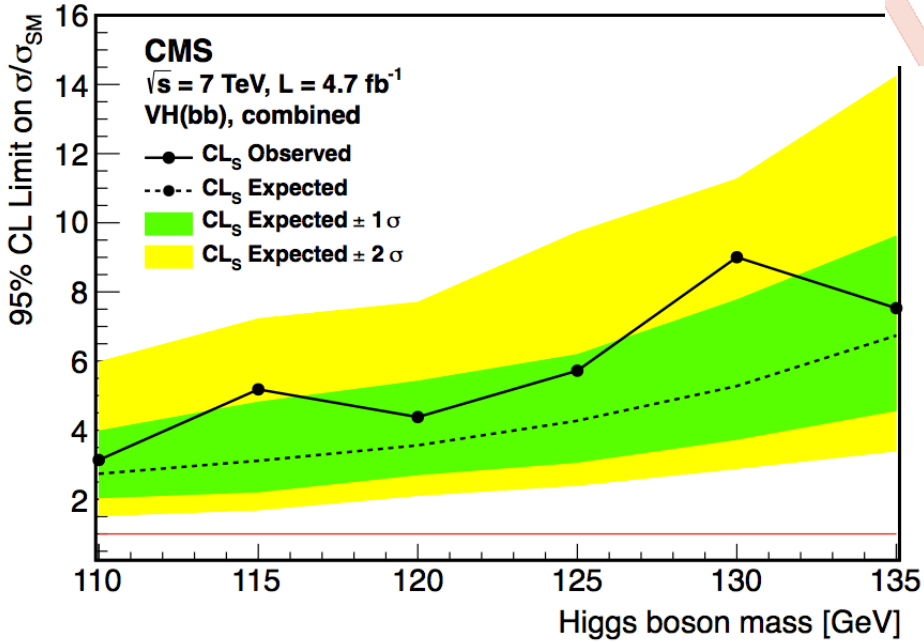


H→bb

- BKG too large, needs additional tag
- Exploit VH associated production with W and Z decaying leptonically
  - W->e,μ
  - Z->ee,μμ, νν
- Require boosted bb system
- Mass resolution ~10%

ZH→ννbb candidate





Excluded at 95% CL
3-9xSM

arXiv:1202.4083
Published by JHEP

CMS Higgs combination

search-range: [110 - 600] GeV

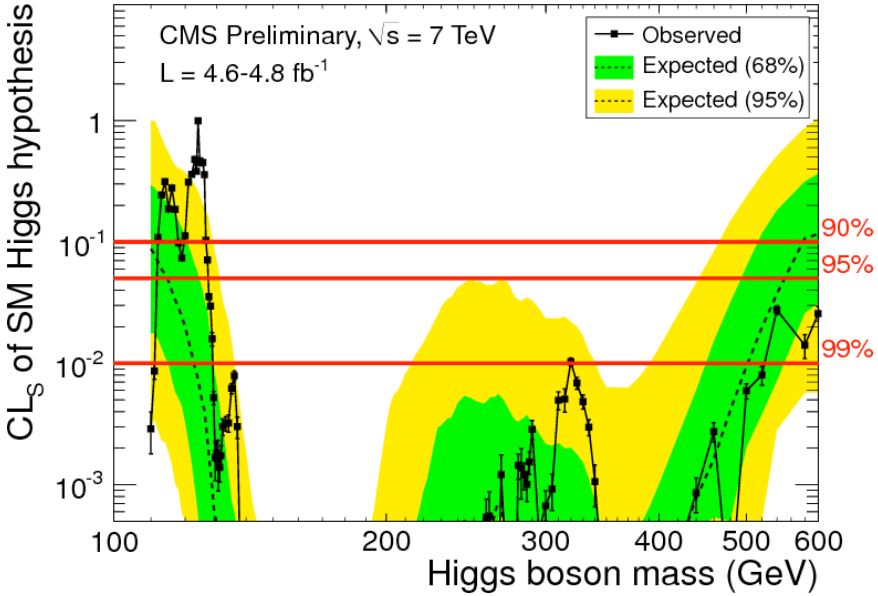
all the modes analyzed by CMS with 4.7 fb^{-1} are included

Combination strategy

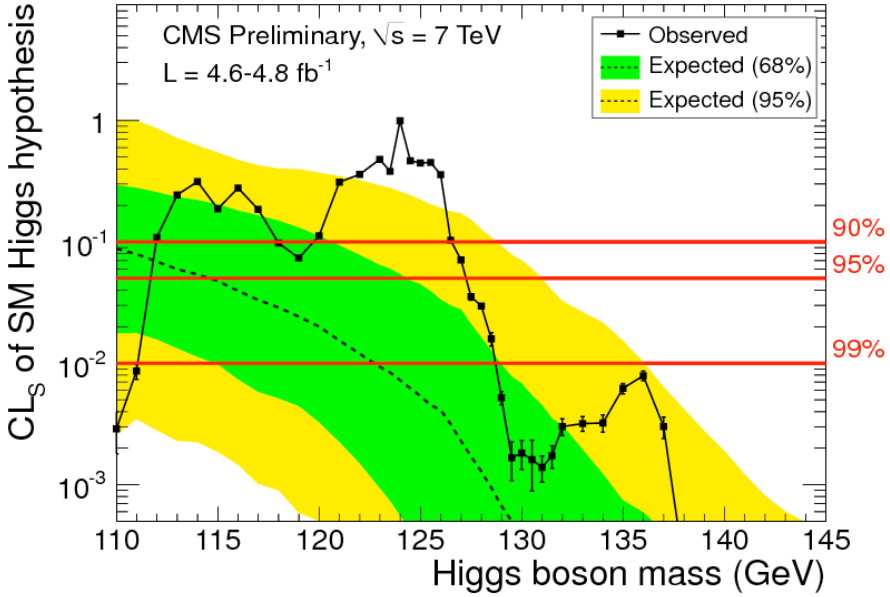
- SM cross sections and branching ratios are assumed with their theoretical uncertainties and an overall signal strength multiplier is fit
- Method for CL calculation is LHC-type CLs
  - Frequentist CLs with profiled likelihood test statistics and log-normal treatment of nuisance parameters
  - ATL-PHYS-PUB/CMS NOTE 2011-11, 2011/005, (2011)
- Here we show the latest results that include the new analyses presented here
- Results based on 2011 data have been recently submitted for publication
  - arXiv:1202.1488, published by PLB

Combined exclusion C.L.

Full mass range



Low mass region

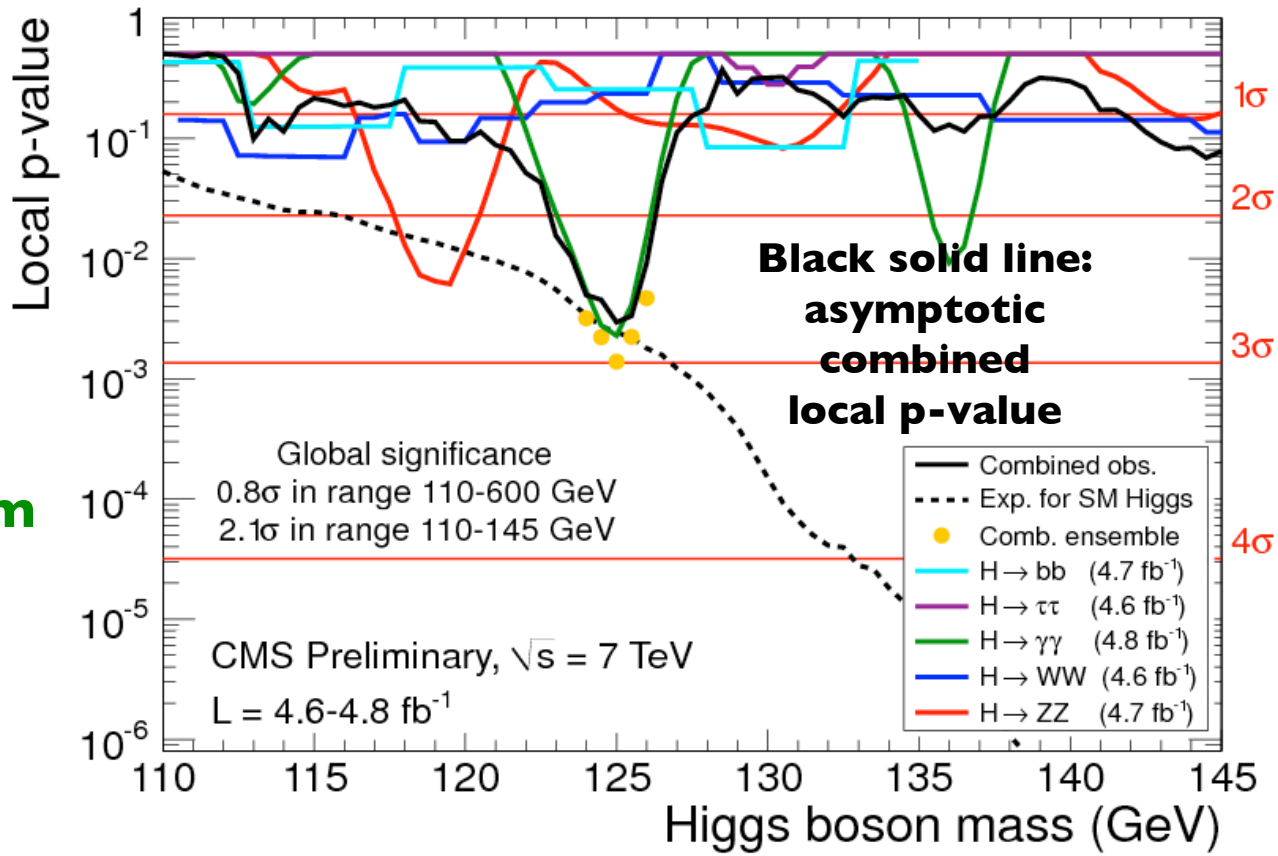


- Expected: 95% exclusion M_H in [114.5-543] GeV
- Observed: 95% exclusion M_H in [127.5-600] GeV
  99% exclusion M_H in [129-525] GeV
- 95% allowed mass range: 114.4-127.5 GeV
- Observed lower limit higher than expected because of excess in data at low mass

CMS document
HIG-12-008

Combined p-values

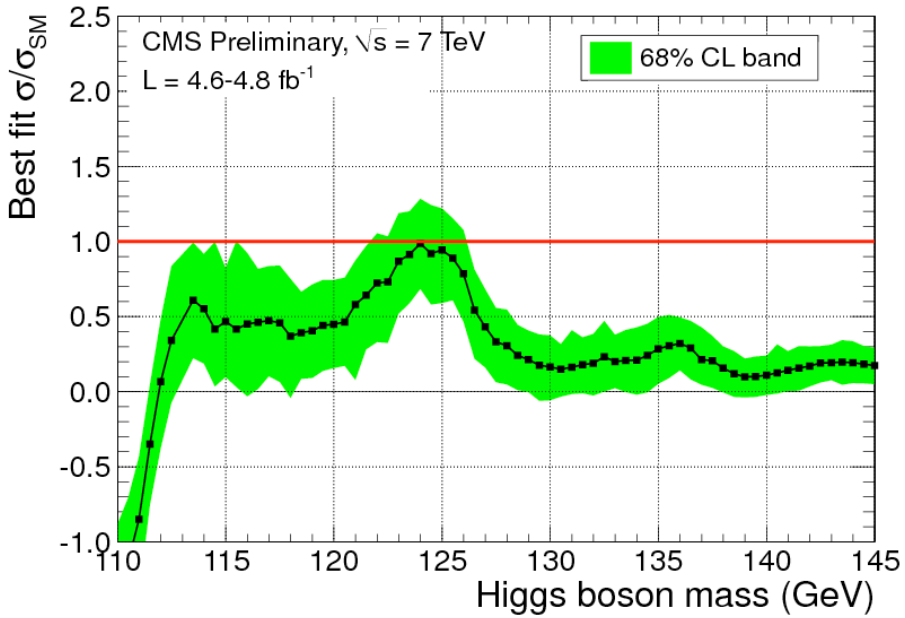
Low mass region zoom

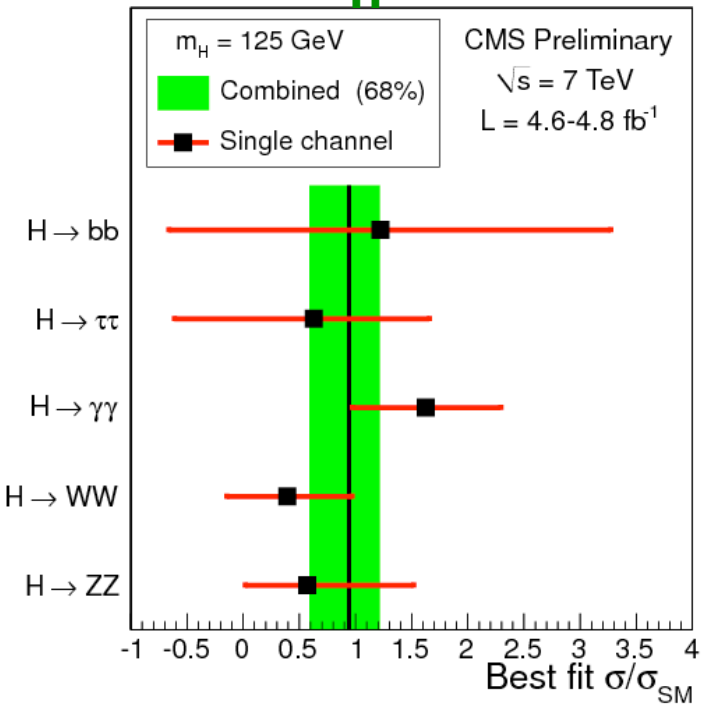


- Minimum p-value observed at 125 GeV with local significance: 2.8 σ
  - Similar significance expected from signal
- Estimated global significance:
  - 0.8 σ in [110-600] GeV, 2.1 σ in [110-145] GeV,

Fitted signal strength σ/σ_{SM}

Comparison of channels for M_H=125 GeV





- The fitted σ of the excess near 125 GeV is consistent with the SM Higgs boson expectation
- At low mass several channels show modest excess
  - At 125 GeV the mass-sensitive channels show an excess consistent with signal expectations
  - consistent with signal and bkg-only hypothesis: more data needed

Conclusions

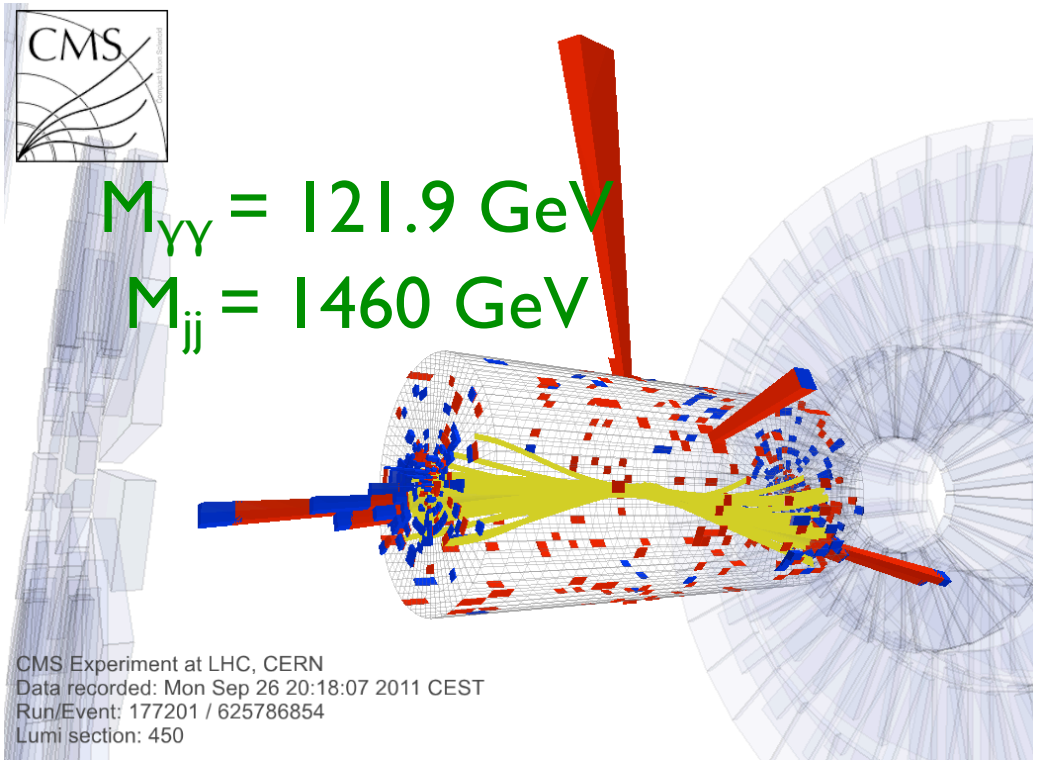
- SM scalar boson search in >10 independent channels with the full 2011 dataset
- SM Higgs excluded in M_H in [127.5-600] GeV at 95% CL
- Observe a modest excess around 125 GeV:
  - Local significance 2.8 σ, global significance 0.8 σ (in the full search range) and 2.1 σ (in 110-145 GeV)
  - The excess is consistent both with background fluctuation and also with a Higgs boson with mass about 125 GeV
  - if the Higgs exists, limited between 114.5 and 127.5 GeV

- In 2012 LHC is running at 8 TeV CM energy and high luminosity
  - should be able to discover or exclude the SM scalar boson down to the LEP limit

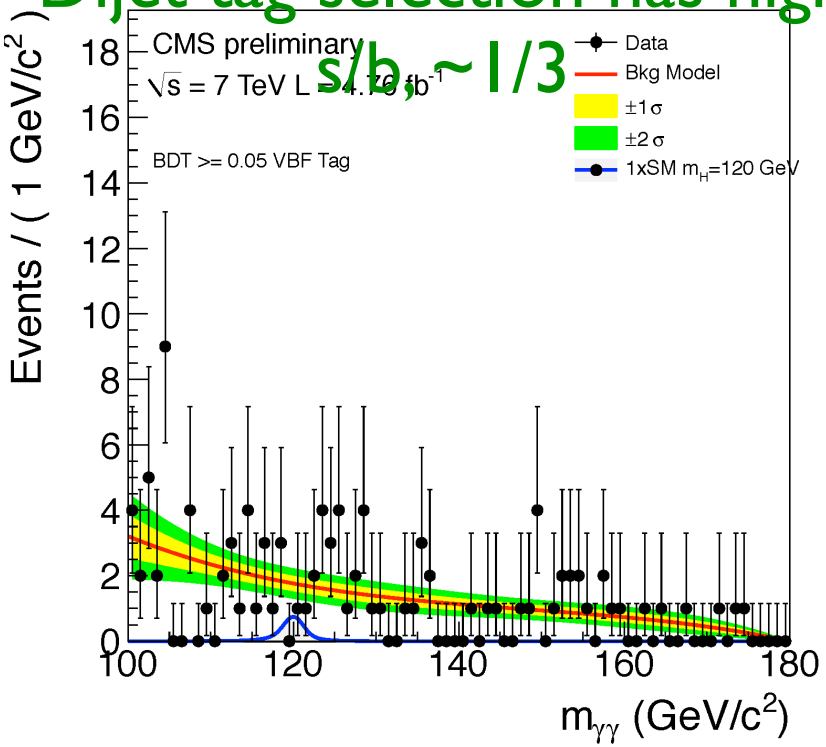
Backup

H→γγ VBF analysis

- Exclusive dijet tag improves sensitivity by ~10%
- Photon identification is the same
  - tighter lead photon E_t cut (E_t lead/$M_{\gamma\gamma}$ > 55/120)
- Dijet tag selection on dijet variables
  - exploits two additional VBF high p_T jets at large rapidity



Dijet tag selection has high s/b ~1/3

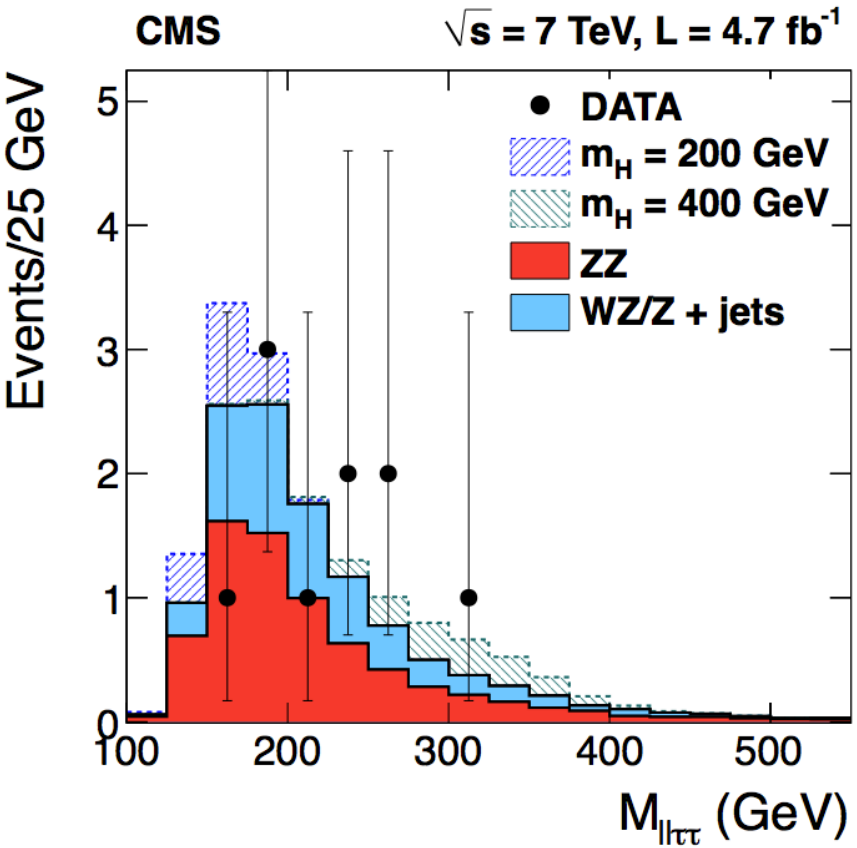


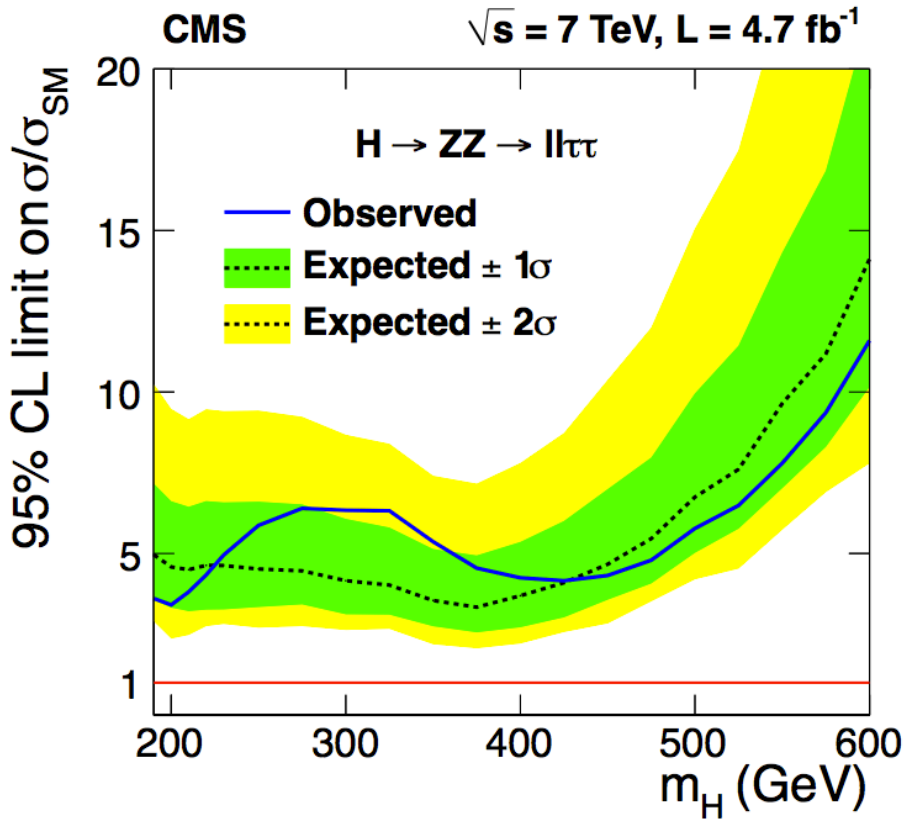
H→ZZ→2l2τ channel

arXiv:1202.3617
Published by JHEP

- All τ decays used
- Sensitivity about 4xSM
- Expect 10.2 events from BG
- Observe 10 events in data

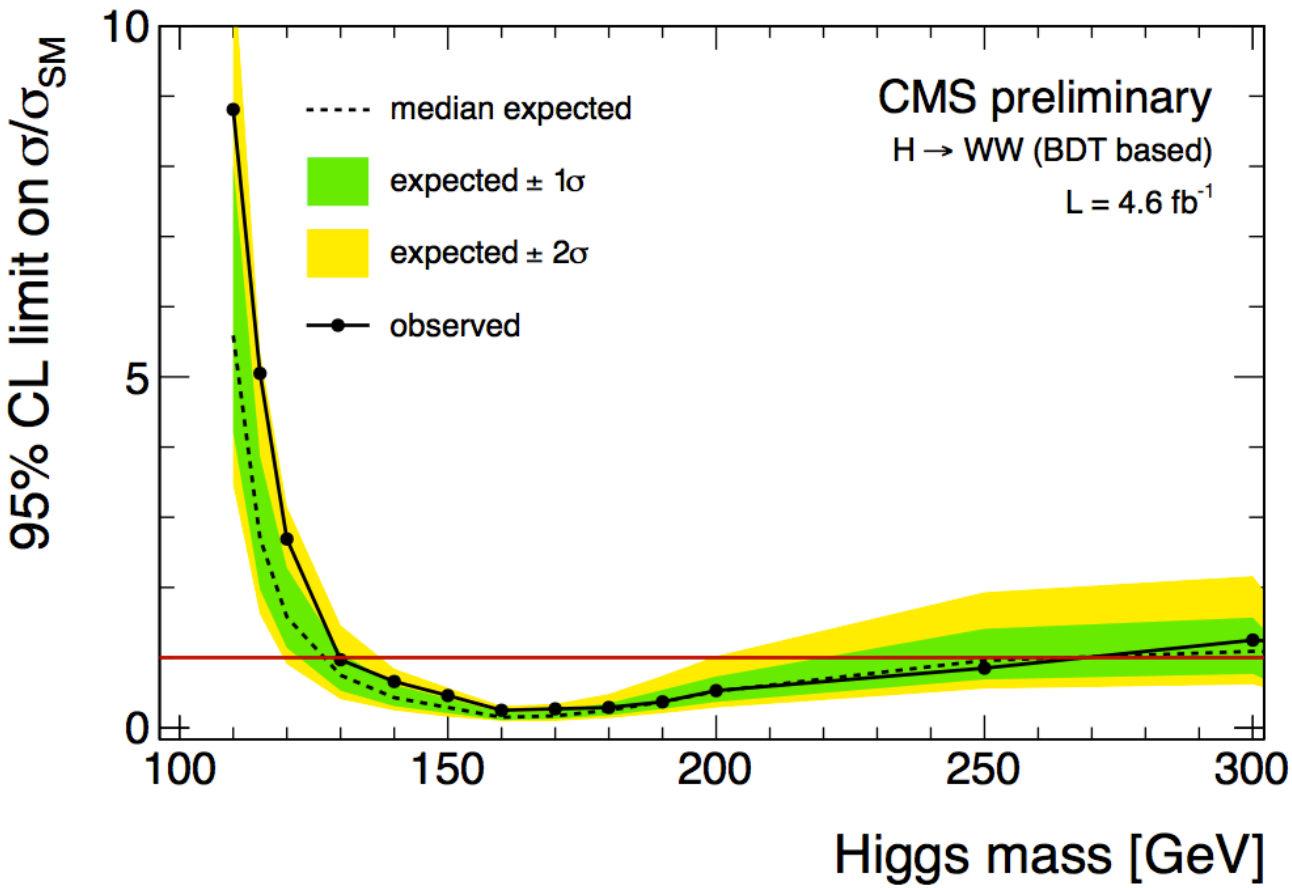
95% Exclusion





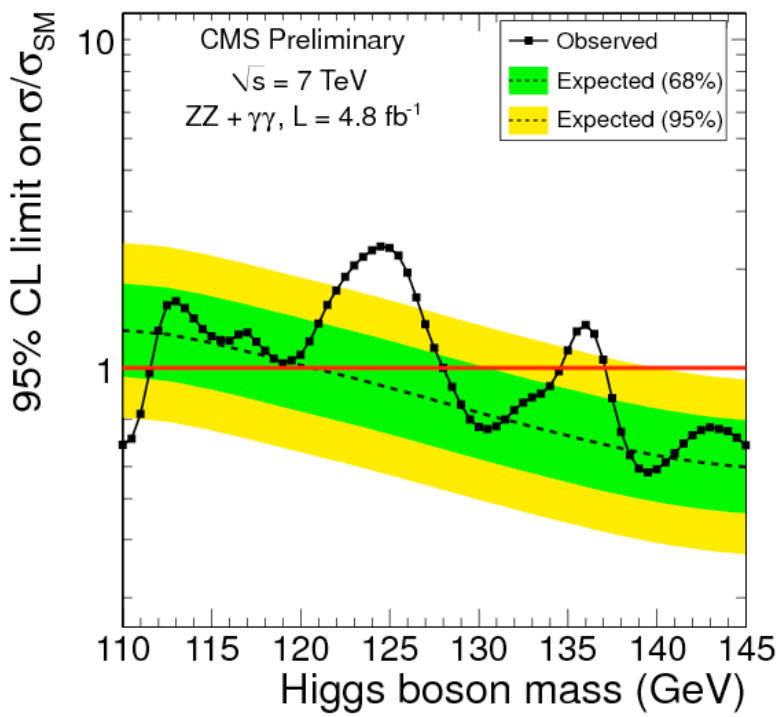
H→WW→2l2ν: exclusion 100-300 GeV

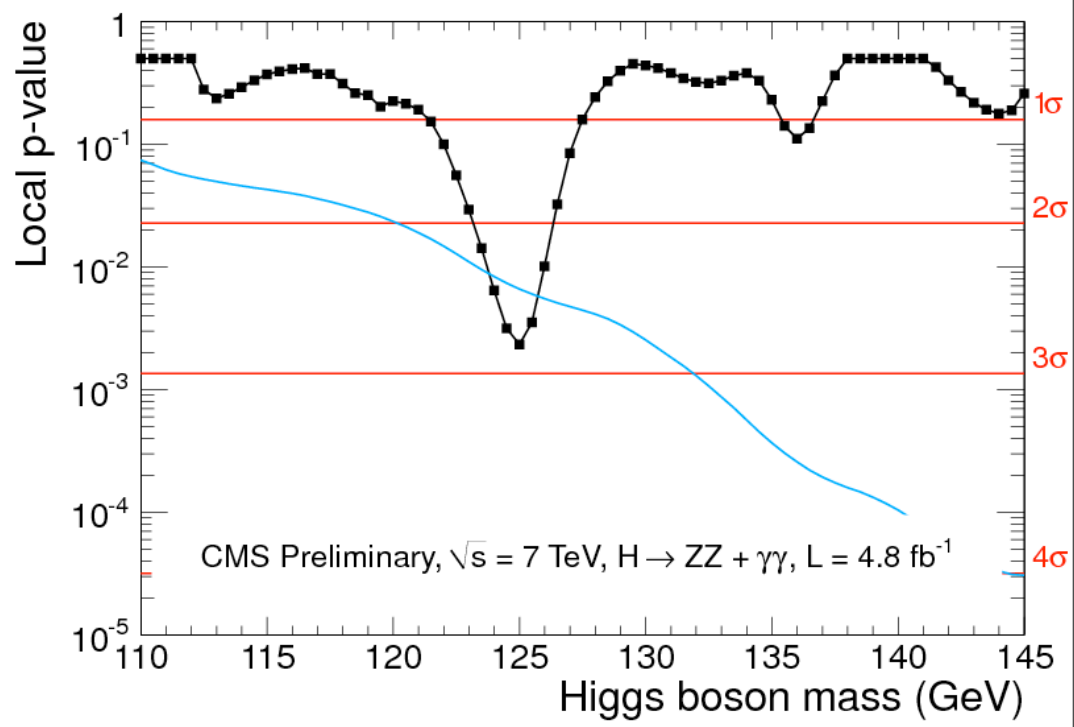
arXiv:1202.1489
Accepted by PLB



high-resolution channels combination(ZZ+γγ)

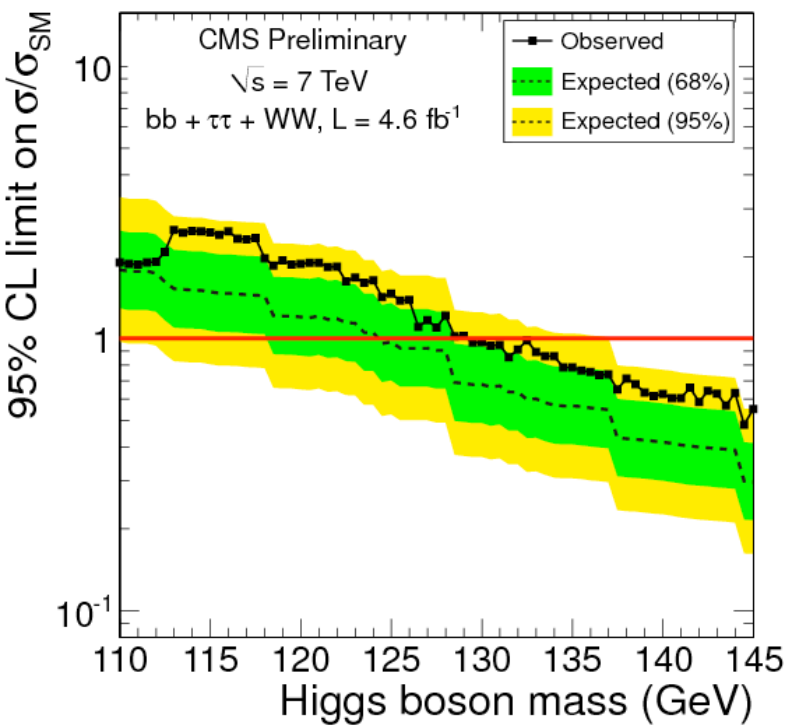
CMS document
HIG-12-008

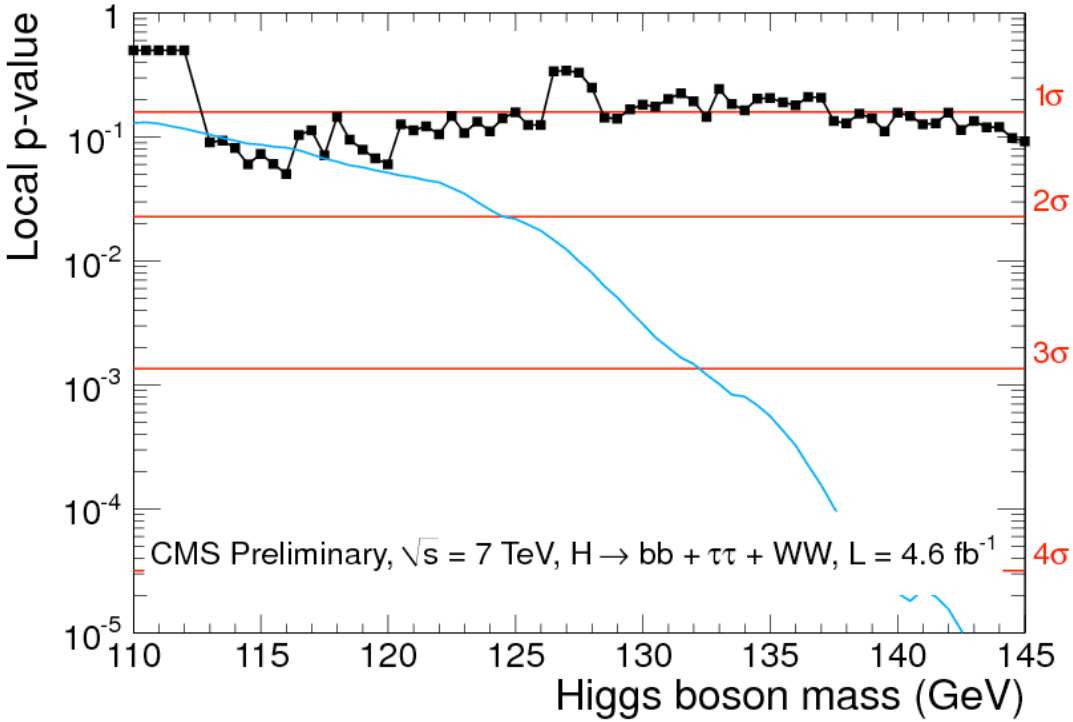




low-resolution channels combination(WW+bb+ττ)

CMS document
HIG-12-008

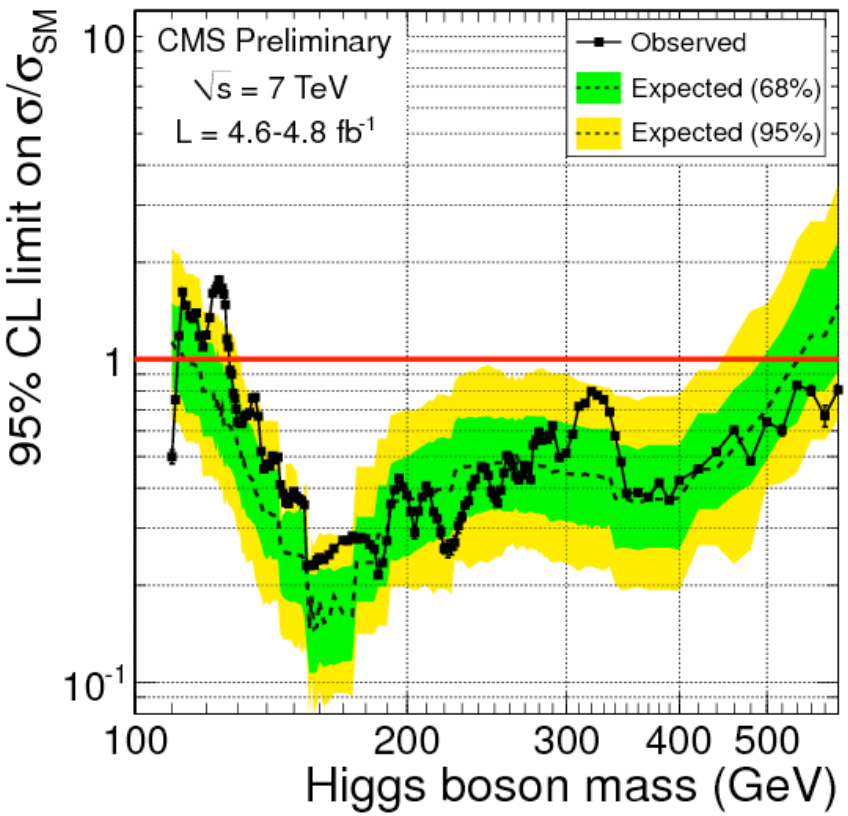




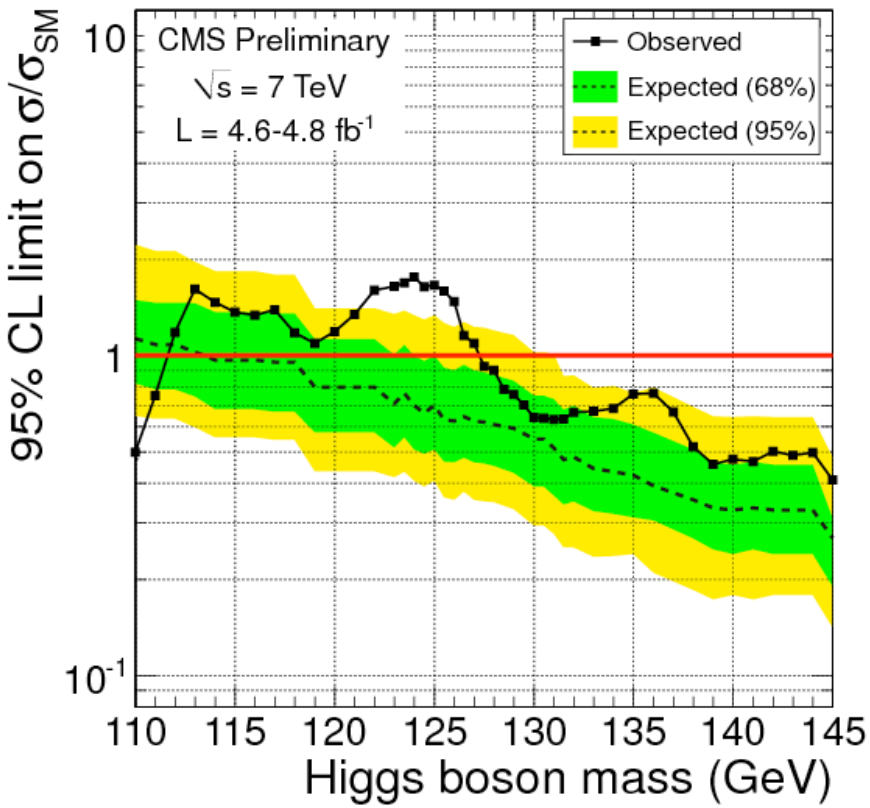
95 % CL exclusion

CMS document
HIG-12-008

Full mass range



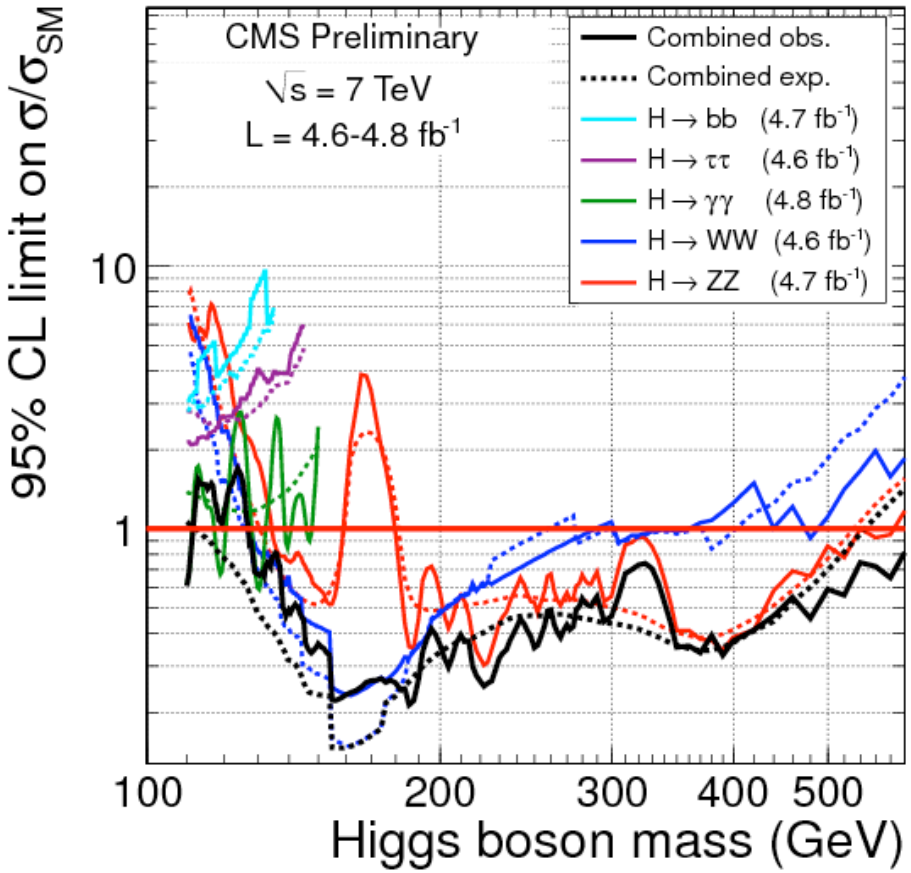
Low mass region



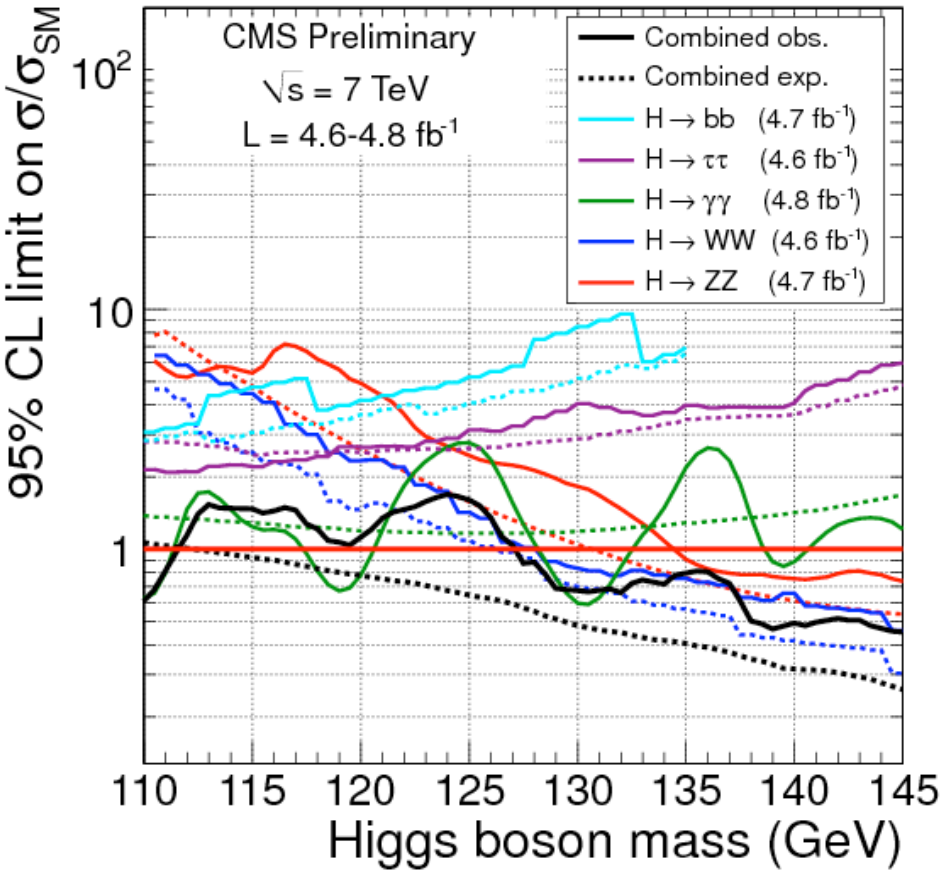
95 % CL exclusion for different channels

CMS document
HIG-12-008

Full mass range

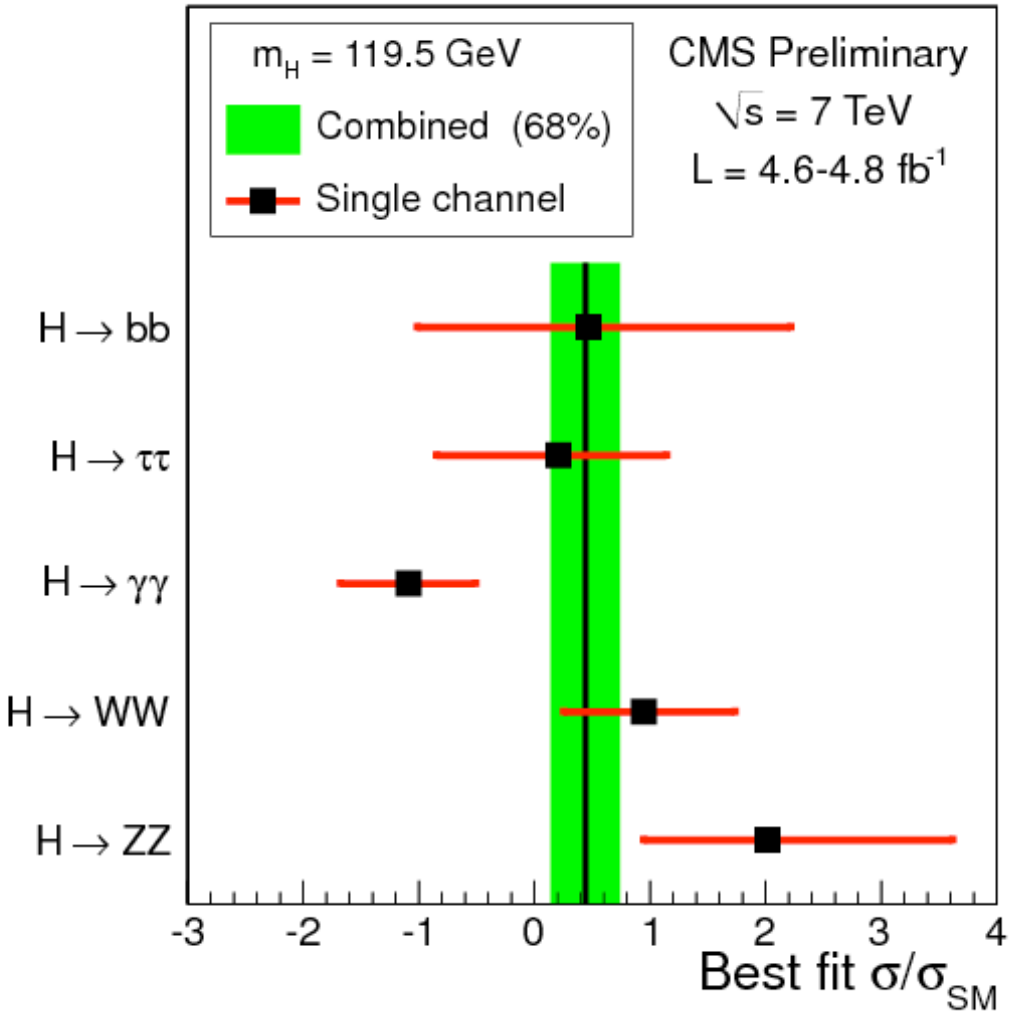


Low mass region



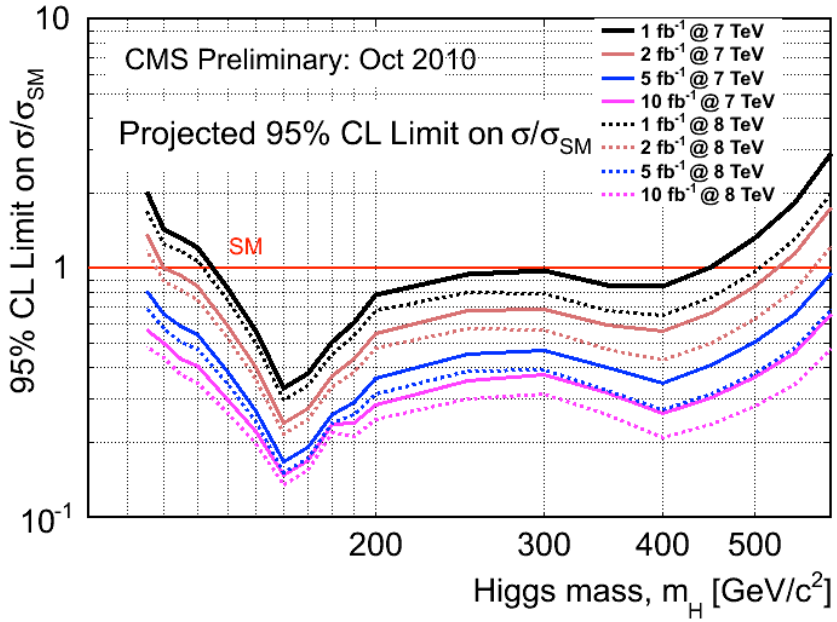
Fitted signal strengths at 119.5 GeV

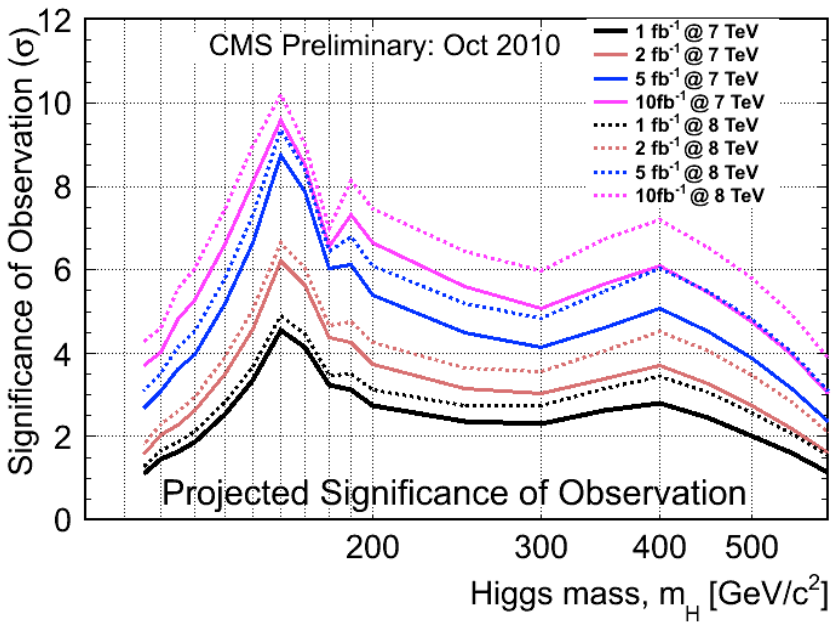
CMS document
HIG-12-008



Projections for 2012 run

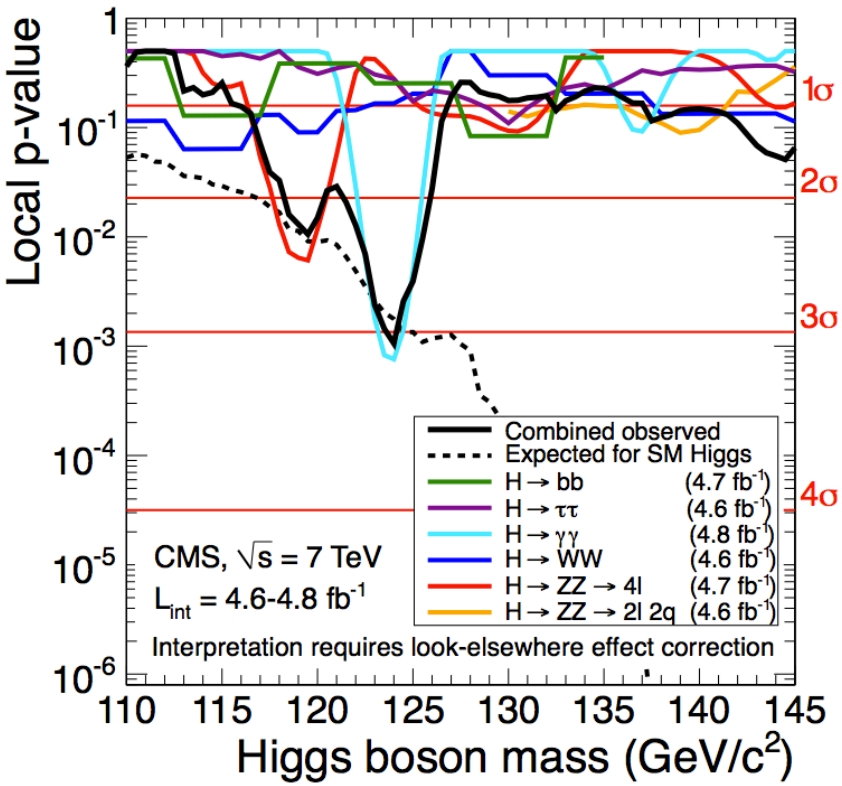
- 8 TeV vs 7 TeV is expected to increase the sensitivity by 10-20%

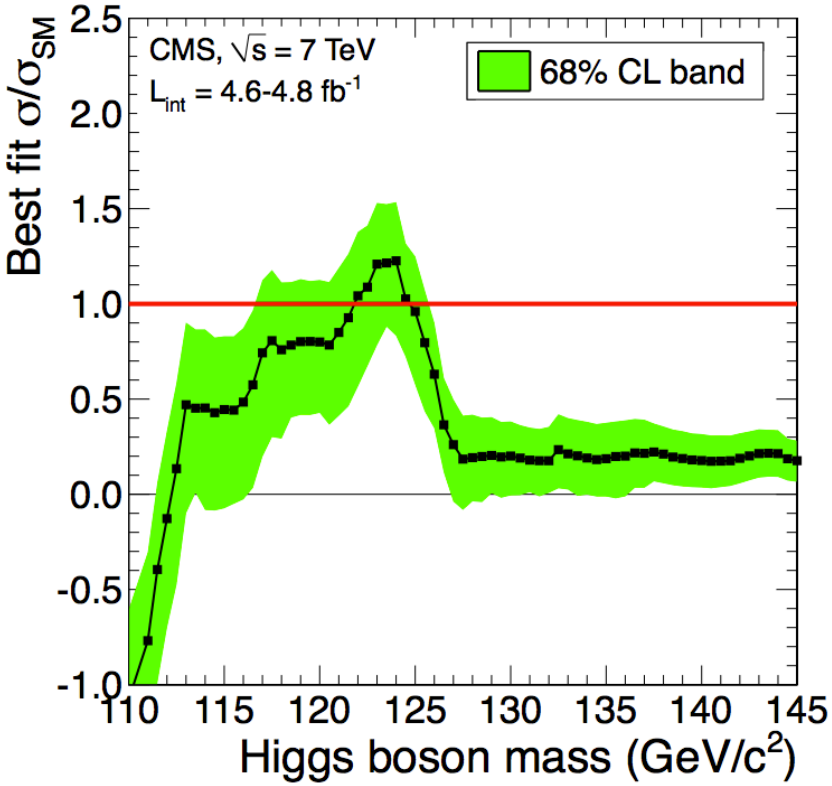




p-values and fitted μ

arXiv:1202.1488
Accepted by PLB





Local significance: 3.1 σ

Global p-value significance:

2.1 σ in [110-145] GeV

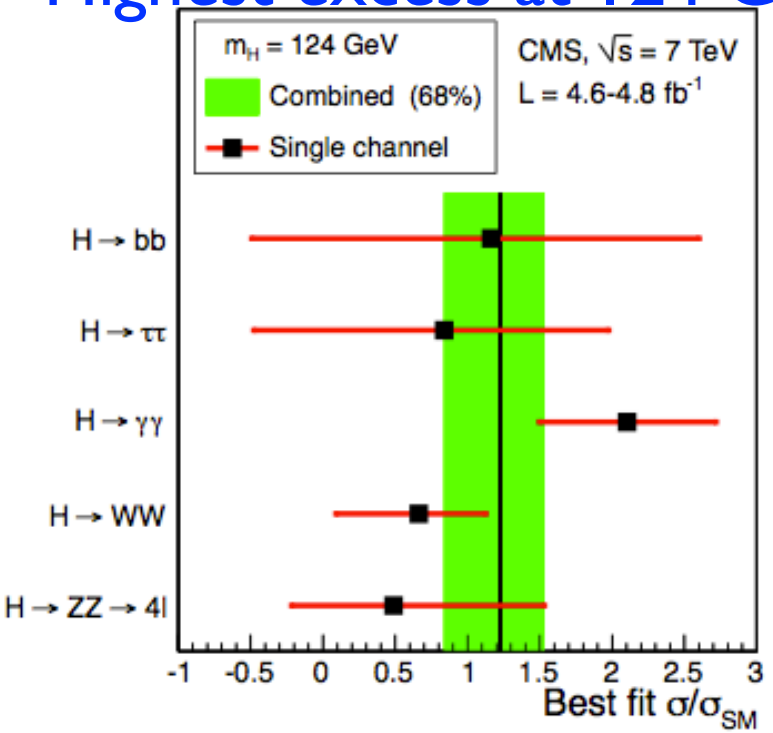
1.5 σ in [110-600] GeV

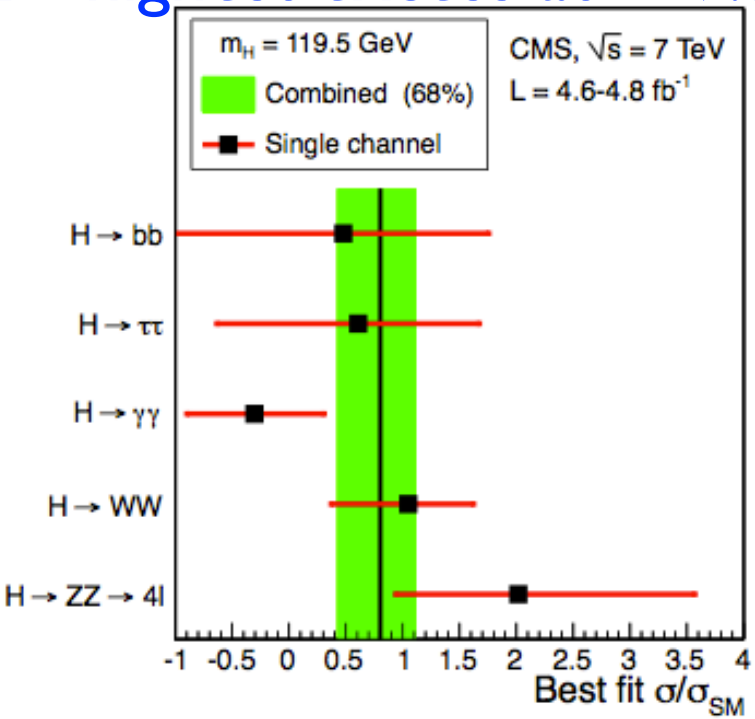
Largest excess observed

- Fitted signal strength in the different channels and in combination

arXiv:1202.1488
Published by PLB

Highest excess at 124 GeV 2nd highest excess at 119.5 GeV





- All investigated channels show consistent excess but the observed excess is also consistent with BG only

- Local p-value significance ~2.3 σ
- Excess mainly coming from the 3 ZZ->4l events
- Deficit in that region in H->γγ