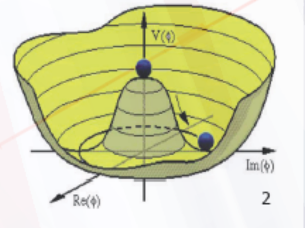
Search for the Standard Model Higgs in H→ZZ→4l channel with the CMS experiment

Nicola De Filippis
Politecnico and INFN Bari

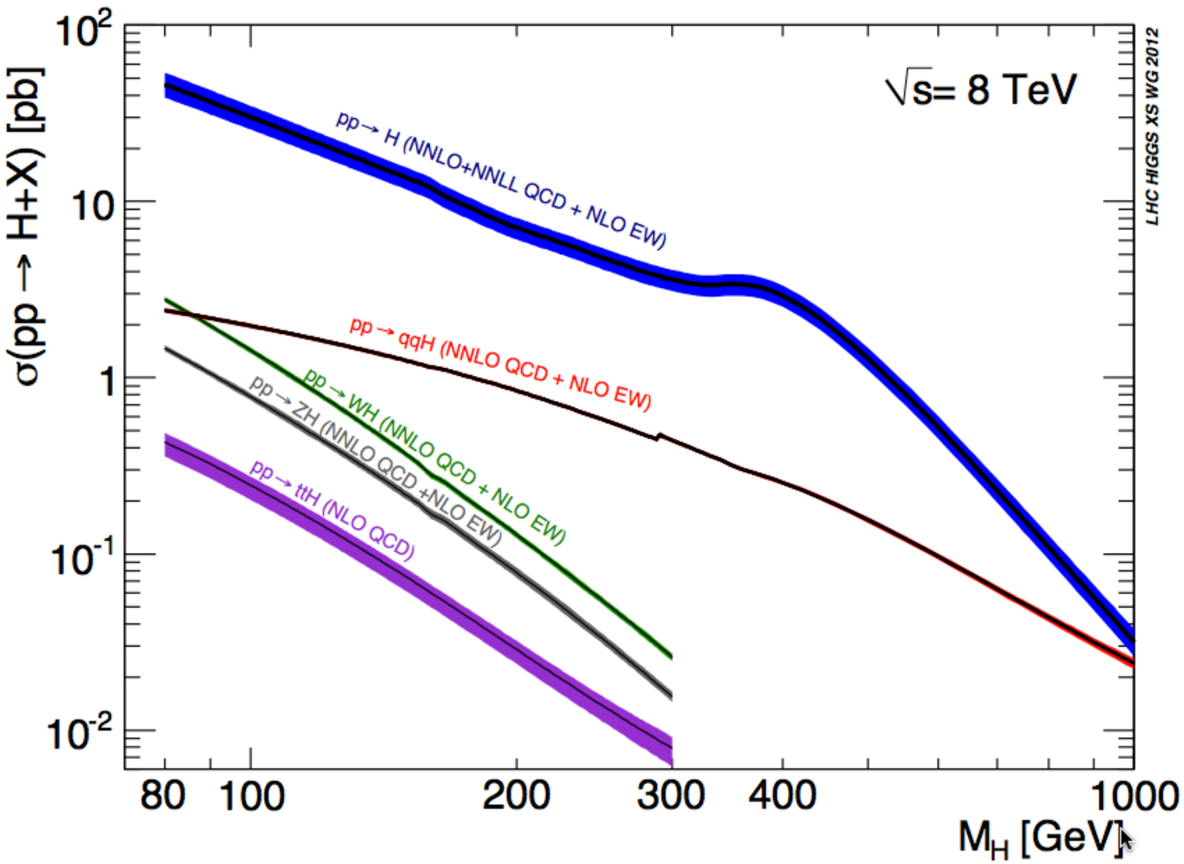


CMS Experiment at LHC, CERN
Data recorded: Wed May 23 21:09:26 2012 CEST
Run/Event: 194789 / 164079659

Outline

- SM Higgs production and decay at LHC
- Searches for SM Higgs in CMS in
  - H→ZZ→ 4l (e,μ,τ)
  - Data/Lumi: 5.1 fb^{-1} (7 TeV, 2011) + 12.2 fb^{-1} (8 TeV, 2012)
  - Statistical interpretation
- Mass measurement and J^{CP} studies

SM Higgs production at LHC

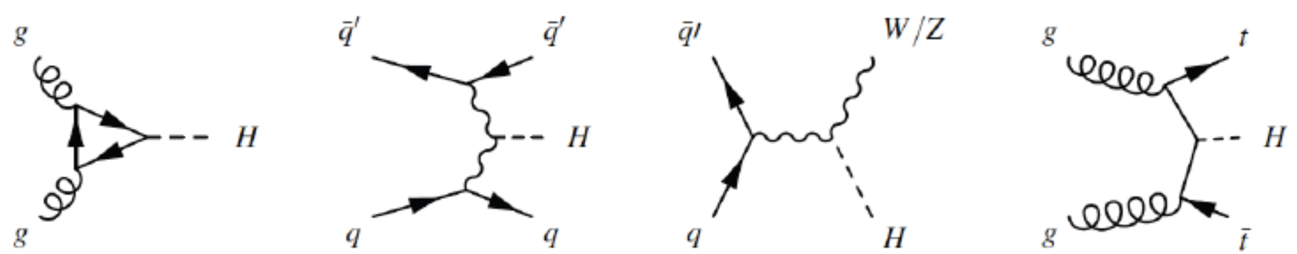


Gluon-gluon fusion:

→ radiative corrections at:

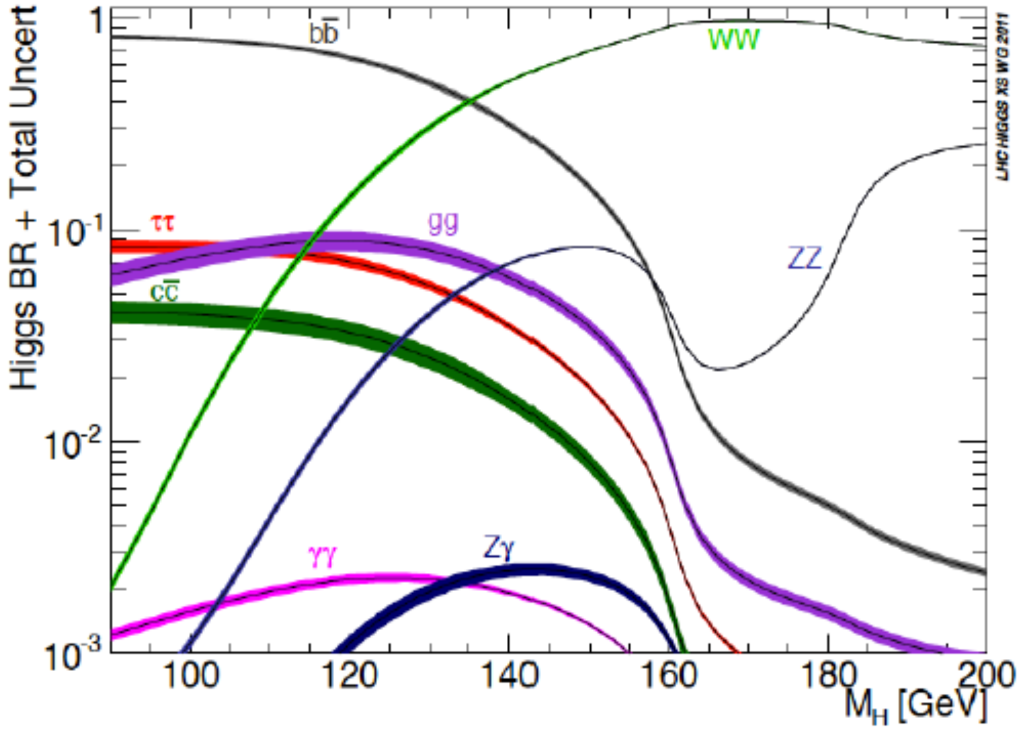
- NLO QCD
- NNLO QCD
- NNLL QCD
- NLO EW

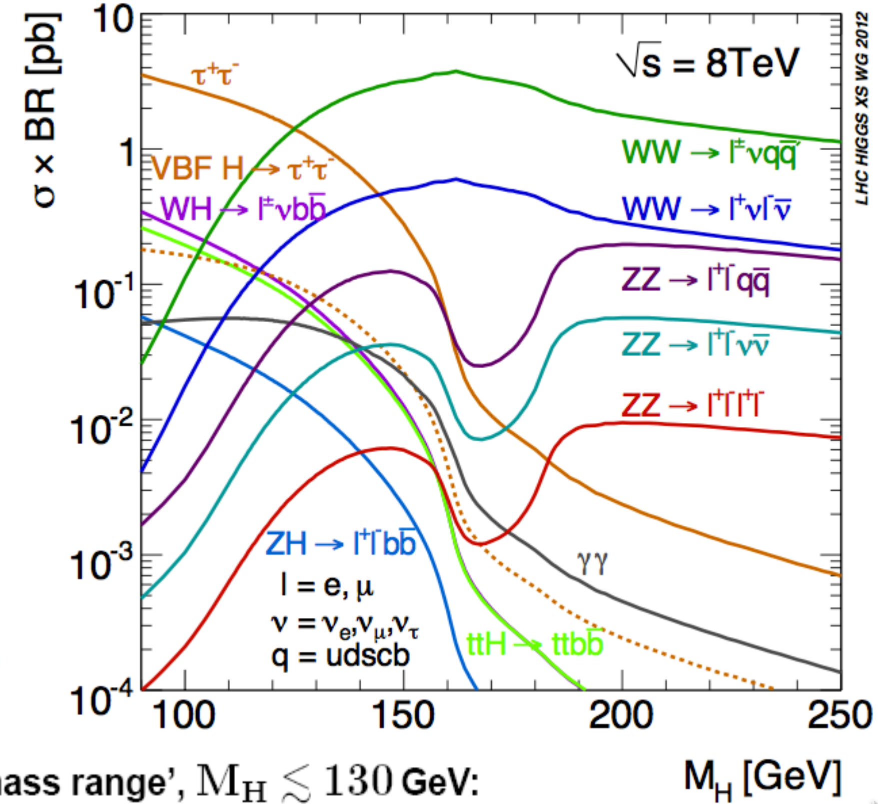
	$K_{NNLO/NLO}$ ($K_{NLO/LO}$)	Scale	PDF+a_s	Total error
ggF	+25% (+100%)	+12% -7%	±8%	+20 -15%
VBF	<1% (+5-10%)	±1%	±4%	±5%
WH/ ZH	+2-6% (+30%)	±1%	±4%	±5%
ttH	- (+5-20%)	+4% -10%	±8%	+12 -18%

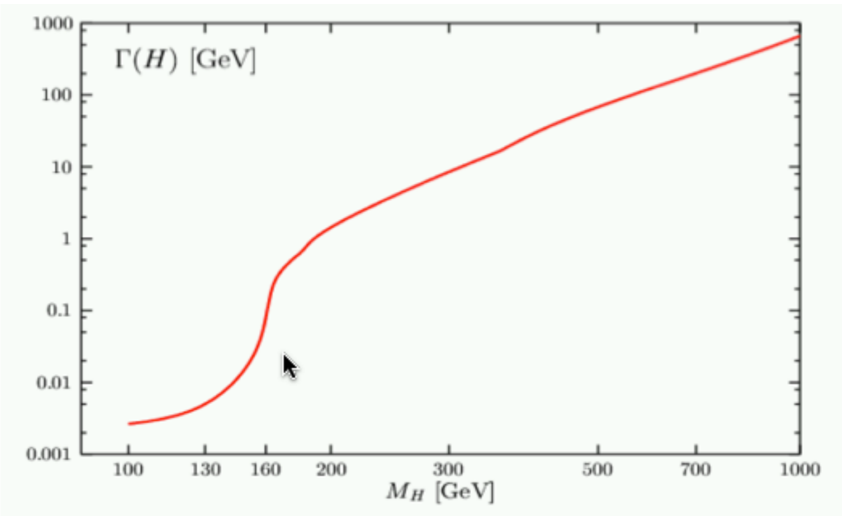


(a) $gg \to H$ (b) VBF (c) VH (d) $t\bar{t}H$

Higgs decay channels

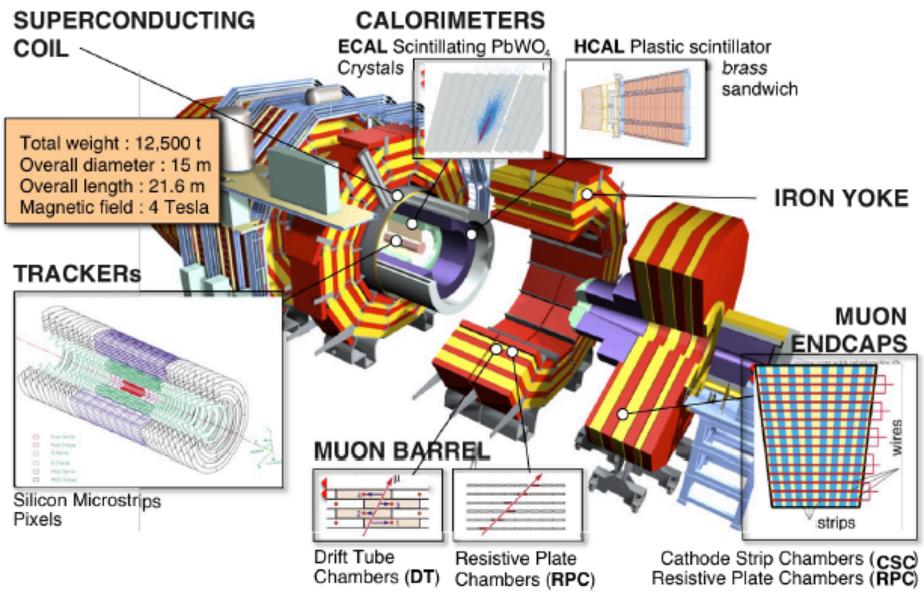


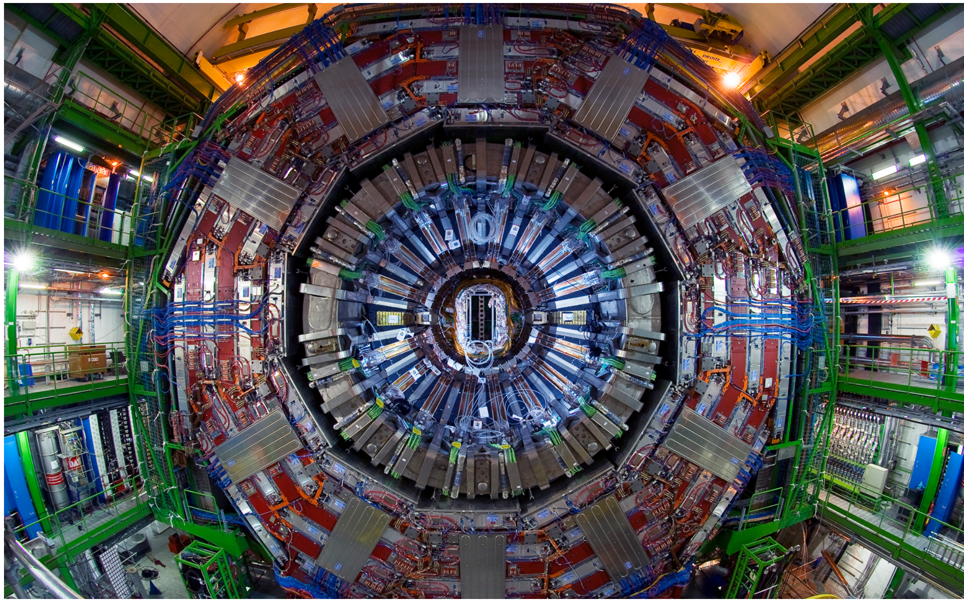




- 'Low mass range', $M_H \lesssim 130$ GeV:
  - $H \to b\bar{b}$ dominant, BR = 60–90%
  - $H \to \tau^+\tau^-, c\bar{c}, gg$ BR= a few %
  - $H \to \gamma\gamma, \gamma Z$, BR = a few permille.
- 'High mass range', $M_H \gtrsim 130$ GeV:
  - $H \to WW^*, ZZ^*$ up to $\gtrsim 2M_W$
  - $H \to WW, ZZ$ above (BR $\to \frac{2}{3}, \frac{1}{3}$)
  - $H \to t\bar{t}$ for high M_H; BR $\lesssim 20\%$.

CMS in a nutshell





|η|<2.5 : Tracker

$\sigma/p_T \approx 10^{-4} p_T \oplus 0.005$

|η|<4.9 : EM Calorimeter

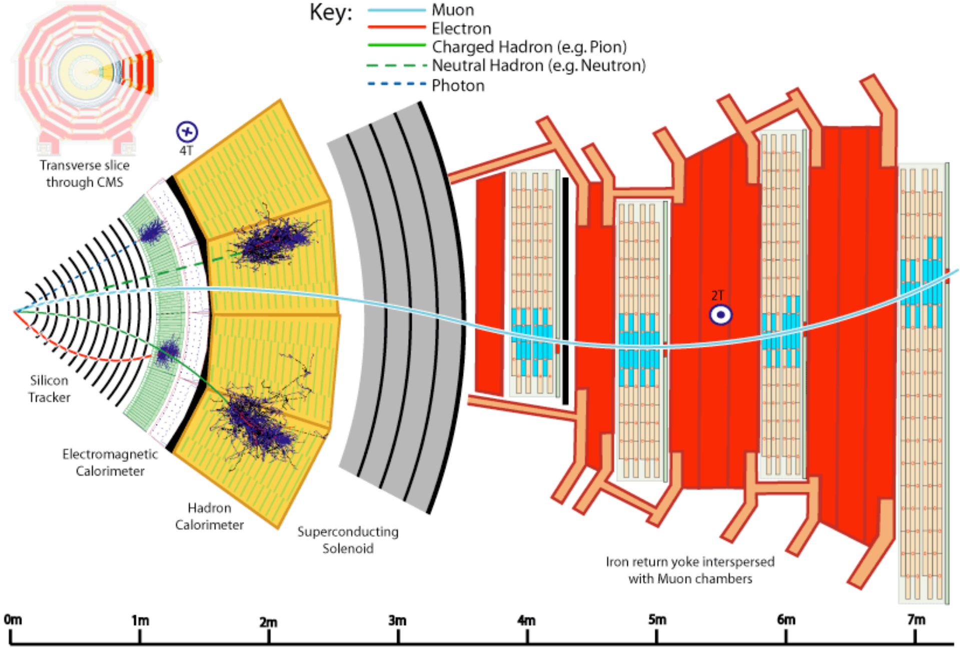
$\sigma/E \approx 0.03/\sqrt{E} + 0.003$

|η|<4.9 : HAD Calorimeter

$\sigma/E \approx 1.0/\sqrt{E} + 0.05$

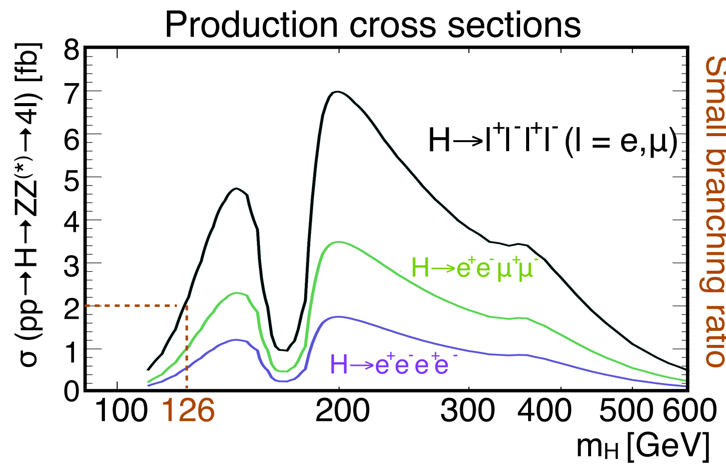
|η|<2.4 : Muon spectrometer

$\sigma/p_T \approx 0.10$ *(1TeV muons)*



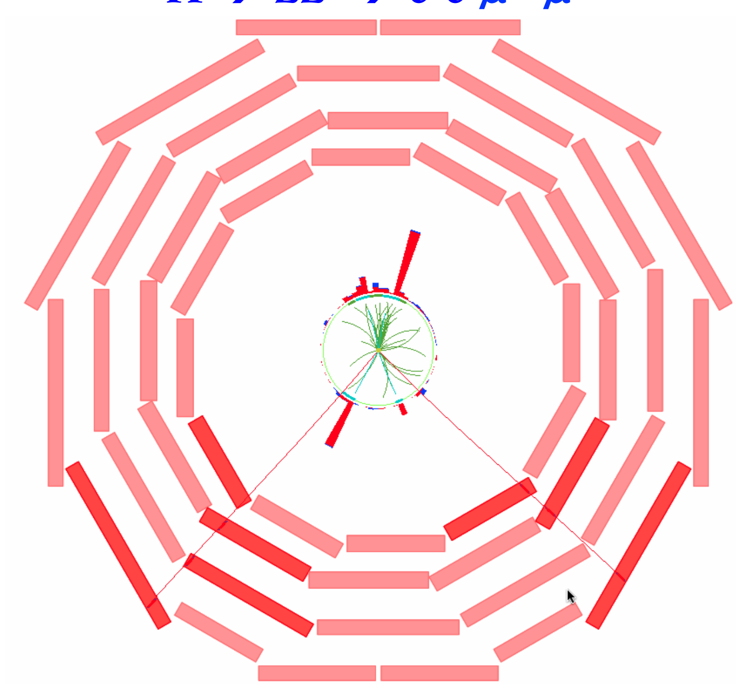
H→ZZ→4l in a nutshell

- Signatures: **4e, 4μ and 2e2μ** final state
  - clean but extremely demanding channel for requiring the **highest possible efficiencies (lepton Reco/ID/Isolation)**.
  - σ x BR small ≈ few fb



- Backgrounds:
  - Irreducible: ZZ*
  - Reducible: Zbb, tt+jets, Z+light jets, WZ+jets

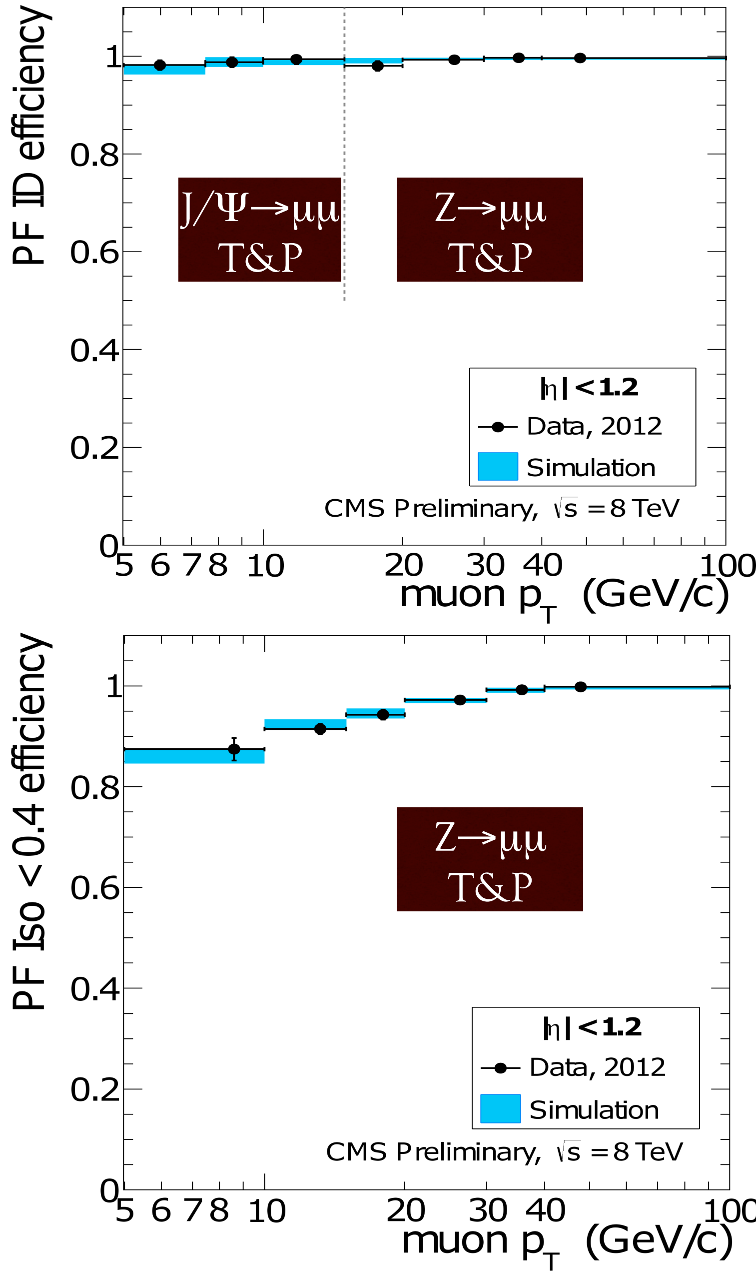
H → ZZ* → e⁺e⁻μ⁺μ⁻



- Sensitivity: $115 < m_H < 1000$ GeV
- Selection strategy:
  - triggering on double leptons
  - applying reco, id and isolation of leptons
  - recovery of FSR photons
  - use of impact parameter
  - m_Z and m_{Z^*} constraint
  - kinematical discriminant / scalarity of the Higgs

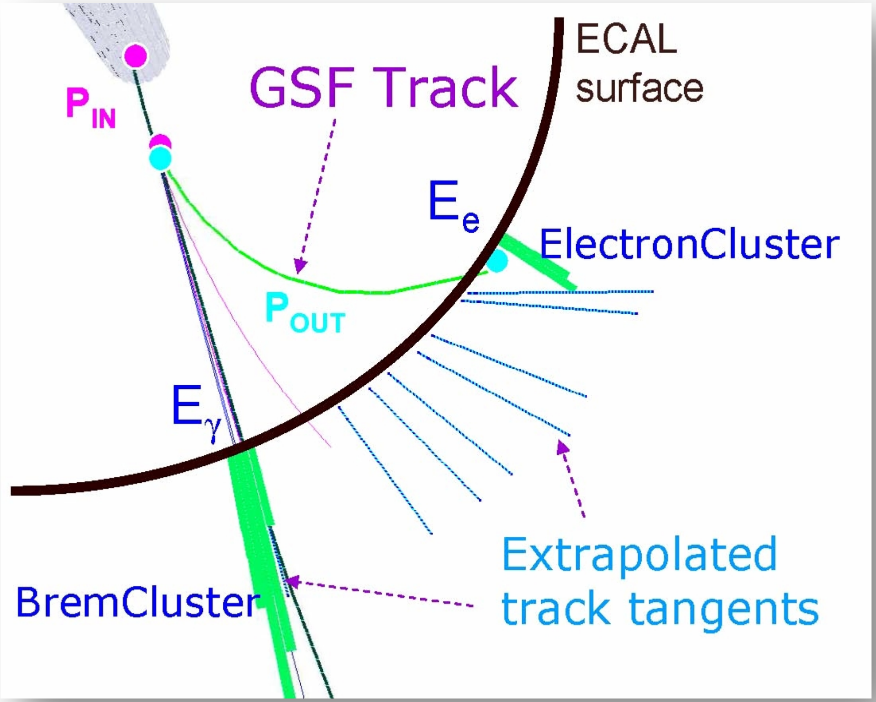
Muon reco and identification

- Particle Flow Muon Identification in 2012 ZZ analysis
  - Exploit information from all subdetectors
- High efficiency down to p_T = 5 GeV
  - Exploit also tracker-based muon ID
  - Efficiency controlled in data with J/Ψ and Z T&P
- Tighter quality criteria applied in analysis handling final state with τ to further suppress reducible backgrounds



Electron reconstruction

- From Superclusters in ECAL
  - Collect energy spread in ϕ due to bremsstrahlung (E_T>4 GeV)
- Dedicated track finding and GSF fit
  - Gaussian Sum Filter to cope with change of curvature and enable hit collection up to ECAL
- ECAL-seeded reconstruction
  - Complemented by a tracker-seeded reconstruction to gain efficiency at low p_T
- Electron classes
  - Separate "simple" and "more complicated" electron patterns due to bremsstrahlung
- Energy/Momentum
  - A weighted combination of E and p from ECAL and Tracker information
  - ECAL information obtained by a Regression technique, such as in $H \rightarrow \gamma\gamma$

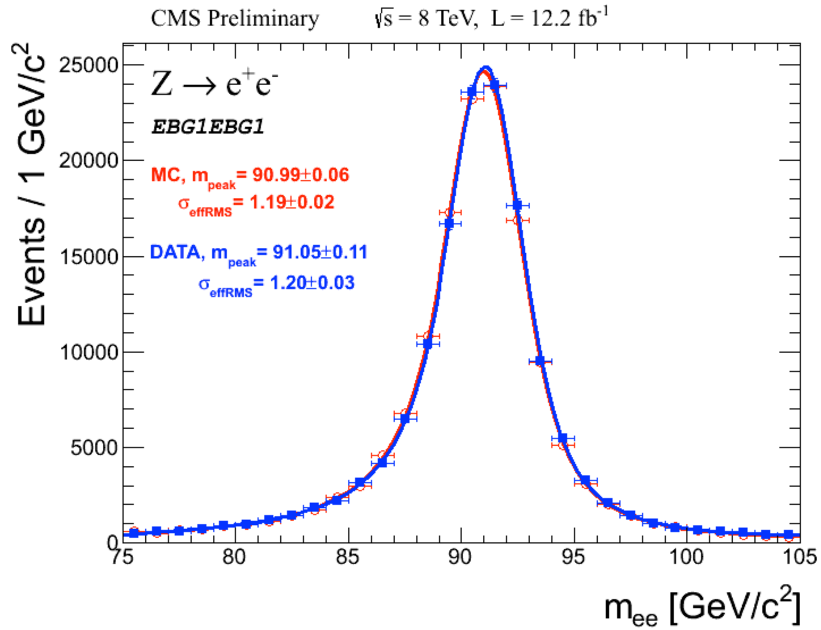


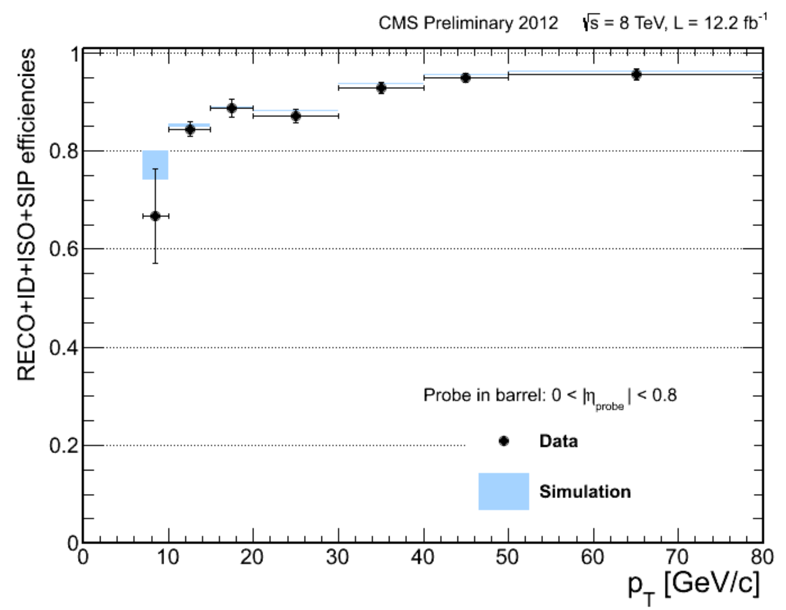
- Energy scale and resolution
  - Determined at the Z peak for all of the different electron categories
  - Also control at low p_T via $J/\psi \rightarrow ee$

→ 10% improvement in Higgs mass resolution

Electron identification

- BDT MVA analysis for electron ID:
  - combine ECAL, tracker, ECAL-tracker-HCAL matching (shower shape observables) and impact parameter (IP) observables
- Training samples
  - signal and W+1 jet (fake) for training / Z+ 1 jet to optimize the working point
- Performance:
  - 30% efficiency improvement in H→ZZ-→4e wrt cut based ID
- Efficiencies
  - Via tag-and-probe at the Z→ee peak





Leptons selection for 4l analysis

Loose leptons used for reducible bkg estimation:

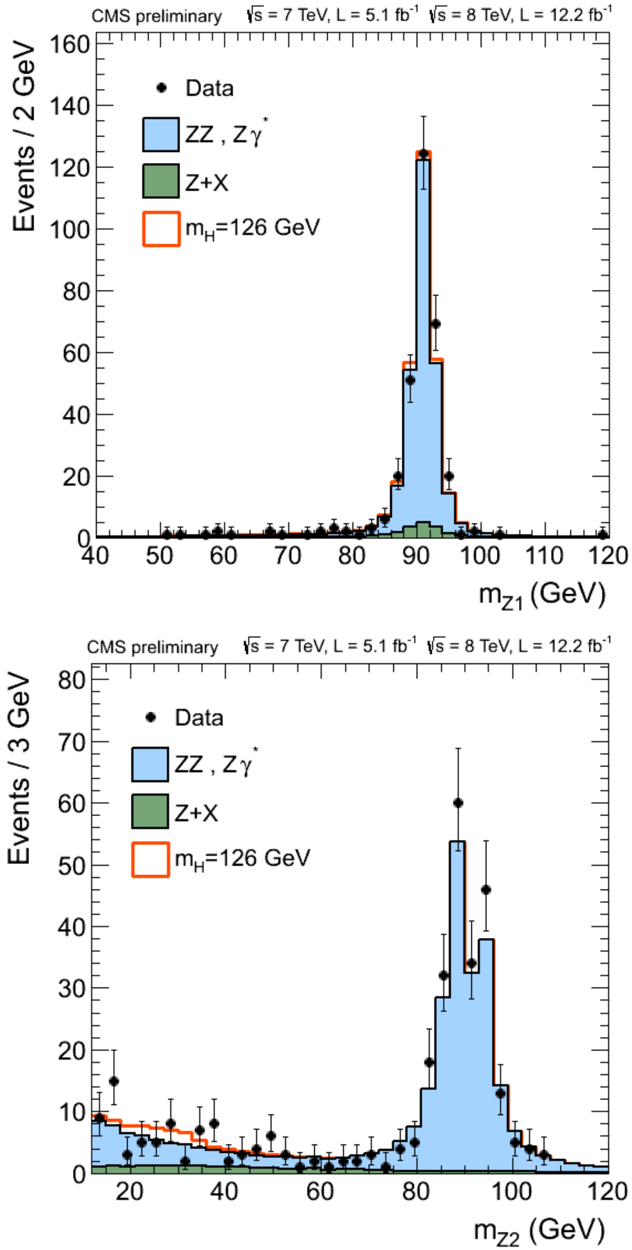
- Electrons:
  - $|\eta| < 2.5$
  - $p_T > 7$ GeV
  - having 0 or 1 expected missing inner hits
- muons (global or tracker)
  - $|\eta| < 2.4$
  - $p_T > 5$ GeV
  - arbitrated and with a requirement on the shared segments
- loose requirements on the transverse ($d_{xy} < 0.5$ cm) and longitudinal ($d_z < 1$ cm) IP
- $\Delta R > 0.02$ between the leptons.

Good leptons used for baseline selection:

- loose leptons with more criteria:
  - electrons passing the electron ID
  - muons passing the Particle Flow Muon ID
  - Relative PFIso < 0.4 (loose requirement)
  - the significance of the impact parameter SIP3D = $IP/\sigma_{IP} < 4$
  - FSR photons combined with leptons

Event selection

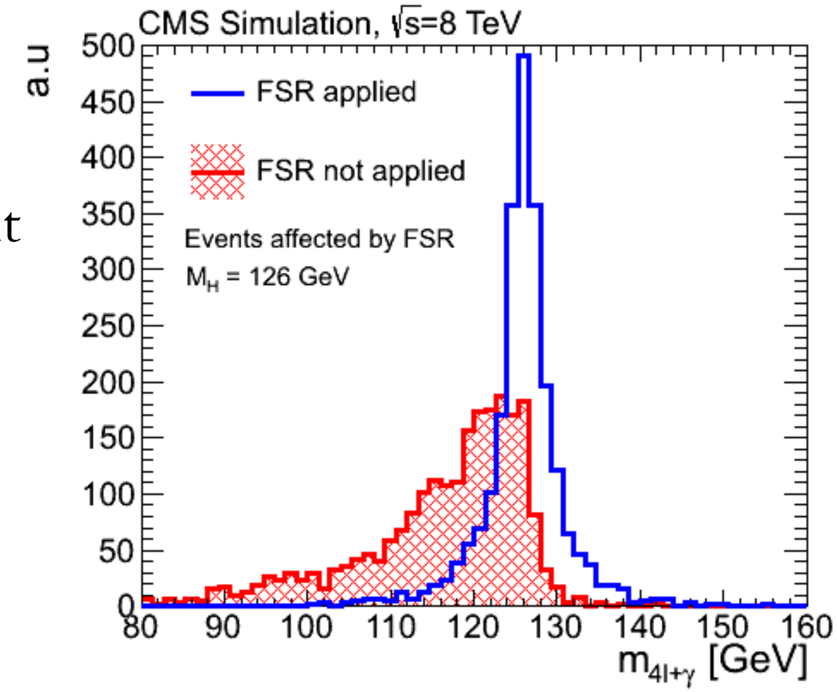
- Trigger: di-lepton signatures (ee, eμ or μμ) + tri-electron trigger
- Leptons
  - muons: pT > 5 GeV, |η| < 2.4, isolated, compatible with PV
  - electrons: pT > 7 GeV, |η| < 2.5, isolated, compatible with PV
- Lepton selection
  - at least one lepton with pT > 20 GeV
  - at least two leptons with pT > 10 GeV
- First Z candidate (Z1)
  - chosen as di-lepton pair with m(ll) closest to mZ
  - apply: 40 < m(ll) < 120 GeV
- Second Z candidate (Z2)
  - build from remaining highest pT leptons
  - apply: 4 < m(ll) < 120 GeV
- mll > 4 GeV of opposite-sign and same flavor pairs
- Kinematics
  - Higgs: m(4l) > 100 & mZ2 > 12 GeV



Final State Radiation recovery

FSR photon with p_T>2 GeV

- affect 8% (15%) of muons (eles)
- often collinear to the lepton and low energy
- collected with electron reco algo but not for muon reco



Selection:

- Particle Flow ID
- E_T> 2 GeV
- $|\eta|$<2.4
- PF Isolation < 1.0
- Associated to a lepton if ΔR<0.5
- Associates photon with Z if:
  - $M(ll+\gamma)$< 100 GeV
  - $|M(ll+\gamma) - M_Z| < |M(ll) - M_Z|$
- photons removed from lepton isolation calculation

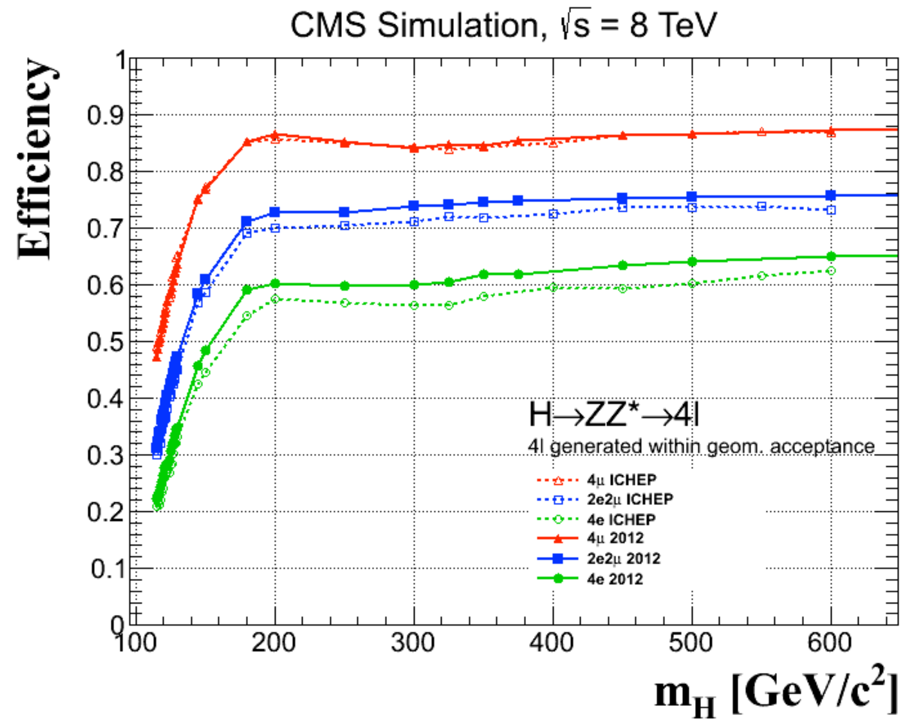
Performance for m_H=126 GeV

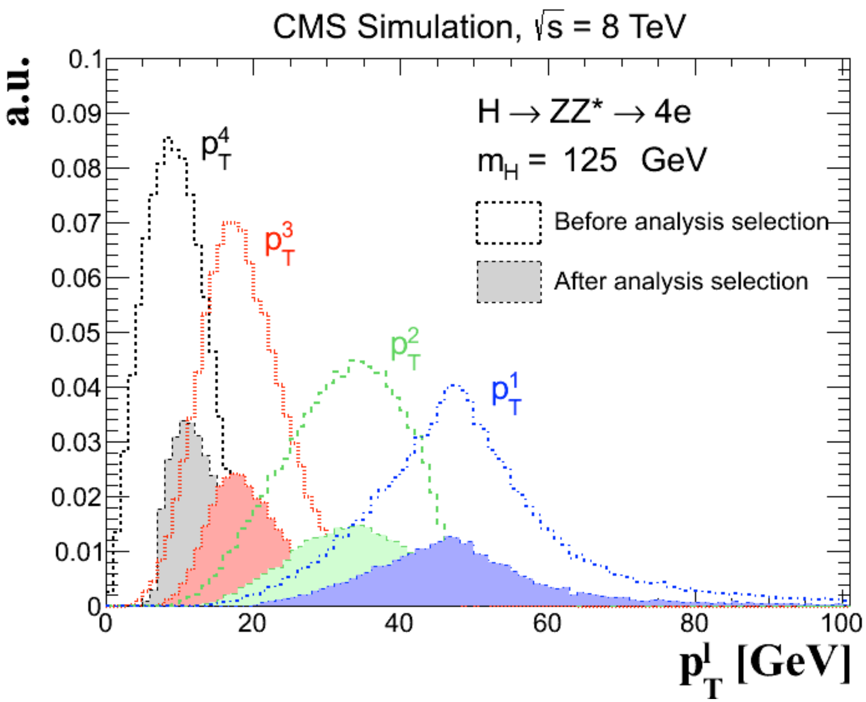
- 6% of events affected by FSR
- Average purity of 80%
- 2% events added in analysis (mostly 4μ)

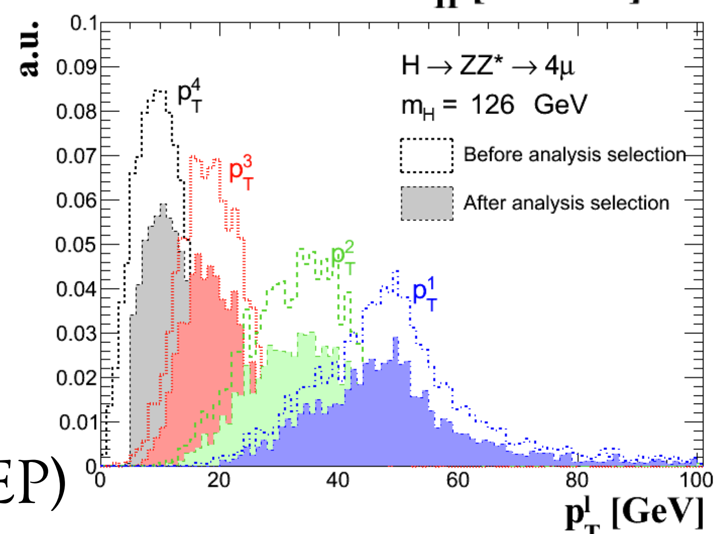
Studies with data via Z→llγ and Z→4lγ

Signal efficiency

Best effort to recover efficiency lor low p_T electrons and muons







At m_H=125 GeV:

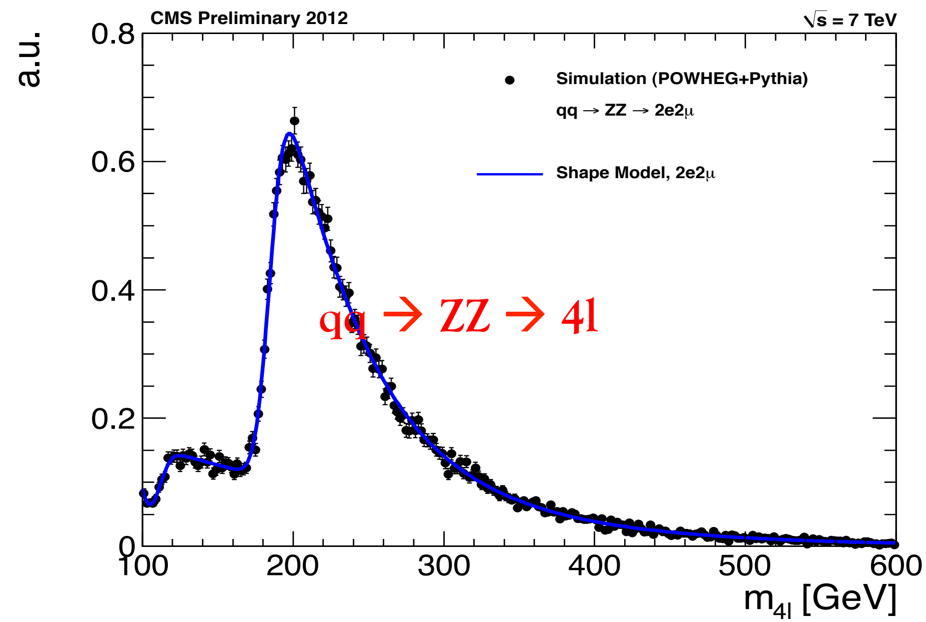
Eff = 30%, 58%, 42% for 4e,4μ,2e2μ

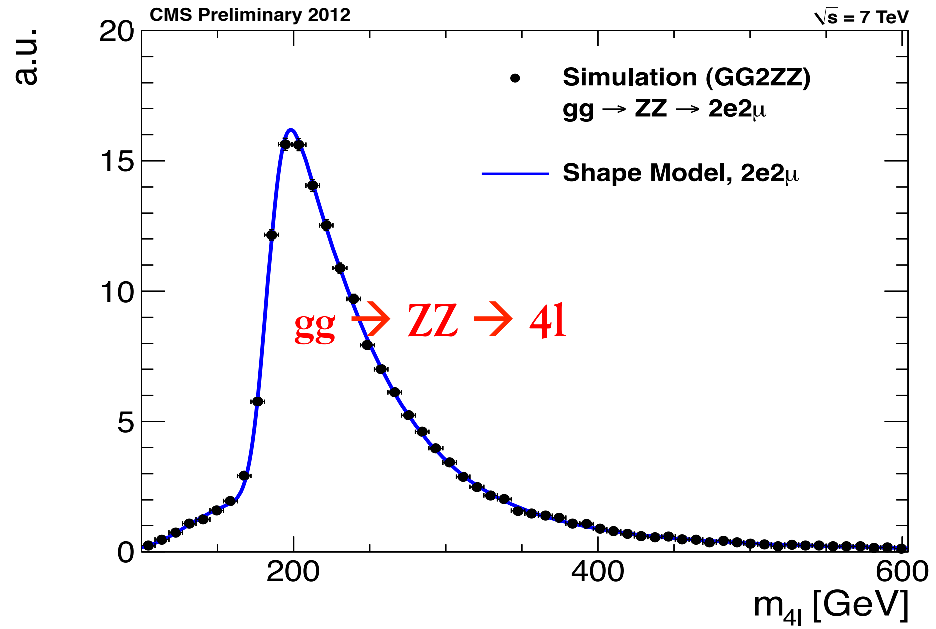
(2% gain in electron channels . w.r.t to ICHEP)

Background estimate: ZZ

Irreducible background: ZZ → 4l

- Estimated **using simulation**
- Phenomenological shape models
- Events yield uncertainty related to PDF+α_s and QCD scale evaluated using MCFM
- Corrected for data/MC scale





Background estimate: Z+X

REDUCIBLE background estimated **from data**:

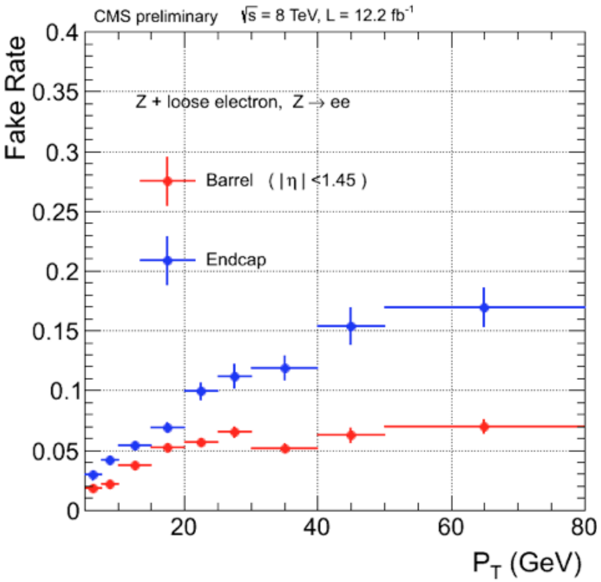
Z→ll + leptons from b-decays or from mis–id of light jets

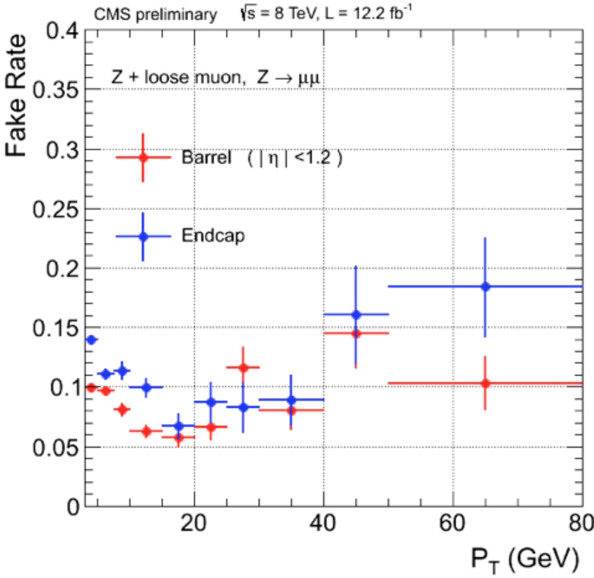
- Measure probabilities for lepton mis-identification
- Control samples:

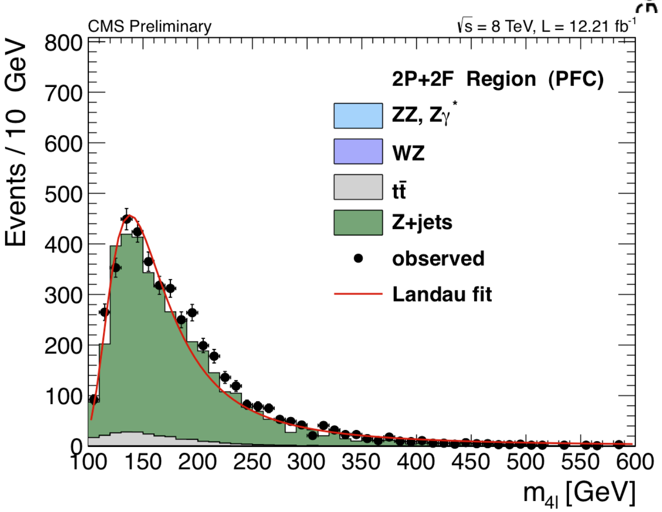
  a) Z + 1 good lepton + 1 loose lepton + MET<25 GeV (3P+1F)

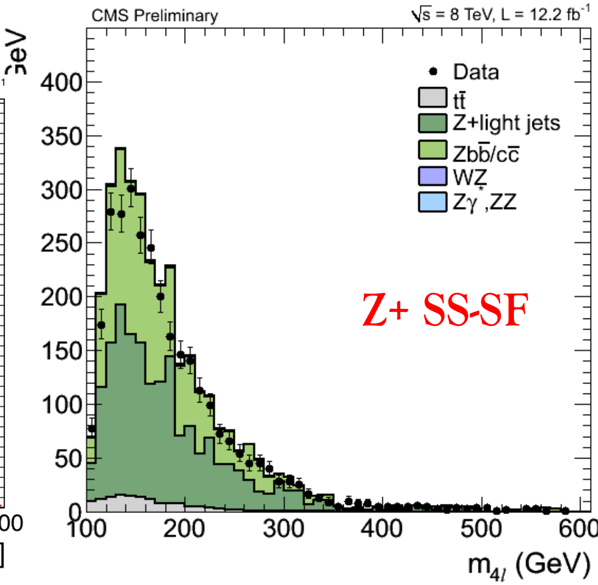
  b) Z + 2 loose leptons + MET<25 GeV (2P+2F)

  c) Z + 2 loose lepton with same sign and same flavour (SS-SF) + m_{4l}>100 and m_{z1}, m_{z2} cut
- Events yield extrapolated in signal region









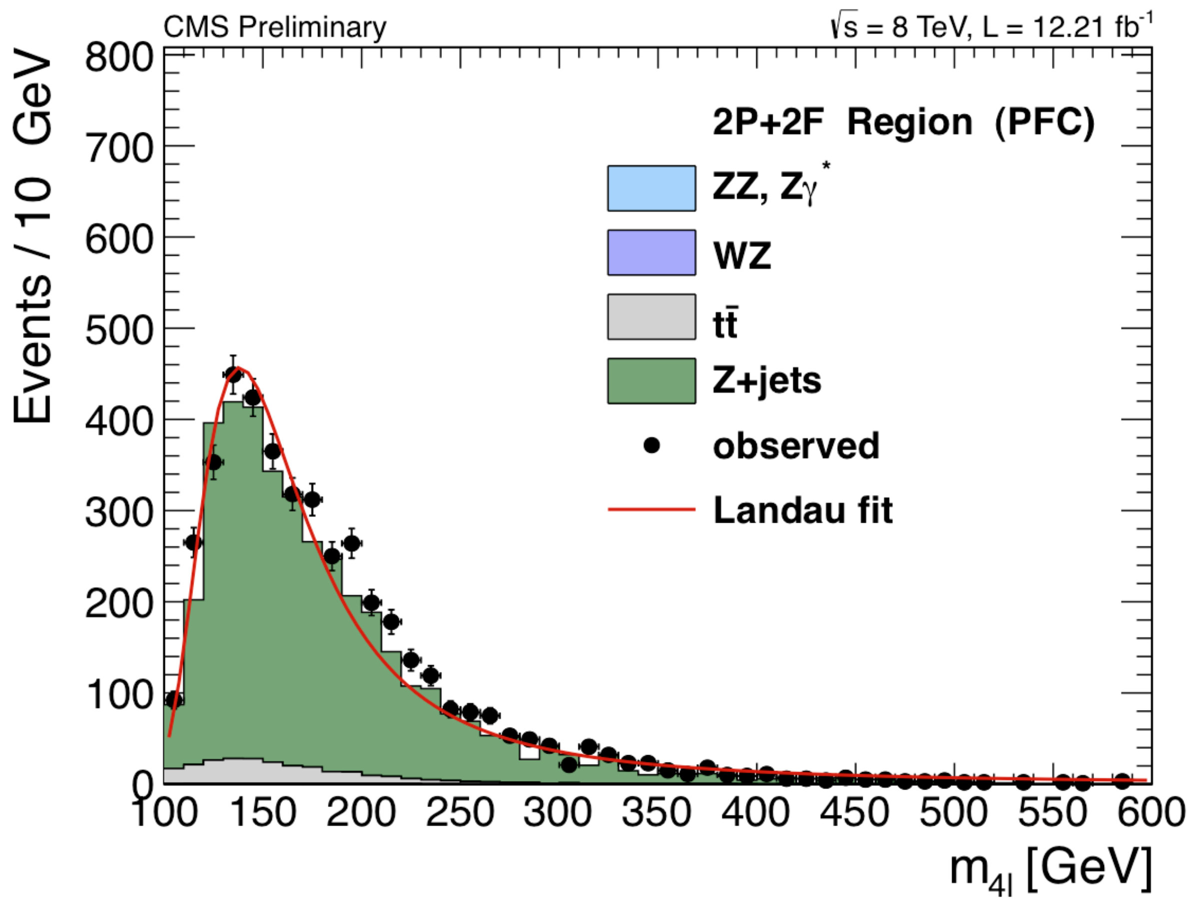
Background control: closure test

Validation in data using "wrong flavors & charges" events (WFC) in control region 2P+2F

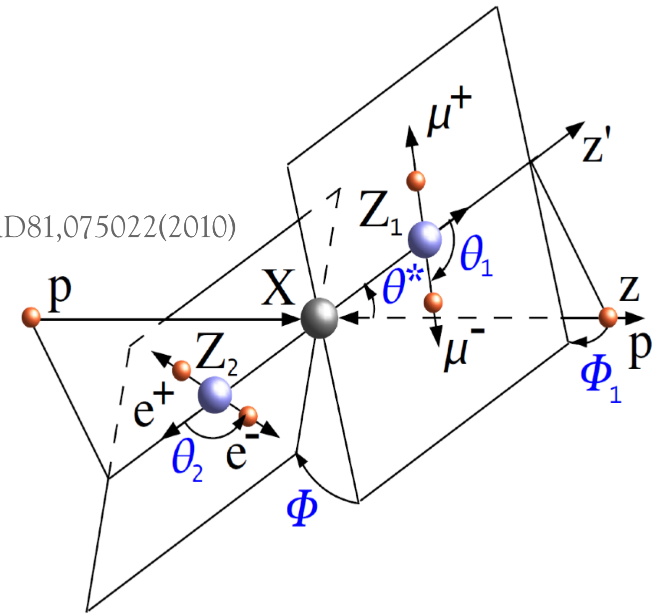
- Extrapolation to WFC signal region

The difference between the predicted and the observed n. events in WFC signal region is taken as the **measurement of the uncertainty of the method**

TOTAL uncertainty <50% dominated by the statistical uncertainty in signal region (30%)

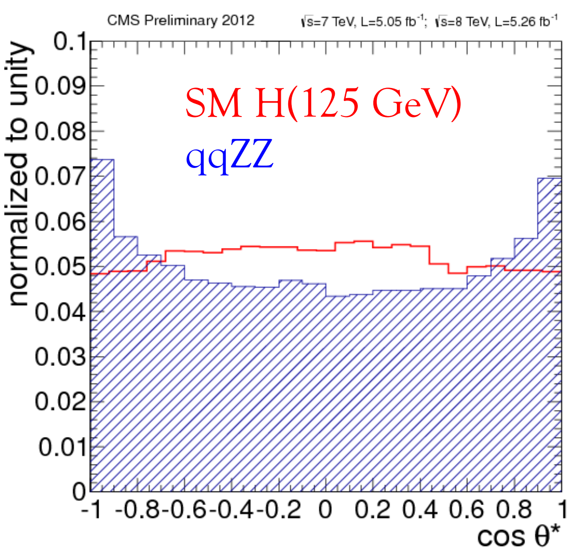


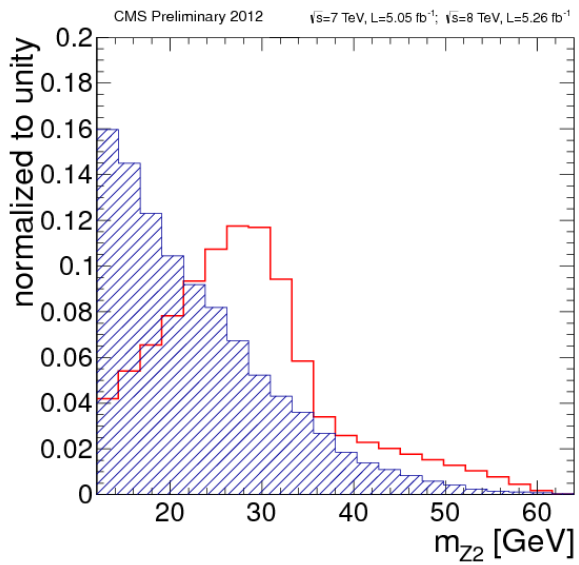
Kinematical discriminant - MELA

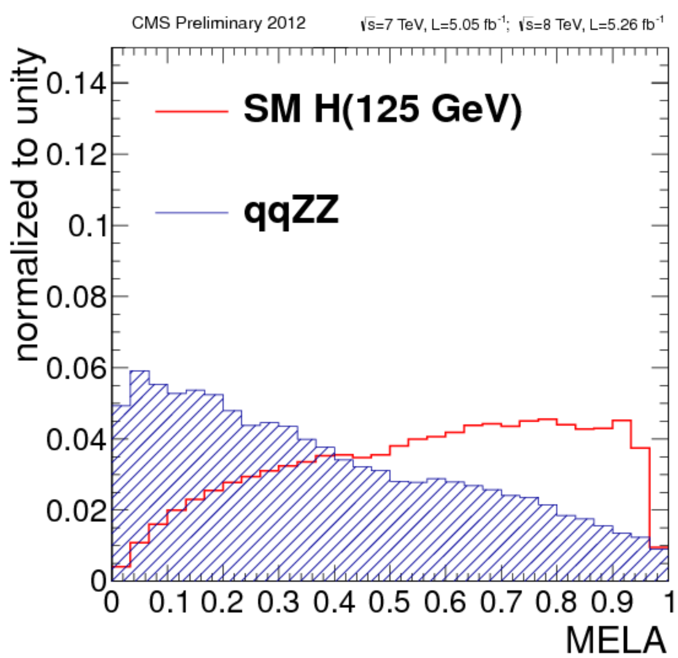


Matrix Element Likelihood Analysis:
uses kinematic inputs for
signal to background discrimination
$\{m_1, m_2, \theta_1, \theta_2, \theta^*, \Phi, \Phi_1\}$

$$\mathrm{MELA} = \left[1 + \frac{\mathcal{P}_{\mathrm{bkg}}(m_1, m_2, \theta_1, \theta_2, \Phi, \theta^*, \Phi_1 | m_{4\ell})}{\mathcal{P}_{\mathrm{sig}}(m_1, m_2, \theta_1, \theta_2, \Phi, \theta^*, \Phi_1 | m_{4\ell})}\right]^{-1}$$







Systematic uncertainties

Theoretical uncertainties

Source of uncertainties	Error for different processes						
	ggH	VBF	WH	ZH	ttH	ZZ	ggZZ
gg partonic luminosity	7.2-9.2				0-9.8		10
$qq/q\bar{q}$ partonic luminosity		1.2-1.8	0-4.5	0-5.0		5	
QCD scale uncert. for $gg \to H$	5.5-7.9						
QCD scale uncert. for VBF qqH		0.1-0.2					
QCD scale uncert. for VH			0-0.6	0-1.5			
QCD scale uncert. for ttH					0-8.8		
4ℓ-acceptance for $gg \to H$	negl.	negl.	negl.	negl.	negl.		
Uncertainty on $BR(H \to 4\ell)$	2	2	2	2	2		
QCD scale uncert. for ZZ(NLO)						2.6-6.7	
QCD scale uncert. for $gg \to ZZ$							24-44
CB mean, parametric	0.2 for electrons, 0.1 for muons						
CB sigma, parametric	20						
CB tail parameter, parametric	5.0						

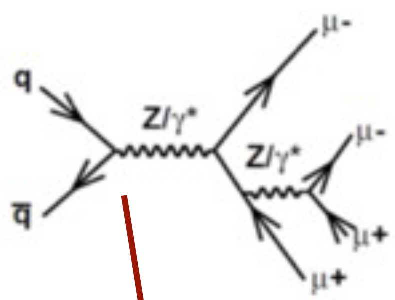
Experimental uncertainties

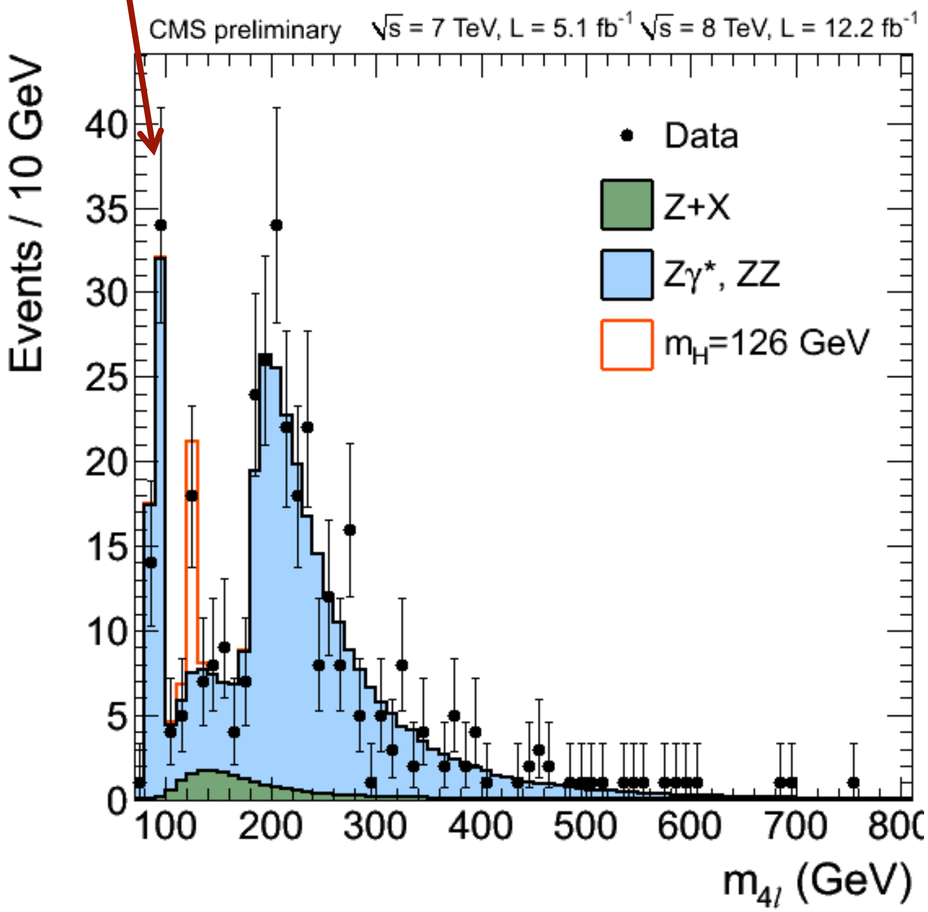
Source of uncertainties	Error for different processes					
	$H \to ZZ \to 4\ell$			$ZZ/ggZZ \to 4\ell$		
	$4e$	4μ	$2e2\mu$	$4e$	4μ	$2e2\mu$
Luminosity	2.2 (4.4 for 8 TeV)					
Trigger	1.5%					
electron reco/ID/isolation (4e)	5.5-11					
muon reco/ID/isolation (4mu)	1.9-3.8					

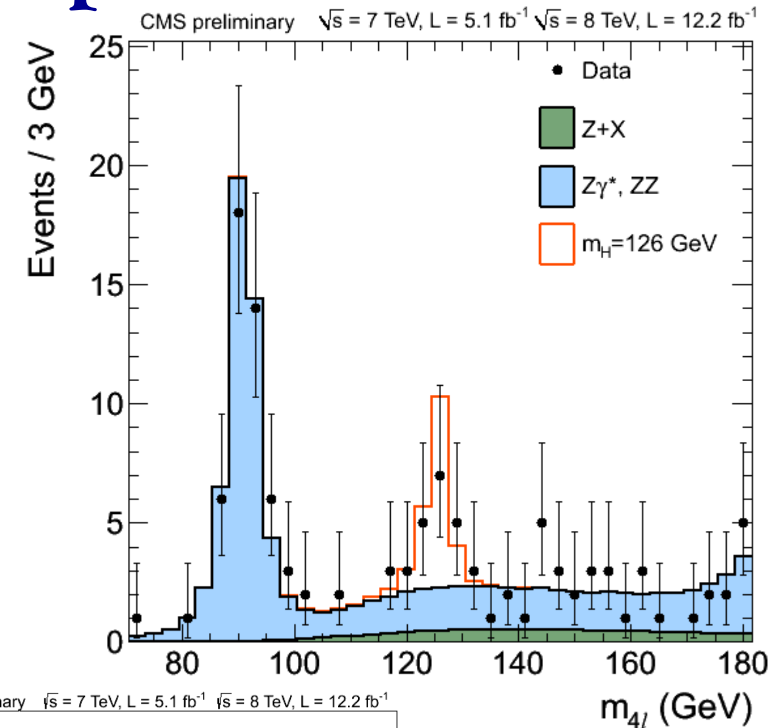
Uncertainty on background estimation

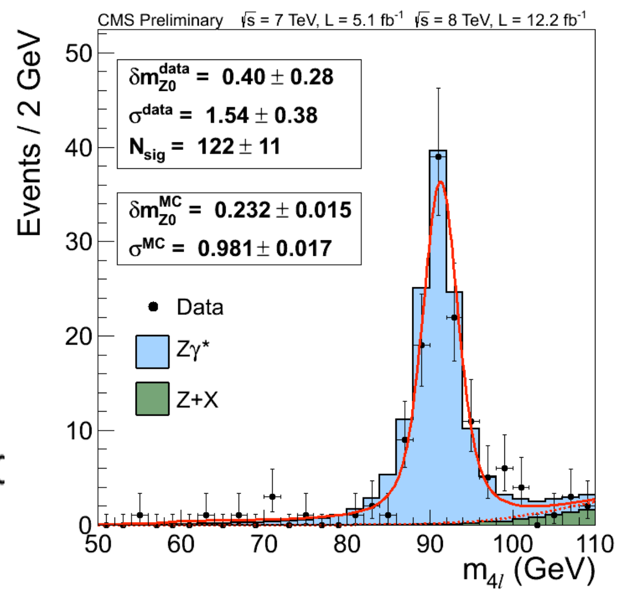
$4e$	4μ	$2e2\mu$
-40..+90 (-30..+90)	-	-
-	-40..+60 (-40..+80)	-
-	-	-50..+60 (-30..+100)

4l invariant mass spectrum

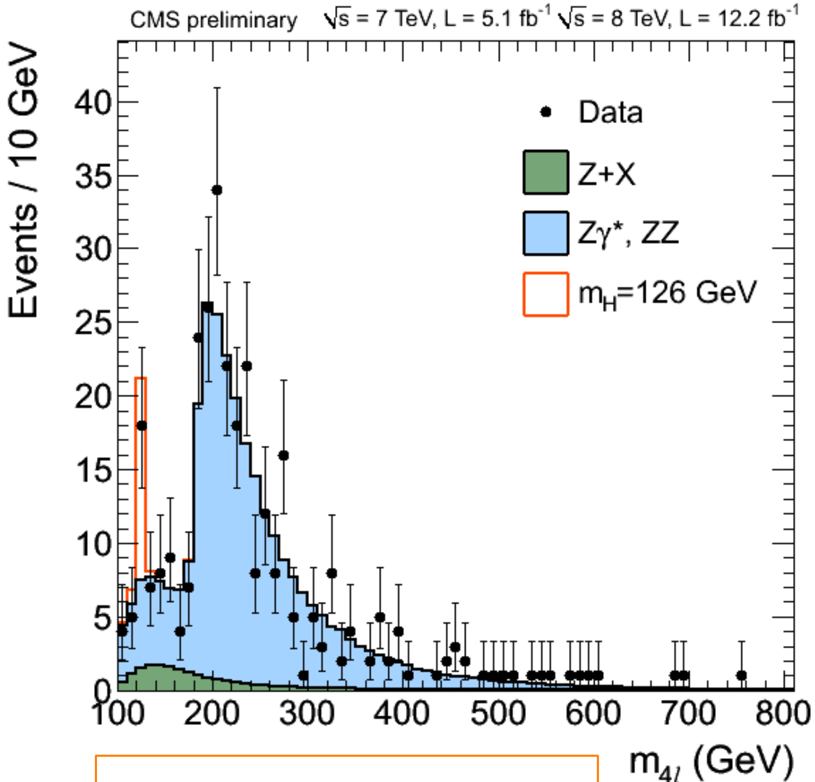






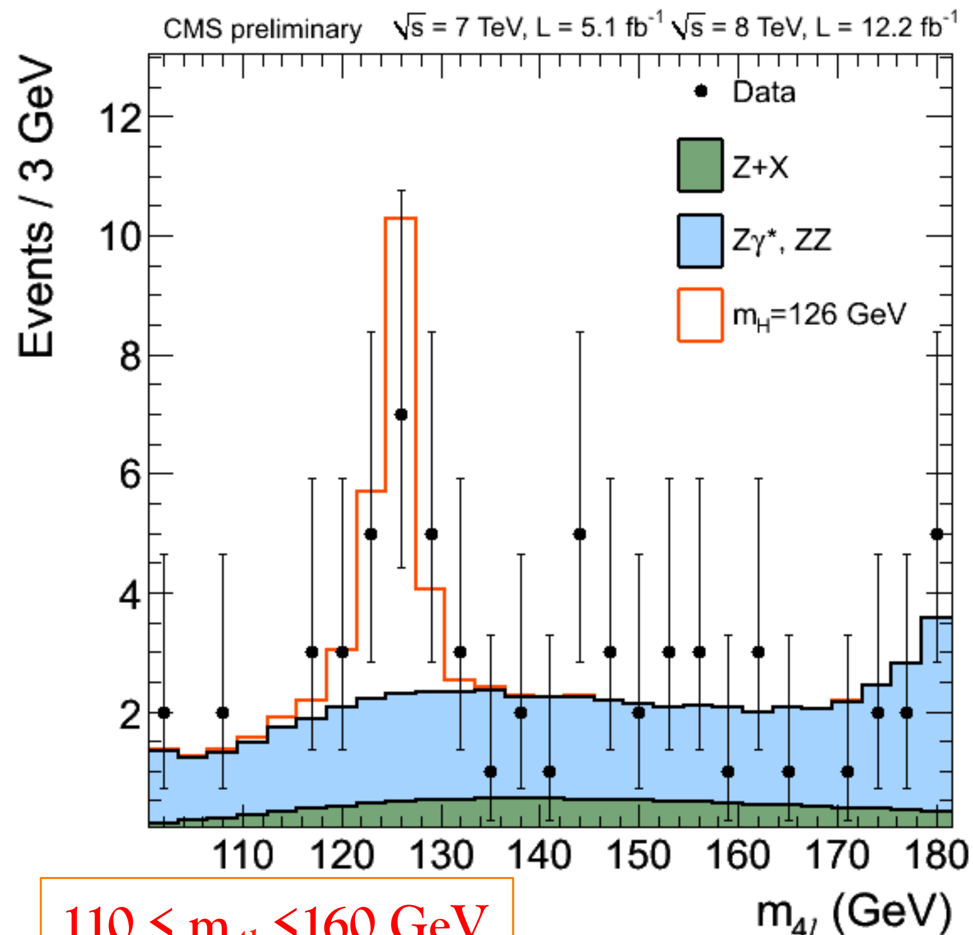


4l invariant mass spectrum



$100 < m_{4l} < 1000$ GeV

Channel	4e	4μ	$2e2\mu$	$2\ell 2\tau$
ZZ background	53.0 ±6.3	82.7 ±8.9	131.1 ±14.3	19.0 ±2.3
Z+ X	$7.6^{+6.9}_{-5.2}$	$2.9^{+2.2}_{-1.6}$	$10.1^{+9.9}_{-6.5}$	20.4 ±6.2
All background expected	$60.7^{+9.3}_{-8.2}$	$85.6^{+9.2}_{-9.1}$	$141.3^{+17.3}_{-15.7}$	39.4 ±6.6
m_H = 125 GeV	2.4 ±0.4	4.6 ±0.5	6.0 ±0.7	–
m_H = 126 GeV	2.7 ±0.4	5.1 ±0.6	6.6 ±0.8	–
m_H = 200 GeV	15.5 ±1.9	23.1 ±2.6	38.5 ±4.3	5.6 ±0.6
m_H = 350 GeV	9.5 ±1.2	13.6 ±1.5	23.2 ±2.7	5.7 ±0.6
m_H = 500 GeV	3.3 ±0.4	4.7 ±0.6	8.1 ±0.9	2.8 ±0.3
m_H = 800 GeV	0.5 ±0.1	0.6 ±0.1	1.1 ±0.1	0.3 ±0.1
Observed	59	95	162	45

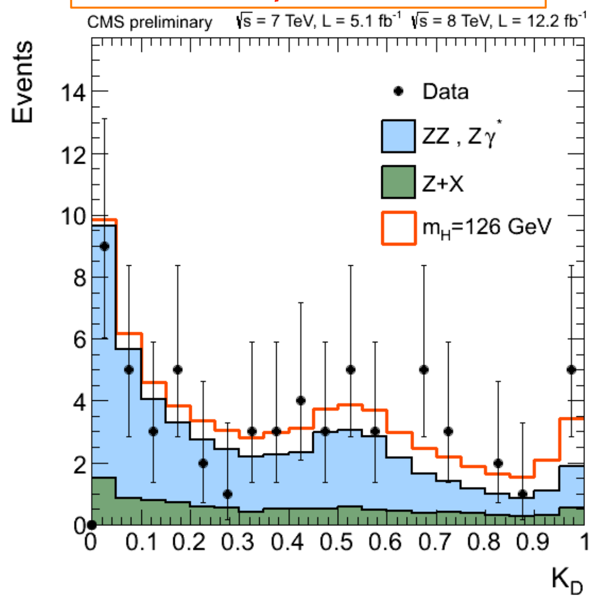


$110 < m_{4l} < 160$ GeV

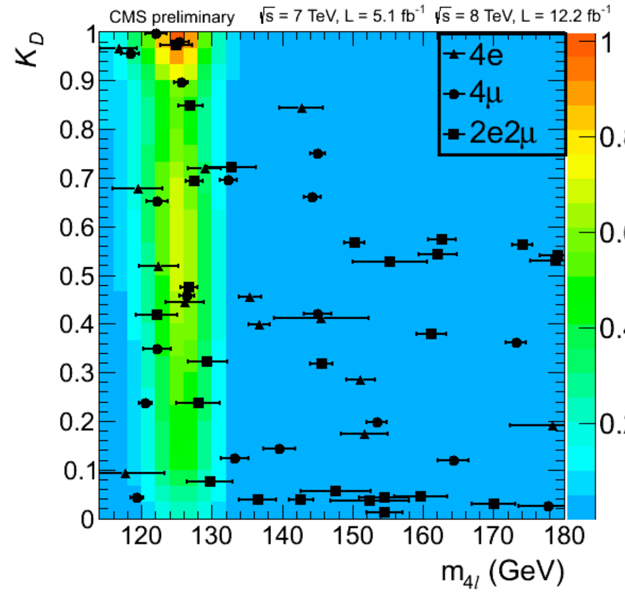
Channel	4e	4μ	$2e2\mu$	4ℓ
ZZ background	4.7 ±0.6	9.6 ±1.0	12.5 ±1.4	26.8 ±1.8
Z+ X	$3.4^{+3.0}_{-2.3}$	$1.6^{+1.2}_{-0.9}$	$5.6^{+5.4}_{-3.6}$	$10.6^{+5.3}_{-4.4}$
All backgrounds	$8.0^{+3.1}_{-2.3}$	$11.2^{+1.6}_{-1.4}$	$18.1^{+5.6}_{-3.8}$	$37.3^{+6.6}_{-4.7}$
m_H = 125 GeV	2.4 ±0.4	4.6 ±0.5	5.9 ±0.7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

Kinematical - MELA discriminant

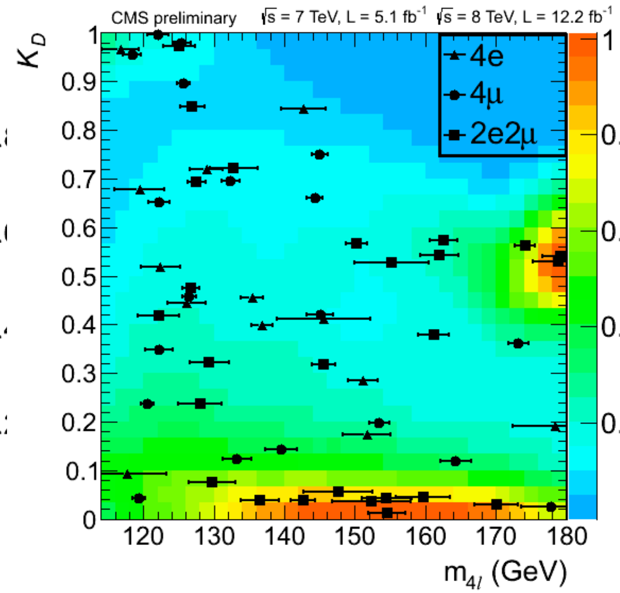
$100 < m_{4l} < 180$ GeV

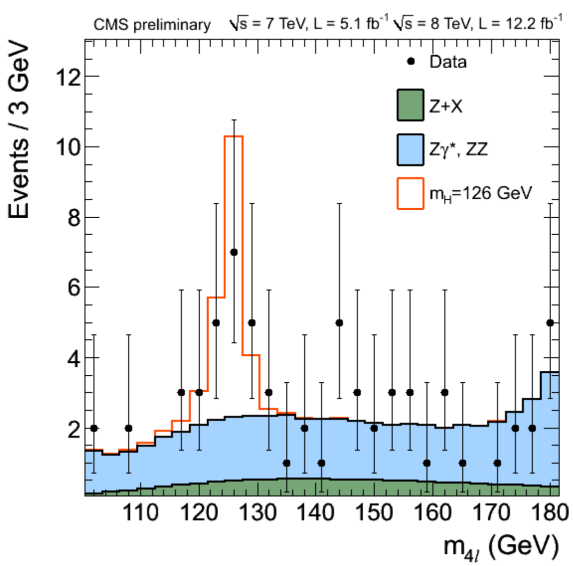


Data w.r.t 126 GeV Higgs Exp.



Data w.r.t. background exp.

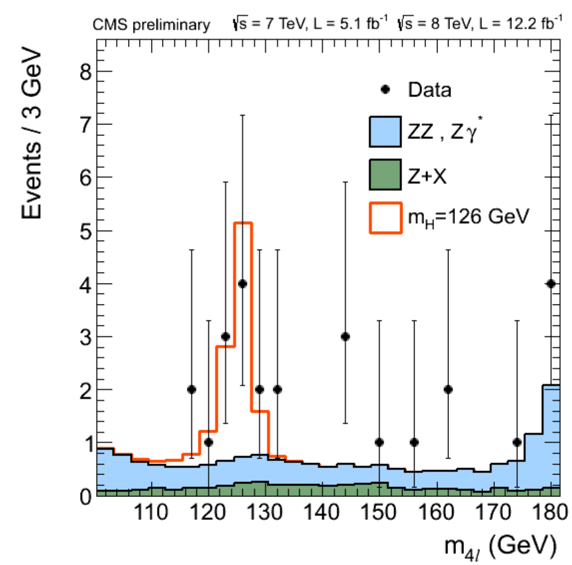




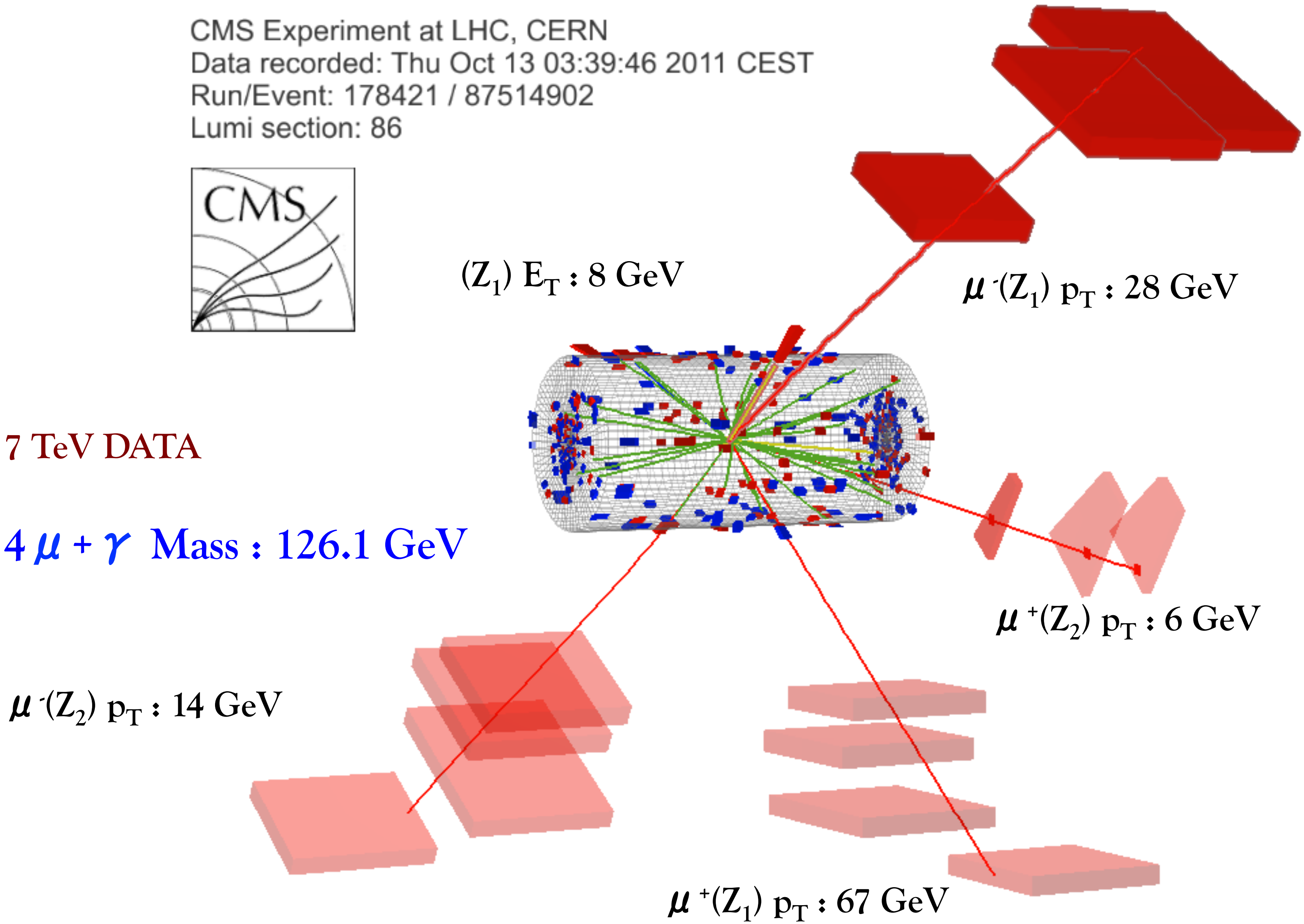
Enrich the signal content

MELA > 0.5

Cut value chosen such that signal probability > background probability



Candidate event

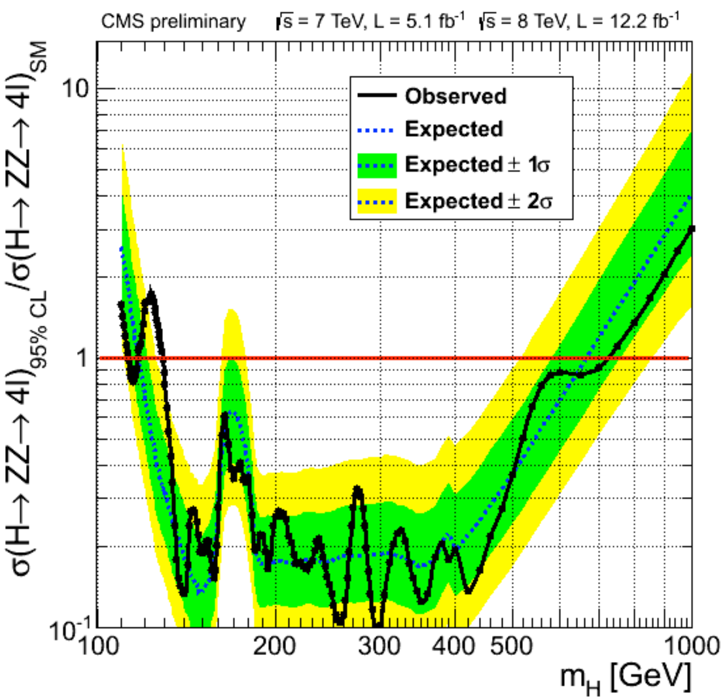


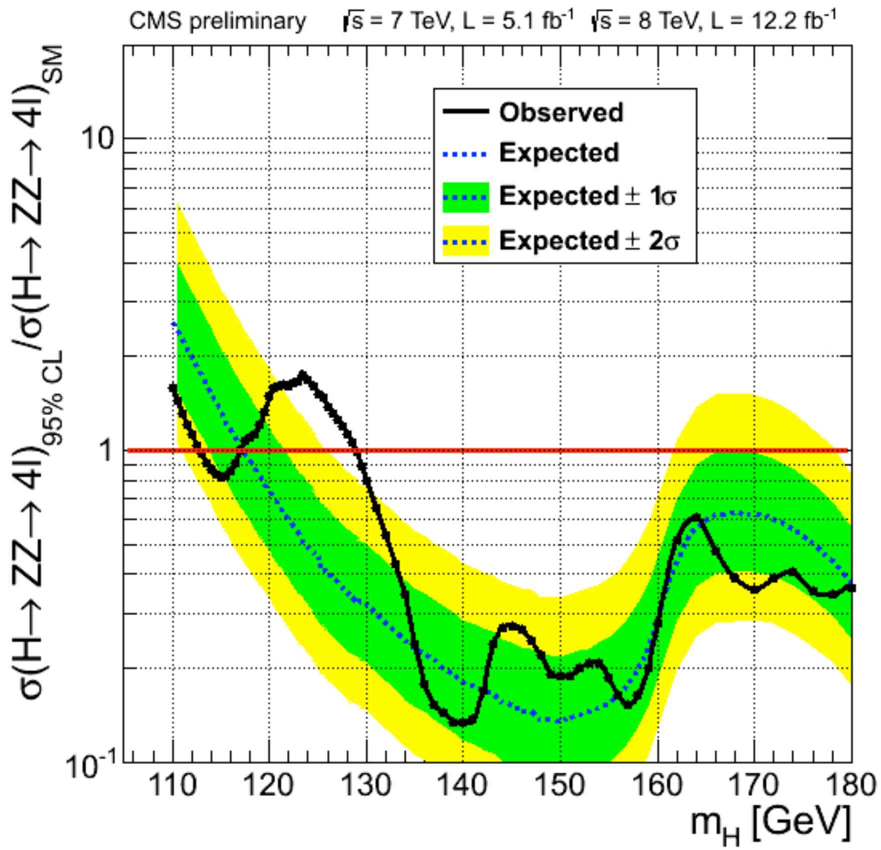
Statistical treat. : exclusion limits

2D pdf built (KD,m4l):

$$\mathcal{P}_{\rm sig}(m_{4\ell},{\rm KD}) = \mathcal{P}^{\rm 1D}_{\rm sig}(m_{4\ell}) \times \mathcal{P}_{\rm sig}(KD|m_{4\ell})$$

- ✓ Test statistic: profile likelihood ratio
- ✓ nuisance parameters included
- ✓ CL_s method for exclusion limit

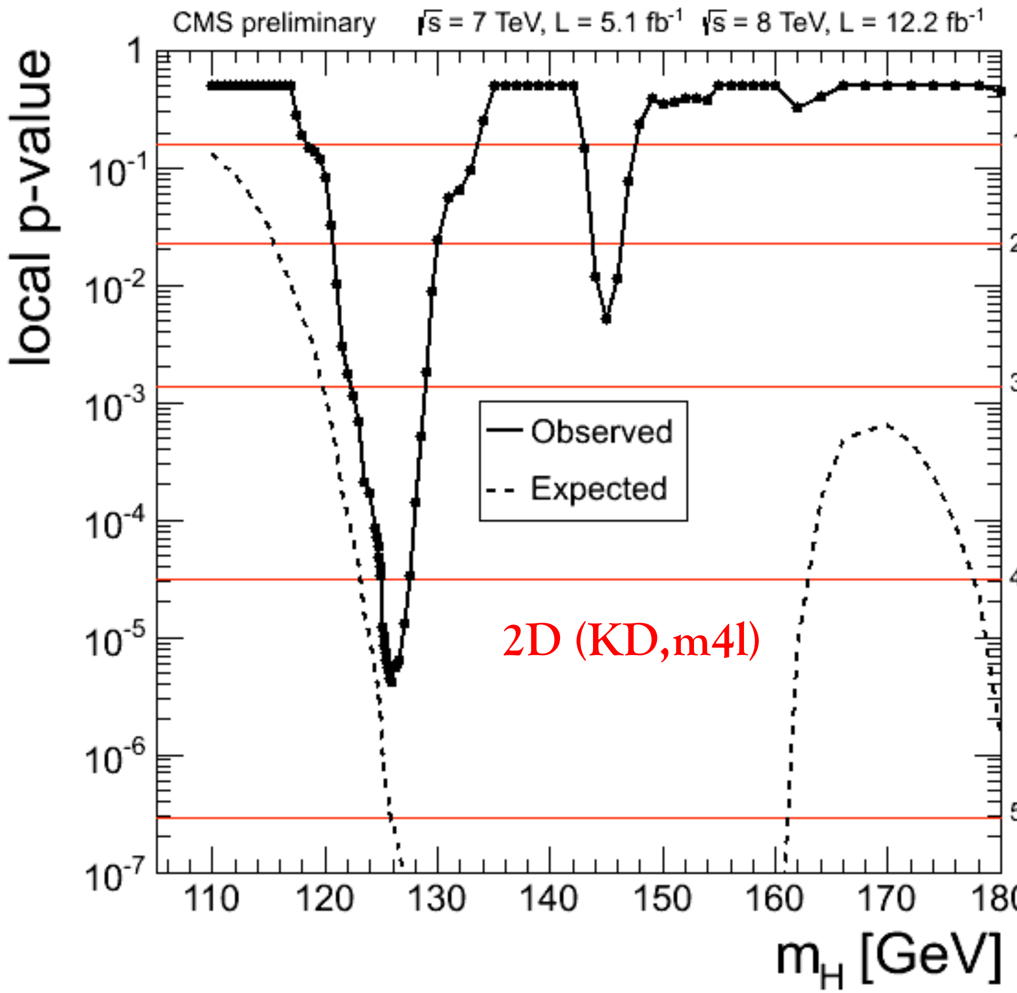




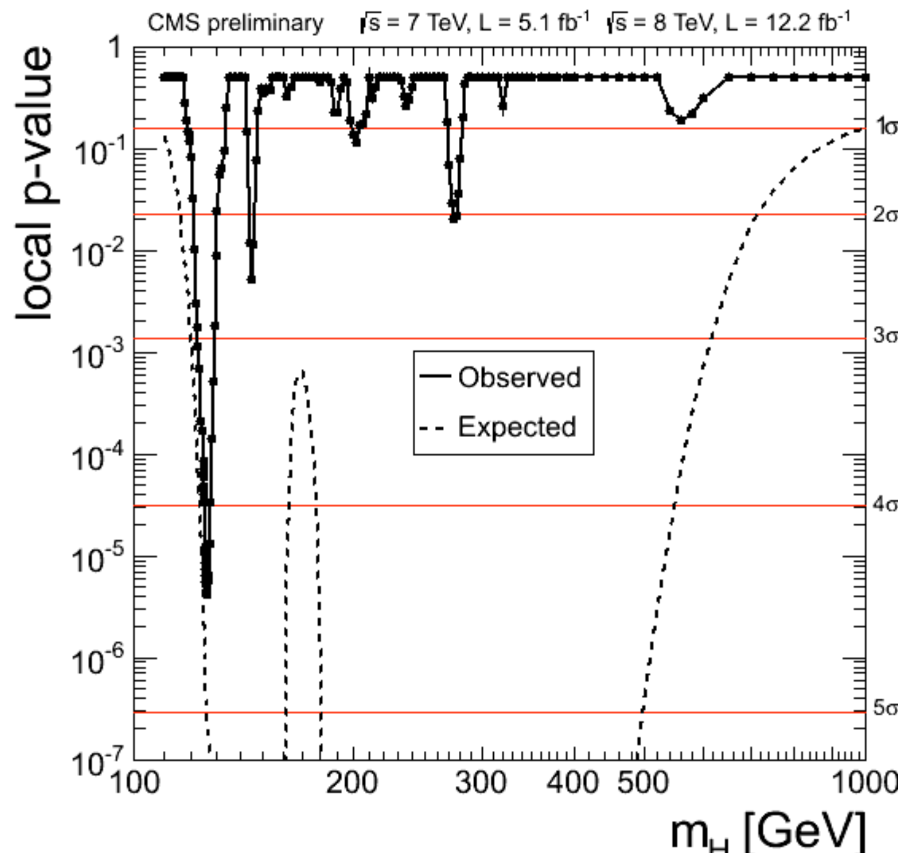
Observed limit:

95% CL exclusion in ranges 113-116 and 129–720 GeV

Statistical treat. : local significance



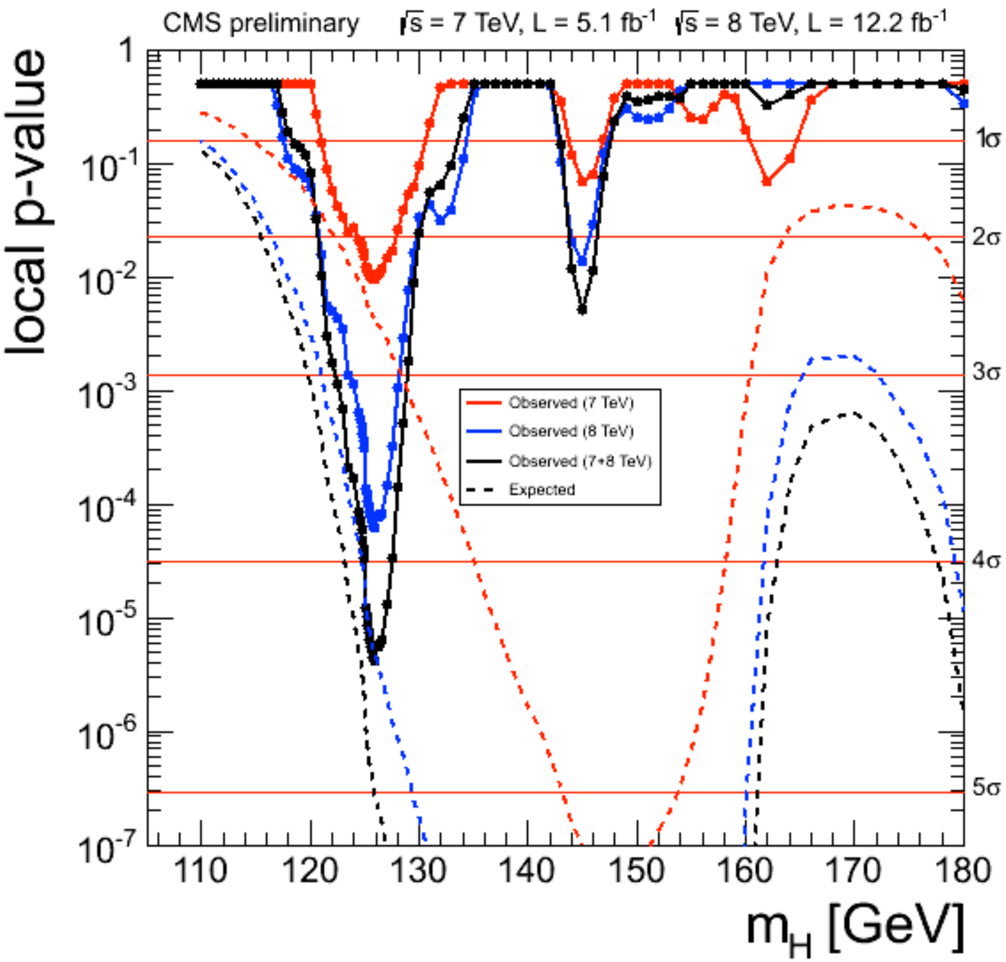
p-value: probability that the background can fluctuate to give an excess of events equal or larger than what observed

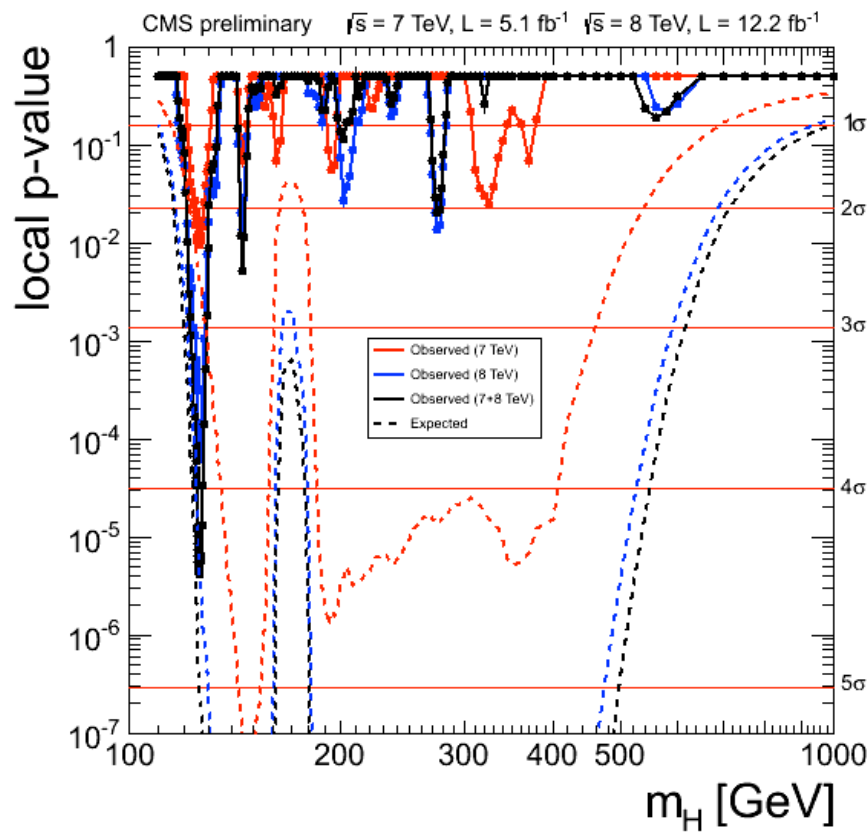


Minimum observed p-value ≈ 4.5σ (5.0σ expected)

(ICHEP was: 3.2 σ)

4l analysis: 2011 vs 2012



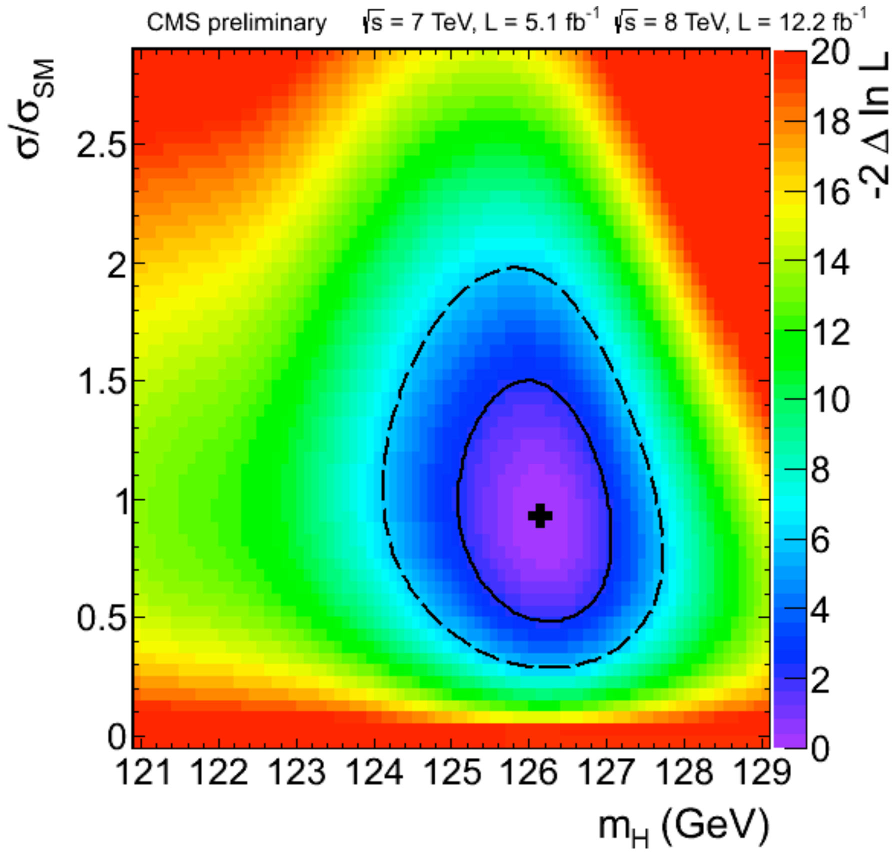


Excess consistent in 2011 and 2012

Mass measurement

3D pdf built (KD,m4l, EBE):

$$\mathcal{P}_{\text{sig}}(m_{4\ell}, EBE, KD) = \mathcal{P}^{1D}_{\text{sig}}(m_{4\ell}) \times \mathcal{P}_{\text{sig}}(EBE|m_{4\ell}) \times \mathcal{P}_{\text{sig}}(KD|m_{4\ell})$$



- Event by Event mass error (EBE) included
  - from muon track fit error matrix
  - from electron momentum error
- 3% of better significance by using the EBE
- 10% improvement on error on m_X

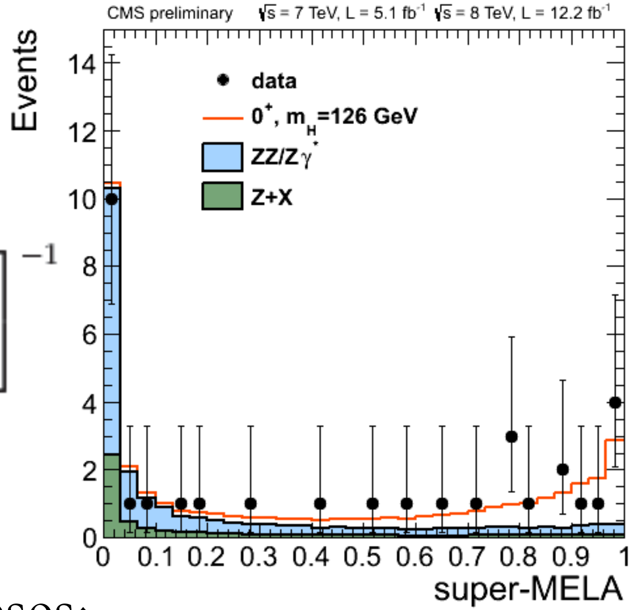
The combined best-fit mass is

m_X = 125.8 +/- 0.4 (stat) +/- 0.4 (syst) GeV

J^{CP} studies and results

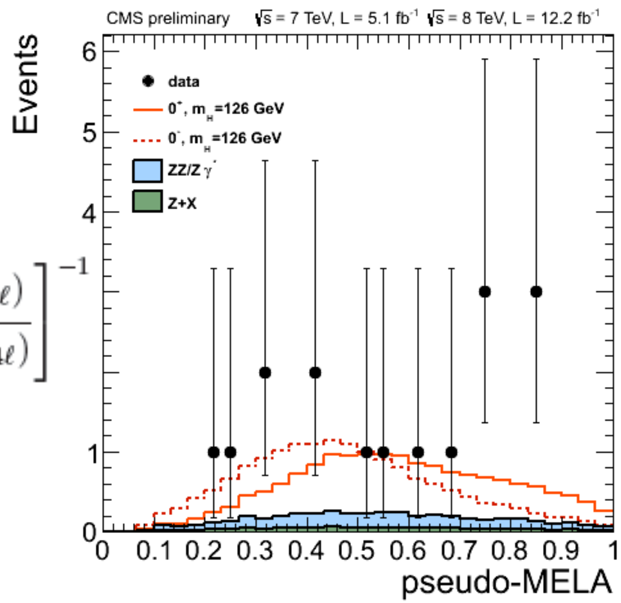
To improve the signal to ZZ discrimination:

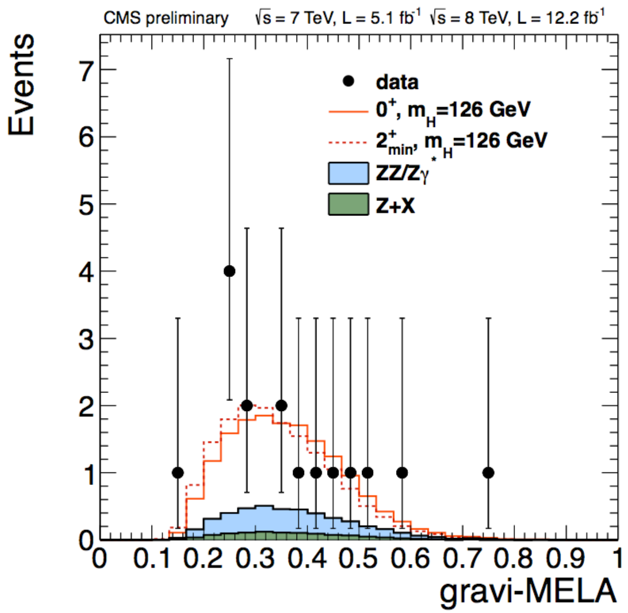
$$SMD = \frac{\mathcal{P}_{\rm sig}}{\mathcal{P}_{\rm sig} + \mathcal{P}_{\rm bkg}} = \left[1 + \frac{\mathcal{P}_{\rm bkg}(m_1, m_2, \vec{\Omega}|m_{4\ell}) \times \mathcal{P}_{\rm bkg}(m_{4\ell})}{\mathcal{P}_{\rm sig}(m_1, m_2, \vec{\Omega}|m_{4\ell}) \times \mathcal{P}_{sig}(m_{4\ell})}\right]^{-1}$$



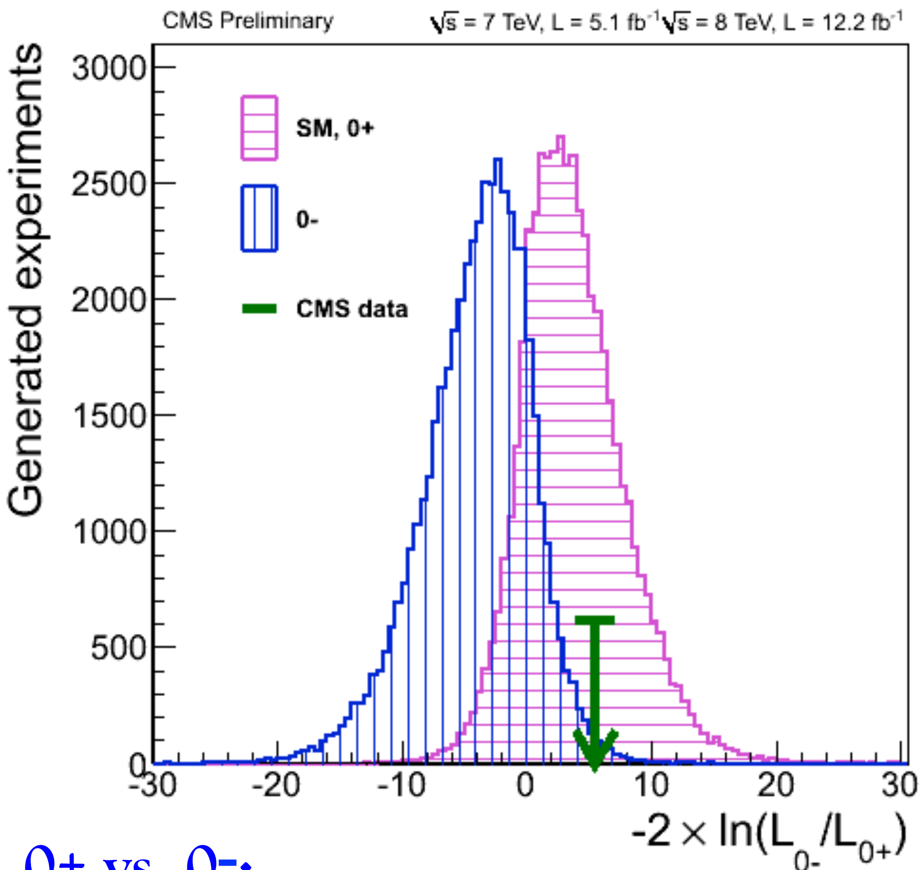
To separate J^{CP} 0^+ from 0^- and 2^+ hypotheses:

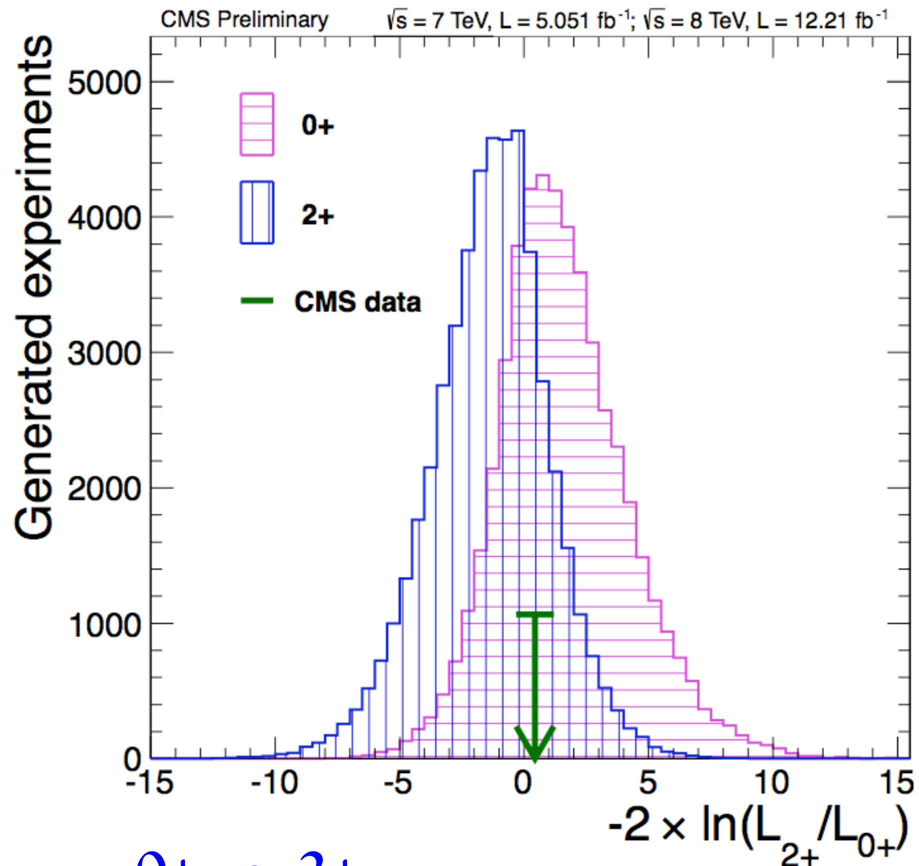
$$\mathcal{D}_{J^P} = \frac{\mathcal{P}_{\rm SM}}{\mathcal{P}_{\rm SM} + \mathcal{P}_{J^P}} = \left[1 + \frac{\mathcal{P}_{J^P}(m_1, m_2, \vec{\Omega}|m_{4\ell})}{\mathcal{P}_{\rm SM}(m_1, m_2, \vec{\Omega}|m_{4\ell})}\right]^{-1}$$





Statistical analysis: J^{CP}





0+ vs. 0-:

Expected separation: 2.0 σ

Observed:

- 0^- is consistent with observation within 2.45 σ (<3% using CLs)
- 0^+ is within 0.5 σ

0+ vs. 2+:

More data needed to exploit 0+ vs 2+ separation

Conclusions

- ✓ **4.5 σ** of local significance for the m(X) ≈126 GeV
- ✓ measured mass: **125.8 ± 0.4(stat.) ± 0.4(syst.) GeV**
- ✓ 0^- is consistent with observation within 2.45 σ (2.4% using CLs) while **0+ is consistent within 0.5** σ
- ✓ upper limits at 95% confidence level exclude the standard model Higgs boson in the range 113-116 and 129–720 GeV (together with 2l2τ)
- ✓ **Agreement with the SM prediction in the whole mass range**

Backup

H→ZZ→2l2τ in a nutshell

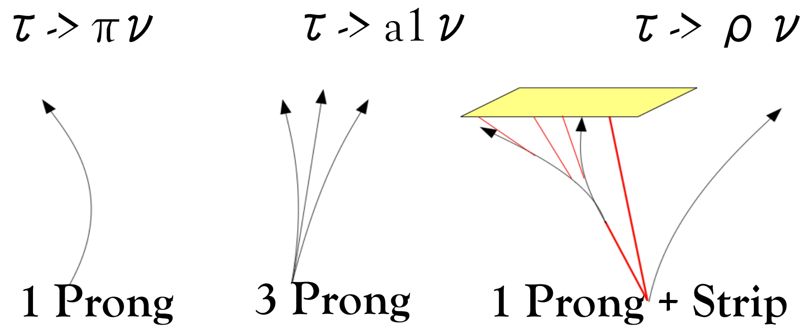
- Signature: 8 final states

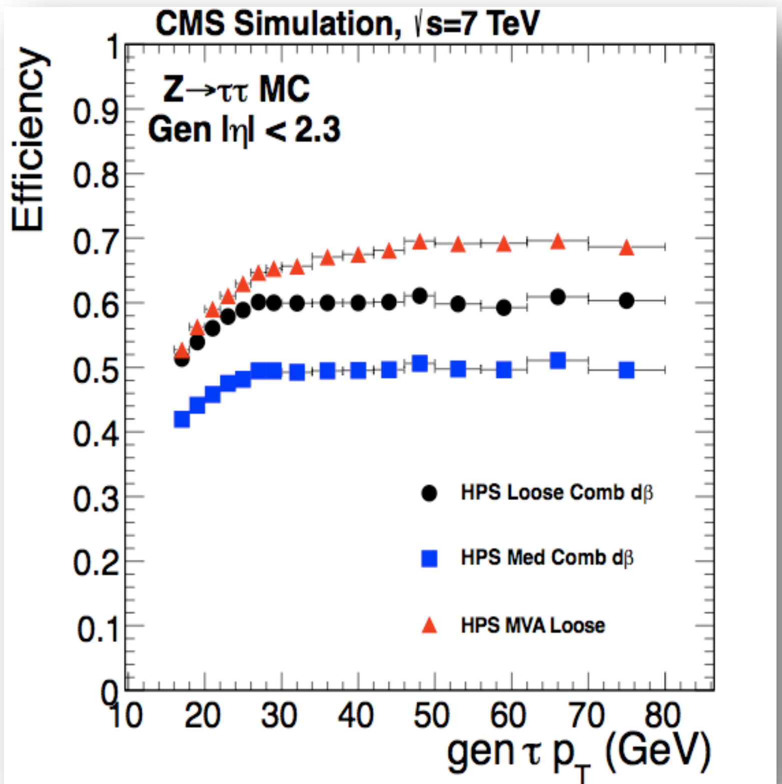
  μμττ, μμμτ, μμeτ, eeττ, eeeτ, eeμτ, μμμe, eeeμ

- Backgrounds:
  - ZZ, WZ and Z associated with additional jets
- Selection Strategy:
  - Lepton trigger as in 4l analysis 2011/2012
  - Leptons Identification
  - τ reconstruction via Hadron Plus Strip algorithm
  - Leading Z selection (mass close to nominal Z mass)
  - Mass window for leading and sub-leading Z bosons
  - e/μ/τ Isolation

Tau reconstruction and identification

- Tau identification – Hadron Plus Strip algorithm:
  - Reconstructs individual decay modes
  - Charged hadrons combined with electromagnetic objects
    - EM strips used to account for material effects
- Tau isolation:
  - Multivariate discriminator using sum of energy deposits in dR rings around the tau





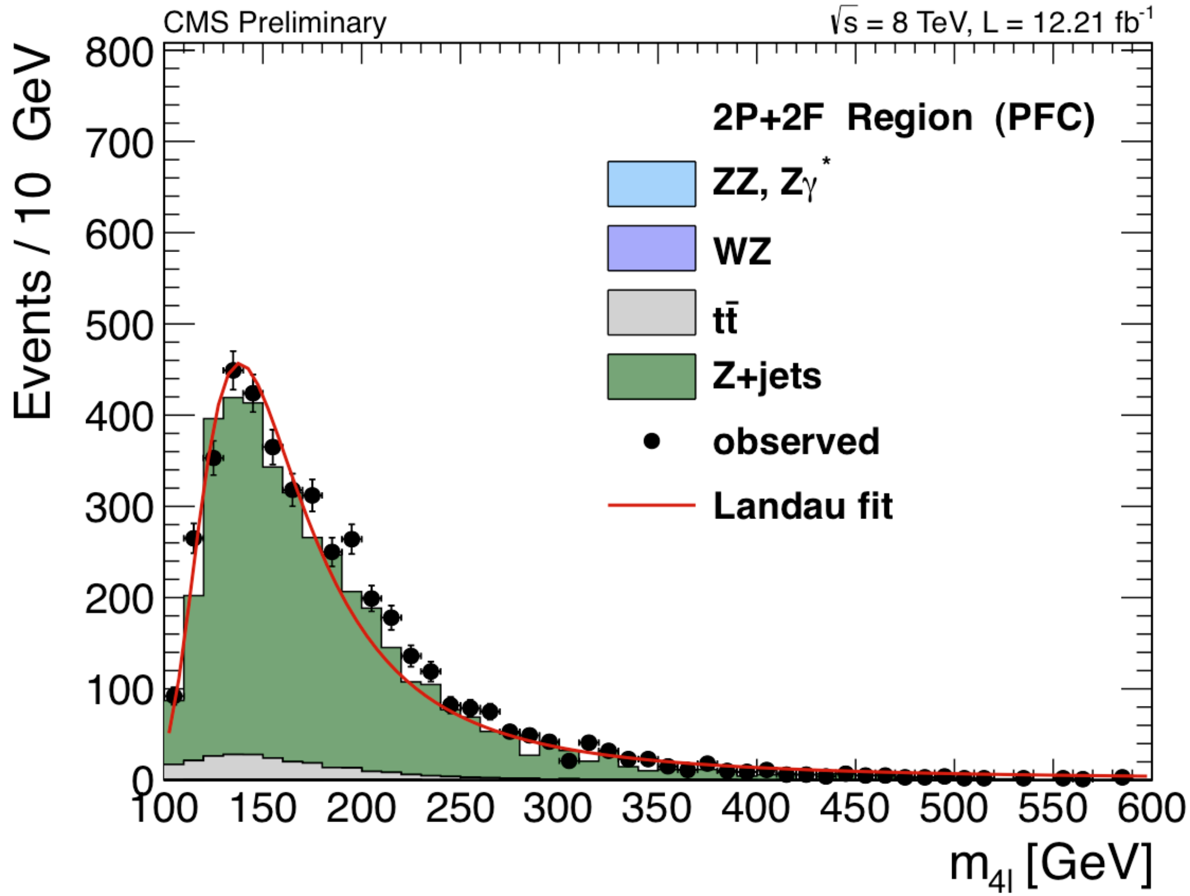
Background control: closure test

Validation in data using "wrong flavors & charges" events (WFC) in control region 2P+2F

- Extrapolation to WFC signal region

The difference between the predicted and the observed n. events in WFC signal region is taken as the **measurement of the uncertainty of the method**

TOTAL uncertainty <50% dominated by the statistical uncertainty in signal region (30%)



Background control: uncertainty

Sources:

- Statistical uncertainty due to the limited size of the samples in the control regions where we measure and where we apply the fake ratio method

- Different compositions of reducible background processes (DY, tt¯, WZ, Zg()) in the region where we measure and where we apply the fake ratio method,

- Choice of the functional form for the m4l shape that is used to extrapolate from the full range of the invariant m4l mass to the range of interest.

Estimate the systematic uncertainty for the prediction method using the MC closure test,

Estimate directly the systematic uncertainty for the prediction method using the "orthogonal"
4l data samples with the "wrong combination of charge and flavour".

Systematic uncertainties

- Luminosity:
  - 2.2% @ 7 TeV, 5% @ 8 TeV
- Signal
  - BR: 2%, mH independent
  - Acceptance: 0.1-0.2%
  - PDF: 2%, mH independent
  - No HqT re-weighting: <2% (signal acceptance effect)
  - High-mass line shape: up to ~30%
  - Data to MC correction factors uncertainties

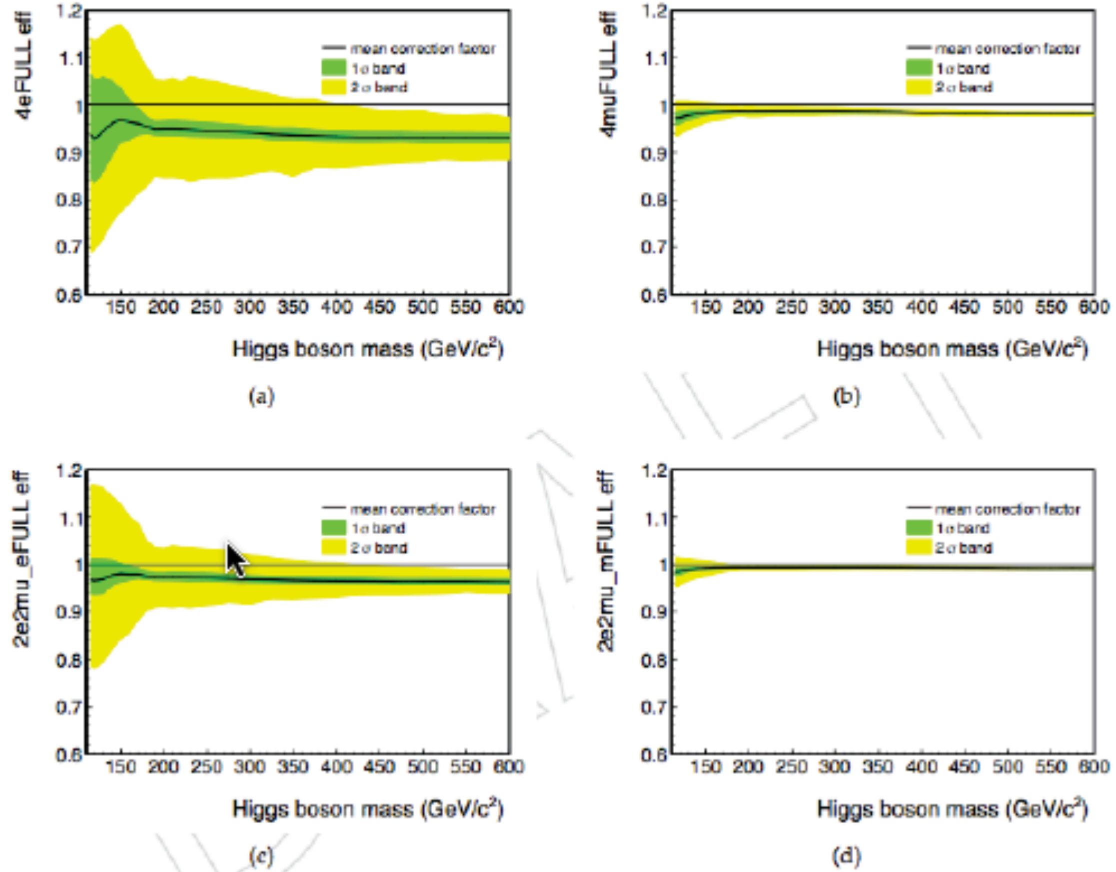


Figure 48: Data-to-MC corrections and their uncertainties for $4e$ (a), 4μ (b) and $2e2\mu$ (c-d) events associated with the electron (left) and muon (right) RECO/ID/ISO/SIP full selection efficiencies.

Systematic uncertainties (2)

- qq/gg→ZZ→4l
  - PDF and QCD scale (plots)
- Reducible backgrounds
  - Statistical (control region (CR) finite size)
    - ~5-10%
  - Different composition of backgrounds in CR and signal region (SR)
    - <40%
  - Choice of the functional form for the m(4l) shape of the red. Background
    - ~10-15%
  - Total: ~50%

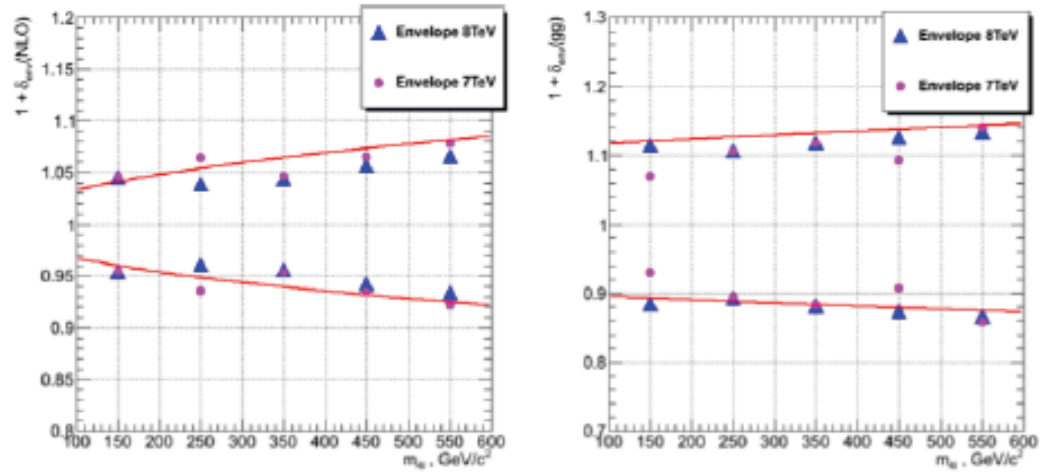


Figure 65: PDF+α_s uncertainties for (left) $pp \to ZZ \to 4\ell$ at NLO and (right) $gg \to ZZ \to 4\ell$ processes. The points are evaluated uncertainties. The curves are the fit systematic error $\kappa(m_{4\ell})$ to be used in the statistical analysis. These errors are driven by two independent nuisance parameters *pdf_qqbar* for $pp \to ZZ \to 4\ell$ at NLO and *pdf_gg* for $gg \to ZZ \to 4\ell$.

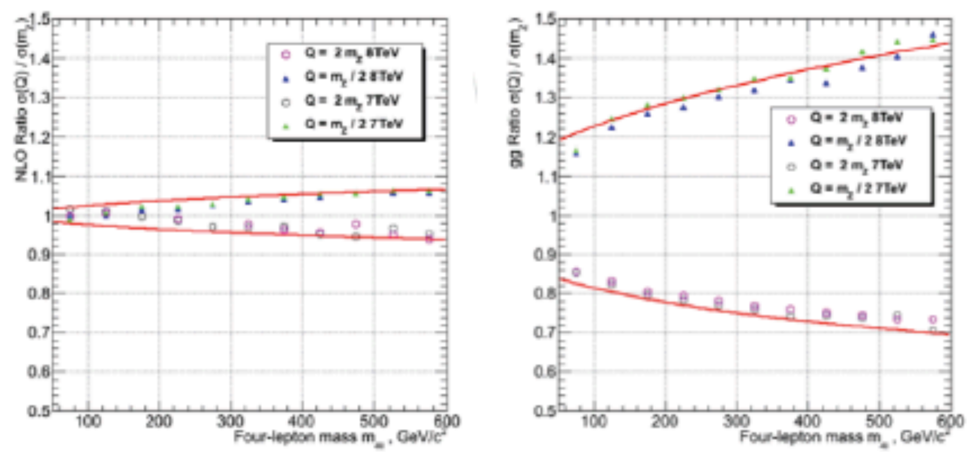
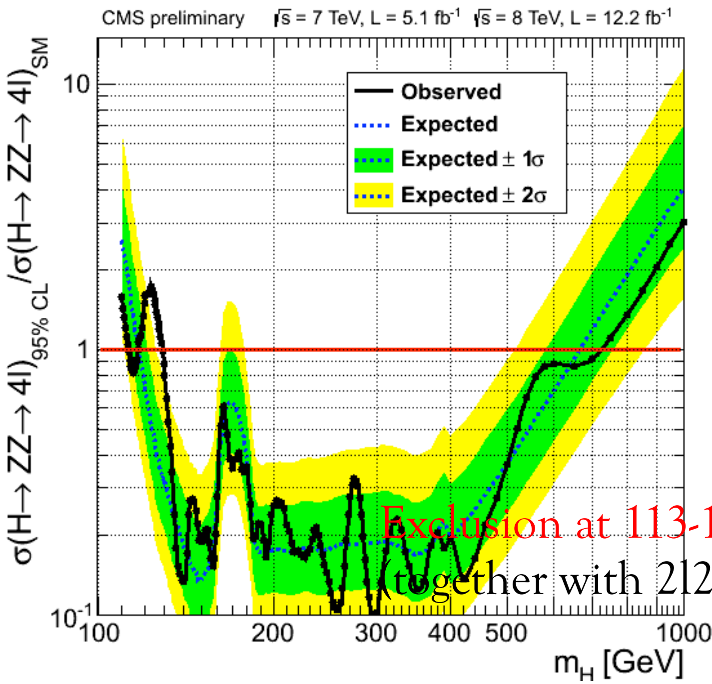
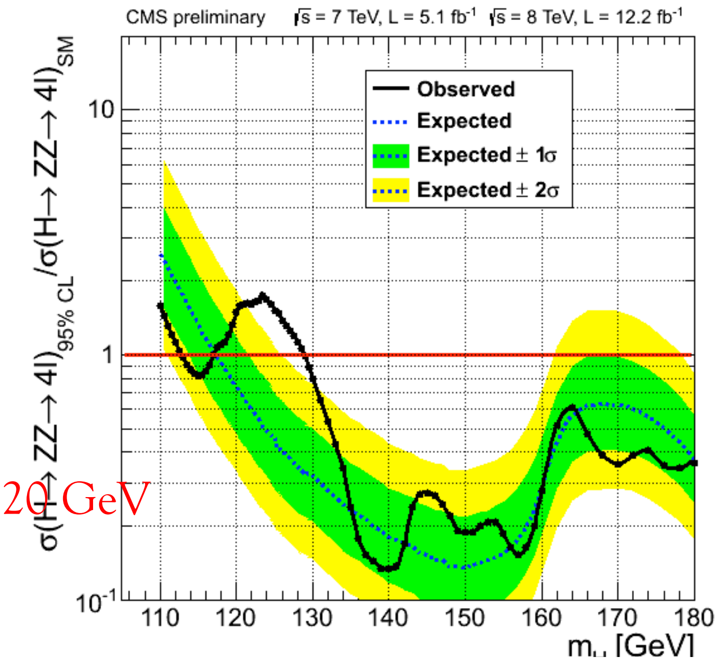


Figure 66: QCD scale uncertainties for (left) $pp \to ZZ \to 4\ell$ at NLO and (right) $gg \to ZZ \to 4\ell$ processes. The points are evaluated uncertainties. The curves are the fit systematic error $\kappa(m_{4\ell})$ to be used in the statistical analysis. These errors are driven by two independent nuisance parameters *QCDscale_VV* for $pp \to ZZ \to 4\ell$ at NLO and *QCDscale_ggVV* for $gg \to ZZ \to 4\ell$.

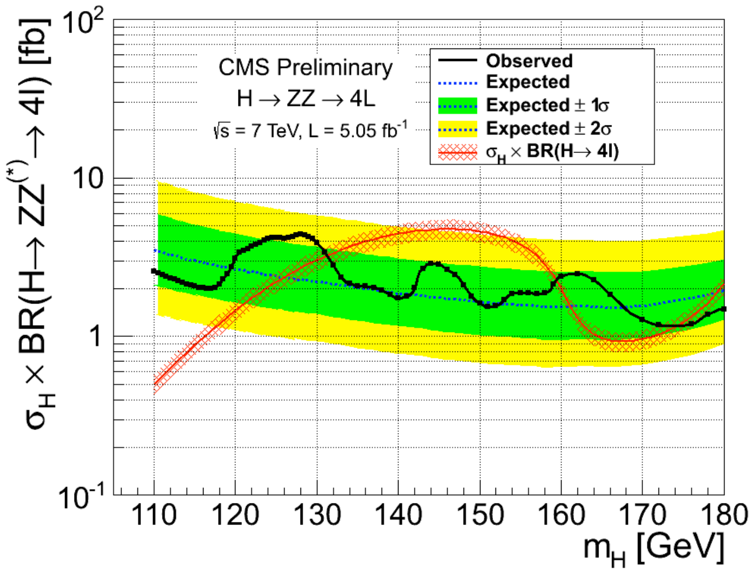
4l analysis: exclusion limits

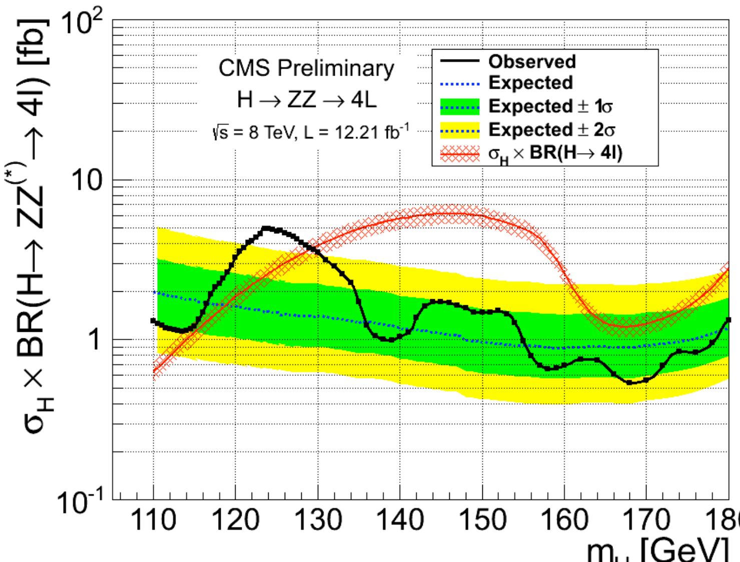




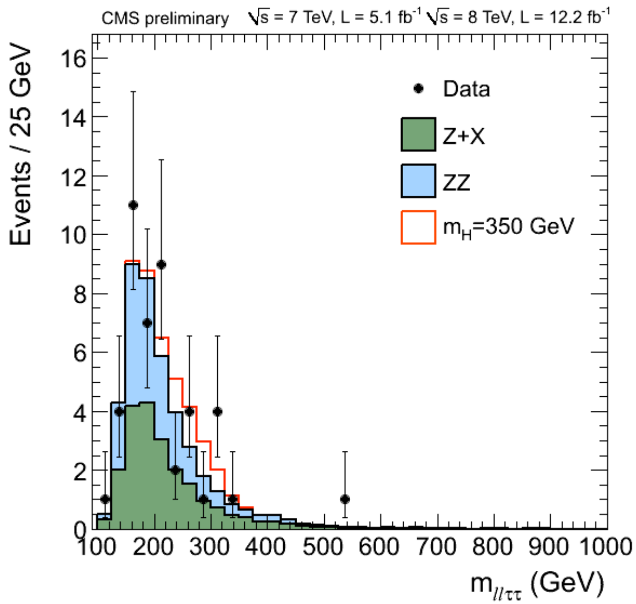
Exclusion at 113-116 and 129–720 GeV

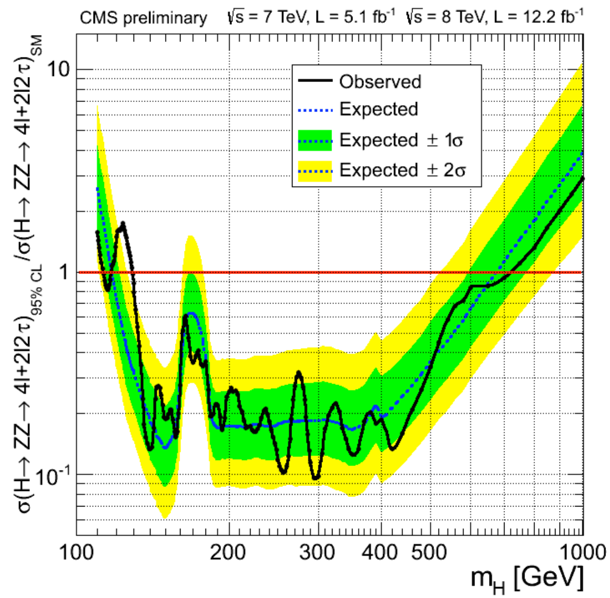
(together with 2l2tau)

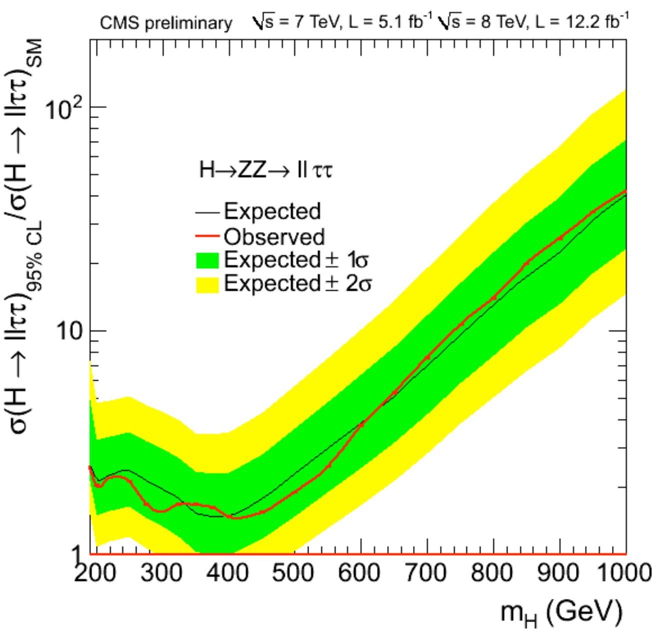


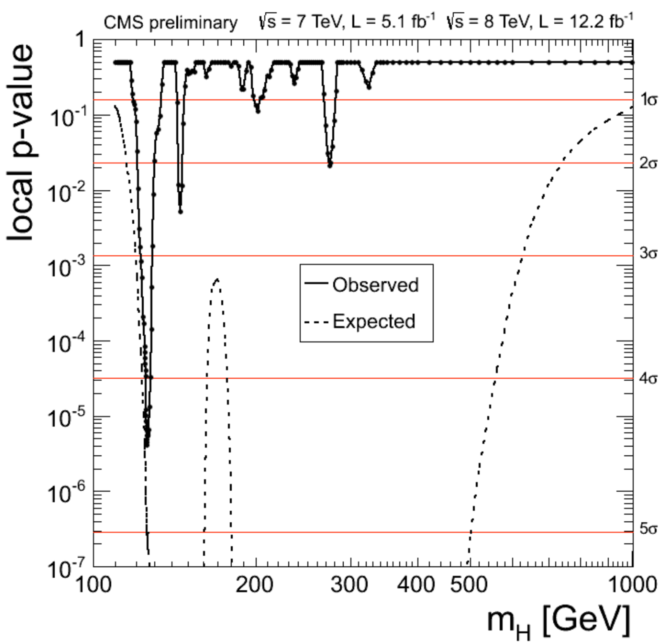


4l + 2l2tau combined

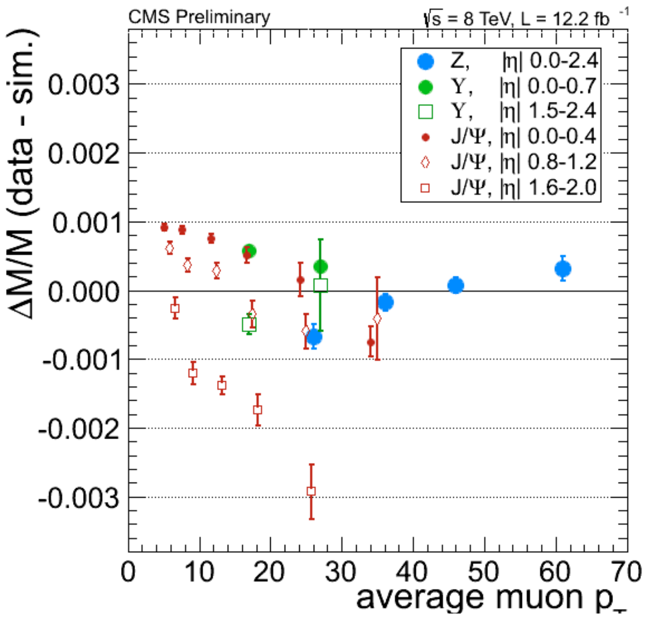


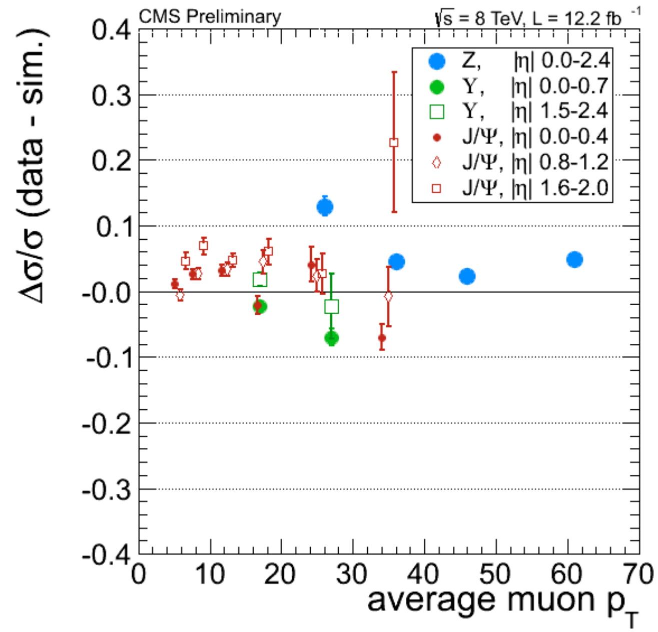




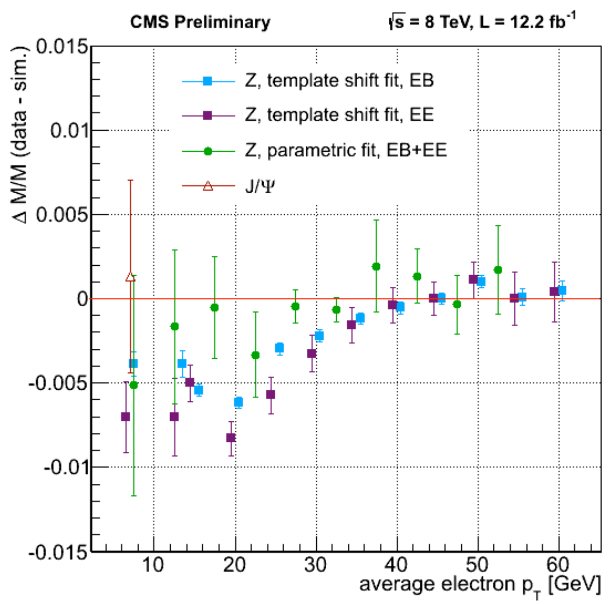


Mass measurements:

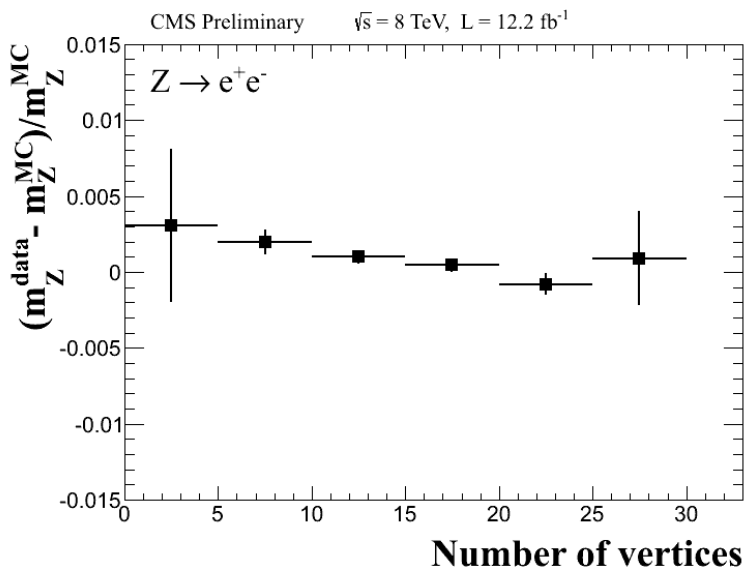




Scale

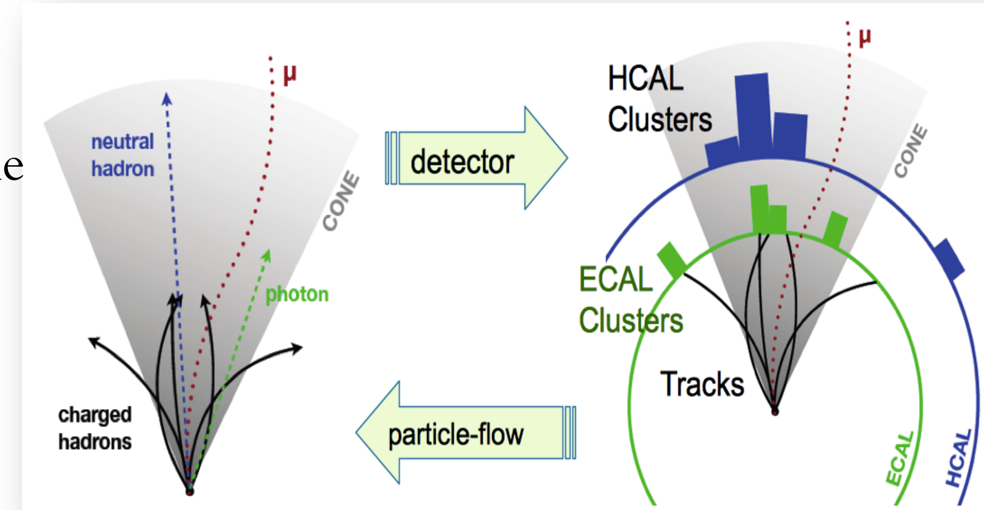


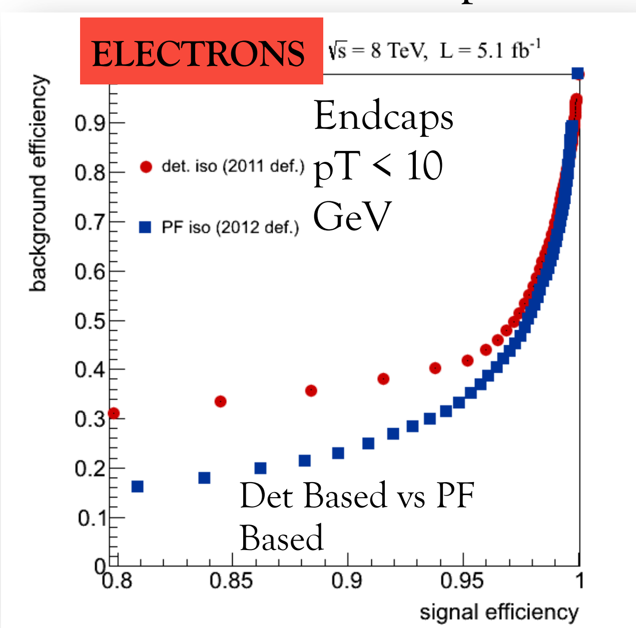
Resolution



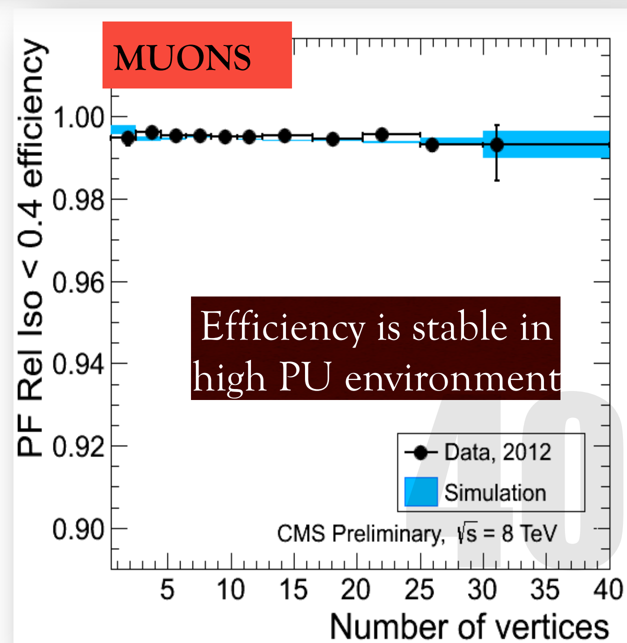
Particle Flow isolation

- Created by summing energy deposits from charged hadrons, neutral hadrons and photons in $\Delta R=0.4$ cone around the lepton
  - Avoids double counting of the energy deposit in the calorimeters from charged particles
  - Automatic footprint removal





- Pile-up contribution:
  - Negligible for charged hadrons from primary vertex
  - Neutral contribution corrected using the average energy density (ρ) from the pile-up and UE (FastJET algorithm)



M_{Z1}, M_{Z2}

For m(4l) = 121.5..130.5 GeV

	Exp. Bkg	m_H = 126	Obs
4e	1.25	2.20	3
4μ	2.09	4.26	6
2e2μ	3.14	5.97	8
Total	6.48	12.43	17