Search for the Standard Model Higgs boson at CMS in the 4-lepton channel

ICFP 2012

Marco Meneghelli

for the CMS collaboration

Phys. Rev. Lett. **108**, 111804 (2012)

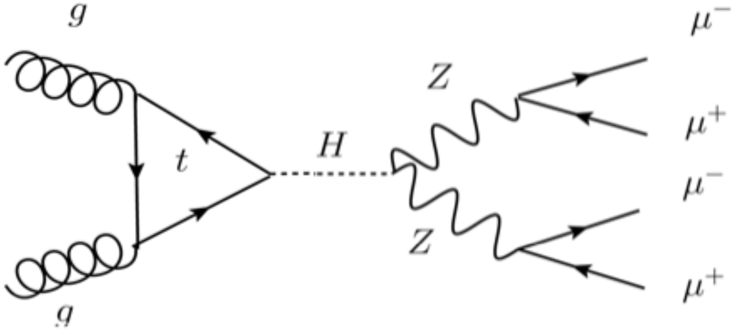
Outline

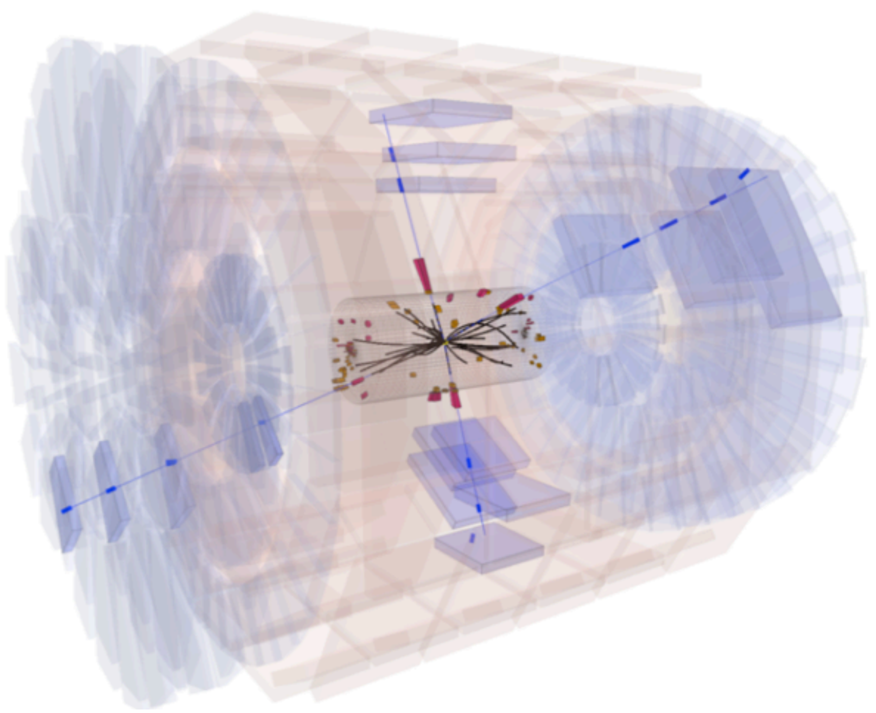
- The 4-lepton channel
- Data and Montecarlo
- Event selection
- Lepton identification
- Background evaluation and control
- Systematics
- Results (Final distributions and statistical analysis)

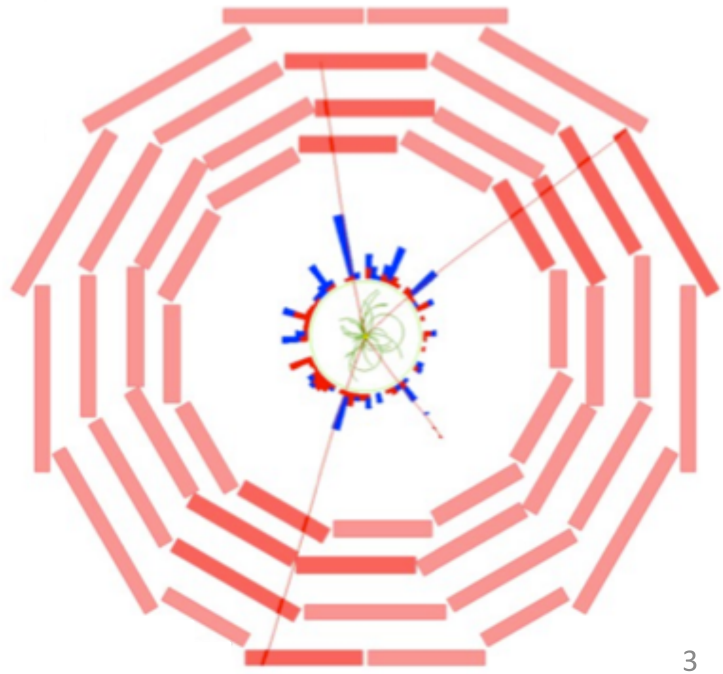
- *Ref.* Phys. Rev. Lett. **108**, 111804 (2012)

The 4-lepton channel

- Decay channel: $H \rightarrow ZZ^* \rightarrow l_1^+ l_1^- l_2^+ l_2^-$ $l_i = \mu, e$
- **Experimental signature:**
  4 isolated leptons coming from the Primary Interaction Vertex

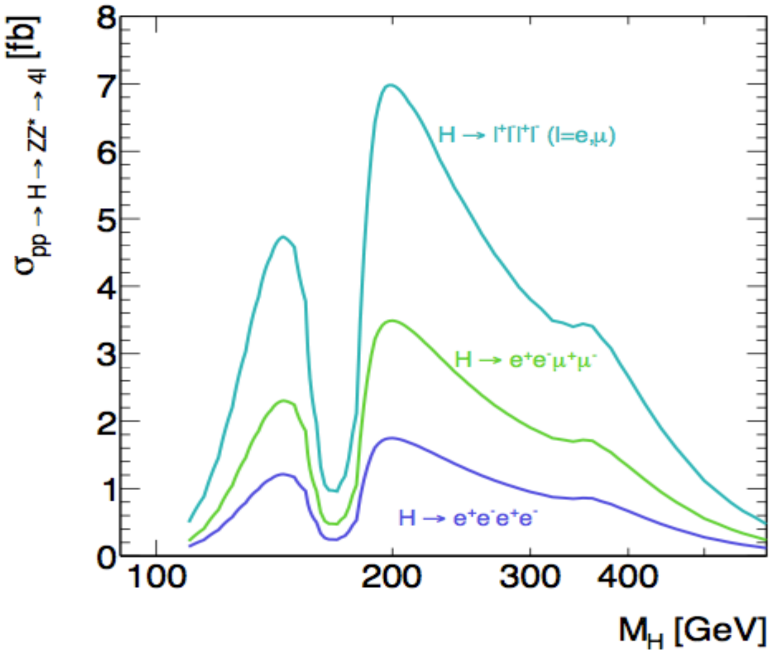


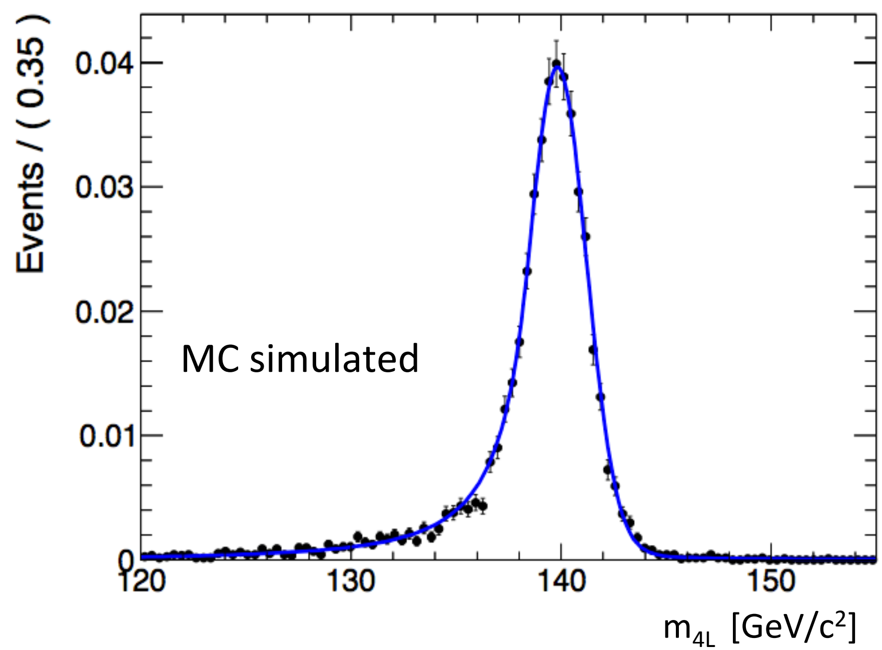




The 4-lepton channel

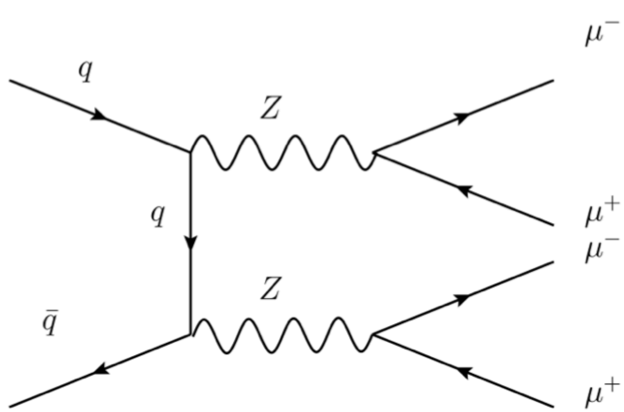
- Relatively small signal cross section: 0.5 – 7 fb
- Very clean signature and a resonance peak in the 4-lepton mass distribution with excellent resolution



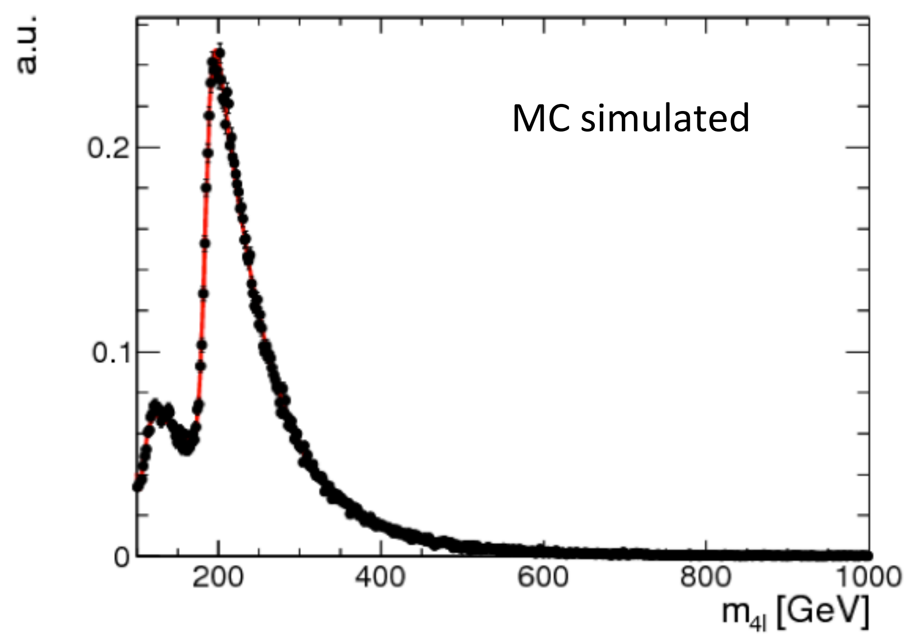


The 4-lepton channel

- Ratio signal:background almost 1:1
  - Irreducible background: the ZZ continuum   yield: ~14 fb
  - Reducible backgrounds: Z+jets, tt, WZ, QCD   yield: ~1.1 fb
- Good sensitivity for a wide mass range 110 < m_H < 600 GeV/c^2

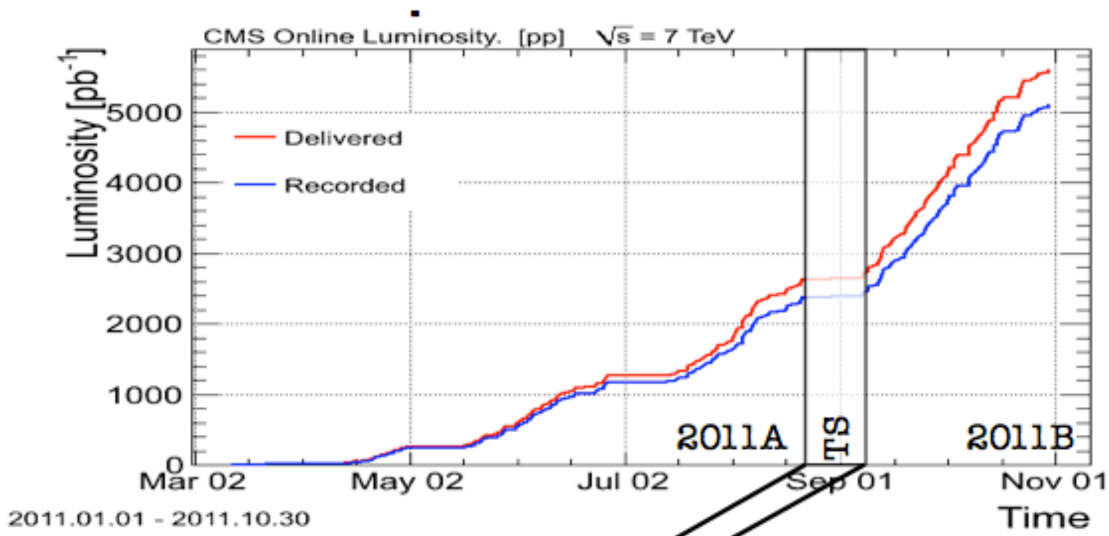


qq →ZZ→4L



Experimental Data

- Available statistics: **4.7 fb^{-1}** collected during 2010 and 2011



2010-2011A periods

HLT: 2×10^{32} - $>2\times10^{33}$

RunRange: [136033-149442] **April11 ReReco**
[160329-168437] **July5 ReReco**
[170053-172619] **Aug5 ReReco**
[172620-173692] **Oct3 ReReco**

L = 2.2 fb^{-1}

2011B period

HLT: 3×10^{33}

RunRange: [175860-177051] **Prompt Reco**

L = 2.5 fb^{-1}

Montecarlo samples

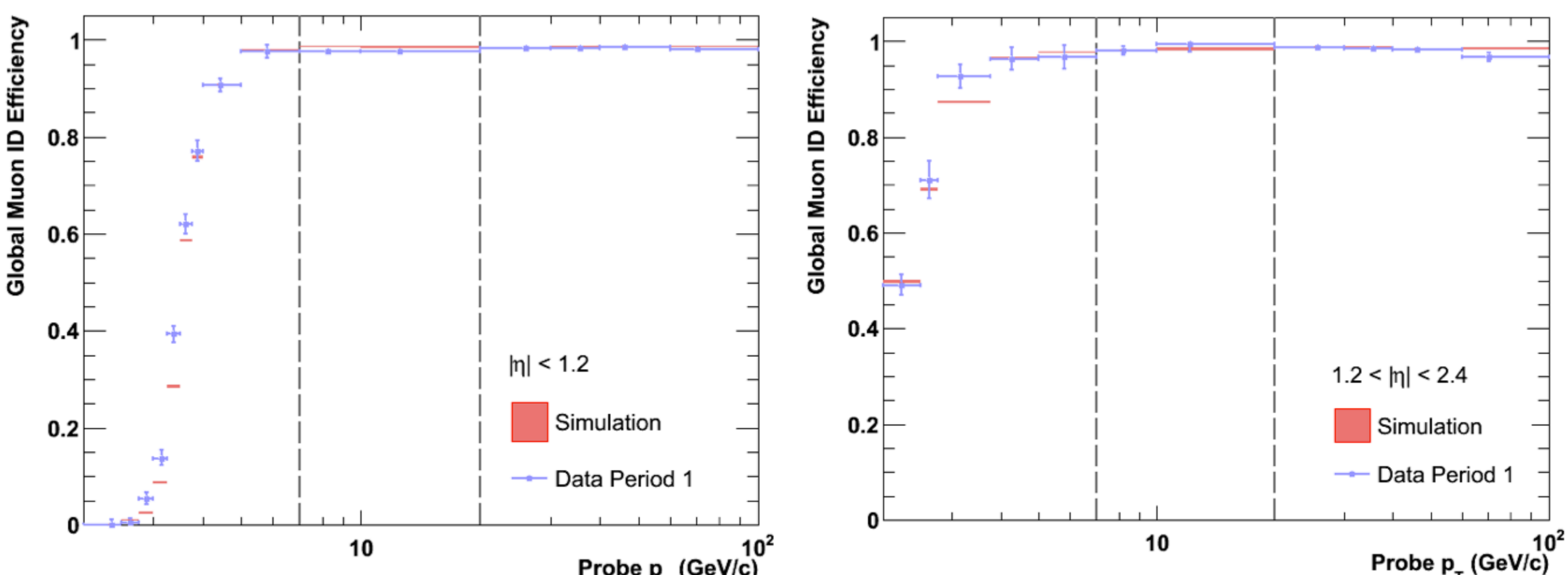
- Signal:
  - gg→H→ZZ→4l [m_H = 100 – 600 GeV/c^2] powheg
- Backgrounds:
  - qq→ZZ→4l powheg
  - gg→ZZ→4l gg2zz
  - tt →2l2nu2b powheg
  - Z+jets madgraph
  - W+jets madgraph
  - WZ , WW powheg
  - QCD pythia

Event selection

- **Trigger**
  - identification of same-flavor lepton pairs passing thresholds in p_t
- **Skim**
  - 2 reconstructed leptons above a given p_t, with an invariant mass > 40 GeV/c^2
- **Pre-selection**
  - identification of a first, on shell, Z boson (a opposite-sign, same-flavor lepton pair with mass close to 91 GeV/c^2)
  - a third lepton passing quality cuts
  - a fourth lepton passing quality cuts, with same-flavor opposite-sign wrt the third
  - choice of the best Z_2 as the opposite-sign same-flavor lepton pair with the highest p_t
- **Selection**
  - Cut on lepton isolation
  - Cut on significance of impact parameter
  - Cuts on kinematics

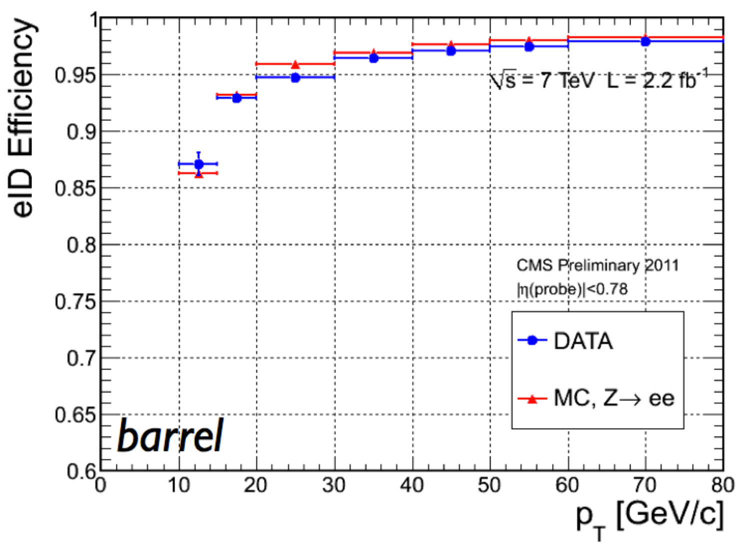
Lepton identification - Muons

- Algorithm: Global muon reconstruction
  - fit hits in the muon system and in the silicon central tracker
  - four-momentum is derived from the inner track
- ID cuts:
  - pt > 5 GeV/c   , | eta | < 2.4
  - tracker hits > 10
- Efficiencies are evaluated with Tag&Probe technique at the Z and J/ψ peaks
  - this technique exploits the known presence of two leptons in a resonance to computed almost-unbiased efficiencies

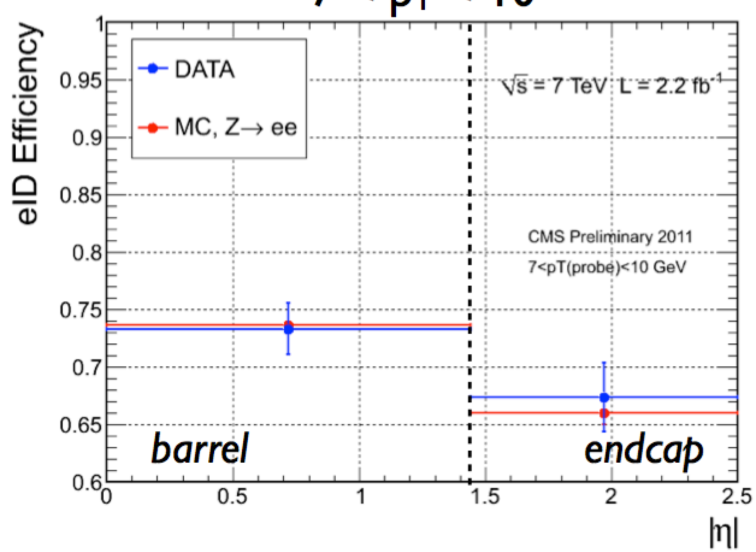


Lepton identification - Electrons

- Algorithm: Gaussian-Sum-Filter (gsf) electron
  - combines ECAL and tracker information, matching (super) clusters of energy in the ECAL with hits in the inner tracker
  - four momentum is given by the combination of ECAL and tracker info
- ID cuts:
  - pt > 7 GeV/c , | eta | < 2.5
  - tight electron selection to enhance purity → best s/b
  - conversions rejected cutting on expected missing inner hits
- Efficiencies are evaluated with Tag&Probe technique at the Z peak

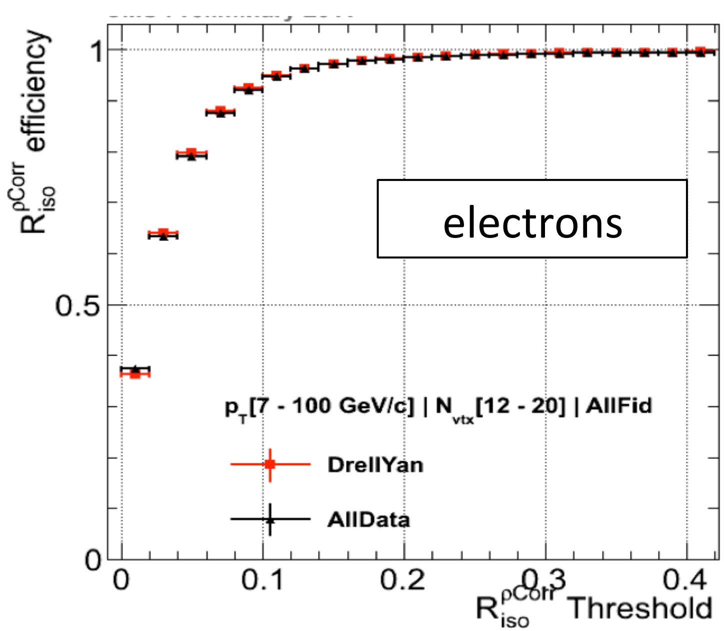


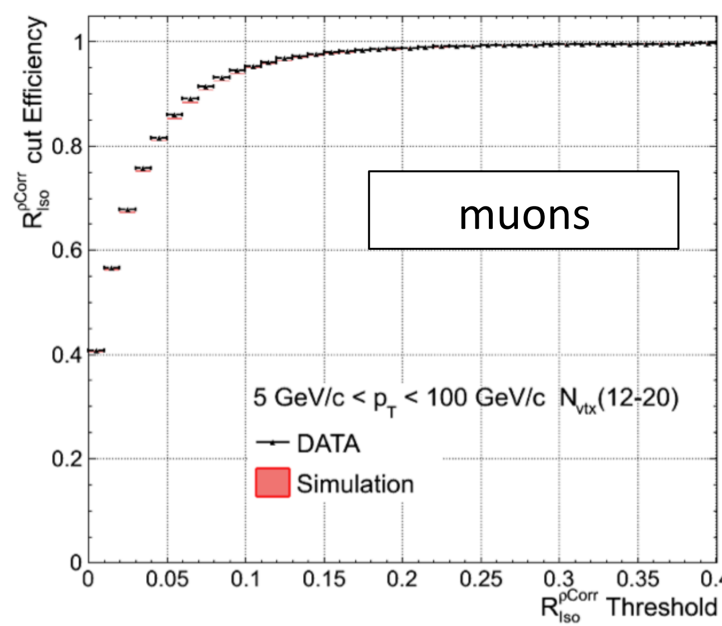
7 < p_T < 10



Lepton isolation

- The leptons used in the analysis are required to be isolated, i.e. non contained in a jet, in order to discriminate against QCD processes
- Lepton isolation is evaluated looking at the p_t of tracks or energy deposits in a geometrical cone around the track of each lepton:
  - lepton track isolation Iso_{track}: sum of p_t of tracks within a $\Delta R < 0.3$ cone
  - lepton ECAL/HCAL isolation Iso_{ECAL} Iso_{HCAL}: sum of ET from energy deposits within a $\Delta R < 0.3$ cone
  - a small "veto" cone around the lepton is excluded from the computation
  - Iso_{ECAL} Iso_{HCAL} are corrected for the effect of pile-up increasing
  - normalization to p_t → Final variable $R_{iso} = (Iso_{track} + Iso_{ECAL} + Iso_{HCAL})/p_t$
- Final cut is on the sum $R^i_{iso} + R^j_{iso}$ (< 0.35) over all the lepton couples *ij*

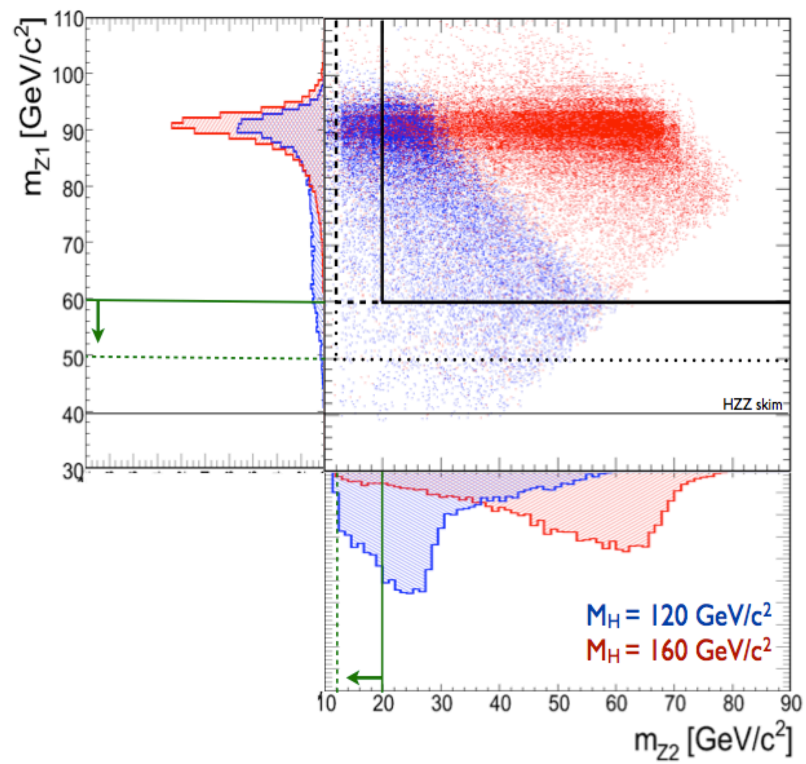




Primary leptons choice & Kinematics

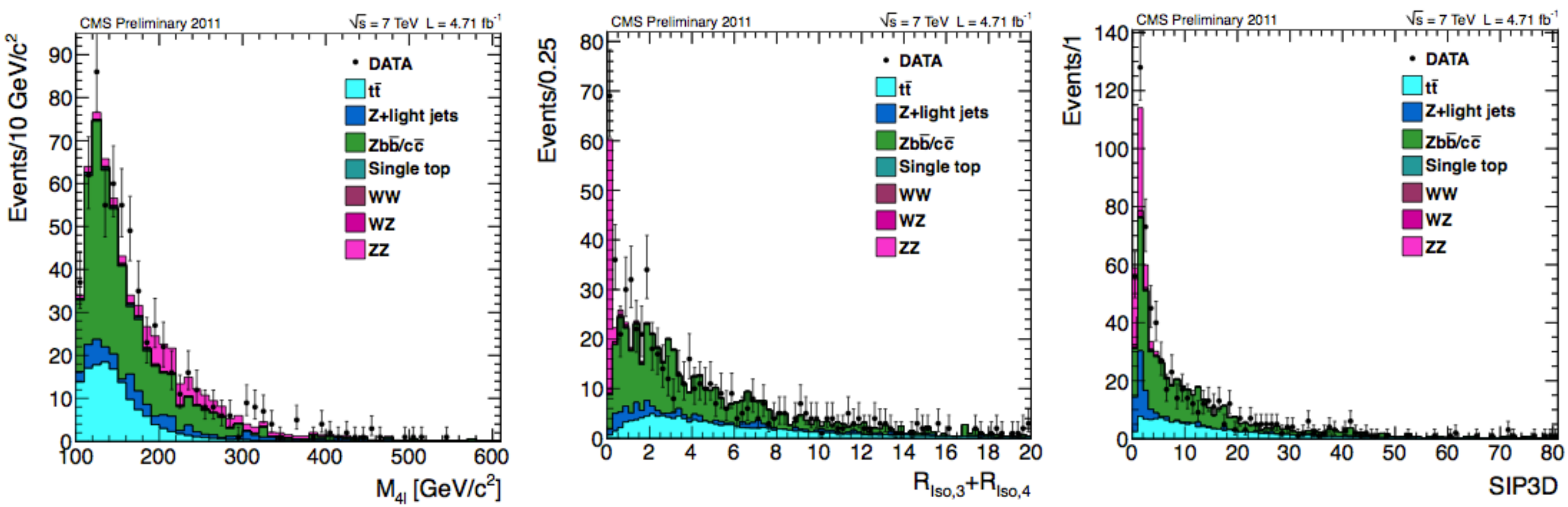
- All the leptons are required to come from a common primary vertex
- This allows to discriminate against reducible background (Zbb,Zcc,tt) by rejecting leptons coming from secondary vertices (b-jets)
- Cut on the significance of impact parameter (the distance of closest approach of lepton track with respect to the primary interaction vertex, normalized to its uncertainty): $|SIP_{3D}| = |IP_{3D}/s_{IP}| < 4$

- Kinematical cuts are posed to discriminate against QCD resonances:
  - $50 < m_{Z1} < 120\ GeV/c^2$
  - $12 < m_{Z2} < 120\ GeV/c^2$
  - $m_{4l} > 100\ GeV/c^2$



Event selection

- A comparison done at the early stage of selection, with relaxed cuts on flavor/sign, allow to see the difference between signal-like events and reducible background ones
- **Cuts on isolation and significance of IP are crucial to discriminate signal against reducible background**



Background control: irreducible

- ZZ continuum yield is directly estimated from MC

$$\left(\sigma_{NLO}^{q\bar{q}\to ZZ\to 4l} \times \varepsilon_{MC}^{q\bar{q}\to ZZ\to 4l} + \sigma_{LO}^{gg\to ZZ\to 4l} \times \varepsilon_{MC}^{gg\to ZZ\to 4l}\right) \times L$$

- A data-driven cross check has been done by normalizing to the measured Z rate. With this procedure the uncertainty on luminosity is canceled out by the ratio:

$$\frac{\sigma_{NLO}^{q\bar{q}\to ZZ\to 4l} + \sigma_{LO}^{gg\to ZZ\to 4l}}{\sigma_{NNLO}^{q\bar{q}\to Z\to 2l}} \times \frac{\varepsilon_{MC}^{ZZ\to 4l}}{\varepsilon_{MC}^{Z\to 2l}} \times N_{data}^{Z\to ll}$$

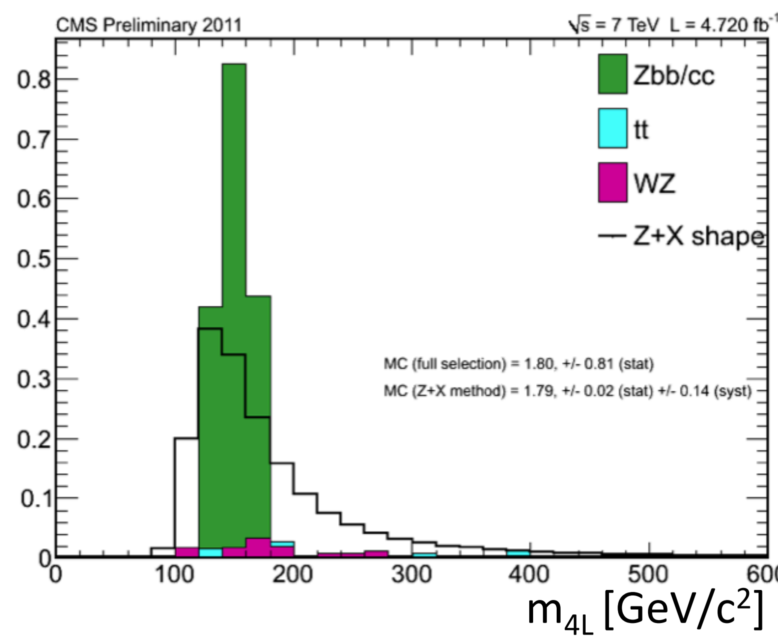
- The number of expected events, for the available integrated luminosity, is:

	Channel	Normalization to Z Rate	Direct measurement from MC
$q\bar{q}$	$N^{ZZ\to 4\mu}$	18.2 ± 1.6	18.0 ± 1.5
	$N^{ZZ\to 4e}$	11.7 ± 1.1	11.5 ± 1.0
	$N^{ZZ\to 2\mu 2e}$	29.0 ± 2.5	28.4 ± 2.4
gg	$N^{ZZ\to 4\mu}$	1.12 ± 0.34	1.11 ± 0.34
	$N^{ZZ\to 4e}$	0.79 ± 0.24	0.77 ± 0.24
	$N^{ZZ\to 2\mu 2e}$	1.8 ± 0.58	1.85 ± 0.56

Background control: reducible

- A signal-free control region CR is defined as Z_1 + 2 loose SS-SF leptons (Same Sign – Same Flavor)
- Every event in the CR is weighted with the probability for it to pass the tight selection, using a single-lepton fake rate. The fake rate is estimated from data, using a Z+1 lepton sample
- More direct methods are limited by the statistics available in Z+jets MC sample
- A closure test on MCs has been done, resulting in a fair agreement with expectations
- The number of expected events, for the available integrated luminosity, is:

Channel	Events expected (4.7 /fb)
4e	**1.67** ± 0.05 ± 0.5
4μ	**1.13** ± 0.09 ± 0.46
2e2μ	**2.71** ± 0.08 ± 0.88



Systematics

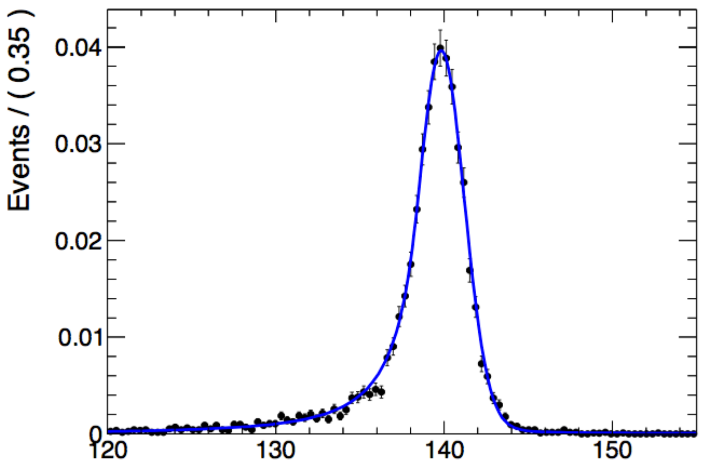
- **Theoretical uncertainties** on Higgs production and ZZ cross sections are considered. They come from PDF systematic errors and from uncertainties on QCD scale and renormalization factors μ_R and μ_F

Source	Uncertainty on XS_H	Uncertainty on XS_{ZZ}
gg partonic luminosity	8 %	10 %
qq partonic luminosity	3 – 5 %	5 %
QCD scale	5 – 12 %	2 – 40 %
Branching Ratio	2 %	

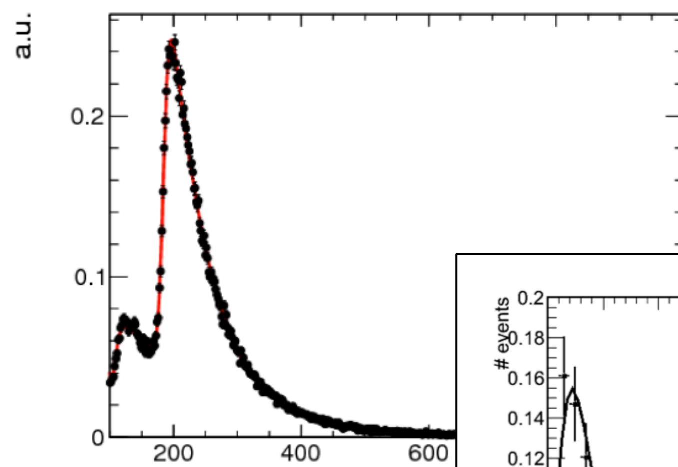
- Uncertainty on **integrated luminosity** is 4.5 %
- A 1.5 % uncertainty is assigned to **Trigger**
- Tag&Probe techniques, applied both on data and MC, allow to compute **data-to-simulation scale factors**, to be used as systematic uncertainties:
  - Lepton reconstruction/identification: 0.5 – 3.8 %
  - Lepton isolation: 1 – 2 %

Fit to mass distributions

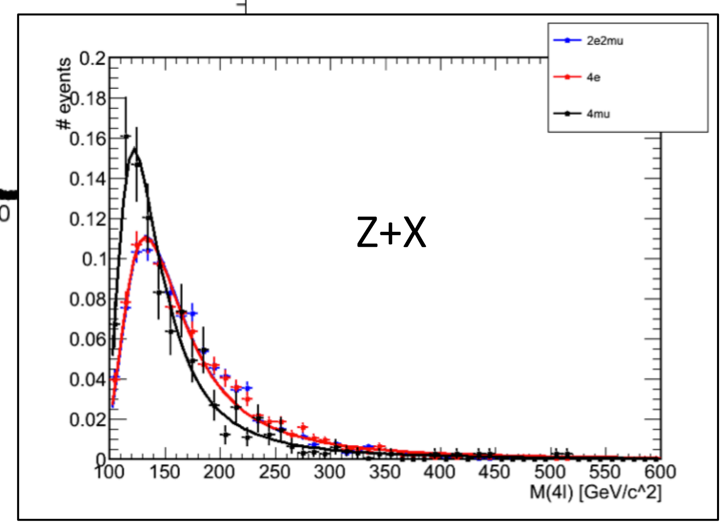
- Signal distribution is fitted with a Breit-Wigner function convoluted with a Crystall-ball function
- Irreducible ZZ continuum is fitted with an "empirical" p.d.f.
- Reducible background is modeled with a Landau function



Higgs m_H = 140 GeV/c^2



continuum ZZ



Results – Mass distribution

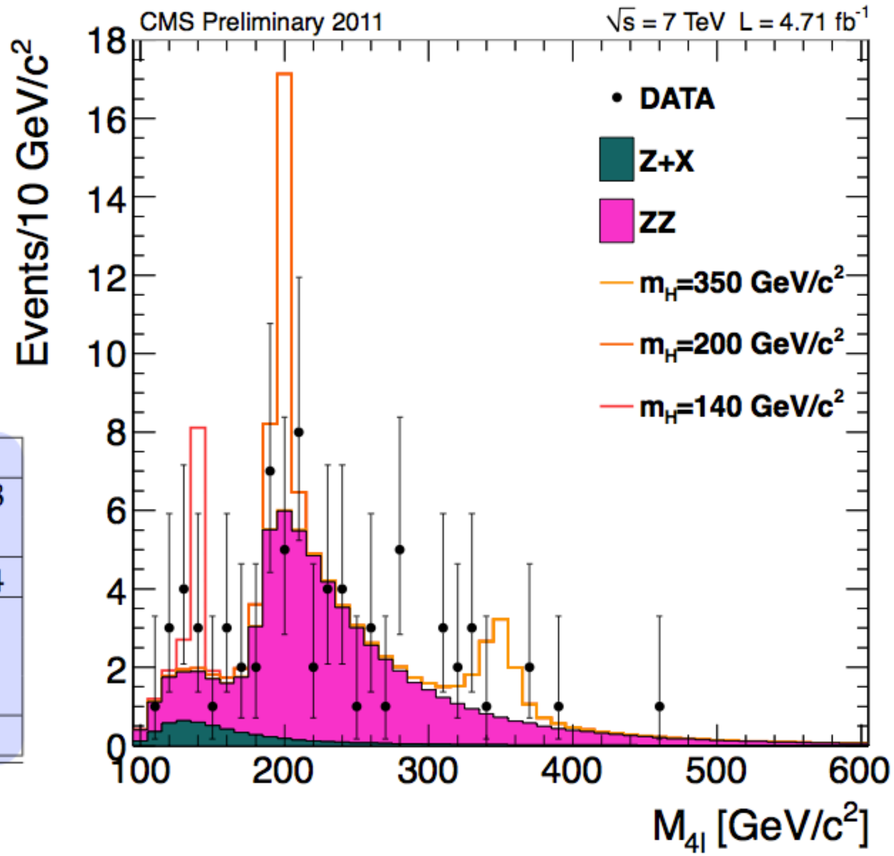
- Final mass distribution for events passing the full selection
- Expected distributions for background and (some) signals are shown

4mu+4e+2e2mu

Observed events: **72**

Expected events: 67.1 ± 6.0

Baseline	$4e$	4μ	$2e2\mu$
ZZ	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



Results – Low mass region

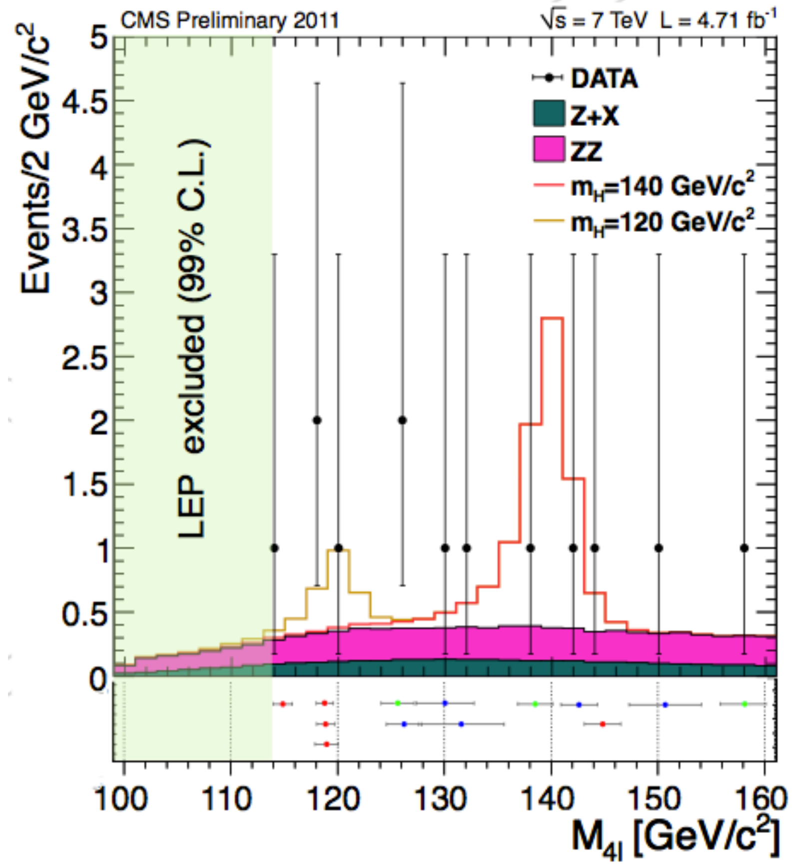
- Final mass distribution for events passing the full selection
  - $m_{4L} < 160\ GeV/c^2$
- Expected distributions for background and (some) signals are shown

4mu+4e+2e2mu

Observed events: **13**

Expected events: 9.5 ± 1.3

Final state:	4e	4μ	2e2μ
Obs. events:	3	5	5
Exp. events:	1.7	3.3	4.5



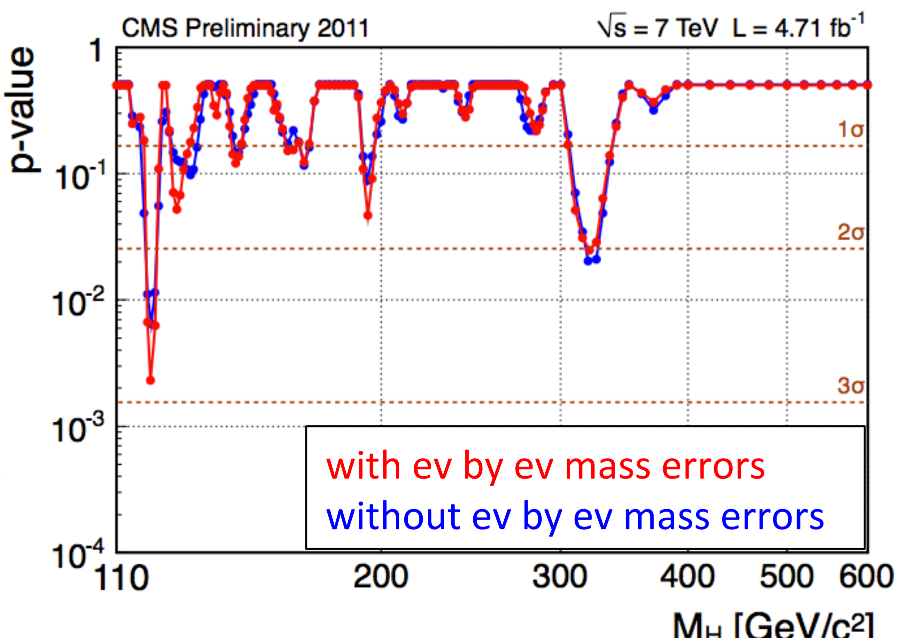
Results – Local p-values

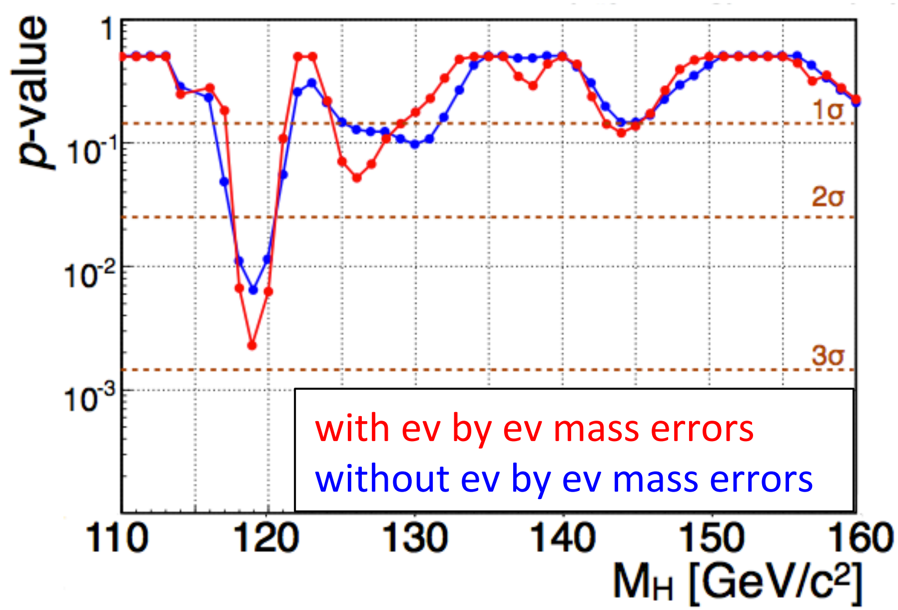
- **To quantify an excess** the p-value of q_0 , the (unbinned) profile likelihood discriminator is used:

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

  – μ is the "cross section modifier" XS/XS_{SM}
  – θ are the nuisance parameters, used to model systematics

- Local excesses are observed, up to ~3 sigma for m_H = 119 GeV $/c^2$ **compatible with background only hypothesis**

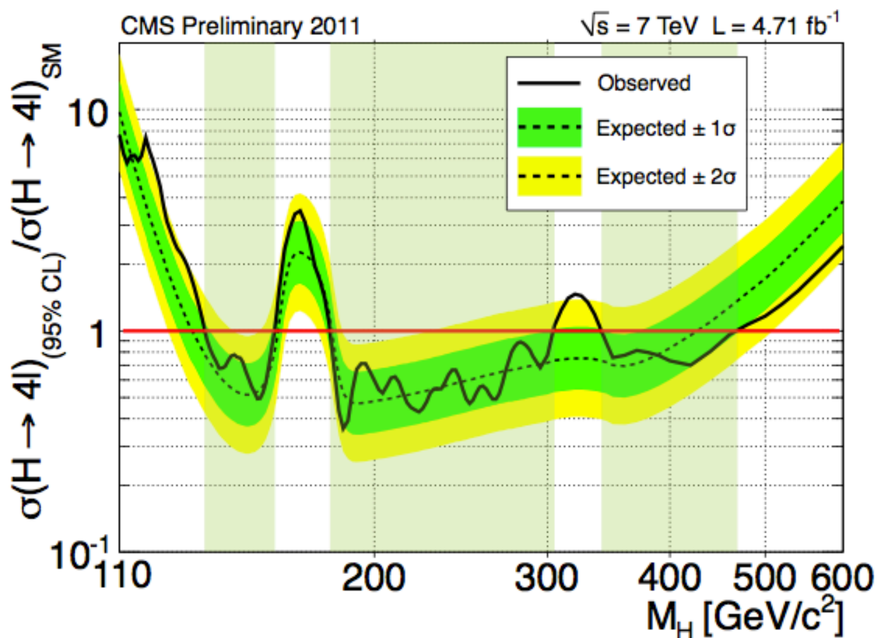




Results – Limits and exclusion

- **Limits** on μ are computed with the (frequentist) CLs method
  - this method uses the profile likelihood ratio q_{μ} as a discriminant between signal +background (SB) and background only (B) hypotheses, and it is designed to exclude the signal presence only if the analysis is able to distinguish it from the background
  - the signal presence is excluded at the 95% CL limit on signal cross-section goes below Standard Model predicted-value
- Within the 4-lepton channel the Standard Model Higgs boson is excluded at the 95% CL in the mass range:

[134-158], [180-305], [340-460] GeV/c^2



Conclusions

- The results for the SM Higgs search at CMS in the 4-lepton channel with 4.7 fb^{-1} have been presented
- The channel is sensitive to the Higgs search over a wide mass range
- 72 events have been observed, while 67.1 ± 6.0 were expected from SM background
- **The results are consistent with the background-only hypothesis**
- A local excess below 3-sigma is observed for $m_H \sim 119$ GeV/c^2
- The SM Higgs boson is excluded at the 95% CL in the mass range:

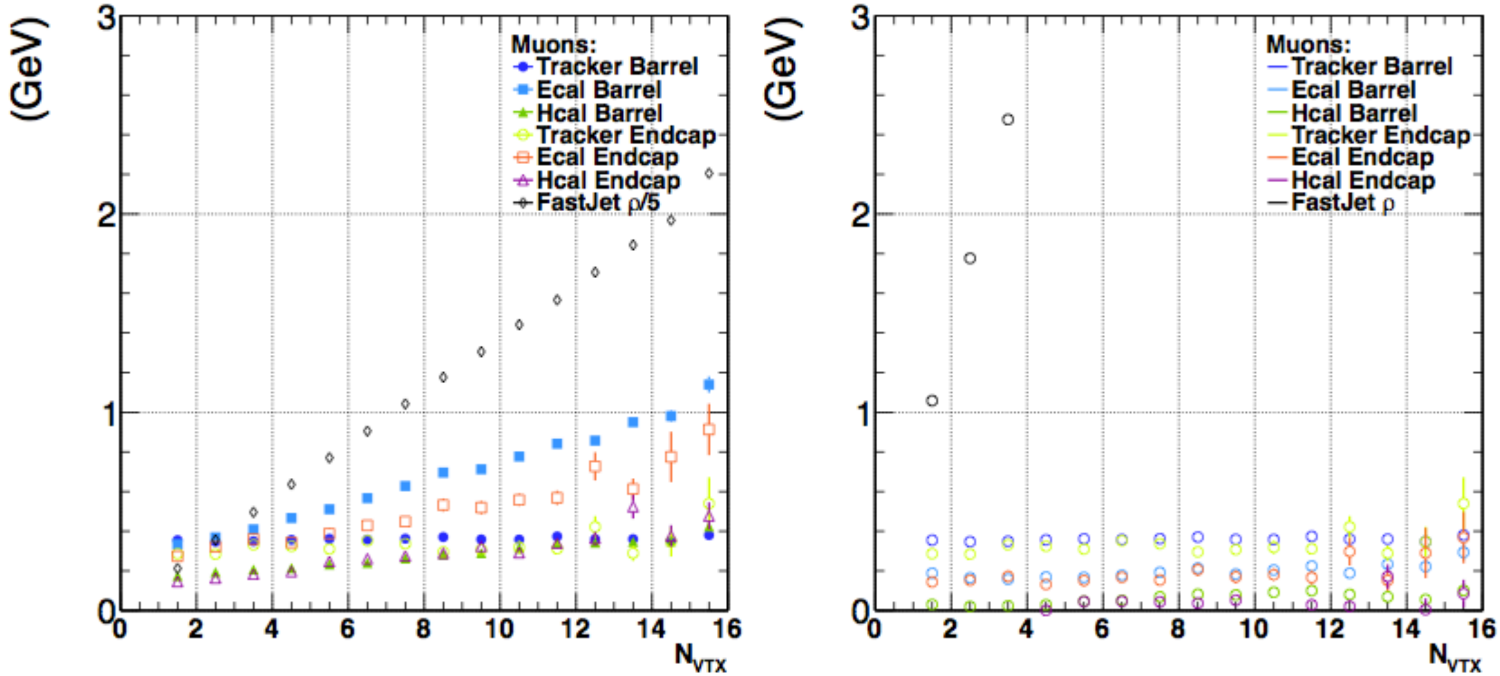
  [134-158], [180-305], [340-460] GeV/c^2
- With data expected by the end of 2012 (20 fb^{-1}) a local excess up to 4-sigma is expected in the 4-lepton channel for a ~120 GeV/c^2 Higgs mass, while the 5-sigma discovery will be accessible by the combination of the various analysis of CMS

Backup

Pile-up correction

- Isolation variables are the most sensitive to pile-up. In particular ECAL/HCAL
- The correction uses the median of Jet energy deposit in the event: ρ
- The effective cone-areas A_i are computed from data, both in barrel and endcap region

$$\sum Iso_{corrected} = \sum Iso - \rho \cdot A$$



Ev by Ev mass errors

- An uncertainty on each m4L entry can be assigned by propagating errors on track fit
- This allows a more accurate description of the oscillations in the *p-value*