

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

Meneghelli Marco^{1,a} for the CMS collaboration

¹*Bologna University and INFN sez. Bologna*

Abstract. One of the main targets of the CMS experiment is to search for the Standard Model Higgs boson. The 4-lepton channel (from the decay $H \rightarrow ZZ \rightarrow 4l, l = e, \mu$) is one of the most promising. The analysis is based on the identification of two opposite-sign, same-flavor lepton pairs: leptons are required to be isolated and to come from the same primary vertex. The Higgs would be statistically revealed by the presence of a resonance peak in the 4-lepton invariant mass distribution. The Higgs mass is a free parameter of the Standard Model, and the 4-lepton channel search is sensitive almost in the entire mass range. With data collected in 2010 and 2011 (4.7 fb^{-1} at $\sqrt{s} = 7 \text{ TeV}$) the Higgs boson has been excluded in a wide region of mass at 95% CL.

1 Search for the Higgs boson at CMS: the 4-lepton channel

The Standard Model (SM) of electroweak interactions [1][2] predicts the existence of a scalar particle, the Higgs boson, associated with the field responsible for the spontaneous electroweak symmetry breaking. This particle would be responsible also for masses of SM fermions. Its own mass, m_H , is a free parameter of the theory. Experimentally, direct searches for the SM Higgs boson at LEP [3] and Tevatron [4] colliders have led to a lower mass bound of $m_H > 114.4 \text{ GeV}$, and to an exclusion in the range $162 - 166 \text{ GeV}$, at 95% Confidence Level (CL). Indirect constraints from precision measurements [5] favor the mass range $m_H < 158 \text{ GeV}$ at 95% CL.

The Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider [6] is designed to search for the Higgs boson and possible signals of “New Physics” arising from the proton-proton collisions at high energy. The experimental apparatus is described elsewhere [7]: its main features are a superconducting solenoidal magnet, core of a cylindrical structure that sees from inside out the silicon tracker detector, the electromagnetic and hadron calorimeters (ECAL and HCAL) and, externally, the muon system.

The $H \rightarrow ZZ \rightarrow 4l$ analysis [24] exploits the Higgs decay channel $H \rightarrow ZZ \rightarrow l_1^+ l_1^- l_2^+ l_2^-$, with $l_i = e, \mu$. It looks for events containing four isolated leptons coming from the primary interaction vertex. Such events are quite rare in SM processes, so the analysis has an advantageous signal-to-background ratio, despite the small cross section for the $pp \rightarrow H \rightarrow ZZ \rightarrow 4l$ process, which ranges from 0.5 fb to 7 fb depending on the Higgs mass. The main background is the irreducible non-resonant ZZ production, with Zs decaying leptonically, with an yield of $\sim 14 \text{ fb}$. A small contamination of Z+X reducible background (that is highly suppressed by lepton ID cuts) survives at small masses, around

^ae-mail: marco.meneghelli@cern.ch

130 GeV. The events have a very clean signature and the 4-lepton mass can be reconstructed with excellent resolution. The Higgs boson would be revealed by the presence of a resonance peak in the 4-lepton mass distribution. The analysis has a good sensitivity for Higgs discovery in a wide mass range, from ~ 120 GeV to ~ 1 TeV.

1.1 Data and MC

Data correspond to an integrated luminosity of 4.7 fb^{-1} from 7 TeV pp collisions, collected during 2010 and 2011. Events are selected by the trigger system that requires the presence of a pair of electrons (a pair of muons) with transverse energy (transverse momenta) for the first and second lepton above 17 and 8 GeV respectively. The trigger efficiency within the acceptance of this analysis is of the order of 99% for signal.

The Higgs boson signals from gluon-fusion ($gg \rightarrow H$), and vector-boson fusion ($qq \rightarrow qqH$), are generated with POWHEG [8] at next-to-leading-order (NLO). Contributions from gluon fusion up to next-to-next-to-leading-order (NNLO) and next-next-to-leading-log, taken from [9], are used to rescale the event to get the total cross section $\sigma(pp \rightarrow H)$. The cross section is scaled by the branching fraction $BR(H \rightarrow 4l)$ [10].

The SM background contribution from ZZ production via qq is generated at NLO with POWHEG, other di-boson processes (WW, WZ) are generated with PYTHIA [12] with cross sections rescaled to NLO predictions. The $gg \rightarrow ZZ$ contribution is generated with GG2ZZ [11]. The $Zb\bar{b}$, $Zc\bar{c}$, $Z\gamma$, and Z+light-jets samples are generated with MADGRAPH [13] with cross sections rescaled to NNLO. The $t\bar{t}$ events are generated at NLO with POWHEG.

All events are processed through a detailed simulation of the CMS detector based on GEANT4 [14] and are reconstructed with the same algorithms that are used for data.

2 Leptons and event selection

The events selected by the trigger system (as described in section 1.1) are required to pass an offline selection, studied to maximize the analysis sensitivity for the Higgs, thus to maximize the final signal yield over the background one. The reducible backgrounds, like $Zb\bar{b}$, $Zc\bar{c}$, $Z\gamma$, Z+light-jets, $t\bar{t}$, are almost totally killed by selection cuts, so at the end the dominant background is the irreducible ZZ non-resonant production via quark or gluon fusion.

2.1 Lepton identification

The leptons used in the analysis are identified by stringent ID requirements in order to minimize the contamination from fakes, like undecayed charged hadrons. The variables of “isolation” and “significance of impact parameter” are assigned using dedicated algorithms.

Muons are reconstructed using the “Global Muon” algorithm [15], that fits the hits both in the muon system and in the silicon central tracker and combines the information in a unique fit using a Kalman-Filter technique. The four-momentum to be assigned to the muon object is derived from the inner track. Some ID cuts are applied:

- $p_T > 5 \text{ GeV}$
- $|\eta| < 2.4$
- number of tracker hits > 10

Electrons are reconstructed using the ‘‘Gaussian Sum Filter’’ (GSF) algorithm [16][17], that combines ECAL and tracker information, matching (super)clusters of energy in the ECAL with hits in the inner tracker. The four-momentum is given by a combination of ECAL and tracker information. Some ID cuts are applied:

- $p_T > 7$ GeV
- $|\eta| < 2.5$
- electrons are required to be ‘‘tight’’, with the identification sensitive to the amount of bremsstrahlung emitted along the trajectory in the inner tracker, as described in [16]
- electrons from photon conversions are rejected cutting on expected missing inner hits

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. *The isolation* is evaluated looking to p_T of tracks and to energy deposits in a geometrical cone around the lepton track. A small ‘‘veto’’ cone around the lepton is excluded from the computation in order to avoid the self-contribution. Three quantities are used:

- lepton track isolation ISO_{track} : sum of p_T of tracks within a $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.3$ cone
- lepton ECAL/HCAL isolation ISO_{ECAL}, ISO_{HCAL} : sum of E_T from energy deposits within a $\Delta R < 0.3$ cone

ISO_{ECAL} and ISO_{HCAL} are corrected for the effect of pile-up increasing during data taking. The three quantities are combined in a unique variable, normalized to lepton p_T , that is assigned to each lepton:

$$R_{iso} = (ISO_{track} + ISO_{ECAL} + ISO_{HCAL})/p_T$$

All the leptons are required to come from a common primary vertex. This allows to discriminate against reducible background (in particular $Zb\bar{b}, Zc\bar{c}, t\bar{t}$), by rejecting leptons coming from secondary vertices (b-jets). The key quantity is *the significance of impact parameter*, the distance of closest approach of lepton track with respect to the primary interaction vertex, normalized to its uncertainty:

$$|SIP_{3D}| = |IP_{3D}/s_{IP}|$$

2.2 Event selection

Firstly a loose event *skim* is applied, requiring two reconstructed leptons above a given p_T , with an invariant mass greater than 40 GeV. Then follow the *selection* steps:

- the identification of a first, on shell, Z boson: the opposite-sign, same-flavor, lepton pair with the mass closest to 91.19 GeV, with each lepton passing ID quality cuts
- a third lepton passing ID quality cuts
- a fourth lepton passing ID quality cuts, with same-flavor and opposite-sign with respect to the third lepton
- the choice of the best Z_2 as the opposite-sign same-flavor lepton pair with the highest p_T . For the $4e$ and 4μ final states, at least three of the four combinations of opposite-sign pairs must satisfy $m_{ll} > 12$ GeV
- cut on lepton isolation: $R_{iso}^i + R_{iso}^j < 0.35$ over all the lepton couples (i, j)
- cut on significance of impact parameter: $\max_i \{SIP_{3D}^i\} < 4$ over all the leptons i

- cut on kinematics, it is required:

$$50 < m_{Z_1} < 120 \text{ GeV}$$

$$12 < m_{Z_3} < 120 \text{ GeV}$$

$$m_{4l} > 100 \text{ GeV}$$

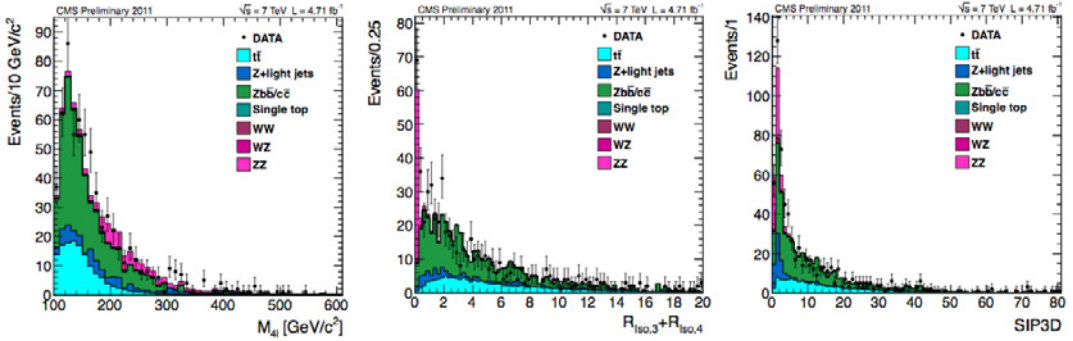


Figure 1. Distributions of m_{4l} (left), isolation (center) and SIP_{3D} (right) variables at an early stage of selection. The isolation and SIP_{3D} variables have a highly discriminant power against the reducible background, that would otherwise mask any signal in the m_{4l} spectrum.

A comparison done at the early stage of selection, with relaxed cuts on ID, isolation, SIP_{3D} and flavor-sign of leptons, allow to see the difference between signal-like events and reducible background ones: see Fig. 1. Cuts on isolation and SIP_{3D} are crucial to discriminate signal against reducible background, which would otherwise give a significant yield, dramatically lowering the analysis sensitivity.

3 Signal and backgrounds

The signal cross section $\sigma(pp \rightarrow H \rightarrow ZZ \rightarrow 4l)$ ranges from 0.5 fb to 7 fb depending on the Higgs mass as reported in [9]. Signal efficiency strongly depends on the Higgs mass and on the final sub-channel. It is typically low, around $\sim 20 - 30\%$ for low masses ~ 115 GeV. The efficiency grows with Higgs mass reaching a high efficiency plateau, around $\sim 80 - 90\%$ depending on the sub-channel, for masses greater than ~ 200 GeV. The signal invariant mass distribution has a peaked shape, broadened by resolution effects, and it is fitted with a Breit-Wigner pdf convoluted with a Crystal-ball function.

The irreducible ZZ continuum efficiency is directly estimated from MC, while the cross section is theoretically predicted up to NLO [18][19]. The number of expected events, for the available integrated luminosity, is reported in table 1 for the three sub-channels. The ZZ invariant mass shape is fitted with an empirical pdf.

The reducible Z+X yield is estimated with a data-driven technique. A signal-free Control Region (CR) is defined as $Z_1 + 2$ loose same-sign same-flavor leptons. Then every event in the CR is weighted with the probability for it to pass the tight selection, using a single-lepton fake rate, estimated from data, using a Z + 1 lepton sample. The number of expected events, for the available integrated luminosity, is reported in table 2. The Z + X invariant mass shape is fitted with a Landau function.

Table 1. Number of irreducible ZZ yield, $L_{int} = 4.7 \text{ fb}^{-1}$

Channel	$q\bar{q} \rightarrow ZZ$	$gg \rightarrow ZZ$
$4e$	11.5 ± 1.0	0.77 ± 0.24
4μ	18.0 ± 1.5	1.11 ± 0.34
$2e2\mu$	28.4 ± 2.4	1.85 ± 0.56

Table 2. Number of irreducible Z+X yield, $L_{int} = 4.7 \text{ fb}^{-1}$

Channel	Yield \pm stat. \pm syst.
$4e$	$1.67 \pm 0.05 \pm 0.5$
4μ	$1.13 \pm 0.09 \pm 0.46$
$2e2\mu$	$2.71 \pm 0.08 \pm 0.88$

3.1 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 [20][21]:

- gg partonic luminosity: 8% on $\sigma(pp \rightarrow H)$ and 10% on $\sigma(pp \rightarrow ZZ)$
- qq partonic luminosity: 3 - 5 % on $\sigma(pp \rightarrow H)$ and 5% on $\sigma(pp \rightarrow ZZ)$
- QCD scale: 5 - 12% on $\sigma(pp \rightarrow H)$ and 2 - 40% on $\sigma(pp \rightarrow ZZ)$
- Branching Ratio: 2%

Instrumental uncertainties are:

- integrated luminosity: 4.5%
- trigger efficiency: 1.5%
- 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 and identification: 0.5-3.8%

lepton isolation: 1-2%

4 Results

The final $4l$ invariant mass distribution, for events passing the full selection, is reported in Fig. 2. Distributions for irreducible and reducible SM backgrounds are shown as full histograms, while empty histograms are three possible signals, with different m_H .

72 events pass the selection, 67.1 ± 6.0 are expected from background.

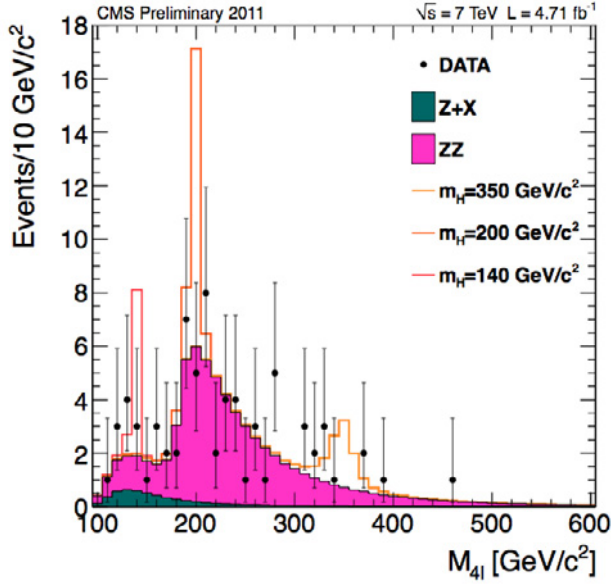


Figure 2. 4-lepton invariant mass spectrum, $L_{int} = 4.7 \text{ fb}^{-1}$, $\sqrt{s} = 7 \text{ TeV}$

4.1 Statistical analysis

Limits on Higgs cross section, for different mass hypotheses, are computed with the (frequentist) CLs method [22]:

- This method uses the profile likelihood ratio 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 sensitive to it
- The signal presence is excluded at 95% CL if the limit on signal cross section goes below the SM predicted value

Within the 4-lepton channel, the SM Higgs boson is excluded at the 95% CL in the mass range:

$$[134 - 158], [180 - 305], [340 - 460] \text{ GeV}$$

In Fig. 3, the 95% CL limits on SM Higgs cross section are reported, as functions of the hypothesized Higgs mass.

To quantify an excess over the background, the p-value of the (unbinned) profile likelihood discriminator is used, as described in [23]. The lower the p-value is, the more improbable the background-only hypothesis is. In Fig. 4, the background-only local p-values are reported, as functions of the hypothesized Higgs mass. Two approaches are shown: taking into account an event-by-event mass error (red), and without doing it (blue). A local excess, still below 3-sigma, is observed for $m_H \sim 119 \text{ GeV}$.

In general, the results are compatible with the background-only hypothesis.

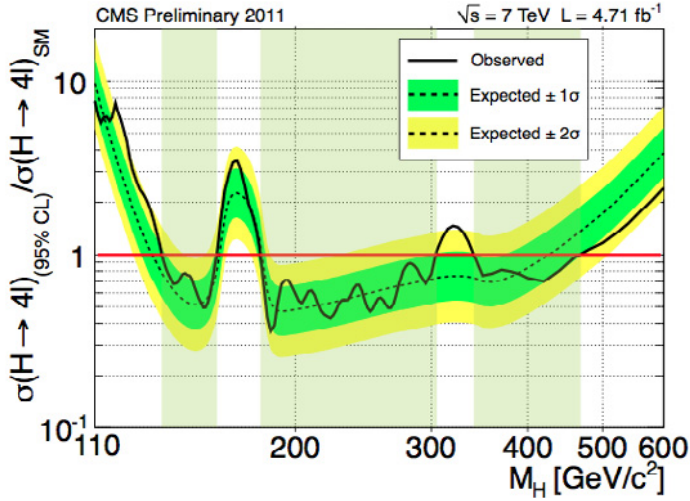


Figure 3. 95 % CL limit on SM Higgs cross section, as a function of the Higgs mass

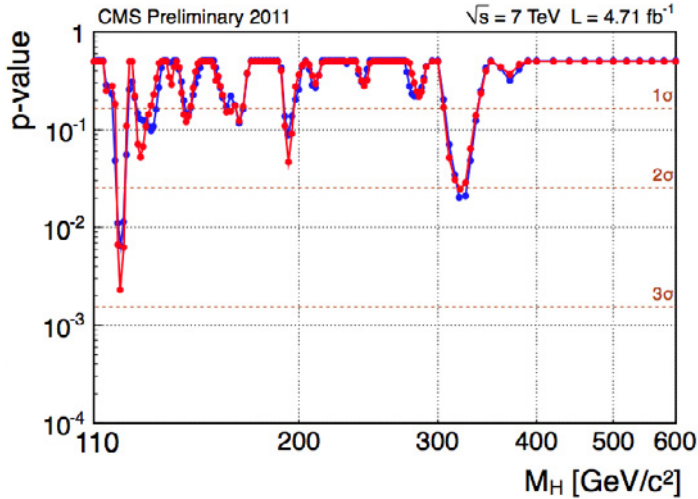


Figure 4. p-values on background-only hypothesis, as a function of the Higgs mass. Red: with ev. by ev. mass errors, Blue: without ev. by ev. mass errors

5 Conclusions

The results for the SM Higgs search at CMS in the 4-lepton channel, with 4.7 fb^{-1} , have been briefly presented. This channel is sensitive to the Higgs search over a wide mass range. After having applied the selection on the 2010 and 2011 data, 72 events are observed, while 67.1 ± 6.0 are 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. The SM Higgs boson is excluded at the 95% CL in the mass range [134 – 158], [180 – 305], [340 – 460] GeV.

References

- [1] S. L. Glashow, Nucl. Phys. **22** (1961) 579-588
- [2] S. Weinberg, Phys. Rev. Lett. **19** (1967) 1264-1266
- [3] ALEPH, DELPHI, L3 and OPAL Collaborations, CERN-EP/2003-011
- [4] The TEVNPH Working Group, FERMILAB-PUB-09-060-E, CDF Note **9713**, D0 Note **5889**
- [5] Andreas Hoecker, arXiv:0909.0961v2
- [6] L. Evans, P. Bryant, et. al., JINST **3** (2008) S08001
- [7] CMS Physics Technical Design report, Vol I, CERN/LHCC 2006-001
- [8] S. Frixione, P. Nason, and C. Oleari, JHEP **11** (2007) 070
- [9] C. Anastasiou, R. Boughezal, and F. Petriello, JHEP **04** (2009) 003
- [10] LHC Higgs Cross Section Working Group, CERN Report CERN-2011-002, (2011)
- [11] T. Binoth, N. Kauer, and P. Mertsch, in Proceedings of the XVI Int. Workshop on Deep-Inelastic Scattering and Related Topics (DIS'07) 2008
- [12] T.Sjostrand, S.Mrenna, and P.Z.Skands, JHEP **05** (2006) 026
- [13] J. Alwall et al., JHEP **09** (2007) 028
- [14] J. Allison et al., IEEE Trans. Nucl. Sci. **53** (2006) 270.
- [15] CMS Collaboration, CMS Physics Analysis Summary CMS-PAS-MUO-10-002, (2010)
- [16] S. Baffioni et al., Eur. Phys. J. C **49** (2007) 1099
- [17] CMS Collaboration, Physics Analysis Summary CMS-PAS-EGM-10-004, (2010)
- [18] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60** (1999) 113006
- [19] J. M. Campbell, R. Ellis, and C. Williams, JHEP **07** (2011) 018
- [20] H.-L. Lai et al., Phys. Rev. D **82** (2010) 074024
- [21] A. Martin et al., Eur. Phys. J. C **63** (2009) 189
- [22] T. Junk, Nucl. Instrum. Meth. A **434** (1999) 435
- [23] E. Gross and O. Vitells, Eur. Phys. J. C **70** (2010) 525-530
- [24] CMS collaboration, Phys. Rev. Lett. **108**, 111804 (2012)