

Analysis strategy for the SM Higgs boson search in the four-lepton final state in CMS

Alberto Graziano^{1,a}

Department of Experimental Physics, University of Torino/INFN, Torino, Italy

Abstract. The current status of the searches for the SM Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ decay channel with the CMS experiment [1] is presented. The selection cuts for suppressing the backgrounds while keeping very high signal efficiencies are described, along with the data-driven algorithms implemented to estimate the background yields and the systematic uncertainties. With an integrated luminosity of 1.66 fb^{-1} , upper limits at 95% CL on the SM-like Higgs cross section \times branching ratio exclude cross sections from about one to two times the expected value from the Standard Model in the range $150 < m_H < 420 \text{ GeV}$. No evidence for the existence of the SM Higgs boson has been found so far.

1 Signal and backgrounds

The $H \rightarrow ZZ \rightarrow 4\ell$ analysis described here considers three final states ($4e$, 4μ , $2e2\mu$) over a Higgs boson mass range $100 < m_H < 600 \text{ GeV}/c^2$. The signature of signal events is very clean: two pairs of same-flavour, opposite-charge, high- p_T isolated leptons pointing to the same reconstructed vertex are looked for. Mass constraints can be set on the dilepton invariant mass: also at low m_H , at least one Z boson is on-shell, therefore at least one pair of leptons has $m_{\ell^+\ell^-} \approx m_Z$.

The $ZZ \rightarrow 4\ell$ background is often referred to as the *ir-reducible* one, because of its signal-like kinematics. Since the SM Higgs boson is a scalar particle, angular correlations among the final-state leptons can be exploited to discriminate between the signal and the ZZ background.

The *reducible* backgrounds are $Zb\bar{b}$, $Zc\bar{c}$, $t\bar{t}$, with heavy-flavour quarks decaying semileptonically. Leptons originating from these decays usually point to displaced vertices and they are soft and not isolated. The *instrumental* backgrounds are processes with jets misidentified as leptons, such as $Z + \text{jets}$, $W + \text{jets}$, QCD events.

2 Event selection

The analysis consists of a set of cuts aiming to reduce the background contributions by preserving a high signal efficiency, as described in [2] [3] [4].

– Requirements on muons and electrons

First of all, electrons and muons must satisfy some trigger, reconstruction and identification requirements. Depending on the *trigger menu* deployed during data taking, reconstructed leptons are required to be matched to online trigger objects passing a single- or double-lepton trigger selection. They also have to pass loose p_T and isolation cuts.

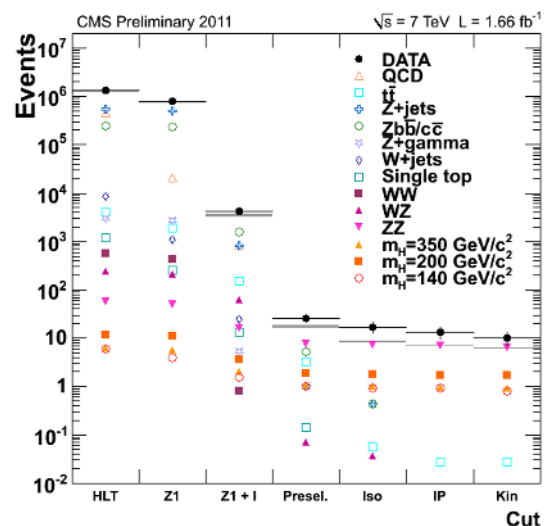


Fig. 1. Event yields in the 4μ channel as a function of the event selection steps. Black points with uncertainties represent the data, other symbols represent the MC expectations. The samples correspond to an integrated luminosity of $L = 1.66 \text{ fb}^{-1}$.

– First Z candidate selection

The first Z candidate is defined as the one with the dilepton invariant mass closest to the nominal Z mass, in the mass window $60 < m_{\ell^+\ell^-} < 120 \text{ GeV}/c^2$, after a selection including cuts on the p_T of both leptons, on their isolation and on the significance of their 3D impact parameter with respect to the primary event vertex.

– Z₁ + 1 lepton

The presence of a third high-quality lepton of any flavour and charge is required. At this stage of the selection the phase space of the main reducible backgrounds is preserved for data-driven background estimation and control.

^a e-mail: alberto.graziano@cern.ch

– $Z_1 + 2$ leptons

The requirement of a fourth lepton with matching flavour and opposite charge with respect to the third one is added.

– ‘Best 4ℓ candidate’ selection

The second Z candidate is reconstructed from the two highest- p_T leptons not associated to Z_1 and passing $m_{Z_2} > 12 \text{ GeV}/c^2$. At this stage the ambiguity due to combinatorics in events with extra fake leptons is limited and the ‘best 4ℓ candidate’ is chosen.

The 4ℓ candidate must satisfy $m_{4\ell} > 100 \text{ GeV}/c^2$. Moreover, in the $4e$ and 4μ final states only, at least three out of the four possible $\ell^+\ell^-$ combinations are required to have $m_{\ell^+\ell^-} > 12 \text{ GeV}/c^2$ in order to reject background events with leptonic J/ψ decays.

– Cut on relative lepton isolation

The two leptons with the largest isolation variable, which is the sum of tracker, ECAL and HCAL isolation divided by the lepton p_T , are then considered. The sum of their isolation values is required to be lower than a threshold.

– Cut on the 3D impact parameter significance of leptons

The displaced vertex of leptons originating from b -quark decays can be a handle for further background rejection. A cut is therefore applied on the significance of the 3D impact parameter of the lepton track with respect to the reconstructed primary vertex. This significance is defined as $SIP_{3D} = IP_{3D}/\sigma_{IP_{3D}}$ and its absolute value is required to be less than 4 for all selected leptons.

– Cut on Z_1, Z_2 kinematics

Finally, additional constraints are imposed on the p_T of the selected leptons ($p_T^{\ell_1, \ell_2, \mu_3, \mu_4} > 20, 10, 5, 5 \text{ GeV}/c$ for muons, $p_T^{\ell_1, \ell_2, e_3, e_4} > 20, 10, 7, 7 \text{ GeV}/c$ for electrons) and on the invariant mass of the second Z candidate, which is required to be $20 < m_{Z_2} < 120 \text{ GeV}/c^2$ in the *baseline* selection and $60 < m_{Z_2} < 120 \text{ GeV}/c^2$ in the *high mass* selection. The $m_{4\ell}$ distribution for events passing the baseline selection is shown in Fig. 2.

3 $ZZ \rightarrow 4\ell$ cross section measurement

The $ZZ \rightarrow 4\ell$ inclusive cross section has been measured after the *high mass* selection cuts $60 < m_{Z_1} < 120 \text{ GeV}/c^2$, $60 < m_{Z_2} < 120 \text{ GeV}/c^2$ as

$$\begin{aligned} \sigma(pp \rightarrow ZZ + X) \times BR(ZZ \rightarrow 4\ell) &= \frac{\sum (N_{obs}^{ich} - N_{bkg}^{ich})}{\mathcal{A}_{4\ell} \times \epsilon_{ZZ \rightarrow 4\ell} \times \mathcal{L}} = \\ &= 20.84_{-4.0}^{+6.8} \text{ (stat.)} \pm 0.54 \text{ (syst.)} \pm 0.94 \text{ (lumi.) fb} \end{aligned} \quad (1)$$

where the sum runs over the three final states ($4e, 4\mu, 2e2\mu$). This result should be compared with the theoretical value:

$$\sigma_{TH}(pp \rightarrow ZZ + X) \times BR(ZZ \rightarrow 4\ell) = 28.32 \pm 2.57 \text{ fb}$$

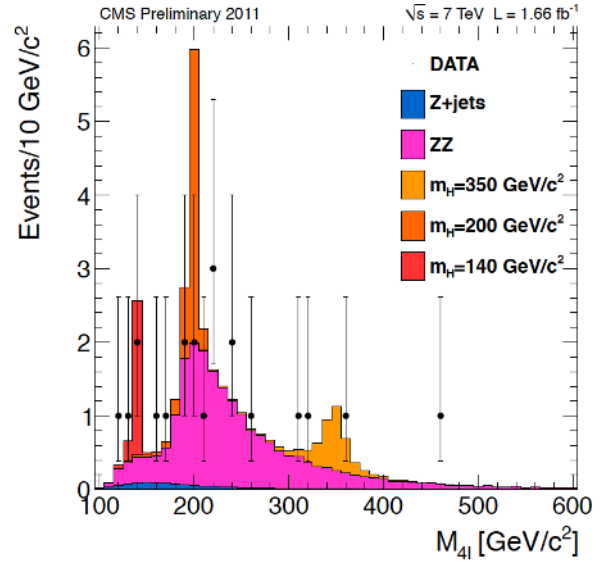


Fig. 2. Distribution of $m_{4\ell}$ after the baseline selection for the sum of the 4ℓ channels. Points represent the data, shaded histograms represent the signal and background expectations. The samples correspond to an integrated luminosity of $L = 1.66 \text{ fb}^{-1}$.

4 ZZ background control

The contribution from ZZ background, which is the main one after the whole event selection, can be estimated from the number of $Z \rightarrow 2\ell$ events observed in data with the following formula:

$$N_{ZZ \rightarrow 4\ell} = \frac{\sigma_{q\bar{q} \rightarrow ZZ \rightarrow 4\ell}^{NLO} + \sigma_{gg \rightarrow ZZ \rightarrow 4\ell}^{LO}}{\sigma_{pp \rightarrow Z \rightarrow 2\ell}^{NNLO}} \cdot \frac{\epsilon_{ZZ \rightarrow 4\ell}^{MC}}{\epsilon_{Z \rightarrow 2\ell}^{MC}} \cdot N_{Z \rightarrow 2\ell}^{observed} \quad (2)$$

This method exploits the fact that most Feynman diagrams are shared by the two processes. The results are compatible with those obtained directly from MC, but the systematic uncertainties are smaller because most of them cancel out in the ratio.

5 $Zb\bar{b}, Zc\bar{c}, t\bar{t}$ background control

In order to perform a data-driven measurement of the $Zb\bar{b}, Zc\bar{c}, t\bar{t}$ background yield, the first Z candidate is defined as in the signal selection, whereas the flavour, charge and isolation requirements on the leptons from Z_2 are relaxed and the cut on their impact parameter significance reversed: $|SIP_{3D}| > 5$ (see Fig. 3).

To propagate the event yield from this control region to the signal phase space, correction factors are introduced that account for the relaxed isolation and kinematical cuts, for the reversed impact parameter cut and for the combinatorics related to considering pairs of leptons of any flavour and charge.

6 Z + jets background control

6.1 Single lepton fake rate measurement

Prior to measuring the Z+jets background yield, the *single lepton fake rate* must be evaluated. This is done from a

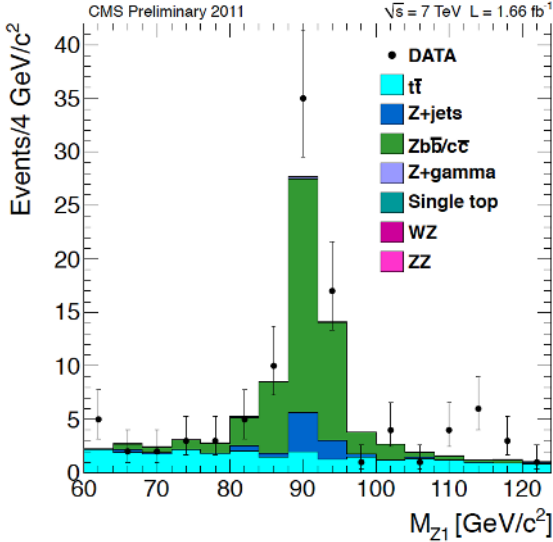


Fig. 3. Distribution of the best reconstructed Z candidate invariant mass for the events in the 4ℓ background control region. Solid points represent the data, shaded histograms represent the MC expectations. The signal and the ZZ background contribute negligibly. The samples correspond to an integrated luminosity of $L = 1.66 \text{ fb}^{-1}$.

sample of exactly 3 leptons, therefore signal-free, in which the contamination from WZ events is removed with a cut on the missing transverse energy ($E_T < 25 \text{ GeV}/c^2$). A Z_1 candidate is looked for like in the signal selection and the remaining lepton (a ‘fakeable object’) is required to pass identification and isolation cuts. The fraction of fakeable objects passing this selection, as a function of lepton p_T and pseudorapidity, is the single lepton fake rate:

$$\varepsilon(p_T^\ell, \eta^\ell) = \frac{N(\text{passing ID and isolation cuts})}{N(\text{fakeable objects})} \quad (3)$$

6.2 Definition of the control region

A Z_1 candidate is reconstructed as in the signal selection. The control region is signal-free because the third and fourth leptons are required to have the same flavour and charge, $\ell_3^\pm \ell_4^\pm$ (the lepton charge misassignment is negligible). No identification and isolation cuts are applied on these two leptons, whereas the kinematical cuts $m_{\ell_3 \ell_4} > 12 \text{ GeV}/c^2$, $m_{4\ell} > 100 \text{ GeV}/c^2$ are.

6.3 Extrapolation to the signal region

The number of Z+jets events in the signal region (SR) can be extrapolated from the one in the control region (CR) by means of the following formula:

$$N_{SR}^{Z+jets} = N_{CR}^{Z+jets} \times \frac{1}{2} \times \frac{\varepsilon(p_T^{\ell_3}, \eta^{\ell_3}) \times \varepsilon(p_T^{\ell_4}, \eta^{\ell_4})}{[1 - \varepsilon(p_T^{\ell_3}, \eta^{\ell_3})] \times [1 - \varepsilon(p_T^{\ell_4}, \eta^{\ell_4})]} \quad (4)$$

where N_{CR}^{Z+jets} is scaled by the ratio of the joint probability of both ℓ_3 and ℓ_4 passing the ID and isolation selection (as expected in the SR) to the joint probability of neither of them passing it (as expected in the CR). The factor 1/2 accounts for the triangular area of the signal region in the $(ISO_{\ell_3}, ISO_{\ell_4})$ plane. See Fig. 4.

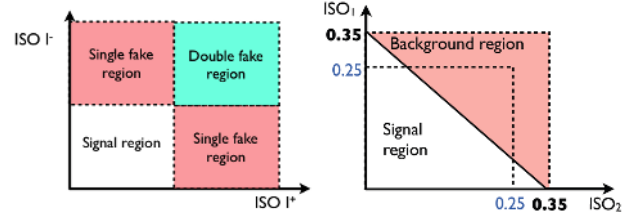


Fig. 4. Two sketches about the extrapolation from the control region to the signal one for the evaluation of the Z+jets event yield. Left: the CR, called ‘Double fake region’, is shown in green, the SR in white. Right: the SR is defined by $ISO_{\ell_3} + ISO_{\ell_4} < 0.35$.

7 Exclusion limits for $\sqrt{s} = 7 \text{ TeV}$, $L = 1.66 \text{ fb}^{-1}$

With $\sqrt{s} = 7 \text{ TeV}$ and $L = 1.66 \text{ fb}^{-1}$, the search for a SM-like Higgs boson performed in the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ channel allows to set 95% CL upper limits on $\sigma \cdot BR$ that exclude cross sections from about one to two times the expected SM ones in the mass range $150 < m_H < 420 \text{ GeV}/c^2$. See Fig. 5.

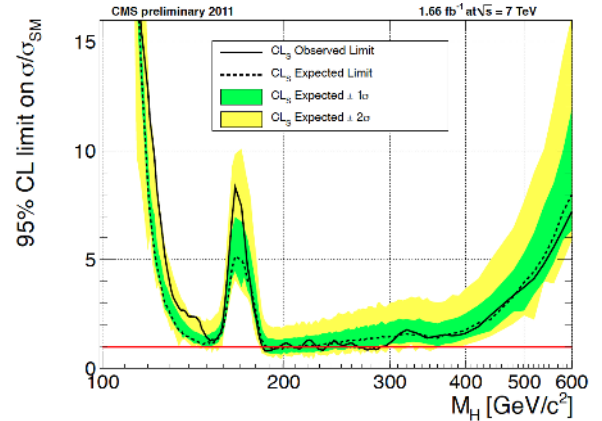


Fig. 5. The mean expected and the observed upper limits at 95% CL on $\sigma(pp \rightarrow H + X) \times BR(ZZ \rightarrow 4\ell)$ for a Higgs boson in the mass range $120 \div 600 \text{ GeV}/c^2$, for an integrated luminosity of 1.66 fb^{-1} using the CL_s approach. The expected ratio for the SM is presented. The results are obtained using a shape analysis method.

References

1. The CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3 (2008) S08004.476
2. The CMS Collaboration, *Search for a Standard Model Higgs boson in the decay channel $H \rightarrow ZZ \rightarrow 4\ell$* , CMS Physics Analysis Summary 2011/015
3. The CMS Collaboration, *Search for a Standard Model Higgs boson in the decay channel $H \rightarrow ZZ \rightarrow 4\ell$* , CMS Physics Analysis Summary 2011/004
4. N. Amapane et al., *Search for a Standard Model Higgs boson in the decay channel $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$* , CMS AN 2011/123