EPJ Web of Conferences **49**, 12008 (2013)
DOI: 10.1051/epjconf/20134912008

Search for the standard model Higgs boson decaying to two photons with CMS at the LHC

Nancy Marinelli, [a] University of Notre Dame, IN, USA. On behalf of the CMS Collaboration.

Abstract. The search for the Higgs boson is presented on 10.4 fb^{-1} of data recorded with CMS at the Large Hadron Collider (LHC) during proton-proton interaction runs at $\sqrt{s}$ = 7 and 8 TeV. The search is performed in the mass range between 110 and 150 GeV and limits are set on the cross section in the context of the Standard Model. An excess of events above the expected background is observed at about 125 GeV which is consistent with the observation of a new boson statistically compatible with the Standard Model Higgs boson hypothesis.

1 Introduction

The Standard Model [1], [2], [3] has been so far one of the most successful theories ever formulated and received multiple experimental confirmations. The mechanism of the electroweak symmetry breaking (EWSB), responsible for the generation of the masses of elementary particles, however, is still in need of experimental proof. The EWSB is achieved by introducing in the theory a complex scalar doublet which leads to the prediction of the existence of the Higgs boson [4], [5], [6], [7], [8], [9].

The Standard Model (SM) predicts the existence of four different production mechanisms for the Higgs boson in proton-proton interactions, the gluon-gluon fusion being by far the dominant one. The other mechanisms, in order of decreasing cross sections are the vector-boson fusion (VBF), the associated production with a Z or W boson and the $t\bar{t}$ fusion. The Higgs boson can be sought for inclusively or exclusively in these production modes by looking for one of the predicted decay channels. The decay to two photons has always been expected to be of extreme relevance for a low-mass Higgs boson because despite the expected very small branching fraction (between 0.14% and 0.23%), it features an extremely clean signature due to a pair of isolated, medium-high transverse energy photons. The existence of the boson is hence expected to be revealed by the presence of a narrow, small peak over a continuum background mostly due to the irreducible direct two-photon QCD production.

The search for the Higgs boson described here was performed in the context of the Standard Model, in the mass range between 110 and 150 GeV, combining both the inclusive and the exclusive searches, the latter in the VBF production mode. Primary interaction vertex, photon selection and event classification were based on multivariate techniques (MVA); the result was extracted by a fit to the photon-photon invariant mass distribution, $m_{\gamma\gamma}$, in several event categories.

[a]e-mail: nancy.marinelli@cern.ch

Results shown here, and published in [10],[11], were obtained with the analysis of 5.1 fb^{-1} of data collected with CMS at $\sqrt{s}$ = 7 TeV and 5.3 fb^{-1} of data collected at $\sqrt{s}$ = 8 TeV. In this paper the analysis on the 8 TeV dataset is described ; details for the analysis on 7 TeV data are available in [10].

The Compact Muon Solenoid (CMS) detector is described in detail elsewhere [12]. The central feature of the CMS apparatus is a superconducting solenoid, of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the field volume are a silicon pixel and strip tracker, a lead-tungstate crystal electromagnetic calorimeter ECAL and a brass/scintillator hadron calorimeter HCAL. Muons are measured in gas-ionization detectors embedded in the steel return yoke. Extensive forward calorimetry complements the coverage provided by the barrel and endcap detectors.

2 Pillars of the analysis

The success of the search of the Higgs boson in the two-photon channel decay relies on four main points: high invariant-mass resolution, vital to distinguish a tiny peak on a huge background; effective photon identification, essential to separate prompt photons from photons secondary to neutral meson decays produced in jets; event categorization, effective with probing different signal-to-background or different mass resolution signal categories; and finally the background modeling, which after photon identification is mostly ($\sim$ 70%) irreducible, due to genuine isolated two-photon QCD events. The inter-channel calibration of the electromagnetic calorimeter (ECAL) as well as the corrections for the crystal transparency loss due to radiation, are the fundamental components to insure good energy resolution. Energy corrections are further applied to clustered electromagnetic energy in order to insure complete shower and converted photon containment as well as to correct the dependence on the pile-up event rate. The energy corrections are derived with a multivariate energy regression trained on simulated Higgs bo-

son signal. The validity of the results is checked with Z→ $e^+ e^-$ events.

The di-photon invariant mass resolution also depends on the opening angle between the two photons, which in turn is related to the directions of the two photons and the accuracy in the identification of the primary vertex where they originate. The natural spread in the longitudinal position of the interaction point is about 6 cm and and the uncertainty on the position of the primary vertex worsens as the pile-up event rate increases. In CMS it was checked that if the vertex is known within 1 cm there is no impact on the invariant mass resolution. Techniques involving multivariate methods, based on boosted decision trees, were hence developed in order to meet this target. A first multivariate analysis, MVA_1, gives high score to likely primary vertex according to the information on the event topology extracted from all tracks present in the event and from the pointing provided from converted photons if the $e^+ e^-$ tracks are reconstructed. A second multivariate analysis, MVA_2, is hence built from the score MVA_1 attributed to each vertex plus the $\gamma\gamma$-system transverse momentum, the number of vertices reconstructed in the event and the longitudinal distance between the highest MVA_1-score vertex and the next two. Both multivariate analyses are trained on simulated Higgs boson events and their validity is tested in Z→ $\mu^+\mu^-$ events and in γ+jet events, the latter to verify specifically events with converted photons. The primary vertex selection achieved is about 80% as shown in Fig. 1 as a function of the $\gamma\gamma$-system transverse momentum.

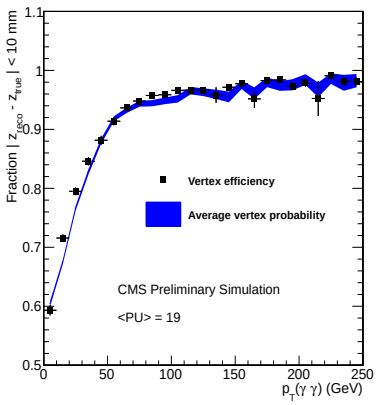

Figure 1. Comparison between the true vertex identification efficiency and the average estimated vertex probability as a function of the reconstructed di-photon p_T. The plots are obtained on Higgs boson signal events with m_H =120 GeV selected by the analysis.

Photon identification is based on isolation measured both in the calorimeters and in the tracking devices as well as on the shower shape and extension, the latter being the major discriminant between primary and secondary photons. Additional information are the photon pseudo-rapidity and the event density per unit area, ρ which is necessary to adjust the dependence on the pile-up event rate. The photon isolation was measured using the particle flow algorithm [13] information, which is by construction less sensitive to the increasing pile-up rate. All this information is combined into another multivariate analysis which attributes high score to genuine photon-like objects.

3 Analysis method

For the inclusive analysis, pairs of photon candidates are selected in each event on the base of the ratio between their transverse momentum and the $\gamma\gamma$-system invariant mass ($p^\gamma_{T1} > m_{\gamma\gamma}/3$ and $p^\gamma_{T2} > m_{\gamma\gamma}/4$). The information from the primary vertex and from the photon identification MVAs are combined into a final di-photon multivariate analysis which assigns high score to di-photon candidates with signal-like kinematics, good mass resolution, high probability of correct vertex identification and with each photon having high score from the identification. The distribution of the MVA output is shown in Fig. 2 for simulated signal and background.

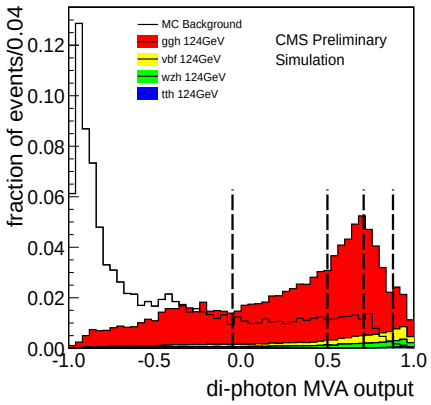

Figure 2. Di-photon ID MVA output distribution for simulated signal and background. The region below -0.05, where the background dominates, is not used in the analysis and the vertical dashed lines indicate the class boundaries.

For the inclusive analysis, candidates are divided in four categories according to the MVA output with boundaries chosen as shown in Fig. 2. The boundaries are optimized to achieve the best expected limit and the region below -0.05, where the background dominates and gives negligible gain in sensitivity, is not used in the analysis. The MVA is trained on simulated Higgs boson ($m_{\gamma\gamma}$ = 123 GeV) events and γ+jet events and validated on Z→e^+e^- data and simulation. The exclusive analysis in the VBF production channel exploits the forward jet tagging: the highest energy jets in the event with large pseudo-rapidity gap are selected to enrich the VBF content and reject the background. The di-jet mass, m_{jj}, is required to be greater than 250 GeV. The photons are subjected to the same requirements as in the inclusive analysis except a harder cut on the transverse momentum of the leading photon, $p^\gamma_{T1} > m_{\gamma\gamma}/2$.

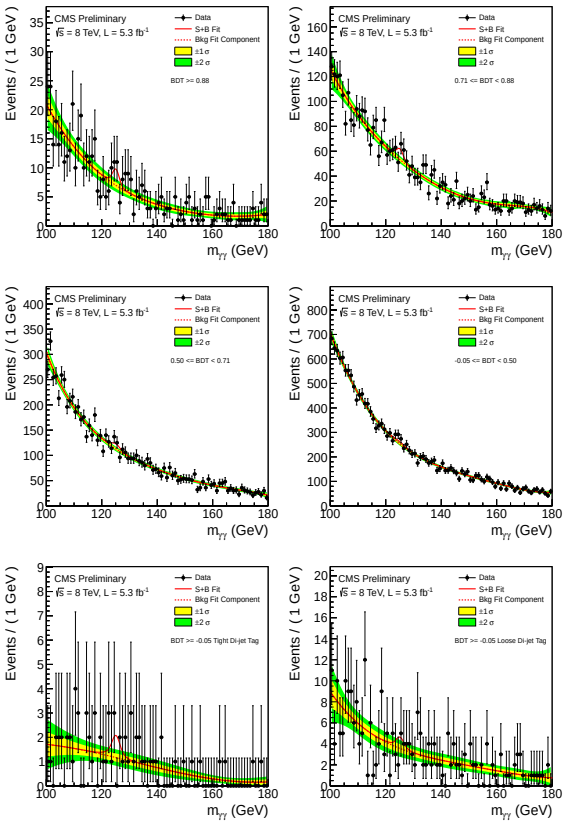

Figure 3. The result of signal-plus-background model fit to the $m_{\gamma\gamma}$ distribution for the six event categories of the 8 TeV dataset. The first four distributions are for the inclusive analysis. From left to right and top to bottom they are in decreasing signal-to-background order. The last two plots are for the two VBF categories, with respectively tight and loose m_{jj} cut.

Two more categories are defined according to m_{jj}; they have a signal-to-background ratio an order of magnitude larger than the inclusive analysis leading to about 10% increase in the final sensitivity. Table 1 shows the number of expected signal events from a SM Higgs boson with m_H = 125 GeV as well as the estimated background at $m_{\gamma\gamma}$ =125 GeV for each of the eleven classes in the 7 and 8 TeV datasets.

The final result is extracted by a fit to the signal plus background $m_{\gamma\gamma}$ distribution in the six categories (five for the 7 TeV dataset). The background level is measured from data in the mass range between 110 and 180 GeV; it is modeled by fitting data with polynomial functions with order ranging from three to five (Fig. 3). Care was taken to verify that the potential level of bias introduced in the analysis was below 20% of the statistical uncertainty.

4 Results

The statistical interpretation and the limit extraction is based on the asymptotic CL_S [14] using the profile likelihood ratio as a test-statistic [15]. The combination of the results obtained with the 7 and 8 GeV datasets has been

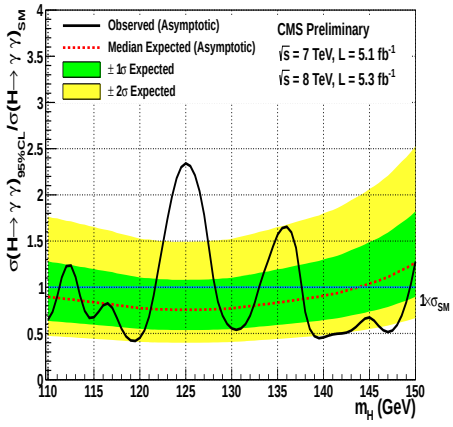

Figure 4. 95% CL limits on the cross section of a Higgs boson decaying to two photons relative to the SM expectation for the combined 7 and 8 TeV datasets.

achieved by considering the analyses on the two datasets as two separate sub-channels. The uncertainties include theoretical uncertainties on the expected cross sections and acceptances for signal and background processes, experimental uncertainties on the modeling of the detector response (event reconstruction and selection efficiencies, energy scale and resolution). All known sources of systematic uncertainties are included in the likelihood model which is used for the limit setting. Systematic errors which are correlated between event classes (theory, luminosity, photon and trigger efficiency, etc) are modeled as common nuisance parameters.

Figure 4 shows the 95% CL exclusion limits combined for the 7 and 8 TeV data samples. The Higgs boson exclusion was expected in the mass range between 110 and 143 GeV. The observed limit shows the presence of a large excess at about 125 GeV and excludes the Higgs boson at 95% CL in the three mass ranges 114 to 121 GeV, 129 to 132 GeV and 138 to 149 GeV. The corresponding local p-values observed are shown in Fig. 5; the minimum, corresponding to the largest upward fluctuation of the observed limit, falls at about 125 GeV, with a value of 1.8×10^{-5} corresponding to 4.1σ local significance. The global significance in the whole mass range 110 to 150 GeV is 3.5 σ. This result is consistent with the observation of a new boson with mass $\sim$ 125 GeV.

The combined best-fit signal strength, σ/σ_{SM}, for a Standard Model Higgs boson with mass 125 GeV gives $\sigma/\sigma_{SM} = 1.56 \pm 0.43$ (Fig. 6). The results for the different signal categories are reported in Fig. 7 which shows the compatibility, within the uncertainty, across all categories and data samples.

The mass of the observed boson is measured to be $125.1 \pm 0.4(stat) \pm 0.6(syst)$. The calibration of the energy scale is performed at the peak of the Z boson mass. The uncertainty on the mass value is dominated by the systematic uncertainty which arises mostly from the imperfect simulation of the differences between photons and elec-

Table 1. Expected number of SM Higgs boson events (m_H=125 GeV) and estimated background (at $m_{\gamma\gamma}$=125 GeV) for all event classes in the 7 and 8 TeV datasets. The composition of the SM Higgs boson signal in terms of the production processes and its mass resolution is also given.

Expected signal and estimated background									
Event classes		SM Higgs boson expected signal (m_H=125 GeV)							Background $m_{\gamma\gamma}$ = 125 GeV (ev./GeV)
		Total	ggH	VBF	VH	ttH	σ_{eff} (GeV)	FWHM/2.35 (GeV)	
7 TeV 5.1 fb^{-1}	Untagged 0	3.2	61%	17%	19%	3%	1.21	1.14	3.3 ± 0.4
	Untagged 1	16.3	88%	6%	6%	1%	1.26	1.08	37.5 ± 1.3
	Untagged 2	21.5	91%	4%	4%	–	1.59	1.32	74.8 ± 1.9
	Untagged 3	32.8	91%	4%	4%	–	2.47	2.07	193.6 ± 3.0
	Dijet tag	2.9	27%	73%	1%	–	1.73	1.37	1.7 ± 0.2
8 TeV 5.3 fb^{-1}	Untagged 0	6.1	68%	12%	16%	4%	1.38	1.23	7.4 ± 0.6
	Untagged 1	21.0	88%	6%	6%	1%	1.53	1.31	54.7 ± 1.5
	Untagged 2	30.2	92%	4%	3%	–	1.94	1.55	115.2 ± 2.3
	Untagged 3	40.0	92%	4%	4%	–	2.86	2.35	256.5 ± 3.4
	Dijet tight	2.6	23%	77%	–	–	2.06	1.57	1.3 ± 0.2
	Dijet loose	3.0	53%	45%	2%	–	1.95	1.48	3.7 ± 0.4

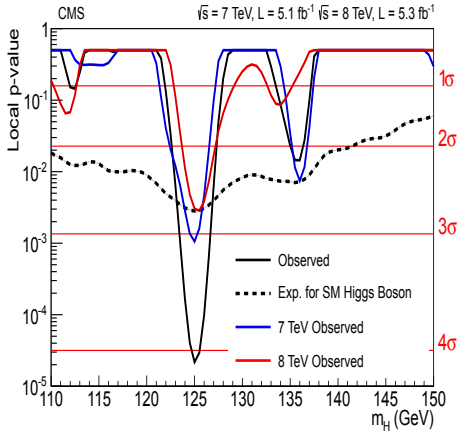

Figure 5. The local p-value as a function of m_H for the combined 7 and 8 TeV data sets. The additional lines show the values for the two data sets taken individually. The dashed line shows the expected local p-value for the combined data sets, should a SM Higgs boson exist with mass m_H.

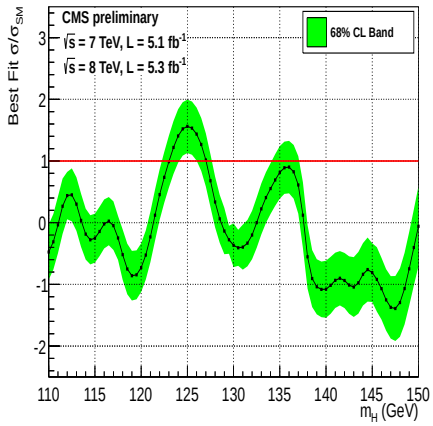

Figure 6. The best fit signal strength, σ/σ_{SM}, shown as a function of the Higgs boson mass hypothesis.

trons and from the necessary extrapolation from the Z boson mass to $m_H \sim 125$ GeV.

Finally, and only for visualization purposes, the $m_{\gamma\gamma}$ distribution was produced by summing the contributions from all event categories, re-weighting each of them by the corresponding signal-to-background ratio. The result is shown in Fig. 8, where a clear peak over the smoothly falling down background, is visible at about 125 GeV. The inset also shows the un-weighted distribution. Figure 9 shows the same distribution after background subtraction.

5 Conclusions

The Higgs boson search in the context of the Standard Model interpretation was performed on a total of 10.4 fb^{-1} of data collected with CMS at the LHC during the pp runs at $\sqrt{s}$=7 and 8 TeV. And excess of events above the SM predictions was observed at about 125 GeV, consistently in 7 and 8 TeV data. From the combined analysis on the 7 and 8 TeV data the observed limit is 2.3×SM, where the expected limit is 0.76×SM, corresponding to a local significance of 4.1 σ and to a global significance of 3.2 σ. The result is evidence of the existence of a new boson decaying to two photons, compatible, within experimental uncertainties with the Standard Model Higgs hypothesis. The observation of the decay in the two photon channel allows already to exclude the spin-1 nature of the new state. The mass of the new boson measured from the two-photon decay channel is $125.1 \pm 0.4(stat) \pm 0.6(syst)$.

References

[1] S.L. Glashow, "Partial Symmetries of Weak Interactions", Nucl. Phys. 22 (1961) 579

[2] S. Weinberg, "A Model of Leptons", Phys. Rev. Lett. 19 (1967) 1264

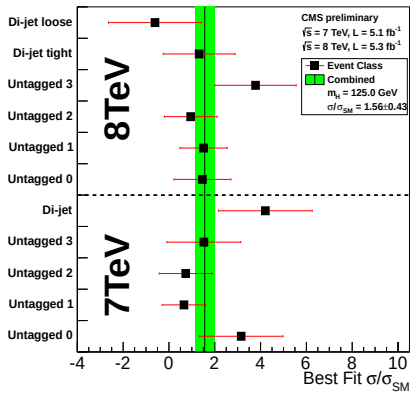

Figure 7. The best fit signal strength, σ/σ_{SM}, for the combined fit to the eleven classes (vertical line) and broken down to the individual contributing classes (points) for the hypothesis of a SM Higgs boson mass of 125.0 GeV. The band corresponds to ±1 standard deviation uncertainty on the overall value. The horizontal bars indicate ±1 standard deviation uncertainty on the values for individual classes.

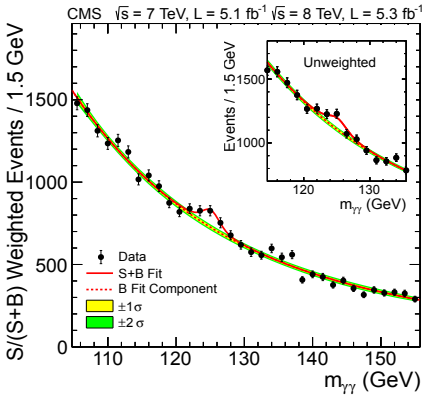

Figure 8. The di-photon invariant mass distribution with each event weighted by the S/(S+B) value of its category. The lines represent the fitted background and signal, and the coloured bands represent the ±1 and ±2 standard deviation uncertainties in the background estimate. The inset shows the central part of the unweighted invariant mass distribution.

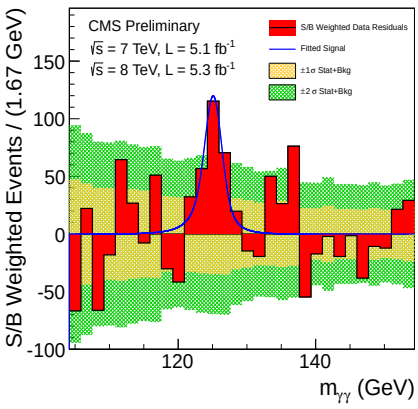

Figure 9. The di-photon invariant mass distribution with each event weighted by the S/(S+B) value of its category as in Fig. 8 but with background subtraction. The uncertainty bands correspond to the sum in quadrature of the estimated uncertainties on the background and the weighted data.

[3] A. Salam, "Weak and electromagnetic interactions", Elementary particle physics: relativistic groups and analyticity, N. Svartholm, ed., Almquvist and Wiskell (1968) 367, Proceedings of the eighth Nobel symposium.

[4] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", Phys. Rev. Lett. 13 (1964) 321.

[5] P.W. Higgs, "Broken symmetries, massless particles and gauge fields", Phys. Rev. Lett. 12 (1964) 132.

[6] P.W. Higgs, "Broken symmetries and the masses of gauge bosons", Phys. Rev. Lett. 13 (1964) 508.

[7] G.S. Guralnik, C.R. Hagen, and T.W.B. Kibble, "Global conservation laws and massless particles", Phys. Rev. Lett. 13 (1964) 585.

[8] P.W. Higgs, "Spontaneous symmetry breakdown without massless bosons", Phys. Rev. 145 (1966) 1156.

[9] T.W.B. Kibble, "Symmetry breaking in non-Abelian gauge theories", Phys. Rev. Lett. 155 (1967) 1554.

[10] The CMS Collaboration, "Search for the standard model Higgs boson decaying into two photons in pp collisions at $\sqrt{s}$=7 TeV", Phys. Lett. B710, Vol 3, 403 (2012)

[11] The CMS Collaboration,"Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC" Phys. Lett. B716, 30-61 (2012)

[12] The CMS Collaboration, "The CMS experiment at the CERN LHC" JINST 3 (2008) S08004

[13] The CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", CMS-PAS-PFT-09-001.

[14] G. Cowan, K. Cranmer, E. Gross and O. Vitells "Asymptotic formulae for likelihood-based tests of new physics" Eur. Phys. J. C71 (2011) 1554.

[15] ATLAS and CMS Collaborations, "Procedure for the LHC Higgs boson search combination in Summer 2011" ATL-PHYS-PUB-2011-11, CMS NOTE-2011/005, CDS Record 1363354 (2011)