Search for a Standard Model Higgs Boson decaying into two photons at CMS

Serguei Ganjour∗†
CEA-Saclay/IRFU, Gif-sur-Yvette, France
E-mail: ganzhur@cea.fr

A search for a Higgs boson decaying into two photons in proton-proton collisions at centre-of-mass energies of 7 TeV and 8 TeV is reported. The data analysed correspond to an integrated luminosity of $3.1\text{ fb}^{-1}$, recorded in 2011 at $\sqrt{s}=7\text{ TeV}$ and of $5.3\text{ fb}^{-1}$, recorded in 2012 at $\sqrt{s}=8\text{ TeV}$ by the CMS experiment at the LHC. The Higgs boson mass range explored is 110–150 GeV. Limits are set on the cross section of the Standard Model Higgs boson decaying to two photons. An excess of events above the expected background is observed for a Higgs boson mass hypothesis of 125 GeV, with a local significance of 4.1σ. The global significance of observing such an excess anywhere in the search range is estimated to be 3.2σ. This result constitutes evidence for the existence of a new state that decays into two photons. A fit to this signal gives a mass of $125.1\pm0.4(\text{stat})\pm0.6(\text{syst})$ GeV. The decay to two photons indicates that the new particle is a boson with spin different from one.
1. Introduction

Despite the great success of the standard model (SM) [1, 2, 3], one of its fundamental principals, electroweak symmetry breaking (EWSB) mechanism, requires experimental confirmation. In the SM, the EWSB responsible for generating the masses of the gauge bosons is achieved by introducing a complex scalar doublet, leading to the prediction of the Higgs boson [4, 5, 6, 7, 8, 9]. Precision electroweak measurements, not taking into account the results from direct searches, constrain the Higgs boson mass to be less than 158 GeV [10]. To date, the direct experimental searches exclude at 95% CL the SM Higgs boson in the mass range below 600 GeV, except for a small mass window 122-128 GeV [11, 12, 13, 14].

The $H \rightarrow \gamma\gamma$ decay channel provides a clean final-state signature with two high $E_T$ isolated photons. Hence, the signal can be identified as a narrow peak at a Higgs boson mass on top of a continuous background. A good resolution of CMS electromagnetic calorimeter (ECAL) would allow the mass to be reconstructed with high precision. In the mass range $110 < m_H < 150$ GeV, $H \rightarrow \gamma\gamma$ is one of the discovery channels for a Higgs boson at the LHC [15, 16]. The primary production mechanism of the Higgs boson at the LHC is gluon fusion with additional contributions from vector boson fusion (VBF) and production in association with a W or Z boson, or a $t\bar{t}$ pair. The decay of the Higgs boson into two photons is induced via virtual loops, with the W boson and top quark being the dominant contributors. The SM $H \rightarrow \gamma\gamma$ branching fraction varies between 0.14% and 0.23% [17]. Previous searches for the SM Higgs boson in this channel have been conducted by ATLAS and CMS experiments at the LHC [18, 19]. Likewise the production cross sections the Higgs boson branching fraction into two photons is strongly affected in beyond the SM phenomena. Example of such searches can be found in Refs. [20, 21].

The results presented here build upon the searches reported previously [19, 22], and take advantage of updated detector calibration constants for the 7 TeV dataset and improved analysis methods for the 8 TeV dataset to overcome the increase in the number of simultaneous collisions (pile-up). Throughout this report emphasis has been placed on the description of the analysis of the 8 TeV dataset corresponding to an integrated luminosity of 5.3 fb$^{-1}$ recorded in 2012. Finally, the combined result with 7 TeV data, for the SM Higgs search in the $H \rightarrow \gamma\gamma$ decay mode is presented. The reported result corresponds to a total integrated luminosity of 10.4 fb$^{-1}$ and is described in detail elsewhere [23].

The Compact Muon Solenoid (CMS) detector is described in detail elsewhere [24]. The central feature of the CMS apparatus is a superconducting solenoid, of 6 m internal diameter, providing a 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. Event Selection and Classes

Events are collected with a set of diphoton triggers which use looser identification criteria than those used in the analysis and $p_T$ thresholds of 26 and 18 GeV, yielding signal efficiency of almost 100%. The event selection requires two photon candidates with $p_T^\gamma(1) > m_{\gamma\gamma}/3$ and
\( p_T^\gamma(2) > m_{\gamma\gamma}/4 \) within the ECAL fiducial region: \(|\eta| < 2.5\) excluding the barrel-endcap transition region \(1.44 < |\eta| < 1.57\). Photon identification using a boosted decision trees (BDT) is applied after a preselection, to distinguish prompt photons from non-prompt photon background due to electromagnetic showers originating in jets – mainly due to single and multiple \( \pi^0 \)'s. The electromagnetic shower width, electron veto and the ratio of hadronic energy in HCAL towers behind the supercluster to the ECAL energy of the supercluster are included in preselection requirements. The BDT is trained using simulated pp \( \to \gamma+\)jet event samples. Prompt photons are considered as the signal, while non-prompt photons emerged from jet are assumed as the background. Shower topology, isolation, supercluster pseudorapidity and the event density per unit area, \( \rho \) are used as input to the photon identification BDT. The later quantity \( \rho \) is introduced to adjust the pile-up dependence in the isolation variables, computed based on the particle flow algorithm.

The Higgs boson candidate is formed with two photons satisfying the preselection requirements, including a requirement on the photon identification BDT output. An event classifier variable (diphoton BDT) is then constructed based on the multivariate discriminator. This variable is designed to be mass independent and classifies with a high score events with: signal-like kinematic characteristics, good diphoton mass resolution events, photon-like values from the photon identification BDT. The diphoton BDT incorporates the kinematic properties of the diphoton system (excluding \( m_{\gamma\gamma} \)), a per-event estimate of the diphoton mass resolution, and a per-photon identification measure (the photon identification BDT output value). It is trained using simulated MC background and Higgs boson signal events (\( m_H=123 \) GeV).

The analysis is performed in the mutually exclusive sub-channels: the untagged and the dijet tag. The diphoton BDT output value provides an indication of signal-to-background, including the effect of mass resolution. This output value has been used to divide events into classes prior to fitting the diphoton invariant mass spectrum. The events are split into classes by introducing boundary values that give rise to the best expected exclusion limit. The mass value of 124 GeV is used to optimize the boundary values. Lowest score events \(< -0.05 \) are dropped as they give negligible (< 1%) gain in sensitivity. Events having score above \(-0.05 \) are split beyond four classes.

Additional classification of events into dijet-tagged classes enriches the VBF content and improves the overall sensitivity by about 10%. Presence of two jets well-separated in rapidity significantly reject backgrounds. Candidate diphoton events for the dijet-tagged classes have the same selection requirements imposed on the photons as for the other classes with the exception of the \( p_T \) threshold on the leading photon, which is increased to \( p_T^\gamma(1) > m_{\gamma\gamma}/2 \), to reduce background with negligible loss of acceptance. The selection variables for the jets use the two highest \( p_T \) jets in the event with pseudorapidity \(|\eta| < 4.7\). The selection requirements are optimized to obtain the best expected limit at 95% CL on the VBF signal cross section using fully simulated VBF signal events and the diphoton background estimation from data. The \( p_T \) thresholds for the two jets are 30 GeV and 20 GeV, and the pseudorapidity separation between them is required to be greater than 3.0. The dijet mass \( m_{jj} \) is required to be greater than 250 GeV. The selected events are subdivided into two classes according to amount of background contamination using the dijet mass. The class boundary value \( m_{jj} = 500 \) GeV is chosen. In addition for the high dijet mass events, the \( p_T \) threshold for the subleading jet is required to be 30 GeV. Two additional selection criteria, relating the dijet to the diphoton system, are applied for all selected events: the difference between the average pseudorapidity of the two jets and the pseudorapidity of the diphoton system is required to be less
than 2.5 [25], and the difference in azimuthal angle between the diphoton system and the dijet system is required to be greater than 2.6 radians. Events passed the dijet tag selections are excluded from untagged classes, so a statistically independent search is performed.

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 table also shows the fraction of each Higgs boson production process (as predicted by MC simulation) as well as the mass resolution, represented both as $\sigma_{\text{eff}}$, half-the-width of the narrowest interval containing 68.3% of the distribution, and as the full width at half maximum (FWHM) of the invariant mass distribution divided by 2.35.

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 of 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.

<table>
<thead>
<tr>
<th>Event classes</th>
<th>SM Higgs boson expected signal ($m_H = 125$ GeV)</th>
<th>Background $m_{\gamma\gamma} = 125$ GeV (ev./GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>ggH</td>
</tr>
<tr>
<td>7 TeV 5.1 fb$^{-1}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untagged 0</td>
<td>3.2</td>
<td>61.4%</td>
</tr>
<tr>
<td>Untagged 1</td>
<td>16.3</td>
<td>87.6%</td>
</tr>
<tr>
<td>Untagged 2</td>
<td>21.5</td>
<td>91.3%</td>
</tr>
<tr>
<td>Untagged 3</td>
<td>32.8</td>
<td>91.3%</td>
</tr>
<tr>
<td>Dijet tag</td>
<td>2.9</td>
<td>26.8%</td>
</tr>
<tr>
<td>8 TeV 5.3 fb$^{-1}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untagged 0</td>
<td>6.1</td>
<td>68.2%</td>
</tr>
<tr>
<td>Untagged 1</td>
<td>21.0</td>
<td>87.6%</td>
</tr>
<tr>
<td>Untagged 2</td>
<td>30.2</td>
<td>92.1%</td>
</tr>
<tr>
<td>Untagged 3</td>
<td>40.0</td>
<td>91.6%</td>
</tr>
<tr>
<td>Dijet tight</td>
<td>2.6</td>
<td>22.8%</td>
</tr>
<tr>
<td>Dijet loose</td>
<td>3.0</td>
<td>53.1%</td>
</tr>
</tbody>
</table>

3. Signal and Background Modeling

The description of the Higgs boson signal is obtained from samples of Monte Carlo (MC) events which are fully simulated using GEANT4 [26]. The gluon fusion and VBF productions of the SM Higgs signal are simulated using the next-to-leading order (NLO) matrix-element generator POWHEG [27, 28] interfaced with PYTHIA [29]. The production in association with a W or Z boson and a $t\bar{t}$ pair is simulated with PYTHIA. For the dominant gluon-gluon fusion process, the Higgs boson transverse momentum spectrum has been reweighted to the next-to-next-to-leading logarithmic (NNLL) + NLO distribution computed by the HQT program [30, 31, 32]. The gluon-gluon fusion process cross-section is reduced by 2.5% for all values of $m_H$ to account for interference with background diphoton final states [33]. The magnitude of the variation of the interference effect in our acceptance is smaller than the systematic uncertainty on the effect. The simulated
events are reweighted to reproduce the distribution of the number of interactions taking place in each bunch crossing. About 20 multiple collisions in each bunch crossing was measured in 2012 data which is about a factor of 2 more than in 2011 data.

The energy scale corrections are derived from $Z^0 \rightarrow e^+e^-$ data events by comparing the di-electron invariant mass peak position in the data and MC simulated events. Once energy scale corrections are applied for the data, the amount of random energy smearing is measured from the width difference of the $Z^0$ boson peak in data and MC simulation. The photon energy corrections and smearing factors are derived in four pseudo-rapidity regions (two in the barrel and two in the endcap) and two $R_9$ categories ($R_9$ variable, defined as the energy sum of 3x3 crystals centered on the crystal with maximum energy deposit divided by the total clustered energy) separately and per photon object. The uncertainties of these corrections and factors including the difference between photons and electrons are taken as systematic errors in the limit setting procedure.

The background model is derived from data, by fitting the diphoton mass distributions over the range $100 < m_{\gamma\gamma} < 180$ GeV. The choice of the function used to fit the background and the choice of the range are made based on a study of the possible bias introduced by the two choices. Bias studies are performed using a number of generated pseudo-experiments with background only and signal plus background hypotheses. The maximum bias is found to be at least five times smaller than the statistical uncertainties of the fit. The polynomial model was finally chosen. The required order of polynomial ranges from 3 to 5.

4. Results

The statistical approach considered in evaluating the limit is the asymptotic CL$_S$ [34] using the profile likelihood ratio as a test-statistic [35]. Given the narrowness of the Higgs mass peak in $H \rightarrow \gamma\gamma$ channel, which has a resolution approaching 1 GeV in the classes with the best resolution, the search is carried out with steps of 0.5 GeV in the signal hypothesis mass in range between 110 and 150 GeV.

For the combined search results of the $H \rightarrow \gamma\gamma$ decay mode, the analyses of the 7 TeV and 8 TeV datasets are taken as separate sub-channels. The six sub-channels analyzed with 8 TeV dataset and described in Section 2, and five sub-channels analyzed with 7 TeV dataset and described in Ref. [22], are combined using the techniques from Ref. [36] to account for all statistical and systematic uncertainties and their correlations. The uncertainties include: theoretical uncertainties on the expected cross sections and acceptances for signal and background processes, experimental uncertainties in 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 1 (left) shows the limit, at 95% confidence level (CL), for the combined datasets relative to the SM model expectation. The expected sensitivity (dashed line) allows exclusion of the SM Higgs boson in the mass range 110–144 GeV. A prominent excess can be seen in the observed limit (solid line) at around $m_H=125$ GeV. Figure 1 (right) shows the combined local $p$-value, calculated using the asymptotic approximation. The local $p$-value quantifies the probability for the
Search for SM Higgs decaying into two photons at CMS

Serguei Ganjour

Figure 1: (Left) The observed and expected 95% CL combined upper limit on the signal strength, $\sigma/\sigma_{SM}$, of a SM Higgs boson, as a function of the Higgs boson mass. (Right) The observed and expected $p$-values as a function of the Higgs boson mass. Contributions of each period are also shown.

background to produce a fluctuation at least as large as observed, and assumes that the relative signal strength between the event classes follows the MC signal model for the SM Higgs boson. The local $p$-value corresponding to the largest upwards fluctuation of the observed limit, at 125 GeV, has been computed to be $1.8 \times 10^{-5}$, corresponding to a global significance of 4.1σ. Since a fluctuation of the background could occur at any point in the mass range, there is a look-elsewhere effect [37]. When this is taken into account, the probability under the background only hypothesis of observing a similar or larger excess in the full analysis mass range 110–150 GeV is $7.2 \times 10^{-4}$ (3.2σ). The excess thus constitutes evidence of a new particle with a mass near 125 GeV. As this new particle decays into photons it must be a boson with spin different from one.

Figure 2 shows the combined best fit signal strength as a function of the Higgs boson mass hypothesis and compared to the best fit signal strengths in each of the event classes at $m_H = 125$ GeV. At this mass the combined best fit signal strength is $1.56 \pm 0.43$ times the SM Higgs boson cross section.

Figure 3 (left) shows the result of performing a weighted sum over all event classes, where the weights are the ratio of signal to background as derived from Table 1. The weighted data, the weighted signal model, and the weighted background model are normalized such that the integral of the weighted signal model matches the number of signal events from the best fit. The bin size of the weighted distributions is chosen to match the $\sigma_{eff}$ of the weighted signal model and the bins are centered at the best fit $m_H$ value. Figure 3 (right) shows the weighted data after subtracting the background model. The uncertainty shown around the horizontal axis corresponds to the sum in quadrature of the estimated uncertainties on the weighted background model and the weighted data.

The mass of the observed boson is measured to be $125.1 \pm 0.4(\text{stat.}) \pm 0.6(\text{syst.})$ GeV. The calibration of the energy scale is achieved with reference to the known Z boson mass. The systematic uncertainty includes two dominant sources: imperfect simulation of the differences between electrons and photons and uncertainty of the kinematic extrapolation from $m_Z$ to $m_H \approx 125$ GeV. The two uncertainties, which together amount to 0.5%, are assumed to be fully correlated between all the event categories in the 7 and 8 TeV data.
Search for SM Higgs decaying into two photons at CMS

Serguei Ganjour

Figure 2: The best fit signal strength relative to the SM Higgs boson cross section. (Left) Scan as a function of the Higgs boson mass hypothesis. (Right) The combined fit to the eleven classes (vertical line) and for the individual contributing classes (points with error bars) for the hypothesis of a SM Higgs boson mass of 125 GeV. The band corresponds to $\pm 1\sigma$ uncertainties on the overall value. The horizontal bars indicate $\pm 1\sigma$ uncertainties on the values for individual classes.

Figure 3: (Left) S/B-weighted $m_{\gamma\gamma}$ distribution for the data and the signal-plus-background fits. (Right) Background-subtracted S/B-weighted $m_{\gamma\gamma}$ distribution. The uncertainties shown around the horizontal axis correspond to the sum in quadrature of the estimated uncertainties on the background and the weighted data.
5. Conclusion

A search has been performed for the SM Higgs boson decaying into two photons using data obtained from 5.1 fb$^{-1}$ of pp collisions at $\sqrt{s} = 7$ TeV and 5.3 fb$^{-1}$ at $\sqrt{s} = 8$ TeV. The expected sensitivity on the limit on the production cross section times branching fraction, at 95% confidence level, is between 0.7 and 1.3 times the standard model prediction in the mass range 110–150 GeV. An excess of events above the expected SM background is observed for a Higgs boson mass hypothesis of 125 GeV, where the expected limit is 0.76 times the SM expectation. The local significance of this excess is 4.1\,\sigma. The global significance of observing an excess with a local significance $\geq 4.1\,\sigma$ anywhere in the search range 110–150 GeV is estimated to be 3.2\,\sigma. This result constitutes evidence for the existence of a new massive state that decays into two photons. Within the experimental uncertainties, the result is consistent with the expectations for a SM Higgs boson. The decay to two photons indicates that the new particle is a boson with spin different from one. More data are needed to test whether the properties of this new state are indeed those of the SM Higgs boson or whether some differ, implying new physics beyond the standard model.

References

Search for SM Higgs decaying into two photons at CMS

Serguei Ganjour


[21] CMS Collaboration, “Search for the fermiophobic model Higgs boson decaying into two photons in pp collisions at √s = 7 and 8 TeV” CMS HIG-12-022, CDS Record 1461397 (2012)


[36] CMS Collaboration, "Combined results of searches for the Higgs boson in pp collisions at $\sqrt{s} = 7$ and 8 TeV" CMS HIG-12-020, CDS Record 1460438 (2012)