Search for the Higgs boson in the diphoton decay channel with ATLAS detector

Marcos Jimenez Belenguer[1,a]

[1] *On behalf of the ATLAS Collaboration*

Abstract. This document reports on a search for the Standard Model Higgs boson in the diphoton decay channel in proton-proton collisions at center-of-mass energies of $\sqrt{s}$ =7 TeV and $\sqrt{s}$ = 8 TeV using integrated luminosities of 4.8 fb^{-1} and 5.9 fb^{-1}, respectively, recorded with the ATLAS detector at the Large Hadron Collider. The search is performed for Higgs boson masses between 110 and 150 GeV. The expected exclusion limit at 95% confidence level varies between 0.8 and 1.6 times the Standard Model cross section over the studied mass range, and results in an expected exclusion range from 110 GeV to 139.5 GeV. The observed exclusion ranges for a Standard Model Higgs boson are found to be (112-122.5) GeV and (132-143) GeV at 126.5 GeV with a local significance of 4.5 σ. An excess is observed at 126.5 GeV with a signal strength of 1.8 over the SM expectation.

1 Introduction

The search for the Higgs boson is one of the principal tasks of the LHC and is part of the effort to unravel the mechanism of electroweak symmetry breaking within the SM. These proceedings present the analysis strategy and the results of the search for the SM Higgs boson in the diphoton decay channel using 4.8 fb^{-1} of 7 TeV and 5.9^{-1} of 8TeV data collected with the ATLAS detector. The analysis closely follows that of the July publication [1], with some minor changes in the way the photon energy and mean number of interactions per bunch crossing (pileup) are treated in the MC simulation. The results presented here were combined with those of SM Higgs seaches in other decay channels, and published by ATLAS in November 2012 [3].

2 The ATLAS detector and data sample

The ATLAS detector is described in detail elsewhere [2]. After the application of data-quality requirements, the data samples amount to 4.8 fb^{-1} of 7 TeV and 5.9^{-1} of 8 TeV data. The mean number of interactions per bunch crossing was of 9.1 in the data sample acquired during 2011, and of 19.5 for the data taken up to June 2012. The simulation is corrected to reflect the distribution of interactions per bunch crossing and the spread of the z position of the primary vertex observed in data. The data sample considered in this analysis was selected using a diphoton trigger. In the last step of the triggering chain, two clusters formed from energy depositions in the electromagnetic calorimeter are required, which fulfill loose criteria on the shapes

of the electromagnetic clusters. The trigger has an efficiency greater than 99% for events passing the final event selection.

3 Reconstruction and selection of $H \rightarrow \gamma\gamma$ candidates

The photon reconstruction is seeded from clusters of energy deposits in the electromagnetic calorimeter, which, for converted photon candidates, are matched to tracks and to conversion vertex candidates reconstructed in the inner detector. The energies of the clusters are calibrated, separately for unconverted and converted candidates, to account for energy losses upstream of the calorimeter and for energy leakage outside of the cluster. The calibration is refined by applying η-dependent correction factors, which are of the order of $\pm1\%$, determined from $Z \rightarrow e^+e^-$ events. The simulation is corrected to reflect the energy resolution observed using $Z \rightarrow e^+e^-$ events in the data, which requires an energy smearing of about 1% in the calorimeter barrel region and between 1.2% and up to 2.1% in the calorimeter end-caps. The identification of photons is based on shower shapes measured in the electromagnetic calorimeter. An initial loose cut-based selection, used also at the trigger level, is based on shower shapes in the second layer of the electromagnetic calorimeter, as well as the energy deposition in the hadronic calorimeter. A tight identification adds information from the finely segmented strip layer of the calorimeter, which provides good rejection of hadronic jets where a neutral meson carries most of the jet energy. Two variants of the tight photon identification are used. For the $\sqrt{s}$ =7 TeV data, a neural-network based selection is used. For

[a] e-mail: mjimenez@mail.cern.ch

the $\sqrt{s}$ =8 TeV data, a cut-based selection is used, which has been tuned for robustness against pile-up effects. To further suppress hadronic background a cut on the isolation transverse energy, defined as the sum of the transverse energy of positive-energy three-dimensional clusters reconstructed in the electromagnetic and hadronic calorimeters in a cone of $\Delta R = 0.4$ around the photon candidate, is applied. The isolation is corrected for leakage of the photon energy outside of the excluded region around the photon barycenter and for contributions from pile-up as well as the underlying event, on an event-by-event basis. Events are required to contain at least two reconstructed photon candidates in the fiducial region of the calorimeter, $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. The barrel-endcap transition regions,$1.37 < |\eta| < 1.52$ are excluded. To ensure well-reconstructed photon candidates, further quality requirements are applied to the reconstructed clusters. Similarly, converted photon candidates reconstructed from tracks passing through dead modules of the innermost pixel layer are rejected, strongly decreasing the misidentification of electrons as converted photons. Further criteria are applied to the two highest-pT photon candidates. The leading (subleading) photon candidate is required to have $p_T > 40(30)$ GeV. Tight identification criteria are applied to both photon candidates. Furthermore, both photon candidates are required to have an isolation transverse energy of less than 4 GeV. With this selection, 23788(35251) events are observed in the diphoton invariant mass range between 100 and 160 GeV in the $\sqrt{s}$ =7(8) TeV data sample. The primary vertex of the hard interaction is identified by combining the following elements in a global likelihood: the directions of flight of the photons as determined by the measurements using the longitudinal segmentation of the calorimeter, the average beam spot position, and the $\sum p_T^2$ of the tracks associated with each reconstructed vertex. In case of the $\sqrt{s}$ =7 TeV data, the conversion vertex is also used in the likelihood for converted photons with tracks containing silicon hits. The calorimeter information alone already provides a diphoton mass resolution very close to the optimal value which is based on the true hard scattering primary vertex position. The addition of the tracking information from the inner detector is necessary to improve the identification of the hard-interaction primary vertex needed for the jet selection. Jets are reconstructed from three-dimensional clusters of energy in the electromagnetic and hadronic calorimeters using the anti-kt algorithm [4] with a distance parameter of R = 0.4. Jet candidates are required to have a transverse momentum of greater than 25 GeV (30 GeV) for $|\eta^{jet}| < 2.5(|\eta^{jet}| > 2.5)$. Jets within $|\eta^{jet}| < 2.5$ must fulfill a requirement, based on tracking information, that they originate from the diphoton production vertex ("jet-vertex fraction requirement").

4 Event categorization

The selected events are classified into 10 exclusive categories, which differ in the invariant mass resolution as well as in the signal-to-background ratio and thus increase the sensitivity of the search. One category is dedicated
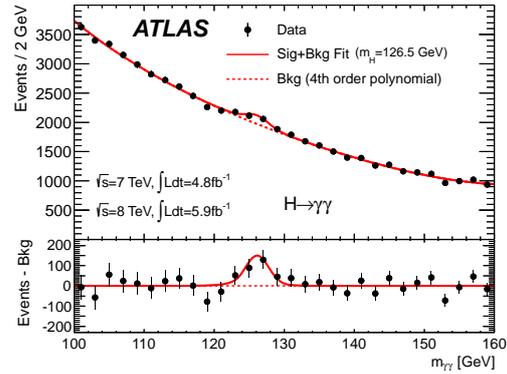

Figure 1. Invariant mass distribution for the combined $\sqrt{s}$ ==7 TeV and $\sqrt{s}$ =8 TeV data samples. Superimposed is the result of a fit including a signal component fixed to a hypothesized mass of 126.5 GeV and a background component described by a fourth-order Bernstein polynomial. The bottom inset displays the residual of the data with respect to the fitted background [1].

to increase the sensitivity to the Higgs boson production via vector-boson fusion. Vector-boson fusion events are characterized by two forward jets with little hadronic activity between the two jets. Events in the high-mass 2-jets category are required to have a dijet invariant mass greater than 400 GeV and a η separation greater than 2.8. In addition, the azimuthal angle difference $\Delta\phi$ between the diphoton and the dijet systems is required to be larger than 2.6. The remaining events are classified by whether both photon candidates are unconverted photons ("unconverted") or at least one photon candidate is a converted photon ("converted"), whether both photon candidates are within $|\eta| < 0.75$ ("central") or at least one photon candidate is outside of this region ("rest"). For events with at least one converted photon candidate, a separate "converted transition" category is defined, where at least one photon candidate has $1.3 < |\eta| < 1.75$, which corresponds to the region of transition between the barrel and the endcaps of the calorimeters. Except for the "converted transition" category, all categories are further divided by a p_{Tt} [5] cut at 60 GeV into the "low p_{Tt}" and "high p_{Tt}" categories.

5 Background composition and modeling

The main processes contributing to the background in the $H \rightarrow \gamma\gamma$ search can be divided into two classes: the irreducible background consisting of the QCD diphoton production, and the reducible background consisting of associated production of a photon with jets and processes with several jets in the final state, when one or two jets are misidentified as prompt photons. Several methods based on the variation of the photon identification and isolation criteria are used to determine the composition of the diphoton candidate events. The fraction of diphoton events in the selected sample has been estimated to be (80 ± 4)% in the $\sqrt{s}$ =8 TeV data and $(75^{+3}_{-2})\%$ in the $\sqrt{s}$=7 TeV data. Background from Drell-Yan processes arises

through the mis-reconstruction of electrons as photons, mostly through reconstruction of electrons as converted photons. The number of Drell-Yan events is measured by using $Z \rightarrow e^+e^-$ data events reconstructed as dielectron and e-γ pairs to be $N_{\gamma\gamma}^{DY} = 325 \pm 3(stat) \pm 30(syst)(N_{\gamma\gamma}^{DY} = 270 \pm 4(stat) \pm 24(syst))$ for the $\sqrt{s} =7(8)$ TeV data in the mass region of (100-160) GeV. For the statistical analysis of the measured diphoton spectrum, the background is parametrized by an analytic function for each category, where the normalization and the shape are obtained from fits to the diphoton invariant mass distribution. Different parametrizations are chosen for the different event categories to achieve a good compromise between limiting the size of a potential bias introduced by the chosen parametrization and retaining good statistical power. Depending on the category, an exponential function, a fourth-order Bernstein polynomial or an exponential function of a second-order polynomial is used. For the analysis of the inclusive sample not divided into categories, a fourth-order Bernstein polynomial is used. Potential biases from the choice of background parametrization are estimated using three different sets of high statistics background-only MC models having different event generators for the prompt diphoton background. For a given parametrization, the potential bias is estimated by performing a maximum likelihood fit in the mass range of (100-160) GeV using the sum of a signal and the background parametrization. Parametrizations for which the estimated potential bias is smaller than 20% of the uncertainty on the fitted signal yield, or where the bias is smaller than 10% of the number of expected signal events for a SM Higgs boson for each of the background models are selected for further studies. Among these selected parametrizations, the parametrization with the best expected sensitivity at $m_H = 125$ GeV is selected as the background parametrization. For the chosen parametrization, the largest absolute signal yield obtained for a given MC model over the full mass range studied (from 110 GeV to 150 GeV) is assigned as a systematic uncertainty.

6 Signal modeling

Higgs boson production and decay are simulated using several MC generators, with a subsequent full detector simulation using GEANT4 [6]. Powheg [7], interfaced to Pythia6 [8] for $\sqrt{s} =7$ TeV and Pythia8 [9] for $\sqrt{s} =8$ TeV for showering and hadronization, is used for the generation of gluon fusion and vector-boson fusion production. Pythia6 for $\sqrt{s} =7$ TeV and Pythia8 for $\sqrt{s} =8$ TeV are used to generate Higgs bosons produced in association with W/Z and $t\bar{t}$. In addition, QCD soft-gluon resummation up to next-to-next-to-leading logarithmic order improves the NNLO calculation [10, 11] and is taken into account by means of event reweighting for the simulation of the gluon fusion production mode at $\sqrt{s} =7$ TeV. For the $\sqrt{s} =8$ TeV simulation, a Higgs boson p_T tuning and finite mass effects are taken into account directly in Powheg [12, 13]. In total, 79.3 (111.6) events are expected for a SM Higgs boson in the 7(8) TeV data.

The signal shape is modeled by the sum of a Crystal Ball function (taking into account the core resolution and a non-Gaussian tail towards lower mass values) and a small, wider Gaussian component (taking into account outliers in the distribution). In the inclusive sample, the width of the Crystal Ball component is 1.63 GeV.
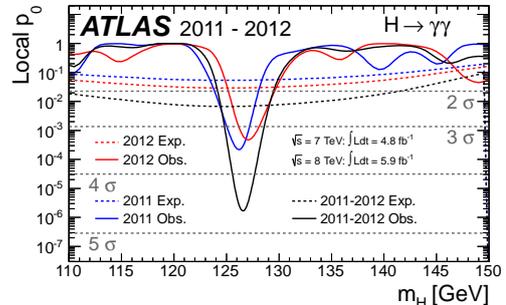

Figure 2. Expected and observed local p_0-value for the analysis using 10 categories, compared to an analysis using only 9 categories (no 2-jets category) and a fully inclusive analysis for the combined $\sqrt{s} ==7$ TeV and $\sqrt{s} =8$ TeV data [1].

Systematic uncertainties on the global signal yield arise from uncertainties on the integrated luminosity (1.8% (3.6%)) for the $\sqrt{s} =7(8)$ TeV data, on the trigger efficiency (1%), on the photon identification efficiency (8.4% (10.8%)), on the isolation cut efficiency (0.4% (0.5%)), from pile-up effects (4%), from the uncertainty on the photon energy scale (0.3%), the predicted Higgs boson production cross section [12, 13] and the Higgs boson decay branching fraction (5%). Systematic uncertainties due to event migration between different categories are due to the modeling of the Higgs boson kinematics (1.1% in the low-p_T categories, 12.5% in the high-p_T categories, and 9% in the 2-jets category), due to pile-up effects (3% (2%) for categories with unconverted photons, 2% (2%) for categories with converted photons, and 2% (12%) for the 2-jets category), due to the material description (4% for categories with unconverted and 3.5% for categories with converted photons), due to the jet energy scale uncertainties (up to 19% for the 2-jets category, and up to 4% for the other categories), due to perturbative uncertainty on the gluon fusion contribution to the 2-jets category (25%), due to the modeling of the underlying event (for the 2-jets category, 30% on gluon fusion and the associated production processes, and 6% uncertainty on the vector-boson fusion process) and due to the jet-vertex-fraction requirement (13% for the $\sqrt{s} =8$ TeV data). Systematic uncertainties on the invariant mass resolution are arising from the uncertainty on the electron energy resolution (12%), from the extrapolation of the electron calibration to photons (6%) and from pile-up effects (4%). The uncertainty on the peak position of the signal invariant mass distribution arises from the uncertainty on the energy scale in the calorimeter presampler, due to material effects when extrapolating the electron energy scale to photons and due to the in-situ calibration method with $Z \rightarrow e^+e^-$ events

(0.6%). All systematic uncertainties, except for the uncertainty on the integrated luminosity, are treated as fully correlated between the $\sqrt{s}$ =7 TeV and the $\sqrt{s}$ =8 TeV analyses.

7 Results

The shape parameters and the normalization of the background are determined by a fit to the data events selected in the mass range (100-160) GeV, performed separately for each category and separately for the $\sqrt{s}$ =7 TeV and $\sqrt{s}$ =8 TeV data. As one example, Fig. 1 shows the invariant mass of diphoton candidates in the inclusive combined $\sqrt{s}$ =7 TeV and $\sqrt{s}$ =8 TeV data samples. The statistical procedures used to test the background-only hypothesis and to set exclusion limits are described in detail in [14]. The minimal p_0-value, quantifying the compatibility of the data with the background-only hypothesis, is found at the m_H = 126(127) GeV for the $\sqrt{s}$ =7(8) TeV data and corresponds to a local significance of $3.4\sigma(3.2\sigma)$. The positions of the two minima are compatible within their uncertainties. The combined p_0-value is shown in at m_H = 126 GeV, 2.5σ are expected for a SM Higgs boson. As a cross-check, the analysis is repeated without dividing the datasets into categories. Results for 95% CL exclusions on the Higgs boson production cross section are obtained in the mass range 110-150 GeV. The expected CL_s limit ranges from 0.8 to 1.6 times the SM expectation in the considered mass range. Under the background-only assumption, the observed(expected) range for the exclusion of a SM Higgs boson is (112-123) GeV and (132-143) GeV ((110-140)GeV). The best fit value for the signal strength μ is obtained from a simultaneous fit to all categories in the $\sqrt{s}$ =7 TeV and $\sqrt{s}$ =8 TeV data samples. At m_H = 126 GeV, where the largest deviation from the background-only hypothesis is found when combining the results presented here with those from other analyses [3], the best fit vaue is $\hat{\mu}$ = 1.8±0.5.

References

[1] The ATLAS Collaboration, Phys. Rev. Lett. **108**, 111803 (2012).
[2] The ATLAS Collaboration, JINST **3** (2008) S08003.
[3] The ATLAS Collaboration, Phys. Lett. **B 716**, 1-29 (2012).
[4] M. Cacciari et. al., JHEP **0804** (2008).
[5] OPAL Collaboration., Eur. Phys. J. **C4** (1998).
[6] GEANT4 Collaboration, Nucl. Instrum. Meth. **A 506** (2003).
[7] S. Alioli et.al, JHEP **0904** (2009).
[8] T. Sjstrand et. al., JHEP **0605** (2006).
[9] T. Sjstrand et. al, Comput.Phys.Commun. **178** (2008).
[10] S. Catani et. al., JHEP **0307** (2003).
[11] D. de Florian and M. Grazzini, Phys. Lett. **B 674** (2009).
[12] LHC Higgs Cross Section Working Group, CERN-2011-002 (2011).
[13] LHC Higgs Cross Section Working Group, CERN-2012-002 (2012).
[14] ATLASandCMSCollaborations,ATLAS-PHYS-PUB-2011-011, CMS NOTE-2011/005 (2011).