Available on the CERN CDS information server **CMS PAS HIG-12-015**

CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2012/07/08

Evidence for a new state decaying into two photons in the search for the standard model Higgs boson in pp collisions

The CMS Collaboration

Abstract

A search for a Higgs boson decaying into two photons is described. The analysis is performed using datasets recorded by the CMS experiment at the LHC from pp collisions at a centre-of-mass energies of 7 TeV and 8 TeV. The datasets correspond to an integrated luminosity of 5.1 and 5.3 fb^{-1}, respectively. Limits are set on the cross section of the standard model Higgs boson decaying to two photons. An excess of events above the expected standard model background is observed for a Higgs boson mass hypothesis of 125 GeV, with a local significance of 4.1σ. 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 3.2σ. This result constitutes evidence for the existence of a new state that decays into two photons.

1 Introduction

The standard model (SM) [1–3] of particle physics has been very successful in explaining experimental data. The origin of the masses of the W and Z bosons that arise from electroweak symmetry breaking remains to be identified. In the SM the Higgs mechanism is postulated, which leads to an additional scalar field whose quantum, the Higgs boson, should be experimentally observable [4–9].

Direct searches by the experiments at LEP ruled out a SM Higgs boson lighter than 114.4 GeV at 95% confidence level (CL) [10]. Limits at 95% CL on the SM Higgs boson mass have also been placed by experiments at the Tevatron, excluding the range 162–166 GeV [11], and by the ATLAS collaboration at the Large Hadron Collider (LHC), excluding the ranges 111.4–116.6, 119.4–122.1, and 129.2–541 GeV [12]. The CMS collaboration at the LHC, using part of the data employed in the present search, has previously excluded at 95% CL the range 127–600 GeV [13]. Fits to the precision electroweak measurements indirectly constrain the mass of the SM Higgs boson to be less than 158 GeV at 95% CL [14].

The $H \to \gamma\gamma$ decay channel provides a clean final-state topology which would allow the mass to be reconstructed with high precision. In the mass range $110 < m_H < 150$ GeV, $H \to \gamma\gamma$ is one of the more promising channels for a Higgs search at the LHC. The primary production mechanism of the Higgs boson at the LHC is gluon fusion [15] with additional small contributions from vector boson fusion (VBF) [16] and production in association with a W or Z boson [17], or a $t\bar{t}$ pair [18–30]. In the mass range $110 < m_H < 150$ GeV the SM $H \to \gamma\gamma$ branching fraction varies between 0.14% and 0.23% [31]. Previous searches for the SM Higgs boson in this channel have been conducted by the D0 experiment [32], and also at the LHC by ATLAS and CMS experiments [33, 34].

This document describes a search for a Higgs boson decaying into two photons in pp collisions at centre-of-mass energies of 7 TeV and 8 TeV, collected in 2011 and 2012. The datasets correspond to an integrated luminosity of 5.1 fb^{-1} at 7 TeV and 5.3 fb^{-1} at 8 TeV. The search results are presented for three analysis procedures: 1) a cut-based analysis where the background model is derived from a fit to the diphoton mass spectrum, 2) an analysis that uses MVA techniques both for photon identification and event classification, and uses the same background model from a fit to the diphoton mass spectrum, and 3) an analysis that uses the same MVA techniques as the previous one but estimates the background from side bands of the diphoton mass distribution. Comparisons of the results from the different analysis procedures provide an important cross-check of the validity of the techniques.

The results presented here build upon the searches reported previously [34, 35], and take advantage of updated detector calibration constants for the 7 TeV dataset and improved analysis methods for the 8 TeV dataset to cope with the increase in the number of simultaneous collisions (pile-up). Throughout this document emphasis has been placed on the description of the analysis of the 8 TeV dataset. For the 7 TeV data set, changes to the analysis described previously [34, 35] are presented where appropriate. For the combined search results, the analyses of the 7 TeV and 8 TeV datasets are taken as separate sub-channels.

2 The CMS detector

A detailed description of the CMS detector can be found elsewhere [36]. The central feature is a superconducting solenoid, 13 m in length and 6 m in diameter, which provides an axial magnetic field of 3.8 T. The bore of the solenoid is instrumented with particle detection

systems. The steel return yoke outside the solenoid is instrumented with gas detectors used to identify muons. Charged particle trajectories are measured by the silicon pixel and strip tracker, with full azimuthal coverage within $|\eta| < 2.5$, where the pseudorapidity η is defined as $\eta = -\ln[\tan(\theta/2)]$, with θ being the polar angle of the trajectory of the particle with respect to the counterclockwise beam direction. A lead-tungstate crystal electromagnetic calorimeter (ECAL) and a brass/scintillator hadron calorimeter (HCAL) surround the tracking volume and cover the region $|\eta| < 3$. The ECAL barrel extends to $|\eta| \approx 1.48$. A lead/silicon-strip preshower detector is located in front of the ECAL endcap. The preshower detector includes two planes of silicon sensors measuring the x and y coordinates of the impinging particles. A steel/quartz-fibre Cherenkov forward calorimeter extends the calorimetric coverage to $|\eta| < 5.0$. In the region $|\eta| < 1.74$, the HCAL cells have widths of 0.087 in both pseudorapidity and azimuth (ϕ). In the (η, ϕ) plane, and for $|\eta| < 1.48$, the HCAL cells map on to 5×5 ECAL crystal arrays to form calorimeter towers projecting radially outwards from points slightly offset from the nominal interaction point. In the endcap, the ECAL arrays matching the HCAL cells contain fewer crystals. Calibration of the ECAL uses $\pi^0 \to \gamma\gamma$, $\mathrm{W} \to e\nu$, and $\mathrm{Z} \to ee$ decays. Deterioration of transparency of the ECAL crystals due to irradiation during the LHC running periods and their subsequent recovery is monitored continuously and corrected for using light injected from a laser and LED system.

3 Data sample and reconstruction

The datasets consist of events collected with diphoton triggers and correspond to integrated luminosities of 5.1 fb^{-1} at 7 TeV and 5.3 fb^{-1} at 8 TeV. Diphoton triggers with asymmetric transverse energy, E_{T}, thresholds and complementary photon selections were used. One selection required a loose calorimetric identification using the shower shape and very loose isolation requirements on photon candidates, and the other required only that the photon candidate had a high value of the R_9 variable. This variable is defined as the energy sum of 3×3 crystals centred on the most energetic crystal in the supercluster (described below) divided by the energy of the supercluster. Its value is used in the analysis to identify photons undergoing a conversion. The E_{T} thresholds used were at least 10% lower than the final selection thresholds. As the instantaneous luminosity delivered by the LHC increased, it became necessary to tighten the isolation cut applied in the trigger. To maintain high trigger efficiency, all four possible combinations of threshold and selection criterion were deployed (i.e., with both photon candidates having the R_9 condition, with the high threshold candidate having the R_9 condition applied and the low threshold candidate having the loose ID and isolation, and so on). Accepting events that satisfy any of these triggers results in a $> 99\%$ trigger efficiency for events passing the offline selection.

Photon candidates are reconstructed from clusters of ECAL channels around significant energy deposits, which are merged into superclusters. The clustering algorithms result in almost complete recovery of the energy of photons that convert in the material in front of the ECAL. In the barrel region, superclusters are formed from five-crystal-wide strips in η centred on the locally most energetic crystal (seed) and have a variable extension in ϕ. In the endcaps, where the crystals are arranged according to an x-y rather than an η-ϕ geometry, matrices of 5×5 crystals (which may partially overlap) around the most energetic crystals are merged if they lie within a narrow road in η.

The photon energy is computed starting from the raw supercluster energy. In the region covered by the preshower detector ($|\eta| > 1.65$) the energy recorded in that detector is added. In order to obtain the best resolution, the raw energy is corrected for the containment of the

shower in the clustered crystals, and the shower losses for photons which convert in material upstream of the calorimeter. These corrections are computed using a multivariate regression technique based on the boosted decision tree (BDT) implementation in TMVA [37]. The regression is trained on photons in a sample of simulated events using the ratio of the true photon energy to the raw energy as the target variable. The input variables are the global η and ϕ coordinates of the supercluster, a collection of shower-shape variables, and a set of local cluster coordinates. A second BDT, using the same input variables, is trained on fully simulated photons to provide an estimate of the Gaussian uncertainty on the energy value provided by the first BDT.

Jets, whose use will be described below, are reconstructed using a particle-flow algorithm [38, 39], which uses the information from all CMS sub-detectors to reconstruct different types of particles produced in the event. The basic objects of the particle-flow reconstruction are the tracks of charged particles reconstructed in the central tracker, and energy deposits reconstructed in the calorimetry. These objects are clustered with the anti-k_T algorithm [40] using a value of 0.5 for the "distance parameter" ΔR. The jet energy measurement is calibrated to correct for detector effects using samples of dijet, $\gamma +$ jet, and $\mathrm{Z} +$ jet events [41]. Energy from pile-up interactions overlapping the one that produced the diphoton, and from the underlying event, is also included in the reconstructed jets. This energy is subtracted using the FASTJET technique [42–44], which is based on the calculation of the η-dependent transverse momentum density, evaluated on an event-by-event basis. Finally, particles produced in pile-up interactions may be clustered into (fake) jet objects of relatively large transverse momentum. These fake jets are removed using selection criteria based on the compatibility of the tracks in a jet with the primary vertex and on the width of the jet.

Samples of Monte Carlo (MC) events used in the analysis are fully simulated using GEANT [45]. The simulated events include the effects of pile-up, and the events are reweighted to reproduce the expected distribution of the number of interactions taking place in each bunch crossing.

4 Vertex location

The mean number of pp interactions per bunch crossing is 9.5 (18.7) in the 7 TeV (8 TeV) dataset. The interaction vertices reconstructed using the tracks of charged particles are distributed in the longitudinal direction, z, with an RMS spread of 6 cm ($\sim$5 cm) in the 7 TeV (8 TeV) dataset. If the interaction point is known to better than about 10 mm, then the resolution on the opening angle between the photons makes a negligible contribution to the mass resolution, compared to the contribution from the ECAL energy resolution. Thus the mass resolution can be preserved by correctly assigning the reconstructed photons to one of the interaction vertices reconstructed from the tracks. The techniques used to achieve this are described below.

The reconstructed primary vertex having the highest probability of being the interaction vertex of the diphoton event can be identified using the kinematic properties of the tracks associated with the vertex and their correlation with the diphoton kinematic properties. In addition, if either of the photons converts and the tracks from the conversion are reconstructed and identified, the direction of the converted photon, determined by combining the conversion vertex position and the position of the ECAL supercluster, can be used to point to and so identify the diphoton interaction vertex.

For the determination of the primary vertex position using kinematic properties, three discriminating variables are constructed from the measured scalar, p_{T}, or vector, $\vec{p}_T$, transverse momenta of the tracks associated with each vertex, and the transverse momentum of the diphoton

system, $p_{\mathrm{T}}^{\gamma\gamma}$. These three variables are: $\sum p_{\mathrm{T}}^2$, and two variables which quantify the p_{T} balance with respect to the diphoton system: -$\sum(\vec{p}_T \cdot \frac{\vec{p}_T^{\gamma\gamma}}{|\vec{p}_T^{\gamma\gamma}|})$ and $(|\sum \vec{p}_T| - p_{\mathrm{T}}^{\gamma\gamma})/(|\sum \vec{p}_T| + p_{\mathrm{T}}^{\gamma\gamma})$. An estimate of the "pull" to each vertex from the longitudinal location on the beam axis pointed to by any reconstructed tracks (from a photon conversion) associated with the two photon candidates is also computed: $pull_{conv} = |z_{conversion} - z_{vertex}|/\sigma_{conversion}$. These variables are used in a multivariate system based on boosted decision trees (BDT) to choose the reconstructed vertex to associate with the photons.

The vertex-finding efficiency, defined as the efficiency to locate the vertex to within 10 mm of its true position, has been studied with $\mathrm{Z} \to \mu\mu$ events where the algorithm is run after the removal of the muon tracks. The use of tracks from a converted photon to locate the vertex is studied with $\gamma +$ jet events. In both cases the ratio of the efficiency measured in data to that in MC simulation is close to unity. The value is measured as a function of the Z boson p_{T}, as measured by the reconstructed muons, and is used as a correction to the Higgs boson signal model. An uncertainty of 0.4% is ascribed to the knowledge of the vertex finding efficiency coming from the statistical uncertainty in the efficiency measurement from $\mathrm{Z} \to \mu\mu$ (0.2%) and the uncertainty related to the Higgs boson p_{T} spectrum description, which is estimated to be 0.3%. The overall vertex-finding efficiency for a Higgs boson of mass 120 GeV, integrated over its p_{T} spectrum, is computed to be 83.0 ± 0.2(stat)$\pm$0.4(syst)% in the 7 TeV dataset, and $\sim$79.0$\pm$0.2(stat)% in the 8 TeV dataset.

A second vertex-related multivariate discriminant was designed to estimate, event-by-event, the probability for the vertex assignment to be within 10 mm of the diphoton interaction point. This, used in conjunction with the event-by-event estimate of the photons' energy resolution, allows an estimate of the diphoton mass resolution. A BDT was trained, using simulated $\mathrm{H} \to \gamma\gamma$ events, to separate events where the chosen vertex lies within 10 mm of the generated interaction point. The inputs of the BDT were:

- the values of the vertex BDT output for the three most likely vertices in each event,
- the number of vertices in each event,
- $p_{\mathrm{T}}^{\gamma\gamma}$,
- the distances between the chosen vertex and the second and third choices,
- the number of photons with an associated conversion track.

The performance of the BDT has been verified by comparison of the true vertex identification probability in simulated $\mathrm{H} \to \gamma\gamma$ events with the average value of the estimated probability as a function of the reconstructed diphoton p_{T}. The performance on data has been compared to that in MC simulation by using $\mathrm{Z} \to \mu\mu$ and $\gamma +$ jet events, and the distribution of the output in simulated events is found to match what is observed in data.

5 Photon selection

The event selection requires two photon candidates with $p_{\mathrm{T}}^{\gamma}(1) > m_{\gamma\gamma}/3$ and $p_{\mathrm{T}}^{\gamma}(2) > m_{\gamma\gamma}/4$ within the ECAL fiducial region: $|\eta| < 2.5$ and excluding the barrel-endcap transition region $1.44 < |\eta| < 1.57$. The fiducial region requirement is applied to the supercluster position in the ECAL, and the p_{T} threshold is applied after the vertex assignment. The excluded barrel-endcap transition region removes from the acceptance the last two rings of crystals in the barrel, to ensure complete containment of accepted showers, and the first ring of trigger towers in the endcap which is obscured by cables and services exiting between the barrel and endcap. In the rare

case where the event contains more than two photons passing all the selection requirements, the pair with the highest summed (scalar) p_{T} is chosen.

The dominant backgrounds to $\mathrm{H} \to \gamma\gamma$ consist of 1) the irreducible background from the prompt diphoton production, and 2) the reducible backgrounds from $\mathrm{pp} \to \gamma + \text{jet}$ and $\mathrm{pp} \to \text{jet} + \text{jet}$ where one or more of the objects reconstructed as a photon corresponds to a particle in a jet that deposit substantial energy in the ECAL, typically a photon from the decay of neutral light mesons. These reconstructed objects are generally referred to as *fake* photons.

Photon identification using a BDT is applied after a preselection, to distinguish prompt photons from non-prompt photon background. The preselection requirements consist of an electron veto, isolation criteria using the ratio of hadronic energy in HCAL towers behind the supercluster to the ECAL energy in the supercluster, and criteria based on the shower width. The latter two requirements are chosen such that the preselected photons have a very high efficiency to pass the trigger selection and the MC generator level enrichment filter.

The photon identification BDT is trained using simulated $\mathrm{pp} \to \gamma + \text{jet}$ event samples. Prompt photons are used as the signal and non-prompt photons used as the background for the BDT training.

The following variables are used as input to the photon identification BDT:

1. Shower topology variables, where the shower shape variables of the MC simulation are scaled to match the MC simulation to data.
   (a) $\sigma_{i\eta i\eta}$, the energy weighted (single crystal energy over the supercluster energy) standard deviation of single crystal η within the 5×5 crystals centred at the crystal with maximum energy.
   (b) $\mathrm{cov}(i\eta, i\phi)$, the off-diagonal element of the energy weighted covariance matrix of single crystal η and ϕ within the 5×5 crystals centred at the crystal with maximum energy.
   (c) $\mathrm{E}_{2\times2}/\mathrm{E}_{5\times5}$, the ratio of the energy in the 2×2 groups of crystals which contain the crystal with maximum energy and which have the maximum energy sum, to the energy in the 5×5 crystals centered on the crystal with maximum energy.
   (d) R_9, as previously defined.
   (e) σ_η, the energy weighted (crystal energy over supercluster energy) standard deviation of single crystal η within the supercluster.
   (f) σ_ϕ, the energy weighted (crystal energy over supercluster energy) standard deviation of single crystal ϕ within the supercluster.
   (g) σ_{RR}, the square root of the sum of the squares of the standard deviations of the shower spread in the x and y planes of the preshower detector.
2. Isolation variables, based on the particle flow algorithm.
   (a) Particle flow photon isolation sum within a $\Delta R < 0.3$ cone.
   (b) Particle flow charged hadron isolation sum within a $\Delta R < 0.3$ cone, calculated with respect to the selected vertex.
   (c) Particle flow charged hadron isolation sum within a $\Delta R < 0.3$ cone, with respect to the vertex for which this isolation sum is greatest.

3. ρ, the event energy density per unit area. This variable is introduced to adjust the pile-up dependence in the isolation variables.

4. Supercluster pseudorapidity, the η of the supercluster corresponds to the reconstructed photon. This variable is introduced to adjust the η dependence of the shower topology variables and isolation variables.

The efficiency of the photon preselection is measured in data using "tag and probe" techniques [46] using two event samples. The efficiency of the complete selection excluding the electron veto requirement is determined using $Z \to ee$ events. Table 1 shows the results for data and MC simulation, and the ratio of efficiency in data to that in the simulation, $\epsilon_{data}/\epsilon_{MC}$. The efficiency for photons to pass the electron veto has been measured using $Z \to \mu\mu\gamma$ events, where the photon is produced by final-state radiation, which provide a rather pure ($> 99\%$) source of prompt photons. The efficiency ranges from 97 to 100% and the ratio $\epsilon_{data}/\epsilon_{MC}$ for the electron veto is close to unity in all categories. The quadratic sum of the statistical and systematic uncertainties for the measurements of efficiencies using data are propagated to the uncertainties on the ratios. The ratios are used as corrections to the signal efficiency simulated in the MC model of the signal. The uncertainties on the ratios are taken as systematic uncertainties in setting limits.

Table 1: Photon preselection efficiencies for the 8 TeV dataset measured in four photon categories, obtained with a tag and probe technique using $Z \to ee$ events after applying all requirements except the electron veto.

Category	ϵ_{data} (%)	ϵ_{MC} (%)	$\epsilon_{\mathrm{data}}/\epsilon_{\mathrm{MC}}$
Barrel, $R_9 > 0.9$	98.94±0.30	99.16±0.01	0.998±0.003
Barrel, $R_9 < 0.9$	93.27±0.56	93.67±0.03	1.003±0.006
Endcap, $R_9 > 0.9$	98.32±0.90	97.71±0.02	1.005±0.009
Endcap, $R_9 < 0.9$	92.98±1.71	92.98±0.03	1.003±0.018

The cut-based photon identification algorithm used prevously [34] was also updated to cope with the increased number of pile-up collisions. Using isolation sums based on the particle flow algorithm has the effect of reducing the variation of the selection efficiency with pile-up to less than 10% over the range from 10 to 40 reconstructed primary vertices.

The efficiency of the trigger has also been measured using $Z \to ee$ events, with the events classified as described below. For events passing the analysis preselection the trigger efficiency is found to be above 99.5% for all selected events.

6 The diphoton BDT

The analysis uses events with two photons satisfying the preselection requirements, including a requirement on the photon identification BDT output which retains more than 99% of the signal events passing the other preselection requirements, while removing 27% of the data events in the region of $100 < m_{\gamma\gamma} < 180$ GeV.

An event classifier variable is constructed that fulfills the following criteria:

1. The variable should classify with a high score events with:

(a) signal-like kinematic characteristics ,

(b) good diphoton mass resolution events,

(c) photon-like values from the photon identification BDT,

2. The variable should be mass independent; it should not select events according to the invariant mass.

The multivariate discriminator 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). This choice of inputs is justified by the fact that the signal-to-background ratio, and the relative magnitude of the contribution of background "photons" from jets, varies as a function of the photon kinematic properties. In addition, the diphoton mass resolution depends on the location of the associated energy deposits in the calorimeter, whether or not one or both photons converted in the detector volume in front of the calorimeter, and the probability that the correct primary vertex has been used to reconstruct the diphoton mass.

The following variables are used:

- the relative transverse momenta of both photons, $p_{\mathrm{T}}^{1(2)}/m_{\gamma\gamma}$,
- the pseudo-rapidities of both photons, $\eta^{1(2)}$,
- the cosine of the angle between the two photons in the transverse plane, $\cos(\phi_1 - \phi_2)$,
- the relative diphoton mass resolution, $\sigma_m^{right}/m_{\gamma\gamma}$ (right vertex),
- the photon identification BDT output value for both photons.

The relative mass resolution estimate is computed from the photon energy resolution estimate assuming Gaussian resolution functions as:

$$\sigma_m^{right}/m_{\gamma\gamma} = \frac{1}{2}\sqrt{(\sigma_{E1}/E_1)^2 + (\sigma_{E2}/E_2)^2} \tag{1}$$

This computation assumes the correct primary vertex has been selected, in which case the energy measurement of the photons is the dominant contribution to the mass resolution. Since the correct primary vertex is not always selected, we compute and input to the BDT additionally the relative mass resolution computed under the assumption that the incorrect primary vertex was chosen. In this case the distance between the true vertex and the selected vertex is distributed as a Gaussian with width $\sqrt{2}\,\sigma_Z^{beamspot}$, and the contribution to the resolution σ_m^{vtx} can be computed analytically given the impact positions of the two photons in the calorimeter. The relative resolution under the incorrect vertex hypothesis is then computed as:

$$\sigma_m^{wrong}/m_{\gamma\gamma} = \sqrt{\left(\sigma_m^{right}/m_{\gamma\gamma}\right)^2 + (\sigma_m^{vtx}/m_{\gamma\gamma})^2} \tag{2}$$

In the training, information needs to be provided that signal-to-background is inversely proportional to mass resolution. This is achieved by weighting the signal events used to train the BDT:

$$w_{sig} = \frac{p_{vtx}}{\sigma_m^{right}/m_{\gamma\gamma}} + \frac{1 - p_{vtx}}{\sigma_m^{wrong}/m_{\gamma\gamma}} \tag{3}$$

The BDT will thus tend to assign a high score classifier value to events with better resolution.

The BDT is trained using simulated MC background and Higgs boson signal events ($m_{\mathrm{H}} = 123\,\mathrm{GeV}$).

In order to verify that there is good agreement between data and MC simulation for the diphoton BDT output value, studies have been performed using both $\mathrm{Z} \to ee$ and the leading photons in high mass, $m_{\gamma\gamma} > 160\,\mathrm{GeV}$, diphoton events, whose purity is high. These studies examine both the inputs to the diphoton BDT, and its output value. They show that a systematic uncertainty of ± 0.01 on the photon identification BDT output value, together with an uncertainty on the per-photon energy resolution estimate, parameterized as a rescaling of the resolution estimate by $\pm 10\%$ about its nominal value, fully covers the differences observed between data and MC simulation.

7 Event classes

The diphoton BDT output value provides an indication of signal-to-background, including the effect of mass resolution. This output value has been used in two ways, the first of which, described in this section, is to divide events into classes prior to fitting the diphoton invariant mass spectrum.

The method described below is used to choose the values of the diphoton BDT to be used as class boundaries. The procedure is to successively split the events into classes by introducing a boundary value. The first boundary results in two classes, and then these classes are further split. Each split is introduced using the boundary value that gives rise to the best expected exclusion limit. The optimized quantity is the median expected limit, computed using the asymptotic approximation [47] to the modified frequentist $\mathrm{CL_S}$ procedure against the MC simulated background prediction.

To avoid possible overtraining effects, since the BDT was trained against the $m_{\mathrm{H}} = 123\,\mathrm{GeV}$ hypothesis, the nearby mass value of 124 GeV is used to optimize the classification.

Negligible ($< 1\%$) gain in sensitivity, as measured by the expected limit on the Higgs boson cross section at 95% CL, is found for splitting beyond five classes, and the lowest score class can be dropped as it does not contribute significantly to the sensitivity. Dropping the lowest score class is equivalent to applying a selection cut, $\mathrm{BDT}^{\gamma\gamma}_{\mathrm{low}}$, on the diphoton BDT output value, and removes 76% of the diphoton events, $100 < m_{\gamma\gamma} < 180\,\mathrm{GeV}$, passing the preselection. The fraction of Higgs boson events ($m_{\mathrm{H}} = 120\,\mathrm{GeV}$) passing the preselection which are removed is 22%.

Investigating the properties of the simulated signal events in these classes reveals that the best class (class 0) contains, almost exclusively, those events where $p_{\mathrm{T}}^{\gamma\gamma} > 40\,\mathrm{GeV}$, while the second best class (class 1) is dominated by events where both photons are unconverted and are situated in the central barrel region of the ECAL.

Events passing a dijet tag defined to select Higgs bosons produced by the VBF process are analysed separately. Higgs boson events produced by VBF have two forward jets, originating from the two scattered quarks. Higgs bosons produced by this mechanism have a harder transverse momentum spectrum than those produced by the gluon-gluon fusion process or the photon pairs produced by the background processes [48]. By using a dijet tag it is possible to define small classes of events which have an expected signal-to-background ratio more than an order of magnitude greater than events in the four classes defined by the diphoton BDT. The additional classification of events into dijet-tagged classes improves the sensitivity of the analysis by about 10%, as measured by the expected limit, at 95% CL, on the Higgs boson cross section.

While in the 7 TeV dataset one single class of dijet-tagged events is used, in the 8 TeV dataset analysis two classes are defined using the dijet invariant mass.

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 $p_{\mathrm{T}}^{\gamma}(1) > m_{\gamma\gamma}/2$, to reduce background with negligible loss of acceptance.

The selection variables for the jets use the two highest transverse energy (E_{T}) jets in the event with pseudorapidity $|\eta| < 4.7$. The pseudorapidity restriction avoids the use of jets for which the energy corrections are less reliable and is found to have only a small effect (<2% change) on the signal efficiency. The selection requirements have been optimized using simulated events of VBF signal and diphoton background, to minimize the expected limit at 95% CL on the VBF signal cross section, using the only the dijet-tagged event classes.

The pseudorapidity separation between the two jets is required to be greater than 3.0. The E_{T} thresholds for the two jets are 30 GeV and dijet invariant mass is greater than 500 GeV. An event is also selected if the leading jet E_{T} is above 30 GeV, the second leading jet E_{T} is above 20 GeV, and the dijet invariant mass is greater than 250 GeV. By requiring these criteria in sequence, two independent classes of events are defined. Two additional selection criteria, relating the dijet to the diphoton system, have been applied: 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 [49], and the difference in azimuthal angle between the diphoton system and the dijet system is required to be greater than 2.6 ($\approx 150^{\circ}$).

Table 2 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 σ_{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.

Significant systematic uncertainties on the efficiency of dijet tagging of signal events arise from the uncertainty on the MC modelling of jet-energy corrections and jet-energy resolution, and from uncertainties in predicting the presence of the jets and their kinematic properties. These uncertainties arise from the effect of different underlying event tunes, and from the uncertainty on parton distribution functions and QCD scale factor. Overall, an uncertainty of 10% is assigned to the efficiency for VBF signal events to enter the dijet-tag class, and an uncertainty of 50%, which is dominated by the uncertainty on the underlying event tune is assigned to the efficiency for signal events produced by gluon-gluon fusion to enter the dijet-tag classes. The uncertainty on the underlying event tunes was investigated by comparing the DT6 [50], P0 [51], ProPT0 and ProQ20 [52] tunes to the Z2 tune [53] in PYTHIA [54].

For the cut-based analysis, the classification scheme previously used [34] and based on the pseudo-rapidity and R_9 of the photons is supplemented with the two dijet-tagged event classes defined above.

8 Signal modelling

The description of the Higgs boson signal used in the search is obtained from MC simulation using the next-to-leading order (NLO) matrix-element generator POWHEG [55, 56] interfaced with PYTHIA [54]. For the dominant gluon-gluon fusion process, the Higgs boson transverse

Table 2: 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.

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	$\pm$ 0.4
	Untagged 1	16.3	88%	6%	6%	1%	1.26	1.08	37.5	$\pm$ 1.3
	Untagged 2	21.5	91%	4%	4%	–	1.59	1.32	74.8	$\pm$ 1.9
	Untagged 3	32.8	91%	4%	4%	–	2.47	2.07	193.6	$\pm$ 3.0
	Dijet tag	2.9	27%	73%	1%	–	1.73	1.37	1.7	$\pm$ 0.2
8 TeV 5.3 fb^{-1}	Untagged 0	6.1	68%	12%	16%	4%	1.38	1.23	7.4	$\pm$ 0.6
	Untagged 1	21.0	88%	6%	6%	1%	1.53	1.31	54.7	$\pm$ 1.5
	Untagged 2	30.2	92%	4%	3%	–	1.94	1.55	115.2	$\pm$ 2.3
	Untagged 3	40.0	92%	4%	4%	–	2.86	2.35	256.5	$\pm$ 3.4
	Dijet tight	2.6	23%	77%	–	–	2.06	1.57	1.3	$\pm$ 0.2
	Dijet loose	3.0	53%	45%	2%	–	1.95	1.48	3.7	$\pm$ 0.4

momentum spectrum has been reweighted to the next-to-next-to-leading logarithmic (NNLL) + NLO distribution computed by the HQT program [57–59]. 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 [60]. 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. The uncertainty on the signal cross section due to PDF uncertainties has been determined using the PDF4LHC prescription [61–65]. The SM Higgs boson cross sections and branching fractions used are taken from ref. [66].

The uncertainty on the relative event class yields due to scale uncertainty has been estimated by varying the renormalization and factorization scales used by MC@NLO [67] and MCFM [68], between $m_H/2$ and $2 \times m_H$, and examining the effect on the Higgs boson kinematic properties in terms of rapidity and transverse momentum. By deriving weights defined as the ratio of the differential cross section in the varied case to that in the nominal, the effect on the event class yields can be determined. This effect is a migration of signal events between classes, which can be as large as 12.5%. The effect of PDF uncertainties on the relative event class yields has been determined in a similar fashion, by varying the Higgs boson kinematic properties according to the variations of the 26 eigenvalues of the CT10 [69] PDF set within their uncertainties. The largest migration of signal events between classes due to the PDF uncertainties is 1.3%.

Two other causes of migration of events between classes have also been assessed. These are the uncertainties on the diphoton BDT output value due to uncertainty on the photon identification BDT output value, and due to uncertainty on the per-photon energy resolution estimate. The resulting uncertainties are propagated to the final statistical analysis.

Table 3 lists the sources of systematic uncertainty on the signal considered in the analysis, together with the magnitude of the variation of the source that has been applied.

Corrections are made to the measured energy of the photons based on detailed study of the

Table 3: Separate sources of systematic uncertainties accounted for in the analysis of the 8 TeV data set. The magnitude of the variation of the source that has been applied to the signal model is shown.

Sources of systematic uncertainty		**Uncertainty**	
Per photon		Barrel	Endcap
Photon selection efficiency		0.8%	2.2%
Energy resolution ($\Delta\sigma/E_{MC}$)	$R_9 > 0.94$ (low η, high η)	0.22%, 0.60%	0.90%, 0.34%
	$R_9 < 0.94$ (low η, high η)	0.24%, 0.59%	0.30%, 0.52%
Energy scale ($(E_{data} - E_{MC})/E_{MC}$)	$R_9 > 0.94$ (low η, high η)	0.19%, 0.71%	0.88%, 0.19%
	$R_9 < 0.94$ (low η, high η)	0.13%, 0.51%	0.18%, 0.28%
Photon identification BDT (Effect of up to 4.3% event class migration.)		$\pm$0.01 (shape shift)	
Photon energy resolution BDT (Effect of up to 8.1% event class migration.)		$\pm$10% (shape scaling)	
Per event			
Integrated luminosity		4.4%	
Vertex finding efficiency		0.2%	
Trigger efficiency	One or both photons $R_9 < 0.94$ in endcap	0.4%	
	Other events	0.1%	
Dijet selection			
Dijet-tagging efficiency	VBF process	10%	
	Gluon-gluon fusion process	50%	
(Effect of up to 15% event migration among dijet classes.)			
Production cross sections		Scale	PDF
Gluon-gluon fusion		+12.5% -8.2%	+7.9% -7.7%
Vector boson fusion		+0.5% -0.3%	+2.7% -2.1%
Associated production with W/Z		1.8%	4.2%
Associated production with $t\bar{t}$		+3.6% -9.5%	8.5%
Scale and PDF uncertainties (Effect of up to 12.5% event class migration.)		(y, p_T)-differential	

mass distribution of $Z \to ee$ events and comparison with MC simulation. After the application of these corrections the $Z \to ee$ events are re-examined and values are derived for the additional smearing that needs to be made to the MC simulation to account for the energy resolution observed in the data. These smearings are derived for photons separated into four η regions (two in the barrel and two in the endcap) and two categories of R_9. The uncertainties on the measurements of the photon scale and resolution are taken as systematic uncertainties in the limit setting.

9 Mass-fit background model

In the analysis the background is exclusively evaluated from the data, without reference to the MC simulation, which is used for BDT training. However, the diphoton mass spectrum that is observed after the full event selection is found to agree with the distribution predicted by MC simulation, within the uncertainties, on the cross sections of the contributing processes which are estimated to be about 15%. The background components have been scaled by K-factors obtained from CMS measurements [70–72]. The contribution to the background in the diphoton mass range $110 < m_{\gamma\gamma} < 150\,\text{GeV}$ from processes giving non-prompt photons is about 28%.

The mass-fit background model is obtained by fitting the observed diphoton mass distributions in each of the five event classes over the range $100 < m_{\gamma\gamma} < 180\,\text{GeV}$. The choice of function used to fit the background and the choice of the fit range are based on a study of the possible bias introduced by the choice on both the limit, in the case of no signal, and the measured signal strength, in the case of a signal. Provided that the potential bias is substantially smaller than the statistical accuracy of the fit prediction it can be safely neglected. An acceptable maximum bias on the fitted signal strength has been taken as five times smaller than the statistical accuracy.

The study of bias is performed as follows. Five functional forms are fitted to the simulated background distribution of the diphoton invariant mass spectrum and taken as "truth models" for the background. The functional forms are: exponential ($\sum a_i e^{b_i x}$), power law ($\sum a_i x^{-b_i}$), polynomial ($\sum a_i x^i$), Laurent series and the form $x^{-a} \times e^{(N-1)P}$ with $P = \sum a_i x^i$. The functional forms are intended to span the possibilities of the true distribution (i.e. the distribution to which signal-free data would tend if the number of events were indefinitely increased). The order of the functional forms, defining the number of degrees of freedom in the fit, is chosen after a study of fits to the data and using an F-test. This is done for each of the event classes in turn. The order of the fit, for each functional form, is raised until raising it further does not significantly improve the goodness of fit. The order that is required to achieve this is then used for the truth model.

Pseudo-data are then randomly generated from the truth model fits to the data, and these pseudo-data are fitted with candidate fit functions. Only polynomial shapes seem able to describe both exponential and power law truth models, and especially only polynomial models are able to describe polynomial truth models without significant bias. The set of fit models under consideration is thus restricted to polynomials. The required order of polynomial ranges from 3 to 5.

The $m_{\gamma\gamma}$ distributions for the data in the eleven event classes, together with the results of the simultaneous fit of the signal plus background model to the eleven classes, are shown in Figs. 1 and 2 for the 7 TeV and the 8 TeV data samples respectively. The uncertainty bands shown in the background component of the fit are computed from the fit uncertainty on the background yield within bins corresponding to those used for the data points. The fit is performed with

a floating signal strength (relative to the SM expectation) that has a best fit value of $\sigma/\sigma_{\mathrm{SM}} = 1.56 \pm 0.43$.

It should be noted that for the cut-based analysis, the background model procedure described above is also utilised, albeit applied to different event classes. In the cut-based analysis the order of the polynomials also ranges from 3 to 5.

10 Background model from mass side bands

An alternative approach to constructing the background model has also been studied. This approach uses data in the sidebands of the invariant mass distribution. This method causes the result to be less sensitive to the parametric form used to describe the diphoton mass spectrum, and allows the explicit inclusion of a systematic uncertainty for the possible bias in the background mass fit. The result is extracted from a fit to the output distribution of a BDT which takes two dimensionless inputs: the diphoton BDT output, and the mass, in the form $\Delta m/m_{\mathrm{H}}$, where $\Delta m = m_{\gamma\gamma} - m_{\mathrm{H}}$, and m_{H} is the signal mass hypothesis. The output value of this mass window BDT is binned. The only assumption made concerning the shape of the background is that the fraction of events in each bin varies linearly as a function of invariant mass. A parametric fit to the mass spectrum is used only to obtain the normalization, and not the shape, of the background model.

The mass window BDT is trained using simulated Higgs boson events ($m_{\mathrm{H}} = 123\,\mathrm{GeV}$) and simulated background including prompt-prompt, prompt-fake and fake-fake processes. The training samples are not used in any other part of the analysis, except as input to the binning algorithm, whose role is described below, thus avoiding any biases that may emerge due to over-training on fluctuations in the samples.

The event sample is identical to that for the analysis using the mass-fit background model. The four event classes defined by the diphoton BDT are not used. Instead, an optimized binning procedure is used to rebin the mass window BDT output distributions such that a small number of bins can be used without degrading the sensitivity. The mass window BDT is not used for events passing the dijet selection. Instead, the dijet-tagged events constitute an additional bin appended to the bins of the mass window BDT output value.

The data in the sidebands of the invariant mass distribution are used to model the background. The sidebands have the same kinematic distributions as the background within the signal region, but are free of signal for a given signal hypothesis. The signal region is defined to be the range $\pm 2\%$ on either side of the mass hypothesis m_{H}, which for m_{H}=125 GeV contains 93% of the signal.

In order for the kinematic variables used by the BDT to cover the same phase space in the sidebands as in the signal region, each sideband is defined to have the same width of $\pm 2\%$ relative to the mass hypothesis corresponding to the centre of the sideband. Multiple sidebands are used in order to increase the statistics.

Further details on the method can be found in Ref. [34].

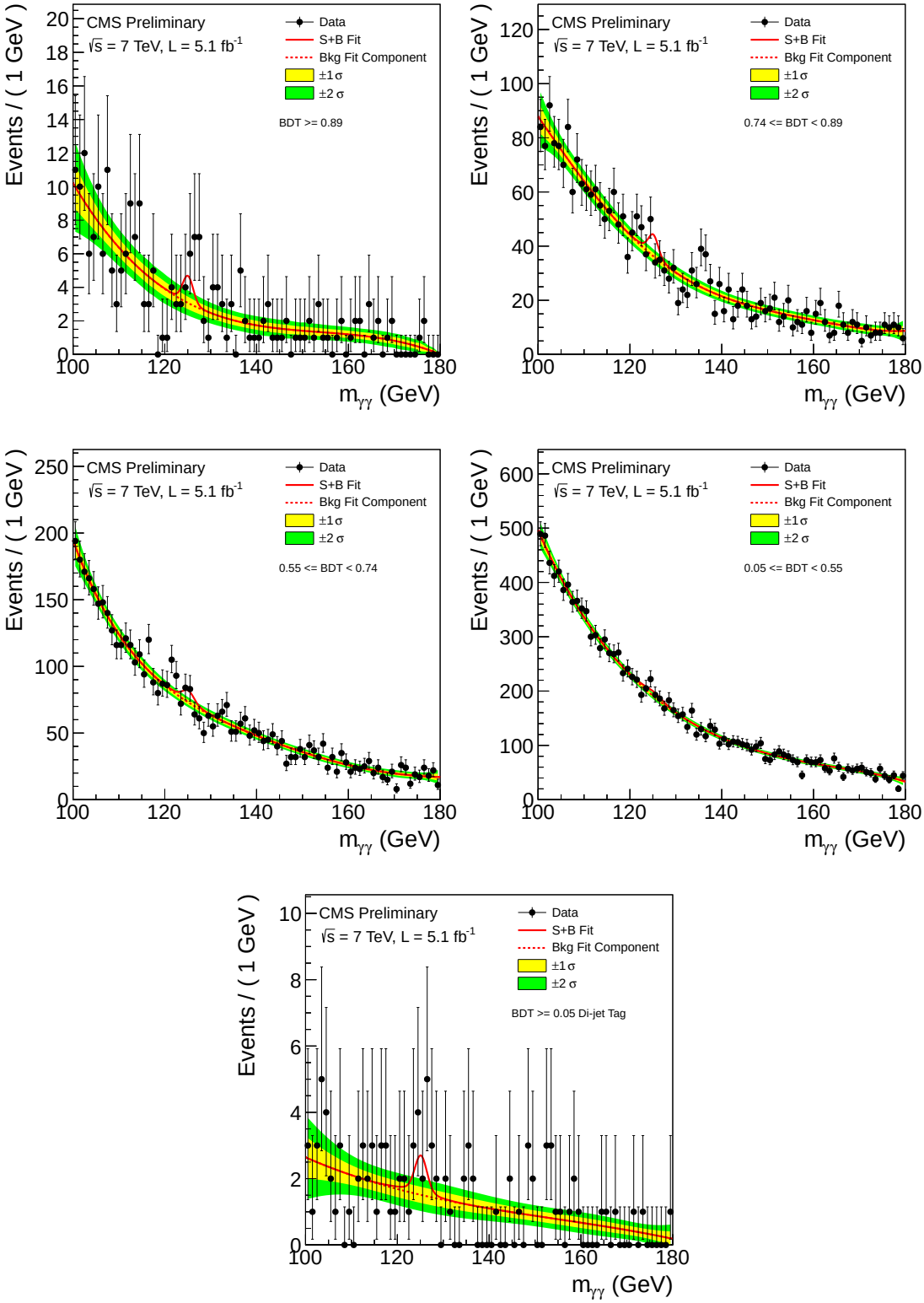


Figure 1: A comparison of data and the result of the signal-plus-background model fit to the $m_{\gamma\gamma}$ distribution, for the five event classes of the 7 TeV dataset.

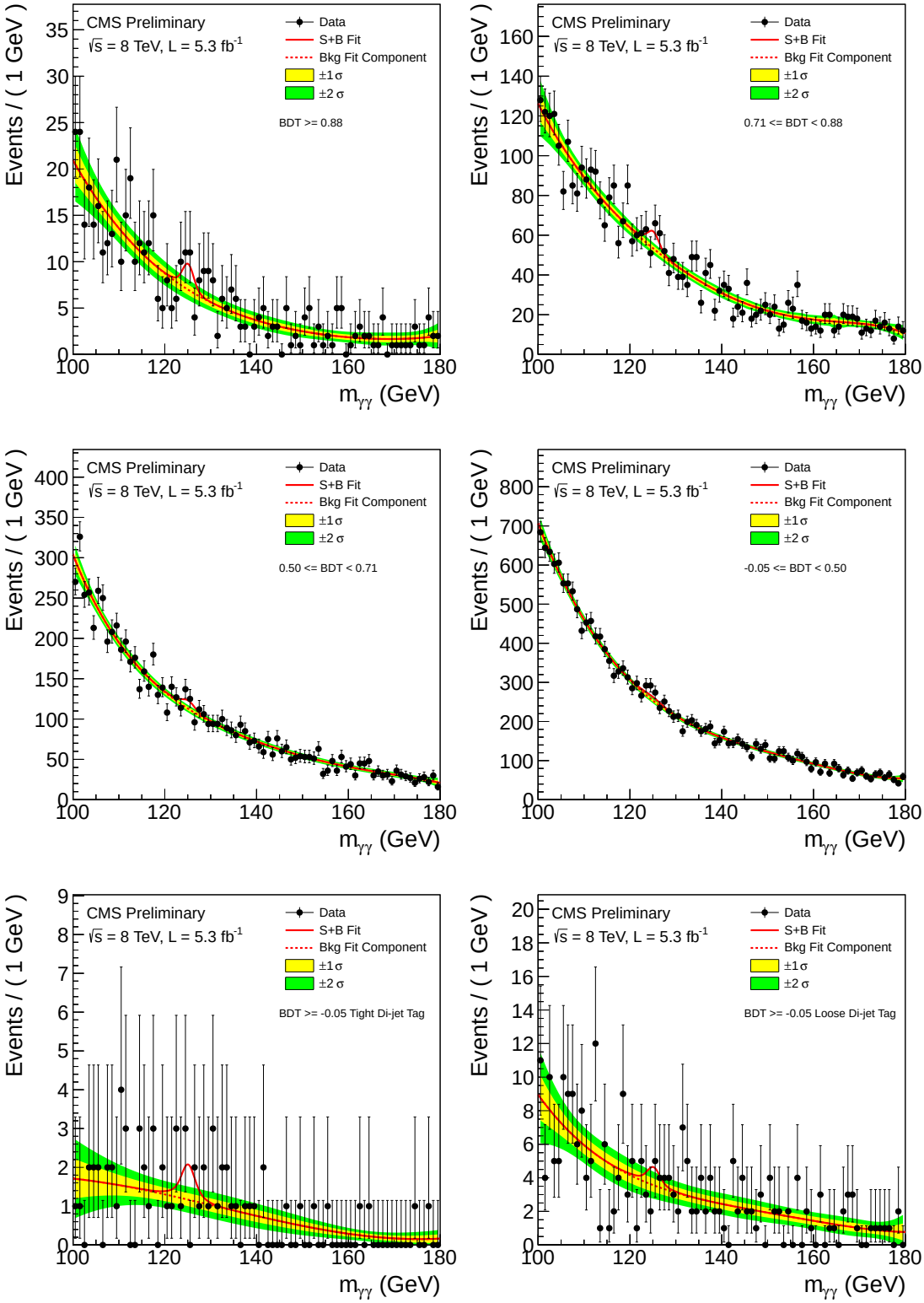


Figure 2: A comparison of data and the result of the signal-plus-background model fit to the $m_{\gamma\gamma}$ distribution, for the six event classes of the 8 TeV dataset.

11 Results

The confidence level (CL) for exclusion or discovery of a SM Higgs boson signal is evaluated from a simultaneous fit to the diphoton invariant mass distributions for the eleven event classes.

The limits are evaluated using a modified frequentist approach, CL_S, taking the profile likelihood as a test statistic [73–75]. Both a binned and an unbinned evaluation of the likelihood are considered. While most of the analysis and determination of systematic uncertainties are common for these two approaches, there are differences at the final stages which make a comparison useful. The signal model is taken from MC simulation after applying the corrections determined from data/simulation comparisons of $Z \to ee$ and $Z \to \mu\mu\gamma$ events mentioned above. The background is evaluated from a fit to the data without reference to the MC simulation. Since a Higgs boson signal would be reconstructed with a mass resolution approaching 1 GeV in the classes with best resolution, the limit and signal significance evaluation is carried out in steps of 0.5 GeV. The limits set on the cross section of a Higgs boson decaying to two photons relative to the SM expectation, using the CL_S computation, are shown for the 7 and 8 TeV datasets in Fig. 3 using the MVA selection and event classification as in ref. [35].

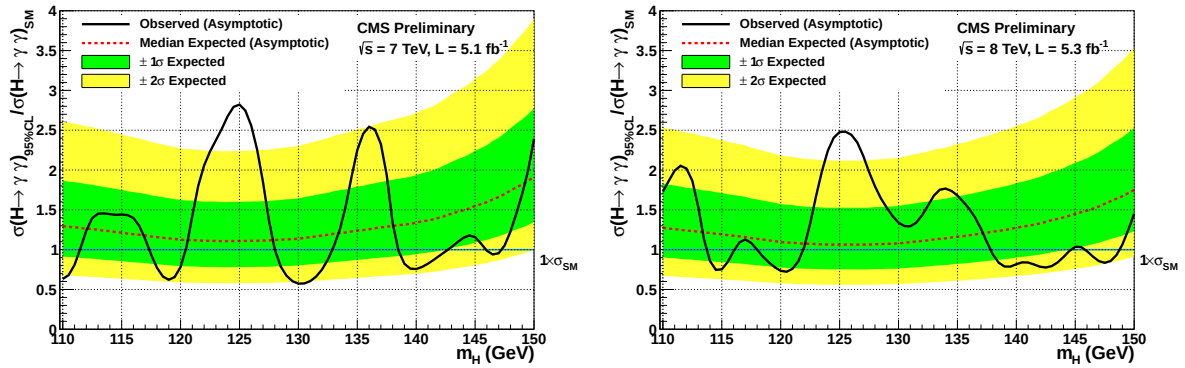
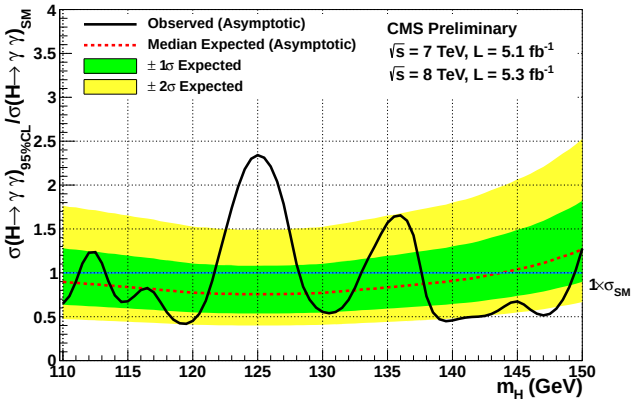


Figure 3: For the mass-fit MVA, the exclusion limit on the cross section of a SM Higgs boson decaying into two photons as a function of the boson mass relative to the SM cross section, where the theoretical uncertainties on the cross section have been included in the limit setting. The limit is calculated using the modified frequentist CL_S method. The panels show the results for the 7 and 8 TeV datasets.

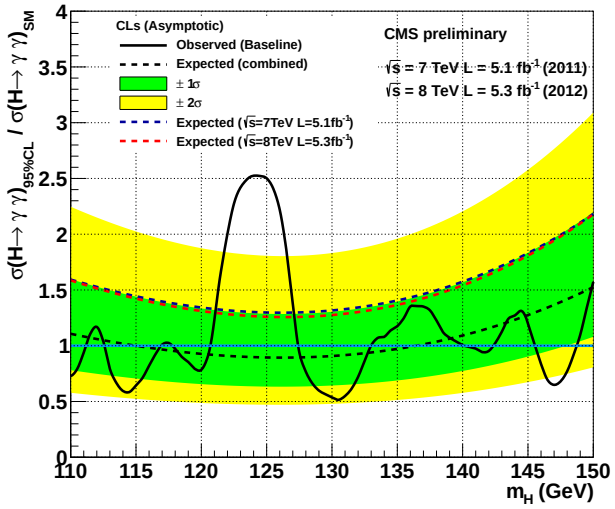
For the combined datasets, Fig. 4 shows the limit set on the cross section of a Higgs boson decaying to two photons relative to the SM expectation. For all three analysis procedures, a prominent excess can be seen in the limit plot at around m_H=125 GeV.

Figure 5 shows the local p-value calculated using the asymptotic approximation at 0.5 GeV intervals in the mass range $110 < m_H < 150$ GeV. The local p-values for the 7 and 8 TeV data samples are also shown. The local p-value quantifies the probability for the background to produce a fluctuation at least as large as that observed, and assumes that the relative signal strength between the event classes follows the MC signal model for the standard model Higgs boson.

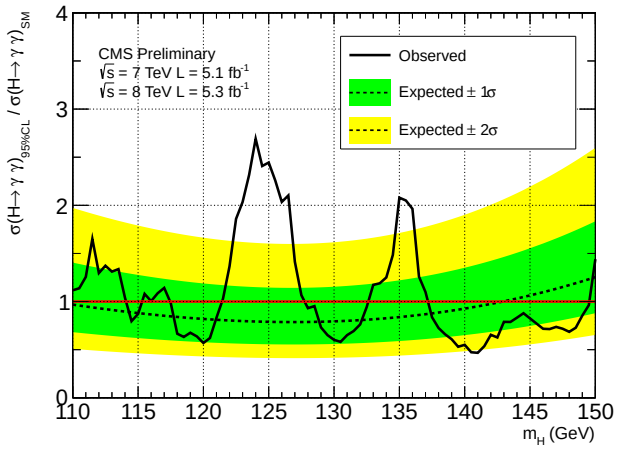
For the mass-fit MVA, the local p-value corresponding to the largest upwards fluctuation of the observed limit, at 125 GeV, has been computed to be 1.8×10^{-5} (4.1σ) in the asymptotic approximation. At this mass the best fit signal strength is 1.56 ± 0.43 times the standard model Higgs boson cross section. In Fig. 6 this combined best fit signal strength is shown as a function



(a) mass-fit MVA.

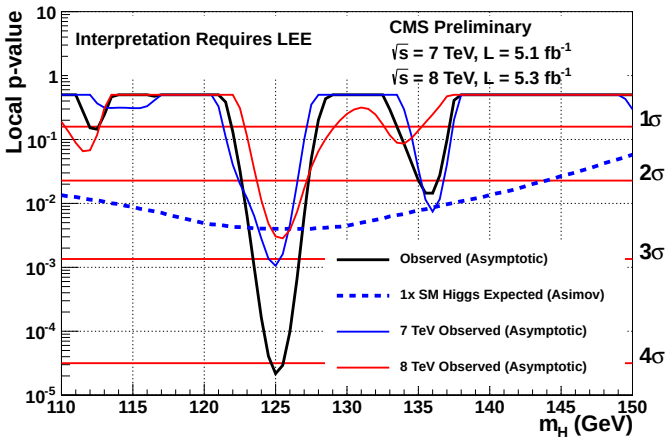


(b) Cut-based analysis.

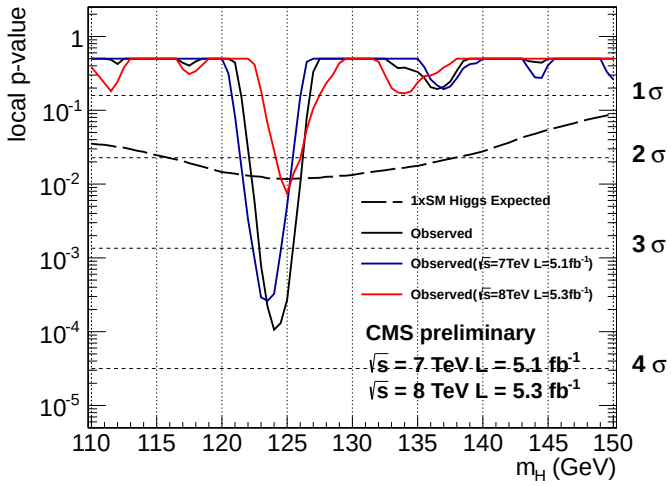


(c) mass window MVA.

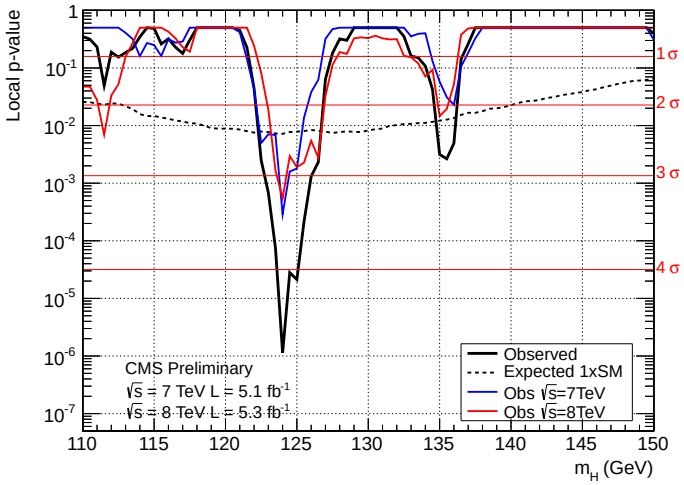
Figure 4: Limits on the cross section of a Higgs boson decaying to two photons relative to the SM expectation for the combined 7 and 8 TeV datasets, obtained with the three analysis methods. The primary result is shown in (a).



(a) mass-fit MVA.

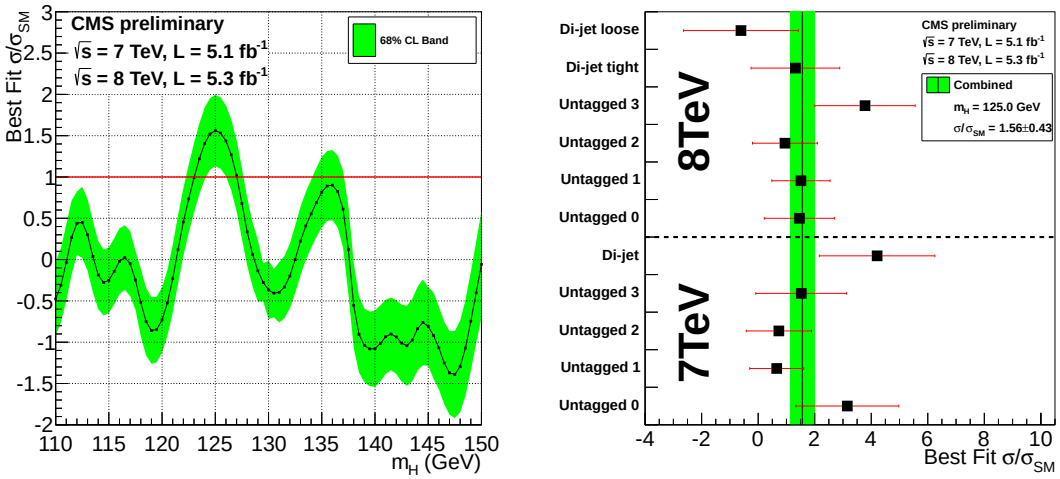


(b) Cut-based analysis.



(c) mass window MVA.

Figure 5: Observed local p-values as a function of m_{H} obtained with the three analysis methods. The primary result is shown in (a).



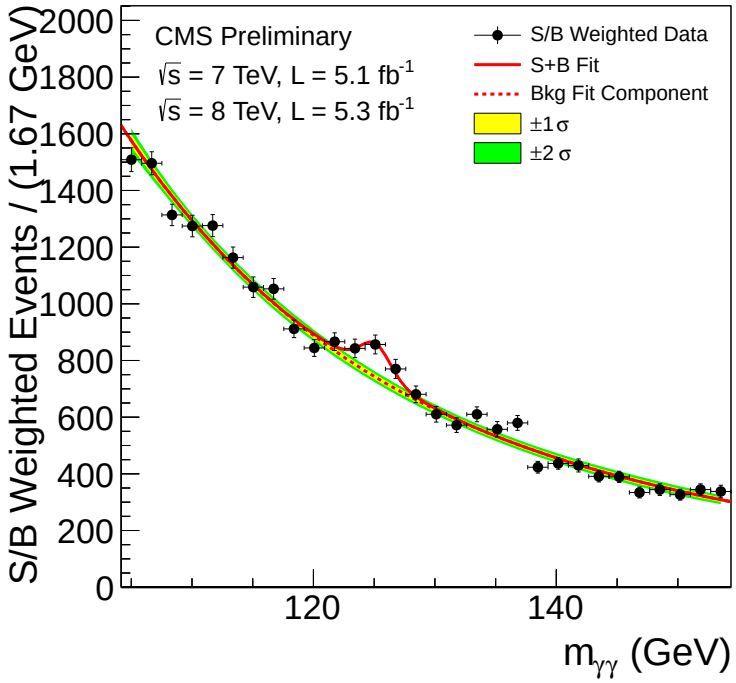
(a) Best fit signal strength as a function of m_{H}.

(b) Best fit signal strength at m_{H}=125 GeV for each event class.

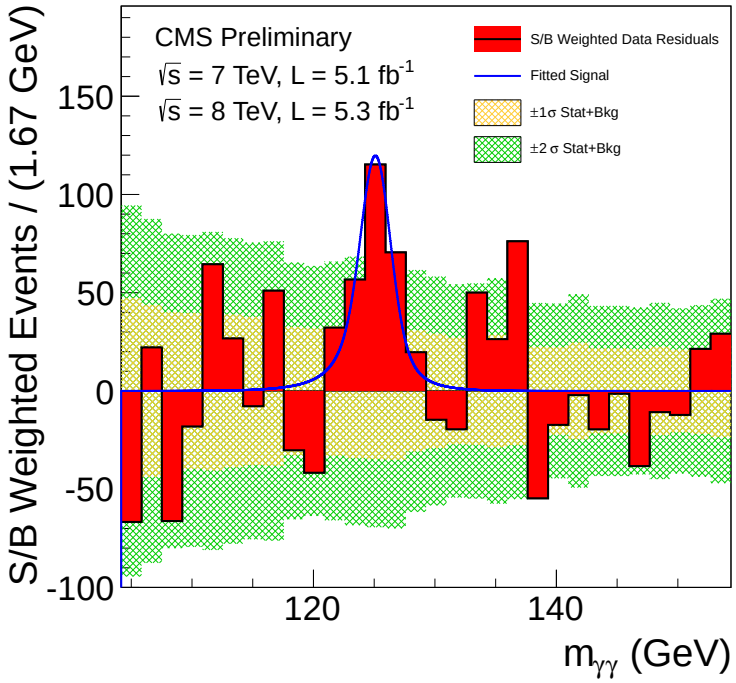
Figure 6: The best fit signal strength relative to the SM Higgs boson cross section. Left: scan as a function of the Higgs boson mass hypothesis. Right: the combined fit to the eleven classes. (vertical line) and for the individual contributing classes (points) for the hypothesis of a SM Higgs boson mass of 125.0 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.

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. Since a fluctuation of the background could occur at any point in the mass range, there is a look-elsewhere effect [76]. 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 ($110 < m_{\mathrm{H}} < 150$ GeV) is 7.2×10^{-4}, corresponding to a global significance of 3.2σ.

In order to visualize the significance yielded by the statistical methods in the $m_{\gamma\gamma}$ distribution, it is necessary to take into account the large differences in the signal-to-background ratio between the event classes. Figure 7 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 2. 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 σ_{eff} of the weighted signal model and the bins are centered at the best fit m_{H} value. Figure 7 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.



(a) S/B-weighted $m_{\gamma\gamma}$ distribution



(b) Background=subtracted S/B-weighted $m_{\gamma\gamma}$ distribution.

Figure 7: For the mass-fit MVA and combined datasets, result of summing the data and the signal-plus-background fits weighted by the ratio of signal-to-background in each event class as shown in Table 2. In the background-subtracted plot, the uncertainties shown around the horizontal axis correspond to the sum in quadrature of the estimated uncertainties on the background and the weighted data.

12 Conclusion

A search has been performed for the standard model Higgs boson decaying into two photons using data obtained from 5.1 fb^{-1} of pp collisions at $\sqrt{s} = 7\,\text{TeV}$ and 5.3 fb^{-1} at $\sqrt{s} = 8\,\text{TeV}$. The selected events are subdivided into classes according to a diphoton boosted decision tree output trained to quantify the signal-to-background ratio, and the results of a search in each class are combined. 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 standard model background is observed for a Higgs boson mass hypothesis of 125 GeV, where the expected limit is 0.76 times the standard model expectation. The local significance of this excess is 4.1σ. 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 3.2σ.

This result constitutes evidence for the existence of a new massive state that decays into two photons.

References

[1] S. L. Glashow, "Partial-symmetries of weak interactions", *Nucl. Phys.* **22** (1961) 579, doi:10.1016/0029-5582(61)90469-2.

[2] S. Weinberg, "A model of leptons", *Phys. Rev. Lett.* **19** (1967) 1264, doi:10.1103/PhysRevLett.19.1264.

[3] A. Salam, "Weak and electromagnetic interactions", in *Elementary particle physics: relativistic groups and analyticity*, N. Svartholm, ed., p. 367. Almquvist & Wiskell, 1968. 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, doi:10.1103/PhysRevLett.13.321.

[5] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132, doi:10.1016/0031-9163(64)91136-9.

[6] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (1964) 508, doi:10.1103/PhysRevLett.13.508.

[7] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles", *Phys. Rev. Lett.* **13** (1964) 585, doi:10.1103/PhysRevLett.13.585.

[8] P. W. Higgs, "Spontaneous symmetry breakdown without massless bosons", *Phys. Rev.* **145** (1966) 1156, doi:10.1103/PhysRev.145.1156.

[9] T. W. B. Kibble, "Symmetry breaking in non-Abelian gauge theories", *Phys. Rev.* **155** (1967) 1554, doi:10.1103/PhysRev.155.1554.

[10] ALEPH, DELPHI, L3, OPAL Collaborations, and The LEP Working Group for Higgs Boson Searches, "Search for the Standard Model Higgs boson at LEP", *Phys. Lett. B* **565** (2003) 61, doi:10.1016/S0370-2693(03)00614-2.

[11] CDF and D0 Collaborations, "Combination of Tevatron Searches for the Standard Model Higgs Boson in the W^+S^- Decay Mode", *Phys. Rev. Lett.* **104** (2010) 061802, doi:10.1103/PhysRevLett.104.061802. A more recent, unpublished, limit is given in preprint arXiv:1103.3233.

[12] ATLAS Collaboration Collaboration, "Combined search for the Standard Model Higgs boson in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector", arXiv:1207.0319.

[13] CMS Collaboration, "Combined results of searches for the standard model Higgs boson in pp collisions at sqrt(s) = 7 TeV", *Phys. Lett. B* **710** (2012) 26, doi:10.1016/j.physletb.2012.02.064, arXiv:1202.1488.

[14] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD Electroweak and heavy flavour groups, "Precision electroweak measurements and constraints on the Standard Model", (2010). arXiv:1012.2367.

[15] H. Georgi, S. Glashow, M. Machacek et al., "Higgs Bosons from Two Gluon Annihilation in Proton Proton Collisions", *Phys. Rev. Lett.* **40** (1978) 692, doi:10.1103/PhysRevLett.40.692.

[16] R. N. Cahn, S. D. Ellis, R. Kleiss et al., "Transverse Momentum Signatures for Heavy Higgs Bosons", *Phys.Rev.* **D35** (1987) 1626, doi:10.1103/PhysRevD.35.1626.

[17] S. Glashow, D. V. Nanopoulos, and A. Yildiz, "Associated Production of Higgs Bosons and Z Particles", *Phys. Rev. D* **18** (1978) 1724–1727, doi:10.1103/PhysRevD.18.1724.

[18] S. Dawson, "Radiative corrections to Higgs boson production", *Nucl. Phys. B* **359** (1991) 283, doi:10.1016/0550-3213(91)90061-2.

[19] M. Spira, A. Djouadi, D. Graudenz et al., "Higgs boson production at the LHC", *Nucl. Phys. B* **453** (1995) 17, doi:10.1016/0550-3213(95)00379-7.

[20] R. V. Harlander and W. B. Kilgore, "Next-to-next-to-leading order Higgs production at hadron colliders", *Phys. Rev. Lett.* **88** (2002) 201801, doi:10.1103/PhysRevLett.88.201801.

[21] C. Anastasiou and K. Melnikov, "Higgs boson production at hadron colliders in NNLO QCD", *Nucl. Phys. B* **646** (2002) 220, doi:10.1016/S0550-3213(02)00837-4.

[22] V. Ravindran, J. Smith and W.L. van Neerven, "NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions", *Nucl. Phys. B* **665** (2003) 325, doi:10.1016/S0550-3213(03)00457-7.

[23] S. Actis, G. Passarino, C. Sturm et al., "NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders", *Phys. Lett. B* **670** (2008) 12, doi:10.1016/j.physletb.2008.10.018.

[24] C. Anastasiou, R. Boughezal, and F. Petriello, "Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion", *JHEP* **04** (2009) 003, doi:10.1088/1126-6708/2009/04/003.

[25] D. L. Rainwater and D. Zeppenfeld, "Searching for H to gamma gamma in weak boson fusion at the LHC", *JHEP* **9712** (1997) 005, arXiv:hep-ph/9712271.

[26] P. Bolzoni, F. Maltoni, S.-O. Moch et al., "Higgs production via vector-boson fusion at NNLO in QCD", *Phys. Rev. Lett.* **105** (2010) 011801, doi:10.1103/PhysRevLett.105.011801.

[27] D. de Florian and M. Grazzini, "Higgs production through gluon fusion: updated cross sections at the Tevatron and the LHC", *Phys. Lett. B* **674** (2009) 291, doi:10.1016/j.physletb.2009.03.033.

[28] M. Ciccolini, A. Denner, and S. Dittmaier, "Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC", *Phys. Rev. Lett.* **99** (2007) 161803, doi:10.1103/PhysRevLett.99.161803.

[29] M. Ciccolini, A. Denner, and S. Dittmaier, "Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC", *Phys. Rev. D* **77** (2008) 013002, doi:10.1103/PhysRevD.77.013002.

[30] T. Han and S. Willenbrock, "QCD correction to the $pp \to WH$ and ZH total cross-sections", *Phys. Lett. B* **273** (1991) 167, doi:10.1016/0370-2693(91)90572-8.

[31] S. Actis, G. Passarino, C. Sturm et al., "NNLO Computational Techniques: the Cases $H \to \gamma\gamma$ and $H \to gg$", *Nucl. Phys. B* **811** (2009) 182, doi:10.1016/j.nuclphysb.2008.11.024, arXiv:0809.3667.

[32] D0 Collaboration, "Search for Resonant Diphoton Production with the D0 Detector", *Phys. Rev. Lett.* **102** (2009) 231801, doi:10.1103/PhysRevLett.102.231801.

[33] ATLAS Collaboration, "Search for the Standard Model Higgs boson in the two photon decay channel with the ATLAS detector at the LHC", *Phys. Lett. B* **705** (2011) 452, doi:10.1016/j.physletb.2011.10.051.

[34] CMS Collaboration, "Search for the standard model Higgs boson decaying into two photons in pp collisions at $\sqrt{s} = 7$ TeV", *Phys.Lett.* **B710** (2012), no. 3, 403–425, doi:10.1016/j.physletb.2012.03.003, arXiv:1202.1487v1.

[35] CMS Collaboration, "A search using multivariate techniques for a standard model Higgs boson decaying into two photons", *CDS Record* **1429931** (2011).

[36] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[37] A. Hoecker, P. Speckmayer, J. Stelzer et al., "TMVA: Toolkit for Multivariate Data Analysis", *PoS* **ACAT** (2007) 040, arXiv:physics/0703039.

[38] CMS Collaboration, "Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, (2009).

[39] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, (2010).

[40] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063.

[41] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **06** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002.

[42] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett.* **B659** (2008) 119, doi:10.1016/j.physletb.2007.09.077.

[43] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005.

[44] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", arXiv:1111.6097.

[45] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[46] CMS Collaboration, "Measurement of the inclusive W and Z production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the CMS experiment", *JHEP* **2011** (2011) 1, doi:10.1007/JHEP10(2011)132.

[47] G. Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727.

[48] A. Ballestrero, G. Bevilacqua, and E. Maina, "A complete parton level analysis of boson-boson scattering and ElectroWeak Symmetry Breaking in lv + four jets production at the LHC", *JHEP* **05** (2009) 015, doi:10.1088/1126-6708/2009/05/015.

[49] D. L. Rainwater, R. Szalapski, and D. Zeppenfeld, "Probing color singlet exchange in Z + two jet events at the CERN LHC", *Phys. Rev. D* **54** (1996) 6680, doi:10.1103/PhysRevD.54.6680.

[50] e. Bartalini, Paolo, R. Field, R. Chierici et al., "Multiple partonic interactions at the LHC. Proceedings, 1st International Workshop, MPI'08, Perugia, Italy, October 27-31, 2008", arXiv:1003.4220.

[51] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes", *Phys.Rev.* **D82** (2010) 074018, doi:10.1103/PhysRevD.82.074018, arXiv:1005.3457.

[52] A. Buckley, H. Hoeth, H. Lacker et al., "Systematic event generator tuning for the LHC", *Eur. Phys. J. C* **65** (2010) 331–357, doi:10.1140/epjc/s10052-009-1196-7, arXiv:0907.2973.

[53] R. Field, "Early LHC Underlying Event Data - Findings and Surprises", arXiv:1010.3558.

[54] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **0605** (2006) 026, doi:10.1088/1126-6708/2006/05/026.

[55] S. Alioli, P. Nason, C. Oleari et al., "NLO Higgs boson production via gluon fusion matched with shower in POWHEG", *JHEP* **04** (2009) 002, doi:10.1088/1126-6708/2009/04/002.

[56] P. Nason and C. Oleari, "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG", *JHEP* **02** (2010) 037, doi:10.1007/JHEP02(2010)037.

[57] G. Bozzi, S. Catani, D. de Florian et al., "The q(T) spectrum of the Higgs boson at the LHC in QCD perturbation theory", *Phys. Lett. B* **564** (2003) 65, doi:10.1016/S0370-2693(03)00656-7.

[58] G. Bozzi, S. Catani, D. de Florian et al., "Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC", *Nucl. Phys. B* **737** (2006) 73, doi:10.1016/j.nuclphysb.2005.12.022.

[59] D. de Florian, G. Ferrera, M. Grazzini et al., "Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC", *JHEP* **1111** (2011) 064, doi:10.1007/JHEP11(2011)064.

[60] Dixon, L. and Siu, S., "Resonance-continuum interference in the di-photon Higgs signal at the LHC", doi:10.1103/PhysRevLett.90.252001, arXiv:0302233v1.

[61] M. Botje, J. Butterworth, A. Cooper-Sarkar et al., "The PDF4LHC Working Group Interim Recommendations", (2011). arXiv:1101.0538.

[62] S. Alekhin, S. Alioli, R. D. Ball et al., "The PDF4LHC Working Group Interim Report", (2011). arXiv:1101.0536.

[63] H.-L. Lai, M. Guzzi, J. Huston et al., "New parton distributions for collider physics", *Phys. Rev. D* **82** (2010) 074024, doi:10.1103/PhysRevD.82.074024.

[64] A. Martin, W. Stirling, R. Thorne et al., "Parton distributions for the LHC", *Eur. Phys. J. C* **63** (2009) 189, doi:10.1140/epjc/s10052-009-1072-5.

[65] NNPDF Collaboration Collaboration, "Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology", (2011). arXiv:1101.1300.

[66] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables", CERN Report CERN-2011-002, (2011).

[67] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton showers simulations", *JHEP* **0206** (2002) 029, doi:10.1088/1126-6708/2002/06/029.

[68] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC", *Nucl.Phys.Proc.Suppl.* **205-206** (2010) 10–15, doi:10.1016/j.nuclphysbps.2010.08.011, arXiv:1007.3492. Talk presented by R.K Ellis at Loops and Legs in Quantum Field Theory 2010, Woerlitz, Germany, April 25-30, 2010, (6 pages and 4 figures).

[69] M. Guzzi, P. Nadolsky, E. Berger et al., "CT10 parton distributions and other developments in the global QCD analysis", arXiv:1101.0561.

[70] CMS Collaboration, "Measurement of the Drell-Yan Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV", *JHEP* **10** (2011) 7, doi:10.1007/JHEP10(2011)007.

[71] CMS Collaboration, "Measurement of the Production Cross Section for Pairs of Isolated Photons in pp collisions at $\sqrt{s} = 7$ TeV", (2011). arXiv:1110.6461. Submitted to *JHEP*.

[72] CMS Collaboration, "Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **700** (2011) 187, doi:10.1016/j.physletb.2011.05.027.

[73] A. Read, "Modified frequentist analysis of search results (the CL_s method)", Technical Report CERN-OPEN-2000-005, CERN, (2000).

[74] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[75] ATLAS and CMS Collaborations, LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", ATL-PHYS-PUB/CMS NOTE 2011-11, 2011/005, (2011).

[76] E. Gross and O. Vitells, "Trial factors for the look elsewhere effect in high energy physics", *Eur. Phys. J. C* **70** (2010) 525, doi:10.1140/epjc/s10052-010-1470-8, arXiv:1005.1891.