CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2013/03/21

Updated measurements of the Higgs boson at 125 GeV in the two photon decay channel

The CMS Collaboration

Abstract

Results are reported from a search for the Standard Model Higgs boson in the two photon channel in the mass range $110 < m_H < 150$ GeV, using the full dataset recorded by the CMS experiment at the LHC from pp collisions at centre-of-mass energies of 7 and 8 TeV. The most sensitive, MVA, analysis observes an excess of events at a mass of 125 GeV, with a local significance of 3.2σ (standard deviations), where a local significance of 4.2σ is expected from a standard model Higgs boson. The best-fit signal strength, σ/σ_{SM}, is 0.78 ± 0.27 at $m_H = 125$ GeV, and the mass is fitted to be $125.4 \pm 0.5(stat.) \pm 0.6(syst.)$. The cut-based analysis observes a corresponding excess with a local significance of 3.9σ (3.5σ expected), and $\sigma/\sigma_{SM} = 1.11 \pm 0.31$ at $m_H = 124.5$ GeV. All measurements are compatible among themselves and with a Higgs boson with a mass of 125.4 GeV.

1 Introduction

Within the Standard Model (SM) of particle physics [1–3], the masses of the particles arise from the spontaneous breaking of the electroweak symmetry which is implemented through the Higgs-mechanism. In its minimal version, this is realized through the introduction of a doublet of complex scalar fields. After breaking of the electroweak symmetry, only one scalar field is present in the theory and the corresponding quantum, the Higgs boson, should be experimentally observable [4–9]. In 2012 both the ATLAS [10] and the CMS [11] collaborations observed a new boson with an invariant mass of about 125 GeV whose properties are at present compatible with the Standard Model Higgs boson.

The $H \to \gamma\gamma$ decay channel provides a clean final-state topology which allows the mass to be reconstructed with high precision. In the mass range $110 < m_H < 150\,\text{GeV}$, $H \to \gamma\gamma$ is one of the most promising channels for the Higgs search at the LHC despite its low branching fraction varying between 0.14% and 0.23% [12]. The primary production mechanism of the Higgs boson at the LHC is gluon fusion [13] with additional smaller contributions from vector boson fusion (VBF) [14] and production in association with a W or Z boson [15], or with a $t\bar{t}$ pair [16–28].

This note presents an updated measurement [29–31] performed on the full dataset collected at the centre-of-mass energies of 7 and 8 TeV in the years 2011 and 2012. The datasets correspond to an integrated luminosity of $5.1\,\text{fb}^{-1}$ at 7 TeV and $19.6\,\text{fb}^{-1}$ at 8 TeV.

The analysis searches for a localized excess of diphoton events over a smoothly falling background due to prompt diphoton production and to events with at least one jet misidentified as photon. To achieve the best sensitivity to a Standard Model Higgs boson decaying to two photons, the events are separated into classes. The search results are presented for an analysis that uses Multi-Variate-Analysis (MVA) techniques both for photon identification and event classification, and extracts the signal from the background using a fit to the diphoton mass spectrum (referred to in the following as "mass-fit-MVA"). An independent analysis is also presented in which the photon identification and the events classification is cut-based and in which the background model is derived as before from a fit to the diphoton mass spectrum (referenced as "cut-based").

Additional event classes are defined to identify the events from specific production mechanisms, selecting events based on the presence of additional objects in the final state. The presence of two forward jets selects events produced by the vector boson fusion (VBF) production mechanism while events with a muon, an electron, or missing transverse energy (E_T^{miss}) target production in association with a vector boson (VH).

2 The CMS detector

A detailed description of the CMS detector can be found elsewhere [32]. Its 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 both the tracker and the calorimetry. The steel return yoke outside the solenoid is instrumented with gas detectors used to reconstruct and 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 $|\eta| < 1.479$ while

the ECAL endcaps cover the region $1.479 < |\eta| < 3.0$ A lead/silicon-strip preshower detector is located in front of the ECAL endcap in the region $1.653 < |\eta| < 2.6$. The preshower detector includes two planes of silicon sensors measuring the x and y coordinates of the impinging particles. A steel/quartz-fibre Čerenkov 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 photons from $\pi^0 \to \gamma\gamma$ and electrons from $W \to e\nu$, and $Z \to e^+e^-$ decays. Changes in the transparency of the ECAL crystals due to irradiation during the LHC running periods and their subsequent recovery are monitored continuously and corrected for, using light injected from a laser and LED system [33].

3 Data sample

The dataset consists of events collected with diphoton triggers and corresponds to an integrated luminosity of $5.1\,\text{fb}^{-1}$ at 7 TeV and $19.6\,\text{fb}^{-1}$ at 8 TeV. Diphoton triggers with asymmetric transverse energy (E_T) thresholds and complementary photon selections are used. One selection requires a loose calorimetric identification based on the shape of the electromagnetic shower and loose isolation requirements on the photon candidates, while the other requires only that the photon candidate has a high value of the R_9 shower shape variable.[1] The E_T thresholds at trigger level are 26 (18) GeV and 36 (22) GeV on the leading (trailing) photon depending on the running period. As the instantaneous luminosity delivered by the LHC increased, it became necessary to tighten the isolation requirement applied in the trigger. To maintain high trigger efficiency, all four combinations of threshold and selection criteria are deployed (i.e. with both photon candidates fulfilling the R_9 condition, with the high threshold candidate fulfilling the R_9 condition and the low threshold candidate fulfilling the loose ID and isolation, and so on). Accepting events that satisfy any of these triggers results in a measured trigger efficiency greater than 99.5% for events satisfying the preselection (see below). The total trigger efficiency is found to be uniform in the analysed datasets. Samples of Monte Carlo (MC) events, used in the analysis to describe the signal and to train the MVA discriminants, are fully simulated using GEANT [34]. The simulated events include the effects of pile-up (overlapping pp interactions in a bunch crossing), and are reweighed to reproduce the expected distribution of the number of interactions taking place in each bunch crossing.

4 Photon reconstruction and identification

Photon candidates are reconstructed from the energy deposits in the ECAL, grouping its channels into a supercluster. The superclustering algorithms achieve an almost complete collection of the energy of photons (and electrons) that convert into electron-positron pairs (emit bremstrahlung) 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 ϕ road. The photon candidates

[1] The R_9 variable is defined as the energy sum of 3×3 crystals centred on the most energetic crystal in the supercluster divided by the energy of the supercluster. Unconverted photons, having a narrow shower shape will tend to have high values of R_9, while converted photons, with wide shower shapes will tend to have lower values.

are collected within the ECAL fiducial region $|\eta| < 2.5$, excluding the barrel-endcap transition region $1.4442 < |\eta| < 1.566$. The fiducial region requirement is applied to the supercluster position (defined as the barycenter of the supercluster's active channels) in the ECAL, and a p_T threshold is applied after the vertex assignment (see section 6). The exclusion of the barrel-endcap transition region ensures complete containment of the accepted showers in either the ECAL barrel or endcaps .
About half of the photons convert in the material upstream the ECAL. Conversion track pairs are reconstructed from a combination of Gaussian-Sum-Filter (GSF) electron tracks and ECAL-seeded tracks fit to a common vertex and then matched to the photon candidate.

4.1 Photon energy

The photon energy is computed starting from the raw crystals energies recorded by the ECAL. 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 energy resolution, the crystal signals are calibrated to compensate several detector effects [33]. The variation of crystal transparency during the run is continuously monitored and corrected for using a correction factor based on the change in response to the light from the laser system. The single-channel response of the ECAL is equalized exploiting the ϕ-symmetry of the energy flow, the mass constraint on the energy of the two photons in π^0/η decays, and the momentum constraint on the energy of isolated electrons from W and Z decays. Finally, the containment of the shower in the clustered crystals, and the shower losses for photons which convert in the material upstream of the calorimeter are corrected using a multivariate regression technique based on a Boosted Decision Tree (BDT). 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. Among the regression input variables are the global η and ϕ coordinates of the supercluster, and a collection of shower shape variables: R_9 of the supercluster, the ratio of the 5x5 crystal energy centred around the seed crystal to the raw supercluster energy, the energy weighted η-width and ϕ-width of the supercluster and the ratio of hadronic energy behind the supercluster to the electromagnetic energy of the cluster. In the endcap, the ratio of preshower energy to raw supercluster energy is additionally included. Finally the number of primary vertices and median energy density ρ in the event are included in order to correct residual energy scale effects from pileup. A second BDT provides an event-by-event estimate of the energy uncertainty. It is trained on an independent sample of MC photons. It uses the same input variables as the first BDT but its target is the absolute deviation between the correction predicted by the first regression and the true correction to generator-level energy. The absolute energy scale and the residual long term drifts in the response are corrected using $Z \to e^+e^-$ decays. The photon energy resolution predicted by the MC simulation is corrected through the addition of a constant gaussian (smearing) term determined from the comparison of the $Z \to e^+e^-$ line-shape in data and MC. The smearing term is extracted differentially in the electrons η, R_9 and in run periods. Figure 1 shows that the energy resolution after corrections, estimated with $Z \to e^+e^-$ decays, is constant at the per mill level.

4.2 Photon identification

The dominant backgrounds to $H \to \gamma\gamma$ consist of an irreducible fraction from the prompt diphoton production, and a reducible one from $pp \to \gamma + \text{jet}$ and $pp \to \text{jet} + \text{jet}$ where one or more of the objects reconstructed as a photon corresponds to a jet. Typically, these jets contain a neutral meson taking a substantial fraction of the total jet p_T thus appearing as isolated. At large transverse momenta, the photons from the neutral mesons are collimated and they are reconstructed as a single photon. These reconstructed objects are generally referred to as *fake* photons. A photon identification BDT is used to distinguish prompt photons from the fake

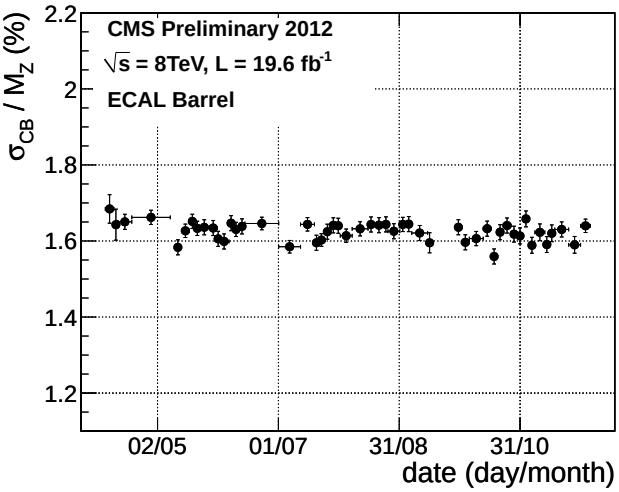


Figure 1: Instrumental mass resolution as a function of time, estimated with $Z \to e^+e^-$ decays, after corrections for the dataset at 8 TeV. This shows that the energy resolution of ECAL is stable at the per mill level.

ones for the mass-fit-MVA while a cut-based photon identification is used for the cut-based analysis.

The photons entering the analysis need to satisfy some preselection criteria matching the trigger requirements. These consist of an electron veto (removing the photon candidate if its supercluster is matched to a GSF-electron with no missing hits on the innermost tracker layers and it is not matched to a reconstructed conversion), a selection on the hadronic leakage of the shower (measured as the ratio of hadronic energy in HCAL towers behind the supercluster to the ECAL energy in the supercluster) and a loose selection based on isolation and on the shape of the shower.

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 those observed in $Z \to e^+e^-$ data samples, and cross-checked with $Z \to \mu^+\mu^-\gamma$ data events.

   (a) $\sigma_{i\eta i\eta}$, defined as:

   $$\sigma^2_{i\eta i\eta} = \frac{\sum(\eta_i - \bar{\eta})^2 w_i}{\sum w_i} \quad , \quad \text{where } \bar{\eta} = \frac{\sum w_i \eta_i}{\sum w_i} \quad \text{and} \quad w_i = \max\left[0 \quad ; \quad 4.7 + \log\left(\frac{E_i}{E_{5\times5}}\right)\right]$$

   where the sum runs over the 5×5 crystal matrix around the most energetic crystal in the supercluster, and the η distances are measured in units of the crystal size in the η direction.

   (b) $\mathrm{cov}(i\eta, i\phi)$, the off-diagonal element of the energy-weighed covariance matrix of single crystals η and ϕ within the 5×5 crystals centred at the crystal with maximum energy.

(c) $E_{2\times2}/E_{5\times5}$, the ratio of the energy in the 2×2 group 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, defined as the energy sum of 3×3 crystals centred on the most energetic crystal in the supercluster divided by the energy of the supercluster.

(e) σ_η, the energy weighed (crystal energy over supercluster energy) second moment of single crystals η within the supercluster.

(f) σ_ϕ, the energy weighed (crystal energy over supercluster energy) second moment of single crystals ϕ within the supercluster.

(g) σ_{RR}, the sum in quadrature of the second moment of the shower spread in the x and y planes of the preshower detector (for the range of pseudorapidities covered by the preshower detector)

2. Isolation variables, based on the particle flow algorithm [35].

   (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 energy density per unit area in the event. This variable is introduced to account for the pile-up dependence in the isolation variables [36].

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

The cut-based photon identification is instead performed by applying selection criteria on a set of discriminating variables. To obtain the best prompt-fake discrimination the criteria are optimized separately in four categories defined in terms of the photon's pseudorapidity and R_9. The photons candidates are categorized according to whether they are reconstructed within the ECAL barrel or endcap and on their value of R_9 (greater or smaller than 0.94). The variables used are:

1. the ratio of hadronic energy in HCAL towers behind the supercluster to the ECAL energy in the supercluster.

2. $\sigma_{i\eta i\eta}$ as described above

3. the three particle-flow based isolation variables (charged isolation, photon isolation, neutral isolation).

The isolation variables are corrected to maintain high efficiency under high pile-up conditions: the contribution of the pile-up and underlying event is estimated, on an event-by-event basis, as the product of the measured energy density ρ and an effective area corresponding to the size of the isolation cone.

The selection criteria were optimized on a Monte Carlo sample of γ+jet events to give the highest efficiency of accepted photons for a chosen signal to background ratio.

The efficiency of the photon preselection is measured in data using "tag and probe" technique [37]. The efficiency of all preselection criteria except the electron veto requirement is determined using $Z \to e^+e^-$ events. The efficiency for photons to satisfy the electron veto is measured using $Z \to \mu^+\mu^-\gamma$ events, where the photon is produced by final-state radiation. This provides a more than 99% pure source of prompt photons. The efficiency of the preselection criteria, for events with an Higgs boson of 125 GeV, ranges from 92% to 99% and the ratio $\epsilon_{data}/\epsilon_{MC}$ is consistent with 1 within uncertainties in all categories shown in Table 1. The measured $\epsilon_{data}/\epsilon_{MC}$ ratios are used to correct the simulated signal sample and the associated uncertainties are taken into account in the signal extraction procedure.

Table 1: Photon preselection efficiencies for the 8 TeV dataset measured in four photon categories.

Preselection category	ϵ_{data} (%)	ϵ_{MC} (%)	$\epsilon_{\mathrm{data}}/\epsilon_{\mathrm{MC}}$
Barrel; R_9 >0.90	0.9831 ± 0.0030	0.9867 ± 0.0001	0.996 ± 0.003
Barrel; R_9 <0.90	0.9257 ± 0.0055	0.9233 ± 0.0003	1.003 ± 0.006
Endcap; R_9 >0.90	0.9873 ± 0.0090	0.9784 ± 0.0002	1.009 ± 0.009
Endcap; R_9 <0.90	0.9375 ± 0.0170	0.9405 ± 0.0003	0.997 ± 0.017

5 Other reconstructed objects

The study of Higgs-production mechanisms other than the gluon fusion is performed by separating events with particular signatures. In the case of vector boson fusion (VBF) production, the Higgs boson is accompanied by jets separated by a large rapidity gap. In the case of the associated production (VH) the Higgs-boson is radiated from a heavy vector boson (W/Z). The selected events are accompanied by at least one charged lepton (a muon or an electron) or large $E_{\mathrm{T}}^{\mathrm{miss}}$ due to neutrinos from the decay of the W or Z boson.

5.1 Jets

Jets are reconstructed using a particle-flow algorithm [38, 39]. The basic objects of the particle-flow reconstruction are the tracks of charged particles reconstructed in the central tracker and in the muon system, and energy deposits reconstructed in the calorimeters. 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, γ + jet, and Z + jet events [41]. The energy from pile-up (PU) interactions and from the underlying event are also included in the reconstructed jets. This energy is subtracted using the FASTJET technique [36, 42, 43], which is based on the calculation of the η-dependent transverse momentum density, evaluated on an event-by-event basis. Particles produced in PU interactions may be clustered into jets of relatively large transverse momentum, referred to as PU jets. These PU jets are removed using selection criteria based on the compatibility of the tracks in a jet with the primary vertex or on the width of the jet. Finally, jets within $\Delta R < 0.5$ (where $\Delta R = \sqrt{\phi^2 + \eta^2}$) with respect to any of the two photons are rejected.

5.2 Muons

Muons are reconstructed with the particle-flow algorithm and are required to have a $p_{\mathrm{T}} >$ 20 GeV and to be within $|\eta| < 2.4$. A tight selection is applied, based on the quality of the track

and the number of tracker and muon spectrometer hits. A strict match between the tracker and the muon spectrometer segments is also applied, to reduce the contamination from muons produced in decays of hadrons and from beam halo. Finally, a loose particle-flow isolation requirement is applied [44].

5.3 Electrons

Electrons are identified as clusters of energy deposited in the ECAL matched to GSF tracks. Electron candidates are required to have a cluster with $E_{\mathrm{T}} > 20$ GeV and $|\eta| < 1.4442$ or $1.566 < |\eta| < 2.5$. The electron identification is based on a multivariate technique [45]. The electron track has to fulfill requirements on the transverse and longitudinal impact parameter with respect to the electron vertex and cannot have more than one missing hit in the innermost layers of the tracker. Fake electrons from conversions are excluded as described in [46] and a loose particle-flow isolation condition is applied.

5.4 Missing transverse energy

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed by the particle-flow algorithm [38]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from signal events overlaps with a component due to detector effects like noise, inaccurate jet energy reconstruction and relative mis-alignment among the subdetectors, which cannot be reproduced with sufficient accuracy in the MC simulations. To reach the desired level of data/MC agreement (both in magnitude and ϕ-direction) a set of corrections derived from data have been applied. These corrections on the jet energy resolution (smearing) and on the average difference between reconstructed and generated p_{T} (scale) are applied to the reconstructed jets. The corrected particle-flow $E_{\mathrm{T}}^{\mathrm{miss}}$ is then required to be above 70 GeV.

6 Diphoton vertex

The mean number of pp interactions per bunch crossing is 9.5 (19.9) in the 7 TeV (8 TeV) dataset with a corresponding variance of 4.8(7.3). The interaction vertices are built using the reconstructed tracks from the produced charged particles. Their distribution in the longitudinal direction (z), has an RMS spread of about 6 cm (5 cm) in the 7 TeV (8 TeV) dataset.

The diphoton mass resolution is driven by the photon energy resolution and the knowledge about the direction of the photons, which is dominated by the knowledge of the vertex where they originate. The relative contribution from the vertex assignment to the mass resolution becomes negligible with respect to the photon energy resolution when the distance between the chosen vertex and the true one is below 1 cm.

6.1 Diphoton vertex identification

Since photons are neutral particles, and therefore do not leave ionization signal in the tracker, the diphoton vertex is identified indirectly. The vertex can be identified using the kinematic properties of the diphoton system and its correlations with the kinematic properties of the recoiling tracks. If either of the photons converts, the direction of the converted photon tracks can be used to identify the diphoton interaction vertex.

For the determination of the diphoton 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 ($p_{\mathrm{T}}^{\gamma\gamma}$). The three variables are: $\sum p_{\mathrm{T}}^2$, the sum of the transverse momentum squared

of the tracks associated with each vertex; -$\sum(\vec{p}_T \cdot \frac{\vec{p}_T^{\gamma\gamma}}{|\vec{p}_T^{\gamma\gamma}|})$, the balance between all the tracks associated with each vertex and the diphoton system in the transverse plane; and $(|\sum \vec{p}_T| - p_{\mathrm{T}}^{\gamma\gamma})/(|\sum \vec{p}_T| + p_{\mathrm{T}}^{\gamma\gamma})$, the asymmetry between the sum of the transverse momenta of all the tracks associated with each vertex and the diphoton system. The conversion information is used, whenever available, through the estimation of the compatibility of each vertex from the longitudinal location on the beam axis pointed to by any reconstructed tracks associated with the diphoton candidates: $pull_{conv} = |z_{conversion} - z_{vertex}|/\sigma_{conversion}$. The variables are then input to a multivariate system based on a BDT to choose the reconstructed vertex to be associated with the diphoton. The vertex is assigned following the presented strategy for all event classes as defined in section 7.

The vertex-finding efficiency, defined as the efficiency to locate the vertex to within 1 cm of its true position, has been measured with $\mathrm{Z} \to \mu^+\mu^-$ events. The algorithm is run after the removal of the muon tracks to mimic the presence of the two photons. The use of tracks from a converted photon to locate the vertex is validated 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 vertex finding efficiency in simulated Higgs boson signal events. The overall vertex-finding efficiency for a Higgs boson of mass 120 GeV, integrated over its p_{T} spectrum, is computed to be 80% in the 8 TeV dataset.

6.2 Per event vertex probability

For the mass-fit-MVA analysis a second vertex-related multivariate discriminant has been designed to estimate, event-by-event, the probability for the vertex assignment to be within 1 cm of the diphoton interaction point. This, used in conjunction with the event-by-event estimate of the energy resolution of each photon is used to estimate the diphoton mass resolution (see section 7). A BDT was trained, using simulated $\mathrm{H} \to \gamma\gamma$ events, to separate events where the chosen vertex lies within 1 cm of the generated interaction point. The inputs of the BDT are:

- the values of the vertex BDT output for the three most likely vertices in each event,
- the number of vertices in each event,
- the transverse momentum of the diphoton system,
- the distances between the chosen vertex and the second and third choices,
- the number of photons with an associated conversion.

The vertex probability is estimated as a linear fit to the BDT output score and it is validated in data using $\mathrm{Z} \to \mu\mu$ and $\gamma +$ jet events.

7 Event classification

The analysis uses events with two photon candidates satisfying the preselection requirements (see section 4.2) and with $p_{\mathrm{T}}^{\gamma}(1) > m_{\gamma\gamma}/3$ and $p_{\mathrm{T}}^{\gamma}(2) > m_{\gamma\gamma}/4$ (where $m_{\gamma\gamma}$ is the event by event diphoton invariant mass). In case of multiple diphoton candidates, the one with the highest scalar sum of the photons transverse momenta is selected.

To achieve the best analysis performance, the events are first separated into classes based on their mass resolution and signal to background ratio, analyzed separately and the individual results combined in a simultaneous statistic treatment of all event classes.

The first step in the classification of the events extracts those with specific signatures: first the

muon tagged events are selected then, from the remainder, the electron events, then the dijet events and lastly the $E_{\mathrm{T}}^{\mathrm{miss}}$ events. The events remaining untagged are further subdivided into classes based on a BDT classifier in the case of the mass-fit-MVA analysis, and based on the minimum R_9 and the maximum pseudorapidity of the two photons for the cut-based analysis.

In this section the classification of the untagged and tagged events is detailed.

7.1 Untagged events

The mass-fit-MVA analysis use a multivariate event classifier to discriminate diphoton Higgs events from the diphoton, photon-fake, and fake-fake continua. For this analysis an additional requirement is then applied on the photon identification BDT output which retains more than 99% of the simulated signal events fulfilling the preselection requirements, while removing 23.5% of the data events in the region of $100 < m_{\gamma\gamma} < 180\,\mathrm{GeV}$.
The multivariate event classifier is constructed to satisfy the following criteria:

1. The BDT 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 mass of the Higgs signal used for the training sample.

The multivariate classifier incorporates the kinematic properties of the diphoton system (excluding $m_{\gamma\gamma}$), a per-event estimate of the diphoton mass resolution and the photon identification BDT output value. This choice of inputs is justified by the fact that the signal-to-background ratio varies as a function of the photons kinematics. In addition, the diphoton mass resolution depends on several factors: 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 true diphoton vertex has been identified. For the complete list of variables used in the BDT see reference [31].

The multivariate classifier assigns a score to each event in data. High scores correspond to signal-like events, while low-scores to background-like ones. For this reason the events in the mass-fit-MVA analysis are split into classes depending on the score of the classifier. An iterative procedure has been devised to sequentially split the sample into several classes. Each split is introduced setting the boundary value that gives rise to the best expected exclusion limit. The optimized quantity is the median expected 95% confidence level exclusion limit on the signal strength modifier computed using frequentist CLs approach [47]. The optimization is performed using signal and background simulated events.

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. The lowest score class, contributing to less than 1% to the sensitivity has been dropped, leaving only four categories for the analysis. This is equivalent to set a lower bound to the diphoton BDT output value. Out of the preselected events, this selection removes 73% of diphoton events in the mass window $100 < m_{\gamma\gamma} < 180\,\mathrm{GeV}$ while removing only 28% of simulated Higgs boson events (at $m_{\mathrm{H}} = 125\,\mathrm{GeV}$).

Investigating the properties of the simulated signal events in these classes reveals, as expected, 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 situated in the central barrel region of the ECAL.

In contrast, the cut-based analysis uses a simple classification based on the the minimum R9 and the maximum pseudorapidity of the two photons. Both variables are effective in separating diphotons with good mass resolution from those with worse resolution and in separating events with a higher signal to background ratio from those with a lower one. The class boundary values are chosen to match those used in the photon categories for the photon identification cuts: both photons in barrel and the minimum of R_9 greater than 0.94; both photons in barrel and the minimum of R_9 smaller than 0.94; one or more photons in endcap and the minimum of R_9 greater than 0.94; one or more photons in endcap and the minimum of R_9 smaller than 0.94.

7.2 Events tagged by exclusive signatures

Vector boson fusion (VBF) events are produced together with 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 selection it is possible to define classes of events which have an expected signal-to-background ratio more than an order of magnitude larger than events in the four classes previously defined.

The two highest transverse energy (E_{T}) jets in the event are required to be within pseudorapidity $|\eta| < 4.7$. The pseudorapidity restriction avoids the use of jets for which the energy corrections are less reliable and it is found to have only a small effect (<2% change) on the signal efficiency.

While in the 7 TeV dataset one single class of dijet-tagged events is used [31], in the 8 TeV dataset analysis two classes are defined. The mass-fit-MVA analysis uses a multivariate approach (BDT) to classify the events while the cut-based splits the events into two classes depending on requirements imposed on a few selection variables.

For the mass-fit-MVA the highest p_{T} photon is required to have $p_{\mathrm{T}}^{\gamma}(1) > m_{\gamma\gamma}/2$ while the second highest $p_{\mathrm{T}}^{\gamma}(2) > m_{\gamma\gamma}/4\,\mathrm{GeV}$. The variables used in the dijet BDT are: the transverse momenta of the leading and sub-leading photons divided by the invariant mass of the diphoton system ($p_{\mathrm{T}}^{\gamma}(1)/m_{\gamma\gamma}$, $p_{\mathrm{T}}^{\gamma}(2)/m_{\gamma\gamma}$), the transverse momenta of the leading and sub-leading jets, the dijet invariant mass, the difference between the jets pseudorapidities, the difference between the average pseudorapidity of the two jets and the pseudorapidity of the diphoton system, and the difference in azimuthal angle between the diphoton system and the dijet system. The events are separated into two classes based on the output of the BDT, and they have a signal to background ratio of about 0.5 and 0.2 respectively.

For the cut-based analysis the highest p_{T} photon is required to have $p_{\mathrm{T}}^{\gamma}(1) > m_{\gamma\gamma}/2$ while the second photon $p_{\mathrm{T}}^{\gamma}(2) > 25\,\mathrm{GeV}$; the difference in pseudorapidity between the two jets has to be greater than 3, the Zeppenfeld variable ($Z = \eta(\gamma_1 + \gamma_2) - \frac{\eta(j_1)+\eta(j_2)}{2}$) less than 2.5 and the difference in azimuthal angle between the dijet system and the diphoton system greater than 2.6. The dijet selected events are then split into two exclusive categories based on the invariant mass of the dijet system. Events in the first category are required to have $m_{j_1 j_2} > 500$ GeV and to have both jets with $p_{\mathrm{T}} > 30$ GeV; the remaining events are required to have $m_{j_1 j_2} > 250$ GeV with a loosened selection on the second jet p_{T} at 20 GeV.

Events satisfying the muon, electron or $E_{\mathrm{T}}^{\mathrm{miss}}$ selection criteria are sensitive to Higgs bosons produced by the associate production mechanism (VH). The selected Higgs boson events pro-

duced in association with a $W^{\pm}$ or a Z may have at least one charged lepton in the final state or, if the Z decays into two neutrinos or the lepton from the $W^{\pm}$ is not reconstructed, they will have large $E_{\mathrm{T}}^{\mathrm{miss}}$. The two photons in the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ tagged events are required to have $p_{\mathrm{T}}^{\gamma}(1)/m_{\gamma\gamma} > 45/120$ and $p_{\mathrm{T}}^{\gamma}(2) > m_{\gamma\gamma}/4$ for the mass-fit-MVA analysis, while for the cut-based analysis the requirement on the second photon is $p_{\mathrm{T}}^{\gamma}(2) > 25$ GeV. For the $E_{\mathrm{T}}^{\mathrm{miss}}$ tagged events, only events with both photons in the ECAL barrel are retained for this analysis because of the better mass resolution. It has been verified that the resulting loss in sensitivity is negligible. Additional selection criteria are applied to the angular separation between the $E_{\mathrm{T}}^{\mathrm{miss}}$ direction and the diphoton system and the leading jet. Leptons (muons and electrons) are required to be separated from the closest photon by $\Delta R(\gamma, lepton) > 1.0$, and the invariant mass of the lepton-photon pair is required to be more than 10 GeV away from the Z-pole mass. In addition a conversion veto is applied to the electrons to reduce the photon fakes.

8 Signal model

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 [49, 50] interfaced with PYTHIA [51]. For the dominant gluon fusion process, the Higgs boson transverse momentum spectrum has been reweighted to the NNLL + NLO distribution computed by the HQT program [52–54] for the 7 TeV sample. At 8 TeV, POWHEG has been tuned following the recommendations of the LHC Higgs Cross-Section-Working-Group [55] to reproduce the HQT (NNLO+NNLL) spectrum. The gluon fusion process cross-section is reduced by 2.5% for all values of m_{H} to account for the interference with the QCD diphoton production [56]. The simulated events are reweighted to reproduce the observed distribution of the number of interactions taking place in each bunch crossing. The SM Higgs boson cross sections and branching fractions used are taken from ref. [57].

A parametric signal model is constructed separately for each event class and for each production mechanism from a fit of the MC mass peak to the sum of two or three Gaussians, after applying the corrections determined from data/MC comparisons of $Z \to e^+e^-$ and $Z \to \mu^+\mu^-\gamma$ events.

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 from data at $m_{\gamma\gamma}$ =125 GeV for each of the 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 (gaussian equivalent).

9 Background model

In the analysis the background estimation is entirely data driven. The contribution to the background in the diphoton mass range $110 < m_{\gamma\gamma} < 150$ GeV is dominated by the diphoton continuum which is predicted to be about 70% of the total in simulations.

The background models for the mass-fit-MVA and the cut-based analyses are obtained by fitting the observed diphoton mass distributions in each of the event classes over the range $100 < m_{\gamma\gamma} < 180$ GeV. The choice of background parametrisation in each event class is fully data-driven and starts with considering families of functions that could a priori describe the

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 for the mass-fit-MVA analysis. 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	61.4%	16.8%	18.7%	3.1%	1.21	1.14	3.3	$\pm$ 0.4
	Untagged 1	16.3	87.6%	6.2%	5.6%	0.5%	1.26	1.08	37.5	$\pm$ 1.3
	Untagged 2	21.5	91.3%	4.4%	3.9%	0.3%	1.59	1.32	74.8	$\pm$ 1.9
	Untagged 3	32.8	91.3%	4.4%	4.1%	0.2%	2.47	2.07	193.6	$\pm$ 3.0
	Dijet tag	2.9	26.8%	72.5%	0.6%	–	1.73	1.37	1.7	$\pm$ 0.2
8 TeV 19.6 fb^{-1}	Untagged 0	17.0	72.9%	11.6%	12.9%	2.6%	1.36	1.27	22.1	$\pm$ 0.5
	Untagged 1	37.8	83.5%	8.4%	7.1%	1.0%	1.50	1.39	94.3	$\pm$ 1.0
	Untagged 2	150.2	91.6%	4.5%	3.6%	0.4%	1.77	1.54	570.5	$\pm$ 2.6
	Untagged 3	159.9	92.5%	3.9%	3.3%	0.3%	2.61	2.14	1060.9	$\pm$ 3.5
	Dijet tight	9.2	20.7%	78.9%	0.3%	0.1%	1.79	1.50	3.4	$\pm$ 0.2
	Dijet loose	11.5	47.0%	50.9%	1.7%	0.5%	1.87	1.60	12.4	$\pm$ 0.4
	Muon tag	1.4	0.0%	0.2%	79.0%	20.8%	1.85	1.52	0.7	$\pm$ 0.1
	Electron tag	0.9	1.1%	0.4%	78.7%	19.8%	1.88	1.54	0.7	$\pm$ 0.1
	E_T^{miss} tag	1.7	22.0%	2.6%	63.7%	11.7%	1.79	1.64	1.8	$\pm$ 0.1

background distribution, including sums of exponentials, sums of power law terms, Laurent series, and polynomials. In each family, the number of degrees of freedom (number of exponentials, number of terms in the series, degree of the polynomial, etc) is increased until the F-test between N+1 degrees of freedom and N degrees of freedom for the fit to data shows no significant improvement ($p-$value<0.05), and the function with N degrees of freedom is retained as representative of that family of functions. The value of N is determined separately both for each functional family and for each event class. These "truth" functions are used to generate sets of toy MC. The same functional families are then used as "fit" functions for the toy MCs. For each pair of "truth-fit" functions a potential bias on the signal estimation is computed. The final background model, for each event class, is chosen to be the one giving a maximum potential bias on the fitted signal strength less than five times the statistical uncertainty on the background. Provided that the potential bias fulfils this condition, the systematic associated to the background shape can be safely neglected.

Polynomials of orders from 2 to 5 are found to fulfil the requirements above and are used to model the background distributions in the various categories for both the mass-fit-MVA and the cut-based analyses. More details about the procedure used to choose the appropriate functional form are reported in [31].

The $m_{\gamma\gamma}$ distributions for the mass-fit-MVA analysis in the different event classes, together with the results of the simultaneous fit of the signal plus background model to the fourteen classes, are shown in Figures 2-4 for the 7 and 8 TeV data samples respectively. The uncertainty bands shown on the background component of the fit are computed from the fit uncertainty on the background yield in bins corresponding to those used to display the data.

The signal model shown are the result of the fit to an overall signal strength at each mass

hypothesis, simultaneously in each event class.

10 Systematic uncertainties

The systematic uncertainties affecting the signal considered in the analysis performed on the 8 TeV dataset are summarized in Table 3. The systematic uncertainties on the 7 TeV dataset are reported in reference [30]. The methods used to estimate them are reported here below. Most of the systematic uncertainties are common to all analyses; analysis specific ones are detailed where needed.

The systematic uncertainties calculated at the single photon level are:

- *Energy scale and resolution:* scale and resolutions are studied with electrons from $Z \to e^+e^-$ and then applied to photons. The main source of systematic uncertainty is the different interactions of electrons and photons with material upstream the ECAL. Uncertainties are assessed by changing the rescaling of R_9 distributions, changing the R_9 selection, the regression training and the electron selection used.
- *Photon identification:* taken as the largest uncertainty on the data/MC scale factors computed on $Z \to e^+e^-$ events using a tag-and-probe technique. This systematic uncertainty is applied to the photon identification in the cut-based analysis and to the loose photon preselection in the MVA analysis.
- *R_9 selection (cut-based):* taken from the data/MC comparison of the photon R_9 categorization in $Z \to \mu\mu\gamma$ events. The statistical uncertainty on the single photon is propagated to the diphoton categories and the result is assigned as systematic uncertainty on the category migration between low and high R_9 categories.
- *Photon identification BDT and photon energy resolution BDT (MVA analysis):* the agreement between data and simulation is assessed using $Z \to e^+e^-$ candidates, $Z \to \mu^+\mu^-\gamma$ candidates and the highest transverse energy photon in the diphoton invariant mass region where $m_{\gamma\gamma} > 160\,\text{GeV}$ (the fake photon contribution becomes smaller at high diphoton invariant mass). Both the inputs to the diphoton BDT and its output value are compared. A variation of ± 0.01 on the photon identification BDT output value, together with an uncertainty on the per-photon energy resolution estimate, parametrized as a rescaling of the resolution estimate by $\pm 10\%$ about its nominal value, fully covers the differences observed between data and MC simulation.

The systematic uncertainties calculated at the event level are:

- *Integrated luminosity:* the luminosity uncertainty is estimated as described in [58].
- *Vertex finding efficiency:* taken from the statistical uncertainty on the data/MC scale factor on $Z \to \mu^+\mu^-$ and the uncertainty on the signal p_{T} distribution arising from theory uncertainties.
- *Trigger efficiency:* extracted from $Z \to e^+e^-$ using a tag-and-probe technique and rescaling them to take into account the different R_9 distributions for electrons and photons.
- *Global energy scale:* an uncertainty of 0.25% to account for the imperfect modelling by the MC of electron/photon differences and 0.4% to account for possible non-linearity when extrapolating from the Z-mass scale to the $m_H \sim 125$ GeV, are added. The two uncertainties, which together amount to 0.47%, are fully correlated between all the

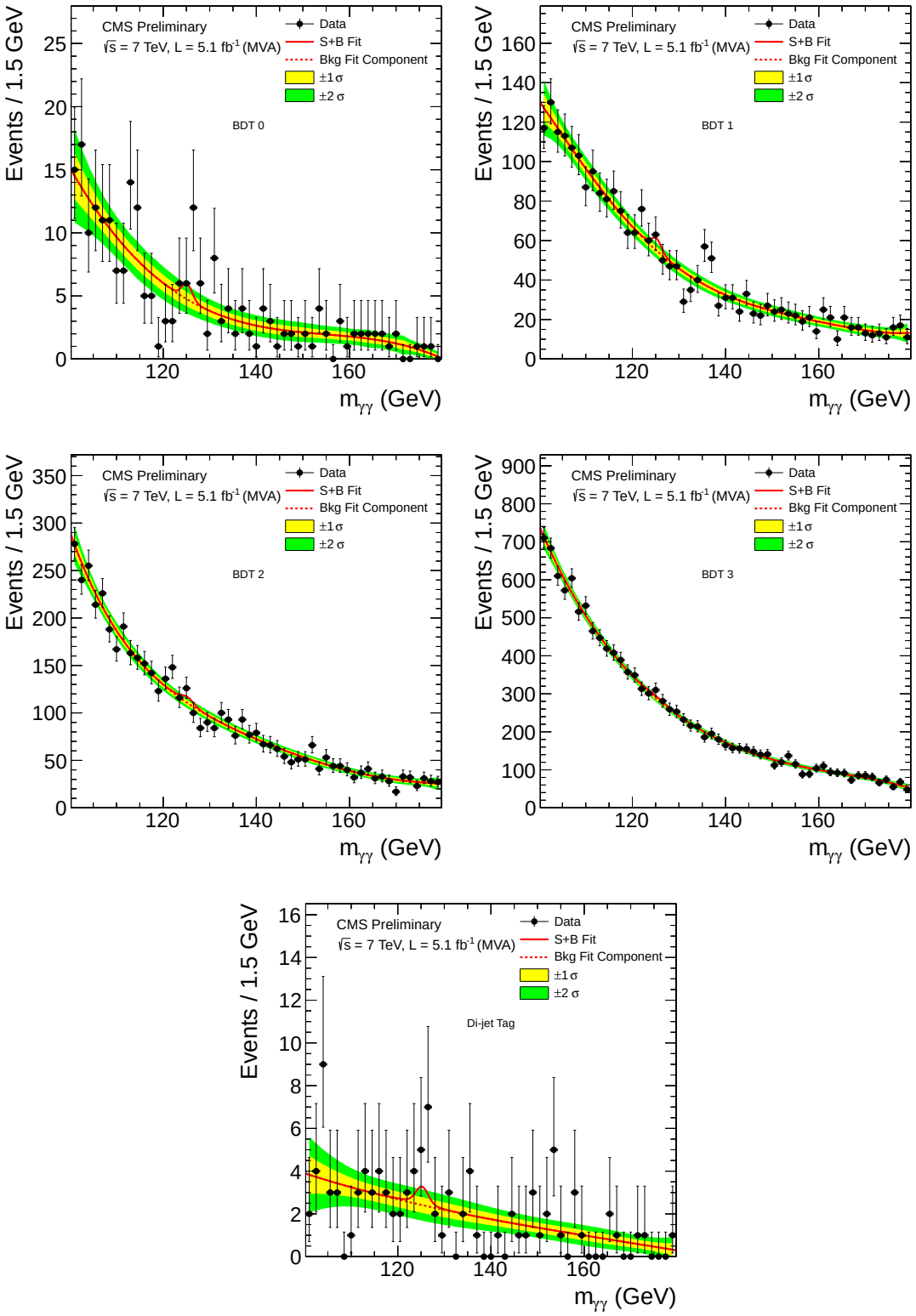


Figure 2: Signal plus background model fits to the four inclusive classes and the dijet class for the 7 TeV dataset of the mass-fit-MVA analysis.

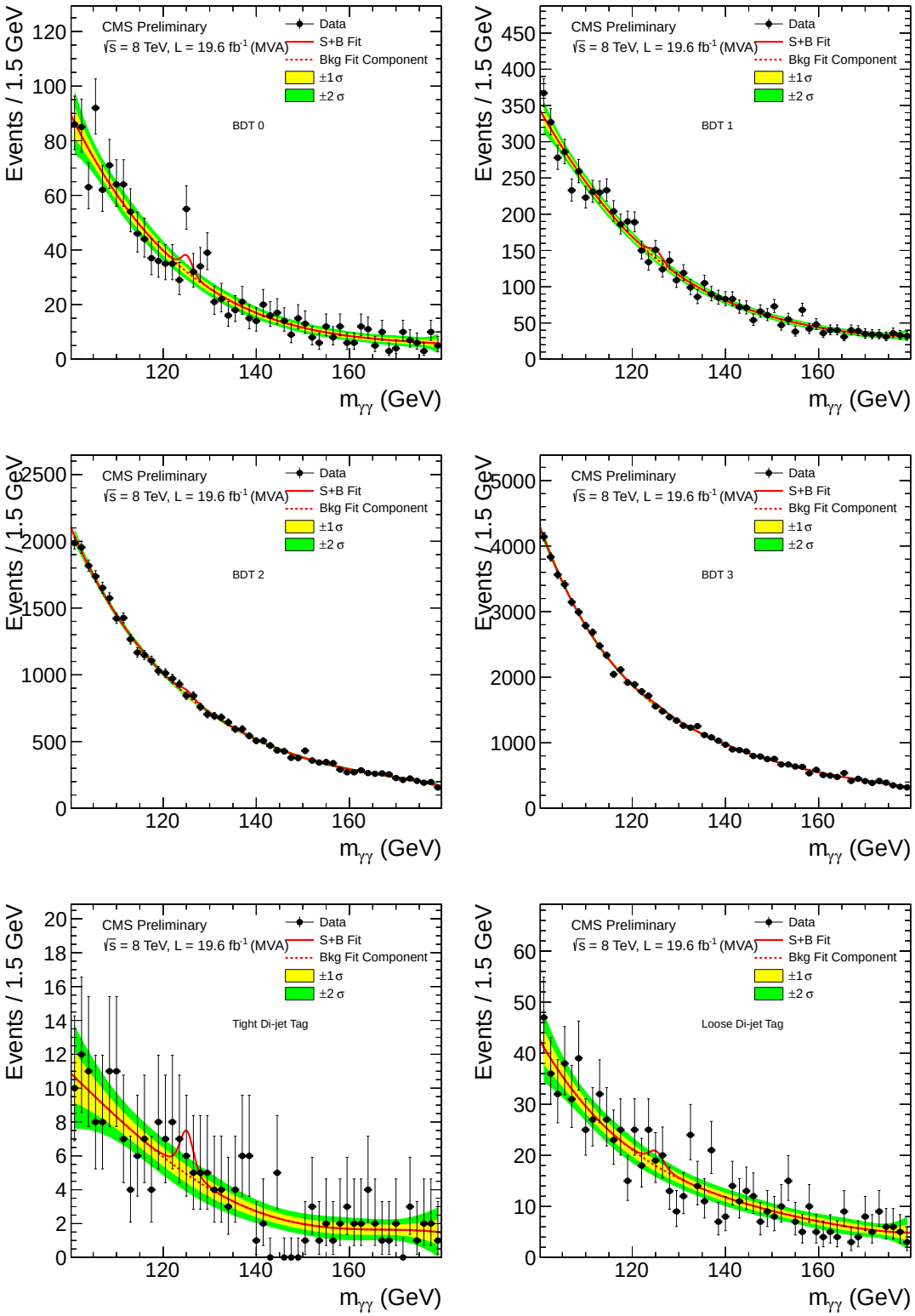


Figure 3: Signal plus background model fits to the four inclusive classes and the two dijet classes for the 8 TeV dataset of the mass-fit-MVA analysis.

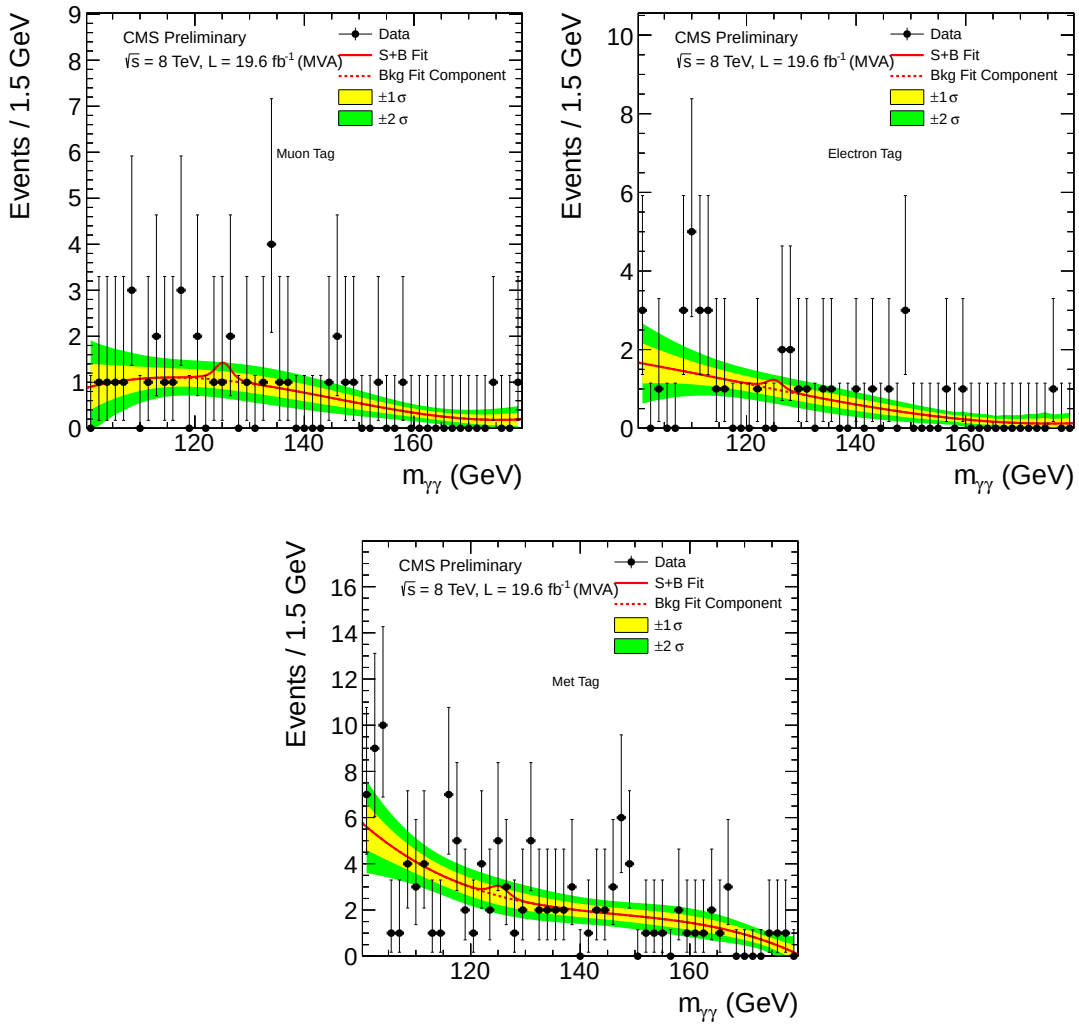


Figure 4: Signal plus background model fits to the muon, electron and MET classes for the 8 TeV dataset of the mass-fit-MVA analysis.

categories of the analysis.

The systematic uncertainties for the events with exclusive signatures are:

- *dijet tagging efficiency:* two effects are taken into account: the uncertainty on the MC modelling of the jet-energy corrections and resolution; and the uncertainty in predicting the presence of jets and their kinematic properties. They are calculated using different underlying event tunes and from the uncertainty on parton distribution functions and QCD scale factor. The uncertainty on the underlying event tunes was investigated by comparing the DT6 [59], P0 [60], ProPT0 and ProQ20 [61] tunes to the Z2 tune [62] in PYTHIA [51]. The effect of PDF uncertainties has been determined by varying the Higgs boson kinematic properties according to the variations of the eigenvalues of the CT10 [63] PDF set within their uncertainties. While the systematic uncertainties are assessed separately for the mass-fit-MVA and the cut-based analyses, the numerical values turns out to be numerically similar.
- *Lepton identification efficiency:* for both electrons and muons, the uncertainty on the identification efficiency is computed varying the data/simulation efficiency scale factor by its uncertainty. The resulting difference in the signal efficiency estimated in the MC simulation is taken as systematic uncertainty.
- E_{T}^{miss} *selection efficiency:* Systematic uncertainties due to $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are estimated in both signal events where real $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected (i.e. VH production) and the other Higgs production modes. For VH signal events the uncertainty is estimated by applying or not the $E_{\mathrm{T}}^{\mathrm{miss}}$ corrections and taking the difference in efficiency as a systematic uncertainty. For the other modes (gluon-gluon fusion, VBF and $\mathrm{t\bar{t}}$) the uncertainty is mainly due to a different fraction of events in the tail of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. The systematic uncertainty is evaluated by comparing diphoton data and MC in a control sample enriched in γ+jet events which looks similar (in terms of $E_{\mathrm{T}}^{\mathrm{miss}}$) to the Higgs signal events.

The theoretical systematic uncertainties considered are:

- *Production cross section:* the systematic uncertainty on the production cross section follows the recommendation of the LHC Higgs Cross Section working group [57].

Table 3: Separate sources of systematic uncertainties accounted for in the analysis of the 8 TeV data set.

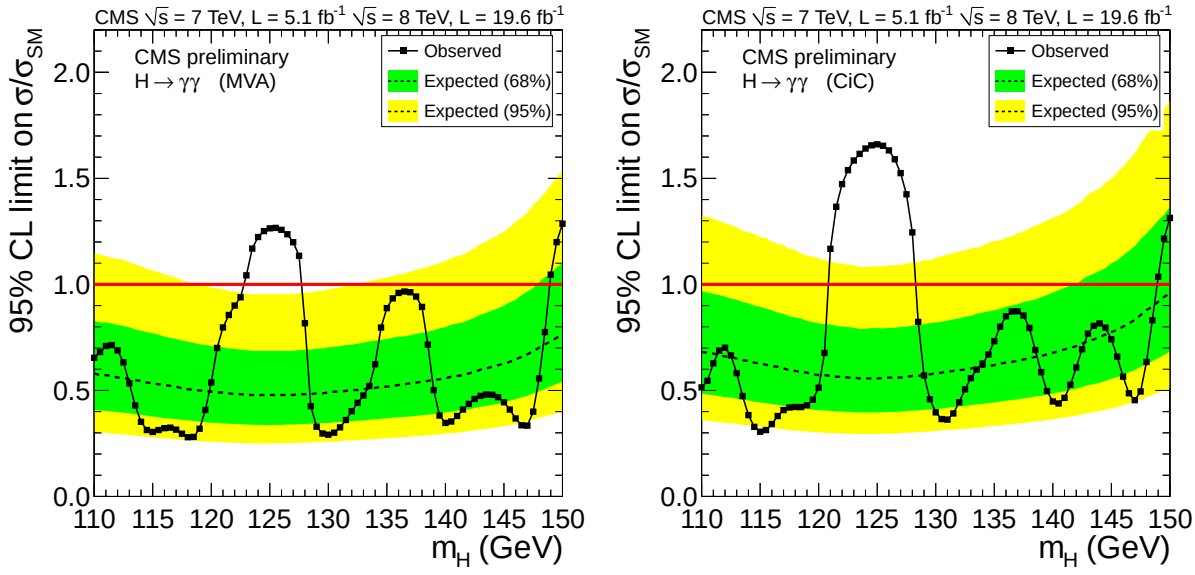
Sources of systematic uncertainty		**Uncertainty**	
Per photon		Barrel	Endcap
Energy resolution ($\Delta\sigma/E_{MC}$)	$R_9 > 0.94$ (low η, high η)	0.23%, 0.72%	0.93%, 0.36%
	$R_9 < 0.94$ (low η, high η)	0.25%, 0.60%	0.33%, 0.54%
Energy scale ($(E_{data} - E_{MC})/E_{MC}$)	$R_9 > 0.94$ (low η, high η)	0.20%, 0.71%	0.88%, 0.12%
	$R_9 < 0.94$ (low η, high η)	0.20%, 0.51%	0.18%, 0.12%
Photon identification efficiency		1.0%	2.6%
Cut-based			
$R_9 >$ 0.94 efficiency (results in class migration)		4.0%	6.5%
MVA analyses			
Photon identification BDT		$\pm$0.01 (shape shift)	
(Effect of up to 4.3% event class migration.)			
Photon energy resolution BDT		$\pm$10% (shape scaling)	
(Effect of up to 8.1% event class migration.)			
Per event			
Integrated luminosity		4.4%	
Vertex finding efficiency		0.2%	
Trigger efficiency		1.0%	
Global energy scale		0.47%	
Dijet selection			
Dijet-tagging efficiency	VBF process	10%	
	Gluon-gluon fusion process	30%	
(Effect of up to 15% event migration among dijet classes.)			
Muon selection			
Muon identification efficiency		1.0%	
Electron selection			
Electron identification efficiency		1.0%	
$E_{\mathrm{T}}^{\mathrm{miss}}$ **selection**			
$E_{\mathrm{T}}^{\mathrm{miss}}$ cut efficiency	Gluon-gluon fusion	15%	
	Vector boson fusion	15%	
	Associated production with W/Z	4%	
	Associated production with $\mathrm{t\bar{t}}$	4%	
Production cross sections		Scale	PDF
Gluon-gluon fusion		+7.6% -8.2%	+7.6% -7.0%
Vector boson fusion		+0.3% -0.8%	+2.6% -2.8%
Associated production with W/Z		+2.1% -1.8%	4.2%
Associated production with $\mathrm{t\bar{t}}$		+4.1% -9.4%	8.0%

11 Results

The SM Higgs boson hypothesis is tested against the background-only hypothesis performing a simultaneous fit to the diphoton invariant mass distributions in the various event classes under each of the two hypotheses. The results on the 8 TeV dataset are combined with the 7 TeV dataset as described in [31].

The 95% confidence level exclusion limits on the signal strength modifier are evaluated using a modified frequentist approach, CL_S, taking the profile likelihood ratio as a test statistic [64–66]. The signal model is taken from MC simulation after applying the corrections determined from data/MC comparisons of $Z \to e^+e^-$ and $Z \to \mu^+\mu^-\gamma$ events as described in Section 8. The background is evaluated from a fit to the data without reference to the MC simulation as described in Section 9. The theoretical uncertainties on the cross section have been included in the limit setting.

The limits on the production cross section times branching ratio of a Higgs boson decaying to two photons relative to the SM expectation, using the CL_S computation, are shown in Figure 5. The mass range between 110 and 149 GeV is excluded at 95% confidence level, except the region between 122.8 and 127.8 where an excess of events is found. Similar regions are obtained for the cut-based analysis.

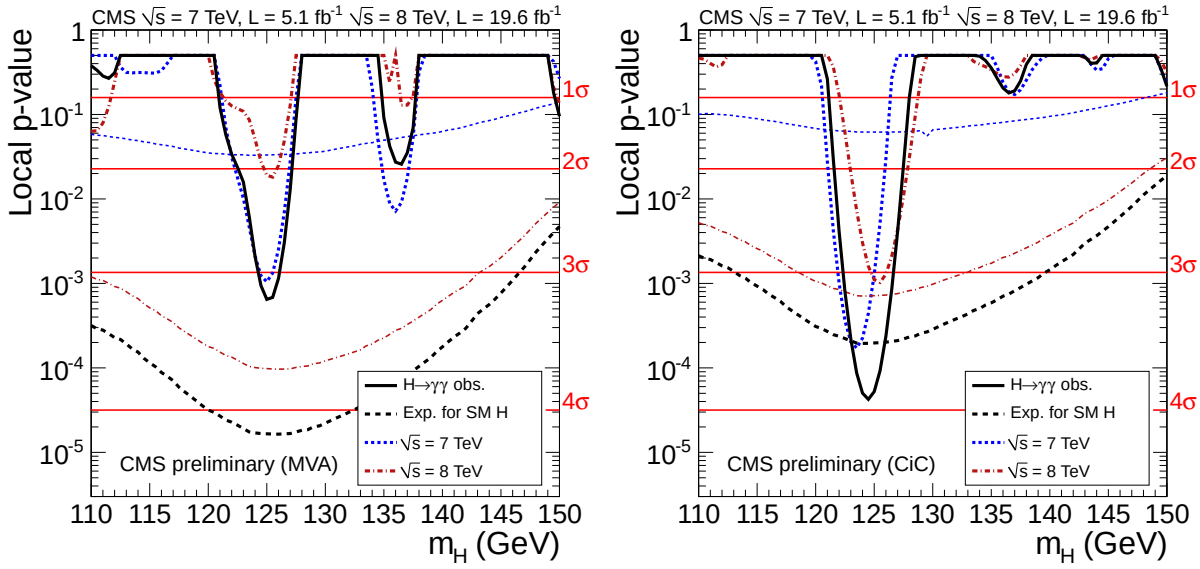


(a) Mass-fit-MVA.

(b) Cut-based analysis.

Figure 5: The exclusion limit on the cross section of a SM Higgs boson decaying into two photons as a function of the boson mass hypothesis relative to the SM cross section. The results are shown for the mass-fit-MVA analysis (left) and the cut-based analysis (right).

Figure 6 shows the local p-values for the different analyses, calculated using the asymptotic approximation in the mass range $110 < m_H < 150$ GeV for the 7 TeV + 8 TeV datasets combined as well as for the 7 TeV and the 8 TeV datasets separately. The local p-value quantifies the probability for the background to produce a fluctuation as large as the observed one or larger, 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 analysis, the local p-value corresponding to the largest signal-like fluctuation of the observed limit, at 125 GeV, has been computed to be 3.2 σ in the asymptotic approximation where a local significance of 4.2σ is



(a) Mass-fit-MVA.

(b) Cut-based analysis.

Figure 6: Observed local p-values as a function of m_{H}. The results are shown for the mass-fit-MVA analysis (left) and the cut-based analysis (right).

	mass-fit-MVA (at m_H = 125 GeV)	cut-based (at m_H = 124.5 GeV)
7 TeV	$1.69^{+0.65}_{-0.59}$	$2.27^{+0.80}_{-0.74}$
8 TeV	$0.55^{+0.29}_{-0.27}$	$0.93^{+0.34}_{-0.30}$
7 + 8 TeV	$0.78^{+0.28}_{-0.26}$	$1.11^{+0.32}_{-0.30}$

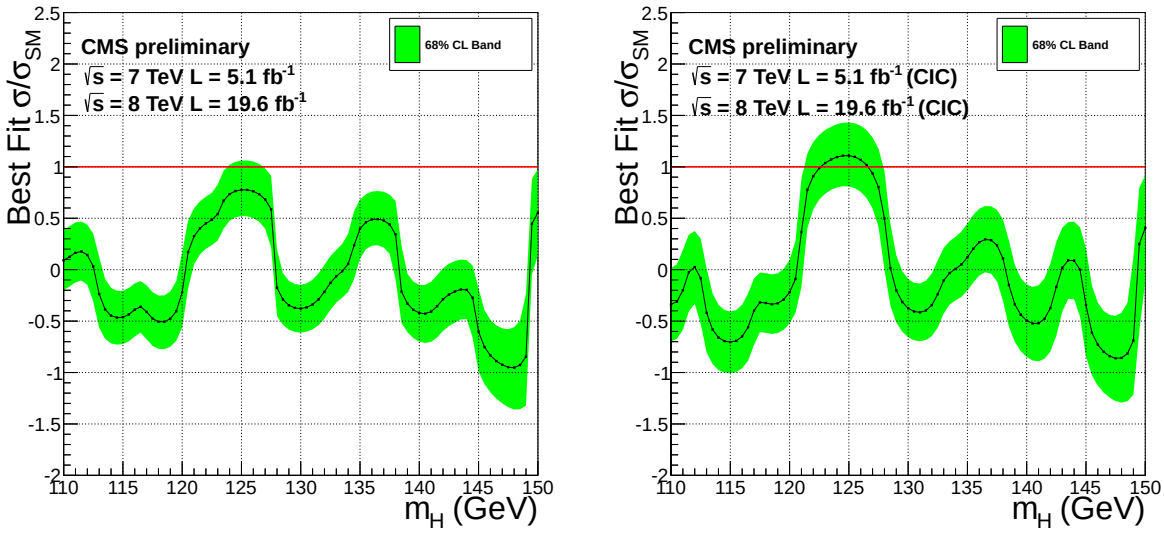
Table 4: The values of the best fit signal strength for the different datasets and analyses.

expected from a Standard Model Higgs boson. For the cut-based analysis, the largest signal-like fluctuation is observed at 124.5 GeV with a corresponding value of 3.9σ ($3.5\ \sigma$ expected).

In Figure 7 the combined best fit signal strength is shown as a function of the Higgs boson mass hypothesis, for both the mass-fit-MVA analysis and the cut-based one.

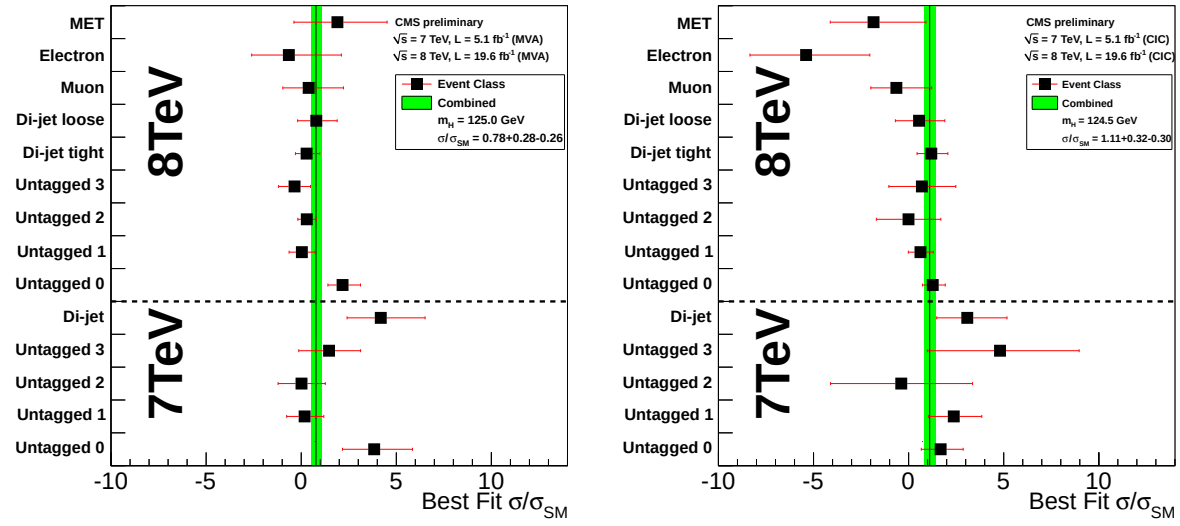
The best fit signal strength corresponding to the largest signal like fluctuation at 125 GeV is $\sigma/\sigma_{SM} = 0.78^{+0.28}_{-0.26}$ for the mass-fit-MVA analysis and $\sigma/\sigma_{SM} = 1.11^{+0.32}_{-0.30}$ at the mass of 124.5 GeV for the cut-based analysis. The compatibility between these results is reported in section 11.3. The values of the best fit signal strength, derived separately for the 7 and 8 TeV datasets for the two analyses, are reported in table 4. As a further cross-check, a second MVA-based analysis which uses a background model derived from the signal sidebands [11], has also been performed giving compatible results.

In Figure 8 the best fit signal strengths is shown in each of the event classes and separately for the 7 TeV and 8 TeV datasets. The vertical line corresponds to the SM Higgs boson mass hypothesis corresponding to the largest signal-like fluctuation in Figure 7: 125 GeV for the mass-fit-MVA analysis and 124.5 GeV for the cut-based analysis. The band corresponds to $\pm 1\sigma$ uncertainties on the overall value.



(a) Mass-fit-MVA (b) Cut-based analysis

Figure 7: The best fit signal strength relative to the SM Higgs boson cross section. The results are shown for the mass-fit-MVA analysis (left) and the cut-based analysis (right).



(a) Mass-fit-MVA (b) Cut-based analysis

Figure 8: The combined fit to the fourteen classes (vertical line) and for the individual contributing classes (points) for the hypothesis of a SM Higgs boson mass of 125 GeV (left) for the mass-fit-MVA and 124.5 GeV (right) for the cut-based. 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.

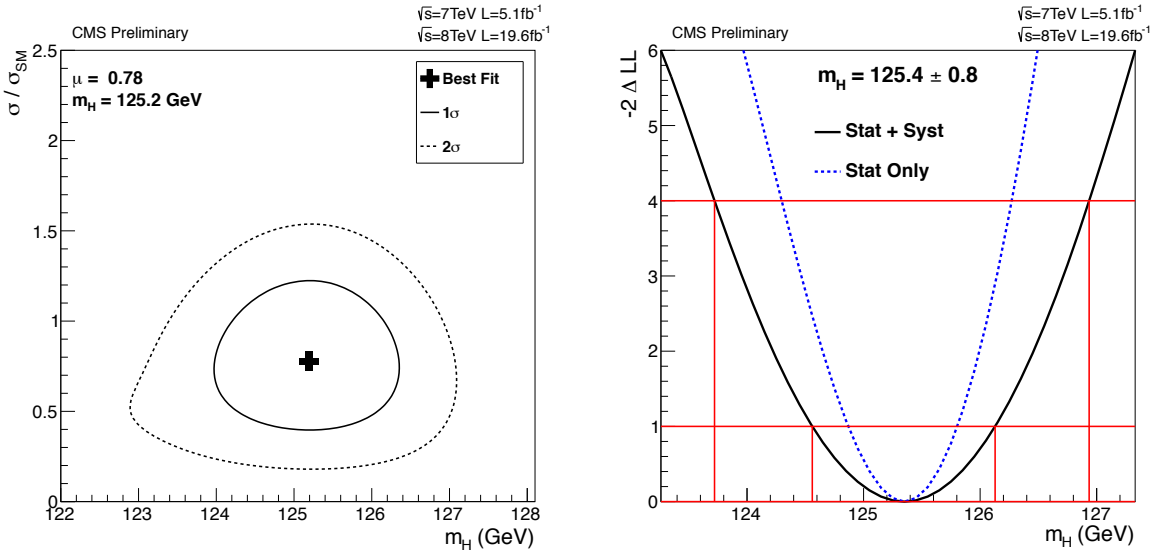


Figure 9: (left) The 2D 68% confidence level region for the signal strength modifier μ and the mass of the observed particle. (right) The scan of the negative-log-likelihood as a function of the hypothesised mass.

11.1 Mass results

The mass of the observed boson is measured to be $125.4 \pm 0.5(stat.) \pm 0.6(syst.)$ GeV. The calibration of the energy scale is achieved with reference to the known Z boson mass, as described in Section 4.1. The two main sources of systematic uncertainty are: (i) imperfect simulation of the detector response to electrons/photons and (ii) the mis-modeling of the detector linearity in the extrapolation from the Z to the Higgs energy scale. The systematic uncertainties are evaluated by making comparisons between data and simulated samples of $Z \to e^+e^-$ and $H \to \gamma\gamma$ with $m_H = 125$ GeV. The two uncertainties, which together amount to 0.47%, are assumed to be fully correlated between all the event categories in the 7 and 8 TeV data. Additive scale uncertainties between event categories are also included but have a sub-dominant effect on the overall systematic uncertainty.

Figure 9 shows the 2D 68% and 95% CL contours for the signal strength modifier μ and the mass of the observed particle and the scan of the negative-log-likelihood as a function of the hypothesised mass.

Figure 10 and Figure 11 show the weighted diphoton invariant mass spectrum and the corresponding background subtracted one for the mass-fit-MVA and the cut-based analyses respectively. The weights are the ratio of signal to signal plus background in a mass window of $\pm$ 1 σ_{eff} for each category, around their respective best fit mass.

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 roughly match the σ_{eff} of the weighted signal model and the bins are centred at the best fit value of m_{H}.

The uncertainty shown around the horizontal axis in the background subtracted plots corresponds to the sum in quadrature of the estimated uncertainties on the weighted background model and the weighted data.

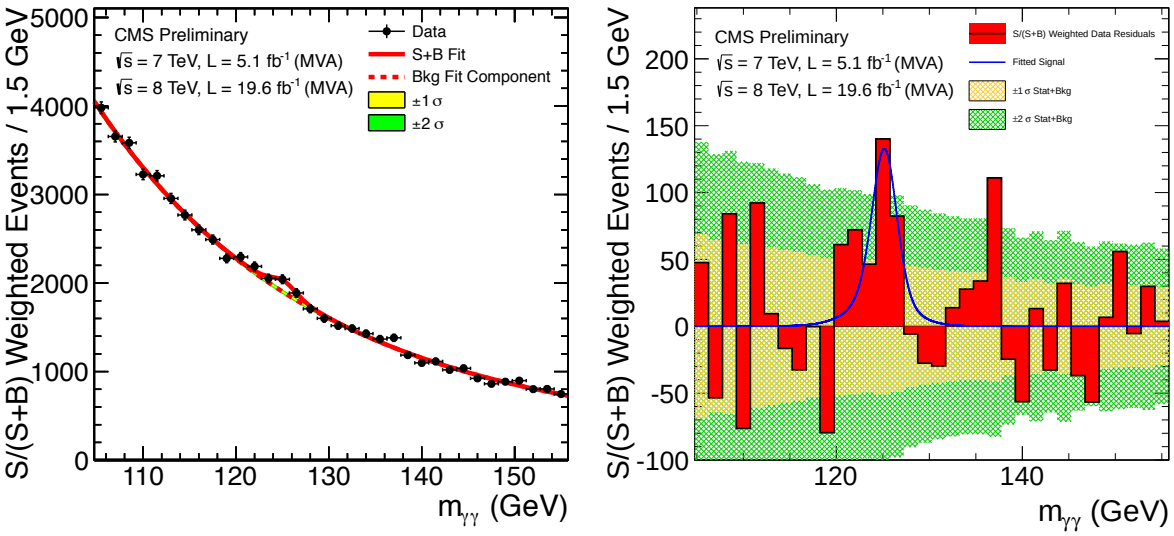


Figure 10: (left) The diphoton mass spectrum weighted by the ratio of signal-to-background in each event class for the mass-fit-MVA analysis. (right) The background-subtracted weighted mass spectrum.

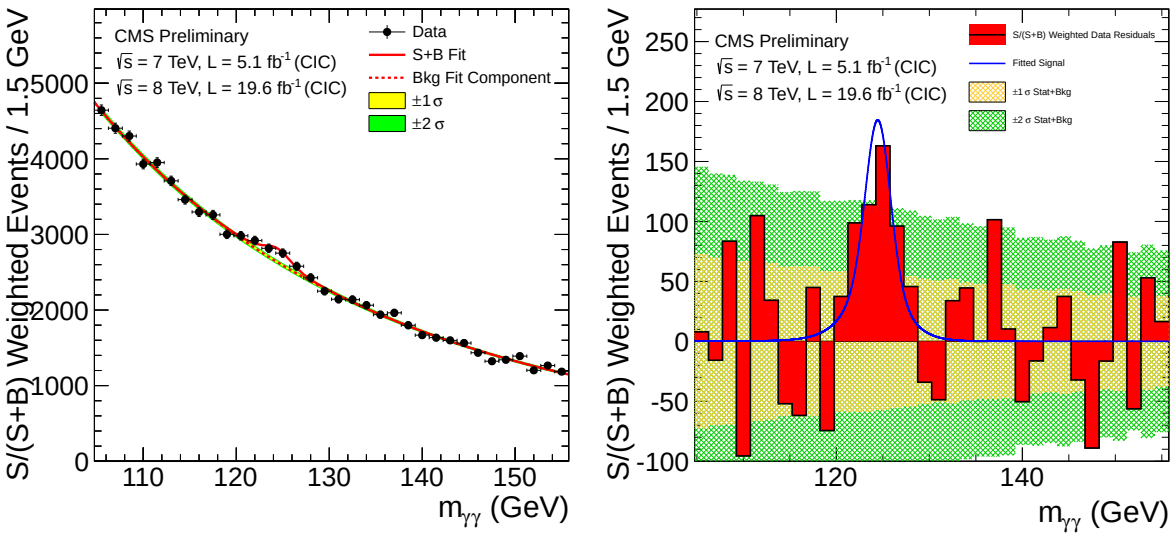


Figure 11: (left) The diphoton mass spectrum weighted by the ratio of signal-to-background in each event class for the cut-based analysis. (right) The background-subtracted weighted mass spectrum.

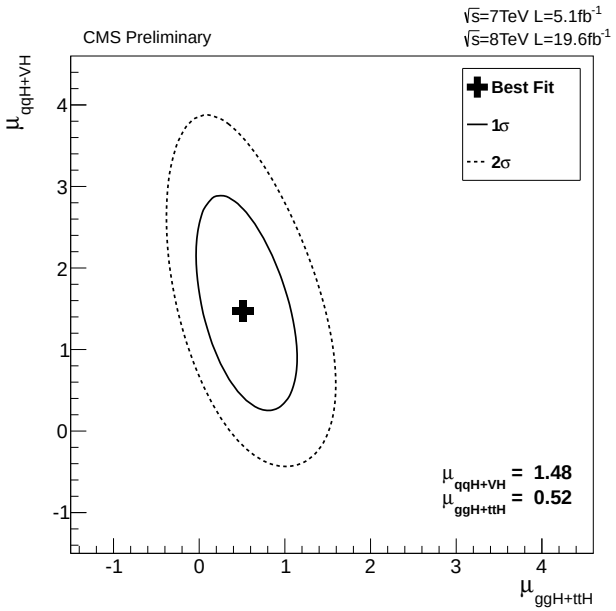


Figure 12: The 68% CL intervals for the signal strength modifiers associated with the gluon-gluon-fusion-plus-$t\bar{t}$H and for VBF-plus-VH production mechanisms, respectively. The cross indicates the best-fit value.

11.2 Production properties

The four main Higgs boson production mechanisms can be associated with either top-quark couplings (gluon fusion and $t\bar{t}$H) or vector boson couplings (VBF and VH). Figure 12 shows the 68% and 95% CL contours (computed as the variations around the likelihood maximum) for the signal strength modifiers associated with the gluon-fusion-plus-$t\bar{t}$H ($\mu_{ggH+t\bar{t}H}$) and for VBF-plus-VH production mechanisms ($\mu_{q\bar{q}H+VH}$). The best fit values are found to be $(\mu_{ggH+t\bar{t}H}, \mu_{q\bar{q}H+VH}) = (0.52, 1.48)$ for the mass-fit-MVA analysis.

11.3 Compatibility studies

The statistical compatibility of the mass-fit-MVA and the cut-based analyses on the whole of the 7 and 8 TeV has been tested using a non parametric technique, commonly known as "jackknife delete-d resampling" [67, 68].

The method can be applied to the estimation of the variance of statistical estimators and it proceeds as follows: given an estimator θ and a sample of n random variables $s = x_1..x_n$, several pseudo-samples $s^{(j)}$ are extracted. Each pseudo-sample contains all the events in s except for a subset of size d. Each pseudo sample is generated such that each event is removed once and only once. A total of $g = n/d$ can be extracted in this way. The estimator θ is then evaluated on each pseudo sample and the corresponding values are labelled as $\theta_{(j)}$. The jackknife estimator for the variance of θ is computed as:

$$var_J(\theta) = \frac{g-1}{g} \sum_{j=1}^{g} \left(\bar{\theta} - \theta_{(j)}\right)^2 \quad \text{where} \quad \bar{\theta} = \frac{1}{g} \sum_{j=1}^{g} \theta_{(j)}$$

The jackknife technique can be used to compute the statistical correlation between any pair of estimators. In particular, it was applied to strength modifier from the mass-fit-MVA analysis and the cut-based analysis. In general: two analyses "A" and "B", performed on the same

dataset will select different, but possibly overlapping set of events s_A and s_B depending on their specific selection criteria. The problem of determining the correlation between the signal strength modifiers μ_A and μ_B determined by the two analyses can be reduced to the computation of three variances:

- $var_J(\mu_I)$ (I=A,B), obtained applying the jackknife resampling to the sample s_I ;
- $var_J(\delta\mu := \mu_A - \mu_B)$, obtained applying the jackknife resampling to the sample $s_{AB} := s_A \cup s_B$.

When the union of two datasets is used to generate pseudo-samples, μ_I (I=A,B) is computed using the events in the set $s_I \cap s_{AB}^{(j)}$.

The technique described above was applied in particular to the 8TeV dataset. Here, out of all the events selected by the mass-fit-MVA analysis and the cut-based one, roughly 50% of the events are selected by both analyses, 32% are selected only by the cut-based analysis and 18% only by the mass-fit-MVA one. The corresponding numbers on the MC signal events are 81% selected by both analyses, 11% only by the cut-based analysis and 8% only by the mass-fit-MVA analysis.

On the 8 TeV dataset, the jackknife estimator for $\sigma_J(\delta\mu) = \sqrt{var_J(\delta\mu)}$ for $m_H = 125\,\text{GeV}$ was evaluated to be 0.23 ± 0.02, where the uncertainty was estimated varying the deleted sample size d. The statistical correlation between the signal strength estimators for the two analyses can in turn be evaluated to be roughly 75%.

Given this, the significance of the difference between the best fit signal strength obtained by the mass-fit-MVA analysis and the cut-based one can be estimated to be 1.8σ, considering the 8TeV dataset alone. The results of the two analyses obtained on the full dataset are estimated to be compatible within 1.5σ, assuming the same statistical correlation on the 7 and 8TeV datasets.

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 19.6 fb^{-1} at $\sqrt{s} = 8\,\text{TeV}$. The selected events are subdivided into classes, and the results of the search in each class are combined.

The mass range between 110 and 149 GeV is excluded as 95% confidence level, except the region between 122.8 and 127.8 where an excess of events is found.

For the mass-fit-MVA analysis, the local significance of the excess is 3.2 σ with a corresponding expected value of 4.2 σ and the best fit signal strength is $0.78^{+0.28}_{-0.26}$ times the Standard Model Higgs boson cross section. For the cut-based analysis, the local significance of the excess is 3.9 σ with a corresponding expected value of 3.5 σ and a best fit signal strength of $1.1^{+0.32}_{-0.30}$ times the Standard Model Higgs boson cross section. The two results on the complete 2011 and 2012 datasets are found to be compatible at the 1.5σ level after taking into account the correlations between the two analysis.

The best fit value for the signal strength modifiers associated with the gluon-fusion-plus-$t\bar{t}$H and for VBF-plus-VH production mechanisms are found to be $(\mu_{ggH+t\bar{t}H}, \mu_{q\bar{q}H+VH}) = (0.52, 1.48)$ for the mass-fit-MVA analysis.

The mass of the observed Higgs boson is measured to be $125.4 \pm 0.5(stat.) \pm 0.6(syst.)$ GeV.

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] ATLAS Collaboration Collaboration, "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC", *Phys.Lett.* **B716** (2012) 1–29.

[11] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30–61.

[12] S. Actis 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.

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

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

[15] 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.

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

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

[18] 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.

[19] 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.

[20] 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.

[21] S. Actis 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.

[22] 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.

[23] 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.

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

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

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

[31] CMS Collaboration, "Evidence for a new state decaying into two photons in the search for the standard model Higgs boson in pp collisions", *CDS Record* **1460419** (2012).

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

[33] CMS Collaboration, "Electromagnetic calorimeter calibration with 7 TeV data", *CDS Record* **1279350** (2010).

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

[35] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", *CDS Record* **1247373** (2010).

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

[37] 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.

[38] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", *CDS Record* **1194487** (2010).

[39] CMS Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", *CDS Record* **1279341** (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, G. P. Salam, and G. Soyez, "The Catchment Area of Jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005.

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

[44] CMS Collaboration, "Higgs to gamma gamma, Fermiophobic", *CDS Record* **1461937** (2012).

[45] CMS Collaboration, "Evidence for a new state in the search for the standard model Higgs boson in the $H \to ZZ \to 4l$ channel in pp collisions at $\sqrt{s}$ = 7 and 8 TeV", *CDS Record* **1460664** (2012).

[46] CMS Collaboration, "Studies of Tracker Material", *CDS Record* **1279138** (2010).

[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] S. Alioli 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.

[50] 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.

[51] 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.

[52] G. Bozzi 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.

[53] G. Bozzi 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.

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

[55] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 2. Differential Distributions", technical report, (2012).

[56] 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.

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

[58] CMS Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2012 Update", *CDS Record* **1482193** (2012).

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

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

[61] A. Buckley 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.

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

[63] M. Guzzi et al., "CT10 parton distributions and other developments in the global QCD analysis", arXiv:1101.0561.

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

[65] 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.

[66] 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).

[67] M. H. Quenouille, "Approximate tests of correlation in time-series", *J. Roy. Statist. Soc. Ser. B.* **11** (1949) 68–84.

[68] J. Tukey, "Bias and confidence in not quite large samples", *Annals of Mathematical Statistics* **29** (1958), no. 614, 1261–1295.