Proceedings of the PHOTON-2017 Conference, Geneva, Switzerland, 22 – 26 May 2017, edited by David d'Enterria, Albert de Roeck and Michelangelo Mangano, CERN Proceedings, Vol. 1/2018, CERN-Proceedings-2018-001 (CERN, Geneva, 2018)

Latest results on Higgs boson $\rightarrow \gamma\gamma$ in the CMS experiment

Hamed BAKHSHIANSOHI[*]
Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium

Abstract
The photonic decay of the Higgs boson is one of the most prominent decay modes for the observation and measurement of the Higgs boson properties, although its branching fraction is as low as 0.002. The CMS collaboration has analyzed 35.9fb^{-1} of proton-proton collision data delivered by the LHC at a center of mass energy of 13 TeV. The inclusive results of the search in different Higgs boson production channels and the differential fiducial cross section measurement results are presented in this talk.

Keywords
CMS Experiment, Higgs Boson, Photon, Inclusive, Differential

1 Introduction

The Higgs boson is the last discovered elementary particle predicted by the standard model of particles (SM). Since it was discovered by the CMS [1] and ATLAS experiments in 2012 [2, 3], many studies have been performed to understand its interaction with other elementary particles. So far all the measurements have been consistent with the standard model predictions.

As the photon is a massless particle, it does not couple to Higgs boson directly. The Higgs boson decays to two photons via a loop involving top-quarks or W-bosons and hence the branching fraction of the $H \rightarrow \gamma\gamma$ is as low as 0.002. Having a clean final state with an invariant mass peak that can be reconstructed with very high precision makes this rare decay of the Higgs boson very important.

The CMS experiment recorded 35.9 fb^{-1} of proton-proton collision data, to be used for physics analysis, delivered by the LHC at a center of mass energy of 13 TeV in 2016. This data has been analyzed to measure the cross section for different Higgs boson production mechanisms, where the Higgs boson decays to two photons. Results have been documented in Ref.[4]. A dedicated analysis was also performed to measure the differential and fiducial production cross section of the Higgs boson in the photonic decay mode and the results have been documented in Ref.[5]. Both of these analyses are reviewed in this report.

2 Event reconstruction

Within the superconducting solenoid which is the central feature of the CMS experiment, silicon pixel and strip trackers, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections, are located. The trajectory of charged particles are measured by the silicon pixel and strip tracker. The ECAL extends up to $|\eta| < 1.48$ in the barrel, while the endcaps cover the region $1.48 < |\eta| < 3.0$. ECAL crystal arrays projecting radially outwards from the nominal interaction point, with a slight offset.

The global event reconstruction algorithm, known as particle-flow event reconstruction [6], reconstructs and identifies each individual particle using information from the various elements of the CMS detector. The energies of photons is directly obtained from the ECAL measurement. While the energy

[*]on behalf of the CMS collaboration

https://doi.org/10.23727/CERN-Proceedings-2018-001.235

of electrons, muons, and hadrons are determined from a combination of the tracker and calorimeter information. Hadronic jets are clustered from these reconstructed particles using the anti-k_T algorithm [7], with a size parameter of 0.4. Jets including B-hadrons are tagged using the information of the displaced decay vertex, using the combined secondary vertex (CSV algorithm) [8]. The missing transverse momentum vector (MET) is taken as negative vector sum of all reconstructed particle candidate transverse momenta.

2.1 Photon reconstruction

Photons are identified as ECAL energy clusters not linked to the extrapolation of any charged-particle trajectory to the ECAL. The clustering algorithm has been optimized to recollect the total energy of the photon, including conversions in the material upstream of the calorimeter. Clustering starts based on local energy peaks above a given threshold, as the "seeds". Then they are grown by aggregating crystals with at least one side in common with a clustered crystal and with an energy in excess of a given threshold. Finally, clusters are dynamically merged into "superclusters", to allow good energy containment, accounting for geometrical variations of the detector along η, and optimizing robustness against pileup.

The energy of photons is computed from the sum of the energy of the ECAL reconstructed hits, calibrated and corrected for several detector effects [9]. The energy is corrected to contain the electromagnetic showers in the clustered crystals and the energy losses of converted photons. A multivariate regression technique is used to compute the correction. It allows estimating simultaneously the energy of the photon and its median uncertainty.

The photon candidates used in this analysis are required to satisfy preselection criteria similar to, but slightly more stringent than, the trigger requirements. To reject ECAL energy deposits incompatible with a single isolated electromagnetic shower, such as those coming from neutral mesons, a selection on shower shape variables is applied. To reject hadrons, the ratio of energy in HCAL cells behind the supercluster to the energy in the supercluster is checked. If the supercluster matches to an electron track with no missing hits in the innermost tracker layers, the photon candidate is vetoed. To discard photons within jets, the photon is requested to be isolated in tracker and calorimeters.

The efficiency of all preselection criteria, except the electron veto requirement, is measured with a tag and probe technique using $Z \rightarrow$ee events. The efficiency for photons to satisfy the electron veto requirement is measured with $Z \rightarrow \mu\mu\gamma$ events, in which the photon is produced by final-state radiation and provides a sample of prompt photons with purity higher than 99%. Simulation is corrected to data to consider the differences in efficiencies of simulation and data.

A Boosted Decision Tree (BDT) is trained to separate prompt photons from photon candidates satisfying the preselection requirements. The photon identification BDT is trained using simulated γ+jets events where prompt photons are considered as signal and non-prompt photons as background. Shower shape and isolation variables are used as the input variables. Fig. 1 (Left) shows the output of the identification BDT score.

2.2 Diphoton reconstruction

Determination of the vertex from which the diphoton originated has a large impact on the diphoton mass resolution. A BDT is trained using observables related to tracks recoiling against the diphoton system. In case of conversion in the tracker material, the track information is also helpful in vertex assignment. The probability that the assigned vertex is within 1cm of the diphoton interaction point is then estimated using a second BDT.

Events with two preselected photon candidates with $p_T^{\gamma 1} > \mathrm{m}_{\gamma\gamma}/3$ and $p_T^{\gamma 2} > \mathrm{m}_{\gamma\gamma}/4$, in the mass range $100 < \mathrm{m}_{\gamma\gamma} < 180$ GeV are selected.

A dedicated diphoton BDT is trained to evaluate the diphoton mass resolution per event. Higher

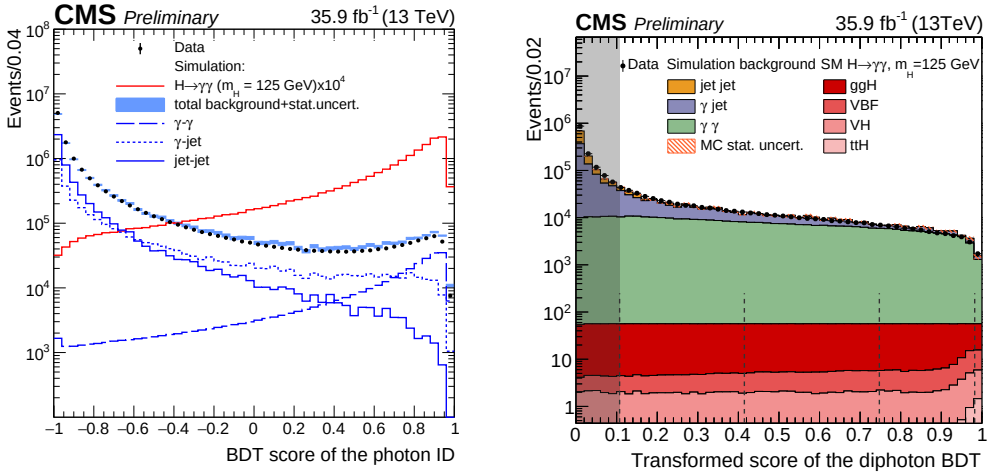

Fig. 1: (Left) Photon identification BDT score in the 13 TeV data set (points), and for simulated background events (blue histogram). The red histogram corresponds to simulated Higgs boson signal events. (Right) Transformed score of the diphoton multivariate classifier for events with two photons satisfying the preselection requirements in data (points), simulated signal (red shades), and simulated background (coloured histograms). Both signal and background are stacked together. The grey shade indicates events discarded from the analysis (see Ref.[4]).

values of the diphoton BDT show that the event has two photons with signal-like kinematics, good mass resolution, and high photon identification BDT score. The input variables to the classifier are:

- $p_T^{\gamma}/\mathrm{m}_{\gamma\gamma}$ for each photon;
- the pseudorapidity of the two photons;
- the cosine of the angle between the two photons in the transverse plane;
- photon ID BDT scores for both photons;
- two per-event relative mass resolution estimates, one under the hypothesis that the mass has been reconstructed using the correct primary vertex, and the other under the hypothesis that the mass has been reconstructed using an incorrect vertex;
- the per-event probability estimate that the correct primary vertex has been assigned to the diphoton.

The distribution of the diphoton BDT score is shown in Fig. 1(Right).

3 Categorization

To increase the sensitivity of the analysis, events are classified according to the production mechanisms and their mass resolution and predicted signal-to-background ratio. Higgs boson production mechanisms considered in this analysis are gluon fusion (ggH), vector boson fusion (VBF), production associated with a vector boson (VH) or with a top quark pair ($\mathrm{t\bar{t}H}$).

In total 14 exclusive categories are defined. These categories and their rankings are shown in Fig 2. A dedicated BDT is trained to separate $\mathrm{t\bar{t}H}$ events, based on the jet information. The cut on the BDT output is optimized together with the cut on the diphoton BDT score to maximize the expected sensitivity to this production mechanism. Events with one lepton, either electron or μ and at least 2 jets, one of which is tagged as b-jet are tagged under the $\mathrm{t\bar{t}H}$ Leptonic category. Events with no lepton and at least 3 jets that have at least one b-tagged jet are tagged as hadronic decays of $\mathrm{t\bar{t}H}$.

There are 5 VH categories, where V stands for either Z-boson or W-boson. Events with two leptons with invariant mass consistent with the Z-boson mass are tagged as ZH Leptonic event. The next two VH categories are dedicated to events with 1 lepton, among which events with MET>45 GeV and

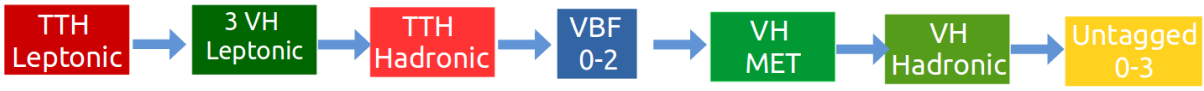

Fig. 2: Schematic view of event categorization and the order of exclusive categories.

less than 3 jets are good candidates for leptonic decays of WH events. The VH-Leptonic loose category contains events with MET<45 GeV, to include ZH events with one missing lepton. The VH(MET) category is defined to tag ZH events, where Z decays to neutrinos, or WH events when the lepton of the W-boson decay is missing. Events with MET>85 GeV, where $\Delta\phi(\mathrm{MET}, \gamma\gamma)$>2.4 are tagged in VH(MET) category. To include the hadronic decays of vector bosons, events with more than 2 jets when the invariant mass of one of the pair of jets is between 60 GeV and 120 GeV are tagged as VH(Hadronic) category.

Having two forward jets is the distinct feature of the VBF production. Events with two forward jets where the invariant mass of them is greater than 250 GeV are selected. Forward information is also used to train a BDT. Events are categorized in 3 categories based on the BDT output to obtain the best expected significance.

All the remaining untagged events are divided into 4 categories according to their mass resolution.

4 Results

In each category, Higgs boson events appear as a peak on a falling background of non-Higgs events. The shape of the falling background is modeled fitting on data, while the signal shape is obtained by fitting on simulation.

4.1 Background Model

The discrete profiling method [10] is used to describe the background. Exponentials, Bernstein polynomials, Laurent series and power law function families are considered. The F-test [11] technique is used to determine the maximum order for each family of functions to be used.

4.2 Signal Shape Modeling

Given that the distribution of $m_{\gamma\gamma}$ depends significantly on the correct vertex assignment of the candidate diphoton, distributions with the correct vertex and wrong vertex are fitted separately when constructing the signal model. For each process, category, and vertex scenario, the $m_{\gamma\gamma}$ distributions are fitted using a sum of at most five Gaussian functions.

Parameter values of the signal shape for each process, category, and vertex scenario are found as a function of the Higgs boson mass in the range from 120 to 130 GeV.

The efficiency times acceptance of the signal model as a function of m_H for all categories combined is shown in Fig. 3.

4.3 Yields and inclusive results

The expected signal yield and the composition of different production mechanisms in each category and also the ratio of signal events is shown in Fig. 4.

A fit is performed to obtain the signal strength for each production in all categories. The results are shown in Fig. 5. The expected sensitivity for different production modes are compatible with the observation.

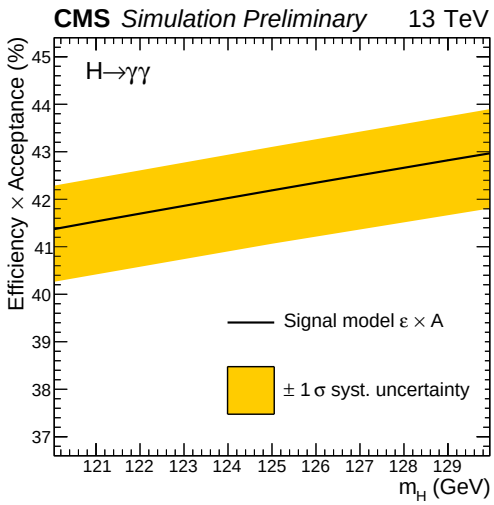

Fig. 3: The efficiency × acceptance of the signal model as a function of m_H for all categories combined. The black line represents the yield from the signal model. The yellow band indicates the effect of the systematic uncertainties for trigger, photon identification and selection, photon energy scale and modelings of the photon energy resolution, and vertex identification (see Ref.[4]).

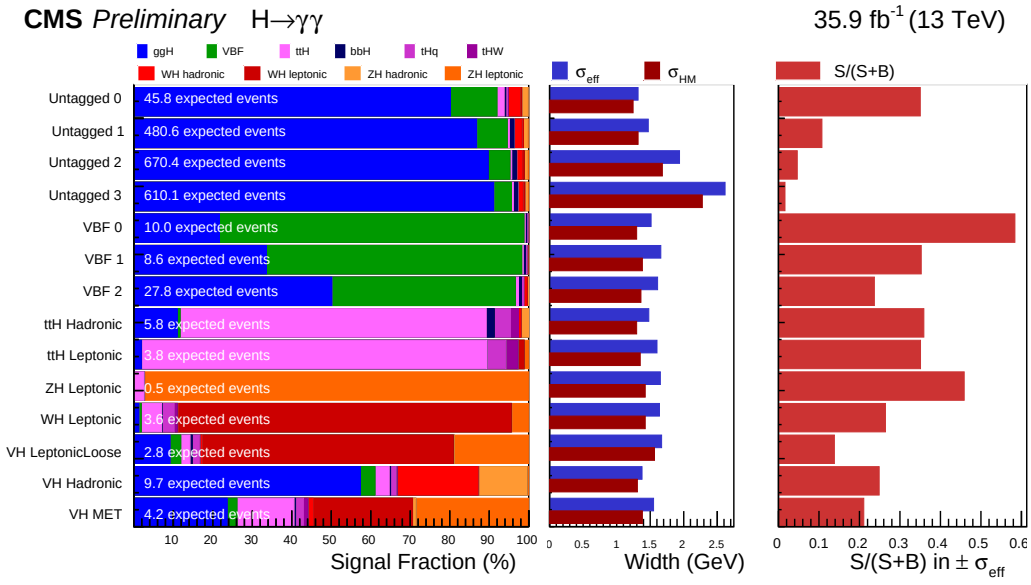

Fig. 4: Expected composition of signal events per production mechanism in different categories. The ratio of the number of signal events. (S) to the number of signal plus background events (S+B) is plotted on the right hand side (see Ref.[4]).

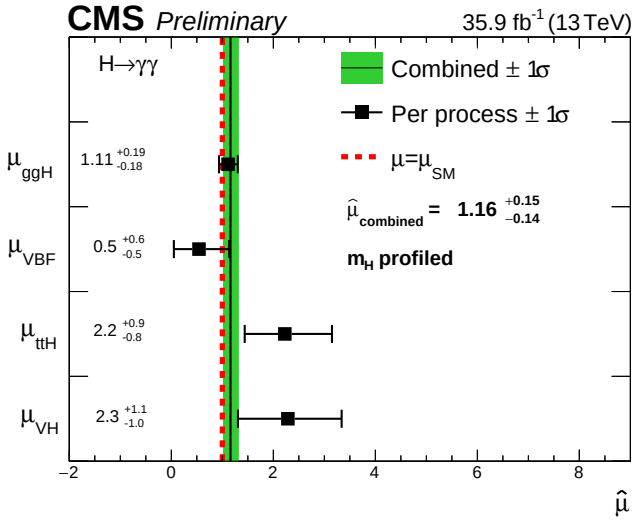

Fig. 5: Signal strength modifiers measured for each process (black points), with m_H profiled, compared to the overall signal strength (green band) and to the SM expectation (dashed red line), (see Ref.[4]).

4.4 Differential and Fiducial results

A dedicated analysis is performed to measure the cross section of the Higgs boson production vs. different properties of Higgs boson and event. Diphoton events are categorized based on an uncorrelated mass resolution estimator defined as

$$y(\sigma_m/m|m) = \int_0^{\sigma_m/m} f(\sigma_m/m'|m)d\sigma_m/m'$$

where the conditional distribution of σ_m/m, $f(\sigma_m/m|m)$, is constructed by sorting the values of σ_m/m in bins of m. The number of categories was optimized and found to be 3 in order to saturate the maximum achievable sensitivity.

Total and differential cross sections are unfolded to particle-level using a likelihood fit. Total cross section in fiducial phase space with $p_T^{\gamma 1} = m_{\gamma\gamma}/3$, $p_T^{\gamma 2} = m_{\gamma\gamma}/4$ and $|\eta^\gamma|$<2.5 is measured and found to be 84^{+13}_{-12} fb which is consistent with the theory prediction.

The differential cross sections as a function of the diphoton transverse momentum and the jet multiplicity are reported in Figure 6, together with the corresponding theoretical predictions. Two sets of predictions are shown in each plot. For the first, shown in orange, MADGRAPH_aMC@NLO was used to simulate all the Higgs boson production processes. The second, shown in green, was obtained using POWHEG-generated ggH events, while taking other production mechanisms from MADGRAPH_aMC@NLO. The plots show in blue the sum of the contributions from VBF, VH and ttH (labeled HX).

References

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

[2] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716:30, 2012. arXiv:1207.7235, doi:10.1016/j.physletb.2012.08.021.

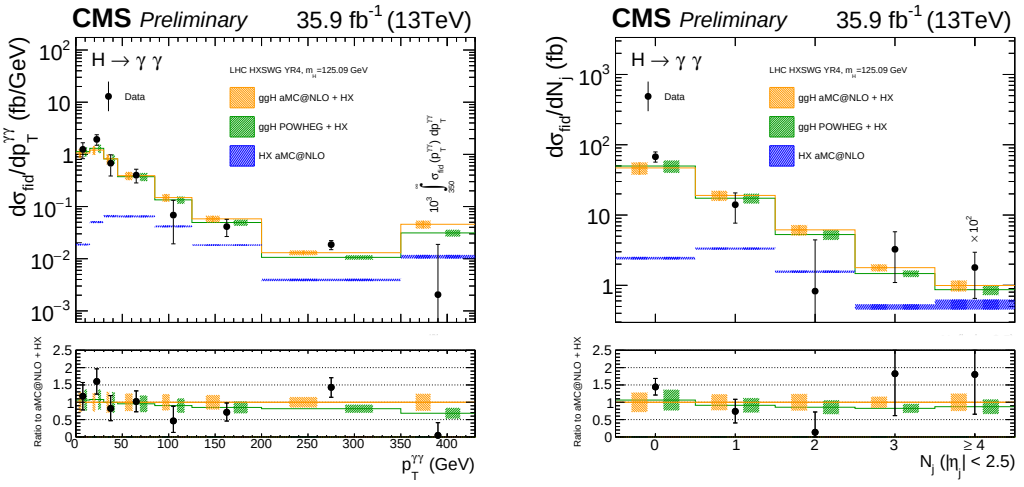

Fig. 6: Measured $H \to \gamma\gamma$ differential cross-section (black points) for (Left) $p_{T,\gamma\gamma}$, (Right) N_{jets}. The measurements are compared to the theoretical predictions, combining the Higgs boson cross sections and branching fraction as in the LHC Higgs Cross Section Working Group [12] with two different generators for the gluon-gluon fusion process: MADGRAPH_aMC@NLO (in orange) and POWHEG (in green) (see Ref.[5]).

[3] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1, 2012. arXiv:1207.7214, doi:10.1016/j.physletb.2012.08.020.

[4] CMS Collaboration. Measurements of properties of the Higgs boson in the diphoton decay channel with the full 2016 data set. Technical Report CMS-PAS-HIG-16-040, CERN, Geneva, 2017. URL: https://cds.cern.ch/record/2264515.

[5] CMS Collaboration. Measurement of differential fiducial cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-HIG-17-015, CERN, Geneva, 2017. URL: http://cds.cern.ch/record/2257530.

[6] CMS Collaboration. Particle-flow reconstruction and global event description with the CMS detector. Submitted to *JINST*, 2017. URL: https://cds.cern.ch/record/2270046, arXiv:1706.04965.

[7] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008. arXiv:0802.1189, doi:10.1088/1126-6708/2008/04/063.

[8] CMS Collaboration. Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013. arXiv:1211.4462, doi:10.1088/1748-0221/8/04/P04013.

[9] CMS Collaboration. Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7TeV$. *JINST*, 8:09009, 2013. arXiv:1306.2016, doi:doi:10.1088/1748-0221/8/09/P09009.

[10] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies. Handling uncertainties in background shapes: the discrete profiling method. *JINST*, 10(04):P04015, 2015. arXiv:1408.6865, doi:10.1088/1748-0221/10/04/P04015.

[11] R. A. Fisher. On the interpretation of χ^2 from contingency tables, and the calculation of p. *Journal of the Royal Statistical Society*, 85(1):87, 1922. doi:10.2307/2340521.

[12] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016. URL: https://arxiv.org/abs/1610.07922, arXiv:1610.07922.