PoS(ICHEP2016)385

Measurements of Higgs boson production and properties in the diphoton decay channel using the CMS detector

Vittorio Raoul Tavolaro[*][†]

Institute for Particle Physics, Eidgenössische Technische Hochschule Zürich, Switzerland.
E-mail: vittorio.tavolaro@cern.ch

The Higgs boson production is observed in the diphoton channel with 12.9 fb^{-1} of proton-proton collision data collected by the LHC in 2016 at centre-of-mass energy of 13 TeV. A maximum significance of 6.1 σ is measured at 126.0 GeV, with a best-fit signal strength relative to the standard model prediction of $0.95 \pm 0.20 = 0.95 \pm 0.17$ (stat) $^{+0.10}_{-0.07}$ (syst) $^{+0.08}_{-0.05}$ (theo). The fiducial cross section is measured to be $\hat{\sigma}_{fid} = 69^{+18}_{-22}$ fb $= 69^{+16}_{-22}$(stat)$^{+8}_{-6}$(syst) fb, where the standard model theoretical prediction is 73.8 ± 3.8 fb.

38th International Conference on High Energy Physics
3-10 August 2016
Chicago, USA

---

[*]Speaker.

[†]On behalf of the CMS collaboration.

1. Introduction

The discovery of a new particle during the Run 1 period of the LHC, broadly compatible with the Standard Model (SM) Higgs boson, has been followed by a comprehensive set of studies of its properties in different decay channels and production modes. Despite the small branching ratio predicted by the SM ($\approx 0.2\%$), the H $\rightarrow \gamma\gamma$ decay channel provides a clean final state with an invariant mass peak that can be reconstructed with great precision. As a consequence, the H $\rightarrow \gamma\gamma$ is one of the highest sensitivity channels for the discovery and study the Higgs boson. The sensitivity of the dataset analysed allows to study individual production mechanisms, through the definition of exclusive analysis categories, targeting gluon-gluon fusion (ggH), vector boson fusion (VBF) and associated production with a $t\bar{t}$ pair (ttH). The analysis [1] described in this paper presents the standalone observation in the diphoton decay channel of a particle consistent with being the Higgs boson in the data collected in the LHC Run 2 during 2016 at 13 TeV, corresponding to an integrated luminosity of 12.9 fb^{-1}. Per-production mechanism signal strength modifiers relative to the SM prediction and the fiducial production cross-section are also measured.

2. Di-photon mass reconstruction

The signal extraction is based on the identification of a narrow fully reconstructed peak over a large smoothly falling background in the diphoton invariant mass distribution. This is obtained through a fit of the diphoton invariant mass in several event categories, to improve the sensitivity of the analysis and to allow the study of specific production mechanisms. Since the expected width of the Higgs boson is negligible compared to the detector resolution, the signal to background ratio scales with the inverse of the mass resolution. It is therefore of primary importance to optimise the detector resolution.

The response of the CMS electromagnetic calorimeter, described in [2], is corrected for its change in time during the operation and calibrated at the crystal level to obtain a uniform response across the azimuthal and longitudinal angular coordinates. Multivariate energy corrections are derived to account for the electromagnetic shower global and local containment, providing a per-event and per-photon correction simultaneously with an estimate of the energy resolution. Residual energy scale and resolution corrections are obtained from a detailed data to simulation comparison on a pure sample of Z $\rightarrow$ ee events, with same calibrations and multivariate corrections applied.

The electromagnetic calorimeter does not provide information on the direction of flight of the incident photons, so it is necessary to employ the information of the primary vertices in order to reconstruct the angle between the two photons and thus their invariant mass. The absence of ionisation tracks in the tracker system for unconverted photons and the high pile-up conditions of the collisions pose a significant challenge in the identification of the primary vertex in diphoton events. The primary vertex assignment is achieved through a multivariate classifier (MVA) which exploits information from the kinematic correlations and track momentum imbalance created by the presence of the photon pair, combined with the direction of the tracks resulting from photon conversion to an electron-positron pair, when present. For the purpose of this analysis, a vertex assignment is considered correct when its longitudinal displacement with respect to the true primary vertex of the event has a negligible impact on the diphoton invariant mass resolution, compared to the impact

on the resolution of the calorimetric measurement of the photon energies. This condition translates into a requirement on the longitudinal displacement to be smaller than 1 cm. The probability of a successful per-event primary vertex multivariate assignment is estimated with a second MVA. The vertex identification and probability algorithms are validated using $Z \rightarrow \mu\mu$ events from data and simulation, where the muon tracks are used to tag the true vertex and then removed from the event to test the performance of the algorithm. The average vertex identification efficiency is about 80%.

3. Photon selection

The events are selected using a double-photon trigger selection based on the transverse energy of the photons, their invariant mass and a combination of isolation variables and variables describing the shape of the electromagnetic shower in the calorimeter.

The main backgrounds are resulting from irreducible diphoton production, and reducible photon+jets events, where one jet is mis-reconstructed as a photon. Events from multi-jets QCD processes, where two jets are reconstructed as photons, bring smaller contribution to the backgrounds. The photon selection aims at selecting prompt, isolated photons while rejecting jets, typically with a large fraction of electromagnetic energy, mimicking a photon in the detector. In order to do so, after a loose preselection slightly tighter than the trigger selection, a multivariate photon identification is applied, exploiting information from shower-shape, kinematic and isolation variables, and the median energy density of the event.

A diphoton MVA is subsequently built, with the aim of selecting signal-like, good mass resolution diphoton pairs and rejecting diphoton pairs resulting from background events. The information exploited in order to build this MVA are a per-event mass resolution estimate, kinematic variables of the two photons and of the diphoton system, the photon identification MVA scores and the probability that the correct vertex was selected.

4. Event categories

The selected events are assigned exclusively to eight different categories, defined and ordered as follows:

- categories enriched in events produced via the ttH mechanism. Two categories are defined, targeting respectively events where the *W* bosons resulting from the *t* decays are both decaying hadronically (labeled *ttH Hadronic Tag*) and events where at least one *W* is decaying leptonically (labeled *ttH Leptonic Tag*). The selection is performed using the photon identification and diphoton MVAs, the study of additional leptons, jets and missing transverse energy in the event;

- categories enriched in events produced via the VBF mechanism. The selection is performed exploiting a MVA trained to identify events with two jets having the characteristic kinematic configuration of the VBF process. The output of this MVA is combined with the diphoton MVA into a VBF MVA, selecting events with two signal-like photons and two VBF-like jets. Two VBF-enriched categories are defined on the output of the latter MVA, labeled *VBF Tag 0* and *1*, in descending order of VBF MVA score;

- 4 categories are defined on the output of the diphoton MVA, populated mostly by ggH events, labeled *Untagged Tag 0,1,2,3*, in descending order of diphoton MVA score. The number and definition of categories is optimised maximising the expected sensitivity of the analysis.

The output of the diphoton MVA and the final categorisation are shown in figure 1, together with the yields of simulated signal events in the analysis categories.
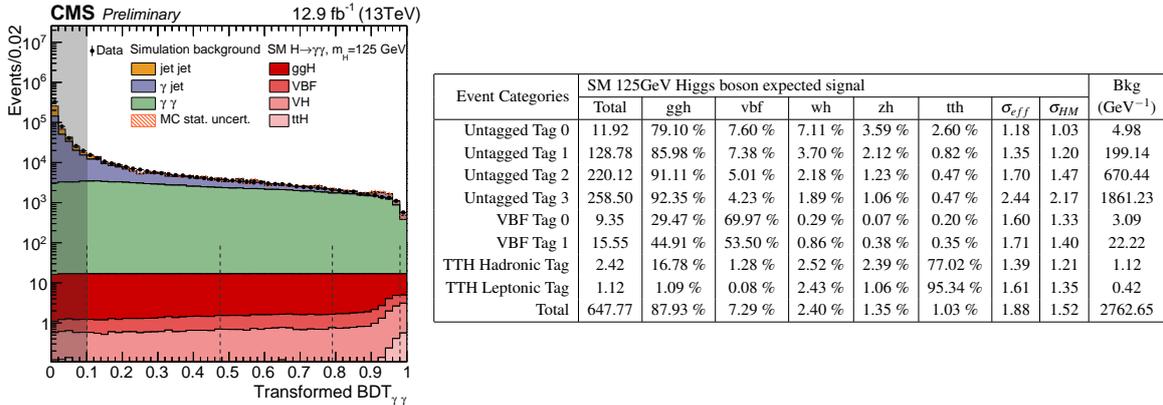

| Event Categories | SM 125GeV Higgs boson expected signal | | | | | | | | Bkg |
	Total	ggh	vbf	wh	zh	tth	σ_{eff}	σ_{HM}	(GeV^{-1})
Untagged Tag 0	11.92	79.10 %	7.60 %	7.11 %	3.59 %	2.60 %	1.18	1.03	4.98
Untagged Tag 1	128.78	85.98 %	7.38 %	3.70 %	2.12 %	0.82 %	1.35	1.20	199.14
Untagged Tag 2	220.12	91.11 %	5.01 %	2.18 %	1.23 %	0.47 %	1.70	1.47	670.44
Untagged Tag 3	258.50	92.35 %	4.23 %	1.89 %	1.06 %	0.47 %	2.44	2.17	1861.23
VBF Tag 0	9.35	29.47 %	69.97 %	0.29 %	0.07 %	0.20 %	1.60	1.33	3.09
VBF Tag 1	15.55	44.91 %	53.50 %	0.86 %	0.38 %	0.35 %	1.71	1.40	22.22
TTH Hadronic Tag	2.42	16.78 %	1.28 %	2.52 %	2.39 %	77.02 %	1.39	1.21	1.12
TTH Leptonic Tag	1.12	1.09 %	0.08 %	2.43 %	1.06 %	95.34 %	1.61	1.35	0.42
Total	647.77	87.93 %	7.29 %	2.40 %	1.35 %	1.03 %	1.88	1.52	2762.65

Figure 1: Left: distribution of the diphoton MVA output for simulated events (stacked histograms) and data (black dots). The variable is built to have a flat distribution on signal events (red shades histograms). The vertical dashed lines show the boundaries for the definition of the four Untagged categories. Right: signal event yields for the analysis categories, total and split by production mechanism. The width of the signal distribution is given as σ_{eff}, defined as the interval containing the 68% of the signal in the category and as σ_{FWHM}, defined as the interval corresponding to the full width at half maximum of the signal distribution.

5. Signal and background models

In order to perform the signal extraction in the different categories, parametric models are built both for signal and background components.

The signal model is extracted from simulation, using a sum of gaussian functions to describe the different components contributing to the diphoton invariant mass. The model is constructed after applying all the data-to-simulation correction factors and efficiencies scale factors. The model is fit to several Higgs mass hypotheses in a mass range between 120 and 130 GeV.

The extraction of the background model is completely data driven. The choice of the functional form for the description of the background diphoton mass distribution in each category is treated as a discrete nuisance parameter floating over an initial set of possible functional forms. This technique is described in detail in [3].

6. Results

The signal and background models are combined to fit the diphoton mass distributions in all the analysis categories. Figure 2 shows the sum of events in all categories, weighted for the signal to signal-plus-background yields ratio of each category, together with the signal-plus-background

fit. The maximum observed significance is 6.1 standard deviations, at a best-fit mass of 126.0 GeV, yielding a signal strength modifier relative to the standard model of $0.95 \pm 0.20 = 0.95 \pm 0.17$ (stat) $^{+0.10}_{-0.07}$ (syst) $^{+0.08}_{-0.05}$ (theo).
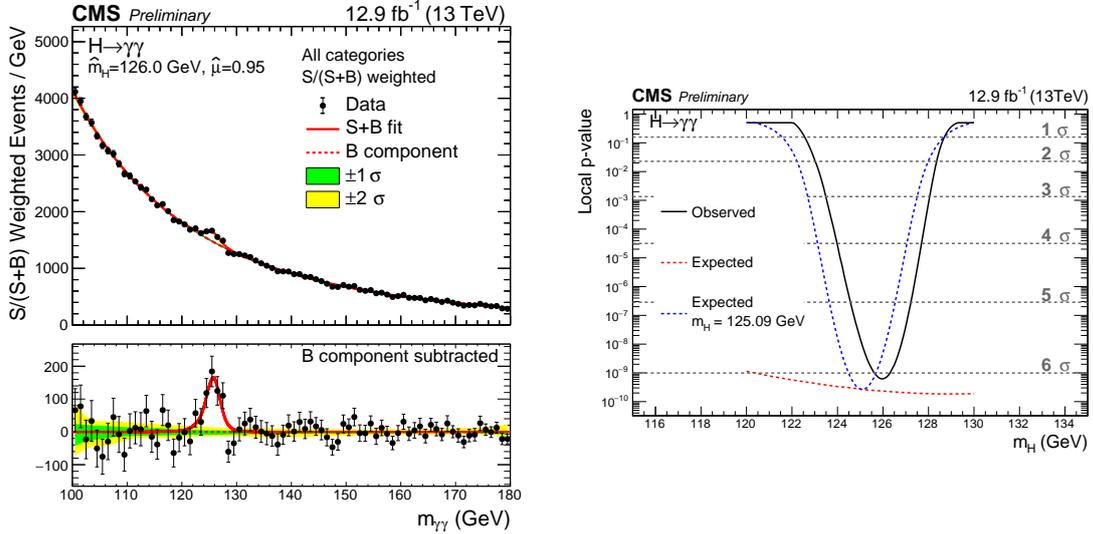

Figure 2: Left: diphoton invariant mass distribution in data, weighted for their signal to signal-plus-background ratio. The solid red line shows the signal-plus-background fit, while the dashed red line indicates the background component of the fit. Right: local p-value observed in data (solid black line), expected local p-value for a Higgs boson with a mass of 125.09 GeV (blue dashed line) and minimum expected local p-value as function of the mass hypothesis of the Higgs boson (red dashed line).

Signal strength modifiers are also extracted for ggH, VBF and ttH production mechanisms, with the SM prediction within uncertainties, as summarised in figure 3. The bosonic ($\mu_{\mathrm{VBF,VH}}$) and fermionic ($\mu_{\mathrm{ggH,ttH}}$) components of the signal strength modifiers are also compatible with the SM prediction and are shown in figure 3.

7. Fiducial cross section measurement

A different categorisation of the events is employed to perform the measurement of the fiducial cross section for the production of the Higgs boson in the diphoton decay channel. The categorisation relies on a diphoton mass resolution estimator, built removing the correlation present between mass and mass resolution as a consequence of the calorimetric correlation between energy and energy resolution. Using the de-correlated estimator prevents shaping of the mass distribution after categorisation. Three categories in de-correlated mass resolution are defined. A fiducial phase space for the measurement is defined in terms of pseudo-rapidity, transverse momentum, generator level isolation and photon identification MVA score of the particles. The measurement of the fiducial cross section yields $\hat{\sigma}_{fid} = 69^{+18}_{-22}$ fb $= 69^{+16}_{-22}$(stat)$^{+8}_{-6}$(syst) fb, showing very good agreement with the standard model theoretical prediction [4] [5] of 73.8 ± 3.8 fb.
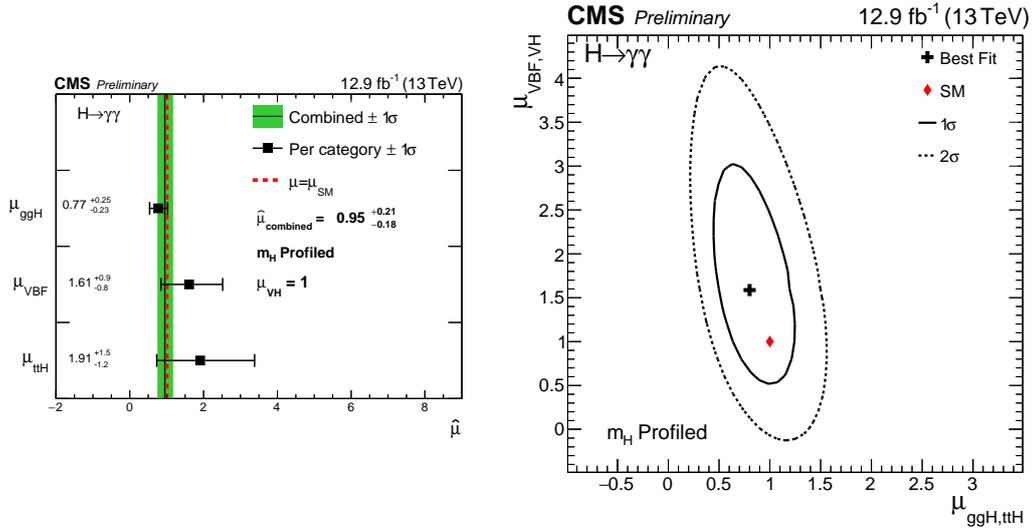
Figure 3: Left: fit of the signal strength modifiers relative to the SM prediction for gluon fusion (μ_{ggH}), vector boson fusion (μ_{VBF}) and in association with a $t\bar{t}$ pair (μ_{ttH}) production mechanisms. The signal strength of the associated production with a vector boson (VH) is set to the SM prediction. The measurements are shown by the black squares, and compared to the prediction indicated by the dashed vertical red line. Right: fit of the bosonic ($\mu_{VBF,VH}$) and fermionic ($\mu_{ggH,ttH}$) components of the signal strength modifiers. The black cross shows the best-fit value, while the solid (dashed) black contours indicate the 1 (2) standard deviations uncertainties. The SM prediction is shown by the red diamond.

8. Summary

This paper reports the observation of the Higgs boson decaying in the diphoton channel, using a dataset corresponding to 12.9 fb^{-1} collected in 2016 at 13 TeV. The maximum significance of the observation is 6.1σ at a mass of 126.0 GeV. The signal strength is measured also in its bosonic and fermionic components, as well per individual ggH, VBF and ttH production mechanisms, indicating good compatibility with the SM predictions. The fiducial cross section is measured and found to be in agreement with the theoretical predctions [4] [5].

References

[1] Updated measurements of Higgs boson production in the diphoton decay channel at $\sqrt{s} = 13$ TeV in pp collisions at CMS. Technical Report CMS-PAS-HIG-16-020, CERN, Geneva, 2016.

[2] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[3] 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.

[4] J R Andersen et al. Handbook of LHC Higgs Cross Sections: 4. Higgs Properties. Technical report.

[5] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.