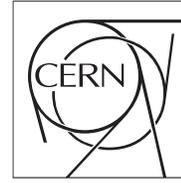


The Compact Muon Solenoid Experiment  
**Conference Report**

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



15 September 2016 (v4, 22 September 2016)

# Higgs boson production at the CMS experiment

Somnath Choudhury  
(for the CMS collaboration)

## Abstract

After the discovery of the Higgs boson, the ATLAS and CMS experiments have performed combined measurements of the mass of the Higgs boson and also measurements of the production and decay rates, as well as constraints on its couplings to vector bosons and fermions. These combined LHC measurements from the proton-proton collision Run-1 data will be summarized. In addition, the first results from the CMS experiment, on the way to the rediscovery of the Higgs boson in different production and decay modes with the early Run-2 data will also be presented in this paper.

Presented at *QCD16 19th High-Energy Physics International Conference in Quantum Chromodynamics*



## Higgs boson production at the CMS experiment \*

Somnath Choudhury  
(for the CMS collaboration)

*Indian Institute of Science, Bangalore, India*

---

### Abstract

After the discovery of the Higgs boson, the ATLAS and CMS experiments have performed combined measurements of the mass of the Higgs boson and also measurements of the production and decay rates, as well as constraints on its couplings to vector bosons and fermions. These combined LHC measurements from the proton-proton collision Run-1 data will be summarized. In addition, the first results from the CMS experiment, on the way to the rediscovery of the Higgs boson in different production and decay modes with the early Run-2 data will also be presented in this paper.

*Keywords:* Higgs boson, mass, coupling, signal strength, cross section

---

### 1. Introduction

In July 2012, the ATLAS and CMS experiments [1] reported the observation of a new particle at a mass of approximately 125 GeV with Higgs boson-like properties using proton-proton collision data from the LHC at CERN corresponding to integrated luminosities of around  $5 \text{ fb}^{-1}$  at 7 TeV and  $5 \text{ fb}^{-1}$  at 8 TeV centre-of-mass energies [2]. Subsequent results from both experiments established that the measurements of the properties of the new particle, including its spin, parity, and coupling strengths to other standard model (SM) particles, are consistent with those expected for the SM Higgs boson within uncertainties. The LHC experiments had to establish methods to cope with the challenge of a high number of simultaneous collisions per beam crossing (pileup), which occurs at high luminosity. The successful mitigation of pileup was demonstrated and an almost uniform response of the missing transverse energy resolution as a function of the number of primary vertices in the event has been achieved.

The SM Higgs boson analyses utilise the major Higgs boson production mechanisms in proton-proton collisions which are in order of decreasing cross sections namely gluon fusion, vector boson fusion (VBF), associated production with vector bosons (VH), and production with top quark pair. A measurement of the rate of Higgs boson production in association with top quark pair provides a direct test of the coupling between the top quark and the Higgs boson and has also been studied by the CMS experiment. The measurement of the properties of the discovered Higgs boson is crucial to understand the nature of the new particle and study its compatibility with the standard model. Any deviations from SM expectations would definitely indicate new physics at the TeV scale signalling scenarios beyond the standard model (BSM). Combined ATLAS and CMS measurements of the Higgs boson mass, production and decay rates, as well as constraints on its couplings to vector bosons and fermions have been performed using the LHC Run-1 data. With the start of LHC Run-2 in the year 2015, the CMS experiment is on its way towards rediscovery of the Higgs boson in 13 TeV proton collisions. The Higgs boson production in the gluon fusion process has a cross section enhanced 2.3 times in the

---

\*Talk given at 19th International Conference in Quantum Chromodynamics (QCD 16), 4 - 8 July 2016, Montpellier - France  
Email address: [somnath.choudhury@cern.ch](mailto:somnath.choudhury@cern.ch)

13 TeV collisions with respect to the 8 TeV collisions and this factor is 3.9 for the associated production with top quark pair. With the analysis of up to  $2.8 \text{ fb}^{-1}$  of data collected in the year 2015, the Run-1 sensitivity would be established with more data from the larger data set of LHC in the year 2016.

## 2. LHC combined measurements of Higgs boson from Run-1

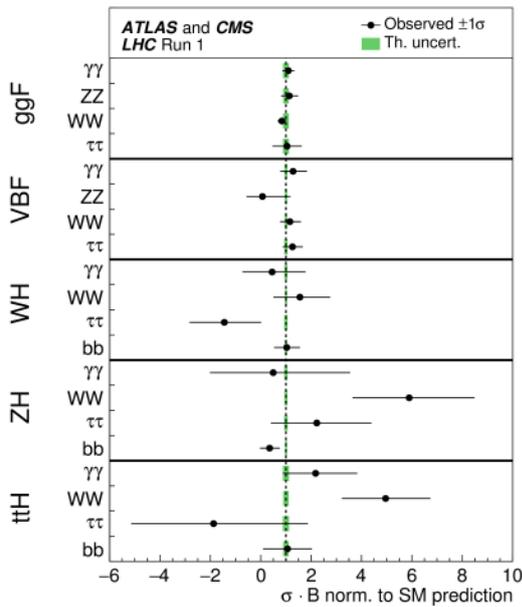


Figure 1: Best fit values of  $\sigma_i \cdot B^f$  for each specific channel  $i \rightarrow H \rightarrow f$ , as obtained from the generic parameterisation with 23 parameters for the combination of the ATLAS and CMS measurements. The error bars indicate the  $1\sigma$  intervals. The fit results are normalised to the SM predictions for the various parameters and the shaded bands indicate the theoretical uncertainties in these predictions.

ATLAS and CMS, have the ability to perform a very precise measurement of the Higgs boson mass due to the significant relevance of the high resolution boson decay modes  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  with an excellent invariant mass resolution of 1-2% in the mass range of interest. The measurement of the mass of the discovered boson in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  channel relied largely on the maximization of the efficiency for low transverse momentum leptons and on the precise calibration of the lepton momentum scale and resolution using  $J/\Psi$ ,  $\Upsilon$  and Z boson decays to di-leptons. In the case of the  $H \rightarrow \gamma\gamma$  channel, the mass measurement depended largely both on the possibility to maintain a good mass resolution in

high pileup regime (from both the energy and the direction contribution) and on the understanding of the electron/photon extrapolation for the energy scale (depending on the detector material, the shower shapes, etc). Recently the ATLAS and CMS collaborations have published a combined measurement of the Higgs boson mass [3], using the high resolution  $H \rightarrow ZZ$ , and  $H \rightarrow \gamma\gamma$  data from LHC Run 1, where Run 1 indicates the LHC data-taking period in 2011 and 2012 at pp centre-of-mass energies  $\sqrt{s} = 7$  and 8 TeV respectively. The combined ATLAS and CMS measured mass of the observed Higgs boson is  $125.09 \pm 0.24 \text{ GeV}$ .

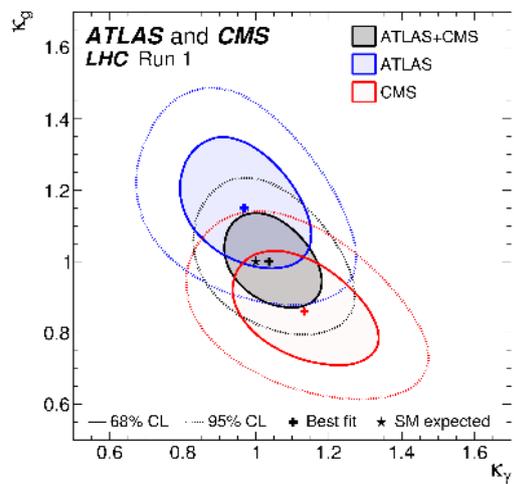


Figure 2: Likelihood scan for the loop induced coupling modifiers  $\kappa_g$  and  $\kappa_\gamma$  for ATLAS and CMS experiment after the LHC Run 1, and their combination.

Combined ATLAS and CMS measurements of the Higgs boson production and decay rates, as well as constraints on its couplings to vector bosons and fermions have been performed [4]. The combination is based on the analysis of different Higgs boson production processes and its decays to bosons and fermions. All results are reported assuming a value of 125.09 GeV for the Higgs boson mass, the result of the combined Higgs boson mass measurement by ATLAS and CMS experiments. The analysis uses the LHC proton-proton collision datasets recorded by the ATLAS and CMS detectors in 2011 and 2012, corresponding to integrated luminosities per experiment of approximately  $5 \text{ fb}^{-1}$  at  $\sqrt{s} = 7 \text{ TeV}$  and  $20 \text{ fb}^{-1}$  at  $\sqrt{s} = 8 \text{ TeV}$ . The Higgs boson production and decay rates of the two experiments are combined within the context of two generic parameterisations: one based on ratios of cross sections and branching ratios and the other based on ratios of cou-

pling modifiers ( $\kappa$ ), introduced within the context of a leading order Higgs boson coupling framework.

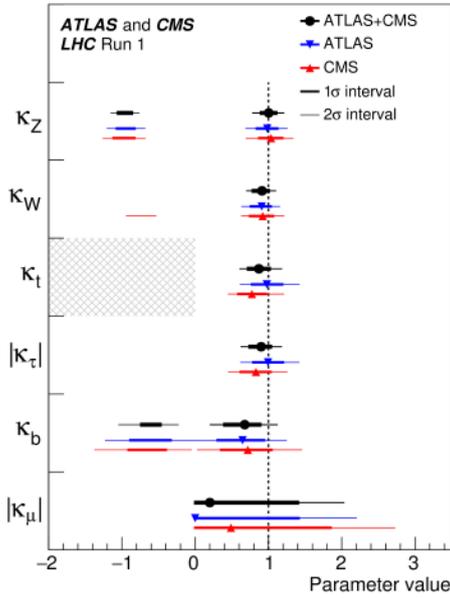


Figure 3: Fitted coupling strength modifiers  $\kappa_i$  values assuming no BSM decay.

The signal strength parameter  $\mu$ , defined as the ratio between the measured Higgs boson rate and its SM expectation, has been extensively used to characterise the Higgs boson yields. Assuming the SM values for the Higgs boson branching ratios, the five main Higgs boson production processes are explored with independent signal strengths. The signal strengths at the two energies are assumed to be the same for each production process. The p-value of the compatibility between the data and SM predictions is 24%. Similarly to the production case, Higgs boson decays can be studied with five independent signal strengths, one for each decay channel included in the combination, assuming that the Higgs boson production cross sections are the same as in the SM. Unlike the production, these decay-based signal strengths are independent of the collision centre-of-mass energy and therefore the 7 and 8 TeV datasets can be combined without additional assumptions. The p-value of the compatibility between the data and SM predictions is 60%.

The potential presence of BSM physics is also probed using specific parameterisations. This constraint applies to invisible decays into BSM particles, decays into BSM particles that are not detected as such, and modifications of the decays into SM particles that are not directly

measured by the experiments. Thus, assuming coupling modifier to vector boson  $|\kappa_V| \leq 1$  and  $\text{BR}_{\text{BSM}} \geq 0$ , an upper limit of  $\text{BR}_{\text{BSM}} = 0.34$  at 95% CL is obtained, compared to an expected upper limit of 0.39 which leaves enough room for BSM physics in the Higgs boson sector. In the photon and gluon loop induced couplings, fixing all other couplings at tree level to the SM, we find the coupling strength compatible with SM expectations.

The Higgs boson production and decay rates of the two experiments in all searched modes are combined with two generic parameterisations based on ratios of cross sections and branching ratios, and based on ratios of coupling modifiers. Results show a general consistency of data with SM predictions for all parameterisations considered. The combined signal yield relative to the SM expectation is measured to be  $1.09 \pm 0.07$  (stat)  $\pm 0.08$  (syst), where the systematic uncertainty is dominated by the theoretical uncertainty in the inclusive cross section calculations. The combined observed significance for the vector boson fusion production process is 5.4 standard deviations ( $\sigma$ ) with 4.6 $\sigma$  expected, while the observed significance for the tau pair decay is 5.5 $\sigma$  (5.0 $\sigma$  expected).

### 3. Towards rediscovery of the SM Higgs boson at LHC Run-2

The Run-2 analysis presented here has been performed with the 13 TeV pp collision data collected in the year 2015.

#### 3.1. $H \rightarrow \gamma\gamma$

The first measurements of the Higgs boson production in the two photons decay channel [5] with pp collisions at centre-of-mass energy of 13 TeV has been presented. The analysis is based on 2.7 fb<sup>-1</sup> of data collected with the CMS experiment at the LHC in the year 2015. Despite the small branching ratio predicted by the SM (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. In the LHC Run-2, this channel remains one of the best to perform a precise characterisation of the Higgs boson properties.

The analysis has been performed with an event classification based on the diphoton mass resolution, signal over background (S/B) ratio and the different Higgs boson production modes. Multivariate (MVA) techniques have been employed to discriminate the selection of photons, the photon pairs and the event vertex in the analysis. The determination of the primary vertex associated with  $H \rightarrow \gamma\gamma$  decay has a direct impact on the

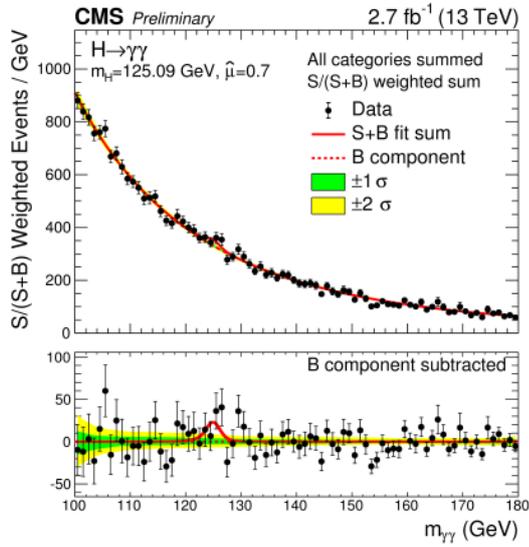


Figure 4: The diphoton invariant mass distribution combining all event categories.

diphoton mass resolution. This is dominated by the electromagnetic calorimeter energy contribution only if the interaction point is known to better than about 1 cm. The diphoton vertex assignment relies on a multivariate technique based on a Boosted Decision Tree (BDT). The inputs to the BDT are based on observables related to tracks recoiling against the diphoton system. In the presence of tracks from photons converted in the tracker, two additional variables are used as input to the BDT: the number of conversions and the pull between the longitudinal position of the reconstructed vertex and the longitudinal position of the vertex estimated using conversion tracks. The photon identification BDT uses the following inputs: shower shape observables; isolation variables based on sum of the  $p_T$  of photons and of charged hadrons within regions of  $R < 0.3$  around the candidate. Two such charged hadron isolations are used: one considering hadrons coming from the chosen vertex in the event, and one considering hadrons coming from the vertex with the largest  $p_T$  sum among remaining vertices. The latter is effective in rejecting photon candidates originating from mis-identification of jets from a vertex other than the chosen one; the energy median density per unit area in the event,  $\rho$ , makes the BDT independent of pileup; photon kinematic observables (pseudorapidity and energy): to account for the dependence of the shower topology and isolation variables on  $\eta$  and  $p_T$ . The MVA discriminator for  $\gamma\gamma$  selection takes as inputs variables related to the kinemat-

ics of the diphoton system (excluding  $m_{\gamma\gamma}$ ), a per-event estimate of  $m_{\gamma\gamma}$  resolution, and a per-photon identification measurement. This choice of inputs is justified by the fact that the relative magnitude of the contribution of fake photons from jets varies as a function of the photon kinematics, and so does the S/B ratio. In addition, the  $m_{\gamma\gamma}$  resolution depends on the location of the associated energy deposits in the calorimeter, whether or not one or both photons converted in the detector volume in front of electromagnetic calorimeter, and the probability that the correct primary vertex has been used to reconstruct the diphoton mass. A simultaneous fit for signal on top of smoothly falling background is done in the diphoton invariant mass spectrum. The observed significance for the SM Higgs boson is  $1.7\sigma$ , where a significance of  $2.7\sigma$  is expected. The best fit signal strength relative to the standard model prediction is  $0.69^{+0.47}_{-0.42}$ . The rediscovery of the Higgs boson in the diphoton decay mode would be established with more data collected in 2016.

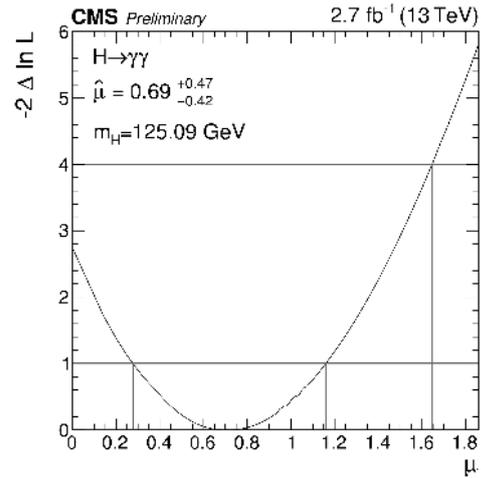


Figure 5: The likelihood scan for the signal strength modifier, where the value of the SM Higgs boson mass is fixed to  $m_H = 125.09$  GeV.

### 3.2. $H \rightarrow ZZ^* \rightarrow 4\ell$

The  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay mode [6] at the LHC is considered as the golden channel as it is a very powerful search channel that is sensitive to a SM Higgs boson throughout its entire allowed mass range with a clean signature of four isolated leptons from a common primary vertex. The analysis focuses on the electron and muon final states where the 4-lepton invariant mass resolution is excellent and the background due to jets misidentified as leptons is strongly suppressed. The major backgrounds to the signal include the di-boson  $ZZ$ ,

Z+jets, and top pair production. The SM di-boson irreducible ZZ background is estimated from Monte Carlo (MC) simulation. The Z+jets and top backgrounds are estimated from control samples in data. In such background estimation method, a control region is defined by reversing at least one signal selection requirement, and additional criteria can be added so that the control sample has a particular type of background contribution enhanced and does not overlap with the signal sample where all events passed the final selection. The background normalization in the control sample can be related to its normalization in the signal sample via a transfer factor. This transfer factor and its uncertainty are usually determined from the simulation samples. The background contribution in the signal sample can therefore be determined by the product of the background contribution in the data control sample and the transfer factor. Such method provides a data driven estimation of the background, and is widely used in the LHC experiments, including many Higgs boson related analyses.

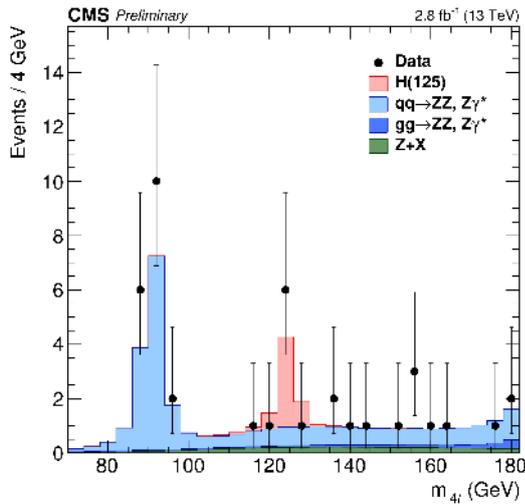


Figure 6: The four lepton invariant mass distribution combining all event categories.

The selected events are classified to  $4\mu$ ,  $2\mu 2e$  and  $4e$  categories, depending on the flavour of the leading pair of leptons. Events are further classified to enhance the sensitivity to various Higgs boson production modes. The observed significance for the standard model Higgs boson with  $m_H = 125.09$  GeV is  $2.5\sigma$ , where the expected significance is  $3.4\sigma$ . The signal strength modifier  $\mu$ , defined as the production cross section of the Higgs boson times its branch-

ing fraction to four leptons relative to the standard model expectation, is measured to be  $\mu = 0.99^{+0.33}_{-0.26}$  at  $m_H = 125.09$  GeV. The signal strength modifiers for the main Higgs boson production modes have also been constrained. The model independent fiducial cross section is measured to be  $2.29^{+0.74}_{-0.64}$  (stat)  $^{+0.30}_{-0.23}$  (syst)  $^{+0.01}_{-0.05}$  (model dep.) fb and differential cross sections as a function of the  $p_T$  of the Higgs boson and the number of associated jets are determined. The mass is measured to be  $m_H = 124.50^{+0.48}_{-0.46}$  GeV and the width is constrained to be  $\Gamma_H < 41$  MeV.

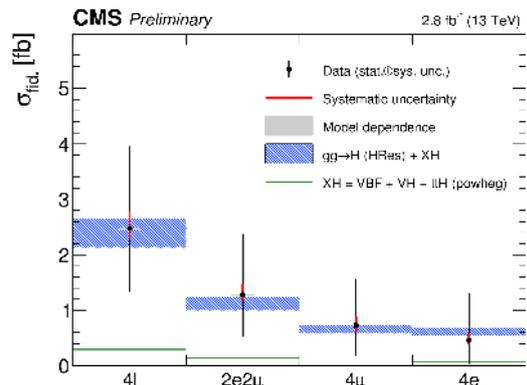


Figure 7: The measured fiducial cross section in each final state from simultaneous fit to the  $m_{4l}$  spectrum.

### 3.3. VBF $H \rightarrow b\bar{b}$

The  $H \rightarrow b\bar{b}$  decay in the vector boson fusion production process [7] is a challenging all-hadronic final state, dominant background coming from multijet production. Triggering is a big challenge for this decay mode in the dense hadron environment of the LHC. The strategy is similar to Run-1 analysis with triggering on forward jets + central bottom quark jet(s) targeting production in the vector boson fusion mode and categorisation of events based on MVA discriminant. The dominant background to this search is from QCD production of multijet events. Other backgrounds arise from: (i) hadronic decays of Z or W bosons produced in association with additional jets, (ii) hadronic decays of top quark pairs, and (iii) hadronic decays of singly-produced top quarks. The search is performed on selected four-jet events that are characterized by the response of a multivariate discriminant trained to separate signal events from background without making use of kinematic information on the two b-jet candidates. First, the resolution of the invariant mass of the two

b-quark jets is improved by applying multivariate regression techniques. Then jet composition properties are used to separate jets originating from light quarks or gluons. Third, soft QCD activity outside the jets is quantified and used as a discriminant between QCD processes with strong colour flow and the VBF signal without colour flow. Subsequently, the invariant mass distribution of two bottom quark jets is analyzed in each category in the search for a signal bump above the smoothly falling contribution from the SM background. The combination of the results obtained in this search with the similar CMS search with Run 1 data yields an observed (expected) upper limit of 3.4 (2.3) times the SM prediction. The combined fitted  $H \rightarrow b\bar{b}$  signal strength from Run-1 and Run-2 is  $\mu = 1.3^{+1.2}_{-1.1}$ , with a signal significance of  $1.2\sigma$  for  $m_H = 125$  GeV.

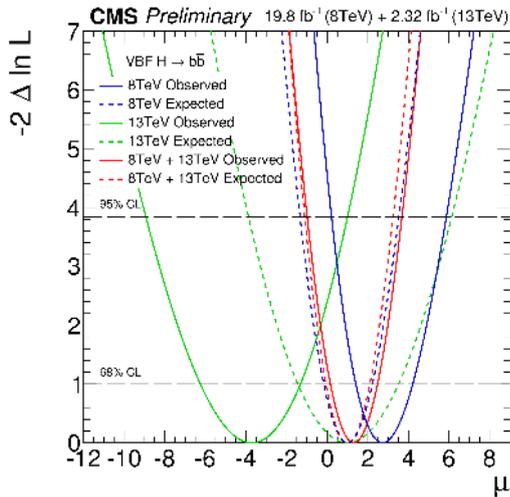


Figure 8: Observed and SM-expected likelihood profile of the signal strength with  $m_H = 125$  GeV, using Run-1 8 TeV data, Run-2 13 TeV data, and for the combination of 8 TeV and 13 TeV data.

### 3.4. $t\bar{t}H$ production

The top-Higgs interaction vertex is directly accessible when the Higgs boson is produced in association with one or more top quarks, since the Higgs boson is too light to decay to top quarks directly. The associated production with a single top quark is suppressed by the SM by interference from couplings of the Higgs boson to the top quark and the W boson; in certain BSM theories this contribution is enhanced and needs to be considered also in corresponding  $t\bar{t}H$  interactions. On the other hand, the associated production of a Higgs boson and a top quark-antiquark pair ( $t\bar{t}H$  production) is a

direct probe of the top-Higgs coupling. The  $t\bar{t}H$  analyses with 13 TeV data are separated by the Higgs boson decays into  $t\bar{t}H(b\bar{b})$ ,  $t\bar{t}H(\gamma\gamma)$ ,  $t\bar{t}H(\text{multileptons})$  [8–10] and further subdivided by the  $t\bar{t}$  decays. For each of the channels an independent analysis is performed. In each, the corresponding dominant reducible and irreducible background processes are  $t\bar{t}$  associated with other particles. In each analysis, events are classified in two major categories according to the lepton multiplicities and then further subdivided into more categories for better S/B ratio separation.

The bottom pair decay has the largest Higgs boson branching ratio, but with an overwhelming background coming from  $t\bar{t} + \text{jets}$ . The main strategy is to obtain a good signal separation and constrain the backgrounds. The events are categorised into two categories targeting different  $t\bar{t}$  decay topologies: lepton+jets and dileptons. The lepton+jets category has high statistics whereas the dilepton category has the minimal non top pair backgrounds. Further classification of events by the number of jets associated, b-jet multiplicities and boosted jets (lepton+jets) is done with 13 subcategories formed in the analysis. The combined best-fit value of the signal strength is  $\mu = -2.0 \pm 1.8$  that is  $1.7\sigma$  below the SM expectation.

Despite the small branching ratio predicted by the SM (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 due to the excellent photon identification and energy resolution. The analysis follows the same strategy as the  $H \rightarrow \gamma\gamma$  analysis. Diphoton selection to separate signal from background is done using BDT approach. Event classification is performed via mass resolution, signal over background ratio and production mechanism. The analysis is performed with two  $t\bar{t}H$  categories with leptonic and hadronic tags. In the dominant hadronic category, a small excess of events appears around the mass of the Higgs boson, while in the leptonic category, only three events pass the event selection criteria, and none of them are observed in the Higgs mass region. The observed signal strength for the combination of the two categories is  $\mu = 3.8^{+4.5}_{-3.6}$ , estimated for  $m_H = 125.09$  GeV, and is in agreement with the SM within the large uncertainties. The results are statistically limited and thus better precision is expected with increase of statistics.

The analysis of Higgs boson decays to multileptons targets leptonic final states from Higgs boson decays to WW, ZZ, or tau pair, with at least one Z, W, or tau decaying leptonically. Despite the small branching ratio, the presence of one or two additional leptons from top quark decays, leads to clean experimental sig-

natures having two same-sign leptons or at least three leptons (electrons or muons), plus b-tagged jets. The main backgrounds are irreducible  $t\bar{t}+V$  obtained from MC and reducible  $t\bar{t}+jets$  estimated from data. The two categories in the analysis are with 2 same sign leptons and  $\geq 3$  leptons. The sub-categorisation is based on the lepton charge, presence of hadronic tau, lepton flavour, presence of 2 b-tags and signal over background bins. The separation of signal from top pair,  $t\bar{t}+V$  via BDTs with categories division in signal over background bins. The modelling of fake lepton backgrounds from control region relaxing lepton selection. The combined best fit of all sub-categories is  $0.6^{+1.4}_{-1.1}$ .

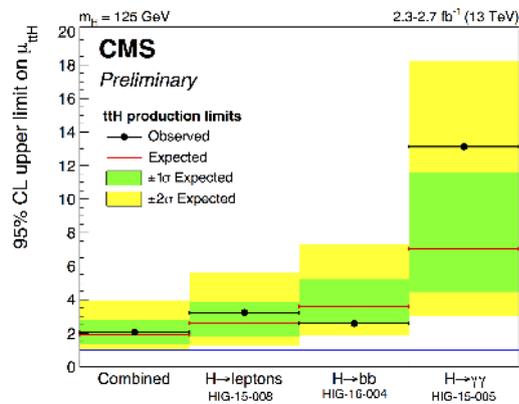


Figure 9: 95% confidence level upper limits on the rate of  $t\bar{t}H$  production relative to the standard model expectation.

Combining the three statistically independent analysis channels [11] and assuming a Higgs boson mass of  $m_H = 125$  GeV, and correlating the common systematic uncertainties, a signal strength of  $\mu = 0.15^{+0.95}_{-0.85}$  is obtained which is in agreement with the SM expectation.

#### 4. Conclusion

A combination of CMS and ATLAS measurements of the Higgs boson mass, production and decay rates have been performed using Run-1 LHC data. The combined mass measurement of ATLAS and CMS in the  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay channels yield a Higgs boson mass 125.09 GeV. A combination of CMS and ATLAS measurements of the Higgs boson production and decay rates have been performed utilising all accessible production and decay modes. Any deviations in the coupling of the Higgs boson to fermions and vector bosons from that expected in the standard model have been explored in different models. No significant deviation from SM has been observed. The global signal

strength has been measured  $1.09 \pm 0.11$ . The year 2015 has marked the start of the CMS experiment taking the 13 TeV LHC proton collision data under the 25ns bunch crossing conditions and physics objects have been very well reconstructed and commissioned. The first results with the 13 TeV data collected in the year 2015 have been reported in the Higgs boson decay to bosons and fermions. No large excess over SM is found in Higgs boson production with top quark pair and a signal strength of  $\mu = 0.15^{+0.95}_{-0.85}$  is obtained at a Higgs boson mass of  $m_H = 125$  GeV. The CMS experiment is ready to measure Higgs boson properties with more precision from the additional 13 TeV data in 2016, with measurements like the Higgs boson cross section being part of the standard analysis in LHC Run-2.

#### References

- [1] The ATLAS Collaboration, JINST 3 (2008) S08003; The CMS Collaboration, JINST 3 (2008) S08004.
- [2] The ATLAS Collaboration, Phys. Lett. B 716 (2012) 1; The CMS Collaboration, Phys. Lett. B 716 (2012) 30.
- [3] The ATLAS and CMS Collaborations, Phys. Rev. Lett. 114 (2015) 191803.
- [4] The ATLAS and CMS Collaborations, JHEP 08 (2016) 045.
- [5] The CMS Collaboration, CMS-PAS-HIG-15-005; and all references therein.
- [6] The CMS Collaboration, CMS-PAS-HIG-15-004; and all references therein.
- [7] The CMS Collaboration, CMS-PAS-HIG-16-003; and all references therein.
- [8] The CMS Collaboration, CMS-PAS-HIG-15-003; and all references therein.
- [9] The CMS Collaboration, CMS-PAS-HIG-15-008; and all references therein.
- [10] The CMS Collaboration, CMS-PAS-HIG-16-004; and all references therein.
- [11] [twiki.cern.ch/twiki/bin/view/CMSPublic/TTHCombMoriond2016](http://twiki.cern.ch/twiki/bin/view/CMSPublic/TTHCombMoriond2016)