

Measurements of the Higgs $H(125)$ boson at CMS

Valeria Botta^{1,a}

on behalf of the CMS Collaboration

¹Deutsches Elektronen-Synchrotron (DE)

Abstract. The most recent measurements of the Higgs $H(125)$ boson in several final states, including decays to bosons, fermions, and the associated production of a Higgs boson with top quarks, are reviewed. Results have been obtained analysing the proton-proton collision data collected with the CMS detector at the Large Hadron Collider (LHC) in 2016, corresponding to an integrated luminosity of 35.9 fb^{-1} at a centre-of-mass energy of 13 TeV.

1 Introduction

In July 2012 the ATLAS and CMS Collaborations announced the discovery of a new particle [1–3], compatible with the neutral scalar boson predicted by the Brout-Englert-Higgs mechanism of electroweak symmetry breaking [4–6], commonly known as the Higgs boson. From a combination of ATLAS and CMS measurements [7], the mass of the Higgs boson has been determined to be $125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst.}) \text{ GeV}$. For a Higgs boson with this mass, several decay modes are experimentally accessible at the LHC, and a large number of physics analyses are being performed to characterise the discovered particle. In this Contribution, a selection of the latest results on the Higgs $H(125)$ boson obtained by the CMS Collaboration are presented, including Higgs decays to bosons, fermions, and the production of Higgs bosons in association with top quarks. The results reported here have been obtained analysing the LHC proton-proton collision data collected with the CMS detector [8] in 2016, corresponding to an integrated luminosity of 35.9 fb^{-1} at a centre-of-mass energy of 13 TeV.

2 Higgs boson decays to bosons

2.1 $H \rightarrow ZZ^*$

The Higgs boson decay to a pair of Z bosons $H \rightarrow ZZ^*$ is one of the most sensitive channels to study the $H(125)$ boson at the LHC. In fact, events where both Z bosons decay to a pair of leptons ($\ell = e, \mu$) are efficiently recorded and their kinematics can be precisely reconstructed at CMS. The analysis of the full 2016 dataset [9] targets the final states with 4μ , $2e2\mu$ and $4e$. The reconstructed invariant mass of the four-lepton system can be seen in Fig. 1 (left). The $H \rightarrow ZZ^*$ signal is extracted from a bi-dimensional maximum-likelihood fit that uses both the four-lepton invariant mass and a kinematic

^ae-mail: valeria.botta@cern.ch

discriminant built with matrix element techniques to discriminate against the background. Dedicated event categories are included to target the VBF, VH and $t\bar{t}H$ Higgs boson production modes. The cross section for the production and decay $pp \rightarrow H \rightarrow ZZ^* \rightarrow 4\ell$ is measured in a fiducial phase space and it is found to be $2.92^{+0.48}_{-0.44}(\text{stat})^{+0.28}_{-0.24}(\text{syst})$ fb, which is compatible with the SM prediction of 2.76 ± 0.14 fb. Fiducial cross sections are also measured differentially as a function of the transverse momentum of the Higgs boson, as it can be seen in Fig. 1 (right), and as a function the number of associated jets, and the transverse momentum of the leading associated jet. The comparison to the theory predictions shows good agreement within the uncertainties. The dominant sources of systematic uncertainty are experimental, and arise from the lepton identification efficiency and from the measurement of the integrated luminosity.

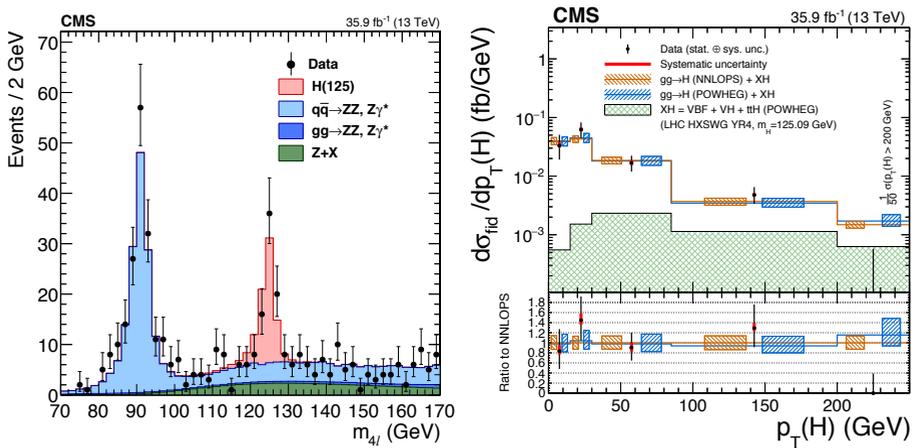


Figure 1. Left: distribution of the reconstructed four-lepton invariant mass in the mass range from 70 to 170 GeV. Points with error bars represent the data and stacked histograms represent expected signal and background distributions. The SM Higgs boson signal and the ZZ backgrounds are normalised to the SM expectation, whilst the $Z+X$ background is normalised to the estimation from data. Right: fiducial differential cross section as a function of the Higgs boson p_T . The component of the signal produced via VBF, VH and $t\bar{t}H$ processes is denoted as XH . The last bin represents the integrated cross section for $p_T(H) > 200$ GeV and is scaled by $1/50$. No events are observed with $p_T(H) > 200$ GeV [9].

2.2 $H \rightarrow \gamma\gamma$

Despite its low branching fraction of about 0.2% predicted by the SM, the $H \rightarrow \gamma\gamma$ decay has a very clean event signature and the diphoton invariant mass $m_{\gamma\gamma}$ can be reconstructed with high resolution. Events with two highly energetic isolated photons are categorised based on the estimated mass resolution, and the Higgs boson signal is extracted from a simultaneous maximum likelihood fit to the $m_{\gamma\gamma}$ distribution in all the categories. Figure 2 (left) shows the combined $m_{\gamma\gamma}$ spectrum for all categories, together with the fitted signal and background components. The inclusive cross section for the production and decay $pp \rightarrow H \rightarrow \gamma\gamma$ is measured in a fiducial phase space having a signal acceptance of 60% and it is found to be $84 \pm 11(\text{stat.}) \pm 7(\text{syst.})$ fb, in agreement with the SM prediction of 75 ± 4 fb within the uncertainties. This measurement is currently the most precise one and it is limited by the data statistics. Fiducial cross sections have been measured also differentially as a function of the re-

constructed Higgs boson p_T (Fig. 2, right) and of the number of associated jets. The measurements agree well with the theory predictions within their uncertainties.

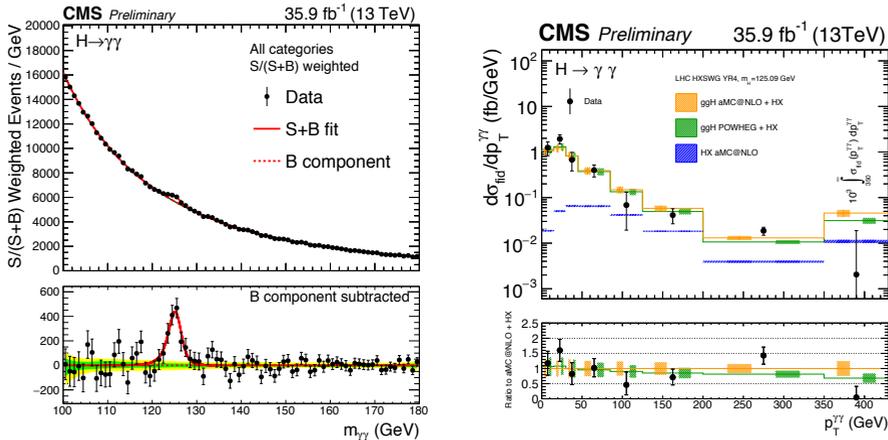


Figure 2. Left: observed diphoton invariant mass spectrum weighted by the ratio $S/(S + B)$ for all event categories, together with the fitted signal and background components. The bottom panel shows the weighted mass spectrum after the background subtraction. Right: measured $H \rightarrow \gamma\gamma$ differential cross section (black points) as a function of the reconstructed Higgs boson p_T , compared to the theoretical predictions (coloured rectangles). The sum of the contributions from VBF, VH and $t\bar{t}H$ processes is labeled as HX [10].

3 Higgs boson decays to fermions

3.1 $H \rightarrow \tau\tau$

The Higgs boson decay to a pair of τ leptons is the most sensitive channel to probe the direct coupling of the Higgs boson to fermions, due to the expected large event rate in the SM and small contribution from background events, with respect to other fermionic decay modes. The analysis of the full 2016 dataset [11] considers the four most likely final states for $\tau\tau$ pairs, namely $\tau_h\tau_h$, $\mu\tau_h$, $e\tau_h$ and $e\mu$, where τ_h indicates a tau lepton decaying to hadrons. Events are categorised in three mutually exclusive categories: the 0-jet category, targeting gluon-fusion Higgs boson production, contains events with no jets with $p_T > 30$ GeV and $|\eta| < 4.7$; the VBF category, targeting vector-boson fusion production, contains events with two jets with invariant mass $m_{jj} > 300$ GeV; the boosted category, that includes all other events not entering one of the previous categories, to account for gluon fusion production accompanied by jets, softer VBF events and VH production with the vector boson decaying hadronically. The signal is extracted with a maximum likelihood fit to two-dimensional distributions of two sensitive variables, namely the mass of the di-tau system and a second variable chosen to disentangle the signal from the most important background processes in that channel and category. The main irreducible background is the $Z \rightarrow \tau\tau$ process, estimated from simulation with data driven corrections. The QCD multi-jet production is the main background in the $\tau_h\tau_h$ channel and it is fully estimated from data. The normalisations of the QCD, W +jets and $t\bar{t}$ backgrounds are derived from data thanks to dedicated control regions included in the final fit. An excess of events corresponding to a significance of 4.9σ is observed in data, the expectation for the SM Higgs boson being 4.7σ . The left panel of Fig. 3 shows

the observed excess of events in the di-tau mass distribution for some channels and categories. The measured signal strength $\mu = 1.09^{+0.27}_{-0.26}$ is compatible among final states and with the SM expectation within the uncertainties, as it can be seen in the right panel of Fig. 3. The main contribution to the signal strength uncertainty is statistical. Among the experimental uncertainties that affect the shapes of the di-tau mass, the most important are the ones related to the hadronic tau reconstruction. In the combination of the 7, 8 and 13 TeV results from CMS, the decay of the Higgs boson to a pair of τ leptons is observed with a significance of 5.9σ (expected 5.9σ). This provides the first observation of the Higgs boson decay to τ leptons by a single experiment.

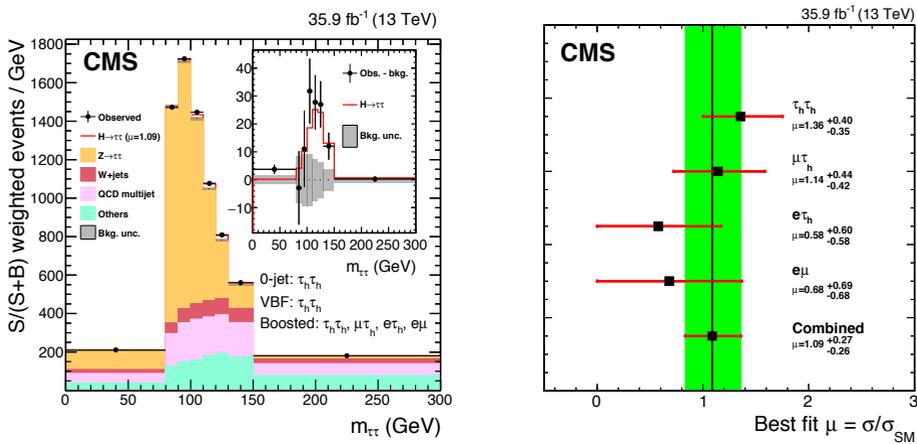


Figure 3. Left: observed and predicted $m_{\tau\tau}$ distribution, combined for the categories of the $\tau_h\tau_h$ channel and the boosted category of the $\mu\tau_h$, $e\tau_h$ and $e\mu$ channels, and weighted according to $S/(S+B)$. The inset shows the corresponding difference between the observed data and expected background distributions, together with the SM Higgs signal expectation. Right: best fit signal strength per channel, for $m_H = 125.09$ GeV. The combined best fit signal strength is $\mu = 1.09^{+0.27}_{-0.26}$ [11].

3.2 $VH, H \rightarrow b\bar{b}$

The decay into a b-quark pair is the main decay mode for the SM Higgs boson, with a branching fraction $\mathcal{B} \approx 58\%$. However, the search for $H \rightarrow b\bar{b}$ decays is particularly challenging at the LHC due to the extremely large multi-jet background. The associated production of the Higgs boson with a vector boson provides a handle to select the events and to reduce the background. The $VH, H \rightarrow b\bar{b}$ analysis targets final states with the vector boson decaying leptonically, namely $Z(\nu\nu)H, W(\ell\nu)H,$ and $Z(\ell\ell)H,$ with $\ell = e, \mu$. Boosted decision trees (BDT) are trained separately for each channel to search for a Higgs boson with a mass of 125 GeV, and the results are obtained from a maximum-likelihood fit to the BDT output distributions. An example of fitted BDT output can be seen in Fig. 4 (left). The most discriminating variable is the reconstructed invariant mass of the $b\bar{b}$ system, whose resolution is improved by $\approx 10\%$ thanks to a dedicated b-jet energy regression. The main background for the analysis is the $V+b$ -jets production, that is modelled from simulation with data-driven corrections. Several control regions are included in the final fit to derive the background normalisation from the data. The analysis strategy is validated performing a search for $VZ(b\bar{b})$ production, that presents the same event topology and a comparable cross section as the $H \rightarrow b\bar{b}$ process, and it is observed with a significance

of 5.0σ (expected 4.9σ). An excess of events compatible with the SM Higgs boson decaying to a $b\bar{b}$ pair is observed with a significance of 3.3σ (expected 2.8σ) for $m_H = 125$ GeV. The measured signal strength is $\mu = 1.19^{+0.40}_{-0.38}$, compatible with the SM expectation and among different channels, as can be seen in Fig. 4 (right). When combining the latter result with the analyses performed by CMS on the 7 and 8 TeV data, the significance of $VH, H \rightarrow b\bar{b}$ process amounts to 3.3σ (expected 3.8σ) and the signal strength $\mu = 1.06^{+0.31}_{-0.29}$. This provides a strong evidence of the $H \rightarrow b\bar{b}$ decay. The final result of this analysis from CMS became available after this Conference and can be found in [13].

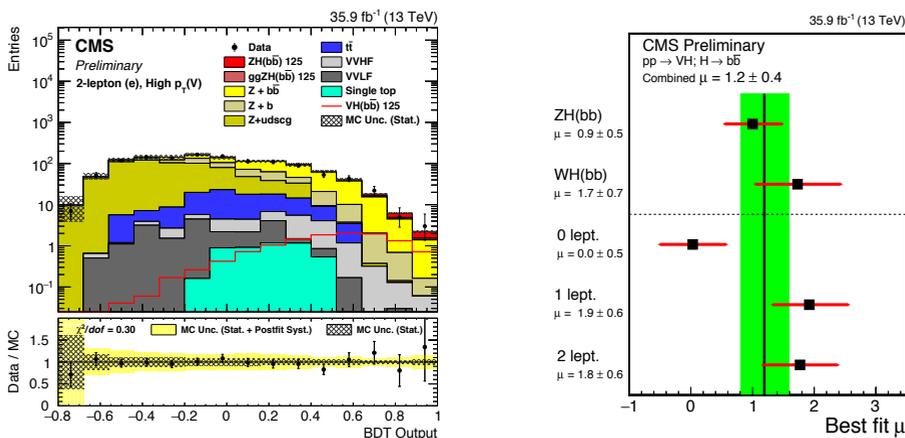


Figure 4. Left: example of post-fit BDT output distribution, for the $Z(ee)$ channel in the high- $p_T(Z)$ region. The bottom inset shows the ratio of the number of events observed in data to that of the prediction from simulated samples for signal and backgrounds. Right: best-fit value of the signal strength μ for a 125 GeV Higgs boson separately for the WH and ZH production modes as well as for final states with different number of leptons [12].

3.3 Boosted $H \rightarrow b\bar{b}$

The first inclusive search for $H \rightarrow b\bar{b}$ decays at the LHC has been performed by CMS on the 2016 dataset [14]. In order to improve the analysis sensitivity, the search focuses on the boosted topology, where the Higgs boson has $p_T > 450$ GeV and recoils against an initial-state radiation jet, that is used for the online selection of the events. A highly boosted Higgs boson decaying to $b\bar{b}$ is reconstructed as a single, large radius jet and identified using novel techniques to probe its two-prong and double b-tagging signature. The two-prong discriminator has an efficiency of $\approx 58\%$ and a mistag rate of $\approx 26\%$, while the double b-tagging identification has an efficiency of $\approx 33\%$ for signal events and $\approx 1\%$ for multi-jet background events. The main background process is the QCD multi-jet production, that is estimated from data, using a control region with events failing the double b-tag requirement. Figure 5 shows the mass distribution for large jets failing (left) and passing (right) the double b-tag identification, in which the W and Z boson mass peaks can be seen when subtracting the multi-jet and $t\bar{t}$ background. Events are further categorised depending on the p_T of the Higgs boson candidate and the signal is extracted from a combined binned maximum likelihood fit to the jet mass distribution in the passing and failing regions and in a $t\bar{t}$ -enriched control region. The analysis is validated using $Z \rightarrow b\bar{b}$ events, resulting in the first observation of the latter process in the single-jet topology with a local significance of 5.1σ (expected 5.8σ). The $H \rightarrow b\bar{b}$ signal is observed with a significance of

1.5σ (expected 0.7σ), the best-fit signal strength being $\mu = 2.3^{+1.8}_{-1.6}$. The measured cross section times branching fraction for gluon fusion production of $H \rightarrow b\bar{b}$ with reconstructed $p_T > 450$ GeV and $|\eta| < 2.5$ is $74 \pm 48(\text{stat})^{+17}_{-10}(\text{syst})$ fb, which is consistent within uncertainties with the SM prediction. The final result of this analysis became available after this Conference and can be found in [15].

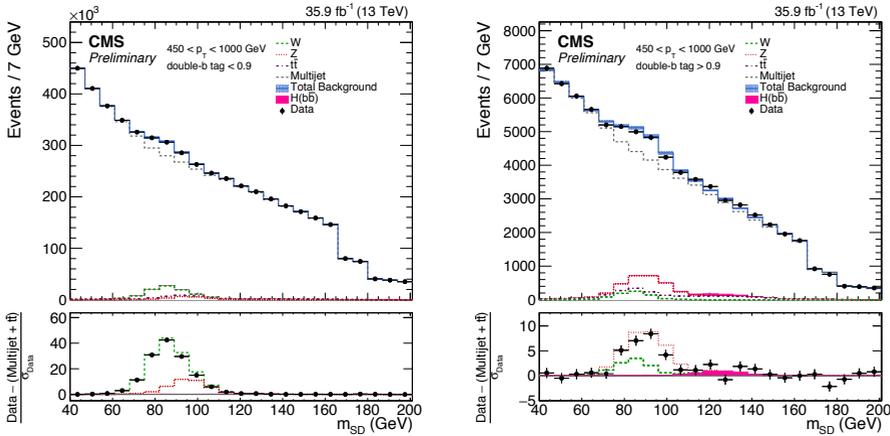


Figure 5. Post-fit jet mass distribution for data and measured SM background, for events failing (left) and passing (right) the double b-tag requirement. The contributions from W and Z boson production are clearly visible [14].

4 Higgs boson production in association with top quarks

4.1 $t\bar{t}H$

The production of a Higgs boson in association with top quark pairs $t\bar{t}H$ provides a direct probe to measure the top quark Yukawa coupling, that in some new physics models is predicted to be enhanced with respect to the SM expectation. The $t\bar{t}H$ production mode is a low rate process, with a SM cross section $\sigma = 507$ fb at $\sqrt{s} = 13$ TeV. Therefore, many background sources need to be considered in the analyses, from the large $t\bar{t} + V$ and $t\bar{t} + \text{jets}$ production to the more rare backgrounds like production of same-sign W boson pairs and tri-boson production. Small experimental effects like lepton charge misidentification and jets misidentified as leptons are taken into account as well. In order to maximise the sensitivity, multivariate analysis techniques are largely employed in $t\bar{t}H$ searches. The most promising channels are the ones with more than one lepton in the final state, targeting $H \rightarrow WW^*, ZZ^*$ and $\tau\tau$ decays. CMS has analysed the full 2016 dataset searching for $t\bar{t}H$ production in final states with 2,3 or 4 leptons [16]. In this search, events are classified based on the number of leptons and their flavour and the number of b-tagged jets. The background originating from non-prompt leptons is rejected thanks to a dedicated BDT discriminant. The signal is extracted with a maximum likelihood fit of two BDT outputs, trained separately against the $t\bar{t} + \text{jets}$ and $t\bar{t} + V$ backgrounds. The contribution of non-prompt and charge-misidentified leptons is estimated from data, with an uncertainty of $\approx 30\%$, while other backgrounds are estimated from simulation, with an uncertainty of $\approx 10\%$. An excess of events is observed in data, compatible with the $t\bar{t}H$ signal. The significance of the excess is 3.3σ (expected 2.4σ), and the best-fit signal strength for a Higgs boson of 125 GeV is $\mu = 1.5 \pm 0.3(\text{stat}) \pm 0.4(\text{syst})$ compatible among different final states, as shown in Fig. 6 (right).

References

- [1] ATLAS Collaboration, Phys. Lett. B **716** 1 (2012)
- [2] CMS Collaboration, Phys. Lett. B **716** 30 (2012)
- [3] CMS Collaboration, JHEP **06** 081 (2013)
- [4] F. Englert and R. Brout, Phys. Rev. Lett. **13** 321 (1964)
- [5] P. W. Higgs, Phys. Lett. **12** 132 (1964)
- [6] P. W. Higgs, Phys. Rev. Lett. **13** 508 (1964)
- [7] ATLAS and CMS Collaborations, Phys. Rev. Lett. **114** 191803 (2015)
- [8] CMS Collaboration, JINST **3** S08004 (2008)
- [9] CMS Collaboration, CMS-HIG-16-041, <https://cds.cern.ch/record/2272260>, sub. to JHEP
- [10] CMS Collaboration, CMS-PAS-HIG-17-015, <https://cds.cern.ch/record/2257530>
- [11] CMS Collaboration, CMS-HIG-16-043, <https://cds.cern.ch/record/2276465>, sub. to PLB
- [12] CMS Collaboration, CMS-PAS-HIG-16-044, <https://cds.cern.ch/record/2278170>
- [13] CMS Collaboration, CMS-HIG-16-044, <https://cds.cern.ch/record/2285286>, sub. to PLB
- [14] CMS Collaboration, CMS-PAS-HIG-17-010, <https://cds.cern.ch/record/2266164>
- [15] CMS Collaboration, CMS-HIG-17-010, <https://cds.cern.ch/record/2284496>, sub. to PRL
- [16] CMS Collaboration, CMS-PAS-HIG-17-004, <https://cds.cern.ch/record/2256103>
- [17] CMS Collaboration, CMS-PAS-HIG-17-005, <https://cds.cern.ch/record/2264553>