Evidence of Higgs boson decay to bottom quarks: a review

Daniel Stefanian
School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel

December 12, 2024

Abstract

This paper reviews the observation of the Standard Model Higgs boson decay to bottom quarks ($H \to b\bar{b}$), particularly focusing on processes involving Higgs bosons produced in association with W or Z bosons (VH). Analyzed final states included 0, 1, or 2 charged leptons along with two bottom quark jets. Data from the CMS measurement in 2017, totaling 41.3 fb^{-1} from proton-proton collisions at a center-of-mass energy of 13 TeV, is presented. This, along with previous measured data at a center-of-mass energy of 7 and 8 TeV, combined, revealed an event excess at a Higgs mass of 125 GeV, yielding a significance of 4.8σ, close to the expected 4.9σ, and a measured signal strength of $\mu = 1.01 \pm 0.22$ relative to the standard model expectation. When combined with other production processes for $H \to b\bar{b}$, the observed and expected significance increased to 5.6σ and 5.5σ respectively, with a signal strength of $\mu = 1.04 \pm 0.20$.

1 Introduction

Predicted and later discovered in 2012 [1, 25, 26], the Higgs boson is a fundamental particle that confirms the existence of the Higgs field, which gives mass to other particles [2]. Through the *Higgs mechanism*, particles acquire mass by interacting with this field, with stronger interactions resulting in greater mass. This discovery supports the Standard Model's (SM) explanation of mass generation. This study centers on the observation of the Higgs boson decay into bottom (b) quarks ($H \to b\bar{b}$) using newly analyzed 2017 data at 13 TeV [28], complemented by earlier CMS measurements from Run 1 at 7 and 8 TeV [13, 32] and Run 2 in 2016 at 13 TeV [19], as well as results from other production processes [10, 30, 31]. The SM, through the Higgs mass generation mechanism and Yukawa couplings, predicts that the Higgs boson decays to a bottom quark-antiquark pair with a branching fraction of approximately 58% [27].

Decay is the process by which an unstable particle transforms into lower-mass particles to reach a lower-energy state, driven by fundamental forces such as the weak nuclear force. Decay is a statistical process, so the timing and manner of decay are random. Each possible decay channel, referring to a specific particle decay pathway, has a probability, called a branching ratio, which represents the likelihood of this particular decay channel occurring [6].

The study of $H \to b\bar{b}$ decay is crucial for understanding the Yukawa coupling [5], which quantifies the strength of the interaction between the Higgs field and fermions.

The measurement uses the data collected at the LHC (Large Hadron Collider) and CMS (Compact Muon Solenoid). The LHC, located at CERN in Geneva, is a 27-kilometer-long ring with superconducting magnets. It includes 1,232 dipole magnets to bend proton beams and 392 quadrupole magnets to focus them. Just before collisions, specialized magnets focus the beams to increase collision probability [3]. The CMS detector captures the stable particles produced. It features a 6-meter diameter superconducting solenoid generating a 3.8 Tesla magnetic field, silicon trackers for detecting charged particles, calorimeters to measure electron/photon and hadron energy, muon spectrometer and a two-tier trigger system for event selection [4].

The Higgs bosons in this measurement are produced in high-energy proton-proton collisions at the LHC with a center-of-mass (CM) energy of $\sqrt{s} = 13$ TeV [28], representing the total data collected and the potential collision events contributing to the analysis. These collisions, through gluon fusion, vector boson fusion, and VH production, create the Higgs bosons. Gluon fusion, the dominant mechanism, involves two gluons fusing to produce a Higgs boson [9]. In vector boson fusion, two quarks exchange W or Z bosons, which then produce a Higgs boson [10]. The VH production process, where a Higgs boson is produced alongside a W or Z boson, as illustrated in Figure 1, is the focus of this measurement due to its sensitivity in studying Higgs boson decay to b quarks. For two particles, the center-of-mass energy E_{CM} is given by:

$$E_{CM} = \sqrt{(E_1 + E_2)^2 - (\vec{p_1} + \vec{p_2})^2} = \sqrt{s} \quad (1)$$

where E_1 and E_2 are the energies of the two colliding particles, and $p\vec{_1}$ and $p\vec{_2}$ are their momenta.

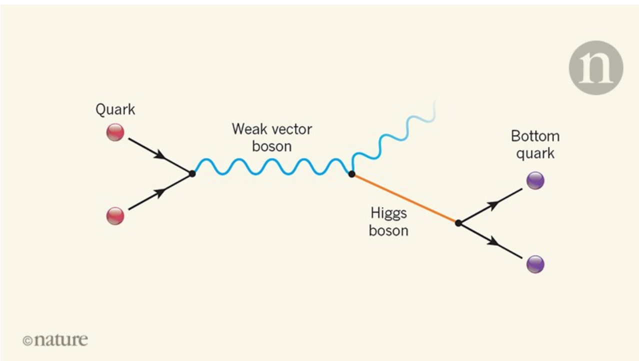

Figure 1: Diagram showing two quarks interacting to produce both a charged weak vector boson W and a Higgs boson, with the Higgs subsequently decaying into two b quarks ($H \to b\bar{b}$). Reproduced from [7].

2 Methods and analysis

The analysis considers five final states, which are specific configurations of particles resulting from the collisions: $Z(\nu\nu)H, W(\mu\nu)H, W(e\nu)H, Z(\mu\mu)H$ and $Z(ee)H$. These states correspond to decay channels with 0, 1, or 2 charged leptons produced from the decay of a vector boson.

The analysis specifically searched for two $b-jets$, which are collimated streams of particles originating from the hadronization of b quarks produced in the decay of a Higgs boson into b quarks [11]. Key background processes include the production of W and Z bosons accompanied by jets, top (t) quark pair production, diboson events that generate two vector bosons such as WW, and multijet events involving multiple jets from various interactions.

Each channel has a specific signal region with a higher concentration of VH events and several control regions designed to isolate and measure distinct background processes. A binned-likelihood fit is applied to the distributions of the deep neural network (DNN) score [20] in the signal regions and the dijet invariant mass (m_{jj}) distribution in the control regions to identify a potential Higgs boson signal, normalizing and comparing data across all channels, where dijet mass refers to the invariant mass calculated from two jets produced in a particle collision.

To ensure the reliability of the analysis, the same method is applied to extract a signal from the associated production process VZ, where the Z boson decays into b quarks ($Z \to b\bar{b}$). This process shares the same final state as the VH process, where the Higgs boson decays into b quarks ($H \to b\bar{b}$), but VZ has a much larger production cross section—5 to 15 times greater—depending on the interaction conditions, such as the center-of-mass energy ($\sqrt{s}$) and momentum transfer during proton-proton collisions, representing the likelihood of the process occurring during collisions. The VH signal is identified through techniques like multivariate discriminants, enriched signal region selection targeting VH events, and control regions to model and subtract background contributions, enabling the extraction of the VH signal despite the larger VZ background.

Both signal and background processes are simulated using Monte Carlo (MC) event generators. The detector's response is modeled by GEANT4 [15], which simulates particle interactions with matter. Signal processes, ZH and WH production, are simulated with high precision using next-to-leading order (NLO) techniques, providing a more accurate description by including corrections beyond the simplest calculations. ZH production from gluons is simulated at leading order (LO), the simplest approximation. The signal and background processes are simulated with event generators, ensuring consistency with real data by accounting for additional proton-proton interactions, known as pileup.

The CMS particle-flow (PF) algorithm is essential for reconstructing and identifying individual particles [16] like electrons, muons, photons, and hadrons from collision data. It operates in both the initial triggering and later detailed analysis stages. This process is important because accurately reconstructing and identifying individual particles and their momenta is crucial for correct analysis of the events produced in the measurements. In the current analysis, the CMS PF algorithm and the identification of jets and missing transverse momentum are used to select and characterize events that are relevant for Higgs boson searches, allowing for precise measurements of kinematic variables. Jets, clusters of particles, are identified using the anti-kT algorithm [17], and missing transverse momentum p_T^{miss} is calculated by summing the transverse momenta of all identified particles, where p^T is the transverse momentum of an object. By determining the primary collision point (vertex) as the one with the highest summed transverse momenta of jets and excluding pileup particles, the analysis ensures that the data reflects genuine interactions, improving the reliability of the signal extraction and background discrimination.

Events of interest are selected using a two-step trigger system [18] focusing on missing transverse momentum and lepton isolation [16] to maintain high efficiency across channels, which include the 0-lepton, 1-lepton and 2-lepton channels.

Events are selected based on the presence of 0, 1, or 2 leptons (muons or electrons) and two identified b jets, formed from b quarks. A b jet is identified by analyzing jet features, such as the distance of particle tracks from the main collision point and the presence of extra particles like soft leptons. A high score from this analysis classifies the jet as a b jet, with lower scores qualifying it under less strict conditions. Muons and electrons from W or Z boson decays are identified using the criteria outlined in [19].

Each channel has specific requirements for the transverse momentum p^T of the leptons and jets. The reconstruction of vector boson decays varies by channel, with missing transverse momentum p_T^{miss} serving as a proxy for the Z boson's transverse momentum in the 0-lepton channel. Higgs boson candidates are formed from jets likely originating from b quarks, assessed using the deepCSV algorithm for b-tagging, which identifies jets from b quarks by analyzing their decay patterns.

Background in the event selection process is significantly reduced by imposing high transverse momentum p^T requirements for the vector boson (V). Events are categorized based on their kinematic properties, such as missing transverse momentum and the transverse momentum of the vector boson, to enhance the separation of signal and background processes. To further reduce background from t quark ($t\bar{t}$) and quantum chromodynamics (QCD) processes, Events are refined using criteria on jet multiplicity and angular separation to improve the selection of signal over background.

After applying the selection criteria, the resolution of m_{jj} for reconstructed Higgs boson decays is about 15%. The dijet invariant mass measures the combined mass of two jets resulting from particle decays, providing critical information about the originating particle. This resolution improves to an average of 10–13% using a DNN [20], a type of machine learning model trained on b jets from simulated t quark events. The DNN considers various jet properties and incorporates the momenta of nearby jets to account for additional radiation emitted during the decay.

Signal regions are defined based on specific dijet mass ranges to enhance VH event selection in each analysis channel. DNNs are trained for each channel using optimized input variables, including m_{jj} and $p_T(V)$, to improve signal-background separation. Control regions normalize major background processes, focusing on $t\bar{t}$ production and W/Z boson production with heavy- or light-flavor jets. For $t\bar{t}$ production, events with characteristics typical of t quark pair production are selected. For W/Z boson production, control regions are divided into those enriched with heavy-flavor (HF) jets and light-flavor (LF) jets, based on jet flavors, the number of additional jets, and kinematic properties.

3 Results

The significance of the observed excess of events is calculated using the profile likelihood asymptotic approximation [21–24], which assesses how likely the observed data is under different hypotheses. For the 2017 data, the observed significance is 3.3σ (standard deviations) above the background-only hypothesis, compared to an expected significance of 3.1σ for the SM Higgs boson, with a measured signal strength (measured production cross section times branching fraction, normalized to the expected SM value) of $\mu = 1.08 \pm 0.34$. The uncertainty in this measurement arises from various sources, including background normalization, sample size, b-tagging efficiency, and modeling of V+jets, all treated as independent nuisance parameters in the fit.

The VZ process, where the Z boson decays into two b quarks ($Z \to b\bar{b}$), validates the detection method for the VH process, where the Higgs boson decays into two b quarks ($H \to b\bar{b}$), as both share the same final state. DNNs are trained on simulated VZ events to identify the signal, with other processes, including VH production, considered background.

The significant excess of events in the combined WZ and ZZ production, with an observed significance of 5.2σ, is measured against the background-only hypothesis, which assumes no signal is present and 5.0σ is expected. The signal strength is $\mu = 1.05 \pm 0.22$.

Measurements of the VH process with $H \to b\bar{b}$ are combined with similar results from CMS data collected at 13 TeV, totaling 35.9 fb^{-1} [19]. The combination of Run 2 data sets produces an observed and expected signal significance of 4.4σ and 4.2σ respectively, with a signal strength of $\mu = 1.06 \pm 0.26$

An additional combination is performed of Run 1 and Run 2 CMS data sets, provides an observed and expected signal significance of 4.8σ and 4.9σ, respectively, and reports a signal strength of $\mu = 1.01 \pm 0.22$

Fitting the dijet mass m_{jj} distribution (see Figure 2), although not as sensitive as the DNN score, gives a clearer view of the Higgs signal. Much like the VZ analysis, the study [8] focuses on the signal region between 60 and 160 GeV, using data from 2016 and 2017. Events are grouped into bins based on their signal-to-background ratio (S/B), which measures the relative strength of the signal (S) compared to the background (B), and a higher S/B ratio indicates a clearer distinction between the signal and background, enhancing the likelihood of accurately identifying the signal amidst other processes. The combined m_{jj} distributions are weighted by the ratio $S/(S+B)$, applying a factor to each entry based on the signal-to-background ratio. This adjustment, using fitted normalizations, enhances the signal by reducing background and provides a clearer representation of the data.

Combined CMS measurements from multiple Higgs boson production processes, including VH, gluon fusion, vector boson fusion, and top-quark associated production, at collision energies of 7, 8, and 13 TeV is performed. Most uncertainties are treated as uncorrelated, except for the jet energy scale uncertainty and theory uncertainties, which are correlated across processes. The combination shows an observed signal significance of 5.6σ (expected 5.5σ) and a measured signal strength of $\mu = 1.04 \pm 0.20$. Figure 3 provides a summary of all results.

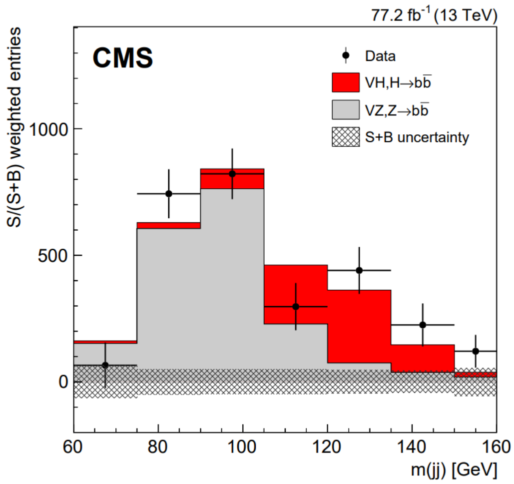

Figure 2: The dijet invariant mass distribution, weighted by $S/(S+B)$, is shown for 2016 and 2017 data. Weights come from a fit to m_{jj} displaying data with the fitted VH signal (red) and VZ background (grey), after subtracting other backgrounds. Error bars show the pre-subtraction 1σ statistical uncertainty, and the grey hatching indicates the 1σ total uncertainty. Reproduced from [8].

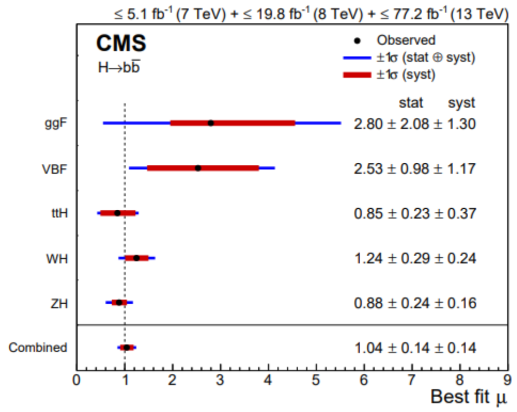

Figure 3: The best-fit value for the $H \to b\bar{b}$ signal strength is shown with its 1σ uncertainties—systematic (in red) and total (in blue)—for each of the five production modes, as well as the combined result. The dashed vertical line represents the SM prediction. All results come from a single fit using data from all analyses, assuming a Higgs boson mass of 125.09 GeV. Reproduced from [8].

4 Summary and discussion

This study presents the measurement of the SM Higgs boson decaying to b quarks. By combining all CMS measurements of the VH process ($H \to b\bar{b}$) from proton-proton collisions at CM energies of 7, 8, and 13 TeV, an observed significance of 4.8σ and an expected significance of 4.9σ are found at a Higgs boson mass of 125.09 GeV, with a signal strength of $\mu = 1.01 \pm 0.22$. When including previous CMS measurements of $H \to b\bar{b}$ from other production methods, the observed significance rises to 5.6σ and the expected significance to 5.5σ, with a signal strength of $\mu = 1.04 \pm 0.20$. This confirms the observation of the $H \to b\bar{b}$ decay by the CMS Collaboration. The results reinforce the SM predictions regarding the Higgs boson and its interactions with fermions, highlighting the observed decay as crucial for probing the Higgs mechanism and exploring potential new physics beyond the SM, with future studies aimed at enhancing precision measurements and investigating the implications of these findings on theoretical models.

References

[1] 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. B716 (2012) 1-29.

[2] Higgs, P. W., *Broken Symmetries and the Masses of Gauge Bosons* , American Physical Society, Vol. 13, Iss. 16 — 19, October 1964.

[3] CERN, *The Large Hadron Collider*, https://home.cern/science/accelerators/large-hadron-collider

[4] CERN, *CMS Detector*, https://cms.cern/detector

[5] Q. Bi, K. Chai, J. Gao, Y. Liu, and H. Zhang., *Investigating Bottom-Quark Yukawa Interaction at Higgs Factory* , Chinese Physics C, September 2020

[6] Griffiths, D., *Introduction to elementary particles*, 2nd edition, Wiley-VCH, October 2008

[7] nature, *Long-sought decay of the Higgs boson seen*, https://www.nature.com/articles/d41586-018-07405-x

[8] CMS Collaboration, *Observation of Higgs boson decay to bottom quarks*, Phys. Rev. Lett. 121, 121801 (2018), https://doi.org/10.48550/arXiv.1808.08242

[9] CMS Collaboration, *Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair*, Phys. Rev. Lett. 120 (2018) 071802, doi:10.1103/PhysRevLett.120.071802, arXiv:1709.05543.

[10] CMS Collaboration, *Search for the standard model Higgs boson produced through vector boson fusion and decaying to bb*, Phys. Rev. D 92 (2015) 032008, doi:10.1103/PhysRevD.92.032008, arXiv:1506.01010.

[11] Cottingham, W. N., & Greenwood, D. A. (2007), *An Introduction to the Standard Model of Particle Physics*, (2nd ed.). Cambridge University Press, https://doi.org/10.1017/CBO9780511791406

[12] ATLAS Collaboration, *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*, JHEP 12 (2017) 24, https://doi.org/10.1007/JHEP12(2017)024

[13] ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{7}$ and 8 TeV*, JHEP 08 (2016) 045, https://doi.org/10.48550/arXiv.1606.02266

[14] CMS Collaboration, *Search for the standard model Higgs boson produced through vector boson fusion and decaying to bb*, Phys. Rev. D 92 (2015) 032008, doi:10.1103/PhysRevD.92.032008, arXiv:1506.01010.

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

[16] CMS Collaboration, *Particle-flow reconstruction and global event description with the CMS detector*, JINST 12 (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[17] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-kT jet clustering algorithm*, JHEP 04 (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[18] CMS Collaboration, *The CMS trigger system*, JINST 12 (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[19] CMS Collaboration, *Evidence for the Higgs boson decay to a bottom quark-antiquark pair*, Phys. Lett. B 780 (2018) 501, https://doi.org/10.1016/j.physletb.2018.02.050

[20] CMS Collaboration, *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*, JINST 13 (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[21] 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.

[22] A. L. Read, *Presentation of search results: The CLs technique*, J. Phys. G 28 (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[23] ATLAS and CMS Collaborations, LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011*, ATL-PHYS-PUB 011-11/CMS NOTE 2011-005, 2011.

[24] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C 71 (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].

[25] 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, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[26] CMS Collaboration, *Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV*, JHEP 06 (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[27] D. de Florian et al., *Handbook of LHC Higgs cross sections: 4. deciphering the nature of the Higgs sector*, CERN Report CERN-2017-002-M, 2016, https://doi.org/10.23731/CYRM-2017-002

[28] CMS Collaboration, *CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s}$ = 13 TeV*, CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.

[29] CMS Collaboration, *Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$*, Phys. Rev. D 92 (2015) 032008, https://doi.org/10.1103/PhysRevD.92.032008

[30] CMS Collaboration, *Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair*, Phys. Rev. Lett. 120 (2018) 071802, https://doi.org/10.1103/PhysRevLett.120.071802

[31] CMS Collaboration, *Observation of $t\bar{t}H$ production*, Phys. Rev. Lett. 120 (2018) 231801, https://doi.org/10.1103/PhysRevLett.120.231801

[32] CMS Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, Eur. Phys. J. C 75 (2015) 212, https://doi.org/10.48550/arXiv.1412.8662