CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2017/08/10

Evidence for the decay of the Higgs Boson to Bottom Quarks

The CMS Collaboration

Abstract

A search for the standard model (SM) Higgs boson (H) decaying to $b\bar{b}$ when produced in association with a weak vector boson (V) is reported for the following processes: $Z(\nu\nu)H$, $W(\mu\nu)H$, $W(e\nu)H$, $Z(\mu\mu)H$, and $Z(ee)H$. The search is performed in data samples corresponding to an integrated luminosity of 35.9 fb^{-1} at $\sqrt{s} = 13$ TeV recorded by the CMS experiment at the LHC during Run 2 in 2016. An excess of events is observed in data compared to the expectation in the absence of a $H \to b\bar{b}$ signal. The significance of this excess is 3.3 standard deviations, where the expectation from SM Higgs boson production is 2.8. The signal strength corresponding to this excess, relative to that of the SM Higgs boson production is 1.2 ± 0.4. This result is combined with the one from the search for the same processes performed by the CMS experiment in Run 1 of the LHC (using proton-proton collisions at $\sqrt{s} = 7$ and $\sqrt{s} = 8$ TeV with data samples corresponding to luminosities of up to 5.1 fb^{-1} and 18.9 fb^{-1}, respectively). The observed combined signal significance is 3.8 standard deviations, where 3.8 are expected from a SM signal. The corresponding signal strength, relative to that of the SM Higgs boson, is $1.06^{+0.31}_{-0.29}$.

1 Introduction

In 2012, at the Large Hadron Collider (LHC), the ATLAS and CMS collaborations reported the discovery of a new boson [1–3]. With the datasets collected in proton-proton collisions at $\sqrt{s} = 7\,\text{TeV}$ and 8 TeV, corresponding to total integrated luminosities of up to approximately $25\,\text{fb}^{-1}$, significant signals have been observed in channels where the boson decays into $\gamma\gamma$, ZZ, WW, or $\tau\tau$ [4–12]. The measured production and decay rates and spin-parity properties [13–17] of this boson are compatible with those of the standard model (SM) Higgs boson (H) [18–23] and consistent with the measured mass of $m_{\text{H}} = 125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst.})$ GeV [24].

The dataset analyzed in this document has also been used by CMS to observe the Higgs boson decaying to $\tau\tau$ [25] and to determine its mass in the ZZ channel to be $m_{\text{H}} = 125.26 \pm 0.20(\text{stat.}) \pm 0.08(\text{syst.})$ GeV [26]. The latter result is the best precision achieved to date.

The $\text{H} \to \text{b}\bar{\text{b}}$ branching fraction for that mass is approximately 58% [27], by far the largest in the SM. The $\text{H} \to \text{b}\bar{\text{b}}$ decay tests directly the Higgs boson coupling to fermions, and more specifically to down-type quarks, and has not yet been established experimentally with a sufficiently large statistical significance. This search is notable because it gives the best constraints of the Higgs boson couplings present in both its production (H-W, H-Z) and its decay (H-b). Moreover, for a consistent picture of the Higgs boson at 125 GeV, an observation in this channel is necessary and is important to solidify the Higgs boson as the source of mass generation in the fermion sector of the SM [28, 29].

In their final result on the search for the SM Higgs boson at the Tevatron $\text{p}\bar{\text{p}}$ collider, the CDF and D0 collaborations reported evidence for an excess of events in the 115–140 GeV mass range [30]. For masses below 130 GeV, the channels in which the Higgs boson is produced in association with a weak vector boson and decaying to $\text{b}\bar{\text{b}}$ [31] dominate the search sensitivity. The measured local significance of this excess, for the combination of the search results from both collaborations, is 3.0 standard deviations at $m_{\text{H}} = 125\,\text{GeV}$, with the expected value being 1.9 standard deviations.

At the LHC, the search for $\text{H} \to \text{b}\bar{\text{b}}$ has been performed by studying different production channels, including the production of the Higgs boson in association with a top-quark pair [32–35], through vector boson fusion [36, 37], and in association with a weak vector boson (VH production) [17, 38, 39]. VH production has the greatest expected sensitivity. The search in this channel by the ATLAS Collaboration [17, 38], using data samples corresponding to an integrated luminosity of up to $4.7\,\text{fb}^{-1}$ at $\sqrt{s} = 7\,\text{TeV}$ and up to $20.3\,\text{fb}^{-1}$ at $\sqrt{s} = 8\,\text{TeV}$, resulted in an observed (expected) excess of events from the background-only hypothesis corresponding to a local significance of 1.7 (2.7) standard deviations. The ratio of the measured signal yield to the expectation from the SM (denoted as signal strength, μ) was found to be $\mu = 0.62 \pm 0.37$. The corresponding search by the CMS Collaboration [17, 37, 39] in data samples corresponding to integrated luminosities of up to $5.1\,\text{fb}^{-1}$ at $\sqrt{s} = 7\,\text{TeV}$ and up to $18.9\,\text{fb}^{-1}$ at 8 TeV resulted in an excess of events with a local significance of 2.0 standard deviations, consistent with the 2.5 standard deviations expected from the production of the SM Higgs boson. The corresponding signal strength was $\mu = 0.9 \pm 0.4$. These results have been evaluated for a Higgs boson mass of 125.09 GeV.

The ATLAS collaboration has recently presented preliminary results for the search for $\text{H} \to \text{b}\bar{\text{b}}$ produced in association with a weak boson with approximately $36\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ data [40]. An excess of events above background with a 3.5σ significance is observed, corresponding to a signal strength of $\mu = 1.20^{+0.42}_{-0.36}$. In combination with the results from the same search in Run 1, the total significance is 3.6σ and the signal strength is $\mu = 0.90^{+0.28}_{-0.26}$.

This document reports on the search at the Compact Muon Solenoid (CMS) experiment for the decay of the SM Higgs boson to bottom quarks, $H \to b\bar{b}$, when produced through the $pp \to VH$ process, where V is either a W or a Z boson. This search is performed with data samples from the LHC Run 2, recorded during 2016, corresponding to an integrated luminosity of $35.9\,fb^{-1}$ at $\sqrt{s} = 13\,TeV$ from pp collisions at the LHC. Many of the experimental techniques used are similar to those documented in [39] but are described here for completeness.

The following five processes are considered in the search: $Z(\nu\nu)H$, $W(\mu\nu)H$, $W(e\nu)H$, $Z(\mu\mu)H$, and $Z(ee)H$, all with the Higgs boson decaying to $b\bar{b}$. The final states that predominantly correspond to these processes are characterized by the number of leptons required in the event selection, and are referred to as the 0-lepton, 1-lepton, and 2-lepton channels.

Throughout this document the term "lepton" (symbol ℓ) refers solely to muons and electrons, but not to taus. The leptonic decays of taus in WH processes are implicitly included in the $W(\mu\nu)H$ and $W(e\nu)H$ channels. Background processes originate from the production of W and Z bosons in association with jets: W+jets and Z+jets (from gluons and from light- or heavy-flavor quarks), from singly and pair-produced top quarks ($t\bar{t}$), from diboson production (VV), and from quantum chromodynamics (QCD) multijet events.

Simulated samples of signal and background events are used to optimize the extraction of signal events from data. For each channel, several control regions, each enriched in events from individual background processes, are selected to study the agreement between simulated samples and data. These regions test the accuracy of the simulated samples' modeling for the variables relevant to the analysis. Also for each channel, a signal region enriched in VH events is selected, and a joint fit to the shape and normalization of specific distributions for the signal and control regions for all channels combined, is used to determine the results of the analysis. The distribution used in the signal region is the output of a boosted-decision-tree (BDT) event discriminant [41, 42] that helps isolate signal from background. For the control regions, the distribution used is the value of the variable that identifies jets originating from b quarks for the jet with the lowest such value among those used to reconstruct the $H \to b\bar{b}$ decay. This variable helps to distinguish between the different background processes because there are significant differences in this variable's shape between backgrounds with one and two b jets. Details on how these variables are defined are presented in Section 5.

The result of the fitting procedure is used to measure the presence of the Higgs boson signal over the expectation from SM background processes alone. The significance of any excess of events, and the corresponding event yield, is compared with the expectation from a SM Higgs boson signal.

To validate the analysis procedure, the same methodology is used to extract a signal for the VZ process, with $Z \to b\bar{b}$ which has a nearly identical final state to VH with $H \to b\bar{b}$ but with a production cross section a few times larger.

2 Detector and simulated samples

A detailed description of the CMS detector can be found elsewhere [43]. The momenta of charged particles are measured using a silicon pixel and strip tracker that covers the pseudorapidity range $|\eta| < 2.5$ and is immersed in a 3.8 T axial magnetic field. The pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, where θ is the polar angle of the trajectory of a particle with respect to the direction of the counterclockwise proton beam. Surrounding the tracker are a crystal electromagnetic calorimeter (ECAL) and a brass/scintillator hadron calorimeter (HCAL), both

used to measure particle energy deposits and both consisting of a barrel assembly and two endcaps. The ECAL and HCAL extend to a pseudorapidity range of $|\eta| < 3.0$. A steel/quartz-fiber Cherenkov forward detector extends the calorimetric coverage to $|\eta| < 5.0$. The outermost component of the CMS detector is the muon system, consisting of gas-ionization detectors placed in the steel return yoke of the magnet to measure the momenta of muons traversing through the detector. The two-level CMS trigger system selects events of interest for permanent storage. The first trigger level, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events in less than 3.2 μs. The high-level trigger software algorithms, executed on a farm of commercial processors, further reduce the event rate using information from all detector subsystems. The variable $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is used to measure the separation between reconstructed objects in the detector, where ϕ is the angle (in radians) of the trajectory of the object in the plane transverse to the direction of the proton beams.

Samples of simulated signal and background events are produced using the Monte Carlo (MC) event generators listed below. The CMS detector response is modeled with GEANT4 [44]. The Higgs boson signal samples, with $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, are produced at next-to-leading order (NLO) using the POWHEG+MiNLO [45, 46] event generator. The three production processes considered as signal are associated production with vector bosons, gluon fusion, and vector-boson fusion. The MG5AMC@NLO [47] generator is used at NLO with FxFx merging [48] for the diboson samples. The same generator is used at leading-order (LO) accuracy with MLM matching [49] for the W+jets and Z+jets in inclusive and b-quark enriched configurations, as well as the QCD multijet sample. The $\mathrm{t\bar{t}}$ [50] production process, as well as the single-top-quark sample for the t-channel [51] are produced with POWHEG V2. The single-top-quark sample for the tW- [52], and s-channel [53] are instead produced with POWHEG V1. The production cross sections for the signal samples are rescaled to next-to-next-to-leading-order (NNLO) QCD + NLO electroweak accuracy combining the VHNNLO [54–56], VH@NNLO [57, 58] and HAWK [59] generators as described in the documentation produced by the LHC Working Group on Higgs boson cross sections [60], and applied as a function of the vector boson transverse momentum (p_{T}). The production cross sections for the diboson and $\mathrm{t\bar{t}}$ samples are rescaled to NNLO with the MCFM generator [61], while the cross sections for the W+jets and Z+jets samples are rescaled to NNLO cross sections calculated using the FEWZ program [62–64]. The parton distribution functions (PDF) set used to produce the NLO samples is the NLO NNPDF3.0 set [65], while the leading-order NNPDF3.0 set is used for the LO samples. For parton showering and hadronization the POWHEG and MG5AMC@NLO samples are interfaced with PYTHIA8.212 [66]. The PYTHIA8 parameters for the underlying event description correspond to the CUETP9M1 set derived in [67] based on the work described in [68].

During the 2016 data-taking period the LHC instantaneous luminosity reached approximately $1.5 \times 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$ and the average number of pp interactions per bunch crossing was approximately twenty four. The simulated samples include these additional pp interactions, denoted as pileup interactions (or pileup), that overlap with the event of interest in the same bunch crossing.

3 Triggers

Several triggers are used to collect events with final-state objects consistent with the signal processes in the five channels under consideration.

For the 0-lepton channel, the quantities used in the trigger are derived from the reconstructed objects identified by an online particle-flow (PF) algorithm [69] that combines the information

from all CMS subsystems to identify and reconstruct individual particles emerging from the proton-proton collisions: charged hadrons, neutral hadrons, photons, muons, and electrons. The main trigger used requires that both the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, and the MHT in the event be above a threshold of 110 GeV. $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined online as the magnitude of the negative vector sum of the transverse momenta of all reconstructed objects identified by this PF algorithm, while MHT is defined as the magnitude of the vector sum of the transverse momenta of all reconstructed jets (with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 5.2$) identified by the same algorithm. For $\mathrm{Z}(\nu\nu)\mathrm{H}$ events with a $E_{\mathrm{T}}^{\mathrm{miss}} > 170$ GeV, evaluated offline, the trigger efficiency is approximately 92%, and near 100% efficient at 200 GeV.

For the 1-lepton channels, single-lepton triggers are used. The muon trigger p_{T} threshold is set at 18 GeV and the electron p_{T} threshold is set at 23 GeV. For the 2-lepton channels, dilepton triggers are used. The muon p_{T} thresholds are 24 and 8 GeV, and the electron p_{T} thresholds are 27 and 12 GeV. All leptons in these triggers are required to pass tight lepton-identification criteria. In addition, to maintain an acceptable trigger rate, and to be consistent with what is expected from signal events, leptons are also required to be isolated from other tracks and calorimeter energy deposits. For $\mathrm{W}(\mu\nu)\mathrm{H}$ events that pass all offline requirements described in Section 5, the single-muon trigger efficiency is ≈95%. The corresponding efficiency for $\mathrm{W}(\mathrm{e}\nu)\mathrm{H}$ events recorded with the single-electron trigger is ≈90%. For $\mathrm{Z}(\ell\ell)\mathrm{H}$ signal events the dilepton triggers are nearly 100% efficient.

4 Event reconstruction

The characterization of VH events in the channels studied here requires the reconstruction of the following objects, all originating from a common interaction vertex: electrons, muons, neutrinos (reconstructed as $E_{\mathrm{T}}^{\mathrm{miss}}$), and jets –including those that originate from the hadronization of b quarks, referred to as "b jets."

The reconstructed vertex with the largest value of summed charged physics-object p_{T}^2 is taken to be the primary pp interaction vertex. The physics objects are the objects returned by a jet finding algorithm [70, 71] applied to all charged tracks associated with the vertex, plus the corresponding associated missing transverse momentum. This vertex is used as the reference vertex for all relevant objects in the event, which are reconstructed with the PF algorithm. The pileup interactions affect jet momentum reconstruction, missing transverse energy reconstruction, lepton isolation, and b-tagging efficiencies. To mitigate these effects, all charged hadrons that do not originate from the primary interaction are removed from consideration in the event. In addition, the average neutral energy density from pileup interactions is evaluated from PF objects and subtracted from the reconstructed jets in the event and from the summed energy in the isolation criteria used for leptons [72]. These pileup-mitigation procedures are applied on an object-by-object basis.

Muons are reconstructed using two algorithms [73]: one in which tracks in the silicon tracker are matched to hits in the muon detectors, and another in which a track fit is performed using hits in the silicon tracker and in the muon systems. In the latter, the muon is seeded by hits in the muon systems. The muon candidates used in the analysis are required to be successfully reconstructed by both algorithms. Further identification criteria are imposed on the muon candidates to reduce the fraction of tracks misidentified as muons. These include the number of hits in the tracker and in the muon systems, the fit quality of the global muon track, and its consistency with the primary vertex. Muon candidates are required to be in the $|\eta| < 2.4$ region.

Electron reconstruction requires the matching of a set of ECAL clusters, named super-clusters (SC), to a track in the silicon tracker. Electron identification [74] relies on a multivariate technique that combines observables sensitive to the amount of bremsstrahlung along the electron trajectory, such as the geometrical matching and momentum consistency between the electron trajectory and the associated calorimeter clusters, as well as various shower-shape observables in the calorimeters. Additional requirements are imposed to remove electrons that originate from photon conversions. Electrons are required to be in the pseudorapidity range $|\eta| < 2.5$, excluding candidates for which the SC lies in the $1.444 < |\eta_{\mathrm{SC}}| < 1.566$ transition region between the ECAL barrel and endcap, where electron reconstruction is suboptimal.

Charged leptons from W and Z boson decays are expected to be isolated from other activity in the event. For each lepton candidate, a cone is constructed around the track direction at the event vertex. The scalar sum of the transverse momentum of each reconstructed particle compatible with the primary vertex and contained within the cone is calculated, excluding the contribution from the lepton candidate itself. This sum is called isolation. In the presence of pileup, isolation is contaminated with particles from the other interactions. A quantity proportional to the pileup is used to correct on average the isolation to mitigate reductions in signal efficiency at larger values of pileup. In the 1-lepton channel, if the corrected isolation sum exceeds 6% of the lepton candidate p_{T}, the lepton is rejected. In the 2-lepton channel, the threshold is looser; the isolation of each candidate can be up to 20% (15%) of the muon (electron) p_{T}. Including the isolation requirement, the total efficiency to reconstruct muons is in the range of 85-100%, depending on p_{T} and η. The corresponding efficiency for electrons is in the range of 40-90%.

Jets are reconstructed from PF objects using the anti-k_{T} clustering algorithm [70], with a distance parameter of 0.4, as implemented in the FASTJET package [71, 75]. Each jet is required to lie within $|\eta| < 2.4$, to have at least two tracks associated with it, and to have electromagnetic and hadronic energy fractions of at least 1%. The last requirement removes jets originating from instrumental effects. Jet energy corrections are applied as a function of pseudorapidity and transverse momentum of the jet [76]. The missing transverse energy vector is calculated offline as the negative of the vectorial sum of transverse momenta of all PF objects identified in the event, and the magnitude of this vector is referred to as $E_{\mathrm{T}}^{\mathrm{miss}}$ in the rest of this article.

The identification of b jets is performed using the CMVA b-tagging algorithm [77]. This algorithm combines, in a likelihood discriminant, information within jets that helps differentiate between b jets and jets originating from light quarks, gluons, or charm quarks. This information includes track impact parameters, secondary vertices, and information related to low p_{T} leptons if contained within a jet. The output of this discriminant has continuous values between -1.0 and 1.0. A jet with a CMVA value above a certain threshold is said to be "tagged". The efficiency to tag b jets and the rate of misidentification of non-b jets depend on the threshold chosen, and are typically parameterized as a function of the p_{T} and η of the jets. These performance measurements are obtained directly from data in samples that can be enriched in b jets, such as $\mathrm{t\bar{t}}$ and multijet events (where, for example, requiring the presence of a muon in the jets enhances the heavy-flavor content of the events). Several thresholds for the CMVA output discriminant are used in this analysis. Depending on the threshold used, the efficiencies to tag jets originating from b quarks, c quarks, and light quarks or gluons are in the 50–75%, 5–25%, and 0.15–3.0% ranges, respectively. Three working points are used in this analysis. The loose (tight) has the highest (lowest) efficiency and most (least) contamination in the ranges previously mentioned. There is also an intermediate working point used denoted medium.

To measure the additional hadronic activity excluding the vector boson and Higgs boson decay

products, only reconstructed charged tracks can be used. This is done to measure the hadronic activity associated with the main primary vertex (PV) as defined above. A collection of "additional tracks" is assembled using reconstructed tracks that: (i) satisfy the high purity quality requirements defined in Ref. [78] and $p_{\mathrm{T}} > 300\,\mathrm{MeV}$; (ii) are not associated with the vector boson, nor with the selected b jets in the event; (iii) have a minimum longitudinal impact parameter, $|d_z(\mathrm{PV})|$, with respect to the main PV, rather than to other pileup interaction vertices; (iv) satisfy $|d_z(\mathrm{PV})| < 2\,\mathrm{mm}$; and (v) are not in the region between the two selected b-tagged jets. This is defined as an ellipse in the η-ϕ plane, centered on the midpoint between the two jets, with major axis of length $\Delta R(\mathrm{bb}) + 1$, where $\Delta R(\mathrm{bb}) = \sqrt{(\Delta\eta_{\mathrm{bb}})^2 + (\Delta\phi_{\mathrm{bb}})^2}$, oriented along the direction connecting the two b jets, and with minor axis of length 1. The additional tracks are then clustered into "soft track jets" using the anti-k_{T} clustering algorithm with a distance parameter of 0.4. The use of track jets represents a clean and validated method [79] to reconstruct the hadronization of partons with very low energies down to a few GeV [80]; an extensive study of the soft track jet activity can be found in Refs. [81, 82].

Events from data and from the simulated samples are required to satisfy the same trigger and event reconstruction requirements. Corrections that account for the differences in the performance of these algorithms between data and simulated samples are computed from data and used in the analysis.

5 Event selection

A signal region enriched in VH events is determined separately for each channel. Simulated events in this region are used to train a BDT event discriminant to help differentiate between signal and background events. Also for each channel, different control regions, each enriched in events from individual background processes, are selected in order to both study the agreement between simulated samples and data and to provide a distribution that is fit in combination with the output distribution of the signal region BDT discriminant to extract a potential $\mathrm{H} \to \mathrm{b\bar{b}}$ signal.

Background processes to VH production with $\mathrm{H} \to \mathrm{b\bar{b}}$ are the production of vector bosons in association with one or more jets (V+jets), $\mathrm{t\bar{t}}$ production, single-top-quark production, diboson production, and QCD multijet production. These processes have production cross sections that are several orders of magnitude larger than that of the Higgs boson production, with the exception of diboson production whose production cross section for the VZ process, where $\mathrm{Z} \to \mathrm{b\bar{b}}$, is only a few times larger than the VH production cross section. Given the nearly identical final state, this process provides a benchmark against which the Higgs boson search strategy can be tested. The results of this test are discussed in Section 7.2.

Below we describe the selection criteria used to define the signal regions and the variables used to construct the BDT discriminant. Also described are the criteria used to select appropriate background-specific control regions and the corresponding distributions used in the signal-extraction fits.

5.1 Signal regions

Signal events are characterized by the presence of a vector boson recoiling from two b jets with an invariant mass near 125 GeV. The event selection for the signal region therefore relies on the reconstruction of the decay of the Higgs boson into two b-tagged jets and on the reconstruction of the leptonic decay modes of the vector boson. To further exclude background events, several other requirements are imposed on each channel. The signal region requirements are listed in

Table 1 and described below.

The reconstruction of the $H \to b\bar{b}$ decay is based on the selection of the pair of jets in the event for which the values of the output of the CMVA discriminant are the two highest amongst all jets in the event. Both jets are required to be central (with $|\eta| < 2.4$), have a p_T above a minimum threshold that is different for the different channels, and pass some standard requirements to remove jets from pileup [83]. The background from V+jets and diboson production is reduced significantly when the b-tagging requirements listed in Table 1 are applied. Moreover, processes where the two jets originate from genuine b quarks dominate the final selected data sample. In what follows the additional selection criteria, aside from the reconstruction of the $H \to b\bar{b}$ decay, are described for each channel.

5.1.1 0-lepton channel

This channel targets mainly $Z(\nu\nu)H$ events in which the E_T^{miss} is interpreted as the transverse momentum of the Z boson in the $Z \to \nu\bar{\nu}$ decay. In order to overcome large QCD multijet backgrounds, a relatively high threshold of $E_T^{miss} > 170\,GeV$ is required. The QCD multijet background is further reduced to negligible levels in this channel when requiring that the E_T^{miss} does not originate from the direction of (mismeasured) jets. To that end if there is a central ($|\eta| < 2.5$) jet with $p_T > 25\,GeV$, whose azimuthal angle is within 0.5 radian of the E_T^{miss} direction, the event is rejected. Moreover, this requirement to reject E_T^{miss} produced by mismeasured jets is further enforced by applying the same selection substituting E_T^{miss} by an alternative missing transverse energy reconstruction, named $E_T^{miss}{}_{trk}$, obtained by using only charged tracks with $p_T > 0.5\,GeV$ and $|\eta| < 2.5$. To reduce background events from $t\bar{t}$ and WZ production channels, events with any additional isolated leptons with $p_T > 20\,GeV$ are rejected. The sum of these additional leptons is denoted, N_{al}.

5.1.2 1-lepton channel

This channel targets mainly $W(\ell\nu)H$ events in which candidate $W \to \ell\nu$ decays are identified by the presence of one isolated lepton as well as missing transverse energy (implicitly required in the $p_T(V)$ selection criteria mentioned below, where $p_T(V)$ is calculated from the vectorial sum of the E_T^{miss} and the lepton p_T). Muons (electrons) are required to have $p_T > 25(30)\,GeV$. It is also required that the azimuthal angle between the E_T^{miss} direction and the lepton be less than 2.0. The lepton isolation for either flavor of lepton is required to be smaller than 6% of the lepton p_T. These requirements significantly reduce possible contamination from QCD multijet production. With the same motivation as in the 0-lepton channel, events with any additional isolated leptons are rejected. To substantially reject $t\bar{t}$ events, the number of additional central jets with $p_T > 25\,GeV$, N_{aj}, is allowed to be at most one.

5.1.3 2-lepton channel

This channel targets candidate $Z \to \ell\ell$ decays which are reconstructed by combining isolated, oppositely-charged pairs of electrons or muons and requiring the dilepton invariant mass to satisfy $75 < M(\ell\ell) < 105\,GeV$. The p_T for each lepton is required to be $>20\,GeV$. Isolation requirements are relaxed in this channel as the QCD multijet background is practically eliminated after requiring compatibility with the Z boson mass.

5.1.4 Vector boson p_T requirements and $H \to b\bar{b}$ mass reconstruction

Background events are substantially reduced by requiring significant large transverse momentum of the reconstructed vector boson (p_T boost) or of the Higgs boson candidate [84]. In this

kinematic region the V and H bosons recoil from each other with a large azimuthal opening angle, $\Delta\phi(\mathrm{V},\mathrm{H})$, between them. For each channel, different $p_{\mathrm{T}}(\mathrm{V})$ regions are selected. Because of different signal and background content, each $p_{\mathrm{T}}(\mathrm{V})$ region has different sensitivity and the analysis is performed separately in each region. For the 0-lepton channel, a single boost region requiring $E_{\mathrm{T}}^{\mathrm{miss}} > 170\,\mathrm{GeV}$ is studied. The 1-lepton channels also study a single region, with $p_{\mathrm{T}}(\mathrm{V}) > 100\,\mathrm{GeV}$. The 2-lepton channels consider two regions: a low- and a high-p_{T} regions defined by $50 < p_{\mathrm{T}}(\mathrm{V}) < 150\,\mathrm{GeV}$ and $p_{\mathrm{T}}(\mathrm{V}) > 150\,\mathrm{GeV}$. The results from these two regions are combined. In the rest of the text the term "boost region" is used to refer to these $p_{\mathrm{T}}(\mathrm{V})$ regions.

After all event selection criteria described in this section are applied, the dijet invariant-mass resolution of the two b jets from the Higgs boson decay is approximately 10%, depending on the p_{T} of the reconstructed Higgs boson, with a few percent shift on the value of the mass peak, relative to 125 GeV. The Higgs boson mass resolution is further improved by applying multivariate regression techniques similar to those used at the CDF experiment [85] and used for several Run 1 $\mathrm{H} \to \mathrm{b\bar{b}}$ analyses by ATLAS and CMS [38, 39]. The regression estimates a correction that is applied after the aforementioned jet-energy corrections. It is computed for individual b jets in an attempt to improve the accuracy of the measured energy with respect to the b-quark energy. To this end, a specialized BDT is trained on b jets from simulated $\mathrm{t\bar{t}}$ events with inputs that include detailed jet structure information, which differs in jets from b quarks from that of jets from light-flavor quarks or gluons. These inputs include variables related to several properties of the secondary vertex (when reconstructed), information about tracks, jet constituents, and other variables related to the energy reconstruction of the jet. Because of semileptonic b-hadron decays, jets from b quarks contain, on average, more leptons and a larger fraction of missing energy than jets from light quarks or gluons. Therefore, in the cases where a low-p_{T} lepton is found in the jet or in its vicinity, the following variables are also included in the BDT regression: the p_{T} of the lepton, the ΔR distance between the lepton and the jet directions, and the momentum of the lepton transverse to the jet direction.

The average improvement on the mass resolution, measured on simulated signal samples, when the corrected jet energies are used is approximately 15%, depending on the p_{T} of the reconstructed Higgs boson. The usage of the regression technique increases the sensitivity of the analysis by approximately 10%. The performance of these corrections is shown in Fig. 1 for simulated samples of $\mathrm{Z}(\ell\ell)\mathrm{H}(\mathrm{bb})$ events where the improvement in the reconstructed mass resolution is approximately 15%. The validation of the regression technique in data is done with samples of $\mathrm{Z} \to \ell\ell$ events with two b-tagged jets and in $\mathrm{t\bar{t}}$-enriched samples in the lepton+jets final state. In both cases, after the jets are corrected, the RMS values of the distributions decrease, which is interpreted as improvement in jet resolution, and the central value of the distribution moves toward the expected values of one and the top quark mass, respectively. The distributions for data and the simulated samples are in very good agreement before and after the regression correction is applied. Importantly, the reconstructed dijet invariant mass distributions for background processes do not develop a peak structure when the regression correction is applied to the jets in the event.

To help separate signal from background, an event BDT discriminant is trained using simulated samples for signal and all background processes. The set of event input variables used, listed in Table 2, is chosen by iterative optimization from a larger number of potentially discriminating variables. Among the most discriminant variables for all channels are the dijet invariant mass distribution ($M(\mathrm{jj})$), the number of additional jets (N_{aj}), the value of CMVA for the Higgs boson daughter with the second largest CMVA value ($\mathrm{CMVA}_{\mathrm{min}}$), and the distance between Higgs boson daughters ($\Delta R(\mathrm{jj})$). As described in Section 7, the shape of the output distribution of this

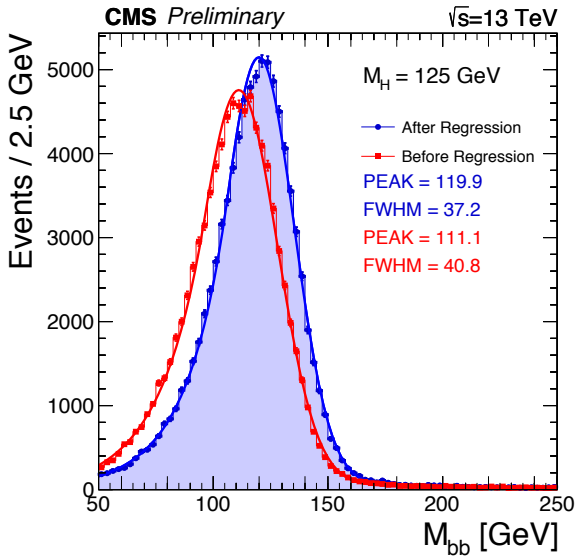

Figure 1: Dijet invariant mass distributions for simulated samples of $Z(\ell\ell)H(bb)$ events ($m_H = 125\,GeV$), before (red) and after (blue) the energy correction from the regression procedure is applied. A combination of a Bernstein polynominal and a Crystal-Ball function is used to fit the distribution. The fitted mean and width of the core of the distribution are displayed on the figure.

event BDT discriminant is used to search for events resulting from Higgs boson production.

5.2 Background control regions

Several control regions, each enriched with one of the main background processes, are selected in data and used to validate the simulated samples' modeling of the distributions of variables most relevant to the analysis —including those used as input to the event BDT discriminants– and to determine the normalization of the main background processes. Separate control regions are defined for $t\bar{t}$ production and for the production of W and Z bosons in association with either predominantly heavy-flavor (HF) or light-flavor (LF) jets. Tables 3–5 list the selection criteria used to define these control regions for the 0-lepton, 1-lepton, and 2-lepton channels, respectively.

Different background processes feature different b jet compositions, e.g. two real b jets for $t\bar{t}$ and V+bb, one real b jet for V+b, no real b jet for V+light. This feature, together with different jet kinematic distributions and a careful choice of the selection criteria used to define the control regions, provides quite different spectra in $CMVA_{min}$. While some control regions are very pure in targeted background, others contain several backgrounds. The distinct shapes in $CMVA_{min}$can be used to extract the normalization scale factors of the various simulated background samples.

All simulated processes are simultaneously fit to data in control regions using a binned likelihood function and allowing the scale factor of each process to float freely. In these fits the shape and normalization of the $CMVA_{min}$ distribution for each background component is allowed to vary within the systematic and statistical uncertainties described in Section 6. These uncertainties are treated as independent nuisance parameters.

The simulated samples for the V+jets processes are split into sub-processes according to how many of the Monte Carlo generator-level jets (with $p_T > 20$ GeV and $|\eta| < 2.4$) contain at least

Table 1: Selection criteria that define the signal region. Entries marked with "-" indicate that the variable is not used in the given channel. If different, the entries in square brackets indicate the selection for the different boost regions as defined in the first row of the table. The p_{T} thresholds for the highest and second highest p_{T} jets are $p_{\mathrm{T}}(\mathrm{j}_1)$ and $p_{\mathrm{T}}(\mathrm{j}_2)$, respectively. $\mathrm{CMVA}_{\mathrm{max}}$ and $\mathrm{CMVA}_{\mathrm{min}}$ are the b-tagging requirements for the jets with the highest and second-highest values of the output of the CMVA discriminant. Anti-QCD refers to rejection of events where $E_{\mathrm{T}}^{\mathrm{miss}}$ points in the same or opposite direction of a high p_{T} jet. The values listed for kinematic variables are in units of GeV, and for angles in units of radians.

Variable	0-lepton	1-lepton	2-lepton
$p_{\mathrm{T}}(\mathrm{V})$	> 170	> 100	$[50, 150], > 150$
$M(\ell\ell)$	-	-	$[75, 105]$
p_{T}^{ℓ}	-	$(> 25, > 30)$	> 20
$p_{\mathrm{T}}(j_1)$	> 60	> 25	> 20
$p_{\mathrm{T}}(j_2)$	> 35	> 25	> 20
$p_{\mathrm{T}}(\mathrm{jj})$	> 120	> 100	-
$M(\mathrm{jj})$	$[60, 160]$	$[90, 150]$	$[90, 150]$
$\mathrm{CMVA}_{\mathrm{max}}$	> 0.9432	> 0.9432	> -0.5884
$\mathrm{CMVA}_{\mathrm{min}}$	> -0.5884	> -0.5884	> -0.5884
N_{aj}	< 2	< 2	-
N_{al}	$= 0$	$= 0$	-
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 170	-	-
Anti-QCD	Yes	-	-
$\Delta\phi(\mathrm{V}, \mathrm{H})$	> 2.0	> 2.5	> 2.5
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, E_{\mathrm{T}}^{\mathrm{miss}}{}_{\mathrm{trk}})$	< 0.5	-	-
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \ell)$	-	< 2.0	-
Lepton Isolation	-	< 0.06	-
Event BDT	> -0.8	> 0.3	> -0.8

Table 2: Variables used in the training of the event BDT discriminant. Jets are counted as additional jets if they satisfy the following: $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.4$ for 0-lepton, $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.9$ for 1-lepton, and $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.4$ for 2-lepton.

Variable	Channels utilizing
$M(\mathrm{jj})$: dijet invariant mass	All
$p_{\mathrm{T}}(\mathrm{jj})$: dijet transverse momentum	All
$p_{\mathrm{T}}(\mathrm{V})$: vector boson transverse momentum	All
$\mathrm{CMVA}_{\mathrm{max}}$: value of CMVA for the Higgs boson daughter with largest CSV value	2-lepton, 0-lepton
$\mathrm{CMVA}_{\mathrm{min}}$: value of CMVA for the Higgs boson daughter with second largest CSV value	All
$\mathrm{CMVA}_{\mathrm{add}}$: value of CMVA for the additional jet with largest CSV value	0-lepton
$\Delta\phi(\mathrm{V},\mathrm{H})$: azimuthal angle between V and dijet	All
$p_{\mathrm{T}}(j)$: transverse momentum of each Higgs boson daughter	2-lepton, 0-lepton
$p_{\mathrm{T}}(\mathrm{add.})$: transverse momentum of leading additional jet	0-lepton
$\Delta\eta(\mathrm{jj})$: difference in η between Higgs boson daughters	2-lepton, 0-lepton
$\Delta R(\mathrm{jj})$: distance in η–ϕ between Higgs boson daughters	2-lepton
N_{aj}: number of additional jets N.B. definition slightly different per channel	1-lepton, 2-lepton
$p_{\mathrm{T}}(\mathrm{jj})/p_{\mathrm{T}}(\mathrm{V})$: p_{T} balance between Higgs boson candidate and vector boson	2-lepton
M_Z: Z boson mass	2-lepton
SA5: number of soft activity jets with $p_{\mathrm{T}} > 5$ GeV	All
M_{t}: reconstructed top quark mass	1-lepton
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \ell)$: azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and lepton	1-lepton
$E_{\mathrm{T}}^{\mathrm{miss}}$: missing transverse energy	1-lepton, 2-lepton
$m_T(W)$: W transverse mass	1-lepton
$\Delta\phi(\mathrm{jj})$: difference in ϕ between Higgs boson daughters	0-lepton
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jet.})$: azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and the closest jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$	0-lepton

one b hadron. The scale factors of the sub-processes are allowed to vary independently in the fit. The notation used is: V + udscg for the case where none of those jets contains a b hadron, V + b for the case where only one of the jets contains at least one b hadron, and V + $b\bar{b}$ for the case where both jets contain at least one b hadron.

These scale factors account not only for possible cross section discrepancies, but also for potential residual differences in the selection efficiency of physics objects. Given that each channel probes the tail of a different kinematic distribution, it is expected that for specific selections (e.g. very large E_T^{miss}, high dijet p_T, etc), the scale factors for the same physics process could vary significantly among the various analysis channels. This motivates creating independent control regions for the same process in different channels where feasible. Indeed, there generally is some variation among the channels, although most scale factors are near 1 and all are between 0.78 and 1.7.

Figure 2 shows $p_T(V)$ and examples of distributions for variables in different control regions and for different channels after the scale factors described above have been applied to the corresponding simulated samples.

The $CMVA_{min}$ distributions in all the control regions are fit in combination with the event BDT output distributions in the signal region to extract the VH signal. Examples of the $CMVA_{min}$ distributions after the control region plus signal region fit with full systematic treatment are shown in Fig. 3.

In this fit, discussed further in Section 7, the scale factors for the various background simulated samples are allowed to float freely and observed to be consistent with those obtained in the control regions alone. Table 6 summarizes the scale factors obtained in the final, simultaneous fit of control and signal regions.

Table 3: Definition of the control regions for the 0-lepton channel. The values listed for kinematic variables are in units of GeV, and for angles in units of radians. Entries marked with "-" indicate that the variable is not used in the given control region.

Variable	$t\bar{t}$	Z+LF	Z+HF
V Decay Category	$W(\ell\nu)$	$Z(\nu\nu)$	$Z(\nu\nu)$
$p_T(j_1)$	> 60	> 60	> 60
$p_T(j_2)$	> 35	> 35	> 35
$p_T(jj)$	> 120	> 120	> 120
E_T^{miss}	> 170	> 170	> 170
$\Delta\phi(V, H)$	> 2	> 2	> 2
N_{al}	≥ 1	$= 0$	$= 0$
N_{aj}	≥ 2	≤ 1	≤ 1
$M(jj)$	–	–	$\notin [60 - 160]$
$CMVA_{max}$	> 0.4432	< 0.4432	> 0.9432
$CMVA_{min}$	> -0.5884	> -0.5884	> -0.5884
$\Delta\phi(j, E_T^{miss})$	–	> 0.5	> 0.5
$\Delta\phi(E_T^{miss}, E_T^{miss}{}_{trk})$	–	< 0.5	< 0.5
$\min \Delta\phi(j, E_T^{miss})$	$< \pi/2$	–	–

Table 4: Definition of the control regions for the 1-lepton channels. The same selection is used for all boost regions. LF and HF refer to light- and heavy-flavor jets. METsig is $E_{\mathrm{T}}^{\mathrm{miss}}$ divided by the square root of the scalar sum of jet p_{T} where jet $p_{\mathrm{T}} > 30\,\mathrm{GeV}$. The values listed for kinematic variables are in units of GeV. Entries marked with "-" indicate that the variable is not used in the given control region.

Variable	$\mathrm{t\bar{t}}$	W+LF	W+HF
$p_{\mathrm{T}}(j_1)$	> 25	> 25	> 25
$p_{\mathrm{T}}(j_2)$	> 25	> 25	> 25
$p_{\mathrm{T}}(\mathrm{jj})$	> 100	> 100	> 100
$p_{\mathrm{T}}(\mathrm{V})$	> 100	> 100	> 100
$\mathrm{CMVA}_{\mathrm{max}}$	> 0.9432	$[-0.5884, 0.4432]$	> 0.9432
N_{aj}	> 1	-	$= 0$
N_{al}	$= 0$	$= 0$	$= 0$
METsig	-	> 2.0	> 2.0
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \ell)$	< 2	< 2	< 2
$M(\mathrm{jj})$	< 250	< 250	< 90 (low) or $[150, 250]$ (high)

Table 5: Definition of the control regions for the 2-lepton channels. The same selection is used for both the low- and high-boost regions. The values listed for kinematic variables are in units of GeV. Entries marked with "-" indicate that the variable is not used in the given control region.

Variable	$\mathrm{t\bar{t}}$	Z+LF	Z+HF
$p_{\mathrm{T}}(\mathrm{jj})$	> 100	> 100	-
$p_{\mathrm{T}}(\mathrm{V})$	$[50, 150]$,> 150	$[50, 150]$,> 150	$[50, 150]$,> 150
$\mathrm{CMVA}_{\mathrm{max}}$	> 0.9432	< 0.9432	> 0.9432
$\mathrm{CMVA}_{\mathrm{min}}$	> -0.5884	< -0.5884	> -0.5884
N_{aj}	-	-	-
N_{al}	-	-	-
$E_{\mathrm{T}}^{\mathrm{miss}}$	-	-	< 60
$\Delta\phi(\mathrm{V}, \mathrm{H})$	-	-	> 2.5
$M(\ell\ell)$	$\notin [0, 10]$, $\notin [75, 120]$	$[75, 105]$	$[85, 97]$
$M(\mathrm{jj})$	-	-	$\notin [90, 150]$

Table 6: Data/MC scale factors for each of the main background processes in each channel, as obtained from the combined fit to control and signal region distributions described in Section 7. Electron and muons samples in the 1-lepton and 2-lepton channels are fit simultaneously to determine average scale factors.

Process	0-lepton	1-lepton	2-lepton low-p_{T}	2-lepton high-p_{T}
W0b	1.14 ± 0.07	1.14 ± 0.07	-	-
W1b	1.66 ± 0.12	1.66 ± 0.12	-	-
W2b	1.49 ± 0.12	1.49 ± 0.12	-	-
Z0b	1.03 ± 0.07	-	1.01 ± 0.06	1.02 ± 0.06
Z1b	1.28 ± 0.17	-	0.98 ± 0.06	1.02 ± 0.11
Z2b	1.61 ± 0.10	-	1.09 ± 0.07	1.28 ± 0.09
$\mathrm{t\bar{t}}$	0.78 ± 0.05	0.91 ± 0.03	1.00 ± 0.03	1.04 ± 0.05

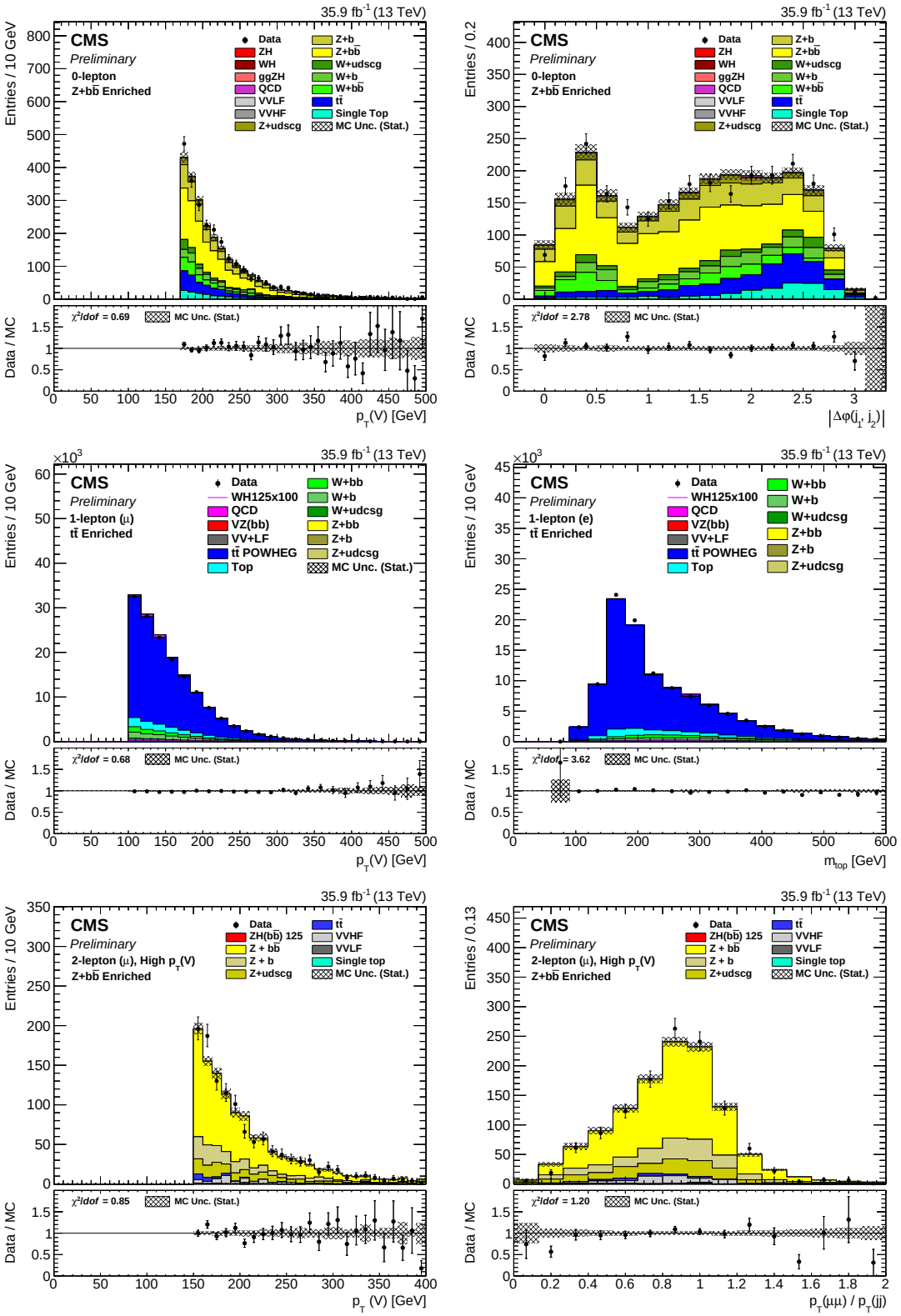

Figure 2: Examples of distributions for variables in the simulated samples and in data for different control regions and for different channels after applying the data/MC scale factors in Table 6. The top row of plots is from the 0-lepton Z+HF control region. The middle row shows variables in the 1-lepton $t\bar{t}$ control region. The bottom row shows variables in the 2-lepton Z+HF control region. The plots on the left are always $p_T(V)$. On the right is a key variable that is validated in that control region. They are, from top to bottom, the azimuthal angle between the two jets that comprise the Higgs boson, the reconstructed top quark mass, and the ratio of $p_T(V)$ and $p_T(jj)$.

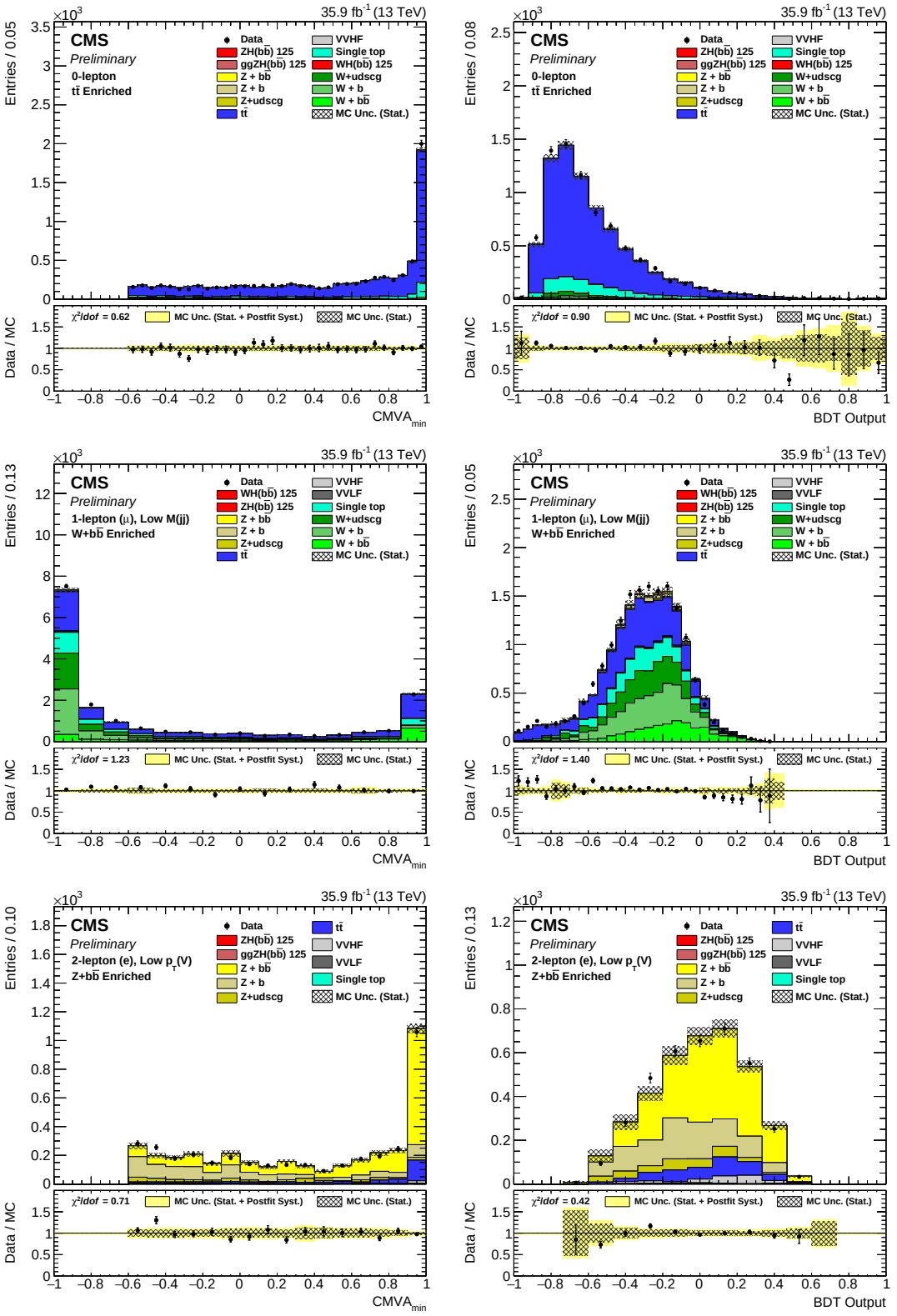

Figure 3: On the left there are examples of $\mathrm{CMVA_{min}}$ distributions in control regions after simulated samples are fit to the data. On the right are corresponding BDT distributions of the same control regions as the plots on the left. Note that the BDT distributions are not part of the fit and are primarily for validation. The control regions shown from top to bottom are: $\mathrm{t\bar{t}}$ for the 0-lepton channel, low mass HF for the single-muon channel, and HF for the dielectron channel.

5.2.1 Background event re-weighting

In inclusive vector boson samples the $p_T(V)$ spectrum in data is observed to be softer than in simulated samples, as expected from higher order electroweak corrections to the production processes [86]. To account for this, the events in all three channels are re-weighted. The correction is negligible for low $p_T(V)$ but becoming sizable at high $p_T(V)$, reaching -10% around 400 GeV.

After these corrections, a residual discrepancy in $p_T(V)$ between data and simulated samples is observed in $t\bar{t}$ and W+jets control regions. In the 0-lepton channel, $t\bar{t}$ samples are re-weighted as a function of the generated top quark's p_T according to observed discrepancies in data and simulated samples in differential top quark cross section measurements [87]. This re-weighting resolves the discrepancy in $p_T(V)$ in $t\bar{t}$ control regions. In the 1-lepton channel, additional corrections are needed for W+jets, and corrections are derived from the data in 1-lepton control regions for these processes: $t\bar{t}$, W+LF, and the sum of W+b, W+bb and single top. A re-weighting of simulated events in $p_T(V)$ is derived for each such that the shape of the sum of simulated processes matches the data. The correction functions are extracted through a simultaneous fit of linear functions in $p_T(V)$. The uncertainties in the fit parameters are used to assess the systematic uncertainty. The $p_T(V)$ spectra resulting from re-weighting in either the top quark p_T or $p_T(V)$ are equivalent.

The V+jets LO simulated samples are used in the analysis because considerably more events are available than for the NLO samples. A normalization K-factor is applied to the LO samples to account for the difference in cross sections. Kinematic distributions between the two samples are found to be consistent after applying to the LO samples a correction derived from the comparison with the NLO samples of the distribution of the pseudorapidity separation between the two jets in the event that are candidates for the reconstructed $H \to b\bar{b}$ decay, $\Delta\eta(jj)$. Different corrections are derived depending on whether these two jets are matched to zero, one, or two b quarks. Both the $\Delta\eta(jj)$ distributions of the NLO samples and the corrected LO samples agree well with data in control regions.

6 Uncertainties

Systematic effects impact mass resolution, BDT shape, and signal and background normalizations in the most sensitive region of the BDT. The uncertainties associated with the fitted scale factors have the largest impact on the uncertainty of the fitted signal strength, μ. The next largest effect comes from the size of the simulated samples and uncertainties from correcting mismodeling of kinematic variables, both in signal and in background simulated samples. The next group of very significant systematic uncertainties are related to b-tagging uncertainties and uncertainties in jet energy. All systematic uncertainties are listed in Table 7 and are described in more detail below.

The sizes of simulated samples are sometimes limited, and in some cases the statistical uncertainty in the simulated samples is non-negligible compared to the statistical uncertainty of the data. In this case, nuisance parameters are fitted within poissonian statistical uncertainty in the likelihood function per histogram bin. V+jets samples often require these nuisance parameters, and so this is among the leading systematic uncertainties for this analysis.

The corrections to the $p_T(V)$ spectra in the $t\bar{t}$ and W+jets samples are applied per sample according to the uncertainty in the simultaneous fit previously described. This uncertainty on the correction is at most 5% on the background yield near $p_T(V)$ of 400 GeV. For V+jets, the

difference between the shape of the BDT output distribution for events generated with the MADGRAPH and the MC@NLO ++ Monte Carlo generators is considered as a shape systematic uncertainty. For $t\bar{t}$ the differences in the shape of the BDT output distribution between the nominal sample generated with POWHEG and that obtained from the MC@NLO [88] generator are considered as shape systematic uncertainties. Uncertainties due to the renormalisation and factorisation of the QCD scale and PDF uncertainties are also considered for the simulated backgrounds.

The b-tagging efficiencies and the probability to tag as a b jet a jet originating from a different flavor (mistag) are measured in heavy-flavor enhanced samples of jets that contain muons and are applied consistently to jets in signal and background events. The measured uncertainties for the b-tagging scale factors are: 1.5% per b-quark tag, 5% per charm-quark tag, and 10% per mistagged jet (originating from gluons and light u, d, or s quarks) [77]. These uncertainties are propagated to the $\mathrm{CMVA}_{\mathrm{min}}$ distributions by re-weighting events. The shape of the BDT output distribution is also affected by the shape of the CMVA distributions because $\mathrm{CMVA}_{\mathrm{min}}$ is an input to the BDT. For the 2-lepton channel $\mathrm{CMVA}_{\mathrm{max}}$is also an input to the BDT. The signal strength uncertainty increases by 8% and 5%, respectively, due to b-tagging and mistag scale factor uncertainties propagated through the CMVA and finally to the BDT.

The uncertainties in the jet energy scale and resolution have an effect on the shape of the event BDT output distribution because the dijet invariant mass is crucial BDT input. The impact of the jet energy scale uncertainty is determined by recomputing the BDT output distribution after shifting the energy scale up and down by its uncertainty. Similarly, the impact of the jet energy resolution is determined by recomputing the BDT output distribution after increasing or decreasing the jet energy resolution. The individual contribution to the increase in signal strength uncertainty is found to be around 6% for the jet energy scale and 4% for the jet energy resolution uncertainty. The uncertainty on the jet energy scale and resolution vary as a function of jet p_{T} and η. There are several individual sources of uncertainty, which are derived independently and are fully uncorrelated between themselves for jet energy scale [89], while a single resolution shape systematic is evaluated.

The total VH signal cross section has been calculated to next-to-next-to-leading-order together with next-to-next-to-leading-log (NNLO+NNLL) QCD accuracy combined with NLO electroweak, and the associate systematic uncertainties include the effect of scale variations and PDF uncertainties [60]. The estimated uncertainties of the NLO electroweak corrections are 7% for the WH and 5% for the ZH production processes, respectively. The estimate for the NNLO QCD correction results in an uncertainty of 1% for the WH and 4% for the ZH production processes, respectively.

An uncertainty of 15% is assigned to the event yields obtained from simulated samples for single-top-quark production. For the diboson backgrounds, a 15% cross section uncertainty is assumed. These uncertainties are consistent with the CMS measurements of these processes [90–92].

In 0- and 1-lepton channel we consider the uncertainty related to the $E_{\mathrm{T}}^{\mathrm{miss}}$ estimate.

Muon and electron trigger, reconstruction, and identification efficiencies in simulated samples are corrected for differences in data and simulate using samples of leptonic Z-boson decay. These corrections are affected by uncertainties coming from the efficiency measurement method, the lepton selection and the limited size of the Z-boson samples. They are measured and propagated as a function of lepton p_{T} and η. The parameters describing the $\mathrm{Z}(\nu\nu)\mathrm{H}$ trigger efficiency turn-on curve have been varied within their statistical uncertainties and also

Table 7: Effect of each source of systematic uncertainty on the signal strength μ (defined as the ratio of the best-fit value for the production cross section for a 125 GeV Higgs boson, relative to the SM cross section). The third column shows the uncertainty in μ from each source when only that particular source is considered. The last column shows the percentage decrease in the uncertainty when removing that specific source of uncertainty. Due to correlations, the total systematic uncertainty is less than the sum in quadrature of the individual uncertainties. The second column shows whether the source affects only the normalization or both the shape and normalization of the event BDT output distribution. See text for details.

Source	Type	Individual contribution to μ uncertainty (%)	Effect of removal on μ uncertainty (%)
Scale factors ($t\bar{t}$,V+jets)	norm.	9.4	3.5
Size of simulated samples	shape	8.1	3.1
Simulated samples' modeling	shape	4.1	2.9
Btag	shape	7.9	1.8
Jet energy scale	shape	4.2	1.8
Signal cross sections	norm.	5.3	1.1
Cross section uncertainties (single-top, VV)	norm.	4.7	1.1
Jet energy resolution	shape	5.6	0.9
Mistag	shape	4.6	0.9
Luminosity	norm.	2.2	0.9
Missing transverse energy	shape	1.3	0.2
Lepton efficiency and trigger	norm.	1.9	0.1

estimated for different assumptions on the methods used to derive the efficiency. The total individual impact of the two sources of uncertainty is about 2% on signal strength uncertainty.

The uncertainty in the CMS luminosity measurement is estimated to be 2.5% [93]. Events in simulated samples must be re-weighted such that the distribution of pileup in the simulated samples matches that estimated in data. A 5% uncertainty is assigned and in practice the impact of this uncertainty is negligible.

The combined effect of the systematic uncertainties results in a reduction of 25% on the expected significance of an observation when the Higgs boson is present in the data at the predicted SM rate.

7 Results

Results are obtained from combined signal and background binned-likelihood fits, simultaneously for all channels, to both the shape of the output distribution of the event BDT discriminants in the signal region and to the $\mathrm{CMVA}_{\mathrm{min}}$ distributions for the control regions corresponding to each channel. The BDT discriminants are trained separately for each channel to search for a Higgs boson with a mass of 125 GeV. To remove the background-dominated portion of the BDT output distribution, only events with a BDT output value above a certain threshold are considered. To achieve a better sensitivity in the search, this threshold was optimized separately for each channel.

In this fit the shape and normalization of all distributions for signal and for each background component are allowed to vary within the systematic and statistical uncertainties described in Section 6. These uncertainties are treated as independent nuisance parameters in the fit. Nuisance parameters, the signal strength and the scale factors described in Section 5.2 are allowed to float freely and are adjusted by the fit.

In total, seven event BDT output distributions are included in the fit: one for the 0-lepton channel, one for each lepton flavor for the 1-lepton channels, and two for each lepton flavor for the 2-lepton channels (corresponding to the two $p_{\mathrm{T}}(\mathrm{V})$ regions). The number of $\mathrm{CMVA}_{\mathrm{min}}$ distributions included is 21: three for the 0-lepton channel, three for each lepton flavor for the 1-lepton channels, and six for each lepton flavor for the 2-lepton channels (each corresponding to one of two $p_{\mathrm{T}}(\mathrm{V})$ regions). Figure 4 shows the seven BDT output distributions after they have been adjusted by the fit.

Table 8 lists, for the 20% most-sensitive region of the BDT output distribution, the total number of events for the main backgrounds, for the 125 GeV SM Higgs boson signal, and for data. Simulation is normalized using the results of the simultaneous fit of signal plus background to data. An excess compatible with the presence of the SM Higgs boson is observed. Figure 5 combines the BDT output values of all channels where the events are gathered in bins of similar expected signal-to-background ratio, as given by the value of the output of their corresponding BDT discriminant (trained with a Higgs boson mass hypothesis of 125 GeV). The observed excess of events in the bins with the largest signal-to-background ratio is consistent with what is expected from the production of the SM Higgs boson.

Table 8: The total number of events in each channel, for the 20% most-sensitive region of the BDT output distribution, for the expected backgrounds, for the 125 GeV SM Higgs boson VH signal, and for data. The signal-to-background ratio (S/B) is also shown.

	0-lepton	1-lepton	2-lepton	
Process			Low $p_{\mathrm{T}}(\mathrm{V})$	High $p_{\mathrm{T}}(\mathrm{V})$
Vbb	216.8	102.5	617.5	113.9
Vb	31.8	19.9	141.1	17.2
V + udscg	10.2	9.8	58.4	4.1
$\mathrm{t\bar{t}}$	34.7	98.0	157.7	3.2
Single-top-quark	11.8	44.6	2.0	0.2
VV(udscg)	0.4	1.5	6.4	0.6
$\mathrm{VZ(b\bar{b})}$	7.7	6.9	22.9	3.8
Total backgrounds	267.0	283.3	1005.9	142.9
VH	34.7	26.0	33.5	22.1
Data	334	320	1030	179
S/B	0.13	0.11	0.033	0.156

For $m_{\mathrm{H}} = 125$ GeV, the excess of observed events corresponds to a local significance of 3.3 standard deviations away from the background-only hypothesis. Significance is computed using the standard LHC profile likelihood asymptotic approximation [94]. This excess is consistent with the SM prediction for Higgs boson production with signal strength, as the best-fit value of the production cross section for a 125 GeV Higgs boson, relative to the SM cross section, $\mu_{H,SM} = \sigma/\sigma_{\mathrm{SM}}$, is $1.19^{+0.21}_{-0.20}(\mathrm{stat.})^{+0.34}_{-0.32}(\mathrm{syst.})$. With $\mu_{H,SM} = 1.0$ the expected significance is 2.8 standard deviations.

The relative sensitivity of the channels that are topologically distinct is shown in Table 9 for $m_{\mathrm{H}} = 125$ GeV. The table lists the expected and observed significances for the 0-lepton channel, for the 1-lepton channels combined, and for the 2-lepton channels combined.

The best-fit values of the production cross section for a 125 GeV Higgs boson, relative to the SM cross section (signal strength, μ), are shown in the lower portion of Fig. 6 for 0-, 1- and

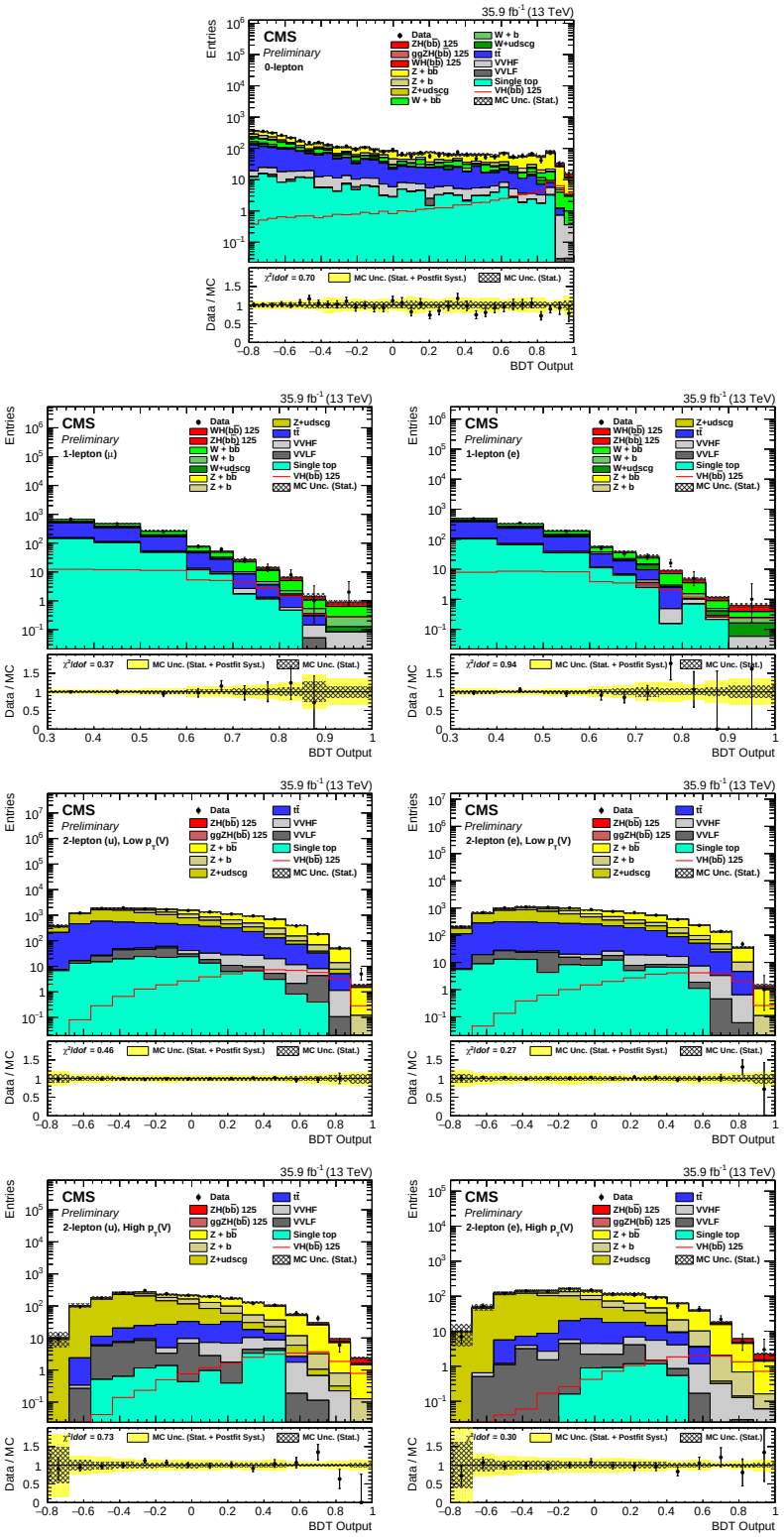

Figure 4: Post-fit BDT output distributions for the 13 TeV data (points with error bars), for the 0-lepton channel (top), for the 1-lepton channels (middle), and for the 2-lepton low-$p_{\mathrm{T}}(\mathrm{V})$ and high-$p_{\mathrm{T}}(\mathrm{V})$ regions (bottom). The bottom inset shows the ratio of the number of events observed in data to that of the prediction from simulated samples for signal and backgrounds.

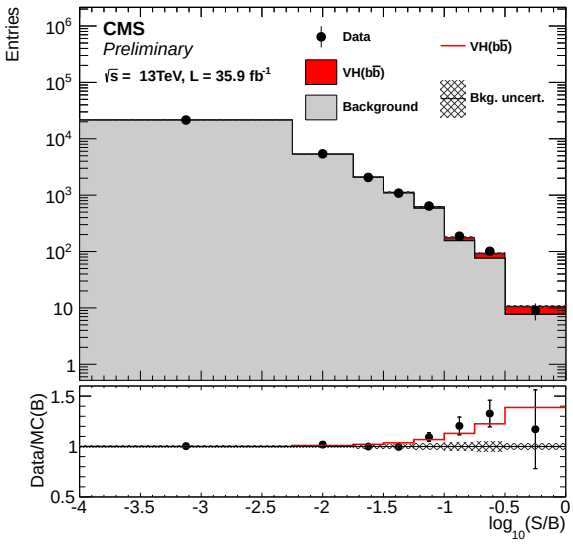

Figure 5: Combination of all channels into a single event BDT distribution. Events are sorted in bins of similar expected signal-to-background ratio, as given by the value of the output of the value of their corresponding BDT discriminant (trained with a Higgs boson mass hypothesis of 125 GeV). The bottom inserts show the ratio of the data to the background-only prediction.

Table 9: The expected and observed significances for VH production with $H \to b\bar{b}$ are shown for each channel fit individually as well as for the combination of all three channels.

$m_H = 125$ GeV	Significance expected	Significance observed
0-lepton	1.5	0.0
1-lepton	1.5	3.2
2-lepton	1.8	3.1
All channels	2.8	3.3

2-lepton channels. The observed signal strengths are consistent with each other at the 5% level. In the upper portion of Fig. 6 the signal strengths for the separate WH and ZH production processes are shown. The two production modes are consistent with the SM expectations within uncertainties.

The fit for the WH and ZH production modes is not simply correlated to the analysis channels because the analysis channels contain mixed processes. The WH process contributes approximately 15% of the Higgs boson signal event yields in the 0-lepton channel, resulting from events in which the lepton is outside the detector acceptance, and the 2-lepton process contributes less than 3% to the 1-lepton channel when one of the leptons is outside the detector acceptance.

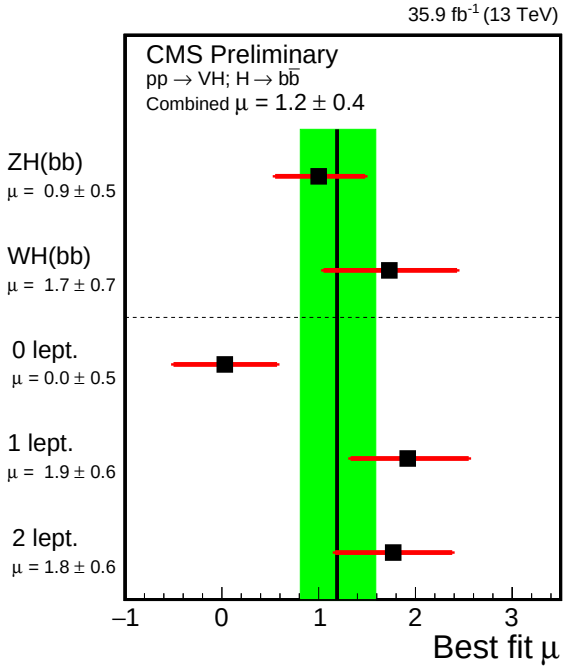

Figure 6: The best-fit value of the production cross section for a 125 GeV Higgs boson relative to the SM cross section–i.e., signal strength μ–is shown in black with green error band. Above the dashed line are the WH and ZH signal strengths when each production mode has an independent signal strength parameters in the fit. When each channel is fit with its own signal strength parameter, the results are shown below the dashed line.

7.1 Combined results with Run 1 data for VH with $H \to b\bar{b}$

The combination of this result with similar searches performed by the CMS experiment during Run 1 of the LHC [17, 37, 39] (using using proton-proton collisions at $\sqrt{s} = 7$ and $\sqrt{s} = 8$ TeV with data samples corresponding to luminosities of up to 5.1 fb^{-1} and 18.9 fb^{-1}, respectively) yields an observed signal significance of 3.79 standard deviations, where 3.75 are expected from a SM signal. The corresponding signal strength is $\mu = \sigma/\sigma_{\text{SM}} = 1.06^{+0.31}_{-0.29}$. In the combination, all systematic uncertainties are assumed to be uncorrelated and all uncertainties from theory are assumed to be fully correlated. Table 10 lists these results.

Table 10: The expected and observed significances and the observed signal strengths for VH production with H $\to$ b$\bar{\text{b}}$ for Run 1 data [17], Run 2 2016 data, and for the combination of the two.

$m_{\text{H}} = 125\,\text{GeV}$	Significance expected	Significance observed	Signal strength observed
Run 1	2.5	2.1	$0.89^{+0.44}_{-0.42}$
Run 2	2.8	3.3	$1.19^{+0.40}_{-0.38}$
combined	3.8	3.8	$1.06^{+0.31}_{-0.29}$

7.2 Extraction of VZ with Z $\to$ b$\bar{\text{b}}$

The VZ process with Z $\to$ b$\bar{\text{b}}$, having a nearly identical final state as VH with H $\to$ b$\bar{\text{b}}$ serves as a validation of the methodology used in the search for the latter process. Event BDT discriminants are trained using as signal the diboson sample for the VZ with Z $\to$ b$\bar{\text{b}}$ process. All other processes, including VH production (at the predicted SM rate for a 125 GeV Higgs boson mass), are treated as background. The only modification made is the requirement that in the signal region $M(\text{jj})$ be in the $[60, 160]$ GeV range.

The results from the combined fit for all channels of the control and signal region distributions, as defined in Sections 5.1 and 5.2, are summarized in Table 11 for the data recorded during 2016, corresponding to an integrated luminosity of 35.9 fb^{-1} at $\sqrt{s} = 13$ TeV. The observed excess of events for the combined WZ and ZZ processes, with Z $\to$ b$\bar{\text{b}}$, has a significance of 5.0 standard deviations from the background-only event yield expectation. The corresponding signal strength, relative to the prediction from the diboson MADGRAPH generator mentioned in Section 2, and rescaled to the cross section from the NLO MCFM generator, is measured to be $\mu_{\text{VV}} = 1.02^{+0.22}_{-0.23}$.

Table 11: Validation results for VZ production with Z $\to$ b$\bar{\text{b}}$. Expected and observed signal strengths, and expected and observed local significances of the excess of events above the estimated background. Values are given in numbers of standard deviations.

Channel	Expected signal strength VZ(b$\bar{\text{b}}$)	Observed signal strength VZ(b$\bar{\text{b}}$)	Expected significance VZ(b$\bar{\text{b}}$)	Observed significance VZ(b$\bar{\text{b}}$)
0-lepton	1.00 ± 0.33	0.57 ± 0.32	3.1	2.0
1-lepton	1.00 ± 0.38	1.67 ± 0.47	2.6	3.7
2-lepton	1.00 ± 0.31	1.33 ± 0.34	3.2	4.5
Combined	1.00 ± 0.22	1.02 ± 0.22	4.9	5.0

Figure 7 shows the combined event BDT output distribution for all channels, with the content of each bin, for each channel, weighted by the expected signal-to-background ratio. The excess of events in data, over background, is shown to be compatible with the yield expectation from VZ production with Z $\to$ b$\bar{\text{b}}$

8 Summary

A search for the 125 GeV standard model (SM) Higgs boson (H) when produced in association with an electroweak vector boson and decaying to b$\bar{\text{b}}$ is reported for the Z($\nu\nu$)H, W($\mu\nu$)H, W(eν)H, Z($\mu\mu$)H and Z(ee)H processes. The search is performed in data samples corresponding to integrated luminosities of 35.9 fb^{-1} at $\sqrt{s} = 13$ TeV, recorded by the CMS experiment

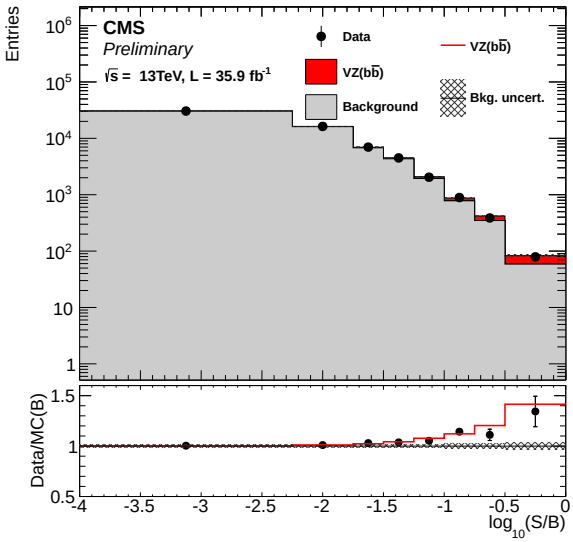

Figure 7: Combination of all channels in the VZ search, with $Z \to b\bar{b}$ into a single event BDT distribution. Events are sorted in bins of similar expected signal-to-background ratio, as given by the value of the output of their corresponding BDT discriminant. The bottom inset shows the ratio of the data to the predicted background, with a red line overlaying the expected SM contribution from VZ with $Z \to b\bar{b}$.

at the LHC. The observed signal significance is 3.3 standard deviations, where the expectation from the SM Higgs production is 2.8. The corresponding signal strength is $\mu = \sigma/\sigma_{SM} = 1.2 \pm 0.4$.

The combination of this result with the one from the same search performed by the CMS experiment in Run 1 of the LHC using using proton-proton collisions at $\sqrt{s} = 7$ and $\sqrt{s} = 8$ TeV with data samples corresponding to luminosities of up to 5.1 fb^{-1} and 18.9 fb^{-1}, respectively yields an observed signal significance of 3.8 standard deviations, where 3.8 are expected from a SM signal. The corresponding signal strength is $\mu = \sigma/\sigma_{SM} = 1.06^{+0.31}_{-0.29}$.

This result provides strong evidence for the decay of the Higgs boson into a pair of b quarks.

References

[1] 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.

[2] 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** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[3] 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.

[4] ATLAS Collaboration, "Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS

detector", *Phys. Rev. D* **90** (2014), no. 11, 112015, doi:10.1103/PhysRevD.90.112015, arXiv:1408.7084.

[5] CMS Collaboration, "Observation of the diphoton decay of the Higgs boson and measurement of its properties", *Eur. Phys. J. C* **74** (Jul, 2014) 3076. 79 p, doi:10.1140/epjc/s10052-014-3076-z, arXiv:1407.0558.

[6] ATLAS Collaboration, "Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector", *Phys. Rev. D* **91** (2015), no. 1, 012006, doi:10.1103/PhysRevD.91.012006, arXiv:1408.5191.

[7] CMS Collaboration, "Measurement of the properties of a Higgs boson in the four-lepton final state", *Phys. Rev. D* **89** (2014) 092007, doi:10.1103/PhysRevD.89.092007, arXiv:1312.5353.

[8] ATLAS Collaboration, "Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector", *Phys. Rev. D* **92** (2015), no. 1, 012006, doi:10.1103/PhysRevD.92.012006, arXiv:1412.2641.

[9] ATLAS Collaboration, "Study of (W/Z)H production and Higgs boson couplings using $H \to WW^*$ decays with the ATLAS detector", *JHEP* **08** (2015) 137, doi:10.1007/JHEP08(2015)137, arXiv:1506.06641.

[10] CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states", *JHEP* **01** (2014) 096, doi:10.1007/JHEP01(2014)096, arXiv:1312.1129.

[11] ATLAS Collaboration, "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector", *JHEP* **04** (2015) 117, doi:10.1007/JHEP04(2015)117, arXiv:1501.04943.

[12] CMS Collaboration, "Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons", *JHEP* **05** (2014) 104, doi:10.1007/JHEP05(2014)104, arXiv:1401.5041.

[13] ATLAS Collaboration, "Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment", *Eur. Phys. J. C* **76** (2016), no. 1, 6, doi:10.1140/epjc/s10052-015-3769-y, arXiv:1507.04548.

[14] 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), no. 5, 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.

[15] CMS Collaboration, "Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs", *Phys. Rev. Lett.* **110** (2013), no. 8, 081803, doi:10.1103/PhysRevLett.110.081803, arXiv:1212.6639.

[16] ATLAS Collaboration, "Evidence for the spin-0 nature of the Higgs boson using ATLAS data", *Phys. Lett. B* **726** (2013) 120–144, doi:10.1016/j.physletb.2013.08.026, arXiv:1307.1432.

[17] ATLAS, CMS Collaboration, "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 $\sqrt{s} = 7$ and 8 TeV", *JHEP* **08** (2016) 045, doi:10.1007/JHEP08(2016)045, arXiv:1606.02266.

[18] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", *Phys. Rev. Lett.* **13** (1964) 321, doi:10.1103/PhysRevLett.13.321.

[19] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132, doi:10.1016/0031-9163(64)91136-9.

[20] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (1964) 508, doi:10.1103/PhysRevLett.13.508.

[21] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles", *Phys. Rev. Lett.* **13** (1964) 585, doi:10.1103/PhysRevLett.13.585.

[22] P. W. Higgs, "Spontaneous symmetry breakdown without massless bosons", *Phys. Rev.* **145** (1966) 1156, doi:10.1103/PhysRev.145.1156.

[23] T. W. B. Kibble, "Symmetry breaking in non-Abelian gauge theories", *Phys. Rev.* **155** (1967) 1554, doi:10.1103/PhysRev.155.1554.

[24] ATLAS, CMS Collaboration, "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments", *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.

[25] CMS Collaboration, "Observation of Higgs boson decays to a pair of tau leptons",. To be submitted to Phys. Lett. B.

[26] CMS Collaboration, "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at sqrt(s) = 13 TeV", arXiv:1706.09936.

[27] S. Heinmeyer et al., "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", CERN Report CERN-2013-004, 2013. doi:10.5170/CERN-2013-004, arXiv:1307.1347.

[28] S. Weinberg, "A model of leptons", *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266, doi:10.1103/PhysRevLett.19.1264.

[29] Y. Nambu and G. Jona-Lasinio, "Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. I", *Phys. Rev.* **122** (1961) 345, doi:10.1103/PhysRev.122.345.

[30] CDF, D0 Collaboration, "Higgs Boson Studies at the Tevatron", *Phys. Rev. D* **88** (2013), no. 5, 052014, doi:10.1103/PhysRevD.88.052014, arXiv:1303.6346.

[31] CDF, D0 Collaboration, "Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron", *Phys. Rev. Lett.* **109** (2012) 071804, doi:10.1103/PhysRevLett.109.071804, arXiv:1207.6436.

[32] ATLAS Collaboration, "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector", *Eur. Phys. J. C* **75** (2015), no. 7, 349, doi:10.1140/epjc/s10052-015-3543-1, arXiv:1503.05066.

[33] ATLAS Collaboration, "Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *JHEP* **05** (2016) 160, doi:10.1007/JHEP05(2016)160, arXiv:1604.03812.

[34] CMS Collaboration, "Search for the standard model Higgs boson produced in association with a top-quark pair in pp collisions at the LHC", *JHEP* **05** (2013) 145, doi:10.1007/JHEP05(2013)145, arXiv:1303.0763.

[35] CMS Collaboration, "Search for the associated production of the Higgs boson with a top-quark pair", *JHEP* **09** (2014) 087, doi:10.1007/JHEP09(2014)087, 10.1007/JHEP10(2014)106, arXiv:1408.1682. [Erratum: JHEP10,106(2014)].

[36] ATLAS Collaboration, "Search for the Standard Model Higgs boson produced by vector-boson fusion and decaying to bottom quarks in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector", *JHEP* **11** (2016) 112, doi:10.1007/JHEP11(2016)112, arXiv:1606.02181.

[37] 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), no. 3, 032008, doi:10.1103/PhysRevD.92.032008, arXiv:1506.01010.

[38] ATLAS Collaboration, "Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector", *JHEP* **01** (2015) 069, doi:10.1007/JHEP01(2015)069, arXiv:1409.6212.

[39] CMS Collaboration, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks", *Phys. Rev. D* **89** (2014), no. 1, 012003, doi:10.1103/PhysRevD.89.012003, arXiv:1310.3687.

[40] ATLAS Collaboration, "Evidence for the Hbb decay with the ATLAS detector", ATLAS CONFERNCE NOTE ATLAS-CONF-2017-041, CERN, Geneva, Jul, 2017.

[41] B. P. Roe et al., "Boosted decision trees, an alternative to artificial neural networks", *Nucl. Instrum. Meth. A* **543** (2005) 577, doi:10.1016/j.nima.2004.12.018, arXiv:physics/0408124.

[42] A. Hocker et al., "TMVA—Toolkit for Multivariate Data Analysis", *PoS* **ACAT** (2007) 040, arXiv:physics/0703039.

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

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

[45] S. Frixione, P. Nason, and C. Oleari, "Matching nlo qcd computations with parton shower simulations: the powheg method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[46] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, "$HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO", *JHEP* **10** (2013) 083, doi:10.1007/JHEP10(2013)083, arXiv:1306.2542.

[47] 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** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[48] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[49] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[50] S. Frixione, P. Nason, and G. Ridolfi, "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[51] R. Frederix, E. Re, and P. Torrielli, "Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO", *JHEP* **09** (2012) 130, doi:10.1007/JHEP09(2012)130, arXiv:1207.5391.

[52] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[53] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1007/JHEP02(2010)011, 10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: JHEP02,011(2010)].

[54] G. Ferrera, M. Grazzini, and F. Tramontano, "Higher-order QCD effects for associated WH production and decay at the LHC", *JHEP* **04** (2014) 039, doi:10.1007/JHEP04(2014)039, arXiv:1312.1669.

[55] G. Ferrera, M. Grazzini, and F. Tramontano, "Associated ZH production at hadron colliders: the fully differential NNLO QCD calculation", *Phys. Lett. B* **740** (2015) 51–55, doi:10.1016/j.physletb.2014.11.040, arXiv:1407.4747.

[56] G. Ferrera, M. Grazzini, and F. Tramontano, "Associated WH production at hadron colliders: a fully exclusive QCD calculation at NNLO", *Phys. Rev. Lett.* **107** (2011) 152003, doi:10.1103/PhysRevLett.107.152003, arXiv:1107.1164.

[57] O. Brein, R. V. Harlander, and T. J. E. Zirke, "vh@nnlo - Higgs Strahlung at hadron colliders", *Comput. Phys. Commun.* **184** (2013) 998–1003, doi:10.1016/j.cpc.2012.11.002, arXiv:1210.5347.

[58] R. V. Harlander, S. Liebler, and T. Zirke, "Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-Doublet Model", *JHEP* **02** (2014) 023, doi:10.1007/JHEP02(2014)023, arXiv:1307.8122.

[59] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, "HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders", *Comput. Phys. Commun.* **195** (2015) 161–171, doi:10.1016/j.cpc.2015.04.021, arXiv:1412.5390.

[60] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", arXiv:1610.07922.

[61] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 10, doi:10.1016/j.nuclphysbps.2010.08.011, arXiv:1007.3492.

[62] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", *Comput. Phys. Commun.* **182** (2011) 2388, doi:10.1016/j.cpc.2011.06.008, arXiv:1011.3540.

[63] Y. Li and F. Petriello, "Combining QCD and electroweak corrections to dilepton production in FEWZ", *Phys. Rev. D* **86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.

[64] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "W Physics at the LHC with FEWZ 2.1", *Comput. Phys. Commun.* **184** (2013) 208, doi:10.1016/j.cpc.2012.09.005, arXiv:1201.5896.

[65] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[66] T. Sjöstrand, S. Mrenna, and P. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[67] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016), no. 3, 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[68] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune", *Eur. Phys. J. C* **74** (2014), no. 8, 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.

[69] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", arXiv:1706.04965. Submitted to JINST.

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

[71] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[72] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[73] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{7}$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[74] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[75] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett. B* **641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[76] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[77] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[78] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[79] CMS Collaboration, "Commissioning of trackjets in pp collisions at $\sqrt{s}$=7 TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-006, CERN, 2010.

[80] CMS Collaboration, "Performance of jet reconstruction with charged tracks only", CMS Physics Analysis Summary CMS-PAS-JME-08-001, CERN, 2009.

[81] CMS Collaboration, "Measurement of the hadronic activity in events with a Z and two jets and extraction of the cross section for the electroweak production of a Z with two jets in pp collisions at $\sqrt{s}$ = 7 TeV", *JHEP* **10** (2013) 062, doi:10.1007/JHEP10(2013)062, arXiv:1305.7389.

[82] CMS Collaboration, "Measurement of electroweak production of two jets in association with a Z boson in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J. C* **75** (2015), no. 2, 66, doi:10.1140/epjc/s10052-014-3232-5, arXiv:1410.3153.

[83] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS Physics Analysis Summary CMS-PAS-JME-16-003, CERN, Geneva, 2017.

[84] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC", *Phys. Rev. Lett.* **100** (2008) 242001, doi:10.1103/PhysRevLett.100.242001, arXiv:0802.2470.

[85] CDF, D0 Collaboration, "Improved b-jet Energy Correction for $H \to b\bar{b}$ Searches at CDF", arXiv:1107.3026.

[86] S. Kallweit et al., "NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging", *JHEP* **04** (2016) 021, doi:10.1007/JHEP04(2016)021, arXiv:1511.08692.

[87] CMS Collaboration, "Measurement of differential cross sections for top quark pair production using the lepton + jets final state in proton-proton collisions at 13 tev", *Phys. Rev. D* **95** (May, 2017) 092001, doi:10.1103/PhysRevD.95.092001.

[88] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations", *JHEP* **06** (2002) 029, doi:10.1088/1126-6708/2002/06/029, arXiv:hep-ph/0204244.

[89] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017), no. 02, P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[90] CMS Collaboration, "Cross section measurement of t-channel single top quark production in pp collisions at sqrt(s) = 13 TeV", *Submitted to: Phys. Lett. B* (2016) arXiv:1610.00678.

[91] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions at $\sqrt{(s)} = 13$ TeV", *Phys. Lett. B* **766** (2017) 268–290, doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.

[92] CMS Collaboration, "Measurement of the ZZ production cross section and Z $\to \ell^+\ell^-\ell'^+\ell'^-$ branching fraction in pp collisions at $\sqrt{s}$=13 TeV", *Phys. Lett. B* **763** (2016) 280–303, doi:10.1016/j.physletb.2016.10.054, arXiv:1607.08834.

[93] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

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