Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks

Citation

As Published
http://dx.doi.org/10.1103/PhysRevD.89.012003

Publisher
American Physical Society

Version
Final published version

Accessed
Sun Feb 10 04:53:03 EST 2019

Citable Link
http://hdl.handle.net/1721.1/87750

Terms of Use
Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.

Detailed Terms

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Search for the standard model Higgs boson produced in association with a $W$ or a $Z$ boson and decaying to bottom quarks

S. Chatrchyan et al.*
(CMS Collaboration)

(Received 14 October 2013; published 21 January 2014)

A search for the standard model Higgs boson ($H$) decaying to $b\bar{b}$ when produced in association with a weak vector boson ($V$) is reported for the following channels: $W(\mu\nu)H$, $W(e\nu)H$, $W(\tau\nu)H$, $Z(\mu\mu)H$, $Z(\tau\tau)H$, and $Z(\nu\nu)H$. The search is performed in data samples corresponding to integrated luminosities of up to 5.1 inverse femtobarns at $\sqrt{s} = 7$ TeV and up to 18.9 fb$^{-1}$ at $\sqrt{s} = 8$ TeV, recorded by the CMS experiment at the LHC. An excess of events is observed above the expected background with a local significance of 2.1 standard deviations for a Higgs boson mass of 125 GeV, consistent with the expectation from the production of the standard model Higgs boson. The signal strength corresponding to this excess, relative to that of the standard model Higgs boson, is $1.0 \pm 0.5$.

DOI: 10.1103/PhysRevD.89.012003 PACS numbers: 13.85.Rm, 14.80.Bn

I. INTRODUCTION

At the Large Hadron Collider (LHC), the ATLAS and CMS collaborations have reported the discovery of a new boson [1,2] with a mass, $m_H$, near 125 GeV and properties compatible with those of the standard model (SM) Higgs boson [3–8]. To date, significant signals have been observed in channels where the boson decays into $\gamma\gamma$, $ZZ$, or $WW$. The interaction of this boson with the massive $W$ and $Z$ vector bosons indicates that it plays a role in electroweak symmetry breaking. The interaction with the fermions and whether the Higgs field serves as the source of mass generation in the fermion sector, through a Yukawa interaction, remains to be firmly established.

At $m_H = 125$ GeV the standard model Higgs boson decays predominantly into a bottom quark-antiquark pair ($b\bar{b}$) with a branching fraction of $= 58\%$ [9]. The observation and study of the $H \rightarrow b\bar{b}$ decay, which involves the direct coupling of the Higgs boson to down-type quarks, is therefore essential in determining the nature of the newly discovered boson. The measurement of the $H \rightarrow b\bar{b}$ decay will be the first direct test of whether the observed boson interacts as expected with the quark sector, as the coupling to the top quark has only been tested through loop effects.

In their combined search for the SM Higgs boson [10], the CDF and D0 collaborations at the Tevatron $p\bar{p}$ collider have reported evidence for an excess of events in the 115–140 GeV mass range, consistent with the mass of the Higgs boson observed at the LHC. In that search, the sensitivity below a mass of 130 GeV is dominated by the channels in which the Higgs boson is produced in association with a weak vector boson and decaying to $b\bar{b}$ [11]. The combined local significance of this excess is reported to be 3.0 standard deviations at $m_H = 125$ GeV, while the expected local significance is 1.9 standard deviations. At the LHC, a search for $H \rightarrow b\bar{b}$ by the ATLAS experiment using data samples corresponding to an integrated luminosity of 4.7 fb$^{-1}$ at $\sqrt{s} = 7$ TeV resulted in exclusion limits on Higgs boson production, at the 95% confidence level (C.L.), of 2.5 to 5.5 times the standard model cross section in the 110–130 GeV mass range [12].

This article reports on a search at the Compact Muon Solenoid (CMS) experiment for the standard model Higgs boson in the $pp \rightarrow VH$ production mode, where $V$ is either a $W$ or a $Z$ boson and $H \rightarrow b\bar{b}$. The previous Higgs boson search in this production mode at CMS used data samples corresponding to integrated luminosities of up to 5.1 fb$^{-1}$ at $\sqrt{s} = 7$ TeV and up to 5.3 fb$^{-1}$ at $\sqrt{s} = 8$ TeV [13]. The results presented here combine the analysis of the 7 TeV data sample in Ref. [13] with an updated analysis of the full 8 TeV data sample corresponding to a luminosity of up to 18.9 fb$^{-1}$.

The following six channels are considered in the search: $W(\mu\nu)H$, $W(e\nu)H$, $W(\tau\nu)H$, $Z(\mu\mu)H$, $Z(\tau\tau)H$, and $Z(\nu\nu)H$, all with the Higgs boson decaying to $b\bar{b}$. Throughout this article the term “lepton” refers only to charged leptons and the symbol $\ell$ is used to refer to both muons and electrons, but not to taus. For the $W(\tau\nu)H$ final state, only the 8 TeV data are included and only taus with 1-prong hadronic decays are explicitly considered; the $\tau$ notation throughout this article refers exclusively to such decays. The leptonic decays of taus in $WH$ processes are implicitly accounted for in the $W(\mu\nu)H$ and $W(e\nu)H$ channels. Backgrounds arise from production of $W$ and $Z$ bosons in association with jets (from gluons and from light- or heavy-flavor quarks), singly and pair-produced top quarks ($tt$), dibosons, and quantum chromodynamics (QCD) multijet processes.

Simulated samples of signal and background events are used to provide guidance in the optimization of the

*Full author list given at the end of the article.

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.
analysis. Control regions in data are selected to adjust the event yields from simulation for the main background processes in order to estimate their contribution in the signal region. These regions also test the accuracy of the modeling of kinematic distributions in the simulated samples.

Upper limits at the 95% C.L. on the $pp \rightarrow VH$ production cross section times the $H \rightarrow bb$ branching fraction are obtained for Higgs boson masses in the 110–135 GeV range. These limits are extracted by fitting the shape of the output distribution of a boosted-decision-tree (BDT) discriminant [14,15]. The results of the fitting procedure allow one to evaluate the presence of a Higgs boson signal over the expectation from the background components. The significance of any excess of events, and the corresponding event yield, is compared with the expectation from a SM Higgs boson signal.

II. DETECTOR AND SIMULATED SAMPLES

A detailed description of the CMS detector can be found elsewhere [16]. 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 $\theta$ 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 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 $\mu$s. 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 $\phi$ is the angle (in radians) of the trajectory of the object in the plane transverse to the direction of the proton beams.

Simulated samples of signal and background events are produced using various Monte Carlo (MC) event generators, with the CMS detector response modeled with GEANT4 [17]. The Higgs boson signal samples are produced using the POWHEG [18] event generator. The MADGRAPH 5.1 [19] generator is used for the diboson, $W + jets$, $Z + jets$, and $t\bar{t}$ samples. The single-top quark samples, including the $tW$, $t$, and $s$-channel processes, are produced with POWHEG and the QCD multijet samples with PYTHIA 6.4 [20]. The production cross sections for the diboson and $t\bar{t}$ samples are rescaled to the cross sections from the next-to-leading-order (NLO) MC@NLO generator [21], while the cross sections for the $W + jets$ and $Z + jets$ samples are rescaled to next-to-next-to-leading-order (NNLO) cross sections calculated using the FEWZ program [22–24]. The default set of parton distribution functions (PDF) used to produce the NLO POWHEG samples is the NLO MSTW2008 set [25], while the leading-order (LO) CTEQ6L1 set [26] is used for the other samples. For parton showering and hadronization the POWHEG and MADGRAPH samples are interfaced with HERWIG++ [27] and PYTHIA, respectively. The PYTHIA parameters for the underlying event description are set to the Z2 tune for the 7 TeV samples and to the Z2* tune for the 8 TeV samples [28]. The TAUOLA [29] library is used to simulate tau decays.

During the 2011 data-taking period the LHC instantaneous luminosity reached up to $3.5 \times 10^{33}$ cm$^{-2}$ s$^{-1}$ and the average number of $pp$ interactions per bunch crossing was approximately 9. During the 2012 period the LHC instantaneous luminosity reached $7.7 \times 10^{33}$ cm$^{-2}$ s$^{-1}$ and the average number of $pp$ interactions per bunch crossing was approximately 21. Additional simulated $pp$ interactions overlapping with the event of interest in the same bunch crossing, denoted as pileup events, are therefore added in the simulated samples to reproduce the pileup distribution measured in data.

III. TRIGGERS

Several triggers are used to collect events consistent with the signal hypothesis in the six channels under consideration.

For the $W(\mu\nu)H$ and $W(\ell\nu)H$ channels, the trigger paths consist of several single-lepton triggers with tight lepton identification. Leptons are also required to be isolated from other tracks and calorimeter energy deposits to maintain an acceptable trigger rate. For the $W(\mu\nu)H$ channel and for the 2011 data, the trigger thresholds for the muon transverse momentum, $p_T$, are in the range of 17 to 24 GeV. The higher thresholds are used for the periods of higher instantaneous luminosity. For the 2012 data the muon trigger $p_T$ threshold for the single-isolated-muon trigger is set at 24 GeV. For both the 2011 and 2012 data, a single-muon trigger with a 40 GeV $p_T$ threshold, but without any isolation requirements, is also used for this channel. The combined single-muon trigger efficiency is $\approx 90\%$ for $W(\mu\nu)H$ events that pass all offline requirements that are described in Sec. V.

For the $W(\ell\nu)H$ channel and for the 2011 data, the electron $p_T$ threshold ranges from 17 to 30 GeV. To maintain acceptable trigger rates during the periods of high instantaneous luminosity, the lower-threshold triggers.
also require two central ($|\eta| < 2.6$) jets, with a $p_T$ threshold in the 25–30 GeV range, and a minimum requirement on the value of an online estimate of the missing transverse energy, $E_T^{\text{miss}}$, in the 15–25 GeV range. $E_T^{\text{miss}}$ is defined online as the magnitude of the vector sum of the transverse momenta of all reconstructed objects identified by a particle-flow algorithm [30,31]. This algorithm combines the information from all CMS subdetectors to identify and reconstruct online individual particles emerging from the proton-proton collisions: charged hadrons, neutral hadrons, photons, muons, and electrons. These particles are then used to reconstruct jets, $E_T^{\text{miss}}$ and hadronic $\tau$-lepton decays, and also to quantify the isolation of leptons and photons. For the 2012 data, the electron trigger uses a 27 GeV threshold on the $p_T$ and no other requirements on jets or $E_T^{\text{miss}}$ are made. The combined efficiency for these triggers for $W(\nu)H$ events to pass the offline selection criteria is $>95\%$.

For the $W(\tau\nu)H$ channel trigger, a 1-prong hadronically decaying tau is required. The $p_T$ of the charged track candidate coming from the tau decay is required to be above 20 GeV and the $p_T$ of the tau (measured from all reconstructed charged and neutral decay products) above 35 GeV. Additionally, the tau is required to be isolated inside an annulus with inner radius $\Delta R = 0.2$ and outer radius $\Delta R = 0.4$, where no reconstructed charged candidates with $p_T > 1.5$ GeV must be found. A further requirement of a minimum of 70 GeV is placed on the $E_T^{\text{miss}}$. The efficiency of this trigger for $W(\tau\nu)H$ events that pass the offline selection criteria is $>90\%$.

The $Z(\mu\mu)H$ channel uses the same single-muon triggers as the $W(\mu\nu)H$ channel. For the $Z(ee)H$ channel, dielectron triggers with lower $p_T$ thresholds, of 17 and 8 GeV, and tight isolation requirements are used. These triggers are nearly 100% efficient for all $Z(\ell\ell)H$ signal events that pass the final offline selection criteria.

For the $Z(\nu\nu)H$ channel, combinations of several triggers are used, all requiring $E_T^{\text{miss}}$ to be above a given threshold. Extra requirements are added to keep the trigger rates manageable as the instantaneous luminosity increases and to reduce the $E_T^{\text{miss}}$ thresholds in order to improve signal acceptance. A trigger with $E_T^{\text{miss}} > 150$ GeV is used for the complete data set in both 2011 and 2012. During 2011 additional triggers that require the presence of two central jets with $p_T > 20$ GeV and $E_T^{\text{miss}}$ thresholds of 80 or 100 GeV, depending on the instantaneous luminosity, were used. During 2012 an additional trigger that required two central jets with $p_T > 30$ GeV and $E_T^{\text{miss}} > 80$ GeV was used. This last trigger was discontinued when the instantaneous luminosity exceeded $3 \times 10^{33}$ cm$^{-2}$ s$^{-1}$ and was replaced by a trigger that required $E_T^{\text{miss}} > 100$ GeV, at least two central jets with vectorial sum $p_T > 100$ GeV and individual $p_T$ above 60 and 25 GeV, and no jet with $p_T > 40$ GeV closer than 0.5 in azimuthal angle to the $E_T^{\text{miss}}$ direction. In order to increase signal acceptance at lower values of $E_T^{\text{miss}}$, triggers that require jets to be identified as coming from $b$ quarks are used. For these triggers, two central jets with $p_T$ above 20 or 30 GeV, depending on the luminosity conditions, are required. It is also required that at least one central jet with $p_T$ above 20 GeV be tagged by the online combined secondary vertex (CSV) $b$-tagging algorithm described in Sec. IV. This online $b$-tagging requirement has an efficiency that is equivalent to that of the tight offline requirement, CSV $> 0.898$, on the value of the output of the CSV discriminant. The $E_T^{\text{miss}}$ is required to be greater than 80 GeV for these triggers. For $Z(\nu\nu)H$ events with $E_T^{\text{miss}} > 130$ GeV, the combined trigger efficiency for $Z(\nu\nu)H$ signal events is near 100% with respect to the offline event reconstruction and selection, described in the next sections. For events with $E_T^{\text{miss}}$ between 100 and 130 GeV the efficiency is 88%.

IV. 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, taus, neutrinos, and jets (including those originating from $b$ quarks). The charged leptons and neutrinos (reconstructed as $E_T^{\text{miss}}$) originate from the vector boson decays. The $b$-quark jets originate from the Higgs boson decays.

The reconstructed interaction vertex with the largest value of $\sum_i p_T^{\text{jet} i}$, where $p_T^{\text{jet} i}$ is the transverse momentum of the $i$th track associated with the vertex, is selected as the primary event vertex. This vertex is used as the reference vertex for all relevant objects in the event, which are reconstructed with the particle-flow 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 identified by a particle-flow-based algorithm and removed from consideration in the event. In addition, the average neutral energy density from pileup interactions is evaluated from particle-flow objects and subtracted from the reconstructed jets in the event and from the summed energy in the isolation cones used for leptons, described below [32]. These pileup-mitigation procedures are applied on an event-by-event basis.

Jets are reconstructed from particle-flow objects using the anti-$k_T$ clustering algorithm [33], with a distance parameter of 0.5, as implemented in the FASTJET package [34,35]. Each jet is required to lie within $|\eta| < 2.5$, 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 [36]. The missing transverse energy vector is calculated offline as the negative of the vectorial sum of transverse momenta.
of all particle-flow objects identified in the event, and the magnitude of this vector is referred to as $E_{T}^{\text{miss}}$ in the rest of this article.

Muons are reconstructed using two algorithms [37]: one in which tracks in the silicon tracker are matched to signals in the muon detectors, and another in which a global track fit is performed, seeded by signals 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 measurements 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 considered in the $|\eta| < 2.4$ range.

Electron reconstruction requires the matching of an energy cluster in the ECAL with a track in the silicon tracker [38]. Identification criteria based on the ECAL shower shape, matching between the track and the ECAL cluster, and consistency with the primary vertex are imposed. Electron identification relies on a multivariate cluster, and consistency with the primary vertex are imposed. Electron identification relies on a multivariate technique that combines observables sensitive to the amount of bremsstrahlung along the electron trajectory, the geometrical and momentum matching between the electron trajectory and associated clusters, as well as shower-shape observables. Additional requirements are imposed to remove electrons produced by photon conversions. In this analysis, electrons are considered in the pseudorapidity range $|\eta| < 2.5$, excluding the $1.44 < |\eta| < 1.57$ transition region between the ECAL barrel and endcap, where electron reconstruction is suboptimal.

Charged leptons from the $W$ 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. If this sum exceeds approximately 10% of the candidate $p_T$, the lepton is rejected; the exact requirement depends on the lepton $\eta$, $p_T$, and flavor. Including the isolation requirement, the total efficiency to reconstruct muons is in the 87%–91% range, depending on $p_T$ and $\eta$. The corresponding efficiency for electrons is in the 81%–98% range.

The hadronically decaying taus are reconstructed using the hadron plus strips (HPS) algorithm [39] which uses charged hadrons and neutral electromagnetic objects (photons) to reconstruct tau decays. Reconstructed taus are required to be in the range $|\eta| < 2.1$. In the first step of reconstruction, charged hadrons are reconstructed using the particle-flow algorithm. Since neutral pions are often produced in hadronic tau decays, the HPS algorithm is optimized to reconstruct neutral pions in the ECAL as objects called “strips.” The strip reconstruction starts by centering one strip on the most energetic electromagnetic particle and then looking for other particles in a window of 0.05 in $\eta$ and 0.20 in $\phi$. Strips satisfying a minimum transverse momentum of $p_T(\text{strip}) > 1$ GeV are combined with the charged hadrons to reconstruct the hadronic tau candidate. In the final step of reconstruction, all charged hadrons and strips are required to be contained within a narrow cone size of $\Delta R = 2.8/p_T(\tau)$, where $p_T(\tau)$ is measured from the reconstructed hadronic tau candidate and is expressed in GeV. Further identification criteria are imposed on the tau candidate to reduce the fraction of electron and muons misidentified as taus. These include the tau candidate passing an antielectron discriminator and an antimuon discriminator. The isolation requirement for taus is that the sum of transverse momenta of particle-flow charged hadron and photon candidates, with $p_T > 0.5$ GeV and within a cone of $\Delta R < 0.5$, be less than 2 GeV. The tau reconstruction efficiency is approximately 50% while the misidentification rate from jets is about 1%.

Jets that originate from the hadronization of $b$ quarks are referred to as “$b$ jets.” The CSV $b$-tagging algorithm [40] is used to identify such jets. The algorithm combines the information about track impact parameters and secondary vertices within jets in a likelihood discriminant to provide separation between $b$ jets and jets originating from light quarks, gluons, or charm quarks. The output of this CSV discriminant has values between zero and one; a jet with a CSV value above a certain threshold is referred to as being “$b$ tagged.” The efficiency to tag $b$ jets and the rate of misidentification of non-$b$ jets depend on the threshold chosen, and are typically parametrized as a function of the $p_T$ and $\eta$ of the jets. These performance measurements are obtained directly from data in samples that can be enriched in $b$ jets, such as $t\bar{t}$ 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 CSV 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.

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 simulations are computed from data and used in the analysis.

V. EVENT SELECTION

The background processes to $VH$ production with $H \rightarrow bb$ are the production of vector bosons in association with one or more jets ($V + \text{jets}$), $t\bar{t}$ production, single-top quark production, diboson production ($VV$), and QCD multijet production. Except for dibosons, these processes have production cross sections that are several orders of
magnitude larger than Higgs boson production. The production cross section for the VZ process, where \(Z \rightarrow b\bar{b}\), is only a few times larger than the VH production cross section, and given the nearly identical final state this process provides a benchmark against which the Higgs boson search strategy can be tested.

The event selection is based on the reconstruction of the vector bosons in their leptonic decay modes and of the Higgs boson decay into two b-tagged jets. Background events are substantially reduced by requiring a significant boost of the \(p_T\) of the vector boson, \(p_T(V)\), or of the Higgs boson [41]. In this kinematic region the V and H bosons recoil away from each other with a large azimuthal opening angle, \(\Delta \phi(V, H)\), between them. For each channel, different \(p_T(V)\) boost regions are selected. Because of different signal and background content, each \(p_T(V)\) region has different sensitivity and the analysis is performed separately in each region. The results from all regions are then combined for each channel. The low-, intermediate-, and high-boost regions for the \(W(e\nu)H\) and \(W(\tau\nu)H\) channels are defined by \(100 < p_T(V) < 130\) GeV, \(130 < p_T(V) < 180\) GeV, and \(p_T(V) > 180\) GeV. For the \(W(\tau\nu)H\) a single \(p_T(V) > 120\) GeV region is considered. For the \(Z(\ell\ell)H\) channels, the low- and high-boost regions are defined by \(50 < p_T(V) < 100\) GeV and \(p_T(V) > 100\) GeV. For the \(Z(\nu\nu)H\) channel \(E_T^{\text{miss}}\) is used to define the low-, intermediate-, and high-boost \(p_T(V)\) regions as \(100 < E_T^{\text{miss}} < 130\) GeV, \(130 < E_T^{\text{miss}} < 170\) GeV, and \(E_T^{\text{miss}} > 170\) GeV, respectively. In the rest of the article the term “boost region” is used to refer to these \(p_T(V)\) regions.

Candidate \(W \rightarrow \ell\nu\) decays are identified by requiring the presence of a single-isolated lepton and additional missing transverse energy. Muons are required to have \(p_T > 20\) GeV; the corresponding thresholds for electrons and taus are 30 and 40 GeV, respectively. For the \(W(\ell\nu)H\) and \(W(\tau\nu)H\) channels, \(E_T^{\text{miss}}\) is required to be \(>45\) and \(>80\) GeV, respectively, to reduce contamination from QCD multijet processes. To further reduce this contamination, it is also required for the \(W(\ell\nu)H\) channels that the azimuthal angle between the \(E_T^{\text{miss}}\) direction and the lepton be \(<\pi/2\), and that the lepton isolation for the low-boost region be tighter.

Candidate \(Z \rightarrow \ell\ell\) decays 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.

The identification of \(Z \rightarrow \nu\bar{\nu}\) decays requires the \(E_T^{\text{miss}}\) in the event to be within the boost regions described above. The QCD multijet background is reduced to negligible levels in this channel when requiring that the \(E_T^{\text{miss}}\) does not originate from mismeasured jets. To that end three event requirements are made. First, for the high-boost region, \(\Delta \phi(E_T^{\text{miss}}, \text{jet}) > 0.5\) radians requirement is applied on the azimuthal angle between the \(E_T^{\text{miss}}\) direction and the closest jet with \(|\eta| < 2.5\) and \(p_T > 20\) GeV for the 7 TeV analysis or \(p_T > 25\) GeV for the 8 TeV analysis (where more pileup interactions are present). For the low- and intermediate-boost regions the requirement is tightened to \(\Delta \phi(E_T^{\text{miss}}, \text{jet}) > 0.7\) radians. The second requirement is that the azimuthal angle between the missing transverse energy direction as calculated from charged tracks only (with \(p_T > 0.5\) GeV and \(|\eta| < 2.5\) and the \(E_T^{\text{miss}}\) direction, \(\Delta \phi(E_T^{\text{miss}}, E_T^{\text{miss}}(\text{tracks}))\), should be smaller than 0.5 radians. The third requirement is made for the low-boost region where the \(E_T^{\text{miss}}\) significance (defined as the ratio between the \(E_T^{\text{miss}}\) and the square root of the total transverse energy in the calorimeter, measured in GeV) should be larger than 3. To reduce background events from \(t\bar{t}\) and WZ production in the \(W(\ell\nu)H\), \(W(\tau\nu)H\), and \(Z(\nu\nu)H\) channels, events with an additional number of isolated leptons, \(N_{\ell\ell} > 0\), with \(p_T > 20\) GeV are rejected.

The reconstruction of the \(H \rightarrow b\bar{b}\) decay proceeds by selecting the pair of jets in the event, each with \(|\eta| < 2.5\) and \(p_T\) above a minimum threshold, for which the value of the magnitude of the vectorial sum of their transverse momenta, \(p_T(\ell\ell))\), is the highest. These jets are then also required to be tagged by the CSV algorithm, with the value of the CSV discriminator above a minimum threshold. The background from \(V + \) jets and diboson production is reduced significantly when the \(b\)-tagging requirements are applied and processes where the two jets originate from genuine b quarks dominate the final selected data sample.

After all event selection criteria described in this section are applied, the dijet invariant mass resolution of the two \(b\) jets from the Higgs 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. The Higgs boson mass resolution is further improved by applying multivariate regression techniques similar to those used at the CDF experiment [42]. An additional correction, beyond the standard CMS jet energy corrections, is computed for individual \(b\) jets in an attempt to recalibrate to the true \(b\)-quark energy. For this purpose, a specialized BDT is trained on simulated \(H \rightarrow b\bar{b}\) signal 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 \(\Delta R\) distance between the lepton and the jet directions, and the transverse momentum of the lepton relative to the jet direction. For the \(Z(\ell\ell)H\) channels the \(E_T^{\text{miss}}\) in the event
and the azimuthal angle between the $E_T^{\text{miss}}$ and each jet are also considered in the regression. The output of the BDT regression is the corrected jet energy. The average improvement on the mass resolution, measured on simulated signal samples, when the corrected jet energies are used is $\approx 15\%$, resulting in an increase in the analysis sensitivity of 10%–20%, depending on the specific channel. This improvement is shown in Fig. 1 for simulated samples of $Z(\ell\ell)H(bb)$ events where the improvement in resolution is $\approx 25\%$. The validation of the regression technique in data is done with samples of $Z \to \ell\ell$ events with two $b$-tagged jets and in $t\bar{t}$-enriched samples in the lepton + jets final state. In the $Z \to \ell\ell$ case, when the jets are corrected by the regression procedure, the $p_T$ balance distribution, between the $Z$ boson, reconstructed from the leptons, and the $b$-tagged dijet system is improved to be better centered at zero and narrower than when the regression correction is not applied. In the $t\bar{t}$-enriched case, the reconstructed top quark mass distribution is closer to the nominal top quark mass and also narrower than when the correction is not applied. In both cases the distributions for data and the simulated samples are in very good agreement after the regression correction is applied.

The signal region is defined by events that satisfy the vector boson and Higgs boson reconstruction criteria described above together with the requirements listed in Table I. In the final stage of the analysis, to better separate signal from background under different Higgs boson mass hypotheses, an event BDT discriminant is trained separately at each mass value using simulated samples for signal and all background processes. The training of this BDT is performed with all events in the signal region. The set of event input variables used, listed in Table II, 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 $m(jj)$, the number of additional jets ($N_q$), the value of CSV for the Higgs boson daughter with the second largest CSV value (CSV$_\text{min}$), and the distance between Higgs boson daughters [$\Delta R(jj)$]. It has been suggested that variables related to techniques that study in more detail the substructure of jets could help improve the sensitivity of the $H \to b\bar{b}$ searches [41]. In this analysis, several combinations of such variables were considered as additional inputs to the BDT discriminant. However they did not yield significant gains in sensitivity and are not included in the final training used.

A fit is performed to the shape of the output distribution of the event BDT discriminant to search for events resulting from Higgs boson production. Before testing all events through this final discriminant, events are classified based on where they fail in the output distributions of several other background-specific BDT discriminants that are trained to discern signal from individual background processes. This technique, similar to the one used by the CDF Collaboration [44], divides the samples into four distinct subsets that are enriched in $t\bar{t}$, $V +$ jets, dibosons, and $VH$. The increase in the analysis sensitivity from using this technique in the $Z(\nu\nu)H$ and $W(\nu\nu)H$ channels is 5%–10%. For the $Z(\ell\ell)H$ channel the improvement is not as large and therefore the technique is not used for that case. The technique is also not used in the $W(\tau\nu)H$ channel because of the limited size of the simulated event samples available for training multiple BDT discriminants. The first background-specific BDT discriminant is trained to separate $t\bar{t}$ from $VH$, the second one is trained to separate $V +$ jets from $VH$, and the third one separates diboson events from $VH$. The output distributions of the background-specific BDTs are used to separate events in four subsets: those that fail a requirement on the $t\bar{t}$ BDT are classified as $t\bar{t}$-like events, those that pass the $t\bar{t}$ BDT requirement but fail a requirement on the $V +$ jets BDT are classified as $V +$ jets-like events, those that pass the $V +$ jets BDT requirement but fail the requirement on the diboson BDT are classified as dibosonlike events and, finally, those that pass all BDT requirements are considered $VH$-enriched events. The events in each subset are then run through the final event BDT discriminant and the resulting distribution, now composed of four distinct subsets of events, is used as input to the fitting procedure.

As a validation of the multivariate approach to this analysis, these BDT discriminants are also trained to find diboson signals ($ZZ$ and $WZ$, with $Z \to b\bar{b}$) rather than the $VH$ signal. The event selection used in this case is identical to that used for the $VH$ search.

FIG. 1 (color online). Dijet invariant mass distribution 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 Bukin function [66] is fit to the distribution and the fitted width of the core of the distribution is displayed in the figure.

S. CHATRCHYAN et al. PHYSICAL REVIEW D 89, 012003 (2014)
As a cross-check to the BDT-based analysis, a simpler analysis is done by performing a fit to the shape of the dijet invariant mass distribution of the two jets associated with the reconstructed Higgs boson, \( m(jj) \). The event selection for this analysis is more restrictive than the one used in the BDT analysis and is optimized for sensitivity in this single variable. Table III lists the event selection of the \( m(jj) \) analysis. Since the diboson background also exhibits a peak in the \( m(jj) \) distribution from \( Z \) bosons that decay into \( b \) quark pairs, the distribution is also used to measure

### Table I
Selection criteria that define the signal region. Entries marked with “⋯⋯” indicate that the variable is not used. 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 \( p_T(j_1) \) and \( p_T(j_2) \), respectively. The transverse momentum of the leading tau track is \( p_T(\tau) \). The values listed for kinematic variables are in units of GeV, and for angles in units of radians.

<table>
<thead>
<tr>
<th>Variable</th>
<th>( W(\ell \nu)H )</th>
<th>( W(\nu \nu)H )</th>
<th>( Z(\ell \ell)H )</th>
<th>( Z(\nu \nu)H )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_{\ell\ell} )</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>[75–105]</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
</tr>
<tr>
<td>( p_T(j_1) )</td>
<td>( &gt;30 )</td>
<td>( &gt;30 )</td>
<td>( &gt;20 )</td>
<td>( &gt;60 )</td>
</tr>
<tr>
<td>( p_T(j_2) )</td>
<td>( &gt;30 )</td>
<td>( &gt;30 )</td>
<td>( &gt;20 )</td>
<td>( &gt;30 )</td>
</tr>
<tr>
<td>( p_T(j) )</td>
<td>( &gt;100 )</td>
<td>( &gt;120 )</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>( &gt;130 )</td>
</tr>
<tr>
<td>( m(jj) )</td>
<td>( &lt;250 )</td>
<td>( &lt;250 )</td>
<td>[40–250]</td>
<td>&lt;250</td>
</tr>
<tr>
<td>( E_T^{\text{miss}} )</td>
<td>( &gt;45 )</td>
<td>( &gt;80 )</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
</tr>
<tr>
<td>( p_T(\tau) )</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>[100–130]</td>
<td>[130–170]</td>
</tr>
<tr>
<td>( p_T ) (track)</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>[100–130]</td>
<td>[130–170]</td>
</tr>
<tr>
<td>( \text{CSV}_{\text{max}} )</td>
<td>( &gt;0.40 )</td>
<td>( &gt;0.40 )</td>
<td>[0.50]</td>
<td>[0.244]</td>
</tr>
<tr>
<td>( \text{CSV}_{\text{min}} )</td>
<td>( &gt;0.40 )</td>
<td>( &gt;0.40 )</td>
<td>&gt;0.244</td>
<td>&gt;0.244</td>
</tr>
<tr>
<td>( N_{aj} )</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
</tr>
<tr>
<td>( N_{at} )</td>
<td>=0</td>
<td>=0</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>=0</td>
</tr>
<tr>
<td>( \Delta \phi(V, H) )</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>&gt;2.0</td>
</tr>
<tr>
<td>( \Delta \phi(E_T^{\text{miss}}, \text{jet}) )</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>( &gt;0.7 )</td>
</tr>
<tr>
<td>( \Delta \phi(E_T^{\text{miss}}, E_T^{\text{miss}}(\text{tracks})) )</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>&lt;0.5</td>
</tr>
<tr>
<td>( E_T^{\text{miss}} ) significance</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>( &gt;3 )</td>
</tr>
<tr>
<td>( \Delta \phi(E_T^{\text{miss}}, \ell) )</td>
<td>&lt;( \pi/2 )</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
<td>⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯</td>
</tr>
</tbody>
</table>

### Table II
Variables used in the training of the event BDT discriminant. Jets are counted as additional jets if they satisfy the following: \( p_T > 20 \) GeV and \( |\eta| < 4.5 \) for \( W(\ell \nu)H \), \( p_T > 20 \) GeV and \( |\eta| < 2.5 \) for \( Z(\ell \ell)H \), and \( p_T > 25 \) GeV and \( |\eta| < 4.5 \) for \( Z(\nu \nu)H \).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_T(j_1), p_T(j_2) )</td>
<td>Transverse momentum of each Higgs boson daughter</td>
</tr>
<tr>
<td>( m(jj) )</td>
<td>Dijet invariant mass</td>
</tr>
<tr>
<td>( p_T(V) )</td>
<td>Vector boson transverse momentum (or ( E_T^{\text{miss}} ))</td>
</tr>
<tr>
<td>( N_{aj} )</td>
<td>Number of additional jets (see caption)</td>
</tr>
<tr>
<td>( \text{CSV}_{\max} )</td>
<td>CSV for the Higgs boson daughter with largest CSV value</td>
</tr>
<tr>
<td>( \text{CSV}_{\min} )</td>
<td>CSV for the Higgs boson daughter with second largest CSV value</td>
</tr>
<tr>
<td>( \Delta \phi(V, H) )</td>
<td>Azimuthal angle between ( V ) and ( E_T^{\text{miss}} ) and dijet</td>
</tr>
<tr>
<td>( \Delta \eta(jj) )</td>
<td>Difference in ( \eta ) between Higgs boson daughters</td>
</tr>
<tr>
<td>( \Delta R(jj) )</td>
<td>Distance in ( \eta - \phi ) between Higgs boson daughters</td>
</tr>
<tr>
<td>( \delta_{\text{pull}} )</td>
<td>Color pull angle [43]</td>
</tr>
<tr>
<td>( \Delta \phi(E_T^{\text{miss}}, \text{jet}) )</td>
<td>Azimuthal angle between ( E_T^{\text{miss}} ) and the closest jet [only for ( Z(\nu \nu)H )]</td>
</tr>
<tr>
<td>( \text{max CSV}_{aj} )</td>
<td>Maximum CSV of the additional jets in an event [only for ( Z(\nu \nu)H ) and ( W(\ell \nu)H )]</td>
</tr>
<tr>
<td>( \text{min CSV}_{aj} )</td>
<td>Minimum distance between an additional jet and the Higgs boson candidate [only for ( Z(\nu \nu)H ) and ( W(\ell \nu)H )]</td>
</tr>
<tr>
<td>( \text{InV} ) mass</td>
<td>Invariant mass of the VH system [only for ( Z(\ell \ell)H )]</td>
</tr>
<tr>
<td>( \text{Cosine of the angle between the direction of the V boson in the rest frame of the VH system and the direction of the VH system in the laboratory frame} ) [only for ( Z(\ell \ell)H )]</td>
<td></td>
</tr>
<tr>
<td>( \text{Cosine of the angle between the direction of one of the leptons in the rest frame of the Z boson and the direction of the Z boson in the laboratory frame} ) [only for ( Z(\ell \ell)H )]</td>
<td></td>
</tr>
<tr>
<td>( \text{Cosine of the angle between the direction of one of the jets in the rest frame of the reconstructed Higgs boson and the direction of the reconstructed Higgs boson in the laboratory frame} ) [only for ( Z(\ell \ell)H )]</td>
<td></td>
</tr>
</tbody>
</table>
the consistency of the diboson rate with the expectation from the standard model. A consistent rate measurement would support the validity of the estimate of the background processes in the Higgs boson search.

VI. BACKGROUND CONTROL REGIONS

Appropriate control regions are identified in data and used to validate the simulation modeling of the distributions used as input to the BDT discriminants, and to obtain scale factors used to adjust the simulation event yield estimates for the most important background processes: production of W and Z bosons in association with jets and $t\bar{t}$ production. For the W and Z backgrounds the control regions are defined such that they are enriched in either heavy-flavor (HF) or light-flavor (LF) jets. Furthermore, these processes are split according to how many of the two jets selected in the Higgs boson reconstruction originate from $b$ quarks, and separate scale factors are obtained for each case. The notation used is $V + u ds c g$ for the case where none of the jets originate from a $b$ quark, $V + b$ for the case where only one of the jets is from a $b$ quark, and $V + b \bar{b}$ for the case where both jets originate from $b$ quarks.

To obtain the scale factors by which the simulated event yields are adjusted, a set of binned likelihood fits is simultaneously performed to CSV distributions of jets for events in the control regions. These fits are done separately for each channel. Several other distributions are also fit to

<table>
<thead>
<tr>
<th>Variable</th>
<th>W + LF</th>
<th>$t\bar{t}$</th>
<th>W + HF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T(V)$</td>
<td>[100–150]</td>
<td>[100–130]</td>
<td>[100–150]</td>
</tr>
<tr>
<td>$m_{\ell\ell}$</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$p_T(V)$</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$p_T(j_1)$</td>
<td>&gt;30</td>
<td>&gt;30</td>
<td>&gt;30</td>
</tr>
<tr>
<td>$p_T(j_2)$</td>
<td>&gt;30</td>
<td>&gt;30</td>
<td>&gt;30</td>
</tr>
<tr>
<td>$p_T(j)$</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>$m_{jj}$</td>
<td>&gt;250</td>
<td>&lt;250</td>
<td>&lt;250, $\not E_T$ [90–150]</td>
</tr>
<tr>
<td>CSV$_{\text{max}}$</td>
<td>$\in$ [0.244–0.898]</td>
<td>&gt;0.898</td>
<td>&gt;0.898</td>
</tr>
<tr>
<td>$N_{aj}$</td>
<td>&lt;2</td>
<td>&gt;1</td>
<td>=0</td>
</tr>
<tr>
<td>$N_{sf}$</td>
<td>=0</td>
<td>=0</td>
<td>=0</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$</td>
<td>&gt;45</td>
<td>&gt;45</td>
<td>&gt;45</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$ signif.</td>
<td>&gt;2.0($\mu$), &gt;3.0(e)</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$W(\ell\nu)H$</td>
<td>[100–150]</td>
<td>[100–130]</td>
<td>[100–150]</td>
</tr>
<tr>
<td>$W(\tau\nu)H$</td>
<td>[100–150]</td>
<td>[100–130]</td>
<td>[100–150]</td>
</tr>
<tr>
<td>$Z(\ell\ell)H$</td>
<td>[50–100]</td>
<td>[50–100]</td>
<td>[50–100]</td>
</tr>
<tr>
<td>$Z(\nu\nu)H$</td>
<td>[100–150]</td>
<td>[100–130]</td>
<td>[100–150]</td>
</tr>
</tbody>
</table>

Table IV. Definition of the control regions for the $W(\ell\nu)H$ and the $W(\tau\nu)H$ channels. Entries marked with “...” indicate that the variable is not used. The same selection is used for all boost regions. Here, LF and HF refer to light- and heavy-flavor jets. The values listed for kinematic variables are in units of GeV. Because of the limited size of the simulated samples, the scale factors derived in these control regions are also applied to the $W(\nu\nu)H$ channel.

Table V. Definition of the control regions for the $Z(\ell\ell)H$ channel. Entries marked with “...” indicate that the variable is not used. The same selection is used for both the low- and high-boost regions. The values listed for kinematic variables are in units of GeV.
TABLE VI. Definition of the control regions for the $Z(\nu\nu)H$ channel. Entries marked with "\ldots" indicate that the variable is not used. If different, the entries in square brackets indicate the selection for the different boost regions as defined by the $E_T^{\text{miss}}$ in the first row of the table. The values listed for kinematic variables are in units of GeV, and for angles in units of radians.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$Z + LF$</th>
<th>$Z + HF$</th>
<th>$t\bar{t}$</th>
<th>$W + LF$</th>
<th>$W + HF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_T^{\text{miss}}$</td>
<td>[100–130]</td>
<td>[130–170]</td>
<td>[100–130]</td>
<td>[130–170]</td>
<td>[100–130]</td>
</tr>
<tr>
<td>$p_T^{(j)}$</td>
<td>&gt;60</td>
<td>&gt;60</td>
<td>&gt;60</td>
<td>&gt;60</td>
<td>&gt;60</td>
</tr>
<tr>
<td>$p_T^{(\ell)}$</td>
<td>&gt;30</td>
<td>&gt;30</td>
<td>&gt;30</td>
<td>&gt;30</td>
<td>&gt;30</td>
</tr>
<tr>
<td>$m^{(jj)}$</td>
<td>&lt;250</td>
<td>&lt;250, $\notin$ [100–140]</td>
<td>&lt;250</td>
<td>&lt;250, $\notin$ [100–140]</td>
<td>&lt;250</td>
</tr>
<tr>
<td>$CSV_{\text{max}}$</td>
<td>[0.244–0.898]</td>
<td>&gt;0.679</td>
<td>&gt;0.898</td>
<td>[0.244–0.898]</td>
<td>&gt;0.679</td>
</tr>
<tr>
<td>$N_{\ell}$</td>
<td>$\ldots$</td>
<td>&gt;0.244</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>&gt;0.244</td>
</tr>
<tr>
<td>$N_0$</td>
<td>$\ldots$</td>
<td>&gt;0.244</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>&gt;0.244</td>
</tr>
<tr>
<td>$\Delta\phi(V, H)$</td>
<td>$\ldots$</td>
<td>&gt;2.0</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>&gt;2.0</td>
</tr>
<tr>
<td>$\Delta\phi(E_T^{\text{miss}}, \text{jet})$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$ significance</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
</tbody>
</table>

TABLE VII. Data/MC scale factors for 8 TeV data derived from the control regions, where the first quoted uncertainty is statistical and the second is systematic. The muon and electron channels in $Z(\ell\ell)H$ and $W(\ell\nu)H$ are simultaneously fit to determine average scale factors. For the $Z(\ell\ell)H$ channel only four scale factors are derived, valid for both the low and high $p_T^{(V)}$ boost regions. Because of the limited size of the simulated event samples, the scale factors obtained for the $W(\ell\nu)H$ channels are also applied to the $W(\tau\nu)H$ channel.

<table>
<thead>
<tr>
<th>Process</th>
<th>$W(\ell\nu)H$</th>
<th>$Z(\ell\ell)H$</th>
<th>$Z(\nu\nu)H$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low $p_T^{(V)}$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$W + udscg$</td>
<td>1.03 ± 0.01 ± 0.05</td>
<td>⋮</td>
<td>0.83 ± 0.02 ± 0.04</td>
</tr>
<tr>
<td>$W + b$</td>
<td>2.22 ± 0.25 ± 0.20</td>
<td>⋮</td>
<td>2.30 ± 0.21 ± 0.11</td>
</tr>
<tr>
<td>$W + bb$</td>
<td>1.58 ± 0.26 ± 0.24</td>
<td>⋮</td>
<td>0.85 ± 0.24 ± 0.14</td>
</tr>
<tr>
<td>$Z + udscg$</td>
<td>⋮</td>
<td>1.11 ± 0.04 ± 0.06</td>
<td>1.24 ± 0.03 ± 0.09</td>
</tr>
<tr>
<td>$Z + b$</td>
<td>⋮</td>
<td>1.59 ± 0.07 ± 0.08</td>
<td>2.06 ± 0.06 ± 0.09</td>
</tr>
<tr>
<td>$Z + bb$</td>
<td>⋮</td>
<td>0.98 ± 0.10 ± 0.08</td>
<td>1.25 ± 0.05 ± 0.11</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>1.03 ± 0.01 ± 0.04</td>
<td>1.10 ± 0.05 ± 0.06</td>
<td>1.01 ± 0.02 ± 0.04</td>
</tr>
<tr>
<td><strong>Intermediate $p_T^{(V)}$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$W + udscg$</td>
<td>1.02 ± 0.01 ± 0.07</td>
<td>⋮</td>
<td>0.93 ± 0.02 ± 0.04</td>
</tr>
<tr>
<td>$W + b$</td>
<td>2.90 ± 0.26 ± 0.20</td>
<td>⋮</td>
<td>2.08 ± 0.20 ± 0.12</td>
</tr>
<tr>
<td>$W + bb$</td>
<td>1.30 ± 0.23 ± 0.14</td>
<td>⋮</td>
<td>0.75 ± 0.26 ± 0.11</td>
</tr>
<tr>
<td>$Z + udscg$</td>
<td>⋮</td>
<td>⋮</td>
<td>1.19 ± 0.03 ± 0.07</td>
</tr>
<tr>
<td>$Z + b$</td>
<td>⋮</td>
<td>⋮</td>
<td>2.30 ± 0.07 ± 0.08</td>
</tr>
<tr>
<td>$Z + bb$</td>
<td>⋮</td>
<td>⋮</td>
<td>1.11 ± 0.06 ± 0.12</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>1.02 ± 0.01 ± 0.15</td>
<td>⋮</td>
<td>0.99 ± 0.02 ± 0.03</td>
</tr>
<tr>
<td><strong>High $p_T^{(V)}$</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$W + udscg$</td>
<td>1.04 ± 0.01 ± 0.07</td>
<td>⋮</td>
<td>0.93 ± 0.02 ± 0.03</td>
</tr>
<tr>
<td>$W + b$</td>
<td>2.46 ± 0.33 ± 0.22</td>
<td>⋮</td>
<td>2.12 ± 0.22 ± 0.10</td>
</tr>
<tr>
<td>$W + bb$</td>
<td>0.77 ± 0.25 ± 0.08</td>
<td>⋮</td>
<td>0.71 ± 0.25 ± 0.15</td>
</tr>
<tr>
<td>$Z + udscg$</td>
<td>⋮</td>
<td>1.11 ± 0.04 ± 0.06</td>
<td>1.17 ± 0.02 ± 0.08</td>
</tr>
<tr>
<td>$Z + b$</td>
<td>⋮</td>
<td>1.59 ± 0.07 ± 0.08</td>
<td>2.13 ± 0.05 ± 0.07</td>
</tr>
<tr>
<td>$Z + bb$</td>
<td>⋮</td>
<td>0.98 ± 0.10 ± 0.08</td>
<td>1.12 ± 0.04 ± 0.10</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>1.00 ± 0.01 ± 0.11</td>
<td>1.10 ± 0.05 ± 0.06</td>
<td>0.99 ± 0.02 ± 0.03</td>
</tr>
</tbody>
</table>
verify consistency. These scale factors account not only for cross section discrepancies, but also for potential residual differences in physics object selection. Therefore, separate scale factors are used for each background process in the different channels. The uncertainties in the scale factor determination include two components: the statistical uncertainty due to the finite size of the samples and the systematic uncertainty. The latter is obtained by subtracting, in quadrature, the statistical component from the full uncertainty which includes the effect of various sources of systematic uncertainty such as $b$ tagging, jet energy scale, and jet energy resolution.

Tables IV, V, and VI list the selection criteria used to define the control regions for the $W(\ell\nu)H$, $Z(\ell\ell)H$, and $Z(\nu\nu)H$ channels, respectively. Because of the limited size of the simulated event samples the scale factors obtained for the $W(\ell\nu)H$ channels are applied to the $W(\tau\nu)H$ channel. Table VII summarizes the fit results for all channels for the 8 TeV data. The scale factors are found to be close to unity for all processes except for $V + b$ for which the scale factors are consistently found to be close to 2. In this case, most of the excess occurs in the region of low CSV$_{\text{min}}$ values in which events with two displaced vertices are found relatively close to each other, within a distance.

![Figure 2](color online). 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 VII. Top left: dijet $p_T$ distribution in the $Z + b$ control region for the $Z(ee)H$ channel. Top right: $p_T$ distribution in the $t\bar{t}$ control region for the $W(\mu\nu)H$ channel. Bottom left: CSV$_{\text{min}}$ distribution for the $W + HF$ high-boost control region for the $Z(\nu\nu)H$ channel. Bottom right: $E_T^{\text{miss}}$ distribution for the $Z + HF$ high-boost control region for the $Z(\nu\nu)H$ channel. The bottom inset in each figure shows the ratio of the number of events observed in data to that of the Monte Carlo prediction for signal and backgrounds.
ΔR < 0.5 defined by the directions of their displacement trajectories with respect to the primary vertex. This discrepancy is interpreted as arising mainly from mismodeling in the generator parton shower of the process of gluon-splitting to $b$-quark pairs. In this process the dominant contribution typically contains a low-$p_T$ $b$ quark that can end up not being reconstructed as a jet above the $p_T$ threshold used in the analysis, or that is merged with the jet from the more energetic $b$ quark. These discrepancies are consistent with similar observations in other studies of the production of vector bosons in association with heavy-flavor quarks by the ATLAS and CMS experiments [45–47].

Figures 2 and 3 show examples of distributions for variables in the simulated samples and in data for different control regions and for different channels. The scale factors described above have been applied to the corresponding simulated samples.

VII. UNCERTAINTIES

The systematic uncertainties that affect the results presented in this article are listed in Table VIII and are described in more detail below.

The uncertainty in the CMS luminosity measurement is estimated to be 2.2% for the 2011 data [48] and 2.6% for...
the 2012 data [49]. Muon and electron trigger, reconstruction, and identification efficiencies are determined in data from samples of leptonic Z-boson decays. The uncertainty on the event yields resulting from the trigger efficiency estimate is 2% per lepton and the uncertainty on the identification efficiency is also 2% per lepton. The parameters describing the $Z(\nu \nu)H$ trigger efficiency turn-on curve have been varied within their statistical uncertainties and also estimated for different assumptions on the methods used to derive the efficiency. This results in an event yield uncertainty of about 3%.

The jet energy scale is varied within its uncertainty as a function of jet $p_T$ and $\eta$. The efficiency of the analysis selection is recomputed to assess the variation in event yields. Depending on the process, a 2%–3% yield variation is found. The effect of the uncertainty on the jet energy resolution is evaluated by smearing the jet energies according to the measured uncertainty. Depending on the process, a 3%–6% variation in event yields is obtained. The uncertainties in the jet energy scale and resolution also have an effect on the shape of the BDT output distribution. 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. An uncertainty of 3% is assigned to the event yields of all processes in the $W(\ell\nu)H$ and $Z(\nu\nu)H$ channels due to the uncertainty related to the missing transverse energy estimate.

Data/MC $b$-tagging scale factors 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: 3% per $b$-quark tag, 6% per charm-quark tag, and 15% per mistagged jet (originating from gluons and light $u$, $d$, or $s$ quarks) [40]. These translate into yield uncertainties in the 3%–15% range, depending on the channel and the specific process. The shape of the BDT output distribution is also affected by the shape of the CSV distributions and an uncertainty is assigned according to a range of variations of the CSV distributions.

The total $VH$ signal cross section has been calculated to NNLO accuracy, and the total theoretical uncertainty is $\approx 4\%$ [50], including the effect of scale variations and PDF uncertainties [25,51–54]. This analysis is performed in the boosted regime, and differences in the $p_T$ spectrum of the $V$ and $H$ bosons between data and MC introduce systematic effects in the signal acceptance and efficiency estimates. Two calculations are available that evaluate the NLO electroweak (EW) [55–57] and NNLO QCD [58] corrections to $VH$ production in the boosted regime. Both the electroweak and QCD corrections are applied to the signal samples. The estimated uncertainties of the NLO electroweak corrections are 2% for both the $ZH$ and $WH$ production processes. The estimate for the NNLO QCD correction results in an uncertainty of 5% for both the $ZH$ and $WH$ production processes.

The uncertainty in the background event yields estimated from data is approximately 10%. For $V + jets$, the difference between the shape of the BDT output
distribution for events generated with the MADGRAPH and the HERWIG++ 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 one obtained from the nominal MADGRAPH samples and those obtained from the POWHEG and MC@NLO [59] generators are considered as shape systematic uncertainties.

An uncertainty of 15% is assigned to the event yields obtained from simulation 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 [60,61]. The limited number of MC simulated events is also taken into account as a source of uncertainty.

The combined effect of the systematic uncertainties results in an increase of about 15% on the expected upper limit on the Higgs boson production cross section and in a reduction of 15% on the expected significance of an observation when the Higgs boson is present in the data at the predicted standard model rate.

VIII. RESULTS

Results are obtained from combined signal and background binned likelihood fits to the shape of the output distribution of the BDT discriminants. These are trained separately for each channel and for each Higgs boson mass hypothesis in the 110–135 GeV range. In the simultaneous fit to all channels, in all boost regions, the BDT shape and normalization for signal and for each background component are allowed to vary within the systematic and statistical uncertainties described in Sec. VII. These uncertainties are treated as independent nuisance parameters in the fit. All nuisance parameters, including the scale factors described in Sec. VI, are adjusted by the fit.

In total 14 BDT distributions are considered. Figure 4 shows an example of these distributions after the fit for the high-boost region of the \( Z(\nu\nu)H \) channel, for the \( m_H = 125 \) GeV mass hypothesis. The four partitions in the left panel correspond to the subsets enriched in \( t\bar{t}, V + \text{jets}, \text{diboson}, \text{and } VH \text{ production}, \) as described in Sec. V. The right panel shows the rightmost, \( VH \)-enriched, partition in more detail. For completeness, all 14 BDT distributions used in the fit are shown in Figs. 10–14 in the Appendix. Table IX lists, for partial combinations of channels, the total number of events in the four highest bins of their corresponding BDT for the expected backgrounds, for the 125 GeV SM Higgs boson signal, and for data. An excess compatible with the presence of the SM Higgs boson is observed. Figure 5 combines the BDT outputs 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 standard model Higgs boson.

The results of all channels, for all boost regions and for the 7 and 8 TeV data, are combined to obtain 95%
TABLE IX. The total number of events for partial combinations of channels in the four highest bins of their corresponding BDT for the expected backgrounds (B), for the 125 GeV SM Higgs boson VH signal (S), and for data. Also shown is the signal-to-background ratio (S/B).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$V + b\bar{b}$</td>
<td>25.2</td>
<td>22.4</td>
<td>15.9</td>
<td>4.3</td>
<td>158.6</td>
<td>36.2</td>
<td>177.3</td>
<td>98.3</td>
</tr>
<tr>
<td>$V + b$</td>
<td>3.1</td>
<td>2.9</td>
<td>9.6</td>
<td>1.2</td>
<td>95.8</td>
<td>14.6</td>
<td>84.7</td>
<td>58.3</td>
</tr>
<tr>
<td>$V + udscg$</td>
<td>4.5</td>
<td>8.5</td>
<td>10.0</td>
<td>2.5</td>
<td>62.3</td>
<td>8.7</td>
<td>57.6</td>
<td>31.0</td>
</tr>
<tr>
<td>$\ell\ell$</td>
<td>113.2</td>
<td>106.5</td>
<td>50.3</td>
<td>22.6</td>
<td>107.0</td>
<td>6.9</td>
<td>153.8</td>
<td>87.4</td>
</tr>
<tr>
<td>Single-top quark</td>
<td>24.1</td>
<td>20.3</td>
<td>14.7</td>
<td>7.4</td>
<td>2.9</td>
<td>0.4</td>
<td>54.5</td>
<td>20.1</td>
</tr>
<tr>
<td>$VV(udscg)$</td>
<td>0.3</td>
<td>1.3</td>
<td>1.2</td>
<td>0.2</td>
<td>2.4</td>
<td>0.4</td>
<td>2.3</td>
<td>1.5</td>
</tr>
<tr>
<td>$VZ(b\bar{b})$</td>
<td>1.1</td>
<td>1.4</td>
<td>2.3</td>
<td>1.1</td>
<td>11.0</td>
<td>2.7</td>
<td>9.5</td>
<td>6.9</td>
</tr>
<tr>
<td>Total backgrounds</td>
<td>171.7</td>
<td>163.4</td>
<td>104.1</td>
<td>39.4</td>
<td>439.8</td>
<td>69.8</td>
<td>539.7</td>
<td>303.5</td>
</tr>
<tr>
<td>VH</td>
<td>3.0</td>
<td>6.0</td>
<td>8.3</td>
<td>1.4</td>
<td>5.5</td>
<td>6.3</td>
<td>8.5</td>
<td>8.5</td>
</tr>
<tr>
<td>Data</td>
<td>185</td>
<td>182</td>
<td>128</td>
<td>35</td>
<td>425</td>
<td>77</td>
<td>529</td>
<td>322</td>
</tr>
<tr>
<td>S/B (%)</td>
<td>1.7</td>
<td>3.7</td>
<td>8.0</td>
<td>3.4</td>
<td>1.3</td>
<td>9.0</td>
<td>1.6</td>
<td>2.8</td>
</tr>
</tbody>
</table>

Confidence level (C.L.) upper limits on the product of the VH production cross section times the $H \rightarrow b\bar{b}$ branching fraction, with respect to the expectations for a standard model Higgs boson $(\sigma/\sigma_{SM})$. At each mass point the observed limit, the median expected limit, and the 1 and 2 standard deviation bands are calculated using the modified frequentist method CL$_{s}$ [62–64]. Figure 6 displays the results.

For a Higgs boson mass of 125 GeV the expected limit is 0.95 and the observed limit is 1.89. Given that the resolution for the reconstructed Higgs boson mass is $\approx 10\%$, these results are compatible with a Higgs mass of 125 GeV. This is demonstrated by the red dashed line in the left panel of Fig. 6, which is the expected limit obtained from the sum of expected background and the signal of a SM Higgs boson with a mass of 125 GeV.

For all channels an excess of events over the expected background contributions is indicated by the fits of the BDT output distributions. The probability ($p$ value) to observe data as discrepant as observed under the background-only hypothesis is shown in the right panel of Fig. 6 as a function of the assumed Higgs boson mass. For $m_H = 125$ GeV, the excess of observed events corresponds to a local significance of 2.1 standard deviations away from the background-only hypothesis. This is consistent with the 2.1 standard deviations expected when assuming the standard model prediction for Higgs boson production.

The relative sensitivity of the channels that are topologically distinct is demonstrated in Table X for $m_H = 125$ GeV. The table lists the expected and observed limits and local significance for the $W(\ell\nu)H$ and $W(\tau\nu)H$ channels combined, for the $Z(\ell\ell)H$ channels combined, and for the $Z(\nu\nu)H$ channel.

The best-fit values of the production cross section for a 125 GeV Higgs boson, relative to the standard model cross section (signal strength, $\mu$), are shown in the left panel of Fig. 7 for the $W(\ell\nu)H$ and $W(\tau\nu)H$ channels combined, for the $Z(\ell\ell)H$ channels combined, and for the $Z(\nu\nu)H$ channel. The observed signal strengths are consistent with each other, and the value for the signal strength for the combination of all channels is 1.0 $\pm$ 0.5. In the right panel of Fig. 7 the correlation between the signal strengths for the separate $WH$ and $ZH$ production processes is shown. The
FIG. 6 (color online). Left: expected and observed 95% C.L. upper limits on the product of the $VH$ production cross section times the $H \rightarrow bb$ branching fraction, with respect to the expectations for the standard model Higgs boson. The limits are obtained combining the results of the searches using the 2011 (7 TeV) and 2012 (8 TeV) data. The red dashed line represents the expected limit obtained from the sum of expected backgrounds and the SM Higgs boson signal with a mass of 125 GeV. Right: local $p$ values and corresponding significance (measured in standard deviations) for the background-only hypothesis to account for the observed excess of events in the data.

two production modes are consistent with the SM expectation, within uncertainties. This figure contains slightly different information than the one on the left panel as some final states contain signal events that originate from both $WH$ and $ZH$ production processes. The $WH$ process contributes approximately 20% of the Higgs boson signal event yields in the $Z(\nu\nu)H$ channel, resulting from events in which the lepton is outside the detector acceptance, and the $Z(\ell\ell)H$ process contributes less than 5% to the $W(\ell\nu)H$ channel when one of the leptons is outside the detector acceptance. The dependency of the combined signal strength on the value assumed for the Higgs boson mass is shown in the left panel of Fig. 8.

In the right panel of Fig. 8 the best-fit values for the $\kappa_V$ and $\kappa_b$ parameters are shown. The parameter $\kappa_V$ quantifies the ratio of the measured Higgs boson couplings to vector bosons relative to the SM value. The parameter $\kappa_b$ quantifies the ratio of the measured Higgs boson partial width into $bb$ relative to the SM value. They are defined as $\kappa_V^2 = \sigma_{vH}/\sigma_{vH}^{SM}$ and $\kappa_b^2 = \Gamma_{bb}/\Gamma_{bb}^{SM}$, with the SM scaling of the total width [65]. The measured couplings are consistent with the expectations from the standard model, within uncertainties.

A. Results for the dijet mass cross-check analysis

The left panel of Fig. 9 shows a weighted dijet invariant mass distribution for the combination of all channels, in all boost regions, in the combined 7 and 8 TeV data, using the event selection for the $m(jj)$ cross-check analysis described in Sec. V. For each channel, the relative event weight in each boost region is obtained from the ratio of the expected number of signal events to the sum of expected signal and background events in a window of $m(jj)$ values between 105 and 150 GeV. The expected signal used corresponds to the production of the SM Higgs boson with a mass of 125 GeV. The weight for the highest-boost region is set to 1.0 and all other weights are adjusted proportionally. Figure 9 also shows the same weighted dijet invariant mass distribution with all backgrounds, except diboson production, subtracted. The data

<table>
<thead>
<tr>
<th>$m_H$ (GeV)</th>
<th>$\sigma/\sigma_{SM}$ (95% C.L.)</th>
<th>Significance expected</th>
<th>Significance observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W(\ell\nu, \tau\nu)H$</td>
<td>1.6</td>
<td>2.3</td>
<td>1.3</td>
</tr>
<tr>
<td>$Z(\ell\ell)H$</td>
<td>1.9</td>
<td>2.8</td>
<td>1.1</td>
</tr>
<tr>
<td>$Z(\nu\nu)H$</td>
<td>1.6</td>
<td>2.6</td>
<td>1.3</td>
</tr>
<tr>
<td>All channels</td>
<td>0.95</td>
<td>1.89</td>
<td>2.1</td>
</tr>
</tbody>
</table>

TABLE X. The expected and observed 95% C.L. upper limits on the product of the $VH$ production cross section times the $H \rightarrow bb$ branching fraction, with respect to the expectations for the standard model Higgs boson, for partial combinations of channels and for all channels combined, for $m_H = 125$ GeV. Also shown are the expected and observed local significances.
are consistent with the presence of a diboson signal from ZZ and WZ channels, with $Z \rightarrow b\bar{b}$, with a rate consistent with the standard model prediction from the MADGRAPH generator, together with a small excess consistent with the production of the standard model Higgs boson with a mass of 125 GeV. For the $m(jj)$ analysis, a fit to the dijet invariant mass distribution results in a measured Higgs boson signal strength, relative to that predicted by the standard model, of $\mu = 0.8 \pm 0.7$, with a local significance of 1.1 standard deviations with respect to the background-only hypothesis. For a Higgs boson of mass 125 GeV, the expected and observed 95% C.L. upper limits on the production cross section, relative to the standard model prediction, are 1.4 and 2.0, respectively.

### B. Diboson signal extraction

As a validation of the multivariate technique, BDT discriminants are trained using the diboson sample as signal, and all other processes, including $VH$ production (at the predicted standard model rate for a 125 GeV Higgs mass), as background. This is done for the 8 TeV data set only. The observed excess of events for the combined WZ and ZZ processes, with $Z \rightarrow b\bar{b}$, differs by over 7 standard deviations with respect to the background-only hypothesis. For a Higgs boson of mass 125 GeV, the expected and observed 95% C.L. upper limits on the production cross section, relative to the standard model prediction, are 1.4 and 2.0, respectively.
for a single experiment thus far. The signal strength represented by the expected significance, is the highest 2.1 standard deviations. The sensitivity of this search, as the production of the standard model Higgs boson is also 1.89.

125 GeV the expected limit is 0.95 and the observed limit is 1.9 corresponding to this excess, relative to that of the standard model Higgs boson, is $\mu = 1.0 \pm 0.5$. The measurements presented in this article represent the first indication of the $H \rightarrow b\bar{b}$ decay at the LHC.

IX. SUMMARY

A search for the standard model Higgs boson when produced in association with an electroweak vector boson and decaying to $b\bar{b}$ is reported for the $W(\mu\nu)H, W(e\nu)H, W(\tau\nu)H, Z(\mu\nu)H, Z(e\nu)H$, and $Z(\nu\nu)H$ channels. The search is performed in data samples corresponding to integrated luminosities of up to 5.1 fb$^{-1}$ at $\sqrt{s} = 7$ TeV and up to 18.9 fb$^{-1}$ at $\sqrt{s} = 8$ TeV, recorded by the CMS experiment at the LHC.

Upper limits, at the 95% confidence level, on the $VH$ production cross section times the $H \rightarrow b\bar{b}$ branching fraction, with respect to the expectations for a standard model Higgs boson, are derived for the Higgs boson in the mass range 110–135 GeV. For a Higgs boson mass of 125 GeV the expected limit is 0.95 and the observed limit is 1.89.

An excess of events is observed above the expected background with a local significance of 2.1 standard deviations. The expected significance when taking into account the production of the standard model Higgs boson is also 2.1 standard deviations. The sensitivity of this search, as represented by the expected significance, is the highest for a single experiment thus far. The signal strength corresponding to this excess, relative to that of the standard model Higgs boson, is $\mu = 1.0 \pm 0.5$. The measurements presented in this article represent the first indication of the $H \rightarrow b\bar{b}$ decay at the LHC.

ACKNOWLEDGMENTS

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: the Austrian Federal Ministry of Science and Research and the Austrian Science Fund; the Belgian Fonds de la Recherche Scientifique, and Fonds voor Wetenschappelijk Onderzoek; the Brazilian Funding Agencies (CNPq, CAPES, FAPERJ, and FAPESP); the Bulgarian Ministry of Education and Science; CERN; the Chinese Academy of Sciences, Ministry of Science and Technology, and National Natural Science Foundation of China; the Colombian Funding Agency (COLCIENCIAS); the Croatian Ministry of Science, Education and Sport; the Research Promotion Foundation, Cyprus; the Ministry of Education and Research, Recurrent financing contract SF0690030s09 and European Regional Development
Appendix: Postfit BDT Distributions

Figures 10–14 show all the postfit BDT distributions, for the $m_H = 125$ GeV training, for all channels and for all boost regions. In order to better display the different shapes of the signal and background BDT distributions, Fig. 15 shows these distributions for the highest-boost region in each channel, normalized to unity. See Sec. VIII for more details.
FIG. 10 (color online). Postfit BDT output distributions for $W(\mu \nu)H$ in the low-boost region (left), the intermediate-boost (right), and the high-boost (bottom), for 8 TeV data (points with error bars), all backgrounds, and signal, after all selection criteria have been applied. Bottom right: $VH$-enriched partition of the high-boost region is shown in more detail. The bottom inset in each figure shows the ratio of the number of events observed in data to that of the Monte Carlo prediction for signal and backgrounds.
FIG. 11 (color online). Postfit BDT output distributions for $W(e^\nu)$ in the low-boost region (left), the intermediate-boost (right), and the high-boost (bottom) for 8 TeV data (points with error bars), all backgrounds, and signal, after all selection criteria have been applied. Bottom right: $VH$-enriched partition of the high-boost region is shown in more detail. The bottom inset in each figure shows the ratio of the number of events observed in data to that of the Monte Carlo prediction for signal and backgrounds.
FIG. 12 (color online). Postfit BDT output distributions for $W(\tau\nu)H$ for 8 TeV data (points with error bars), all backgrounds, and signal, after all selection criteria have been applied. The bottom inset shows the ratio of the number of events observed in data to that of the Monte Carlo prediction for signal and backgrounds.
FIG. 13 (color online). Postfit BDT output distributions for $Z(\ell\ell)H$ in the low-boost region (left) and high-boost region (right), for 8 TeV data (points with error bars), all backgrounds, and signal, after all selection criteria have been applied. Top: $Z(\mu\mu)H$, bottom: $Z(\ell\ell)H$. The bottom inset in each figure shows the ratio of the number of events observed in data to that of the Monte Carlo prediction for signal and backgrounds.
FIG. 14 (color online). Postfit BDT output distributions for $Z(\nu\nu)H$ in the low-boost region (left), the intermediate-boost (right), and the high-boost (bottom) for 8 TeV data (points with error bars), all backgrounds, and signal, after all selection criteria have been applied. Bottom right: $VH$-enriched partition of the high-boost region is shown in more detail. The bottom inset in each figure shows the ratio of the number of events observed in data to that of the Monte Carlo prediction for signal and backgrounds.
FIG. 15 (color online). BDT output distributions, normalized to unity, for the highest-boost region in each channel, for all backgrounds and signal, after all selection criteria have been applied.
SEARCH FOR THE STANDARD MODEL HIGGS BOSON


(CMS Collaboration)

1Yerevan Physics Institute, Yerevan, Armenia
2Institut für Hochenergiephysik der ÖeAW, Wien, Austria
3National Centre for Particle and High Energy Physics, Minsk, Belarus
4Universiteit Antwerpen, Antwerpen, Belgium
5Vrije Universiteit Brussel, Brussel, Belgium
6Université Libre de Bruxelles, Bruxelles, Belgium
7Ghent University, Ghent, Belgium
8Université Catholique de Louvain, Louvain-la-Neuve, Belgium
9Université de Mons, Mons, Belgium
10Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil
11Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
12aUniversidade Estadual Paulista, São Paulo, Brazil
12bUniversidade Federal do ABC, São Paulo, Brazil
13Institute for Nuclear Research and Nuclear Energy, Sofia, Bulgaria
14University of Sofia, Sofia, Bulgaria
15Institute of High Energy Physics, Beijing, China
16State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
17Universidad de Los Andes, Bogota, Colombia
18Technical University of Split, Split, Croatia
19University of Split, Split, Croatia
20Institute Rudjer Boskovic, Zagreb, Croatia
21University of Cyprus, Nicosia, Cyprus
22Charles University, Prague, Czech Republic
23Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt
24National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
25Department of Physics, University of Helsinki, Helsinki, Finland
26Helsinki Institute of Physics, Helsinki, Finland
27Lappeenranta University of Technology, Lappeenranta, Finland
28DSM/IRFU, CEA/Saclay, Gif-sur-Yvette, France
29Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France
30Institut Pluridisciplinaire Hubert Curien, Université de Strasbourg, Université de Haute Alsace Mulhouse, CNRS/IN2P3, Strasbourg, France
31Centre de Calcul de l’Institut National de Physique Nucleaire et de Physique des Particules, CNRS/IN2P3, Villeurbanne, France
32Université de Lyon, Université Claude Bernard Lyon 1, CNRS-IN2P3, Institut de Physique Nucléaire de Lyon, Villeurbanne, France
33Institute of High Energy Physics and Informatization, Tbilisi State University, Tbilisi, Georgia
34RWTH Aachen University, I. Physikalisches Institut, Aachen, Germany
35RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany
36RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany
37Deutsches Elektronen-Synchrotron, Hamburg, Germany
38University of Hamburg, Hamburg, Germany
39Institut für Experimentelle Kernphysik, Karlsruhe, Germany
40Institute of Nuclear and Particle Physics (INPP), NCSR Demokritos, Aghia Paraskevi, Greece
41University of Athens, Athens, Greece
42University of Ioannina, Ioannina, Greece
43KFKI Research Institute for Particle and Nuclear Physics, Budapest, Hungary
44Institute of Nuclear Research ATOMKI, Debrecen, Hungary
45University of Debrecen, Debrecen, Hungary
46National Institute of Science Education and Research, Bhubaneswar, India
47Panjab University, Chandigarh, India
SEARCH FOR THE STANDARD MODEL HIGGS BOSON ... PHYSICAL REVIEW D 89, 012003 (2014)
aDeceased.
bAlso at Vienna University of Technology, Vienna, Austria.
cAlso at CERN, European Organization for Nuclear Research, Geneva, Switzerland.
dAlso at Institut Pluridisciplinaire Hubert Curien, Université de Strasbourg, Université de Haute Alsace Mulhouse, CNRS/IN2P3, Strasbourg, France.
eAlso at National Institute of Chemical Physics and Biophysics, Tallinn, Estonia.
fAlso at Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia.
gAlso at Universidade Estadual de Campinas, Campinas, Brazil.
hAlso at California Institute of Technology, Pasadena, CA, USA.
iAlso at Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France.
jAlso at Zewail City of Science and Technology, Zewail, Egypt.
kAlso at Suez Canal University, Suez, Egypt.
lAlso at Cairo University, Cairo, Egypt.
mAlso at Fayoum University, El-Fayoum, Egypt.
nAlso at British University in Egypt, Cairo, Egypt.
oNow at Ain Shams University, Cairo, Egypt.
pAlso at National Centre for Nuclear Research, Swierk, Poland.
qAlso at Université de Haute Alsace, Mulhouse, France.
rAlso at Universidad de Antioquia, Medellin, Colombia.
sAlso at Joint Institute for Nuclear Research, Dubna, Russia.
tAlso at Brandenburg University of Technology, Cottbus, Germany.
uAlso at The University of Kansas, Lawrence, KS, USA.
vAlso at Institute of Nuclear Research ATOMKI, Debrecen, Hungary.
wAlso at Eötvös Loránd University, Budapest, Hungary.
xAlso at Tata Institute of Fundamental Research—EHEP, Mumbai, India.
yAlso at Tata Institute of Fundamental Research—HECR, Mumbai, India.
zNow at King Abdulaziz University, Jeddah, Saudi Arabia.
aaAlso at University of Visva-Bharati, Santiniketan, India.
bbAlso at University of Ruhuna, Matara, Sri Lanka.
cAlso at Isfahan University of Technology, Isfahan, Iran.
dAlso at Sharif University of Technology, Tehran, Iran.
eAlso at Plasma Physics Research Center, Science and Research Branch, Islamic Azad University, Tehran, Iran.
fAlso at Università degli Studi di Siena, Siena, Italy.
gAlso at Centre National de la Recherche Scientifique (CNRS)—IN2P3, Paris, France.
hAlso at Purdue University, West Lafayette, IN, USA.
iAlso at Universidad Michoacana de San Nicolas de Hidalgo, Morelia, Mexico.
jjAlso at INFN Sezione di Padova, Università di Padova, Università di Trento (Trento), Padova, Italy.
kkAlso at Faculty of Physics, University of Belgrade, Belgrade, Serbia.
llAlso at Facoltà Ingegneria, Università di Roma, Roma, Italy.
mmAlso at Scuola Normale e Sezione dell’INFN, Pisa, Italy.
nAlso at University of Athens, Athens, Greece.