Search for the standard model Higgs boson decaying to bottom quarks in pp collisions at $\sqrt{s} = 7$ TeV

CMS Collaboration

A V A I L A B L E O N L I N E 5 M A R C H 2 0 1 2
Accepted 28 February 2012
Received in revised form 28 February 2012
Available online 5 March 2012
Editor: M. Doser

A R T I C L E I N F O
Article history:
© 2012 CERN. Published by Elsevier B.V. Open access under CC BY-NC-ND license.

A B S T R A C T

A search for the standard model Higgs boson (H) decaying to $b\bar{b}$ when produced in association with weak vector bosons ($V$) is reported for the following modes: $W(\mu\nu)H$, $W(e\nu)H$, $Z(\mu\mu)H$, $Z(ee)H$ and $Z(\nu\nu)H$. The search is performed in a data sample corresponding to an integrated luminosity of 4.7 fb$^{-1}$, recorded by the CMS detector in proton–proton collisions at the LHC with a center-of-mass energy of 7 TeV. No significant excess of events above the expectation from background is observed. Upper limits on the $VH$ production cross section times the Higgs boson mass range, the observed 95% confidence level upper limits vary from 3.4 to 7.5 times the standard model prediction; the corresponding expected limits vary from 2.7 to 6.7 times the standard model prediction.

1. Introduction

The process by which the electroweak symmetry is broken in nature remains elusive. In the standard model (SM) [1–3] the Higgs mechanism is considered to be the explanation [4–9]. The search for the Higgs boson is currently one of the most important endeavors of experimental particle physics.

Direct searches by experiments at the Large Electron–Positron Collider (LEP) have set a 95% confidence level (CL) lower bound on the Higgs boson mass of $m_H > 114.4$ GeV [10]. Direct searches at the Tevatron exclude at 95% CL the 162–166 GeV mass range [11], and the ATLAS experiment at the Large Hadron Collider (LHC) excludes, also at 95% CL, the following three regions: $m_H \notin 145$–206, 214–224, and 340–450 GeV [12–14]. Measurements of the W boson and top quark masses at LEP and the Tevatron, combined with precision measurements of electroweak parameters at the Z pole, provide an indirect constraint of $m_H < 158$ GeV at 95% CL [15]. The most likely mass for the SM Higgs boson remains near the LEP limit, where the Higgs boson decays predominantly into $b\bar{b}$. Experiments at the Tevatron have set 95% CL upper limits on the production cross section for a Higgs boson in this low-mass region. These limits range from approximately 4 to 10 times the standard model prediction, depending on the channels studied [16–22]. The observation of the $H \rightarrow b\bar{b}$ decay is of great importance in determining the nature of the Higgs boson.

At the LHC the main SM Higgs boson production mechanism is gluon fusion, with a cross section of $\approx 17$ pb for $m_H = 120$ GeV [23–39]. However, in this production mode, the detection of the $H \rightarrow b\bar{b}$ decay is considered nearly impossible due to overwhelming dijet production expected from quantum-chromodynamic (QCD) interactions. The same holds true for the next most copious production mode, through vector-boson fusion, with a cross section of $\approx 1.3$ pb [40–44]. Processes in which a low-mass Higgs boson is produced in association with a vector boson [45] have cross sections of $\approx 0.66$ pb and $\approx 0.36$ pb for WH and ZH, respectively.

In this Letter a search for the standard model Higgs boson in the $pp \rightarrow VH$ production mode is presented, where $V$ is either a W or a Z boson. The analysis is performed in the 110–135 GeV Higgs boson mass range, using a data sample corresponding to an integrated luminosity of 4.7 fb$^{-1}$, collected in 2011 by the Compact Muon Solenoid (CMS) experiment at a center-of-mass energy of 7 TeV. The following final states are included: $W(\mu\nu)H$, $W(e\nu)H$, $Z(\mu\mu)H$, $Z(ee)H$ and $Z(\nu\nu)H$, all with the Higgs boson decaying to $b\bar{b}$. Backgrounds arise from production of $W$ and $Z$ bosons in association with jets (from all quark flavors), singly and pair-produced top quarks ($t\bar{t}$), dibosons and QCD multijet processes. Simulated samples of signal and backgrounds are used to provide guidance in the optimization of the analysis as a function of the Higgs boson mass. Control regions in data are selected to adjust the simulations and estimate the contribution of the main backgrounds in the signal region. Upper limits at the 95% CL on the $pp \rightarrow VH$ production cross section are obtained for Higgs boson masses between 110–135 GeV. These limits are
based on the observed event count and background estimate in signal-enriched regions selected using the output discriminant of a boosted-decision-tree algorithm [46] (BDT analysis). As a cross-check, limits are also derived from the observed event count in the invariant mass distribution of $H \rightarrow bb$ candidates ($m(jj)$) analysis.

2. CMS detector and simulations

A detailed description of the CMS detector can be found elsewhere [47]. The momenta of charged particles are measured using a silicon pixel and strip tracker that covers the pseudorapidity range $|\eta| \leq 2.5$ and is immersed in a 3.8 T solenoidal 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 cryostat electromagnetic calorimeter (ECAL) and a brass-scintillator hadron calorimeter (HCAL), both used to measure particle energy depositions and consisting of a barrel assembly and two endcaps. The ECAL and HCAL extend to a pseudorapidity range $|\eta| \leq 3.0$. A steel/quartz-fiber Cherenkov forward detector (HF) extends the calorimetric coverage to $|\eta| \leq 5.0$. The outermost component of the CMS detector is the muon system consisting of gas detectors placed in the steel return yoke to measure the momentum of muons traversing the detector.

Simulated samples of signal and backgrounds are produced using various event generators, with the CMS detector response modeled with GEANT4 [48]. The Higgs boson signal samples are produced using POWHEG [49] interfaced with the HERWIG [50] event generator. The diboson samples are generated with PYTHIA 6.4 [51]. The MADGRAPH 4.4 [52] generator is used for the $W+\text{jets}$, $Z+\text{jets}$, and $t\bar{t}$ samples. The single-top samples are produced with POWHEG and the QCD multijet samples with PYTHIA. The default set of parton distribution functions (PDF) used to produce these samples is CTEQ6L1 [53]. The PYTHIA parameters for the underlying event are set to the Z2 tune [54].

During the period in which the data for this analysis was recorded, the LHC instantaneous luminosity reached up to $3.5 \times 10^{33} \text{ cm}^{-2} \text{s}^{-1}$ and the average number of pp interactions per bunch crossing was approximately ten. Additional pp interactions overlapping with the event of interest in the same bunch crossing, denoted as pile-up events (PU), are therefore added in the simulated samples to represent the PU distribution measured in data.

3. Triggers and event reconstruction

3.1. Triggers

Several triggers are used to collect events consistent with the signal hypothesis in each of the five channels. For the WH 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 depositions to maintain an acceptable trigger rate. For the $W(\mu\nu)H$ channel, the trigger thresholds for the muon transverse momentum, $p_T$, are in the range of 17 to 40 GeV. The higher thresholds are used for the periods of higher instantaneous luminosity. The combined trigger efficiency is $\approx 90\%$ for signal events that would pass all offline requirements, described in Section 4. For the $W(\ell\nu)H$ channel, the electron $p_T$ threshold ranges from 17 to 30 GeV. The lower-threshold trigger paths require two jets and a minimum requirement on an online estimate of the missing transverse energy, evaluated in the high level trigger algorithm as the modulus of the negative vector sum of the transverse momenta of all reconstructed jets identified by a particle-flow algorithm [55]. These extra requirements help to maintain acceptable trigger rates during the periods of high instantaneous luminosity. The combined efficiency for these triggers for signal events that pass the final offline selection criteria is $>95\%$.

The $Z(\ell\ell)H$ channel uses the same single-muon triggers as the $W(\mu\nu)H$ channel. For the $Z(\nu\nu)H$ channel, dielectron triggers with lower $p_T$ thresholds (17 and 8 GeV) and tight isolation requirements are used. These triggers are $\approx 99\%$ efficient for all ZH signal events that pass the final offline selection criteria. For the $Z(\ell\nu)H$ channel, a combination of four triggers is used. The first one requires missing transverse energy $>150$ GeV and is used for the complete dataset. The other triggers use lower thresholds on the missing transverse energy (evaluated for these cases using all energy deposits in the calorimeter), but require the presence of jets. One of these triggers requires missing transverse energy above 80 GeV and a central ($|\eta| < 2.4$) jet with $p_T$ above 80 GeV, and the other two require the presence of two central jets with $p_T > 20$ GeV and missing transverse energy thresholds of 80 and 100 GeV, depending on the luminosity. The combined trigger efficiency for $Z(\ell\nu)H$ signal events is $\approx 98\%$ with respect to the offline event reconstruction and selection, described below.

3.2. Event reconstruction

The reconstructed interaction vertex with the largest value of $\sum_i p_T^i$, where $p_T^i$ is the transverse momentum of the $i$-th track associated to 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 PU interactions affect jet momentum reconstruction, missing transverse energy reconstruction, lepton isolation and b-tagging efficiency. To mitigate these effects, a track-based algorithm that filters all charged hadrons that do not originate from the primary interaction is used. In addition, a calorimeter-based algorithm evaluates the energy density in the calorimeter from interactions not related to the primary vertex and subtracts its contribution to reconstructed jets in the event [56].

Jets are reconstructed from particle-flow objects [55] using the anti-$k_T$ clustering algorithm [57], as implemented in the FASTJET package [58,59], using a distance parameter of 0.5. Each jet is required to be within $|\eta| < 2.5$, to have at least two tracks associated to it, and to have electromagnetic and hadronic energy fractions of at least 1% of the total jet energy. Jet energy corrections, as a function of pseudorapidity and transverse energy of the jet, are applied [60]. 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 Letter.

Electron reconstruction requires the matching of an energy cluster in the ECAL with a track in the silicon tracker [61]. Identification criteria based on the ECAL shower shape, track-ECAL cluster matching, and consistency with the primary vertex are imposed. 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.

Muons are reconstructed using two algorithms [62]: one in which tracks in the silicon tracker are matched to signals in the muon chambers, and another in which a global track fit is performed seeded by signals in the muon system. The muon candidates used in the analysis are required to be reconstructed successfully 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
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 energy 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 pT the lepton is rejected; the exact requirement depends on the lepton pT, pT and flavor.

The Combined Secondary Vertex (CSV) b-tagging algorithm [63] is used to identify jets that are likely to arise from the hadronization of b quarks. This algorithm combines the information about track impact parameters and secondary vertices within jets in a likelihood discriminant to provide separation of b jets from jets originating from light quarks and gluons, and also from charm quarks. Several working points for the CSV output discriminant are used in the analysis, with different efficiencies and misidentification rates for b jets. For a CSV > 0.90 requirement the efficiencies to tag b quarks, c quarks, and light quarks, are approximately 50%, 6%, and 0.15%, respectively [64]. The corresponding efficiencies for CSV > 0.244 are approximately 82%, 40%, and 12%.

All events from data and from the simulated samples are required to pass the same trigger and event reconstruction algorithms. Scale factors that account for the differences in the performance of these algorithms between data and simulations are computed and used in the analysis.

4. Event selection

The background processes to VH production are vector-boson + jets, tt, single-top, dibosons (VV) and QCD multijet production. These overwhelm the signal by several orders of magnitude. The event selection for the bdt analysis is based first on the kinematic reconstruction of the vector bosons and the Higgs boson decay into two b-tagged jets. Backgrounds are then substantially reduced by requiring a significant boost in the pT of the vector boson and the Higgs boson [65], which can recoil away from each other with a large azimuthal opening angle, Δφ(V, H), between them. The boost requirements in the Z(ℓℓ)H and WH analyses are pT > 100 and pT > 150 GeV, respectively. The fractions of signal events that satisfy these requirements are approximately 25% and 10%. For the Z(νν)H analysis the boost requirement is pT > 160 GeV.

Candidate W → ℓν decays are identified by requiring the presence of a single isolated lepton and additional missing transverse energy. Muons are required to have a pT above 20 GeV; the corresponding value for electrons is 30 GeV. For the W(ℓν)H analysis, Emiss T is required to be greater than 35 GeV to reduce contamination from QCD multijet processes.

Candidate Z → ℓℓ decays are reconstructed by combining isolated, oppositely charged pairs of electrons or muons, each lepton with pT > 20 GeV, and requiring the dilepton invariant mass to satisfy 75 GeV < mℓℓ < 105 GeV. The identification of Z → ℓℓ decays requires Emiss T > 160 GeV. The high threshold is dictated by the trigger and is consistent with a significant boost in the pT of the Z boson. The QCD multijet background is greatly reduced in this channel when requiring that the Emiss T does not originate from mismeasured jets. To that end, a Δφ(Emiss T, jet) > 0.5 radians requirement is applied on the azimuthal angle between the Emiss T direction and the closest jet with pT > 20 GeV and |η| < 2.5. To reduce backgrounds from tt and WZ in the WH and Z(νν)H channels, events with additional isolated leptons, Nηj, with pT > 20 GeV are rejected.

<table>
<thead>
<tr>
<th>Variable</th>
<th>W(ℓν)H</th>
<th>Z(ℓℓ)H</th>
<th>Z(νν)H</th>
</tr>
</thead>
<tbody>
<tr>
<td>pT(ℓ1)</td>
<td>&gt; 30 GeV</td>
<td>&gt; 20 GeV</td>
<td>&gt; 80 GeV</td>
</tr>
<tr>
<td>pT(ℓ2)</td>
<td>&gt; 30 GeV</td>
<td>&gt; 20 GeV</td>
<td>&gt; 20 GeV</td>
</tr>
<tr>
<td>pT(νν)</td>
<td>&gt; 150 (165) GeV</td>
<td>&gt; 100 GeV</td>
<td>&gt; 160 GeV</td>
</tr>
<tr>
<td>pT(V)</td>
<td>&gt; 150 (160) GeV</td>
<td>&gt; 100 GeV</td>
<td>–</td>
</tr>
<tr>
<td>Emiss T &gt; 35 GeV for W(ℓν)H</td>
<td>–</td>
<td>&gt; 160 GeV</td>
<td></td>
</tr>
<tr>
<td>Δφ(V, H) &gt; (–2.95) rad</td>
<td>&gt; (–2.90) rad</td>
<td>&gt; (–2.90) rad</td>
<td></td>
</tr>
<tr>
<td>CSVmax</td>
<td>&gt; 0.40 (0.90)</td>
<td>&gt; 0.244 (0.90)</td>
<td>&gt; 0.50 (0.90)</td>
</tr>
<tr>
<td>CSVmin</td>
<td>&gt; 0.40</td>
<td>&gt; 0.244 (0.50)</td>
<td>&gt; 0.50</td>
</tr>
<tr>
<td>Nηj</td>
<td>= 0</td>
<td>= 0</td>
<td>= 0</td>
</tr>
<tr>
<td>Δφ(Emiss T, jet)</td>
<td>&lt; (–0), &lt; (–2)</td>
<td>&lt; (–0)</td>
<td>&lt; (–0)</td>
</tr>
<tr>
<td>Δφ(FEMiss T, jet)</td>
<td>–</td>
<td>–</td>
<td>&gt; 0.5 (1.5) rad</td>
</tr>
</tbody>
</table>

The reconstruction of the H → b̅b decay is made by requiring the presence of two central (|η| < 2.5) jets above a minimum pT threshold, and tagged by the CSV algorithm. If more than two such jets are found in the event, the pair of jets with the highest total dijet transverse momentum, pT(jjj), is selected. After the b-tagging requirements are applied, the fraction of H → b̅b candidates in signal events that contain the two b jets from the Higgs boson decay is near 100%. The background from V + jets and dibosons is reduced significantly through b tagging, and sub-processes where the two jets originate from genuine b quarks dominate the final selected data sample.

The bdt analysis is implemented in the TMVA framework [66]. To better separate signal from background under different Higgs boson mass hypotheses, the bdt is trained separately at each mass value using simulated samples for signal and background that pass the event selection described above. The final set of input variables is chosen by iterative optimization from a larger number of potentially discriminating variables. The same set is used for all modes and for all Higgs boson mass hypotheses tested. These include the dijet invariant mass mjj, the dijet transverse momentum pTjj, the separation in pseudorapidity between the two jets Δηjj, the transverse momentum of the vector boson pT(V), the maximum and minimum CSV values among the two jets, the azimuthal angle between the vector boson and the dijets Δφ(V, H), and the number of additional central jets Nηj. A signal region, where observed and expected events are counted, is identified in the bdt output distribution by optimizing a figure of merit that takes into account the level of systematic uncertainty on the expected background.

Table 1 summarizes the selection criteria used in each of the five channels for both the bdt and the mjj analyses. For the cross-check mjj analysis more stringent requirements are imposed on several of the variables used for the bdt selection. In addition, explicit requirements are made on Δφ(V, H) and on Nηj. For each Higgs boson mass, mH, tested events are counted in a 30 GeV window centered on the mean of the expected dijet mass peak. For the Z(ℓℓ)H modes the dijet mass distribution is asymmetric and the window is centered 5 GeV lower than mH, while for the WH and Z(νν)H modes the window is centered at mH. For these modes a higher pT boost requirement is made resulting in more collimated b jets and a mass peak more symmetric around mH. For every channel, the mjj analysis was found to be about 10% less sensitive than the bdt analysis.
5. Background control regions

Appropriate control regions that are orthogonal to the signal region are identified in data and used to adjust the Monte Carlo simulation normalization for the most important background processes: W + jets and Z + jets (with light- and heavy-flavor jets), and tZ. For each of the search channels and for each of these background processes, a control region is found such that its composition is enriched in that specific background process. The discrepancies between the expected and observed yields in the data in these control regions are used to obtain a scale factor by which the normalizations of the simulations are adjusted. For each channel, this procedure is performed simultaneously for all control regions. The background yields in the signal region from these sources are then estimated from the adjusted simulation samples. The uncertainties in the scale factor determination include a statistical uncertainty due to the finite size of the samples and an associated systematic uncertainty from the differences in the shapes of the distributions that could affect the estimate of the yields when extrapolating to the signal region. These systematic uncertainties are obtained by varying the control region selection criteria in order to select regions of phase space that are closer or further from the signal region. The systematic uncertainty assigned covers the largest variation in the scale factor value found. The procedures applied in the construction of the control regions include reversing the b-tagging requirements to enhance W + jets and Z + jets with light-flavor jets, enforcing a tighter b-tagging requirement and requiring extra jets to enhance tZ, and requiring low boost in order to enhance Vb̄b over tZ.

Consistent scale factors are found for each background process across the different channels. For tZ, W + udscg, and Zb̄b production the scale factors are compatible with unity within their uncertainties (10–20%). For Wb̄b, the control region selected contains approximately 50% Wb̄b and single-top events, with the remainder being tZ and W + udscg, which are well constrained by their own control regions. A choice is made to assign the observed excess of events in this region all to Wb̄b, leading to a scale factor of 2 for this background, while the estimate of single-top production is taken from the simulation. Reversing this assignment has a negligible effect on the final result of the analysis. The total uncertainty (excluding luminosity) assigned to the Wb̄b yield in the signal region is approximately 30%. This includes a 15% uncertainty on the extrapolation of the yield from the control region to the signal region, determined in data with the method outlined above. The systematic uncertainty assigned to the predicted yield for single-top production is 30%. The diboson background is taken from the simulation and a systematic uncertainty of 30% is assigned.

For Z(νν)H the QCD multijet background in the signal region is estimated from data using control regions of high and low values of two uncorrelated variables with significant discriminating power towards such events. One is the angle between the missing energy vector and the closest jet in azimuth, Δφ(Emiss,T, jet), and the other is the sum of the CSV values of the two b-tagged jets. The signal region is at high values of both discriminants, while QCD multijet events populate regions with low values of either. The method predicts a very small contamination of 0.015 ± 0.008 for these background events, which is considered to be negligible. For all other search channels, after all selection criteria are applied, the QCD multijet backgrounds are also found to be negligible and not discussed in what follows.

6. Yield uncertainties

Table 2 lists the uncertainties on the expected signal and background yields that enter in the limit calculation.

The uncertainty in the CMS luminosity measurement for the dataset used in the analysis is estimated to be 4.5% [67]. Muon and electron trigger, reconstruction, and identification efficiencies are determined in data from samples of leptonic Z boson decays. The uncertainty on the yields due to the trigger efficiency is 2% per charged lepton and the uncertainty on the identification efficiency is also 2% per lepton. The parameters describing the Z(νν)H trigger efficiency turn-on curve have been varied within their statistical uncertainties and for different assumptions on the methodology to derive the efficiency. A yield uncertainty of 2% is estimated.

The jet energy scale is varied within one standard deviation as a function of pT and η. The efficiency of the analysis selection is recomputed to assess the variation in yield. 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 yields due to this effect is obtained. An uncertainty of 3% is assigned to the yields of all processes in the WH and Z(νν)H modes due to the uncertainty related to the missing transverse energy estimate.

Data-to-simulation b-tagging scale factors, measured in tZ events, are applied consistently to jets in signal and background events. The measured uncertainties for the b-tagging scale factors are: 6% per b tag, 12% per charm tag and of 15% per mistagged jet (originating from gluons and light u, d, s quarks). These translate into yield uncertainties in the 3–15% range, depending on the channel and the specific process.

The total VH signal cross section has been calculated to next-to-next-to-leading (NNLO) order accuracy, and the total theoretical uncertainty is 4% [39], including the effect of scale and PDF variations [68–72]. This analysis is performed in the boosted regime, and thus, potential differences in the pT spectrum of the V and H between data and Monte Carlo generators could introduce systematic effects in the signal acceptance and efficiency estimates. Calculations are available that estimate the next-to-leading-order (NLO) electroweak [73–76] and NNLO QCD [77,78] corrections to VH production in the boosted regime. The central value used for the cross section in the analysis was not adjusted for these calculations. The estimated uncertainties from electroweak corrections for a boost of ∼150 GeV are 5% for ZH and 10% for WH. For the QCD correction, a 10% uncertainty is estimated for both ZH and WH, which includes effects due to additional jet activity from initial- and final-state radiation. The finite size of the signal Monte Carlo samples, after all selection criteria are applied, contributes 1–5% uncertainty across all channels.

The uncertainty in the background yields that results from the estimates from data is in the 10–35% range. For the predictions

<table>
<thead>
<tr>
<th>Source</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminosity</td>
<td>4.5%</td>
</tr>
<tr>
<td>Lepton efficiency and trigger (per lepton)</td>
<td>3%</td>
</tr>
<tr>
<td>Z(νν)H triggers</td>
<td>2%</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>2–3%</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>3–6%</td>
</tr>
<tr>
<td>Missing transverse energy</td>
<td>3%</td>
</tr>
<tr>
<td>b-Tagging</td>
<td>3–15%</td>
</tr>
<tr>
<td>Signal cross section (scale and PDF)</td>
<td>4%</td>
</tr>
<tr>
<td>Signal cross section (pT boost, EWK/QCD)</td>
<td>5–10%/10%</td>
</tr>
<tr>
<td>Signal Monte Carlo statistics</td>
<td>1–5%</td>
</tr>
<tr>
<td>Backgrounds (data estimate)</td>
<td>10–35%</td>
</tr>
<tr>
<td>Diboson and single-top (simulation estimate)</td>
<td>30%</td>
</tr>
</tbody>
</table>
obtained solely from simulation, as described in Section 5, an uncertainty of 30% (approximately the uncertainty on the measured cross section) is assigned for single-top. For the diboson backgrounds, a 30% yield uncertainty is assumed.

7. Results

The primary physics result presented in this Letter is an upper limit on the production of a standard model Higgs boson in as-
Fig. 1. Distributions of the \( \text{BDT} \) output, for \( m_H = 115 \) GeV, for each mode after all selection criteria are applied. The solid histograms for the backgrounds and the signal are summed cumulatively. The line histogram for signal is also shown superimposed. The data is represented by points with error bars.

Table 3 lists, for each Higgs boson mass hypothesis considered, the expected signal and background yields in the signal region for the BDT analysis, together with the observed number of events. Table 3 also lists the requirements on the output of the BDT distributions that define the signal region. These distributions are shown in Fig. 1 for the \( m_H = 115 \) GeV case, where data are overlaid with the predicted sample composition. The invariant dijet mass distribution, combined for all channels, for events that pass the \( m(jj) \) analysis selection is shown in Fig. 2. The predicted number of background events are determined in data using the control regions described in Section 5, and from direct expectations from simulation for those backgrounds for which scale factors were not explicitly derived from control regions. Signal yields are determined from the simulations. The uncertainties include all sources listed in Section 6, except for luminosity. Total signal uncertainties are approximately 20%, and total background uncertainties are approximately in the 20 to 30% range.

No significant excess of events is observed in any channel, and the results of all channels are combined to obtain 95% CL upper limits on the Higgs boson production cross section in the \( VH \) modes with \( H \rightarrow b\bar{b} \), relative to the standard model prediction. This is done separately for both the BDT and \( m(jj) \) analyses for assumed Higgs boson masses in the 110–135 GeV range. The observed limits at each mass point, the median expected limits and the 1\( \sigma \) and 2\( \sigma \) bands are calculated using the modified frequentist
method \text{CL}_s [79–81]. The inputs to the limit calculation include the number of observed events \(N_{\text{obs}}\), and the signal and background estimates \(B_{\text{exp}}\), which are listed in Table 3 for the \text{bdt} analysis. The systematic and statistical uncertainties on the signal and background estimates, listed in Section 6, are treated as nuisance parameters in the limit calculations, with appropriate correlations taken into account.

Table 4 summarizes, for the \text{bdt} and m(jj) analyses, the expected and observed 95% CL upper limits on the product of the VH production cross section times the \(H \rightarrow bb\) branching ratio, with respect to the expectations for a standard model Higgs boson, for the \text{bdt} analysis, determined to be superior and it is considered to be the main result in this Letter. The \text{bdt} results are displayed in Fig. 3.

8. Summary

A search for the standard model Higgs boson decaying to bb when produced in association with weak vector bosons is reported for the following channels: \(W(\ell\nu)H\), \(W(e\nu)H\), \(Z(\mu\nu)H\), \(Z(ee)H\) and \(Z(\nu\nu)H\). The search is performed in a data sample corresponding to an integrated luminosity of 4.7 fb\(^{-1}\). No significant excess of events above the expectation from background is observed. Upper limits on the VH production cross section times the \(H \rightarrow bb\) branching ratio, with respect to the expectations for a standard model Higgs boson, are derived for a Higgs boson in the mass range 110–135 GeV. In this range, the observed 95% confidence level upper limits vary from 3.4 to 7.5 times the standard model prediction; the corresponding expected limits vary from 2.7 to 6.7. This Letter reports the first upper limits from the LHC in these channels.

Acknowledgements

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC machine. We thank the technical and administrative staff at CERN and other CMS institutes, and acknowledge support from: FMSR (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES (Croatia); MEERKAT; MoER, SF0690030s09 and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); CSTRT (Greece); OTKA and NKTH (Hungary); DAIE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); NRF and WCU (Republic of Korea); LAS (Lithuania); CINVESTAV, CONACYT, SEP, and UASLP-FAI (Mexico); MSI (New Zealand); NSHE and NSC (Poland); FCT (Portugal); JINR (Armenia, Belarus, Georgia, Ukraine, Uzbekistan); MON, RosAtom, RAS and RFBR (Russia); MSTD (Serbia); MIPAN and CPAN (Spain); Swiss Funding Agencies (Switzerland); NSC (Taipei); TUBITAK and TAEK (Turkey); STFC (United Kingdom); DOE and NSF (USA). Individuals have received support from the Marie-Curie programme and the European Research Council (European Union); the Leventis Foundation; the A.P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l’Industrie et dans l’Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Council of Science and Industrial Research, India; and the HOMING PLUS programme of Foundation for Polish Science, cofinanced from European Union, Regional Development Fund.

Open access

This article is published Open Access at sciencedirect.com. It is distributed under the terms of the Creative Commons Attribution License 3.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

CMS Collaboration

S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan
Yerevan Physics Institute, Yerevan, Armenia

Institut für Hochenergiephysik der ÖAW, Wien, Austria

V. Mossolov, N. Shumeiko, J. Suarez Gonzalez
National Centre for Particle and High Energy Physics, Minsk, Belarus

Universiteit Antwerpen, Antwerpen, Belgium

Vrije Universiteit Brussel, Brussel, Belgium

Université Libre de Bruxelles, Bruxelles, Belgium

Ghent University, Ghent, Belgium

Université Catholique de Louvain, Louvain-la-Neuve, Belgium

N. Belyi, T. Caerbergs, E. Daubie
Université de Mons, Mons, Belgium

Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil

RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany


Deutsches Elektronen-Synchrotron, Hamburg, Germany


University of Hamburg, Hamburg, Germany


Institut für Experimentelle Kernphysik, Karlsruhe, Germany


Institute of Nuclear Physics “Demokritos”, Aghia Paraskevi, Greece

L. Gouskos, T.J. Mertzimekis, A. Panagiotou, N. Saoulidou, E. Stiliaris

University of Athens, Athens, Greece

I. Evangelou, C. Foudas 1, P. Kokkas, N. Manthos, I. Papadopoulos, V. Patras, F.A. Triantis

University of Ioannina, Ioannina, Greece

A. Arányi, G. Bencze, L. Boldizsár, C. Hajdu 1, P. Hidas, D. Horváth 15, A. Kapusi, K. Krajcžar 16, F. Sikler 1, V. Veszprémi, G. Vesztergombi 16

KFKI Research Institute for Particle and Nuclear Physics, Budapest, Hungary


Panjab University, Chandigarh, India
S. Ahuja, B.C. Choudhary, A. Kumar, A. Kumar, S. Malhotra, M. Naimuddin, K. Ranjan, V. Sharma, R.K. Shivpuri
University of Delhi, Delhi, India

Saha Institute of Nuclear Physics, Kolkata, India

R.K. Choudhury, D. Dutta, S. Kailas, V. Kumar, A.K. Mohanty, L.M. Pant, P. Shukla
Bhabha Atomic Research Centre, Mumbai, India

Tata Institute of Fundamental Research – EHEP, Mumbai, India

S. Banerjee, S. Dugad, N.K. Mondal
Tata Institute of Fundamental Research – HEGR, Mumbai, India

Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

INFN Sezione di Bari, Bari, Italy

INFN Sezione di Bologna, Bologna, Italy

S. Albergo, G. Cappello, M. Chiorboli, S. Costa, R. Potenza, A. Tricomi, C. Tuve
INFN Sezione di Catania, Catania, Italy

G. Barbagli, V. Ciulli, C. Civinini, R. D’Alessandro, E. Focardi, S. Frosali, E. Gallo, S. Gonzalez, M. Meschini, S. Paololetti, G. Sguazzoni, A. Tropiano
INFN Sezione di Firenze, Firenze, Italy

L. Benussi, S. Bianco, S. Colafranceschi, F. Fabbri, D. Piccolo
INFN Laboratori Nazionali di Frascati, Frascati, Italy

P. Fabbricatore, R. Musenich
INFN Sezione di Genova, Genova, Italy
S.G. Heo, S.K. Nam

Kangwon National University, Chuncheon, Republic of Korea


Kyungpook National University, Daegu, Republic of Korea

J.Y. Kim, Zero J. Kim, S. Song

Chonnam National University, Institute for Universe and Elementary Particles, Kwangju, Republic of Korea

H.Y. Jo

Konkuk University, Seoul, Republic of Korea


Korea University, Seoul, Republic of Korea

M. Choi, S. Kang, H. Kim, J.H. Kim, C. Park, I.C. Park, S. Park, G. Ryu

University of Seoul, Seoul, Republic of Korea


Sungkyunkwan University, Suwon, Republic of Korea

M.J. Bilinskas, I. Grigelionis, M. Janulis

Vilnius University, Vilnius, Lithuania


Centro de Investigacion y de Estudios Avanzados del IPN, Mexico City, Mexico

S. Carrillo Moreno, F. Vazquez Valencia

Universidad Iberoamericana, Mexico City, Mexico

H.A. Salazar Ibarguen

Benemerita Universidad Autonoma de Puebla, Puebla, Mexico

E. Casimiro Linares, A. Morelos Pineda, M.A. Reyes-Santos

Universidad Autónoma de San Luis Potosí, San Luis Potosí, Mexico

D. Krofcheck

University of Auckland, Auckland, New Zealand

A.J. Bell, P.H. Butler, R. Doesburg, S. Reucroft, H. Silverwood

University of Canterbury, Christchurch, New Zealand


National Centre for Physics, Quaid-I-Azam University, Islamabad, Pakistan

G. Brona, M. Cwiok, W. Dominik, K. Doroba, A. Kalinowski, M. Konecki, J. KROLIKOWSKI

Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland

L. Levchuk

National Scientific Center, Kharkov Institute of Physics and Technology, Kharkov, Ukraine


University of Bristol, Bristol, United Kingdom


Rutherford Appleton Laboratory, Didcot, United Kingdom


Imperial College, London, United Kingdom


Brunel University, Uxbridge, United Kingdom

K. Hatakeyama, H. Liu, T. Scarborough

Baylor University, Waco, TX, USA

C. Henderson

The University of Alabama, Tuscaloosa, AL, USA


Boston University, Boston, MA, USA


Brown University, Providence, RI, USA


University of California, Davis, Davis, CA, USA


University of California, Los Angeles, Los Angeles, CA, USA