Search for the Higgs Boson in the $H \rightarrow WW \rightarrow l\nu jj$ Decay Channel in $pp$ Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector

G. Aad et al.*
(ATLAS Collaboration)

(Received 16 September 2011; published 30 November 2011)

A search for a Higgs boson has been performed in the $H \rightarrow WW \rightarrow l\nu jj$ channel in 1.04 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 7$ TeV recorded with the ATLAS detector at the Large Hadron Collider. No significant excess of events is observed over the expected background and limits on the Higgs boson production cross section are derived for a Higgs boson mass in the range $240 \text{ GeV} < m_H < 600 \text{ GeV}$. The best sensitivity is reached for $m_H = 400 \text{ GeV}$, where the 95% confidence level upper bound on the cross section for $H \rightarrow WW$ production is 3.1 pb, or 2.7 times the standard model prediction.

DOI: 10.1103/PhysRevLett.107.231801

In the standard model (SM [1–3]), a scalar field vacuum expectation value breaks the electroweak symmetry, gives masses to the $W$ and $Z$ bosons [4–6], and manifests itself directly as the so-called Higgs boson. A primary goal of the Large Hadron Collider (LHC) is to test the SM mechanism of electroweak symmetry breaking by searching for Higgs boson production in high energy proton-proton collisions. Thanks in part to the large gluon luminosity at LHC energies [7,8], the Higgs boson is predominantly produced via gluon fusion ($gg \rightarrow H$) [9–12] and to a lesser extent via vector boson fusion ($qq \rightarrow qgH$) [13–15]. Current limits from direct searches at LEP and the Tevatron exclude Higgs boson masses $m_H < 114.4$ GeV [16] and $156 \text{ GeV} < m_H < 177 \text{ GeV}$ [17] at 95% C.L.

For $m_H \approx 135$ GeV, the dominant decay mode of the Higgs boson is $H \rightarrow WW^{(*)}$ [18,19]. The most sensitive Higgs boson search channel in the mass region around $m_H = 160 \text{ GeV}$ is the purely leptonic mode $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$. For $m_H \approx 200$ GeV, the $H \rightarrow WW \rightarrow \ell\nu jj$ channel, where one $W$ decays to a pair of jets ($W \rightarrow jj$), also becomes important. The advantage of $H \rightarrow WW \rightarrow \ell\nu jj$ over $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ is the ability to fully reconstruct the Higgs boson mass.

This Letter describes a search for a Higgs boson in the $H \rightarrow WW \rightarrow \ell\nu jj$ channel using the ATLAS detector at the LHC, based on 1.04 fb$^{-1}$ of $pp$ collision data at a center-of-mass energy $\sqrt{s} = 7$ TeV collected during 2011. In this analysis, the distribution of the $\ell\nu jj$ invariant mass $m(\ell\nu jj)$, reconstructed using the charged-lepton neutrino invariant mass constraint $m(\ell\nu) = m(W)$ and the requirement that two of the jets in the event are consistent with a $W \rightarrow jj$ decay, is used to search for a Higgs boson signal.

The results of a similar search for $H \rightarrow WW \rightarrow \ell\nu jj$ based on 35 pb$^{-1}$ of data recorded during the 2010 LHC run were presented in Ref. [20].

The present search, based on the measured shape of the $m(\ell\nu jj)$ distribution, is restricted to $m_H > 240$ GeV, in order to ensure a smoothly varying nonresonant background, well clear of the effective kinematic cutoff $m(\ell\nu jj) \sim 160$ GeV. For $m_H \approx 600 \text{ GeV}$, the jets from $W \rightarrow jj$ decay begin to overlap due to the large $W$ boost, and the natural width of the Higgs boson becomes large. A detailed treatment of these issues is beyond the scope of the present analysis. The best sensitivity in this analysis is expected for $m_H \sim 400$ GeV.

The ATLAS detector [21] is a multipurpose particle physics apparatus with forward-backward symmetric cylindrical geometry covering the pseudorapidity range $|\eta| < 2.5$ for track and $|\eta| < 4.9$ for jet measurements [22]. The inner tracking detector (ID) consists of a silicon pixel detector, a silicon microstrip detector, and a transition radiation tracker. The ID is surrounded by a thin superconducting solenoid providing a 2 T magnetic field, and by a high-granularity liquid-argon (LAr) sampling electromagnetic (EM) calorimeter. An iron-scintillator tile calorimeter provides hadronic coverage in the central rapidity range. The end-cap and forward regions are instrumented with LAr calorimetry for both electromagnetic and hadronic measurements. The muon spectrometer surrounds the calorimeters and consists of three large superconducting toroids, each with eight coils, a system of precision tracking chambers, and detectors for triggering.

Detailed Monte Carlo (MC) studies of signal and backgrounds have been carried out [23]. The interaction with the ATLAS detector is modeled with GEANT4 [24] and the events are processed using the same reconstruction that is used to perform the reconstruction on data. The effect of multiple $pp$ interactions in the same bunch crossing (pile-up) at the high luminosities achieved by the LHC in 2011 is modeled by superimposing, at the generation stage, several simulated minimum-bias events on the simulated signal.
and background events. MC samples were generated with different pile-up levels and subsequently reweighted to match the pile-up conditions observed in the data.

The data used in this analysis were recorded during periods when all ATLAS subdetectors were operating under nominal conditions. The events were triggered by requiring the presence of an electron candidate with transverse energy $E_T > 20$ GeV or a muon candidate with transverse momentum $p_T > 18$ GeV.

Electron candidates are selected from clustered energy deposits in the EM calorimeter with an associated track and are required to satisfy a tight set of identification cuts [25] with an efficiency of 71 ± 1.6% for electrons with $E_T > 20$ GeV. While the energy measurement is taken from the EM calorimeter, the pseudorapidity $\eta$ and azimuthal angle $\phi$ are taken from the associated track. The cluster is required to be inside the range $|\eta| < 2.47$, excluding $1.37 < |\eta| < 1.52$ and small calorimeter regions affected by temporary operational problems. The track associated with the electron candidate is required to point back to a reconstructed primary vertex with a transverse impact parameter significance $d_0/\sigma_{d_0} \leq 10$ and an impact parameter along the beam direction $z_0 \leq 10$ mm. Electrons are required to be isolated: the sum of the transverse energies in cells inside a cone $\Delta R < 0.3$ around [26] the cluster barycenter (excluding the electron itself) must satisfy $\Sigma(E_T^{\text{calo}}) < 4$ GeV.

Muons are reconstructed by combining tracks in the inner detector and muon spectrometer, with efficiency 92 ± 0.6% for muons with $p_T > 20$ GeV. Muons are required to pass basic quality cuts on the number and type of hits in the inner detector. They must lie in the range $|\eta| < 2.4$, and satisfy the same impact parameter cuts as electrons. They must also be isolated, with the sum of the transverse momenta of all tracks in a cone $\Delta R < 0.2$ around the muon satisfying $\Sigma(p_T^{\text{track}})/p_T^\mu < 0.1$.

Jets are reconstructed from topological clusters using the anti-$k_t$ algorithm [27] with radius parameter $R = 0.4$. The reconstructed jets are calibrated using $E_T$ and $\eta$ dependent correction factors based on MC simulation and validated with data [28]. They are required to have $E_T > 25$ GeV and $|\eta| < 4.5$. Jets are considered $b$-tagged if they contain a reconstructed displaced secondary vertex consistent with a $b$-decay [29]. The operating point chosen for this $b$-tag selection has an efficiency of 50% for $b$-jets in $t\bar{t}$ events in MC, and the mistag rate for non-$b$-jets has been measured to be between 0.1% and 2.0%, depending on the $|\eta|$ and $p_T$ of the jet [29]. The event missing transverse momentum $E_T^m$ is reconstructed starting from topological energy clusters in the calorimeters calibrated according to the type of the object to which they are associated. The momenta of any muons in the event are also taken into account in the $E_T^m$ measurement.

For this analysis, events are required to have at least one vertex with at least three associated tracks with $p_T > 400$ MeV. There must be exactly one reconstructed lepton candidate (electron or muon) with $p_T > 30$ GeV. In order to ensure that this analysis is statistically independent of the ATLAS $H \to ZZ \to \ell\ell\nu\nu$ analysis, events are vetoed if there is an additional lepton with $p_T > 20$ GeV, including electrons which only satisfy the looser identification cuts used in the $H \to ZZ \to \ell\ell\nu\nu$ analysis [30].

Events are required to have $E_T^m > 30$ GeV to account for an unobserved neutrino from $W \to \ell\nu$ decay. There must be exactly two jets ($H + 0$ jet sample) or exactly three jets ($H + 1$ jet sample) with $E_T > 25$ GeV and $|\eta| < 4.5$. The two jets with invariant mass $(m_{jj})$ closest to the mass of the $W$ boson are required to satisfy $71$ GeV < $m_{jj}$ < $91$ GeV. These two jets are taken as the $W$ decay jets and are required to lie in the range $|\eta| < 2.8$, where the jet energy scale (JES) is best known (to better than ±(4–8)% for $E_T > 25$ GeV [28].)

After this event selection, the background is expected to be dominated by $W +$ jets production. Other important backgrounds are $Z +$ jets, multijets (MJ) from QCD processes, top quark, and diboson ($WW$, $WZ$, and $ZZ$) production. In order to further reject backgrounds from top quark production, events are rejected if any of the jets is $b$-tagged.

Although the MC is not used to model the background in the final fit used to obtain limits, a combination of MC calculations and data-driven methods is used to better understand the background yields at this intermediate stage. Backgrounds due to $W/Z +$ jets, $t\bar{t}$, and diboson production are modeled using the ALPGEN [31], MC@NLO [32], and HERWIG [33] generators, respectively. A small contribution from $W/Z + \gamma$ events is generated using MADEVENT [34]. The shape of the MJ background is modeled using histograms derived from data samples selected in an identical way to the $H \to WW \to \ell\nu jj$ selection except that the electron identification requirements are loosened and the isolation requirement on muons is inverted. In the loosened selection, electrons satisfying the complete set of identification criteria are not included. Expected contributions from non-QCD processes to the MJ shape histogram are subtracted using MC predictions.

To normalize the MJ shape histogram, the loose lepton control sample selection is further relaxed by removing the $E_T^m$ cut to construct a shape template for the $E_T^m$ distribution for the MJ background. The normalizations of this MJ template and the corresponding template for $W/Z +$ jets taken from MC are fit to the observed $E_T^m$ distribution, and the resulting scale factors are then used to normalize the MJ and $W/Z +$ jets processes in comparisons between data and expectations. Both the gluon fusion and the vector boson fusion signal production processes are simulated using the POWHEG [35,36] event generator interfaced to PYTHIA [37], normalizing to the NNLO cross sections [19] shown in Table I.
In order to reconstruct the invariant mass \( m(\ell \nu jj) \) of the \( WW \) system, the mass constraint \( m(\ell \nu) = m(W) \) is used, where the neutrino transverse momentum \( p_T^\nu \) is taken from the event \( E_T \). This equation can have real or complex solutions. In the case of complex solutions, the event is rejected. This requirement rejects 45% of background events in both data and MC, but only 36% of MC signal events with \( m_H = 400 \) GeV. In the case of two real solutions, the solution with smaller neutrino longitudinal momentum \( |p_T^\nu| \) is taken, based on simulation studies. Table II shows the observed and expected numbers of events for signal and background after this full selection.

Figure 1 (top) shows the \( m(\ell \nu jj) \) distribution for this final sample. The expected signal for \( m_H = 400 \) GeV is also shown, scaled up by a factor of 2.7. The \( m(\ell \nu jj) \) resolution is \((7.5 \pm 0.6)\%\) at \( m_H = 400 \) GeV, depending mostly on the jet energy resolution as checked in data versus MC by various jet-balance techniques [38], and shows a \( 1/\sqrt{m_H} \) dependence over the range of this analysis. Limits are set using a maximum likelihood fit to the shape of the observed \( m(\ell \nu jj) \) distribution in the range \( 200 < m(\ell \nu jj) < 2000 \) GeV. The nonresonant background in this fit is modeled by the sum of two exponential functions. The normalization and slope of each exponential are unconstrained parameters in the fit. The double-exponential form for the total background is well justified by fits to the \( m(\ell \nu jj) \) distributions obtained by selecting events with \( m_{jj} \) just below \((50 < m_{jj} < 60 \) GeV\) or just above \((100 < m_{jj} < 110 \) GeV\) the \( W \) peak, respectively.

As a consistency check, the background parametrization was altered to use three exponentials and the shift in signal yield as compared to the nominal background shape was found to be small as compared to other uncertainties. The \( m(\ell \nu jj) \) distribution for the expected signal at each hypothesized \( m_H \) is modeled using the signal MC samples. The fit includes nuisance parameters which account for the uncertainty in the efficiency of the electron, muon, and jet reconstruction. The electron and muon efficiencies are varied within their uncertainties, leading to an uncertainty in the signal efficiency of \( \pm 1.6\% \) and \( \pm 0.6\% \), for electrons and muons, respectively. Varying the jet energy scale within its uncertainties yields a corresponding uncertainty of \( \pm 17\% \) in the expected signal, and smearing the jet energies within the uncertainty on their resolutions results in a signal uncertainty of \( \pm 8.6\% \). The limits also take into account a \( \pm 3.7\% \) uncertainty on the luminosity determination [39] and a \( \pm 19.4\% \) uncertainty on the predicted cross section [19], taken to be independent of mass. The off-shell effects and interference between the signal and backgrounds, which are discussed in Refs. [19,40,41], have been neglected. A conservative estimate of this uncertainty would be \( 150 \times m_H^3 \) (\( m_H \) in TeV), where the \( m_H^3 \) form is motivated by the scaling of the Higgs width with \( m_H \) and the normalization factor is chosen to give \( \sim 30\% \) at \( m_H = 600 \) GeV, based on Fig. 6 of Ref. [41]. If this were included, it would increase the total systematic error by less than 6% for \( m_H \leq 500 \) GeV, and as much as 15% for \( m_H = 600 \) GeV where the limit would be increased by \( \sim 18\% \).

### Table II

<table>
<thead>
<tr>
<th>( m_H ) [GeV]</th>
<th>( \sigma(gg \rightarrow H) ) [pb]</th>
<th>( \sigma(qq \rightarrow H) ) [pb]</th>
<th>( \text{BR}(H \rightarrow \ell^\pm \nu jj) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>2.4 ± 0.4</td>
<td>0.30 + 0.014 − 0.008</td>
<td>0.202</td>
</tr>
<tr>
<td>400</td>
<td>2.0 + 0.31 − 0.34</td>
<td>0.162 + 0.010 − 0.005</td>
<td>0.170</td>
</tr>
<tr>
<td>500</td>
<td>0.85 ± 0.15</td>
<td>0.095 + 0.0068 − 0.0032</td>
<td>0.160</td>
</tr>
<tr>
<td>600</td>
<td>0.33 + 0.063 − 0.058</td>
<td>0.058 + 0.005 − 0.002</td>
<td>0.164</td>
</tr>
</tbody>
</table>

### Table II

<table>
<thead>
<tr>
<th>( H(e\nu jj) + 0j )</th>
<th>( H(\mu\nu jj) + 0j )</th>
<th>( H(e\nu jj) + 1j )</th>
<th>( H(\mu\nu jj) + 1j )</th>
<th>( H + 0j ) or ( 1j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>W/Z + jets</td>
<td>5580 ± 150</td>
<td>6690 ± 430</td>
<td>3610 ± 140</td>
<td>3940 ± 360</td>
</tr>
<tr>
<td>Multijet</td>
<td>592 ± 16</td>
<td>222 ± 15</td>
<td>420 ± 16</td>
<td>195 ± 17</td>
</tr>
<tr>
<td>Top</td>
<td>87 ± 17</td>
<td>87 ± 18</td>
<td>271 ± 54</td>
<td>280 ± 56</td>
</tr>
<tr>
<td>Dibosons</td>
<td>226 ± 45</td>
<td>222 ± 45</td>
<td>89 ± 18</td>
<td>112 ± 23</td>
</tr>
<tr>
<td>Expected background</td>
<td>6490 ± 160</td>
<td>7220 ± 440</td>
<td>4390 ± 150</td>
<td>4530 ± 370</td>
</tr>
<tr>
<td>Data</td>
<td>6446</td>
<td>7201</td>
<td>4134</td>
<td>4380</td>
</tr>
</tbody>
</table>

| Expected Signal (\( m_H = 400 \) GeV) | 11 ± 3.0 | 9 ± 2.4 | 13 ± 3.5 | 10 ± 2.7 | 43 ± 12 |
FIG. 1 (color online). Top: the reconstructed invariant mass $m(\ell\nu jj)$ in the data summed over lepton flavor and jet multiplicity. The expected backgrounds are also shown. Bottom: the difference between data and the fitted nonresonant background. The expected contribution from Higgs boson decays for $m_H = 400$ GeV in the SM is also shown, multiplied by a factor of 2.7.

Figure 1 (bottom) shows the difference between the $m(\ell\nu jj)$ distribution in data and the fitted background. There is no indication of any significant excess. Limits are extracted using the Profile Likelihood [42] as a test statistic and following the $CL_s$ procedure described in Ref. [43].

Figure 2 shows the 95% CL upper bound on the cross-section times branching ratio for Higgs production in units of the Standard Model prediction, $\sigma \times BR_{H \rightarrow WW}/(\sigma \times BR_{H \rightarrow WW})_{SM}$, as a function of $m_H$. The observed cross section limit for $m_H = 400$ GeV is 3.1 pb, or 2.7 times the SM prediction, while the corresponding expected limits are 5.2 pb or 4.5 times the SM expectation. In the SM with an additional heavy fourth generation [44,45], the gluon fusion mechanism for production of a Higgs boson is expected to be substantially enhanced. Within the four generation context, a Higgs boson is excluded at 95% CL by the present data over the range $m_H = 310–430$ GeV.

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFi, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; M00T CR, MPO CR and VSC CR, Czech Republic; DNRF, DSNRC and Lundbeck Foundation, Denmark; ARTEMIS, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNAS, Georgia; BMBF, DFG, HGF, MPG and AvH Foundation, Germany; GSRT, Greece; ISF, MINEVA, GIF, DIP and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNI5W, Poland; GdINES and ICT, Portugal; MERSYS (MECTS), Romania; MES of Russia and ROSATOM, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; AARRS and MVZT, Slovenia; DST/NRF, South Africa; MIECN, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular, from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASG (Taiwan), RAL (UK) and BNL (USA) and in the Tier-2 facilities worldwide.
