Search for single production of a vector-like quark via a heavy gluon in the 4b final state with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV

ATLAS Collaboration

DOI
10.1016/j.physletb.2016.04.061

Publication date
2016

Document Version
Final published version

Published in
Physics Letters B

License
CC BY

Citation for published version (APA):
Search for single production of a vector-like quark via a heavy gluon in the 4b final state with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV

ATLAS Collaboration*

1. Introduction

Composite Higgs [1–4] models interpret the Higgs boson discovered at the Large Hadron Collider (LHC) [5] as a pseudo-Goldstone boson resulting from spontaneous symmetry breaking in a new strongly coupled sector, thus addressing the naturalness problem, the extreme fine tuning required in the Standard Model (SM) to cancel quadratically divergent radiative corrections to the Higgs boson mass. A generic prediction of these models is the existence of massive vector-like quarks (VLQ). These VLQs are expected to mix mainly with the third family of quarks of the SM [6–8], leading to partial compositeness. Colour octet resonances (massive gluons) also occur naturally in these models [6,7,9,10].

Searches for vector-like quarks in the ATLAS and CMS experiments, in both the pair and single production processes [11–22], constrain their mass to be above 700–900 GeV. This analysis is a search for single production of a vector-like quark $b_H$ of charge $-1/3$ via the $s$-channel exchange of a heavy colour octet vector resonance $G^*$, using data recorded by the ATLAS detector at the LHC. The search is performed for the process $Hbb\to b_Hb\to bbb$ (see Fig. 1), based on Ref. [23] and using the benchmark model of Ref. [9]. This simplified minimal composite Higgs model has a composite sector with a global $\text{SU}(3)_c \times \text{SU}(2)_{L} \times \text{SU}(2)_{R} \times \text{U}(1)_Y$ symmetry and an elementary sector which contains the SM particles but not the Higgs boson. Physical states of the composite sector include the heavy gluon $G^*$, a composite Higgs boson and heavy vector-like quarks of charge $5/3, 2/3, -1/3$ and $-4/3$. Among these heavy quarks, there is one singlet of charge $2/3$ which mixes with the right-handed top quark of the SM with an angle $\theta_{Ht}$, and similarly one singlet of charge $-1/3$ which mixes with the right-handed bottom quark of the SM with an angle $\theta_{Hb}$. After mixing between the gluons from the elementary and composite sectors by an angle $\theta_{H}$, the physical state of the heavy gluon has a coupling $g_{t} \cos \theta_{H}$ to composite states, where $g_{t} = g_{s}/\sin \theta_{H}$ and $g_{s}$ is the coupling of the SM gluon. The other parameters of the model are the composite fermion masses, assumed to be universal, the heavy gluon mass $m_{G^*}$ and two Yukawa couplings $Y_{\gamma}$ and $Y_{\rho}$. In a large part of the parameter space, the lightest of the new heavy quarks is $b_H$, of charge $-1/3$, and in this model it decays exclusively to $Hb$. In Ref. [23], the condition $m_{b_H} = m_{G^*}/2$ is applied, with the result that pair production of the heavy partners is kinematically forbidden and the width of $G^*$ is consequently not too large. In the search presented here, the phase space is extended to $m_{b_H} \geq m_{G^*}/2$. When $m_{b_H} < m_{G^*}/2$, present results on pair produc-

---

* E-mail address: atlas.publications@cern.ch.

1 Charge conjugate states are implied in the following text.
tion of vector-like quarks can be recast in a model with a massive colour octet [24].

For high masses of the $G^*$ and $B_H$ resonances, the Higgs boson is highly boosted and the decay products are reconstructed in a single large-radius (large-$R$) jet in the detector, whereas for lower masses the four $b$-quarks are reconstructed as separate small-radius jets. The analysis uses two sets of selection criteria to target these two cases.

2. The ATLAS detector

The ATLAS detector, located at the LHC, is described in detail in Ref. [25]. It covers nearly the full solid angle around the collision point. The inner detector is surrounded by a solenoid that produces a 2 T axial magnetic field. The tracks of charged particles are reconstructed with a high-granularity silicon pixel and microstrip detector for $|\eta| < 2.5$. A straw-tube transition radiation detector extends the tracking to larger radii and provides electron/pion discrimination. The electromagnetic calorimeter consists of a barrel and end-cap lead/liquid-argon (LAr) sections with an accordion geometry covering $|\eta| < 3.2$, preceded by a thin presampler, covering $|\eta| < 1.8$, which allows corrections for fluctuations in upstream energy losses. A copper/LAr electromagnetic calorimeter covers the very forward angles. Hadronic calorimetry is installed in the barrel region, $|\eta| < 1.7$, using steel as the absorber and scintillator tiles as the active material. In the endcaps, copper/LAr calorimeters cover 1.5 $< |\eta| < 3.2$ followed by a forward calorimeter based on tungsten absorbers in LAr as sensitive medium, up to $|\eta| = 4.9$. Surrounding the hadronic calorimeters are large toroidal magnets whose magnetic fields deflect the trajectories of charged particles exiting the barrel and end-cap calorimeters. The muon spectrometer uses monitored drift tubes for tracking in $|\eta| < 2.7$ with cathode strip chambers in the innermost station for $|\eta| > 2.0$. A dedicated muon trigger is provided by resistive plate chambers in the barrel and thin-gap chambers in the end-cap, covering $|\eta| < 2.4$.

A three-level trigger system, consisting of a hardware Level-1 trigger and two software-based trigger levels reduce the event rate to be recorded to less than about 400 Hz.

3. Data and simulation

Data used in this analysis correspond to an integrated luminosity of 19.5 fb$^{-1}$ of pp collisions collected at the LHC at a centre-of-mass energy of $\sqrt{s} = 8$ TeV, with all the essential elements of the ATLAS detector fully operational and stable.

Simulated signal and background samples are produced by Monte Carlo (MC) event generators and passed through a GEANT4 [26] simulation of the ATLAS detector [27]. Additional events from the same and neighbouring bunch crossings (pile-up) are included by adding simulated diffractive and non-diffractive pp collisions to hard-scattering events. The pile-up rate is reweighted in accordance with the luminosity profile of the recorded data. All simulated events are then reconstructed using the same reconstruction software as the data.

Signal samples based on the model discussed in Ref. [23] are generated with MADGRAPH5_aMC@NLO [28], using CTETQ6L1 [29].

2 ATLAS experiment uses a right-handed coordinate system with its origin at the nominal interaction point in the centre of the detector and the $z$-axis along the beam pipe. The $x$-axis points from the IP to the centre of the LHC ring, and the $y$-axis points upward. Cylindrical coordinates ($\rho, \phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. The distance in $\eta$-$\phi$ space is referred to as $\Delta R = \sqrt{\Delta \eta^2 + (\Delta \phi)^2}$. parton distribution functions (PDFs), in the mass region $m_G/2 \leq m_{b}\leq m_G$, with 1 TeV $< m_{b} < 3$ TeV, in steps of 250 GeV in $m_{G}$ and in steps of 125 GeV in $m_{b}$. The Higgs boson mass is set to 126 GeV and its branching ratio BR ($H \rightarrow bb$) to 56.1% [30].

The parameters of the model are set as in Ref. [23]: $\gamma = 3$, $Y_T = Y_r = 3$, $\sin \theta_H = \sin \theta_B = 0.6$.

The event selection requires at least two $b$-jets in the final state. Multijet events from strong interactions have a large cross section and are the dominant background. Due to the large number of events required to simulate this background and the difficulty of modelling it accurately, it is evaluated using a data-driven method, as described in Section 6. Other background contributions include top-pair and single-top-quark production, generated with POWHEG-Box [31–33] interfaced to PYTHIA [34] using CT10 PDFs [35]. The $t\bar{t}$ sample is normalised to the theoretical calculation performed at next-to-next-to-leading order (NNLO) including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms with Top++2.0 [36,37], giving an inclusive cross section of 253.1$^{+12.3}_{-11.4}$ pb [38]. Samples of $t\bar{t} + Z$ and $t\bar{t} + H$ events are generated with PYTHIA and CTEQ6L1 PDFs. The SHERPA [39] generator, with CT10 PDFs, is used to simulate $W/Z + \text{jets}$ samples with lepton decay of the vector bosons. SHERPA is also used to generate $Z + \text{jets}$ events, with $Z \rightarrow bb$, where the extra jets are produced inclusively. Contributions from diboson backgrounds—$WW$, $WZ$ and $ZZ$—are estimated to be negligible.

4. Object reconstruction

The final state consists of four jets from $b$-quarks ($b$-jets), two of which come from the Higgs boson decay. If the Higgs boson is sufficiently boosted, having a transverse momentum $p_T \gtrsim 300$ GeV, the two $b$-jets may be merged into a single jet with a large radius parameter (large-$R$ jet) and therefore two different jet definitions are used.

Jets with smaller radius parameter, or small-$R$ jets, are reconstructed from calibrated calorimeter energy clusters [40,41] using the anti-$k_t$ algorithm [42] with a distance parameter $R = 0.4$. The high $p_T$ threshold used in the event selection ensures that the contamination of jets from pile-up is small. To ensure high-quality reconstruction of central jets while rejecting most jets not coming from hard-scattering events, criteria as described in Ref. [43] are applied. Jets are corrected for pile-up by a jet-area subtraction method and calibrated by a jet energy scale factor [44]. They are required to have $p_T > 50$ GeV and $|\eta| < 2.5$.

Small-$R$ jets are identified as containing a $b$-hadron ($b$-tagged) by a multivariate algorithm [45]. This algorithm was configured to give a $b$-tagging efficiency of 70% in simulated $t\bar{t}$ events, with a mistag probability of about 1% for gluon and light-quark jets and of about 20% for $c$-quark-initiated jets. The $b$-tagging efficiency in simulated events is corrected to account for differences observed between data and simulation.

Large-$R$ jets are reconstructed using the anti-$k_t$ algorithm with $R = 1.0$. Jet trimming [46,47] is applied to reduce the contamination from pile-up and underlying-event activity: subjects are formed using the $k_t$ algorithm [48] with $R = 0.3$ and subjects with $p_T/(\text{subj})/p_T/(\text{jet}) < 5\%$ are removed.

Leptons are vetoed in this analysis to reduce background involving leptonically decaying vector bosons. Electron candidates with $p_T > 7$ GeV are identified in the range $|\eta| < 2.47$ from energy clusters in the electromagnetic calorimeter, matched to a track in the inner detector. Requirements of ‘medium’ quality, as defined in Ref. [49], are applied together with two isolation criteria: the scalar sum of the transverse momentum (energy) within a radius $\Delta R = 0.2$ around the electron candidate has to be less than
15% (14%) of the electron $p_T(E)$. Muons with $p_T > 7$ GeV and $|\eta| < 2.4$ are reconstructed from matched tracks in the muon spectrometer and the inner detector. Quality criteria are applied, as described in Ref. [50], and an isolation requirement is applied: the scalar sum of the transverse momentum of tracks within a radius $\Delta R = 0.2$ around the muon candidate has to be less than 10% of the muon $p_T$.

### 5. Event selection

Because of the very high hadronic background at the LHC, it is not possible to have adequate Monte Carlo statistics for multi-jet events. The uncertainties in the quality of simulation of b-jets at high-$p_T$ can also be large. For these reasons, for each mass pair $(m_{G^*}, m_{B_H})$ being tested, a data-driven technique was used to evaluate the expected background, as described in Section 6. The technique requires that we define control regions orthogonal to the signal regions. A blind analysis is performed, in which the background is first evaluated without initial knowledge of the data in the signal regions. In order to test the large number of mass pair hypotheses, all signal region cuts are applied except the Higgs mass window which is blinded when evaluating the background in the signal regions.

#### 5.1. Event preselection

Events in the signal region are first preselected according to the following criteria (see end of Section 5.2 for the signal region definition).

- They satisfy a combination of six triggers requiring multiple jets and b-jets for various $p_T$ thresholds, where b-jets are identified by a dedicated online b-tagging algorithm. This combination of triggers is >99% efficient for signal events passing the offline selection, across the $B_H$ and $G^*$ mass ranges considered in this analysis.
- They are vetoed if they contain reconstructed isolated leptons (e or $\mu$) in order to reduce the contribution from $W/Z +$ jets and t$\bar{t}$ backgrounds.
- At least three small-$R$ b-tagged jets must be present in the signal region.
- The invariant mass of the system composed of all selected $R = 0.4$ jets is required to be greater than 600 GeV.

Two event topologies are considered for the signal, depending on the boost of the Higgs boson. Highly boosted Higgs bosons are reconstructed using large-$R$ jets as described in Section 4 and this topology corresponds to the merged scenario (see Section 5.2). If no large-$R$ jet is found, an attempt is made to reconstruct the Higgs boson from two small-$R$ jets (see Section 5.3). The acceptance times reconstruction efficiency for the combined yields of the two topologies varies from 5% to 20% depending on the masses of the $G^*$ and $B_H$.

#### 5.2. Merged selection

The signal region for the merged case consists of the following requirements.

- A large-$R$ jet must be present with $p_T > 300$ GeV and $|\eta| < 2.0$ and mass in the range [90, 140] GeV. The mass window was optimised based on the signal sensitivity. If more than one such large-$R$ jet is present, the Higgs candidate is chosen to be the one with mass closest to 126 GeV. At least one b-tagged jet must be matched to it within a distance $\Delta R = 1.0$.
- There must be at least two additional b-tagged jets separated from the Higgs boson candidate, $\Delta R (H, j) > 1.4$. The two with the highest $p_T$ are used to reconstruct the $G^*$ and $B_H$ candidates.

Once the Higgs boson candidate has been identified as above, there remains an ambiguity in assigning the other jets to the vector-like quark $B_H$. The four-momentum of the $B_H$ candidate is reconstructed as the four-momentum sum of the Higgs boson candidate and either the next-to-leading-$p_T$ (category 1) or the leading-$p_T$ (category 2) b-jet away from it, depending on the assumed mass difference between $G^*$ and $B_H$. For large $G^*-B_H$ mass difference, the $B_H$ and b-quark from $G^*$ splitting have high momentum and therefore the jet from the subsequent $B_H$ decay is likely to be the next-to-leading jet. For a small mass difference the opposite is true since in this latter case the $B_H$ decay products are more boosted than the $G^*$ splitting products. For each $(m_{G^*}, m_{B_H})$ pair, the category which has the higher probability that the correct pairing is formed is chosen, based on the simulated signal events. Finally, the $G^*$ four-momentum is reconstructed as the four-momentum sum of the Higgs boson jet and the two leading-$p_T$ b-jets not matched to the Higgs boson candidate.

Different signal regions are defined for the different $(m_{G^*}, m_{B_H})$ mass pair hypotheses. They are characterised by the choice of category defined above as well as by lower cuts on the reconstructed masses of $G^*$ and $B_H$ candidates. Five inclusive signal regions were defined, with the minimum mass of the $G^*$ candidate ranging from 0.8 to 1.8 TeV and of the $B_H$ candidate from 0.5 to 1 TeV; these are shown in Table 1. No upper cut on the resonance masses was set since the multijet background distribution falls rapidly and the resonance widths become larger for high masses. For each mass pair considered, the signal region that gives the maximum signal sensitivity, the ratio of the expected number of signal events to the square root of the number of background events, is chosen.

#### 5.3. Resolved selection

Events in the resolved signal region are required to satisfy the following criteria.

- In order to be able to later combine the results with the merged channel, events are required to fail the merged selection criteria.
- Events are required to have exactly four small-$R$ jets with $p_T > 50$ GeV and $|\eta| < 2.5$, with at least three of these jets being b-tagged. The Higgs boson candidate is reconstructed using the two jets with invariant mass nearest to 126 GeV. The invariant mass is required to be in the interval [90, 140] GeV and the transverse momentum of the dijet system $p_T(jj) > 200$ GeV.

The four-momentum of the $B_H$ candidate is reconstructed from the four-momentum sum of the Higgs candidate and either the
leading or the next-to-leading-\pt jet away from the Higgs boson jets, depending on the \Gstar-\BH mass splitting. As in the merged case, for each pair of masses considered the category is chosen to be the one with the lower mis-assignment rate of jets, based on samples of simulated signal events. Inclusive signal regions are defined by lower minimum mass values identical to the merged case, and shown in Table 1. Each mass pair is assigned to the same SR for the merged and resolved analysis. The four-momentum of the \Gstar candidate is reconstructed from the four-momentum sum of the four jets in the event.

6. Modelling of the multijet background

The 'ABCD' data-driven method is used to estimate the multijet background. For each of the ten signal regions, three control regions orthogonal to the signal region are defined: region B has all the signal region selection criteria mentioned in Section 5 applied, including the lepton vetoes and lower cuts on the masses of \BH and \Gstar candidates, but the Higgs boson candidate mass is required to be outside the interval \(90, 140\) GeV; region C has all the signal region selection requirements, but requires exactly two jets to be b-tagged; and region D has the Higgs boson candidate outside the Higgs boson mass window and exactly two b-tagged jets. In regions C and D, only one of the two jets not associated with the Higgs boson candidate is b-tagged. The number of multijet (MJ) events expected in the signal region (SR) is then evaluated according to

\[
N_{\text{SR}}^{\text{MJ}} = N_{\text{B}}/N_{\text{D}} \times N_{\text{C}},
\]

where \(N_{\text{X}}\) is the number of events in region X, after having removed the top-quark, diboson and other electroweak background contributions as determined from MC simulations.

This estimate assumes that no bias results from the choice of control regions. To evaluate and potentially correct for the effect of any biases, a re-weighting is performed on two kinematic distributions, the leading-jet \pt and the \Delta\pt between the reconstructed Higgs boson candidate and the leading jet not associated with it. Control regions C and D (B and D) are re-weighted, using a method similar to Ref. [51], to have the same shape as in control region B (C) with weights obtained from \(N_{\text{B}}/N_{\text{D}} (N_{\text{B}}/N_{\text{D}})\) per bin. The effect of this re-weighting is found to be negligible and therefore no correction is applied.

A validation region is defined as the 15 GeV sideband regions outside the Higgs boson candidate mass window, i.e. 75–90 GeV and 140–155 GeV, for each signal region. The contribution from multijet background is estimated as above, but with the control regions B and D excluding these validation regions and region C now being the two sidebands. It is then compared to the number of observed data events, after adding back the simulation-based background, in these regions. Table 2 shows that the expected and observed numbers of events agree well in the validation regions for the merged- and resolved-channel signal regions.

7. Systematic uncertainties

Systematic uncertainties from several sources affect the expected numbers of background and signal events. Table 3 shows the estimated size of the different components.

The statistical uncertainty in the data control regions used for the estimation of the multijet background is considered as part of the statistical error.

There is an uncertainty in the number of background events due to the difference between the observed and estimated numbers of events in each of the validation regions. In each validation region.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>VR1</th>
<th>VR2</th>
<th>VR3</th>
<th>VR4</th>
<th>VR5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected</td>
<td>563 ± 16</td>
<td>213 ± 10</td>
<td>1680 ± 29</td>
<td>135 ± 8</td>
<td>45 ± 4</td>
</tr>
<tr>
<td>Observed</td>
<td>558</td>
<td>184</td>
<td>1666</td>
<td>137</td>
<td>35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>VR1</th>
<th>VR2</th>
<th>VR3</th>
<th>VR4</th>
<th>VR5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected</td>
<td>1065 ± 21</td>
<td>337 ± 11</td>
<td>3758 ± 50</td>
<td>242 ± 10</td>
<td>63 ± 5</td>
</tr>
<tr>
<td>Observed</td>
<td>1073</td>
<td>124</td>
<td>3906</td>
<td>238</td>
<td>56</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>Systematic uncertainty</th>
<th>SR1</th>
<th>SR2</th>
<th>SR3</th>
<th>SR4</th>
<th>SR5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background estimation</td>
<td>5%</td>
<td>15%</td>
<td>8%</td>
<td>10%</td>
<td>27%</td>
</tr>
<tr>
<td>(\bar{\text{f}}) cross section</td>
<td>+1.0% − 1.1%</td>
<td>+0.8% − 0.9%</td>
<td>+1.2% − 1.4%</td>
<td>+0.8% − 0.9%</td>
<td>+0.6% − 0.7%</td>
</tr>
<tr>
<td>JES small-R</td>
<td>+0.29%</td>
<td>+0.15%</td>
<td>+0.01%</td>
<td>−0.32%</td>
<td>+0.20%</td>
</tr>
<tr>
<td>JES small-R</td>
<td>+0.9% − 0.8%</td>
<td>+1.6% − 0.7%</td>
<td>+1.0% − 1.0%</td>
<td>+0.9% − 1.0%</td>
<td>+1.5% − 1.0%</td>
</tr>
<tr>
<td>JES/JMS large-R</td>
<td>+0.31% − 1.5%</td>
<td>+1.3% − 1.5%</td>
<td>+0.13% − 1.9%</td>
<td>+0.9% − 0.8%</td>
<td>+1.6% − 0.20%</td>
</tr>
<tr>
<td>b-tagging</td>
<td>+0.18% − 0.18%</td>
<td>+0.23% − 0.33%</td>
<td>+0.24% − 0.18%</td>
<td>&lt;0.01%</td>
<td>+1.6% − 0.01%</td>
</tr>
<tr>
<td>Luminosity</td>
<td>0.3%</td>
<td>0.3%</td>
<td>0.3%</td>
<td>0.2%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Data/MC statistical (CR)</td>
<td>2.2%</td>
<td>4%</td>
<td>1.3%</td>
<td>4%</td>
<td>8%</td>
</tr>
<tr>
<td>Total (stat.)</td>
<td>2.7%</td>
<td>5%</td>
<td>1.5%</td>
<td>6%</td>
<td>10%</td>
</tr>
<tr>
<td>Total (syst.)</td>
<td>6%</td>
<td>15%</td>
<td>4%</td>
<td>11%</td>
<td>28%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Systematic uncertainty</th>
<th>SR1</th>
<th>SR2</th>
<th>SR3</th>
<th>SR4</th>
<th>SR5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background estimation</td>
<td>3.5%</td>
<td>6%</td>
<td>4%</td>
<td>8%</td>
<td>16%</td>
</tr>
<tr>
<td>(\bar{\text{f}}) cross section</td>
<td>+0.24% − 0.27%</td>
<td>+0.20% − 0.23%</td>
<td>+0.31% − 0.4%</td>
<td>+0.23% − 0.26%</td>
<td>+0.17% − 0.20%</td>
</tr>
<tr>
<td>JER small-R</td>
<td>+0.17%</td>
<td>+0.32%</td>
<td>+0.18%</td>
<td>−0.37%</td>
<td>−0.5%</td>
</tr>
<tr>
<td>JES small-R</td>
<td>+0.8% − 0.6%</td>
<td>+0.7% − 0.6%</td>
<td>+0.6% − 0.7%</td>
<td>+0.8% − 0.7%</td>
<td>+1.0% − 0.8%</td>
</tr>
<tr>
<td>b-tagging</td>
<td>+0.5% − 0.4%</td>
<td>+0.5% − 0.30%</td>
<td>+0.5% − 0.4%</td>
<td>+0.4% − 0.4%</td>
<td>+0.7% − 0.7%</td>
</tr>
<tr>
<td>Luminosity</td>
<td>0.13%</td>
<td>0.13%</td>
<td>0.15%</td>
<td>0.15%</td>
<td>0.11%</td>
</tr>
<tr>
<td>Data/MC statistical (CR)</td>
<td>1.6%</td>
<td>2.7%</td>
<td>1.0%</td>
<td>3.3%</td>
<td>6%</td>
</tr>
<tr>
<td>Total (stat.)</td>
<td>2.1%</td>
<td>4%</td>
<td>1.0%</td>
<td>4%</td>
<td>8%</td>
</tr>
<tr>
<td>Total (syst.)</td>
<td>4%</td>
<td>7%</td>
<td>4%</td>
<td>8%</td>
<td>17%</td>
</tr>
</tbody>
</table>
region, if the observed number of events is compatible with the estimated number within one standard deviation (calculated as the sum in quadrature of the relative statistical errors of the two), this standard deviation is considered to be the background estimation uncertainty. Otherwise, the background uncertainty is considered to be the fractional difference between the observed and estimated numbers of events. This is the largest uncertainty, ranging from 5% in SR1 to 27% in SR5 for the merged case, and from 3.5% in SR1 to 16% in SR5 for the resolved case.

The $t\bar{t}$ contribution dominates the simulation-based background. The theoretical uncertainty on its cross section is taken to be 6%, as discussed in Section 3.

Uncertainties due to the calibration and modelling of the detector affecting the simulation-based background estimates in the control and signal regions are principally due to the jet energy scale (JES) and jet energy resolution (JER). JES uncertainties for small-$R$ jets include contributions from detector reconstruction and from different physics modelling and evaluation methods [52]. Uncertainties leading to a higher (lower) yield than the nominal value are added in quadrature to the total JES up (down) uncertainty. To evaluate the impact of JER for small-$R$ jets, energies of simulated jets are smeared to be consistent with the JER measured in data. The JER systematic uncertainty is the difference between the nominal and smeared values.

JES uncertainties for large-$R$ jets in the central region are evaluated as described in Ref. [47]. The jet mass scale (JMS) uncertainty is 4–5% for $p_T < 700$ GeV and increases linearly with $p_T$ to about 8% in the range $900 \leq p_T \leq 1000$ GeV.

The total uncertainty in the measured $b$-tagging efficiency was evaluated in Ref. [53] and is $p_T$ and $\eta$ dependent. For high-$p_T$ jets, the systematic uncertainty is derived from simulation. It is estimated here for the simulation-based backgrounds, accounting for the statistical uncertainty, the error on the generator-dependent scale factors, the track momentum scale, resolution and efficiency systematic uncertainties, and the extrapolation uncertainties for light jets. It is at or below the percent level and always dominated by the background estimation.

The predicted signal is not confined to the signal region: it could also constitute a fraction of the observed data in the control regions. The effect of this potential contamination on the statistical procedure is described in Section 8.

Systematic uncertainties due to detector effects also affect the VLQ signal yields. They are dominated by the $b$-tagging uncertainties, ranging from 16% to 40% depending on $p_T$, while other sources of systematic uncertainties listed above are below 5%. Theoretical uncertainties in the signal cross section due to the choice of PDFs are estimated from CTEQ6.6 with its 22 eigenvector sets [29].

8. Results

After applying all selection criteria in the signal regions, the multijet background in the Higgs boson candidate mass window is evaluated according to Eq. (1). Mass distributions of reconstructed Higgs boson candidates are shown in Fig. 2 for the merged and resolved cases in SR3. The observed data and the background predictions are consistent within statistical and systematic uncertainties.

For each pair of mass points considered, the expected signal yield, based on the benchmark model, is evaluated in the corresponding signal region. These yields result from the signal $\sigma \times (A \times \epsilon)$, where $\sigma$ is the cross section including all the branching fractions and $(A \times \epsilon)$ is the acceptance times reconstruction
Table 4
Observed data and background yields in the different signal regions for the merged and resolved cases. The first error is statistical and the second is systematic, while for individual background contributions only the statistical error is shown. Statistical errors on the numbers of data events in the control regions used to estimate the multijet background are included in the total statistical error. The row $t\bar{t}$/top includes $t\bar{t}$, single-top and $t\bar{t} + V/H$ backgrounds while $W/Z +$ jets includes leptonic and hadronic decays of the vector boson.

<table>
<thead>
<tr>
<th>Background</th>
<th>SR1</th>
<th>SR2</th>
<th>SR3</th>
<th>SR4</th>
<th>SR5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>1104 ± 27</td>
<td>398 ± 16</td>
<td>3372 ± 49</td>
<td>259 ± 12</td>
<td>85 ± 7</td>
</tr>
<tr>
<td>$t\bar{t}$/top</td>
<td>107 ± 4</td>
<td>30.0 ± 2.3</td>
<td>398 ± 8</td>
<td>18.3 ± 1.9</td>
<td>4.2 ± 1.0</td>
</tr>
<tr>
<td>$W/Z +$ jets</td>
<td>10.5 ± 1.3</td>
<td>4.4 ± 0.9</td>
<td>30.1 ± 1.9</td>
<td>2.6 ± 0.8</td>
<td>0.8 ± 0.5</td>
</tr>
<tr>
<td>Total BG</td>
<td>1222 ± 33 ± 70</td>
<td>432 ± 20 ± 60</td>
<td>3800 ± 60 ± 150</td>
<td>280 ± 16 ± 30</td>
<td>90 ± 9 ± 25</td>
</tr>
<tr>
<td>Data</td>
<td>1310</td>
<td>456</td>
<td>3827</td>
<td>287</td>
<td>89</td>
</tr>
</tbody>
</table>

Resolved

<table>
<thead>
<tr>
<th>Background</th>
<th>SR1</th>
<th>SR2</th>
<th>SR3</th>
<th>SR4</th>
<th>SR5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multijet</td>
<td>1985 ± 34</td>
<td>639 ± 18</td>
<td>8580 ± 90</td>
<td>523 ± 18</td>
<td>141 ± 9</td>
</tr>
<tr>
<td>$t\bar{t}$/top</td>
<td>64.2 ± 3.2</td>
<td>17.7 ± 1.8</td>
<td>353 ± 8</td>
<td>15.4 ± 1.6</td>
<td>3.3 ± 0.7</td>
</tr>
<tr>
<td>$W/Z +$ jets</td>
<td>35.0 ± 3.3</td>
<td>12.7 ± 1.8</td>
<td>142 ± 6</td>
<td>12.8 ± 2.2</td>
<td>2.6 ± 0.4</td>
</tr>
<tr>
<td>Total BG</td>
<td>2080 ± 40 ± 80</td>
<td>669 ± 25 ± 50</td>
<td>9080 ± 90 ± 340</td>
<td>551 ± 23 ± 50</td>
<td>147 ± 12 ± 25</td>
</tr>
<tr>
<td>Data</td>
<td>2106</td>
<td>706</td>
<td>8927</td>
<td>568</td>
<td>122</td>
</tr>
</tbody>
</table>

Table 5
Combined limits, in fb, on $\sigma (pp \rightarrow G^* \rightarrow B_{t\bar{t}}b) \times BR (B_{t\bar{t}} \rightarrow Hb) \times BR (H \rightarrow b\bar{b})$. First and second entries in each cell give the expected and observed limits, respectively. The third entry gives the cross section in fb predicted by the benchmark model. Red cells are excluded and green cells are not excluded at 95% C.L. Cases where $m_{G^*} > 2m_{B_{t\bar{t}}}$ are not considered in this analysis and are marked in yellow. (For interpretation of the references to colour in this table legend, the reader is referred to the web version of this article.)

| $m_{G^*}$ [TeV] | | |
|-----------------|------------------|------------------|------------------|------------------|------------------|
| 2.0             | 62^{+102}_{-22}  | 68               | 62               | 64               | 74               |
| 1.875           |                  |                  |                  |                  |                  |
| 1.75            |                  |                  |                  |                  |                  |
| 1.625           |                  |                  |                  |                  |                  |
| 1.50            |                  |                  |                  |                  |                  |
| 1.375           |                  |                  |                  |                  |                  |
| 1.25            |                  |                  |                  |                  |                  |
| 1.125           |                  |                  |                  |                  |                  |
| 1.0             |                  |                  |                  |                  |                  |
| 0.875           |                  |                  |                  |                  |                  |
| 0.75            | 232^{+130}_{-72} | 232              | 172              | 59               | 90               |
| m_{G^*} [TeV] → | 1.0              | 1.25             | 1.5              | 1.75             | 2.0              | 2.25             | 2.50             | 2.75             | 3.0              |

The efficiency of the signal selection cuts. The amount of contamination, defined as the expected ratio of the number of signal events in control regions B, C, or D to that in the signal region, is also estimated.

Table 4 shows the expected and observed background event yields in each of the signal regions for the merged and resolved cases. No significant excess of data events is found compared to the expected SM background. Taking into account the number of
expected background events in each of the signal regions and the yield of signal events for each test mass pair, together with all statistical and systematic uncertainties, upper limits at the 95% confidence level (CL), using the CL$_s$ prescription [54] and RooStats [55], are set on the cross section times the branching fraction of a signal, combining results from the merged and resolved analyses. To account for possible contamination of the control regions by signal, an iterative procedure is used: a 95% CL limit is first obtained assuming no contamination in the control regions. The contamination in regions B, C, D is then calculated, assuming a signal corresponding to that limit, and the multijet background is then re-evaluated. The procedure is repeated until it converges to a stable value. Expected and observed limits on the cross section $\sigma(pp \rightarrow G^* \rightarrow B_H\overline{b}) \times BR(B_H \rightarrow Hb) \times BR(H \rightarrow b\bar{b})$, where $\sigma(pp \rightarrow G^* \rightarrow B_H\overline{b})$ represents the cross section of the process $pp \rightarrow G^* \rightarrow B_H\overline{b}$ and its complex conjugate, as well as the theoretical cross section for the benchmark model, with its theoretical uncertainty, are shown in Table 5. Limits for the particular cases where $m_{BH} = m_{G^*}/2$ and $m_{BH} = m_{G^*} - 250$ GeV are shown in Figs. 3 and 4.

9. Conclusion

A search for a heavy gluon and a charge $-1/3$ vector-like quark in the process $pp \rightarrow G^* \rightarrow B_H\overline{b}$, with $B_H \rightarrow Hb$ and $H \rightarrow b\bar{b}$, has been performed using an integrated luminosity of 19.5 fb$^{-1}$ of pp collision data recorded at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC. The main background, multijet production, is estimated with a data-driven technique. Five signal regions are defined based on the choice of jet assignment to the $B_H$ candidate and on lower mass requirements for the reconstructed $G^*$ and $B_H$. No significant excess over the SM predictions is observed and upper limits have been set at the 95% confidence level on the total cross section times branching ratio in the two-dimensional plane of $m_{G^*}$ vs. $m_{BH}$ with $m_{G^*} \leq 2m_{BH}$. Using a benchmark model presented in

Ref. [23], a lower limit of 2.0 TeV on the $G^*$ mass is obtained when $m_{G^*} = 2m_{BH}$.

Acknowledgements

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; BMWFV and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/Irfu, France; GNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNISW and NCN, Poland; FCT, Portugal; MINEA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, Canarie, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, FP7, Horizon 2020 and Marie Sklodowska-Curie Actions, European Union; Investissements d'Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; the Royal Society and Leverhulme Trust, United Kingdom.

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), ASGC (Taiwan), RAL (UK) and BNL (USA) and in the Tier-2 facilities worldwide.

References


Also at Department of Physics, King’s College London, London, United Kingdom.

Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan.

Also at Novosibirsk State University, Novosibirsk, Russia.

Also at TRIUMF, Vancouver, BC, Canada.

Also at Department of Physics & Astronomy, University of Louisville, Louisville, KY, United States.

Also at Department of Physics, California State University, Fresno, CA, United States.

Also at Department of Physics, University of Fribourg, Fribourg, Switzerland.

Also at Departament de Fisica de la Universitat Autonoma de Barcelona, Barcelona, Spain.

Also at Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Portugal.

Also at Tomsk State University, Tomsk, Russia.

Also at Università di Napoli Parthenope, Napoli, Italy.

Also at Institute of Particle Physics (IPP), Canada.

Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia.

Also at Department of Physics, The University of Michigan, Ann Arbor, MI, United States.

Also at Ohio University, Athens, OH, United States.

Also at Louisiana Tech University, Ruston, LA, United States.

Also at Instituto Catalana de Recerca i Studis Avancats, ICREA, Barcelona, Spain.

Also at Graduate School of Science, Osaka University, Osaka, Japan.

Also at Department of Physics, National Tsing Hua University, Taiwan.

Also at Department of Physics, The University of Texas at Austin, Austin, TX, United States.

Also at Institute of Theoretical Physics, Ila State University, Tbilisi, Georgia.

Also at CERN, Geneva, Switzerland.

Also at Georgian Technical University (GTU), Tbilisi, Georgia.

Also at Ochanomizu University, Ochanomizu University, Tokyo, Japan.

Also at Manhattan College, New York, NY, United States.

Also at Hellenic Open University, Patras, Greece.

Also at Institute of Physics, Academia Sinica, Taipei, Taiwan.

Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan.

Also at School of Physics, Shandong University, Shandong, China.

Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia.

Also at Section de Physique, Université de Genève, Geneva, Switzerland.

Also at International School for Advanced Studies (SISSA), Trieste, Italy.

Also at Department of Physics and Astronomy, University of South Carolina, Columbia, SC, United States.

Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China.

Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria.

Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia.

Also at National Research Nuclear University MEPhI, Moscow, Russia.

Also at Department of Physics, Stanford University, Stanford, CA, United States.

Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary.

Also at Flinders University of Applied Sciences, Flensburg, Germany.

Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia.

Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France.

* Deceased.