Measurement of the cross section of high transverse momentum $Z \rightarrow b\bar{b}$ production in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector


Published in:
Physics Letters B

DOI:
10.1016/j.physletb.2014.09.020

Link to publication

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).
Measurement of the cross section of high transverse momentum \(Z \rightarrow b\bar{b}\) production in proton–proton collisions at \(\sqrt{s} = 8\) TeV with the ATLAS detector

**ATLAS Collaboration**

**Abstract**

This Letter reports the observation of a high transverse momentum \(Z \rightarrow b\bar{b}\) signal in proton–proton collisions at \(\sqrt{s} = 8\) TeV and the measurement of its production cross section. The data analysed were collected in 2012 with the ATLAS detector at the LHC and correspond to an integrated luminosity of 19.5 fb\(^{-1}\). The \(Z \rightarrow b\bar{b}\) decay is reconstructed from a pair of \(b\)-tagged jets, clustered with the anti-\(k_t\) jet algorithm with \(R = 0.4\), that have low angular separation and form a dijet with \(p_T > 200\) GeV. The signal yield is extracted from a fit to the dijet invariant mass distribution, with the dominant, multi-jet background mass shape estimated by employing a fully data-driven technique that reduces the dependence of the analysis on simulation. The fiducial cross section is determined to be

\[
\sigma_{Z \rightarrow b\bar{b}} = 2.02 \pm 0.20 \text{ (stat.)} \pm 0.25 \text{ (syst.)} \pm 0.06 \text{ (lumi.)} \text{ pb} = 2.02 \pm 0.33 \text{ pb},
\]

in good agreement with next-to-leading-order theoretical predictions.

Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/). Funded by SCOAP3.

1. Introduction

High transverse momentum \((p_T)\), hadronically decaying, electroweak-scale bosons have already been used in searches at the LHC [1–5], and are expected to play an increasingly significant role as the LHC moves to higher centre-of-mass energies in 2015. Therefore it is important to study them directly. This Letter presents the observation of a high-\(p_T\) \(Z \rightarrow b\bar{b}\) signal in a fully hadronic final state, and a measurement of its production cross section. The measurement is compared to the next-to-leading-order (NLO) matrix element plus parton-shower predictions of POWHEG [6–9] and aMC@NLO [10], where the parton-shower, hadronisation and underlying-event modelling are provided by PYTHIA-8.165 [11] and HERWIG++ [12], respectively. This first measurement of a high-\(p_T\) electroweak-scale boson in an all-hadronic final state at the LHC demonstrates the validity of both the analysis techniques used and of the state-of-the-art NLO plus parton-shower particle-level predictions for electroweak-scale bosons decaying to \(b\bar{b}\). It is therefore of great relevance for the search for the \(H \rightarrow b\bar{b}\) signal in the (most sensitive) high Higgs boson \(p_T\) range [13], as well as for searches for TeV-scale resonances decaying to \(b\bar{b}b\bar{b}\) via \(ZZ, ZH\) or \(HH\) [14,15]. A high-\(p_T\) \(Z \rightarrow b\bar{b}\) signal can also provide a useful benchmark for validating the performance of the ATLAS detector (for example, the \(b\)-jet energy scale\(^1\)); and for testing and optimising analysis methods relevant for physics studies involving high-\(p_T\) jets that contain \(b\)-hadrons (\(b\)-jets).

The analysis described in this Letter is designed to select \(b\bar{b}\) decays of \(Z\)-bosons with \(p_T > 200\) GeV, in proton–proton collisions at \(\sqrt{s} = 8\) TeV. The high-\(p_T\) requirement helps to enhance the signal relative to \(b\bar{b}\) production in multi-jet events (predominantly gluon splitting to \(b\bar{b}\) in this high-\(p_T\) regime), which is the dominant source of background and has a more steeply-falling \(p_T\) spectrum. In order to minimise the dependence on simulation, the analysis employs a fully data-driven technique for the determination of the invariant mass spectrum of the multi-jet background. This is especially important given that Monte Carlo (MC) generators have not been tested thoroughly in this region of the \(b\bar{b}\) production phase space.

2. The ATLAS detector

ATLAS is a multi-purpose particle physics experiment [17] at the LHC. The detector layout\(^2\) consists of inner tracking devices...
surrounded by a superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner tracking system provides charged-particle tracking in the pseudorapidity region \(\eta < 2.5\) and vertex reconstruction. It consists of a silicon pixel detector, a silicon microstrip tracker, and a strawtube transition radiation tracker. The system is surrounded by a solenoid that produces a 2 T axial magnetic field. The central calorimeter system consists of a liquid-argon electromagnetic (EM) sampling calorimeter with high granularity and an iron/scintillator-tile calorimeter providing hadronic energy measurements in the central pseudorapidity range \(\eta < 1.7\). The endcap and forward regions are instrumented with liquid-argon calorimeters for both electromagnetic and hadronic energy measurements up to \(|\eta| = 4.9\). The muon spectrometer is operated in a magnetic field provided by air-core superconducting toroids and includes tracking chambers for precise muon momentum measurements up to \(|\eta| = 2.7\) and trigger chambers covering the range \(\eta < 2.4\). A three-level trigger system is used to select interesting events [18]. The Level-1 trigger reduces the event rate to below 75 kHz using hardware-based trigger algorithms acting on a subset of detector information. Two software-based trigger levels, referred to collectively as the High-Level Trigger (HLT), further reduce the event rate to about 400 Hz using information from the entire detector.

3. Data, simulated samples, and event reconstruction

The data sample used in this analysis, after requiring that certain quality criteria be satisfied, corresponds to an integrated luminosity of \(\mathcal{L} = 19.5 \pm 0.5\) fb\(^{-1}\), and was recorded by ATLAS in 2012. The uncertainty on the integrated luminosity is derived, following the same methodology as that detailed in Ref. [19], from a calibration of the luminosity scale using beam-separation scans performed in November 2012.

MC event samples simulated with the GEANT4-based [20] ATLAS detector simulation [21] are used to model the \(Z \to b\bar{b}\) signal and the small \(t\bar{t}\), \(Z \to c\bar{c}\) and \(W \to q\bar{q}'\) background contributions. In addition, multi-jet MC event samples are used for studying the trigger modelling in simulation. The effect of multiple proton–proton interactions in the same bunch crossing (pile-up) is included in the simulation.

The \(Z \to b\bar{b}\) signal is modelled using SHERPA-1.4.3 [22], with the CT10 [23] LO parton distribution function (PDF) set. An alternative \(Z \to b\bar{b}\) model was generated with PYTHIA-8.165 [11] and the CTEQ6L1 [24] leading-order (LO) PDF set and is used to determine the systematic uncertainty associated with \(Z \to b\bar{b}\) modelling. The \(Z \to c\bar{c}\) background is also generated with SHERPA-1.4.3 and the CT10 PDF set. The \(t\bar{t}\) background is simulated with MC@NLO-4.06 [25] interfaced to HERWIG-6.520 [26] for the fragmentation and hadronisation processes, including JIMMY-4.31 [27] for the underlying-event description. The top quark mass is fixed at 172.5 GeV, and the PDF set CT10 is used. The \(W \to q\bar{q}'\) and multi-jet MC samples are generated using PYTHIA-8.165 with the CT10 PDF set.

Jets are reconstructed using the anti-\(k_t\) jet clustering algorithm [28], with radius parameter \(R = 0.4\). The inputs to the reconstruction algorithm are topological calorimeter cell clusters [29] calibrated at the EM energy scale. The effects of pile-up on jet energies are accounted for by a jet-area-based correction [30]. Jets are then calibrated to the hadronic energy scale using \(p_T\)- and \(\eta\)-dependent calibration factors based on MC simulations and the combination of several in situ techniques applied to data [29]. To remove jets with a significant contribution from pile-up interactions, it is required that at least 50% of the summed scalar \(p_T\) of tracks matched to a jet belongs to tracks originating from the primary vertex.\(^4\)

The flavour of jets in the ATLAS simulation is defined by matching jets to hadrons with \(p_T > 5\) GeV. A jet is labelled as a \(b\)-jet if a \(b\)-hadron is found within \(\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} = 0.3\) of the jet axis; otherwise, if a \(c\)-hadron is found within the same distance the jet is labelled as a \(c\)-jet; and if neither is the case then the jet is labelled as a light (quark) jet. The lifetime and other properties of \(b\)-hadrons are used to identify \([b\text{-tag}]\) \(b\)-jets with \(|\eta| < 2.5\), by exploiting the properties and topology of their decay products, such as the impact parameter of tracks (defined as a track’s distance of closest approach to the primary vertex in the transverse plane), the presence of displaced vertices, and the reconstruction of \(c\)-hadron and \(b\)-hadron decays. The \(b\)-tagging algorithm used in this analysis [31] combines the above information using multivariate techniques and is configured to achieve an efficiency of 70% for tagging \(b\)-jets in a MC sample of \(t\bar{t}\) events, while rejecting 80% of \(c\)-jets and more than 99% of light jets in the same sample.

4. Event selection

The events of interest in this analysis were triggered by a combination of six jet-based triggers. The most efficient of these triggers (accepting about 70% of the signal events) requires two jets identified as \(b\)-jets by a dedicated HLT \(b\)-tagging algorithm, with transverse energies \(E_T\) above 35 GeV, and a jet with \(E_T > 145\) GeV that may or may not be one of the \(b\)-tagged jets. The trigger efficiency for the SHERPA signal events passing the full offline event selection is 88.1%.

The event selection requires that there be at least three but no more than five jets with \(|\eta| < 2.5\) and \(p_T > 30\) GeV, and that exactly two of them be \(b\)-tagged. The \(b\)-tagged jets must each have \(p_T > 40\) GeV. The angular distance, \(\Delta R\), between them must be smaller than 1.2 and the transverse momentum of the dijet system they form, \(p_T^{\text{di jets}}\), must be greater than 200 GeV.

The final step of the event selection uses two variables with significant discrimination between signal and background to define two sets of events, one signal-enriched and the other signal-depleted, referred to hereafter as “Signal Region” and “Control Region”. The two variables, which are combined with an artificial neural network (ANN) into a single discriminant, \(S_{\text{NN}}\), are: (1) the dijet pseudorapidity, \(\eta_{\text{di jets}}\); and (2) the pseudorapidity difference, \(\Delta \eta\), between the dijet and the balancing jet, where the balancing jet is chosen to be the one that, when added to the dijet, gives the three-jet system with the smallest transverse momentum. The correlation of these two variables with the dijet invariant mass, \(m_{\text{di jets}}\), is very small, allowing the ANN to be trained using selected data events outside the mass window [80, 110] GeV as background and \(Z \to b\bar{b}\) MC events as signal. Fig. 1 depicts the distributions of \(\eta_{\text{di jets}}\), \(\Delta \eta\) and \(S_{\text{NN}}\) in the signal MC sample and in the data. The data shown here include all events with \(60 < m_{\text{di jets}} < 160\) GeV, and are representative of the background as the signal contribution is estimated to be only about 1%. The Signal Region is defined by \(S_{\text{NN}} > 0.58\) and the Control Region by \(S_{\text{NN}} < 0.45\). The discriminating power of \(\eta_{\text{di jets}}\) and \(\Delta \eta\) stems from the fact that signal production proceeds predominantly via a quark–gluon hard scatter, as opposed to the dominant multi-jet background which is largely initiated by gluon–gluon scattering. Due to the differences between

\(^4\) Amongst all reconstructed proton–proton collisions in a bunch crossing, the primary vertex is defined as the vertex with the highest summed track \(p_T\).
the gluon and quark PDFs, the $Z + \text{jet}$ system tends to be more boosted along the beam axis; hence the $Z$-boson is produced with higher $\eta$ and smaller $\Delta \eta$ separation from its recoil, compared to the background.

Since $S_{NN}$ is minimally correlated with $m_{\text{dijet}}$, the Control Region can be used as an unbiased model of the background in the Signal Region. Fig. 2 shows the normalised ratio of the $m_{\text{dijet}}$ distributions in the Signal and Control Regions, excluding the $Z$ mass window. A first-order polynomial fit to this distribution gives a good $\chi^2$ probability, 0.18, and a gradient consistent with zero, $(−1.37 \pm 1.10) \times 10^{-4}$ GeV$^{-1}$. In addition, the validity of assuming minimal correlation is supported by a test, performed with events from a PYTHIA 8 multi-jet MC sample satisfying the above analysis requirements, giving a ratio of the above distributions consistent with being flat. The impact of possible differences in the background $m_{\text{dijet}}$ shape between the Signal and Control Regions is considered as one of the systematic uncertainties on the measurement.

The total number of data events satisfying the full analysis selection is 236 712 in the Signal Region and 474 810 in the Control Region. The signal-to-background ratio in a 30 GeV window around the $Z$-boson mass is expected to be about 6% (2%) in the Signal (Control) Region. The $t\bar{t}$ events are estimated to represent about 0.5% of the total background in both the Signal and Control Regions, and the $Z \to c\bar{c}$ and $W \to q\bar{q}$ backgrounds are approximately 8% and 6% of the signal, respectively.

5. Cross-section definition

The fiducial cross section of resonant $Z$-boson production, with $Z$ decaying to $b\bar{b}$, $\sigma_{Z\to b\bar{b}}^{\text{fid}}$, is defined as follows. Particle-level jets in MC $Z \to b\bar{b}$ events are reconstructed from stable particles (particles with lifetime in excess of 10 ps, excluding muons and neutrinos) using the anti-$k_t$ algorithm with radius parameter $R = 0.4$. There must be two particle-level $b$-jets in the event that satisfy the following fiducial conditions: $p_T > 40$ GeV, $|\eta| < 2.5$ for the individual jets; and $\Delta R(\text{jet1, jet2}) < 1.2$, $p_T^{\text{dijet}} > 200$ GeV, $60 < m_{\text{dijet}} < 160$ GeV for the dijet system. The cross section is extracted from the measured yield of $Z \to b\bar{b}$ events in the data, $N_{Z\to b\bar{b}}$, as

$$
\sigma_{Z\to b\bar{b}}^{\text{fid}} = \frac{N_{Z\to b\bar{b}}}{C_{Z\to b\bar{b}}},
$$

where $C_{Z\to b\bar{b}}$ is the efficiency correction factor to correct the detector-level $Z \to b\bar{b}$ yield to the particle level. The value of $C_{Z\to b\bar{b}}$ in the SHERPA MC signal is found to be 16.2%, which can be factorised into the product of: trigger efficiency (88.1%), $b$-tagging and kinematic selection efficiency (52.7%), and the efficiency of the $S_{NN}$ requirement that defines the Signal Region (35.0%). The uncertainties on $C_{Z\to b\bar{b}}$ are discussed in Section 7.

6. Signal extraction

The signal yield is extracted by fitting simultaneously the $m_{\text{dijet}}$ distributions of the Signal and Control Regions in the range [60, 160] GeV with a binned, extended maximum-likelihood (EML) fit, using a bin width of 1 GeV.
The $Z \to b\bar{b}$ signal shape is modelled in the EML fit as the sum of three Gaussians. This empirical model describes well both the SHERPA and the PYTHIA signal MC samples, albeit with slightly different parameters. Given this, the SHERPA-based model is used as the baseline for the fit, and the PYTHIA-based model is used for an estimate of the systematic uncertainty on the measurement due to the signal shape. The only free parameters of the signal model in the EML fit are the yield in the Signal Region and the shift, $\delta M_Z$, of the mean of the narrowest Gaussian from its MC-predicted value. The widths and relative contributions of the three Gaussians, as well as the differences between the means of each of the two wider Gaussians and the narrowest one, are fixed to the values determined by a separate fit to the signal MC $m_{\text{dijet}}$ distributions. Given that the Control Region is not signal-free, the simultaneous fit includes a signal component in both the Signal Region and the Control Region. The relative proportion of signal in the two regions, $R_Z = (N_{\text{Control}})/(N_{\text{Signal}})$, is fixed to the value predicted by SHERPA, $R_Z = 0.62$. This choice is supported by the good agreement found between SHERPA and data in the $S_{NN}$ distribution obtained from a pure sample of high-$p_T$ $Z \to \mu^+\mu^-$ events, as discussed in Section 7.

The dominant multi-jet background is modelled in the EML fit using a seventh-order Bernstein polynomial [32]. This is purely an empirical model and the order of the polynomial is chosen by a $\chi^2$ probability saturation test by fitting the $m_{\text{dijet}}$ distribution in the Control Region with the background-only hypothesis and an increasing polynomial order, until no improvement is seen in the fit quality when adding higher-order terms. The coefficients of the Bernstein polynomial are determined by the simultaneous EML fit and are identical for the Signal and Control Regions. In this way, the signal-depleted Control Region constrains the background prediction in the Signal Region. The only additional parameters of the fit are the background normalisations in the Signal and Control Regions.

The small $Z \to c\bar{c}$, $t\bar{t}$ and $W \to q\bar{q}'$ backgrounds are included as separate components in the EML fit for both the Signal and Control Regions, with their $m_{\text{dijet}}$ shapes being parameterised from MC simulation as follows. The $Z \to c\bar{c}$ and $W \to q\bar{q}'$ components are each modelled as three-Gaussian sums like the signal, with all parameters fixed to values from fits to MC simulation. The means of the Gaussians are expressed with respect to the mean of the narrowest $Z \to b\bar{b}$ Gaussian: this couples the position of these backgrounds to the $Z \to b\bar{b}$ signal. The $W \to q\bar{q}'$ component is normalised absolutely to its PYTHIA LO cross section, corrected to NLO by a $K$-factor derived using MCFM [33]. The acceptance of the $Z \to c\bar{c}$ background is taken from the simulation, but its yield is linked to the fitted $Z \to b\bar{b}$ yield, since the $Z \to c\bar{c}$ production differs from the signal only in the well-known branching fractions of the $Z$ decays. All properties of the $t\bar{t}$ component are fixed using the $t\bar{t}$ simulation, with normalisation from the NNLO+NNLL prediction of the $t\bar{t}$ production cross section [34–39]. The contribution from Higgs decays to $b\bar{b}$ is expected to be $\sim 3\%$ of the $Z \to b\bar{b}$ signal and localised away from the signal peak: therefore no such component is included in the EML fit.

The fit procedure has been validated using a comprehensive set of tests based on pseudo-experiments, which have demonstrated that the yield and its uncertainty are accurately determined by the fit procedure over a wide range of input signal yields. In particular, the fit procedure is robust against fitting artefacts like false dips or peaks: a consequence of fitting both signal and control regions simultaneously, with the ratio of $Z \to b\bar{b}$ in each region fixed.

Fig. 3 shows the result of the simultaneous fit to the $m_{\text{dijet}}$ distributions of the Signal and Control Regions, as well as the corresponding background-subtracted data distributions. The rather
complex shape of the background invariant mass distribution results from the use of the six jet-based triggers, all of which have different jet $p_T$ thresholds and hence shape differently the invariant mass distributions. The fitted function models the data well, with a signal peak compatible with $Z \rightarrow bb$ decays. The fitted signal yield is $6420 \pm 640$ (stat.) events.

### 7. Systematic uncertainties

The sources of systematic uncertainties considered in this analysis, which may affect the fitted signal yield, the efficiency correction factor or both, are listed in Table 1.

The jet energy scale (JES) and jet energy resolution (JER) uncertainties are determined using the techniques described in Refs. [29, 40]. The JES uncertainty has a relatively large impact on the signal efficiency, due to the $p_T$ requirements on the individual jets and the dijet system, but a comparatively small impact on the fitted yield, because of the data-driven approach for the background determination and the fact that the location of the signal peak is a free parameter of the EML fit. The JER uncertainty affects predominantly the fitted yield, since it modifies the MC-derived signal shape.

The $b$-tagging efficiency in the simulation is scaled to reproduce the one in data and its uncertainty is evaluated by varying the data-to-MC scale factor applied to each jet in the simulation within a range that reflects the systematic uncertainty on the measured tagging efficiency for $b$-jets in ATLAS [31–41]. The $Z \rightarrow cc$ relative normalisation uncertainty is estimated in a similar way by varying the corresponding scale factors for charm jets in the simulation.

The uncertainty on $C_{Z \rightarrow bb}$ due to a potential mis-modelling of the trigger efficiency is assessed using data events collected with a prescaled trigger that is fully efficient with respect to the analysis event selection. The full offline event selection is applied to these events and the efficiency for passing the analysis trigger requirements is compared to the corresponding efficiency in the multi-jet MC sample, as a function of various kinematic variables. It is found that the two trigger efficiencies are consistent to within 6%. Furthermore, the trigger efficiency in the multi-jet MC sample, when considering only those events where the two $b$-tagged jets are labelled as true $b$-jets, is fully consistent with the trigger efficiency in the signal MC events. Based on these studies, a $\pm 6\%$ trigger efficiency modelling uncertainty is propagated to the cross-section measurement.

The uncertainty on the extracted signal yield due to potential differences in the background $m_{\text{dijet}}$ shape between the Signal and Control Regions (“Control Region bias”) is assessed by repeating the EML fit for a range of $S_{\text{NN}}$ values around the one used in the baseline selection to define the Control Region. These variations of the Control Region definition lead to small biases in the $m_{\text{dijet}}$ shape relative to the Signal Region, resulting in non-zero slopes in the first-order polynomial fits to the distributions equivalent to the one in Fig. 2. The non-zero slopes of these fits bracket the statistical uncertainty with which the slope of the first-order polynomial fit to Fig. 2 is determined. The largest upwards and downwards variations in the fitted signal yield from the EML fits following this procedure are propagated as systematic uncertainty to the cross-section measurement.

The impact on $C_{Z \rightarrow bb}$ of a possible mis-modelling of the distributions of the analysis selection variables, except $S_{\text{NN}}$, in the MC signal is assessed by comparing the SHERPA and PYTHIA MC signal samples. It is found to be less than 1% and therefore it is considered negligible. There is a 15% discrepancy between the PYTHIA and SHERPA predictions for the efficiency of the Signal Region $S_{\text{NN}}$ requirement. Since the input variables to $S_{\text{NN}}$ depend primarily on the dynamics of $Z$-boson production, the modelling by SHERPA is tested by comparing a sample of events in the ATLAS 2012 data containing high-$p_T$ $Z \rightarrow \mu^+\mu^-$ decays to a corresponding SHERPA MC sample, with the dimuon system replacing the dijet system. The agreement is found to be very good, at the level of 2%, and the residual discrepancies are propagated as the “Signal $S_{\text{NN}}$ modelling” uncertainty on both $C_{Z \rightarrow bb}$ and the $R_Z$ fit parameter. This uncertainty also covers the impact from possible differences between the PDFs used in the SHERPA signal sample and the data, given that the above $Z \rightarrow \mu^+\mu^-$ SHERPA sample uses the same PDF set as the $Z \rightarrow bb$ SHERPA signal sample.

The difference obtained in the fitted signal yield when using the PYTHIA signal model rather than the SHERPA one is used as an estimate of the uncertainty on the measurement due to possible mis-modelling of the $m_{\text{dijet}}$ Shape in the MC signal.

The impact on the measurement from the uncertainty on the $W \rightarrow q\bar{q}$ and $t\bar{t}$ normalisations is assessed by varying the fixed number of events of each background in the Signal and Control Regions independently by 50% and repeating the EML fit.

### 8. Results

Using the extracted $Z \rightarrow bb$ yield, the estimated signal-efficiency correction factor and the integrated luminosity of the dataset, the cross section in the fiducial region defined in Section 5 is measured to be

$$\sigma_{Z \rightarrow bb}^{\text{fid}} = 2.02 \pm 0.20 \text{ (stat.)} \pm 0.25 \text{ (syst.)} \pm 0.06 \text{ (lumi.)} \text{ pb}.$$  

The total systematic uncertainty is the result of adding in quadrature all the individual systematic uncertainties on $\sigma_{Z \rightarrow bb}^{\text{fid}}$, listed in Table 1. It is further found that the signal $m_{\text{dijet}}$ peak position is consistent with the $Z \rightarrow bb$ expectation: $4M_Z = -1.5 \pm 0.7 \text{ (stat.)} \pm 2.5 \text{ (syst.)} \text{ GeV}$. The good agreement with zero provides an independent confirmation of the good agreement between data and MC on the energy scale of b-jets in ATLAS.
The robustness of the measurement is supported by several cross-checks and complementary studies. In particular, a consistent cross-section measurement is obtained by applying a tighter $b$-tagging selection (with an efficiency of 60% for tagging $b$-jets in a MC sample of $\ell\ell$ events) or when the requirement on $p_T^{\text{jet}}$ is raised to 250 GeV or 300 GeV. Furthermore, when the same methodology is repeated on two independent classes of events, those accepted by the dominant trigger described above and all other events, both measured cross sections (1.99 ± 0.25 (stat.) pb and 1.87 ± 0.44 (stat.) pb, respectively) are fully consistent with the baseline measurement, even though the $m_{\text{dijet}}$ distributions are significantly different in the two classes of events. In addition, when the background shape obtained from the baseline EML fit is used to fit for a signal in the sample of events with $0.45 < S_{NN} < 0.58$, the fitted signal yield in this sample is consistent with the number of signal events calculated based on the measured cross section and the shape of $S_{NN}$ predicted by Sherpa.

Finally, repeating the analysis with a number of alternative functional forms for the empirical description of the background shape (such as a log-normal function convolved with a fourth-order Bernstein polynomial) leads to negligible variations in the measured cross section compared to the systematic uncertainties of the measurement.

The measured cross section is compared to the particle-level, NLO-plus-parton-shower predictions of two MC generators, POWHEG and aMC@NLO, in the same fiducial region. In both cases, the cross section of the $Z + 1$-jet process is calculated to NLO accuracy. For aMC@NLO, the $Z$ decay is simulated with MadSpin [42]. POWHEG is interfaced to PyTHIA for parton showering, hadronisation and underlying-event contributions, whilst aMC@NLO is interfaced to HERWIG++. The particle-level predictions are then derived by applying to the generated events the fiducial selection defined in Section 5. The predicted cross sections are:

POWHEG: $\sigma_{Z^* \to b\bar{b}} = 2.02^{+0.25}_{-0.19}\,(\text{scale})^{+0.03}_{-0.04}\,(\text{PDF})\,\text{pb}$

aMC@NLO: $\sigma_{Z^* \to b\bar{b}} = 1.98^{+0.16}_{-0.08}\,(\text{scale})^{+0.03}_{-0.04}\,(\text{PDF})\,\text{pb}$.

Both generators use the CT10 PDF set for the central value of the prediction, and the renormalisation and factorisation scales are set to the $p_T$ of the $Z$ boson. The uncertainty due to the ambiguity in the renormalisation and factorisation scales is estimated by doubling or halving them simultaneously. The PDF uncertainty is evaluated by varying the 52 PDFs in the CT10 NLO error set, following the Hessian method and rescaling to the 68% confidence level. Within the experimental and theoretical uncertainties, both predictions are completely consistent with the measured cross section.

POWHEG and aMC@NLO can also be used to provide an estimate of the fraction of the total cross section for $Z \to b\bar{b}$ production at the LHC, with $p_T > 200$ GeV, that is contained within the measured fiducial region. The ratio of the above cross sections to the cross sections calculated without applying any particle-level requirements, only requiring $p_T > 200$ GeV for the $Z$-boson before parton showering, is 0.53 for POWHEG and 0.47 for aMC@NLO.

In conclusion, the high-$p_T$ $Z \to b\bar{b}$ signal has been observed and its production cross section measured in a fully hadronic final state, in 19.5 fb$^{-1}$ of proton-proton collisions at $\sqrt{s} = 8$ TeV recorded in 2012 by the ATLAS detector at the LHC. Within the fiducial region defined in Section 5, the production cross section is measured to be

$$\sigma_{Z^* \to b\bar{b}} = 2.02^{+0.20}_{-0.20}\,(\text{stat.})^{+0.05}_{-0.06}\,(\text{lumi.})\,\text{pb}$$

and is found to be in good agreement with the NLO-plus-parton-shower predictions from POWHEG and aMC@NLO.

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; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MOSMT, CR, VSC CR and Czech Republic; DNRF, DNSRC and Lundbeck Foundation, Denmark; EPLANET, ERC and NSRF, European Union; IN2P3–CNRS, CEA-DSM/JIFRJ, France; GNSF, Georgia; BMBF, DFG, MPG and AvH Foundation, Germany; GSKT and NSRF, Greece; ISF, MINEVA, GIF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; BRF and RCN, Norway; MNiSW and NCN, Poland; GRICES and FCT, Portugal; MINE/IFA, Romania; MES of Russia and ROSATOM, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MIŽS, Slovenia; DST/NRF, South Africa; MINECO, 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.

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


ATLAS Collaboration
