First observation of forward $Zbb^-\bar{b}$ production in pp collisions at $s=8$ TeV


DOI: 10.1016/j.physletb.2017.11.066

License: Creative Commons: Attribution (CC BY)

Citation for published version (Harvard):
https://doi.org/10.1016/j.physletb.2017.11.066

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Published in Physics Letters B on 05/12/2017
DOI: 10.1016/j.physletb.2017.11.066

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• User may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (??)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document. When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 31. Jul. 2019
First observation of forward $Z \to b\bar{b}$ production in $pp$ collisions at $\sqrt{s} = 8$ TeV

LHCb Collaboration

Abstract

The decay $Z \to b\bar{b}$ is reconstructed in $pp$ collision data, corresponding to 2 fb$^{-1}$ of integrated luminosity, collected by the LHCb experiment at a centre-of-mass energy of $\sqrt{s} = 8$ TeV. The product of the $Z$ production cross-section and the $Z \to b\bar{b}$ branching fraction is measured for candidates in the fiducial region defined by two particle-level $b$-quark jets with pseudorapidities in the range $2.2 < \eta < 4.2$, with transverse momenta $p_T > 20$ GeV and dijet invariant mass in the range $45 < m_{jj} < 165$ GeV. From a signal yield of $5462 \pm 763$ $Z \to b\bar{b}$ events, where the uncertainty is statistical, a production cross-section times branching fraction of $332 \pm 46 \pm 59$ pb is obtained, where the first uncertainty is statistical and the second systematic. The measured significance of the signal yield is 6.0 standard deviations. This measurement represents the first observation of the $Z \to b\bar{b}$ production in the forward region of $pp$ collisions.

© 2017 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/). Funded by SCOAP3.

1. Introduction

Measurements of $Z$-boson production in $pp$ collisions constitute an important test of the Standard Model (SM), since they allow the electroweak sector to be precisely probed [1-3]. The LHCb experiment can be used to measure the decay of the $Z$ boson into a $b\bar{b}$ quark pair in the forward region that is inaccessible at other LHC experiments.

The decay $Z \to b\bar{b}$ provides a standard candle for searches in final states with a $b\bar{b}$ quark pair. The inclusive search for the SM Higgs decay to two $b$ quarks at the LHC is of great interest, since the measurement of the Higgs boson coupling to $b$ quarks is an important test of the SM [4]. Several extensions of the SM predict that new heavy particles that decay to two energetic $b$ quarks could be accessible at LHC collision energies [5-7]. A sizeable $Z \to b\bar{b}$ event sample will enable the measurement of the $b\bar{b}$ forward-central asymmetry at the $Z$ pole, which could be enhanced by the contributions from new physics processes [8]. The forward-central asymmetry in inclusive $b\bar{b}$ events has previously been measured by the LHCb collaboration [9].

The measurements of this decay can also be used to demonstrate that no biases are induced by the $b$-jet reconstruction procedure and that the reconstruction efficiencies are evaluated correctly. In addition, the $Z \to b\bar{b}$ decay is important to determine the so-called $b$-jet energy scale. This is the factor that has to be applied to the reconstructed $b$-jet energy in simulated events in order to reproduce the actual detector response.

The reconstruction of the $Z \to b\bar{b}$ decay is challenging at hadron colliders, due to the large QCD background. Many techniques to reconstruct the $Z \to b\bar{b}$ decay channel have been developed by the CDF [10], ATLAS [11] and CMS [12] collaborations. The CDF collaboration reconstructed the $Z \to b\bar{b}$ decay in $p\bar{p}$ collisions at 1.96 TeV and determined the $b$-jet energy scale, obtaining a relative uncertainty on the product of the cross-section and the branching fraction of 29%. The analysis of the ATLAS collaboration reconstructed boosted $Z \to b\bar{b}$ candidates in the central region of $pp$ collisions at 8 TeV, with pseudorapidity $|\eta| < 2.5$, and determined the cross-section with a relative uncertainty of 16%. The CMS collaboration made the first observation of the $Z \to b\bar{b}$ decay in a single-jet topology in the same pseudorapidity region, with a significance of 5.1 standard deviations.

This Letter describes a new method to study the $Z \to b\bar{b}$ decay, performed on $pp$ collision data collected at a centre-of-mass energy of $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of 2 fb$^{-1}$. The low trigger thresholds on the particle energies that are employed at LHCb and the excellent $b$-jet identification performance make it possible to select candidates within a large invariant mass range, including those with masses below the $Z$-boson pole. Events are selected requiring two $b$-jet candidates, referred to as a $b$ dijet, and an additional jet that balances the transverse momentum of the $b\bar{b}$ system. The invariant mass distribution of the $b$ dijet is used to determine the $Z \to b\bar{b}$ yield and the $b$-jet energy scale. The invariant mass distribution of the QCD background is determined using a control region that is defined through observables related to the $b$-dijet system and to the associated balancing...
jet. Simulated data are used to evaluate the reconstruction efficiency and the detector acceptance, enabling a measurement of the Z production cross-section multiplied by the $Z \rightarrow bb$ branching fraction.

2. The LHCb detector, trigger and simulation

The LHCb detector [13,14] is a single-arm forward spectrometer fully instrumented in the pseudorapidity range $2 < \eta < 5$, which is designed for the study of $b$ and $c$ hadrons. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the $pp$ interaction region, a silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of momentum, $p$, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV. The minimum distance of a track to a primary vertex, the impact parameter, is measured with a resolution of $(15 + 29/p_T)$ μm, where $p_T$ is the component of the momentum transverse to the beam, in GeV. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad (SPD) and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers. The online event selection is performed by a trigger system, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction.

Events are required to satisfy at least one of the following hardware trigger requirements: contain a muon with $p_T > 1.86$ GeV, a hadron with transverse energy in the calorimeters $E_T > 3.7$ GeV, an electron with $E_T > 3$ GeV, a photon with $E_T > 3$ GeV or a pair of muons with $p_{T1} \cdot p_{T2} > 1.6$ GeV$^2$. A global event cut (GEC) on the number of hits in the SPD is applied in order to prevent high-multiplicity events from dominating the processing time. At the software trigger stage events are required to have a two-, three- or four-track secondary vertex (SV) with significant displacement from any primary vertex. A multivariate algorithm [15] is used for the identification of secondary vertices consistent with the decay of a $b$ hadron, strongly suppressing the contamination from charmed hadrons.

Simulated events generated with PYTHIA [16], with a specific LHCb configuration [17], are used to model the properties of the signal $Z \rightarrow bb$ events and backgrounds such as $Z \rightarrow cc$, $W \rightarrow qq'$ decays and $t\bar{t}$ events. Decays of hadronized particles are described by EvtGen [18], where the final-state radiation is generated using PhoRors [19]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [20] as described in Ref. [21].

3. Candidate selection

Candidates are selected by requiring the presence of at least three jets, which are reconstructed as detailed in Refs. [22–25]. Jets are reconstructed using a particle flow algorithm [25] and are clustered with the anti-$k_T$ algorithm [26] with a distance parameter 0.5, as implemented in the FASTJET software package [27]. A jet energy correction [25] determined from simulation is applied to recover the jet energy at particle level and jet quality requirements are applied [25]. Jets are heavy-flavour tagged, i.e. as containing a $b$ or $c$ hadron, if a SV is found with a distance $\Delta R < 0.5$ from the jet axis, where $\Delta R$ is the distance in the $(\eta, \phi)$ plane and $\phi$ is the azimuthal angle between the jet axis and the vector that points from the $pp$ interaction point to the SV. The details of the flavour-tagging algorithm are described in Ref. [28]. Two heavy-flavour tagged jets are required to form a $Z \rightarrow bb$ candidate. At least one of the two $b$-jet candidates must be tagged by a SV selected by the software trigger requirements. The two heavy-flavour jets are each required to have transverse momenta $p_T > 20$ GeV, pseudorapidities in the range $2.2 < \eta < 4.2$, and a combined invariant mass $(m_{jj})$ in the range $45 < m_{jj} < 165$ GeV. The fiducial region of the measurement within which the cross-section is determined is defined by the kinematical requirements described above applied to particle-level jets, which are jets reconstructed in the simulation from stable particles (i.e. particles with lifetime in excess of 10 ps, excluding neutrinos) using the default reconstruction algorithm.

In order to increase the signal-to-background ratio, the absolute azimuthal angle between the two $b$-jets is required to be greater than 2.5 radians. The presence of a balancing jet is required to help discriminate $Z \rightarrow bb$ events from the QCD multijet background. The $Z + jet$ signal is predominantly produced via quark–gluon scattering, while the QCD multijet background is produced via gluon–gluon interactions [29]. The balancing jet is defined as that which minimises the total $p_T$ of the Z boson and the jet. This jet is required to have $p_T > 10$ GeV and $2.2 < \eta < 4.2$. Given the SM cross-sections [30] and the selection efficiencies, which are evaluated using simulation, about $17 \times 10^3$ $Z \rightarrow bb$ candidates, 600 $Z \rightarrow cc$ candidates, 200 $W \rightarrow qq'$ candidates and 50 $t\bar{t}$ candidates are expected after the application of the selection criteria. A sample of around $6 \times 10^5$ candidates is selected in data, dominated by the combinatorial background from the multijet QCD events.

A multivariate classifier is trained to discriminate $Z \rightarrow bb$ events from combinatorial QCD events. A uniform Gradient Boost Boosted Decision Tree technique [31] is adopted, in order to ensure a selection efficiency with a low dependence on the dijet invariant mass. The classifier is trained using four kinematical variables of the three-jet system, chosen for both their low correlation with the dijet invariant mass and for their discriminating power. The variables are the absolute pseudorapidity difference between the two heavy-flavour jets, the $p_T$ of the balancing jet, the angle between the balancing-jet momentum and the $Z$-boson candidate momentum in the azimuthal plane with respect to the beam axis, and the polar angle between the balancing-jet momentum and the $Z$-boson flight direction in the $Z$-boson rest frame. The classifier is trained using 5% of the data sample to represent the combinatorial QCD background. This training sample has a negligible $Z \rightarrow bb$, $Z \rightarrow cc$, $W \rightarrow qq'$ and $t\bar{t}$ contamination and it is not used in the dijet invariant mass fit described below. The signal process is modelled using simulated $Z \rightarrow bb$ events. The distributions of the input observables related to the balancing-jet kinematics are validated by comparing the high purity $Z(\rightarrow \mu^+ \mu^-) + jet$ data sample described in Ref. [24] with the corresponding simulation sample.

The output of the classifier (uGB) is shown in Fig. 1. Candidates are selected in two different regions of uGB: the signal region ($uGB > x_s$), which has enhanced $Z \rightarrow bb$ contribution, and a control region ($uGB < x_c$), which has a larger contribution from QCD combinatorial events. The two regions are fitted simultaneously to determine the $Z \rightarrow bb$ yield, and the values of $x_s$ and $x_c$ are chosen in order to achieve the best signal significance.

---

[1] In this Letter natural units where $\hbar = c = 1$ are used.
4. Signal yield determination

A simultaneous fit to the b-dijet invariant mass distributions in the signal and control regions is performed to determine the $Z \rightarrow bb$ yield and the jet energy scale factor, $k_{\text{ES}}$. A triple-Gaussian model is used to describe the $Z \rightarrow bb$ dijet invariant mass distribution. The parameters of this model are obtained separately for the candidates in the signal and control regions using simulation, and are fixed in the fit to the data. The $k_{\text{ES}}$ factor is also introduced in the $Z \rightarrow bb$ invariant mass distribution model in order to account for differences between simulation and data in the jet four-momentum. This is achieved by substituting $m_{jj}$ with $m_{jj}/k_{\text{ES}}$ in the model. The reconstructed invariant mass of dijets in $Z \rightarrow bb$ simulated events has a mean of 80 GeV, i.e. below the known $Z$-boson mass [30], and a resolution of 16%. The reduced mean is due to parton radiation outside the jet cone, missing energy, and residual biases in the reconstructed jet energy that are not recovered by the jet energy correction.

The invariant mass distribution of the combinatorial background is parametrized with a Pearson IV distribution, as is typical to describe the multijet combinatorial background [10]. The four parameters of the Pearson IV function are free to vary in the fit and they have approximately the same values in the signal and control regions, since the uGB is trained to be as uniform as possible with respect to the dijet invariant mass. To take into account the residual correlation with the dijet invariant mass, the Pearson IV distribution is multiplied in the signal (control) region by a linear transfer function $t^{s(c)}(m_{jj})$, defined as

$$t^{s(c)}(m_{jj}) = a^{s(c)} + b^{s(c)} \cdot m_{jj},$$

where the superscript $s$ ($c$) indicates the signal (control) region, and $a^{s(c)}$ and $b^{s(c)}$ are parameters fixed in the invariant mass fit. The parameters $a^{s(c)}$ and $b^{s(c)}$ are determined by fitting the transfer function to the selection efficiency after the requirement that $uGB > x_c$ ($uGB < x_c$) as a function of the dijet invariant mass in the 45 < $m_{jj}$ < 60 GeV and 100 < $m_{jj}$ < 165 GeV intervals, where the $Z \rightarrow bb$ contribution is negligible. As a cross-check, data events with $uGB < x_c$ are fitted with only the QCD background model, ignoring the small $Z \rightarrow bb$ contribution, and a good fit quality is obtained.

The invariant mass model used to fit the signal region is

$$f^s(m_{jj}) = N_Q^s \cdot Q(m_{jj}) \cdot t^s(m_{jj}) + N_Z^s \cdot Z^s(m_{jj}; k_{\text{ES}}),$$

where $N_Q^s$ and $N_Z^s$ are the number of QCD events and the number of $Z$-boson events ($Z \rightarrow bb$ plus $Z \rightarrow c\bar{c}$) in the signal region respectively, and $Q(m_{jj})$, $t^s(m_{jj})$ and $Z^s(m_{jj})$ are the Pearson IV distribution, the transfer function and the $Z$-boson invariant mass distribution model in the signal region, respectively. The $Z \rightarrow c\bar{c}$ invariant mass distribution is assumed to be identical to that of $Z \rightarrow bb$ events. This assumption is verified using the simulation and the two components are therefore fitted together. Backgrounds other than $Z \rightarrow c\bar{c}$ and QCD multijet events are neglected in the fit. Since the $uGB > x_s$ requirement is applied, the expected value of $N_c^s$ is lower than the $17 \times 10^3 Z \rightarrow bb$ events expected before the uGB selection.

The invariant mass model that describes the control region is

$$f^c(m_{jj}) = N_G^c \cdot Q(m_{jj}) \cdot t^c(m_{jj}) + R \cdot N_Z^c \cdot Z^c(m_{jj}; k_{\text{ES}}),$$

where $N_G^c$ is the number of QCD events in the control region and $Q(m_{jj})$, $t^c(m_{jj})$ and $Z^c(m_{jj})$ are the Pearson IV distribution, the transfer function and the $Z$-boson invariant mass distribution model in the control region. The parameter $R$ is the ratio of the efficiency for $Z$-boson candidates selected with $uGB < x_c$ and $uGB > x_s$ and is determined from simulation and fixed in the fit. A simultaneous unbinned maximum likelihood fit is performed with $N_Q^c$, $N_G^c$, $N_Z^c$, $k_{\text{ES}}$ and the Pearson IV parameters free to vary. Pseudoexperiments are used to verify that the fit is stable and estimate any bias. The parameter $N_Q^c$ is determined with a bias of about 2% and the value returned by the fit is corrected accordingly in the cross-section determination.

The fit result is shown in Fig. 2 and the background-subtracted data and result of the fit are shown in Fig. 3. The $Z$-boson yield in the signal region is $5462 \pm 763$ and the jet energy scale factor is measured to be $1.009 \pm 0.015$. Using Wilks’ theorem [32], the $Z \rightarrow bb$ statistical significance is found to be 7.3 standard deviations.

As an additional cross-check to validate the technique, a fit to the dijet invariant mass distribution for candidates with $x_c < uGB < x_s$ is performed, with a model analogous to that used in the signal and control regions. In this case, the parameters of the QCD background are fixed to the values returned by the default fit, but the $Z \rightarrow bb$ yield in this region, $N_c^c$, is left free. The goodness of this fit is acceptable and the ratio $N_c^c/N_c^s$ is compatible with the expectation from simulation.

5. Cross-section determination and systematic uncertainties

The product of the $Z$-boson production cross-section and the $Z \rightarrow bb$ branching fraction is determined using

$$\sigma(pp \rightarrow Z)B(Z \rightarrow \bar{b}b) = \frac{N_c^c}{L \cdot (1 - f_{uGB}) \cdot \epsilon^c \cdot (1 + f_{Z \rightarrow c\bar{c}})},$$

where $L$ is the integrated luminosity, $\epsilon^c$ is the efficiency of the selection requirements, including $uGB > x_s$, for events in the fiducial region, $f_{uGB}$ is the fraction (5%) of data events removed for the multivariate classifier training and $1 + f_{Z \rightarrow c\bar{c}}$ is a factor applied to correct for the small $Z \rightarrow c\bar{c}$ contamination. The selection efficiency is obtained from simulation, but correction factors are applied to account for differences in the heavy-flavour tagging efficiencies between data and simulation [28]. By using a small sample with a looser trigger requirement and a technique similar to that described in Ref. [24], the GEC efficiency is also corrected for differences in data and simulation. The balancing-jet selection efficiency is corrected at Next-to-Leading-Order (NLO) using simulated $Z \rightarrow bb$ events produced with amC@NLO [33] plus Pythia for parton showers. The $f_{Z \rightarrow c\bar{c}}$ fraction is obtained by multiplying the $Z \rightarrow c\bar{c}$ and $Z \rightarrow bb$ branching fraction ratio [30] by the acceptance and the efficiency ratios, both determined using simulation.

The sources of systematic uncertainty considered for the measurement are given in Table 1. Systematic effects that are asso-
associated with differences between data and simulation can affect the signal invariant mass distribution model and the selection efficiency. The impact of these differences is evaluated by repeating the fit with a modified signal model and by recalculating the cross-section varying $\sigma_Z$. Other sources of systematic uncertainties are related to the signal extraction procedure.

The method described in Ref. [28] is used to assess the systematic uncertainty due to the heavy-flavour tagging efficiency which amounts to 5%–10% per jet, depending on the $p_T$ range. This uncertainty is dominated by the size of the calibration samples used in the heavy-flavour tagging efficiency measurement. Since one of the two $b$-jet candidates must be tagged by a SV selected by the software trigger, the uncertainty on this trigger efficiency is included in this contribution. The systematic uncertainty associated with the hardware trigger efficiency is determined by measuring the efficiency with a tag-and-probe technique, using the high purity $Z(\rightarrow \mu^+\mu^-)+\text{jet data sample}$ [24]. In order to avoid trigger bias on the jet selection, the tag is the muon that triggered the event and the probe is the associated jet. The hardware trigger efficiency measured on probe jets is compared between data and simulation and the maximum difference in intervals of the jet $p_T$ is taken as an uncertainty. The latter does not take into account the systematic uncertainty on the GEC efficiency, which is determined separately by studying its dependence on the $b$-dijet invariant mass and assigning the largest variation as the uncertainty. The systematic uncertainty on jet energy correction includes biases due to jet flavour dependence, reconstruction of tracks which are not associated to a real particle, the track momentum resolution and residual differences between simulation and data, as described in Refs. [24, 25]. The jet energy resolution is modelled in simulation with an uncertainty measured in Refs. [23, 25]. The uncertainties related to the jet reconstruction and identification are taken from Ref. [25]. The systematic uncertainty associated with the balancing-jet se-

<table>
<thead>
<tr>
<th>Systematic source</th>
<th>$\sigma_Z$ [%]</th>
<th>$\Delta\mu_5$ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy-flavour tagging efficiency</td>
<td>16.6</td>
<td>0.5</td>
</tr>
<tr>
<td>Hardware trigger efficiency</td>
<td>1.9</td>
<td>–</td>
</tr>
<tr>
<td>GEC efficiency</td>
<td>1.7</td>
<td>–</td>
</tr>
<tr>
<td>Jet energy correction</td>
<td>2.7</td>
<td>0.3</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>1.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Jet identification efficiency</td>
<td>2.0</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>Balancing-jet selection efficiency</td>
<td>1.8</td>
<td>–</td>
</tr>
<tr>
<td>Signal model</td>
<td>2.0</td>
<td>0.3</td>
</tr>
<tr>
<td>QCD model</td>
<td>1.1</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>Transfer functions</td>
<td>1.5</td>
<td>0.8</td>
</tr>
<tr>
<td>$R$ efficiencies ratio</td>
<td>0.3</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>Fit bias</td>
<td>2.1</td>
<td>–</td>
</tr>
<tr>
<td>Subdominant backgrounds ($t\bar{t}, W \rightarrow qq'$)</td>
<td>1.9</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td>Final-state radiation</td>
<td>0.9</td>
<td>–</td>
</tr>
<tr>
<td>$f_Z\leftrightarrow \ell$ fraction</td>
<td>0.1</td>
<td>–</td>
</tr>
<tr>
<td>Luminosity</td>
<td>1.2</td>
<td>–</td>
</tr>
<tr>
<td>Total</td>
<td>17.7</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Fig. 2. Simultaneous fit to the dijet invariant mass distribution of $Z \rightarrow b\bar{b}$ candidates in the (left) signal and (right) control regions.

Fig. 3. Background-subtracted distribution compared with the $Z \rightarrow b\bar{b}$ mass model in the (left) signal and (right) control regions. The one standard deviation total uncertainty band in the background-only hypothesis is also shown. This band includes statistical and systematic uncertainties.
lection efficiency is evaluated by measuring this efficiency in the $Z\rightarrow \mu^+\mu^-$ jet data and simulation samples and taking the difference as a systematic uncertainty.

The uncertainty on the model of the signal invariant mass distribution is determined by repeating the fit with an alternative distribution, consisting of the sum of two modified Gaussians. The uncertainty on the QCD model is determined by considering an alternative parametrization, consisting of an exponential decay model multiplied by a function that describes the effect of the jet $p_T$ requirements on the invariant mass distribution. It has been verified, by generating pseudoexperiments with this alternative model and by fitting them with the default model, that the choice of the QCD distribution model introduces a small bias in the measurement. This bias is taken as the systematic uncertainty.

The systematic uncertainty associated with the transfer functions is evaluated by repeating the fit using second-order polynomial functions instead of linear functions. In these fits the coefficients of the quadratic terms are varied in a range consistent with the data in the invariant mass sidebands used in the determination of the transfer functions. The maximum variation with respect to the default measurement is taken as the uncertainty. The efficiency ratio $R$ is determined using both $Z\rightarrow \mu^+\mu^-$ jet data and simulation, and the observed difference is taken as a systematic uncertainty. The uncertainty associated with a possible bias introduced by the fit procedure is determined using pseudoexperiments.

The fit is repeated introducing contributions from the subdominant backgrounds, $tt$ and $W\rightarrow q\bar{q}$, fixed to their SM expectations [30] and modelled with the simulation. The difference in the results is assigned as a systematic uncertainty. The final-state radiation systematic uncertainty is determined as described in Ref. [16]. The systematic uncertainty due to the $Z\rightarrow c\bar{c}$ contribution is dominated by the knowledge of the $Z\rightarrow c\bar{c}$ branching fraction [30] used in the evaluation of the $f_{Z\rightarrow c\bar{c}}$ parameter. The systematic uncertainty on the luminosity is determined as in Ref. [34].

The different sources of systematic uncertainties are considered to be uncorrelated and the total, relative systematic uncertainty is 17.7% for the cross-section measurement, dominated by the heavy-flavour tagging efficiency uncertainty (16.6%). The total systematic uncertainty for the jet energy scale measurement is 1.1% and is dominated by the uncertainty on the transfer functions (0.8%). The signficance of the signal yield, including all statistical and systematic uncertainties, is 6.0 standard deviations.

6. Results and conclusions

The product of the $Z$-boson production cross-section and the $Z\rightarrow bb$ branching fraction in $pp$ collisions at a centre-of-mass energy of 8 TeV is

$$\sigma(pp \rightarrow Z)f(Z \rightarrow b\bar{b}) = 332 \pm 46 \pm 59 \text{ pb},$$

where the first uncertainty is statistical and the second is systematic. The measurement is made in the fiducial region defined by two particle-level $b$ jets with $p_T > 20$ GeV, $2.2 < \eta < 4.2$, and $45 < m_{jj} < 165$ GeV.

The expected cross-section in the fiducial region of the experimental measurement is calculated at NLO using aMC@NLO plus Pythia for the parton showers and the NNPDF3.0 Parton Distribution Functions (PDFs) set [35]. The theoretical prediction determined in this way is

$$\sigma(pp \rightarrow Z)f(Z \rightarrow b\bar{b}) = 272^{+92}_{-72}(\text{scale}) \pm 5 \text{ (PDFs) pb},$$

where the first uncertainty is related to the missing higher-order corrections and to the error of the strong coupling constant, and the second uncertainty is related to the PDFs. The uncertainty due to missing higher-order corrections is evaluated by varying the renormalization and factorization scales by a factor of two around the nominal choice, and taking the maximum differences with respect to the nominal values. The uncertainty on the strong coupling is included by varying it within its uncertainty and re-calculating the cross-section. The uncertainty on the PDFs is estimated by taking the variance of the cross-section predictions, where each replica of the NNPDF3.0 set is used in turn. The prediction and the measurement are compatible within one standard deviation. The additional data being collected by the LHCb collaborators will allow a more stringent comparison with the theoretical prediction in the future. Moreover, the systematic uncertainty on the heavy-flavour tagging efficiency will be reduced by collecting more data [28].

The measured jet energy scale factor is

$$k_{E_S} = 1.009 \pm 0.015 \pm 0.011,$$

where the first uncertainty is statistical and the second uncertainty is systematic. The $k_{E_S}$ factor is compatible with unity, which demonstrates that the LHCb simulation reproduces accurately the $b$-jet energy in data for $bb$-jet pairs with about 100 GeV of invariant mass. Since a jet energy correction evaluated using simulation is already applied on $b$ jets, $k_{E_S}$ represents the residual correction obtained using the data.

Acknowledgements

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and PASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom), RRCIK and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (USA). We are indebted to the communities behind the multiple open-source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Sklodowska-Curie Actions and ERC (European Union), ANR, Labex P2IO, ENIGMASS and OCEVU, and Région Auvergne-Rhône-Alpes (France), RBFR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, the Royal Society, the English-Speaking Union and the Leverhulme Trust (United Kingdom).

References

LHCb Collaboration
