Measurement of b-quark fragmentation properties in jets using the decay B-+/- -> J/psi K-+/- in pp collisions at root s=13 TeV with the ATLAS detector

Aad, G.; Abbott, B.; Abbott, DC; Abud, AA; Abeling, K.; Abhayasinghe, DK; Abidi, S.H.; AbouZeid, Ossama Sherif Alexander; Abramowicz, H.; Abreu, H.; Camplani, Alessandra; Dam, Mogens; Ignazzi, Rosanna; Hansen, Jørgen Beck; Hansen, Peter Henrik; Hansen, Jørn Dines; Petersen, Troels Christian; Wiglesworth, Graig; Xella, Stefania; ATLAS Collaboration

Published in:
Journal of High Energy Physics (Online)

DOI:
10.1007/JHEP12(2021)131

Publication date:
2021

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
https://doi.org/10.1007/JHEP12(2021)131

Download date: 07. apr., 2022
Measurement of $b$-quark fragmentation properties in jets using the decay $B^\pm \rightarrow J/\psi K^\pm$ in $pp$ collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

ABSTRACT: The fragmentation properties of jets containing $b$-hadrons are studied using charged $B$ mesons in 139 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 13$ TeV, recorded with the ATLAS detector at the LHC during the period from 2015 to 2018. The $B$ mesons are reconstructed using the decay of $B^\pm$ into $J/\psi K^\pm$, with the $J/\psi$ decaying into a pair of muons. Jets are reconstructed using the anti-$k_t$ algorithm with radius parameter $R = 0.4$. The measurement determines the longitudinal and transverse momentum profiles of the reconstructed $B$ hadrons with respect to the axes of the jets to which they are geometrically associated. These distributions are measured in intervals of the jet transverse momentum, ranging from 50 GeV to above 100 GeV. The results are corrected for detector effects and compared with several Monte Carlo predictions using different parton shower and hadronisation models. The results for the longitudinal and transverse profiles provide useful inputs to improve the description of heavy-flavour fragmentation in jets.

KEYWORDS: Hadron-Hadron scattering (experiments)

ArXiv ePrint: 2108.11650
1 Introduction

The fragmentation of heavy quarks is a crucial aspect of quantum chromodynamics (QCD). Detailed studies and precision measurements of the heavy-quark fragmentation properties allow a deeper understanding of QCD. Furthermore, reliable modelling of the fragmentation is of great importance for measurements of the production of Higgs bosons [1, 2], top quarks [3, 4] and their associated production [5, 6], whose hadronic decays predominantly feature heavy quarks. Uncertainties related to the modelling of the fragmentation processes of $b$-quarks into hadrons are significant in the most precise top quark mass determinations [7–11], and are also the subject of theoretical study [12–15]. This subject has been studied in $e^+e^-$ collisions for charm [16–19] and bottom quarks [20–24]. In hadron-hadron collisions, measurements of observables sensitive to the heavy-flavour fragmentation functions have been reported for $D^*$ mesons [25–27], $D^0$ mesons [28] and $J/\psi$ quarkonia [29].

The Monte Carlo (MC) predictions used at the LHC are tuned to describe the measurements in $e^+e^-$ collisions at relatively low centre-of-mass energies. Therefore, new measurements of $b$-quark fragmentation can be used to improve the MC simulation at LHC energy scales.

This analysis presents a measurement of the fragmentation of $b$-quarks into charged $B$ mesons using the ATLAS full Run 2 dataset, containing 139 fb$^{-1}$ of $pp$ collisions at
$\sqrt{s} = 13$ TeV. To this end, the $B^\pm$ meson is reconstructed using the decay chain $B^\pm \rightarrow J/\psi K^\pm \rightarrow \mu^+\mu^- K^\pm$. The procedure is similar to those in previous ATLAS measurements involving $B^\pm \rightarrow J/\psi K^\pm$ final states [30, 31]. Jets are reconstructed using the anti-$k_t$ algorithm with radius parameter $R = 0.4$. The reconstructed $B$ mesons are matched to jets and the longitudinal and transverse momentum profiles, $z$ and $p_T^{rel}$, are defined as

$$z = \frac{\vec{p}_B \cdot \vec{p}_j}{|\vec{p}_j|^2}; \quad p_T^{rel} = \frac{|\vec{p}_B \times \vec{p}_j|}{|\vec{p}_j|},$$

where $\vec{p}_B$ is the three-momentum of the $B$ hadron and $\vec{p}_j$ is the three-momentum of the jet. The longitudinal profile, $z$, quantifies the fraction of the jet momentum carried by the $B$ meson in the direction parallel to the jet axis. On the other hand, the transverse profile, $p_T^{rel}$, quantifies the momentum of the $B$ meson in the direction orthogonal to the jet axis. These variables are sensitive to the fragmentation function $D_h^b(x, Q^2)$, which is defined as the probability of a quark $q$ to fragment into a hadron $h$ with an energy fraction $x$ at a scale $Q$ [32, 33]. The measurement is performed in different intervals of the jet transverse momentum, which provides a probe of the scaling of the fragmentation functions. Furthermore, as discussed in the following, these observables are also sensitive to the contributions of gluons producing a $b\bar{b}$ pair. Since, in many cases, $b\bar{b}$ pairs arising from the splitting of high-$p_T$ gluons are not resolved into two different jets, the reconstructed $B$ meson carries a smaller fraction of the jet energy, resulting in a flatter distribution of $z$ and $p_T^{rel}$ compared to jets with a single hard-scattering $b$-quark.

The paper is structured as follows. Section 2 is dedicated to the description of the ATLAS detector. A summary of the MC samples and the dataset used throughout the analysis is included in section 3, and the object and event selection is described in detail in section 4. Section 5 is dedicated to the estimation of the purity corrections, while a discussion of the subsequent corrections for detector effects is found in section 6. The uncertainties affecting this measurement are discussed in section 7 and the results are presented in section 8. Section 9 provides the summary and conclusions.

## 2 The ATLAS detector

The ATLAS detector [34] at the LHC covers nearly the entire solid angle around the collision point. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting toroidal magnets with eight coils each.

The inner-detector system is immersed in a 2T axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track, the first hit

---

1. ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) 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 upwards. Cylindrical coordinates $(r, \phi)$ 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 $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. 

---

- 2 -
normally being in the insertable B-layer installed before Run 2\cite{35, 36}. It is followed by the silicon microstrip tracker, which usually provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$. The TRT also provides electron identification information based on the fraction of hits (typically 30 in total) above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic measurements respectively.

The muon spectrometer comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated by the superconducting air-core toroids. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. A set of precision chambers covers the region $|\eta| < 2.7$ with three layers of monitored drift tubes, complemented by cathode-strip chambers in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions.

Interesting events are selected to be recorded by the first-level trigger system implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger\cite{37}. The first-level trigger accepts events from the 40 MHz bunch crossings at a rate below $100 \text{ kHz}$, which the high-level trigger reduces in order to record events to disk at about $1 \text{ kHz}$.

An extensive software suite\cite{38} is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

## 3 Data and Monte Carlo samples

The dataset used in this analysis comprises the data taken from 2015 to 2018 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. After applying quality criteria to ensure good ATLAS detector operation\cite{39}, the total integrated luminosity used in this analysis is 139 fb$^{-1}$\cite{40, 41}. The uncertainty in the combined 2015–2018 integrated luminosity is 1.7\%\cite{40}, obtained using the LUCID-2 detector\cite{42} for the primary luminosity measurements. The measurements of the fragmentation properties are unaffected by this uncertainty, given that they are normalised to the total cross section, thus cancelling out the contribution of the luminosity uncertainty. The average number of inelastic $pp$ interactions produced per bunch crossing for the dataset considered, hereafter referred to as ‘pile-up’, is $\langle \mu \rangle = 33.6$.

Several different models of multijet production are used in this analysis. The MC samples were generated using the *Pythia* 8\cite{43, 44}, *Sherpa*\cite{45}, and *Herwig* 7\cite{46––}
48] generators. These models differ in the matrix element (ME) calculation, the parton shower (PS) and the hadronisation model (HM). The main features of these samples are summarised in table 1.

The Pythia 8 samples were generated using Pythia 8.240. The matrix element was calculated at leading order for the $2 \rightarrow 2$ process. The PS algorithm includes initial- and final-state radiation based on the dipole-style $p_T$-ordered evolution, including $g \rightarrow q\bar{q}$ branchings and a detailed treatment of the colour connections between partons [43]. The renormalisation and factorisation scales were set to the geometric mean of the squared transverse masses of the two outgoing particles (labelled 3 and 4), i.e.

$$\mu^2_r = \mu^2_f = \sqrt{m^2_{T3} \cdot m^2_{T4}} = \sqrt{(p^2_T + m^2_3) \cdot (p^2_T + m^2_4)}.$$  

Two different sets of hadronisation and underlying-event parameter values (tunes) were used in the generation of the Pythia 8 samples. Two of the samples make use of the ATLAS A14 tune [49], for which the CTEQ6L1 PDF set [50] was used for the ME generation, the PS, and the simulation of multi-parton interactions (MPI). Two additional samples make use of the Monash tune [51], interfaced to the NNPDF2.3lo PDF set [52] for the ME generation, PS and MPI. The hadronisation is modelled using the Lund string model [53, 54] in all samples. The two samples using the A14 tune make use of the Lund-Bowler parameterisation of the fragmentation function [55], differing in the value of the parameter $r_b$, which controls the shape of the fragmentation function for $b$-quarks: while the first of them uses the nominal value $r_b = 0.855$, the second one uses the so-called A14-rb tune, for which $r_b = 1.05$, as obtained from a combined fit to LEP and SLD data [20–23]. The two samples using the Monash tune make use of the Peterson [56] and Lund-Bowler [55] parameterisations for the fragmentation functions, where the $r_b$ parameter is set to $r_b = 0.855$ for the Lund-Bowler sample. The Pythia 8 sample making use of the A14-rb tune is taken as the nominal MC sample, used across the analysis for unfolding and uncertainty estimation.

The Sherpa samples were generated using Sherpa 2.2.5. The matrix element was calculated at leading order (LO) for the $2 \rightarrow 2$ process, and the default Sherpa CSS dipole PS [57, 58]. Matrix element renormalisation and factorisation scales for $2 \rightarrow 2$ processes were set to the harmonic mean of the Mandelstam variables $s$, $t$ and $u$ [59]. The CT14nnlo [60] PDF set was used for the matrix element calculation, as well as for the modelling of the PS and MPI. The different Sherpa samples use different hadronisation models. The first one makes use of the SHERPA AHADIC model for hadronisation [61], which is based on the cluster hadronisation algorithm [62], while the second one uses the Lund string model [43, 53, 54, 63] for the modelling of the hadronisation.

Finally, two Herwig 7 samples were generated at leading order using Herwig 7.2.1. The matrix element for the $2 \rightarrow 2$ process was calculated at LO with the MMHT2014lo PDF [64]. The renormalisation and factorisation scales were set as

$$\mu^2_r = \mu^2_f = \frac{2stu}{s^2 + t^2 + u^2},$$

where $s$, $t$ and $u$ are the Mandelstam variables [59]. The first sample uses the default Herwig 7 angle-ordered PS, while the second sample uses a dipole-based PS [65]. For both
<table>
<thead>
<tr>
<th>Generator</th>
<th>ME order</th>
<th>Scales $\mu_r, \mu_f$</th>
<th>Parton shower</th>
<th>PDF set</th>
<th>Tune</th>
<th>Hadronisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pythia 8</td>
<td>$2 \rightarrow 2 \oplus 0 \text{ LO}$</td>
<td>$m_{T3} \cdot m_{T4}^{1/2}$</td>
<td>$p_T$-ordered</td>
<td>CTEQ6L1</td>
<td>A14</td>
<td>Lund-Bowler</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A14-RB</td>
<td>Monash</td>
<td>Lund-Bowler</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NNPDF2.3</td>
<td></td>
<td>Lund-Bowler</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Peterson</td>
</tr>
<tr>
<td>Sherpa</td>
<td>$2 \rightarrow 2 \oplus 0 \text{ LO}$</td>
<td>$H(s,t,u)$</td>
<td>CSS (dipole)</td>
<td>CT14</td>
<td>—</td>
<td>Cluster model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lund string model</td>
</tr>
<tr>
<td>Herwig 7</td>
<td>$2 \rightarrow 2 \oplus 0 \text{ LO}$</td>
<td>$\sqrt{\frac{2m_{T}}{s+t+u}}$</td>
<td>Angle-ordered</td>
<td>MMHT2014</td>
<td>—</td>
<td>Cluster model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dipole</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Properties of the Monte Carlo samples used in the analysis, including the perturbative order in $\alpha_s$, the number of final-state partons, the PDF set, the parton shower algorithm, the renormalisation and factorisation scales, the tune and the hadronisation model. Further details can be found in the text.

Herwig 7 samples, the MMHT2014LO [64] PDF was used for the modelling of the MPI, and the hadronisation was modelled by means of the default Herwig 7 cluster hadronisation algorithm.

The decays of the $B$ mesons, of particular importance for this analysis, were modelled using the EvtGen 1.6.0 generator [66] for the totality of the samples described above. The samples simulated using Pythia 8 A14-RB, Pythia 8 A14, and Sherpa with string hadronisation were passed through the GEANT4-based [67] ATLAS detector-simulation program [68] since they were also used to correct the measurements for detector effects, as described in section 6. They are reconstructed and analysed with the same processing chain as the data. These fully simulated samples include the effect of multiple pp interactions per bunch crossing, simulated using Pythia 8.186 interfaced to the A3 tune [69], as well as the effect on the detector response of interactions from bunch crossings before or after the one containing the hard interaction. In addition, during the data-taking, some modules of the ATLAS hadronic calorimeter were disabled for some periods of time. The resulting non-functioning regions in the hadronic calorimeter were not necessarily included in the simulation for all the samples. However, the effect of removing jets pointing towards these regions has been found to be negligible.

4 Object and event selection

Events are selected using triggers optimised for $J/\psi$ meson identification and selection in its decay into muon pairs [70]. These triggers select events with two muons with $p_T > 6$ GeV. During the 2017 and 2018 data-taking periods, the trigger also required the invariant mass of the dimuon pair to satisfy $2$ GeV < $m_{\mu\mu}$ < $9$ GeV and the angular separation between the muons to satisfy $\Delta R_{\mu\mu} < 1.5$. The trigger selection efficiency is about 60%.

The primary vertex of the event is reconstructed as the vertex maximising the value of $\sum p_T^2$ for all tracks originating from it. The reconstruction of $B^{\pm}$ mesons is done by re-fitting a pair of oppositely charged muons with $p_T > 6$ GeV and $|\eta| < 2.5$ to a common vertex, with a fit quality fulfilling $\chi^2 < 10$. Both muons are reconstructed using information...
from both the inner detector and the muon spectrometer. They must pass the Medium quality requirements [71, 72], including at least one hit in the precision chambers of the muon spectrometer, and are required to have an invariant mass consistent with the mass of the $J/\psi$ meson, i.e. to lie within the range 2.6–3.6 GeV. Moreover, both muons must spatially match the muon objects used in the trigger selection within $\Delta R = 0.01$. Given the short $J/\psi$ lifetime following the $B^\pm \to J/\psi K^\pm$ decay, a third track with $p_T > 4$ GeV and $|\eta| < 2.5$ is fitted to a common vertex together with the two muon tracks, requiring $\chi^2/N_{dof} < 2.0$. In the fit, the invariant mass of the pair of muon tracks is constrained [73] to match the world average value for the $J/\psi$ mass, $m_{J/\psi} = 3096.900$ MeV [74]. Moreover, each of the three tracks in the triplet must have at least one hit in the pixel detector and at least four hits in the silicon microstrip detector, while the invariant mass of the triplet is required to lie within the range 5.0–5.7 GeV. Finally, the pseudo-proper lifetime of the $B^\pm$ candidate, $\tau = m_B L_{xy}/p_T$, where $m_B$ is the PDG value of the $B^\pm$ mass, $m_B = 5279.320$ MeV [74], and $L_{xy}$ is the transverse flight distance from the primary vertex, is required to be above 0.20 ps. The efficiency of this cut has been shown to be well described by the MC simulations, and hence to have only a small impact on the results after the corrections for detector effects are applied. The mass of the charged kaon, $m_K = 493.677$ MeV [74], is assumed for the third track for the sake of calculating the three-body invariant mass.

Jet reconstruction is performed using particle-flow objects [75, 76] as inputs to the anti-$k_T$ algorithm as implemented in FastJet [77, 78] with a radius parameter $R = 0.4$. Muon tracks are not considered in the particle-flow algorithm and, thus, the energy of muons from the $B^\pm$ decay chain is not accounted for in the jet energy. In order to correct for the presence of muons inside the jets, muons passing the same quality requirements as those used in the $B^\pm$ reconstruction are matched to the jet by means of the ghost-association method [79]. This procedure defines objects with infinitesimal energy and the same direction as the track momentum (ghosts), which are then used as inputs to the jet reconstruction algorithm. A track is considered matched to a given jet if the ghost associated with it is clustered inside the jet. The jet four-momentum is then corrected by vectorially adding to it the four momenta of the muons associated with the jet, after subtracting the muon energy loss in the calorimeter. After this correction is applied, the jet energy is calibrated and, in order to ensure full inner-detector acceptance, jets with $p_T > 20$ GeV and $|\eta| < 2.1$ are preselected. The jet calibration procedure includes energy corrections for pile-up, as well as angular corrections. Effects due to energy losses in inactive material, shower leakage, the magnetic field and inefficiencies in energy clustering and jet reconstruction are taken into account. This is done using a simulation-based correction, in bins of $\eta$ and $p_T$, derived from the relation of the reconstructed jet energy to the energy of the corresponding particle-level jet. In a final step, an in situ calibration corrects for residual differences in the jet response between the MC simulation and the data using $p_T$-balance techniques for dijet, $\gamma + \text{jet}$, $Z + \text{jet}$ and multijet final states. In order to reject pile-up jets, the so-called ‘jet vertex tagger’ (JVT) algorithm is used [80]. Moreover, in order to avoid overlaps between jets (and thus, incorrect matchings with $B$ mesons), jets at a distance $\Delta R < 0.8$ from any other jet with $p_T > 20$ GeV are discarded.
The reconstructed $B^{\pm}$ mesons are then matched to the corresponding jet by requiring both objects to be within $\Delta R = 0.4$ of each other. If there is more than one $B$ meson candidate for a single jet, the one with lowest $\chi^2$ to the three-track vertex is selected. If any of the three tracks arising from the $B^{\pm}$ decay is found to lie at $\Delta R > 0.4$ from the jet, its four-momentum is added to the jet four-momentum. The basic object for this analysis is, therefore, a jet containing a $B^{\pm}$ meson. The selection criteria described above are met by 1,413,684 jets in 1,404,620 events.

To correct the simulation for differences in the muon identification and JVT acceptance efficiencies from those measured in the data, simulated events are weighted using dedicated corrections provided to that end [71, 80]. The differences in the trigger efficiencies between data and the simulation are accounted for as a systematic uncertainty, as described in section 7.

Figure 1 shows the detector-level distributions of the jet $p_T$ and the $p_T$ of the $B^{\pm}$ meson within the jet, together with the MC predictions. The data distributions include the purity corrections described in section 5. While the SHERPA sample with string-based fragmentation describes the jet $p_T$ within 10%, it fails to describe the $B$ meson $p_T$, with discrepancies of over 50%. The PYTHIA samples with the A14 and A14-rb tunes give similar descriptions of the jet $p_T$, with differences from data of up to 40% in the tail of the distribution. The differences between the two tunes are clearly observed in the distribution of the $B$ meson $p_T$, where A14 gives a more accurate description within 10% while A14-rb shows discrepancies of up to 30% in the tail of the distribution.

For the fragmentation measurement, the data are binned in three intervals of the jet $p_T$, namely [50, 70), [70, 100) and $p_T \geq 100$ GeV. This choice ensures that the bin width is larger than the resolution in the different $p_T$ intervals [75]. For each of these intervals, the
$z$ distributions are binned in intervals with a width of 0.07, while the $p_T^{\text{rel}}$ distributions use a variable bin width. This choice allows finer binning, while retaining sufficient statistical precision in each bin, as needed for the estimation of the signal purity described in section 5. The lower limit of the $z$ distribution is extended towards lower values with increasing jet $p_T$, where a larger fiducial phase space becomes available. The $z$ and $p_T^{\text{rel}}$ distributions are normalised to unit area by dividing by the number of entries, allowing the systematic uncertainties to be reduced.

Since the momentum resolution for charged particles is much better than the jet energy resolution, the measured $z$ distribution extends beyond unity. The values larger than one are absorbed as an overflow into the last bin of the $z$ distribution. Similarly, values of $z$ below the lower limit of the first bin are absorbed into the first bin as an underflow, and values of $p_T^{\text{rel}}$ above the upper limit of the last bin are absorbed into it as an overflow.

### 5 Signal extraction

The selected sample of $B^\pm$ meson candidates does not only contain $B^\pm$ mesons, but also backgrounds arising from different sources. For each $(p_T, x)$ bin, with $p_T$ being the jet $p_T$ and $x = z$ or $p_T^{\text{rel}}$, only a fraction $p(p_T, x)$ of the reconstructed entries, referred to as the purity, are real $B^\pm$ mesons. The number of $B^\pm$ mesons, $R$, reconstructed in a given bin can thus be expressed as

$$ R(p_T, x) = N(p_T, x) \times p(p_T, x), $$

where $N$ is the total number of candidates in a given bin. In order to determine $p(p_T, x)$, a binned maximum-likelihood fit to the invariant mass distribution of the $B^\pm$ candidates is performed. The probability density function (pdf) of the model describing the invariant mass distribution can be written as a combination of the signal and background pdfs as

$$ F(m) = \lambda_s F_s(m) + \lambda_{B_x} F_{B_x}(m) + \lambda_{B_\pi} F_{B_\pi}(m) + \lambda_c F_c(m), $$

where $F_s$, $F_{B_x}$, and $F_c$ are the pdfs for each of the components, signal or background, and $\lambda_s$, $\lambda_{B_x}$, $\lambda_{B_\pi}$ and $\lambda_c$ are coefficients representing their relative fractions, and thus $\lambda_s(p_T, x) = p(p_T, x)$. The closure relation $\sum_i \lambda_i = 1$ must be satisfied by these coefficients. Each of the fit components is described below.

- The signal $F_s$, arising from the real $B^\pm$ meson contribution, is modelled by a double-Gaussian function

$$ F_s(m|\mu, \sigma_1, \sigma_2) = \frac{1}{\sqrt{2\pi}} \frac{\beta}{\sigma_1} \exp \left[ \frac{-(m-\mu)^2}{2\sigma_1^2} \right] + \frac{1-\beta}{\sigma_2} \exp \left[ \frac{-(m-\mu)^2}{2\sigma_2^2} \right], $$

where $\beta$ is the relative normalisation of the two Gaussian components.

- The misreconstructed background, $F_{B_x}$, arising from the decays $B^{\pm/0} \to J/\psi K^{*\pm/0} \to J/\psi(K\pi)^{\pm/0}$ and $B^{\pm/0} \to J/\psi(K\pi)^{\pm/0}$, creates a low-mass structure, displaced from the $B^\pm$ mass. It is modelled by the following function

$$ F_{B_x}(m|b, s) = 1 - \tanh \left( \frac{m - s}{b} \right). $$
The resonant background $F_{B^\pm}$, arising from the decays $B^\pm \to J/\psi \pi^\pm$, creates a peaking structure displaced from the $B^\pm$ mass towards higher masses. It is modelled by the sum of a Gaussian and an asymmetric Gaussian function,

$$F_{B^\pm}(m|\mu_1, \mu_2, \sigma_1, \sigma_2, \sigma_3, \gamma) = \frac{1}{\sqrt{2\pi}} \left\{ \frac{\gamma}{\sigma_1} \exp \left[ -\frac{(m - \mu_1)^2}{2\sigma_1^2} \right] + (1 - \gamma)G_{\text{asym}}(m|\mu_2, \sigma_2, \sigma_3) \right\}, \quad (5.1)$$

where $\gamma$ is a relative normalisation factor and

$$G_{\text{asym}}(m|\mu_2, \sigma_2, \sigma_3) = \begin{cases} \frac{1}{\sigma_2} \exp \left[ -\frac{(m - \mu_2)^2}{2\sigma_2^2} \right] & \text{if } m \leq \mu_2 \\ \frac{1}{\sigma_3} \exp \left[ -\frac{(m - \mu_2)^2}{2\sigma_3^2} \right] & \text{if } m > \mu_2. \end{cases}$$

The combinatorial background $F_c$, arising from random combinations of real $J/\psi$ mesons with additional tracks, is modelled by a first-order polynomial

$$F_c(m|p_0, p_1) = p_0 + p_1 m.$$ 

All the parameters in eq. (5.1) ($\mu_1, \mu_2, \sigma_1, \sigma_2, \sigma_3, \gamma$) are obtained from fits to MC simulations, performed using PYTHIA8 A14-rb, of $B^\pm \to J/\psi \pi^\pm$ in bins of the jet $p_T$ and independently of $z$ or $p_{T1}^{\text{rel}}$. The normalisation of this background, $\lambda_{B^\pm}$, is fixed relative to the signal as the ratio $R(\pi, K)$ of the branching fractions of signal and background [74], i.e.

$$\frac{\lambda_{B^\pm}}{\lambda_s} = R(\pi, K) = \frac{B(B^\pm \to J/\psi \pi^\pm)}{B(B^\pm \to J/\psi K^\pm)} = 3.84\%.$$ 

The fits are performed individually for each ($p_T, z$) and ($p_T, p_{T1}^{\text{rel}}$) bin in the mass range [5.0, 5.7] GeV, and each of them has ten free parameters: four for the signal ($\mu_1, \sigma_1, \sigma_2, \beta$), two for the misreconstructed background ($b, s$), two for the combinatorial background ($p_0, p_1$), and two for the free normalisations of each component, $\lambda_i$. Figure 2 shows the results of such fits in four representative bins of the longitudinal profile $z$ and transverse profile $p_{T1}^{\text{rel}}$ for the lowest and highest jet-$p_T$ selections.

The quality of the fits is quantified using $\chi^2$ functions, defined as

$$\chi^2 = \sum_i [(x_i - f_i)^2/(\Delta x_i)^2]$$

for each fit. Here, $x_i$ is the value of the data distribution in bin $i$; $\Delta x_i$ is its statistical uncertainty, and $f_i$ is the corresponding value of the fit model. The values of the $\chi^2$ divided by the number of degrees of freedom, $N_{\text{dof}} = 30$, range from 0.6 to 2.2. The fit results are studied as a function of the jet $p_T$ and of the fragmentation variables $z$ and $p_{T1}^{\text{rel}}$. The results show that the signal component dominates the fit models with an average, global purity close to 69%, followed by the misreconstructed background, with a global value of about 15% and the combinatorial background, amounting to 12% of the sample. Finally, the $B^\pm \to J/\psi \pi^\pm$ background amounts to 3% of the sample. As a general trend, the purity of the sample increases as a function of $z$, while it is relatively constant as a function of $p_{T1}^{\text{rel}}$, particularly at low $p_T$. Figure 3 shows the signal purity and the fractions...
of the different backgrounds resulting from the fit as a function of $z$ and $p_{\text{T}}^{\text{rel}}$, for the lowest and highest jet-$p_T$ selections. The hatched bands include the sum in quadrature of the statistical and systematic uncertainties of the fit, as discussed in section 7.

6 Unfolding to particle level

The detector-level distributions, from which the background processes have been subtracted following the methodology described in section 5, are corrected for detector effects such as inefficiencies and migrations caused by the finite resolution of the detector. The particle-level observables are defined using requirements equivalent to the detector-level selection.
Figure 3. Fit results as a function of \( z \) for all jet-\( p_\text{T} \) bins (left) and as a function of \( p_\text{T}^{\text{rel}} \) (right) for all jet-\( p_\text{T} \) bins. The black circular markers represent the purity \( p(p_\text{T}, z) \) for each bin. The error bars represent the statistical uncertainty of the fitted values, while the hatched bands represent the sum in quadrature of the statistical and systematic uncertainties of the fits.

described in section 4. Candidate \( B^\pm \to J/\psi K^\pm \) decays are selected if the muons from the \( J/\psi \) decay have \( p_\text{T} > 6 \) GeV and the kaon has \( p_\text{T} > 4 \) GeV. The kinematics of the \( B^\pm \) hadrons are reconstructed before QED radiation. Jets are reconstructed using the anti-\( k_t \) algorithm with \( R = 0.4 \), using all particles with average lifetime \( \tau > 10 \) ps, including muons and neutrinos. Jets with \( p_\text{T} > 20 \) GeV and \( |\eta| < 2.1 \) are selected, and those overlapping with another jet with \( p_\text{T} > 20 \) GeV within \( \Delta R = 0.8 \) are discarded. While the measurement is performed in the fiducial phase space with \( p_\text{T} > 50 \) GeV, an underflow bin with \( 35 \) GeV < \( p_\text{T} < 50 \) GeV is used to take into account the migrations from phase space.
regions below the jet $p_T$ threshold. The $B^\pm$ mesons are matched to jets if the angular distance between them fulfils $\Delta R < 0.4$. Finally, the four-momenta of any decay products of the $B^\pm$ meson lying outside the jet ($\Delta R > 0.4$) are added to the jet four-momentum.

The unfolding procedure can be parameterised using a transfer matrix $A_{ij}$, which contains the information about the bin-by-bin migrations from the detector-level distribution to the particle-level distributions, and is written as

$$R_i = \sum_{j=1}^N \frac{E_j}{P_i} A_{ij} T_j.$$ 

The unfolding can thus be regarded as a system of linear equations, for which the solution is the particle-level distribution $T_j$. The transfer matrix $A_{ij}$, the efficiency $E_j$ and the purity $P_i$ are determined using the Pythia 8 A14-rb MC sample, and $R_i$ are the values of the detector-level distribution. An iterative matrix inversion method based on Bayes’ theorem [81] is used, as implemented in the RooUnfold program [82]. The number of iterations is chosen so that the sum in quadrature of the statistical uncertainty and the uncertainty associated with the mismodelling in the unfolding, described in section 7, is minimised. This results in four iterations for the unfolding of both the longitudinal and transverse profiles. The unfolding is performed in such a way that the migrations in both dimensions, between the fragmentation variable ($z$ or $p_T^{rel}$) and the jet $p_T$ are taken into account. The correlations in the statistical uncertainties, as well as those between the particle-level and detector-level MC distributions, are treated by using pseudo-experiments [83] in the calculation of each component of the unfolding, including matrices, efficiencies and purities as well as in the detector-level data.

The efficiency $E$ is estimated as the fraction of particle-level jets, associated with a $B$ meson, which are matched within $\Delta R = 0.4$ to a detector-level jet that contains a reconstructed $B$ meson. Its value ranges from 10% to 30% depending on $z$ and $p_T^{rel}$. The efficiency decreases as a function of $z$ for all $p_T$ bins, although this behaviour is more pronounced for the higher $p_T$ bin. This is due to the fact that, in this boosted regime, the dimuon trigger cannot efficiently separate the trajectories of the two muons produced in the $J/\psi$ decay, and the events containing $B^\pm$ mesons are therefore not recorded efficiently. The values of $E$ as a function of $p_T^{rel}$ are significantly flatter.

The purity $P$ is estimated as the fraction of detector-level jets which are matched to a particle-level jet within $\Delta R = 0.4$. Its value ranges from 60% to 100%, systematically increasing as a function of both $z$ and the jet $p_T$. As a function of $p_T^{rel}$, the purity exhibits a decrease in the low $p_T$ region. This is due to the fact that reconstructed mesons flying in the same direction as the jet, as well as hadrons with high $p_T$, are more likely to arise from a true $B$ meson. In the high $p_T$ region, $P$ is constant as a function of $p_T^{rel}$.

The values of the matrix elements $A_{ij}$ are estimated using particle-level $B$-meson jets, geometrically matched to the corresponding detector-level objects within $\Delta R = 0.4$. Figure 4 shows the two-dimensional transfer matrices as a function of $(p_T, z)$ and $(p_T, p_T^{rel})$.

\footnote{Data falling in this bin are also corrected for background sources as described in section 5.}

\footnote{This purity is unrelated to the purity described in section 5 and corrected for before the unfolding.}
Figure 4. Transfer matrices, obtained using the Pythia 8 A14-rb MC sample, as a function of the jet $p_T$ and the longitudinal profile $z$ (top) and the transverse profile $p_T^{rel}$ (bottom). Each small square represents a bin in $(p_T, z)$ or $(p_T, p_T^{rel})$, while each large box represents one of the three jet-$p_T$ intervals used in the analysis. The first bin, containing jets with $p_T < 50$ GeV, accounts for the underflow.

7 Systematic uncertainties

The systematic uncertainties in the measurement are classified into four main categories. The first of them concerns the identification of $B^\pm$ mesons, and includes the systematic uncertainties in muon reconstruction as well as the uncertainties in the fit procedure for the determination of the signal purity described in section 5; the second category concerns jet reconstruction and identification, including the jet energy scale and resolution, the jet angular resolution and the efficiencies of the JVT algorithm. The third category is related
to the unfolding procedure, including both the uncertainty in the MC model used for
the corrections and the uncertainty due to mismodelling of the data by the nominal MC
simulation. The fourth category takes into account systematic effects in the MC description
of the pile-up dependence of the measurement.

The systematic uncertainties from all experimental sources are estimated by using
the detector-level MC events. The resulting shifted distributions are then corrected for
detector effects following the unfolding procedure described in section 6, normalised to
unit area, and compared with the nominal MC distributions, also normalised to unit area.
The results of this comparison constitute the systematic uncertainties of the unfolded
results for the normalised differential cross sections, and exploit the correlations between
the numerator and the denominator when applying the normalisation. The full set of
systematic uncertainties are described below, together with a detailed description of how
they are estimated.

**B meson reconstruction uncertainties:**

- The uncertainties in the purity corrections for the $B^\pm$ candidates are evaluated by
  considering other fit models as alternatives to those described in section 5. The
  probability density function for each component of the fit is substituted by an alter-
  native function. For the signal model, one of the Gaussian functions in the
double-Gaussian pdf is replaced by an asymmetric Gaussian function. The hyper-
  bolic tangent shape in the misreconstructed-background pdf, $1 - \tanh(x)$, is replaced
  by $1 - \text{erf}(x) = (2/\sqrt{\pi}) \int_x^\infty e^{-t^2} dt$. For the $B^\pm \to J/\psi \pi^\pm$ background, one of the
  Gaussian functions in the double-Gaussian shape is replaced by a Crystal Ball func-
  tion. Finally, the first-order polynomial describing the shape of the combinatorial
  background is replaced by an exponential function. The total uncertainty in the pu-
  rity corrections is defined as the sum in quadrature of the deviations of the varied fit
  results, evaluated separately, from the nominal fit model described in section 5. The
  impact of this uncertainty ranges from 1% to a maximum of 17% for the measurement
  of $z$, and from 1% to 8% for the measurement of $p_T^{\text{rel}}$, systematically increasing with
  the jet $p_T$.

- Muon momentum scale and resolution uncertainties [71] include variations in the
  smearing of the inner detector and muon spectrometer tracks, variations of the muon
  momentum scale, and additional charge-dependent variations related to the correction
  for the sagitta bias. These variations are applied to the muon selection and both
  when adding the muon four-momenta to the jets using the ghost-association method
described in section 4 and when reconstructing the mass and four-momenta of the $B^\pm$
  mesons. These are applied in a fully correlated way. The impact of this uncertainty
  is below 2% in all regions of the phase space, for both the $z$ and $p_T^{\text{rel}}$ measurements.

- Muon identification uncertainties [72] cover variations of the muon efficiency correc-
  tions described in section 4. The variations include $1\sigma$ shifts of the statistical and
  systematic uncertainties in the efficiency corrections. As in the muon reconstruction
case, the variations are applied, in a fully correlated way, both when adding muons to the jets and when reconstructing the $B$ hadron momenta. The impact of this uncertainty is small, with values below 3% in all regions of the phase space.

- The uncertainty in the dimuon trigger efficiency [70] is evaluated by using correction factors to take into account the uncertainty in the mismodelling of the trigger efficiency. These correction factors are applied as per-jet weights depending on the $p_T$ and rapidity of the muons, as well as the angular distance $\Delta R$ between the two muons in the $B^\pm$ decay. The difference between the weighted distribution and the default one defines the systematic uncertainty. The values of this uncertainty range from a few per mille to a maximum of 4% in the lower tails of the $z$ distribution, and from a few per mille to a maximum of 2% as a function of $p_T^{\text{rel}}$.

- The uncertainty in the kaon reconstruction efficiency is estimated by randomly rejecting 2% of the kaon tracks from the reconstruction, following a uniform random distribution [84]. The impact on the fragmentation properties is at the per mille level for both the $z$ and $p_T^{\text{rel}}$ measurements.

**Jet-related uncertainties:**

- The jet energy scale (JES) and jet energy resolution (JER) uncertainties are estimated as described in ref. [76]. The JES is calibrated on the basis of the simulation and in situ corrections obtained from data. The JES uncertainties are estimated using a correlation scheme comprising a set of 29 independent components, which depend on the jet $p_T$ and $\eta$. The total JES uncertainty in the $p_T$ of individual jets ranges from 2% at $p_T = 50$ GeV to approximately 1% at $p_T = 100$ GeV, with a mild dependence on $\eta$. The JER uncertainty is estimated using a correlation scheme involving 13 independent variations. Each of these variations ranges from about 0.8% at $p_T = 50$ GeV to about 0.5% at $p_T = 100$ GeV. In this measurement, the JES and JER uncertainties are propagated by varying the energy and $p_T$ of each jet by one standard deviation of each of the independent components. The impact of the JES uncertainty varies from a few per mille for low $p_T^{\text{rel}}$ values at high jet-$p_T$ to 10% for medium values of $z$ at low jet-$p_T$, and typically decreases with increasing jet $p_T$. The JER has a similar behaviour, varying from approximately 1% for low $p_T^{\text{rel}}$ values at high jet $p_T$ to 30% at medium values of $z$ at low jet-$p_T$.

- The impact of the jet angular resolution (JAR) uncertainty is estimated by smearing the angular coordinates ($\eta, \phi$) of the jets by 10% of the angular resolution of jets reconstructed from topological clusters [85], estimated in MC simulation. The $\eta$ and $\phi$ smearing is applied with the $p_T$ component of the jets held constant. The JAR uncertainty is negligible for the measurement of $z$. On the other hand, due to a linear dependence of $p_T^{\text{rel}}$ on the sine of the angle between the $B$ meson and the jet, this uncertainty can have an impact of up to 3% for the measurement of $p_T^{\text{rel}}$.

- The uncertainty from the efficiency corrections for the JVT algorithm described in section 4, used to mitigate the impact of pile-up jets, is evaluated by using alternative
correction factors, shifted by the corresponding uncertainties. This variation has two components, each of which shifts the correction factors in opposite directions. The resulting uncertainty is very small, with values below 0.1%.

**Unfolding-related uncertainties:**

- The effects of the mismodelling of the data by the MC simulation on the unfolding is accounted for as an additional source of uncertainty. This is assessed by reweighting the particle-level distributions so that the detector-level fragmentation variables predicted by the MC samples match those in the data. The modified detector-level distributions are then unfolded using the method described in section 6. The difference between the modified unfolded distribution and the reweighted particle-level distribution is taken as the uncertainty. The resulting uncertainty ranges from 0.2% to 10%, and typically increases with increasing values of $z$ and decreasing values of $p_{T}^{\text{rel}}$.

- The uncertainty due to using a particular MC model in the unfolding is estimated from the impact of two independent variations. The first accounts for the different amounts of gluon splitting in samples generated with different tunes. It is evaluated by comparing the results of the unfolding using the A14 samples with those using the different gluon splitting fractions in the Monash samples. These two particular tunes, which differ in the value of $\alpha_s$ for final-state radiation, are chosen because fits to the gluon splitting fraction in data are compatible with the uncertainty band spanned by them. The second is related to the description of the measured observables made by different MC models, after disentangling the effect of different gluon splitting fractions. It is evaluated by repeating the unfolding using two alternative samples. While the nominal result is obtained using the PYTHIA8 A14-RB sample, the alternative samples are simulated using PYTHIA8 A14 and SHERPA, in which the fragmentation is simulated by the string model. To avoid double counting of the differences arising from the description of the gluon splitting, the SHERPA and PYTHIA8 A14 samples are reweighted so that the fraction of the selected jets arising from gluon splittings $g \rightarrow b\bar{b}$ is the same as in the PYTHIA8 A14-RB sample. The envelope of the differences between the results using PYTHIA8 A14-RB and both alternative samples covers the second source of systematic uncertainty. The total modelling uncertainty is defined by the sum in quadrature of these two variations. This uncertainty ranges from 4% to 30% and it is, together with the JES uncertainty, dominant for these measurements.

**Pile-up-related uncertainties:**

- The uncertainty due to pile-up is derived from the ratios of the unfolded results for low ($\mu < 32$) and high ($\mu \geq 32$) pile-up subsamples to the results for the nominal measurement. The uncertainty is defined as the envelope of the two ratios, and is below 10% for both $z$ and $p_{T}^{\text{rel}}$.

---

4The value of the strong coupling constant for time-like parton showers in the A14 sample is $\alpha_s = 0.1260$, while the Monash tune uses $\alpha_s = 0.1365$. 
The total uncertainty is then determined as the sum in quadrature of all the systematic uncertainties described above. The total uncertainty is larger for low values of the jet $p_T$, where it can reach values up to 30% in some bins. For the higher $p_T$ bins, the uncertainty is systematically smaller, mainly because of the smaller jet energy scale components. Figure 5 shows the values of the systematic uncertainties as a function of $z$ and $p_T^{\text{rel}}$, for the lowest and highest bins of the jet $p_T$.

The dependence of the particle-level results on the choice of PDF was tested by using two alternative PDF sets for the SHERPA predictions with the string fragmentation model. Predictions using NNPDF3.0 and MMHT2014 were compared with the nominal prediction, which uses CT14 (see Table 1). A maximum deviation of 2% from the nominal results is observed. Moreover, the CT14 uncertainty, estimated using the full set of eigenvectors, produces a total variation of the order of a few per mille. The effect of varying the PDF is thus much smaller than the experimental uncertainties, and is therefore neglected.

8 Results

The particle-level results are presented and compared with the predictions described in Section 3. Figures 6 to 8 show the distributions of the longitudinal and transverse profiles for each $p_T$ bin.

The results show important differences between the low and high $p_T$ bins. In particular, the lower tails of the $z$ distributions contain a larger fraction of the data at high $p_T$, which translates into more high-$p_T^{\text{rel}}$ events at high $p_T$. This is understood to be due to the fact that gluon splittings, $g \rightarrow b\bar{b}$, occur with a larger probability at high values of the jet $p_T$. The resulting $b$-quarks are reconstructed in the same jet, while the $B^\pm$ meson originates from the fragmentation of one of them, thus leading to smaller values of $z$ and higher values of $p_T^{\text{rel}}$.

The results for the longitudinal profile show reasonable agreement with the HERWIG 7 prediction with the angle-ordered parton shower, while large discrepancies are observed with the dipole parton shower. In particular, the prediction largely overestimates the data in the low $z$ tails at low $p_T$, with smaller discrepancies for higher values of the jet $p_T$. This is understood to be a result of the larger fraction of jets arising from gluon splittings in the HERWIG 7 sample with the dipole parton shower.

The SHERPA predictions give a reasonable description of the $z$ distributions in the low and medium $p_T$ bins, although both predictions differ from data for very high values of $z$, particularly for the cluster model. At high $p_T$, the description worsens for the cluster model, which tends to underestimate the data at low $z$ and significantly overestimate the data at high $z$, while the Lund string model confirms the deviation from data at high $z$.

The PYTHIA8 samples tend to give descriptions that are qualitatively similar, and these are compatible with data within the systematic uncertainties across the different jet $p_T$ bins. However, the Monash+Peterson model overestimates the data at intermediate $z$ and low $p_T$, while it tends to conform with the rest at high $p_T$. The Monash+Lund-Bowler and A14+Lund-Bowler models both overestimate the data at very high values of $z$ in all $p_T$ ranges.
The discrepancies observed for the $z$ distributions have their counterparts in the $p_T^{\text{rel}}$ distributions as follows. Due to the larger gluon splitting fractions, the HERWIG 7 sample with the dipole parton shower significantly overestimates the data for $p_T^{\text{rel}}$ values between 1.5 and 4.0 GeV at low $p_T$, while the differences are smaller with increasing $p_T$. The HERWIG angle-ordered parton shower gives a better description of the $p_T^{\text{rel}}$ distributions, although non-negligible discrepancies are also observed.

The SHERPA predictions, particularly the one from the sample interfaced with the cluster hadronisation model, show large discrepancies for low values of $p_T^{\text{rel}}$, which increase when moving towards higher bins of jet $p_T$. The description of the $p_T^{\text{rel}}$ tails is within uncertainties at low $p_T$, but for higher $p_T$, both SHERPA samples tend to underestimate
the data for high $p_T^{rel}$. In general, the description from the sample interfaced to the Lund string model is observed to be better than the one provided by the cluster model.

All Pythia 8 samples give a good description of the $p_T^{rel}$ distributions across the different jet-$p_T$ bins.

In order to explicitly test the scale dependence of the longitudinal and transverse profiles, the average values of the longitudinal and transverse profiles, $\langle z \rangle$ and $\langle p_T^{rel} \rangle$, are studied as a function of the jet $p_T$. The results, together with the MC predictions, are shown in figure 9. All Pythia samples describe the scale dependence reasonably well for both $\langle z \rangle$ and $\langle p_T^{rel} \rangle$, although the samples using the A14 and A14-rb tunes predict slightly larger values of $\langle z \rangle$ (and slightly lower values of $\langle p_T^{rel} \rangle$) than measured in data. The Sherpa sample making use of the cluster hadronisation model fails to describe the $\langle p_T^{rel} \rangle$ data, disagreeing by 10% to 25%, but describes the $\langle z \rangle$ distribution reasonably well, except at high $p_T$. The Sherpa sample interfaced to the Lund string hadronisation model describes the $\langle z \rangle$ data well, while showing small discrepancies for $\langle p_T^{rel} \rangle$, although much smaller than when using the cluster model. The Herwig 7 sample making use of the angle-ordered parton shower describes the $\langle z \rangle$ scale dependence very well, while showing discrepancies of up to 15% for the $\langle p_T^{rel} \rangle$ data. Finally, the Herwig 7 sample implementing the dipole-based parton shower fails to describe both the $\langle z \rangle$ and $\langle p_T^{rel} \rangle$ profiles, showing discrepancies of up to 10% for the former, and up to 20% for the latter.

Some of the discrepancies described above can be traced to the modelling of the gluon splittings $g \rightarrow b\bar{b}$ in the parton shower algorithms. Jets in the MC simulation are considered to contain a gluon splitting into a $b\bar{b}$ pair if they contain two weakly decaying $B$ hadrons.

Figure 6. Distributions of the longitudinal profile $z$ and the transverse profile $p_T^{rel}$ for 50 GeV < $p_T$ < 70 GeV, together with different predictions from Pythia 8, Sherpa and Herwig 7. The vertical error bars represent the total experimental uncertainties. The lower panels show the ratios of MC predictions to the data, where the gray bands represent the total uncertainties.
Figure 7. Distributions of the longitudinal profile $z$ and the transverse profile $p_T^\text{rel}$ for $70 \text{ GeV} < p_T < 100 \text{ GeV}$, together with different predictions from PYTHIA 8, SHERPA and HERWIG 7. The vertical error bars represent the total experimental uncertainties. The lower panels show the ratios of MC predictions to the data, where the gray bands represent the total uncertainties.

Figure 8. Distributions of the longitudinal profile $z$ and the transverse profile $p_T^\text{rel}$ for $p_T > 100 \text{ GeV}$, together with different predictions from PYTHIA 8, SHERPA and HERWIG 7. The vertical error bars represent the total experimental uncertainties. In the first bin of the $p_T^\text{rel}$ distribution, the prediction by the SHERPA sample using the cluster model is outside the range of the ratio panel. The lower panels show the ratios of MC predictions to the data, where the gray bands represent the total uncertainties.
Figure 9. Average values of the longitudinal profile \(\langle z\rangle\) and of the transverse profile \(\langle p_{\text{rel}}^T\rangle\) as a function of the jet \(p_T\), compared with MC predictions by Pythia 8, Sherpa and Herwig 7. The points are in the bin centre. The vertical error bars represent the total experimental uncertainties. The lower panels show the ratios of MC predictions to the data, where the gray bands represent the total uncertainties.

The fragmentation properties of \(b\)-jets have been measured by using the decay of charged \(B\) mesons into \(J/\psi\) charmonia and charged kaons, using 139 fb\(^{-1}\) of \(pp\) collisions at \(\sqrt{s} = 13\) TeV recorded by the ATLAS detector at the LHC. To this end, the decay channel of the \(B\) mesons is fully reconstructed using a three-track fit to a common vertex. The
Figure 10. Fraction of jets arising from gluon splittings $g \to b\bar{b}$ as a function of the jet $p_T$ for different MC models.

$B^{\pm}$ candidates are matched to jets, reconstructed using the anti-$k_t$ algorithm with $R = 0.4$, and the longitudinal and transverse profiles of the $B$ mesons over the jet momentum are measured. The measurement spans three bins of the jet transverse momentum, and the yield of $B$ mesons is extracted for each bin of the fragmentation variable ($z$ or $p_T^{\text{rel}}$) and the jet $p_T$. The results are then corrected for detector resolution effects using an iterative Bayesian algorithm, and the systematic uncertainties due to the muon, jet and event reconstruction properties are evaluated.

The results are compared with different MC predictions, including PYTHIA 8, HERWIG 7 and SHERPA samples using different models for the parton shower and hadronisation. Generally, the best description of the longitudinal profile is provided by the PYTHIA 8 and SHERPA samples making use of the string hadronisation model, which provide similar descriptions for all values of the jet $p_T$. For medium and high values of the jet $p_T$, the SHERPA sample with cluster hadronisation shows large deviations from data at very high values of $z$. The HERWIG 7 sample with the angle-ordered parton shower also gives a good description of the longitudinal profile for all values of the jet $p_T$. In contrast, the HERWIG 7 samples using the dipole-based parton shower show large deviations from the data in the lower tails of the $z$ distribution. For the transverse profile, similar comments apply. While the PYTHIA 8 and SHERPA samples using the string hadronisation model give fair and similar descriptions of the data for all values of the jet $p_T$, HERWIG 7 fails to describe the data. The angle-ordered parton shower, however, provides a better description than the dipole-based parton shower. At medium and high $p_T$, the SHERPA sample with the cluster hadronisation model gives a poor description for low values of $p_T^{\text{rel}}$. 
This analysis provides key measurements which help to better understand the fragmentation functions of heavy quarks. As has been shown, significant differences among different MC models are observed, and also between the models and the data. Some of the discrepancies are understood to arise from poor modelling of the $g \to b\bar{b}$ splittings, to which the present analysis has substantial sensitivity. Including the present measurements in a future tune of the MC predictions may help to improve the description and reduce the theoretical uncertainties of processes where heavy-flavour quarks are present in the final state, such as top quark pair production or Higgs boson decays into heavy quark pairs.

Acknowledgments

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; CERN; ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRI, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; JINR; MES of Russia and NRC KI, Russian Federation; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DSI/NRF, South Africa; MICINN, 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, CANARIE, Compute Canada and CRC, Canada; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BŞF-NSF and GIF, Israel; Norwegian Financial Mechanism 2014-2021, Norway; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, 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 (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [86].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.
References


[6] CMS collaboration, Search for $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV, JHEP 03 (2019) 026 [arXiv:1804.03682] [inSPIRE].


Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern; Switzerland

School of Physics and Astronomy, University of Birmingham, Birmingham; United Kingdom

Facultad de Ciencias y Centro de Investigaciónes, Universidad Antonio Nariño, Bogotá; Colombia

Departamento de Física, Universidad Nacional de Colombia, Bogotá; Colombia

Dipartimento di Fisica e Astronomia A. Righi, Università di Bologna, Bologna; INFN Sezione di Bologna; Italy

Physikalisches Institut, Universität Bonn, Bonn; Germany

Department of Physics, Boston University, Boston MA; United States of America

Department of Physics, Brandeis University, Waltham MA; United States of America

Transilvania University of Brasov, Brăila; Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest; Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi; National Institute for Research and Development of Isotopic and Molecular Technologies, Bucharest; University Politehnica Bucharest; Bucharest; West University in Timisoara; Timișoara; Romania

Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava; Department of Nuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice; Slovak Republic

Physics Department, Brookhaven National Laboratory, Upton NY; United States of America

Departamento de Física (FCEN) and IFIBA, Universidad de Buenos Aires and CONICET, Buenos Aires; Argentina

California State University, CA; United States of America

 Cavendish Laboratory, University of Cambridge, Cambridge; United Kingdom

Department of Physics, University of Cape Town, Cape Town; iThemba Labs, Western Cape; Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg; National Institute of Physics, University of the Philippines Diliman (Philippines); University of South Africa, Department of Physics, Pretoria; University of Zululand; Kwalsanga; School of Physics, University of the Witwatersrand, Johannesburg; South Africa

Department of Physics, Carleton University, Ottawa ON; Canada

Faculté des Sciences Ain Chock, Université de Casablanca; Faculté des Sciences, Université Ibn-Tofail, Kénitra; Faculté des Sciences Semlalia, Université Mohammed V, Rabat; Mohammed VI Polytechnic University, Ben Guerir; Morocco

CERN, Geneva; Switzerland

Enrico Fermi Institute, University of Chicago, Chicago IL; United States of America

LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand; France

Nevis Laboratory, Columbia University, Irvington NY; United States of America

Niels Bohr Institute, University of Copenhagen, Copenhagen; Denmark

Dipartimento di Fisica, Università della Calabria, Rende; INFN Gruppo Collegato di Cosenza; Laboratori Nazionali di Frascati, Italy

Physics Department, Southern Methodist University, Dallas TX; United States of America

Physics Department, University of Texas at Dallas, Richardson TX; United States of America

National Centre for Scientific Research ‘Demokritos’, Agia Paraskevi; Greece

Department of Physics, Stockholm University; Oskar Klein Centre, Stockholm; Sweden

Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen; Germany

Lehrstuhl für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund; Germany

Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden; Germany

Department of Physics, Duke University, Durham NC; United States of America

SUPA — School of Physics and Astronomy, University of Edinburgh, Edinburgh; United Kingdom

INFN e Laboratori Nazionali di Frascati, Frascati; Italy

Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg; Germany
Ochanomizu University, Otsuka, Bunkyo-ku, Tokyo; Japan
Ohio State University, Columbus OH; United States of America
Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman OK; United States of America
Department of Physics, Oklahoma State University, Stillwater OK; United States of America
Palacký University, Joint Laboratory of Optics, Olomouc; Czech Republic
Institute for Fundamental Science, University of Oregon, Eugene, OR; United States of America
Graduate School of Science, Osaka University, Osaka; Japan
Department of Physics, University of Oslo, Oslo; Norway
Department of Physics, University of Pennsylvania, Philadelphia PA; United States of America
Konstantinov Nuclear Physics Institute of National Research Centre "Kurchatov Institute", PNPI, St. Petersburg; Russia
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh PA; United States of America
Laboratório de Instrumentação e Física Experimental de Partículas — LIP(a), Lisboa; Departamento de Física(b), Faculdade de Ciências, Universidade de Lisboa, Lisboa; Departamento de Física(c), Universidade de Coimbra, Coimbra; Centro de Física Nuclear da Universidade de Lisboa(d), Lisboa; Departamento de Física Teórica y del Cosmos(e), Universidad de Granada, Granada (Spain); Dep Física and CEFITEC de Faculdade de Ciências e Tecnologia(f), Universidade Nova de Lisboa, Caparica; Instituto Superior Técnico(g), Universidade de Lisboa, Lisboa; Portugal
Institute of Physics of the Czech Academy of Sciences, Prague; Czech Republic
Charles University, Faculty of Mathematics and Physics, Prague; Czech Republic
Particle Physics Department, Rutherford Appleton Laboratory, Didcot; United Kingdom
IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette; France
Santa Cruz Institute for Particle Physics, University of California Santa Cruz, Santa Cruz CA; United States of America
Departamento de Física(a), Pontificia Universidad Católica de Chile, Santiago; Millennium Institute for Subatomic physics at high energy frontier (SAPHIR)(b), Santiago; Universidad de la Serena(c), La Serena; Universidad Andres Bello(d), Department of Physics, Santiago; Instituto de Alta Investigación(e), Universidad de Tarapacá, Arica; Departamento de Física(f), Universidad Técnica Federico Santa María, Valparaíso; Chile
Universidade Federal de São João del Rei (UFSJ), São João del Rei; Brazil
Department of Physics, University of Washington, Seattle WA; United States of America
Department of Physics and Astronomy, University of Sheffield, Sheffield; United Kingdom
Department of Physics, Shinshu University, Nagano; Japan
Department Physik, Universität Siegen, Siegen; Germany
Department of Physics, Simon Fraser University, Burnaby BC; Canada
SLAC National Accelerator Laboratory, Stanford CA; United States of America
Department of Physics, Royal Institute of Technology, Stockholm; Sweden
Departments of Physics and Astronomy, Stony Brook University, Stony Brook NY; United States of America
Department of Physics and Astronomy, University of Sussex, Brighton; United Kingdom
School of Physics, University of Sydney, Sydney; Australia
Institute of Physics, Academia Sinica, Taipei; Taiwan
E. Andronikashvili Institute of Physics(a), Iv. Javakhishvili Tbilisi State University, Tbilisi; High Energy Physics Institute(b), Tbilisi State University, Tbilisi; Georgia
Department of Physics, Technion, Israel Institute of Technology, Haifa; Israel
Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv;
Israel
Department of Physics, Aristotle University of Thessaloniki, Thessaloniki; Greece
International Center for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo; Japan

Department of Physics, Tokyo Institute of Technology, Tokyo; Japan
Tomsk State University, Tomsk; Russia
Department of Physics, University of Toronto, Toronto ON; Canada
TRIUMF(a), Vancouver BC; Department of Physics and Astronomy(b), York University, Toronto ON; Canada
Division of Physics and Tomonaga Center for the History of the Universe, Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba; Japan
Department of Physics and Astronomy, Tufts University, Medford MA; United States of America
Department of Physics and Astronomy, University of California Irvine, Irvine CA; United States of America
Department of Physics and Astronomy, University of Uppsala, Uppsala; Sweden
Department of Physics, University of Illinois, Urbana IL; United States of America
Instituto de Física Corpuscular (IFIC), Centro Mixto Universidad de Valencia — CSIC, Valencia; Spain
Department of Physics, University of British Columbia, Vancouver BC; Canada
Department of Physics and Astronomy, University of Victoria, Victoria BC; Canada
Fakultät für Physik und Astronomie, Julius-Maximilians-Universität Würzburg, Würzburg; Germany
Department of Physics, University of Warwick, Coventry; United Kingdom
Waseda University, Tokyo; Japan
Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot; Israel
Department of Physics, University of Wisconsin, Madison WI; United States of America
Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal; Germany
Department of Physics, Yale University, New Haven CT; United States of America

a Also at Borough of Manhattan Community College, City University of New York, New York NY; United States of America
b Also at Bruno Kessler Foundation, Trento; Italy
c Also at Center for High Energy Physics, Peking University; China
d Also at Centro Studi e Ricerche Enrico Fermi; Italy
e Also at CERN, Geneva; Switzerland
f Also at CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille; France
g Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève; Switzerland
h Also at Departament de Física de la Universitat Autònoma de Barcelona, Barcelona; Spain
i Also at Department of Financial and Management Engineering, University of the Aegean, Chios; Greece
j Also at Department of Physics and Astronomy, Michigan State University, East Lansing MI; United States of America
k Also at Department of Physics and Astronomy, University of Louisville, Louisville, KY; United States of America
l Also at Department of Physics, Ben Gurion University of the Negev, Beer Sheva; Israel
m Also at Department of Physics, California State University, East Bay; United States of America
n Also at Department of Physics, California State University, Fresno; United States of America
o Also at Department of Physics, California State University, Sacramento; United States of America
p Also at Department of Physics, King’s College London, London; United Kingdom
q Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg; Russia