Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

ATLAS Collaboration

DOI
10.1140/epjc/s10052-021-09402-3

Publication date
2021

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):
Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

ATLAS Collaboration

CERN, 1211 Geneva 23, Switzerland

Received: 6 July 2020 / Accepted: 30 June 2021 / Published online: 3 August 2021
© CERN for the benefit of the ATLAS collaboration 2021

Abstract Jet energy scale and resolution measurements with their associated uncertainties are reported for jets using 36–81 fb$^{-1}$ of proton–proton collision data with a centre-of-mass energy of $\sqrt{s} = 13$ TeV collected by the ATLAS detector at the LHC. Jets are reconstructed using two different input types: topo-clusters formed from energy deposits in calorimeter cells, as well as an algorithmic combination of charged-particle tracks with those topo-clusters, referred to as the ATLAS particle-flow reconstruction method. The anti-$k_t$ jet algorithm with radius parameter $R = 0.4$ is the primary jet definition used for both jet types. This result presents new jet energy scale and resolution measurements in the high pile-up conditions of late LHC Run 2 as well as a full calibration of particle-flow jets in ATLAS. Jets are initially calibrated using a sequence of simulation-based corrections. Next, several in situ techniques are employed to correct for differences between data and simulation and to measure the resolution of jets. The systematic uncertainties in the jet energy scale for central jets ($|\eta| < 1.2$) vary from 1% for a wide range of high-$p_T$ jets ($250 < p_T < 2000$ GeV), to 5% at very low $p_T$ (20 GeV) and 3.5% at very high $p_T$ ($> 2.5$ TeV). The relative jet energy resolution is measured and ranges from $(24 \pm 1.5\%)$ at 20 GeV to $(6 \pm 0.5\%)$ at 300 GeV.

Contents

1 Introduction ........................................... 1
2 The ATLAS detector ................................... 2
3 Data and Monte Carlo simulated samples .......... 3
4 Jet reconstruction .................................... 4
5 Jet energy scale calibration .......................... 5
  5.1 Simulation-based jet calibrations ................. 6
    5.1.1 Pile-up corrections .......................... 6
    5.1.2 Jet energy scale and $\eta$ calibration ...... 7
    5.1.3 Global sequential calibration ............... 9
  5.2 In situ jet calibrations .......................... 11
  5.2.1 Relative calibration measurement in $\eta$ using dijet events .......... 12
  5.2.2 Calibration measurement using $Z$+jet and $\gamma$+jet events ........... 15
  5.2.3 High-$p_T$ jet calibration using multijet balance .................. 18
  5.2.4 Pile-up and the in situ analyses ............ 19
  5.2.5 In situ combination ......................... 20
5.3 Systematic uncertainties .......................... 23
  5.3.1 Uncertainty correlations and reductions ...... 25
  5.3.2 Uncertainties for EMtopo and PFlow jets .... 26
6 Jet energy resolution .................................. 26
  6.1 Resolution measurement using dijet events ....... 27
  6.2 Noise measurement using random cones .......... 28
  6.3 Combination of in situ jet energy resolution .... 31
  6.4 Application of JER and its systematic uncertainties .............. 31
7 Conclusions ......................................... 32
References ............................................. 33

1 Introduction
The energetic proton–proton ($pp$) collisions produced by the Large Hadron Collider (LHC) yield final states that are predominantly characterized by jets, or collimated sprays of charged and neutral hadrons. Jets constitute an essential piece of the physics programme carried out using the ATLAS detector due to their presence in the signal processes being measured and searched for, the various background processes that hide those signals, and the additional activity due to simultaneous $pp$ collisions. Measurements of the energy scale and resolution of these complex objects, as well as their associated systematic uncertainties, are therefore essential both for precision measurements of the Standard Model (SM) and for sensitive searches for new physics beyond it. This paper presents the strategy used for the determination of the jet energy scale (JES) and resolution (JER) by the ATLAS experiment and its implementation as it pertains to the analysis of data from Run 2 of the LHC. Results for the JES and JER are presented using data collected dur-
ing 2015–2017, corresponding to integrated luminosities in the range 36–81 fb\(^{-1}\), depending on the analysis method and its goals. This publication focuses on calibrating jets reconstructed with the anti-\(k_t\) [1] algorithm with radius parameter \(R = 0.4\).

The ATLAS Collaboration has published previous calibrations and uncertainties of the energy scale and resolution for this jet definition with data taken in 2010 [2–4], 2011 [5], 2012 [6], and 2015 [7]. Additionally, some ATLAS publications have targeted different jet definitions. In particular, the Run 1 papers include dedicated calibrations\(^1\) of jets reconstructed with the anti-\(k_t\) algorithm with \(R = 0.6\) and \(R = 1.0\), and a dedicated in situ calibration of large-radius jets has also been completed in Run 2 data [9]. This publication extends and improves on previous calibrations of anti-\(k_t\) \(R = 0.4\) jets, taking full advantage of the larger dataset recorded over the period of 2015–2017. The significant increase in the number of proton collisions per bunch crossing in 2016 and 2017 data-taking leads to a correspondingly more difficult environment for jet reconstruction, and this result presents new jet energy scale and resolution measurements in these unique high pile-up conditions.

Section 2 describes the ATLAS detector, and Sect. 3 describes the recorded data and the Monte Carlo (MC) simulation samples used in this paper. Section 4 presents the inputs and algorithms used to reconstruct the jets. Section 5 and Sect. 6 present the methods used and the result of both the calibration and the resulting systematic uncertainties of the JES and the JER, respectively.

2 The ATLAS detector

The ATLAS detector [10] at the LHC covers nearly the entire solid angle around the collision point.\(^2\) 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. The inner-detector system (ID) is immersed in a 2 T axial magnetic field and provides charged-particle tracking in the range \(|\eta| < 2.5\).

\(^1\) Comparisons in Run 1 between \(R = 0.4\) and \(R = 0.6\) jets confirm the need for dedicated calibrations for different jet radii [8].

\(^2\) ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point in the centre of the detector. The positive \(x\)-axis is defined by the direction from the interaction point to the centre of the LHC ring, with the positive \(y\)-axis pointing upwards, while the beam direction defines the \(z\)-axis. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, \(\phi\) being the azimuthal angle around the \(z\)-axis. The pseudorapidity \(\eta\) is defined in terms of the polar angle \(\theta\) by \(\eta = -\ln\tan(\theta/2)\). Rapidity is defined as \(y = 0.5\ln((E + p_t)/(E - p_t))\), where \(E\) denotes the energy and \(p_t\) is the component of the momentum along the beam direction. The angular distance \(\Delta R\) is defined as \(\sqrt{(\Delta y)^2 + (\Delta \phi)^2}\).

The silicon pixel detector covers the vertex region and typically provides four measurements per track, with the innermost space-point provided by the insertable B-layer that was installed before Run 2 [11,12]. The pixel detector is followed by the silicon microstrip tracker, which usually yields eight measurements per track. The silicon-based 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 \(|\eta| < 3.2\), high-granularity lead/liquid-argon (LAr) calorimeters with both barrel and endcap sections provide electromagnetic calorimetry. An additional thin LAr presampler covers \(|\eta| < 1.8\), and is used to correct for energy loss in materials traversed by particles prior to reaching 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 cover the range 1.5 < \(|\eta| < 3.2\). The solid angle coverage between 3.2 < \(|\eta| < 4.9\) is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimized for electromagnetic and hadronic measurements respectively. Interfaces that exist between each of these components, in particular between the barrel and endcap regions, provide for space to route various services and infrastructure, such as electrical and fiber-optic cabling, cooling, and support structures. However, these so-called transition regions also create discontinuities in the response of the calorimeter to both charged and neutral particles due to energy absorption in the inactive materials and changes in the geometry of the active materials of the calorimeters. The calibrated response and resolution of the calorimeter must therefore either correct for these features, or account for them when establishing systematic uncertainties. Figure 1 shows the many components of the calorimeter system, with reference pseudorapidities and various relevant transition regions marked as well [10,13,14].

The muon spectrometer (MS) comprises separate trigger and high-performance tracking chambers measuring the deflection of muons in a magnetic field generated by superconducting air-core toroids. The field integral of the toroids ranges between 2.0 and 6.0 Tm 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, fol-
Fig. 1 Layout of the ATLAS calorimeters with pseudorapidity ($\eta$) values marked for reference. The inner detector systems can be seen in black-and-white in the center of the diagram; tracking is provided up to $\eta = 2.5$. The electromagnetic (EM) barrel and endcap calorimeters are shown in green. The EM barrel has consistent performance throughout, but has a seam in the construction at $\eta = 0$ which can impact jet energy resolution. The EM endcap has a precision region marked in darker green and an extended region in light green, and the transition from one to the other around $\eta \sim 2.5$ involves a dramatic change in the material layers. The hadronic Tile calorimeter is shown in light blue while the hadronic endcap calorimeters based on liquid argon are illustrated in light orange. The forward calorimeters are shown in dark orange. Pink filled regions represent the tile plug calorimeter, often referred to as TileGap1 and TileGap2. The thin hot pink line marks the location of the very narrow gap and cryostat scintillators (TileGap3). The regions corresponding to the transition from barrel to endcap ($\eta \sim 1.4$) and from endcap to forward calorimeter ($\eta \sim 3.1$) are given for reference.

followed by selections made by algorithms implemented in software in the high-level trigger [15]. The first-level trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the high-level trigger reduces in order to record events to disk at about 1 kHz.

3 Data and Monte Carlo simulated samples

The data used for the measurements presented here were collected in $pp$ collisions at the LHC with a centre-of-mass energy of 13 TeV and a 25 ns proton bunch crossing interval during 2015–2017. The integrated luminosities of the datasets used are in the range 36–81 fb$^{-1}$ after requiring that all detector subsystems were operational during data recording.

Additional $pp$ collisions in the same and nearby bunch crossings are referred to as pile-up. The number of reconstructed primary vertices ($N_PV$) and the mean number of interactions per bunch crossing ($\mu$) are optimal observables to quantify the level of pile-up activity. The average value of $\mu$ is 13.7, 24.9, and 37.8 in the 2015, 2016, and 2017 datasets, respectively [16]. As described below, these conditions are accounted for in the production and reconstruction of simulated data.

Simulated dijet, multijet, $Z$+jet, and $\gamma$+jet samples are used in determining the jet energy scale and its uncertainties. Table 1 summarizes the MC generators, adjustable sets of parameters (tunes), and parton distribution function (PDF) sets used for all nominal and alternative samples of the various simulated processes. The nominal samples for the majority of analyses were generated with PYTHIA 8.186 [17] (from now on referred to as PYTHIA8) or Powheg+PYTHIA8.186 [17,20,21]. The multijet balance analysis uses SHERPA 2.1.1 [22] as the nominal generator since it incorporates up to three jets in the matrix element and is thus more suitable for multijet processes that have more than two jets in the final state. The dijet, multijet, and $\gamma$+jet nominal samples use the NNPDF2.3LO PDF set [19] and the A14 set of tuned parameters [18]. For the $Z$+jet analysis, the dedicated AZNLO tune [26] is used instead. Alternative samples for defining systematic variations use various generators and tunes.

Stable particles, defined as those with $c\tau > 10$ mm, output by the generators were passed through the GEANT4-based
simulations of the ATLAS detector [27,28]. This step simulates the interactions of the particles with matter in the detector and generates outputs which can be reconstructed in the same way as data. Hadronic showers were simulated using the FTFP BERT model as described in Ref. [29]. A set of simulated dijet events using the less detailed Atlassoft-II (AFII) are also studied to determine the difference in performance between full and fast simulation and provide appropriate calibrations for AFII samples in analyses [27].

Pile-up is incorporated in the MC samples by overlaying simulated inelastic interactions on the generated hard-scatter interaction. The inelastic interactions were simulated in Pythia 8.210 using the A3 tune and the NNPDF2.3LO PDF set [19,30]. To determine the number of simulated pp collisions to overlay onto a particular hard-scattering process, a random value is drawn from a Poisson distribution of the number of pp collisions per bunch crossing with a mean given by the desired average number of collisions per crossing for a particular data period. Events simulated with a particular pile-up profile are then compared with data from the corresponding data period. One set of MC samples was created using the pile-up profile of 2015 + 2016 data (average number of collisions 23.7) while a second independent set of samples used the profile of 2017 data. When data and simulation are compared in this paper, both sets of MC samples are used unless otherwise specified and are normalized to the luminosity of 2015+2016 data and 2017 data separately.

### 4 Jet reconstruction

The primary jet definition used in the majority of physics analyses by the ATLAS Collaboration and in the studies presented here is the anti-
$k_t$[1] algorithm with a radius parameter $R = 0.4$ as implemented in the FastJet 3.2.2 [31,32] software package. Four-vector objects are used as inputs to the algorithm, and may be stable particles defined by MC generators, charged-particle tracks, calorimeter energy deposits, or algorithmic combinations of the latter two, as in the case of the particle-flow reconstruction technique [33].

For use in jet reconstruction, calorimeter cells are first clustered into three-dimensional, massless, topological clusters (topo-clusters) using a nearest-neighbour algorithm [34]. Cells are added to a topo-cluster according to the ratio of the cell energy to the expected noise in each cell using thresholds that control the growth of each topo-cluster. The resulting energy of the topo-cluster is defined at the electromagnetic (EM) scale, which is the baseline calorimeter scale that correctly measures energy depositions from electromagnetic showers. Only positive-energy topo-clusters are used as inputs to the jet reconstruction. A jet produced in the hard-scatter process is expected to originate from the primary vertex, defined as the reconstructed vertex with at least two associated tracks and the largest sum of squared track momenta. Therefore, an event-by-event correction to account for the position of the primary vertex in each event – referred to as an origin correction – is applied to every topo-cluster, based on its depth within the calorimeter and pseudorapidity. This method is to be contrasted with earlier approaches [7] that applied this correction only to the jet four-momentum rather than to its constituents.

Jets reconstructed using only calorimeter-based energy information use the origin-corrected EM scale topo-clusters and are referred to as EMtopo jets. This was the primary jet definition used in ATLAS physics results prior to the end of Run 2. EMtopo jets exhibit robust energy scale and resolution characteristics across a wide kinematic range, and are independent of other reconstruction algorithms such as tracking at the jet-building stage.

Hadronic final-state measurements can be improved by making more complete use of the information from both the tracking and calorimeter systems. The particle flow (PFlow) algorithm is based on Ref. [33] and updated as described below. Particle flow directly combines measurements from both the tracker and the calorimeter to form the input signals for jet reconstruction, which are intended to approxi-
mate individual particles. Specifically, energy deposited in the calorimeter by charged particles is subtracted from the observed topo-clusters and replaced by the momenta of tracks that are matched to those topo-clusters. These resulting PFlow jets exhibit improved energy and angular resolution, reconstruction efficiency, and pile-up stability compared to calorimeter jets [33]. EMtopo and PFlow jets are retained for the analyses discussed in this paper only if they have an uncalibrated $p_T > 7$ GeV and $|\eta| < 4.5$.

The updates to the PFlow algorithm since its description in Ref. [33] as are follows. The expected mean value of the energy deposited by pions, $\langle E_{\text{dep}} \rangle$, and its expected standard deviation, $\sigma (E_{\text{dep}})$, were recomputed using the updated simulation, geometry, and topo-cluster noise thresholds for Run 2 [7]. The shower profiles were similarly updated. The only algorithmic change was an improvement in the transition between using track energy and cluster energy in high-$p_T$ jets. Since energetic particles are often in the core of jets and thus poorly isolated from nearby activity, accurate removal of the calorimeter energy associated with the track can be difficult. Therefore, the PFlow algorithm prevents energy subtraction in these cases. Formerly this was managed by applying a simple $p_T^{\text{trk}} < 40$ GeV cut in the track selection. In the updated algorithm, a more sophisticated procedure is used to prevent the subtraction in cases where the advantages of the tracker are smaller and where the particle shower falls used to prevent the subtraction in cases where the advantages of the tracker are smaller and where the particle shower falls.

In the updated algorithm, a more sophisticated procedure is applied only when no subtraction has been performed for a selected track. In this case the PFlow object includes both the full topo-cluster energy and the track momentum. To avoid double-counting the energy while maintaining the contribution from the calorimeter measurement, the track momentum is scaled by a factor $(1 - \langle E_{\text{dep}} \rangle / p_{\text{trk}})$. The resulting PFlow object uses the desired information and has a final energy of approximately $p_{\text{trk}}$, matching the response for the subtracted case. The second scale factor is applied in both the subtracted and non-subtracted cases for all PFlow objects created from selected tracks below 100 GeV. It smooths the transition between the lower-energy PFlow objects which are at the scale of the tracks and the higher-energy objects at the electromagnetic scale of the clusters. The energy of these PFlow objects is scaled by unity for $p_T^{\text{trk}}$ below 30 GeV, by $(1 - \langle E_{\text{dep}} \rangle / p_{\text{trk}})$ for objects with 60 GeV < $p_T^{\text{trk}}$ < 100 GeV, and by a linearly descending scale factor in between. This ensures that all objects are at the electromagnetic scale by 60 GeV.

Tracks used in PFlow objects and in deriving calibrations for both EMtopo and PFlow jets are reconstructed within the full acceptance of the inner detector ($|\eta| < 2.5$), required to have a $p_T > 500$ MeV, and satisfy quality criteria based on the number of hits in the ID subdetectors. To suppress the effects of pile-up, tracks must satisfy $|z_0 \sin \theta| < 2$ mm, where $z_0$ is the distance of closest approach of the track to the hard-scatter primary vertex along the $z$-axis and $\theta$ is the polar angle. Tracks are matched to jets using ghost association [35], a procedure that treats them as four-vectors of infinitesimal magnitude during the jet reconstruction and assigns them to the jet with which they are clustered.

MC simulation is used to determine the energy scale and resolution of jets by comparing PFlow and EMtopo jets with particle-level truth jets. Truth jets are reconstructed using stable final-state particles and exclude muons, neutrinos, and particles from pile-up interactions. Truth jets are selected with the same $p_T > 7$ GeV and $|\eta| < 4.5$ thresholds as EMtopo and PFlow jets, and are geometrically matched to those jets using the angular distance $\Delta R$ with the requirement $\Delta R < 0.3$.

5 Jet energy scale calibration

The jet energy scale calibration restores the jet energy to that of jets reconstructed at the particle level. The full chain
of corrections is illustrated in Fig. 2. All stages correct the four-momentum, scaling the jet $p_T$, energy, and mass.

At the beginning of the chain, the pile-up corrections remove the excess energy due to additional proton–proton interactions within the same (in-time) or nearby (out-of-time) bunch crossings. These corrections consist of two components: a correction based on the jet area and transverse momentum density of the event, and a residual correction derived from MC simulation and parameterized as a function of the mean number of interactions per bunch crossing ($\mu$) and the number of reconstructed primary vertices in the event ($N_{PV}$). These corrections are discussed in Sect. 5.1.1. The absolute JES calibration corrects the jet so that it agrees in energy and direction with truth jets from dijet MC events, and is detailed in Sect. 5.1.2. Furthermore, the global sequential calibration (derived from dijet MC events) improves the jet $p_T$ resolution and associated uncertainties by removing the dependence of the reconstructed jet response on observables constructed using information from the tracking, calorimeter, and muon chamber detector systems, as introduced in Sect. 5.1.3. All these calibrations are applied to both data and MC simulation. Finally, a residual in situ calibration is applied to data only to correct for remaining differences between data and MC simulation. It is derived using well-measured reference objects, including photons, Z bosons, and calibrated jets, and for the first time benefits from a low-$p_T$ measurement using the missing-$E_T$ projection fraction method for better pile-up robustness. It is described in Sect. 5.2. The full treatment and reduction of the systematic uncertainties is discussed in Sect. 5.3.

5.1 Simulation-based jet calibrations

The derivation of the calibrations derived exclusively from MC simulation samples is described below.

5.1.1 Pile-up corrections

As a result of the increase of the topo-clustering $p_T$ thresholds (to suppress electronic and pile-up noise) and in the instantaneous luminosity, the contribution from pile-up to the JES in the 2015–2017 data-taking period differs from the one observed in 2015. The pile-up corrections are therefore evaluated using updated MC simulations of the software reconstruction and pile-up conditions. These corrections are derived using the same methods employed in 2015 [7] and are summarized in the following paragraphs.

First, a jet $p_T$-density-based subtraction of the per-event pile-up contribution to the jet $p_T$ is performed. The jet area $A$ is a measure of the susceptibility of the jet to pile-up and is calculated by determining the relative number of ghost particles associated with a jet after clustering. Next, the pile-up contribution is estimated from the median $p_T$ density, $\rho$, of jets in the $y$–$\phi$ plane, $\langle p_T/A \rangle$. The calculation of $\rho$ uses jets reconstructed using the $k_t$ algorithm [36] with radius parameter $R = 0.4$ from positive-energy topo-clusters with $|\eta| < 2$. The computation of $\rho$ in the central region of the detector gives a more meaningful measure of the pile-up activity than the median over the entire $\eta$ range, and this is because $\rho$ drops to nearly zero beyond $|\eta| \sim 2$. This drop is due to the lower occupancy in the forward region relative to the central region, which is a result of a coarser segmentation in the forward region. The $k_t$ algorithm is chosen due to its tendency to naturally reconstruct jets including an uniform soft background [35], while $\rho$ is used to reduce the bias from hard-scatter jets which populate the high-$p_T$ tails of the distribution. The distribution of $\rho$ in MC simulation for representative $N_{PV}$ values is shown in Fig. 3. The ratio of the $\rho$-subtracted jet $p_T$ to the uncorrected jet $p_T$ is applied as a scale factor to the jet four-momentum and does hence not affect its direction.

The $\rho$ calculation is derived from the central, lower-occupancy regions of the calorimeter and does not fully describe the pile-up sensitivity in the forward calorimeter...
The dependences of the $p_T$-density-based and residual corrections on $N_{\text{PV}}$ and $\mu$ as a function of $|\eta_{\text{det}}|$ for PFlow jets are shown in Fig. 4. The negative dependence on $\mu$ for out-of-time pile-up is a result of the liquid-argon calorimeter’s pulse shape, which is negative during the period shortly after registering a signal [37]. These corrections are similar to those derived for EMtopo jets, although the $N_{\text{PV}}$-dependent corrections for PFlow jets in the $|\eta_{\text{det}}| < 2.5$ region are reduced by about 60% relative to EMtopo due to the usage of tracks in the PFlow algorithm. For EMtopo jets, the shape of the residual corrections is comparable to that found in 2015 MC simulation, except in the forward region ($|\eta_{\text{det}}| > 2.5$), where it is found to be smaller by 0.1 GeV. This difference is primarily caused by higher topo-cluster noise thresholds used in the full Run 2 data.

Four systematic uncertainties are introduced to account for MC mis-modelling of $N_{\text{PV}}$, $\mu$, the $p_T$ topology, and the $p_T$ dependence of the residual pile-up corrections. The last of these is derived from the full logarithmic fits to $\alpha$ and $\beta$, as discussed previously. Two in situ methods are used to estimate uncertainties in the modelling of $N_{\text{PV}}$ and $\mu$. The first method uses jets reconstructed from tracks to provide a measure of the jet $p_T$ independent of pile-up. This is only used for $|\eta| < 2.1$. The second method exploits the $p_T$ balance between a reconstructed jet and a $Z$ boson and is used for $2.1 < |\eta| < 4.5$. These systematic uncertainties are described in more detail in Ref. [38]. Finally, the $\mu$ topology uncertainty accounts for the uncertainty in the underlying event’s contribution to $\rho$, and is discussed in detail in Sect. 5.2.4.

5.1.2 Jet energy scale and $\eta$ calibration

The absolute jet energy scale and $\eta$ calibrations correct the reconstructed jet four-momentum to the particle-level energy scale accounting for non-compensating calorimeter response, energy losses in passive material, out-of-cone effects and biases in the jet $\eta$ reconstruction. Such biases are primarily caused by the transition between different calorimeter technologies and sudden changes in calorimeter granularity. The calibration is derived for $R = 0.4$ anti-$k_t$ jets from a Pythia8 MC simulation of dijet events after the application of the pile-up corrections. Reconstructed jets are geometrically matched to truth jets within $\Delta R = 0.3$. In addition, reconstructed (truth) jets are required to have no other reconstructed (truth) jet of $p_T > 7$ GeV within $\Delta R = 0.6$ ($\Delta R = 1.0$).

The average jet energy response $R$, defined as the mean of a Gaussian fit to the core of the $E_{\text{rec}}/E_{\text{true}}$ distribution, is measured in $E_{\text{true}}$ and $\eta_{\text{det}}$ bins. The decision to calculate the response as a function of $E_{\text{true}}$ instead of $E_{\text{rec}}$ is motivated by the fact that for fixed $E_{\text{true}}$ ($E_{\text{rec}}$) bins the response distribution is (not) Gaussian. The average response is parameterized

\begin{equation}
\rho^{\text{corr}} = \rho^{\text{reco}} - \rho \times A - \alpha \times (N_{\text{PV}} - 1) - \beta \times \mu,
\end{equation}

where $\rho^{\text{reco}}$ refers to the $p_T$ of the reconstructed jet before any pile-up correction is applied. Reconstructed jets with $p_T > 7$ GeV are geometrically matched to truth jets within $\Delta R = 0.3$. The residual $p_T$ dependences on $N_{\text{PV}}$ ($\alpha$) and $\mu$ ($\beta$) are observed to be fairly linear and independent of one another. Independent linear fits are used to derive $\alpha$ and $\beta$ coefficients in bins of $p_T^{\text{true}}$ and $|\eta_{\text{det}}|$, where $p_T^{\text{true}}$ is the $p_T$ of the truth jet that matches the reconstructed jet. The jet $\eta$ pointing from the geometric centre of the detector, $\eta_{\text{det}}$, is used to remove any ambiguity as to which region of the detector is measuring the jet. Both the $\alpha$ and $\beta$ coefficients are seen to have a logarithmic dependence on $p_T^{\text{true}}$, and logarithmic fits are performed in the range $20$ GeV $< p_T^{\text{true}} < 200$ GeV for each bin of $|\eta_{\text{det}}|$. In each $|\eta_{\text{det}}|$ bin, the fitted values of the $\alpha$ and $\beta$ coefficients at $p_T^{\text{true}} = 25$ GeV are taken as their nominal values, reflecting their behaviour in the $p_T$ region where pile-up effects are most relevant. The differences between the logarithmic fits over the full $p_T^{\text{true}}$ range and the nominal fits are used for a $p_T$-dependent systematic uncertainty in the residual pile-up dependence. Finally, linear fits are performed to the binned coefficients as a function of $|\eta_{\text{det}}|$. This reduces the effects of statistical fluctuations and allows the $\alpha$ and $\beta$ coefficients to be smoothly sampled in $|\eta_{\text{det}}|$, particularly in regions of varying dependence.

Fig. 3  Per-event median $p_T$ density, $\rho$, at $N_{\text{PV}} = 15$ (solid), $N_{\text{PV}} = 25$ (long dashed), and $N_{\text{PV}} = 35$ (short dashed) for $37 < \mu < 38$ as found in MC simulation.
Fig. 4 Dependence of PFlow jet $p_T$ on (a) in-time pile-up ($N_{PV}$ averaged over $\mu$) and (b) out-of-time pile-up ($\mu$ averaged over $N_{PV}$) as a function of $|\eta_{det}|$ for $p_T^{true} = 25$ GeV. Errors are taken from the fit results and are too small to be visible on the scale of the plot.

Fig. 5 The average energy response as a function of reconstructed jet (a) $\eta_{det}$ and (b) energy $E^{reco}$. Each value is obtained from the corresponding parameterized function derived with the PYTHIA8 MC sample and only jets satisfying $p_T > 20$ GeV are shown.

As a function of $E^{reco}$ using a numerical inversion procedure, as detailed in Ref. [2], and the jet calibration factor is taken as the inverse of the average energy response. The response is higher for PFlow jets than for EMtopo jets at low energies since tracking information is used. The response for PFlow jets as a function of $E^{reco}$ ($\eta_{det}$) for representative $\eta_{det}$ ($E^{reco}$) bins is shown in Fig. 5.

After the JES calibration based on the results in Fig. 5 is applied, the response diverges from 1 by a maximum of about 5% (3%, 1%) at $p_T^{true} = 20$ (30, 50) GeV. This level of non-closure is observed across entire $\eta_{det}$ range. These small non-closures are seen for low-$p_T$ jets due to a slightly non-Gaussian energy response and jet reconstruction threshold effects, both of which impact the response fits. The closure in this result is an improvement with respect to the 2015 calibration and is thanks to advances in the fitting method and parameters.

A bias in the reconstructed jet $\eta$, defined as a significant deviation from zero in the signed difference between the reconstructed and truth jet $\eta$, denoted by $\eta^{reco}$ and $\eta^{true}$ respectively, is observed and shown in Fig. 6 as a function of $|\eta_{det}|$ for PFlow jets. The bias for EMtopo jets is similar, showing the same features. It is largest in jets that encompass two calorimeter regions with different energy
responses caused by changes in calorimeter geometry or technology. This artificially increases the energy of one side of the jet relative to the other, altering the reconstructed four-momentum. The barrel–endcap ($|\eta_{\text{det}}| \sim 1.4$) and endcap–forward ($|\eta_{\text{det}}| \sim 3.1$) transition regions can be clearly seen in Fig. 5a as susceptible to this effect. A second correction is therefore derived as the difference between the reconstructed and truth $\eta$ ($E_{\text{true}}$ and $E_{\text{reco}}$ respectively) parameterized as a function of $E_{\text{true}}$ and $|\eta_{\text{det}}|$ to remove such bias. A numerical inversion procedure is again used to derive corrections in $E_{\text{reco}}$ from $E_{\text{true}}$. This calibration only alters the jet $p_T$ and $\eta$, not the full four-momentum. EMtopo and PFlow jets calibrated with the full jet energy scale and $\eta$ calibration are considered to be at the EM+JES scale and PFlow+JES scale, respectively.

The absolute JES and $\eta$ calibrations are also derived for a PYTHIA8 MC sample using AFIJ. An additional systematic uncertainty is considered for these samples to account for a small non-closure in the calibration beyond $|\eta_{\text{det}}| \sim 3.2$, due to the approximate treatment of hadronic showers in the forward calorimeters. This uncertainty is below 0.5% for all central jets and is about 3% for a forward jet of $p_T = 20$ GeV, falling rapidly with increasing $p_T$.

### 5.1.3 Global sequential calibration

Even after the application of the previous jet calibrations (from now on referred to as MCJES), for a given ($p_T^{\text{true}}$, $|\eta_{\text{det}}|$) bin, the response can vary from jet to jet depending on the flavour and energy distribution of the constituent particles, their transverse distribution, and the fluctuations of the jet development in the calorimeter. Furthermore, the average particle composition and shower shape of a jet varies between initiating particles, most notably between quark- and gluon-initiated jets. A quark-initiated jet will often include hadrons with a higher fraction of the jet $p_T$ that penetrate further into the calorimeter, while a gluon-initiated jet will typically contain more particles of softer $p_T$, leading to a lower calorimeter response and a wider transverse profile. The global sequential calibration (GSC), a procedure used in the 2012 [6] and 2015 [7] calibrations, is a series of multiplicative corrections to reduce the effects from these fluctuations and improve the jet resolution without changing the average jet energy response. The jet resolution $s_R$ is given by the standard deviation of a Gaussian fit to the jet $p_T$ response distribution, where the $p_T$ response is defined similarly to jet energy response as the ratio of $p_T^{\text{reco}}$ to $p_T^{\text{true}}$.

The GSC is based on global jet observables such as the longitudinal structure of the energy depositions within the calorimeters, tracking information associated with the jet, and information related to the activity in the muon chambers behind a jet. For these studies, reconstructed jets are geometrically matched to truth jets and a numerical inversion procedure is used, as explained in Sect. 5.1.2. Six observables are identified that improve the resolution of the JES through the GSC. For each observable, an independent jet four-momentum correction is derived as a function of $p_T^{\text{true}}$ and $|\eta_{\text{det}}|$ by inverting the reconstructed jet response in PYTHIA8 MC simulation events. Corrections for each observable are applied independently and sequentially to the jet four-momentum for jets with $|\eta_{\text{det}}| < 3.5$ (unless stated otherwise). No improvement in resolution was found from altering the sequence of the corrections.

The six stages of the GSC account for the dependence of the jet response on (in the order in which they are applied):

- $f_{\text{charged}}$, the fraction of the jet $p_T$ measured from ghost-associated tracks with $p_T > 500$ MeV ($|\eta_{\text{det}}| < 2.5$);
- $f_{\text{Tile0}}$, the fraction of jet energy measured in the first layer of the hadronic Tile calorimeter ($|\eta_{\text{det}}| < 1.7$);
- $f_{\text{LAr3}}$, the fraction of jet energy measured in the third layer of the electromagnetic LAr calorimeter ($|\eta_{\text{det}}| < 3.5$);
- $n_{\text{trk}}$, the number of tracks with $p_T > 1$ GeV ghost-associated with the jet ($|\eta_{\text{det}}| < 2.5$);
- $w_{\text{trk}}$, also known as track width, the average $p_T$-weighted transverse distance in the $\eta$–$\phi$ plane between the jet axis and all tracks of $p_T > 1$ GeV ghost-associated with the jet ($|\eta_{\text{det}}| < 2.5$);
- $n_{\text{segments}}$, the number of muon track segments ghost-associated with the jet ($|\eta_{\text{det}}| < 2.7$).

The first correction is only applied to PFlow jets. The $n_{\text{segments}}$ correction, also known as the punch-through correction, reduces the tails of the response distribution caused by high-$p_T$ jets that are not fully contained in the calorimeter. All corrections are derived as a function of jet $p_T$, except for the punch-through correction, which is derived as a function of jet energy since this effect is more correlated with the energy escaping the calorimeters.

The underlying distributions of these observables are shown for PFlow jets in MC simulation and bins of equal statistics in Fig. 7. Each observable has been studied in data and simulation and is found to be well modelled [6,7,33]. The spike at zero in the $f_{\text{Tile0}}$ distribution at low $p_T^{\text{true}}$, shown in Fig. 7b, corresponds to jets that are fully contained in the electromagnetic calorimeter and do not deposit energy in the Tile calorimeter. The tail towards negative values in the $f_{\text{Tile0}}$ and $f_{\text{LAr3}}$ distributions at low $p_T^{\text{true}}$, shown in Fig. 7b, c, respectively, reflects calorimeter noise fluctuations. Slight differences with respect to data have a negligible impact on the GSC since the dependence of the average jet response on the observables is well modelled in MC simulation, as observed by an in situ dijet tag-and-probe method described in Ref. [2].

In this method, the average $p_T$ asymmetry between back-to-back jets is measured as a function of each observable.
Fig. 7  Jet response for PFlow jets in four broad $p_T$ ranges as a function of each of the six observables used in the GSC: a) the fraction of the jet $p_T$ carried by charged particles, b) the fraction of energy in the first layer of the Tile calorimeter, c) the fraction of energy in the third layer of the electromagnetic calorimeter, d) the number of tracks, e) the track width, and f) the number of muon spectrometer track segments associated with the jet. Jets at the PFlow+JES scale with $0.2 < |\eta| < 0.3$ (except for $n_{\text{segments}}$ which is shown for $|\eta| < 1.3$ due to low statistics) are selected from a sample of PYTHIA8 dijet MC events and the corresponding preceding GSC steps have been applied accordingly. The error bars show only the statistical uncertainty. The bottom panels show the normalized distributions of the variables.
The average jet $p_T$ response for PFlow jets in MC simulation as a function of each of the GSC observables is shown in Fig. 7 for representative $p_T^{\text{true}}$ ranges. The dependence of the jet response on each observable is reduced to less than 2% after the full GSC is applied, with small deviations from unity reflecting the correlations between observables that are unaccounted for in the corrections.

The fractional jet resolution, defined as $\sigma_R/R$, is used to determine the size of the fluctuations in the jet energy reconstruction and is shown for PFlow jets with $0.2 < |\eta_{\text{det}}| < 0.3$ in MC simulation in Fig. 8. As more corrections are applied, the fractional jet resolution improves and the jet response dependence on the jet flavour is reduced. No improvement is observed in Fig. 8 from the punch-through correction since only a small fraction of jets received this calibration, but there are analyses where their region of interest has a large fraction of jets that would receive this correction [39,40].

5.2 In situ jet calibrations

Once jets are corrected to the particle level using the MCJES and GSC, they require one final calibration step to account for differences between the jet response in data and simulation. These differences are caused by imperfect simulation of both the detector materials and the physics processes involved: the hard scatter and underlying event, jet formation, pile-up, and particle interactions with the detector. The final in situ calibration measures the jet response in data and MC simulation separately and uses the ratio as an additional correction in data.

Jet response is calculated by balancing the $p_T$ of a jet against that of a well-calibrated reference object or system. The response $R_{\text{in situ}}$ is defined as the average ratio of the jet $p_T$ to the reference object $p_T$ in bins of reference object $p_T$, where that average is taken from the peak location found by fitting the distribution with a Gaussian function. $R_{\text{in situ}}$ is sensitive to effects such as the presence of additional radiative jets or the transition of energy into or out of the jet cone, although these effects can be mitigated through careful event selection. A better method is to form the double ratio from the response in data and MC simulation:

$$c = \frac{R_{\text{data}}}{R_{\text{MC}}},$$

which is robust to secondary effects so long as they are well modelled in simulation and is therefore a reliable measure of the jet energy scale difference between data and MC simulation. The double ratio $c$ is transformed via numerical inver-

3 Requirements on the angle between the leading jet and the well-calibrated reference object, as well as cuts on the maximum $p_T$ of the second jet in the event, help suppress additional radiation to ensure the events are as clean as possible.

---

**Fig. 8** Resolution of jets at the PFlow+JES scale with $0.2 < |\eta_{\text{det}}| < 0.3$ measured in Pythia8 dijet MC simulation after each stage of the global sequential calibration (GSC). All jet flavours, including $b$-jets, are considered. The lower panel shows the difference in quadrature between the resolution before any GSC correction is applied ($\sigma$) and after the corresponding GSC step is applied ($\sigma'$) from a function of reference object $p_T$ to a function of jet $p_T$ (and jet $\eta$ where applicable). This is the final in situ calibration.

There are three stages of in situ analyses. First, the $\eta$ intercalibration analysis corrects the energy scale of forward ($0.8 \leq |\eta_{\text{det}}| < 4.5$) jets to match those of central ($|\eta_{\text{det}}| < 0.8$) jets using the $p_T$ balance in dijet events. Second, the $Z$+jet and $\gamma$+jet analyses balance the hadronic recoil in an event against the $p_T$ of a calibrated $Z$ boson or photon. The missing-$E_T$ projection fraction (MPF) method uses the full hadronic recoil instead of a jet to compute the balance to help mitigate effects of pile-up and jet reconstruction threshold which otherwise make low-$p_T$ measurements challenging [41]. Finally, the multijet balance (MJB) analysis uses a system of well-calibrated low-$p_T$ jets to calibrate a single high-$p_T$ jet [42]. The $Z/\gamma$+jet and MJB analyses are computed only for central jets, but are also applicable to forward jets due to the effect of the $\eta$ intercalibration. Each measurement is translated from a function of reference object $p_T$ into jet $p_T$. A statistical combination of the $Z/\gamma$+jet and MJB analyses provides a single smooth calibration applicable across the full momentum range.

Since the three in situ analyses ($\eta$ intercalibration, $Z/\gamma$+jet MPF, and MJB) are performed sequentially, systematic uncertainties are propagated from each to the next. Within each analysis, systematic uncertainties arise from three sources: modelling of physics processes in simulation, uncertainties in the measurement of the reference object, and uncertainties in the expected $p_T$ balance due to the event’s topology. Mis-modelling is accounted for by comparing the
predictions of two MC generators and taking their difference as the uncertainty. Systematic uncertainties in the measurement of the reference object are taken from the ±1σ uncertainties in each object’s calibration and propagated through the analysis. Event topology uncertainties are estimated by varying the event selections used and observing the impact on the final MC simulation to data ratio.

A rebinning procedure is applied to each systematic uncertainty to ensure that the features represented in the final result are statistically significant and not the result of fluctuations in small numbers of simulated or data events. This is only performed where the response does not vary sharply with \( p_T \), ensuring it does not obscure real physics effects. The rebinning procedure follows a bootstrapping method: pseudo-experiment datasets are created by sampling from a Poisson distribution with a mean of one for each event in the data or MC simulation [43]. The pseudo-experiments are therefore statistically correlated yet unique, and the root mean square distribution across the pseudo-experiments provides a measure of the statistical uncertainty of the analysis. The measured result for each systematic uncertainty is then rebinned as appropriate for each analysis to ensure that the final shapes are statistically significant.

The \( Z/\gamma + \text{jet} \) and MJB calibrations and uncertainties are derived from the full 2015–2017 combined datasets with a total luminosity of 80 fb\(^{-1}\). The \( \eta \) intercalibration analysis uses a dataset of total size 81 fb\(^{-1}\), but since this analysis is more sensitive than the others to year-by-year fluctuations, the dataset is split into two blocks and a time-dependent result is computed instead. One \( \eta \) intercalibration is derived from and applies to the 2015 + 2016 dataset while a second independent calibration is derived from the 2017 dataset and applies to 2017 + 2018 data. These two data periods are treated separately due to a change in LAr calorimeter read-out that occurred between 2016 and 2017 data taking and affected jet reconstruction in the endcap regions. With no changes of similar scale made between 2017 and 2018 data taking, the 2017 calibration can be reasonably applied to 2018 as well. The post-calibration jet performance is consistent between these two different data periods and therefore a single set of uncertainties based on the 2015 + 2016 dataset is used for the \( \eta \) intercalibration in all years, with only a small localized additional uncertainty added for 2018 as described in Sect. 5.2.1.

Certain common selection criteria are applied to all three in situ analyses. Each event must have a reconstructed vertex with at least two associated tracks of \( p_T > 500 \) MeV. All jets must satisfy quality criteria to reject non-collision background, calorimeter noise, and cosmic rays [44]. Furthermore, each jet with 20 GeV < \( p_T < 60 \) GeV and \( |\eta_{\text{det}}| < 2.4 \) must pass jet vertex tagging, or JVT, requirements with selection criteria that are specific to the jet definition [45]. These requirements match jets to the primary vertex and are 92% efficient for EMtopo jets and 97% efficient for PFlow jets.

### 5.2.1 Relative calibration measurement in \( \eta \) using dijet events

The \( \eta \) intercalibration analysis produces a correction which is applied to forward jets (0.8 ≤ \( |\eta_{\text{det}}| < 4.5 \)) to bring them to the same energy scale as central jets (|\( \eta_{\text{det}} \) | < 0.8). Jets in the central region of the detector are taken to be well-calibrated, while jets in the forward regions vary in response and must be corrected accordingly. Events are selected with exactly two jets in different \( \eta \) regions of the detector. To maximize statistics, neither jet need be in the central region: instead, all regions will be calibrated relative to one another.

For these dijet events, momentum balance requires that the transverse momentum of the two jets must be equal and opposite. Therefore, the momentum asymmetry of the two jets is a metric for the response difference between the two detector regions (left and right for simplicity):

\[
A = \frac{p_T^{\text{left}} - p_T^{\text{right}}}{p_T^{\text{avg}}},
\]

where \( p_T^{\text{avg}} = (p_T^{\text{left}} + p_T^{\text{right}})/2 \). The response ratio \( R \) of the two jets defines the calibration factor \( c \) for each jet and is then:

\[
R = \frac{c^{\text{left}}}{c^{\text{right}}} = \frac{2 + \langle A \rangle}{2 - \langle A \rangle} \approx \frac{p_T^{\text{left}}}{p_T^{\text{right}}},
\]

The average response ratio \( \langle R_{ijx} \rangle \) is measured in each bin of \( \eta^{\text{left}} \), \( \eta^{\text{right}} \), and \( x \) of \( p_T^{\text{avg}} \); \( \Delta (R_{ijx}) \) is the statistical uncertainty in each bin. All \( \eta \) values are in detector coordinates rather than corrected jet coordinates (\( \eta_{\text{det}} \)) since the properties of interest correlate to specific regions of detector hardware. The following function relates the correction factors and responses in each of the \( N \) bins:

\[
S(c_1, \ldots, c_N) = \sum_{j=2}^{N} \sum_{i=1}^{j-1} \frac{1}{\Delta(R_{ijx})} \left( \frac{1}{(R_{ijx})^2} - \frac{1}{c_{ix}} \right) + X(c_{ix})
\]

Here, the function \( X(c_{ix}) \) quadratically imposes a penalty on correction factors deviating from 1.\(^4\) Minimizing this function produces the correction factors to be used in the calibration.

Previous iterations of the jet energy scale have used a fit in MINUIT to minimize \( S(c_{ix}) \). The current calibration

\(^4\) This penalty function takes the form \( X(c_i) = \lambda \left( \frac{1}{x} \sum_{i=1}^{N} x_i - 1 \right)^2 \), where \( \lambda \) introduces the Lagrange multiplier visible in Eq. (2). The purpose of the penalty function is to ensure that the appropriate minimum is selected by suppressing local minima with large values of \( c_{ix} \), and as such its exact form is somewhat arbitrary.
instead minimizes the function analytically. Suppressing the $x$ indices for clarity and setting the derivative of $S$ with respect to some correction factor $c_x$ equal to zero, the following equation defines the correction factor values which minimise $S$:

$$
\sum_{i=1}^{N} \left( \frac{1}{\Delta x^2} \frac{(R_{i} - \bar{R}_{i})^2}{\bar{R}_{i}^2} + \lambda \right) c_i + \left( \sum_{i=1}^{N} \frac{1}{\Delta x^2} \frac{(R_{i} - \bar{R}_{i})^2}{\bar{R}_{i}^2} + \lambda \right) c_0 \\
+ \sum_{i=1}^{N} \left( \frac{1}{\Delta x^2} \frac{(R_{i} - \bar{R}_{i})^2}{\bar{R}_{i}^2} + \lambda \right) c_i = 0 .
$$

Equation (2) can then be expressed as a matrix system of linear equations. This matrix system is solved independently for each $p_T^{\text{avg}}$ bin $x$ to obtain values for the correction factors $c_{ix}$ for each $\eta_{\text{det}}$ bin $i$ in this momentum range. Solving analytically for the $c_{ix}$ in this way allows the result to be found approximately a thousand times more quickly than using a fit. This large reduction in computational requirements in turn allows the analysis to use a finer binning in $\eta_{\text{det}}$, capturing more details of the detector structure. The two methods agree well and each shows good closure when tested in simulation. Finally, the full set of correction factors are normalized such that the average correction factor in the central region $|\eta_{\text{det}}| < 0.8$ is unity.

Events are selected using a combination of single-jet triggers, with each trigger only considered in the jet $p_T$ range for which it is at least 99% efficient [15,46]. Events may pass either a central jet trigger or a forward jet trigger, or both. In the case that a trigger is prescaled, the passing event is weighted by the appropriate amount. Jets with $|\eta_{\text{det}}| < 2.4$ are also required to satisfy JVT criteria to minimize contributions from pile-up and must pass basic cleaning requirements [38,44]. Each selected event must have two jets with $p_T > 25$ GeV and $|\eta| < 4.5$. To ensure a clean dijet topology, events are further required to have no third jet with significant $p_T$: $p_T^{\text{third}}/p_T^{\text{avg}} < 0.25$, where $p_T^{\text{avg}}$ is the average momentum of the two leading jets. The two leading jets are required to be back-to-back in the azimuthal plane such that $\Delta \phi > 2.5$ rad.

Like the other in situ analyses, the goal of the $\eta$ intercalibration is to correct for data–simulation differences, so the quantity of interest is the ratio of the measured calorimeter response in MC simulation to the response in data. The nominal calibration is derived by comparison with POWHEG+PYTHIA8 simulated events. The analysis binning in $p_T^{\text{avg}}$ and $\eta_{\text{det}}$ is selected to balance the requirements of both sufficient statistics in sparse regions and resolution of narrow detector features. As such, it varies for different values of $\eta_{\text{det}}$. Remaining statistical fluctuations in the final calibration are smoothed using a two-dimensional Gaussian kernel with parameters selected to preserve significant structures.

Figure 9 shows the measured response in data and POWHEG+PYTHIA8 MC simulation for the 2017 dataset as a function of $\eta_{\text{det}}$ for three different $p_T^{\text{avg}}$ ranges (Fig. 9a–c) and as a function of $p_T^{\text{avg}}$ for three different $\eta_{\text{det}}$ ranges (Fig. 9d–f). The simulation can be seen to approximately reproduce the $\eta_{\text{det}}$-dependent features of the response observed in data, although the response in data is consistently higher than the response in simulation. The simulation/data response ratio as directly measured is shown in discrete points in the bottom panel, while the calibration derived from smoothing the response ratio is overlaid as the solid curve. The dashed curve shows the extrapolation to $p_T$ ranges beyond the available data, taken from the Gaussian smoothing results. Since the smoothing is stronger in the $p_T$ direction and weaker in $\eta_{\text{det}}$ to preserve detector features, this sets each extrapolated value to approximately the value of the last populated bin at lower $p_T$. Above $p_T = 2$ TeV the value is kept constant.

Uncertainties are derived as a function of $\eta_{\text{det}}$ and $p_T$ and account for mis-modelling of physics, detector, and event topology effects on the momentum balance of dijet events. The dominant uncertainty is in MC mis-modelling and is taken to be the difference between the smoothed calibration curves derived from the POWHEG+PYTHIA8 and SHERPA dijet samples. Additional uncertainties in the physics and topology modelling are assessed by varying the $p_T^{\text{third}}$, $\Delta \phi$, and pile-up suppression cuts and using a bootstrapping method to ensure observed shapes are statistically significant as discussed in Sect. 5.2. Similarly, the JVT uncertainty is determined by comparison with tighter and looser working points. These uncertainties can take positive or negative values. The statistical uncertainty is strictly positive and is taken from the data and MC simulation sample sizes. Finally, a non-closure uncertainty is assessed by comparing the response in data with that in POWHEG+PYTHIA8 after applying the derived $\eta$ intercalibration. This uncertainty is largest for $|\eta_{\text{det}}| \sim 2.1–2.6$, where detector transitions make modelling of the LAr pulse shape particularly difficult [34], and for jets near the kinematic limit, where they have the maximum possible $p_T$ for a given $\eta_{\text{det}}$ subject to the constraint of a 13 TeV centre-of-mass energy. The non-closure uncertainty is treated as three independent nuisance parameters, two covering the regions around $\pm 2.4$ in $\eta_{\text{det}}$ and one at the kinematic limit, since these two non-closure uncertainties are uncorrelated.

After being corrected each with their dedicated calibration, the 2015+2016 and 2017 datasets are in good agreement, and therefore a single set of uncertainties is sufficient to cover both cases. The uncertainties calculated with the 2015+2016 dataset are selected for this role. The only dataset-dependent uncertainty is an additional small non-closure uncertainty used for 2018 data only. It covers the region around $\eta = \pm 1.5$ to account for the difference in Tile calorimeter calibration.
Fig. 9 Relative response of jets calibrated with PFlow+JES in data (black circles) and POWHEG+PYTHIA8 MC simulation (red squares). Response is shown as a function of $\eta_{\text{det}}$ for jets of (a) $40 \text{ GeV} < p_{\text{jet}}^T < 60 \text{ GeV}$, (b) $85 \text{ GeV} < p_{\text{jet}}^T < 115 \text{ GeV}$, and (c) $270 \text{ GeV} < p_{\text{jet}}^T < 330 \text{ GeV}$, and as a function of $p_T$ for jets of (d) $1.2 < \eta_{\text{det}} < 1.4$, (e) $2.6 < \eta_{\text{det}} < 2.8$, and (f) $3.0 < \eta_{\text{det}} < 3.2$. The lower panel shows the response ratio of simulation to data (red squares) as well as the smoothed in situ calibration factor derived from the ratio (solid curve) which is used to perform the $\eta$ intercalibration. Dotted lines show the extrapolation of the in situ calibration to the regions without data points. The dashed red and blue horizontal lines provide reference points for the viewer during this year of data-taking and has a maximum size of 2%.

The method uncertainties are shown in Fig. 10. Three illustrative $p_T$ values are selected. The uncertainties decrease slightly as a function of $p_T$ and increase significantly as a function of $\eta_{\text{det}}$ outside of the central detector region, while in the central region they are zero by construction. For practical use the various systematic uncertainty terms are summed in quadrature to produce one single systematic uncertainty dominated by the modelling term. In the cases where up and down variations differ, the largest absolute value of the two is used at each point. The total systematic uncertainty and the statistical uncertainty are both symmetrized in $\eta_{\text{det}}$. The non-closure uncertainties, not included in Fig. 10 as they are not method uncertainties, are instead shown in Fig. 22 where it can be seen that they are kept asymmetric to reflect real differences in the detector.
Fig. 10 Systematic uncertainties associated with the $\eta$ intercalibration procedure as a function of $\eta_{\text{det}}$ for (a) $p_T = 50$ GeV, (b) $p_T = 100$ GeV, and (c) $p_T = 300$ GeV. The solid purple band shows the total systematic uncertainty, while the grey band shows the statistical uncertainty alone. Individual sources of uncertainty are marked by coloured lines. These have been smoothed to remove the impact of statistical fluctuations. Thus the visible shapes are statistically significant. The MC modelling term is the dominant source of uncertainty.

The calibrations are similar in size and shape between PFlow and EMtopo jets. Systematic uncertainties are also similar in size and shape since the dominant MC modelling component does not differ meaningfully between the two jet collections.

5.2.2 Calibration measurement using $Z$+jet and $\gamma$+jet events

The next stage of the in situ calibration corrects for the differences between data and MC simulation using the momentum balance between the measured hadronic activity in the event and the $p_T$ of a well-calibrated photon or $Z$ boson. Only the central region of the detector ($|\eta| < 0.8$) is used for this analysis: the $\eta$ intercalibration ensures that a correction derived centrally translates directly to forward jets as well.

The $Z/\gamma$+jet analyses rely on the energy scale of the photon or the electrons and muons from the $Z$ decay being well measured. All three objects are cleanly measured in the ATLAS detector and the uncertainties in their energy scales are small [47,48]. The response is calculated separately in $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ events since the sources of uncertainties propagated from $e$ and $\mu$ calibration are independent, and the three channels are combined at a later stage.

The $Z$+jet response measurement is limited at moderate to high $p_T$ by low statistics and thus covers a range in jet $p_T$ from 17 GeV to 1 TeV with large uncertainties in the final bin. The $\gamma$+jet response measurement benefits from much higher statistics and extends to 1.2 TeV with little loss in sensitivity. However, it is limited at low jet $p_T$ by both the trigger prescales and the prevalence of soft jets misidentified as photons and so begins at 25 GeV.

The missing-$E_T$ projection fraction technique is used for both of the $Z/\gamma$+jet analyses and balances the reference object $p_T$ against the full hadronic recoil in an event. By doing so, it is possible to compute the calorimeter response to hadronic showers directly. This approach is robust to both pile-up and the underlying event, which each cancel out directionally on average over a large collection of events, and is not strongly affected by jet definitions since these become relevant only in the application of the calibration. The showering and topology effects in moving from a recoil-level quantity to a jet-level quantity are studied and found to be small, as discussed below. Taking $\vec{p}_T^{\text{recoil}}$ as the total transverse momentum of the hadronic activity in a clean $Z/\gamma$+jet event and $p_T^{\text{jet}}$ as the transverse momentum of the photon or $Z$ boson, conservation of transverse momentum means that at the particle level:

\[ \vec{p}_T^{\text{recoil}} + p_T^{\text{jet}} = \vec{p}_T. \]
The balance could be altered by the presence of initial- or final-state radiation. To suppress the effects of such additional radiation, a cut is placed on the azimuthal angle $\Delta \phi$ between the jet and the reconstructed photon or Z boson in the event and an uncertainty due to the topology is evaluated by varying the event selection requirements. If the calorimeter response to the hadronic activity in this event is $r_{\text{MPF}}$ and the response for the calibrated reference object is 1, and assuming any missing energy in the event is due to the low response to the hadronic recoil ($r_{\text{MPF}} < 1$), then at the detector level Eq. (3) becomes:

$$\vec{p}_{\text{T,truth}}^{\text{ref}} + r_{\text{MPF}} \vec{p}_{\text{T,truth}}^{\text{recoil}} = -\vec{E}_{\text{T}}^{\text{miss}}$$

After taking the projection of each term in the direction of the reference object, defined by a unit vector $\hat{n}_{\text{ref}}$, the response to the hadronic recoil is then seen to depend only on the missing energy in the event and the momentum of the reference object. The MPF response $R_{\text{MPF}}$ is defined by measuring the average of $r_{\text{MPF}}$ across events binned in the reference object $p_T$. Thus,

$$R_{\text{MPF}} = \left(1 + \frac{\hat{n}_{\text{ref}} \cdot \vec{E}_{\text{T}}^{\text{miss}}}{\vec{p}_{\text{T}}^{\text{ref}}} \right).$$

This peak location is taken to be the average response in that bin, and the response is mapped from reference to jet $p_T$ by finding the average jet $p_T$ in the events in each bin after $\eta$ intercalibration but before the application of any other in situ steps.

Missing energy in each event is reconstructed from calorimeter topo-clusters in the case of EMtopo jet calibration and from particle-flow objects in the case of PFlow jet calibration, ensuring that the energy scale is consistent. The $Z \rightarrow ee$ events are required to pass a dielectron trigger with $p_T^{\text{ele,e2}} > 15$ GeV; $Z \rightarrow \mu\mu$ events must pass a similar dimuon trigger with $p_T^{\mu1,\mu2} > 14$ GeV [49,50]. Electrons entering the analysis must have $p_T > 20$ GeV, ensuring that the trigger is fully efficient, must be contained within the tracker such that $|\eta_e| < 2.47$, and must not fall in the calorimeter transition region ($1.37 < |\eta| < 1.52$). Muons entering the analysis are required to have $p_T > 20$ GeV and to fall within $|\eta| < 2.4$. Both electron and muon candidates must also pass loose identification and isolation requirements [47,48]. All $Z$+jet events are selected such that the reconstructed mass calculated from the electron or muon pair must be close to the Z boson mass: $66$ GeV $< m_{ee/\mu\mu} < 116$ GeV.

A combination of single-photon triggers are used for the $\gamma$+jet analysis, with the lowest trigger threshold corresponding to $E_T^\gamma > 15$ GeV. Offline photons must have $E_T^\gamma > 25$ GeV and $|\eta^\gamma| < 1.37$ and must satisfy tight identification and isolation criteria [47].

Both the $Z$+jet and $\gamma$+jet analyses have further selection requirements on the jets and event topology to suppress pile-up and initial- and final-state radiation. All jets within $\Delta R = 0.2$ of a photon or $\Delta R = 0.35$ of a lepton are removed. Jets must satisfy basic cleaning requirements and pass the JVT selection to suppress pile-up. Selected events must have one jet with $p_T > 10$ GeV and $|\eta| < 0.8$. Additional event activity is suppressed by requiring that any subleading jet must have $p_T < \max(0.3 \times p_T^{\text{ref}}, 12$ GeV and that the leading jet and reference object must be relatively back-to-back with $\Delta \phi^{\text{ref}, \text{jet}} > 2.9$. The relatively loose $p_T$ cut on subleading jets is shown to be acceptable for the MPF analysis due to its intrinsic robustness to pile-up effects.

Figures 11 and 12 show the MPF response calculated in $Z$+jet and $\gamma$+jet events for data and for two MC samples using different generators. The lower panels show the MC simulation to data ratio for both generators. The results using POWHEG+PYTHIA8 ($Z$+jet) and PYTHIA8 ($\gamma$+jet) constitute the nominal calibration while SHERPA is used to define an uncertainty due to the generator choice. In the lowest $p_T$ bin of the $\gamma$+jet measurement, the discrepancy between the MC predictions is caused by a generator-level cut at 35 GeV present in the SHERPA sample. This point is included in the final in situ combination, but due to its large generator uncertainty it contributes very little to the overall weighted-average-based result (see Sect. 5.2.5 and Fig. 19a) and the total effect is negligible. The $\gamma$+jet generator uncertainty at this point has therefore been set to its value in the second-lowest bin for display purposes in Fig. 14 to better reflect its actual contribution to the total systematic uncertainties. The apparent dip near the lowest $p_T$ range of each measurement is due to the interplay of two factors: an asymmetry in the $R_{\text{MPF}}$ distribution near the low $p_T$ reconstruction threshold which causes the measured response to increase for the lowest $p_T$ values, and the natural increase in response with higher jet $p_T$. One motivation for the use of the MPF technique is increased resilience to this threshold effect.

Two small correction factors are derived in simulation and use the true calorimeter response, defined as the ratio of measured energy in the calorimeter deposited by particles belonging to a particle-level jet to the total energy of the particle-level jet. The topology correction accounts for the differences in calorimeter response for sparse energy depositions versus those in the dense cores of jets, and is found by taking the average of the ratio of $R_{\text{MPF}}$ to the true calorimeter response in each $p_T$ bin. The showering correction accounts for the flow of particles entering or exiting across the boundaries of the jet definition and is calculated from the ratio.

---

5 The balance can also be affected by leptonic processes in heavy flavour decays, though heavy flavour production is rare in $Z$+jet events. Electrons are included in the balance calculation but muons and neutrinos are not. However, this effect is not a problem like additional radiation, since the balance in heavy flavour events actually contributes to measuring the energy loss and correctly calibrating their momenta.

---

Springer
Fig. 11 Average PFlow jet response as a function of reference $p_T$ for $Z$+jet events where the $Z$ boson decays into $\text{e}^+\text{e}^-$ and into $\mu^+\mu^-$ events are combined at a later stage. The black points correspond to 2015–2017 data while the pink diamonds and blue triangles correspond to independent Monte Carlo samples from two different generators, and their error bars show the statistical uncertainties. The ratio of MC simulation to data for both generators is shown in the bottom panel and defines the in situ correction to be applied. The dotted lines at 1 and 1.05 serve as a reference of the true calorimeter response to the measured response of the reconstructed jet, therefore varying with the jet algorithm and size. The total correction factor is the product of the two and is found to be less than 2% for jets of $p_T < 50$ GeV and negligible above that. This correction factor would in principle be applied identically to $R_{\text{MPF}}$ in both data and simulation to better estimate jet response, but since the ratio of $R_{\text{MPF}}$ in data and simulation is the quantity of interest for the in situ calibration, the correction would cancel out in the ratio and only the uncertainty in its derivation is relevant. This uncertainty is taken from a comparison of two different physics lists (FTFP BERT [29] and QGSP BIC [51]) in the simulation of the particle/detector interactions and is found to be $\sim 2\%$ for jets with $p_T < 20$ GeV, $\sim 0.5\%$ for jets with $20 < p_T < 40$ GeV and zero for jets with $p_T > 40$ GeV.

Fig. 12 Average PFlow jet response as a function of reference $p_T$ for $\gamma$+jet events calculated using the MPF technique. The black points correspond to 2015–2017 data. The red and blue triangles correspond to independent Monte Carlo samples from two different generators. Error bars show the statistical uncertainties. The ratio of MC simulation to data for both generators is shown in the bottom panel and defines the in situ correction to be applied. The dotted lines at 1 and 1.05 serve as a reference of the true calorimeter response to the measured response of the reconstructed jet, therefore varying with the jet algorithm and size. The total correction factor is the product of the two and is found to be less than 2% for jets of $p_T < 50$ GeV and negligible above that. This correction factor would in principle be applied identically to $R_{\text{MPF}}$ in both data and simulation to better estimate jet response, but since the ratio of $R_{\text{MPF}}$ in data and simulation is the quantity of interest for the in situ calibration, the correction would cancel out in the ratio and only the uncertainty in its derivation is relevant. This uncertainty is taken from a comparison of two different physics lists (FTFP BERT [29] and QGSP BIC [51]) in the simulation of the particle/detector interactions and is found to be $\sim 2\%$ for jets with $p_T < 20$ GeV, $\sim 0.5\%$ for jets with $20 < p_T < 40$ GeV and zero for jets with $p_T > 40$ GeV.

The full set of uncertainties is shown for the $Z \to \text{ee} + \text{jet}$ and $Z \to \mu\mu + \text{jet}$ analyses in Fig. 13 and for the $\gamma$+jet analysis in Fig. 14. The dominant systematic uncertainties are due to generator differences at lower $p_T$ and to the photon energy scale at higher $p_T$. Uncertainties in the $e^-, \mu^-$, and $\gamma$ energy scales and resolutions are taken from the calibrations provided for each physics object and are propagated through the analysis [47,48]. The $\Delta\phi$ and second-jet veto uncertainties are estimated by varying the cuts and comparing the resulting response measurements. As in the $\eta$ intercalibration, the JVT uncertainty is determined by comparison with tighter and looser working points. A photon purity uncertainty is estimated for the $\gamma$+jet analysis using control regions dominated by dijet events where one of the jets can be misidentified as a photon. The uncertainty on the final state modelling is taken, as discussed, from the generator comparison. Limited data and MC statistics contribute to the statistical uncertainty, which is largest for the lowest and highest bins of each analysis. A bootstrapping procedure is applied to the uncertainties to suppress statistical fluctuations as previously described.

Similar analyses in the $Z/\gamma$+jet final states but explicitly balancing the reference $p_T$ against the $p_T$ of a reconstructed jet (direct balance) are used to cross-check the jet energy
Fig. 13 Systematic uncertainties for PFlow jets as a function of reference $p_T$ for a $Z \rightarrow ee + jet$ events and b $Z \rightarrow \mu\mu + jet$ events calculated using the MPF technique. Uncertainties due to the JVT, second-jet veto, and $\Delta \phi$ cuts derive from the analysis technique. Electron or muon (as appropriate) scale and resolution uncertainties are propagated through the analysis from the uncertainties associated with the individual objects. The statistical uncertainties come from the MC simulation/data ratio and reach a maximum value of 0.083 in b while the difference between the Pythia8 and Sherpa samples defines the MC generator uncertainty. All uncertainties are smoothed to ensure that the visible fluctuations are statistically significant.

Fig. 14 Systematic uncertainties on PFlow jets as a function of reference $p_T$ for $\gamma + jet$ events calculated using the MPF technique. Uncertainties due to the JVT, second-jet veto, and $\Delta \phi$ cuts derive from the analysis technique. Photon scale and resolution uncertainties are propagated through the analysis from the uncertainties associated with the individual objects. The statistical uncertainties come from the MC simulation/data ratio while the difference between the Pythia8 and Sherpa samples defines the MC generator uncertainty. All uncertainties are smoothed to ensure that the visible fluctuations are statistically significant.

scale calibration. The JES results computed using direct balance showed good agreement with those derived via MPF.

The innate difference in response between EMtopo and PFlow jets can be seen by comparing their measured MPF responses. Since the MPF method uses topo-clusters and PFlow objects in computing the missing energy, the measured responses are independent of the MCJES calibration and reflect the precalibration response for each jet input type. The MPF responses measured in the $\gamma + jet$ analysis for EMtopo and PFlow jets are shown in Fig. 15. The shape of the EMtopo measurement follows the form of the Groom’s function, which corresponds to the response expected from a hadronic calorimeter [52]. The PFlow measurement does not follow the same shape but instead shows an improvement over the baseline calorimeter response at low $p_T$ thanks to the inclusion of information from tracks.

5.2.3 High-$p_T$ jet calibration using multijet balance

The final stage of in situ calibration derives a correction for jets with $p_T$ above the range of the $Z/\gamma + jet$ analyses using the multijet balance (MJB) technique. Events are selected with a single high-$p_T$ jet balanced against a system of lower-$p_T$ jets (the recoil system). The jets of the recoil system are selected to ensure they are well calibrated using a combination of the $Z/\gamma + jet$ results (Sect. 5.2.2), while the leading jet is left at the scale of the $\eta$ intercalibration. The response of the system is defined as:

$$R_{MJB} = \frac{\langle p_{\text{lead}} \rangle}{p_{\text{ref}}},$$

where $p_{\text{ref}}$ is taken from the vector sum of all jets in the recoil system. In a procedure parallel to that used for the $Z/\gamma + jet$ analyses, the response is measured in bins of $p_{\text{ref}}$ and the correction is then mapped to the uncalibrated leading jet by finding the average $p_{\text{lead}}$ of the events in each bin.
Since the MJB analysis can only include events where all jets of the recoil system can already be well-calibrated, events with very high $p_T$ are often excluded as their second and third leading jets can have momenta outside the range of calibration by the $Z/\gamma +$jet analyses. To address this, MJB proceeds via two iterations. In the first iteration, a combination of the $Z/\gamma +$jet results is used to calibrate the recoil system, so only events with subleading jets of $p_T < 1.2$ TeV are included. In the second iteration, events with subleading jets up to $p_T = 1.8$ TeV are included and calibrated using the MJB results from the first iteration. This extends the range of the calibration to $p_T^{\text{lead}} = 2.4$ TeV.

Events are selected for the MJB analysis using a variety of single-jet triggers with each corresponding to a unique range of $p_T^{\text{lead}}$. To suppress dijet topologies and ensure that only true multijet events are used, events must have at least three jets with $p_T > 25$ GeV and $|\eta| < 2.8$ and the subleading jet must not have a momentum above $0.8 p_T^{\text{lead}}$. Jets are as usual required to pass JVT selections, limiting the effects of pile-up. Isolation of the leading jet from contamination by the recoil system is ensured by requiring that the azimuthal angle $\Delta \phi$ between the leading jet and the direction of the recoil system is at least 0.3 radians and that the $\Delta \phi$ between the leading jet and any individual jet in the recoil system with $p_T > 0.05 p_T^{\text{lead}}$ is at least 1.0 radians.

The MJB response in data and in four MC samples with different generators is shown in Fig. 16a. In both data and MC simulation, the response decreases at lower $p_T$ due to the intrinsic bias in $R_{\text{MJB}}$ from the combined effects of the leading jet isolation and $p_T$ asymmetry requirements. This bias is greater for lower- $p_T$ leading jets, but is well modelled in simulation, leaving the calibration unbiased. The lower panel shows the ratio of the response of each MC sample to data. Here, the ratio of the SHERPA sample to data defines the nominal correction while the ratio based on PYTHIA defines an uncertainty on the generator choice. This response ratio is constant and approximately 2% for jets above 1 TeV; below this point the calculated correction is slightly smaller.

All uncertainties in the MJB analysis are shown in Fig. 16b. The dominant term at low $p_T^{\text{lead}}$ is the uncertainty from jet flavour, derived in simulation and reflecting the difference in jet response for quark-initiated and gluon-initiated jets. Two terms contribute, one reflecting the uncertainty in the fraction of gluon-initiated jets in the sample, the other based on the difference in MC simulation-derived gluon response between generators. Other independently derived uncertainties correspond to pile-up and punch-through effects and are propagated through the MJB analysis via the recoil system. The $Z +$jet, $\gamma +$jet, and $\eta$ intercalibration uncertainties are propagated from the previous stages of in situ analysis. Event selection uncertainties are determined by varying each of the analysis cuts and determining the effects on the measured response ratio. Finally, the MC generator uncertainty is derived as described above by comparing the response ratio of SHERPA with PYTHIA as an alternative. Results using HERWIG and POWHEG+PYTHIA are shown for reference but are not included in the uncertainty definition as they are less reliable for this measurement. All uncertainties are smoothed via the bootstrapping procedure to ensure statistical significance, and the total uncertainty is found to be below 1.5% for all considered values of $p_T^{\text{lead}}$. The MC generator uncertainty, which is defined in a one-sided fashion from the response ratios, is symmetrised by the in situ combination process along with the other uncertainties. However, its full one-sided size is shown in Fig. 16b for easier comparison with Fig. 16a.

For EMtopo jets the intrinsic bias at low $p_T$ is slightly smaller and more closely tracked by simulation, leading in turn to slightly reduced systematic uncertainties for jets below $p_T \sim 700$ GeV. Above $p_T > 1$ TeV, in situ uncertainties propagated from lower-$p_T$ jets have a greater impact, and therefore the uncertainty is smaller for PFlow jets than for EMtopo jets.

5.2.4 Pile-up and the in situ analyses

One of the primary changes in LHC run conditions over the course of Run 2 was an increase in pile-up. The average number of interactions per crossing ($\mu$) during 2015+2016 data taking was 23.7, which increased to 37.8 in 2017. The data taken during 2018 and to which the calibrations in this paper are also applied has an average of 36.1 interactions per crossing [16]. The consistency of the calibrations for events...
with different pile-up conditions is therefore an important feature of the methods.

Figure 17 shows individual bins in the response ratios of the Z+jet and γ+jet analyses separated out as a function of number of primary vertices in the event. The Z+jet results are shown for 25 GeV < $p_T^{\text{ref}}$ < 30 GeV and the γ+jet results for 45 GeV < $p_T^{\text{ref}}$ < 65 GeV, in the regions where each has appropriate statistical significance. The multijet balance analysis is not shown: due to the higher $p_T$ regime in which it operates it is more robust to pile-up effects. A linear fit to the data/simulation ratio has a slope compatible with zero within the fit uncertainties in each plot, demonstrating the stability of the in situ calibration as a function of $N_{\text{PV}}$. This in turn illustrates the efficacy of the pile-up corrections described in Sect. 5.1.1 and shows that the inclusively derived calibration is applicable to events with a range of pile-up conditions.

The in situ JES measurements can be used to calculate the dependence of the measured median $p_T$ density $\rho$ on the event topology in simulation and data and to derive an uncertainty, as mentioned in Sect. 5.1.1. The density $\rho$ is computed as a function of $\mu$ for each of the Z+jet, γ+jet, and dijet topologies as shown in the top panels of Fig. 18. Taking $\rho_{2017}$ as the value of $\rho$ for the average pile-up conditions during 2017 data taking and $t_1$ and $t_2$ as any two in situ measurement topologies out of Z+jet, γ+jet, and dijet, then the following metric of consistency can be defined:

$$\Delta = \left( \frac{\rho_{2017}^{t_1} - \rho_{2017}^{t_2}}{\rho_{2017}^{t_1} - \rho_{2017}^{t_2}} \right)_{\text{MC}} - \left( \frac{\rho_{t_1}^{\text{data}} - \rho_{t_2}^{\text{data}}}{\rho_{t_1}^{\text{data}} - \rho_{t_2}^{\text{data}}} \right)_{\text{data}}.$$

The quantity $\max(|\Delta|)$ is then the largest value of $\Delta$ across the various topology comparisons. The total $\rho$ topology systematic uncertainty is given by

$$\Delta_{p_T} = \max(|\Delta|) \times C_{p_T}^{\text{JES}} \times \pi R^2,$$

where $C_{p_T}^{\text{JES}}$ is the size of the MCJES correction for a jet with the relevant $p_T$. The second panels in Fig. 18 show $\rho_{t_1}^{\text{data}} - \rho_{t_2}^{\text{data}}$ for the comparisons (Z+jet, dijet) and (γ+jet, dijet) in both MC simulation and data. The lower panels show the difference of these two quantities between data and MC simulation, that is, $\Delta_{Z+jet, dijet}$ and $\Delta_{\gamma+jet, dijet}$. The input to the systematic uncertainty $\max(|\Delta|)$ is the most discrepant of the two lines in the lower panel evaluated at $\mu = 37.8$, the value in 2017 data. As Fig. 18 illustrates, this uncertainty is larger for PFFlow jets than for EMtopo jets. This is understood to be due to a greater sensitivity to the underlying event when tracking information is included, which leads to greater differences among the simulated samples.

5.2.5 In situ combination

The data/MC simulation response ratios,

$$\left( \frac{p_T^{\text{jet}}}{p_T^{\text{ref}}} \right)_{\text{data}} / \left( \frac{p_T^{\text{jet}}}{p_T^{\text{ref}}} \right)_{\text{MC}},$$

from the four different ‘absolute’ in situ measurements of $Z(\to ee)+\text{jet}$, $Z(\to \mu\mu)+\text{jet}$, $\gamma+\text{jet}$, and the multijet balance must be combined to produce a single calibration covering the full range of jet $p_T$ from 17 GeV to 2.4 TeV. The

---

Fig. 16 a Response for the leading PFFlow jet in multijet events as a function of $p_T^{\text{ref}}$ and b the systematic uncertainties on the response. Subleading jets in the event are calibrated using the $Z/\gamma+\text{jet}$ MPF corrections, while the leading jet is calibrated only up to the $\eta$ intercalibration. The response is shown for data and for simulation using four different MC generators, and the MC simulation-to-data response ratios in the bottom panel correspond to the derived in situ calibration. The error bars show the statistical uncertainties. The nominal calibration is defined by the comparison with SHERPA; its difference from the PYTHIA result defines the ‘MC generator’ uncertainty in b. This uncertainty is defined in a single-sided way by the measured response difference and therefore it is not symmetrised for display in b but instead its full one-sided value is shown. Other uncertainties come from the event selection and MC simulation/data statistics or are propagated from the Z+jet, γ+jet, flavour, pile-up, $\eta$ intercalibration, and punch-through studies.
Fig. 17 Average PFlow MPF jet response as a function of $N_{PV}$ for (a) Z+jet events with reference $p_T$ derived from the reconstructed Z boson in the range 25 GeV $< p_T^\text{ref} < 30$ GeV and for (b) $\gamma$+jet events with reference $p_T$ defined from the photon in the range 45 GeV $< p_T^\text{ref} < 65$ GeV. For the Z+jet analysis, results from the $Z \rightarrow \mu\mu$ and $Z \rightarrow e\mu$ channels are combined to reduce statistical fluctuations. The black points correspond to 2015–2017 data while the pink and blue points correspond to Monte Carlo samples from two different generators. The error bars reflect the statistical uncertainties. The ratio of MC simulation to data for both generators is shown in the bottom panel.

Fig. 18 Inputs to the $\rho$ topology uncertainty derived in the Z+jet, $\gamma$+jet, and dijet in situ analyses. The error bars show the statistical uncertainties. The top panels relate the $p_T$ density $\rho$ to the mean number of interactions per bunch crossing $\mu$ in data and MC simulation for the three input analyses. The second panels show the difference between the Z+jet and dijet and between the $\gamma$+jet and dijet measurements. The lowermost panels show the difference between the data and MC simulation lines in the second panels: this defines the size of the topology uncertainty. The two plots show a E_Mtopo and a PFlow jets, illustrating why this uncertainty is larger for PFlow jets than for E_Mtopo jets.

The combination procedure is briefly summarized here; for a detailed description see Ref. [5]. Each of the absolute in situ measurements is converted from a parameterisation in terms of reference object $p_T$ into jet $p_T$ and divided into finer bins of 1 GeV using second-order polynomial splines. A $\chi^2$ minimization is performed in each bin, taking as inputs the measurements available in that $p_T$ range and their uncertainties. This minimisation functions as a weighted average, with the weight given to each input measurement decreasing as
Fig. 19 a The weight assigned to different techniques in the combination of in situ measurements of the relative jet $p_T$ response of anti-$k_t$, $R = 0.4$ particle-flow jets in data and simulation, as a function of the jet $p_T$. For each $p_T$ bin, the weights of the $Z$+jet, $\gamma$+jet, and multi-jet balance methods are shown. b The $\chi^2/N_{\text{dof}}$ metric, illustrating the compatibility of the in situ measurements being combined, as a function of jet $p_T$. In the low $p_T$ range, the combination is between three measurements ($Z(\rightarrow ee)$+jet, $Z(\rightarrow \mu\mu)$+jet, and $\gamma$+jet) of which the two $Z$+jet measurements have several correlated uncertainties, resulting in increased tension compared to previous calibrations.

The weights of each input measurement in this combination are shown in Fig. 19a as a function of jet $p_T$. The $Z$+jet measurements dominate for jet $p_T$ below $\sim 500$ GeV where the statistical uncertainties on these measurements grow dramatically; the $Z(\rightarrow \mu\mu)$+jet is the more powerful of the two in the upper half of this range due to the size of the electron scale and resolution uncertainties affecting the $Z(\rightarrow ee)$+jet channel. The combination is then dominated by $\gamma$+jet until jet $p_T$ of above 1 TeV, where the lower statistics in this channel and the decreased flavour uncertainties in the multi-jet balance analysis allow the latter to dominate. The final calibration curve is determined by smoothing the outputs from the minimization with a Gaussian kernel.

The $\sqrt{\chi^2/N_{\text{dof}}}$ across the measurements, before any scaling is applied, is shown in Fig. 19b. This metric shows the degree of tension between the input measurements at each point: when they are in agreement well within uncertainties the value will be below 1, while when they differ relative to their uncertainties it will be above 1. Following PDG guidelines, in bins where tension between the input measurements, quantified by $\sqrt{\chi^2/N_{\text{dof}}}$, is found to be greater than 1, the uncertainties in the measurements in that bin are scaled by the same tension factor to ensure that the overall level of agreement between methods is acceptable within uncertainties for all $p_T$ values [53]. However, since the tensions visible at low $p_T$ are primarily between the two $Z$+jet measurements, and since the MC generator and showering and topology uncertainties are fully correlated between the two channels and therefore cannot contribute to this tension, these two components are excluded from the scaling procedure. The components which are not scaled are the dominant uncertainties.

Figure 20 shows the final in situ combination as a function of jet $p_T$. To complete the calibration, the inverse of the curve ($R_{\text{MC}}/R_{\text{data}}$) is taken as the scaling factor and applied to data.
The combined measurement (solid line) for PFlow+JES jets is compared with each of the four absolute in situ analyses (empty shapes) in Fig. 20a. The total size of the correction is approximately 3% at low \( p_T \) and decreases to around 2% for jets above 200 GeV. A comparison between the results for EM+JES and PFlow+JES jets is shown in Fig. 20b, where the overall size of both the correction and its uncertainty is seen to be slightly larger for EM+JES jets.

Each uncertainty component from the in situ analyses is individually propagated through the combination procedure. First, the relevant measured response is varied by 1\( \sigma \) in the uncertainty component within its standard binning. The finer rebinning, \( \chi^2 \) minimization, and combination procedure is repeated, although using the weights as determined for the nominal result to prevent the varied uncertainty from decreasing the contribution of the measurement. The difference between the combined calibration curve with the systematically shifted input and the nominal calibration curve is taken as 1\( \sigma \) in the varied uncertainty. Throughout this process, each individual uncertainty source is treated as fully correlated across \( \eta \) and \( p_T \) but entirely uncorrelated with all other uncertainty sources. After this step, the uncertainties from the Z+jet analyses are taken to be fully correlated with the same uncertainties propagated through the multijet balance. Other assumptions of correlation between components can similarly be made and altered after their propagation, allowing multiple different assumptions.

5.3 Systematic uncertainties

The full uncertainty in the jet energy scale consists of 125 individual terms derived from the in situ measurements, pileup effects, flavour dependence, and estimates of additional effects as summarized in Table 2. The majority of the individual terms stem from the in situ measurements and cover the effects of analysis selection cuts, event topology dependence, and MC mis-modelling and statistical limitations, as well as the uncertainties associated with the calibration of the electrons, muons, and photons.

The \( \eta \) intercalibration analysis results in five nuisance parameters, with a sixth for 2018 data only, as discussed in Sect. 5.2.1: one covers systematic effects, one covers statistical uncertainty, and three (four in 2018) are used to parameterize the non-closure. Pile-up effects are described by four nuisance parameters which account for offsets and \( p_T \) dependence in \( \langle \mu \rangle \) and \( N_{PV} \), as well as event topology dependence of the density metric \( \rho \). The offset and \( p_T \) dependence terms are derived in data using a combination of Z+jet measurements and measurements comparing reconstructed jets with track-jets. The \( \rho \) topology term is the largest of the pile-up uncertainties and is determined by the maximum deviation in measured density between different in situ measurements under the same pile-up conditions.

The two flavour dependence uncertainties are derived from simulation and account for relative flavour fractions and differing responses to quark- and gluon-initiated jets [5,6]. The flavour composition uncertainty accounts for the differing response of quark- and gluon-initiated jets in a sample with some uncertainty on the fraction of gluon-initiated jets \( f_g \). Where \( \mathcal{R}_q \) and \( \mathcal{R}_g \) are the responses measured in PYTHIA and \( \sigma^f_g \) is the uncertainty on \( f_g \) in the sample, the flavour composition uncertainty is defined as:

\[
\sigma_{\text{composition}} = \frac{|\mathcal{R}_q - \mathcal{R}_g|}{f_g \mathcal{R}_g + (1 - f_g) \mathcal{R}_q}.
\]

The flavour response uncertainty accounts for the fact that, unlike quark-initiated jet response, gluon-initiated jet response is found to differ significantly between generators. This uncertainty is defined by comparison between the nominal PYTHIA sample and an alternative HERWIG sample:

\[
\sigma_{\text{response}} = f_g \left( \mathcal{R}_g^{\text{PYTHIA}} - \mathcal{R}_g^{\text{HERWIG}} \right).
\]

Figure 21 shows the gluon-jet response and the difference between quark-jet and gluon-jet responses using both PYTHIA and HERWIG for PFlow jets. The samples are the same as those used for the multijet balance analysis and are dominated by gluon jets at low \( p_T \). For HERWIG, \( \mathcal{R}_q - \mathcal{R}_g \) becomes negative in the 90–600 \( p_T \) region (which appears as a bump in the \( |\mathcal{R}_q - \mathcal{R}_g| \) curve).

An additional uncertainty applied only to \( b \)-initiated jets covers the difference in response between jets from light-versus heavy-flavour quarks and replaces the flavour composition and response uncertainties for these heavy-flavour jets. The punch-through uncertainty accounts for mis-modelling of the GSC correction to jets which pass through the calorimeter and into the muon system, taking the difference in jet response between data and MC simulation in bins of muon detector activity as the systematic uncertainty. Both are discussed in more detail in Ref. [6]. Finally, the high-\( p_T \) ‘single particle’ uncertainty is derived from studies of the response to individual hadrons and is used to cover the region beyond 2.4 TeV, where the MJN analysis no longer has statistical power [29]. When calibrating MC samples simulated using AFI, an additional non-closure uncertainty is applied to account for the difference in jet response between these samples and those which used full detector simulation.

The total jet energy scale uncertainty is shown in Fig. 22a as a function of jet \( p_T \) for fixed \( \eta_{\text{jet}} = 0 \) and in Fig. 22b as a function of jet \( \eta \) for fixed \( p_T^{\text{jet}} = 60 \) GeV. A dijet-like composition of the sample (that is, predominantly gluons) is assumed in computing the flavour uncertainties. The uncertainties in the \( \eta \) intercalibration analysis are labelled ‘relative in situ JES’ with the non-closure uncertainty creating the asymmetric peaks around \( \eta = \pm 2.5 \). Uncertainties in all other in situ measurements are combined into the ‘abso-
<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$ intercalibration</td>
<td>Envelope of the generator, pile-up, and event topology variations</td>
</tr>
<tr>
<td>Systematic mis-modelling</td>
<td>Statistical uncertainty (single component)</td>
</tr>
<tr>
<td>Statistical component</td>
<td>Three components describing non-closure at high energy and at $\eta \sim \pm 2.4$</td>
</tr>
<tr>
<td>Non-closure</td>
<td>Single component describing non-closure at $\eta \sim \pm 1.5$ due to Tile calibration</td>
</tr>
<tr>
<td>Z + jet</td>
<td>Electron scale uncertainty in the electron energy scale</td>
</tr>
<tr>
<td>Photon scale</td>
<td>Uncertainty in the electron energy scale</td>
</tr>
<tr>
<td>Electron resolution</td>
<td>Uncertainty in the electron energy resolution</td>
</tr>
<tr>
<td>Muon scale</td>
<td>Uncertainty in the muon momentum scale</td>
</tr>
<tr>
<td>Muon resolution (ID)</td>
<td>Uncertainty in muon momentum resolution in the ID</td>
</tr>
<tr>
<td>Non-closure, 2018 only</td>
<td>Uncertainty in muon momentum resolution in the MS</td>
</tr>
<tr>
<td>MC generator</td>
<td>Difference between MC event generators</td>
</tr>
<tr>
<td>JVT cut</td>
<td>Jet vertex tagger uncertainty</td>
</tr>
<tr>
<td>$\Delta\phi$ cut</td>
<td>Variation of $\Delta\phi$ between the jet and Z boson</td>
</tr>
<tr>
<td>Showering &amp; topology</td>
<td>Radiation suppression through second-jet veto</td>
</tr>
<tr>
<td>Statistical</td>
<td>Modelling energy flow and distribution in and around a jet</td>
</tr>
<tr>
<td>Statistical</td>
<td>Statistical uncertainty in 28 discrete $p_T$ terms</td>
</tr>
<tr>
<td>$\gamma +$ jet</td>
<td>Uncertainty in the photon energy scale</td>
</tr>
<tr>
<td>Photon scale</td>
<td>Uncertainty in the photon energy scale</td>
</tr>
<tr>
<td>Photon resolution</td>
<td>Uncertainty in the photon energy resolution</td>
</tr>
<tr>
<td>MC generator</td>
<td>Difference between MC event generators</td>
</tr>
<tr>
<td>JVT cut</td>
<td>Jet vertex tagger uncertainty</td>
</tr>
<tr>
<td>$\Delta\phi$ cut</td>
<td>Variation of $\Delta\phi$ between the jet and photon</td>
</tr>
<tr>
<td>Showering &amp; topology</td>
<td>Radiation suppression through second-jet veto</td>
</tr>
<tr>
<td>Subleading jet veto</td>
<td>Modelling energy flow and distribution in and around a jet</td>
</tr>
<tr>
<td>Photon purity</td>
<td>Purity of sample used for $\gamma +$ jet balance</td>
</tr>
<tr>
<td>Statistical</td>
<td>Statistical uncertainty in 16 discrete $p_T$ terms</td>
</tr>
<tr>
<td>Multijet balance</td>
<td>Angle between leading jet and recoil system</td>
</tr>
<tr>
<td>$\Delta\phi$ (lead, recoil system)</td>
<td>Difference between MC event generators</td>
</tr>
<tr>
<td>$\Delta\phi$ (lead, any sublead)</td>
<td>Second jet’s $p_T$ contribution to the recoil system</td>
</tr>
<tr>
<td>MC generator</td>
<td>Jet $p_T$ threshold</td>
</tr>
<tr>
<td>$p_T$ asym selection</td>
<td>Statistical uncertainty in 28 discrete $p_T$ terms</td>
</tr>
<tr>
<td>Jet $p_T$</td>
<td>Statistical uncertainty in 28 discrete $p_T$ terms</td>
</tr>
<tr>
<td>Statistical</td>
<td>Statistical uncertainty in 28 discrete $p_T$ terms</td>
</tr>
<tr>
<td>Pile-up</td>
<td>Uncertainty in the $\mu$ modelling in MC simulation</td>
</tr>
<tr>
<td>$N_{PV}$ offset</td>
<td>Uncertainty in the $N_{PV}$ modelling in MC simulation</td>
</tr>
<tr>
<td>$p_T$ topology</td>
<td>Uncertainty in the per-event $p_T$ density modelling in MC simulation</td>
</tr>
<tr>
<td>$p_T$ dependence</td>
<td>Uncertainty in the residual $p_T$ dependence</td>
</tr>
<tr>
<td>Jet flavour</td>
<td>Uncertainty in the proportional sample composition of quarks and gluons</td>
</tr>
<tr>
<td>Flavour composition</td>
<td>Uncertainty in the response of gluon-initiated jets</td>
</tr>
<tr>
<td>Flavour response</td>
<td>Uncertainty in the response of $b$-quark-initiated jets</td>
</tr>
<tr>
<td>$b$-jets</td>
<td>Uncertainty in GSC punch-through correction</td>
</tr>
<tr>
<td>Punch-through</td>
<td>Uncertainty in the response of single-particle and test-beam measurements</td>
</tr>
<tr>
<td>Single-particle response</td>
<td>High-$p_T$ jet uncertainty from single-particle and test-beam measurements</td>
</tr>
<tr>
<td>AFII non-closure</td>
<td>Difference in the absolute JES calibration for simulations in AFII</td>
</tr>
</tbody>
</table>
Fig. 21 a Measured gluon-initiated jet response and b difference between quark- and gluon-initiated jet responses for PFlow jets using two different generators. These define the flavour response and composition uncertainties respectively.

Fig. 22 Fractional jet energy scale systematic uncertainty components for anti-$k_t$, $R = 0.4$ jets a as a function of jet $p_T$ at $\eta = 0$ and b as a function of $\eta$ at $p_T = 60$ GeV, reconstructed from particle-flow objects. The total uncertainty, determined as the quadrature sum of all components, is shown as a filled region topped by a solid black line. Flavour-dependent components shown here assume a dijet flavour composition.

5.3.1 Uncertainty correlations and reductions

The detail contained in 125 independent nuisance parameters is far more than is required by most analyses, so it is necessary to reduce the uncertainty description to a smaller number of terms. One could imagine a single ‘Jet energy scale’ nuisance parameter constructed by adding in quadrature all of the independent components. However, a meaningful set of correlations exist between the jet energy scale uncertainties for two jets at different $\eta$ and $p_T$ as a result of the structures of the nuisance parameters. In the case of reduction to a single component, the entirety of this correlation information would be lost and an unrealistic assumption – that of full correlation between the jet energy scale uncertainties for any values of $\eta$ and $p_T$ – would be enforced. In practice, a variety of reduced uncertainty schemes are provided to allow simplified descriptions with a minimum loss of correlation information.

The 98 uncertainty components stemming from the absolute in situ analyses are functions only of $p_T$ and thus their behaviour can be easily represented by a smaller number of orthogonal terms. An eigenvector decomposition is performed on a covariance matrix of these uncertainty components and the largest of the resulting orthogonal terms are kept separate as new effective nuisance parameters [5]. The remaining terms are combined into a single residual nuisance parameter. To determine how many components to keep independently and how many to combine in the residual term, the covariance matrix for the reduced set is also computed and the difference in correlation in each jet $\eta$ and $p_T$ between the reduced set and the full set is calculated. This difference is taken as a measure of the information loss and the number of combined terms is adjusted so that the difference is below an acceptable bound (usually 0.05). Two different reduction schemes are produced: the global reduction combines all $p_T$-dependent in situ uncertainty components regardless of their sources and results in 8 reduced components for a total of 23 once the two-dimensional terms (not arising from the in situ analyses and not reduced) are included. The category reduction combines the $p_T$-dependent in situ uncertainty components in separate groups based on their origin (detector, statistical, modelling, or mixed) and results in 15 reduced components for a total of 30. The JES correlation matrix for the full set of nuisance parameters is shown in Fig. 23a. The bin-by-bin correlation loss between the full set of nuisance parameters and the category reduction is shown in Fig. 23b and is below 0.05 everywhere as required.
Fig. 23  a The jet energy scale correlation matrix for two PFlow+JES jets at $\eta = 0$ using the full set of 98 $p_T$-dependent in situ nuisance parameters and b the difference in correlation information between the full description and the category reduction. The maximum loss of correlation information is 0.02 and occurs at the $(p_T^1, p_T^2)$ location specified by the text at the bottom of the plot.

While the same procedure could in principle be used for the components which depend on both $p_T$ and $\eta$, the complexity added by the second dimension means that nearly as many eigenvectors would be needed to adequately describe the correlations as there were original terms and so the gain would be minimal. However, many analyses still require fewer than 25 nuisance parameters and are not affected by loss of correlation information. To provide suitable uncertainties for these, a strong reduction procedure is used to group the globally reduced versions of the absolute in situ uncertainties together with the two-dimensional uncertainties into three effective nuisance parameters as detailed in Ref. [7]. The three terms of the $\eta$ intercalibration non-closure uncertainty are kept separate because their two-dimensional shapes are especially difficult to reduce and would cause an unacceptably large correlation loss.

Four different sets (scenarios) of the three effective nuisance parameters are created by varying the combinations of terms they contain. The varied sets are chosen such that the correlation loss in each is constrained to an $\eta$–$p_T$ range which is well described by a different set. The metric for assessing performance of the four scenarios is the uncovered correlation loss, defined as the maximum difference in correlation between any two reduced scenarios minus the minimum difference in correlation between any reduced scenario and the full set of nuisance parameters. The uncovered correlation loss is calculated for a fine grid of points in $\eta$ and $p_T$, ensuring no small-scale structures are missed. Contents of the effective nuisance parameters are varied, keeping systematic uncertainties with similar behaviours mostly grouped together, until a set of scenarios is found in which the maximum uncovered correlation loss is kept below 0.25 and confined to sufficiently small regions that the average correlation loss in an $\eta$–$p_T$ plane does not exceed 0.02. A detailed discussion of the application of strongly reduced uncertainties within physics analyses can be found in Ref. [7].

5.3.2 Uncertainties for EMtopo and PFlow jets

Although the scale of individual calibrations may vary between EMtopo and PFlow jets, the final uncertainties are similar in size. A slightly larger pile-up uncertainty contribution in PFlow jets due to the impact of the underlying event is offset by smaller in situ uncertainties, leading to a comparable total overall uncertainty. Figure 24 shows the total uncertainty in EMtopo and PFlow jets for a range of $p_T$ values at fixed $\eta = 0$ and for a range of $\eta$ values at fixed $p_T = 60$ GeV. The level of agreement is representative of other $p_T$ and $\eta$ ranges.

6 Jet energy resolution

Precise knowledge of the jet energy resolution (JER) is important for detailed measurements of SM jet production, measurements and studies of the properties of the SM particles that decay to jets (e.g. $W/Z$ bosons, top quarks), as well as searches for physics beyond the SM involving jets. The JER also affects the missing transverse momentum, which plays an indispensable role in many searches for new physics and measurements involving particles that decay into neutrinos, and thus rely on well-reconstructed missing momentum.

The dependence of the relative JER on the transverse momentum of the jet may be parameterized using a functional form expected for calorimeter-based resolutions, with three independent contributions, namely the noise ($N$), stochastic ($S$) and constant ($C$) terms [54]:

$$\frac{\sigma(p_T)}{p_T} = \frac{N}{p_T} + \frac{S}{\sqrt{p_T}} + C. \quad (4)$$
The noise ($N$) term is due to the contribution of electronic noise to the signal measured by the detector front-end electronics, as well as that due to pile-up. Since both contribute directly to the energy measured in the calorimeter but are approximately independent of the energy deposited by the showing particles, the contribution to the JER scales like $1/p_T$. The noise term is expected to be significant in the low-$p_T$ region, below $\sim$30 GeV. Statistical fluctuations in the amount of energy deposited are captured by the stochastic ($S$) term, which represents the limiting term in the resolution up to several hundred GeV in jet $p_T$. The $S$ term contribution to the JER scales like $1/\sqrt{p_T}$. The constant ($C$) term corresponds to fluctuations that are a constant fraction of the jet $p_T$, such as energy depositions in passive material (e.g. cryostats and solenoid coil), the starting point of the hadron showers, and non-uniformities of response across the calorimeter. The constant term is expected to dominate the high-$p_T$ region, above approximately 400 GeV.

In order to measure the JER, jet momentum must be measured precisely. This implies that the jets must either recoil against a reference object whose momentum can be measured precisely, or be balanced against one another in a well-defined dijet system [5,6]. Measurements using the latter approach are presented here, as well as a method for measuring the contributions to the resolution from the noise term ($N$) due to both pile-up and electronics. The 2017 data, corresponding to an integrated luminosity of 44 fb$^{-1}$ is used for these measurements.

6.1 Resolution measurement using dijet events

Dijet events are both plentiful and produced via a set of $2 \rightarrow 2$ processes that are theoretically well-understood. JER measurements using these events for the dijet balance method rely on the approximate scalar balance between the transverse momenta of the two leading jets. Deviations from exact balance, measured via the asymmetry, given by

$$A = \frac{p_T^{\text{probe}} - p_T^{\text{ref}}}{p_T^{\text{avg}}}$$

are due to a combination of experimental resolution, the presence of additional radiation in the event, and biases due to the event selection used in the measurement. In Eq. (5), $p_T^{\text{avg}}$ is the $p_T$ of a reference jet which is required to be located in a well-calibrated region of the detector ($p_T^{\text{ref}}$, taken here to be the central region of the calorimeter $0.2 \leq |\eta^{\text{ref}}| < 0.7$, where the seam at $\eta^{\text{calo}} = 0$ is excluded to ensure the reference jet energy is as cleanly measured as possible. The probe jet, with transverse momentum $p_T^{\text{probe}}$, may be located either within this central reference region or beyond it, with $|\eta^{\text{probe}}| < 4.5$. The probe jet is the jet for which the resolution is to be measured and $p_T^{\text{avg}}$ is the mean of the probe and reference jet momenta, $p_T^{\text{avg}} = (p_T^{\text{probe}} + p_T^{\text{ref}})/2$. The standard deviation of $A$ for a particular $(p_T^{\text{avg}}, \eta^{\text{probe}})$ bin is denoted by $\sigma_A$, and in the case of a measurement of the probe jet asymmetry may be expressed as

$$\sigma_A^{\text{probe}} = \frac{\sigma_{\text{probe}}}{\langle p_T^{\text{avg}} \rangle} = \left\{ \sigma_{p_T^{\text{probe}}} / p_T \right\} \oplus \left\{ \sigma_{p_T^{\text{ref}}} / p_T \right\}$$

where $\sigma_{p_T^{\text{probe}}}$ and $\sigma_{p_T^{\text{ref}}}$ are the standard deviations of $p_T^{\text{probe}}$ and $p_T^{\text{ref}}$, respectively, and are used to denote the JER for each of the relevant objects. For calibrated jets, $\langle p_T^{\text{probe}} \rangle = \langle p_T^{\text{ref}} \rangle = \langle p_T^{\text{avg}} \rangle$ in the reference region. The reference jet relative resolution, $\sigma_{p_T^{\text{ref}}}/p_T^{\text{ref}}$, must therefore be subtracted from the measured asymmetry distribution in order to extract the resolution of the probe jet as

$$\left\{ \sigma_{p_T^{\text{probe}}} / p_T \right\} = \sigma_A^{\text{probe}} \ominus \left\{ \sigma_{p_T^{\text{ref}}} / p_T \right\}$$

Equation (6) is valid in the probe region as well, up to a correction factor that accounts for the potential overall imbalance between the reference jet and the probe jet in that region. This correction factor is found to be negligible ($<1\%$) for the measurements performed here. However, the $p_T$ balance of the measured jets, and thus the measured asymmetry distribution, is measurably affected on an event-by-event basis by physics effects such as additional radiation, non-perturbative processes including hadronization and multi-parton interactions, and others that may lead to particle losses and addi-
tions in the measured jets. Consequently, the measured dijet balance asymmetry distribution represents a convolution of the intrinsic detector resolution and the particle-level balance affected by the aforementioned effects. The determination of $\sigma_{\Delta pT}^{\text{probe}}$ must therefore account for such effects, for example by subtracting the particle-level quantity from the measured quantity in quadrature:

$$\left(\sigma_{\Delta pT}^{\text{probe}}\right)_{\text{det}} = \left(\sigma_{\Delta pT}^{\text{probe}}\right)_{\text{meas}} \ominus \sigma_{\Delta pT}^{\text{truth}}.$$ 

The results presented here use an iterative fitting procedure to extract the impact of these effects and to isolate the intrinsic detector resolution, $\left(\sigma_{\Delta pT}^{\text{probe}}\right)_{\text{det}}$, by assuming a Gaussian convolution of detector effects with the particle-level balance. First, the asymmetry distribution measured at particle level in MC simulation is fitted with an ad hoc function $A^{\text{truth}}$ based on exponential curves and found to describe it well. Second, the measured asymmetry distribution, $A^{\text{meas}}$, is fitted by the function

$$A^{\text{meas}} = A^{\text{truth}} \otimes R(\mu_{\Delta pT}^{\text{det}}, \sigma_{\Delta pT}^{\text{det}}),$$

taking $A^{\text{truth}}$ from the particle-level fit and where $R(\mu_{\Delta pT}^{\text{det}}, \sigma_{\Delta pT}^{\text{det}})$ is a Gaussian distribution with width $\sigma_{\Delta pT}^{\text{det}}$, representing the detector resolution for the probe jet and offset $\mu_{\Delta pT}^{\text{det}}$ accounting for any residual non-closure in the JES calibration.

Collision data used for the dijet balance measurement are collected using specific combinations of central and forward jet triggers for each of the 11 $p_T^{\text{avg}}$ ranges used in the measurement. Trigger selections are required to be at least 99% efficient in the range of $p_T^{\text{avg}}$ in which a particular combination is used. Jets must also pass JVT selection requirements as described in Sect. 5.2.2.

Topology criteria are applied to select well-defined dijet production processes with minimal contributions due to additional radiation or higher-order processes. The azimuthal angle, $\Delta \phi$, between the two leading jets in the event and the maximum $p_T$ of a third jet, $p_T^{\text{max}}$, are constrained by the following two criteria:

$$\Delta \phi (j_1, j_2) \geq 2.7 \text{ rad.,}$$

$$p_T^{\text{max}} < \min (25 \text{ GeV, } 0.25 \cdot p_T^{\text{avg}}).$$

Example asymmetry distributions are shown in Fig. 25 in two representative bins of $p_T^{\text{avg}}$ and probe jet $\eta_{\text{det}}$. An iterative Gaussian fit to the core of the asymmetry distribution is used to extract the JER. The result of the measurement of the relative JER and its systematic uncertainty is shown in Fig. 26 for a single narrow range of $\eta_{\text{det}}$ and as a function of $p_T^{\text{jet}}$. The JER is observed to be slightly underestimated by MC simulation in this central region of the detector.

Systematic uncertainties are dominated by imprecise knowledge of the scale of the jets at low $p_T$, which results in an approximate 1.5% uncertainty at 40 GeV, whereas the non-closure of the dijet balance method itself is largely dominant at higher $p_T$. The non-closure uncertainty is evaluated as the difference between the resolution measured using the in situ procedure applied to MC simulation and the particle-level resolution, $\sigma(R)/R$, where $R = p_T^{\text{reco}}/p_T^{\text{true}}$. Good agreement is found, resulting in an uncertainty in the relative resolution that is approximately 0.4% and generally increases with $p_T$ due to the non-Gaussian jet response. At lower $p_T$ the uncertainties propagated from the JES dominate. The increase in JES uncertainty around 800 GeV is a result of the single-particle uncertainty (see Sect. 5.3): the jet energy scale calibration used for the dijet energy resolution measurement is necessarily based on a smaller dataset than the one presented in this paper, allowing the two measurements to converge simultaneously, and as a result the statistics were lower and the single-particle uncertainty became dominant at a lower $p_T$ value than in Fig. 22a. Additional systematic uncertainties are estimated by varying the analysis cuts and the JVT selection and by comparing the result with one obtained from an alternative MC generator (SHERPA 2.1.1).

6.2 Noise measurement using random cones

Direct estimates of the noise term of Eq. (4) are obtained by measuring the fluctuations in the energy deposits due to pile-up using data samples that are collected by random unbiased triggers. These measurements are performed using the random cones method in which energy deposits in the calorimeter are summed at the energy scale of the constituents in circular areas analogous to the jet area for anti-$k_t$. $R = 0.4$ jets. This approach is adopted due to its ability to account for any non-Gaussian behaviour of the noise contributions to the JER. Two such random cone sums, $p_T^{\text{c1}}$ and $p_T^{\text{c2}}$, are obtained at random $\phi$ values and within opposite $\pm \Delta \eta$ regions and the difference between them, $\Delta p_T^{\text{RC}}$, provides a measure of the random fluctuations of deposited energy. Multiple non-overlapping cones are selected within each event to maximize statistical power; this is demonstrated to cause no bias in the overall result. This random cone balance is given by

$$\Delta p_T^{\text{RC}} = p_T^{\text{c1}} - p_T^{\text{c2}},$$

and the estimated pile-up noise is determined by the central 68% confidence interval of the distribution of $\Delta p_T^{\text{RC}}$, $\sigma_{\text{RC}}$, sampled over many events as a function of both $\eta$ and pile-up levels, as indexed by $\mu$. Specifically, the noise term due to pile-up, $N_{\text{PU}}$, is determined as

$$N_{\text{PU}} = \frac{\sigma_{\text{RC}}}{2\sqrt{2}},$$

where the width of the distribution is divided by 2 to obtain the half-width of the distribution, and by $\sqrt{2}$ to obtain the fluctuations corresponding to just a single random cone. The distribution of $\Delta p_T^{\text{RC}}$ is shown in Fig. 27a. Updates to the
Fig. 25 Asymmetry distribution measured in data and particle-level Pythia8 for PFlow jets in two example $p_T$ and $\eta$ ranges. Error bars represent the statistical uncertainty. a. The measured asymmetry is shown for probe jets with $80 \text{ GeV} < p_{\text{avg}}^T < 110 \text{ GeV}$ in the range $0.2 < |\eta_{\text{probe}}^{\text{det}}| < 0.7$, where the distributions are symmetric by construction. b. The measured asymmetry is shown for probe jets with $300 \text{ GeV} < p_{\text{avg}}^T < 400 \text{ GeV}$ in the range $1.3 < |\eta_{\text{probe}}^{\text{det}}| < 1.8$. In this $\eta_{\text{probe}}^{\text{det}}$ range the distributions can be asymmetric. Two fits are performed iteratively: the particle-level asymmetry is modelled with an ad hoc function which is subsequently convolved with a Gaussian function in order to describe the reconstructed asymmetry. The detector resolution is then extracted from the Gaussian fit parameter.

Fig. 26 a Relative jet energy resolution and b absolute uncertainty in the relative resolution as a function of $p_T$ for PFlow jets in the central region of the detector, measured using the dijet balance method. The resolution in data is shown in black points with error bars indicating statistical uncertainties; the resolution in detector-level simulated events is shown by the blue curve with total systematic uncertainty given by the blue band. The systematic uncertainty is dominated by terms propagated from the JES uncertainty, while additional terms arise from the analysis selection, pile-up rejection (JVT), physics modelling (comparison with alternative generator), and non-closure effects. The bump in uncertainty around 800 GeV comes from the single-particle uncertainty

The energy scale of the noise estimated by $N_{\text{PU}}$ in Eq. (7) is the constituent energy scale and not that of the jets measured in Sect. 6.1. In order to compare the measurement of the noise term $N_{\text{PU}}$ using the random cone method with the JER measured at the fully calibrated scale (e.g. PFlow+JES) a conversion factor is required. The nominal JES calibration factor is used to perform this conversion to the appropriate energy scale. The result is an estimate of the noise due to pile-up that may be directly compared with the measured JER.

A closure test of the random cone measurements is performed by comparing the in situ measurement of the calibrated $N_{\text{PU}}$ with the expectation from MC simulation. Results are reported here for PFlow jets. To isolate the contribution to the JER from pile-up noise in the MC simulation, the
The difference in the random cone sums, $\Delta p_{T}^{RC}$, measured in the central region ($|\eta_{\text{det}}| < 0.7$) in randomly triggered data using PFlow objects. Comparison between the pile-up noise term $N^{\text{PU}}$ determined using the random cone method (black solid circles) and the expectation from MC simulation (orange squares) as extracted from the difference in quadrature of MC simulation with (red downward triangles) and without (blue upward triangles) pile-up. Results are shown at the PFlow+JES energy scale for jets in the central region of the detector ($|\eta_{\text{det}}| < 0.7$).

JER is determined in simulated events both with and without pile-up and a subtraction in quadrature is performed between the extracted resolutions. The two JER determinations in MC simulation events with and without pile-up are shown in Fig. 27b and their quadratic difference is compared directly to the in situ measurement from the random cones method. Each is fitted, as shown by the dotted lines in Fig. 27b: the random cone measurement is fitted with $N/p_{T}$ while the quadratic difference is fitted with $N/p_{T} \oplus S/\sqrt{p_{T}}$ to account for non-negligible stochastic contributions. The non-closure of the method is largely due to the differences in topo-cluster formation sensitivity to pile-up and electronic noise in the presence versus absence of hard-scatter particles, and is taken as a systematic uncertainty in the result. This non-closure uncertainty is the dominant uncertainty in the JER noise term, ranging from approximately 17% in the most central region to 75% in the endcap transition region ($2.5 < |\eta| < 3.2$).

The total noise contribution to the JER includes not just pile-up but also electronic noise, to which the random cones are not sensitive due to the topo-clustering process. To estimate this electronic contribution, a fit is performed to the JER measured in a dedicated MC simulation sample with $\mu = 0$ and the electronic noise term is extracted as $N^{\mu=0}$. The total noise term used in the JER combination is therefore taken to be $N = N^{\text{PU}} \oplus N^{\mu=0}$ and is shown as a function of $\eta$ in Fig. 28 along with its systematic uncertainties. The dominant systematic uncertainty in the random cone measurement of $N^{\text{PU}}$ is the previously discussed non-closure uncertainty, but additional terms arise from varying the quantile of the confidence interval used to extract $\sigma_{\text{RC}}$ and from using a different estimate of the conversion factor to the calibrated JES scale. Two systematic uncertainties apply to $N^{\mu=0}$: a 20% relative uncertainty conservatively estimating the differences in JER between data and MC simulation and an uncertainty due to the fit parameterization and stability. The systematic uncertainties enter the combined JER fit unsymmetrized in $\eta$ but are symmetrized during the statistical combination, and so the one-sided components are symmetrized in Fig. 28 to illustrate their final contribution to the total uncertainty.
Fig. 29 a The relative jet energy resolution as a function of $p_T$ for fully calibrated PFlow+JES jets. The error bars on points indicate the total uncertainties on the derivation of the relative resolution in dijet events, adding in quadrature statistical and systematic components. The expectation from Monte Carlo simulation is compared with the relative resolution as evaluated in data through the combination of the dijet balance and random cone techniques. b Absolute uncertainty on the relative jet energy resolution as a function of jet $p_T$. Uncertainties from the two in situ measurements and from the data/MC simulation difference are shown separately.

6.3 Combination of in situ jet energy resolution

A combined measurement of the JER is obtained by performing a fit to the dijet balance measurements (Sect. 6.1) using a constraint on the noise term ($N$) derived from the random cones measurement and $\mu = 0$ simulation sample (Sect. 6.2). The implementation of this statistical combination is performed in a manner nearly identical to that for the JES (Sect. 5.2.5), propagating uncertainties from the dijet measurement in the same way and using a similar eigenvalue decomposition to reduce the final number of nuisance parameters.

Instead of using polynomial splines to interpolate across $p_T$, the JER combination uses the functional form from Eq. (4). A fit to the dijet measurement data is performed, fixing the noise term to the value measured by the random cone analysis. Dijet measurement uncertainties are taken to be fully correlated between $\eta$ bins. Uncertainties due to the random cones measurements are determined by propagating the noise term uncertainties and repeating the fit with different $N$. These uncertainties are taken to be decorrelated between central ($|\eta| < 2.5$) and forward ($|\eta| > 2.5$) regions.

The resulting combined measurement of the JER for PFlow+JES jets is shown in Fig. 29a. The dijet measurement data points are shown along with the total in situ combination, while the constraint on the noise term derived from random cones and included in that combination is demonstrated by plotting $N/p_T$ and its uncertainties as a separate curve for illustrative purposes. Figure 29b shows the absolute uncertainties on the combined JER measurement. For each value of $p_T^\text{jet}$ and $\eta$, a toy jet is created and the size of each JER nuisance parameter corresponding to it is retrieved and plotted.

Comparisons of the JER measurements for PFlow+JES and EM+JES jets, as a function of both $p_T^\text{jet}$ and $\eta$, are provided in Fig. 30. The fit to the resolution as a function of $p_T$ for the PFlow+JES jets shows an improvement in resolution over EM+JES jets at low $p_T$.

Figure 31 shows the total JER uncertainty in EMtopo and PFlow jets for a range of $p_T$ values at fixed $\eta = 0.2$ and for a range of $\eta$ values at fixed $p_T = 30$ GeV. The level of agreement is representative of other $p_T$ and $\eta$ ranges.

6.4 Application of JER and its systematic uncertainties

In order to ensure that the resolution of the jet energy scale in simulation matches that in data wherever possible, a smearing procedure is recommended. For regions of jet $p_T$ in which the resolution in data is larger than in MC simulation, the simulation sample should be smeared until its average resolution matches that of data. In regions of jet $p_T$ where resolution is smaller in data than in MC simulation, no smearing is performed, since the data should remain unaltered.

JER systematic uncertainties are propagated through physics analyses by smearing jets according to a Gaussian function with width $\sigma_{\text{smear}}$. If $\sigma_{\text{nom}}$ is the nominal JER of the sample, after MC simulation smearing if necessary, and $\sigma_{\text{NP}}$ is the one-standard-deviation variation in the uncertainty component to be evaluated, then:

$$\sigma_{\text{smear}}^2 = (\sigma_{\text{nom}} + |\sigma_{\text{NP}}|)^2 - \sigma_{\text{nom}}^2.$$

Application of JER systematic uncertainties must account for two factors: first, anti-correlations across a single uncertainty component, and second, differences in resolution between data and MC simulation.

Anti-correlation becomes an issue when a single JER component is positive in some regions of phase space and negative in others. To propagate such systematic uncertainties to analyses, smearing should be applied to the simulation when
Fig. 30 The relative jet energy resolution for fully calibrated PFlow+JES jets (blue curve) and EM+JES jets (green curve) a as a function of $p_T$ and b as a function of $\eta$. The fit to the resolution as a function of $p_T$ for the PFlow+JES jets shows an improvement in resolution over EM+JES jets at low-$p_T$.

Fig. 31 Fractional jet energy resolution systematic uncertainty summed across all components for anti-$k_t$, $R = 0.4$ jets a as a function of jet $p_T$ at $\eta = 0.2$ and b as a function of $\eta$ at $p_T = 30$ GeV. The total JER uncertainty is shown for both EM+JES and PFlow+JES jets.

$\sigma_{NP} > 0$ and applied to the data when $\sigma_{NP} < 0$. It should be noted that the nominal data remains unchanged as this applies only to the application of systematic uncertainties. In the case that data statistics are too low to safely smear, pseudo-data may be smeared instead.

Differences in resolution between data and MC simulation are already accounted for by the application of additional smearing to the simulation when the resolution in simulation is better than in data. When the JER is smaller in data, this difference is accounted for by applying its full value as an additional systematic uncertainty:

$$\sigma_{NP} = \sigma_{\text{data}}^{\text{nom}} - \sigma_{\text{MC}}^{\text{nom}}.$$

This term is defined by the dijet asymmetry measurements of Sect. 6.1 and is zero for the central $\eta$ slice shown in Fig. 29b, but for some $p_T$ ranges in more forward detector regions it can be significant. A large value of this uncertainty for PFlow jets at $\eta \sim 3.2$ is the source of the peaks visible in Fig. 31b.

7 Conclusions

The calibration of the jet energy scale and resolution for jets reconstructed with the anti-$k_t$ algorithm with radius parameter $R = 0.4$ is presented. Jets are built from either the energy deposits that form topological clusters of calorimeter cells or a combination of charged-particle tracks and topological clusters. The measurements discussed here use 36–81 fb$^{-1}$ of data recorded with the ATLAS detector during 2015–2017 in $pp$ collisions at a centre-of-mass energy of 13 TeV at the Large Hadron Collider. It is the first full calibration of PFlow jets performed by the ATLAS collaboration, the first jet energy scale measurement in the high pile-up conditions of late Run 2 data-taking, and the first jet energy resolution measurement in 13 TeV data.

A sequence of simulation-based corrections removes the contribution to the jet energy from additional proton–proton interactions in the same or nearby bunch crossings, corrects the jet so that it agrees in energy and direction with particle-level jets and, improves the jet energy resolution.
Any remaining difference between simulation and data is removed with in situ techniques using well-measured reference objects, including photons, Z bosons, and other jets, such that the energy scale of fully calibrated jets is unity within uncertainties. The jet energy resolution is measured in a dijet balance system, and the contribution to the resolution from the noise term due to pile-up and electronics is also measured. The relative jet energy resolution ranges from 0.25 (0.35) to 0.04 for PFlow (EMtopo) jets as a function of jet $p_T$.

Systematic uncertainties in the jet energy scale for central jets ($|\eta| < 1.2$) vary from 1% for a large range of high-$p_T$ jets ($250 < p_T < 2000$ GeV), to 5% at very low $p_T$ (20 GeV) and 3.5% at very high $p_T$ (> 2.5 TeV). The absolute uncertainty on the relative jet energy resolution is found to be 1.5 at 20 GeV decreasing to 0.5 at 300 GeV.

Acknowledgements We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMWF, HGF and MPG, Germany; GSRT, 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; DST/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, CRC and IVADO, Canada; Beijing Municipal Science & Technology Commission, China; COST, ECRD, ERFD, 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 Aristides programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafsson Foundation, 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), NLT1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [55].

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors’ comment: All ATLAS scientific output is published in journals, and preliminary results are made available in Conference Notes. All are openly available, without restriction on use by external parties beyond copyright law and the standard conditions agreed by CERN. Data associated with journal publications are also made available: tables and data from plots (e.g. cross section values, likelihood profiles, selection efficiencies, cross section limits, ...) are stored in appropriate repositories such as HEpdata (http://hepdata.cedar.ac.uk/). ATLAS also strives to make additional material related to the paper available that allows a reinterpretation of the data in the context of new theoretical models. For example, an extended encapsulation of the analysis is often provided for measurements in the framework of RIVET (http://rivet.hepforge.org/). This information is taken from the ATLAS Data Access Policy, which is a public document that can be downloaded from http://opendata.cern.ch/record/413 [opendata.cern.ch].]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Funded by SCOAP³.

References


Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
Faculty of Science, Kyoto University, Kyoto, Japan
Kyoto University of Education, Kyoto, Japan
Research Center for Advanced Particle Physics and Department of Physics, Kyushu University, Fukuoka, Japan
Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina
Physics Department, Lancaster University, Lancaster, UK
Oliver Lodge Laboratory, University of Liverpool, Liverpool, UK
Department of Experimental Particle Physics, Jožef Stefan Institute and Department of Physics, University of Ljubljana, Ljubljana, Slovenia
School of Physics and Astronomy, Queen Mary University of London, London, UK
Department of Physics, Royal Holloway University of London, Egham, UK
Department of Physics and Astronomy, University College London, London, UK
Louisiana Tech University, Ruston, LA, USA
Fysiska institutionen, Lunds universitet, Lund, Sweden
Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France
Departamento de Física Teorica C-15 and CIAFF, Universidad Autónoma de Madrid, Madrid, Spain
Institut für Physik, Universität Mainz, Mainz, Germany
School of Physics and Astronomy, University of Manchester, Manchester, UK
CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
Department of Physics, University of Massachusetts, Amherst, MA, USA
Department of Physics, McGill University, Montreal, QC, Canada
School of Physics, University of Melbourne, Melbourne VIC, Australia
Department of Physics, University of Michigan, Ann Arbor, MI, USA
Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA
B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Belarus
Research Institute for Nuclear Problems of Byelorussian State University, Minsk, Belarus
Group of Particle Physics, University of Montreal, Montreal, QC, Canada
P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, Russia
National Research Nuclear University MEPhI, Moscow, Russia
D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
Fakultät für Physik, Ludwig-Maximilians-Universität München, Munich, Germany
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Munich, Germany
Nagasaki Institute of Applied Science, Nagasaki, Japan
Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan
Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM, USA
Institute for Mathematics, Astrophysics and Particle Physics, Radboud University/NIKHEF, Nijmegen, The Netherlands
Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, The Netherlands
Department of Physics, Northern Illinois University, DeKalb, IL, USA
(a)Budker Institute of Nuclear Physics and NSU, SB RAS, Novosibirsk, Russia; (b)Novosibirsk State University Novosibirsk, Novosibirsk, Russia
Institute for High Energy Physics of the National Research Centre Kurchatov Institute, Protvino, Russia
Institute for Theoretical and Experimental Physics named by A.I. Alikhanov of National Research Centre “Kurchatov Institute”, Moscow, Russia
Department of Physics, New York University, New York, NY, USA
Ochanomizu University, Otsuka, Bunkyo-ku, Tokyo, Japan
Ohio State University, Columbus, OH, USA
Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman, OK, USA
Department of Physics, Oklahoma State University, Stillwater, OK, USA
Palacký University, RCPTM, Joint Laboratory of Optics, Olomouc, Czech Republic
Institute for Fundamental Science, University of Oregon, Eugene, OR, USA
Graduate School of Science, Osaka University, Osaka, Japan
Department of Physics, University of Oslo, Oslo, Norway

Springer
a Also at Borough of Manhattan Community College, City University of New York, New York, NY, USA
b Also at Centro Studi e Ricerche Enrico Fermi, Rome, Italy
c Also at CERN, Geneva, Switzerland
d Also at CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
e Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Geneva, Switzerland
f Also at Departamento de Física de la Universidad Autónoma de Barcelona, Barcelona, Spain
g Also at Department of Financial and Management Engineering, University of the Aegean, Chios, Greece
h Also at Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA
i Also at Department of Physics and Astronomy, University of Louisville, Louisville, KY, USA
j Also at Department of Physics, Ben Gurion University of the Negev, Beer Sheva, Israel
k Also at Department of Physics, California State University, East Bay, USA
l Also at Department of Physics, California State University, Fresno, USA
m Also at Department of Physics, California State University, Sacramento, USA
n Also at Department of Physics, King’s College London, London, UK
o Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia
p Also at Department of Physics, University of Fribourg, Fribourg, Switzerland
q Also at Dipartimento di Matematica, Informatica e Fisica, Università di Udine, Udine, Italy
r Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
s Also at Giresun University, Faculty of Engineering, Giresun, Turkey
t Also at Graduate School of Science, Osaka University, Osaka, Japan
u Also at Hellenic Open University, Patras, Greece
v Also at IJCLab, Université Paris-Saclay, CNRS/IN2P3, 91405 Orsay, France
w Also at Instituto Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain
x Also at Instytut für Experimentalphysik, Universität Hamburg, Hamburg, Germany
y Also at Institute for Mathematics, Astrophysics and Particle Physics, Radboud University/Nikhef, Nijmegen, The Netherlands
z Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria
aa Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
ab Also at Institute of Particle Physics (IPP), Victoria, Canada
ac Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
ad Also at Instituto de Fisica Teorica, IFT-UAM/CSIC, Madrid, Spain
ae Also at Istanbul University, Dept. of Physics, Istanbul, Turkey
af Also at Joint Institute for Nuclear Research, Dubna, Russia
ag Also at Louisiana Tech University, Ruston, LA, USA
ah Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
ai Also at National Research Nuclear University MEPhI, Moscow, Russia
aj Also at Physics Department, An-Najah National University, Nablus, Palestine
ak Also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
al Also at The City College of New York, New York, NY, USA
am Also at TRIUMF, Vancouver, BC, Canada
an Also at Universita di Napoli Parthenope, Naples, Italy
ao Also at University of Chinese Academy of Sciences (UCAS), Beijing, China
* Deceased