ATLAS CONF Note
ATLAS-CONF-2017-062
18th July 2017

Jet reclustering and close-by effects in ATLAS Run 2

The ATLAS Collaboration

The reconstruction of hadronically-decaying, high-p_{T} W, Z and Higgs bosons and top quarks is instrumental to exploit the physics potential of the ATLAS detector in pp collisions at a centre-of-mass energy of 13 TeV at the Large Hadron Collider (LHC). The jet reclustering procedure reconstructs such objects by using calibrated anti-k_t jets as inputs to the anti-k_t algorithm with a larger distance parameter. The performance of these reclustered large-radius jets during LHC Run 2 is studied, and compared with that of trimmed anti-k_t large-radius jets directly constructed from locally calibrated topological clusters. The propagation of calibrations to reclustered jets is found to be sufficient to restore their average energy and mass scales to particle level. The modelling of small-radius anti-k_t jets in each other's vicinity is studied using methods which combine tracking and calorimeter information. Systematic uncertainties resulting from the propagation of the uncertainties on the input jets to the reclustering procedure are studied. Comparisons between 33.2 fb^{-1} of data collected during 2016 operations and simulation are shown, and the relative jet mass scale and resolution for reclustered and conventional jets are extracted using the forward-folding technique.

1 Introduction

The shower of particles originating from the hadronisation and fragmentation of a strongly-interacting particle produced in a high-energy particle collision is reconstructed experimentally as a jet. At the Large Hadron Collider (LHC), jets are reconstructed using recombination procedures [1, 2] such as the anti-k_t algorithm [3], Cambridge-Aachen algorithm [4, 5] and k_t algorithm [6, 7]. Jets may be built using a collection of four-vectors which corresponds to a set of simulated particles, tracks of charged particles in the inner detector (section 2), or calibrated topological clusters built from energy depositions in the calorimeter [8]. Recombination algorithms accept one user-determined distance parameter, R, which determines the ultimate size of the output jets. Jets built from topological clusters using the anti-k_t algorithm and $R = 0.4$ are widely-used by ATLAS, and are supported by a complex energy- and mass-calibration procedure and corresponding suite of systematic uncertainties [9]. 'Conventional' large-R anti-k_t jets built directly from calorimeter clusters with $R = 1.0$ are also supported with their own dedicated energy and mass calibrations and uncertainties [10], and are often used to reconstruct high-transverse-momentum (high-p_{T}), hadronically-decaying W, Z and Higgs bosons and top quarks ("boosted" bosons and top quarks).

The use of calibrated, small-R jets as inputs to large-R jet reconstruction has been applied in many ATLAS searches for supersymmetry (SUSY) [11–17] and exotica [18–20] throughout Runs 1 and 2 of the LHC. This procedure, known as jet reclustering [21], propagates calibrations and uncertainties from the input jets through to the final large-R jets, exploiting our detailed understanding of the input objects in order to provide greater flexibility at the physics analysis level. The performance of reclustered and conventional large-R jets is directly compared in these studies.

The propagation of calibrations and uncertainties to reclustered jets eliminates the need to prepare a specific calibration for the large-R jets considered for use in the final data analysis. This allows individual searches and measurements to optimise the large-R jet distance parameter and the details of the grooming procedure, i.e., the dedicated algorithm to remove soft and collinear radiation from jets, provided that the input jets used for reclustering are themselves not modified. As only specific large-R jet sizes and grooming configurations are supported by ATLAS (currently, only anti-k_t $R = 1.0$ jets trimmed with f_{cut} =5% and R_{sub} =0.2 [10, 22]), this can lead to large gains in analysis sensitivity depending on the final state in question. When searching for high-mass resonances or massive supersymmetric particles decaying into high-p_{T} W, Z and Higgs bosons and top quarks, a jet built with $R = 1.0$ may not always be optimal. Reclustered jets with radii as small as $R = 0.8$ and as large as $R = 1.4$ have been previously used in ATLAS searches. Some care must be taken when discussing the finer details of jet substructure in the context of reclustered jets, particularly when defining observables which are sensitive to soft radiation within a large-R jet, such as the N-subjettiness [23, 24]. As such, a discussion of substructure observables other than the calorimeter-based jet mass is beyond the scope of the studies presented here.

The calibrations and uncertainties applied to the anti-k_t $R = 0.4$ jets used as inputs when reclustering are based on analyses of well-understood, isolated systems of dijet, multijet, γ+jet and Z+jet systems [9]. As these systems do not contain large amounts of additional activity near the selected jets, it is possible that calibrations and uncertainties derived in these topologies will break down as the amount of radiation (hadronic activity) near a jet increases. As dense systems of many close-by jets are often reclustered when reconstructing a massive hadronically decaying particle, it is necessary to understand these potential limitations. The applicability of jet calibrations intended for a single jet on this merged object in the proximity of additional radiation was previously unclear, and is a focal point of the studies presented here.

Figure 1 provides a specific example of a scenario where close-by effects may be important to consider, for the case of boosted top quarks. A toy two-step decay chain with scalar particles $X \to YZ(\to AB)$ illustrates the qualitative behavior similar to a boosted top quark, since $m_X = 175$ GeV $\approx m_t$, $m_Z = 80$ GeV $\approx m_W$ and $m_Y = m_A = m_B = 0$. The top-quark-like particle X is in a fully resolved regime when all of the particles are further apart than the small-radius jet radius, $R = 0.4$. As the transverse momentum of the mother particle X increases, the number of decay products which are separated by more than $\Delta R = 0.4$ decreases. For intermediate values of p_{T} (600-1200 GeV), two of the three particles produced by the decay of particle X are likely to have merged into a single $\Delta R = 0.4$ region ('jet') with significant mass which is also usually within $\Delta R < 1.0$ of the remaining parton.

This note is a collection of studies associated with the procedure of jet reclustering in ATLAS, performed during 2016 datataking at the LHC. A brief overview of the ATLAS detector is provided in section 2, and a summary of the data and simulated event samples used in the various studies is provided in section 3. Section 4 details the general performance of reclustered large-R jets and draws comparisons with conventional trimmed anti-k_t large-R jets directly constructed from topological clusters. Section 5 presents studies of the modelling of jets in the vicinity of nearby radiation. A breakdown of the effects of systematic uncertainties propagated from the input jets to reclustered jets is shown in section 6. Comparisons between data and simulation, and other *in situ* studies are shown in section 7.

2 The ATLAS detector

ATLAS[1] is a cylindrical, multi-purpose detector with nearly 4π coverage in solid angle, built around interaction point 1 of the LHC at CERN [25]. The inner detector (ID) provides measurements of charged particle momenta up to psuedorapiditites of $|\eta| < 2.5$ using silicon pixel and microstrip detectors. A transition radiation tracking (TRT) detector provides enhanced electron identification capabilities within $|\eta| < 2.0$. A thin superconducting solenoid surrounds the ID, immersing it within a constant 2 T axial magnetic field.

The ATLAS electromagnetic calorimeter resides between the solenoid and its cryostat vessel wall. It utilises lead absorbing and liquid argon sampling layers to contain and measure showers of particles occurring within pseudorapidities of $|\eta| < 3.2$, at depths between 24 – 27 radiation lengths (X_0). Beyond the cryostat vessel wall, the ATLAS scintillating-tile/steel hadronic tile calorimeter provides energy measurement and shower containment for hadronic activity within $|\eta| < 1.7$. The end-cap regions ($1.5 < |\eta| < 3.2$) of the detector are instrumented with liquid argon calorimeters with either stainless steel or copper absorbing layers, respectively in the electromagnetic and hadronic endcap calorimeter modules (EMEC, HEC). The forward calorimeter provides both electromagnetic and hadronic energy measurements beyond the coverage of the EMEC and HEC, extending coverage from $|\eta| > 3.2$ to $|\eta < 4.9|$ with liquid argon systems utilising either copper or tungsten absorbers, depending on the calorimeter layer. Altogether, the combined ATLAS calorimetry systems provide containment for hadronic activity to shower depths greater than 10 hadronic interaction lengths (λ_0) across nearly the entire detector acceptance (up to $|\eta| < 4.9$).

---

[1] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.

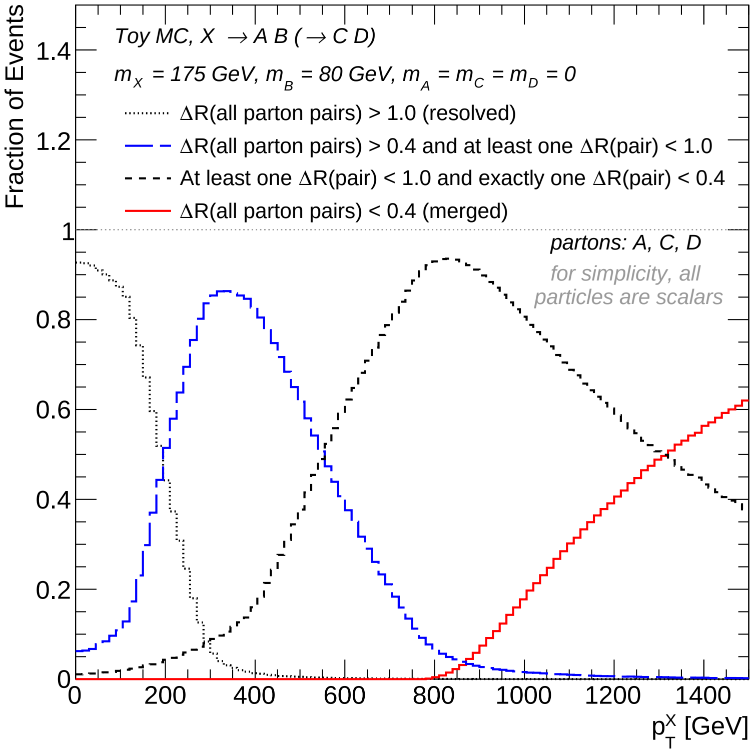


Figure 1: A scalar particle (X) decays to two particles A and B. The particle B subsequently decays into particles C and D. To resemble a boosted top decay, $m_X = 175$ GeV $\approx m_t$ (the mass of the top quark), $m_B = 80$ GeV $\approx m_W$ (the mass of the W boson), and $m_A = m_C = m_D = 0$. Plotted is the fraction of events that are resolved, merged and in-between resolved and merged, as a function of p_T of particle X. The intermediate regime in which reclustering interpolates between single and multiple resolved jets is shown as the broken black distribution: in this region, close-by effects are important to consider as there is a massive small-radius jet very close to another jet. The intermediate regime before the dashed line has all three particles close enough to form one large-radius jet, but still sufficiently far apart that three small-radius jets are resolved inside the large-radius jet.

The ATLAS muon spectrometer is rigged around the tile calorimeter, and contains the three air-cored superconducting toroidal magnets after which ATLAS is named. Three layers of tracking chambers allow the full muon momentum to be reconstructed up to pseudorapidities inside of $|\eta| < 2.7$, and separate, dedicated chambers with a faster readout time allow triggering on muons within $|\eta| < 2.4$.

The ATLAS trigger system [26] consists of a hardware-based Level-1 system, which reduces the rate of event collection from up to 40 MHz to 100 kHz; and a software-based high level trigger (HLT), which processes events in order to further reduce the rate to 1 kHz. Events passing selection in the HLT are written to disk for use in analysis.

3 Data and simulation

The data presented within this note were collected during 2016 operation of the LHC, corresponding to an integrated luminosity of 33.2 fb^{-1} following data quality requirements. During data-taking in 2016, the average number of interactions per bunch crossing was 24.9 and the proton bunch spacing in the LHC was 25 ns. In section 5, a chain of single-jet triggers selects dijet and multijet events with leading-jet transverse momenta as low as 220 GeV, in which the effects of close-by radiation on jet modelling is studied. Within the analysis presented in section 7, $t\bar{t}$ events in the lepton+jets channel are selected using single-muon triggers, which are fully efficient following a requirement that reconstructed muons in these events have a $p_{\mathrm{T}} > 25$ GeV. These events provide a convenient environment to study the properties of high-p_{T}, hadronically decaying W bosons and top quarks *in situ*.

Studies documented within this note are performed using a variety of Monte Carlo (MC) simulated samples produced by ATLAS. Dijet events are generated at LO in Pythia 8.1 [27] with the 2 → 2 matrix element convolved with the NNPDF2.3LO PDF set [28] and using the A14 event tune [29]. The parton shower and fragmentation is simulated in Pythia 8.1 using the leading-logarithmic approximation through p_{T}-ordered parton showers [30]. Additional dijet events are simulated using different generators, in order to provide several comparisions to data when studying the effects of close-by radiation on jets. Sherpa 2.1 [31] generates events using multi-leg $2 \rightarrow N$ matrix elements, which are matched to parton showers following the CKKW prescription [32]. These Sherpa events are simulated using the CT10 PDF set [33] and default Sherpa event tune. Herwig++ 2.7 [34, 35] provides a sample of events where additional radiation is showered using angular-ordering. These events are generated with the 2 → 2 matrix element, convolved with the CTEQ6L1PDF set [36] and configured with the UE-EE-5 tune [37].

In order to study the resolution of jet reclustering in boosted W boson and top quark topologies, exotic W' and Z' bosons with a wide range of masses (from 400 GeV to 5 TeV) are simulated using Pythia 8.1, then forced to decay hadronically, respectively through W and Z bosons or top quarks[2]. This approach provides a source of boosted objects across a wide range of p_{T} for study.

For the *in situ* studies presented in section 7, several additional samples of Standard Model physics processes are used. The details of the $t\bar{t}$ and single top [38], W/Z+jets [39] and diboson [40] samples which are included in these studies may be found in their respective references.

All simulated events have been reconstructed using a full simulation of the ATLAS detector [41] implemented in Geant 4 [42], which describes the interactions of particles with the detector. Hadronic showers are simulated using the FTFP BERT model [43]. The effects of multiple simultaneous pp collisions (pileup) are simulated with unbiased pp collisions using Pythia 8.1 and overlaid on the nominal dijet interactions.

4 MC-based performance

The response of a jet observable such as its transverse momentum (p_{T}) or mass (m) is defined as the ratio of the calorimeter jet observable to that of the associated jet at particle-level, reconstructed with the same jet algorithm:

$$\mathcal{R}_{p_{\mathrm{T}}} = \frac{p_{\mathrm{T}}^{\mathrm{reco}}}{p_{\mathrm{T}}^{\mathrm{true}}}, \tag{1}$$

$$\mathcal{R}_{m} = \frac{m^{\mathrm{reco}}}{m^{\mathrm{true}}}. \tag{2}$$

---

[2] The requirement on hadronic decays is enforced during sample generation for the W' samples and at the analysis level for the Z' samples.

Following calibration, the average response of a reconstructed jet observable for an ensemble of jets should be restored to its average particle-level value. As the input anti-k_t $R = 0.4$ jets have already been calibrated, the large-R jet produced should also be calibrated. Any deviation of the average response value from that at particle-level following calibration is referred to as "non-closure," and could be indicative that additional calibration is necessary. The closures of reclustered jets are compared to those of conventional, trimmed anti-k_t $R = 1.0$ jets built directly from topological clusters, in the dijet topologies which are used to create the calibrations applied to both the small-R and conventional large-R jets.

Conventional ATLAS large-R jets are built directly from topological clusters calibrated at the local hadronic scale [8]. Large-R jets and their substructure are susceptible to the effects of pile-up contamination[3], and so these jets are trimmed by reclustering their input constituents using the k_t algorithm with a distance parameter of $R_{\text{sub}} = 0.2$, then discarding any of the k_t sub-jets which fail to possess at least 5% of the original jet's transverse momentum ($f_{\text{cut}} = 0.05$) [22]. This procedure also improves the mass resolution of large-R jets which reconstruct massive objects, and improves the signal/background performance of the widely-used electroweak boson and top tagging techniques which rely on the jet mass for discrimination power.

Reclustered jets are built from anti-k_t $R = 0.4$ jets with at least 25 GeV of transverse momentum and $|\eta| < 2.5$, built from topological clusters which have been calibrated at the electromagnetic scale [8, 45]. Reconstructed jets are calibrated to particle-level through the application of a jet energy scale calibration (JES) derived from $\sqrt{s} = 13$ TeV collision data and simulation, a global sequential calibration (GSC) used to improve the resolution and flavour response [9], and a jet mass scale calibration (JMS) which is derived using simulated dijet events. These calibrated small-R jets are used as inputs to the anti-k_t algorithm with $R = 1.0$ in order to produce large-R reclustered jets. To follow the procedure applied to standard large-R jets in ATLAS, these reclustered jets are also trimmed by discarding any constituents (here, the input small-R jets) which fail to possess at least 5% of the untrimmed reclustered jet's transverse momentum[4]. The choice of the trimming f_{cut} parameter is also based on the value selected in previous ATLAS results [16, 17]. No additional calibration is applied to the energy or mass of these new objects.

Truth jets are built from particle four-vectors obtained directly from the Monte Carlo generator, before the ATLAS detector simulation is applied. Only the subset of these particles which are stable within the ATLAS detector and which may leave a significant deposition of energy in the calorimeter (particles with $c\tau > 10$ mm which are not muons or neutrinos) are selected for truth jet reconstruction. These stable particles are given as inputs to the anti-k_t algorithm with a distance parameter of either $R = 1.0$ or $R = 0.4$, to produce large- and small-R truth jets. Large-R truth jets are trimmed using the same procedure as the conventional, reconstructed large-R jets and matched to these directly using the criterion that the angular separation (ΔR) between the large-R reconstructed and truth jet be less than 0.6 ($\Delta R < 0.6$). The $R = 0.4$ truth jets are matched within $\Delta R < 0.3$ to the reconstructed small-R jets which form the constituents of large-R reclustered jets. These matched small-R truth jets are used to build 'reclustered' truth jets: the four-vectors of all truth jets matched to constituents of reclustered jets are summed in order to obtain a four-vector which represents the full set of input jets for each reclustered jet. This approach guarantees that the number of reconstructed and truth jet constituents associated with reclustered jets will never differ, which in turn ensures that the jet response will appropriately describe the p_{T} or mass of the jet in question. Events for which the reclustered jet does not include a matched truth-level jet for each reconstructed jet are not considered in these studies. All reclustered and conventional large-R jets are

---

[3] The amount of pileup contamination which may impact a jet scales as the square of the distance parameter R which is used to reconstruct the jet [44].

[4] Note that the p_{T} threshold on the input jets acts as an effective trimming as well.

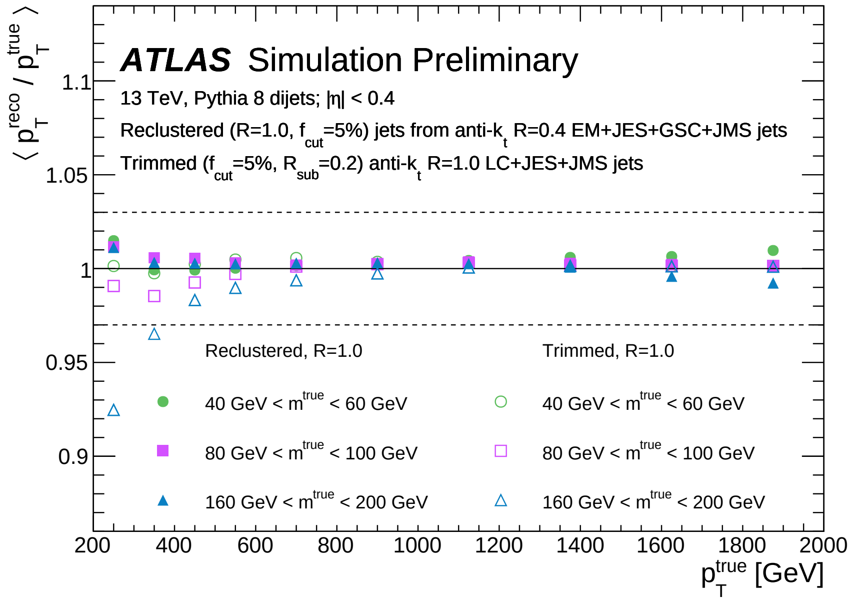
required to be well-isolated at both truth- and detector-level: large-R jets with another large-R truth jet or large-R reconstructed jet within $\Delta R < 2.5$ satisfying $p_{\mathrm{T}} > 150$ GeV and $|\eta| < 2.0$ are not selected for modelling studies. Truth jets may be matched with W bosons or top quark jets by requiring that the truth jet is matched to within $\Delta R < 0.75$ of a hadronically-decaying W boson or top quark at parton-level [46].

Response distributions for both conventional and reclustered large-R jets are produced, then their cores are fit with a Gaussian distribution in the range of $\pm 2\sigma$ around the peak of the distribution. The fitted position of the Gaussian peak is shown as a function of $p_{\mathrm{T}}^{\mathrm{true}}$ and m^{true} in figures 2 and 3 for representative kinematic regions. Good closure is observed for both reclustered and conventional large-R jets in dijet topologies, except occasionally at low jet masses. As the performance of reclustered large-R jets is consistent with that of conventional large-R jets, it may be concluded that the propagation of energy and mass calibrations from the input objects when reclustering is sufficient to restore the average p_{T} and mass values of reclustered jets to their particle-level quantities, and that additional calibration of the large-R reclustered objects is not necessary.

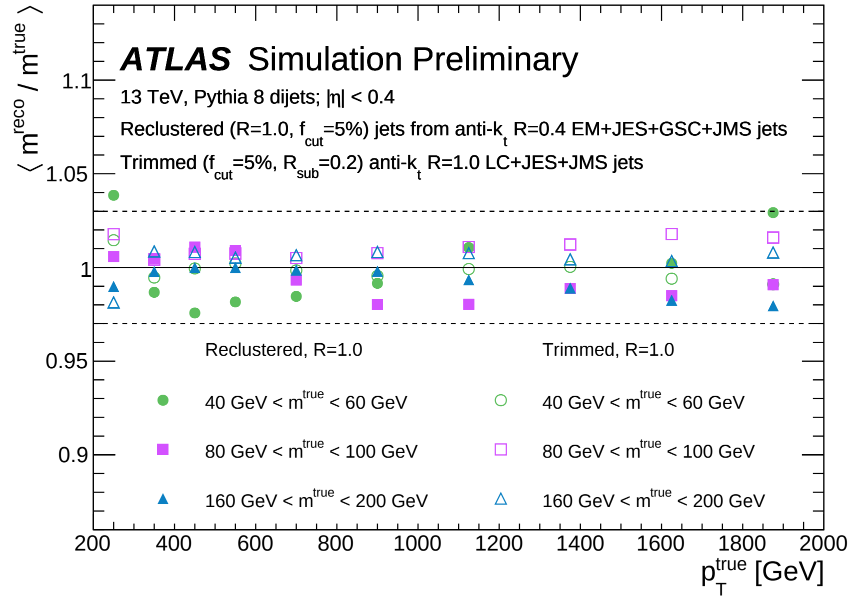
The large-R jet calorimeter mass resolution, defined as the 68% inter-quantile range of the response distribution divided by twice the median of the response distribution[5], is shown in Figure 4 as a function of the truth jet transverse momentum. The mass resolution is shown for high-p_{T} W bosons and top quarks, which produce jets with substantial mass, which may in turn be used to differentiate these jets from other large-R jets produced by light quarks and gluons [46, 47]. A smaller mass resolution implies increased performance when attempting to reconstruct and identify the origin of these objects. Reclustered jets are observed to have a smaller calorimeter mass resolution than conventional trimmed large-R jets for both W and top jets, across nearly the entire examined p_{T} range. This improvement is likely attributable to the separate calibration of each input constituent used to produce reclustered jets, and may be as much as 22% for boosted W bosons and 29% for boosted top quarks within a representative kinematic region ($|\eta| < 0.4$). The large-R jet track-assisted mass [10] is not considered for comparison in these studies, as track-based augmentation of the input anti-k_t $R = 0.4$ jets used in jet reclustering is unavailable at this time.

---

[5] If a jet response distribution is perfectly Gaussian, this measure of resolution simplifies to the ratio of the Gaussian standard deviation to mean of the distribution.

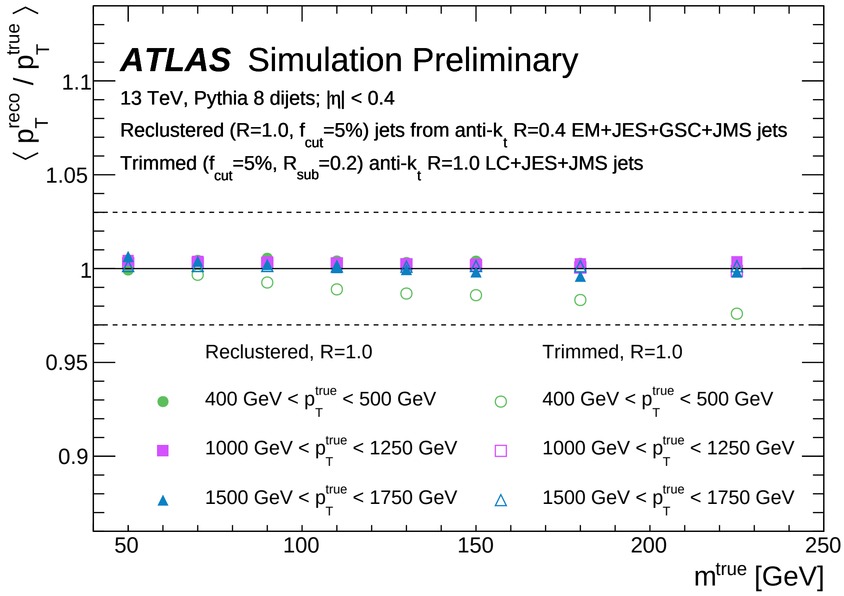


(a)

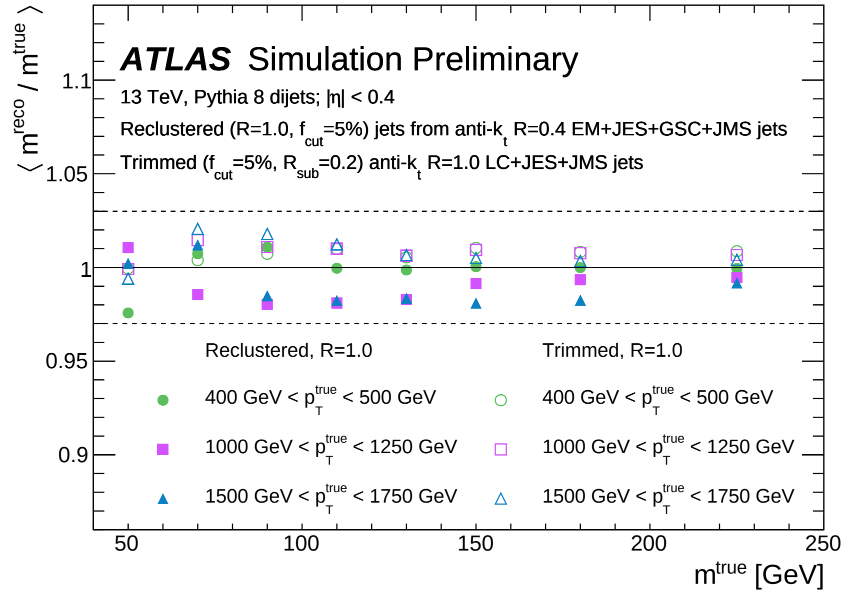


(b)

Figure 2: Reclustered and conventional trimmed large-R jet (a) p_{T} and (b) mass responses, shown as a function of the matched truth jet transverse momentum ($p_{\mathrm{T}}^{\mathrm{true}}$) in representative mass bins. Reclustered jets are built from anti-k_t $R = 0.4$ jets whose energies and masses have been calibrated, which are in turn built from topological clusters calibrated at the electromagnetic scale. Conventional trimmed jets are built from topological clusters calibrated to the local hadronic scale using the anti-k_t algorithm with $R = 1.0$, and subsequently trimmed using $R_{\mathrm{sub}} = 0.2$ and $f_{\mathrm{cut}} = 0.05$. The dashed lines indicate deviations from unity of 3%.

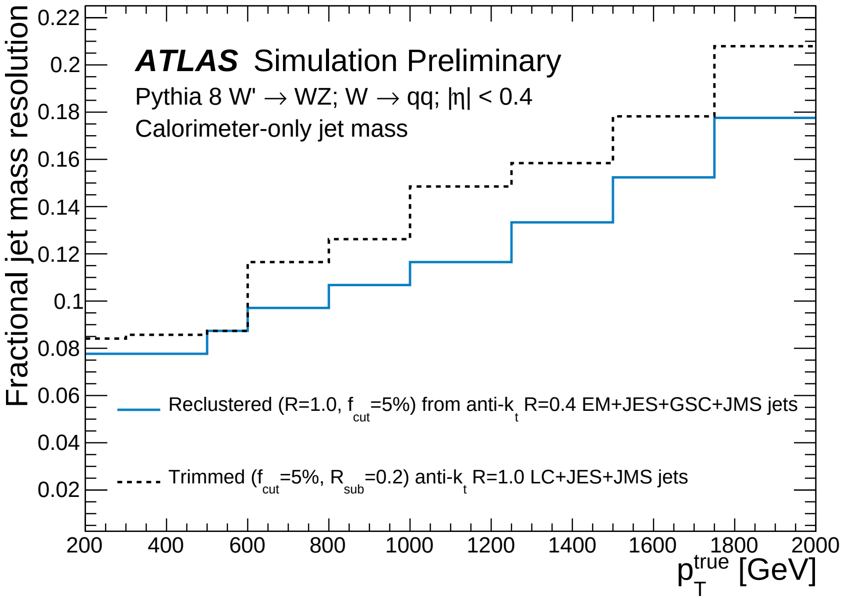


(a)

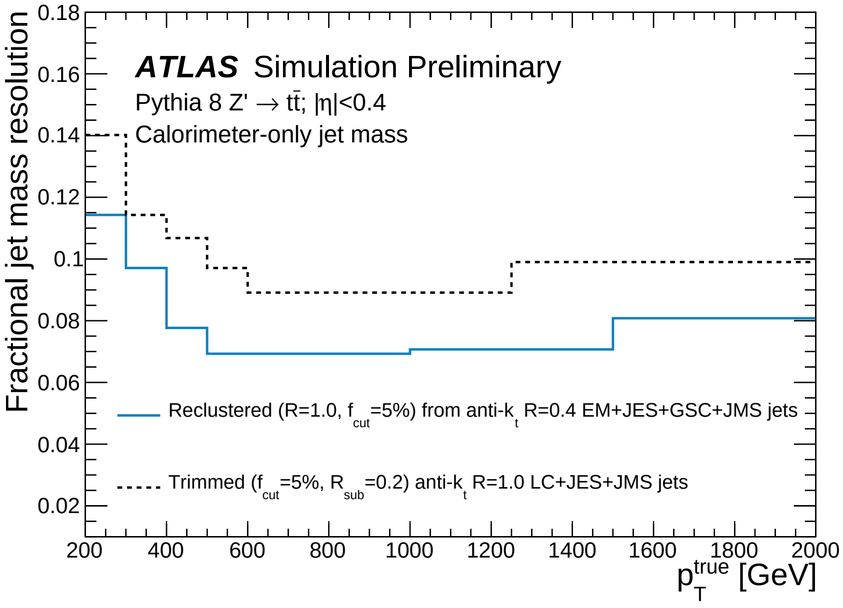


(b)

Figure 3: Reclustered and trimmed large-R jet (a) p_{T} and (b) mass responses, shown as a function of the matched truth jet mass in representative p_{T} bins. Reclustered jets are built from anti-k_t $R = 0.4$ jets whose energies and masses have been calibrated, which are in turn built from topological clusters calibrated at the electromagnetic scale. Conventional trimmed jets are built from topological clusters calibrated to the local hadronic scale using the anti-k_t algorithm with $R = 1.0$, and subsequently trimmed using $R_{\mathrm{sub}} = 0.2$ and $f_{\mathrm{cut}} = 0.05$. The dashed lines indicate deviations from unity of 3%.



(a)



(b)

Figure 4: Reclustered and trimmed large-R jet calorimeter mass resolutions for (a) high-p_{T} W boson jets and (b) high-p_{T} top quark jets within $|\eta| < 0.4$. Reclustered jets are built from anti-k_t $R = 0.4$ jets with at least 25 GeV of p_{T} and whose energies and masses have been calibrated, which are in turn built from topological clusters calibrated at the electromagnetic scale. Conventional trimmed jets are built from topological clusters calibrated to the local hadronic scale using the anti-k_t algorithm with $R = 1.0$, and subsequently trimmed using $R_{\mathrm{sub}} = 0.2$ and $f_{\mathrm{cut}} = 0.05$. No additional calibration is applied to the reclustered large-R jets.

5 Modelling of close-by effects

When constructing reclustered large-R jets, uncertainties and calibrations are propagated from the input anti-k_t $R = 0.4$ jets to the final objects. These calibrations and uncertainties are derived on isolated, well-understood dijet, multijet, γ+jet and Z+jet systems [9]. It is thus important to understand their applicability in the dense systems of radiation which are selected to reconstruct high-p_{T} electroweak bosons and top quarks at the LHC. With enough precision, the standard calibrations and uncertainties may not be sufficient to ensure a good description of jets in data and simulation. Whether there exists a point at which these close-by effects become significant enough that they lead to deviations larger than those covered by existing jet uncertainties has not been previously determined. An investigation of the effects of nearby hadronic radiation on jet modelling is presented below in order to determine whether it is necessary to derive extra uncertainties for the p_{T} and mass of jets in close-by systems during jet reclustering, as a function of nearby activity in the event.

The leading and subleading anti-k_t $R = 0.4$ jets whose modelling in the face of nearby radiation is studied. These jets are the same as those which are taken as the inputs to jet reclustering. The goal of these studies is to determine how robust the energy and mass calibrations applied to these jets are within the dense environments within reclustered jets. Probe jets are required to possess a minimum p_{T} of 220 GeV, a minimum mass of 40 GeV, to be within $|\eta| < 2.1$ and to be matched within $\Delta R < 0.3$ of a truth jet with $p_{\mathrm{T}} > 15$ GeV. The minimum jet mass requirement has been included to ensure that reclustered large-R jets with a single $R = 0.4$ jet constituent and extremely low values of m/p_{T}, which will not be selected by data analyses applying jet reclustering, will also not enter into these studies.

Track jets are built from inner detector tracks using the anti-k_t algorithm with distance parameter $R = 0.4$. These tracks are reconstructed using an iterative algorithm seeded on combinations of measurements from the silicon detectors and combining a combinatorial Kalman fitter with a stringent ambiguity solver [48, 49]. Tracks must have a transverse momentum above 400 MeV, be associated with the primary vertex of the event, pass a requirement on the transverse impact parameter that $\sin\theta \cdot d_0 < 3$ mm, and satisfy the 'loose' quality criteria outlined in Ref. [50]. Track jets with a minimum p_{T} of 15 GeV are matched to probe jets within $\Delta R < 0.4$, in order to construct the track-to-calorimeter jet momentum and mass responses,

$$r_{\mathrm{trk}}^{p_{\mathrm{T}}} = \frac{p_{\mathrm{T}}^{\mathrm{calo}}}{p_{\mathrm{T}}^{\mathrm{trk}}}, \tag{3}$$

$$r_{\mathrm{trk}}^{m} = \frac{m^{\mathrm{calo}}}{m^{\mathrm{trk}}}. \tag{4}$$

Track-to-calorimeter responses are an often-used tool when studying the modelling of jets: the average r_{trk} value of an ensemble of jets is proportional to that of the reco-to-truth response (equations 1 and 2) of the same ensemble, allowing comparisons with data to be performed [51]. These quantities are examined as a function of a pair of close-by metrics, in order to study their modelling. The first metric, the distance between the probe jet and the nearest reconstructed jet passing a minimum p_{T} requirement of 20 GeV is simply given by the ΔR_{min} between the two objects. The second, f_{closeby}, is the projection of the p_{T} of all jets with $p_{\mathrm{T}} > 20$ GeV within some predefined distance R (here taken to be 1.0) on the parallel component of the probe jet axis, normalised to the magnitude of the probe jet three-momentum:

$$f_{\mathrm{Closeby}} = \sum_j \frac{\vec{p}_j \cdot \vec{p}}{|\vec{p}|^2}. \tag{5}$$

All calorimeter jets in the event passing a minimum p_{T} requirement of 20 GeV may be considered when calculating the ΔR and f_{Closeby}; these baseline jets are not required to be associated with a truth jet.

One-dimensional distributions of the ΔR, f_{Closeby}, $r_{\mathrm{trk}}^{p_T}$ and r_{trk}^{m} quantities are shown in Figure 5 for data and simulation, for jets with p_{T} between 220 GeV and 330 GeV. Good agreement is observed in all cases between data and the various generators within the bulk of the r_{trk} distributions. Poor modelling is observed between data and simulation for the f_{Closeby} and ΔR distributions. As these observables are used only to subdivide the r_{trk} responses as a function of close-by activity, this level of agreement is acceptable. The relevant sources of systematic uncertainty on these quantities are those related to the scale and resolution of the jet energy and mass. The jet energy scale uncertainty is calculated by varying a set of strongly-reduced nuisance parameters [9], while the jet mass scale uncertainty includes tracking, modelling and statistical effects. As these systematic uncertainties are largely correlated between different generators, only the central Herwig++ and Sherpa values are shown for comparison.

The median values of the probe jet calorimeter-to-track p_{T} and mass responses in data and simulation are shown in figures 6-9, as functions of ΔR and f_{Closeby}. Excellent agreement between the median r_{trk} value in data and simulation is observed for both p_{T} and mass responses, as a function of both close-by metrics, in each examined p_{T} bin. The double-ratio of the median calorimeter-to-track response between data and simulation provides an estimate of the uncertainty on the calorimeter response. All ratios are found to be consistent with unity within uncertainties (propagated from the input distributions), and so no additional uncertainty is necessary to account for the effects of close-by radiation on jet modelling for input jets during the reclustering procedure.

It is possible that systematic mismodelling of jets due to close-by effects in simulation may manifest itself differently depending on the amount of close-by radiation which is present. In order to study this potential effect, triple ratios sensitive to these differences are defined by:

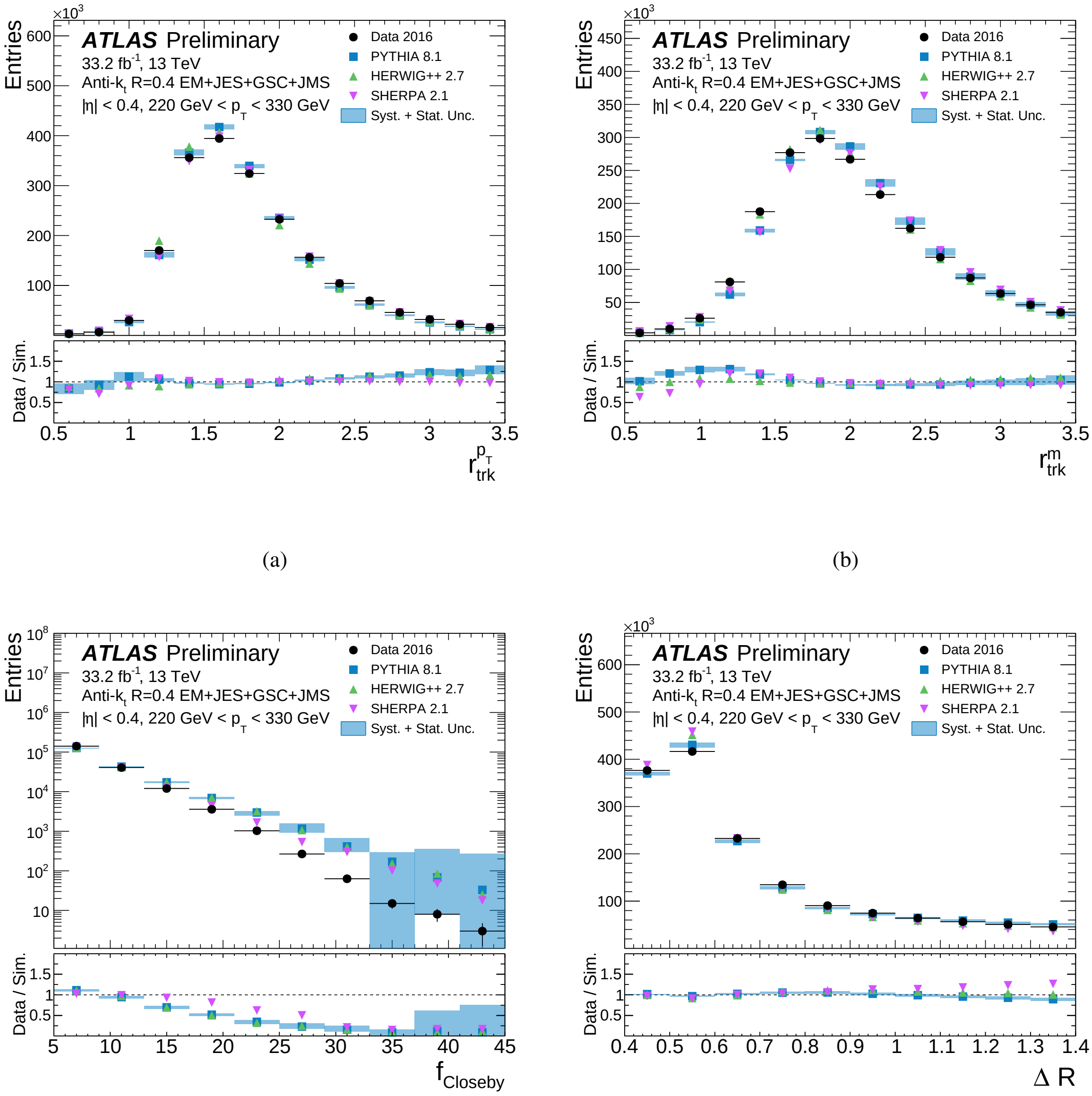
1. Constructing the r_{trk} ratio for a jet's p_{T} or mass.
2. Taking the ratio of the median r_{trk} value in environments with *high* and *low* amounts of close-by activity. High and low are defined respectively as above and below the median f_{Closeby} value from data within a given p_{T} and m/p_{T} bin.
3. Taking the ratio of the previous ratio between data and simulation.

This leads to the observables $R_{\mathrm{triple}}^{p_T}$ and R_{triple}^{m}:

$$R_{\mathrm{triple}}^{p_T} = \frac{\left(\left\{ \frac{p_T^{\mathrm{calo}}}{p_T^{\mathrm{trk}}} \right\}_{f_{\mathrm{closeby}}^{\mathrm{low}}} \Big/ \left\{ \frac{p_T^{\mathrm{calo}}}{p_T^{\mathrm{trk}}} \right\}_{f_{\mathrm{closeby}}^{\mathrm{high}}} \right)_{\mathrm{DATA}}}{\left(\left\{ \frac{p_T^{\mathrm{calo}}}{p_T^{\mathrm{trk}}} \right\}_{f_{\mathrm{closeby}}^{\mathrm{low}}} \Big/ \left\{ \frac{p_T^{\mathrm{calo}}}{p_T^{\mathrm{trk}}} \right\}_{f_{\mathrm{closeby}}^{\mathrm{high}}} \right)_{\mathrm{MC}}}, \tag{6}$$

$$R_{\mathrm{triple}}^{m} = \frac{\left(\left\{ \frac{m^{\mathrm{calo}}}{m^{\mathrm{trk}}} \right\}_{f_{\mathrm{closeby}}^{\mathrm{low}}} \Big/ \left\{ \frac{m^{\mathrm{calo}}}{m^{\mathrm{trk}}} \right\}_{f_{\mathrm{closeby}}^{\mathrm{high}}} \right)_{\mathrm{DATA}}}{\left(\left\{ \frac{m^{\mathrm{calo}}}{m^{\mathrm{trk}}} \right\}_{f_{\mathrm{closeby}}^{\mathrm{low}}} \Big/ \left\{ \frac{m^{\mathrm{calo}}}{m^{\mathrm{trk}}} \right\}_{f_{\mathrm{closeby}}^{\mathrm{high}}} \right)_{\mathrm{MC}}}. \tag{7}$$

These quantities should take a value near unity if close-by effects are consistently modelled as a function of close-by activity. Many experimental systematic effects are also expected to cancel when building



(a)

(b)

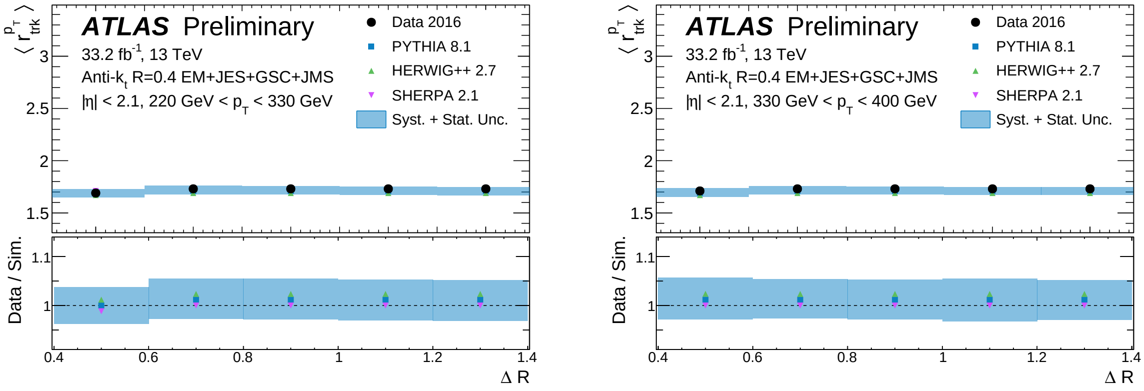
(c)

(d)

Figure 5: Representative distributions of the anti-k_t $R = 0.4$ probe jet (a) $r_{\mathrm{trk}}^{p_{\mathrm{T}}}$, (b) r_{trk}^{m}, (c) f_{Closeby} and (d) ΔR, for jets with p_{T} between 220 GeV and 330 GeV. The energy and mass of these jets have been calibrated, and they are built from topological clusters of calorimeter cells which have been calibrated at the electromagnetic scale. f_{Closeby} and ΔR respectively quantify the amount of activity in the vicinity of a selected jet, and the smallest distance from a selected jet to another reconstructed jet. Each simulated sample is normalised to the integral of the data distribution. Systematic uncertainties arising from the JES, JER and JMS are considered in these studies.

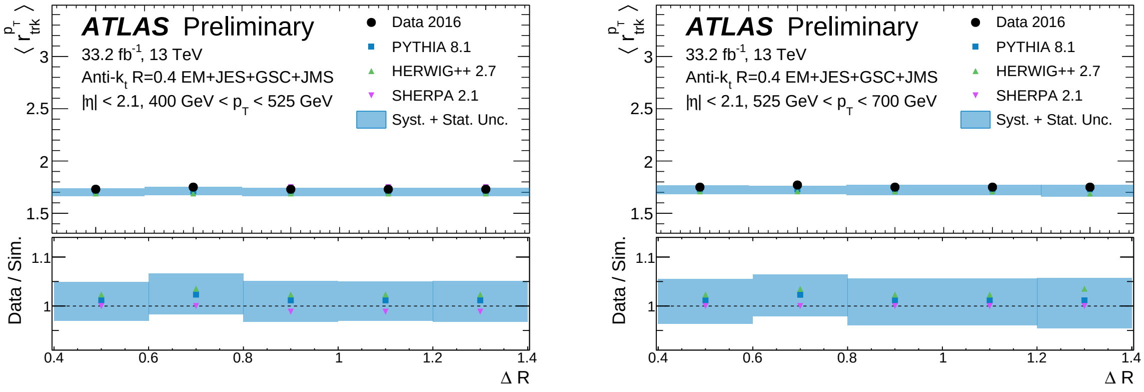
these ratios. Triple ratios of the probe jet p_{T} and mass are respectively presented in figures 10 and 11, as a function of the probe jet m/p_{T}. Agreement with unity is observed within the precision of the uncertainties, propagated from the input distributions to these ratios. Hence, the modelling of close-by effects as a function of the amount of activity is found to be consistently well-handled in Pythia, Herwig++ and Sherpa.

In all cases which have been studied within the context of jet reclustering, the modelling of the anti-k_t $R = 0.4$ jets under study appears robust in the dense environments where significant close-by effects may be present. Pythia, Herwig++ and Sherpa each describe the trends of the median $r_{\mathrm{trk}}^{p_{\mathrm{T}}}$ and r_{trk}^{m} values in data well, as a function of the distance between reconstructed jets and the amount of radiation in the vicinity of a jet. Based on these studies, it is concluded that it is not necessary to apply any additional uncertainties to account for mis-modelling due to close-by effects in the context of jet reclustering.



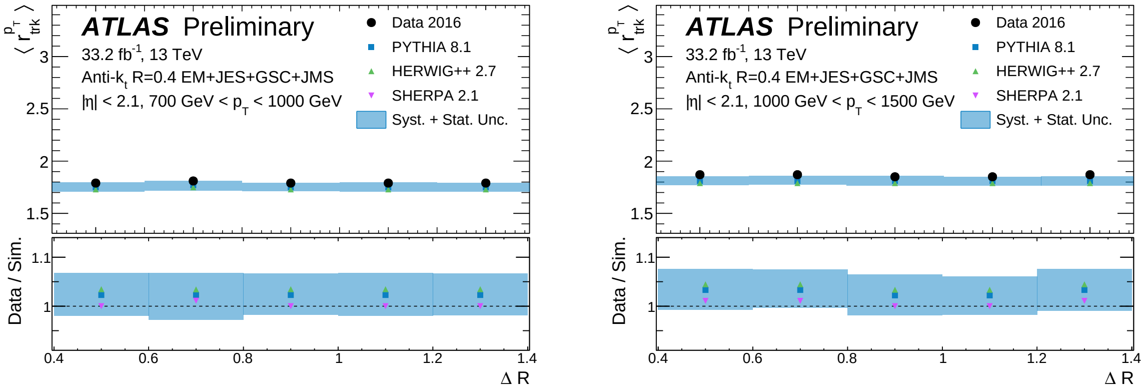
(a) 220 GeV $< p_T <$ 330 GeV

(b) 330 GeV $< p_T <$ 400 GeV



(c) 400 GeV $< p_T <$ 525 GeV

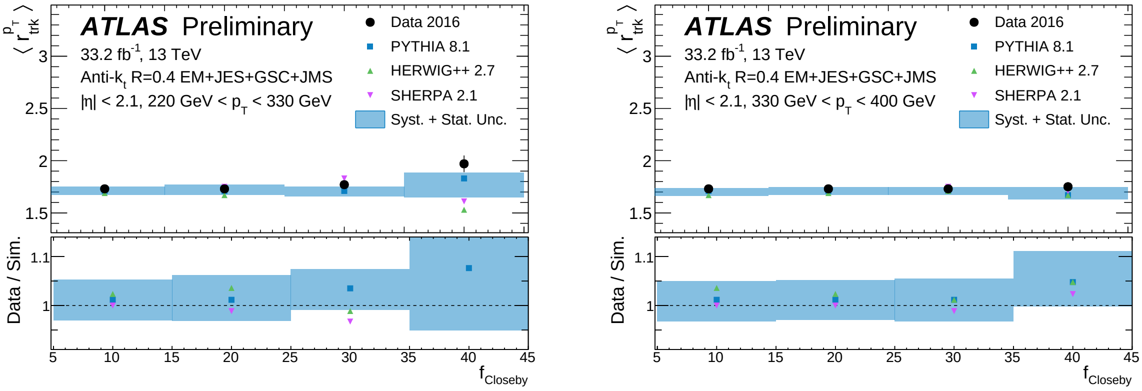
(d) 525 GeV $< p_T <$ 700 GeV



(e) 700 GeV $< p_T <$ 1000 GeV

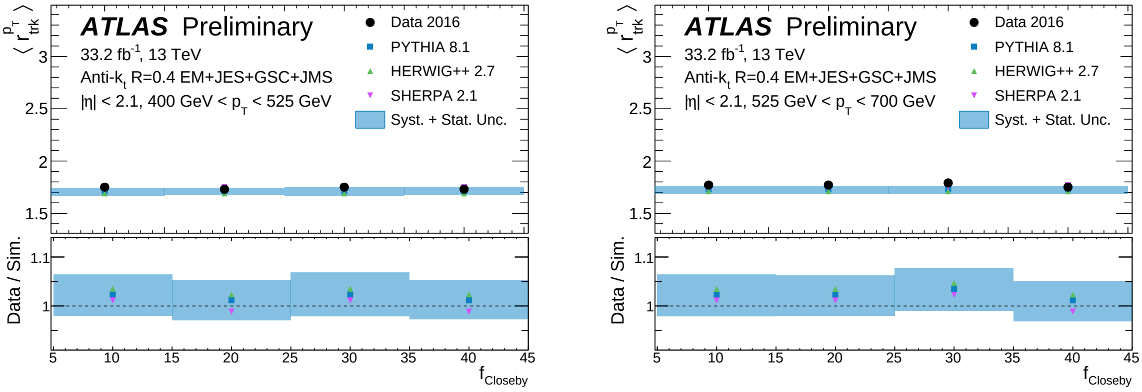
(f) 1000 GeV $< p_T <$ 1500 GeV

Figure 6: The median value of the $r_{\text{trk}}^{p_T}$ response, for anti-k_t $R = 0.4$ probe jets, as a function of ΔR. The double-ratio of $r_{\text{trk}}^{p_T}$ between data and simulation is shown in the lower panel. Statistical uncertainties and systematic uncertainties related to the JES, JER and JMS are shown as a blue band, propagated from the input distributions.



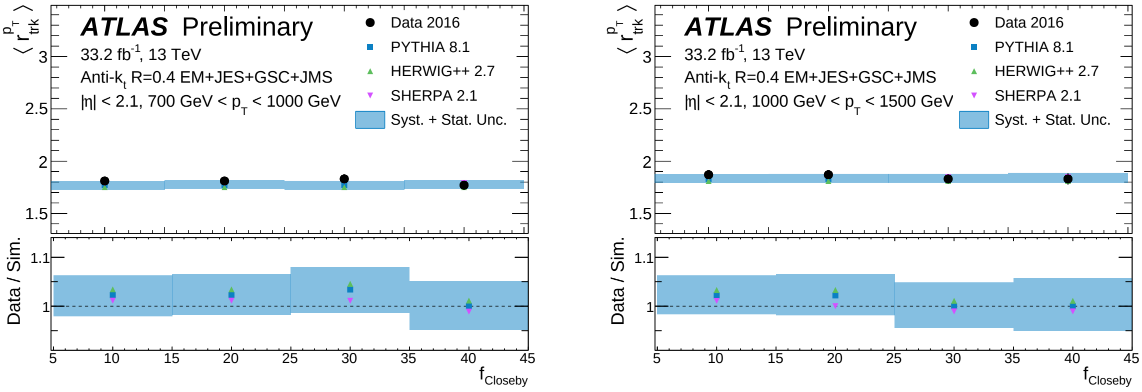
(a) 220 GeV $< p_{\mathrm{T}} <$ 330 GeV

(b) 330 GeV $< p_{\mathrm{T}} <$ 400 GeV



(c) 400 GeV $< p_{\mathrm{T}} <$ 525 GeV

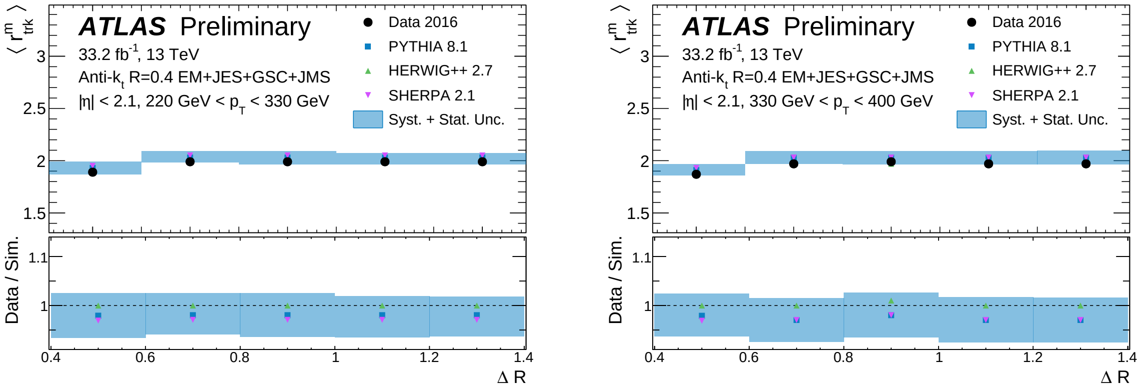
(d) 525 GeV $< p_{\mathrm{T}} <$ 700 GeV



(e) 700 GeV $< p_{\mathrm{T}} <$ 1000 GeV

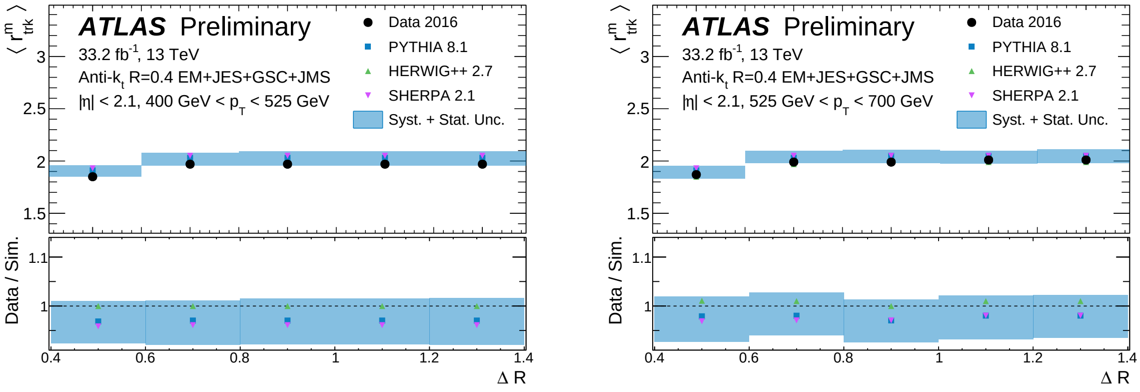
(f) 1000 GeV $< p_{\mathrm{T}} <$ 1500 GeV

Figure 7: The median value of the $r_{\mathrm{trk}}^{p_{\mathrm{T}}}$ response, for anti-k_t $R = 0.4$ probe jets, as a function of f_{Closeby}. The double-ratio of $r_{\mathrm{trk}}^{p_{\mathrm{T}}}$ between data and simulation is shown in the lower panel. Systematic and statistical uncertainties are shown as a blue band, propagated from the input distributions.



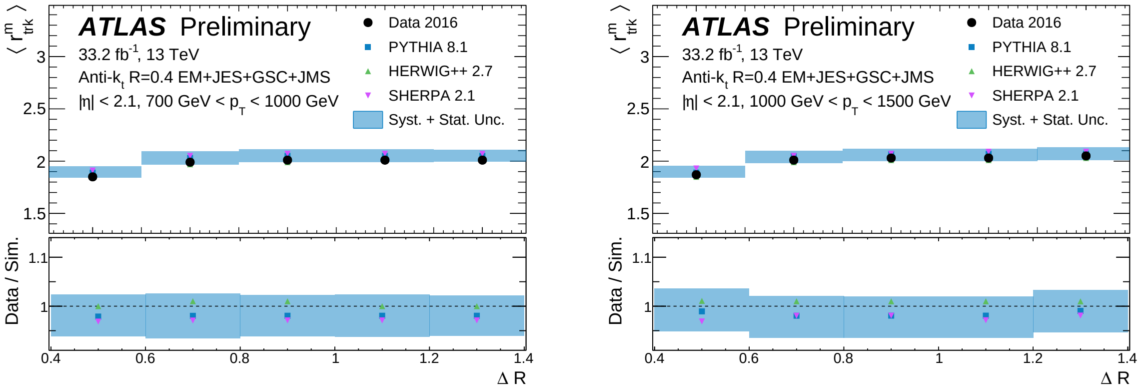
(a) 220 GeV $< p_{\mathrm{T}} <$ 330 GeV

(b) 330 GeV $< p_{\mathrm{T}} <$ 400 GeV



(c) 400 GeV $< p_{\mathrm{T}} <$ 525 GeV

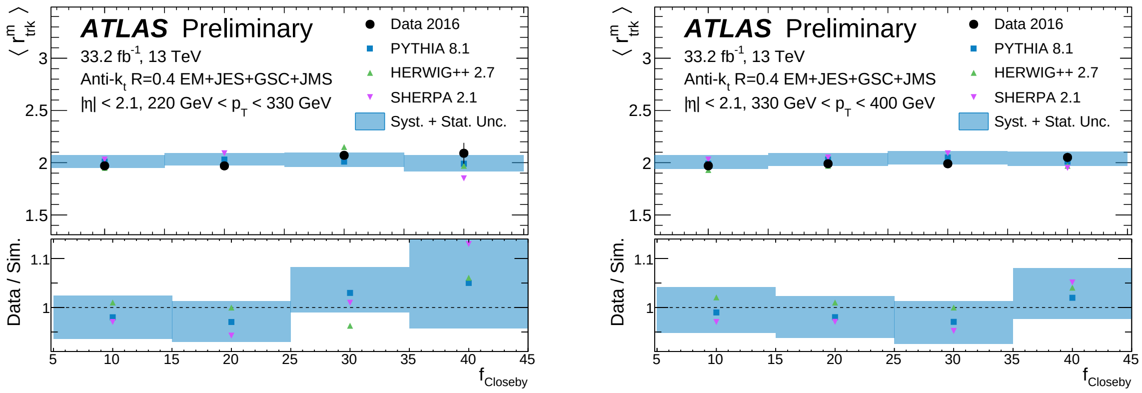
(d) 525 GeV $< p_{\mathrm{T}} <$ 700 GeV



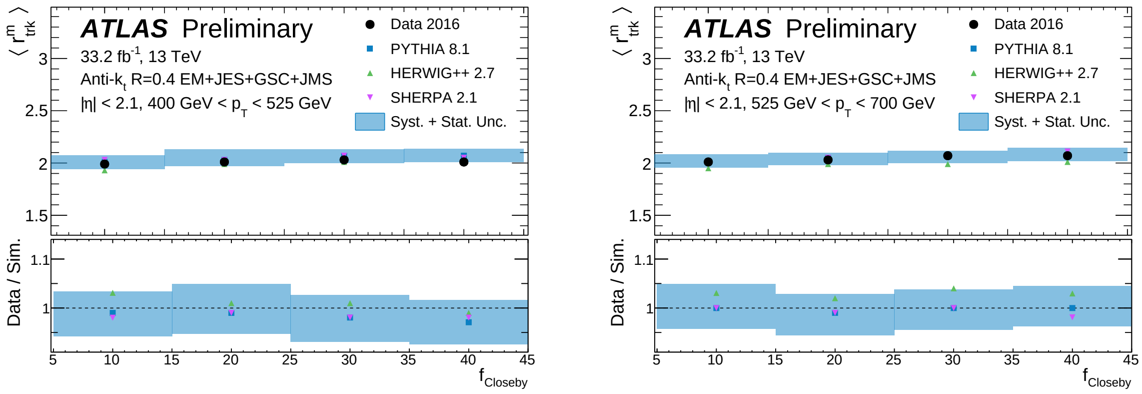
(e) 700 GeV $< p_{\mathrm{T}} <$ 1000 GeV

(f) 1000 GeV $< p_{\mathrm{T}} <$ 1500 GeV

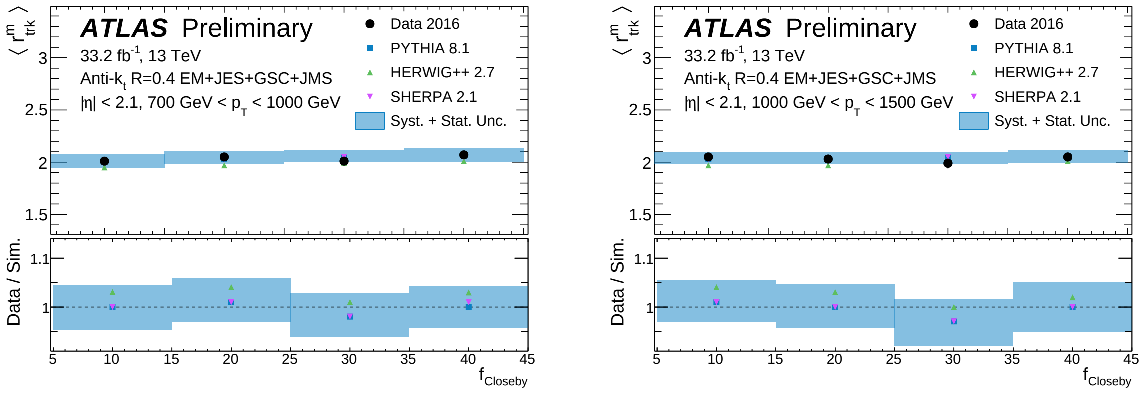
Figure 8: The median value of the r^m_{trk} response, for anti-k_t $R = 0.4$ probe jets, as a function of ΔR. The double-ratio of r^m_{trk} between data and simulation is shown in the lower panel. Statistical uncertainties and systematic uncertainties related to the JES, JER and JMS are shown as a blue band, propagated from the input distributions.



(a) 220 GeV $<$ p_T $<$ 330 GeV (b) 330 GeV $<$ p_T $<$ 400 GeV

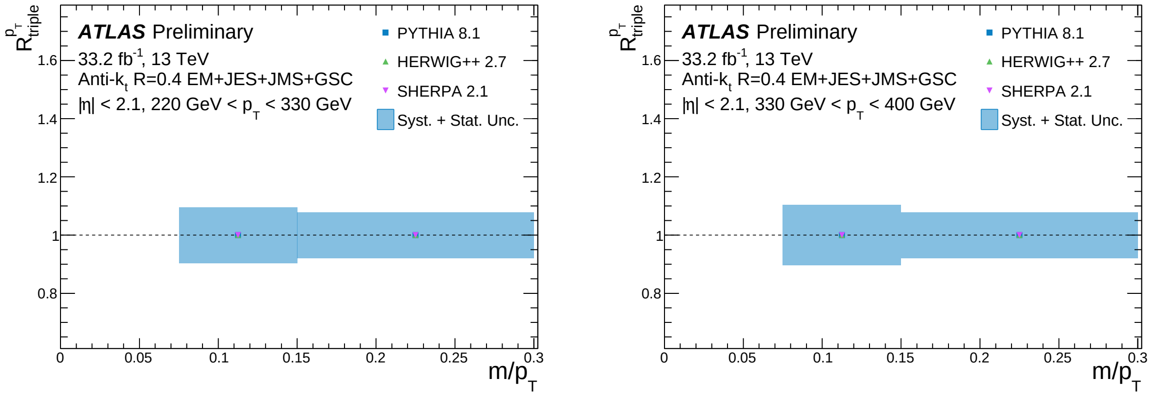


(c) 400 GeV $<$ p_T $<$ 525 GeV (d) 525 GeV $<$ p_T $<$ 700 GeV

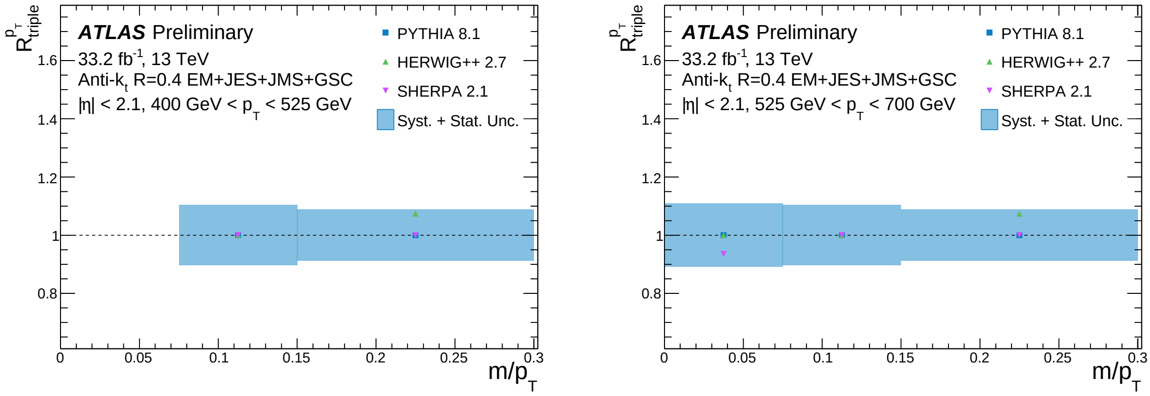


(e) 700 GeV $<$ p_T $<$ 1000 GeV (f) 1000 GeV $<$ p_T $<$ 1500 GeV

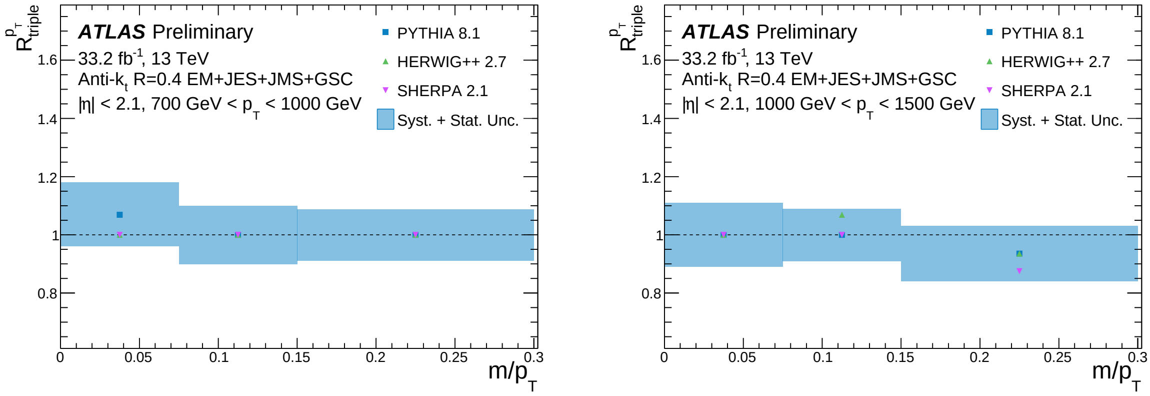
Figure 9: The median value of the r^m_{trk} response, for anti-k_t $R = 0.4$ probe jets, as a function of f_{Closeby}. The double-ratio of r^m_{trk} between data and simulation is shown in the lower panel. Statistical uncertainties and systematic uncertainties related to the JES, JER and JMS are shown as a blue band, propagated from the input distributions.



(a) 220 GeV $< p_{\mathrm{T}} <$ 330 GeV (b) 330 GeV $< p_{\mathrm{T}} <$ 400 GeV

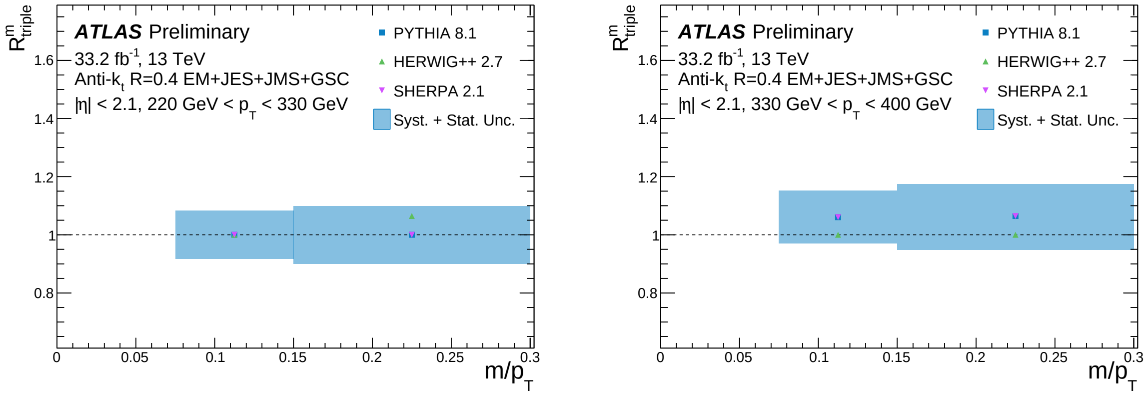


(c) 400 GeV $< p_{\mathrm{T}} <$ 525 GeV (d) 525 GeV $< p_{\mathrm{T}} <$ 700 GeV

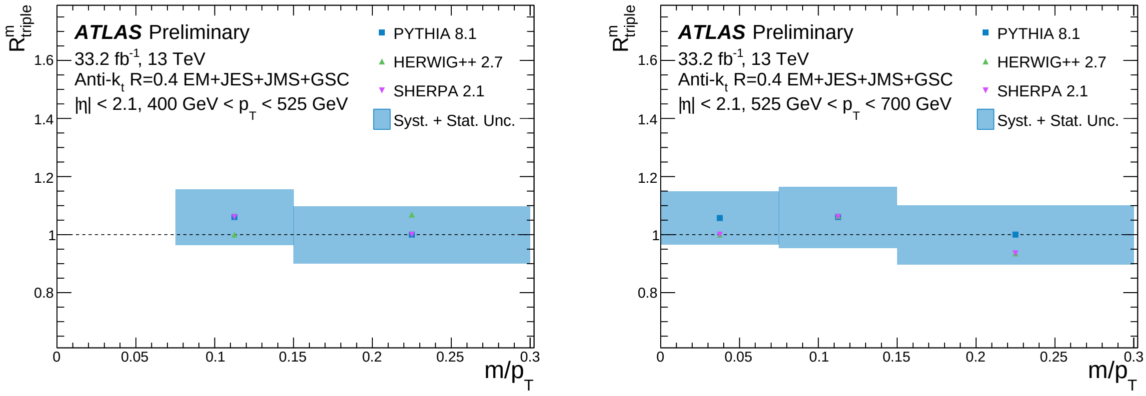


(e) 700 GeV $< p_{\mathrm{T}} <$ 1000 GeV (f) 1000 GeV $< p_{\mathrm{T}} <$ 1500 GeV

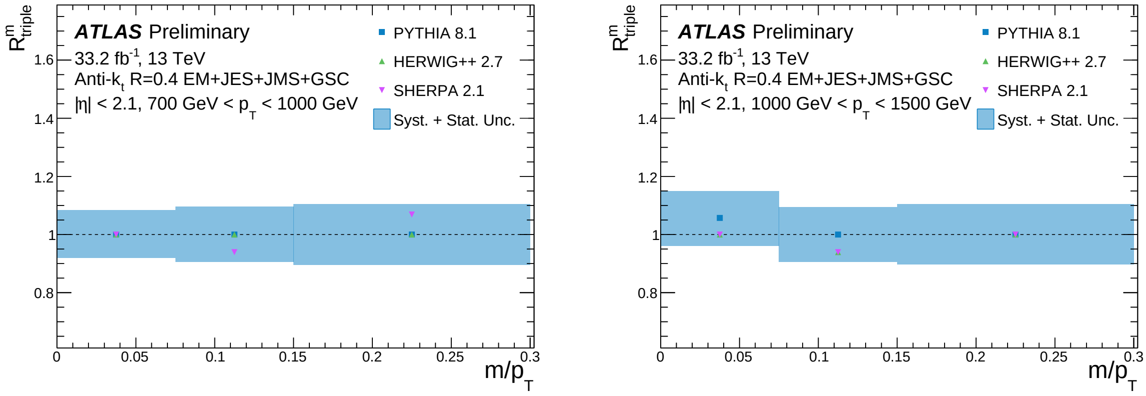
Figure 10: The $r^{p_{\mathrm{T}}}_{\mathrm{trk}}$ triple-ratio (equation 6) values of anti-k_t $R = 0.4$ probe jets, as a function of m/p_{T}. Statistical uncertainties and systematic uncertainties related to the JES, JER and JMS are shown as a blue band, propagated from the one-dimensional input distributions.



(a) 220 GeV $< p_T <$ 330 GeV (b) 330 GeV $< p_T <$ 400 GeV



(c) 400 GeV $< p_T <$ 525 GeV (d) 525 GeV $< p_T <$ 700 GeV



(e) 700 GeV $< p_T <$ 1000 GeV (f) 1000 GeV $< p_T <$ 1500 GeV

Figure 11: The r^{m}_{trk} triple-ratio (equation 7) values of anti-k_t $R = 0.4$ probe jets, as a function of m/p_{T}. Statistical uncertainties and systematic uncertainties related to the JES, JER and JMS uncertainties are shown as a blue band, propagated from the one-dimensional input distributions.

6 Systematic uncertainties

The studies presented in section 5 conclude that there are no significant modelling issues which require additional systematic uncertainties on the response of calibrated anti-k_t $R = 0.4$ jets in the dense topologies selected when reconstructing high-p_{T} electroweak bosons and top quarks using the jet reclustering procedure.

In the absence of additional uncertainties from close-by effects, the jet energy scale (JES) and jet mass scale (JMS) uncertainties on reclustered jets are acquired by propagating the input jet JES and JMS uncertainties through the reclustering procedure. The top row of Figure 12 compares these propagated uncertainties with the uncertainties provided for the conventional trimmed $R = 1.0$ large-R jets. The jet energy scale for small-R jets is well-constrained from a detailed study of *in situ* object balancing analyses including dijets, multijets, γ+jets and Z+jets [9] events, while the JES uncertainty for large-R jets is solely computed using r_{trk} techniques [10]. Due to the more sophisticated chain of *in situ* analyses performed to obtain the uncertainties on the small-R JES, it is not surprising that the JES uncertainty for reclustered jets is smaller than that for conventional trimmed large-R jets across the entire p_{T} range from 200 GeV to 1000 GeV.

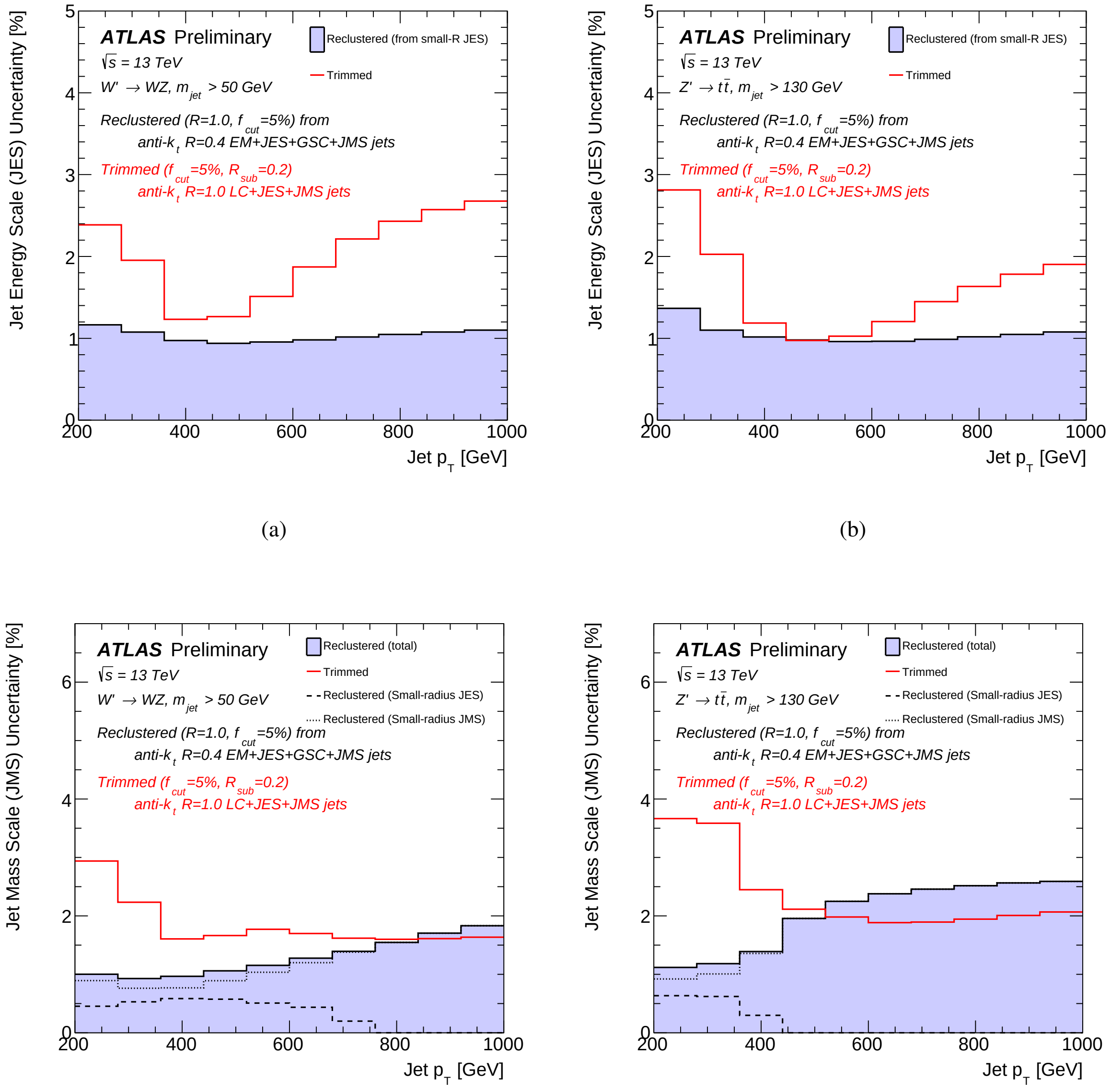
At low-to-moderate p_{T}, when W bosons and top quarks reconstructed by jet reclustering are built from more than a single constituent, their JMS uncertainty is dominated by the JES uncertainty of the individual small-R jets from which they are composed. The JMS uncertainty for reclustered jets is thus smaller than for conventional large-R jets in this regime, as shown in the bottom row of Figure 12. The JMS uncertainty of reclustered jets is smaller than that of conventional jets with transverse momentum below approximately 800 GeV for W boson jets, and 500 GeV for top quark jets. In the high-p_{T} regime, where reclustered jets are more likely to contain a massive small-R jet, the jet mass scale uncertainties propagated from the input jets become dominant and are comparable with the conventional large-R jet uncertainties, since both are computed using the same r_{trk} techniques.

7 In-situ studies

In order to study the modelling of the JES, JER, JMS and JMR in data and MC independently from the sample in which they calibrations were derived, we study hadronic top decays. A sample of high-p_{T} top quarks and W bosons is selected in data and the simulated $t\bar{t}$, single top, W/Z+jets and diboson samples described in section 3 using the approach outlined in Ref. [10]. Single-muon triggers are used to select $t\bar{t}$ events in the lepton+jets channel, which are fully efficient following the requirement of exactly one offline muon which satisfies $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$, as well as additional requirements on its quality and isolation. The $t\bar{t}$ purity of the selection is enhanced by requiring at least 20 GeV of missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$ [52]) and that the sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse mass (m_{T}[6]) of the leptonically-decaying W boson is greater than 60 GeV (*i.e.* $E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}} > 60$ GeV). Events must contain at least one b-tagged jet, identified using the MV2c10 algorithm operating at the 70% tagging effciency working point [53, 54] and at least one trimmed anti-k_t $R = 1.0$ jet with $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2.0$. Finally, at least one anti-k_t $R = 0.4$ jet with $p_{\mathrm{T}} > 25$ GeV must be located within $\Delta R < 1.5$ of the muon. The resulting $t\bar{t}$ purity of this selection is over 70%. Anti-k_t $R = 0.4$ jets with $p_{\mathrm{T}} > 25$ GeV are reclustered, in

---

[6] The transverse mass, m_{T}, is defined as $m_{\mathrm{T}}^2 = 2p_{\mathrm{T}}^{\ell}E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi)$, where $\Delta\phi$ is the azimuthal angle between the lepton and the direction of the $E_{\mathrm{T}}^{\mathrm{miss}}$.



(a)

(b)

(c)

(d)

Figure 12: The jet energy (above) and mass (below) scale uncertainties as a function of p_{T} for reclustered jets and conventional large-R jets from $W' \to WZ$ (left) and $Z' \to t\bar{t}$ (right) events. A requirement that the mass of the large-R jet be > 50 (130) GeV is applied to the plots on the left (right). The reclustered JES uncertainty is removed from the JMS uncertainty for reclustered jets (there is a residual small-radius JES that does still contribute to the JMS).

order to reconstruct the boosted top quark or W boson. These reclustered large-R jets are trimmed using $f_{\mathrm{cut}} = 5\%$.

Following the reclustering procedure, the sample is split into a pair of p_{T} bins based on the conventional large-R jet p_{T}. A low-p_{T} bin from 200 GeV to 350 GeV preferentially contains high-p_{T}, hadronically-decaying W bosons produced by top quarks which are not boosted enough to be fully contained in a single $R = 1.0$ jet. Events where the conventional jet p_{T} is greater than 350 GeV are placed in a second bin, which largely contains hadronically-decaying, high-p_{T} top quarks. The resultant reclustered and conventional trimmed large-R jet mass distributions are shown in Figure 13, in data and simulation. The selected data is consistent with the simulation and its associated systematic uncertainty, which includes the typical JES, JER, JMS and JMR uncertainties as well as systematics related to the modelling of $t\bar{t}$ events [10], and is plotted as a shaded band in this figure.

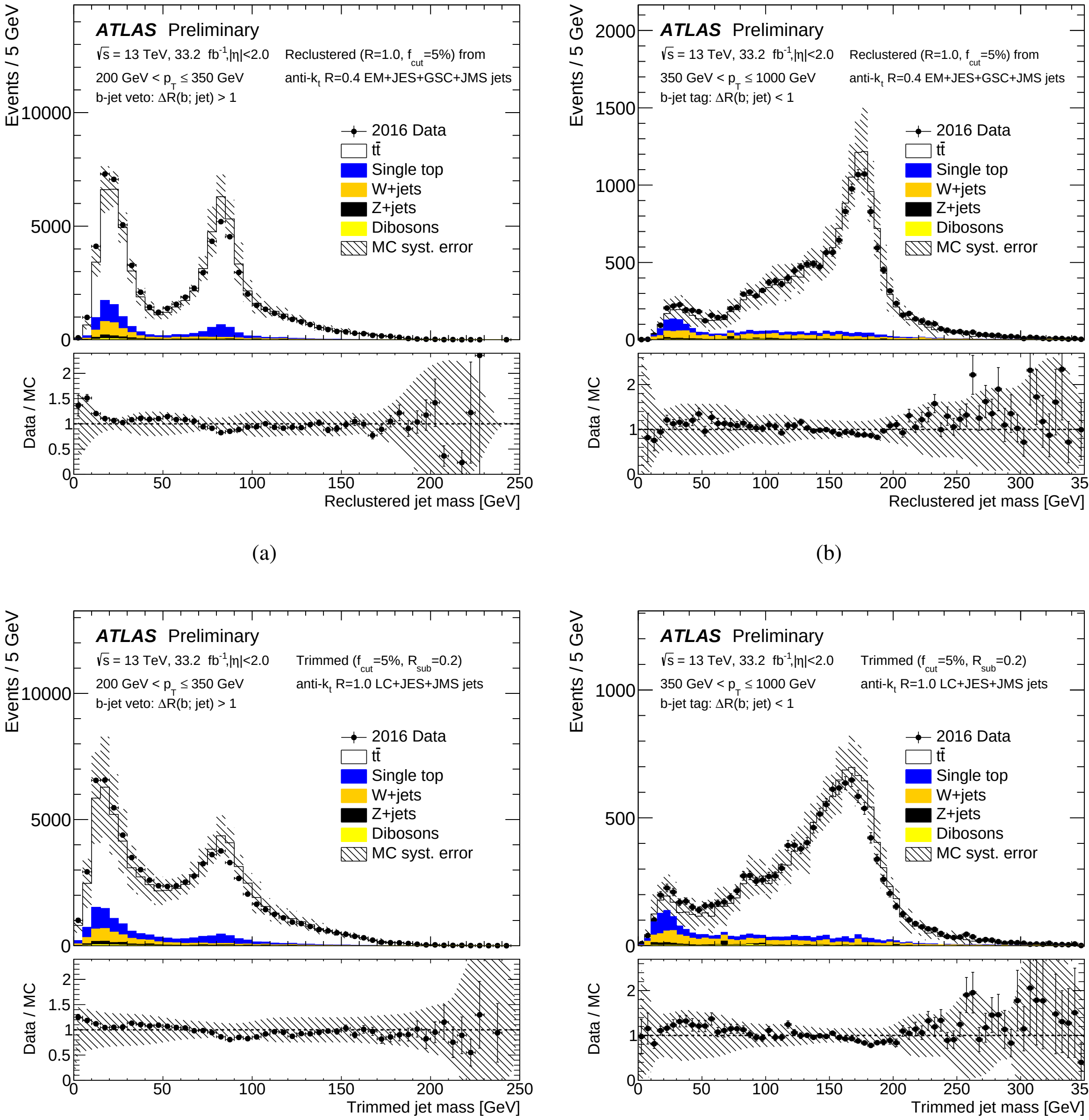
The W mass peaks in the reclustered and conventional trimmed jet mass distributions are separately fit using the forward-folding technique in order to extract the relative *in situ* JMS and JMR [10, 55]. This procedure folds particle-level spectra in simulation by a modified detector response in order to best fit the reconstructed data. The resolution function that describes the transition from a particle-level quantity x^{true} to a calibrated, detector-level quantity x^{reco} is stretched and shifted such that the average value of x^{reco} in a fixed bin $\langle x^{\mathrm{reco}} | x^{\mathrm{true}}, p_{\mathrm{T}}^{\mathrm{reco}} \rangle$ is scaled by s and the resolution is independently scaled by r:

$$\left(x^{\mathrm{folded}} | x^{\mathrm{true}}, p_{\mathrm{T}}^{\mathrm{reco}} \right) = s x^{\mathrm{reco}} + \left(x^{\mathrm{reco}} - \langle x^{\mathrm{reco}} | x^{\mathrm{true}}, p_{\mathrm{T}}^{\mathrm{reco}} \rangle \right) (r - s), \tag{8}$$

where $\left(x^{\mathrm{folded}} | x^{\mathrm{true}}, p_{\mathrm{T}}^{\mathrm{reco}} \right)$ is the folded quantity for a particular jet with some x^{true} and $p_{\mathrm{T}}^{\mathrm{reco}}$. A two-dimensional binned χ^2 fit is performed on the mass distributions to determine the values $s_{\mathrm{data}}^{\mathrm{MC}}$ and $r_{\mathrm{data}}^{\mathrm{MC}}$ such that the folded distribution from simulation best represents the data.

The results of the forward-folding fit are shown in Figure 14 for reclustered and conventional trimmed large-R jets. The relative mass scale is consistent with unity for reclustered jets, while there is a systematic shift observed for conventional trimmed large-R jets. As the calibrated anti-k_t $R = 0.4$ input jets to the reclustering procedure have the full GSC and *in situ* calibrations applied, this is not unexpected. In contrast, the conventional jets have only the MC-based inclusive JES and JMS applied. Dedicated studies of small-radius jets find that their energy scale is lower in data with respect to simulation [9], which is consistent with the jet mass scale for conventional jets in Figure 14. The *in situ* correction applied to anti-k_t $R = 0.4$ jets decreases with p_{T}. This may explain why the scale of conventional top jets is closer to unity than that of conventional W jets, as there is no *in situ* correction available for conventional large-R jets.

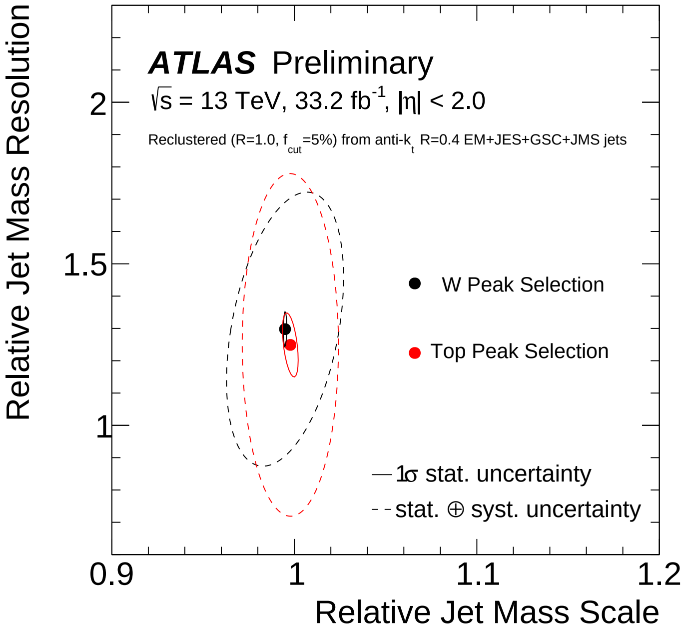
The relative jet mass resolution for reclustered jets is also consistent with unity. Conventional jets have a larger resolution in data than in simulation for both W and top jets. The results for conventional jets are consistent with those in Ref. [10], though there is a slight shift toward a lower scale and a larger resolution. The uncertainties on the relative mass scale and resolution are dominated by theoretical modelling uncertainties and are larger for reclustered jets. This is largely explained by the fact that Figure 14 shows the *relative* jet mass resolution as a multiplicative factor r: $\sigma_{\mathrm{Data}} = r\sigma_{\mathrm{MC}}$. The Gaussian smearing required to match the resolution in data and simulation is given by $\sigma_{\mathrm{smear}} = \sqrt{r^2 - 1}\sigma_{\mathrm{MC}}$. Since the resolution σ_{MC} is smaller for reclustered jets compared with trimmed jets, the amount of smearing required to match the data is actually smaller.



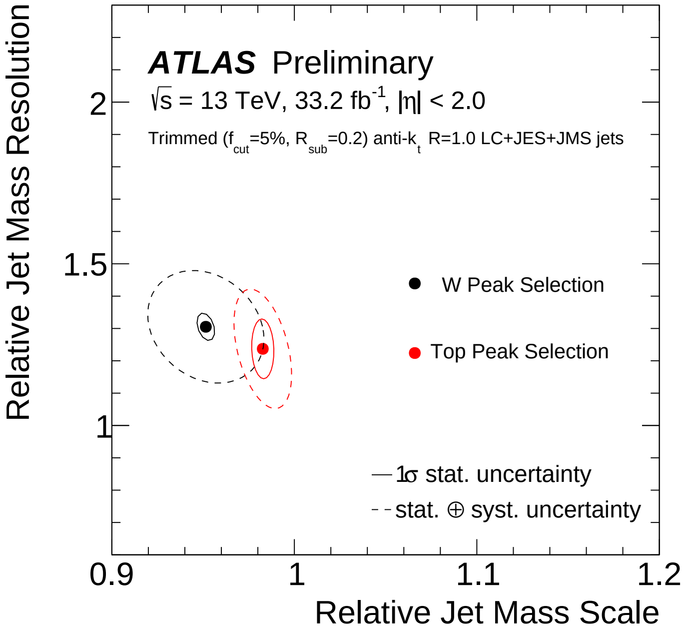
(a) (b)

(c) (d)

Figure 13: The jet mass distributions in data and expected SM simulation for reclustered jets built from $R = 0.4$ jets with the JES+GSC+JMS calibrations applied and for trimmed $R = 1.0$ jets with the appropriate JES+JMS calibrations applied. Events are shown with $p_T > 200$ GeV (left) and $p_T > 350$ GeV (right) using the full 2016 dataset. In addition to the different p_T requirements, a requirement of nearby b-jets is also used to improve the purities. Distributions are shown prior to the forward-folding fit.



(a)



(b)

Figure 14: The fitted values of the (relative) jet mass scale and resolution for JES+GSC+JMS calibrated reclustered (left) and JES+JMS calibrated trimmed jets (right) using the forward-folding method. The values reported use the W and top mass peaks from Figure 13, and the ellipses drawn represent the 2-dimensional 68% confidence intervals.

8 Conclusion

Jet reclustering is a commonly-used technique within ATLAS which uses calibrated anti-k_t $R = 0.4$ jets as inputs to the anti-k_t algorithm with a larger distance parameter R. Common applications of this procedure include the reconstruction of high-p_{T}, hadronically-decaying W, Z and Higgs bosons and top quarks as large-R jets. Reclustering provides the flexibility to optimise the size of large-R jets and the particulars of the grooming procedure which is applied to them, at the cost of more detailed knowledge about substructure observables of the large-R jet other than its calorimeter-based mass. For many ATLAS data analyses, the jet mass is the most discriminating substructure observable, and this compromise is worth making in order to capitalise on the additional benefits reclustering provides. Reclustering serves as a practical alternative to building large-R jets directly from topological clusters of calorimeter cells and producing a dedicated calibration for the particular R value and grooming configuration selected.

Several studies examining the performance and modelling of reclustered jets have been presented. The calibrations propagated from the reclustered input anti-k_t $R = 0.4$ jets are found to restore the average mass and p_{T} values to their particle-level values. Observed non-closure of these values tends to be consistent with the non-closures also observed in conventional trimmed anti-k_t $R = 1.0$ jets. The jet mass resolution of reclustered jets is found to be up to 22% (29%) smaller for high-p_{T} W bosons (top quarks) than the calorimeter mass resolution of conventional trimmed large-R jets.

No significant evidence of systematic mis-modelling of input $R = 0.4$ jet transverse momenta or mass due to close-by effects in the dense topologies characteristic of jet reclustering has been observed in any of the MC generator configurations studied (Pythia 8.1, Herwig++ 2.7 and Sherpa 2.1). Close-by effects in the regime relevant to jet reclustering are well modelled, and an additional uncertainty to account for the mis-modelling of these effects by simulation is not necessary when applying this technique.

The reclustered large-R JES uncertainty propagated from the input jets is found to be smaller than the JES uncertainty on trimmed large-R jets with p_{T} below 1 TeV for both W boson and top quark jets. The reclustered large-R JMS uncertainty is observed to be smaller than that of conventional large-R jets up to 800 GeV for jets from boosted W bosons and 450 GeV for boosted top quark jets.

The relative jet mass scale and jet mass resolution are extracted using a forward-folding procedure for both reclustered and conventional trimmed large-R jets. The relative jet mass scale of reclustered jets is observed to be close to unity, indicating that the input jet calibrations allow for the accurate reconstruction of boosted hadronically decaying particles.

References

[1] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys. Lett. **B641** (2006) 57, arXiv: `hep-ph/0512210 [hep-ph]`.

[2] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, Eur. Phys. J. **C72** (2012) 1896, arXiv: `1111.6097 [hep-ph]`.

[3] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: `0802.1189 [hep-ph]`.

[4] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, JHEP **08** (1997) 001, arXiv: `hep-ph/9707323 [hep-ph]`.

[5] M. Wobisch and T. Wengler, 'Hadronization corrections to jet cross-sections in deep inelastic scattering', *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, 1998 270, arXiv: `hep-ph/9907280 [hep-ph]`.

[6] S. Catani, Y. Dokshitzer, M. Seymour and B. Webber, *Longitudinally-invariant kt-clustering algorithms for hadron-hadron collisions*, Nuclear Physics B **406** (1993) 187 , ISSN: 0550-3213, URL: `http://www.sciencedirect.com/science/article/pii/055032139390166M`.

[7] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. **D48** (1993) 3160, arXiv: `hep-ph/9305266 [hep-ph]`.

[8] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, (2016), arXiv: `1603.02934 [hep-ex]`.

[9] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, (2017), arXiv: `1703.09665 [hep-ex]`.

[10] ATLAS Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, ATLAS-CONF-2016-035, 2016, URL: `https://cds.cern.ch/record/2200211`.

[11] ATLAS Collaboration, *Search for pair production of gluinos decaying via top or bottom squarks in events with b-jets and large missing transverse momentum in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, ATLAS-CONF-2016-052, 2016, URL: `https://cds.cern.ch/record/2206134`.

[12] ATLAS Collaboration, *Search for a Scalar Partner of the Top Quark in the Jets+*$E_{\mathrm{T}}^{\mathrm{miss}}$ *Final State at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, ATLAS-CONF-2016-077, 2016, URL: `https://cds.cern.ch/record/2206250`.

[13] ATLAS Collaboration, *Search for production of supersymmetric particles in final states with missing transverse momentum and multiple b-jets at* $\sqrt{s} = 13$ *TeV proton-proton collisions with the ATLAS detector*, ATLAS-CONF-2017-021, 2017, URL: `https://cds.cern.ch/record/2258143`.

[14] ATLAS Collaboration, *Search for direct pair production of the top squark in all-hadronic final states in proton–proton collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, JHEP **09** (2014) 015, arXiv: `1406.1122 [hep-ex]`.

[15] ATLAS Collaboration, *Search for new phenomena in final states with large jet multiplicities and missing transverse momentum at* $\sqrt{s} = 8$ *TeV proton–proton collisions using the ATLAS experiment*, JHEP **10** (2013) 130, arXiv: `1308.1841 [hep-ex]`.

[16] ATLAS Collaboration, *Search for pair production of gluinos decaying via stop and sbottom in events with b-jets and large missing transverse momentum in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **94** (2016) 032003, arXiv: `1605.09318 [hep-ex]`.

[17] ATLAS Collaboration, *Search for top squarks in final states with one isolated lepton, jets, and missing transverse momentum in* $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, Phys. Rev. D **94** (2016) 052009, arXiv: `1606.03903 [hep-ex]`.

[18] ATLAS Collaboration,
Search for pair production of vector-like top quarks in events with one lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, (2017), arXiv: `1705.10751 [hep-ex]`.

[19] ATLAS Collaboration, *Search for new phenomena in* $t\bar{t}$ *final states with additional heavy-flavour jets in* 3.2 *fb*$^{-1}$ *of pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, ATLAS-CONF-2016-104, 2016, URL: `https://cds.cern.ch/record/2220371`.

[20] ATLAS Collaboration,
Search for pair production of vector-like top partners in events with exactly one lepton and large missing transverse momentum in $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, ATLAS-CONF-2016-101, 2016, URL: `https://cds.cern.ch/record/2217232`.

[21] B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski and C. Wanotayaroj,
Jets from Jets: Re-clustering as a tool for large radius jet reconstruction and grooming at the LHC, JHEP **02** (2015) 075, arXiv: `1407.2922 [hep-ph]`.

[22] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, JHEP **02** (2010) 084, arXiv: `0912.1342 [hep-ph]`.

[23] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP **03** (2011) 015, arXiv: `1011.2268 [hep-ph]`.

[24] J. Thaler and K. Van Tilburg,
Maximizing Boosted Top Identification by Minimizing N-subjettiness, JHEP **02** (2012) 093, arXiv: `1108.2701 [hep-ph]`.

[25] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[26] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: `1611.09661 [hep-ex]`.

[27] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: `0710.3820 [hep-ph]`.

[28] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. **B867** (2013) 244, arXiv: `1207.1303 [hep-ph]`.

[29] ATLAS Collaboration, *ATLAS Pythia 8 tunes to* 7 *TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: `https://cds.cern.ch/record/1966419`.

[30] R. Corke and T. Sjostrand, *Improved Parton Showers at Large Transverse Momenta*, Eur. Phys. J. **C69** (2010) 1, arXiv: `1003.2384 [hep-ph]`.

[31] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: `0811.4622 [hep-ph]`.

[32] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, arXiv: `hep-ph/0109231 [hep-ph]`.

[33] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv: `1007.2241 [hep-ph]`.

[34] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639, arXiv: `0803.0883 [hep-ph]`.

[35] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv: `hep-ph/0011363 [hep-ph]`.

[36] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv: `hep-ph/0201195 [hep-ph]`.

[37] S. Gieseke, C. Rohr and A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. **C72** (2012) 2225, arXiv: `1206.0041 [hep-ph]`.

[38] ATLAS Collaboration, *Simulation of top-quark production for the ATLAS experiment at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2016-004, 2016, URL: `https://cds.cern.ch/record/2120417`.

[39] ATLAS Collaboration, *Monte Carlo Generators for the Production of a W or* Z/γ^* *Boson in Association with Jets at ATLAS in Run 2*, ATL-PHYS-PUB-2016-003, 2016, URL: `https://cds.cern.ch/record/2120133`.

[40] ATLAS Collaboration, *Multi-boson simulation for* 13 *TeV ATLAS analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: `https://cds.cern.ch/record/2119986`.

[41] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: `1005.4568 [hep-ex]`.

[42] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) 250 , ISSN: 0168-9002.

[43] ATLAS Collaboration, *A measurement of the calorimeter response to single hadrons and determination of the jet energy scale uncertainty using LHC Run-1 pp-collision data with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 26, arXiv: `1607.08842 [hep-ex]`.

[44] M. Dasgupta, L. Magnea and G. P. Salam, *Non-perturbative QCD effects in jets at hadron colliders*, JHEP **02** (2008) 055, arXiv: `0712.3014 [hep-ph]`.

[45] P. Speckmayer, T. Carli and C. W. Fabjan, 'Energy Measurement of Hadrons with the CERN ATLAS Calorimeter', Presented on 18 Jun 2008, PhD thesis: Vienna, Tech. U., 2008, URL: `https://cds.cern.ch/record/1112036`.

[46] ATLAS Collaboration, *Identification of high transverse momentum top quarks in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, JHEP **06** (2016) 093, arXiv: `1603.03127 [hep-ex]`.

[47] ATLAS Collaboration, *Identification of Boosted, Hadronically Decaying W Bosons and Comparisons with ATLAS Data Taken at* $\sqrt{s} = 8$ *TeV*, Eur. Phys. J. C **76** (2016) 154, arXiv: `1510.05821 [hep-ex]`.

[48] ATLAS Collaboration, *The Optimization of ATLAS Track Reconstruction in Dense Environments*, ATL-PHYS-PUB-2015-006, 2015, URL: `https://cds.cern.ch/record/2002609`.

[49] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444.

[50] ATLAS Collaboration,
Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ *TeV*,
ATL-PHYS-PUB-2015-018, 2015, URL: `https://cds.cern.ch/record/2037683`.

[51] ATLAS Collaboration, *Performance of jet substructure techniques for large-R jets in proton–proton collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector*, JHEP **09** (2013) 076, arXiv: `1306.4945 [hep-ex]`.

[52] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-023, 2015,
URL: `https://cds.cern.ch/record/2037700`.

[53] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*,
JINST **11** (2016) P04008, arXiv: `1512.01094 [hep-ex]`.

[54] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*,
ATL-PHYS-PUB-2016-012, 2016, URL: `https://cds.cern.ch/record/2160731`.

[55] ATLAS Collaboration, *Measurement of large radius jet mass reconstruction performance at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, ATLAS-CONF-2016-008, 2016,
URL: `https://cds.cern.ch/record/2139642`.