ATLAS NOTE
ATLAS-CONF-2016-035
19th July 2016

Jet mass reconstruction with the ATLAS Detector in early Run 2 data

The ATLAS Collaboration

Abstract

This note presents the details of the ATLAS jet mass reconstruction for groomed large-radius jets. The jet mass scale calibrations are determined from Monte Carlo simulation. An alternative jet mass definition that incorporates tracking information called the track-assisted jet mass is introduced and its performance is compared to the traditional calorimeter-based jet mass definition. Events enriched in boosted W, Z boson and top quark jets are used to directly compare the jet mass scale and jet mass resolution between data and simulation. This in-situ technique is also extended to constrain the jet energy scale and resolution.

ATLAS-CONF-2016-035
19 July 2016

1 Introduction

Jets are collimated streams of particles resulting from the production of high energy quarks and gluons. The distribution of energy inside a jet contains information about the initiating particle. When a massive particle such as a top quark or W/Z bosons is produced with significant Lorentz boost and decays into quarks, the entire hadronic decay may be captured inside a single jet. The mass of such jets (*jet mass*) is one of the most powerful tools for distinguishing massive particle decays from the continuum multijet background; therefore, it is critical to improve the reconstructed jet mass by reducing its resolution (JMR) and any associated systematic uncertainties.

This note documents the reconstruction of the jet mass with the ATLAS detector [1]. Two jet mass definitions are presented, each with strength in specific kinematic regimes. The baseline large-radius jet mass is constructed solely from calorimeter information via the constituents of a jet. A second definition augments the calorimeter information using finer granularity inputs from the tracking detector. Each of these jet mass definitions are described in more detail in Sec. 5, following background information on the ATLAS detector, simulated samples, and object reconstruction respectively in Sec. 2, 3, and 4. After introducing the various methods, their resolutions are quantitatively compared in simulation in Sec. 6 and their systematic uncertainties are documented in Sec. 7. In-situ studies of both jet mass definitions are presented in Sec. 8 and 9, including a measurement of the jet mass and energy scales and resolutions. Methods for improving upon the jet mass definitions are documented in Sec. 10. The note ends with conclusions and future outlook in Sec. 11.

2 ATLAS detector

ATLAS is a multi-purpose particle detector with nearly 4π coverage in solid angle.[1] A lead/liquid-argon sampling electromagnetic calorimeter is split into barrel ($|\eta| < 1.5$) and endcap ($1.5 < |\eta| < 3.2$) sections. A steel/scintillating-tile hadronic calorimeter covers the barrel region ($|\eta| < 1.7$) and two endcap copper/liquid-argon sections extend to higher pseudo-rapidity ($1.5 < |\eta| < 3.2$). Finally, the forward region ($3.1 < |\eta| < 4.9$) is covered by a liquid-argon calorimeter with Cu (W) absorber in the electromagnetic (hadronic) section.

Inside the calorimeters is a 2 T solenoid that surrounds an inner tracking detector which measures charged-particle trajectories covering a pseudorapidity range $|\eta| < 2.5$ with pixel and silicon microstrip detectors (SCT), and additionally covering the region $|\eta| < 2.0$ with a straw-tube transition radiation tracker (TRT). Outside the calorimeter is a muon spectrometer: a system of detectors for triggering up to $|\eta| < 2.4$ and precision tracking chambers up to $|\eta| < 2.7$ inside a magnetic field supplied by three large superconducting toroid magnets.

[1] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (r, 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 $\eta = -\ln\tan(\theta/2)$. Angular distance is measured in units of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

3 Data and Simulated Samples

The 2015 LHC collision data used in later sections gave a mean number of additional proton-proton (pp) interactions per bunch crossing of approximately 14 and a bunch spacing of 25 ns. Following requirements based on beam and detector conditions and data quality, the dataset corresponds to an integrated luminosity of 3.2 fb^{-1}.

Samples of Monte Carlo (MC) simulated events are used for studying the detector response of various mass reconstruction techniques as well as for comparisons with data. Table 1 summarizes the MC samples used for the studies in this note. More details about the $t\bar{t}$ and single top, W/Z+jets, and diboson simulation settings can be found in Ref. [2–5]. QCD multijets, simulated with Pythia 8, are used to derive the calibration factors for the jet mass scale and also to study the jet mass resolution over a wide kinematic range. The exotic production of W' and Z' bosons provides a copious source of simulated boosted W bosons, Z bosons and of top quarks respectively. The signal samples are produced with W' and Z' boson masses ranging from 400 to 5000 GeV. The MC samples are processed through the full ATLAS detector simulation [6] based on Geant 4 [7].

Process	ME Generator	ME PDF	Fragmentation	UE Tune	Cross-section Order
$t\bar{t}$	Powheg-Box v2 [8]	CT10 [9]	Pythia 6 [10]	P2012 [11]	NNLO+NNLL [12–17]
Single top	Powheg-Box	CT10	Pythia 6	P2012	NNLO+NNLL [18–20]
W/Z+jets	Sherpa 2.1.1 [21]	CT10	Sherpa	Default	NNLO [22]
Diboson	Sherpa 2.1.1	CT10	Sherpa	Default	NLO
QCD multijets	Pythia 8 [23]	NNPDF23LO [24]	Pythia 8	A14 [25]	N/A
$W'(\rightarrow WZ)$	Pythia 8	NNPDF23LO	Pythia 8	A14	N/A
$Z'(\rightarrow t\bar{t})$	Pythia 8	NNPDF23LO	Pythia 8	A14	N/A

Table 1: Overview of the simulated samples.

4 Object Definitions

Large-radius jets are formed using FastJet [26] implementation of the anti-k_t jet algorithm [27] with distance parameter $R = 1.0$ and then trimmed [28] by re-clustering the constituents of the jet into subjets using the k_t algorithm with distance parameter $R_{\text{sub}} = 0.2$ and removing any of these subjets with transverse momentum (p_{T}) less than 5% of the original jet p_{T}. Large-radius *calorimeter jets* are built from calorimeter-cell clusters calibrated to the hadronic scale using the local cell signal weighting (LCW) method [29]. The jet energy and pseudorapidity is then further calibrated to account for residual detector effects using energy and pseudorapidity dependent calibration factors derived from simulation [30–32]. The jet mass calibration is the final calibration step for large-radius calorimeter jets and it is described in Sec. 5. All large-radius calorimeter jets are required to have $p_{\text{T}} > 200$ GeV with $|\eta| < 2.0$. Detector-stable simulated particles ($c\tau > 10$ mm), excluding muons and neutrinos, are used as inputs for particle-level *truth jets*. Small-radius jets[2] are formed from calorimeter cell-clusters at the electromagnetic (EM) scale and then corrected on average for the impact of pileup. These jets are then calibrated to the hadronic scale as described in Ref. [31]. A further calibration is applied to data in order to correct for differences in the response measured using in-situ techniques.

[2] Small-radius jets are clustered using the anti-k_t jet algorithm with $R = 0.4$ distance parameter.

Large-radius calorimeter jets are identified as W/Z-jets (top-jets) if there is a truth W/Z-boson (top quark) associated to their untrimmed jet area using the ghost association method [33]. The W/Z-jets (top-jets) are studied in simulated samples of W' boson (Z' boson) events and they are required to have $p_{\mathrm{T}} > 250$ GeV (> 350 GeV).

Tracks are reconstructed in the inner detector using an iterative algorithm seeded on combinations of measurements from the silicon detectors and combining a combinatorial Kalman filter with a stringent ambiguity solver [34]. Reconstructed tracks are required to have transverse momentum of at least 400 MeV, $|\eta| < 2.5$, and have at least seven hits in the pixel and SCT detectors. Not more than one module with such hits is allowed to be shared between multiple tracks in the pixel or SCT detector. Further, the total number of holes[3] is not allowed to exceed two per track, and not more than one in the pixel detector. Additionally, the tracks must have been used in the fit of the primary vertex, or they were not used in any vertex fit. In the latter case, their absolute value of z_0^{BL} times $\sin\theta$ must be less than 3 mm. z_0^{BL} is the difference between the longitudinal position of the primary vertex and the longitudinal position of the track along the beam line at the point where d_0^{BL}, the transverse impact parameter calculated with respect to the measured beam line position, is measured. Tracks are associated to jets using the ghost association method.

5 Jet Mass Definitions and Calibration

5.1 Jet mass definitions

The **calorimeter-based jet mass (m^{calo})** for a large-radius calorimeter jet J with calorimeter-cell cluster constituents i with energy E_i, momentum $\vec{p}_i$ ($|\vec{p}_i| = E_i$) is defined as:

$$m^{\mathrm{calo}} = \sqrt{\left(\sum_{i\in J} E_i\right)^2 - \left(\sum_{i\in J} \vec{p}_i\right)^2}. \tag{1}$$

The angular spread in the decay products of a boosted massive particle scales as $1/p_{\mathrm{T}}$. For a sufficiently high Lorentz boost, this spread is comparable with the calorimeter granularity. Tracking information can be used to maintain performance beyond this granularity limit. One simple method for combining tracking and calorimeter information is the **track-assisted jet mass (m^{TA})**:

$$m^{\mathrm{TA}} = \frac{p_{\mathrm{T}}^{\mathrm{calo}}}{p_{\mathrm{T}}^{\mathrm{track}}} \times m^{\mathrm{track}}, \tag{2}$$

where $p_{\mathrm{T}}^{\mathrm{calo}}$ is the transverse momentum of a large-radius calorimeter jet, $p_{\mathrm{T}}^{\mathrm{track}}$ is the transverse momentum of the four-vector sum of tracks associated to the large-radius calorimeter jet, and m^{track} is the invariant mass of this four-vector sum (the track mass is set to m_π). The ratio of $p_{\mathrm{T}}^{\mathrm{calo}}$ to $p_{\mathrm{T}}^{\mathrm{track}}$ corrects for charged-to-neutral fluctuations, improving the resolution with respect to a track-only jet mass definition

[3] Holes are defined as intersections of the reconstructed track trajectory with a sensitive detector element that do not result in a hit.

(m^{track}). This is illustrated by Fig. 1, which shows that the peak position and width of the track-assisted jet mass (dashed black line) are comparable to the calorimeter-based jet mass (dashed red line) and significantly better than the track-only jet mass (dashed blue line) for 1.6 TeV $< p_{\text{T}} <$ 1.8 TeV.

A procedure for correcting the jet mass as in Eq. 2 was first proposed using hadronic calorimetry to correct electromagnetic-only measurements [35, 36]. The extension to charged particle tracks was introduced in the context of top-quark jet tagging [37] using the HEPTopTagger algorithm [38, 39]. Since that time, there have been phenomenological studies using track-assisted jet mass[4] for ultra boosted ($p_{\text{T}} \gtrsim O(10)$ TeV) boson and top quark jets [40, 41]. This note is the first experimental study of the track-assisted jet mass, including a discussion of its calibration and the associated systematic uncertainties.

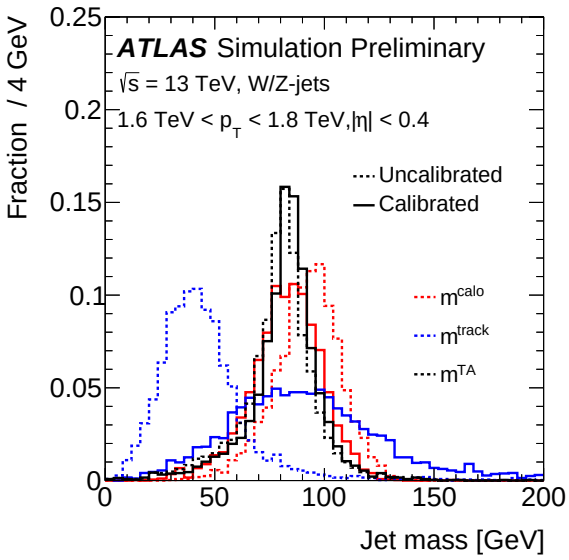


Figure 1: Uncalibrated (dashed line) and calibrated (solid line) reconstructed jet mass distribution for calorimeter-based jet mass, m^{calo} (red), track-assisted jet mass m^{TA} (black) and the invariant mass of four-vector sum of tracks associated to the large-radius calorimeter jet m^{track} (blue) for W/Z-jets.

5.2 Jet mass scale calibration

The jet mass scale (JMS) calibration procedure aims to correct, on average, the reconstructed jet mass to the particle-level jet mass by applying calibration factors derived from a sample of simulated QCD multijet events. The procedure is analogous to the jet energy scale (JES) calibration [30–32].

The calibration is derived using isolated large-radius calorimeter jets that are matched to isolated particle-level truth jets. A particle-level truth jet is considered matched to a large-radius calorimeter jet if it is within $\Delta R < 0.6$ of the calorimeter jet. The isolation criteria is that there should be no other large-radius calorimeter (particle-level truth) jet with $p_{\text{T}} > 100$ GeV within $\Delta R = 1.5$ (2.5).

[4] The phenomenological studies have not given a name to the quantity to Eq. 2, so it is defined here as the track-assisted jet mass.

For each matched pair of large-radius calorimeter and truth jets, the jet mass response for a given jet mass definition is defined as:

$$\mathcal{R}_{\mathrm{m}} = m^{\mathrm{reco}}/m^{\mathrm{truth}}; \qquad m^{\mathrm{reco}} \in c^{\mathrm{JES}} \cdot m^{\mathrm{calo}}, c^{\mathrm{JES}} \cdot m^{\mathrm{TA}}, \tag{3}$$

where c^{JES} is the jet energy scale calibration factor which depends on E^{reco} and η_{det}.[5] The jet mass response is calculated using the reconstructed jet mass with the jet energy scale calibration applied. For each ($p_{\mathrm{T}}^{\mathrm{truth}}$, $|\eta_{\mathrm{det}}|$, m^{truth})-bin, the average jet mass response $\langle\mathcal{R}_{\mathrm{m}}\rangle$ is extracted and defined as the mean of a Gaussian fit to the jet mass response distribution. In order to be able to apply the calibration in data, the procedure must not depend on particle-level quantities. To this end, a *numerical inversion* technique is applied to calibrate a reconstructed jet quantity x^{reco}:

$$x^{\mathrm{reco}}_{\mathrm{calibrated}} = \frac{x^{\mathrm{reco}}}{\langle\mathcal{R}_x\rangle(f^{-1}(x^{\mathrm{reco}}))} \equiv c_x(E^{\mathrm{reco}} \text{ or } p_{\mathrm{T}}^{\mathrm{reco}}, \eta_{\mathrm{det}}, x^{\mathrm{reco}}) \cdot x^{\mathrm{reco}}, \tag{4}$$

where $f(x) = \langle\mathcal{R}_x\rangle(x) \cdot x$, which is the average reconstructed jet quantity given the particle-level quantity and c_{x} is the calibration factor for x^{reco} which is defined as the inverse of the average response $(1/\langle\mathcal{R}_x\rangle)$. In a given $|\eta_{\mathrm{det}}|$-bin, the jet mass scale calibration factors c_{mass} are parameterized as a function of $p_{\mathrm{T}}^{\mathrm{reco}}$ and m^{reco} and the function is constructed by using a two dimensional Gaussian kernel (see Ref. [42] for more detail). For a reconstructed large-radius calorimeter jet with energy E^{reco}, reconstructed transverse momentum $p_{\mathrm{T}}^{\mathrm{reco}}$, detector pseudorapidity η_{det}, calorimeter-based jet mass m^{calo} and track-assisted jet mass m^{TA}, the calibration is applied first for the jet energy and then for the jet mass:

$$m^{\mathrm{calo}}_{\mathrm{calibrated}} = c^{\mathrm{JMS}}_{\mathrm{calo}}(c^{\mathrm{JES}} \cdot p_{\mathrm{T}}^{\mathrm{reco}}, |\eta_{\mathrm{det}}|, c^{\mathrm{JES}} \cdot m^{\mathrm{calo}}) \cdot c^{\mathrm{JES}} \cdot m^{\mathrm{calo}}, \tag{5}$$

$$m^{\mathrm{TA}}_{\mathrm{calibrated}} = c^{\mathrm{JMS}}_{\mathrm{TA}}(c^{\mathrm{JES}} \cdot p_{\mathrm{T}}^{\mathrm{reco}}, |\eta_{\mathrm{det}}|, c^{\mathrm{JES}} \cdot m^{\mathrm{TA}}) \cdot c^{\mathrm{JES}} \cdot m^{\mathrm{TA}}, \tag{6}$$

where $c^{\mathrm{JMS}}_{\mathrm{calo}}$ ($c^{\mathrm{JMS}}_{\mathrm{TA}}$) is the jet mass scale calibration factor for calorimeter-based (track-assisted) jet mass with a dependency on the JES-calibrated $p_{\mathrm{T}}^{\mathrm{reco}}$, $|\eta_{\mathrm{det}}|$ and the JES-calibrated m^{calo} (m^{TA}). The jet mass is said to be calibrated if the average response $\langle\mathcal{R}\rangle = 1$ and the calibration procedure is deemed to fulfil *closure* when the same calibration factor is applied to the same sample from which it is derived.

The jet energy calibration partially mitigates the inhomogeneities in the detector response as a function of η, but the full mass calibration is required to bring the average response close to one. Figure 2 (a, c) shows the calorimeter-based and track-assisted average jet mass response as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ for several m^{truth} bins. Since the $p_{\mathrm{T}}^{\mathrm{calo}}$ term in the definition of the track-assisted jet mass (Eq. 2) is calibrated, the residual jet mass calibration factors are smaller than for the calorimeter-based jet mass.

Figure 3 (a, c) shows the calorimeter-based and track-assisted average jet mass response for several $|\eta_{\mathrm{det}}|$ bins. A dependency on $|\eta_{\mathrm{det}}|$ is observed for the calorimeter-based average jet mass response while there is none for the track-assisted average jet mass response. Figure 4 (a, c) shows the dependency of both jet mass response on m^{truth}.

Following the jet mass scale calibration, a uniform jet mass response is restored to within about 3% across the $p_{\mathrm{T}}^{\mathrm{truth}}$, m^{truth} and $|\eta_{\mathrm{det}}|$ range for both reconstructed jet mass definitions as shown in Fig. 2 (b,

[5] The pseudorapidity of the jet based on the detector geometry.

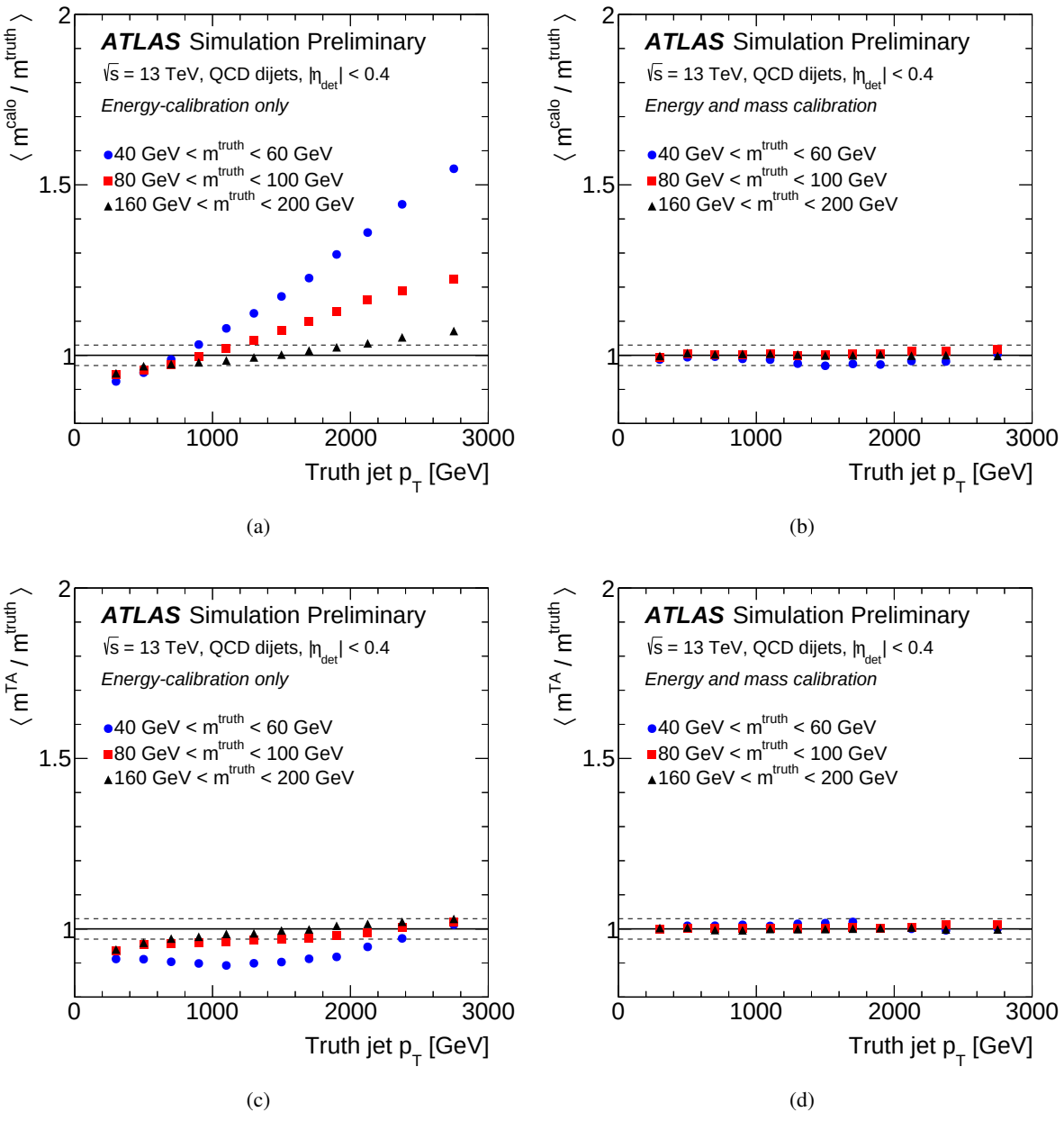


Figure 2: The average calorimeter-based jet mass response (a,b) and the average track-assisted jet mass response (c,d) as functions of $p_{\mathrm{T}}^{\mathrm{truth}}$ for central jets in bins of m^{truth} before (a,c) and after (b,d) the mass calibration is applied. The dashed lines are at 1 ± 0.03.

d), Fig. 3 (b, d) and Fig. 4 (b, d). For high $p_{\mathrm{T}}^{\mathrm{truth}}$ and low m^{truth} jets, there is non-closure for the shape of the calibrated calorimeter-based jet mass response distribution as shown in Fig. 5(c). The calorimeter-based average jet mass response distribution does not retain the gaussian shape after the jet mass calibration is applied. At low m^{truth}, the calorimeter-based average jet mass response decreases rapidly as a function of m^{truth}, as shown in Fig. 4(a); therefore, in a fixed m^{truth} bin, lower mass response jets receive a larger correction than higher mass response jets. The double-peak structure is additionally due to the large

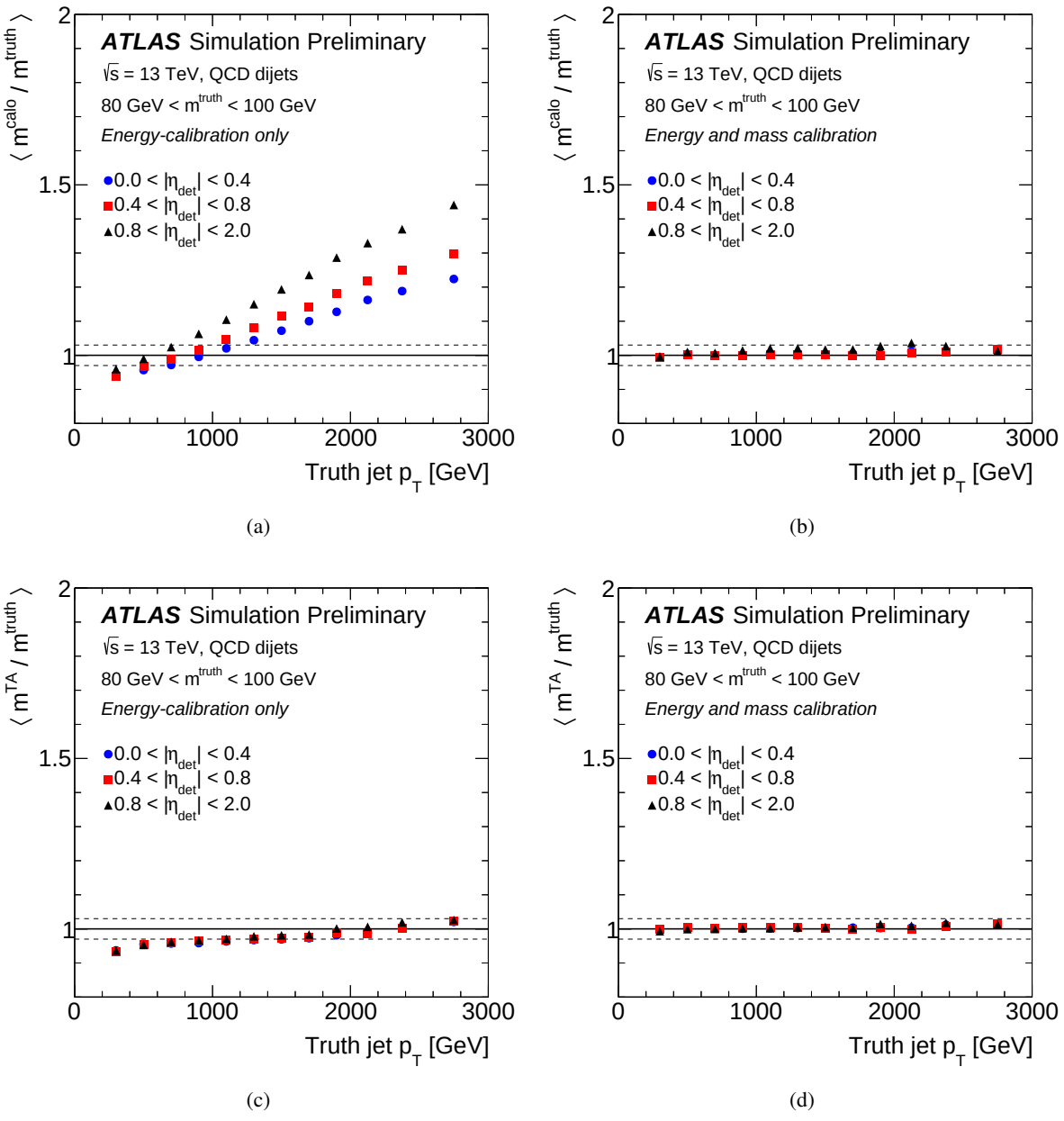
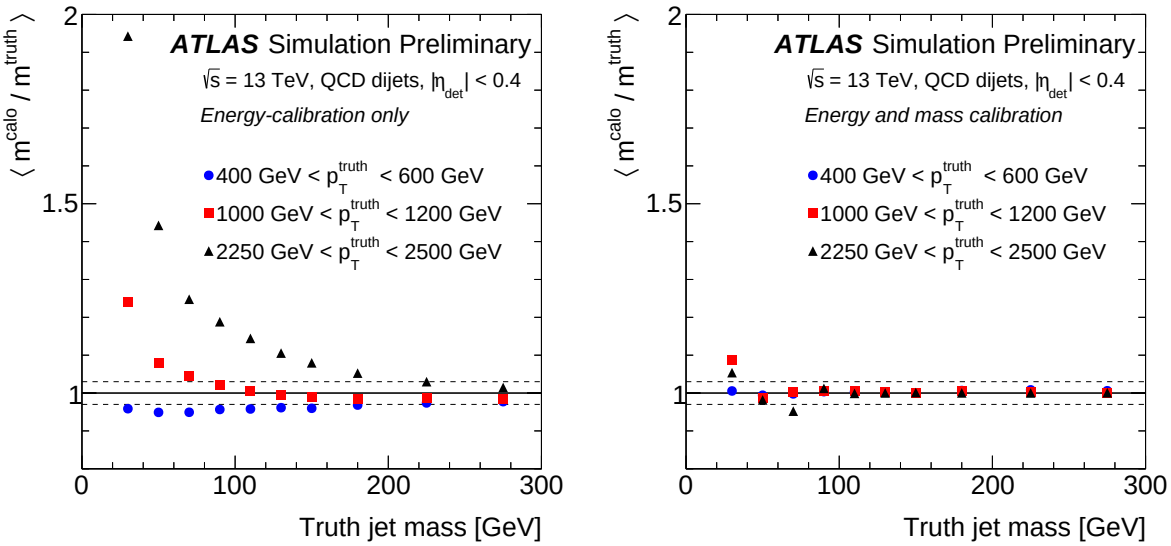


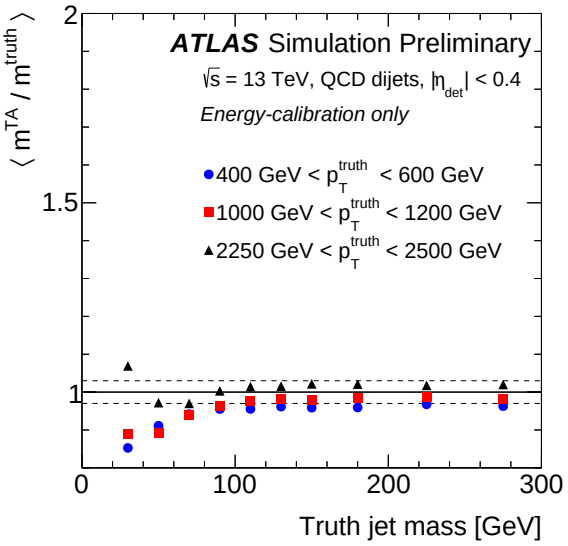
Figure 3: The average calorimeter-based jet mass response (a,b) and the average track-assisted jet mass response (c,d) as functions of $p_{\mathrm{T}}^{\mathrm{truth}}$ for jets with 80 GeV $< m^{\mathrm{truth}} <$ 100 GeV in bins of $|\eta_{\mathrm{det}}|$ before (a,c) and after (b,d) the mass calibration is applied. The dashed lines are at 1 ± 0.03.

resolution of the jet mass so that in a fixed m^{truth} bin, there are two large populations of jets: those with a low reconstructed mass that get a large correction (less than one) and those with a large reconstructed mass that get a correction that is nearly unity (illustrated by Fig. 4(a)). For low $p_{\mathrm{T}}^{\mathrm{truth}}$ jets (Fig. 5(a)) and high m^{truth} jets (Fig. 5(b) and (d)), closure is observed for the shape of both jet mass response distributions as the correction applied on the jet masses is small.

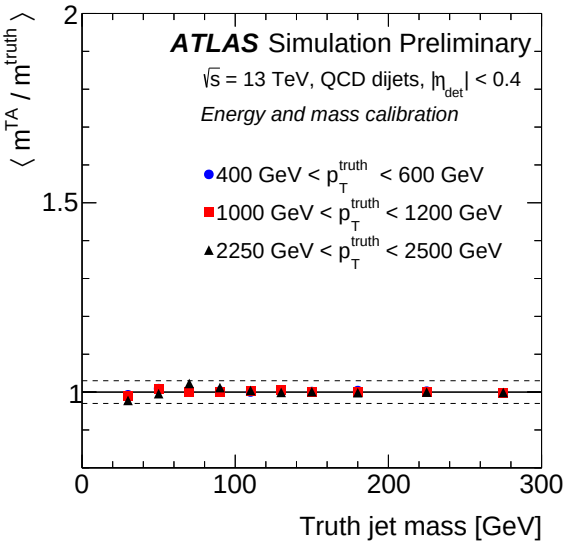


(a)

(b)



(c)



(d)

Figure 4: The average calorimeter-based jet mass response (a,b) and the average track-assisted jet mass response (c,d) as functions of m^{truth} for central jets in bins of $p_{\text{T}}^{\text{truth}}$ before (a,c) and after (b,d) the mass calibration is applied. The dashed lines are at 1 ± 0.03.

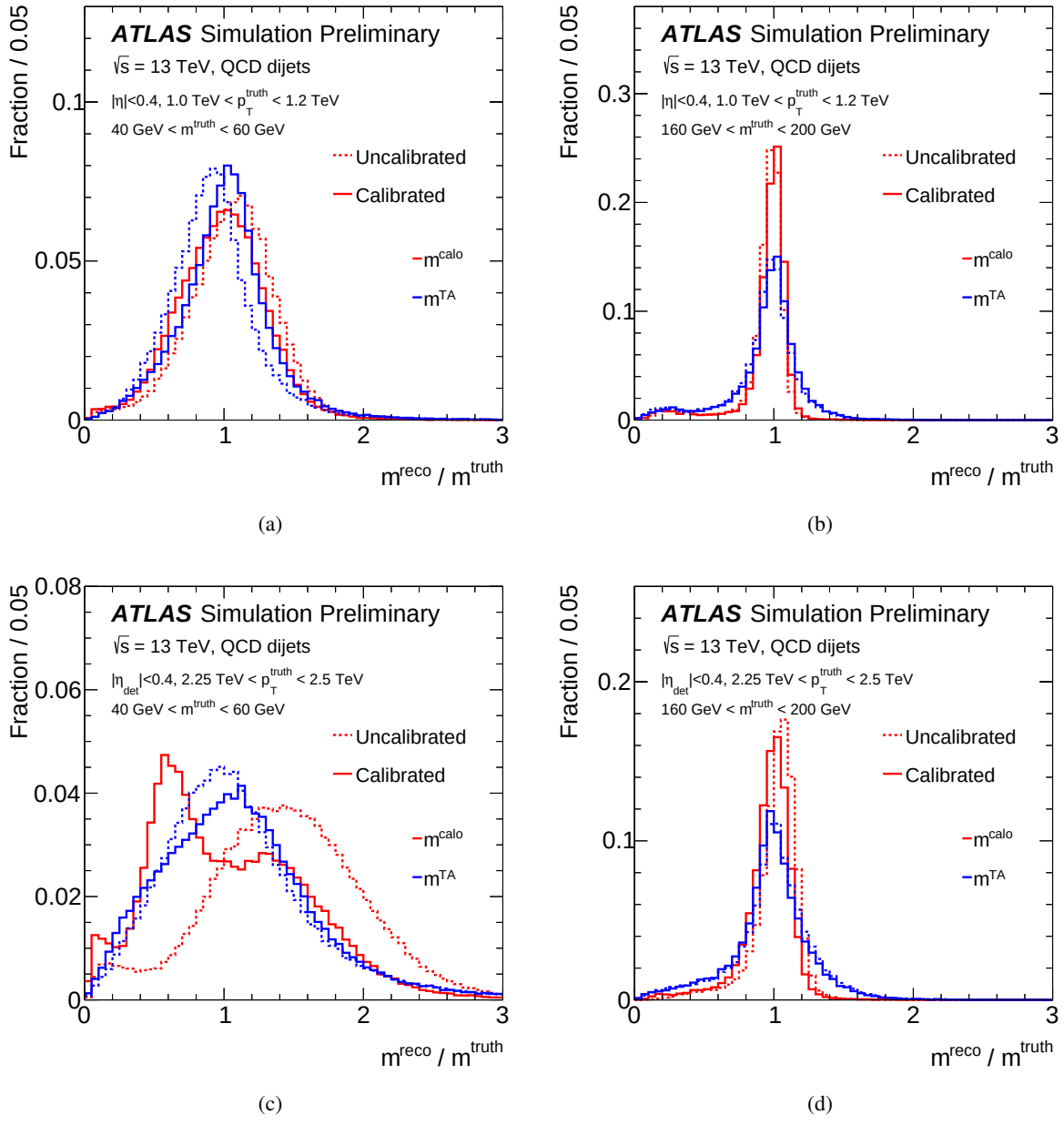
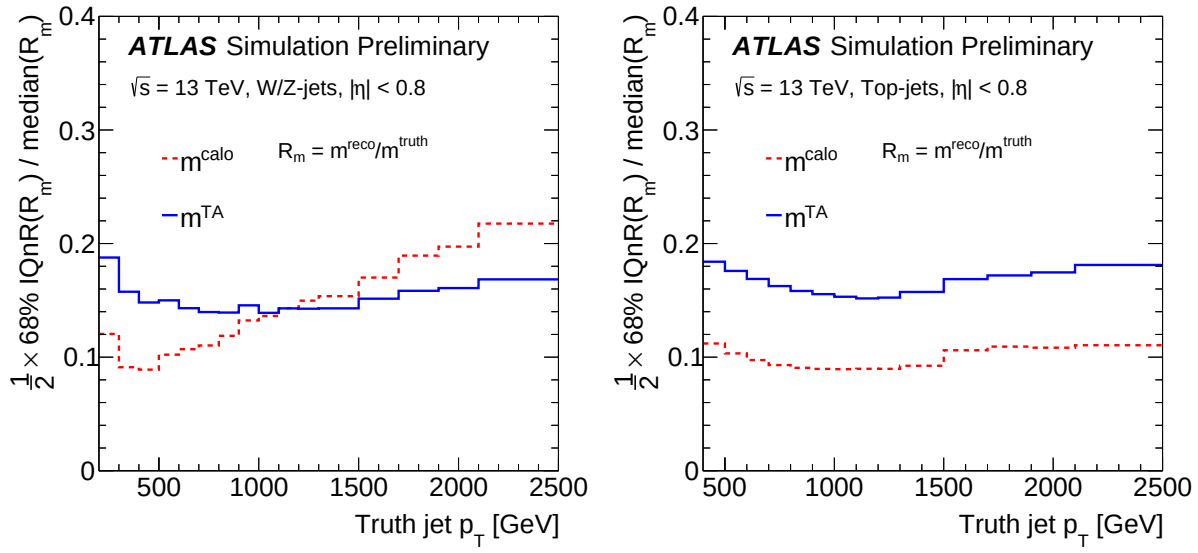


Figure 5: Uncalibrated (dashed line) and calibrated (solid line) jet mass response distributions for calorimeter-based jet mass (red) and track-assisted jet mass (blue) for central jets with 1.0 TeV $< p_{\mathrm{T}}^{\mathrm{truth}} <$ 1.2 TeV (a,b) and 2.25 TeV $< p_{\mathrm{T}}^{\mathrm{truth}} <$ 2.5 TeV (c,d).

6 Jet Mass Performance in Simulation

Figure 6 shows the jet mass resolution as a function of truth jet p_{T} for W and Z bosons jets as well as top quark jets. There are many ways to quantify the resolution of the response distribution, but one robust method that is insensitive to outliers is to use half of the 68% interquantile range (IQnR)[6] divided by the median. In the ideal Gaussian case, this quantity coincides with the standard deviation. The left plot in Fig. 6 shows that both the calorimeter-based and track-assisted jet mass resolution degrade at high p_{T}; for the calorimeter this is due to finite granularity and for the tracker this is due to an increasing track resolution and an increased rate of track merging in the high density jet core. For W and Z boson jets, the track-assisted jet mass has a superior resolution to the calorimeter-based jet mass above about $p_{\mathrm{T}} > 1$ TeV. The charged-to-neutral fluctuations dominate the resolution of the track-assisted jet mass, which is worse than that of the calorimeter-based jet mass resolution below 1 TeV. In contrast, the track-assisted jet mass resolution is larger than the calorimeter-based jet mass resolution over the entire range p_{T} range for top-quark jets. For a fixed p_{T}, the separation between the decay products of primary particles with mass m is proportional to m. Therefore, the point at which the calorimeter granularity makes the calorimeter-based jet mass resolution worse than the track-assisted jet mass resolution is at a much higher p_{T}. This point is also higher for top-quark jets due to the larger subjet multiplicity. The track-assisted jet mass performance can be improved at low p_{T} by reducing the impact of charge-to-neutral fluctuations through local (subjet) corrections. This is explored in more detail in Sec. 10.2. The baseline track-assisted jet mass is most useful for W and Z boson jets with $p_{\mathrm{T}} \gtrsim 1$ TeV, where the resolution is smaller than the calorimeter-based jet mass without any further modification.



(a) (b)

Figure 6: The resolution of the jet mass response as a function of truth jet p_{T} for W and Z boson jets (a) and top-quark jets (b) for calorimeter-based jet mass (red dashed line) and track-assisted jet mass (blue solid line). The half of the 68% interquantile range (IQnR) divided by the median of the jet mass response is used as an outlier insensitive measure of the resolution.

[6] This is defined as $q_{84\%} - q_{16\%}$, whereby $q_{16\%}$ and $q_{84\%}$ are the 16$^{\mathrm{th}}$and 84$^{\mathrm{th}}$ percentiles of a given distribution.

7 Jet Mass Systematic Uncertainties

A variety of methods are used to estimate potential sources of systematic differences in the jet mass scale and resolution between the data and simulation. Since the partonic center of mass energy is unknown at a hadron collider, there is no conservation law to use direct balance techniques to constrain the mass resolution using data, as it can be done for the jet p_{T}. The jet mass scale of calorimeter-based jet mass is probed in the data by studying the ratio $r^{\mathrm{m}}_{\mathrm{track}} = m^{\mathrm{calo}}/m^{\mathrm{track}}$ in an inclusive selection of high p_{T} QCD dijet events [32]. The average value of $r^{\mathrm{m}}_{\mathrm{track}}$ is approximately[7] proportional to the jet mass scale and so $1 - \langle r^{\mathrm{m}}_{\mathrm{track}}\rangle_{\mathrm{Data}}/\langle r^{\mathrm{m}}_{\mathrm{track}}\rangle_{\mathrm{MC}}$ is a measure of the scale uncertainty. Track modeling and fragmentation modeling can also introduce changes in $1 - \langle r_{\mathrm{track}}\rangle_{\mathrm{Data}}/\langle r_{\mathrm{track}}\rangle_{\mathrm{MC}}$, so their uncertainties limit the precision of this method to $\sim 5\%$. The r_{track}-method cannot be used to measure the jet mass resolution because the resolution of $r^{\mathrm{m}}_{\mathrm{track}}$ is dominated by charged-to-neutral fluctuations. Instead, an in-situ method based on the hadronic-decay of W bosons and top quarks is used for this purpose, described in Sec. 9 in more detail.

One key advantage of the track-assisted jet mass over the calorimeter-based jet mass is that the systematic uncertainties can be determined through auxiliary studies. In particular, the jet mass scale and jet mass resolution uncertainty on m^{TA} are estimated by propagating the track reconstruction uncertainties and calorimeter-jet p_{T} uncertainties through the definition in Eq. 2. The calorimeter-jet p_{T} uncertainty is estimated using the p_{T} version of the r_{track} method: $r^{p_{\mathrm{T}}}_{\mathrm{track}} = p^{\mathrm{calo}}_{\mathrm{T}}/p^{\mathrm{track}}_{\mathrm{T}}$, though in the future any of the in-situ methods for small-radius jets could be adapted for this purpose.

The dominating track reconstruction inefficiency for isolated particles is due to hadronic interactions with the inner detector material. Inside the core of high p_{T} jets, an additional inefficiency due to the high particle-density exists. For isolated tracks, the reconstruction efficiency uncertainty is estimated by varying the amount of inner detector material within its measured uncertainty [43]. The uncertainty on the loss of tracks in the core of high p_{T} jets is estimated with a data-driven technique based on the measured energy loss of charged particles in the pixel detector [44]. In addition to the track reconstruction efficiency, the other leading source of uncertainty is due to *fake* tracks resulting from badly reconstructed tracks. The modeling of fake tracks is probed in data by studying the pileup dependence of the number of reconstructed tracks. Based on the assumption that this dependency should be linear without any fake track contribution, the fraction of fakes is estimated by any observed non-linearity. A 30% uncertainty on the fake rate is assigned based on data/MC comparisons of this dependence. Furthermore, possible effects due to the uncertainty in the reconstructed momentum of tracks were assessed using an iterative method based on $Z \to \mu\mu$ events [45–47]. In addition to detector-based uncertainties, the modeling of fragmentation leads to an uncertainty in the track-assisted jet mass resolution. A fragmentation uncertainty is estimated by comparing Pythia 8 and Herwig++.

Figure 7 summarizes the various components of both the calorimeter-based and track-assisted jet mass scale uncertainties. The tracking uncertainties for m^{TA} and for m^{calo} (through r_{track}) are fully correlated. However, the impact of the tracking uncertainties are smaller for m^{TA} compared with m^{calo} because a large extent of the uncertainty cancels in the ratio $m^{\mathrm{track}}/p^{\mathrm{track}}_{\mathrm{T}}$. At high p_{T}, the uncertainty is limited by the size of the dataset used to assess the modeling of r_{track}. Between about $p_{\mathrm{T}} = 300$ GeV and $p_{\mathrm{T}} = 1$ TeV, the uncertainty is about 4% for m^{calo} and about 2% for m^{TA}.

[7] One can decompose $r^{\mathrm{m}}_{\mathrm{track}} = \mathcal{R} \times (m^{\mathrm{truth}}/m^{\mathrm{charged\ truth}}) \times (m^{\mathrm{charged\ truth}}/m^{\mathrm{track}})$. If all three terms on the right hand side of the equation are independent, then $\langle r^{\mathrm{m}}_{\mathrm{track}}\rangle \propto \langle\mathcal{R}\rangle$. However, since the calorimeter response depends on the charged-to-neutral ratio, this factorization is only approximate.

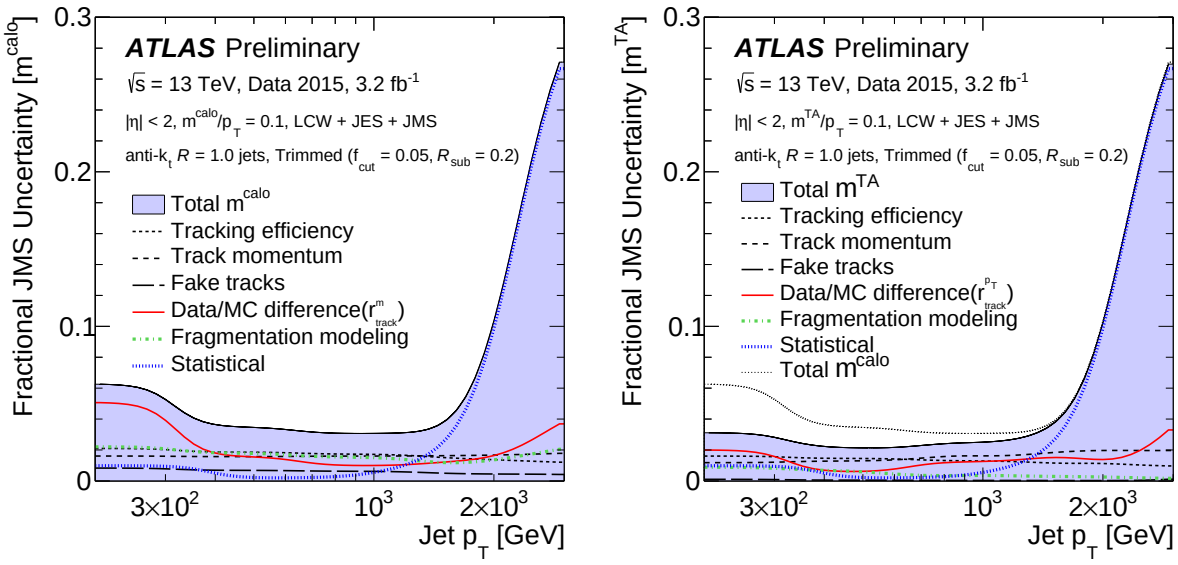


Figure 7: A breakdown of the systematic uncertainties on the jet mass scale for m^{calo} (left) and m^{TA} (right) as a function of jet p_{T} for $m^{\text{reco}}/p_{\text{T}} = 0.1$ and $|\eta| < 2$. In the right figure, the total JMS uncertainty for m^{calo} is included for comparison with the total JMS uncertainty for m^{TA}. The uncertainty is parameterized as a function of m/p_{T} and these two plots show a slice at $m/p_{\text{T}} = 0.1$.

8 Comparisons between Data and Simulation

Figure 8 shows the distribution of the reconstructed jet mass in a sample of events enriched in top quark pair ($t\bar{t}$) events. The selection is based on the Run 1 in-situ JMS/JMR measurement [48] and is summarized here for completeness. Single-muon triggers are used to reject most of the events from QCD multijet background processes. $t\bar{t}$ events are chosen by requiring a muon with $p_{\text{T}} > 25$ GeV and $|\eta| < 2.5$, as well as a missing transverse momentum $E_{\text{T}}^{\text{miss}} > 20$ GeV. The muons are required to satisfy a series of quality criteria, including isolation.[8] The events are rejected if there is an additional muon. In addition, the sum of the $E_{\text{T}}^{\text{miss}}$ and the transverse mass[9] of the W boson, reconstructed from the lepton and $E_{\text{T}}^{\text{miss}}$, is required to be greater than 60 GeV. Events must have at least one b-tagged jet (at the 70% efficiency working point for jets containing b-hadrons) and have at least one large-radius trimmed jet with $p_{\text{T}} > 200$ GeV and $|\eta| < 2$. Furthermore, there must be a small-radius jet with $p_{\text{T}} > 25$ GeV, and $\Delta R < 1.5$ to the selected lepton (targeting the decay chain $t \to bW(\to \mu\nu)$). The resulting event purity is better than 70%. As expected, there are three mass peaks in Fig. 8 corresponding to the top quark mass, W boson mass, and the quark/gluon Sudakov peak. The W boson mass peak is more pronounced than the top quark mass peak due to the relatively low p_{T}. For $p_{\text{T}} \gtrsim 200$ GeV, the track-assisted jet mass resolution is significantly larger than the calorimeter-based jet mass, which is why the W-boson mass peak is broader for the top right plot of Fig. 8. The top quark mass peak is enhanced at higher p_{T} in Fig. 9.

[8] Muon are considered isolated if they are well separated from jets ($\Delta R > 0.4$) and the track/calorimeter energy within a small cone, centered on the lepton direction but excluding the lepton itself, is below a fixed relative value.

[9] The transverse mass, m_{T}, is defined as $m_{\text{T}}^2 = 2p_{\text{T}}^{\mu} E_{\text{T}}^{\text{miss}}(1 - \cos(\Delta\phi))$, where $\Delta\phi$ is the azimuthal angle between the muon and the direction of the missing transverse momentum.

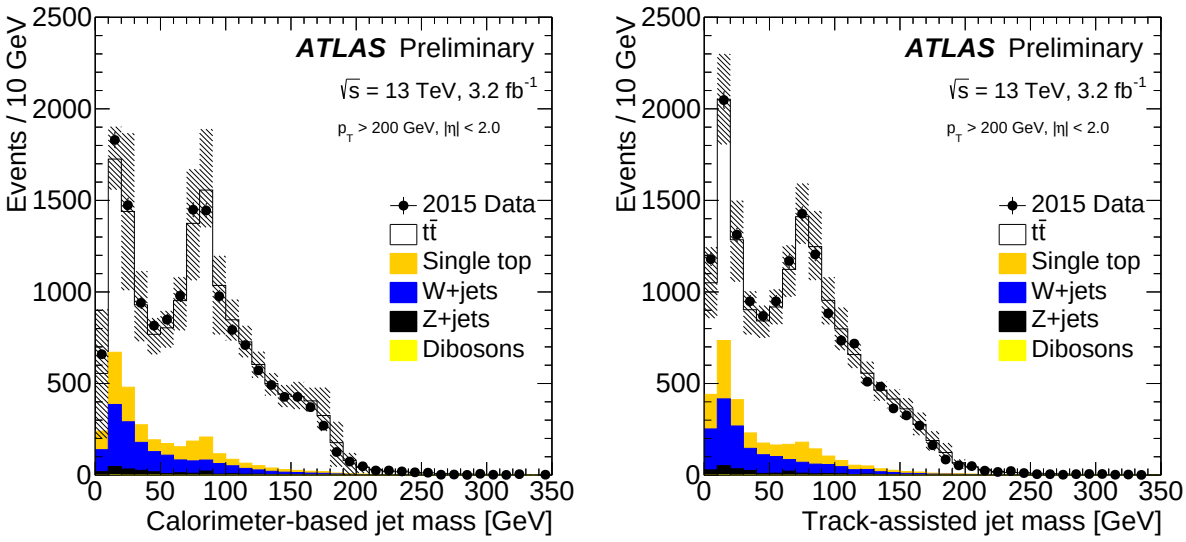


Figure 8: The calorimeter-based jet mass distribution (left) and the track-assisted jet mass distribution (right) for jets with $p_{\mathrm{T}} > 200$ GeV. The MC is normalized to the data event yield. The uncertainty band includes systematic uncertainties related to the modeling of $t\bar{t}$ and the jet energy/mass scale uncertainties. See Sec. 9 for details.

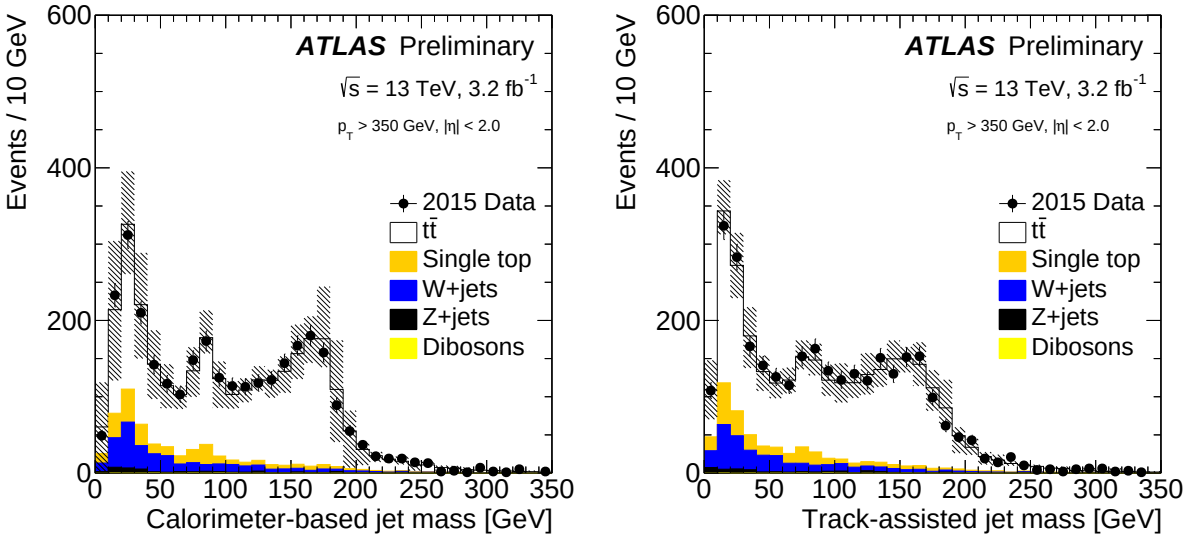


Figure 9: The calorimeter-based jet mass distribution (left) and the track-assisted jet mass distribution (right) for jets with $p_{\mathrm{T}} > 350$ GeV. The MC is normalized to the data event yield. The uncertainty band includes systematic uncertainties related to the modeling of $t\bar{t}$ and the jet energy/mass scale uncertainties. See Sec. 9 for details.

9 In-situ Mass and Energy Scale and Resolution

Top quark pair production provides an abundant source of hadronically decaying top quark and W boson jets that can be used to study the reconstruction of the jet four-vector in-situ.[10] One approach is the forward-folding method [48], in which particle-level spectra are folded by a modified detector response in order to best fit the reconstructed data. The resolution function that describes the transition from particle-level quantities x^{truth} to calibrated detector-level quantities x^{reco} is stretched and shifted so that the average value of x^{reco} in a fixed truth bin $\langle x^{\text{reco}}|x^{\text{truth}}, p_{\text{T}}^{\text{reco}}\rangle$ is scaled by s and the resolution is independently scaled by r:

$$x^{\text{folded}}|x^{\text{truth}}, p_{\text{T}}^{\text{reco}} = s x^{\text{reco}} + \left(x^{\text{reco}} - \langle x^{\text{reco}}|x^{\text{truth}}, p_{\text{T}}^{\text{reco}}\rangle\right)(r - s), \tag{7}$$

where $x^{\text{folded}}|x^{\text{truth}}, p_{\text{T}}^{\text{reco}}$ is the folded quantity for a jet with x^{truth} and $p_{\text{T}}^{\text{reco}}$. By construction, $x^{\text{folded}} = x^{\text{reco}}$ when $s = r = 1$ and $x^{\text{folded}} = x^{\text{truth}}$ if $s = 1$, $r = 0$, and the method closes so that $\langle x^{\text{reco}}|x^{\text{truth}}, p_{\text{T}}^{\text{reco}}\rangle = x^{\text{truth}}$. A two-dimensional χ^2 fit is performed to determine the values $s_{\text{data}}^{\text{MC}}$ and $r_{\text{data}}^{\text{MC}}$ such that the folded distribution best fits the data. By construction, fitting the simulation to itself results in $s_{\text{MC}}^{\text{MC}} = r_{\text{MC}}^{\text{MC}} = 1$. To extract the jet mass scale and resolution, Eq. 7 is applied to the jet mass spectrum in one-lepton $t\bar{t}$ events (see Sec. 8). The W-boson and top-quark resonance peaks are due to the combination of final-state radiation and the detector response. Taking the particle-level jet mass distribution as input, the relative jet mass scale and jet mass resolution can be extracted from forward-folding.

An improvement in the $\sqrt{s} = 13$ TeV measurement with respect to the $\sqrt{s} = 8$ TeV result in Ref. [48] is the addition of jet energy scale and resolution constraints using the same forward-folding method. Top-quarks in $t\bar{t}$ events tend to be produced with similar transverse momenta and so the ratio of the hadronically decaying top quark to leptonically decaying top quark transverse momentum, $p_{\text{T}}^{\text{had top}}/p_{\text{T}}^{\text{lep top}}$, is sensitive to the energy scale and resolution of the top quark jet. The leptonically decaying top quark is reconstructed from the selected lepton, the missing transverse momentum, and the close-by b-jet used in the event selection (see Sec. 8). The balance of the leading large-radius calorimeter jet p_{T} with a partially reconstructed leptonically decaying top quark p_{T} from only the lepton ($p_{\text{T}}^{\text{lep}}$) or the lepton and the close-by jet ($p_{\text{T}}^{\text{lep + b-jet}}$) is also considered due to their reduced systematic uncertainty from the calorimeter energy scale. Figure 10 shows the calorimeter-based jet mass as well as the three reconstructed top-quark transverse momentum ratio distributions. The p_{T} ratio with the fully reconstructed leptonically decaying top quark has the sharpest peak because it contains the most information.

The χ^2 distributions marginalized over s or over r are shown in Fig. 11 and 12, respectively. The minima for the jet mass and energy resolutions are much flatter than for the corresponding scales because the resolutions in Fig. 10 are largely dominated by the physical resolution (from e.g. FSR). As expected from the relative size of the peaks in Fig. 10, the leptonically decaying top quark transverse momenta ratio is the most sensitive to the relative scale and resolution.

Table 2 summarizes the fitted values and systematic uncertainties. The jet mass scale and resolution of calorimeter-based jets are compared with the results obtained on $\sqrt{s} = 8$ TeV data in Fig. 13. The mass scale and resolution of the track-assisted jet mass and reclustered jets are presented in Fig. 14.

[10] The studies presented in this section are not currently used as the default uncertainties applied by analyses - see Sec. 7 for the current jet mass scale uncertainties. These in-situ studies may be used in the future, but additional studies are required to extend their validity beyond the presented kinematic regime.

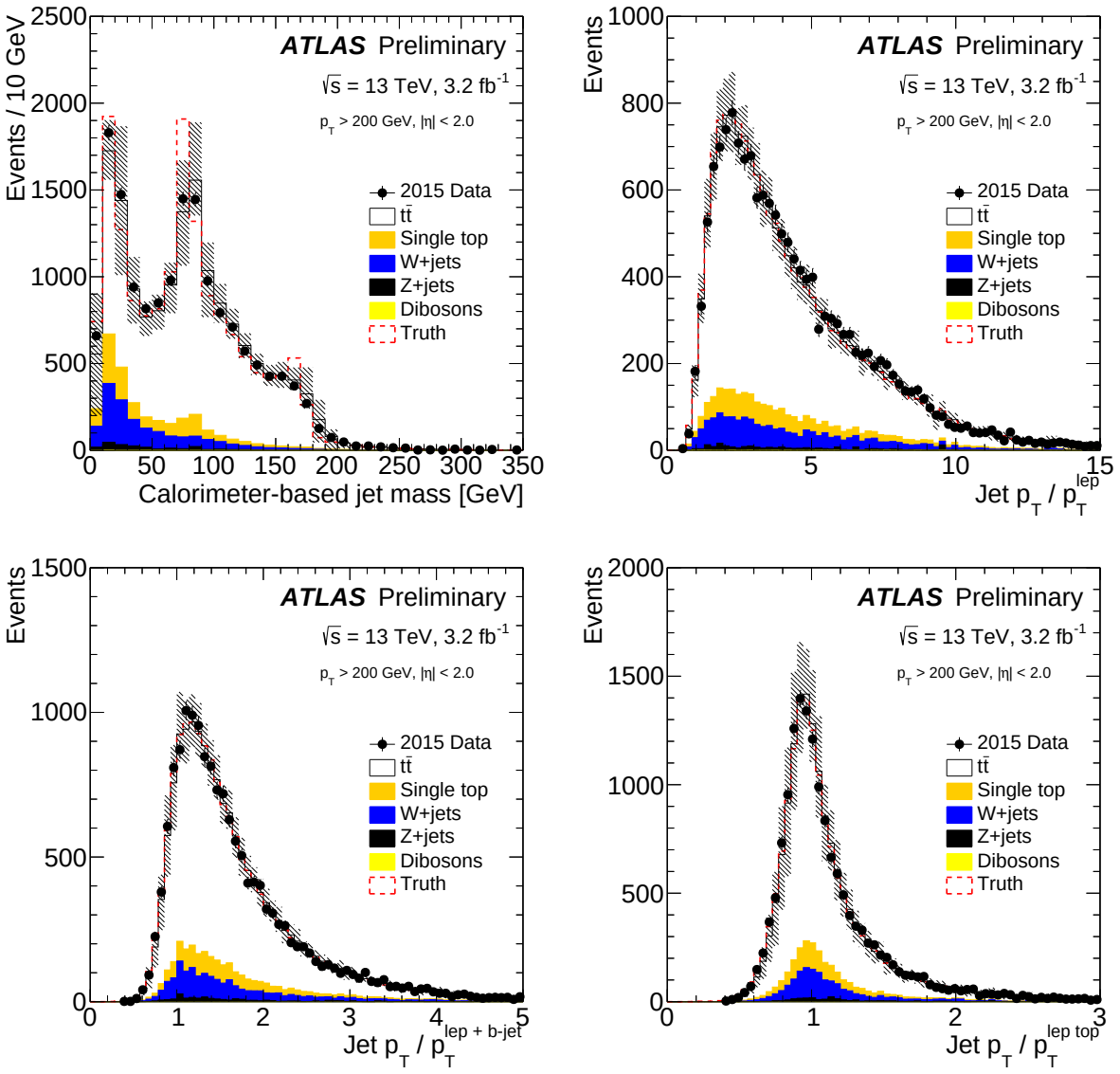


Figure 10: The distribution of the calorimeter-based jet mass (top left) and the jet p_{T} divided by various reference transverse momenta: charged lepton p_{T} (top right), combined with the nearest b-jet momentum (bottom left), further combined with the $E_{\mathrm{T}}^{\mathrm{miss}}$ (bottom right). The truth histograms show the distribution of the particle-level jet mass (top left) or the particle-level jet p_{T} divided by the detector-level reference momentum (top right and bottom). The uncertainty band includes systematic uncertainties related to the modeling of $t\bar{t}$ and the jet energy/mass scale uncertainties. See the text for details.

The jet mass scale and the resolution extracted from the data have a sizeable statistical uncertainty and may be affected by a bias due to imperfections in the modelling of physics process or other aspects of the experimental response.

The most relevant experimental uncertainties are those which change the shape of the jet p_T or jet mass distributions. The jet energy scale and jet energy resolution systematics are the most relevant systematic uncertainties for the jet mass. Pile-up doesn't strongly affect trimmed large-radius jets. Therefore, the

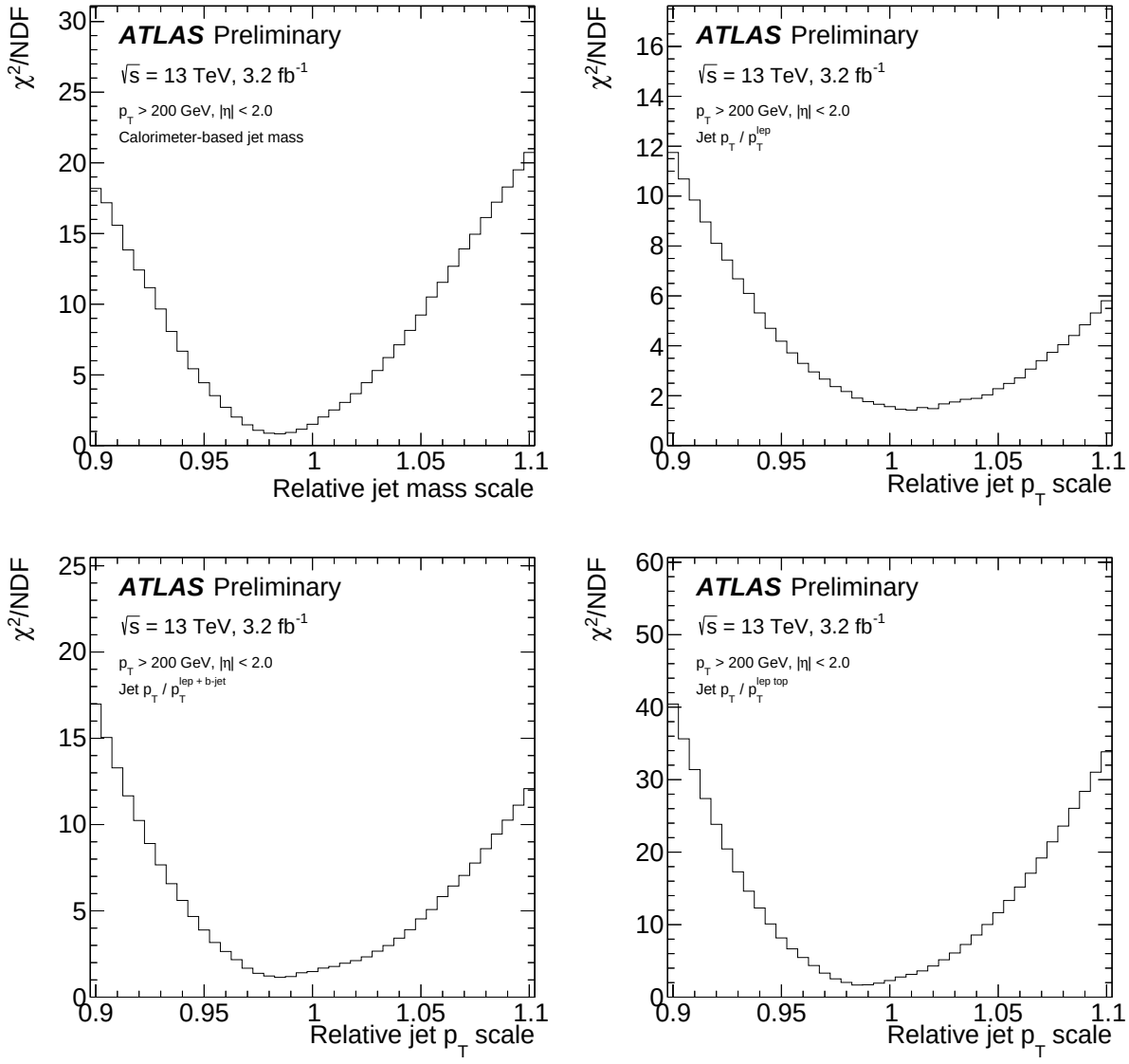


Figure 11: The χ^2 per degrees of freedom, marginalized over the relative jet mass or jet transverse momentum scale for the calorimeter-based jet mass (top left), the jet p_{T} using the ratio $p_{\mathrm{T}}^{\mathrm{had}}/p_{\mathrm{T}}^{\mathrm{lep}}$ (top right), the jet p_{T} using the ratio $p_{\mathrm{T}}^{\mathrm{had}}/p_{\mathrm{T}}^{\mathrm{lep}+b\text{-jet}}$ (bottom left), and the jet p_{T} using the ratio $p_{\mathrm{T}}^{\mathrm{had}}/p_{\mathrm{T}}^{\mathrm{lep\ top}}$ (bottom right).

associated systematic is small enough to be included as a component of JES uncertainty. Three different scenarios are used to study the correlation between small-R jets and large-R jet uncertainties and no large differences are seen. The b-tagging, $E_{\mathrm{T}}^{\mathrm{miss}}$ and lepton systematic uncertainties are small, since they affect mostly the acceptance.

Uncertainties in the modelling of the $t\bar{t}$ production process, top quark decay and fragmentation of the hadronic final state are the largest source of systematic uncertainty. The shape of the particle-level distributions and the distribution of the energy inside a jet depend on the modelling of fragmentation, as well as initial and final state radiation. To take into account these effects, the fit is repeated with several

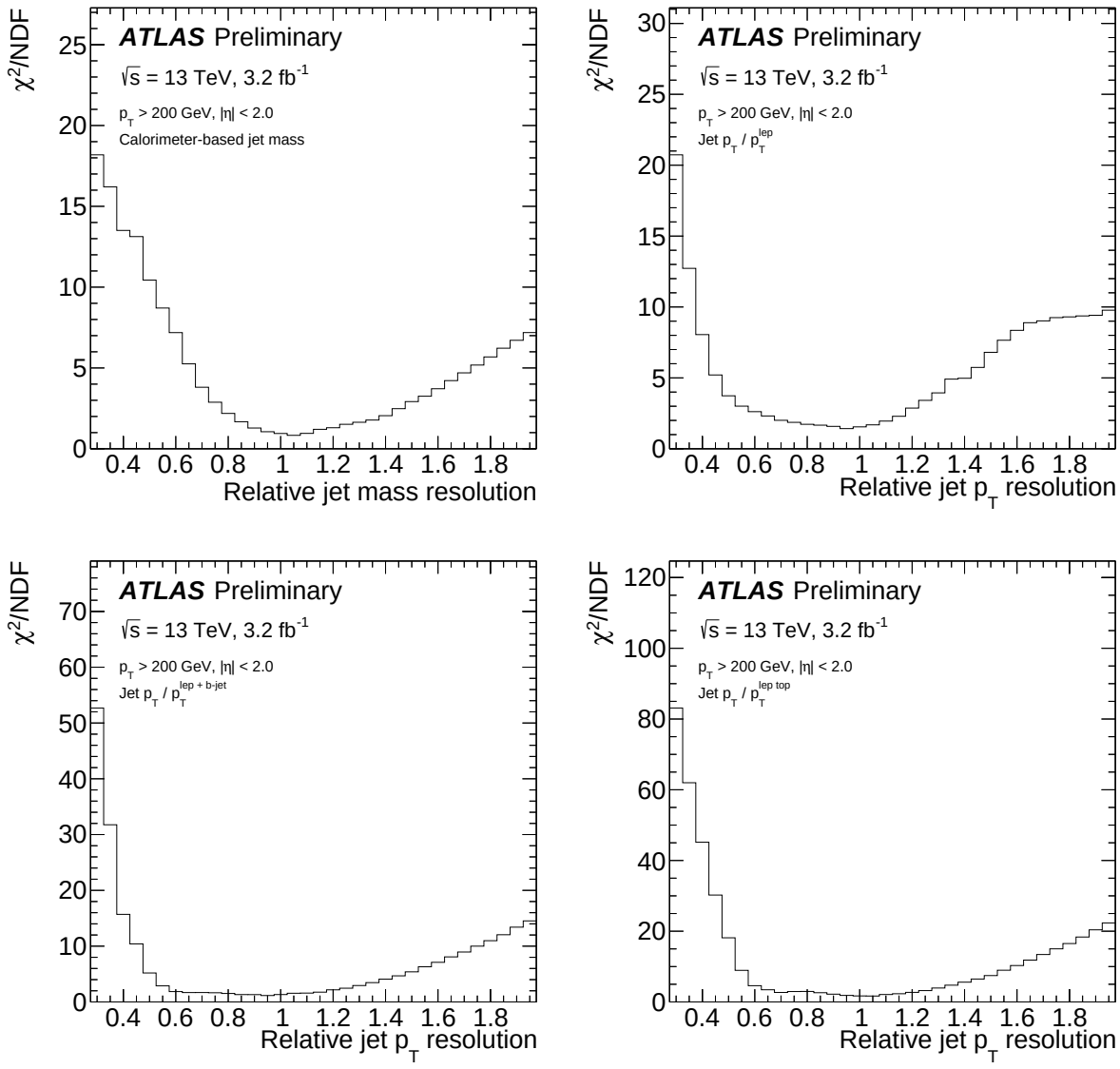


Figure 12: The χ^2 per degrees of freedom, marginalized over the relative jet mass or jet transverse momentum resolution for the calorimeter-based jet mass (top left), the jet p_{T} using the ratio $p_{\mathrm{T}}^{\mathrm{had}}/p_{\mathrm{T}}^{\mathrm{lep}}$ (top right), the jet p_{T} using the ratio $p_{\mathrm{T}}^{\mathrm{had}}/p_{\mathrm{T}}^{\mathrm{lep}+b\text{-jet}}$ (bottom left), and the jet p_{T} using the ratio $p_{\mathrm{T}}^{\mathrm{had}}/p_{\mathrm{T}}^{\mathrm{lep\ top}}$ (bottom right).

Monte Carlo generators. The uncertainty due to the NLO subtraction scheme is estimated by comparing a Powheg+Herwig sample with a MC@NLO+Herwig sample. The effect of using different fragmentation models is estimated by comparing Powheg+Pythia and Powheg+Herwig samples. Uncertainties due to initial and final state radiation are estimated by comparing $t\bar{t}$ Monte Carlo generators with different factorisation/renormalisation scales, as well as the Perugia radLo/radHi tunes.

Quantity		Value	Stat. Uncert	Modeling	Jets	Total Syst.
m^{calo}	$s^{\text{MC}}_{\text{data}}$	0.984	0.6 %	1.7 %	1.6 %	**2.3 %**
m^{calo}	$r^{\text{MC}}_{\text{data}}$	1.047	6.6 %	18.1 %	7.0 %	**19.4 %**
m^{TA}	$s^{\text{MC}}_{\text{data}}$	0.981	1.1 %	2.4 %	4.8 %	**5.3 %**
m^{TA}	$r^{\text{MC}}_{\text{data}}$	1.036	6.1 %	14.6 %	5.0 %	**15.5 %**
$p_{\text{T,jet}}/p_{\text{T}}^{\text{lep}}$	$s^{\text{MC}}_{\text{data}}$	1.011	0.7 %	1.3 %	0.4 %	**1.3 %**
$p_{\text{T,jet}}/p_{\text{T}}^{\text{lep}}$	$r^{\text{MC}}_{\text{data}}$	0.945	4.1 %	6.8 %	2.7 %	**7.3 %**
$p_{\text{T,jet}}/p_{\text{T}}^{\text{lep + b-jet}}$	$s^{\text{MC}}_{\text{data}}$	0.985	0.4 %	0.7 %	1.2 %	**1.4 %**
$p_{\text{T,jet}}/p_{\text{T}}^{\text{lep + b-jet}}$	$r^{\text{MC}}_{\text{data}}$	0.903	6.1 %	5.5 %	4.7 %	**7.2 %**
$p_{\text{T,jet}}/p_{\text{T}}^{\text{lep top}}$	$s^{\text{MC}}_{\text{data}}$	0.987	0.2 %	0.3 %	2.1 %	**2.1 %**
$p_{\text{T,jet}}/p_{\text{T}}^{\text{lep top}}$	$r^{\text{MC}}_{\text{data}}$	1.024	3.1 %	6.2 %	6.0 %	**8.6 %**

Table 2: Summary of the systematic uncertainties for the relative jet mass or energy scales ($s^{\text{MC}}_{\text{data}}$) and resolutions ($r^{\text{MC}}_{\text{data}}$). The first column states which observable is used to extract the relative jet mass (first four rows) or jet energy (rows 5-10) scale and resolutions.

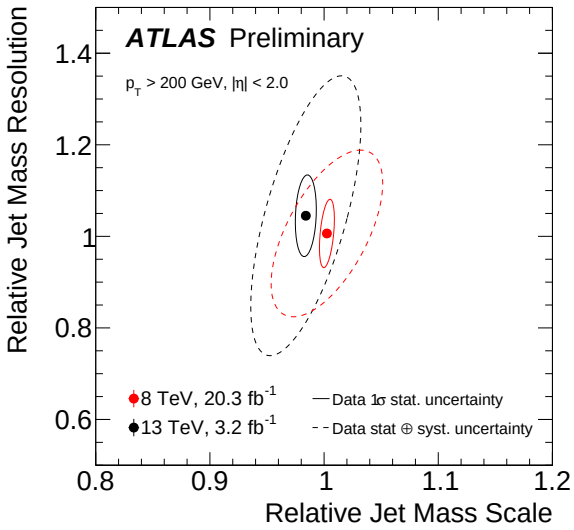


Figure 13: The relative jet mass scale and relative jet mass resolution extracted from the $\sqrt{s}$ = 8 TeV dataset for large-radius trimmed calorimeter jets with a subjet size of R_{sub} = 0.3 [48] and the values extracted from the $\sqrt{s}$ = 13 TeV dataset with R_{sub} = 0.2.

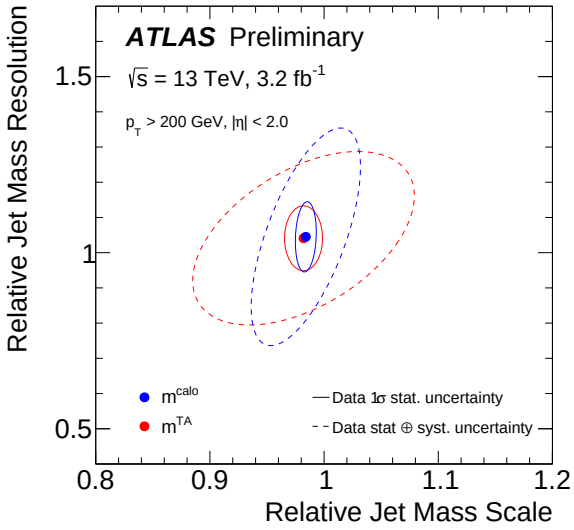


Figure 14: The relative jet mass scale and relative jet mass resolution for calorimeter-based jet mass (blue) and track-assisted jet mass (red)

10 Improving the Track-assisted Jet Mass

As discussed in Section 6, the performance of track-assisted jet mass depends on the p_{T} regime and the topology of the signal jet. In this section, several explored techniques which have the potential to improve the performance of the track-assisted jet mass are discussed.

10.1 Combination with the calorimeter-based jet mass

As the calorimeter-based jet mass is not used explicitly in the construction of the track-assisted jet mass, it may be possible to reduce the response resolution by combining information from both mass definitions. Fluctuations in the calorimeter energy response impact both the calorimeter-based jet mass and p_{T} response. However, the p_{T} response is not as sensitive to the local distribution of fluctuations and therefore the jet p_{T} response and the calorimeter-based jet mass response are nearly independent. The correlation is slightly higher at lower p_{T}, where the the decay products are more spread out and thus the p_{T} and mass are more related. Figure 15 shows the track-assisted jet mass response vs the calorimeter-based jet mass response for jets with $p_{\mathrm{T}} > 1$ TeV where the correlation coefficient is $\sim 10\%$. Due to the approximate independence and Gaussian nature of the p_{T} and mass responses, the optimal combination of the two variables is linear: [11] $m^{\mathrm{comb}} = a \times m^{\mathrm{calo}} + b \times m^{\mathrm{TA}}$. For calibrated m^{calo} and m^{TA}, the combined jet mass is also calibrated if $a + b = 1$. Using this constraint and minimizing the resolution of the m^{comb} response results in the nearly optimal weights:

$$a = \frac{\sigma_{\mathrm{calo}}^{-2}}{\sigma_{\mathrm{calo}}^{-2} + \sigma_{\mathrm{TA}}^{-2}} \qquad b = \frac{\sigma_{\mathrm{TA}}^{-2}}{\sigma_{\mathrm{calo}}^{-2} + \sigma_{\mathrm{TA}}^{-2}}, \tag{8}$$

where σ_{calo} and σ_{TA} are the calorimeter-based jet mass resolution function and the track-assisted mass resolution function respectively. Figure 16 shows the resolution of the calorimeter-based jet mass and the track-assisted jet mass (as in Fig. 6), but additionally shows the resolution of the combined jet mass for W-jets. The combined jet mass smoothly interpolate between $m^{\mathrm{comb}} \sim m^{\mathrm{calo}}$ at low p_{T} and $m^{\mathrm{comb}} \sim m^{\mathrm{TA}}$ at high p_{T}. As expected, the combined jet mass resolution is never larger than either of the input jet masses. For top-jets, since the calorimeter-based jet mass performs best in all p_{T} range, the combined jet mass will be mostly weighted by the calorimeter-based jet mass contribution and the resolution of the combined jet mass is as good as the calorimeter-based jet mass.

Systematic uncertainties on the combined jet mass are estimated by propagating uncertainties on m^{calo} and m^{TA} through the linear combination defining m^{comb}. Just as the tracking r_{track} uncertainties for the calorimeter jet p_{T} are treated as fully correlated with the tracking uncertainties on m^{track} and $p_{\mathrm{T}}^{\mathrm{track}}$, the tracking r_{track} uncertainties for the calorimeter-based jet mass are also fully correlated with all of the tracking uncertainties.

[11] If the joint distribution of the responses is Gaussian, then one can write their probability distribution function as $f(x,y) = h(x,y) \times \exp[A(\mu) + T(x,y)\mu]$, where x is the calorimeter-based jet mass response, y is the track-assisted jet mass response, μ is the common average response, and h, A, T are real-valued functions. This form shows that the distribution is from the *exponential family* and therefore T is a *sufficient statistic*. Since the natural parameter space is one-dimensional, T is also *complete*. Therefore, the unique minimal variance unbiased estimator of μ is the unique unbiased function of $T(x,y) = x/\sigma_x^2 + y/\sigma_y^2$. See e.g. Ref. [49] for details.

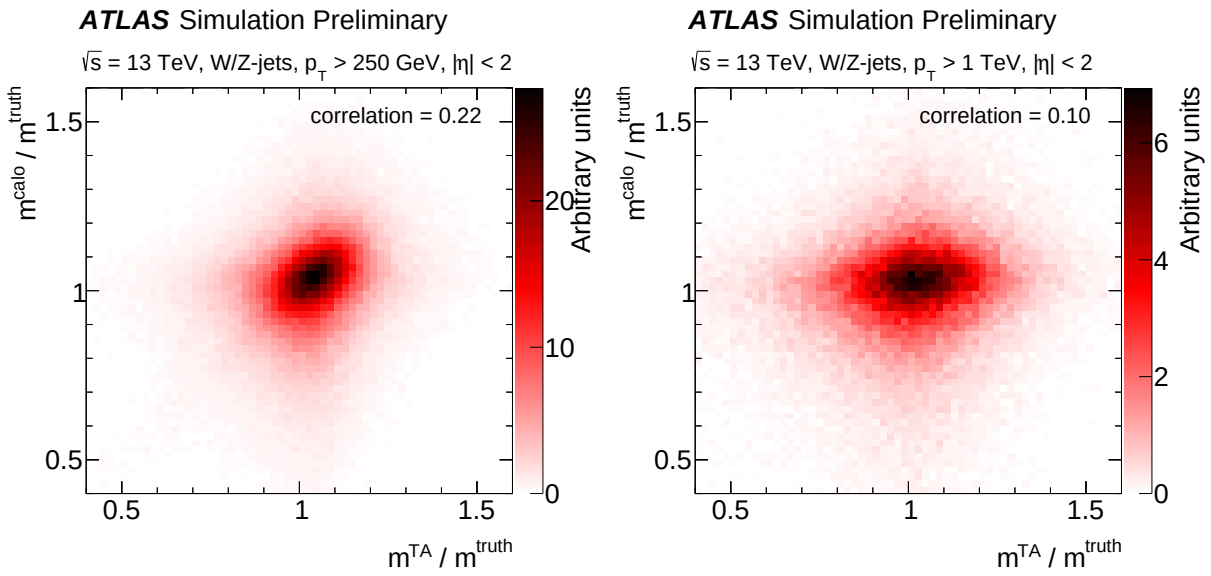


Figure 15: The calorimeter-based jet mass response vs the track-assisted jet mass response for W/Z-jets produced from with $p_{\mathrm{T}} > 250$ GeV (left) and $p_{\mathrm{T}} > 1$ TeV (right).

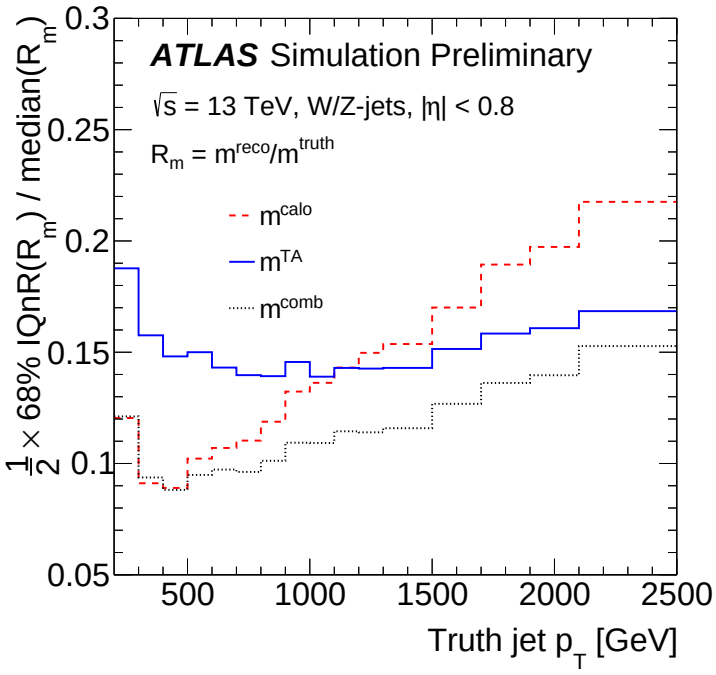


Figure 16: The jet mass resolution vs jet p_{T} for calorimeter-based jet mass (blue solid line), track-assisted jet mass (red dashed line), and the combined jet mass (black dotted line) for W/Z-jets. Note that since the response distributions are non-Gaussian, the linear combination of the two mass definitions is slightly non-optimal in the lowest bins; the combined resolution (in terms of the IQnR) is not exactly $\sigma^2_{\mathrm{combined}} = a\sigma^2_{\mathrm{calo}} + b\sigma^2_{\mathrm{TA}}$.

10.2 Track-assisted subjet mass

Figure 1 illustrated that correcting the track mass by $p_{\mathrm{T}}^{\mathrm{calo}}/p_{\mathrm{T}}^{\mathrm{track}}$ is essential for improving the resolution of the jet mass. The charged-to-neutral fluctuations are large and the physical resolution of the track-mass with respect to the full particle-level jet mass is comparable to the detector-level calorimeter-based jet mass resolution. One way to further reduce the resolution of the track-assisted jet mass is to use *local* charged-to-neutral corrections by applying the track-assisting procedure to subjets. The charge-to-neutral ratio is not constant across the jet and therefore the fluctuations can be suppressed by correcting different regions separately. A set of tracks is matched to each calorimeter subjet[12] using the ghost-association method and the track-assisted four-vector of the subjet as specified by:

$$(p_{\mathrm{T}},\eta,\phi,m)^{\mathrm{subjet}} = \left(p_{\mathrm{T}}^{\mathrm{calo}},\eta^{\mathrm{tracks}},\phi^{\mathrm{tracks}},m^{\mathrm{tracks}} \times p_{\mathrm{T}}^{\mathrm{calo}}/p_{\mathrm{T}}^{\mathrm{tracks}}\right), \tag{9}$$

where the tracker information is used to obtain geometric information and the calorimeter is used for energy information. The m^{TAS} is then given by the invariant mass of the sum of all track-assisted subjet four-vectors. Figure 17 shows the resolution of the m^{TAS} compared with m^{TA} and the calorimeter-based jet mass resolution. For boosted W boson jets at relatively low p_{T}, the m^{TAS} has significantly better resolution than m^{TA} due to the superior angular resolution from the tracker for the resolved subjets. The improvement is less significant at high $p_{\mathrm{T}} \gtrsim 2m_V/R_{\mathrm{sub}} \sim 1$ TeV, where many jets have only one subjet. For top-quark jets, the low p_{T} regime end is much higher as $m_{\mathrm{top}} \sim 2m_V$. Therefore, m^{TAS} is a significant improvement upon m^{TA} over the entire range in Fig. 17 and the resulting resolution is slightly worse to the calorimeter-based jet mass resolution.

The subjet track-assisted jet mass is a promising extension of the m^{TA} with a resolution that is as low or lower than the calorimeter-based jet mass over a wide range of phase space. One key challenge is the calibration of the subjet energy scale. Such a calibration would improve the resolution of the overall jet mass, but requires a careful analysis of the impact of close-by jets on the energy scale. Furthermore, in order to benefit from the factorization of mass and p_{T} systematic uncertainties, it is necessary to measure the subjet energy and mass scale and resolution. In-situ techniques for measuring the subjet energy scale have been developed previously [50], but these must be extended to the energy resolution and possibly the mass scale and resolution in the future.

10.3 Further Improvements to the Jet Mass

One natural extension of m^{TAS} is to combine it with the calorimeter-based jet mass as described in Sec. 10.1. This is not as simple as for m^{TA} because the use of local calorimeter information induces a correlation between the calorimeter-based jet mass response and m^{TAS}. This is shown explicitly in Fig. 18. The correlation is strongest for top-quark jets ($\rho \sim 0.57$) than for W-boson jets ($\rho \sim 0.24$) due to the higher subjet multiplicity. For approximately Gaussian distributed response distributions, a generalized linear combination is still nearly optimal, though Eq. 8 must be generalized to include correlations.

[12] The R = 0.2 subjets which survives the trimming procedure. A method based on independently clustered track-jets matched to calorimeter subjets was also studied, but found to result in a slightly worse resolution (and has an algorithmic complication for unmatched track-jets and calorimeter subjets).

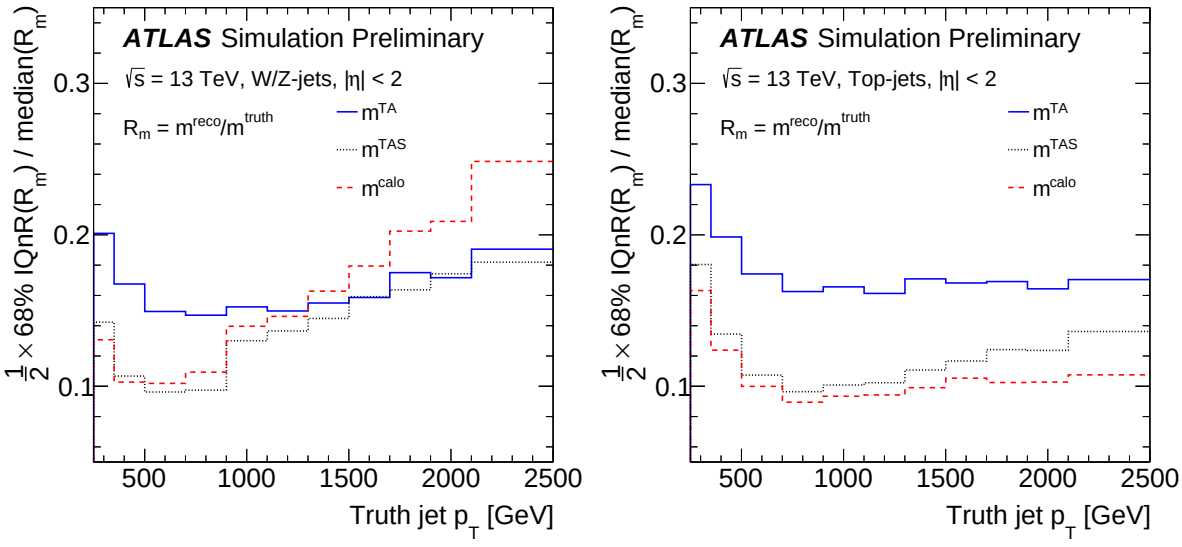


Figure 17: The jet mass resolution as a function of jet p_{T} for jets produced from boosted W bosons (left) and from boosted top quarks (right). Three different jet mass reconstruction algorithms are displayed: the track-assisted jet mass (m^{TA}), the track-assisted subjet mass (m^{TAS}) and the calorimeter-based jet mass (m^{calo}).

It may be possible to further reduce the local charged-to-neutral fluctuations by decreasing the subjet size. A particle-flow like algorithm is recovered in the limit that the subjet size tends to zero. While this introduces challenges like the ones described at the end of Sec. 10.2, it also may degrade performance at very high p_{T} where tracks and calorimeter clusters cannot be uniquely matched with high efficiency.

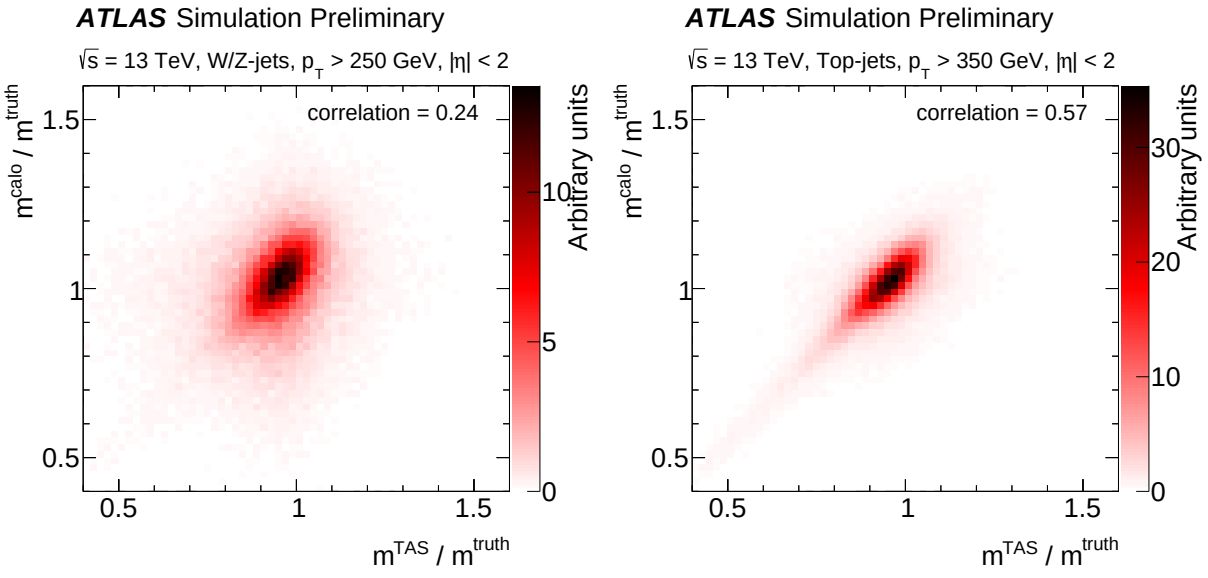


Figure 18: The calorimeter-based jet mass response vs the track-assisted subjet mass response distribution for W/Z-jets (left) and top-jets (right). The linear correlation coefficient is displayed in the top right of both plots.

In addition to reducing the resolution from mitigating local charged-to-neutral fluctuations, there may also

be potential gains from improving the use of the tracks at high p_{T}. Figure 19 shows the track-assisted jet mass response when replacing detector-tracks with the charged-particles in the associated particle-level truth jet. The resolution of the charged-particles m^{TA} response is a lower bound on potential improvements from tracking. For instance, per-track covariance matrices could be used to reduce the impact of poorly reconstructed tracks.

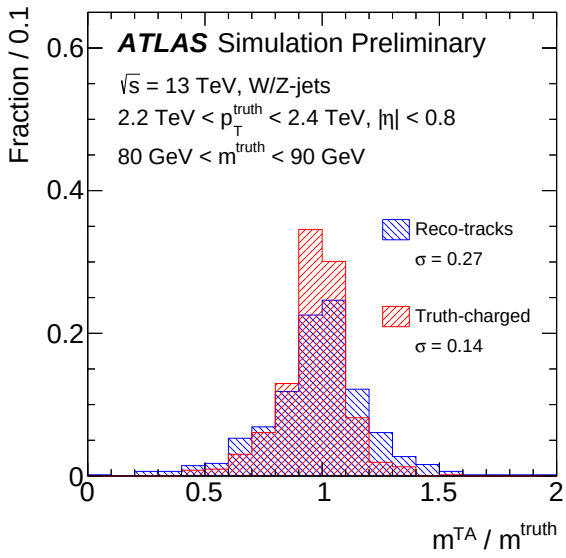


Figure 19: The track-assisted jet mass response distribution for 2.2 TeV$< p_{\mathrm{T}} <$ 2.4 TeV and 80 GeV$< m^{\mathrm{truth}} <$ 90 GeV. The blue histogram filled with lines slanted down to the right uses reconstructed tracks to define m^{TA} while the red histogram with lines slanted up to the right uses promptly produced charged particles with $c\tau > 30$ ps, excluding muons.

11 Conclusions

This note has described the reconstruction of the mass of large-radius jet. Two definitions of the reconstructed jet mass are presented and each of the definitions has its own advantage. The calorimeter-based jet mass is deduced from the calorimeter-cell cluster constituents of the jet while the track-assisted jet mass exploits the fine granularity of the tracker to reconstruct the jet mass. The jet mass scale calibration procedure is described for the calorimeter-based jet mass and the track-assisted jet mass and calibrated jet masses are within 3% of the particle-level truth jet mass. The performance of all jet mass definitions are evaluated by comparing the fractional resolution of each jet mass definition and the systematic uncertainties on the track-assisted jet mass is compared to the calorimeter-based jet mass. To estimate the relative jet mass scale and resolution between simulation and data, an in-situ forward folding method is employed on events enriched with top-jets and W-jets and this method is extended to measure the relative jet energy scale and resolution. Techniques to improve the track-assisted jet mass through a linear combination with the calorimeter-based jet mass and employing the track-assisted jet mass method on subjets are discussed.

References

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

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

[3] 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: http://cdsweb.cern.ch/record/2120133.

[4] ATLAS Collaboration, *Modelling of the* $t\bar{t}H$ *and* $t\bar{t}V(V = W, Z)$ *processes for* $\sqrt{s}$ = 13 *TeV ATLAS analyses*, ATL-PHYS-PUB-2016-005, 2016, URL: https://cds.cern.ch/record/2120826.

[5] ATLAS Collaboration, *Multi-boson simulation for* 13 *TeV ATLAS analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: http://cdsweb.cern.ch/record/2119986.

[6] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823, arXiv: 1005.4568 [physics.ins-det].

[7] S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[8] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **1006** (2010) 043, arXiv: 1002.2581 [hep-ph].

[9] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[10] T. Sjöstrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, arXiv: hep-ph/0603175 [hep-ph].

[11] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. D **82** (2010) 074018, arXiv: 1005.3457 [hep-ph].

[12] M. Czakon, P. Fiedler and A. Mitov,
Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$,
Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph].

[13] M. Czakon and A. Mitov,
NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction,
JHEP **1301** (2013) 080, arXiv: 1210.6832 [hep-ph].

[14] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **1212** (2012) 054, arXiv: 1207.0236 [hep-ph].

[15] P. Bärnreuther, M. Czakon and A. Mitov, *Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$*, Phys. Rev. Lett. **109** (2012) 132001, arXiv: 1204.5201 [hep-ph].

[16] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. B **710** (2012) 612, arXiv: 1111.5869 [hep-ph].

[17] M. Czakon and A. Mitov,
Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders,
Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph].

[18] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. **D83** (2011) 091503, arXiv: 1103.2792 [hep-ph].

[19] N. Kidonakis,
Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-,
Phys. Rev. D **82** (2010) 054018, arXiv: 1005.4451 [hep-ph].

[20] N. Kidonakis, *NNLL resummation for s-channel single top quark production*,
Phys. Rev. D **81** (2010) 054028, arXiv: 1001.5034 [hep-ph].

[21] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007,
arXiv: 0811.4622 [hep-ph].

[22] S. Catani et al.,
Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO,
Phys. Rev. Lett. **103** (2009) 082001, arXiv: 0903.2120 [hep-ph].

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

[24] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **1504** (2015) 040,
arXiv: 1410.8849 [hep-ph].

[25] ATLAS Collaboration, *ATLAS Run 1 Pythia 8 tunes*, ATL-PHYS-PUB-2014-021, 2014,
URL: http://cds.cern.ch/record/1966419.

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

[27] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*,
JHEP **0804** (2008) 063, arXiv: 0802.1189 [hep-ph].

[28] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, JHEP **1002** (2010) 084,
arXiv: 0912.1342 [hep-ph].

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

[30] ATLAS Collaboration,
Jet energy measurement with the ATLAS detector in proton–proton collisions at $\sqrt{s} = 7$ *TeV*, Eur. Phys. J. C **73** (2013) 2304, arXiv: 1112.6426 [hep-ex].

[31] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in roton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 17, arXiv: 1406.0076 [hep-ex].

[32] 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 **1309** (2013) 076, arXiv: 1306.4945 [hep-ex].

[33] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. **B659** (2008) 119, arXiv: 0707.1378 [hep-ph].

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

[35] M. Son, C. Spethmann and B. Tweedie, *Diboson-Jets and the Search for Resonant Zh Production*, JHEP **08** (2012) 160, arXiv: 1204.0525 [hep-ph].

[36] A. Katz, M. Son and B. Tweedie, *Jet Substructure and the Search for Neutral Spin-One Resonances in Electroweak Boson Channels*, JHEP **03** (2011) 011, arXiv: 1010.5253 [hep-ph].

[37] S. Schaetzel and M. Spannowsky, *Tagging highly boosted top quarks*, Phys. Rev. **D89** (2014) 014007, arXiv: 1308.0540 [hep-ph].

[38] T. Plehn et al., *Stop Reconstruction with Tagged Tops*, JHEP **10** (2010) 078, arXiv: 1006.2833 [hep-ph].

[39] T. Plehn, G. P. Salam and M. Spannowsky, *Fat Jets for a Light Higgs*, Phys. Rev. Lett. **104** (2010) 111801, arXiv: 0910.5472 [hep-ph].

[40] A. J. Larkoski, F. Maltoni and M. Selvaggi, *Tracking down hyper-boosted top quarks*, JHEP **06** (2015) 032, arXiv: 1503.03347 [hep-ph].

[41] S. Bressler et al., *Hadronic Calorimeter Shower Size: Challenges and Opportunities for Jet Substructure in the Superboosted Regime*, (2015), arXiv: 1506.02656 [hep-ph].

[42] ATLAS Collaboration, *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at sqrt(s) = 8 TeV*, 2015,
URL: http://cds.cern.ch/record/2001682.

[43] ATLAS Collaboration,
A study of the material in the ATLAS inner detector using secondary hadronic interactions, JINST **7** (2012) P01013, arXiv: 1110.6191 [hep-ex].

[44] ATLAS Collaboration, *Measurement of track reconstruction inefficiencies in the core of jets via pixel dE/dx with the ATLAS experiment using* $\sqrt{s} = 13$ *TeV pp collision data*, ATL-PHYS-PUB-2016-007, 2016, URL: http://cds.cern.ch/record/2140460.

[45] ATLAS Collaboration,
Early Inner Detector Tracking Performance in the 2015 Data at $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-051, 2015, URL: http://cdsweb.cern.ch/record/2110140.

[46] ATLAS Collaboration,
Study of alignment-related systematic effects on the ATLAS Inner Detector track reconstruction,
ATLAS-CONF-2012-141, 2012, URL: http://cdsweb.cern.ch/record/1483518.

[47] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector and its Performance in 2012*,
ATLAS-CONF-2014-047, 2014, URL: http://cdsweb.cern.ch/record/1741021.

[48] ATLAS Collaboration, *Measurement of the jet mass scale and resolution uncertainty for large radius jets at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, 2016,
URL: http://cds.cern.ch/record/2139642.

[49] E. Lehmann and G. Casella, *Theory of Point Estimation*, Springer Verlag, 1998,
ISBN: 0387985026.

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