New techniques for reconstructing, calibrating and identifying hadronic objects with ATLAS

Sven Menke
on behalf of the ATLAS Collaboration
Max-Planck-Institut für Physik,
Boltzmannstr. 8, 85748 Garching bei München, Germany

Experimental uncertainties related to hadronic object reconstruction can limit the precision of physics analyses at the LHC, and so improvements in performance have the potential to broadly increase the impact of results. Recent refinements to reconstruction, calibration and identification procedures for ATLAS jets and missing transverse momentum result in reduced uncertainties, improved pile-up stability and other performance gains. In this contribution, highlights of these developments are presented.

1 Introduction

Hadronic objects in ATLAS [1] are mainly identified with jets, that are made of topological clusters [2] in the calorimeters and charged particle tracks from the inner detector, which are combined to Particle Flow Objects (PFO) [3] for small ΔR [a] jets and Track Calo Clusters (TCC) [4] for large ΔR jets. Recently, Unified Flow Objects (UFO) [5] have been introduced to replace both PFO and TCC, taking a combined approach depending on the environment to make the best of both algorithms.

The presence of pile-up, underlying soft proton-proton interactions, that add to the hard scatter event of interest, make identifying and calibrating hadronic objects challenging tasks. The pile-up itself can be used to improve its treatment in jets (see [6,7]). In the following, a new approach to use time as additional discriminant in the topological clustering is described in Section 2. Recent developments for jet calibration and identification are reported in Sections 3 and 4, respectively and the performance of missing transverse momentum is highlighted in Section 5.

2 Time as new discriminant

The ATLAS calorimeters not only measure the energy of particle showers but also have excellent time resolution. Including the time spread caused by the beam spot distribution around the nominal interaction point, the signal time of high energetic showers can be determined typically well below the 1 ns level. Electronics noise and signals from interactions from bunch crossings other than the triggered one (out-of-time pile-up) distort the energy measurement but also manifest themselves in signal times different from 0 ns. This feature can be used to discriminate signals dominated by out-of-time pile-up directly at the level of the formation of topological clusters [8], which form the basis for jets, electrons, photons and taus.

[a] $\Delta R = \sqrt{\Delta y^2 + \Delta\varphi^2}$ is the distance measure in jet clustering using the quadratic sum of the difference in rapidity, y, and azimuth, φ, between proto-jets.

In addition to the well established scheme of seeding clusters with cells above $4\,\sigma$ in absolute energy, expanding around the seed cell's neighbours and also around their subsequent neighbours and neighbour's neighbours and so on, as long as the cell's absolute energy exceeds $2\,\sigma$ in absolute energy and finally including a guard ring of cells that fail the $2\,\sigma$ cut [2], the time of any cell above $4\,\sigma$ in absolute energy is now also required to be within a window of $\pm 12.5\,\mathrm{ns}$ to join and expand the cluster [8].

Long lived particles could, however, be the source of high energetic deposits in the calorimeter that stem from decays long after the triggered event. In order not to harm the search for such exotic particles, the time cut is restricted to cells with energies below $20\,\sigma$ – i.e. there is no restriction based on the measured time for very significant positive signals.

This clustering scheme is now default in Run-3 data taking and reduces the amount of jets dominated by out-of-time pile-up by $\sim 50\%$ at $p_{\mathrm{T}} \simeq 20\,\mathrm{GeV}$ and by $\sim 80\%$ for $p_{\mathrm{T}} > 50\,\mathrm{GeV}$, with p_{T} being the transverse momentum of the jet. At the same time the number of in-time jets remains unchanged and the energy resolution improves a little (by about 5%). Fake tau-, electron- and photon-candidates are reduced as well.

3 Calibration methods

Local hadronic calibration (a.k.a. local cell weighting or LCW) [2] is a 4-step procedure to bring the energy scale of topological clusters from the raw "EM"-scale to the particle-level "LC"-scale. The four steps are classification as electromagnetic or hadronic in nature, corrections for hadronic non-compensation, corrections for out-of-cluster deposits inside calorimeters and for deposits outside the calorimeters (dead-material corrections).

Jets [9] can use either EM- or LC-scale objects (clusters or flow objects) and are then calibrated in several steps: a jet-area correction accounts for pile-up, the energy scale is corrected to particle-level as deduced from di-jet simulations, the flavour dependency and resolution get improved by a global correction step (keeping the average energy unchanged) and data are corrected in-situ from the measured p_{T}-balance of jets in multi-jet and $\mathrm{Z}^0/\gamma + \mathrm{jet}$ events to match the simulation.

3.1 Cluster calibration with neural networks

To explore the applicability of neural networks to calorimeter calibration, the non-compensation correction step (and an implicit classification step) of the local hadronic calibration is attempted with machine learning instead of the traditional look-up tables [10]. A similar set of cluster moments as in the single-pion-simulation based legacy method is used in the training of neural networks (NN), but on fully simulated di-jet events, including pile-up and with the pile-up sensitive quantities (the number of primary vertices, N_{PV}, and the average number of interactions per bunch-crossing, μ) added to the list of NN inputs. Both, linearity and resolution, improve in the NN-based methods over the entire phase-space. In particular, the impact of pile-up is substantially reduced since it is explicitly accounted for in the training, while the legacy LCW does not protect against pile-up.

3.2 New techniques for jet calibration

Like for the cluster calibration, several jet calibration steps are re-derived with NN. As an alternative to the Global Sequential Calibration (GSC) that was successfully used in Run-2 after setting the average energy scale (MCJES) to reduce flavour dependency and improve resolution, a Global Neural Network Calibration (GNNC) is the new default for Run-3 [11]. The main difference of GNNC compared to GSC is the possibility to use correlated quantities in the training of NN for GNNC, whereas the six correction steps in GSC are performed sequentially and need to rely on largely uncorrelated observables. At the same time, the GNNC approach lifts the restriction to keep the average scale unchanged. As a result, the linearity shows closure within 1% for GNNC (while GSC inherits a small non-closure up to 3% at low p_{T} from MCJES) and the resolution is improved by $15 - 25\%$.

The MCJES and GNNC steps are followed by an in-situ η-intercalibration in multi-jet events to correct the data and $Z^0 + \text{jet}$ and $\gamma + \text{jet}$ data are compared to simulation to deduce the final jet energy scale (JES) and its uncertainty.

The Missing-E_{T} Projection Fraction (MPF) is used to calculate the p_{T}-balance between the Z^0/γ and the full hadronic recoil. This method is preferred for lower p_{T} jets impacted by pile-up.

For high p_{T} jets, the Direct Balance (DB) method, where the p_{T} balance of the photon w.r.t. one jet only is preferred over the MPF method and applicable to measure the uncertainty for individual b-tagged jets. The new calibration improvements show that the additional JES uncertainty on top of the general JES uncertainty is of the order of $O(1\%)$, leading to a total JES uncertainty of $O(1.5\%)$ for b-tagged jets.

3.3 Calibration of E and m for large ΔR-jets

Jets clustered with large ΔR are important for boosted topologies of heavy resonances. The asymmetric response in energy and mass requires dedicated calibrations for both observables, which remain highly correlated. A complex, deep neural network (DNN) is used to obtain both calibrations simultaneously [12].

The inputs to the DNN are jet kinematics, substructure variables, detector-level energy or transverse-momentum fractions and the pile-up-sensitive quantities N_{PV} and μ. To aid the DNN converging faster, the input is annotated by 11 Gaussian η-dependent weights, splitting the phase-space smoothly into regions dominated by different detector-related effects. The loss function is the sum of negative log-likelihood predicting mode and width of Gaussian distributions in energy, E, and mass, m. Initially, both quantities are trained together on 270 million fully simulated di-jet events. A fork allows to optimise E and m separately, while freezing the respective other quantity. A residual connection for m improves the focus on the most important inputs for the mass.

The DNN outperforms the standard calibration in energy- and mass-scale closure and in resolution for both E and m. The closure stays within 2% and the resolution improves by $> 30\%$ for $p_{\mathrm{T}} > 500\,\mathrm{GeV}$. An important cross-check is the performance on topologies not used in the calibration. This is demonstrated for boosted heavy bosons for jets calibrated by the DNN method compared to that of uncalibrated jets and jets after standard calibration.

4 Identification methods

Distinguishing jets initiated by different particles (light quark, gluon, heavy boson, top-quark) is extremely important to identify the final states. ATLAS uses machine learning (with different network architectures) for jet tagging on simulated samples. Trained either on high-level jet-based quantities – restricted to infrared/collinear safe observables for some – or, with additional information from the jet constituents (flow objects). Performance is evaluated by comparing background rejection rate for a given signal efficiency for different taggers while model dependence is probed by applying the standard-sample trained tagger on different simulated samples with alternative showering and harmonisation modelling. Constituent-based quark-gluon tagging [13], W-boson tagging [14] and boosted top-quark tagging [15] are recent ATLAS developments. The taggers employing networks with constituent features (like PartT, DePArtT and P.Net) are best in background rejection but also show the largest model dependence. Restricting to infrared/collinear safe observables (as in EFN) achieves the lowest model dependence but also the lowest background rejection rates.

5 Performance of missing transverse momentum

The two-dimensional missing transverse momentum vector, $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$, is derived from the corresponding vectors for each "hard" object (obj), such as $\mathrm{e}, \gamma, \tau, \mathrm{jet}$ and μ candidates in the event and a remaining "soft" term from tracks not used already for the hard objects:

$$\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} = -\boldsymbol{p}_{\mathrm{T}}^{\mathrm{hard}} - \boldsymbol{p}_{\mathrm{T}}^{\mathrm{soft}}, \text{with } \boldsymbol{p}_{\mathrm{T}}^{\mathrm{hard}} = \sum_{\mathrm{obj}=\mathrm{e},\gamma,\tau,\mu,\mathrm{jet}} \boldsymbol{p}_{\mathrm{T}}^{\mathrm{obj}} \text{ and } \boldsymbol{p}_{\mathrm{T}}^{\mathrm{soft}} = \sum_{\text{unused tracks}} \boldsymbol{p}_{\mathrm{T}}^{\mathrm{track}}. \tag{1}$$

The scalar transverse momentum sum is used to evaluate the scale of $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$:

$$\sum p_{\mathrm{T}} = \sum_{\mathrm{obj=e},\gamma,\tau,\mu,\mathrm{jet}} p_{\mathrm{T}}^{\mathrm{obj}} + \sum_{\text{unused tracks}} p_{\mathrm{T}}^{\mathrm{track}}. \tag{2}$$

The Run-2 performance has been evaluated with the full Run-2 dataset and the use of PFO for jets[16]. In particular the requirements for the unused tracks in the soft term needed refinements to avoid double-counting. $\mathrm{Z}^0 \to \mu^+\mu^-$ and $\mathrm{Z}^0 \to \mathrm{e}^+\mathrm{e}^-$ events without expected real $p_{\mathrm{T}}^{\mathrm{miss}}$ serve to evaluate the performance.

Improvements by up to 76% and 51% with respect to earlier evaluations on smaller datasets are observed in scale and resolution uncertainties, respectively. Object-based uncertainties are now used to compute a true missing transverse momentum significance instead of the prior used, event-based signficance, the ratio of $p_{\mathrm{T}}^{\mathrm{miss}}$ over the square-root of the scalar sum of p_{T} of jets, which assumes calorimeter-like resolution and hence is approximate only. This results in better separation of signal and background for the object-based significance compared to the event-based significance proxy.

6 Conclusions

The reconstruction and calibration of hadronic objects in ATLAS is a very rich and active field. Reducing the impact of pile-up remains the biggest challenge, but new techniques, like using signal time in the calorimeters as discriminant help. Machine-learning based calibration methods start to replace their legacy counterparts for energy and mass of hadronic objects. Improvements in the jet calibration methods lead to $O(1\%)$ precision in the jet energy scale and $O(15-30\%)$ better resolution in energy and mass in Run-2. The additional b-jet energy scale uncertainty was measured to $O(1\%)$ precision. Taggers utilising machine-learning for q/g, heavy-boson and t-quark identification that employ constituent-level information outperform those with high-level jet information. But at the same time, the model dependence is found to be larger for constituent based taggers. Missing transverse momentum reconstruction benefits from the calibration advancements – especially from jets. The object-based significance sharpens the discrimination power of $p_{\mathrm{T}}^{\mathrm{miss}}$ already in Run-2. Future Run-3 analyses will benefit from these improvements.

1. ATLAS Collaboration, *JINST* **3** S08003 (2008).
2. ATLAS Collaboration, *Eur. Phys. J.* C **77** 490 (2017).
3. ATLAS Collaboration, *Eur. Phys. J.* C **77** 466 (2017).
4. ATLAS Collaboration, ATL-PHYS-PUB-2017-015 (2017).
5. ATLAS Collaboration, *Eur. Phys. J.* C **81** 334 (2021).
6. ATLAS Collaboration, arXiv:2407.10819 (2024).
7. A. Pirttikoski, ATL-PHYS-PROC-2024-066 (2024).
8. ATLAS Collaboration, *Eur. Phys. J.* C **84** 455 (2024).
9. ATLAS Collaboration, *Eur. Phys. J.* C **81** 689 (2021).
10. ATLAS Collaboration, ATL-PHYS-PUB-2023-019 (2023).
11. ATLAS Collaboration, *Eur. Phys. J.* C **83** 761 (2023).
12. ATLAS Collaboration, *Mach. Learn. Sci. Technol.* **5** 035051 (2024).
13. ATLAS Collaboration, ATL-PHYS-PUB-2023-032 (2023).
14. ATLAS Collaboration, ATL-PHYS-PUB-2023-020 (2023).
15. ATLAS Collaboration, *JINST* **19** P08018 (2024).
16. ATLAS Collaboration, arXiv:2402.05858 (2024).