XXX International Workshop on Deep-Inelastic Scattering and Related Subjects

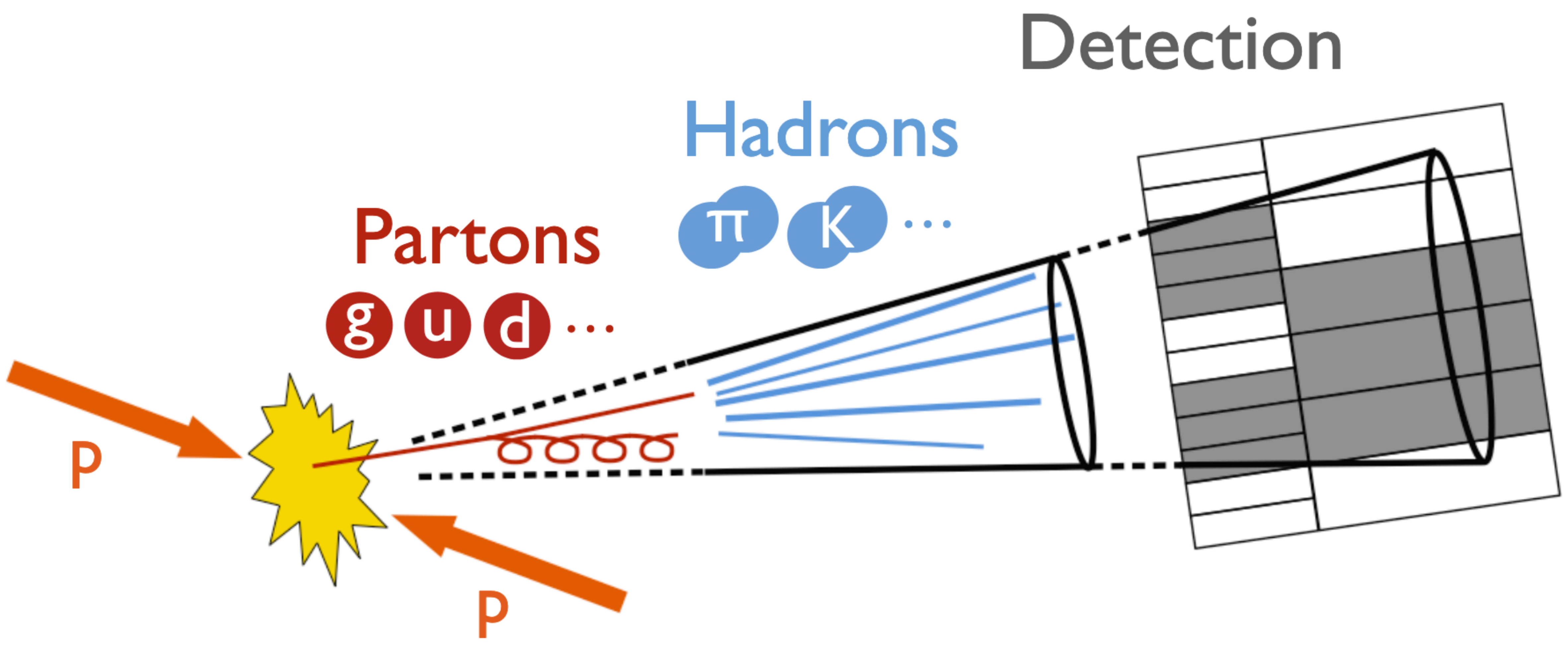
Redefining Performance: New Techniques for ATLAS Jet & MET Calibration

Naseem Bouchhar

29 March 2023

Jets in High Energy Physics

- **Jets are collimated groups of particles that result from the production of high energy quarks and gluons.**
- Jets in the ATLAS experiment are reconstructed using the anti-k_t algorithm, using particle flow (PFlow) jets which consists of reconstructing jets from particle tracks and calorimeter energy deposits.
- These jets are used all through ATLAS so must be studied and understood well and thus they must be calibrated to account for a number of factors.
- The goal of this talk is to present the improvements and new techniques made for ATLAS jet calibration to reduce the jet energy scale (JES) uncertainties.



Calibration Chain

ATLAS has a dedicated jet calibration chain that consists of the following steps:

Jet Reconstruction →

Jet detection using calorimeter and track based inputs.

Area-Based Pile-Up Correction →

Pile-up correction applied as a function of jet area and pile-up density.

Improvement in pile-up density calculation.

Residual Pile-Up Correction

Remove residual pile-up dependency as a function of N_{PV} and μ.

Correlations of functions accounted for.

→ **Absolute MC-based Calibration** →

Calibration of energy and direction to the particle-level scale.

Better fit method used for calibration factor.

Global Calibration →

Reduces flavour dependence and improves jet resolution.

Allowing more inputs for the measurement using neural network.

Residual in-situ Calibration

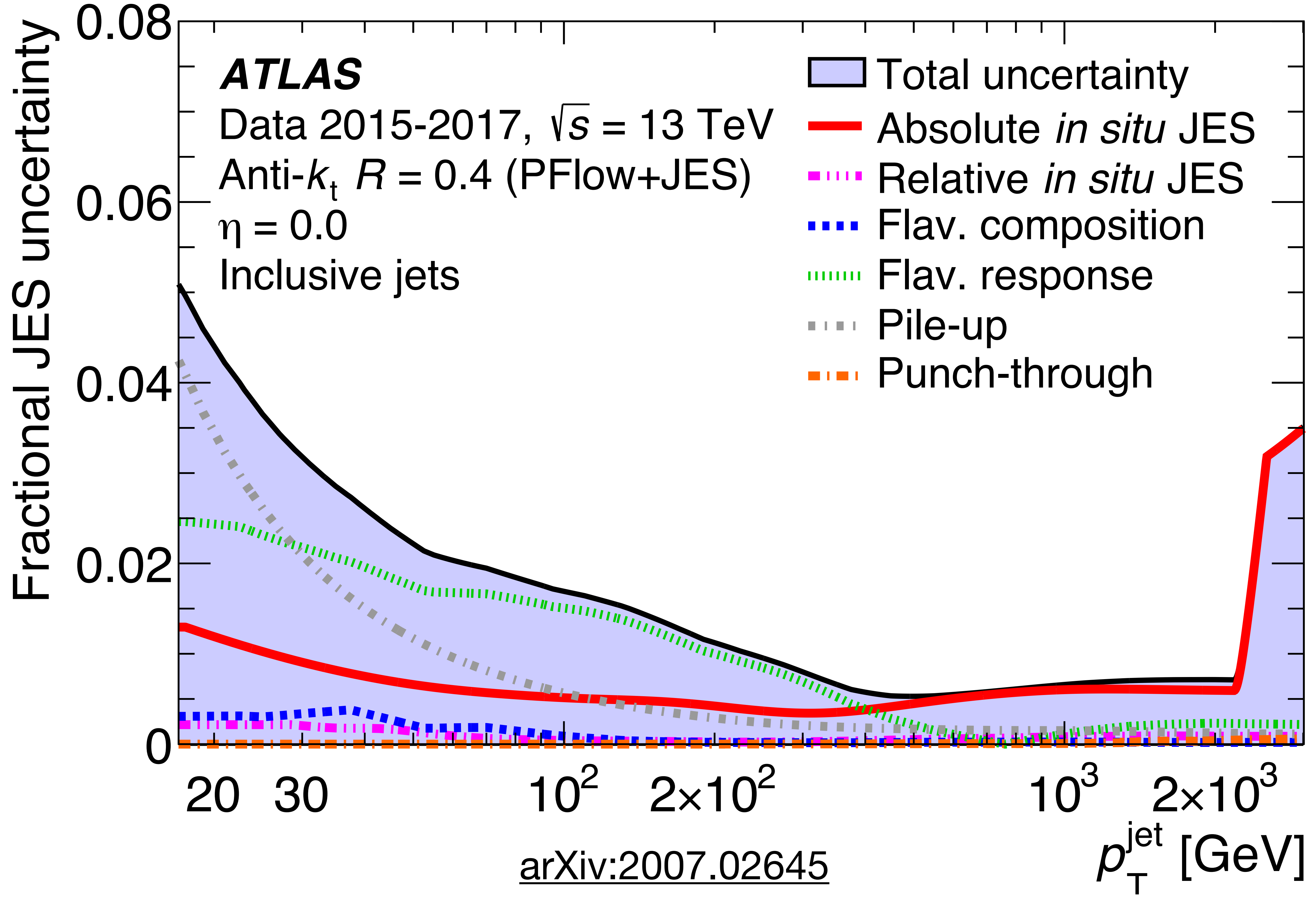
Residual correction to account for differences between MC and data.

- **Modelling uncertainty improvement in η-intercalibration.**
- **bJES studied for γ+jet events.**

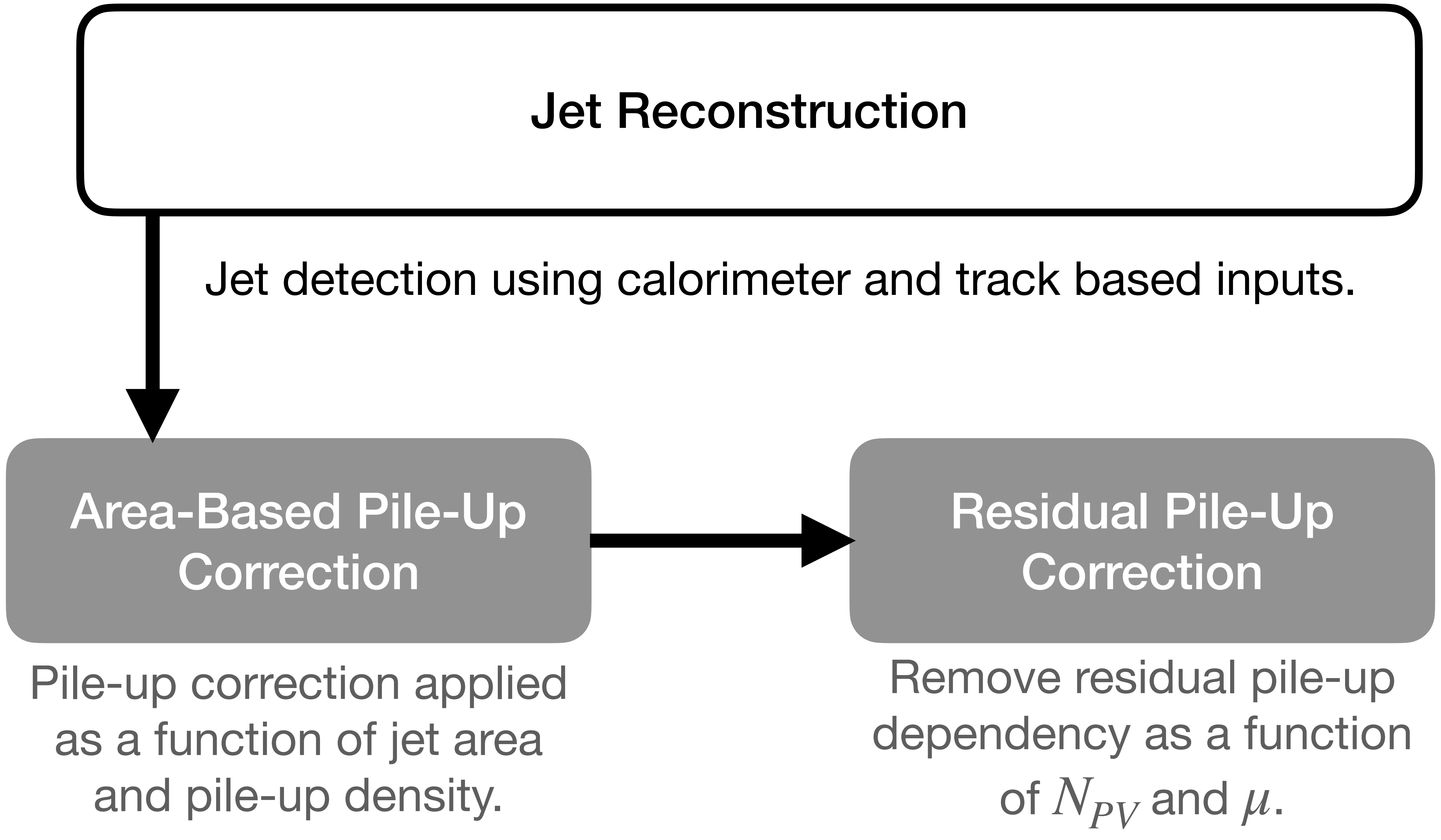
All the calibrations are applied to MC and data except the in-situ which is only applied to data.

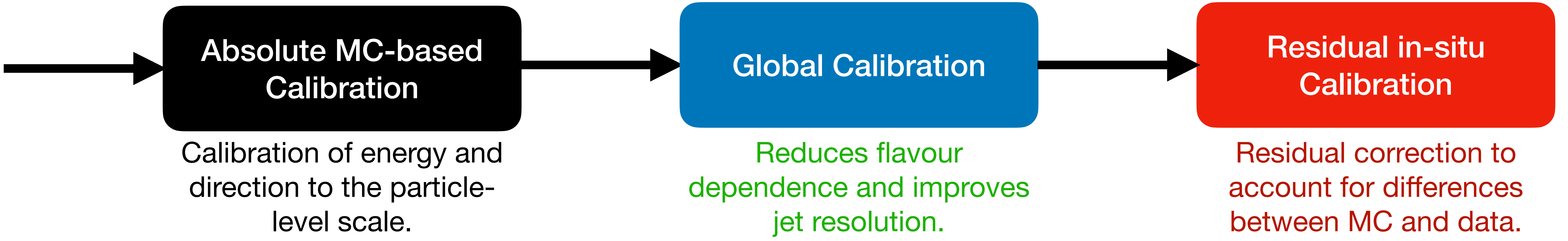
Calibration Chain

ATLAS has a dedicated jet calibration chain that consists of the following steps:



arXiv:2007.02645





Each step has a related JES uncertainty. The idea is to decrease these uncertainties as much as possible.

Area-based Pile-up Correction

- **The jet area *A* is a measure of the susceptibility of the jet to pile-up.**

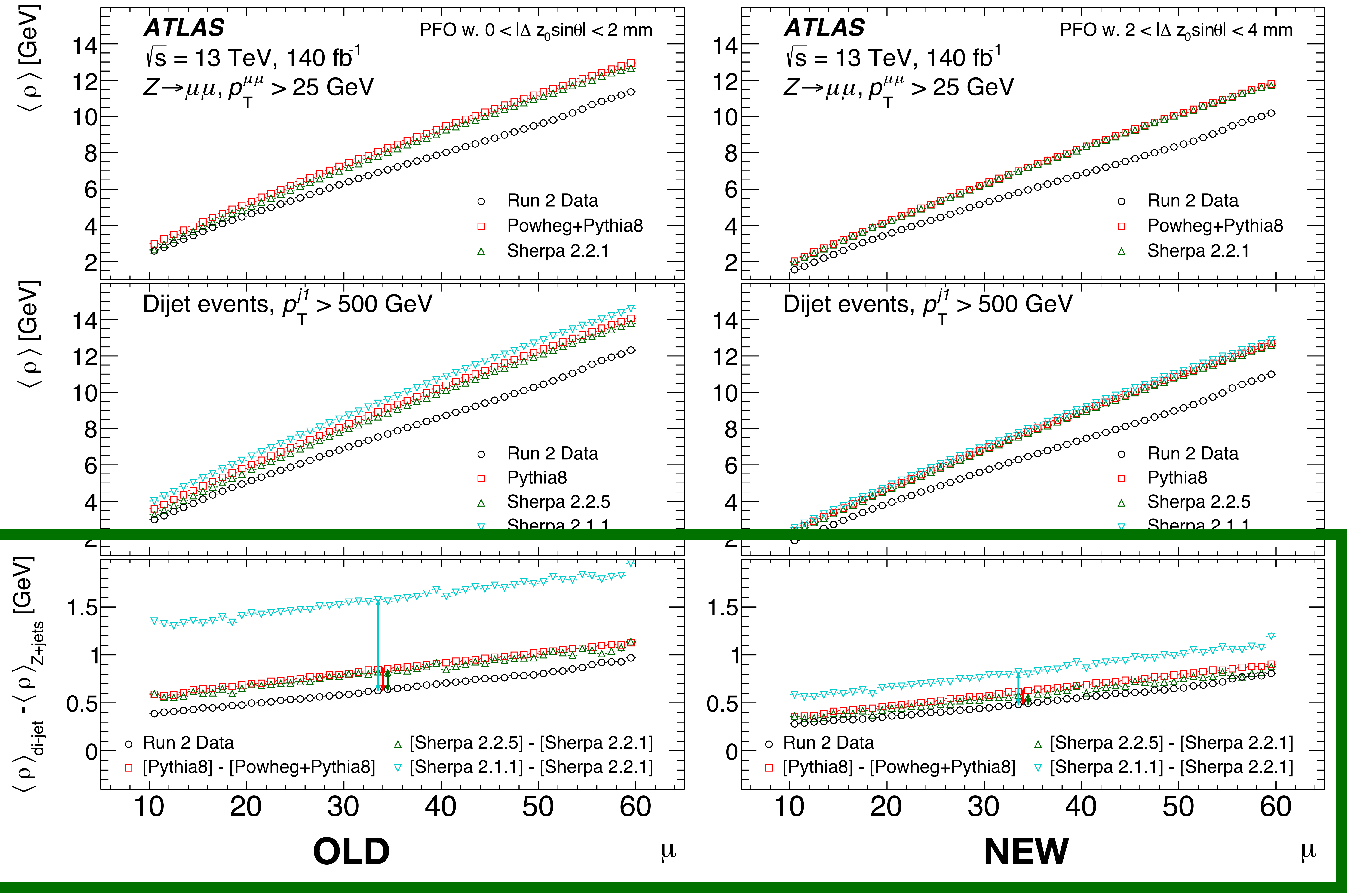
- Want to **mitigate pile-up by subtracting** the **average pile-up density from the energy** of the jet **using the jet area**:

$$< \rho > = \text{median}\left(\frac{p_{T,i}^{jet}}{A_i^{jet}}\right)$$

Where i enumerates over all considered jets.

- **Possible to have a bias in the estimation of ρ due to the event topology** e.g. ttbar events have more hard-scatter activity than Z+jets events
  - To avoid such biases we estimate ρ using a sideband around the primary vertex (being sure it is pileup)

CERN-EP-2023-028



Differences between MC generators greatly reduced when using sideband, and so the **uncertainty is reduced by a factor of 7** compared to what it was in Run 2.

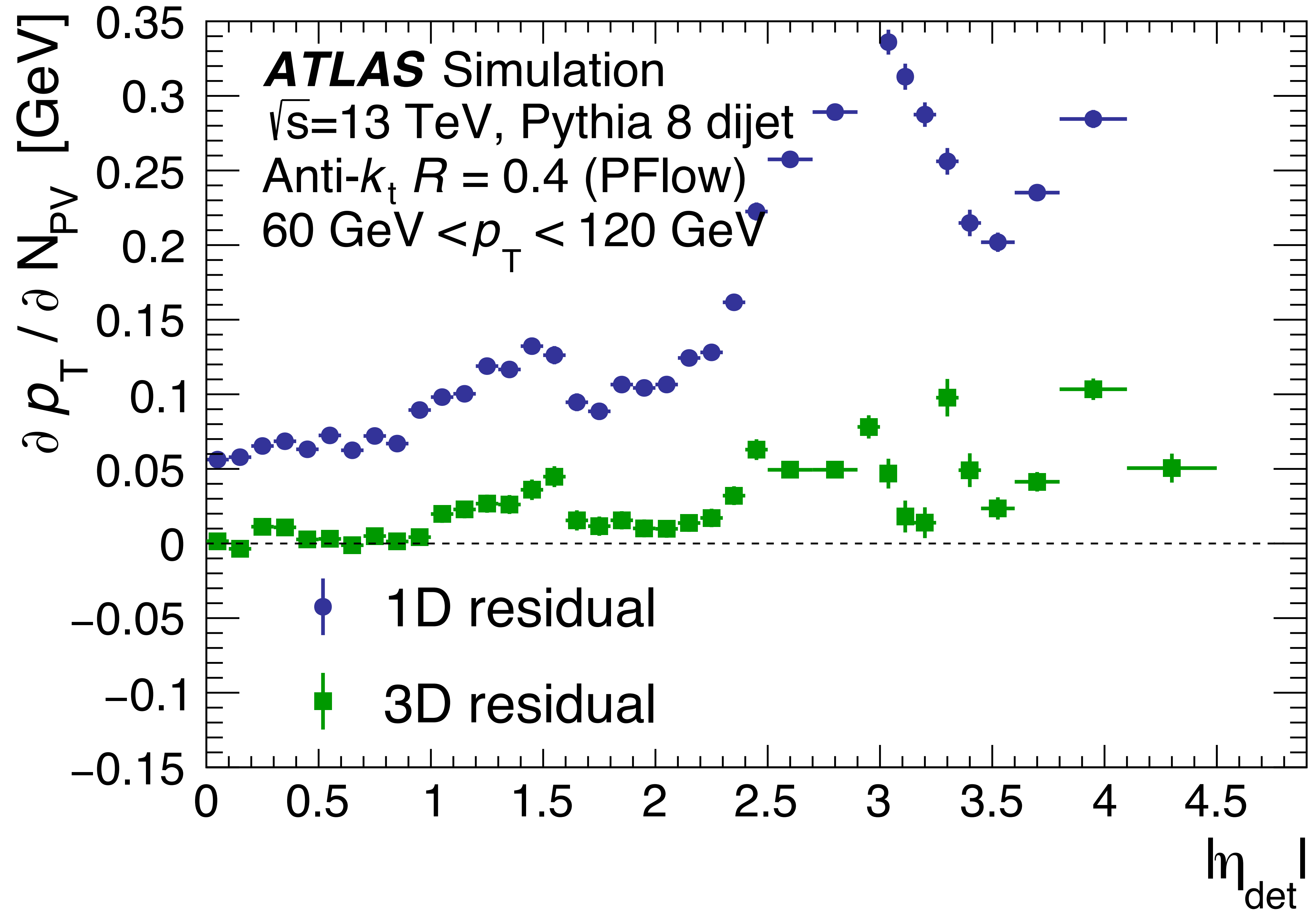
Residual Pileup Correction

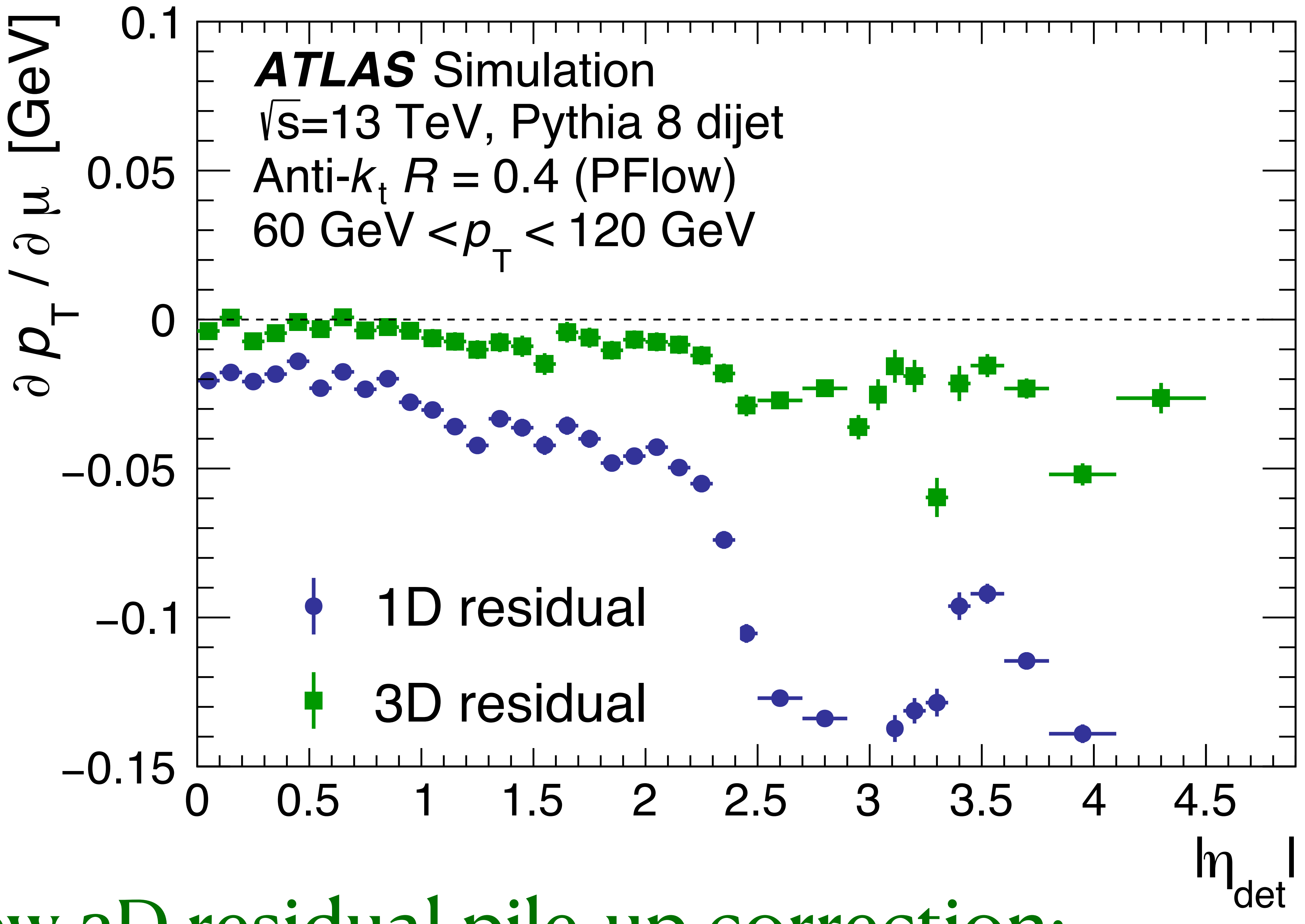
CERN-EP-2023-028

- After the jet-area correction, residual dependencies are found.
- Thus a **residual correction** is also applied to **reduce dependency on pile-up.**

- Original 1D residual pile-up correction:

$$p_T^{1D} = p_T^{area} - \alpha(N_{PV} - 1) - \beta\mu$$





- New 3D residual pile-up correction:

$$p_T^{3D} = p_T^{area} - \Delta p_T(N_{PV}, \mu, p_T^{area})$$

This new correction accounts for **correlations between** N_{PV} **and** μ while also applying an initial correction for detector effects.

Reduction of pile-up dependence.

Absolute MC-Based Calibration (MCJES)

- **MCJES corrects for factors to bring the overall jet energy scale to what it should be at particle-level. Factors such as:**
  - **Calorimeter response.**
  - **Energy losses in dead material.**
  - **Out-of-cone radiation effects.**

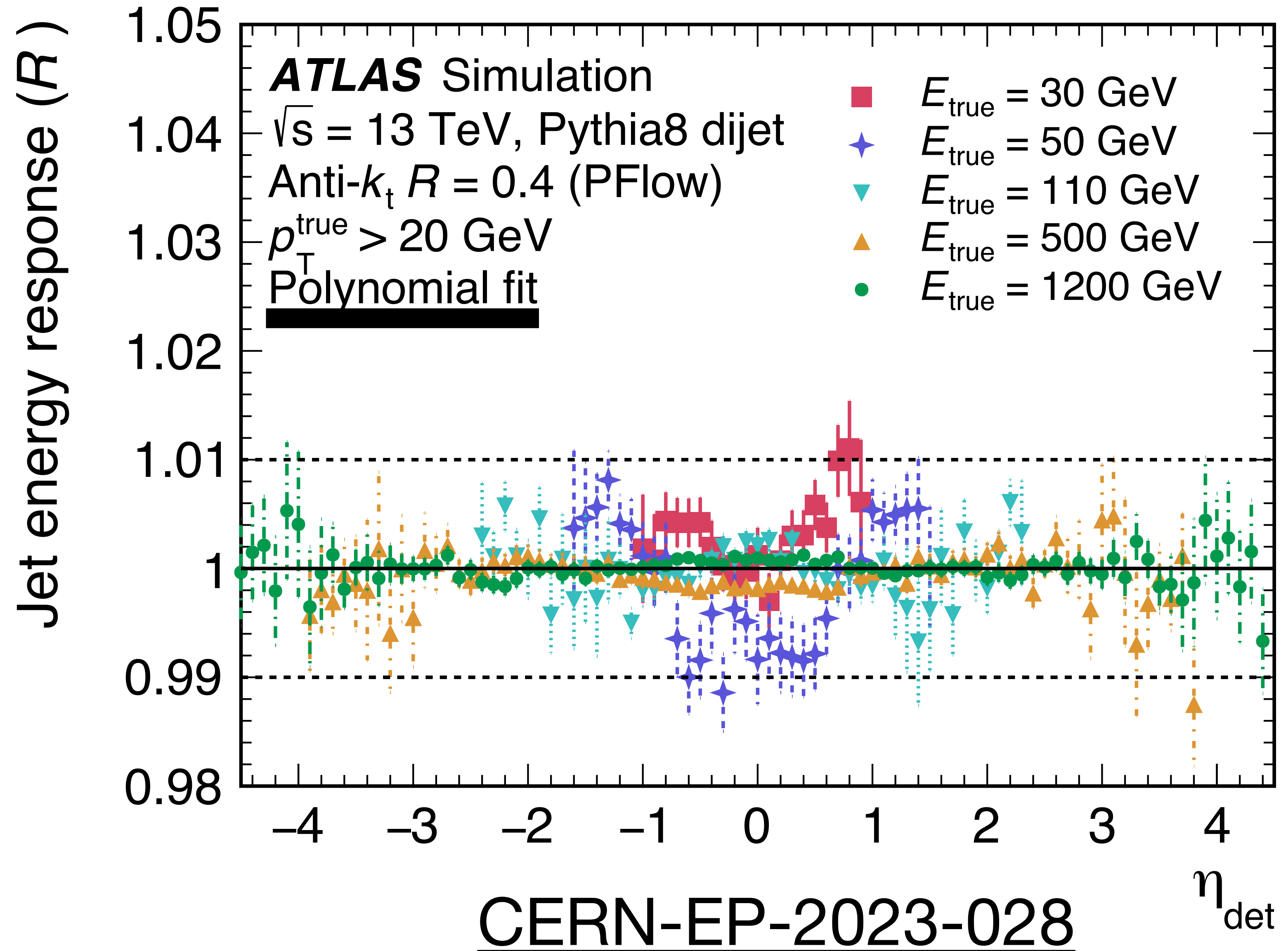
- The average jet energy response R is the mean of a gaussian fit to $\frac{E_{reco}}{E_{truth}}$ in bins of E_{true} and η_{det}.

The inverse of this function is the MCJES calibration factor.

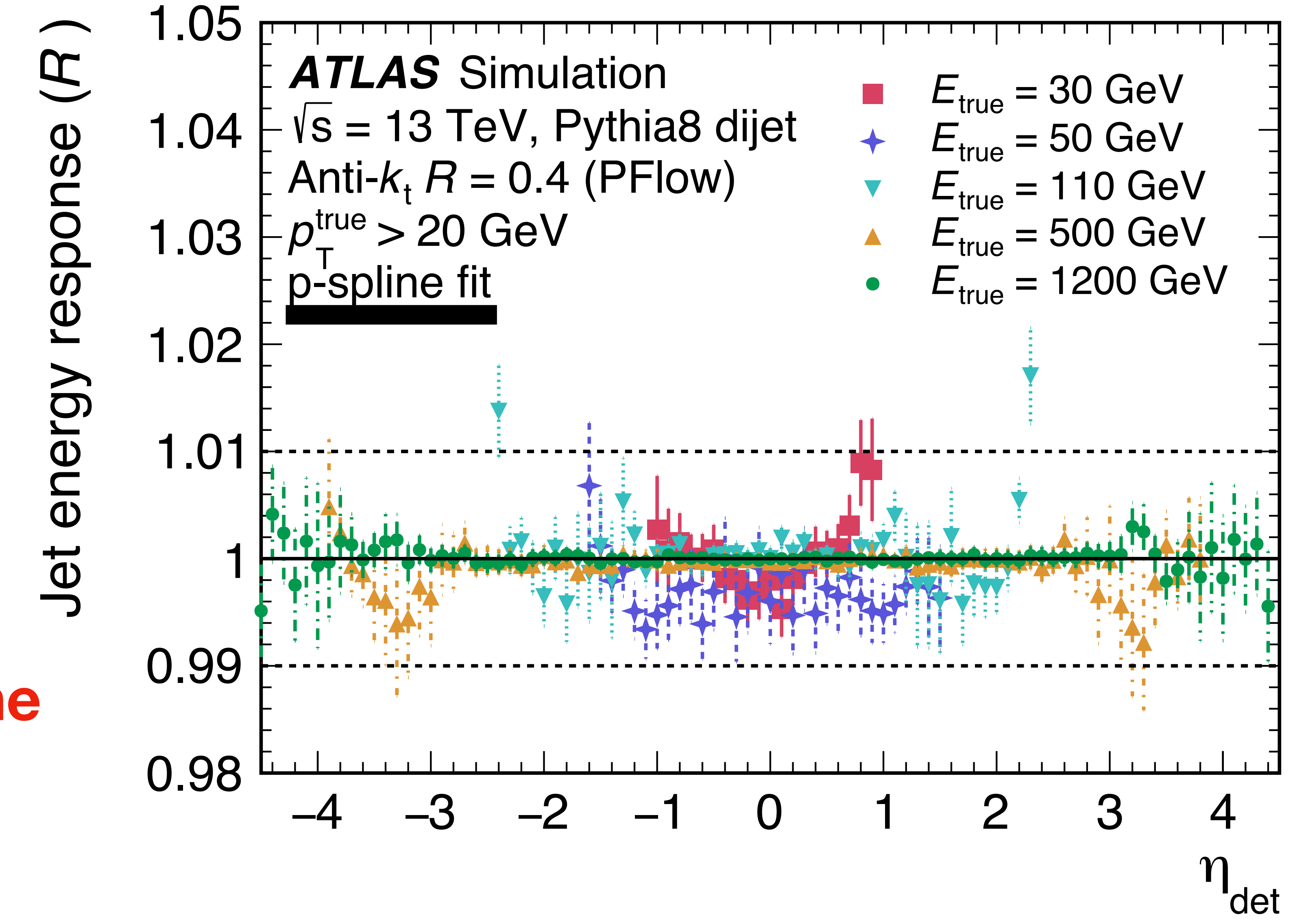
- **Two fit functions were tested to fit R:**
  - **Polynomial fits**
  - **Penalised Splines**

Splines provide better closure at low energies.

The MCJES also corrects the pseudo-rapidity to particle-level. Accounts for **biases in the jet η reconstruction.** Due to detector transitions, jets can have artificial skewed energy distributions.



CERN-EP-2023-028

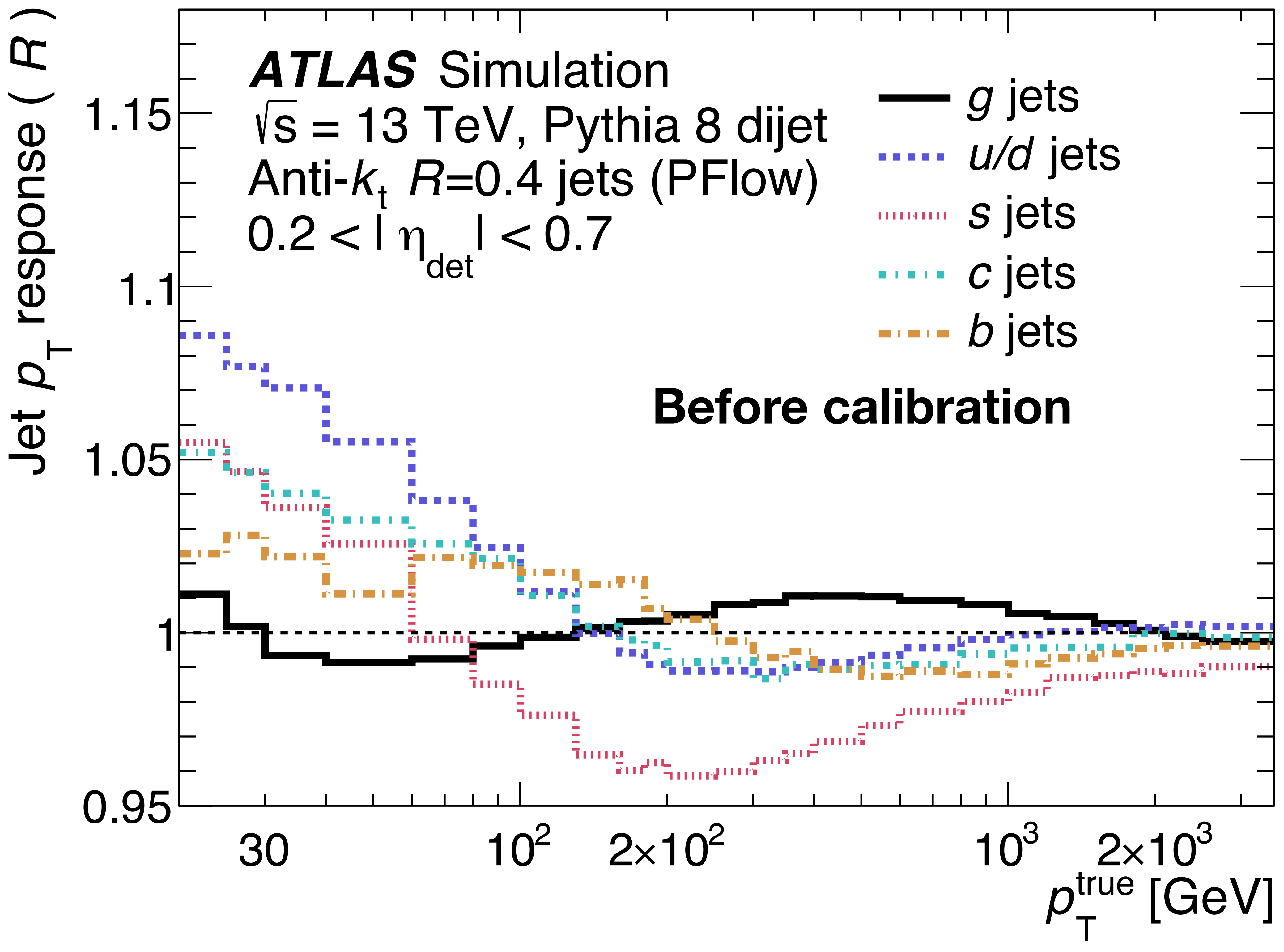


Global Calibration

- **The jet energy response depends on visible characteristics of jets. Such as different responses to quarks/gluons and different MC prediction of these characteristics of the jet.**

  - Two types of global calibration:
  - **Global Sequential Calibration (GSC):**
    - Multiplicative series of corrections relying on global jet measurements, based on six input observables.
  - **Global Neural Network Calibration (GNNC):**
    - Simultaneous correction using a DNN which accounts for input correlations and allows for more input observables.

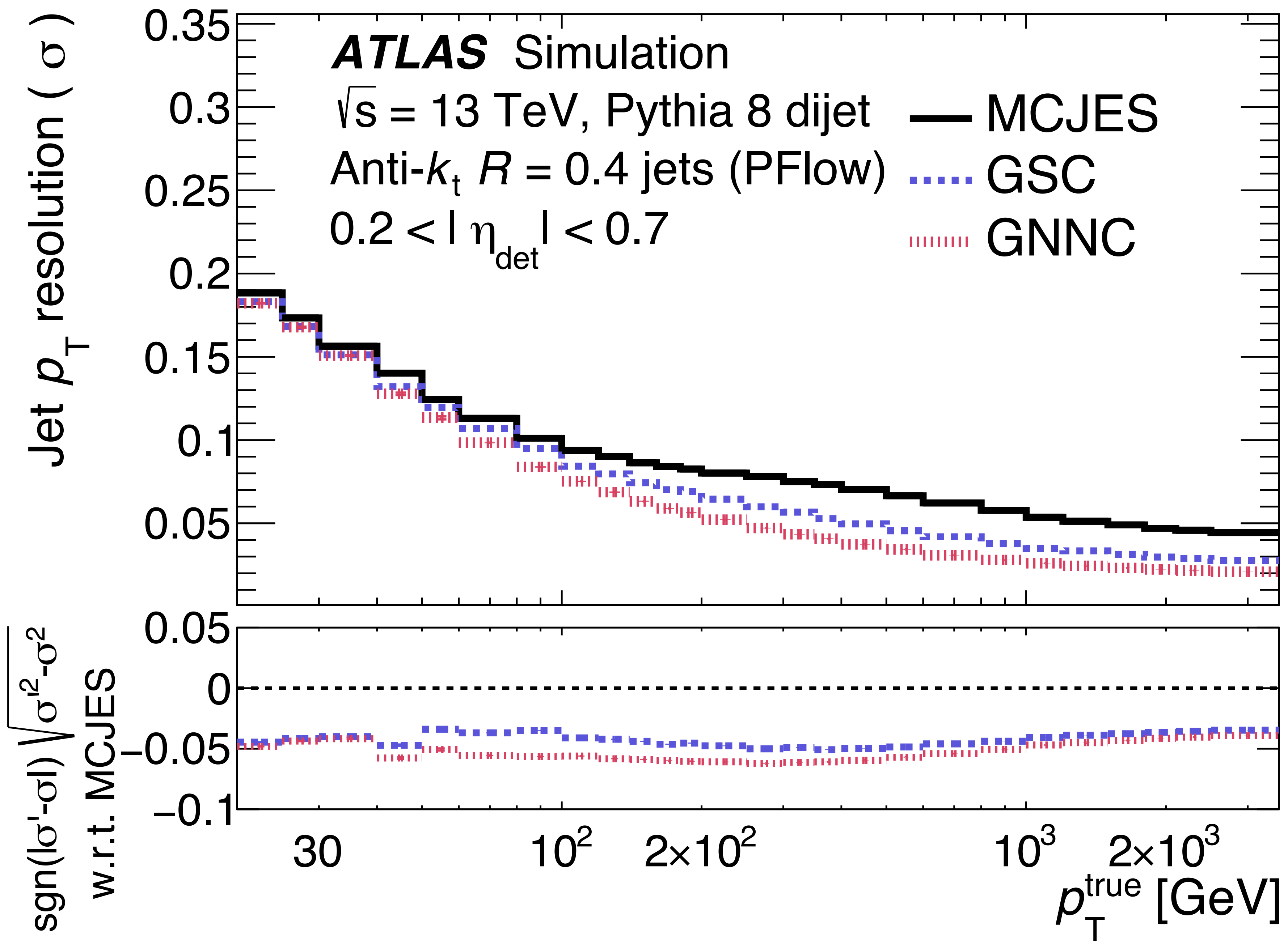
CERN-EP-2023-028



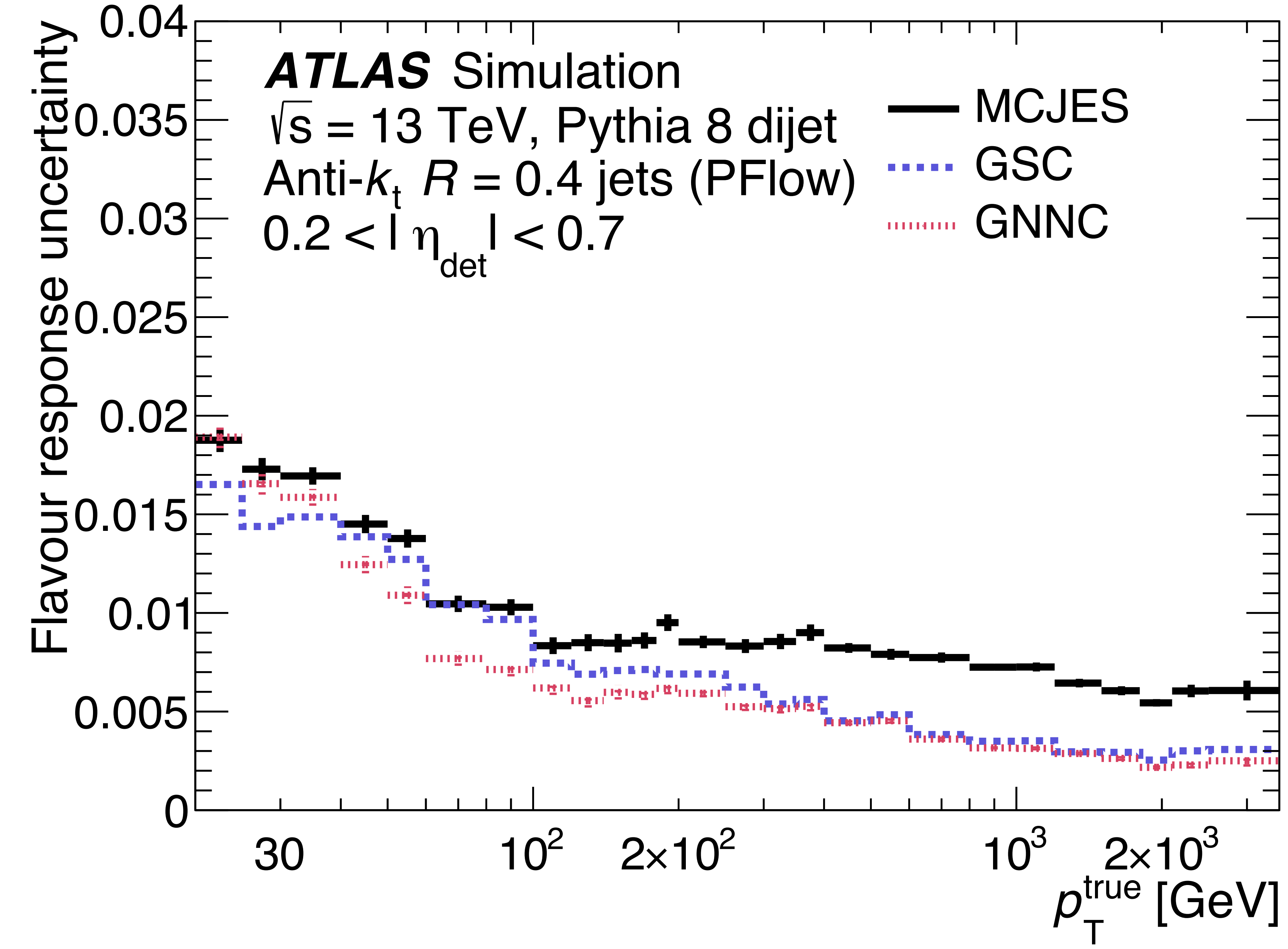
CERN-EP-2023-028

GSC vs GNNC

- **Jet Energy Resolution (JER) average improvement in the GNNC.**



- **The flavour response uncertainty** accounts for the fact that, gluon-initiated jet response is found to differ significantly between generators. **Improved with GNNC.**



- **The flavour composition uncertainty** accounts for the differing response of quark- and gluon-initiated jets. **Improved with GNNC.**

Ben Hodkinson is presenting other areas that ML/AI is being used to improve performance of jets and MET in ATLAS.

In-situ calibration

- **Residual correction to account for residual differences in jet response between MC and data.**

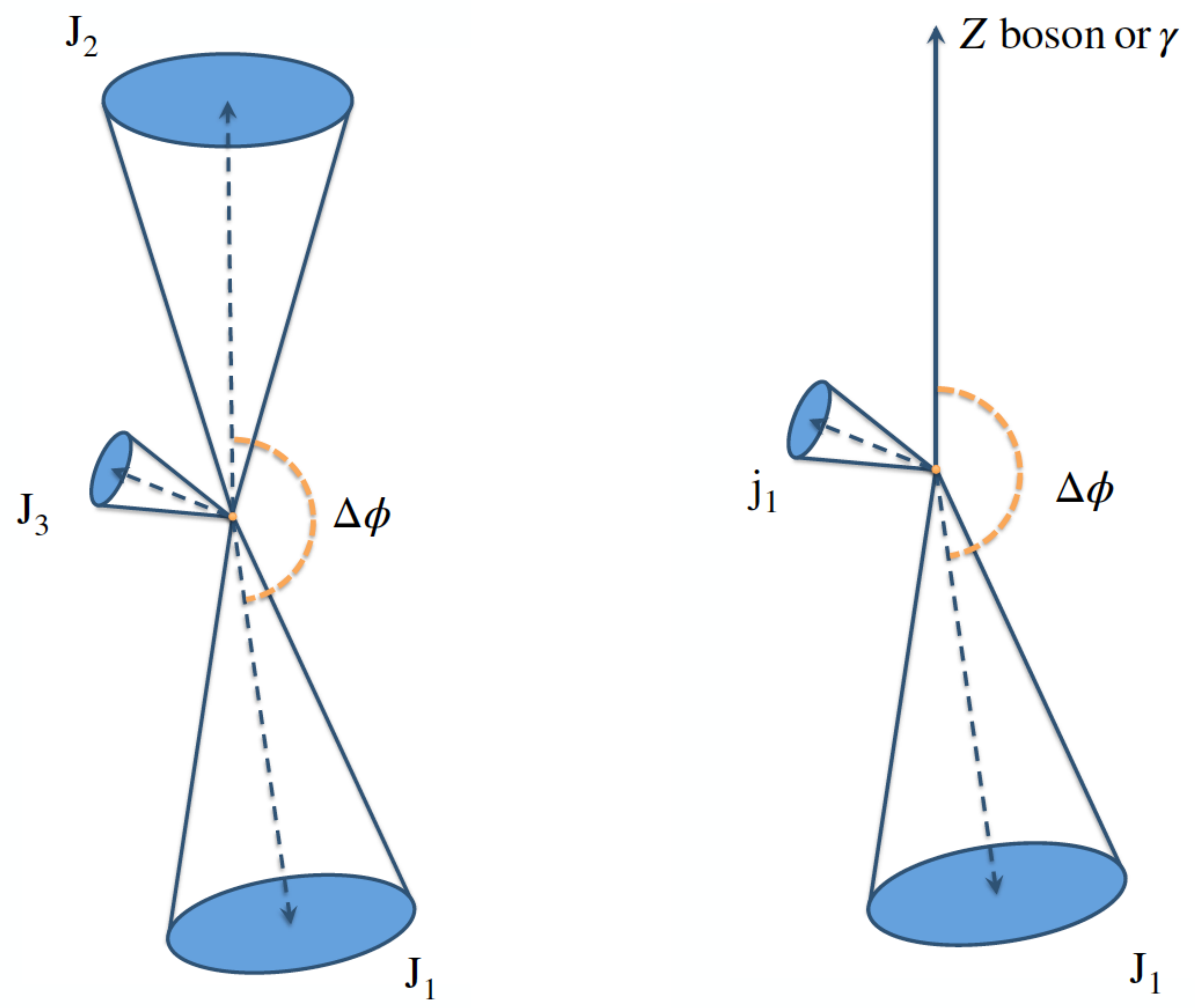
p_T-based balance is measured for a jet recoiling against a well-calibrated object.

The in-situ correction factor is given as $C = \frac{R^{data}_{in-situ}}{R^{MC}_{in-situ}}$ and the inverse is applied to data.

- **η-intercalibration uses di-jet events to correct the response of forward jets in the detector and brings them to the same energy scale of central jets (which are well-calibrated).**

- **bJES calibration uses γ+jet events with the "direct balance" method:**

$$R_{DB} = < \frac{p_T^{ref}}{p_T^{jet}} > , p_T^{ref} = p_T^{Object} \times cos|(\Delta\phi)|$$

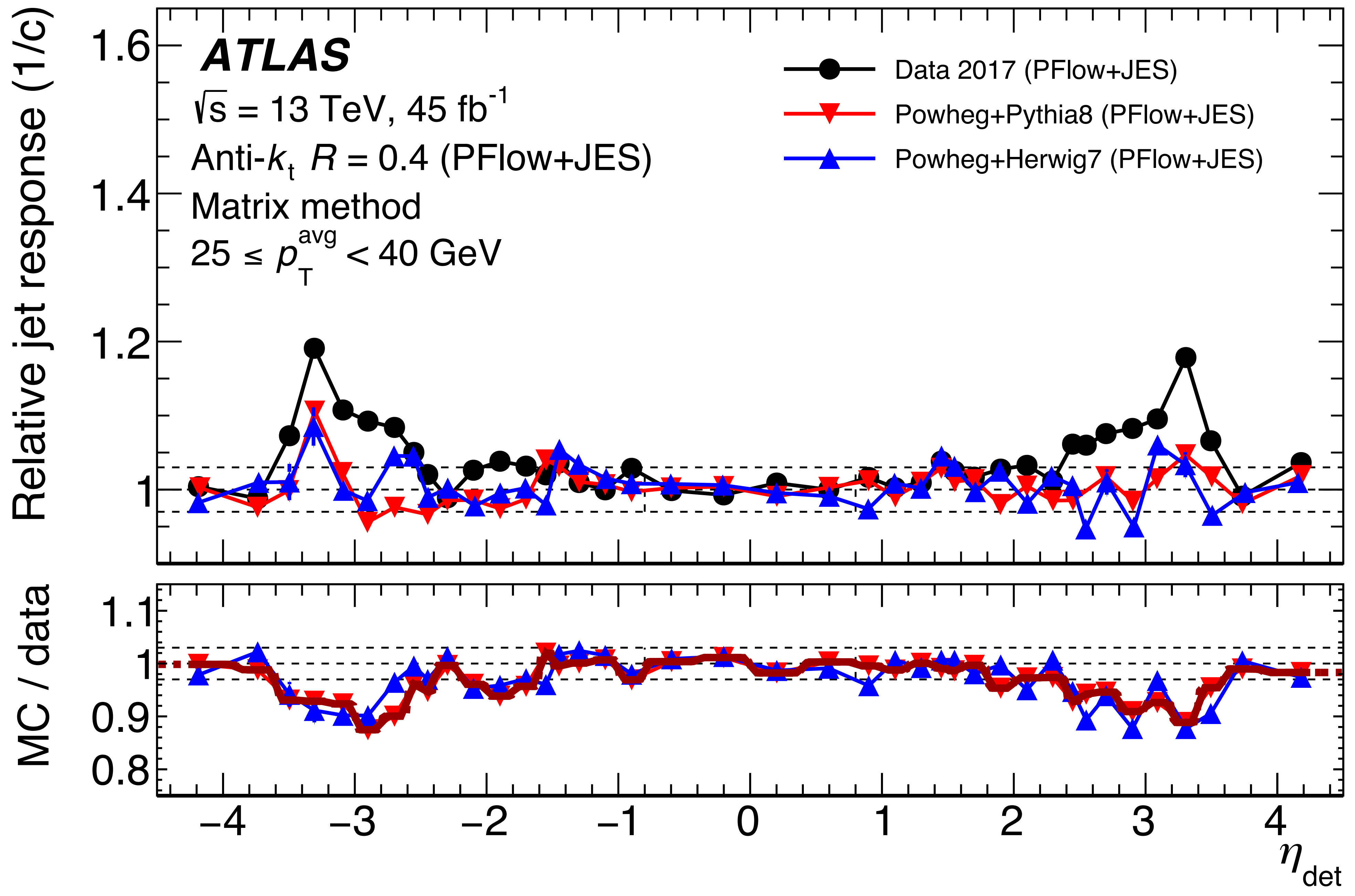


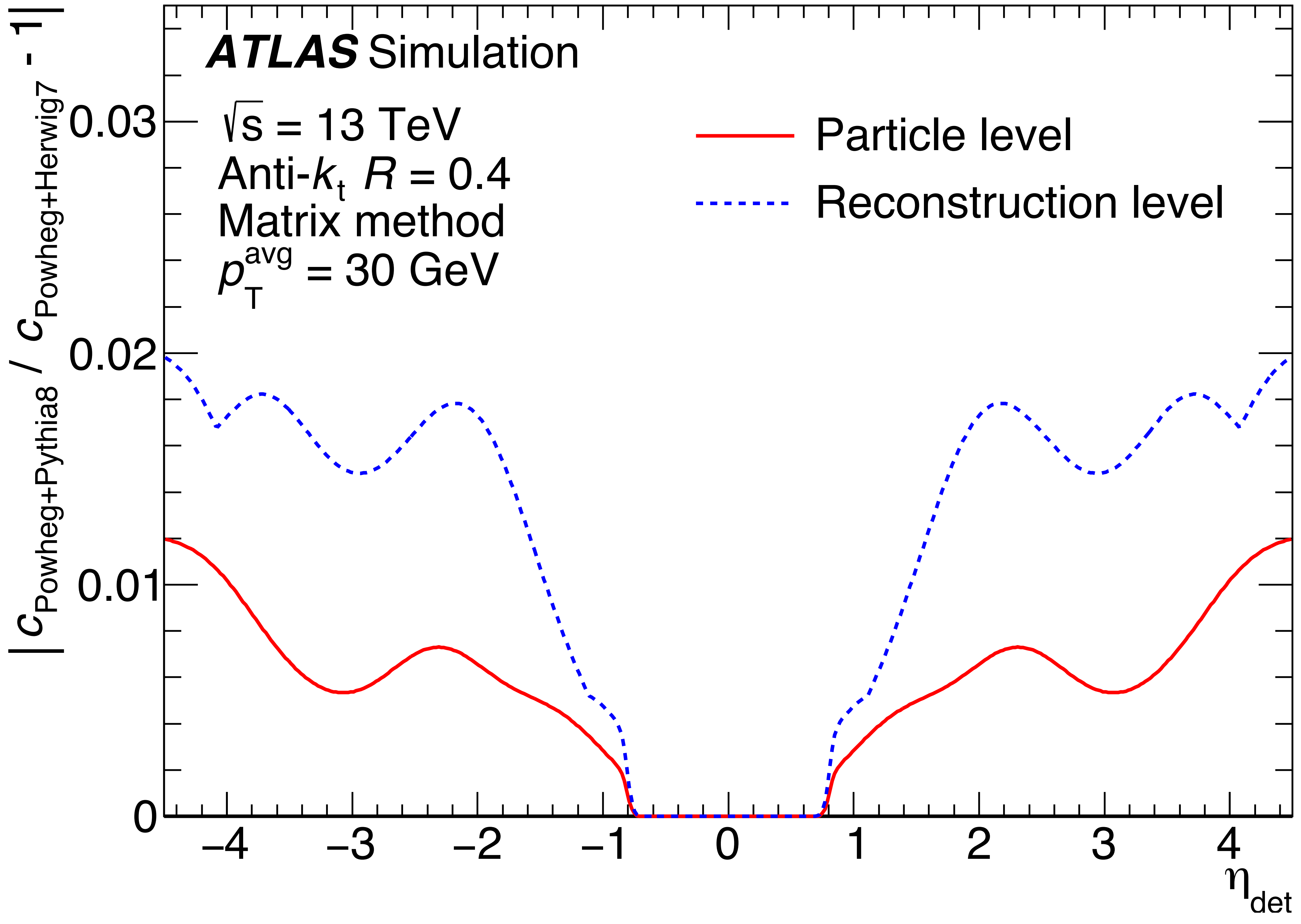
(a) dijet event (b) Z+jet or γ+jet event

Eur. Phys J. C 79. 135 (2019)

η-intercalibration

- **Making the jet response homogeneous across the whole η range of the detector.**
- The correction factor is derived in bins of η and p_T^{avg} (the average p_T of two jets in two distinct detector regions).





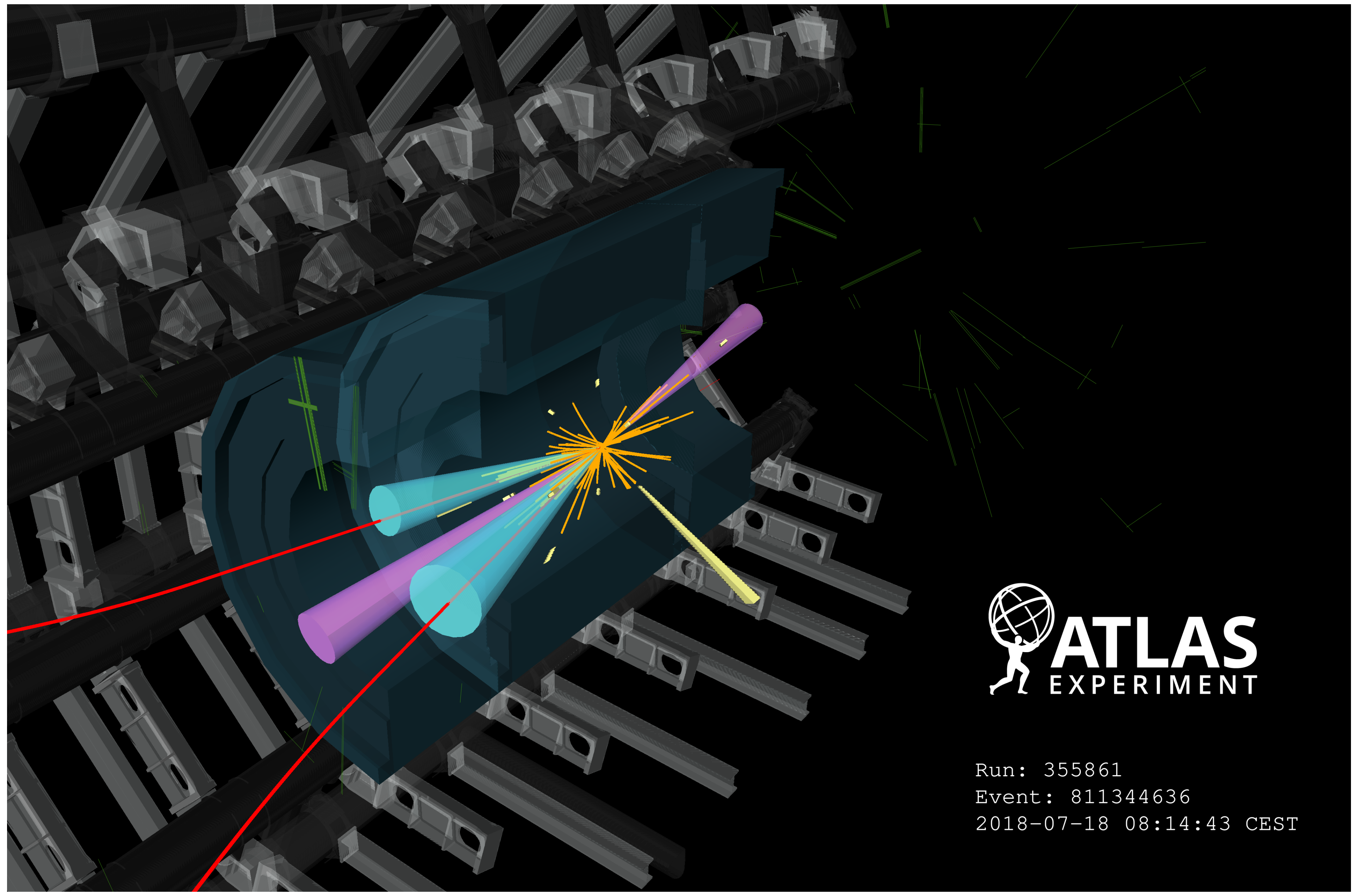
- **New study to separate detector- and particle-level effects for the MC modelling uncertainty:**
  - **Removes possible double-counting of detector effects with the flavour uncertainty.**
  - **MC modelling uncertainty reduced by factor of 2.**

CERN-EP-2023-028

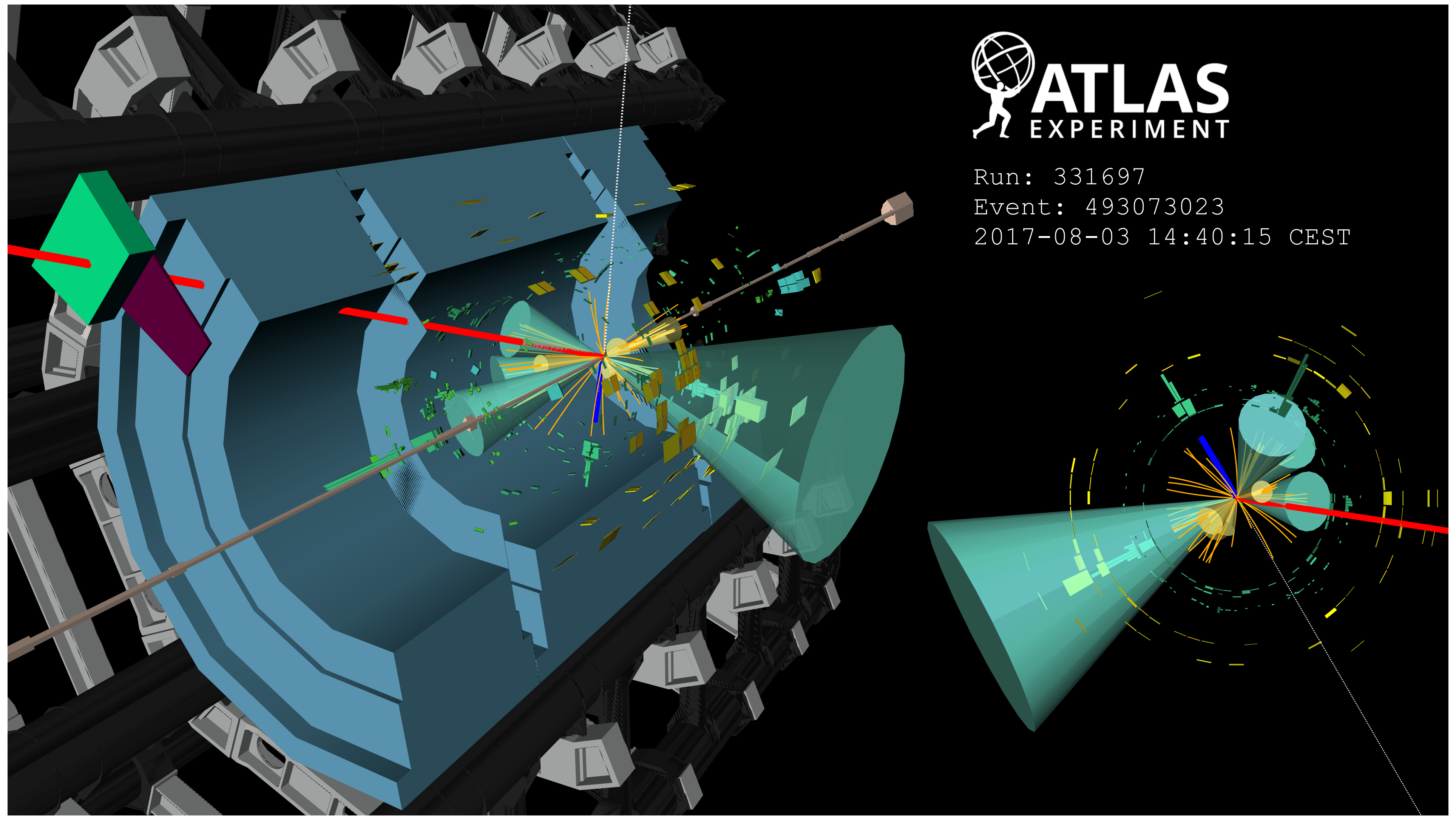
b-Jet Energy Scale

b-Jet Energy Scale (bJES) - the correction of b-jet energy applied after the previous steps to account for response difference due to the nature of b-jets.

- Jets originating from the fragmentation and hadronisation of **bottom quarks (b-jets) play an extremely important role** in many collision processes involved in the ATLAS detector (e.g. the main decay of the Higgs $H -> b\bar{b}$).



arXiv:2010.13651



arXiv:2211.01136

b-jets differ from light-quark and gluon jets:

This may effect the energy response of the jet so it is important to study the differences

bJES with γ+jet events

- Idea is to compare the MC-to-data ratio for the balance of b-tagged jets and inclusive jets in γ+jet events measure if the bJES differs from the JES

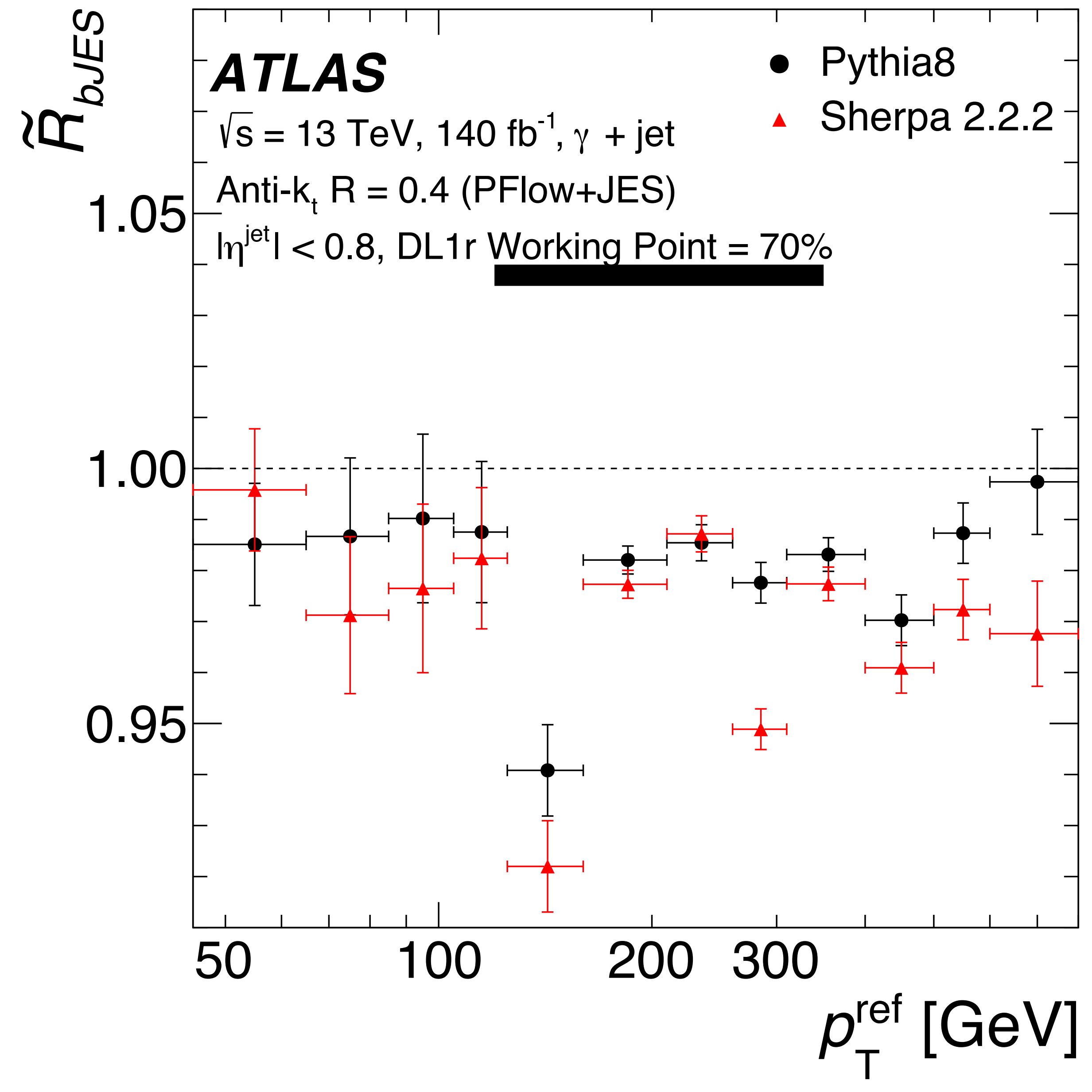
$$\tilde{R}_{bJES} = \frac{R^{MC}_{b-tagged} / R^{data}_{b-tagged}}{R^{MC}_{inclusive} / R^{data}_{inclusive}}$$

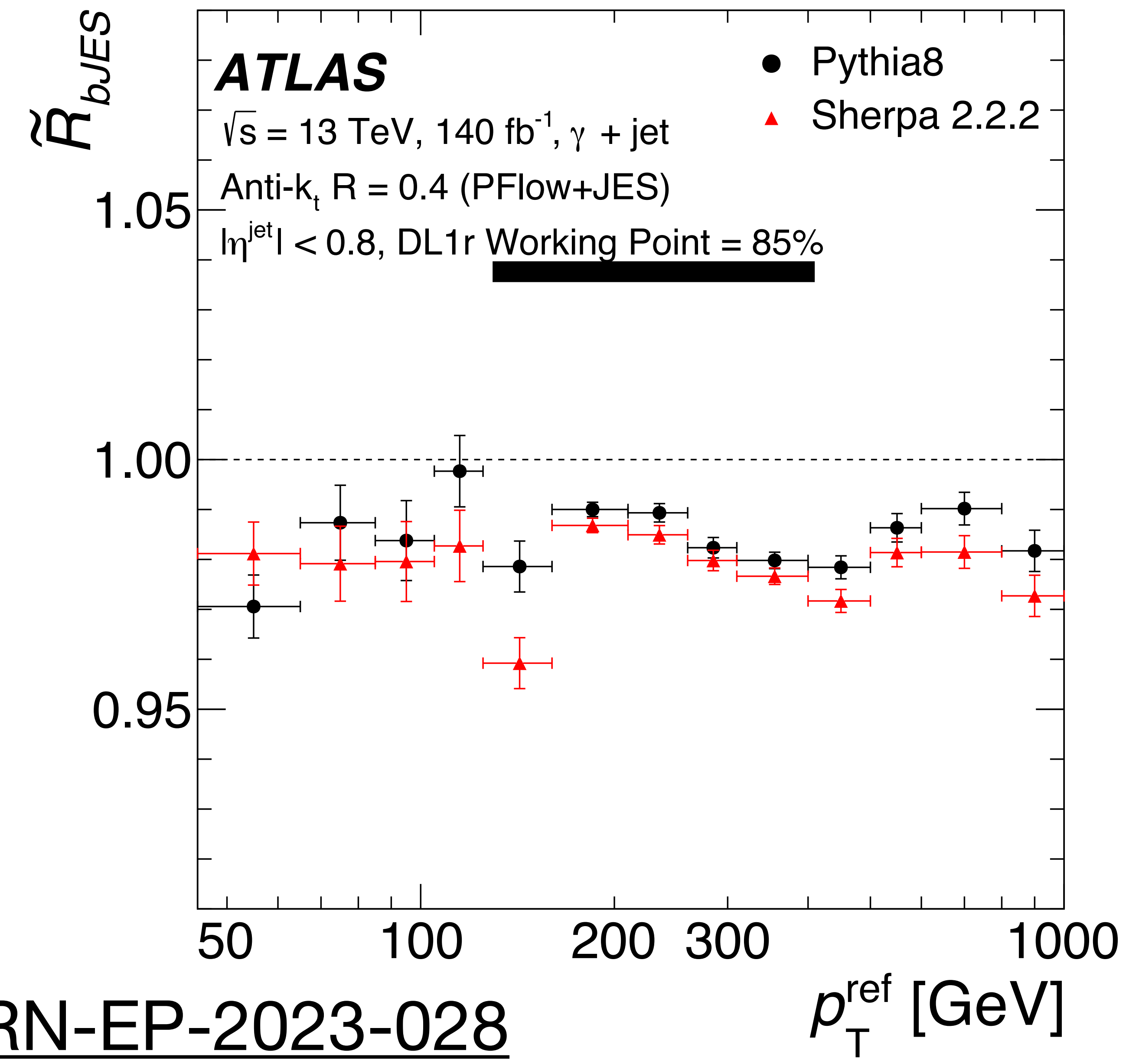
- ATLAS uses a multi-variate NN algorithm called **b-tagging**, which returns mostly b-jets from a samples containing all types of jets. [arXiv:2211.16345]
- B-tagging is applied in terms of efficiencies:
  - Higher efficiency -> Looser cut and lower background rejection (allowing more events through) -> Lower purity but higher statistics.

bJES with γ+jet events

$$\tilde{R}_{bJES} = \frac{R^{MC}_{b-tagged}/R^{data}_{b-tagged}}{R^{MC}_{inclusive}/R^{data}_{inclusive}}$$

- Energy scale determined for the b-jets with respect to inclusive jets for 4 working points: 60%, 70%, 77%, and 85% b-jet efficiencies.
- The energy scale for b-jets is underestimated by ~1% to ~3.5% depending on the efficiency and MC generator.
- In-situ measurements of the bJES will **improve precision of important analyses.**
  - Also trying to **improve b-fragmentation modelling in MCs** with dedicated measurements of b-fragmentation. [arXiv:2108.11650]





CERN-EP-2023-028

Results

- **Jets are used in almost all analyses in ATLAS and so extremely important to calibrate best as possible for more accurate results overall.**
- **JES has to keep up with advances in technology with higher luminosity. New techniques must be constantly developed to keep up with the higher statistics.**
- **Reduction of uncertainty of area-based pile-up correction by a factor of nearly 7.**
- **New 3D residual pile-up correction reduces pile-up dependence.**
- **Better closure in MCJES for new p-spline fitting method.**
- **Global calibration new GNNC technique improves p_T resolution, and flavour uncertainties.**
- **MC modelling uncertainty decreased by a factor of 2 in η-intercalibration.**
- **First time in-situ bJES performance in γ+jet events, finds a discrepancy in the MC-to-data ratio between b-jets and inclusive jets.**

BACKUP

Definitions

- N_{PV} - number of reconstructed primary vertices in an event.
- μ (in pile-up) - interactions per bunch crossing.
- z_0 - the distance of closest approach to the hard-scatter primary vertex along the z-axis.
- E_{reco} - energy of the reconstructed jet.
- E_{true} - the energy of the particle-level jet.
- η_{det} - the jet η which points from the geometric centre of the detector.

Z/γ+jet balance

- **The reference Z/γ object is balanced against the whole hadronic recoil in the event using the MPF technique.**

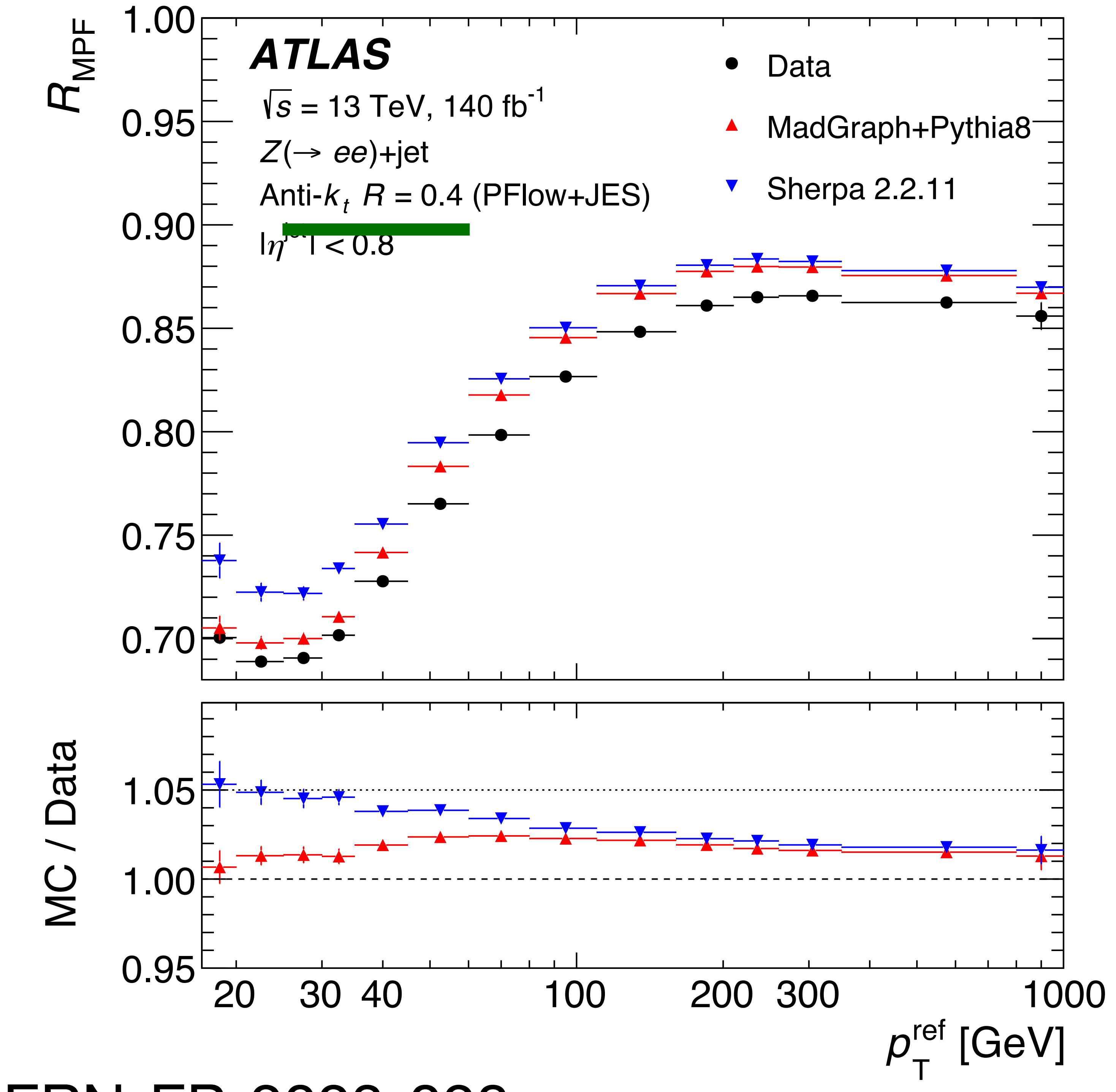
$$R_{MPF} = \left< 1 + \frac{\hat{n}_{ref} \, . \, \overrightarrow{E}_T^{miss}}{p_T^{ref}} \right>$$

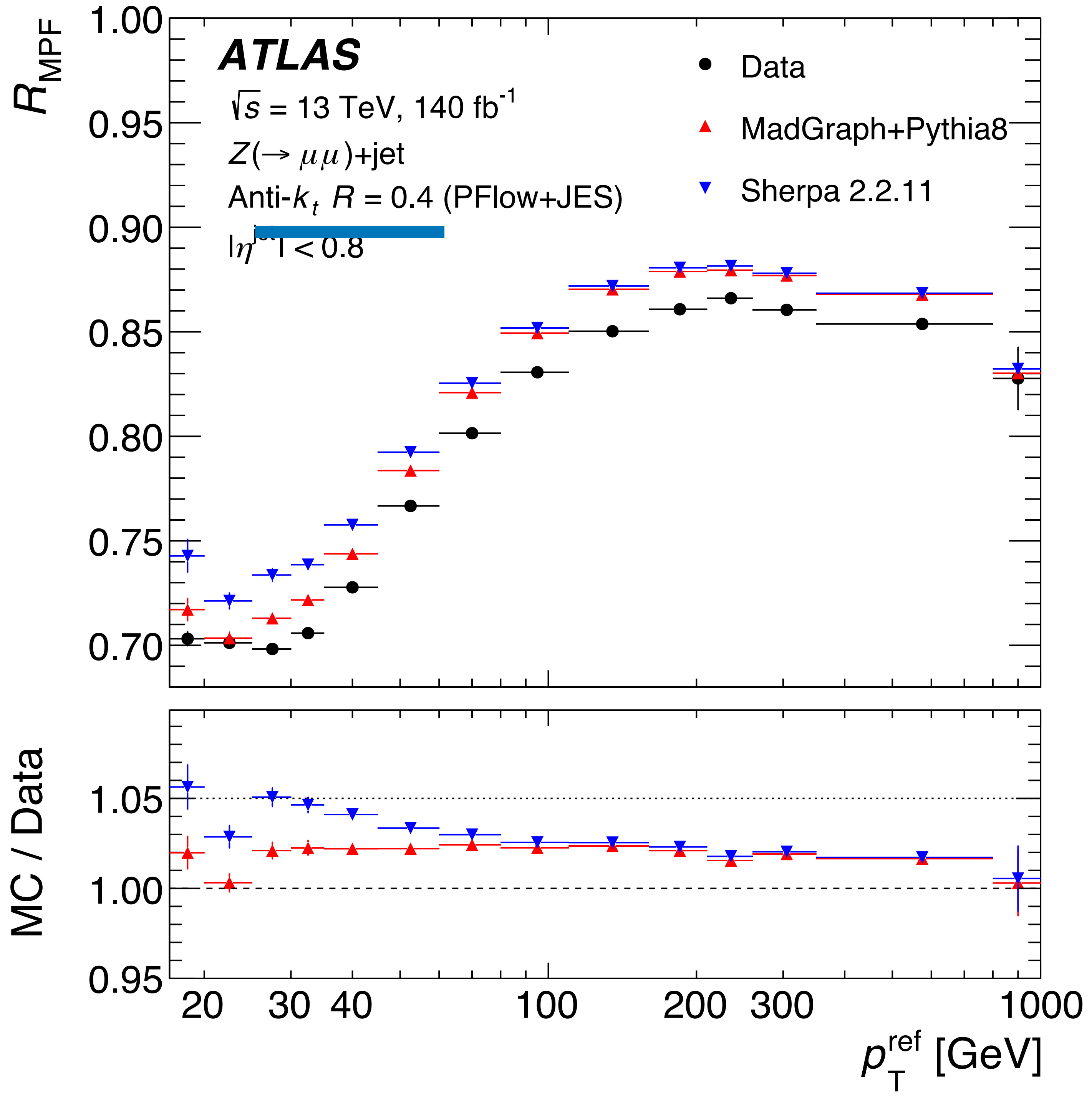
Z+jet(ee) JES is overestimated in MC by ~1-5%.

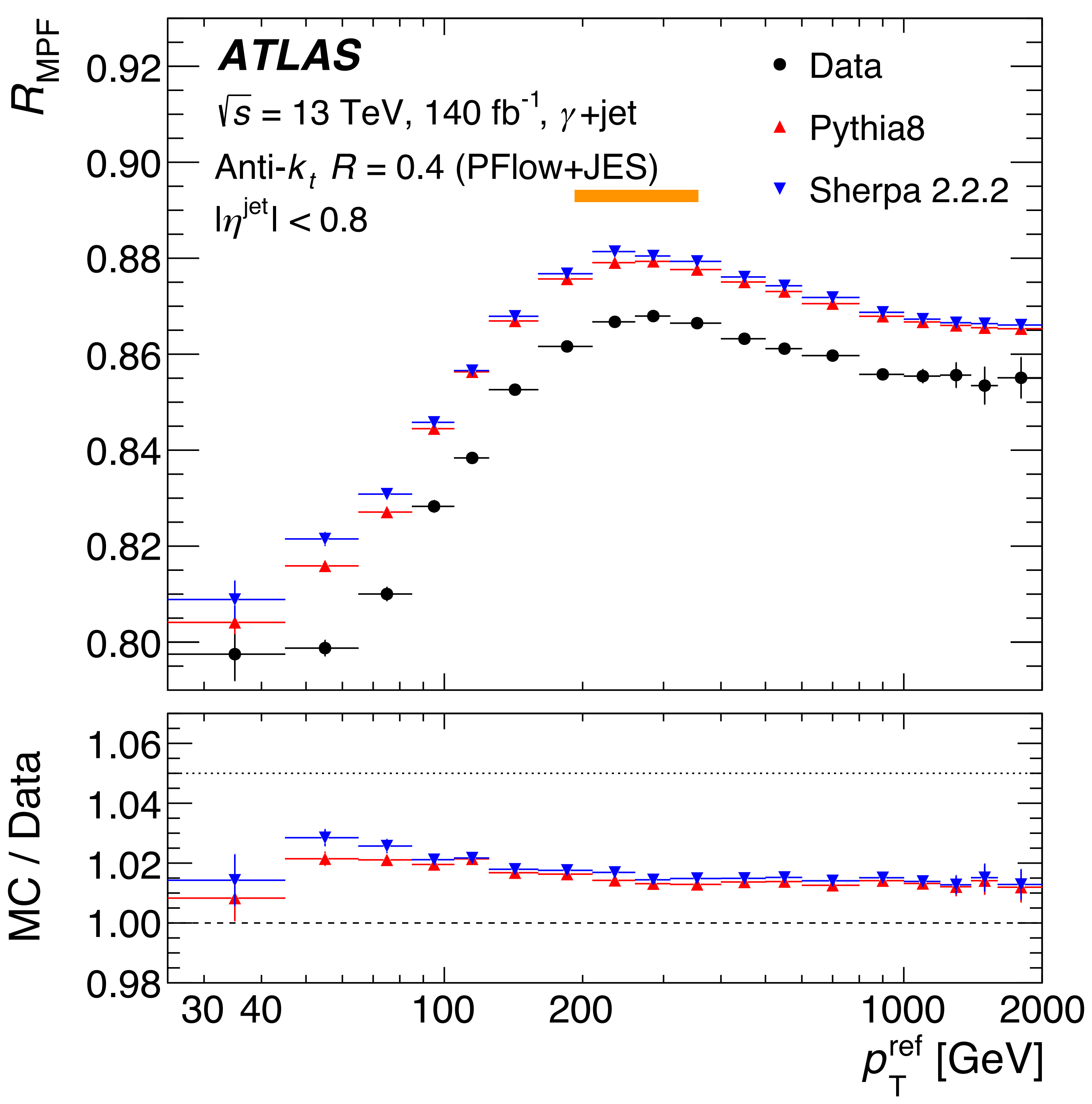
Z+jet(μμ) JES is overestimated in MC by ~1-5%.

γ+jet JES is overestimated in MC by ~1-3%.

Unprecedented precision up to 1% is achieved in the in-situ analysis.







CERN-EP-2023-028

Backup

Anti-kt Algorithm

- The anti-kt algorithm is a jet clustering algorithm. It lies in the definition of distance measures, d_{ij}, between objects i and j and between the object i and beam B, d_{iB}.
- The distance measures for the anti-kt algorithm are calculated as:

$$d_{ij} = min(k_{ti}^{-2}, k_{tj}^{-2})\frac{\Delta_{ij}^2}{R^2}$$

$$d_{iB} = k_{ti}^{-2}$$

Where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i and ϕ_i are the transverse momentum, rapidity, and azimuthal angle of particle i respectively.

- The clustering identifies the smallest of the distances and if it is d_{ij}, the objects i and j and combined, while if it is d_{iB}, i is called a jet and it is removed from the list of objects.
- This is repeated until there are no objects left.

JHEP 0804:063 (2008)

Backup

Particle Flow Jets

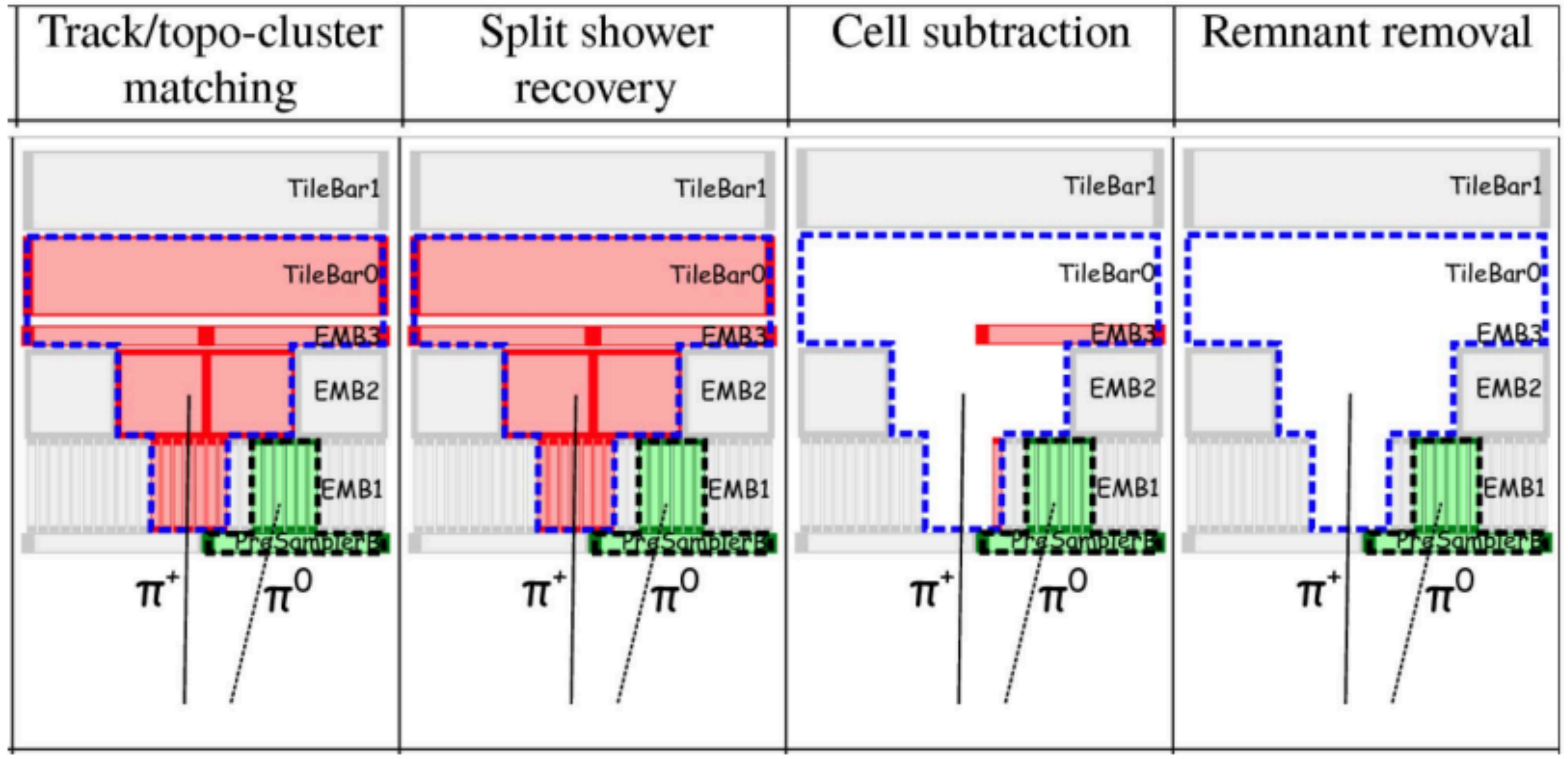
Tracker allows for association to the primary vertex

- Particle-Flow (PFlow) jets are reconstructed by combining track- and calorimeter-based measurements.

Uses the expected energy depositions of single particles to determine contributions of individual tracks to clusters.

Cell-based subtraction prevents double-counting of energy.

- This all leads to improved jet energy, mass resolution and pile-up stability.



Eur. Phys. J. C 77. 466 (2017)

GSC Inputs

The six stages of the GSC, in the order of application, are

- f_{charged}: the fraction of the jet p_T measured from ghost-associated tracks with $p_T > 500$ MeV, $|\eta^{\text{det}}| < 2.5$,
- f_{Tile0}: the fraction of jet energy (E_{frac}) measured in the first layer of the hadronic tile calorimeter, $|\eta^{\text{det}}| < 1.8$,
- f_{LAr3}: the E_{frac} measured in the third layer of the electromagnetic LAr calorimeter, $|\eta^{\text{det}}| < 3.5$,
- N_{track}: the number of tracks with $p_T > 1$ GeV ghost-associated with the jet, $|\eta^{\text{det}}| < 2.5$,
- w_{track}: also known as track width, the average p_T-weighted transverse distance in the η-ϕ plane, between the jet axis and all tracks of $p_T > 1$ GeV ghost-associated with the jet, $|\eta^{\text{det}}| < 2.5$,
- N_{segments}: the number of muon track segments ghost-associated with the jet, $|\eta^{\text{det}}| < 2.8$.

Eur. Phys. J. C 81. 689 (2021)

GNNC Inputs

Calorimeter	$f_{\mathrm{LAr}0-3*}$	The E_{frac} measured in the 0th-3rd layer of the EM LAr calorimeter
	$f_{\mathrm{Tile}0*-2}$	The E_{frac} measured in the 0th-2nd layer of the hadronic tile calorimeter
	$f_{\mathrm{HEC},0-3}$	The E_{frac} measured in the 0th-3rd layer of the hadronic end cap calorimeter
	$f_{\mathrm{FCAL},0-2}$	The E_{frac} measured in the 0th-2nd layer of the forward calorimeter
	$N_{90\%}$	The minimum number of clusters containing 90% of the jet energy.
Jet kinematics	$p_{\mathrm{T}}^{\mathrm{JES}}$ *	The jet p_{T} after the MCJES calibration
	η^{det}	The detector η
Tracking	w_{track}*	The average p_{T}-weighted transverse distance in the η-ϕ plane between the jet axis and all tracks of $p_{\mathrm{T}} > 1$ GeV ghost-associated with the jet
	N_{track}*	The number of tracks with $p_{\mathrm{T}} > 1$ GeV ghost-associated with the jet
	f_{charged}*	The fraction of the jet p_{T} measured from ghost-associated tracks
Muon segments	N_{segments}*	The number of muon track segments ghost-associated with the jet
Pileup	μ	The average number of interactions per bunch crossing
	N_{PV}	The number of reconstructed primary vertices

Table 1: List of variables used as input to the GNNC. Variables with a * correspond to those that are also used by the GSC.

Link to new paper if out by the time I give the talk.