JET/MET RECONSTRUCTION AND PERFORMANCE IN ATLAS

DILIA MARÍA PORTILLO

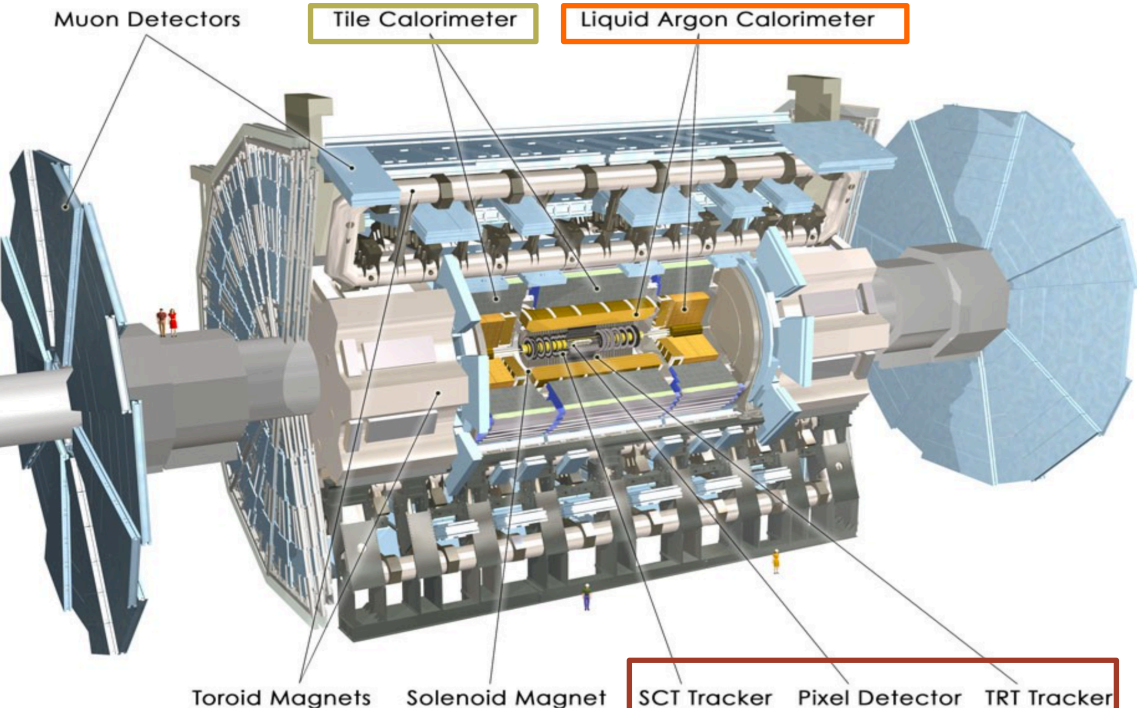
CMS JetMet Workshop

OUTLINE

- Jet inputs/constituents
  - Topological calorimeter cell clustering
  - Particle Flow Objectes
  - Track-Calo-Clusters
  - Unified Flow Objects
- Pile-up Mitigation and grooming techniques
  - Constituent Subtraction
  - Soft Killer
  - Jet Grooming
- Large-R jet optimisation
- Performance Comparison
- UFO taggers
  - Boosted W/Z tagging
  - Boosted top tagging
- Jet Energy Calibration
  - Simulation-based Calibration
  - *In situ* Calibration
- Missing transverse Momentum

ATLAS BASICS

Detectors that are interestina for iet reconstruction:



Inner Detector

- Charged particles tracks
- Decay vertices
- $|\eta| < 2.5$

Electromagnetic Calorimeter

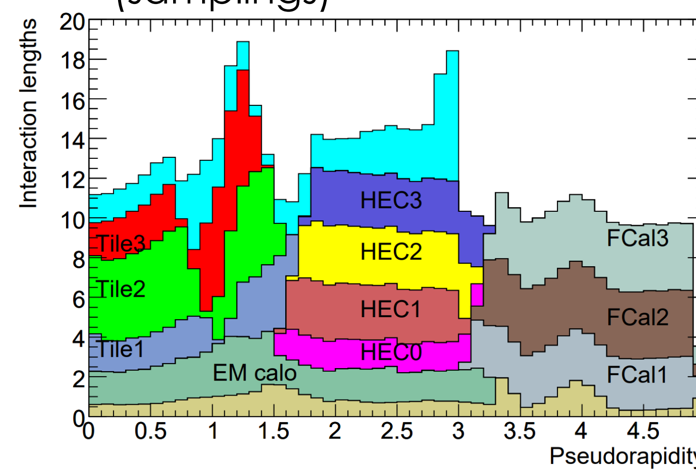
- Electron and photon energy/direction (EM showers)

Hadronic Calorimeter

- Charged and neutral hadrons energy/ direction (HAD showers)

Calorimeters

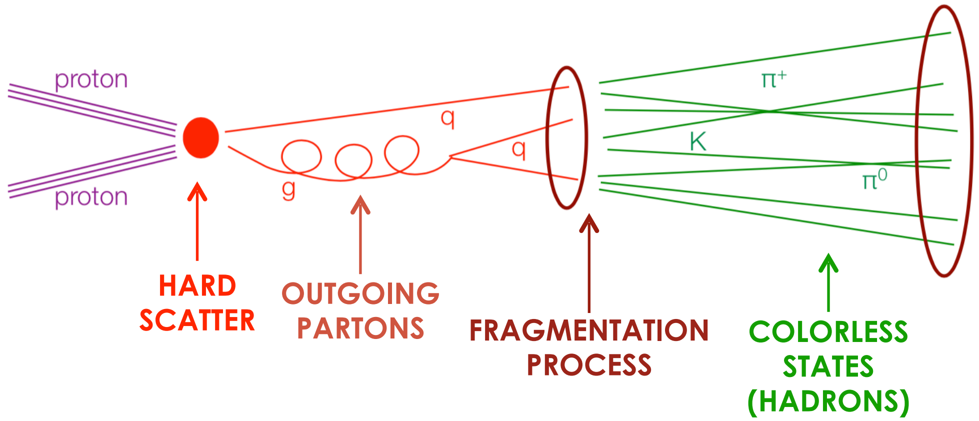
- Full coverage $|\eta| < 4.9$
- High granularity in $\Delta\eta \times \Delta\varphi = 0.025 \times \pi/128$ (central EM)
- Up to seven depth layers (samplings)



- Each read-out unit of the calorimeter defines a cell
- Contain energy/location information
- Each shower deposits energy in many cells

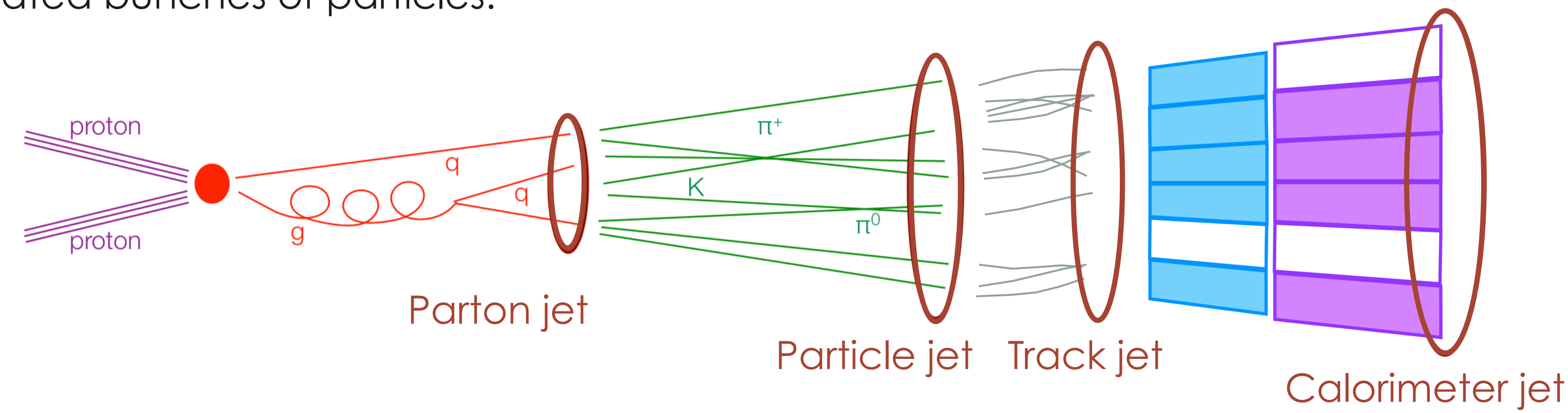
WHAT IS A JET

Due to color confinement, quarks and gluons shower and hadronise immediately into collimated bunches of particles.



WHAT IS A JET

Due to color confinement, quarks and gluons shower and hadronise immediately into collimated bunches of particles.

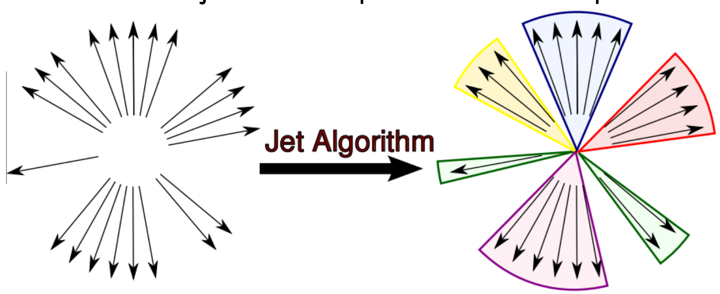


Jets can be formed of any 4-vector (simulated particles: truth jets, ID tracks: Track jets)

- Calorimeter jet is the baseline: calorimeter hits or groups of calorimeter hits

Use a jet algorithm to cluster objects into a jet: It maps final state particle momenta to jet momenta.

1 set of 4-vectors (constituents)



several subsets (jets)

- The Anti-Kt is the most widely used jet algorithm
  - R parameter: 0.4, 1.0

JET CONSTITUENTS

TOPO-CLUSTERS

Topo-clusters: 3D clusters of noise-suppressed calorimeter cells

[Eur. Phys. J. C 77 (2017) 490]

- Calorimeter jet **constituents**
- Baseline and most common **inputs to jet algorithm**

To form a topo-cluster: Use a recursive algorithm to combine cells with related energy deposits

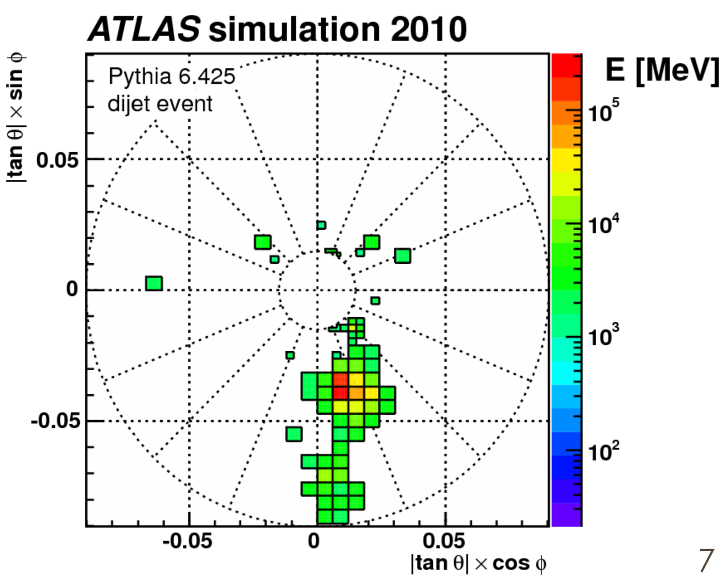
- Define for each cell: **significance**
  Ratio of energy measured to expected average energy due to noise in that cell

$$\zeta_{cell}^{EM} = \frac{E_{cell}^{EM}}{\sigma_{noise,cell}^{EM}}$$

Clustering algorithm

- Clusters are **seeded** by cells with large energy over noise ratio
  - $|\zeta|>4$

Seed cells



TOPO-CLUSTERS

Eur. Phys. J. C 77 (2017) 490

Topo-clusters: 3D clusters of noise-suppressed calorimeter cells

- Calorimeter jet **constituents**
- Baseline and most common **inputs to jet algorithm.**

To form a topo-cluster: Use a recursive algorithm to combine cells with related energy deposits

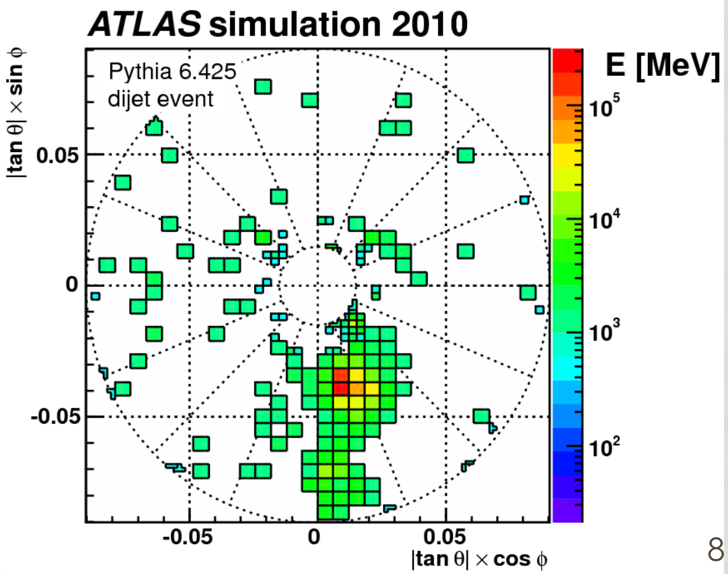
- Define for each cell: **significance**
  Ratio of energy measured to expected average energy due to noise in that cell

$$\zeta^{EM}_{cell} = \frac{E^{EM}_{cell}}{\sigma^{EM}_{noise,cell}}$$

Clustering algorithm

- Clusters are **seeded** by cells with large energy over noise ratio
  - $|\zeta|>4$
- Expanded on neighbouring cells
  - All **Neighbors** with $|\zeta|>2$ are added

Growth cells



TOPO-CLUSTERS

Eur. Phys. J. C 77 (2017) 490

Topo-clusters: 3D clusters of noise-suppressed calorimeter cells

- Calorimeter jet **constituents**
- Baseline and most common **inputs to jet algorithm.**

To form a topo-cluster: Use a recursive algorithm to combine cells with related energy deposits

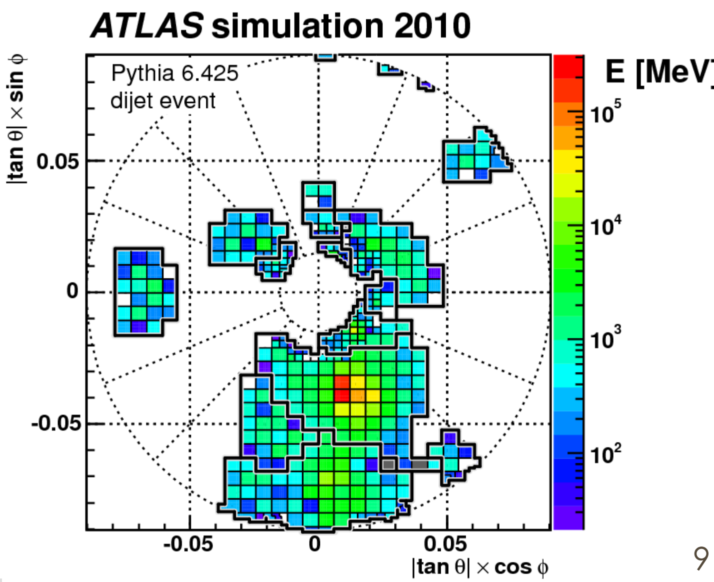
- Define for each cell: **significance**
  Ratio of energy measured to expected average energy due to noise in that cell

$$\zeta_{cell}^{EM} = \frac{E_{cell}^{EM}}{\sigma_{noise,cell}^{EM}}$$

Clustering algorithm

- Clusters are **seeded** by cells with large energy over noise ratio
  - $|\zeta|>4$
- Expanded on neighbouring cells
  - All **Neighbors** with $|\zeta|>2$ are added
- **All neighbouring** cells are added regardless of the significance
  - $|\zeta|>0$
- Final cluster splitting step breaks up large topo-clusters with multiple local maxima

Boundary cells

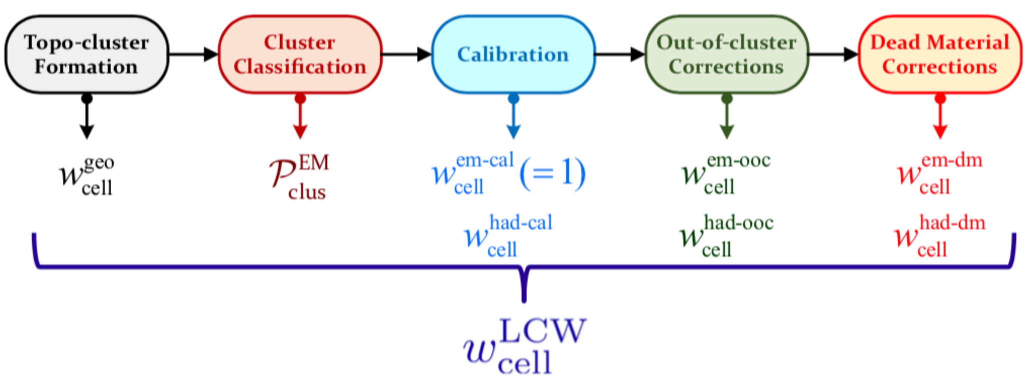


EM AND LCW SCALES

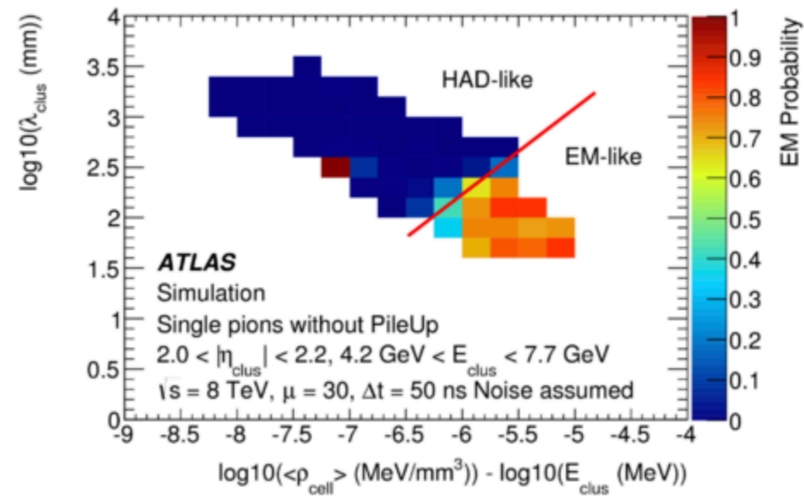
Topo-clusters enter jet finding at one of two **scales**:

CERN-PH-EP-2011-191

- **Electromagnetic** (EM) scale: same scale as the cells.
- **Local cell weighted** (LCW) scale: Topo-clusters calibrated based on their properties. Used for large-R jets.
  - Topo-clusters are identified as either electromagnetic or hadronic by the EM probability $P^{\mathrm{EM}}_{\mathrm{clus}}$
  - Weights are then assigned to account for
    - Differences in detector response due to calorimeter non-compensation
    - Energy falling in unclustered cells
    - Energy deposited in inactive (dead) regions of the detector



$$E^{\mathrm{LCW}}_{\mathrm{clus}} = \sum_{i \in \mathrm{cluster}} w^{\mathrm{LCW}}_{\mathrm{cell},i} \, E^{\mathrm{EM}}_{\mathrm{cell},i}$$

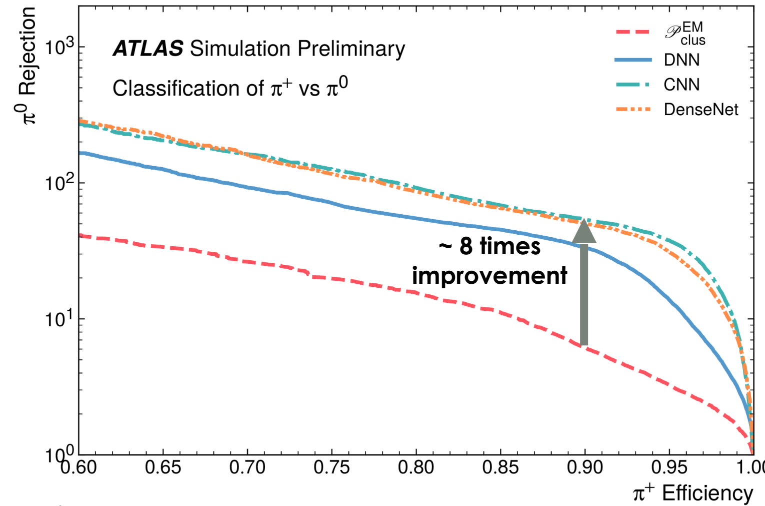


DEEP LEARNING FOR HADRONIC CALIBRATION

ATL-PHYS-PUB-2020-018

Neural Networks trained on calorimeter images can classify clusters and predict their energies using single particle simulation

- Studied DNNs, CNNs, and DenseNet

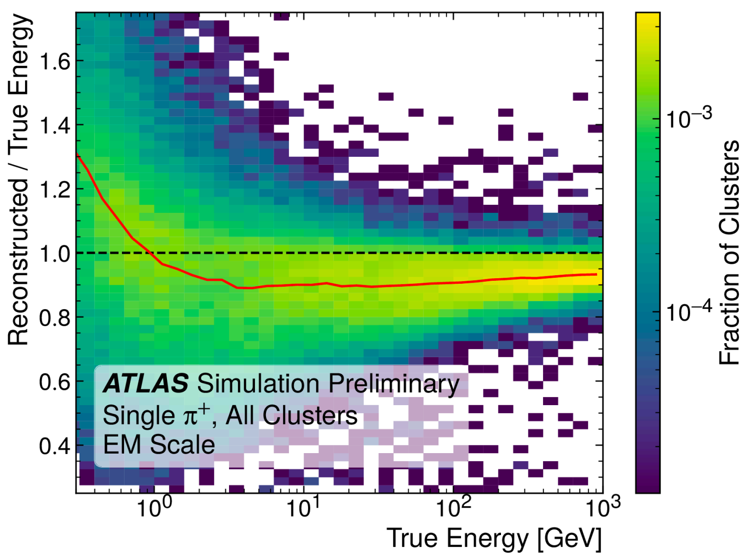


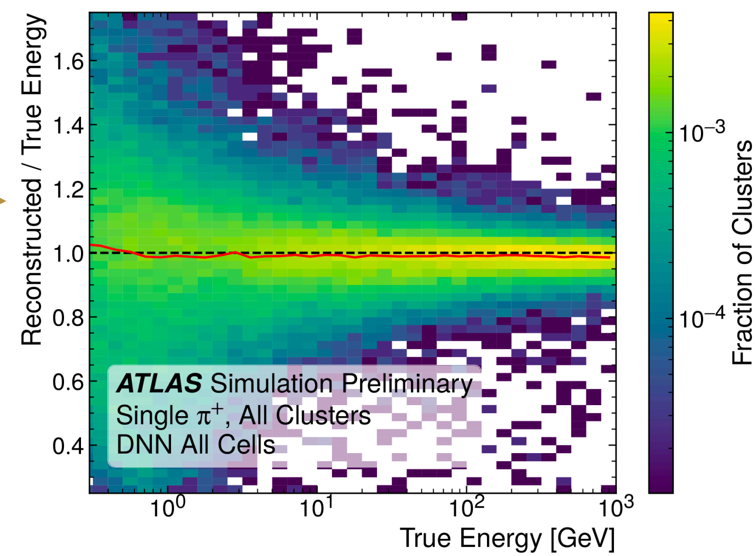
EM vs HAD showers Classification

- The ML techniques all do an excellent job of distinguishing π^0 from $\pi^\pm$ showers

Energy Regression: Predict the true energy deposited in the cluster.

- Energy response ~ 1 after calibration





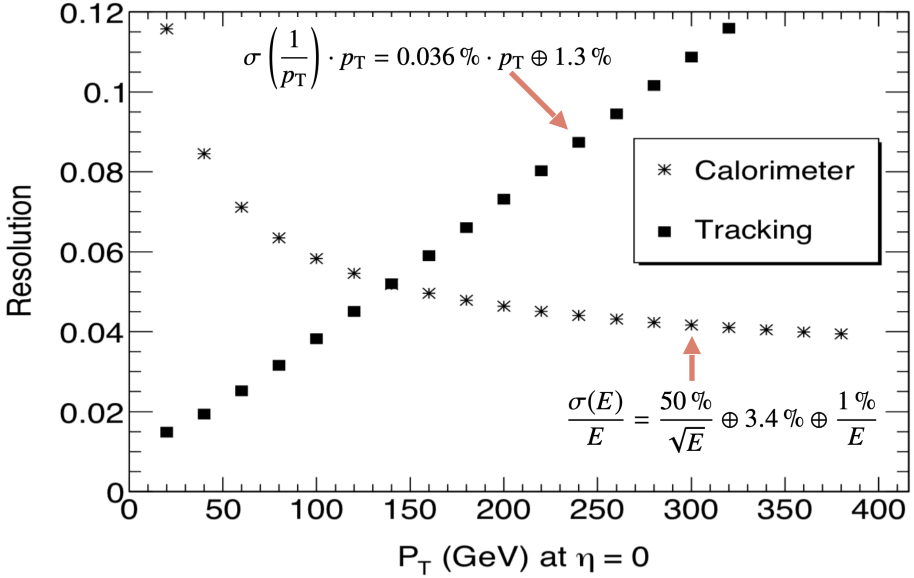
Before Calibration **After calibration**

COMBINING CLUSTERS AND TRACKS

Calorimeter and tracker provide complementary information

Tracker:
- Sensitive to charged particles
- Better angular resolution
- Able to assign tracks to pile-up or hard-scatter vertex.
- Better reconstruction efficiency and momentum resolution at low pT

Calorimeters:
- Sensitive to both neutral and charged particles.
- Better energy resolution at high pT



Alternative jet inputs

Alternative approaches to topo-clusters using measurements from the tracker and calorimeters

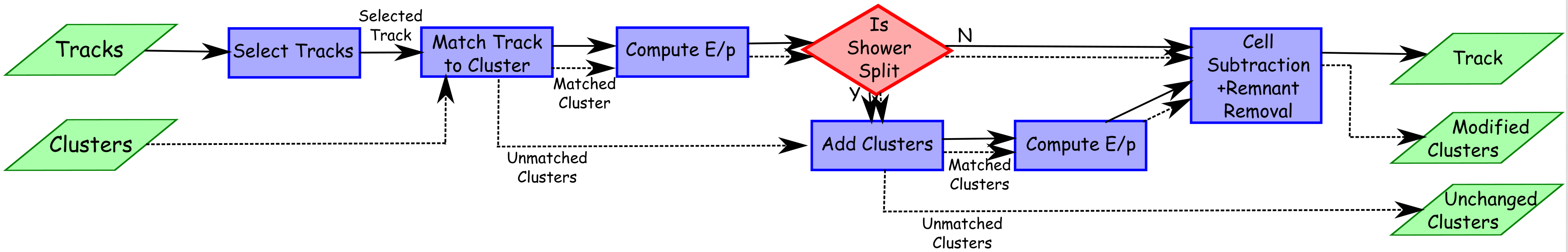
- Particle Flow (PFlow)
- Track Calo Clusters (TCC)
- Unified Flow Objects (UFO)

To be used as jet inputs

PARTICLE FLOW IN ATLAS (PFLOW)

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

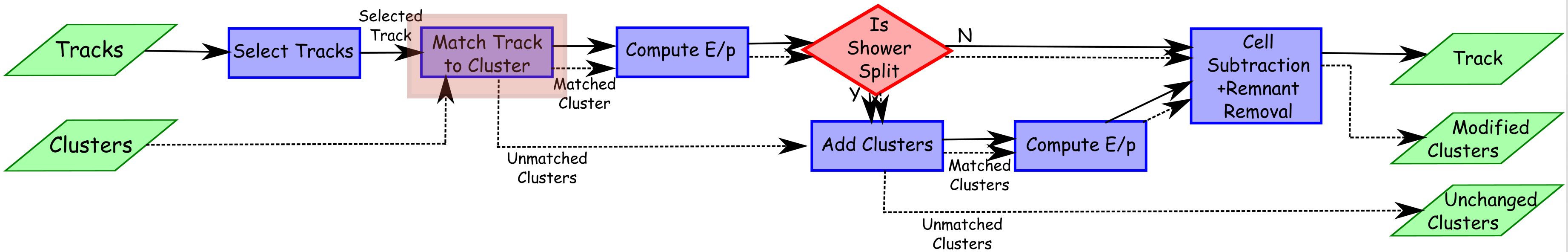
- Measure the energy, without double counting.
- Subtract calorimeter energy deposits that match an extrapolated track.
- Pile-up tracks are removed at the end: Charged Hadron Substraction (CHS)
- Final set of constituents is a combination of tracks and remaining clusters



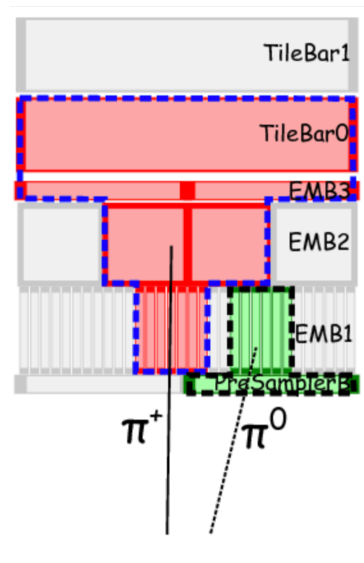
PARTICLE FLOW IN ATLAS (PFLOW)

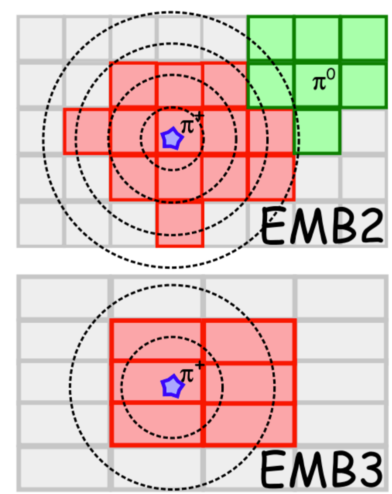
Eur. Phys. J. C 77 (2017) 466

- Measure the energy, without double counting.
- Subtract calorimeter energy deposits that match an extrapolated track.
- Pile-up tracks are removed at the end: Charged Hadron Substraction (CHS)
- Final set of constituents is a combination of tracks and remaining clusters



- Each track is associated with one or more topo-clusters

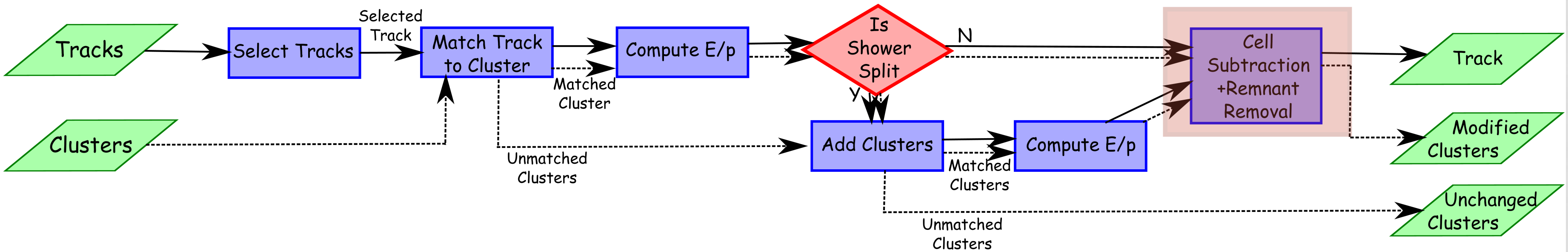




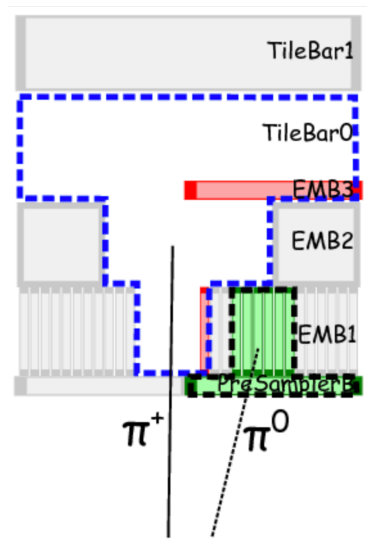
PARTICLE FLOW IN ATLAS (PFLOW)

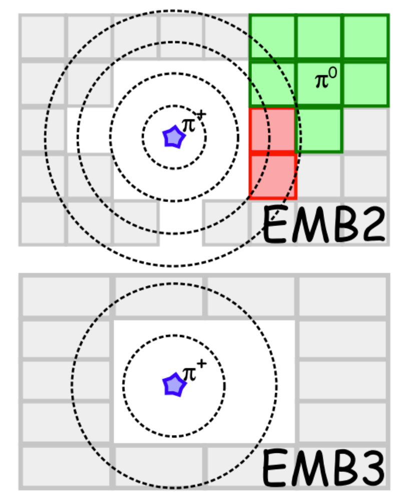
Eur. Phys. J. C 77 (2017) 466

- Measure the energy, without double counting.
- Subtract calorimeter energy deposits that match an extrapolated track.
- Pile-up tracks are removed at the end: Charged Hadron Substraction (CHS)
- Final set of constituents is a combination of tracks and remaining clusters



- Each track is associated with one or more topo-clusters
- Expected calorimeter energy deposit from charged hadron is removed

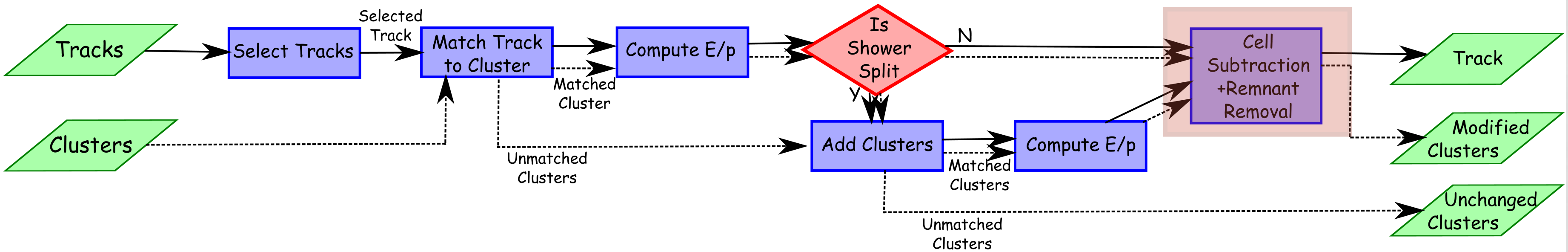




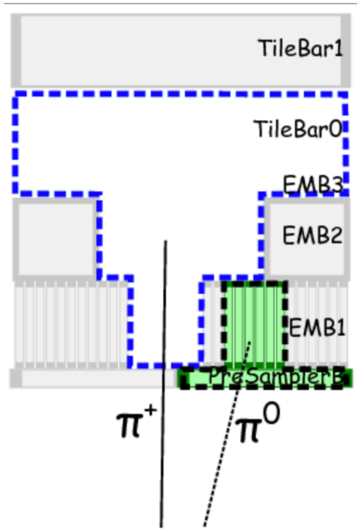
PARTICLE FLOW IN ATLAS (PFLOW)

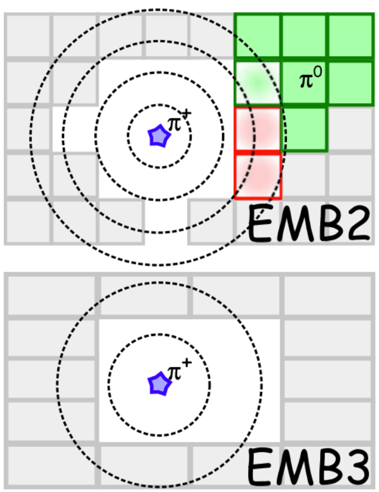
[Eur. Phys. J. C 77 (2017) 466]

- Measure the energy, without double counting.
- Subtract calorimeter energy deposits that match an extrapolated track.
  - Pile-up tracks are removed at the end: Charged Hadron Substraction (CHS)
  - Final set of constituents is a combination of tracks and remaining clusters



- Each track is associated with one or more topo-clusters
- Expected calorimeter energy deposit from charged hadron is removed
- Remnants are removed



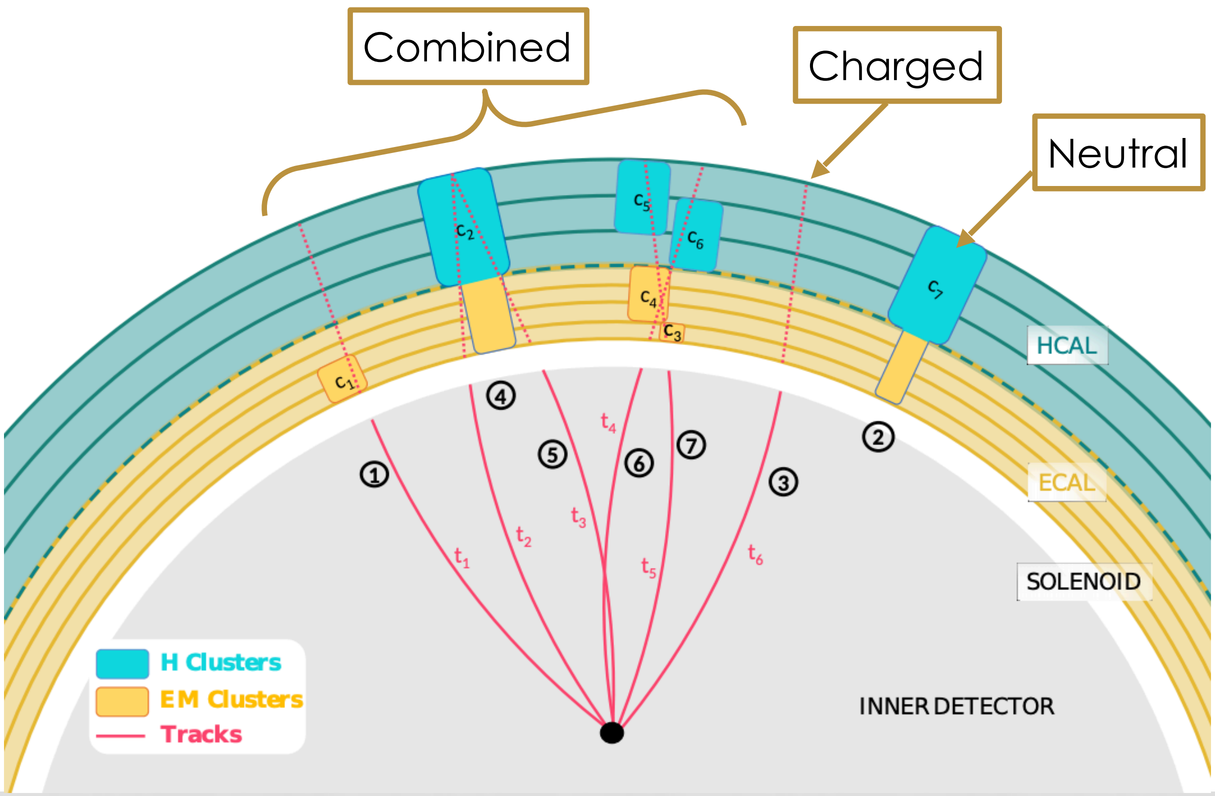


TRACK CALO CLUSTERS (TCC)

ATL-PHYS-PUB-2017-015

- Improve angular resolution of high pT particles
  - Reconstruction of highly boosted t/W/Z/H decays
  - Tracks can resolve structure within clusters due to their spatial resolution.

1. Match tracks to clusters
2. Build 4-vector from matched objects
   - Spatial information from the tracker (eta, phi)
   - Energy measurements from the calorimeter (pt, m)



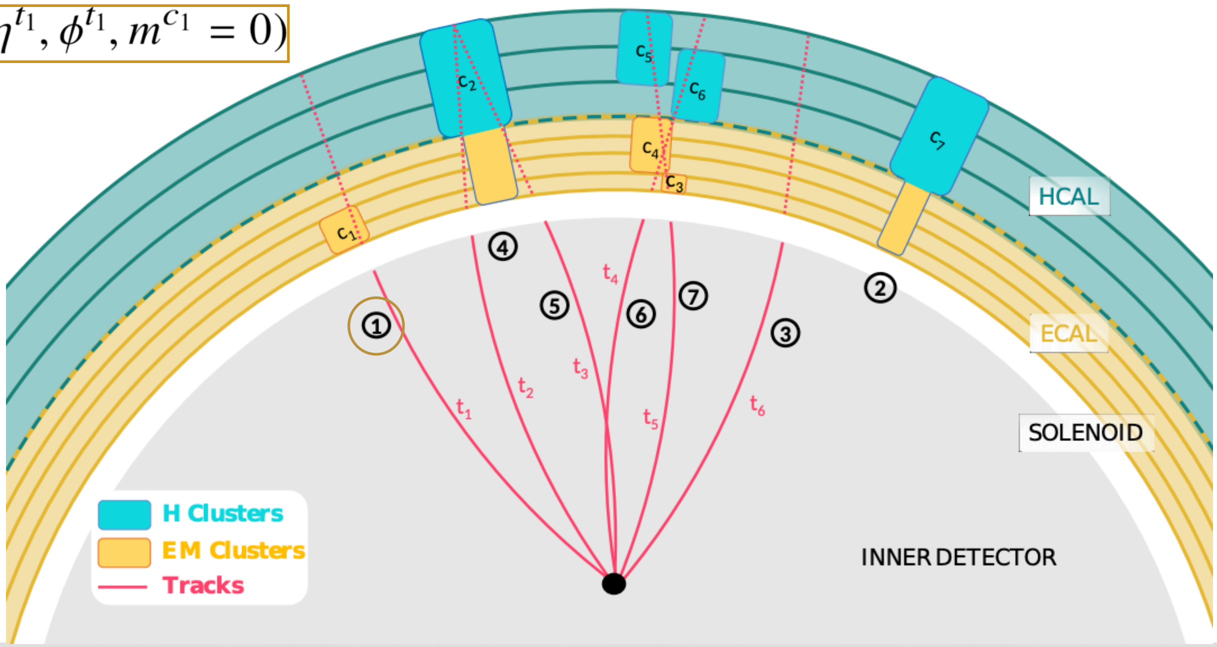
TRACK CALO CLUSTERS (TCC)

ATL-PHYS-PUB-2017-015

- Improve angular resolution of high pT particles
  - Reconstruction of highly boosted t/W/Z/H decays
  - Tracks can resolve structure within clusters due to their spatial resolution.

1. Match tracks to clusters
2. Build 4-vector from matched objects
   - Spatial information from the tracker (eta, phi)
   - Energy measurements from the calorimeter (pt, m)

$$\boxed{\mathrm{TCC}_{①} = (p_{\mathrm{T}}^{c_1}, \eta^{t_1}, \phi^{t_1}, m^{c_1} = 0)}$$



TRACK CALO CLUSTERS (TCC)

ATL-PHYS-PUB-2017-015

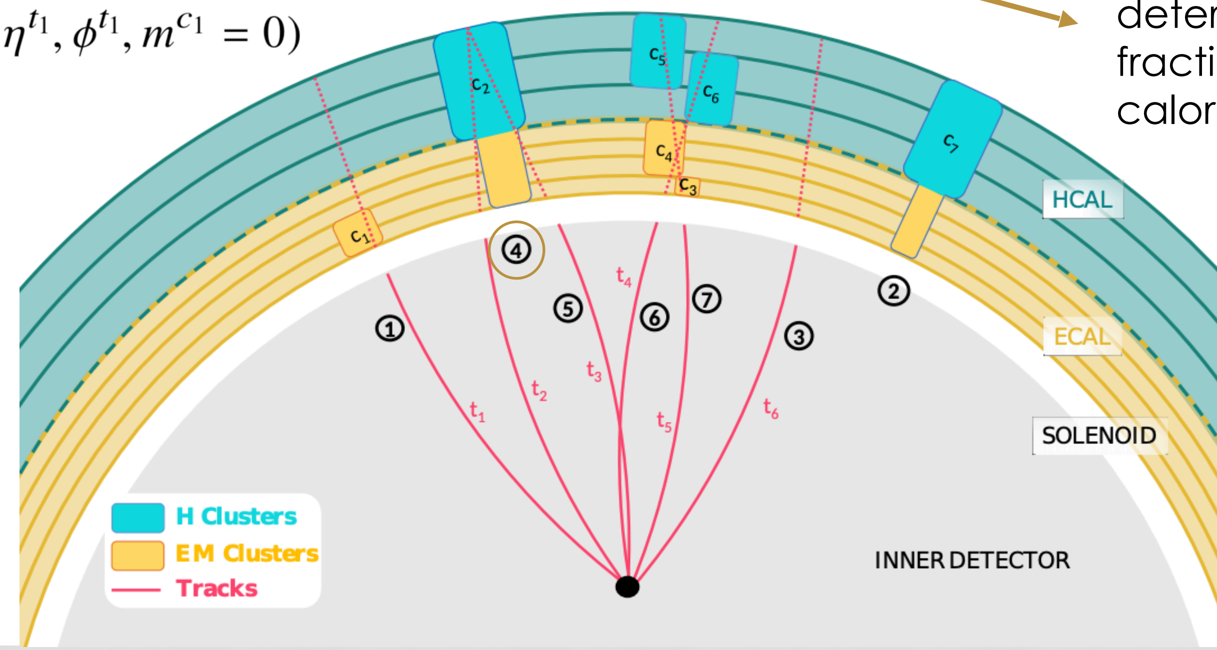
- Improve angular resolution of high pT particles
  - Reconstruction of highly boosted t/W/Z/H decays
  - Tracks can resolve structure within clusters due to their spatial resolution.

1. Match tracks to clusters
2. Build 4-vector from matched objects
   - Spatial information from the tracker (eta, phi)
   - Energy measurements from the calorimeter (pt, m)

$$\mathrm{TCC}_{④} = \left(p_{\mathrm{T}}^{c_2} \frac{p_{\mathrm{T}}^{t_2}}{p_{\mathrm{T}}\left[\mathbf{p}^{t_2} + \mathbf{p}^{t_3}\right]}, \eta^{t_2}, \phi^{t_2}, m^{c_2} \frac{p_{\mathrm{T}}^{t_2}}{p_{\mathrm{T}}\left[\mathbf{p}^{t_2} + \mathbf{p}^{t_3}\right]} = 0 \right)$$

Use track pT to determine sharing fraction of calorimeter energy

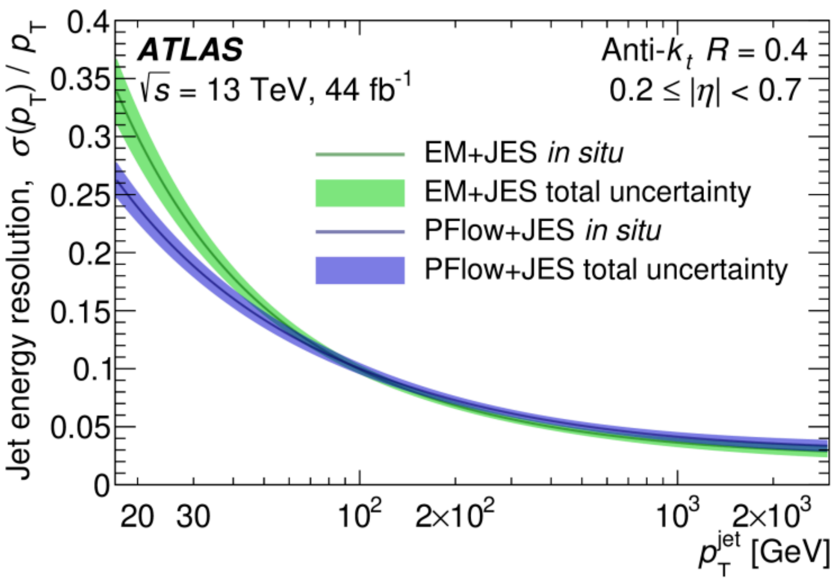
$$\mathrm{TCC}_{①} = (p_{\mathrm{T}}^{c_1}, \eta^{t_1}, \phi^{t_1}, m^{c_1} = 0)$$



PFLOW VS. TCC

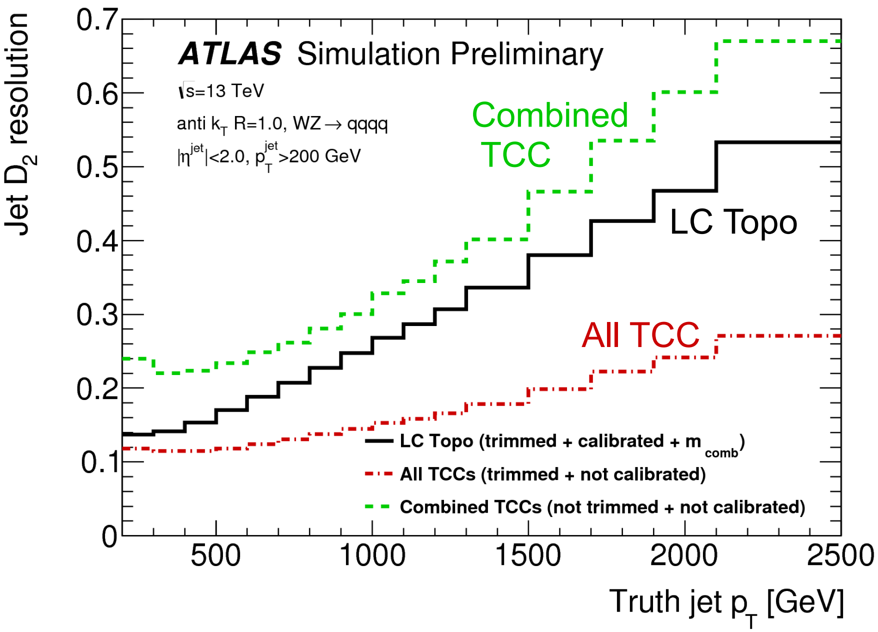
PFlow

- **Start from tracks** selected with stringent quality criteria, subtract track energy from clusters.
- Very good at **low pT.**
  - Improve the **performance/pileup** stability of low pT jets.
  - High pT tracks are excluded.The improvement starts to decrease at very high pT (fewer clusters will be corrected).



TCC

- **Start from calo cluste**r, its energy is split onto the angular position of the tracks.
- Very good at **high pT**
  - Use calorimeter energy scale.
  - Use tracker spatial coordinates: Helps to better understand the structure of the jet



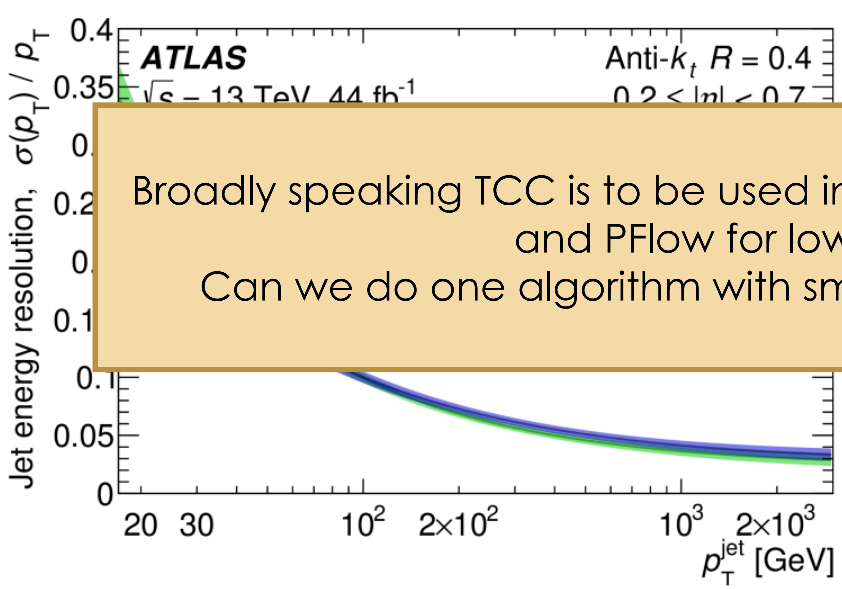
PFLOW VS. TCC

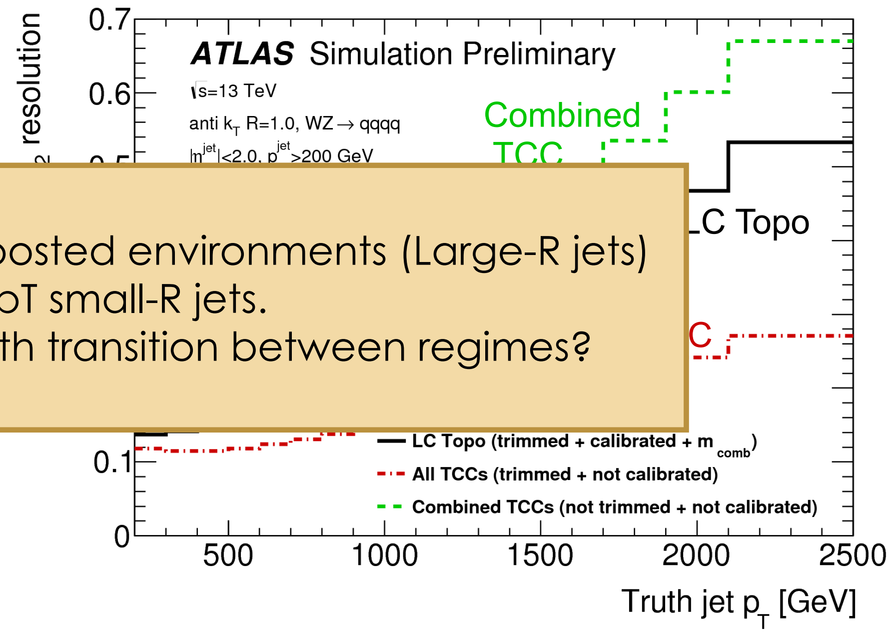
PFlow

- **Start from tracks** selected with stringent quality criteria, subtract track energy from clusters.
- Very good at **low pT.**
  - Improve the **performance/pileup** stability of low pT jets.
  - High pT tracks are excluded.The improvement starts to decrease at very high pT (fewer clusters will be corrected).

TCC

- **Start from calo cluste**r, its energy is split onto the angular position of the tracks.
- Very good at **high pT**
  - Use calorimeter energy scale.
  - Use tracker spatial coordinates: Helps to better understand the structure of the jet

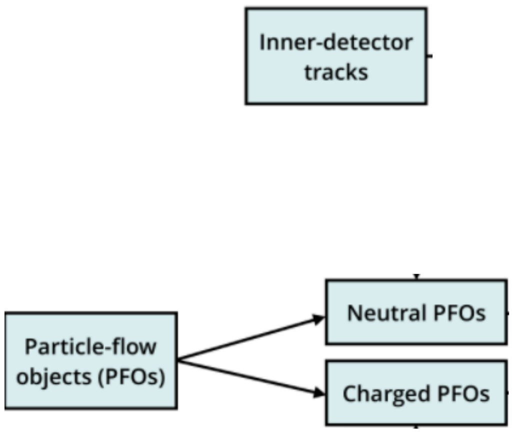


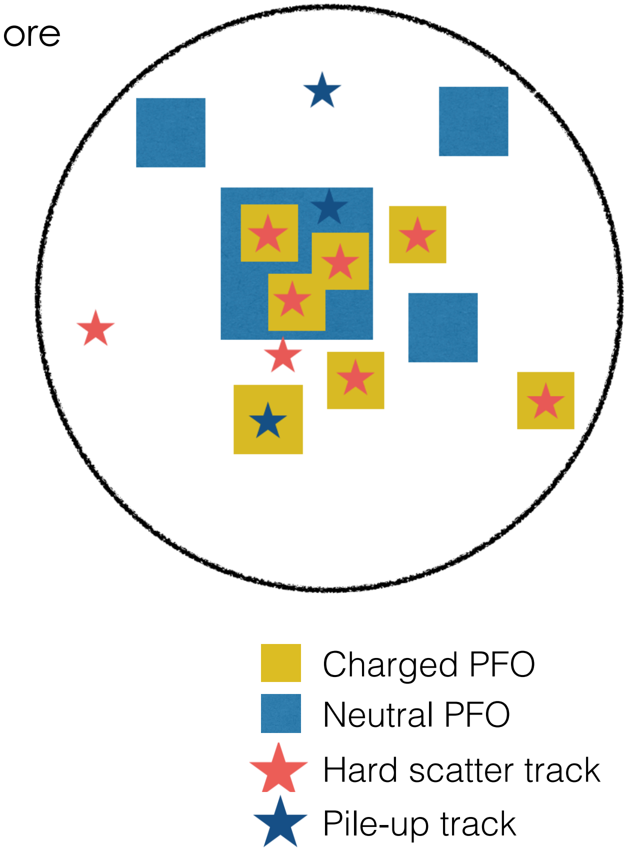


Broadly speaking TCC is to be used in boosted environments (Large-R jets) and PFlow for lower pT small-R jets.
Can we do one algorithm with smooth transition between regimes?

UNIFIED FLOW OBJECTS (UFO)

UFO merges PFO and TCC inputs, aiming to obtain a better and more stable performance in a wide pT range.

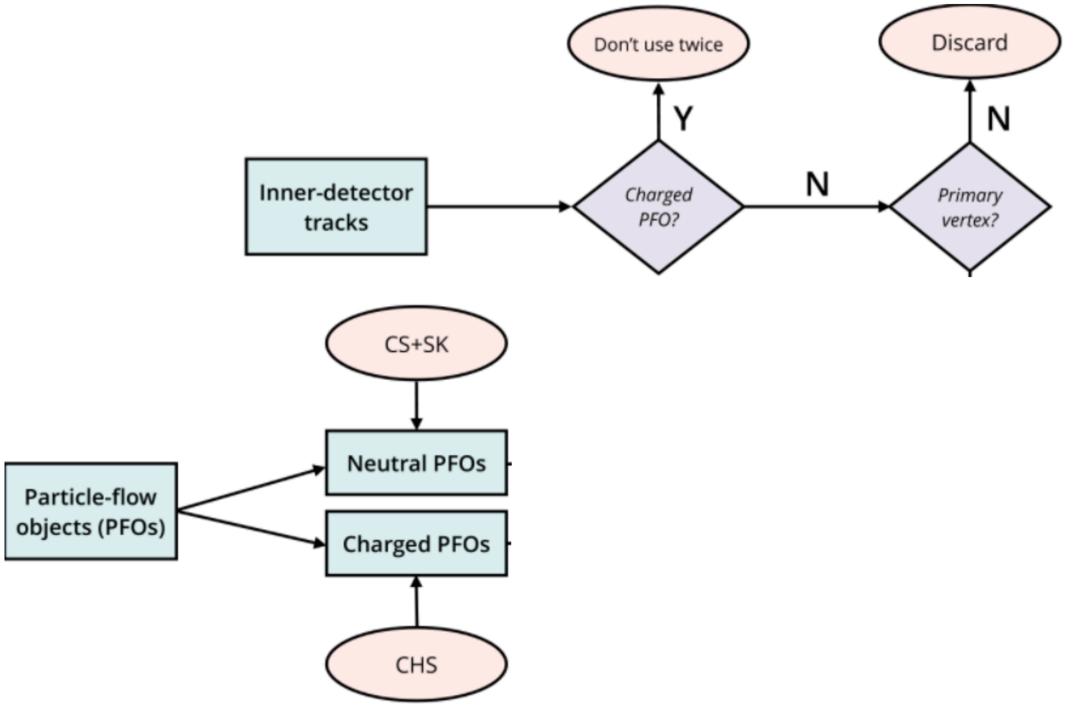


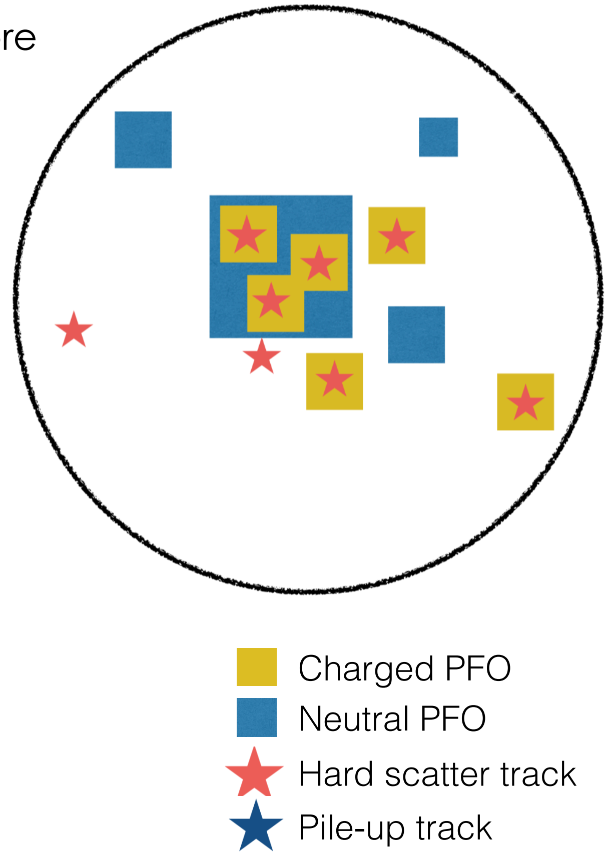


[Eur. Phys. J. C 81 (2021) 334]

UNIFIED FLOW OBJECTS (UFO)

UFO merges PFO and TCC inputs, aiming to obtain a better and more stable performance in a wide pT range.



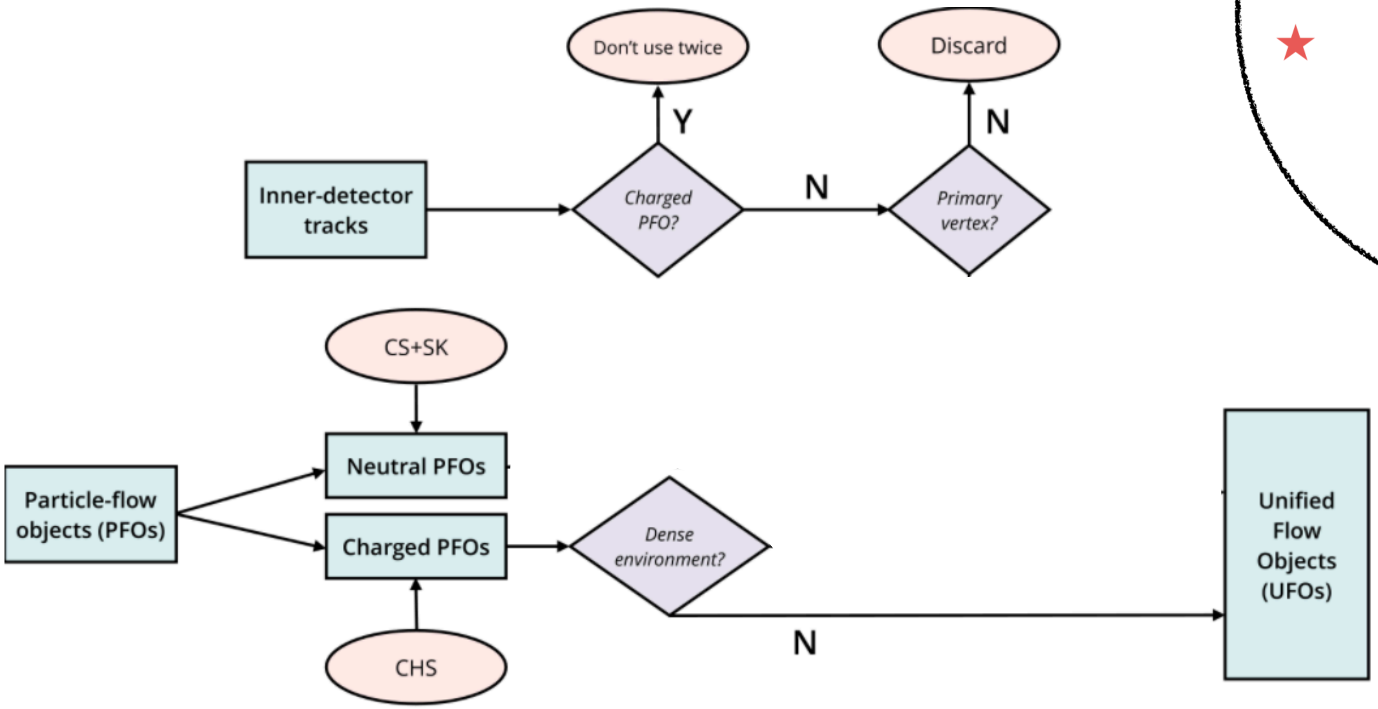


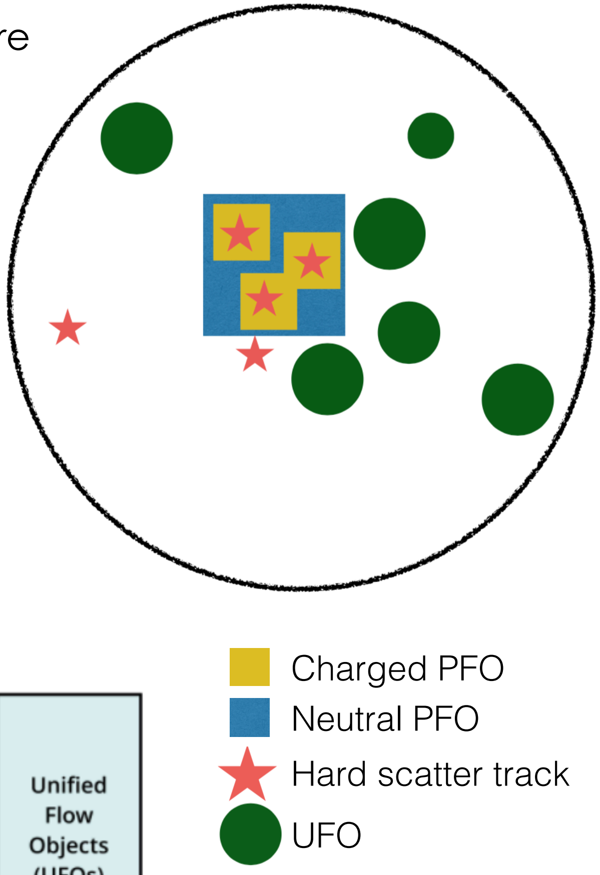
[Eur. Phys. J. C 81 (2021) 334]

- CHS: Charge Hadron Subtraction (Charged PFOs which are not matched to the PV are removed)
- CS+SK: Constituent Subtraction and Soft Killer (next slides)

UNIFIED FLOW OBJECTS (UFO)

UFO merges PFO and TCC inputs, aiming to obtain a better and more stable performance in a wide pT range.

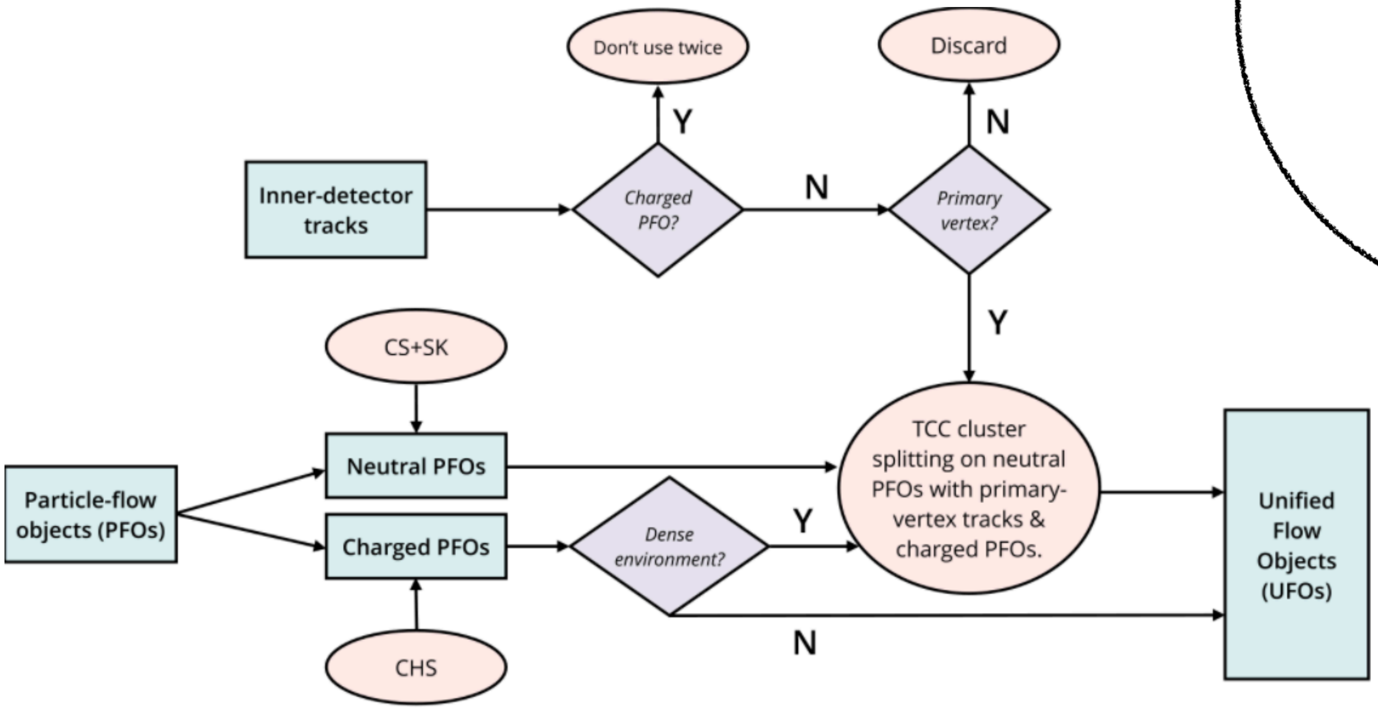


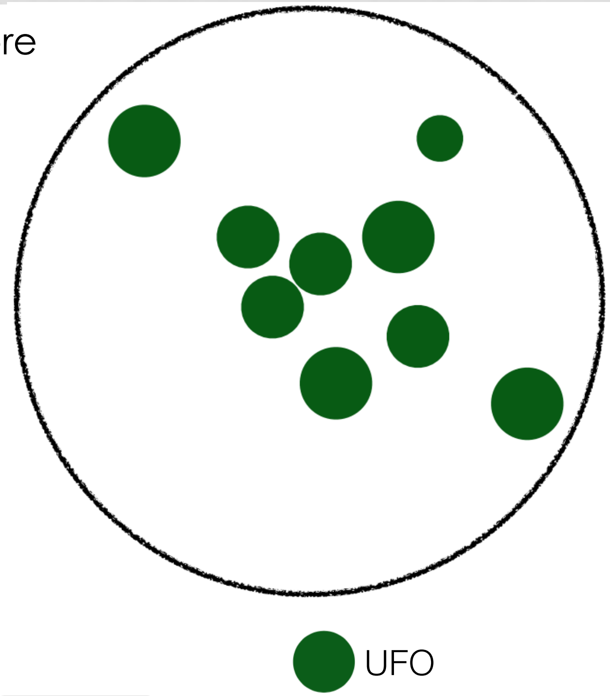


[Eur. Phys. J. C 81 (2021) 334]

UNIFIED FLOW OBJECTS (UFO)

UFO merges PFO and TCC inputs, aiming to obtain a better and more stable performance in a wide pT range.



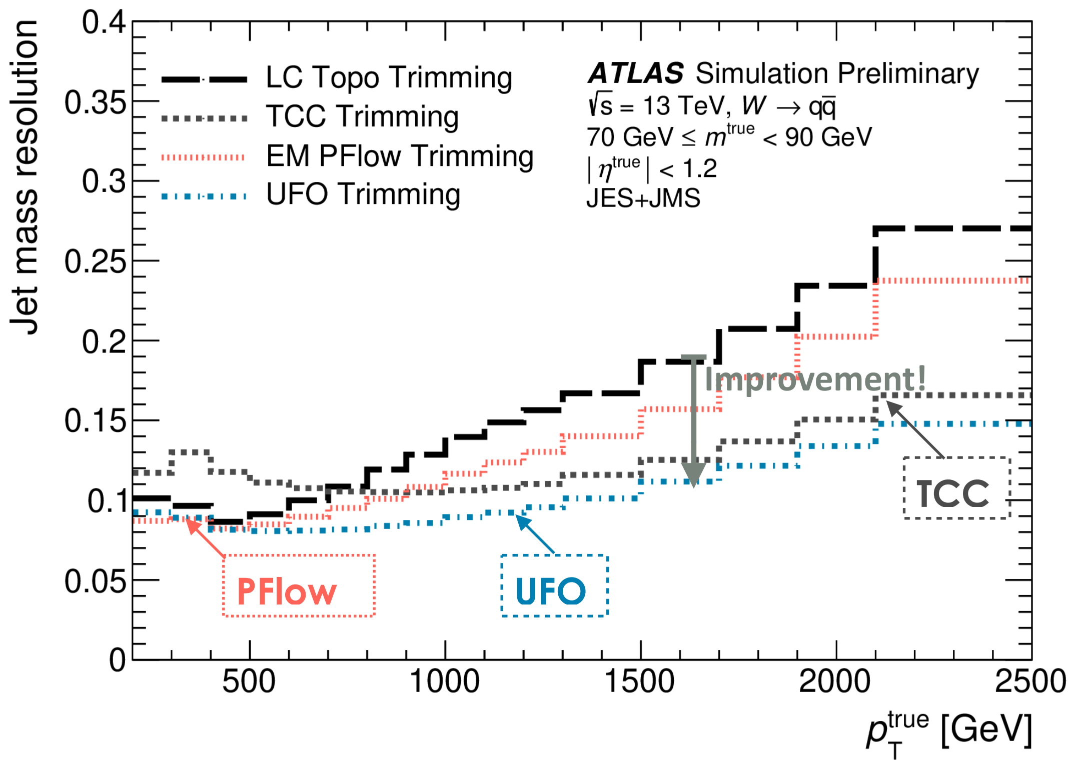


Eur. Phys. J. C 81 (2021) 334

LARGE-RADIUS JETS OPTIMISATION

Jet Mass Resolution

Resolution of the mass Response= $\frac{Mass^{reco}}{Mass^{truth}}$



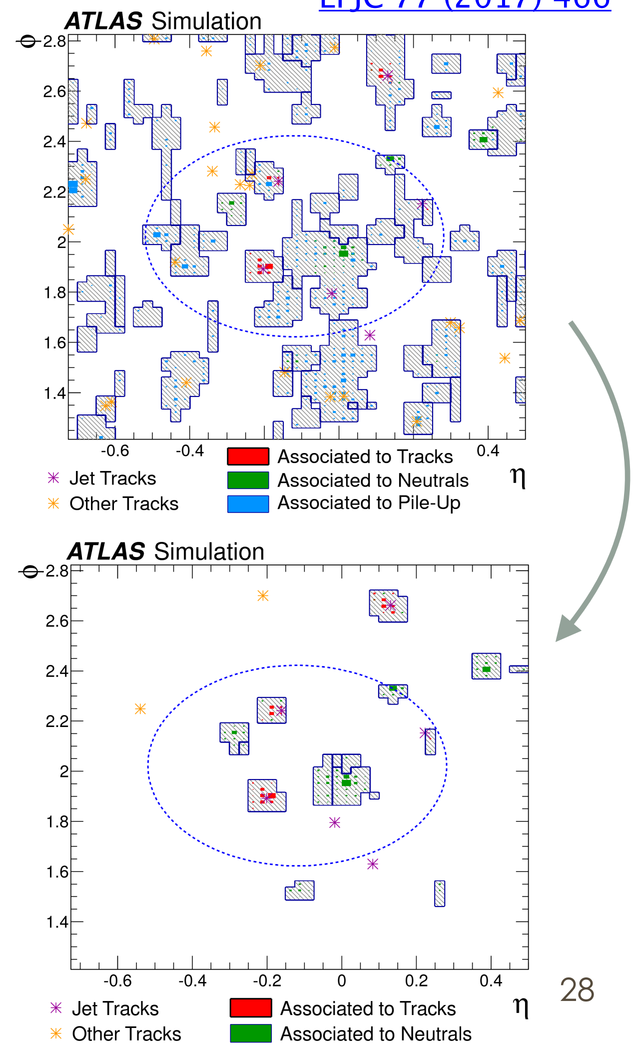
- At low Pt:
  Resolution consistent with PFlow
- At high Pt:
  Better resolution than TCC

PILE-UP MITIGATION AND GROOMING

PILE-UP (PU)

Effects of PU
- Energy contribution to hard scattering objects
  - Biasing total jet energy, and smearing resolution of hard scatter jets
- Objects originating out of PU
  - Adding PU jets in the event

EPJC 77 (2017) 466



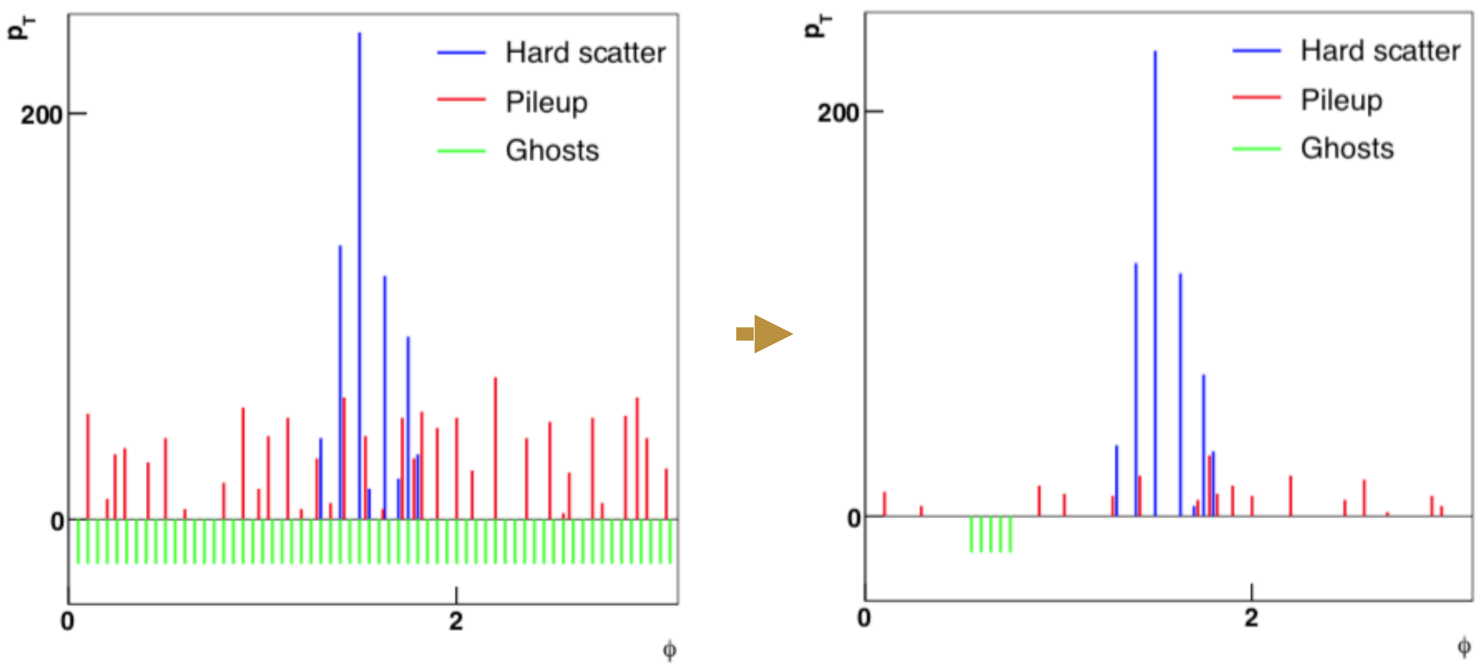
How we can mitigate PU?
- Estimate and subtract PU contribution to object energy
- Filter out objects originating from PU

At different levels
- From the constituent reconstruction
  - Topo-Clusters (noise suppression)
  - Particle Flow/TCC/UFO (no Primary Vertex charged objects can be rejected)
- Constituent-level
  - Voronoi Subtraction
  - Constituent Subtraction
  - SoftKiller
  - PUPPI
- Jet-level
  - Jet-Area Subtraction
  - Grooming (trimming, soft drop)
  - Jet Vertex Tagger (JVT)
  - Forward Jet Vertex Tagger (fJVT)

CONSTITUENT SUBTRACTION (CS)

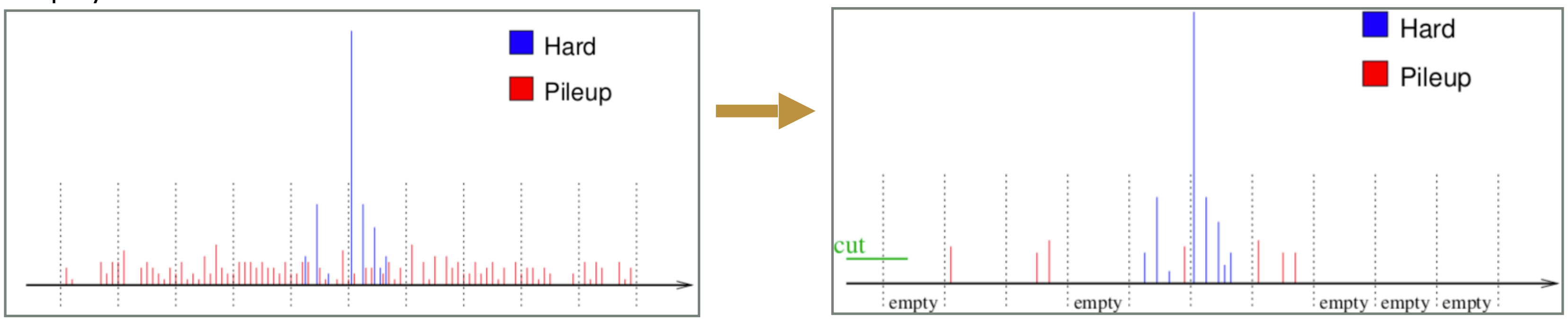
JHEP 1406 (2014) 092

- A "constituent area" subtraction
- Add ghosts to the event with $p_T^g = A_g \times \rho$
  - ✳ ρ is the median energy density
  - ✳ A_g is the area of the ghost (fixed $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$)
- Subtract ghost's contribution from p_T of closest constituent (in ΔR)
  - ✳ Until ΔR(ghost, constituent) > ΔR_{max}
  - ✳ Algorithm built so that constituent's p_T never goes negative

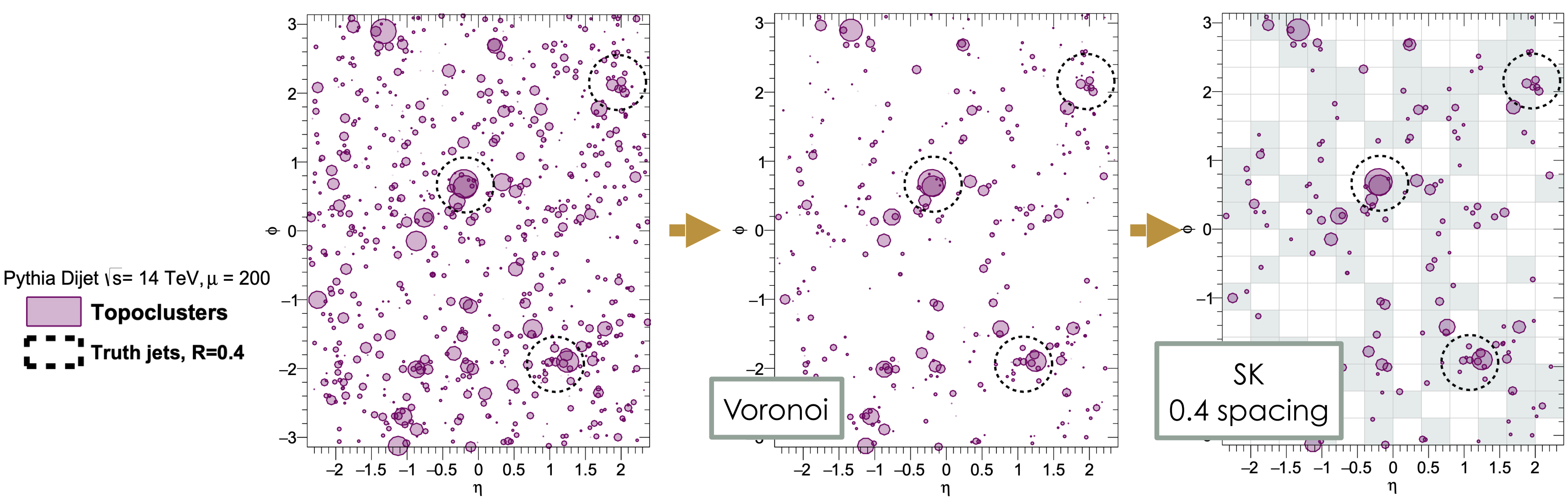


SOFT-KILLER (SK)

- Removes low-p_T constituents
- Applying a p_T cut on particles on an event-by-event basis
- p_T cut determined by putting constituents into an η-φ grid, and requiring half of grid spaces to be empty after the cut

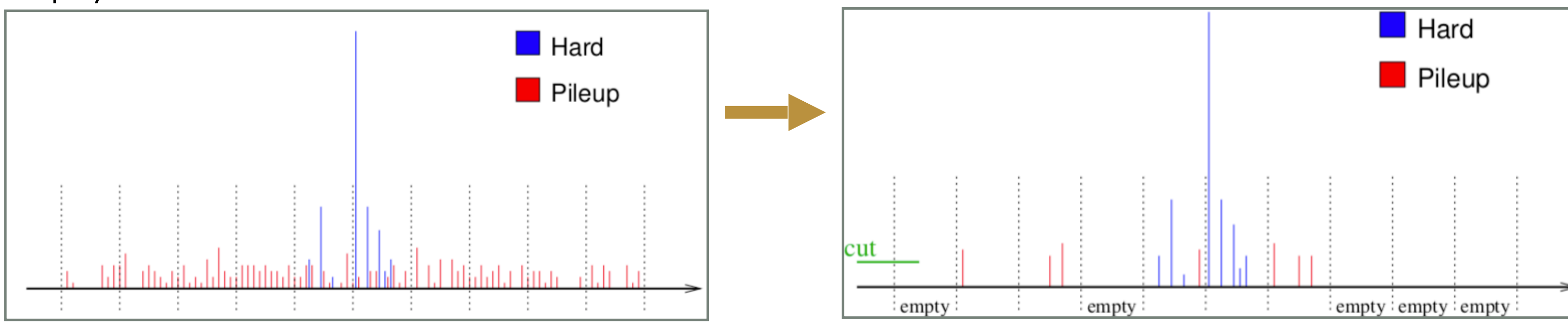


- Soft-Killer does not change remaining constituent's momenta
  - ✳ Combined with 'area' based correction (Vonoroi or CS)

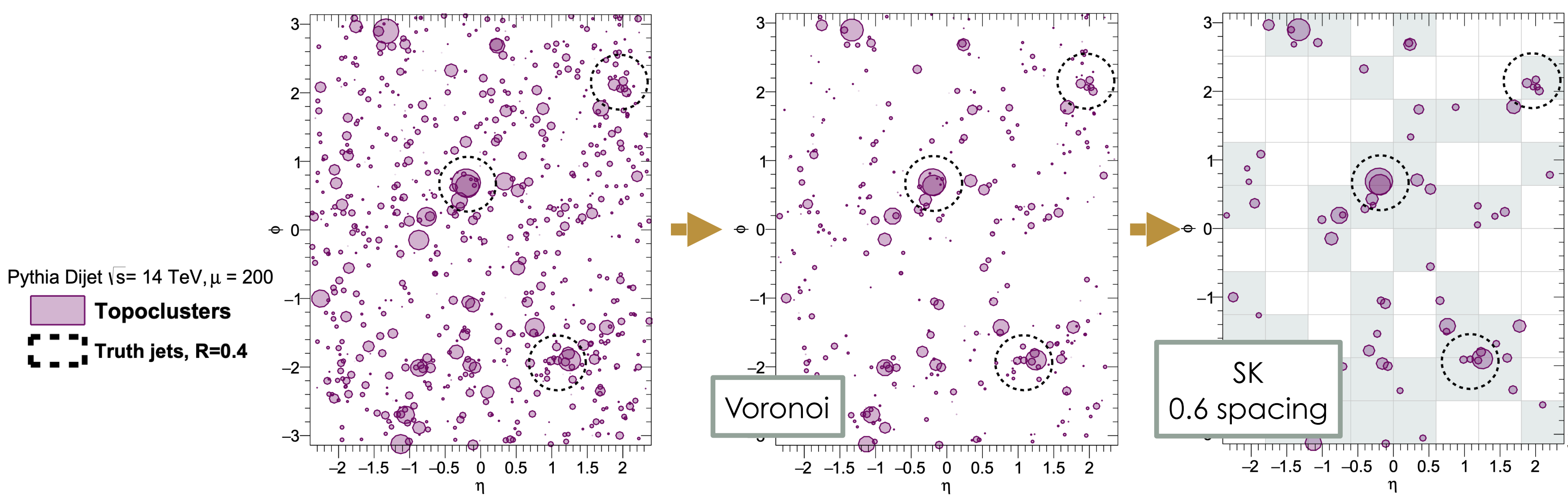


SOFT-KILLER (SK)

- Removes low-p_T constituents
- Applying a p_T cut on particles on an event-by-event basis
- p_T cut determined by putting constituents into an η-φ grid, and requiring half of grid spaces to be empty after the cut



- Soft-Killer does not change remaining constituent's momenta
  - ✳ Combined with 'area' based correction (Voronoi or CS)

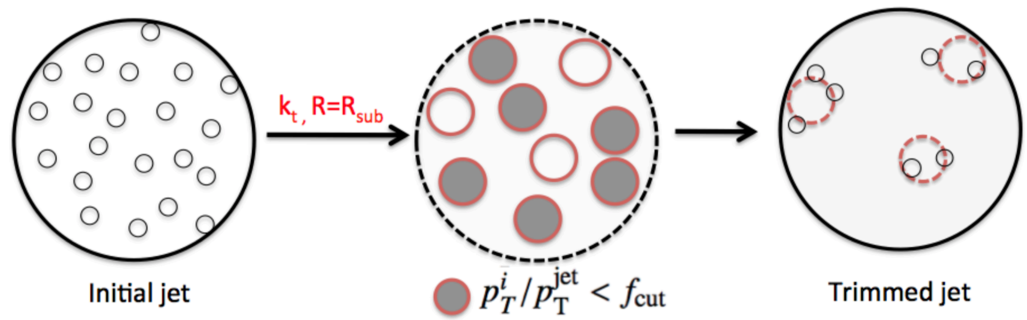


JET GROOMING

- Several groomers in the market (trimming, pruning, mMDT, soft drop…)
  - Historically ATLAS have used trimming. But we recently performed a new scan taking into account different inputs, PU mitigation techniques and grooming

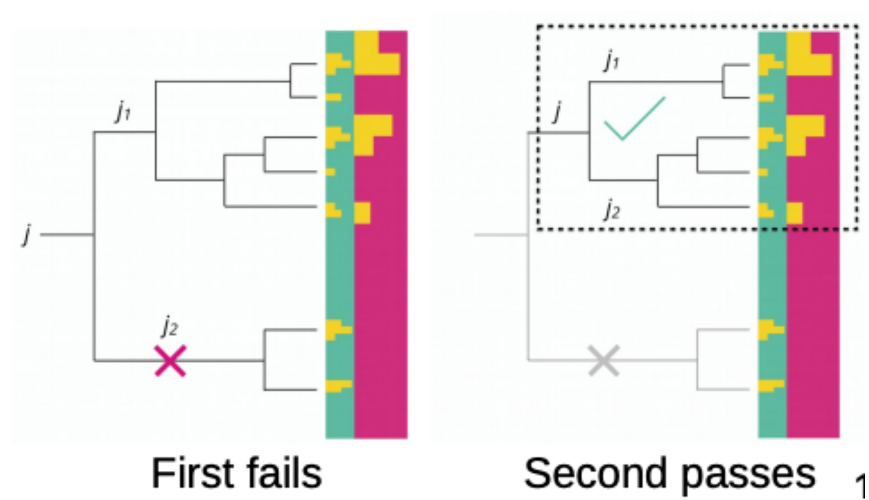
Trimming

- Remove low-pT contributions at the subjet level



Soft Drop

- Define splittings using Cambridge-Aachen (cluster closest stuff first)
- Apply SoftDrop
  - Start from first jet splitting
  - Reach passing condition for both splittings



1

$$\frac{min(p_{T1}, p_{T2})}{p_{T1}+p_{T2}} > z_{cut}\left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}$$

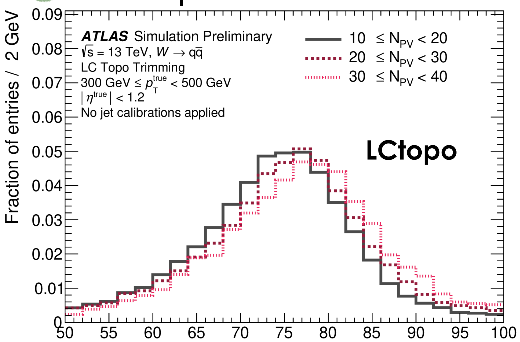
- Keep both
- "Drop" constituent

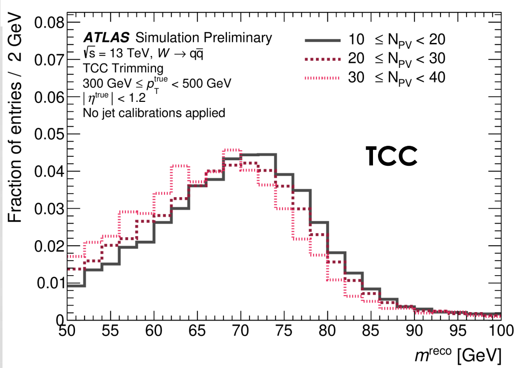
LARGE-R JETS OPTIMISATION AND PERFORMANCE

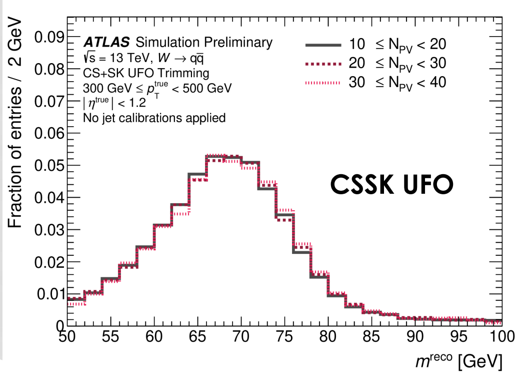
LARGE-RADIUS JETS OPTIMISATION (SCAN)

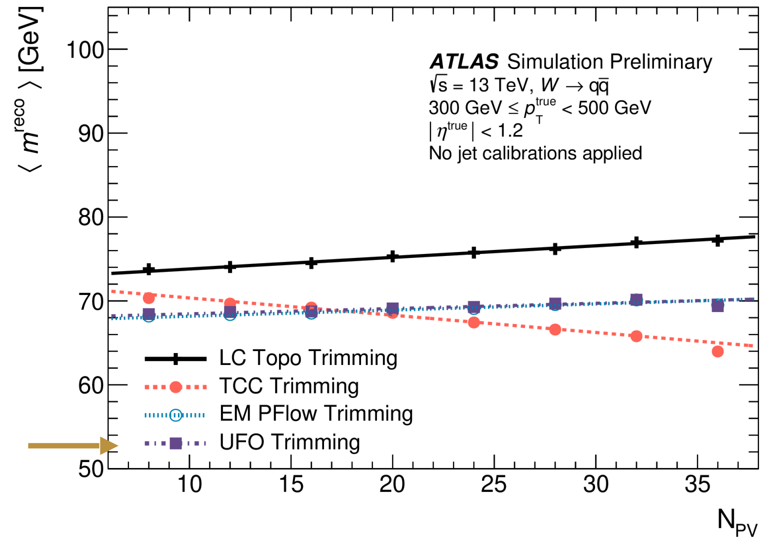
Pile-Up Stability

- Pile-Up affects the reconstructed mass

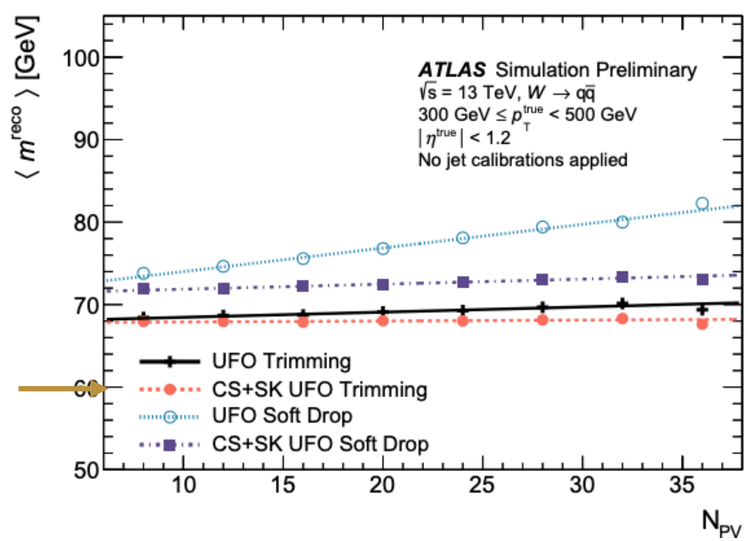








- Clear complementarity between groomers and PU mitigation techniques:

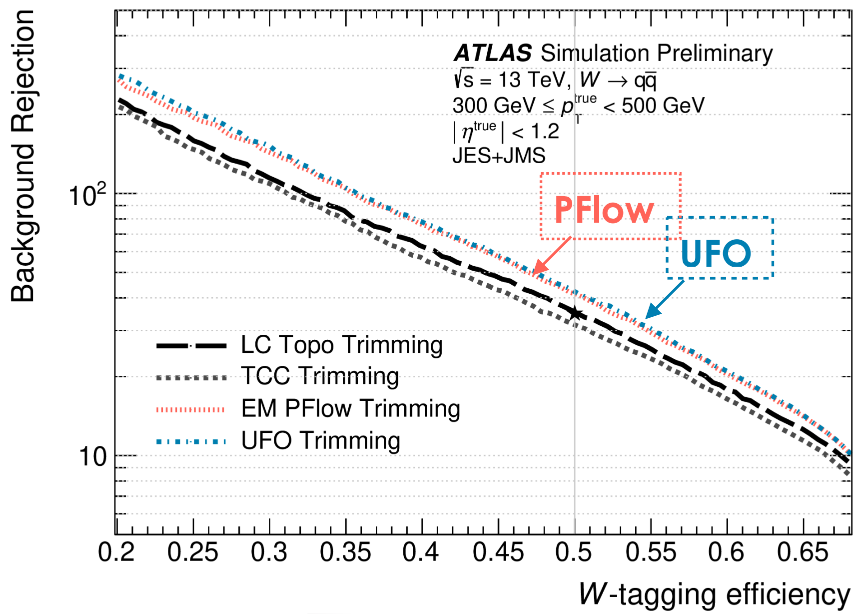


LARGE-RADIUS JETS OPTIMISATION (SCAN)

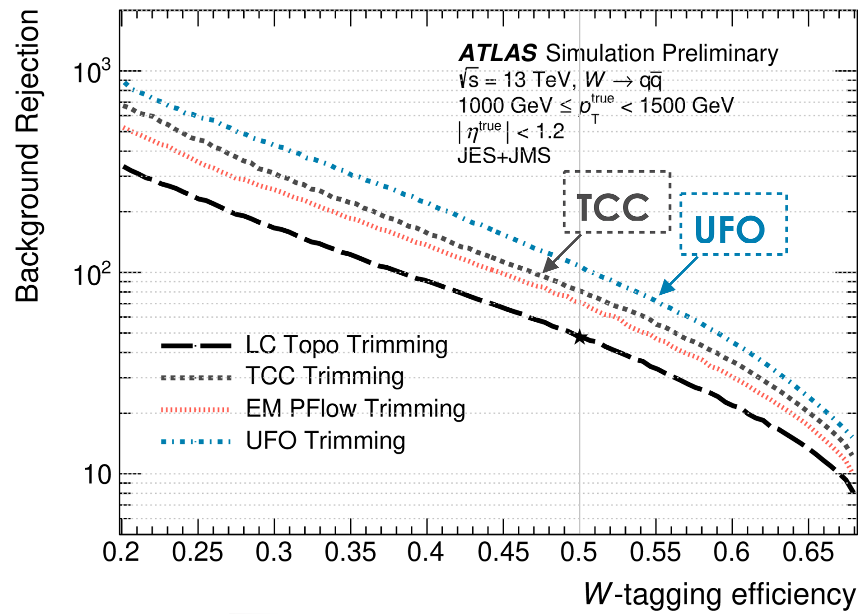
W Tagging Performance

- Signal: W' vrs. Background: Dijet
- for different jet constituents

Based on cuts on mass & jet-substructure variable: 68% mass window + D_2 cut



Low pT

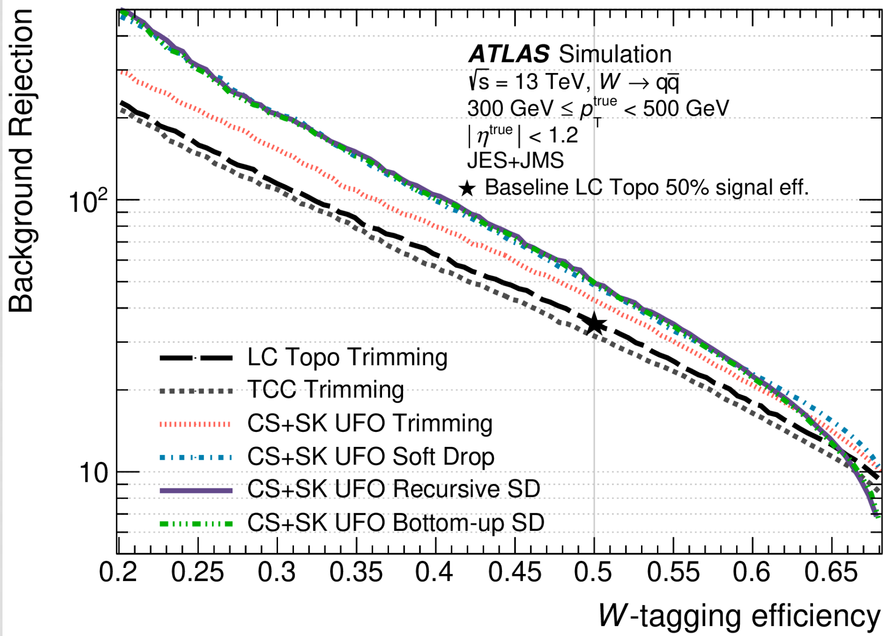


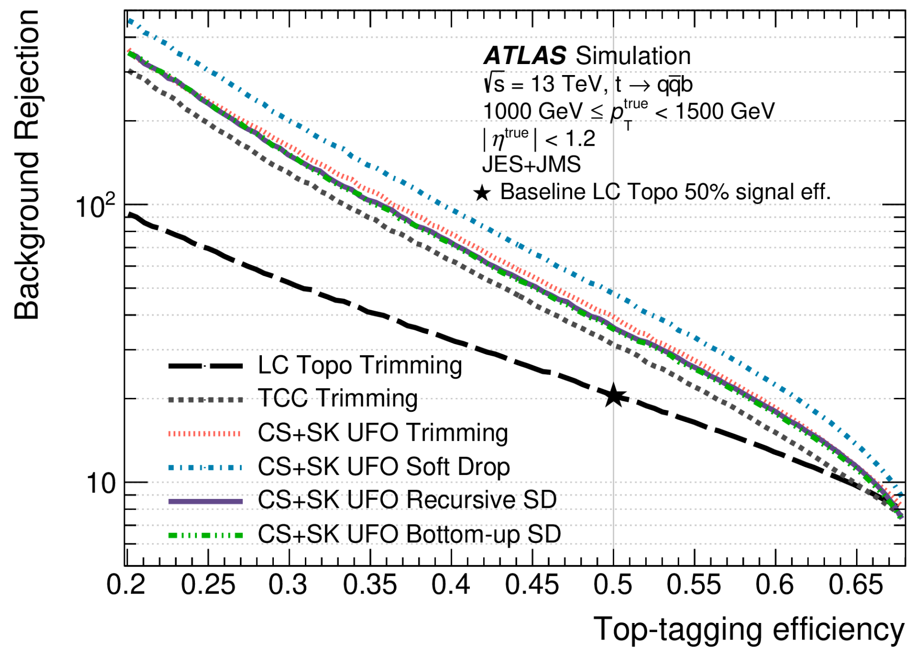
High pT

UFOs already work very well, but we can still improve by surveying jet definitions
ATLAS performed a new **scan of jet definitions**!

LARGE-RADIUS JETS OPTIMISATION (SCAN)

Tagging Performance: finalists





Soft Drop seems to work very well for top, only small losses vs other options for W -tagging

CS+SK UFO SD jets chosen to be the default large radius jet collection for Run3 ATLAS analyses.

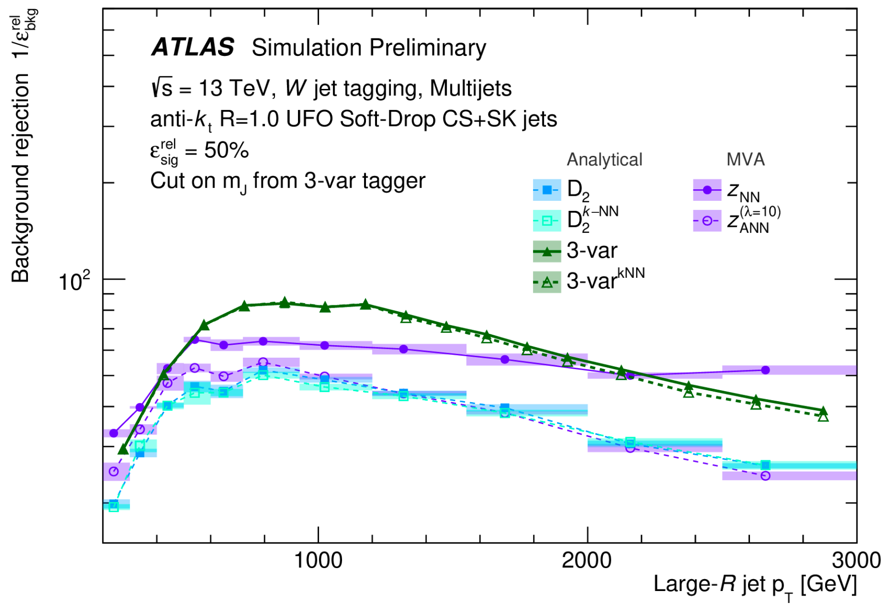
- Currently, the performance of small-R jets is under study

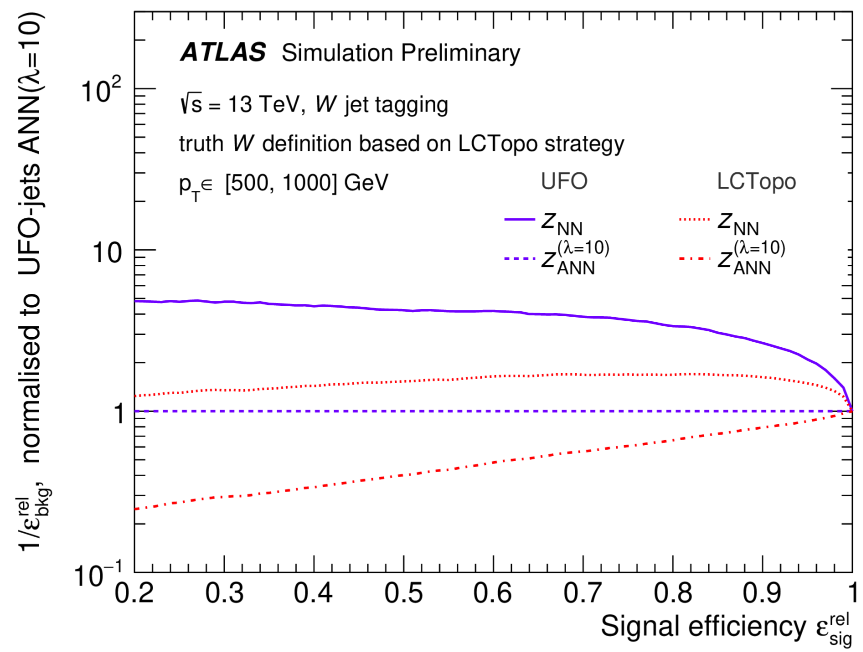
UFO TAGGERS: W/Z BOOSTED DECAYS

ATL-PHYS-PUB-2021-029

- Considered cut-based and ML-based approaches (including ANN version to remove tagger dependence on jet mass)
- 3-var tagger: jet mass, D_2, n_{trk}

Improved performance from LCTopo -> UFO





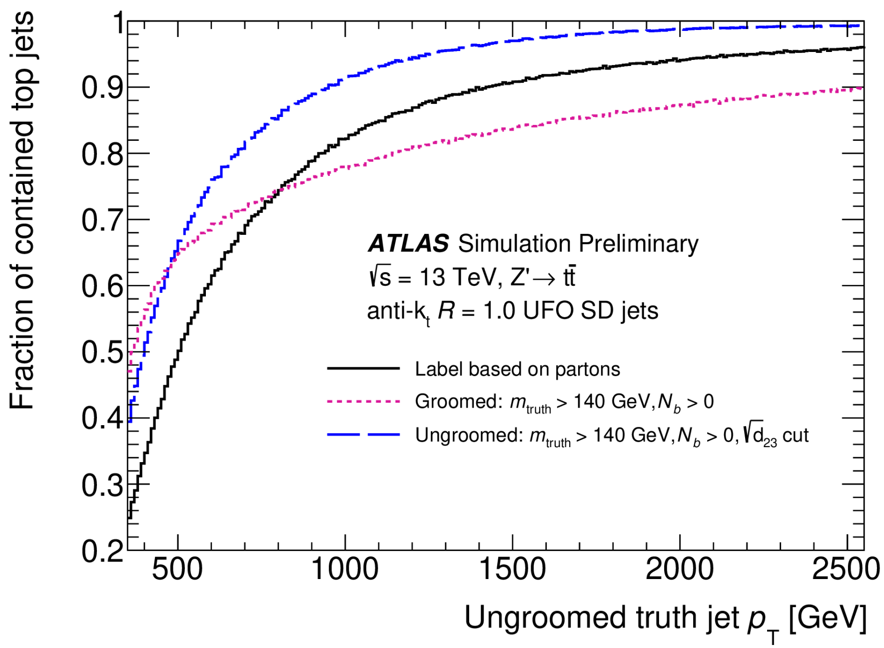
- Dedicated calibrations are derived by comparing tagging efficiencies between data and different MC models.
  - Tagger scale factors:

$$SF = \frac{\epsilon_{Data}}{\epsilon_{MC}} \qquad \epsilon_{Data} = \frac{N^{tag\ pass}_{signal\ fit}}{N^{tag\ pass}_{signal\ fit} + N^{tag\ fail}_{signal\ fit}} \qquad \epsilon_{MC} = \frac{N^{tag\ pass}_{signal}}{N^{tag\ pass}_{signal} + N^{tag\ fail}_{signal}}$$

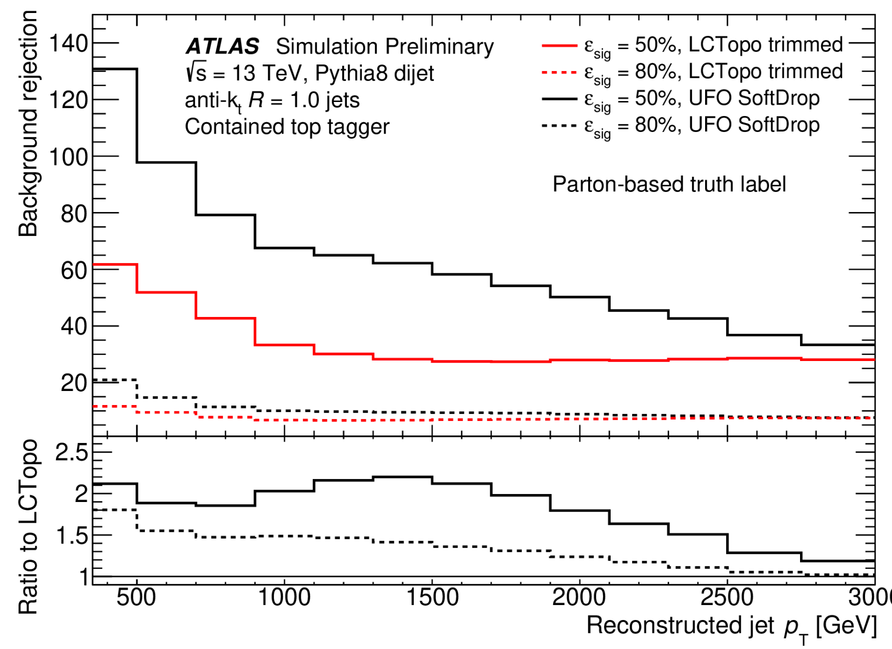
UFO TAGGERS: BOOSTED TOP DECAYS

ATL-PHYS-PUB-2021-028

- New training of DNN tagger exploring different substructure variables
- Improved top-truth labelling has been defined: higher signal efficiency



Improved performance from LCTopo -> UFO



- Dedicated calibrations are derived by comparing tagging efficiencies between data and different MC models.
  - Tagger scale factors:

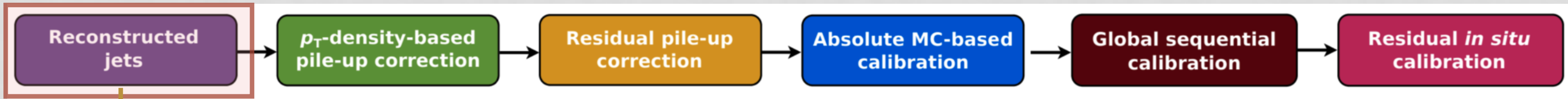
$$SF = \frac{\epsilon_{Data}}{\epsilon_{MC}}$$

$$\epsilon_{Data} = \frac{N^{tag\ pass}_{signal\ fit}}{N^{tag\ pass}_{signal\ fit} + N^{tag\ fail}_{signal\ fit}} \qquad \epsilon_{MC} = \frac{N^{tag\ pass}_{signal}}{N^{tag\ pass}_{signal} + N^{tag\ fail}_{signal}}$$

JET ENERGY SCALE CALIBRATION

SMALL-R JET CALIBRATION

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



PFlow as default with CHS

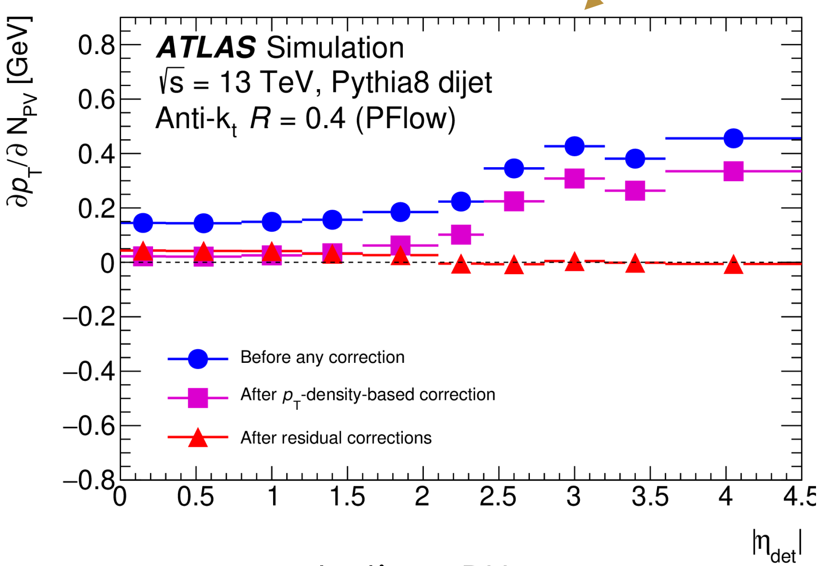
SMALL-R JET CALIBRATION

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

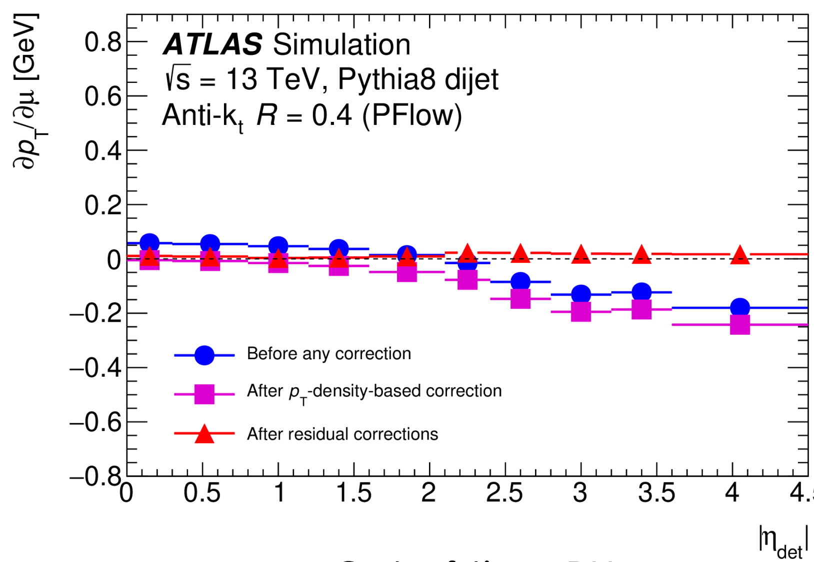
Reconstructed jets → **p_T-density-based pile-up correction** → **Residual pile-up correction** → **Absolute MC-based calibration** → **Global sequential calibration** → **Residual *in situ* calibration**

PFlow as default with CHS

$$p_T^{\mathrm{corr}} = p_T^{\mathrm{reco}} - \overbrace{\rho \times A}^{\text{Jet-area-based PU subtraction}} - \overbrace{\alpha \times (N_{\mathrm{PV}} - 1) - \beta \times \mu}^{\text{Removes residual PU dependency}}$$



In-time PU



Out-of-time PU

SMALL-R JET CALIBRATION

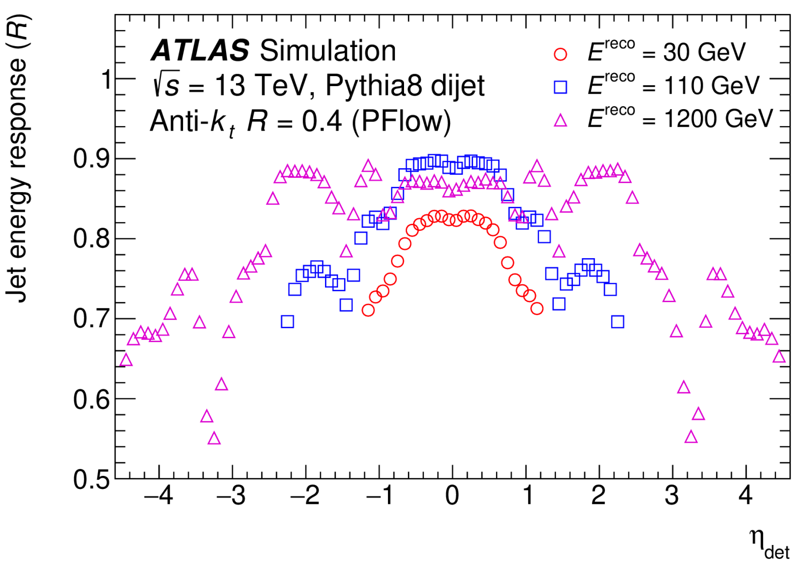
Eur. Phys. J. C 81 (2021) 689

Reconstructed jets → p_T-density-based pile-up correction → Residual pile-up correction → **Absolute MC-based calibration** → Global sequential calibration → Residual *in situ* calibration

Restore the scale of reconstructed jets to particle level
Both **energy** and **direction** are calibrated

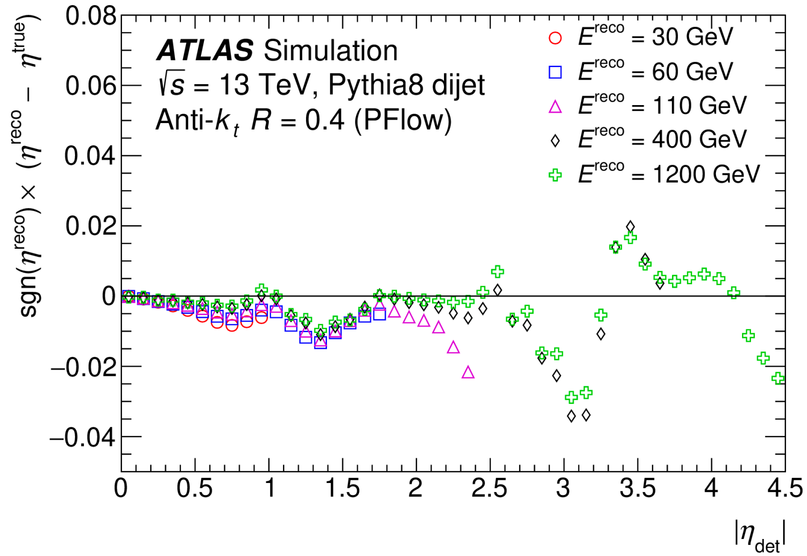
- Determines energy response R

$$R(E_{\text{jet}}^{\text{truth}}, \eta_{\text{det}}) = E_{\text{jet}}^{\text{reco}} / E_{\text{jet}}^{\text{truth}}$$



- Corrects for η mis-reconstruction

$$\Delta\eta(E_{\text{jet}}^{\text{truth}}, |\eta_{det}|) = \eta^{reco} - \eta^{truth}$$



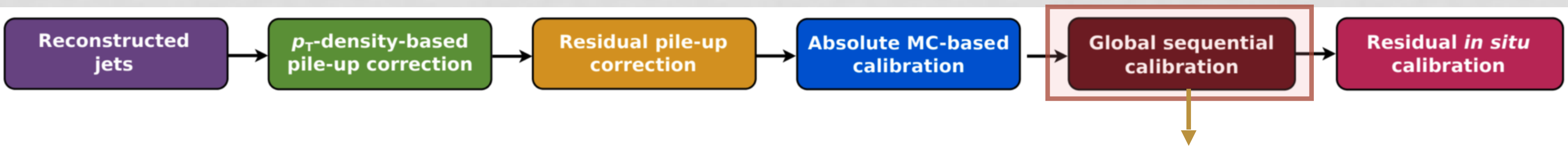
Numerical inversion and calibration:

$$R(E_{\text{jet}}^{\text{truth}}, \eta_{\text{det}}) \rightarrow R^{-1}(E_{\text{jet}}^{\text{reco}}, \eta_{\text{det}})$$

$$\Delta\eta(E_{\text{jet}}^{\text{truth}}, |\eta_{det}|) \rightarrow \Delta\eta(E_{\text{jet}}^{\text{reco}}, |\eta_{det}|)$$

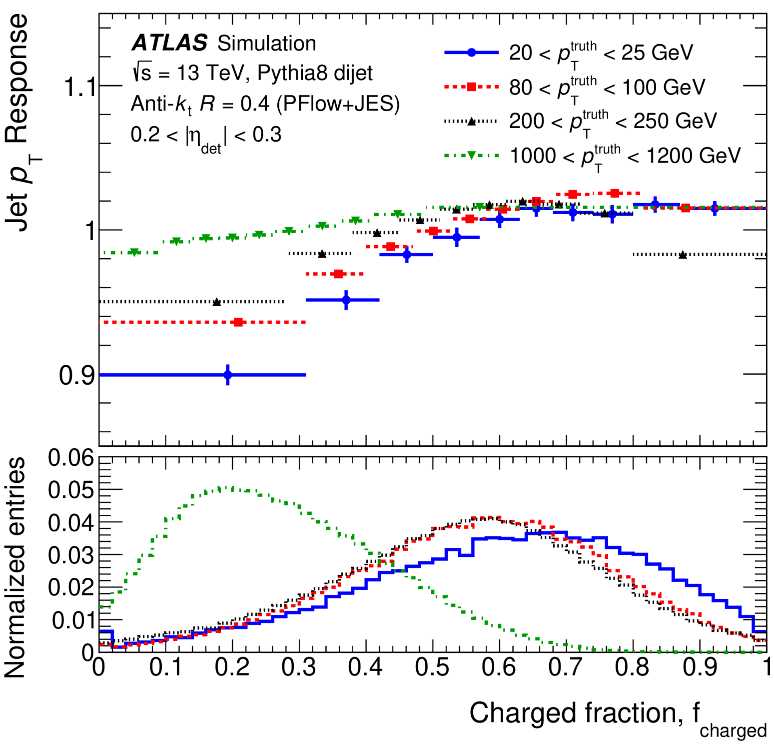
SMALL-R JET CALIBRATION

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

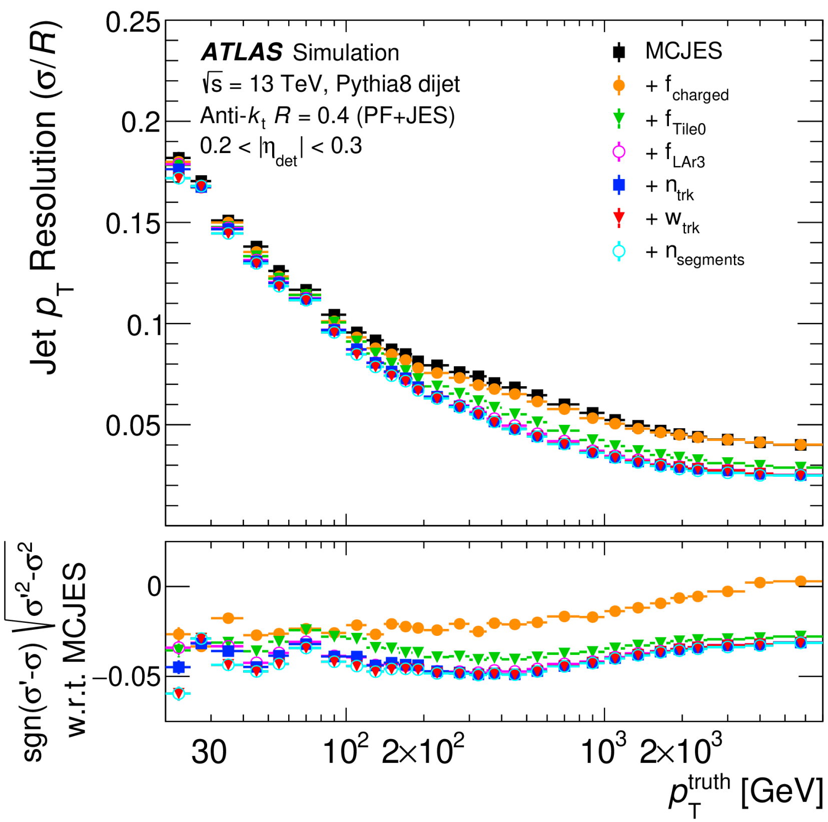


Reduces jet-by-jet response variations

- Using calorimeter, track, and muon segments variables

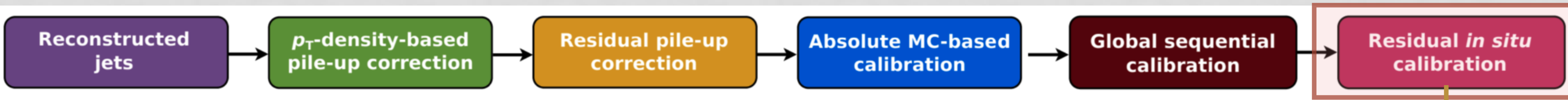


- Reduce jet p_T resolution



SMALL-R JET CALIBRATION

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



Residual calibration using *insitu* measurements applied **only to data**

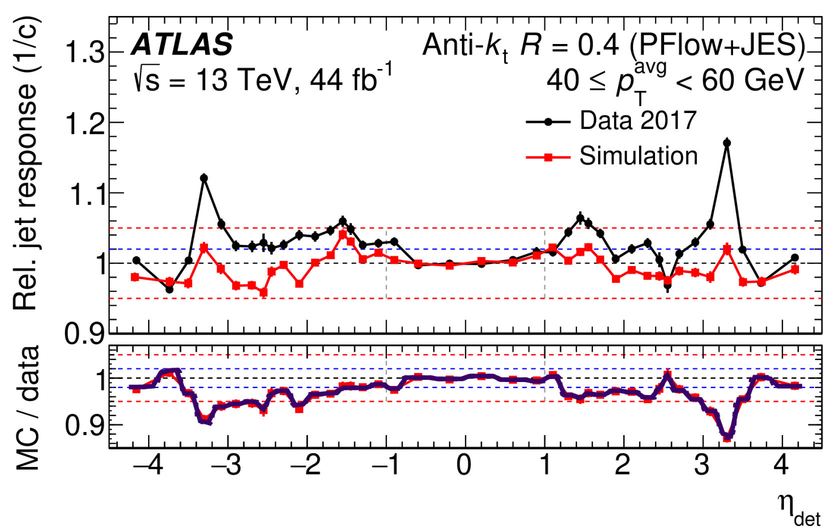
Using p_T balance between well-measured reference objects and probed jet: $r_{in\ situ}(p_T^{ref}) = \frac{p_T^{jet}}{p_T^{ref}}$

Define data-MC differences in terms of double ratios and derive calibrations:

$$c(p_T^{ref}) = \left\langle r_{insitu}^{data} / r_{insitu}^{MC} \right\rangle \rightarrow c^{-1}(p_T^{jet})$$

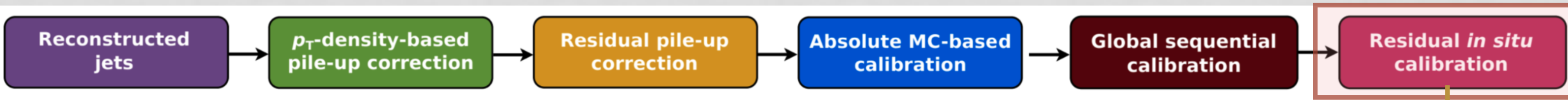
Relative η intercalibration

Dijets with well measured central jet balancing probed (forward) jet



SMALL-R JET CALIBRATION

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



Residual calibration using *insitu* measurements applied **only to data**

Using p_{T} balance between well-measured reference objects and probed jet: $r_{in\,situ}(p_{\mathrm{T}}^{\mathrm{ref}}) = \dfrac{p_{\mathrm{T}}^{\mathrm{jet}}}{p_{\mathrm{T}}^{\mathrm{ref}}}$

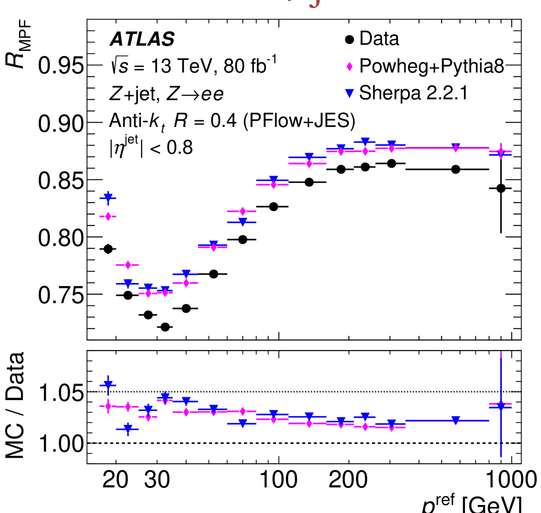
Define data-MC differences in terms of double ratios and derive calibrations:

$$R(p_{\mathrm{T}}^{\mathrm{ref}}) = \left\langle r_{\mathrm{insitu}}^{\mathrm{data}} / r_{\mathrm{insitu}}^{\mathrm{MC}} \right\rangle \rightarrow c = R^{-1}(p_{\mathrm{T}}^{\mathrm{jet}})$$

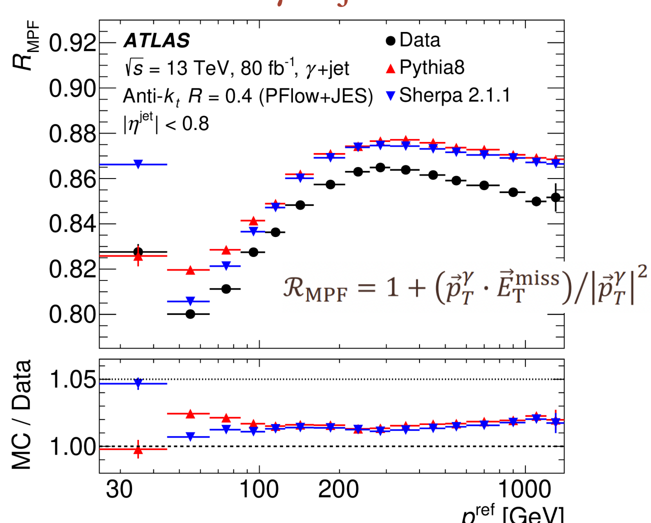
Balance measurement

Absolute response calibration combining various references for full phase space coverage

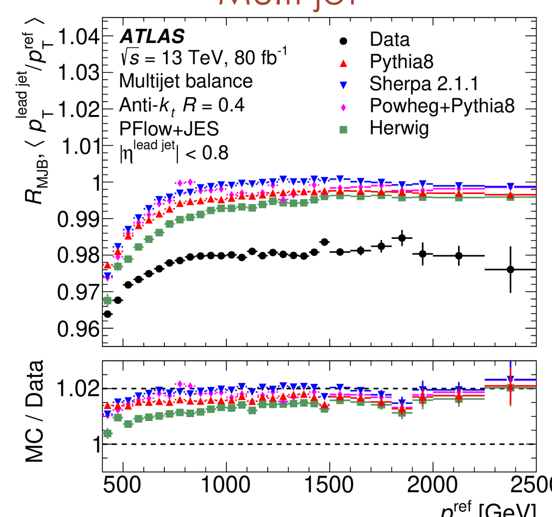
$Z + \mathrm{jet}$



$\gamma + \mathrm{jet}$



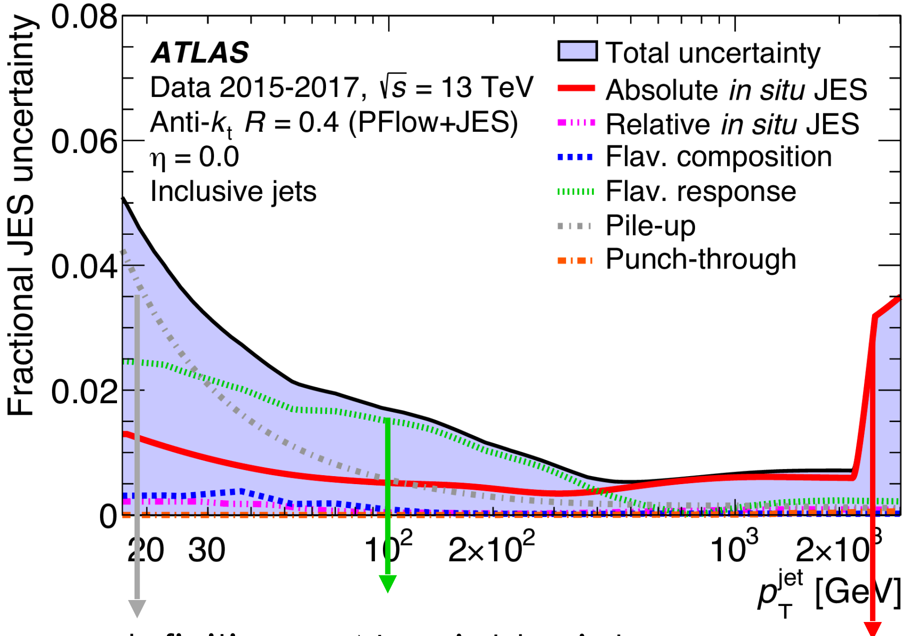
Multi-jet

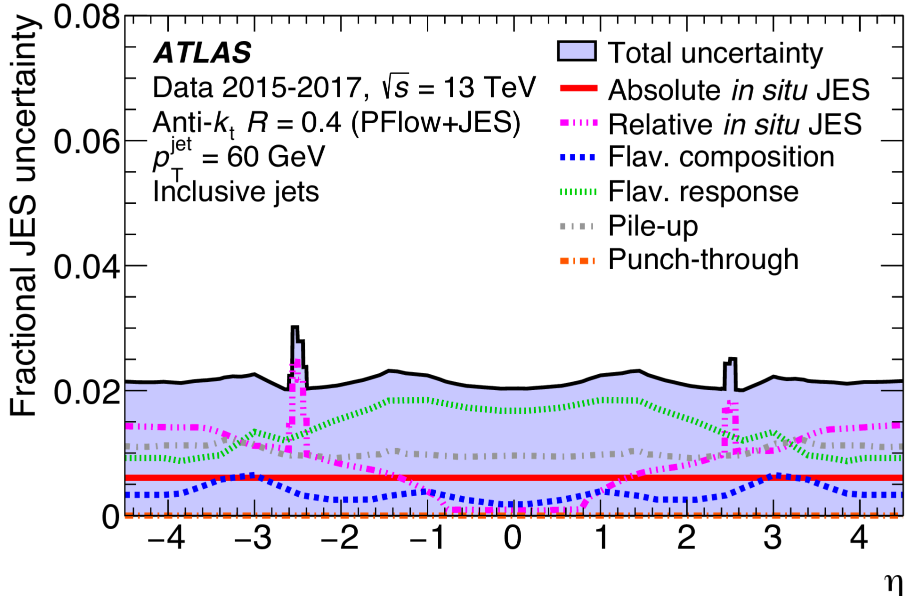


SYSTEMATIC UNCERTAINTIES

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

- Enormous effort from the collaboration to calibrate jet energy and mass with uncertainties below 2% at $p_{\mathrm{T}} \approx 200$ GeV





New ρ definition for PFlow (not using charged PFlow close to Prim. Vertex)

New jet-by-jet flavour uncertainty

Improving single particle uncertainties [Measurement of the energy response to charged pions from W → τ + ντ]

LARGE-R JET CALIBRATION

Input objets to Jet Reconstruction → **Pile-up mitigation algorithms** → **Large-R Jet reconstruction** → **Ungroomed large-R jets** → **Jet grooming**

- Topoclusters
- UFO

- Cell noise suppression
- CS+SK

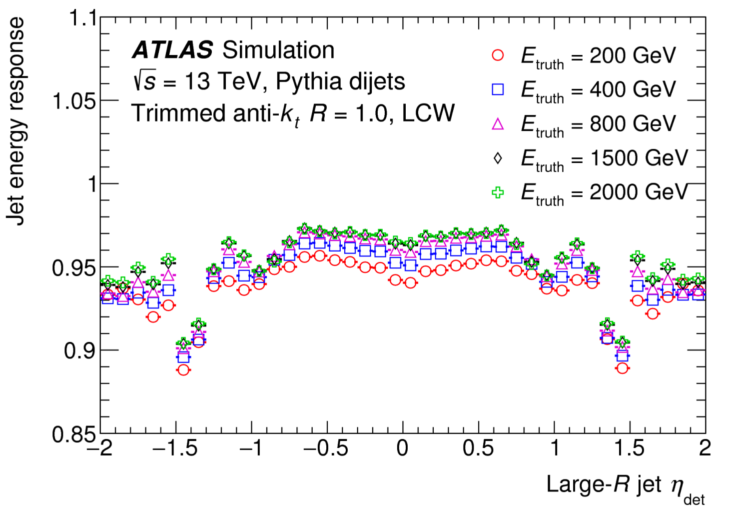
- Anti-kT R=1.0

- Trimming
- Soft drop

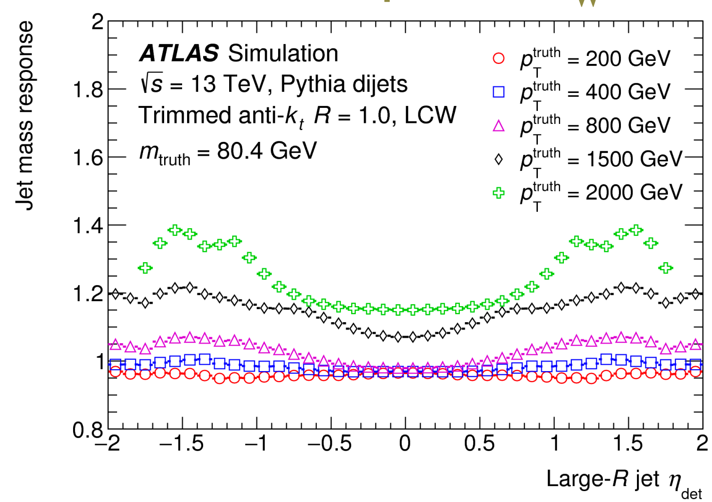
Groomed large-R Jets (LCW scale for topoclusters) → **E, η and m calibration** → **Residual *in situ* Calibration** → **Groomed large-R Jets (LCW (topoclusters)+ JES+JMS)**

MC-based calibration for topoclusters:

Jet Energy Response



Jet Mass Response ~m_W



MISSING TRANSVERSE MOMENTUM

MET RECONSTRUCTION

Object-based Missing Transverse Momentum (MET)

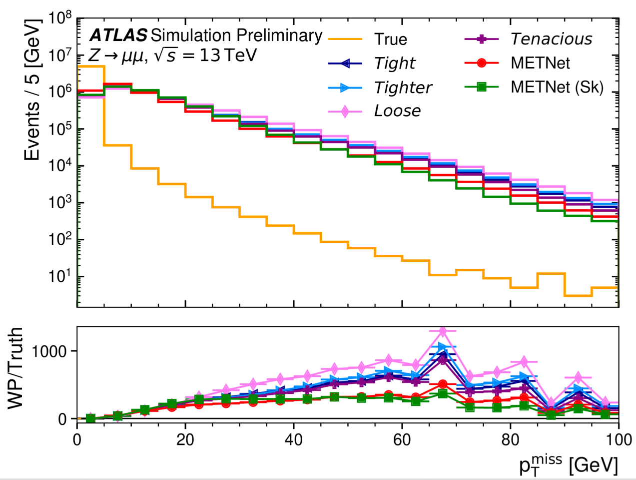
$$\sum_{\text{all}} \vec{p}_{\mathrm{T},i} = 0 \quad \xrightarrow{\text{When imbalance}} \quad \sum_{\text{observable}} \vec{p}_{\mathrm{T},i} + \vec{E}_{\mathrm{T}}^{\text{miss}} = 0 \quad \longrightarrow \quad \boxed{\vec{E}_{\mathrm{T}}^{\text{miss}} = -\sum_{\text{observable}} \vec{p}_{\mathrm{T},i}}$$

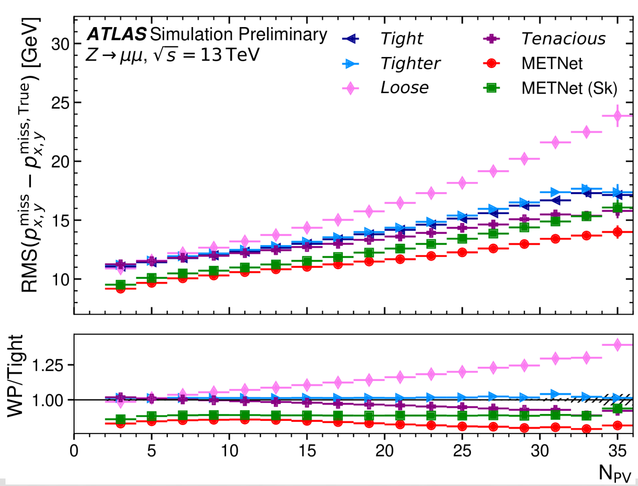
$$\mathbf{E}_{\mathrm{T}}^{\text{miss}} = \underbrace{\underbrace{-\sum_{\substack{\text{selected}\\\text{electrons}}} \mathbf{p}_{\mathrm{T}}^{e}}_{\mathbf{E}_{\mathrm{T}}^{\text{miss},e}} \underbrace{-\sum_{\substack{\text{accepted}\\\text{photons}}} \mathbf{p}_{\mathrm{T}}^{\gamma}}_{\mathbf{E}_{\mathrm{T}}^{\text{miss},\gamma}} \underbrace{-\sum_{\substack{\text{accepted}\\\tau\text{-leptons}}} \mathbf{p}_{\mathrm{T}}^{\tau_{\text{had}}}}_{\mathbf{E}_{\mathrm{T}}^{\text{miss},\tau_{\text{had}}}} \underbrace{-\sum_{\substack{\text{selected}\\\text{muons}}} \mathbf{p}_{\mathrm{T}}^{\mu}}_{\mathbf{E}_{\mathrm{T}}^{\text{miss},\mu}} \underbrace{-\sum_{\substack{\text{accepted}\\\text{jets}}} \mathbf{p}_{\mathrm{T}}^{\text{jet}}}_{\mathbf{E}_{\mathrm{T}}^{\text{miss},\text{jet}}}}_{\text{hard term}} \underbrace{\underbrace{-\sum_{\substack{\text{unused}\\\text{tracks}}} \mathbf{p}_{\mathrm{T}}^{\text{track}}}_{\mathbf{E}_{\mathrm{T}}^{\text{miss},\text{soft}}}}_{\text{soft term}}$$

Machine Learning-based MET: METNet

ATL-PHYS-PUB-2021-025

- Neural Network takes different MET Working Points as inputs (different requirements on the Jets)
- METNet do a regression targeting Truth MET





Improvements in resolution

MET SIGNIFICANCE

How likely is the measured MET to be real or a resolution effect?

[ATL-PHYS-PUB-2021-025]

Object-based MET Significance

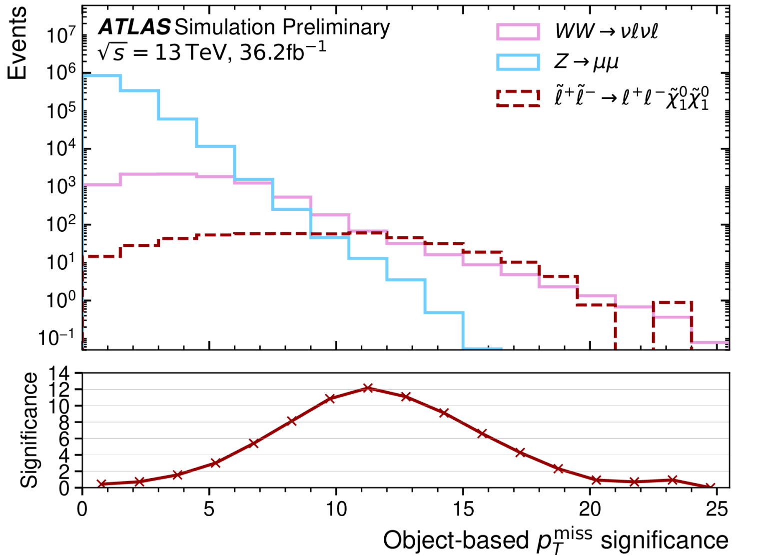
$$S^2 = \frac{\left|\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}\right|^2}{\sigma_L^2\left(1-\rho_{LT}^2\right)}$$

Total longitudinal variance

Correlation factor

Expected resolutions from objects

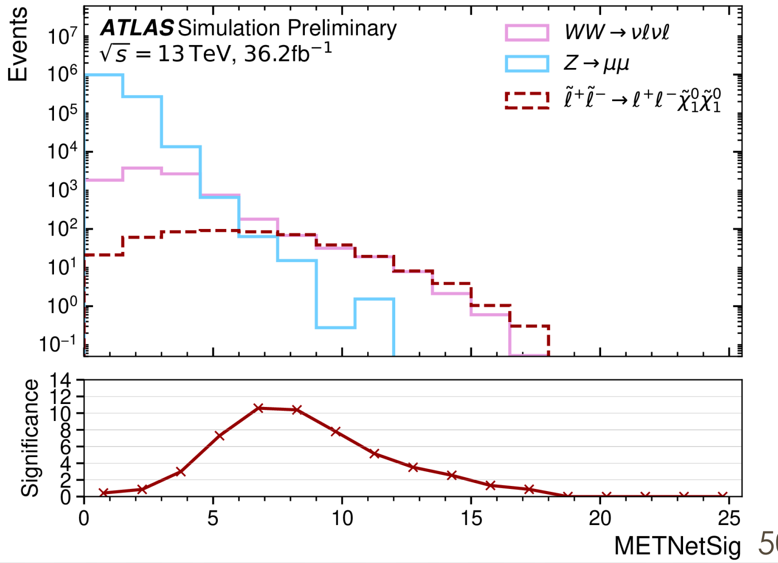
$$\sigma_L^2 = {\sigma_L^{hard}}^2 + {\sigma_L^{soft}}^2$$



Machine Learning-based MET Significance: METNetSig

$$\mathrm{METNetSig} = \frac{\mathrm{METNet}}{\sigma_{\mathrm{MetNet}}}$$

Neural Network's confidence in the METNet prediction



CONCLUSIONS AND OUTLOOK

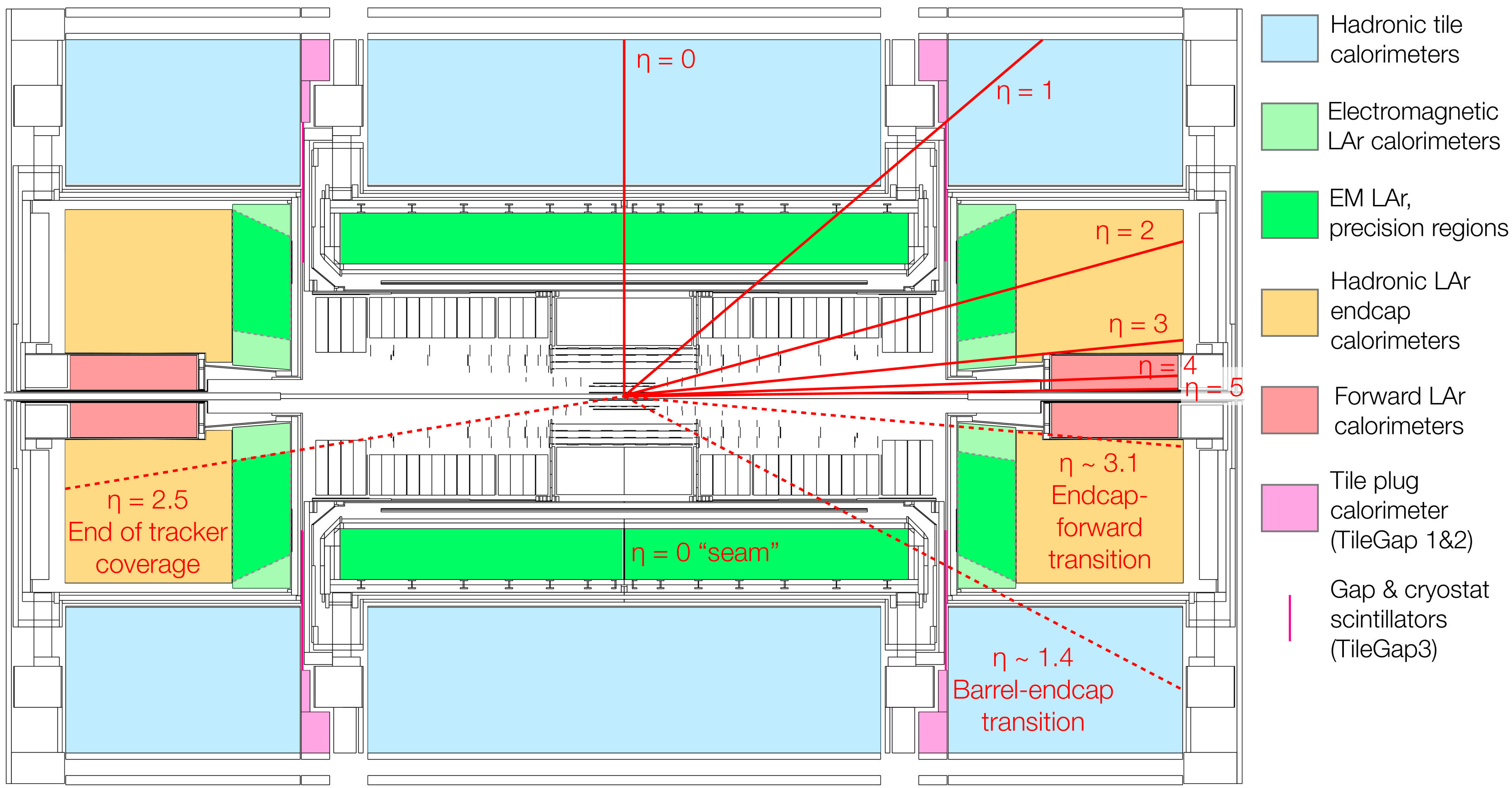
- Jet constituents
  - Great effort in defining jet constituents that combine measurements from the tracker and calorimeters: PFlow, TCC, UFO
  - Promising studies revisiting LCW calibration using modern ML-based approaches
- Jet Reconstruction & performance
  - ATLAS performed a Large-R jet definition optimisation: CS+SK UFO jets with Soft Drop outperform other jet collections in different metrics -> will be the default jet definition in future ATLAS analysis
  - New W/Z and top taggers with UFO jets developed showing improvements
  - Detailed understanding of jet energy calibration requires many detailed studies to converge!
- Missing Transverse Momentum
  - New ML-based MET showing improved resolution

- More improvements will come In the future
- Calibration of jet constituents: LCW for PFlow and UFO, ML approach
- Topo-tower inputs in the forward region
- Simultaneous energy and mass calibration using Neural Networks
- small-R UFO jets
- Topo-clustering time cut for rejecting pile-up
- …

- Inputs
- Calibration
- UFO taggers
- MET

BONUS

ATLAS BASICS

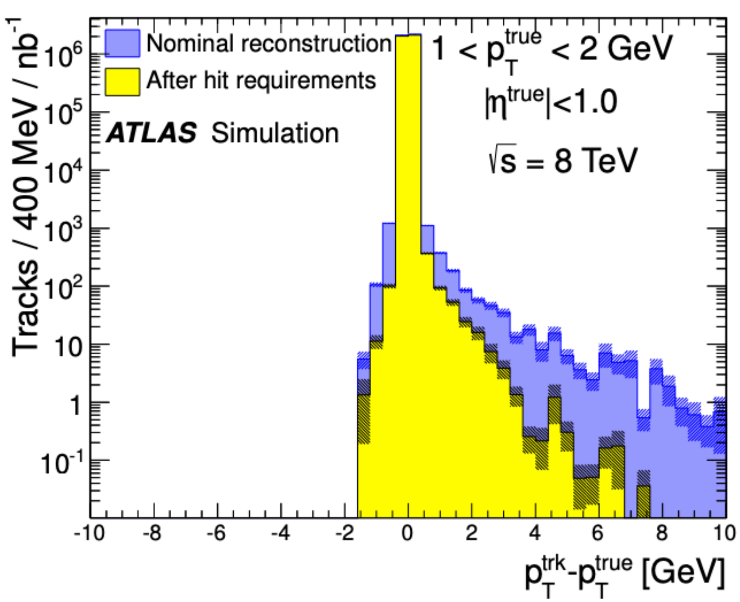


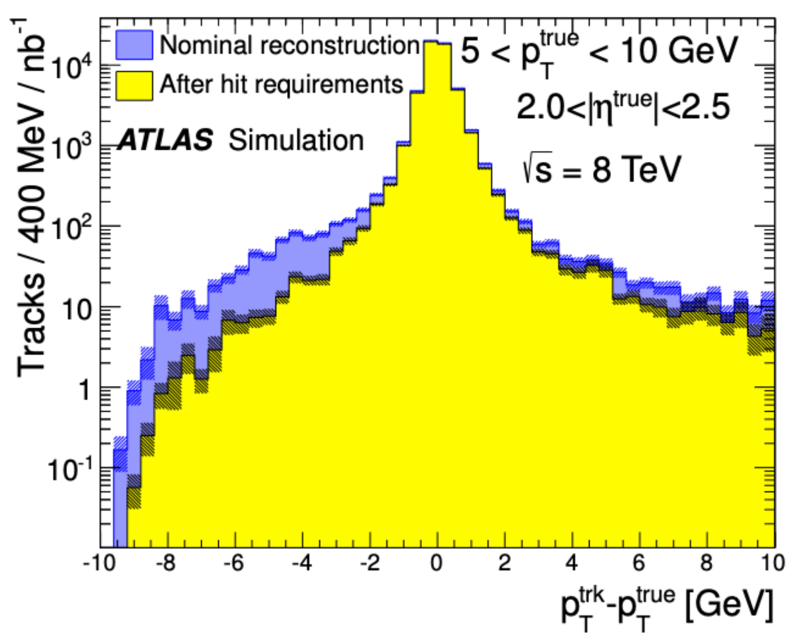
INPUTS + PU

PFLOW: INPUTS

Tracks with a reliable measurement, and where the resolution is better than the calorimeter.
- "tight" tracking criteria: ≥ 9 silicon hits, no pixel holes, $|\eta| < 2.5$ and $0.5 <$ **pT < 40 GeV**
- Tracks associated with electrons or muons are not considered by the particle flow algorithm

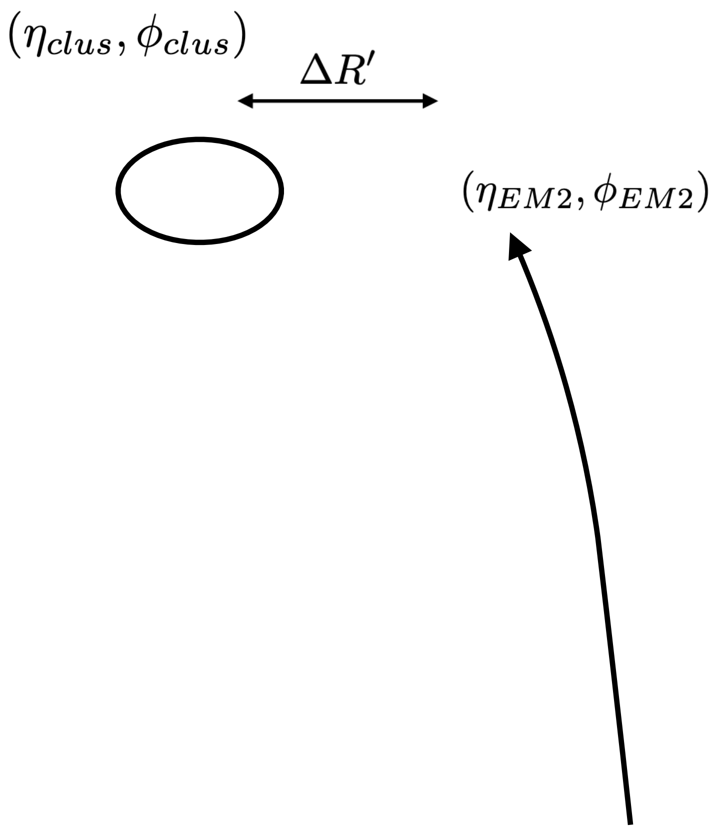
All **topoclusters** are considered

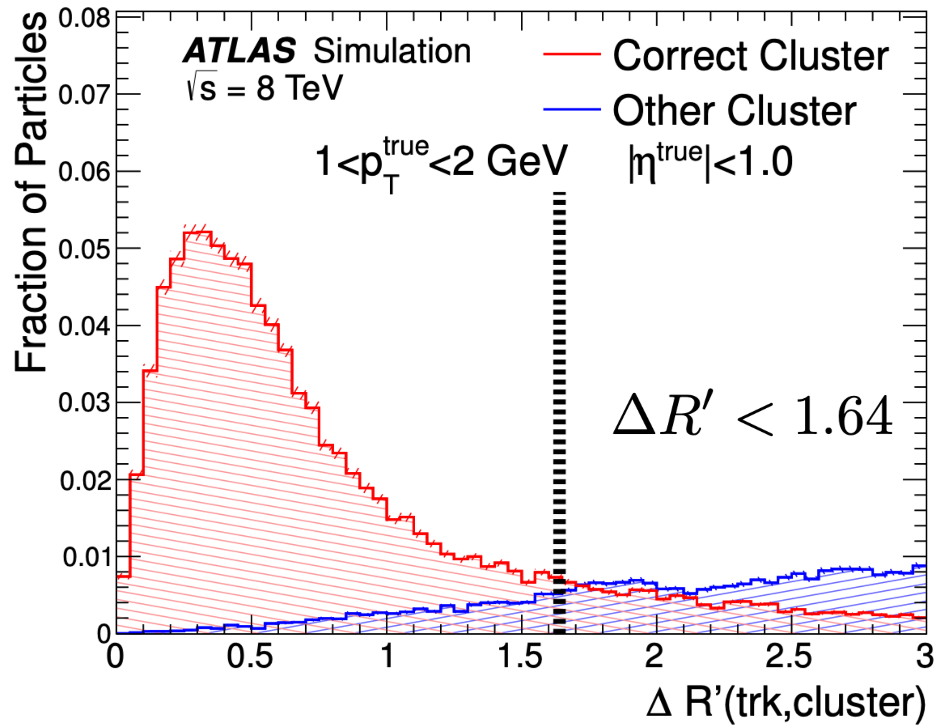




Output:
- **List of tracks**
- **List of Topoclusters**
  - Unmodified Topoclusters
  - New Topoclusters resulting from the energy subtraction

PFLOW: TRACK-CLUSTER ASSOCIATION





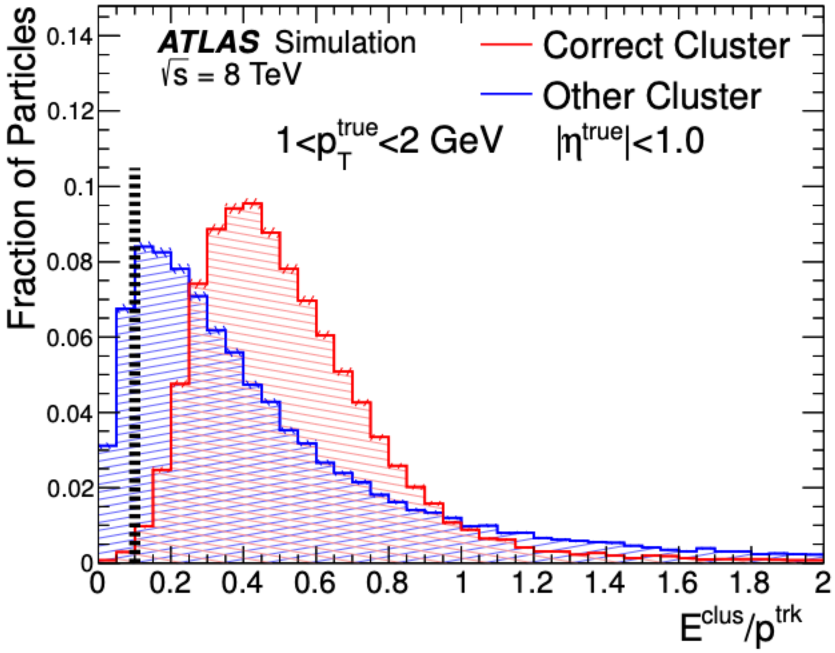
$$\Delta R' = \sqrt{\left(\frac{\eta_{EM2} - \eta_{clus}}{\sigma_\eta^{clus}}\right)^2 + \left(\frac{\phi_{EM2} - \phi_{clus}}{\sigma_\phi^{clus}}\right)^2}$$

σ: Represent the angular toPo-cluster widths.

- The topocluster with the smallest ΔR ′ and E > 0.1p is taken.
- If no topocluster with DeltaR<1.64: particle does not form a topocluster. Retain track

PFLOW: E/P MEASUREMENT

E/p is defined using all topoclusters in a cone of 0.2 around the track extrapolated coordinates in the second layer of the calorimeter.



Selected topocluster to be matched

"Correct" has at least the 90% of the truth energy from the particle matched to the track

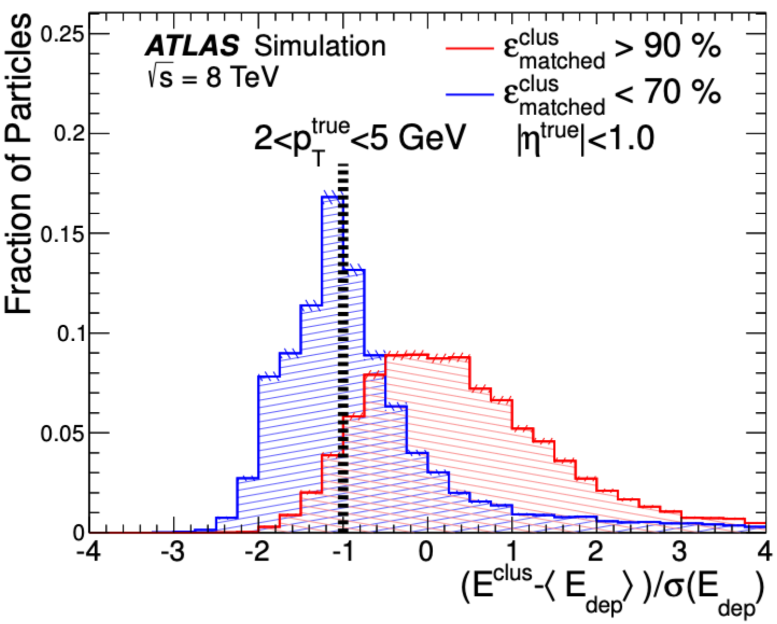
PFLOW: RECOVERING SPLIT SHOWERS

Often particles split their energy across **several clusters**..

Initial track-topocluster match has an **E/P smaller than expected**

The discriminant used to **distinguish** the **single and multiple** topo-cluster cases is the significance of the difference between the expected energy and that of the matched topo-cluster

$$S(E^{\text{clus}}) = \frac{E^{\text{clus}} - \langle E_{\text{dep}} \rangle}{\sigma(E_{\text{dep}})}$$

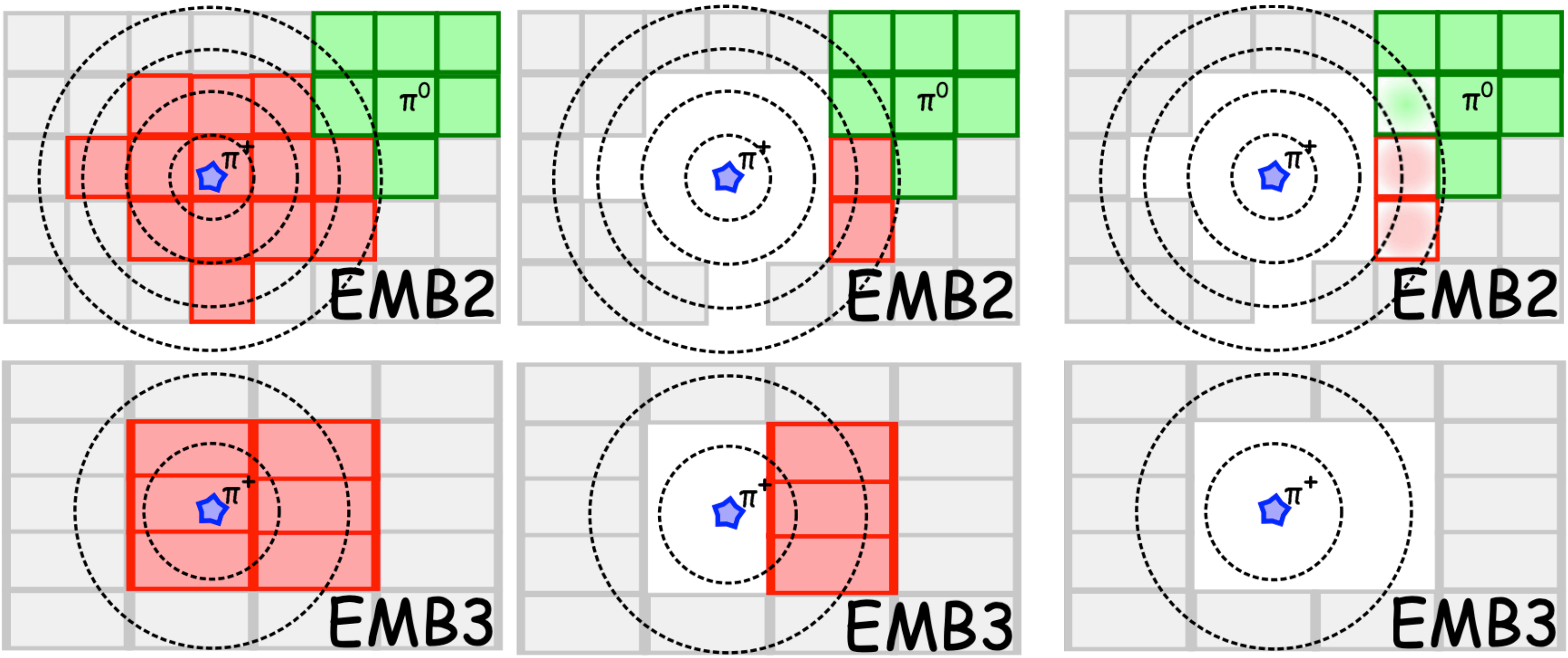


Look for additional clusters if S<-1

Take all clusters in a cone of radius 0.2 around the extrapolated track

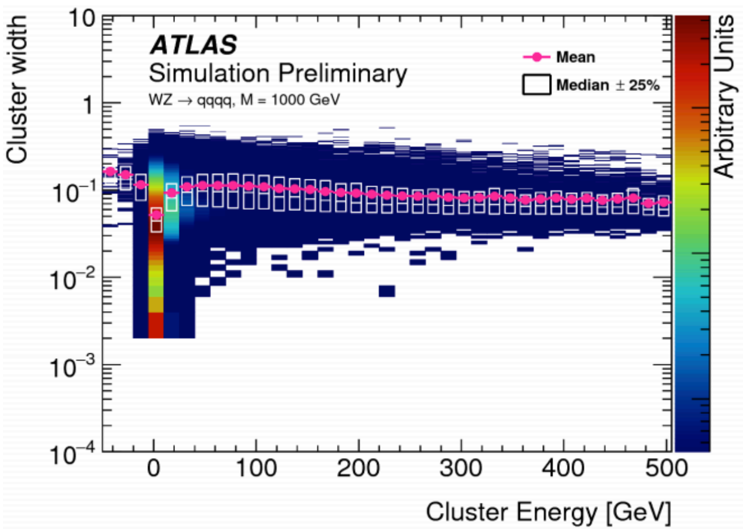
CELL-BY-CELL SUBTRACTION

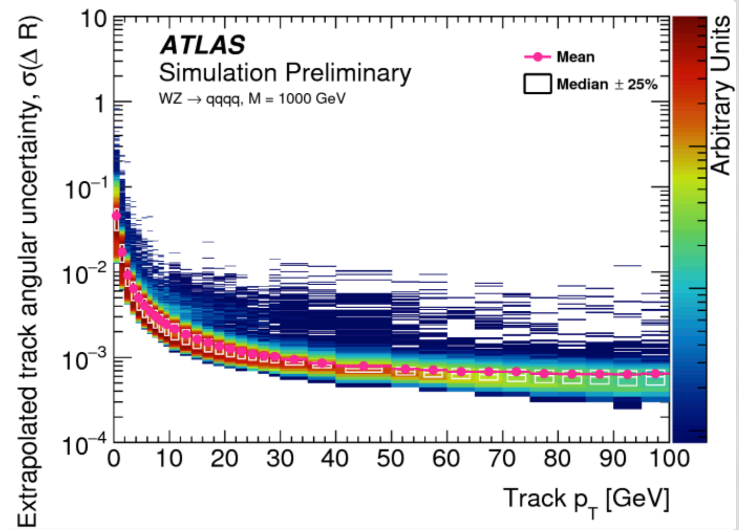
Subtract cell-by-cell if $E_{cl} > p_{trk} \times (E/p)$
Each layer is split into rings around the extrapolated track



TRACKS IN TCC

- The track position is known very precisely
- the uncertainty on its extrapolation to the calorimeter is smaller than the width of a cluster at high pT.





- The matching procedure does not attempt to remove tracks from leptons
- Use tracks originating from any primary vertex (not only HS PV).

Matching:

- 1. Attempt match if σtrack < σcluster (width)
- 2. Match if (ΔR)2 < (σtrack) 2 + (σcluster) 2

PERFORMANCE: JET CONSTITUENT CALIBRATION

Hadronic Calibration with Particle Flow Networks (PFN)

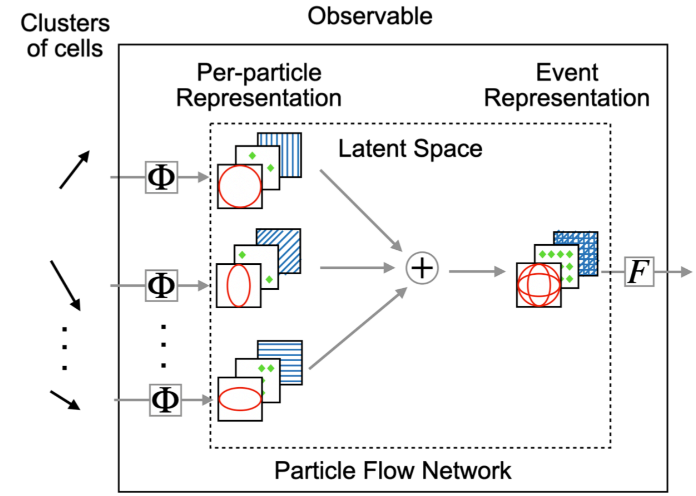
- An event can be seen as a variable-length unordered set of particles
- **Deep Sets Theorem**: A permutation invariant Observable of sets of particles can be parametrised in a general way:

$$\mathcal{O}(\{p_1, ..., p_M\}) = F\left(\sum_{i=1}^{M} \Phi(p_i)\right)$$

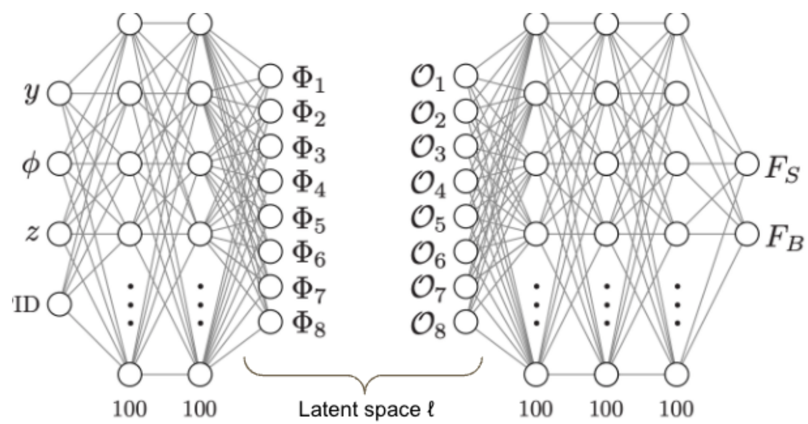
  - Applied to pion classification and energy calibration:

$$\text{PFN} : \mathcal{O} = \sum_{i=1}^{M} \Phi(E, \eta_i, \phi_i, \text{SamplerID})$$

Approximate Φ and F with Neural Networks



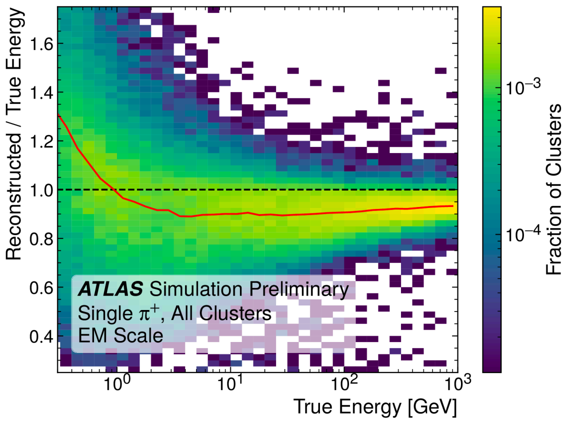
- Use Φ networks to process features in three layers
- Put outputs together with an "**aggregation function**"
- Aggregated outputs are processed by global network "**F**"
- **F output** can be used for **classification** or **regression**



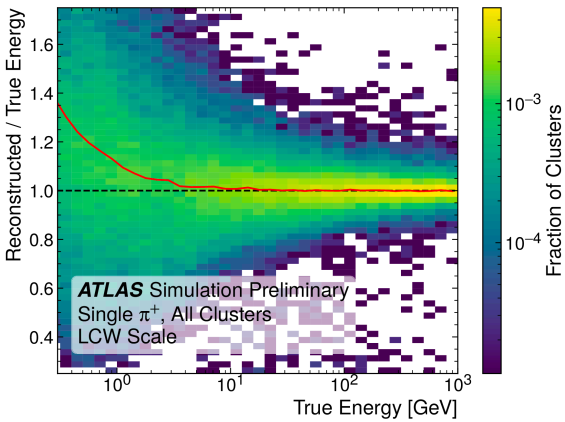
Pion Energy Calibration

- After identifying a cluster as hadronic/EM, need to convert the signal into an energy measurement
- Energy regression goal: Correctly predict the true energy deposited in the cluster.
  - Quantified by measuring the cluster energy response: $R = \frac{E^{\text{reco}}}{E^{\text{truth}}}$ that should be ~ 1

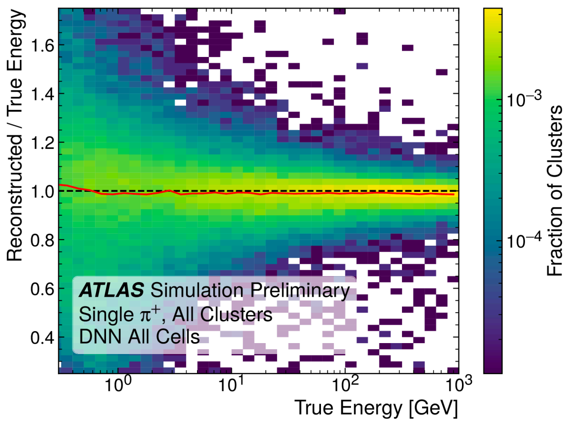
Regression performance for charged pions



Raw "EM" scale under-estimates R

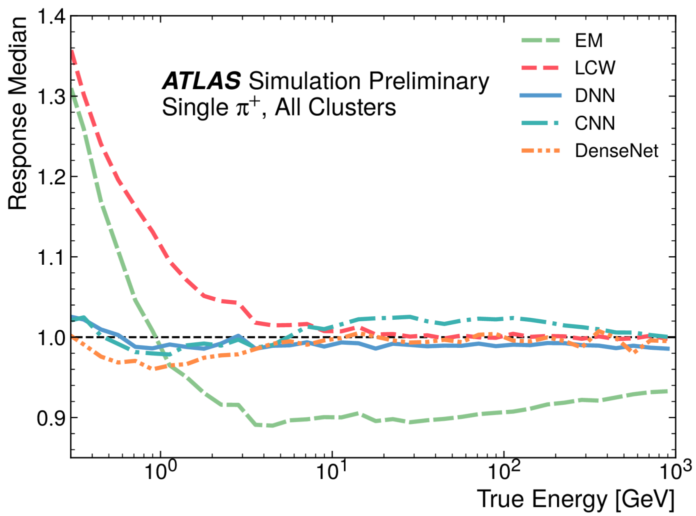


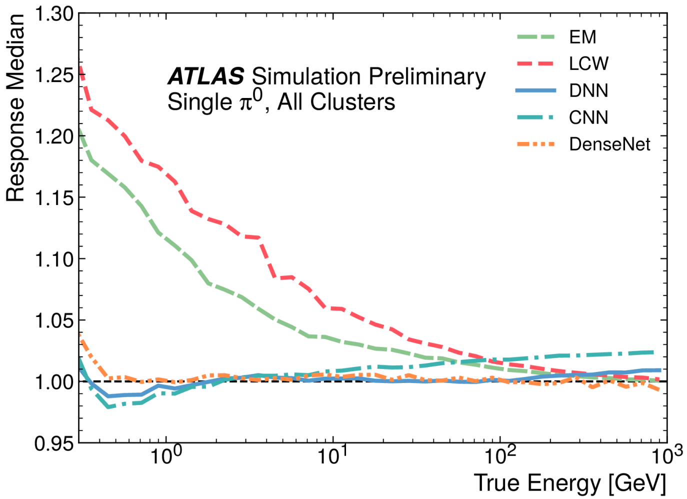
LCW over-estimates R at low-energy

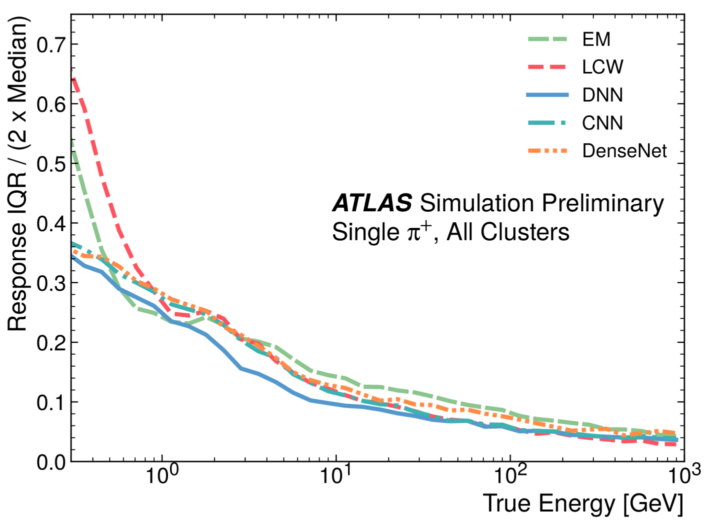


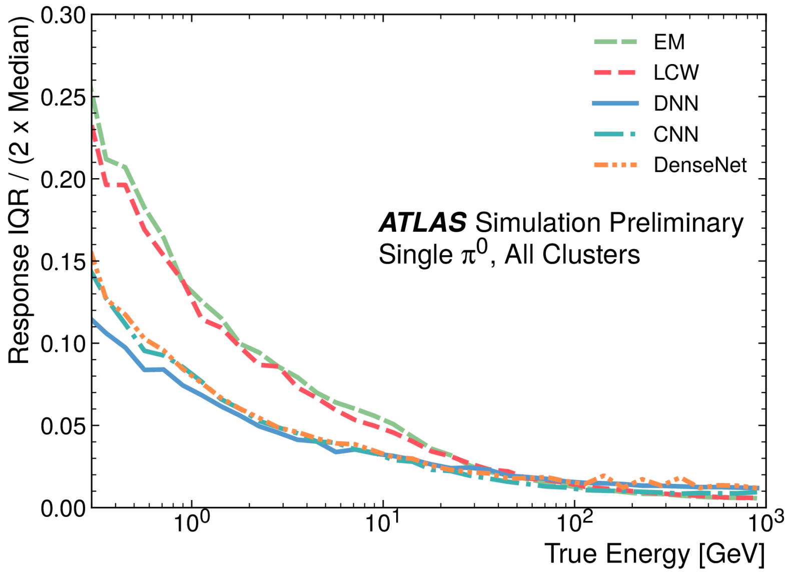
DNN regression does an excellent job nearly everywhere

Pion Energy Calibration

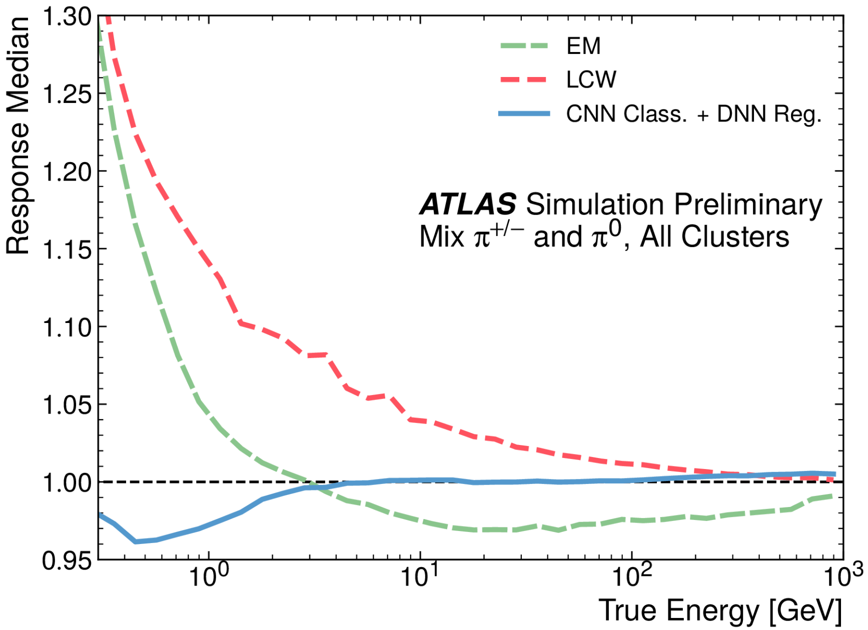


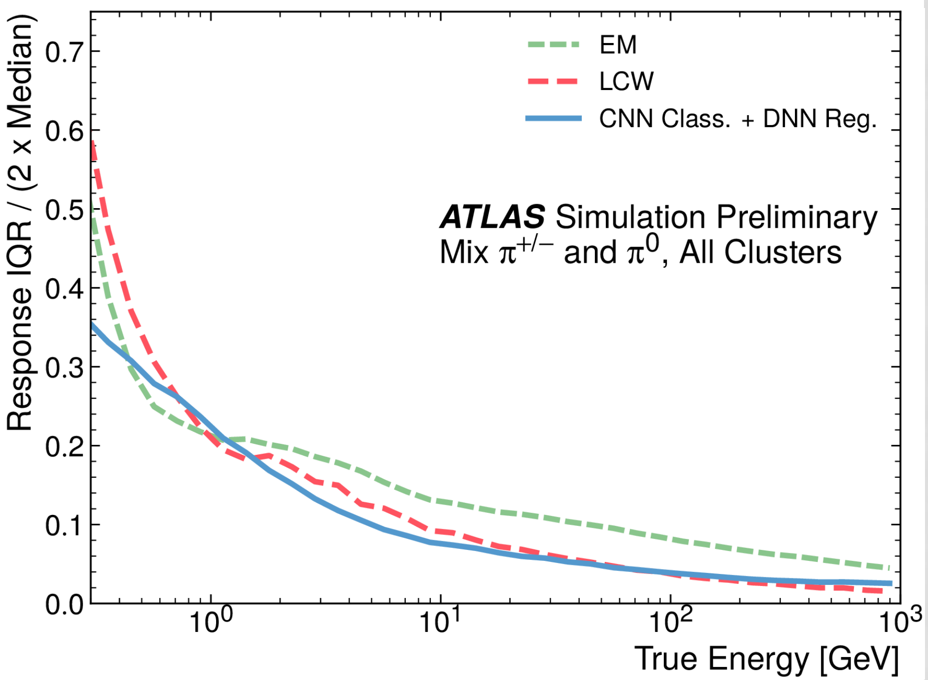






Mixed sample of π ± and π0

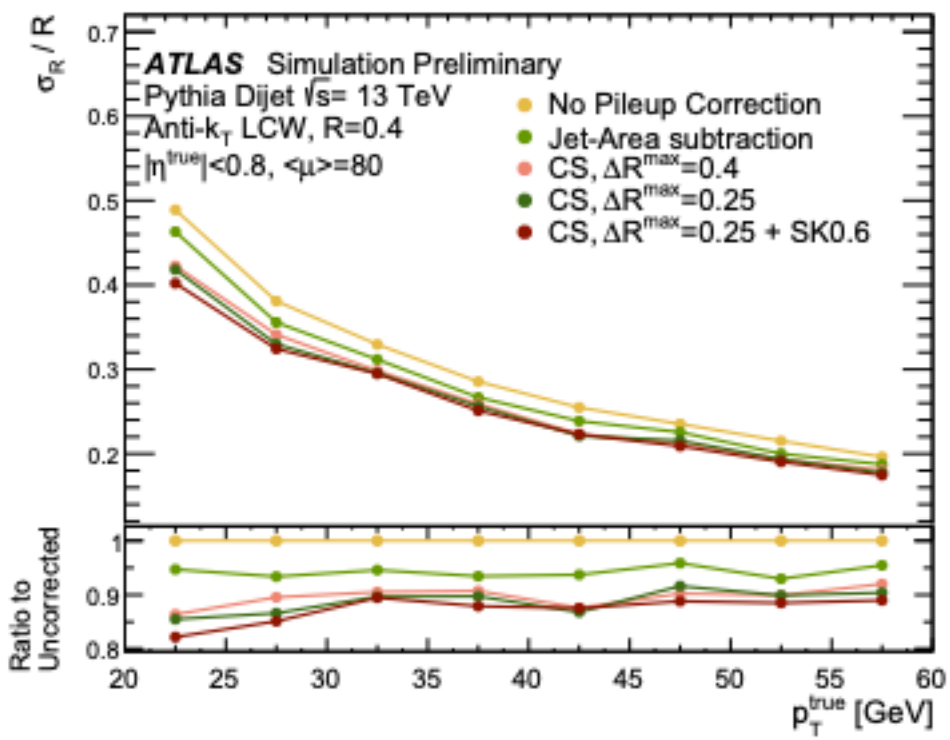




CONSTITUENT SUBTRACTION (CS)

JHEP 1406 (2014) 092

- A "constituent area" subtraction
- Add ghosts to the event with $p_T^g = A_g \times \rho$
- Subtract ghost's contribution from p_T of closest constituent (in ΔR)
  - Until ΔR(ghost, constituent) > ΔR_{max}



ATLAS-CONF-2017-065

VONOROI SUBTRACTION

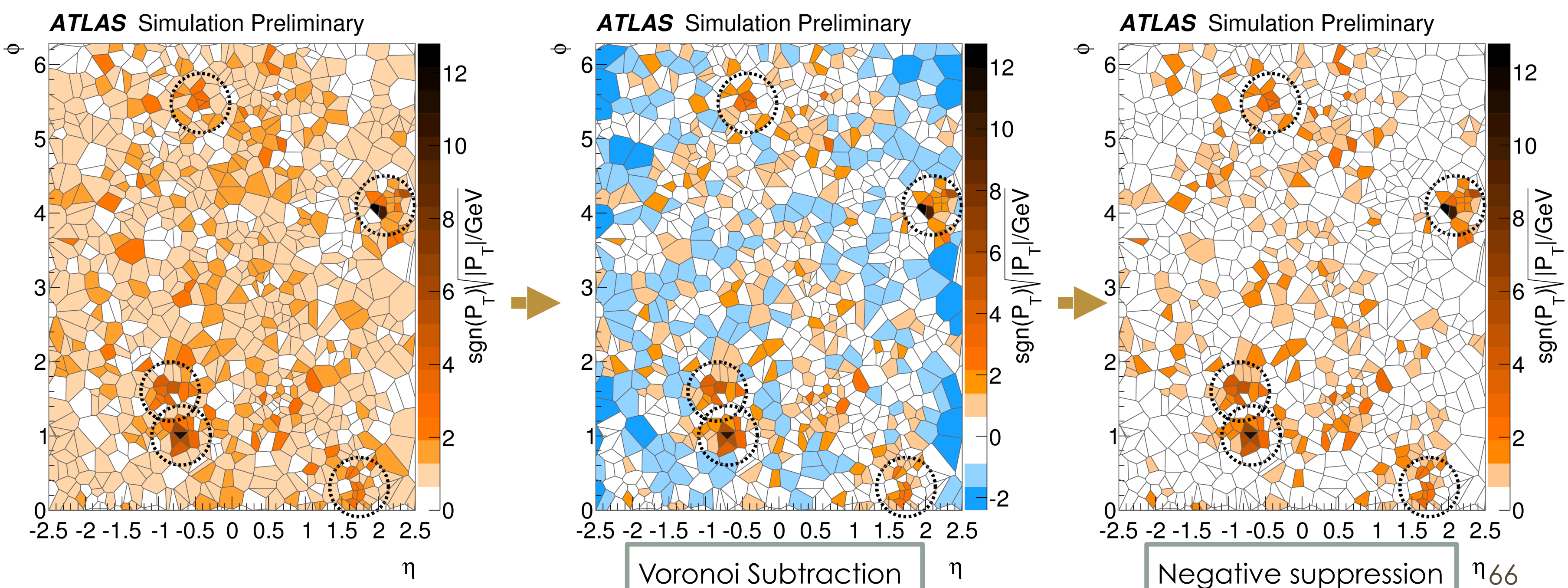
arXiv:1703.09665

- Subtract average PU contribution as a 'constituent area' correction:

$$p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}} - \rho \times A_{\mathrm{Voronoi}}$$

  - Vonoroi Area: All points in η-φ space closer to that constituent than any other
  - ρ: Median energy density

- Different options for handling soft clusters in which p_T^{corr} turns out to be negative
  - Negative suppression: Remove clusters with pT<0
  - 1-σ suppression: Remove clusters with $p_T < 1 - \sigma\sqrt{A}$
  - Spreading: Spread some negative pT to nearby clusters with positive pT



SOFT DROP

$$\underbrace{\frac{\min(p_{T1}, p_{T2})}{p_{T1}+p_{T2}}} > z_{cut}\underbrace{\left(\frac{\Delta R_{12}}{R_0}\right)}^{\beta}$$

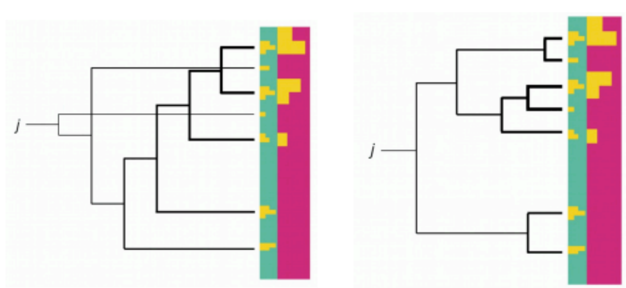
p_T fraction must be sufficiently large

constituents must be close together

Keep both constituents → STOP

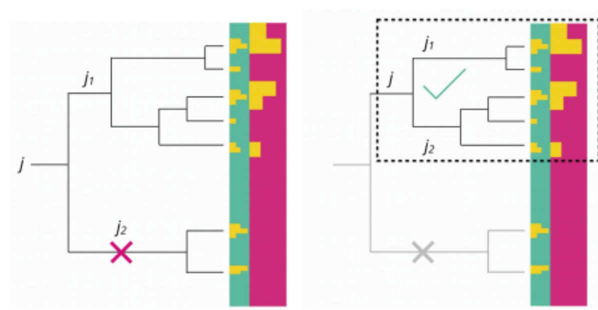
"Drop" constituent → Iterate on next splitting

- Jet type
  - *Create* jets using anti-k_t
    - Cluster hardest stuff first
  - *Define splittings* using Cambridge-Aachen
    - Closest stuff first



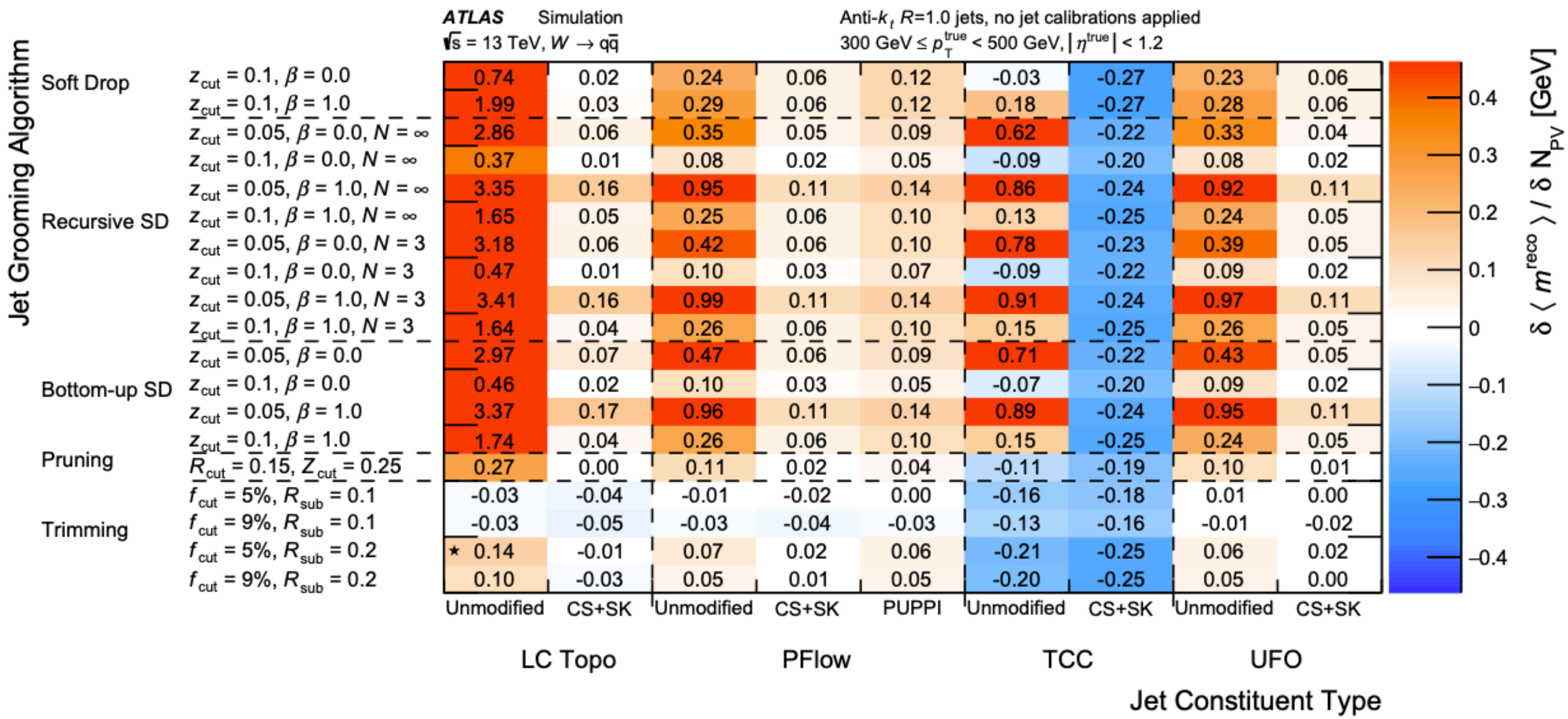
Hard stuff first → Close stuff first

- Apply SoftDrop
  - Start from first jet splitting
  - Reach passing condition for both splittings
  - STOP!



First fails Second passes 18

PILE-UP STABILITY

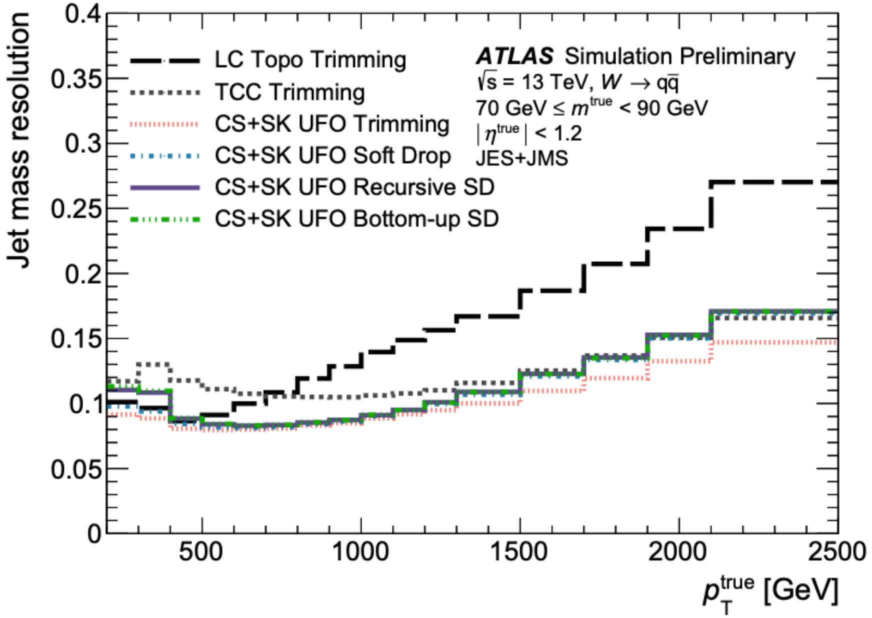


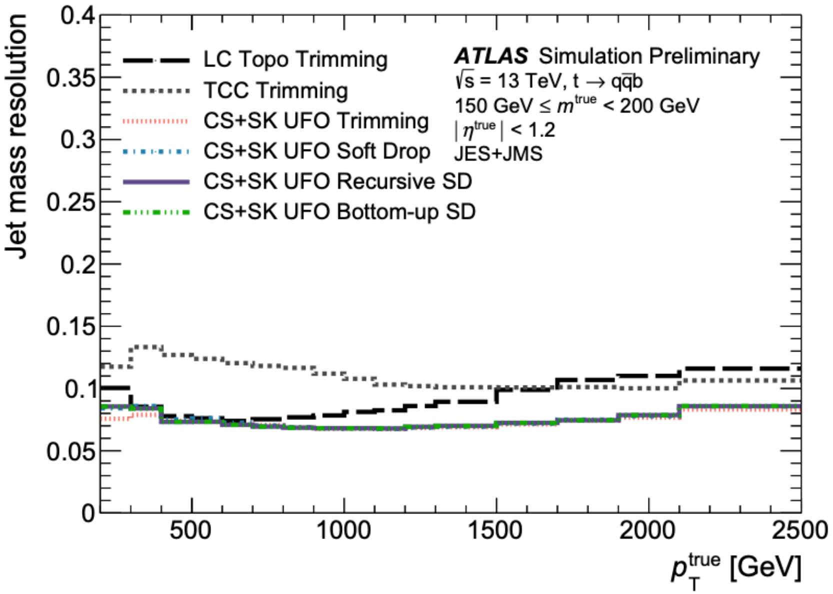
LARGE-RADIUS JETS OPTIMISATION

Jet Mass Resolution

Resolution of the mass Response= $\frac{Mass^{reco}}{Mass^{truth}}$

- Jet definition scan: Tested different jet input type, pileup mitigation, and grooming algorithm
- Finalists:



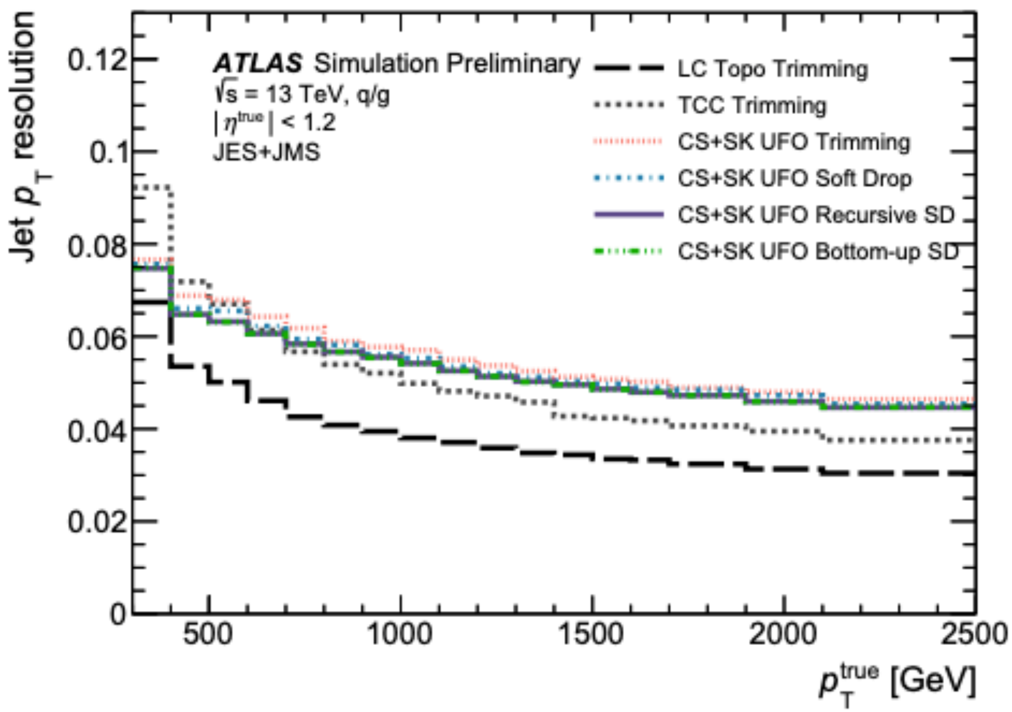


- Winner: CS+SK UFO Soft Drop
  Seems to be working excellent in all of the studied metrics [Eur. Phys. J. C 81 (2021) 334]

LARGE-RADIUS JETS OPTIMISATION

Jet energy resolution

- UFO, CS+SK, Soft Drop seems to be working excellent in all of the metrics - However, there is one downside to UFO UFO uses “EM-scale” topo-clusters
- Degraded JER compared to LC clusters
- All of the improvements make the new definition very useful for many analyses However, still can be improved... An **LC UFO variant**
- could improve JER Could also be mitigated via **dedicated GSC**

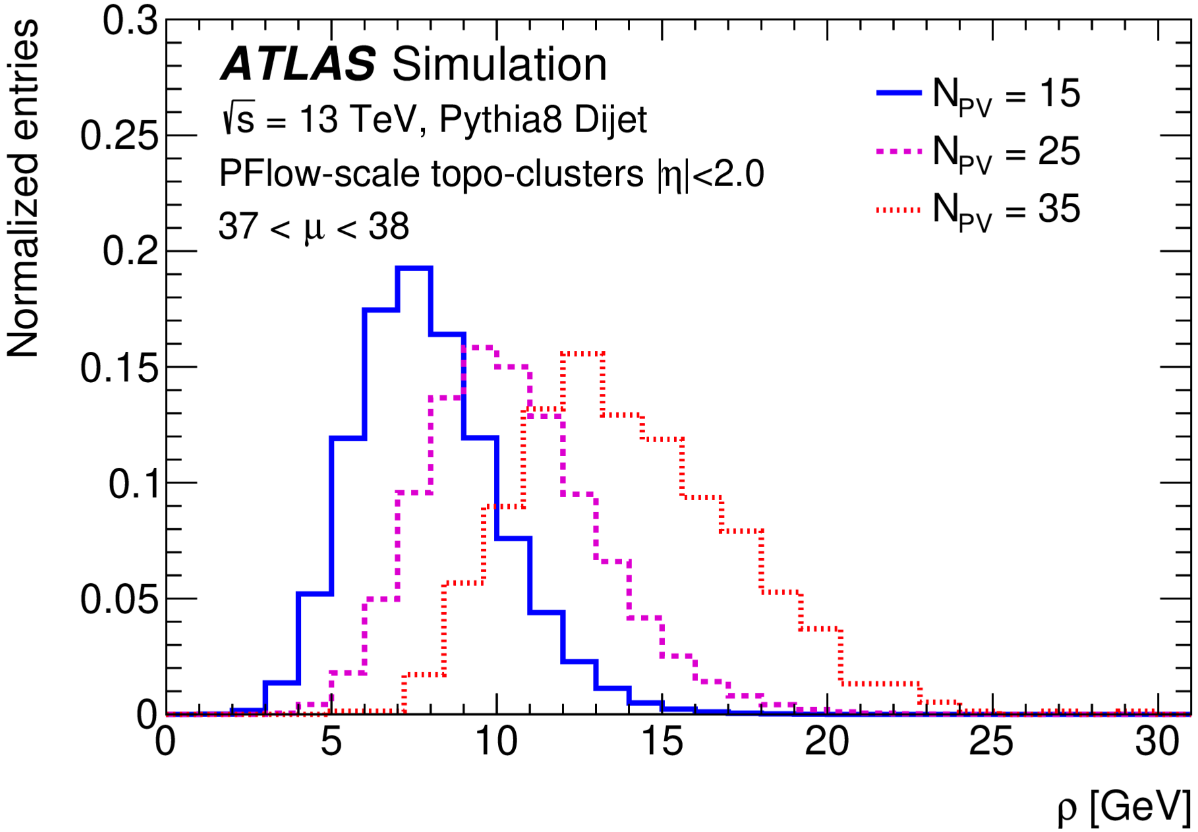


CALIBRATION

PILE UP CORRECTION

ρ = median pT density of jets in the y–ϕ plane:

ρ = <pT/A> of kt 0.4 jets with pT ≥ 0 clustered from topo-clusters within η < 2



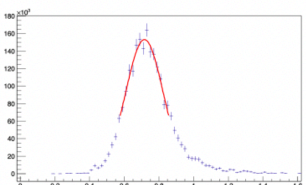
$$\rho_{jet} = \frac{p_{T,jet}}{A_{jet}}$$

MC-JES

Get $C(E^{reco}) = E^{true}$ by first deriving $C^{-1}(E^{true}) = E^{NI}$ (Numerical Inversion)

Later P-A: tested using p_T & y bins

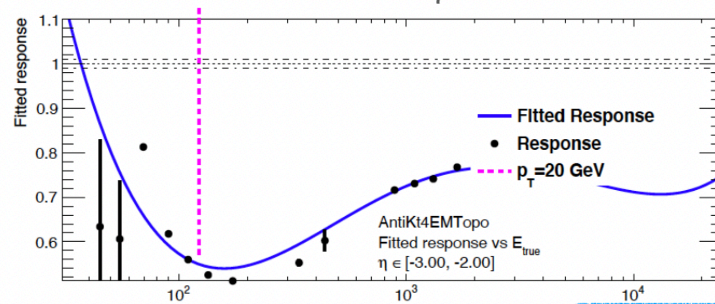
#events



$$r = \frac{E^{reco}}{E^{true}}$$

- Plot **individual response** r in bins of (E^{true}, η)
- Extract **average** R from the **Gaussian** $P(E^{reco} \mid E^{true})$ [2]

R vs E^{true} in η bins



E^{true} [GeV]

New p-spline method by Arthur Linß

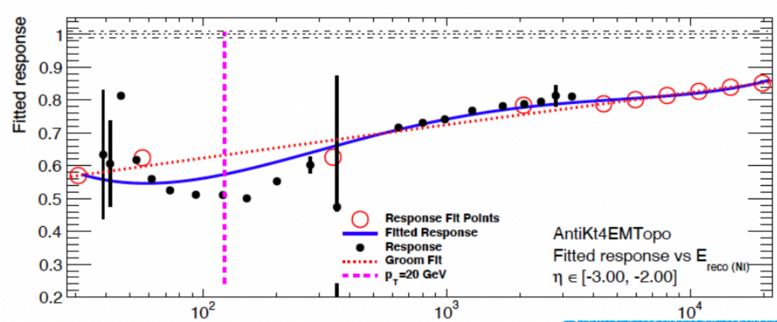
Apply fit

$$R_1(E^{true}) = \sum_{i=0}^{8} \alpha_i ln(E^{true})^i$$

$R_1(E^{true})$ **from fit 1**

$$E^{NI} = C^{-1}(E^{true})$$
$$= R_1(E^{true}) \times E^{true}$$

R vs $E^{NI\ (reco)}$ in η bins

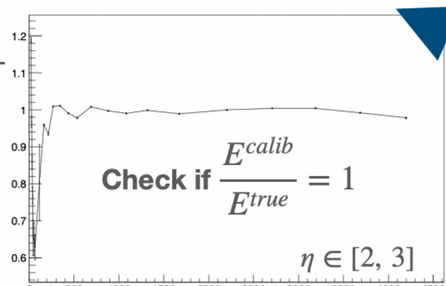


$E^{NI\ (reco)}$ [GeV]

$$E^{calib} = C(E^{reco}) = \frac{E^{reco}}{R_2(E^{NI})}$$

$R_2(E^{NI})$ **from fit 2**

$\frac{E^{calib}}{E^{true}}$



Check if $\frac{E^{calib}}{E^{true}} = 1$

$\eta \in [2, 3]$

E^{true} [GeV]

IN SITU: η-INTERCALIBRATION

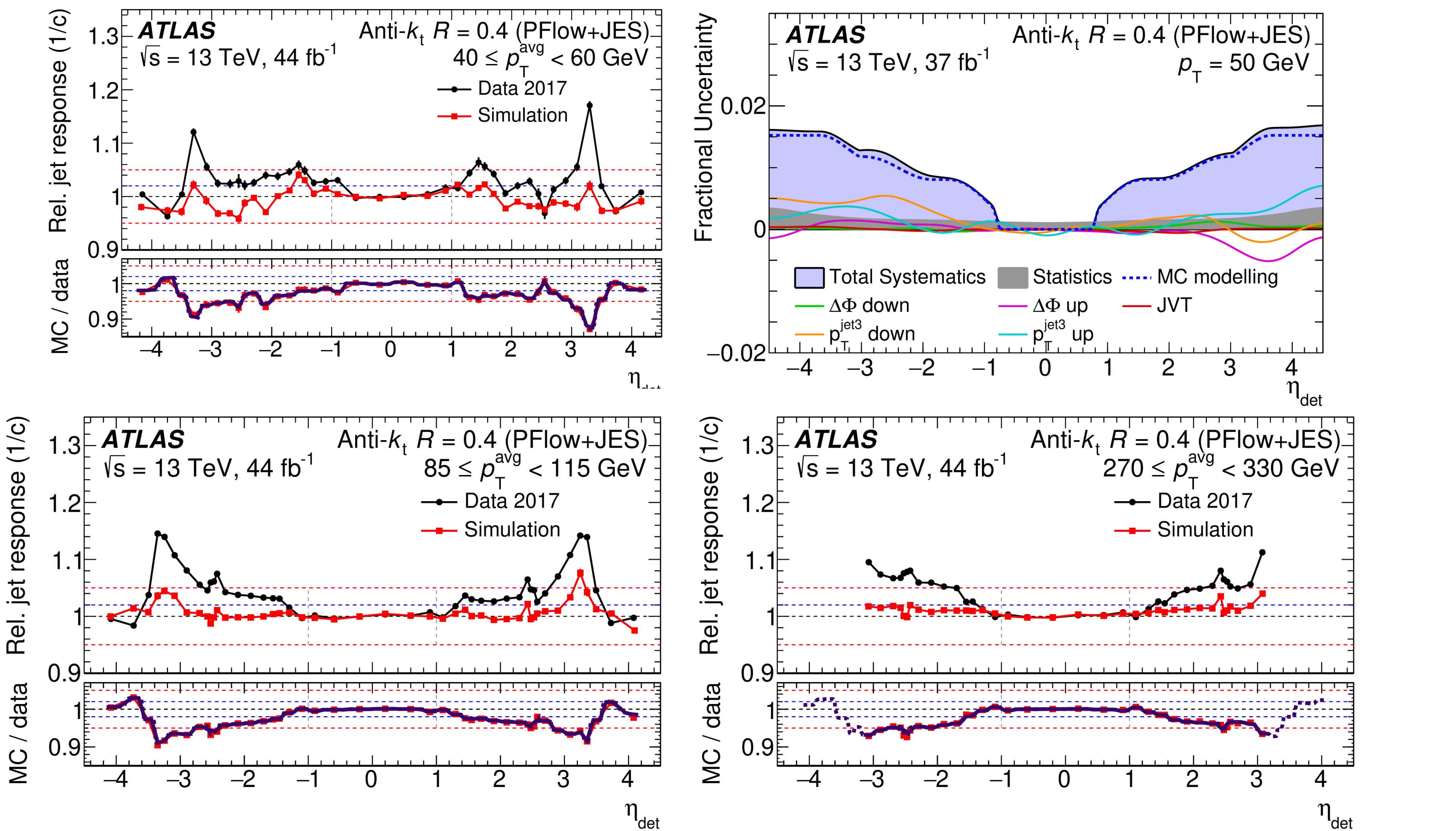
- Correction applied to forward jets ($0.8 \leq |\eta_{det}| < 4.5$) to bring them to the same energy scale as central jets ($|\eta_{det}| < 0.8$).
- All regions will be calibrated relative to one another
- Momentum balance requires that the p_T of the two jets must be equal and opposite
- Asymmetry:

$$\mathcal{A} = \frac{p_T^{\text{left}} - p_T^{\text{right}}}{p_T^{\text{avg}}}$$

- The response ratio defines the calibration factor per jet: $\mathcal{R} = \frac{c^{\text{left}}}{c^{\text{right}}} = \frac{2 + \langle \mathcal{A} \rangle}{2 - \langle \mathcal{A} \rangle} \cong \frac{p_T^{\text{left}}}{p_T^{\text{right}}}$
- calibration minimizes this function analytically: $S(c_{1x}, \ldots, c_{Nx}) = \sum_{j=2}^{N} \sum_{i=1}^{j-1} \left(\frac{1}{\Delta \langle \mathcal{R}_{ijx} \rangle} (c_{ix} \langle \mathcal{R}_{ijx} \rangle - c_{jx}) \right)^2 + X(c_{ix})$
  - Which can be expressed as a matrix system of linear equations.
  - Solved independently for each p_T avg bin x to obtain values for the correction factors c_{ix} for each η_{det} bin i in this momentum range.

The correction factors are normalized such that the average correction factor in the central region $|\eta_{det}| < 0.8$ is unity.

IN SITU: η-INTERCALIBRATION



IN SITU: Z+JET & γ+JET

Balance the **hadronic recoil** in an event against the pT of a calibrated Z **boson or photon**.

Only in central region (| η | < 0.8)

- **Missing-ET projection fraction (MPF)**: uses the full hadronic recoil related to reference
  - Help mitigate effects of pile-up and jet reconstruction threshold at low-pT [ATLAS-CONF-2015-057]
  - symmetry in transverse plane cancels signals from soft emissions

$$\vec{p}_{\mathrm{T,truth}}^{\,\mathrm{ref}} + \vec{p}_{\mathrm{T,truth}}^{\,\mathrm{recoil}} = 0$$

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

Response of had. Recoil

$r_{\mathrm{MPF}} < 1$

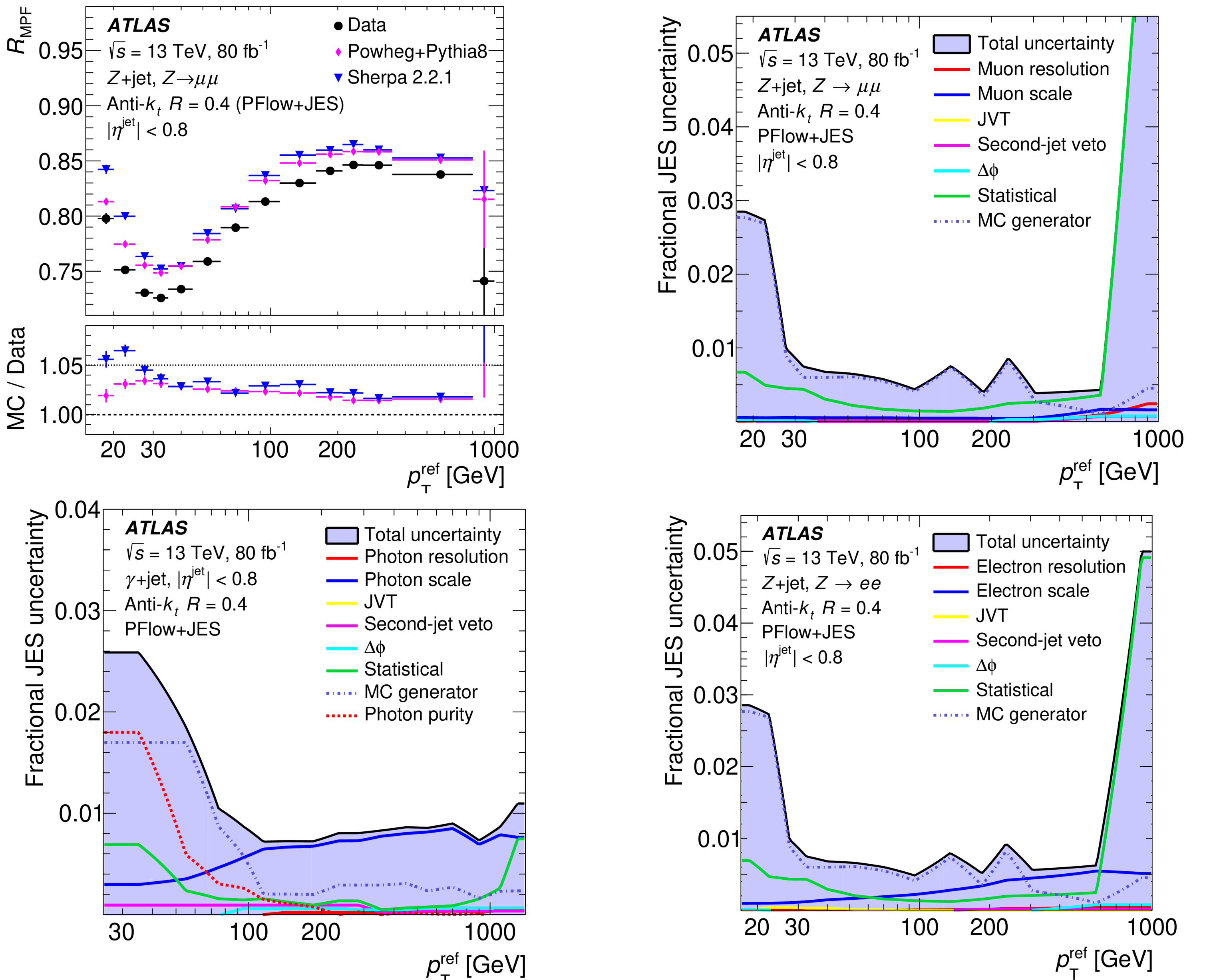
Taking the projection in the direction of the reference object:

$$\mathcal{R}_{\mathrm{MPF}} = \left\langle 1 + \frac{\hat{n}_{\mathrm{ref}} \cdot \vec{E}_{\mathrm{T}}^{\,\mathrm{miss}}}{p_{\mathrm{T}}^{\mathrm{ref}}} \right\rangle$$

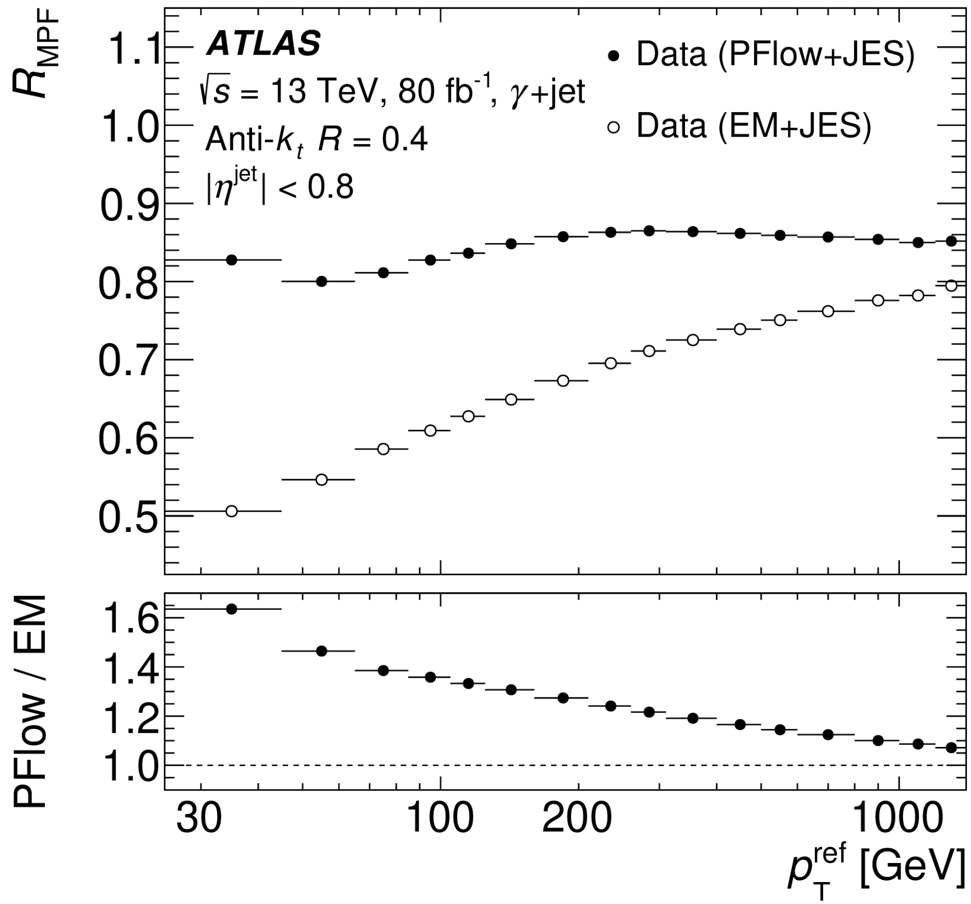
Evaluated as function of η jet and pT ref

Before: Direct balance of leading jet against reference (Z or photon)

IN SITU: *Z*+JET & *γ*+JET



IN SITU: *Z*+JET & *γ*+JET



PFlow measurement shows an improvement over the baseline calorimeter response at low pT thanks to the inclusion of information from tracks

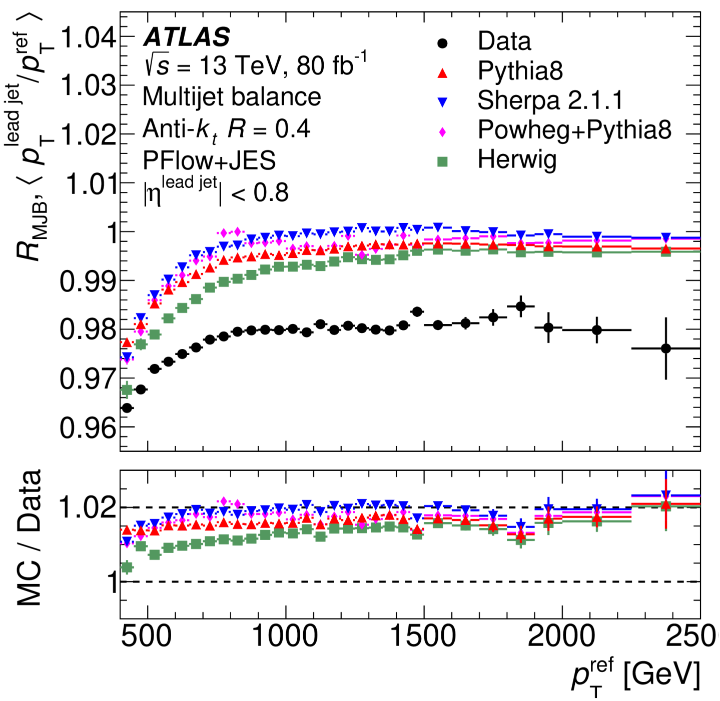
IN SITU: MULTIJET BALANCE

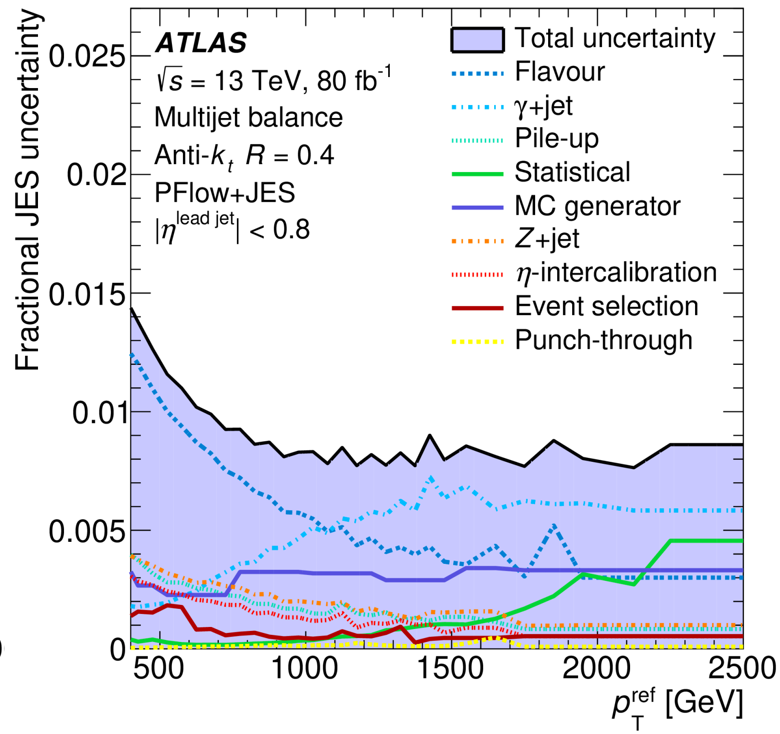
Multijet Balance (MJB) uses a system of well-calibrated low-pT jets to calibrate a single high-pT jet

- Jets of the recoil system are calibrated using Z/γ+jet results, while the leading jet is left at the η intercalibration scale
- Fully calibrated subleading jets require pT sublead < 1 TeV – limits pT of leading jet to about 2.5 TeV

Response:

$$\mathcal{R}_{\text{MJB}} = \left\langle \frac{p_{\text{T}}^{\text{lead}}}{p_{\text{T}}^{\text{ref}}} \right\rangle$$



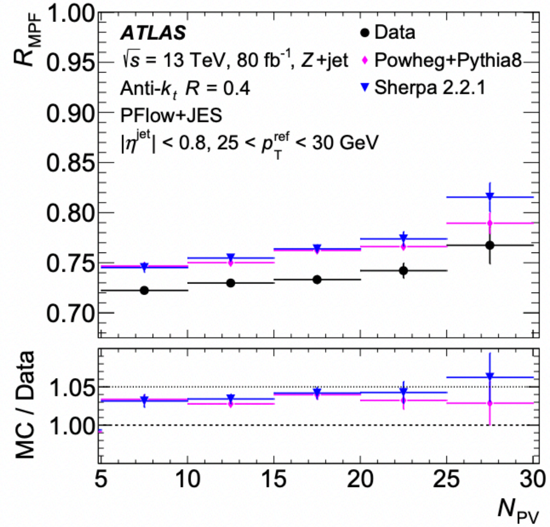


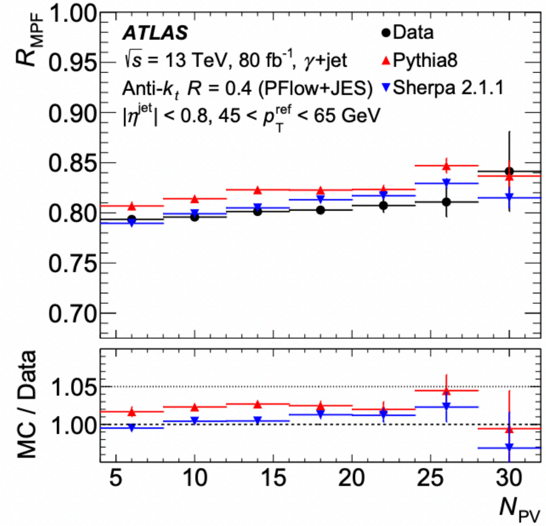
Clean Topology

- Avoid dijet signatures – require second jet not to carry considerable fraction of recoil pT
- Recoil system back to probed jet
- Leading jet isolated/recoil focused. ($\Delta\phi$ selection)

IN SITU: PILE UP

- A linear fit to the data/simulation ratio has a slope compatible with zero within the fit uncertainties in each plot, demonstrating the stability of the in situ calibration as a function of NPV



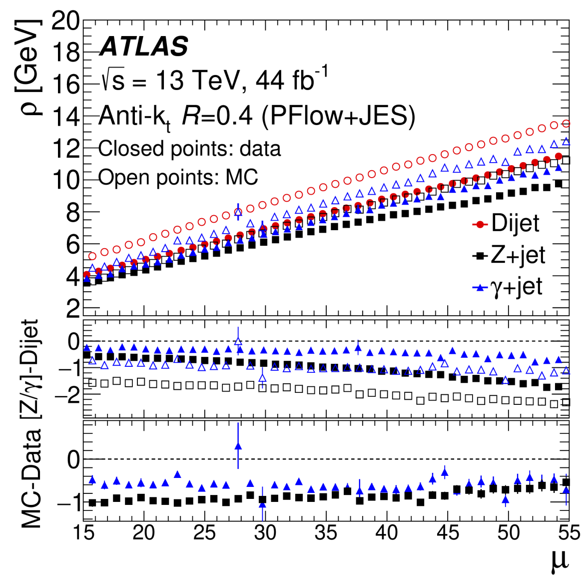


- ρ topology uncertainty

$$\Delta_{p_\mathrm{T}} = \max(|\Delta|) \times C^{\mathrm{JES}}_{p_\mathrm{T}} \times \pi R^2$$

$$\Delta = \left(\rho^{\mathrm{t1}}_{2017} - \rho^{\mathrm{t2}}_{2017}\right)_{\mathrm{MC}} - \left(\rho^{\mathrm{t1}}_{2017} - \rho^{\mathrm{t2}}_{2017}\right)_{\mathrm{data}}$$

evaluated at $\mu = 37.8$



FLAVOUR UNCERTAINTIES

Jet Flavour Response uncertainty

- Difference in MC-derived gluon response between generators
- fgluon is the fraction of gluon-jets in the event

$$\sigma_{flav}^{resp} = f_{gluon}(R_{gluon}^{Py8} - R_{gluon}^{H7})$$

Jet Flavour Composition uncertainty

- uncertainty in the fraction of gluon-initiated jets in the sample

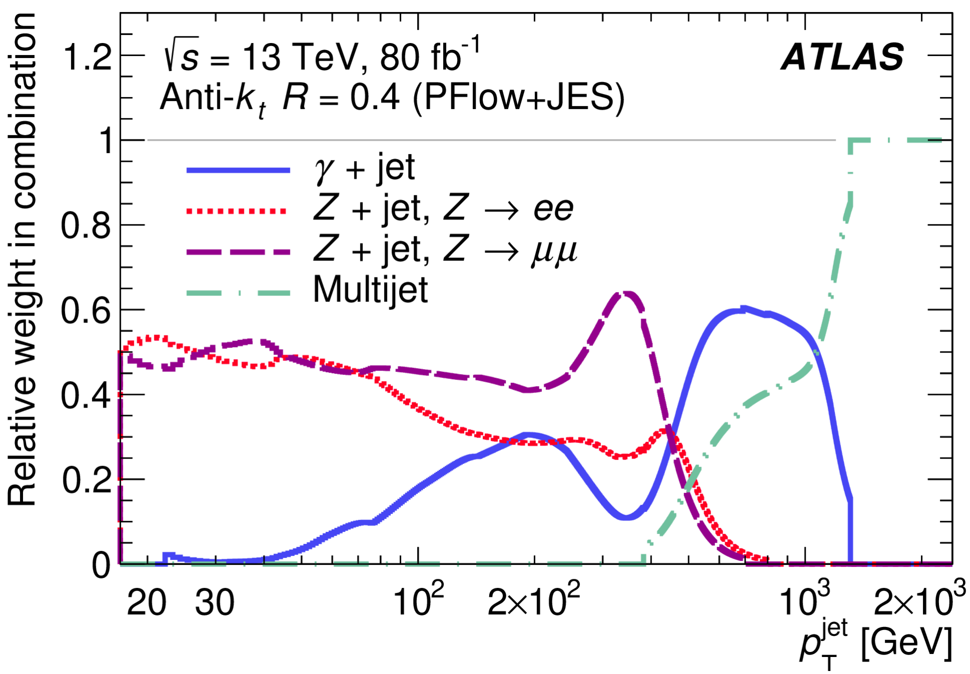
$$\sigma_{flav}^{comp} = \sigma_{f_{gluon}} \frac{|R_g - R_q|}{f_{gluon}R_g + (1 - f_{gluon})R_q}$$

IN SITU: COMBINATION

A statistical combination of the Z/γ+jet and MJB analyses provides a single smooth calibration applicable across the full momentum range.

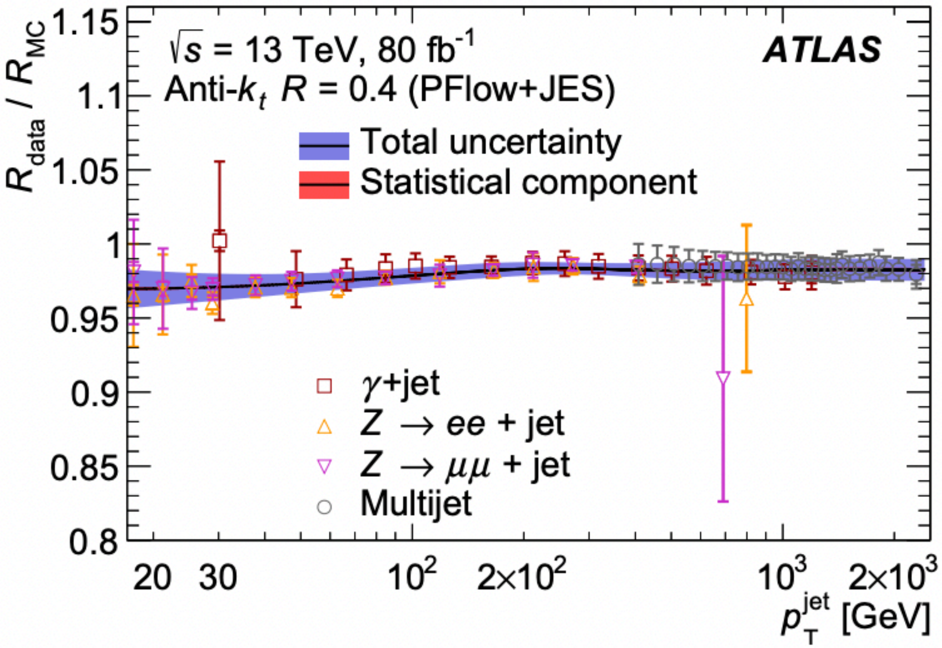
1) Absolute in situ measurements are converted from a parameterisation in pT ref into jet pT and divided into bins of 1 GeV using second-order polynomial splines.
2) A χ 2 minimization is performed in each bin as a weighted average (weight to each input decreasing as its uncertainty grows)
   - Combination favors method with greatest precision in each region

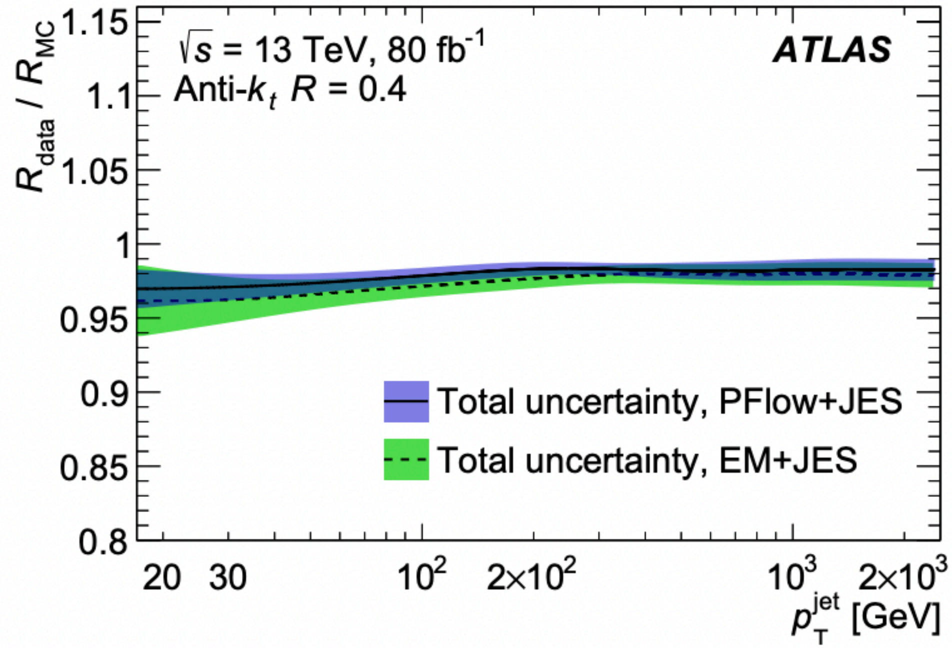
pT- dependent weight applied to calibrations



IN SITU: COMBINATION

To complete the calibration, the inverse of the curve (RMC/Rdata) is taken as the scaling factor and applied to data.
The correction is approximately 3% at low pT and decreases to around 2% for jets above 200 GeV





Systematic uncertainties
Z/γ+jet MPF, and MJB are performed sequentially => systematic uncertainties are propagated from each to the next.
Sources:

- **modelling** of physics processes in simulation (comparing two MC)
- uncertainties in the measurement of the **reference** object ($\pm 1\sigma$ in objects calibration)
- uncertainties in the expected pT balance due to the event's **topology** (varying the event selections and observing the impact on the MC to data ratio)

IN SITU: UNCERTAINTIES

A statistical combination of the Z/γ+jet and MJB analyses provides a single smooth calibration applicable across the full momentum range.

Systematic uncertainties

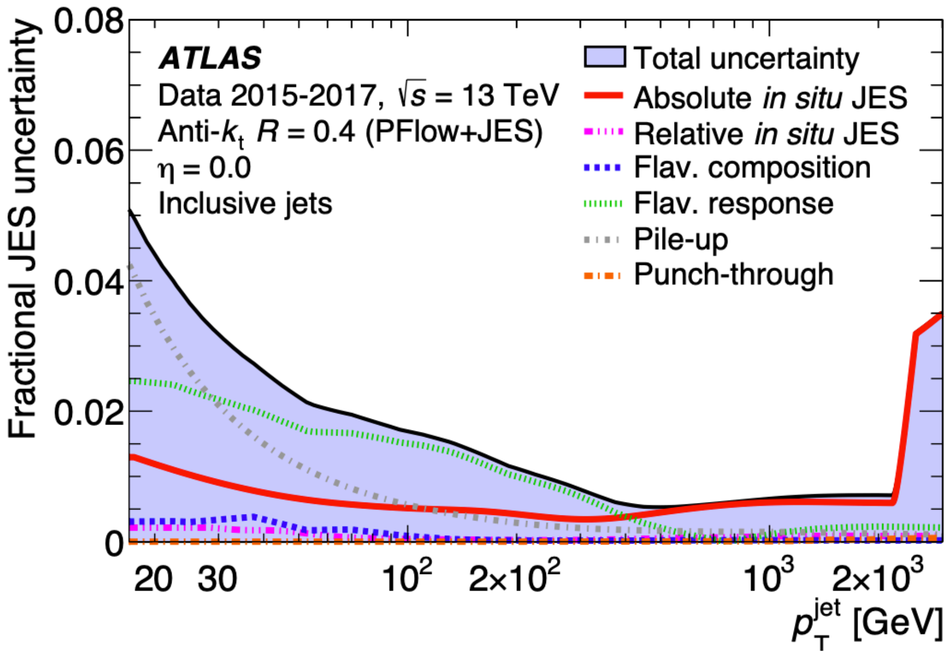
Z/γ+jet MPF, and MJB are performed sequentially => systematic uncertainties are propagated from each to the next.

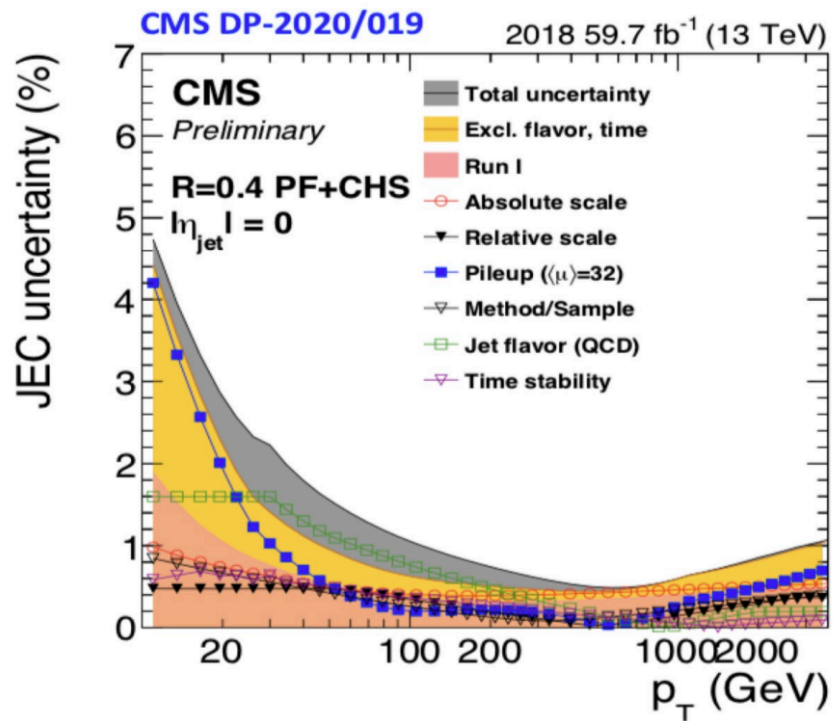
Sources:

- **modelling** of physics processes in simulation (comparing two MC)
- uncertainties in the measurement of the **reference** object ($\pm 1\sigma$ in objects calibration)
- uncertainties in the expected pT balance due to the event's **topology** (varying the event selections and observing the impact on the MC to data ratio)

CALIBRATION UNCERTAINTY

arXiv:2007.02645

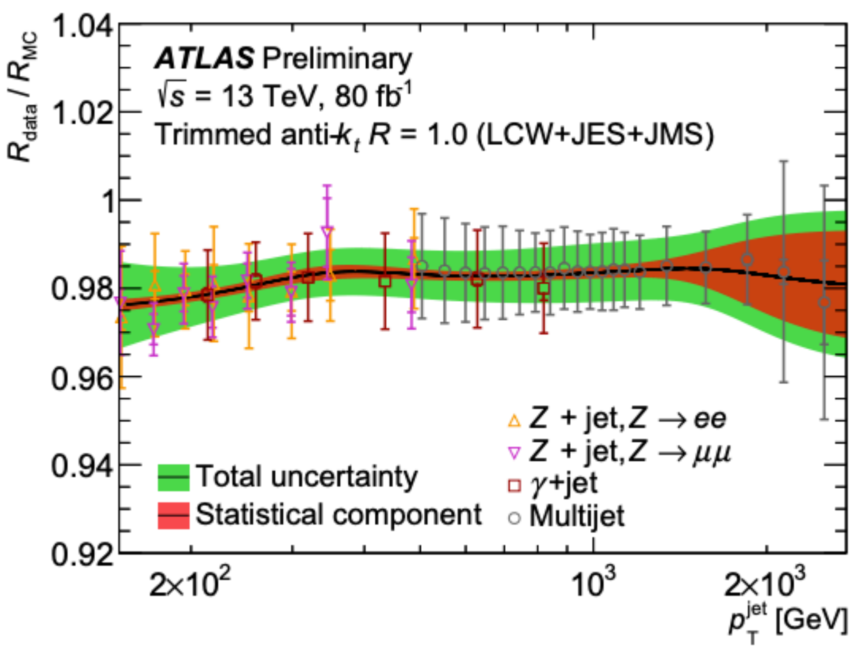


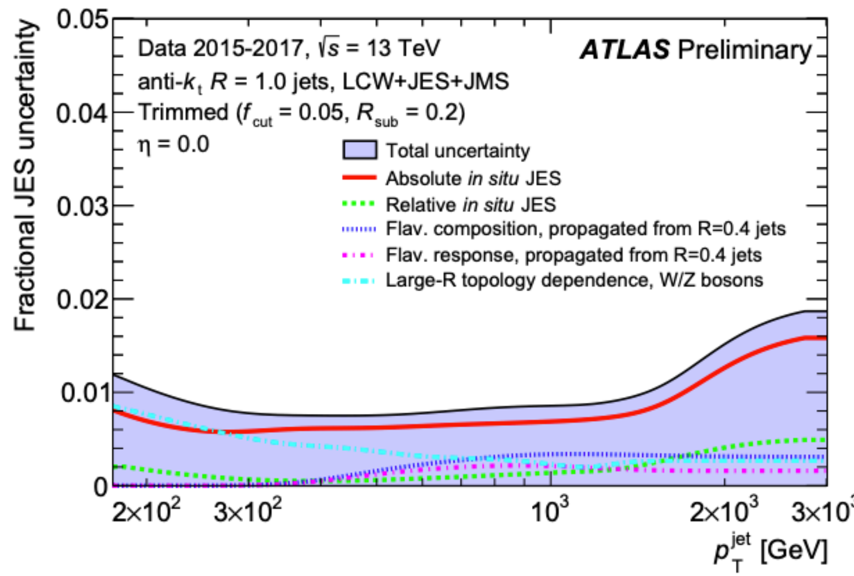


R=1.0 JET CALIBRATION

- A similar calibration procedure is applied to large-R jets.
- The central scale is also 1 − 2% lower in data, and a precision of ~ 1% is achieved.
- Uncertainties also cover the extrapolation between the QCD jets used in the balance methods and hadronic resonances - these are found to be sub-dominant.
- The multi-jet balance propagates the R = 0.4 jet uncertainties, and is the dominant method in the combinations above 600 GeV.

JETM-2019-05



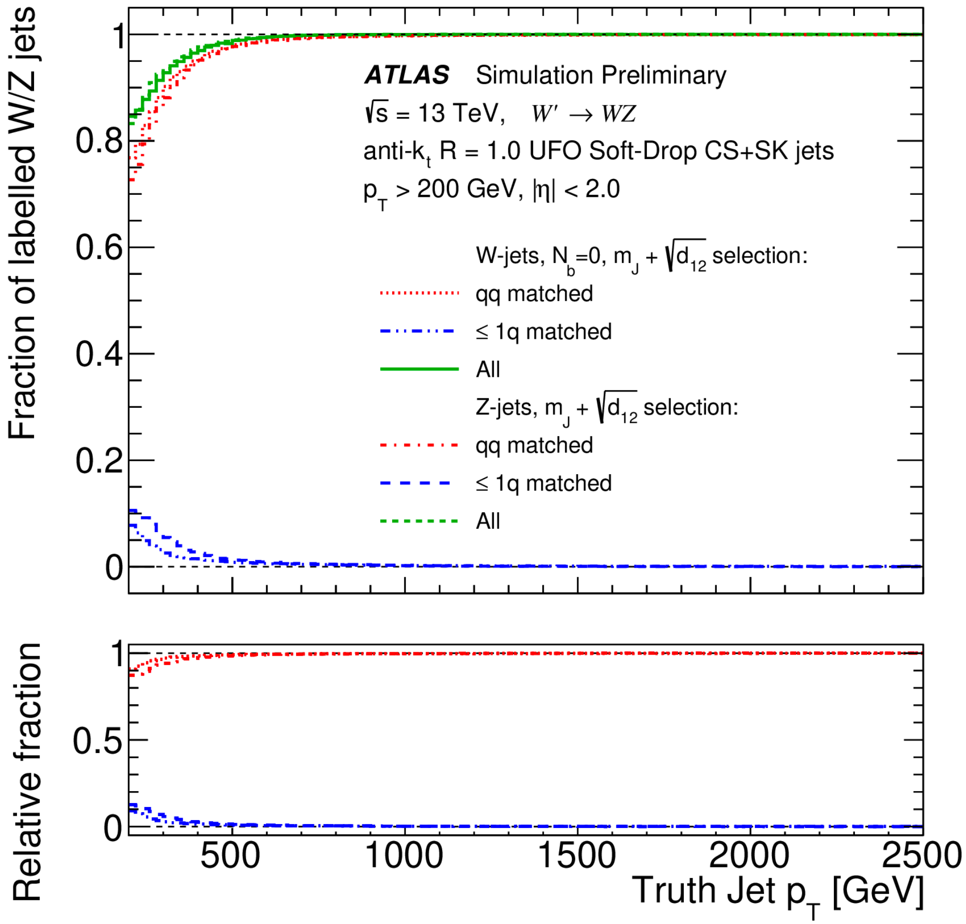


UFO TAGGERS

UFO TAGGERS: W/Z TRUTH LABELLING

ATL-PHYS-PUB-2021-029

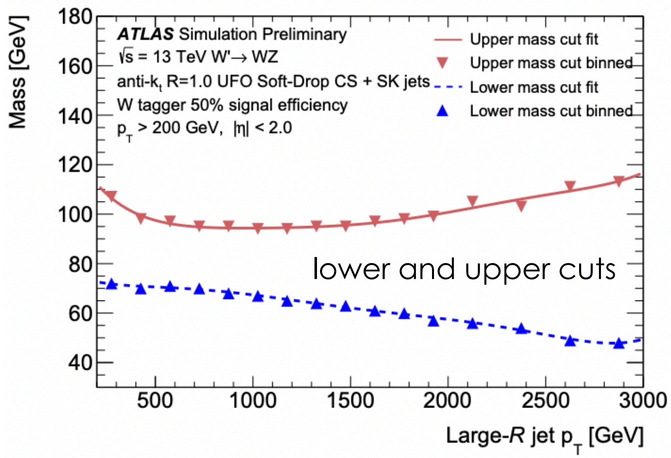
- truth large-R jets: reconstructed from stable particles using the anti-kt algorithm with R = 1.0
- Truth large-R jet is matched to the UFO jet by requiring ΔR < 0.75.
- energy-correlation function ratio, D2, of truth large-R jet: $\sqrt{d_{12}} > 55.25 \cdot \exp\left(\frac{-2.34 \times 10^{-3}}{\text{GeV}} p_{\text{T}}\right)$ GeV

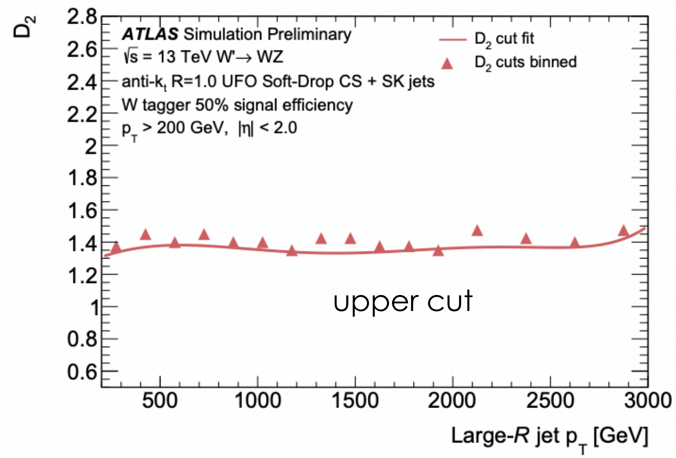


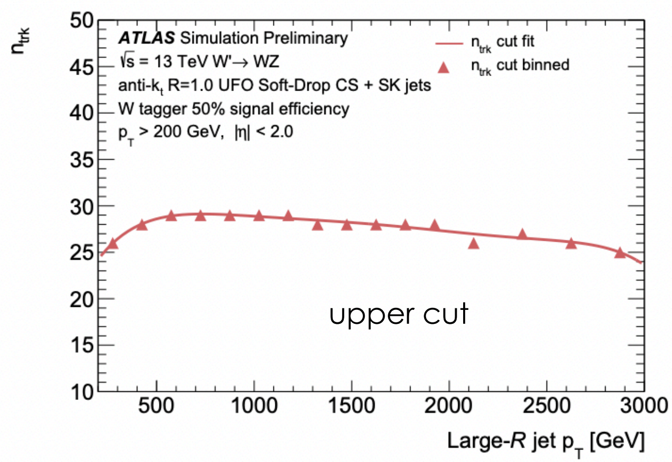
UFO TAGGERS: 3-VAR TAGGER

ATL-PHYS-PUB-2021-029

- Cut-based tagger
  - Based on rectangular cuts on substructure variables, optimized to achieve a flat 50% efficiency independent of pT.
- jet mass, mJ
- energy-correlation function ratio D2 with β = 1
- Number of ID tracks with pT > 500 MeV ghost-associated with the jet before the grooming, ntrk.







almost independent of pT

UFO TAGGERS: DNN TAGGER

ATL-PHYS-PUB-2021-029

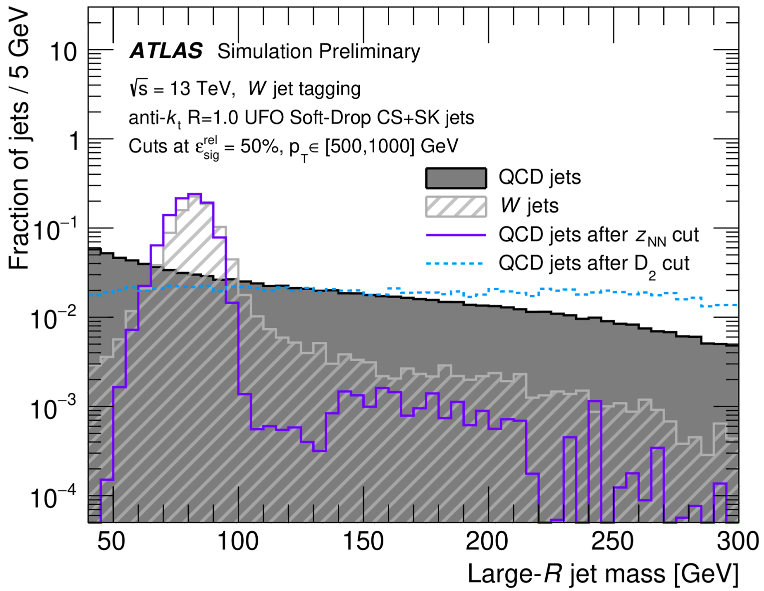
- three fully-connected 32 node dense layers
- **zNN**: resulting DNN score.

Variable	Description
D_2, C_2	Energy correlation ratios
τ_{21}	N-subjettiness
R_2^{FW}	Fox-Wolfram moment
$\mathcal{P}$	Planar flow
a_3	Angularity
A	Aplanarity
Z_{cut}, $\sqrt{d_{12}}$	Splitting scales
$Kt\Delta R$	k_t-subjet ΔR

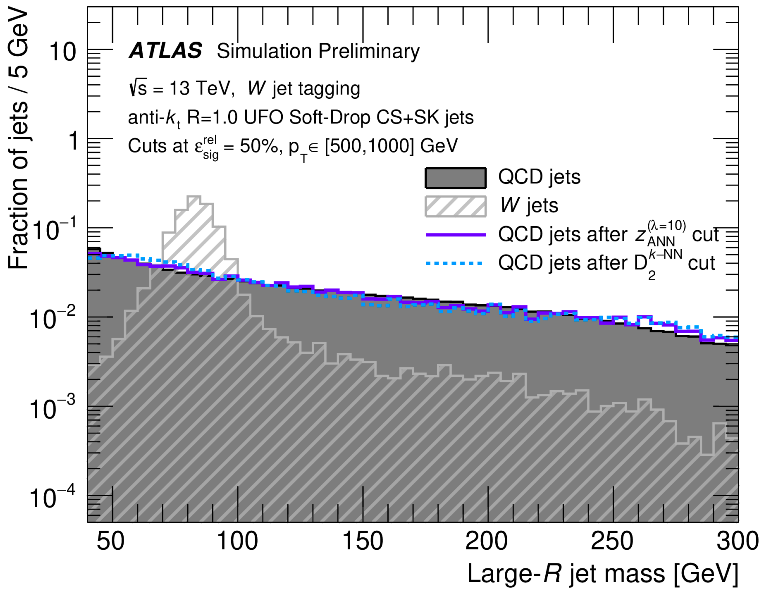
UFO TAGGERS: MASS-DECORRELATED TAGGERS

ATL-PHYS-PUB-2021-029

- Bkg-jet shape quite similar to the W-jets after taggers



- the mass sculpting is suppressed



$$D_2^{k\text{-NN}} = D_2 - D_2^{8\%}$$

- Adversarial Neural Network (ANN) trained to compete with the DNN tagger

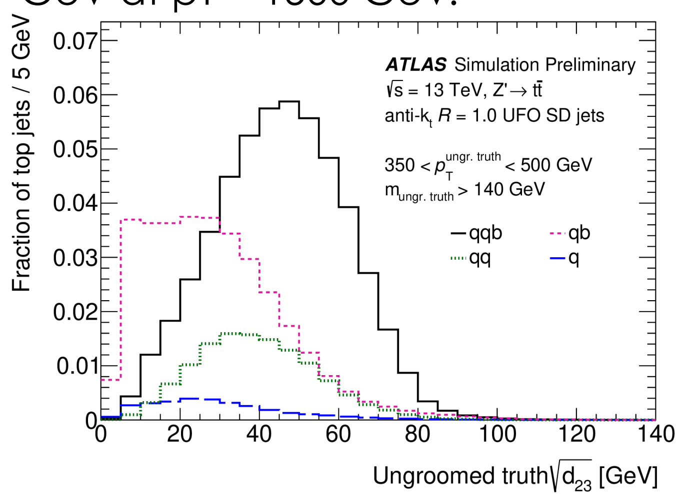
UFO TAGGERS: TOP TRUTH LABELLING

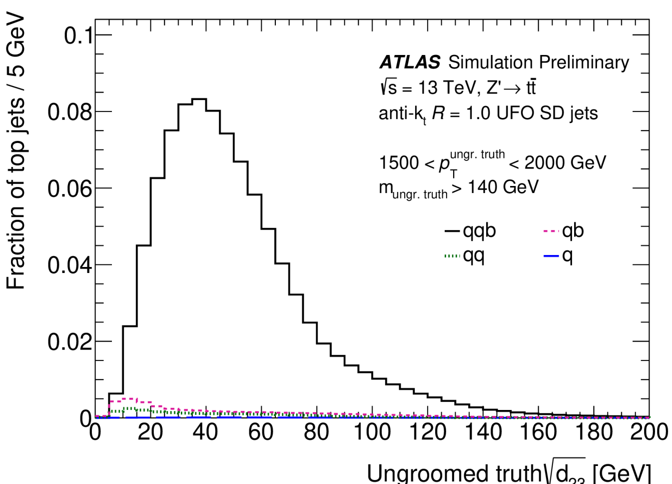
- strategies
  - inclusive: any jet that contains parts of the hadronic top quark decay regardless of full containment.
    - Reco-truth matching: ΔR(J, Jtruth) < 0.75
    - Truth-top matching: ΔR(Jtruth, top) < 0.75
  - contained-top: contain the full decay of the top quark.
    - **Parton based:** truth-Decay products of top (t → W b → qq'b) matched ΔR < 0.75.
    - **Groomed**: one b-hadron is ghost-associated to the truth jet && groomed truth jet mass m > 140 GeV
    - **Ungroomed** (New): one b-hadron is ghost-associated to the ungroomed truth jet Jtruth, its mass is above 140 GeV and fulfills a pT-dependent criteria on the kt splitting scale $\sqrt{d_{32}}$

ATL-PHYS-PUB-2021-028

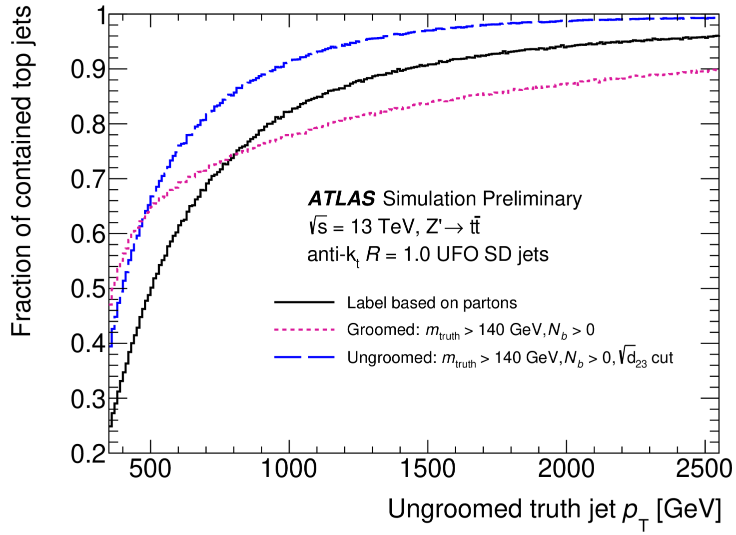
Parton level categories: qqb, qq, qb and q matched to ruth ΔR(Jtruth, q/q'/b) < 0.75.

The optimal cut on $\sqrt{d_{32}}$ decrease exponentially with pT , having values of ≈ 21 GeV at pT = 350 GeV and ≈ 7 GeV at pT = 1500 GeV.

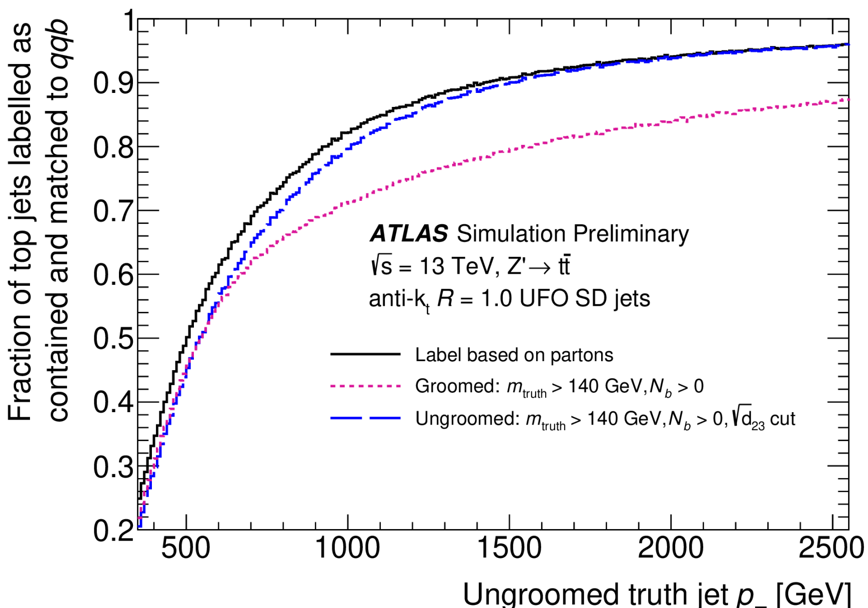




UFO TAGGERS: TOP TRUTH LABELLING



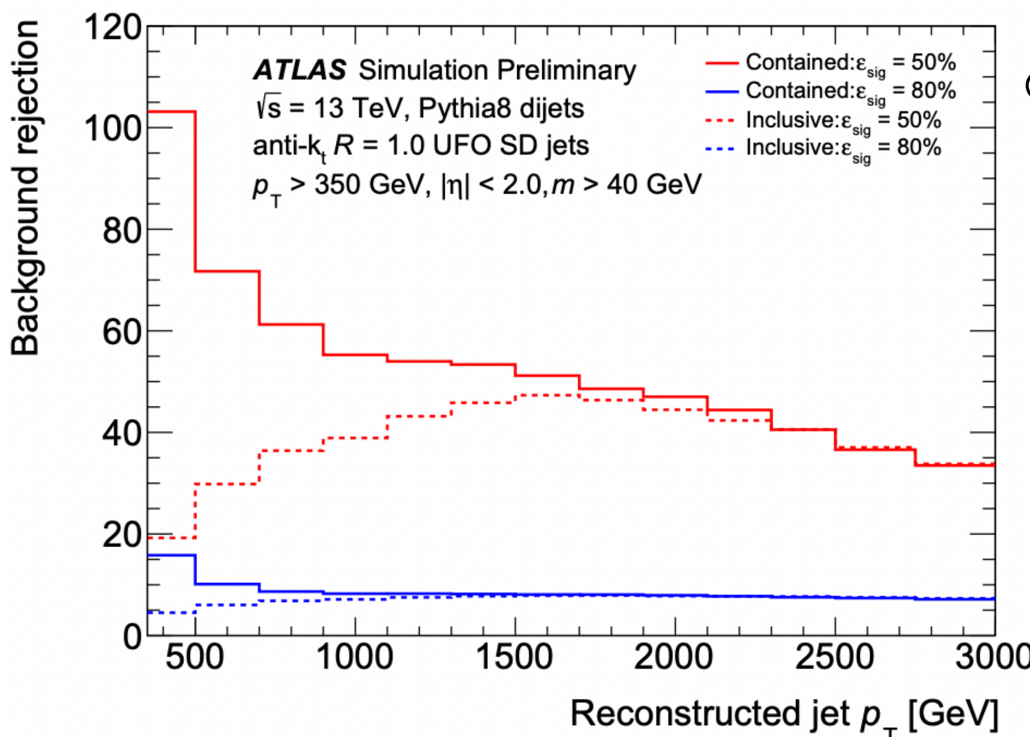
jets are in addition matched to the qq'b partons by requiring ΔR < 0.7:



The ungroomed truth label selects approximately the same number of qq'b-matched jets as the parton-based label, but overall it classifies more jets as contained tops.

UFO TAGGERS: TOP TAGGERS

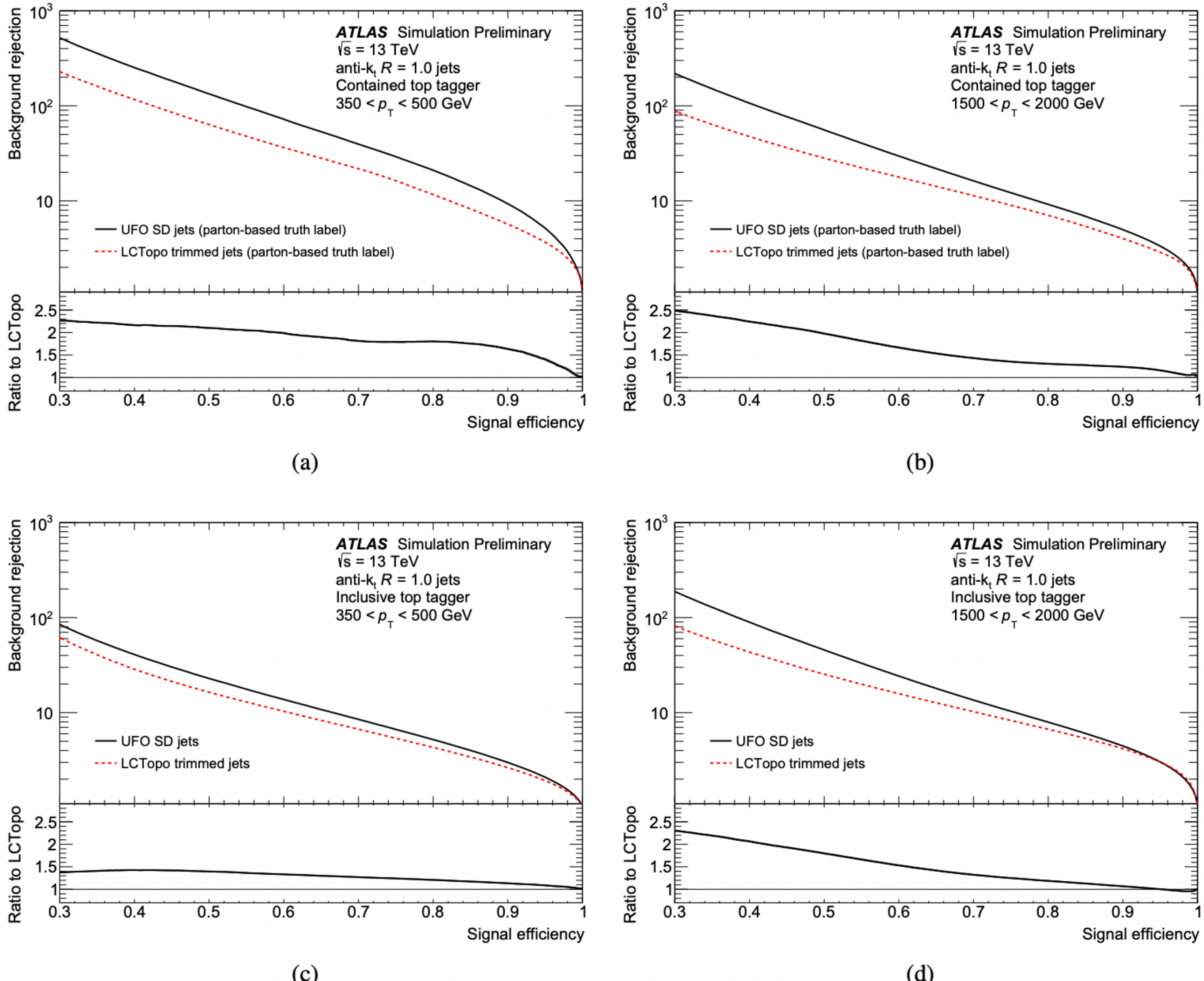
- Two taggers based on deep neural-network (DNN):
  - 1) trained to identify jets labelled as **inclusive** tops
  - 2) to identify jets that are labelled as **contained** tops
- Selection: m > 40 GeV, N const ≥ 3 for good jet substructure features
- Variables: N-subjettiness, τ1 , τ2 , τ3 and τ4 ; kt splitting scales: $\sqrt{d_{12}}$ and $\sqrt{d_{23}}$; (generalized) energy correlation functions and their ratios: ECF1 , ECF2 , ECF3 , C2 , D2 , L2 and L3 ; QW as well as thrust major TM
- Selection criteria applied on DNN score to define two working points with a fixed signal efficiency of 80% and 50% as a function of the pT



Contained: ungroomed truth label

UFO VS LCTOPO: TOP TAGGERS

With parton-based labelling



MET

MET NET

	Selections		
	p_{T} [GeV] for jets with:		
Working point	$\|\eta\| < 2.4$	$\|\eta\| > 2.4$	JVT for jets with $\|\eta\| < 2.4$
Loose	> 20	> 20	> 0.59 for $p_{\mathrm{T}} < 60$ GeV jets
Tight	> 20	> 30	> 0.59 for $p_{\mathrm{T}} < 60$ GeV jets
Tighter	> 20	> 35	> 0.59 for $p_{\mathrm{T}} < 60$ GeV jets
Tenacious	> 20	> 35	> 0.91 for $20 < p_{\mathrm{T}} < 40$ GeV jets > 0.59 for $40 < p_{\mathrm{T}} < 60$ GeV jets > 0.11 for $60 < p_{\mathrm{T}} < 120$ GeV jets

MET SIGNIFICANCE INTRODUCTION

Significance of an observable

- Useful for signal-background discrimination
- Accounts for the observable resolution

$$\mathrm{Signif}(X) = \frac{X}{\sigma_X}$$

A **high value** of **MET significance:**

- Indicates that MET is not well explained by resolution smearing alone
- Suggests that the event is more likely to contain **unseen objects**

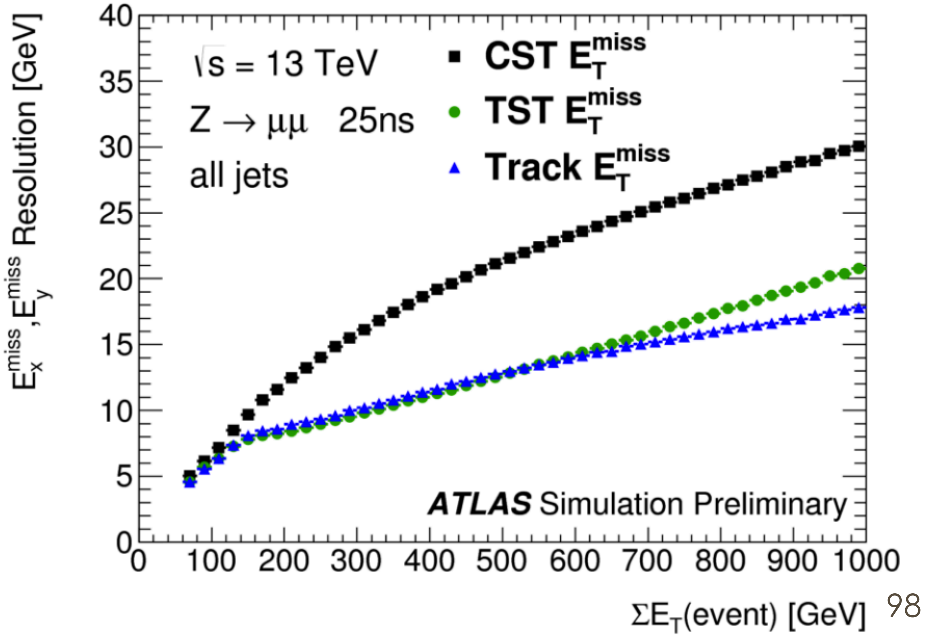
Event-based MET Significance

ATL-PHYS-PUB-2015-023

$$\mathcal{S} = \frac{E_{\mathrm{T}}^{\mathrm{miss}}}{\boxed{\sqrt{\sum E_{\mathrm{T}}}}}$$

Limitations:

- Proxy for the MET resolution
- Event based quantity, neglecting the nature of the objects.
- Do not take into account directional correlations



OBJECT BASED MET SIGNIFICANCE DEFINITION

In a coordinate system with axis parallel (L) and perpendicular (T) to the reconstructed Met direction:

$$S^2 = \frac{\left|\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}\right|^2}{\sigma_L^2\left(1-\rho_{LT}^2\right)}$$

Total Variance Longitudinal

$$\sigma_L^2 = \sigma_L^{hard^2} + \sigma_L^{soft^2}$$

Resolutions

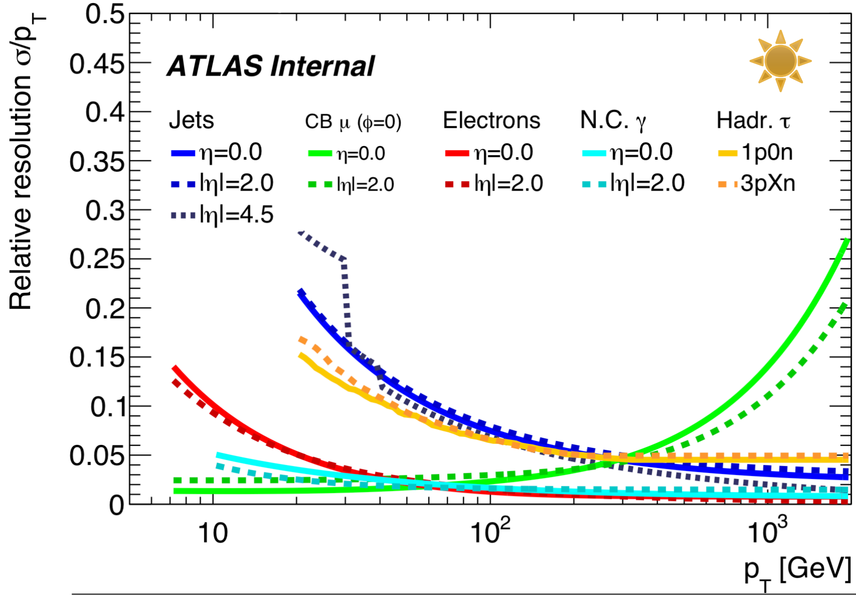
- Hard object variances
- Soft Terms variance

Transverse momentum resolution in the transverse plane for the hard objects

$$\mathbf{V}_i = \begin{pmatrix} \sigma_{p_{T\,i}}^2 & 0 \\ 0 & {p_{T\,i}}^2\sigma_{\phi_i}^2 \end{pmatrix}$$

Resolution parallel to the direction of the object

Resolution perpendicular to the direction of the object



Object	Kinematic	Relative resolution	
		Parallel	Perpendicular
Electrons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	1.7%	0.4%
Photons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	1.9%	0.4%
Hadronic τ	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	5.5% – 6.7%	1%
Jets	$p_{\mathrm{T}} = 20$ GeV, $\eta = 0$	22%	4.6%–7.1%
	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	7%	1.1%–1.6%
Muons	$p_{\mathrm{T}} = 100$ GeV, $\eta = 0$	2–4%	0.1%
Track Soft Term		8.9 GeV	8.9 GeV

OBJECT BASED MET SIGNIFICANCE DEFINITION

In a coordinate system with axis parallel (L) and perpendicular (T) to the reconstructed Met direction:

$$S^2 = \frac{\left|\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}\right|^2}{\sigma_L^2\left(1-\rho_{LT}^2\right)}$$

Resolutions

- Hard object variances
- Soft Terms variance

Total Variance Longitudinal

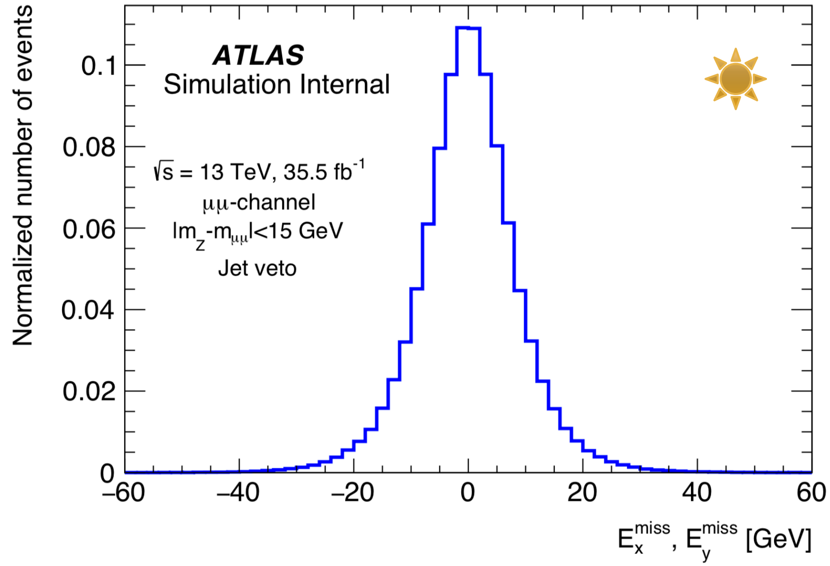
$$\sigma_L^2 = \sigma_L^{hard^2} + \sigma_L^{soft^2}$$

An additional covariance matrix associated to the soft term is also considered:

$$\mathbf{V}_{\mathrm{soft}} = \begin{pmatrix} \sigma_{\mathrm{soft}}^2 & 0 \\ 0 & \sigma_{\mathrm{soft}}^2 \end{pmatrix}$$

Constant soft Term resolution:
8.9 GeV

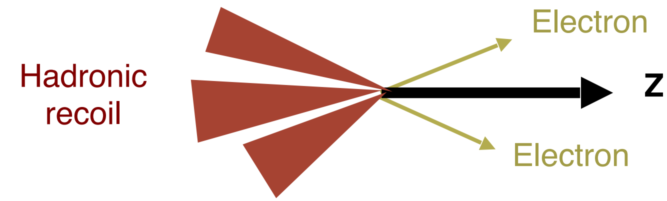
RMS of the MET projection on x in Z→μμ events with jet veto (SumJetEt=0)



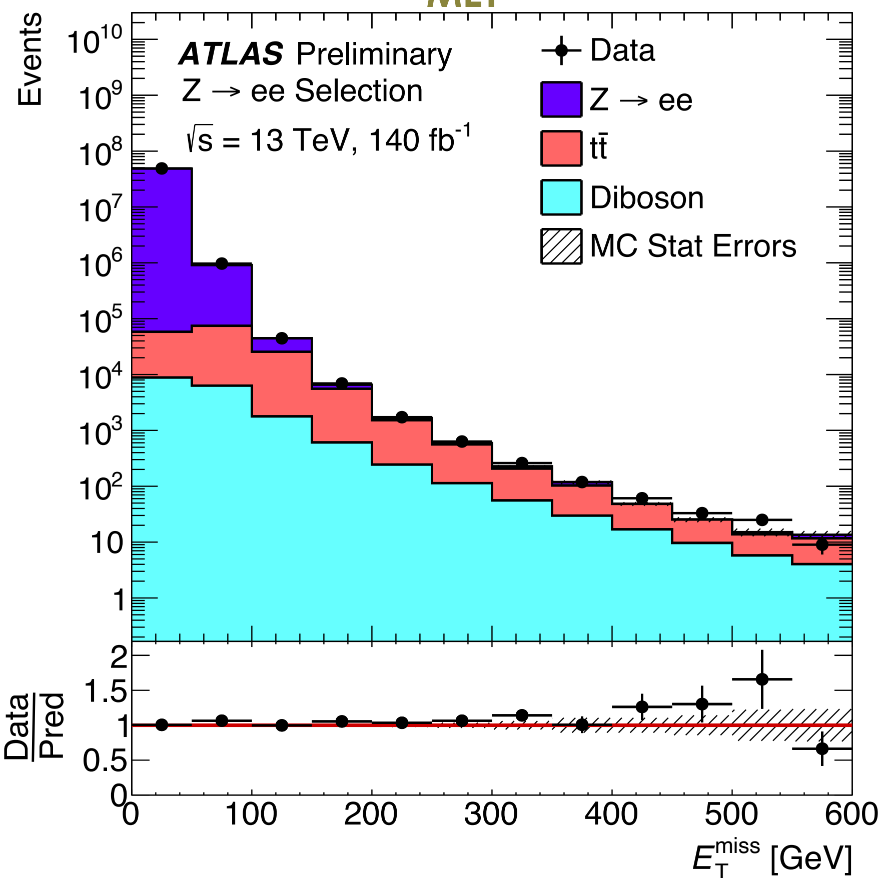
comparison of μμ and ee

PERFORMANCE: MODELLING

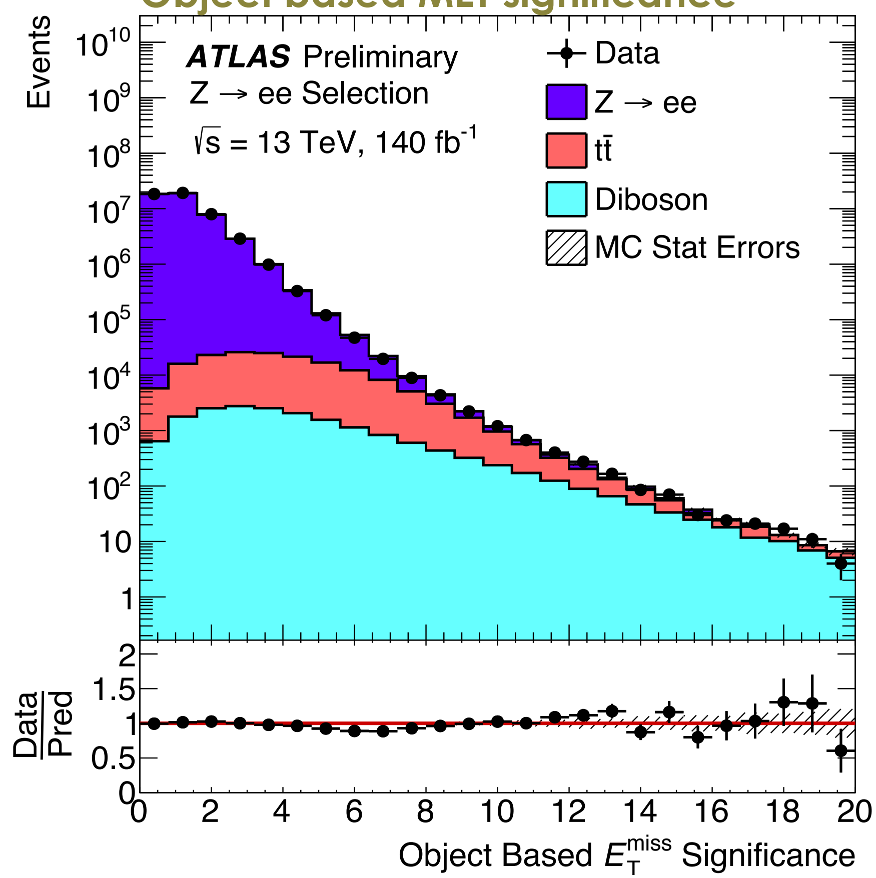
◇ In order to study the performance of the MET… clean process without real MET: **$Z \rightarrow \ell\ell$ + jet selection**



MET



Object based MET significance



Good modelling over the full **Run 2 data-set**!