MANCHESTER 1824
The University of Manchester

ATLAS EXPERIMENT

ATLAS Jet Reconstruction and Calibration

LHC Electroweak Working Group (EWWG) Meeting

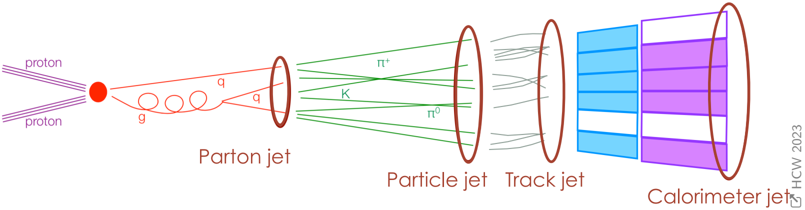
Tobias Fitschen, Ana Peixoto (Jet Definitions and MC Calibration Conveners)
On behalf of the ATLAS Collaboration

26 June 2024

University of Manchester

European Research Council
Established by the European Commission

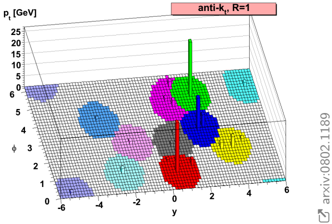
Types of Jets in ATLAS



Jet definitions

- **Particle** (truth) jets (MC only)
- **Track** jets: tracker info only
- **Calo** jets: calo info only
  - LCTopo, EMTopo
- **Combined** track+calo jets
  - ParticleFlow, UFO, TrackCaloCluster

Anti-k_T jet algorithm

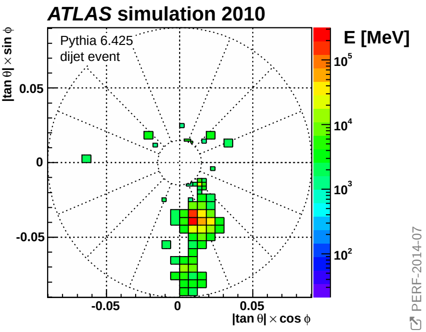


- Mostly circular in $y - \phi$ plane
- Used for most purposes

Constituents:
TopoClusters

Topo Clusters

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



Topological Clusters

of E deposits in calorimeter cells

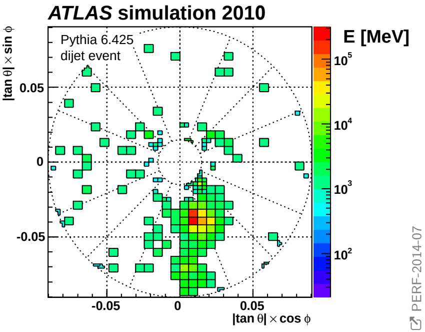
$\rightarrow$ algorithm:

1. **Seed:** Find cells with energy $E > 4 \times |\zeta|$

Cell noise ratio: $\zeta_{\text{cell}}^{\text{EM}} = \dfrac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}$

Topo Clusters

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



Topological Clusters

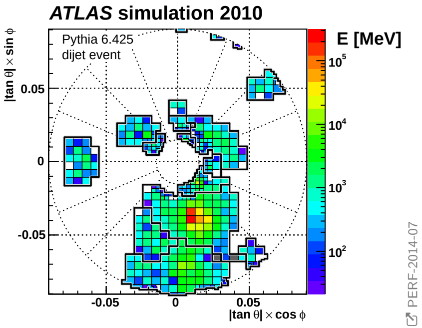
of E deposits in calorimeter cells

→ algorithm:

1. **Seed:** Find cells with energy $E > 4 \times |\zeta|$
2. **Growth:** Neighbors with $E > 2 \times |\zeta|$ are added

Cell noise ratio: $\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}$

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



Topological Clusters

of E deposits in calorimeter cells

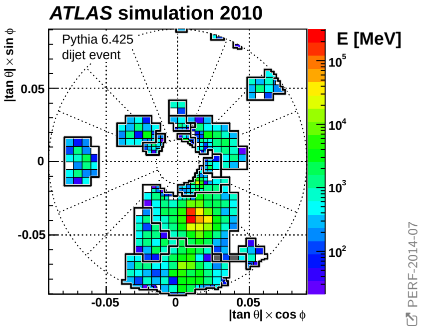
$\rightarrow$ algorithm:

1. **Seed:** Find cells with energy $E > 4 \times |\zeta|$
2. **Growth:** Neighbors with $E > 2 \times |\zeta|$ are added
3. **Boundary:** any neighboring cells are added

(no ζ requirement)

Cell noise ratio: $\zeta_{\text{cell}}^{\text{EM}} = \dfrac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}$

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



Cell noise ratio: $\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}$

Topological Clusters

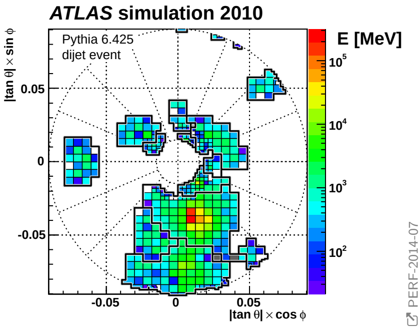
of E deposits in calorimeter cells

$\rightarrow$ algorithm:

1. **Seed:** Find cells with energy $E > 4 \times |\zeta|$
2. **Growth:** Neighbors with $E > 2 \times |\zeta|$ are added
3. **Boundary:** any neighboring cells are added

   (no ζ requirement)
4. **Split:** Breaks up clusters with multiple maxima

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



Cell noise ratio: $\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}$

Topological Clusters

of E deposits in calorimeter cells

$\rightarrow$ algorithm:

1. **Seed:** Find cells with energy $E > 4 \times |\zeta|$
2. **Growth:** Neighbors with $E > 2 \times |\zeta|$ are added
3. **Boundary:** any neighboring cells are added

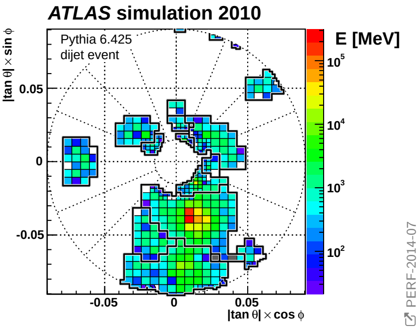
   (no ζ requirement)
4. **Split:** Breaks up clusters with multiple maxima

Jets build from TopoClusters are called **EMTopo** Jets

EM: Electromagnetic scale

Topo Clusters

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



Cell noise ratio: $\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}$

Topological Clusters

of E deposits in calorimeter cells

$\rightarrow$ algorithm:

1. **Seed:** Find cells with energy $E > 4 \times |\zeta|$
2. **Growth:** Neighbors with $E > 2 \times |\zeta|$ are added
3. **Boundary:** any neighboring cells are added

   (no ζ requirement)
4. **Split:** Breaks up clusters with multiple maxima

Jets build from TopoClusters are called **EMTopo** Jets

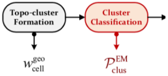
EM: Electromagnetic scale

$\rightarrow$ ATLAS calorimeters are non-compensating

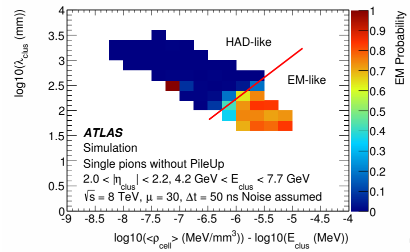
$\rightarrow$ EM response ≈ 1, hadronic response < 1

Local Cluster Weighting (LCW)

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



HCW 2023

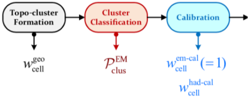


PERF-2014-07

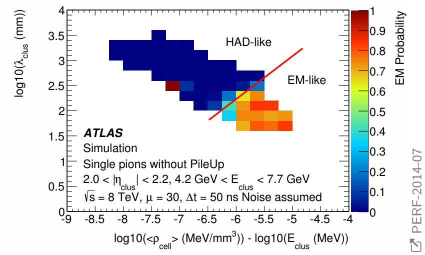
- TopoClusters are identified to be EM or had by likelihood $\mathcal{P}_{\text{clus}}^{\text{EM}}$

Local Cluster Weighting (LCW)

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



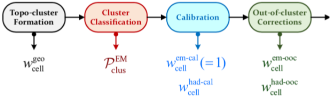
HCW 2023



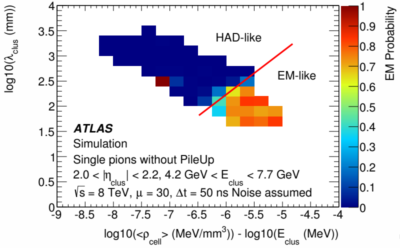
- TopoClusters are identified to be EM or had by likelihood $\mathcal{P}_{\text{clus}}^{\text{EM}}$
- Their momenta are reweighted (ω) by
  - Difference in response due to non-compensating calorimeter

Local Cluster Weighting (LCW)

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



HCW 2023

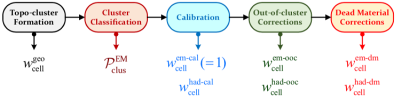


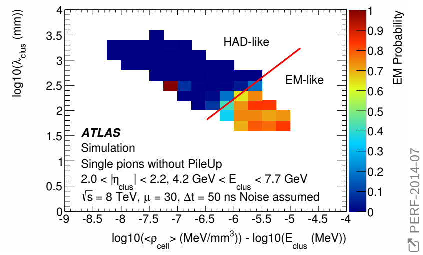
PERF-2014-07

- TopoClusters are identified to be EM or had by likelihood $\mathcal{P}_{\text{clus}}^{\text{EM}}$
- Their momenta are reweighted (ω) by
  - Difference in response due to non-compensating calorimeter
  - Energy falling in unclustered cells

Local Cluster Weighting (LCW)

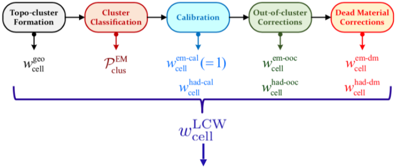
Eur. Phys. J. C 77 (2017) 490





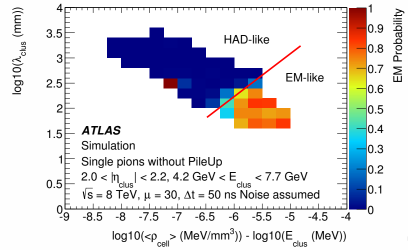
- TopoClusters are identified to be EM or had by likelihood $\mathcal{P}_{\text{clus}}^{\text{EM}}$
- Their momenta are reweighted (ω) by
  - Difference in response due to non-compensating calorimeter
  - Energy falling in unclustered cells
  - Inactive/dead regions of the detector

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



$$E_{\mathrm{clus}}^{\mathrm{LCW}} = \sum_{i \in \mathrm{cluster}} w_{\mathrm{cell},i}^{\mathrm{LCW}} \, E_{\mathrm{cell},i}^{\mathrm{EM}}$$

HCW 2023

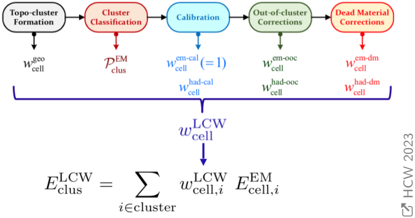


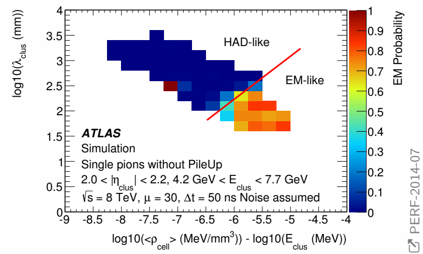
PERF-2014-07

- TopoClusters are identified to be EM or had by likelihood $\mathcal{P}_{\mathrm{clus}}^{\mathrm{EM}}$
- Their momenta are reweighted (ω) by
  - Difference in response due to non-compensating calorimeter
  - Energy falling in unclustered cells
  - Inactive/dead regions of the detector

Local Cluster Weighting (LCW)

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





- TopoClusters are identified to be EM or had by likelihood $\mathcal{P}_{\text{clus}}^{\text{EM}}$
- Their momenta are reweighted (ω) by
  - Difference in response due to non-compensating calorimeter
  - Energy falling in unclustered cells
  - Inactive/dead regions of the detector

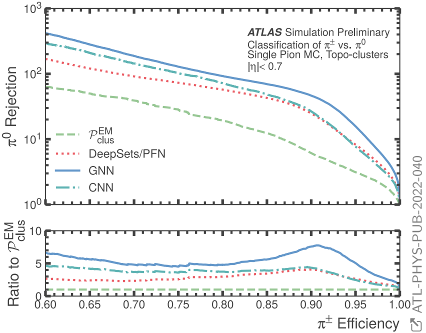
Jets build from TopoClusters+LCW are called **LCTopo** Jets used for large-R ($R = 1.0$) jets in Run 2

Recent Development:
ML Cluster Calibration

First step in cluster calibration: Differentiate EM from hadronic clusters

Non-compensating ATLAS calorimeter requires different calibrations for neutral/charged clusters

π^0 vs $\pi^\pm$ classification performance



Point cloud of energy deposits in calorimeter cells

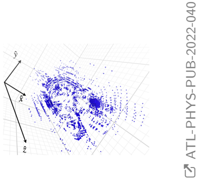
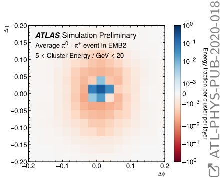


Image: $\pi^0 - \pi^\pm$ difference



Baseline used in LCW: $\mathcal{P}^{EM}_{clus}$

- Binned EM-scale cluster variables
  - Total cluster energy $E^{EM}_{cluster}$
  - Pseudorapidity η
  - Longitudinal depth λ_{clus}
  - 1st cell energy moment $\langle\rho_{cell}\rangle$
- Combined into likelihood $\mathcal{P}^{EM}_{clus}$

Individual calorimeter cell signals

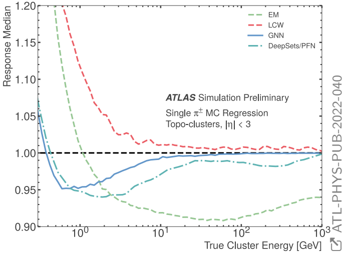
→ As point clouds (**GNN**, **PFN**)

→ Or projected on images (**CNN**)

Observations

- All point cloud methods significantly outperform baseline $\mathcal{P}^{EM}_{clus}$

$\pi^{\pm}$ cluster energy response



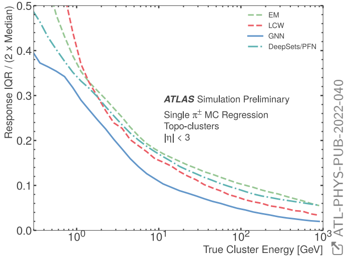
Second step: Energy Calibration

Observations

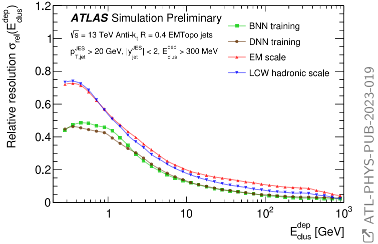
- **GNN** performs best wrt. response and width
- Followed by **Deep Sets**
- New: **Bayesian NN (BNN)**

Cluster Energy Resolution

Point Cloud Methods



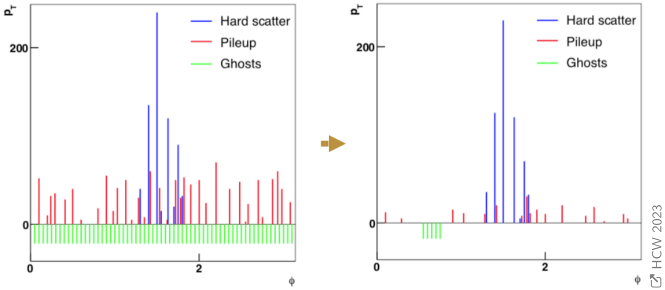
DNN / BNN



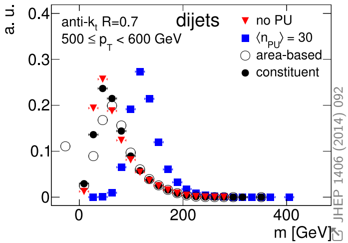
Pileup Mitigation
at Constituent Level

Constituent Subtraction (CS)

JHEP 1406 (2014) 092

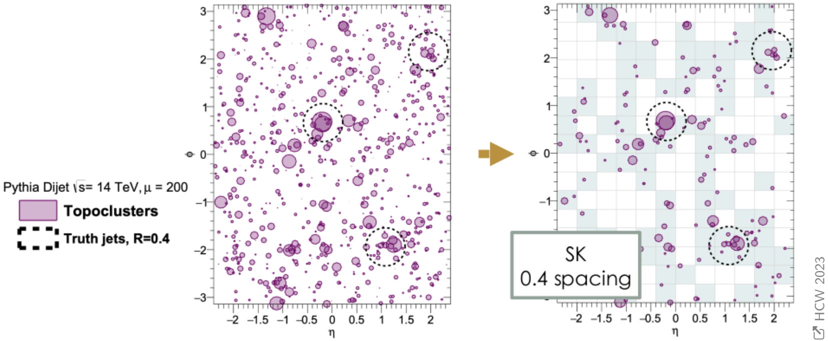


- Add ghosts in grid of $A_g = \eta \times \phi = 0.1 \times 0.1$
- With $p_T^g = A_g \times \rho$
  - $\rho = \text{med}\{\frac{p_T}{A}\}$: median energy density in event
  - Measure of PU in event
- Subtract p_T^g from p_T of constituents c within $\Delta R(g, c)$



Mass profile with CS closer to no-PU than with area-based alone

Soft Killer (SK)



- CS: Scales constituents
- SK: Removes constituents
- Consider constituents in η, ϕ grid
- All constituents with $p_{\mathrm{T}} < p_{\mathrm{T}}^{\mathrm{cut}}$ are removed
- $p_{\mathrm{T}}^{\mathrm{cut}}$ determined such that half of grid cells are empty

ATLAS uses CS+SK for R=1.0 jets

Constituents:
Adding Tracks

Adding Tracker Information

ATLAS EXPERIMENT

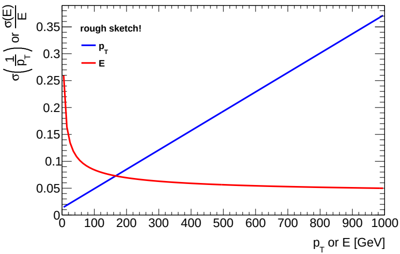
the tracker p_{T} resolution

$$\sigma\left(\frac{1}{p_{\mathrm{T}}}\right) = 0.036\% \cdot p_{\mathrm{T}} \oplus 1.3\%$$

is better than the calorimeter E resolution

$$\frac{\sigma(E)}{E} = \frac{50\%}{\sqrt{E}} \oplus 3.4\% \oplus \frac{1\%}{E}$$

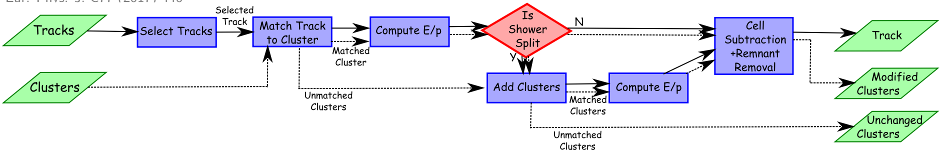
source



Additionally, the tracker has better acceptance (threshold) for soft particles

Jet definitions used in Run 3 (and partly Run 2) rely on calo+track information

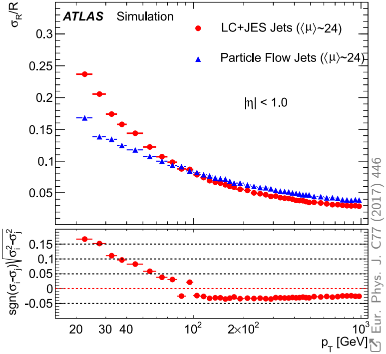
Eur. Phys. J. C77 (2017) 446

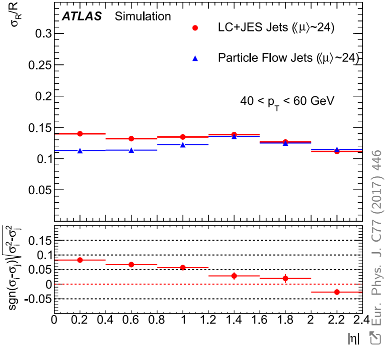


PFlow makes use of tracking information at constituent level

shows great JER improvement over calo jets in low-pT

Especially in the central region





arxiv:1306.4945

In addition to R=0.4 many analyses use R=1.0 jets

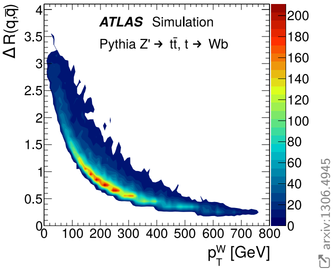
Best option depends on $p_T(V)$:

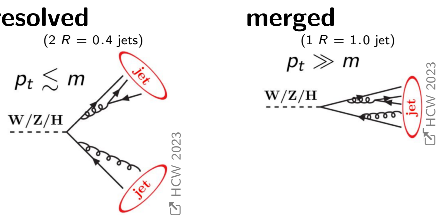
- Separation inversely proportional to transverse momentum p_T

$$\Delta R(q, q') \approx \frac{2m_W}{p_T^W}$$

- For $m_W = 80$ GeV, $R = 0.4$ cones around qq' overlap ($\Delta R < 0.8$) at $p_T > 200$ GeV

→ Reconstruct merged





Jet-tagging can be done to identify initiator of R=1.0 jets

→ **need good mass and substructure resolution**

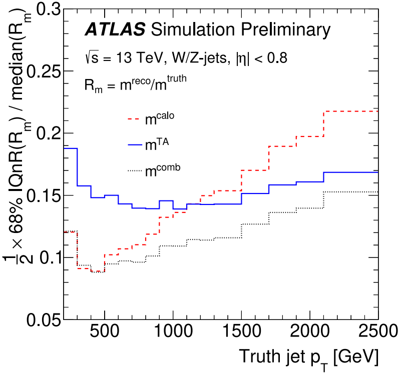
ATLAS-CONF-2016-035

Combine track with calo information for jet-mass definition

- Tracks are ghost-associated to calo-jet, yielding track mass m^{track}
- Scaled by calo/track correction factor accounting for neutral components

$$m^{TA} = \frac{p_T^{calo}}{p_T^{track}} \times m^{track}$$

- Linearly combined with calo mass according to resolution σ



ATLAS-CONF-2016-035

$$m^{comb} = m_{calo}\frac{\sigma_{m_{calo}}^{-2}}{\sigma_{m_{calo}}^{-2} + \sigma_{m_{TA}}^{-2}} + m_{TA}\frac{\sigma_{m_{TA}}^{-2}}{\sigma_{m_{calo}}^{-2} + \sigma_{m_{TA}}^{-2}}$$

$\rightarrow$ **Improved mass resolution over the whole p_T range**

$\rightarrow$ but only for mass, not for variables

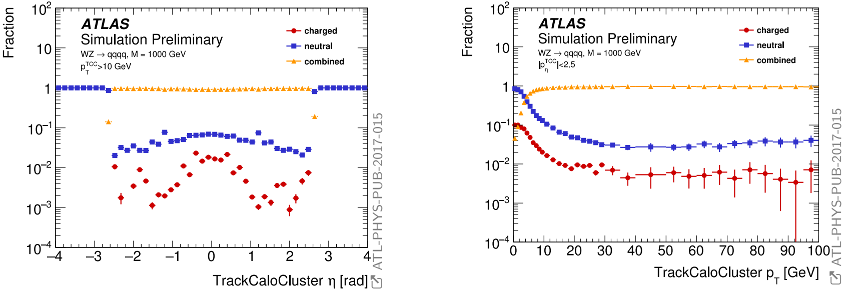
ATL-PHYS-PUB-2017-015

Make use of excellent angular resolution of track for substructure

- Resolution-based track-to-cluster matching

$$\Delta R < \sqrt{\sigma^2_{\text{cluster}} + \sigma^2_{\text{track}}}$$

- resulting in 3 different constituents:
  - **combined:** clusters matched to tracks from primary vertex (PV)
  - **charged:** tracks from PV not matched to any cluster
  - **neutral:** clusters not matched to any track (from the PV)
  - Clusters matched to tracks from PU vertices are **discarded**



Make use of excellent angular resolution of track for substructure

- Each track can be part of multiple **combined** objects
- And any **combined** object can include many tracks
- But each track τ defines only one TCC with the 4-vector

$$p_\tau^{\mathrm{TCC}} = (p_{\mathrm{T}}[\mathcal{M}_\tau], \eta^\tau, \phi^\tau, m[\mathcal{M}_\tau])$$

- η, ϕ purely track-based
- p_{T}, m based on TCC energy-sharing equation:

$$\mathcal{M}_\tau = \sum_c p^c f_\tau^c \mathcal{F}_\tau^{c,\tau}$$

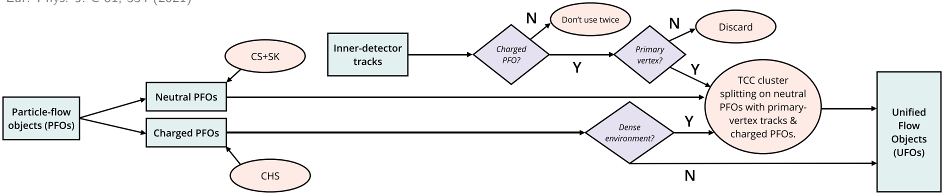
- Sum of momenta p^c of clusters c matched to τ weighted by:

  f_τ^c: how much p_{T} c contributes out of all clusters in τ

  $\mathcal{F}_\tau^{c,\tau}$: how much p_{T} this τ demands out of all τ

Unified Flow Objects (UFO)

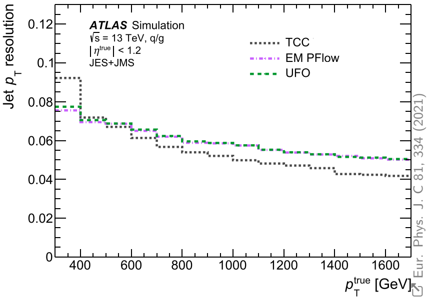
Eur. Phys. J. C 81, 334 (2021)

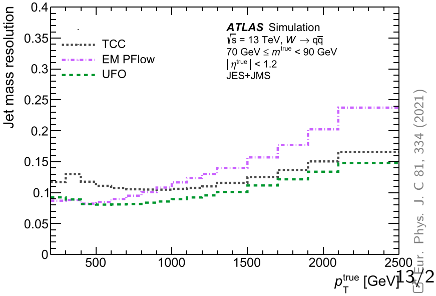


PFlow Shows best jet mass and p_T resolution at low p_T

TCC performs better at high p_T

UFO combines the best of both





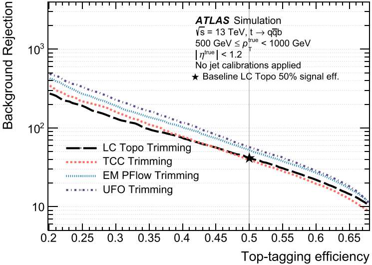
Why UFO Jets?

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

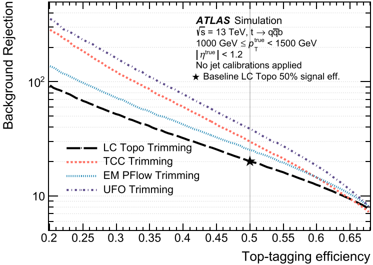
Extensive effort in ATLAS to find best jet definition for tagging:
Eur. Phys. J. C 81, 334 (2021)

- Expected tagger performance evaluated for simple 2-variable cuts:
  - W/Z tagger: m, D_2
  - Top tagger: m, τ_{32}

UFO jets show best performance for simple top tagger:



500 GeV $< p_T^{\text{true}} <$ 1000 GeV



1000 GeV $< p_T^{\text{true}} <$ 1500 GeV

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

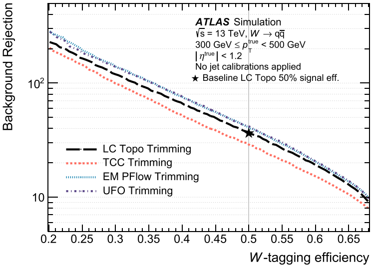
Why UFO Jets?

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

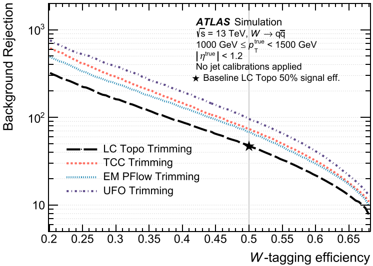
Extensive effort in ATLAS to find best jet definition for tagging:
Eur. Phys. J. C 81, 334 (2021)

- Expected tagger performance evaluated for simple 2-variable cuts:
  - W/Z tagger: m, D_2
  - Top tagger: m, τ_{32}

...as well as simple W tagger:



$500 \text{ GeV} < p_T^{true} < 1000 \text{ GeV}$



$1000 \text{ GeV} < p_T^{true} < 1500 \text{ GeV}$

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

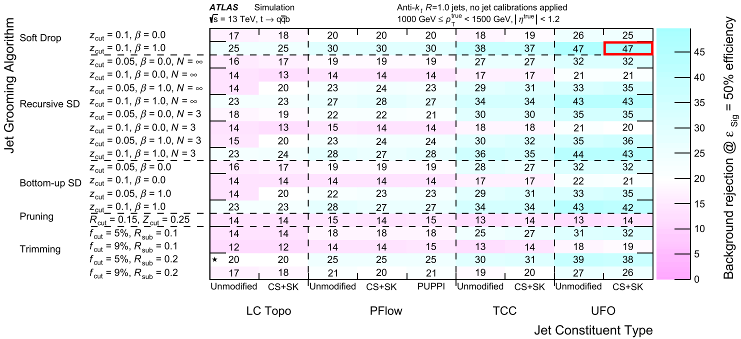
Grooming

Optimisation of Jet Definition

Background rejection for various pileup mitigations and groomings:

Here: 2-variable top tagger, high-p_T range

(plots for W and low-p_T in backup)



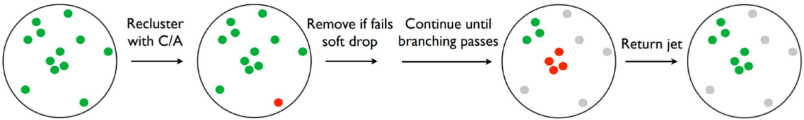
Eur. Phys. J. C 81, 334 (2021)

Best background rejection with:

- $R = 1.0$ anti-k_T **UFO** jets
- Pileup Mitigation: Constituent Subtraction + SoftKiller (**CS+SK**)
- Grooming: Soft Drop (**SD**) with $\beta = 1.0$ $z_{\text{cut}} = 0.1$

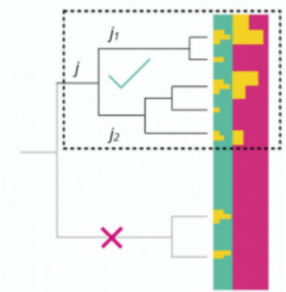
Other factors: Good pileup stability, mass resolution, ...

Grooming: Soft Drop



SVJ workshop 2022

- Re-cluster using Cambridge/Aachen
  (closer constituents first)
- Consider splitting history
- At each split either keep both or reject one branch
- Based on splitting condition:
- Tunable paramters determined empirically: $z_{\text{cut}} = 0.1$, $\beta = 1.0$



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

HCW 2023

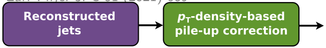
Jet Calibration

Jet Calibration

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

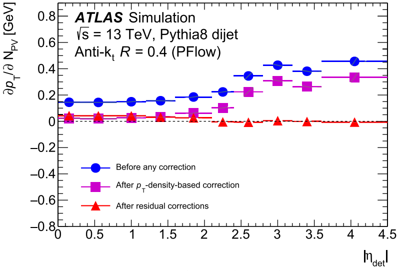
Reconstructed jets →

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

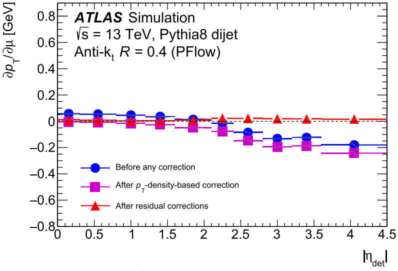


Pileup Correction

$$p_T^{corr} = p_T^{reco} - \rho \times A - \alpha \times N_{PV} - \beta \times \langle\mu\rangle$$



In-time pile-up dependence



Out-of-time pile-up dependence

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

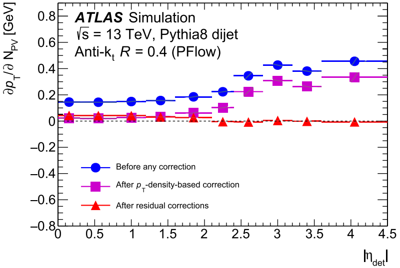
- **Jet-Area based correction**
  - For in-time PU, based on event energy density ρ and jet area A

Jet Calibration

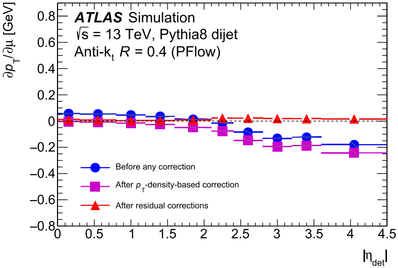
Eur. Phys. J. C 81 (2021) 689

Pileup Correction

$$p_T^{corr} = p_T^{reco} - \rho \times A - \alpha \times N_{PV} - \beta \times \langle \mu \rangle$$



In-time pile-up dependence



Out-of-time pile-up dependence

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

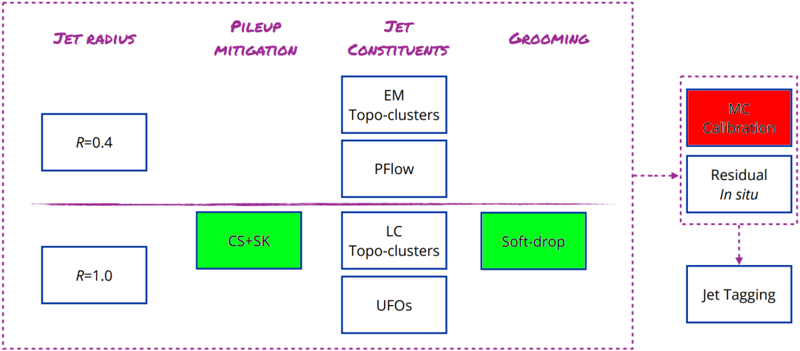
- **Jet-Area based correction**
  - For in-time PU, based on event energy density ρ and jet area A
- **Residual correction**
  - Based on number of primary vertices N_{PV} (in-time) and avg. number of bunch-crossing (out-of-time) over multiple events

Jet Calibration

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

PU correction applied to small (R=0.4) jets only

Large (R=1.0) jets: **CS+SK PU mitigation + SoftDrop**
before MC calibration instead



adapted: Matt LeBlanc, SVJ workshop 2022

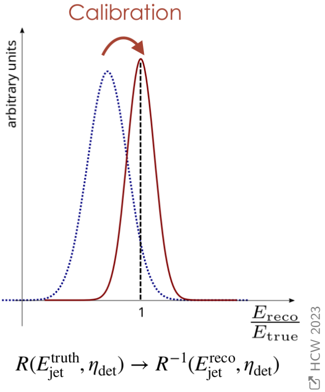
Jet Calibration

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

Reconstructed jets → p_T-density-based pile-up correction → Residual pile-up correction → Absolute MC-based calibration →

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

- Calculate E response in bins of η and E_{true} in MC
- Numerical inversion yields calibration factors
- Origin correction corrects jet η
- Largest calibration step that brings response on average to 1



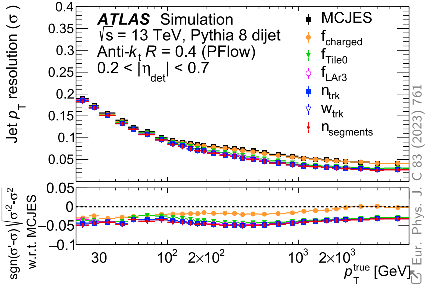
$$R(E_{\text{jet}}^{\text{truth}}, \eta_{\text{det}}) \rightarrow R^{-1}(E_{\text{jet}}^{\text{reco}}, \eta_{\text{det}})$$

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 →

p_T **resolution** after each GSC step



Global Sequential Calibration

- After energy scale calibrated on average, **GSC** corrects for small differences
- E.g. for different jet flavours
- **Sequentially** corrects for each variable
- Only for small (R=0.4) jets

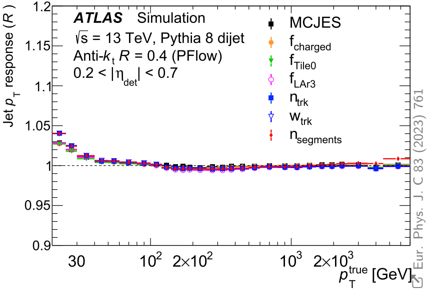
GSC improves JER by applying different corrections for different population of jets (e.g. q/g initiated)

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 →

p_T **response** after each GSC step



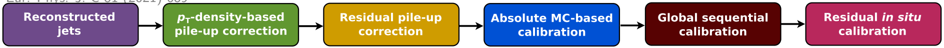
Global Sequential Calibration

- After energy scale calibrated on average, **GSC** corrects for small differences
- E.g. for different jet flavours
- **Sequentially** corrects for each variable
- Only for small (R=0.4) jets

GSC improves JER by applying different corrections for different population of jets (e.g. q/g initiated) but leaves JES on average the same

Jet Calibration

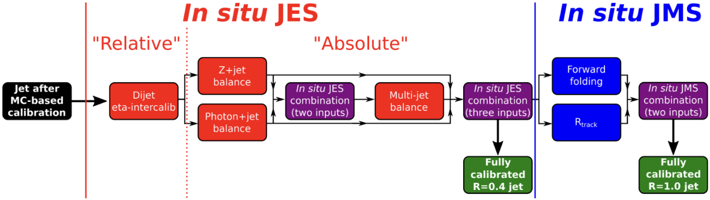
Eur. Phys. J. C 81 (2021) 689



In-situ calibration in data

Corrects jets with high uncertainty (e.g. forward)

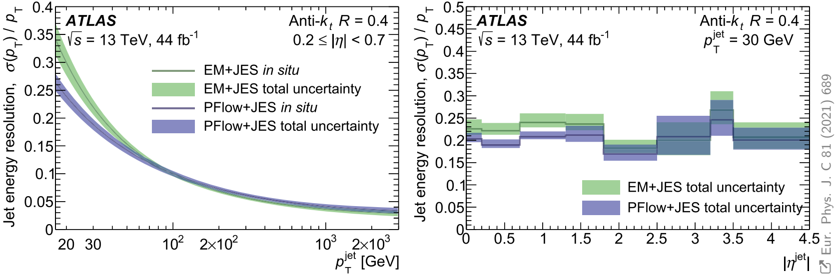
based on well-known (photons, central jets...) objects



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

Jet Energy Resolution (JER)
after full calibration
for EMTopo and PFlow R=0.4 jets

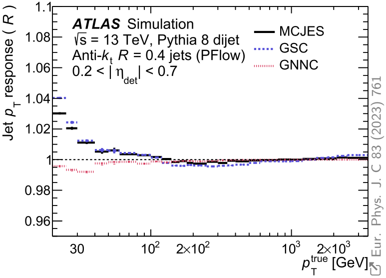


Recent Developments:
ML Jet Calibration

Jet Calibration: GNNC

ATLAS EXPERIMENT

Response

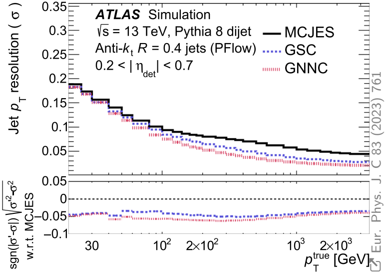


Global NN Calibration (GNNC)

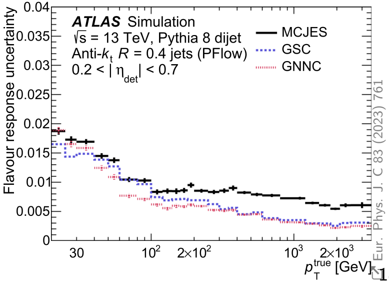
- **GSC** Does not exploit correlations of variables
- New method (**GNNC**) uses MLP trained to predict p_T response

→ Improvement over full p_T range

Resolution

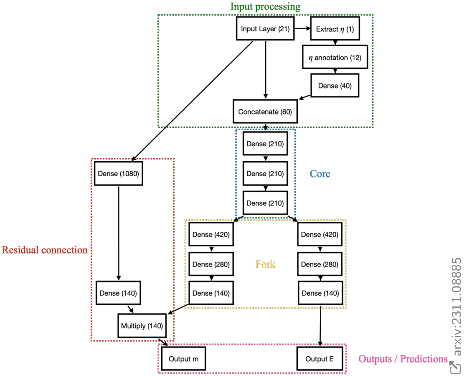


Flavour Uncertainty



Simultaneous Calibration of Jet Energy and Mass using ML

Model Architecture



Method:

- Predict responses

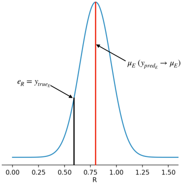
$$R_E = \frac{E_{\text{reco}}}{E_{\text{true}}}, \quad R_M = \frac{M_{\text{reco}}}{M_{\text{true}}}$$

- Modeled by Gaussians

$$y_{\text{pred}} = (\mu^E, \sigma^E, \mu^m, \sigma^m)_{\text{pred}}$$

$\Rightarrow$ Calibration Factors:

$$E_{\text{calib}} = \frac{E_{\text{reco}}}{\mu^E_{\text{pred}}}, \quad M_{\text{calib}} = \frac{M_{\text{reco}}}{\mu^M_{\text{pred}}}$$

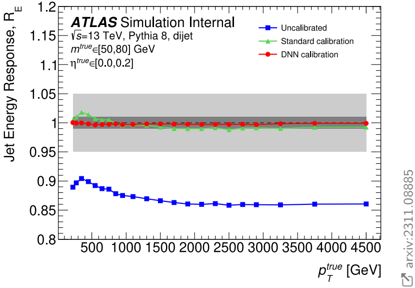


Mixture density network (MDN) loss:

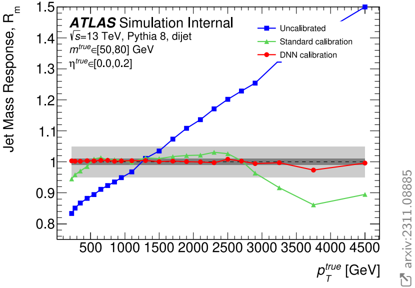
$$\mathcal{L}_{\text{MDN}} = -\log(P(y_{\text{true}}, y_{\text{pred}})) = \log(\sigma_{\text{pred}}) + \frac{1}{2}\frac{(y_{\text{true}} - \mu_{\text{pred}})^2}{\sigma^2_{\text{pred}}}$$

LargeR DNN Calibration: Results

Response: E



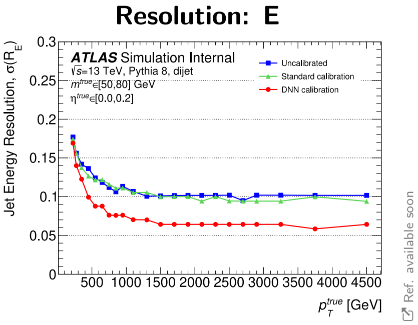
Response: M

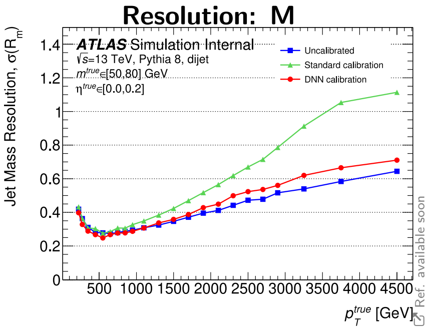


Improvement across the board

- **DNN**: better closure than **standard** calib. in response for E and M
- M response stable even in low and high p_T regime
- Resolution drastically improved
- Less dependence on η, pileup, MC generator for E and M
- More stable across different processes (H, W/Z, top) for E and M
- More stable across different flavours (q/g) for E and M

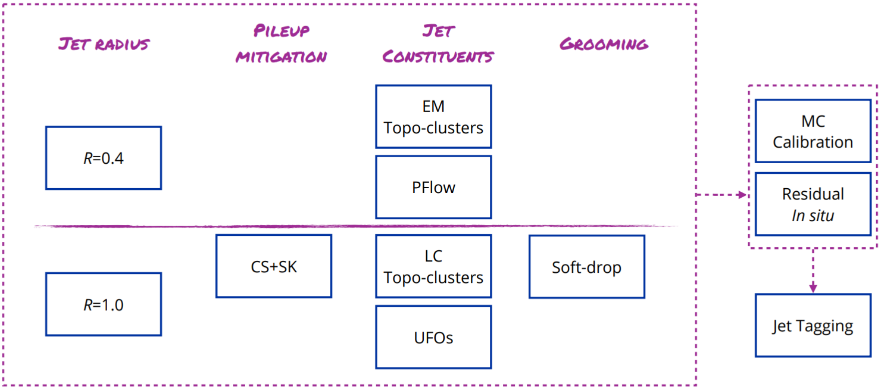
LargeR DNN Calibration: Results





Improvement across the board

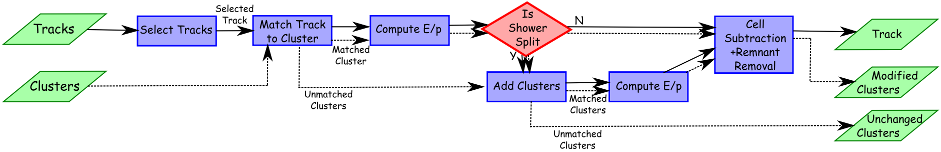
- **DNN**: better closure than **standard** calib. in response for E and M
- M response stable even in low and high p_T regime
- Resolution drastically improved
- Less dependence on η, pileup, MC generator for E and M
- More stable across different processes (H, W/Z, top) for E and M
- More stable across different flavours (q/g) for E and M



adapted: Matt LeBlanc, SVJ workshop 2022

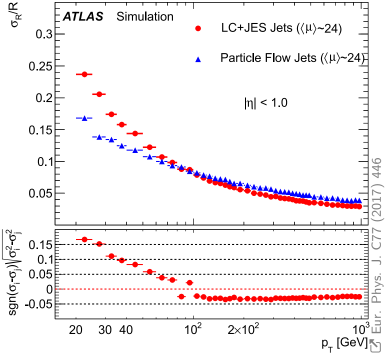
Appendix

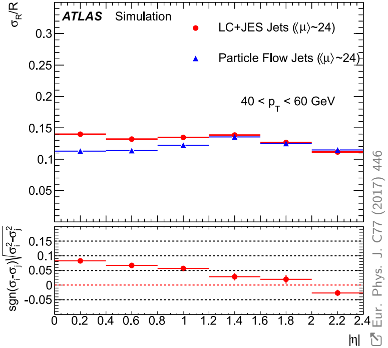
Particle Flow (PFlow)



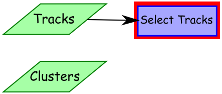
PFlow makes use of tracking information at constituent level shows great JER improvement over calo jets in low-pT

Especially in the central region



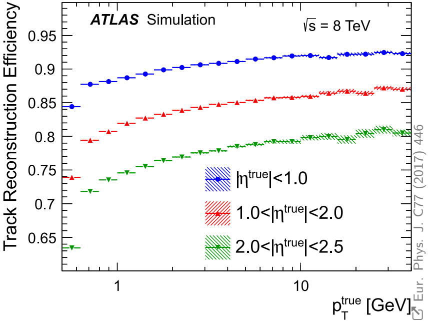


How to Make a PFlow Jet

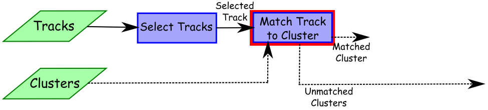


Track Selection

- ≥ 9 hits in Si detectors
- No missing pixels in track
- $|\eta| < 2.5,\ 0.5 > p_{\mathrm{T}} > 40$ GeV
- Not matched to e or μ



How to Make a PFlow Jet



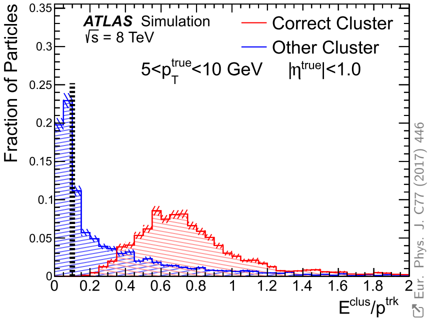
Track-cluster matching

- Matched to cluster with minimum distance metric

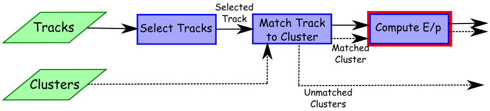
$$\Delta R' = \sqrt{\left(\frac{\Delta\Phi}{\sigma_\Phi}\right)^2 + \left(\frac{\Delta\eta}{\sigma_\eta}\right)^2} < 1.64$$

σ: ang. cluster widths

- And $\frac{E^{\text{clus}}}{p^{\text{trk}}} > 0.1$



How to Make a PFlow Jet

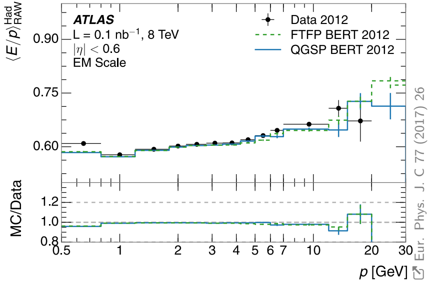


E/p Correction

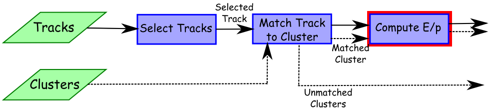
- Avg deposited energy of particle: $\langle E_{\text{dep}} \rangle = p^{\text{trk}} \langle E^{\text{clus}}_{\text{ref}} / p^{\text{trk}}_{\text{ref}} \rangle$
- $\langle E^{\text{clus}}_{\text{ref}} / p^{\text{trk}}_{\text{ref}} \rangle$ measured in isolated single π
- Sum E of clusters in $\Delta R = 0.4$ cone around track
- Binned in $p_T^{\text{trk}}, \eta^{\text{trk}}$, LHED (Layer of Highest Density)

low-p_T:

in isolated single hadrons inclusive *pp* collisions



How to Make a PFlow Jet

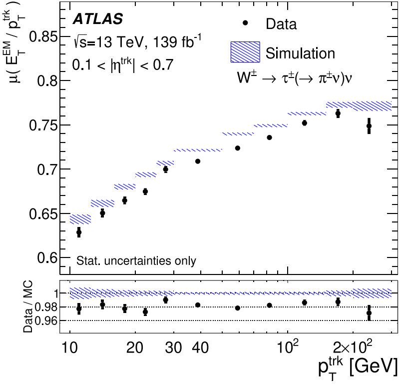


E/p Correction

- Avg deposited energy of particle: $\langle E_{\text{dep}} \rangle = p^{\text{trk}} \langle E^{\text{clus}}_{\text{ref}} / p^{\text{trk}}_{\text{ref}} \rangle$
- $\langle E^{\text{clus}}_{\text{ref}} / p^{\text{trk}}_{\text{ref}} \rangle$ measured in isolated single π
- Sum E of clusters in $\Delta R = 0.4$ cone around track
- Binned in $p_{\text{T}}^{\text{trk}}, \eta^{\text{trk}}$, LHED (Layer of Highest Density)

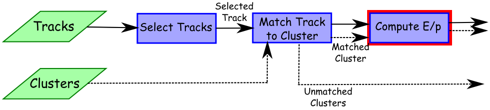
higher $p_{\mathbf{T}}$:

in $W \to \tau(\to \pi\nu)\nu$



Eur. Phys. J. C 82 (2022) 223

How to Make a PFlow Jet



Layer of highest Density (LHED)

- Energy density of jth cell in ith calo layer:

$$\rho_{ij} = \frac{E_{ij}}{V_{ij}} \left(\mathrm{GeV}/X_0^3\right)$$

E: energy, V : volume of cell measured in rad length X

- Weighted based on proximit to track by gaussian with width $\Delta R = 0.035$

- Avg E density for each layer:

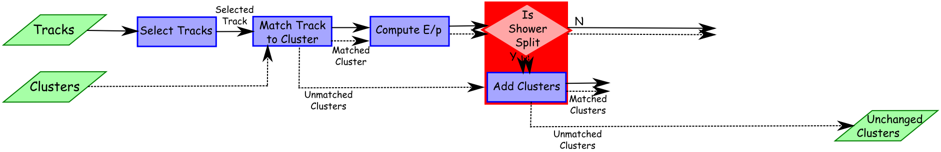
$$\langle\rho'\rangle_i = \sum_j w_{ij}\rho_{ij}$$

$\rightarrow$ LHED is layer with max change of ρ':

$$\Delta\rho'_i = \frac{\langle\rho'\rangle_i - \langle\rho'\rangle_{i-1}}{d_i - d_{i-1}}$$

d_i: depth of layer i

How to Make a PFlow Jet

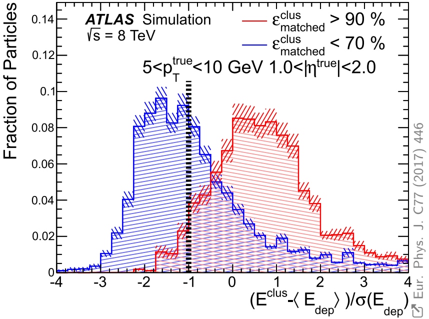


Recover Split Showers

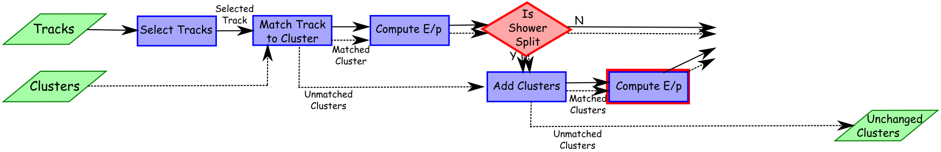
- Often particles deposit energy in more than 1 cluster
- If single/multi cluster discriminant:

$$S(E^{\text{clus}}) = \frac{E^{\text{clus}} - \langle E_{\text{dep}} \rangle}{\sigma(E_{\text{dep}})} < -1$$

recover clusters within $\Delta R < 0.2$ of track



How to Make a PFlow Jet

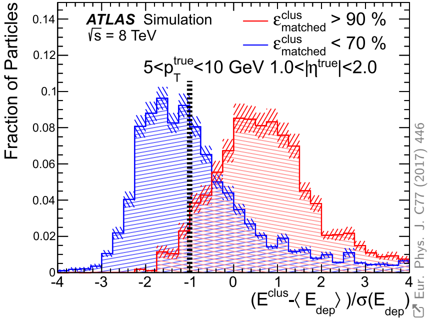


Recover Split Showers

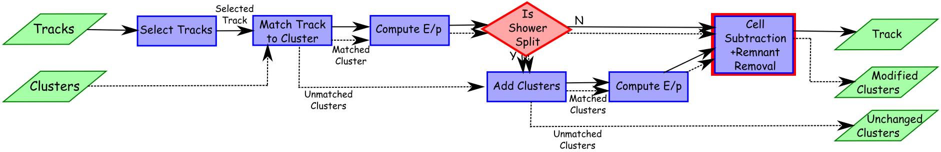
- Often particles deposit energy in more than 1 cluster
- If single/multi cluster discriminant:

$$S(E^{\text{clus}}) = \frac{E^{\text{clus}} - \langle E_{\text{dep}} \rangle}{\sigma(E_{\text{dep}})} < -1$$

recover clusters within $\Delta R < 0.2$ of track



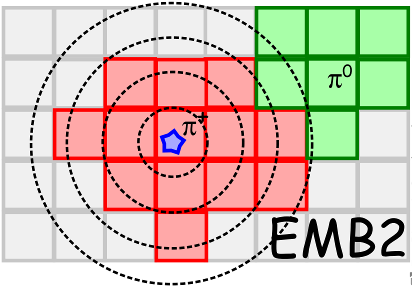
How to Make a PFlow Jet



Cell-by-cell subtraction

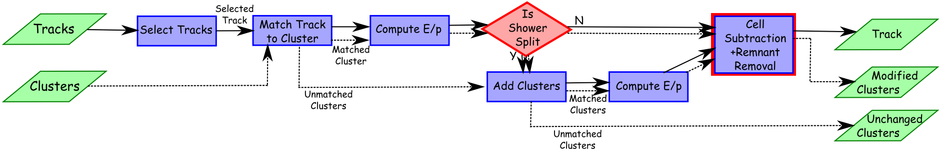
- If $\langle E_{\mathrm{dep}} \rangle = p^{\mathrm{trk}} \langle E^{\mathrm{clus}}_{\mathrm{ref}} / p^{\mathrm{trk}}_{\mathrm{ref}} \rangle$ after correction $> \sum_{i\ \mathrm{matched}} E^{\mathrm{clus}}_{i}$: all clusters are removed

cell-by-cell subtraction



Eur.Phys. J. C77 (2017) 466

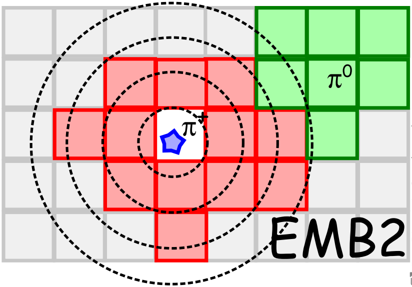
How to Make a PFlow Jet



Cell-by-cell subtraction

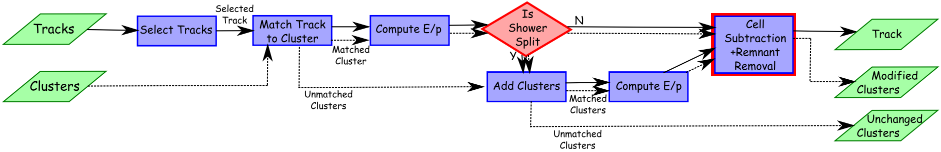
- If $\langle E_{\text{dep}} \rangle = p^{\text{trk}} \langle E^{\text{clus}}_{\text{ref}} / p^{\text{trk}}_{\text{ref}} \rangle$ after correction $> \sum_{i \text{ matched}} E_i^{\text{clus}}$: all clusters are removed
- Else: clusters in rings around track subtracted from highest-to-lowest energy density
- In each layer, starting in LHED
- Until $E_{\text{after subtr}} < \langle E_{\text{dep}} \rangle$

cell-by-cell subtraction



Eur.Phys. J. C77 (2017) 466

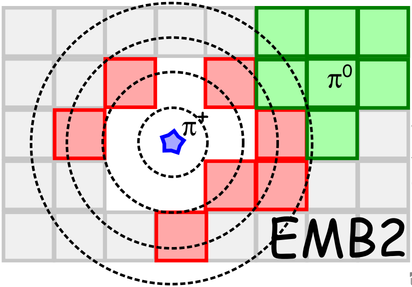
How to Make a PFlow Jet



Cell-by-cell subtraction

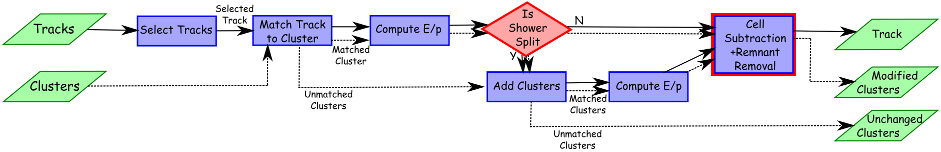
- If $\langle E_{\text{dep}} \rangle = p^{\text{trk}} \langle E_{\text{ref}}^{\text{clus}} / p_{\text{ref}}^{\text{trk}} \rangle$ after correction $> \sum_{i \text{ matched}} E_i^{\text{clus}}$: all clusters are removed
- Else: clusters in rings around track subtracted from highest-to-lowest energy density
- In each layer, starting in LHED
- Until $E_{\text{after subtr}} < \langle E_{\text{dep}} \rangle$

cell-by-cell subtraction



Eur.Phys. J. C77 (2017) 466

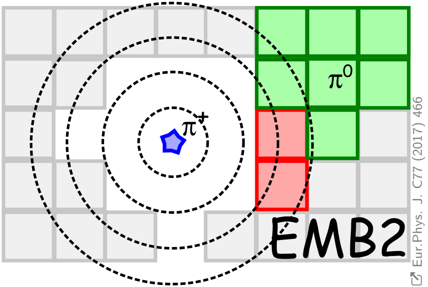
How to Make a PFlow Jet



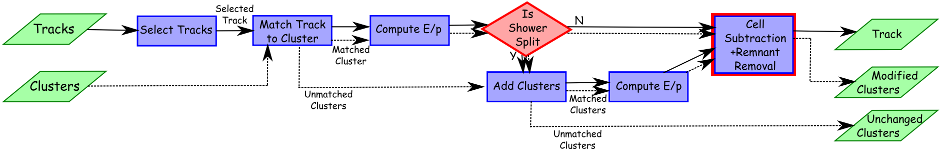
Cell-by-cell subtraction

- If $\langle E_{\mathrm{dep}} \rangle = p^{\mathrm{trk}} \langle E^{\mathrm{clus}}_{\mathrm{ref}} / p^{\mathrm{trk}}_{\mathrm{ref}} \rangle$ after correction $> \sum_{i\text{ matched}} E^{\mathrm{clus}}_i$: all clusters are removed
- Else: clusters in rings around track subtracted from highest-to-lowest energy density
- In each layer, starting in LHED
- Until $E_{\text{after subtr}} < \langle E_{\mathrm{dep}} \rangle$

cell-by-cell subtraction



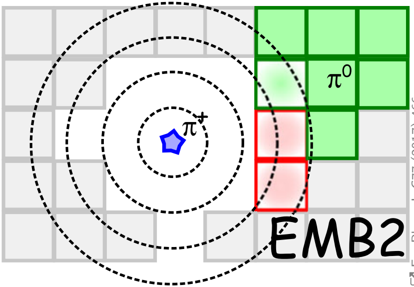
How to Make a PFlow Jet



Cell-by-cell subtraction

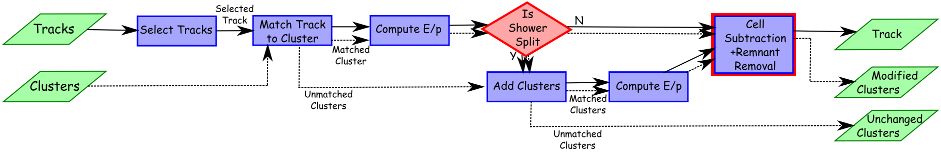
- If $\langle E_{\text{dep}} \rangle = p^{\text{trk}} \langle E^{\text{clus}}_{\text{ref}} / p^{\text{trk}}_{\text{ref}} \rangle$ after correction $> \sum_{i \text{ matched}} E^{\text{clus}}_{i}$: all clusters are removed
- Else: clusters in rings around track subtracted from highest-to-lowest energy density
- In each layer, starting in LHED
- Until $E_{\text{after subtr}} < \langle E_{\text{dep}} \rangle$
- Then scale cluster energies accordingly

cell-by-cell subtraction



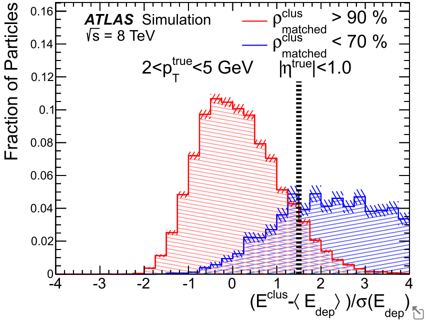
Eur.Phys. J. C77 (2017) 466

How to Make a PFlow Jet

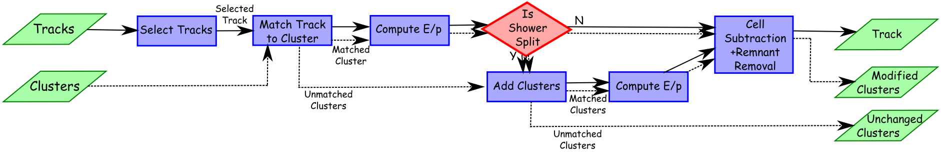


Remnant removal

- If remainging cell $E < 1.5\sigma$ of width of $p^{\mathrm{trk}}\langle E^{\mathrm{clus}}_{\mathrm{ref}}/p^{\mathrm{trk}}_{\mathrm{ref}}\rangle$:
  - Cluster-system likely produced by single particle
  - $\rightarrow$ Remaining E removed
- Else:
  - likely produced by multiple particles
  - $\rightarrow$ Remaining E retained



How to Make a PFlow Jet



Done!

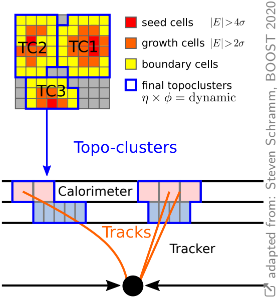
Final constituents: Remaining clusters and tracks

→ **Goal of pflow procedure:** Avoid double-counting between them

Why UFO Jets?

Calorimeter only:

- **LCTopo**: Topological calorimeter clusters



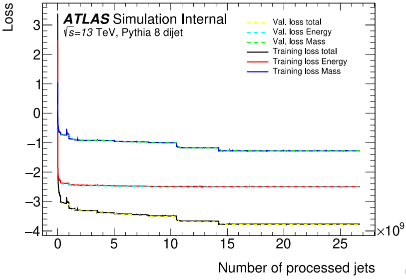
Combined with tracking:

- **PFlow:** Particle Flow Objects
  - Low p_T: Use track 4-vector for charged particles, subtract energy from cluster 4-vectors
  - High p_T: Use cluster 4-vectors, ignore tracks
- **TCC**: Track Calo Clusters
  - Low p_T: Use cluster 4-vectors, ignore tracks
  - High p_T: Split clusters using tracks, get energy from clusters but angles from tracks

Combining PFlow and TCC:

- **UFO** combines **TCC** and **PFlow** to achieve optimal performance over a broad kinematic (p_T) range

LargeR DNN Calibration: Training



Ref. available soon

Steps	N°	Number of epochs	Batch size	Loss
Initialisation	1	2	15000	MDNA
	2	2	25000	MDNA
	3	2	35000	MDNA truncated (4.0σ)
	4	2	15000	MDNA truncated (3.5σ)
Common training	5	6	95000	MDNA truncated (3.5σ)
	6	6	95000	MDNA truncated (3.5σ)
	7	6	125000	MDNA truncated (3.2σ)
	8	6	125000	MDNA truncated (3.2σ)
	9	10	155000	MDNA truncated (3.0σ)
	10	15	95000	MDNA truncated (E: 3.0σ, m : 2.0σ)
Exclusive mass training	11	50	95000	MDN truncated (1.0σ)

Training Strategy

- Multi-stage training process
  - First E & M simultaneous
  - Then only M
- Alternative losses used in some training stages
  - To accommodate for asymmetric response:

Asymmetric MDN:

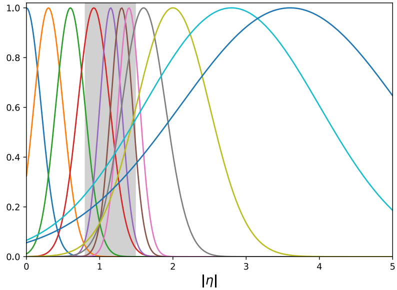
$$P_{\text{MDNA}}(x) = \begin{cases} 1e^{(x-\mu)^2/2\sigma_1} & \text{if} x < \mu \\ 1e^{(x-\mu)^2/2\sigma_2} & \text{if} x \geq \mu \end{cases}$$

Truncated MDN:

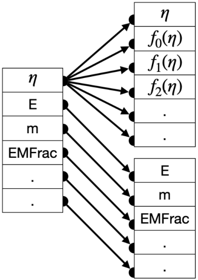
$$P_{\text{trunc}}(x) = \begin{cases} 1e^{(x-\mu)^2/2\sigma} & \text{if}|x < \mu| < N\sigma \\ 0 & \text{otherwise} \end{cases}$$

LargeR DNN Calibration: Eta Annotation

ATLAS EXPERIMENT



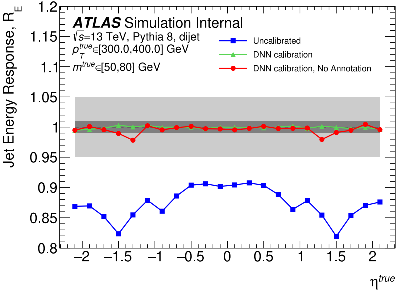
Ref. available soon



Ref. available soon

Complex dependence on η

- With sharp changes from bin-to-bin due to detector geometry/instrumentation
- Difficult for DNN to adapt to this
- Annotation strategy
  - → Add 12 features that are functions of η
  - → Encoding distance to different η regions
- Clear improvement:



Ref. available soon

Substructure Variables for Tagging

W/Z tagger (NN/ANN)

ATL-PHYS-PUB-2021-029

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

Top tagger (DNN)

ATL-PHYS-PUB-2021-028

$\tau_1,\ \tau_2,\ \tau_3,\ \tau_4$	N-subjettiness
$\sqrt{d_{12}},\ \sqrt{d_{23}}$	Splitting scales
$\mathrm{ECF}_1,\ \mathrm{ECF}_2,\ \mathrm{ECF}_3$	Energy correlation (EC) functions
$C_2,\ D_2$	EC ratios
$L_2,\ L_3$	Generalised EC ratios
Q_{W}	Invariant mass / virtuality
T_{M}	Thrust major

Number of Ghost-Associated Tracks n_{trk}

JHEP04(2008)005

n_{trk}: number of tracks

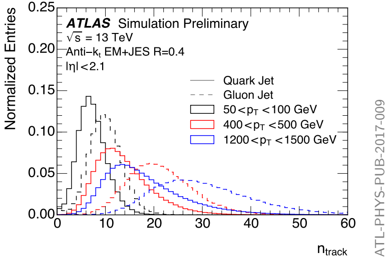
- With $p_T > 500$ MeV
- Ghost-associated to jet

→ Powerful q/g discriminant

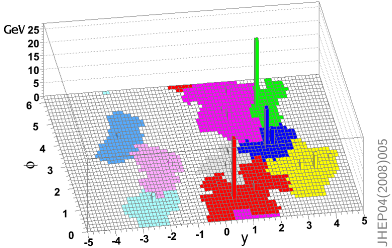
Ghost-associated jet area

- Add dense coverage of 'infinitely' soft 'ghost' constituents
- Count how many are clustered within the jet

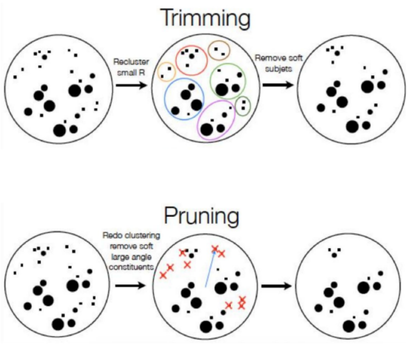
n_{trk} as q/g discriminant

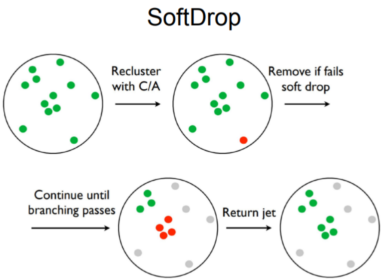


Ghost associated areas of k_t jets



Grooming Techniques





SVJ workshop 2022