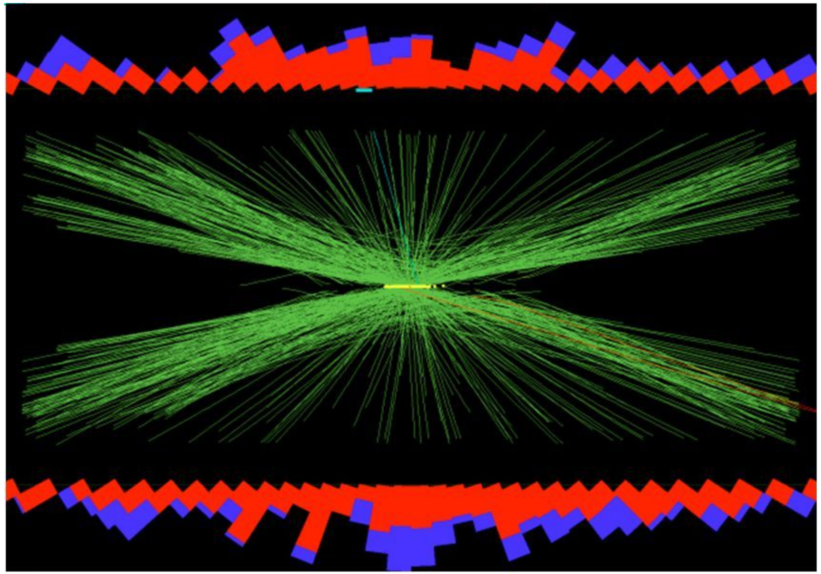
Particle Flow and Jet Reconstruction

Luke Grossman

Physics 290e, 2023-11-08

Contents

- What do we mean by jet reconstruction?
- Yet another ATLAS detector overview
- Tracker vs Calorimeter
- Particle Flow
  - Topo-clusters
  - Track matching
  - Energy subtraction
- Jet reconstruction performance
- Track-CaloClusters
- Newest methods: UFOs



Jet Reconstruction

- Jets are a detector object, especially post-Snowmass
- “Jet reconstruction” algorithms must run on some set of inputs:
  - 4-momentum, pT, or equivalent
  - Spatial track/vertexing
- Some parameters (R) are “free” parameters, but ideally jet reconstruction inputs would work well with them

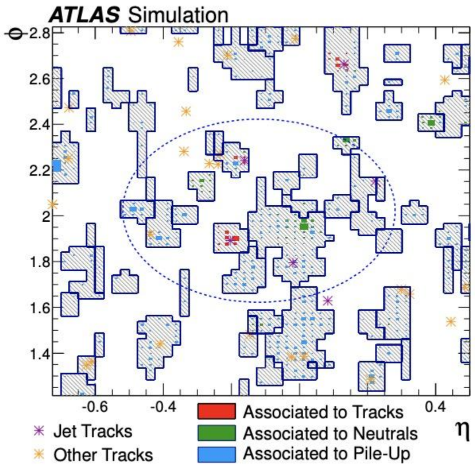
$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}$$

https://doi.org/10.48550/arXiv.1703.10485

Jet Reconstruction

- Jets are a detector object, especially post-Snowmass
- “Jet reconstruction” algorithms must run on some set of inputs:
  - 4-momentum, pT, or equivalent
  - Spatial track/vertexing
- Some parameters (R) are “free” parameters, but ideally jet reconstruction inputs would work well with them
- ATLAS detector data is not a list of particle tracks!
  - Different subdetectors have no correspondence a priori
  - Calorimeter signals in particular are complicated energy distributions with many overlapping hits
  - Only charged particles are “tracked” at all – neutral particles only appear in the calorimeters
  - Calorimeter has no way of separating pileup events

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



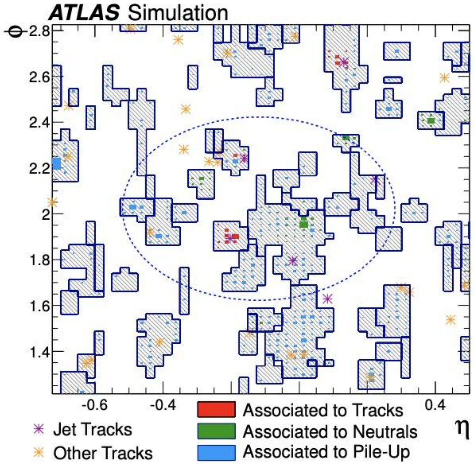
https://doi.org/10.48550/arXiv.1703.10485

Jet Reconstruction

- Jets are a detector object, especially post-Snowmass
- “Jet reconstruction” algorithms must run on some set of inputs:
    - 4-momentum, pT, or equivalent
    - Spatial track/vertexing
- Some parameters (R) are “free” parameters, but ideally jet reconstruction inputs would work well with them
- ATLAS detector data is not a list of particle tracks!
    - Different subdetectors have no correspondence a priori
    - Calorimeter signals in particular are complicated energy distributions with many overlapping hits
    - Only charged particles are “tracked” at all – neutral particles only appear in the calorimeters
    - Calorimeter has no way of separating pileup events

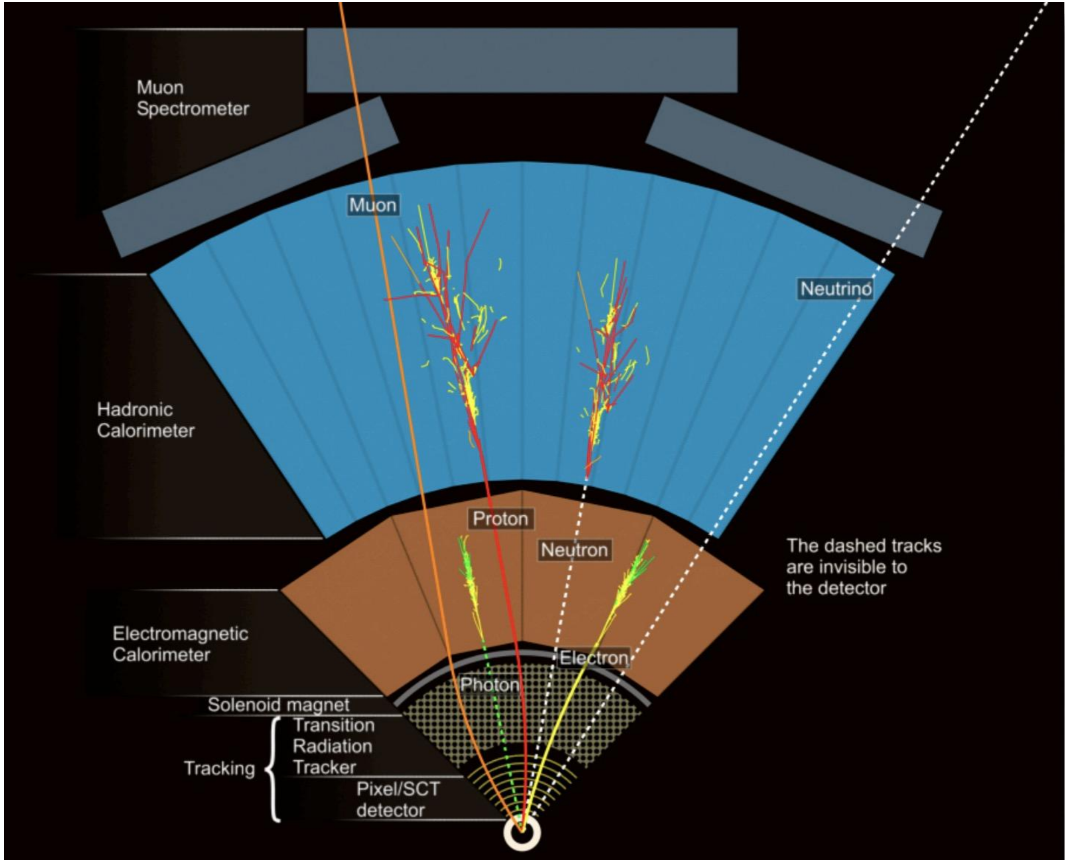
$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}$$

?



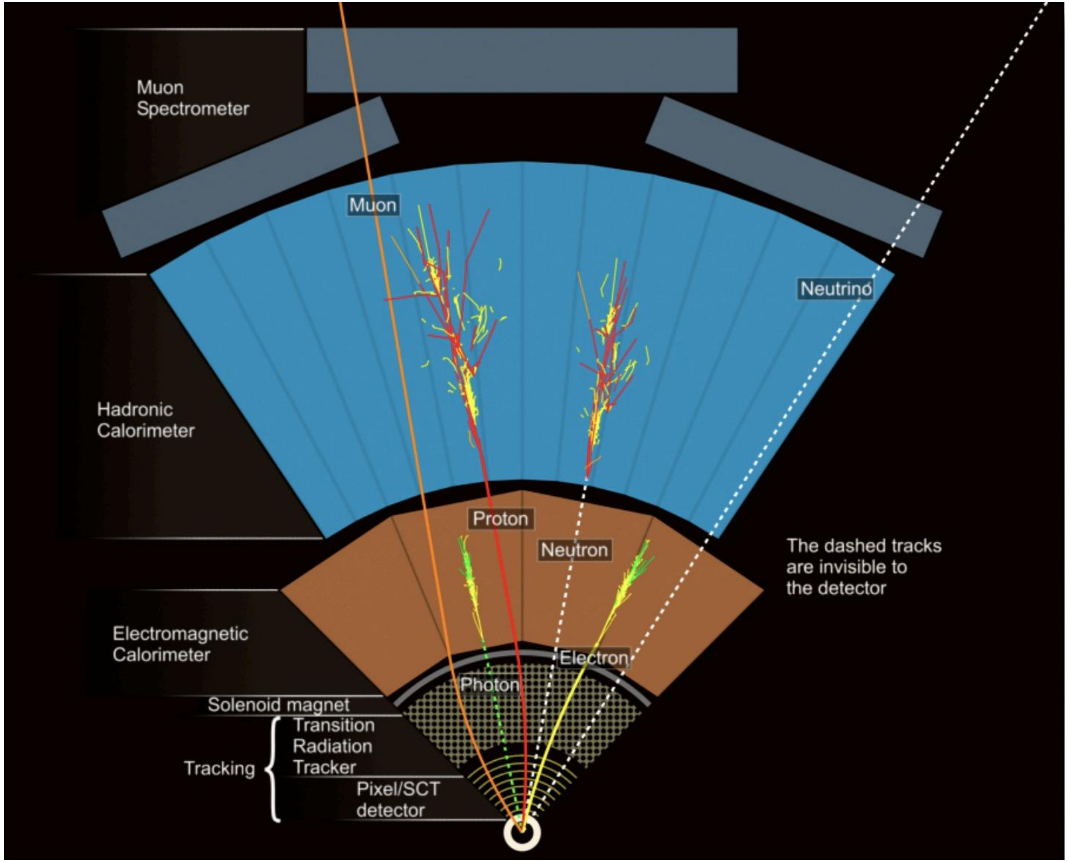
https://doi.org/10.48550/arXiv.1703.10485

Detector Overview



We will focus mainly on the Inner Detector (ID) considered as a whole, and the Calorimeters considered as a whole – for our purposes the EM vs Hadronic calorimeters are like different layers of the same system.

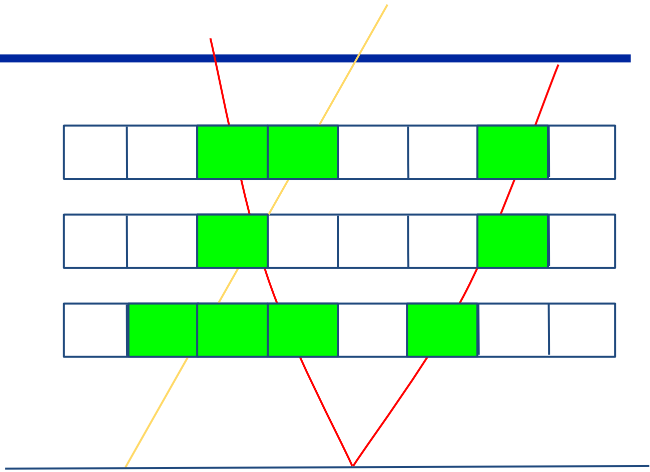
Detector Overview



We will focus mainly on the Inner Detector (ID) considered as a whole, and the Calorimeters considered as a whole – for our purposes the EM vs Hadronic calorimeters are like different layers of the same system.

I'm only going 1-2 levels deep into the processing, so I will ignore e.g. spillover between EM, Had, and muon; Jet energy scale setting; and how ID track reconstruction works

Tracking



- Inner detector covers the region |η| < 2.5
- Charged particles excite electron-hole pairs driven to electrodes
- 4 pixel layers and 8 microstrip layers (crossed)
  - Typically 3 pixel and 4 strip hits are needed for a track
- Transverse momentum resolution scales ~pT, minimum of 500 MeV.

$$\frac{\sigma_{p_T}}{p_T} = 0.032\% \cdot p_T \oplus \frac{1.30\%}{\sqrt{\sin\theta}}$$

- Very good spatial resolution
  - Vertex association typically done at 2 mm
- Before any correspondence with calorimeters, track reconstruction is performed to transform pixel+strip hits into separated, individual tracks

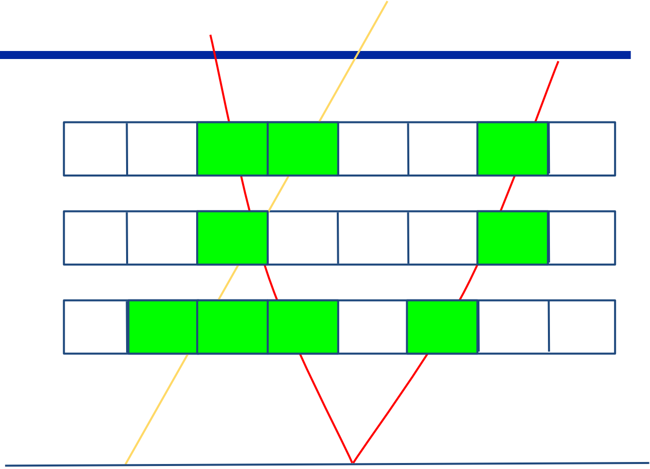
https://doi.org/10.48550/arXiv.1703.10485

Tracking

- Inner detector covers the region |η| < 2.5
- Charged particles excite electron-hole pairs driven to electrodes
- 4 pixel layers and 8 microstrip layers (crossed)
  - Typically 3 pixel and 4 strip hits are needed for a track
- Transverse momentum resolution scales ~pT, minimum of 500 MeV.

$$\frac{\sigma_{p_T}}{p_T} = 0.032\% \cdot p_T \oplus \frac{1.30\%}{\sqrt{\sin\theta}}$$

- Very good spatial resolution
  - Vertex association typically done at 2 mm
- Before any correspondence with calorimeters, track reconstruction is performed to transform pixel+strip hits into separated, individual tracks



Output

List of charged particle tracks

- pT from TRT
- Associated vertex
- Spatial path

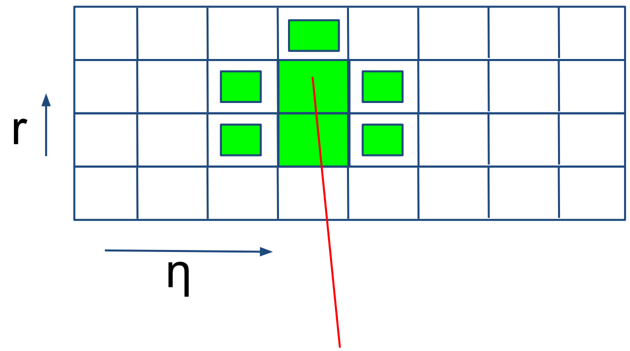
https://doi.org/10.48550/arXiv.1703.10485

Calorimetry

- Calorimeter has much more complicated geometry – generically EM calorimeter surrounded by hadronic calorimeter
    - LAr EM calorimeter, 22 radiation lengths
    - Solid Tile hadronic calorimeter, 9-10 interaction lengths
- Energy resolution scales inverse to E

$$\frac{\sigma_E}{E} = \frac{11.5\%}{\sqrt{E}} \oplus 0.5\%$$

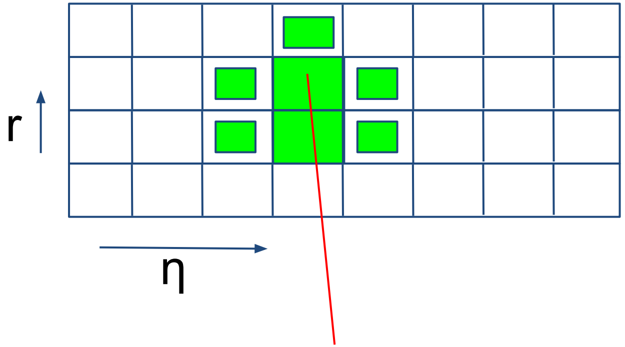
    - Calorimeter better resolution than tracker for hard particles, and vice versa for softer particles
- Each calorimeter segmented into several longitudinal layers, and split in each layer into many ϕ and η sectors.
- "Hits" are defined as energy deposits some number of standard deviations above the typical noise, totaled for a single cell



https://doi.org/10.48550/arXiv.1703.10485

Topo-clusters

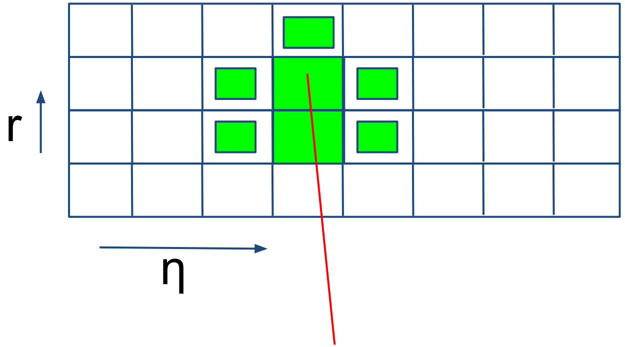
- Step 1 (used to be of 1) in preparing jet reconstruction inputs
- Identify hits:
  - Look for any calorimeter cells above 4σ energy, use them as seeds for a cluster
  - Move outward through all cells above 2σ energy, include them in the cluster
  - Include all boundary cells below 2σ energy
  - Split clusters with more than one energy local maximum
- Not the final output of calorimeters – topo-clusters will continue to be edited as they are made into PFlow objects and eventually UFOs
- In Run 1 topo-clusters were essentially the only input to jet reconstruction!



https://doi.org/10.48550/arXiv.1703.10485

Topo-clusters

- Step 1 (used to be of 1) in preparing jet reconstruction inputs
- Identify hits:
    - Look for any calorimeter cells above 4σ energy, use them as seeds for a cluster
    - Move outward through all cells above 2σ energy, include them in the cluster
    - Include all boundary cells below 2σ energy
    - Split clusters with more than one energy local maximum
- Not the final output of calorimeters – topo-clusters will continue to be edited as they are made into PFlow objects and eventually UFOs
- In Run 1 topo-clusters were essentially the only input to jet reconstruction!



Output

List of topo-clusters

- Energy for all cells
- Directional info

https://doi.org/10.48550/arXiv.1703.10485

Particle Flow

- Due to the limitations of each subdetector, we want to join together tracking and calorimetry to create some unified object
- Ideally it would somewhat “particle-ize” the calorimetry information, associating deposited energy clusters with individual charged-particle tracks

https://doi.org/10.48550/arXiv.1703.10485

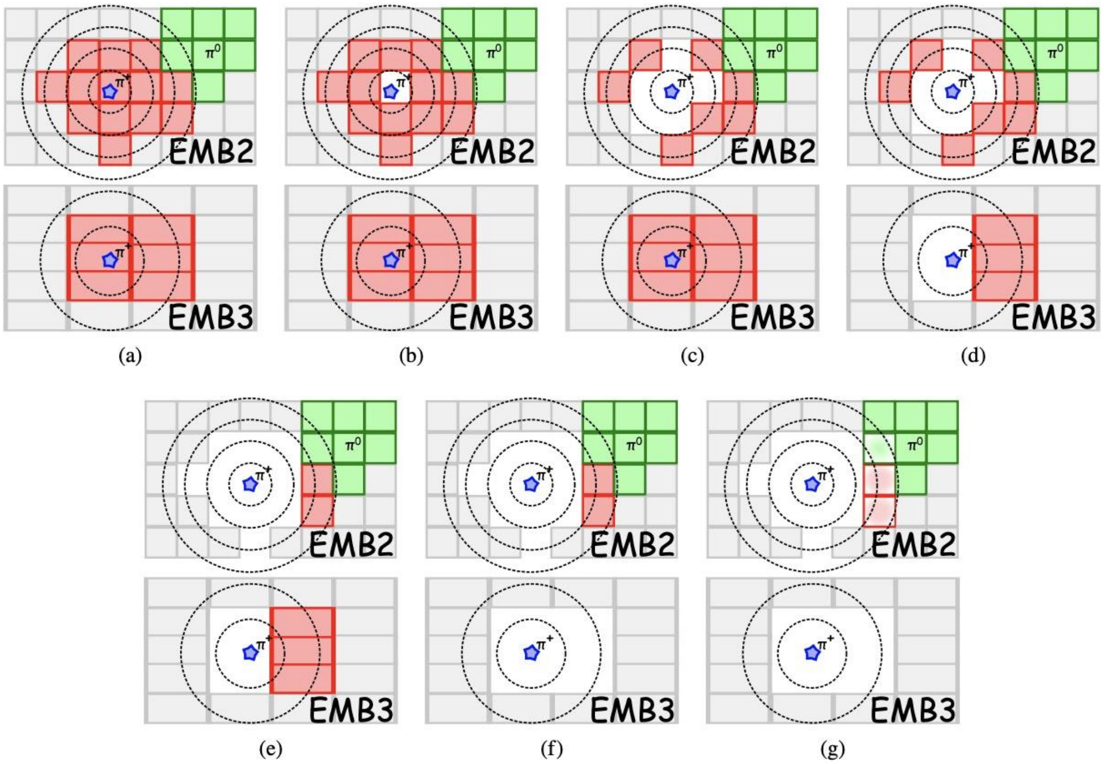
Particle Flow

- Due to the limitations of each subdetector, we want to join together tracking and calorimetry to create some unified object
- Ideally it would somewhat "particle-ize" the calorimetry information, associating deposited energy clusters with individual charged-particle tracks

1. Using ID tracking data, a model of expected calorimeter energy deposition is created for each charged-particle track associated with the primary vertex
2. Angular info is recalculated w.r.t. the PV
3. Charged particles from pileup can be removed from the list of topo-clusters using tracking info
4. PV tracks are matched to topo-clusters using these predictions, and associated energy is subtracted out to prevent double counting
5. Output a group of "Particle Flow Objects" (PFOs) consisting of matched track-cluster pairs and unpaired (neutral) topo-clusters

https://doi.org/10.48550/arXiv.1703.10485

Energy Subtraction

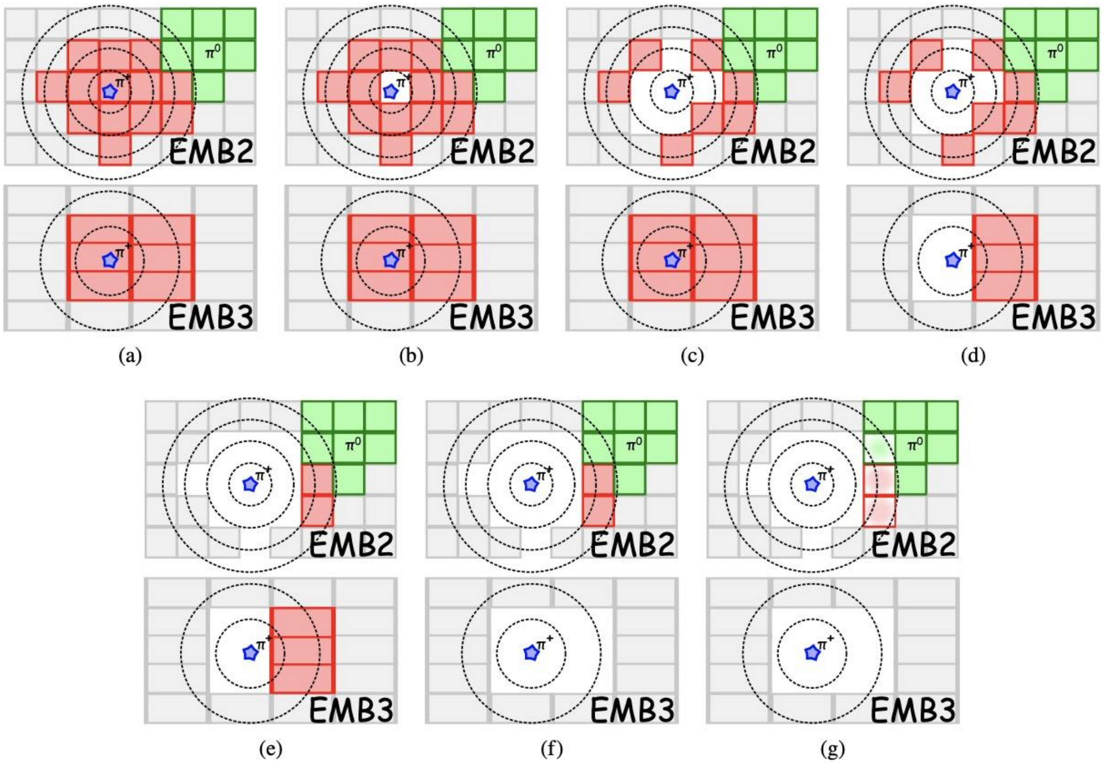


To “associate” track with topo-cluster necessary to remove deposited energy, to avoid double-counting any energy.

The details of track-cluster matching are very complicated – a simple 1:1 matching is shown here.

https://doi.org/10.48550/arXiv.1703.10485

Energy Subtraction



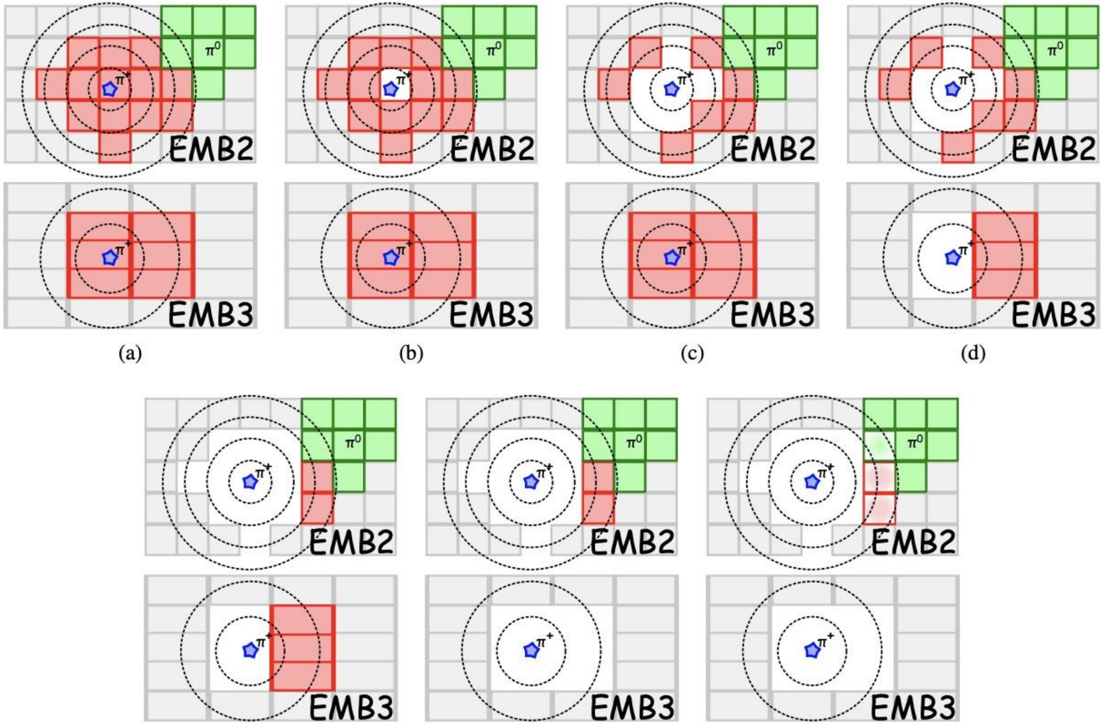
To "associate" track with topo-cluster necessary to remove deposited energy, to avoid double-counting any energy.

The details of track-cluster matching are very complicated – a simple 1:1 matching is shown here.

Energy subtraction proceeds "outwards" until all predicted energy is accounted for.

https://doi.org/10.48550/arXiv.1703.10485

Energy Subtraction



To “associate” track with topo-cluster necessary to remove deposited energy, to avoid double-counting any energy.

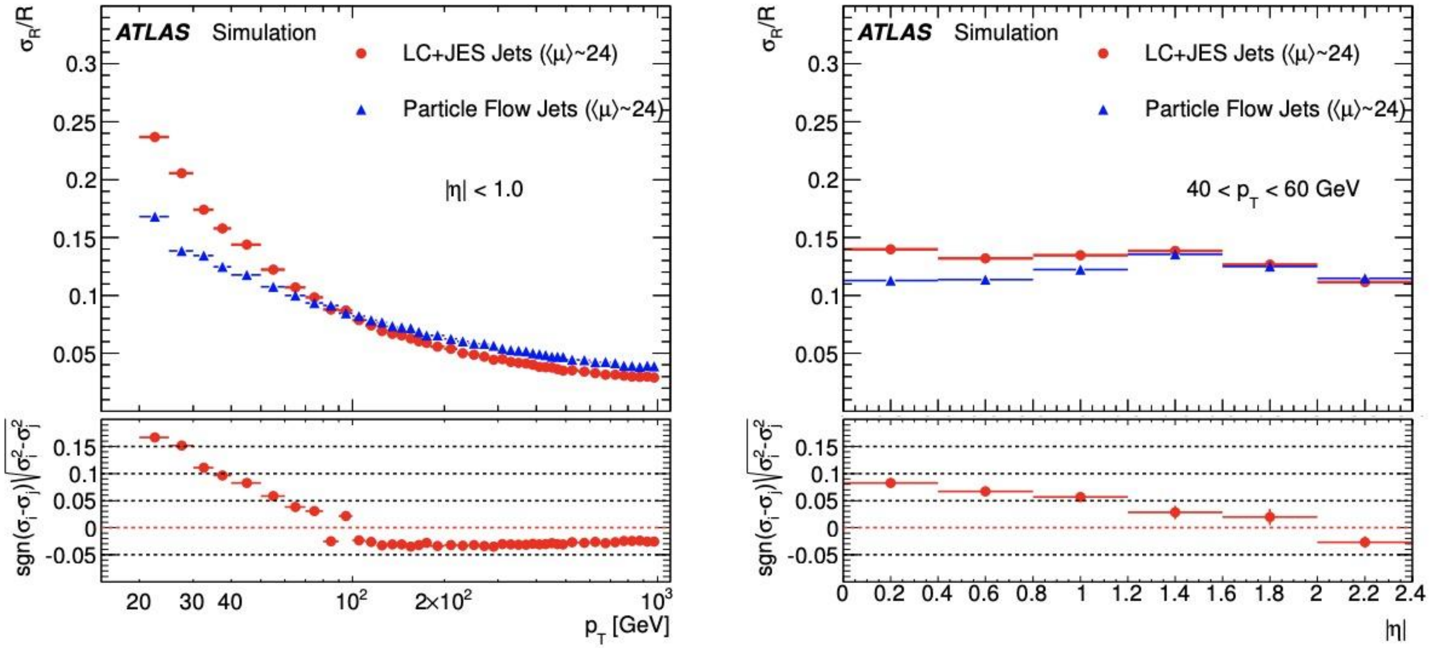
The details of track-cluster matching are very complicated – a simple 1:1 matching is shown here.

Energy subtraction proceeds “outwards” until all predicted energy is accounted for.

In busy environments, extrapolated ID data is less reliable, so there are various cutoffs imposed for heavy calorimeter deposition, and any tracks above 100 GeV are not matched at all.

https://doi.org/10.48550/arXiv.1703.10485

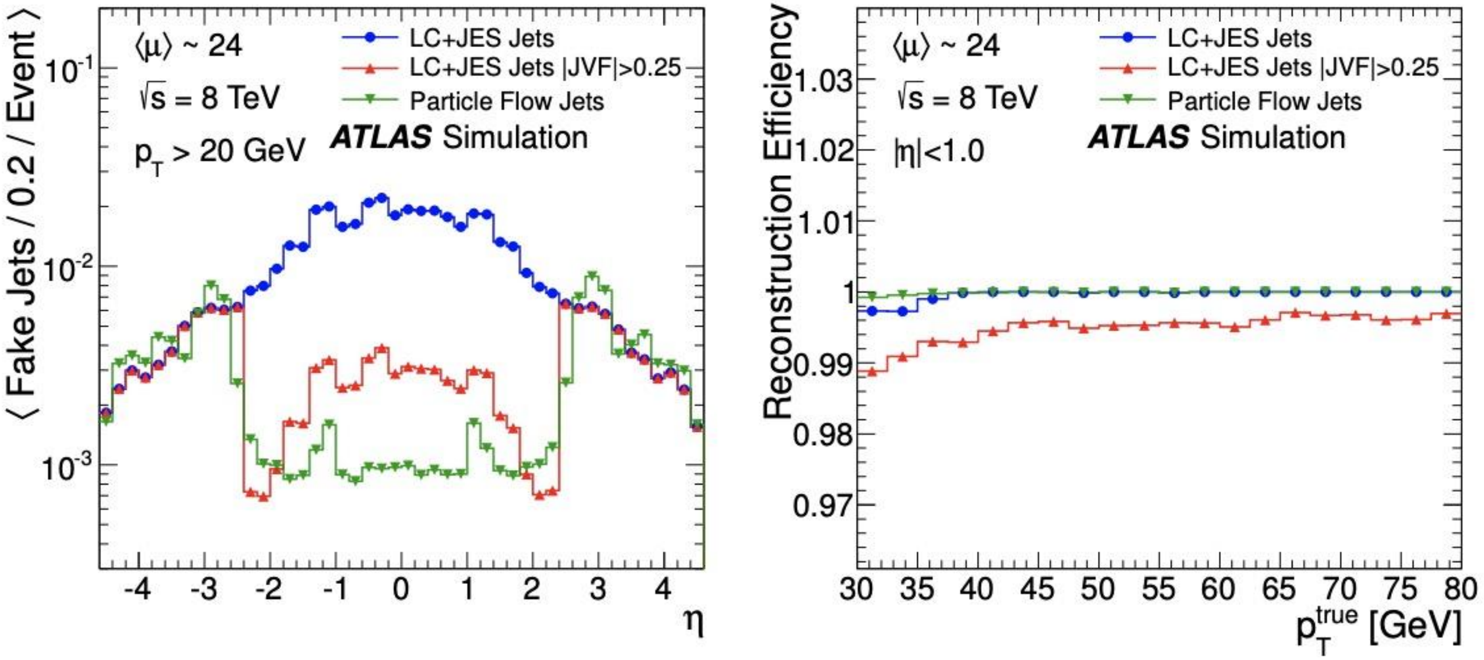
Performance



- Better jet momentum resolution up to ~90 GeV, hampered on the high end by both worse tracking and more overlap in the calorimeter
- Most of this deficit is removed by smoothly disabling PFlow for higher momenta

https://doi.org/10.48550/arXiv.1703.10485

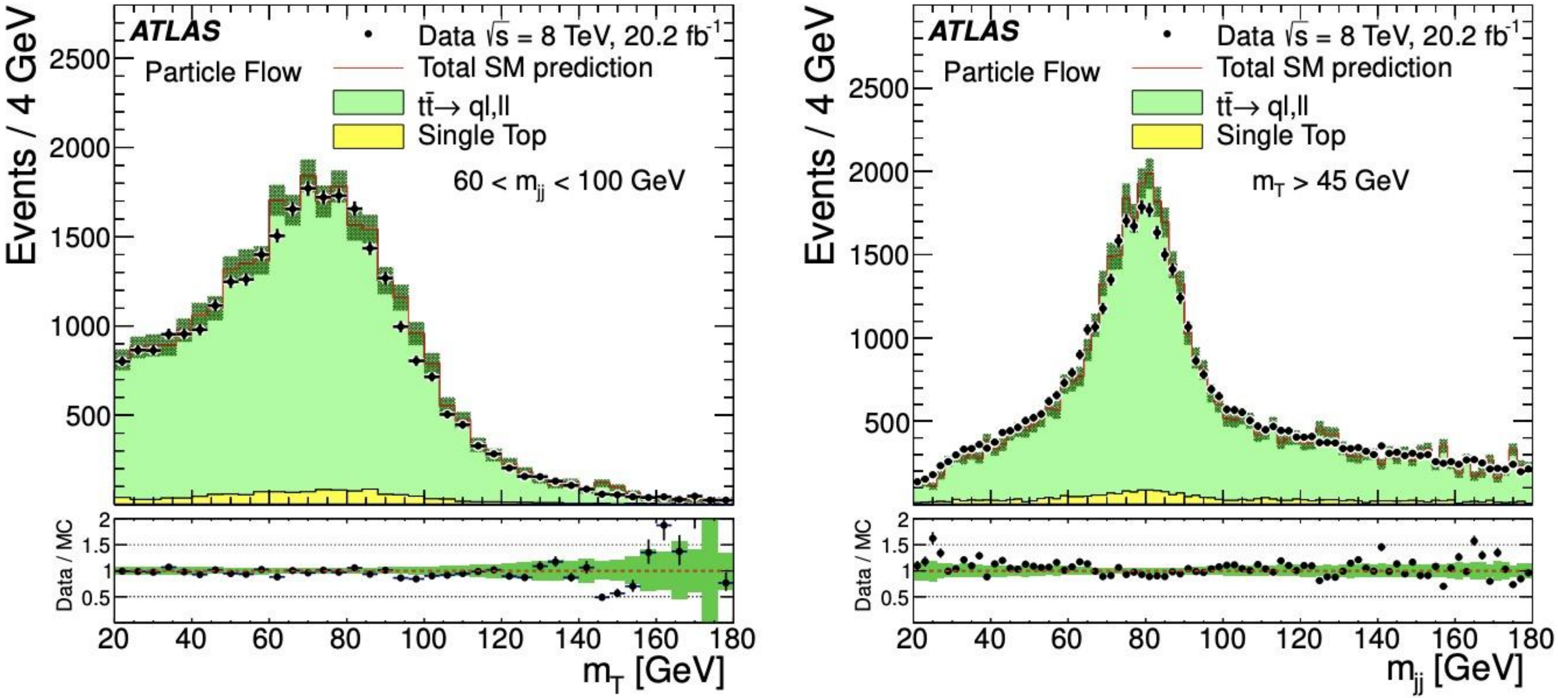
Performance



Particle Flow deals with pileup much better than only topo-clusters, with lower fake rate and better efficiency than previous correction methods

https://doi.org/10.48550/arXiv.1703.10485

Performance



Example comparisons with actual data, top pair events

https://doi.org/10.48550/arXiv.1703.10485

Track-CaloClusters (TCCs)

Particle Flow

Estimates energy contribution from individual tracks to clusters

Uses angular and momentum information from inner detector

Outputs list of PFOs – matched track-cluster pairs (charged) with (un)modified neutral clusters

Optimized for resolution of total jet variables (pT, m, etc.)

https://doi.org/10.48550/arXiv.2009.04986

Track-CaloClusters (TCCs)

Particle Flow		Track-CaloClusters
Estimates energy contribution from individual tracks to clusters		Uses tracks to split up clusters, all energy info from calorimeter
Uses angular and momentum information from inner detector	**vs**	Uses only angular information from inner detector
Outputs list of PFOs – matched track-cluster pairs (charged) with (un)modified neutral clusters		Outputs list of TCCs – separated clusters (charged) with unmodified neutral clusters
Optimized for resolution of total jet variables (pT, m, etc.)		Optimized for jet substructure reconstruction, esp. at high pT

https://doi.org/10.48550/arXiv.2009.04986

Track-CaloClusters (TCCs)

Particle Flow

Estimates energy contribution from individual tracks to clusters

Uses angular and momentum information from inner detector

Outputs list of PFOs – matched track-cluster pairs (charged) with (un)modified neutral clusters

Optimized for resolution of total jet variables (pT, m, etc.)

UFOs

Track-CaloClusters

Uses tracks to split up clusters, all energy info from calorimeter

Uses only angular information from inner detector

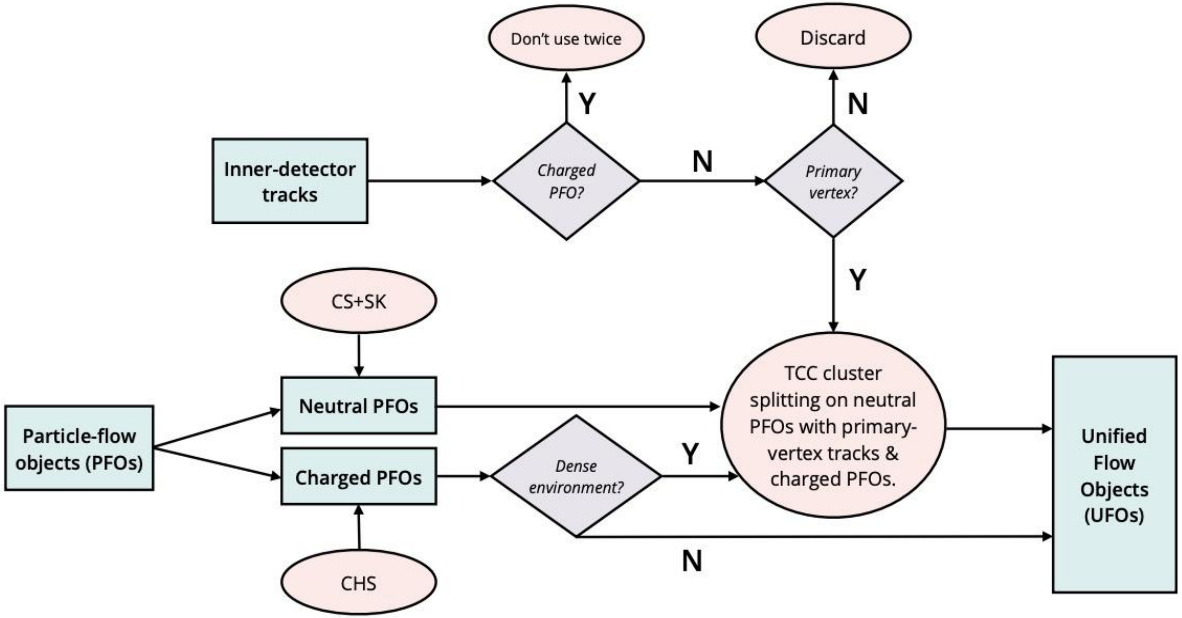
Outputs list of TCCs – separated clusters (charged) with unmodified neutral clusters

Optimized for jet substructure reconstruction, esp. at high pT

https://doi.org/10.48550/arXiv.2009.04986

Unified Flow Objects

UFOs are a more general, versatile way of preparing inputs for jet reconstruction.

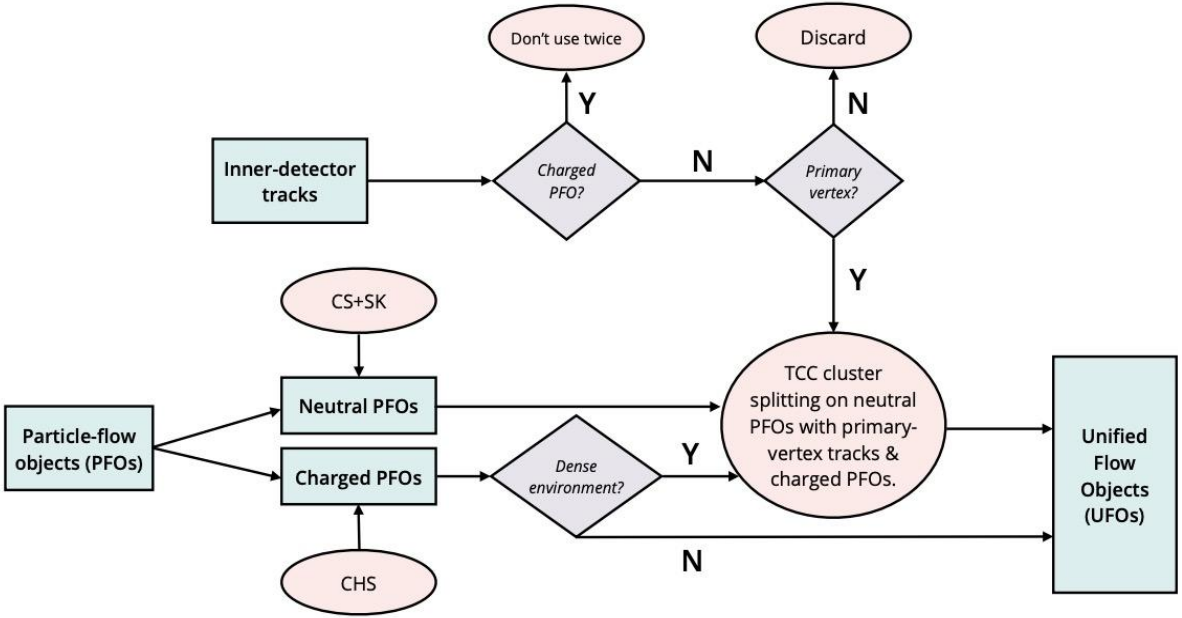


https://doi.org/10.48550/arXiv.2009.04986

Unified Flow Objects

UFOs are a more general, versatile way of preparing inputs for jet reconstruction.

PFOs are used as a basis, and can be further modified by, e.g. soft killer algorithms.



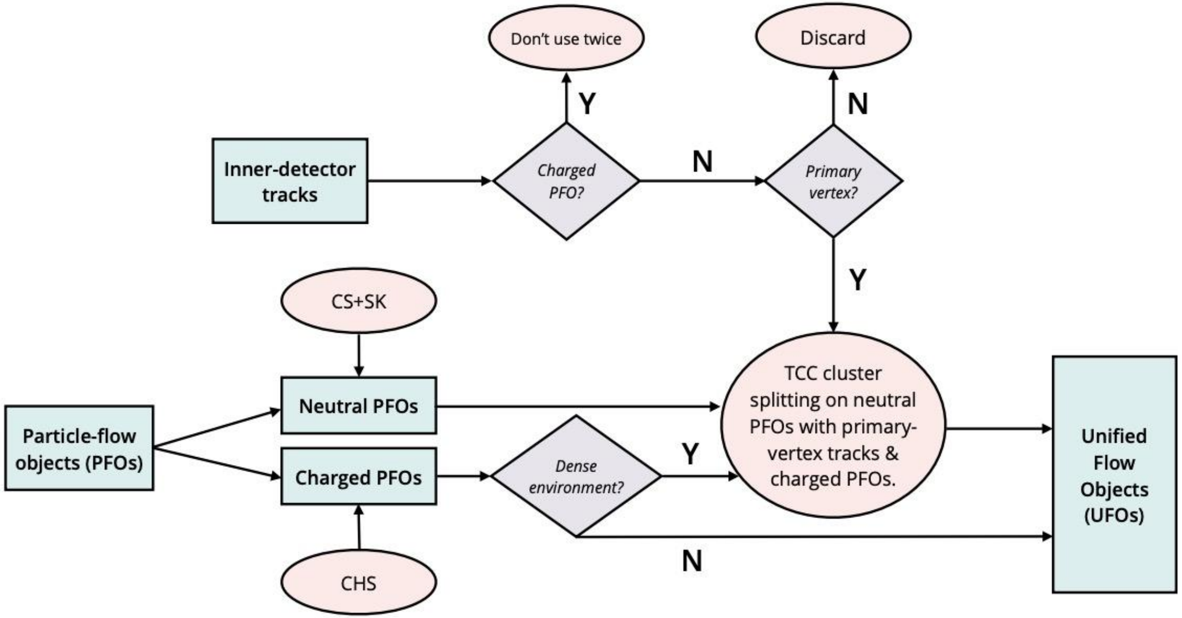
https://doi.org/10.48550/arXiv.2009.04986

Unified Flow Objects

UFOs are a more general, versatile way of preparing inputs for jet reconstruction.

PFOs are used as a basis, and can be further modified by, e.g. soft killer algorithms.

In areas where PFlow performs worse, topo-clusters are split into TCCs, but vertex info is always used.



https://doi.org/10.48550/arXiv.2009.04986

Performance

ATLAS Simulation
$\sqrt{s}$ = 13 TeV, $t \rightarrow q\bar{q}b$

Anti-k_t R=1.0 jets, no jet calibrations applied
500 GeV $\leq p_T^{true} <$ 1000 GeV, $|\eta^{true}| <$ 1.2

Background rejection @ ε_{Sig} = 50% efficiency (colour scale 0–80)

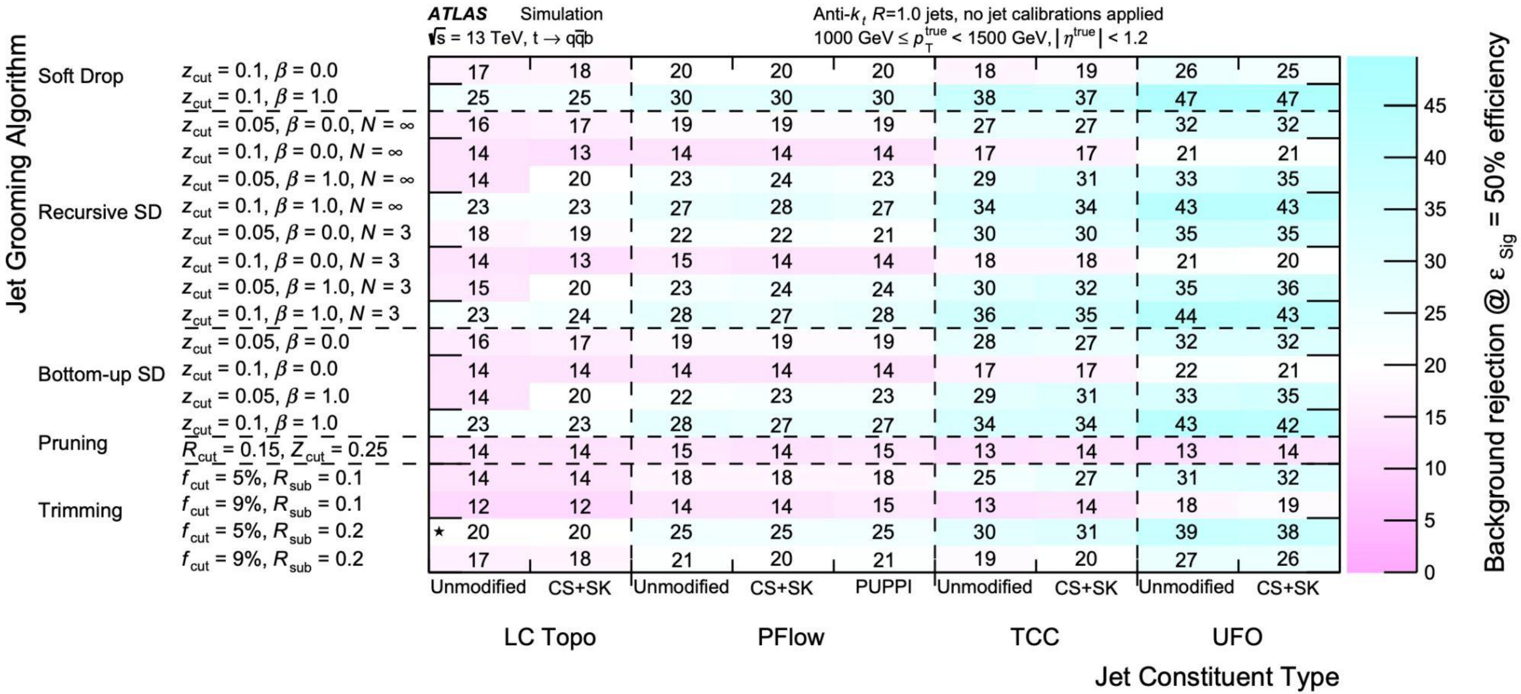
Jet Grooming Algorithm		LC Topo Unmodified	LC Topo CS+SK	PFlow Unmodified	PFlow CS+SK	PFlow PUPPI	TCC Unmodified	TCC CS+SK	UFO Unmodified	UFO CS+SK
Soft Drop	z_{cut} = 0.1, β = 0.0	53	50	68	74	68	51	44	76	78
	z_{cut} = 0.1, β = 1.0	51	54	73	76	70	56	53	76	83
Recursive SD	z_{cut} = 0.05, β = 0.0, $N = \infty$	27	37	45	47	46	38	37	54	57
	z_{cut} = 0.1, β = 0.0, $N = \infty$	31	30	39	36	36	32	29	49	49
	z_{cut} = 0.05, β = 1.0, $N = \infty$	28	42	51	57	55	38	42	55	62
	z_{cut} = 0.1, β = 1.0, $N = \infty$	39	47	60	65	62	47	45	67	73
	z_{cut} = 0.05, β = 0.0, N = 3	31	39	49	52	51	38	39	53	59
	z_{cut} = 0.1, β = 0.0, N = 3	28	30	42	41	39	32	32	51	52
	z_{cut} = 0.05, β = 1.0, N = 3	31	44	52	58	56	39	42	56	61
	z_{cut} = 0.1, β = 1.0, N = 3	40	47	61	67	63	47	45	65	71
Bottom-up SD	z_{cut} = 0.05, β = 0.0	26	36	45	48	46	36	38	54	56
	z_{cut} = 0.1, β = 0.0	30	32	40	39	38	33	30	49	50
	z_{cut} = 0.05, β = 1.0	28	42	52	57	55	38	42	55	61
	z_{cut} = 0.1, β = 1.0	39	48	60	64	62	47	47	65	72
Pruning	R_{cut} = 0.15, Z_{cut} = 0.25	25	21	30	29	28	17	15	31	29
Trimming	f_{cut} = 5%, R_{sub} = 0.1	24	24	36	37	35	33	33	46	46
	f_{cut} = 9%, R_{sub} = 0.1	14	17	25	27	27	17	20	31	32
	f_{cut} = 5%, R_{sub} = 0.2	★ 41	40	52	51	51	39	38	57	60
	f_{cut} = 9%, R_{sub} = 0.2	37	36	52	54	50	31	33	56	56

Jet Constituent Type

Jet tagging performance for simulated top quark jets, lower pT, with a lot of variations

https://doi.org/10.48550/arXiv.2009.04986

Performance



Jet tagging performance for simulated top quark jets, higher pT, with a lot of variations

https://doi.org/10.48550/arXiv.2009.04986

Summary

- Through LHC Run 1, ATLAS jet reconstruction used only calorimetry in the form of topo-clusters

https://doi.org/10.48550/arXiv.2009.04986

Summary

- Through LHC Run 1, ATLAS jet reconstruction used only calorimetry in the form of topo-clusters

- Basic particle-flow algorithms were tested on Run 1 data and found to significantly improve resolution of overall jet variables, at least in the lower-pT regime where tracker data is helpful

https://doi.org/10.48550/arXiv.2009.04986

Summary

- Through LHC Run 1, ATLAS jet reconstruction used only calorimetry in the form of topo-clusters

- Basic particle-flow algorithms were tested on Run 1 data and found to significantly improve resolution of overall jet variables, at least in the lower-pT regime where tracker data is helpful

- Other methods of combining ID and calorimeter data like Track-CaloClusters were tested soon after this, and found to be useful for reconstruction of jet substructure for highly boosted jets

https://doi.org/10.48550/arXiv.2009.04986

Summary

- Through LHC Run 1, ATLAS jet reconstruction used only calorimetry in the form of topo-clusters

- Basic particle-flow algorithms were tested on Run 1 data and found to significantly improve resolution of overall jet variables, at least in the lower-pT regime where tracker data is helpful

- Other methods of combining ID and calorimeter data like Track-CaloClusters were tested soon after this, and found to be useful for reconstruction of jet substructure for highly boosted jets

- As of 2021, the approaches are combined into Unified Flow Objects, augmenting the basic structure of the particle-flow algorithms with TCCs and other widely-used jet grooming techniques where PFOs are less useful

https://doi.org/10.48550/arXiv.2009.04986

Summary

- Through LHC Run 1, ATLAS jet reconstruction used only calorimetry in the form of topo-clusters

- Basic particle-flow algorithms were tested on Run 1 data and found to significantly improve resolution of overall jet variables, at least in the lower-pT regime where tracker data is helpful

- Other methods of combining ID and calorimeter data like Track-CaloClusters were tested soon after this, and found to be useful for reconstruction of jet substructure for highly boosted jets

- As of 2021, the approaches are combined into Unified Flow Objects, augmenting the basic structure of the particle-flow algorithms with TCCs and other widely-used jet grooming techniques where PFOs are less useful

Questions?

https://doi.org/10.48550/arXiv.2009.04986