universitätbonn

Particle flow status & plans

Regina Moles-Valls

IV Single Top Workshop
3-4 December 2015
Bergische Universität Wuppertal

thanks to Tom, Andrew, Peter and Ian for the material

Outline

- Introduction to particle flow
  - Why particle flow in ATLAS?
  - How does it work in general?
- Particle flow performance studies (8 TeV)
- Particle flow in top xAOD and AnalysisTop
  - First plots with mc15 (from P. Falke)

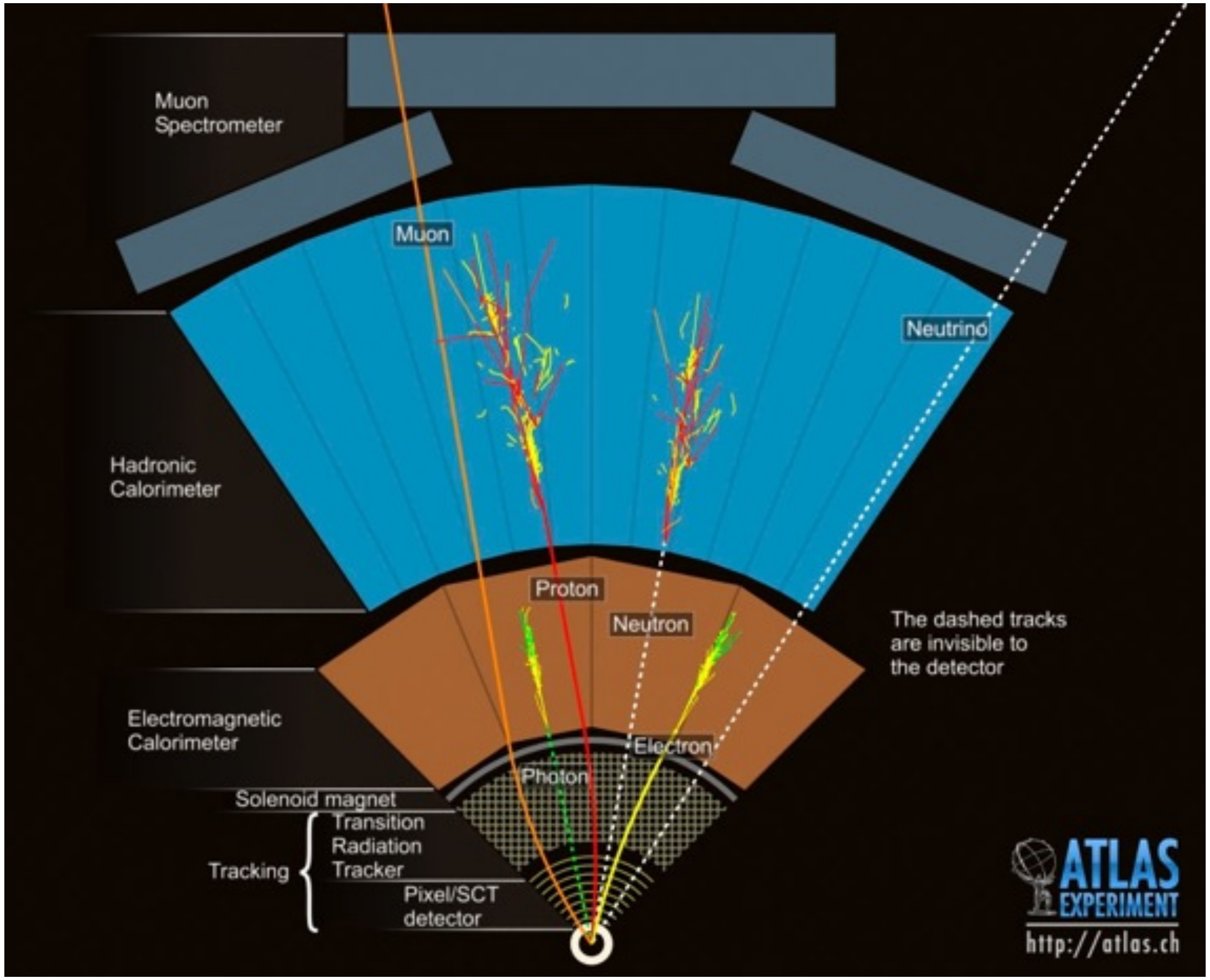
Introduction to Particle Flow

Particle Flow algorithms try to follow the path of the particles through the detector.

Main goal is to improve the energy resolution of the hadronic objects

How to do it? combining the information from different sub-detectors

➡ Emphasise the role of the tracker in jet physics.



Why Particle flow in ATLAS?

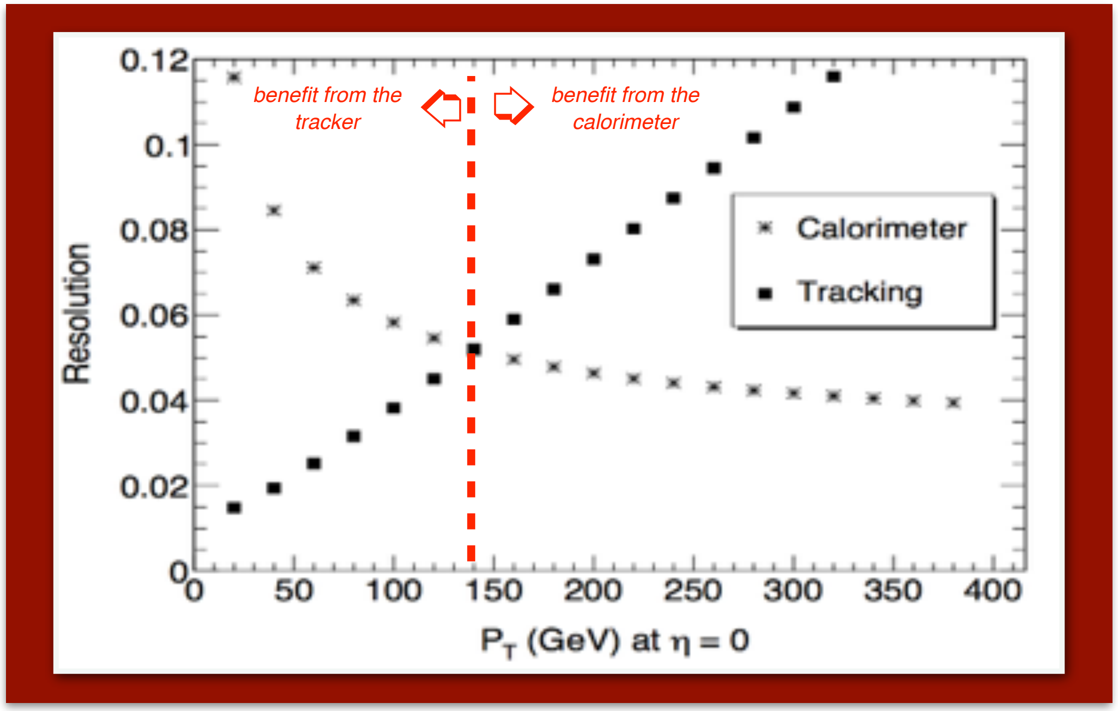
- Reasons for using Particle Flow in ATLAS:

ID tracker

- Tracker resolution significantly better than Calo resolution at low p_T
- Particles that don't create a topocluster (low E) are accessible by the ID
- Better angular resolution of the tracker for single particles
- The vertex information can be used to mitigate the pileup contribution

Calorimeter

- Calorimeter's ability to reconstruct neutral particles
- Better energy resolution at high p_T



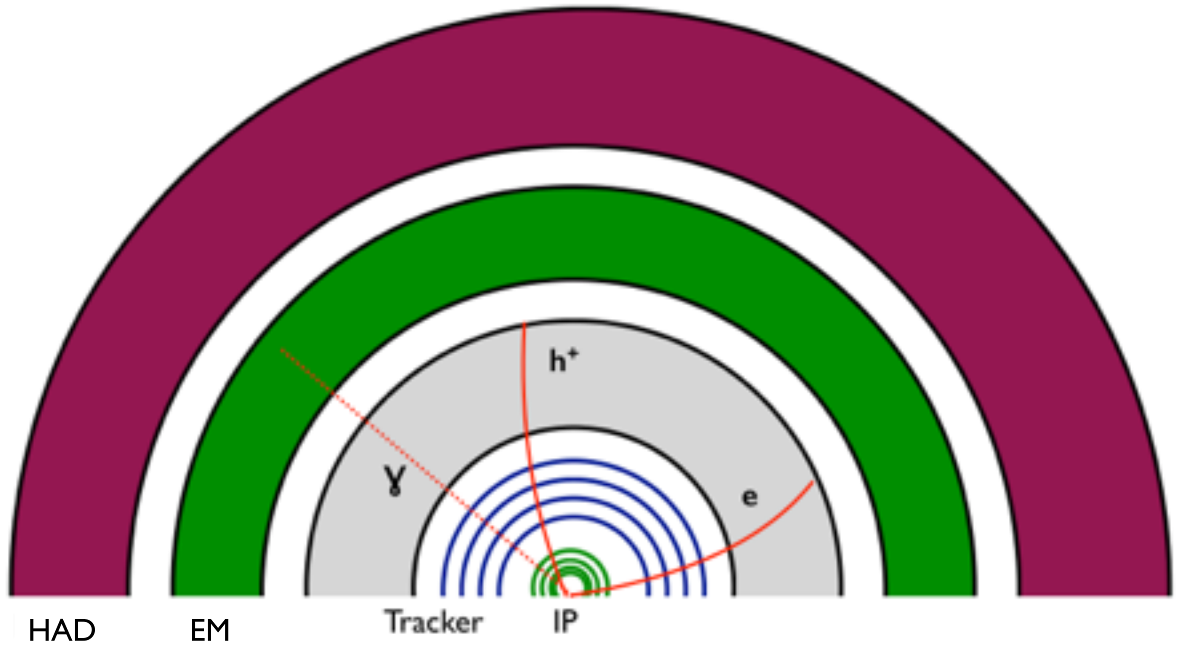
__Particle flow algorithm uses:__

- *the tracker information*
- *the calorimeter information*
- *a combination of both*

Introduction: Particle Flow principle

How does Particle Flow work?

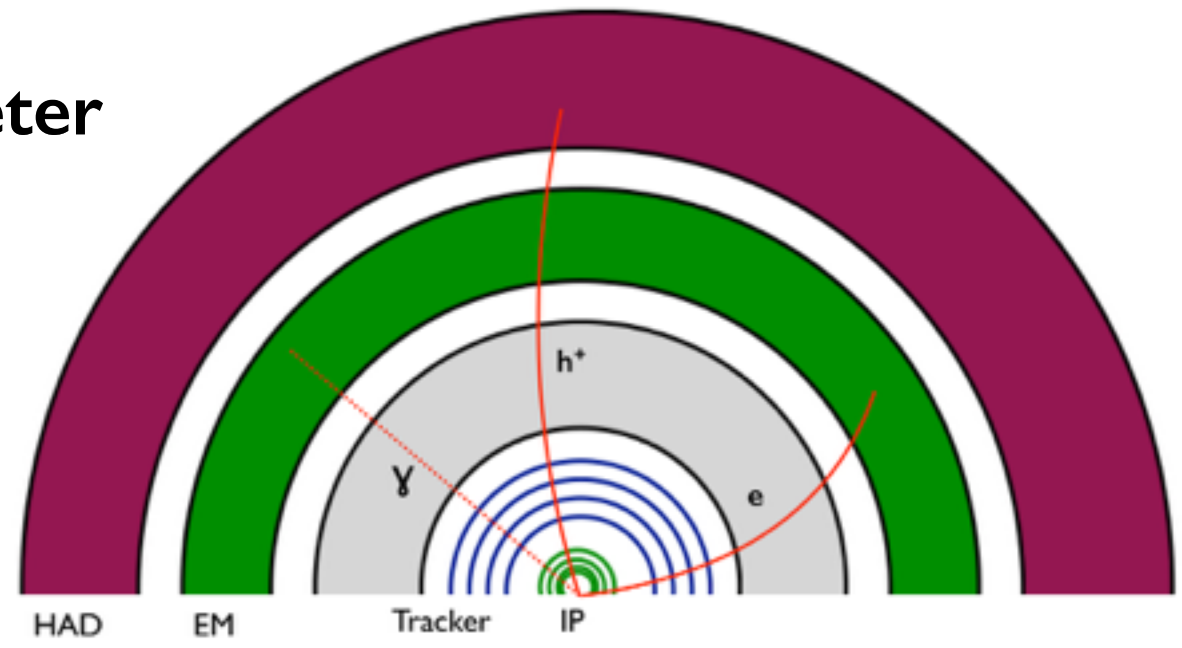
- Track reconstruction in the ID



Introduction: Particle Flow principle

How does Particle Flow work?

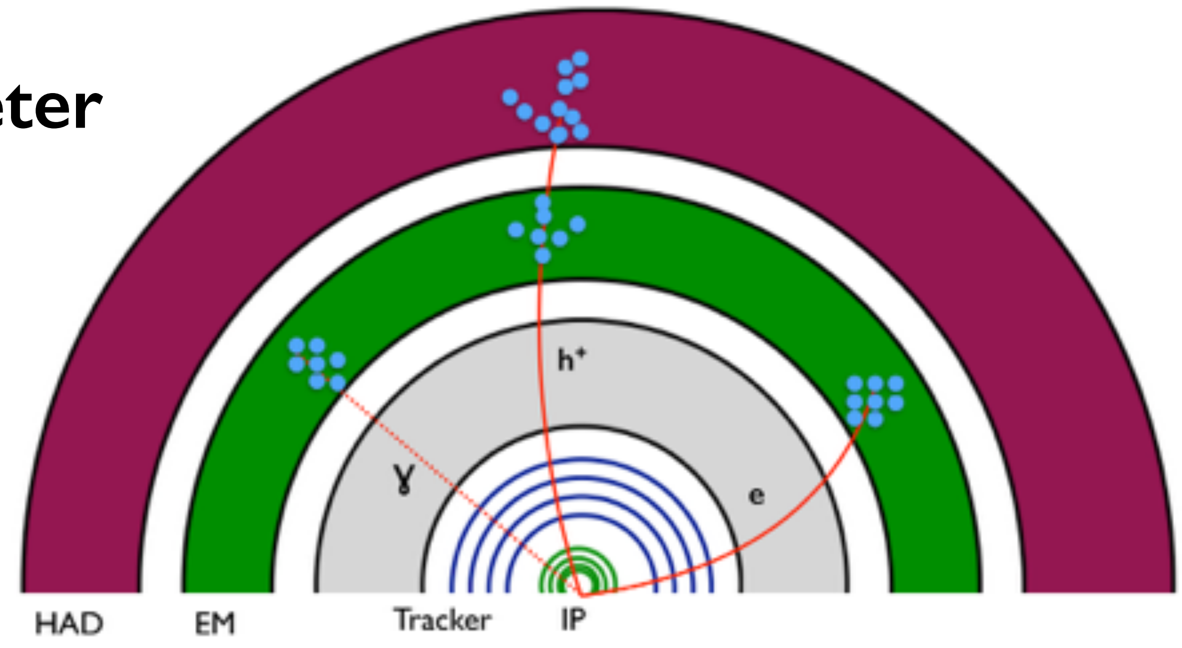
- Track reconstruction in the ID
- Extrapolate the tracks to the Calorimeter



Introduction: Particle Flow principle

How does Particle Flow work?

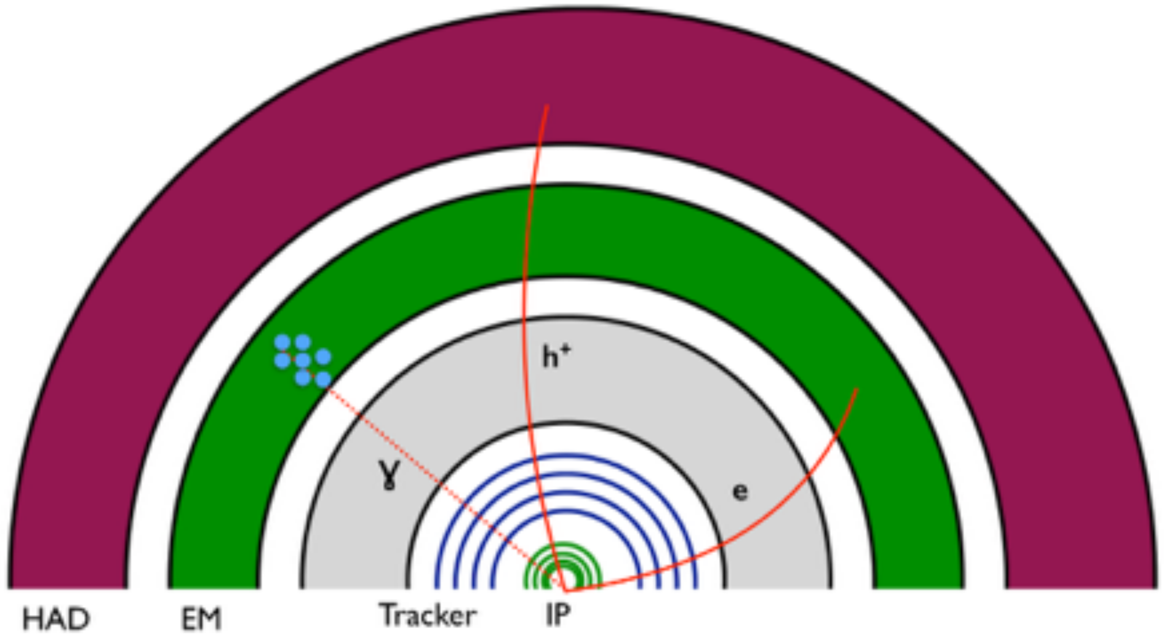
- Track reconstruction in the ID
- Extrapolate the tracks to the Calorimeter
- Match the tracks to the clusters



Introduction: Particle Flow principle

How does Particle Flow work?

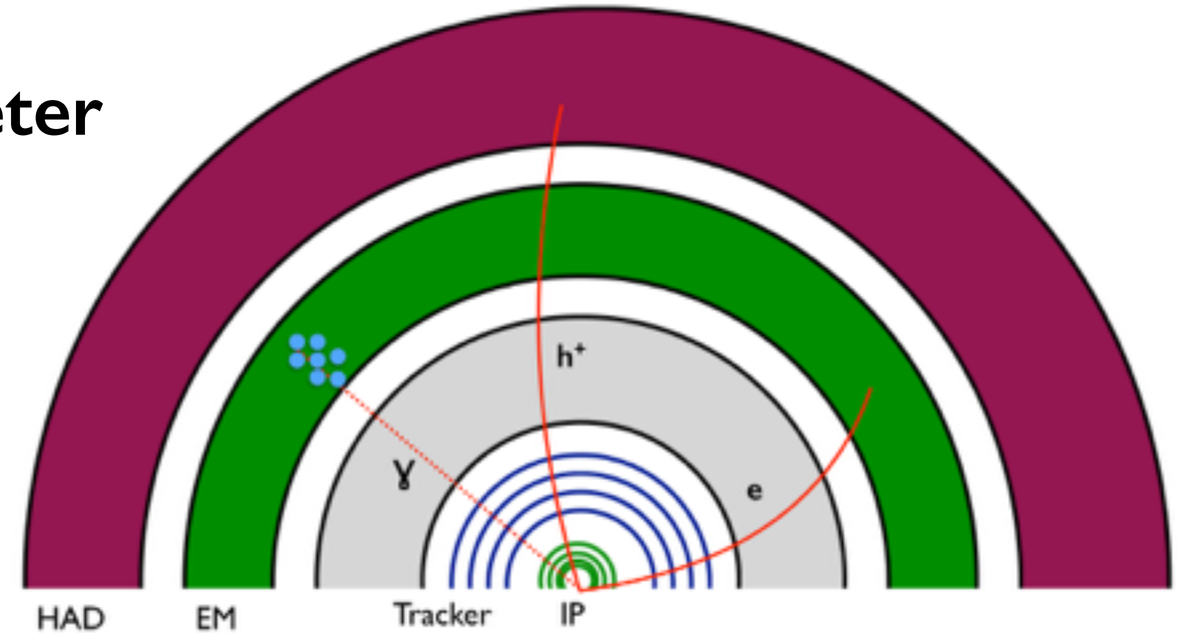
- Track reconstruction in the ID
- Extrapolate the tracks to the Calorime
- Match the tracks to the clusters
- Remove clusters from charge particles
- Finally are kept:
  - tracks and
  - clusters from neutral particles



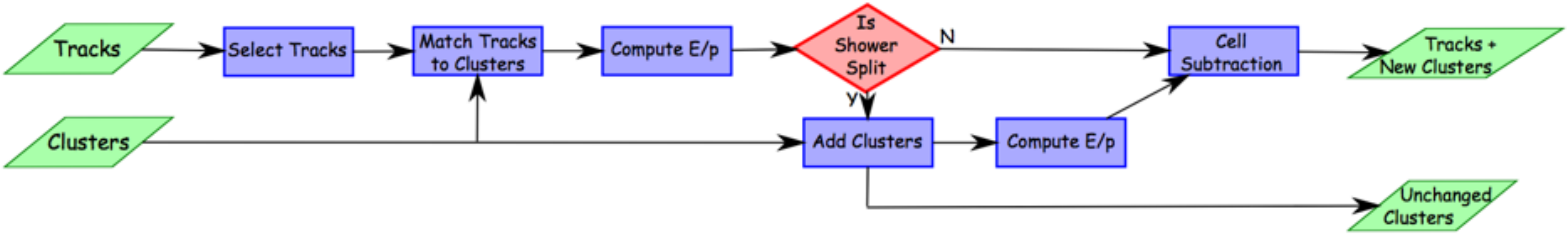
Introduction: Particle Flow principle

How does Particle Flow work?

- Track reconstruction in the ID
- Extrapolate the tracks to the Calorimeter
- Match the tracks to the clusters
- Remove clusters from charge particles
- Finally are kept:
    - tracks and
    - clusters from neutral particles



eflowRec algorithm in ATLAS

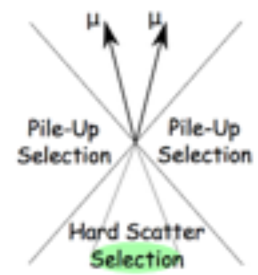


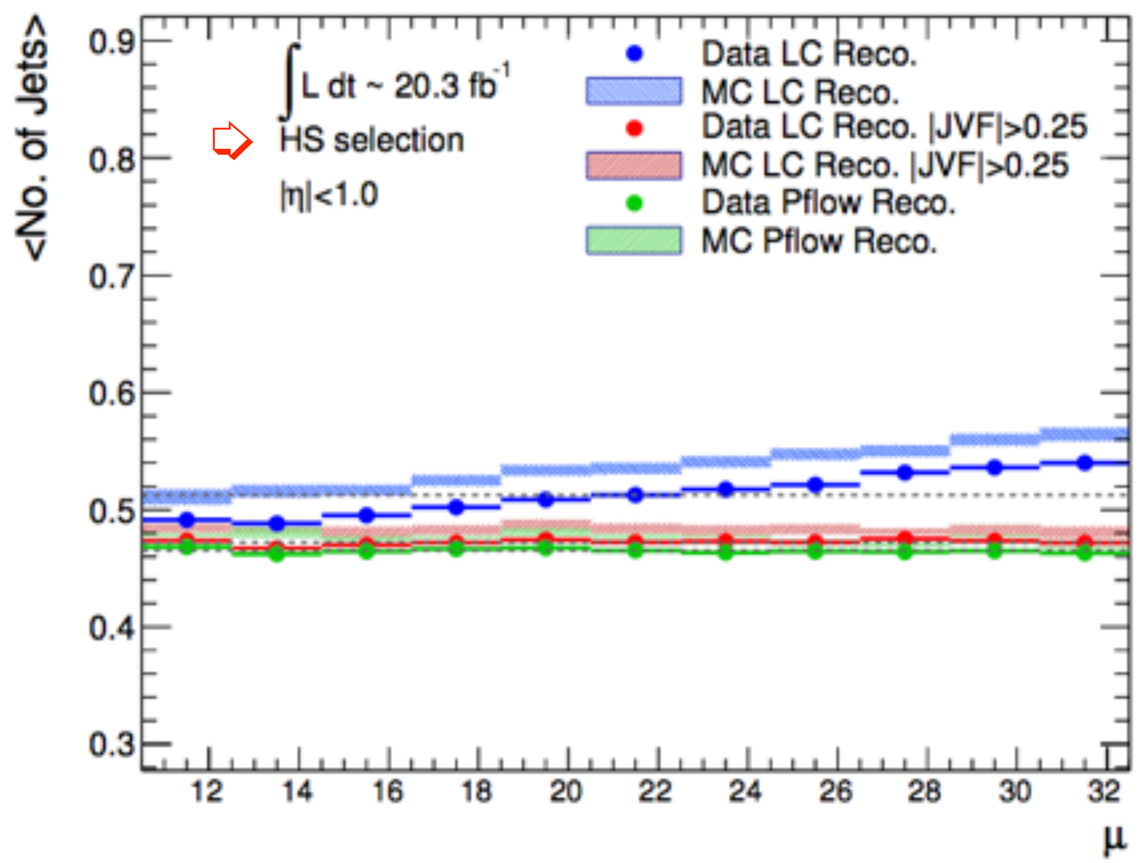
Particle flow performance studies

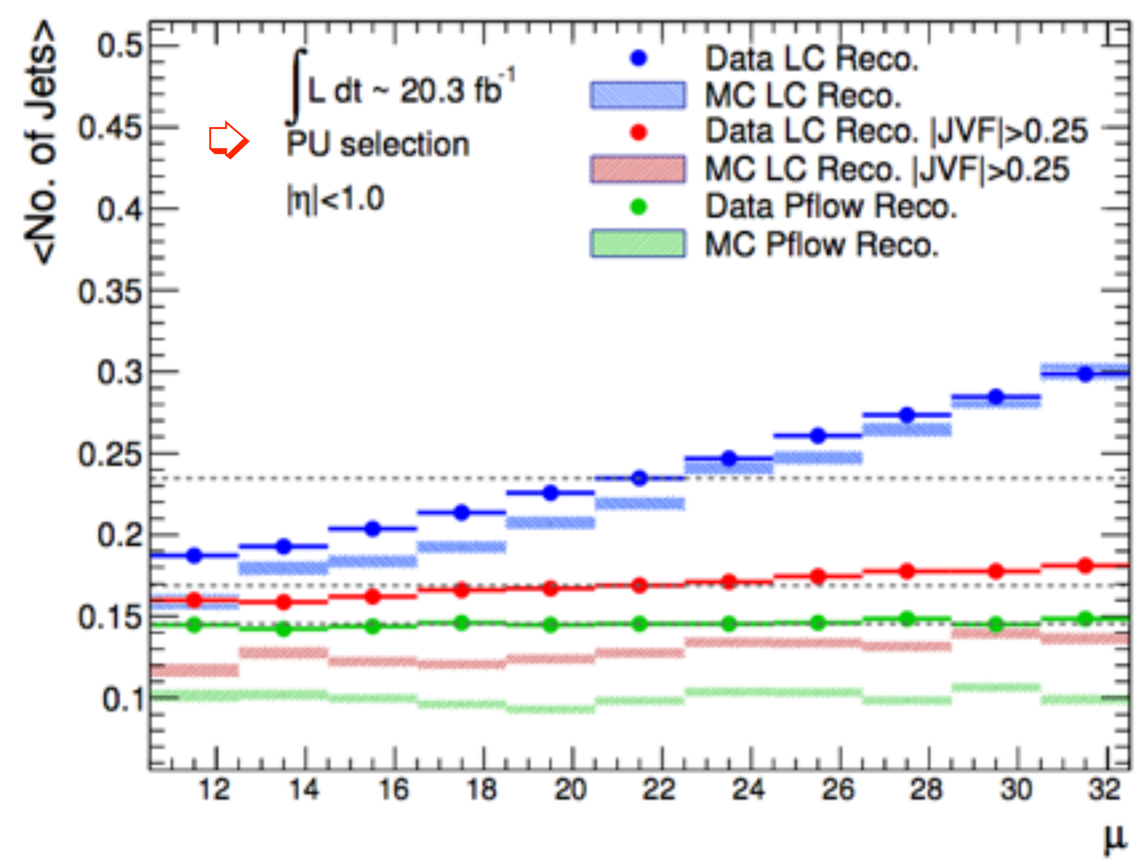
Results based on $\sqrt{s}$ = 8 TeV (rel 17)

Pile-up

- Average number of particle flow jets originating form pile-up is much suppressed.
- Average number of particle flow jets is stable as a function of pile-up.
- Particle flow jets behaviour flatter than the LC and LC+JVF jets
  - Hard scatter (HS) selection
  - Pileup (PI) selection





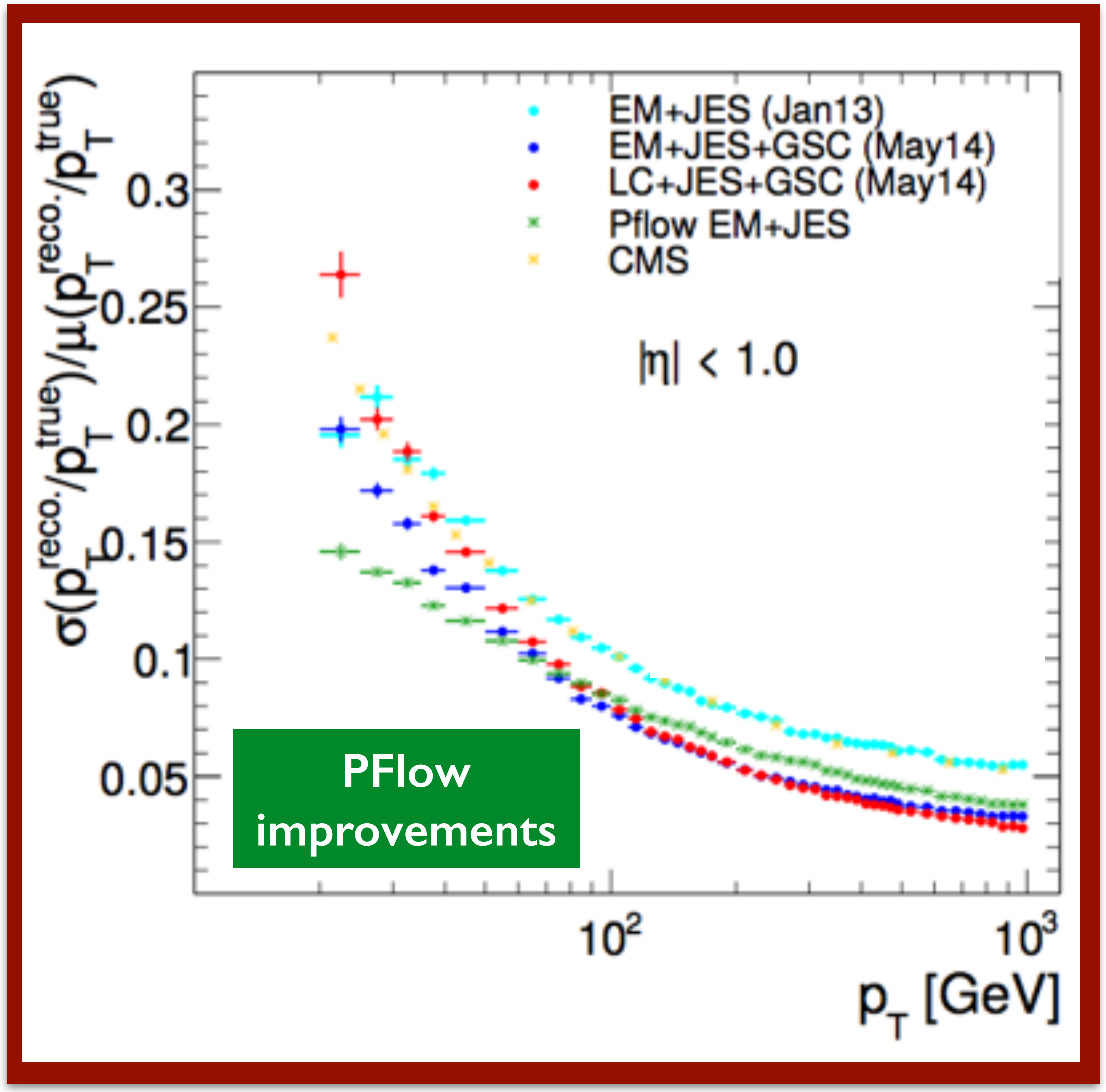


Particle flow performance studies

Results based on $\sqrt{s}$ = 8 TeV (rel 17)

Jet P_T resolution

- Most significant improvements at low p_T and central η region
- Resolution at high p_T is a bit worse
  - It will be fixed in Rel 20.7



Quick look to the fraction of events with p_T<80 GeV in top analysis:

- **tZ l+jets ~ 50%** (I.Cioara)
- **ttbar l+jets ~ 64%** (P.Falke)

(mc15_13TeV.410000.PowhegPythia)

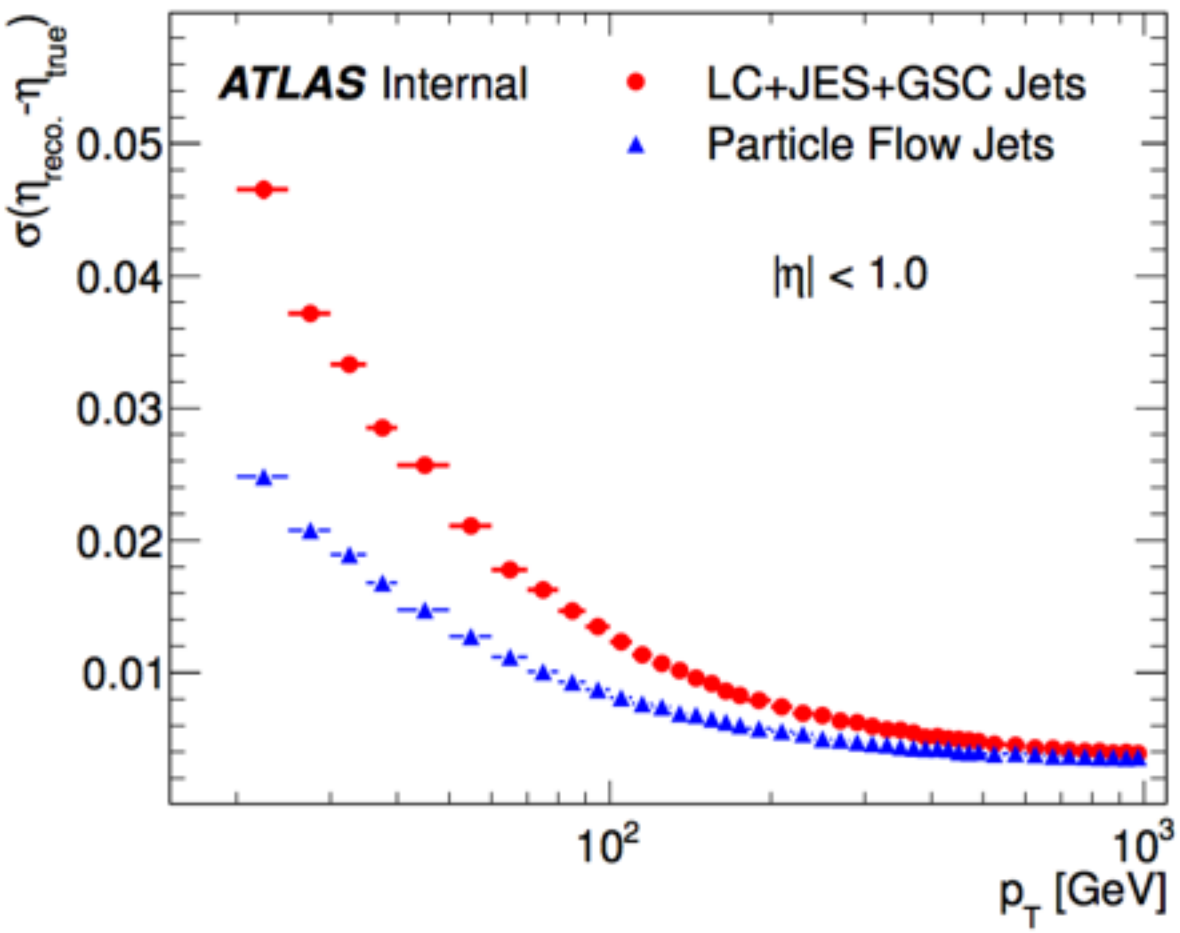
Particle flow performance studies

Results based on $\sqrt{s}$ = 8 TeV (rel 17)

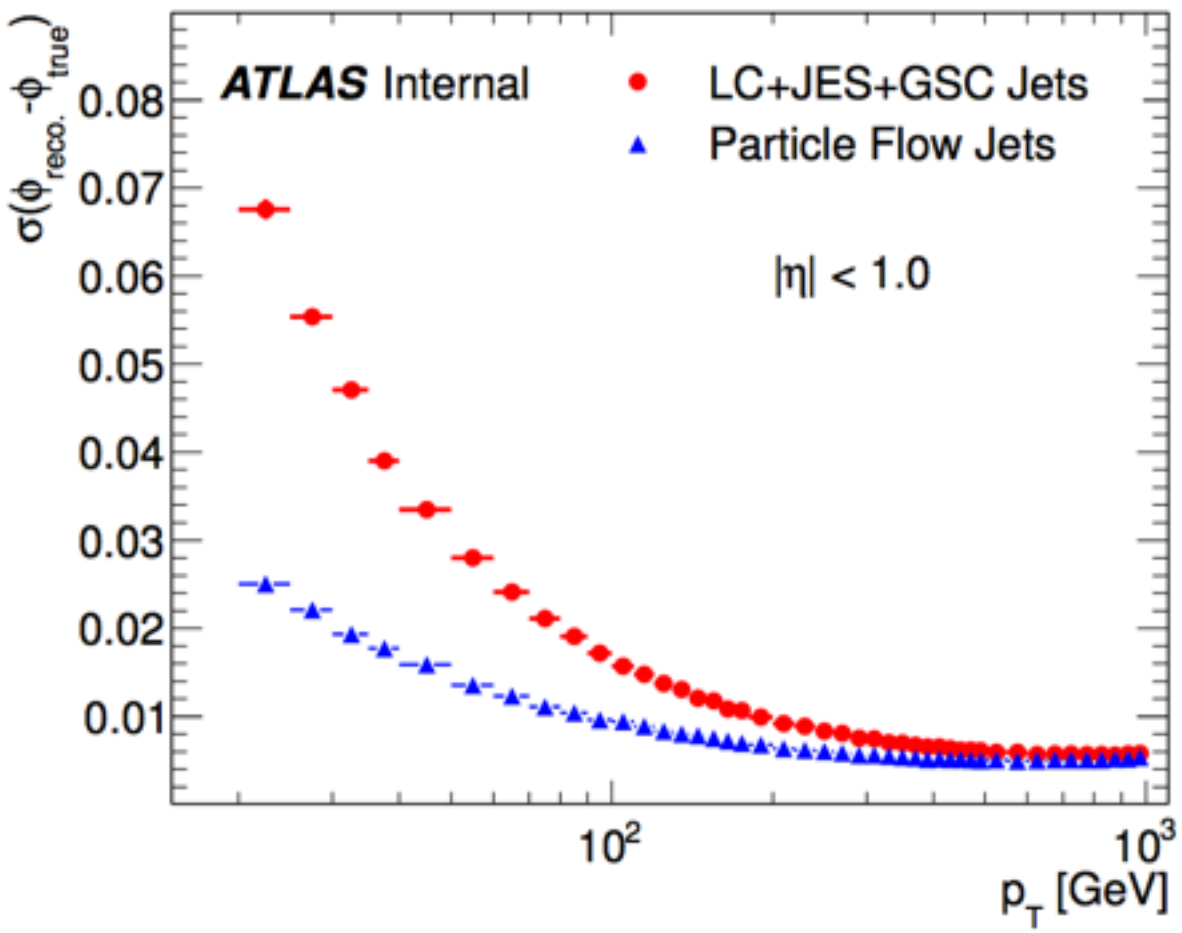
Jet angular resolution

- Improvement in the jet angular resolution for η,Φ

η angular resolution



Φ angular resolution



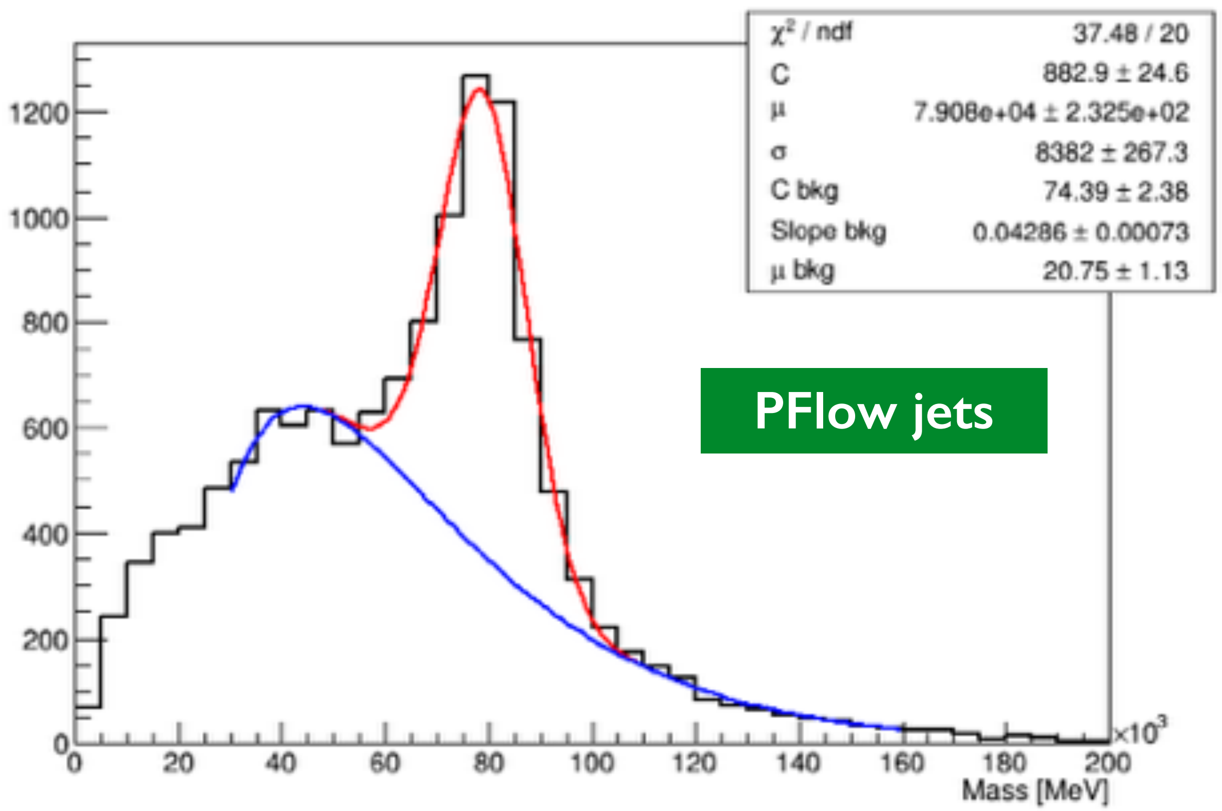
Particle flow performance studies

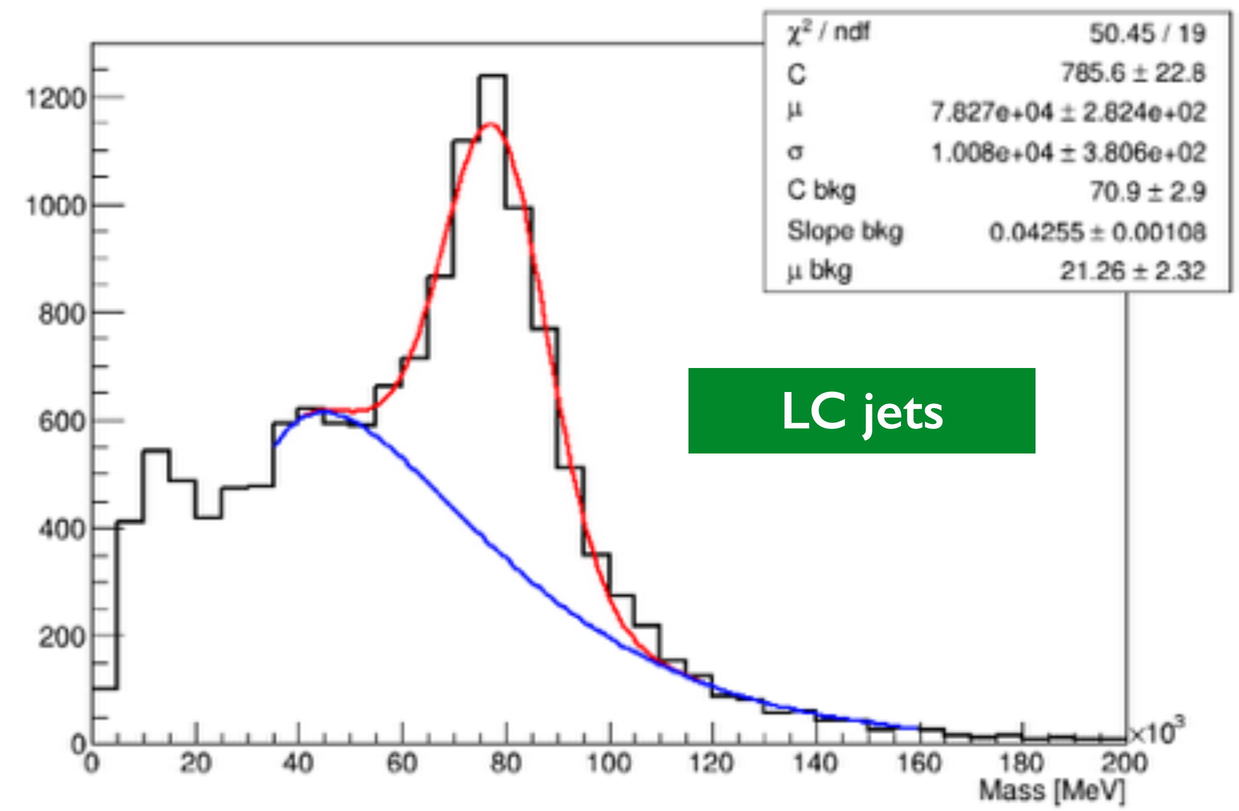
Results based on √s = 8 TeV (rel 17)

Large R-jets in ttbar events

- Invariant mass reconstruction of the hadronic W boson in ttbar events
  - Signal jets' those arising from a W-boson hadronic decay
  - Improvement in the background rejection
  - Improvements in the W mass resolution

σ = 11% (**Pflow**) σ = 13% (**LC**)





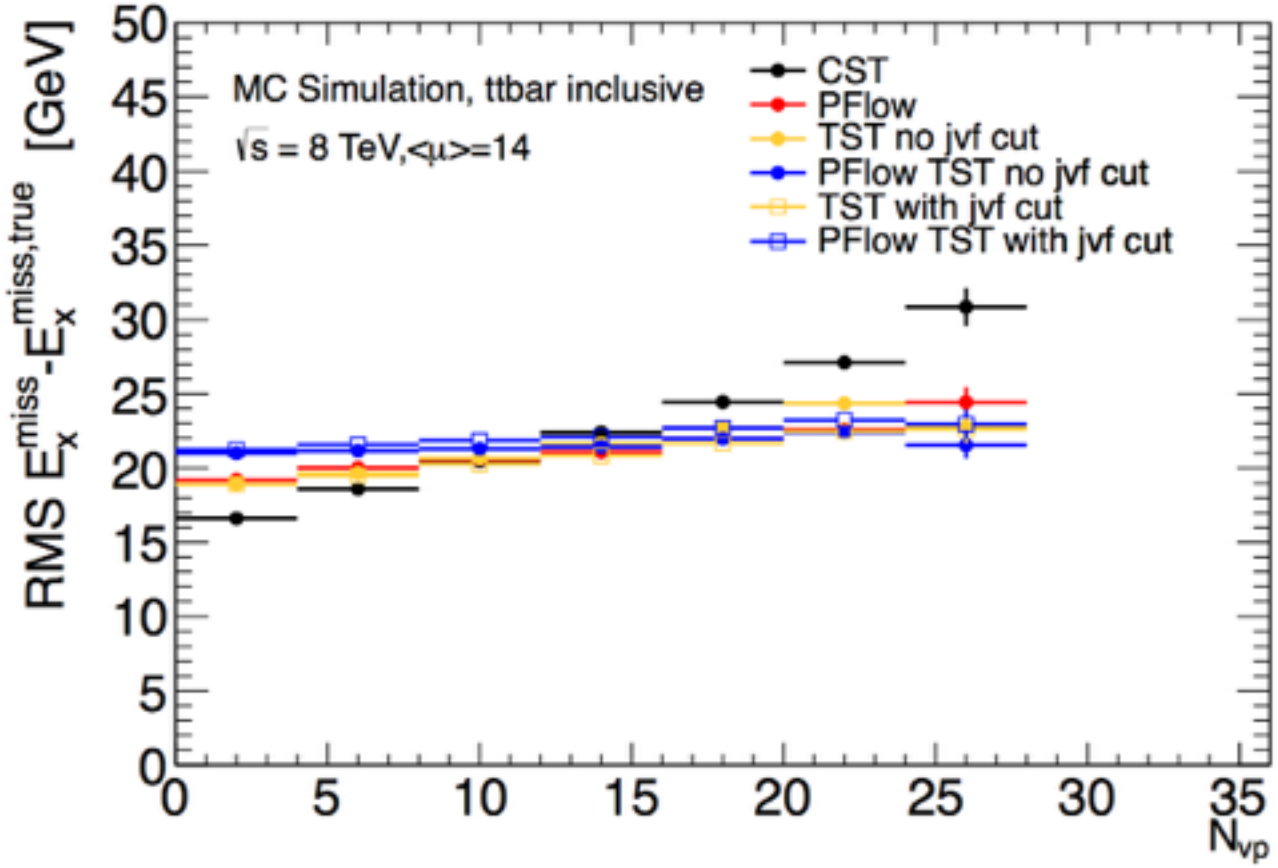
Particle flow performance studies

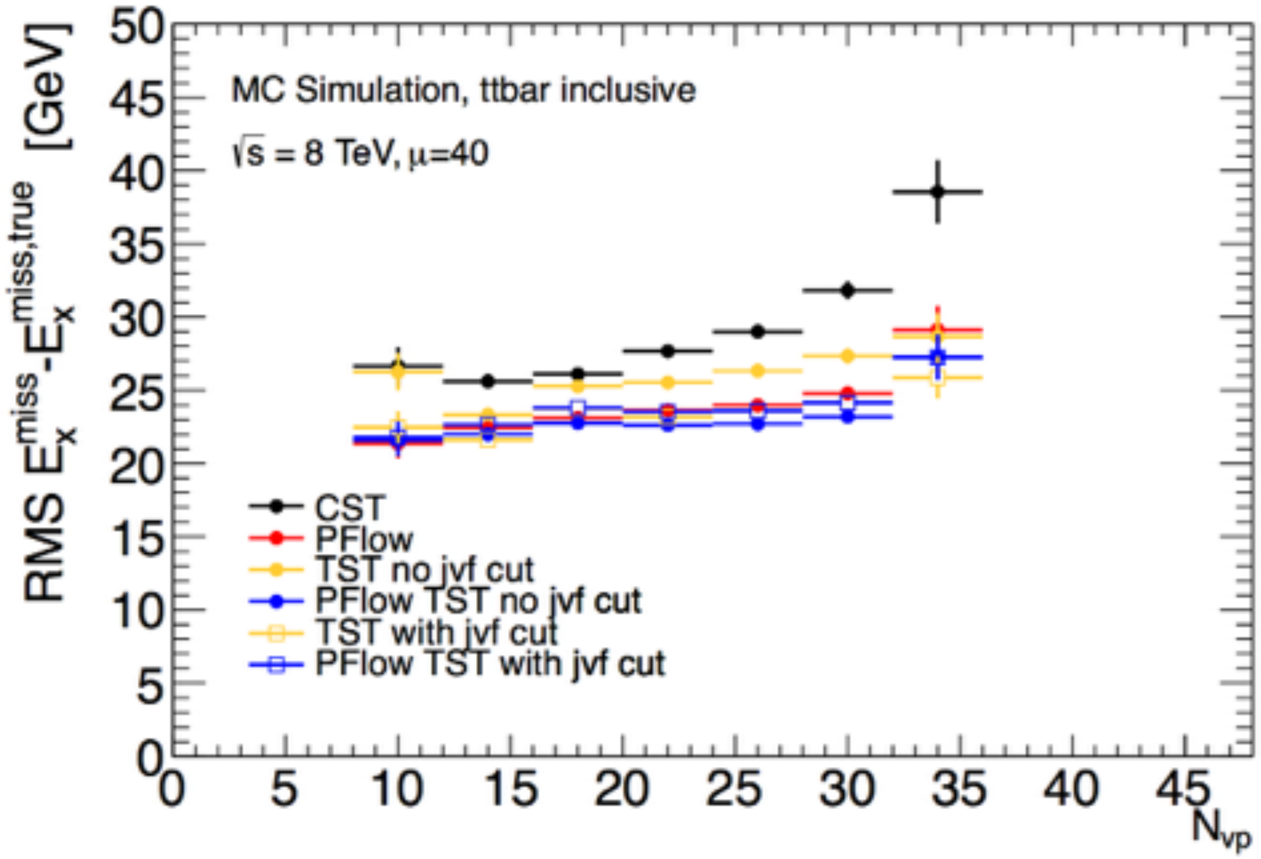
Results based on $\sqrt{s}$ = 8 TeV (rel 17)

E_T^{miss}

$$E_{x,y}^{\text{miss}} = -\left(\Sigma E_{x,y}^{\text{particle flow jets}} + \Sigma p_{x,y}^{e} + \Sigma p_{x,y}^{\gamma} + \Sigma p_{x,y}^{\mu} + \Sigma p_{x,y}^{\text{trk, unassociated}} + \Sigma E_{x,y}^{\text{clu, unassociated}}\right)$$

- ttbar MC events with high and low pileup conditions
  - PFlow
  - **CST** (cluster soft term)
  - TST (track soft term)
  - PFlow TST (PFlow track soft term)





A first look into the ttbar pflow jets...

PFlow in Top xAOD

- PFlow collection available in the primary xAOD
- DerivationFramework (00-02-78)
  - **AntiKt4EMPFlowJets** (1% of the total size in the slimmed xAOD)
    DerivationFramework/DerivationFrameworkCore/trunk/python/AntiKt4EMPFlowJetsCPContent.py
  - **MET_AntiKt4EMPFlow**
    DerivationFrameworkCore/trunk/python/MET_Reference_AntiKt4EMTopoCPContent.py
  - **BTagging_AntiKt4EMPFlow**
    DerivationFrameworkCore/trunk/python/BTagging_AntiKt4EMTopoCPContent.py

- AnalysisTop
  - Apply the right calibration for running PFlow jet collection:
    - JES_MC15Prerecommendation_PFlow_July2015.config
  - No b-tagging SF applied (not available yet)
  - Turn off the jet uncertainties (not available yet)
  - JetCleaning variables not included in the derivation (should use AntiKt4EMTopJets)

First look:

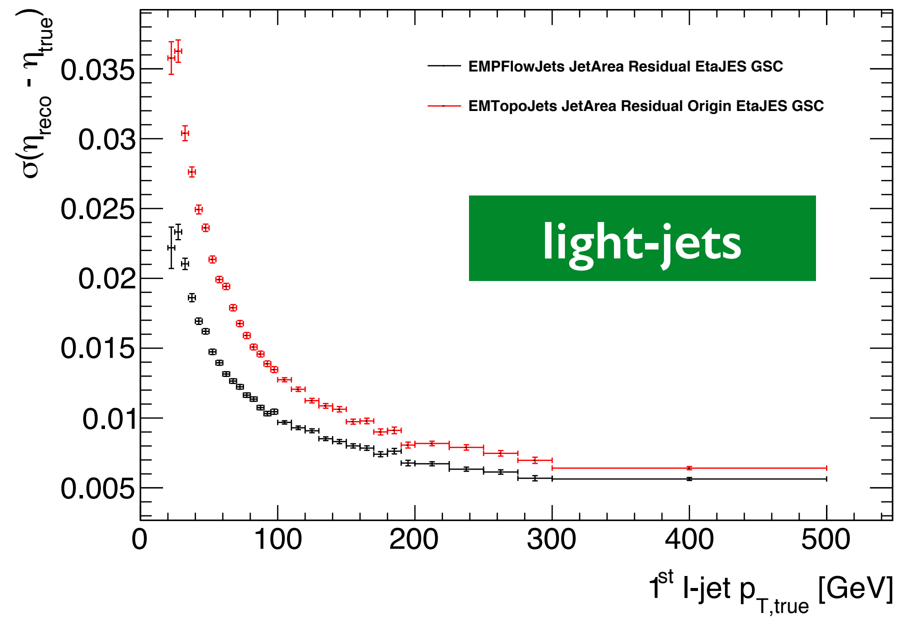
P. Falke

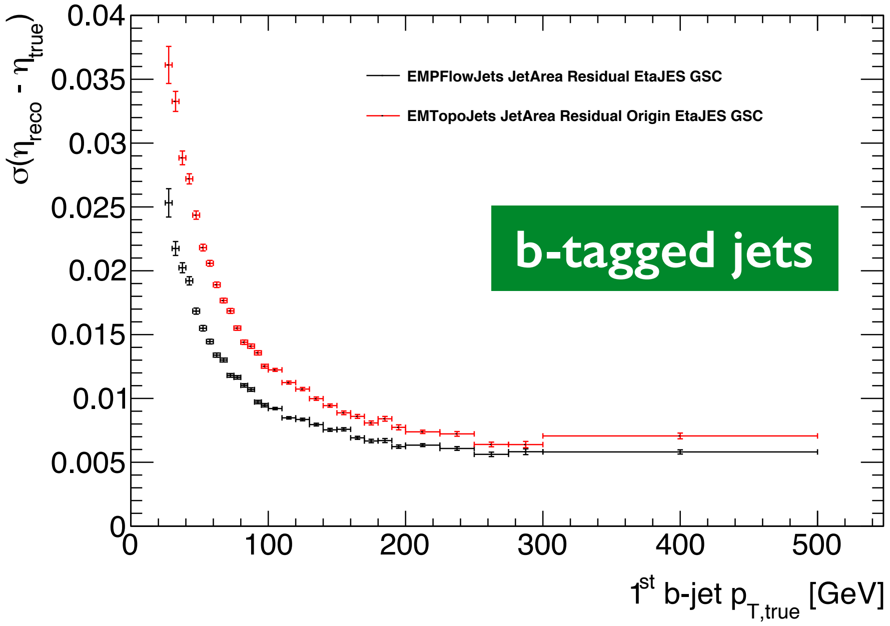
- First look at PFlow jets in AnalysisTop
  - ttbar TOPQ1 derivation

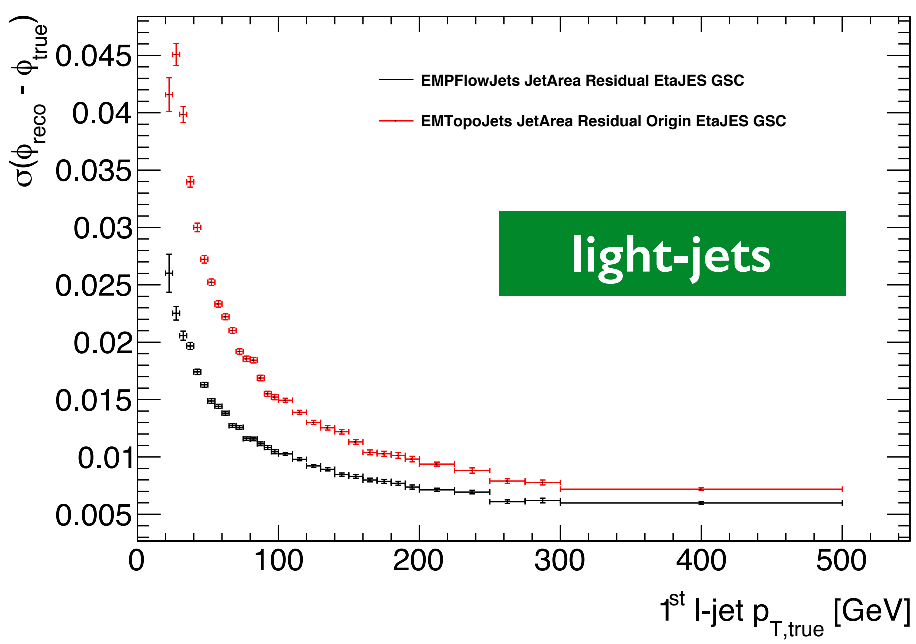
mc15_13TeV.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.DAOD_TOPQ1.e3698_s2608_s2183_r7267_r6282_p2460

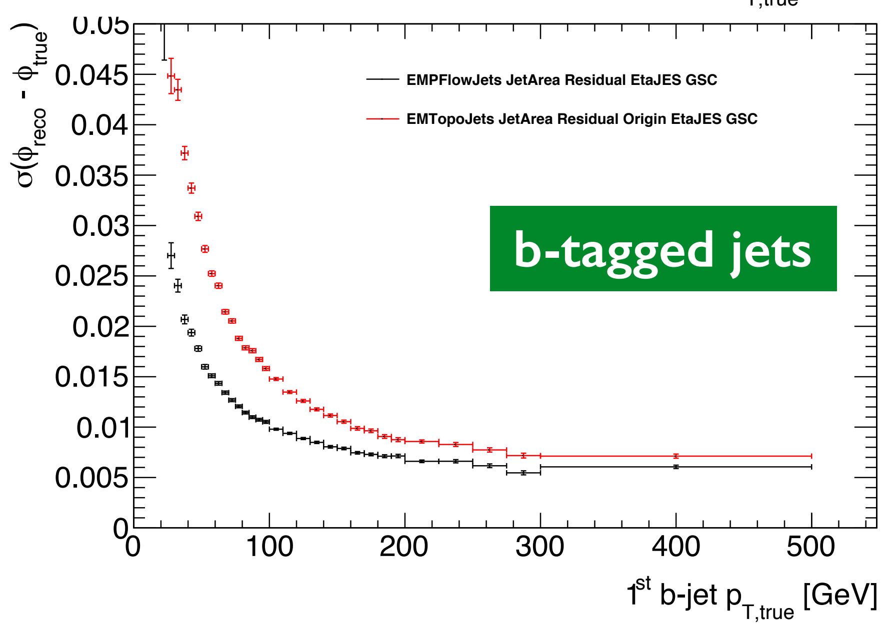
  - l+jets selection implemented

Jet angular resolution





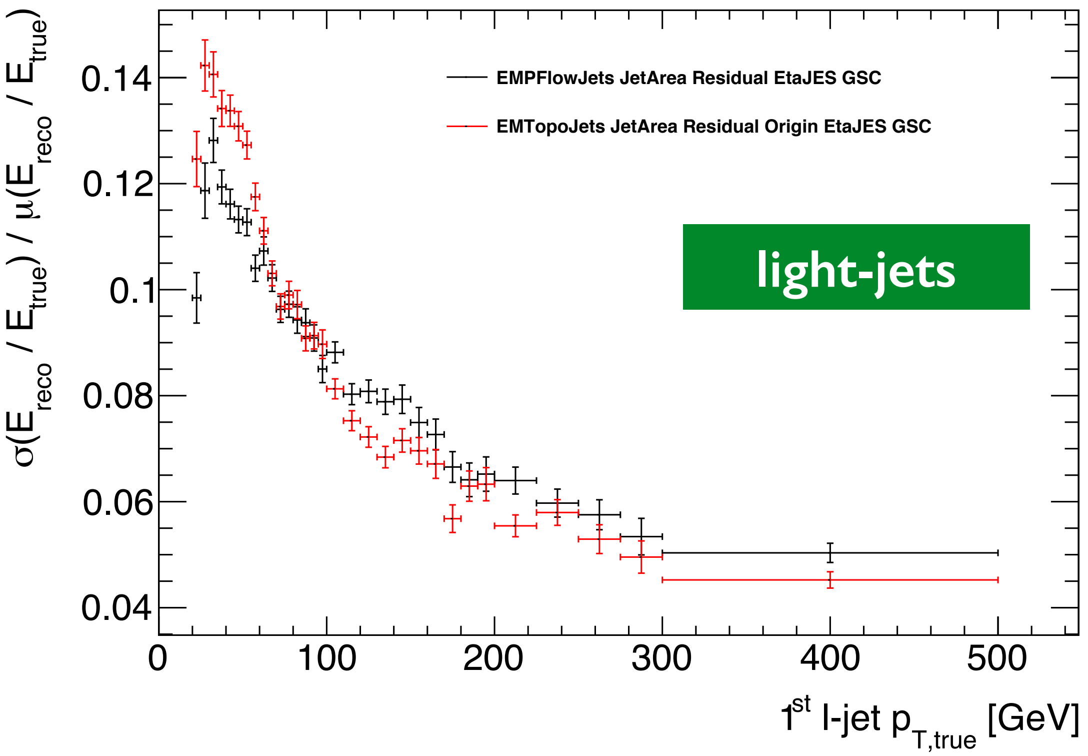


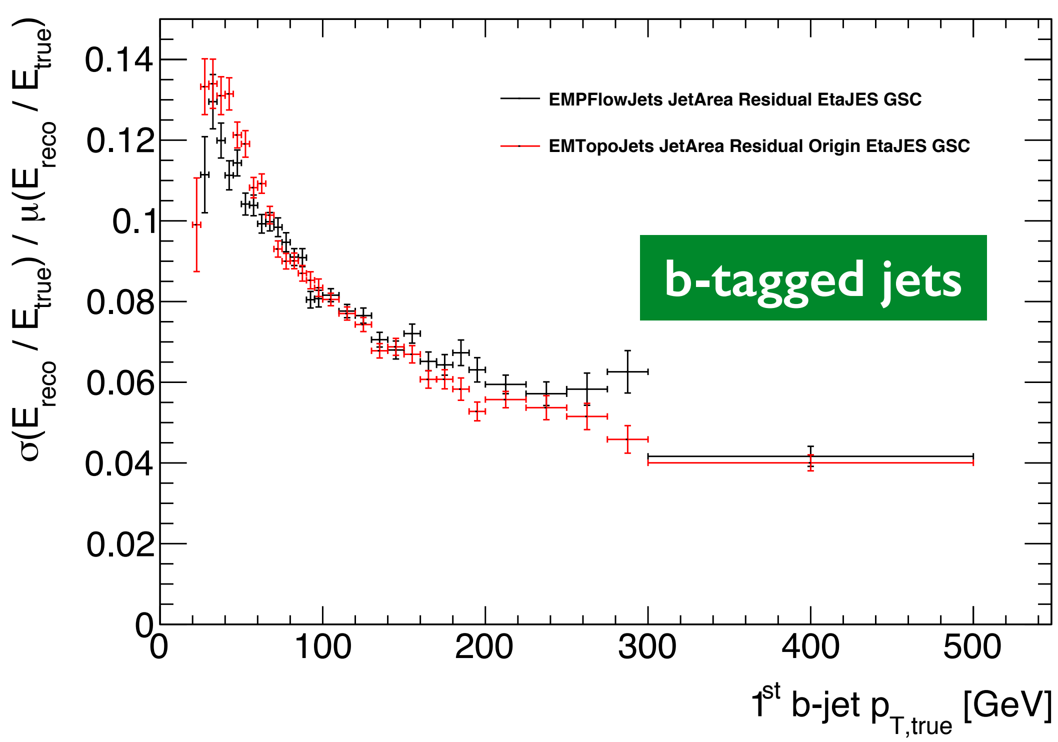


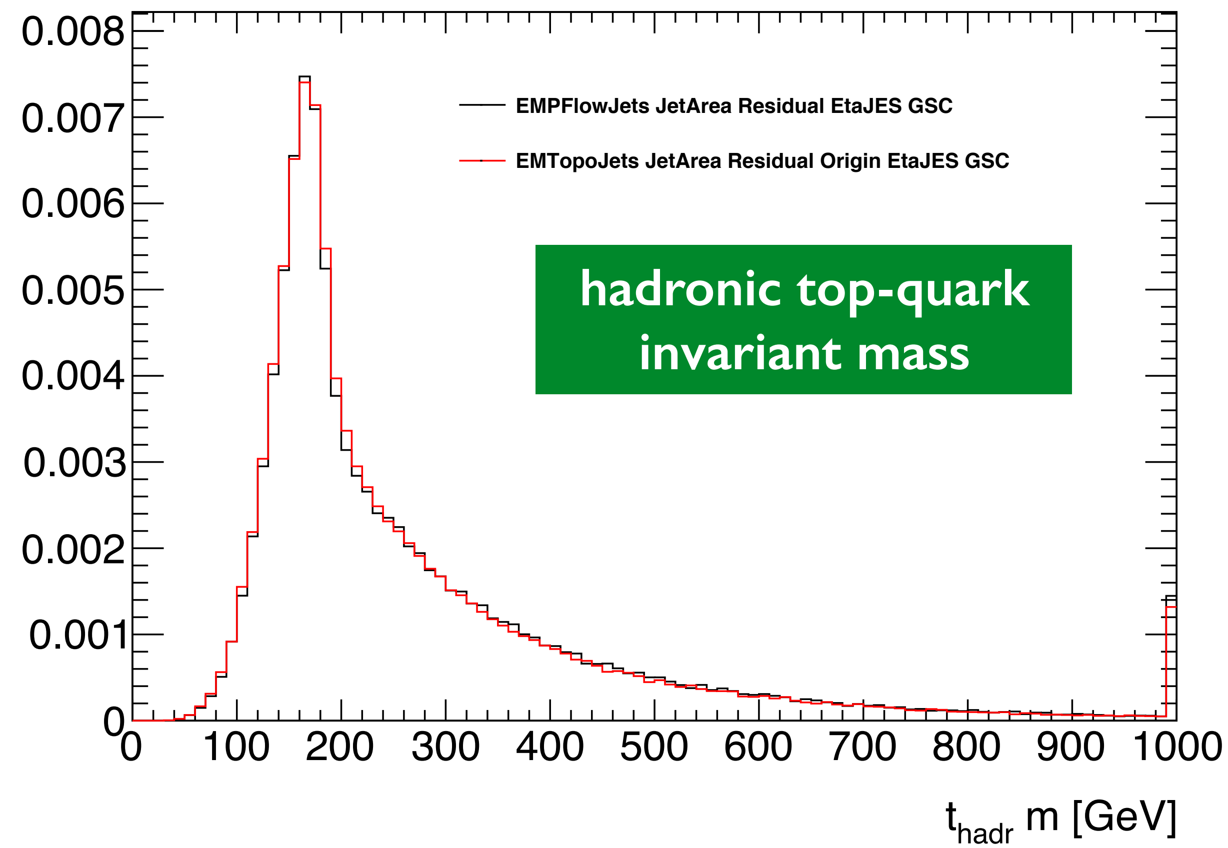
First look:

P. Falke

Jet p_T resolution

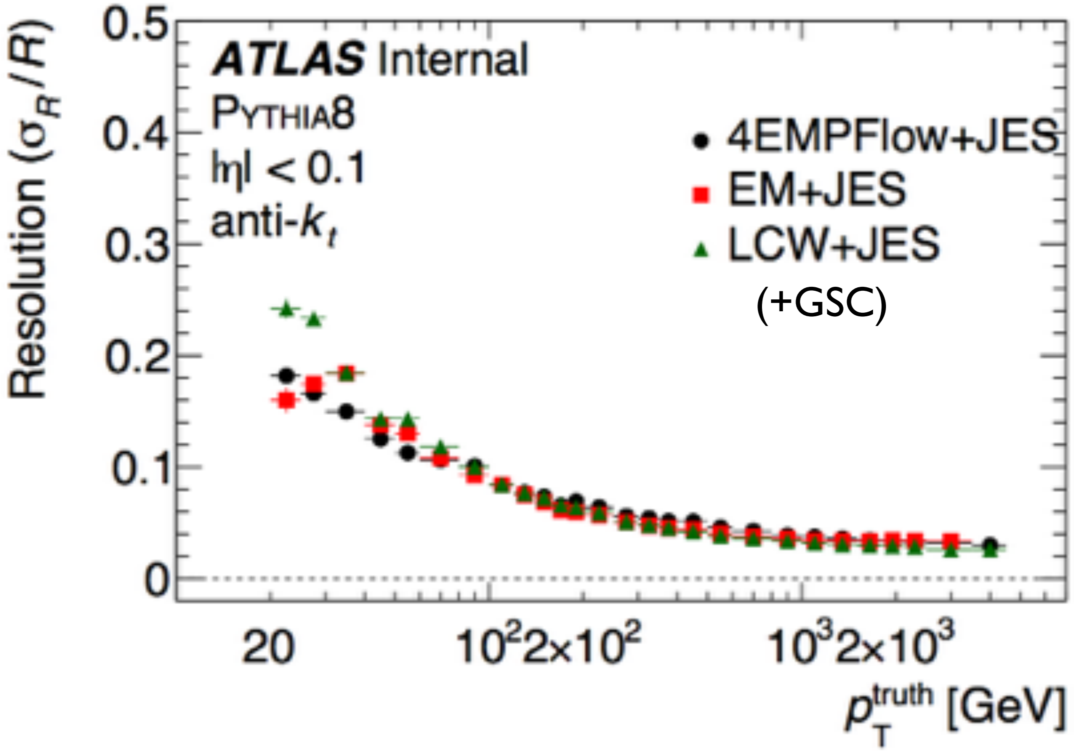






PFlow current/future status

- The PFlow algorithm is in a good shape!
  - Full calibration sequence already in place
  - Optimisation of the b-tagging in progress
  - Systematics are being derived



- **Rel 20.7** (analysis post-Moriond)
  - New charge particle subtraction
  - Updated lepton identification
  - Bug fix for tight tracks included
  - Should be good for physics results

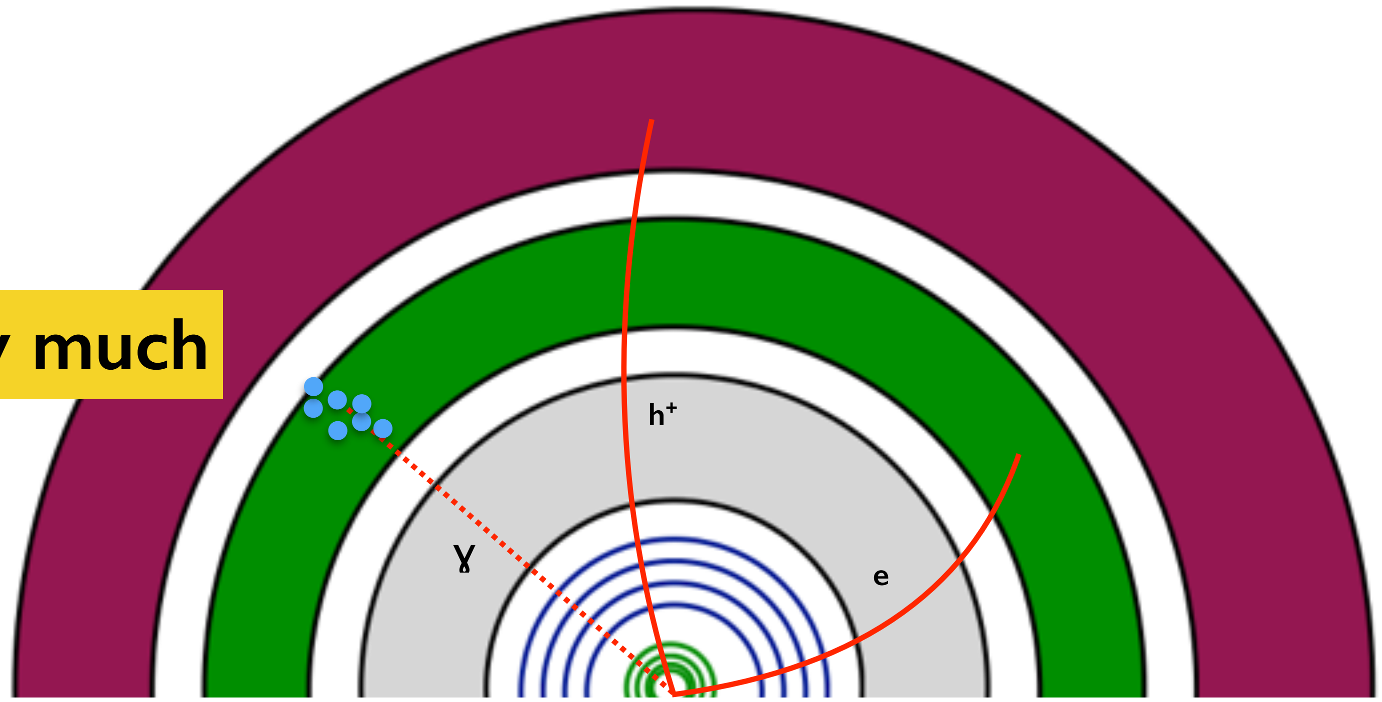
- **Rel. 21**
  - Several improvements in the pipeline:
    - Optimisation of the charge particle subtraction
    - Track-Matching p_T dependent cuts
    - …

Summary

- Performance studies have demonstrated advantages for the particle-flow objects:
  - PFlow jet resolution is better than EM (LC) calorimeter jets at low p_T and comparable above ~80 GeV
  - Suppression and stability of the pileup contribution
  - Better angular determination (in η and Φ)
  - Improvements in E_T^{miss} and large-R jets performance

- Run 2 data allows the particle flow algorithm to be further optimised
- It is time for starting having a look at **particle-flow** in **top physics**! 😉

Thank you very much



Backup

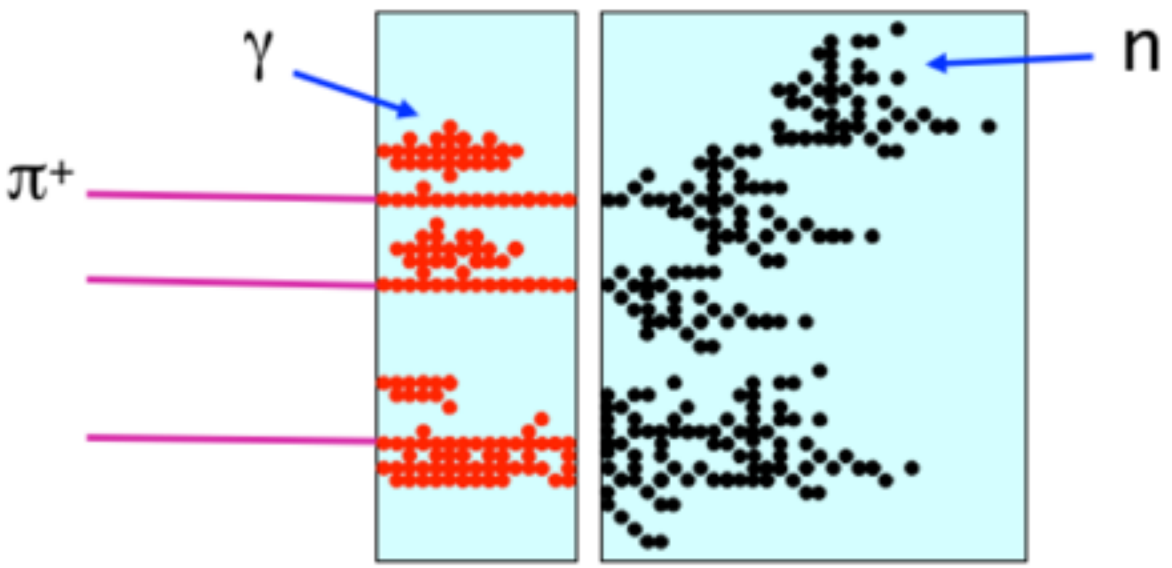
Introduction: What is Particle flow?

First, have a look at the composition of a typical jet:

- 60% of jet energy in charged hadrons
- 30% in photons (mainly from $\pi^0 \rightarrow \gamma\gamma$)
- 10% in neutral hadrons (n, K,…)

Traditional calorimetric approach:

- Measure all components of jet energy in calorimetry system
- 70 % of energy use information form the hadronic calorimeter
- Intrinsically HAD calorimeter resolution limits jet energy resolution
  - HAD Calo. ATLAS: $\frac{\sigma(E)}{E} \approx \frac{50\%}{\sqrt{E}} \oplus 3\%$



$$E_{JET} = E_{ECAL} + E_{HCAL}$$

Introduction: What is Particle flow?

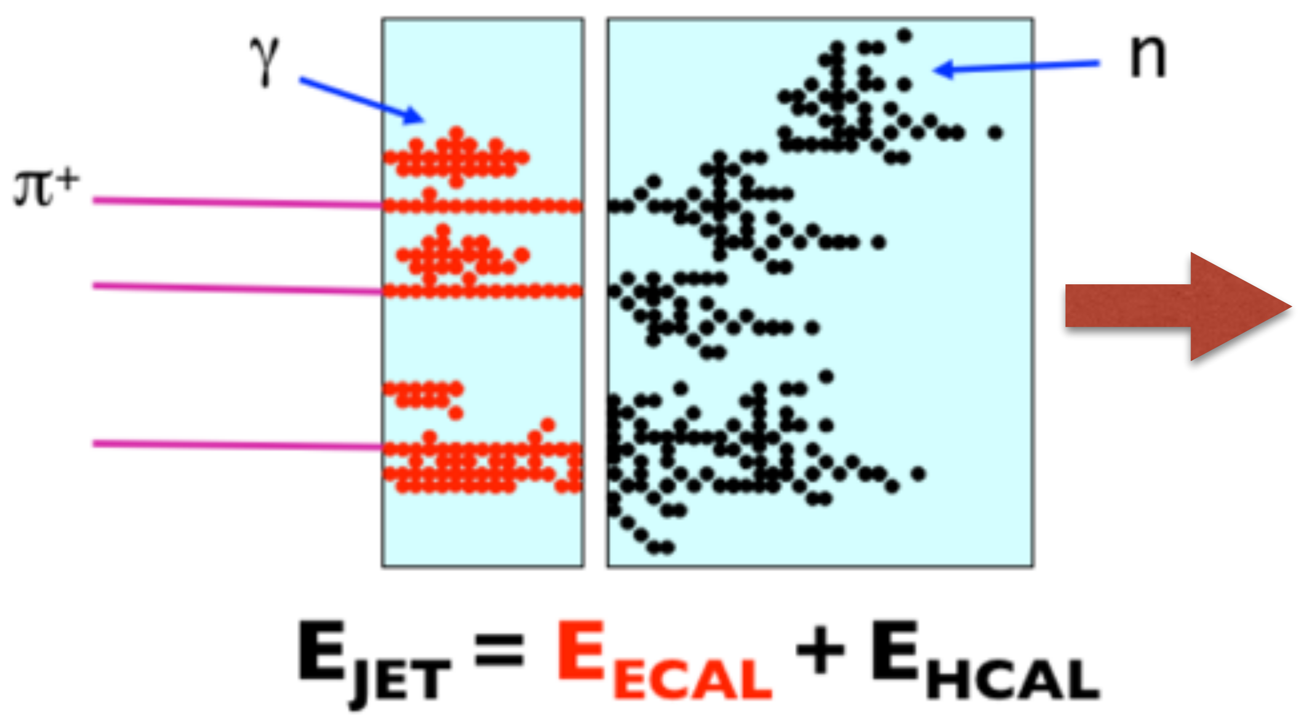
First, have a look at the composition of a typical jet:

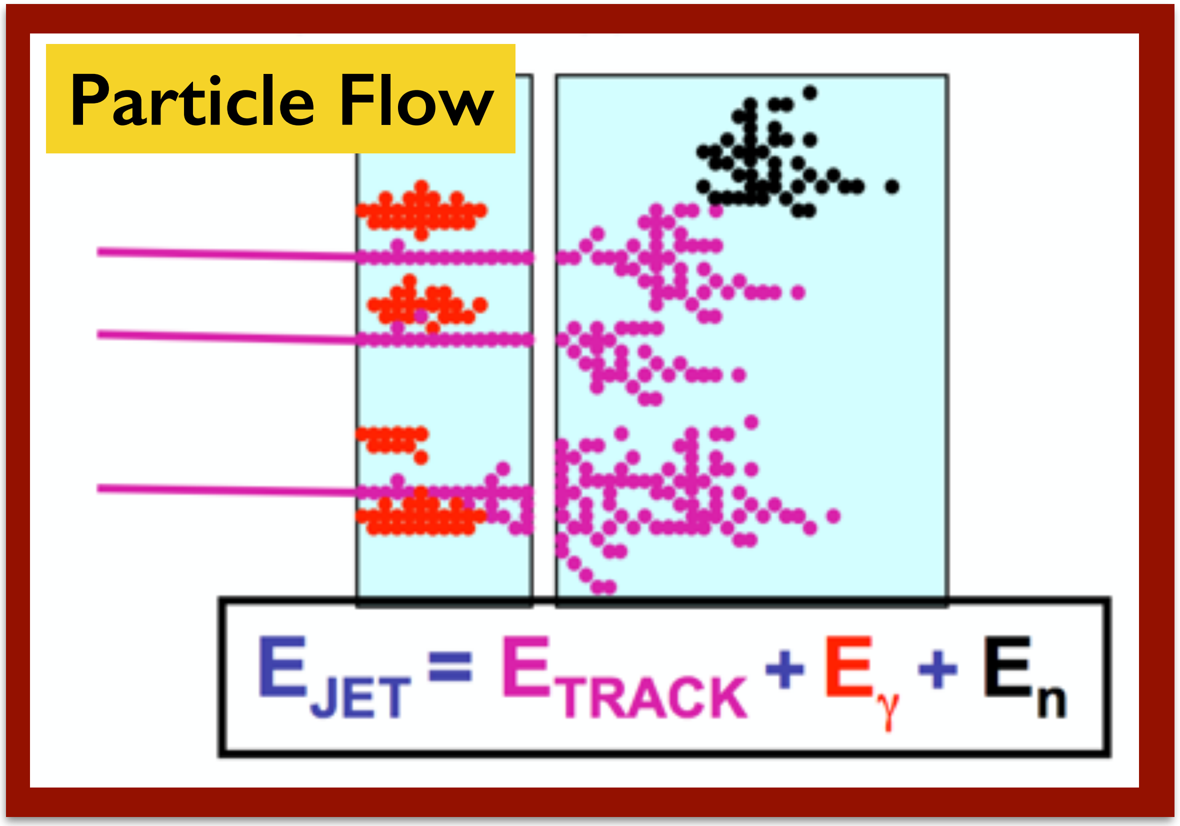
- 60% of jet energy in charged hadrons
- 30% in photons (mainly from $\pi^0 \rightarrow \gamma\gamma$)
- 10% in neutral hadrons (n, K,…)

Particle Flow approach:

- Charged particles measured in tracker
- Photons in EM Calorimeter
- Neutral hadrons in HAD Calorimeter

Only 10 % of jet energy from HAD Calorimeter ➜ much improved resolution





Why Particle flow in ATLAS?

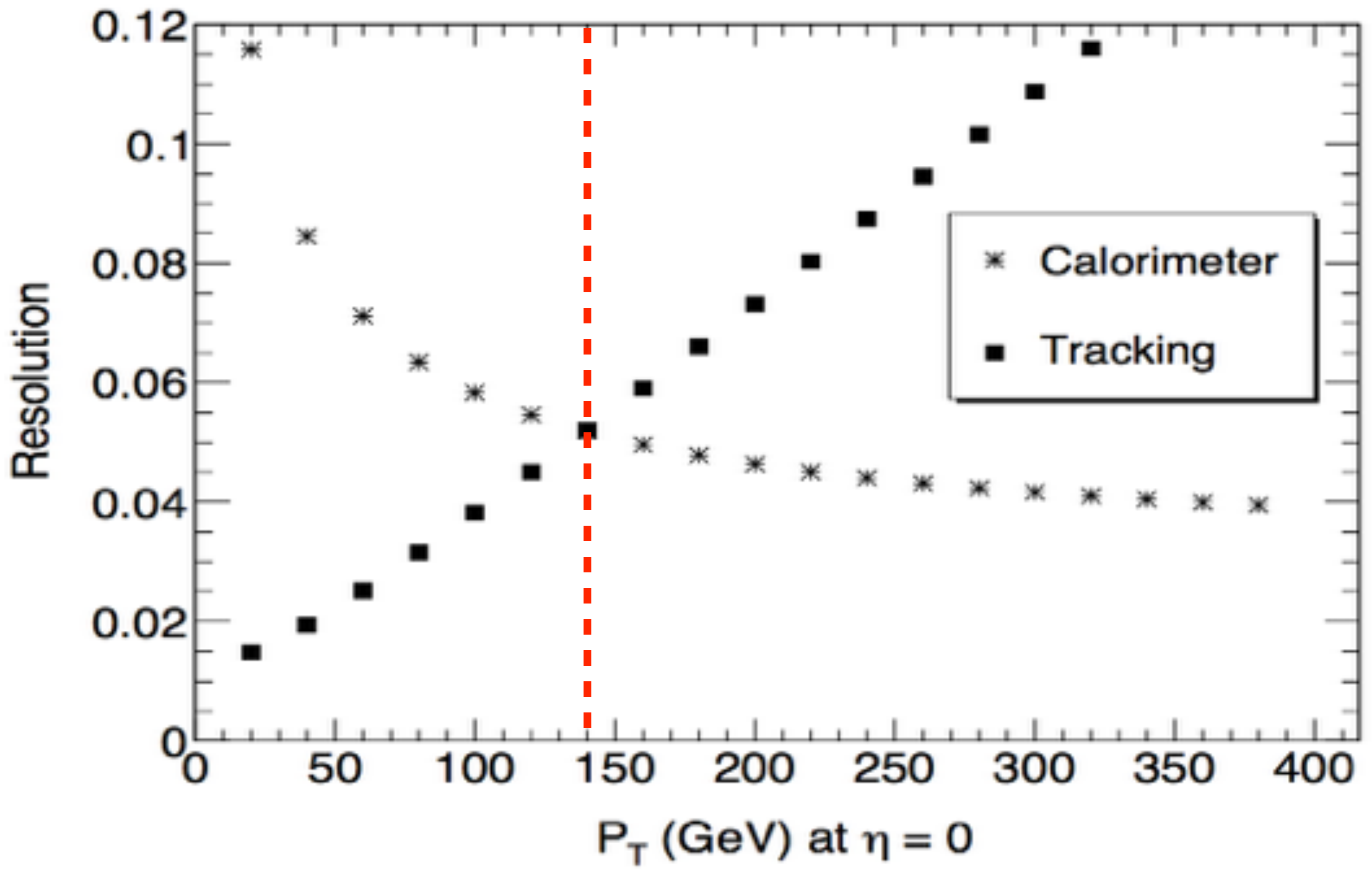
Some reasons for using particle flow in ATLAS:

- The tracker resolution is significantly better than the calorimeter resolution at low p_T
- Low p_T particles, which not pass the threshold required to create a topocluster, can be measured by the tracker
- Better angular resolution of the tracker compared with the calorimeter for single particles
- When the track is reconstructed, the vertex information from which the track comes, can be used to mitigate the pileup contribution
- Calorimeter's ability to reconstruct neutral particles

Particle flow algorithm (eflowRec) decides to use the tracker information or the calorimeter information or a combination of both

$$\frac{\sigma(p_T)}{p_T} \approx 0.036 p_T \% \oplus 1.3\%$$

$$\frac{\sigma(E)}{E} \approx \frac{50\%}{\sqrt{E}} \oplus 3\%$$



Particle flow low-level performance studies

Important quantities for pflow performance studies

Useful for track-cluster matching:

- **Efficiency**: sum of the energy deposited by the true particle in the topo-cluster divided by the total energy deposited by the particle in the calorimeter

$$\epsilon = \frac{\Sigma E_{CalHit}(cluster)}{\Sigma E_{CalHit}(all\ clusters)}$$

- **Purity**: sum of the energy deposited by the true particle in the topo-cluster divided by the sum of the energy of all true particles in this cluster

$$\rho = \frac{\Sigma E_{CalHit}(cluster)}{\Sigma E_{CalHit}(all\ true\ particles\ in\ the\ clusters)}$$

Useful for subtraction studies:

- **R^0**: true neutral energy left after subtraction

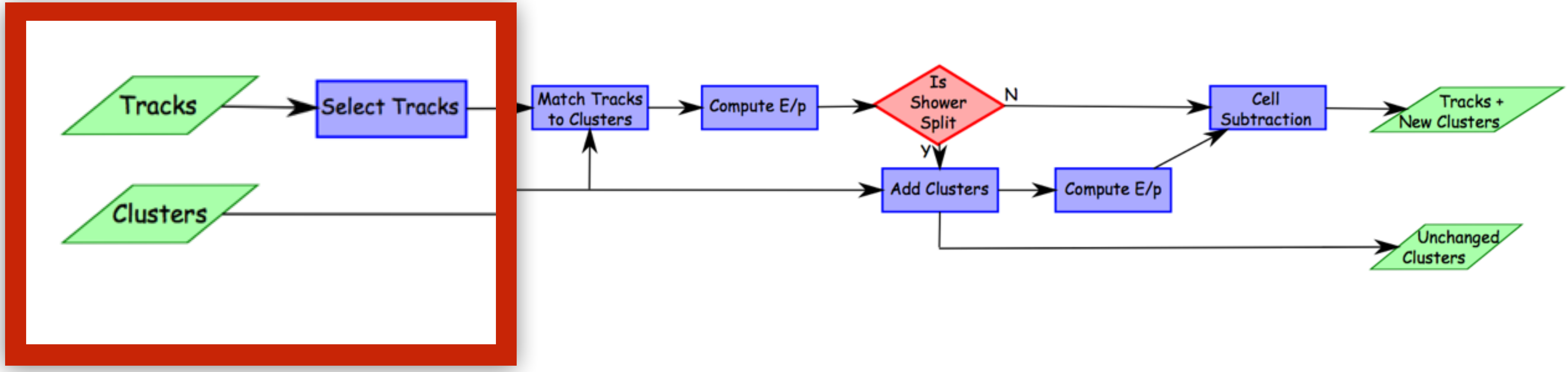
$$R^0 = \frac{\sum_{\text{neutral}} E_{\texttt{CalHit}}(\text{after subtraction})}{\sum_{\text{neutral}} E_{\texttt{CalHit}}(\text{before subtraction})}$$

- **R^+**: true charged energy subtracted

$$(1 - R^+) = 1 - \frac{\sum_{\text{charged}} E_{\texttt{CalHit}}(\text{after subtraction})}{\sum_{\text{charged}} E_{\texttt{CalHit}}(\text{before subtraction})}$$

Overview of the eflowRec chain:

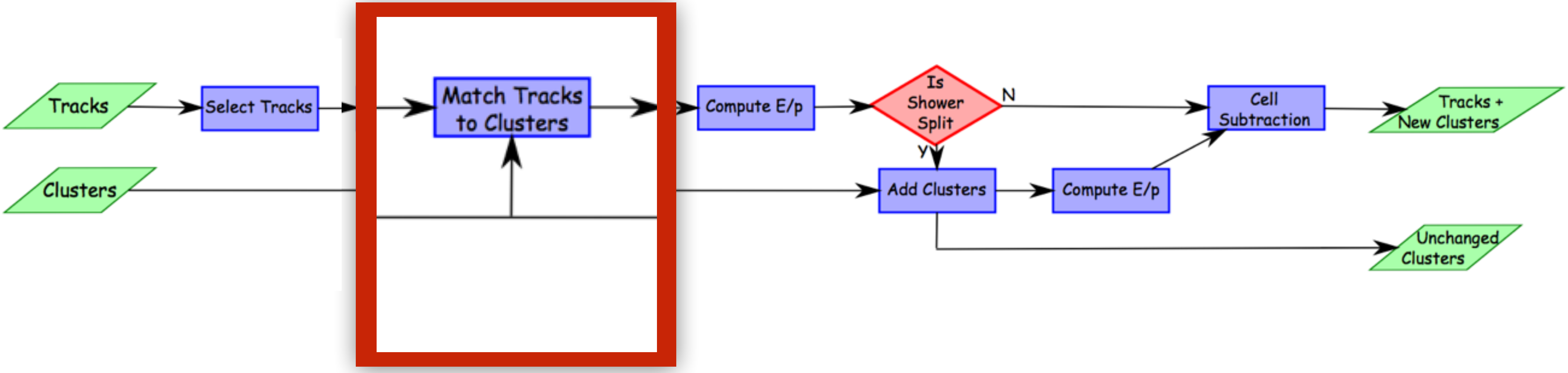
1. Input objects selection:



- **Tracks:** selected tracks with a reliable set of properties
  - Tight track selection (CP group recommendations)
  - $|\eta|<2.5$ & $p_T > 500$ MeV
  - $p_T < 40$ GeV
- **Topological clusters:** all considered

Overview of the eflowRec chain:

2. Track-Cluster matching:



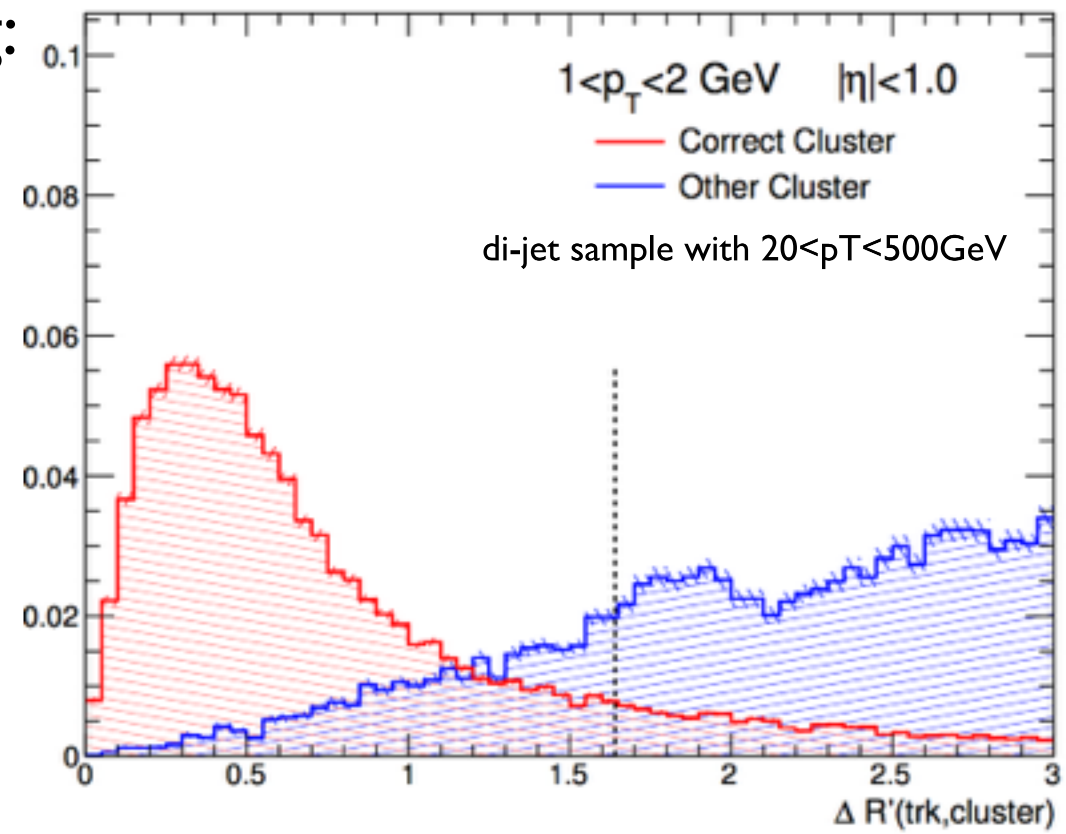
To remove the energy deposited in the calorimeter → match each track to a cluster

- Each track is extrapolated to the 2nd layer of the EM calorimeter
- The nearest topological cluster is found using:

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

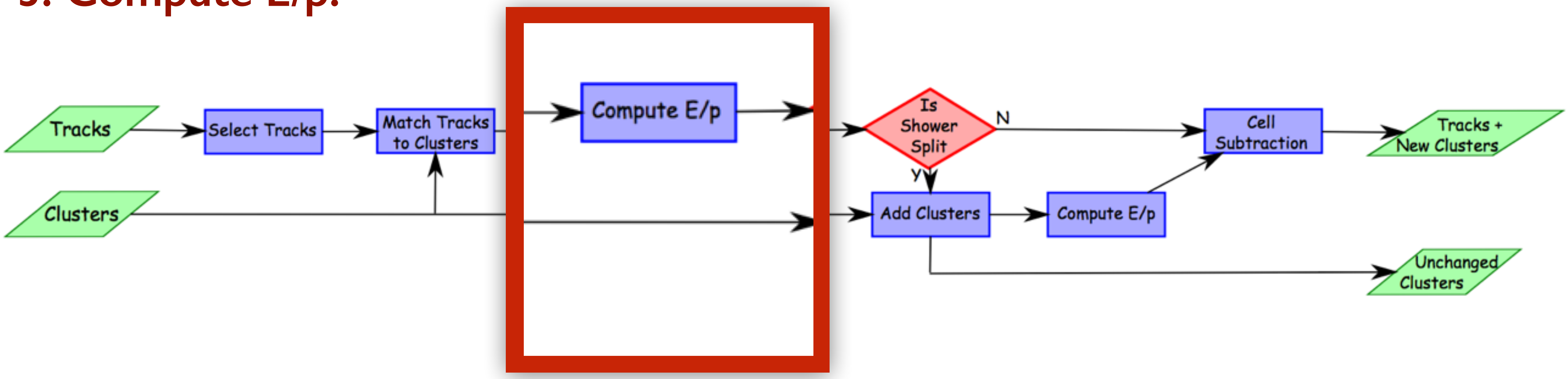
$\sigma_{\eta,\phi}$ of the cluster width

Correct cluster = cluster with E_{cl}>90% E_{true}
Other cluster = next closest cluster



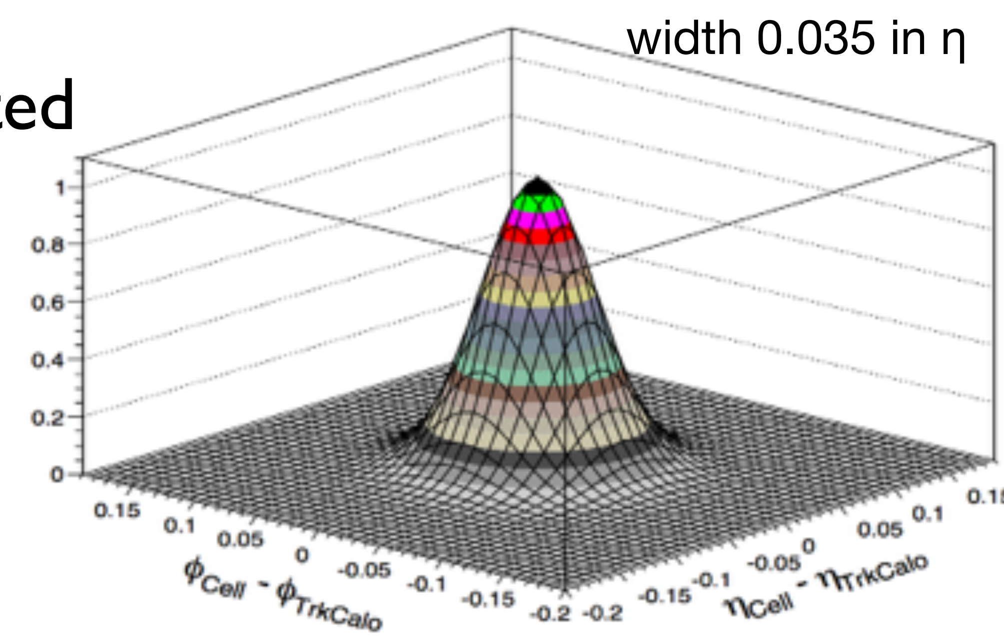
Overview of the eflowRec chain:

3. Compute E/p:



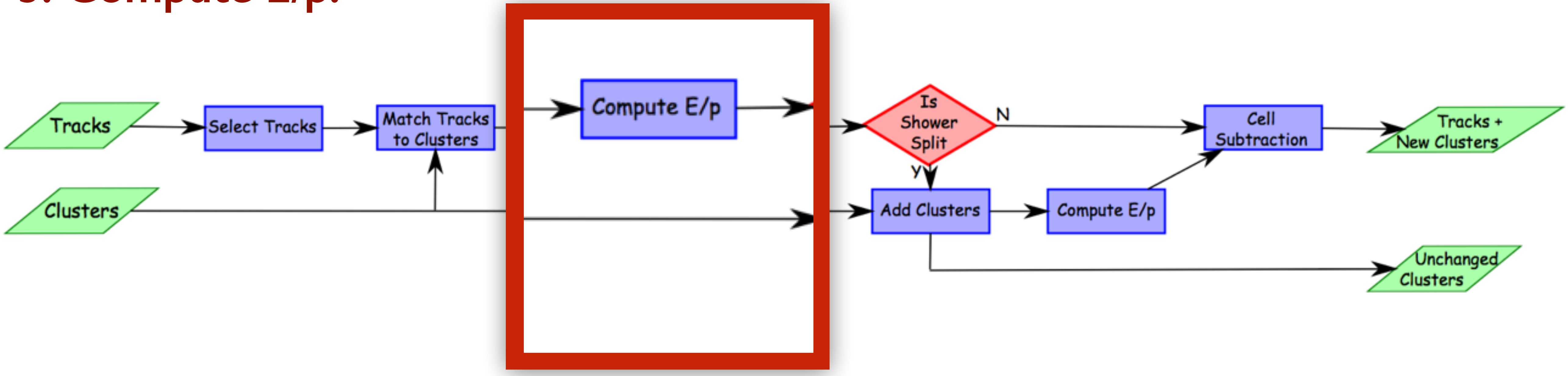
To remove the energy in the calorimeter → how much energy is expected where?

- **LHED** = layer of highest energy density
  - Motivation: find and remove the dense EM core of the shower
  - How to proceed?
    - Scale the cells around the tracks extrapolated position using a Gaussian
    - Calculate the average energy density per radiation length in each layer <ρi'>
    - Take the layer with larges <ρi'>



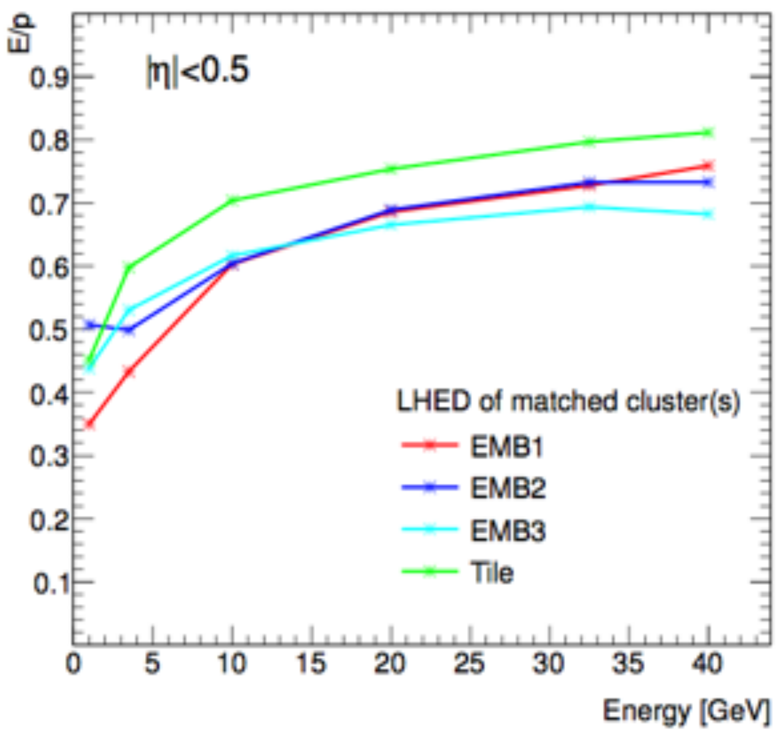
Overview of the eflowRec chain:

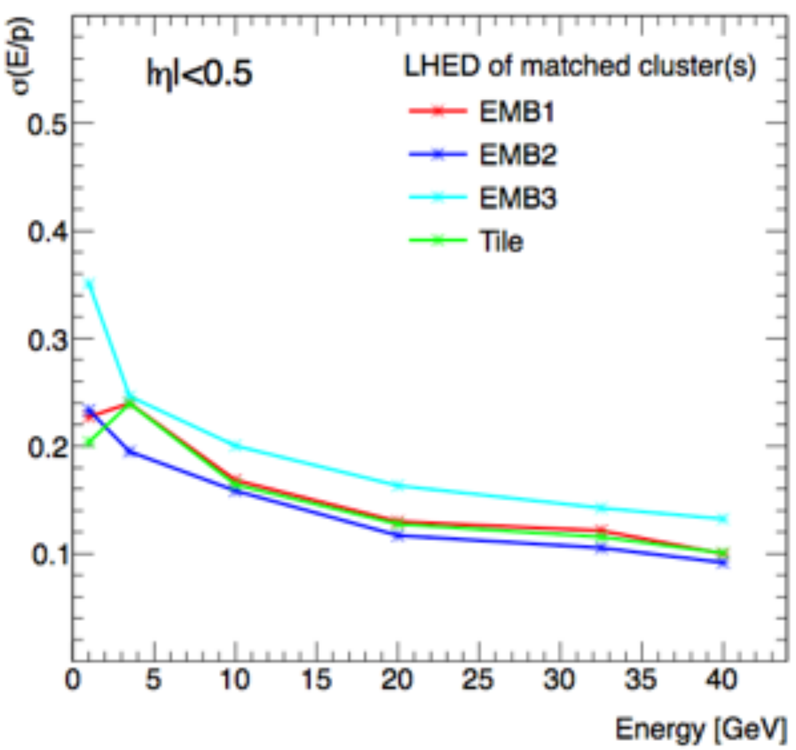
3. Compute E/p:



To remove the energy in the calorimeter → how much energy is expected where?

- E/p: ratio of the E deposited in the calorimeter divided by the p of the track.
  - Single particle samples used to determine E/p
  - E obtained as sum of the clustered E in a cone ΔR<0.2
- The E/p is parametrised in terms of η_{part}, E_{part} , LHED

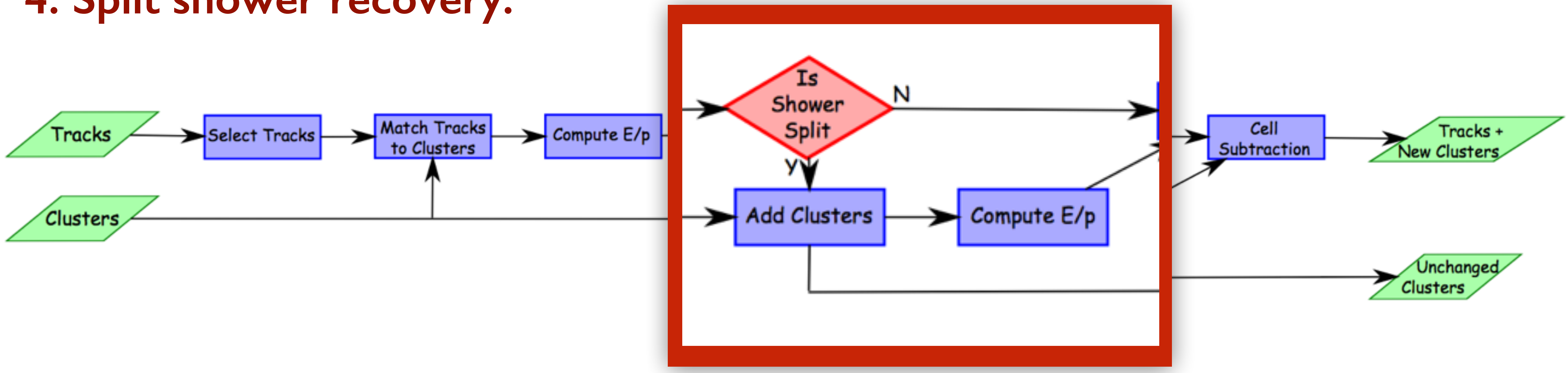




(LHED = layer of highest energy density)

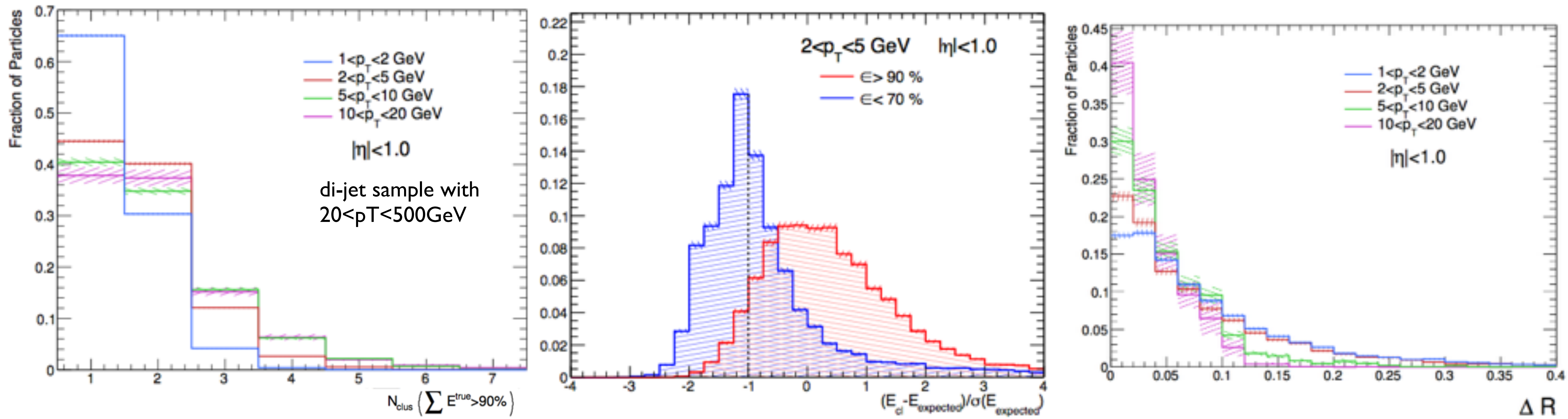
Overview of the eflowRec chain:

4. Split shower recovery:



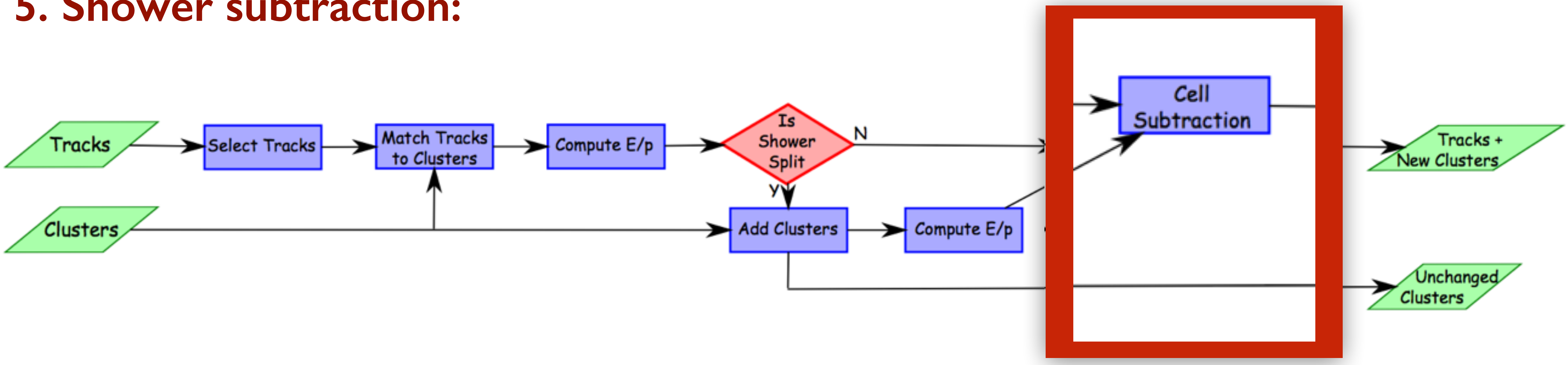
Particles do not always deposit all their energy in a single cluster.

- If E_{clus} is found to be less than E_{exp} by more than $\sigma(E_{exp})$
  - Assume energy has been split over multiple clusters (split shower recovery)
  - All clusters ΔR<0.2 are considered



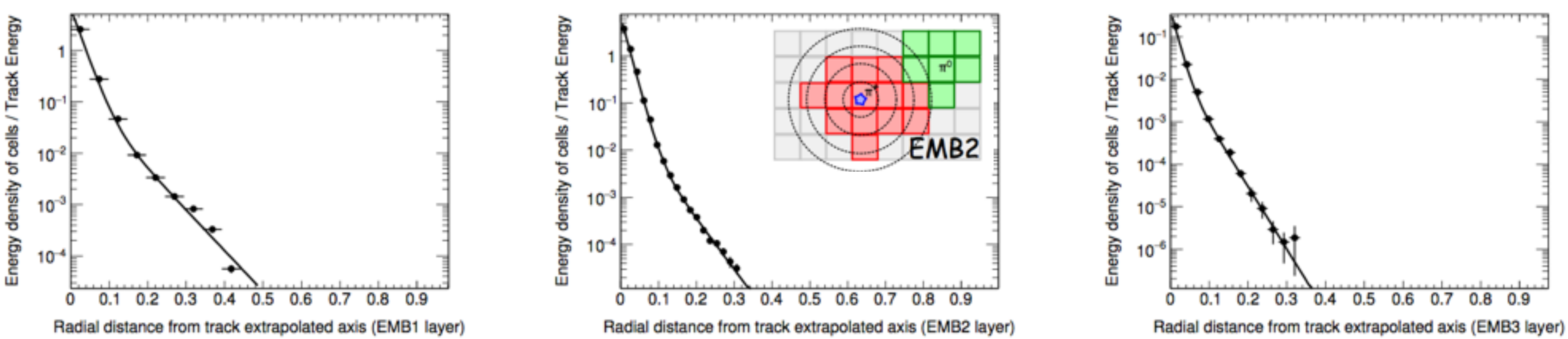
Overview of the eflowRec chain:

5. Shower subtraction:



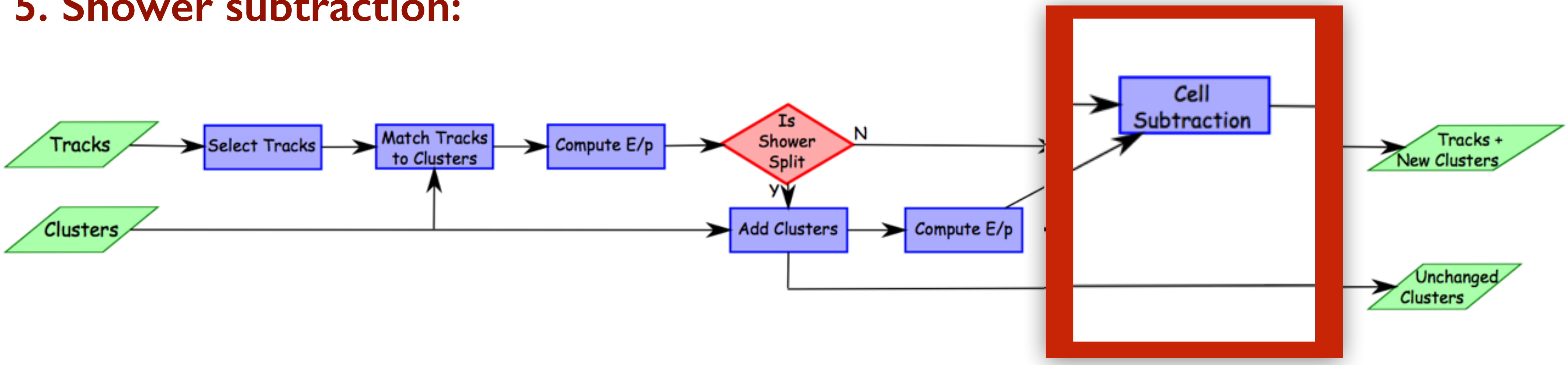
Set of clusters selected → subtraction procedure

- If E_{clus} < expected energy from E/p → completely subtracted
- If E_{clus} > expected energy from E/p → ring-by-ring subtraction
  - Parametrised shower shape in each layer (using single pion sample)
  - Using the cell size, rings in (eta, phi) around track direction defined for each layer
  - Average energy density calculated for each ring



Overview of the eflowRec chain:

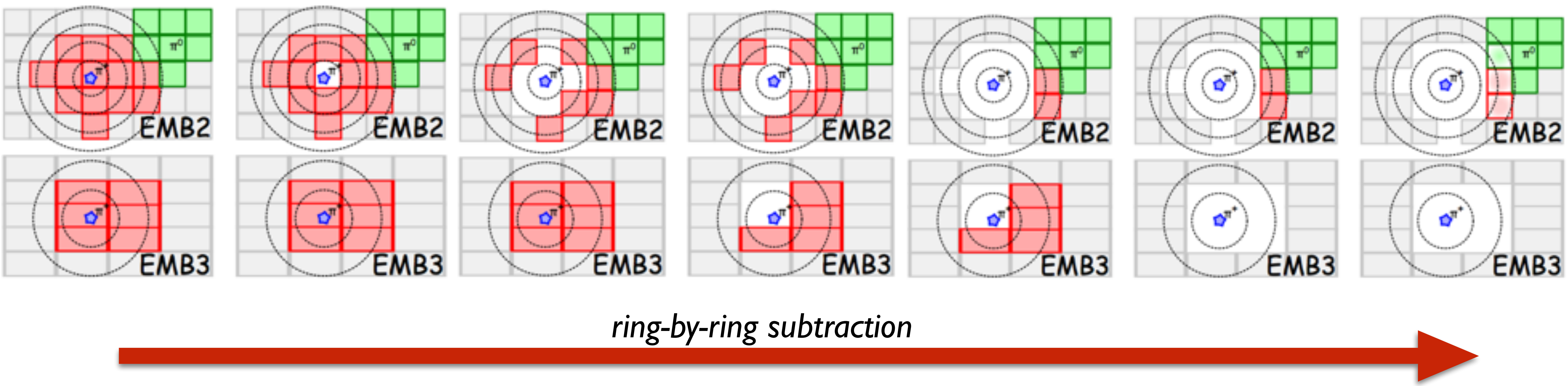
5. Shower subtraction:



Subtraction starts from the highest E density ring

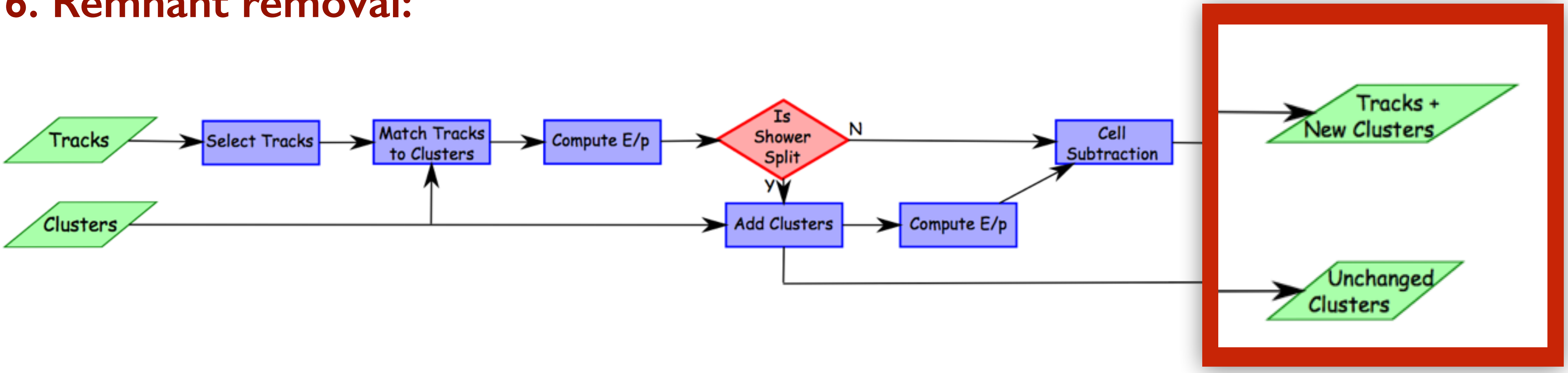
- E_{ring} < E remaining to subtract → removed
- E_{ring} > E remaining to subtract → scaled by the fraction needed

‣ The process of removing cells ring-by-ring is continued until the E_{exp} is subtracted



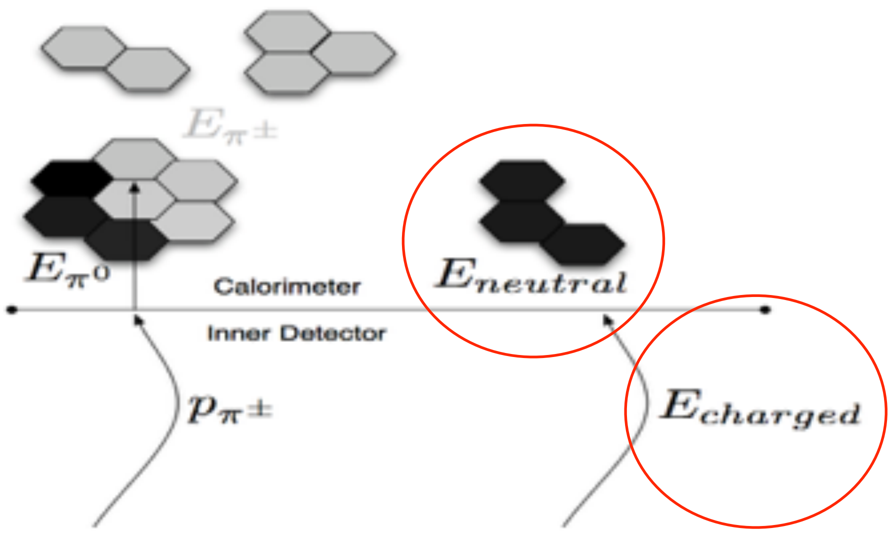
Overview of the eflowRec chain:

6. Remnant removal:

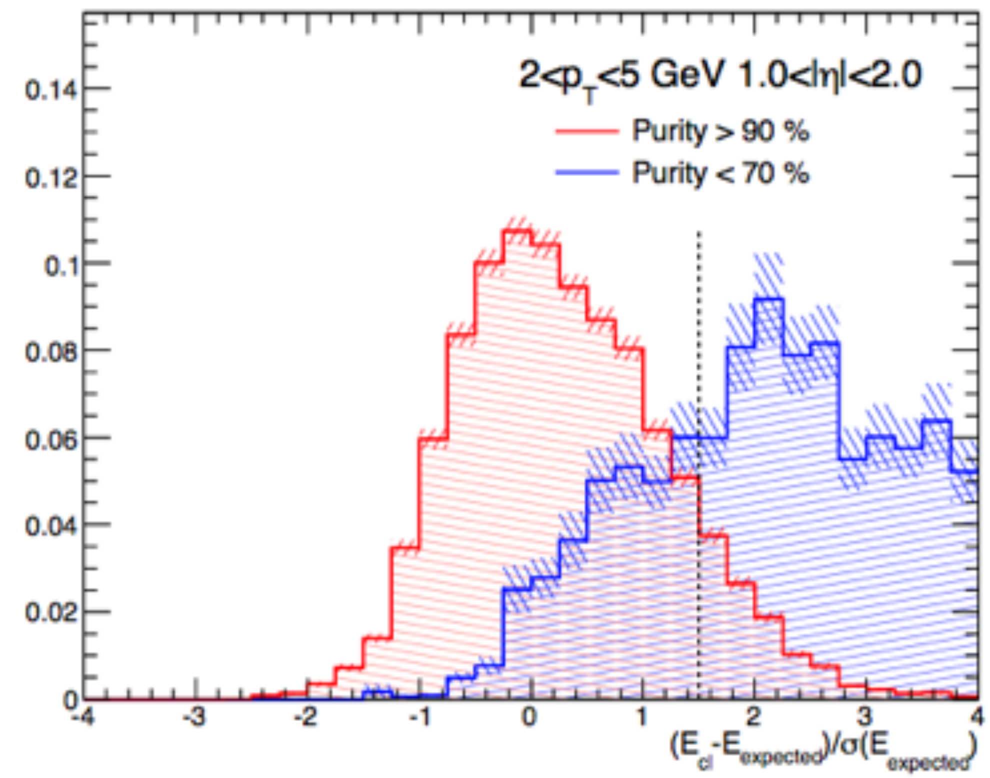


Subtraction stops when E_{exp} has been removed

- If the remaining E is consistent with σ(E/p) → purely noise → totally removed
- If the remaining E larger than σ(E/p) → other particles involved → kept



The selected tracks and the remained clusters represent the reconstructed event

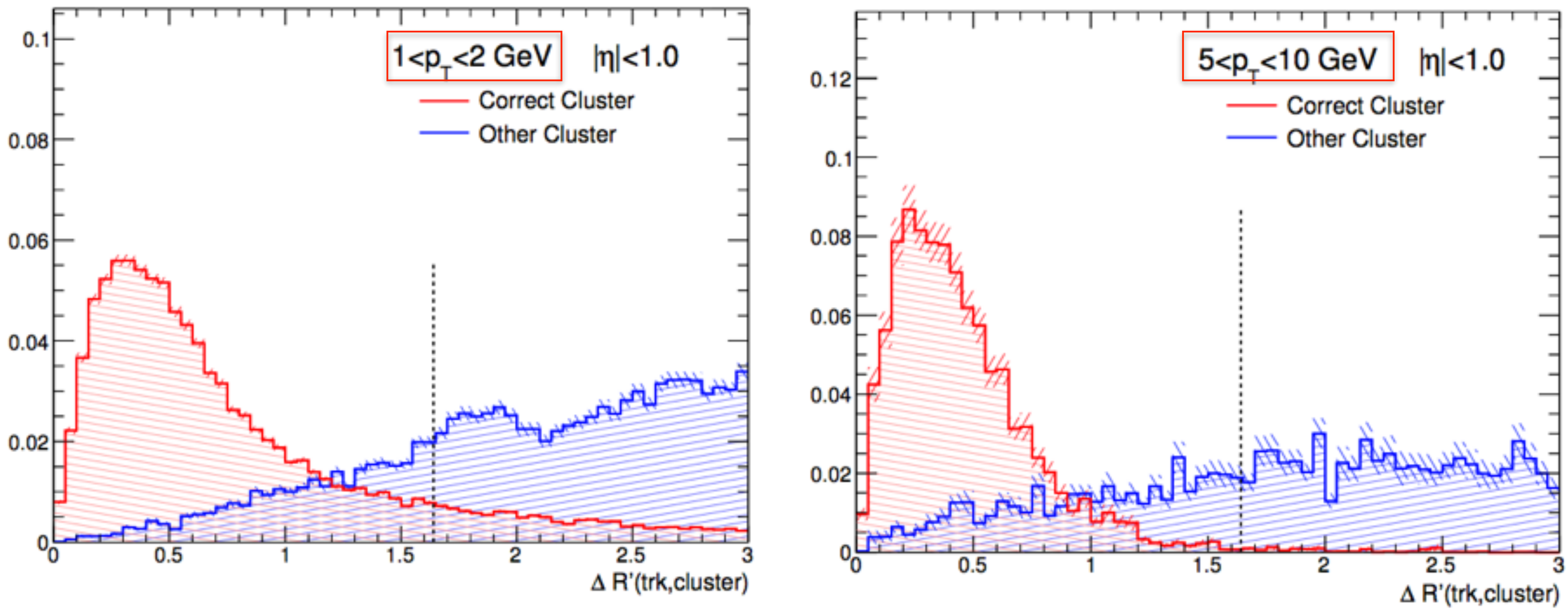


Particle flow low-level performance studies

Track-cluster matching

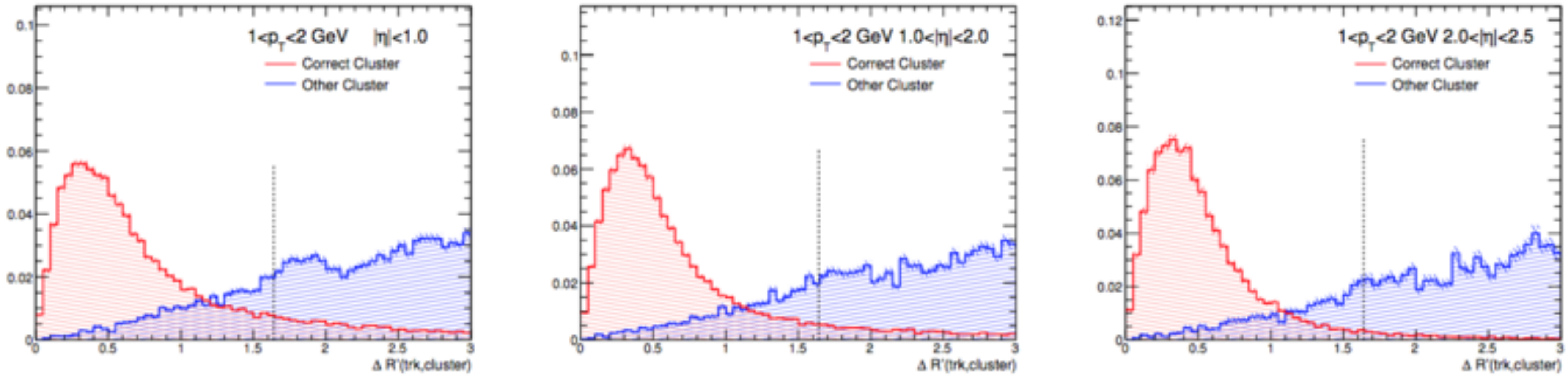
The optimisation of the track-cluster matching at the first step of the algorithm has been studied:

- Extrapolate low momentum tracks to EM1 instead of EM2 ✗
- Extrapolate to the layer with lower ΔR' ✗
- Extrapolate to the layer where most of the energy is deposited ✗
- Use ΔR instead of ΔR' ✗
- Momentum dependence of the ΔR' ✓ (see Marianna's slides)

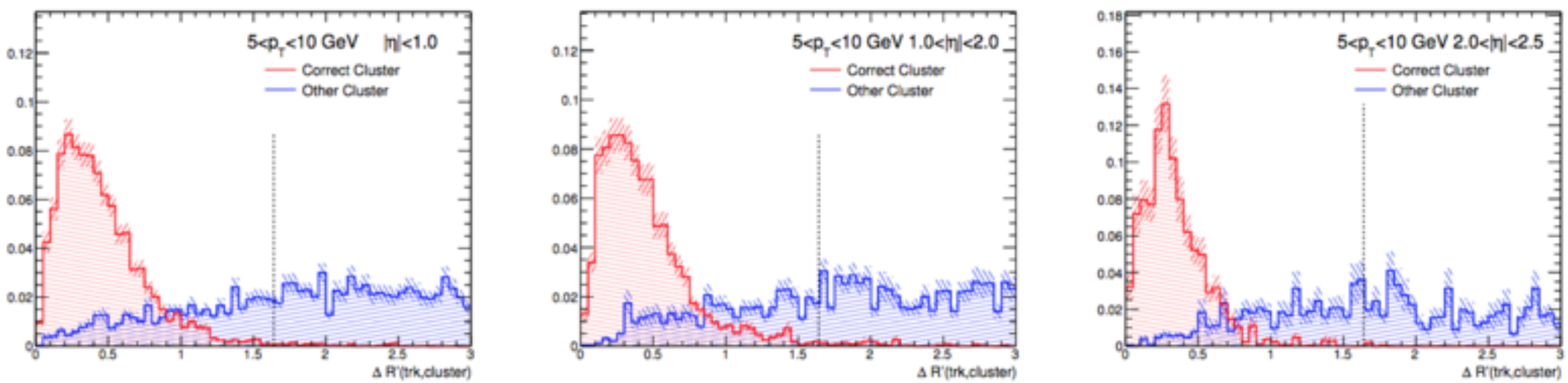


Track-cluster matching

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



(a) $1 < p_{\rm T}^{\rm true} < 2\,{\rm GeV}\ |\eta| < 1.0$ (b) $1 < p_{\rm T}^{\rm true} < 2\,{\rm GeV}\ 1.0 < |\eta| < 2.0$ (c) $1 < p_{\rm T}^{\rm true} < 2\,{\rm GeV}\ 2.0 < |\eta| < 2.5$

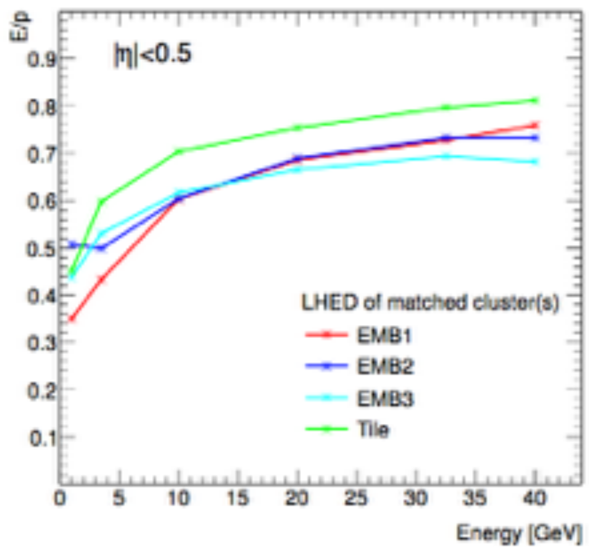


(d) $5 < p_{\rm T}^{\rm true} < 10\,{\rm GeV}\ |\eta| < 1.0$ (e) $5 < p_{\rm T}^{\rm true} < 10\,{\rm GeV}\ 1.0 < |\eta| < 2.0$ (f) $5 < p_{\rm T}^{\rm true} < 10\,{\rm GeV}\ 2.0 < |\eta| < 2.5$

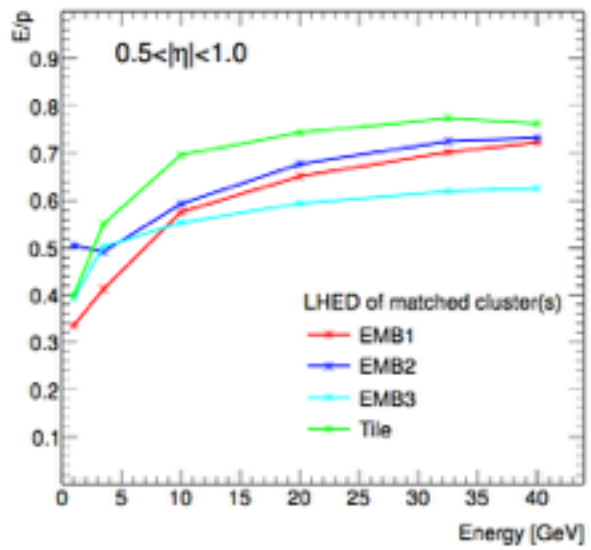
Figure 9: The distribution of $\Delta R'$ for the cluster with the $> 90\,\%$ of the true energy of the particle and the next closest cluster satisfying $E/p_{\rm track} > 0.1$. The data are taken from a dijet sample with $20 < p_{\rm T}^{\rm lead} < 500\,{\rm GeV}$ and the errors shown are MC statistics.

E/p

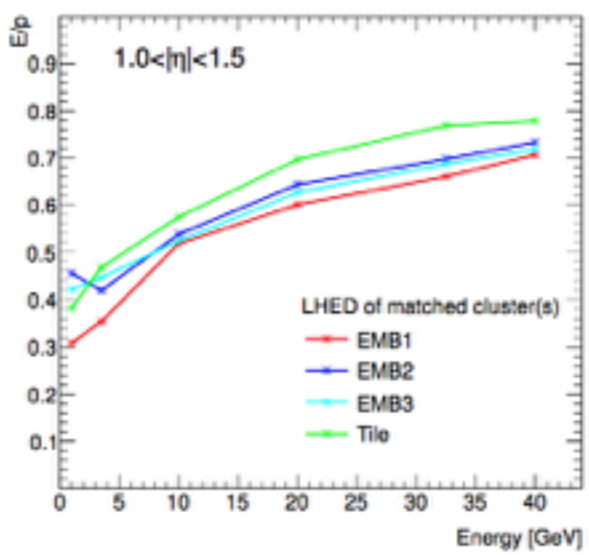
E/p

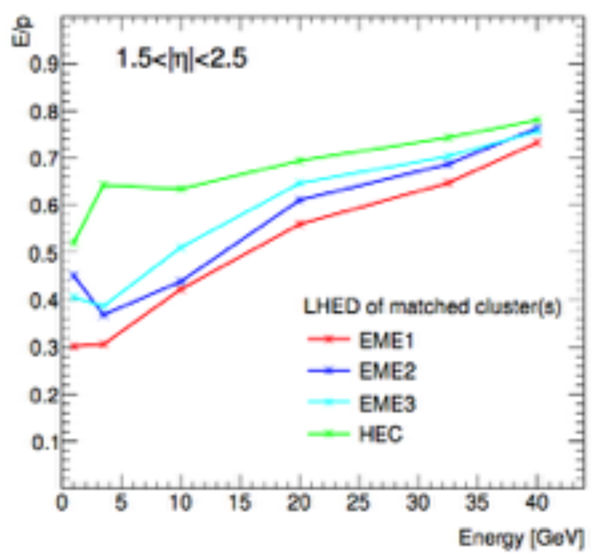


(a) $|\eta| < 0.5$

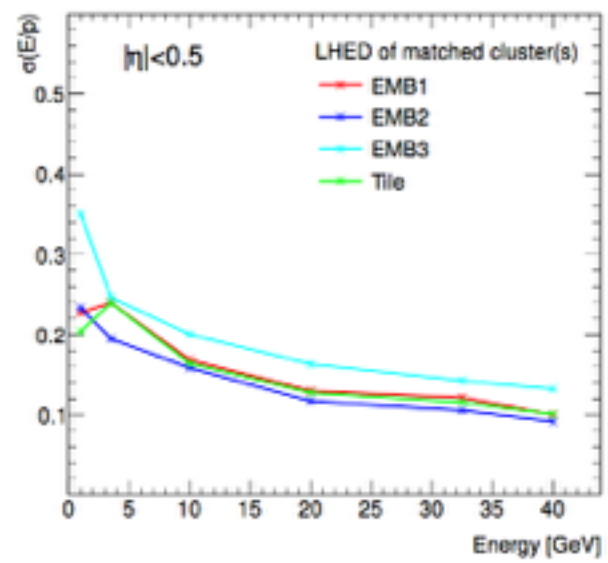


(b) $0.5 < |\eta| < 1.0$

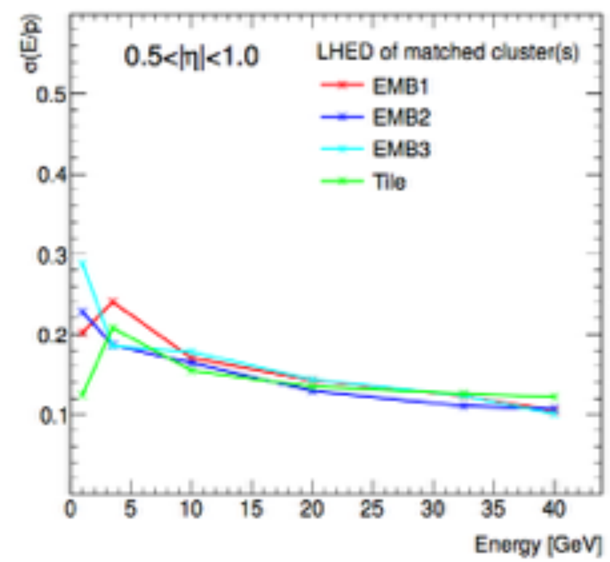




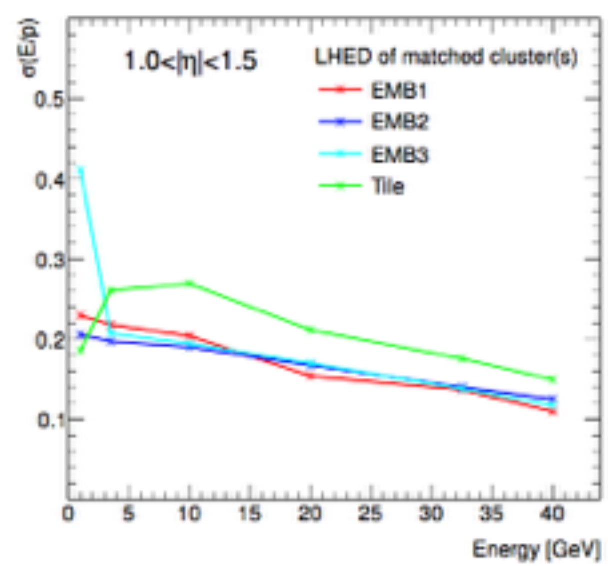
σ(Eexp)

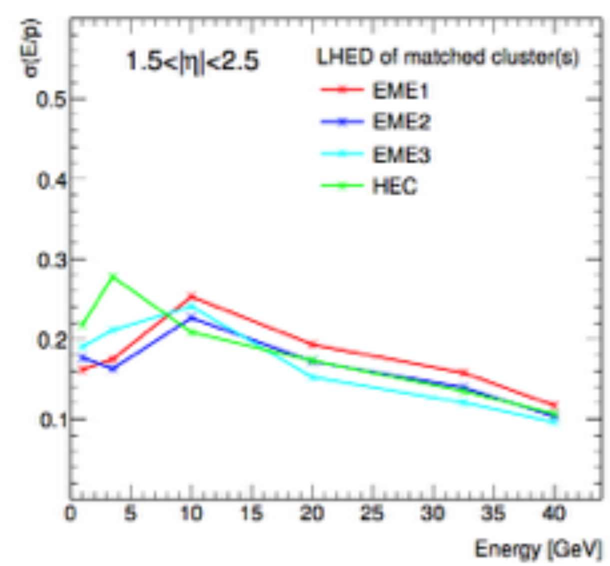


(a) $|\eta| < 0.5$



(b) $0.5 < |\eta| < 1.0$





First look:

P. Falke

Jet angular resolution