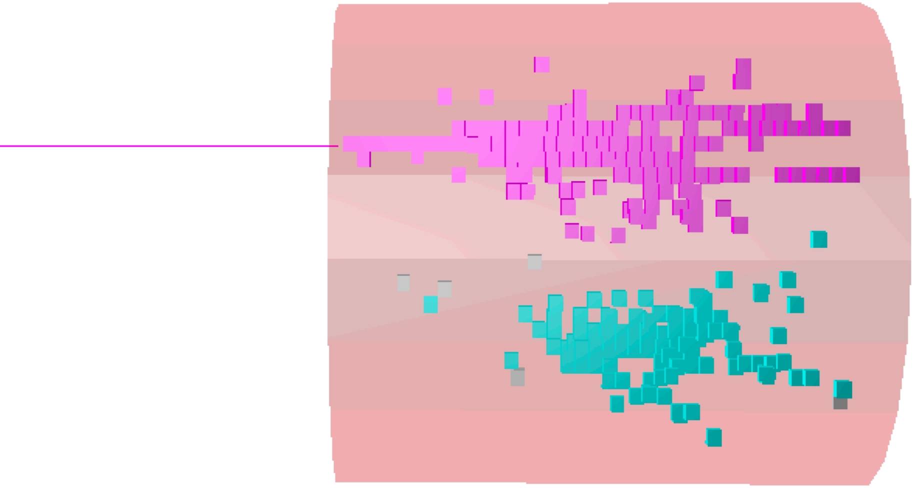
Pandora Particle Flow Algorithm Studies

With AHCAL 2018 Beam Test Data



Magenta: Charged Hadron
Cyan: Neutral Hadron
Grey: Unclustered Hits

Daniel Heuchel (DESY)
daniel.heuchel@desy.de
Virtual CALICE Collaboration Meeting
25th March 2021

Work done in cooperation with
Linghui Liu (University of Tokyo)

Outline

For this Talk

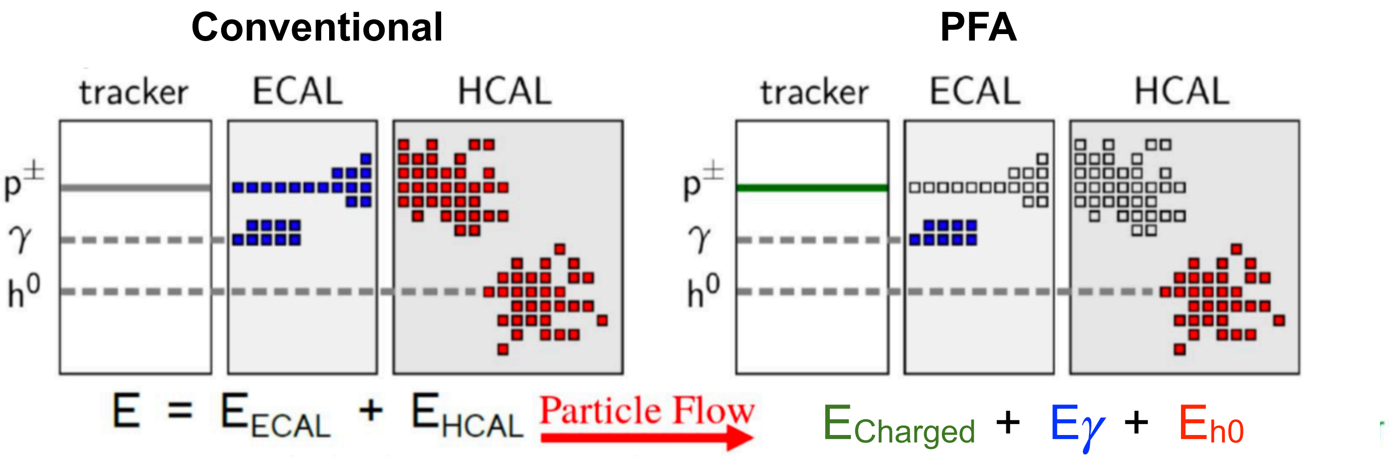
- Recap: Concept of Particle Flow & Confusion
- Motivation & Goals of Study
- Results: PandoraPFA Two Particle Reconstruction

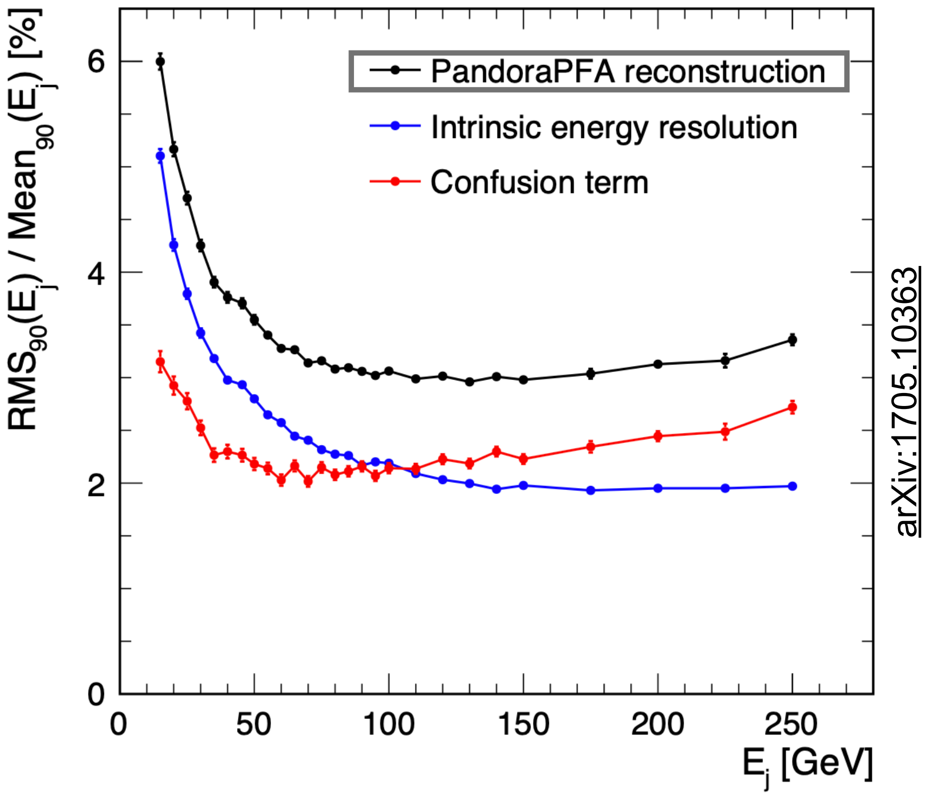
- Summary & Outlook

Particle Flow Approach

The Key to Highest Precision

- Goal at future e^+e^- collider experiments: Jet energy resolution of 3-4% for jet energies between 40-500 GeV
  - ➡ PFA: Measure energy/momentum of each particle with detector providing best resolution
    - ➡ Make use of excellent resolution of tracker (for ~60% charged particles in jets)



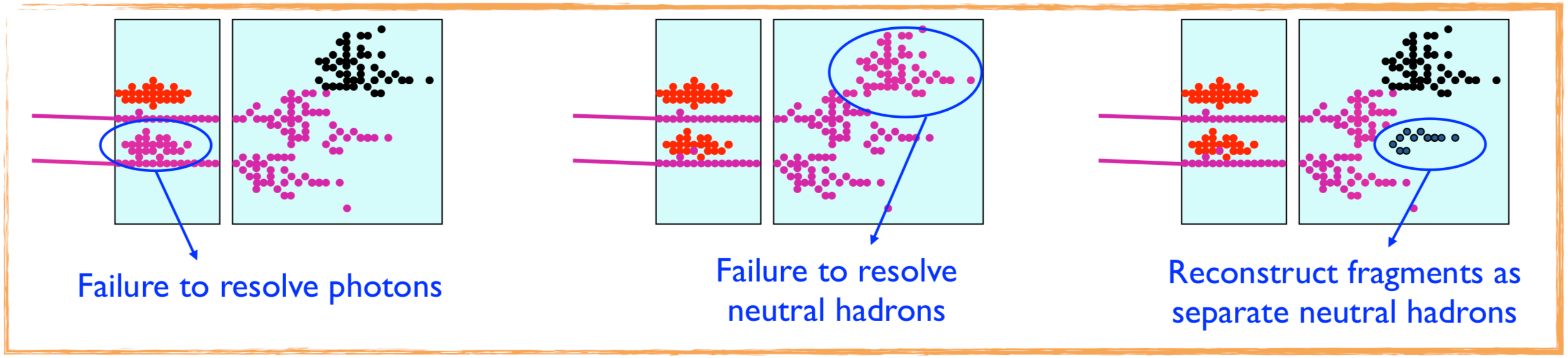


Confusion Scenarios

The Limit of Particle Flow Reconstruction

- **Topologically or energetically confusing** events could cause problems for PFA reconstruction:

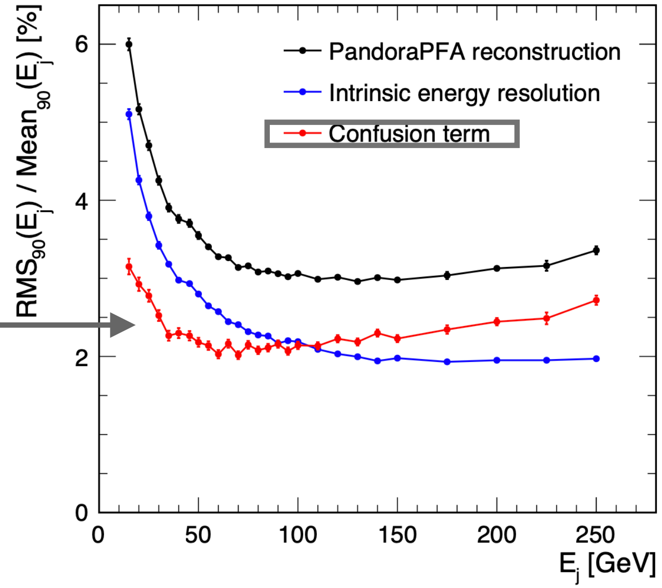
Types of confusion



Missing energy | **Missing energy** | **Double counted energy**

J. S. Marshall: https://indico.in2p3.fr/event/7691/contributions/42712/attachments/34375/42344/3_john_marshall_PFA_marshall_24.04.13.pdf

➡ Missing or double counted energy limiting jet energy resolution at high energies



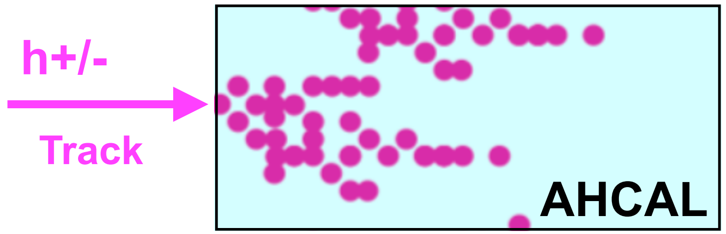
arXiv:1705.10363

Motivation and Goals of Studies I

PandoraPFA on AHCAL 2018 Prototype Data

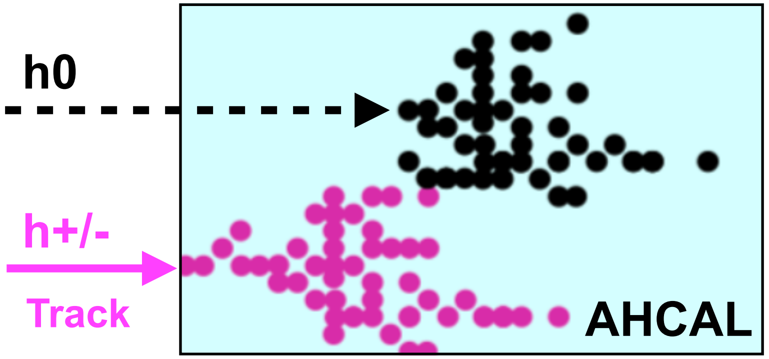
- Apply PandoraPFA on AHCAL 2018 beam test data and MC simulation
  - ➡ Evaluate simulated algorithm performance for standalone application & **provide feedback on beam data**
  - ➡ Compare performance on data & simulation
  - ➡ Study degree of confusion for different scenarios (particle energies, shower separation, leakage etc.)

Scenario 1 (Single Charged Hadron Event)



➡ Scenario is sensitive to double counted energy

Scenario 2 (Charged Hadron + Neutral Hadron Event)



➡ Scenario is sensitive to missing energy

Note: No magnetic field during beam test

Motivation and Goals of Studies II

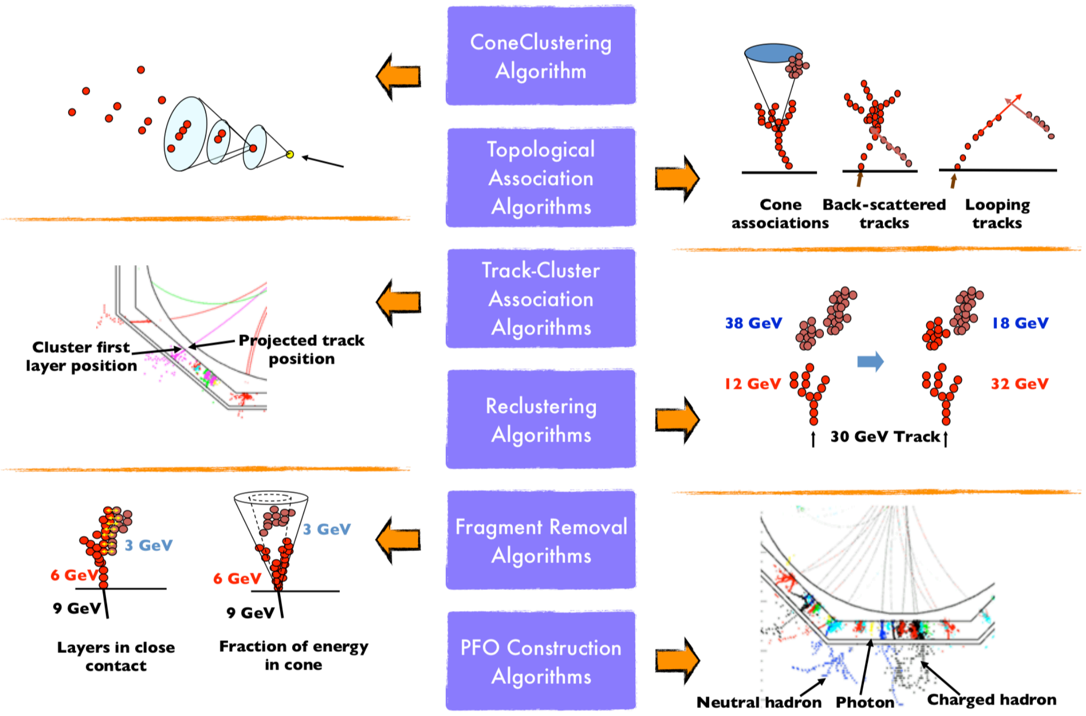
PandoraPFA on AHCAL 2018 Prototype Data

A comparable study was done for the AHCAL 2007 prototype (https://arxiv.org/abs/1105.3417)

Why do it again on AHCAL 2018 prototype data?

- **Significant developments of PandoraPFA until now**
  - ➡ Modular geometry drivers allow standalone application (instead of projection of data to ILD)
  - ➡ Relative easy plugin initialisation and implementation (leakage), etc.
- Latest AHCAL 2018 prototype:
  - ➡ Significant reduction of noise (SiPMs)
  - ➡ Very high and uniform granularity (22k channels)
  - ➡ Timing capabilities
- Single particle studies not done before

Illustration of Key Steps of PandoraPFA



J. S. Marshall: https://indico.in2p3.fr/event/7691/contributions/42712/attachments/34375/42344/3_john_marshall_PFA_marshall_24.04.13.pdf

Motivation and Goals of Studies II

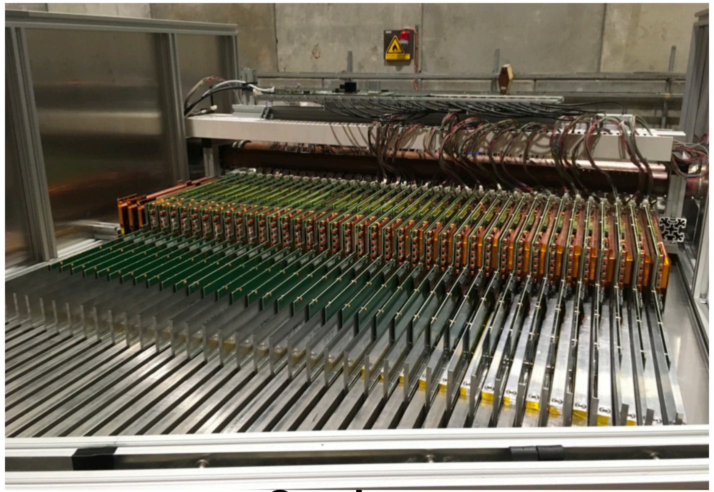
PandoraPFA on AHCAL 2018 Prototype Data

A comparable study was done for the AHCAL 2007 prototype (https://arxiv.org/abs/1105.3417)

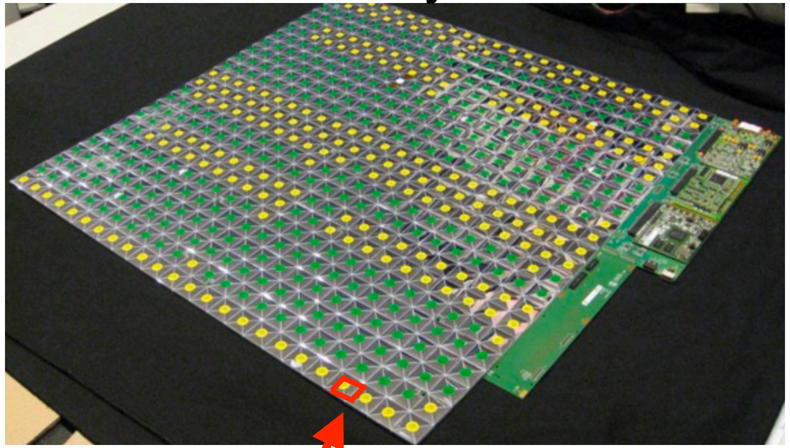
Why do it again on AHCAL 2018 prototype data?

- Significant developments of PandoraPFA until now
  - ➡ Modular geometry drivers allow standalone application (instead of projection of data to ILD)
  - ➡ Relative easy plugin initialisation and implementation (leakage), etc.
- **Latest AHCAL 2018 prototype:**
  - ➡ Significant reduction of noise (SiPMs)
  - ➡ Very high and uniform granularity (22k channels)
  - ➡ Timing capabilities
- Single particle studies not done before

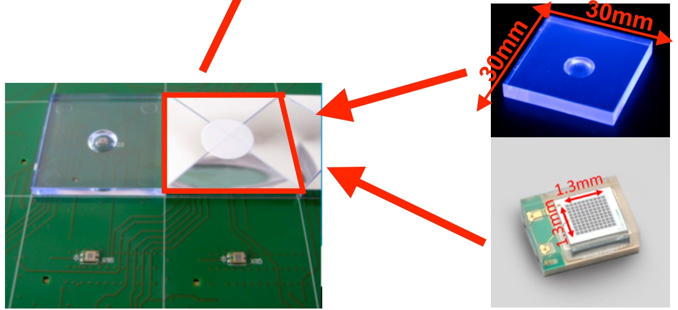
AHCAL 2018 Prototype: 38 layers within steel stack



One layer



One channel: Scintillating tile + SiPM



Overview

Sample Preparation & Analysis Strategy

Note: Preparation and selection tools finished and validated

Reminder: No neutral hadrons in beam tests therefore primary track removal on charged hadrons required to create pseudo-neutrals

AHCAL data & MC →

Unselected charged pion events

→ **Event preparation & selection**

Selected charged pion events

→ **Primary track removal & event overlay**

Selected events with overlaid charged and pseudo-neutral hadron

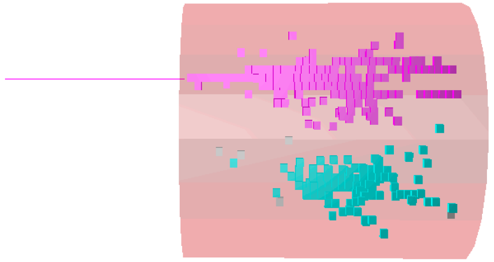
Selected charged pion events ↓ **PandoraPFA** → **Scenario 1: Single particle reconstruction studies**

Selected events with overlaid charged and pseudo-neutral hadron ↓ **PandoraPFA** → **Scenario 2: Two-particle separation studies**

Analysis inspired by first CALICE PFA Paper: https://arxiv.org/abs/1105.3417
&
Remi Ete's ArborPFA Studies on SDHCAL Data
CAN: http://cds.cern.ch/record/2669487/files/fulltext.pdf

Sample and Selection Overview

PandoraPFA Two Particle Reconstruction



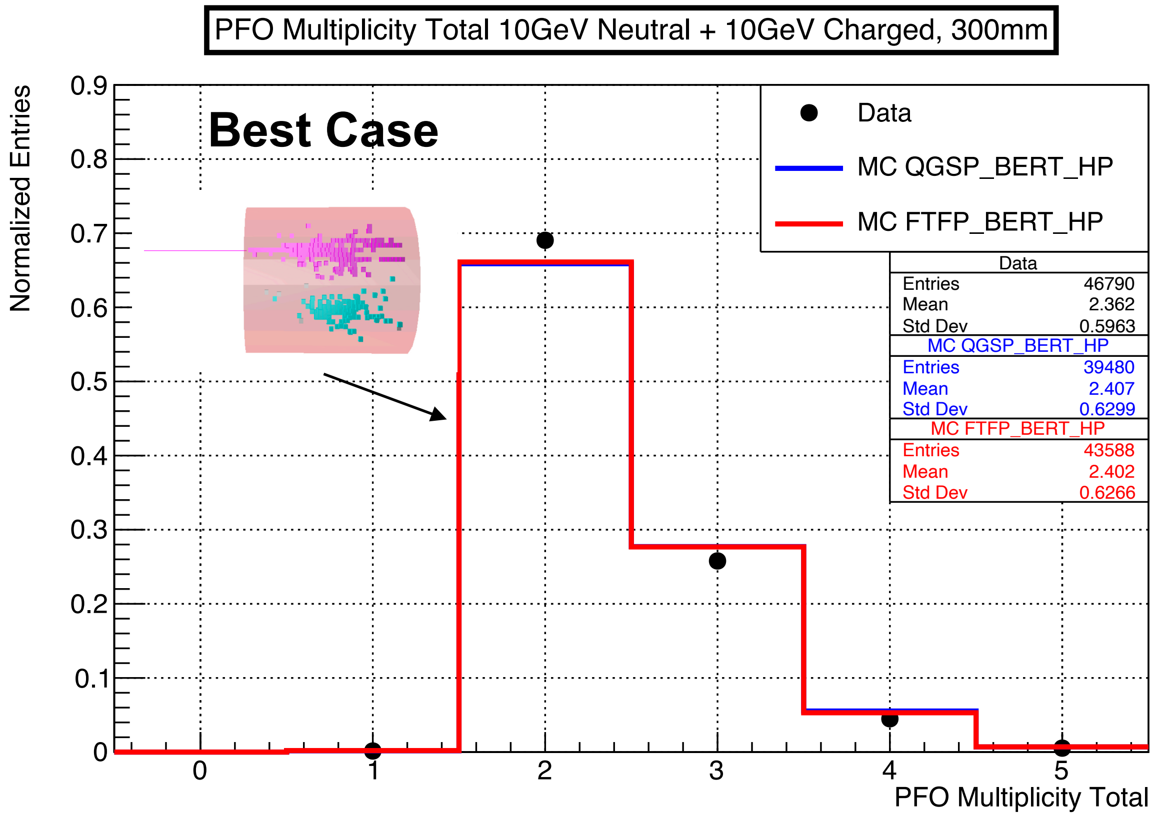
- Scenario: 10 GeV (pseudo-) neutral hadrons overlaid with 10 GeV or 30 GeV charged hadrons
  - ➡ Shower distance bins: 0, 50, 100, 150, 200, 250, 300 mm, width: ±25 mm

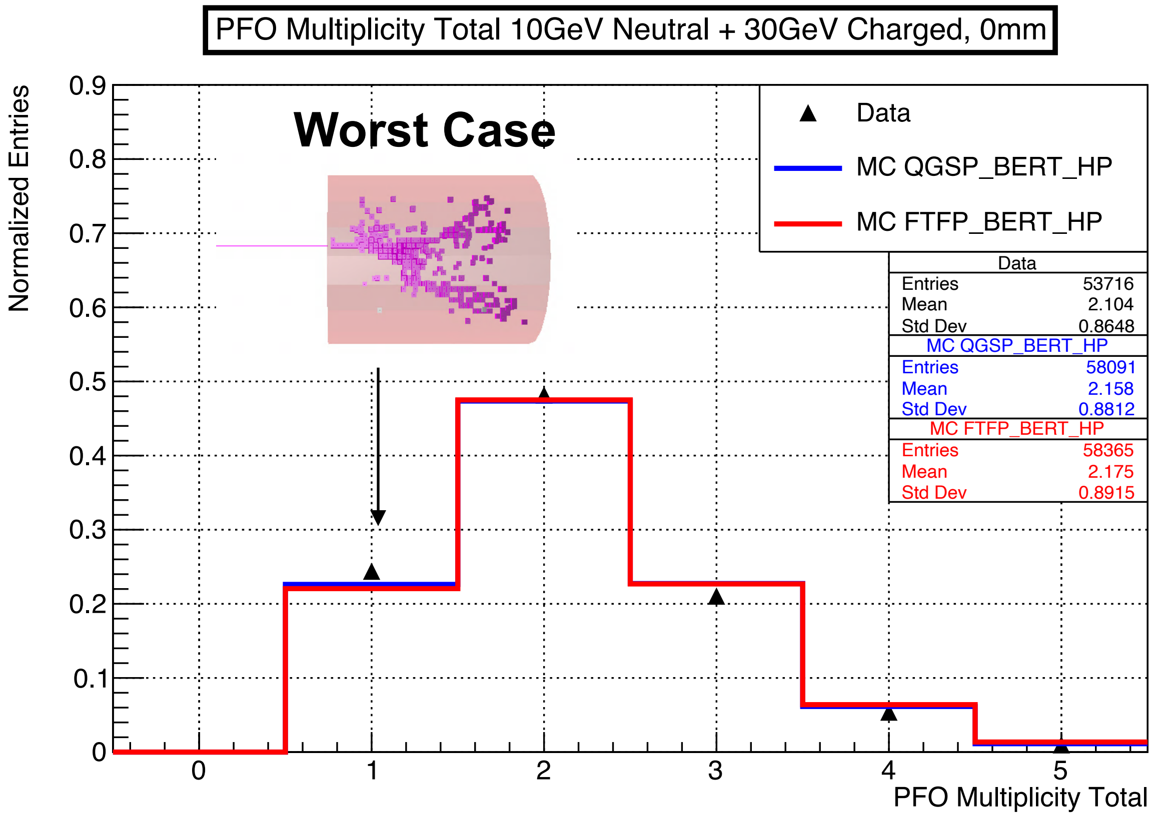
- Data: June Beam Test 2018 @ SPS CERN (~700k events), tracks: Delay wire chamber (by Linghui)
- MC: GEANT4 v.10.03, QGSP_BERT_HP & FTFP_BERT_HP (~1m events each), tracks: MC endpoint extrapo.
- Track momentum: 10 GeV or 30 GeV sharp
- Applied PID (based on BDT by Vladimir) for hadrons to remove beam contamination

- Event selection:
  - ➡ Punch through rejection & no cut on shower start layer (allow long. separation)
  - ➡ Charged hadron: track-hit match layer 1||2||3, track-to-detector-gap rejection
  - ➡ Requiring at least 10% of charged hadron energy associated to track (initial track association, no ECAL)

Total PFO Multiplicity - Examples

How many Reconstructed Particle Flow Objects?



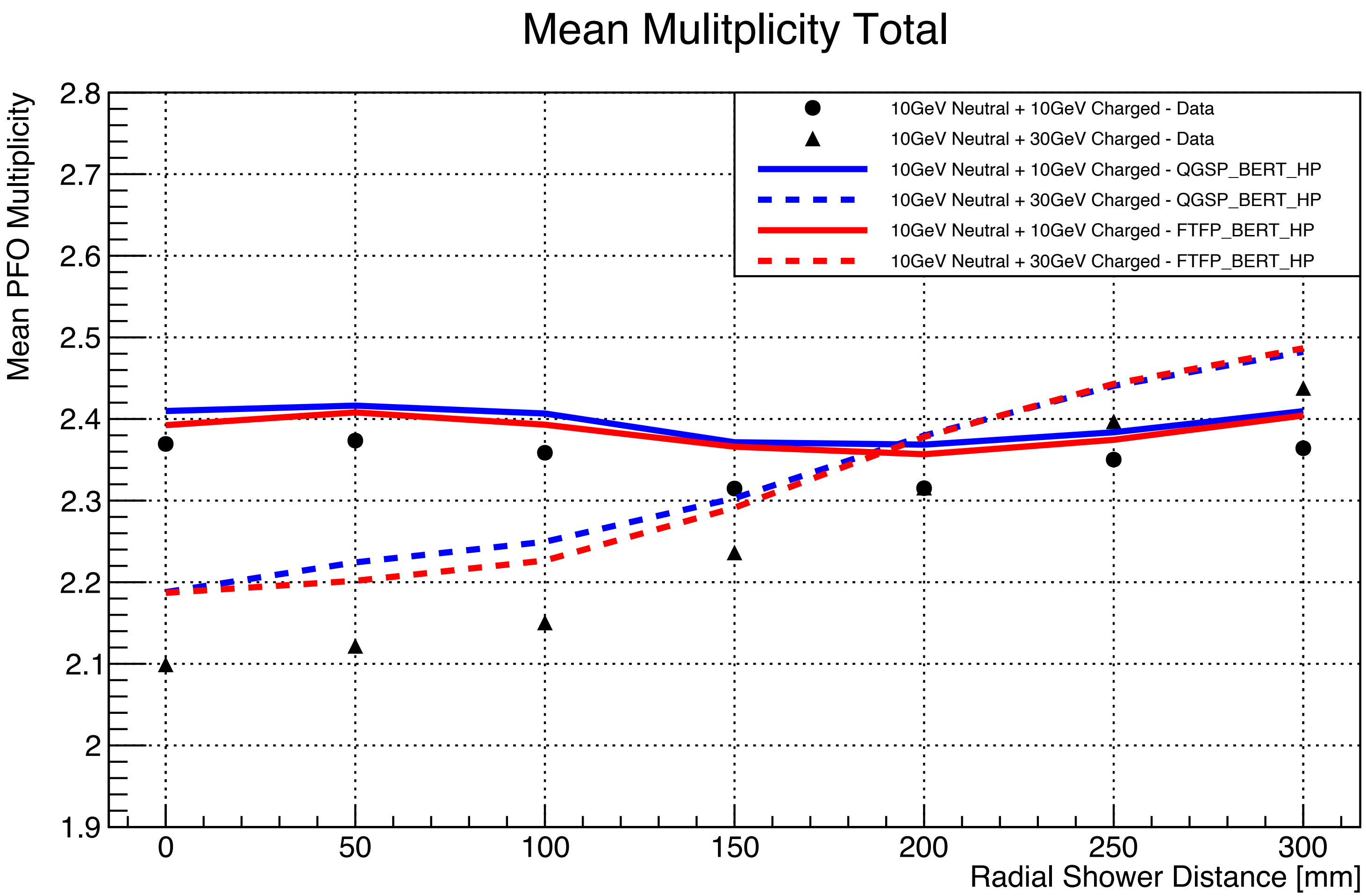


- Excellent data to MC agreement and general trends for PFO multiplicities as expected:
  - ➡ Best case: Clear separation - Almost no events with neutral hadron absorbed, always 2 or more PFO's
  - ➡ Worst case: (Only) ~20% of events with neutral hadron is completely absorbed into charged hadron

Mean PFO Multiplicity

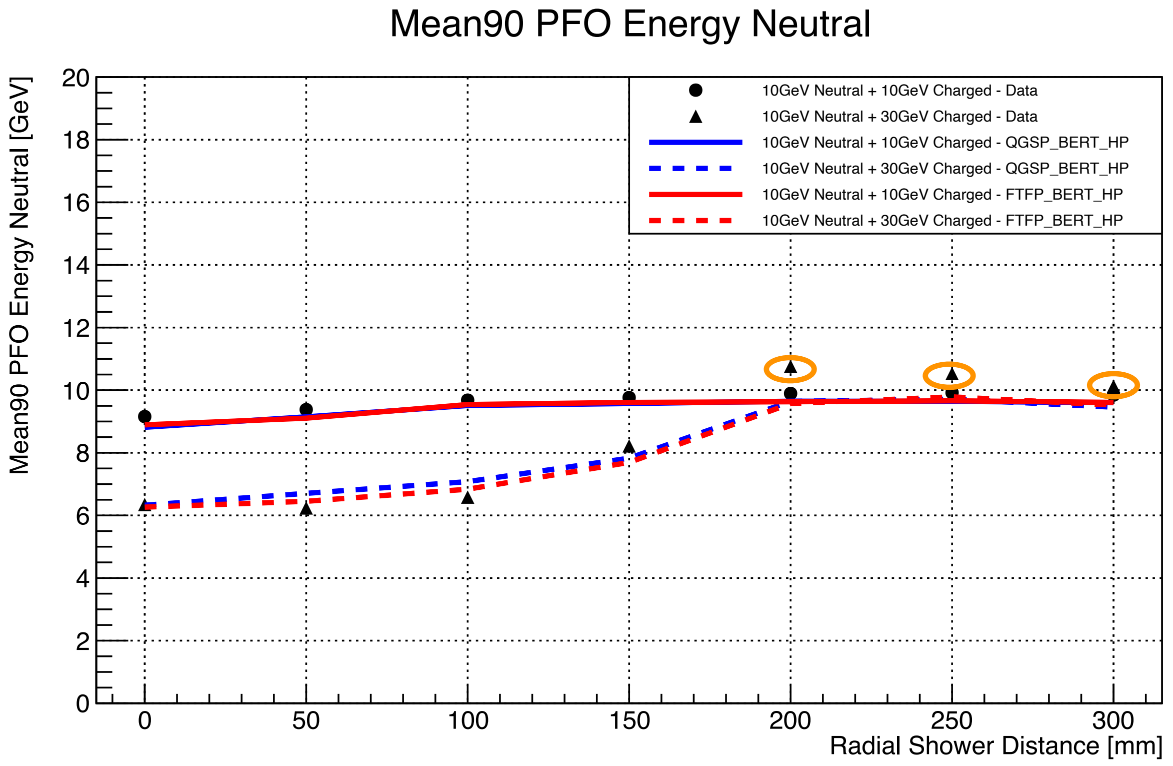
How many Reconstructed Particle Flow Objects on Average?

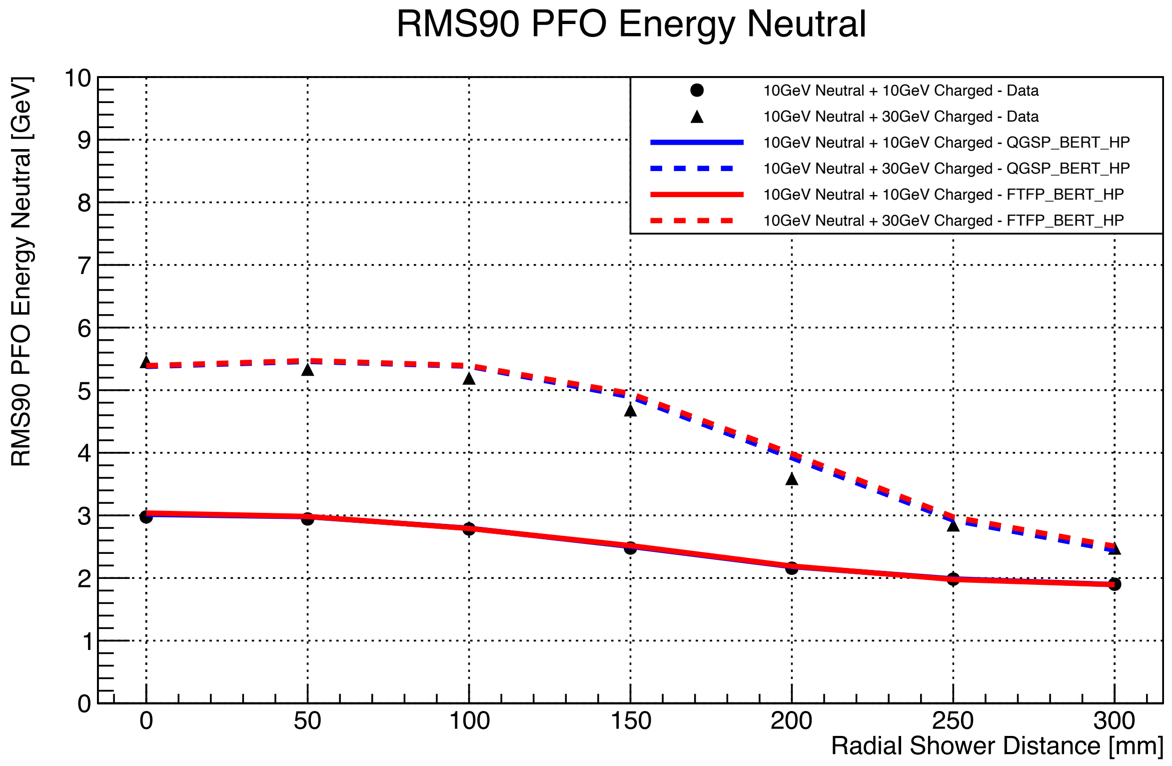
- Good data to MC agreement within 5%
  - ➡ But: Mean multiplicity is systematically lower for data as for single particle reconstruction
- 10GeV + 10GeV: Almost constant over increasing shower distance
- 10GeV + 30GeV: Growing with shower distance
  - ➡ With growing distance less absorption events (1 PFO)
  - ➡ Higher fragmentation within clustering because of higher energy



PFO Energy Neutral Mean90 & RMS90

How well is the 10 GeV Neutral Hadron Reconstructed On Average by PandoraPFA?



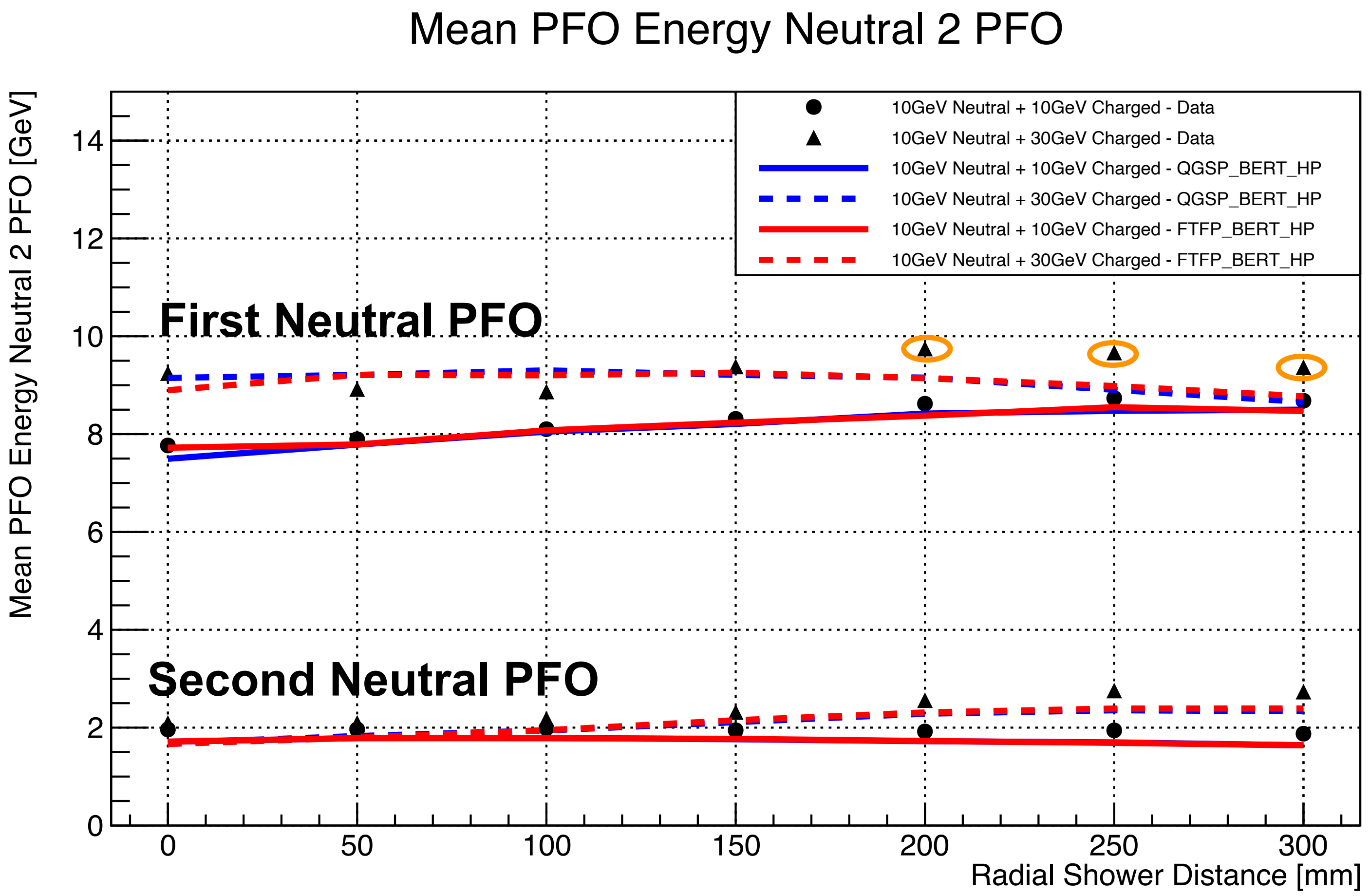


- Vicinity of 10 GeV charged hadron: Sufficient recovery of neutral hadron even at lowest shower distances
- Vicinity of 30 GeV charged hadron: More missing energy at low distances due to absorption, overestimation at far distances for data (currently under investigation)

➡ In general: The larger the distance between showers the preciser the reconstruction, good data to MC agreement

Mean Neutral PFO Energy - Events with 2 Neutral PFOs

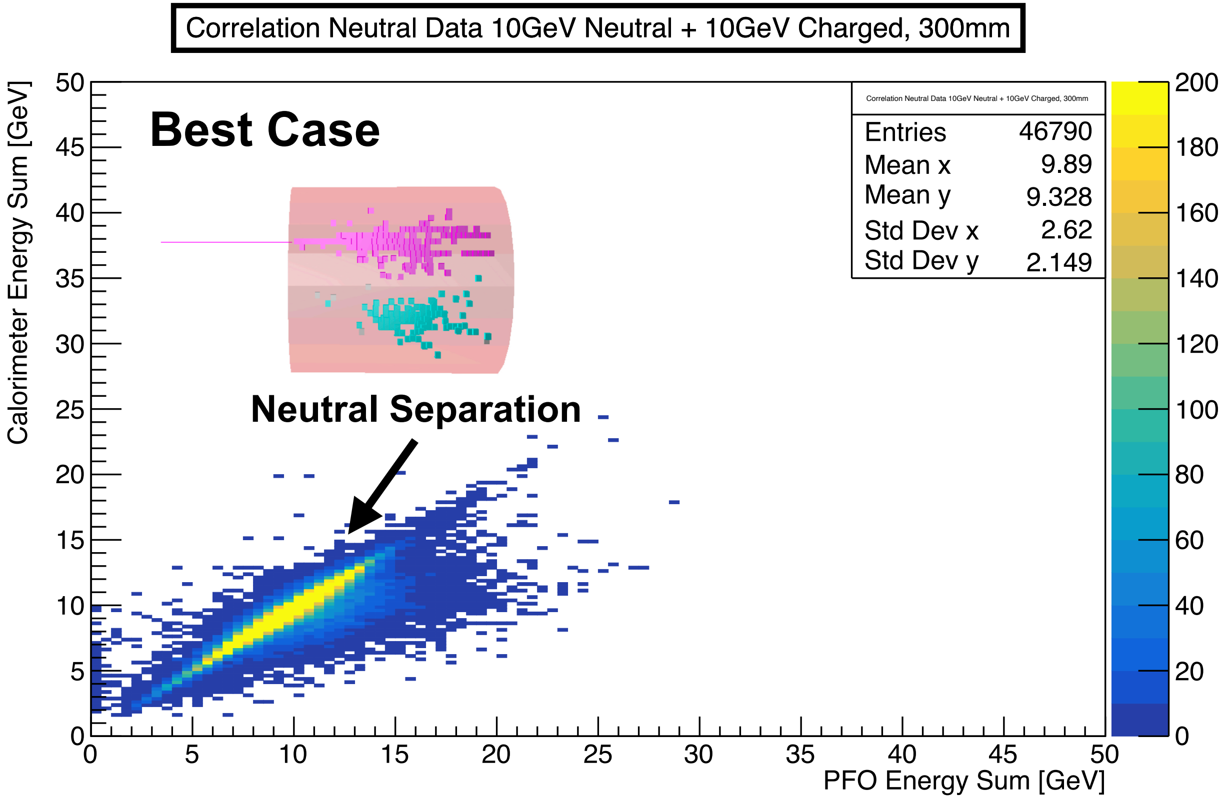
How is the Energy Distributed Among Multiple Neutral PFOs?

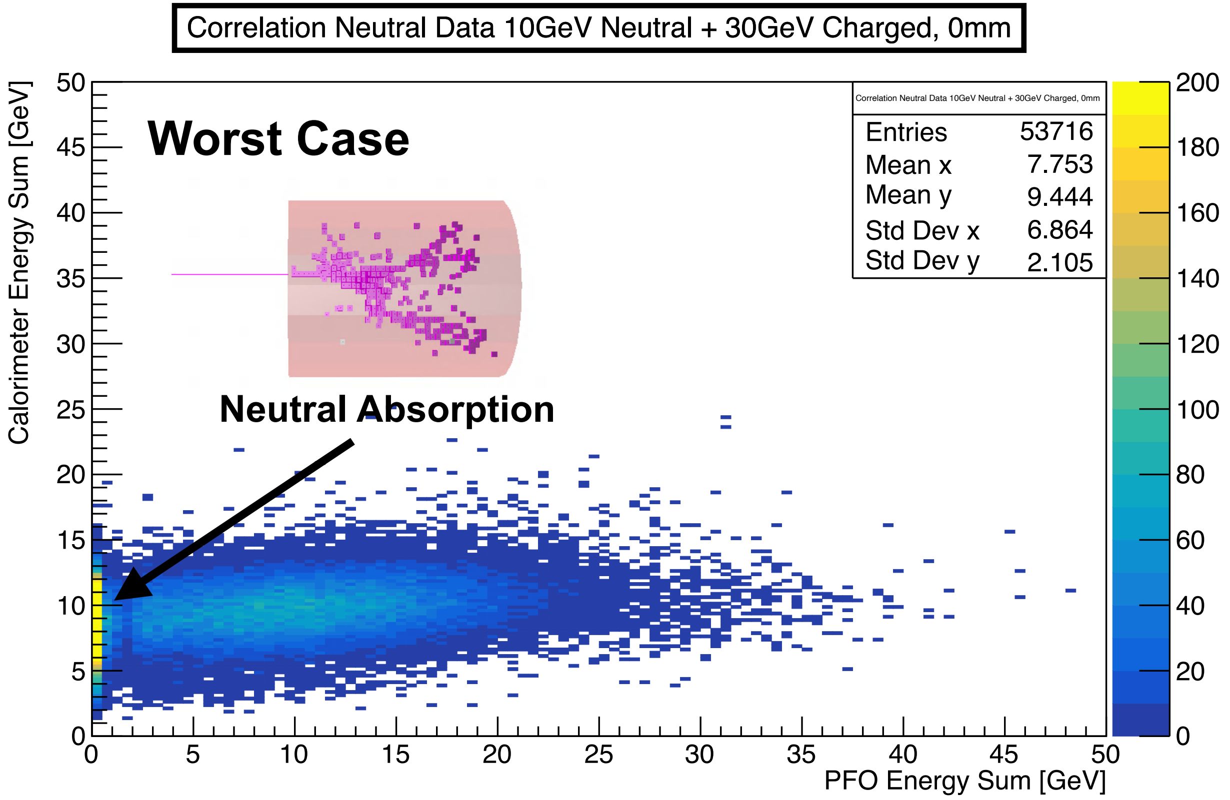


- If 2 neutral PFO’s reconstructed: One PFO carrying most of the identified neutral hadron energy, second one only ~20%
  - ➡ Higher energy for 10 & 30 GeV scenario due to selection bias to double counted energy events
  - ➡ Overestimation of data for large distances for 10 & 30 GeV scenario
- Good data to MC agreement within 5%

Neutral Hadron Energy: PFO vs. Calorimeter Energy Correlation

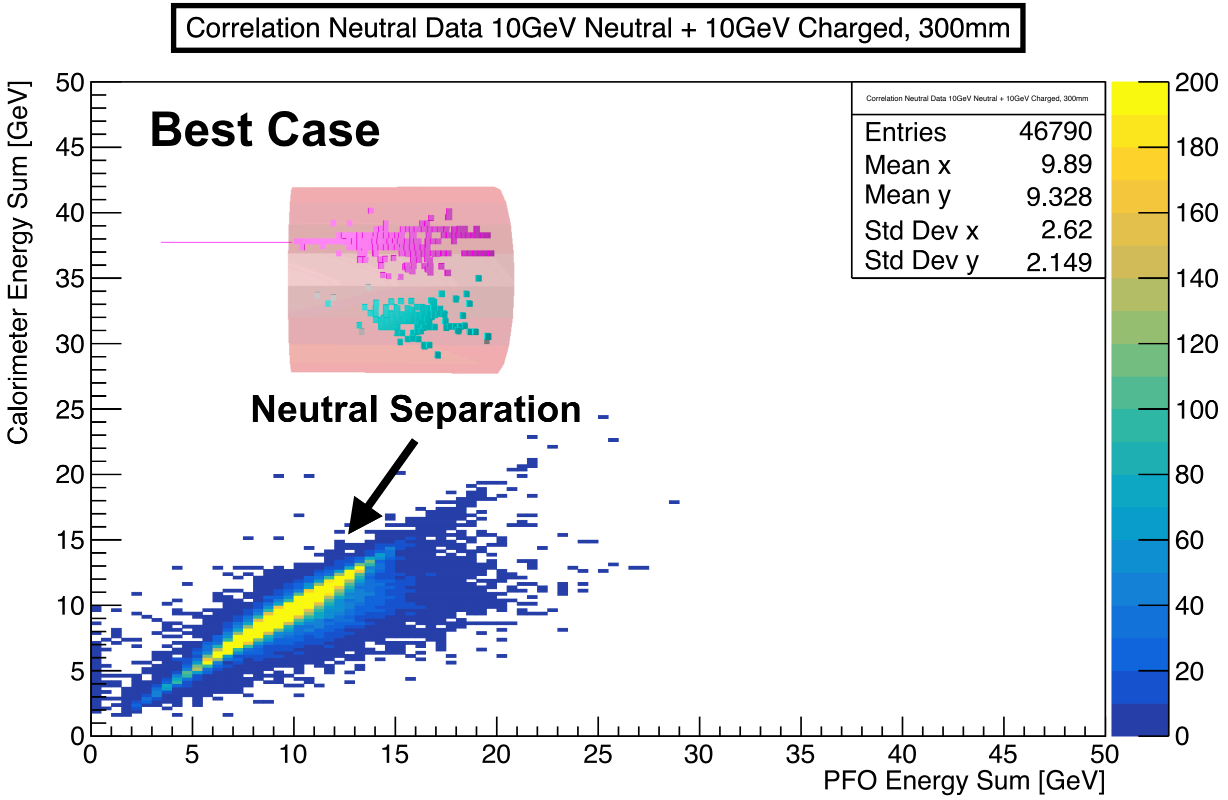
Visualisation Examples for Data

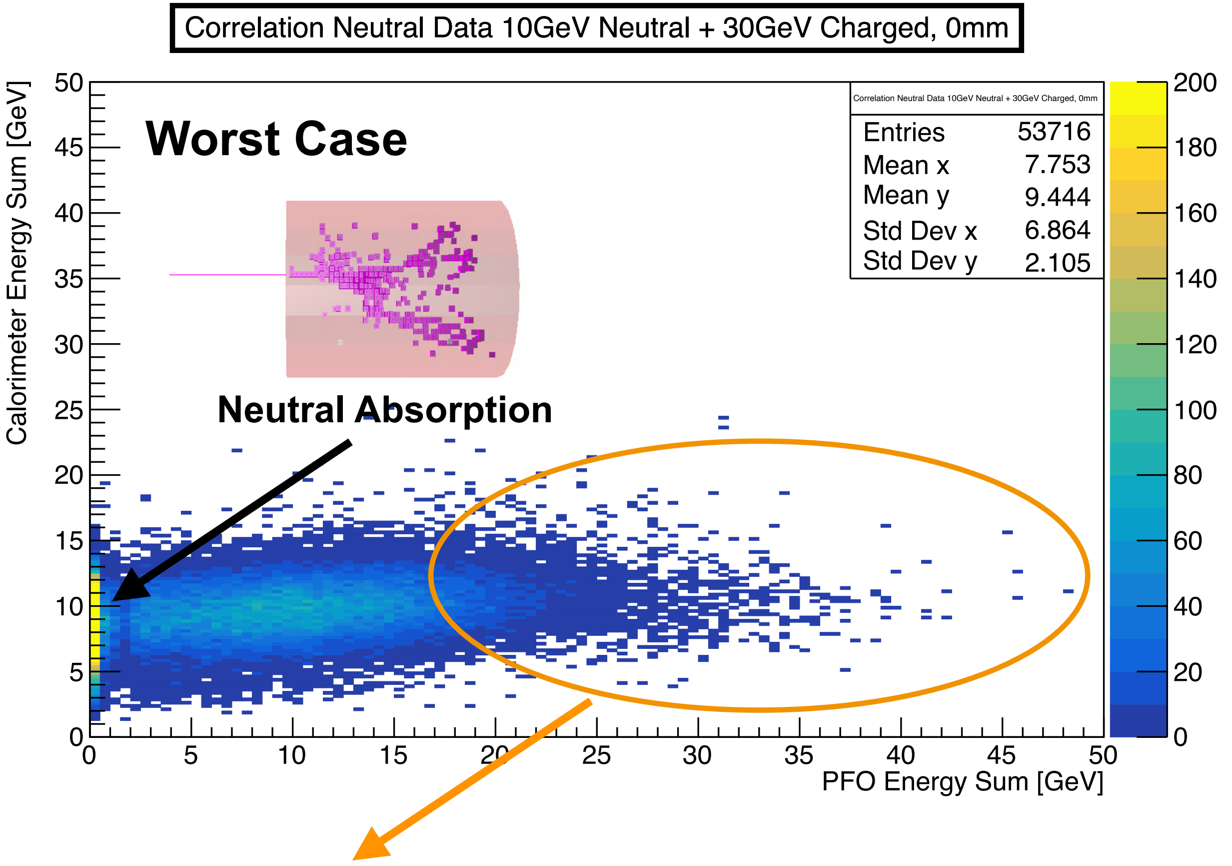




Neutral Hadron Energy: PFO vs. Calorimeter Energy Correlation

Visualisation Examples for Data

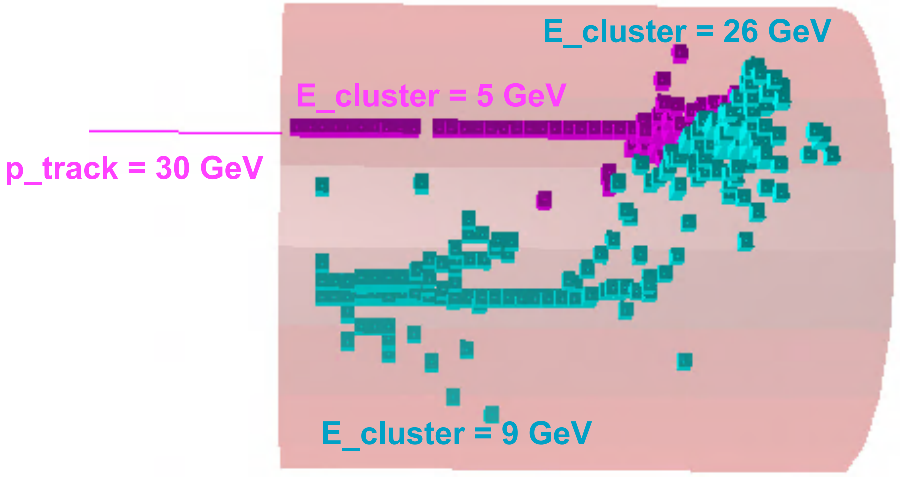


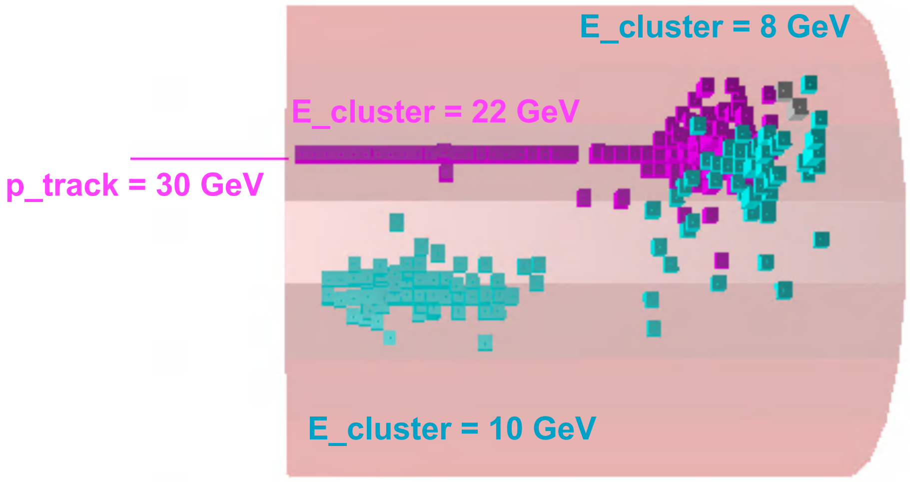


- Overestimation of neutral PFO energy?
  - ➡ Charged hadron hits are absorbed into neutral?

Intermezzo: Events with High Neutral PFO Energy

Illustration of Problematic Events & Solution Status

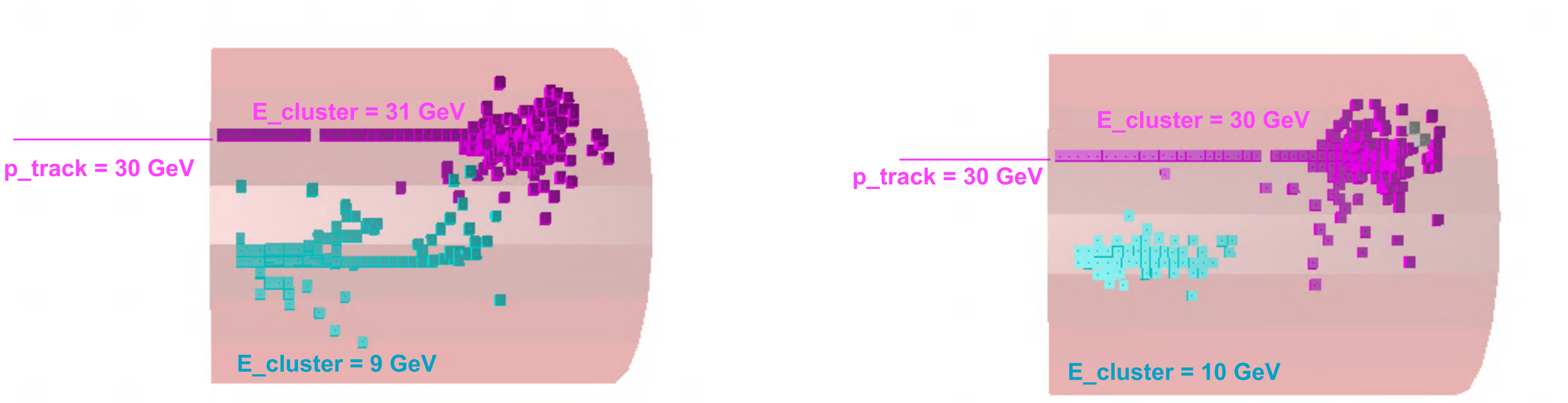




- For around 1% of events bad reconstruction: Re-clustering algorithms are not triggered!
  - ➡ Mostly for 10 GeV neutral + 30 GeV charged hadron scenarios if the charged hadron is transversally closer to calorimeter edge

Intermezzo: Events with High Neutral PFO Energy

Illustration of Problematic Events & Solution Status

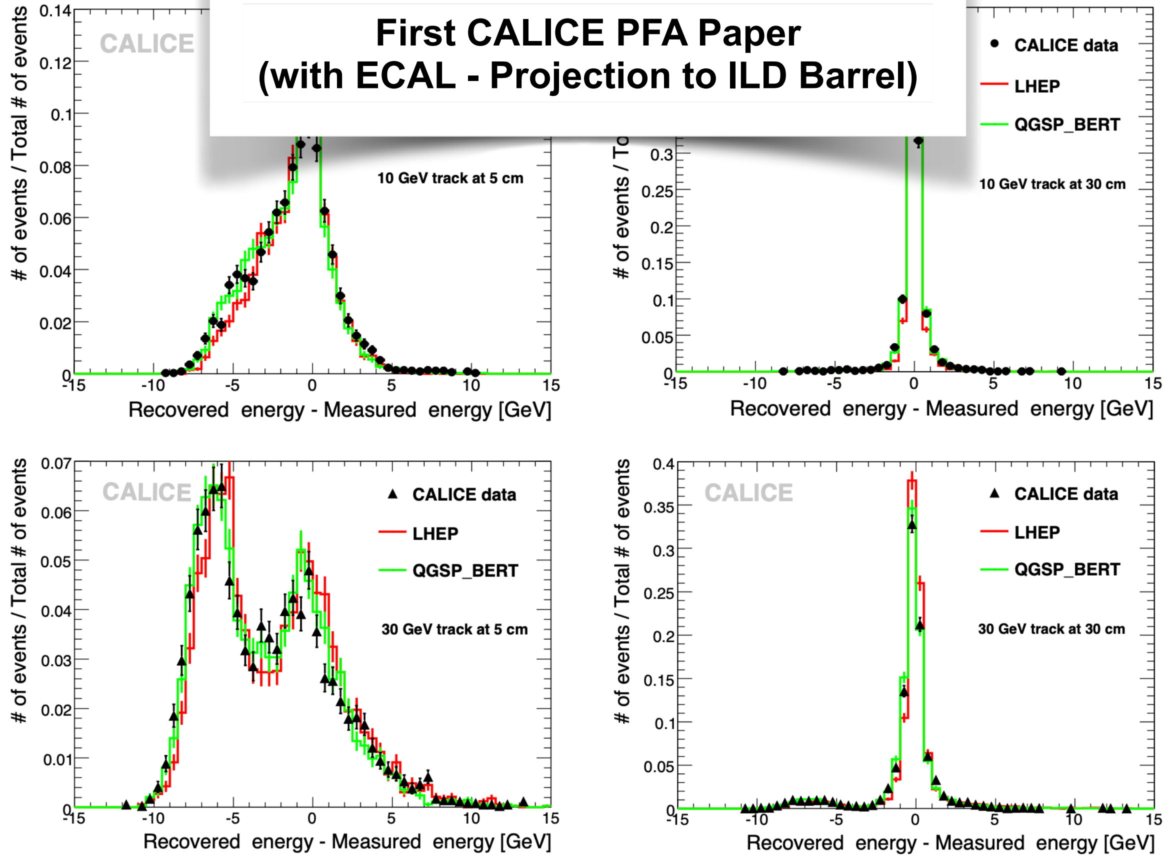


- For around 1% of events bad reconstruction: Re-clustering algorithms are not triggered!
  - ➡ Mostly for 10 GeV neutral + 30 GeV charged hadron scenarios if the charged hadron is transversally closer to calorimeter edge
- Problem tracked down with great support of J. Marshall
  - ➡ Currently „debugging“, optimistic to fully enable the magic of PandoraPFA for this events and potential slight improvement for already „good" events

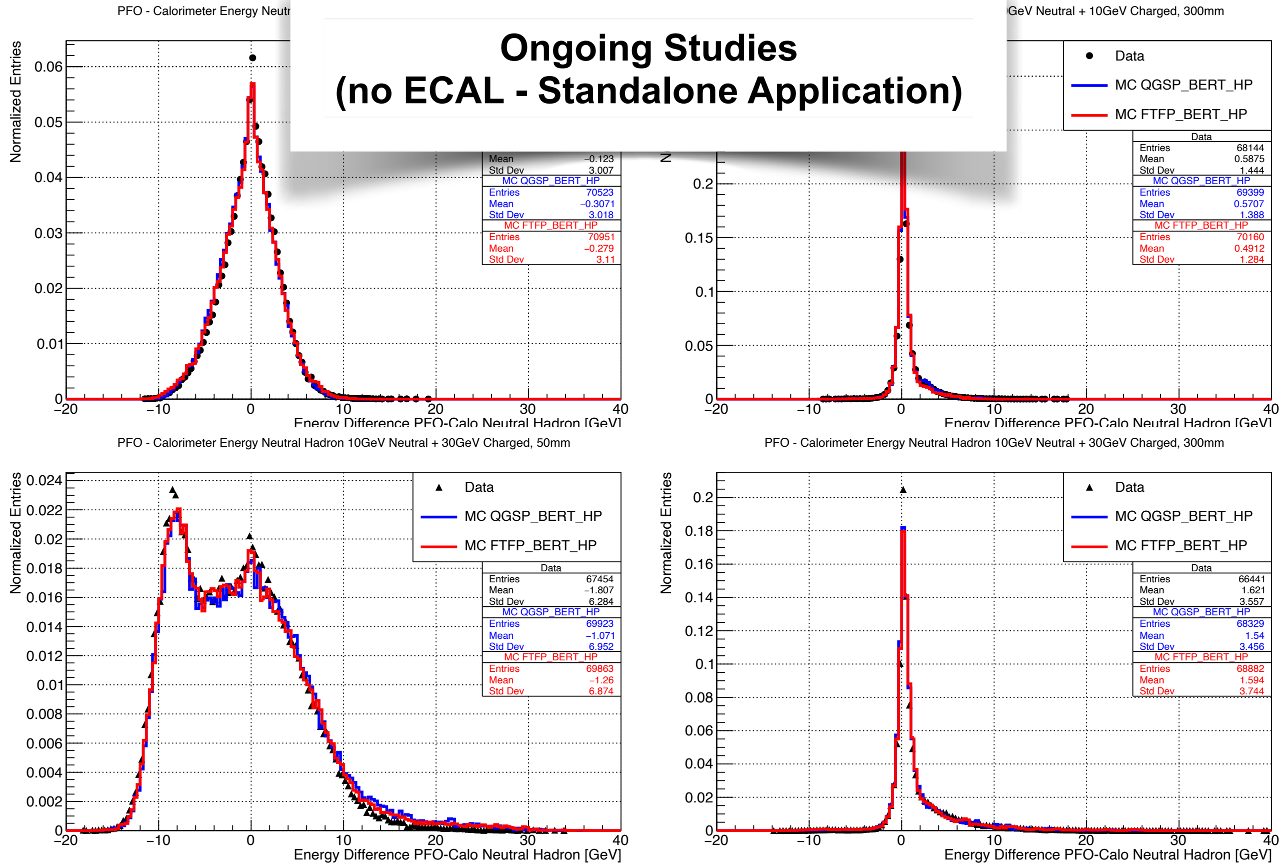
PFO Energy - Calorimeter Energy Neutral Hadron

Is the Energy of the Neutral Hadron Reconstructed Correctly by PandoraPFA?

First CALICE PFA Paper (with ECAL - Projection to ILD Barrel)



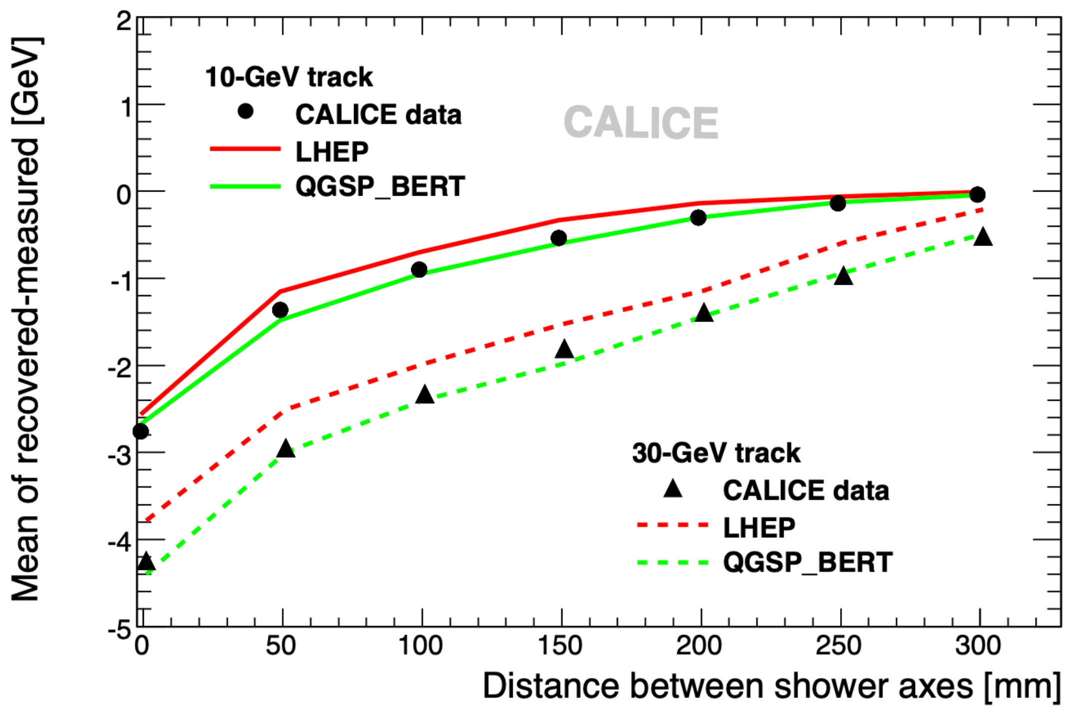
Ongoing Studies (no ECAL - Standalone Application)

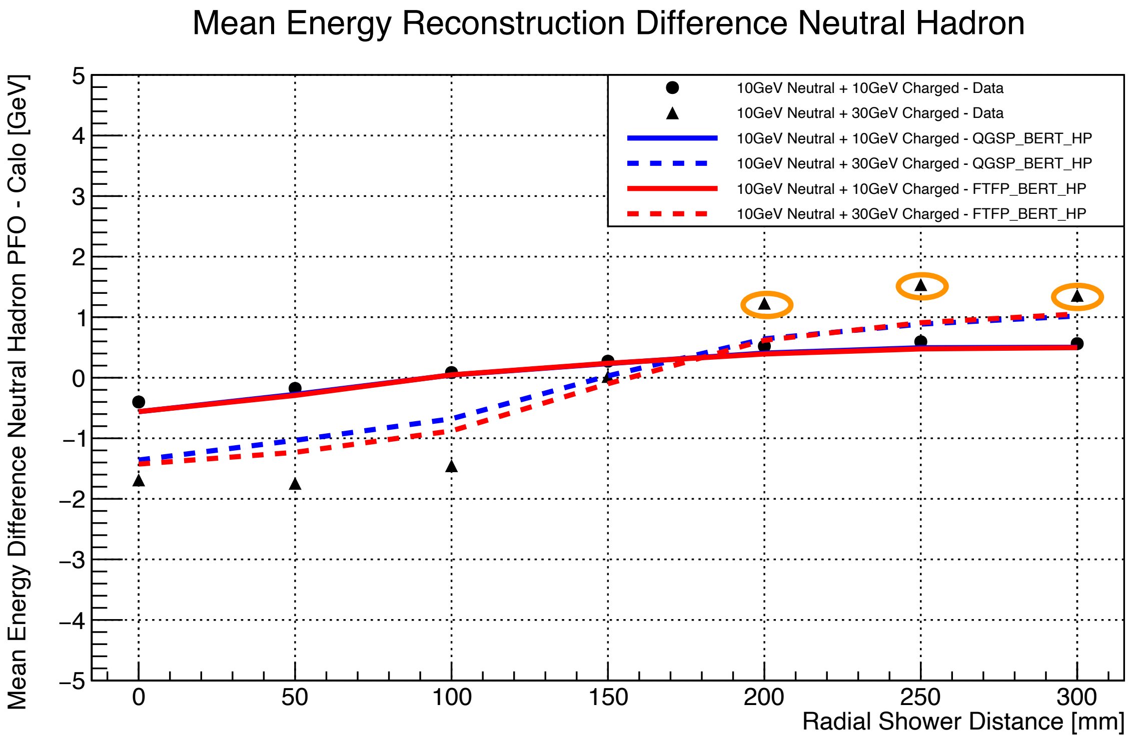


- Same features as for previous studies, but larger width and more pronounced tail to the right (high energy neutral PFO events)
  - ➡ Checked quantitatively for all distances

PFO Energy - Calorimeter Energy Neutral Hadron: Mean

First CALICE PFA Paper vs. Ongoing Studies

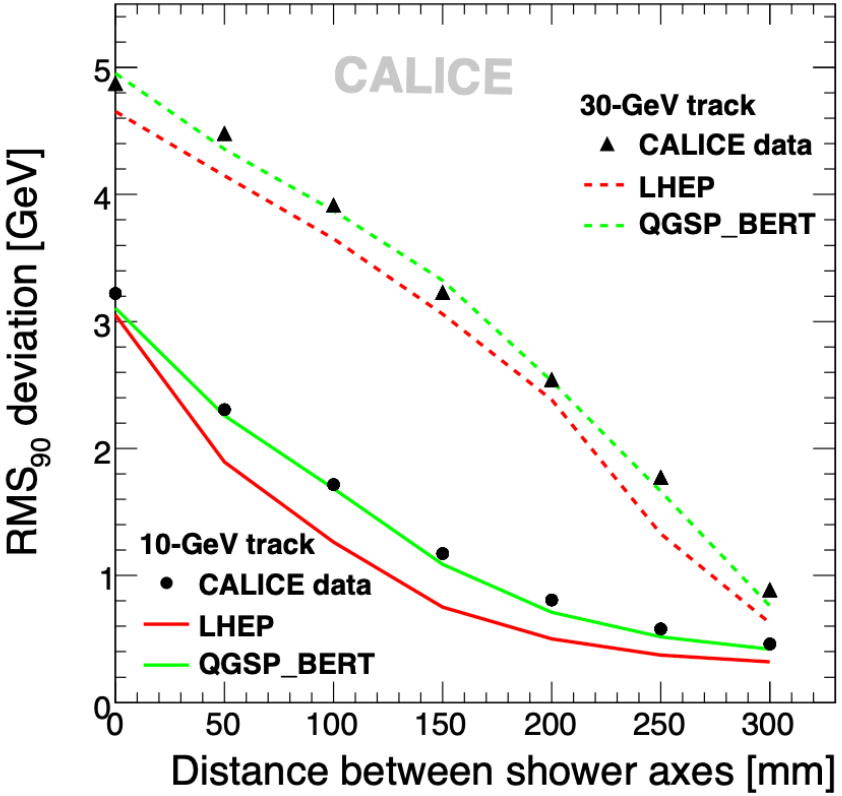


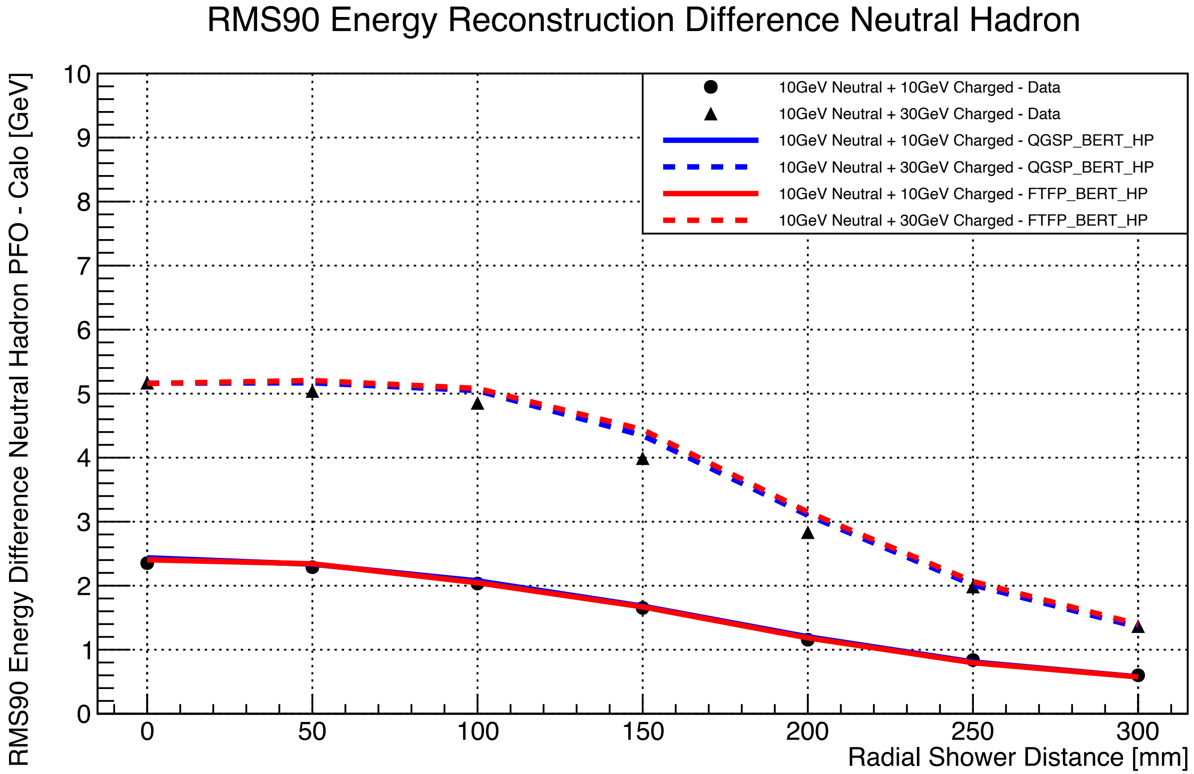


- Same rising trend for growing distance to 10 GeV and 30 GeV charged hadrons (but smaller offset, more flat)
- In general good data to MC agreement (data at large distances: more problematic events)
- In general good reconstruction performance
  - ➡ At low distances missing energy due to absorption, at high distances slight overestimation

PFO Energy - Calorimeter Energy Neutral Hadron: RMS90

First CALICE PFA Paper vs. Ongoing Studies

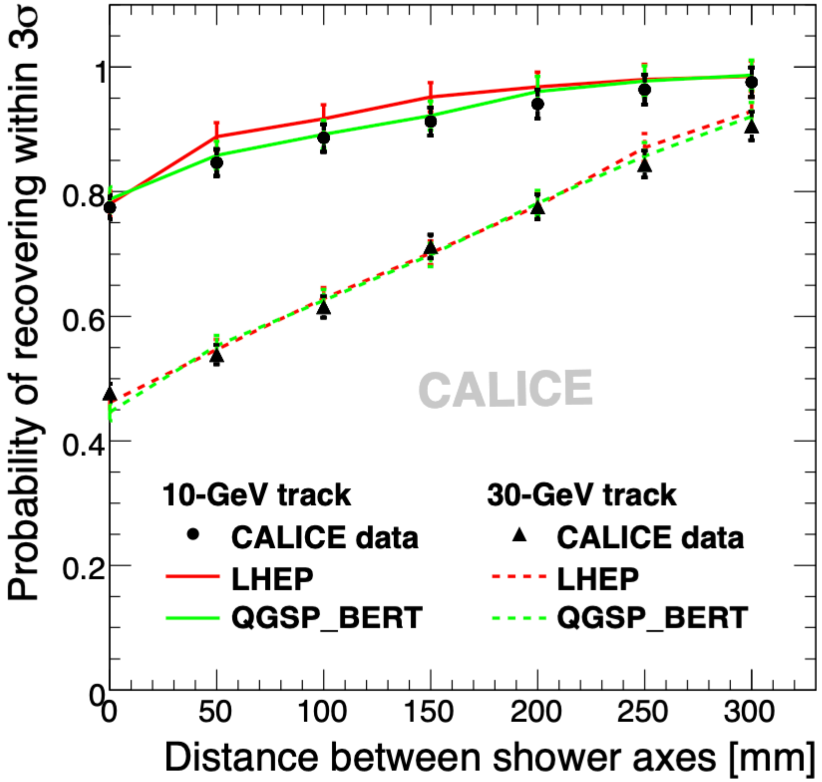


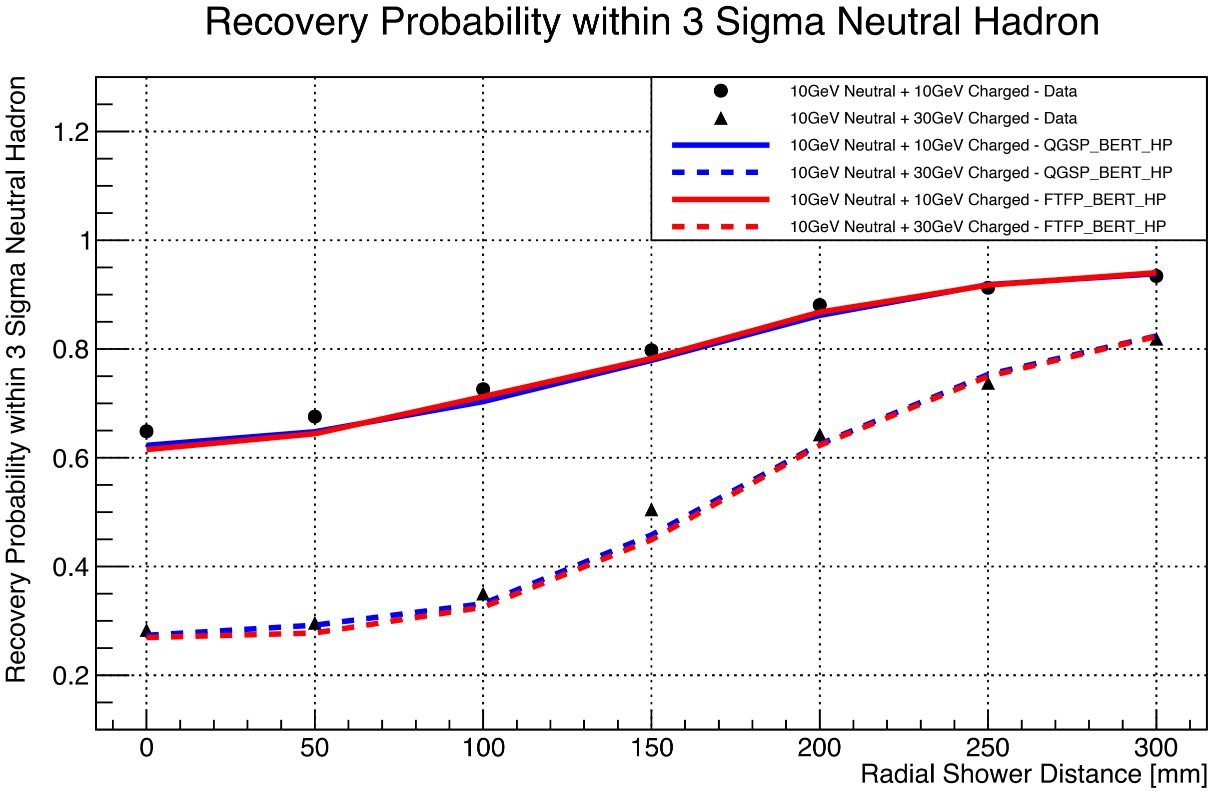


- Excellent data to MC agreement
- Slower falling trend for growing distance to 10GeV and 30GeV charged hadrons
    - ➡ Suspicion: Low distances very tricky without ECAL hits before AHCAL

Neutral Hadron Recovery Probability 3 Sigma

First CALICE PFA Paper vs. Latest Studies

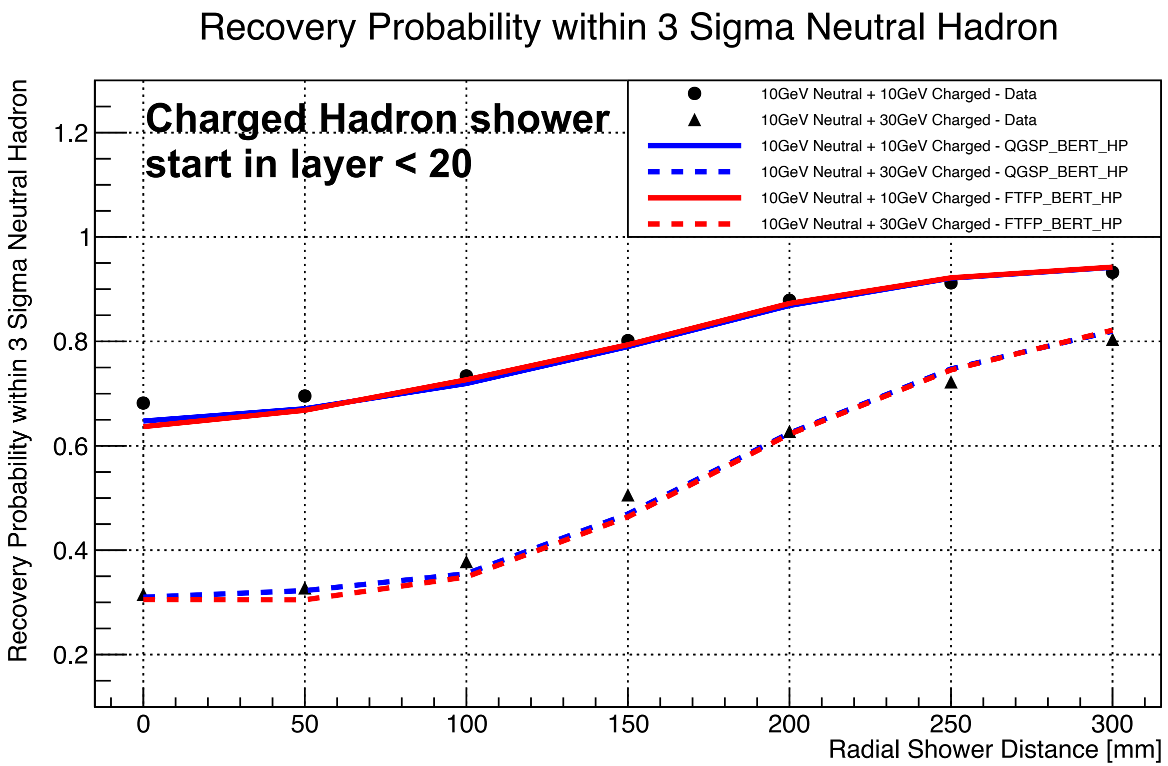


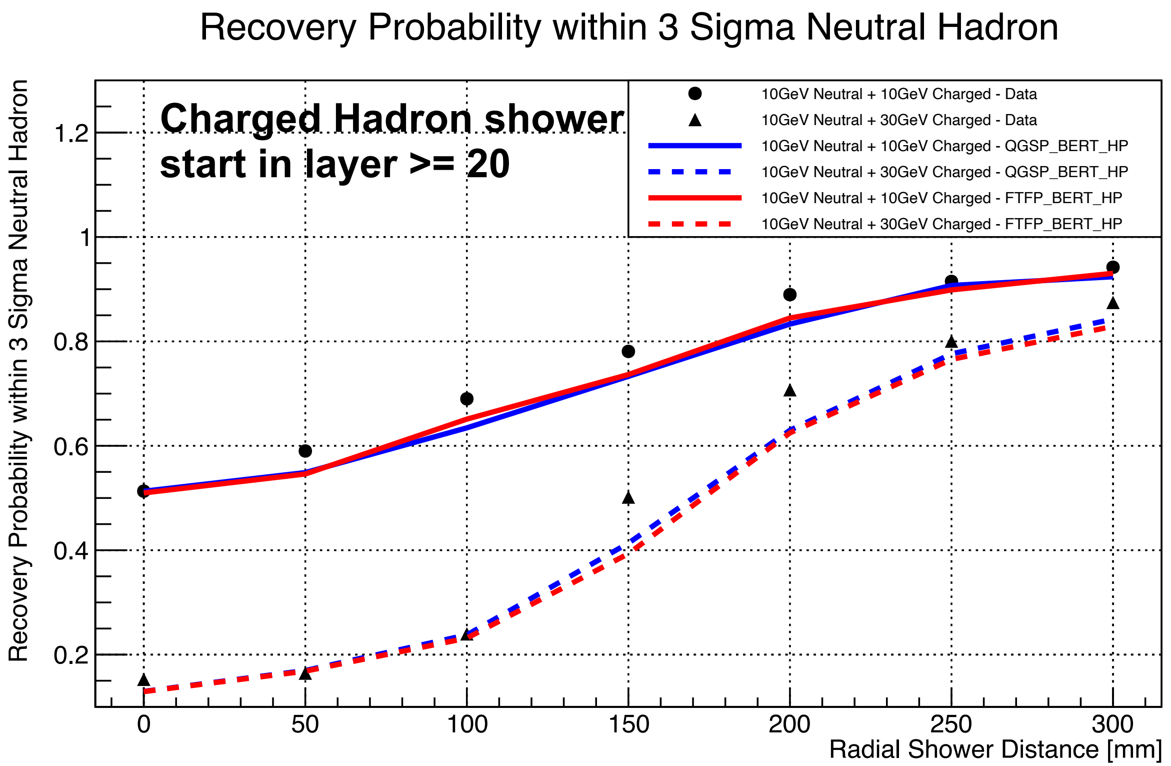


- Definition: Fraction of events for which PandoraPFA recovered neutral hadron energy within **3 sigma (sigma = width of neutral hadron energy sum of calorimeter measurement)**
- Same rising trend for larger separation, but slower growing especially for vicinity of 30 GeV charged hadron
  - ➡ Excellent data to MC agreement, slightly worse performance for current studies (no ECAL?)

Neutral Hadron Recovery Probability 3 Sigma

Influence of Leakage/Longitudinal Separation





- For low distances: Leakage of charged hadron energy due to late shower start dominant
  - ➡ Energy from neutral hadron is absorbed causing significantly lower neutral hadron recovery probability
- For large distances: Additional longitudinal separation of showers pays off
  - ➡ Neutral to far away to suffer from leaking charged hadron, slightly improved neutral hadron recovery probability

Summary & Outlook

- Established well working PandoraPFA environment for reconstruction of AHCAL 2018 standalone events
- Results of PandoraPFA reconstruction for different two particle event scenarios:
  - ➡ In general good algorithm performance in terms of two particle separation
  - ➡ Remarkable agreement between data and MC
  - ➡ Expected trends for confusion observed in investigated scenarios (charged energy, transversal separation)
    - ➡ First look into dependency on longitudinal separation/leakage

Further Plans:

- „Debug" remaining fraction (~1%) of problematic events with feedback of PFA experts (J. Marshall)
- Further investigation of longitudinal separation/leakage, neutral reconstruction efficiency/purity on hit level
- Detailed confusion studies: Comparison to ILD jets (PFA parameter checks), different granularities, different energy thresholds

Summary & Outlook

- Established well working PandoraPFA environment for reconstruction of AHCAL 2018 standalone events
- Results of PandoraPFA reconstruction for different two particle event scenarios:
  - ➡ In general good algorithm performance in terms of two particle separation
  - ➡ Remarkable agreement between data and MC
  - ➡ Expected trends for confusion observed in investigated scenarios (charged energy, transversal separation)
    - ➡ First look into dependency on longitudinal separation/leakage

Thank you!

(And special thanks to J. Marshall for his continuous help to open Pandora's box!)

Further Plans:

- „Debug" remaining fraction (~1%) of problematic events with feedback of PFA experts (J. Marshall)
- Further investigation of longitudinal separation/leakage, neutral reconstruction efficiency/purity on hit level
- Detailed confusion studies: Comparison to ILD jets (PFA parameter checks), different granularities, different energy thresholds

Backup

General

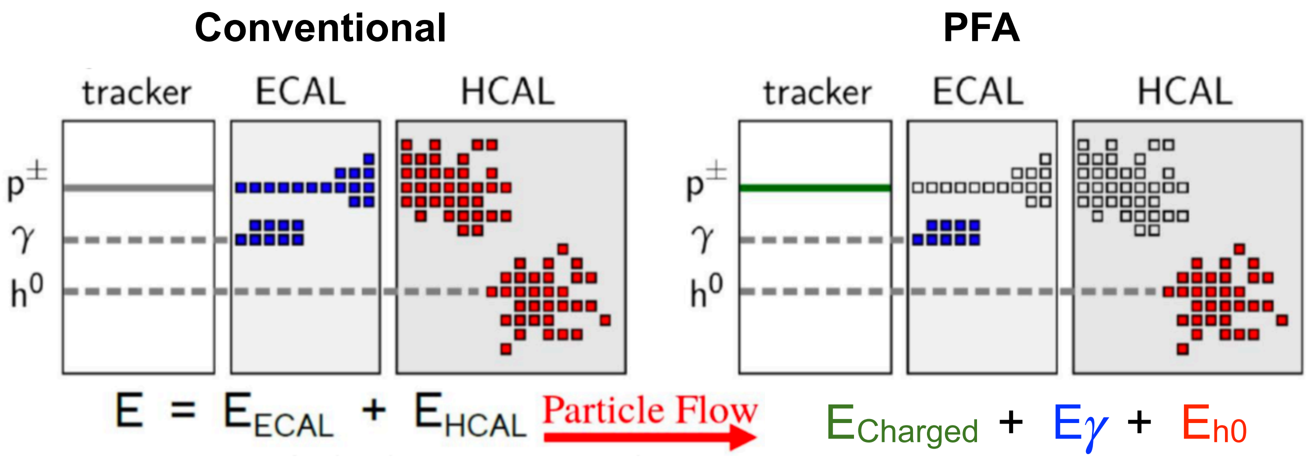
Particle Flow Approach

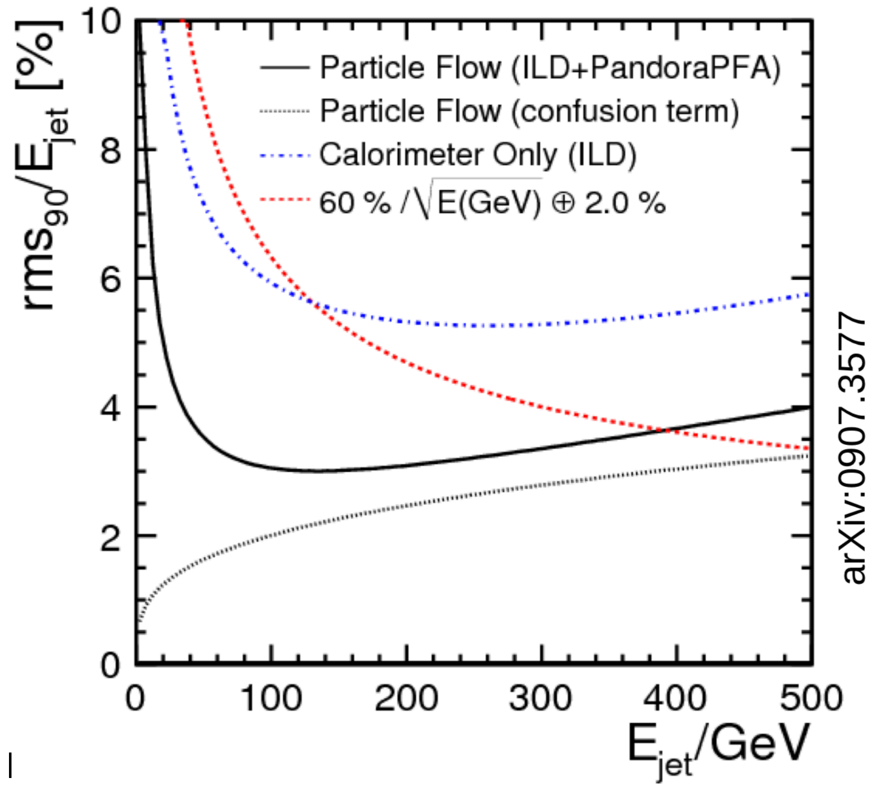
Reaching High Precision

- Goal at the ILC: Jet energy resolution of 3-4% for jet energies between 40-500 GeV
- Typical jet composition of 72% hadrons measured with poor hadronic energy resolution ~60%/√E

➡ PFA: Measure energy/momentum of each particle with detector providing best resolution

- ➡ 62% charged particles ➙ tracker
- ➡ 27% photons ➙ ECAL
- ➡ 10% neutral hadrons ➙ ECAL + HCAL

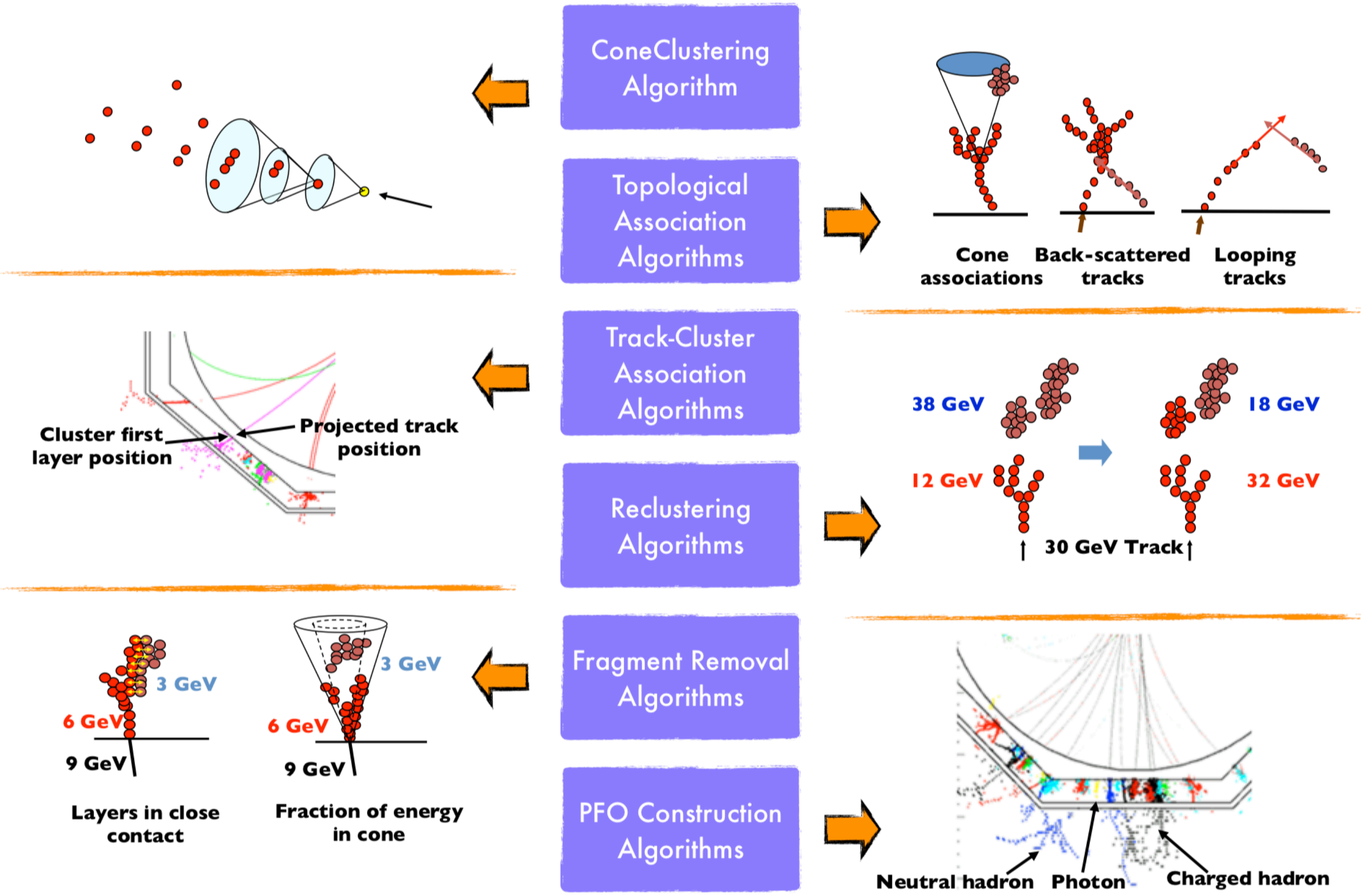




The Pandora Particle Flow Algorithm (PandoraPFA)

A Multi-Algorithm Pattern Recognition Tool

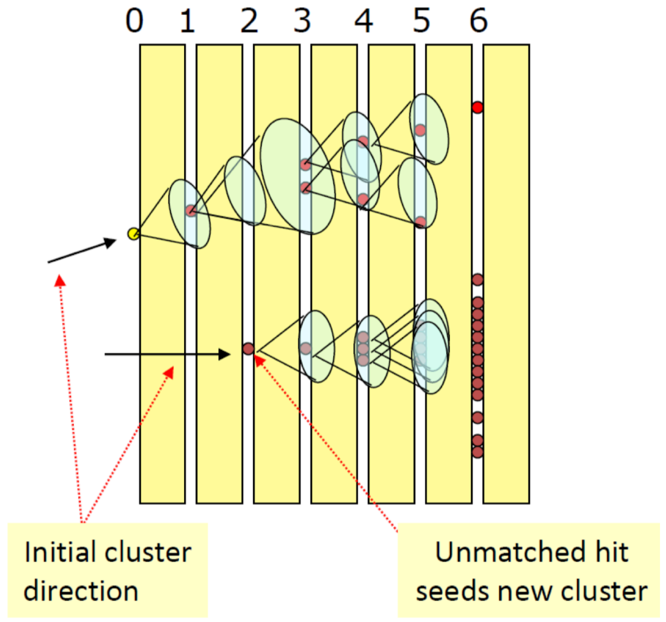
Illustration of Key Steps of PandoraPFA



J. S. Marshall: https://indico.in2p3.fr/event/7691/contributions/42712/attachments/34375/42344/3_john_marshall_PFA_marshall_24.04.13.pdf

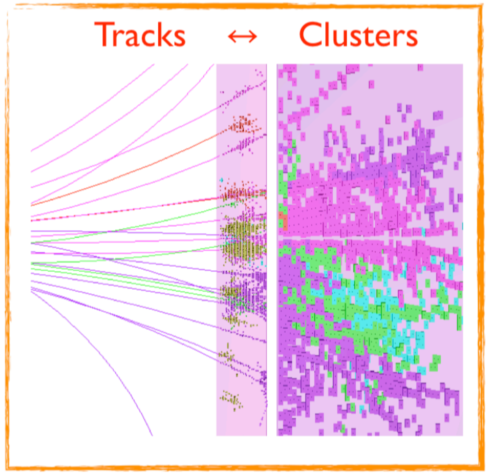
- PandoraPFA: Complex multi-algorithm chain using pattern recognition for event reconstruction
  - ➡ Performs calorimeter hit clustering, topological associations, …
  - ➡ Highly recursive: Find most accurate reconstruction scenario
  - ➡ Overall goal: Distinguish energy depositions originating from charged and neutral particles in calorimeters and avoid **confusion** among those

Clustering



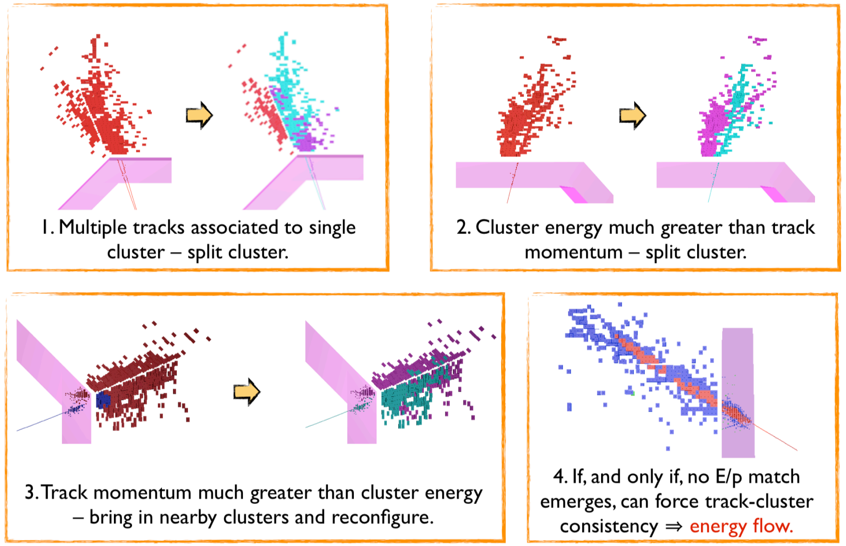
Track to Cluster Association

- Track-cluster association algs match cluster positions and directions with helix-projected track states at calorimeter.
- In very high-density jets, reach limit of "pure" particle flow: can't cleanly resolve neutral hadrons in hadronic showers.
- Identify pattern-recognition problems by looking for significant discrepancies between cluster E and track p.
- Choose to recluster: alter clustering parameters or change alg entirely until cluster splits and consistent E/p achieved.

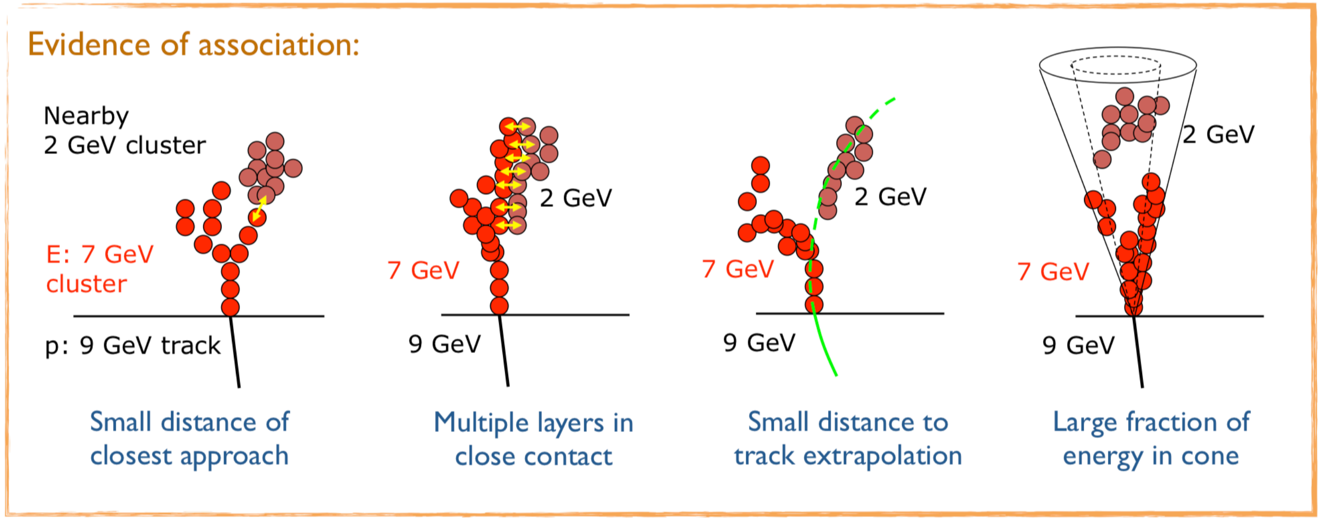


J. S. Marshall: https://indico.in2p3.fr/event/7691/contributions/42712/attachments/34375/42344/3_john_marshall_PFA_marshall_24.04.13.pdf

Re-Clustering



Fragment Removal

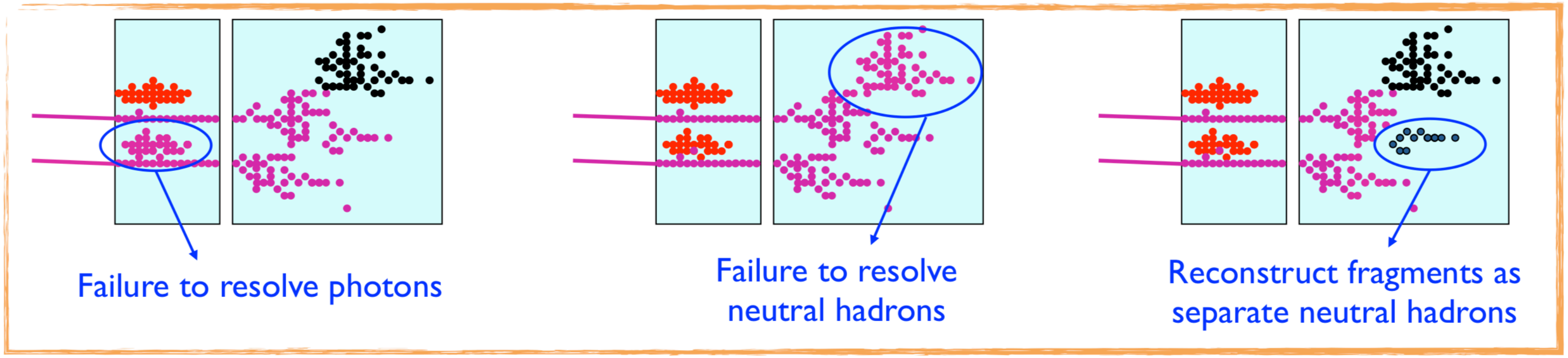


Confusion Scenarios

The Limit of Particle Flow Reconstruction

- **Topologically or energetically confusing** events could cause problems for PFA reconstruction:
  - ➡ Missing or double counted energy limiting jet energy resolution at high energies

Types of confusion



Missing energy | **Missing energy** | **Double counted energy**

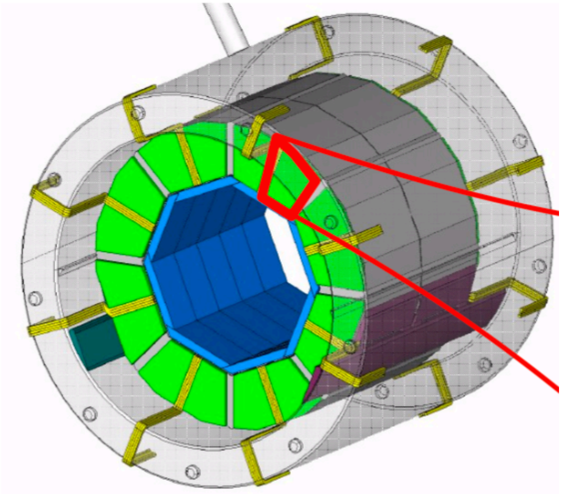
J. S. Marshall: https://indico.in2p3.fr/event/7691/contributions/42712/attachments/34375/42344/3_john_marshall_PFA_marshall_24.04.13.pdf

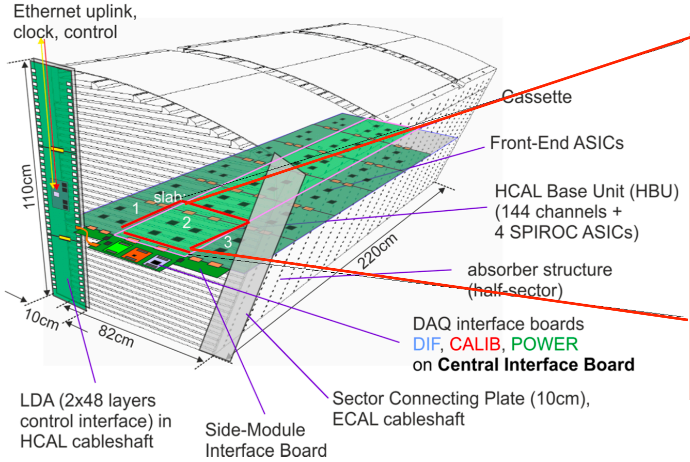
- Crucial requirements for Particle Flow designed detector systems keeping confusion on considerable level:
  - ➡ Calorimeters within magnetic coil for proper track-cluster associations
  - ➡ **High granularity calorimeters** to fully exploit pattern recognition algorithms

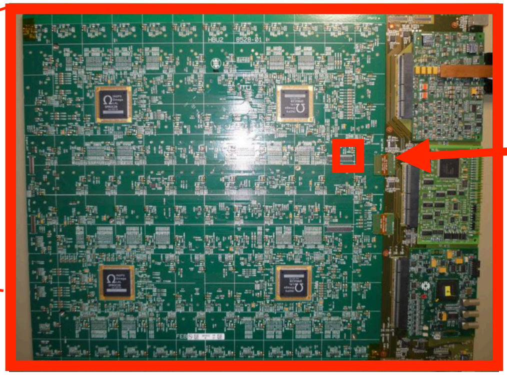
The Analog Hadron Calorimeter (AHCAL) @ ILD

Designed for Particle Flow Reconstruction

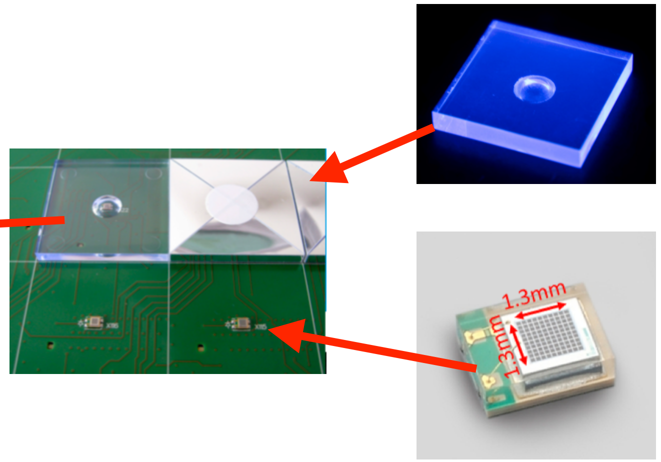
- Highly granular sampling calorimeter for the International Large Detector
  - ➡ Total of ~8 million single channels: Wrapped scintillator tile coupled to SiPM readout
- **H**CAL **B**ase **U**nit: 36 · 36 cm^2 featuring 4 ASICs reading out 144 channels
- Fully integrated detector design to octagonal cylinder
  - ➡ Front-end readout electronics, internal LED calibration system, no cooling within active layers







HBU

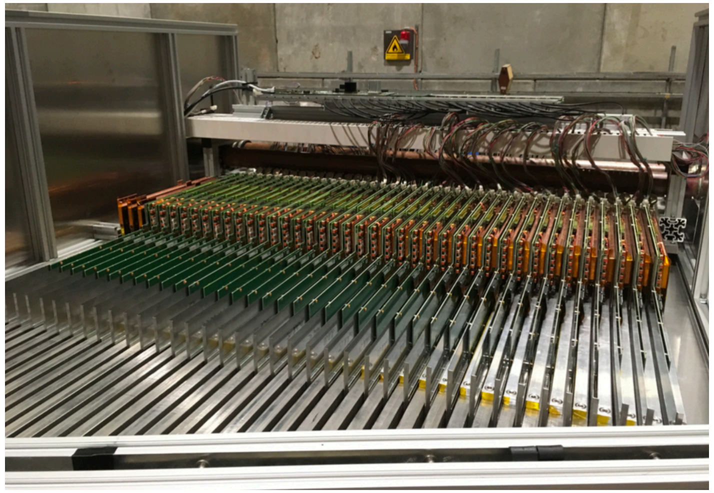


The Analog Hadron Calorimeter Prototype 2018

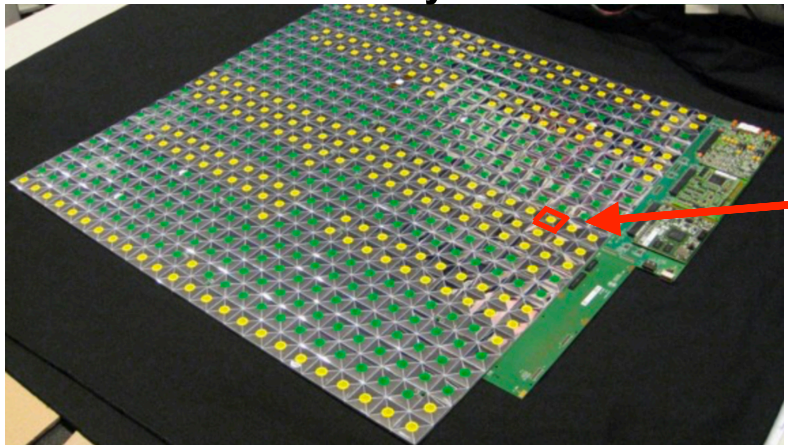
A Highly Granular SiPM-on-tile Sampling Calorimeter

- 38 layer steel sampling calorimeter (~4 λ_n) featuring a total of **~22k channels**
- Active layers (72 x 72 cm^2) consisting of 576 channels
  - ➡ One channel: Silicon-Photomultiplier (SiPM) coupled to wrapped scintillating tile (3 x 3 cm^2)
- **Compact design**: Fully integrated front-end readout electronics, no active cooling
- In 2018: Three successful test beam campaigns at SPS CERN collecting electron/muon/**pion data**

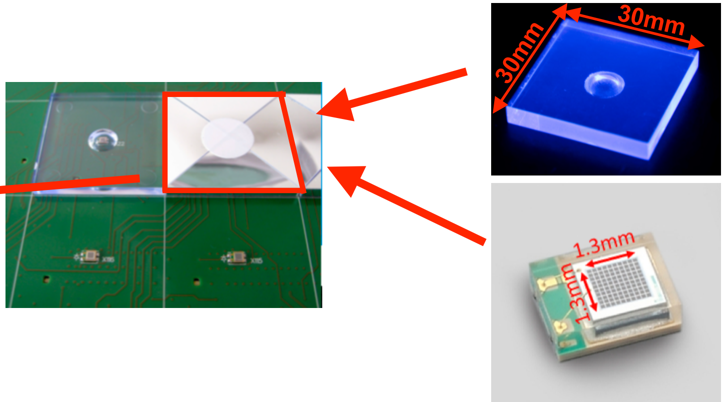
38 layers within steel absorber stack



One layer



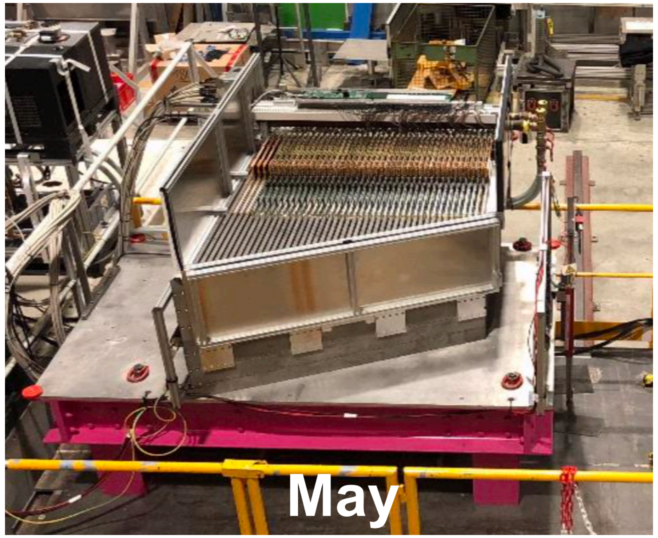
One channel: Scintillating tile + SiPM

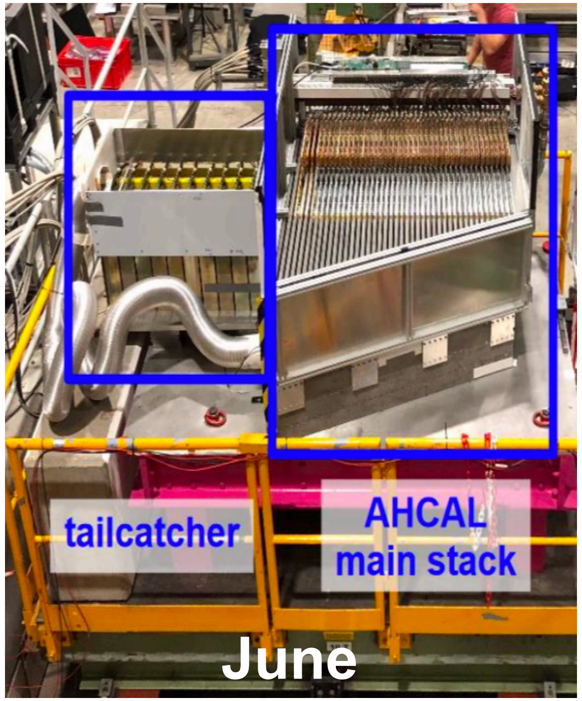


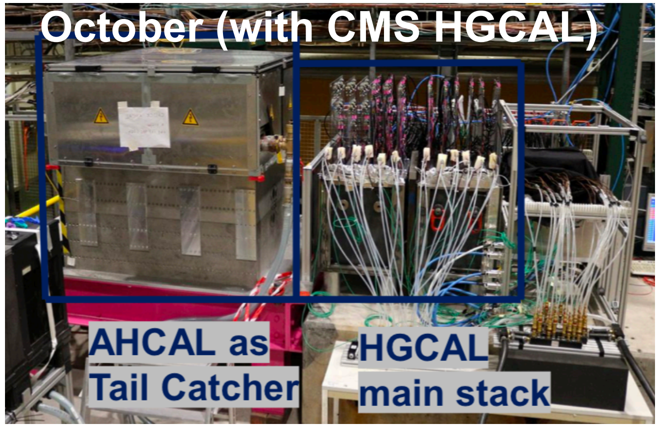
The CALICE AHCAL Beam Test Campaigns 2018

May, June and October @ SPS Cern

- Three successful beam test campaigns at SPS CERN in 2018
- Data sets:
  - ➡ Muons, electrons, **pions**
  - ➡ Energies: 10 - 200 GeV
  - ➡ Events: Multiple 10 million, also at different detector positions
- For this studies: June 2018 beam test data







Analysis Examples

PandoraPFA on AHCAL 2018 Prototype Data

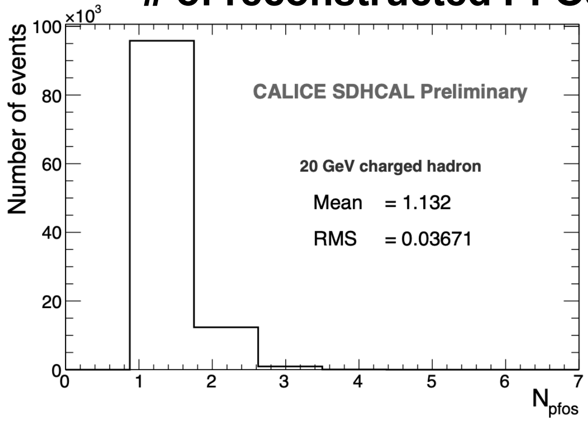
PandoraPFA

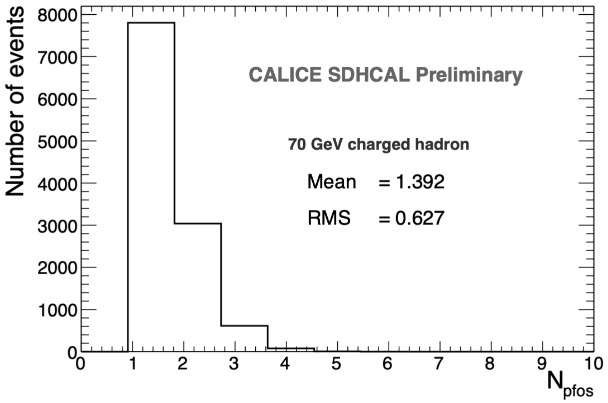
PandoraPFA

Scenario 1: Single particle reconstruction studies

Scenario 2: Two-particle separation studies

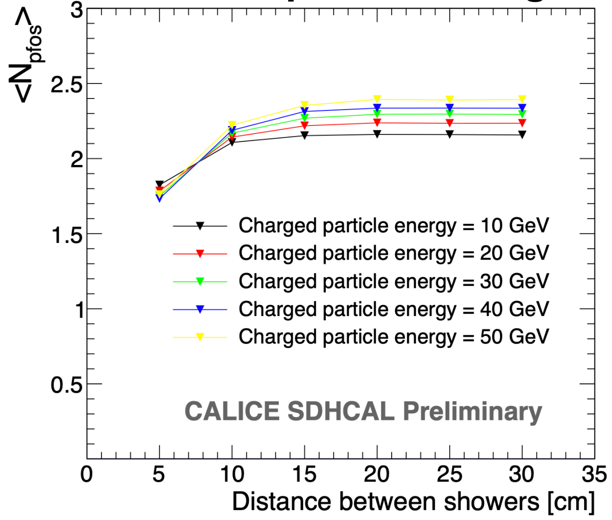
# of reconstructed PFOs for different particle energies



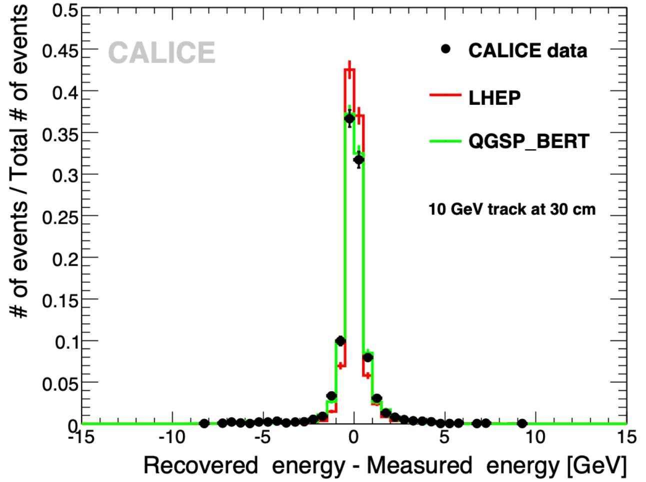


http://cds.cern.ch/record/2669487/files/fulltext.pdf

Mean # of reconstructed PFOs for different particle energies



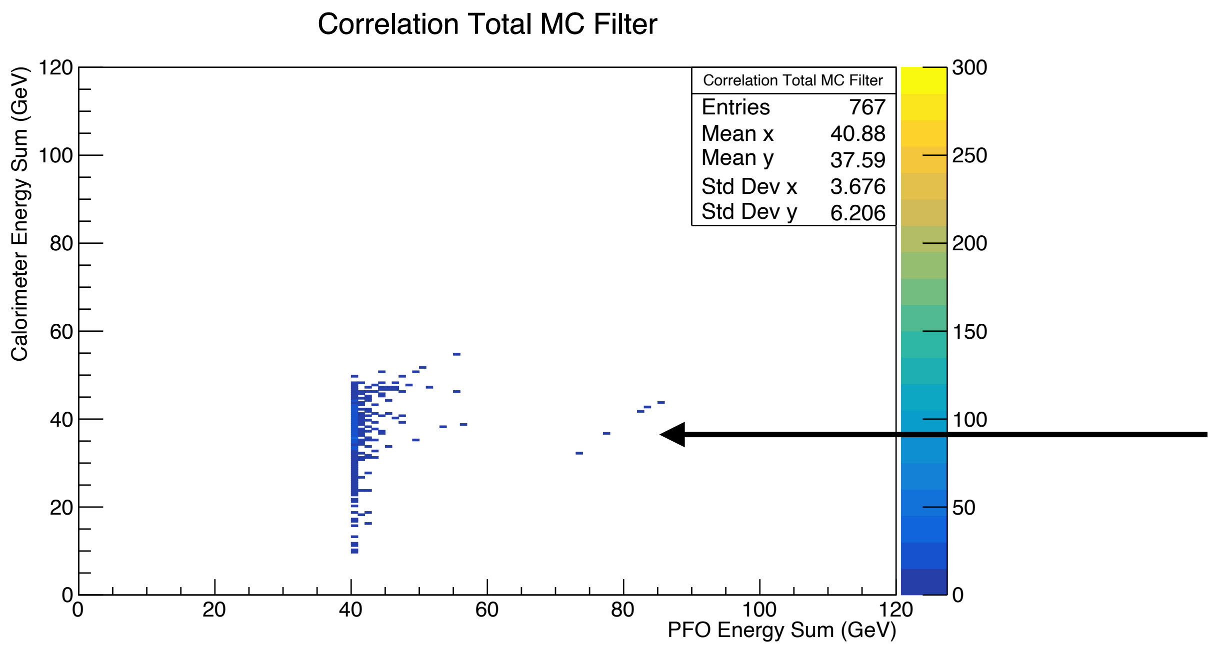
Energy difference of calorimeter pure & PFO reconstruction



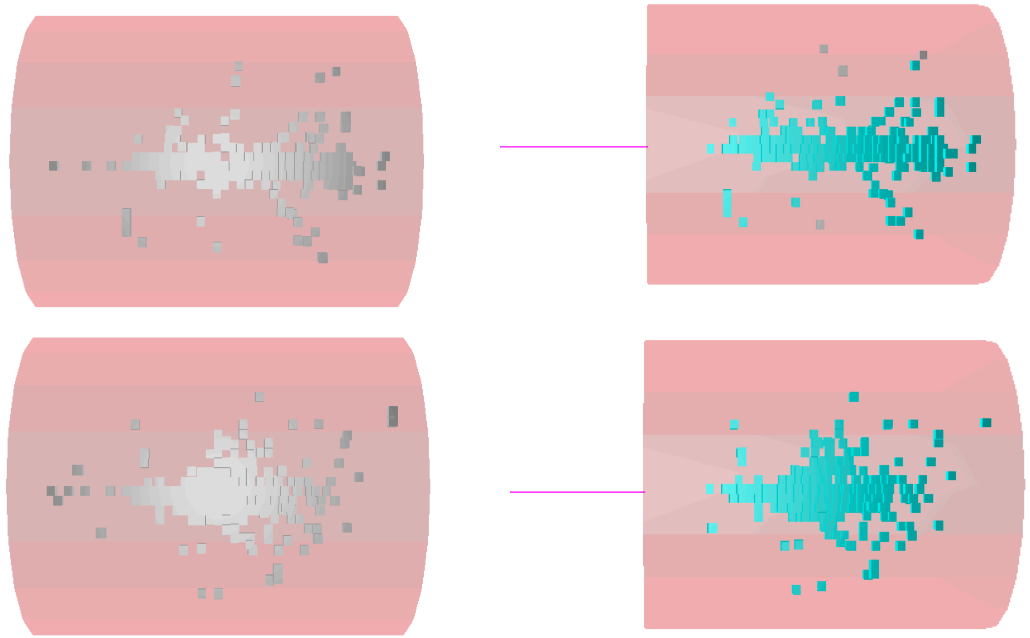
https://arxiv.org/abs/1105.3417

Isolated Hit Merging Algorithm - PandoraPFA

Diagonal „Gap" Events



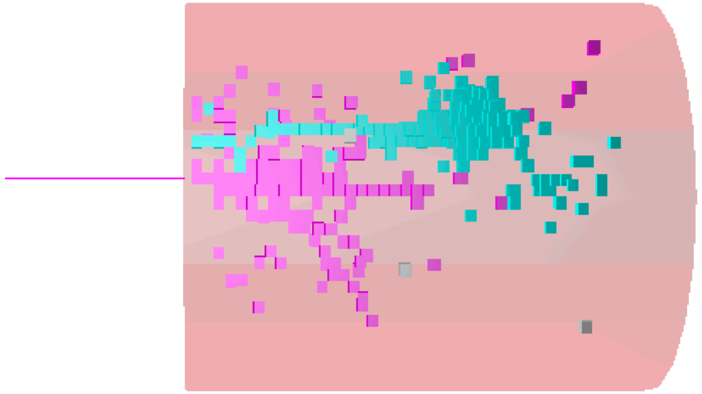
Calorimeter Hit Level **Final PFO Level**



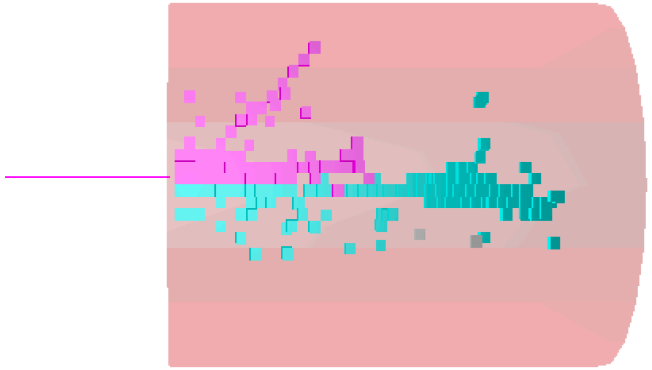
- 98% of diagonal events (non working track-cluster assignment) caused by PandoraPFA's "Isolated Hit Merging“ algorithm
  - Topological gap in primary track of first layers causes algorithm to „cut off“ hits which makes track-cluster assignment impossible
  - ➡ Fine granularity of ECAL in front missing allowing first assignment of track before AHCAL

High Energy Events in 20 GeV Pion Beam Test Data

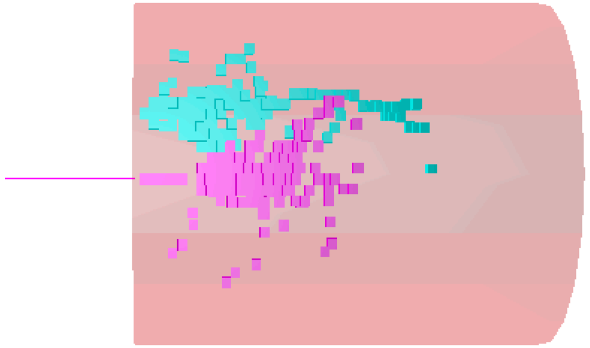
Examples - Multi Particle Events?



Energy in Calorimeter: ~40 GeV



Energy in Calorimeter: ~42 GeV



Energy in Calorimeter: ~38 GeV

Sample Preparation & Selection Tools

Sample Preparation & Selection Tools

Overview & Status

- Event Selection:
  - ➡ Shower start finder algorithm: Implemented and optimised in cooperation with Jonas Mikhaeil
  - ➡ PID (Boosted Decision Tree): Talk by V. Bocharnikov
  - ➡ Event filter: Implemented with selection criteria on shower start layer, shower position, track quality, etc.

- Event Preparation for PandoraPFA:
  - ➡ MIP to GeV conversion: Implemented for EM and HAD scale
  - ➡ Event overlay: Implemented
  - ➡ **Data tracks from DWC and MC tracks**: Implemented and validated
  - ➡ **Primary track removal (based on shower start layer):** Implemented and validated

Illustration of implemented tracks

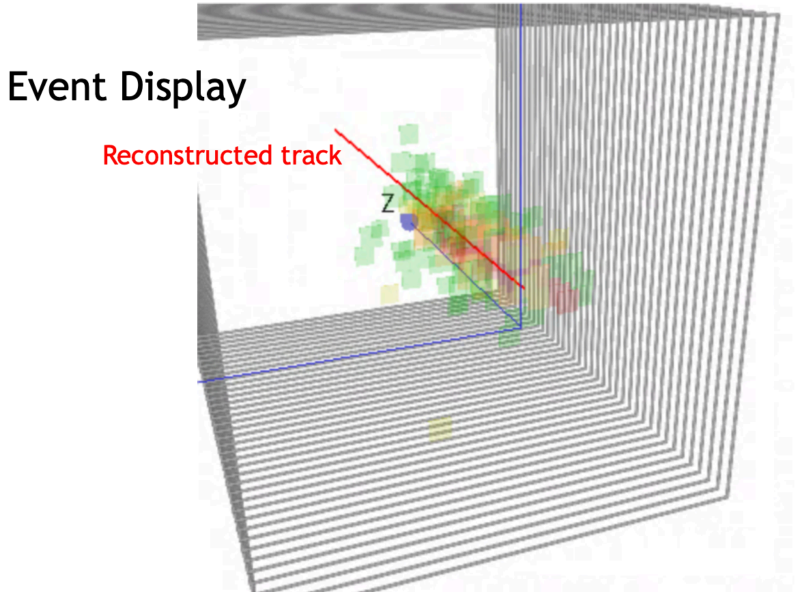
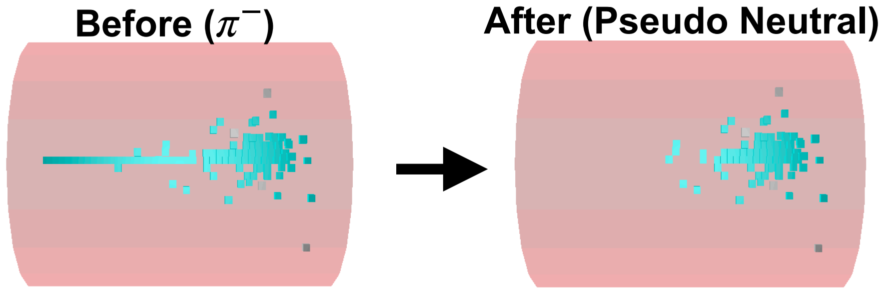


Illustration of pseudo neutral generation

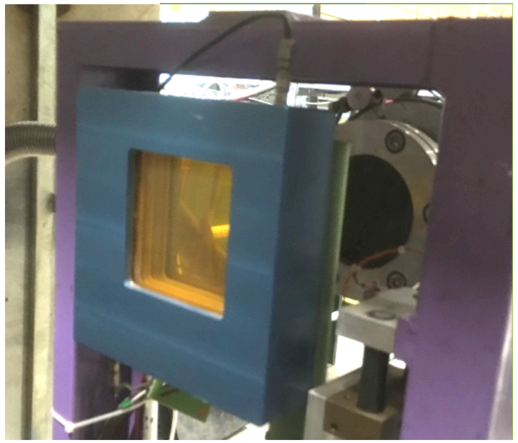


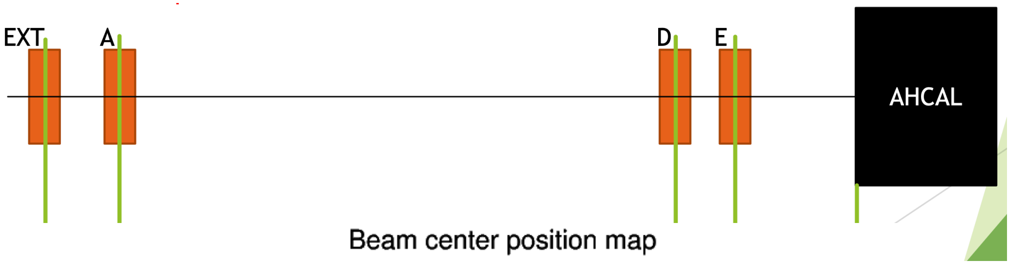
Delay Wire Chambers (DWC)

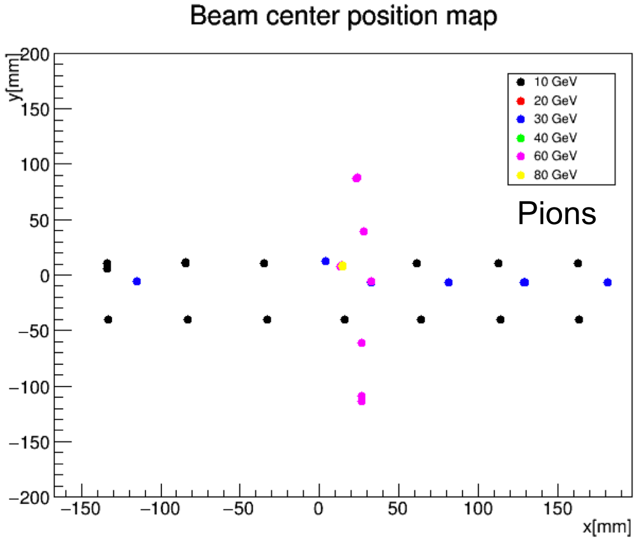
Providing Tracks for Beam Test Events

- Beam Test June 2018 at SPS CERN: Four 100 x 100 mm^2 delay wire chambers (MWPCs)
- Position resolution of each chamber: ~600 µm
  - ➡ **Sub-mm resolution at AHCAL**
- Information extracted:
  - ➡ **Reconstructed track for each event**
  - ➡ Position calibration (Prototype moved on X-Y stage during beam test for position scans)
  - ➡ Measurement of scintillator tile gaps

Work done by Linghui Liu (U. Tokyo)
(https://agenda.linearcollider.org/event/8368/contributions/44971/attachments/35214/54544/LL_AHCALmain_2019.pdf)







Track Quality Check

Implemented MC and Data Tracks for PandoraPFA Studies

Note: Tracks almost completely straight since no B-field present and particles almost only with p_z

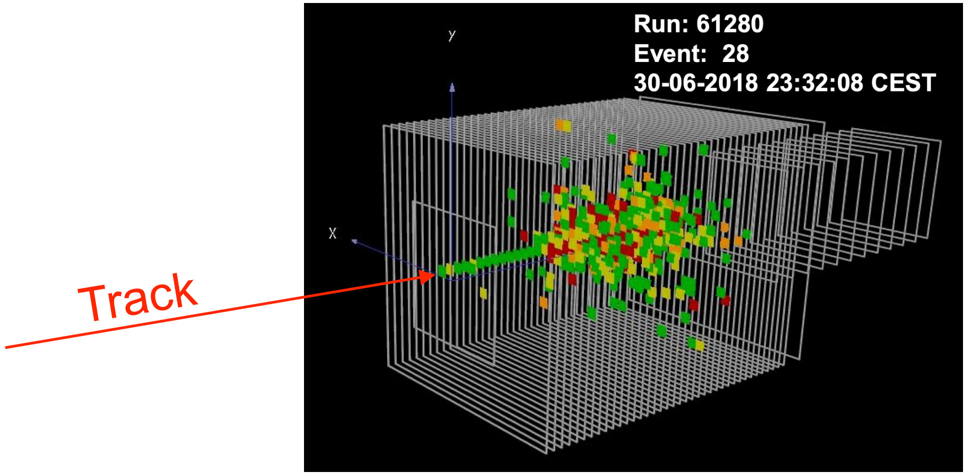
- Data tracks: Reconstructed from DWC of beam test
- MC tracks: MC primary particle endpoint position X/Y extrapolation

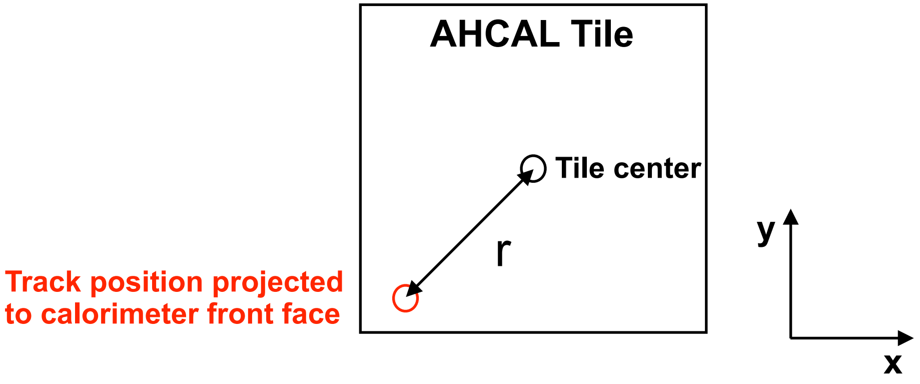
➡ **Track quality?**

How well does track position at calorimeter front face agree with cog in X/Y of event (central shower axis)?

How well does track hit first triggered channel of primary track in layer 1?

Does track hit any triggered channel in layer 1 at all?

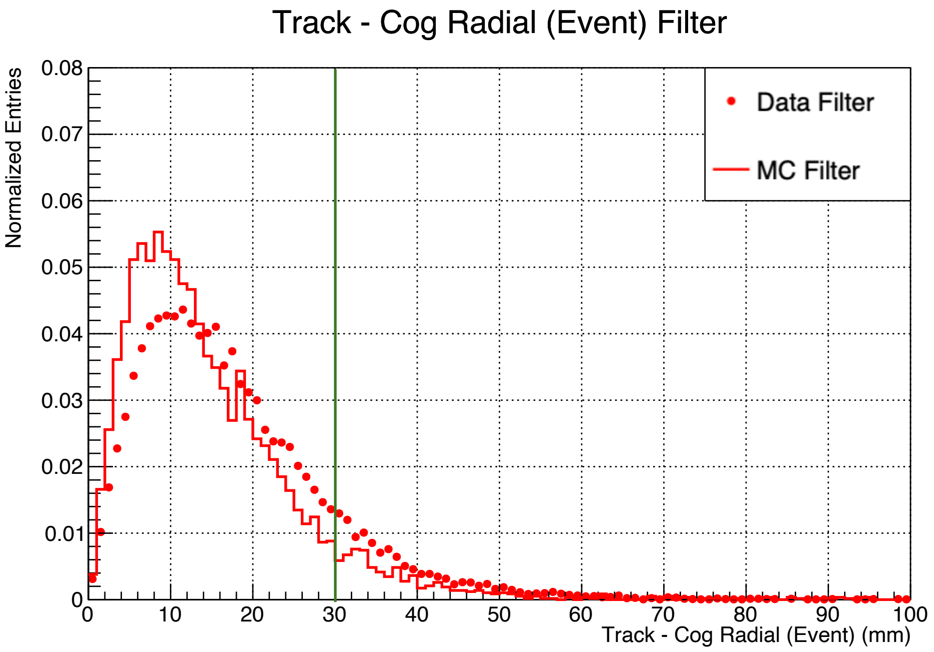




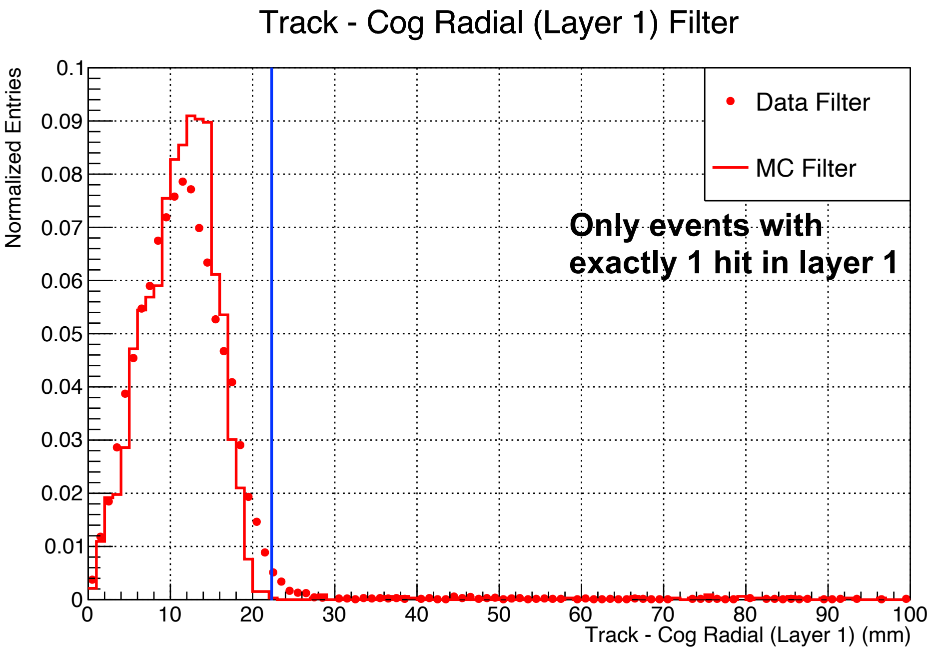
Track Quality Results 20 GeV π^-

Precise Tracks for PandoraPFA Reconstruction

Definition Filter: Applied BDT-PID, Shower start layer < 20, Hit in layer 1+2+3



$$r = \sqrt{(x_{track} - x_{cog})^2 + (y_{track} - y_{cog})^2}$$



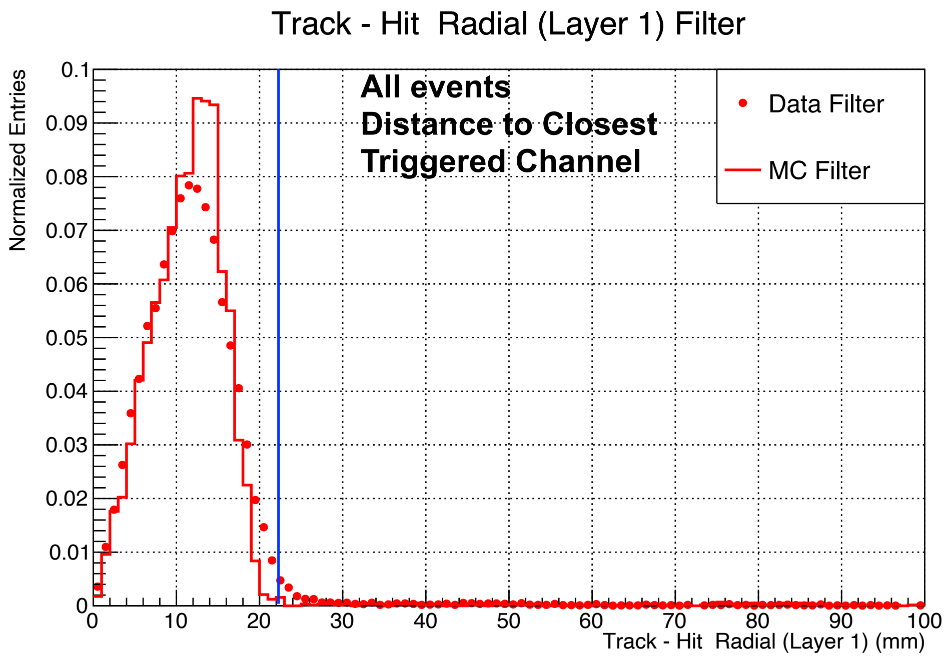
$$r = \sqrt{(x_{track} - x_{hit})^2 + (y_{track} - y_{hit})^2}$$

- Excellent agreement of track and cog (central shower axis) position:
  - ➡ 88.5% (data) and 93% (MC) of events within 30 mm distance (one tile length)
- Most of the tracks hit triggered channel of primary track in layer 1:
  - ➡ 98.2% (data) and 99% (MC) of events within 22 mm radius (tile center - corner distance)

Track Quality Results 20 GeV π^-

Precise Tracks for PandoraPFA Reconstruction

Definition Filter: Applied BDT-PID, Shower start layer < 20, Hit in layer 1+2+3



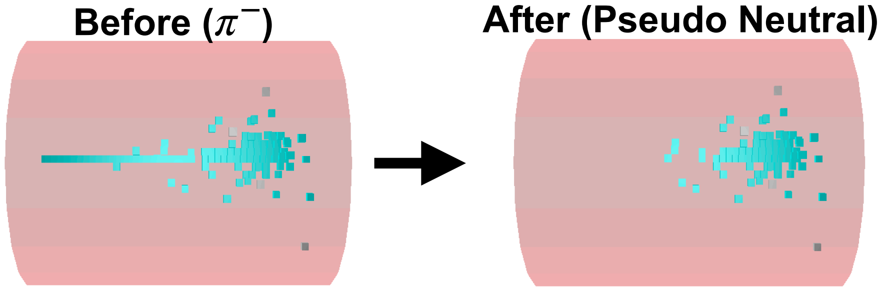
$$r = \sqrt{(x_{track} - x_{hit})^2 + (y_{track} - y_{hit})^2}$$

- Most of the tracks hit a triggered channel in layer 1:
  - ➡ 97.5% (data) and 98.5% (MC) of events within 22 mm radius (tile center - corner distance)
- Similar results achieved for:
  - ➡ Less strict filter options in terms of hit requirements in first layers
  - ➡ Lowest energy scenario of 10 GeV π^-

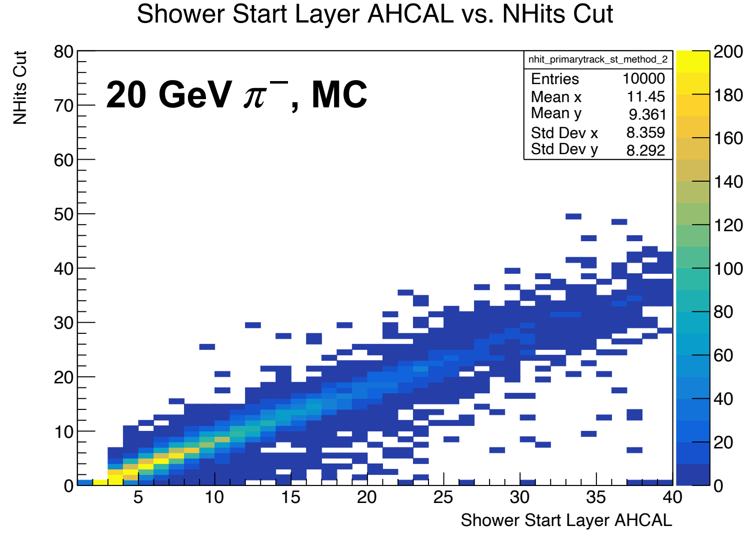
➡ **Excellent track quality validated for data and MC**

Finding and Removing Primary Track

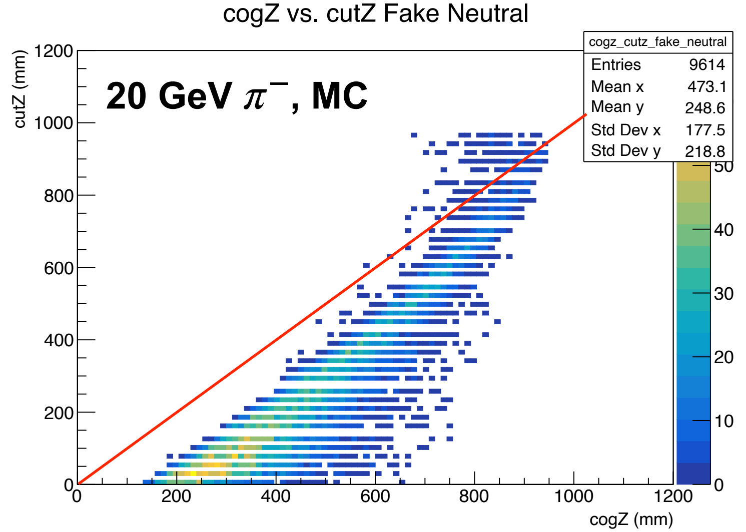
The Method for Creating Pseudo Neutral Hadrons



- Conditions for hit to be considered as primary track hit and being removed:
  - ➡ Hit located in layer before shower start layer - 1
  - ➡ Hit position within r = 60mm to cogX/Y of shower (central shower axis)
  - ➡ Hit energy < 3 MIP

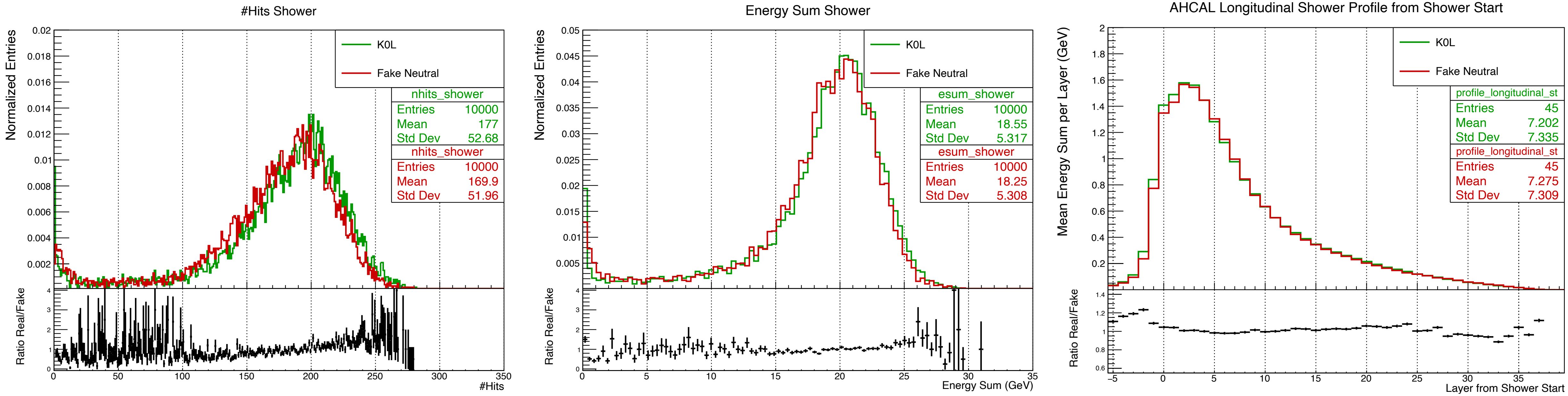


- Method robust and working well:
  - ➡ # cut hits (primary track) well correlated with shower start layer
  - ➡ Z position of potentially last cut hit well before cogZ for most events



Comparison: Real vs. Pseudo Neutrals 20 GeV (MC)

Validation of Primary Track Removal Algorithm

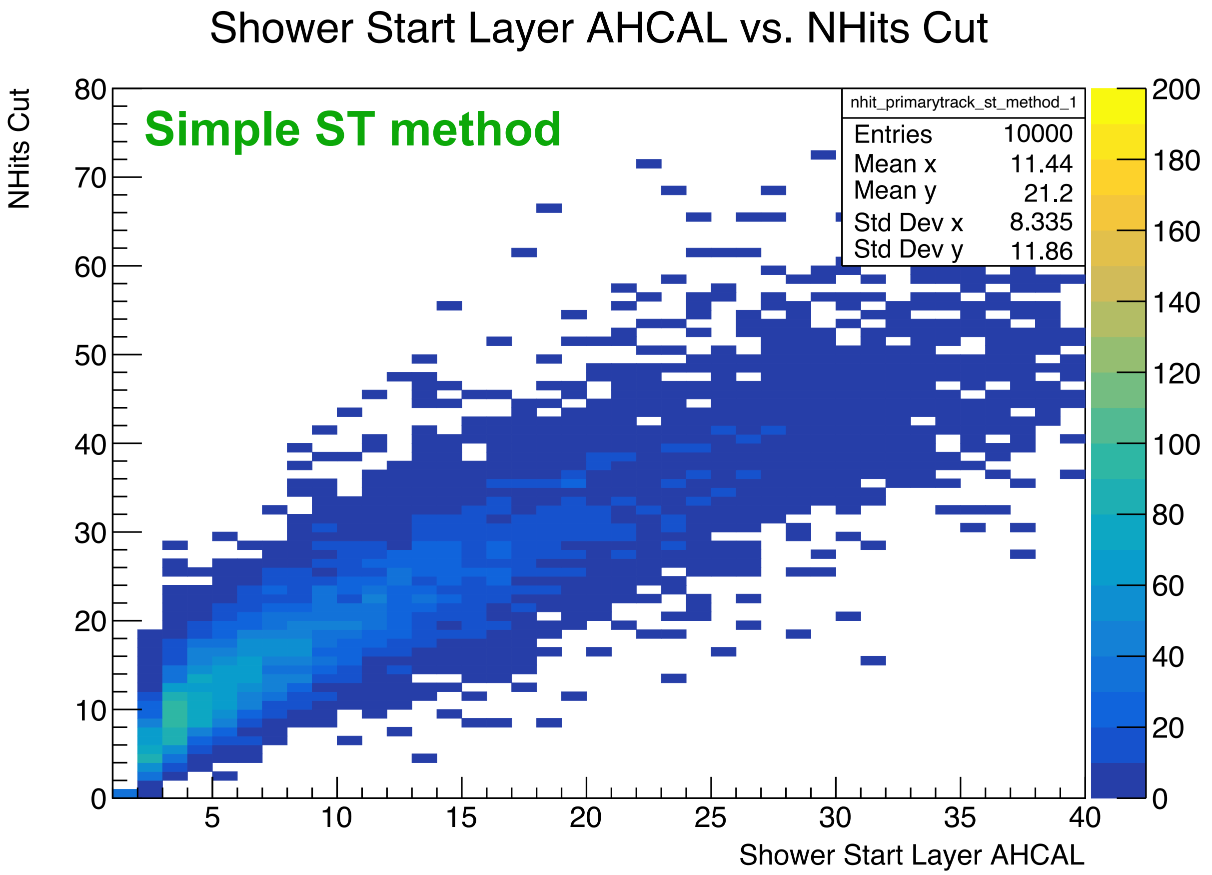


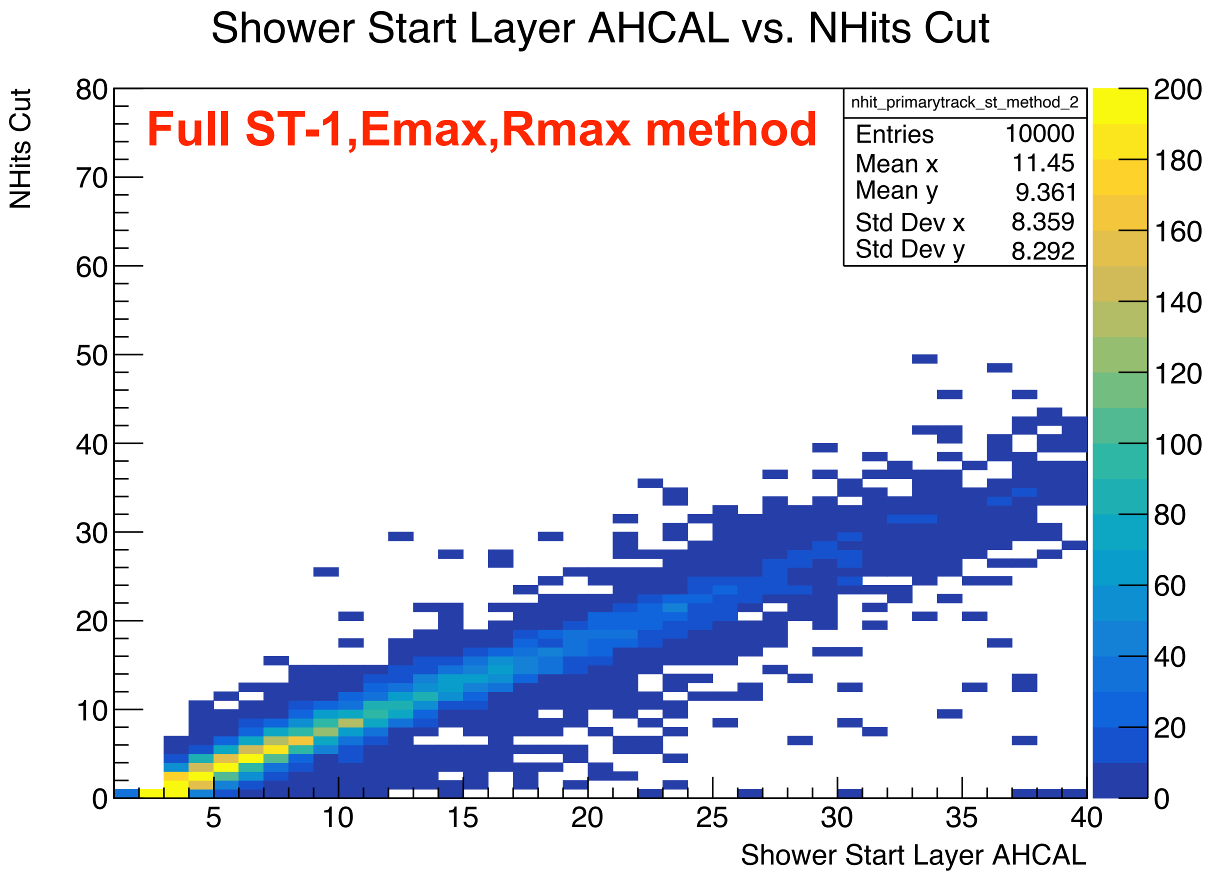
- In general **good agreement** between real neutrals (K0L) and pseudo neutrals (cut π^-) in number of hits, energy sum and longitudinal shower profile

➡ **Pseudo-neutrals validated for charged-neutral separation studies (response and topology)**

Number of Hits (Primary Track) vs. Shower Start Layer

Validation of Method

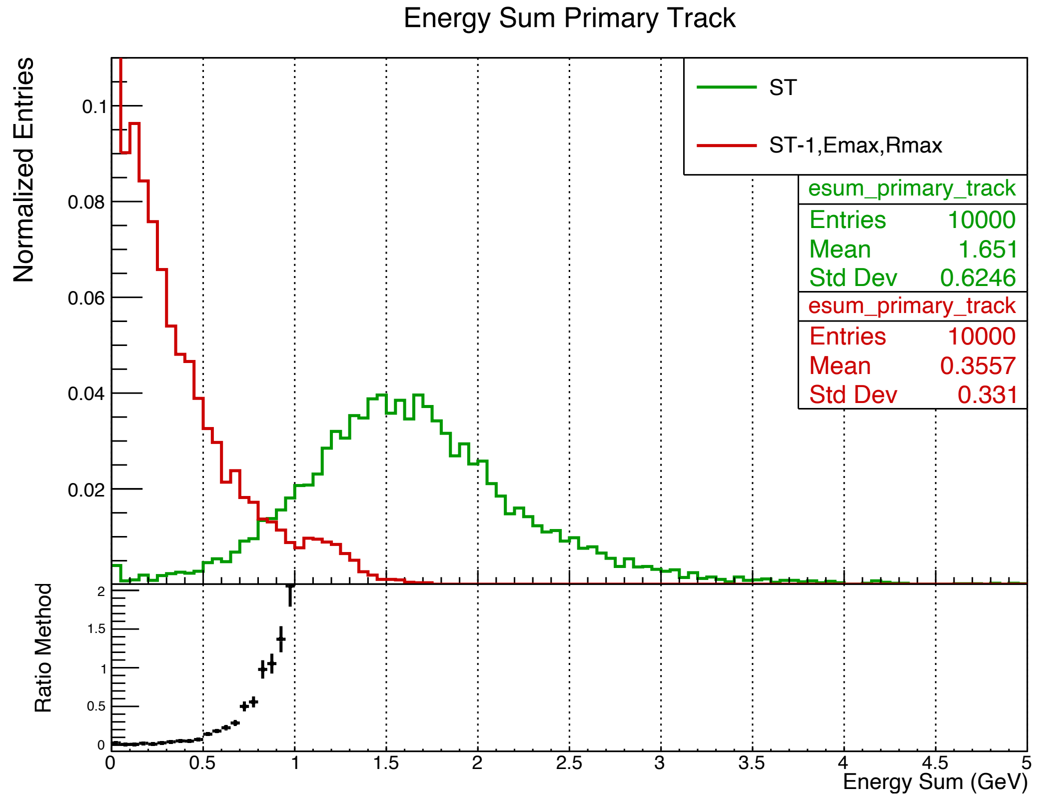


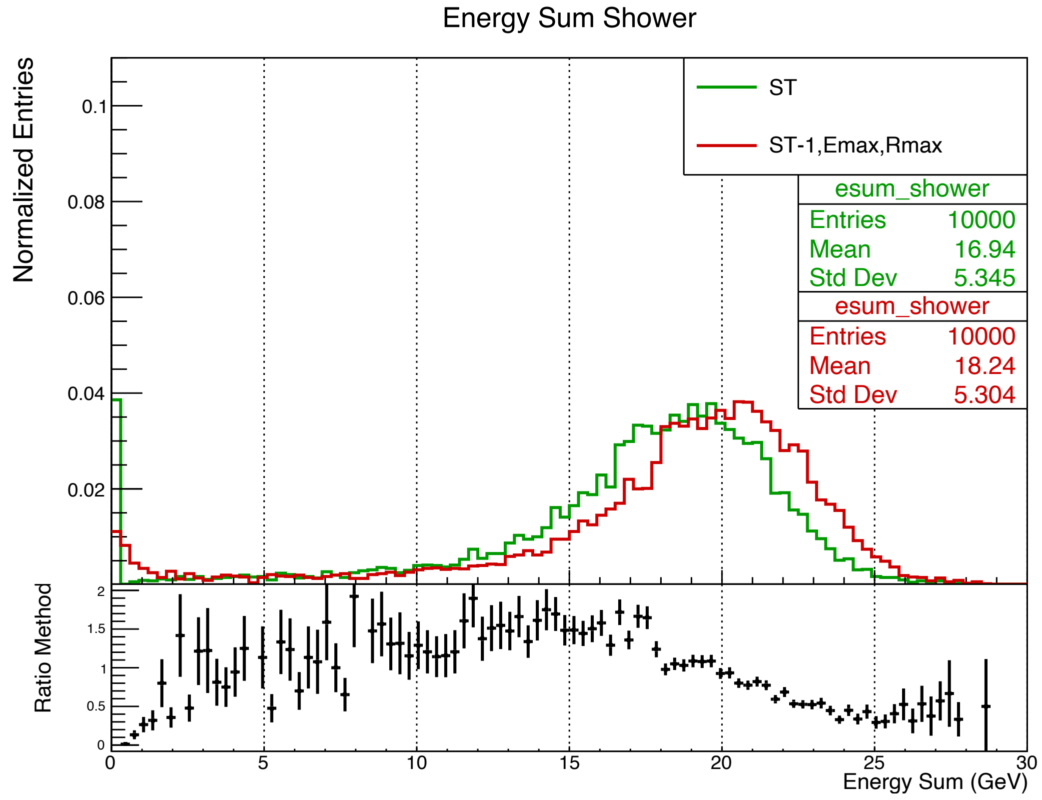


- Too many hits cut away for simple ST method
- Much better correlation of shower start layer and cut nHits of classified primary track for advanced method (#Cut hits ≈ #shower starter layer)

Energy Sum: Primary Track and Shower Hits

Validation of Method

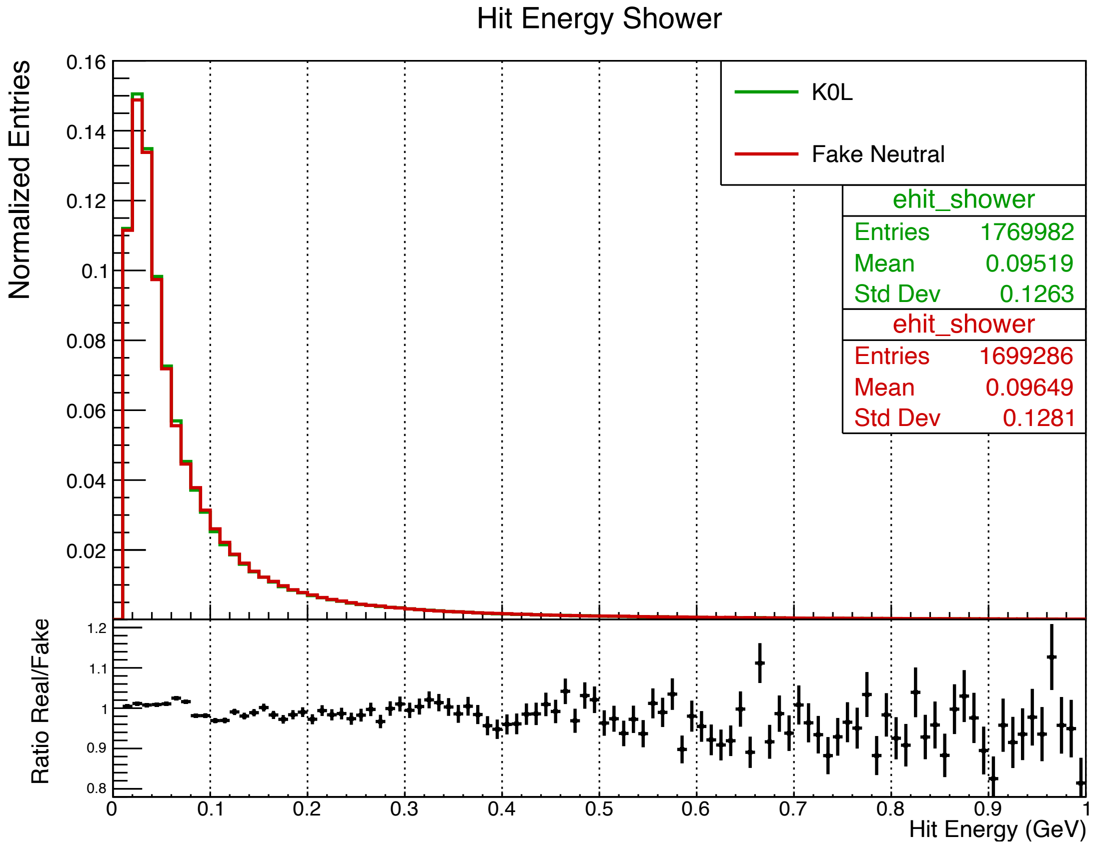


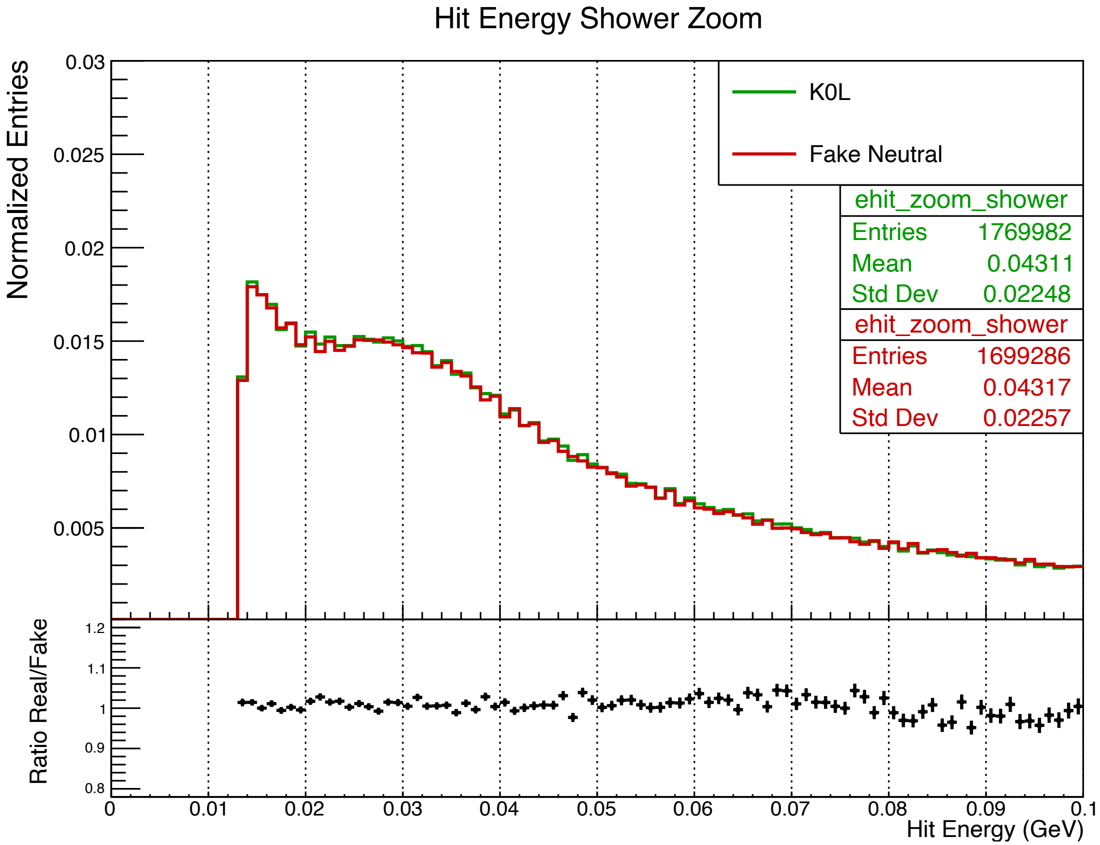


- Shower energy sum much closer to 20 GeV for advanced method
  - ➡ Too much hits and therefore energy cut away with simple method
  - ➡ Simple estimate: Upper primary track energy sum expected for perfect 40 hit MIP track: 0.0268 GeV (1 MIP) * 40 (layers) * 1.4 (landau-gaussian mean) = **~1.5 GeV**

Hit Energy

Real vs. Pseudo Neutrals



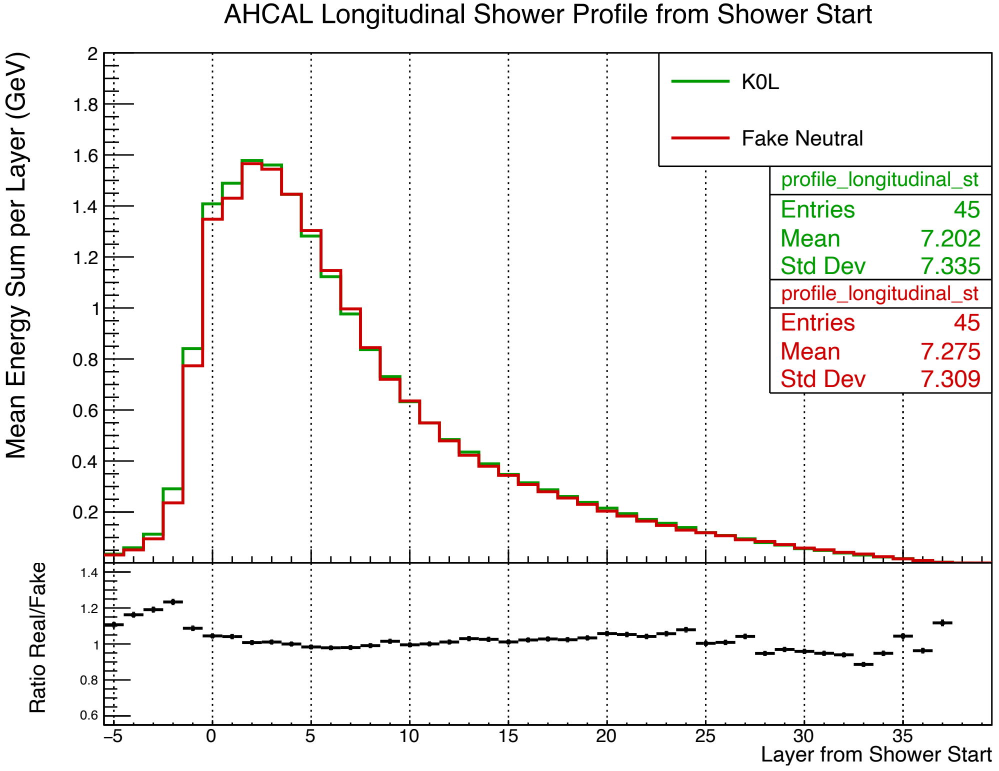


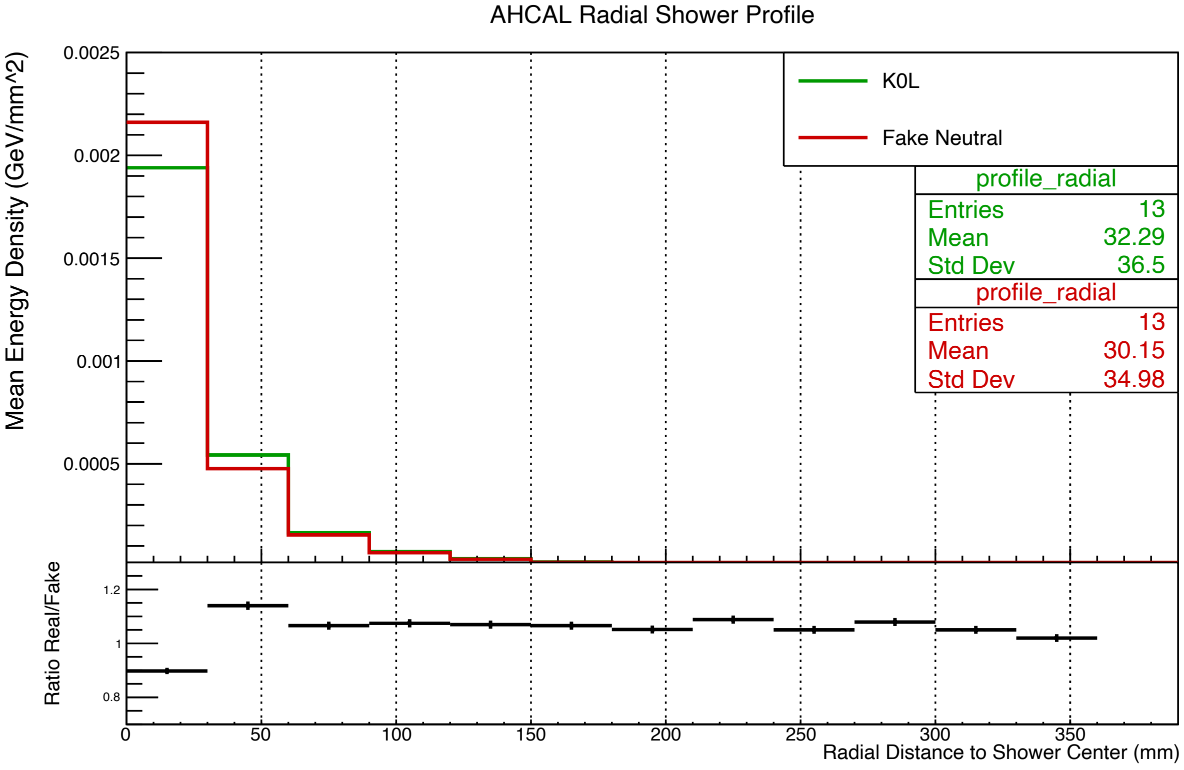
- Very good agreement, even for low energy hits (within 2%)

Shower Profiles: Longitudinal & Radial

Real vs. Pseudo Neutrals

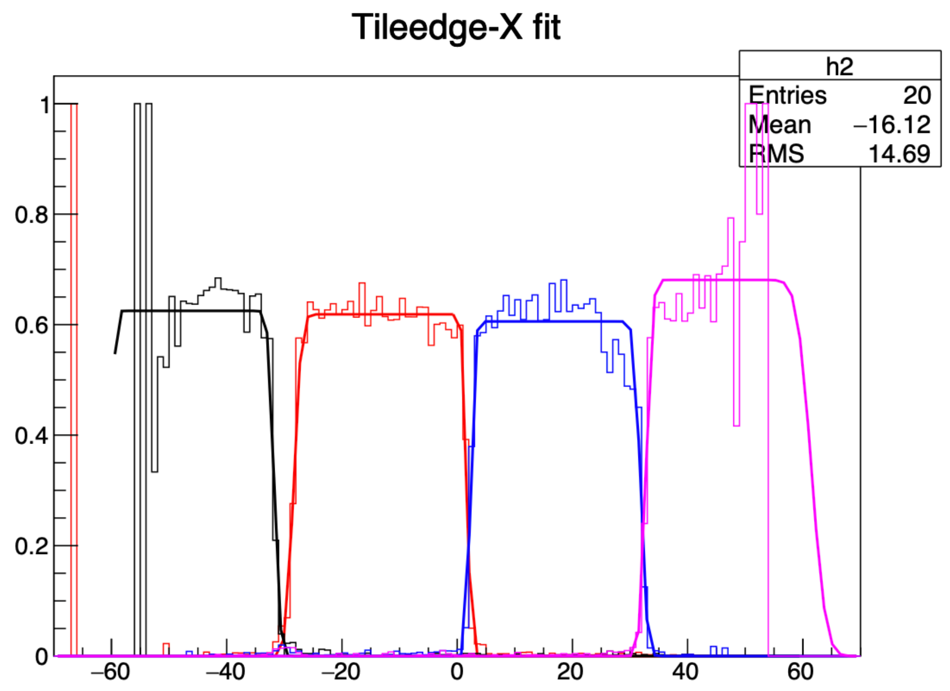
Simple radial profile code:
13 concentric circle areas, no fractional sharing of tile energy between two circle areas if overlap at edge

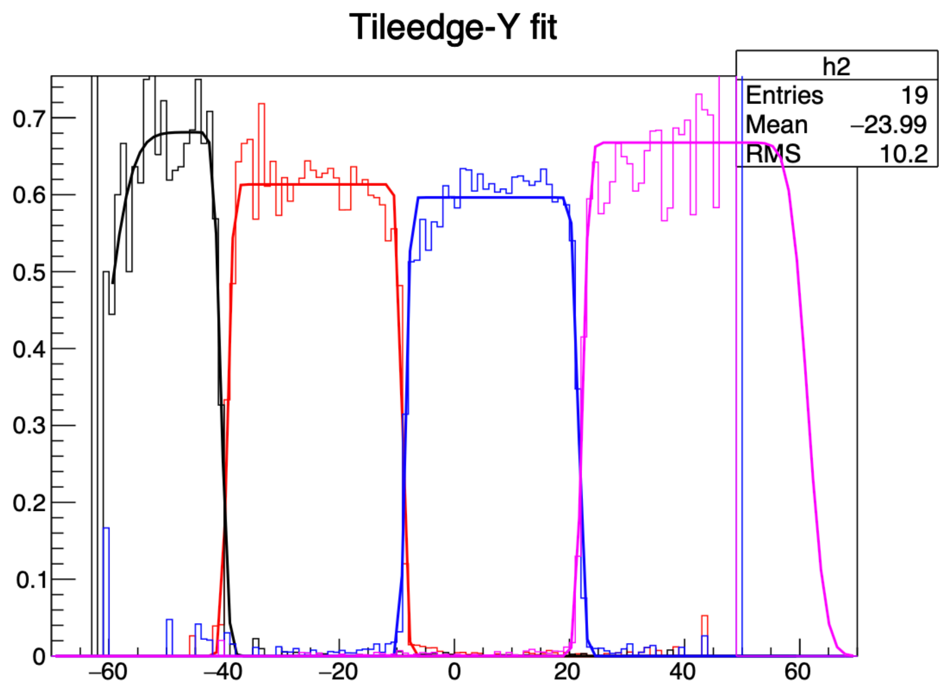




- Reasonable agreement for shower profiles:
  - ➡ Longitudinal: ~20% discrepancy ±2 layer around shower start layer
  - ➡ Radial: ~10-15% discrepancy for first two bins / innermost two circles

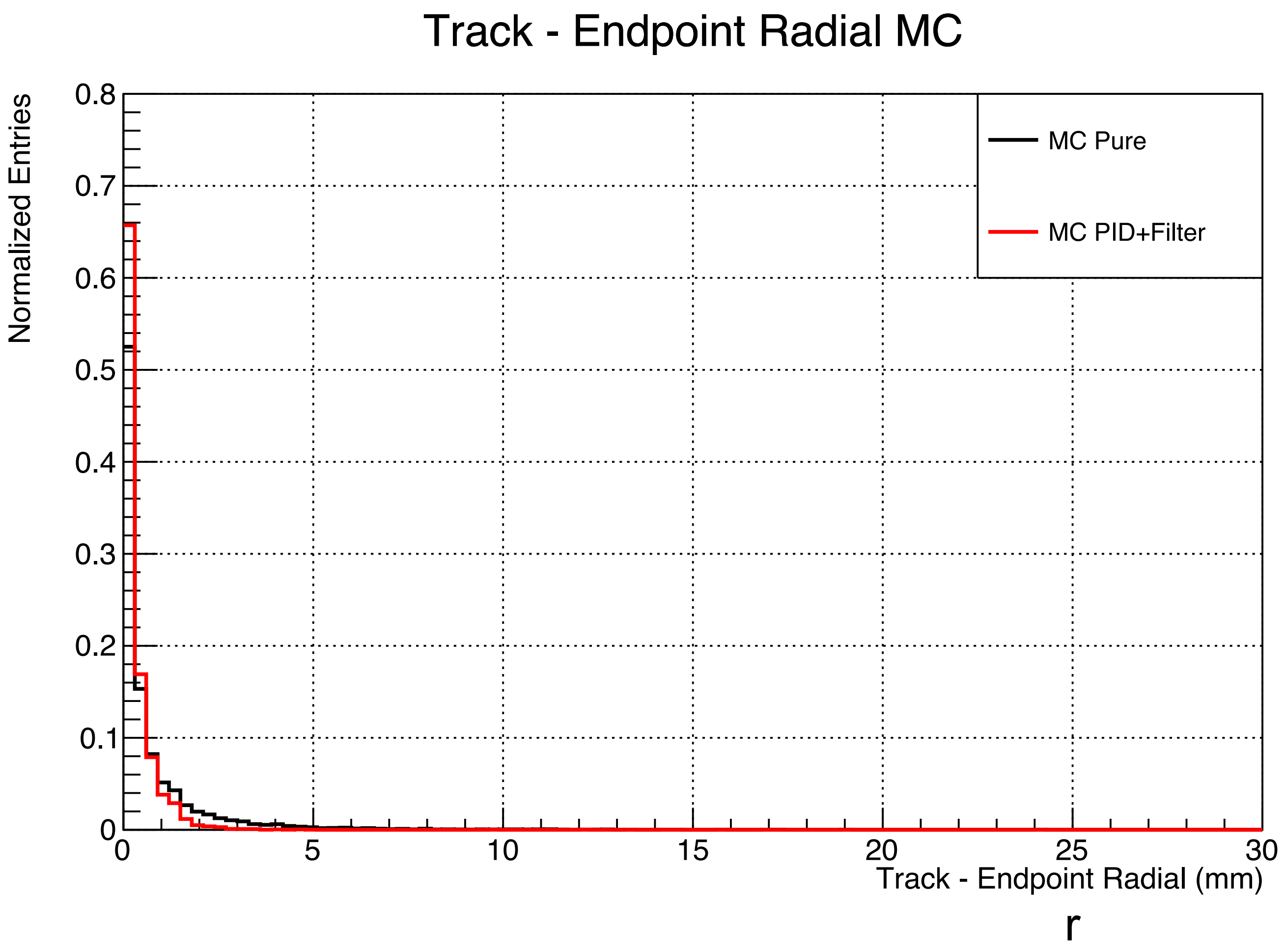
Scintillator Tile Gaps Measurements DWC Example





MC: Track to MC Endpoint Position Comparison

Track Quality Study



- Only events with primary particle endpoint z within calorimeter

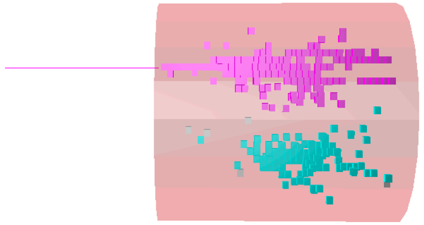
- Radial distance in x-y plane:

$$r = \sqrt{(x_{track} - x_{endpoint})^2 + (y_{track} - y_{endpoint})^2}$$

- Very good agreement between implemented MC track and „truth MC track"
  ➡ 100% of events within 10 mm distance

Basics of Overlay Processor

Estimate of Radial Distance Covering



Magenta: Charged Hadron
Cyan: Neutral Hadron
Grey: Unclustered Hits

- **Overlay processor implemented and working well** (https://stash.desy.de/projects/CALICE/repos/calice_analysis/browse/addonProcs/src/MergeProcessor.cc) - **Big thanks to Linghui for great work and synchronisation on that!** ✔

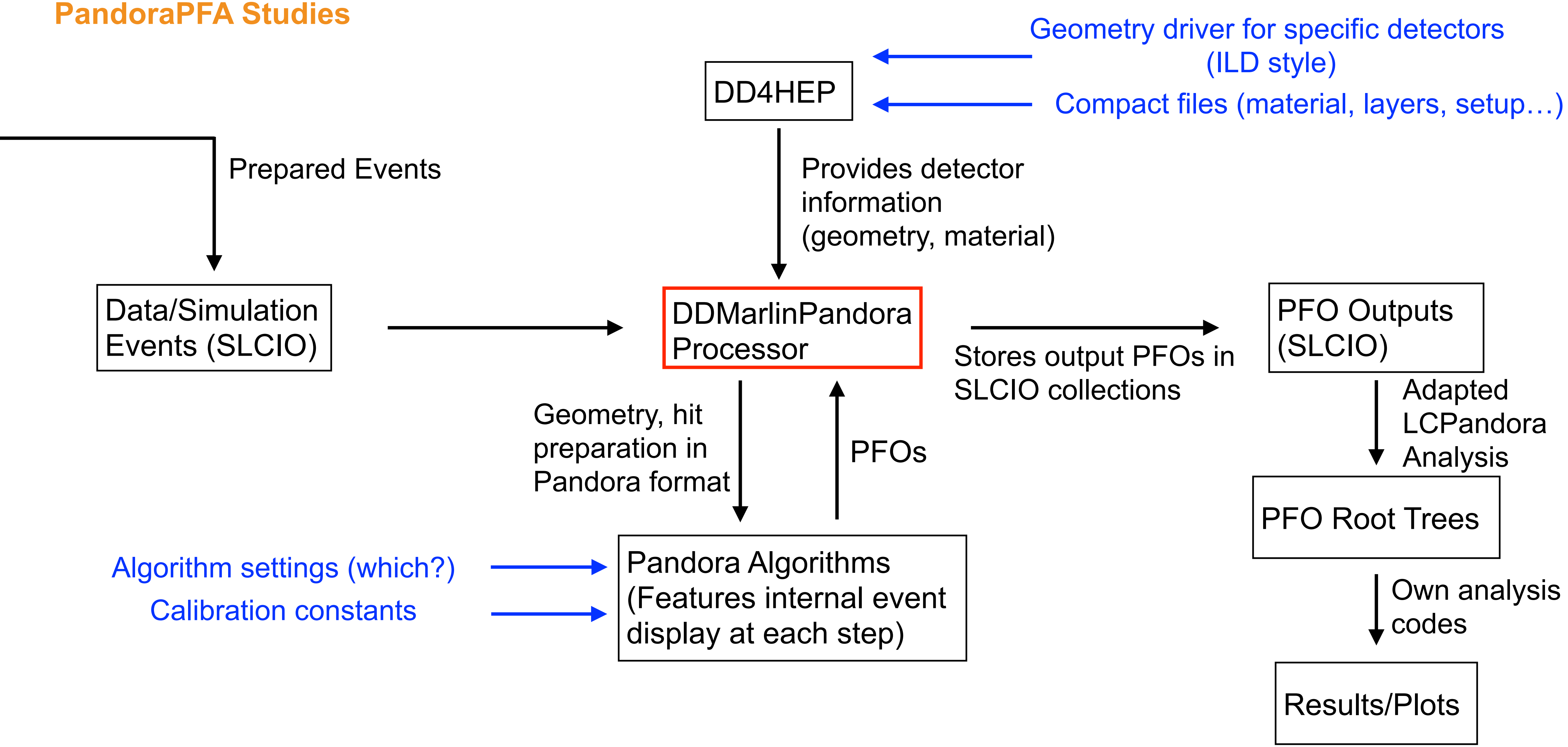
- Requirements (not available in general ILD version):
  - ➡ Proper flagging of merged output hits and saving of individual output collection 1,2 and merged
  - ➡ Proper handling of MIP threshold - Apply 0.5 MIP cut only on overlaid hits
  - ➡ Radial shower distance saving according to cogX,Y of shower pairs
  - ➡ Subsequent event overlay from two input (neutral & charged) LCIO collections

✔

The PandoraPFA Framework: Implementation, Calibration & Basic Checks

Framework / Data Flow Diagram

PandoraPFA Studies



Setting up the PandoraPFA Framework

Technical Challenges & Solutions

Many aspects considered while implementing PandoraPFA from a 4π detector setup (like ILD) to our AHCAL standalone (+tracks) scenario:

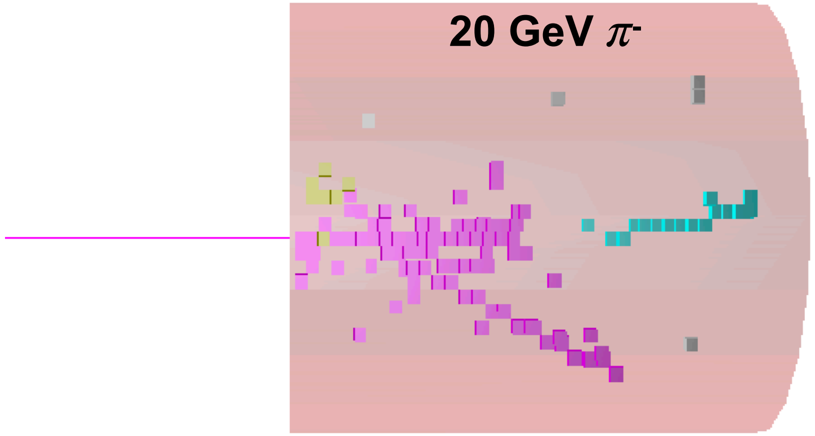
- Simplified detector geometry and related geometry drivers
  - ➡ Careful implementation
- No real tracker, ECAL, muon detector, no B-field
  - ➡ Disable/Re-write related parts code in interface processor
  - ➡ Re-define so-called pseudo layer plugin
  - ➡ Enable algorithm chain step-by-step and check for dependencies, internal cuts & problems (# sub-algorithms/event ~65-90)
- Detector gap implementation
- Internal Pandora energy calibration
- Check available plugins (PID, software compensation,…)

Typical algorithm chain for 1 event

```
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0001, CaloHitPreparation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0002, EventPreparation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0003, ClusteringParent
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0004, ConeClustering
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0005, TopologicalAssociationParent
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0006, LoopingTracks
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0007, BrokenTracks
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0008, ShowerMipMerging
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0009, ShowerMipMerging2
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0010, BackscatteredTracks
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0011, BackscatteredTracks2
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0012, ShowerMipMerging3
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0013, ShowerMipMerging4
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0014, ProximityBasedMerging
[ VERBOSE "MyDDHCALPandora"] ------------> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0016, ConeBasedMerging
[ VERBOSE "MyDDHCALPandora"] ------------> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0017, MipPhotonSeparation
[ VERBOSE "MyDDHCALPandora"] ------------> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0018, SoftClusterMerging
[ VERBOSE "MyDDHCALPandora"] ------------> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] --------> Running Algorithm: Alg0019, IsolatedHitMerging
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0020, SplitTrackAssociations
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0048, SplitMergedClusters
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0050, TrackDrivenMerging
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0051, ResolveTrackAssociations
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0020, SplitTrackAssociations
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0048, SplitMergedClusters
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0053, TrackDrivenAssociation
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0020, SplitTrackAssociations
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0048, SplitMergedClusters
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0054, ExitingTrack
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0056, TrackPreparation
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0015, TrackClusterAssociation
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0057, LoopingTrackAssociation
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0058, TrackRecovery
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0059, TrackRecoveryHelix
[ VERBOSE "MyDDHCALPandora"] ----> Running Algorithm: Alg0060, TrackRecoveryInteractions
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0061, MainFragmentRemoval
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0062, NeutralFragmentRemoval
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0063, PhotonFragmentRemoval
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0064, ClusterPreparation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0065, ForceSplitTrackAssociations
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0066, PfoCreation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0067, PfoPreparation
[ VERBOSE "MyDDHCALPandora"] > Running Algorithm: Alg0068, VisualMonitoring
```

Pandora Visual Monitoring

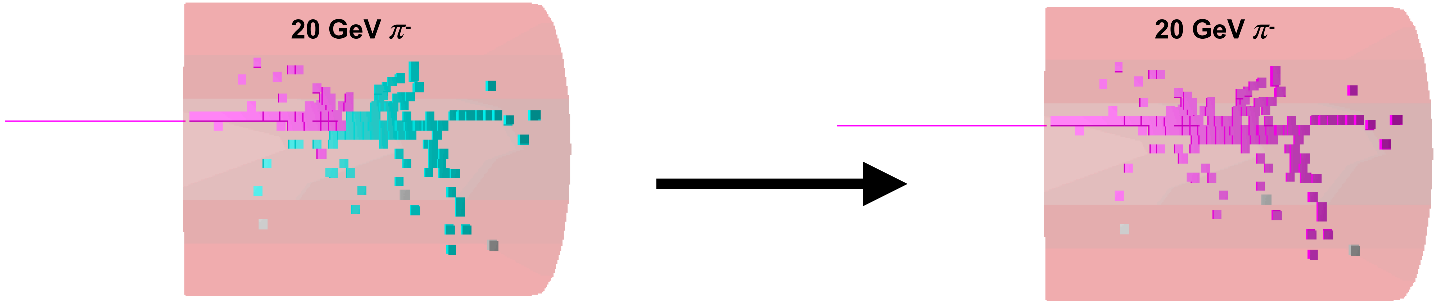
Hits, Clusters & PFOs



Magenta: Charged Hadron
Cyan: Neutral Hadron
Yellow: Photon
Grey: Unclustered Hits

- Cylinder: Existing HCAL end-cap class used for our setup
- Pandora visual monitoring displaying hits, clusters, tracks and PFOs at different reconstruction steps
  - ➡ Great tool to precisely track down technical problems and problematic events

Solved: Non working Track-Cluster association for few events

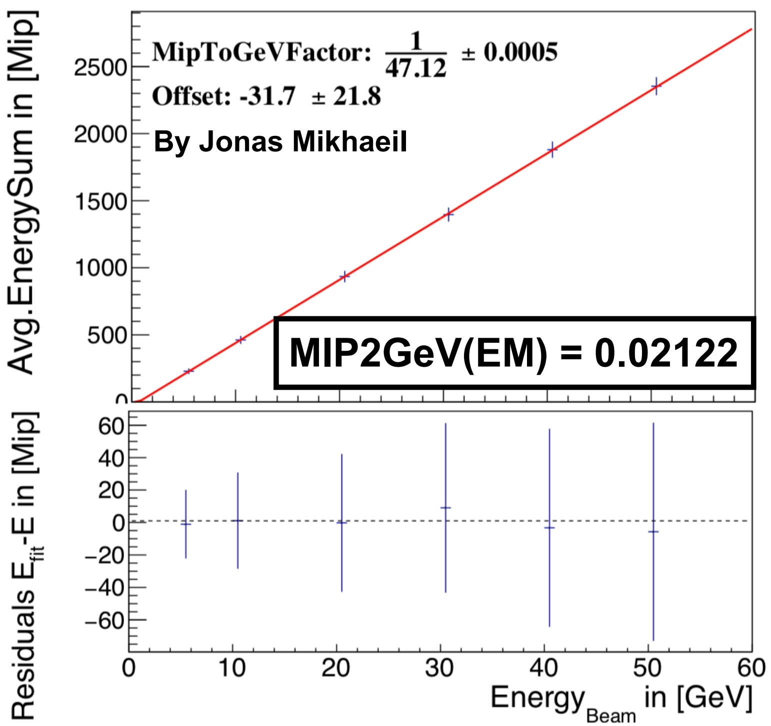


MIP to GeV Conversion

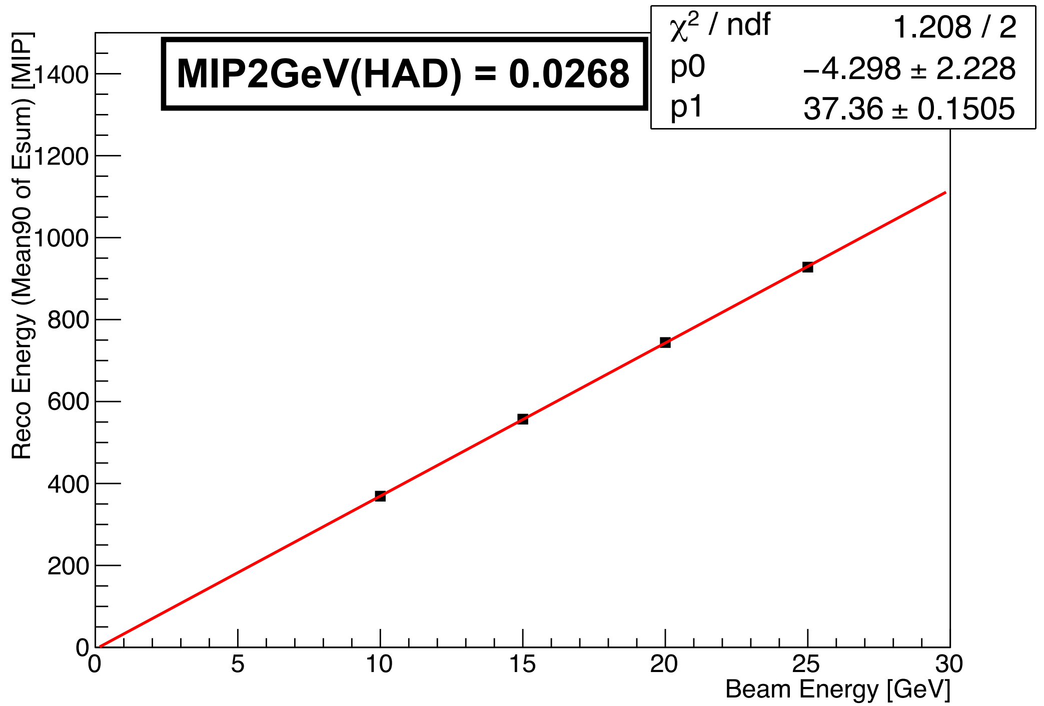
Calibration to EM and HAD Scale

- PandoraPFA framework requires energy depositions in units of GeV
  - ➡ MIP to GeV calibration done on MC samples for EM and HAD energy scale
    - ➡ Extract slope of beam energy vs calorimeter MIP response scan

EM Response Determination (e-)



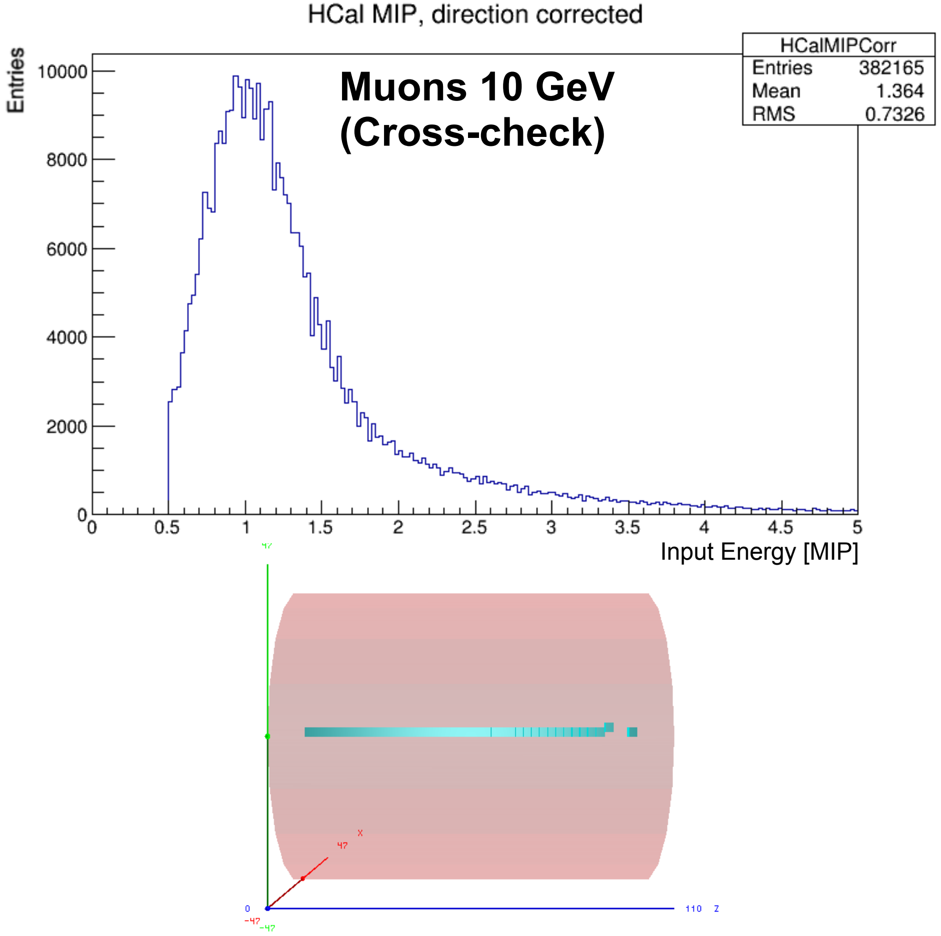
HAD Response Determination (K0L)

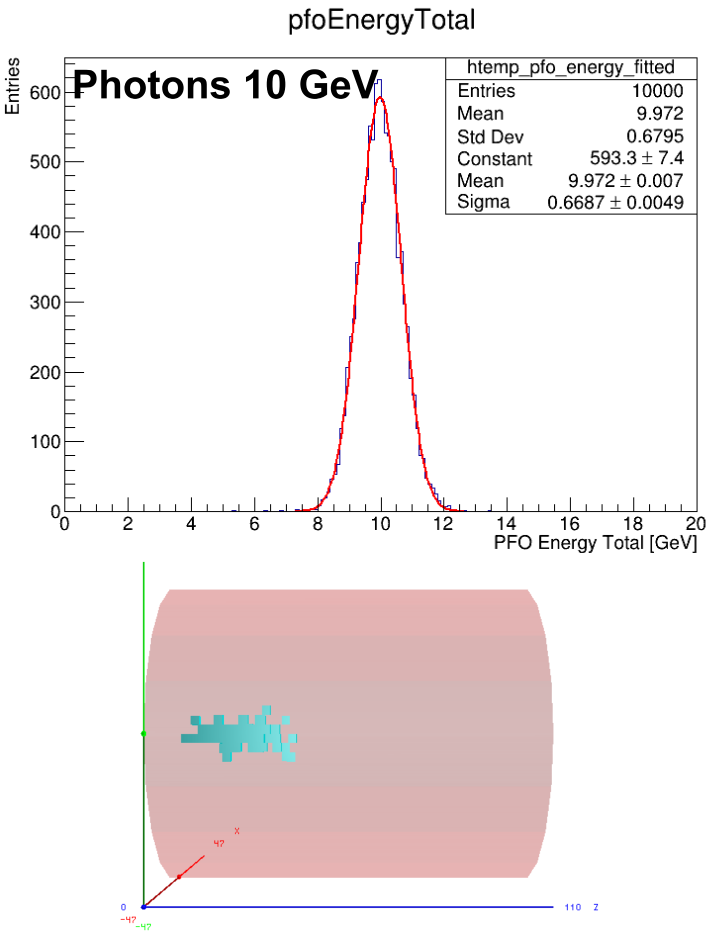


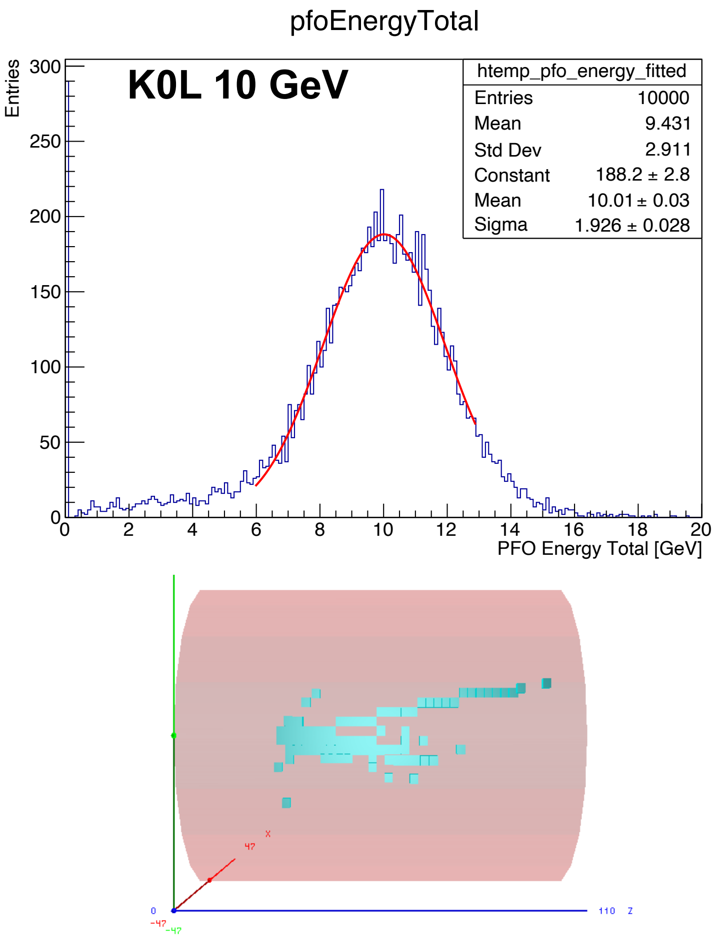
Pandora Energy Calibration

MC Muons, Photons, K0L

Note: Without tracks and ECAL everything classified as neutral hadrons at this step





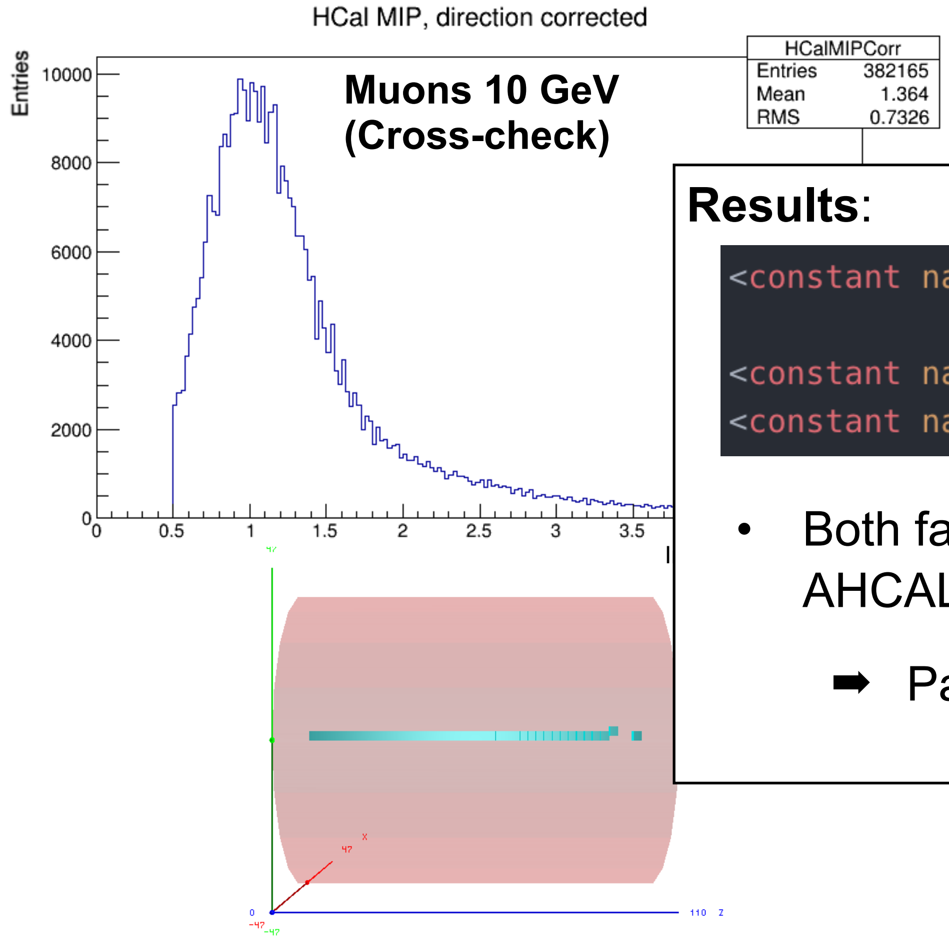


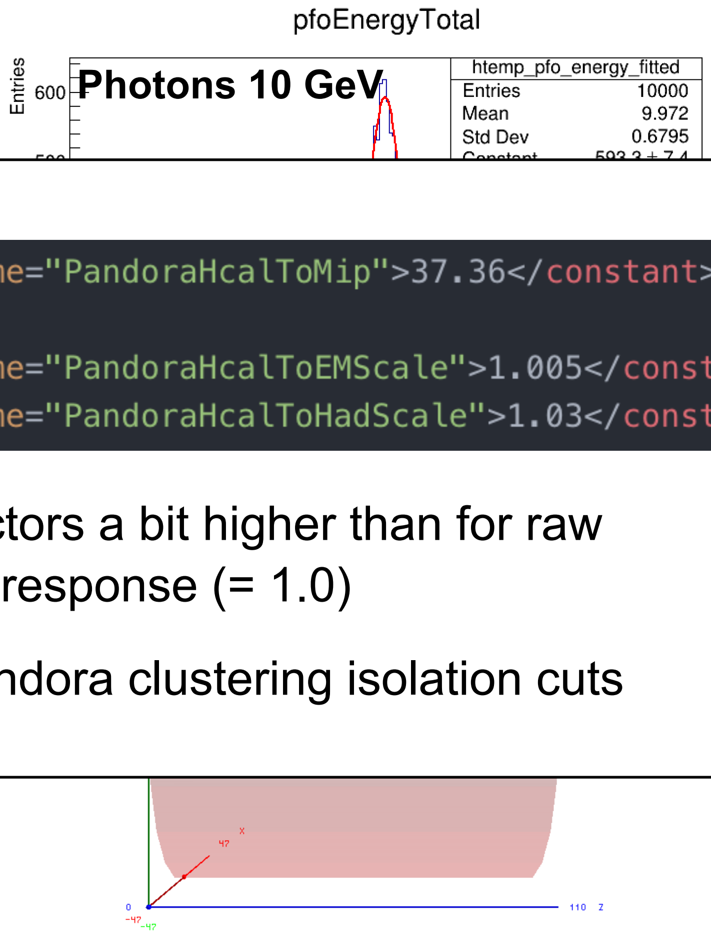
- **Muons**: AHCAL energy GeV -> MIP with negligible angle correction since straight TB tracks
- **Photons and K0L's**: Used to determine EM and HAD response, PFO energy tuned to peak at 10 GeV

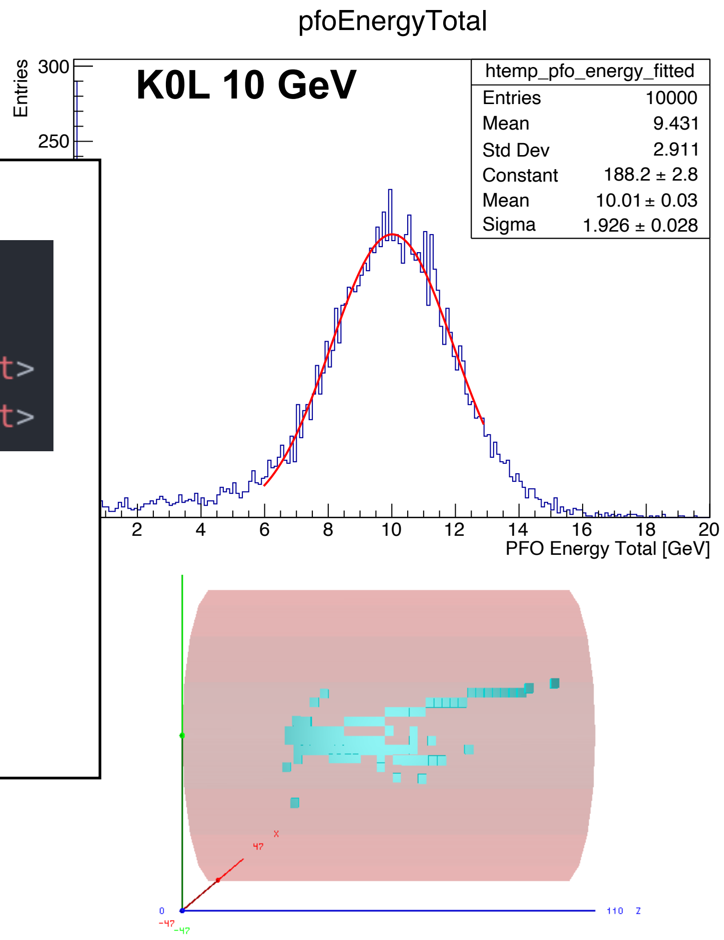
Pandora Energy Calibration

MC Muons, Photons, K0L

Note: Without tracks and ECAL everything classified as neutral hadrons at this step







Results:

```
<constant name="PandoraHcalToMip">37.36</constant>

<constant name="PandoraHcalToEMScale">1.005</constant>
<constant name="PandoraHcalToHadScale">1.03</constant>
```

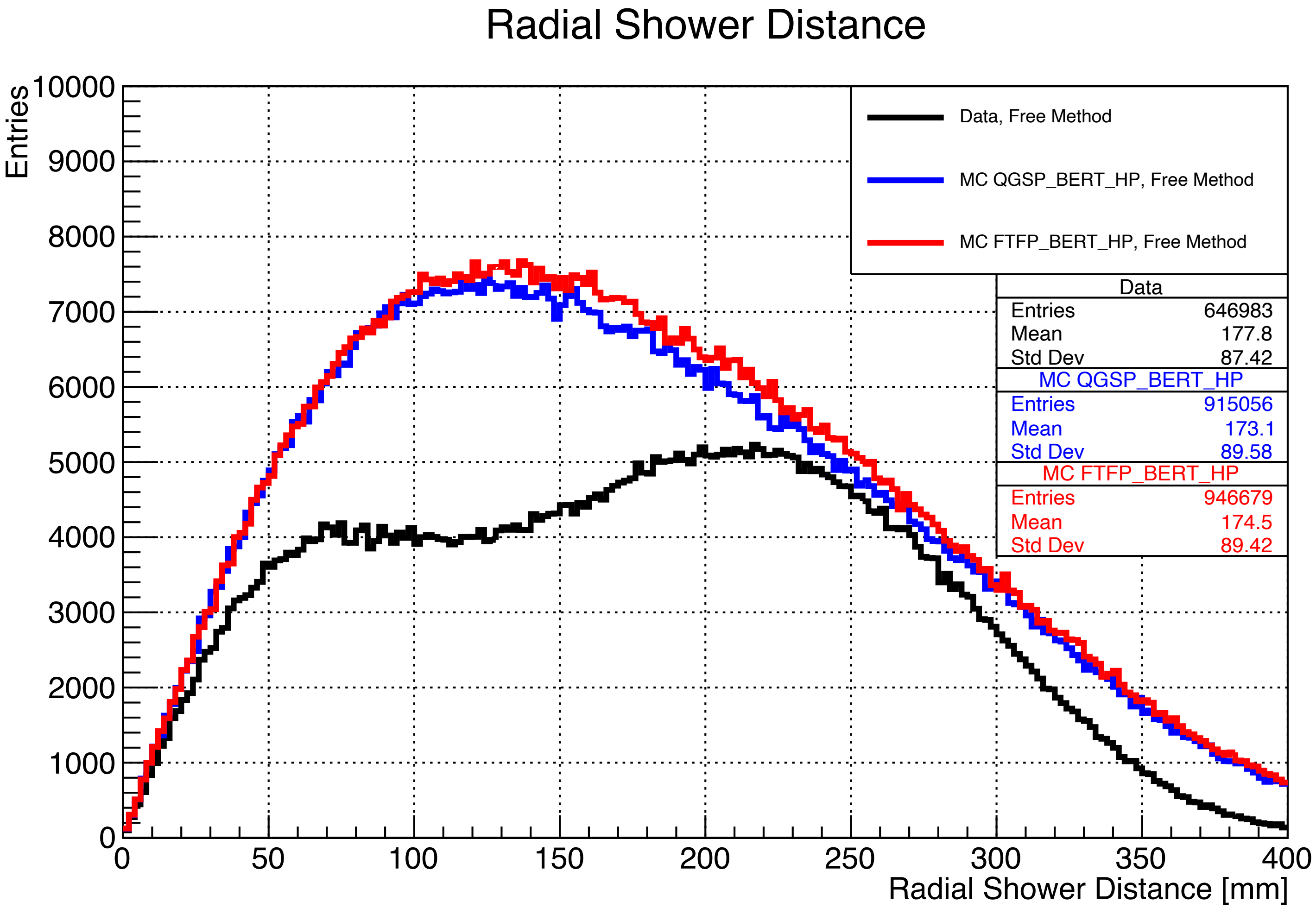
- Both factors a bit higher than for raw AHCAL response (= 1.0)
  - ➡ Pandora clustering isolation cuts

- **Muons**: AHCAL energy GeV -> MIP with negligible angle correction since straight TB tracks
- **Photons and K0L's**: Used to determine EM and HAD response, PFO energy tuned to peak at 10 GeV

Results

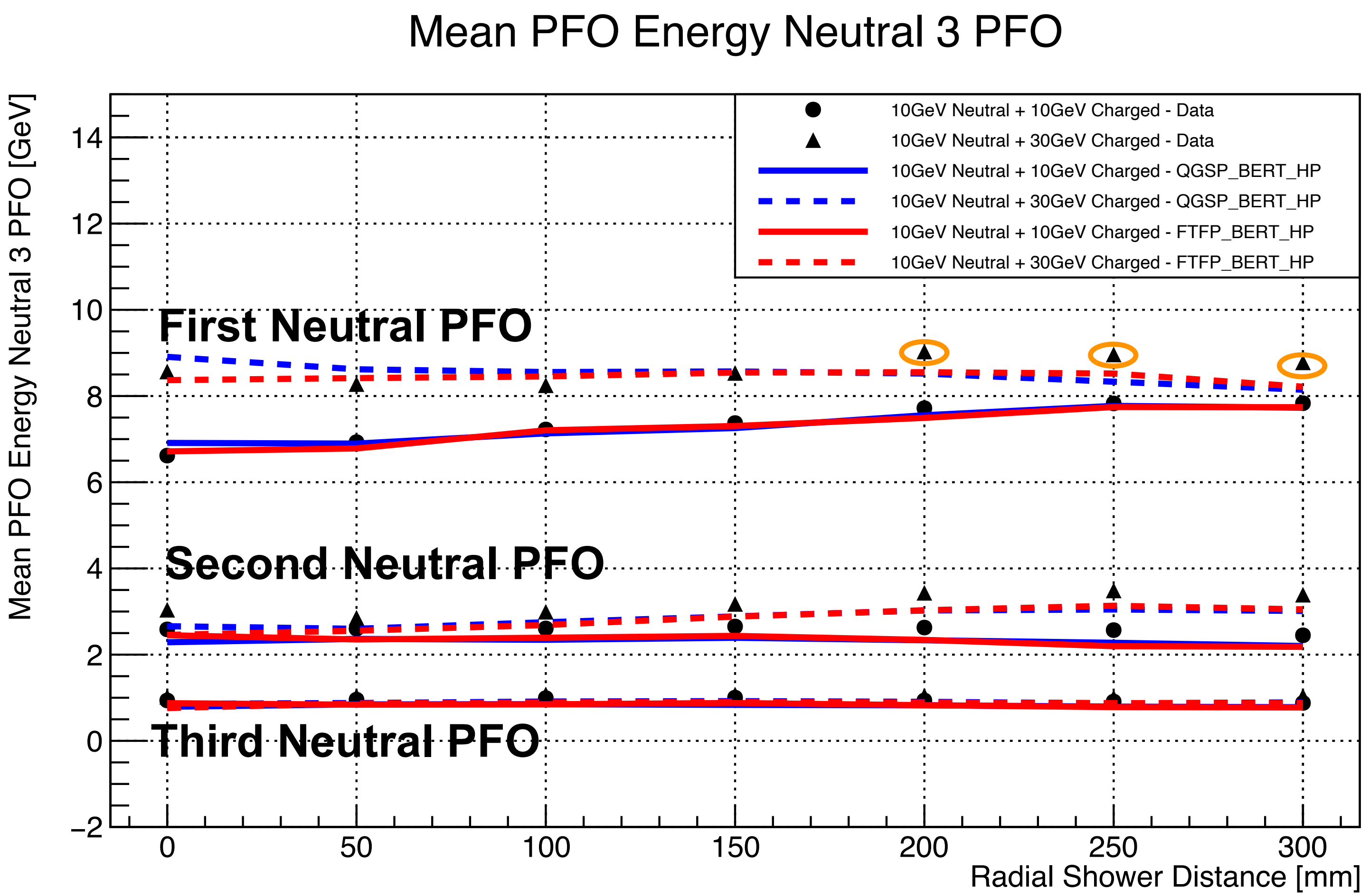
Radial Distance Distribution

Data & MC Samples



Mean Neutral PFO Energy - Events with 3 Neutral PFOs

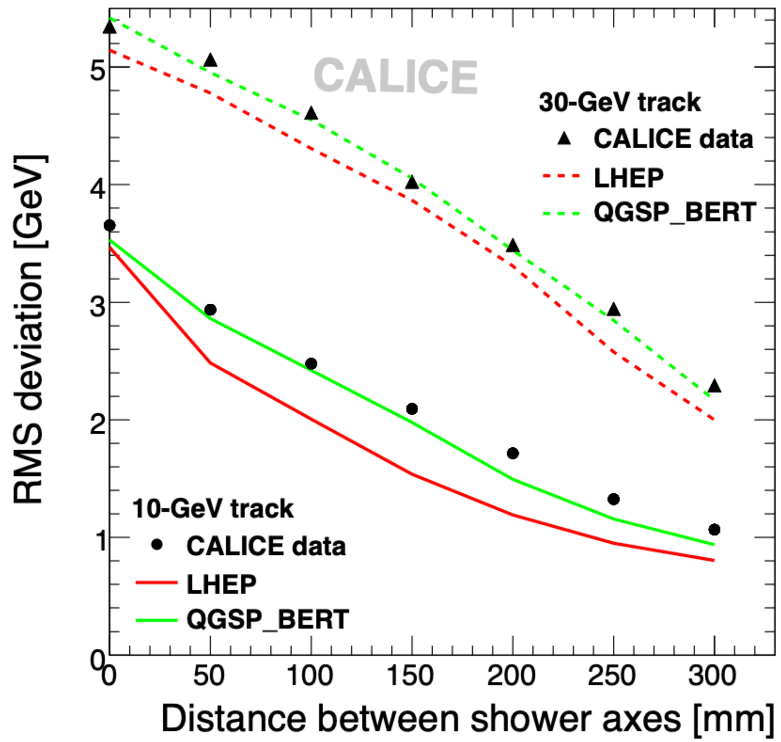
Mean Energy of Neutral PFO's

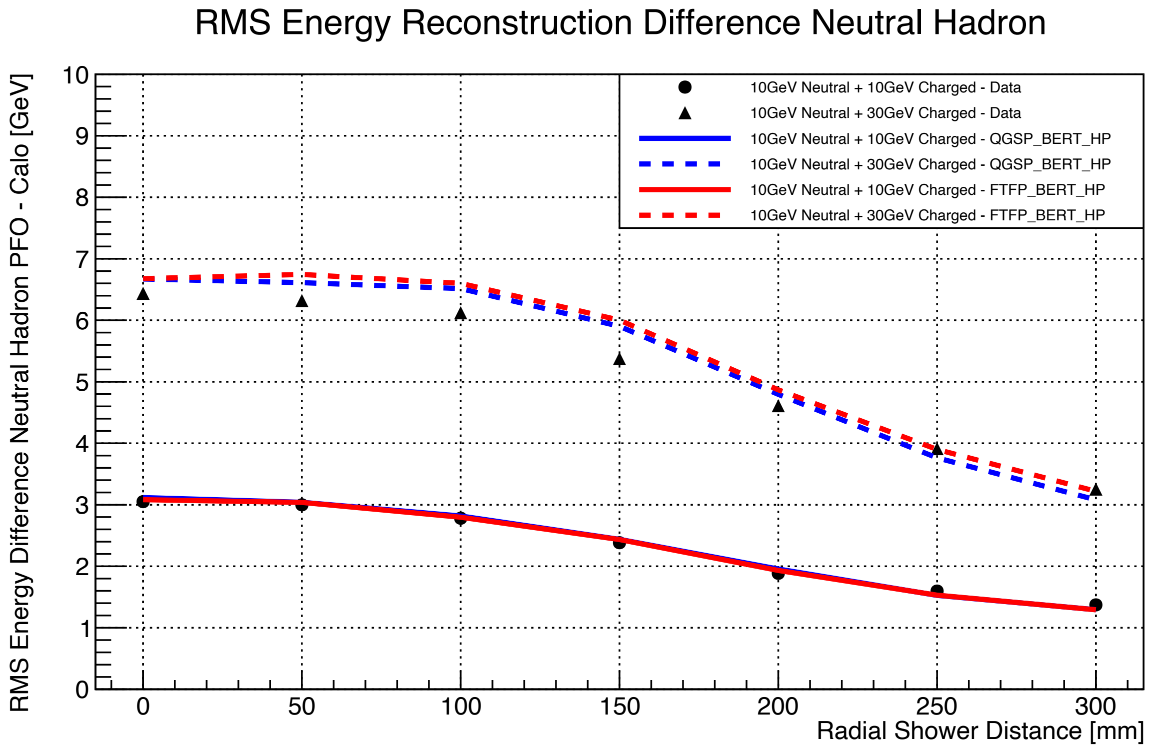


- Reasonable data to MC agreement within 5%
- Same trends as for 2 neutral PFO case for first and second PFO
- Third PFO with even smaller fraction of energy (10%) which is constant for all scenarios
- Careful: Total mean energy higher than 10 GeV, since biased selection of events with trend to confusion

PFO Energy - Calorimeter Energy Neutral Hadron - RMS

First CALICE PFA Paper vs. Latest Studies

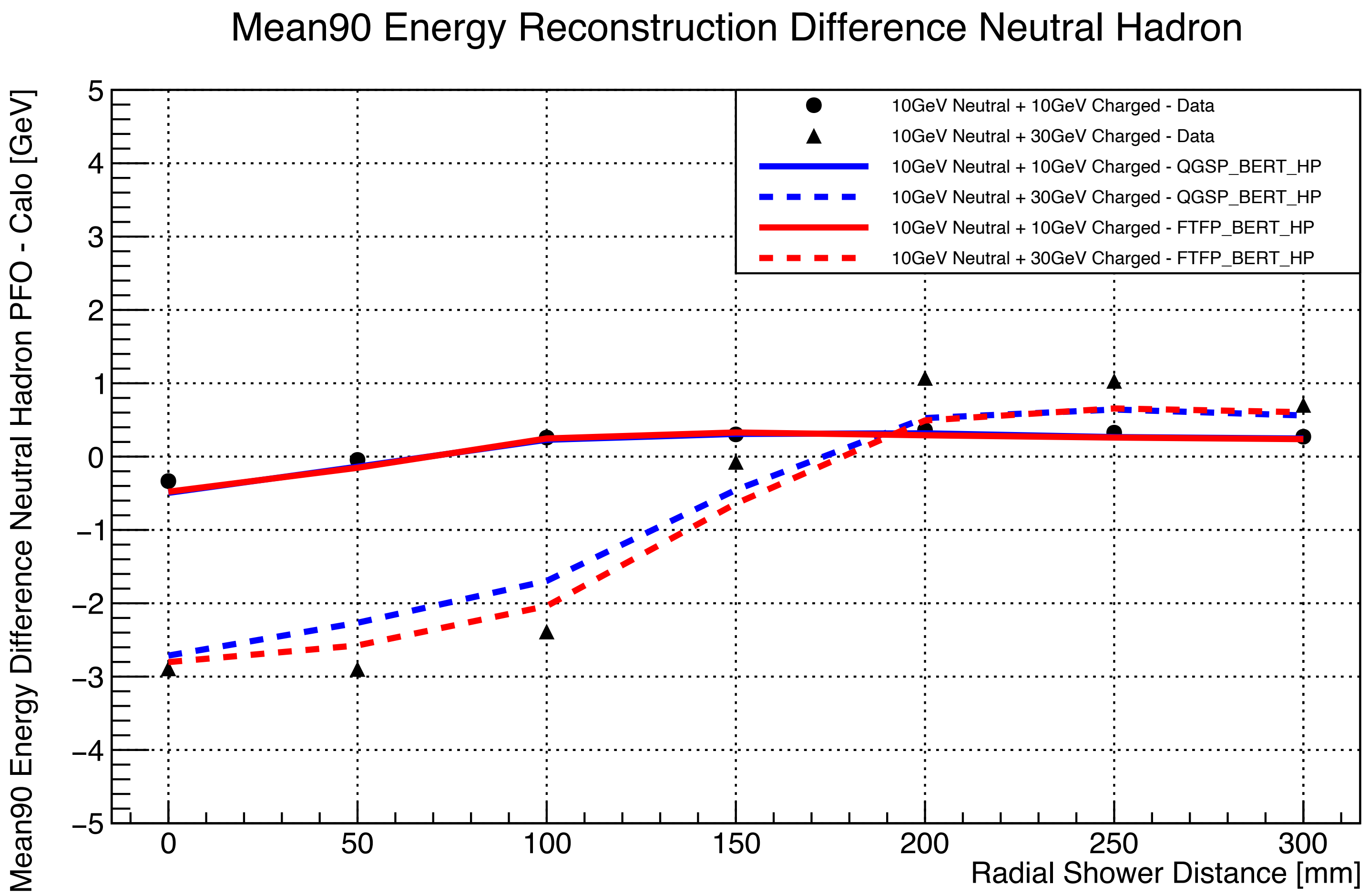




- Falling trend (slower) for growing distance to 10GeV and 30GeV charged hadrons
  - ➡ Vicinity of 10GeV charged hadron: Excellent data/MC agreement
  - ➡ Vicinity of 30GeV charged hadron: In general RMS larger, performance on data slightly better

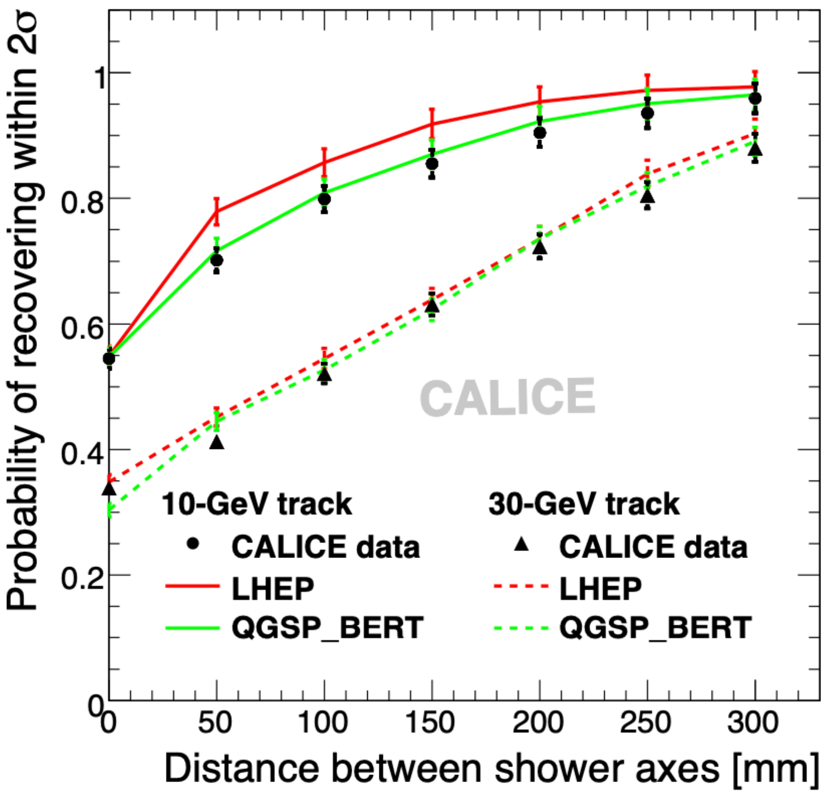
PFO Energy - Calorimeter Energy Neutral Hadron - Mean90

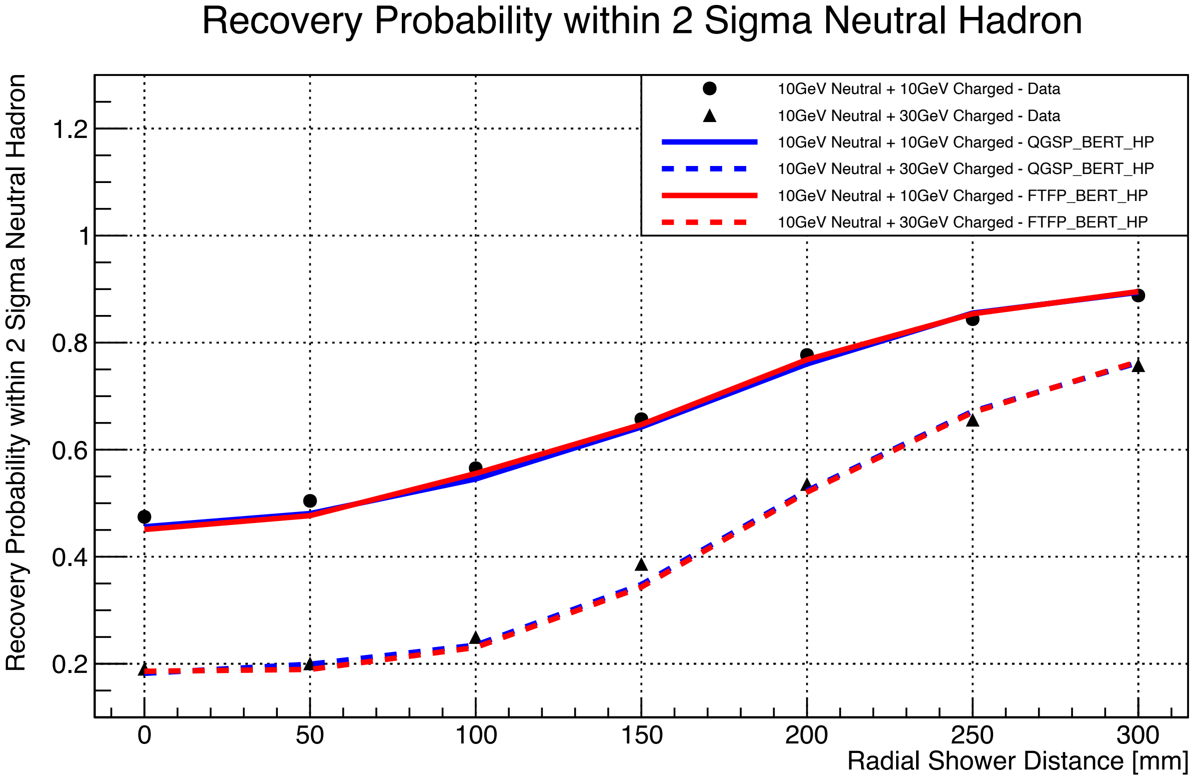
First CALICE PFA Paper vs. Latest Studies



Neutral Hadron Recovery Probability 2sigma

First CALICE PFA Paper vs. Latest Studies





- Definition: Probability that PandoraPFA recovers neutral hadron energy within **2 sigma** (sigma = width of neutral hadron input energy sum)
- Rising trend as for old studies up to 70-90%, but slower growing especially for 30GeV hadron close-by
  - ➡ Suspicion: Low distances very tricky without ECAL hits before