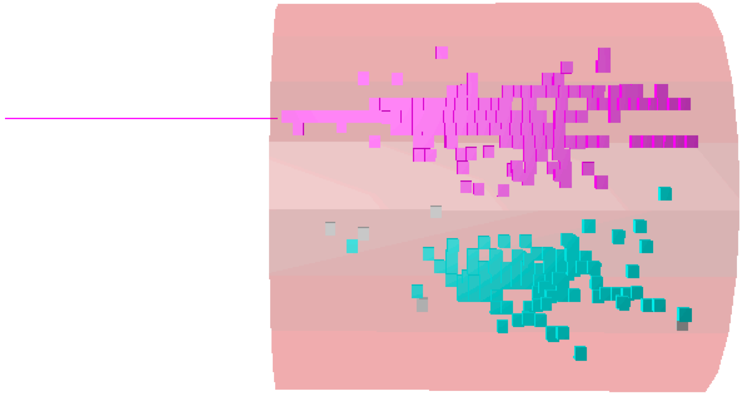
PandoraPFA Studies on AHCAL 2018 Data

Update

PandoraPFA Visual Monitor



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

Daniel Heuchel (DESY)
daniel.heuchel@desy.de
CALICE Analysis Meeting
10th December 2020

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

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES
UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386
HGSFP
CALICE
DESY.

Outline

For this Talk

- Concept of Particle Flow & Motivation for Studies

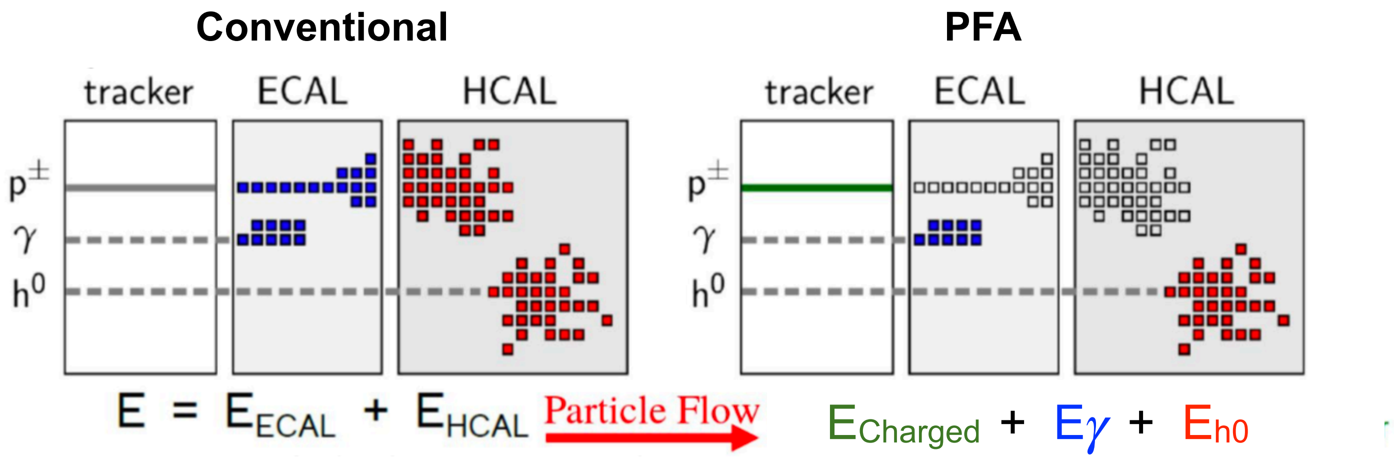
- Quantitative Studies on PandoraPFA **Single** Particle Reconstruction
- Event Overlay Procedure
- First Quantitative Studies on PandoraPFA **Two** Particle Reconstruction

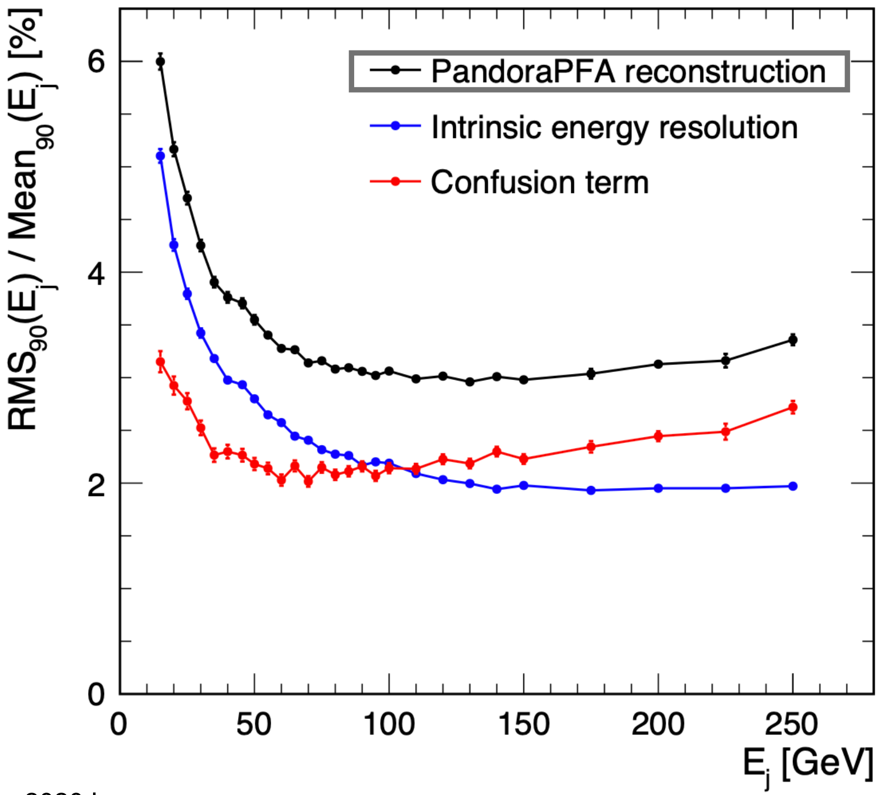
- Summary & Outlook

Particle Flow Approach

Reaching High Precision

- Goal at the ILC: 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)





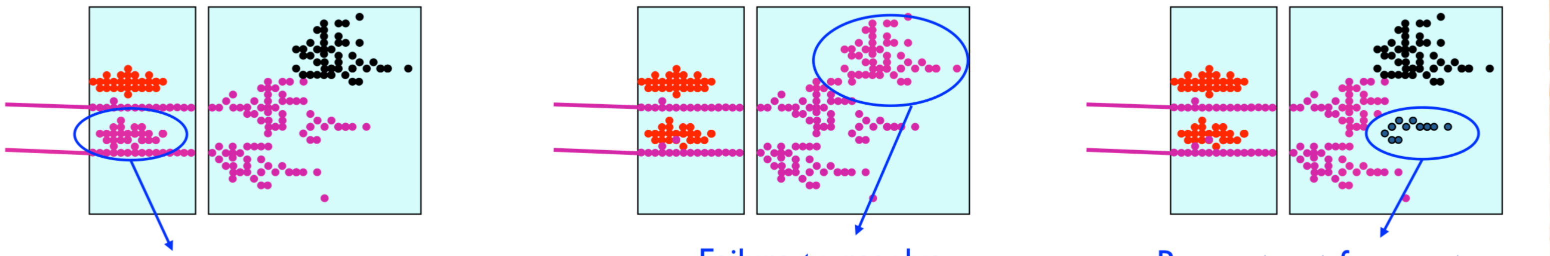
arXiv:1705.10363

The Confusion Term

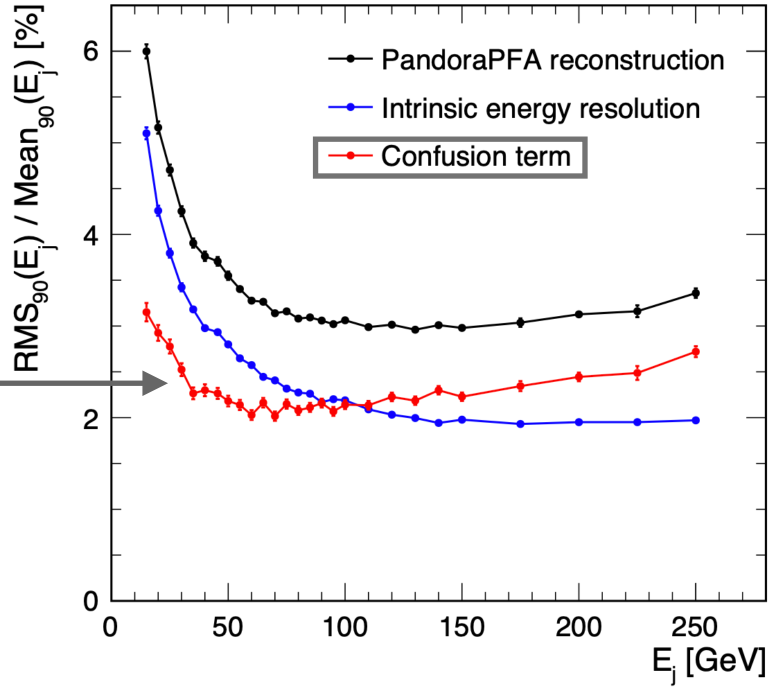
The Limit of PFA Reconstruction

- **Topologically or energetically confusing** scenarios might appear in specific events limiting jet energy resolution:

Types of confusion



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



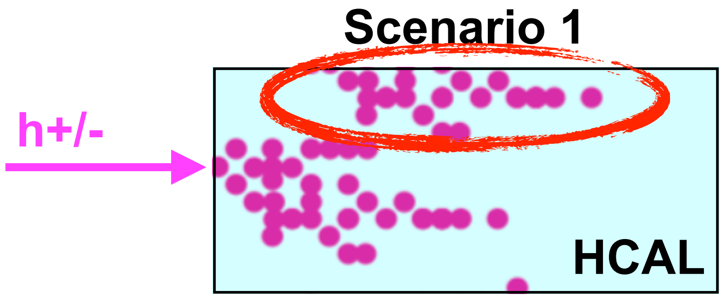
arXiv:1705.10363

Idea and Goals of Studies

PandoraPFA on AHCAL 2018 Prototype Data

- **Evaluate PandoraPFA on AHCAL 2018 beam test data:** Study of **single and two particle reconstruction performance** with regard to **confusion**
  - ➡ Different conditions (energies, transversal/longitudinal separation, amount of leakage, etc.)

- **Scenario 1**: **Single Particle Event (Charged Hadron)**
  - ➡ Confusion: Fragment of charged hadron shower reconstructed as separate neutral hadron
  - ➡ Sensitive to double counted energy

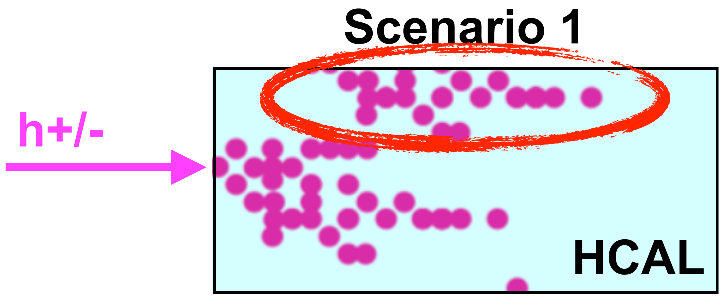


Idea and Goals of Studies

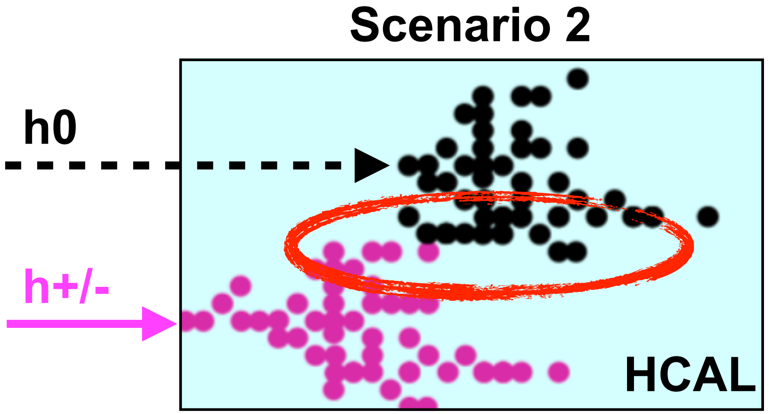
PandoraPFA on AHCAL 2018 Prototype Data

- **Evaluate PandoraPFA on AHCAL 2018 beam test data:** Study of **single and two particle reconstruction performance** with regard to **confusion**
  - ➡ Different conditions (energies, transversal/longitudinal separation, amount of leakage, etc.)

- **Scenario 1**: **Single Particle Event (Charged Hadron)**
  - ➡ Confusion: Fragment of charged hadron shower reconstructed as separate neutral hadron
    - ➡ Sensitive to double counted energy



- **Scenario 2**: **Two Particle Event (Charged + Neutral Hadron)**
  - ➡ Confusion: Failure to resolve neutral hadron; Neutral hadron hits fully or partly added to charged hadron cluster
    - ➡ Sensitive to missing energy

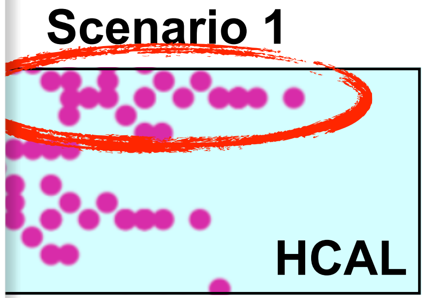


Idea and Goals of Studies

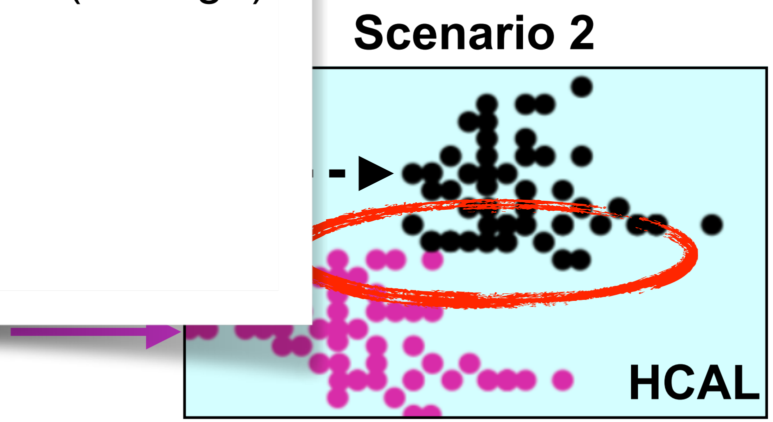
PandoraPFA on AHCAL 2018 Prototype Data

- **Evaluate PandoraPFA on AHCAL 2018 beam test data:** Study of **single and two particle reconstruct**
  - ➡ Diffe … kage, etc.)

- **Scenario 1:**
  - ➡ Confusi … separat
  - ➡ Sens



- **Scenario 2:**
  - ➡ Confusi … fully or
  - ➡ Sensitive to missing energy



A comparable study was done for the 2007 prototype resulting in a CALICE PFA paper in 2011 (https://arxiv.org/abs/1105.3417).

Why do it again on AHCAL 2018 prototype data?

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

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**

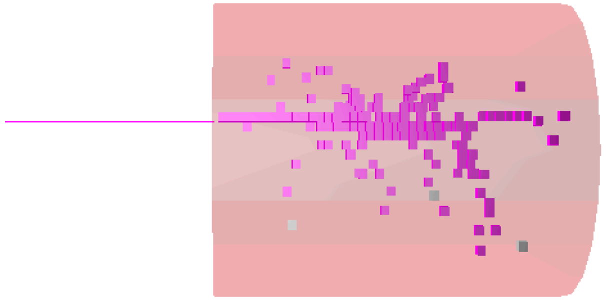
Selected events with overlaid charged and pseudo-neutral hadron → **PandoraPFA** → **Scenario 2: Two-particle separation studies**

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

Data & MC Pion Samples Overview

PandoraPFA Single Particle Reconstruction

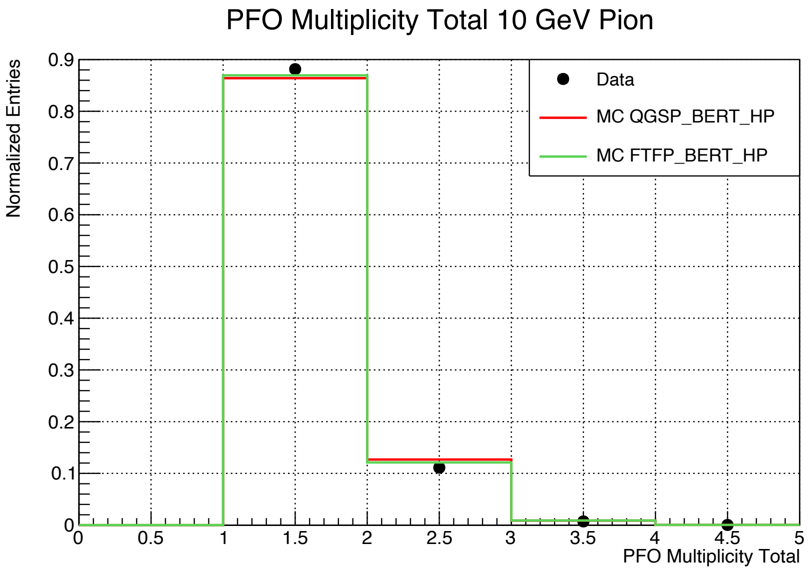
- **Charged pions** events **(10, 20, 40, 60, 80, 120, 160, 200 GeV) with fixed track momentum in PandoraPFA**
  - ➡ Data: **June Beam Test 2018 @ SPS CERN**
  - ➡ MC: **GEANT4 v.10.03, QGSP_BERT_HP & FTFP_BERT_HP**

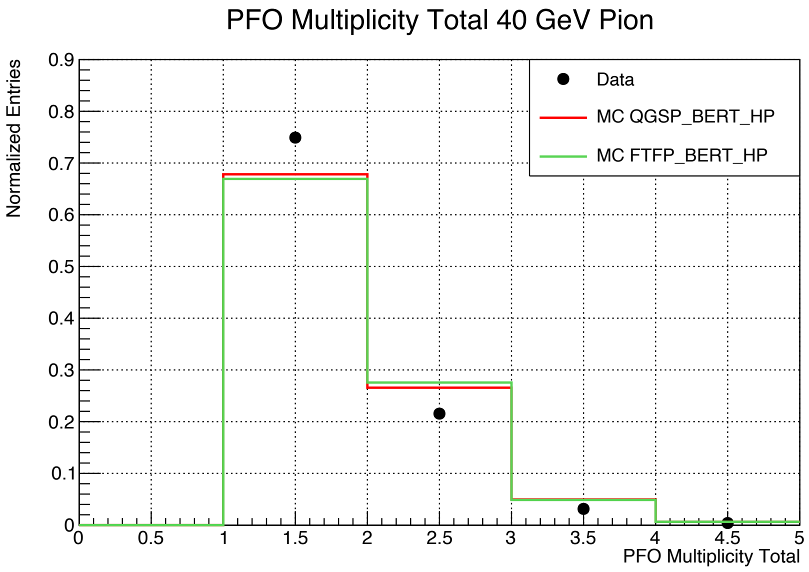


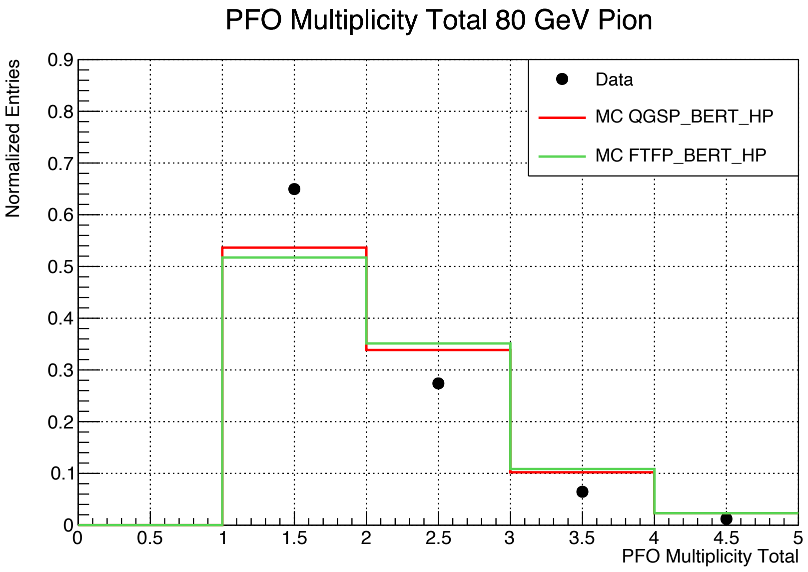
- Applied BDT-PID for hadrons (remove beam contamination)
- Event selection:
  - ➡ At least one hit in layer 1 or 2 or 3 & corresponding track - hit match (for proper track-cluster assignment)
  - ➡ Track to detector crack rejection +- 2mm
  - ➡ Shower start layer < 20 (reject leakage events)
  - ➡ Rejection of remaining events with complete failure of track-cluster association (<1%)
    - ➡ Technical reason within PandoraPFA algorithms: No ECAL before - missing first track association

PFO Multiplicity - How many Reconstructed Particles?

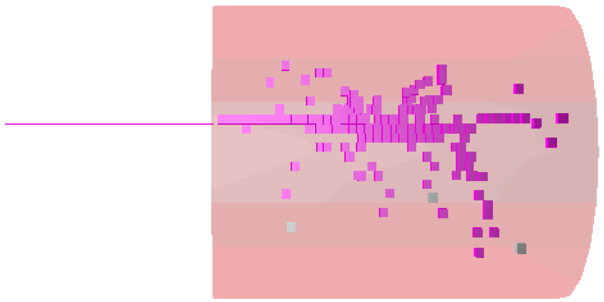
10, 40 and 80 GeV Data and MC



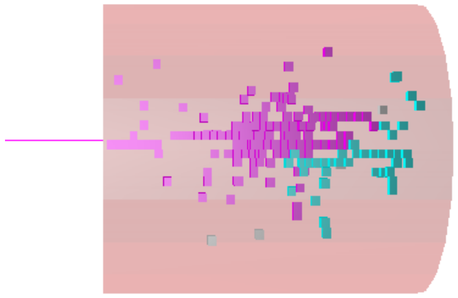




- Excellent agreement of data and MC for 10 GeV
- Mean multiplicity growing with beam energy for data and MC
- Discrepancy between data and MC increasing with growing energy, **performance on data better**



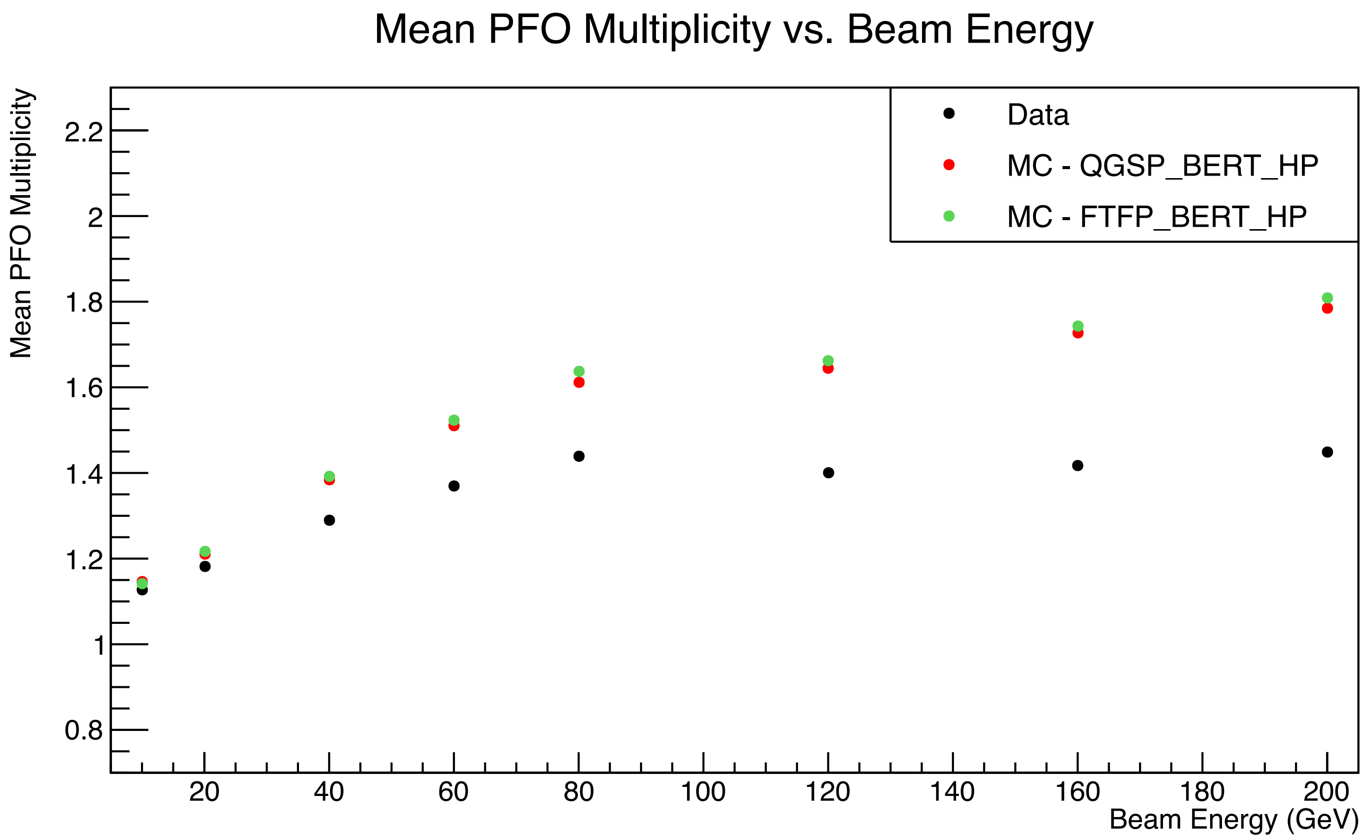
1 PFO



2 PFO

Mean PFO Multiplicity vs. Beam Energy

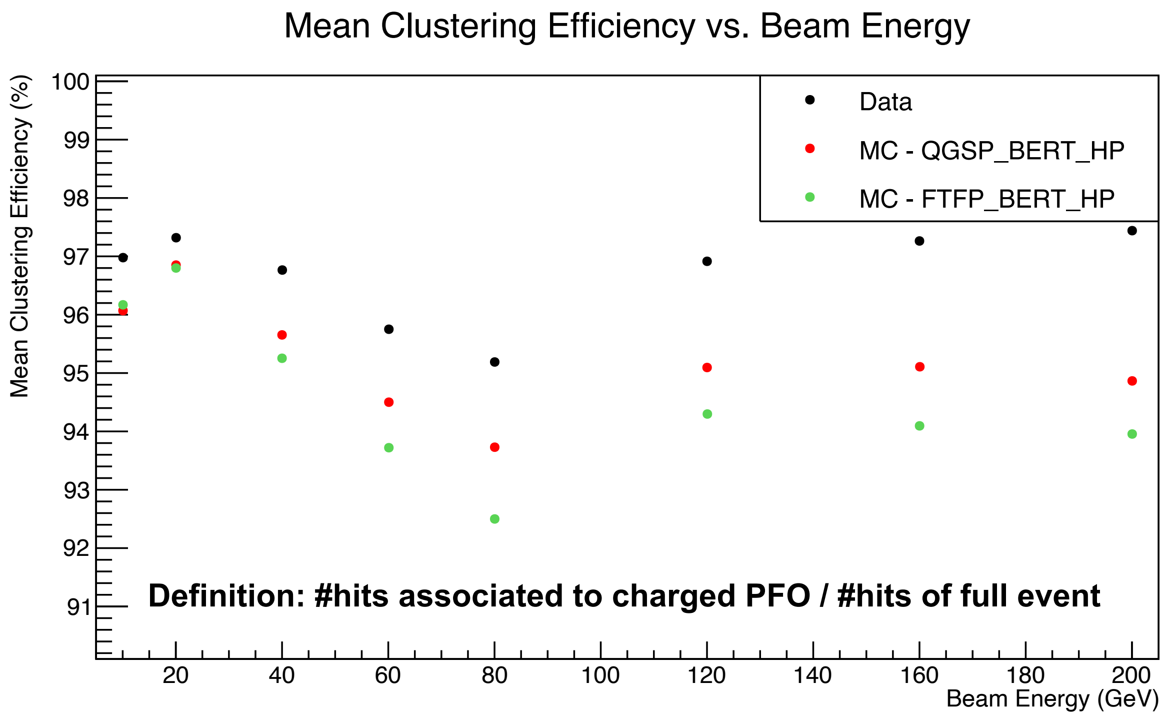
Data and MC

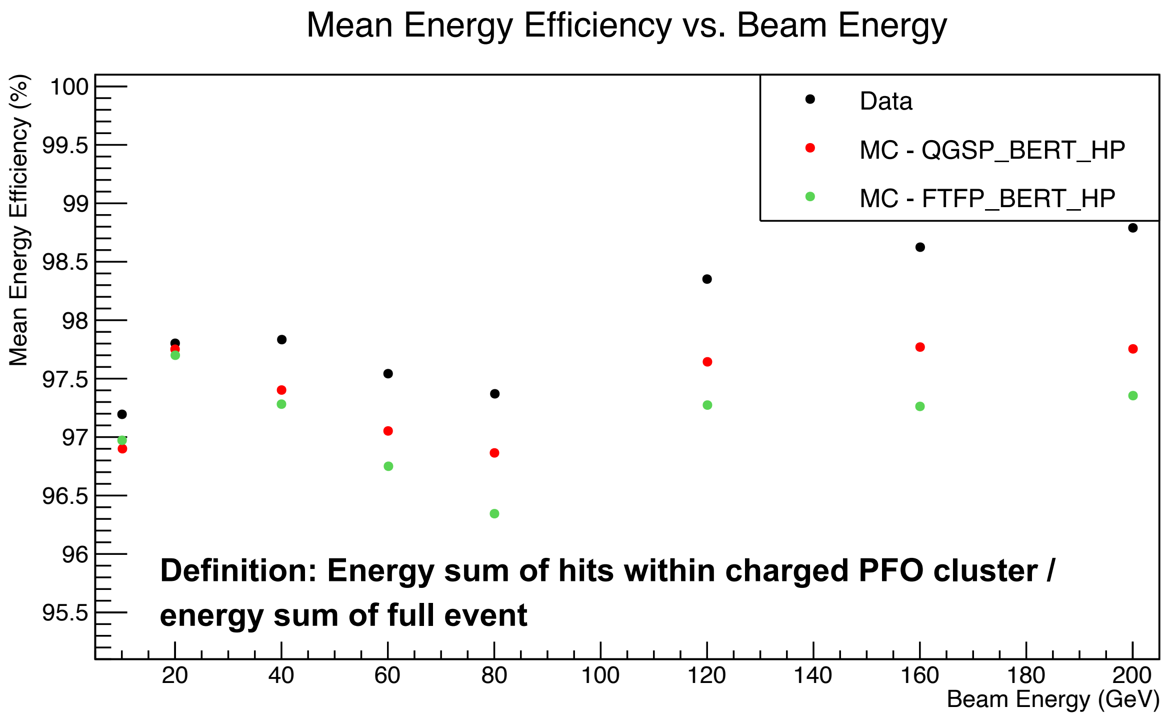


- Growing with beam energy up to 80 GeV reflecting growing confusion
- At 120 GeV reaching a plateau
  - ➡ Dominated by remaining leakage (too soft shower start selection)
    - ➡ „Neutral fragments" more likely to be clustered into charged PFO again to compensate for leakage
- Growing stronger for MC than for data - both physics lists behave similar

Mean Clustering & Energy Efficiency vs. Beam Energy

Data and MC

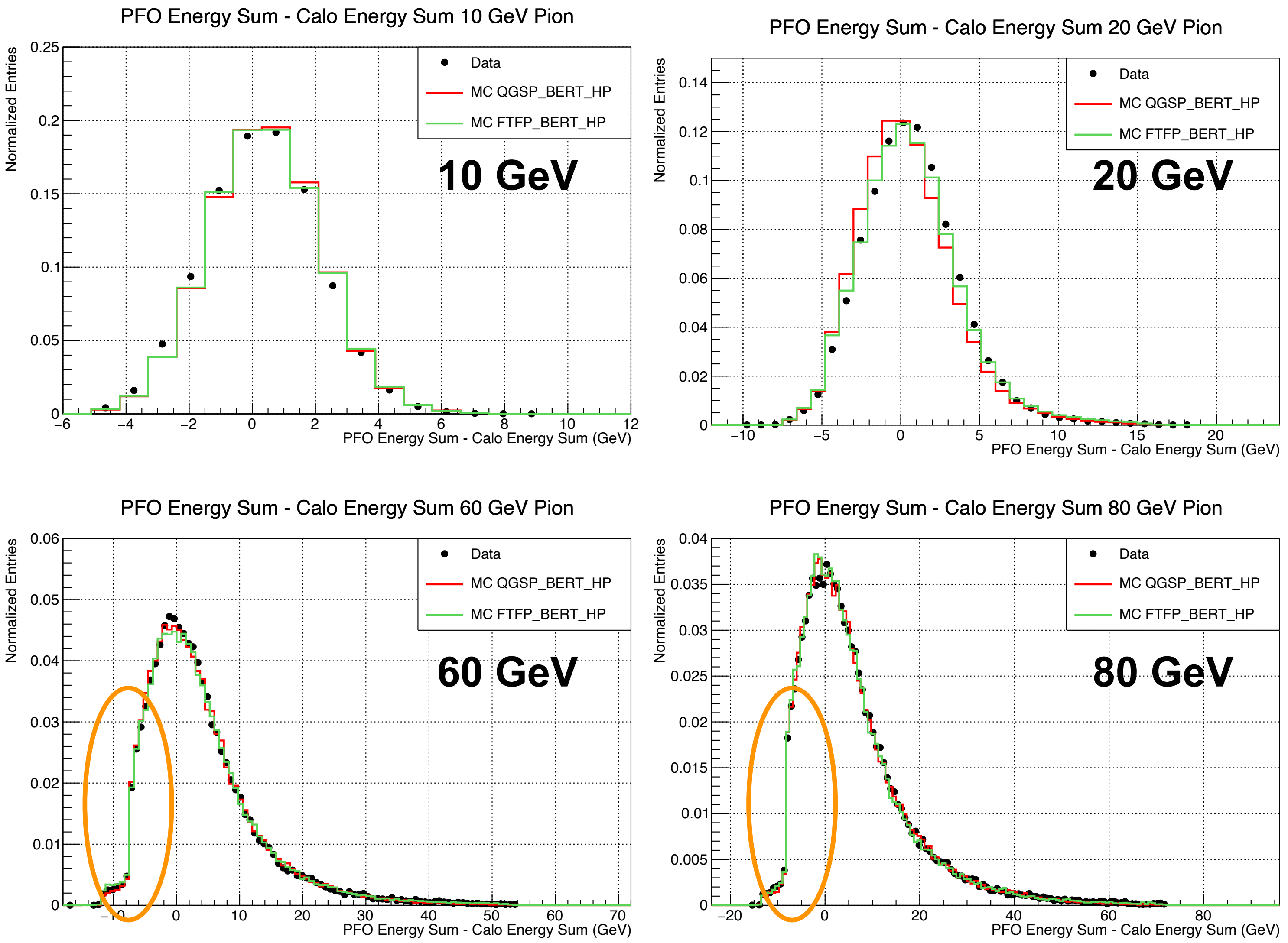




- Both efficiencies on average quite high and comparable with results of ArborPFA studies on SDHCAL data
- Consistent with investigated PFO multiplicity trends
  - ➡ Data best and here: QGSP_BERT_HP better than FTFP_BERT_HP

PFO Energy Sum - Calorimeter Energy Sum

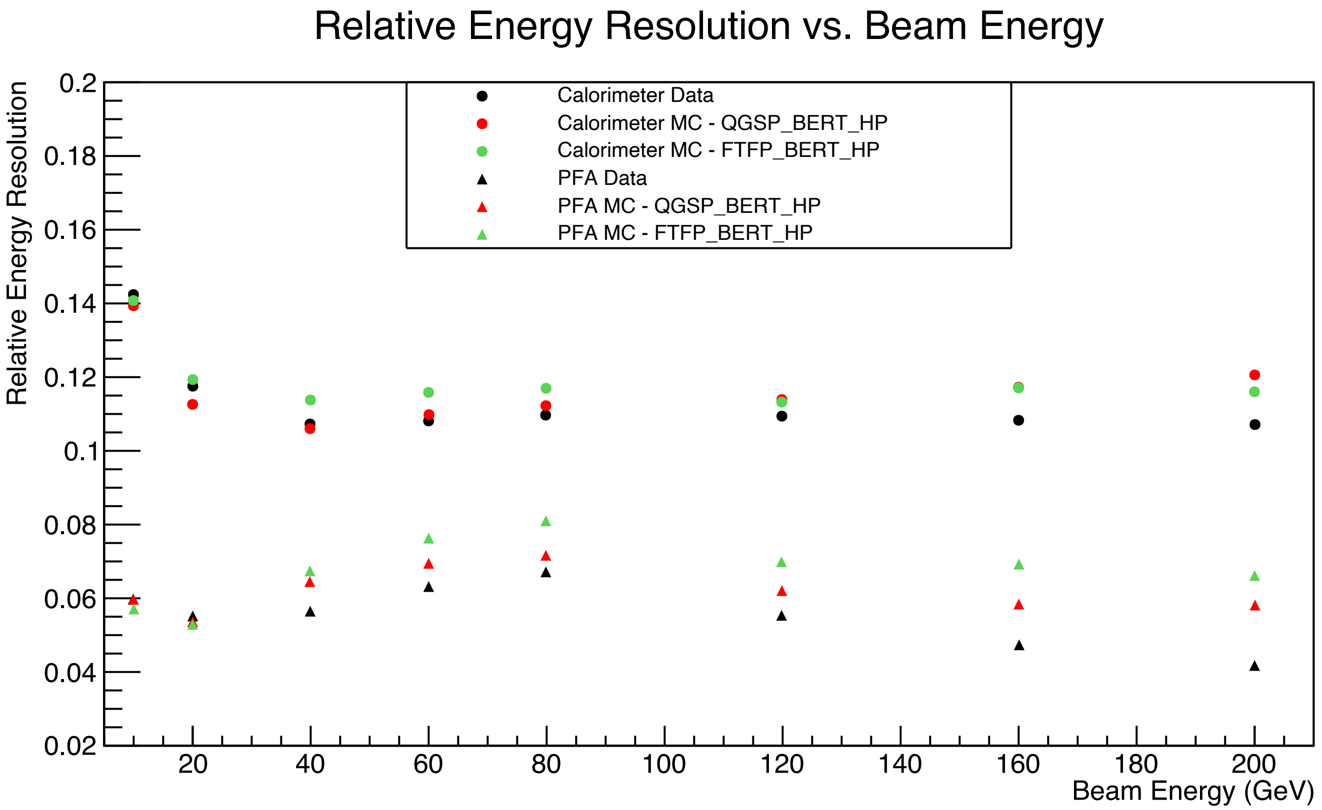
10, 20, 60 & 80 GeV Data and MC



- Overall good agreement between data and MC
- Sharp edge at **-8.8 GeV**
  - ➡ Internal PandoraPFA cut?
  - ➡ John Marshall had few ideas what it might be within PandoraPFA code: Probably related to internally assumed hadronic energy resolution of 60%/sqrt(E)

Relative Energy Resolution vs. Beam Energy

Data and MC



Classical: Mean90 and RMS90

PFA: Full Mean and RMS

- **PFA energy resolution factor of 2 better than classical energy resolution**
- PFA reconstruction performance on data best, at low energies same level as MC

➡ Up to 80 GeV: Growing confusion - degrading of energy resolution

➡ From 120 GeV onwards: Artificial improvement of energy resolution due to remaining leakage

Overview

Sample Preparation & Analysis Strategy

Thanks to Linghui for great shared work and synchronisation effort!

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

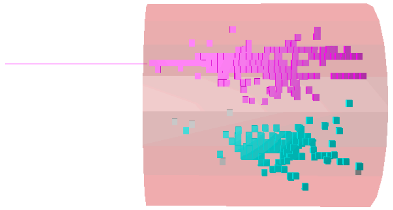
Selected charged pion events ↓ **PandoraPFA** → **Scenario 1: Single particle reconstruction studies**

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

Overlay Processor

Procedure & Requirements



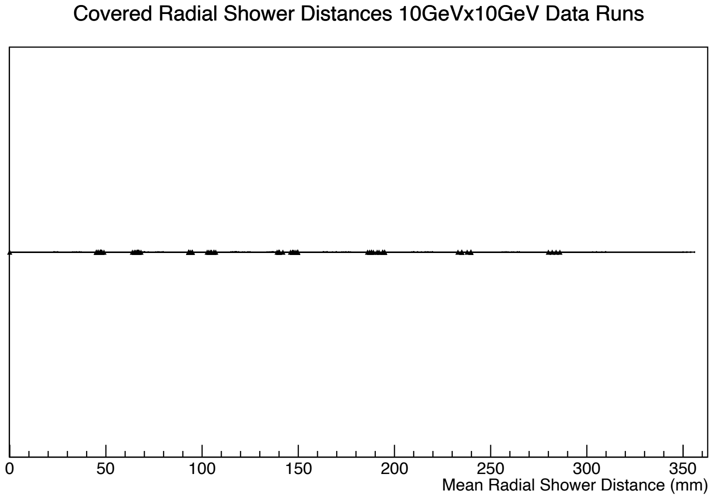
Magenta: Charged Hadron
Cyan: Neutral Hadron
Grey: Unclustered Hits

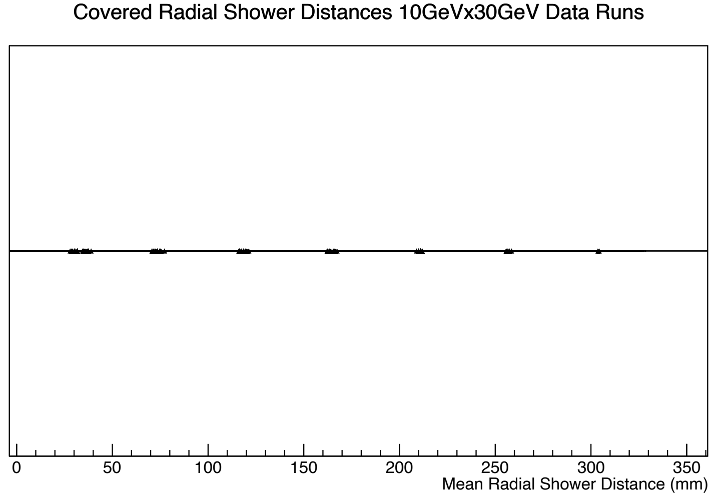
- Initial check: Full radial shower distance range between 0-350mm covered with combinations of test beam data runs at different positions with sufficient statistics for 10 and 30 GeV pions

- Technical goal: Overlay Event A (pseudo neutral hadron) with Event B (charged hadron + track)

- Requirements:
  - ➡ Proper flagging of output hits and saving of individual output collection A, B and A+B
  - ➡ Proper handling of MIP threshold - Apply 0.5 MIP threshold only on overlaid hits
  - ➡ Radial shower distance selection according to cogX,Y of shower pairs - currently range acceptance: +-25mm
  - ➡ All created double particle events unique

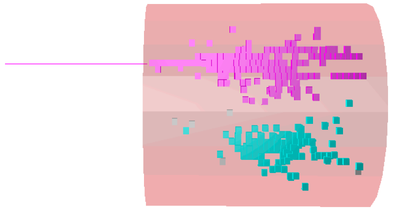
Illustration of Run Combinatorics





Overlay Processor

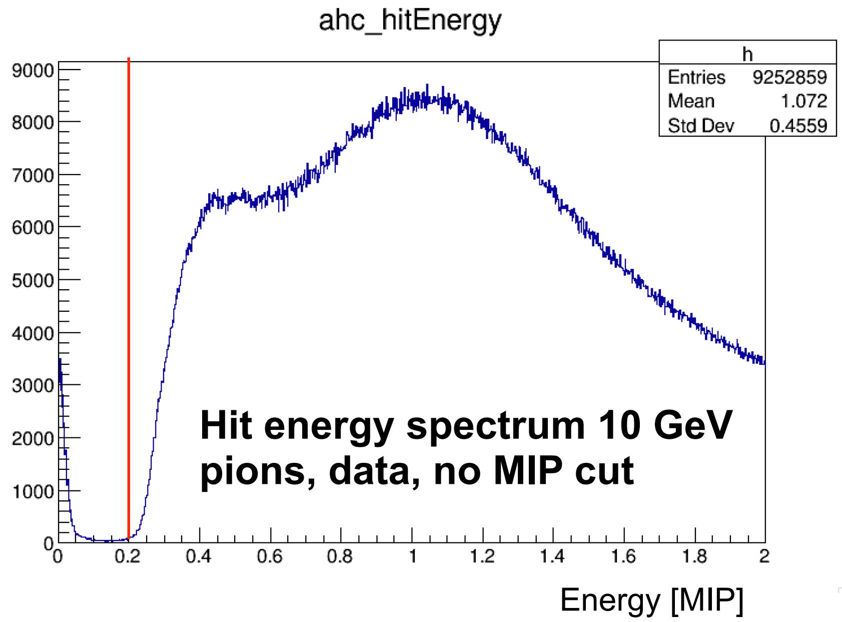
Status and Validation



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

- What is the lowest MIP cut to use for data and MC before overlay?
  - ➡ Hardware MIP cut ~0.2 MIP seems to be a good choice
- Samples used for validation of overlay processor (MC and data):
  - ➡ 10 GeV neutral & 10/30 GeV charged pion (50mm & 200mm distance)
- Results:
  - Overlaid event yield (of initial neutral events) > 94%
  - Fraction of cut low energy hits (lower 0.5 MIP threshold) after/before overlay: > 95%
    - ➡ Most of low energy hits are cut after overlay
  - Mean #new hits after overlay (reaching 0.5 MIP threshold) < 0.25 hits
    - ➡ Negligible for all scenarios

➡ Processor implemented and working well for two particle events



Overview

Sample Preparation & Analysis Strategy

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**

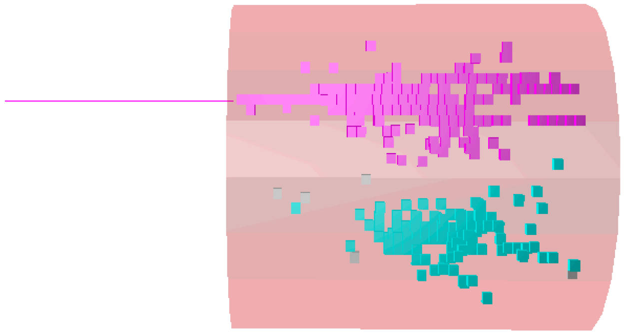
Selected events with overlaid charged and pseudo-neutral hadron ↓ **PandoraPFA** ↓ **Scenario 2: Two-particle separation studies**

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

Data & MC Pion Samples Overview

PandoraPFA Two Particle Reconstruction

- First investigated scenarios:
  - ➡ 10 GeV pseudo-neutral & 10 GeV charged pion (50mm & 200mm distance)
  - ➡ 10 GeV pseudo-neutral & 30 GeV charged pion (50mm & 200mm distance)
  - ➡ Data: **June Beam Test 2018 @ SPS CERN**
  - ➡ MC: **GEANT4 v.10.03, QGSP_BERT_HP & FTFP_BERT_HP**



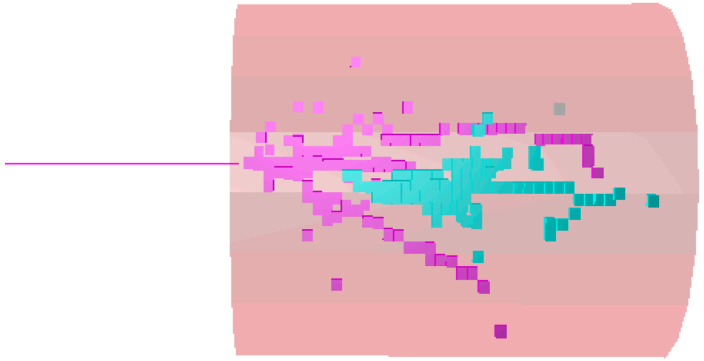
- Applied BDT-PID for hadrons (remove beam contamination)
- Event selection:
  - ➡ Same as for single particle studies except for no cut on shower start layer (allow long. separation)
- Overlay shower axis distance acceptance of +- 25mm

Total PFO Multiplicity - How many Particles Reconstructed?

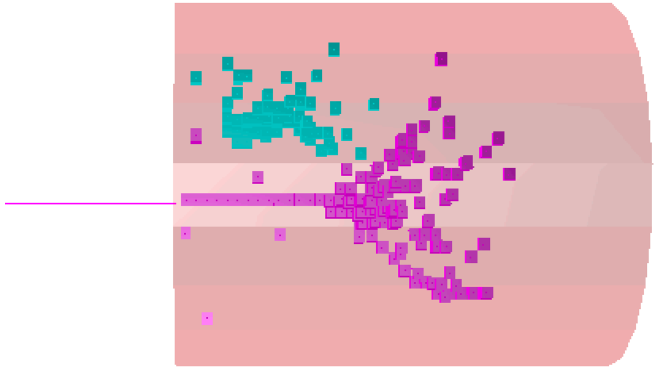
Different Scenarios

Examples of good case: Two PFO’s

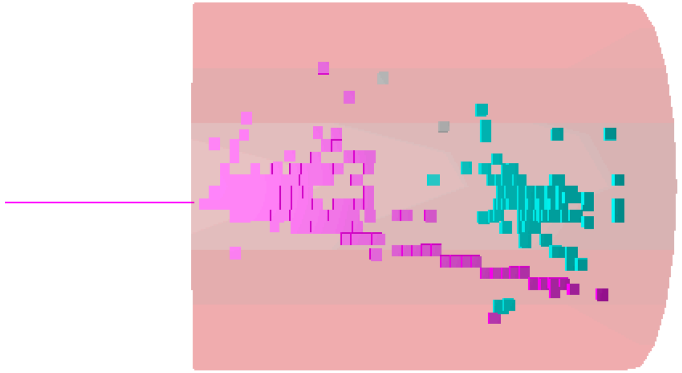
**10GeV Neutral + 10 GeV
Charged Distance: 50mm**



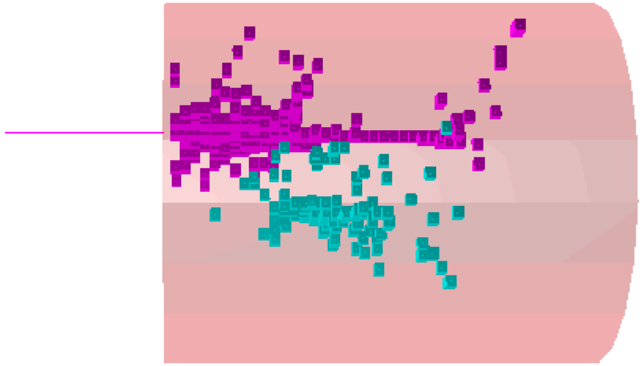
**10GeV Neutral + 10 GeV
Charged Distance: 200mm**



**10GeV Neutral + 30 GeV
Charged Distance: 50mm**



**10GeV Neutral + 30 GeV
Charged Distance: 200mm**

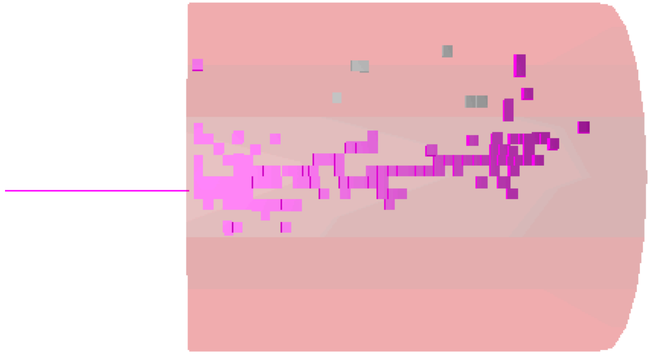


Total PFO Multiplicity - How many Particles Reconstructed?

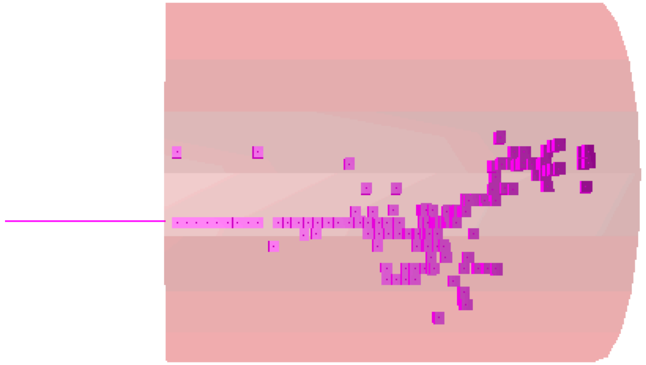
Different Scenarios

Examples of bad case: Only one PFO

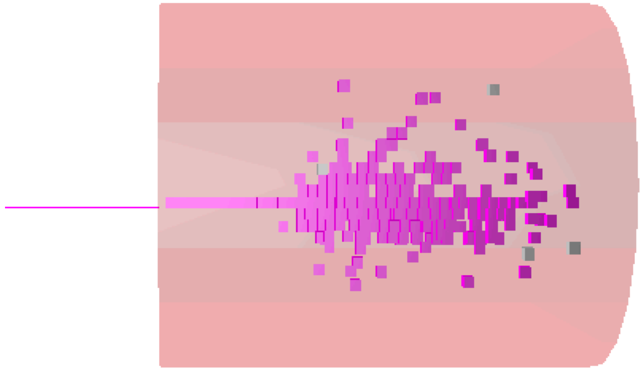
10GeV Neutral + 10 GeV Charged Distance: 50mm



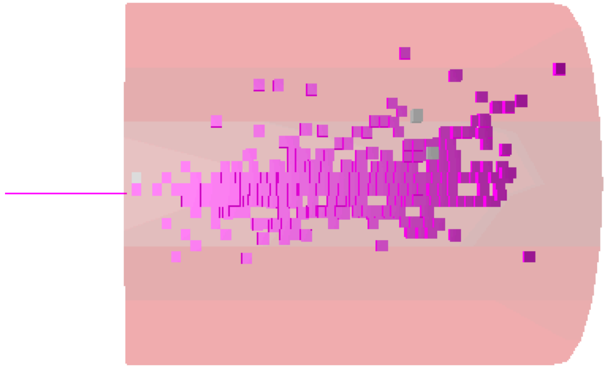
10GeV Neutral + 10 GeV Charged Distance: 200mm



10GeV Neutral + 30 GeV Charged Distance: 50mm



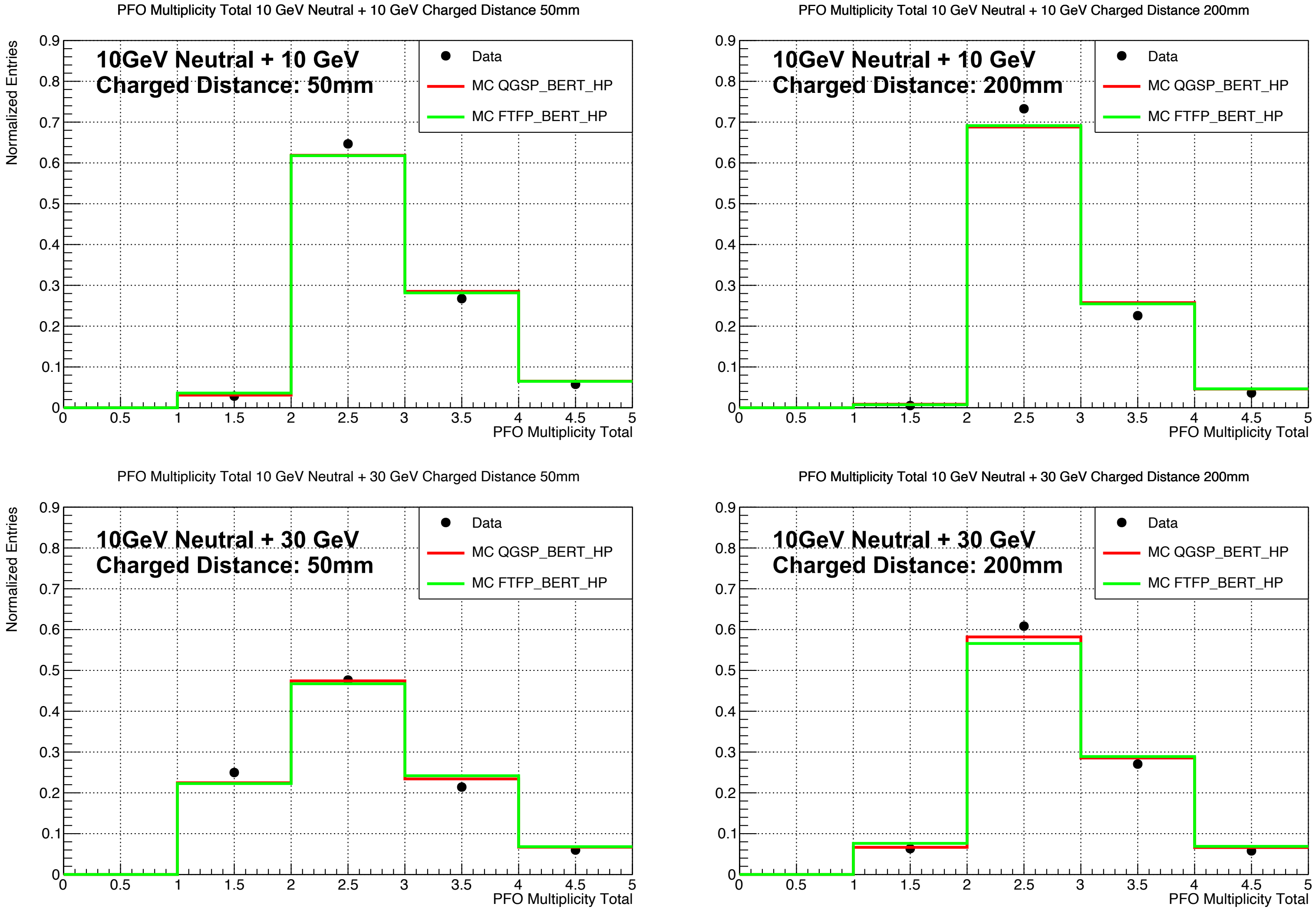
10GeV Neutral + 30 GeV Charged Distance: 200mm



Total PFO Multiplicity - How many Particles Reconstructed?

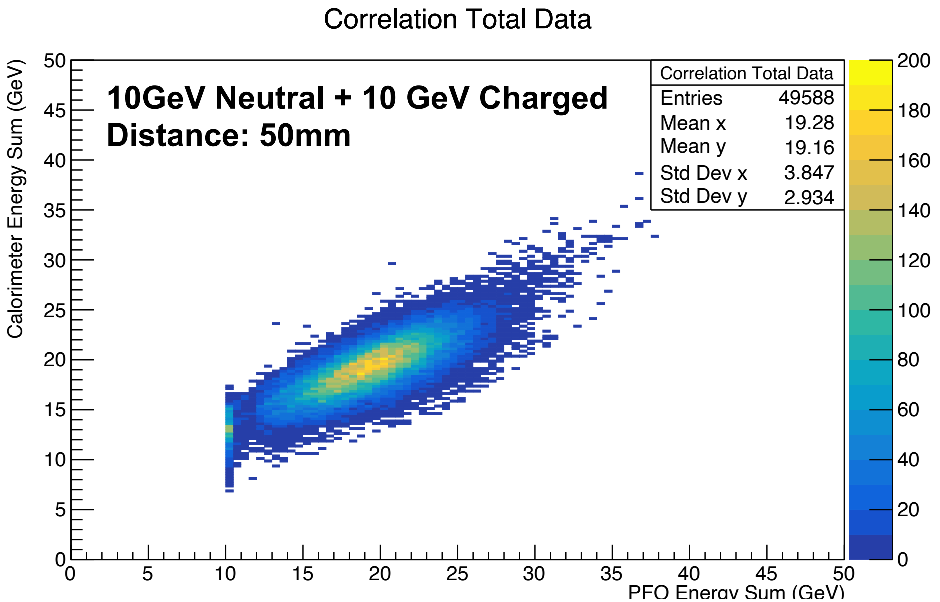
Different Scenarios

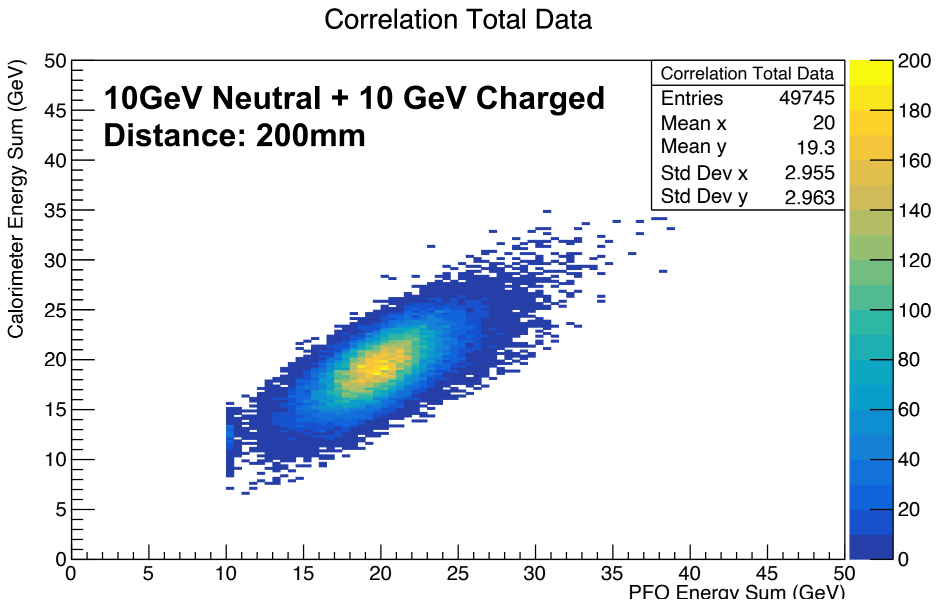
- **Good agreement between data and MC for all scenarios**
- Best case: 10GeV neutral + 10GeV charged hadron with 200mm distance, almost no events with only 1 PFO
- Considerable amount of events with just one 1 PFO in worst case scenario of 10GeV neutral and 30 GeV charged with 50mm distance
  - ➡ All hits get absorbed and associated to track

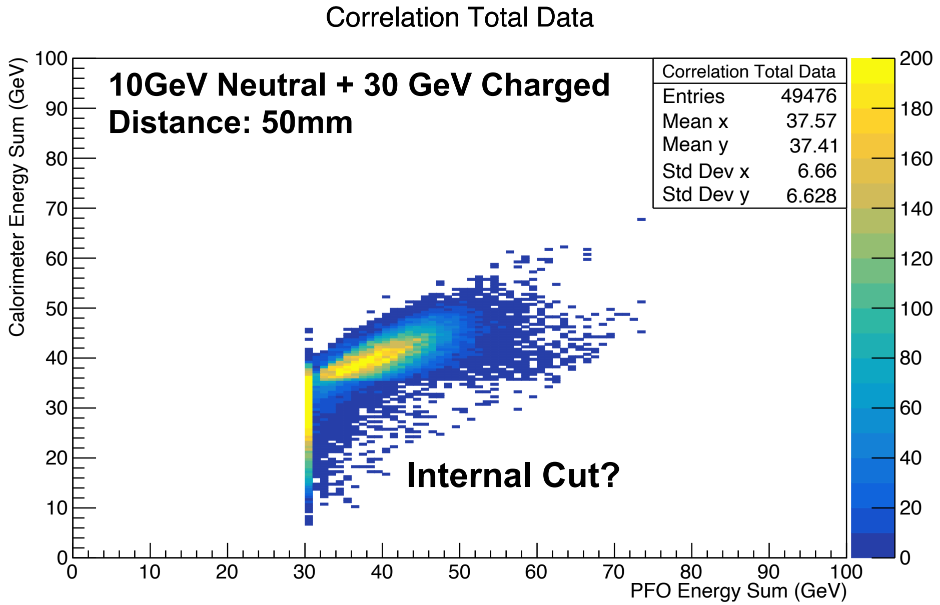


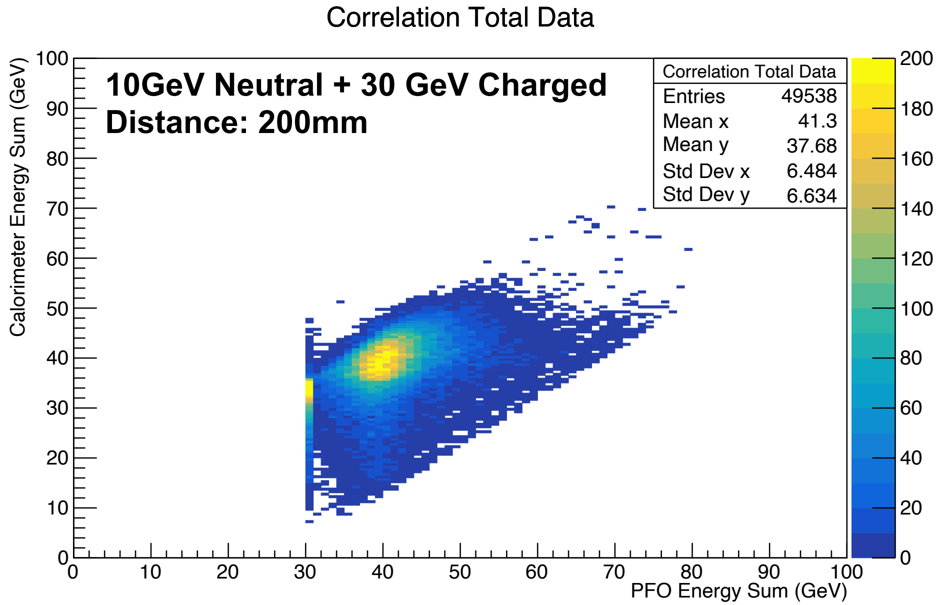
Correlation PFA vs. Calorimeter Energy Sum Total

Data





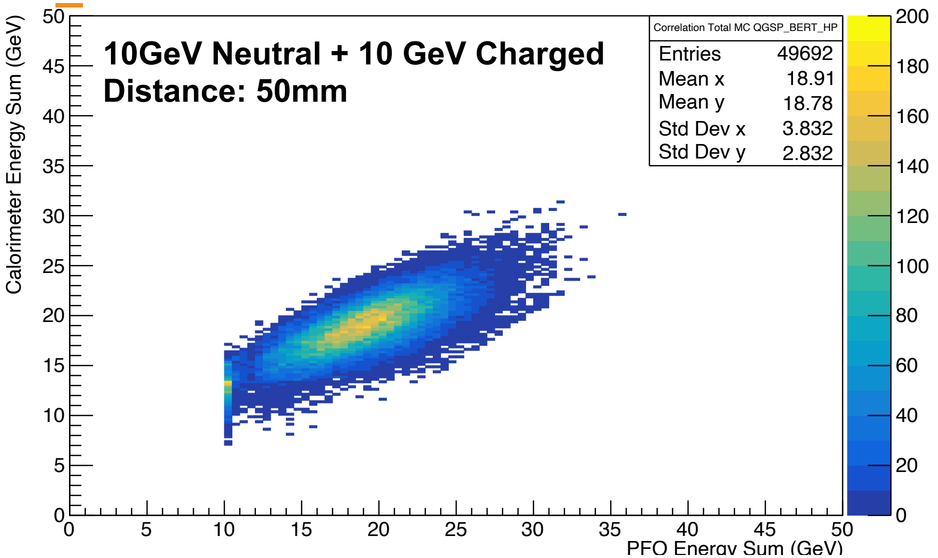




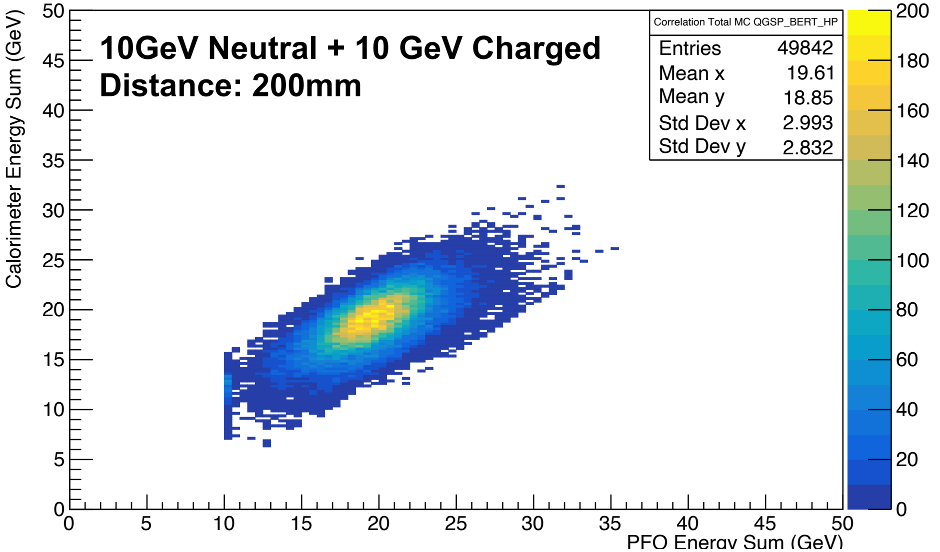
Correlation PFA vs. Calorimeter Energy Sum Total

QGSP_BERT_HP

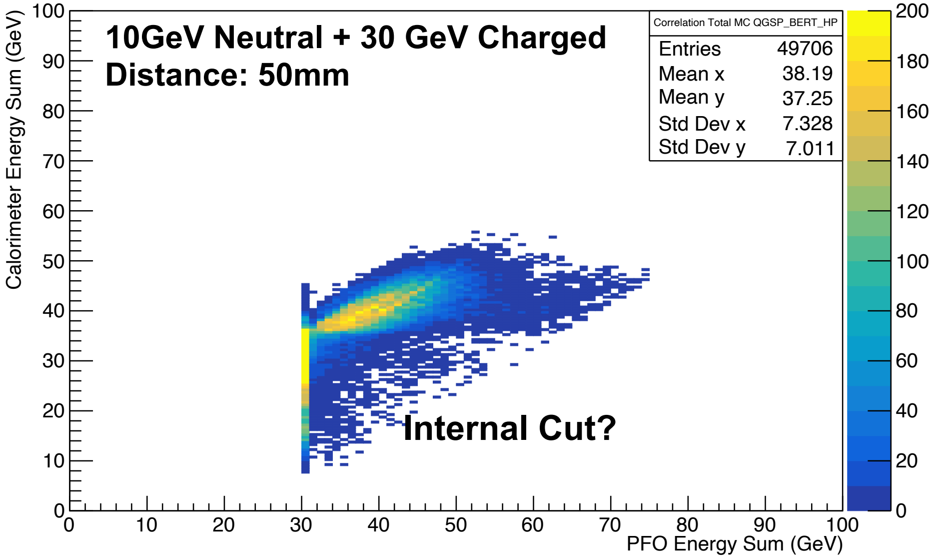
Correlation Total MC QGSP_BERT_HP



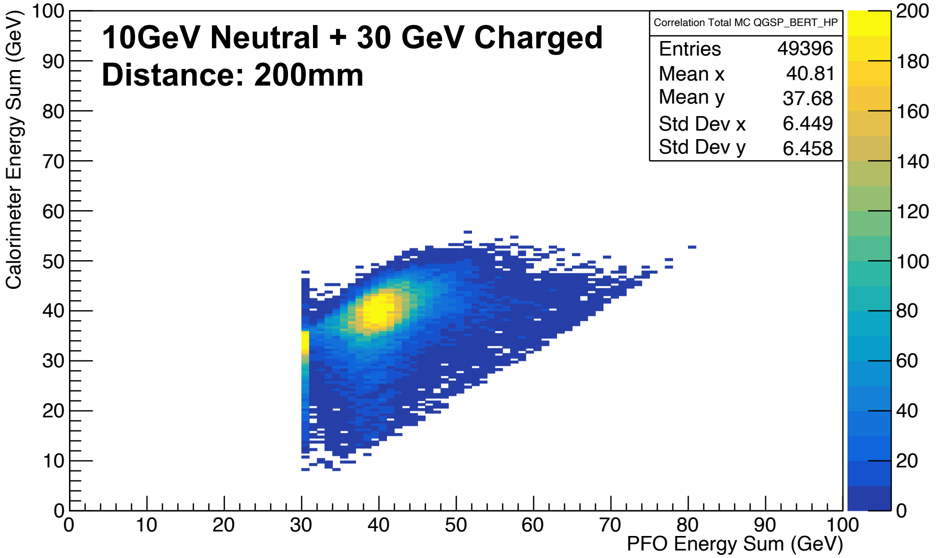
Correlation Total MC QGSP_BERT_HP



Correlation Total MC QGSP_BERT_HP

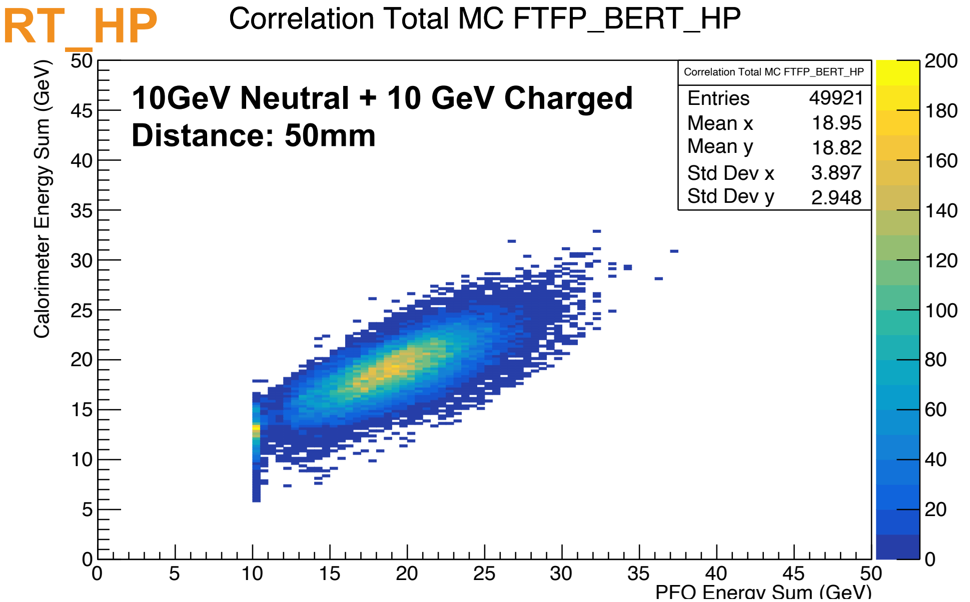


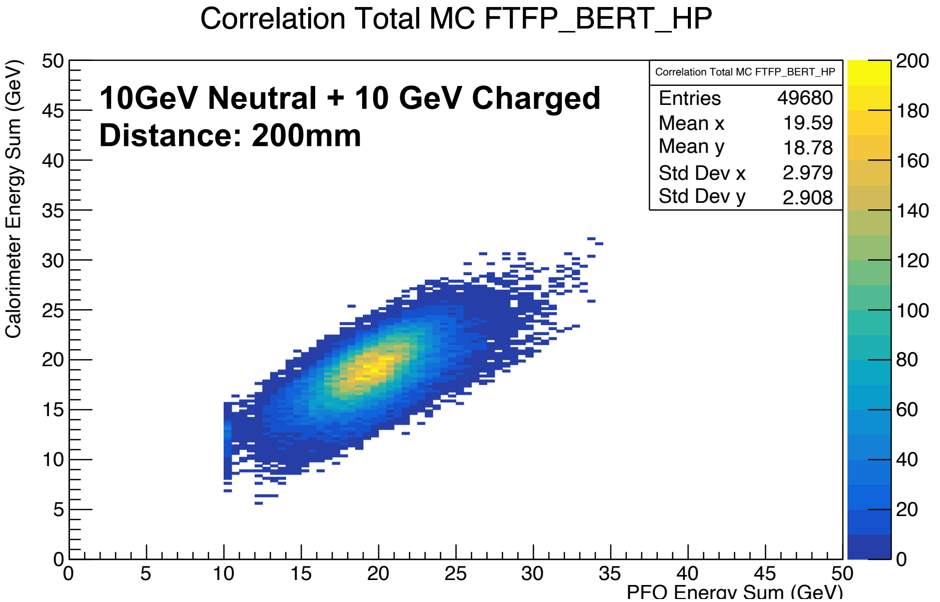
Correlation Total MC QGSP_BERT_HP

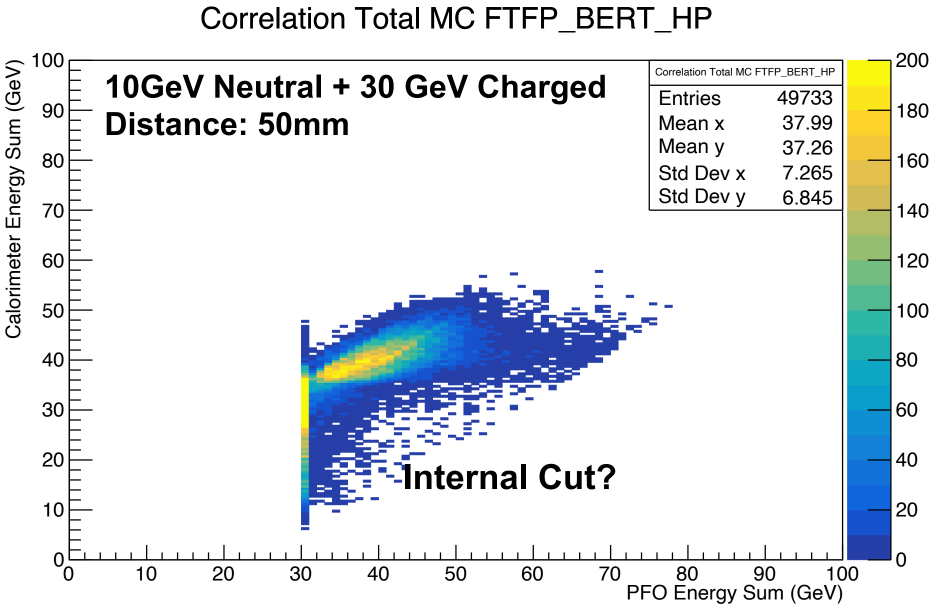


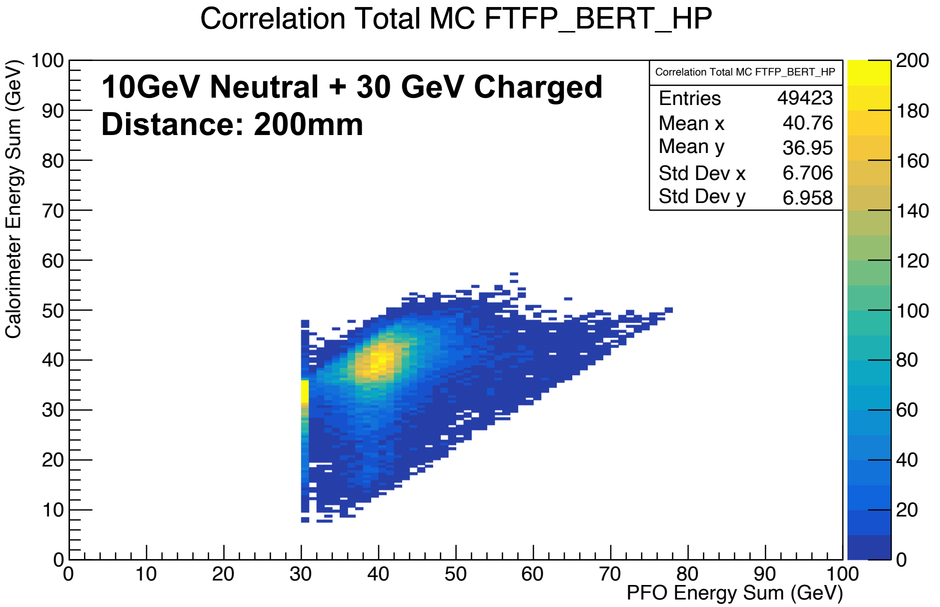
Correlation PFA vs. Calorimeter Energy Sum Total

FTFP_BERT_HP



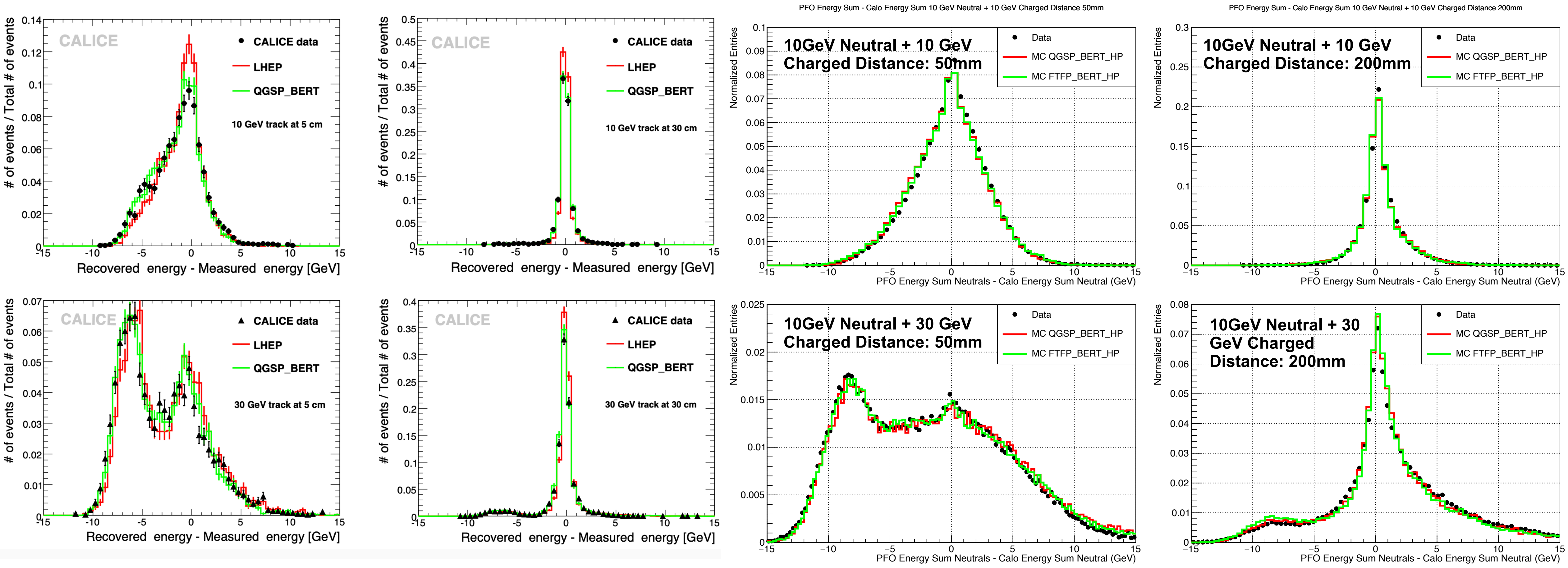






PFO Energy Neutrals - Calorimeter Energy Neutral Hadron

First CALICE PFA Paper vs. Latest Studies



- Reproduction of features and expected trends (very good data to MC agreement)
  - ➡ Distribution of current studies seem to be a bit wider (careful - no ECAL before!)
  - ➡ ToDo: Mean and RMS analysis, recovery probabilities of neutral, etc.

Summary & Outlook

- Sample preparation and event selection tools implemented, validated and finished
- Established well working PandoraPFA environment for reconstruction of AHCAL standalone events
- Quantitative studies on single particle reconstruction advanced
- First quantitative results on several two particle reconstruction scenarios:
  - General trends for energies and distances as expected
  - Data to MC agreement remarkable
  - Reproduction of first plots of old prototype paper

Further Plans:

- Single Particles: Fine-tuning (e.g. leakage cut), test several ideas of J.M. on internal PandoraPFA cuts
- Two Particles: Investigation of full distance range (0-350mm), high level analysis (e.g. separation power,…)
- Leakage correction energy-plugin in PandoraPFA

Summary & Outlook

- Sample preparation and event selection tools implemented, validated and finished
- Established well working PandoraPFA environment for reconstruction of AHCAL standalone events
- Quantitative studies on single particle reconstruction advanced
- First quantitative results on several two particle reconstruction scenarios:
  - General trends for energies and distances as expected
  - Data to MC agreement remarkable
  - Reproduction of first plots of old prototype paper

Thank you!

(And special thanks to J. Marshall for his continuous feedback!)

Pandora's box is open!

Further Plans:

- Single Particles: Fine-tuning (e.g. leakage cut), test several ideas of J.M. on internal PandoraPFA cuts
- Two Particles: Investigation of full distance range (0-350mm), high level analysis (e.g. separation power,…)
- Leakage correction energy-plugin in PandoraPFA

Backup

Particle Flow Approach

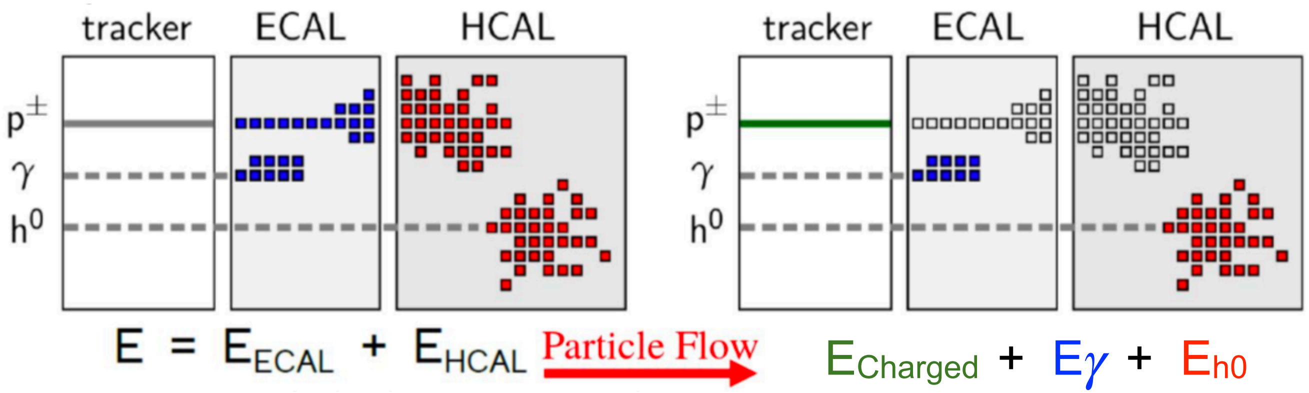
Reaching High Precision

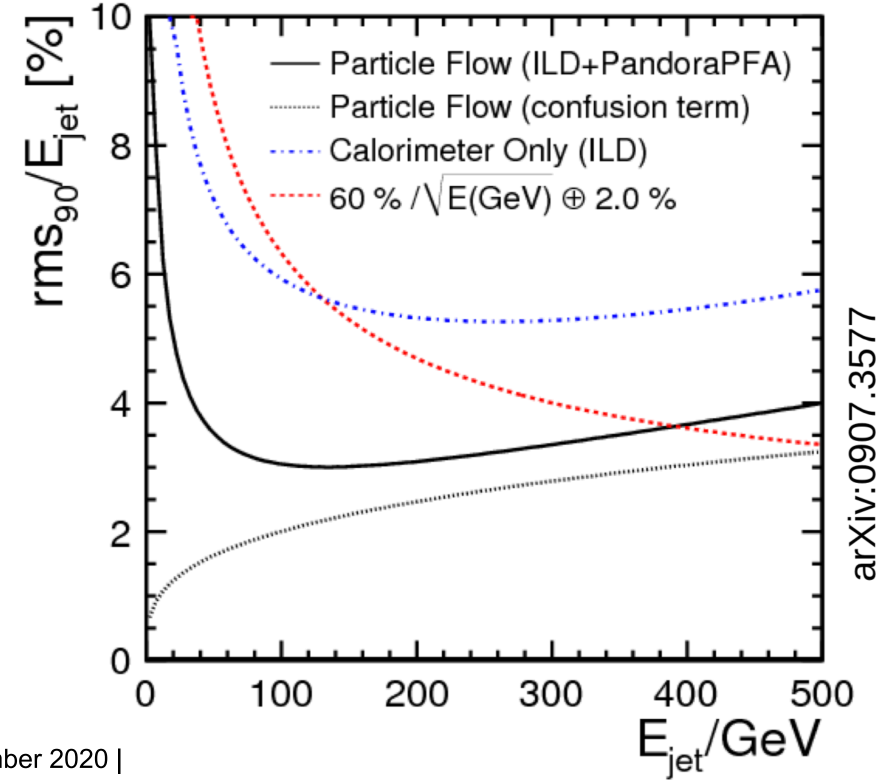
- Goal at the ILC: Jet energy resolution of 3-4% for jet energies between 40-500 GeV
- Typical jet composition of 72% hadrons measured with poor hadronic energy resolution ~60%/√E

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

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

Conventional **PFA**



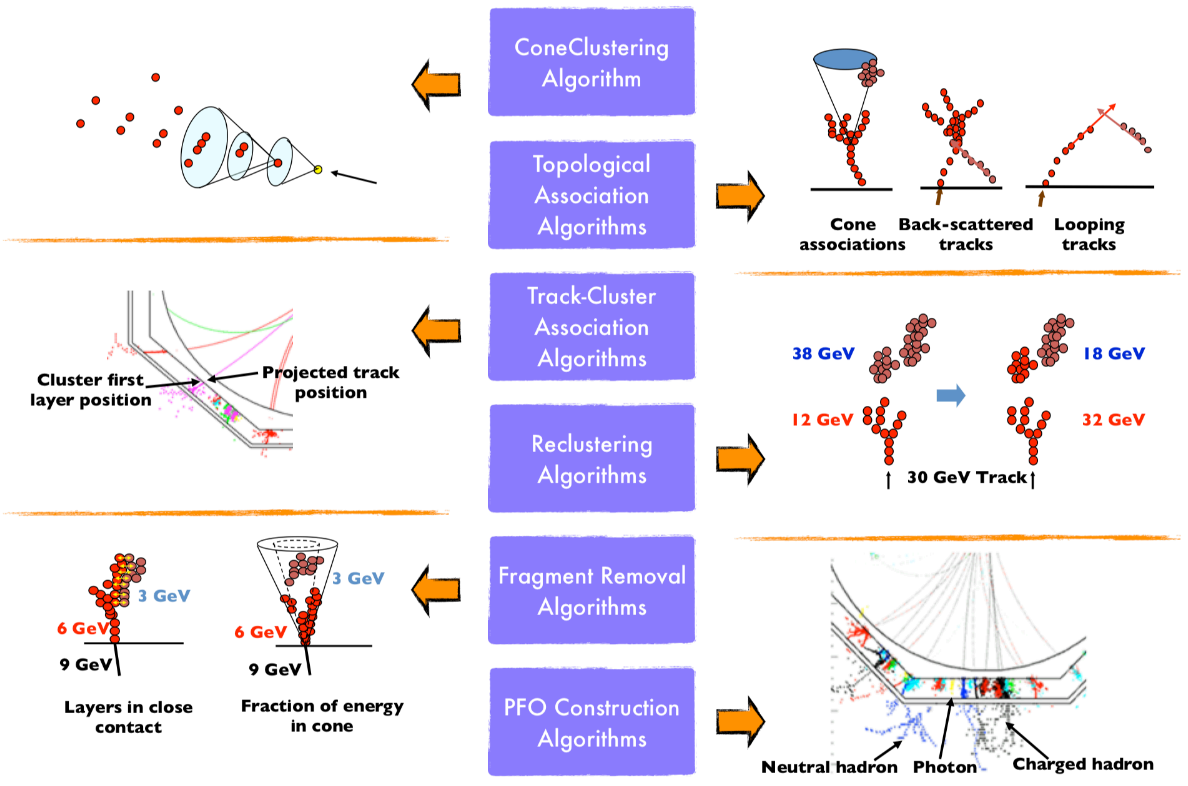


arXiv:0907.3577

The Pandora Particle Flow Algorithm

Multi-Algorithm Pattern Recognition

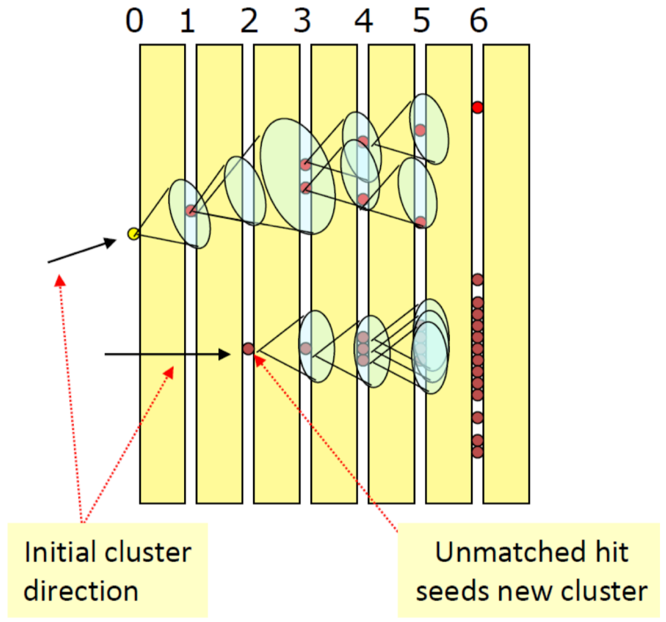
Key Steps of PandoraPFA



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

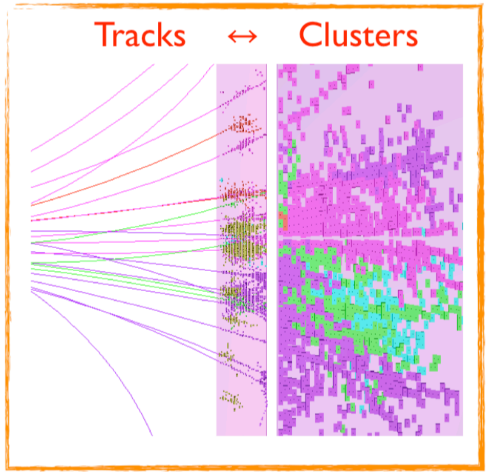
- PandoraPFA: Complex multi-algorithm chain for pattern recognition
  - ➡ Fully reconstruct the path of individual particles through the full detector including their shower substructure
    - ➡ **Correctly** identify charged and neutral particles corresponding energy depositions and take best energy measure for this **Particle Flow Objects (PFOs)**
    - ➡ Charged PFO: Calorimeter cluster + track
    - ➡ Neutral PFO: Calorimeter cluster

Clustering



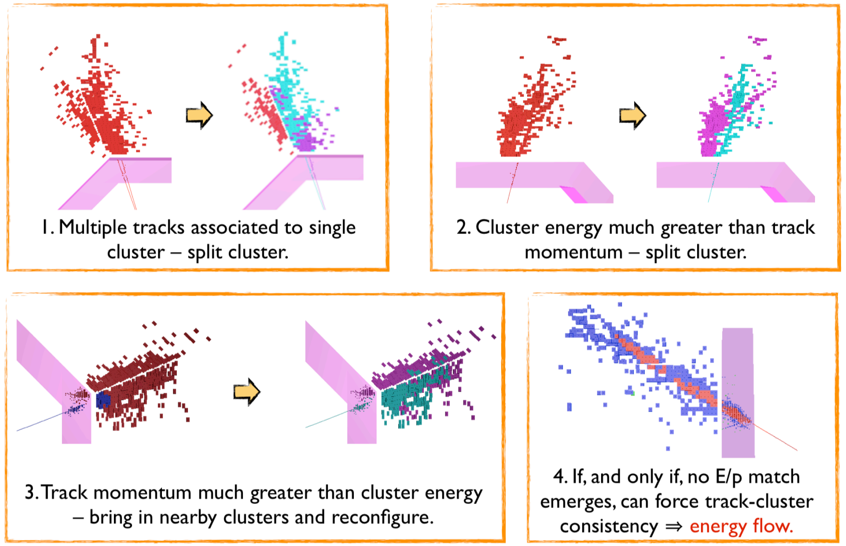
Track to Cluster Association

- 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



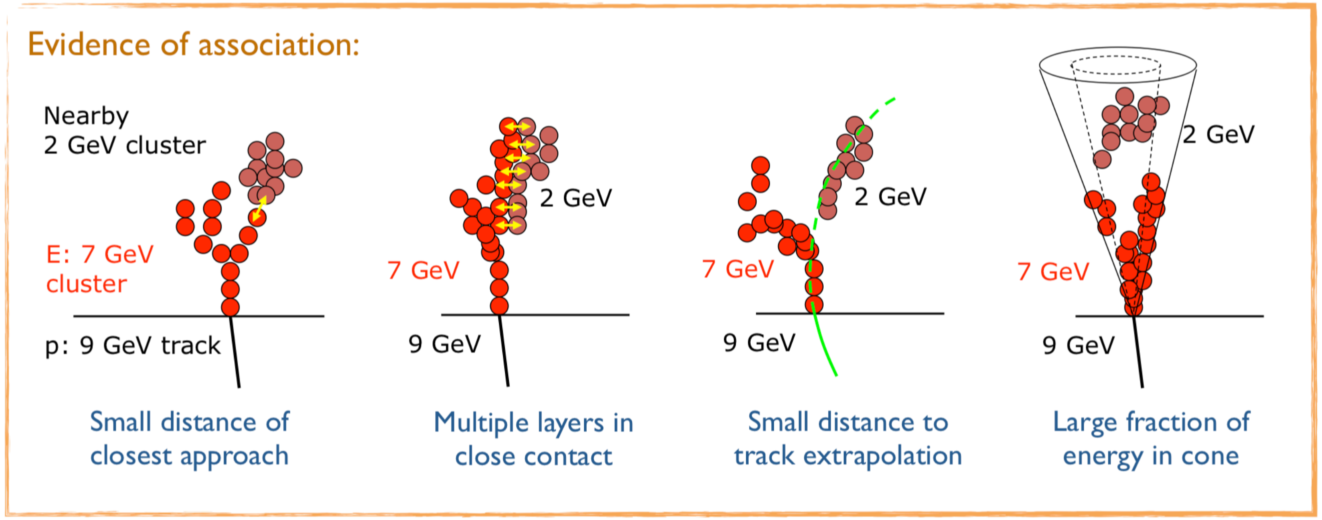
1. Multiple tracks associated to single cluster – split cluster.

2. Cluster energy much greater than track momentum – split cluster.

3. Track momentum much greater than cluster energy – bring in nearby clusters and reconfigure.

4. If, and only if, no E/p match emerges, can force track-cluster consistency ⇒ energy flow.

Fragment Removal

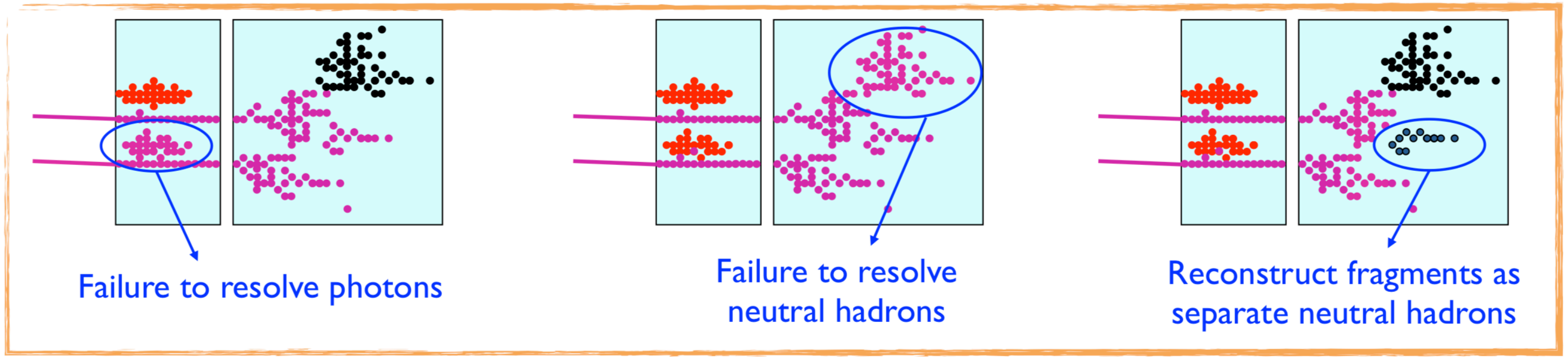


The Confusion Term

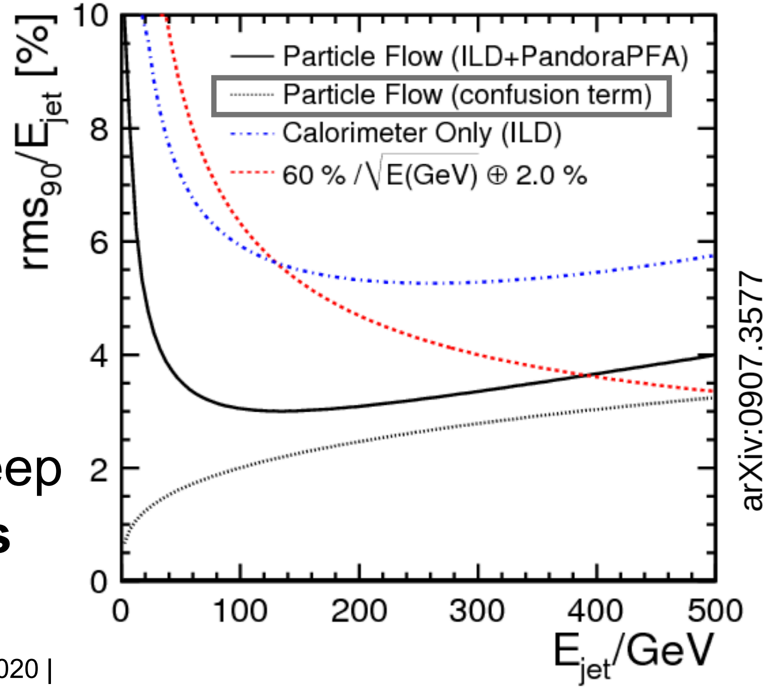
The Limit of PFA Reconstruction

- **Topologically + energetically confusing** scenarios might appear in specific events:
  - ➡ Photons close to a charged particle track
  - ➡ Neutral particle calorimeter hits right next to charged particle calorimeter hits
  - ➡ Neutral fragments in charged hadron shower topologically separated

Types of confusion



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

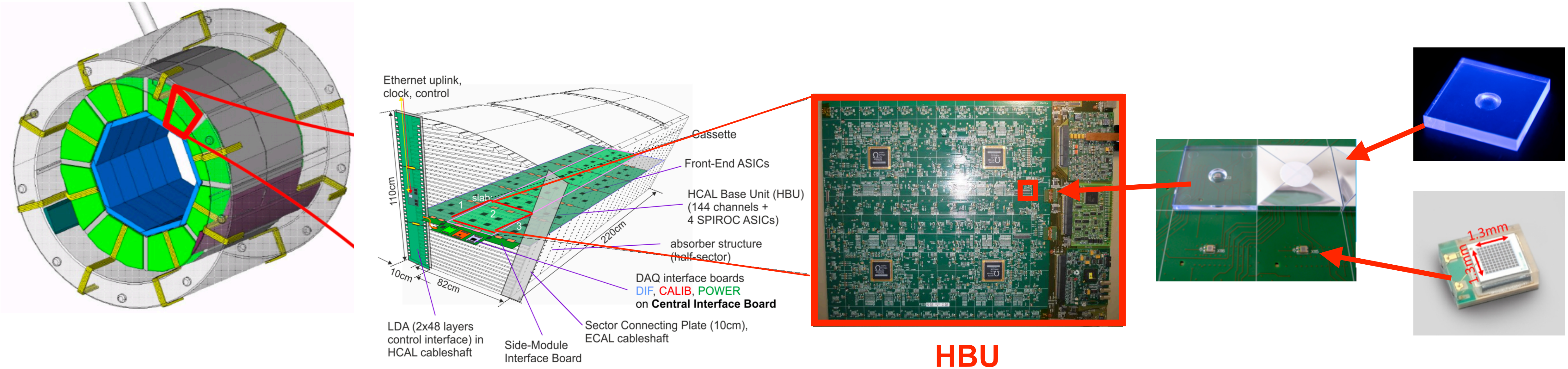


- To be able to assign track to charged particle cluster and keep confusion level low: **High granularity hadron calorimeters required and good energy resolution beneficial**

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



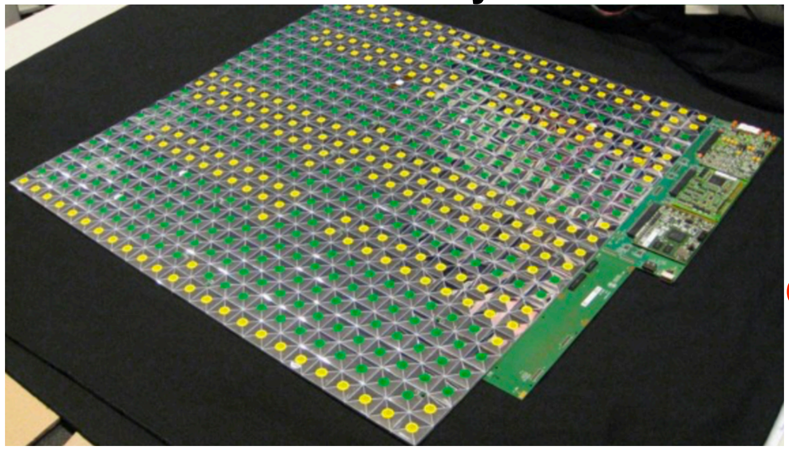
The CALICE AHCAL Technological Prototype 2018

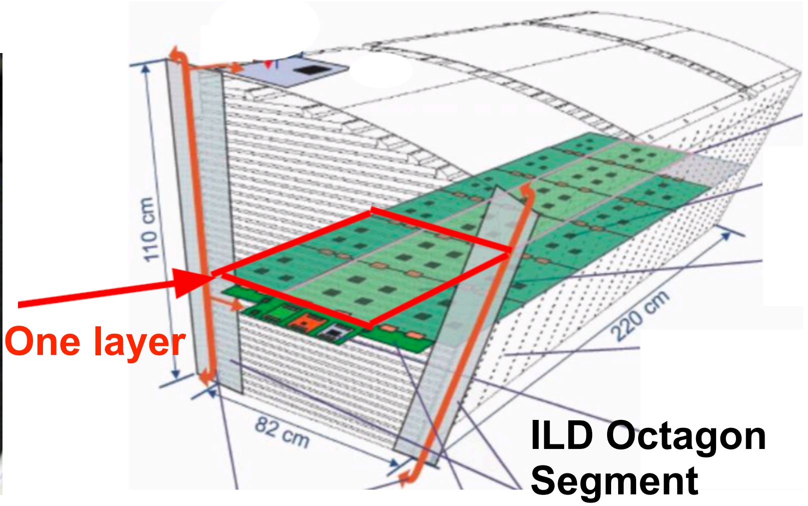
Hardware and Goals

- 38 active layers of 72 · 72 cm^2 alternating with ~1.72 cm thick steel absorbers (~4 λ_n)
- 1 layer = 4 HBU's = 16 ASICS (SPIROC2E) = 576 SiPM-on-tile channels
  - ➡ Total of 608 ASICS, ~22k channels
- **Goals:** Scalability of SiPM-on-tile calorimetry, reliable detector operation and studies on energy linearity, resolution and shower shapes/separation (PFA)

- SiPM (Hamamatsu S13360-1325PE):
  - ➡ 2668 pixels
  - ➡ Operated at ~5 V overvoltage

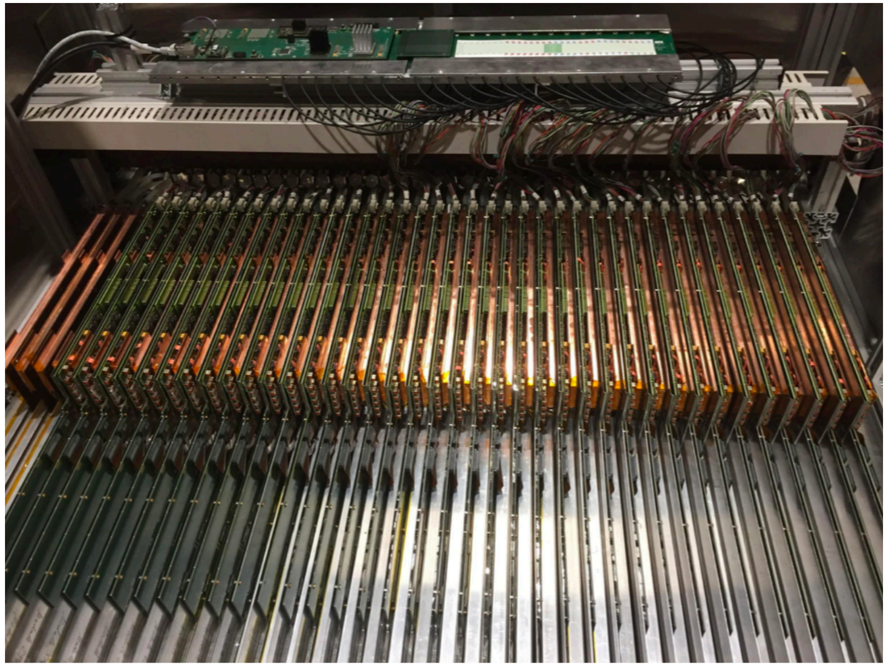
1 AHCAL layer





ILD Octagon Segment

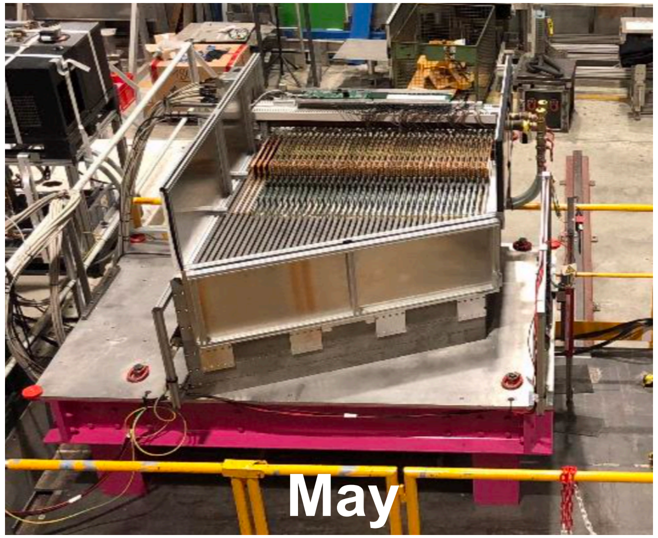
38 layers within steel absorber stack

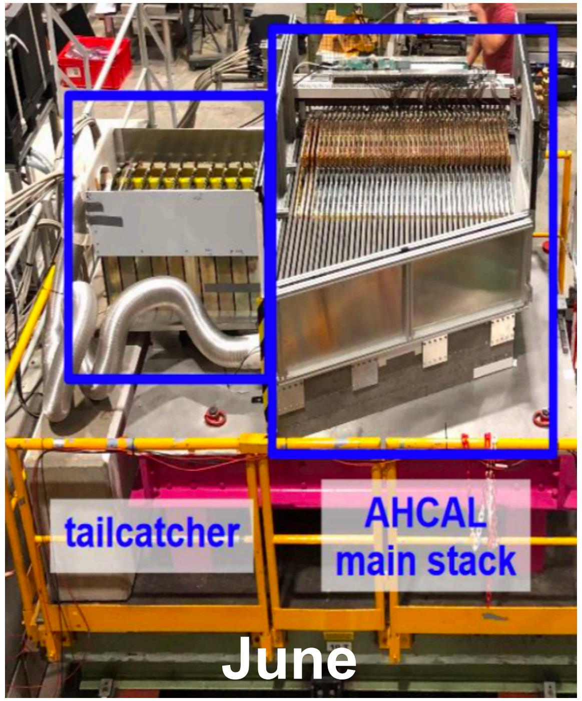


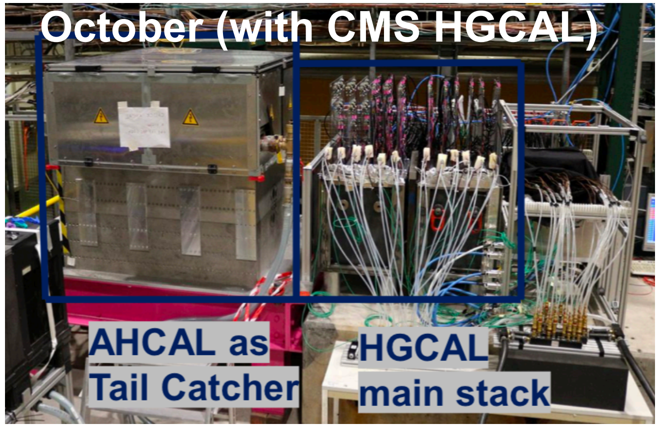
The CALICE AHCAL Beam Test Campaigns 2018

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

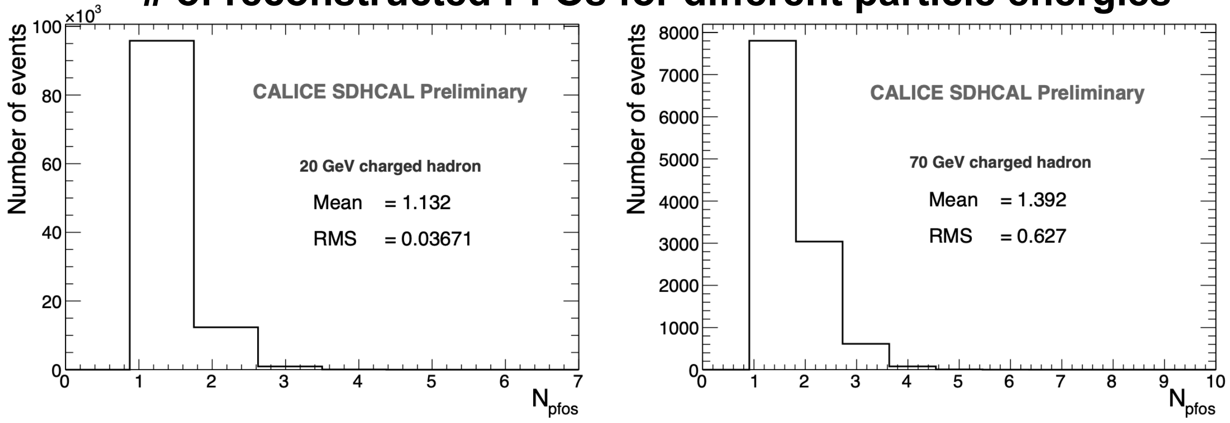
PandoraPFA

Scenario 1:
Single particle reconstruction studies

PandoraPFA

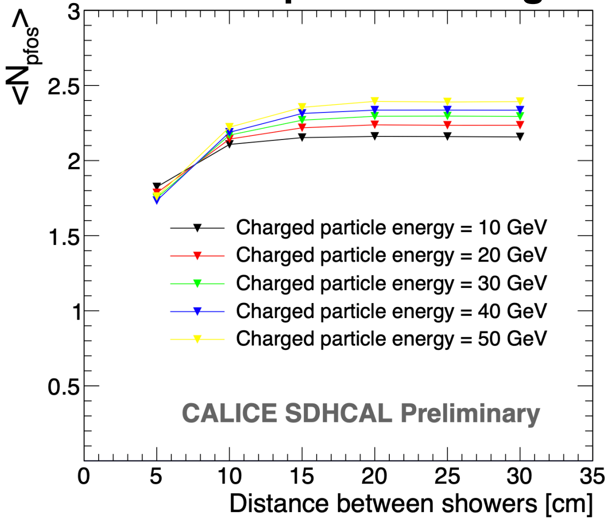
Scenario 2:
Two-particle separation studies

# of reconstructed PFOs for different particle energies

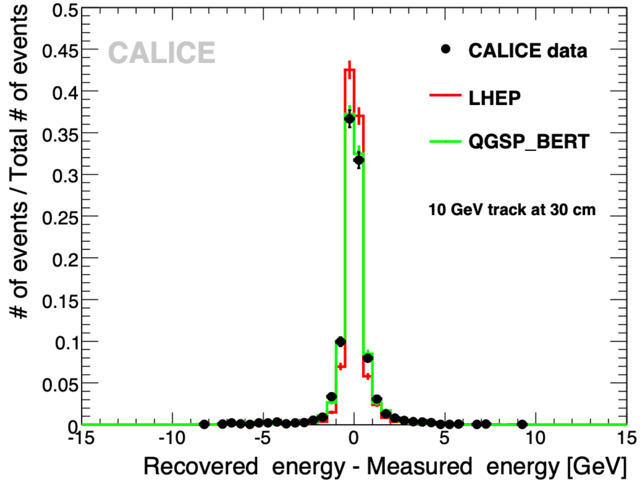


http://cds.cern.ch/record/2669487/files/fulltext.pdf

Mean # of reconstructed PFOs for different particle energies



Energy difference of calorimeter pure & PFO reconstruction



https://arxiv.org/abs/1105.3417

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 and validated
  - ➡ **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

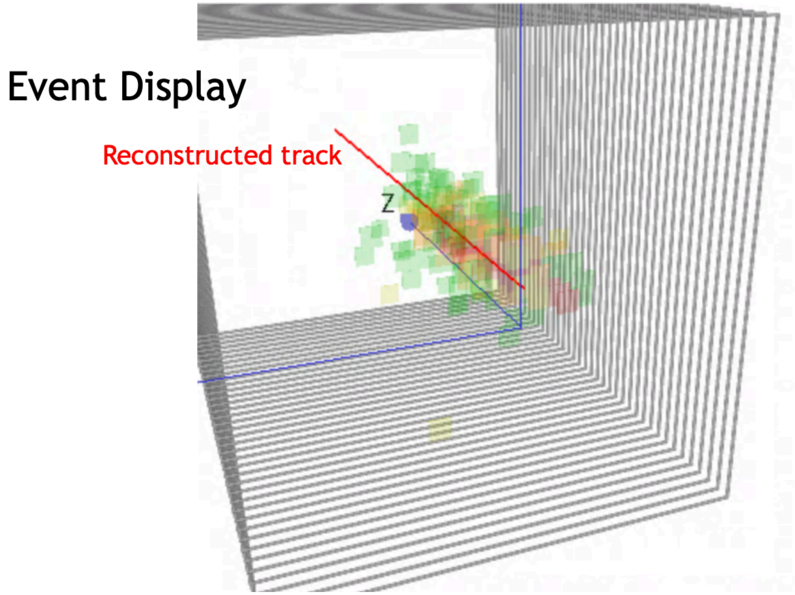
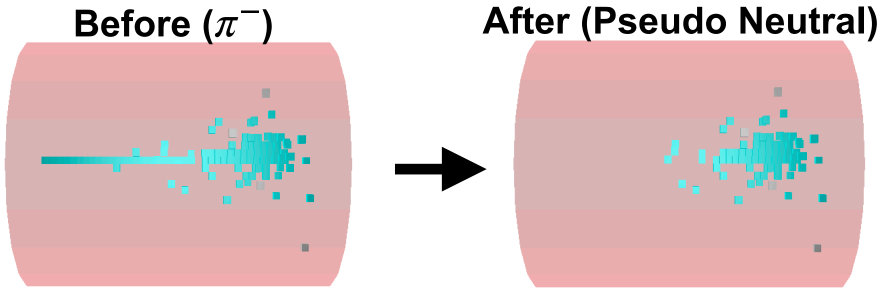


Illustration of pseudo neutral generation

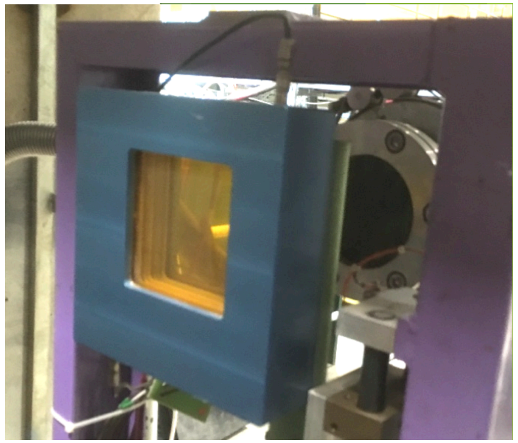


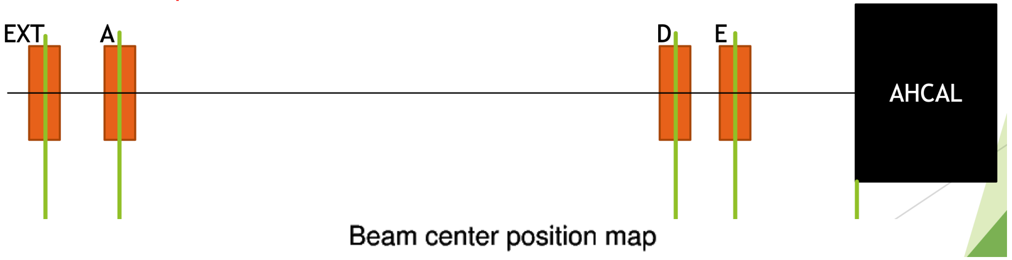
Delay Wire Chambers (DWC)

Providing Tracks for Beam Test Events

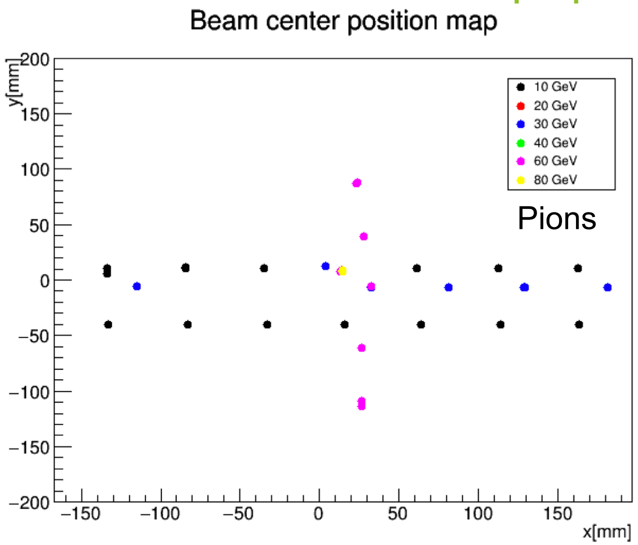
Work done by Linghui Liu (U. Tokyo)
(https://agenda.linearcollider.org/event/8368/contributions/44971/attachments/35214/54544/LL_AHCALmain_2019.pdf)

- 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



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

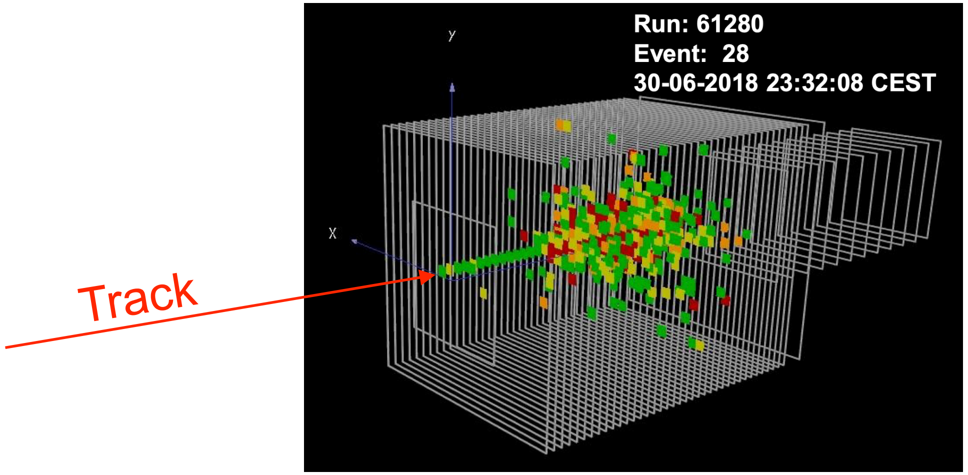
- Data tracks: Reconstructed from DWC of beam test
- MC tracks: MC primary particle endpoint position X/Y extrapolation

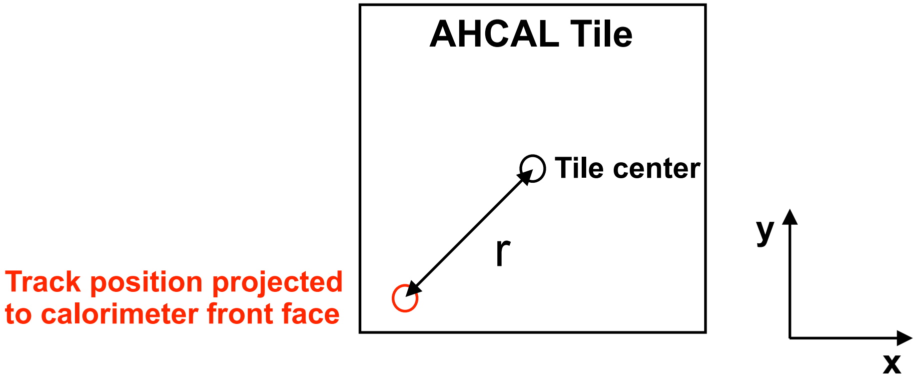
➡ **Track quality?**

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?

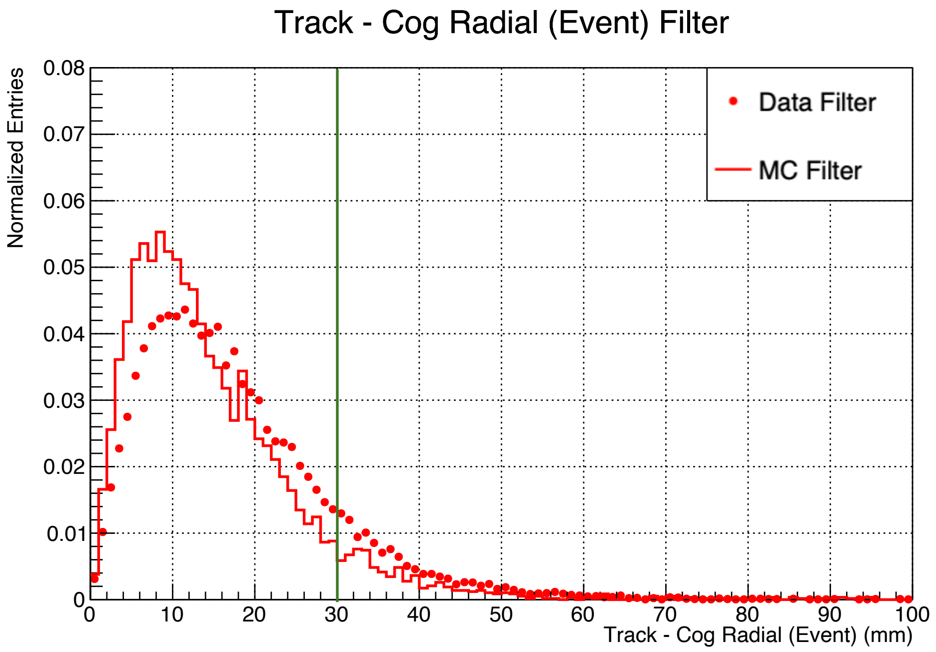




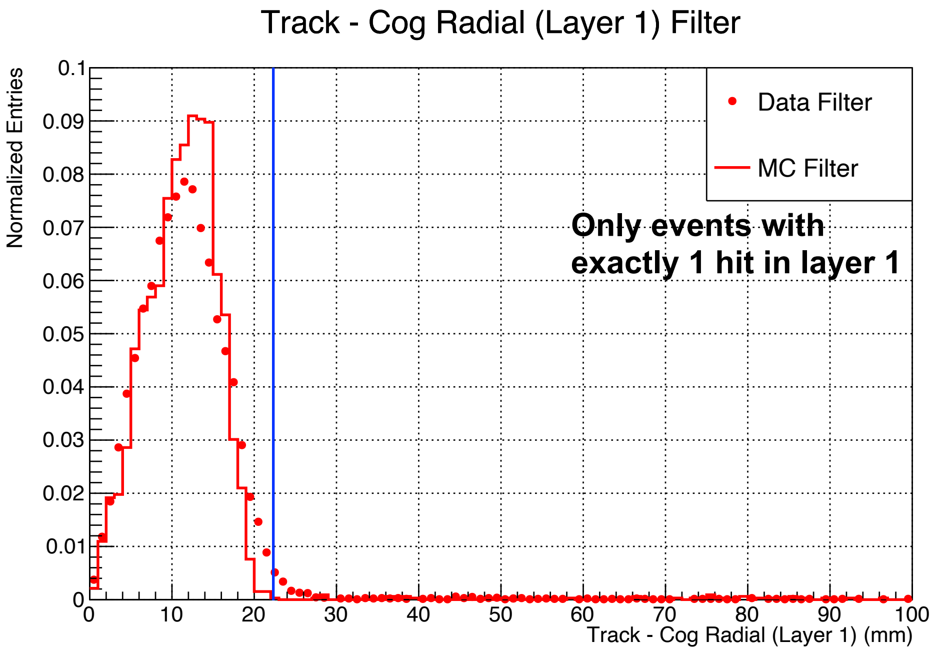
Track Quality Results 20 GeV π^-

Precise Tracks for PandoraPFA Reconstruction

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



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



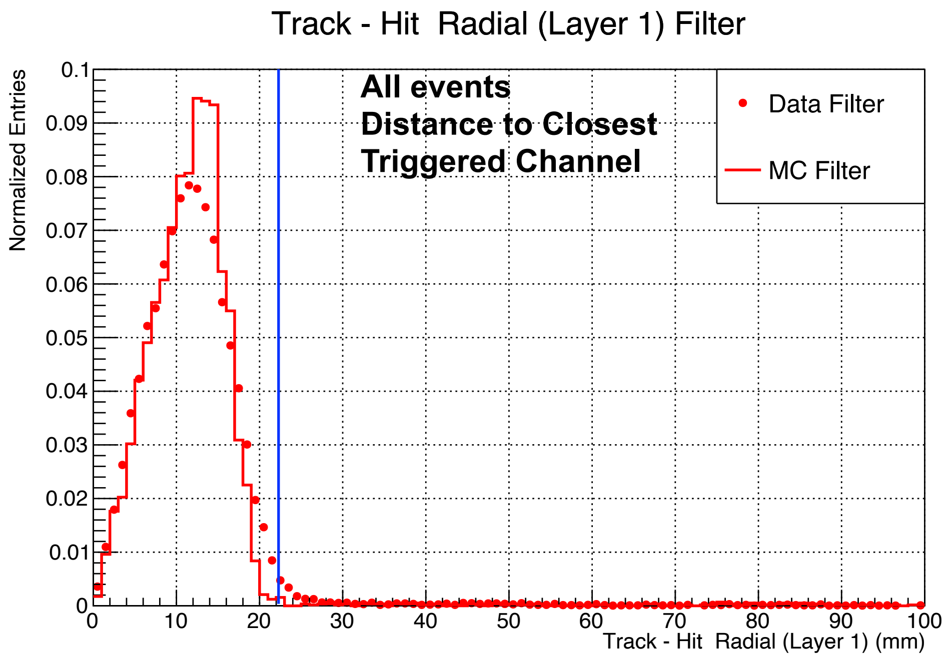
$$r = \sqrt{(x_{track} - x_{hit})^2 + (y_{track} - y_{hit})^2}$$

- 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



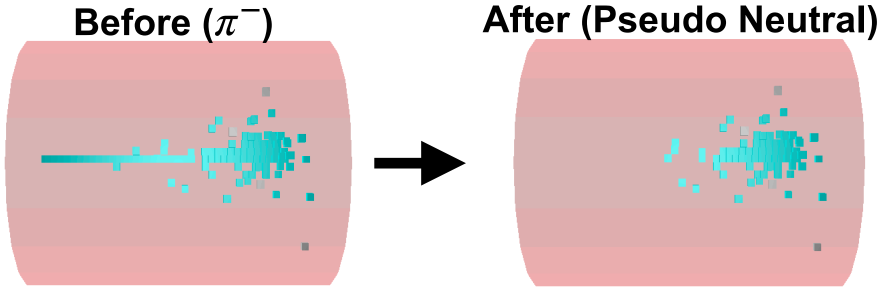
$$r = \sqrt{(x_{track} - x_{hit})^2 + (y_{track} - y_{hit})^2}$$

- 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^-

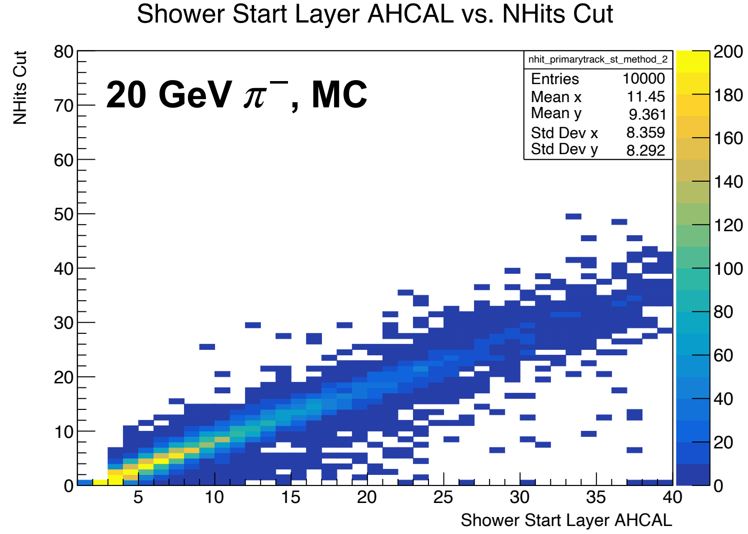
➡ **Excellent track quality validated for data and MC**

Finding and Removing Primary Track

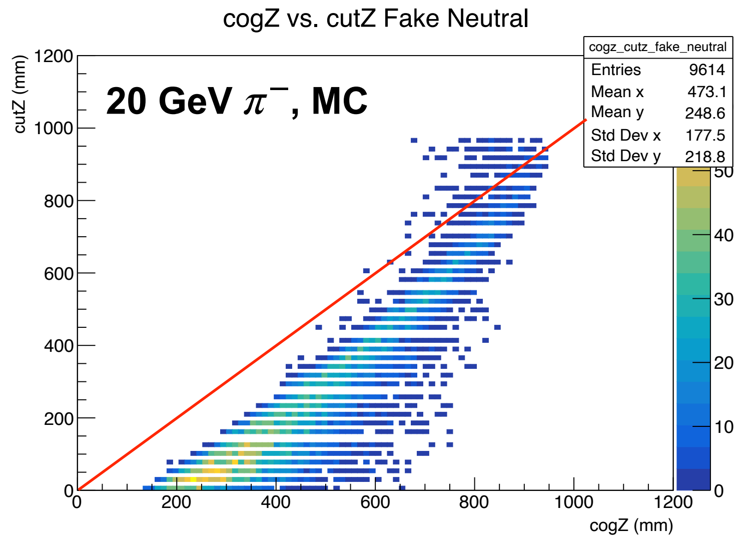
The Method for Creating Pseudo Neutral Hadrons



- 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

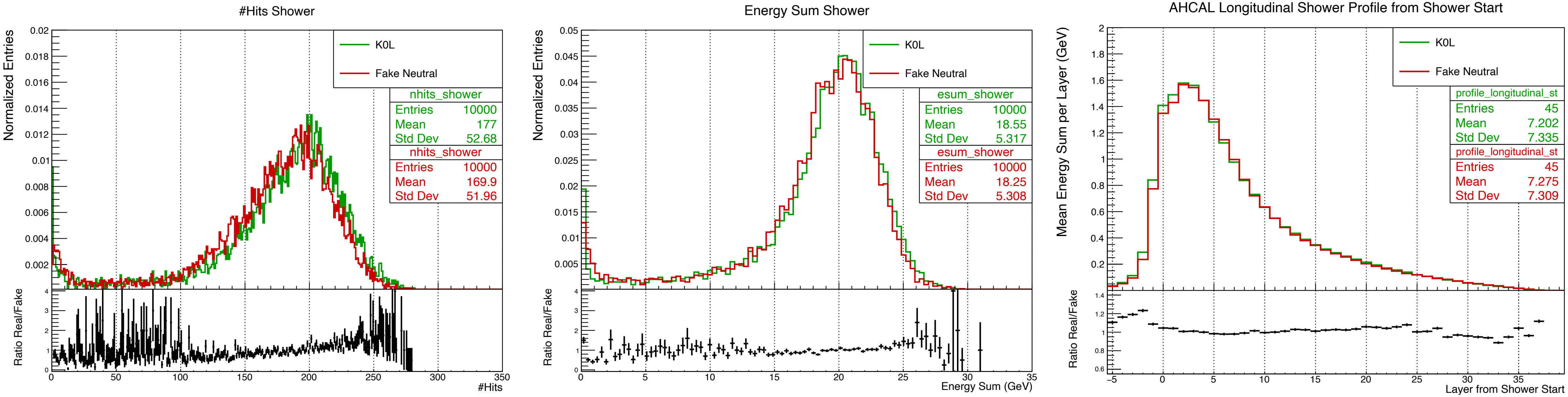


- Method robust and working well:
  - ➡ # cut hits (primary track) well correlated with shower start layer
  - ➡ Z position of potentially last cut hit well before cogZ for most events



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

Validation of Primary Track Removal Algorithm

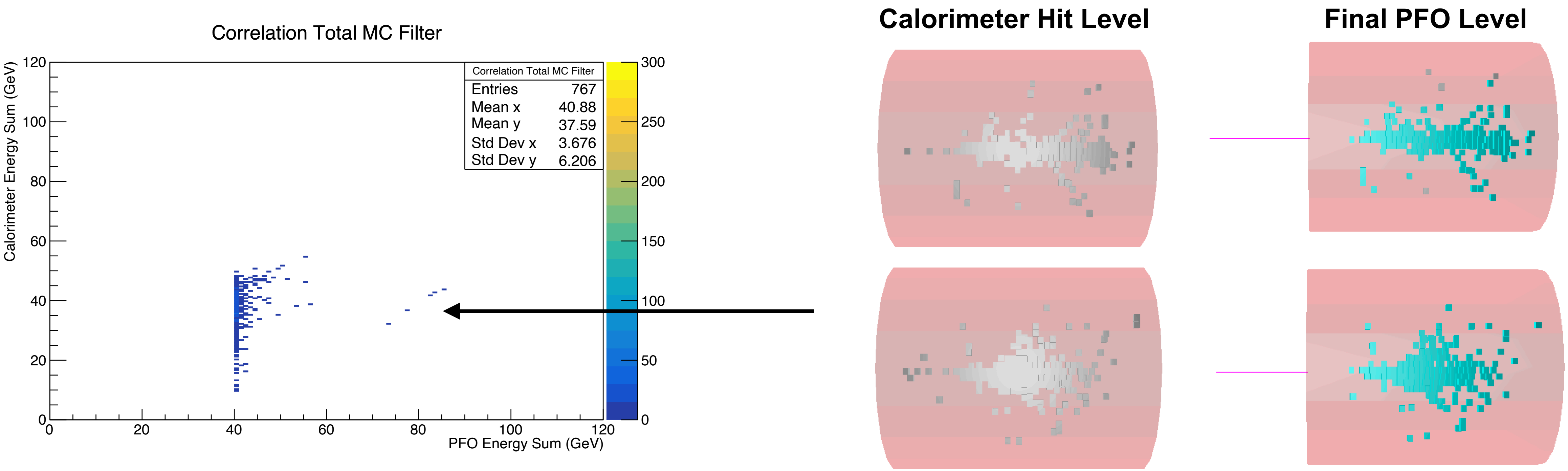


- In general **good agreement** between real neutrals (K0L) and pseudo neutrals (cut π^-) in number of hits, energy sum and longitudinal shower profile

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

Isolated Hit Merging Algorithm - PandoraPFA

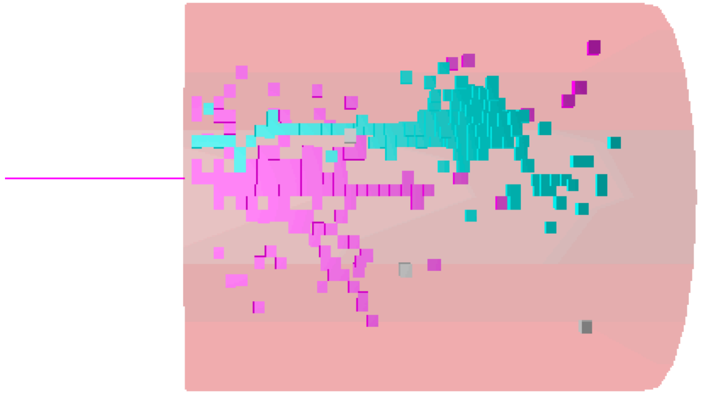
Diagonal „Gap" Events

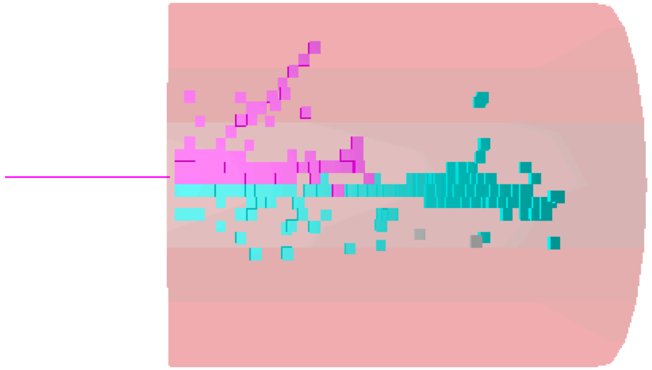


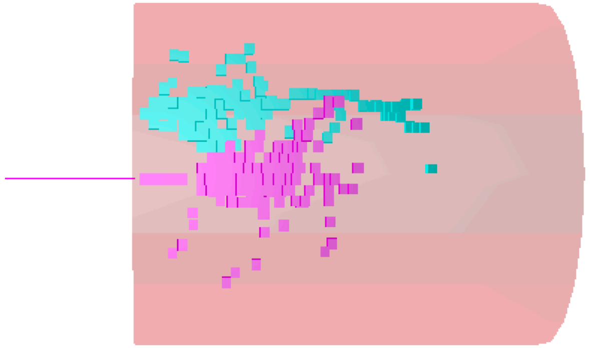
- 98% of diagonal events (non working track-cluster assignment) caused by PandoraPFA's "Isolated Hit Merging“ algorithm
  - Topological gap in primary track of first layers causes algorithm to „cut off“ hits which makes track-cluster assignment impossible
  - ➡ Fine granularity of ECAL in front missing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

High Energy Events in 20 GeV Pion Beam Test Data

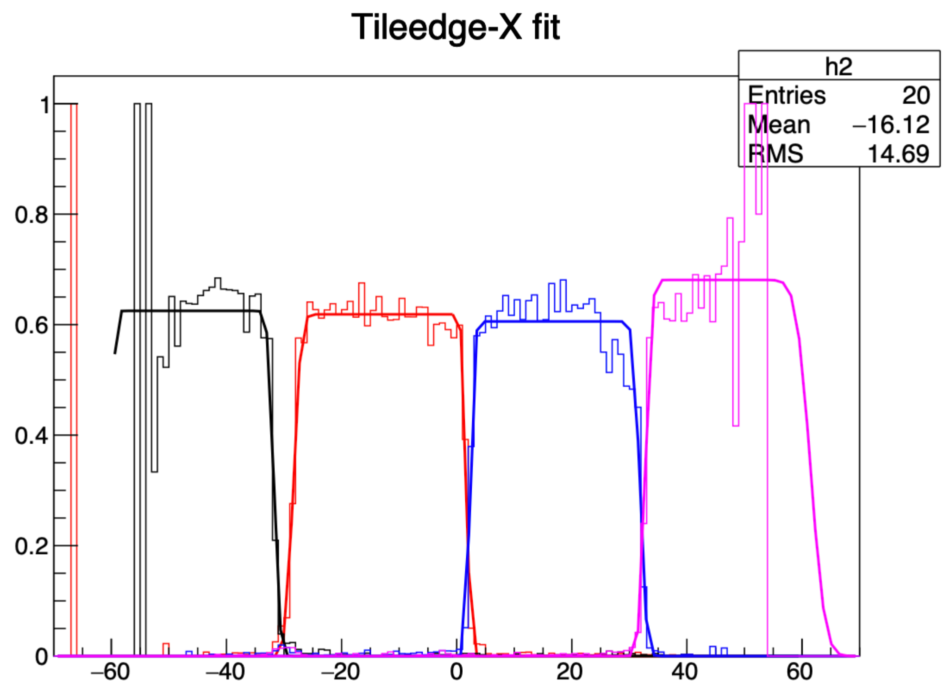
Examples - Multi Particle Events?

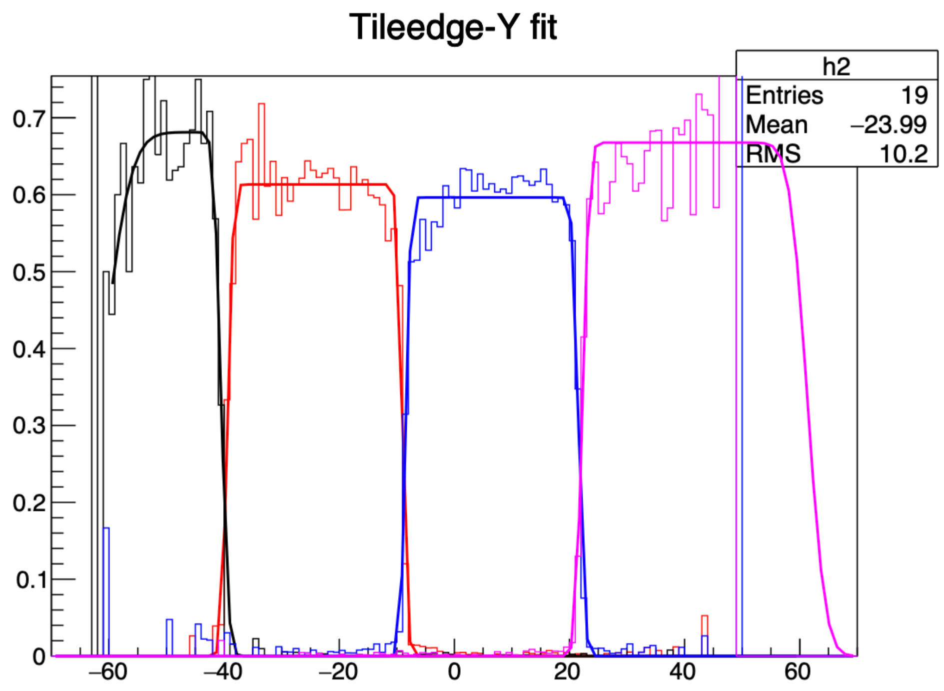

Energy in Calorimeter: ~40 GeV


Energy in Calorimeter: ~42 GeV


Energy in Calorimeter: ~38 GeV

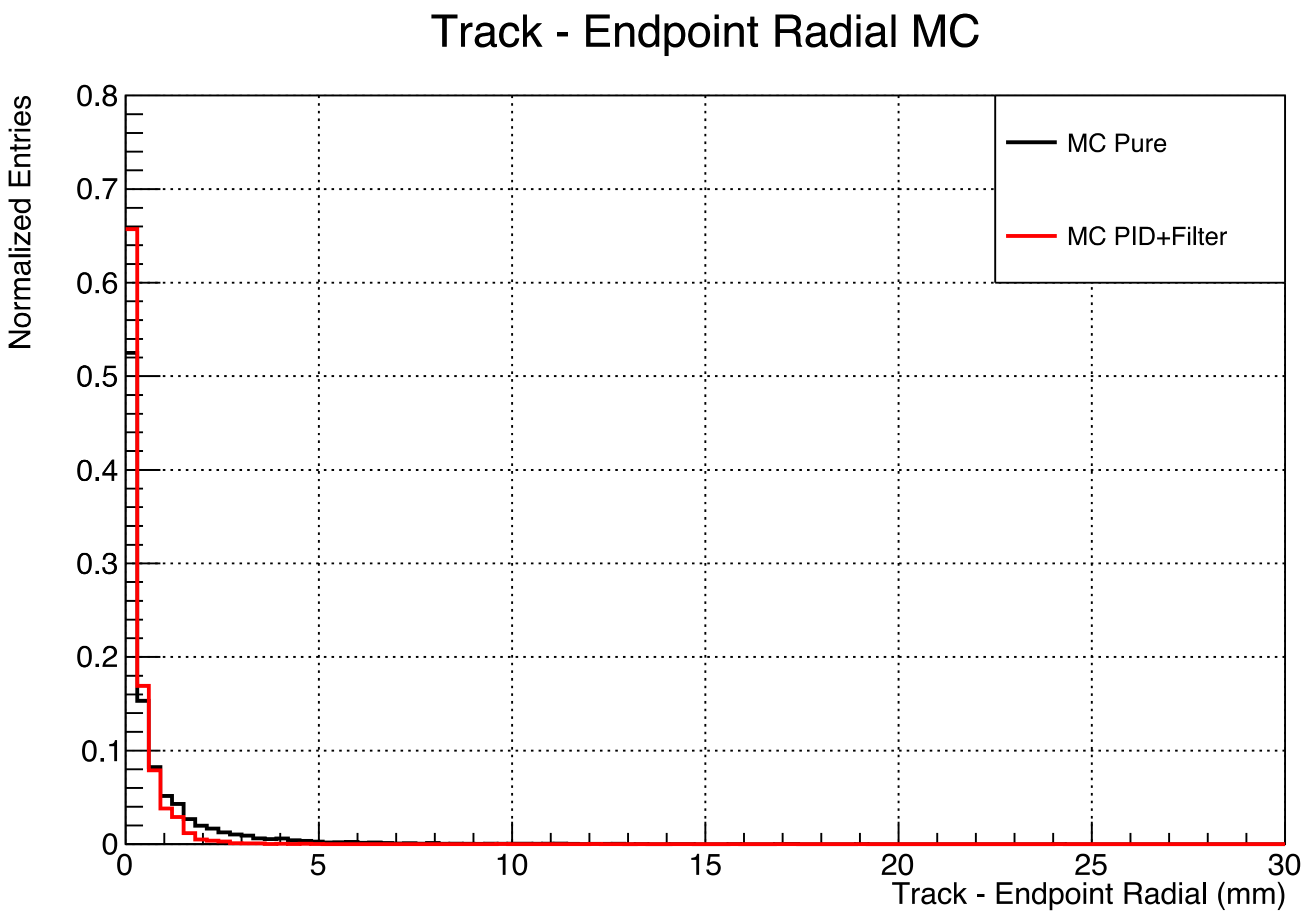
Scintillator Tile Gaps Measurements DWC Example





MC: Track to MC Endpoint Position Comparison

Track Quality Study



r

- 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

Results: Track Agreement 10 GeV Pions

Overview

Data

Agreement [% of events]	Pure (32k)	Hit1+2+3 (18.7k)	Hit1or2or3 (21.1k)	Hit1 (20k)
Track - Event CoG (within 30mm)	76.3	72.2	72.4	72.3
Track - Primary Track Hit Layer 1 (within 22mm)	98.2	98.2	98	98
Track - Closest Hit Layer 1 (within 22mm)	97.1	97.5	97.4	97.4

MC

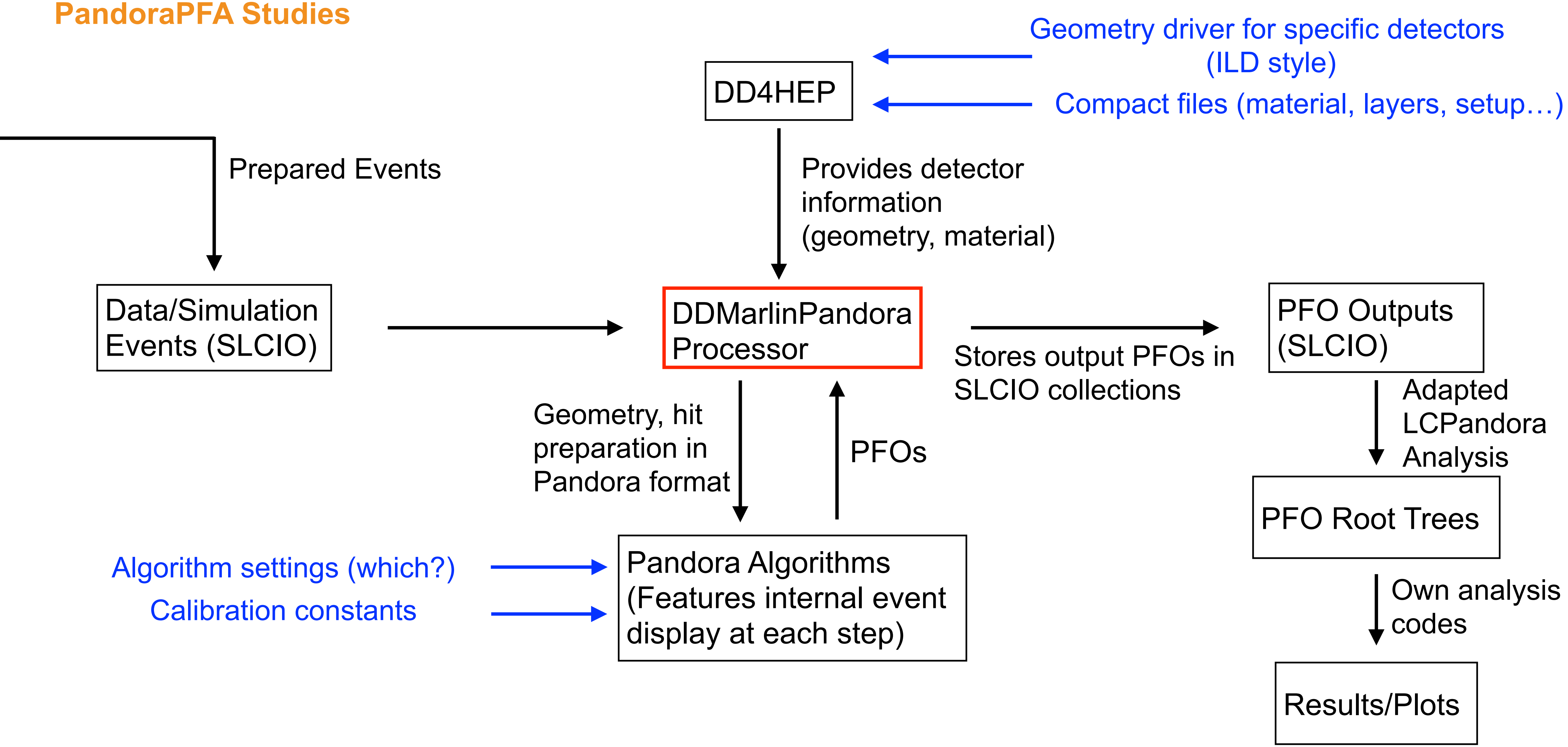
Agreement [% of events]	Pure (46.8k)	Hit1+2+3 (26.7k)	Hit1or2or3 (28.5k)	Hit1 (27.5k)
Track - Event CoG (within 30mm)	72.9	77.2	77.3	77.3
Track - Primary Track Hit Layer 1 (within 22mm)	86.7	99.4	99.2	99.2
Track - Closest Hit Layer 1 (within 22mm)	89.5	98.7	98.6	98.6

- Most of 17.4% MC events with endpoint z < 0 get rejected by PID+filter not depending on hit criterium
- Less strict hit criteria in filter options do not influence fraction of events with good track agreement
  - ➡ Track to CoG (event) agreement worse for 10 GeV pions
  - ➡ Overall good agreement for all filter options and between data and MC
  - ➡ MC tracks on average a bit preciser after PID + Filter also for 10 GeV pions

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:

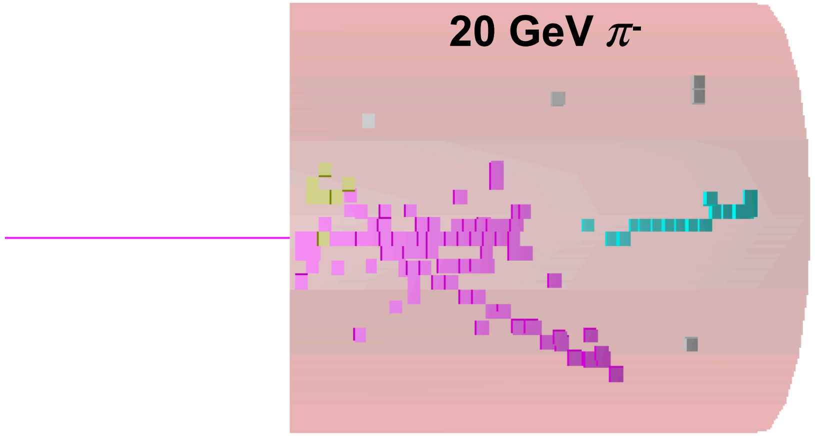
- Simplified detector geometry and related geometry drivers
  - ➡ Careful implementation
- No real tracker, ECAL, muon detector, no B-field
  - ➡ Disable/Re-write related parts code in interface processor
  - ➡ Re-define so-called pseudo layer plugin
  - ➡ Enable algorithm chain step-by-step and check 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

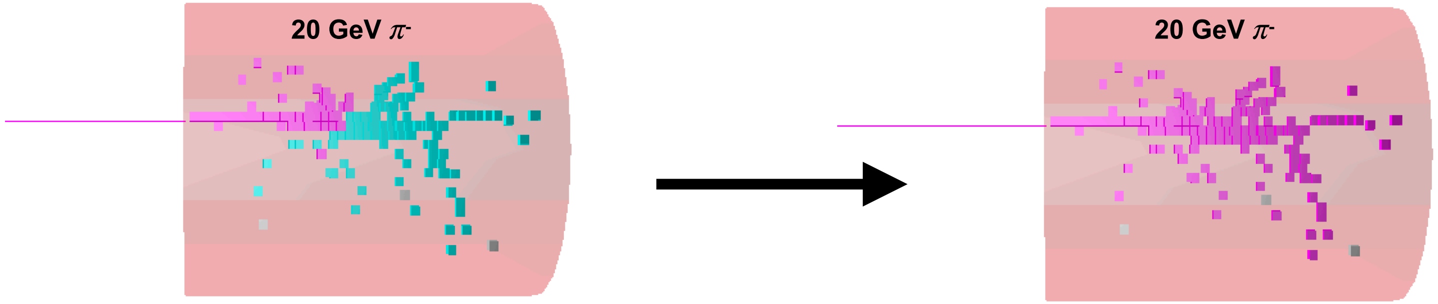
Hits, Clusters & PFOs



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

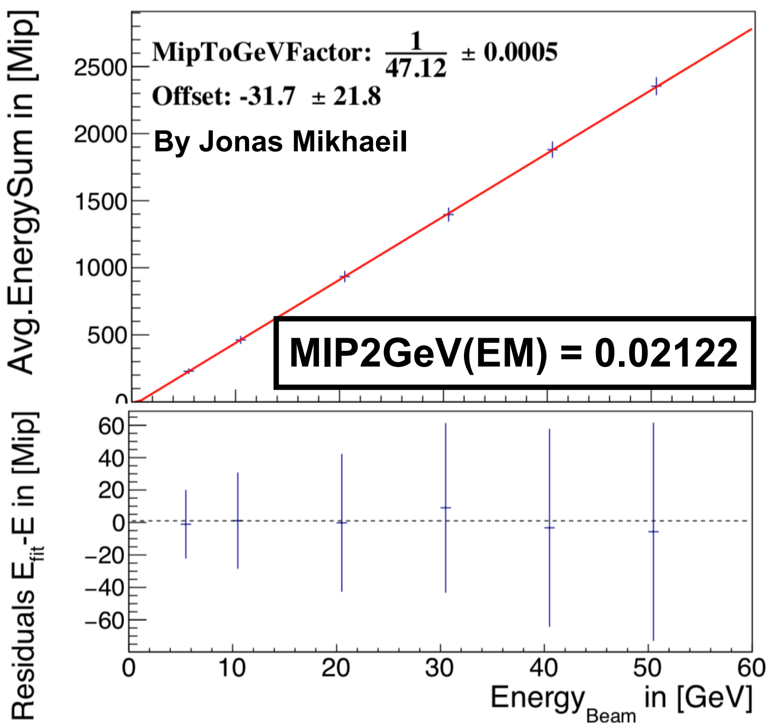


MIP to GeV Conversion

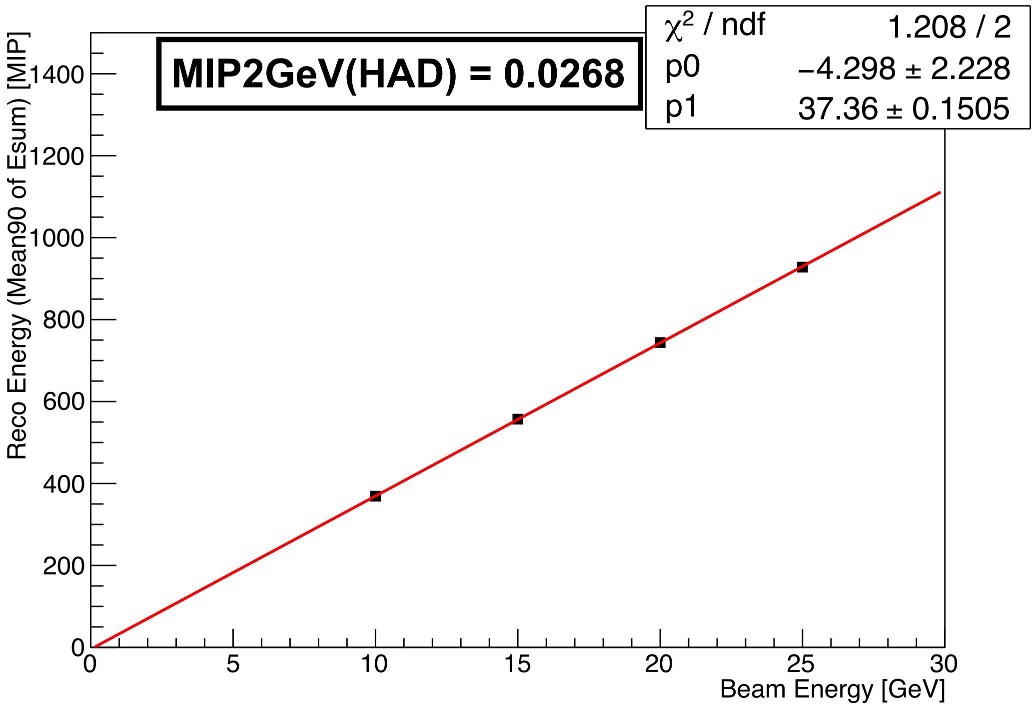
Calibration to EM and HAD Scale

- 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-)



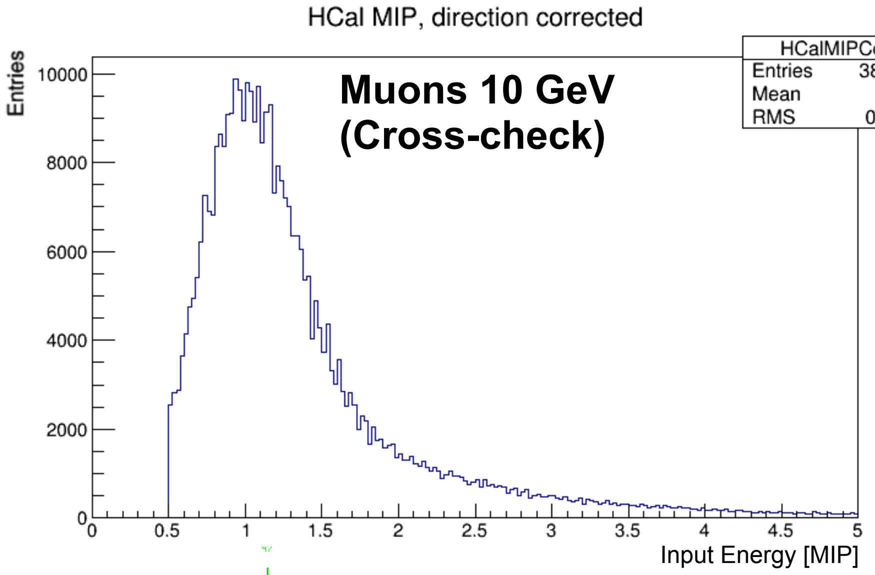
HAD Response Determination (K0L)

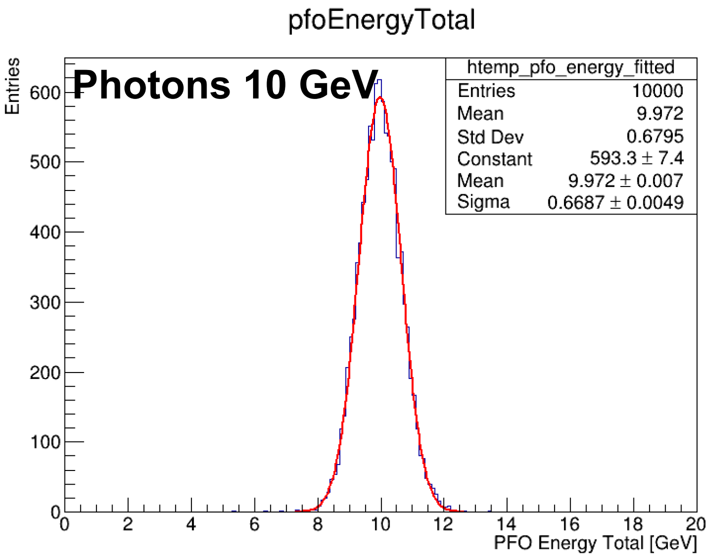


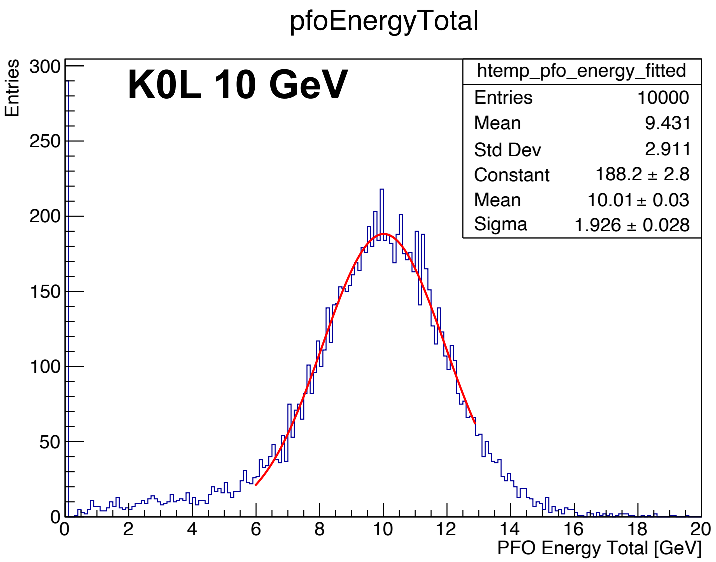
Pandora Energy Calibration

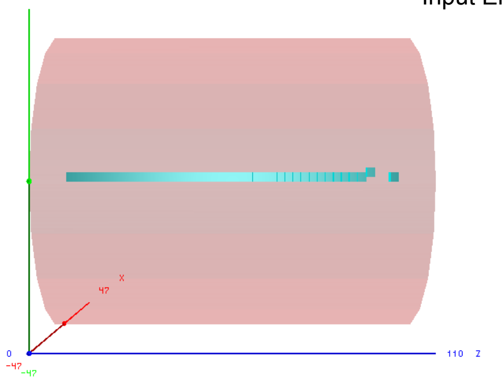
MC Muons, Photons, K0L

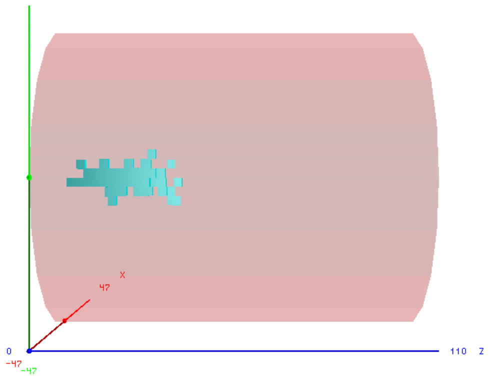
Note: Without tracks and ECAL everything classified as neutral hadrons at this step

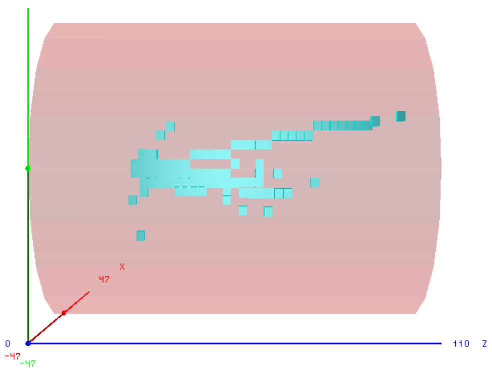










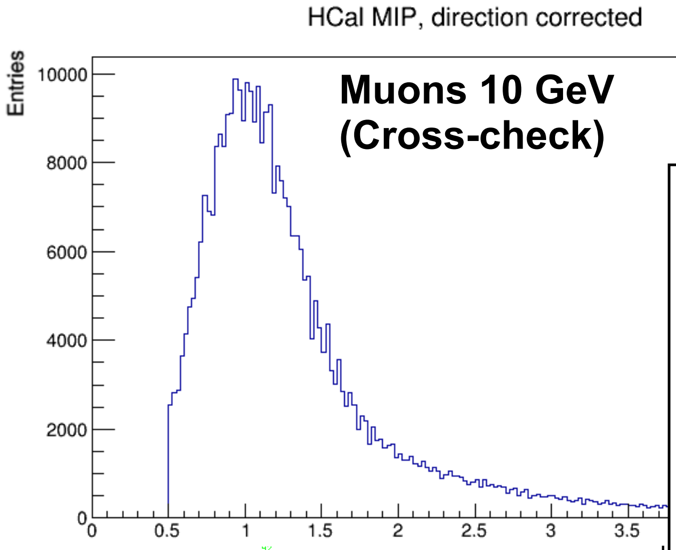


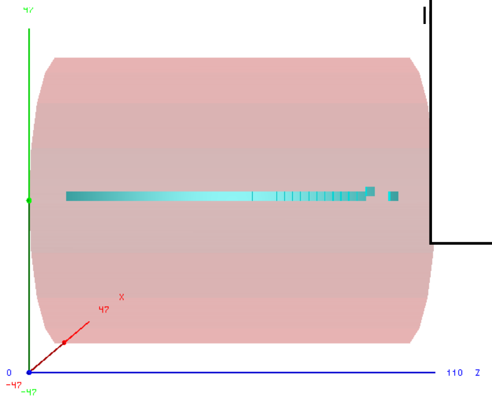
- **Muons**: AHCAL energy GeV -> MIP with negligible angle correction since straight TB tracks
- **Photons and K0L's**: Used to determine EM and HAD response, PFO energy tuned to peak at 10 GeV

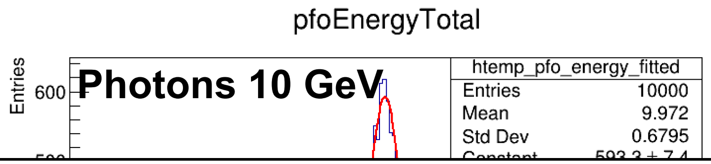
Pandora Energy Calibration

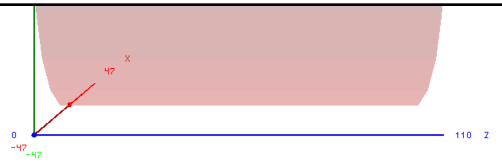
MC Muons, Photons, K0L

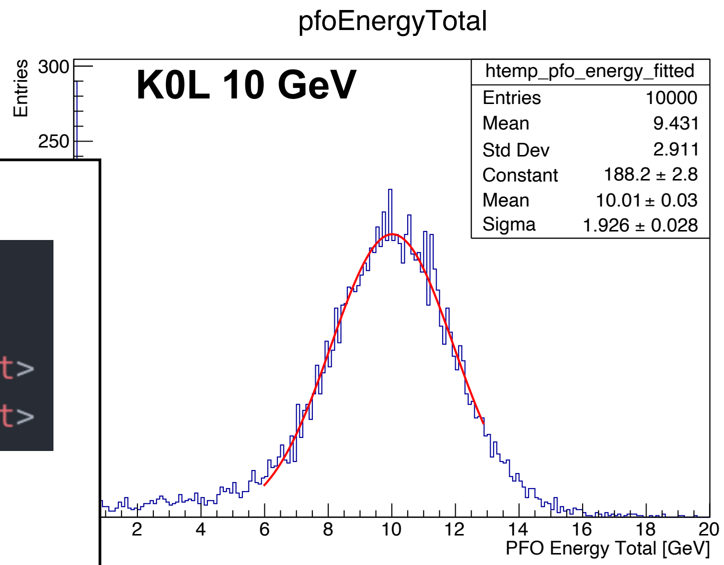
Note: Without tracks and ECAL everything classified as neutral hadrons at this step

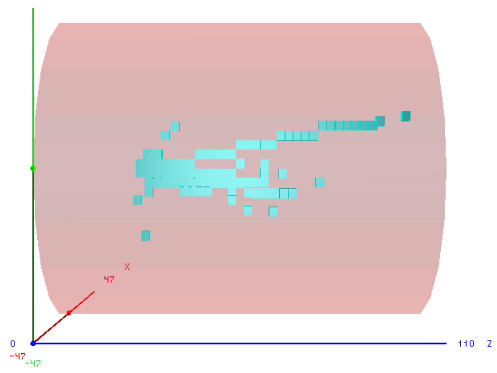












Results:

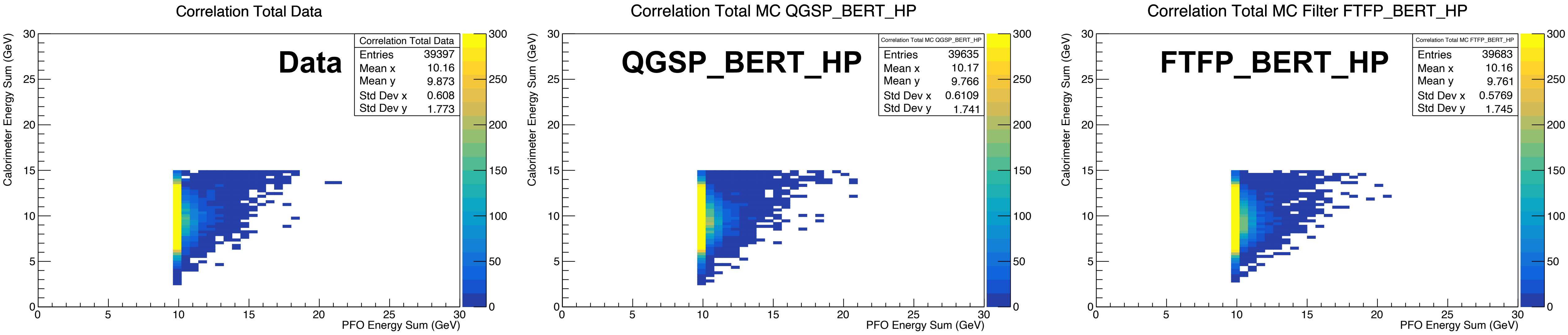
```
<constant name="PandoraHcalToMip">37.36</constant>

<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

PFO Energy Sum to Calorimeter Energy Sum Correlation

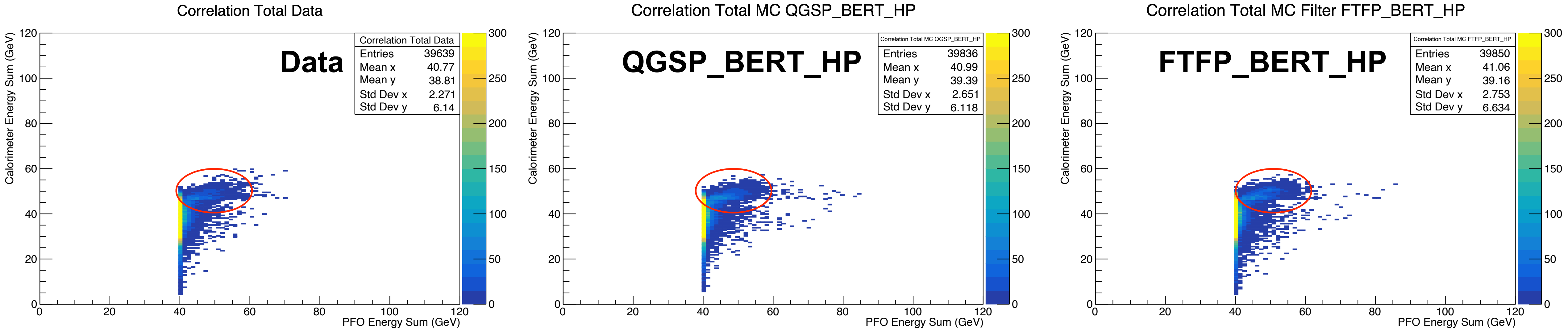
10 GeV



- 10 GeV: Also correlation plots look rather comparable for data and MC

PFO Energy Sum to Calorimeter Energy Sum Correlation

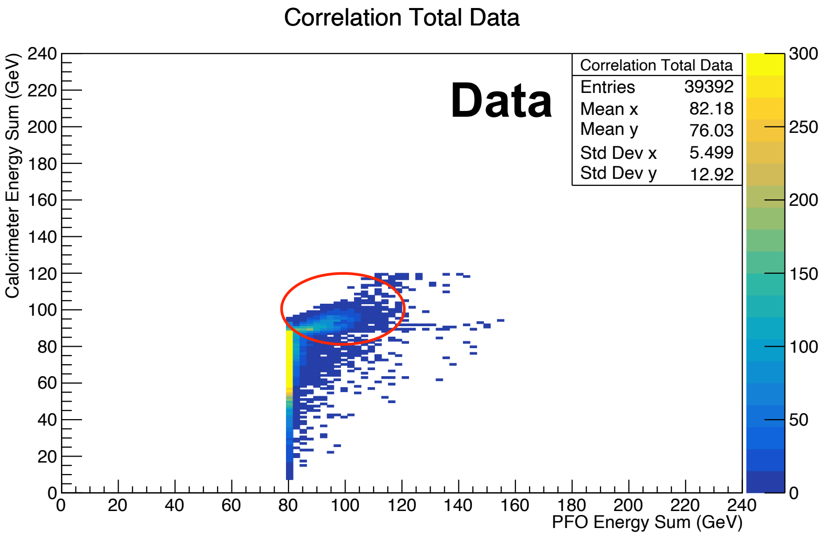
40 GeV

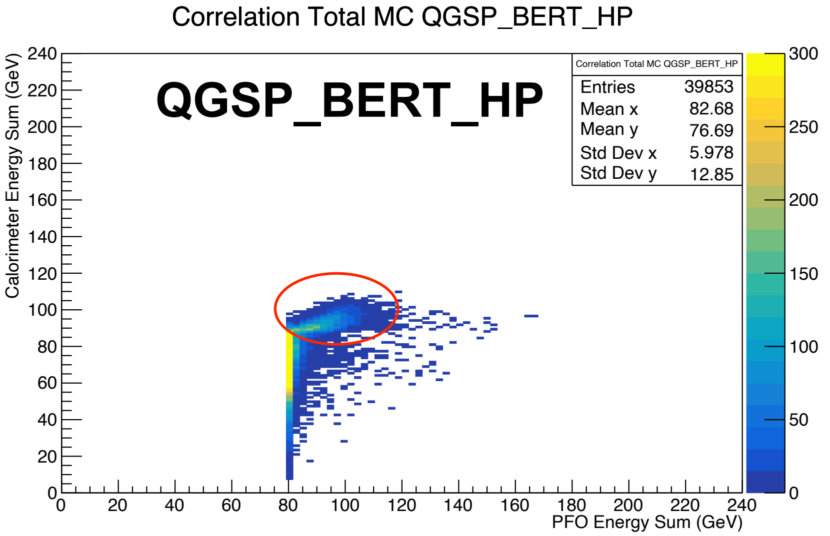


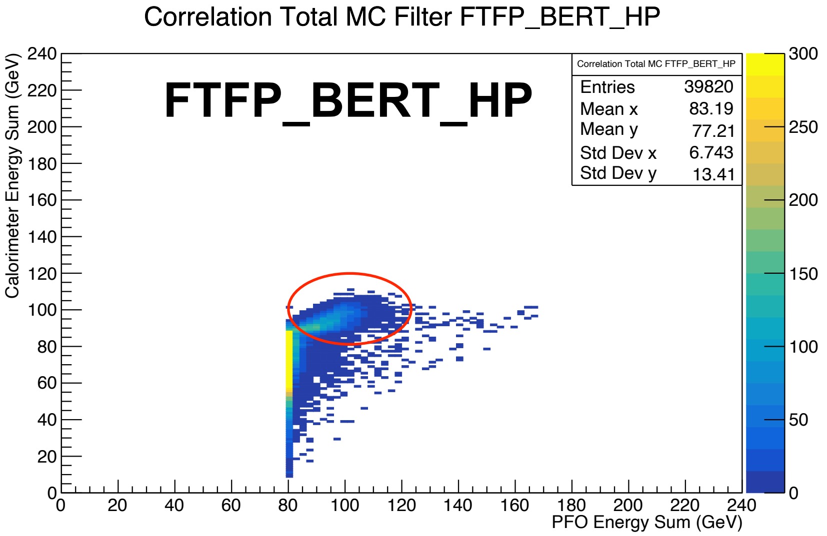
- 10 GeV: Also correlation plots look rather comparable for data and MC
- 40 GeV: Looks also rather comparable, but
  - ➡ Second population starts to appear, especially in MC
  - ➡ Higher fraction of events with multiple PFO's (tail to the right) raising total PFA RMS in MC (FTFP_BERT_HP)

PFO Energy Sum to Calorimeter Energy Sum Correlation

80 GeV





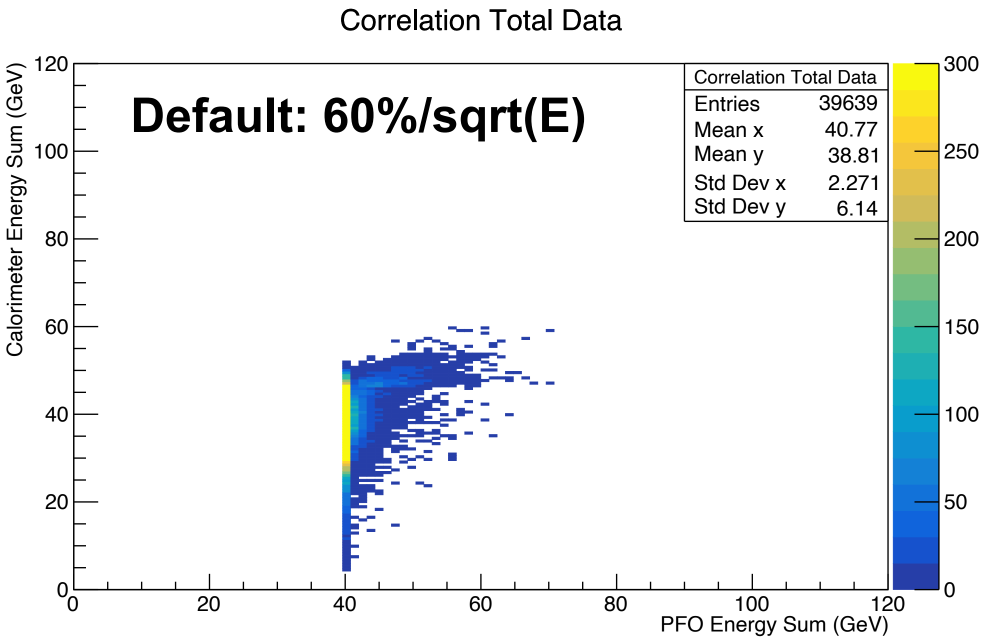


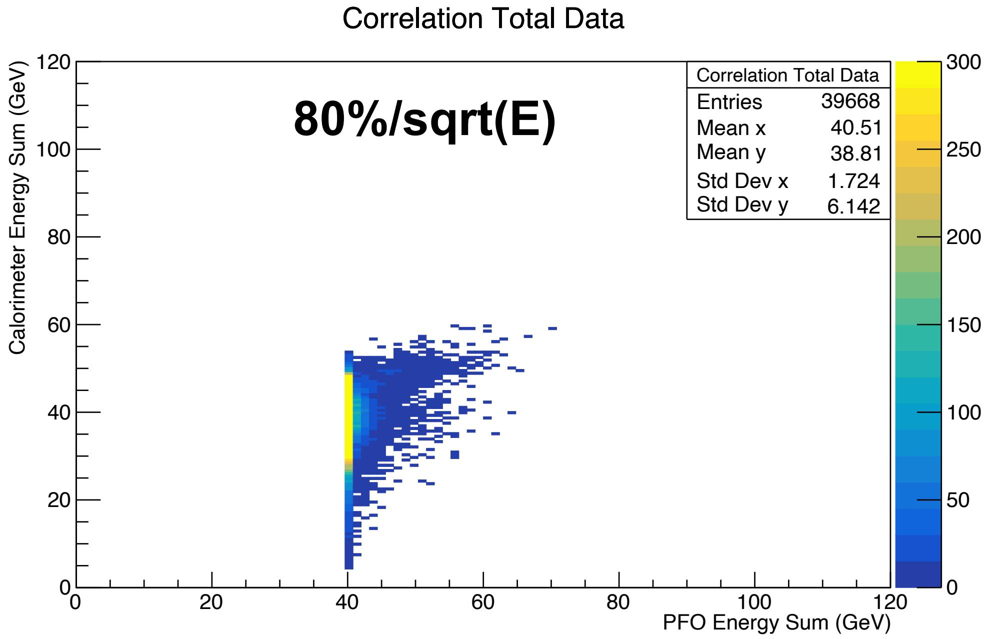
- 10 GeV: Also correlation plots look rather comparable for data and MC
- 40 GeV: Looks also rather comparable, but
  - ➡ Second population starts to appear, especially in MC
  - ➡ Higher fraction of events with multiple PFO's (tail to the right) raising total PFA RMS in MC (FTFP_BERT_HP)
- 80 GeV: Same as for 40 GeV but effects even more dominant

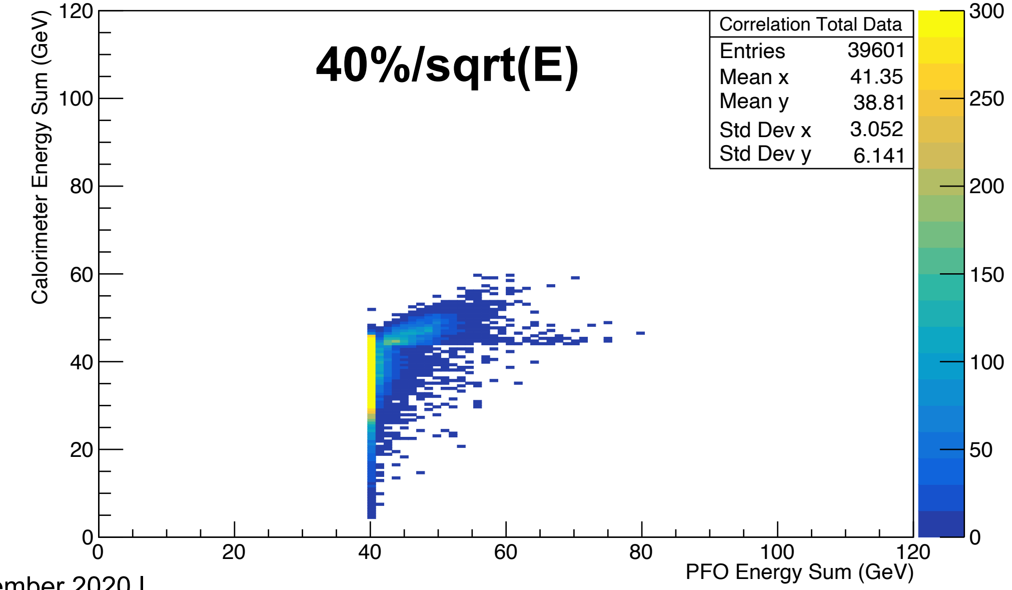
Hadronic Calorimeter Energy Resolution

Influence on „Second Population“ - 40 GeV Data

<HadronicEnergyResolution>x</HadronicEnergyResolution>
in global PandoraSettings.xml







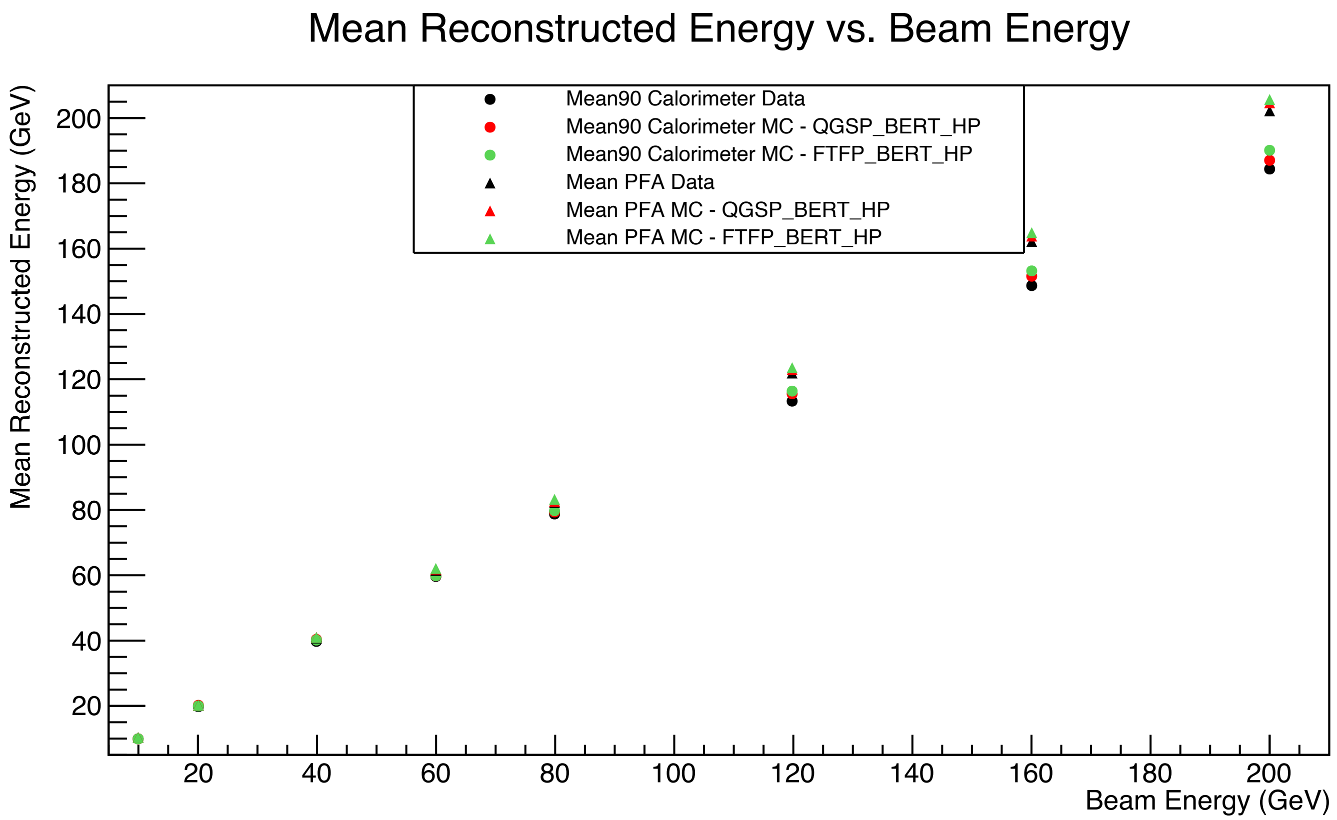
- Assumed **hadronic calorimeter energy resolution** seems to have **large impact** on second population

➡ With less good resolution second populations is vanishing and therefore RMS(PFO) is better

➡ J.M. suggested several more checks on several internal cuts

Mean Energy Sum vs. Beam Energy

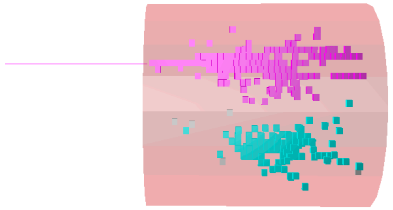
Data and MC



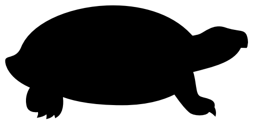
- Calorimeter energy non-linearity mainly originating from leakage (reminder: st < 20 requirement only)
- Linearity seems reasonable for PFA reconstructed energy (FTFP_BERT_HP systematically highest)
  - ➡ Taking track energy is artificially compensating for leakage
  - ➡ Good agreement between classical and PFA reconstruction for low energies

Overlay Processor

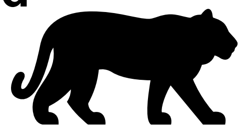
Technical Problems and Solutions



- **Technical goal: Overlay Event A (pseudo neutral hadron) with Event B (charged hadron + track)**
- Technical problems appeared requiring time-consuming implementation for approach of shifting events to profit from more statistics:
  - ➡ What happens to calorimeter hits shifted outside of detector?
    - ➡ Implementation of extra CellIDs in HCAL & Pandora geometry drivers and compact files!
  - ➡ How is the shift performed?
    - ➡ Integer multiple of cell sizes (30mm)? Conflict with Tokyo layer! (60mm cells)
    - ➡ Continuous shifting: Split up energy of channels? Changing topologies?
  - ➡ How to precisely center shower first before performing shift?



- **Go back to original idea of simply using large variety of test beam events (+ MC of course) and search for matching distance pairs within two runs!**

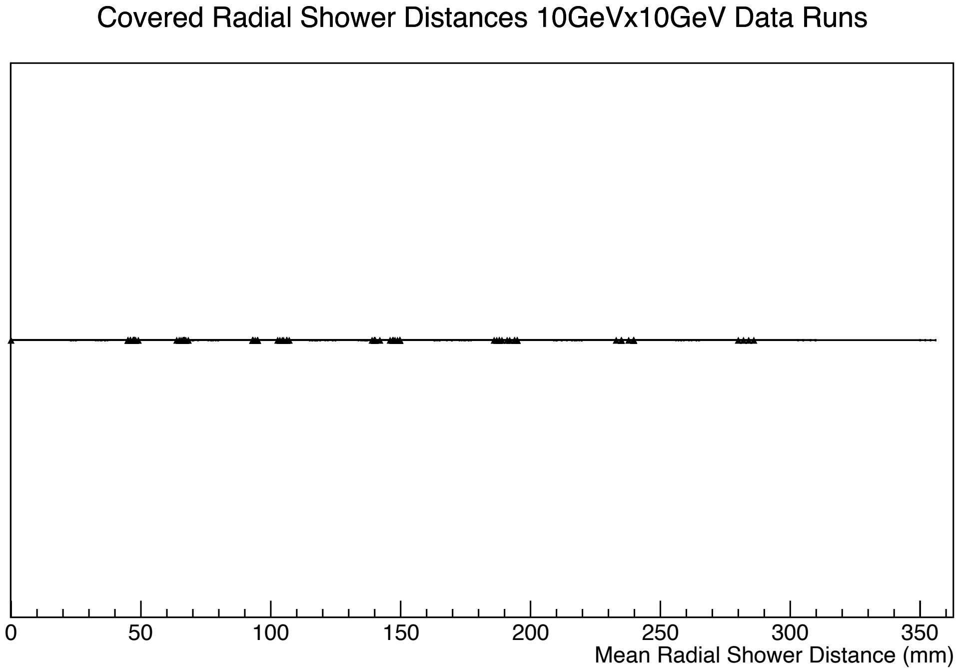


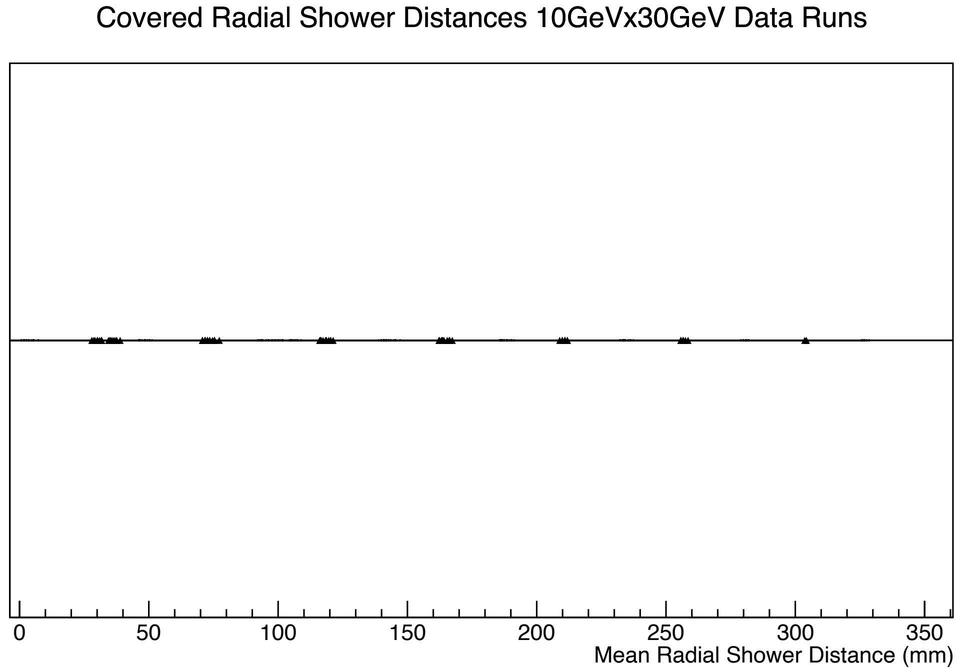
Playing Combinatorics

Estimate/Illustration of Radial Distance Covering

- First goal: 10 GeV (neutrals) overlaid with 10/30 GeV (charged) for different radial shower distances
  - ➡ Studied in first PandoraPFA paper on AHCAL data
- Playing combinatorics with test beam data runs:
  - ➡ Dots: Calculate mean radial difference from mean cogX/Y of all available data run combinations
  - ➡ Error bars: Min-Max propagation with sigma cogX/Y showing overlapping regions with sufficient statistics

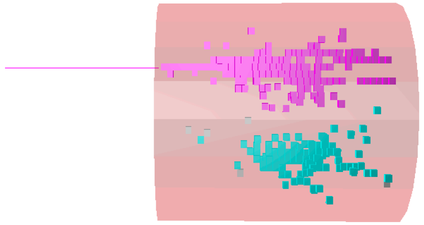
➡ **Full radial shower distance range between 0-350mm covered with test beam data runs!**





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/users/heucheld/repos/calice_analysis/browse/addonProcs/src/MergeProcessor.cc) - **Big thanks to Linghui for great work and synchronisation on that!**

- Former requirements (not available in general ILD version):
  - ➡ Proper flagging of 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 selection according to cogX,Y of shower pairs

- Technical process:
  - ➡ Loop over neutral hadron events - random search for charged hadron event which results in desired radial shower distance -> overlay and save
    - ➡ Currently radial distance agreement range: **+/- 25mm (1 tile total)**
    - ➡ Max attempts for searching matching event within correct distance: **100** (avoid infinite loops for outlier events, processor speed, etc.) -> if not: next neutral hadron event

Overlaid Event Yield

MC: QGSP_BERT_HP, GEANT4.10.3

Validation of Overlay Processor

- Remark: Mean radial distance of used sample combinations very close to desired radial distance of individual pairs for best yield of matching events

Overlaid Event Yield [%]	10GeV+10GeV (50mm)	10GeV+30GeV (50mm)	10GeV+10GeV (200mm)	10GeV+30GeV (200mm)
Data	97.4	96.5	97.9	94.4
MC QGSP_BERT_HP	97.9	96.6	99.5	98.0
MC QGSP_BERT_HP	97.9	96.5	99.4	98.2

- All above 94%: Mainly dependent on exact beam position and profile

Mean #Overlaid Hits for Event Pairs

Validation of Overlay Processor

MC: QGSP_BERT_HP, GEANT4.10.3

Attempts 100, Acceptance range 25mm

<#Overlaid Hits>	10GeV+10GeV (50mm)	10GeV+30GeV (50mm)	10GeV+10GeV (200mm)	10GeV+30GeV (200mm)
Data	5.14	10.9	0.48	0.99
MC QGSP_BERT_HP	5.8	12.2	0.75	1.68
MC FTFP_BERT_HP	5.8	12.3	0.75	1.63

- Decreasing for larger shower distances and increasing with higher energy
- MC seems to have significantly more low energy hits between 0.2 and 0.5 MIP increasing mean
- Also studied: Within overlaid hit energy on average 50/50 from charged/neutral hit

Mean #Overlaid Hits for Event Pairs

Validation of Overlay Processor

MC: QGSP_BERT_HP, GEANT4.10.3

Attempts 10, Acceptance range 15mm

<#Overlaid Hits>	10GeV+10GeV (50mm)	10GeV+30GeV (50mm)	10GeV+10GeV (200mm)	10GeV+30GeV (200mm)
Data (0.2 MIP)	6.7	14.1	0.6	1.2
MC (0.2 MIP)	7.6	16.1	0.9	2.0
Data (0.5 MIP)	6.1	12.8	0.4	1.0
MC (0.5 MIP)	5.6	12.0	0.5	1.1

- Decreasing for larger shower distances and increasing with higher energy
- MC seems to have significantly more low energy hits between 0.2 and 0.5 MIP increasing mean
- Also studied: Within overlaid hit energy on average 50/50 from charged/neutral hit

Mean #Low Energy Hits for Event Pairs

Validation of Overlay Processor

MC: QGSP_BERT_HP, GEANT4.10.3

<#Low Energy (<0.5 MIP) Hits>	10GeV+10GeV (50mm)	10GeV+30GeV (50mm)	10GeV+10GeV (200mm)	10GeV+30GeV (200mm)
Data	15.2	24.2	15.4	24.9
MC QGSP_BERT_HP	50.9	84.7	49.5	81.7
MC FTFP_BERT_HP	50.3	85.7	49.5	81.5

- Proof: MC has significantly more low energy hits between 0.2 and 0.5 MIP especially for higher energies

Mean #Cut Hits (< 0.5 MIP) for Overlaid Event

Validation of Overlay Processor

MC: QGSP_BERT_HP, GEANT4.10.3

<#Cut (<0.5 MIP) Hits> [%]	10GeV+10GeV (50mm)	10GeV+30GeV (50mm)	10GeV+10GeV (200mm)	10GeV+30GeV (200mm)
Data	14.7	23.2	15.3	24.8
MC QGSP_BERT_HP	49.3	81.4	49.2	80.8
MC FTFP_BERT_HP	48.8	82.3	49.2	80.7

- **Conclusion: Most (98% Data, 95% MC) of low energy hits are cut away after overlay by 0.5 MIP threshold**

Mean #New Hits (>0.5 MIP) for Overlaid Event

Validation of Overlay Processor

MC: QGSP_BERT_HP, GEANT4.10.3

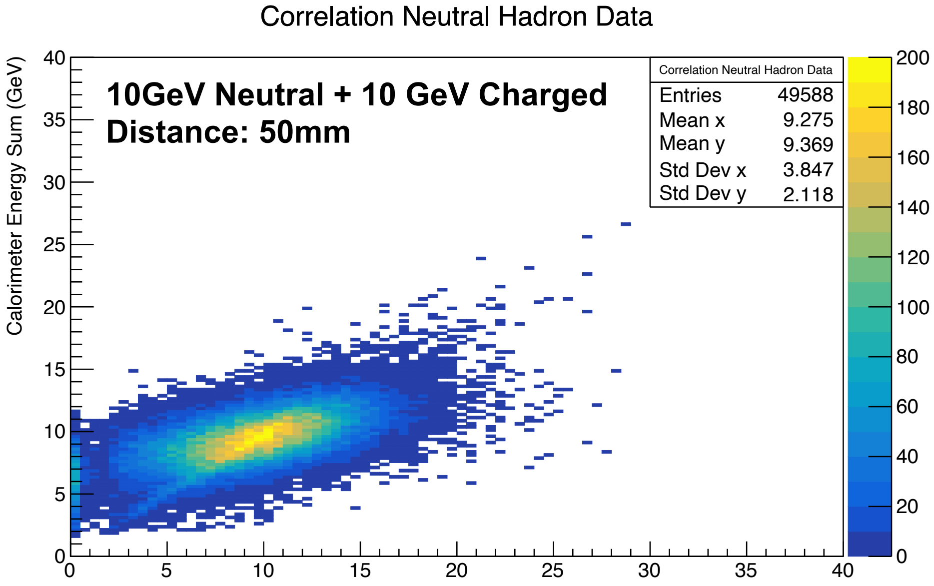
- Question: How many new hits are originating on average from overlaid low energy hits (<0.5 MIP) reaching the 0.5 MIP threshold

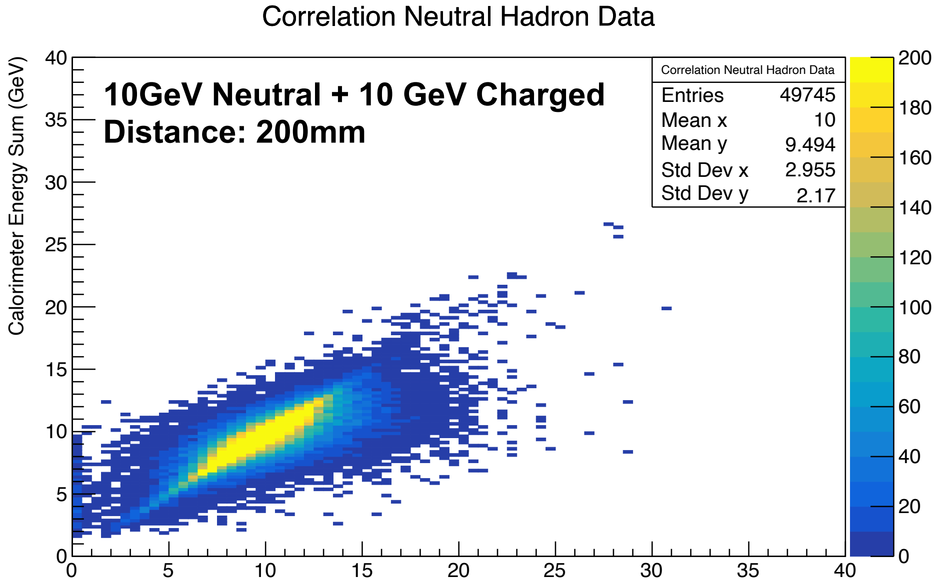
<#New (>0.5 MIP) Hits>	10GeV+10GeV (50mm)	10GeV+30GeV (50mm)	10GeV+10GeV (200mm)	10GeV+30GeV (200mm)
Data	0,015	0,027	0,004	0,008
MC QGSP_BERT_HP	0.1	0,22	0,03	0,08
MC FTFP_BERT_HP	0.1	0,23	0,04	0,08

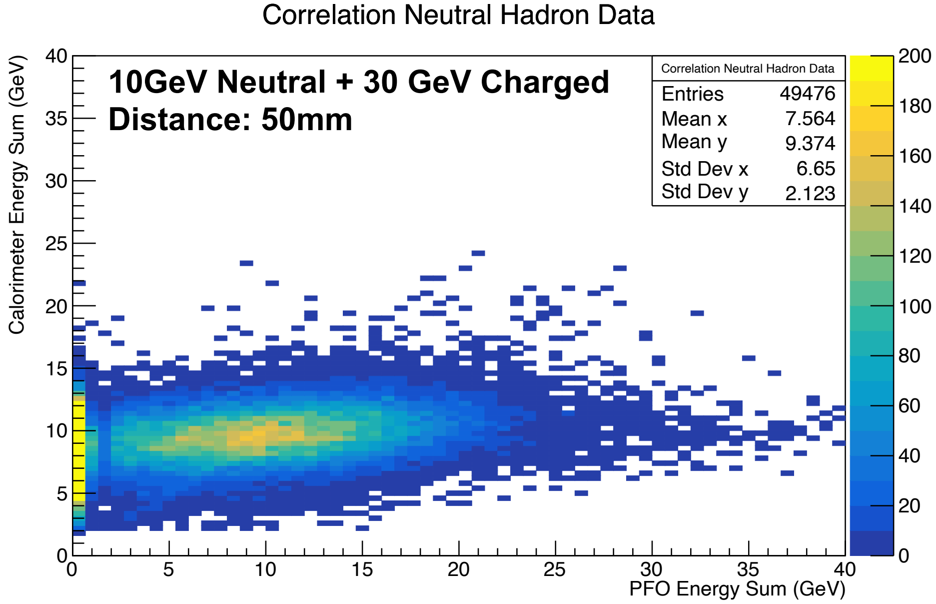
- **Conclusion: Negligible for all scenarios**

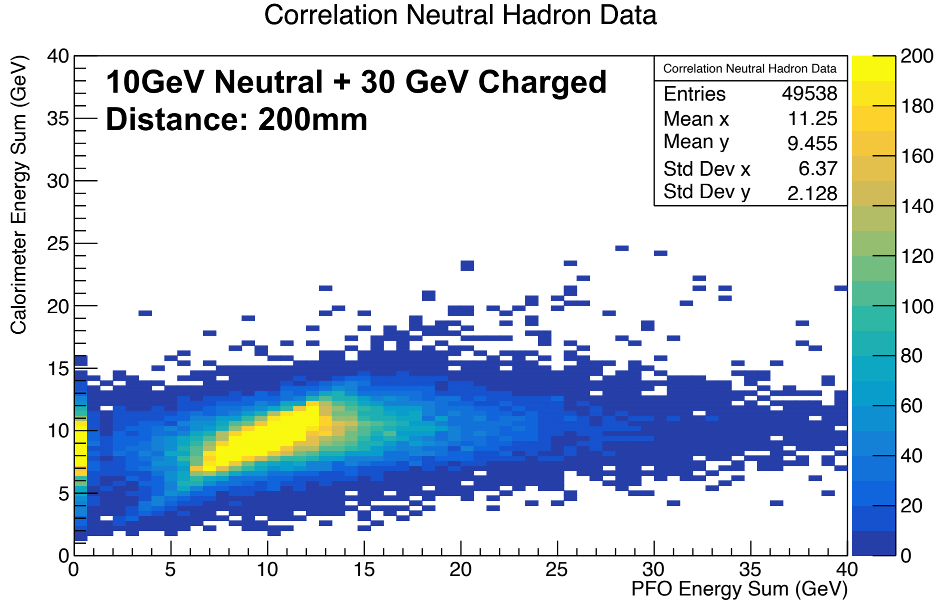
Correlation PFA vs. Calorimeter Energy Sum Neutral Hadron

Data



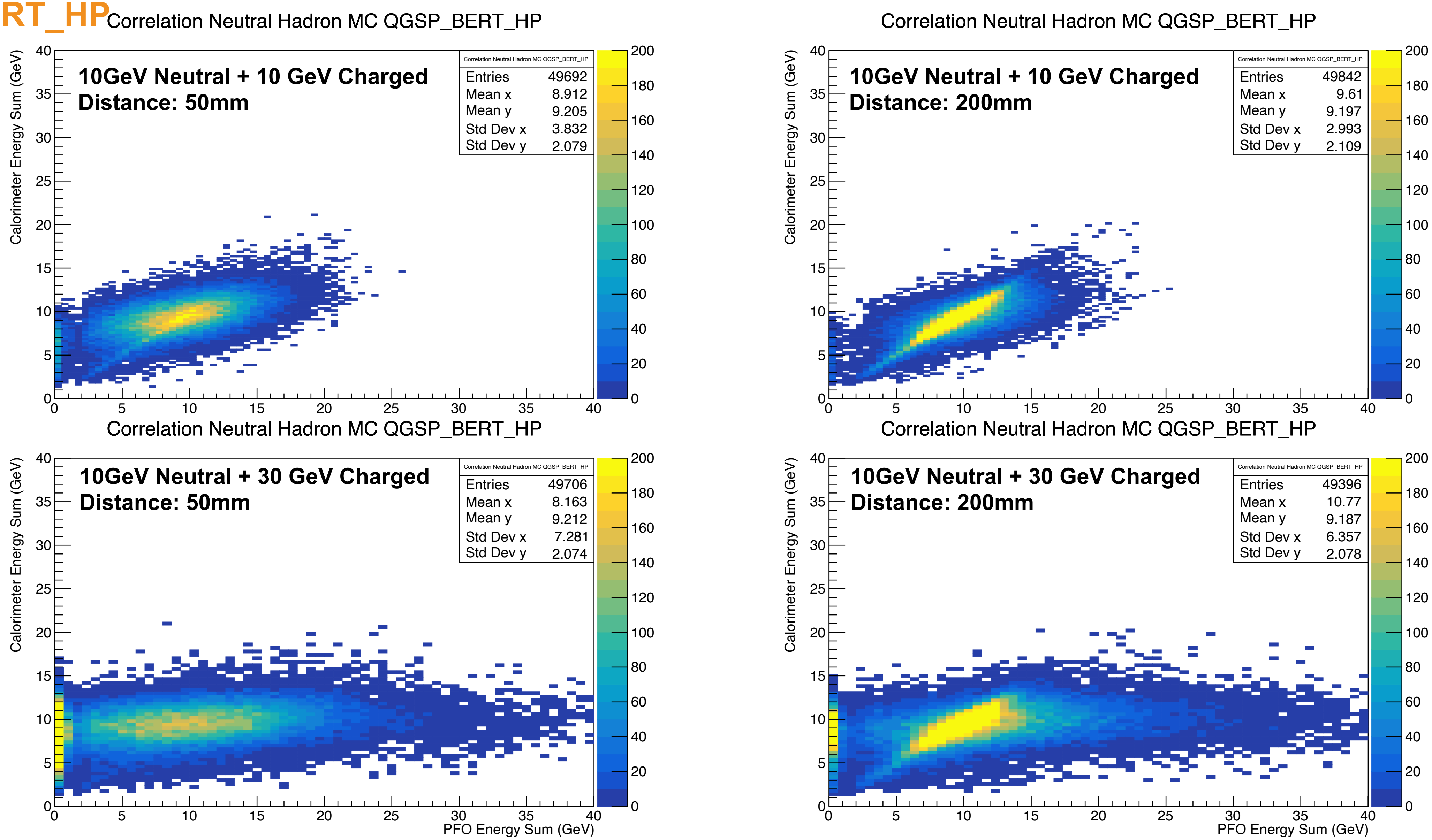






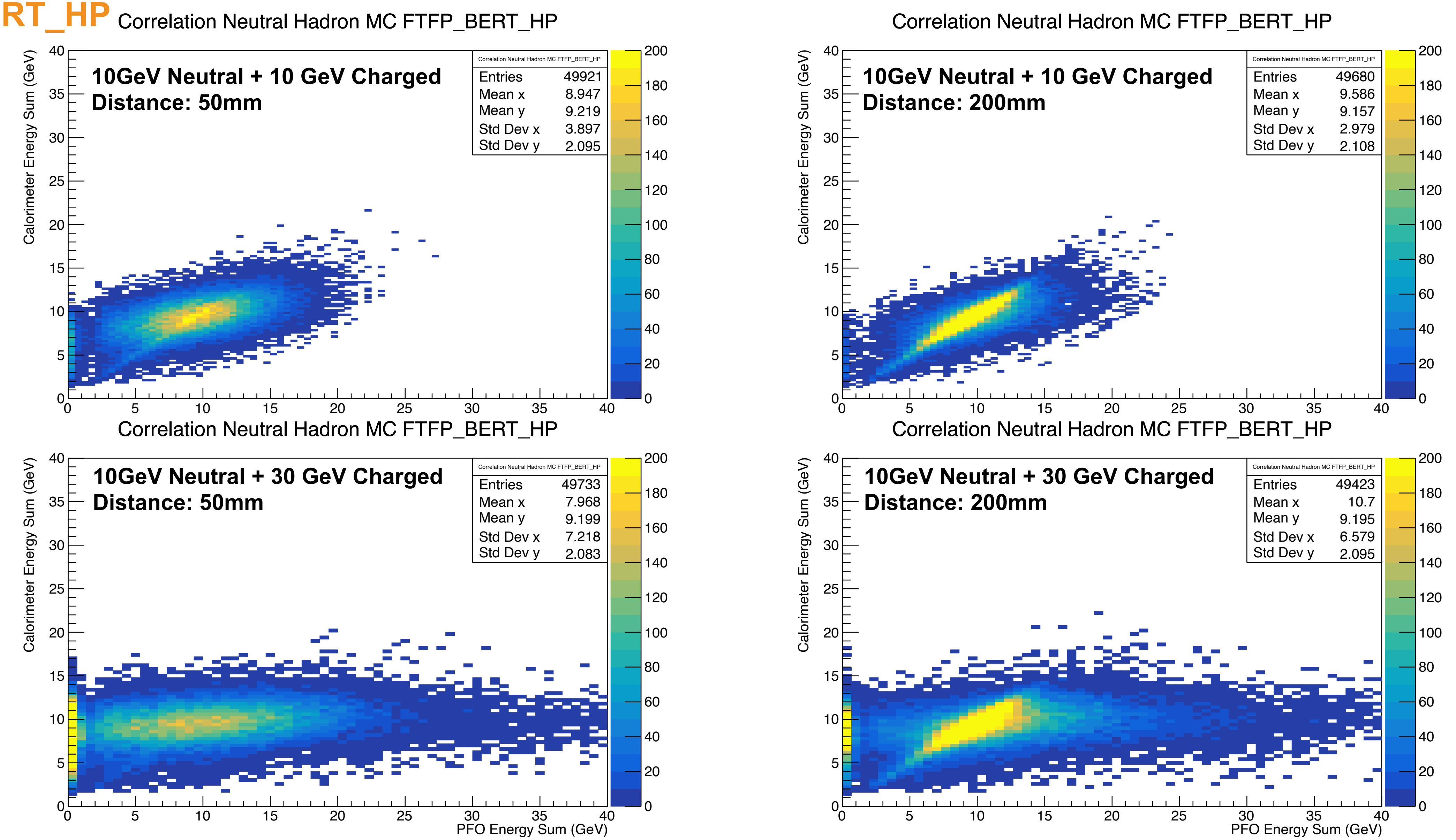
Correlation PFA vs. Calorimeter Energy Sum Neutral Hadron

QGSP_BERT_HP



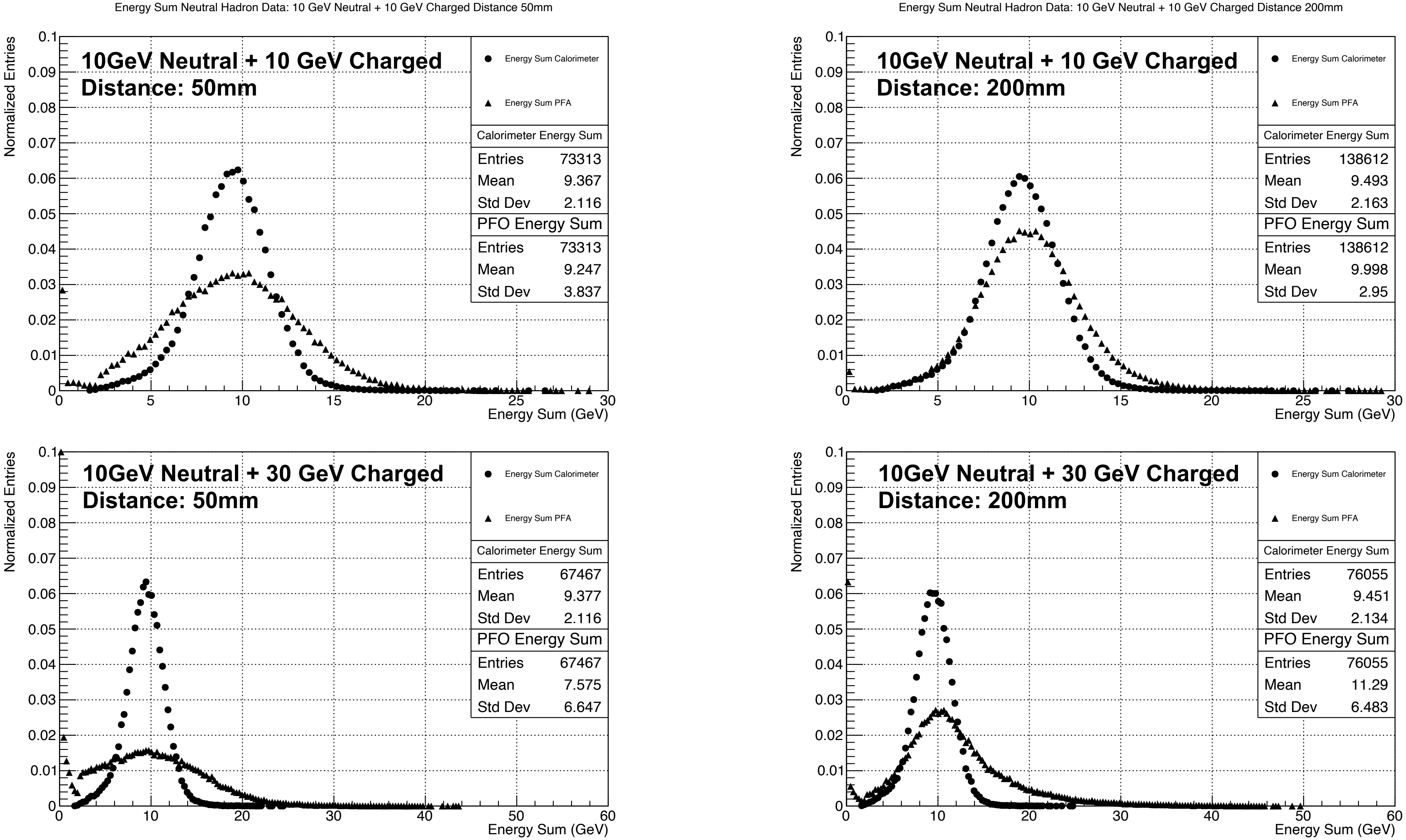
Correlation PFA vs. Calorimeter Energy Sum Neutral Hadron

FTFP_BERT_HP



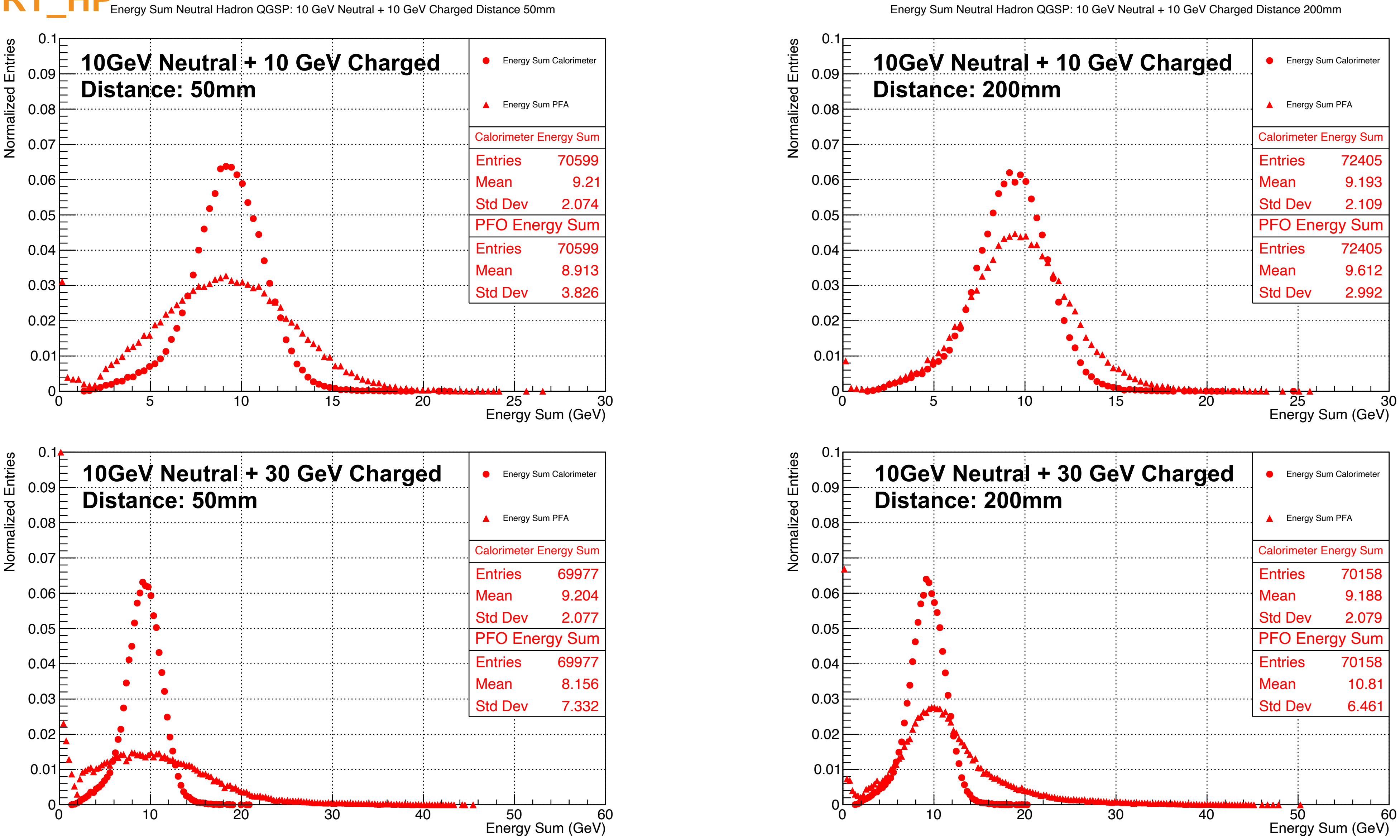
Projection PFA vs. Calorimeter Energy Sum Neutral Hadron

Data



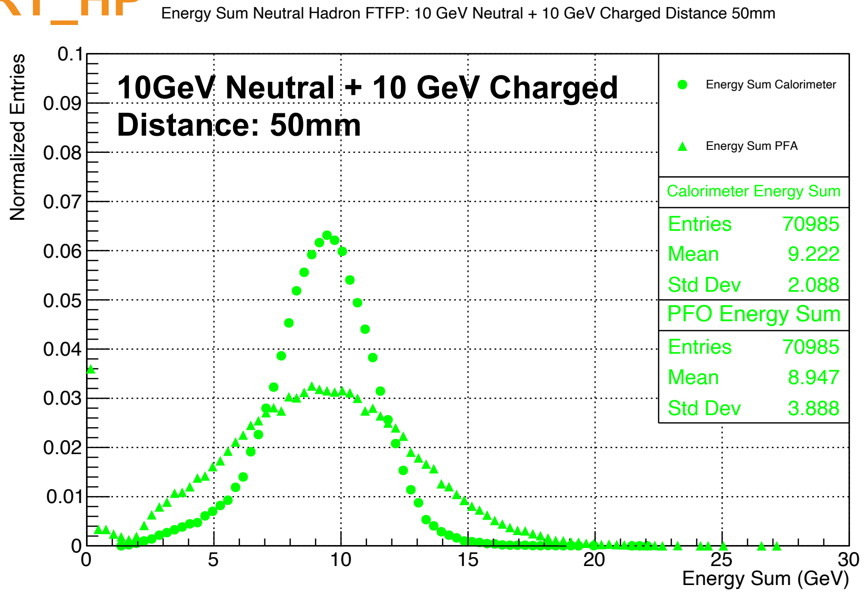
Projection PFA vs. Calorimeter Energy Sum Neutral Hadron

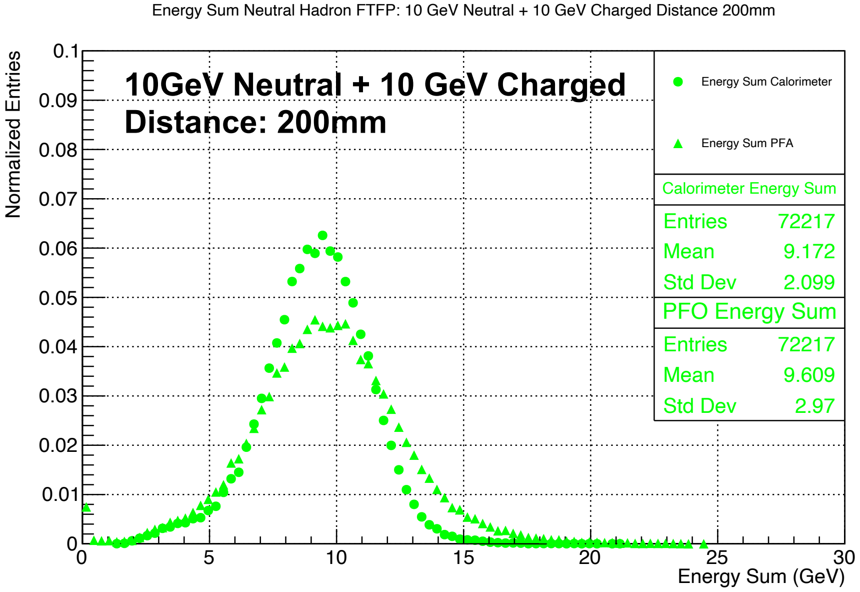
QGSP_BERT_HP

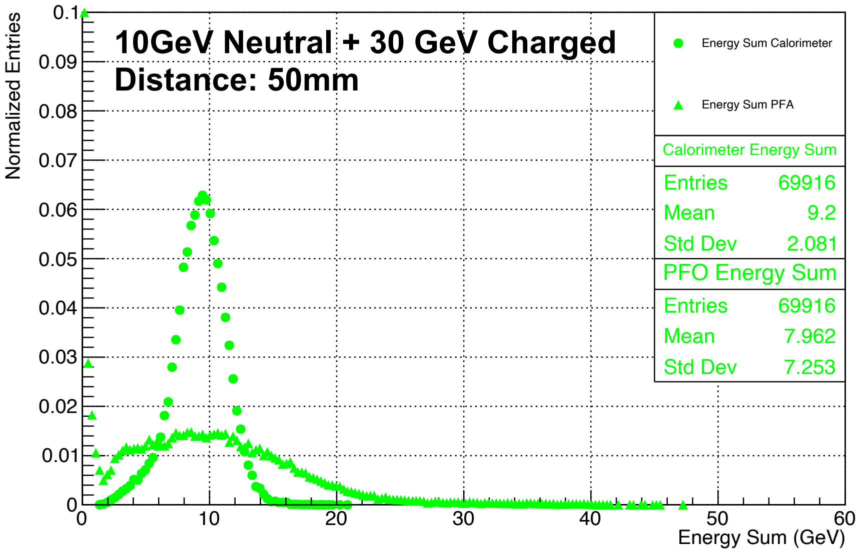


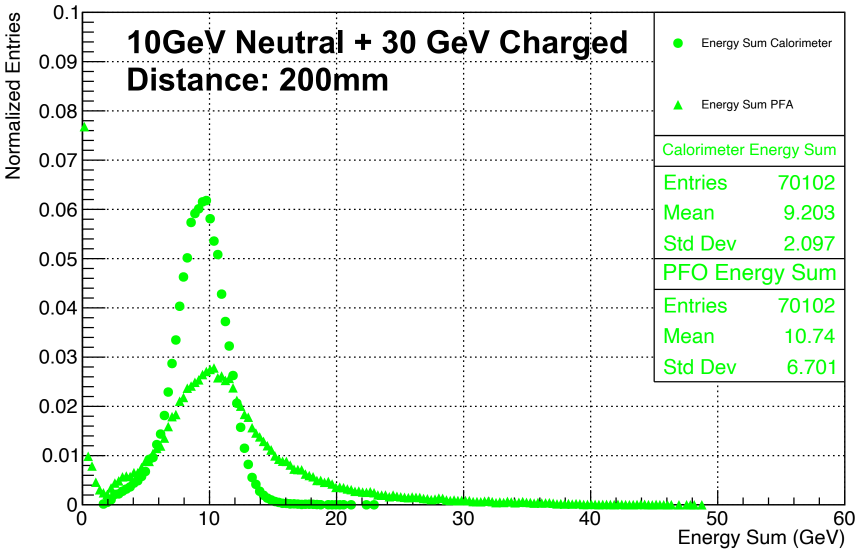
Projection PFA vs. Calorimeter Energy Sum Neutral Hadron

FTFP_BERT_HP



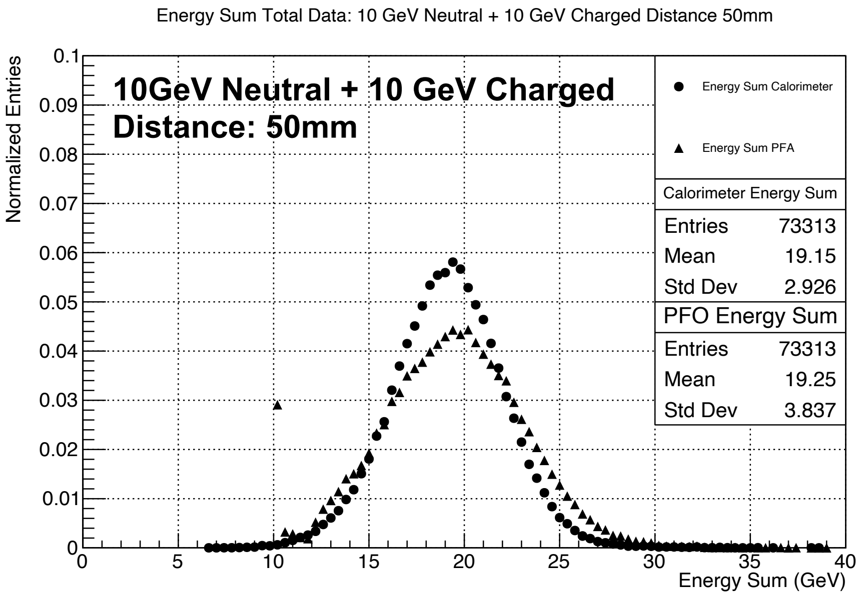


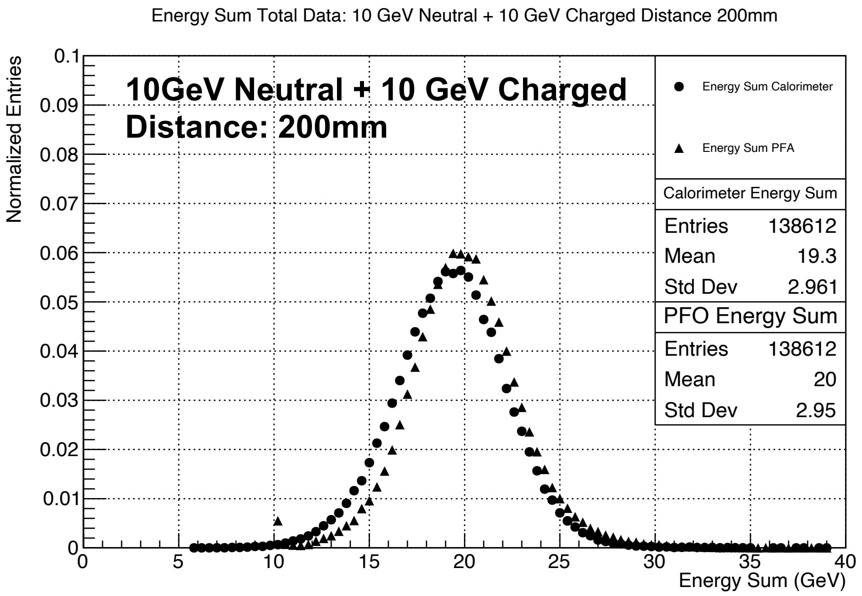


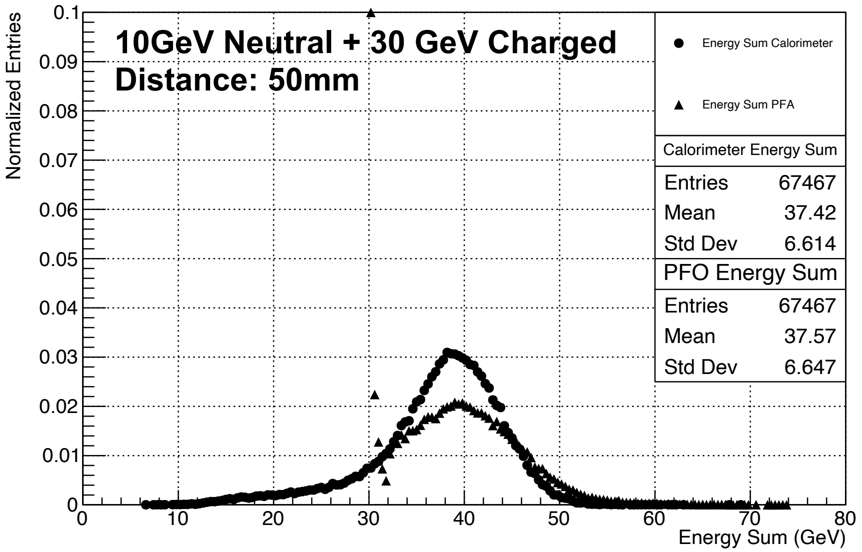


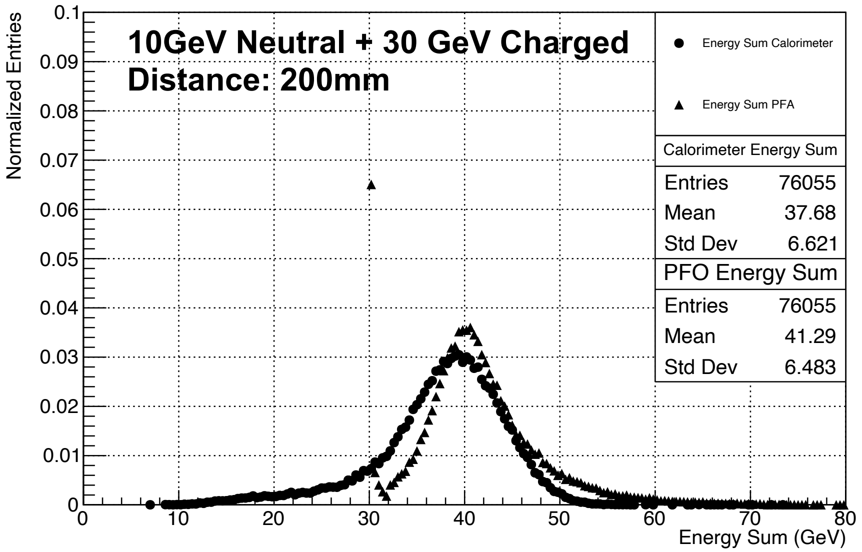
Projection PFA vs. Calorimeter Energy Sum Total

Data



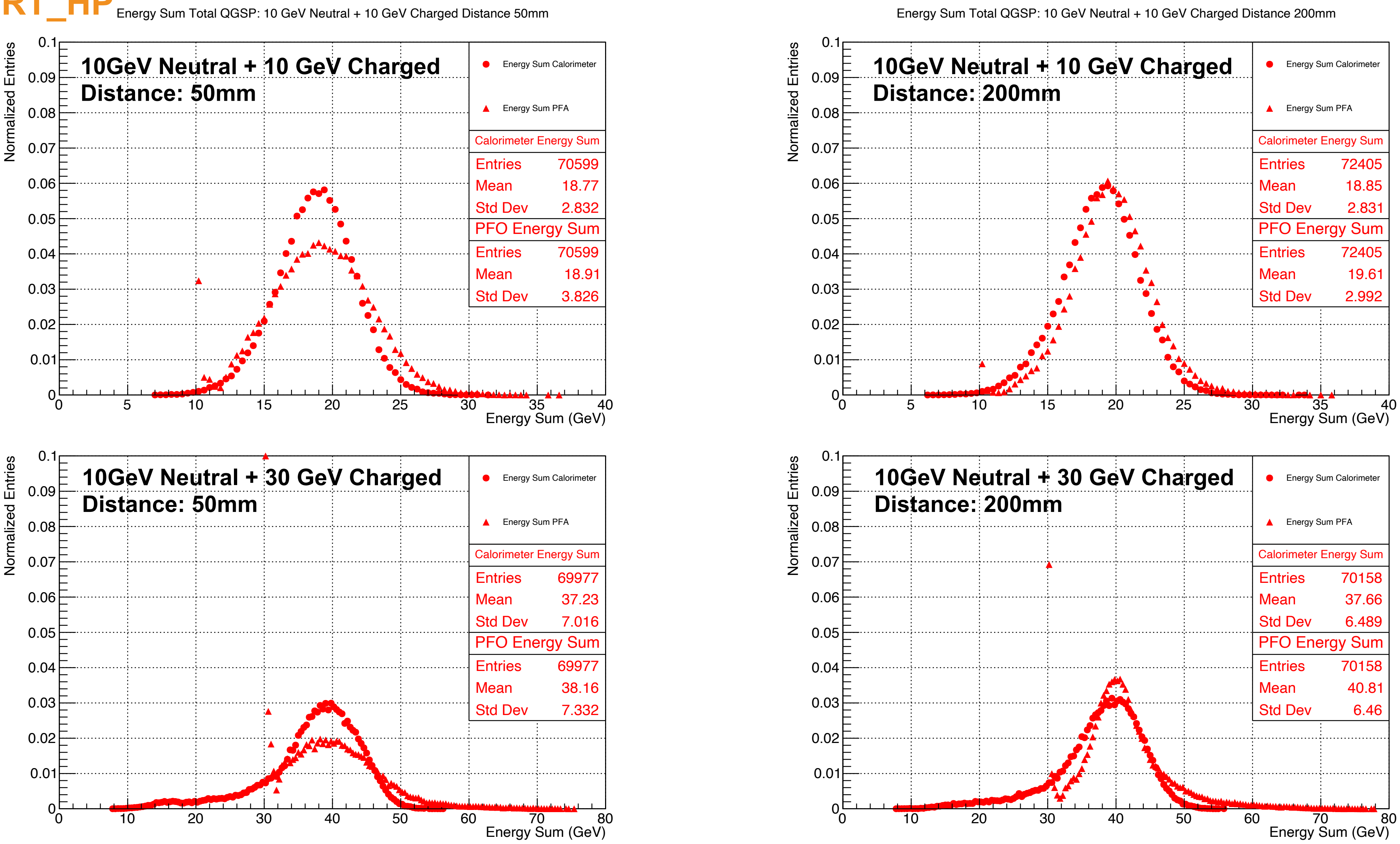






Projection PFA vs. Calorimeter Energy Sum Total

QGSP_BERT_HP



Projection PFA vs. Calorimeter Energy Sum Total

FTFP_BERT_HP