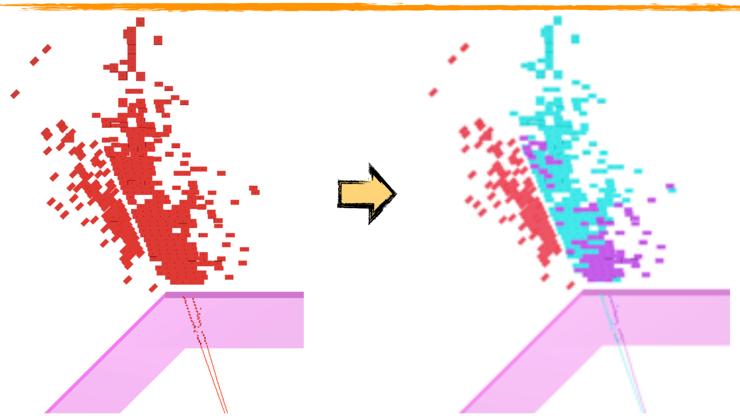
# Calorimetry:

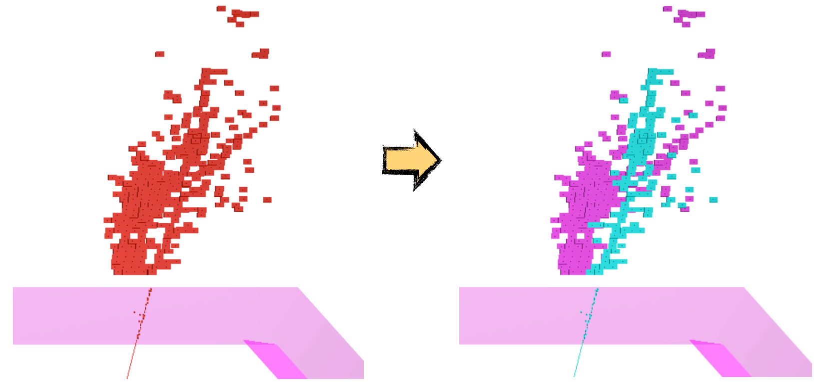
# Validation and Performance of PandoraPFA

Matthias Weber (CERN)

# PandoraPFA algorithm



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

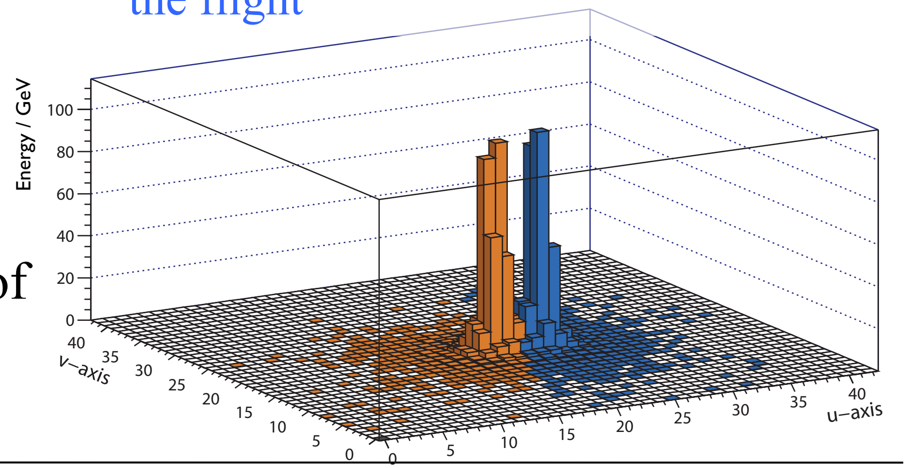


2. Cluster energy much greater than track momentum – split cluster.

- Exploit calorimeter granularity to gradually build up picture of events
- More than 70 algorithms
  - → address different topologies,
  - → correct identification
  - → avoid accidental splitting and merging of particles

NIM A 700 (2013) 153

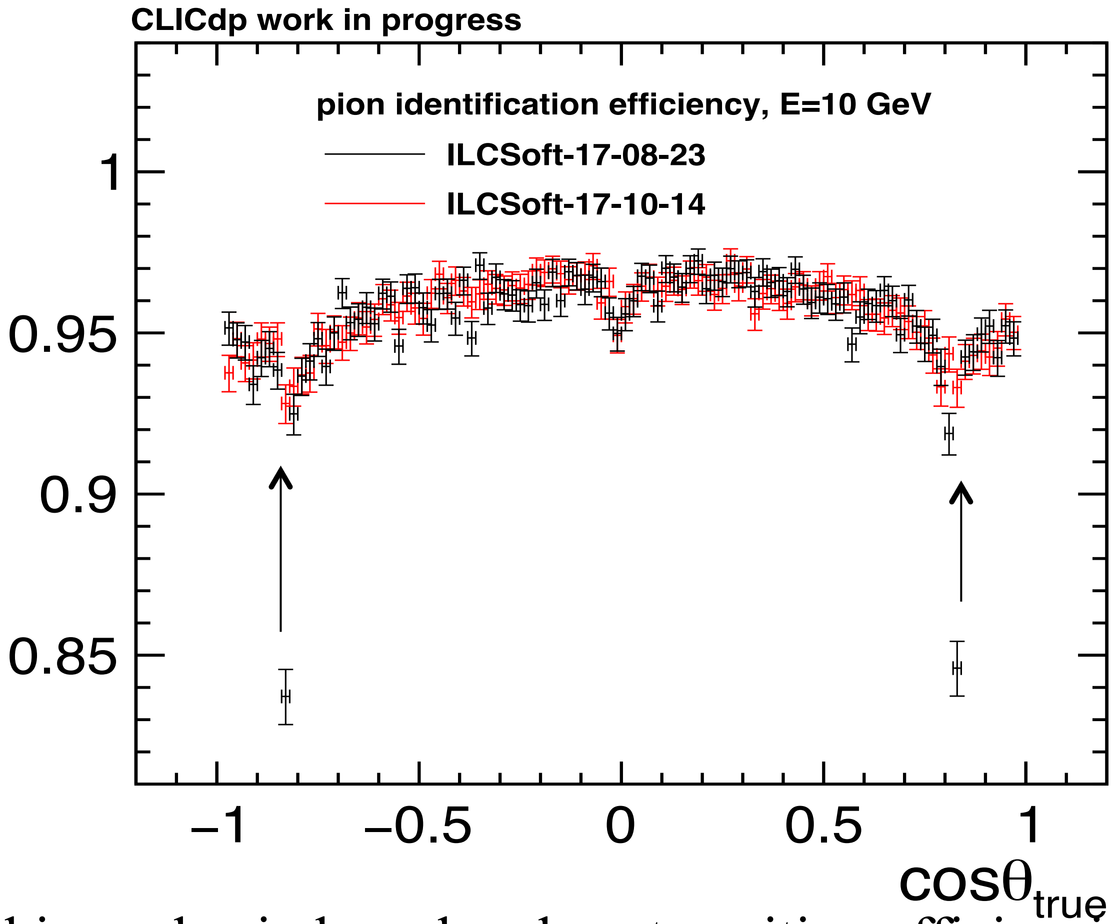
Energy of two nearby photons in transverse plane to the direction of the flight



# Single Particle Identification

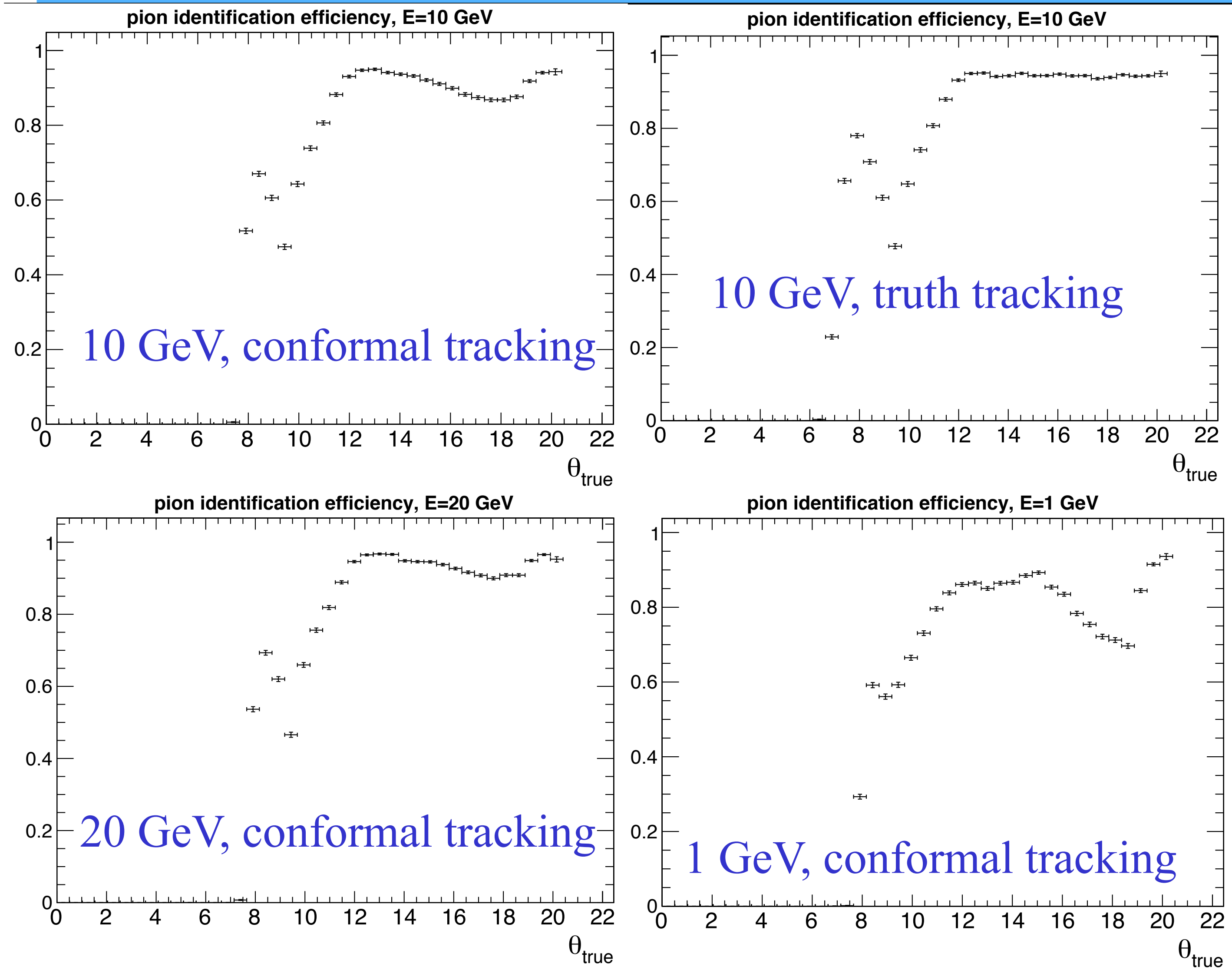
# Pion Identification Efficiency

Studied on isolated single pion events
Example of recent developments in PandoraPFA algorithms:
Modified track cluster matching algorithm in barrel-endcap calorimeter transition




➡ Now achieve also in barrel endcap transition efficiencies beyond 93 %

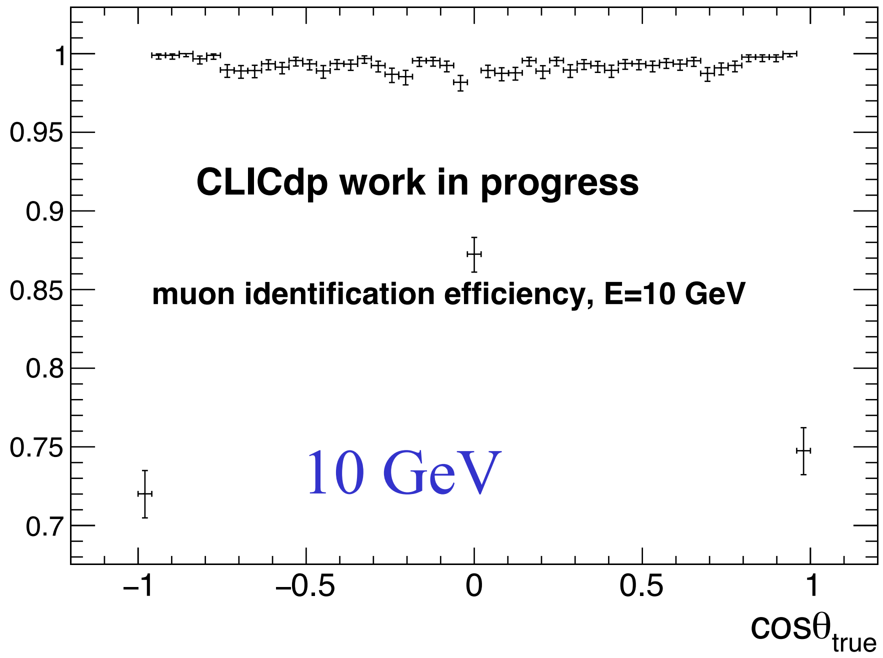
# Forward Pion Identification Efficiency

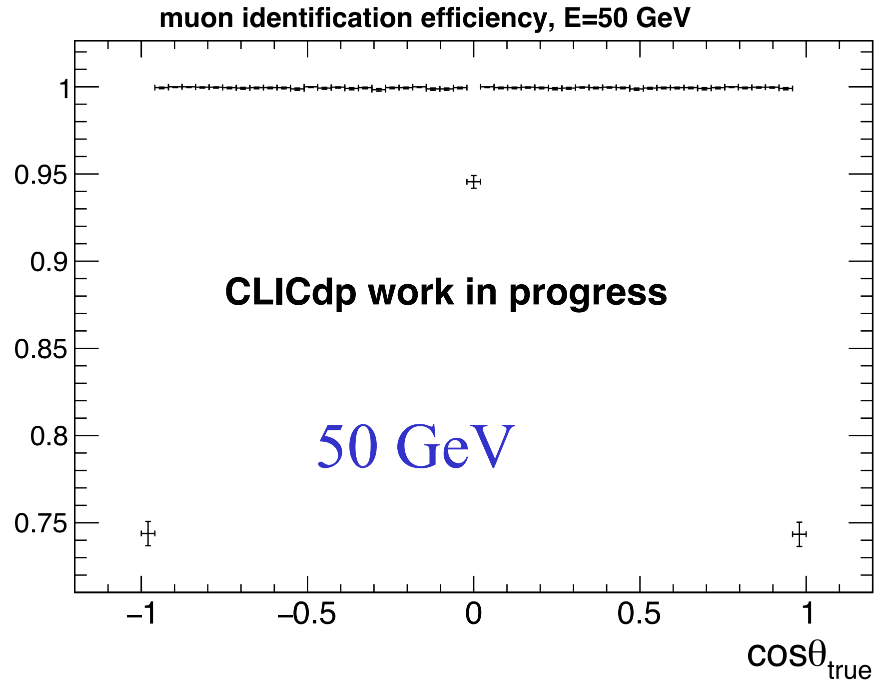



Reconstruct particles, starting at 8 degrees, reach more than 60 % at 10 degrees

# Muon Identification Efficiency

Studied on isolated single muon events
Around 98 % for almost the whole range

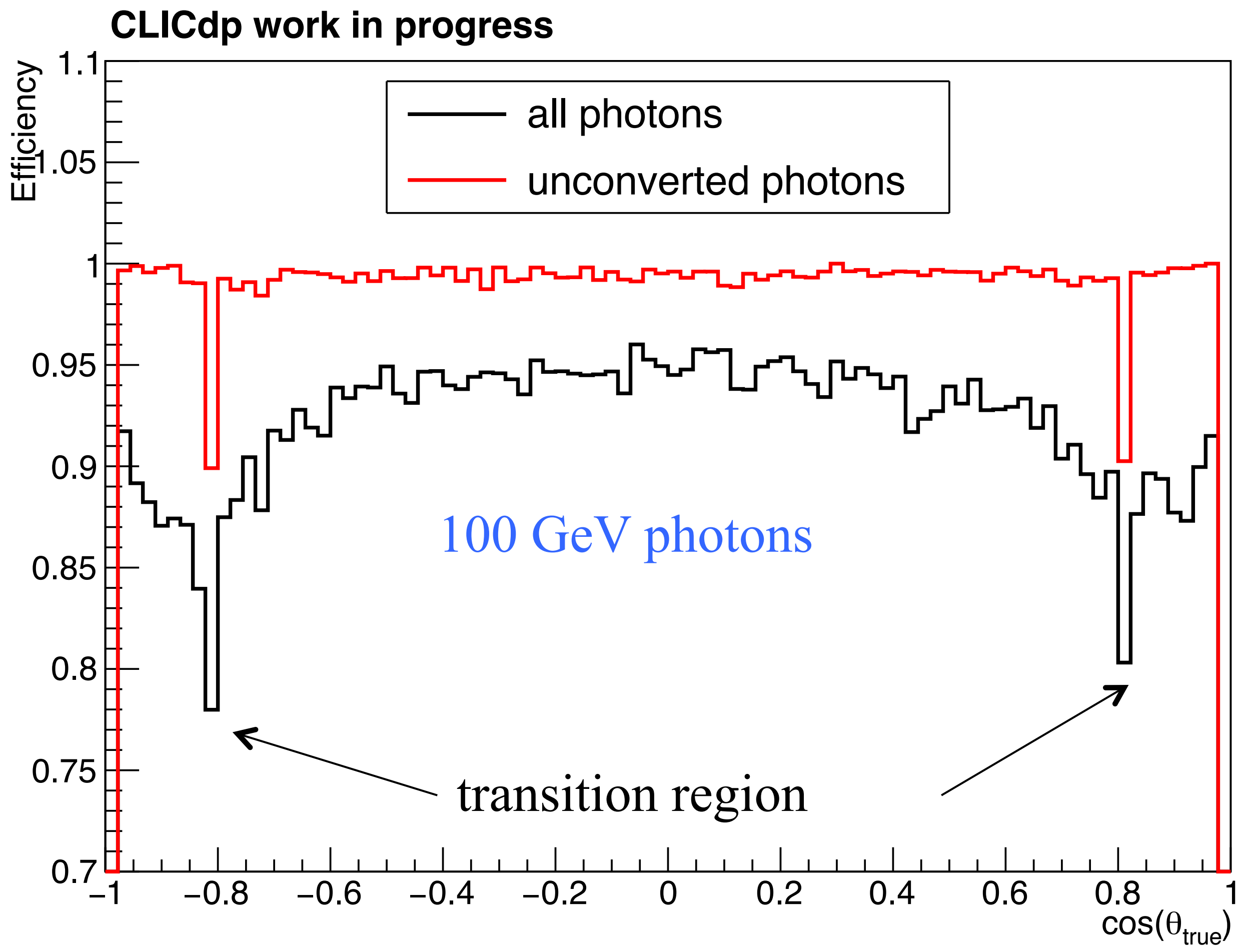



Investigation ongoing for point exactly at 0, where efficiency drops to 87 %, also for higher energies (checked up to 500 GeV), the inefficiency at 90 degrees remains, effect a lot larger than in the pion case
→ muon hits are there, track is also there, PFA expert comment: maybe something wrong in pseudo-layer calculation

# Photon Identification Efficiency

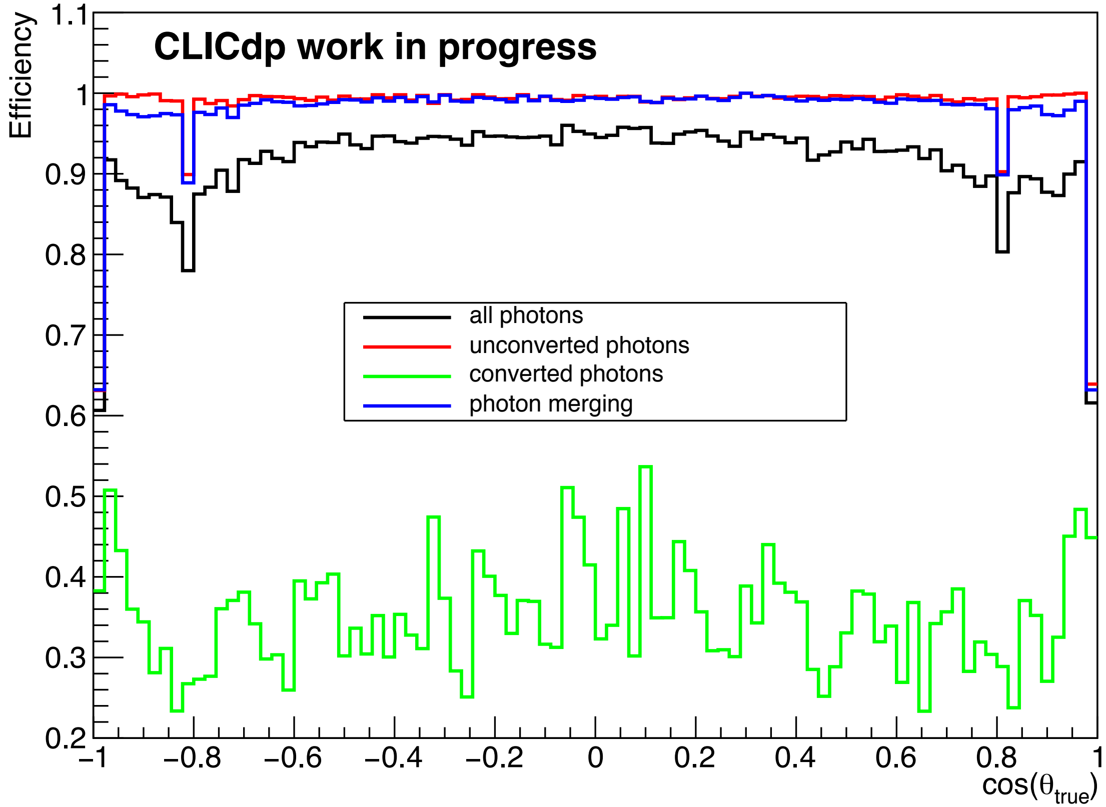
Studied on isolated single photon events for several energies:
Require angular matching with true photon as well as energy matching




➡ For unconverted photons achieve identification beyond 99 %

# Photon Identification Efficiency: Conversions

Photon conversions need to be treated separately, typically conversion rate below 15 %, happens late in the tracking system, thus default track cluster matching requirements in pandoraPFA not met by default →two very close resolved photons, which fails each the energy matching requirement




→ Recover largely performance for **converted photons** by **merging** both resolved reconstructed photons (spatial requirement less hard than for unconverted photons) → **over 90 % of merged pairs** survive energy cut

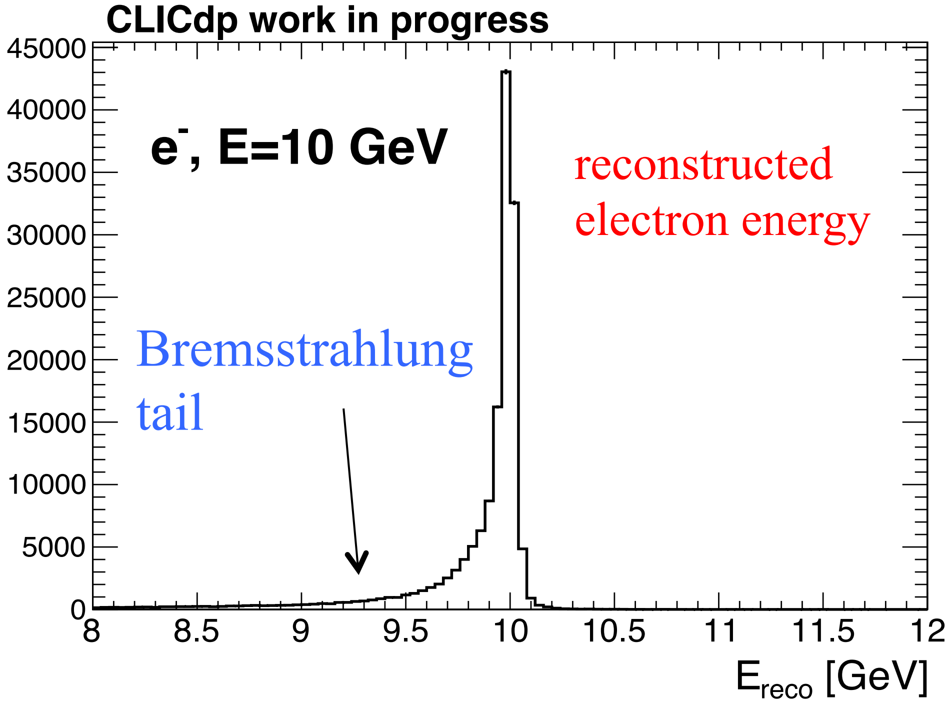
# Electrons and Bremsstrahlung

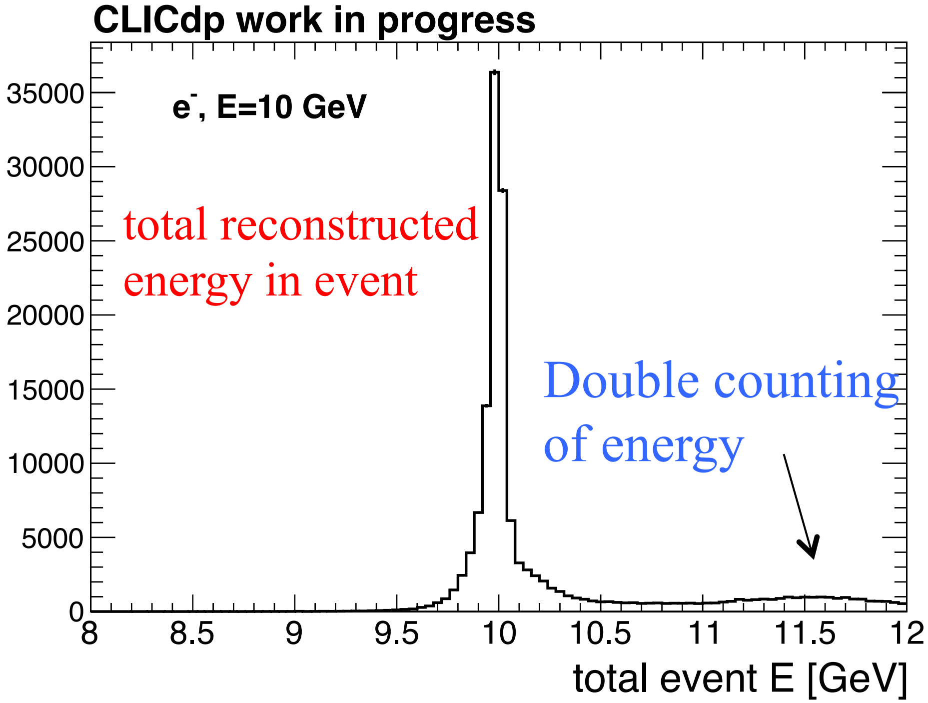
Electron energy resolution impacted by photon bremsstrahlung

→ Track fit also distorted

→ adding all photons up leads to a partial double counting as well

→ At 10 GeV Identification efficiency around 90 % in barrel, in endcap down to 75-80 %, for higher energies identification efficiency above 95 % both in barrel and endcap




⇨ Work ongoing for photon bremsstrahlung recovery algorithm to recover electron response tail and avoid double counting of energy from bremsstrahlung photons

# Software Compensation (SWC)

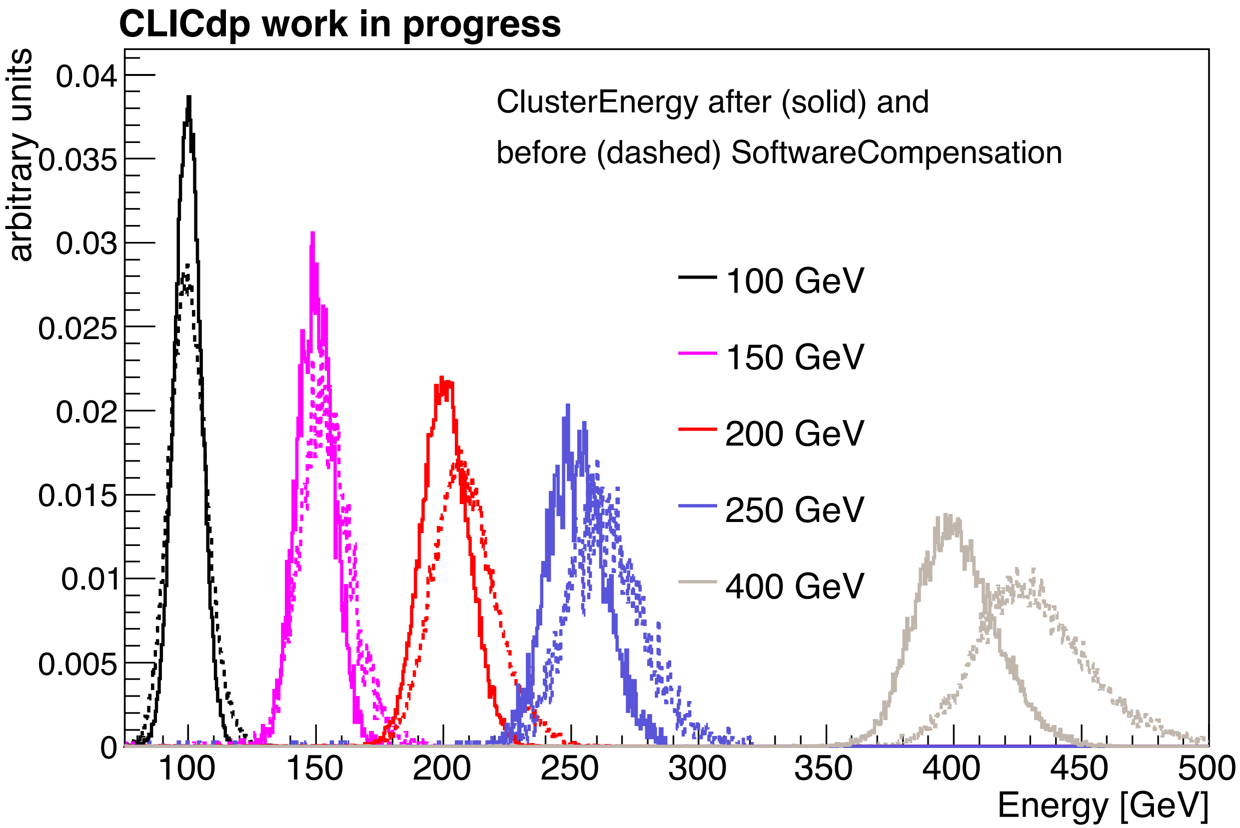
# Neutral Hadrons and Software Compensation

Response of electromagnetic component in hadron showers on average larger than for hadronic component →treating all hits in HCAL the same way leads to a less accurate energy measurement

Follow description of Software Compensation paper **EPJC 77 (2017) 698**
Electromagnetic component of shower **typically** **denser**:
Software compensation at CLIC reweights hits in HCAL depending on the hit energy density, assuming these originate from electromagnetic component
→ assume monotonic falling weight with energy density
→ Weight includes an energy dependence
→ in total 9 different parameters are used in software compensation
→ Software compensation reweighting integrated into PandoraPFA software
CALICE investigates alternative parametrisations of software compensation weights, including hits in ECAL → See talk by *Yasmine*

Default weights tuned for ILD up to 100 GeV → SWC used for $E_{hadrons} < 100$ GeV
at CLIC expect to reach higher hadron energies, at 3 TeV sometimes beyond 500 GeV
→retune parameters for CLIC, use SWC for all hadron energies

# Software Compensation at work at CLIC




Software compensation weights derived from MC using neutron and $K^0_L$ events

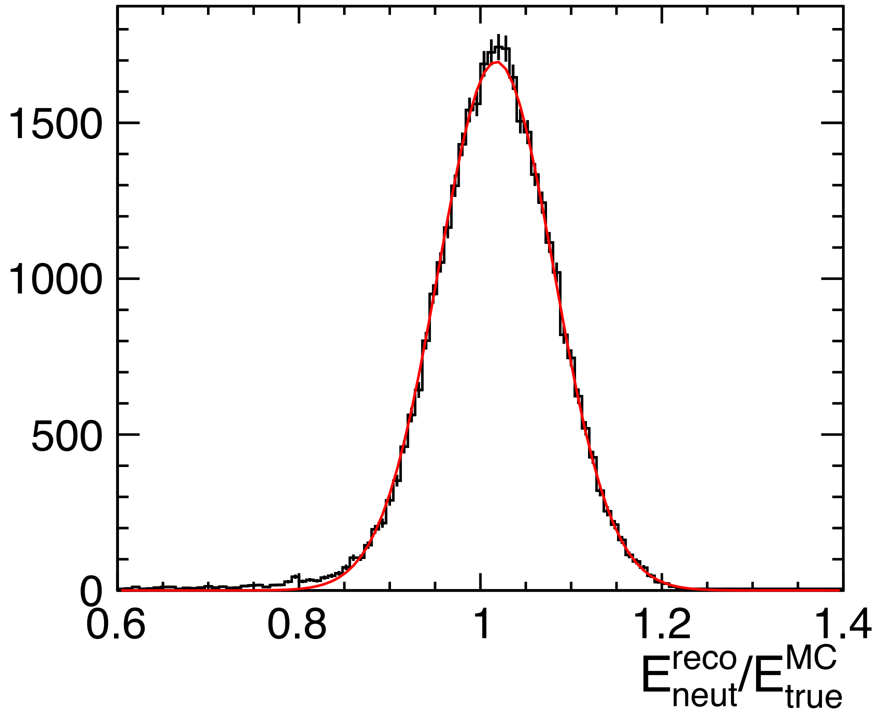
→Mean and resolution after software compensation largely improved

→Software compensation corrects for nonlinear response of hadrons on the fly

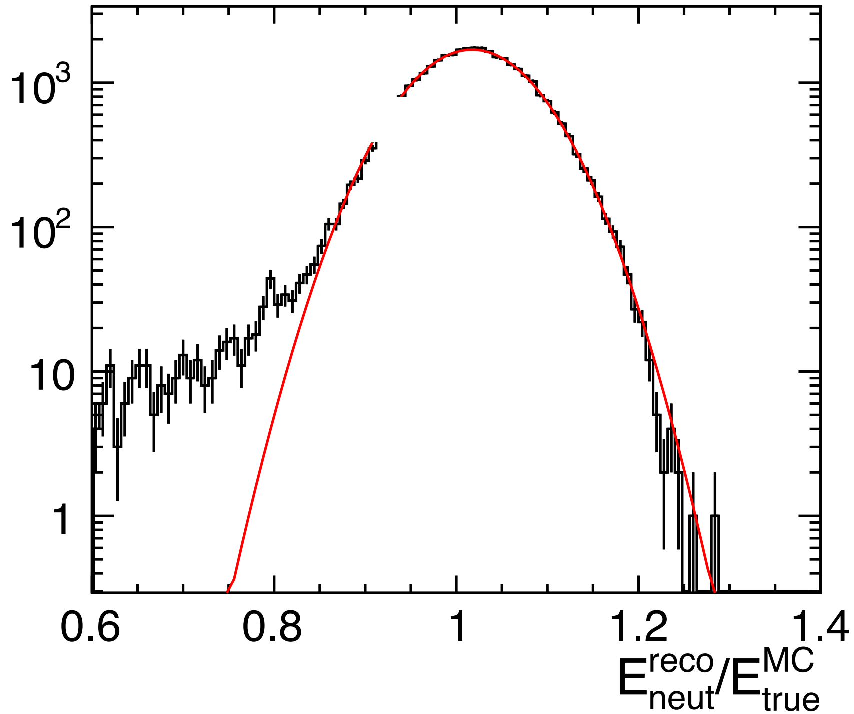
# K0L Energy Resolution

Fit neutral hadron response with a gaussian, several single particle gun energy points→apply CLIC software compensation weights
three regions: barrel, endcap, transition, non gaussian tails to lower responses

Barrel region, K0L, E=75 GeV



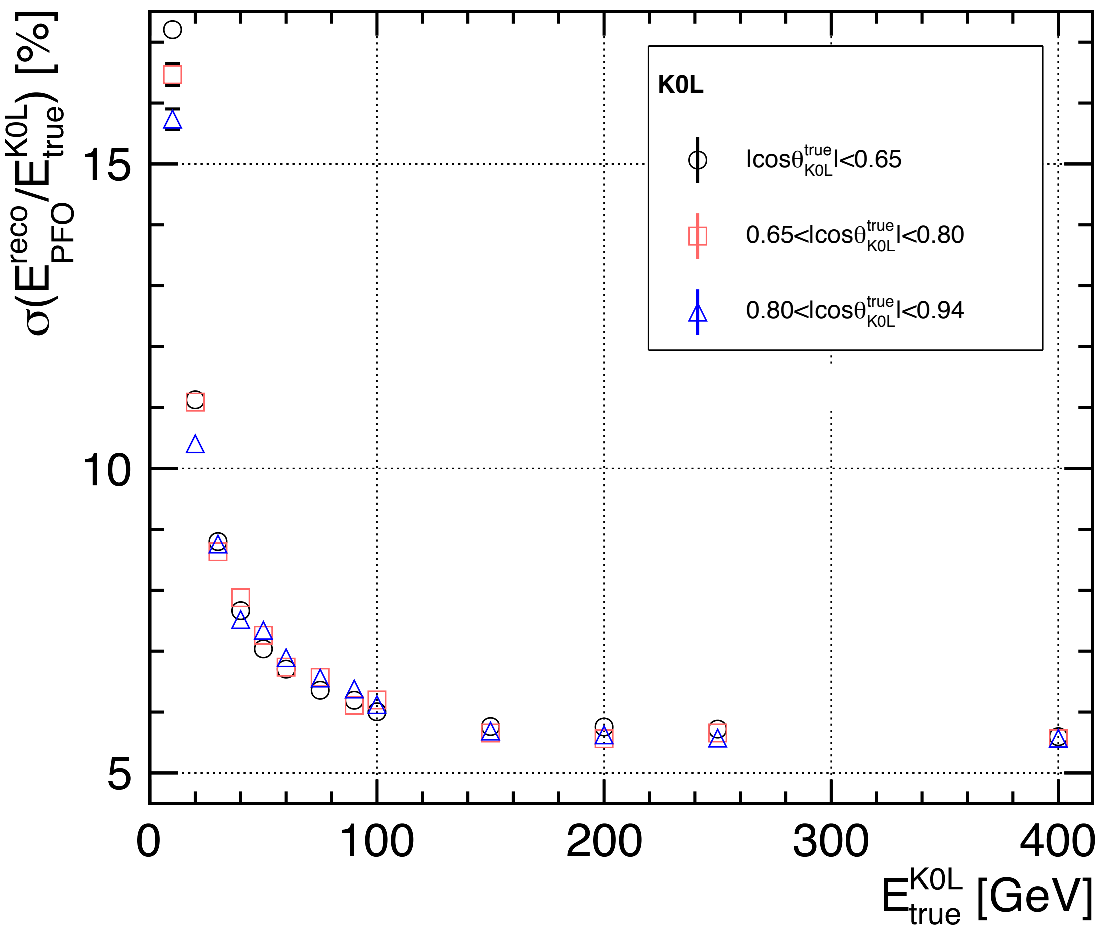

Barrel region, K0L, E=75 GeV




Fit relative response with a gaussian →for 75 GeV 6.4 %, mean of fitted response 1.014

# Neutral Hadrons Energy Resolution Summary

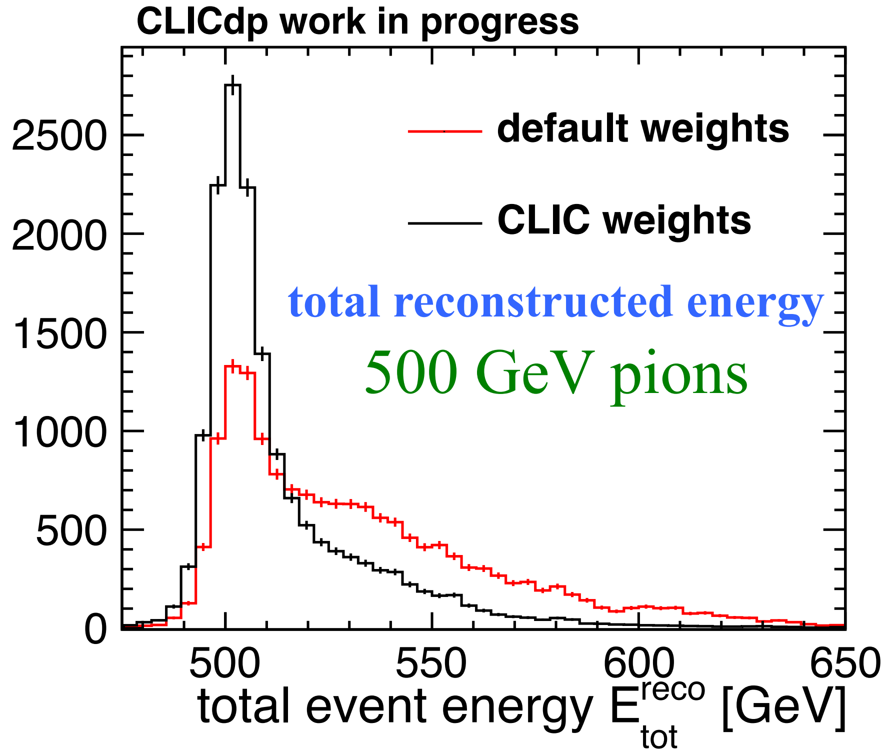
Neutral Hadron energy resolution determined in three regions, barrel, transition region and endcap

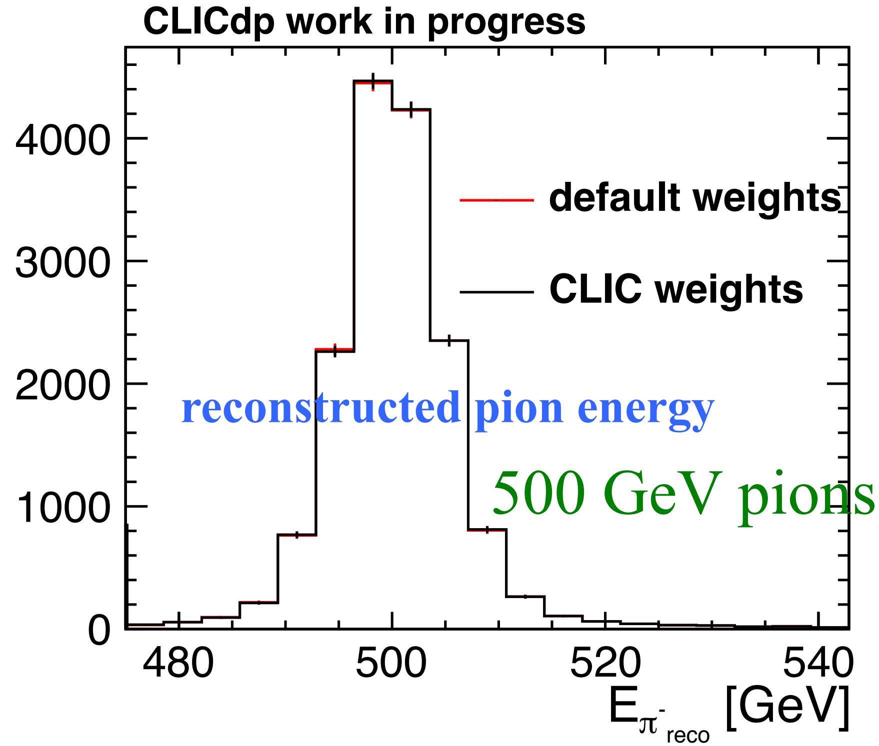



Fit relative response with a gaussian →around 16.2 % at 10 GeV, around 150 GeV we reach a values around 5.5 %

# Software compensation and charged particle

Software compensation applied to all HCAL hits, also impact on charged pions



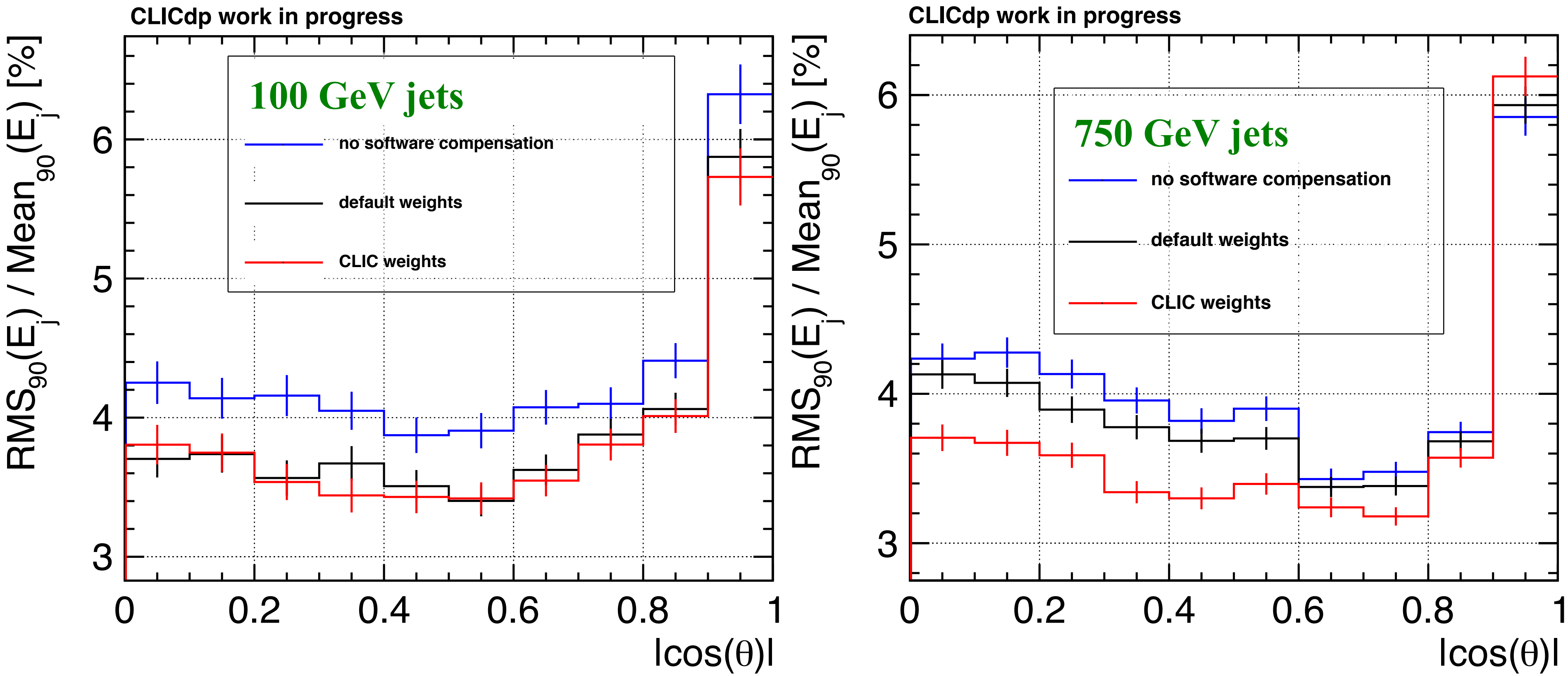




With new CLIC weights significant reduction in reconstructed original cluster energy →less "excessive" energy to split and create additional neutrons (when compared to track energy)

→ Software compensation reduced confusion due to reduced splitting of clusters originating from very high energetic tracks (energies that high relevant for decays of high energetic hadronic taus, pion energies in jets rarely beyond 100 GeV)

# Software compensation at CLIC in dijet events

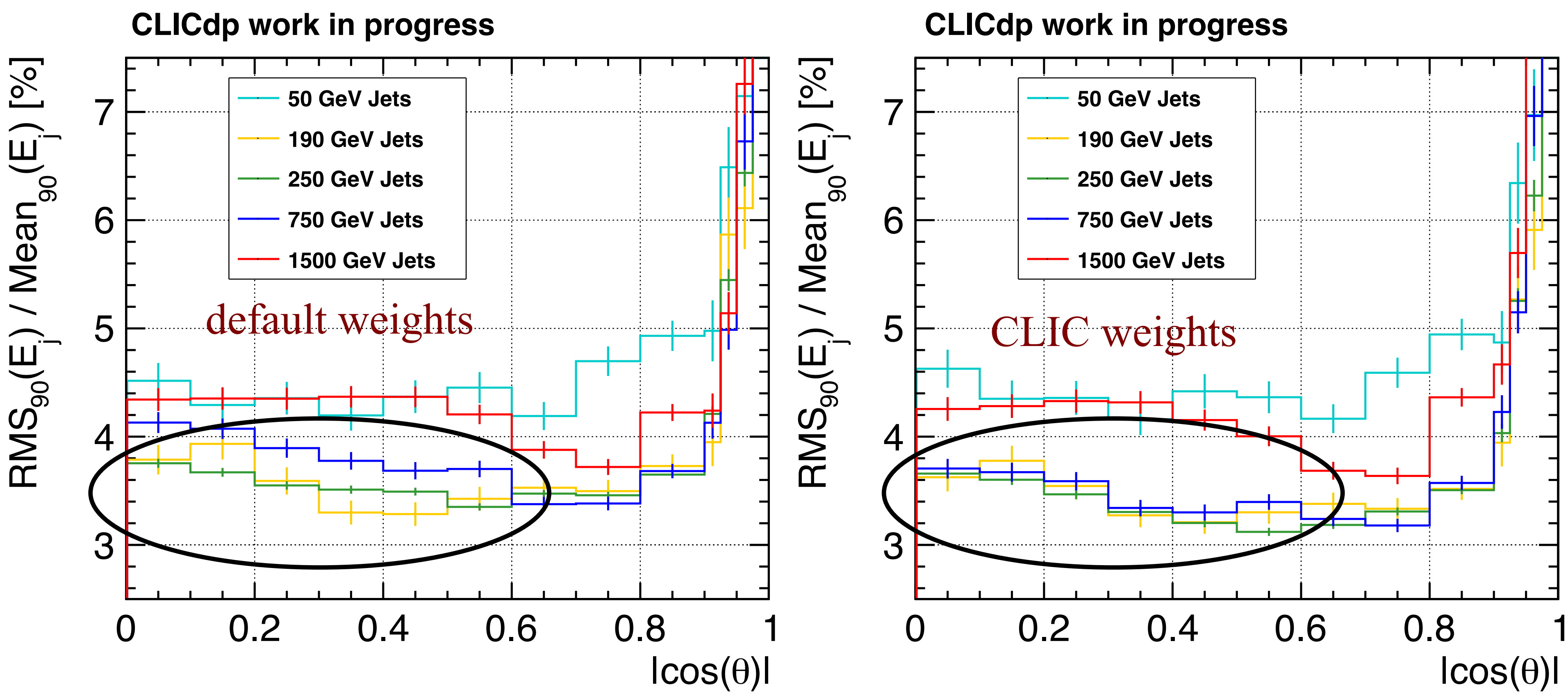
CLIC specific software compensation weights compared to previous default




Significant improved jet energy with new CLIC specific weights for large energetic jets, comparable performance at low jet energies → applying software compensation always improves jet energy resolution

# Jet energy resolution at CLIC

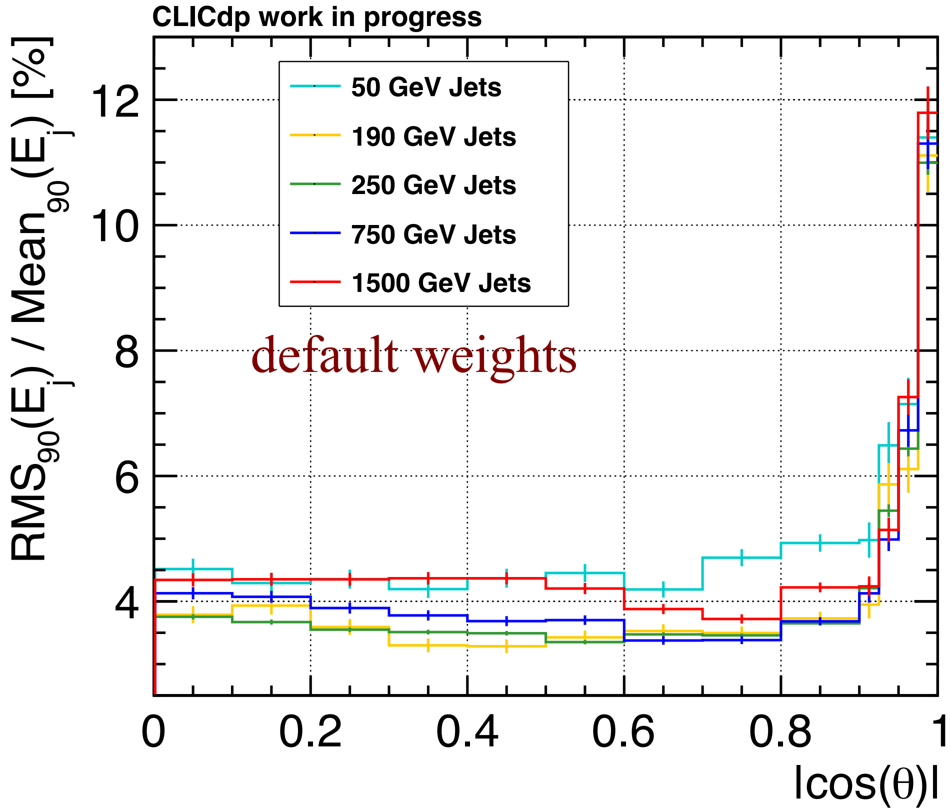
CLIC specific software compensation weights → achieve jet energy resolution between 3.1 % and 4.5 %

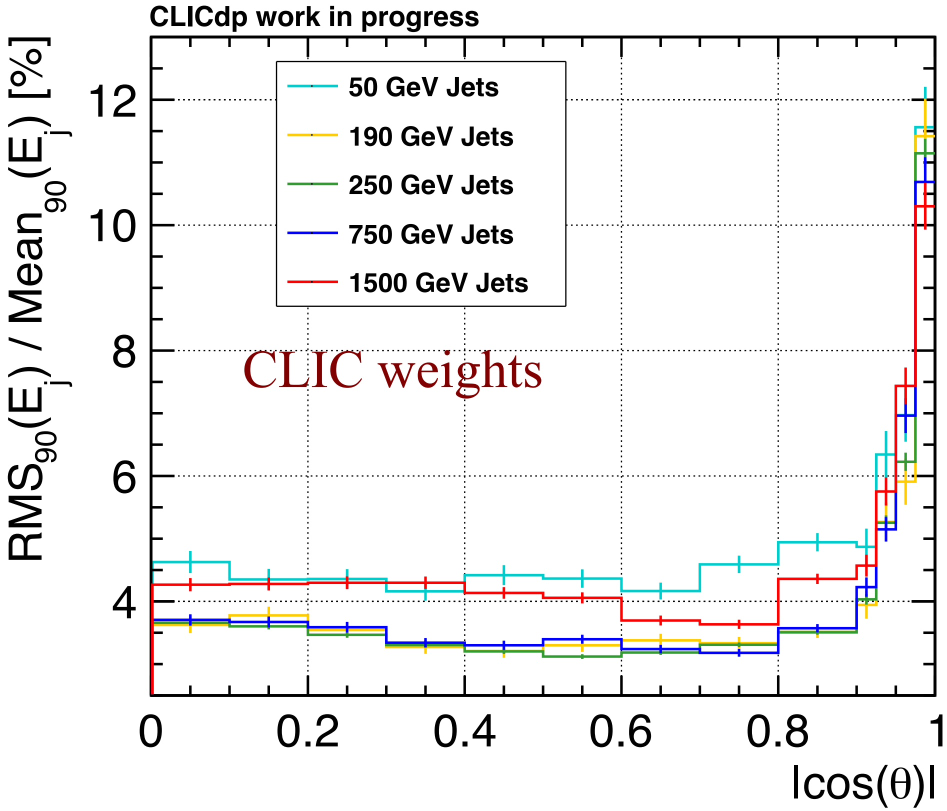



Compared to previous default software compensation weights comparable performance for jets up to 190 GeV, for larger jet energies improvement of resolutions by around 10%

# Jet energy resolution at CLIC: zoomed out

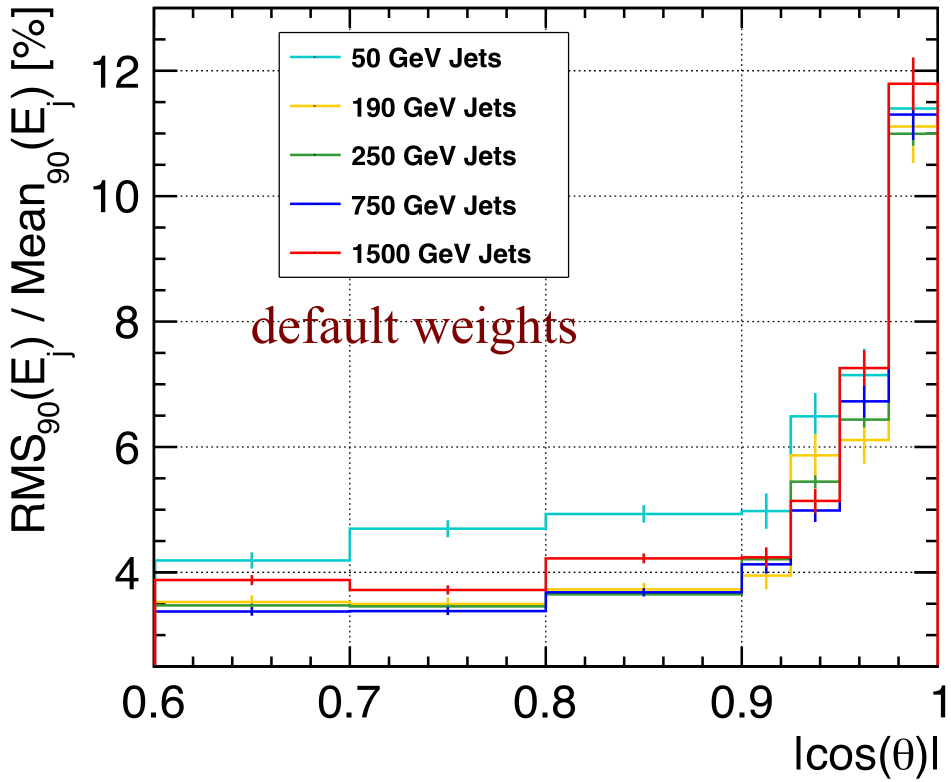
in forward region jet energy values up to 11.5 %, uncertainties of RMS values larger due to less statistics

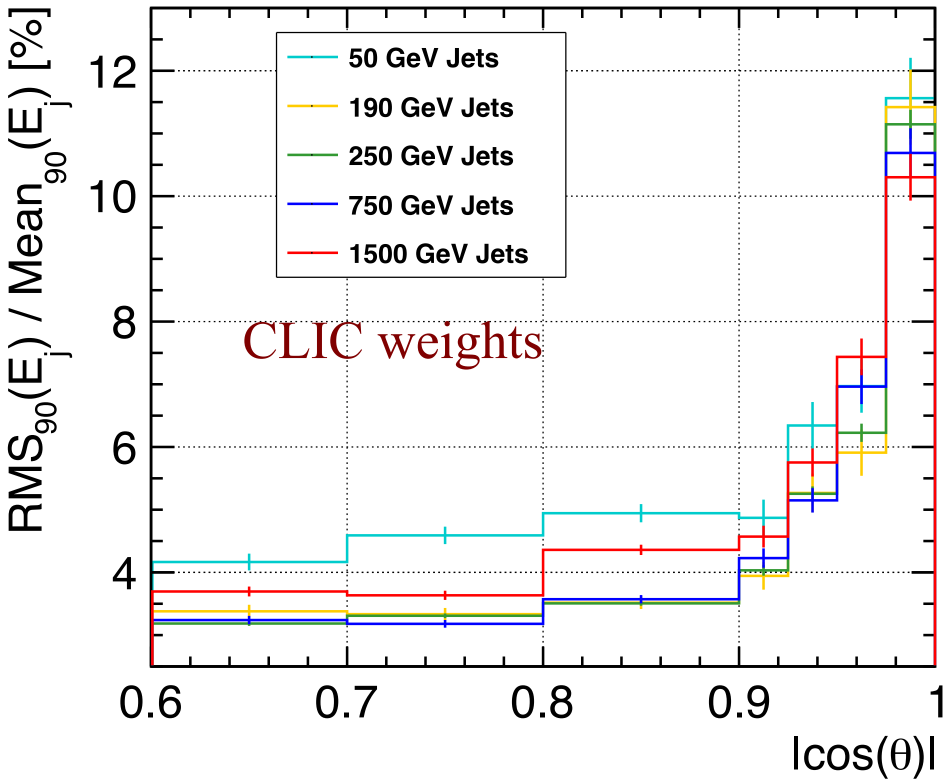


# Jet energy resolution at CLIC: forward jets

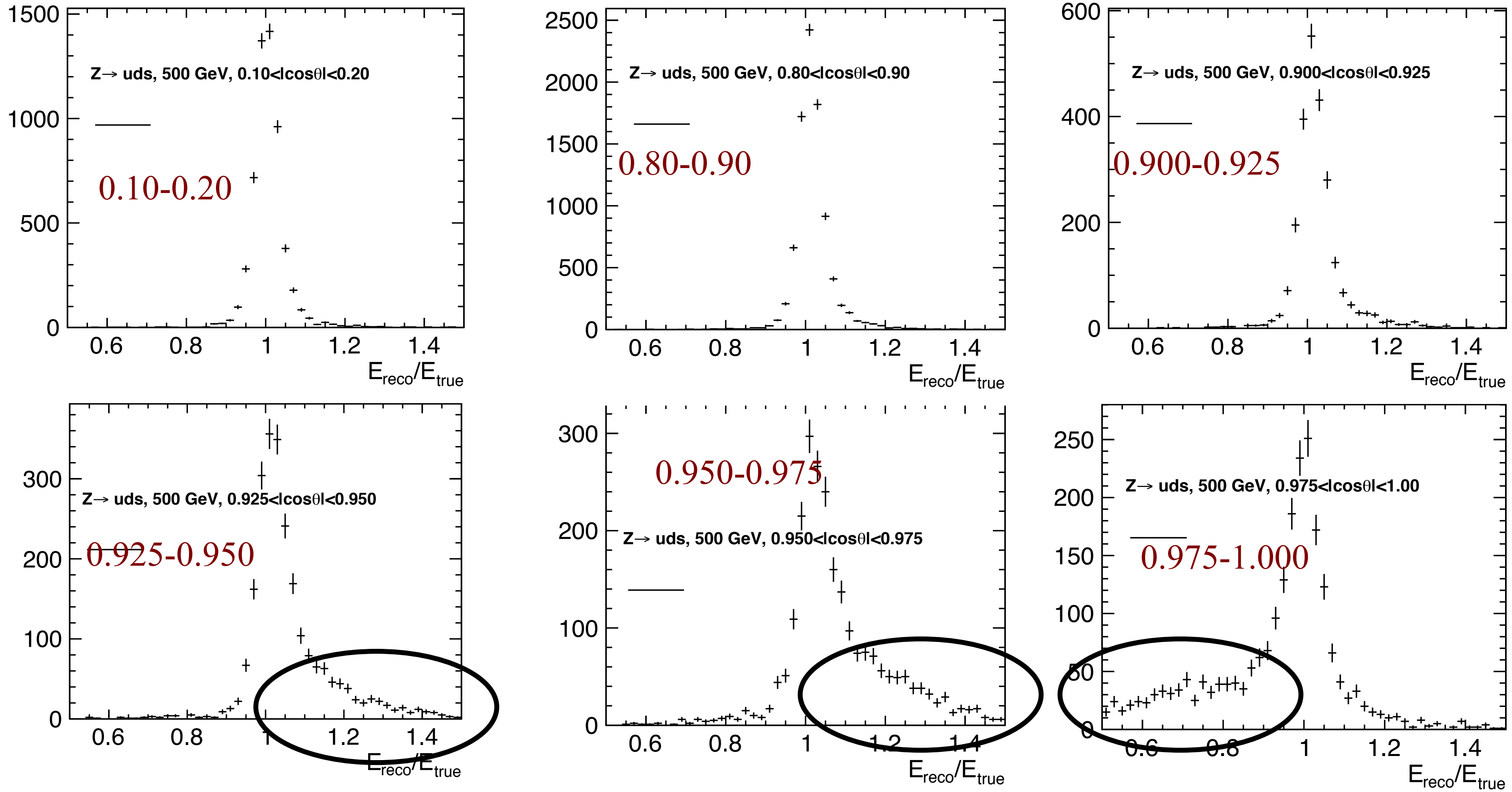
Endcap resolutions a bit worse than resolutions in outer barrel bins, large increase starting from bin at |cos θ| >0.925 (18.2-22 °), a lot worse for jets with |cos θ| >0.975 (12 °)



# Energy Resolution: Z→uds at 500 GeV

Check energy resolution for 500 GeV dataset in different cos θ bins

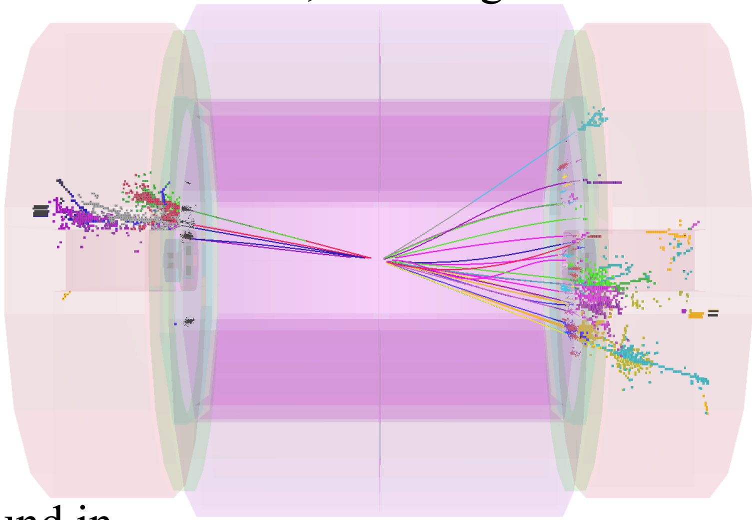



Resolution core not drastically different in forward region (compared to endcap), but long tails

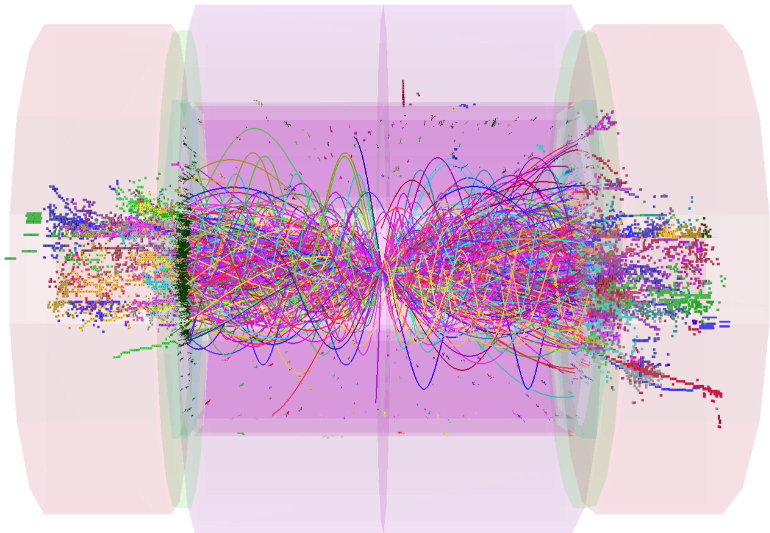
# Effects of beam background

# Beamstrahlung γγ→hadrons background reduction

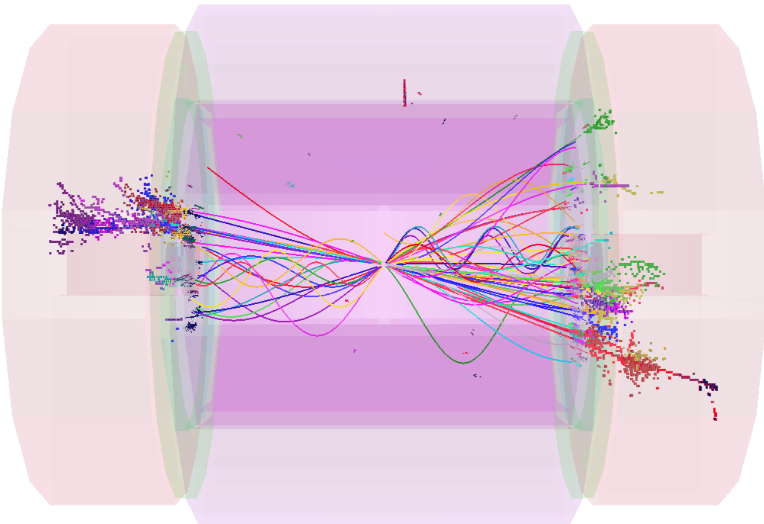
Forward WW, no background



with γγ→hadrons background in reconstruction time window ~1.2 TeV



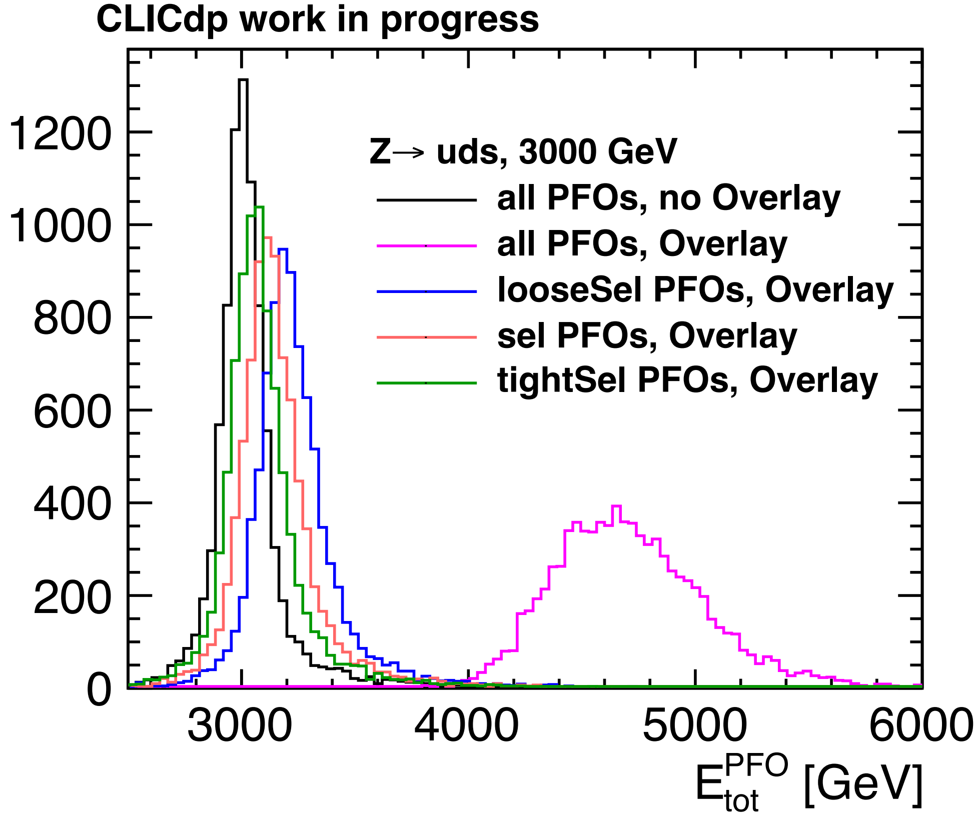
after timing and $p_T$ cuts: ~ 100 GeV remain

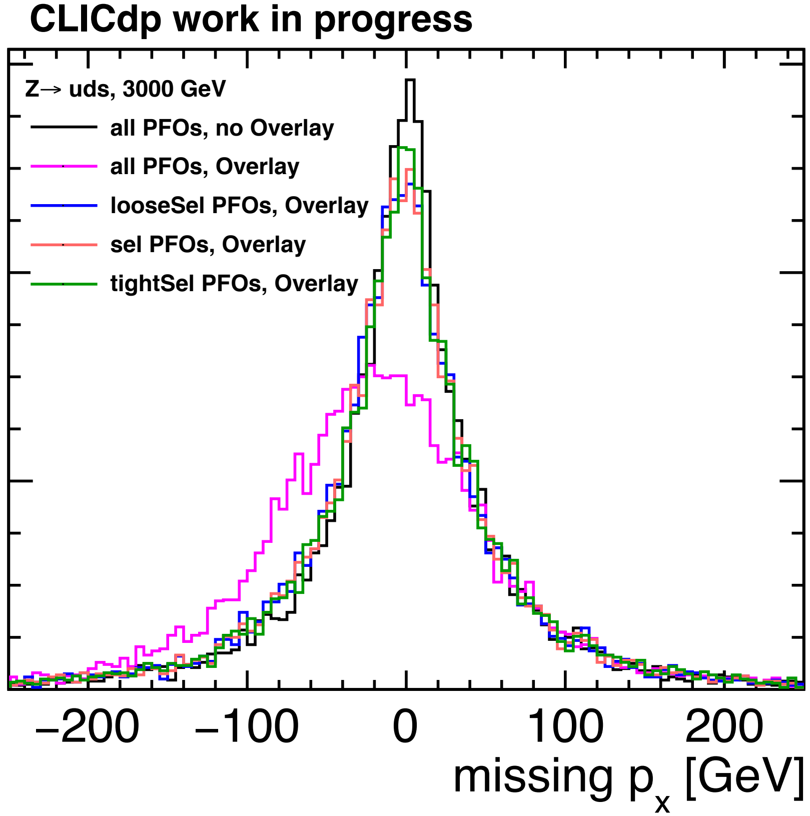


# Background in Z→ light dijets at 3 TeV

At 3 TeV expect 3.2 hadronic event per bunch crossing

Check events of Z→uds quarks, dijet events without genuine missing energy



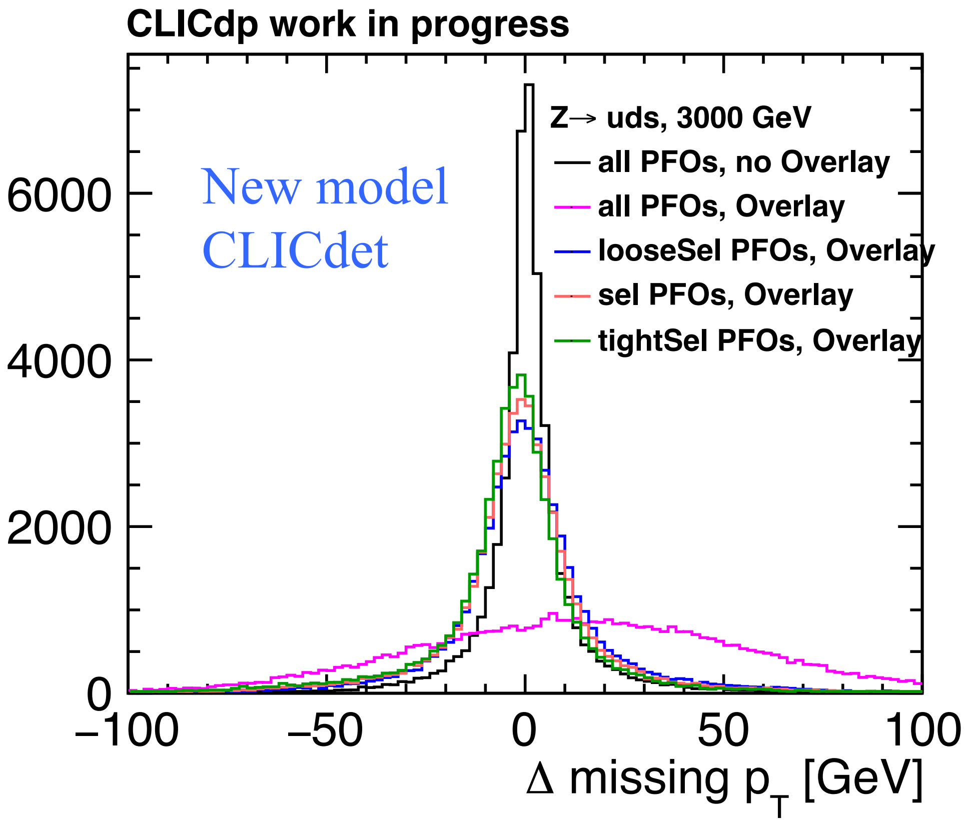



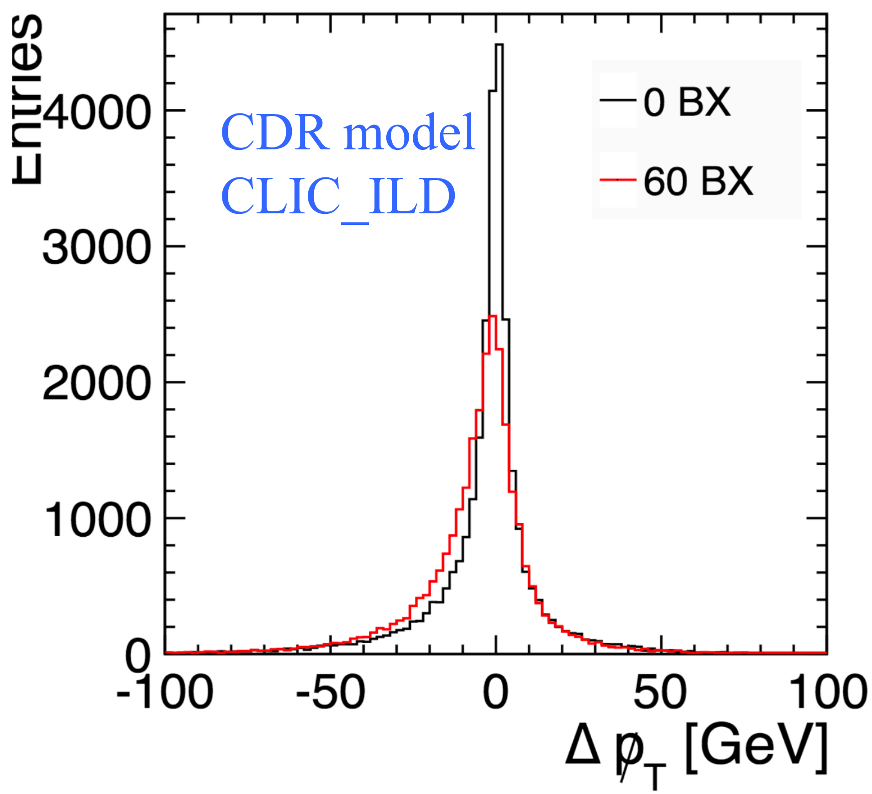

Out of originally almost 2 TeV additional energy after timing and $p_T$ cuts: ~ 100 GeV remain (tight selection → targets to deal with 3 TeV conditions)

No genuine missing momentum component in these events → all selection cuts recover performance before background overlay

# Background in ZZ→qqνν at 3 TeV

At 3 TeV expect 3.2 hadronic event per bunch crossing
Check change in missing transverse momentum



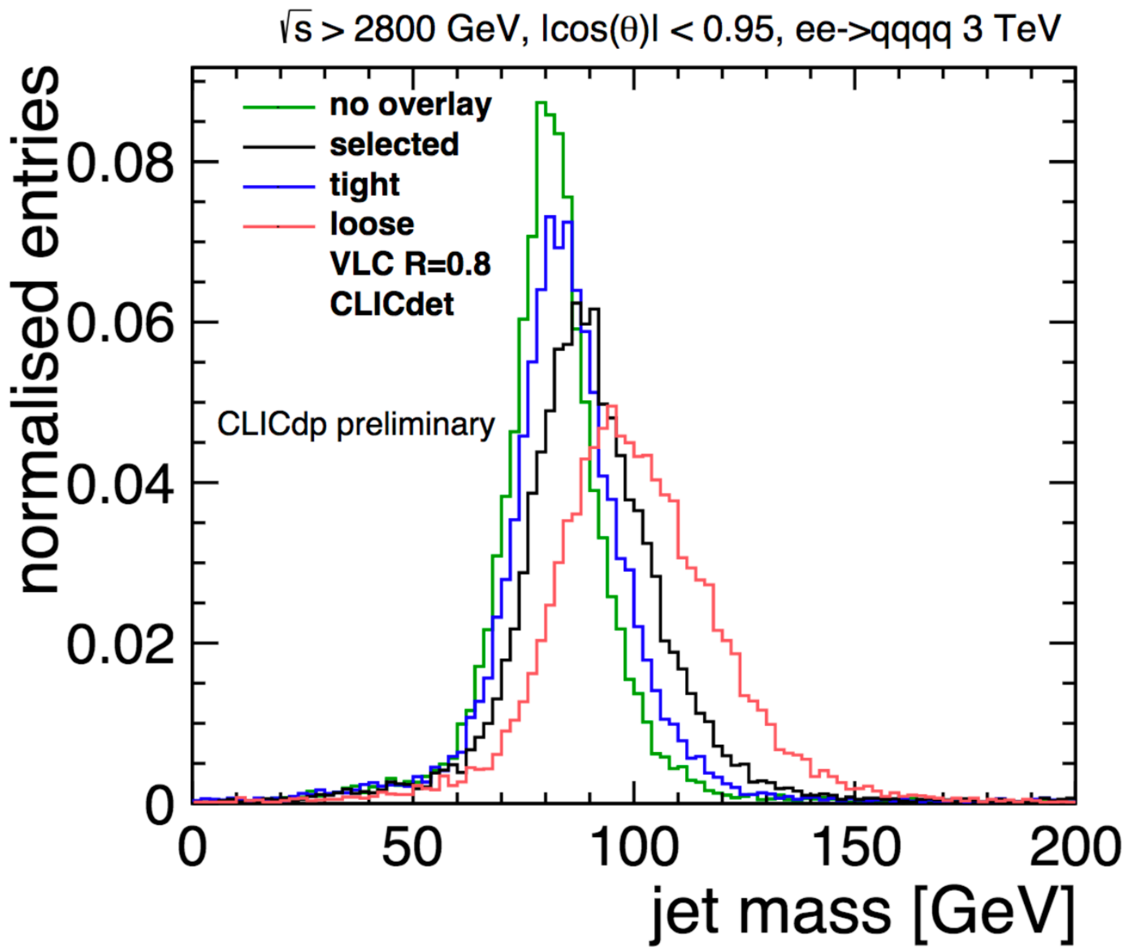




Comparable performance of CLICdet with CDR CLIC_ILD model in presense of background

Here tightSelected PFO cuts applied on red curve

# Background and mass reconstruction

Study timing cuts in 3 TeV WW→qqqq events, boosted topology → cluster event in two jets




Using tight selected PFOs as input for jet clustering almost recovers mass peak at 80 GeV completely

# Summary

Performance of PandoraPFA has been studied with new detector model CLICdet for several particle species, jets and missing transverse momentum distributions

- CLIC specific software compensation parameters have been determined
  - Comparable jet energy resolution for low energetic jets (<250 GeV)
  - Jet resolutions improved by around 10 % for high energetic jets (>250 GeV) when using previous default parameters, in barrel values of 3-4.5 % are achieved, for very forward jets resolution around 7-11 %

  → Confusion term reduced by more accurate cluster energy estimation

- Study impact of $p_T$ and timing cuts on missing energy resolution assuming background levels of a 3 TeV machine
  - Use timing cuts determined for previous detector model CLIC_ILD for new CLICdet model
  - Good performance of tinming cuts show by examples of jet masses in hadronic WW events and missing $p_T$ resolutions in dijet and ZZ→qqνν events

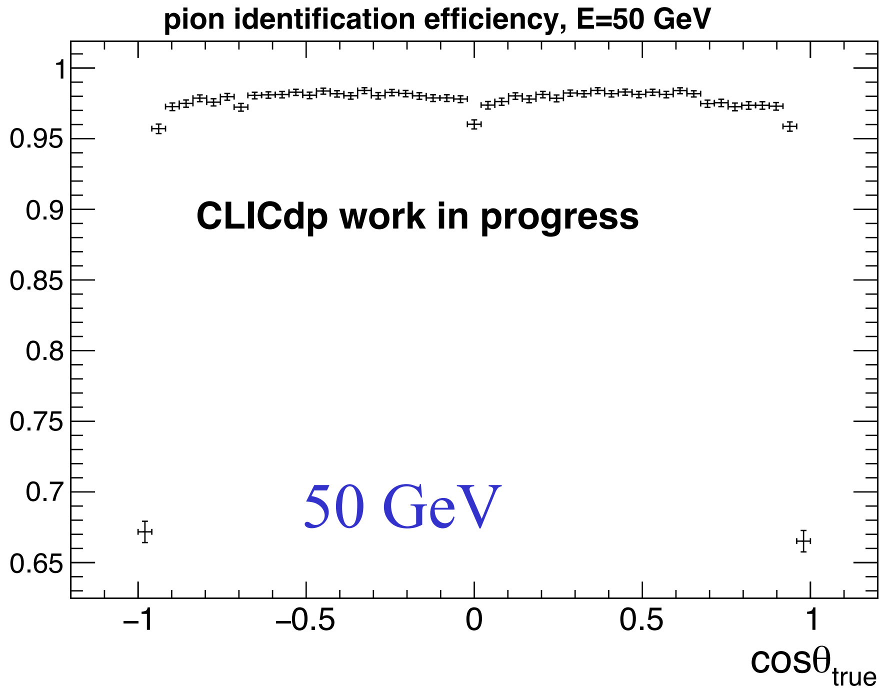
# BACKUP

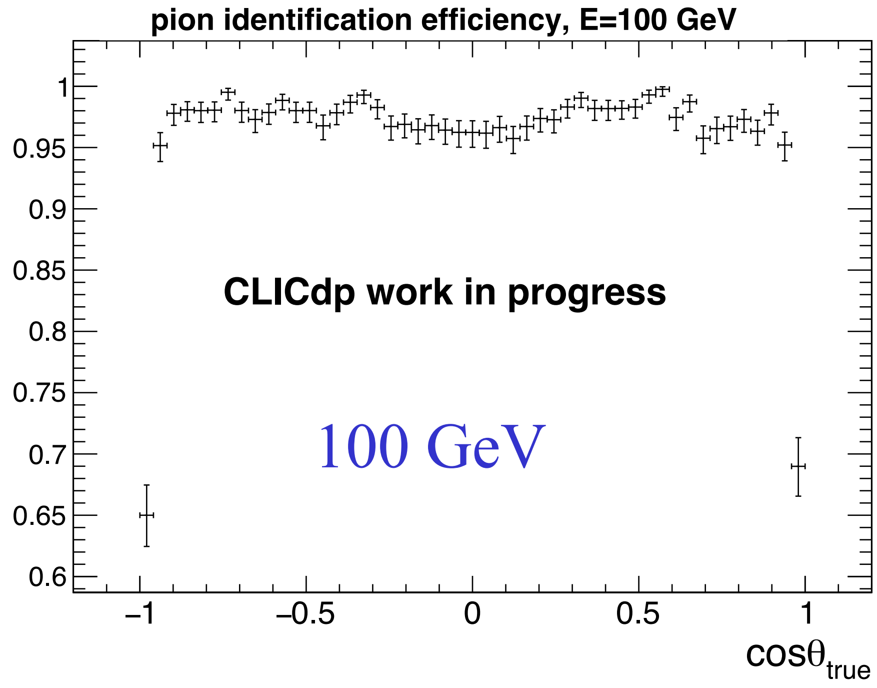
# Definition of PandoraPFA timing cuts

| | Loose configuration | | Tight configuration | |
|---|---|---|---|---|
| **Region** | **$p_T$ range [GeV]** | **Time [ns]** | **$p_T$ range [GeV]** | **Time [ns]** |
| | | **Photons** | | |
| Central | $0.75 \leq p_T < 4.0$ | $t < 2.0$ | $1.0 \leq p_T < 4.0$ | $t < 2.0$ |
| $\lvert\cos(\theta)\rvert \leq 0.975$ | $0 \leq p_T < 0.75$ | $t < 2.0$ | $0.2 \leq p_T < 1.0$ | $t < 1.0$ |
| Forward | $0.75 \leq p_T < 4.0$ | $t < 2.0$ | $1.0 \leq p_T < 4.0$ | $t < 2.0$ |
| $\lvert\cos(\theta)\rvert > 0.975$ | $0 \leq p_T < 0.75$ | $t < 1.0$ | $0.2 \leq p_T < 1.0$ | $t < 1.0$ |
| | | **Neutral hadrons** | | |
| Central | $0.75 \leq p_T < 8.0$ | $t < 2.5$ | $1.0 \leq p_T < 8.0$ | $t < 2.5$ |
| $\lvert\cos(\theta)\rvert \leq 0.975$ | $0 \leq p_T < 0.75$ | $t < 1.5$ | $0.5 \leq p_T < 1.0$ | $t < 1.5$ |
| Forward | $0.75 \leq p_T < 8.0$ | $t < 2.5$ | $1.0 \leq p_T < 8.0$ | $t < 1.5$ |
| $\lvert\cos(\theta)\rvert > 0.975$ | $0 \leq p_T < 0.75$ | $t < 1.5$ | $0.5 \leq p_T < 1.0$ | $t < 1.0$ |
| | | **Charged particles** | | |
| All | $0.75 \leq p_T < 4.0$ | $t < 3.0$ | $1.0 \leq p_T < 4.0$ | $t < 2.0$ |
| | $0 \leq p_T < 0.75$ | $t < 1.5$ | $0 \leq p_T < 1.0$ | $t < 1.0$ |

# Pion Identification Efficiency

Studied on isolated single pion events, around 95 % for almost the whole range, for very high energetic pions in barrel slightly higher inefficiency than for endcap

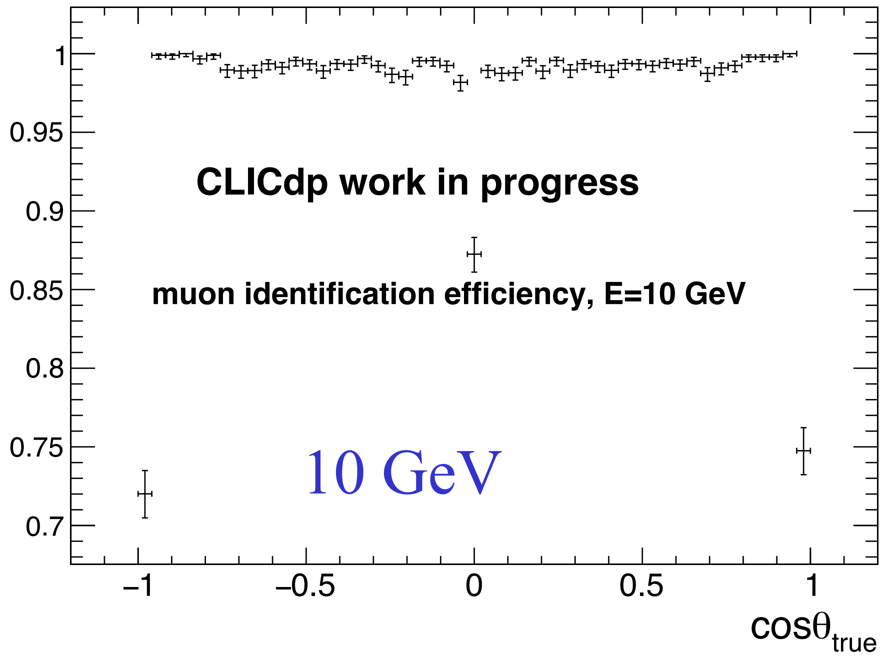


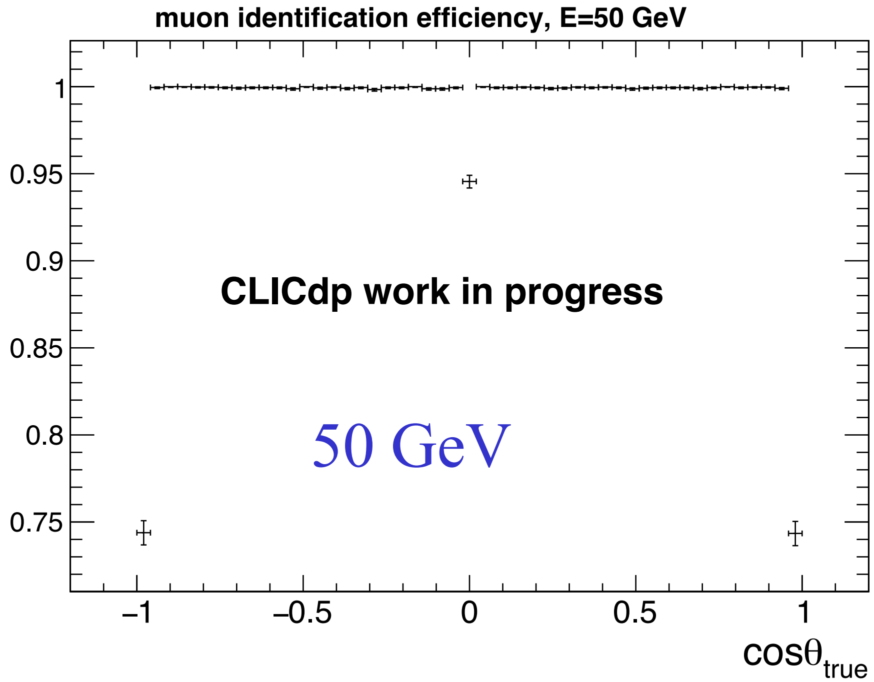




Point at exactly 90 degrees a bit lower than neighbouring region

# Muon Identification Efficiency

Studied on isolated single muon events
Around 98 % for almost the whole range

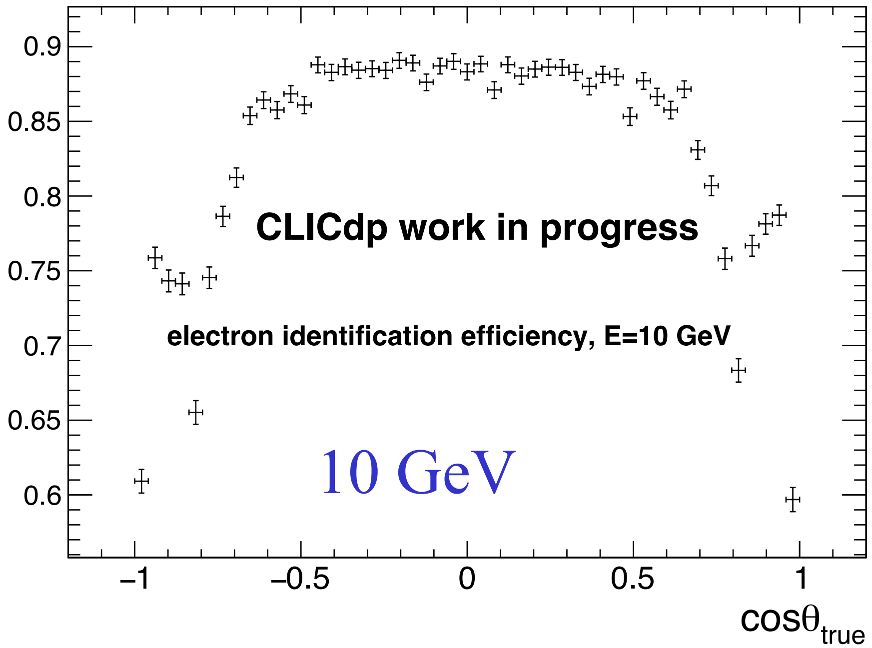


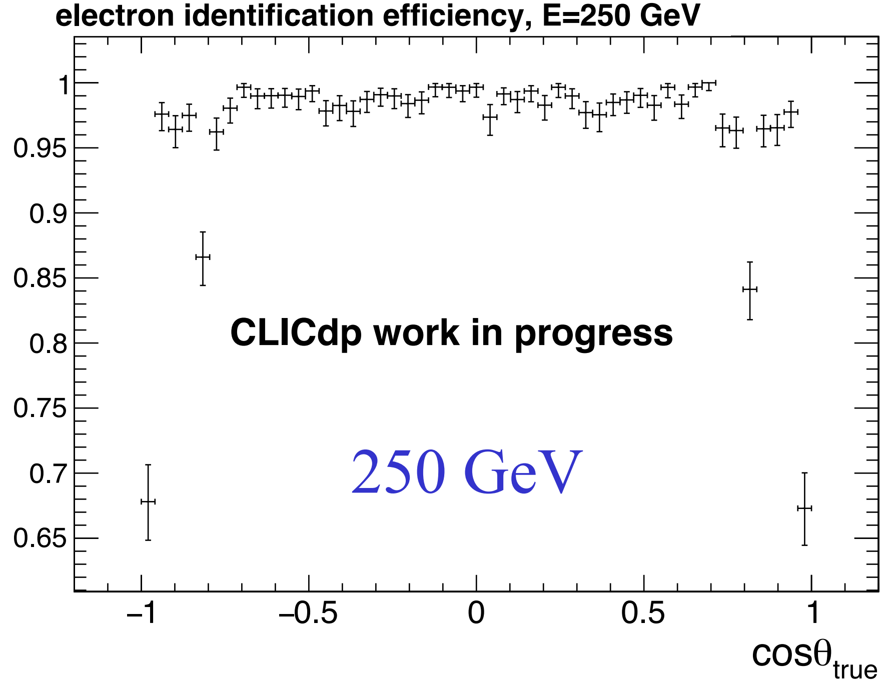




→ Investigation ongoing for point exactly at 0, where efficiency drops to 87 %, also for higher energies (checked up to 500 GeV), the inefficiency at 90 degrees remains, effect a lot larger than in the pion case

# Electron Identification Efficiency

PandoraPFA not designed to perform perfect isolated electron ID, but to find electrons within jets → ID might need further development for low energies



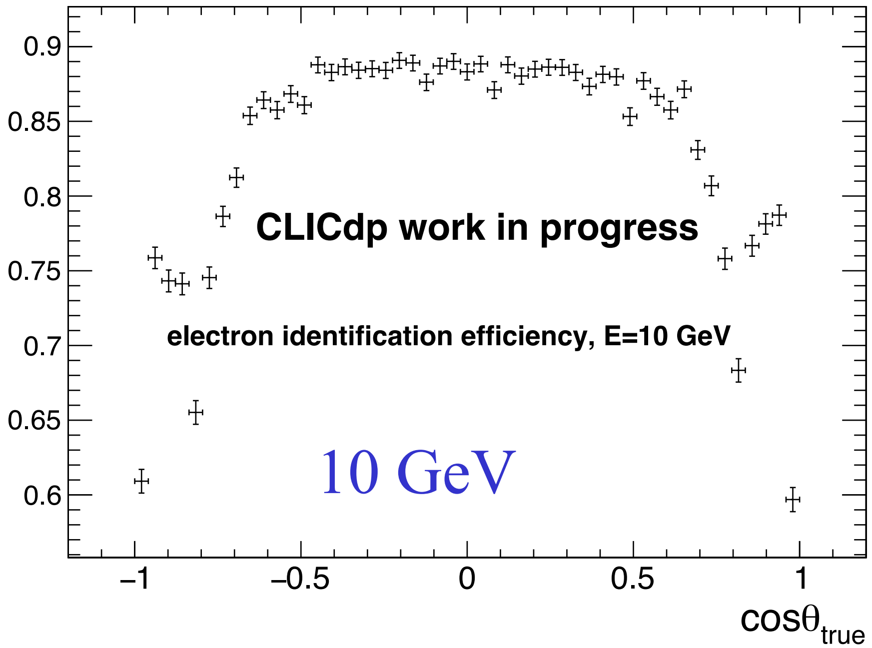



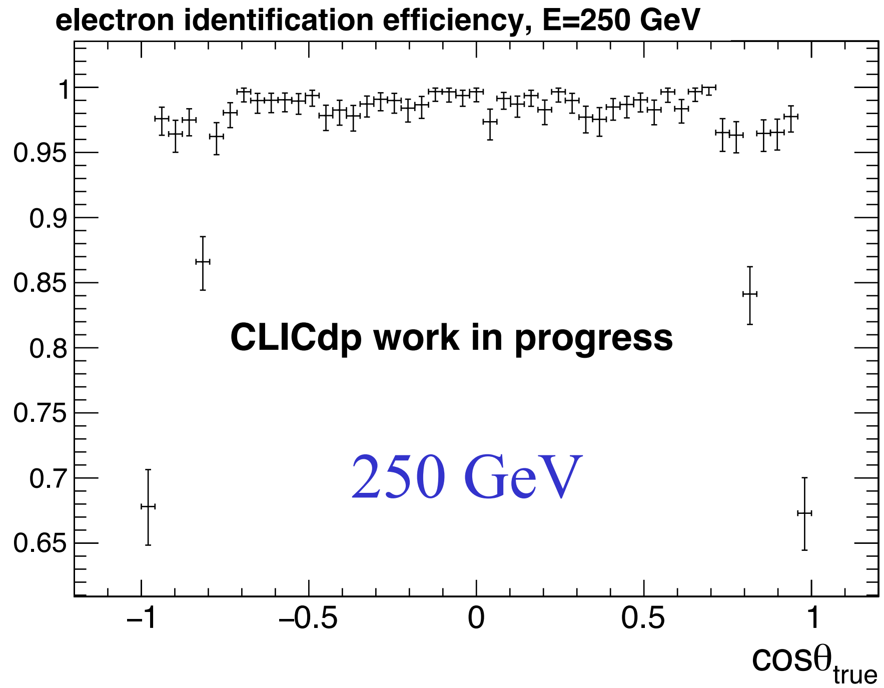

Close to 90 % in inner barrel, drop to 75-80 % in endcap

Above 95 % almost everywhere

# Electron Identification Efficiency

PandoraPFA not designed to perform perfect isolated electron ID, but to find electrons within jets → ID might need further development for low energies



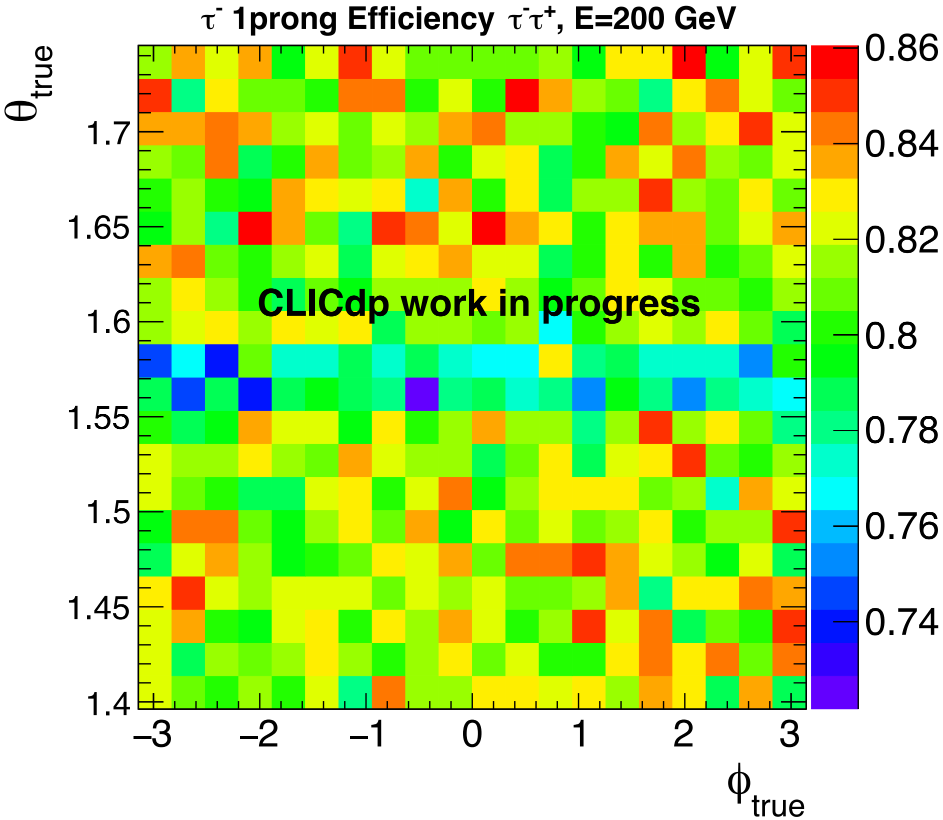



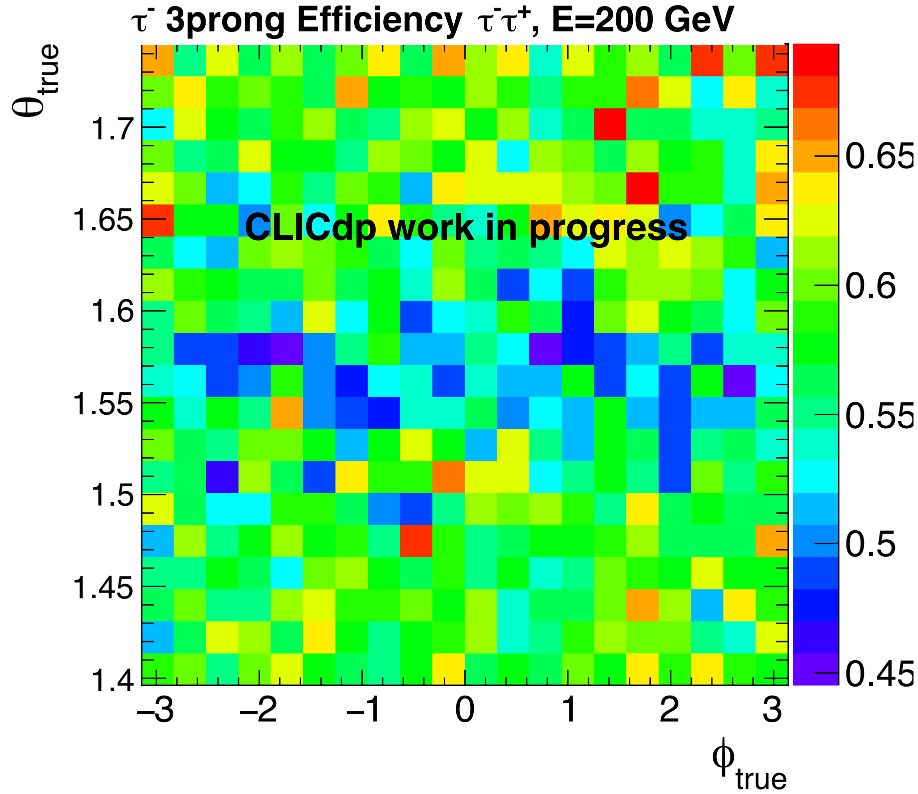

Close to 90 % in inner barrel, drop to 75-80 % in endcap

Above 95 % almost everywhere

# Hadronic Tau Identification Efficiency

Checked on tau tau events in central part of detector → true tau jet spatially matched to reconstructed tau jet, check for correct charge identification and correct tau decay mode



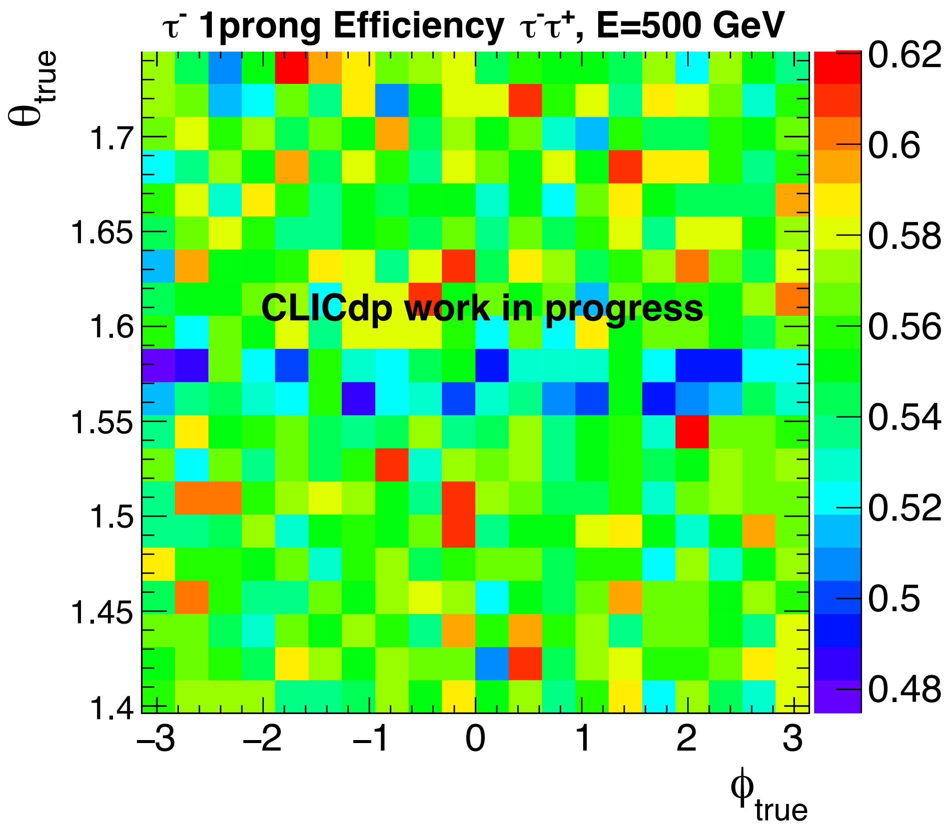



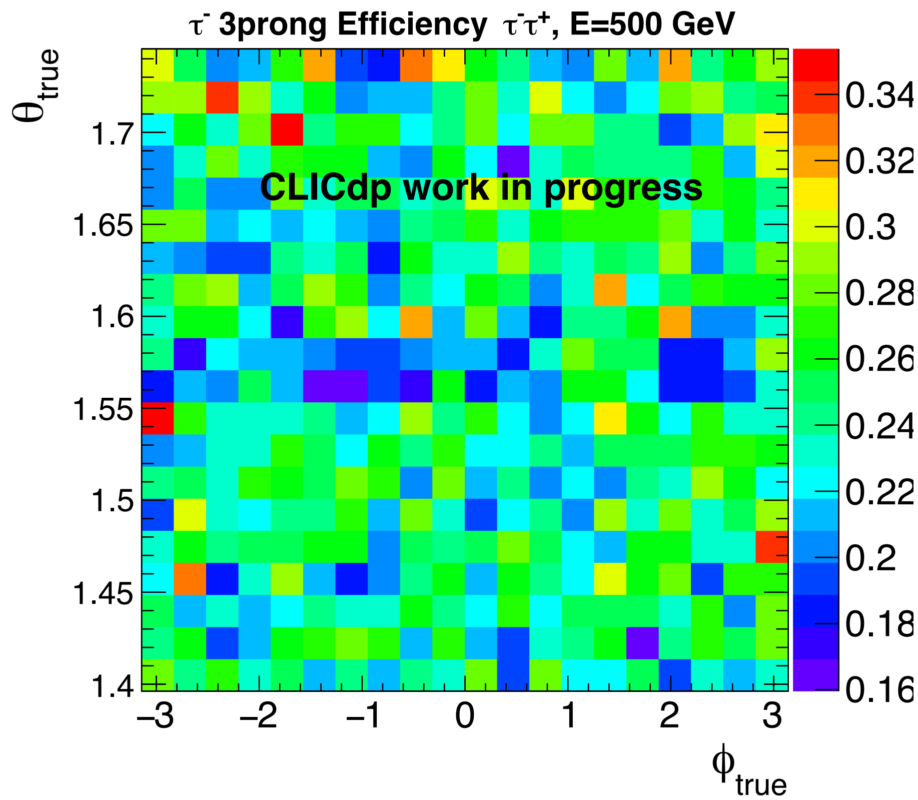

Tau-energy 200 GeV, 1 prong
Around 74-82 %

Tau-energy 200 GeV, 3 prong
Around 48-60 %

# Hadronic Tau Identification Efficiency

Large drop when going to higher energies → not understood so far
Reduced efficiency at polar angles of 90 degrees → not understood so far

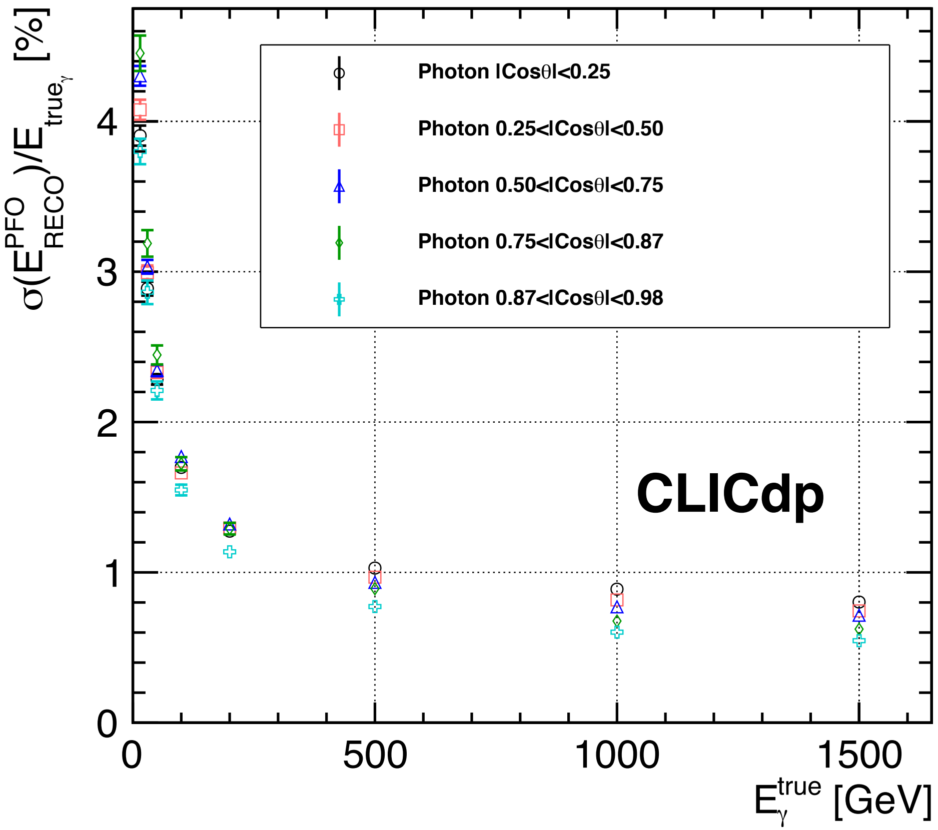


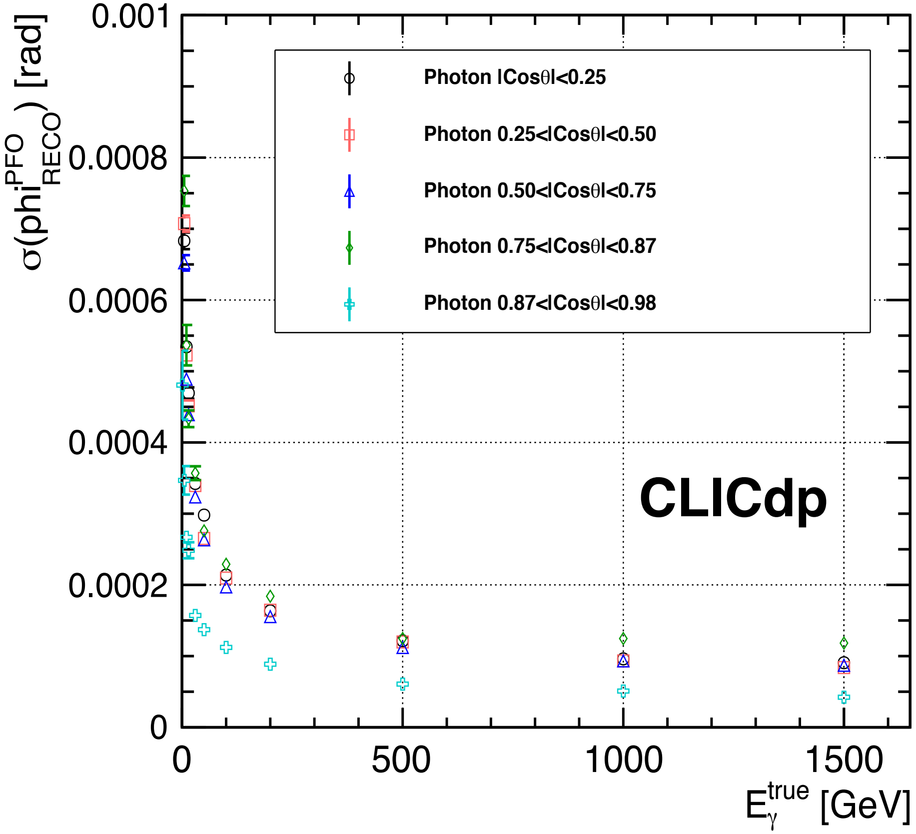

Tau-energy 500 GeV, 1 prong
Around 50-60 %

Tau-energy 500 GeV, 3 prong
Around 18-30 %

# Photon Performance: unconverted photons



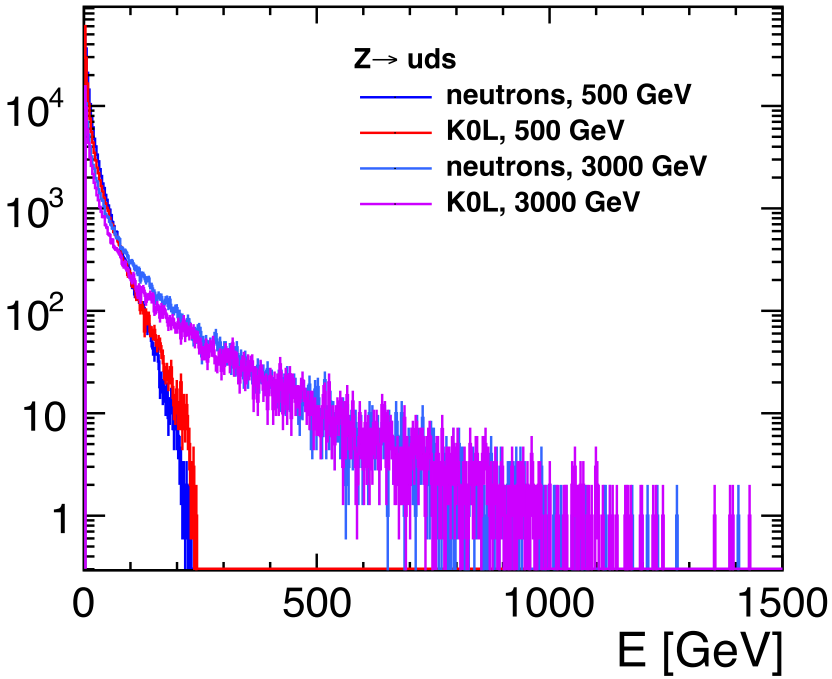

Relative Photon Energy Resolution in different detector regions

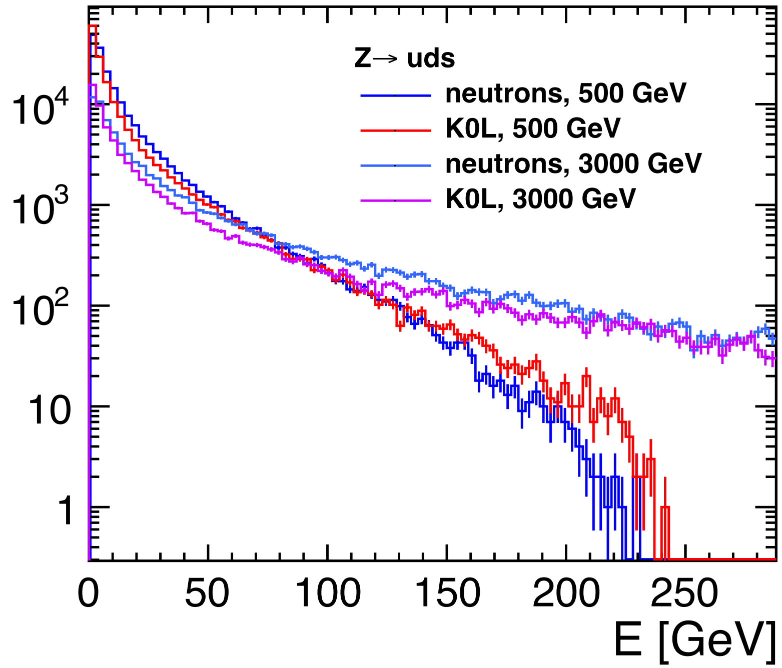



Photon phi position resolution in different detector regions as function of true photon energy

# MC truth Hadron Energy spectrum for CLIC

For 500 GeV dataset neutral hadron energies beyond 90 GeV are 1.9 %, for 3000 dataset 13.7 % → if we want same coverage of neutral hadron energy spectrum need to calculate weights for samples up to far higher energies (1.7 % of hadrons with energies beyond 400 GeV for 3000 GeV sample)




**Z→dijets at 500 vs 3000 GeV**

# Samples and Software used

Produce single particle gun samples of neutrons and K0L's separately, for each point simulate and reconstruct 70000 events
Use the PandoraSettingsSoftwareCompensationTraining script for reconstruction

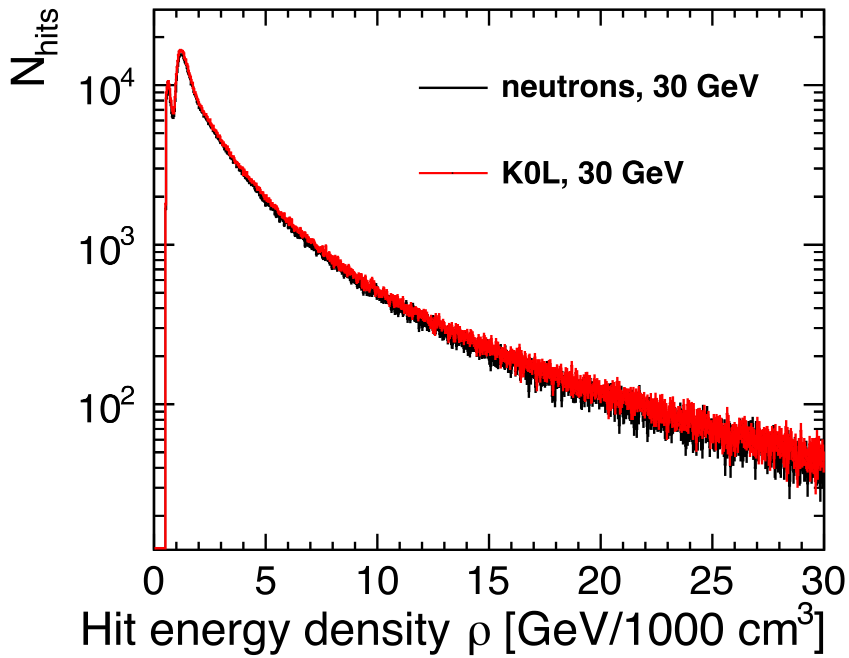
Cleaning of clusters in the Pandora training script identical to cleaning for default reconstruction
→Then run **PandoraPFACalibrate_SoftwareCompensation** script in PandoraAnalysis/calibration

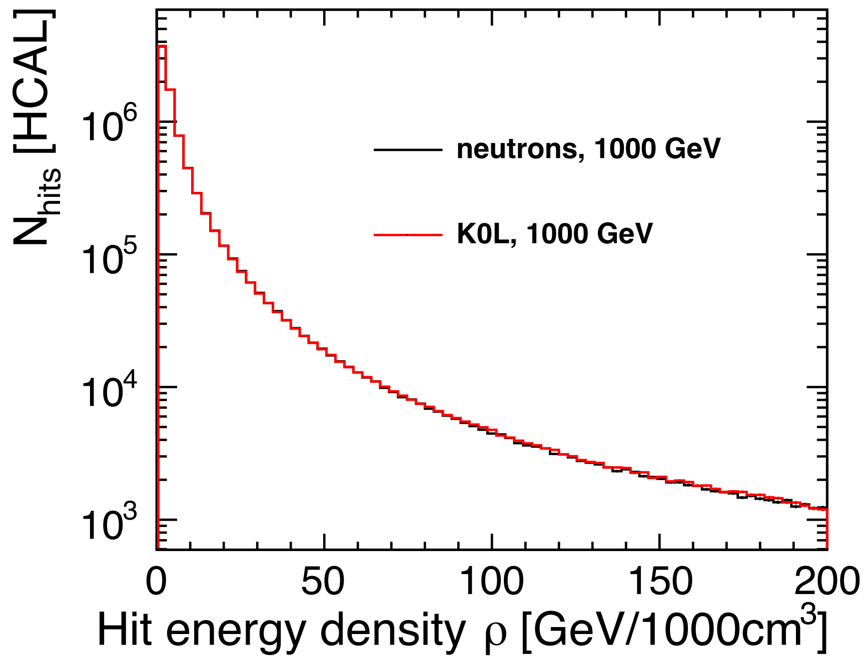
Merge Kaons and neutrons in one sample (relative weight 1:1) and use energy points of 2,5,10,30,75,150,200,400,**1000** for software compensation training

Density binning: 0 2 5 7.5 9.5 13 16 20 23.5 28 33 40 50 75 100, overflow 110

# Check on hit energy densities for very high K0L

So far did not yet change the binning of weights, maybe should extend weights to densities of 100/150 GeV/dm$^3$




The energy of calorimeter clusters are computed as:

$$E_{SC} = \sum_{hits} E_{ECAL} + \sum_{bin\ i} (E^i_{HCAL} \times \omega(\rho_i))$$

$$\text{with } E^i_{HCAL} = \sum_{hits \in bin\ i} E_{hit},$$

$$\omega(\rho) = p_1 \exp(p_2 \rho) + p_3$$

# Software Compensation

Follow the procedure from software compensation paper draft https://arxiv.org/pdf/1705.10363.pdf

$$p_1 = p_{10} + p_{11} \times E_{sum} + p_{12} \times E_{sum}^2$$
$$p_2 = p_{20} + p_{21} \times E_{sum} + p_{22} \times E_{sum}^2$$
$$p_3 = \frac{p_{30}}{p_{31} + e^{p_{32} \times E_{sum}}}.$$

The three parameters depend on the energy of the cluster
→weight shape changes quite a bit for different input hadron energies

→ Use automated procedure from PandoraAnalysis/calibration

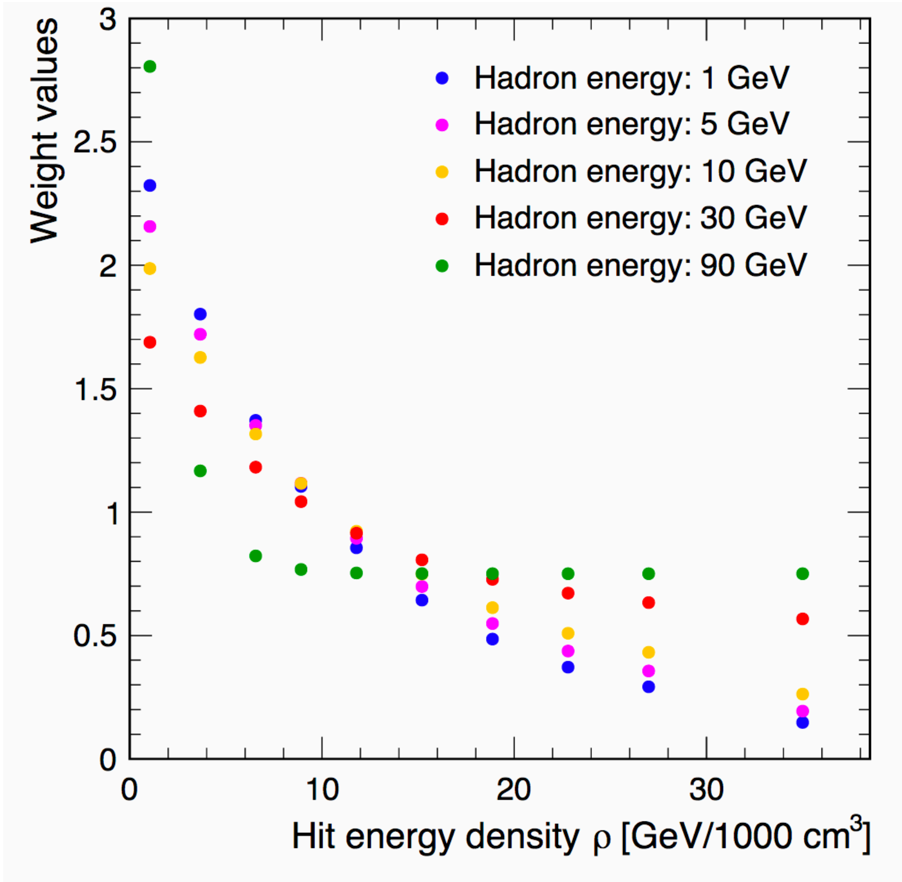



**Fig. 3** Software compensation weights as a function of hit energy density for different hadron energies, derived from Eq. 1.

# Particle Flow Analysis: PandoraPFA

Particle identification at CLIC is based on the Pandora Particle flow identification algorithm

- Pandora aims at reconstructing and identifying each particle reaching the detector
  - The particle types are charged pions, muons, electrons, photons and neutrons
- The main objective of Pandora algorithm is to achieve very excellent jet energy resolution, needed to achieve the desired precision involving hadronic final states

Documentation:

PandoraPFA algorithm: Nucl.Instrum.Meth A 611 (2009) 25

PandoraPFA at CLIC: Nucl.Instrum.Meth A 700 (2013) 153

Documentation of PandoraPFA Software Toolkit: EPJC 75 (2009) 439

# PandoraPFA: Confusion term

**Confusion term:** effect of all pattern recognition on reconstruction, which includes errors in clustering of calorimeter hits and associating of tracks to those cluster

→ Important contribution to jet energy resolution, particularly at large jet energies

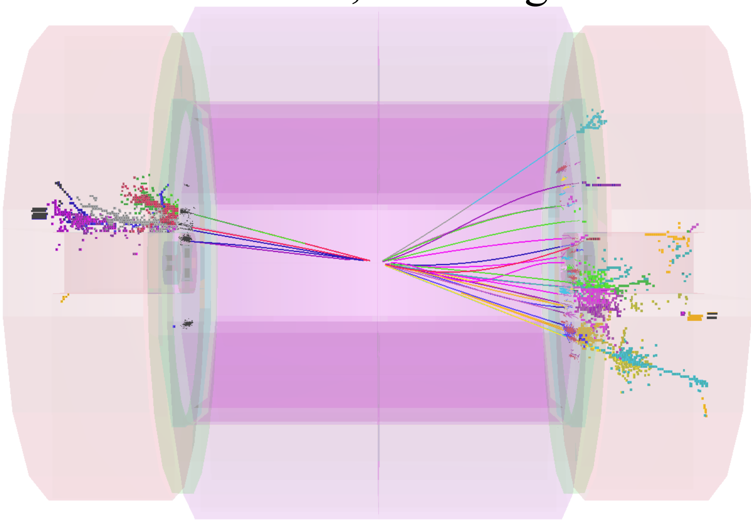
- Example1: energy of pion fluctuates low, then energy of a nearby "true" photon might be assigned to pion energy cluster as the sum matches the charged track momentum better, thus we measure less energy than we should have
- Example2: a lot of bremsstrahlung in the case of electrons: might give rise to additional photons, charged track from electron correctly measured, thus too large energy measured

Nonlinear non compensating natures of hadron calorimeters:

- idea to correct with software for (on average) larger response of hadron showers with large electromagnetic component, improves energy measurement of cluster energies, thus improves the confusion term
- Software compensation technique developed by CALICE → implemented in PandoraPFA now

# Example: Beamstrahlung background reduction

Forward WW, no background



after timing and $p_T$ cuts: ~ 100 GeV remain:

| | Loose configuration | | Tight configuration | |
|---|---|---|---|---|
| **Region** | **$p_T$ range [GeV]** | **Time [ns]** | **$p_T$ range [GeV]** | **Time [ns]** |
| **Photons** | | | | |
| Central | $0.75 \leq p_T < 4.0$ | $t < 2.0$ | $1.0 \leq p_T < 4.0$ | $t < 2.0$ |
| $\lvert\cos(\theta)\rvert \leq 0.975$ | $0 \leq p_T < 0.75$ | $t < 2.0$ | $0.2 \leq p_T < 1.0$ | $t < 1.0$ |
| Forward | $0.75 \leq p_T < 4.0$ | $t < 2.0$ | $1.0 \leq p_T < 4.0$ | $t < 2.0$ |
| $\lvert\cos(\theta)\rvert > 0.975$ | $0 \leq p_T < 0.75$ | $t < 1.0$ | $0.2 \leq p_T < 1.0$ | $t < 1.0$ |
| **Neutral hadrons** | | | | |
| Central | $0.75 \leq p_T < 8.0$ | $t < 2.5$ | $1.0 \leq p_T < 8.0$ | $t < 2.5$ |
| $\lvert\cos(\theta)\rvert \leq 0.975$ | $0 \leq p_T < 0.75$ | $t < 1.5$ | $0.5 \leq p_T < 1.0$ | $t < 1.5$ |
| Forward | $0.75 \leq p_T < 8.0$ | $t < 2.5$ | $1.0 \leq p_T < 8.0$ | $t < 1.5$ |
| $\lvert\cos(\theta)\rvert > 0.975$ | $0 \leq p_T < 0.75$ | $t < 1.5$ | $0.5 \leq p_T < 1.0$ | $t < 1.0$ |
| **Charged particles** | | | | |
| All | $0.75 \leq p_T < 4.0$ | $t < 3.0$ | $1.0 \leq p_T < 4.0$ | $t < 2.0$ |
| | $0 \leq p_T < 0.75$ | $t < 1.5$ | $0 \leq p_T < 1.0$ | $t < 1.0$ |

with background in recronstuction time window ~1.2 TeV