A Particle Flow Algorithm with hit density driven clustering

Lei Xia
ANL-HEP

About this PFA

- What's in it
  - Calibration of calorimeter (independent of PFA)
    - including Ron's angular correction
  - Clustering algorithm
    - hit density driven
  - Track-cluster matching
    - based on track-cluster distance, no E/p check
  - Sum up event energy – 'PFA result'
    - Charged objects (E = sum of P): tracks, clusters matched with tracks
    - Photons (E = sum of cluster energy): (id from MC info)
    - All other clusters (E = sum of cluster energy): nominal neutral hadrons
  - Cleanup of nominal neutral hadrons:
    - After track-cluster matching
    - use geometrical variables to distinguish and remove charge fragments from nominal neutral clusters
  - E/p check for clusters that matched to tracks
- What's still needed (currently using MC information)
  - 'photon ID'
    - Tells me whether a cluster is from an EM or a HAD shower
  - Track finding algorithm
  - Jet algorithm
- Detector mode
  - SiDaug05, SiDaug05_np
    - Si tracker, Si/W EM calorimeter, RPC/SS DHCAL projective/non-projective

Calorimeter calibration, angular correction

- EM showers/clusters

E = 82.5 x E(em,raw) + 0.11 x N(hadhits)

photon(GeV)	1	2	5	10	20	50	100
(σ/M)x√E	0.195	0.183	0.190	0.210	0.217	0.211	0.207

- Hadron showers/clusters

$$E = \frac{0.122 \times N_{HCalHit} + 114. \times E_{EMraw}}{1 + 0.00122 \times N_{HCalHit} + 1.34 \times E_{EMraw}}$$

K_L^0

GeV	1	2	5	10	20
(σ/M)x√E	0.47	0.55	0.60	0.55	0.48

neutron

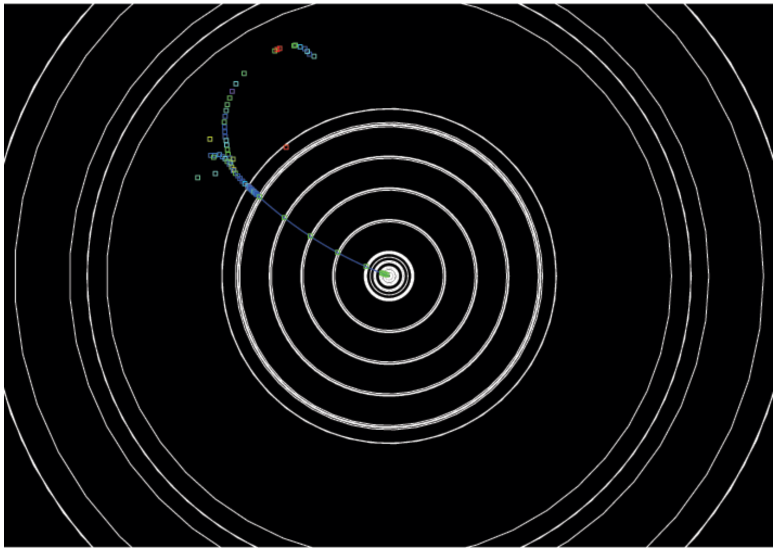
GeV	1	2	5	10	20
(σ/M)x√E	1.25	0.76	0.70	0.68	0.54

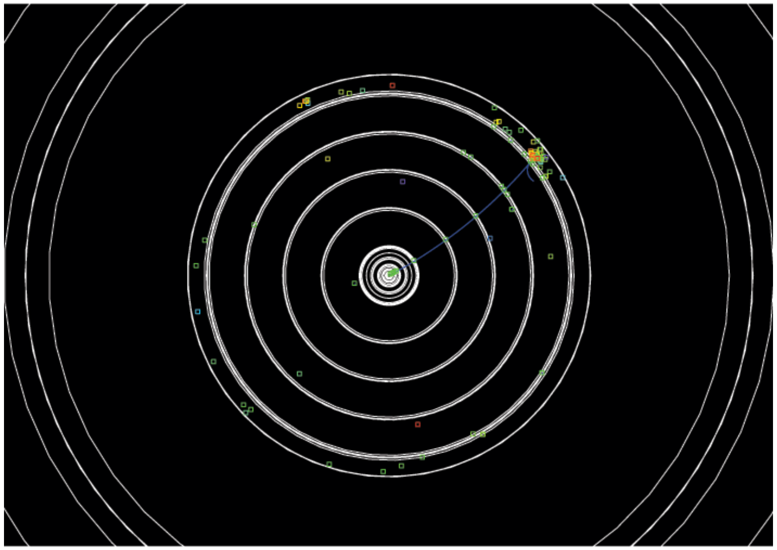
calibration done with single particles at 90 degrees

Ron's angular correction for cluster energy

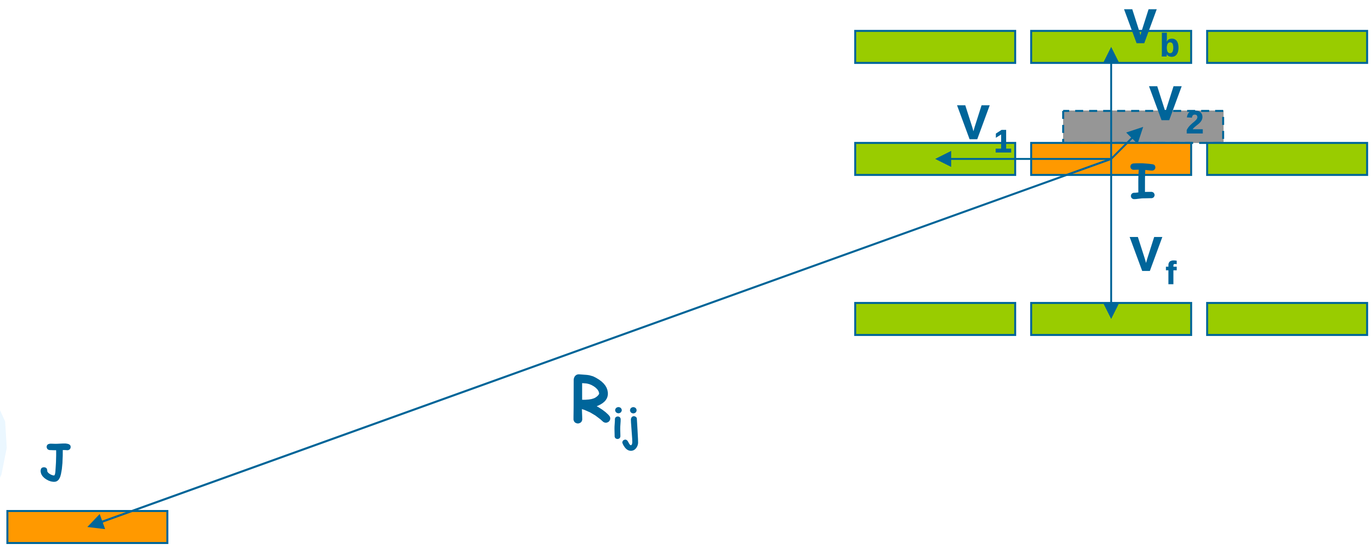
Clustering algorithm

- Goal: to have a clustering algorithm that can
  - Form clusters that can closely represent single particle shower
    - Pick up as many hits as possible for a single particle
    - Distinguish different particles
  - Treat ECal and HCal as one detector
    - Treat cell/layer structure properly
    - Cluster doesn’t break up at boundaries
  - Adjustable parameter for PFA
- Reality: hadron showers have hits all over the detector
  - Impossible to pick up every hits of a shower without messing up different showers
  - Try to pick up only the central part of a shower, and deal with fragments later
  - Use hit density to drive the clustering





Clustering: 2-hit density

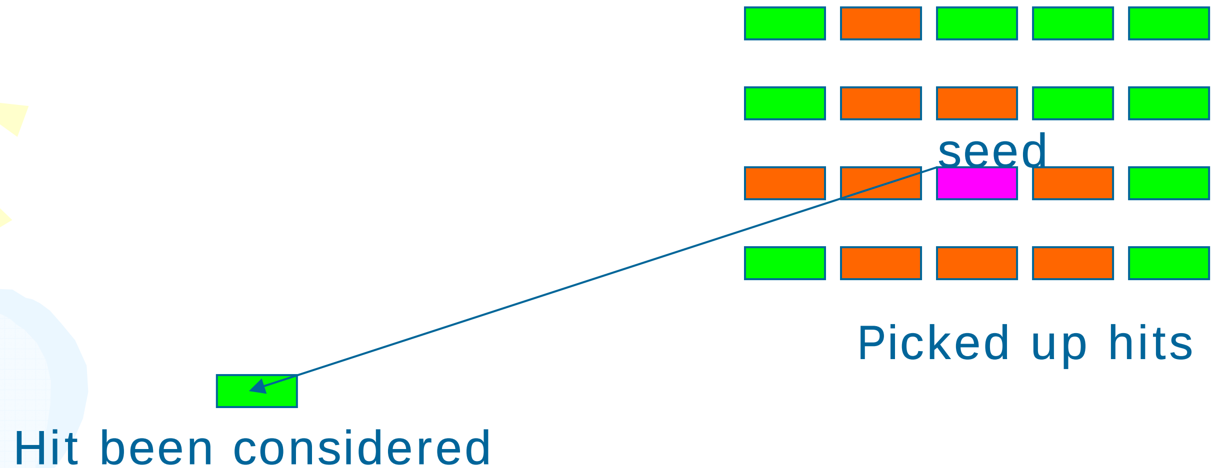


$$D_{ij} = e^{-(|(\hat{V}_1, R_{ij})|/|V_1|)} \times e^{-(|(\hat{V}_2, R_{ij})|/|V_2|)} e^{-(|(\hat{V}_3, R_{ij})|/|V_3|)}$$

With $V_3 = V_f$ (if $(V_f, R_{ij}) > 0$) or V_b (if $(V_b, R_{ij}) > 0$)

- Try to find a two-hit density function which can reflect the closeness of two hits
- Consider cell density variation in different directions
- Use distance normalized to local cell separation to calculate density
- It is a very local density function, only nearby hits contribute

Clustering: grow a cluster



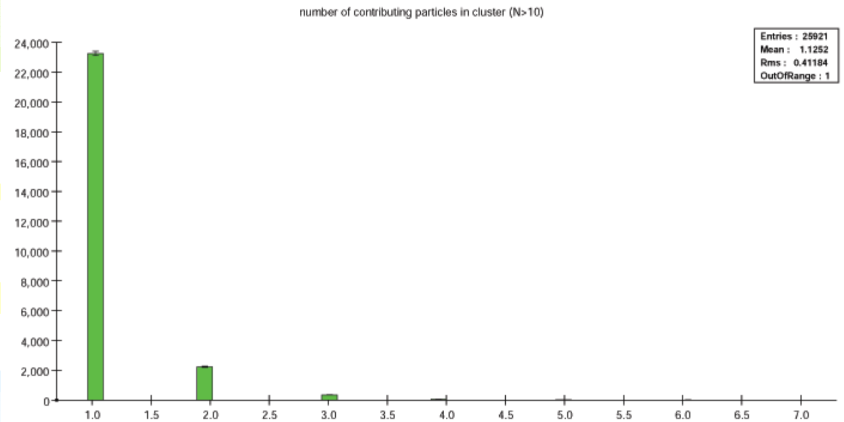
- Find a cluster seed: hit with highest density among remaining hits
- Attach nearby hits to a seed based on 2-hit density: seed cluster
- Attach additional hits based on (hit, seed cluster) density
  - EM hits, D(hit,cluster) > 0.01
  - HAD hits, D(hit,cluster) > 0.001
  - Grow the cluster until no hits can be attached to it
- Find next cluster seed, until run out of hits

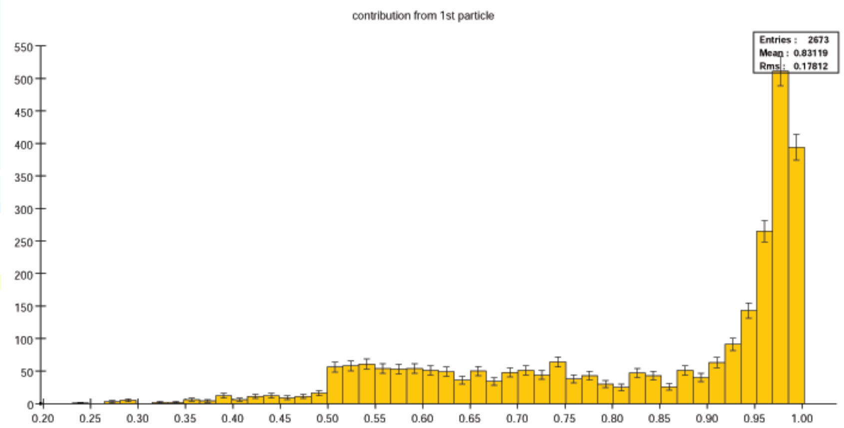
clustering efficiency: single particle

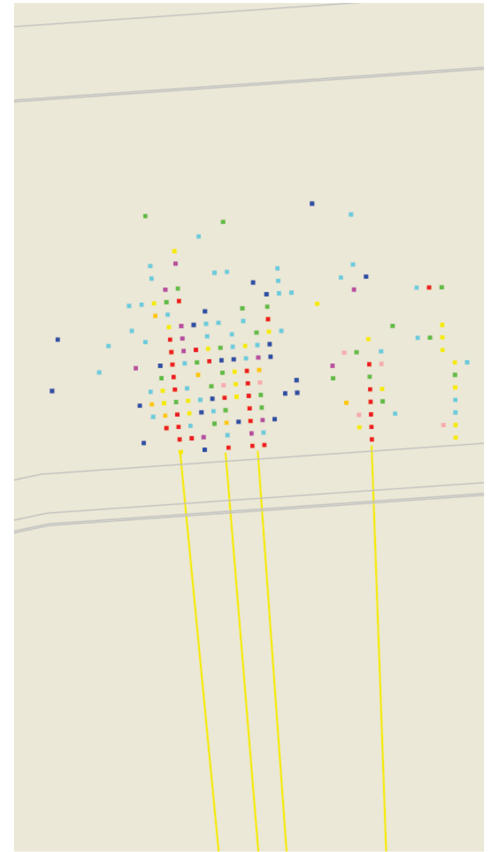
Particle	ECal hit efficiency	HCal hit efficiency	Overall hit efficiency	Overall energy efficiency
Photon (1GeV)	89%	43%	89%	91%
Photon (5GeV)	92%	54%	92%	96%
Photon (10GeV)	92%	61%	92%	97%
Photon (100GeV)	95%	82%	95%	>99%
Pion (2 GeV)	78%	59%	75%	71%
Pion (5 GeV)	81%	70%	79%	80%
Pion (10GeV)	84%	80%	83%	85%
Pion (20GeV)	85%	87%	88%	91%

- Typical electron cluster energy resolution ~ 21%/sqrt(E)
- Typical pion cluster energy resolution ~70%/sqrt(E)
- All numbers are for one main cluster (no other fragments are included)

Z pole uds events: cluster purity



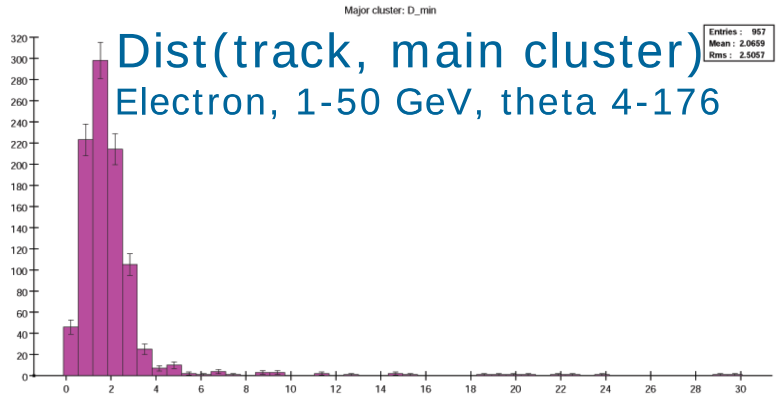


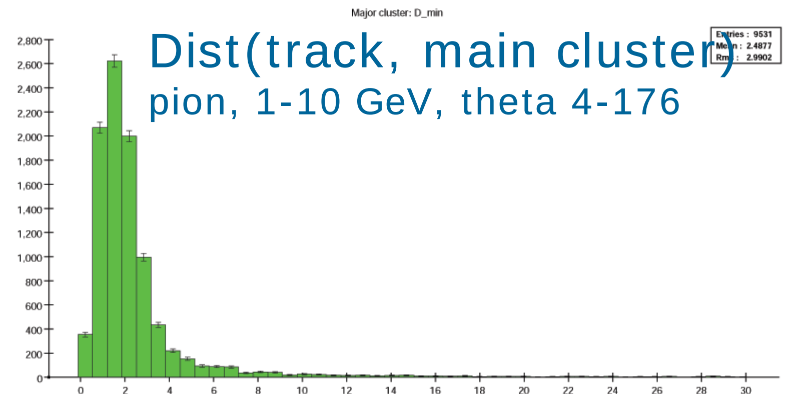


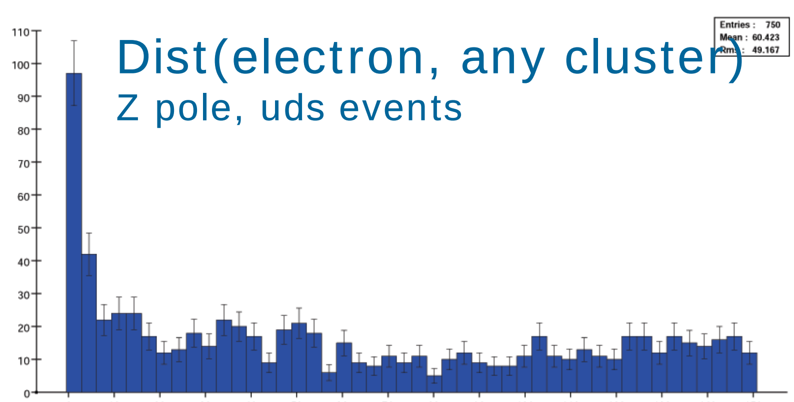
- Most of the clusters (89.7%) are pure (only one particle contributes)
- For the rest 10.3% clusters
  - 55% are almost pure (more than 90% hits are from one particle)
  - The rest clusters contain merged showers, part of them are 'trouble makers'
- On average, 1.2 merged shower clusters/Z pole event

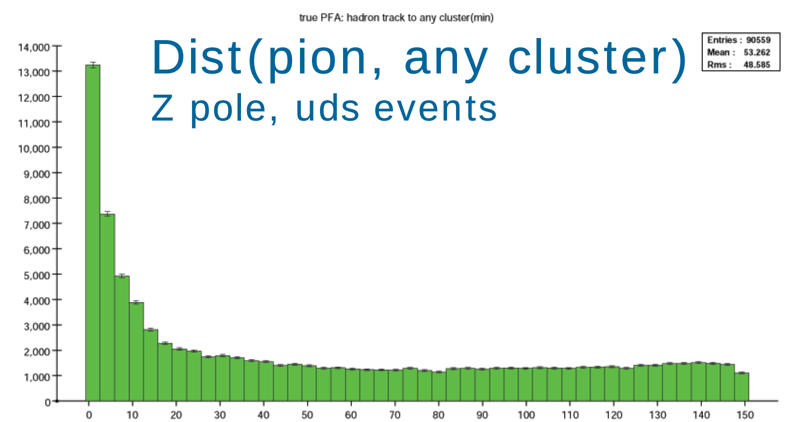
Track – cluster association

- Simple track cluster association
  - Extrapolate tracks into calorimeter by Helixswim (no ionization energy lost calculation)
  - Calculate min(track, cluster)
  - Electrons: min(track, cluster) < 6.mm
  - Hadrons: min(track, cluster) < 35.mm



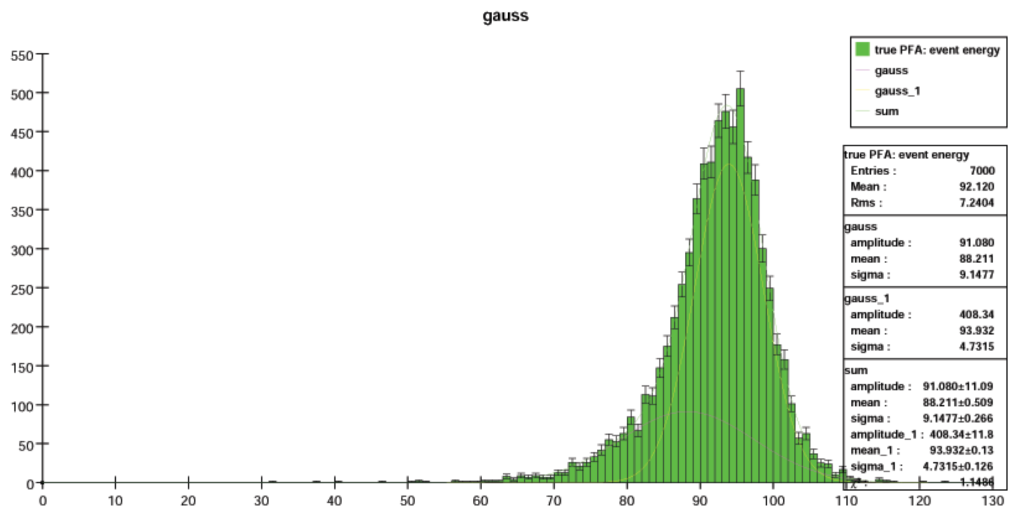






Try putting things together

- Tracking: any charged particle that goes into calorimeter
- Clustering: run clustering algorithm on calorimeter hits
- Associate track with cluster(s): charged clusters
- Look at MC info: photon clusters
- Clusters not associated to any tracks and not identified as photons: nominal neutral hadrons
- Total event energy = P(tracks) + E(photons) + E(neu had)



All 7k events:

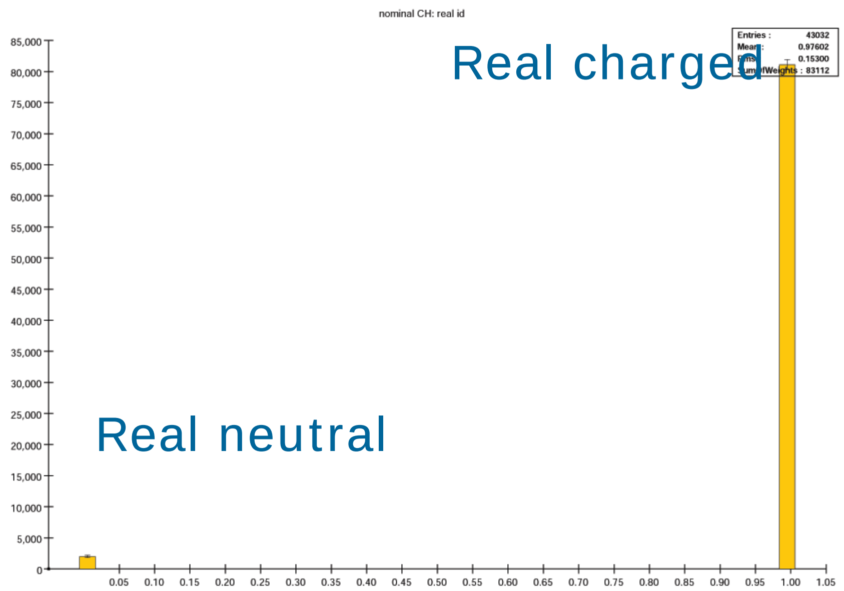
4.73 GeV 70%
at 94 GeV

9.15 GeV 30%
at 91 GeV

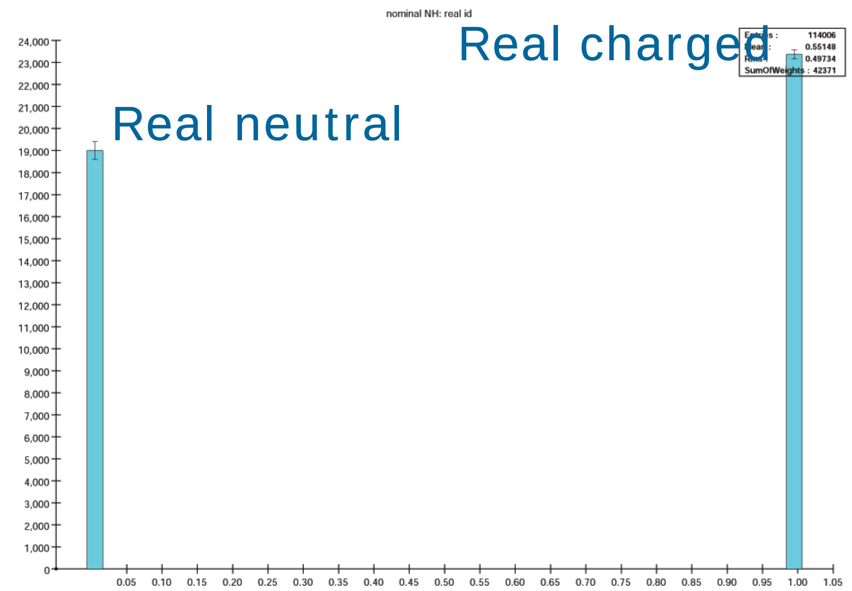
Central part: double counting of charged fragments in ‘neutral’
Big tail on the left: neutrals eaten by charged clusters

How well was the work done?

Nominal charged cluster energy



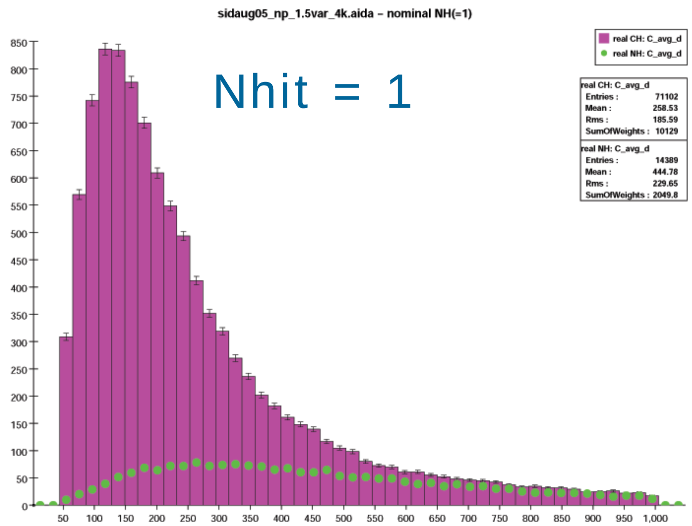
Nominal neutral cluster energy

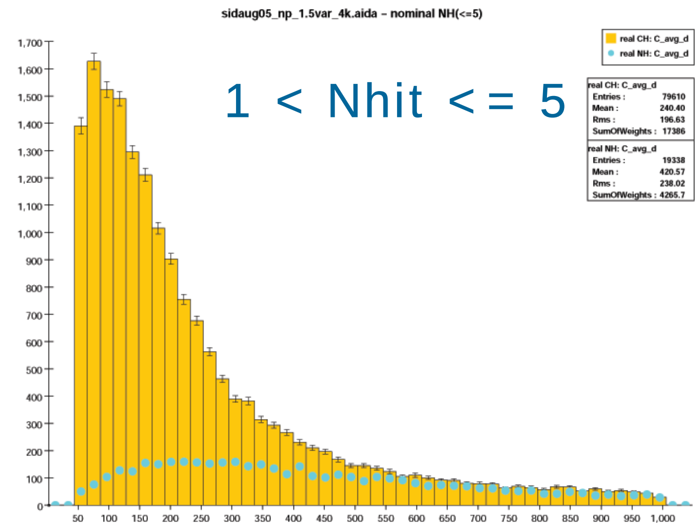


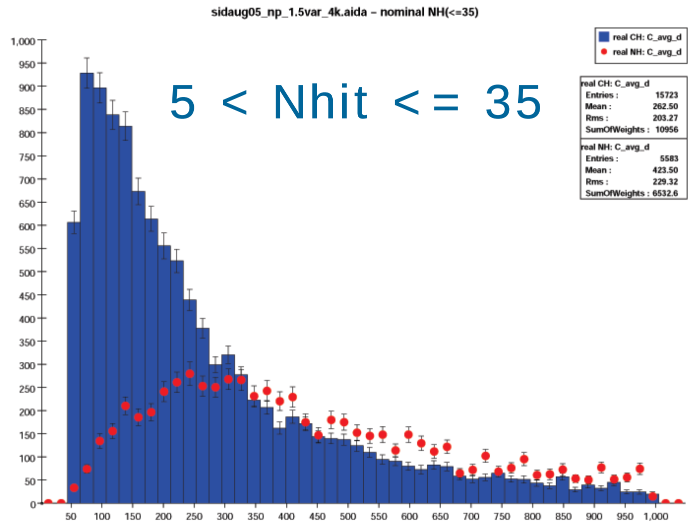
On average
~3% came from neutral

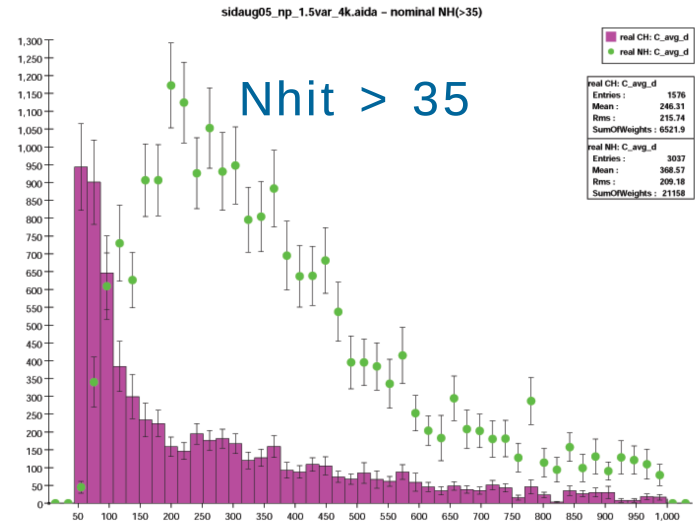
Energy from charged particles
is more than real neutral
-- need to work on it!

Nominal neutral clusters: distance to track



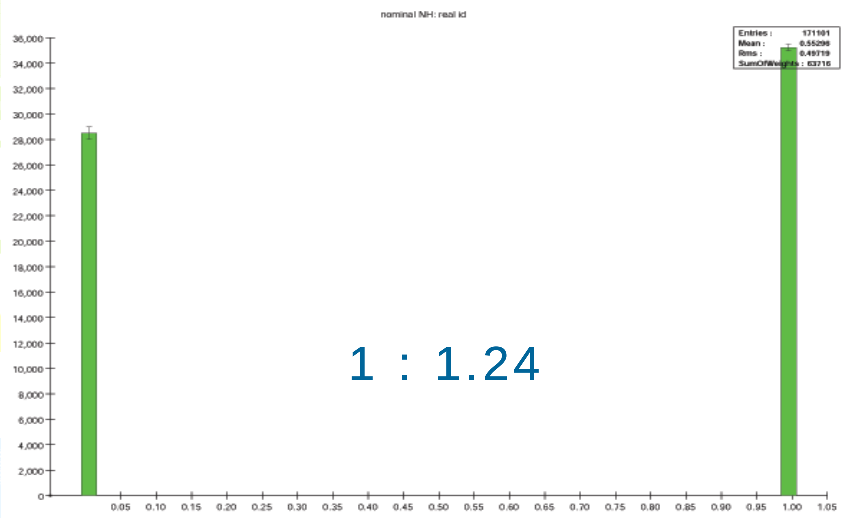




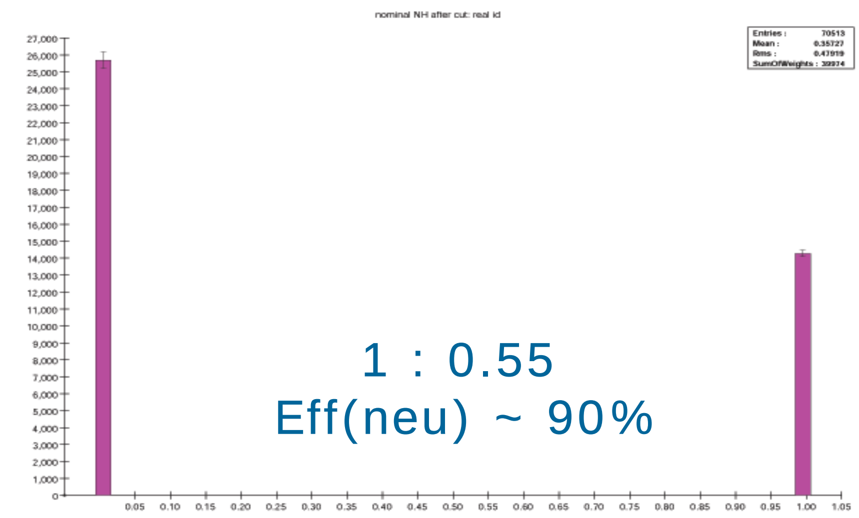


Charge fragment cluster vs. real neutral cluster

After simple cuts: charge fragments and real neutrals



1 : 1.24



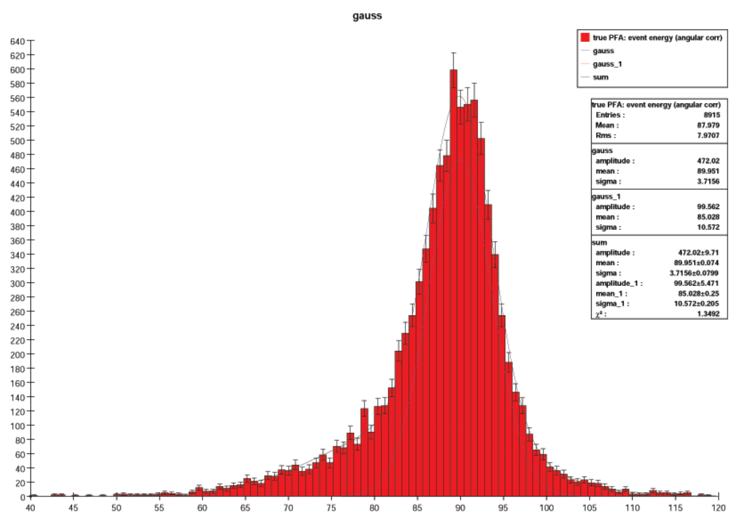
1 : 0.55
Eff(neu) ~ 90%

Cuts:

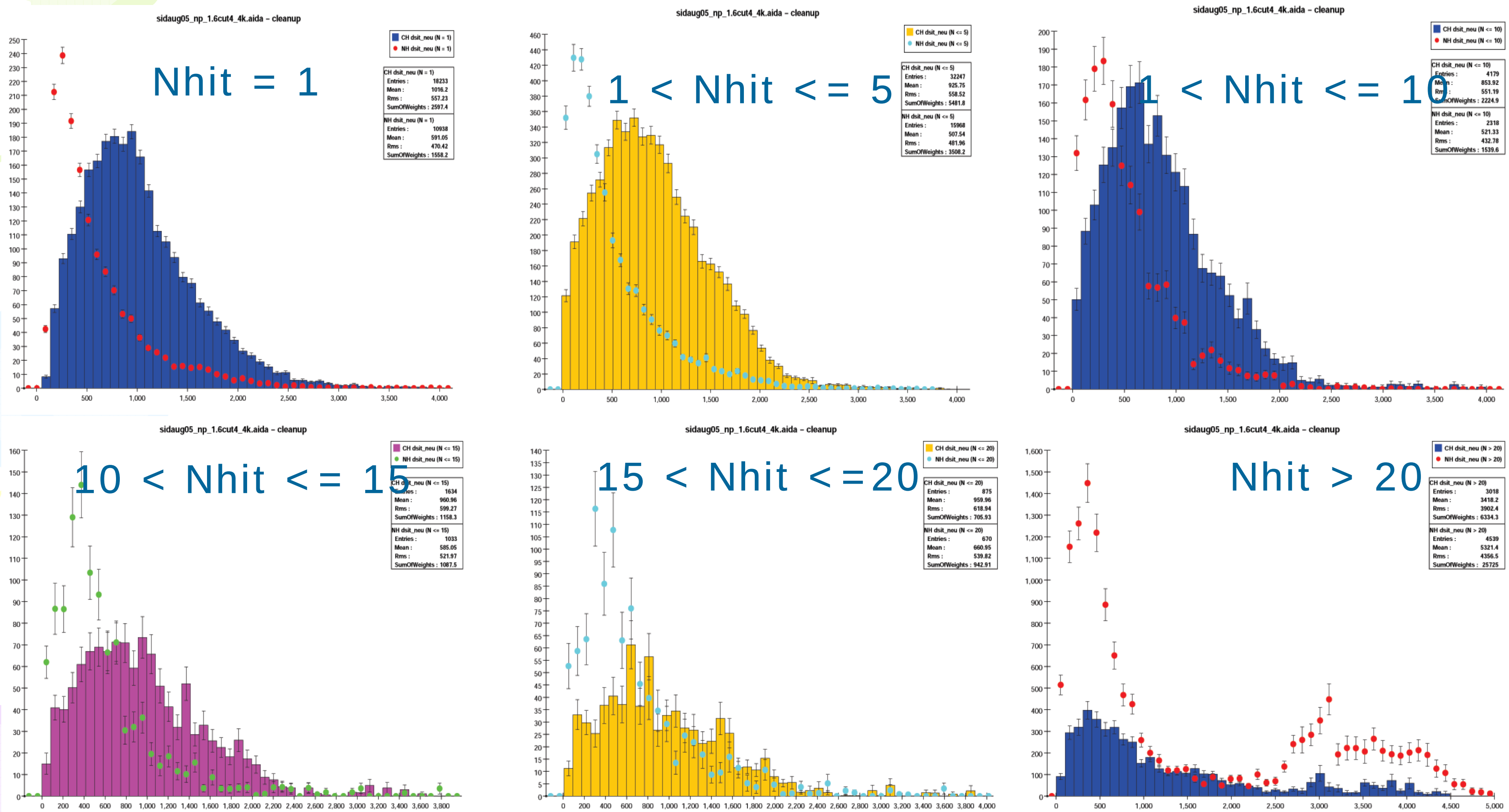
1. For Nhits = 1, cluster-track distance > 350mm
2. For 1<Nhits<=5, 300mm
3. For 5<Nhits<=35, 200mm
4. For Nhits>35, 75mm

All 9k events:

3.72 GeV @90.0GeV 62%
10.6 GeV 38%

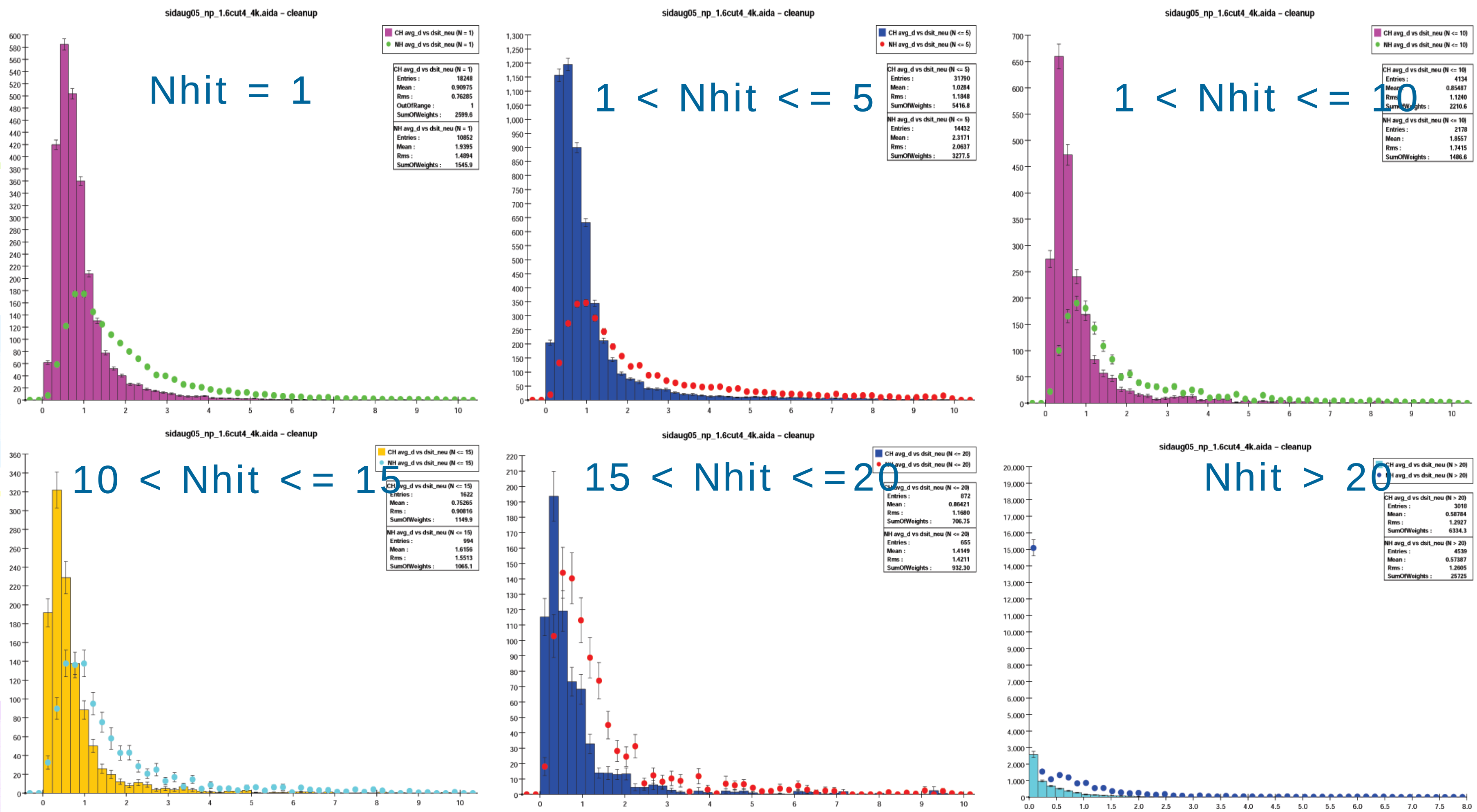


Other variables: distance to neutrals

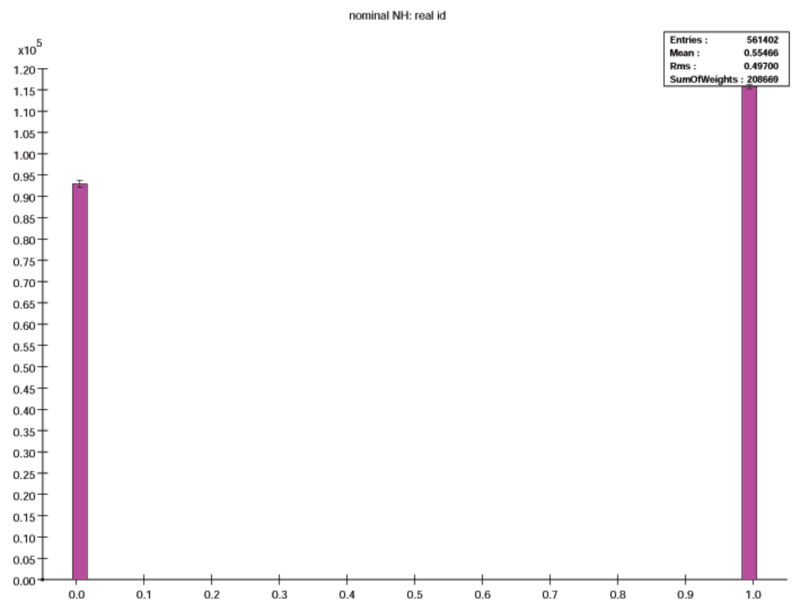


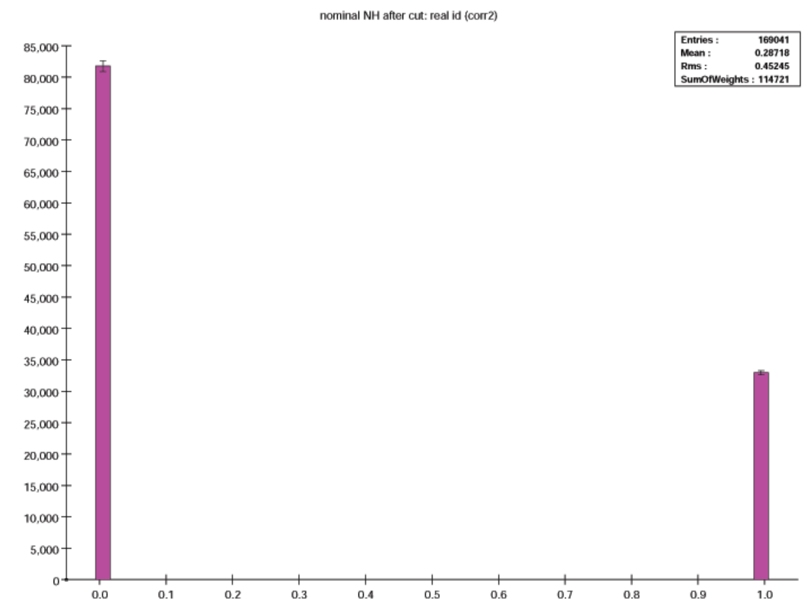
Neutral: any big enough cluster in ‘nominal neutral’

Other variables: ration of distances



Cuts on more variables





1 : 1.24

1 : 0.40

Eff(neu) ~ 88%

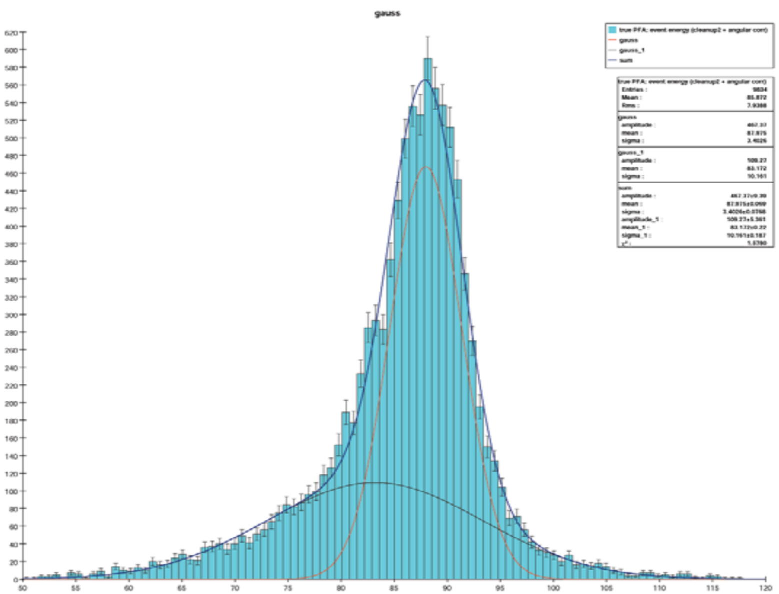
Cuts:

1. For Nhits = 1, ratio of distances < 0.70
2. For 1<Nhits<=5, 0.60
3. For 5<Nhits<=10, 0.50
4. For 10<Nhits<=15, 0.35
5. For 15<Nhits<=20, 0.20
6. For Nhits>20, no cut

Neutral:

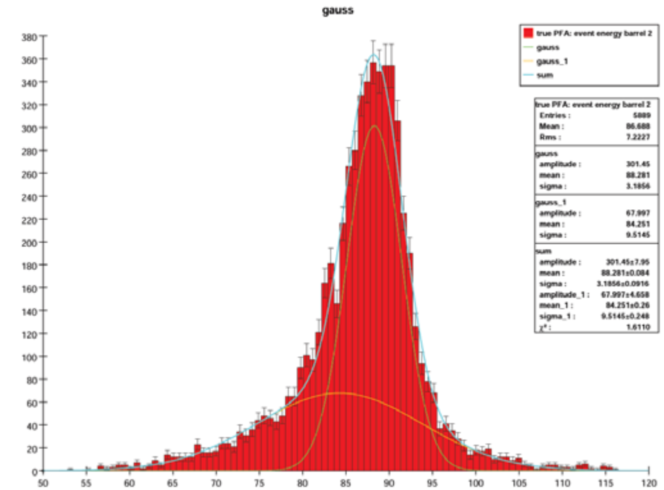
Any cluster with Nhits > max (10, Nhits,current)

PFA: after more cuts

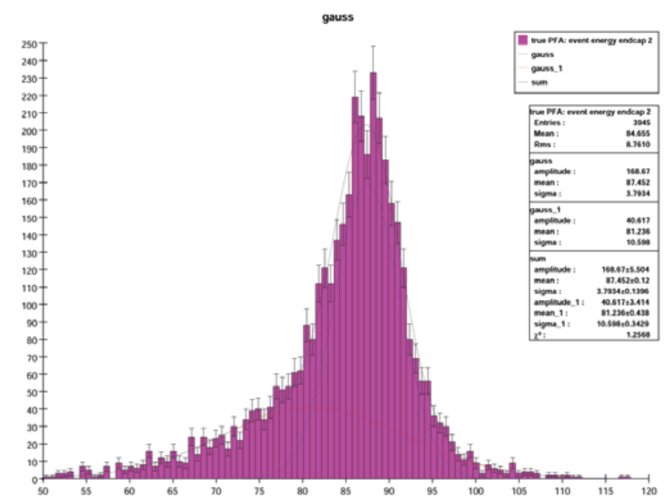


All 10k events:

3.40 GeV @88.0GeV 59%
10.2 GeV 41%



Barrel events: 60%
3.19 GeV @88.3GeV 60%
9.51 GeV 40%



Endcap events: 40%
3.79 GeV @87.5GeV 60%
10.6 GeV 40%

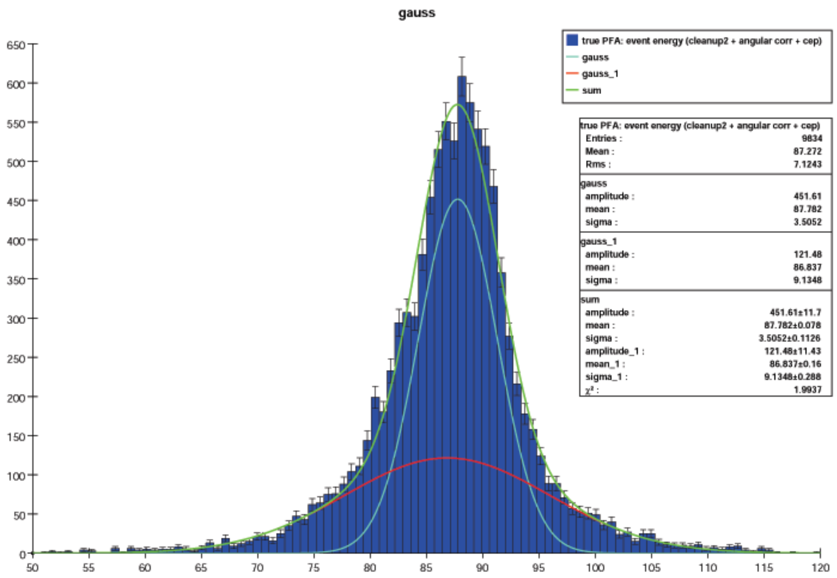
Try to improve tail: E/p correction

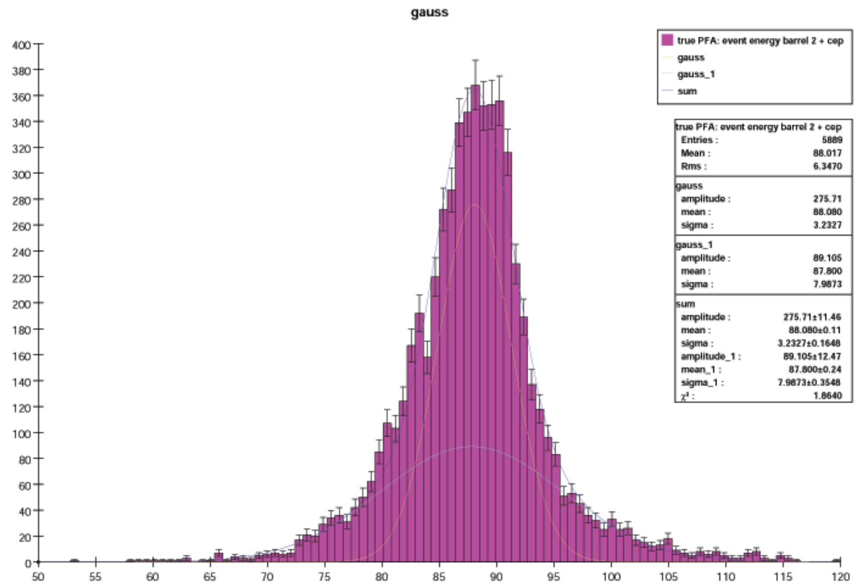
- Assuming 60%/sqrt(E) ECal+HCal combined resolution for charged hadrons

- Look at charge clusters:
  - If E(calorimetry) – P > 3.5sigma
  - If E(calorimetry) – P > 2 GeV
  - If E(calorimetry) > 5 GeV
  - If P > 3 GeV

- Then believe: there must be neutral particle being absorbed into this charge cluster
  - make correction dE = E(calorimetry) – P
  - Equivalent to measure this part of the shower with calorimeter only

E/P correction





All 10k events:

3.51 GeV @87.8GeV 59%
9.14 GeV @86.8GeV 41%

Barrel events:

3.23 GeV @88.1GeV 56%
7.99 GeV @87.8GeV 44%

Improved overall RMS and the wide Gaussian
Two Gaussians at ~ same position
Still big tails, now on both sides

Status

- Photon id
  - will put in Norman’s H-Matrix
- Source code
  - In cvs, but it is a messy single file
  - Will re-write according to recommended template
- Will try to improve the tails on the distribution