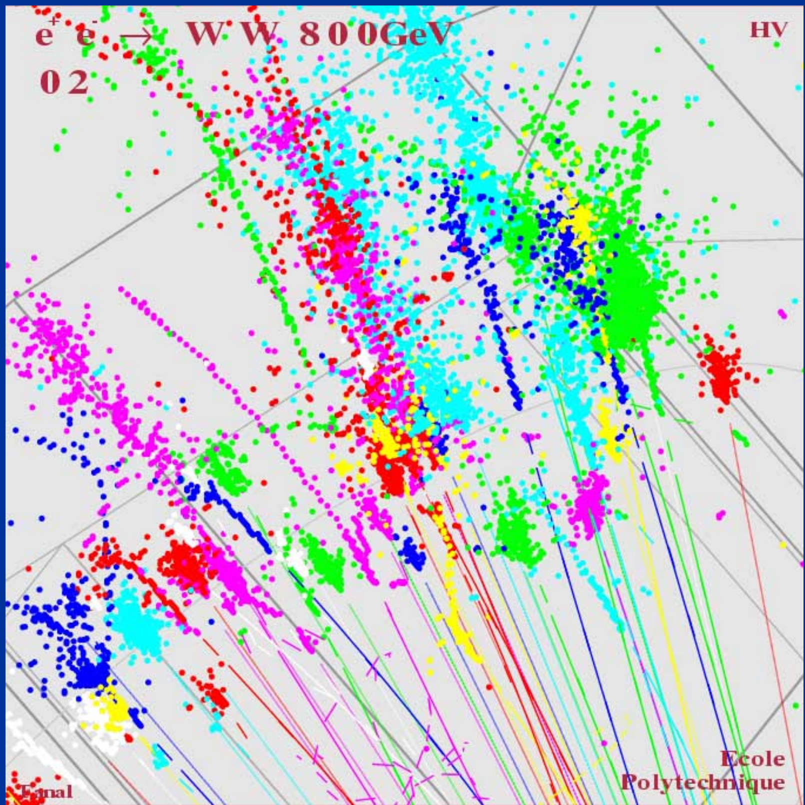
Development of a Particle Flow Algorithms (PFA) at Argonne



Presented by

Lei Xia
ANL - HEP

Introduction

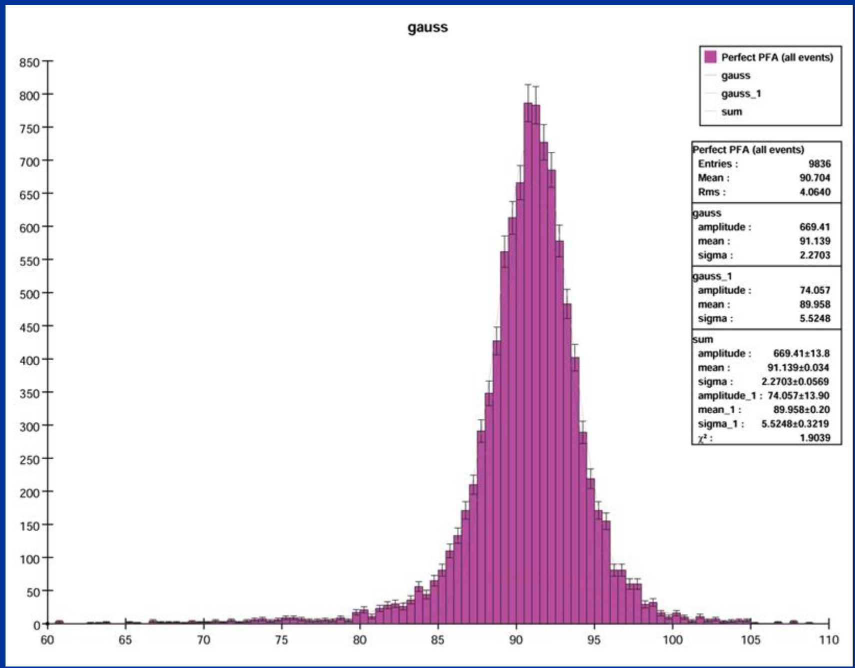
- Measure jets in the PFA way…

Particles in Jets	Fraction of jet energy	Measured with
Charged	65%	Tracker, negligible uncertainty
Photon	25%	ECal, 15%/ √ E
Neutral hadron	10%	ECal + HCal, ~50-60%/ √ E

- Clear separation of the 3 parts is the key issue of PFA
  - Charged particle, photon and neutral hadron: all deposit their energy in the calorimeters
  - Maximum segmentation of the calorimeters is needed to make the separation possible
- One Major R&D issue: development of PFA
  - Show that the ILC goal for jet energy resolution (30%/sqrtE) can be achieved by PFA
  - Develop a PFA that can be used for detector optimization
- Argonne has two parallel efforts on PFA development, this talk shows result from one of them

Perfect PFA: NO algorithm effect

- Take MC track momentum as the energy of charged particles
- Remove calorimeter hits associated with charged particles by looking at MC infomation
- Sum up everything else in the calorimeter as neutral energy
  - Apply appropriate sampling fractions for photon hits and neutral hadron hits
  - Use MC information to separate photon hits and neutral hadron hits
- Z-pole events, just event energy sum, no jet algorithm applied

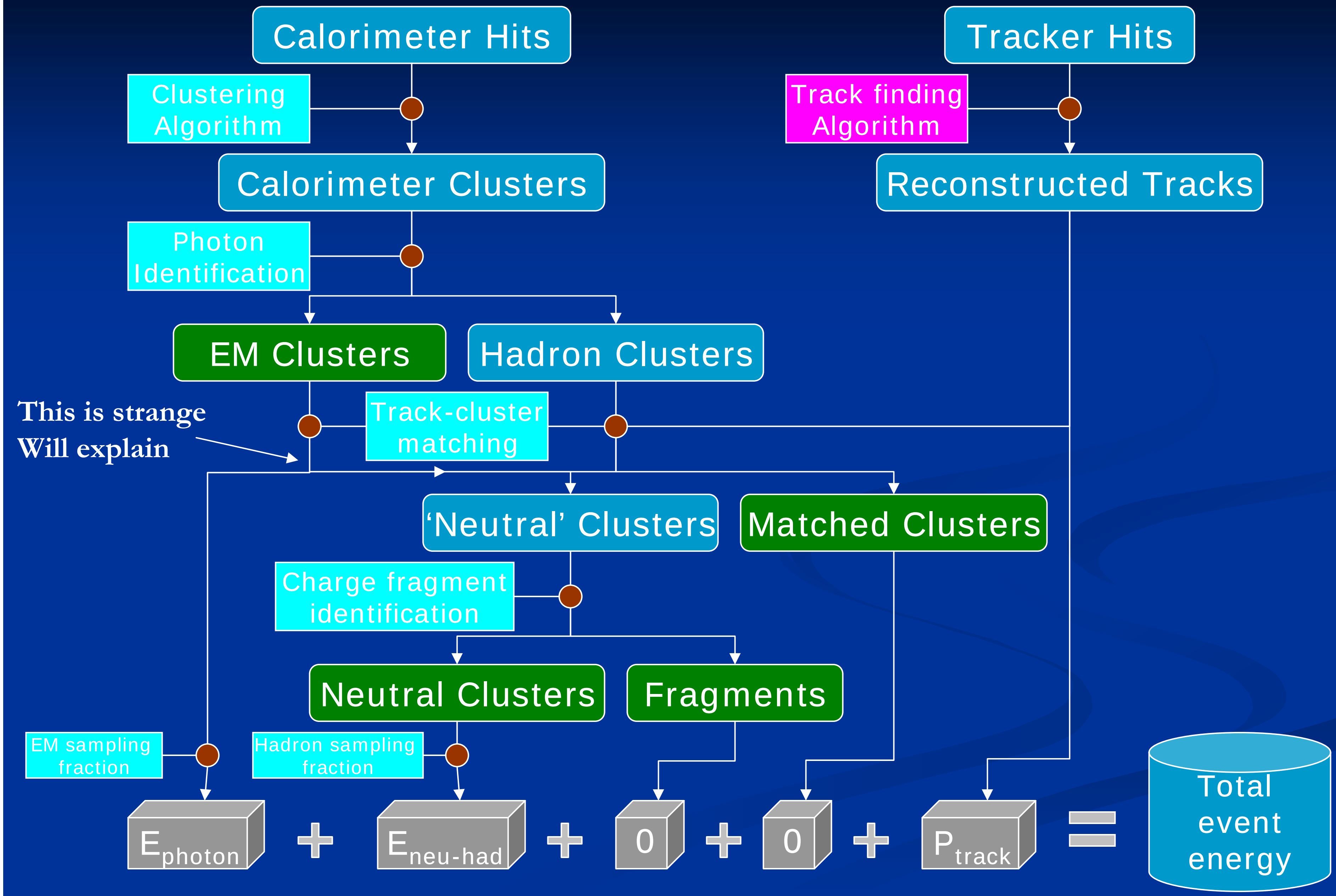


Example: SiD aug05_np

central peak
~2.3 GeV (~80%)
(no event selection)

We have room for PFA development

PFA effort: overview



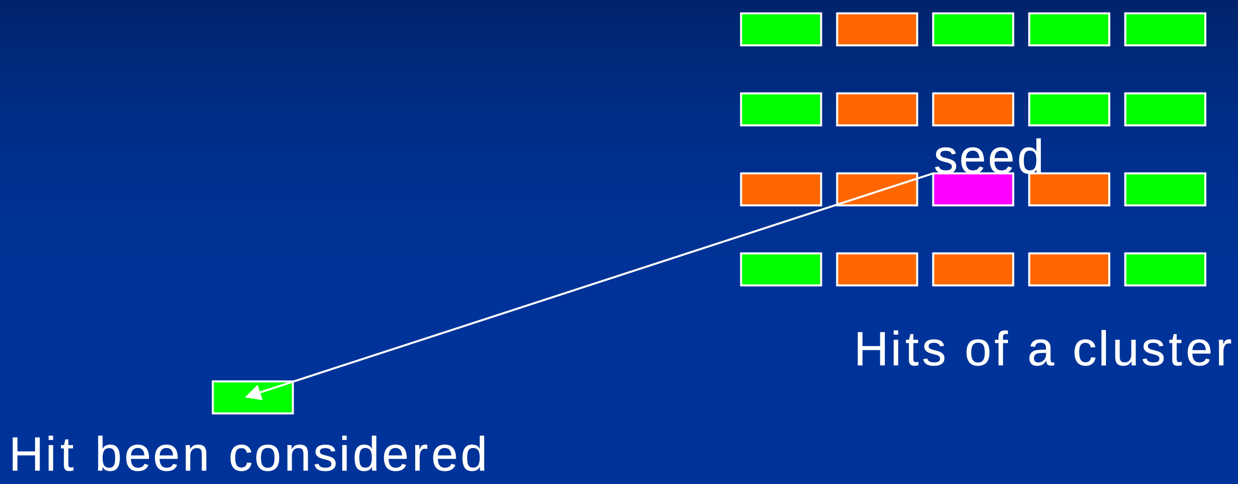
Clustering algorithm: hit density

$$D_i = \sum_{\{j \neq i\}} \left(e^{-(|\vec{V}_1 \bullet \vec{R}_{ij}|/|\vec{V}_1|^2)} \times e^{-(|\vec{V}_2 \bullet \vec{R}_{ij}|/|\vec{V}_2|^2)} \times e^{-(|\vec{V}_3 \bullet \vec{R}_{ij}|/|\vec{V}_3|^2)} \right)$$

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

- Hit density reflects the closeness from one hit i to a group of hits {j}
  - {j} = {all calorimeter hits} to decide if hit i should be a cluster seed
  - {j} = {all hits in a cluster} to decide if hit i should be attached to this cluster
- Consider cell density variation by normalizing distance to local cell separation
  - Density calculation takes care of the detector geometry
  - Clustering algorithm then treat all calorimeter hits in the same way

Clustering algorithm: grow a cluster



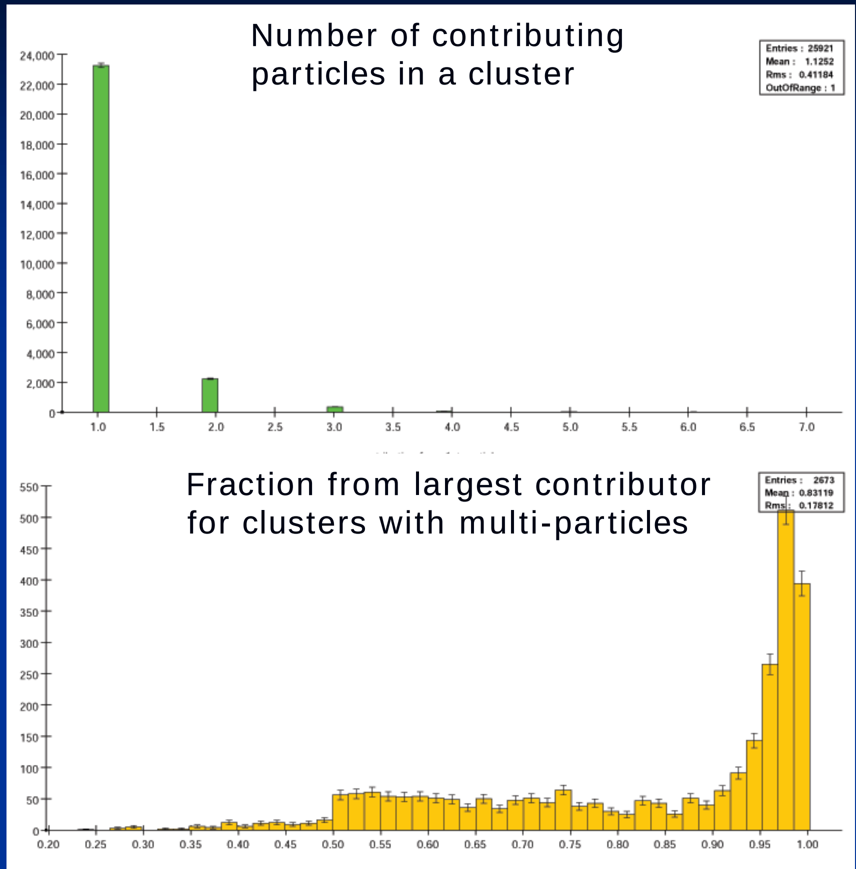
- Find a cluster seed: hit with highest density among remaining hits
- Attach nearby hits to a seed to form a small cluster
- Attach additional hits based on density calculation
  - i = hit been considered, {j} = {existing hits in this cluster}
  - EM hits, $D_i > 0.01$
  - HAD hits, $D_i > 0.001$
  - Grow the cluster until no hits can be attached to it
- Find next cluster seed, until run out of hits

Density driven clustering

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%

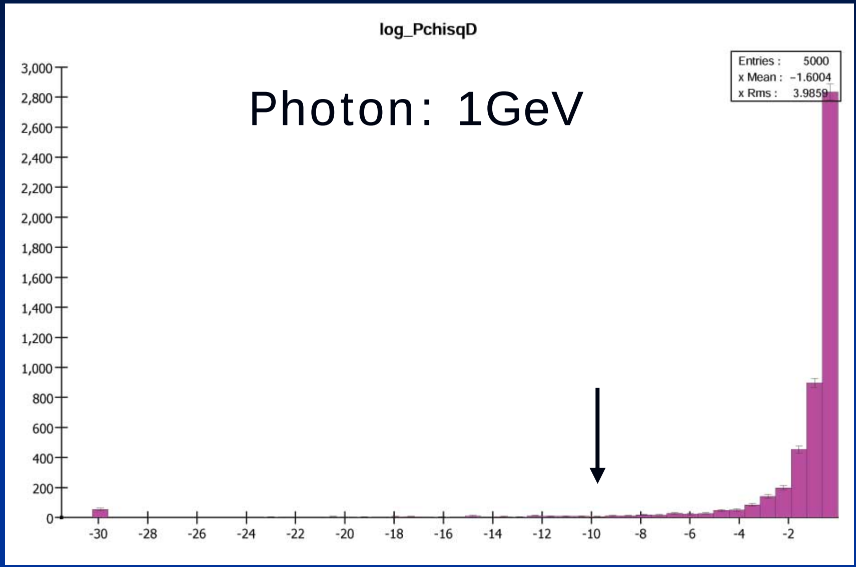
- Typical electron cluster energy resolution ~ 21%/sqrt(E)
- Typical pion cluster energy resolution ~70%/sqrt(E)
- All numbers are for one main cluster (no other fragments are included)

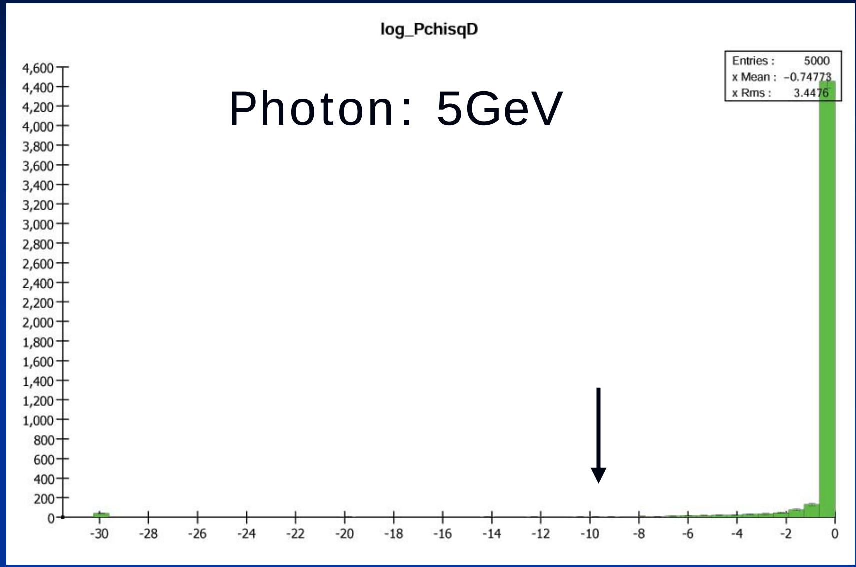
Cluster purity : Z pole (uds) events



- Most of the clusters (89.7%) are pure (only one particle contributes)
- For the remaining 10.3% clusters
  - 55% are almost pure (more than 90% hits are from one particle)
  - The remaining clusters contain merged showers, some of them are ‘trouble makers’
- On average, 1.2 merged shower clusters/Z pole event
  - This will result in double counting or underestimating of jet energy which leads to poor resolution
  - Will re-visit clustering algorithm after other PFA components are more of less settled

Photon id – longitudinal H-matrix





If I use a single cut: log(Prob(chisqD)) > -10

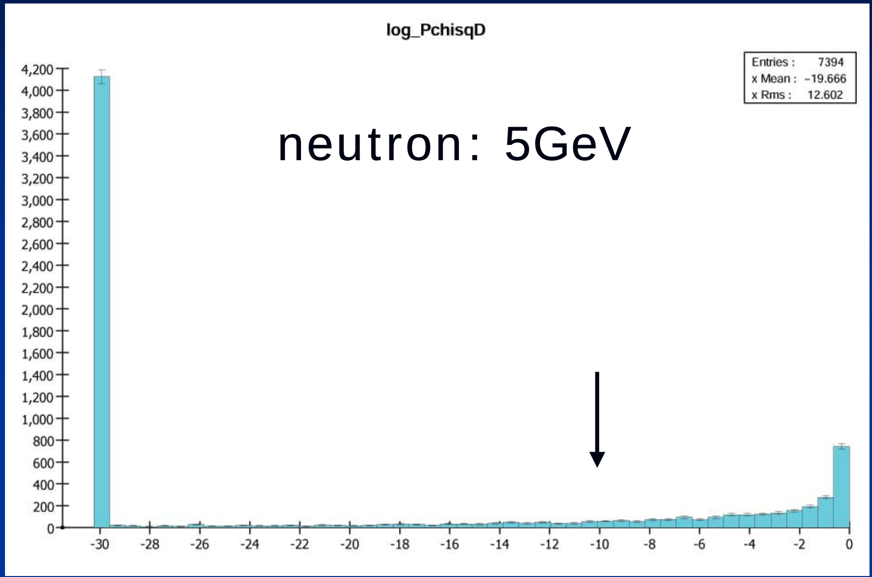
log(Prob(chisqD)) is calculated from default H-matrix accumulated from a wide energy range of photons

E_{photon}	0.10	0.25	0.50	0.75	1.0	2.0	5.0	10.	20.	50.	100.
Eff(%)	1.13	43.9	89.7	95.7	96.8	98.4	98.3	96.6	93.9	83.6	52.5

Trouble comes in at the two ends

But that's not all yet…

Photon id – longitudinal H-matrix



E	1.0	2.0	5.0	10.	20.
eff_{K0} (%)	26.1	30.8	33.2	36.1	35.2
eff_{n} (%)	6.61	18.6	33.0	38.	37.7
eff_{nbar} (%)	49.1	48.8	46.9	40.9	38.5

E(Pi-)	2.0	5.0	10.	20.
Eff (%)	13.3	8.1	7.0	3.8

- Efficiency of photons still need to improve
  - Try to accumulate longitudinal H-matrix(es) at smaller photon energy regions
- Efficiency of hadrons is way to high
  - Take neutral hadrons as photons results in using wrong calibration constant
  - Take charged hadrons as photons results in double counting of energy directly
    - Current (temporary) solution: still subject 'photons' to track-cluster matching which will recover charged hadrons but as the same time will lose some real photons
  - Will use more variables to eliminate hadrons: first IL layer, shower depth/shape, etc.
- A single cut on longitudinal H-matrix is NOT enough to identify photons

Photon id – try H-matrix(E)

Still use a single cut: log(Prob(chisqD)) > -10
But log(Prob(chisqD)) is calculated from individual H-matrix accumulated at the same energies of the photons

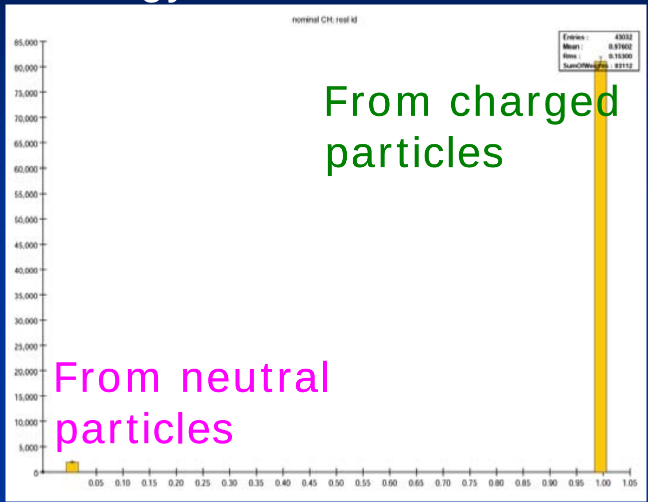
E_{photon}	0.10	0.25	0.50	0.75	1.0	2.0	5.0	10.	20.	50.	100.
Eff(%)	95.6	95.9	96.9	97.4	97.6	97.4	98.5	98.3	97.9	97.1	97.0

- Efficiency of photon looks much better, however, at the cost of accepting even more hadrons (not shown here)
- The energy range that one H-matrix can cover is still to be studied
- I am studying other variables to remove hadrons
  - First interaction layer, shower size/shape, etc

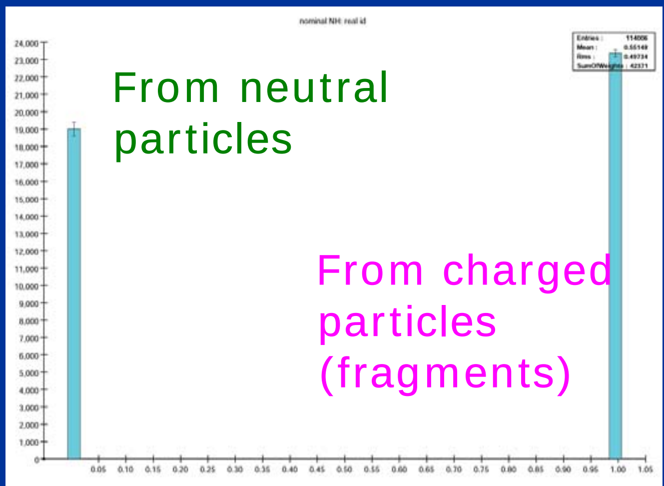
- Eventually, after photon identification is done rather well, I will no longer subject identified photons to track-cluster matching

Charge fragment identification/reduction

Energy of matched clusters



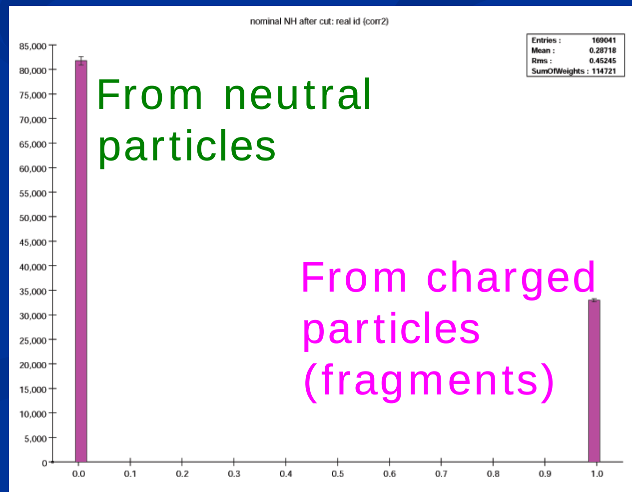
Energy of clusters not matched to any track: neutral candidate



1 : 1.24

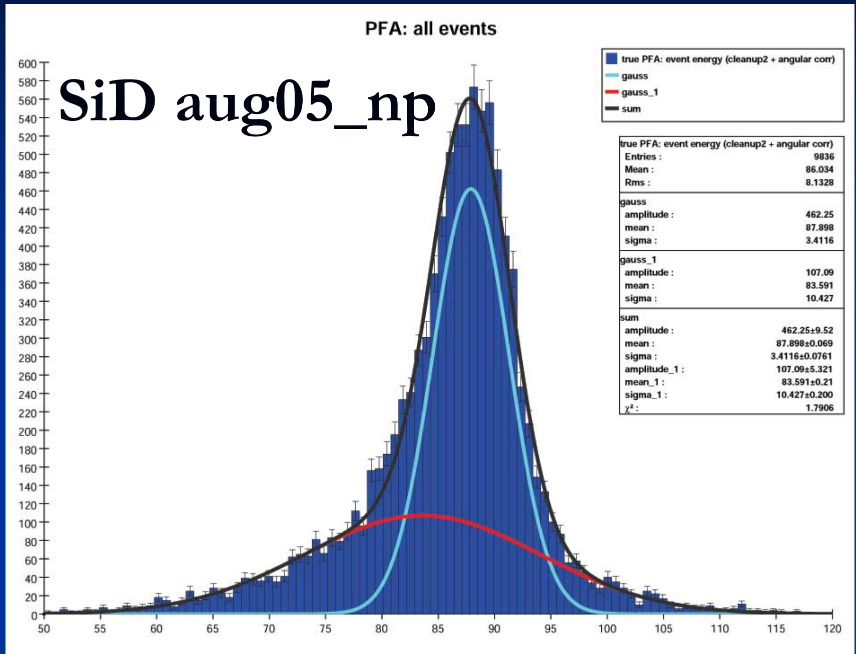
- Use geometrical parameters to distinguish real neutral hadron clusters and charge hadron fragments

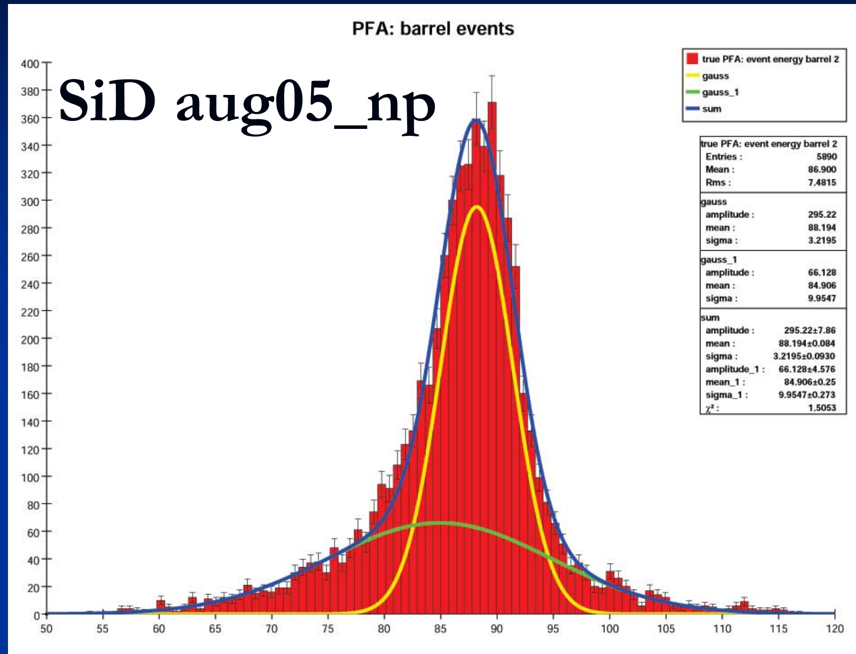
After charge fragment identification/reduction



0.88 : 0.35

PFA: Z-pole (uds) performance





All events:

3.41 GeV @87.9GeV 58.5%
10.4 GeV 41.5%

Barrel events: 60%

3.22 GeV @88.2GeV 59%
9.95 GeV 41%

Barrel: -45 deg < Theta (uds quark) < 45 deg

The broad tails of the distribution need to be reduced significantly, otherwise, physics performance is not going to be good…

Timeline

PFA development

Test Beam

Done → Working PFA

In progress
Should be done
by a few months' effort → Z-pole, 2 jets

Hard to tell at this moment
(doesn't seem to be a trivial job!)

Establish $\sigma_{PFA} \sim E_{jet}$ ⟷ Confirm/tune MC simulation

>= 500 GeV, multiple jets

Optimize detector design

Summery

- Particle Flow Algorithms are being developed at Argonne
  - Two ‘complete’ PFA’s are available to play with
  - Performance of both PFA’s looks promising, and will be improved
- Current performance of this PFA at Z-pole looks promising but not good enough yet
  - Already identified several components that need to be worked on
  - Will continue to work on it in the next few month
  - Need to achieve ~30%/sqrtE with small tail component at this energy
- Need to study PFA performance over the entire ILC interested jet energy range and with more complicate final states
  - Need to show that ILC jet energy resolution goal can be achieved
  - Get PFA ready for optimizing detector design
- Test beam data need to come in time!