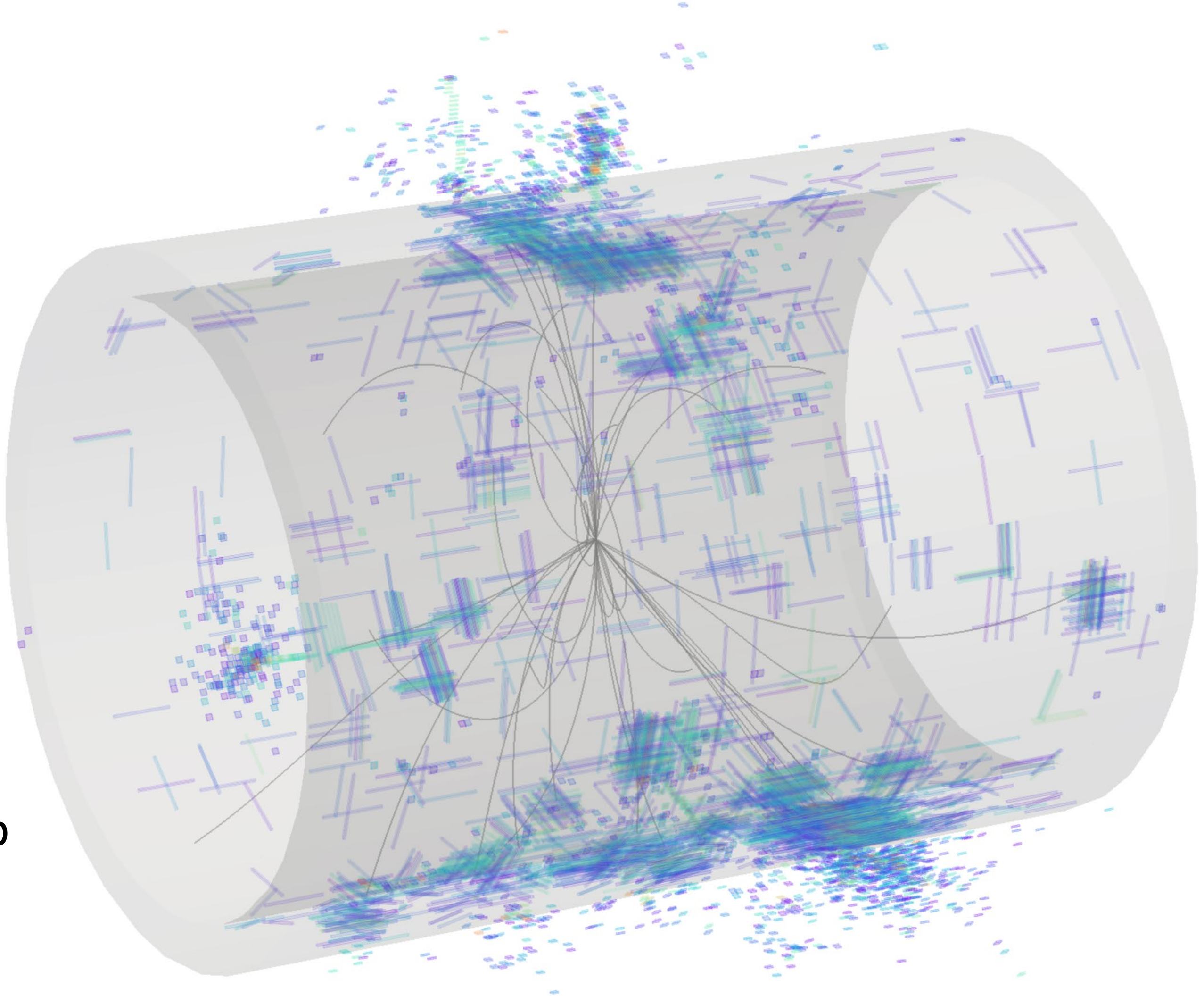
Progress of CyberPFA



Yang Zhang

on behalf of the CEPC ECAL software working group

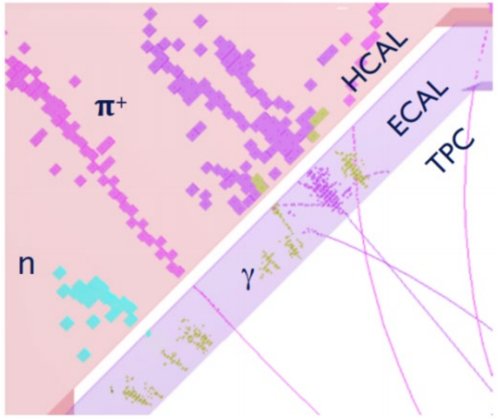
IHEP, CAS

11 Dec. 2024

中国科学院高能物理研究所
Institute of High Energy Physics, Chinese Academy of Sciences

Introduction

- **CEPC: $H(W/Z)$ factory**
  - Precise measurement of Higgs, EW…; search for BSM physics
  - $H/W/Z$ primarily decay into hadrons (jets)
  - Boson Mass Resolution (BMR)<4% is required
    to achieve ~2σ separation of W and Z in their hadronic decay
- **Particle flow approach (PFA)**
  - reconstruct each individual particle with optimal subdetector



Hardware

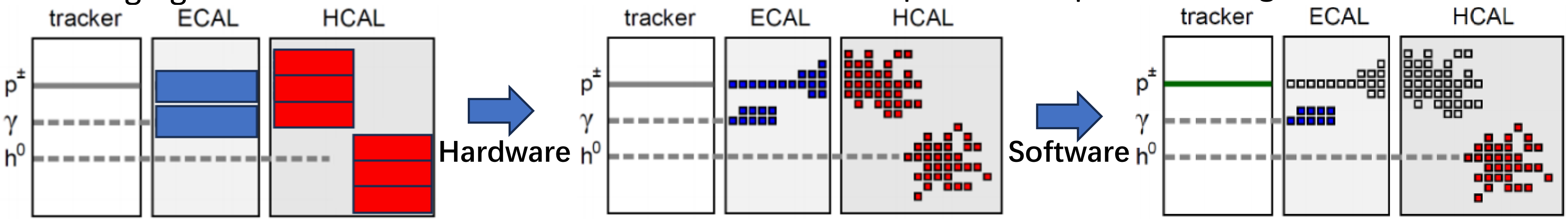
separate clusters from different particles.

- Imaging calorimeter

Software

Assign energy deposits to each particle.

- Sophisticated pattern recognition



- **Confusion term is the limiting factor**
  - Arises from imperfectly associating all energy deposits with the correct particles

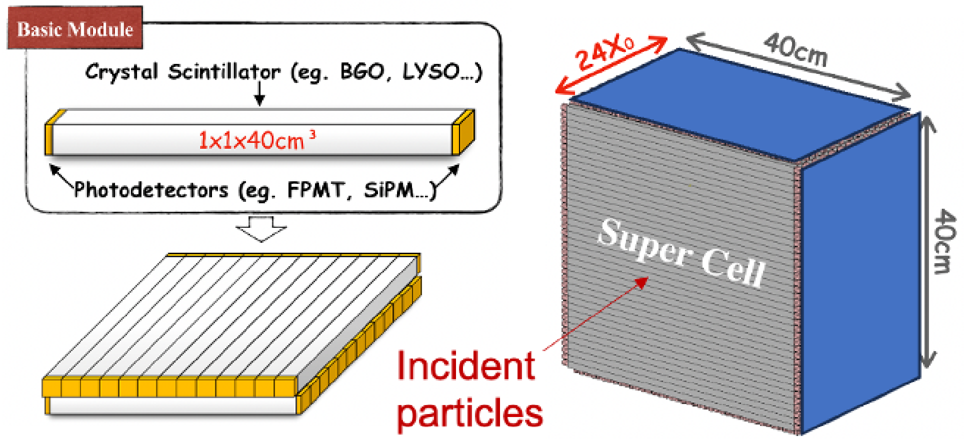
$$\sigma_{jet} = \sigma_{tracker} \oplus \sigma_{ECAL} \oplus \sigma_{HCAL} \oplus \boldsymbol{\sigma_{confusion}}$$

Crystal ECAL design

- **Homogeneous crystal ECAL**
  - Optimal EM resolution: $\sigma_E/E < 3\%/\sqrt{E}$ (better than sampling ECAL: $\sigma_E/E \sim 15\%/\sqrt{S}$)
  - Larger R_M, smaller λ_I/X_0 compare to sampling ECAL (Severer shower overlap between different particles)

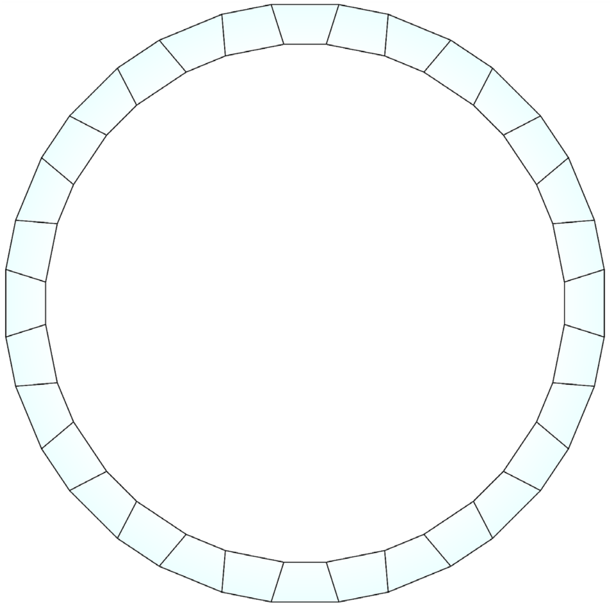
	CsI	$PbWO_4$	LYSO	BGO	W
Light Yield (ph/MeV)	58000	130	30000	7400	-
R_M (cm)	3.57	2.00	2.07	2.23	0.93
X_0 (cm)	1.86	0.89	1.14	1.12	0.35
λ_I (cm)	39.3	20.7	20.9	22.7	9.6
λ_I/X_0	21.1	23.3	18.3	20.3	27.4

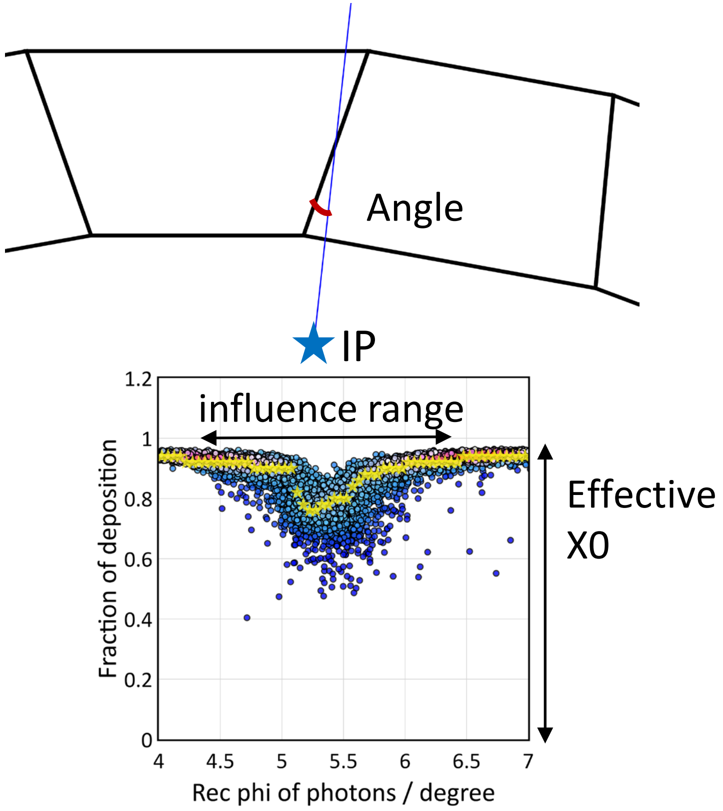
- **Transverse bar design**
  - Compatible for PFA
  - Reduce number of readout channel compare to highly-granular ECAL
  - Minimal longitudinal dead material
  - Ambiguity in reconstruction

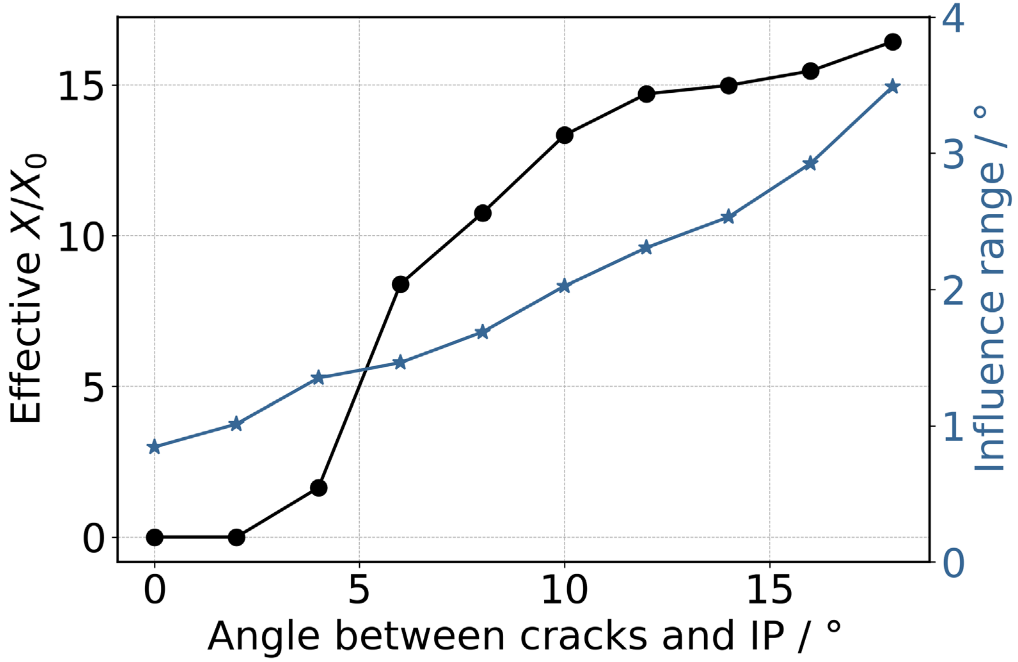


Crystal ECAL barrel design

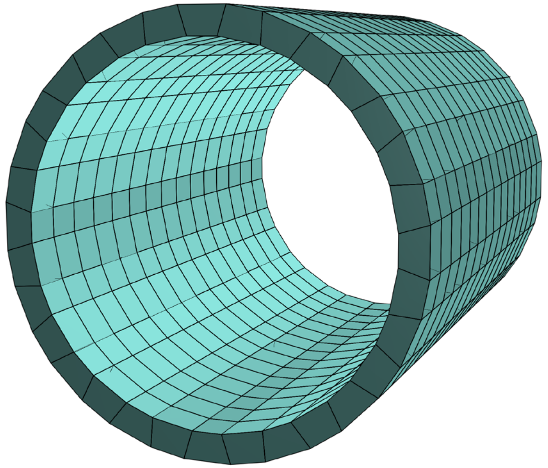
- 32-sided polygon: Interleaved arrangement of regular and inverted trapezoids.
- Crack lead to energy leakage, cannot avoid by hardware, require energy correction.
- A balanced angle 12° between crack and IP.



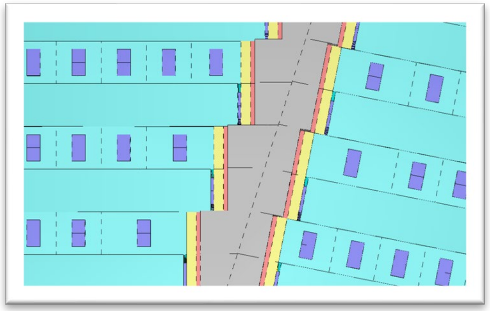


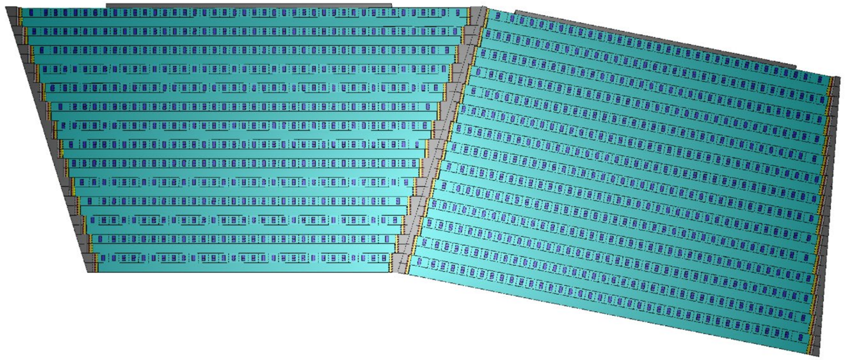


Crystal ECAL barrel implementation

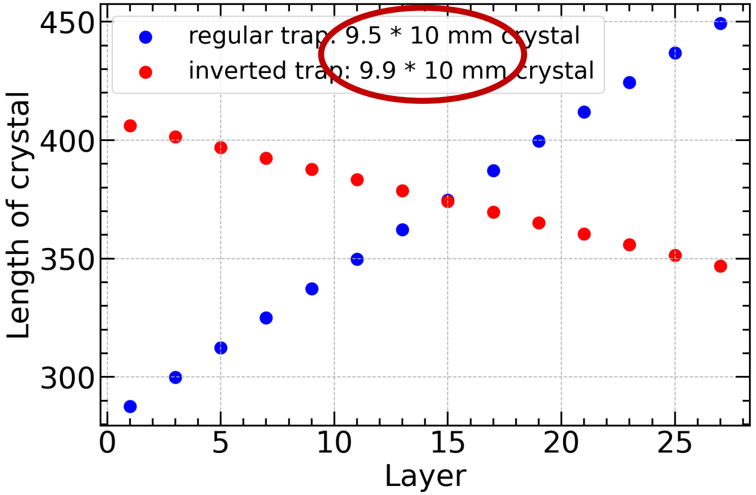


- **Crystal ECAL barrel geometry is implemented in CEPCSW with DD4HEP**
  - 23.7 & 24.8 X_0 for 2 types of module, 28 layers, 30 types of bars
  - Realistic: wrapping, electronic boards, cooling and mechanical supporting.





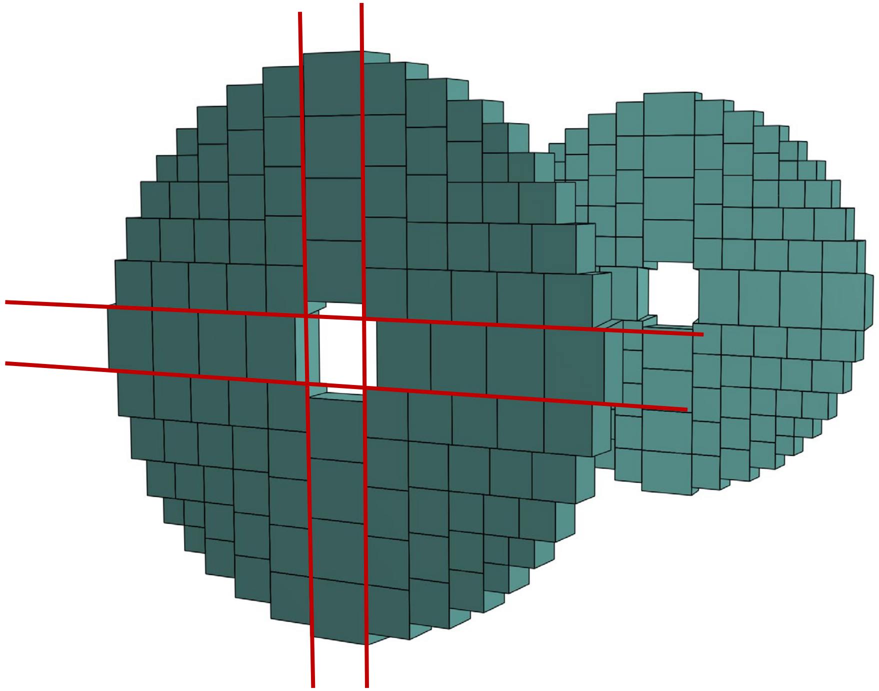
Mechanics is continuously optimized.
Dominant dead material included.

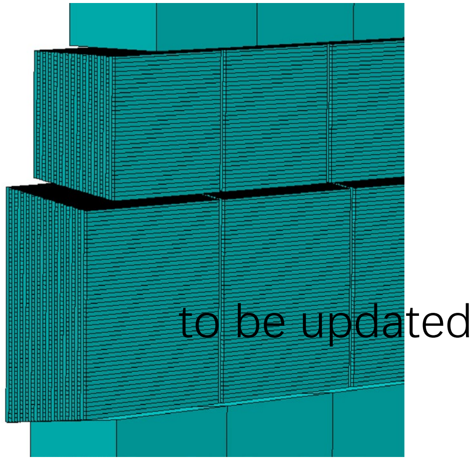


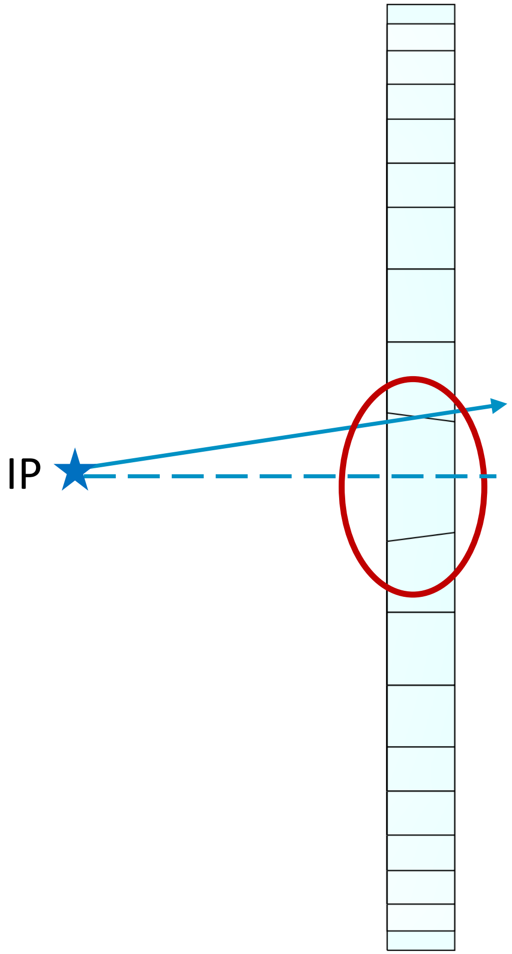
Geometry	Size / mm	Material
Supporting[1]	2.5	carbon fiber
Cooling[2]	1	copper
Electronics front end[3]	1.2+1	PCB+ASIC
Electronic back board[4]	10	PCB
Electro-optical device[5]	3*3*0.8	SiPM
Wrapping[6]	0.1	ESR
Crystal[6]	~10*10*400	BGO

Crystal ECAL endcap implementation

- **Crystal ECAL endcap geometry is implemented in CEPCSW with DD4HEP**
  - >24 X0, 6 types of module, 28 layers, 61 types of bars
  - avoid cracks from pointing into IP
  - ready for development of reconstruction algorithm
  - ➢ further material description will be implemented once mechanical design decided

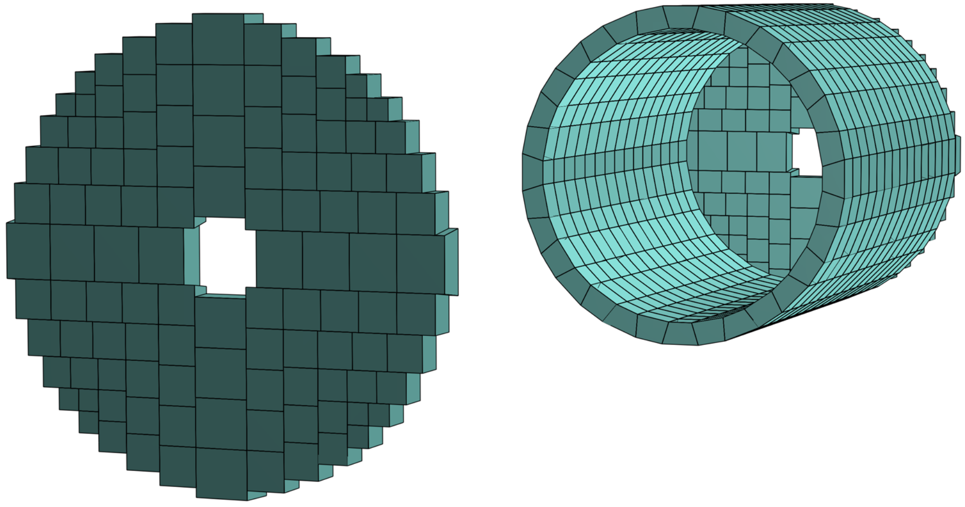


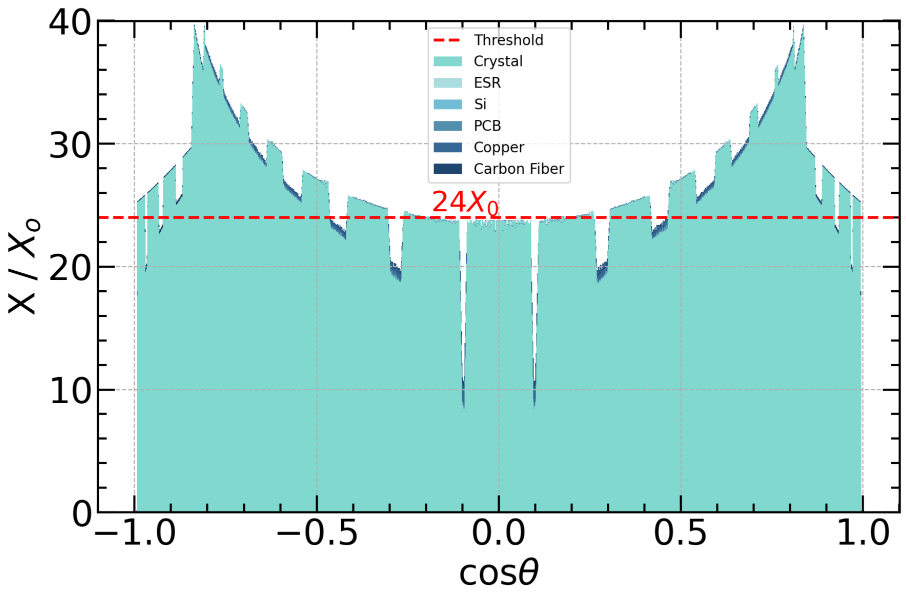


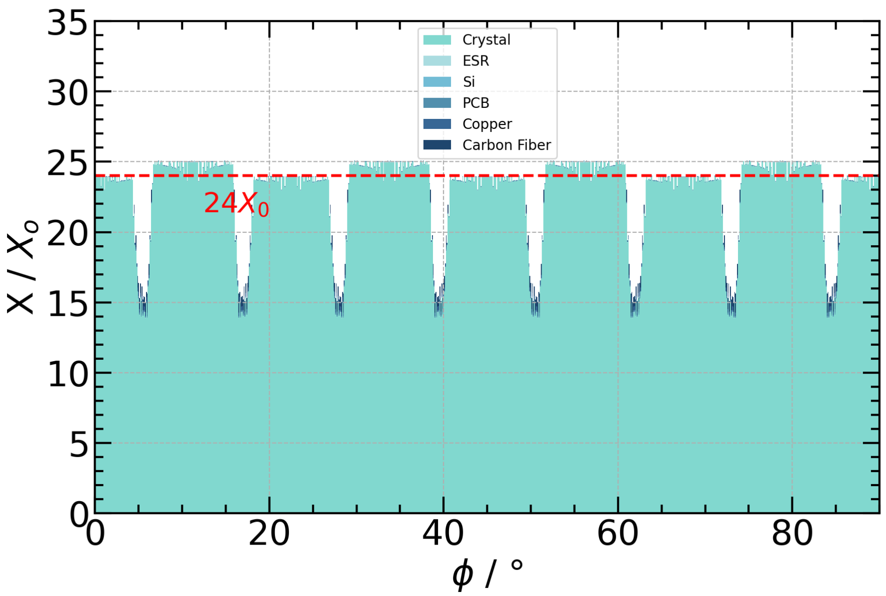


Material Budget of ECAL

Meets the requirement of $> 24X_0$ except at $\cos\theta \sim 0$





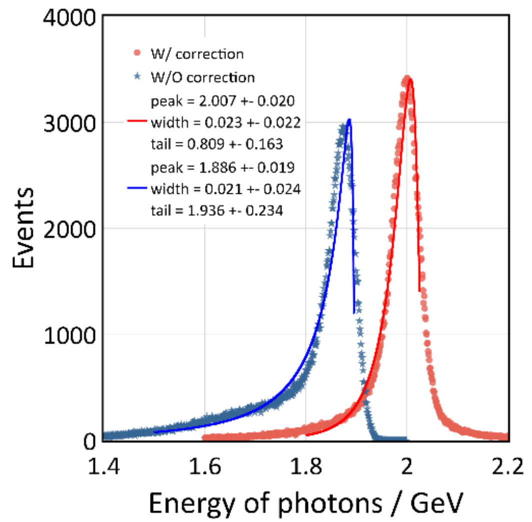


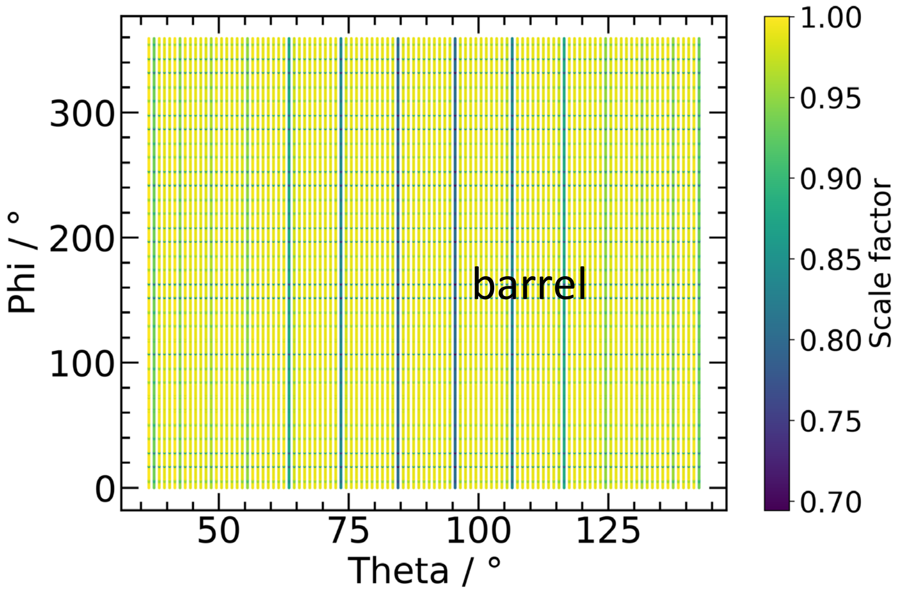
$cos\theta \sim 0$

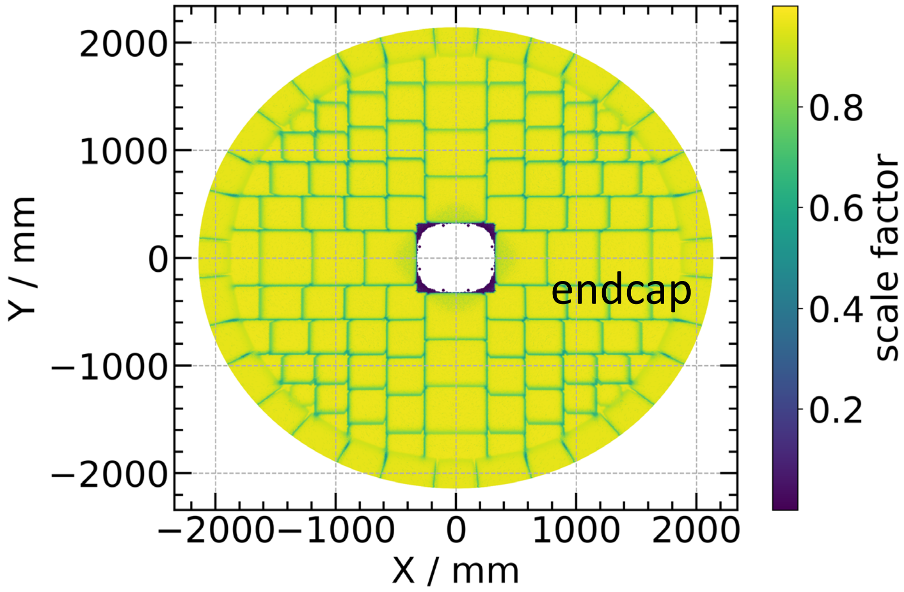
Energy correction of ECAL

$$E_{correction} = E_{depostion} / \frac{E'_{deposition}}{E'_{truth}}$$

Correction factor $\frac{E'_{deposition}}{E'_{truth}}$ is obtained from simulated 2 GeV photon events.

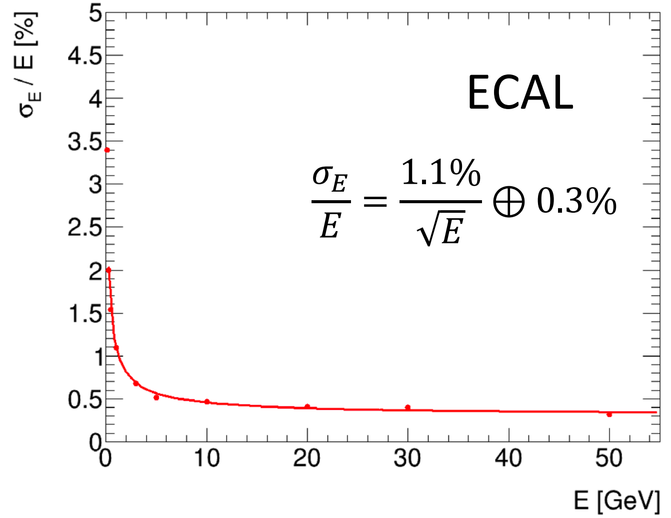


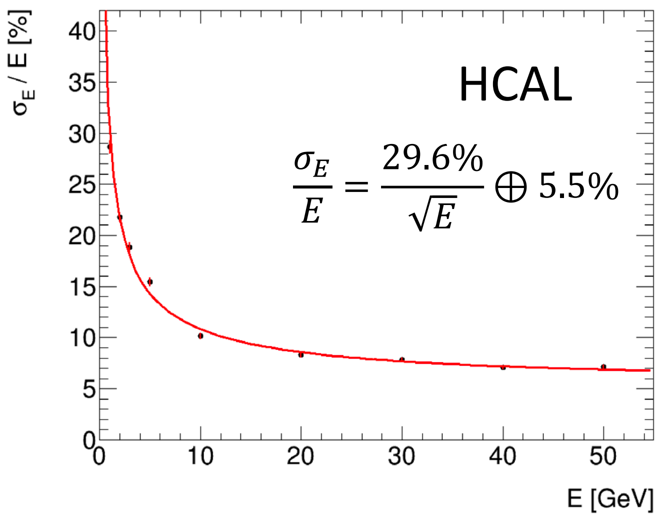




Digitization

- **A preliminary digitization model for ECAL and HCAL:**
  - ECAL:
    - BGO intrinsic light yield 8200 ph/MeV, detected 200 p.e./MIP at single end.
    - Energy threshold 0.05 MIP/ch (~ 10 p.e.), 0.1 MIP / crystal bar.
    - SiPM saturation and correlated correction is applied, noise is not included.
    - Electronics: 13-bit ADC, noise is supposed to be negligible.
  - HCAL:
    - Scintillator glass light yield 80 p.e./ MIP (detected), attenuation is not considered.
    - Energy threshold: 0.1 MIP(~0.7MeV)
    - Effects from SiPM, electronics are not considered.

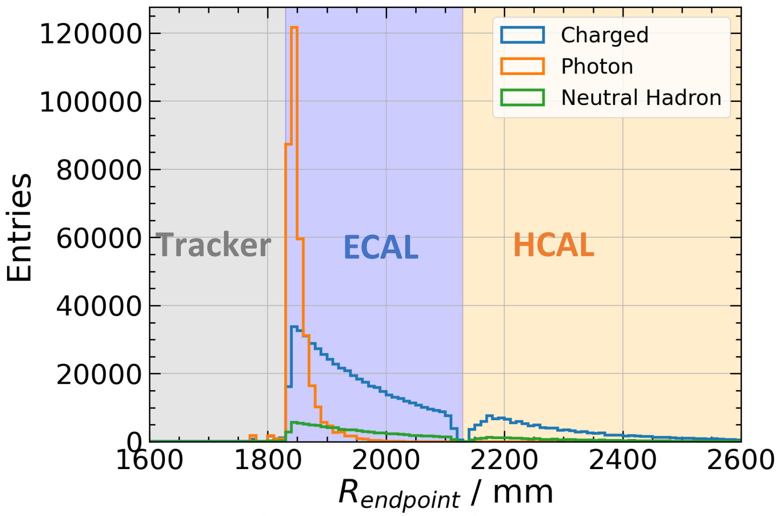




A preliminary digitization model for PFA performance study.
Need to be validated by hardware & electronics colleagues.

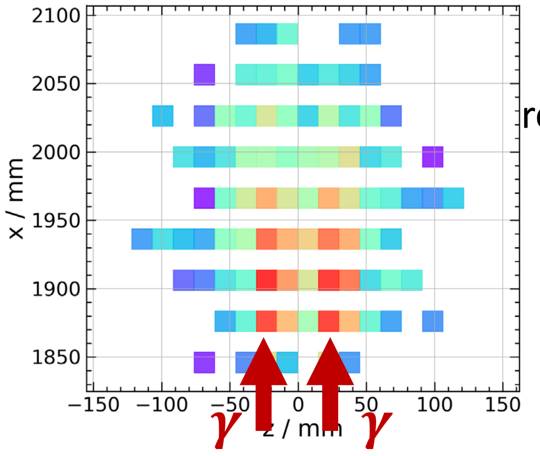
Shower overlap

- Confusion caused by shower overlap from nearby particles
- Pattern recognition
  - Photon shower: compact energy core → Hough transformation
  - Charged particle: track matching
- Energy splitting using lateral profile of EM shower



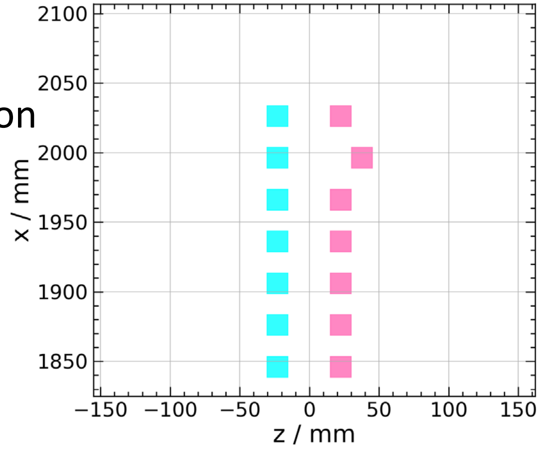
Radius of end-point of particles in jets

~2/3 hadrons shower in ECAL

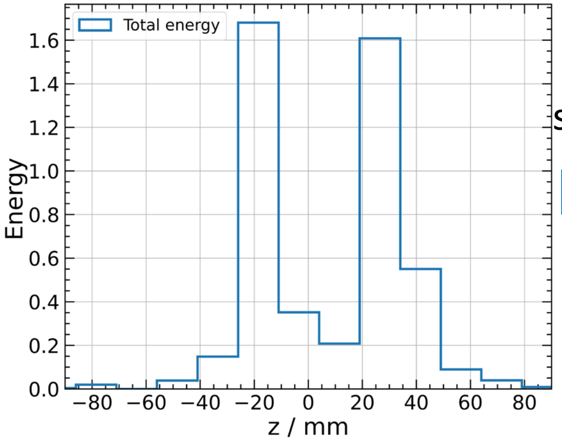


Energy distribution of two overlapped photon

Pattern recognition

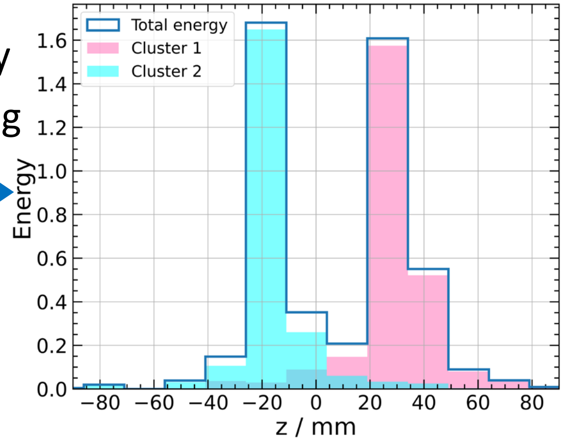


Energy cores of the two photon showers



Lateral energy distribution of two overlapped photon

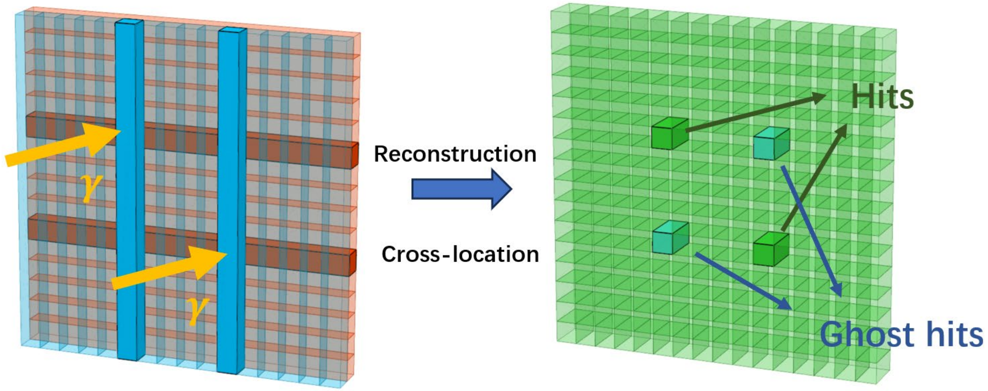
Energy splitting



Lateral energy distribution after energy splitting

Ambiguity problem

- Multiple particles incident in one module arises the ambiguity problem
- Severe in jet events (~5 particles in one module per jet event)

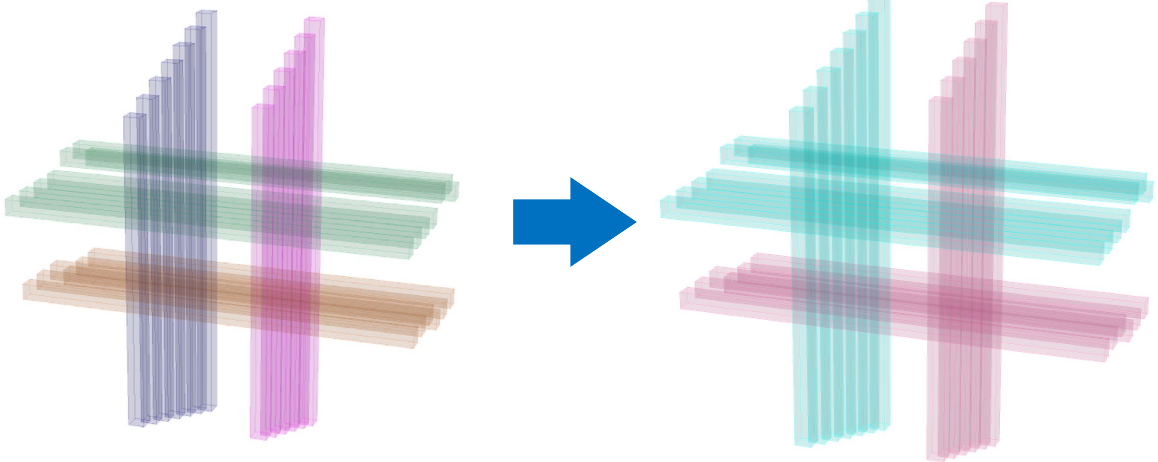


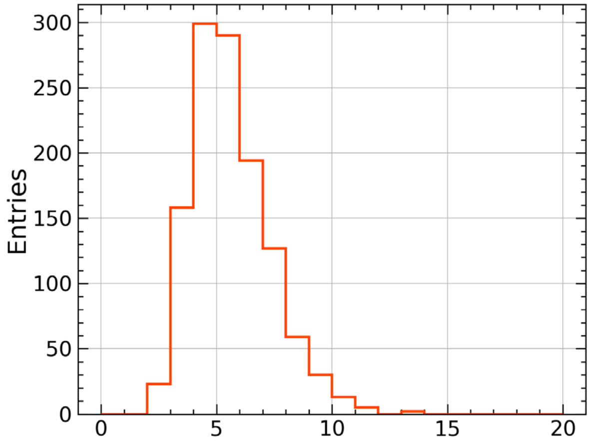
- Matching horizontal and vertical clusters by utilizing energy & time information

$$\chi_E^2 = \frac{\left(E_X - E_Y\right)^2}{\sigma_E^2}$$

$$\chi_T^2 = \frac{(Z_T - Z_Y)^2}{\sigma_S^2 + \sigma_{Z(t)}^2}$$

$$\chi_{point}^2 = \chi_E^2 + \frac{1}{2}\left(\chi_{Tx}^2 + \chi_{Ty}^2\right)$$

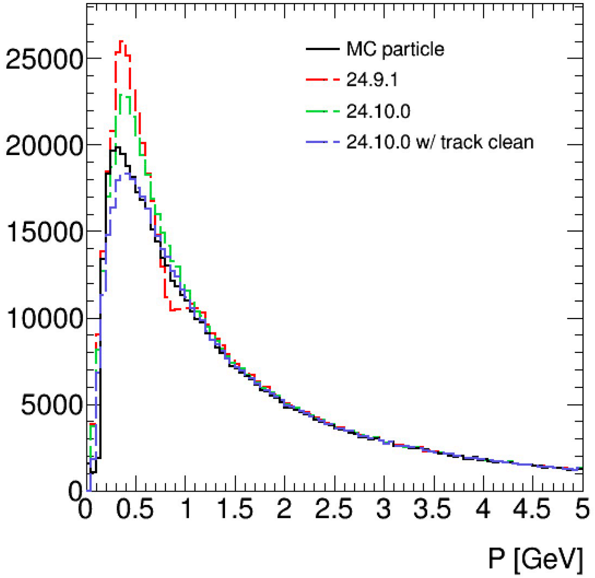


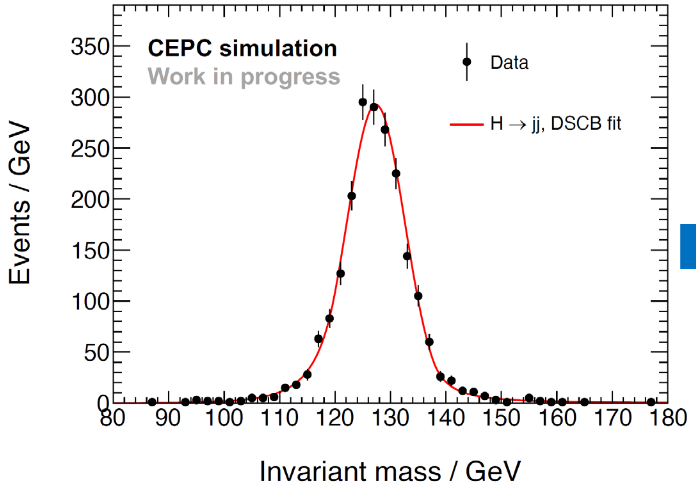


Number of particles in hottest module for $ee \to ZH \to qqgg$ process

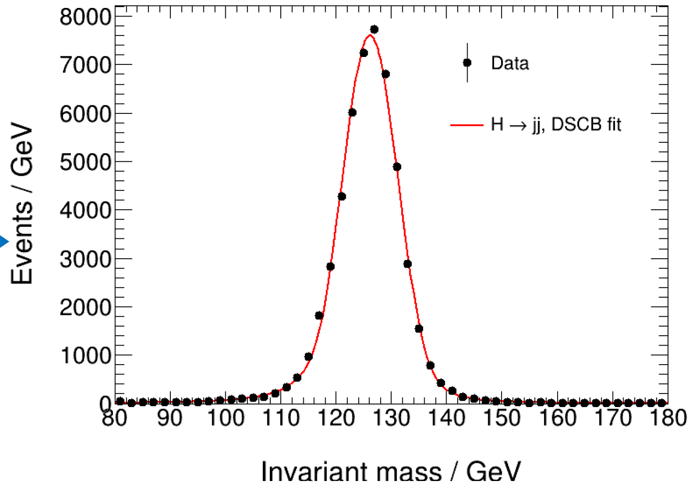
Impact of tracking performance on BMR

- **Optimize tracking performance**
  - REC: fix missing while merging TPC tracks and silicon tracks
  - REC: change sort while combine silicon and TPC
  - REC: improve success rate of merge of silicon and TPC
  - REC: improve efficiency at forward region
  - …





$\mathrm{BMR} = 4.11\%$



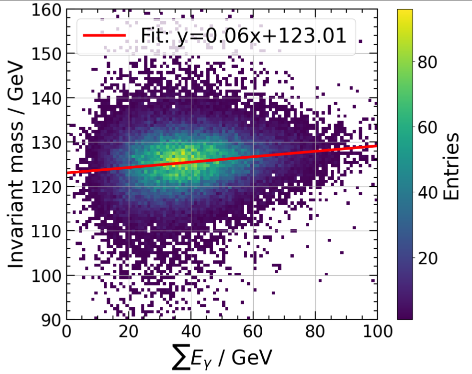
$\mathrm{BMR} = 3.94\%$

Optimization of ECAL calibration

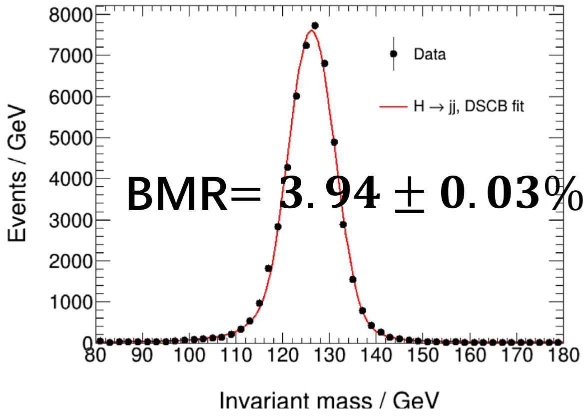
The response of ECAL to electromagnetic(EM) showers and hadronic showers is different, requiring different calibration.

$k_{ECAL} = 1.06$

$k_{HCAL} = 4.00$



m_{jj} correlated with $\sum E_\gamma$

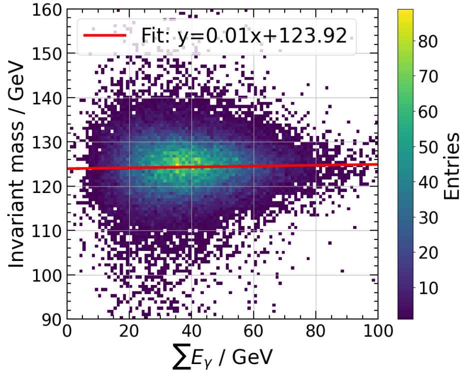


BMR= $\mathbf{3.94 \pm 0.03\%}$

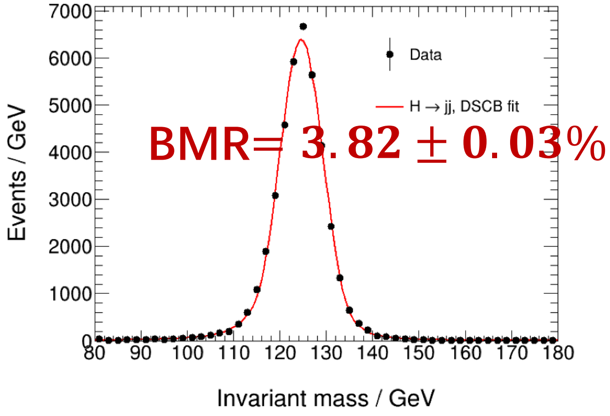
$k_{ECAL}^{hadronic} = \mathbf{1.26}$

$k_{ECAL}^{EM} = \mathbf{1.00}$

$k_{HCAL} = 4.00$



m_{jj} NOT correlated with $\sum E_\gamma$

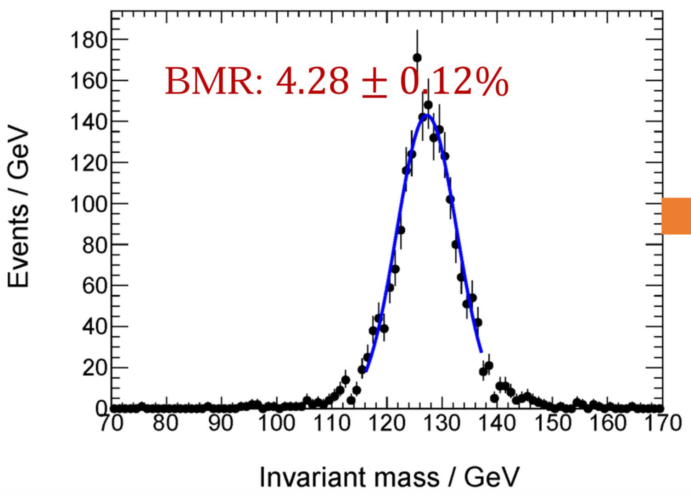


BMR= $\mathbf{3.82 \pm 0.03\%}$

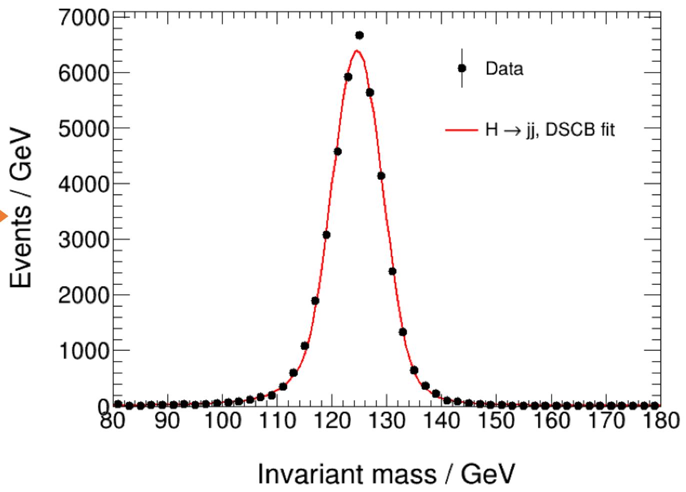
Physics performance

- Full simulation & reconstruction: track + ECAL + GS-HCAL
- Benefit from optimization of track reconstruction and calorimeter calibration
- ~200k events generated, ~ 50k selected for barrel only.

$e^+e^- \to ZH \to \nu\, \nu g g$ in $\sqrt{s} = 240$ GeV



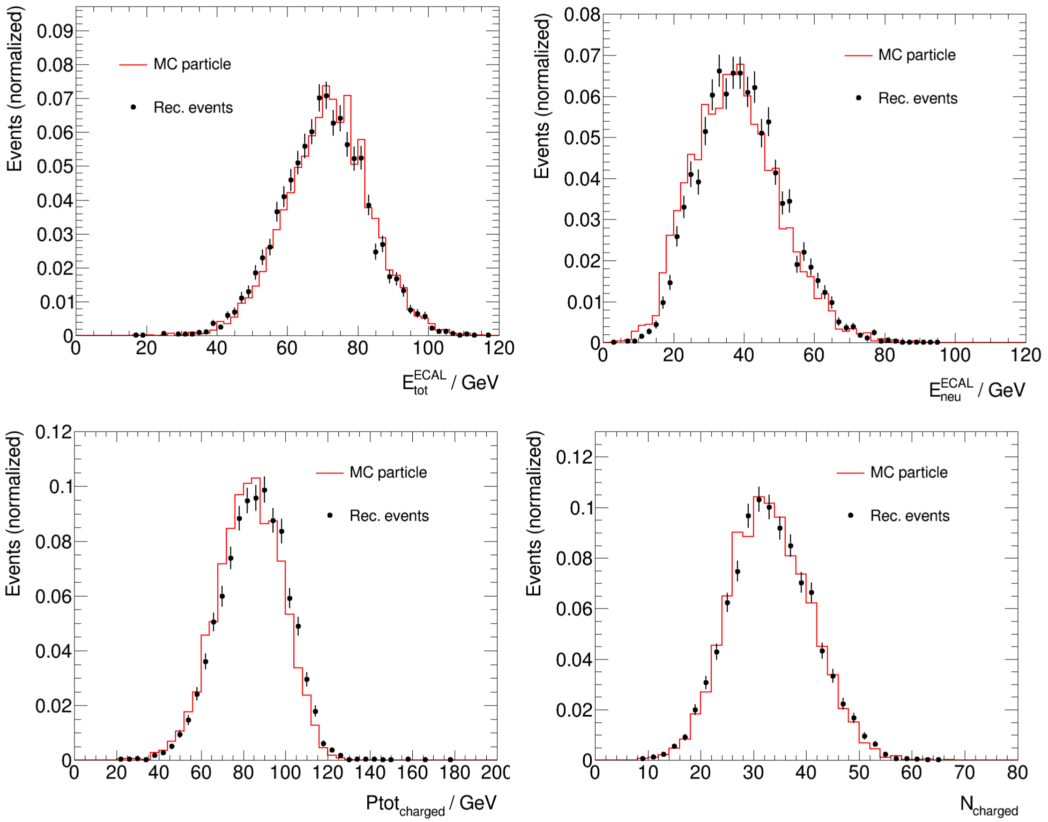
CEPC day in June



$m_{jj} = 124.58 \pm 0.03$GeV

$\sigma_{jj} = 4.76 \pm 0.03$GeV

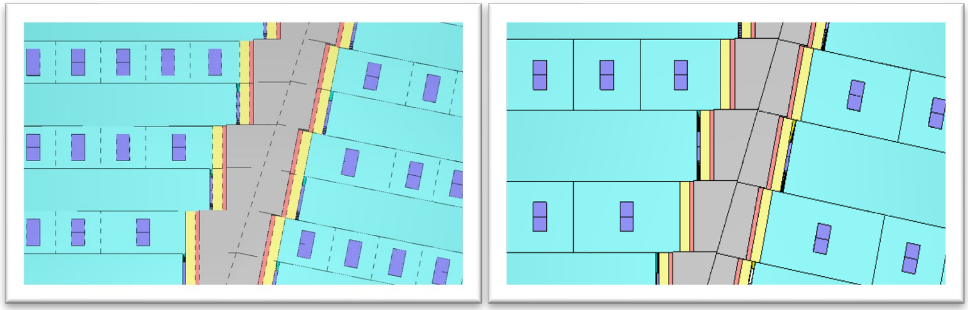
BMR = 3.82 ± 0.03%



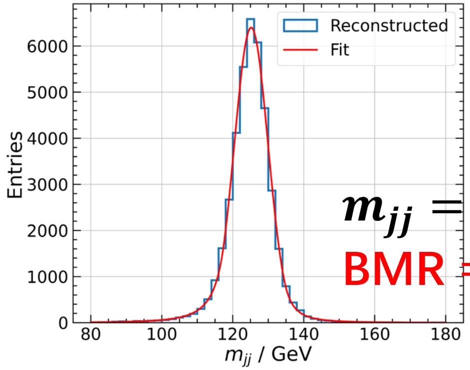
Comparison of granularity

Recommendation of IDRC

- **10mm ➔ 15 mm:**
  - Advantages:
    - Significantly reduce the number of readout channels.
      - Reduce 57.6% (956160 → 405120)
    - Less dead area
      - one step per 2 layers → one step per layer
      - 13.8% → 12.4%
    - Reduce difficulty of production of crystal bars
    - Mechanics(cooling) and electronics benefit from larger granularity
  - Disadvantages:
    - Lager granularity deteriorate particle recognition
- **Both granularity of ECAL meet the requirements of BMR**



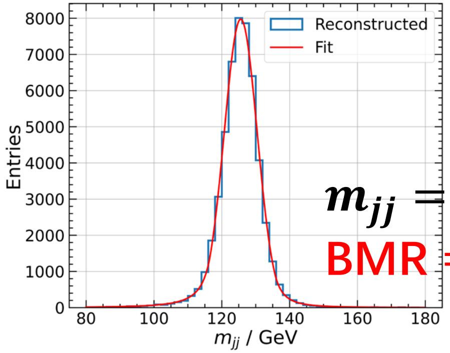
10mm×10mm 15mm×15mm



10 mm ECAL

$m_{jj} = 124.6 \pm 4.8$ GeV

BMR $= 3.82 \pm 0.03\%$

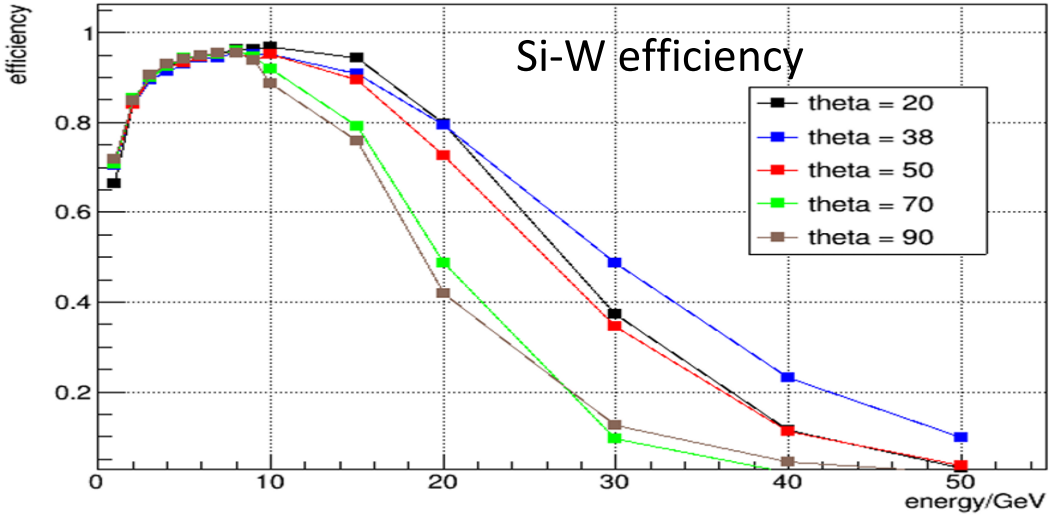


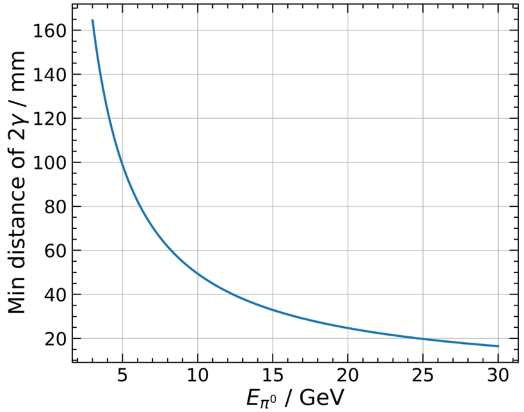
15 mm ECAL

$m_{jj} = 125.0 \pm 4.9$GeV

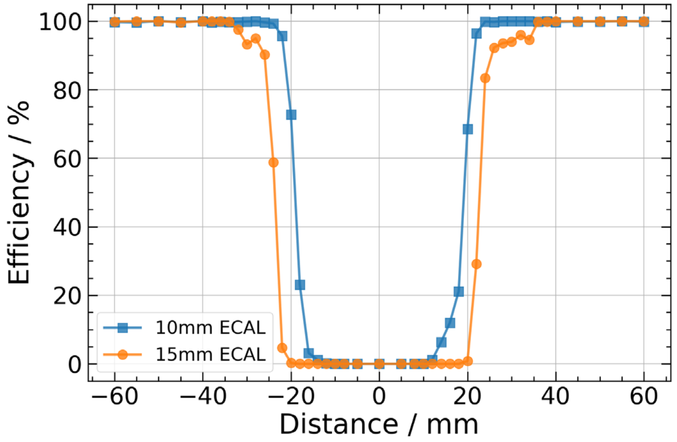
BMR $= 3.89 \pm 0.02\%$

Efficiency of π^0

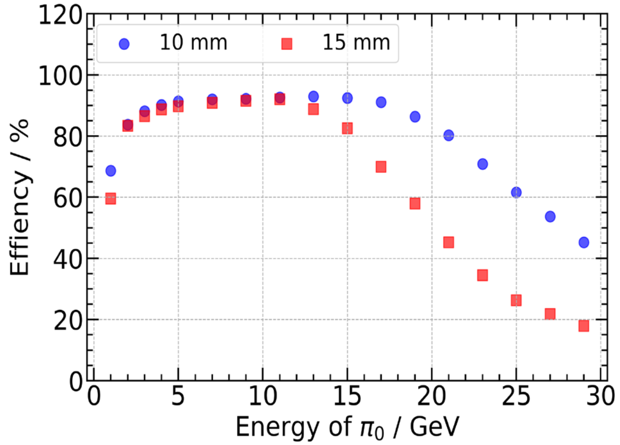




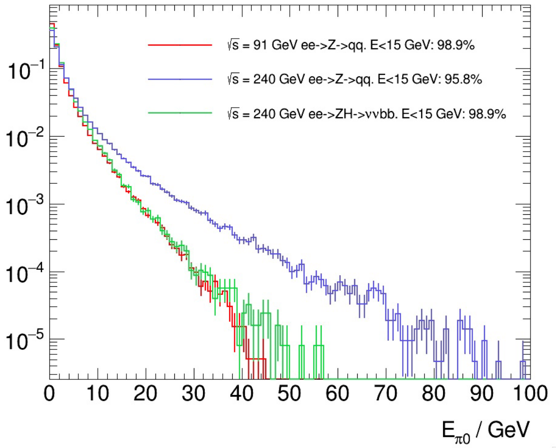
Min distance between the 2γ from π^0

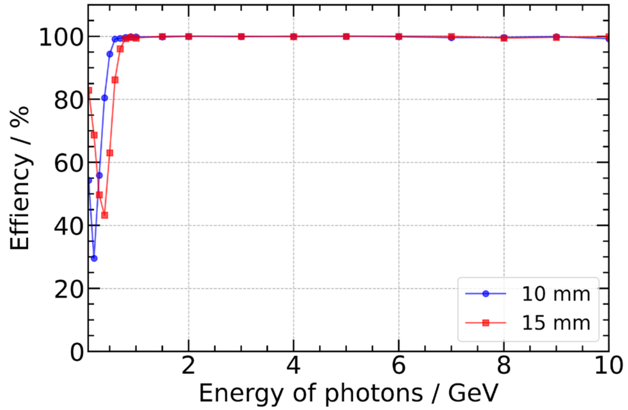


Di-photon separation efficiency



Crystal efficiency

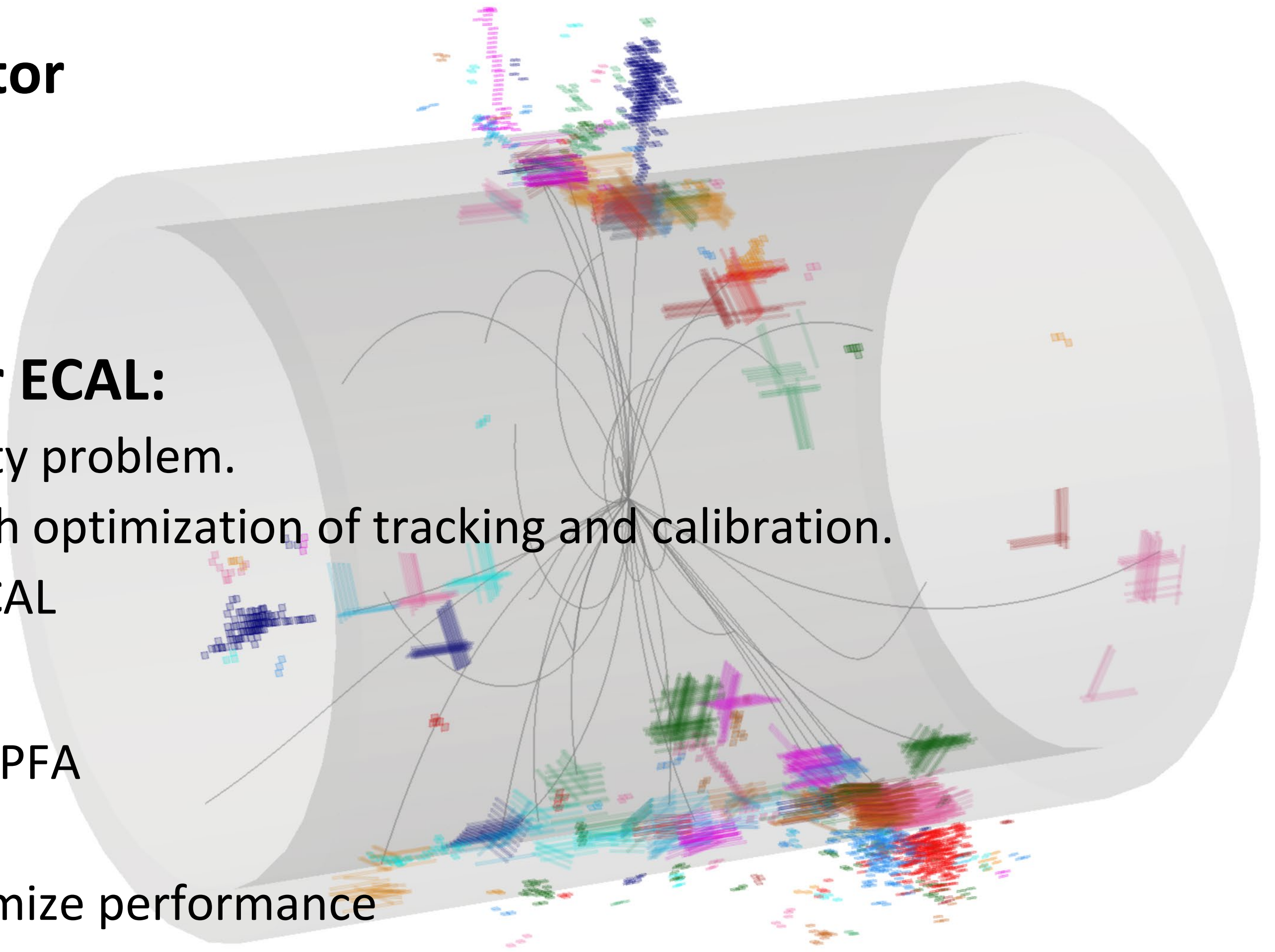




Photon recognition efficiency

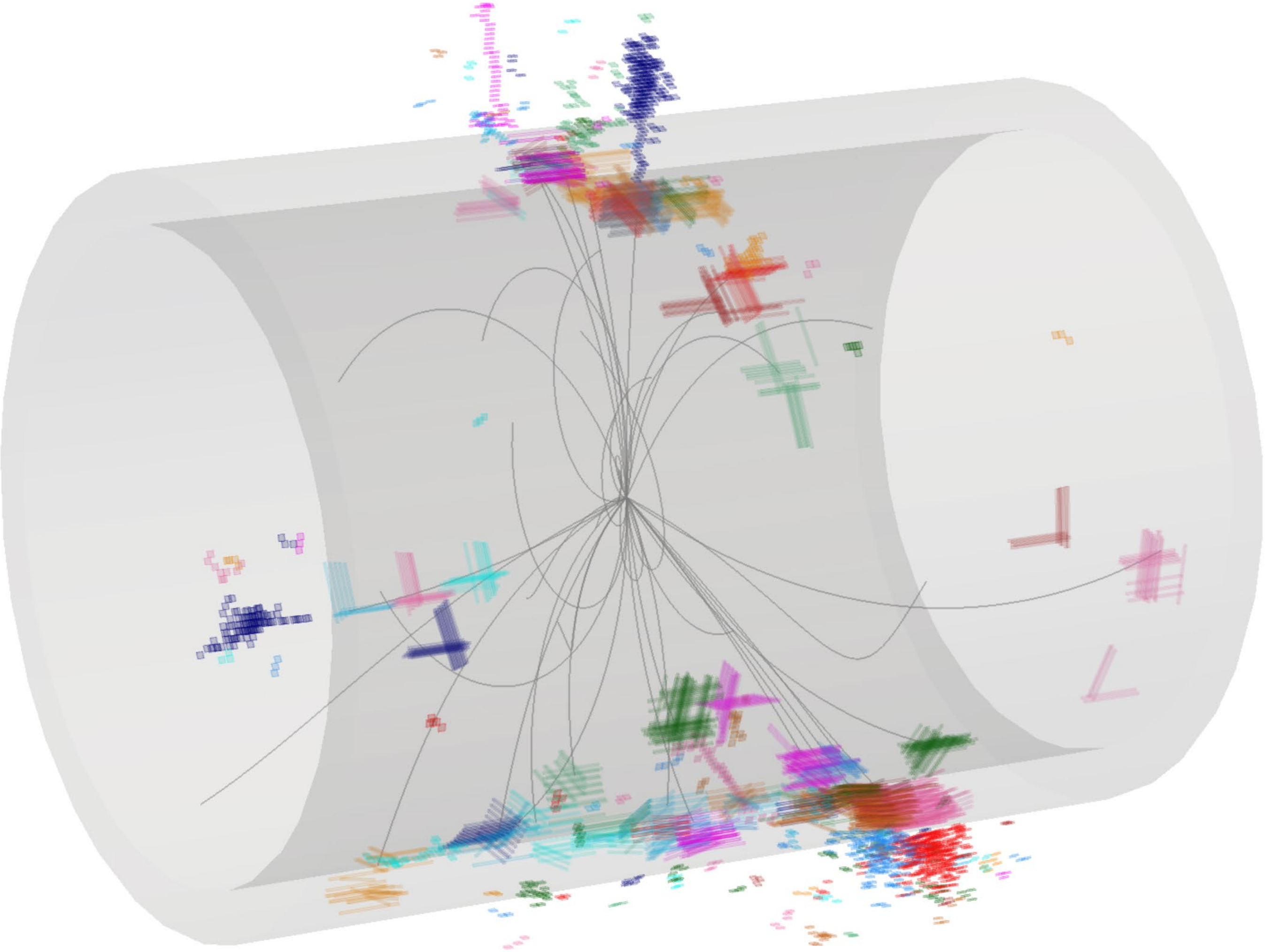
Summary and Plan

- **Crystal bar ECAL design for CEPC detector**
  - Optimal EM resolution
  - Following PFA concept.
  - Barrel and endcap ECAL available
- **CyberPFA developed for the crystal bar ECAL:**
  - Challenges: shower overlapping & ambiguity problem.
  - Performance significantly improved through optimization of tracking and calibration.
  - BMR $< 3.9\%$ for both 10mm and 15mm ECAL
- **Future plan:**
  - Develop the endcap algorithm of the CyberPFA
  - Further optimize the CyberPFA.
  - Introduce machine learning to further optimize performance

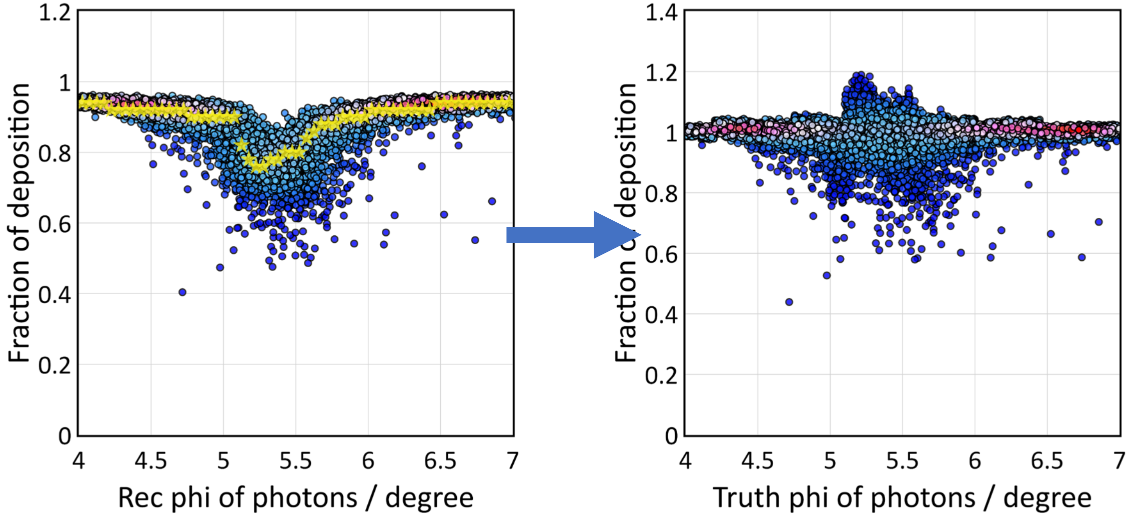


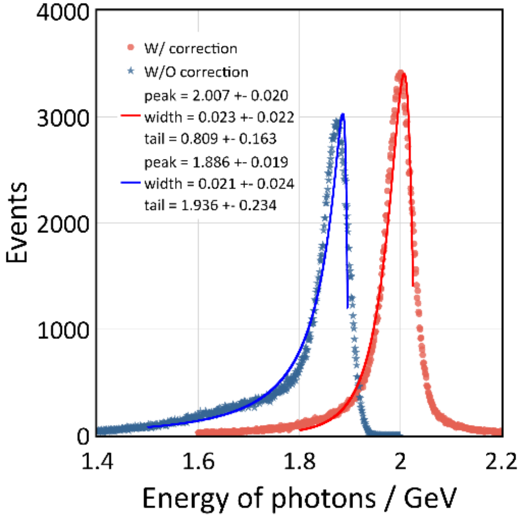
CEPC

Backup



中国科学院高能物理研究所
Institute of High Energy Physics, Chinese Academy of Sciences


Fraction of deposition
Rec phi of photons / degree
Fraction of deposition
Truth phi of photons / degree


W/ correction
W/O correction
peak = 2.007 +- 0.020
width = 0.023 +- 0.022
tail = 0.809 +- 0.163
peak = 1.886 +- 0.019
width = 0.021 +- 0.024
tail = 1.936 +- 0.234
Events
Energy of photons / GeV

CyberPFA process

Topological clustering

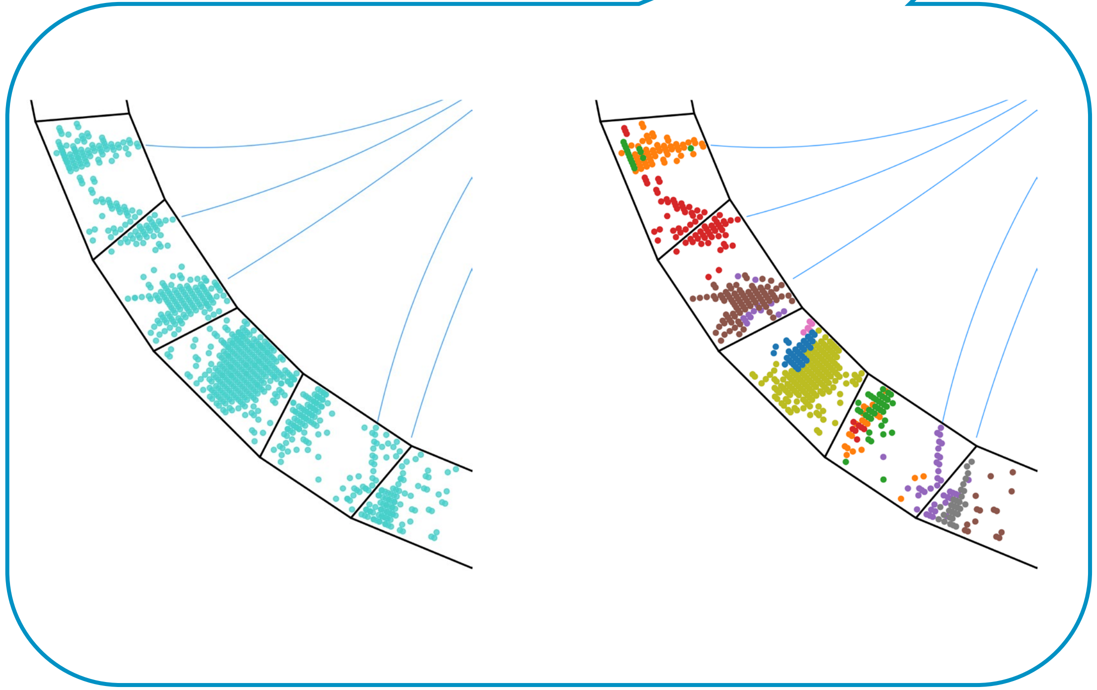
Local max finding

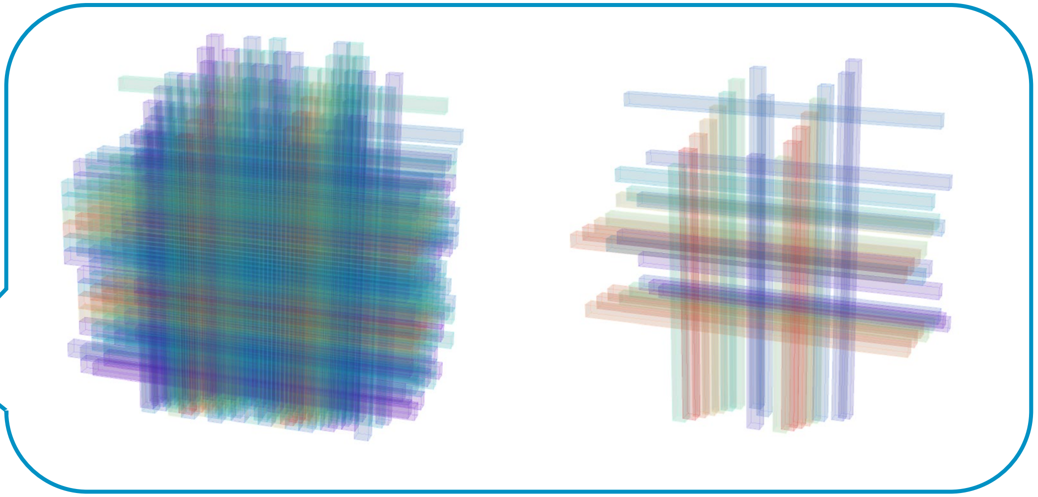
Pattern recognition

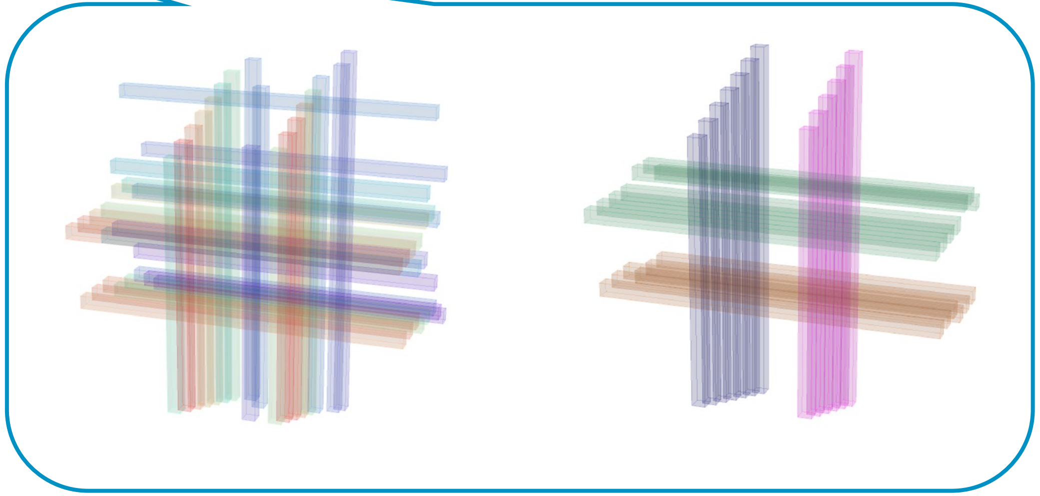
Energy splitting

Matching

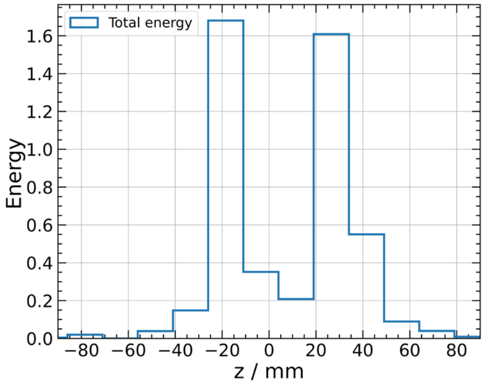
Re-clustering

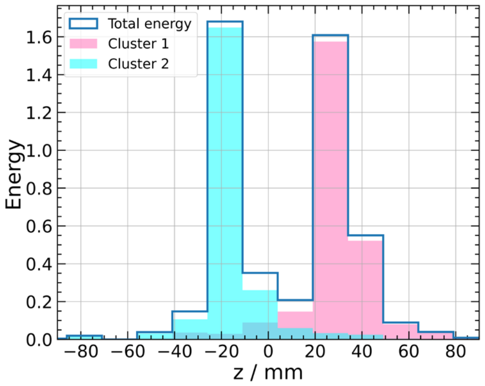




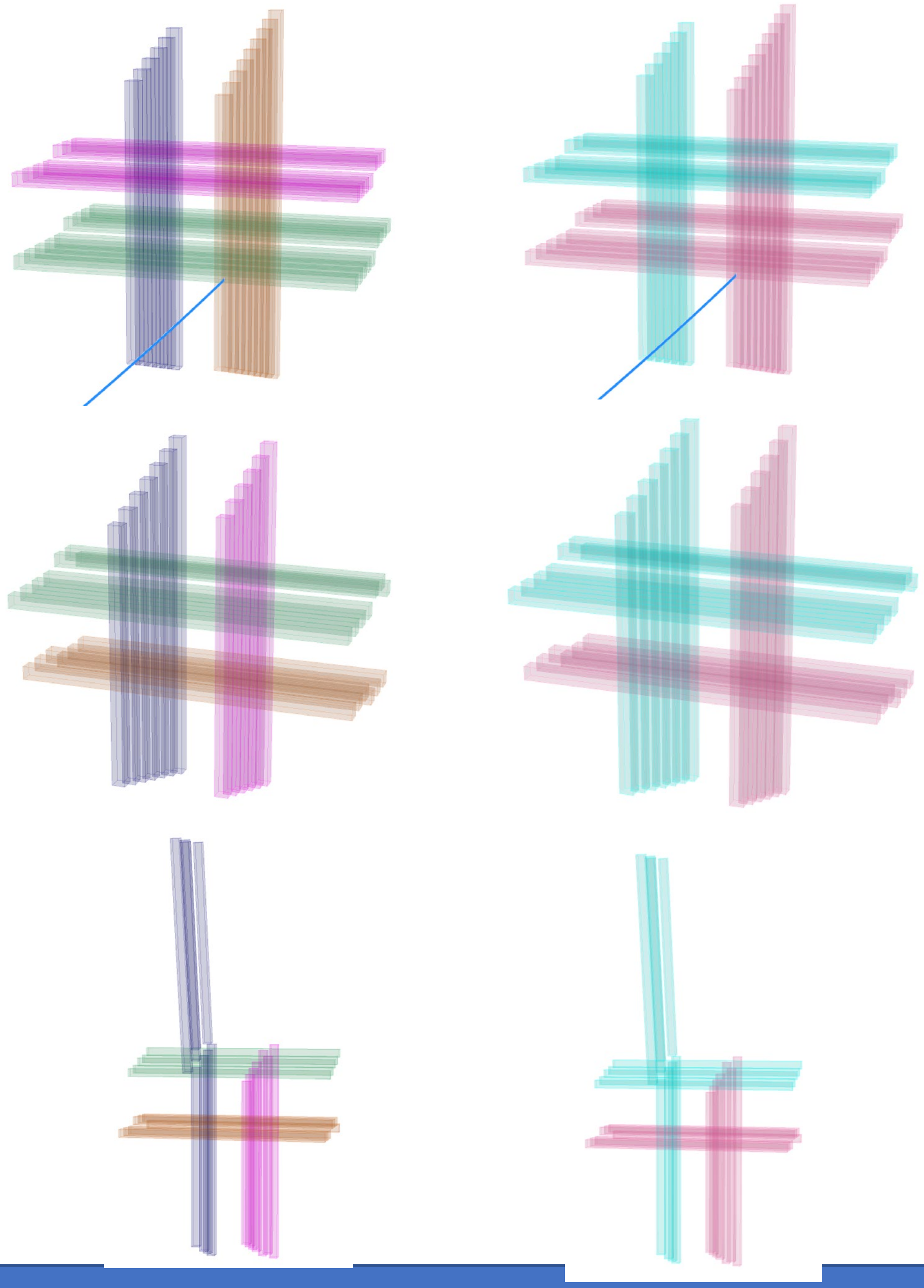


CyberPFA process



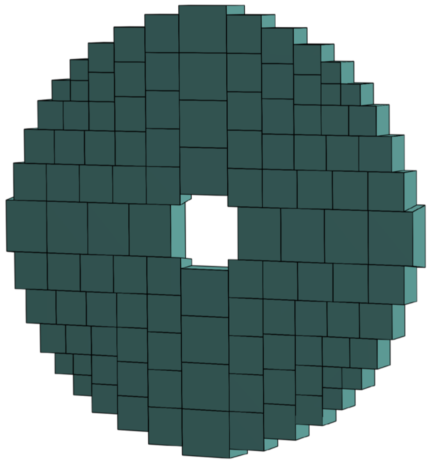


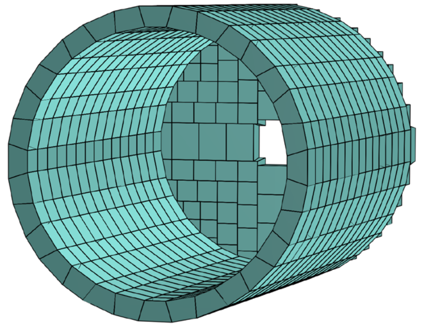
Topological clustering
↓
Local max finding
↓
Pattern recognition
↓
Energy splitting
↓
Matching
↓
Re-clustering

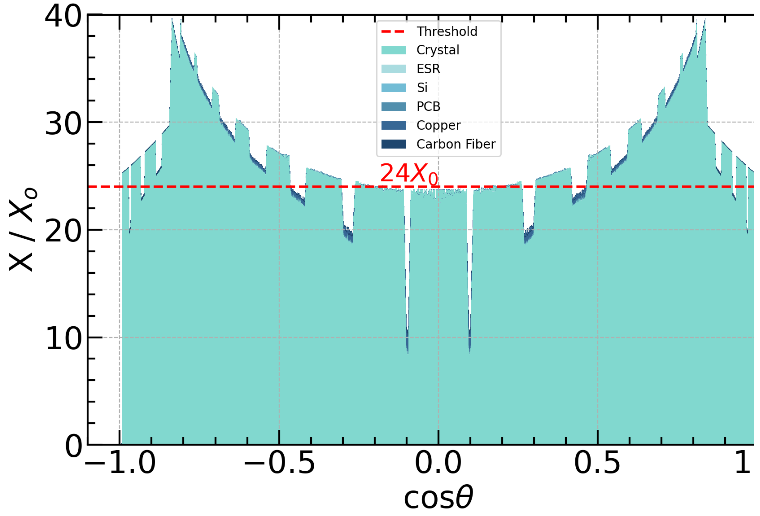


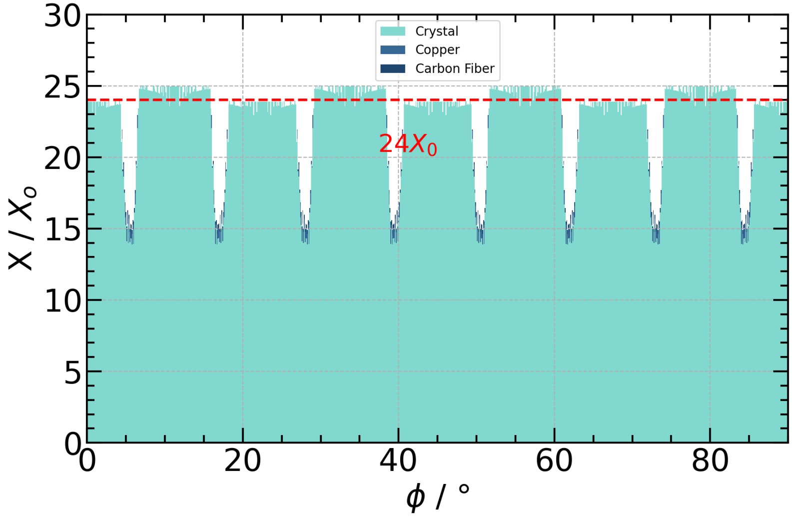
Material Budget of ECAL and tracker

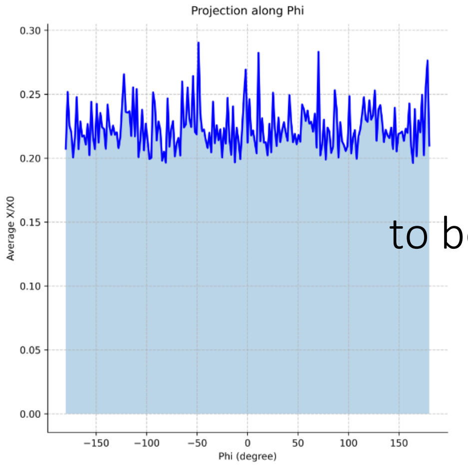
- **Material budget of ECAL:**
  - Essentially meets the requirement of $> 24X_0$
- **Material budget of tracker:**
  - introduce about 5% and 30% photon conversion in ECAL barrel and endcap respectively.



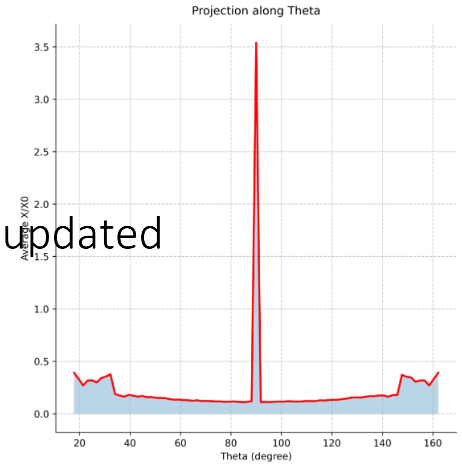






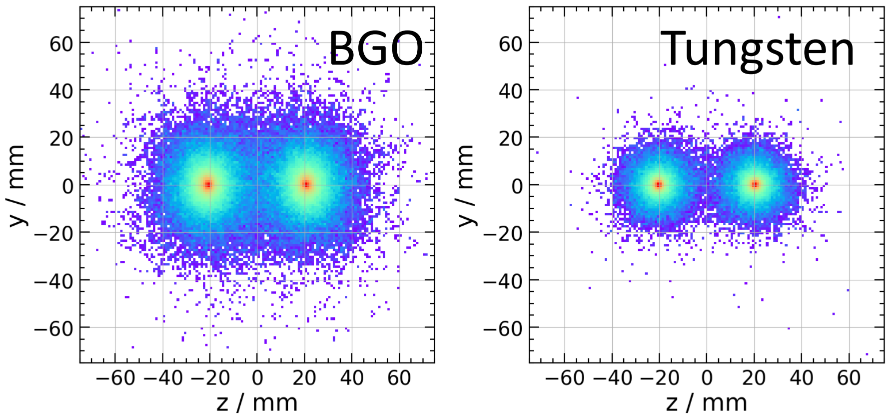


to be updated

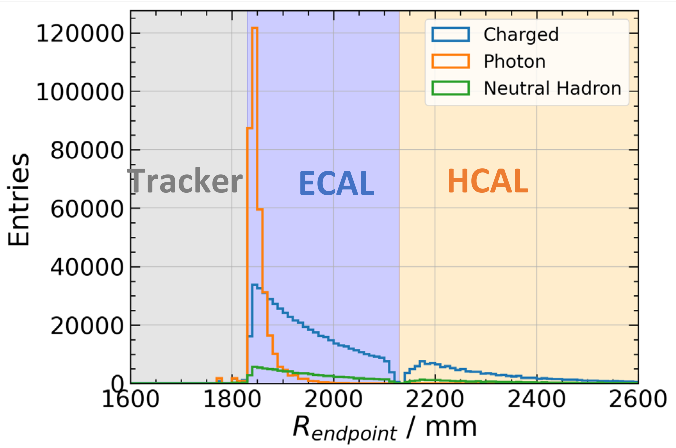


Key issue of software

- **Shower overlap**



Shower profile of two 5GeV photon in BGO & W



Radius of end-point of particles in jets

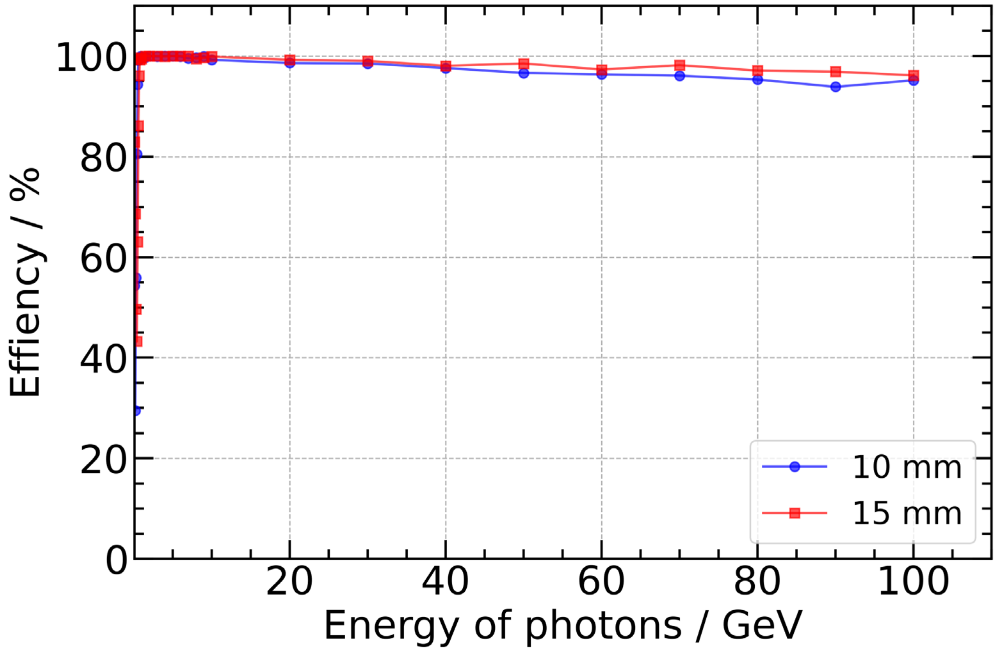
~2/3 hadrons shower in ECAL

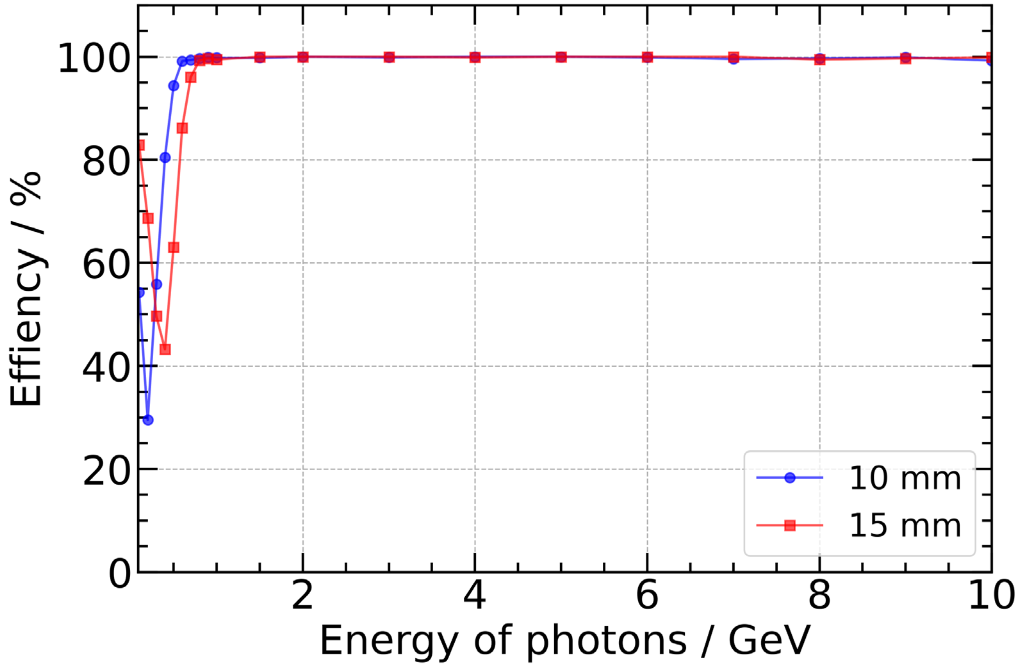
- **Ambiguity problem**

Diagrams illustrating Ambiguity:
1 ptc,
2 ptc,
jet hottest tower n particles

Efficiency of γ Recognition

- $\boldsymbol{\gamma}$
  - $E_{\gamma}: 0.1 \sim 100GeV$
  - $\theta = 91°, \phi = 0°$
  - $N_{cluster} \geq 1$
  - $|E_{cluster} - E_{truth}| < 5\sigma$

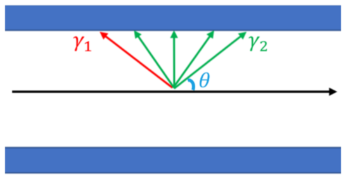


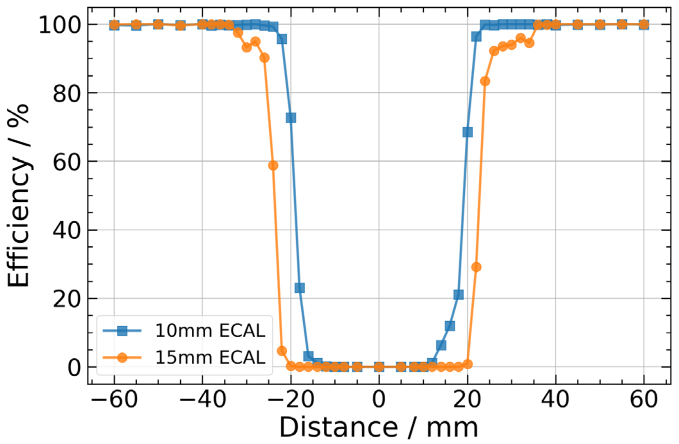


Veto particles with interactions in front of ECAL

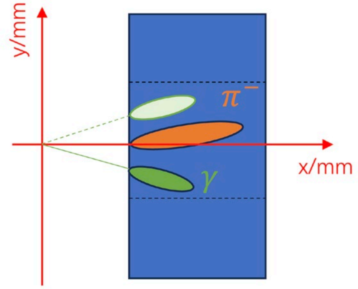
Separation Capability

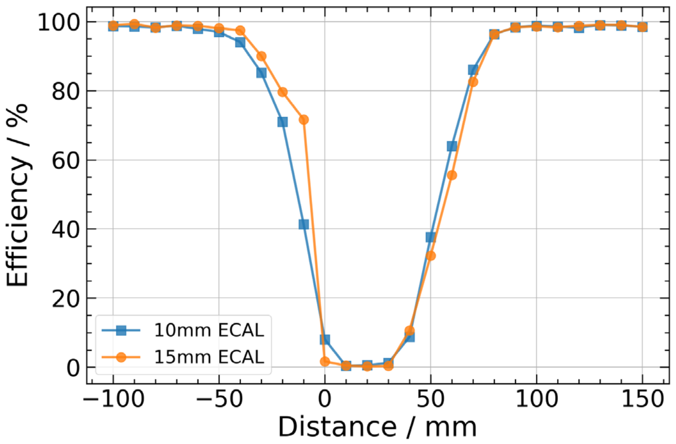
- $\gamma\gamma$
  - $E_{\gamma 1} = E_{\gamma 2} = 5GeV$
  - Success separation:
    - ≥2 PFO,
    - $|E_{\gamma} - E_{PFO}| < \frac{1}{3} E_{\gamma}$
    - $|\theta_{\gamma} - \theta_{PFO}| < 0.3$ for 10mm ECAL, for 15mm ECAL





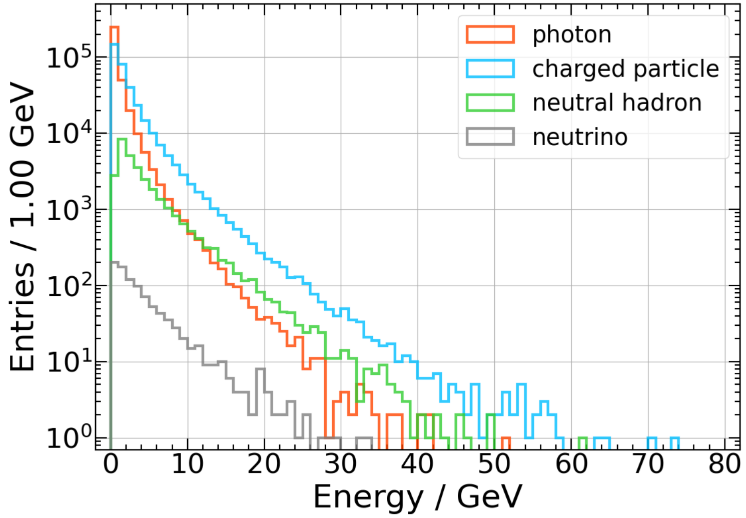
- $\gamma\pi$
  - $E_{\gamma} = E_{\pi^-} = 5GeV$
  - Success separation:
    - 1 charged PFO, ≥1 neutral PFO
    - $|E_{\gamma} - E_{neutral\ PFO}| < \frac{1}{3} E_{\gamma}$
    - <0.45 $|y_{gamma} - y_{PFO}| < 30mm$

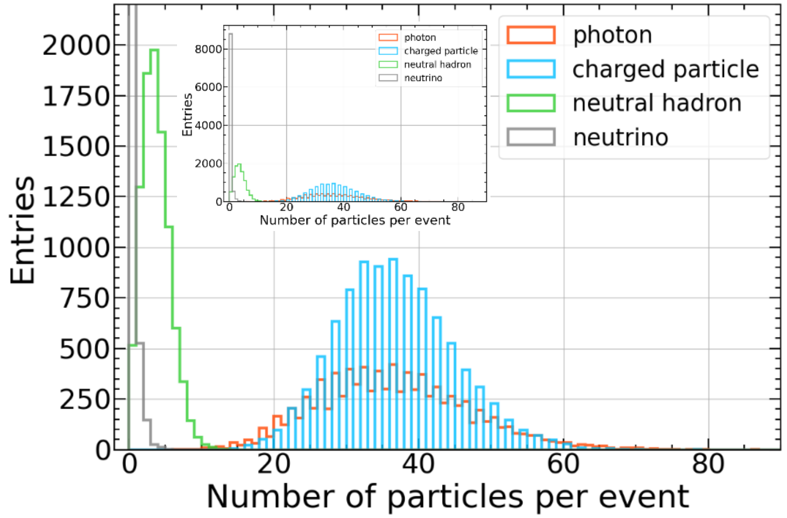


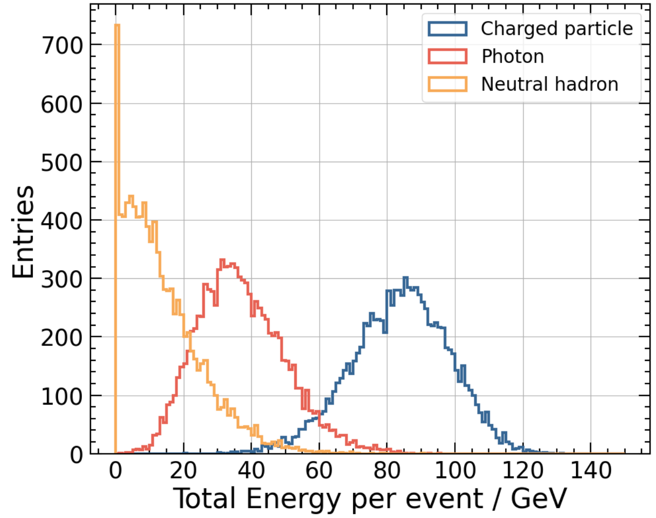


Veto particles with interactions in front of ECAL

$H \rightarrow gg$ MC information



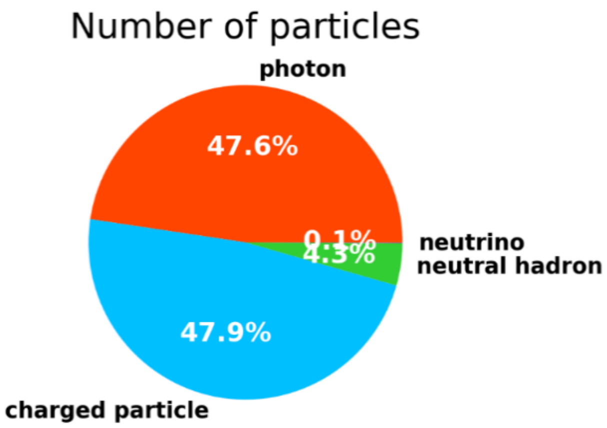


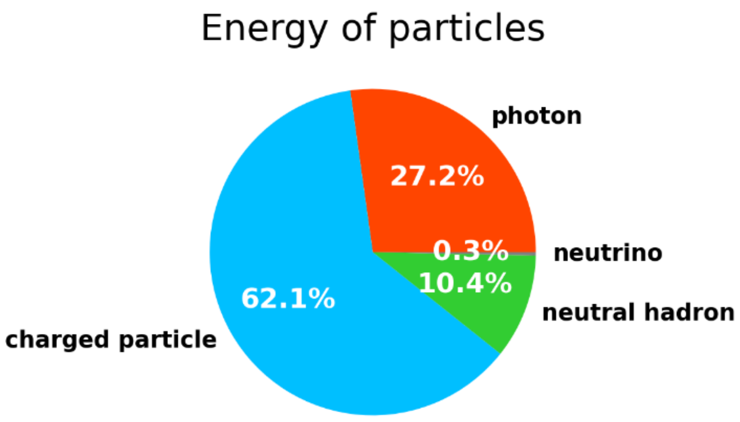


~60% jet energy carried by charged particles (mainly $\pi^{\pm}$)

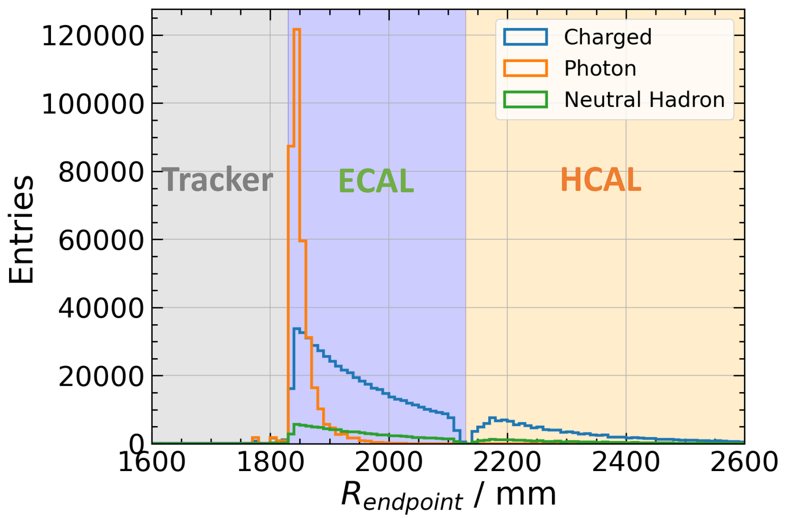
~30% by photons

~10% by neutral hadron

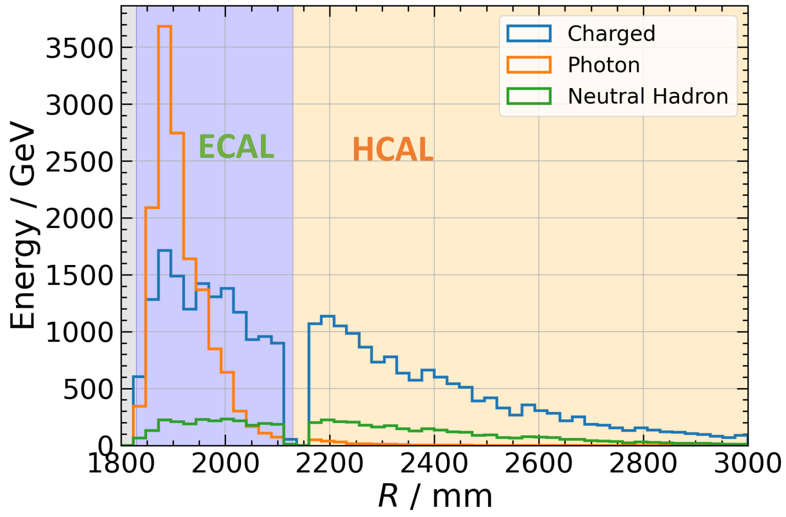




Shower distribution



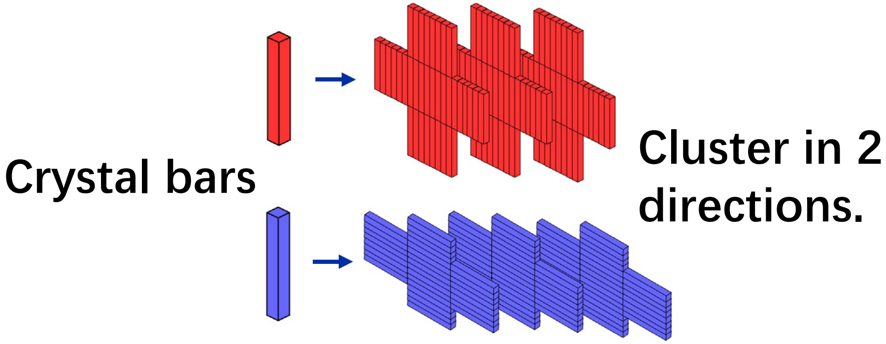
Radius of end-point of particles in jets

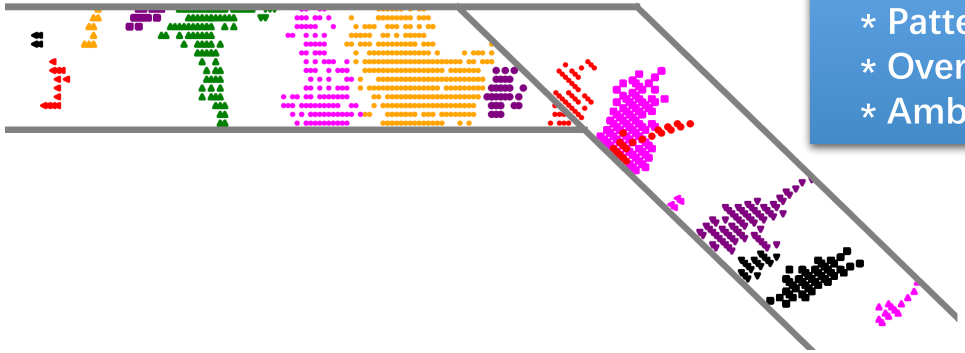


Energy distribution of particles in jets

Clustering & recognition

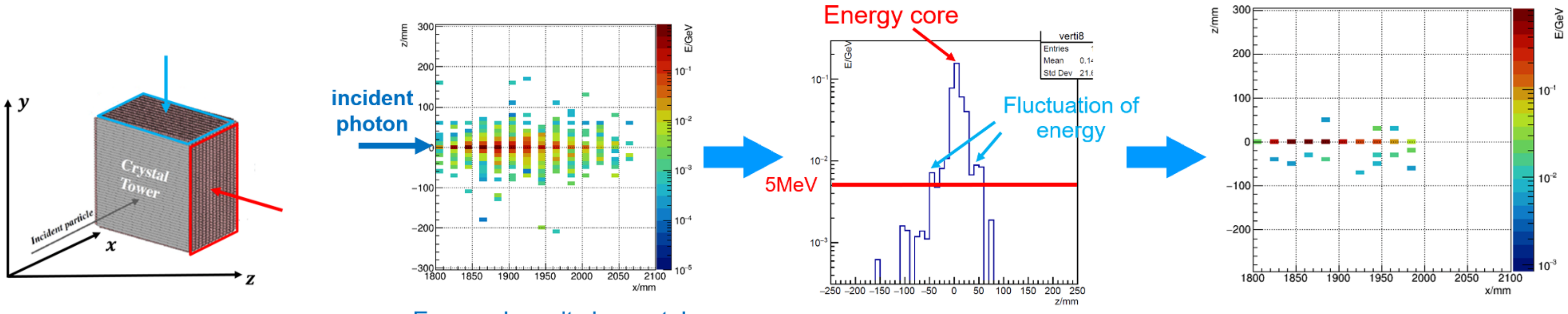
- **Global neighbor clustering for pre-processing.**





Software task:
- ✔ * Clustering
- * Pattern recognition.
- * Overlap: energy splitting.
- * Ambiguity problem.

- **Shower recognition:**
  - Use the local maximum to simplify the pattern in homogeneous ECAL



Energy deposits in crystals

Picking out local-max in each layer

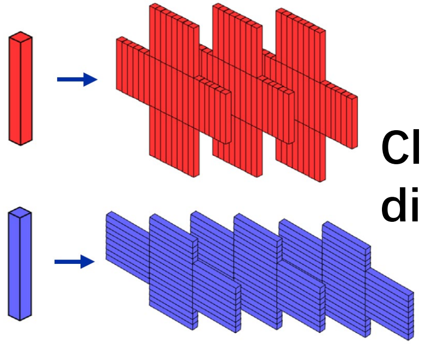
Local Maximum Distribution

Clustering & recognition

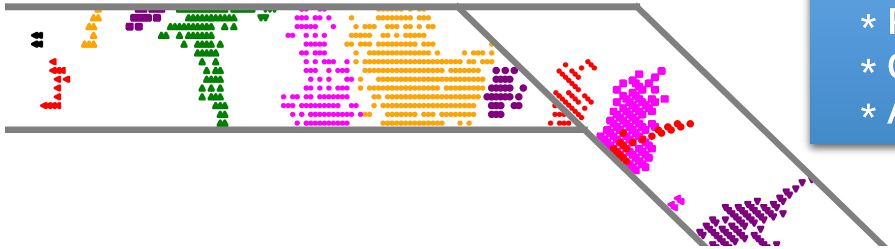
- **Global neighbor clustering for pre-processing.**

Software task:
✔ * Clustering
* Pattern recognition.
* Overlap: energy splitting.
* Ambiguity problem.

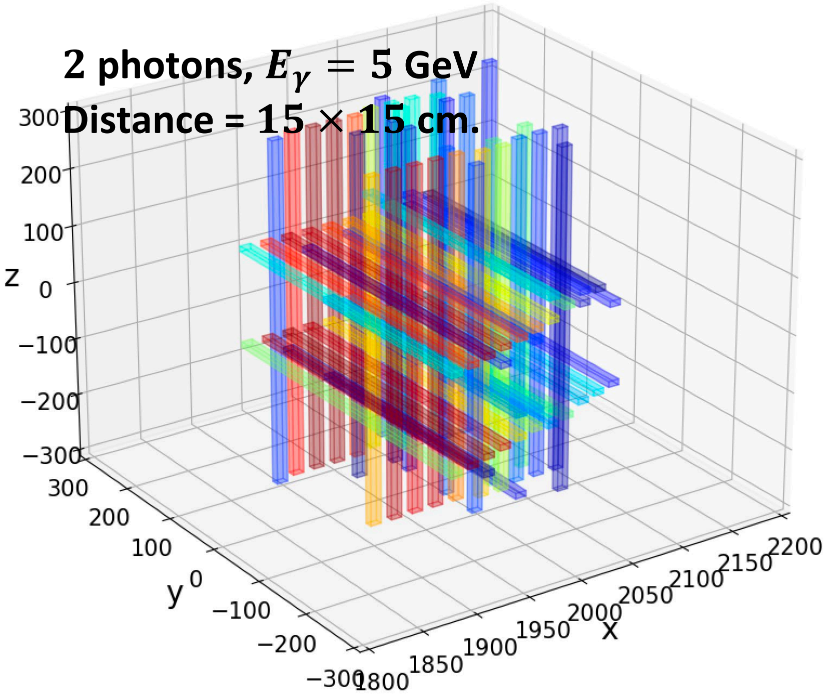
Crystal bars

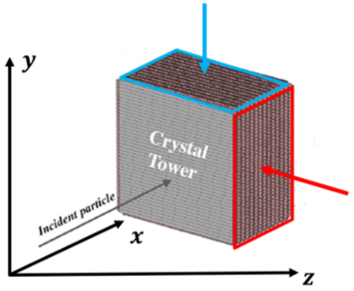


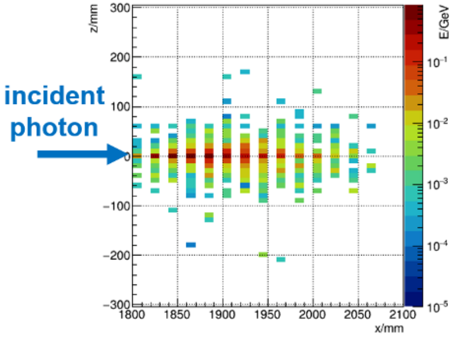
Cluster in 2 directions.



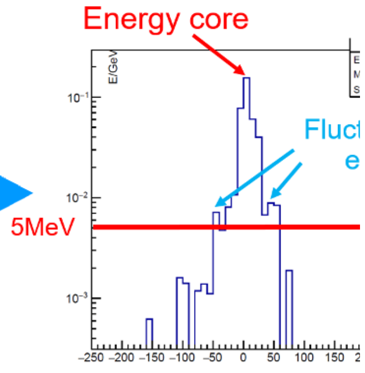
- **Shower recognition:**
  - Use the local maximum to simplify the pattern in homogen







Energy deposits in crystals

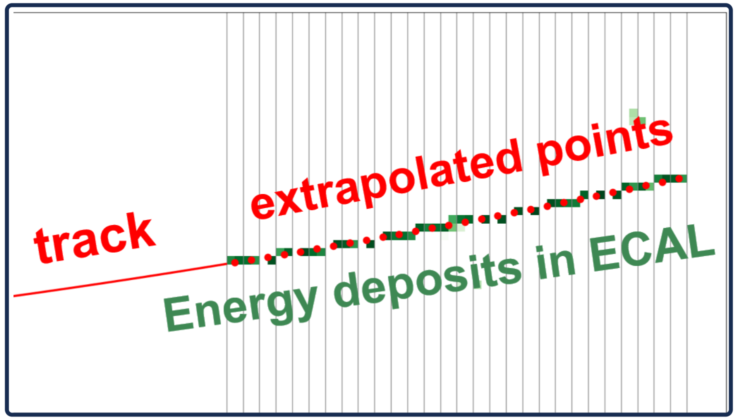


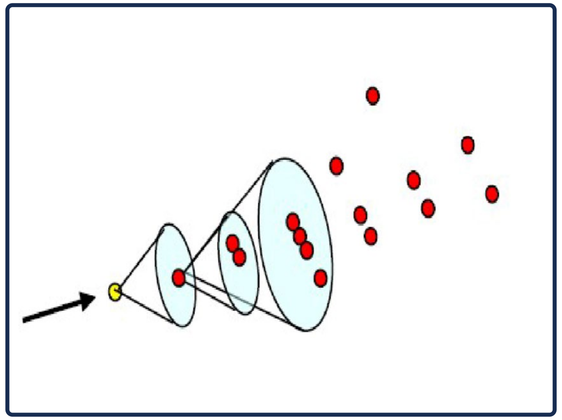
Picking out local-max in ea

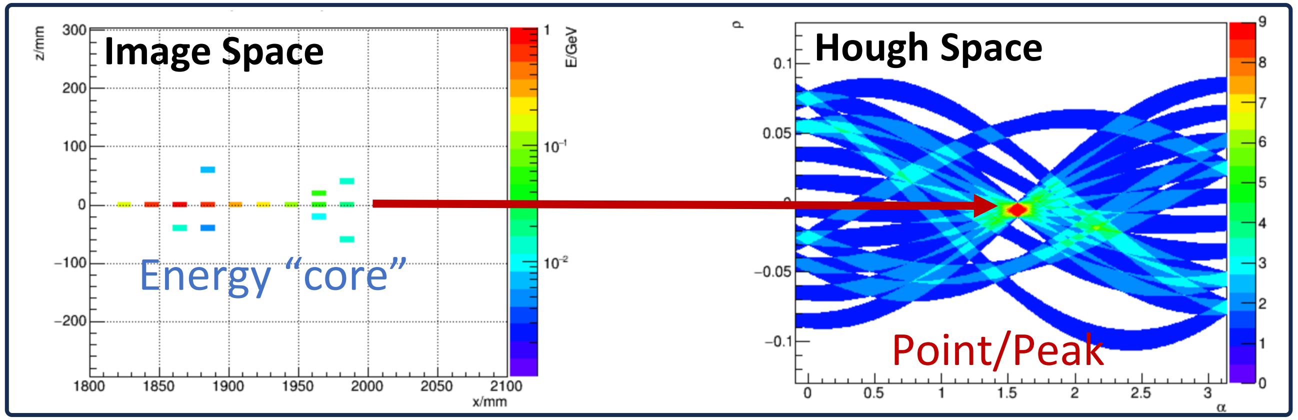
Clustering & recognition

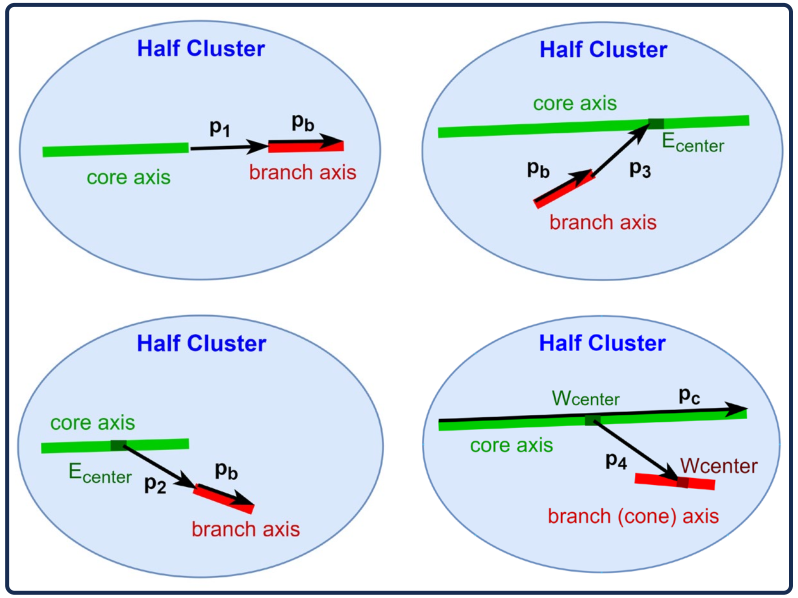
- **Shower recognition:**
  - 3 individual algorithms for different type: track-match, Hough, Cone-clustering.
  - A set of topological cluster merging.

Software task:
✔ * Clustering
✔ * Pattern recognition.
* Overlap: energy splitting.
* Ambiguity problem.





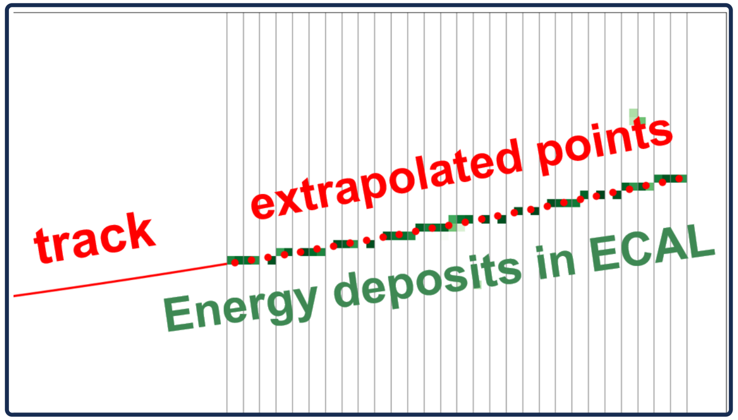


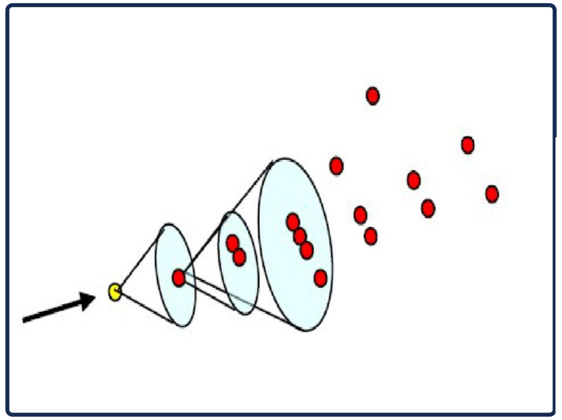


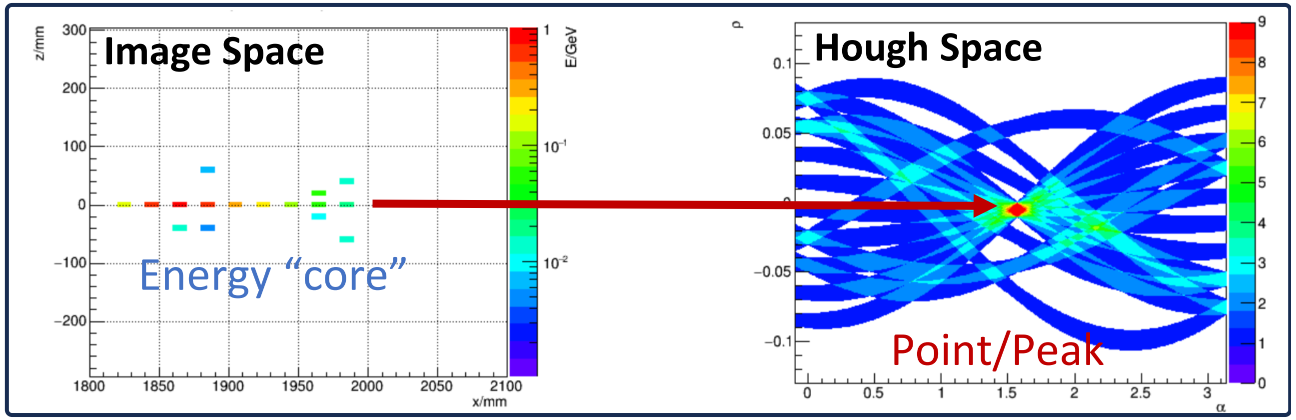
Clustering & recognition

- **Shower recognition:**
  - 3 individual algorithms for different type: track-match, Hough, Cone-clustering.
  - A set of topological cluster merging.

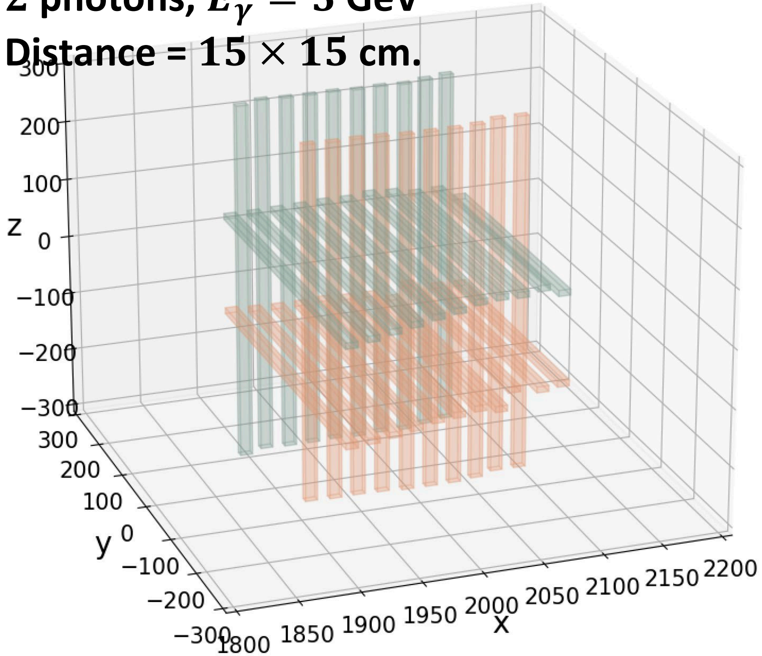
Software task:
✔ * Clustering
✔ * Pattern recognition.
 * Overlap: energy splitting.





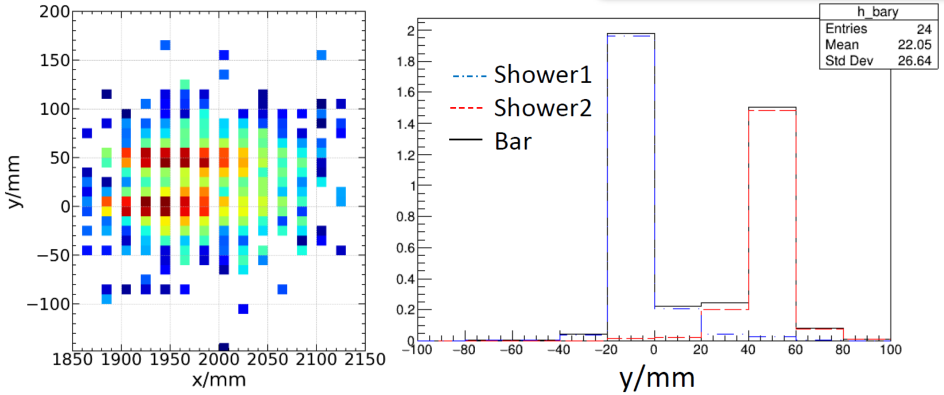


2 photons, $E_\gamma = 5$ GeV
Distance = 15×15 cm.



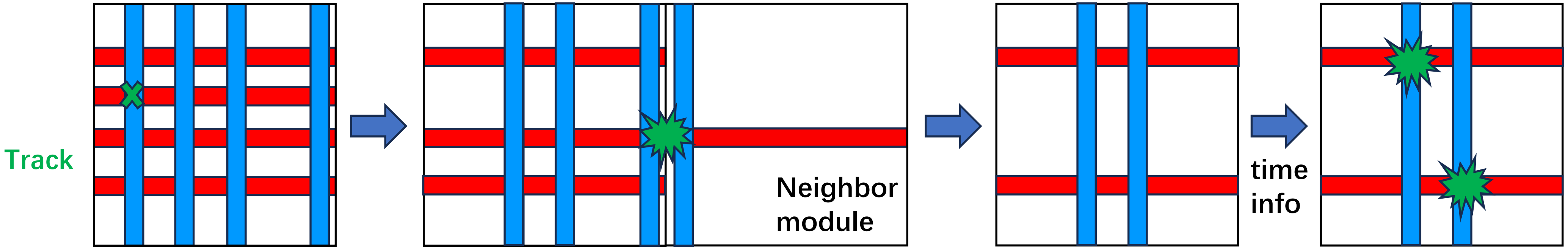
Energy splitting and matching

- **Splitting for the overlapped shower:**
  - Calculate the expected energy deposition from EM profile.
    - Expected energy : $E_{i\mu}^{exp} = E_{\mu}^{seed} \times f(|x_i - x_c|)$
    - Assigned weight: $w_{i\mu} = \frac{E_{i\mu}^{exp}}{\sum_{\mu} E_{i\mu}^{exp}}$



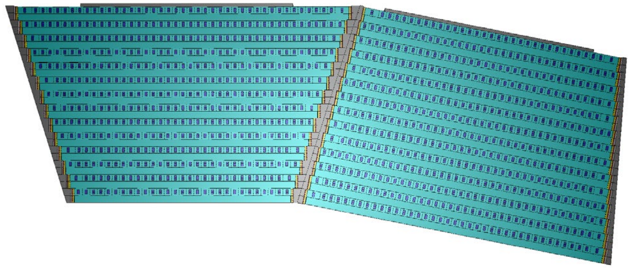
Software task:
✔ * Clustering
✔ * Pattern recognition.
✔ * Overlap: energy splitting.
✔ * Ambiguity problem.

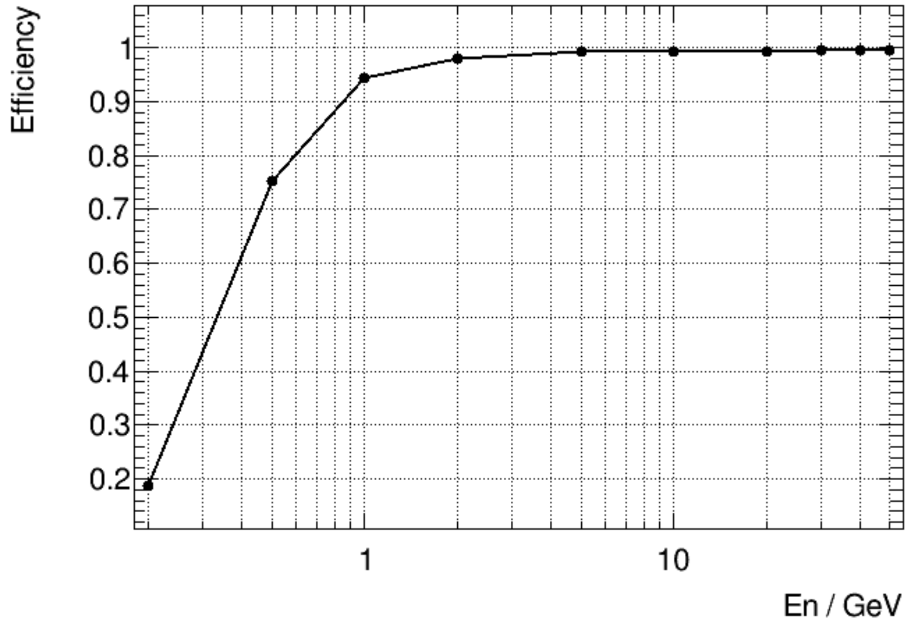
- **Ambiguity removal:**
  - Information from: track, neighbor tower, time.

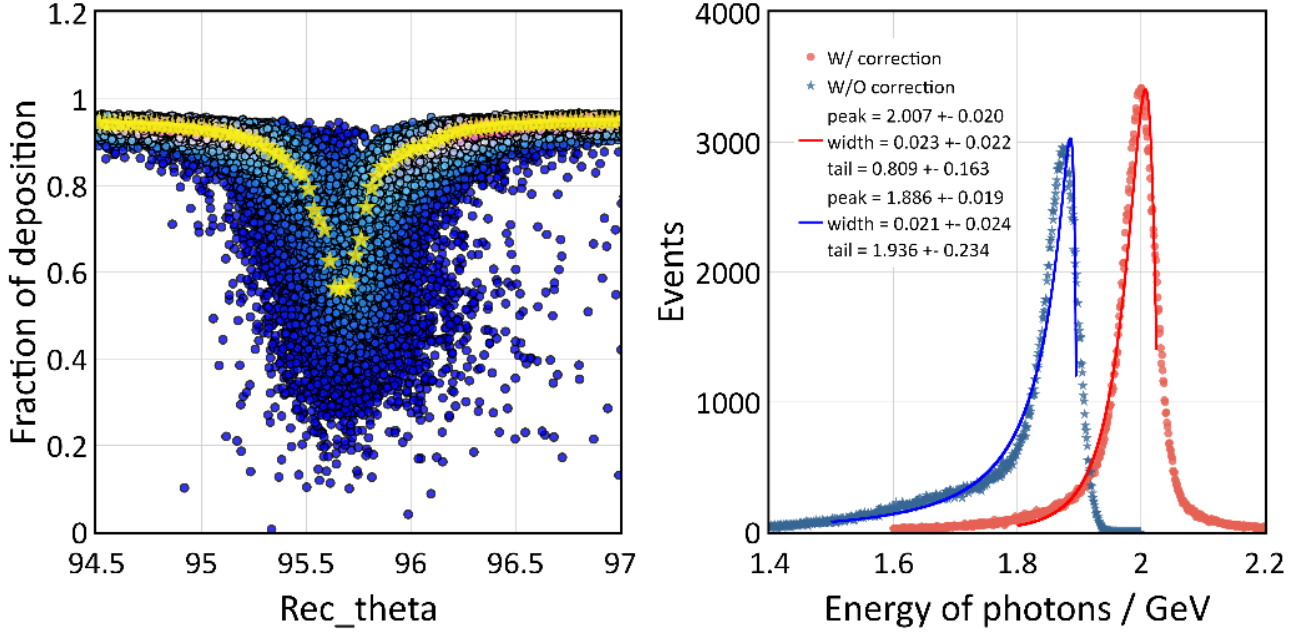


Physics performance: single photon

- **Single photon reconstruction efficiency:**
  - Efficiency: ~100% for >1 GeV photons.
- **Energy correction from simulation:**
  - For the cracks: $E_{corr} = \frac{E'_{truth}}{E'_{deposition}} \times E_{dep}^{mean}$

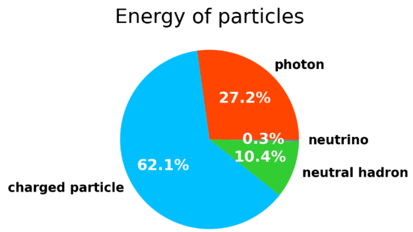






Particle Flow Approach

- **Reconstruct each individual particle with optimal subdetector**
  - Charged particles measured in tracker
  - Photons measured in ECAL
  - Neutral hadrons measured in HCAL



[PandoraPFA](), NIM.A Vol 611, Issue 1, 2009

Contributions from the different particle components to the jet-energy resolution (all energies in GeV).

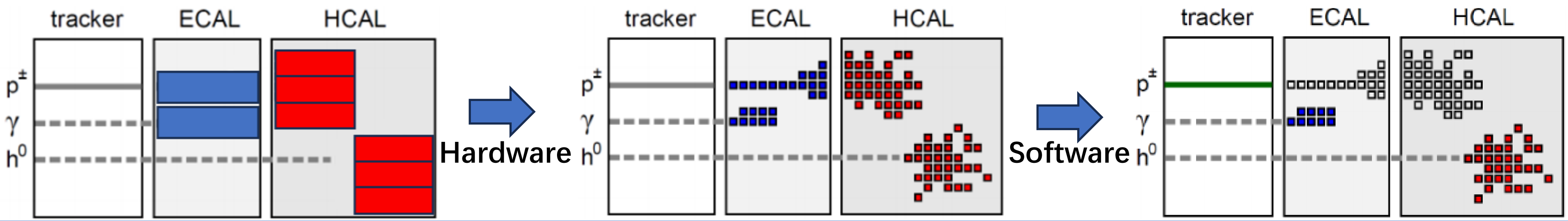
Component	Detector	Energy fract.	Energy res.	Jet energy res.
Charged particles ($X^{\pm}$)	Tracker	$\sim 0.6Ej$	$10^{-4}E_{X^{\pm}}^2$	$< 3.6\times10^{-5}E_j^2$
Photons (γ)	ECAL	$\sim 0.3E_j$	$0.15\sqrt{E_\gamma}$	$0.08\sqrt{E_j}$
Neutral Hadrons (h^0)	HCAL	$\sim 0.1E_j$	$0.55\sqrt{E_{h^0}}$	$0.17\sqrt{E_j}$

$0.03\sqrt{E_\gamma}$ for crystal ECAL

- **Hardware + Software**

Hardware
separate clusters from different particles.
- Imaging calorimeter

Software
Assign energy deposits to each particle.
- Sophisticated pattern recognition



- **Confusion term is the limitation factor**
  - $\sigma_{jet} = \sigma_{tracker} \oplus \sigma_{ECAL} \oplus \sigma_{HCAL} \oplus \sigma_{confusion}$
  - Arises from imperfectly associating all energy deposits with the correct particles

Diagrams illustrating two types of confusion