Electromagnetic Calorimeter Software for Super Tau-Charm Facility

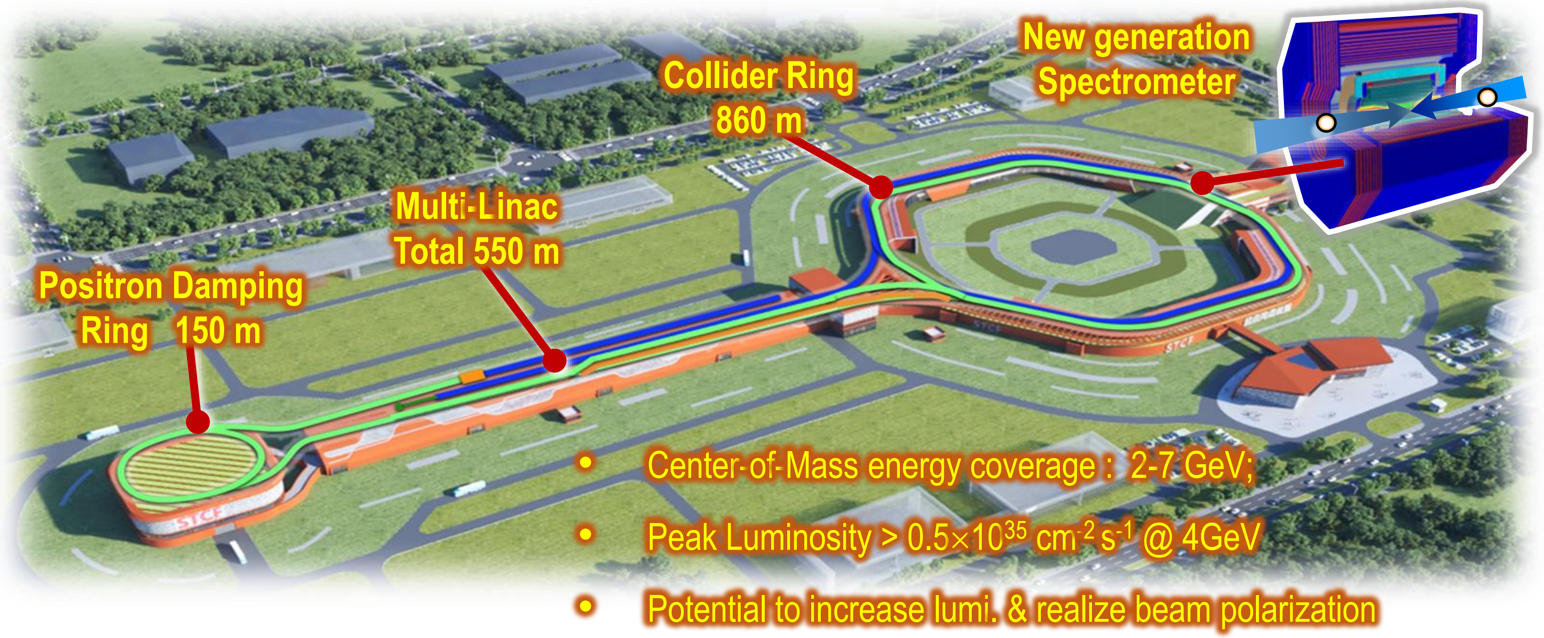
Bo Wang
University of Science and Technology of China
State Key Laboratory of Particle Detection and Electronics
(On behalf of the STCF ECAL working group)

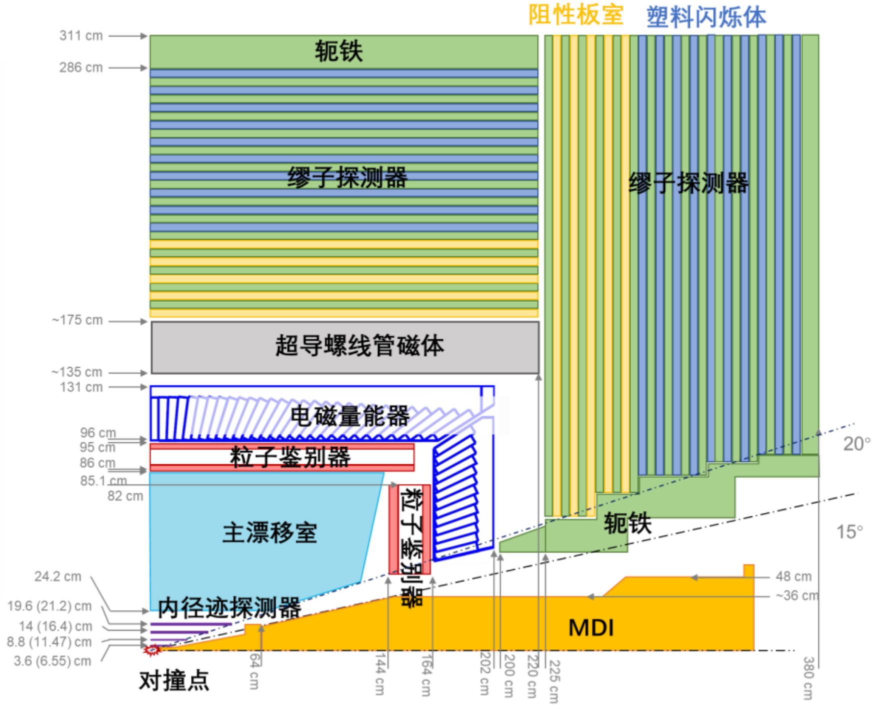
超级陶粲装置研讨会

2026年7月2日，西安

Super Tau-Charm Facility

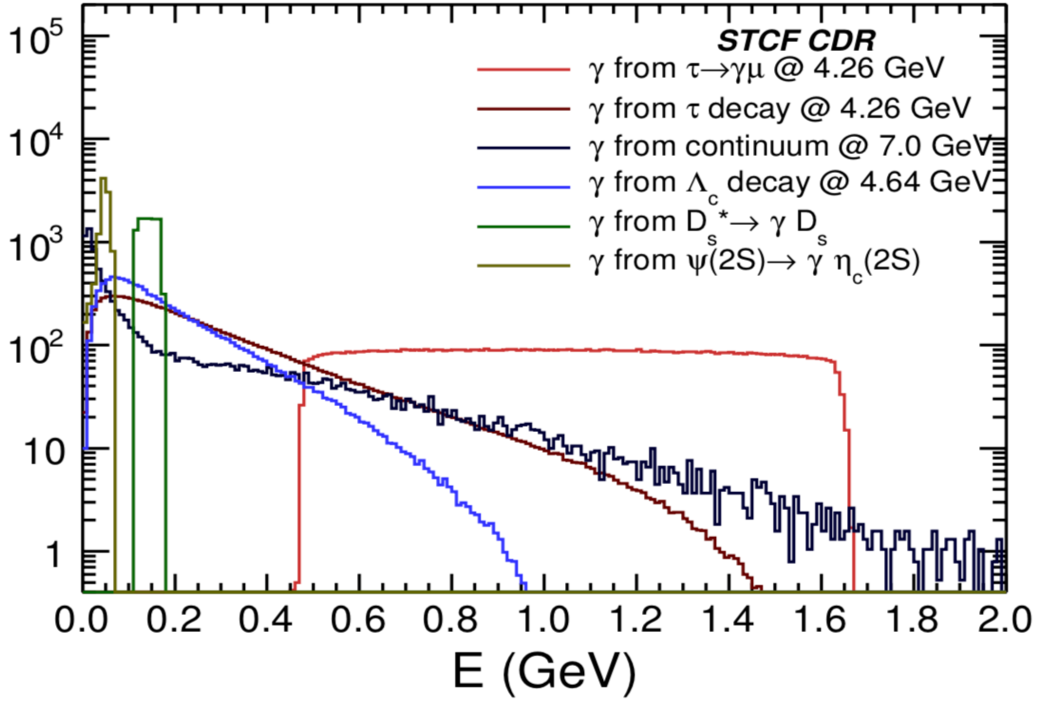
- Electron-positron collider experiment
- High luminosity： beyond $0.5 \times 10^{35}\ cm^{-2} \cdot s^{-1}$ @ 4GeV
- Wide energy region： center-of-mass energy range of 2~7 GeV





Requirements for ECAL

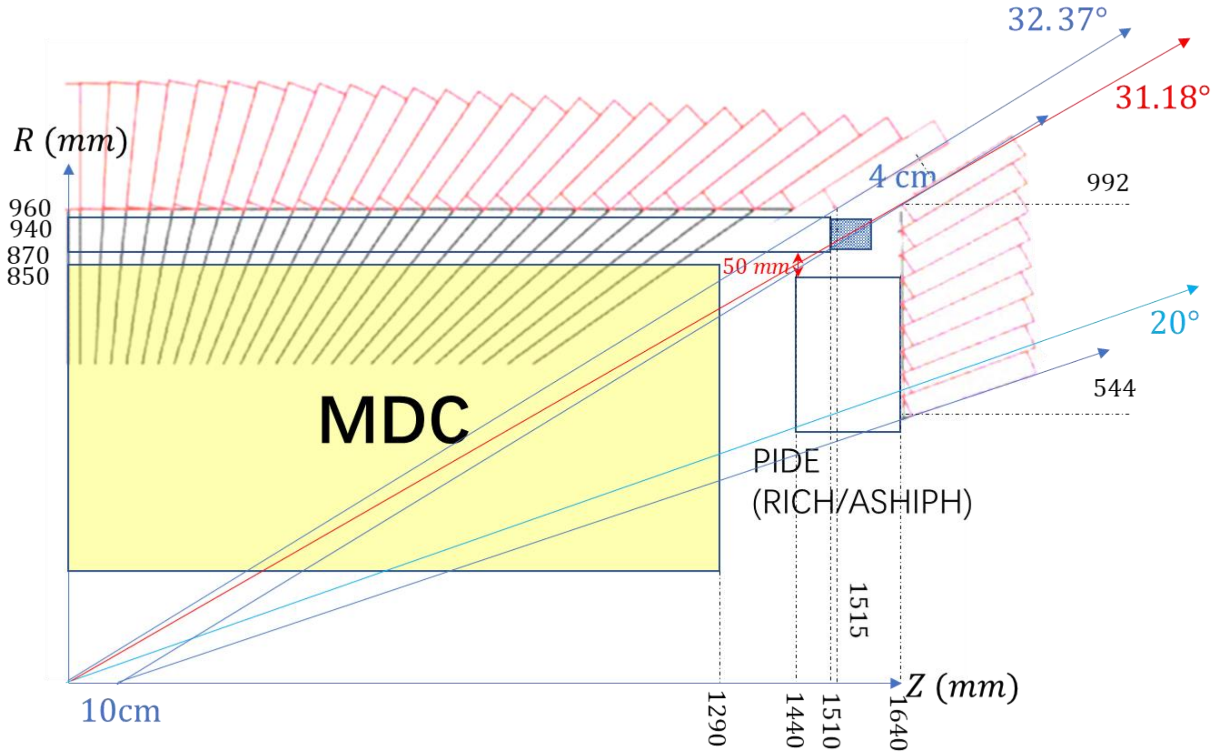
- **Fast response**
  - Challenge of high Luminosity
    - High count rate
    - Extremely high background
- **High precision**
  - Energy resolution
    - Better than 2.5% @1GeV
  - Position resolution
    - Approximately 5mm @1GeV
  - Time resolution
    - Better than 300ps @1GeV

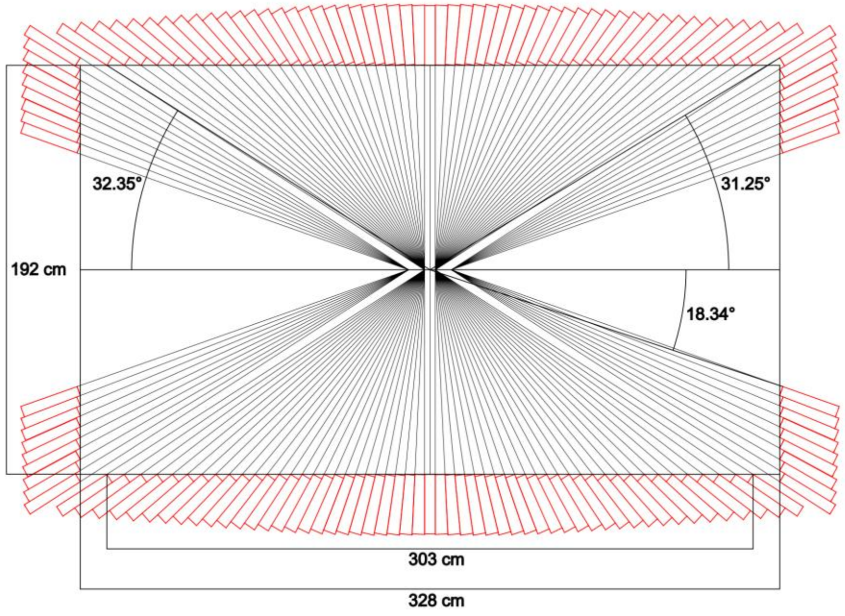


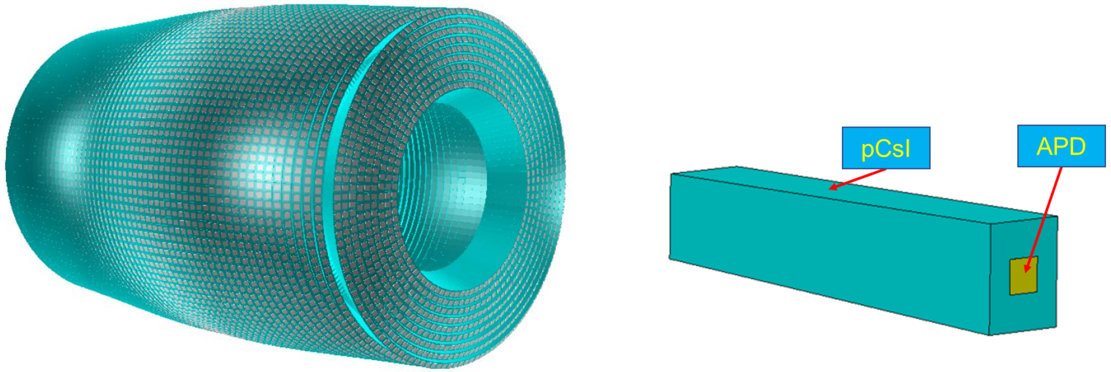
Energy distribution for photons

ECAL Geometry Optimization

- Total absorption calorimeter
  - Barrel: $(51 \rightarrow 53) \times 132 = 6996$
  - Endcap: $3 \times (85 + 105 + 130) \times 2 = 1920$
  - 11% volume reduction

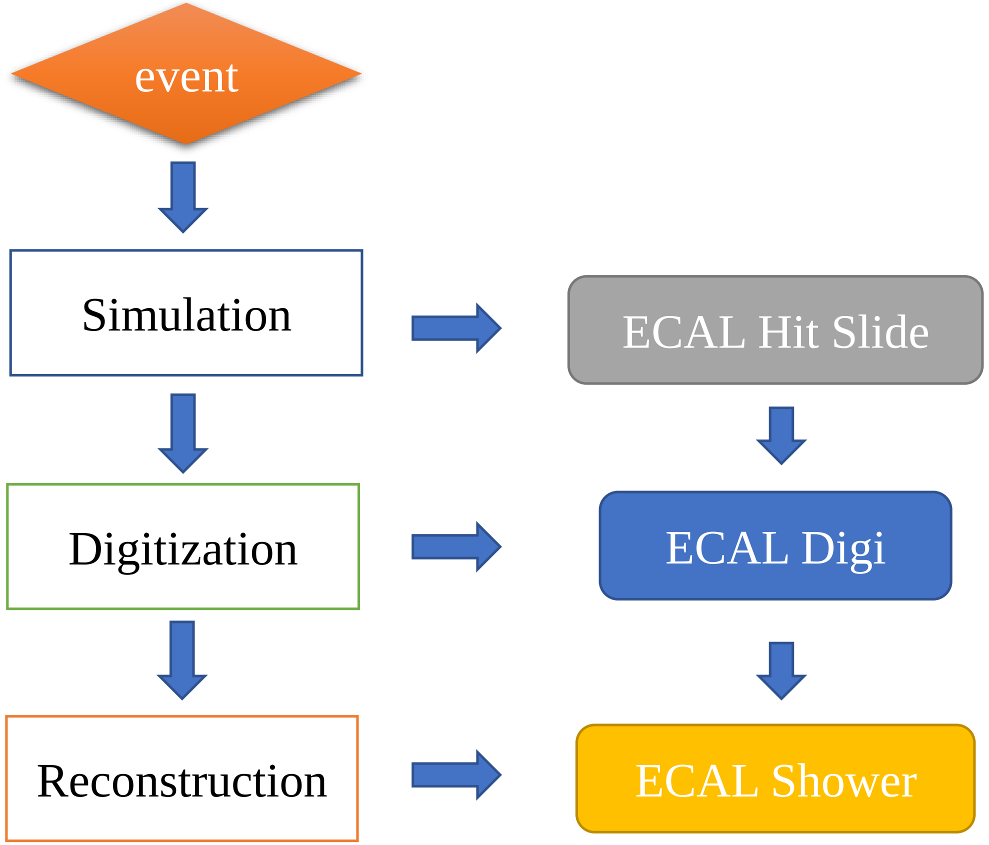






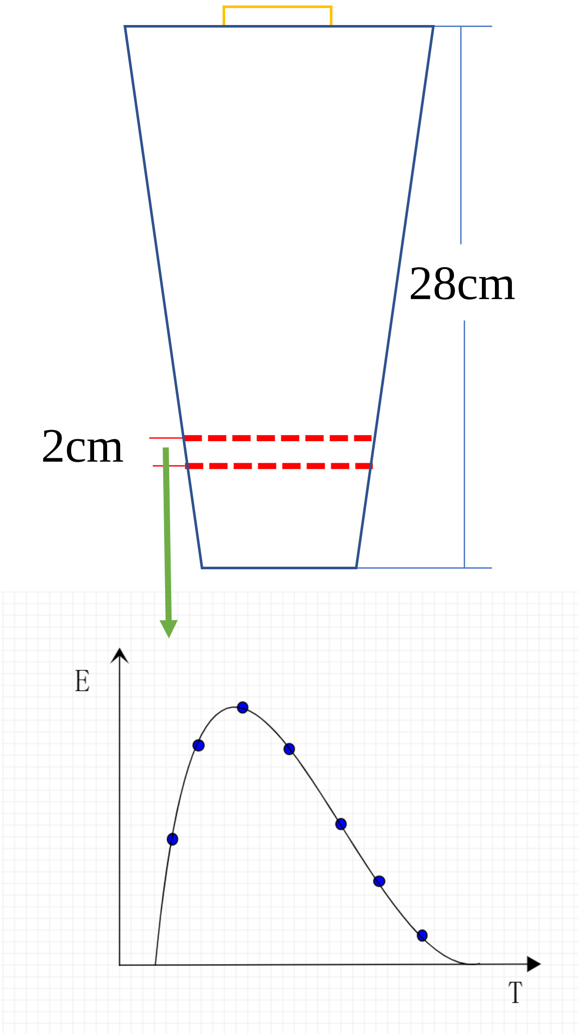
ECAL Full Chain

- Simulation
  - Geant4 based
  - Vertical Layered Crystal Data Structure
- Digitization
  - Waveform shaping
  - Waveform fitting
    - Peaking finding method
    - Template fit method
    - Pipeline fit method
- Reconstruction
  - Classical cluster reconstruction method
  - 4D Cluster reconstruction



Simulation Algorithm

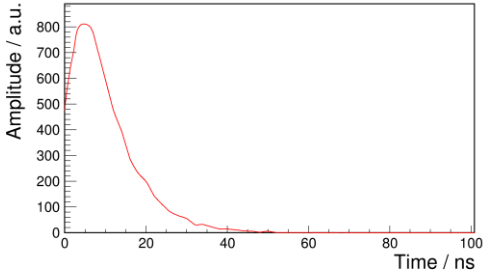
- Based on Geant4 Simulation
  - Process the information for each Geant4 step
- Data Model (layer as a unit)
  - Thickness 2cm as a layer unit
  - Each unit with a energy-time distribution: points with (E,T)
  - Bin width 500ps for energy-time distribution



- ✓ Save storage space
- ✓ Minimal loss of information
- ✓ Facilitate the Digitization

Digitization Algorithm

Waveform shaping



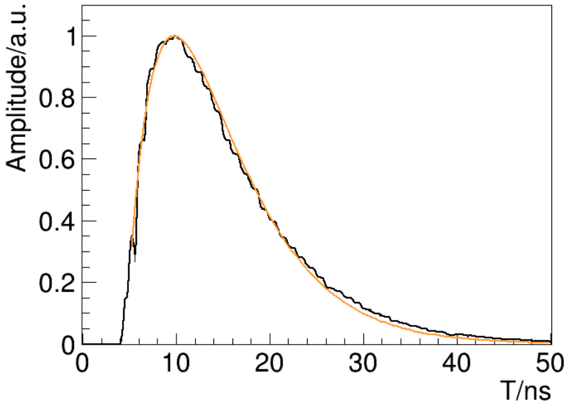
Geant4 optical simulation for each layer unit (standalone)

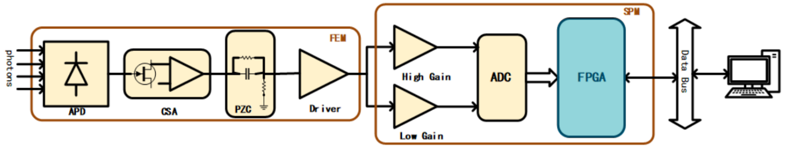
Sampling

Points (E-T) in each layer unit

Fluorescence transfer parameterized simulation

Light output nonuniformity

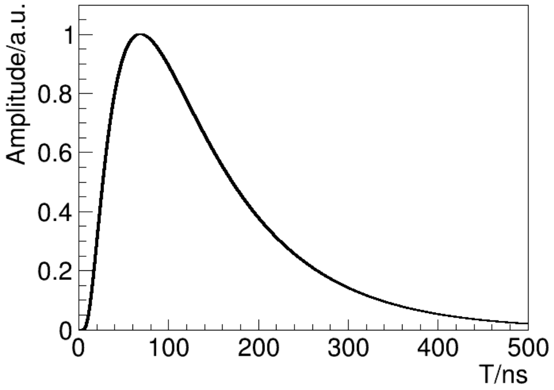




convolution

Scintillation process

Electronics response function



Electronic noise

ADC (12.5ns)

Output waveform

Strategy 0

Digitization Algorithm

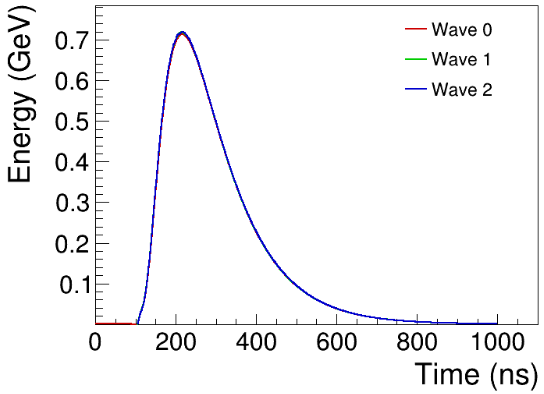
Waveform shaping

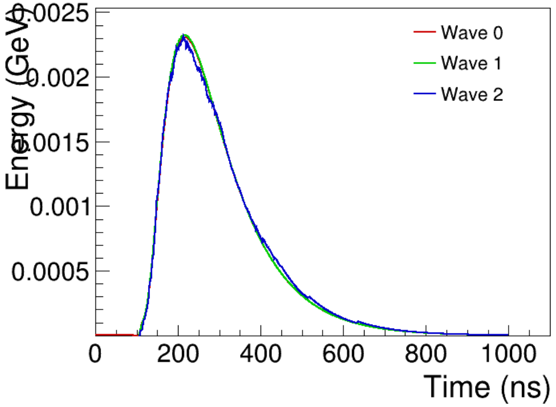
Strategy 1

Time: 15s→ 10s per event

convolution

Points (E-T) in each layer unit → Fluorescence transfer parameterized simulation / Scintillation process / Electronics response function → ADC (12.5ns)





Strategy 2

Time: Comparable

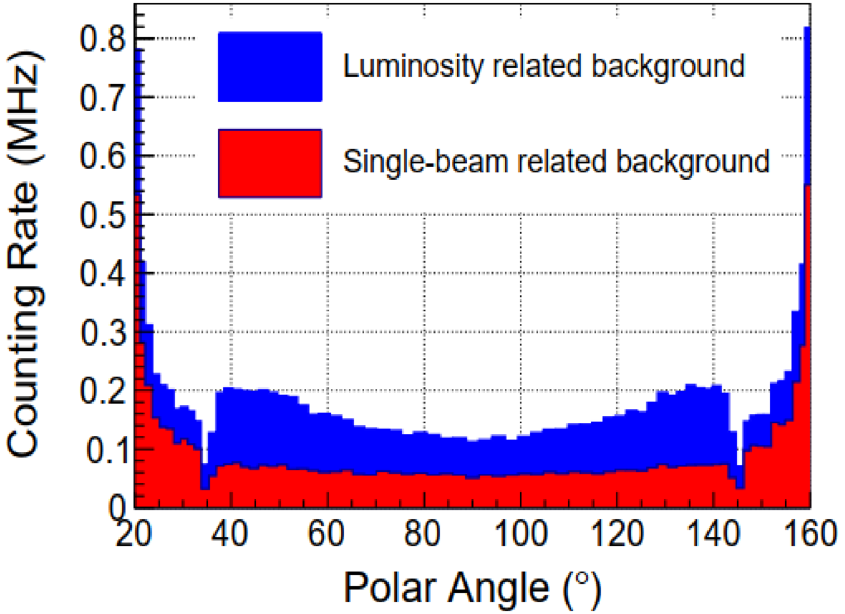
Sampling

convolution

Points (E-T) in each layer unit → Fluorescence transfer parameterized simulation / Scintillation process → Electronics response function → ADC (12.5ns)

Challenges of high background

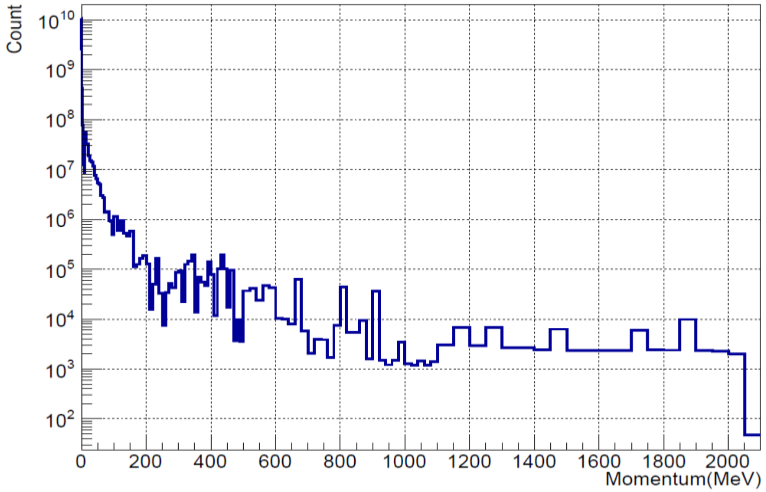
- **Luminosity-related Background**
  - Radiative BhaBha Scattering (RBB)
  - Two Photon Process



Variation of the background counting rate with polar angle

Counting rate reaches the order of MHz

- **Single-beam related Background**
  - Thouschek Effect
  - Coulomb Scattering
  - Bremsstrahlung



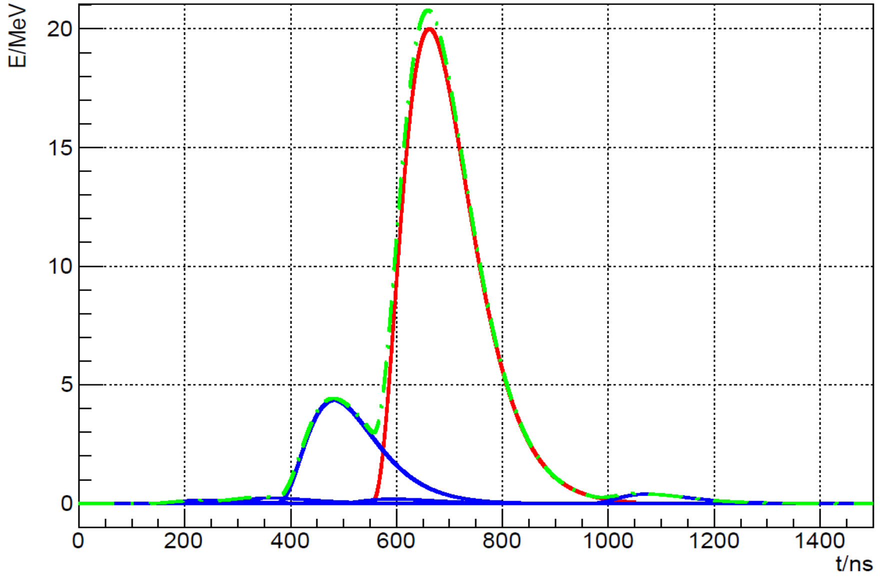
Momentum distribution of background particles

Most background particles concentrate in the low momentum region

Digitization Algorithm

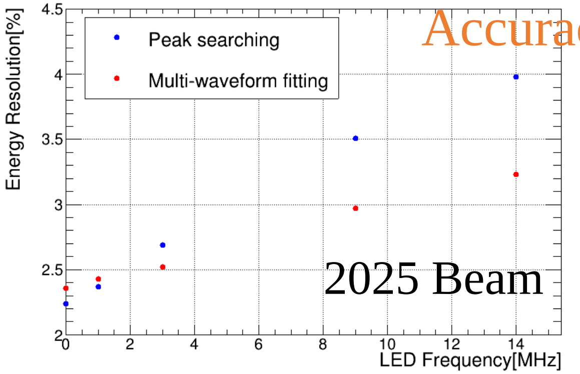
Pile-up recovery

- Pile-up is superimposed on the signal waveform
  - Inaccurate fit results of amplitude and time
  - Larger resolution of energy and time



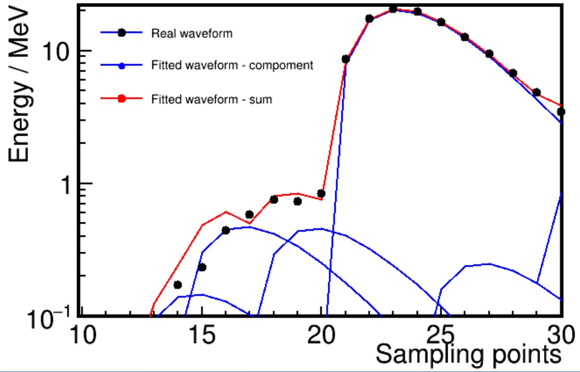
The amplitude of signal is distorted

- Pipeline fitting method



Accuracy unsatisfactory!

- Multi-template fit
  - Has been used to study the capability for pile-up recovery



Digitization Algorithm

Pipeline fit

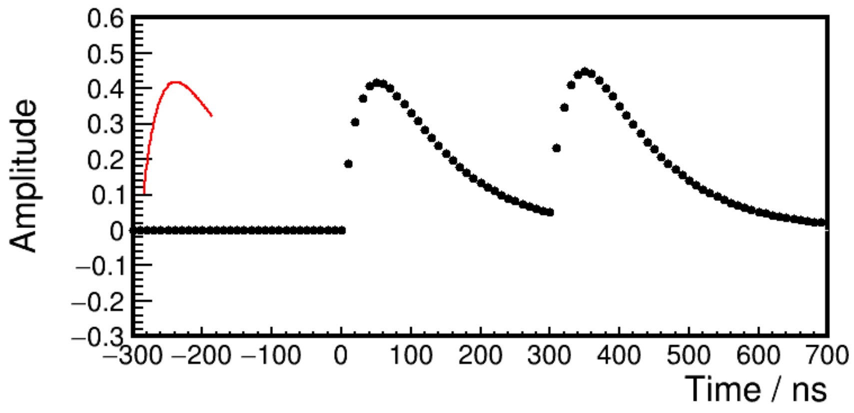
- Interpolation fitting
  - Real-time online processing
  - Template fit once for each fitting
    - Each fitting begin with different ADC point
    - Add one more fitting between two ADC points
  - Fit successful
    - $A > E_{thr}$
    - $\triangle T < 12.5/2$
    - $\chi^2/ndf < 3$:

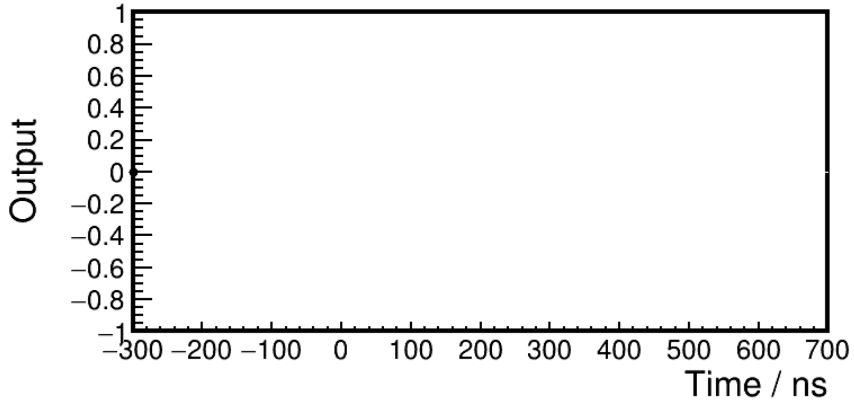
$$\chi^2/ndf = [\sum_{i,j}(y_i - A \cdot f(t_i - \tau) - p) \cdot S_{ij}^{-1} \cdot (y_j - A \cdot f(t_j - \tau) - p)] / ((n-2) \cdot (\sigma_{nos}^2 + (A \cdot 0.01)^2))$$

→ Cache and compared with next fit

→ Remove template

→ Ongoing processing



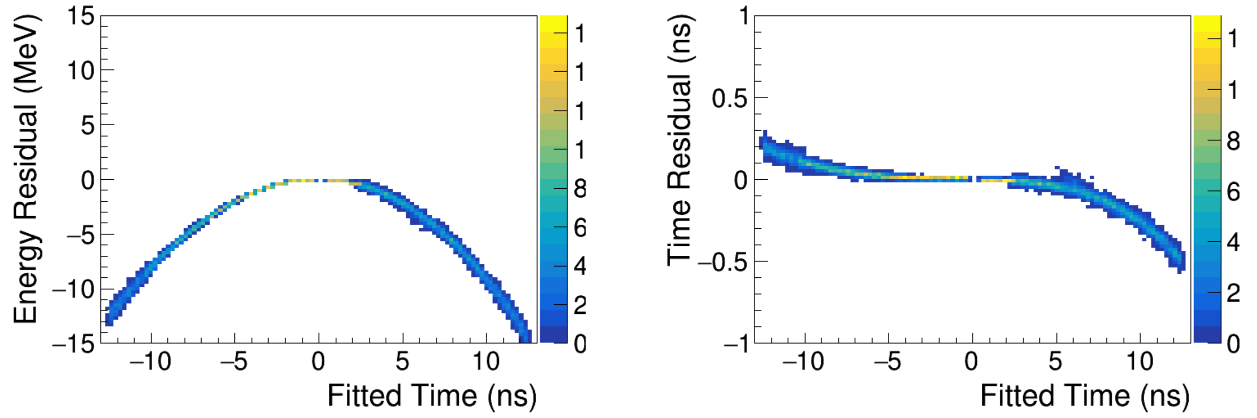


Digitization Algorithm

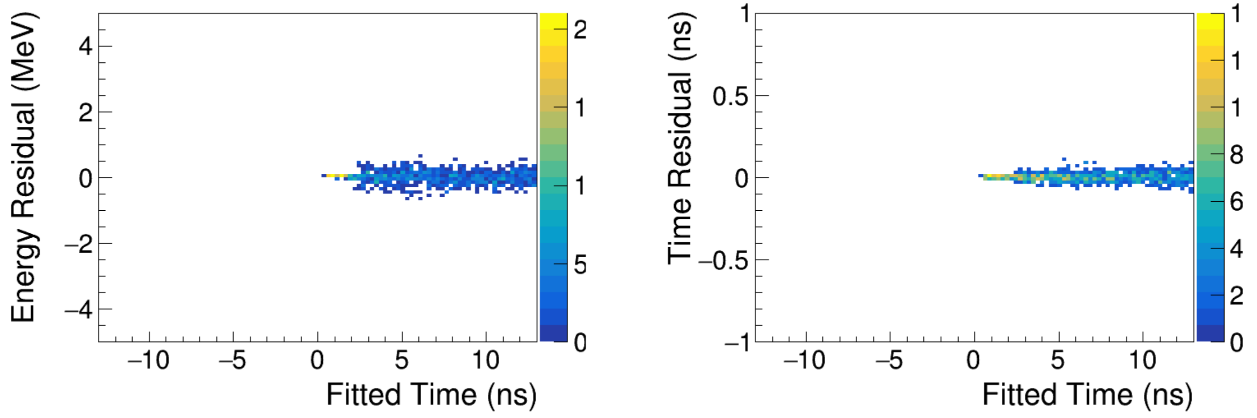
Pipeline fit

- A-T correction
  - Real-time online processing
  - Template fit once for each fitting
  - Fit successful
    - $A > E_{thr}$
    - $\triangle T < 12.5$
  - Correction with control sample
    - $\triangle T \sim (E_{pipe} - E_{single})$
    - $\triangle T \sim (T_{pipe} - T_{single})$

Before Correction



After Correction

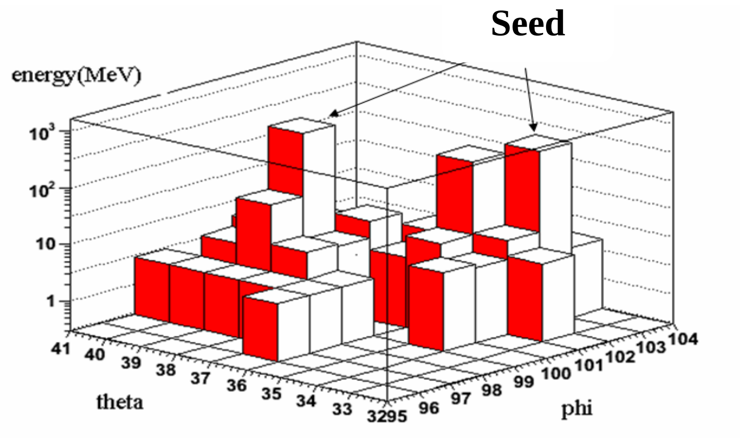


Reconstruction Algorithm

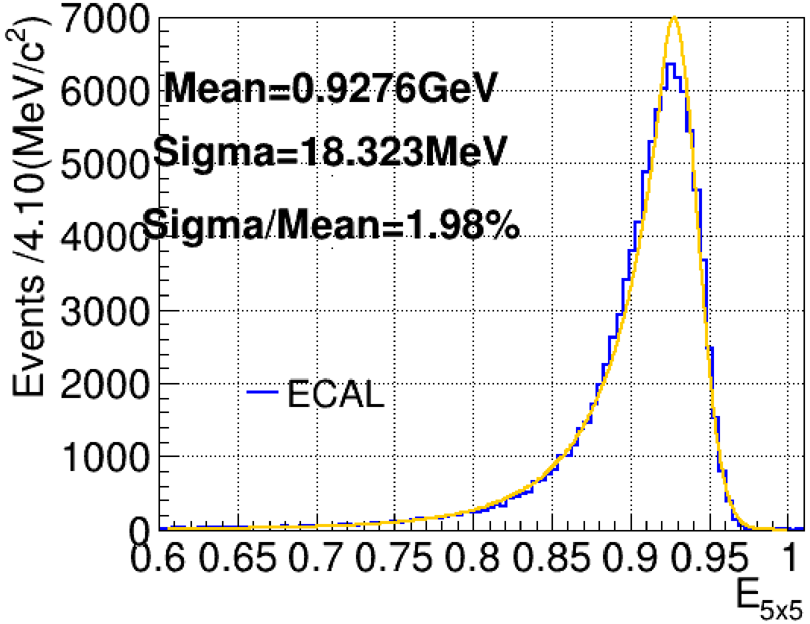
- A complete reconstruction algorithm of ECAL is developed

ECALRecAlg

Cluster search → **Seed search** → **Clusters split into Showers** → **Calculation and correction of shower energy and position**



π^0 Cluster (two photons)



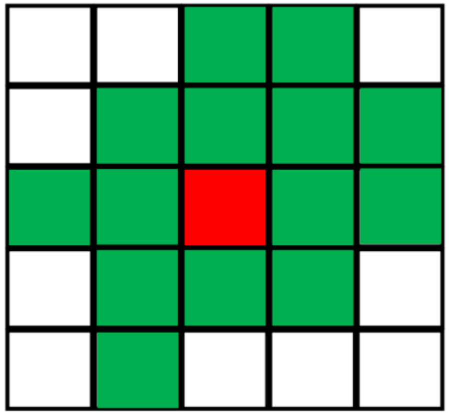
Energy distribution of 1GeV γ with background

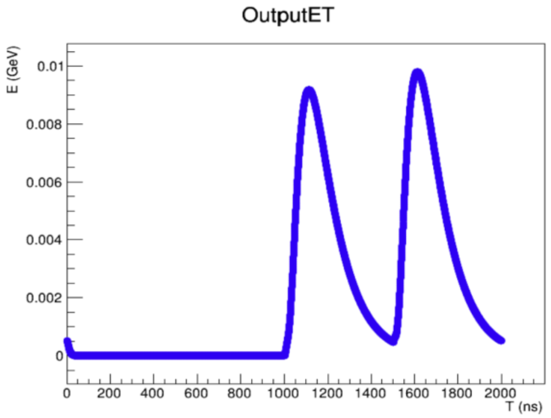
- Fitted by Crystal Ball function
- Energy resolution defined by

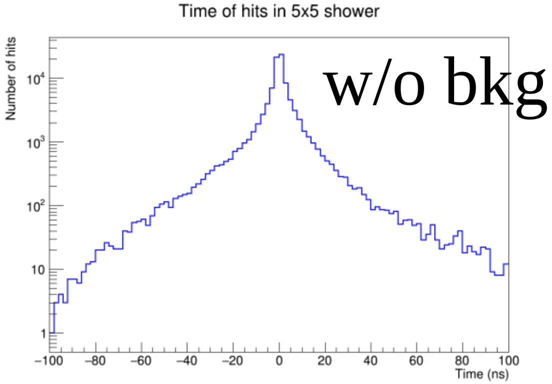
$$\sigma_E = \frac{FWHM}{2.355}$$

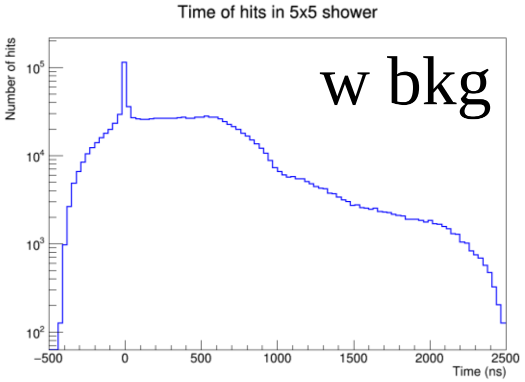
Reconstruction Algorithm

- No Time information





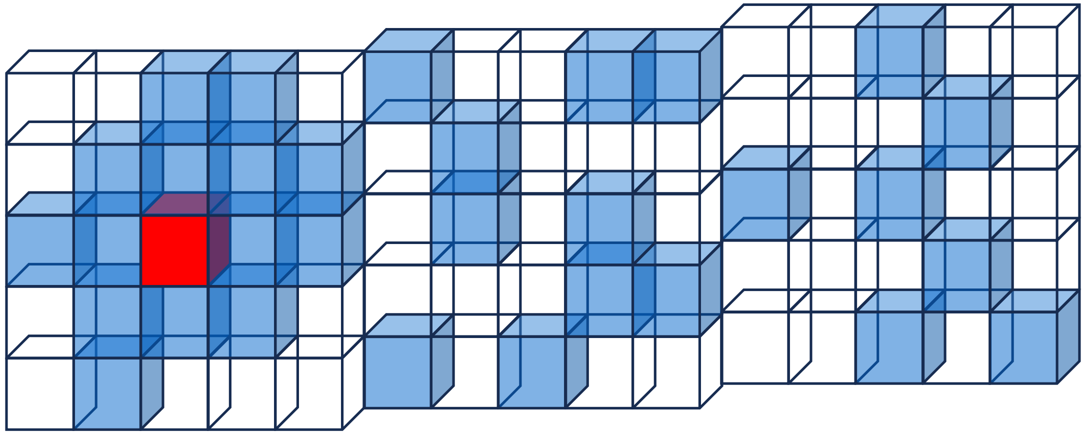


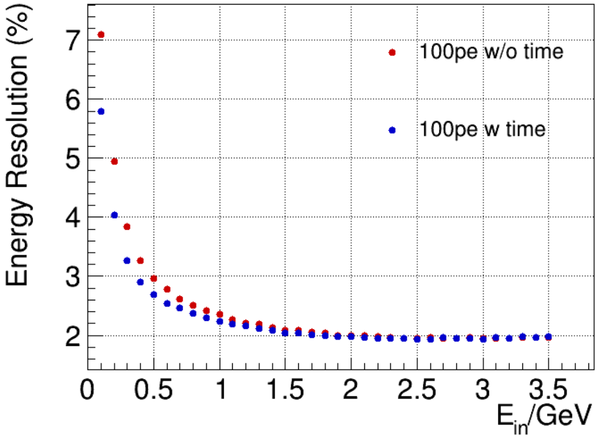


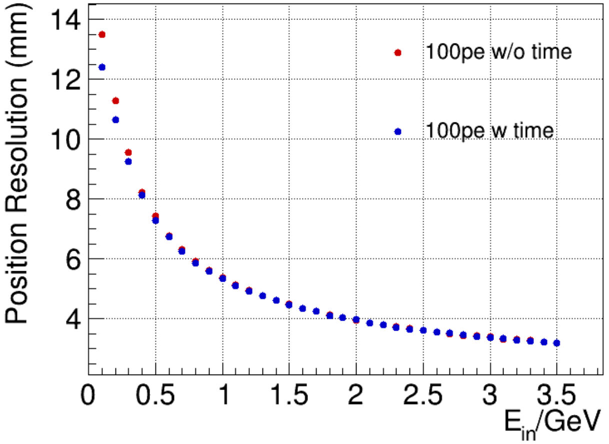
- Digi→Hit
  - The Digi with max energy will be chosen as the Hit to reconstruct the shower
- Hit→Cluster
  - The connected ignited crystals will be treated as a Cluster
- Background
  - Affects signal resolution
  - A large number of fake photons
  - Affects π^0 mass resolution

Reconstruction Algorithm

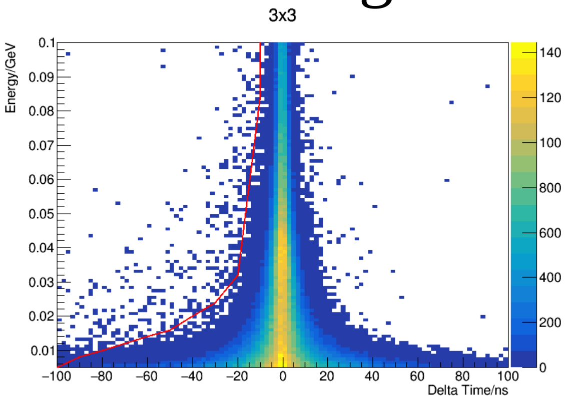
- 4D Reconstruction
  - Seed finding in 3D space with timing information
  - Add time selection while reconstructing Cluster



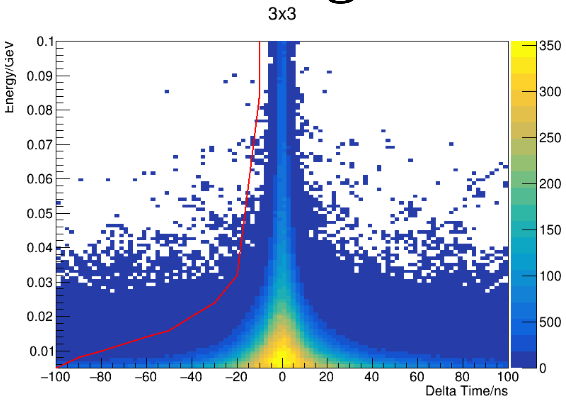




w/o bkg

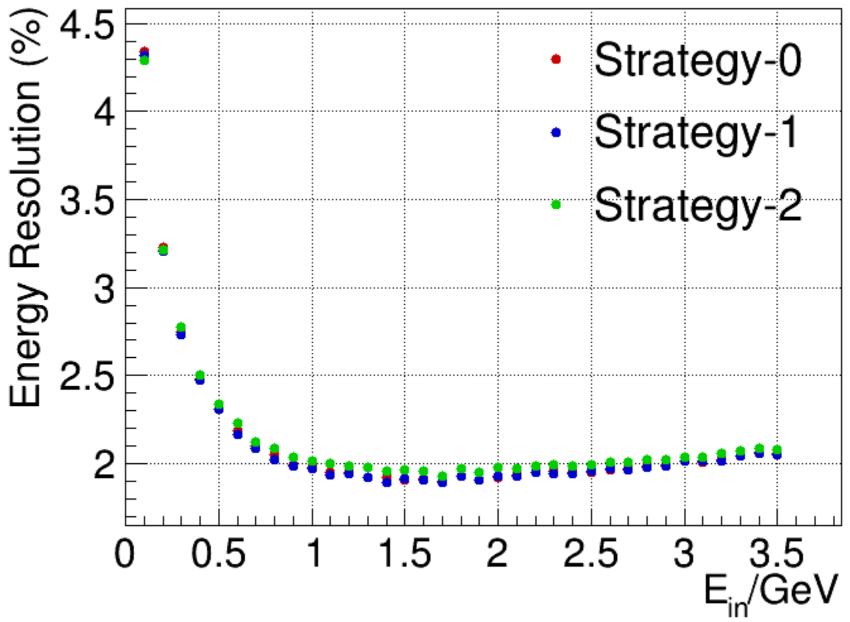


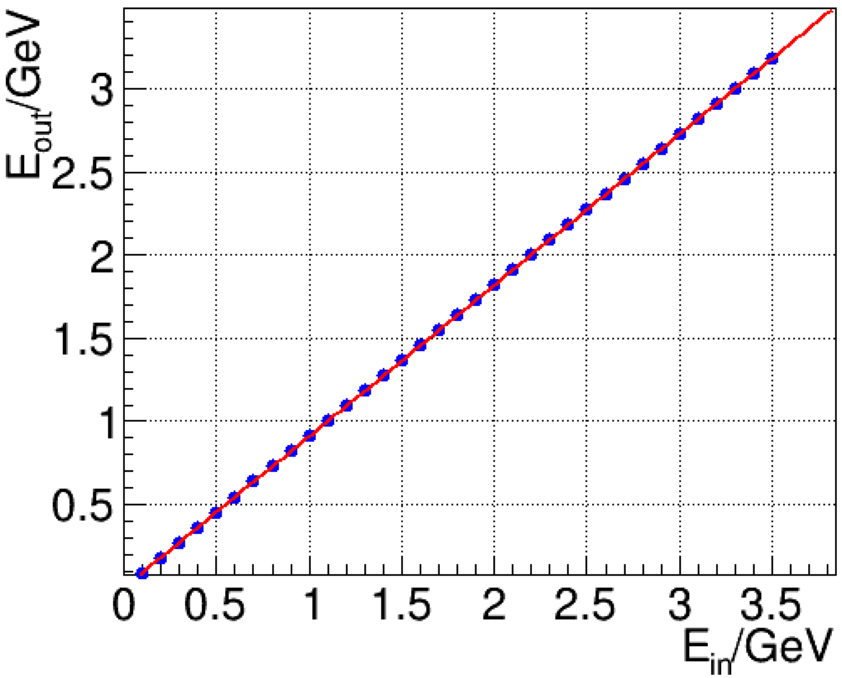
w bkg



Reconstruction Performance

Energy Reconstruction



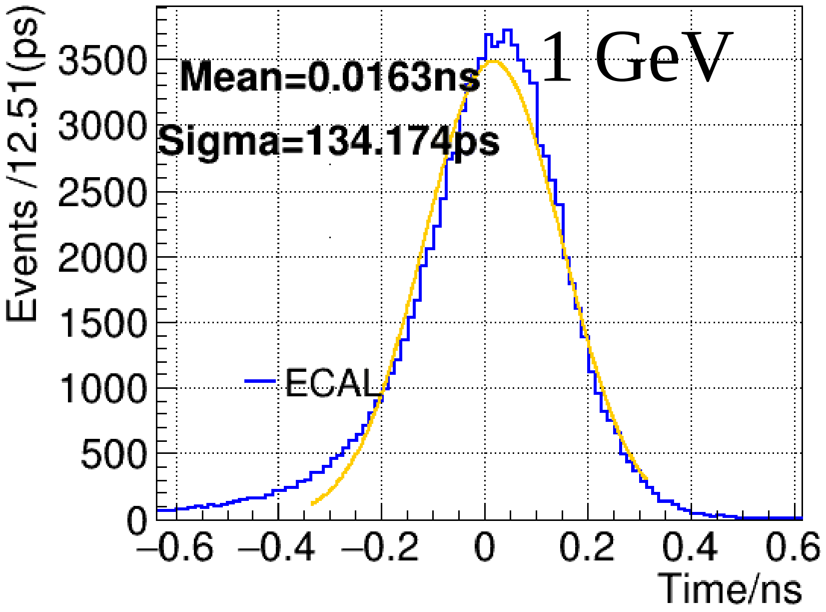


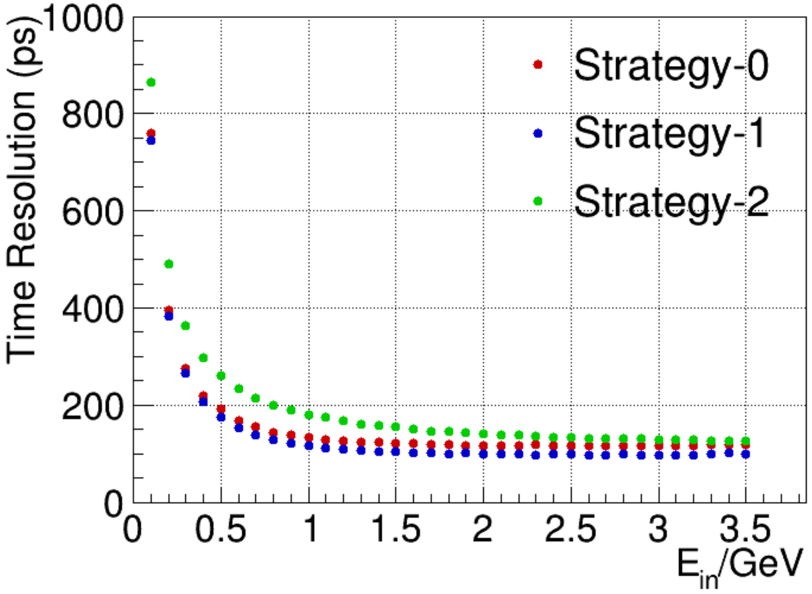
- ✓ $\sigma_E = 1.98\%$ @ 1 GeV
- ✓ $\sigma_E = 4.34\%$ @ 0.1 GeV
- ✓ With background
- ✓ Meet the requirement

- ✓ Good energy linearity in the energy range of 100 MeV ~ 3.5 GeV

Reconstruction Performance

Time Reconstruction





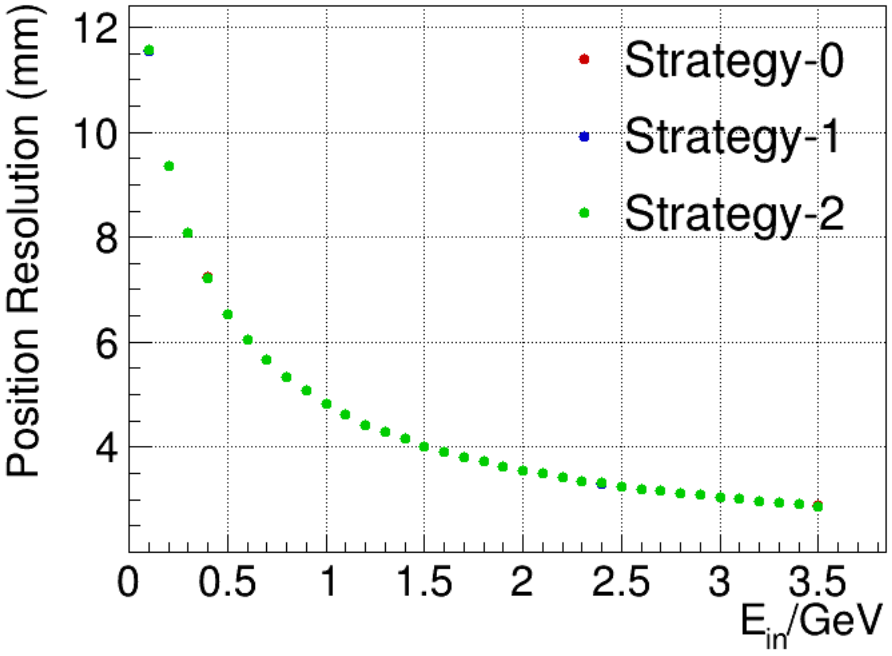
- ✓ σ_T = 134 ps @ 1 GeV
- ✓ Fitted with Gaussian function
- ✓ With background
- ✓ Meet the requirement

Reconstruction Performance

Position Reconstruction

- Splitting algorithms (used by BESIII and Panda)
  - $a_{ik} = \frac{E_k \times \exp(c \times \frac{r_{ik}}{R_M})}{\sum_{j=1}^{m} E_j \times \exp(c \times \frac{r_{ij}}{R_M})}$
- Barycenter method
  - $X_c = \sum_j^N W_j(E_j) \cdot X_j / \sum_j^N W_j(E_j)$

  Where $W_j(E_j) = \max\{0, a + \ln(E_j / \sum_j^N E_j)\}$

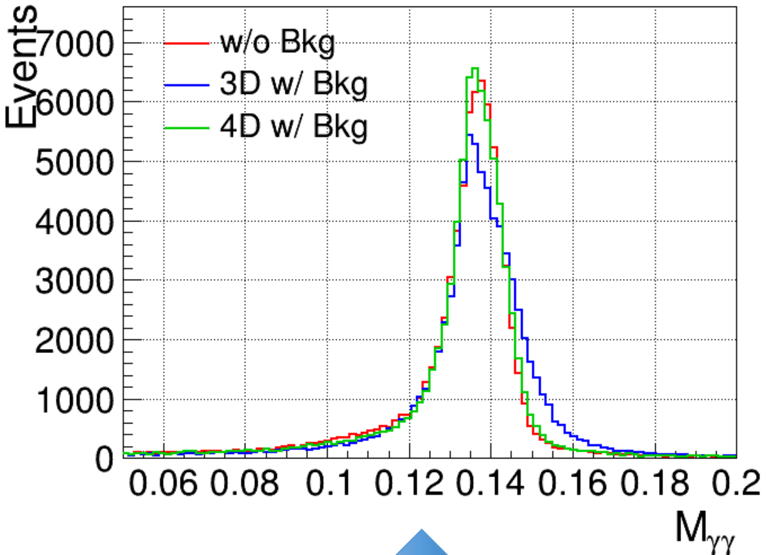


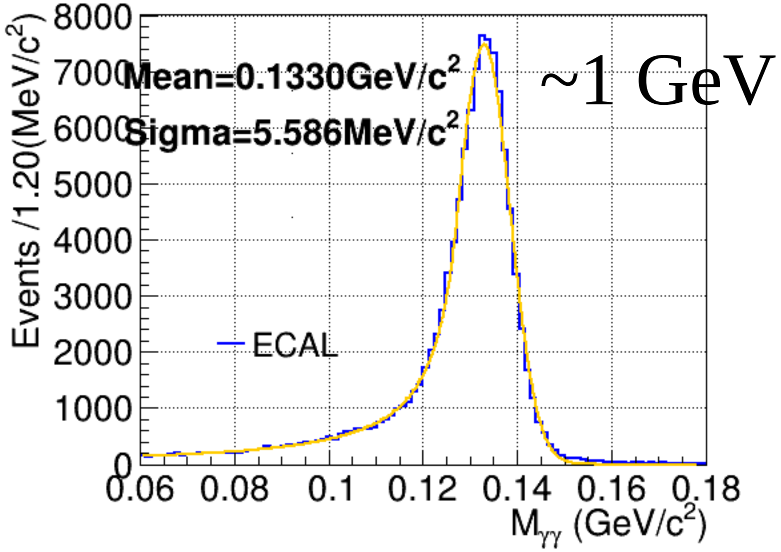
- ✓ σ_{pos} = 4.8 mm @ 1 GeV
- ✓ With background
- ✓ Meet the requirement

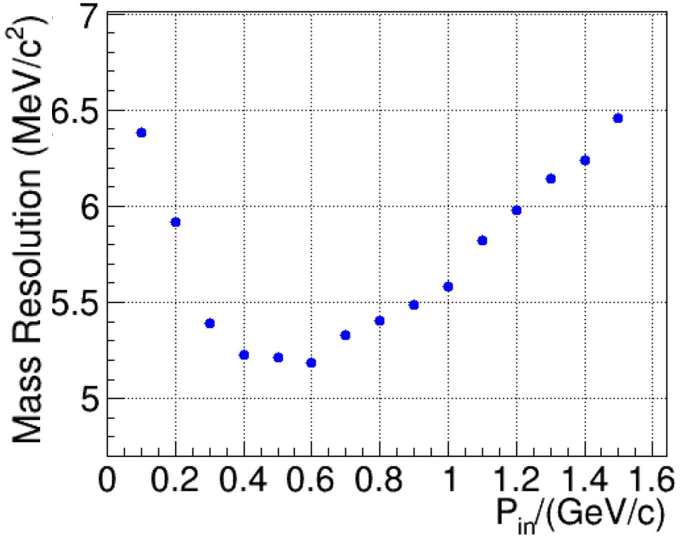
Reconstruction Performance

π^0 Reconstruction

$$m_{\pi^0} = \sqrt{2E_1E_2(1-cos\alpha)}$$







$-10ns < T_{seed} < 10ns$

π^0 with the minimum ΔM

Truth matched π^0

- ✓ $J/\psi \to \rho\pi$ sample
- ✓ σ_{π^0} = 5.6 MeV/c² @ 1 GeV
- ✓ With background

Summary

- The software of ECAL has been established based on OSCAR
- The simulated performance of ECAL meets requirements with the background concerned
  - ✓ Energy measurement with 2.03% @ 1 GeV
  - ✓ Time measurement with 130 ps @ 1 GeV
  - ✓ Position measurement with 5.3 mm @ 1 GeV
  - ✓ π^0 mass measurement with 5.6 MeV/c^2 @ ~1GeV

Prospect:

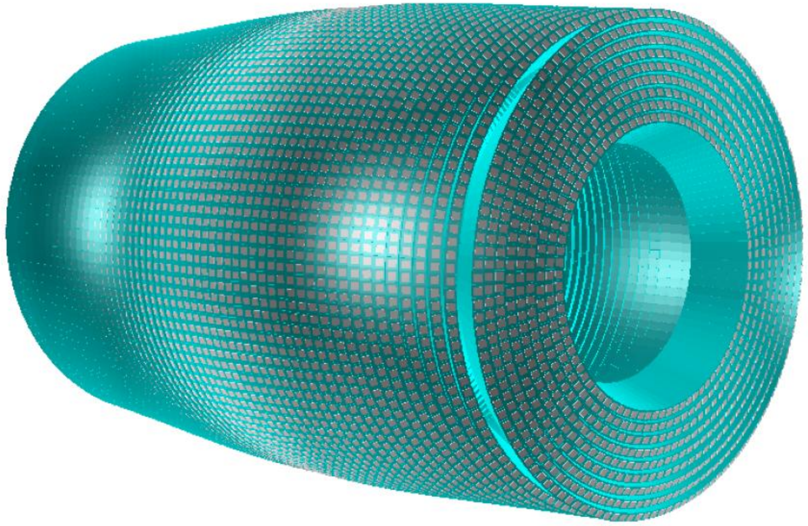
- Verify and update digitization to align with beam test data
- dE/dx for simulation to describe hardon waveform

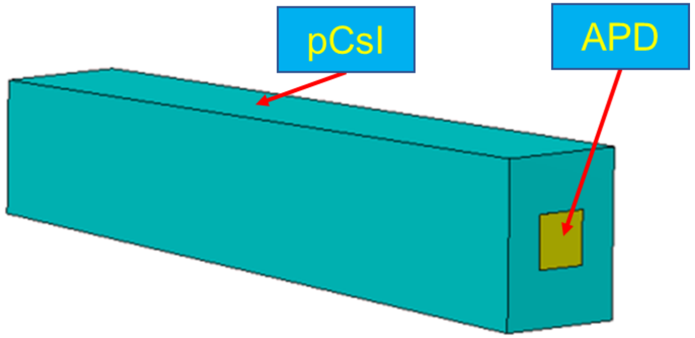
Thanks for your listening!

Back up

ECAL Design

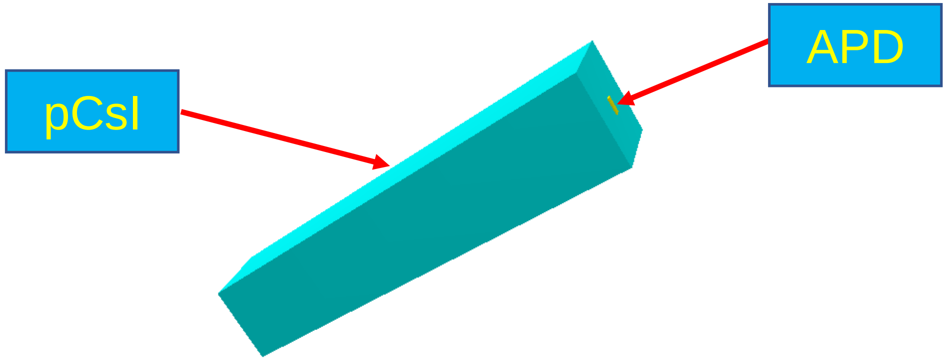
- Total absorption calorimeter
  - Barrel: $53 \times 132 = 6996$
  - Endcap: $3 \times (85 + 105 + 130) \times 2 = 1920$
  - Crystal Size:
    - $5 \times 5 \times 28(15X_0)\ \mathrm{cm}^3$
- Sensitive Unit
  - Pure CsI (pCsI) crystal
    - Fast decay time (~30ns)
    - Good radiation hardness
    - Low light yield
  - Avalanche photodiode (APD)
    - Short wavelength type
    - Large area $(10 \times 10\ mm^2 \times 4)$





ECAL Design —— Sensitive Unit

- **Pure CsI crystal + APD photo-device**
  - Pure CsI (pCsI) crystal
    - ✓ Fast decay time
    - ✓ Good radiation hardness
    - ✓ Low light yield
  - Crystal Size:
    - ✓ Total radiation length
      $15\ X_0$ (28 cm)
    - ✓ End face size
      front end: $\sim 5 \times 5\ cm^2$
      back end: $\sim 6.5 \times 6.5\ cm^2$
  - Avalanche photodiode (APD)
    - ✓ Short wavelength type
    - ✓ Large area ($10 \times 10\ mm^2 \times 4$)

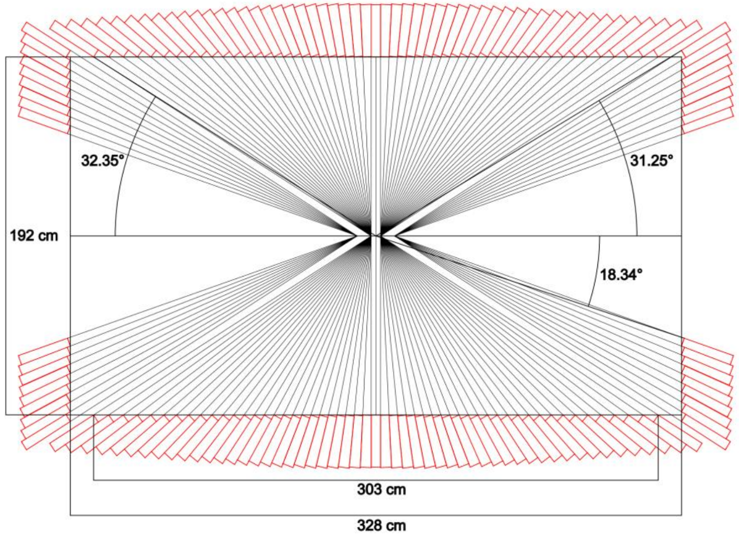


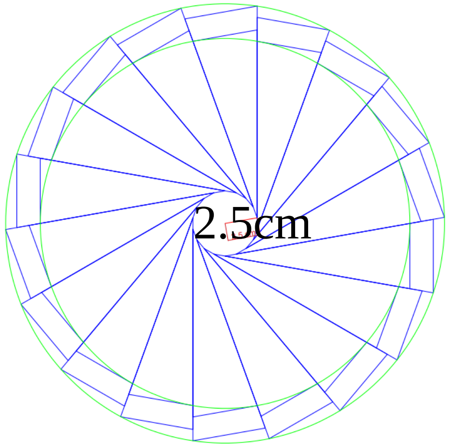
ECAL pCsI crystal unit

Crystal	Pure CsI
Density (g/cm^3)	4.51
Melting Point (°C)	621
Radiation Length (cm)	1.86
Moliere Radius (cm)	3.57
Refractive index	1.95
Hygroscopicity	Slight
Luminescence (nm)	310
Decay time (ns)	30 6
Light yield (%)	3.6 1.1
Dose rate dependent	No
D(LY)/dT (%/°C)	-1.4
Experiment	KTeV Mu2e

ECAL Setup

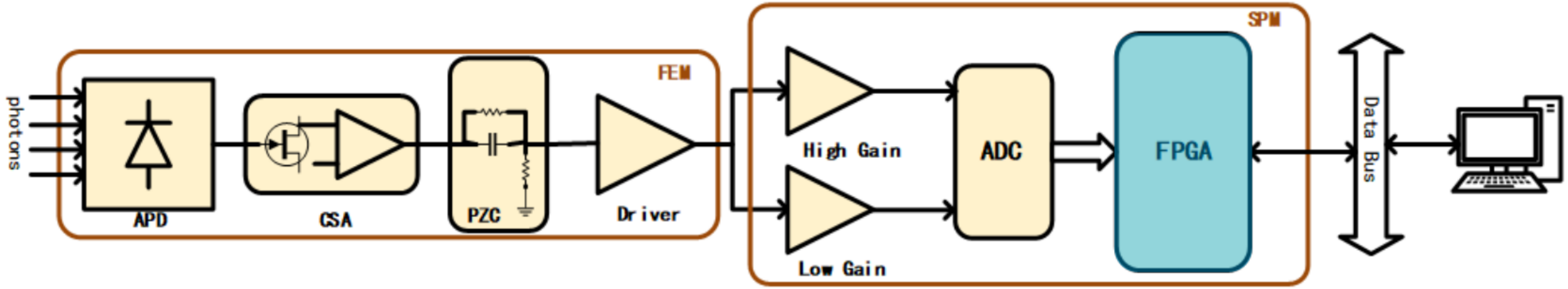
- "Dead Material"
  - 150-μm Teflon reflective film
  - 75-μm polyethylene insulating film
  - 75-μm Al electrostatic shielding film
  - No supporting material
- Light Yield: 300 p.e./MeV
- Light Collection Non-uniformity
  - collection efficiency: $\epsilon(l) = 95\% + l/L \times 5\%$
- $\sigma_{noise} = 0.4$ MeV
- $E_{hit\_thres} = 2.0$ MeV
- Secondary Particles Hit APD





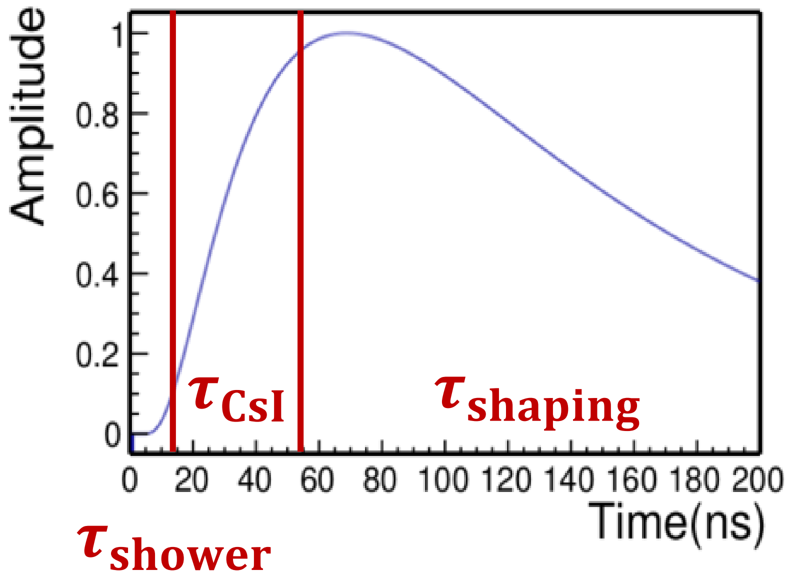
ECAL Electronics

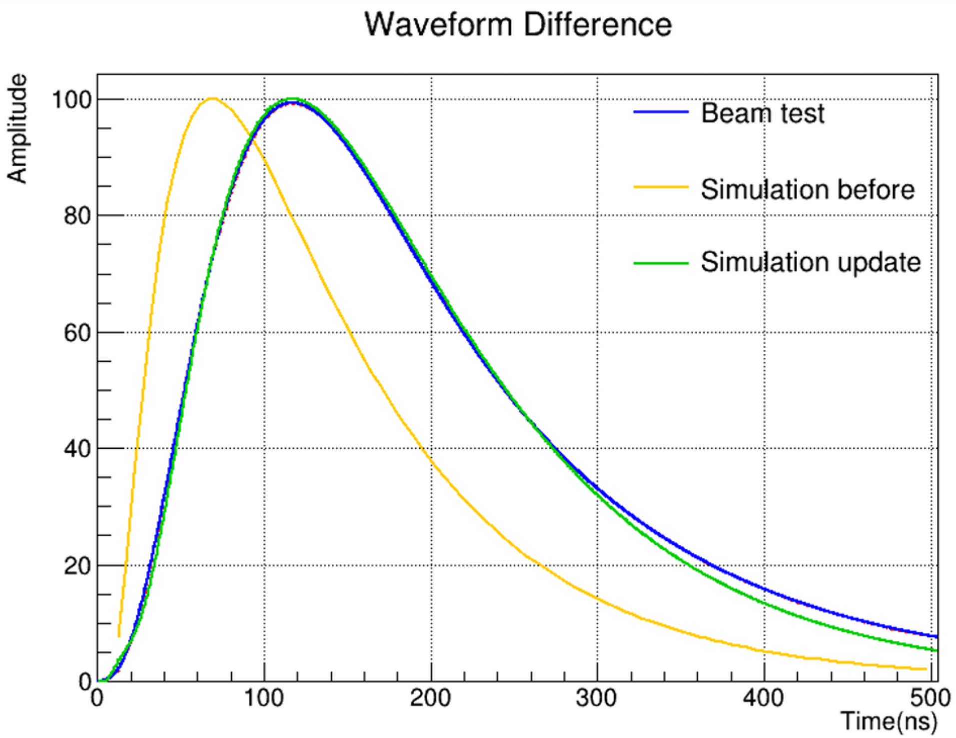
- Charge sensitive amplifier and pole-zero cancellation with shaping time 100 ns
- High and Low Gain (300 MeV/3000MeV)
- 16-bit ADC (~16000 channels) with 80 MHz sampling rate



Digitization Algorithm

Waveform check





- $\tau_{\text{shaping}} = 100$ ns
- $\tau_{\text{CsI}} = 30$ ns
- $\sigma_{\text{noise}} < 1.0$ MeV
- L.Y. > 100 p.e./MeV

→

- $\tau_{\text{shaping}} = 100$ ns
- $\tau_{\text{CsI}} \approx 30$ ns + 85ns
- $\sigma_{\text{noise}} <$ **0.4** MeV
- L.Y. > **300** p.e./MeV

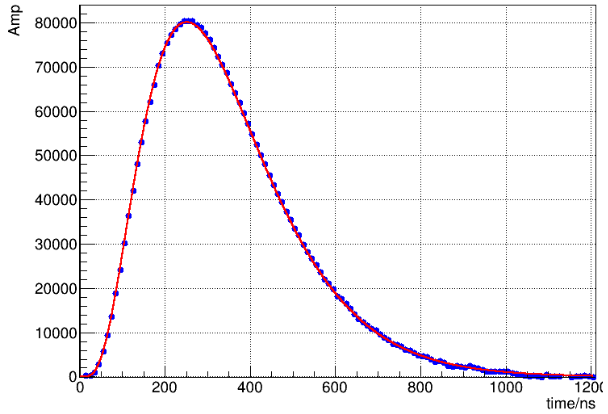
Digitization Algorithm

Waveform fitting

- Template fit to extract Amplitude and Time
  - Find the points around amplitude peak to do the template fitting
  - Template shape function: $f(t) = A \times f(t-\tau) + p$
  - $\chi^2 = \sum_{i,j}(y_i - A \cdot f(t_i - \tau) - p) \cdot S_{ij}^{-1} \cdot \left(y_j - A \cdot f\left(t_j - \tau\right) - p\right)$
  - Apply $\frac{\partial \chi^2}{\partial A} = 0, \frac{\partial \chi^2}{\partial \tau} = 0$

  Where y_i and t_i are from readout waveform; the electronics foundation p in digitization is $p = 0$; A and τ are the amplitude and time from fitting result; S_{ij} is the noise covariance matrix.

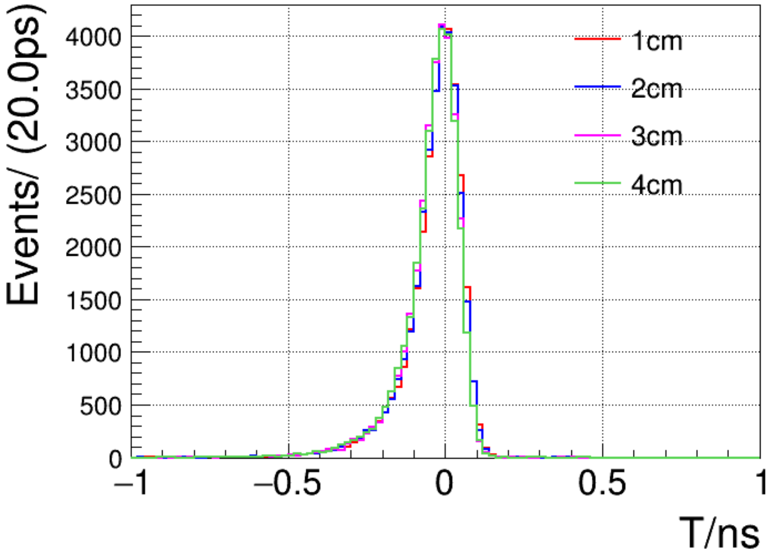
- Pipeline fit for pile-up recovery

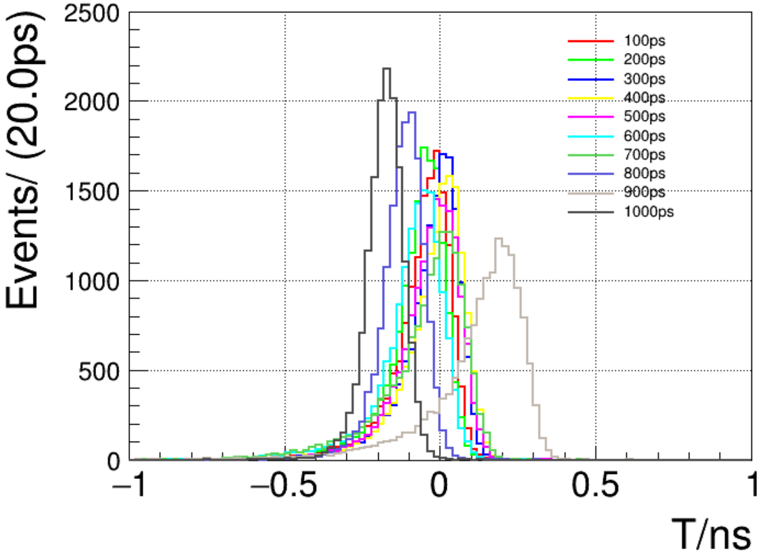


Simulation Algorithm

Data Model

- Sizes of thickness and time bin width have been optimized
  - No large difference of time distribution for different thickness
    - Consider the non-uniformity may vary in the future : 2cm
  - Different time bin width with different time resolution and central value
    - Difference in central value approximately equivalent to a shift of the template
    - Consider the resolution and the similarity to the template: 500ps

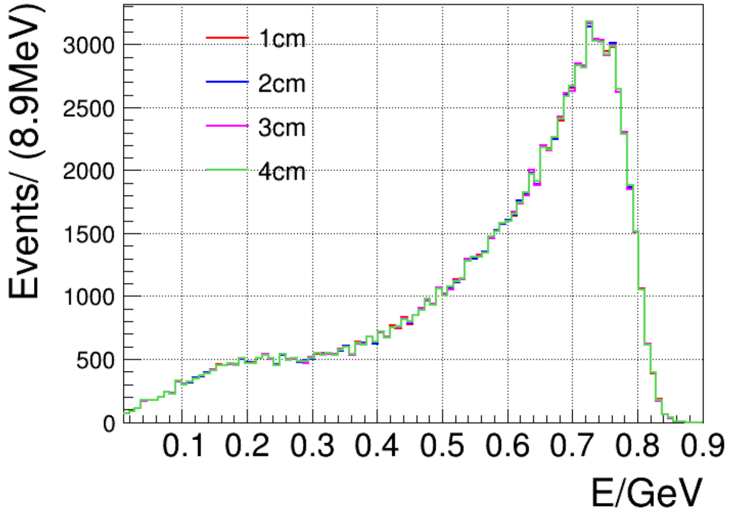


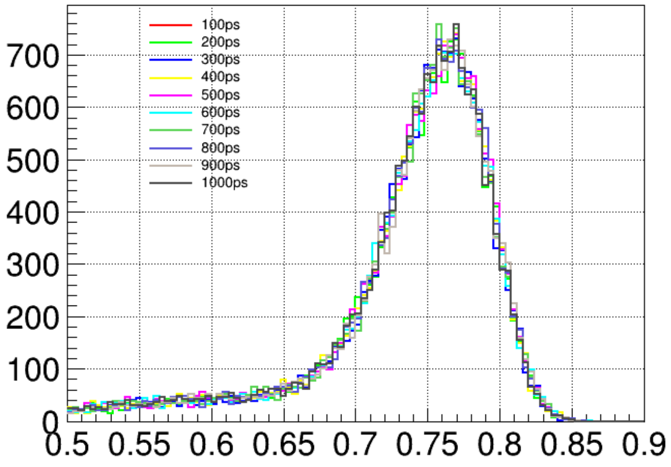


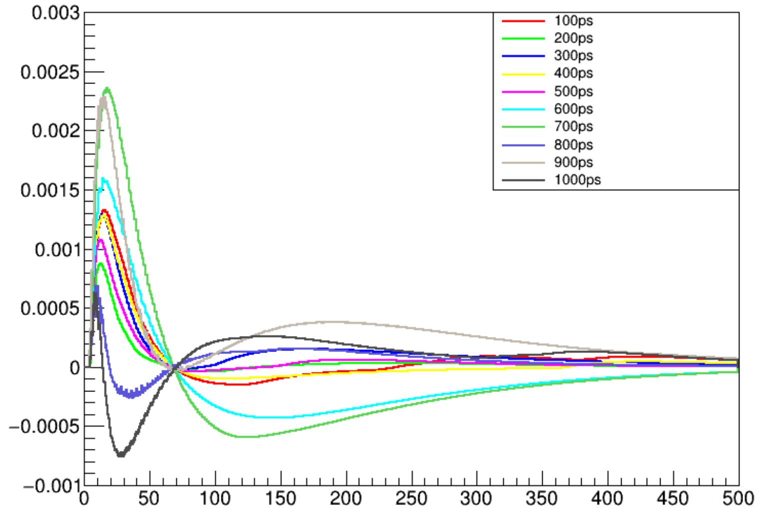
Simulation Algorithm

Data Model

- Energy distribution and comparison with template

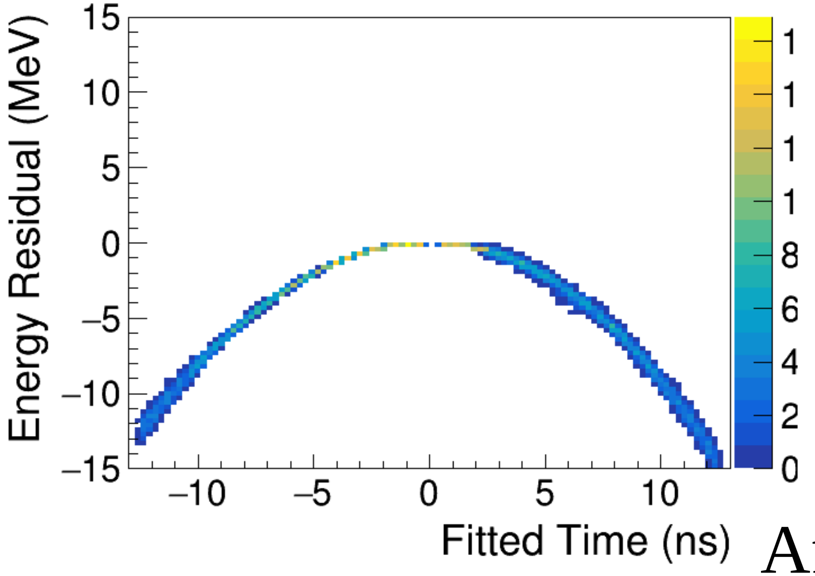


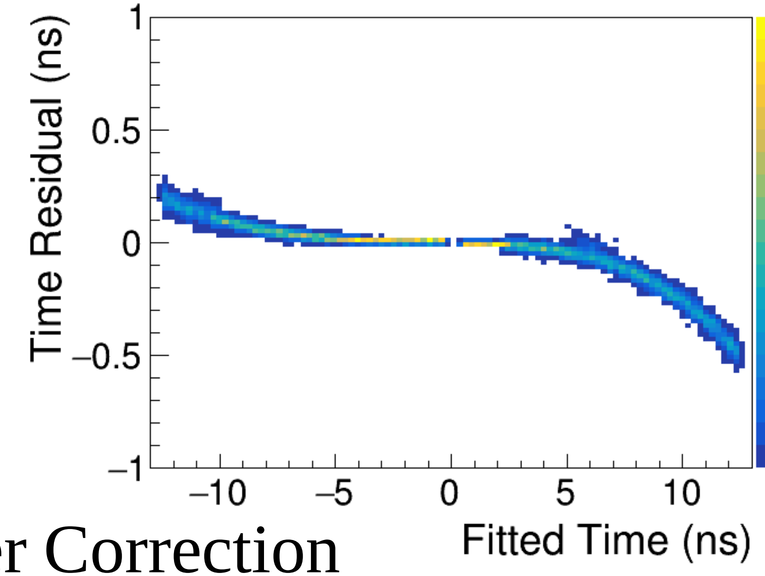




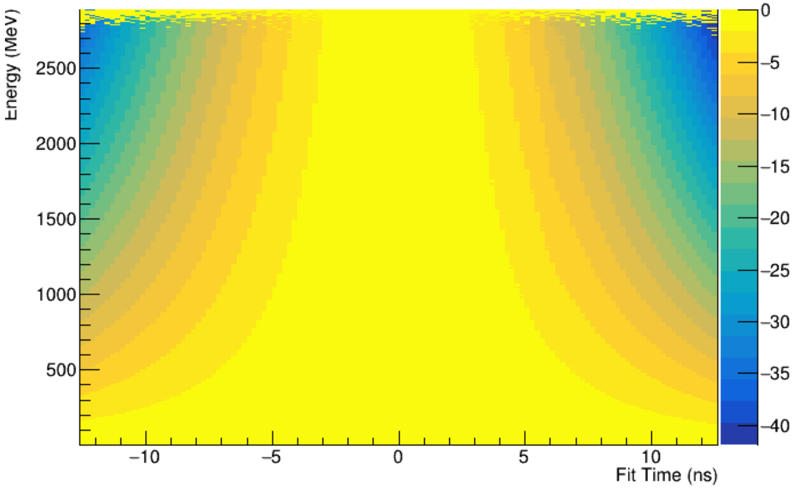
A-T Correction

Before Correction

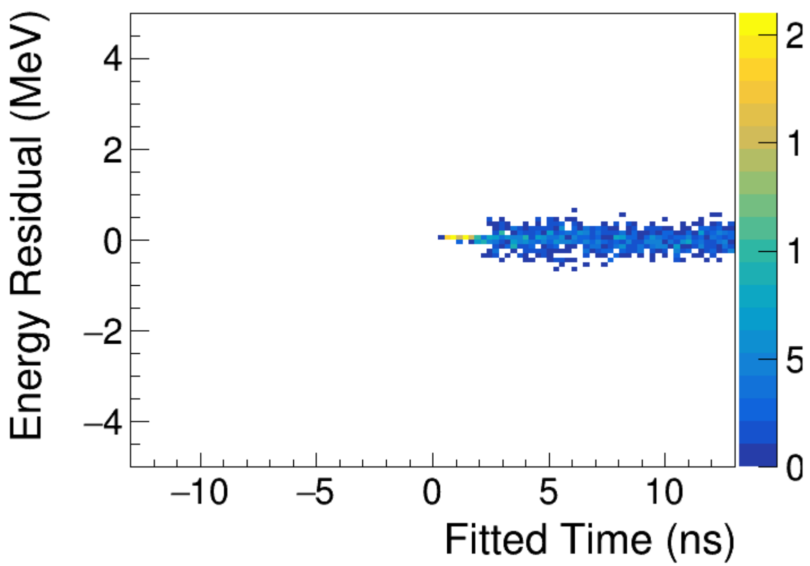


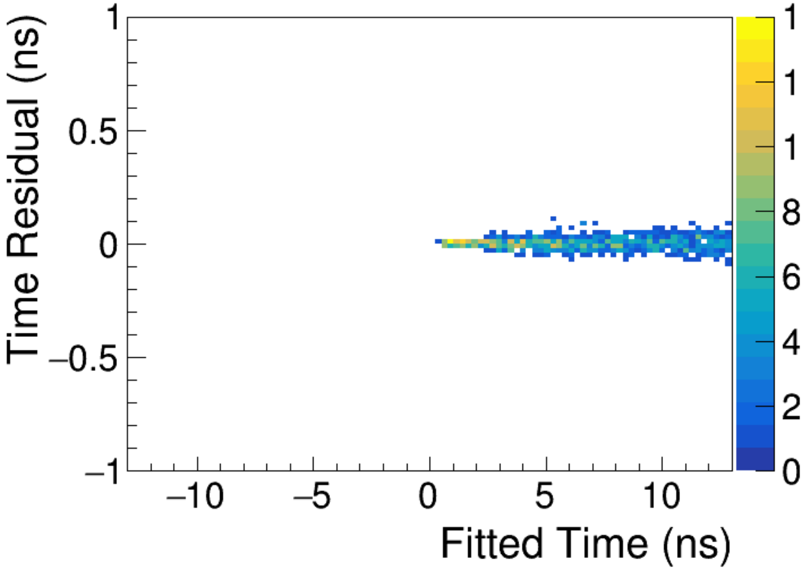


E Correction

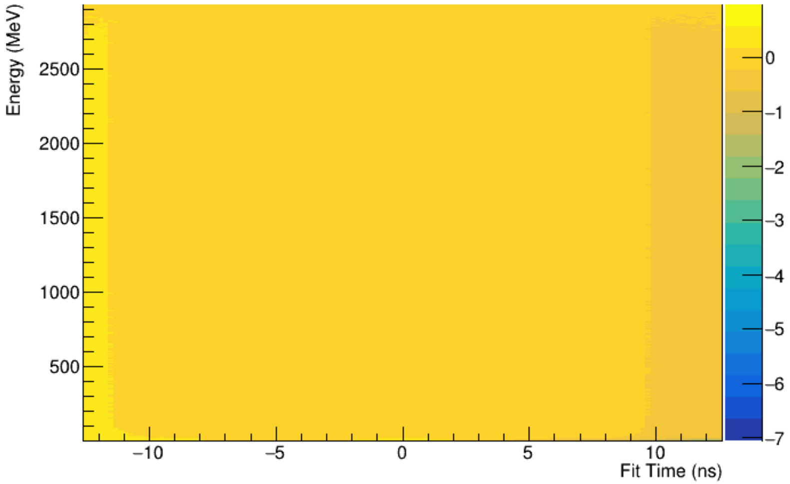


After Correction





T Correction



Template Fitting

- Template shape function： $f(t) = A \times f(t-\tau) + p$
- $\chi^2 = \sum_{i,j}(y_i - A \cdot f(t_i - \tau) - p) \cdot S_{ij}^{-1} \cdot \left(y_j - A \cdot f\left(t_j - \tau\right) - p\right)$
- Apply $\frac{\partial \chi^2}{\partial A} = 0, \frac{\partial \chi^2}{\partial \tau} = 0, \frac{\partial \chi^2}{\partial p} = 0$:

$$\begin{cases} \sum_{i,j} f_{ki} \cdot S_{ij}^{-1} \cdot \left(y_j - Af_{kj} - Bf'_{kj} - p\right) = 0 \\ \sum_{i,j} f'_{ki} \cdot S_{ij}^{-1} \cdot \left(y_j - Af_{kj} - Bf'_{kj} - p\right) = 0 \\ \sum_{i,j} 1 \cdot S_{ij}^{-1} \cdot \left(y_j - Af_{kj} - Bf'_{kj} - p\right) = 0 \end{cases}$$

$$\begin{pmatrix} \boldsymbol{F}_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k^T & \boldsymbol{F}_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k'^T & \boldsymbol{F}_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{I} \\ \boldsymbol{F}'_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k^T & \boldsymbol{F}'_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k'^T & \boldsymbol{F}'_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{I} \\ \boldsymbol{I} \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k^T & \boldsymbol{I} \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k'^T & \boldsymbol{I} \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{I} \end{pmatrix} \cdot \begin{pmatrix} A \\ B \\ p \end{pmatrix} = \begin{pmatrix} \boldsymbol{F}_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{Y} \\ \boldsymbol{F}'_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{Y} \\ \boldsymbol{I} \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{Y} \end{pmatrix}$$

$$\begin{pmatrix} A \\ B \\ p \end{pmatrix} = \begin{pmatrix} \boldsymbol{F}_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k^T & \boldsymbol{F}_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k'^T & \boldsymbol{F}_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{I} \\ \boldsymbol{F}'_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k^T & \boldsymbol{F}'_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k'^T & \boldsymbol{F}'_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{I} \\ \boldsymbol{I} \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k^T & \boldsymbol{I} \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{F}_k'^T & \boldsymbol{I} \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{I} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \boldsymbol{F}_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{Y} \\ \boldsymbol{F}'_k \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{Y} \\ \boldsymbol{I} \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{Y} \end{pmatrix}$$

Nonnegative Least Square (NNLS)

Convention:

- b: A real pulse with m points
- x: fitted amplitudes for n pulses
- A: the ith column of A represents the template for the ith pulse and of course each template has m points
- P: passive set – currently not fixed amps
- R: active set – currently fixed amplitudes

Algorithm *fnnls* **:**

Input: $\mathbf{A} \in \mathbf{R}^{m \times n}$, $\mathbf{b} \in \mathbf{R}^m$
Output: $\mathbf{x}^* \geq 0$ such that $\mathbf{x}^* = \arg\min \|\mathbf{Ax} - \mathbf{b}\|^2$.
Initialization: $P = \varnothing, R = \{1, 2, \cdots, n\}, \mathbf{x} = \mathbf{0}, \mathbf{w} = \mathbf{A}^T\mathbf{b} - (\mathbf{A}^T\mathbf{A})\mathbf{x}$
repeat

1. Proceed if $R \neq \varnothing \wedge [\max_{i \in R}(w_i) > \textit{tolerance}]$
2. $j = \arg\max_{i \in R}(w_i)$
3. Include the index j in P and remove it from R
4. $\mathbf{s}^P = [(\mathbf{A}^T\mathbf{A})^P]^{-1}(\mathbf{A}^T\mathbf{b})^P$
   4.1. Proceed if $\min(\mathbf{s}^P) \leq 0$
   4.2. $\alpha = -\min_{i \in P}[x_i/(x_i - s_i)]$
   4.3. $\mathbf{x} := \mathbf{x} + \alpha(\mathbf{s} - \mathbf{x})$
   4.4. Update R and P
   4.5. $\mathbf{s}^P = [(\mathbf{A}^T\mathbf{A})^P]^{-1}(\mathbf{A}^T\mathbf{b})^P$
   4.6. $\mathbf{s}^R = \mathbf{0}$
5. $\mathbf{x} = \mathbf{s}$
6. $\mathbf{w} = \mathbf{A}^T(\mathbf{b} - \mathbf{Ax})$

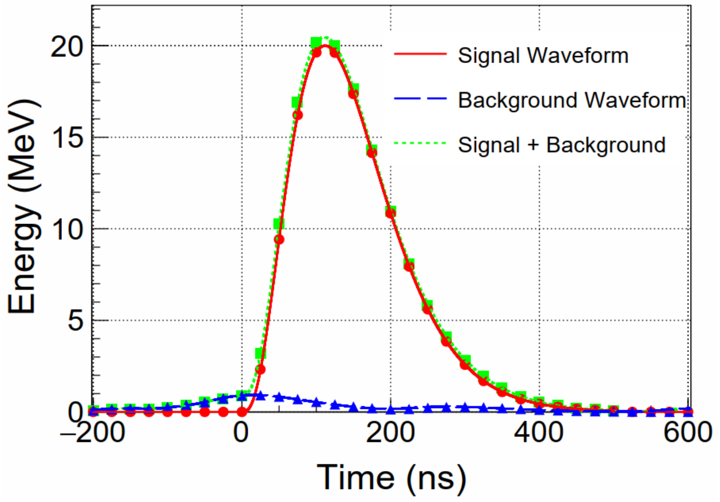
Pile-up Recovery

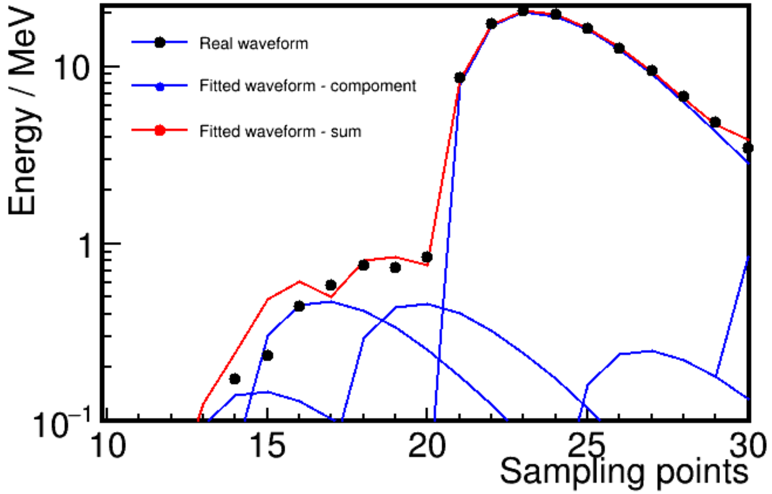
Multi-template fit

- Fit all the potential waveforms with template
- Isolate signals by time
- The fit minimizes the χ^2 defined as:

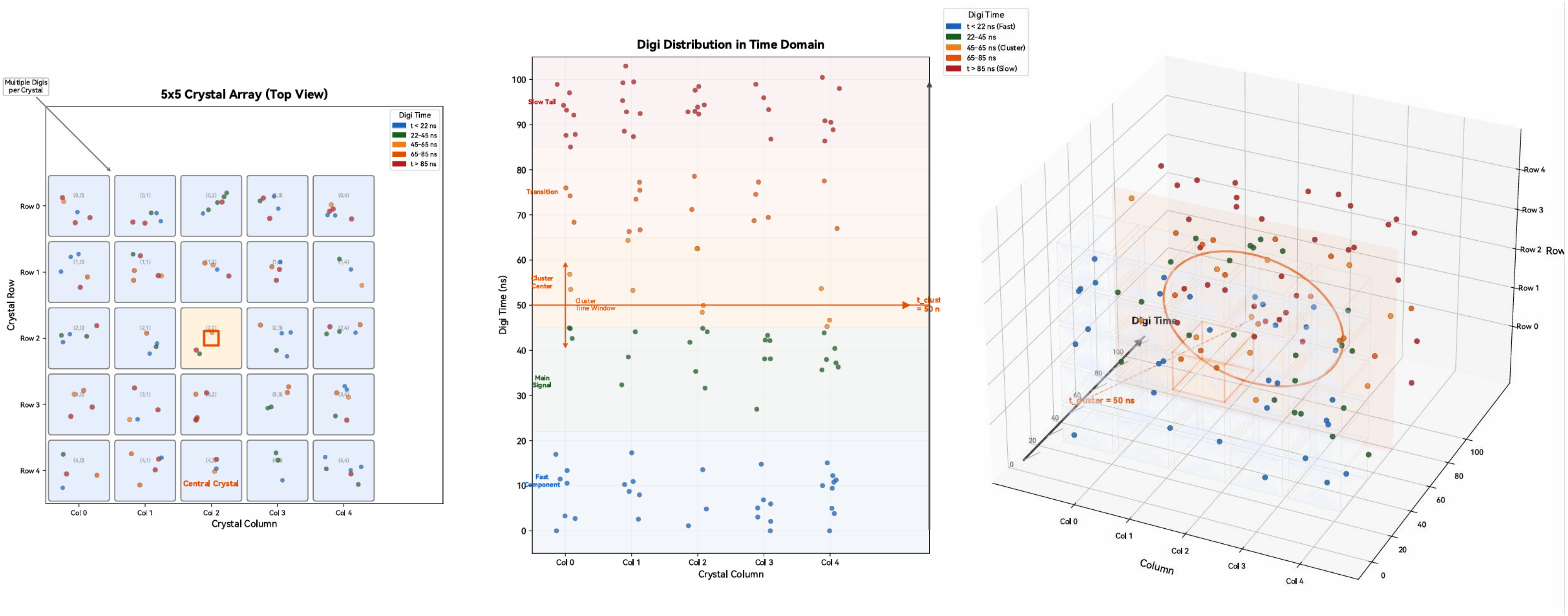
$$\chi^2 = \left(\sum_{j=1}^{N} A_j\overrightarrow{p_j} - \vec{S}\right)^T C^{-1}\left(\sum_{j=1}^{N} A_j\overrightarrow{p_j} - \vec{S}\right)$$

Where:
N is the number of templates;
vector $\vec{S}$ comprise the readout samples;
vector $\overrightarrow{p_j}$ is the waveform template;
A_j are the amplitudes, which are obtained by the fit;
C is the noise covariance matrix.



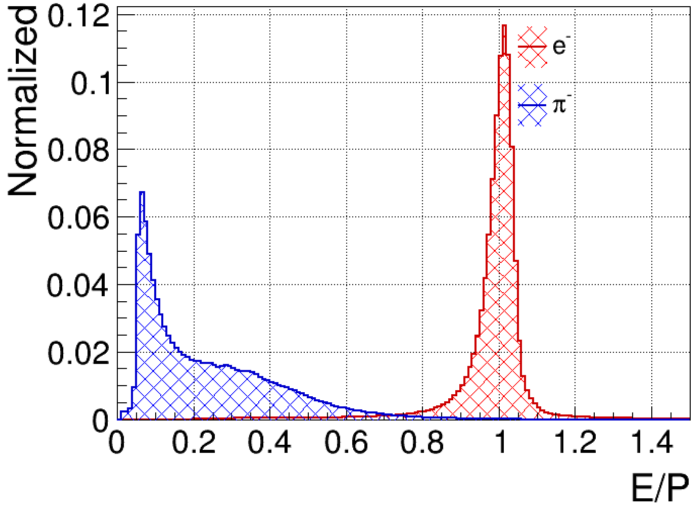


4D Reconstruction Algorithm

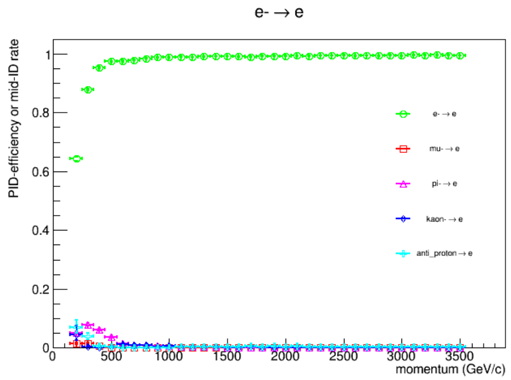


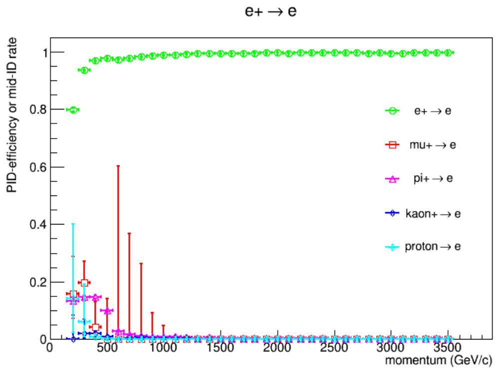
Electron PID Algorithm

- E/P ratio
  - E: Energy of ECAL Shower
  - P: Momentum from Track
  - $L_e = \ln(e) + 0.001; L_{other} = \ln(other) + 0.001$
  - Poor performance in low-momentum

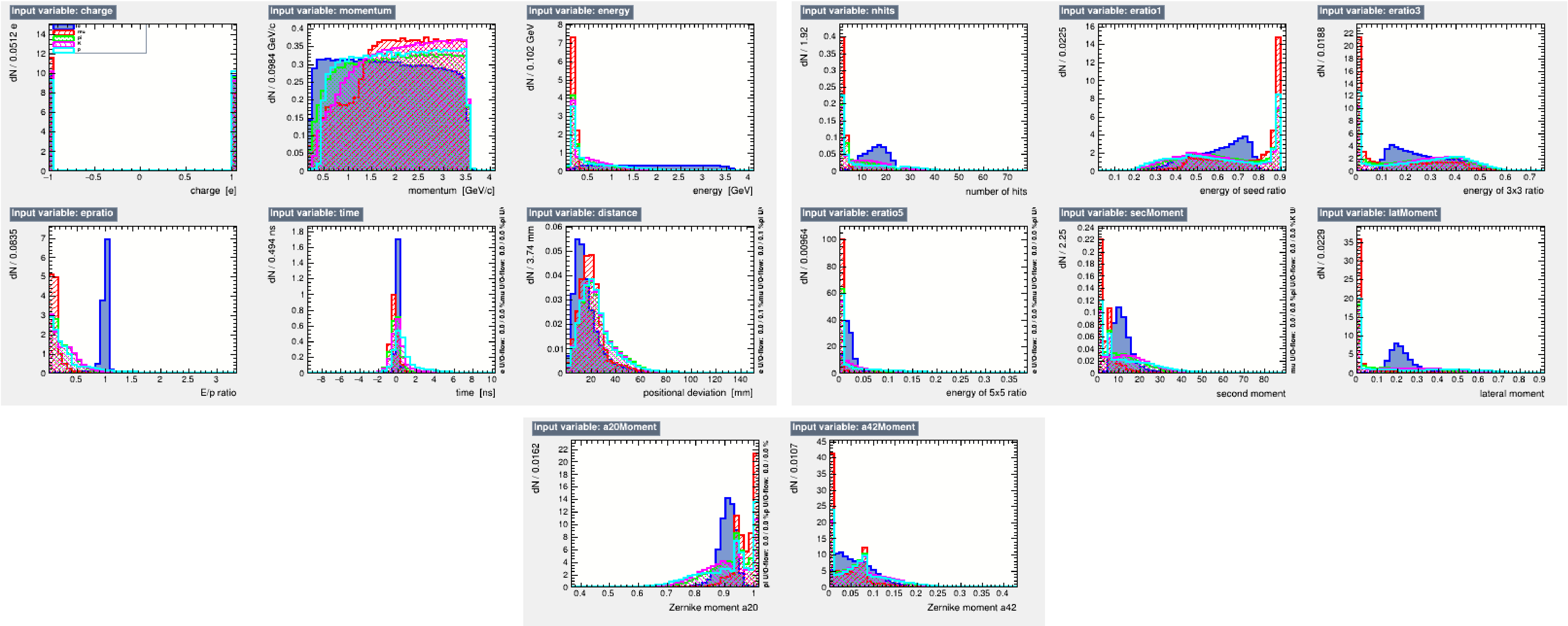


- BDT Method
  - Input variables: charge, P_{track}, E_{shower}, E/P, T_{seed}, $|R_{ext} - R_{shower}|$, N_{hits}, E_{seed}/E_{shower}, $(E_{3x3} - E_{seed})/E_{shower}$, $(E_{5x5} - E_{3x3})/E_{shower}$, second moment, lateral moment, A_{20} moment, A_{42} moment
  - Distinguishing electrons from other charged hadrons (or muon)
  - BDT Method used



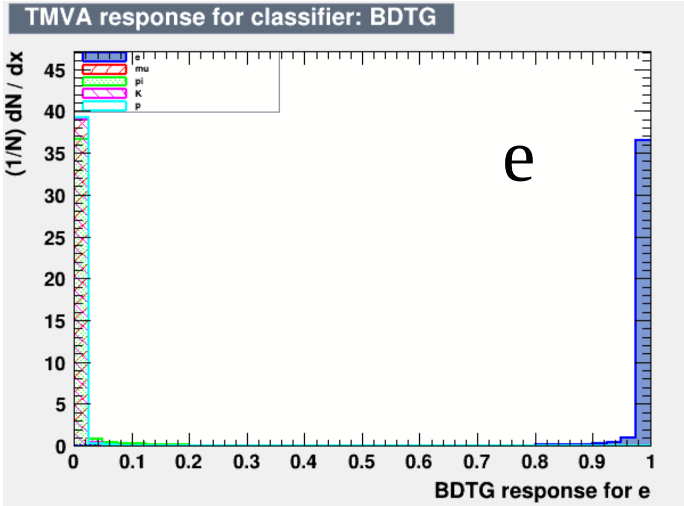


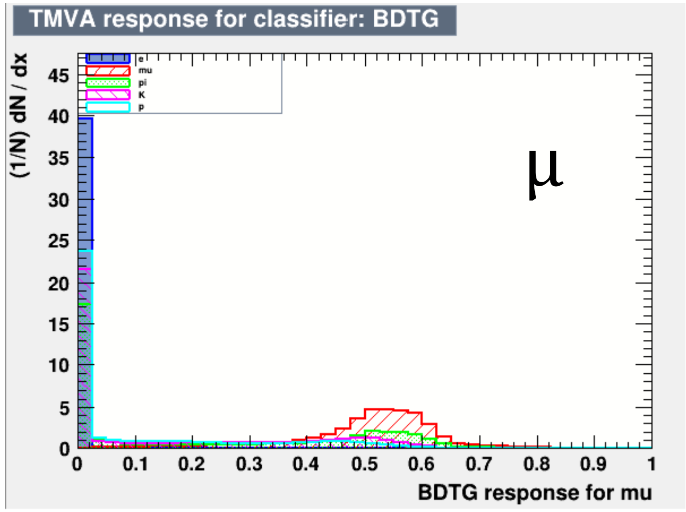
BDT Input

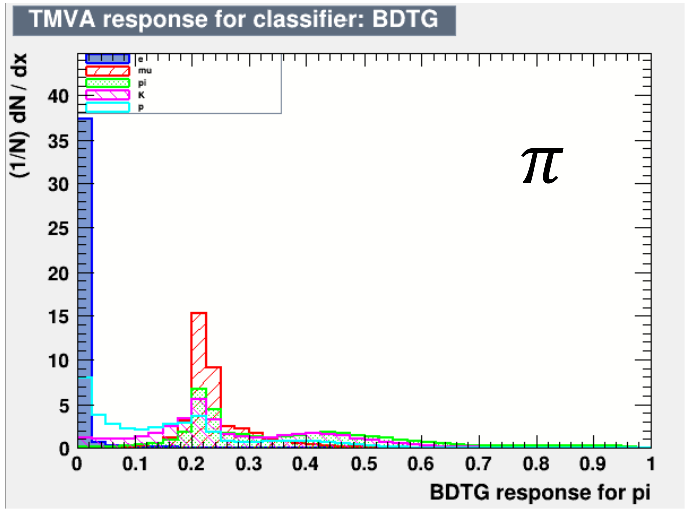


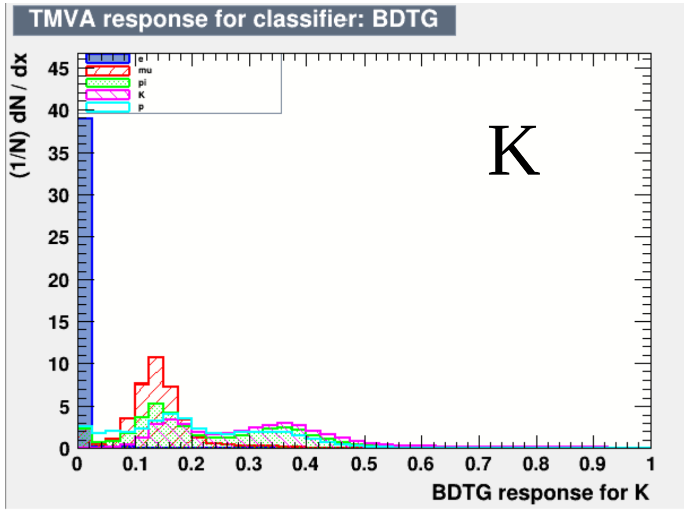
Input variable: charge
dN / 0.0512 e
charge [e]
Input variable: momentum
dN / 0.0984 GeV/c
momentum [GeV/c]
Input variable: energy
dN / 0.102 GeV
energy [GeV]
Input variable: nhits
dN / 1.92
number of hits
Input variable: eratio1
dN / 0.0225
energy of seed ratio
Input variable: eratio3
dN / 0.0188
energy of 3x3 ratio
Input variable: epratio
dN / 0.0835
E/p ratio
Input variable: time
dN / 0.494 ns
time [ns]
Input variable: distance
dN / 3.74 mm
positional deviation [mm]
Input variable: eratio5
dN / 0.00964
energy of 5x5 ratio
Input variable: secMoment
dN / 2.25
second moment
Input variable: latMoment
dN / 0.0229
lateral moment
Input variable: a20Moment
dN / 0.0162
Zernike moment a20
Input variable: a42Moment
dN / 0.0107
Zernike moment a42

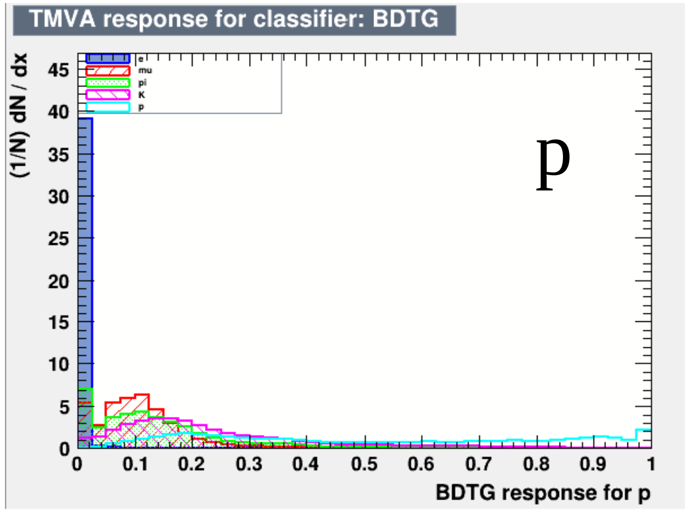
BDT Output


TMVA response for classifier: BDTG
(1/N) dN / dx
e
mu
pi
K
p
BDTG response for e
e


TMVA response for classifier: BDTG
(1/N) dN / dx
e
mu
pi
K
p
BDTG response for mu
μ


TMVA response for classifier: BDTG
(1/N) dN / dx
e
mu
pi
K
p
BDTG response for pi
π


TMVA response for classifier: BDTG
(1/N) dN / dx
e
mu
pi
K
p
BDTG response for K
K


TMVA response for classifier: BDTG
(1/N) dN / dx
e
mu
pi
K
p
BDTG response for p
p

- 桶部角分布相同
  - 共53圈（左右各加一圈）
- 端盖重新排列
  - 最外层与桶部最外层平行
  - 间距4cm
  - 散焦点10cm
  - 尺寸变小
  - 共8→9圈
- 端盖边界1640mm
- 体积 6.47 m^3
- 节约 11%