

# A Cluster Counting Technique Study through Beam Tests

Walaa Elmetenawee (INFN Bari (IT))

on behalf of cluster counting test beam team

ICHEP 2024 Conference - 42<sup>nd</sup> International Conference  
on High Energy Physics

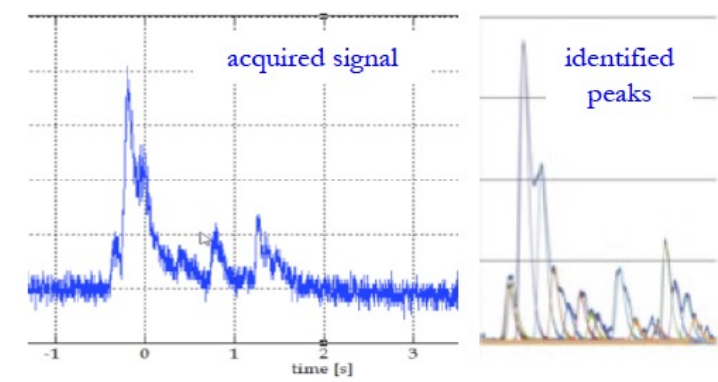
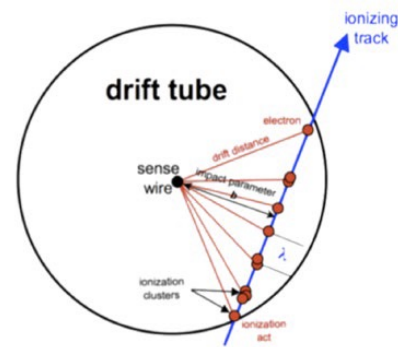
17-24 Jul 2024



ICHEP 2024 | PRAGUE

# Cluster Counting Technique

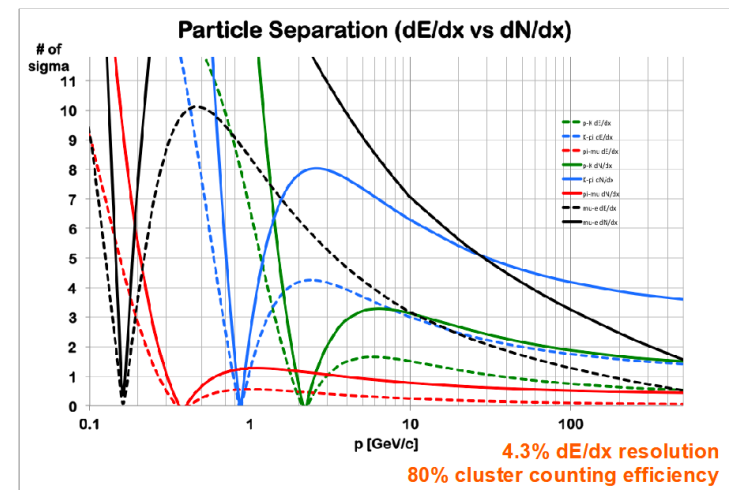
- **Principle:** In He based gas mixtures the signals from each ionization act can be spread in time to few ns. With the help of a fast read-out electronics they can be identified efficiently.
- By counting the number of ionization acts per unit length ( $dN/dx$ ), it is possible to identify the particles (P.Id.) with a better resolution w.r.t the  $dE/dx$  method.



▶ Backup

**dE/dx**  
Truncated mean cut (70-80%) reduces the amount of collected information.  $n = 112$  and a 2m track at 1 atm give  $\sigma \approx 4.3\%$

**$dN_{cl}/dx$**   
 $\delta_{cl} = 12.5/cm$  for He/iC4H10 = 90/10 and a 2m track give  $\sigma \approx 2.0\%$



- **Analytic calculations:** Expected excellent  $K/\pi$  separation over the entire range except  $0.85 < p < 1.05$  GeV (blue lines).
- Despite the fact that the Garfield++ model in GEANT4 reproduces reasonably well the Garfield++ predictions, why particle separation, both with  $dE/dx$  and with  $dN_{cl}/dx$ , in GEANT4 is considerably worse than in Garfield++?
- Despite a higher value of the  $dN_{cl}/dx$  Fermi plateau with respect to  $dE/dx$ , why this is reached at lower values of  $\beta\gamma$  with a steeper slope?

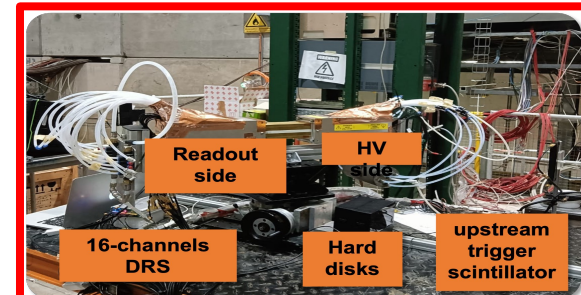
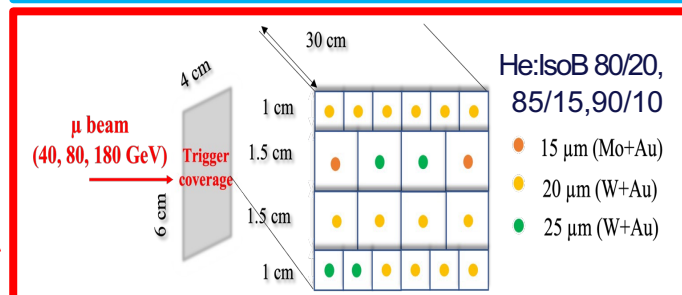
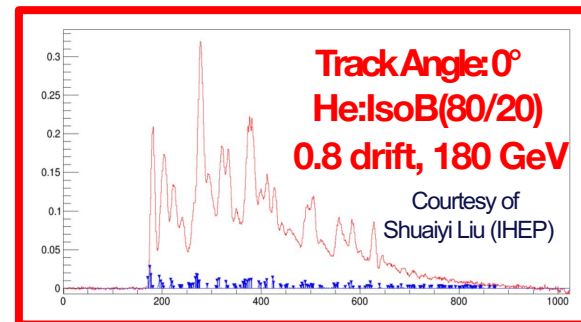
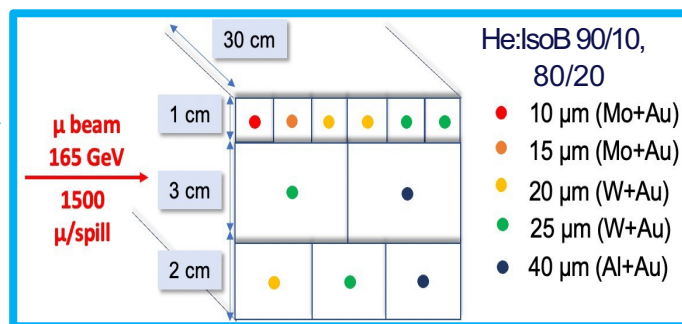
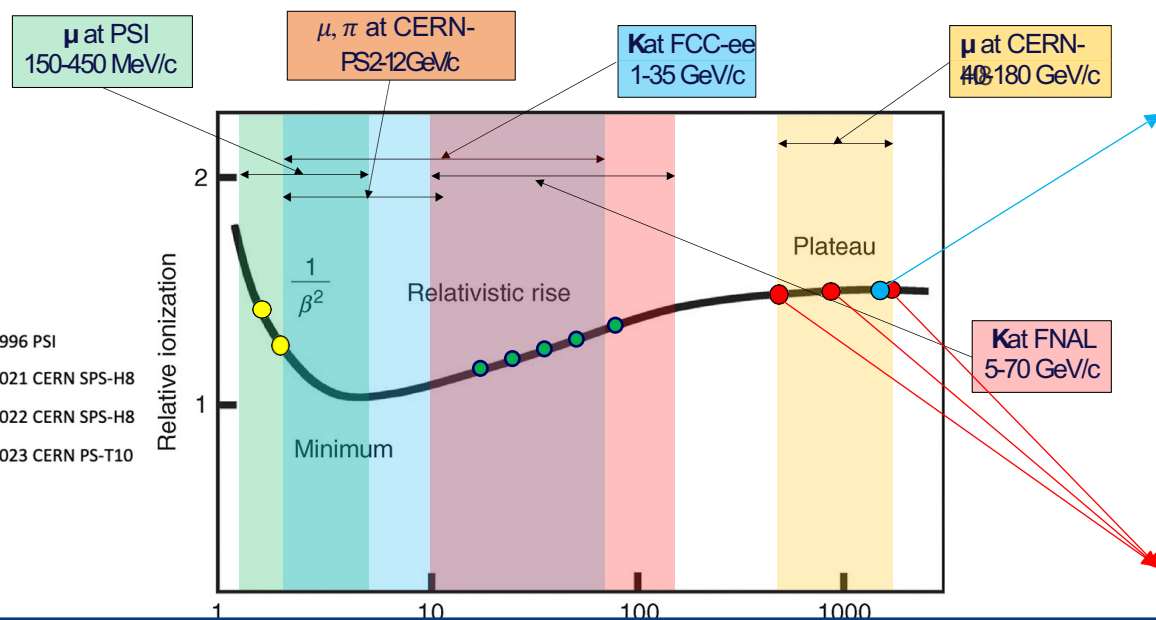
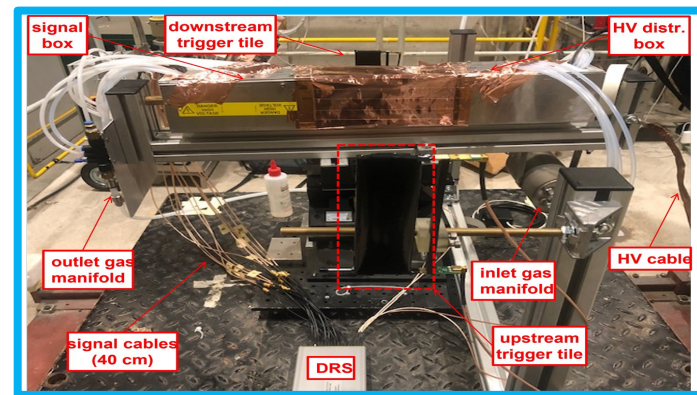
▶ Backup

finding answers by using real data from beam tests!!

# Main Beam Test setup & goals

## Beam tests to experimentally assess and optimize the performance of the cluster counting/timing techniques.

- Two muon beam tests performed at CERN H8 ( $\beta\gamma > 400$ ) in Nov. 2021 and July 2022.
- Two muon beam tests performed at CERN T10 in Jul 2023 and Jul 2024 (**still ongoing**) using  $\mu$  beam (1-12 GeV).
- Another test is planned to be done at FNAL-MT6 with  $\pi$  and K ( $\beta\gamma = 10-140$ ) to fully exploit the relativistic rise.





## ❑ Finding peaks algorithms:

- To accurately identify electron peaks, we have developed and tested Three distinct algorithms:

- ✓ Derivative Algorithm (DERIV)
- ✓ Running Template Algorithm (RTA)
- ✓ NN- based approach (developed by IHEP)

Today talk

▶ Backup

## ❑ Clusterization:

- Merging of electron peaks in consecutive bins in a single electron to reduce fake electrons counting.

## ❑ Different scans done:

- Using the test beam data we evaluated the performance of our algorithms across various conditions: gas mixture, gain, geometrical configuration (cell size, sense wires size), sampling rate, HV, and track angle.

## ❑ Resolution study: $dN/dx$ vs $dE/dx$ :

- Investigated the resolution of the number of detected clusters per unit length ( $dN/dx$ ) versus the energy loss per unit length ( $dE/dx$ ).

## Documentation: Analysis Note is done!

Beam Test Results on Cluster Counting for IDEA  
Drift Chamber

Alessandro Corvaglia<sup>1</sup>, Brunella D'Anzi<sup>2,3\*</sup>,  
Nicola De Filippis<sup>2,4\*</sup>, Francesco De Santis<sup>1</sup>,  
Walaa Elmetenawee<sup>2\*</sup>, Francesco Grancagnolo<sup>1\*</sup>,  
Edoardo Gorini<sup>1</sup>, Marcello Maggi<sup>2</sup>, Alessandro Miccoli<sup>1</sup>,  
Magdy Louka<sup>2,3</sup>, Marco Panareo<sup>1</sup>, Margherita Primavera<sup>1</sup>,  
Francesco Procacci<sup>2</sup>, Andrea Ventura<sup>1</sup>

<sup>1</sup>Istituto Nazionale di Fisica Nucleare Sezione di Lecce, Via Arnesano, 73100, Lecce, Italy.

<sup>2</sup> Istituto Nazionale di Fisica Nucleare Sezione di Bari, Via Giovanni Amendola, 173, 70126, Bari, Italy.

<sup>3</sup> Dipartimento di Fisica, Universit'a di Bari Aldo Moro, Via E. Orabona 4, 70126 Bari, Italy.

<sup>4</sup> Politecnico di Bari, Via E. Orabona 4, 70126 Bari, Italy.

\*Corresponding author(s). E-mail(s): [brunella.danzi@ba.infn.it](mailto:brunella.danzi@ba.infn.it);  
[nicola.defilippis@ba.infn.it](mailto:nicola.defilippis@ba.infn.it); [walaa.elmetenawee@ba.infn.it](mailto:walaa.elmetenawee@ba.infn.it);  
[franco.grancagnolo@le.infn.it](mailto:franco.grancagnolo@le.infn.it);

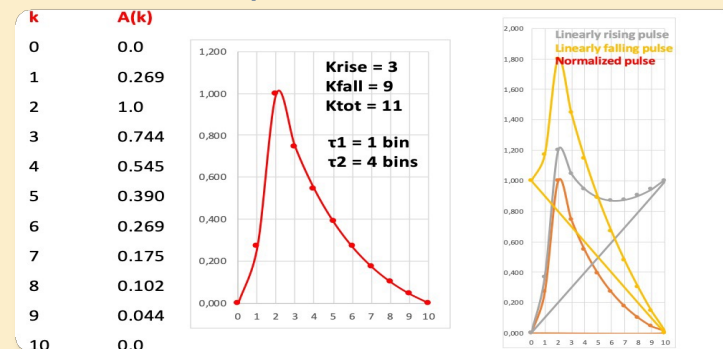
Contributing authors: [alessandro.corvaglia@le.infn.it](mailto:alessandro.corvaglia@le.infn.it);  
[francesco.desantis@le.infn.it](mailto:francesco.desantis@le.infn.it); [edoardo.gorini@le.infn.it](mailto:edoardo.gorini@le.infn.it);  
[marcello.maggi@ba.infn.it](mailto:marcello.maggi@ba.infn.it); [alessandro.miccoli@le.infn.it](mailto:alessandro.miccoli@le.infn.it);  
[magdy.louka@ba.infn.it](mailto:magdy.louka@ba.infn.it); [marco.panareo@le.infn.it](mailto:marco.panareo@le.infn.it);  
[margherita.primavera@le.infn.it](mailto:margherita.primavera@le.infn.it); [f.procacci@phd.poliba.it](mailto:f.procacci@phd.poliba.it);  
[andrea.ventura@le.infn.it](mailto:andrea.ventura@le.infn.it);

## Derivative Algorithm (DERIV)

- Compute the first and second derivative from the amplitude average over two times the timing resolution and require that, at the peak candidate position, they are less than a r.m.s. signal-related small quantity and they increase (decrease) before (after) the peak candidate position of a r.m.s. signal-related small quantity.
- Require that the amplitude at the peak candidate position is greater than a r.m.s. signal-related small quantity and the amplitude difference among the peak candidate and the previous (next) signal amplitude is greater (less) than a r.m.s. signal-related small quantity.
- NOTE: r.m.s. is a measurements of the noise level in the analog signal from first bins.

## Running Template Algorithm (RTA)

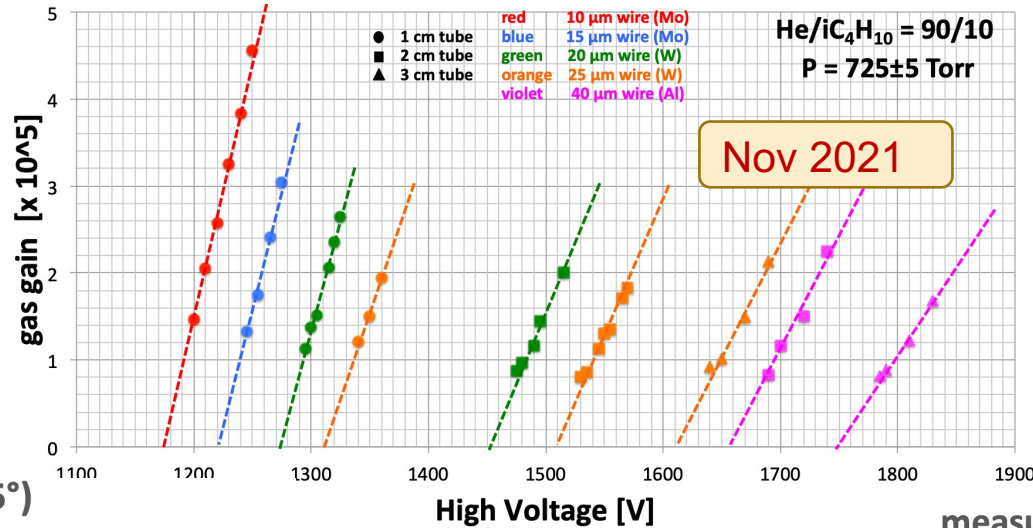
- Define an electron pulse template based on experimental data.
- Raising and falling exponential over a fixed number of bins (Ktot).
- Digitize it (A(k)) according to the data sampling rate.
- The algorithm scan the wave form and run over Ktot bins by comparing it to the subtracted and normalized data (build a sort of  $\chi^2$ ).
- Define a cut on  $\chi^2$ .
- Subtract the found peak to the signal spectrum.
- Iterate the search.
- Stop when no new peak is found.



# Gas gain scan

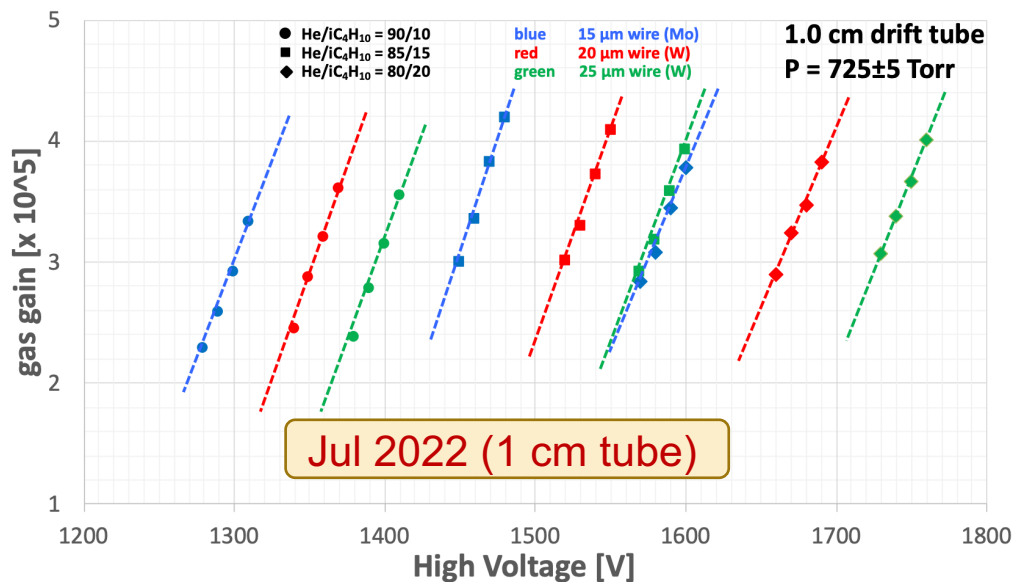
The range of gas gain, independently of the drift tube configuration (drift length, sense wire diameter, gas mixture), lies within  $1 \times 10^5$  and  $5 \times 10^5$ .

measured gas gain vs HV (normal incidence)

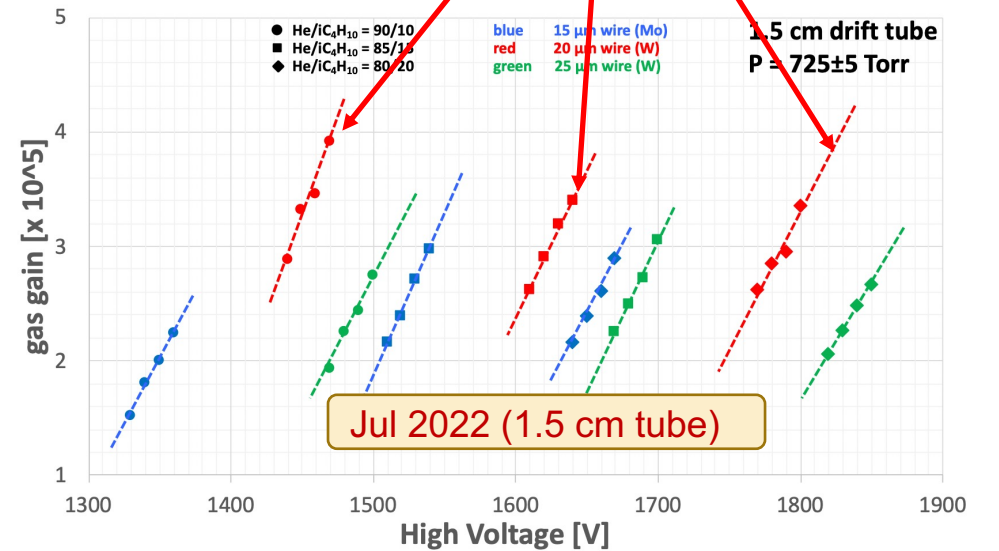


20μm wire excluded from physical quantities mean computation

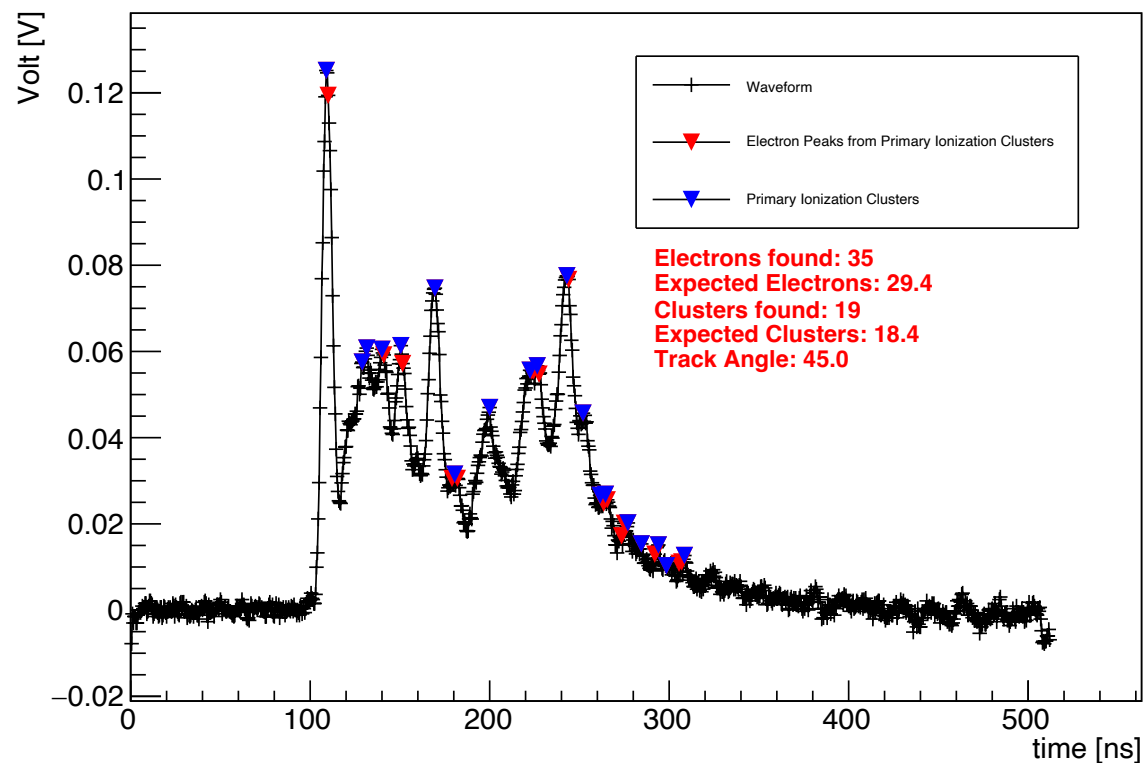
measured gas gain vs HV (45°)



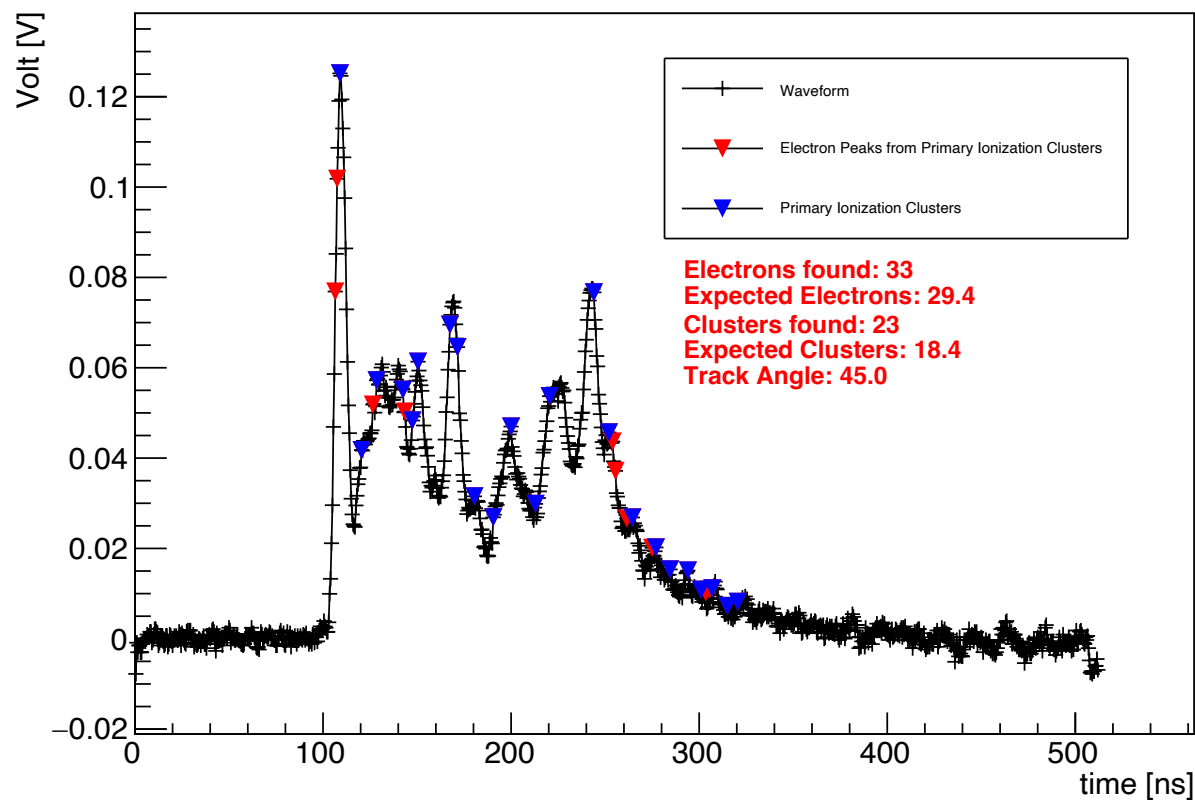
measured gas gain vs HV (45°)



## Peaks found by the RTA algorithm

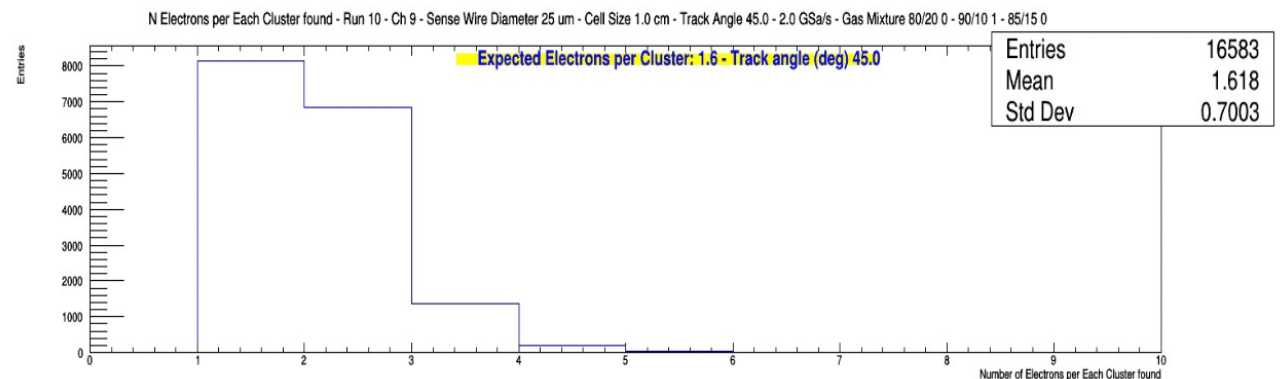
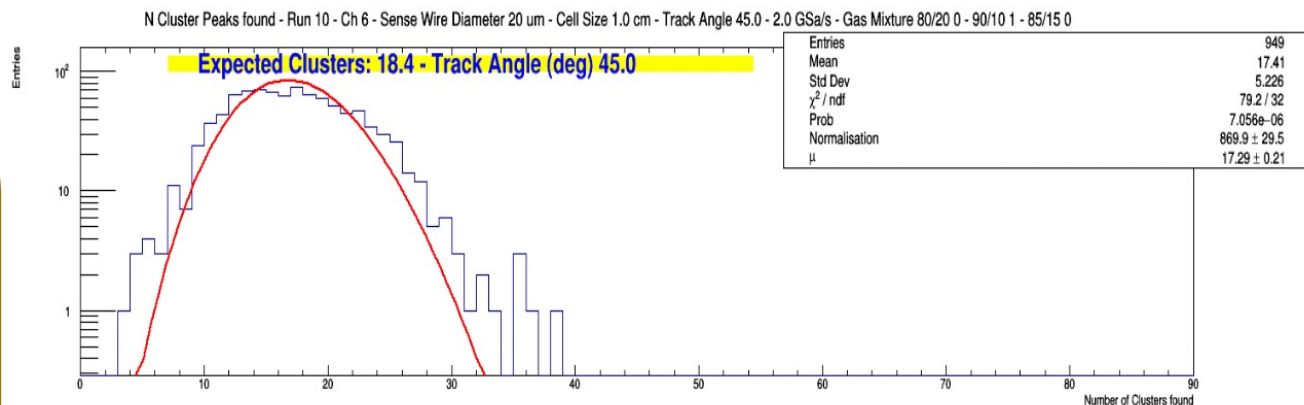


## Peaks found by the DERV algorithm



## Clusterization algorithm

- Merging of electron peaks in consecutive bins in a single electron to reduce fake electrons counting.
- Contiguous electrons peaks which are compatible with the electrons' diffusion time (it has a  $\sim\sqrt{t}ElectronPeak$  dependence, different for each gas mixture) must be considered belonging to the same ionization cluster. For them, a counter for electrons per each cluster is incremented.
- Position and amplitude of the clusters corresponds to the position and height of the electron having the maximum amplitude in the cluster.
- Poissonian distribution for the number of clusters!



$$\text{Expected number of cluster} = \delta \text{ cluster/cm (M.I.P.)} * \text{drift tube size [cm]} * 1.3 \text{ (relativistic rise)} * 1/\cos(\alpha)$$

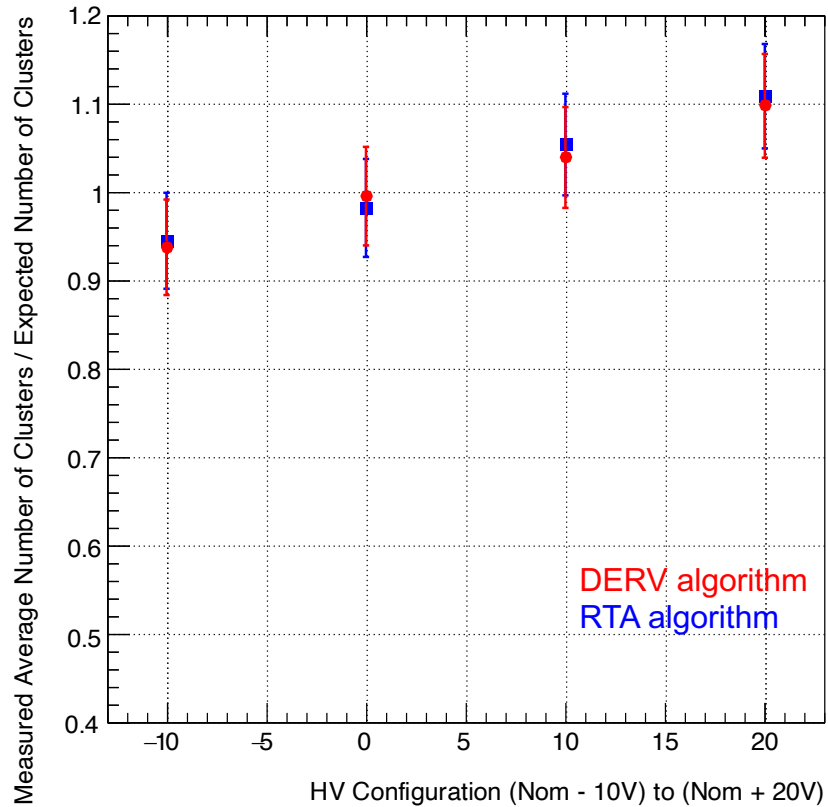
$\alpha$  = angle of the muon track w.r.t. normal direction to the sense wire.

$\delta$  cluster/cm (mip) changes from 12, 15, 18 respectively for He:IsoB 90/10, 85/15 and 80/20 gas mixtures. drift tube size are 0.8, 1.2, and 1.8 respectively for 1 cm, 1.5 cm, and 2 cm cell size tubes.

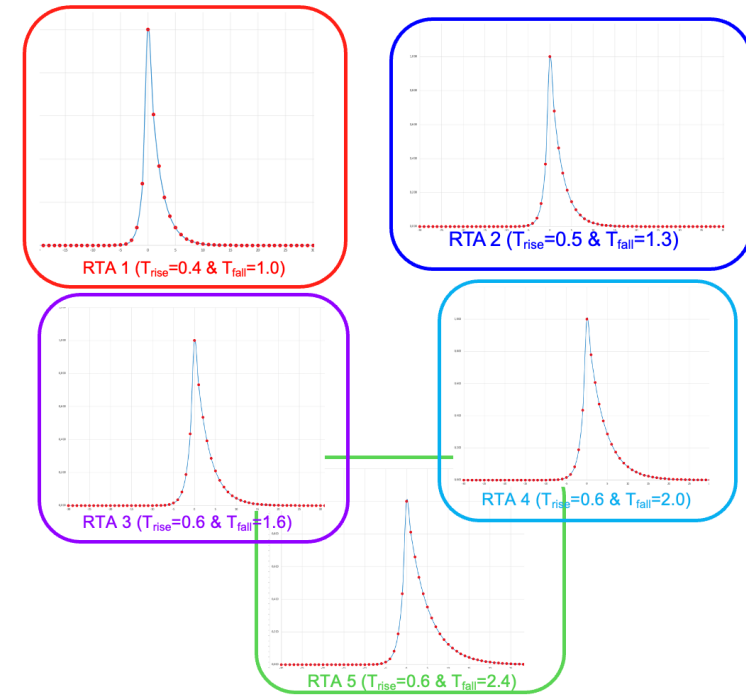
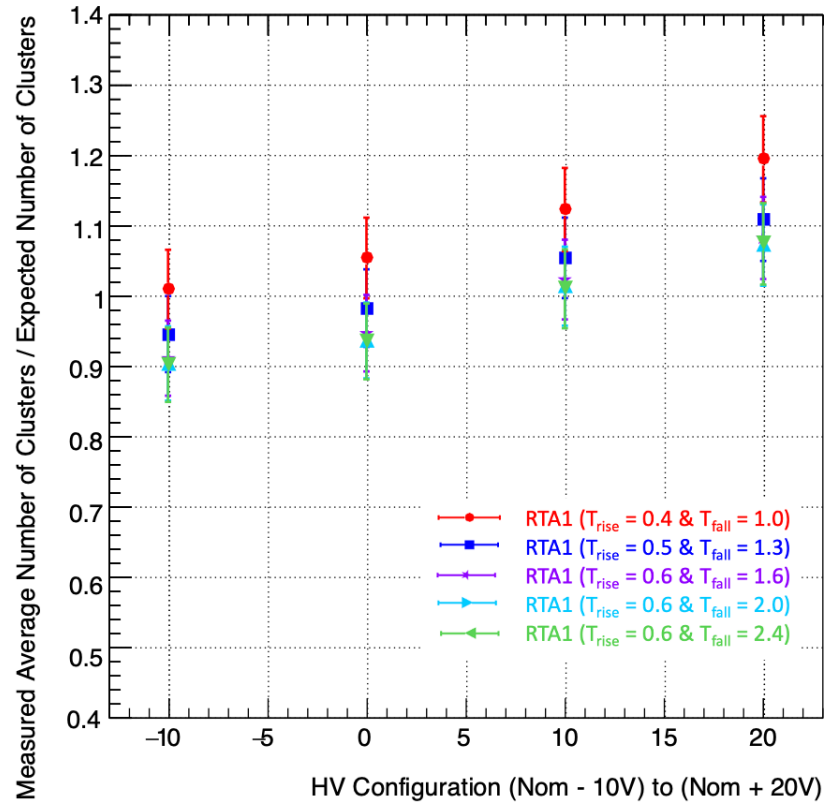




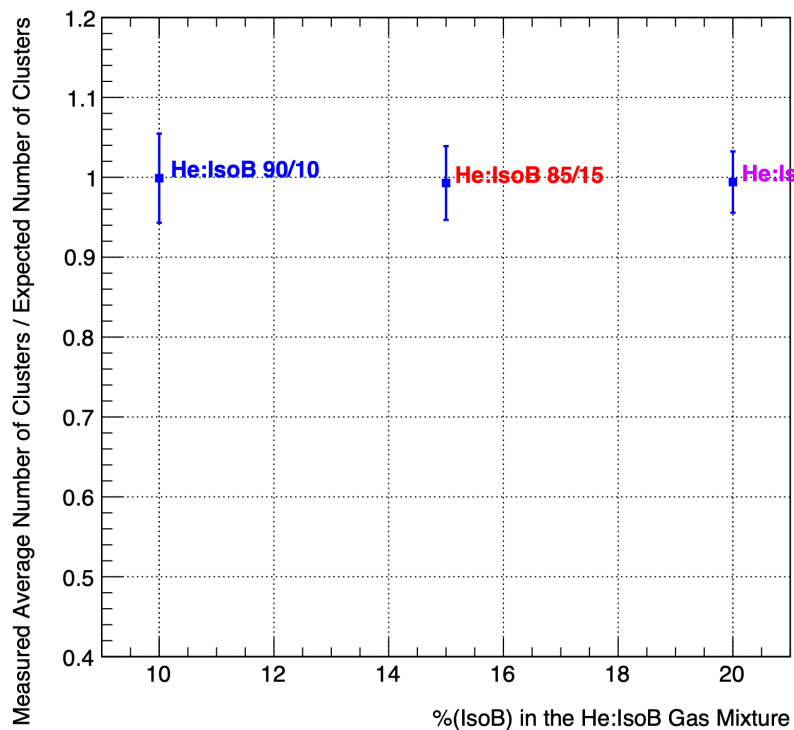
## DERV vs RTA algorithm



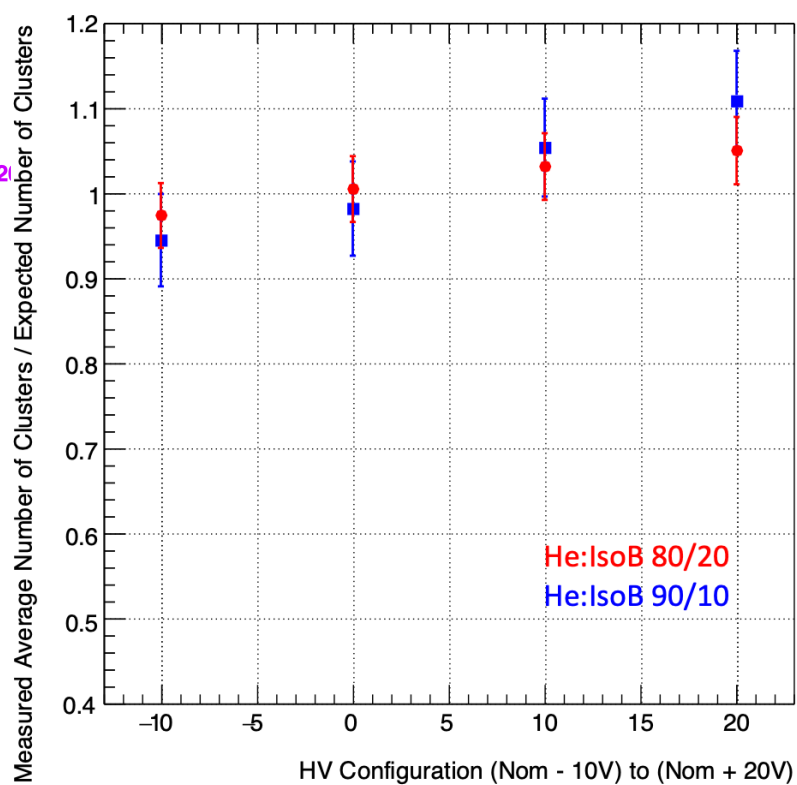
## RTA Templates scan



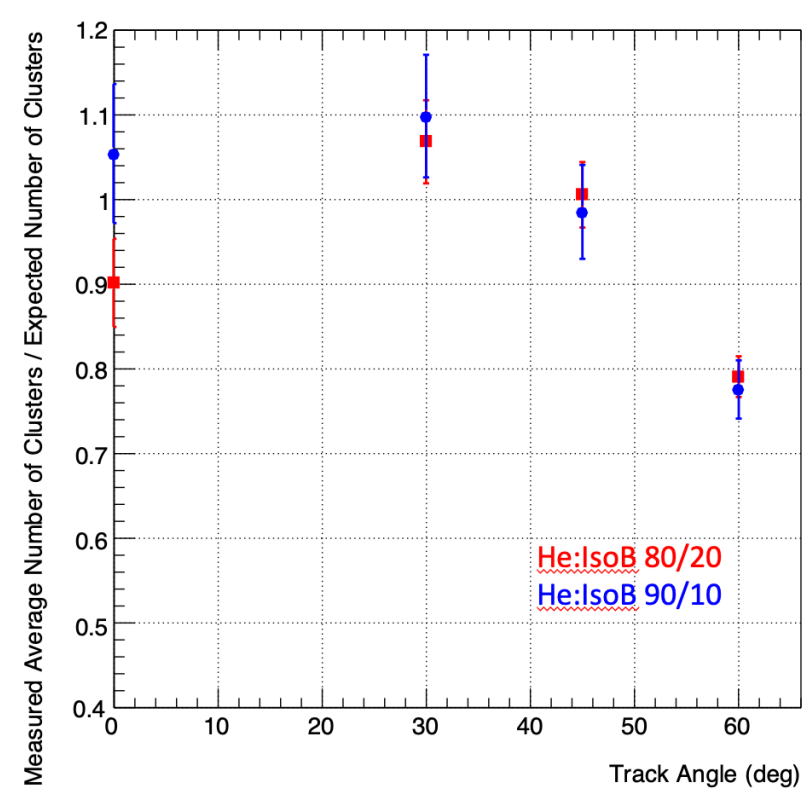
## Gas Mix scan



## HV scan

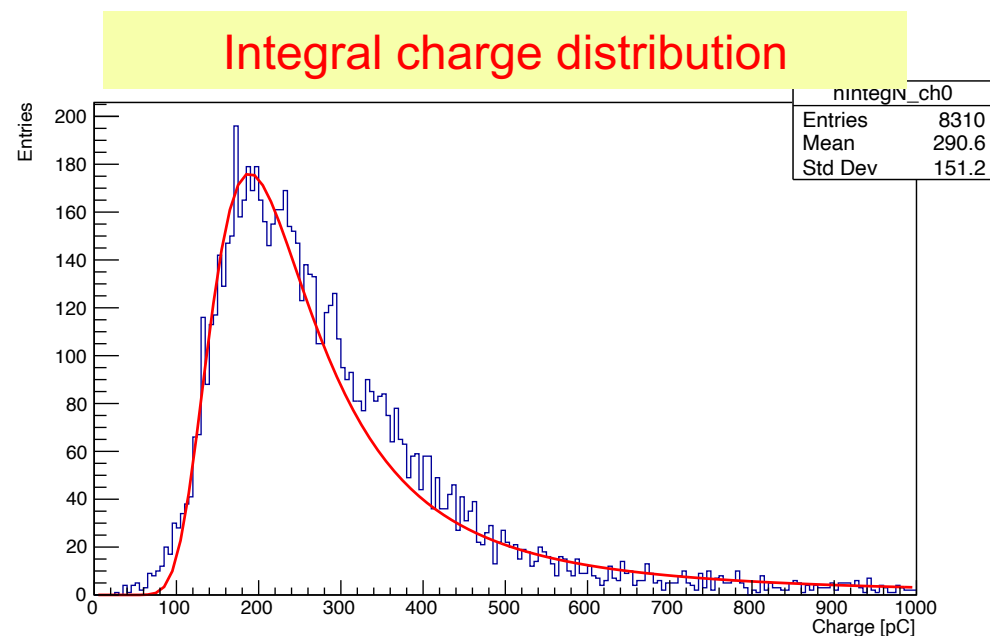


## Angle scan



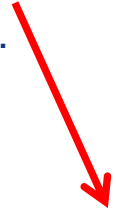
## dE/dx Resolution study:

- Landau distribution for the charges.
- Measure charge of many samples (cells) along track.
- Get "mean" charge over samples = dE/dx.
- Simple "mean" charge subject to large fluctuations  $\Rightarrow$  "Truncated Mean" (robust).
- Reject samples with highest charge (typically) 20-30% and calculate mean ("truncated" mean) of remaining samples.
- Optimize truncation empirically ( $\Rightarrow$  best dE/dx resolution).

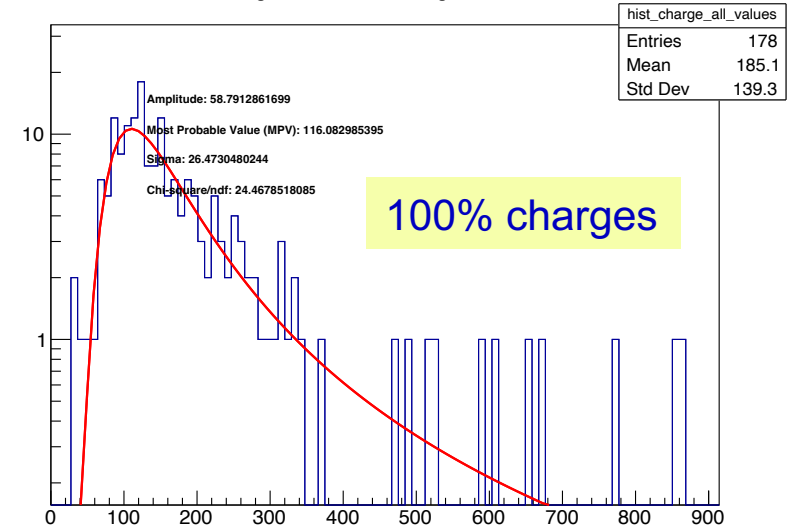


## dE/dx Resolution study:

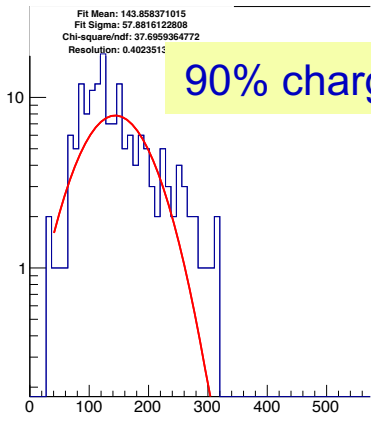
- 2m track length.
- Landau distribution for the charges.
- Optimize truncation empirically ( $\Rightarrow$  best dE/dx resolution).
- Tested the resolution for each.
- Selected the distribution with 80% of the charges to be compared with dN/dx.



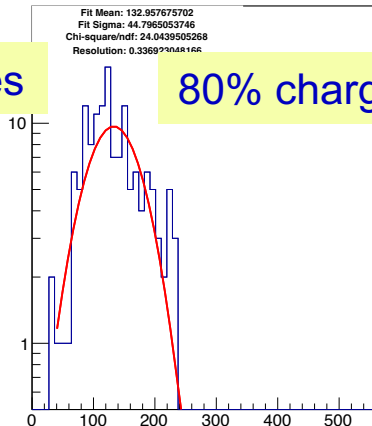
Integral (All) Charge Values



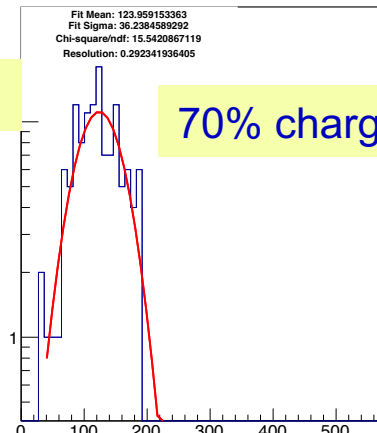
Integral Charge Value:



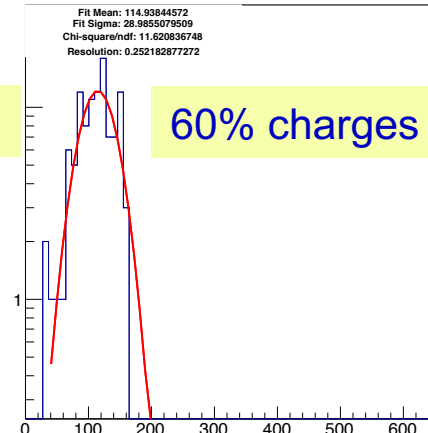
Integral Charge Value



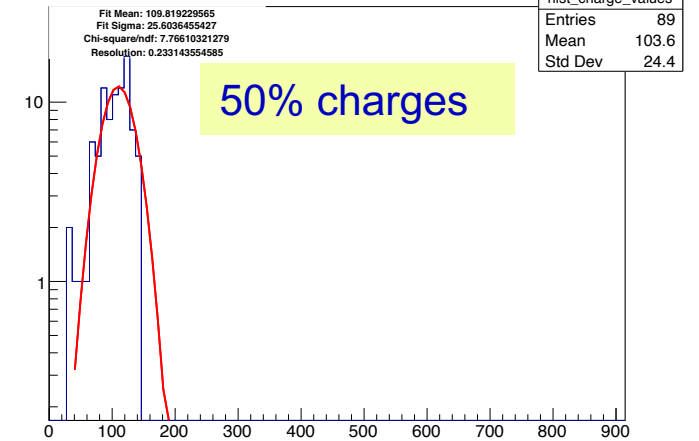
Integral Charge Value:



Integral Charge Values 0.6



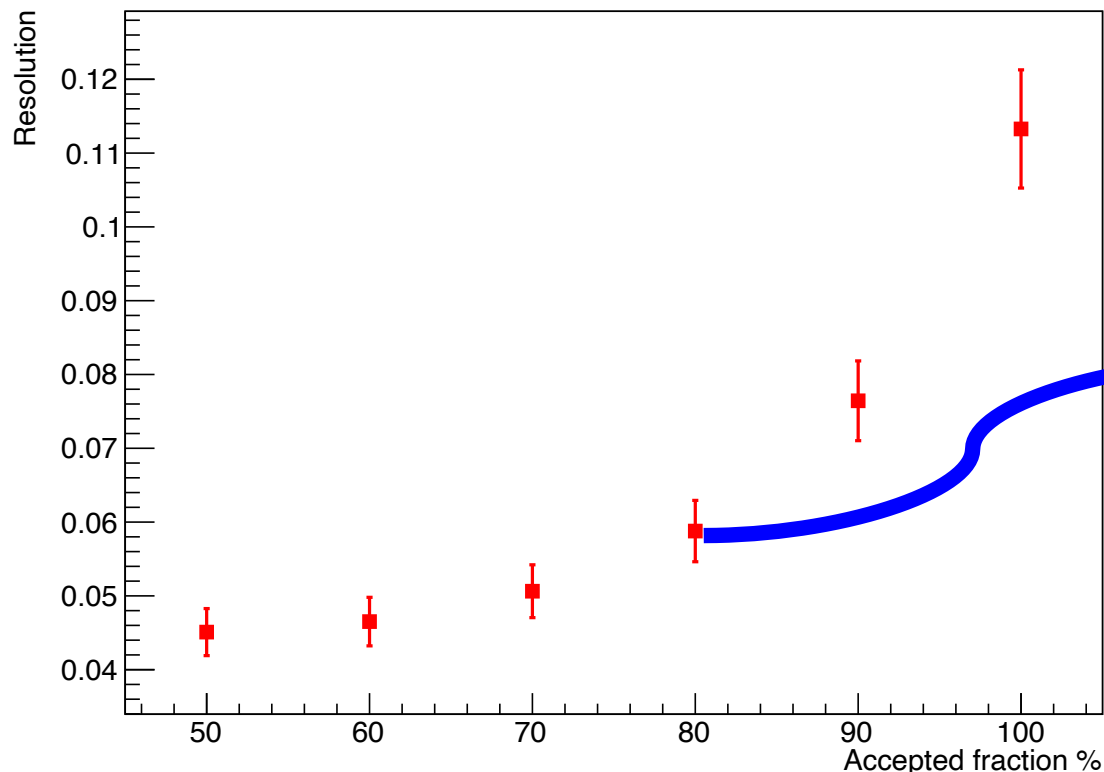
Integral Charge Values 0.5 Cut





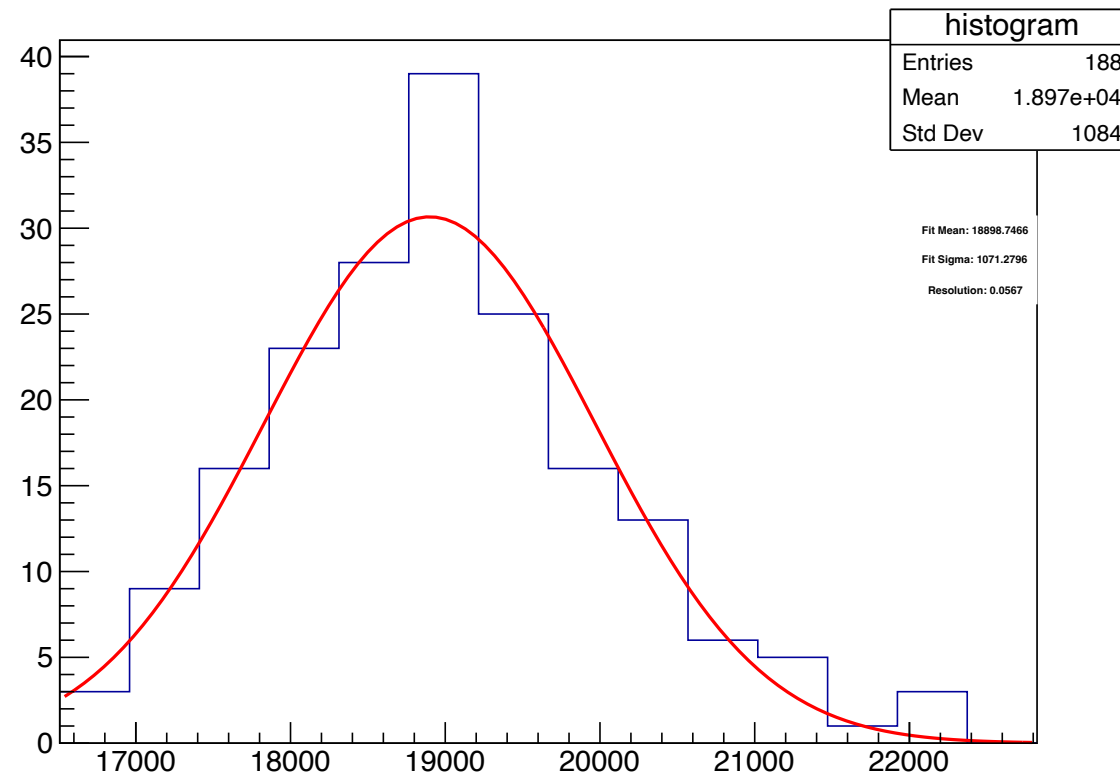
# Resolution study

dE/dx Resolution Scan Vs accepted fraction of charge



dE/dx resolution varies from 4.5% - 11% for 2 m track length relying on the accepted fraction of the charges.

dE/dx Resolution (remove 20% higher charges for each track)

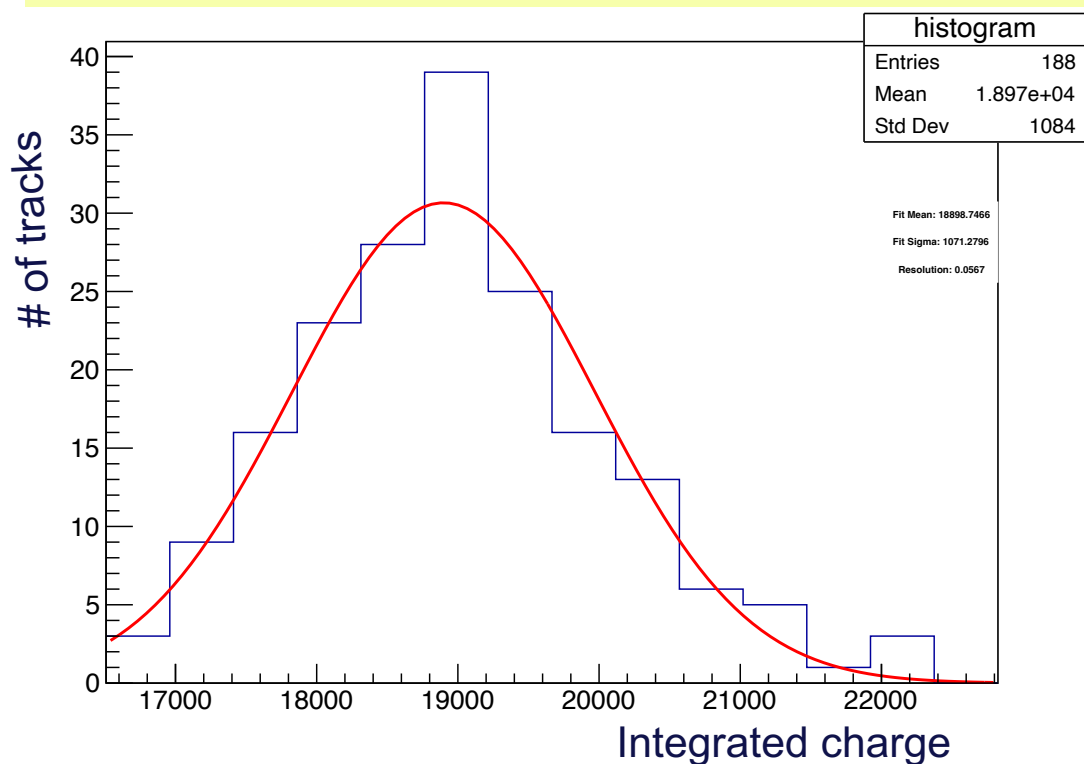


@2m long track we have dE/dx resolution 5.7%

# Resolution study

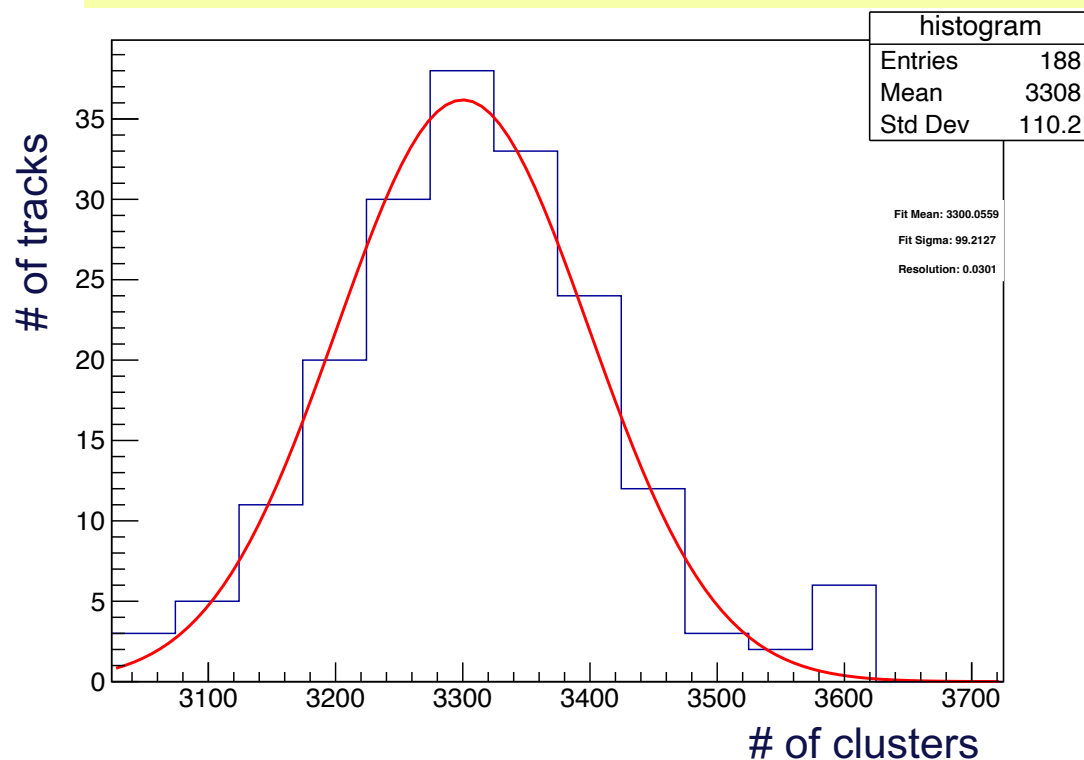
Study done using same tracks (2 m track length) made of the same hits.

## dE/dx Resolution (remove 20% higher charges)



@2m long track we have dE/dx resolution 5.7%

## dN/dx Resolution



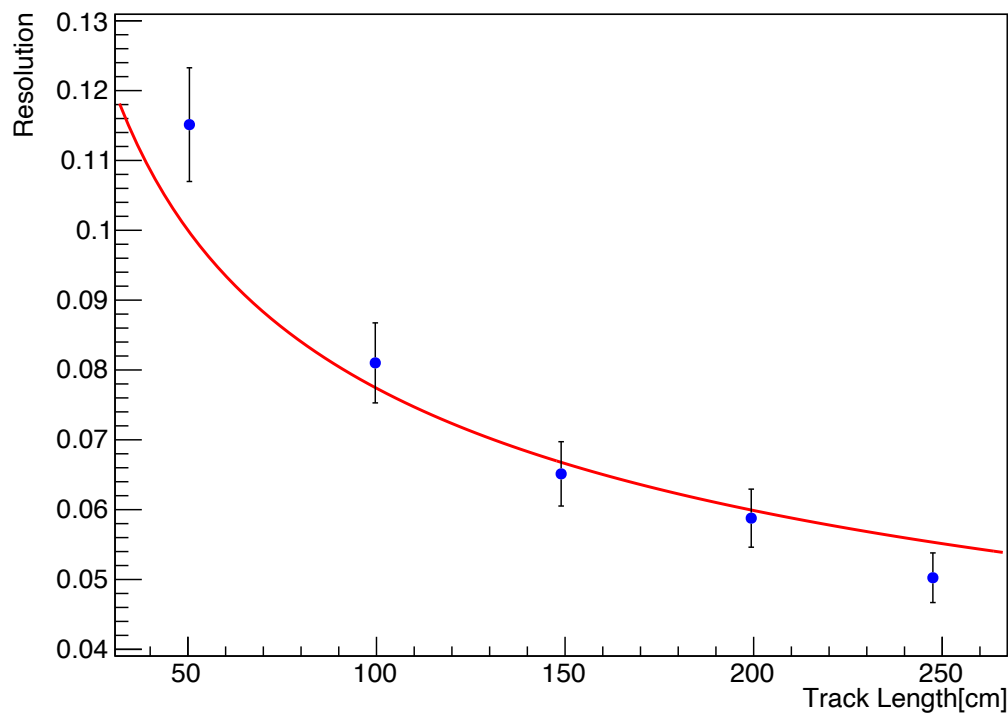
@2m long track we have dN/dx resolution 3%

~ 2 times improvement in the resolution using dN/dx method

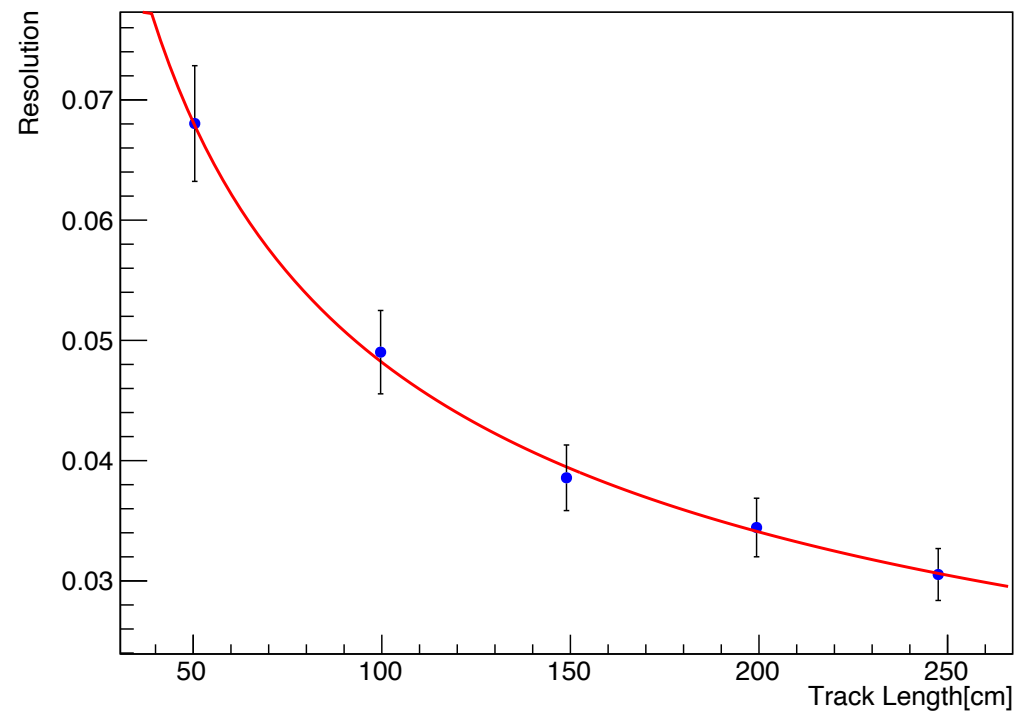
# Resolution study

2m tracks length

dE/dx Resolution



dN/dx Resolution



dE/dx resolution dependence on the track length  $L^{-0.37}$

dN/dx resolution dependence on the track length  $L^{-0.5}$

**~ 2 times improvement in the resolution using dN/dx method**

# Summary

- The cluster counting technique is a high powerful method to improve the particle identification capabilities: analytic evaluation and simulation confirm its potentials.
- Using the test beam data we evaluated the performance of our algorithms across various conditions: gas mixture, gain, geometrical configuration (cell size, sense wires size), sampling rate, HV, and track angle .
- Two different promising algorithms have been developed and used for finding the electron peaks (DERV & RTA algorithms).
- There is a good agreement between the results from the two algorithms and the expectation.
- Using  $dN/dx$  method gives a resolution 2 times better than the  $dE/dx$  method in agreement with the analytical calculation.

Stay tuned for the new results from 2024 test beam  
on the relativistic rise region!



# Backup

From the ordered sequence of the electrons arrival times, considering the average time separation between clusters and their time spread due to diffusion, reconstruct the most probable

sequence of clusters drift times:  $\{t_i^{cl}\} \quad i = 1, N_{cl}$

$dE/dx$

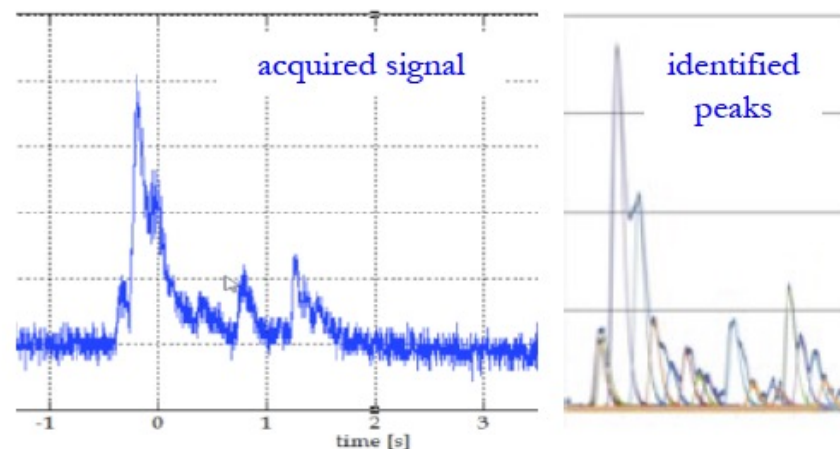
$$\frac{\sigma_{dE/dx}}{(dE/dx)} = 0.41 \cdot n^{-0.43} \cdot (L_{track} [m] \cdot P[atm])^{-0.32}$$

from *Walenta parameterization (1980)*

truncated mean cut (70-80%) reduces the amount of collected information  $n = 112$  and a 2m track at 1 atm give

$$\sigma \approx 4.3\%$$

Increasing P to 2 atm improves resolution by 20% ( $\sigma \approx 3.4\%$ ) but at a considerable cost of multiple scattering contribution to momentum and angular resolutions.



$dN_{cl}/dx$

$$\frac{\sigma_{dN_{cl}/dx}}{(dN_{cl}/dx)} = (\delta_{cl} \cdot L_{track})^{-1/2}$$

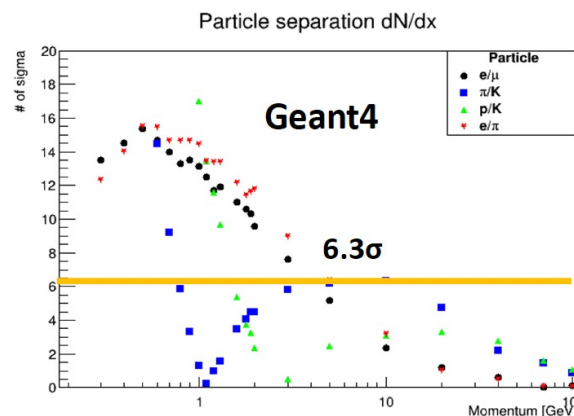
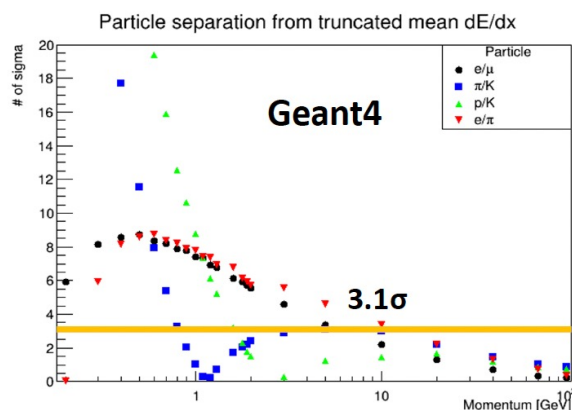
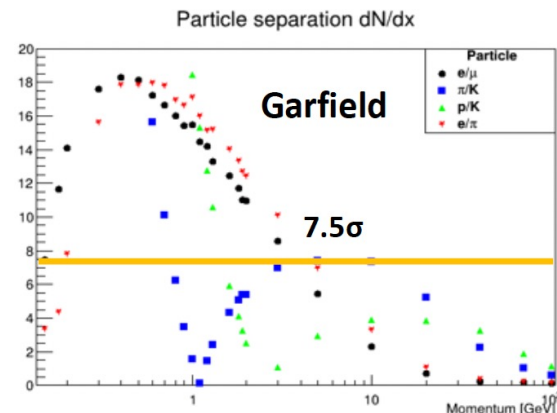
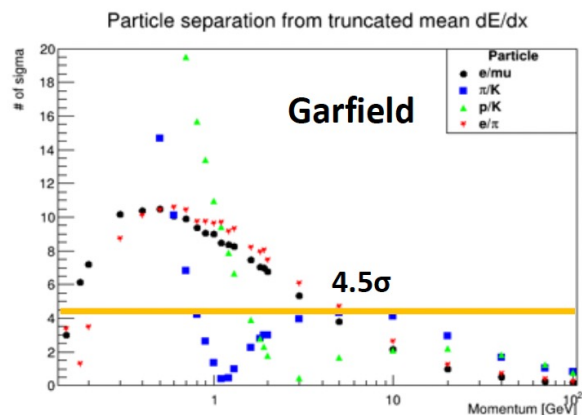
from *Poisson distribution*

$\delta_{cl} = 12.5/cm$  for He/iC<sub>4</sub>H<sub>10</sub>=90/10 and a 2m track give  
 $\sigma \approx 2.0\%$

A small increment of iC<sub>4</sub>H<sub>10</sub> from 10% to 20% ( $\delta_{cl} = 20/cm$ ) improves resolution by 20% ( $\sigma \approx 1.6\%$ ) at only a reasonable cost of multiple scattering contribution to momentum and angular resolutions.

# The simulation of the cluster counting

- We have developed an algorithm, which uses the energy deposit information provided by Geant4, to reproduce, in a fast and convenient way, the clusters density and the cluster size distributions predicted by Garfield++.



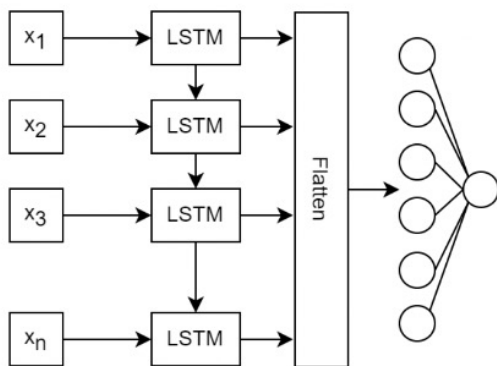
Garfield++ in reasonable agreement with analytical calculations up to 20 GeV/c momentum, then falls much more rapidly at higher momenta.

Despite Geant4 uses the cluster density and the cluster size distributions from Garfield++, it disagrees from Garfield++ and, therefore, from the analytical calculations also.

The algorithm is under development in IHEP, for more information see [this talk](#) by Guang ZHAO.

## Peak finding with LSTM

Why LSTM? Waveforms are time series

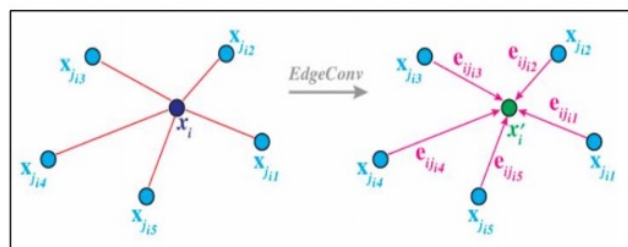


- Architecture: LSTM (RNN-based)
- Method: Binary classification of signals and noises on slide windows of peak candidates

LSTM: Long Short-Term Memory

## Clusterization with DGCNN

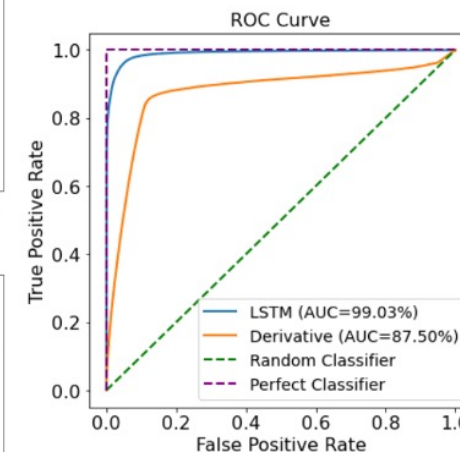
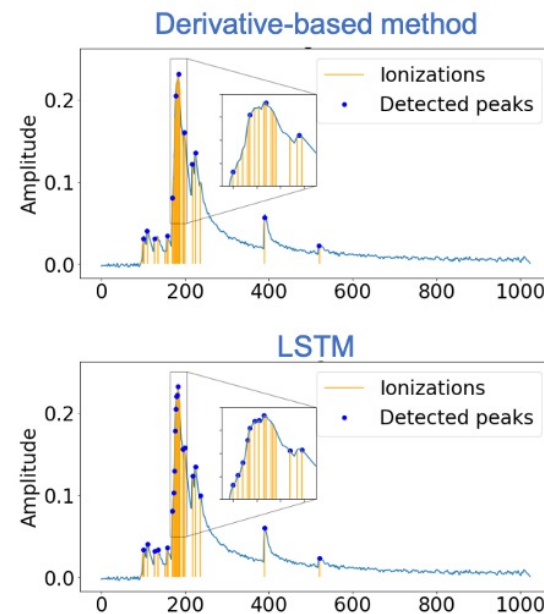
Why DGCNN? Locality of the electrons in the same primary cluster, perform message passing through neighbour nodes in GNN



arXiv: 1801.07829

- Architecture: DGCNN (GNN-based)
- Method: Binary classification of primary and secondary electrons

DGCNN: Dynamic Graph Convolutional neural networks

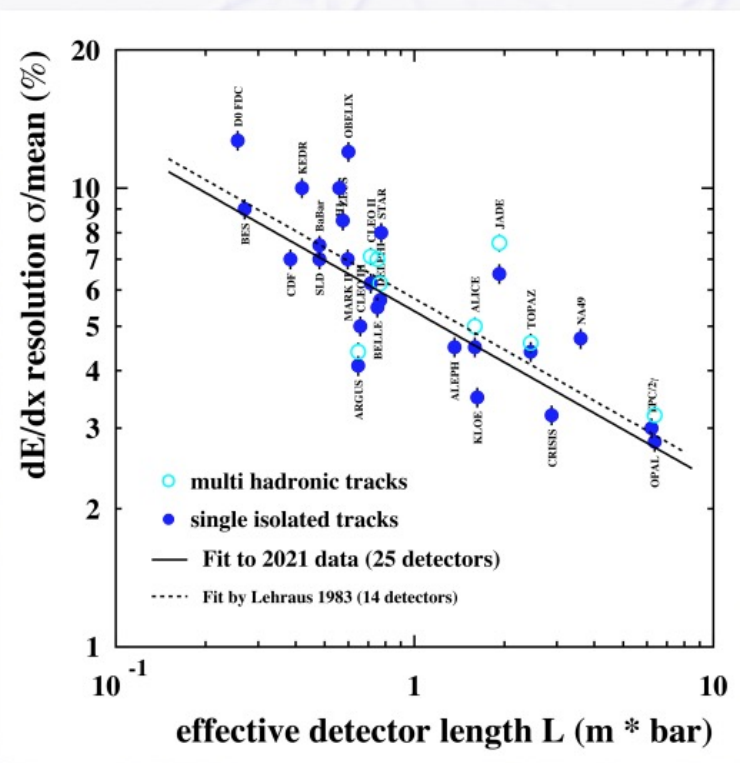


**LSTM model is better classifier compared to derivative-based model**



## “Lehraus” Plot 2021

- dE/dx resolution achieved in large detectors, mainly at e<sup>+</sup>e<sup>-</sup> colliders, at some hadron colliders and fixed target expts.



- Fit by Lehraus 1983:  
dE/dx res. =  $5.7 * L^{-0.37}$  (%)
- Fit in 2021 (25 large detectors):  
dE/dx res. =  $5.4 * L^{-0.37}$  (%)
  - 5.4% typical dE/dx resolution for 1 m track length
  - no significant change to 1983
  - performance of present generation of detectors as predicted ~40 years ago