# Progress on cluster counting for the 4th concept DC


**Guang Zhao**
**for DC PID Group**
zhaog@ihep.ac.cn

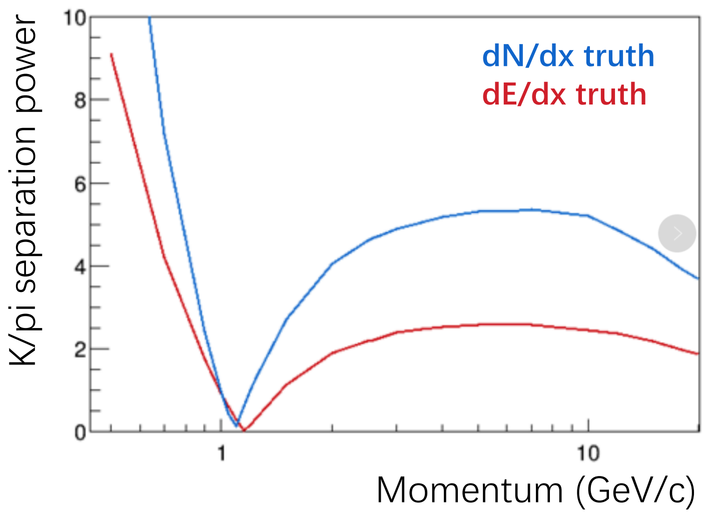

**CEPC Day**
**Jan. 18, 2023**

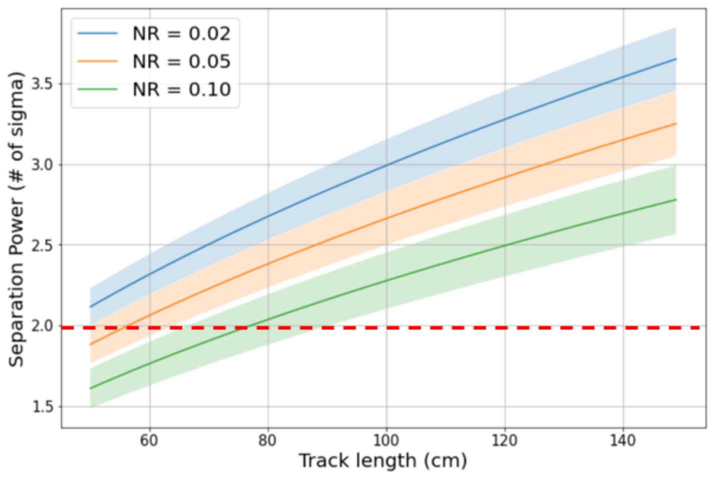
# Motivation

- **Excellent PID is required for CEPC physics programs**
  - Preliminary requirement: > $2\sigma$ for 20 GeV/c pion/kaon separation
- **The cluster counting (dN/dx) is the most promising breakthrough in PID**
  - Resolution has, potentially, a factor of 2 better than the dE/dx
- **Drift chamber with cluster counting can provide sufficient PID power while keeping a reasonable detector size**

# Recaps of simulation

- **Ideal simulation (truth):**
  - K/pi separation with dN/dx is x2 better than dE/dx
- **Full simulation (more realistic, w/ electronics/noises/reconstruction):**
  - $2\sigma$ K/pi separation at 20 GeV/c
- **Determine DC parameters with full simulation**



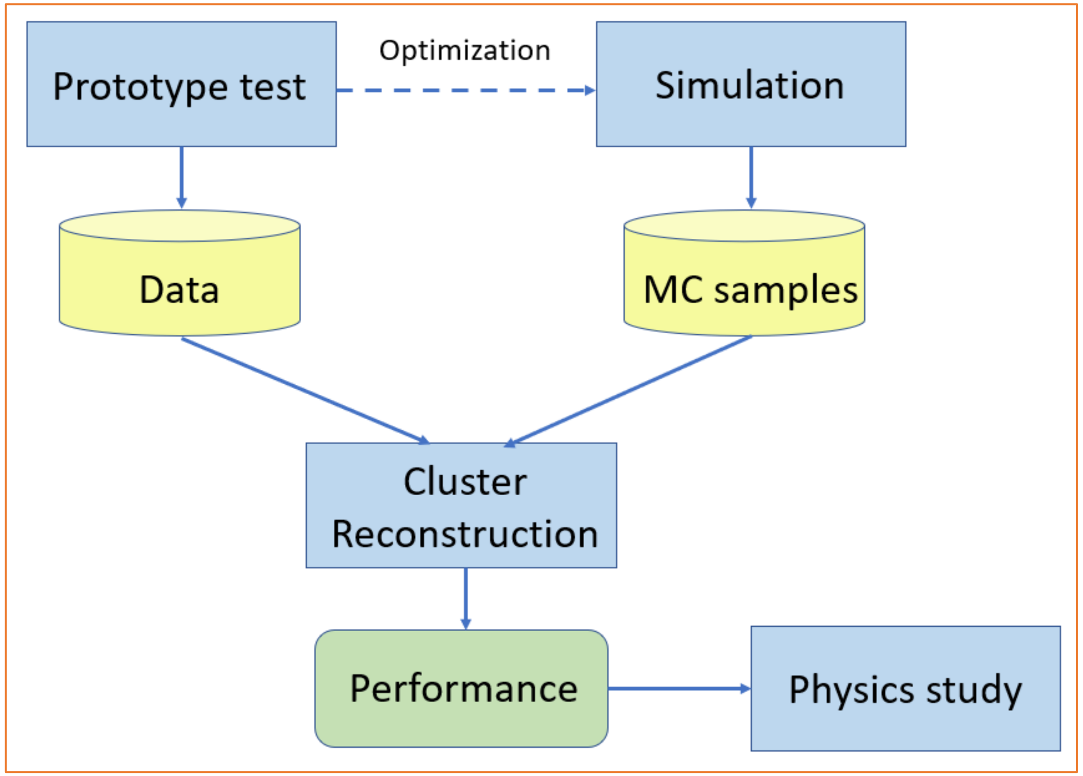

The resolution scales with $L^{-0.5}$
(Noise ratio in beam-test is close to 10%)

| DC Parameters | |
|---|---|
| Radius extension | 800-1800 mm |
| Length of outermost wires（cosθ=0.82） | 5143 mm |
| Thickness of inner CF cylinder | 200 μm |
| Outer CF frame structure | Equivalent CF thickness: 1.63 mm |
| Thickness of end Al plate | 35 mm |
| Cell size | 18 mm × 18 mm |
| # of cells | 24766 |
| Ratio of field wires to sense wires | 3:1 |
| Gas mixture | $He/iC_4H_{10}$=90:10 |

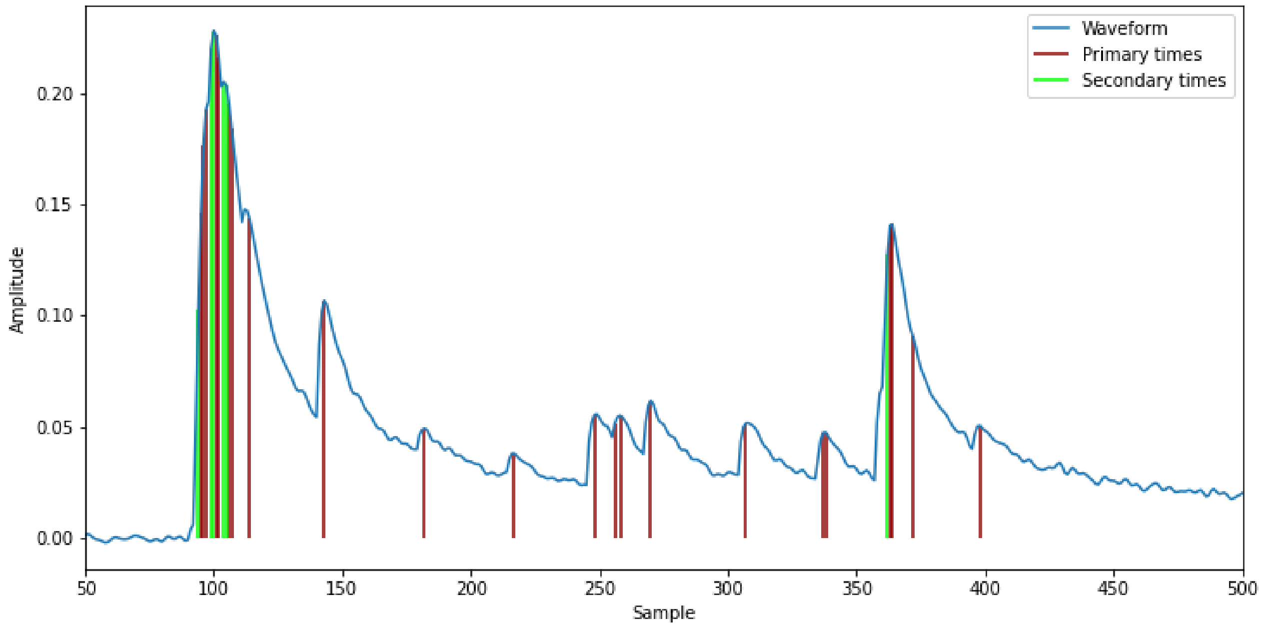
# Major efforts in this talk

- Reconstruction algorithm
- Understanding beam test data

Cluster counting workflow



# Reconstruction algorithm



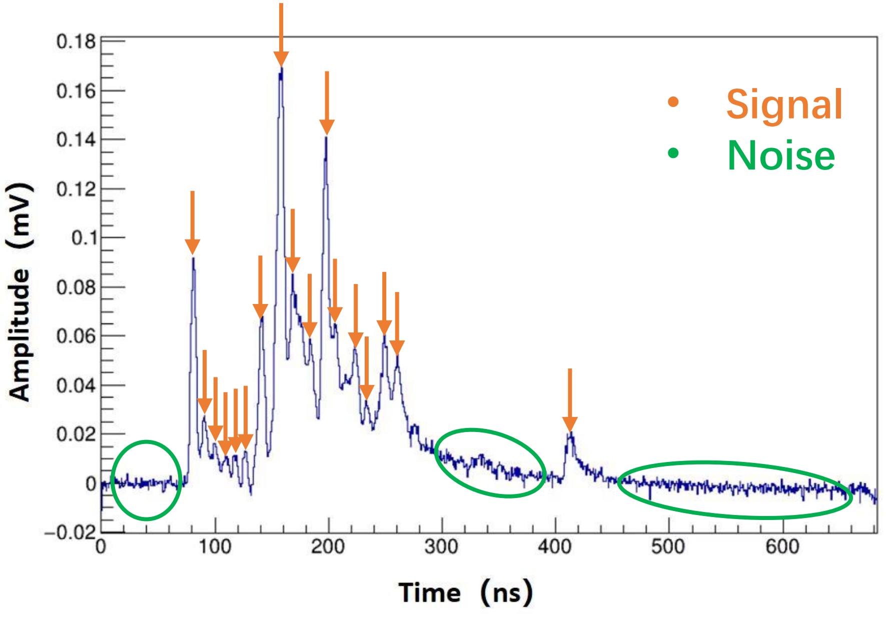

- **Task:**
  - Both **primary electrons** and **secondary electrons** contribute peaks on the waveform
  - Find the number of peaks from **primary electrons**

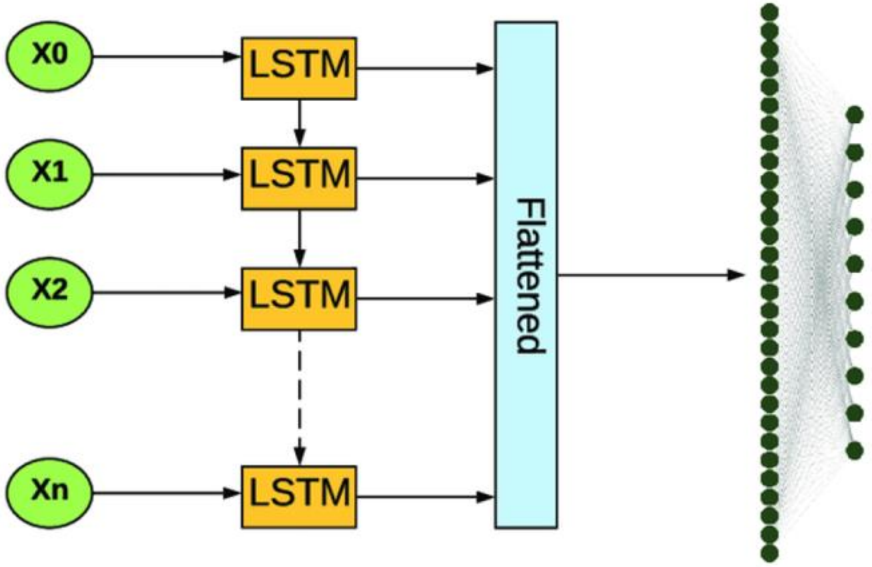
- **Difficulties**
  - High pile-up
  - Could be noisy

- **Machine learning** can make full use of the waveform information, could be effective

# Step 1: Peak finding

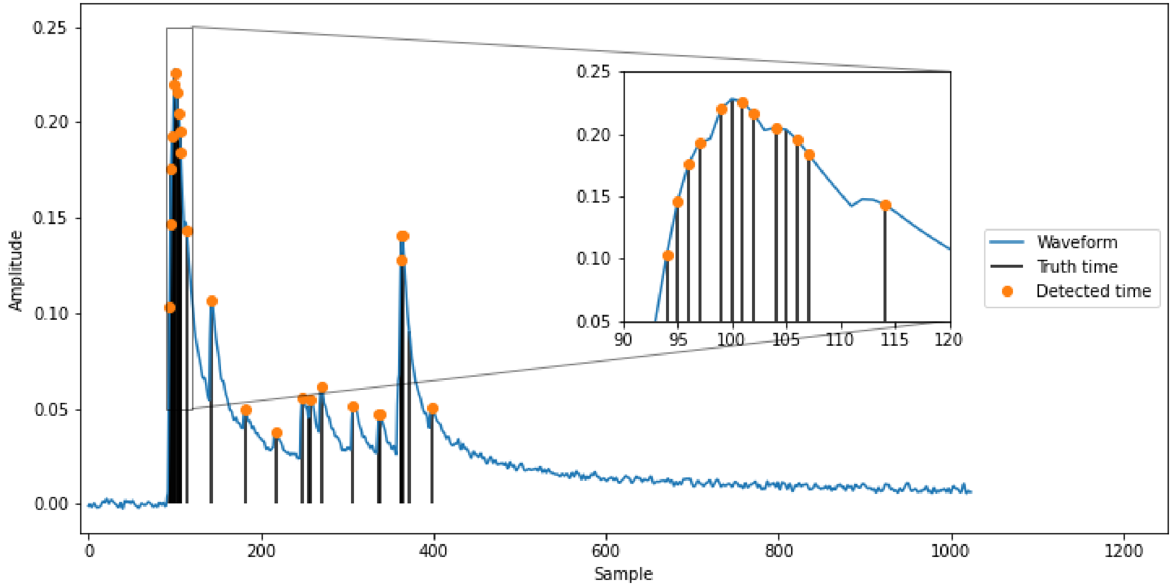
- To detect all peaks from the waveform
- Classification for “signals” and “noises”
- Architecture: Recurrent Neural Network (RNN)



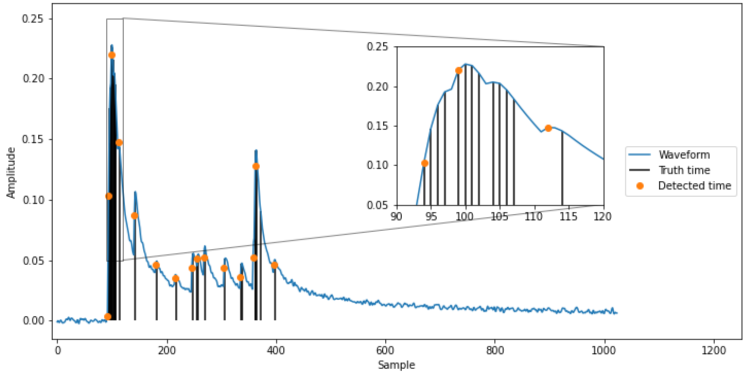



# Compare to the derivative algorithm

## Neural Network

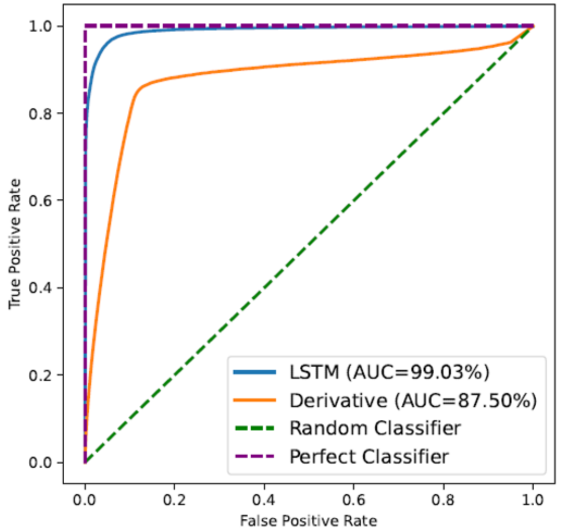



## Derivative




The NN can find the peaks more effective!

## ROC

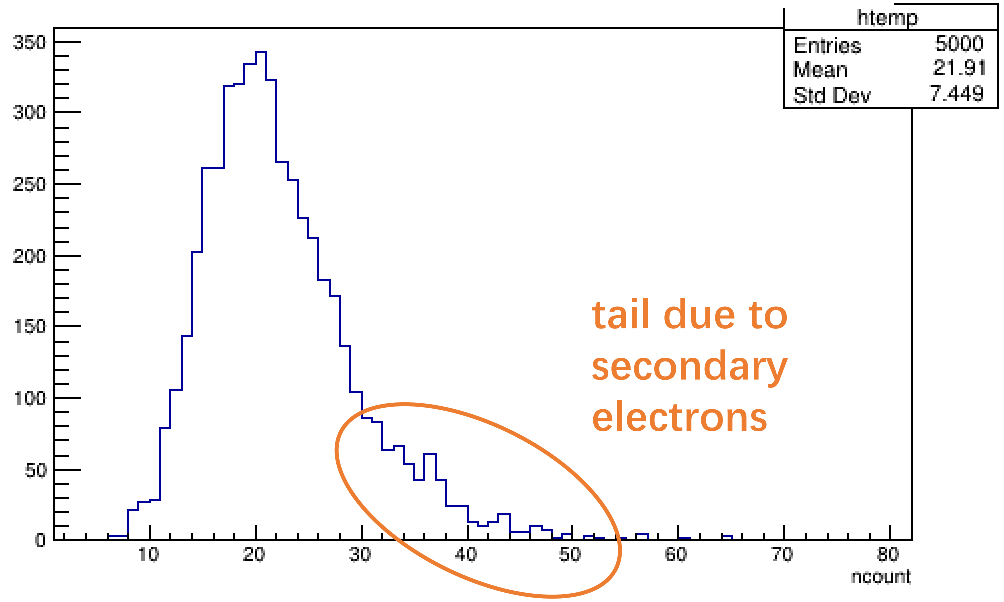


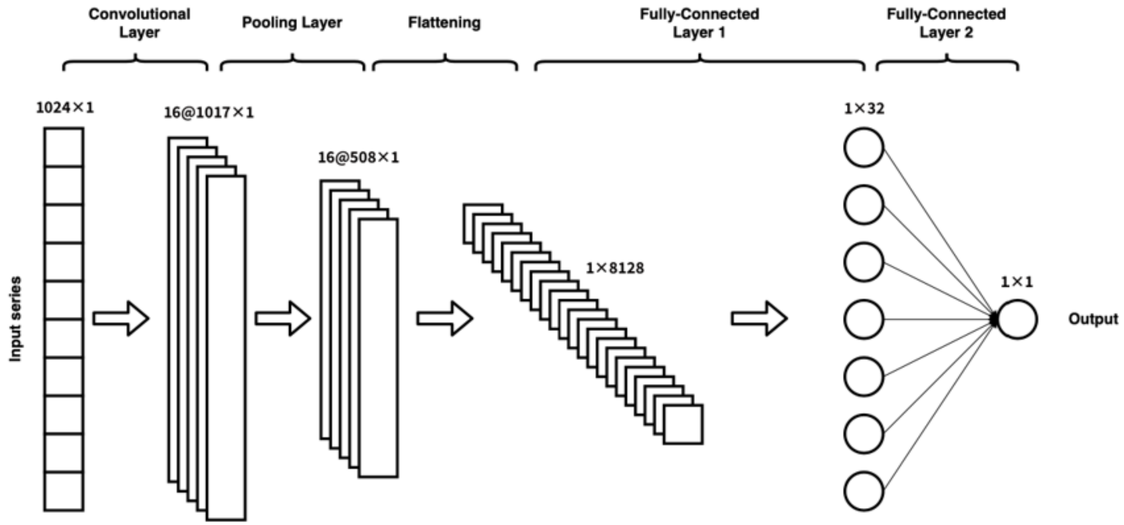

- RNN
- Derivative

# Step 2: Discrimination of the primaries

- To discriminate primary peaks from the secondary ones
- Regression problem
- Architecture: 1D Convolutional Neural Network (CNN)

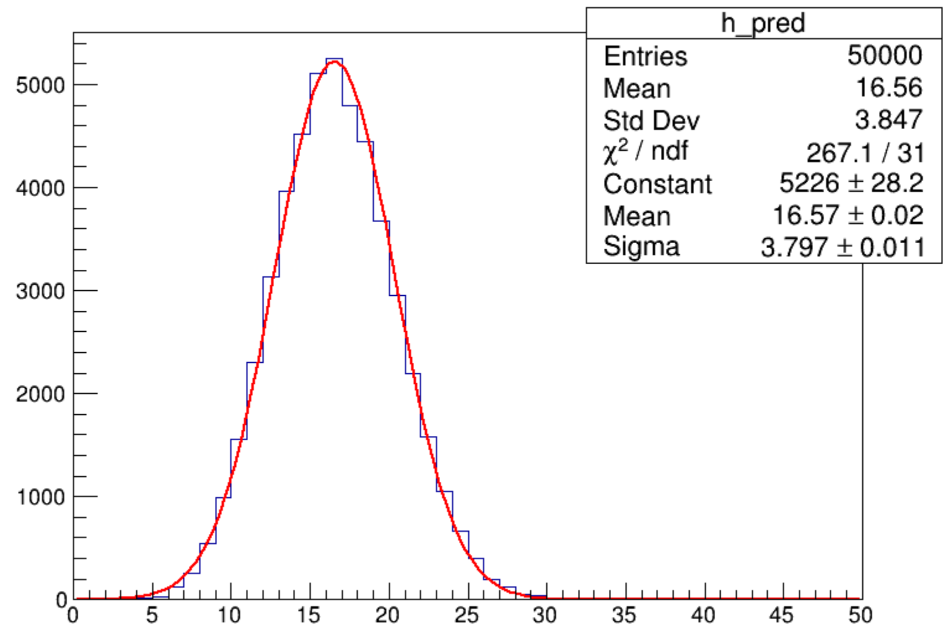
Detected peaks from the peak finding



# $N_{cls}$ distributions

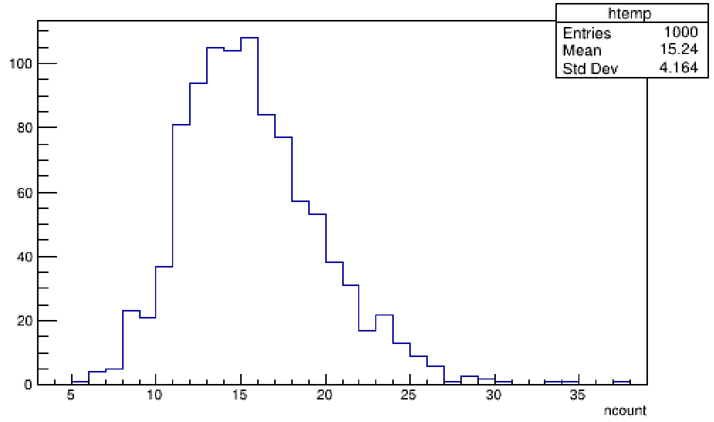
ML




Resolution ~23%

- ✓ Very good Gaussian-like distribution
- ✓ The resolution is very close to the truth value (~21%), which implies possible improvement on PID

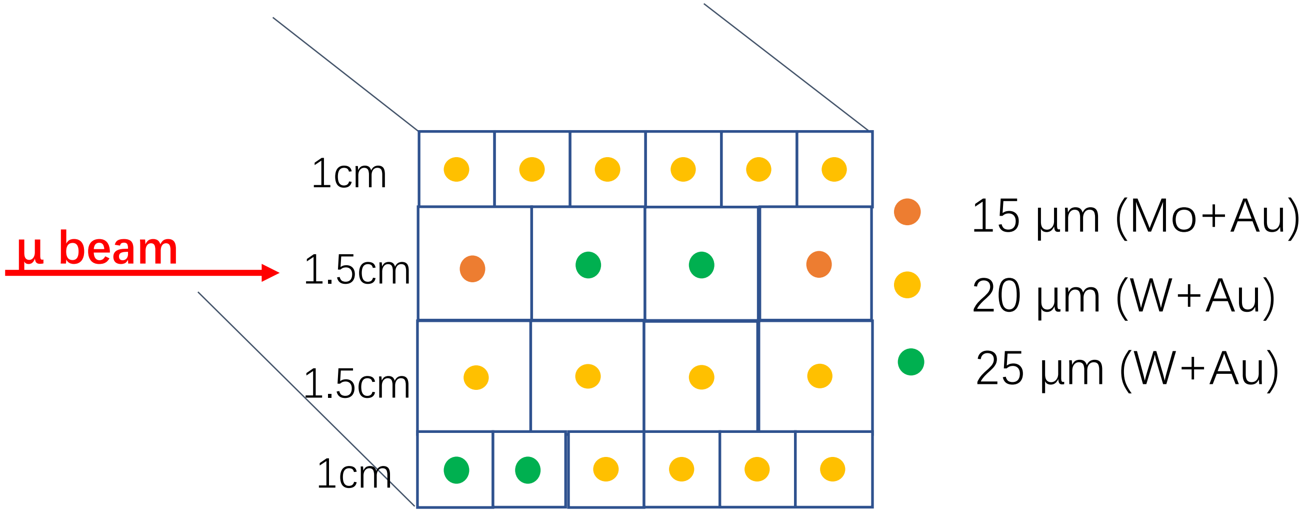
Derivative

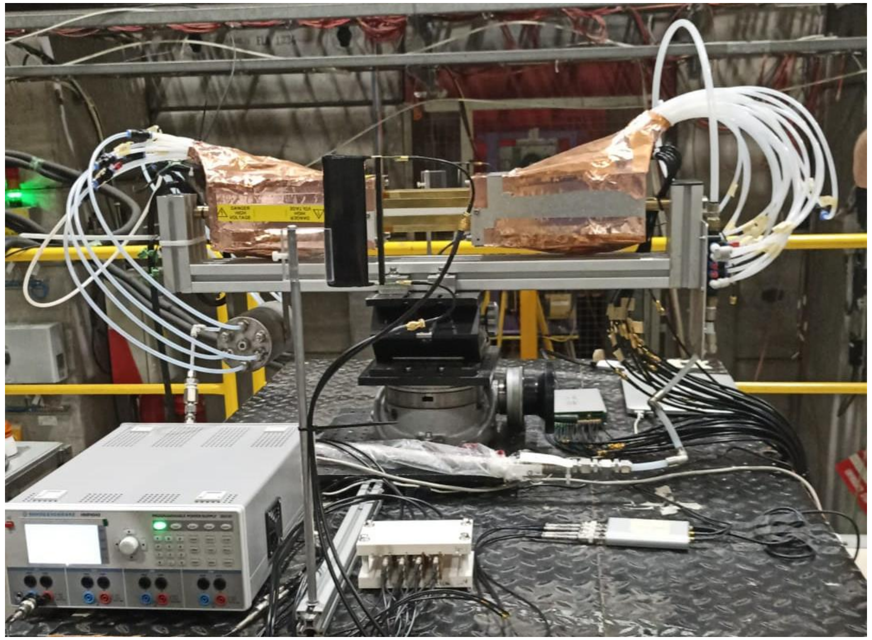



✓ RMS/MEAN ~ 27%

# Beam test

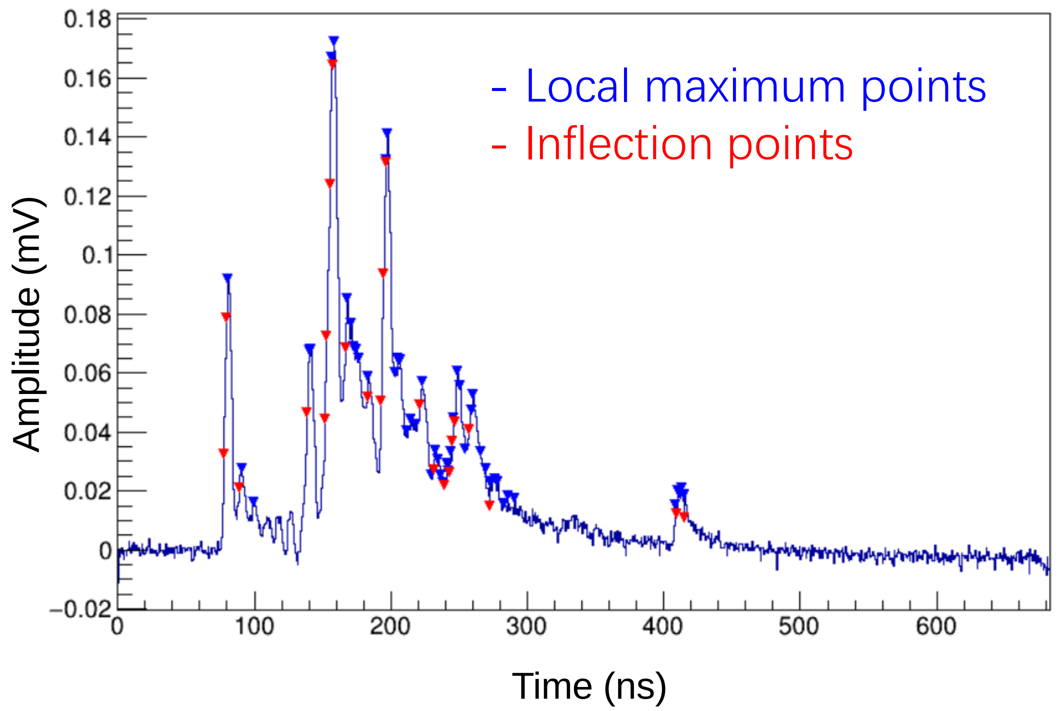
- **Two rounds of beam test organized by INFN group**
  - First round: Nov. 2021 (Shuiting joined)
  - Second round: July 2022 (Shuaiyi joined)
- **Cooperation between INFN and IHEP on data analysis is ongoing**

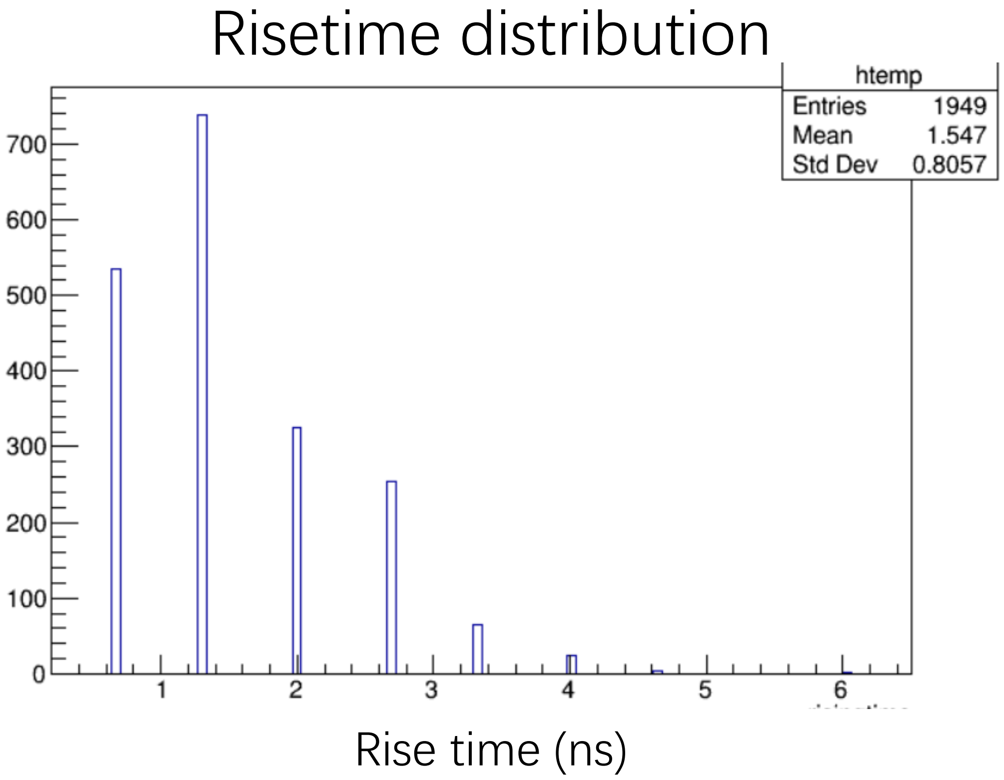


**2nd round of test beam:**
Federica's talk in Meeting on cluster counting in drift chambers, Jul 28, 2022
https://indico.ihep.ac.cn/event/16837/

# Peak risetime estimation

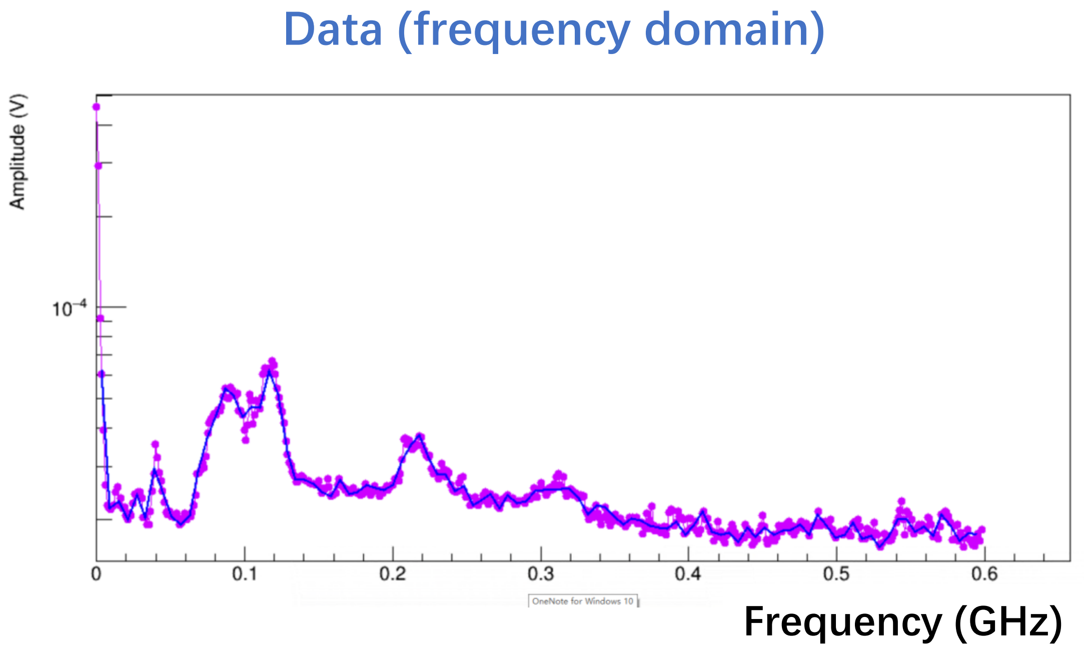


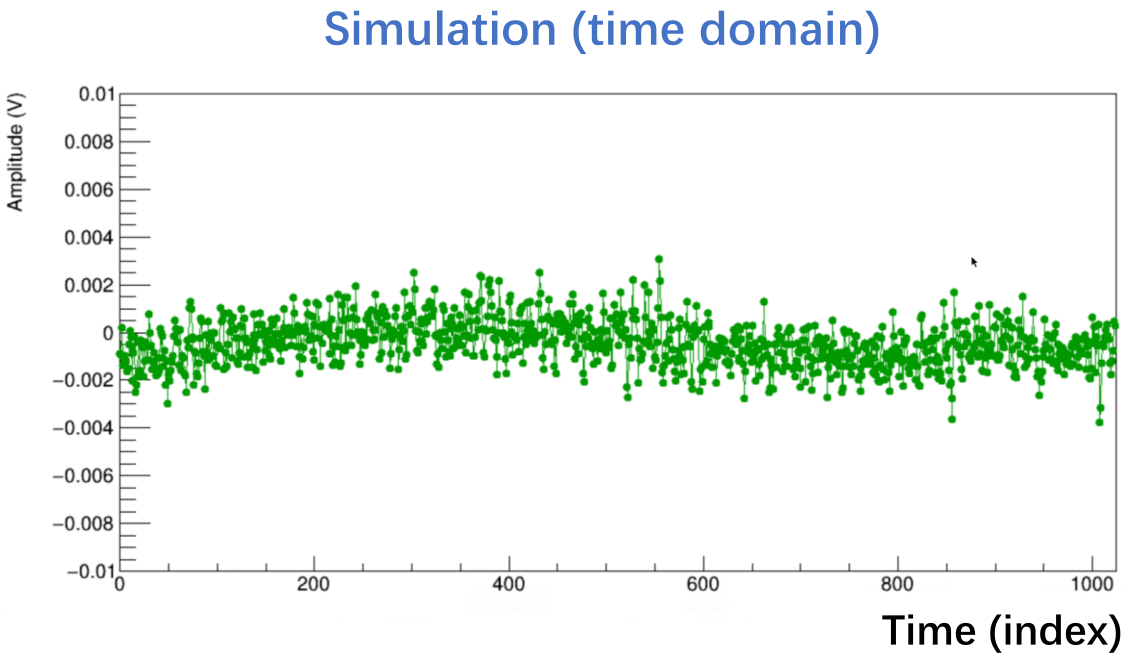

- Require local maximum >0.015mV
- For each local maximum points, find the nearest early point, whose slope starts to decrease (inflection point)
- The time distance between the local maximum and the inflection points is recorded as the risetime.

- Each bin is around 0.67 ns
- Averaged risetime is **1.55 ns**
- 95% of the risetimes less than 2.67 ns

# Noise extraction




IFFT

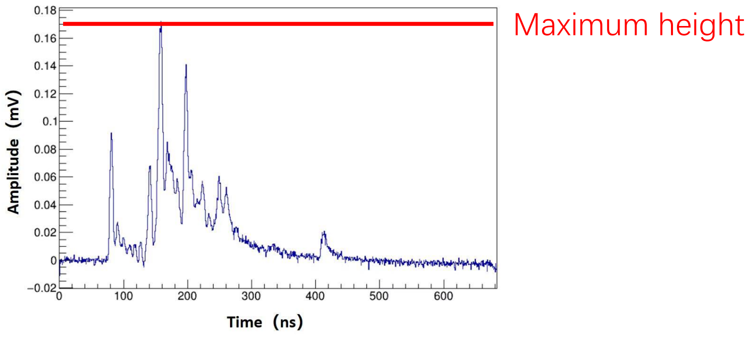



## Noise distribution

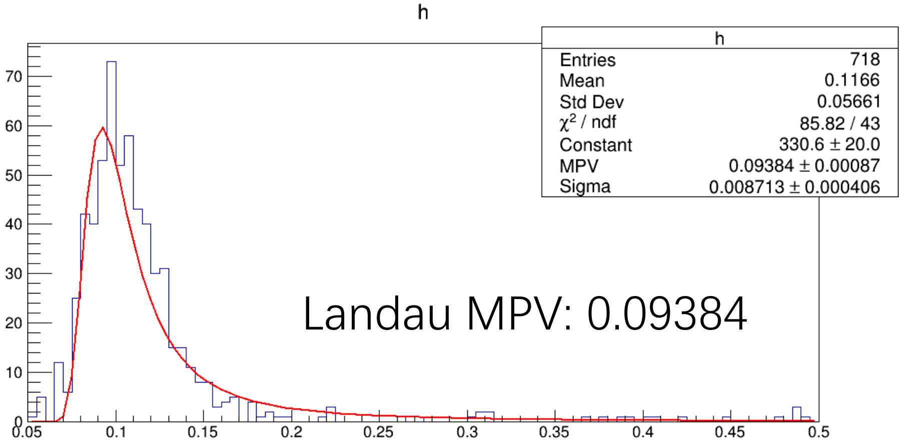
| | Data | Simulation |
|---|---|---|
| Mean | 0.00019 | 0.00058 |
| RMS | 0.00125 | 0.00125 |

# Signal amplitude scale

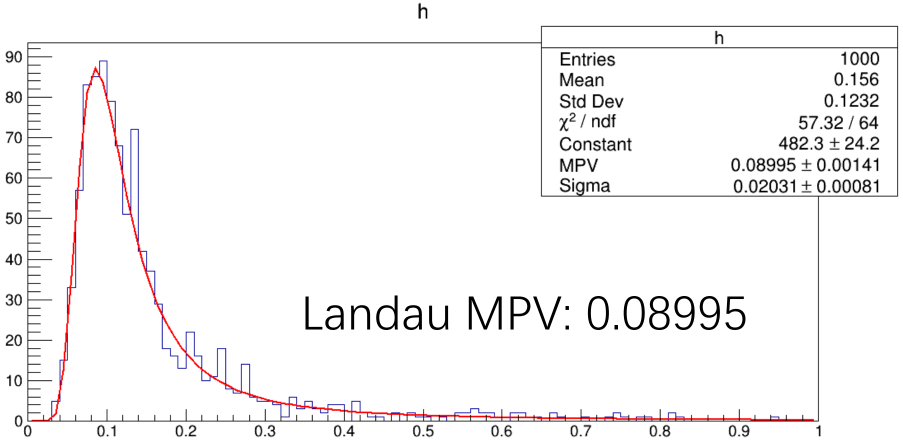
- **Scale the amplitude of MC simulation to data**




Average max. height in Data

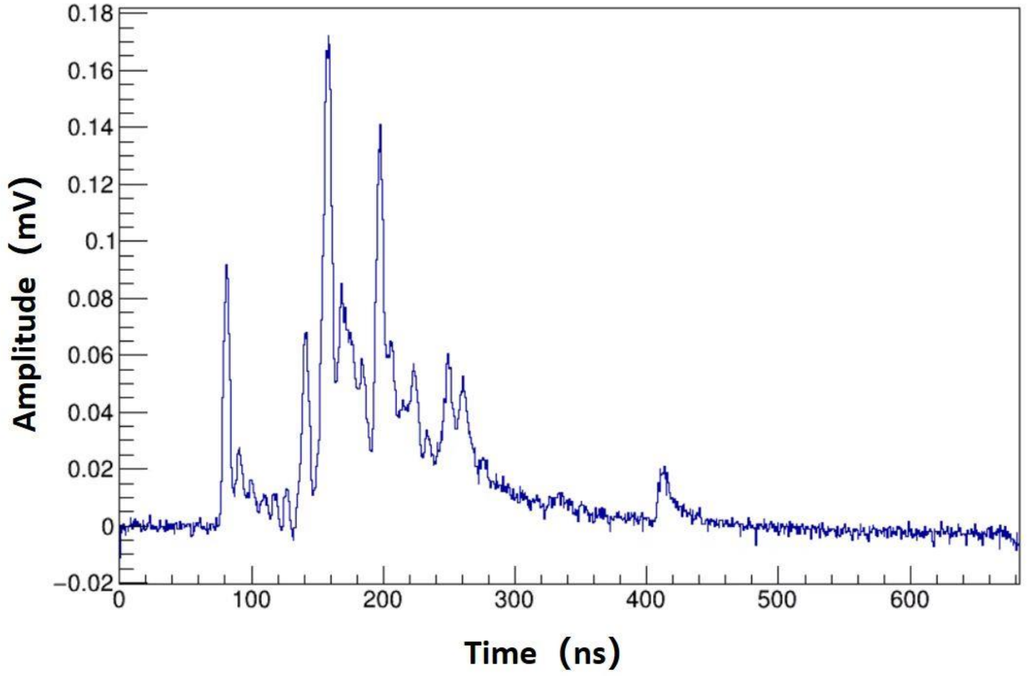



Average max. height in scaled MC

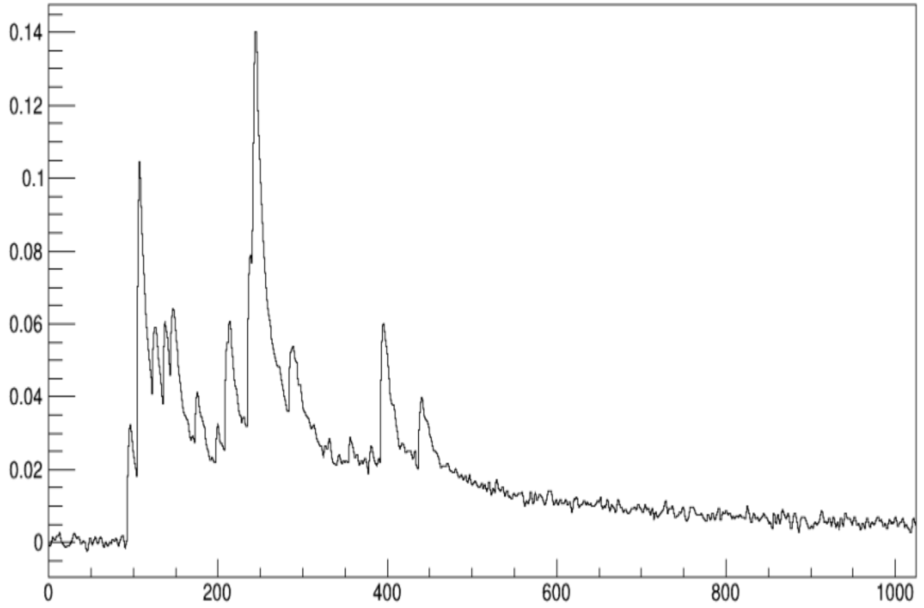


# MC tuning with data

Data

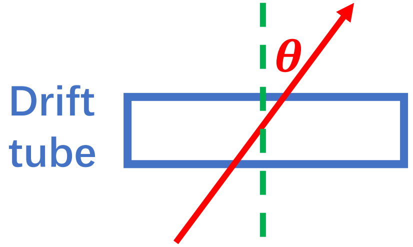


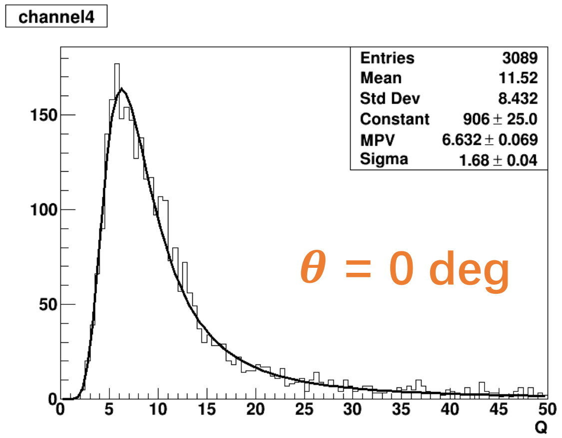

Simulation

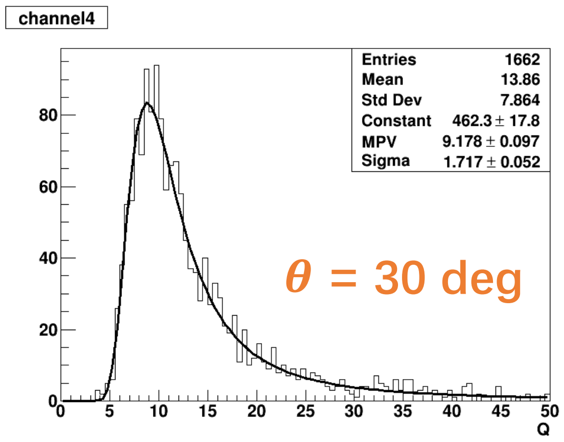


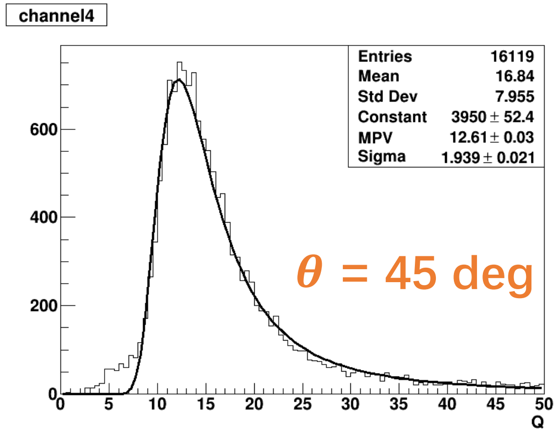

**Tuning MC with data: (1) Rise-time, (2) Noise, (3) Amplitude**
**➔ Better data/MC consistency**

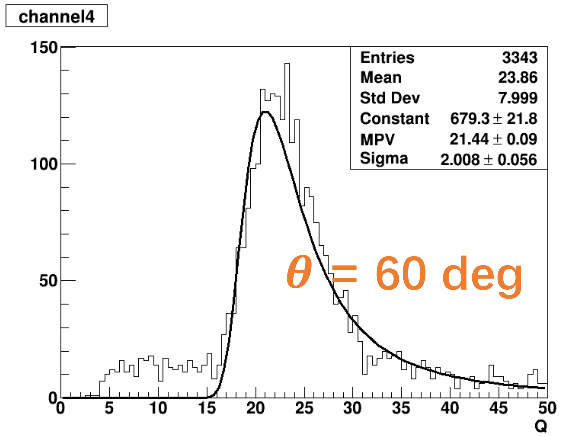
# Charge distribution checks



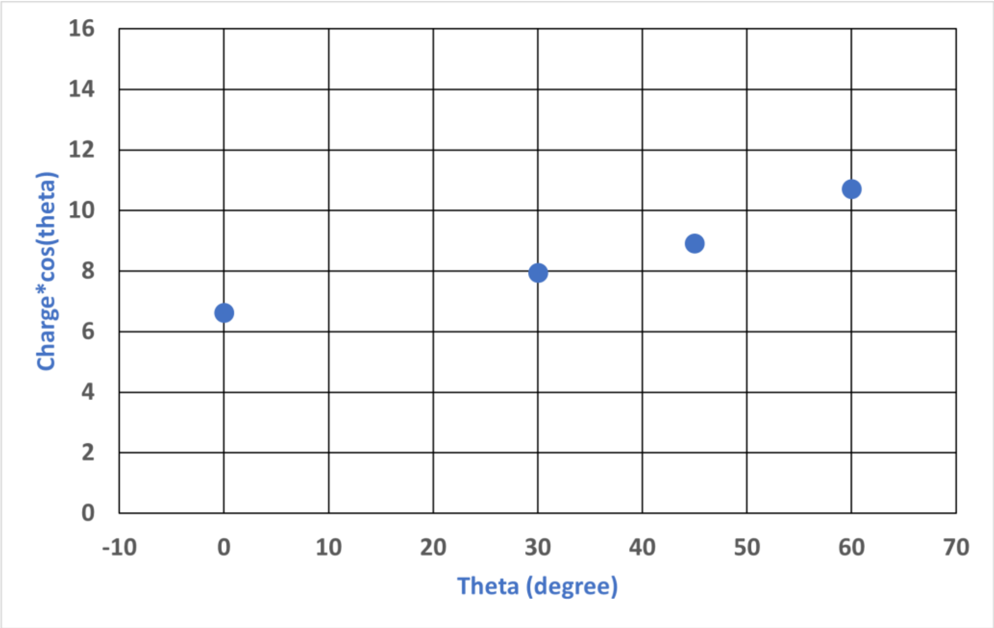



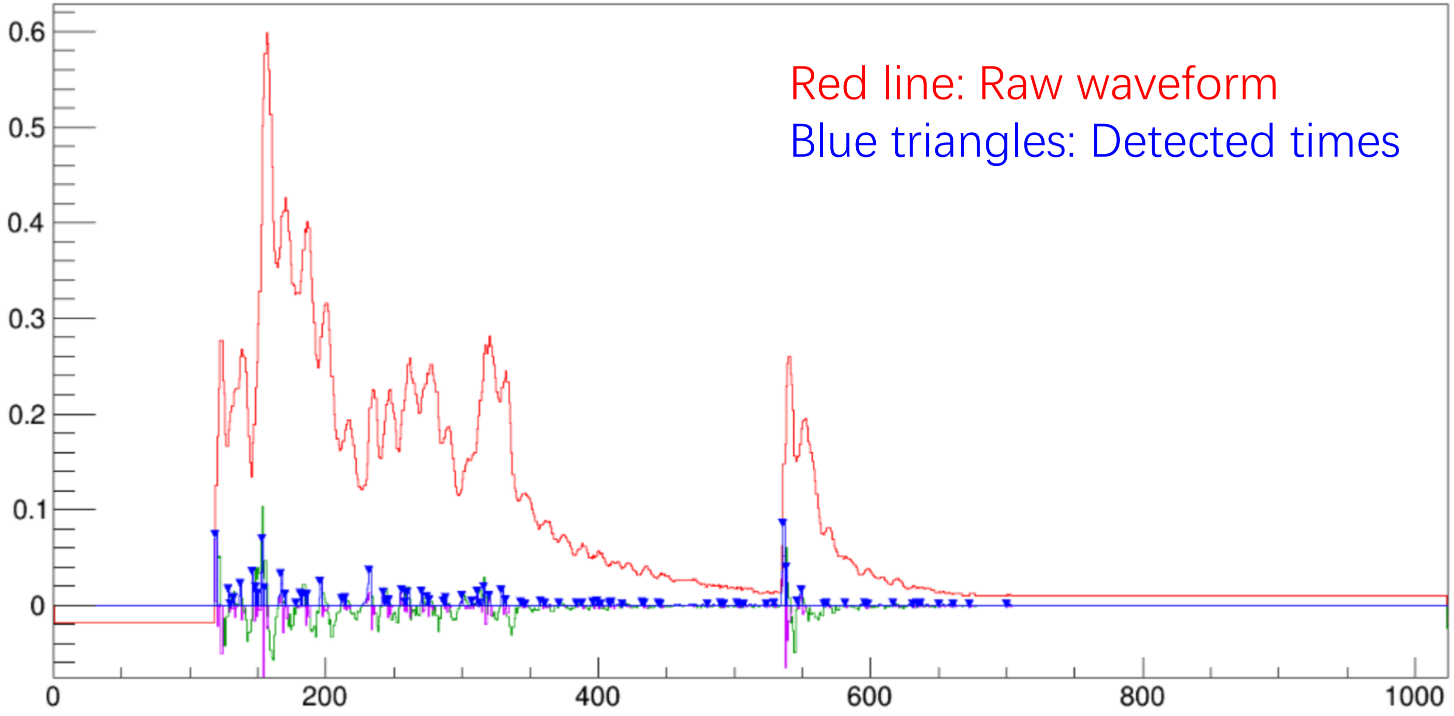

## Charge MPV scaled by track length




➔ Variation due to space charge

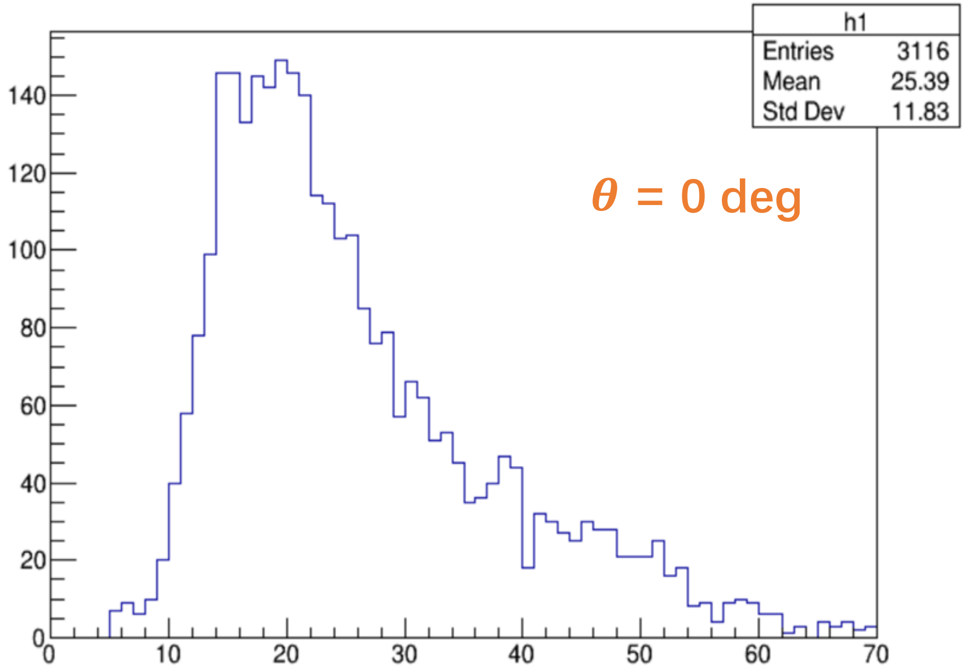
# Preliminary peak finding with derivatives

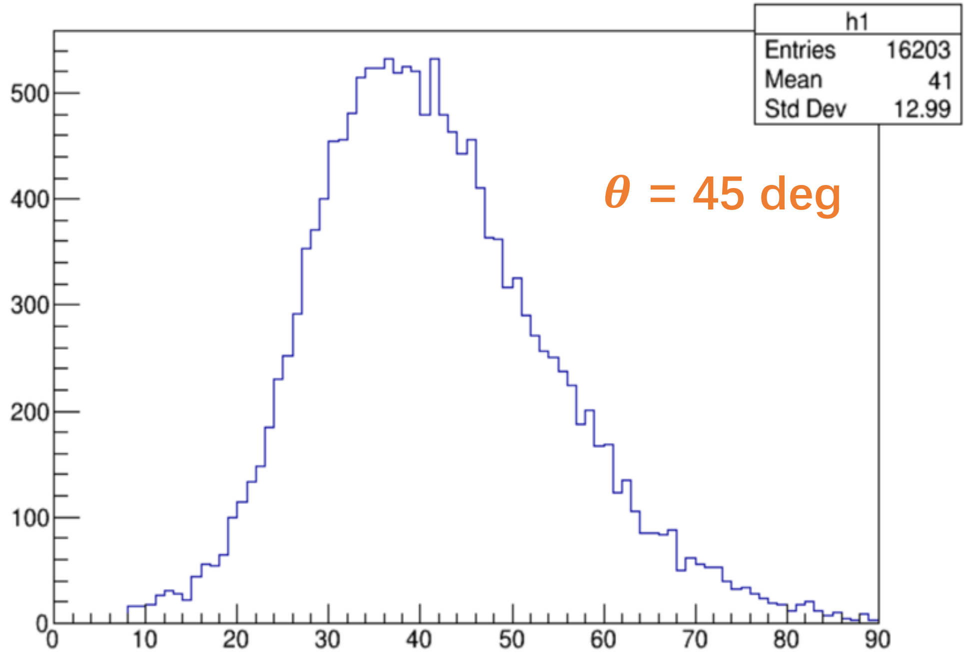
- Apply the algorithm based on 1st and 2nd derivatives on data
- Demonstrated by an example from the 2nd round test beam data



# $N_{cls}$ distributions

- A very preliminary comparison between $\theta = 0° \,\&\, 45°$
- Next to do
  - Optimize the event selection
  - Optimize the parameters of the peak finding algorithm
  - Apply a discrimination algorithm to remove secondary peaks
  - Try ML algorithm



# Collaborations on MOST-MAECI application

(China-Italy Science and Technology Cooperation)

- **Project: INTRePID (INnovative TRacking and Particle IDentification)**
  - Cluster counting and cluster timing for particle identification and improved spatial resolution in drift chambers at the next generation of lepton colliders

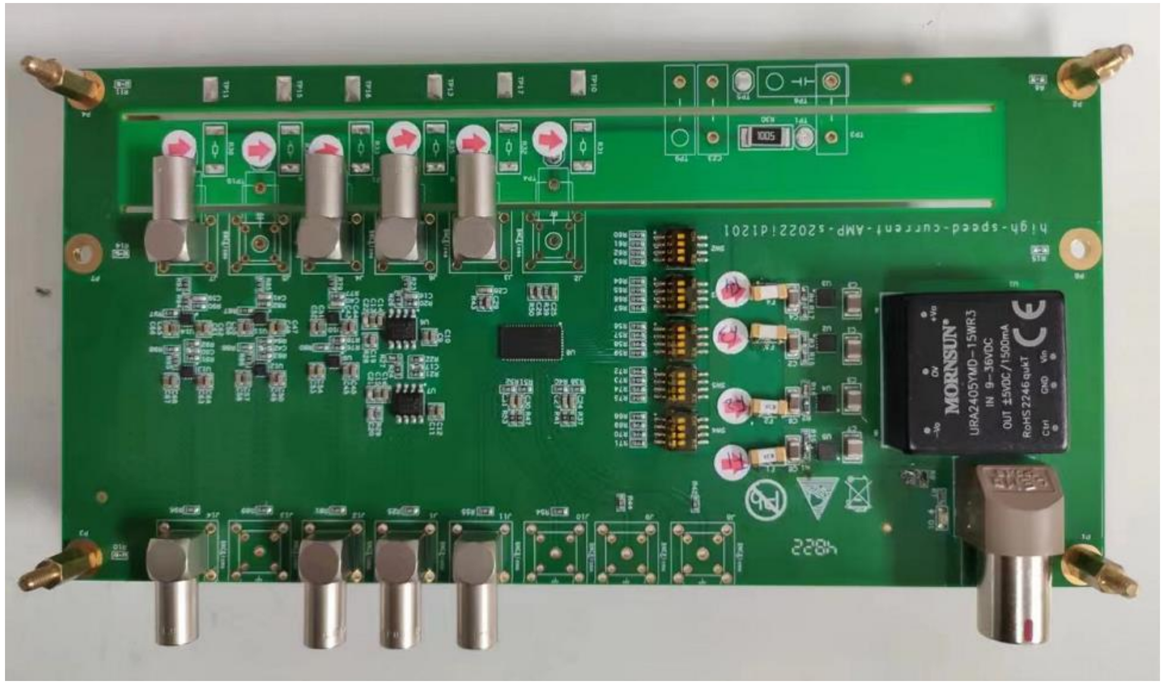
- **Cooperation between INFN and IHEP on**
  - Beam tests for the application of the cluster counting/timing techniques
  - Simulation and reconstruction
  - Design and deployment of real-time ML algorithms on FPGAs

# Progress on hardware: new preamplifiers

- High bandwidth current sensitive preamplifiers based on LMH6522 and AD8099 were designed

- Simulation results show the bandwidth and the gain meet the test requirements

- Test is on going. Preliminary tests verify that the gain is big enough, but noise should be further reduced



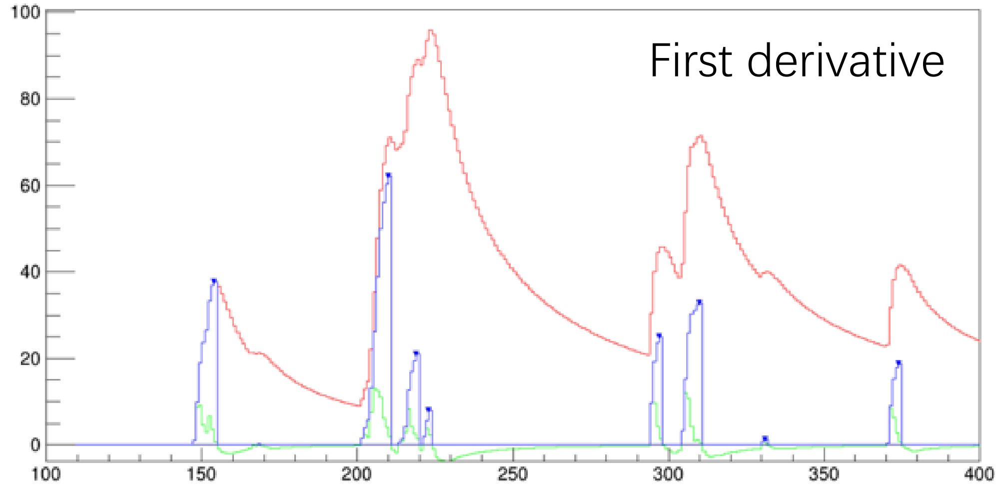
# Summary and outlook

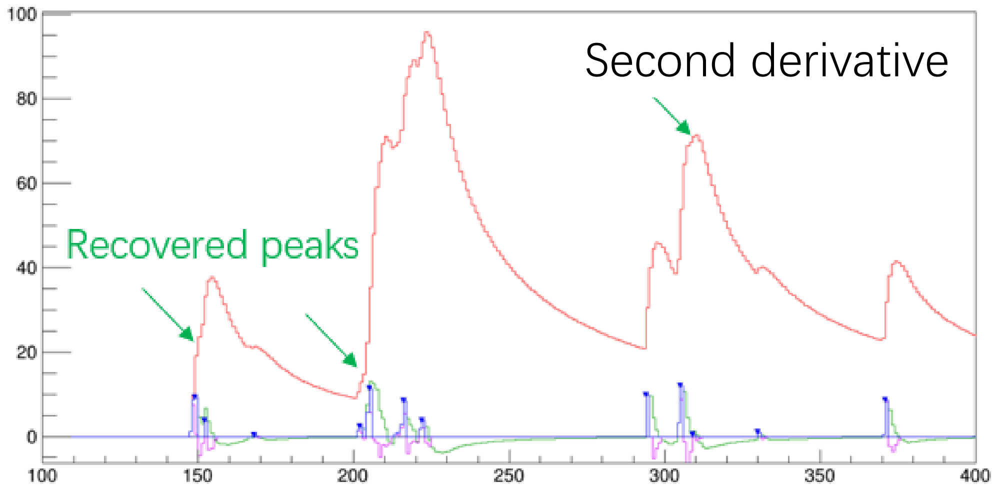
- **A machine learning based reconstruction algorithm is developed**
- **Simulation is tuned with the experimental data**
- **Plans**
  - Better understand the beam test data and fine tuning the MC simulation
  - Optimize the reconstruction algorithm for beam test
  - Extract dN/dx parameters from full simulation and perform physics studies

# Backup

# Reconstruction algorithm with derivatives

- **Peak finding algorithm based on 1st and 2nd order derivatives**
  - Fast and efficient
  - Good pile-up recovery ability on the rising edge




Pile-up on the falling edge is easier to recover. However, it is not the case for pile-up on the rising edge.

# Beam test objectives

## Beam test plans:

1. First of all, need to demonstrate the **ability to count clusters**:
   at a fixed $\beta\gamma$ (e.g. muons at a fixed momentum) count the clusters by
   - doubling and tripling the track length and changing the track angle;
   - changing the gas mixture.
2. Establish the **limiting parameters** for an efficient cluster counting:
   - cluster density (by changing the gas mixture)
   - space charge (by changing gas gain, sense wire diameter, track angle)
   - gas gain saturation
3. In optimal configuration, **measure the relativistic rise as a function of $\beta\gamma$**, both in **$dE/dx$** and in **$dN_{cl}/dx$**, by scanning the muon momentum from the lowest to the highest value (from a few GeV/c to about 250 GeV/c at CERN/H8).
4. Use the experimental results to fine tune the predictions on performance of **cluster counting** for **flavor physics** and for **jet flavor tagging** both in **DELPHES** and in **full simulation**

Slide from Franco

# Space charge from BESIII dE/dx