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

Refined drift chamber simulation in the CEPC experiment

Wenxing Fang, Xingtao Huang, Weidong Li, Tao Lin, Mengyao Liu, Yuan Ye, Yao Zhang, Xueyao Zhang

On behalf of the CEPC drift chamber working group

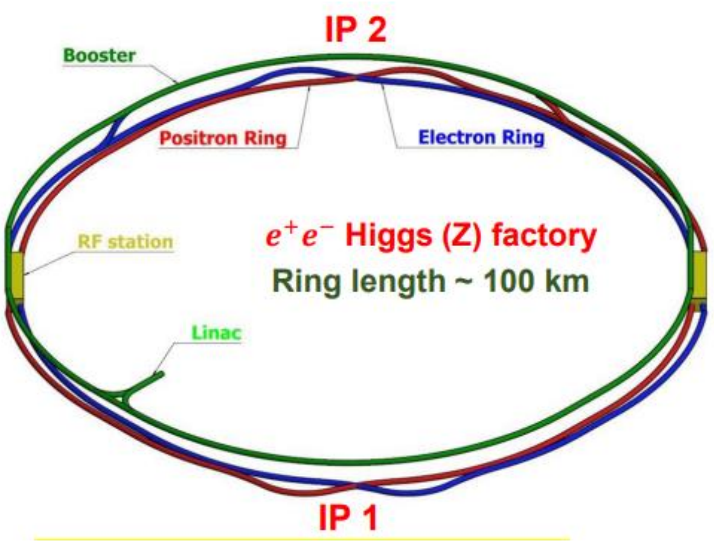
26TH INTERNATIONAL CONFERENCES ON COMPUTING IN HIGH ENERGY PHYSICS & NUCLEAR PHYSICS (CHEP 2023)

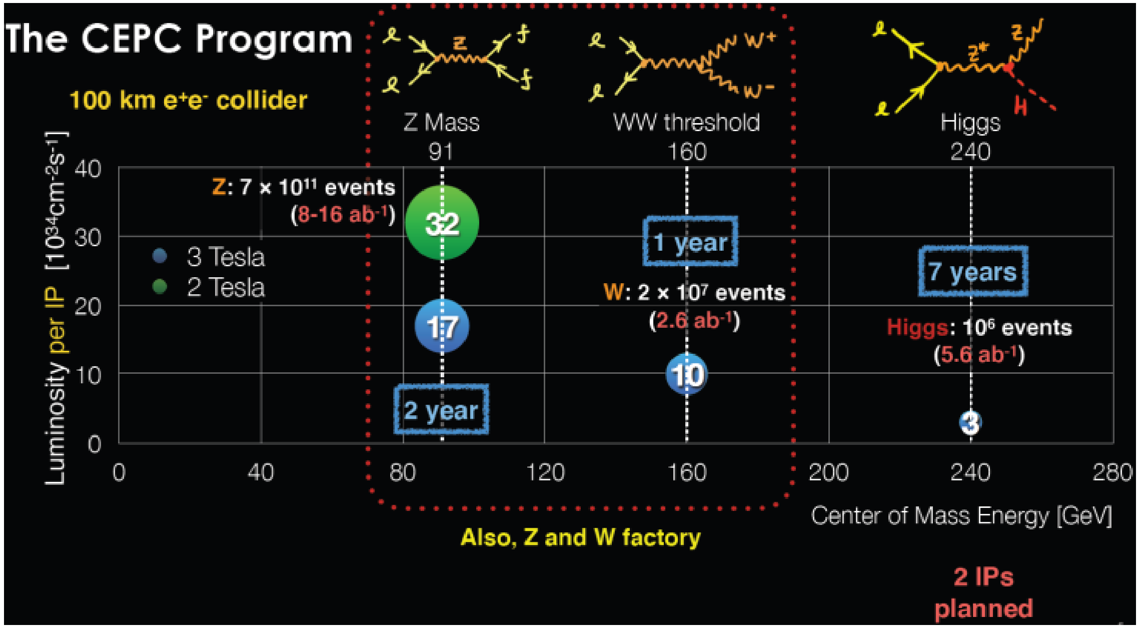
Outline

- Introduction
- Geant4 combined Garfield++ simulation
- Fast waveform simulation using ML
- Method validation using the BESIII experimental data
- Summary and plan

Circular Electron-Positron Collider (CEPC)

- The CEPC is a future circular electron-positron collider in China
- ~1M Higgs at $\sqrt{s} = 240$ GeV (ZH), ~Tera Z at Z pole and ~10M W^+W^- pair and possible $t\bar{t}$ pair production
- Aims at Higgs, EW, flavor physics, BSM

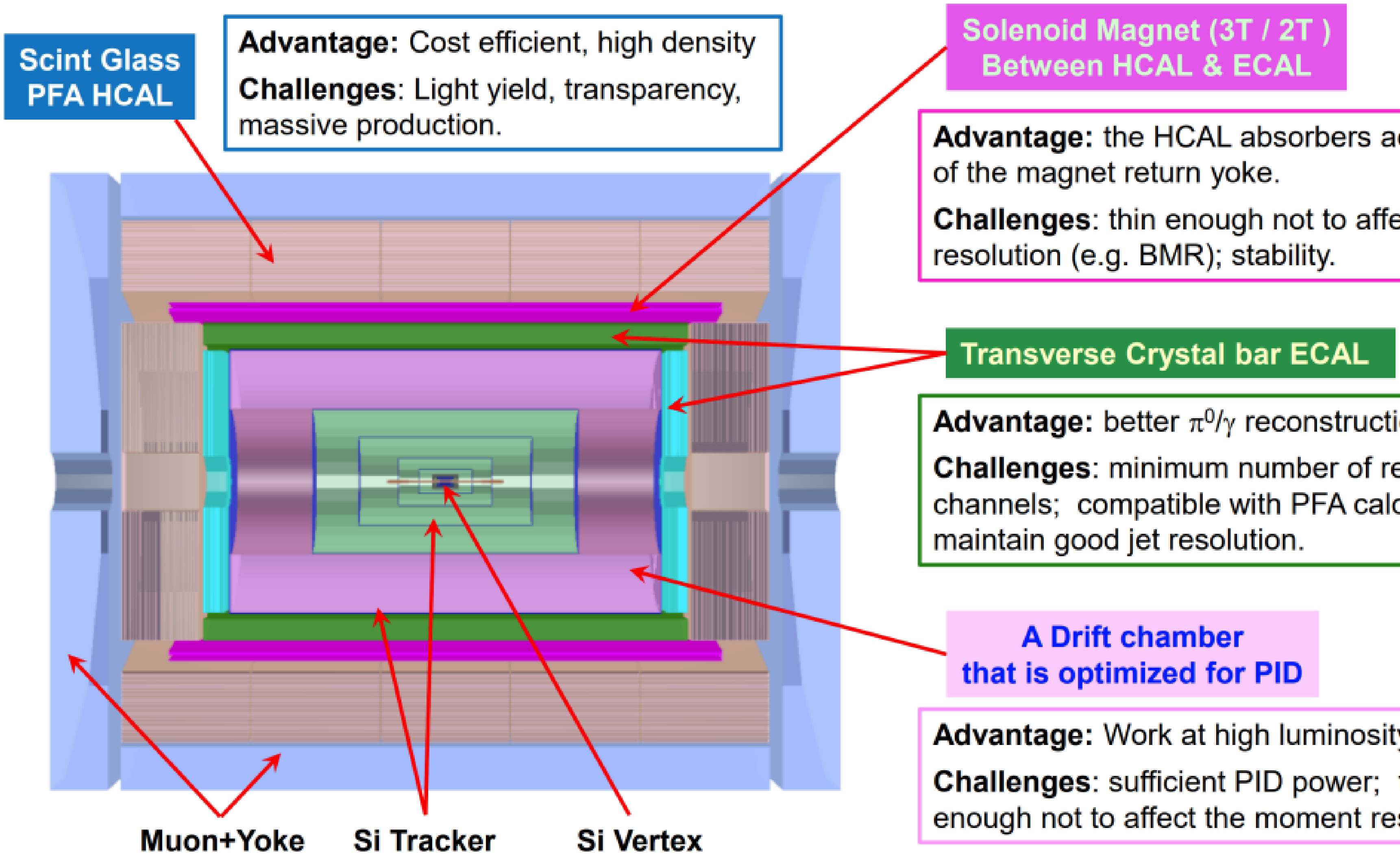




The 4th Conceptual Detector Design

Advantage: Cost efficient, high density

Challenges: Light yield, transparency, massive production.



Advantage: the HCAL absorbers act as part of the magnet return yoke.

Challenges: thin enough not to affect the jet resolution (e.g. BMR); stability.

Advantage: better π^0/γ reconstruction.

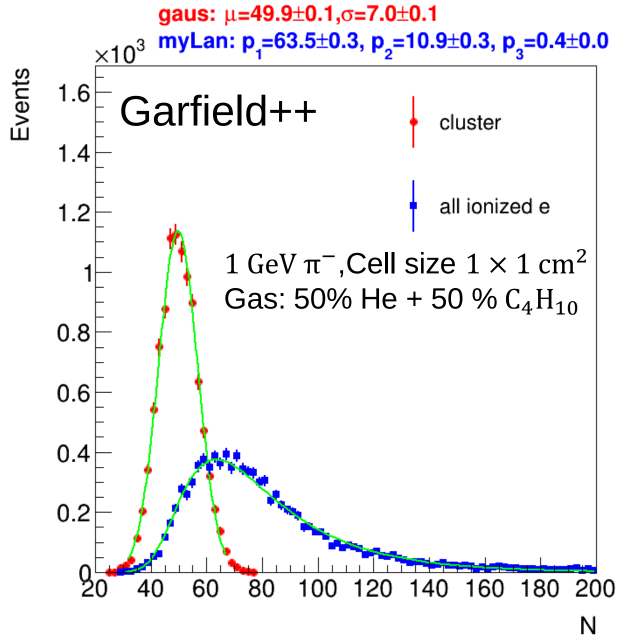
Challenges: minimum number of readout channels; compatible with PFA calorimeter; maintain good jet resolution.

Advantage: Work at high luminosity Z runs

Challenges: sufficient PID power; thin enough not to affect the moment resolution.

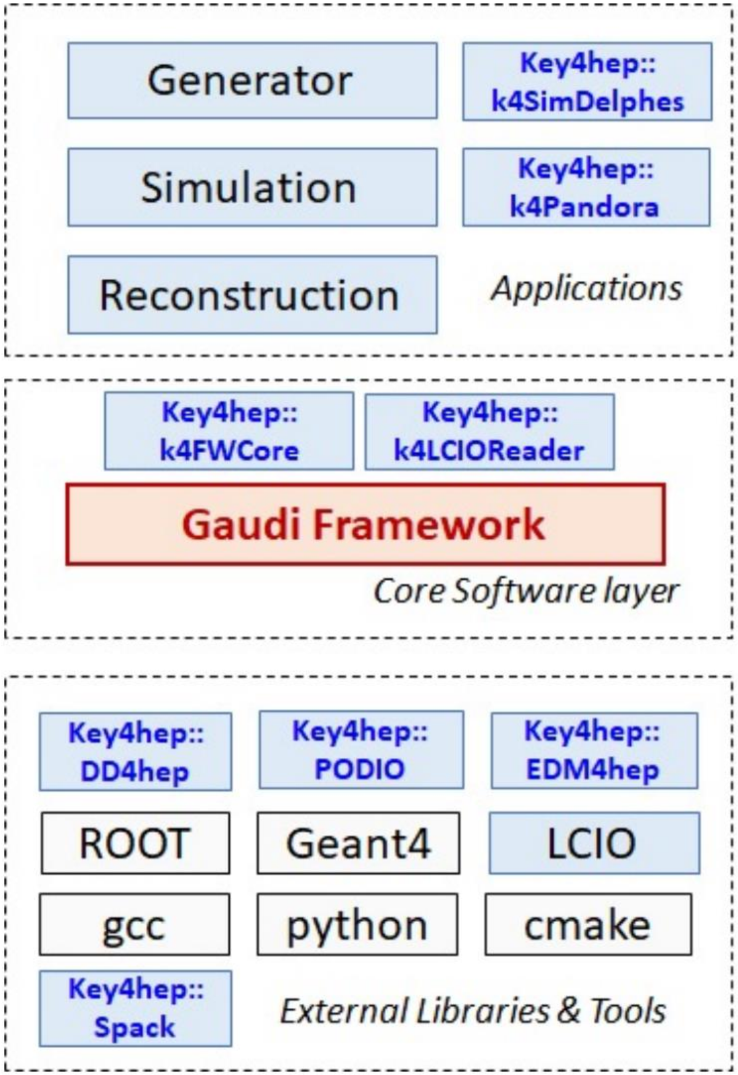
Particle Identification with dE/dx and dN/dx

- Traditionally: using dE/dx
  - The distribution follows Landau distribution (due to the delta-ray production, the gas gain)
  - Using truncated mean to get the expected dE/dx will lose some measured information
  - Resolution is ~6%
- New technique: using Cluster-counting (dN/dx) method
  - The number of clusters of avalanching electrons from primary ionization follows the Poisson distribution
  - Resolution could reach <3%(arXiv:2105.07064)
  - Will be studied in detail in the CEPC drift chamber



The CEPCSW

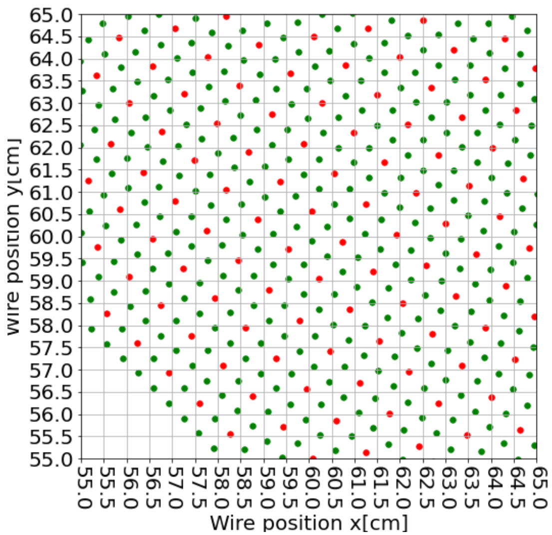
- CEPCSW
  - Offline software for the CEPC experiment
  - Fully integrated with key4hep (talk)
  - Gaudi: defines interfaces to all software components and controls their execution
  - EDM4hep: the generic event data model for HEP experiments
  - DD4hep: geometry and magnetic field description
  - CEPC-specific software: generator, Gean4 simulation, beam background mixing, fast simulation, machine learning interface, etc.



Drift Chamber Parameters in the CEPCSW

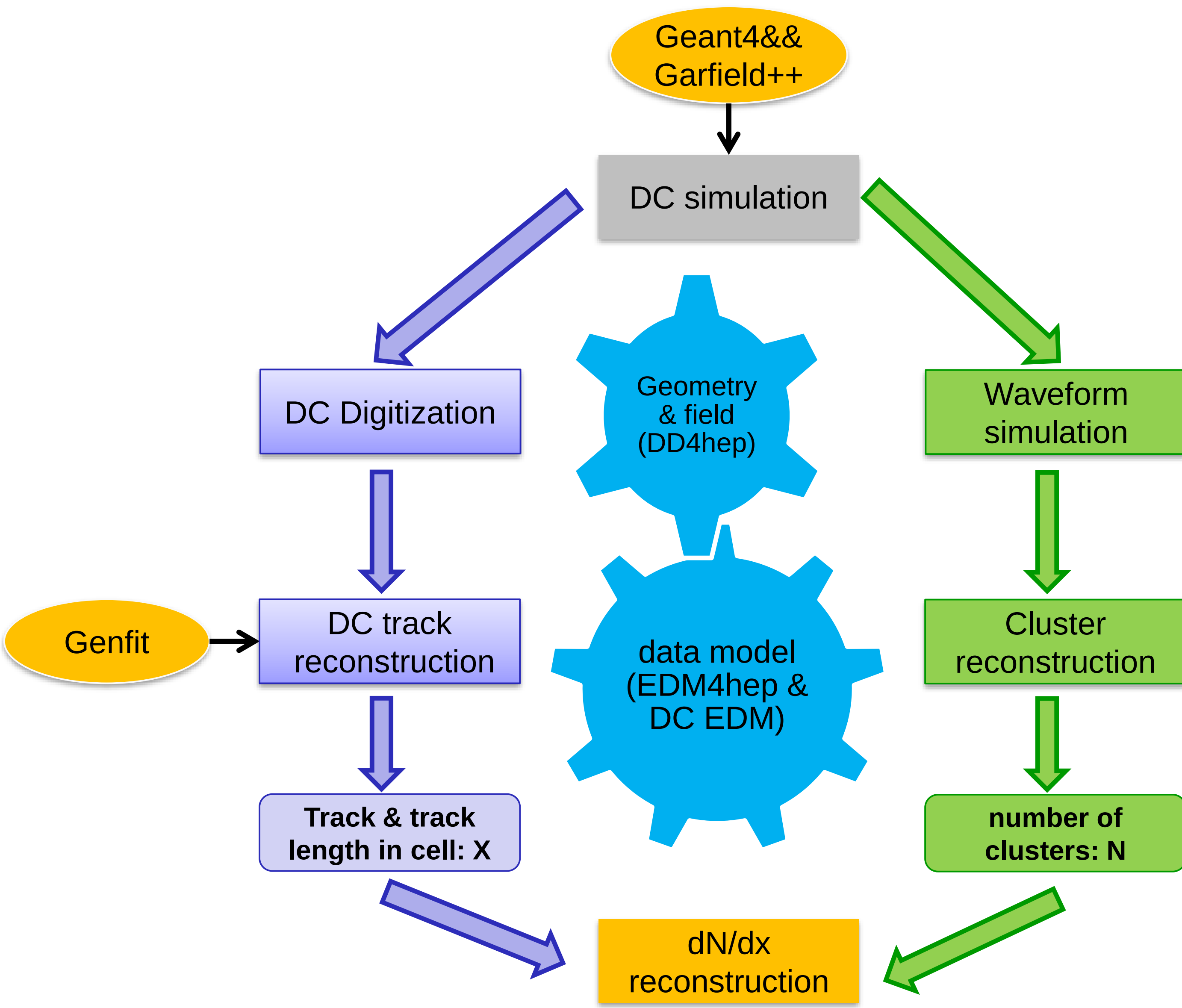
- By using the DD4hep, the DC geometry configuration can be changed easily

Half length	2980 *mm*
Inner and outer radius	800mm to 1800 *mm*
# of Layers	100/55
Cell size	~10mmx10mm/18mmx18mm
Gas	He:iC_4H_{10}=90:10
Single cell resolution	0.11 *mm*
Sense to field wire ratio	1:3
Total # of sense wire	81631/24931
Stereo angle	1.64~3.64 *deg*
Sense wire	Gold plated Tungsten ϕ=0.02*mm*
Field wire	Silver plated Aluminum ϕ=0.04*mm*
Walls	Carbon fiber 0.2 *mm*(inner) and 2.8 *mm*(outer)



Cell structure

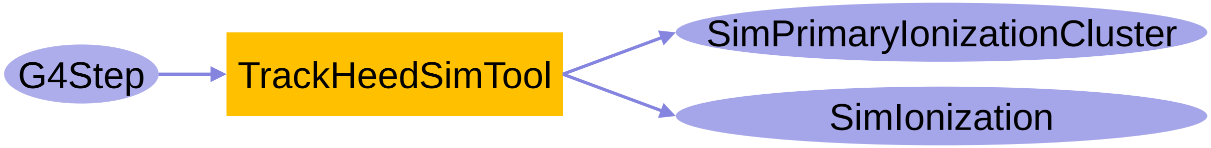
Drift Chamber Simulation and Reconstruction Flow



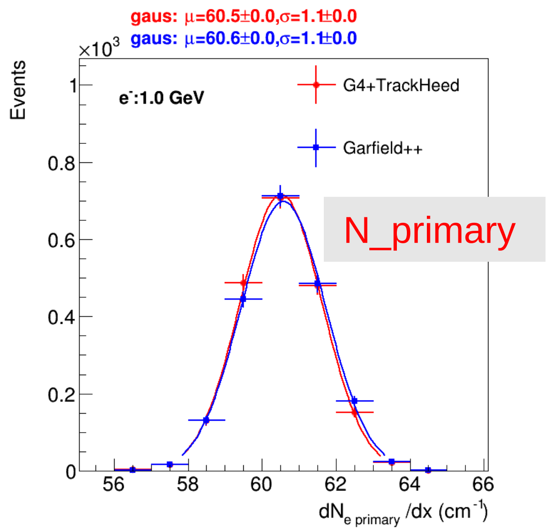
For dN/dx study, both primary ionization reconstruction and track reconstruction are important

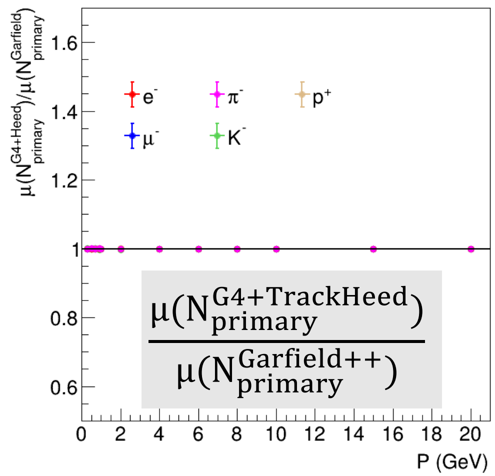
Drift Chamber Simulation

- The precise DC simulation:
  - Garfield++: ionization process, drift time, waveform
  - Geant4: particle transportation, interaction with detector material, decay, …
- Method: combining Geant4 and Garfield++ at G4Step level
- TrackHeedSimTool (Gaudi tool) is created for this task
  - Input: G4Step information (particle type, initial position and momenta, ionization path length)
  - Use TrackHeed (used by Garfield++) to simulate one step length ionization (new API contributed to Garfield++ PR)
  - Output: primary and total ionization information (contains position, time, cell id), saved in the EDM
  - The kinetic energy of G4Track will be updated according to the energy loss in the ionization

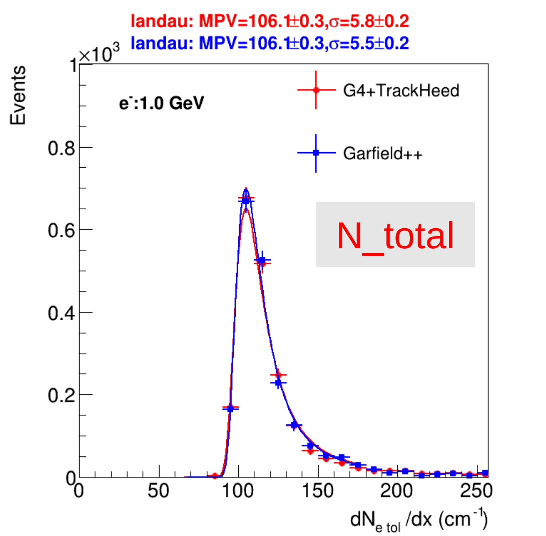


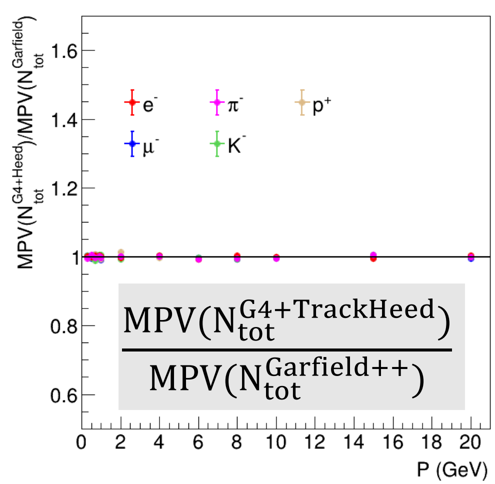
Ionization Simulation Performance Check





- Gas: 50% He + 50 % C_4H_{10}

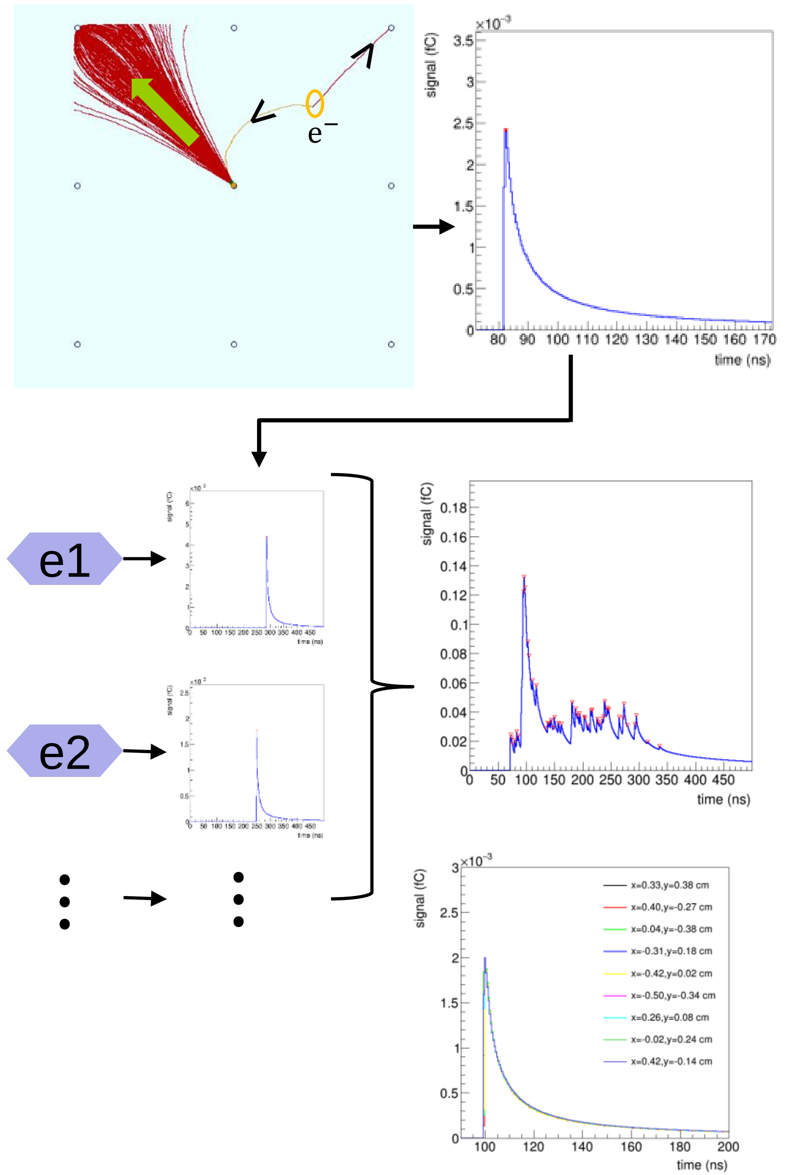




- Consistent with Garfield++ standalone simulation results

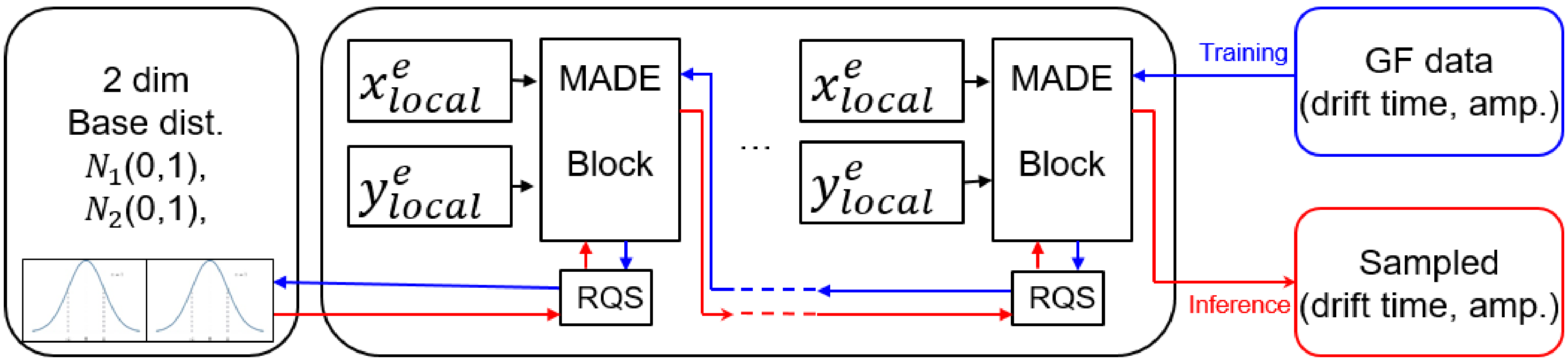
Waveform Simulation

- Simulating the drift and avalanche processes of ionized electron and ions. Getting final waveform
- Using Garfield++: extremely time consuming, could take $\mathcal{O}(1)$ to $\mathcal{O}(10)$ seconds just for one electron (few hours for one track)
- Going to use a fast simulation method
  - For each electron, simulate the drift time and amplitude of the pulse (as the shapes are very similar)
  - Piling up all pulses from the same drift chamber cell gives the final waveform



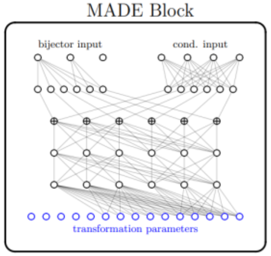
The NN Model

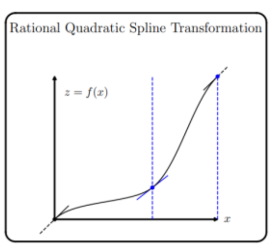
- The Normalizing Flow is adopted:
  - Learning bijective transformation between two distributions
  - Compared to GAN, it is much easier to train (stable and convergent)
  - Reference to the CaloFlow, a similar model is used, RQS (for transformation)+MADE block (for the parameters of RQS)



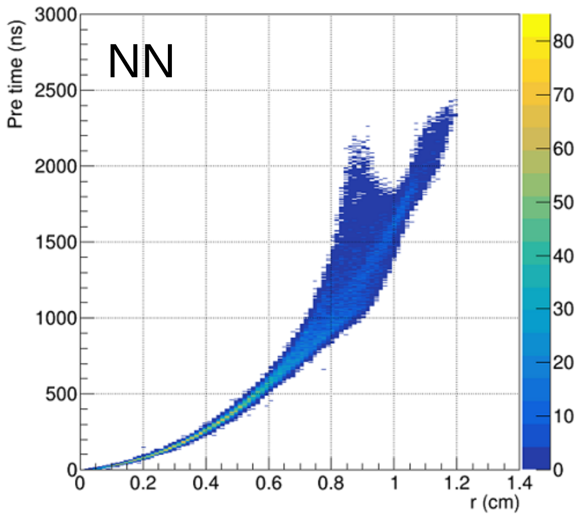
Base distribution	Number of MADE blocks	Layer sizes			Number of RQS bins
		input	hidden	output	
1-dim Standard Normal	6	64	3×64	23	8

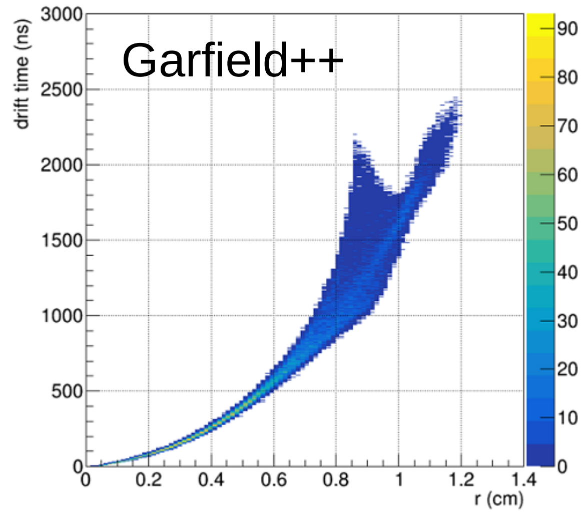
MADE Block



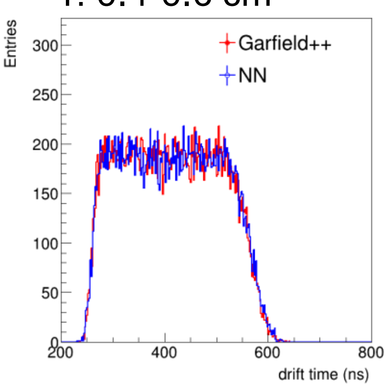


NN Simulation Performance

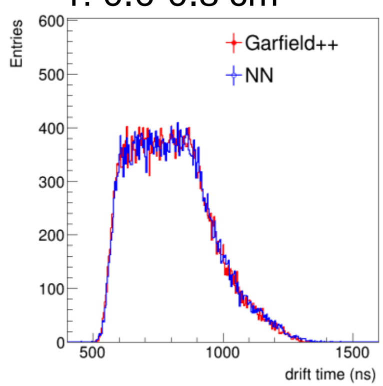




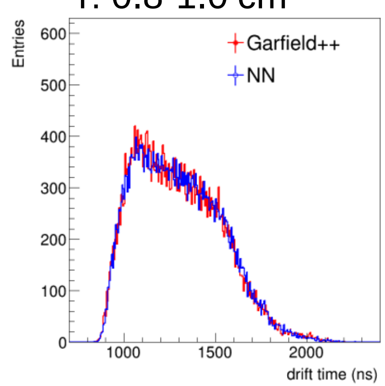
r: 0.4-0.6 cm



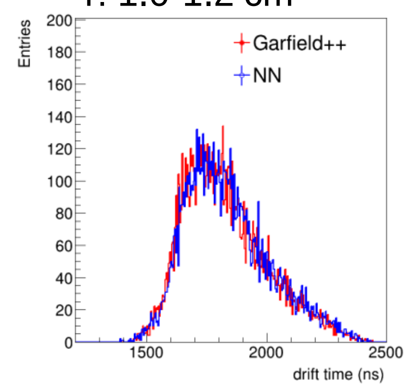
r: 0.6-0.8 cm



r: 0.8-1.0 cm



r: 1.0-1.2 cm



- Good agreement between the NN and Garfield++ simulation

Method Validation Using the BESIII Data

- The BESIII is an experiment studying particle physics in the tau-charm region
- The main drift chamber is used for tracking and PID with dE/dx
- Validate the used NN method for real data:
  - learning drift time distribution
  - dataset: radiative Bhabha event

Prepare experimental data for training

↓

Learn drift time distribution: $D_{drift\ time} = f(doca, eangle, layer)$

↓

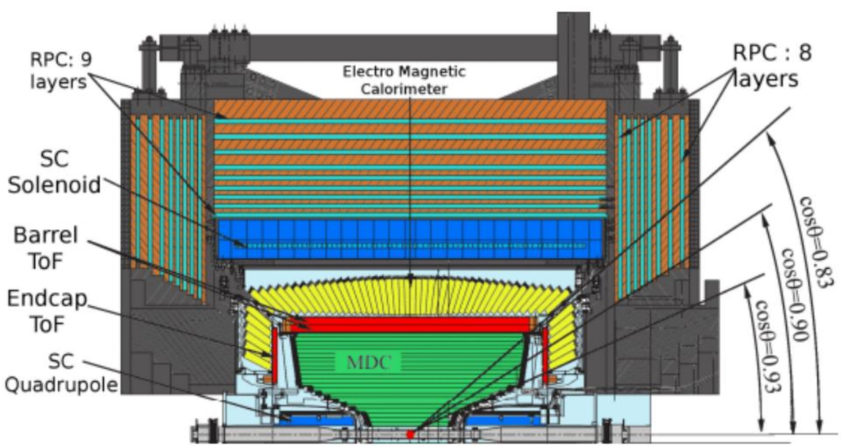
Check $D^{data}_{drift\ time}$, $D^{NN\ sim.}_{drift\ time}$

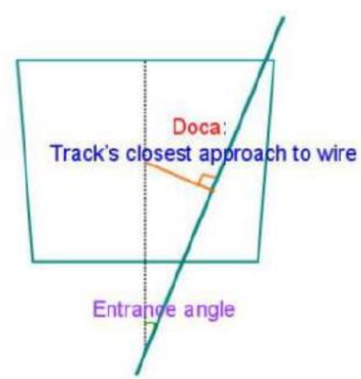
↓

check the X-T cure: C^{data}_{X-T} , $C^{NN\ sim.}_{X-T}$

↓

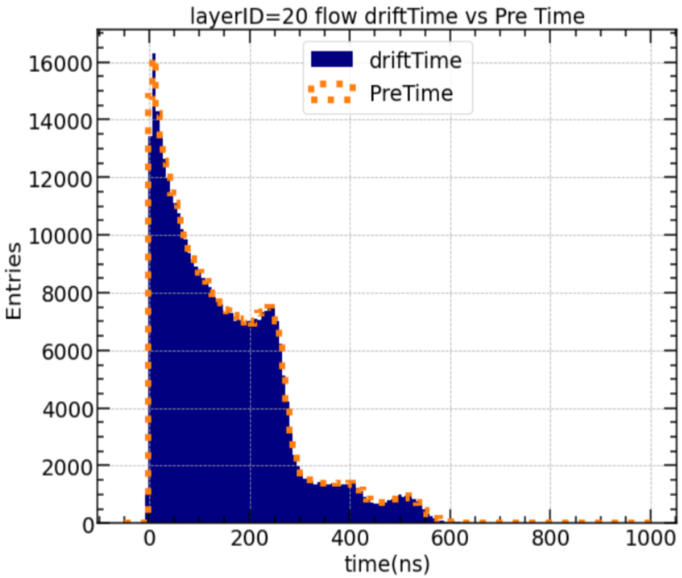
check spatial resolution: using C^{data}_{X-T} , $C^{NN\ sim.}_{X-T}$

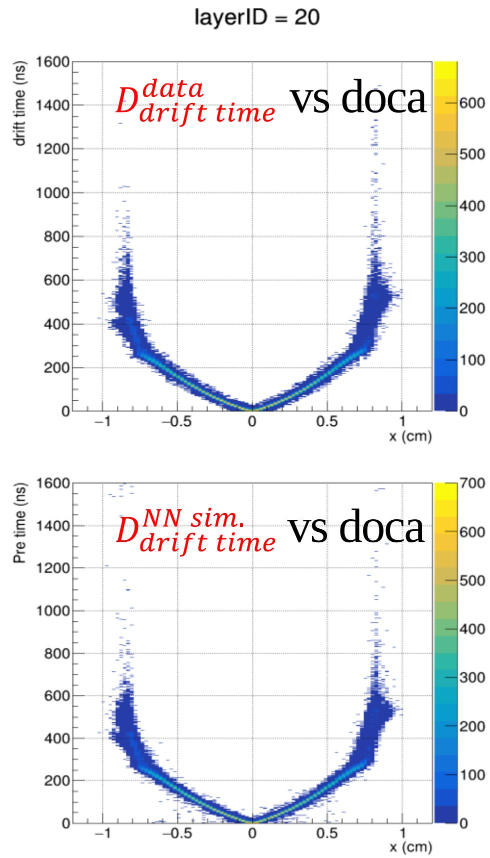


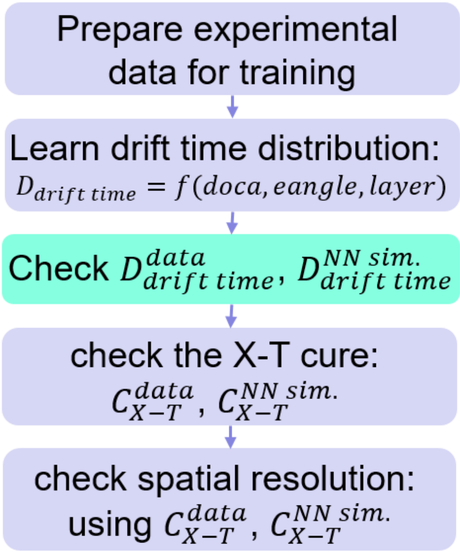


NN Simulation Performance

- Compared the drift time distribution between data and NN simulation



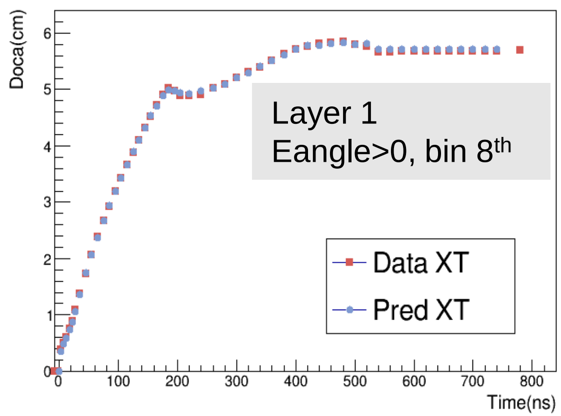


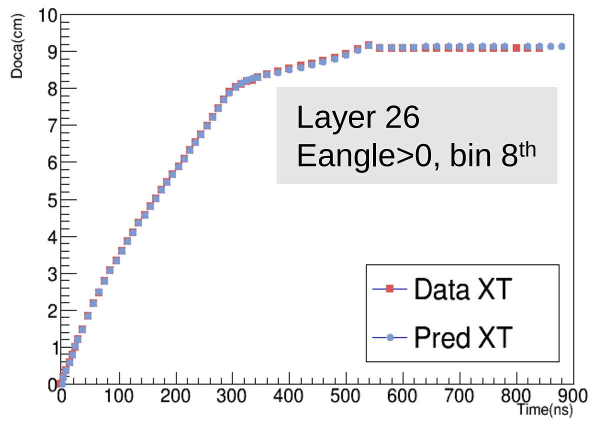


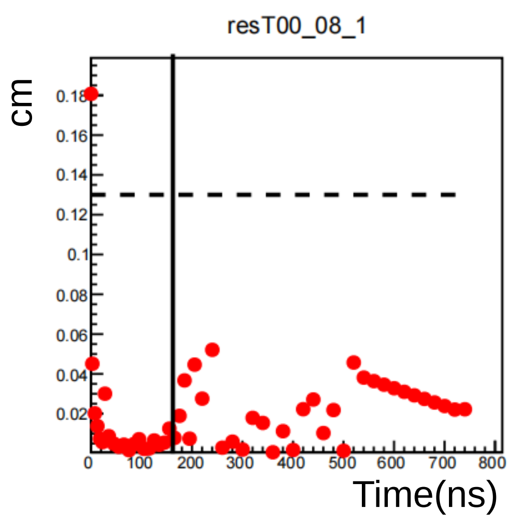
- Consistent between data and NN simulation

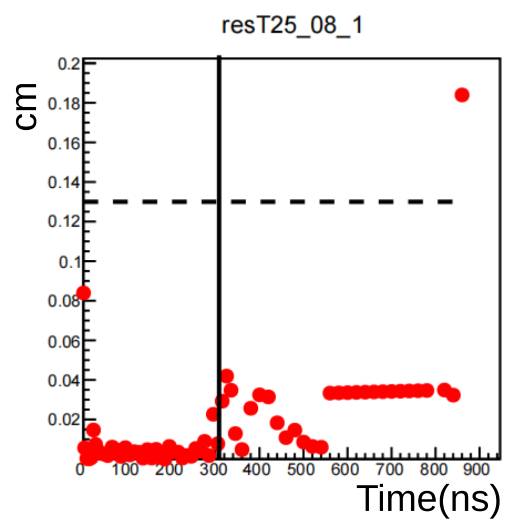
X-T Relationship Curve

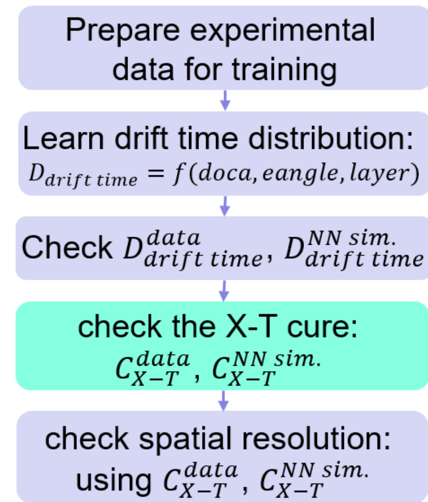
- Check the X-T curve from the experimental data and NN simulation







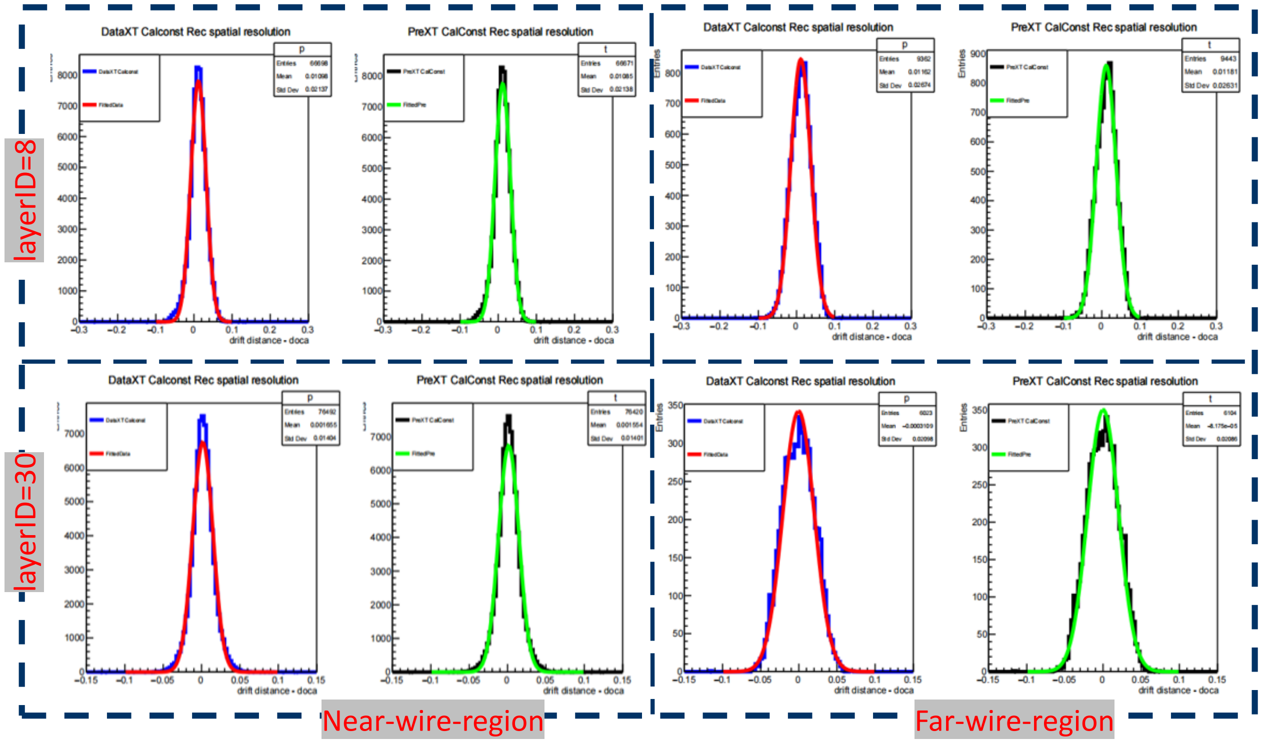




- The mean Docas versus time are obtained for different layers and entrance angle
- red for C^{data}_{X-T}, blue for $C^{NN\ sim.}_{X-T}$
- Difference is < 40 μm

Spatial Resolution Check

- Checked the spatial resolution of reconstructed track, using C_{X-T}^{data} and $C_{X-T}^{NN\ sim.}$



Prepare experimental data for training

Learn drift time distribution: $D_{drift\ time} = f(doca, eangle, layer)$

Check $D_{drift\ time}^{data}$, $D_{drift\ time}^{NN\ sim.}$

check the X-T cure: C_{X-T}^{data}, $C_{X-T}^{NN\ sim.}$

check spatial resolution: using C_{X-T}^{data}, $C_{X-T}^{NN\ sim.}$

layerID	Near-wire-region (C_{X-T}^{data} v.s. $C_{X-T}^{NN\ sim.}$)		Far-wire-region (C_{X-T}^{data} v.s. $C_{X-T}^{NN\ sim.}$)	
	mean	sigma	mean	sigma
8	0.01096/0.01085	0.02137/0.02138	0.01162/0.01181	0.02674/0.02631
30	0.001055/0.001564	0.01404/0.01401	-3.1e-4/-5.1e-5	0.02060/0.02086

- Consistent spatial resolution results by using C_{X-T}^{data} and $C_{X-T}^{NN\ sim.}$
- The NN can learn the drift time from the real data

Summary and Plan

- Precise simulation of the drift chamber is important for the CEPC experiment
- The Garfield++ was integrated with the Geant4-based simulation in the CEPCSW to simulate detector response inside the cell of the drift chamber
- Fast simulation using machine learning method has been developed, 2 orders of magnitude speedup can be achieved
- The method is validated by using real data from the BESIII
- Plan:
  - Consider effects from the electronic response, electronic noise
  - Check the space charge effect, arXiv:2211.06361, RPC simulation with space charge effect (dynamic update of the electric field)
  - Compare the results with test beam data
  - Check track and dN/dx reconstruction performance

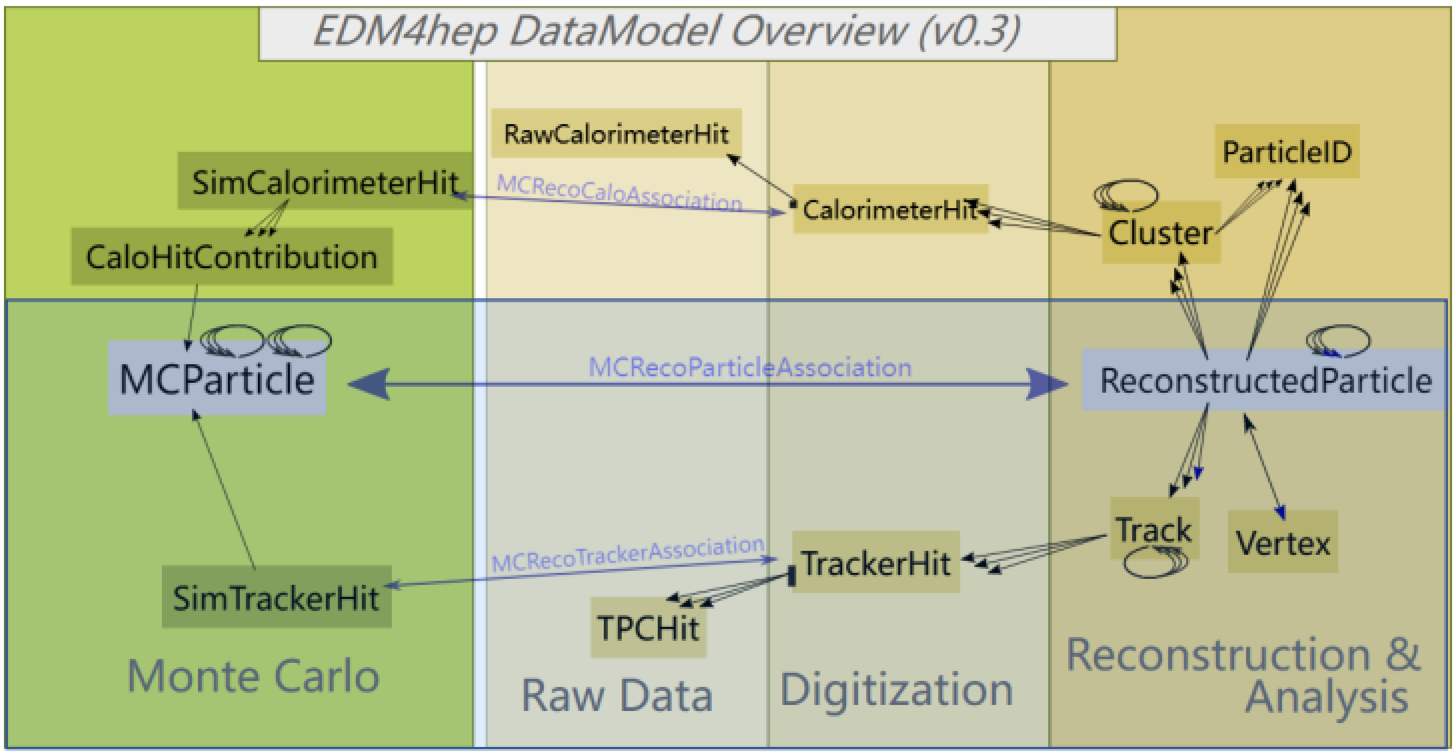
Acknowledgement

- National Natural Science Foundation of China
- The work has been performed in collaboration with AIDAinnova (funded by the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement No 101004761)
- CAS Center for Excellence in Particle Physics
- Ministry of Science and Technology of the People’s Republic of China

Thanks for your attention！

Back up

Event Data Model

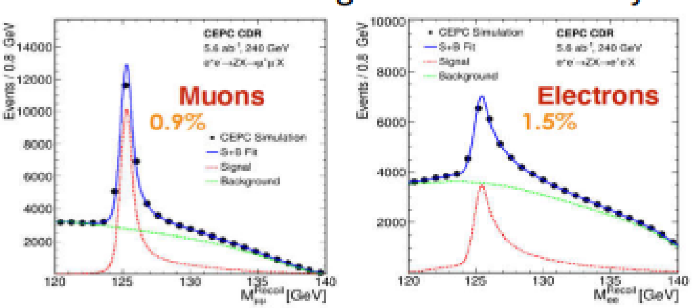


Physics requirement

Detector performance requirements in CDR

Physics process	Measurands	Detector subsystem	Performance requirement
$ZH, Z \to e^+e^-, \mu^+\mu^-$ $H \to \mu^+\mu^-$	$m_H, \sigma(ZH)$ $\mathrm{BR}(H \to \mu^+\mu^-)$	Tracker	$\Delta(1/p_T) = 2 \times 10^{-5} \oplus \frac{0.001}{p(\mathrm{GeV}) \sin^{3/2}\theta}$
$H \to b\bar{b}/c\bar{c}/gg$	$\mathrm{BR}(H \to b\bar{b}/c\bar{c}/gg)$	Vertex	$\sigma_{r\phi} = 5 \oplus \frac{10}{p(\mathrm{GeV}) \times \sin^{3/2}\theta} (\mu\mathrm{m})$
$H \to q\bar{q}, WW^*, ZZ^*$	$\mathrm{BR}(H \to q\bar{q}, WW^*, ZZ^*)$	ECAL HCAL	$\sigma_E^{\mathrm{jet}}/E = 3 \sim 4\%$ at 100 GeV
$H \to \gamma\gamma$	$\mathrm{BR}(H \to \gamma\gamma)$	ECAL	$\Delta E/E = \frac{0.20}{\sqrt{E(\mathrm{GeV})}} \oplus 0.01$

Good EM energy resolution is required for bremsstrahlung radiation recovery



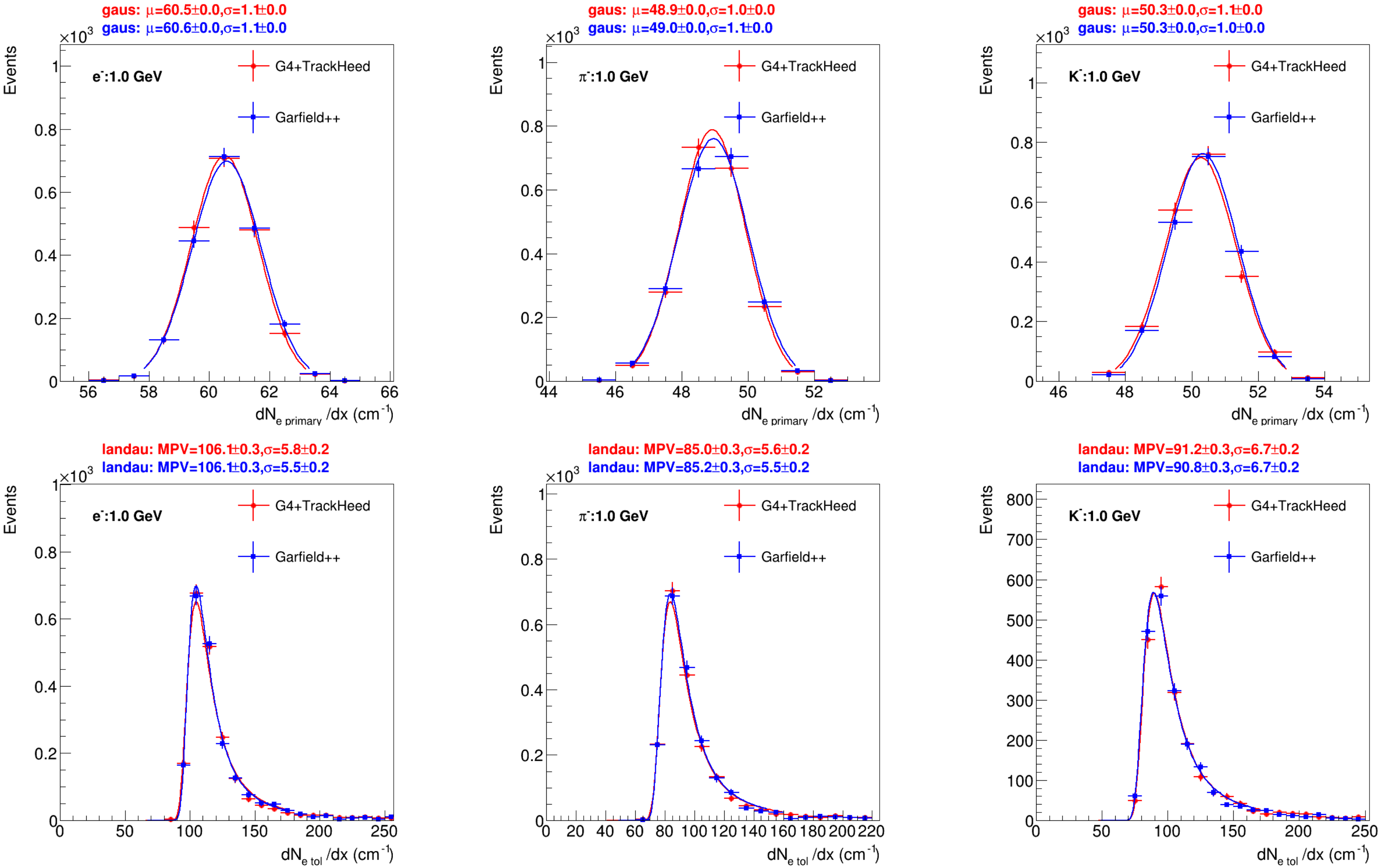
ee → HZ(→ee/μμ), recoil mass against Z boson

EWK physics:
Precise EM measurement

Flavor physics:
Precise EM measurement
Dedicated hadron identification

Ionization simulation performance

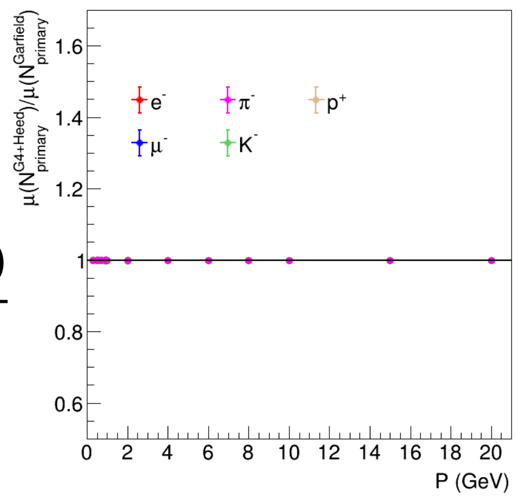
- Gas: 50% He + 50 % C_4H_{10}

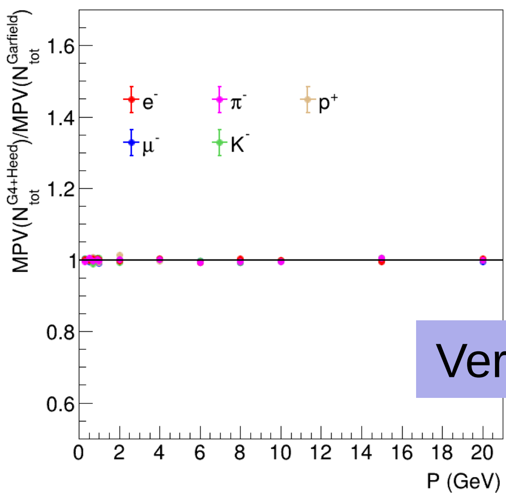


Ionization simulation performance

- Gas: 50% He + 50 % C_4H_{10}

$$\frac{\mu(N_{primary}^{G4+TrackHeed})}{\mu(N_{primary}^{Garfield++})}$$

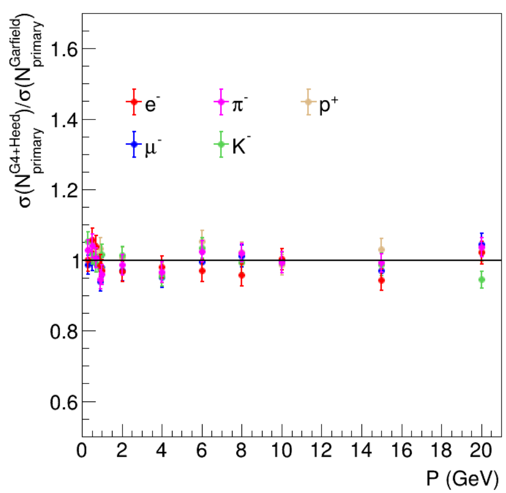


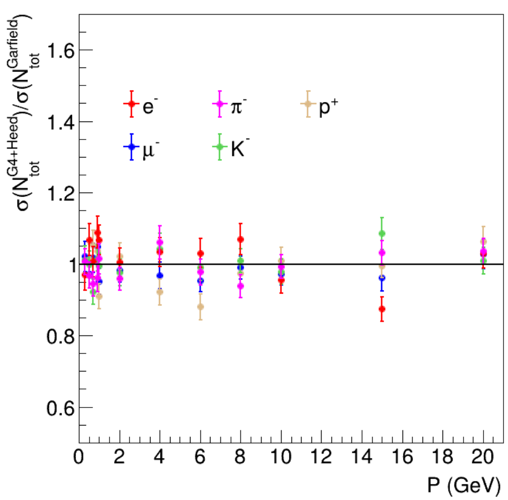


$$\frac{MPV(N_{tot}^{G4+TrackHeed})}{MPV(N_{tot}^{Garfield++})}$$

Very good agreement

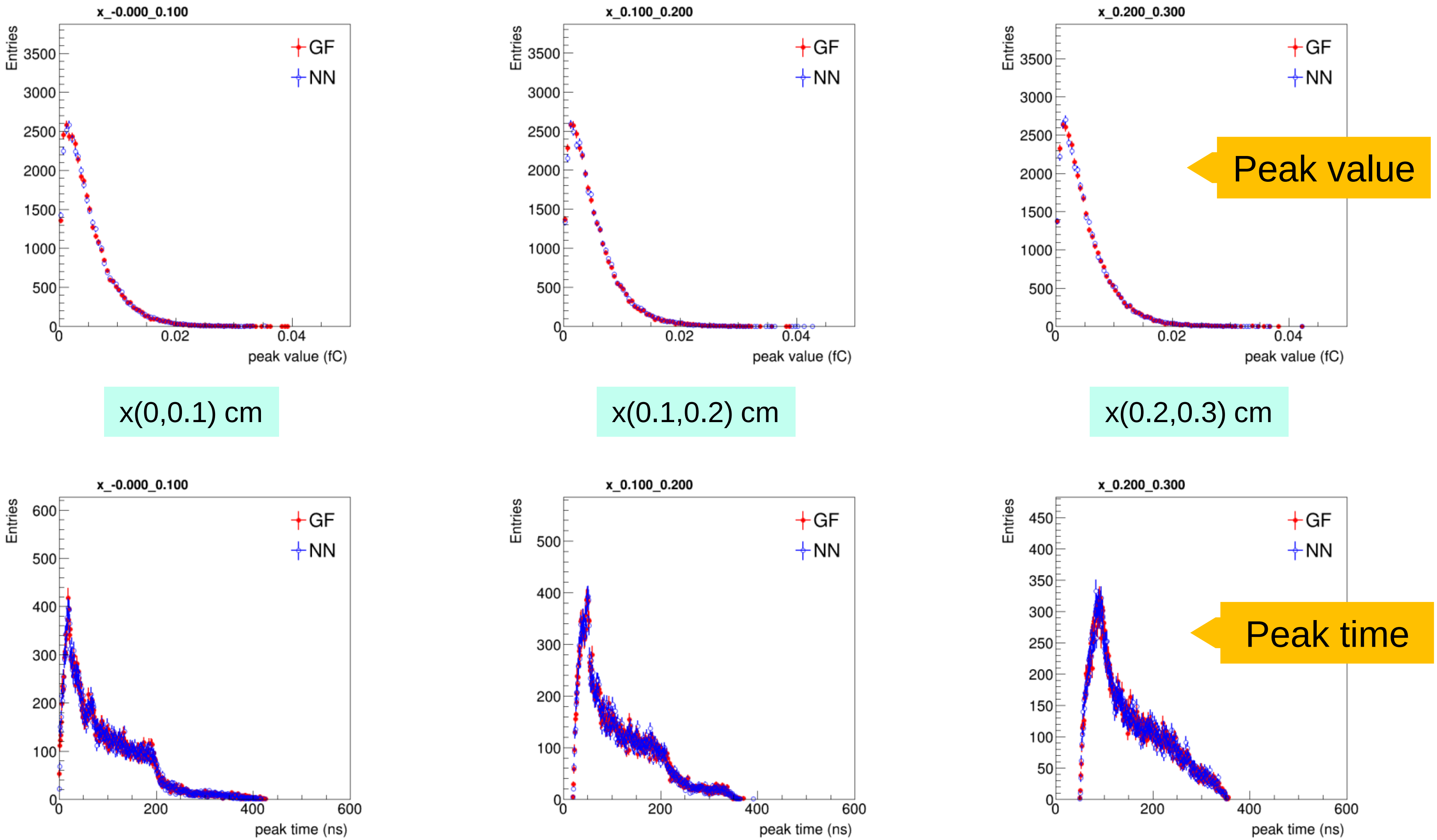
$$\frac{\sigma(N_{primary}^{G4+TrackHeed})}{\sigma(N_{primary}^{Garfield++})}$$





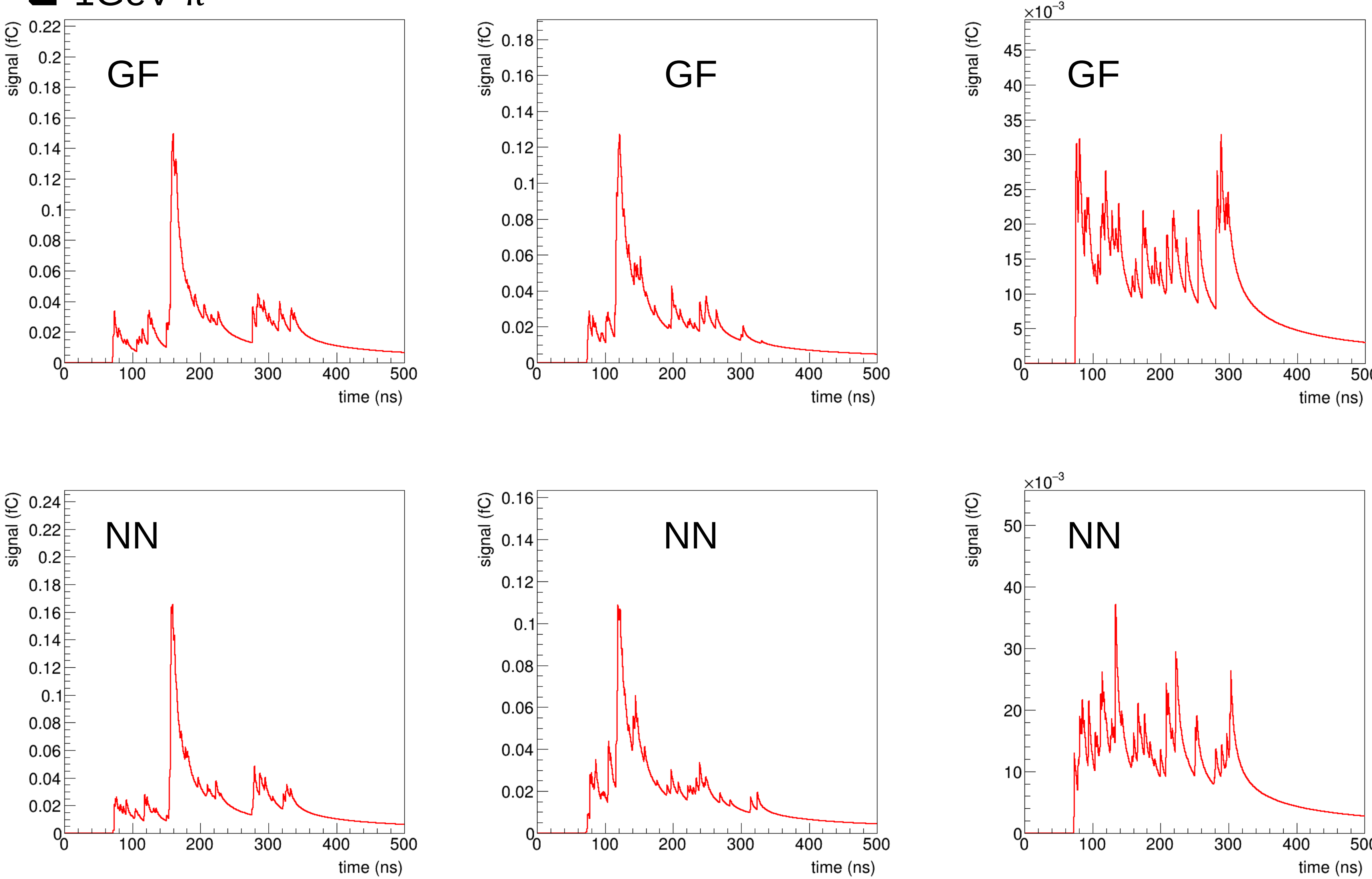
$$\frac{\sigma(N_{tot}^{G4+TrackHeed})}{\sigma(N_{tot}^{Garfield++})}$$

One Dimension Check

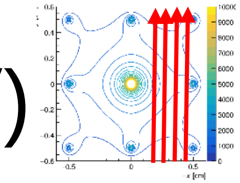


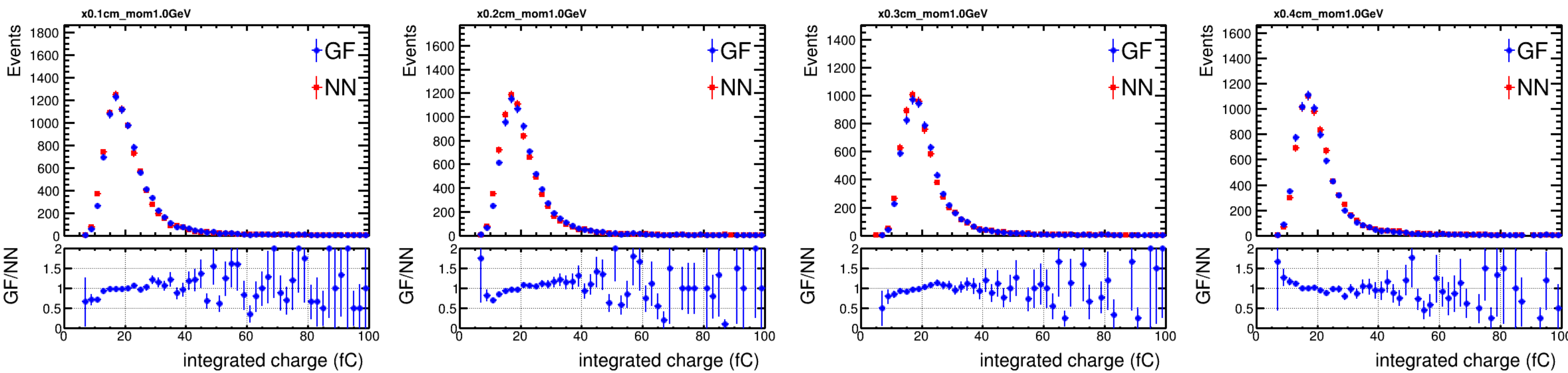
Display of signal response for one cell

- 1GeV π^-



Performance check: electron (1 GeV)



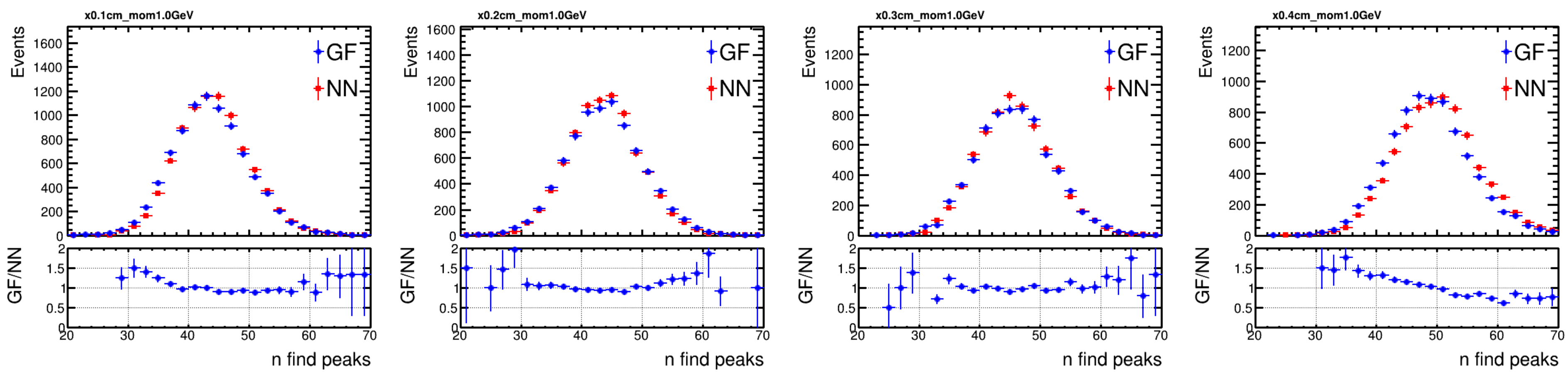


x=0.1 cm

x=0.2 cm

x=0.3 cm

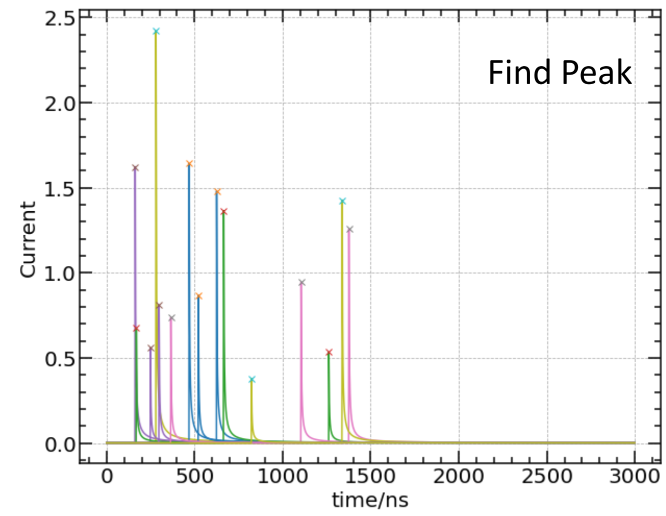
x=0.4 cm



- Checked total integrated charge and number of found peaks(using scipy.signal.find_peaks)
- Good agreement in general (a little bias for the number of found peaks for x=0.4 cm. Could be improved in future), Similar results for other energy points (0.5, 5, and 10 GeV)

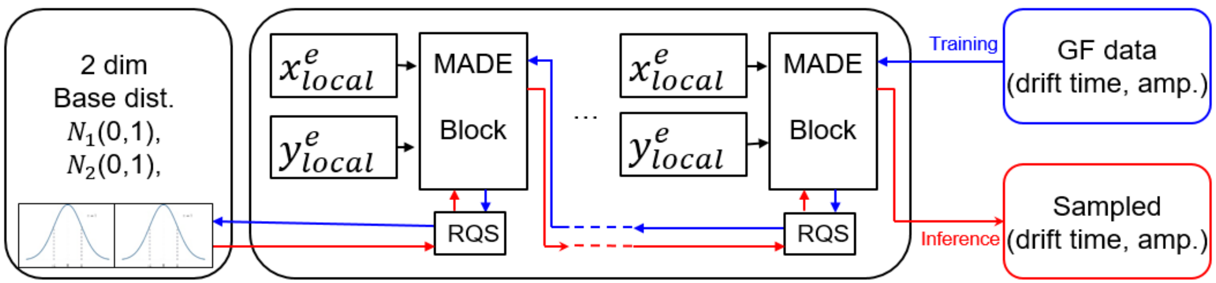
Drift time simulation in CEPC

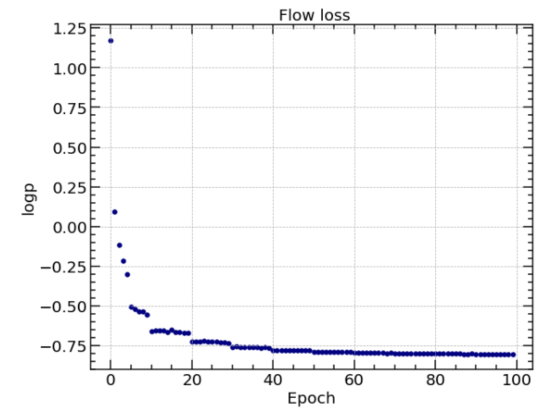
- Using a similar method for the CEPC drift time simulation
- Produce Garfield++ simulation data for training:
  - DC Cell size is 18 mm* 18 mm
  - Gas: 90%He+10%C_4H_{10}
  - Signal wire voltage: 1630 V
  - For each event, an ionized electron is uniformly generated in the cell (x,y) and the pulse is simulated
  - Then peak finding algorithm (scipy.signal.find_peaks()) is used to get pulse time (drift time) and amplitude (used for threshold cut)

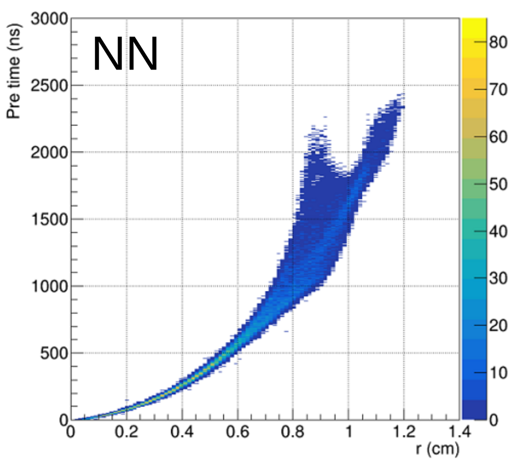


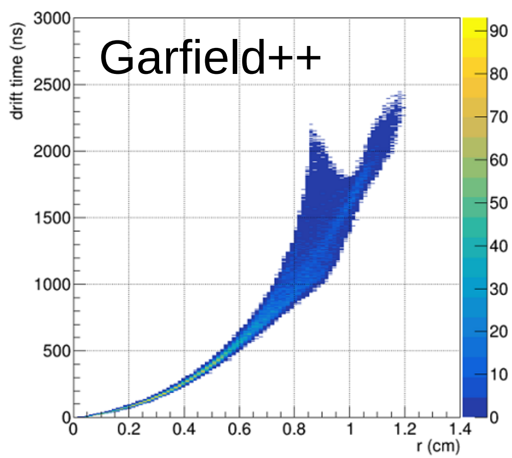
Drift time simulation in CEPC

- Similar NF model is used for training







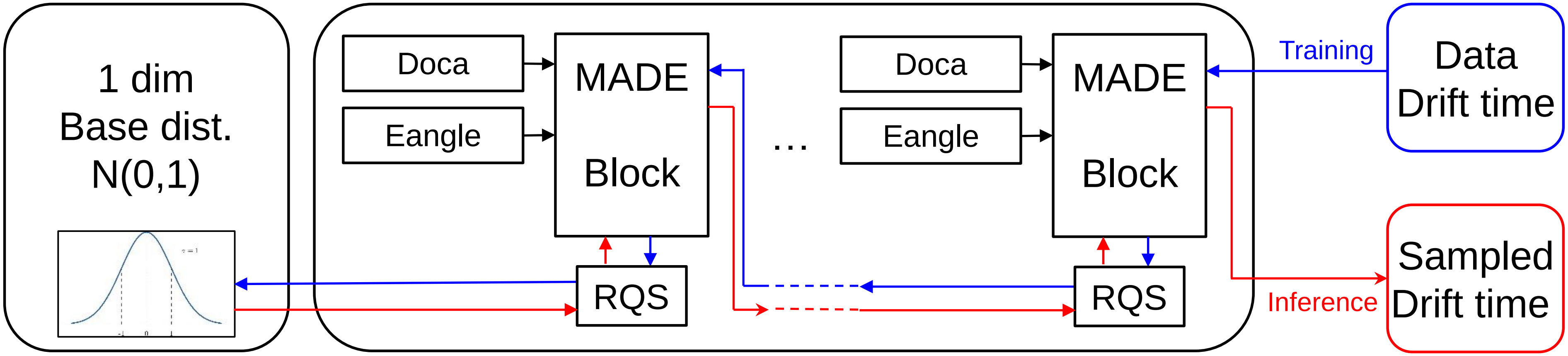


- By using the NN model, one can simulate the drift time and pulse amplitude according to the cell local x,y position of the ionized electron

The NN model

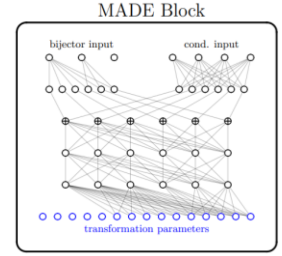
❖ The Normalizing Flow is adopted:

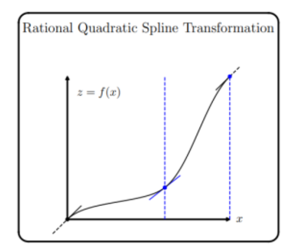
- Learning bijective transformation between two distributions(e.g. drift time ~ N(0,1))
- Compared to GAN, it is much easier to train (stable and convergent)
- Reference to the CaloFlow, a similar model is used, RQS (for transformation)+MADE block (for the parameters of RQS)



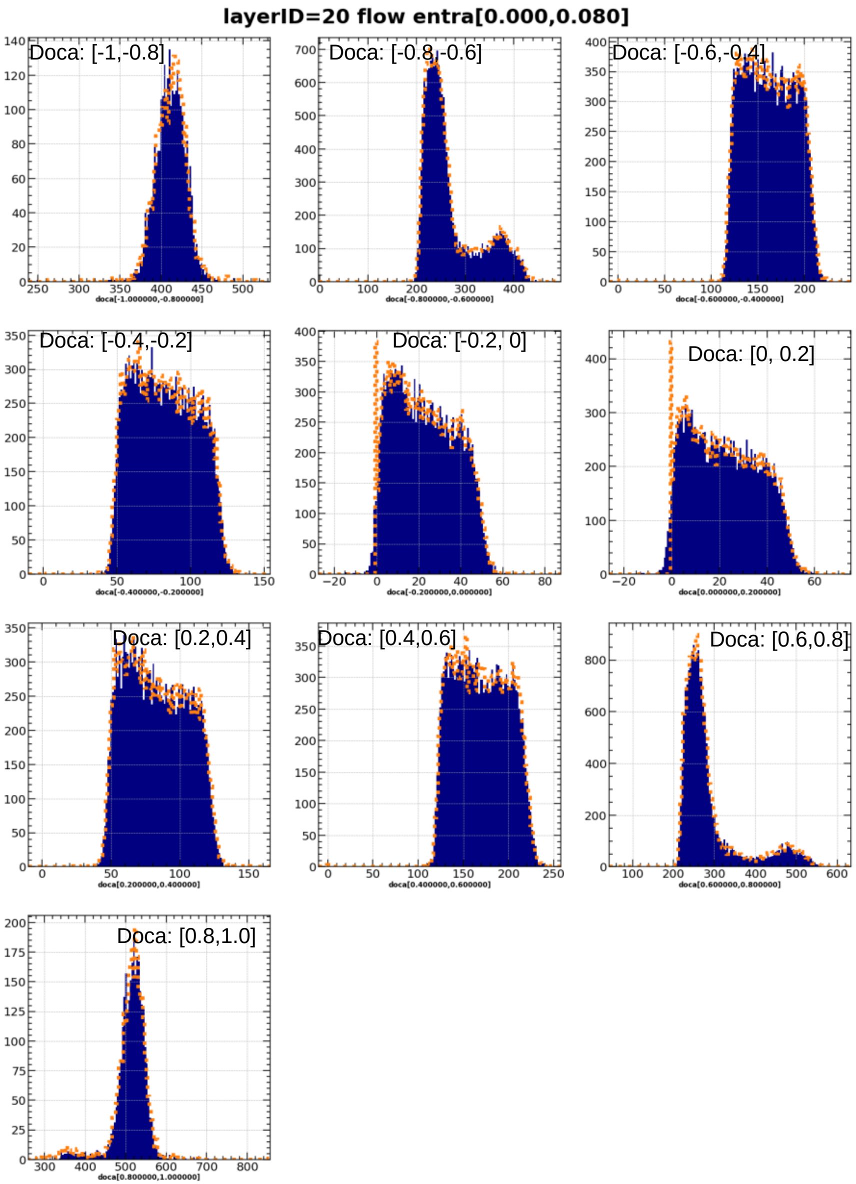
Base distribution	Number of MADE blocks	Layer sizes			Number of RQS bins
		input	hidden	output	
1-dim Standard Normal	6	64	3×64	23	8

MADE Block





NN simulation performance



- A detailed check, in different entrance angles and Doca region.
- In general, good agreement between data and NN simulation
- Drift time distributions are flatter in small Doca regions (|Doca|<0.6) , complex for the large Doca regions (close to cell edge)

Benefits of the new method

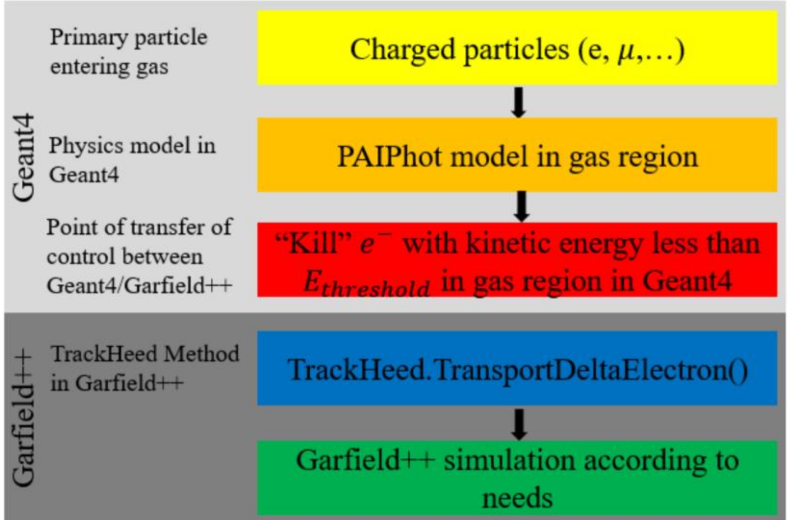
- Average time for one cell ($1 \times 1\mathrm{cm}^2$) simulation for 1 GeV π^- (gas: $50\%\mathrm{He} + 50\%\mathrm{C_4H_{10}}$) :
  - Garfield++: ~250 s
  - NN: ~1 s
- This simulation algorithm is general and applicable for different particles. As for different particles, only the ionization part is different, the signal response simulation keeps the same
- By this way, signal simulation is not related to Geant4 and it is independent between each electron. To further speed up the signal response simulation, GPU or multithreading techniques can be easily used

Drift Chamber Ionization Simulation

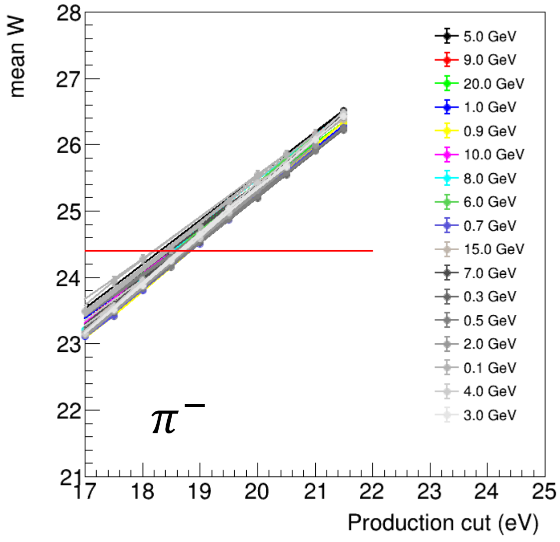
- As we know Geant4 can not simulate the ionization process properly (arXiv:2105.07064)
- Garfield++ is commonly used for precise ionization simulation for simple geometry
- In order to do a detailed drift chamber simulation, including particle interaction with detector materials, ionization in gas, drift and avalanche processes in drift chamber cell, combining Geant4 and Garfield++ is needed
- This paper "Interfacing Geant4, Garfield++ and Degrad for the Simulation of Gaseous Detectors" studied how to combine Geant4 and Garfield++ to get correct energy deposition or total number of ionized electrons (adopted by COMET experiment)

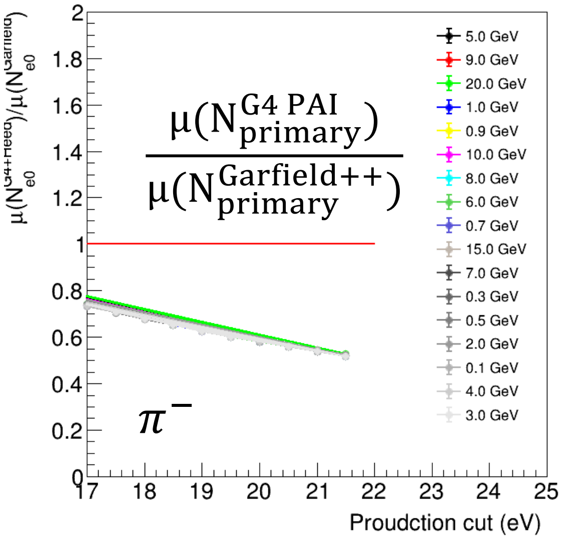
Drift Chamber Ionization Simulation by G4 PAI

- Method:
  - Geant4 PAI(Photo Absorption Ionization) model to simulate primary or secondary ionization
  - TrackHeed (from Garfield++) to simulate ionization from residual delta electron



- It was found that the primary ionization produced by this method is much less than the one from Garfield++ standalone simulation
- Confirmed with authors



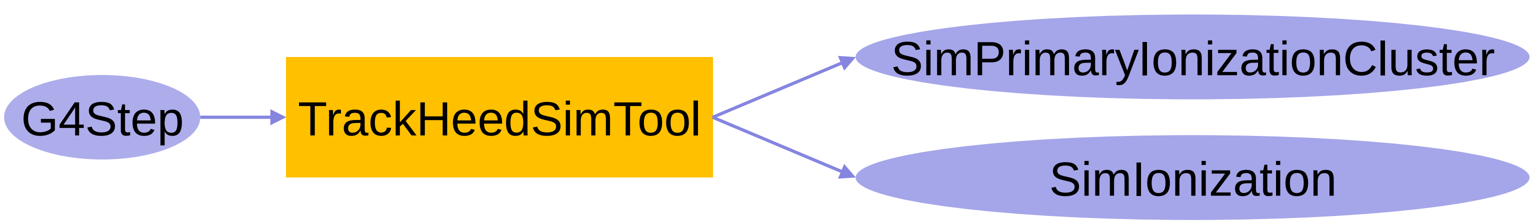


Drift Chamber Simulation

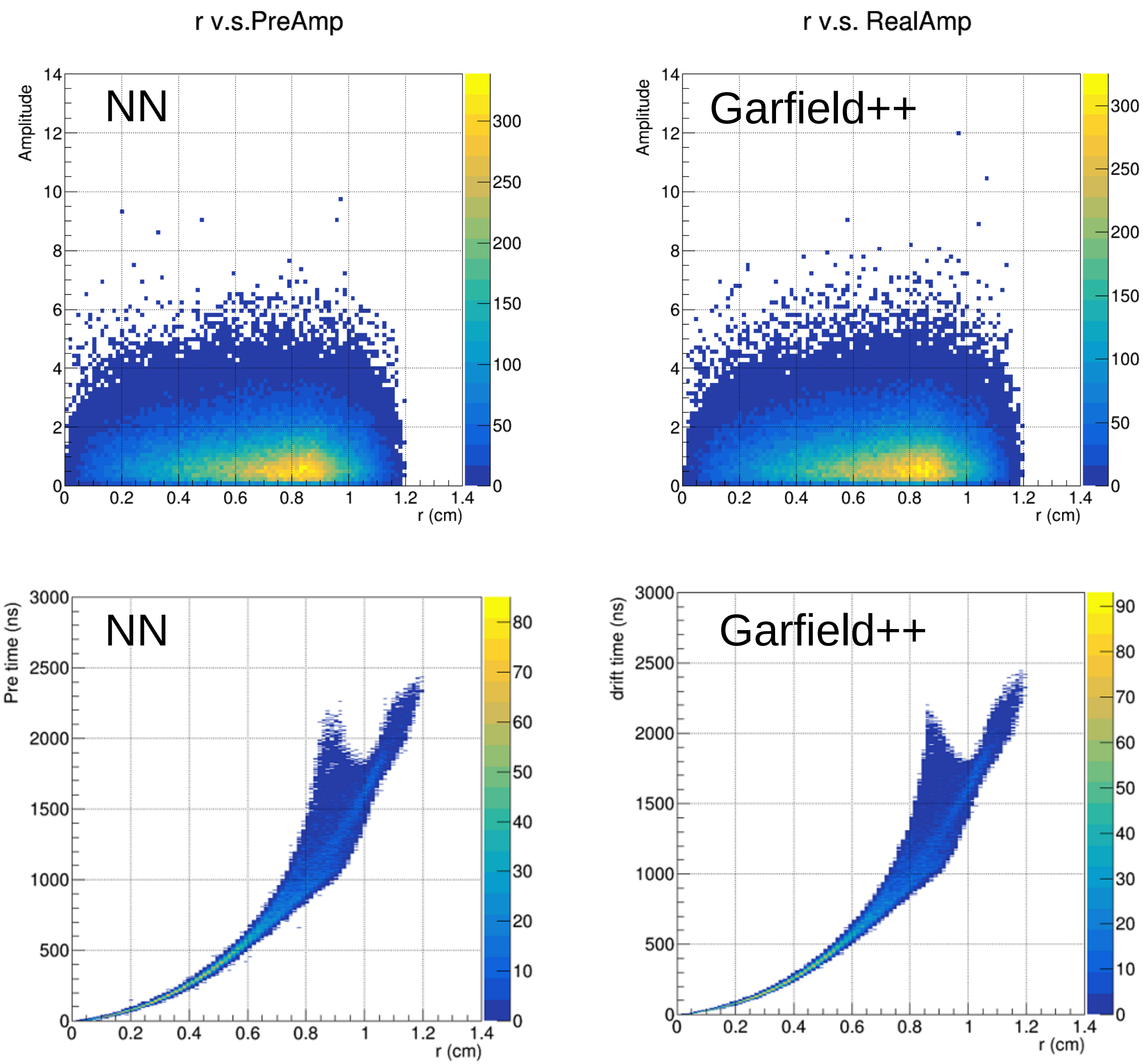
- The precise DC simulation:
  - Garfield++: ionization process, drift time, waveform
  - Geant4: particle transportation, interaction with detector material, decay, …
- ~~As we know Geant4 can not properly simulate the ionization process of charged particles within thin gas as well as the cell response (arXiv:2105.07064)~~
- ~~Garfield++ is commonly used for precise simulation for simple geometry~~
- ~~In order to do a detailed drift chamber simulation, including particle interaction with detector materials, ionization in gas, drift and avalanche processes in drift chamber cell, combining Geant4 and Garfield++ is needed~~

Proposed Drift Chamber Ionization Simulation

- Combining Geant4 and Garfield++ at G4Step level
- TrackHeedSimTool (Gaudi tool) is created for this task
  - Input: G4Step information (particle type, initial position and momenta, ionization path length)
  - Use TrackHeed (used by Garfield++) to simulate one step length (or multi-step length for speed up) ionization (new API contributed to Garfield++ PR)
  - Output: primary and total ionization information (contains position, time, cell id), saved in EDM
  - The kinetic energy of G4Track will be updated according to the energy loss in the ionization
  - Non-uniform magnetic field can be handled easily



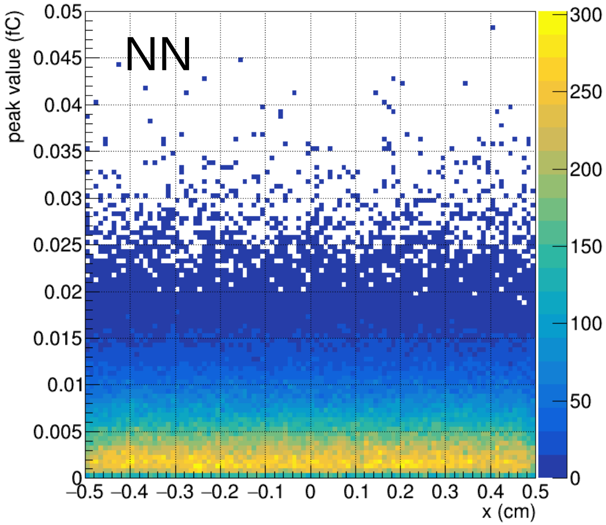
Simulation performance

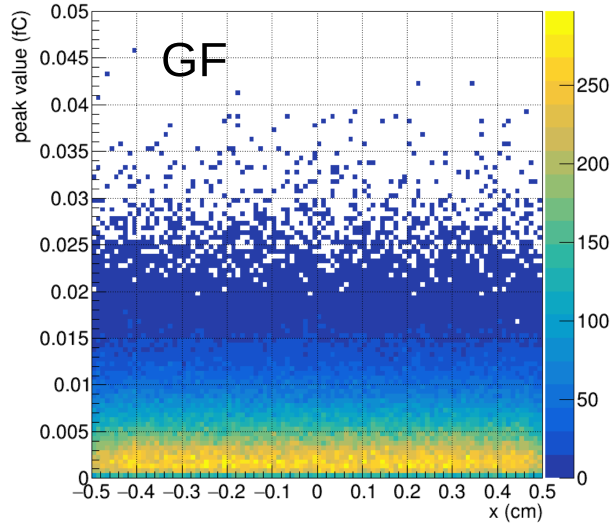


Amplitude value vs r

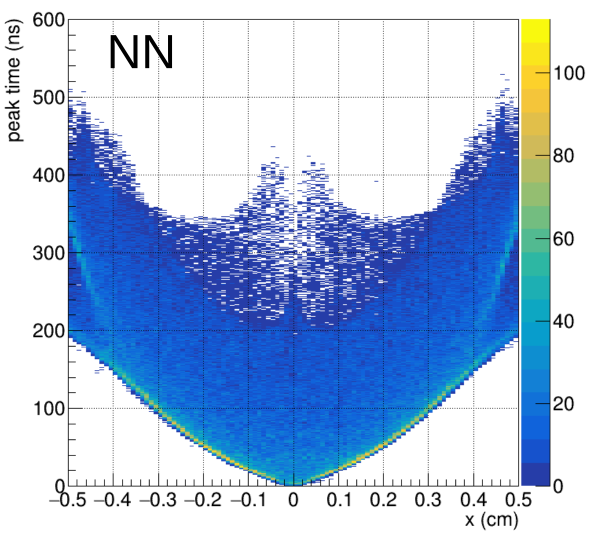
Drift time vs r

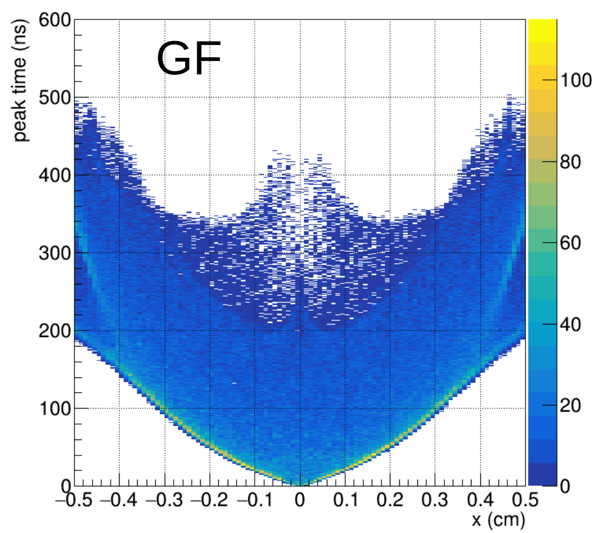
Simulation performance





Peak value vs x





Peak time vs x

Validation using the BESIII data

- The BESIII is a experiment study particle physics in tau-charm region
- The drift chamber is used for …
- Proof the NN model can learn drift time distribution from real data
- Dataset: radiative Bhabha event

Prepare experimental data for training

↓

Learn drift time distribution: $D_{drift\ time} = f(doca, eangle, layer)$

↓

Check $D^{data}_{drift\ time}$, $D^{NN\ sim.}_{drift\ time}$

↓

check the X-T cure: C^{data}_{X-T}, $C^{NN\ sim.}_{X-T}$

↓

check spatial resolution: using C^{data}_{X-T}, $C^{NN\ sim.}_{X-T}$