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

dN/dx simulation software in CEPCSW

Wenxing Fang

On behalf of drift chamber working group

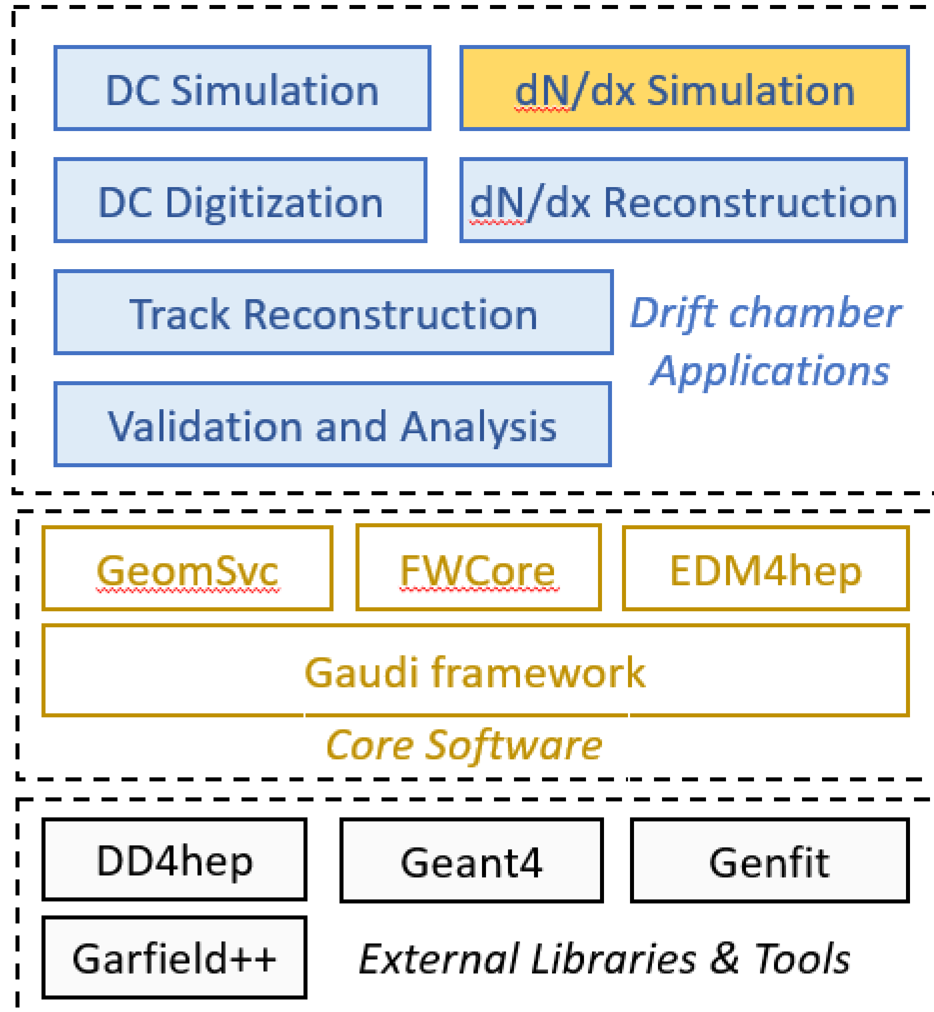
CEPC Physics and Detector Plenary Meeting
2021.08.17

Motivation

- CEPC experiment aims to measure the properties of Higgs, W and Z bosons precisely, which requires very good particle identification ability
- The cluster counting (dN/dx) technique shows its potential on having very good PID performance (<3% resolution)
- The dN/dx technique will be widely explored in CEPC drift chamber detector
- As software part, we aims to support the dN/dx study in detail
- Previous talk:
  - dN/dx simulation software in CEPCSW
- This talk:
  - Induced current simulation
  - Waveform reconstruction

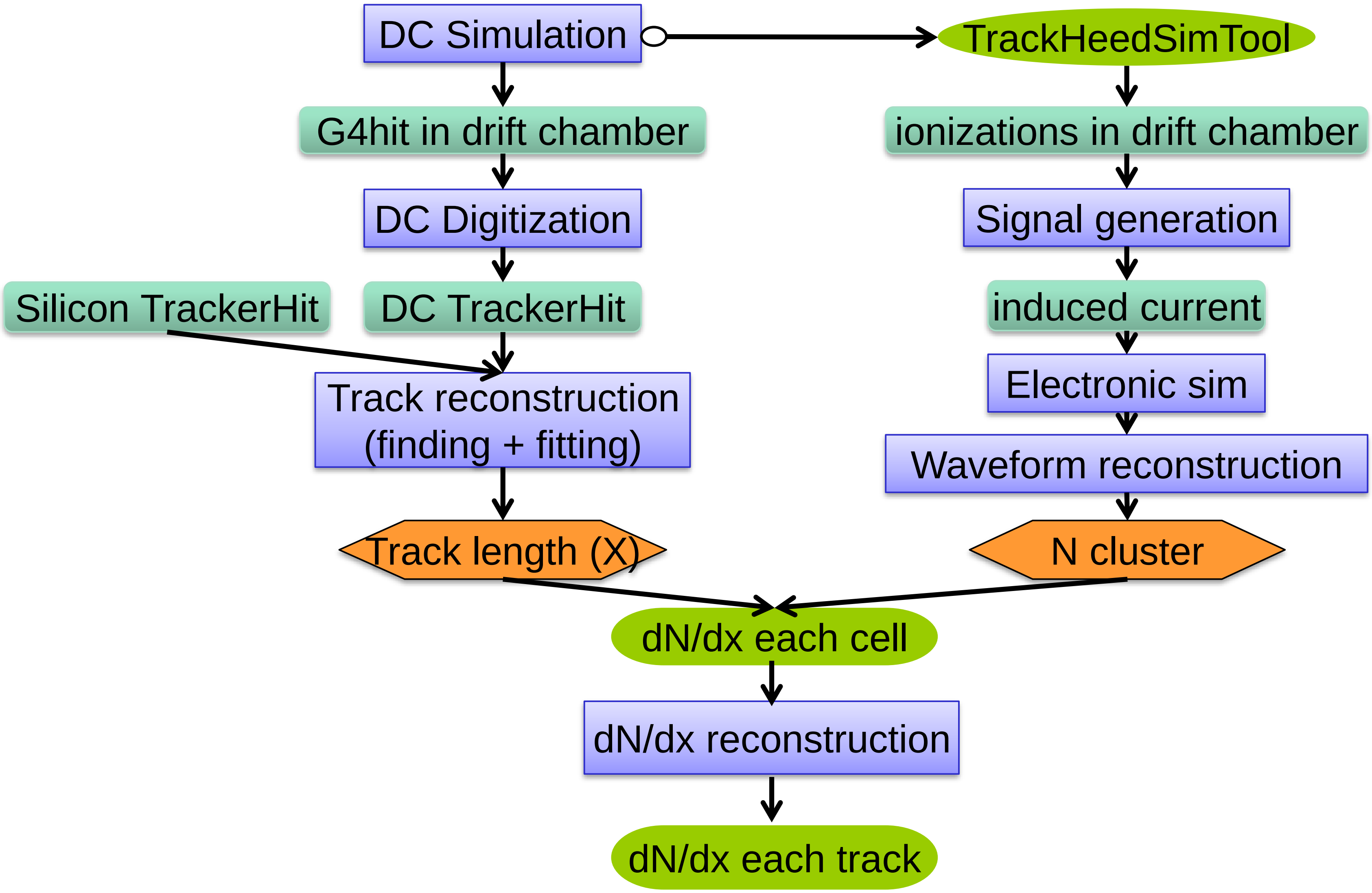
CEPCSW for drift chamber study

- Framework:
  - Gaudi
- EDM:
  - EDM4hep
  - FWCore
- Detector geometry and B field:
  - DD4hep
  - GeomSvc
- Drift chamber:
  - DC simulation (Geant4)
  - DC digitization
  - Track reconstruction (Genfit)
  - dN/dx simulation (Garfield++)
  - dN/dx reconstruction



https://github.com/cepc/CEPCSW

Schema of dN/dx study in CEPCSW



Ionization simulation

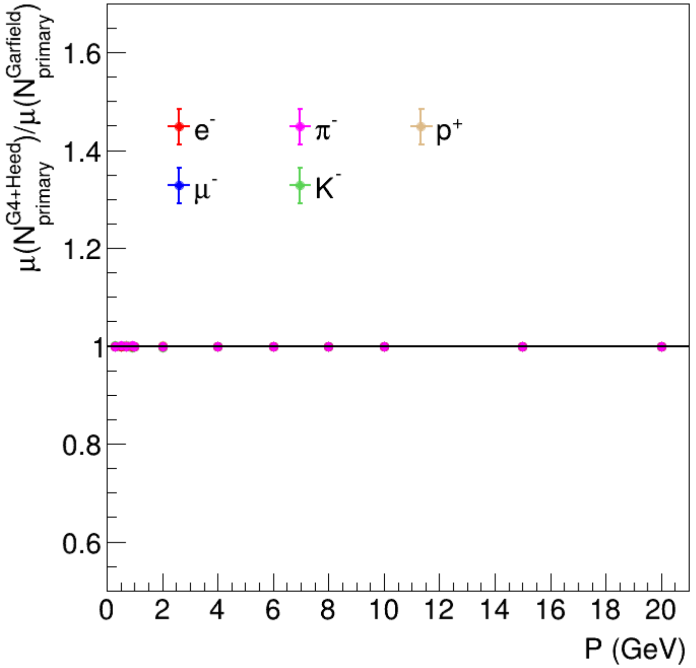
- From previous studies, we know Geant4 can not simulate the ionization process properly (arXiv:2105.07064)
- Garfield++ is commonly used for precisely ionization simulation for simple geometry
- In order to do a detailed drift chamber simulation, including particle interaction will detector materials, 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" introduces some ways to combine Geant4 and Garfield++ to get correct energy deposition or total number of ionized electrons (adopted by COMET experiment)
- However, it can not give both correct number of primary ionization and total number of ionized electrons (see backup)

Ionization simulation in CEPCSW

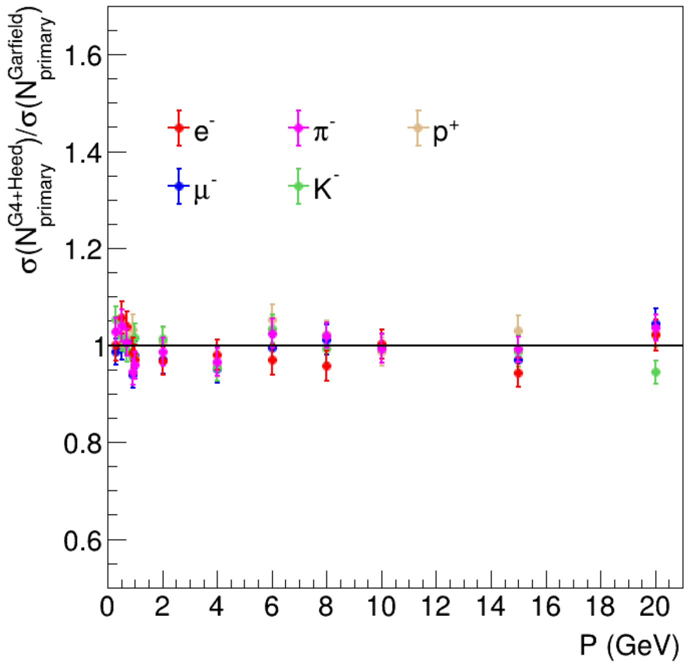
- Combining Geant4 and Garfield++ at G4Step level
- TrackHeedSimTool is created for this task
  - Input: G4Step information (particle type, initial position and momenta, ionization path length)
  - Use TrackHeed (used by Garfield++) to simulate one step length (or multi-step length for speed up) ionization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

Ionization simulation performance

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



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

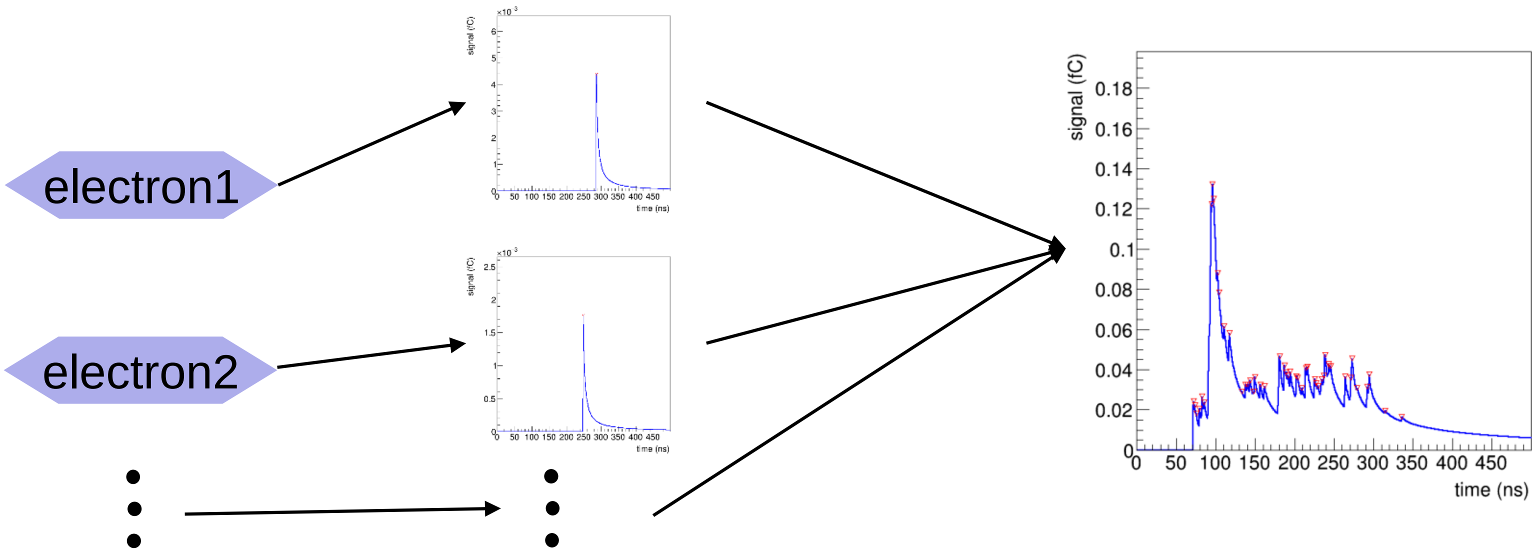


$$\frac{\sigma(N_{\text{primary}}^{\text{G4+TrackHeed}})}{\sigma(N_{\text{primary}}^{\text{Garfield++}})}$$

Consistent with Garfield++ standalone simulation results, see backup for more details

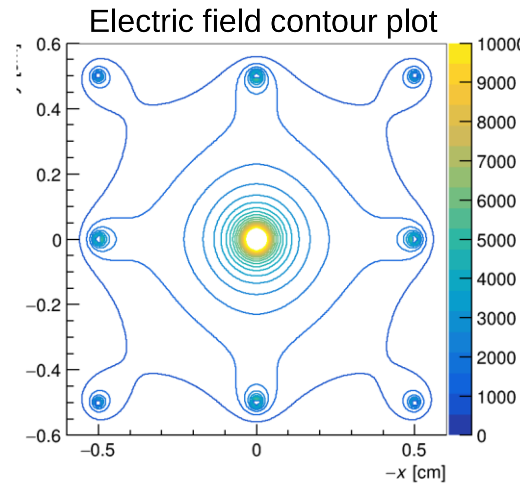
Induced current simulation in CEPCSW

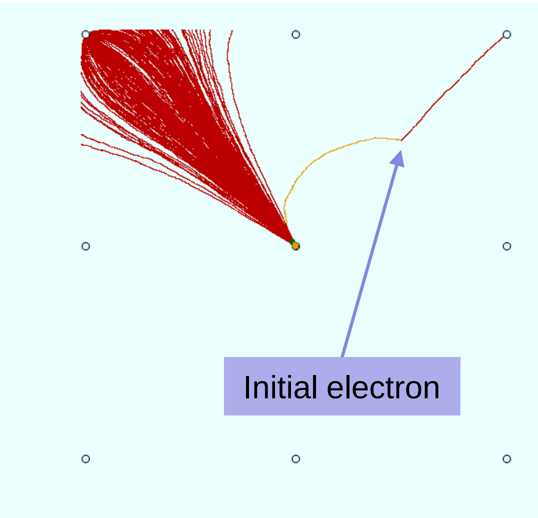
- From ionized electrons to induced currents
  - Using Garfield++, simulate the drift and avalanche of electron and drift of ions. It is slow, $\mathcal{O}(1)$ to $\mathcal{O}(10)$ seconds for different gas for one electron
  - Going to use parameterization method (parameters are based on Garfield++ simulation results), will be much faster
    - For each electron, give one induced current spectrum
    - As done by Garfield++, stacking all spectrums which come from same drift chamber cell gives final induced current for this cell



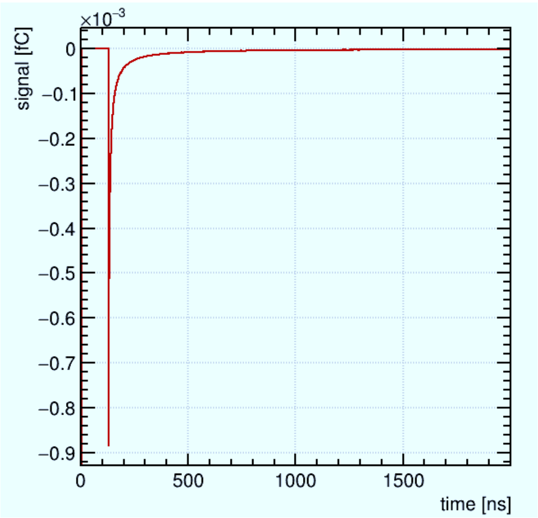
Getting parameters from Garfield++

- Garfield++ simulation:
  - 500k electrons uniformly distributed $1 \times 1\ \mathrm{cm}^2$ drift chamber cell
  - Gas: 50% He + 50 % C_4H_{10}
  - Center signal wire (2000 V), eight field wires (0 V)

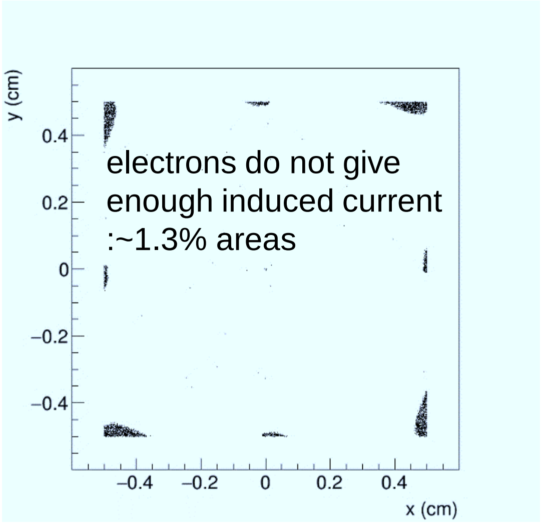




- One electron drift and avalanche
- Ions drift

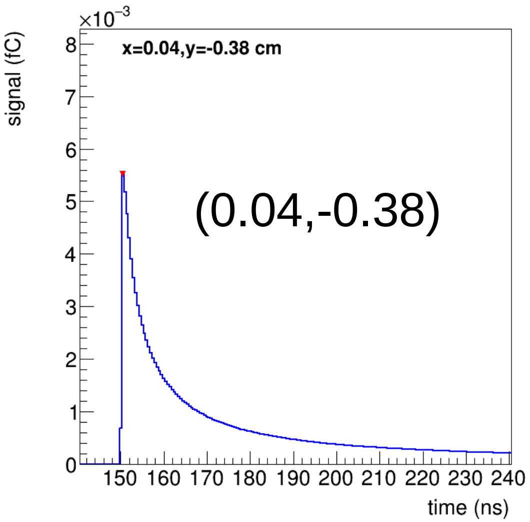


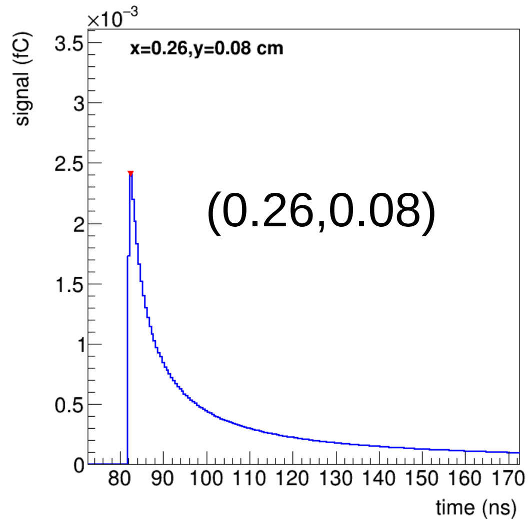
Induced current

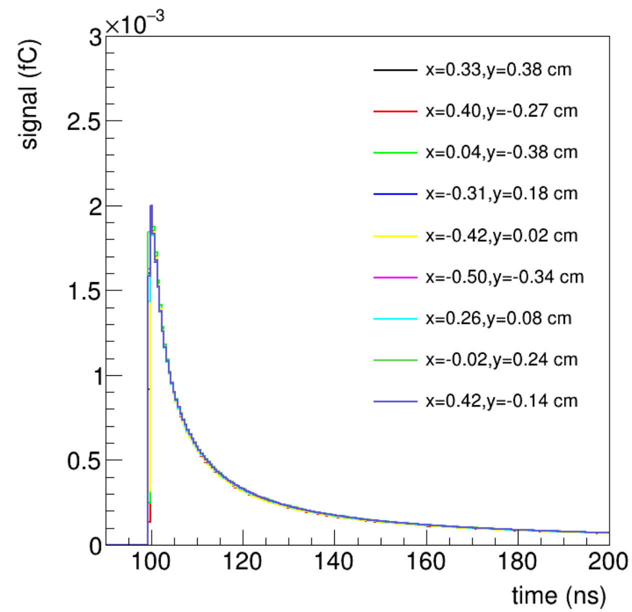


Getting parameters from Garfield++

- After some studies, we find one common feature for the induced currents, which is if we shift the peak position of induced current spectrums to same value (e.g. 100 ns) and scale the peak value to same value (e.g. 2×10^{-3}), all transformed spectrums are similar

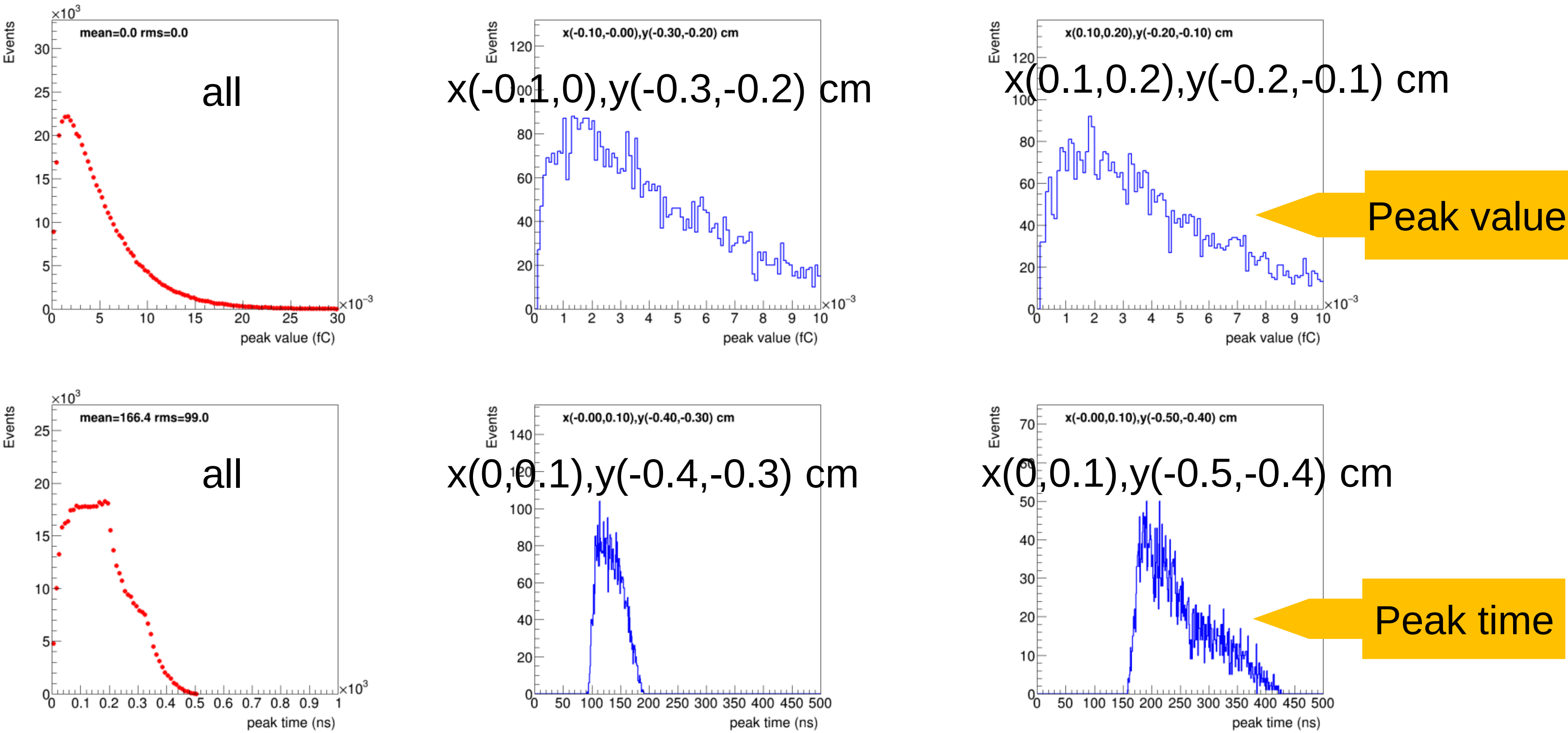






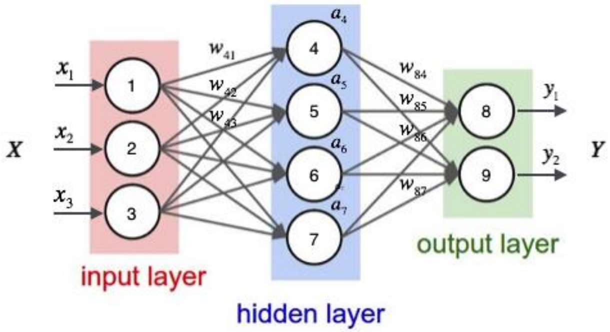
- Simulating induced current ≈ simulating (peak_time, peak_value) + using induced current template

Garfield++ simulation



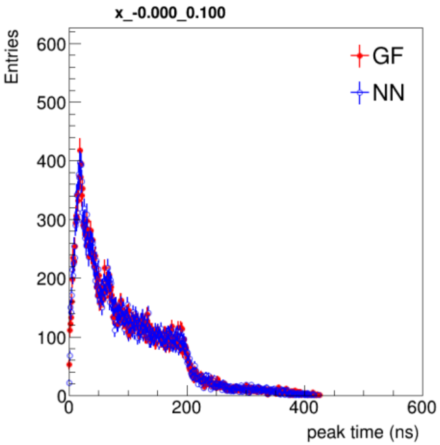
- Simulate (peak_time, peak_value):
  - Sampling method base on which bin the electron (x,y) is located
  - Machine learning method according electron (x, y) without binning

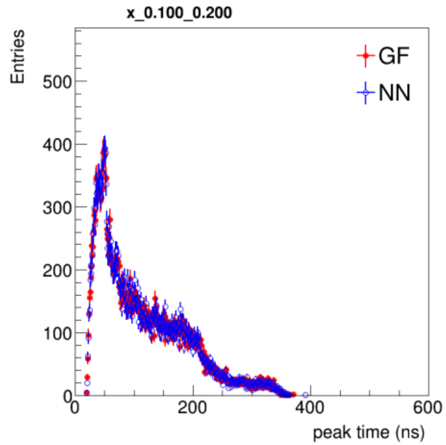
ML for peak time and value simulation

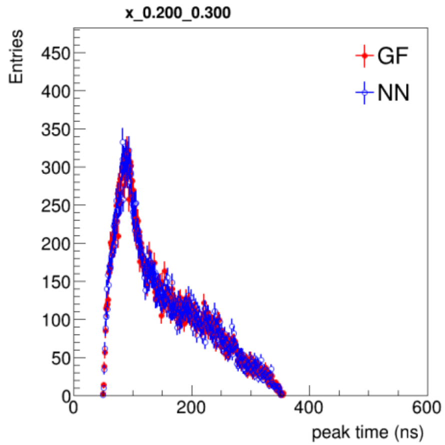


- Model: deep neutral network (DNN)
  - Consist of input, hidden, and output layers
- Input data:
  - Local x and y positions of ionized electrons
  - N(0,1) distribution noise
- Output: peak time and value of induced current
- Loss: two sample test statistics (SmoothKNN) between real data and generated data

Peak time simulation



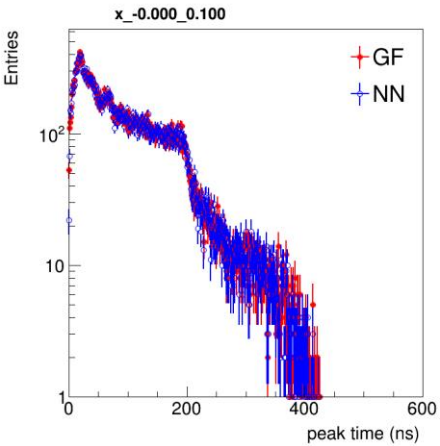


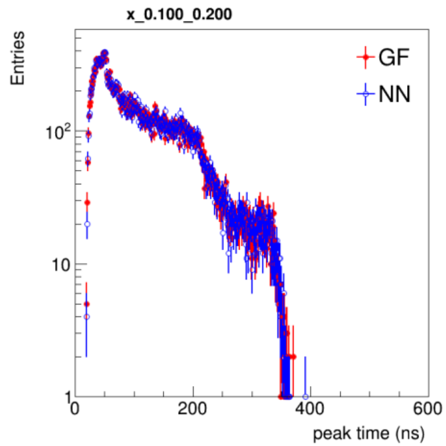


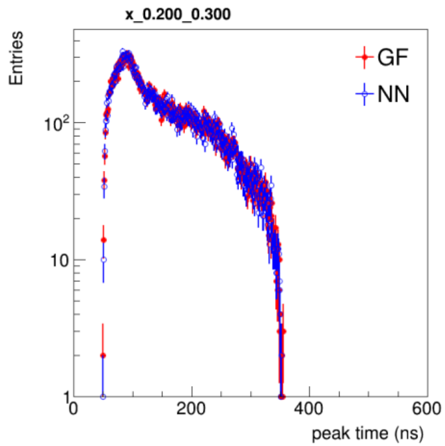
x(0,0.1) cm

x(0.1,0.2) cm

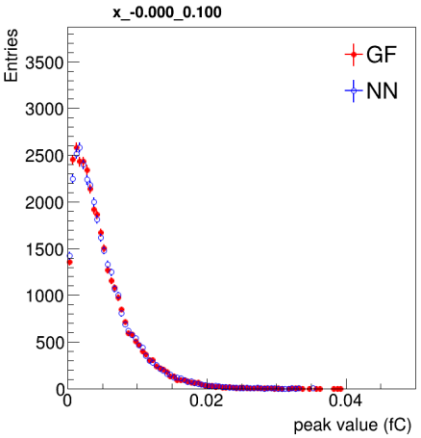
x(0.2,0.3) cm

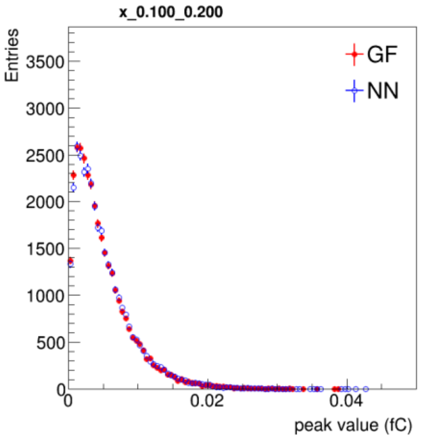


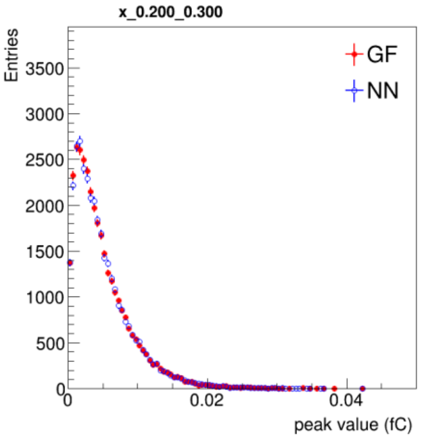




Peak value simulation



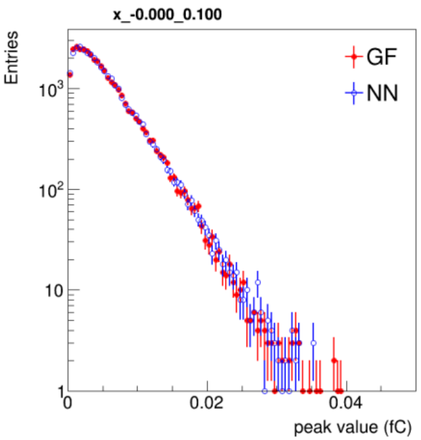


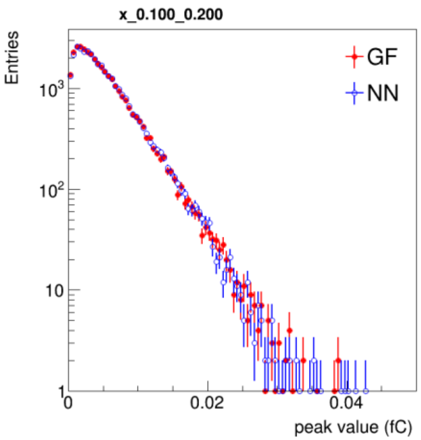


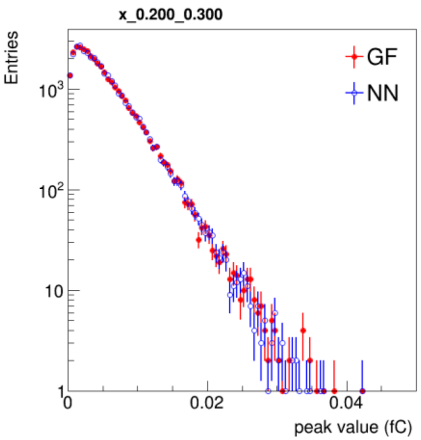
x(0,0.1) cm

x(0.1,0.2) cm

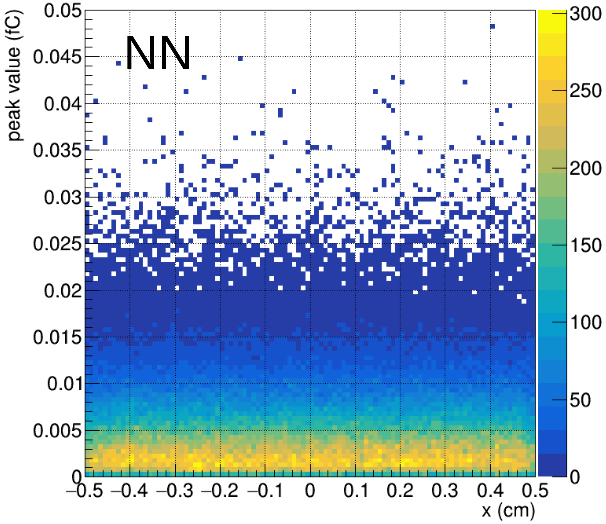
x(0.2,0.3) cm

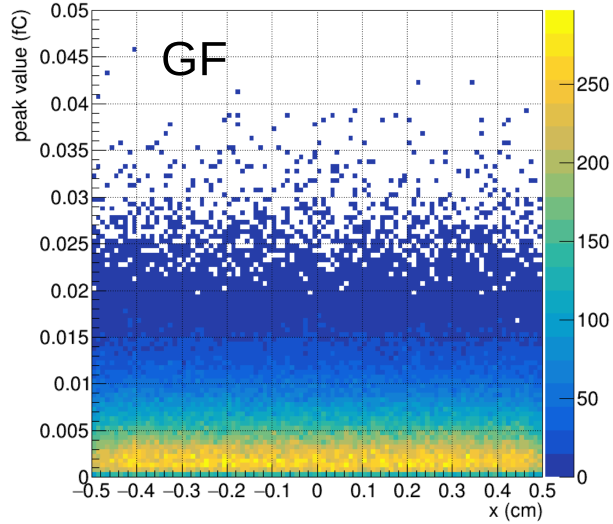




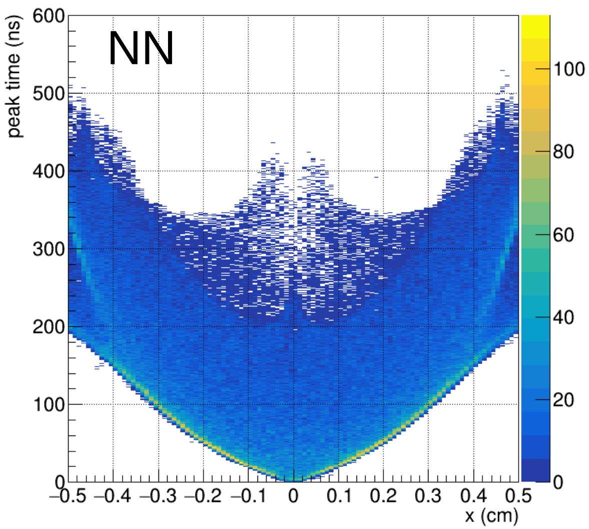


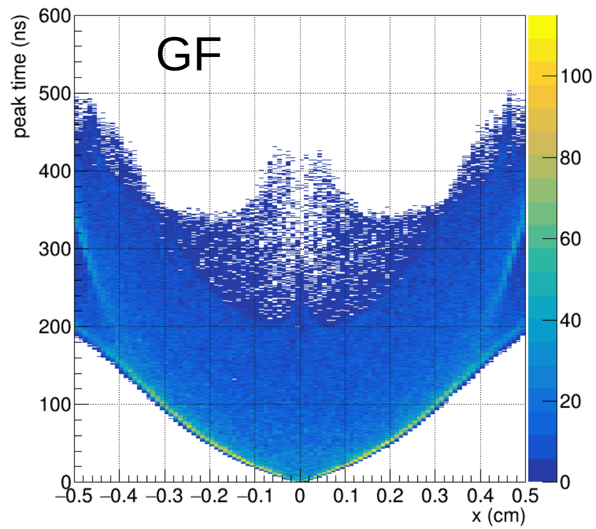
Garfield++ simulation





Peak value vs x

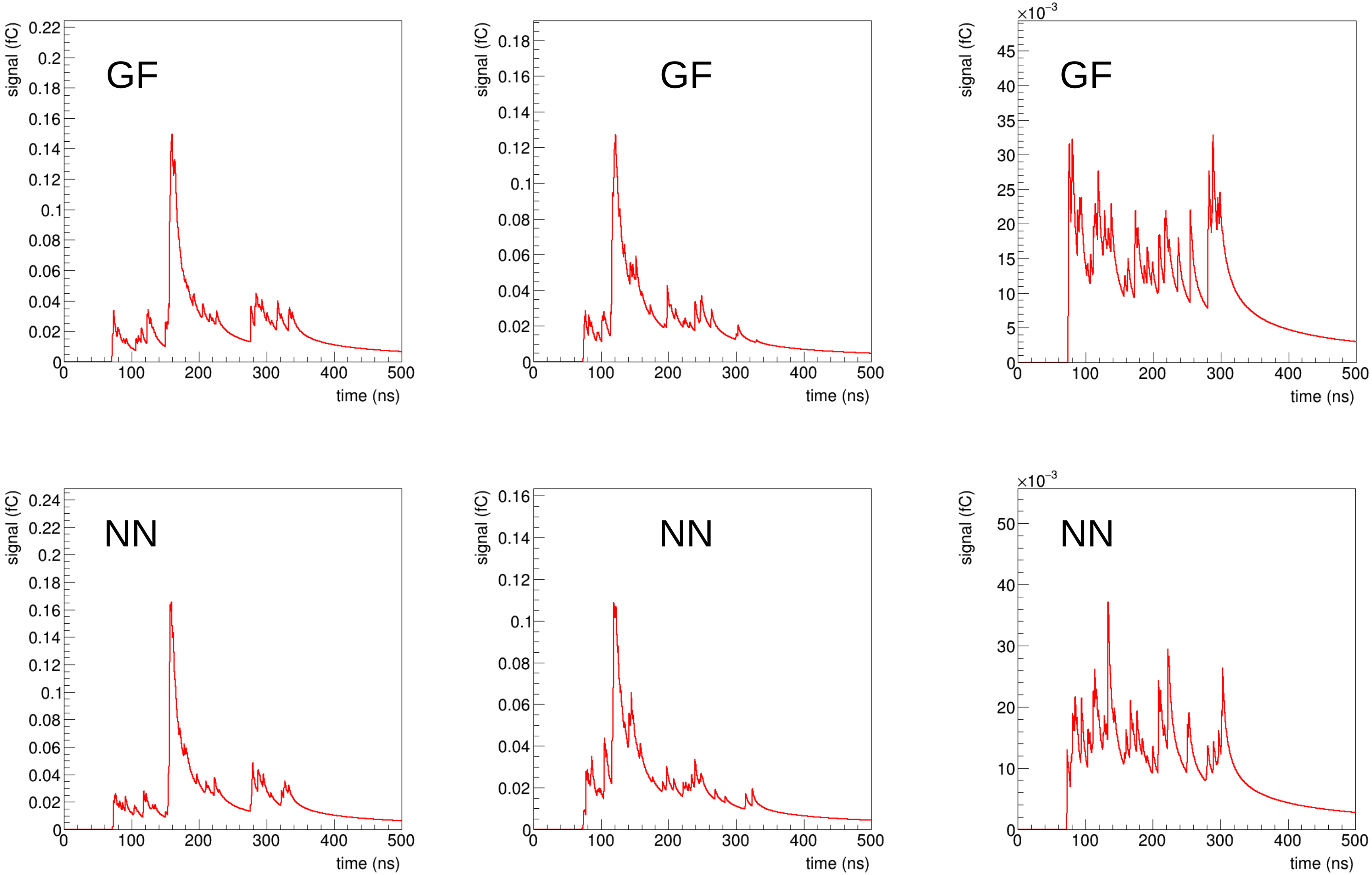




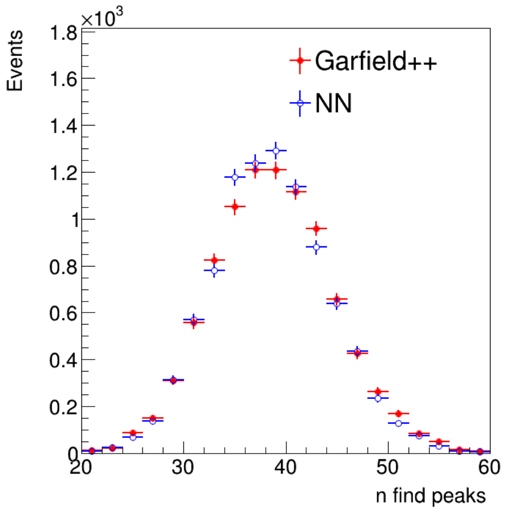
Peak time vs x

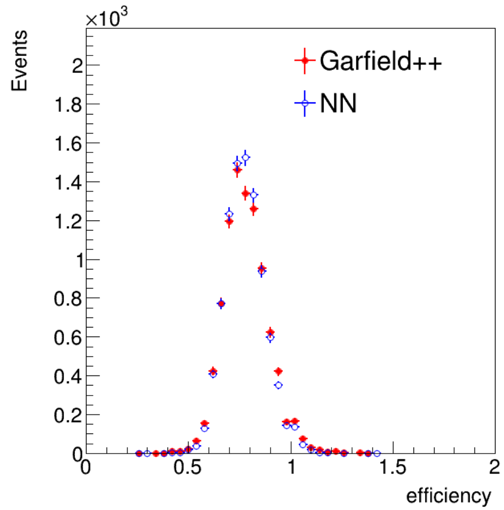
Signal wire induced current simulation

- ❑ 1GeV π^-

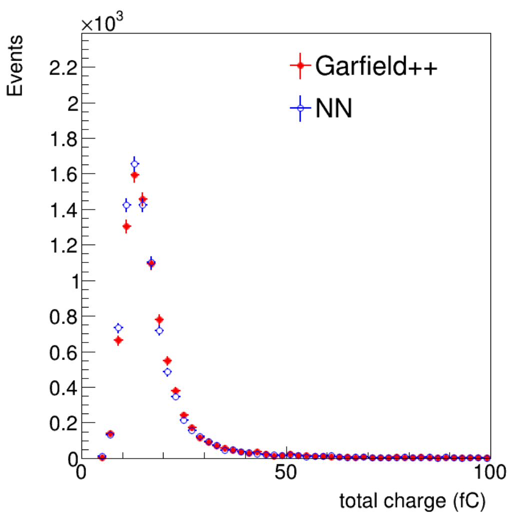


Signal wire induced current simulation





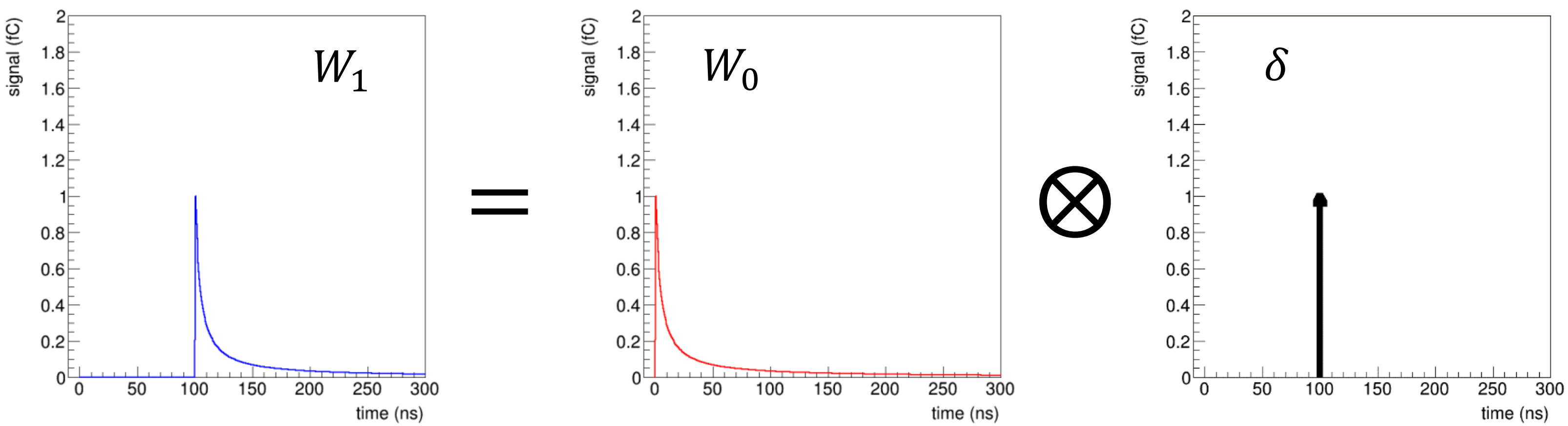
- By this way, the simulation of induced current is not related to Geant4 and is independent between each other. Therefore, it can be easily transported to GPU or using multithreading technique to speed up the simulation



- 1GeV π^-
- Good agreement between NN and Garfield++
- Average time for one cell simulation:
  - Garfield++: ~250 s
  - NN: ~1 s

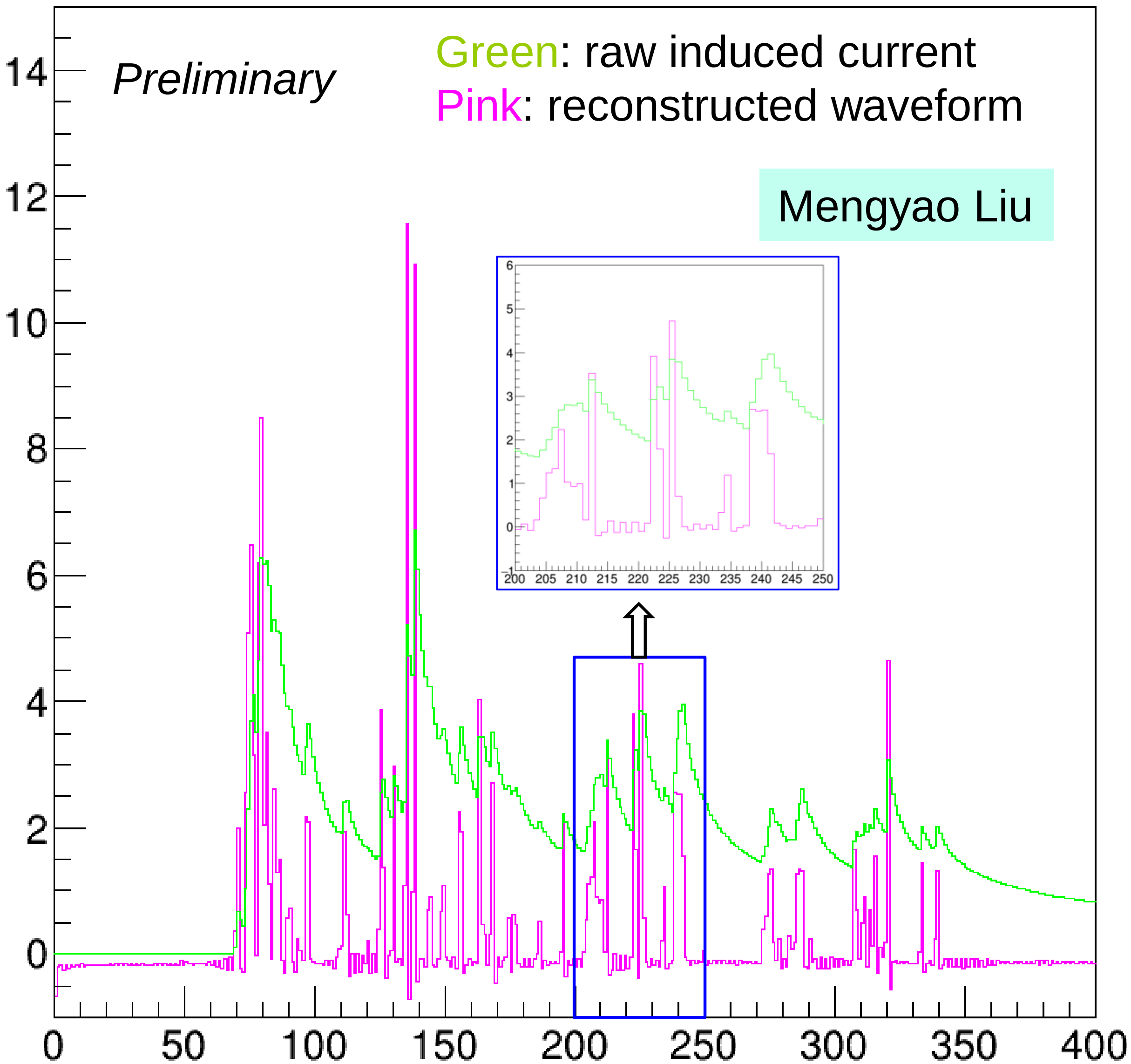
Waveform reconstruction

- Using the deconvolution method from JUNO experiment



- $F(W_1) = F(W_0 \otimes \delta) = F(W_0) * F(\delta)$
- $F(\delta) = F(W_1)/F(W_0)$
- $\delta = F^{-}(F(\delta))$, give information about peak time and value

Waveform reconstruction



- The preliminary results looks promising
- Need further checks, such as cluster find efficiency, the purity of found clusters
- Using this method, we can analyse the components of the waveform and filter electronic noise
- Welcome for other algorithms

User extension data in EDM4hep

- As there is no waveform data format in EDM4hep yet, user extension data is a way to add additional data.
  - WIP: https://github.com/key4hep/EDM4hep/pull/117

Tao Lin

The proposed underlying data structure:

```
edm4hep::UserExt:
    Description: "A simple struct with user defined int/float/double"
    Author : "Tao Lin"
    VectorMembers:
      - int valI // data int
      - float valF // data float
      - double valD // data double
```

The proposed user APIs:

```
ud xyzi;
xyzi.reg("x", 1, 0)
    .reg("y", 1, 1)
    .reg("z", 1, 2)
    .reg("t", 2, 0)
    .reg("i", 0, 0);
```

Runtime Type definition

```
xyzi.from(usrexts[i], 0)
    .get("x", x)
    .get("y", y)
    .get("z", z)
    .get("t", t)
    .get("i", iii);
```

Getters

```
xyzi.put("x", x)
    .put("y", y)
    .put("z", z)
    .put("t", t)
    .put("i", i);
```

```
auto udv = usrexts.create();

xyzi.to(udv);
```

Setters

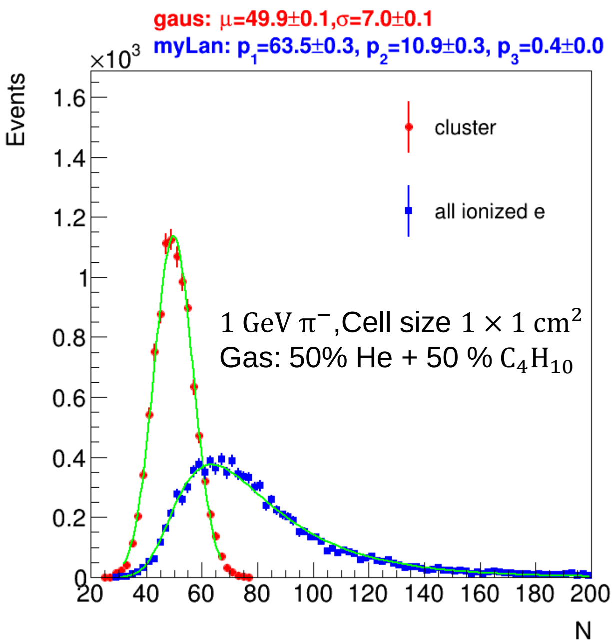
Summary

- A new simulation scheme for drift chamber has been presented
- The ionization simulation using Geant4 combined with TrackHeed have been implemented in CEPCSW. Results are consistent with Garfield++ simulation
- In order to speed up the simulation of induced current, a parameterized method using ML has been described, which gives consistent results with Garfield++ simulation
- The de-convolution with Fourier transforms was introduced to waveform reconstruction, preliminary results are promising and further work is still in progress
- Using user extension data in EDM4hep to store waveform data is proposed. Currently, this is under discussion with key4hep
- Future plan:
  - Studying waveform reconstruction by adding noise and electronics response
  - Development of the reconstruction algorithm for the dN/dx of a track

Back up

Motivation

- The particle identification is very important for CEPC flavor physics study. Good hadron separation up to 20 GeV is essential
- Traditionally: using dE/dx method
  - Due to the production of delta electron, the deposited energy follows Landau distribution
  - Resolution is ~6%
- New technique: using dN/dx (cluster counting) method
  - The number of primary ionization follows Poisson distribution
  - Resolution could reaches <3%
- The dN/dx technique will be widely explored in CEPC drift chamber detector

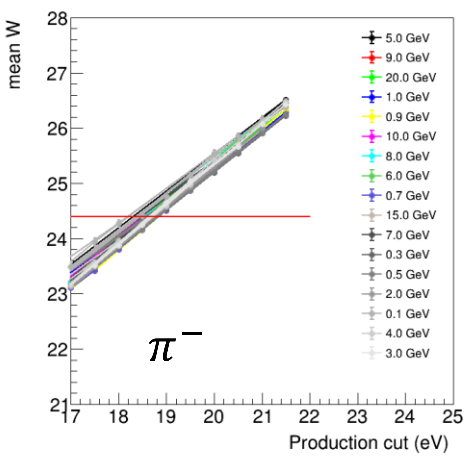


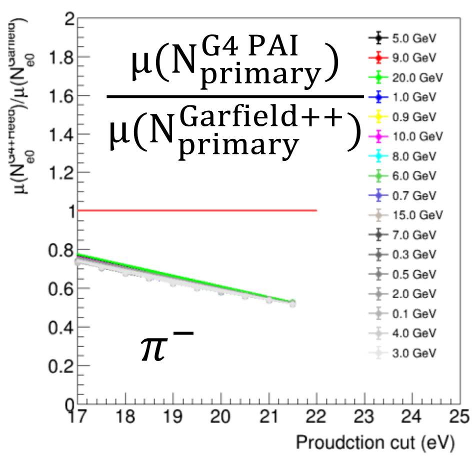
Ionization simulation in gas

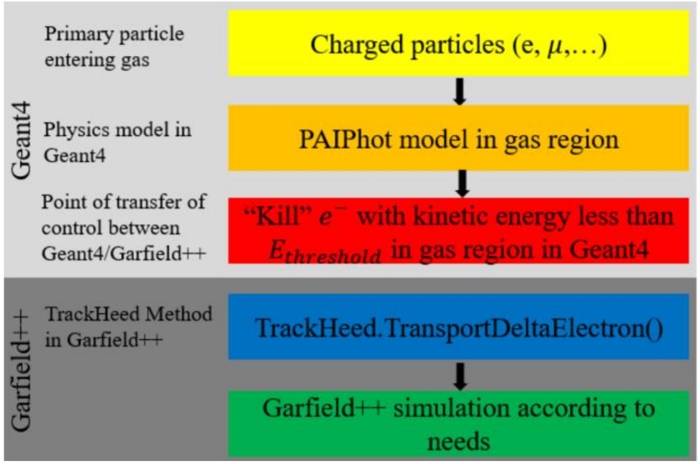
- Garfield++
  - Using Heed PAI model to simulate the ionization in gas precisely
  - Can simulate the drift and avalanche of electrons in gas
  - The drift of ions to cathode can be simulated
  - The induced current can be given
  - It is useful to study and characterize the properties of gas detector with simple geometry but not for full drift chamber detector
- Geant4
  - Can simulate collider events and the interaction between particles and materials in full detector
  - It does not simulate the ionization process properly, neither the drift and avalanche processes
- In order to simulate including particle interaction will detector materials, ionization in gas, drift and avalanche processes in full detector, we try to combined Geant4 and Garfield++ in CEPCSW

Ionization simulation in CEPCSW (G4 PAI)

- First try: according to paper "Interfacing Geant4, Garfield++ and Degrad for the Simulation of Gaseous Detectors" :
  - Geant4 PAI model to simulate primary or secondary ionization
  - TrackHeed to simulate ionization from residual delta electron
- However, it was found that the primary ionization produced by this method is much less than Garfield++.



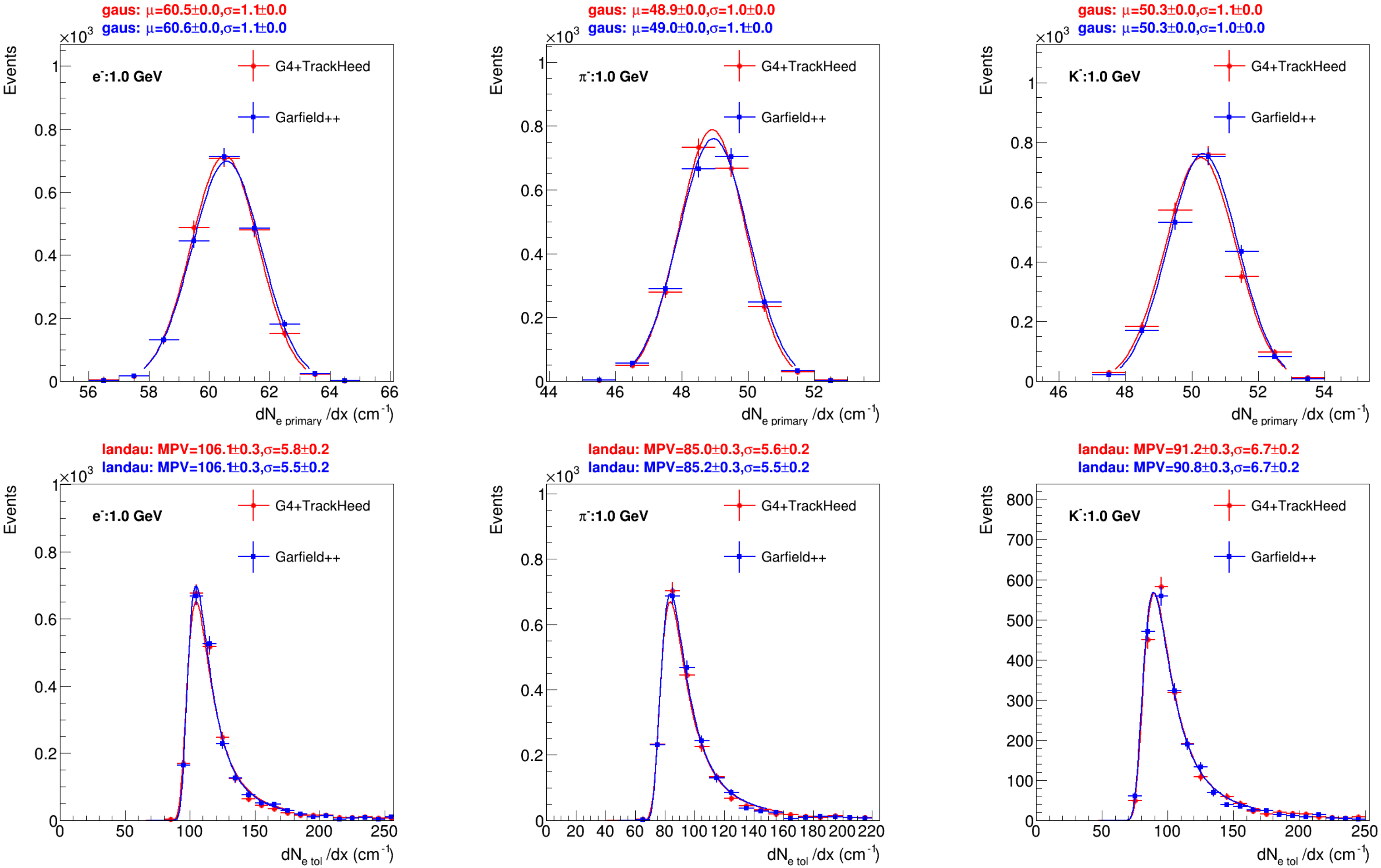




- Checking with authors:
  - This method designed to obtain correct energy deposition (or total ionizations)
  - It is true that this method will give less primary ionizations, so this method is obsoleted

Ionization simulation performance

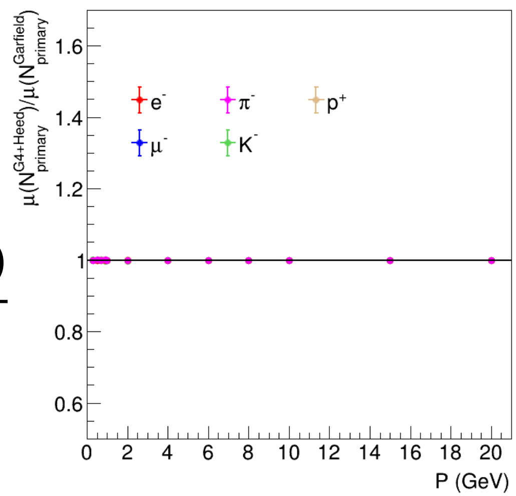
- Gas: 50% He + 50 % C_4H_{10}

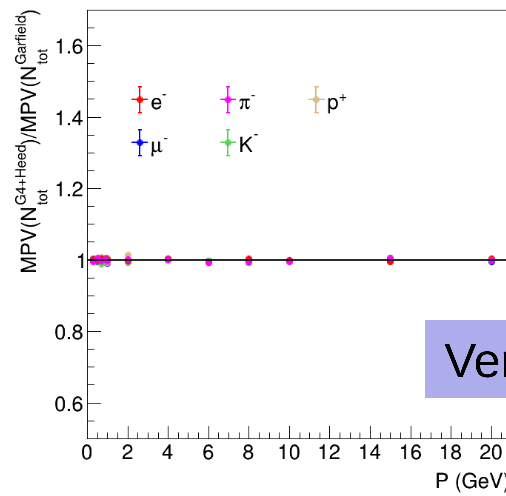


Ionization simulation performance

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

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

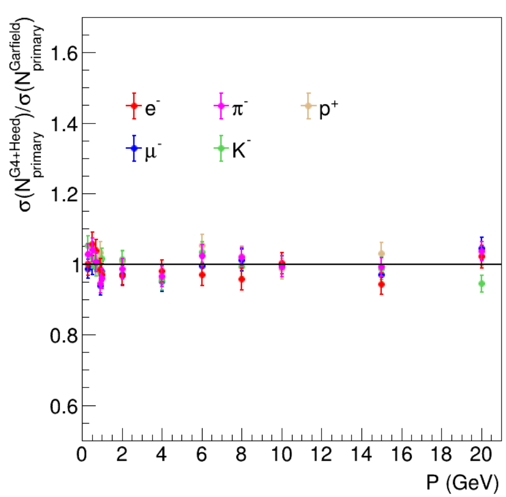


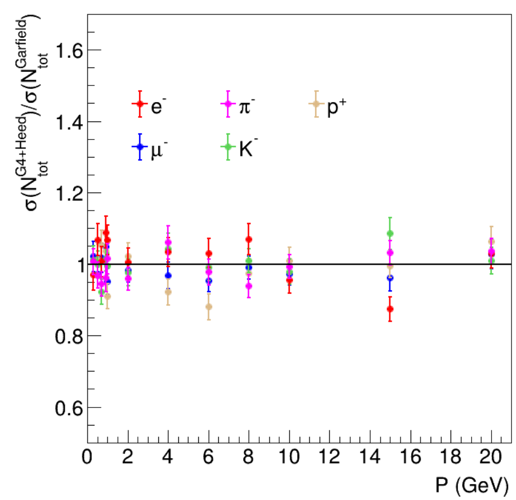


$$\frac{MPV(N^{G4+TrackHeed}_{tot})}{MPV(N^{Garfield++}_{tot})}$$

Very good agreement

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





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