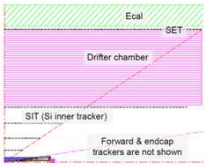
Reconstruction Algorithm of Cluster Counting in Drift Chambers

Shuaiyi Liu[1], Guang Zhao[1], Zhefei Tian[2], Brunella D'Anzi[4], Mingyi Dong[1], Nicola De Filippis[4], Xu Gao[3], Francesco Grancagnolo[4], Fei Huang[1], Shengjie Jin[1], Gang Li[1], Linghui Wu[1], Zhenyu Zhang[2]
[1]Institute of High Energy Physics, Chinese Academy of Science, China
[2]Wuhan University, China
[3]Jilin University, China
[4]Instituto Nazionale di Fisica Nucleare, Italy

Introduction

Drift Chamber in CEPC 4th conceptual detector



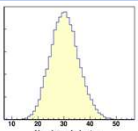
- Tracker with silicon tracker and a drift chamber
- The chamber optimized for PID with cluster counting technique
- Up to 20GeV/c K/π separation power better than 2σ required
- Benefited momentum resolution at low momentum

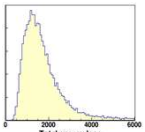
Preliminary DC parameters	
Inner radius	800mm
Outer radius	1800mm
Cell size	18 mm×18 mm
Gas mixture	He/iC_4H_{10}=90:10
Length of outermost wires (cosθ=0.82)	5143mm

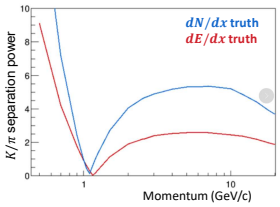
dE/dx vs dN/dx measurement

dN/dx
- Number of primary ionization clusters per unit length
- Poisson distribution
- Small fluctuation

dE/dx
- Energy loss per unit length
- Landau distribution
- Large fluctuation







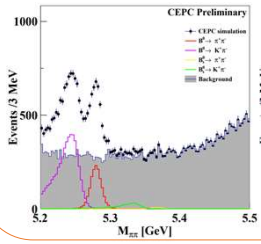
dN/dx has a much better K/π separation power up to 20 GeV/c compared to dE/dx (MC truth)

Advantages of Particle identification

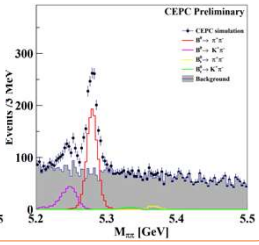
- Reduce combination background
- Improve mass resolution
- Improve jet energy resolution
- Benefit flavor tagging

Simulation of $B^0/B_s^0 \to \pi^+\pi^-$ with Delphes

Without PID

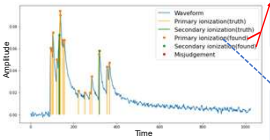


With PID



Algorithm for Simulated Samples

Reconstruction of cluster counting



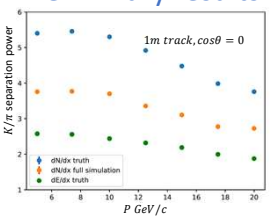
Step1. Peak Finding
Discriminate peaks (both primary and secondary) except from the noises (classification problem)

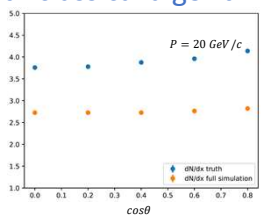
Step2. Clusterization:
Determine the number of clusters (N_{cls}) from the detected peaks (regression problem)

Reconstruction algorithms

- **Classical method**(developed)
  - Derivative-based peak finding + clusterization with peak merge
- **Deep learning based algorithm** (ongoing)
  - Peak finding with LSTM + clusterization with DGCNN

Preliminary results with classical algorithm

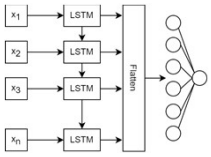




K/π separation power is better than 2 sigma at 20GeV/c in dN/dx full simulation

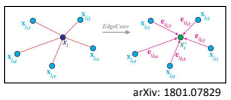
Deep learning based algorithm

Peak finding with LSTM (Long Short Term Memory)



- RNN-based architecture
- Binary classification of signals and noises on slide windows of peak candidates

Clusterization with DGCNN (Dynamic Graphic Convolutional Neural Network)

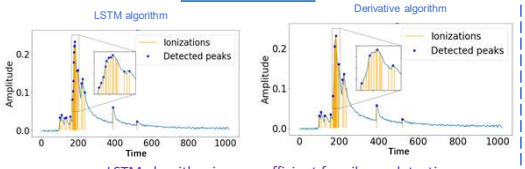


arXiv: 1801.07829

- GNN-based architecture
- Massage passing through neighbor nodes ⇔ Clusterization of electron timings from the same primary cluster
- Binary classification of primary and secondary electrons

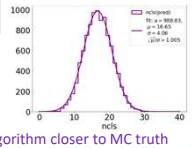
Comparison between Classical & Deep learning algorithms

Peak finding



LSTM algorithm is more efficient for pile-up detection

Clusterization



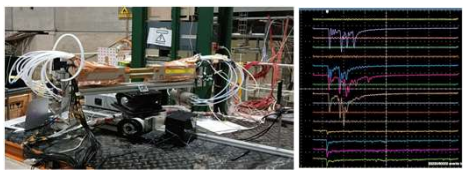
Deep learning algorithm closer to MC truth

Clusterization Method	μ	σ	σ/μ
MC truth	16.53	3.93	23.8%
Classical algorithm	18.67	4.60	24.6%
Deep learning	16.65	4.06	24.4%

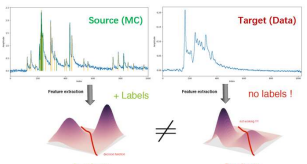
Algorithm of Test Beam Samples

Deep learning algorithm in beam test data analysis

Beam Test for DC prototype



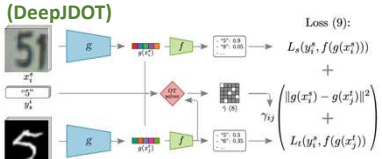
Difficulties for ML in data samples



- Domain shift between MC/data
- Lack of labels in data

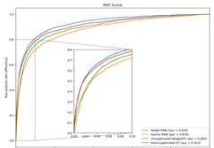
➡ **Semi-supervised DeepJDOT**

Deep Joint Distribution Optimal Transport (DeepJDOT)



- Original DeepJDOT (arxiv:1803.10081)
- We develop the semi-supervised version of DeepJDOT

Validation on MC experiment



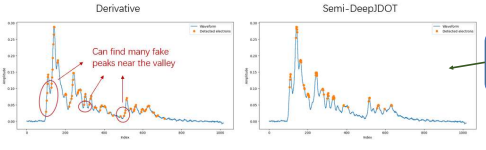
Model	AUC	pAUC (FPR<0.1)
Ideal model *	0.926	0.812
Source RNN	0.878	0.749
Unsupervised DeepJDOT	0.895	0.769
Semi-supervised DeepJDOT	0.912	0.793

* Supervised model assuming there are labels in the training sample

Details of beam test in Guang's talk "**Drift chamber with cluster counting techniques for CEPC**" at "**PID and other detectors:3** " session.

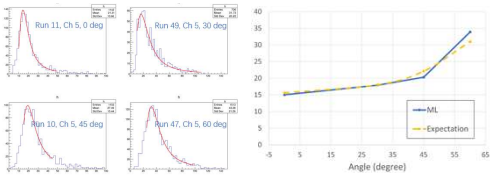
Preliminary results of peak finding

Comparison between classical and ML peak finding algorithms



Deep learning has better performance on finding peaks reasonably on data.

Number of Peaks distributions of data with different angles



The result of deep learning algorithm has consistent tendency with theoretical expectation.

Contact: Shuaiyi Liu e-mail: liusy@ihep.ac.cn
The 2023 International Workshop on the High Energy Circular Electron Positron Collider, 2023 10.23-27, Nanjing