



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

# 2023-2024 年度考核报告

方文兴 (软件组)

实验物理中心

2024年11月22日

# 岗位工作内容

目前聘用的职务	助理研究员	聘用时间	2021.09
研究方向	粒子物理实验软件研究和机器学习		

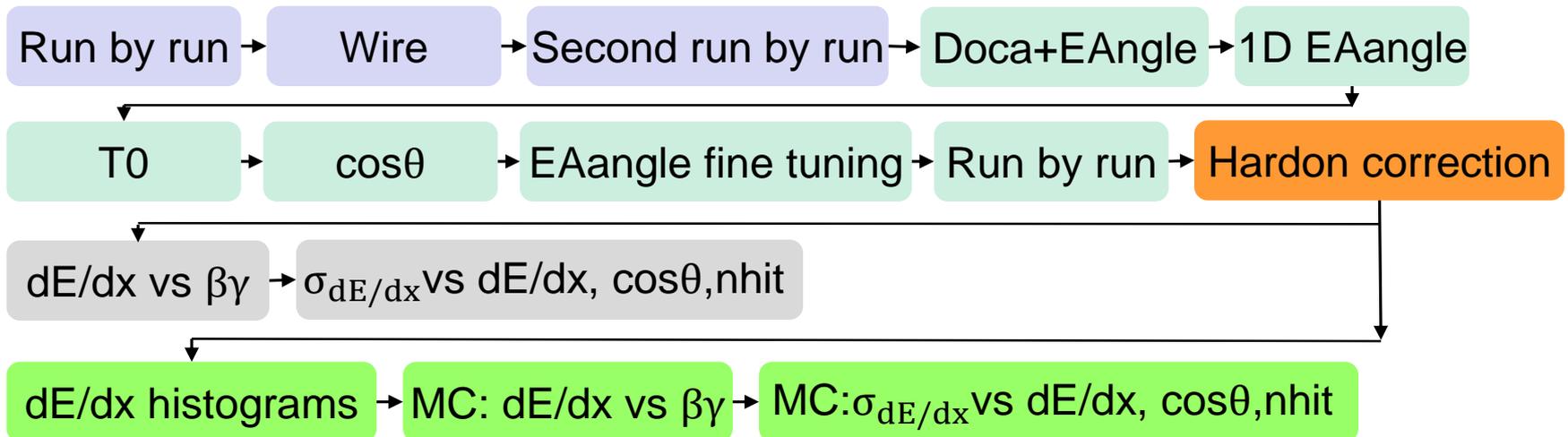
## 主要工作内容：

1. BESIII 实验：dE/dx 软件；机器学习
2. JUNO 实验：OEC 软件；重建算法；机器学习
3. CEPC 实验：漂移室软件；机器学习



# BESIII dE/dx 软件 (1)

- ❖ 背景：粒子鉴别 (PID) 是衡量 BESIII 探测器性能的重要指标之一。其主要由 dE/dx 和 TOF 组成
- ❖ 本人研究工作：
  - 实验数据：基于电子样本和强子样本的 dE/dx 修正，以及刻度
  - 模拟数据：dE/dx 模拟和刻度

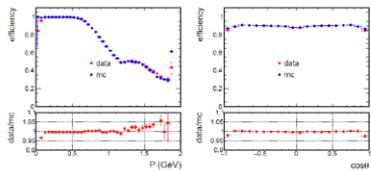


# BESIII dE/dx软件 (2)

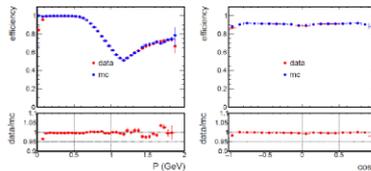
- ❖ 按计划完成 round17 (psi3770, Charmonium, R scan) dE/dx 精细刻度以及模拟的相关工作

## dE/dx Hadron Calibration of Round17 $\psi(3770)$ data

Preliminary results for dE/dx PID efficiency

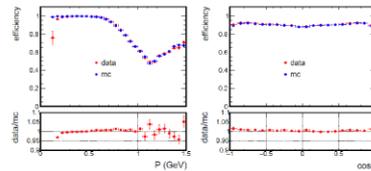


$\pi$  dE/dx PID efficiency

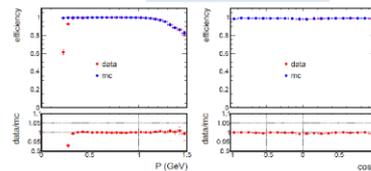


$\pi$  dE/dx PID efficiency, only  $K, \pi$  separation

FANG Wenxing  
QIU Jinfa



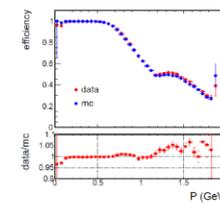
$K$  dE/dx PID efficiency



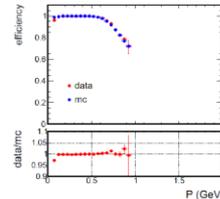
$p$  dE/dx PID efficiency

11

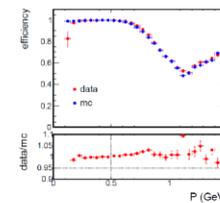
## dE/dx Hadron Calibration of 2<sup>nd</sup> half data of Round17



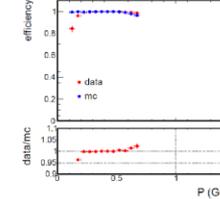
$\pi$  dE/dx PID efficiency



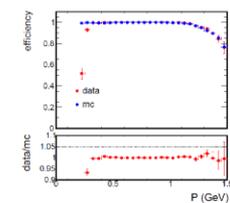
$\pi$  dE/dx PID efficiency



$K$  dE/dx PID efficiency



$K$  dE/dx PID efficiency



$p$  dE/dx PID efficiency

Preliminary results of dE/dx  
PID efficiency of data and MC

12

# BESIII dE/dx软件 (3)

- ❖ BESIII 探测器升级：内室替换为CGEM探测器
- ❖ MDC击中数减少，使得低动量（大角度）径迹的 dE/dx PID 性能下降
- ❖ 研究利用 CGEM 的 dE/dx 信息改善 dE/dx PID 性能

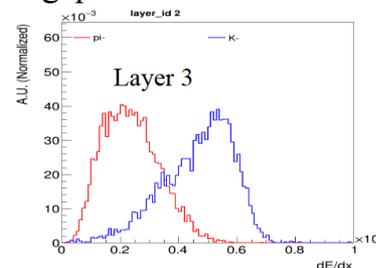
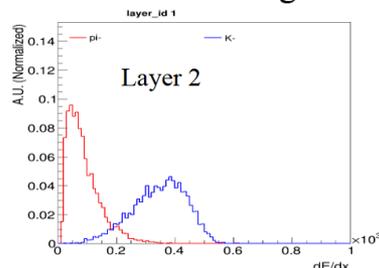
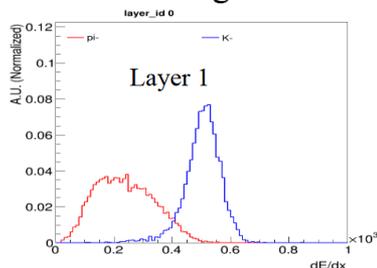
- ❖ CGEM + MDC dE/dx 流程：
  - 1, 对真实数据挑选样本，对MC直接产生
  - 2, CGEM 不同层的增益以及物质修正
  - 3, 分别对CGEM 和漂移室的mean值做修正
  - 4, 利用电子样本，得到漂移室和CGEM dE/dx mean值的相对关系
  - 5, 联合CGEM和漂移室，做截断平均，mean值刻度

## Preliminary dE/dx from CGEM full digitization

- From previous study: for very low momentum particles in CGEM+ODC, fewer (or even no) MDC hits, causes a decrease in dE/dx PID efficiency comparing to MDC case
- It is worth to study if the dE/dx from the CGEM can be used for PID
- MC sample:  $\pi^-$ ,  $K^-$ ,  $p=0.2$  GeV/c,  $\cos\theta = 0$ ,  $\phi=0$
- With CGEM full digitization tuned with run 877~895 of cosmic-ray data
- Selection: MdcKalTrack, momentum difference between the reconstruction and MC truth  $|dP| < 0.02$  GeV/c, angle between them  $dAngle < 0.35$  rad

Wenxin Fang

- ✓ dE: total charge of a 2D-cluster
- ✓ dx: track length in the drift gap



- ✓ Preliminary CGEM dE/dx simulation results show some PID capability
- Plan: prepare calibration, combination codes; need data confirmation

[BESIII Physics and Software Workshop in Autumn 2024, Sep. 23, UCAS](#)

# JUNO $^{14}\text{C}$ pile-up 鉴别

- ❖ JUNO 20 千吨液闪的  $^{14}\text{C}$  放射性活度为4万 Bq (目标值), 导致  $e^+$  能量分辨率变差, 不能忽略
- ❖ 难点:  $^{14}\text{C}$  的信号幅度比  $e^+$  小一个量级以上, 当  $^{14}\text{C}$  和  $e^+$  在时间和空间上接近时,  $^{14}\text{C}$  pile-up 鉴别具有困难
- ❖ 基于机器学习方法的  $^{14}\text{C}$  pile-up 鉴别:
  - 基于 CNN 以及 Transformer 神经网络
  - 大部分区域有很高鉴别效率

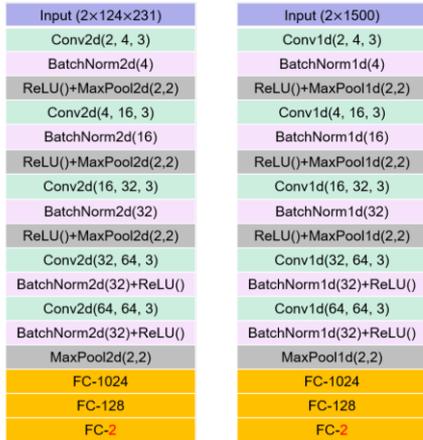


Figure 6. The architecture of the 2DCNN model.

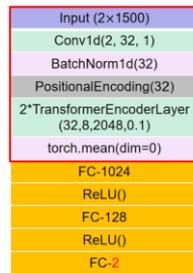


Figure 7. The architecture of the 1DCNN model.

Figure 8. The architecture of the Transformer-based model.

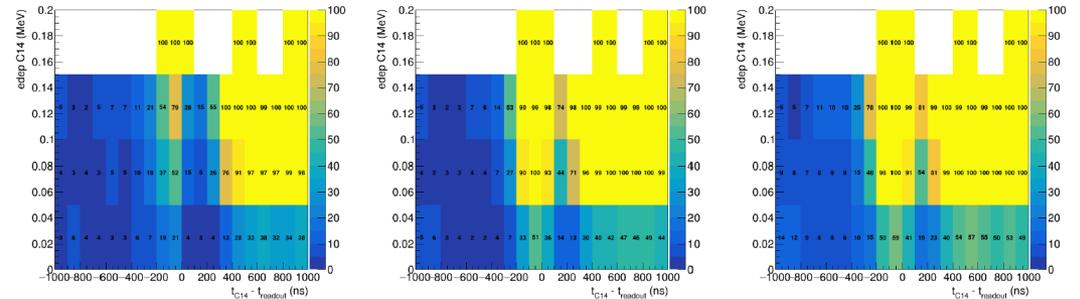
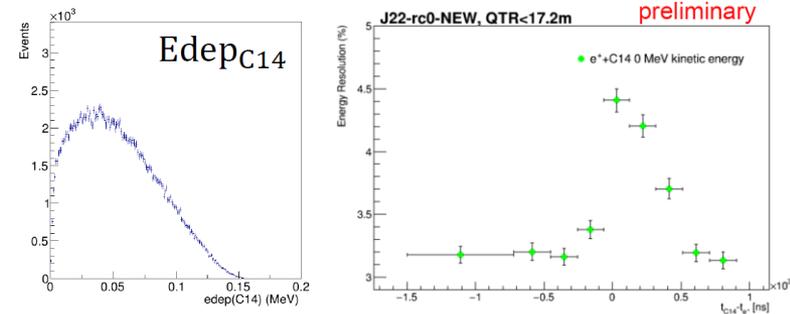
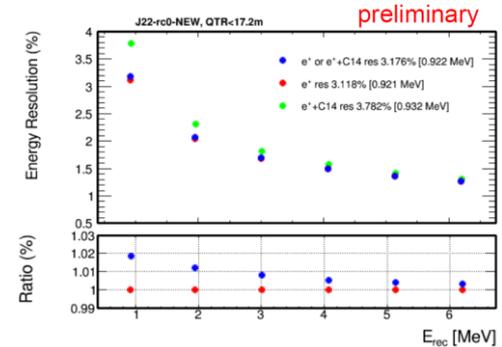


Figure 8. The pile-up event identification efficiency for  $e^+$  with zero kinetic energy. The efficiency for pure  $e^+$  is 99%. The left plot is for 2DCNN model, the middle plot is for 1DCNN model, and the right plot is for Transformer-based model. X axis is time difference between  $^{14}\text{C}$  and electronics readout. Y axis is the deposited energy of  $^{14}\text{C}$  (for  $e^-$  from beta decay of  $^{14}\text{C}$ ).



# 在线事例分类软件 (OEC)

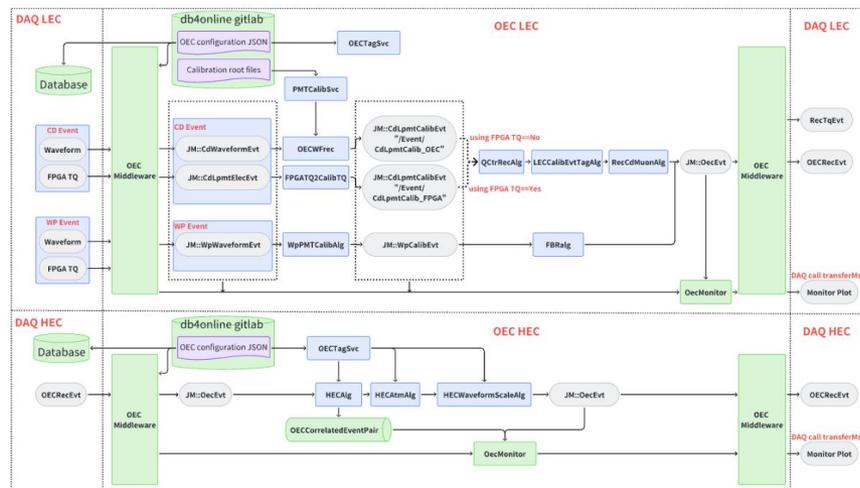
❖ JUNO 经过触发后的事例率为  $\sim 1\text{kHz}$ , 数据量  $\sim 40\text{ GB/s}$  (波形数据全存)

- 需要对事例进行在线分类 (OEC), 判断该事例需不需要保存波形数据, 实现从  $\sim 40\text{ GB/s}$  降低到  $\sim 60\text{ MB/s}$
- OEC 的性能直接影响 JUNO 的数据质量, 十分关键

❖ OEC 软件包括离线 OEC 框架和算法, 在线 OEC middleware (用于将离线框架和算法整合到在线环境中)

❖ 工作内容:

- 系统梳理 OEC 软件
- 开发, 完善 OEC 软件
- 发布 OEC 软件
- 担任 OEC 组协调人之一



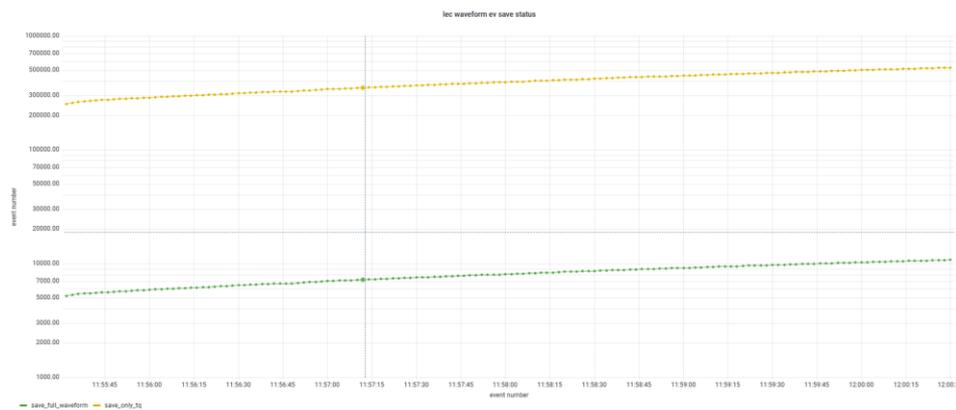
# 在线事例分类软件 (OEC)

## ❖ 新增和完善重要功能:

- 支持挑选刻度事例
- 支持使用 FPGA TQ 做重建
- 支持灵活的 pre-scale
- 完善 OEC 配置文件管理
- 完善数据质量监控图
- 优化 OEC 算法初始化

## ❖ OEC 软件发布:

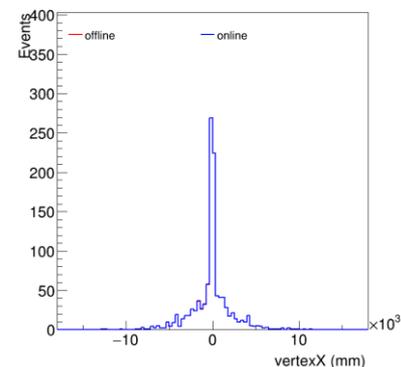
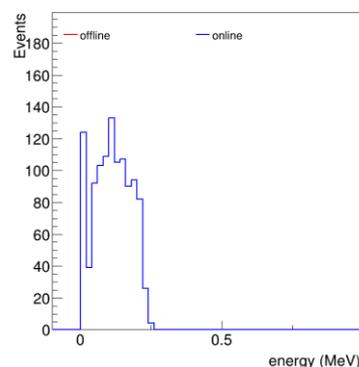
- v0.2.19.data
  - 7月19关灯测试, Run775
- v1.0.2
  - 用于12月份灌水



The proportion of events that require storing the full waveform meet the pre scale of 2%

## □ Run775:

- Pre-scale 结果符合预期
- 检查原始数据 (波形) 在离线和在线环境重建的结果, 结果一致

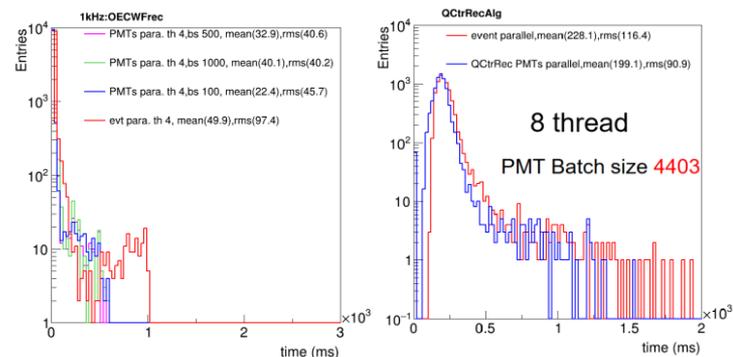


# 基于新版本 Sniper 的 OEC 多线程软件

❖ 新版本 Sniper 支持事例间、事例内（算法间）、算法内并行

❖ 开发 PMT 级别并行的波形重建和事例重建算法

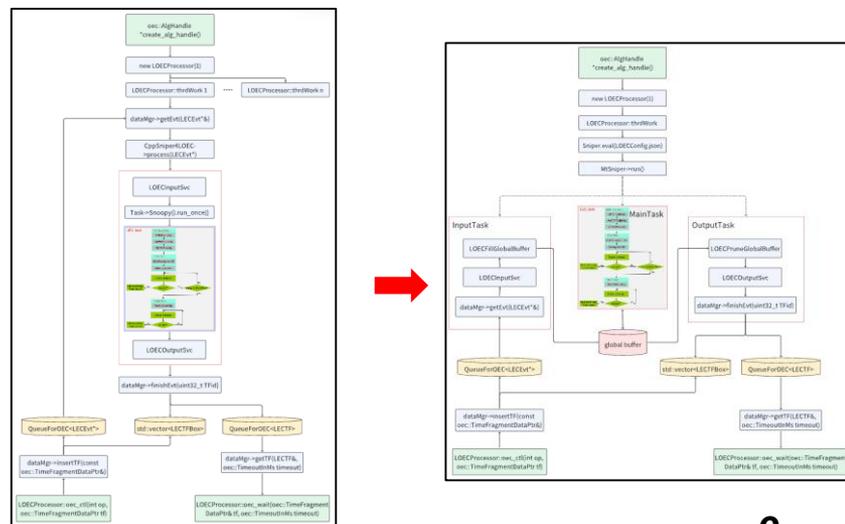
- 测试：物理性能一致，能减小事例最大处理时间
- 对事例重建算法，有时会报错（仍需要解决）



❖ 在线 OEC middleware 的更新

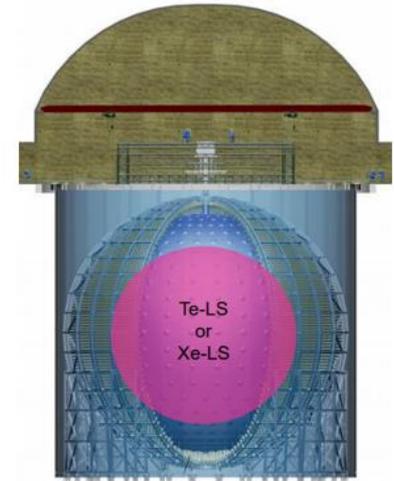
- 保持对 DAQ 的接口不变
- 重新设计 LEC 内部数据处理流程
- 开发新的算法处理数据读入、写出
- 能正常运行，需要更多的性能测试

Updates of OEC offline software,  
24th JUNO Collaboration Meeting



# $0\nu\beta\beta$ 中的太阳中微子本底

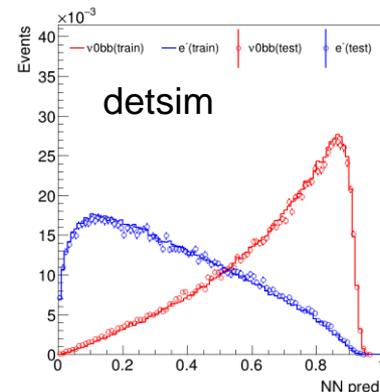
- ❖ 在中微子质量顺序测量完成后，中微子物理的关键问题
  - 中微子的绝对质量？
  - 寻找  $0\nu\beta\beta$  衰变，中微子是 Majorana 粒子？
- ❖ JUNO 将在 2030 年左右升级成世界上最好的寻找  $0\nu\beta\beta$  的实验
- ❖ 本底的排除对  $0\nu\beta\beta$  实验至关重要，其中来自  $^8\text{B}$  太阳中微子的本底是主要本底（~50%）
- ❖ 进一步压低  $^8\text{B}$  太阳中微子本底对该实验具有重要物理意义



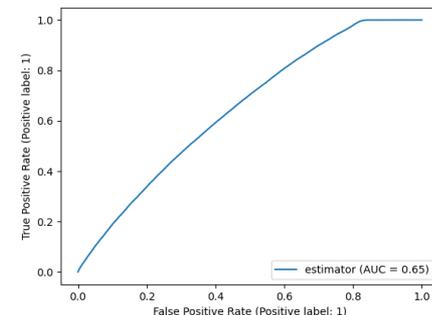
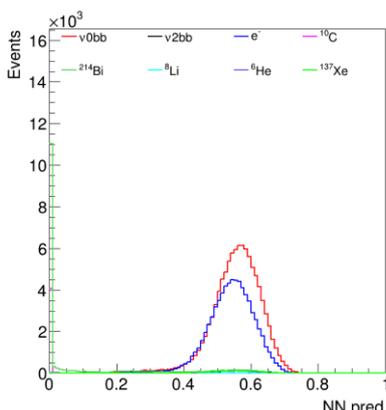
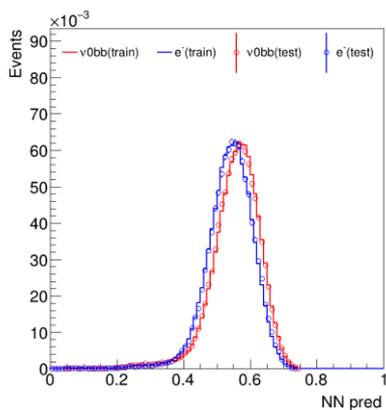
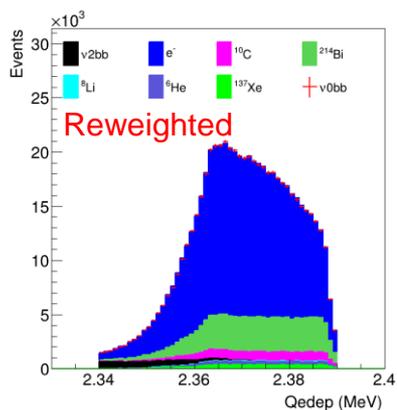
summary of backgrounds in $0\nu\beta\beta$ ROI	
[ROI·(ton $^{136}\text{Xe}$ )·yr] $^{-1}$	
$2\nu\beta\beta$	0.2
$^8\text{B}$ solar $\nu$	0.7
cosmogenic background	
$^{10}\text{C}$	0.053
$^6\text{He}$	0.063
$^8\text{Li}$	0.016
$^{12}\text{B}$	$3.8\times 10^{-4}$
others ( $Z\leq 6$ )	0.01
$^{137}\text{Xe}$	0.07
internal LS radio-purity ( $10^{-17}$ g/g)	
$^{214}\text{Bi}$ ( $^{238}\text{U}$ chain)	0.003
$^{208}\text{Tl}$ ( $^{232}\text{Th}$ chain)	—
$^{212}\text{Bi}$ ( $^{232}\text{Th}$ chain)	0.03
external contamination	
$^{214}\text{Bi}$ (Rn daughter)	0.2
total	1.35

# $0\nu\beta\beta$ 中的太阳中微子本底

- ❖ 难点： $^8\text{B}$  太阳中微子本底和  $0\nu\beta\beta$  事例的区别主要来自事例中的切伦科夫光。而切伦科夫光（没有经过吸收重发射）占总闪烁光比例  $< 1\%$
- ❖ 利用 GNN 网络，将 PMT 击中类比成 Graph 中的 node，区分信号和本底事例
- ❖ 初步结果表明网络能区分太阳中微子本底，性能的瓶颈主要来自于 PMT 的时间分辨

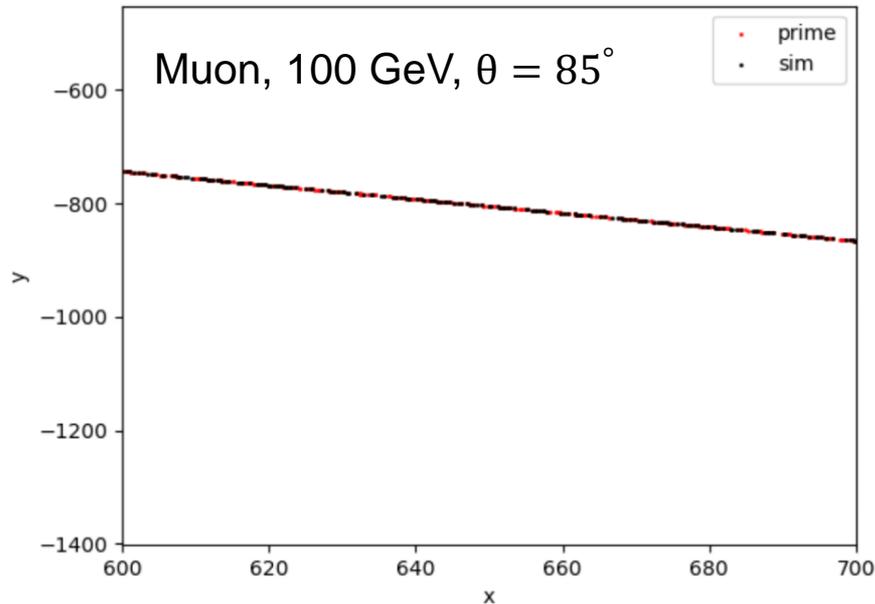


- Detsim data
- Hama 50% earliest
- NNVT 15% earliest
- 1D hit time (tof corrected) for all Hama
- 1D hit time (tof corrected) for all NNVT
- $e^-$  direction smear
- TTS
- T0 smear (2ns)
- Vertex smear
- Energy smear
- Dark Noise



# CEPC

- ❖ CPEC 探测器 baseline 设计采用 TPC (漂移室为 backup)
- ❖ 将之前为漂移室开发的 Geant4 结合 Heed 的模拟方法用于 TPC, 实现更加精确的原初电离模拟, 用于  $dN/dx$  PID 研究



# 论文发表情况

---

## ❖ 期刊文章（第一作者）：

- A data-driven dE/dx simulation with normalizing flow, Nucl.Instrum.Meth.A 1065 (2024) 169544

## ❖ 期刊文章（合作作者）

- Offline filter of data with abnormal high voltage at BESIII drift chamber, JINST 19 (2024) 06, P06042
- Offline data processing system of the BESIII experiment, Eur.Phys.J.C 84 (2024) 9, 937

## ❖ 会议文章（第一作者）：

- Refined drift chamber simulation in the CEPC experiment, EPJ Web Conf. 295 (2024) 03034
- Fast muon simulation in the JUNO experiment with neural networks, EPJ Web Conf. 295 (2024) 09019

## ❖ 会议文章（合作作者）：

- Key4hep Progress Report on Integrations, EPJ Web Conf. 295 (2024) 05010

# 学术交流

---

## ❖ 会议报告：

- Deep Learning-Based C14 Pile-Up Identification in the JUNO Experiment, ACAT 2024, 口头报告
- dE/dx Software in the BESIII Experiment, CHEP 2024, poster
- ML for fast calorimeter simulation, 第十四届全国粒子物理学术会议, 口头报告
- Deep Learning-Based C14 Pile-Up Identification in the JUNO Experiment, Quantum Computing and Machine Learning Workshop
- Updates of OEC offline software, JUNO 24th collaboration meeting
- Updates of OEC offline software, JUNO 23rd collaboration meeting

## ❖ 日常会议：BESIII 运行会、软件组会、数据质量会、OEC 组会、JUNO Physics/Simulation 等

# 负责（参与）的科研项目、经费情况

项目名称	项目类别	经费	本人角色	状态
BESIII漂移室单元 响应模拟研究	NSFC青年基金	30万	主持	执行 2023-2025
CEPC漂移室模拟 研究	谢家麟基金	4万	主持	执行 2023-2024
先进实验数据处理 软件研发	所创新项目	100万	主要参与者	已结题
机器学习在实验高 能物理中的应用	所创新项目	150万	主要参与者	执行 2023-2025

# 公共服务和其它贡献

---

- ❖ 协助指导学生开展研究工作：
  - JUNO C14 鉴别：吴肇祥
  - JUNO OEC Monitor：陈世强
  - JUNO 无中微子双 beta 衰变中的太阳中微子本底研究：客座、科创学生
- ❖ 担任答辩秘书：赵天灏，胡宇翔
- ❖ 参与 CEPC TDR 撰写

# 下一年度工作计划

---

## ❖ BESIII 实验:

- 研究基于 CGEM 的  $dE/dx$  刻度和重建
- 新数据的  $dE/dx$  精细刻度和模拟
- 挖掘机器学习在 BESIII 实验的应用

## ❖ JUNO 实验:

- 主要集中在 OEC 方面的工作

## ❖ CEPC 实验:

- 支持 TDR 的撰写

## ❖ 其它: 中心及组里安排的其它任务

谢谢!