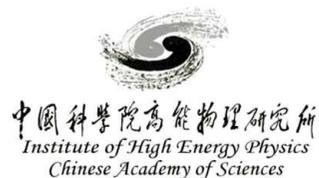


基于电离计数方法的CEPC 漂移室粒子鉴别研究

刘帅毅

代表CEPC漂移室PID工作团队

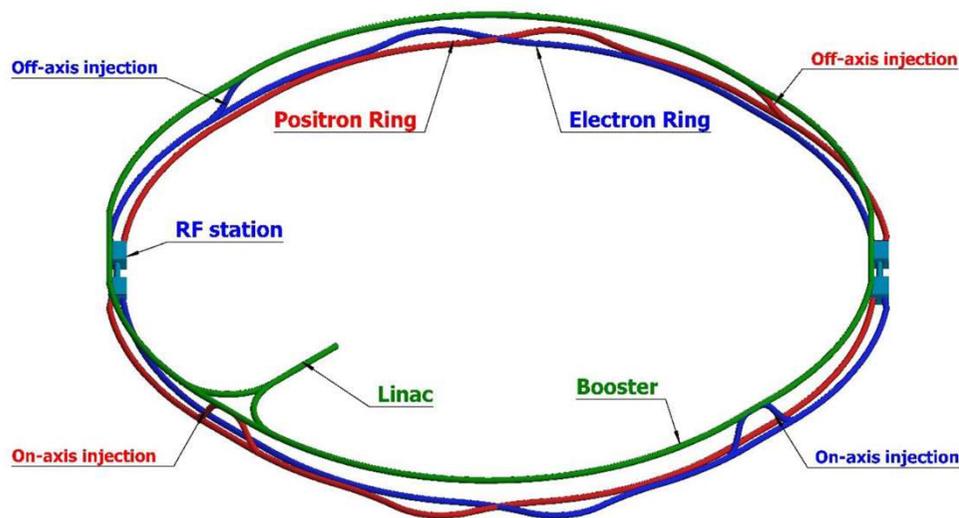


第十一届全国先进气体探测器研讨会，2023年10月12日-14日，北京大兴

目录

- 简介
- 模拟研究
- 模型测试
- 总结

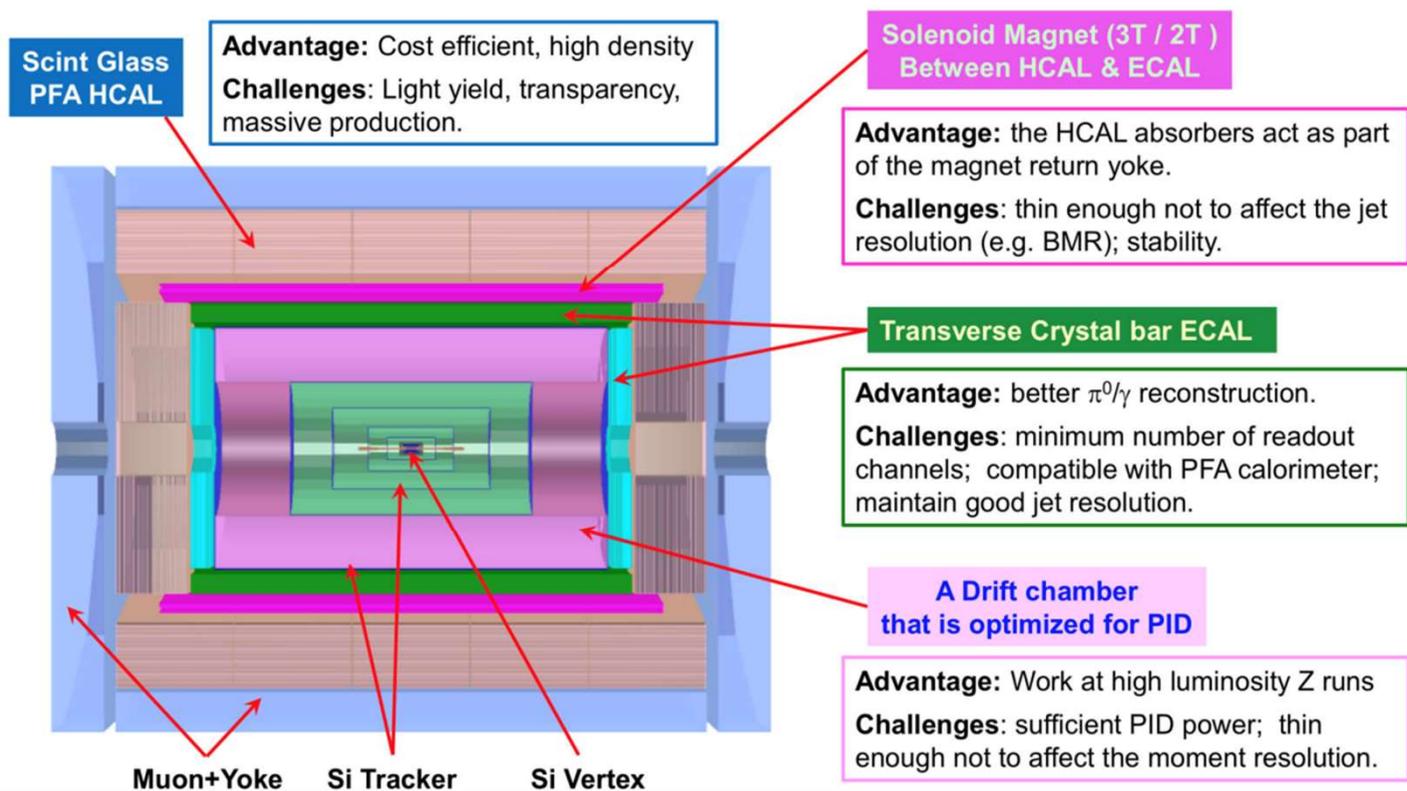
CEPC(Circular Electron Positron Collider)



- CEPC (Circular Electron Positron Collider, 环形正负电子对撞机) 设计为周长100 km的双环双对撞点的对撞机

Operation mode		ZH	Z	W ⁺ W ⁻
\sqrt{s} [GeV]		~240	~91.2	158-172
Run time [years]		7	2	1
CDR	L / IP [$10^{34} \text{ cm}^{-2}\text{s}^{-1}$]	3	32	10
	$\int L dt$ [ab^{-1} , 2 IPs]	5.6	16	2.6
	Event yields [2 IPs]	1×10^6	7×10^{11}	2×10^7
Latest	L / IP [$10^{34} \text{ cm}^{-2}\text{s}^{-1}$]	5	105.5	18.7

CEPC第四个探测器设计方案中的漂移室



- 径迹探测器由硅探测器和漂移室组成
- 漂移室提供基于电离计数的粒子鉴别
- 物理目标要求 K/π 鉴别能力到20GeV/c动量下高于 2σ
- 相比全硅径迹探测器系统，在低动量下有更好的动量分辨

漂移室的优点

- 粒子鉴别

- 减少组合本底
- 提高质量分辨
- 提高喷注能量分辨
- 有助于味标记

- 寻迹

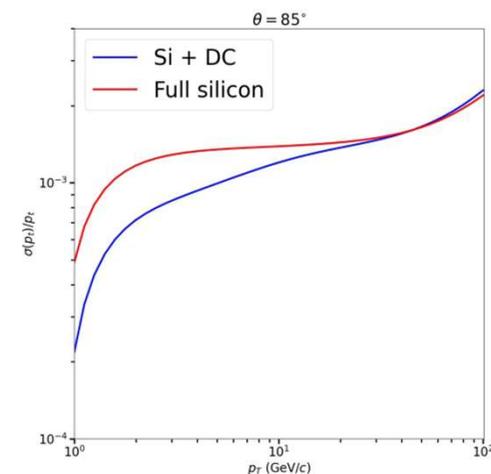
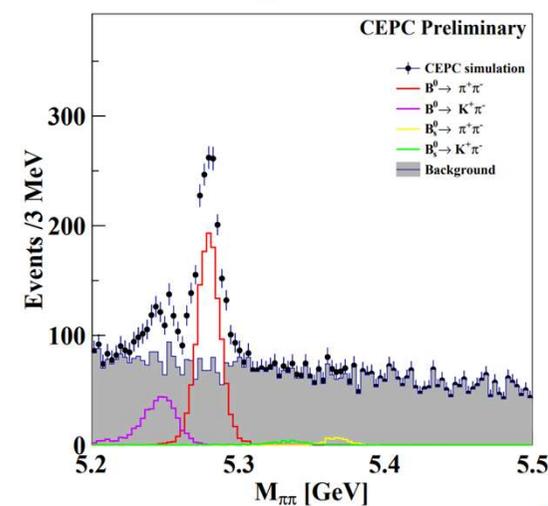
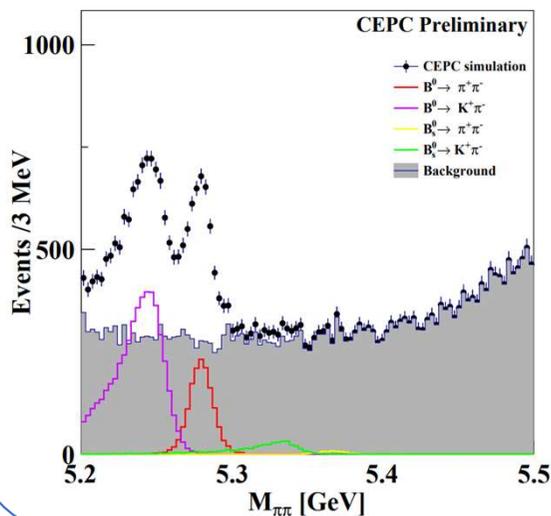
- 在低动量下有更好的动量分辨

2023/10/14

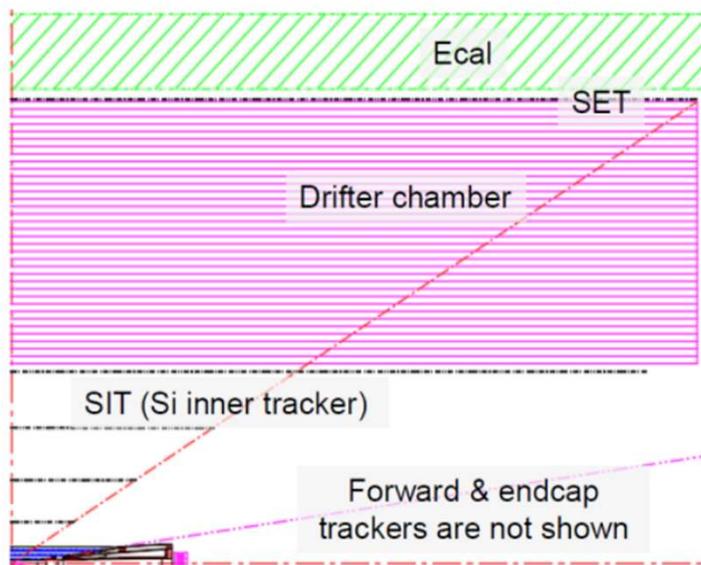
$B^0/B_s^0 \rightarrow \pi^+\pi^-$ 基于Delphes的模拟

没有PID

有PID



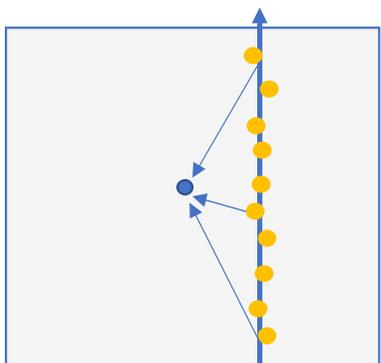
漂移室的初步设计



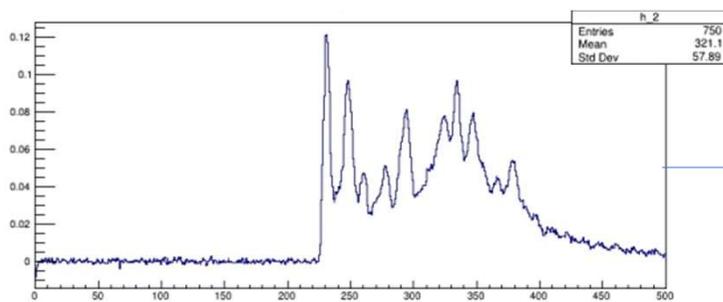
漂移室的初步参数	
内径	800mm
外径	1800mm
单元大小	18 mm×18 mm
气体混合比	He/iC ₄ H ₁₀ =90:10
最外侧丝的长度 (cosθ=0.82)	5143mm

dE/dx 测量

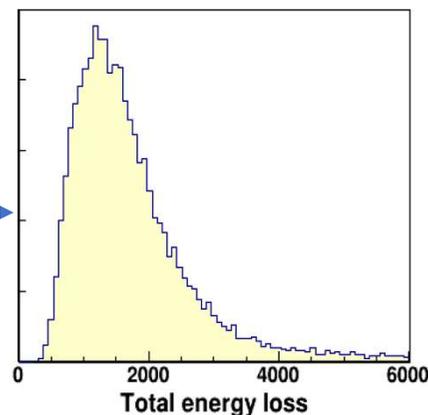
- 电离能损(dE/dx) 通过对波形积分进行测量
- 包含初级电离和次级电离的涨落, 服从朗道分布



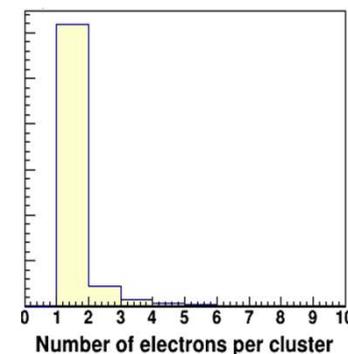
漂移室单元内电离



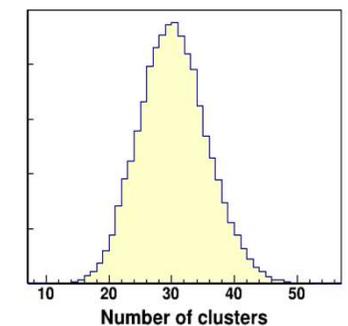
波形



总电离



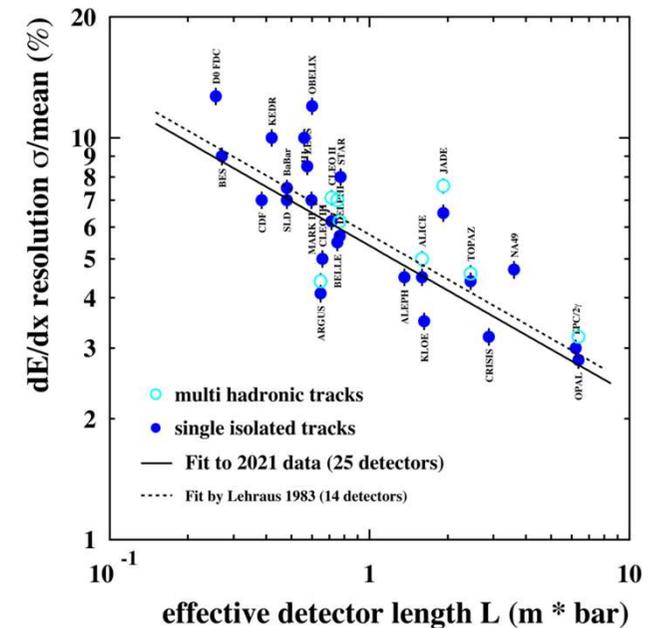
单个电离团的电离数分布



初级电离 7

dE/dx 方法的问题

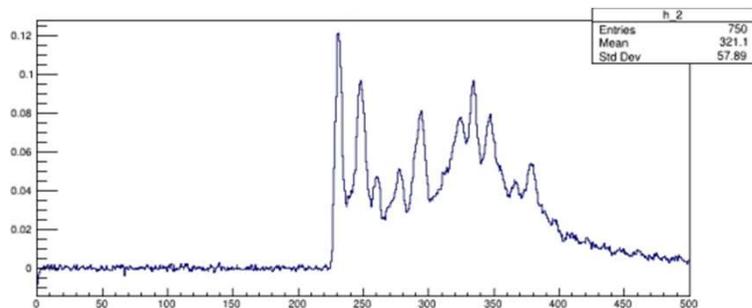
- 多项误差来源
 - 平均电离能的涨落
 - 单个cluster可能释放一个或多个电子
 - 雪崩放大的涨落
- 朗道分布采用截断平均法，丢失统计量
- 在过去40年内分辨没有显著提高
 - 1983: $dE/dx \text{ res.} = 5.7 \times L^{-0.37} (\%)$
 - 2021: $dE/dx \text{ res.} = 5.4 \times L^{-0.37} (\%)$



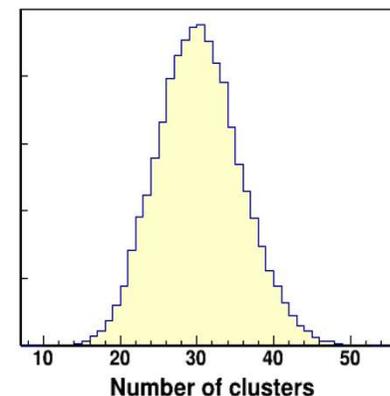
From Micheal Hauschild's talk @RD51 workshop

电离计数方法 (dN/dx 测量)

- 探测单位径迹长度上的初级电离团数目
- 初级电离团数服从泊松分布，有更好的分辨，预期能提高PID性能
- 关键技术：
 - 高采样率低噪声电子学
 - 高性能寻峰算法（处理信号堆积，识别噪声）



电离计数重建

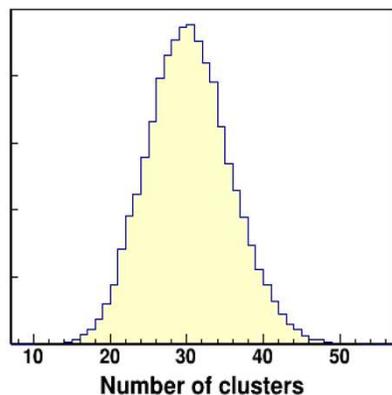


初级电离

dE/dx vs dN/dx

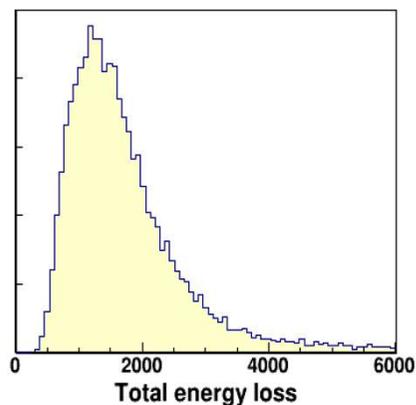
dN/dx

- 单位长度内初级电离数
- 泊松分布
- 更小的涨落

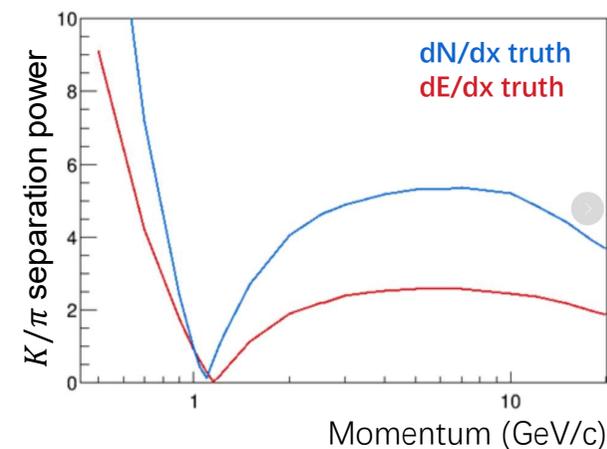


dE/dx

- 单位长度内电离能损
- 朗道分布
- 更大的涨落



K/π 分辨能力 dN/dx vs dE/dx



dN/dx 在 20 GeV/c 动量下分辨能力显著高于 dE/dx 方法(MC truth)

理论上, 1m径迹下, 相比 dE/dx , dN/dx 分辨可以提高一倍:
 dN/dx 分辨: $\sim 2.6\%$ (statistical error only)
 dE/dx 分辨: 5.4% (Lehraus fit)

dN/dx 测量的挑战

• 信号测量的挑战

- 探测器设计, 低漂移速度、低电离团密度气体 \longrightarrow 增加不同电离团信号的时间间隔
- 低扩散、压低多电子电离团的气体 \longrightarrow 减少不同电离团之间的干扰
- 快速电子学
 - 带宽 > 1GHz
 - 前放增益 > 10
 - 采样率 > 1.5GS/s
 - 比特分辨 > 12bit

电离团信号快速成型, 压低噪声

• 重建算法的挑战

- 处理信号堆积
- 去除噪声干扰
- 识别初级和次级电离信号

dN/dx 原理由 A. Davidenko et. al. 在 1969 年首次提出, 由于硬件技术的限制, 尚未在大型实验应用过 (JETP, 1969, Vol. 28, No. 2, p. 223)

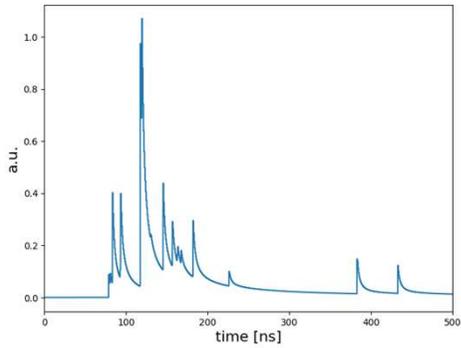
随着电子学技术的发展, 使得 dN/dx 的应用成为可能。目前包括FCC-ee, CEPC, ILC 等均提出了应用 dN/dx 技术的大型气体探测器方案

模拟研究

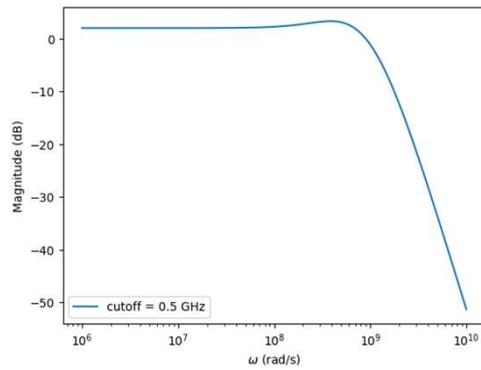
- 传统算法
- 机器学习算法

模拟研究

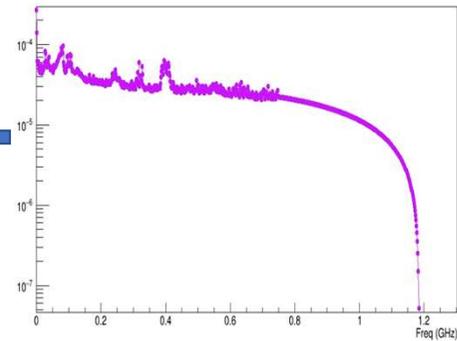
生成信号
(模拟)



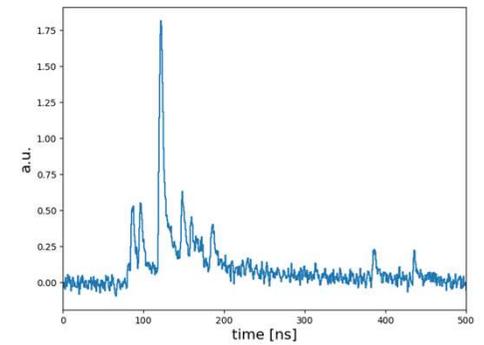
电子学响应
(束流测试的电子学)



噪声
(束流测试)

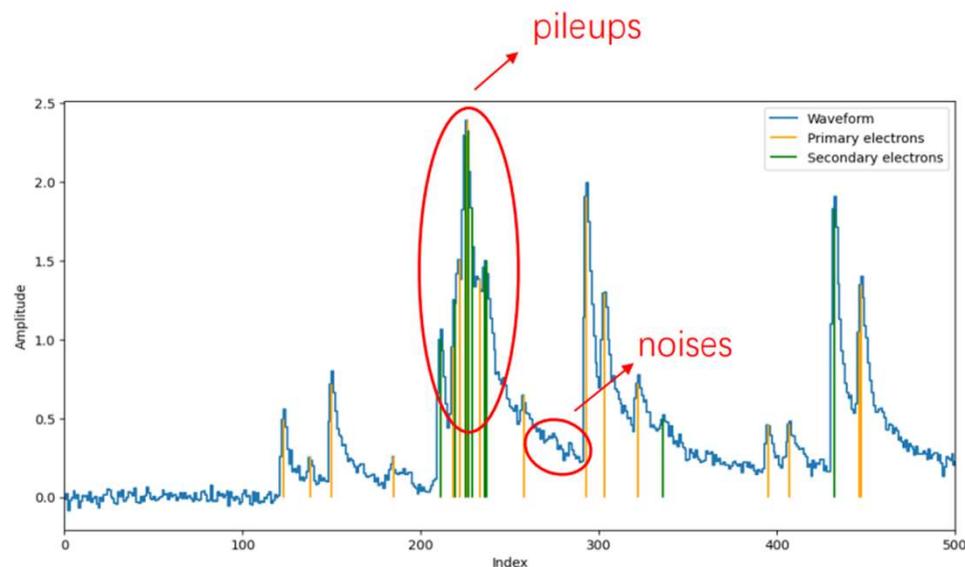


模拟波形



重建算法

- 重建算法的挑战
 - 处理信号堆积
 - 去除噪声干扰
 - 识别初级和次级电离信号
- 两种方案正在研究
 - 传统算法 (已完成)
 - 基于二阶求导的寻峰+次级电离聚团
 - 基于深度学习的算法 (研究中)
 - 基于循环神经网络 (LSTM) 的寻峰 + 基于图神经网络 (DGCNN) 的聚团



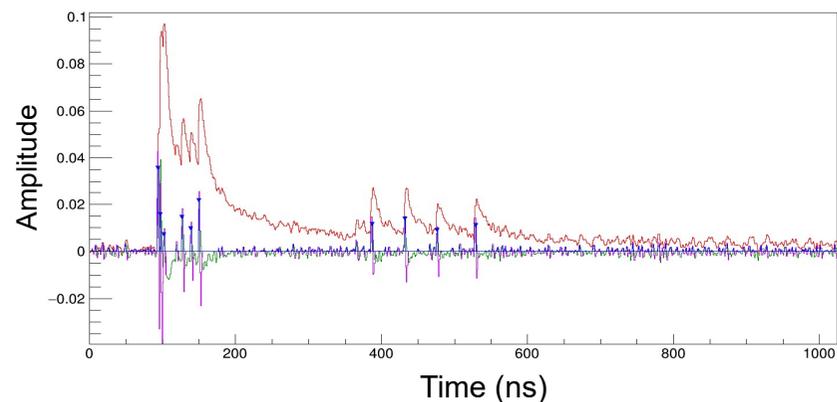
传统算法

寻峰

- 寻找信号峰
- 一阶求导以及二阶求导
- 在上升沿斜率变化处寻峰

聚团

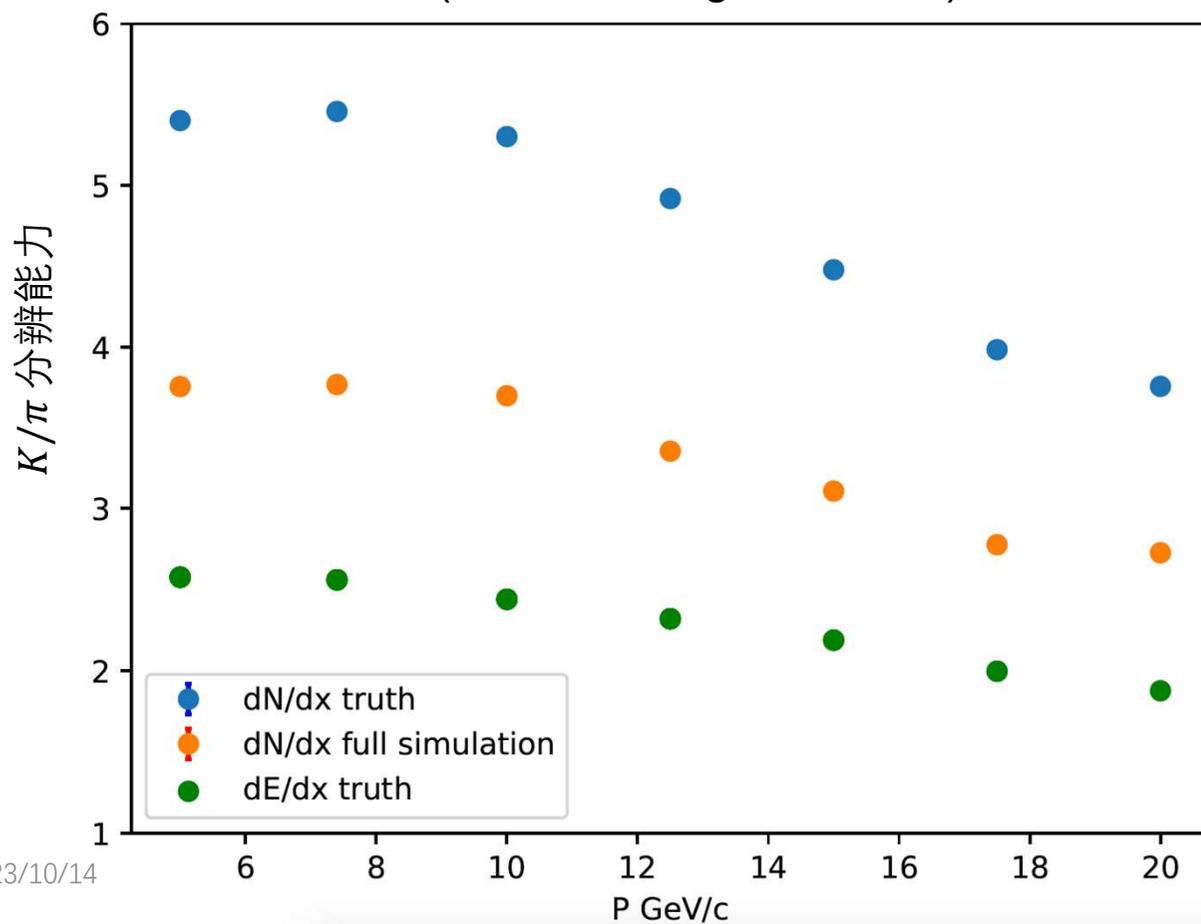
- 通过合并相邻的信号峰得到初级信号数
- 如果 $\Delta t < t_{cut}$, 就将峰合并为电离团
- t_{cut} 与扩散时间有关, 从模拟中估计得到



- 优势: 简单快速
- 劣势: 在高堆叠、高噪声下效率较低
- 开发完成, 已用于数据分析

K/π 分辨能力随动量的变化(传统算法)

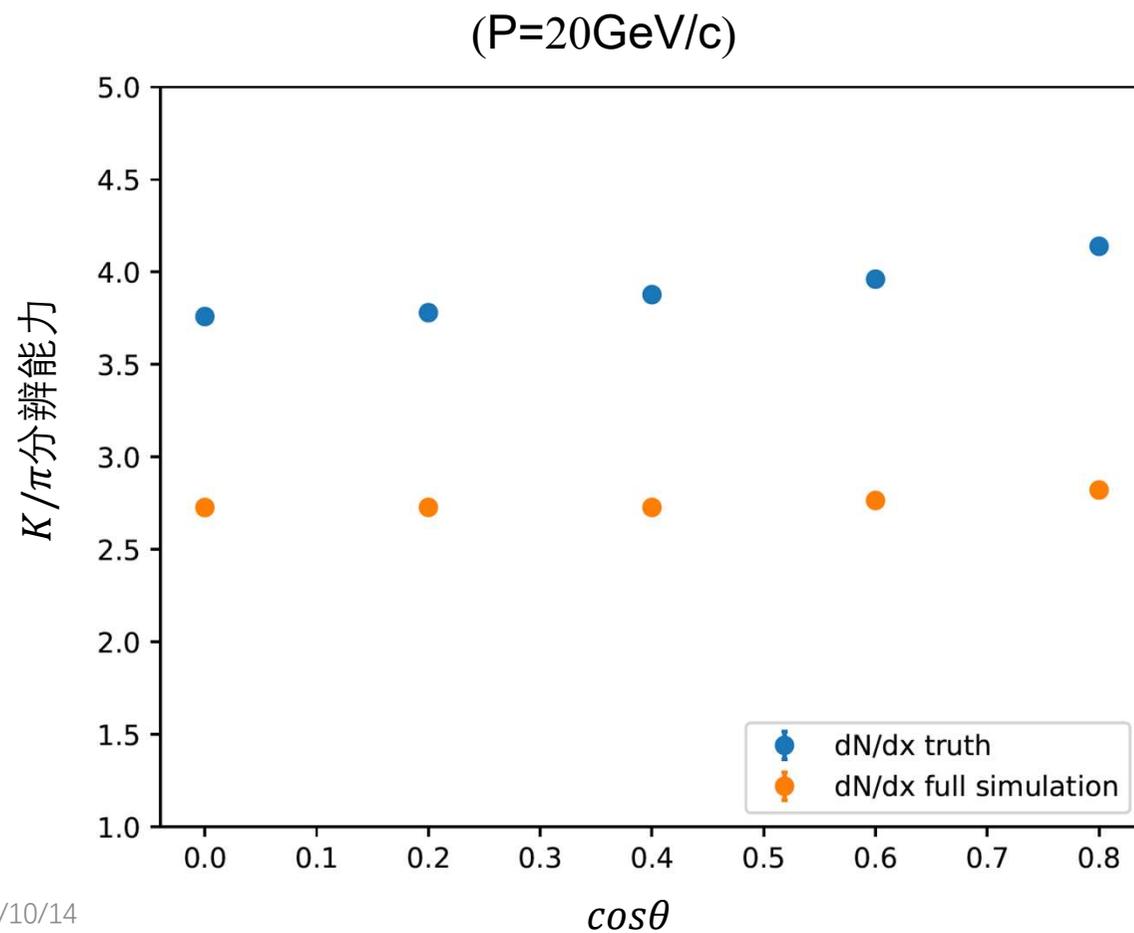
(1m track length, $\cos\theta=0$)



分辨能力

$$S = \frac{\left| \left(\frac{dN}{dx} \right)_{\pi} - \left(\frac{dN}{dx} \right)_{K} \right|}{(\sigma_{\pi} + \sigma_K)/2}$$

K/π 分辨能力随角度 $\cos\theta$ 变化(传统算法)



分辨能力

$$S = \frac{\left| \left(\frac{dN}{dx} \right)_{\pi} - \left(\frac{dN}{dx} \right)_{K} \right|}{(\sigma_{\pi} + \sigma_K)/2}$$

基于深度学习的算法

- 深度学习的优势：
 - 利用大统计量样本进行监督学习，充分提取数据特征
- 算法开发：
 - 寻峰算法：循环神经网络(LSTM)
 - 聚团算法：图神经网络(DGCNN)

LSTM + DGCNN对于 N_{cls} 的重建结果

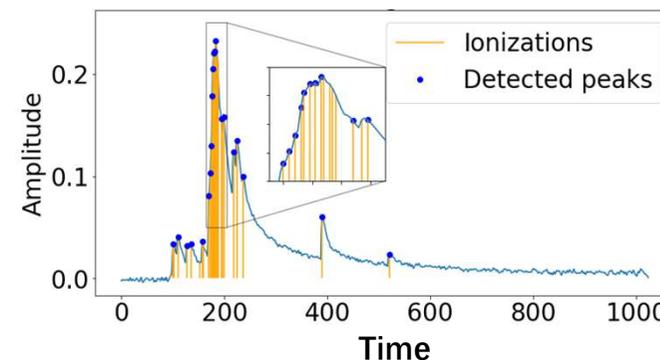
重建方法	μ	σ	σ/μ
MC truth (输入)	16.53	3.93	23.8%
传统算法	18.67	4.60	24.6%
深度学习	16.65	4.06	24.4%

2023/10/14

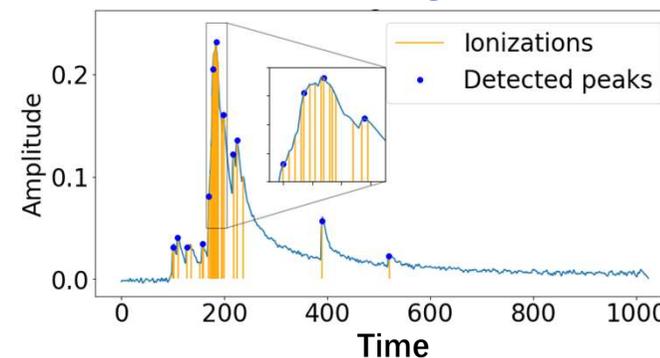
深度学习LSTM + DGCNN算法相比传统算法更加接近MC truth分布

LSTM的寻峰结果

LSTM algorithm



Derivative algorithm



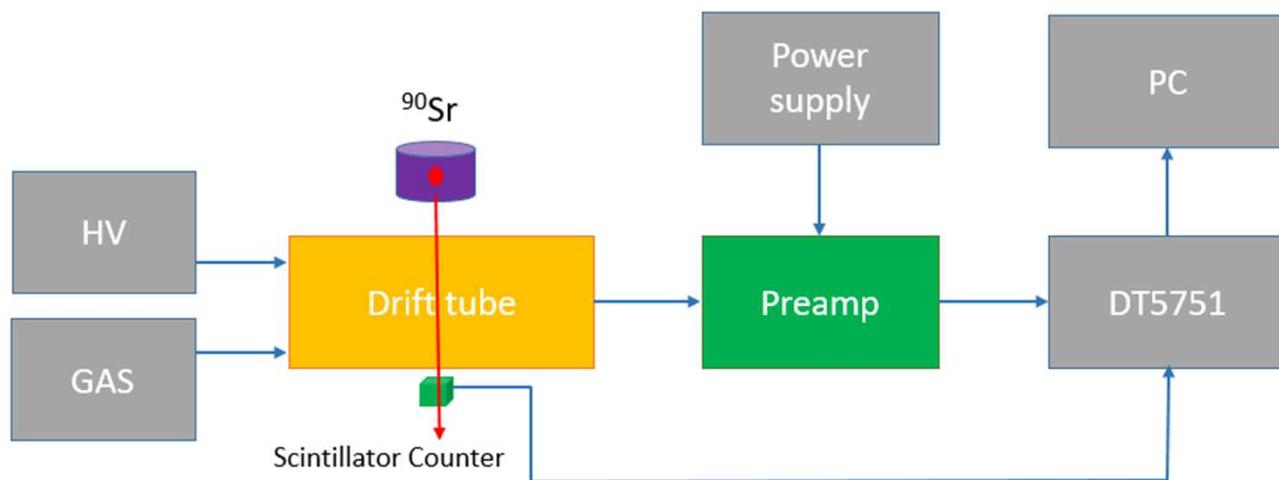
LSTM算法比传统导数算法对于堆积事例更加高效 18

漂移室模型测试

- 目的
 - 验证电离计数方法的可行性
 - 研究电离计数方法的关键技术
- 目前进行的测试
 - 放射源测试
 - 束流测试

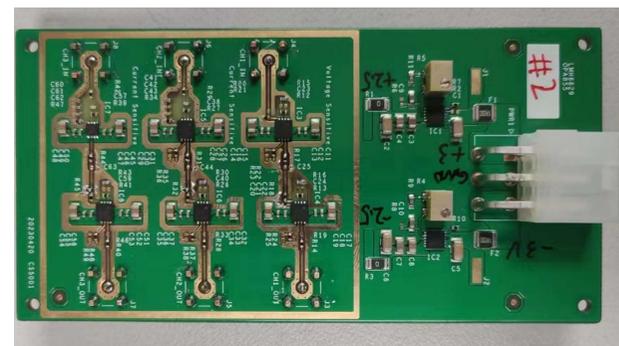
基于Sr90放射源的漂移管测试

测试系统



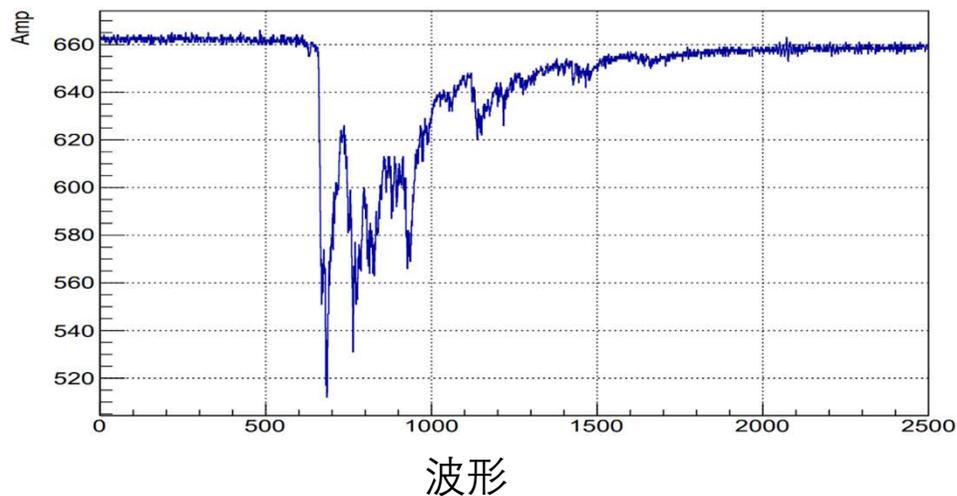
- 漂移管直径: 30mm
- 工作气体: He/iC₄H₁₀=90:10

前置放大器

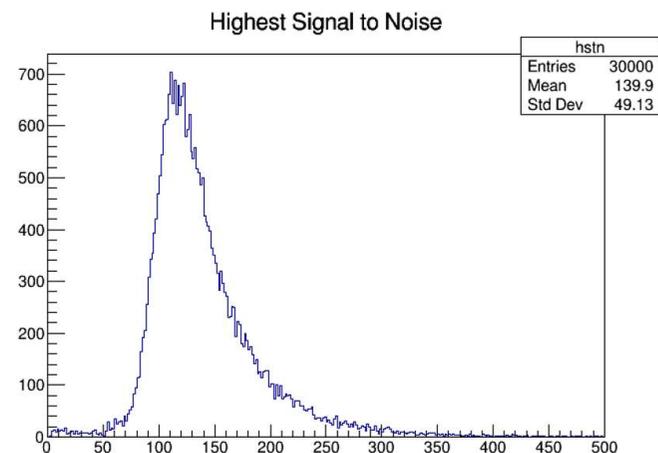


- 放大器带宽~1GHz
- ADC采样率1GHz

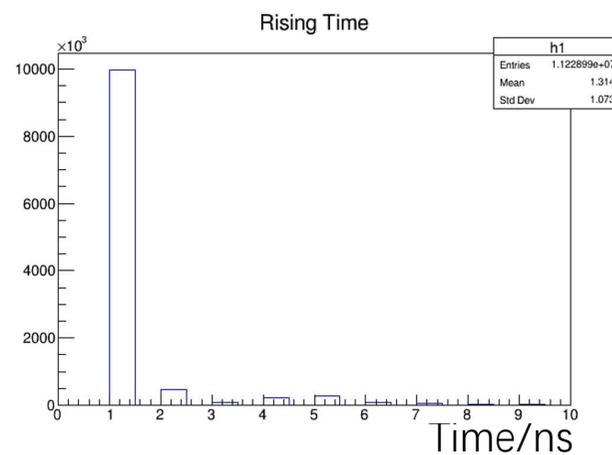
初步测试结果



- 最高信号信噪比 ~140
- 平均上升沿时间~1.3ns

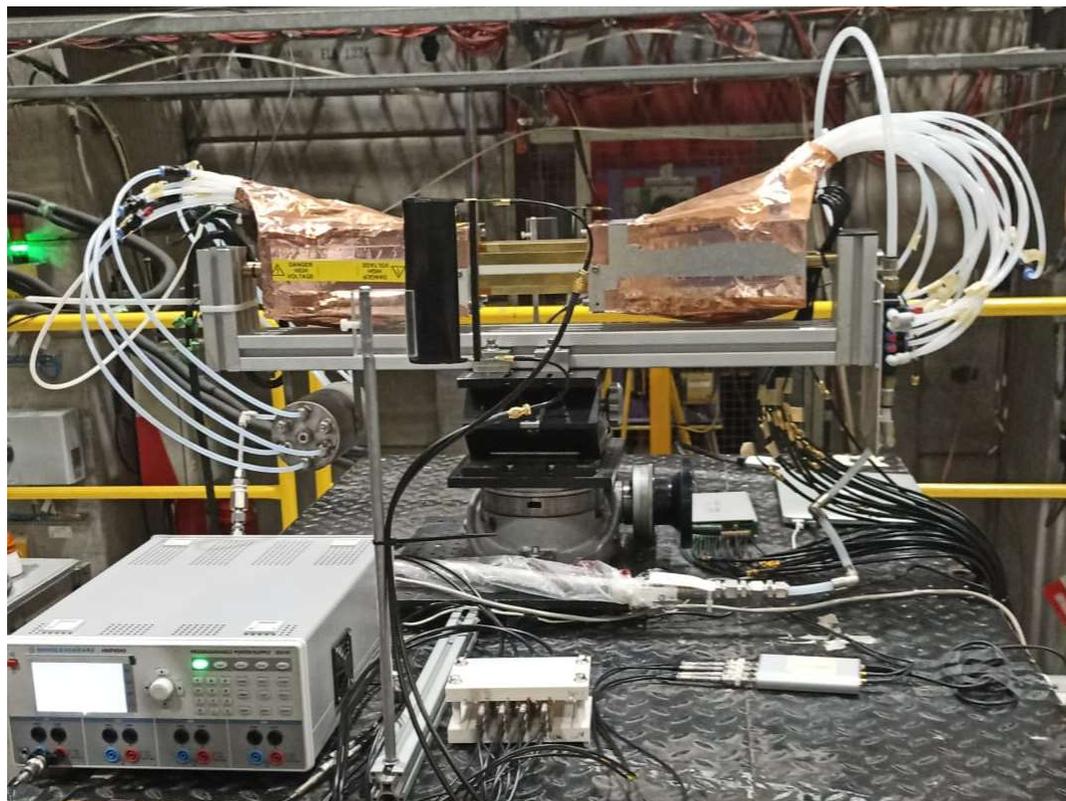
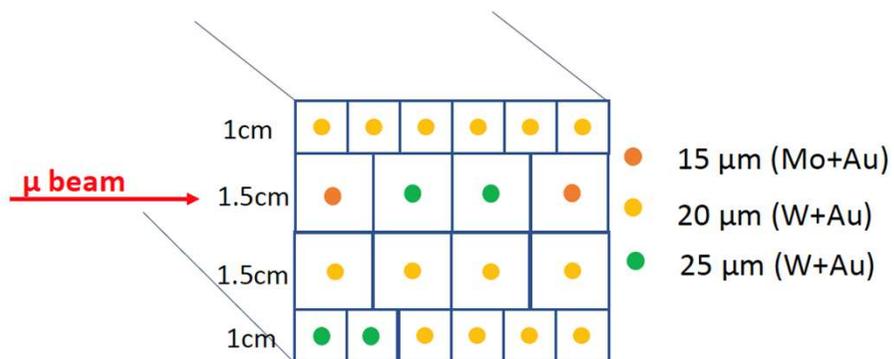


$$STN = \frac{\text{Highest Signal}}{RMS_{noise}}$$



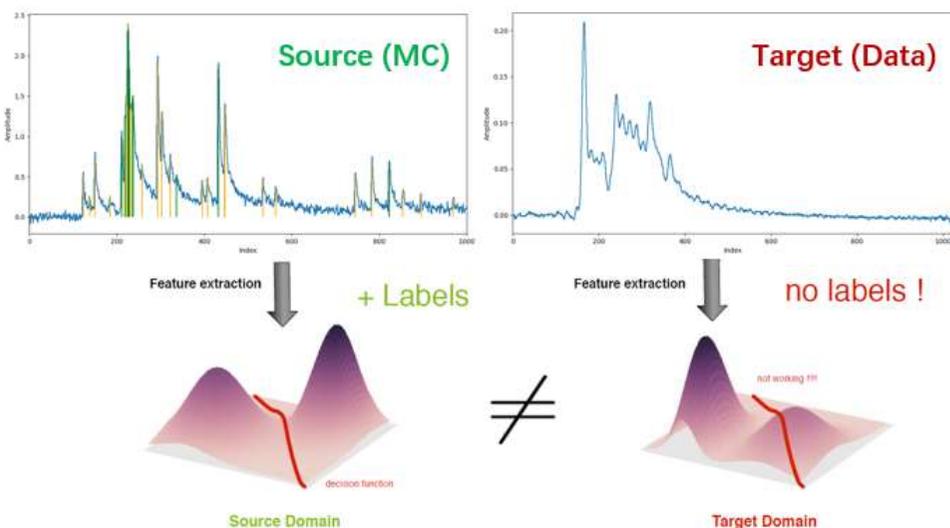
漂移室模型的束流测试

- 与INFN合作完成了3轮束流测试
 - 数据分析正在进行中
 - 已有初步的寻峰结果，算法正在优化中



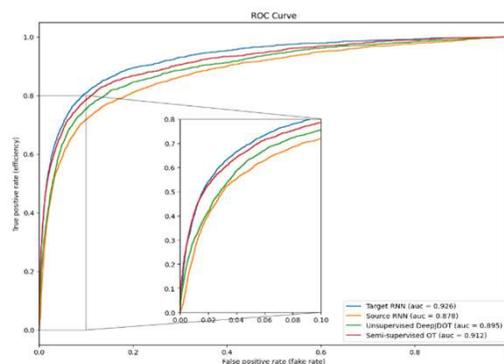
束流测试数据的深度学习重建算法

基于最优传输的模型，解决真实数据的训练问题



- MC与数据之间存在领域偏移
 - 数据缺少标签
- ➔ Semi-supervised DeepJDOT方法

利用MC数值实验进行验证，性能接近理想模型

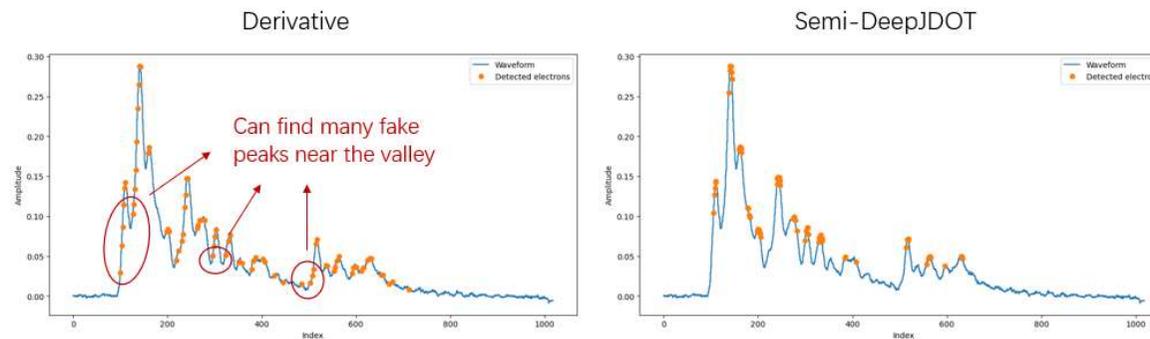


Model	AUC	pAUC (FPR<0.1)
理想模型	0.926	0.812
Source RNN	0.878	0.749
Unsupervised DeepJDOT	0.895	0.769
Semi-supervised DeepJDOT	0.912	0.793

- MC数值实验：利用MC样本和伪数据样本进行测试
- 理想模型：利用MC标注样本的全监督模型

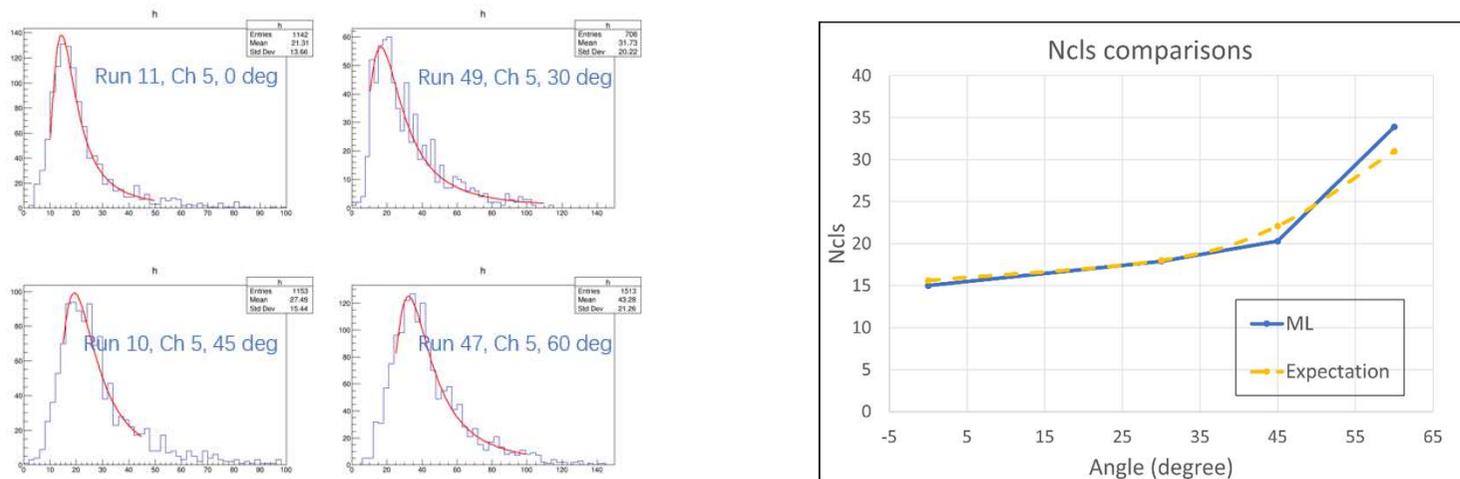
束流测试数据的深度学习重建算法

真实数据的深度学习
寻峰结果



深度学习算法具有更好的信号峰重建和噪声识别能力

不同角度样本
的分布结果



2023/10/14

信号峰重建数目随角度的变化趋势与理论预期基本符合

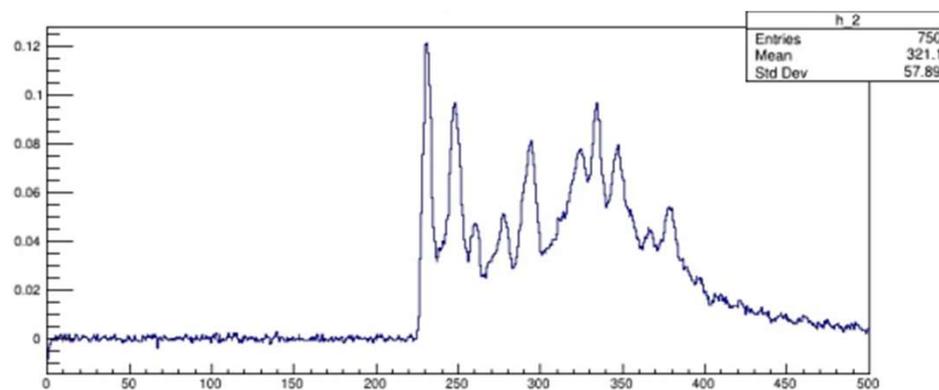
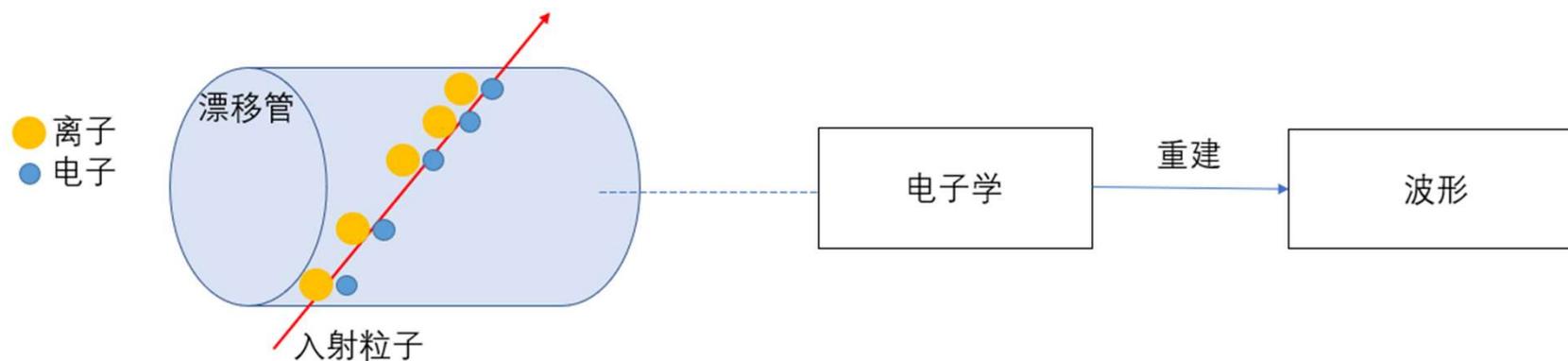
总结

- 在CEPC的漂移室设计中将采用电离计数方法以获得更好的粒子鉴别能力
- 初步模拟研究表明
 - 与传统的 dE/dx 相比, dN/dx 有更好的 K/π 鉴别能力, 20GeV/c达到 2σ 以上
- 开展了基于放射源和束流的模型测试, 正在进行中
 - 初步验证了电离计数方法用于粒子鉴别的可能性
- 工作计划
 - 重建算法的调试和优化
 - 实验数据分析
 - 探测器设计优化

Thanks!

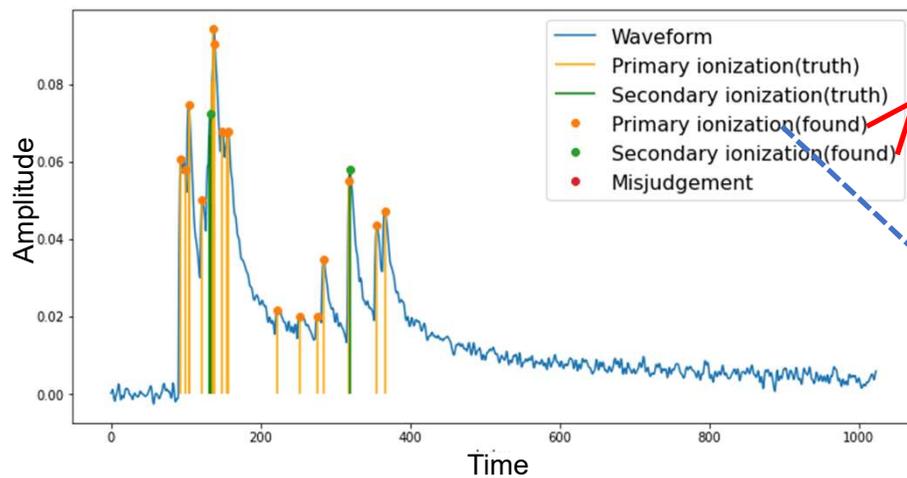
backup

漂移室探测原理-基于电离测量



波形

重建算法的两个步骤

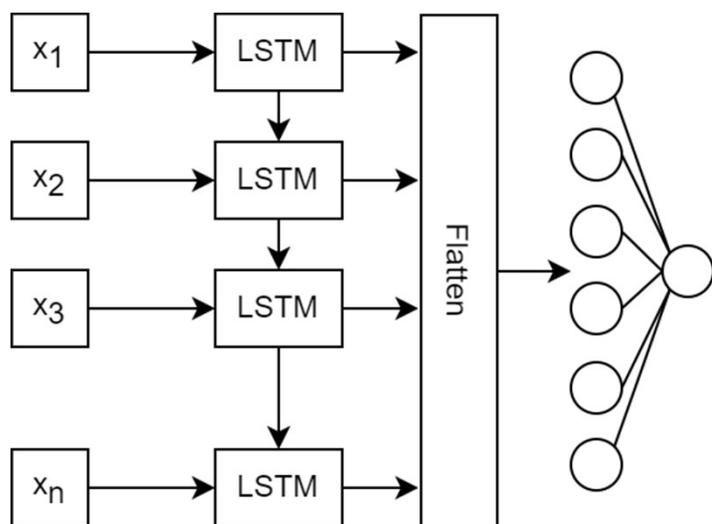


步骤1. 寻峰

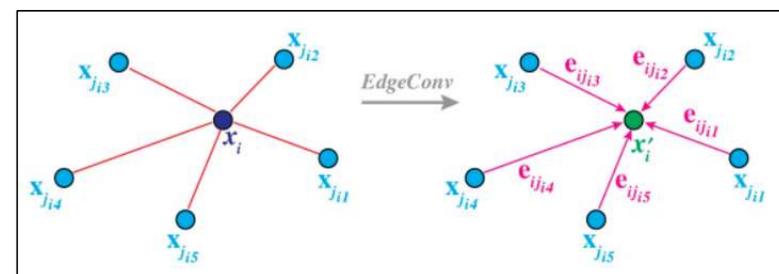
从噪声中辨别出峰的位置 (包括初级电离和次级电离) (分类问题)

步骤2. 簇团化:

从已经寻好的峰中判断簇团的数量 (回归问题)



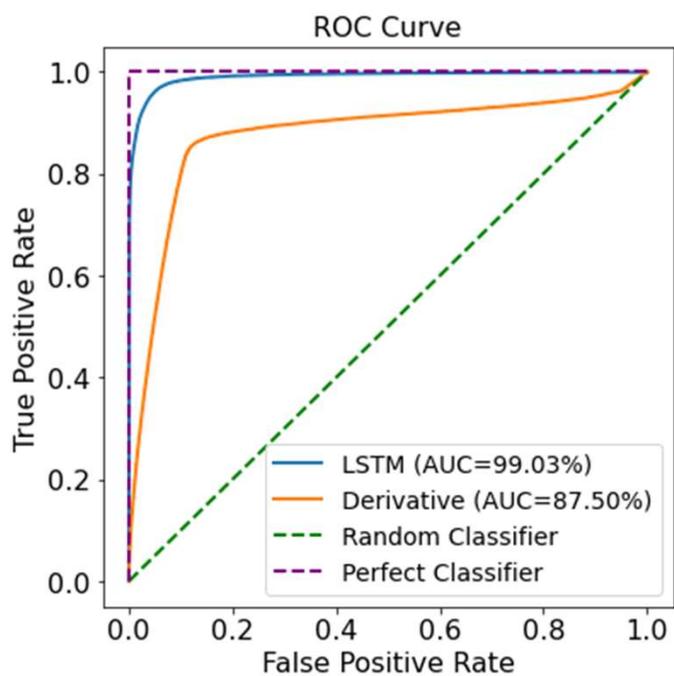
- 基于RNN的数据结构: LSTM
- 在峰的候选者中信号与噪声的分类



arXiv: 1801.07829

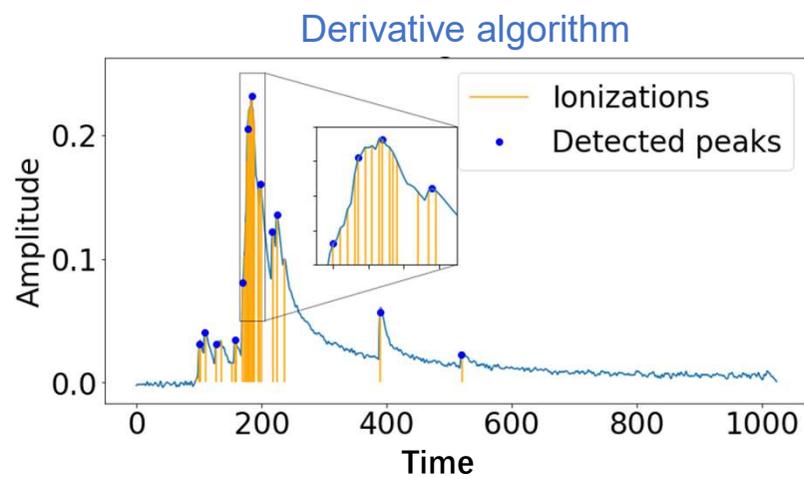
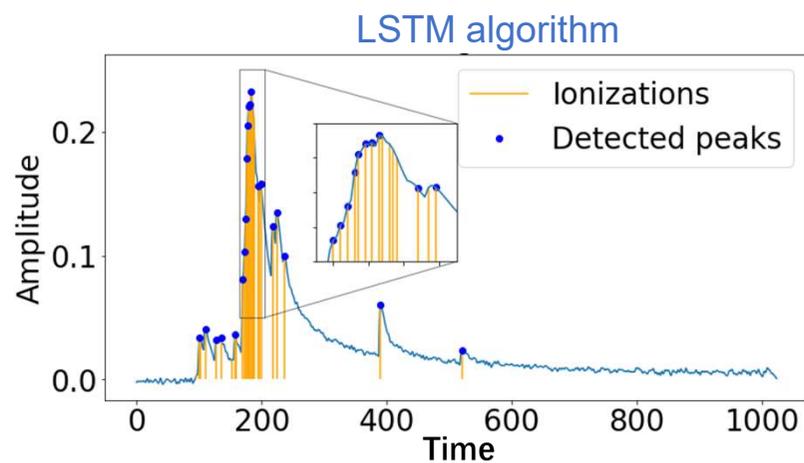
- 基于GNN的数据结构: DGCNN
- 相邻节点间传递信息 \leftrightarrow 时间信息上来自于同一个初级簇团的电子簇团化
- 初级电离电子与次级电离电子的分类

LSTM和导数寻峰算法的比较



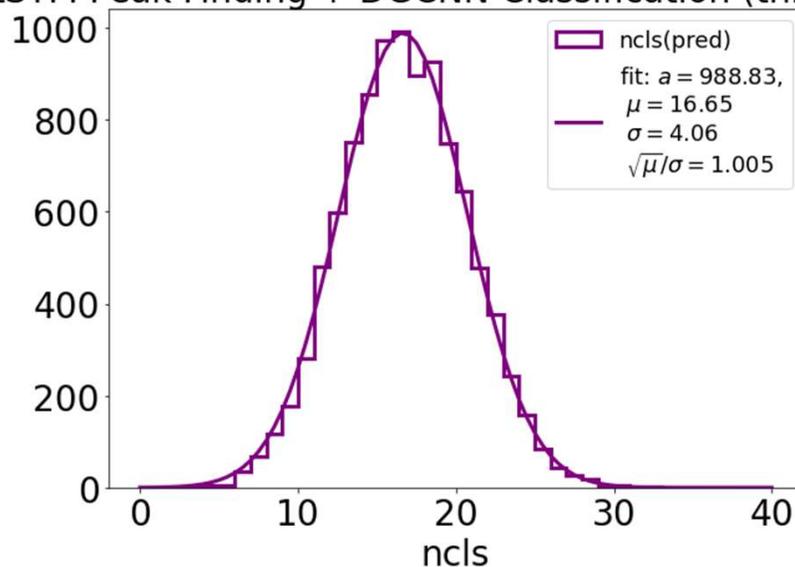
由于更好的堆叠电子的重建能力，LSTM相比导数方法有更好的AUC(area under curve, 曲线下面积)

2023/10/14



深度学习算法的表现

LSTM Peak Finding + DGCNN Classification (thr=0.61)



重建方法	μ	σ	σ/μ
MC truth(输入)	16.53	3.93	23.8%
传统算法	18.67	4.60	24.6%
深度学习	16.65	4.06	24.4%

电离团数分布的中心值和分辨更加MC truth输入一致

上升沿时间计算方法

- 减去基线后的ADC绝对值大于0.02
 - 采用局部最高点减去最近的拐点
 - 局部最高点根据一阶导数由负到正的零点决定
 - 拐点由二阶导数由正到负的零点决定
-
- 一阶导数计算方法 $D1[i] = wf[i+1] - wf[i]$
 - 二阶导数计算方法 $D2[i] = D1[i+1] - D1[i]$

