



# CEPC漂移室基于电离计数方法的粒子鉴别研究

田喆飞<sup>1</sup>, 赵光<sup>2</sup>, 董明义<sup>2</sup>, 高旭<sup>3</sup>, 李刚<sup>2</sup>, 刘帅毅<sup>2</sup>,  
宋维民<sup>3</sup>, 孙胜森<sup>2</sup>, 伍灵慧<sup>2</sup>, 辛水艇<sup>2</sup>, 张振宇<sup>1</sup>, 周详<sup>1</sup>

1. 武汉大学
2. 中国科学院高能物理研究所
3. 吉林大学

第二十一届全国核电子学与核探测技术学术年会  
2023年8月10日

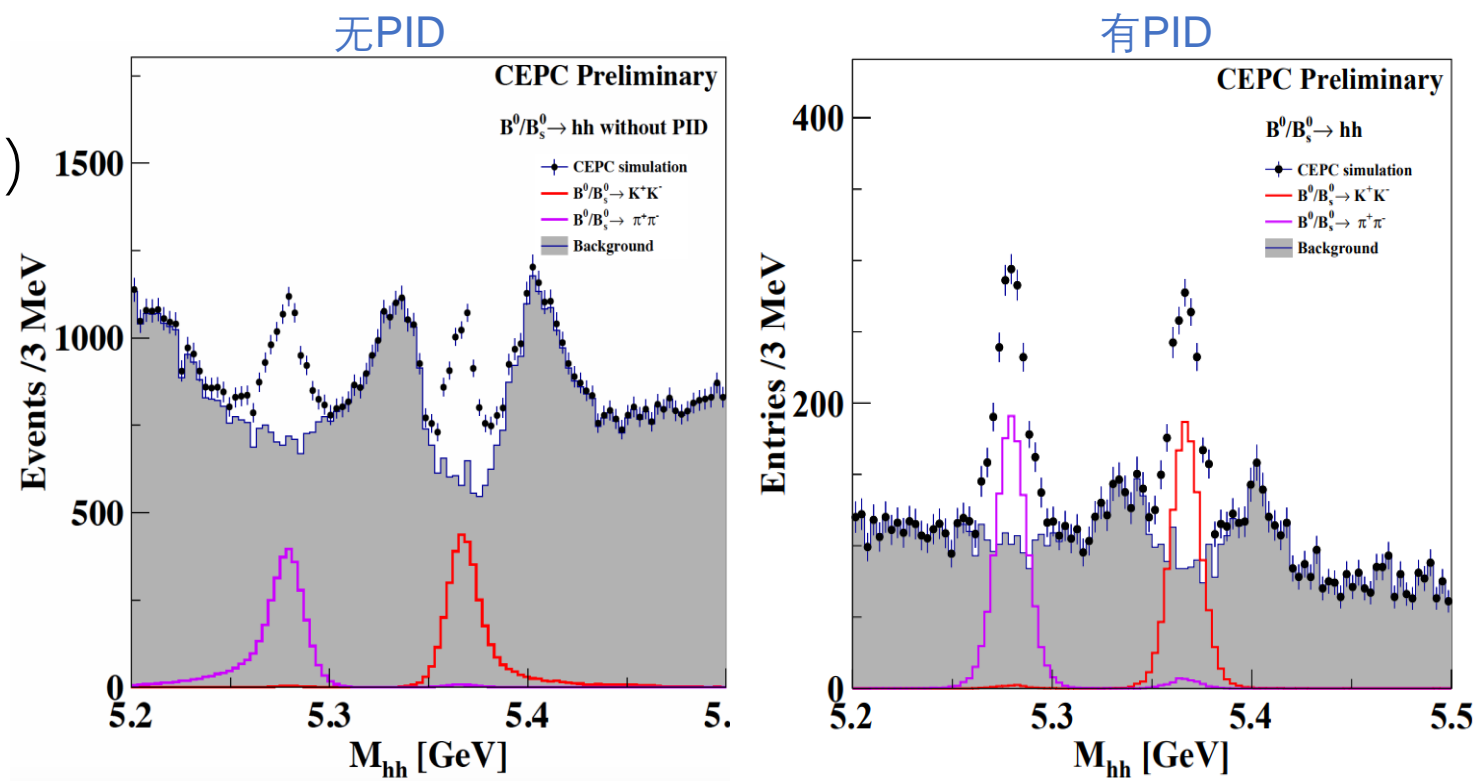
# 大纲

- 研究目的
- 模拟研究
  - 模拟
  - 重建
- 模型测试
- 总结

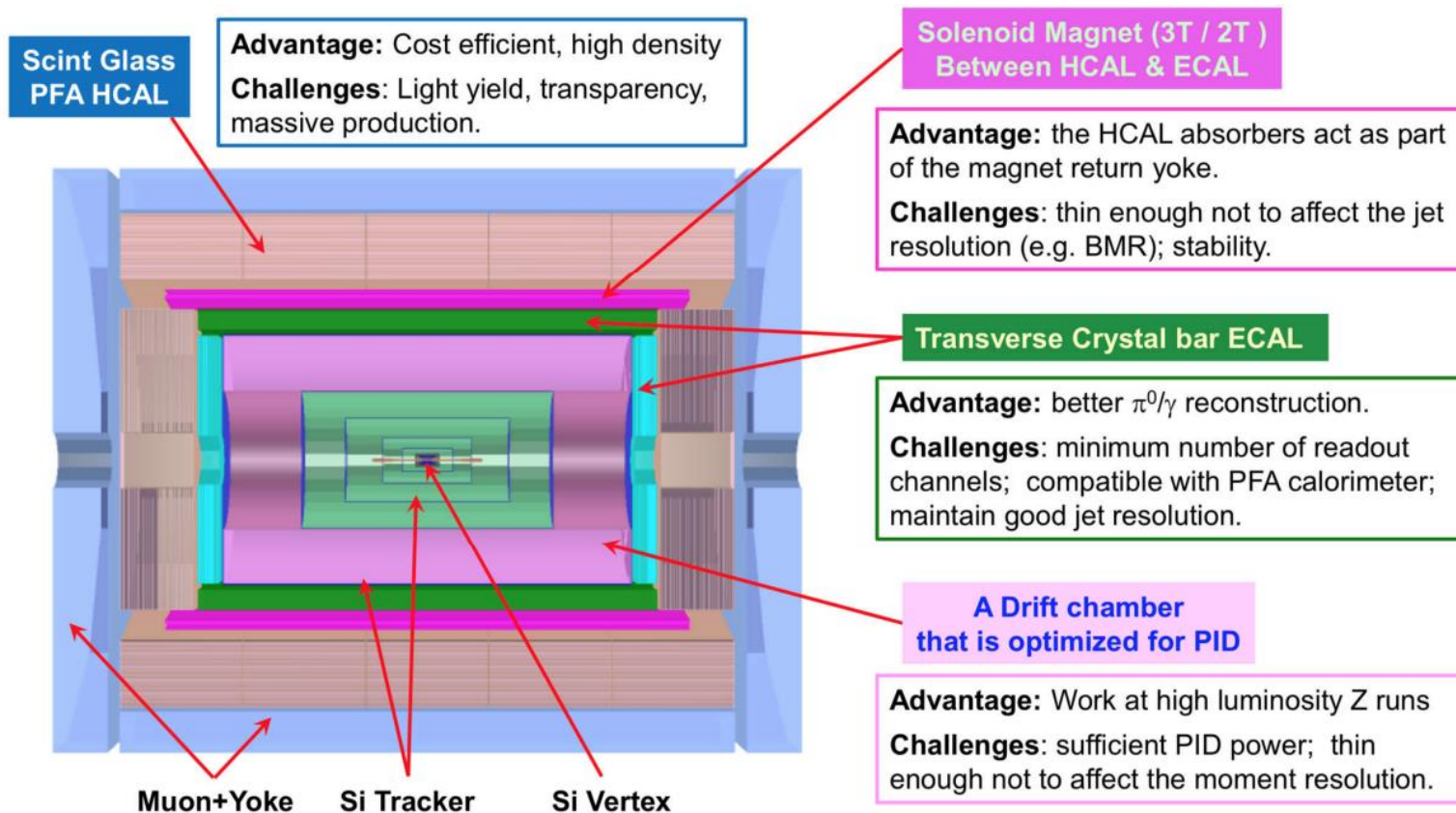
# 研究目的

- 未来轻子对撞机进行味物理和喷注研究的基础是**粒子鉴别 (PID)**
  - 减少误组合本底
  - 提高质量分辨
  - 提高喷注能量分辨
  - 有助于味标记 (flavor tagging)

例:  $B^0/B_s^0$  衰变 (模拟)



# CEPC第四个探测器概念

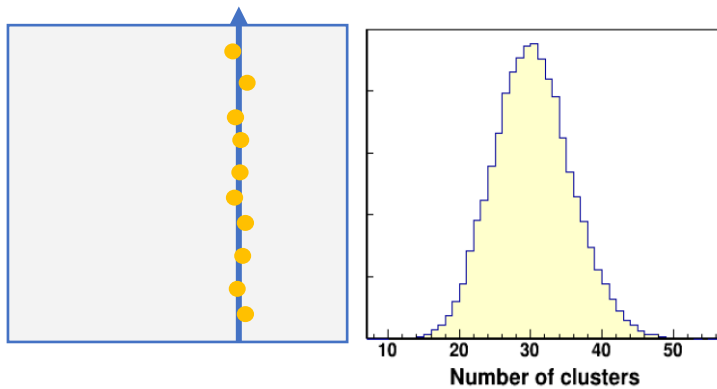


漂移室参数	
内径	800mm
外径	1800mm
单元尺寸	18 mm×18 mm
工作气体	He/iC4H10=90:10

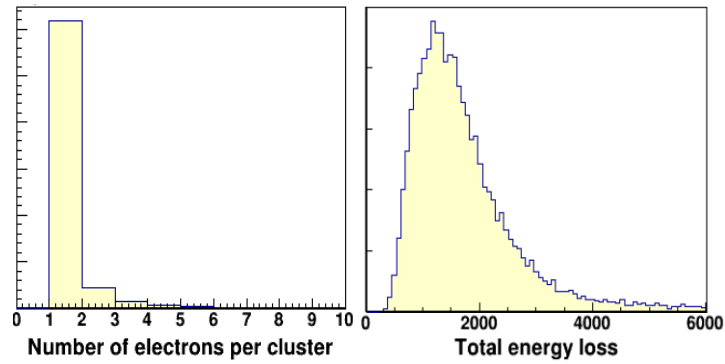
- 漂移室结合硅探测器的探测器设计方案。
- 漂移室应用原初电离计数 (cluster counting) 方法进行粒子鉴别 (PID)。
- CEPC实验的物理目标要求在动量 $<20 \text{ GeV}/c$ 处 $K/\pi$ 鉴别能力达到 $2\sigma$ 。

# $dN/dx$ vs $dE/dx$

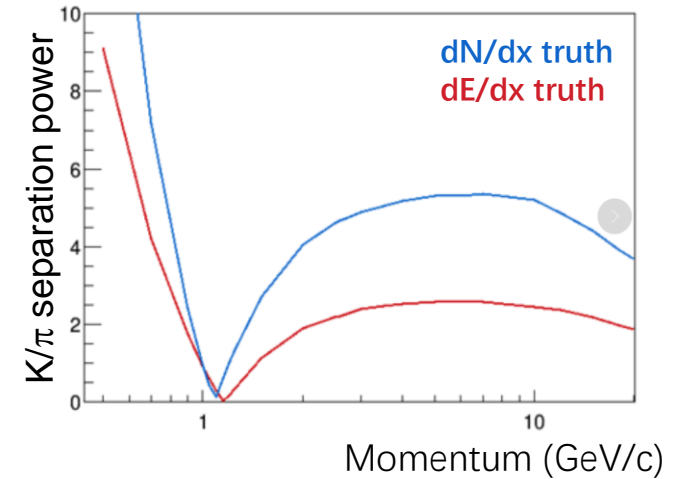
原初电离



次级电离

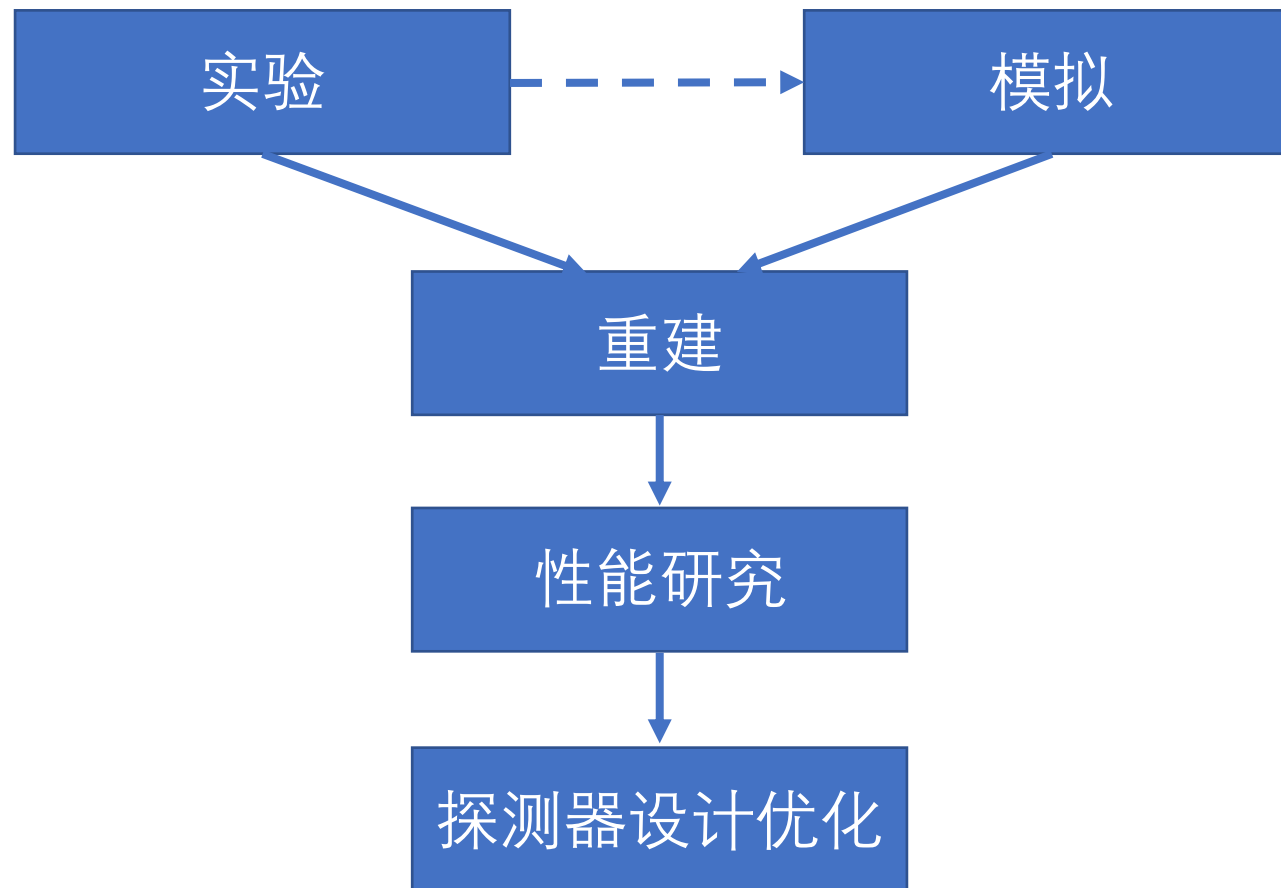


K/ $\pi$  鉴别能力  
 $dN/dx$  vs  $dE/dx$



- 与 $dE/dx$ 方法相比,  $dN/dx$ 方法拥有更好的PID性能
  - $dE/dx$ : 单位径迹长度能损, 朗道分布, 涨落大
  - $dN/dx$ : 单位径迹长度原初电离数, 泊松分布, 涨落小

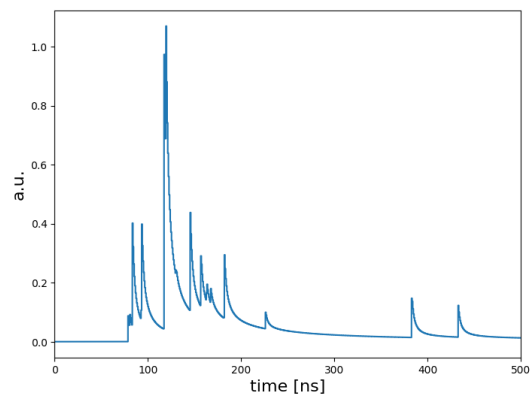
# 研究流程



# 漂移室单元的波形模拟

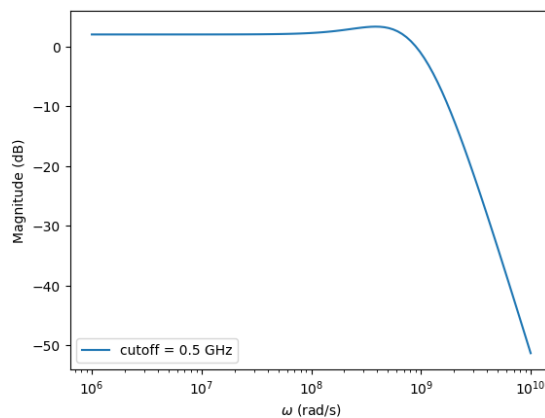
## 感应电流

基于Garfield++研究的快模拟算法



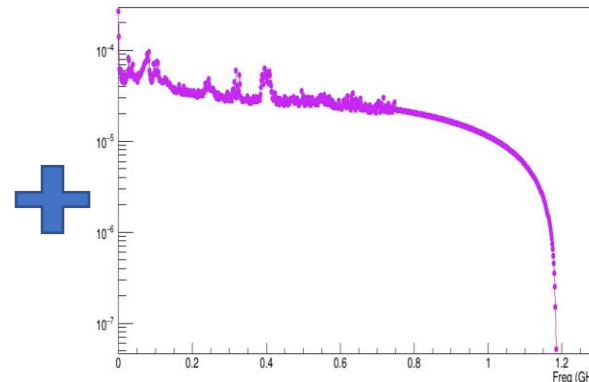
## 电子学响应

来自束流测试

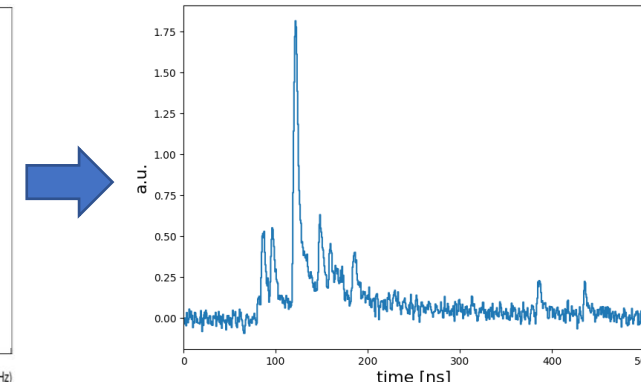


## 噪声

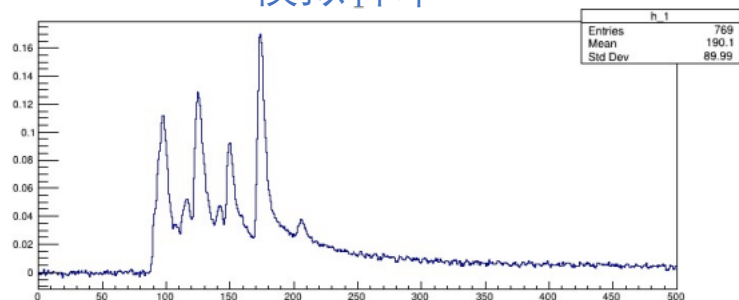
来自束流测试



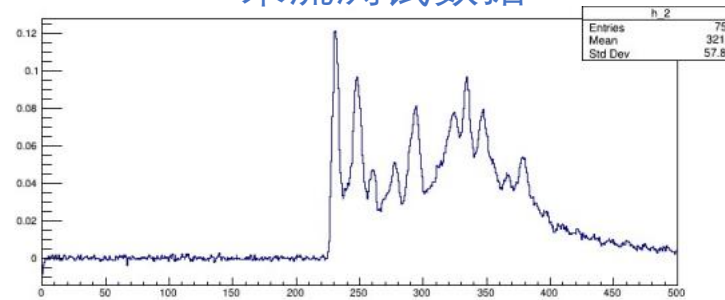
## 实际波形



## 模拟样本



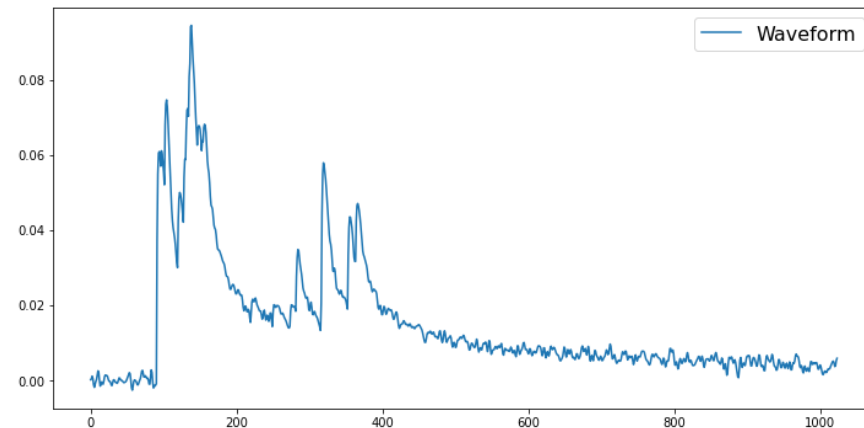
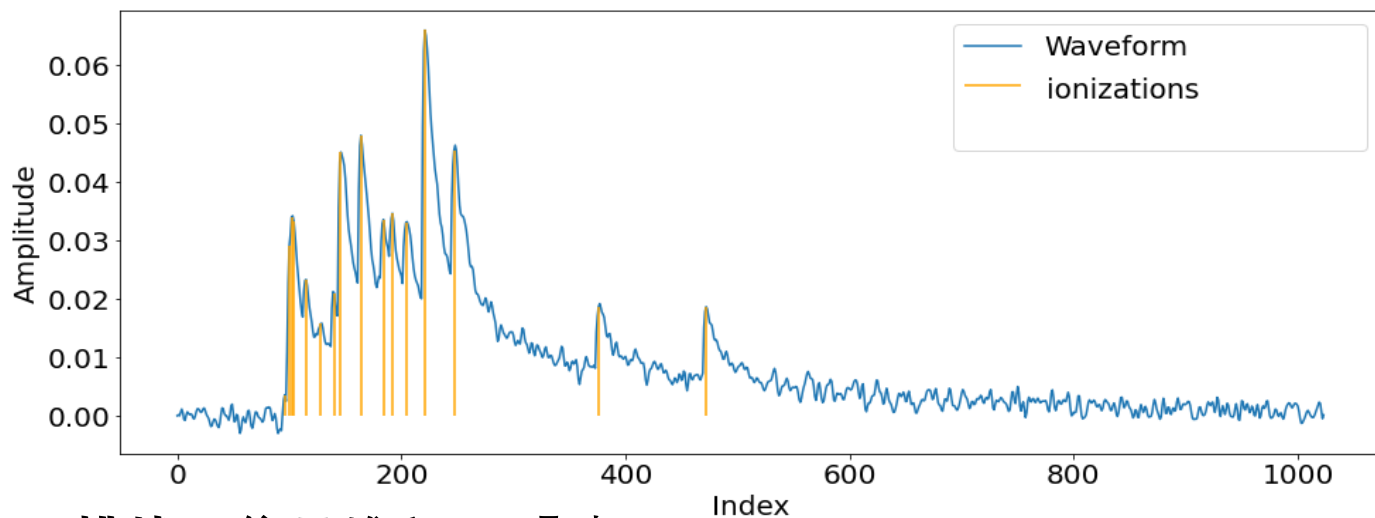
## 束流测试数据



模拟结果和实验有较好的一致性

# 重建算法

- **目的：**从漂移室单元信号丝上产生的感应电流波形中重建得到原初电离数 $N_{cls}$ 。



↓ 重建

- **挑战：**信号堆积 & 噪声

- **重建流程：**两步走

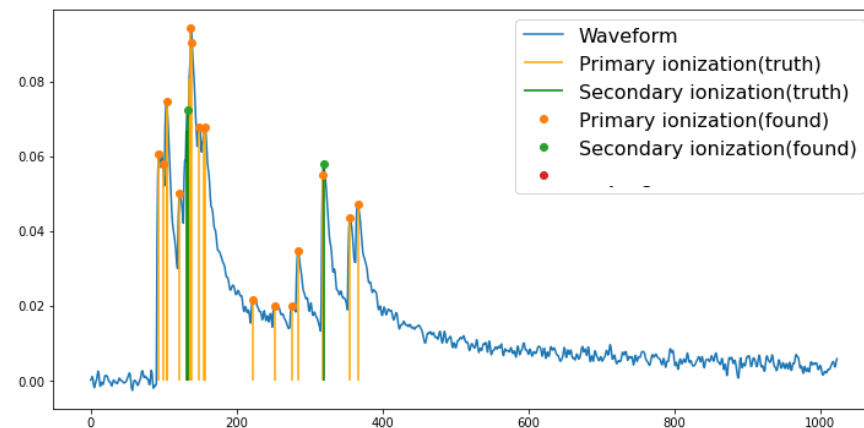
1. 寻峰 (Peak finding): 从波形中寻出电离峰 (包含原初电离和次级电离)

2. 次级电离合并 (Clusterization): 从寻出的电离峰中提取原初电离数

- **两种方法：**

- **传统方法：**基于导数的寻峰 + 基于时间差的峰合并

- **基于神经网络的算法：**LSTM寻峰算法 + DGCNN分类算法





# 传统重建算法及初步结果

## 1. 寻峰 (Peak finding)

- 利用一阶导数和二阶导数。
- 通过上升沿的斜率变化来寻峰。

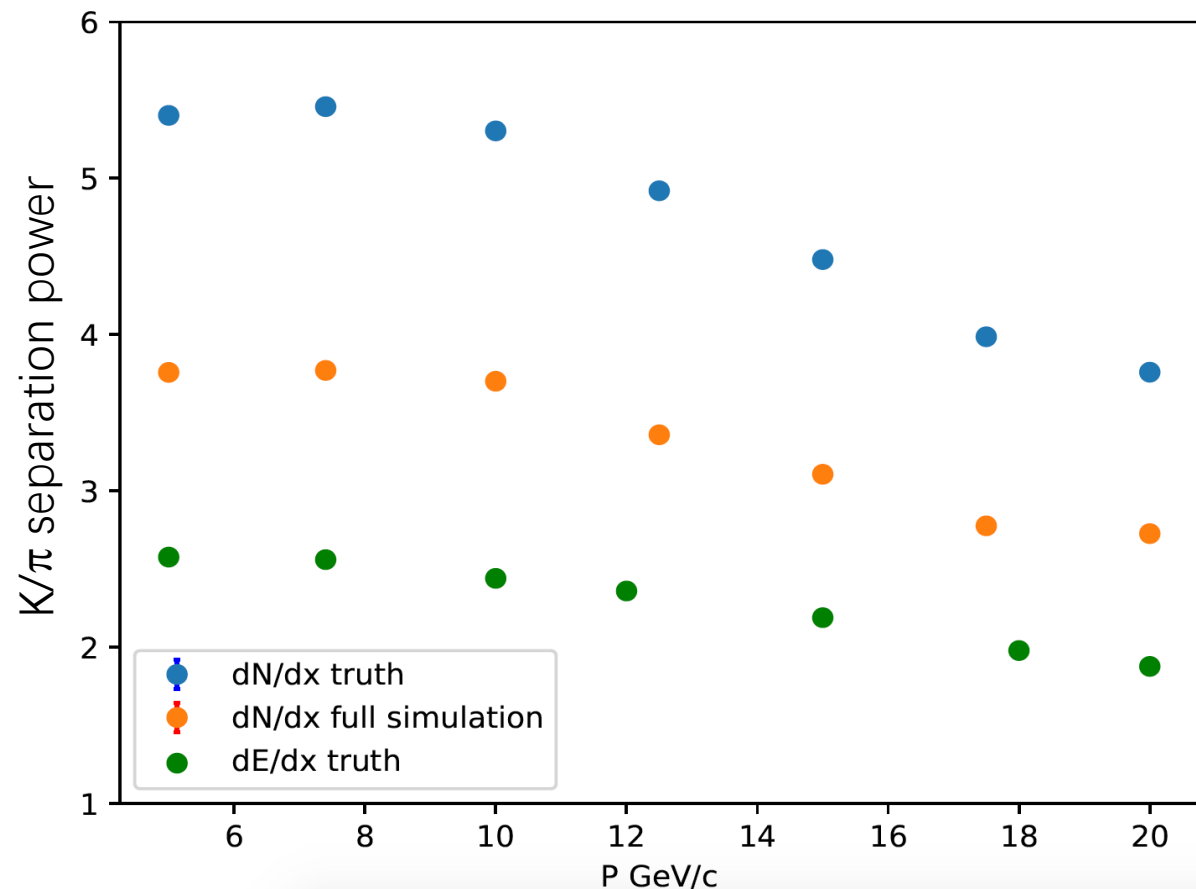
## 2. 次级电离合并 (Clusterization)

- 通过峰的时间差合并初级电离和次级电离。

### 传统算法特点

- 优点: 快速、高效
- 缺点: 难以应对噪声和重叠的信号峰

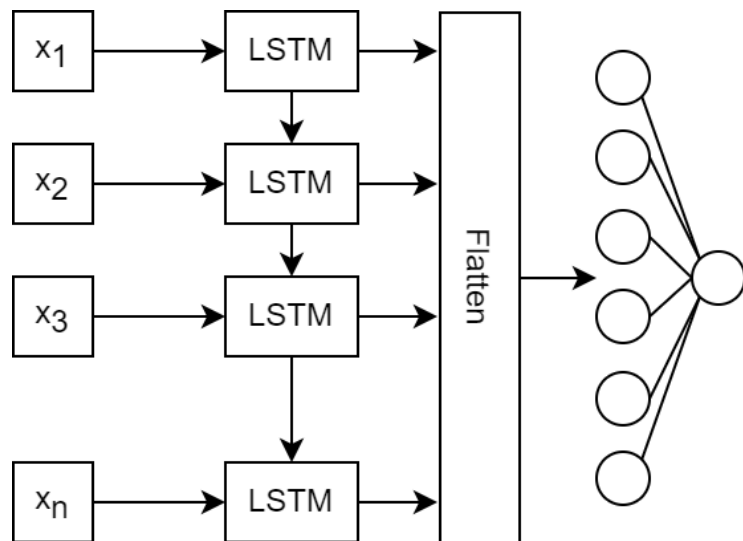
$K/\pi$  鉴别能力 (径迹长度=1m)



$$\text{鉴别能力 (Separation power): } S = \frac{\left| \left( \frac{dN}{dx} \right)_{\pi} - \left( \frac{dN}{dx} \right)_{K} \right|}{(\sigma_{\pi} + \sigma_K)/2}$$

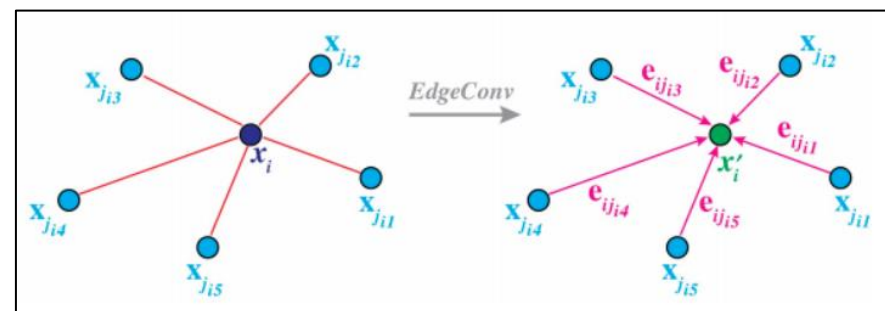
# 基于神经网络的新算法

## 1. 寻峰：长短期记忆网络 (LSTM)



- 循环神经网络 (RNN)
- LSTM适用于时序数据 $\Leftrightarrow$ 从波形中寻峰
- 判断可能为电离峰的波段是信号还是本底的二元分类

## 2. 次级电离合并：动态图卷积神经网络 (DGCNN)

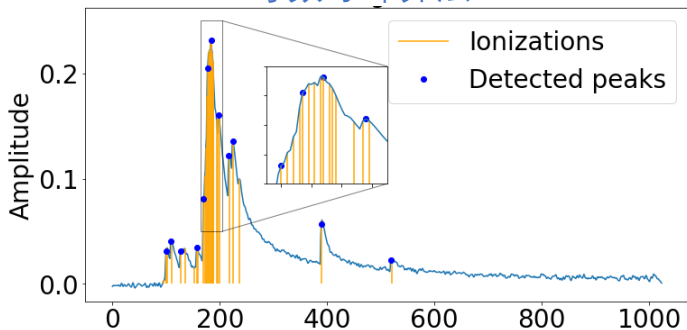


arXiv: 1801.07829

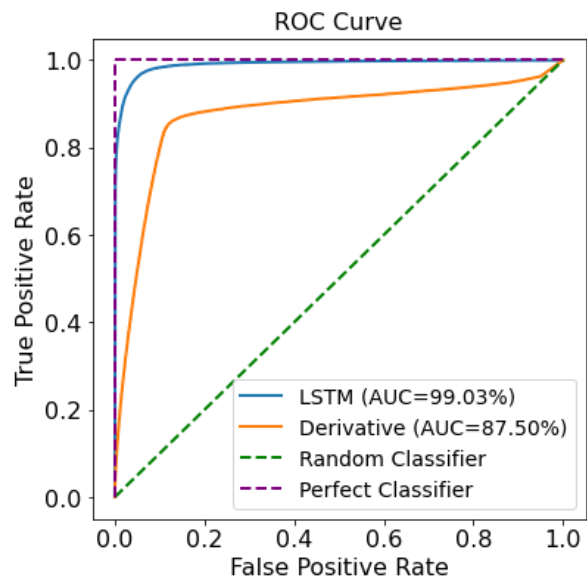
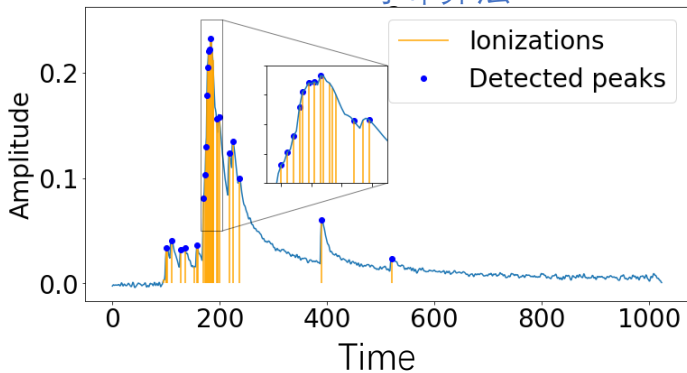
- 图神经网络 (GNN)
- DGCNN捕捉局部信息 $\Leftrightarrow$ 原初电离&次级电离的关联
- 判断寻出的电离峰为原初电离还是次级电离的二元分类。

# 基于神经网络的新算法

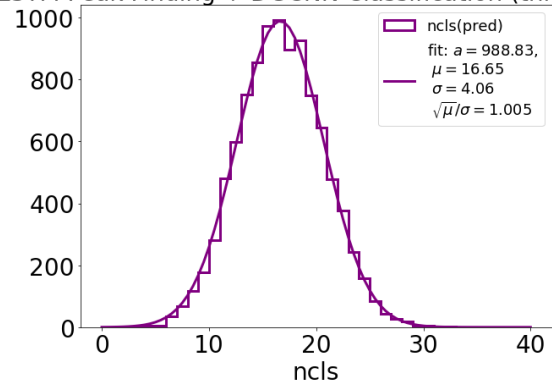
导数寻峰算法



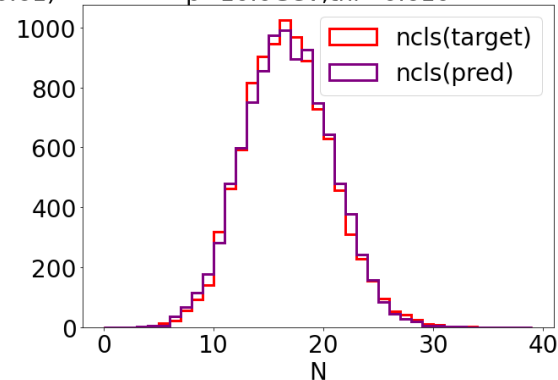
LSTM寻峰算法



LSTM Peak Finding + DGCNN Classification (thr=0.61)



p=10.0GeV, thr=0.610



**LSTM对于信号堆积问题的性能优于传统导数寻峰，作为分类器拥有更优秀的AUC。**

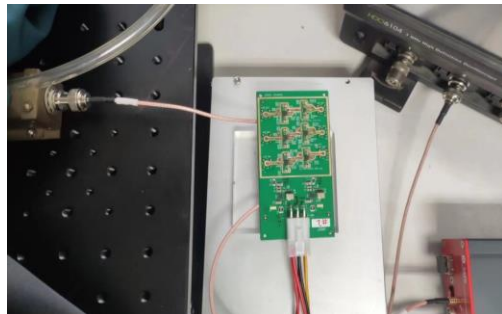
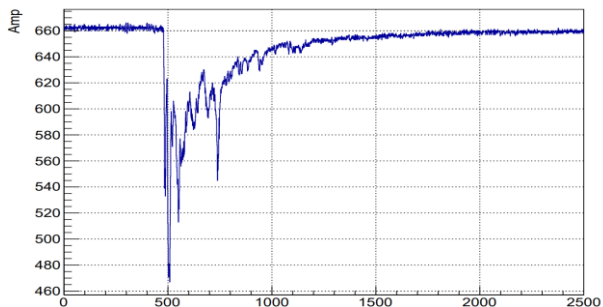
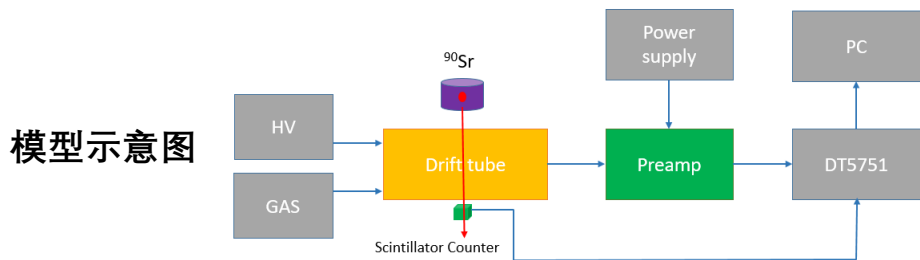
重建方法	$\mu$	$\sigma$	$\sigma/\mu$
MC truth	16.53	3.93	23.8%
传统算法	18.67	4.60	24.6%
神经网络算法	16.65	4.06	24.4%

基于神经网络的新算法得到的原初电离数 $N_{cls}$ 分布最接近Truth信息 (样本:  $p = 10 \text{ GeV } \pi$ )

# 模型测试

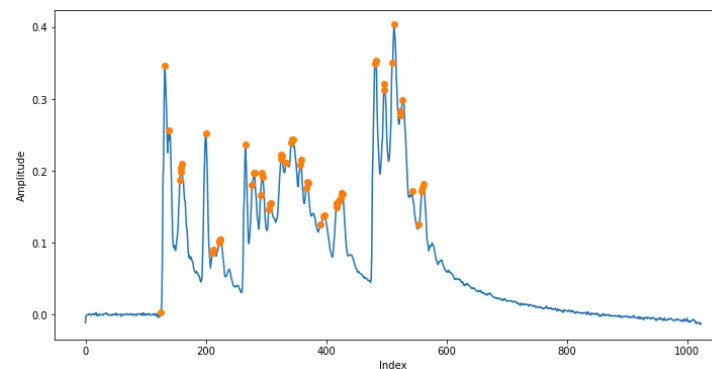
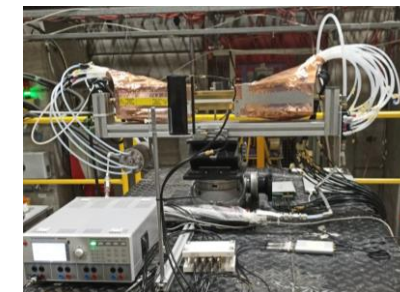
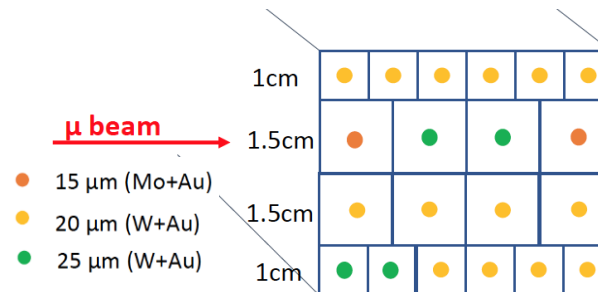
## 基于 $^{90}\text{Sr}$ 放射源的漂移管测试

- 初步结果
  - 低噪声 (5%~10%)
  - 高带宽
  - 电离峰上升时间约为几纳秒
- 进一步测试正在进行



## 束流测试

- 束流测试由INFN小组负责
- 中国-意大利合作
  - 数据获取
  - 数据分析



对束流测试数据  
应用神经网络算  
法寻峰初步结果

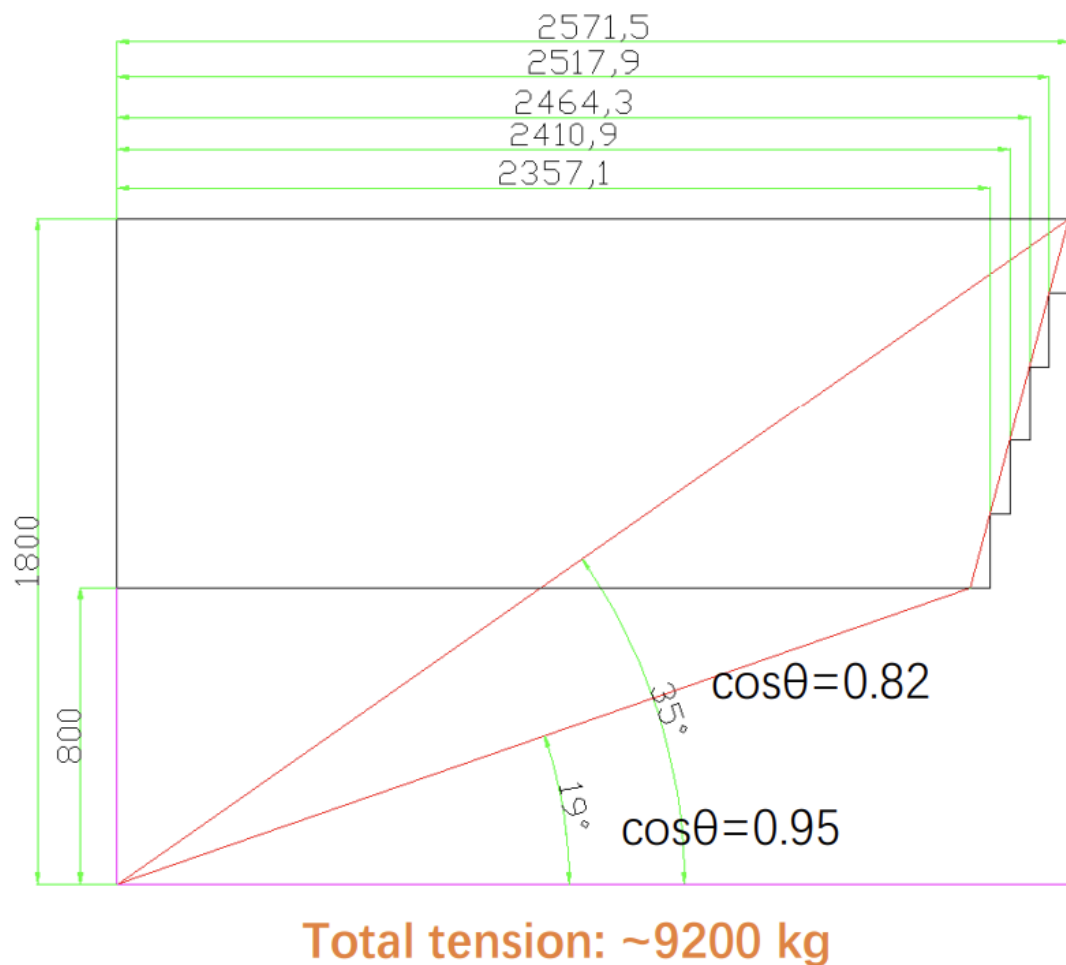
# 总结

- CEPC提出采用漂移室电离计数方法进行粒子鉴别。
- 模拟研究表明，该方法 $K/\pi$ 鉴别能力在20 GeV/c处达到 $2\sigma$ 。
- 为研究该方法可行性，正在开展相关的模型测试。
- 计划：
  - 利用实验数据优化重建算法。
  - 基于实验和模拟的结果，对CEPC漂移室设计方案进行优化。

谢谢！

**BACKUP**

# 漂移室设计参数

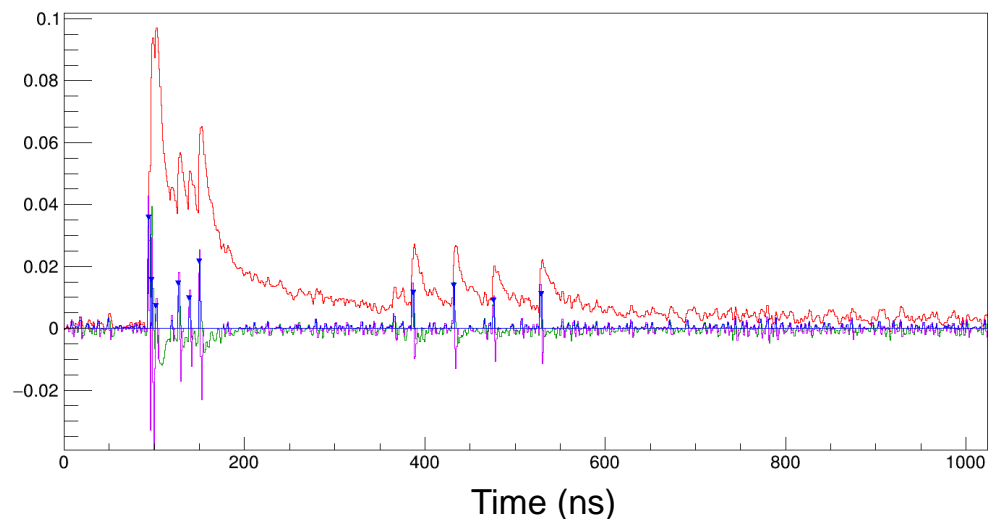


DC Parameters	
R extension	800-1800mm
Length of outermost wires ( $\cos\theta=0.82$ )	5143mm
Thickness of inner CF cylinder	200 $\mu$ m
Outer CF frame structure	Equivalent CF thickness: 1.63mm
Thickness of end Al plate	35mm
Cell size	18 mm $\times$ 18 mm
Diameter of field wire (Al coated with Au)	60 $\mu$ m
Diameter of sense wire (W coated with Au)	20 $\mu$ m
Ratio of field wires to sense wires	3:1
Gas mixture	He/iC4H10=90:10

# 传统方法：导数寻峰+时间差合并

## 寻峰 (Peak finding)

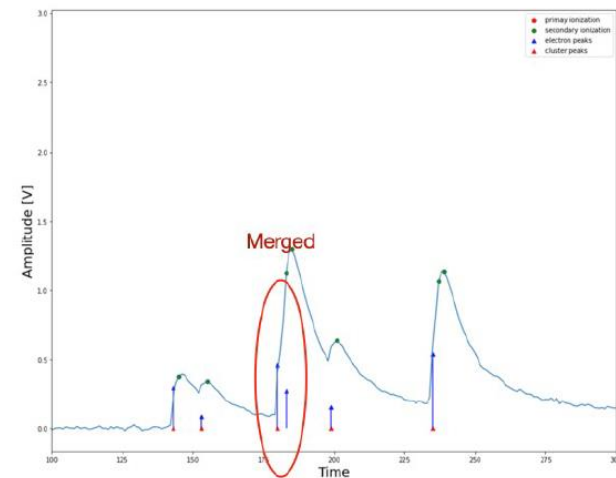
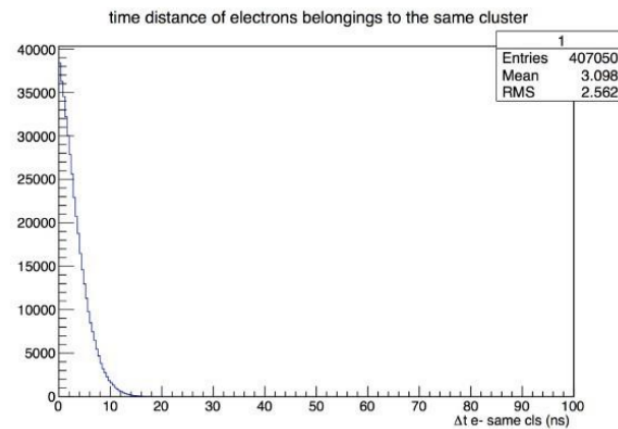
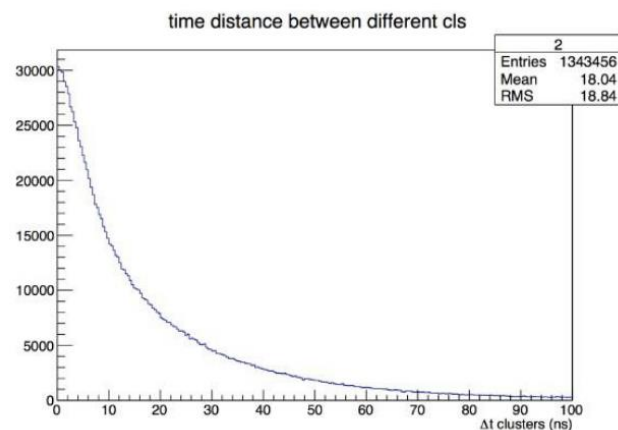
- 利用一阶导数和二阶导数
- 通过上升沿的斜率变化来寻峰



- **优点:** 快速、高效
- **缺点:** 难以应对噪声和重叠的信号峰

## 次级电离合并 (Clusterization)

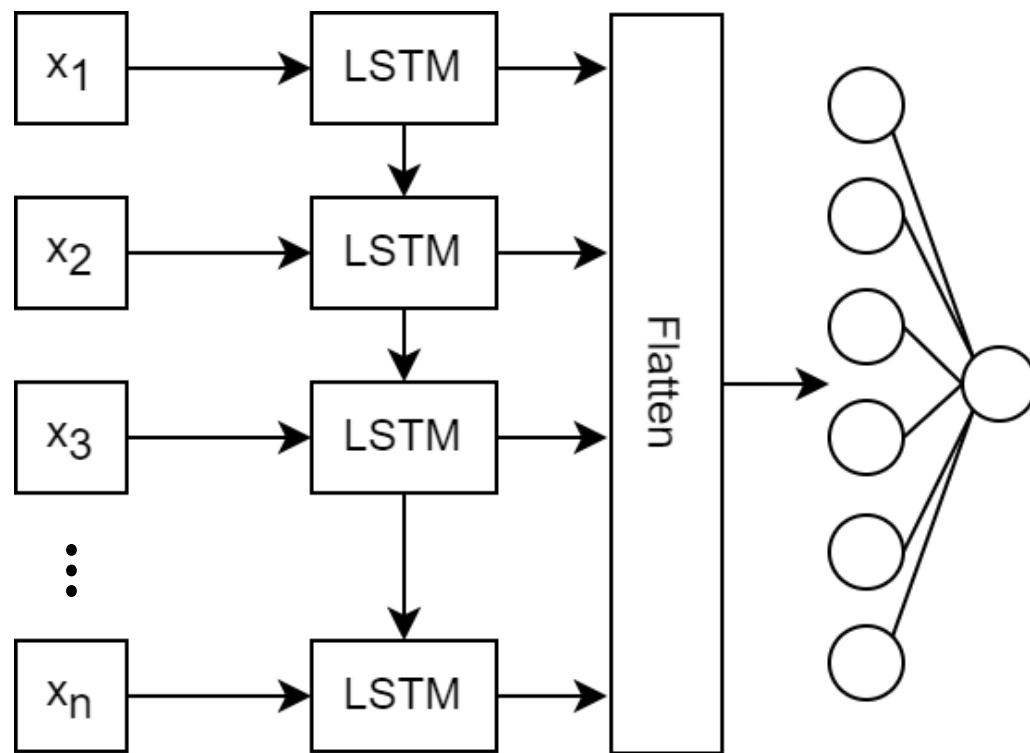
- 通过 $\Delta t$ 合并电离峰
- 通过MC样本估计合并的cut条件





# LSTM寻峰

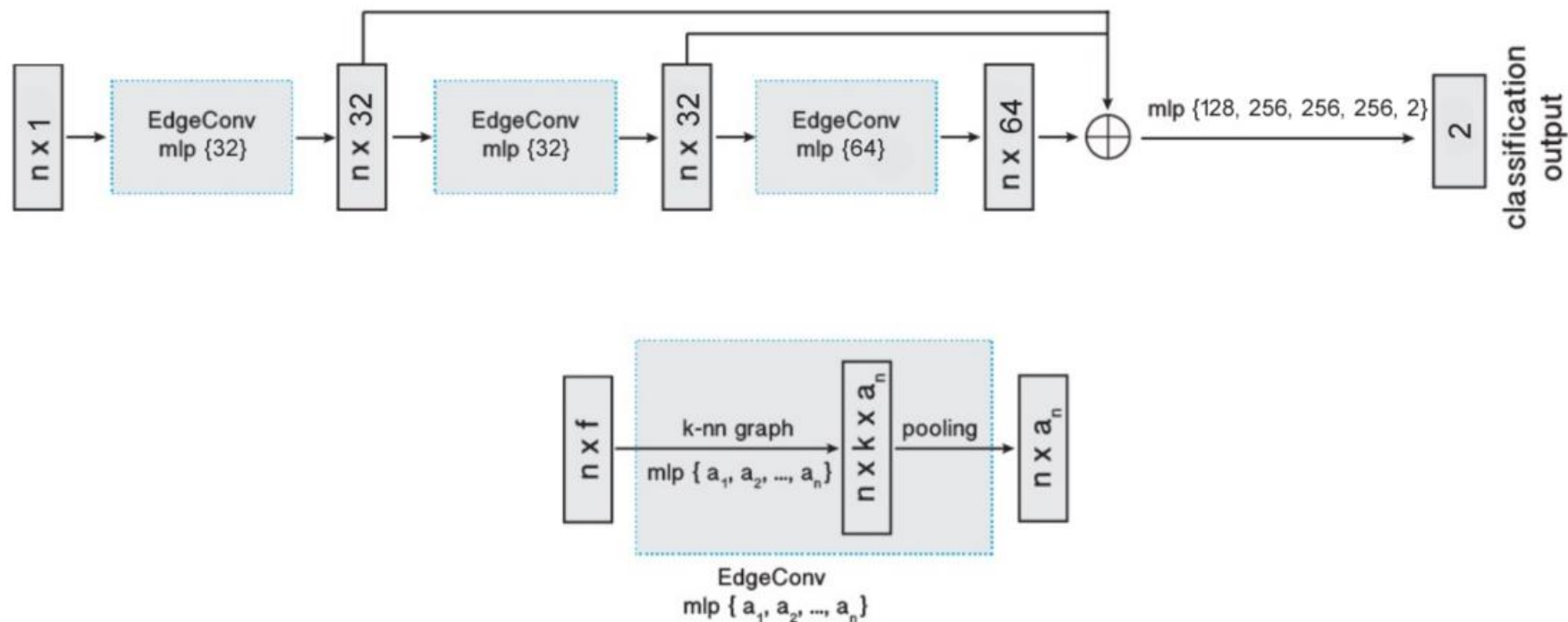
- 网络结构: 长短期记忆 (LSTM) 模型



- 数据集: 波形中的候选信号峰 (斜率降低点)
- 标签: 信号 / 本底
- 特征: 信号峰候选附近的15个点组成的小波段的时间 (x) 和振幅 (y)
- 损失函数: BCE loss

⇒ 二元分类问题

# DGCNN合并



- 图 (Graph): 每个波形对应一张图
- 节点 (Node): 寻峰算法得到的电离峰
- 节点特征 (Node feature): 电离峰的位置 (时间)
- 边 (Edge): EdgeConv层动态计算
- 标签 (Label): 电离峰的种类 (原初电离为1, 非原初电离为0)
- 损失 (Loss): BCE loss

⇒节点分类 (Node classification) 问题

