



人工智能在粒子物理实验 前端电子学中的应用探讨

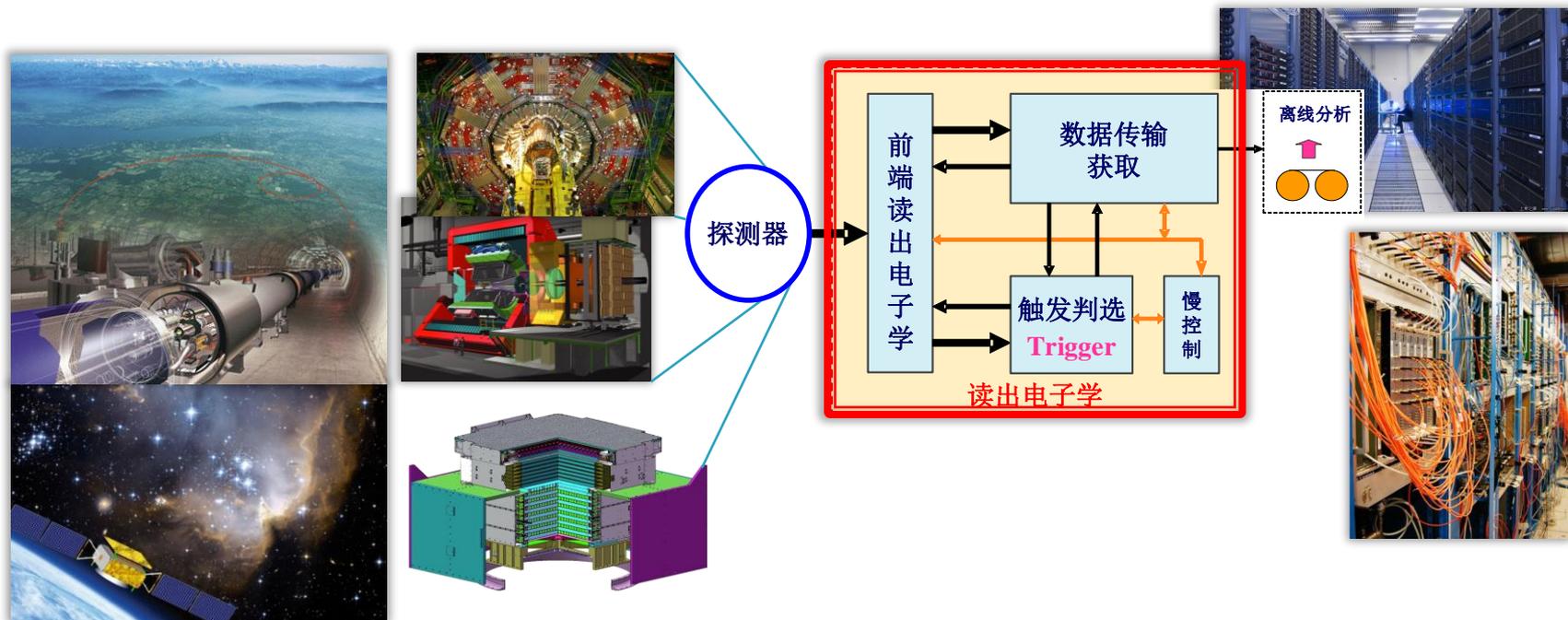
赵雷

中国科学技术大学 核探测与核电子学国家重点实验室

2025年1月12日

粒子物理实验读出电子学

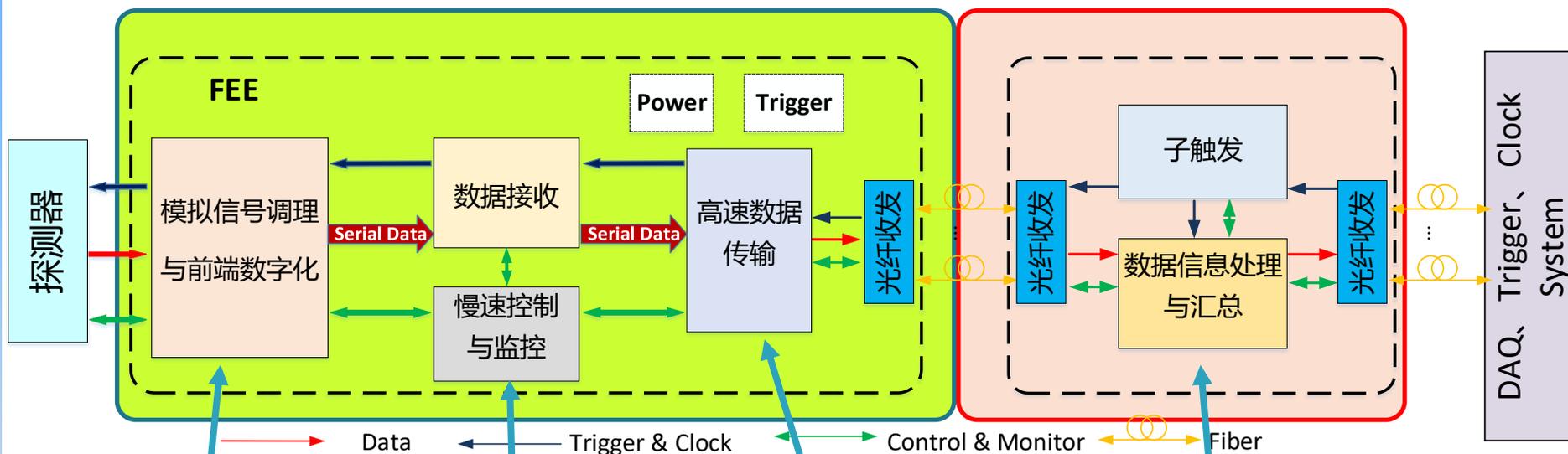
- 在大型核与粒子物理实验中，电子学是其中重要的组成部分和硬件载体。



- 探测器所感知的实验信息，须经电子学提取、处理、传输，才能为后方的分析系统所用。

粒子物理实验读出电子学

前端电子学



前端读出ASIC
电荷、时间
↓
位置、能量...

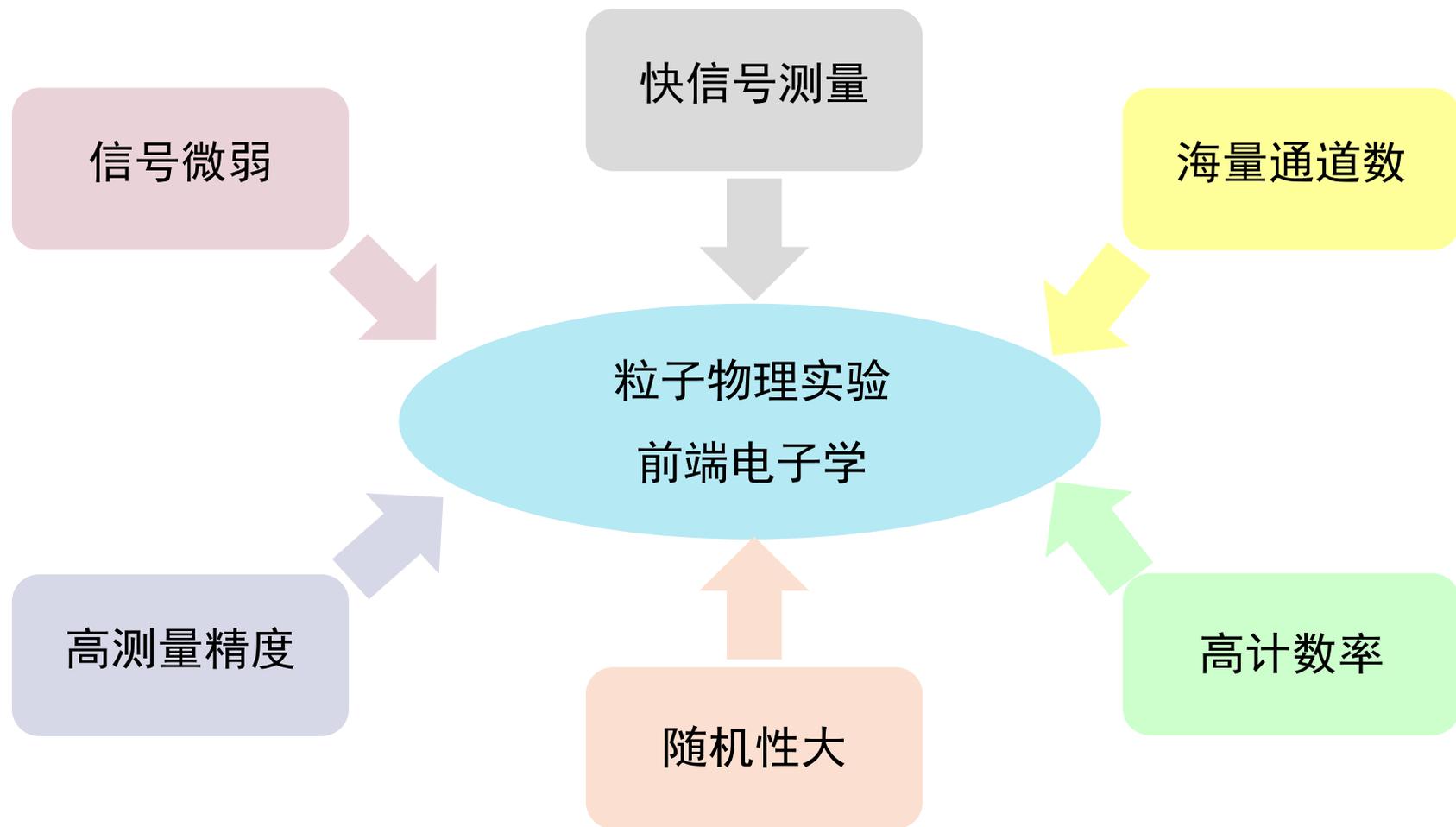
数据汇总芯片
(ROC)
慢控制芯片
(SCA)
...

高速传输芯片
(GBT)
光纤收发芯片
(VTT/VTR)
...

数字信号处理/
触发判选

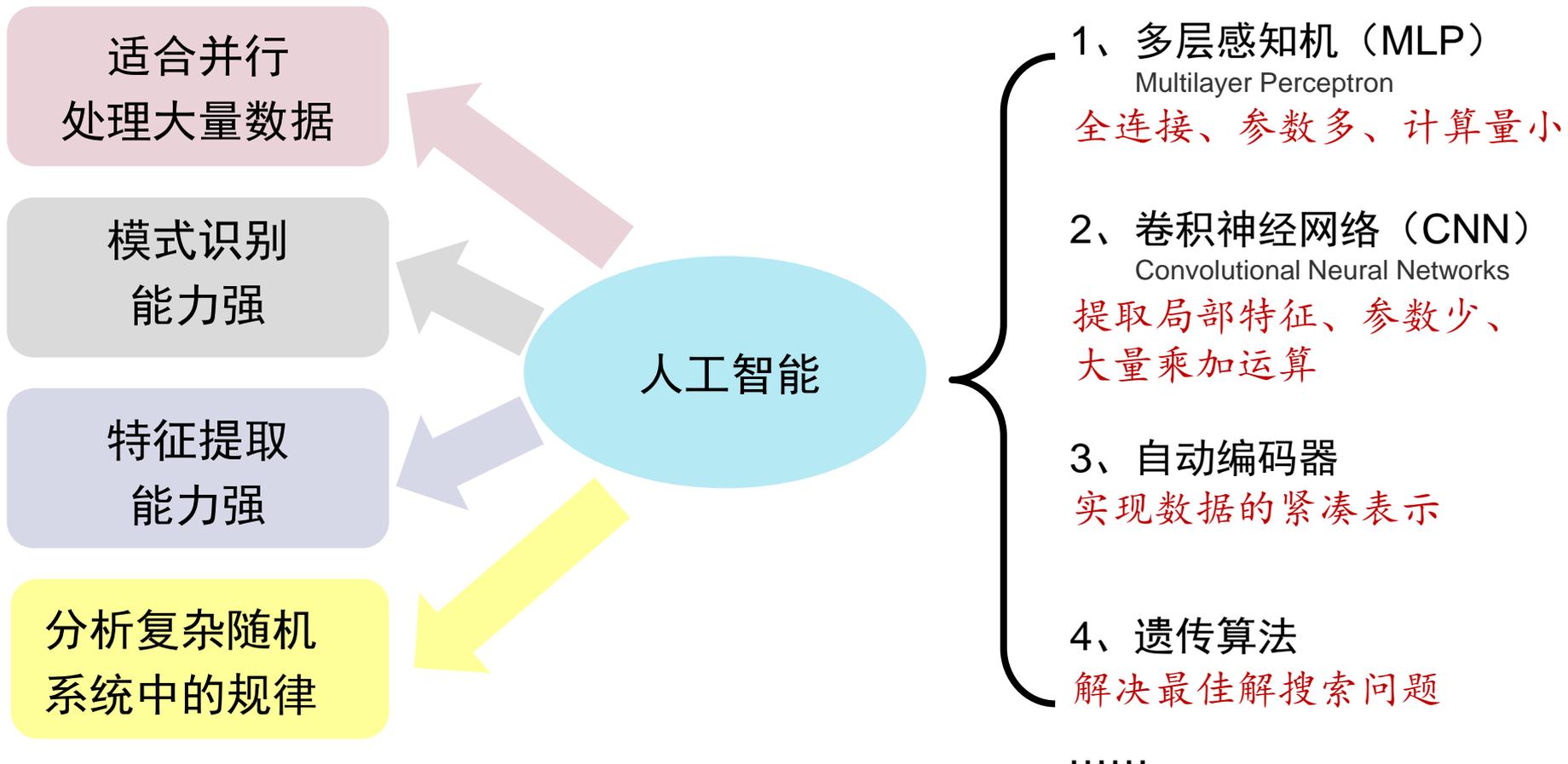
粒子物理实验前端电子学的挑战

- 随着当代核与粒子物理实验的快速发展，对信号测量技术提出了新挑战：

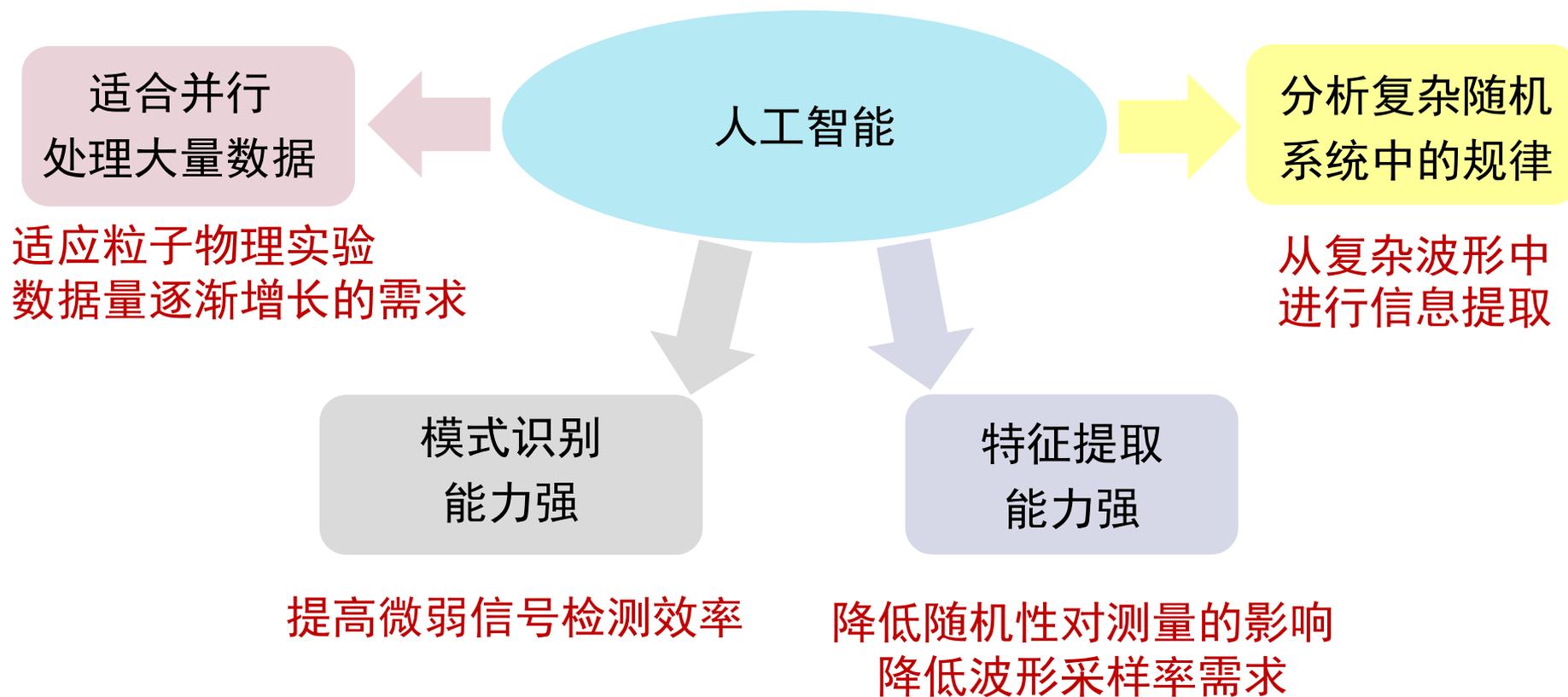


人工智能的优势

- 人工智能作为一种新兴工具，适用于海量的、复杂的信息处理场景。



人工智能的优势



- 人工智能有助于解决和应对粒子物理实验前端电子学发展所面临的挑战

研究方向分类

特征识别与提取

- 提高信号探测效率
- 增加初始有效信息比例、提高信息提取的精度

数据压缩

- 前端电子学
- 后端电子学

研究方向分类

特征识别与提取

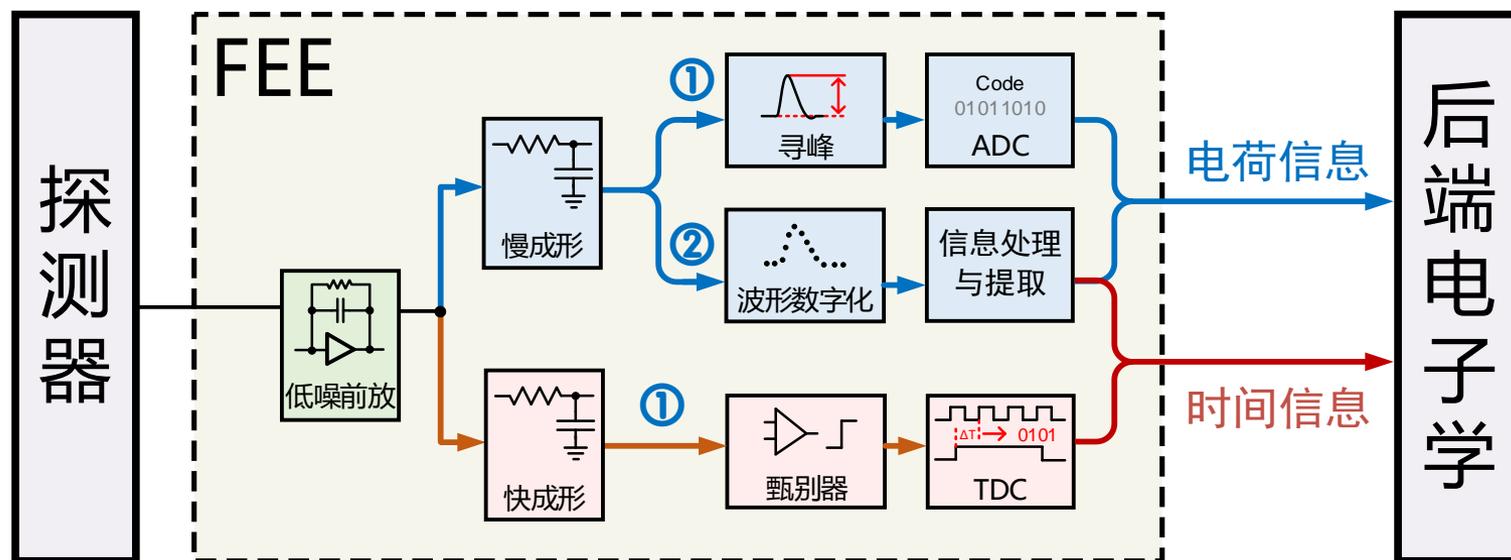
- 提高信号探测效率
- 增加初始有效信息比例、提高信息提取的精度

数据压缩

- 前端电子学
- 后端电子学

前端读出电子学

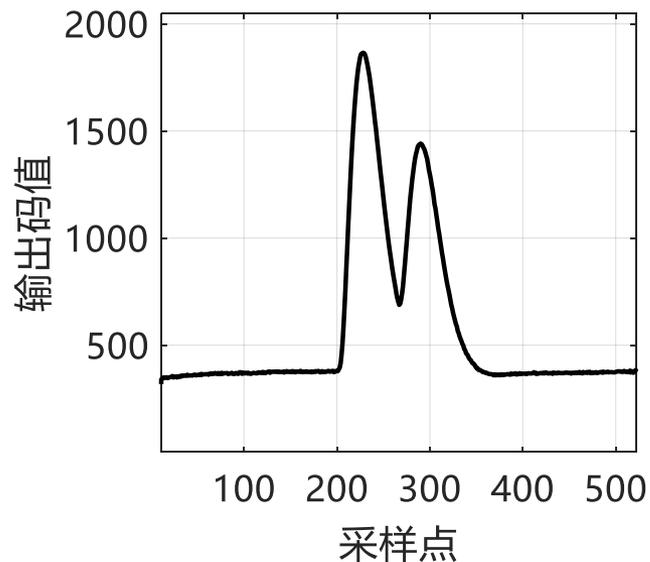
- 以高能物理实验中电荷和时间测量为例，传统的前端电子学的信息提取方式通常分为片内寻峰获取电荷信息+甄别获取时间信息与波形数字化两种技术路线：



粒子物理实验中常用的波形信息处理信号流

波形数字化的优势

- 相比寻峰获取电荷信息+甄别获取时间信息的技术路线，采集全波形信息可以获取最全面的波形信息并对非理想因素进行修正。



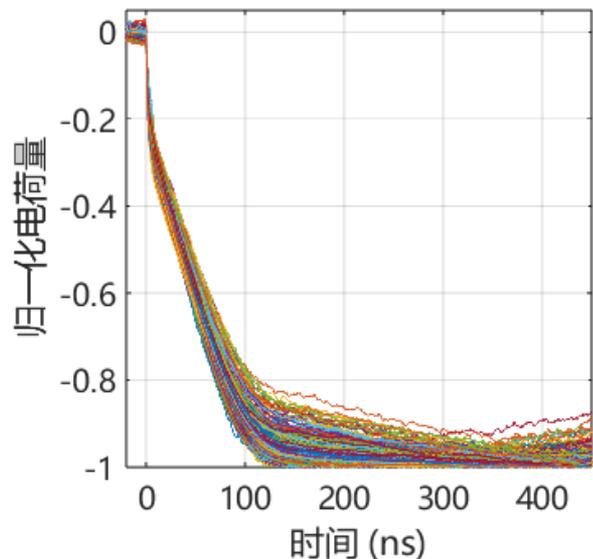
相邻事例发生堆积

波形前沿/全波形分析:

➡ 获得时间信息

通过数字信号处理获得其他信息:

➡ 如面积积分获得电荷信息



探测器电荷收集时间变化

特征识别与提取

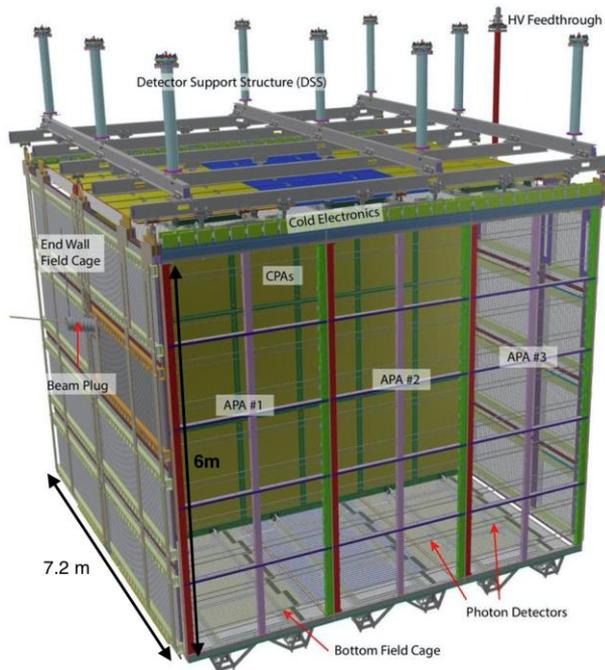
- **提高信号探测效率**
 1. LArTPC探测器低动量信号测量（CERN）
- **提高初始有效信息比例与提取精度**
 1. 基于神经网络的波形峰值检测技术（BNL）
 2. PulseDL（华中师范大学）
 3. 基于机器学习实现不等权重的高精度重心定时（中科大）
 4. 基于机器学习的TIADC失配误差修正方法研究（中科大）
 5. 基于遗传算法的低采样率高精度定时方法研究（中科大）

特征识别与提取

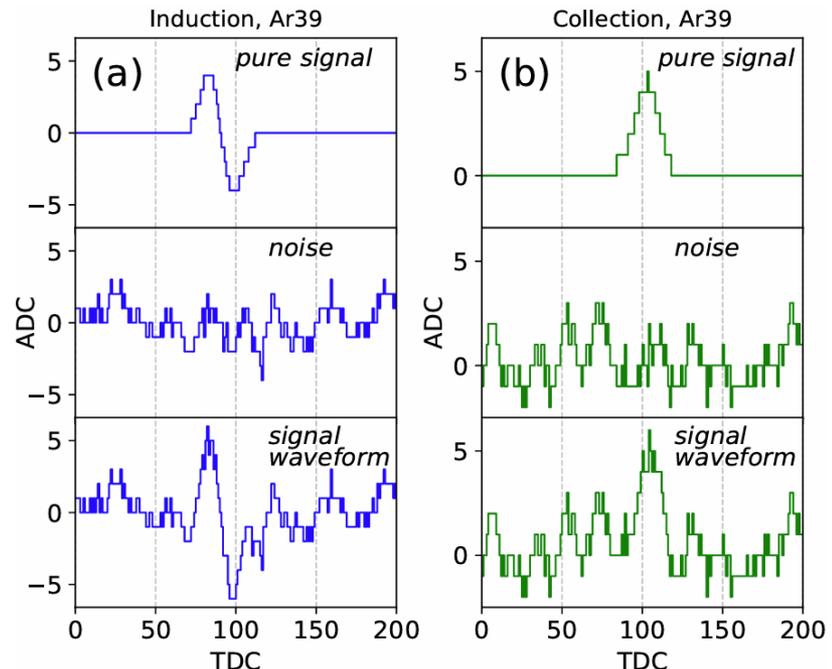
- 提高信号探测效率
 1. LArTPC探测器低动量信号测量 (CERN)
- 提高初始有效信息比例与提取精度
 1. 基于神经网络的波形峰值检测技术 (BNL)
 2. PulseDL (华中师范大学)
 3. 基于机器学习实现不等权重的高精度重心定时 (中科大)
 4. 基于机器学习的TIADC失配误差修正方法研究 (中科大)
 5. 基于遗传算法的低采样率高精度定时方法研究 (中科大)

基于模式识别提高信号探测效率

- 类型：卷积神经网络
- LArTPC (Liquid Argon Time Projection Chamber) 探测器是一种新型的中微子探测器
- 来自恒星的中微子能量低至1~10 MeV，产生接近噪声水平的信号，传统基于阈值进行判选的方法难以对其进行探测



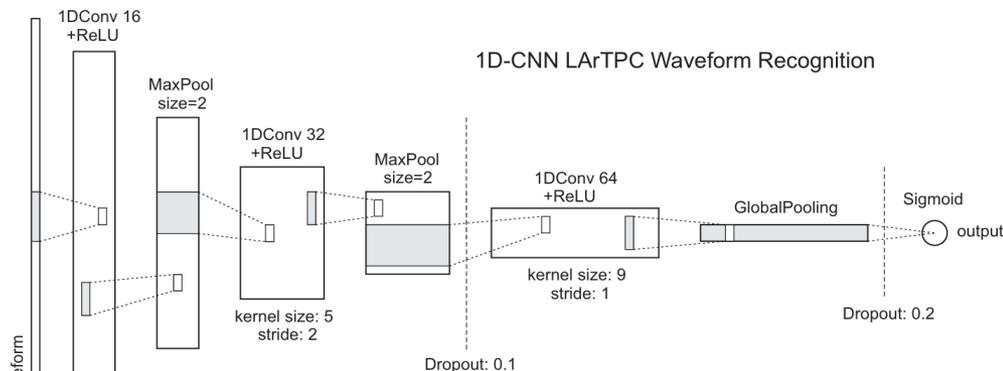
LArTPC探测器结构图



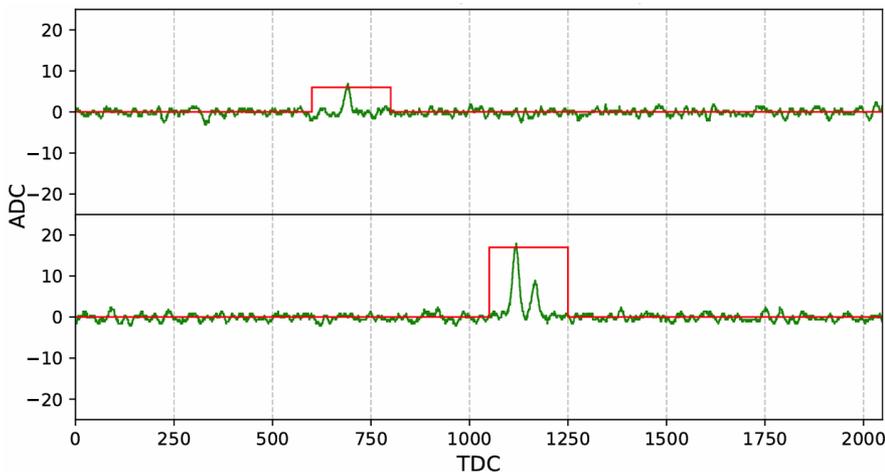
典型小信号波形

基于模式识别提高信号探测效率

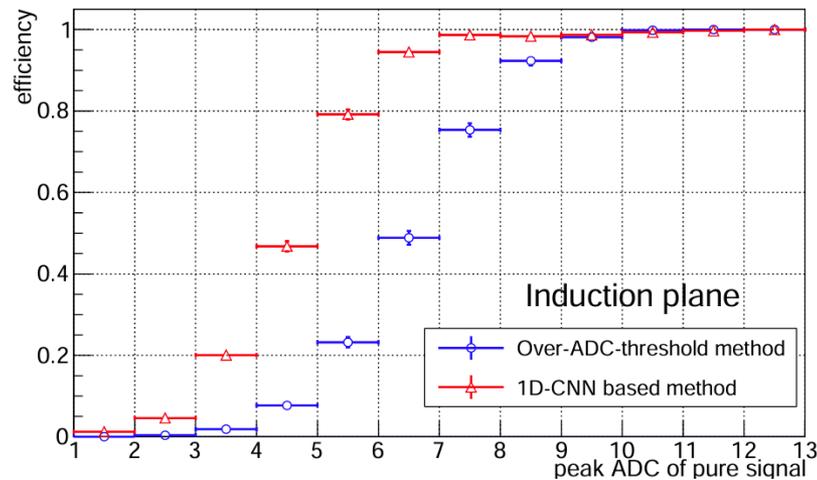
- 使用1D-CNN网络进行模式识别，相较于全连接神经网络参数数目显著减少
- 可测量的中微子能量从原来的0.9 MeV进一步下降至0.5 MeV



波形识别算法结构



信号识别示意



$$\text{效率} = \frac{\text{真阳} + \text{真阴}}{\text{真阳} + \text{假阳} + \text{真阴} + \text{假阴}}$$

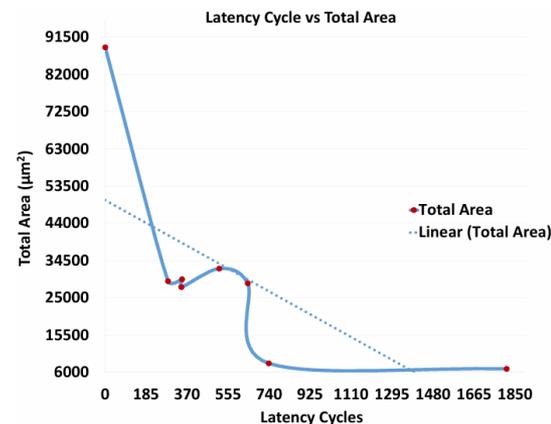
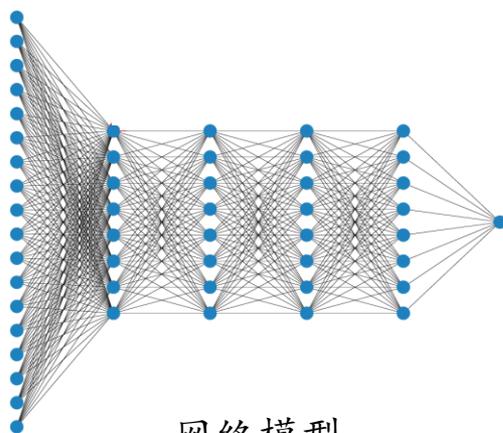
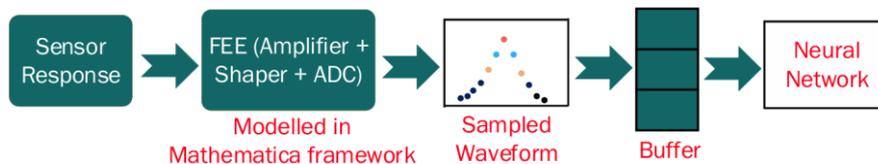
识别准确度对比

特征识别与提取

- 提高信号探测效率
 1. LArTPC探测器低动量信号测量 (CERN)
- 提高初始有效信息比例与提取精度
 1. 基于神经网络的波形峰值检测技术 (BNL)
 2. PulseDL (华中师范大学)
 3. 基于机器学习实现不等权重的高精度重心定时 (中科大)
 4. 基于机器学习的TIADC失配误差修正方法研究 (中科大)
 5. 基于遗传算法的低采样率高精度定时方法研究 (中科大)

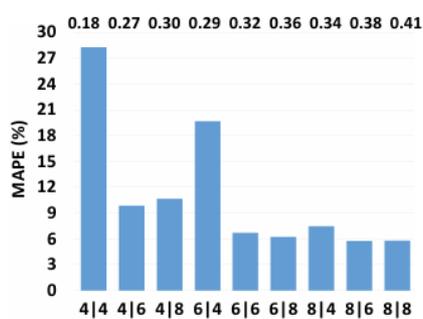
1、基于神经网络的波形峰值检测技术 (BNL)

- 类型：感知机
- BNL实验室探索了通过神经网络提取信号峰值的方法，并基于hls4ml、Catapult工具链进行了实现和仿真分析

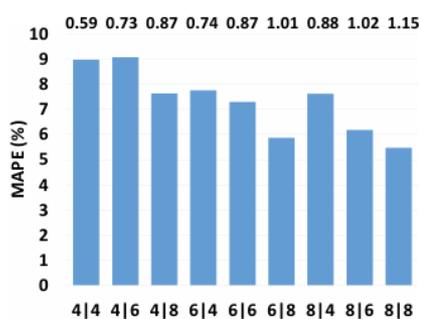


网络模型

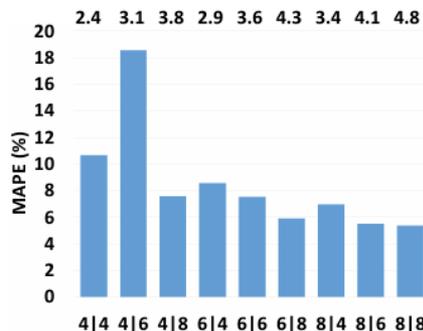
网络面积与延迟的关系



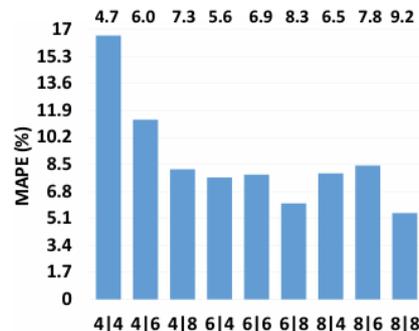
(a) Tiny (8)



(b) Small (16)



(c) Medium (36)

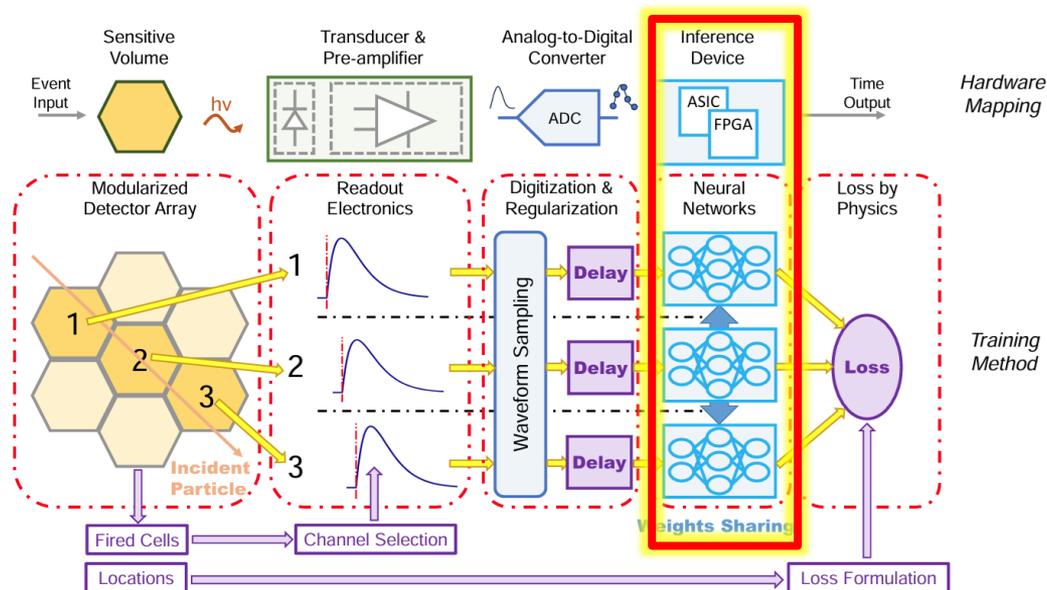


(d) Large (52)

不同网络规模下的归一化误差

2、PulseDL（华中师范大学）

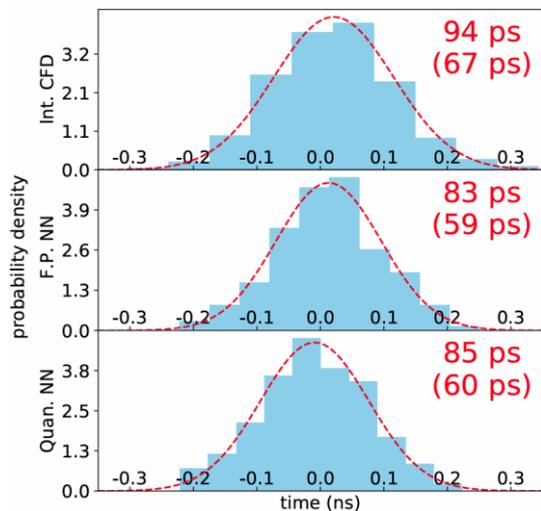
- 类型：自编码器+感知机
- NICA-MPD量能器的读出电子学数据带宽需求高（单板~48 Gbps），并带来高功耗（单板~16 W）
- 由于实际采样率偏低且波形存在长时漂移，传统方法精度受限，华中师范大学基于神经网络方法展开了时间和幅度信息提取的电路研究



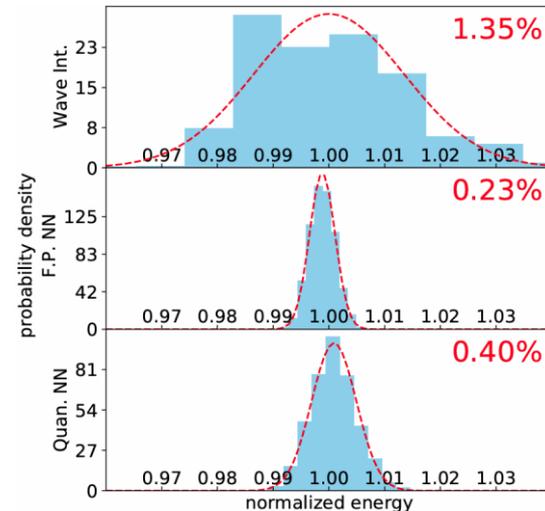
基于物理实验的仿真模型

2、PulseDL (华中师范大学)

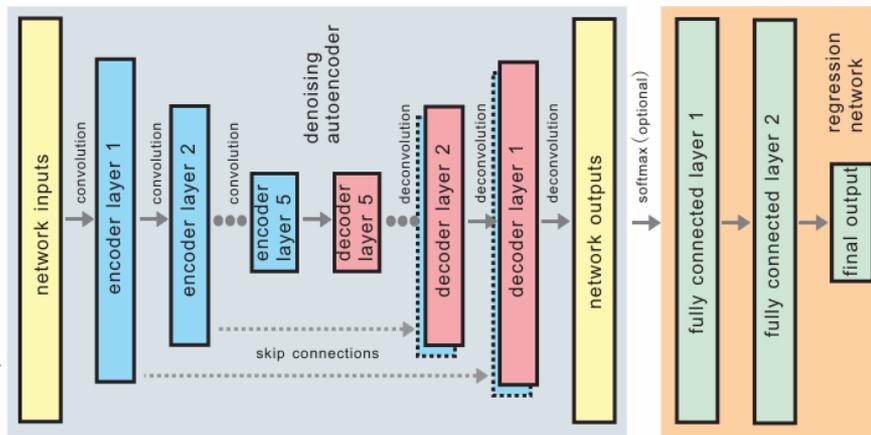
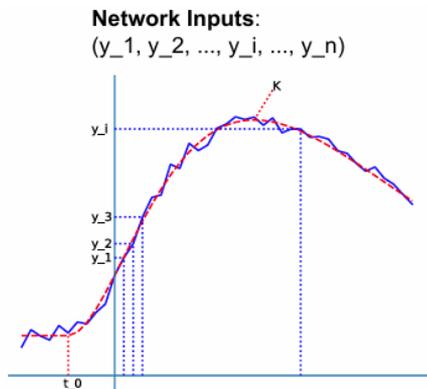
- 神经网络包括降噪自编码器与感知机特征提取两步
- 原型芯片在均衡模式下支持10通道输入、35 Hz计数率，推理总延迟~ms



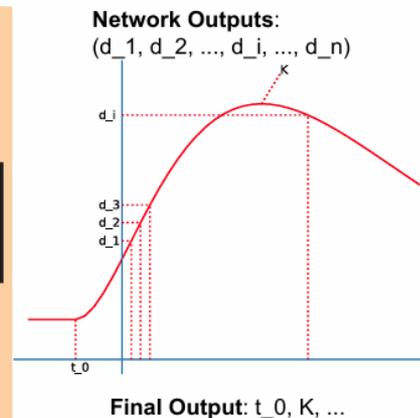
时间精度对比



能量精度对比



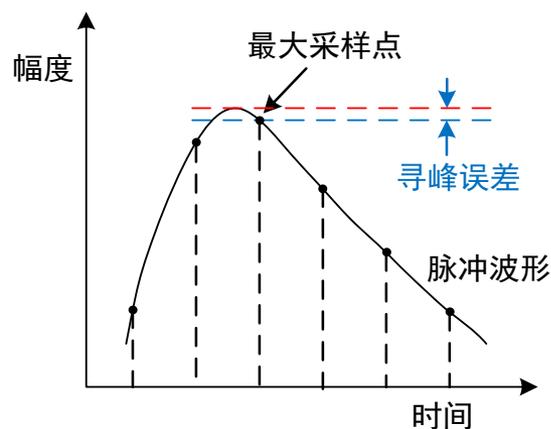
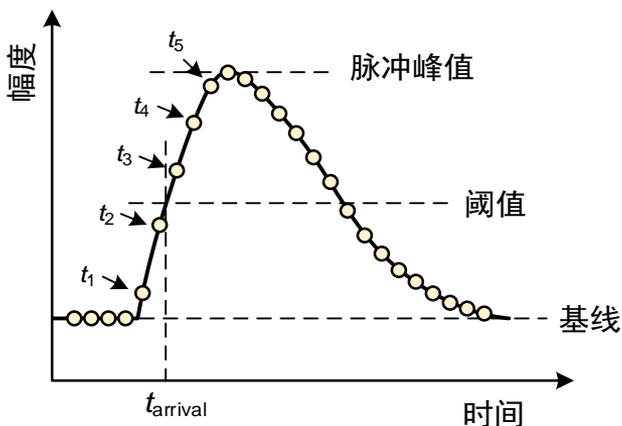
信号处理流程



Final Output: t_0, K, \dots

3、基于机器学习实现不等权重的高精度重心定时（中科大）

- 类型：感知机
- 前沿定时 ➡ 受时间游动效应影响
- 恒比定时 ➡ 受幅度值测量误差影响
- 重心定时 ➡ 等权重而受小斜率点影响，精度下降
- 部分采样点重心定时 ➡ 存在较大的截断误差

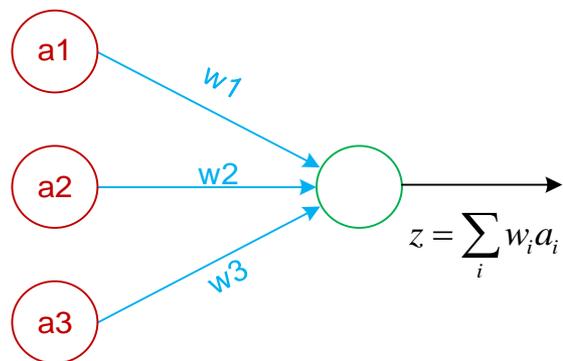


传统的定时方法存在一定的局限性

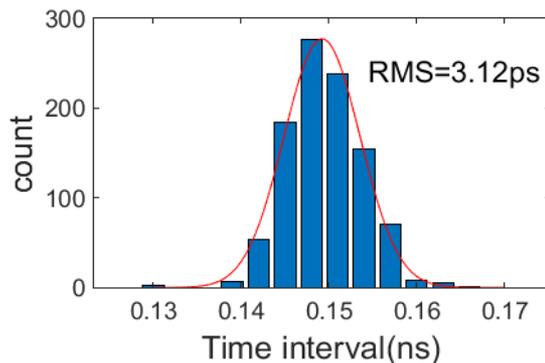
本研究基于全采样点不等权重重心定时展开研究

3、基于机器学习实现不等权重的高精度重心定时（中科大）

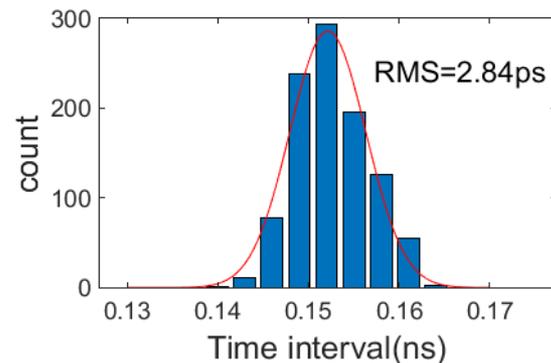
- 类型：感知机
- 不同位置采样点选取不同的权重，基于计算得到理论最佳权重值
- 实际噪声存在相关性且存在非理想因素，采用神经网络进行参数优化



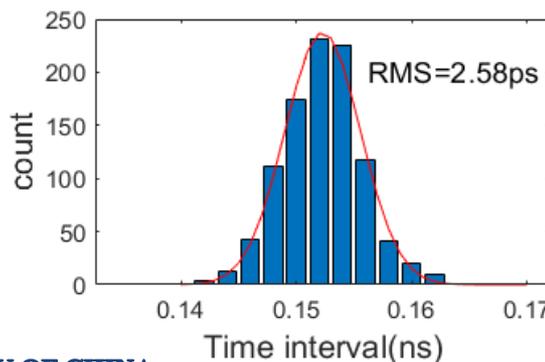
数字恒比定时 3.12 ps



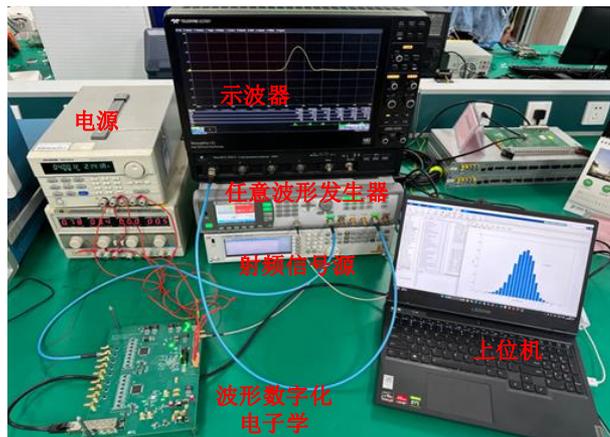
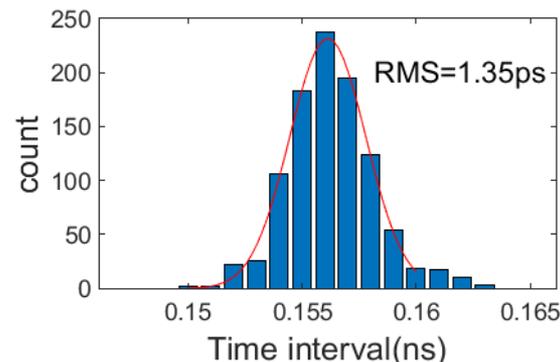
数字重心定时 2.84 ps



数字前沿定时 2.58 ps

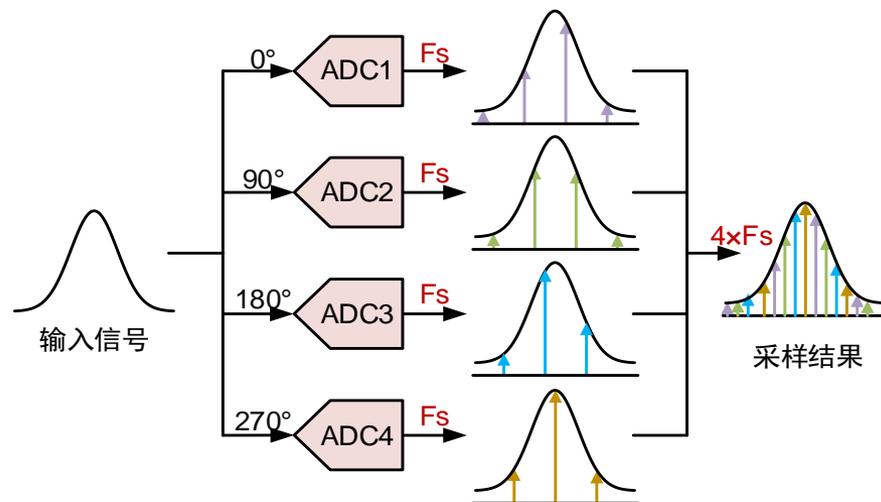


权重重心定时 1.35 ps

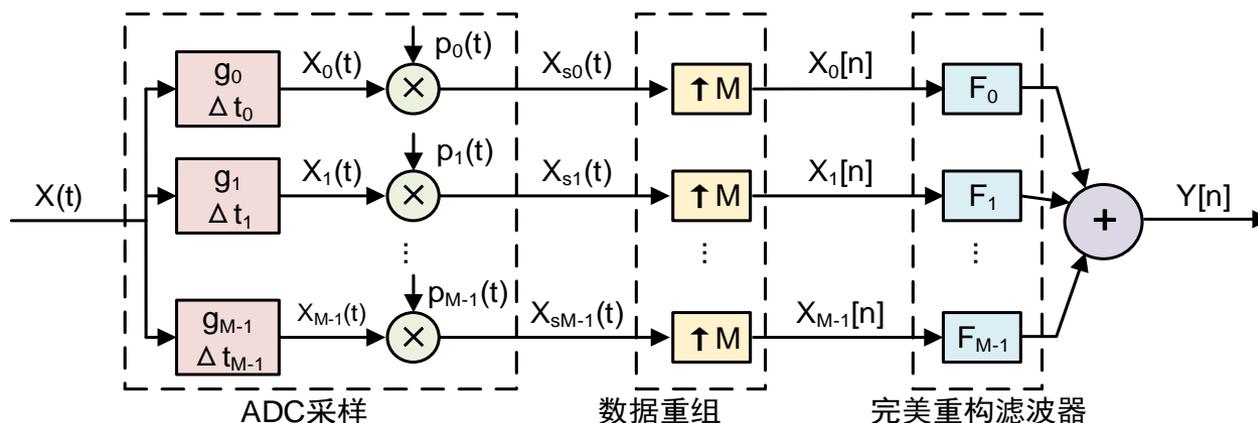


4、基于机器学习的TIADC失配误差修正方法研究（中科大）

- 类型：感知机
- 多个ADC并联可成倍提高采样率，但失配误差会导致精度下降
- 现有的完美重构算法可有效修正失配误差，但易受到IFFT和截断误差的影响而导致资源消耗高

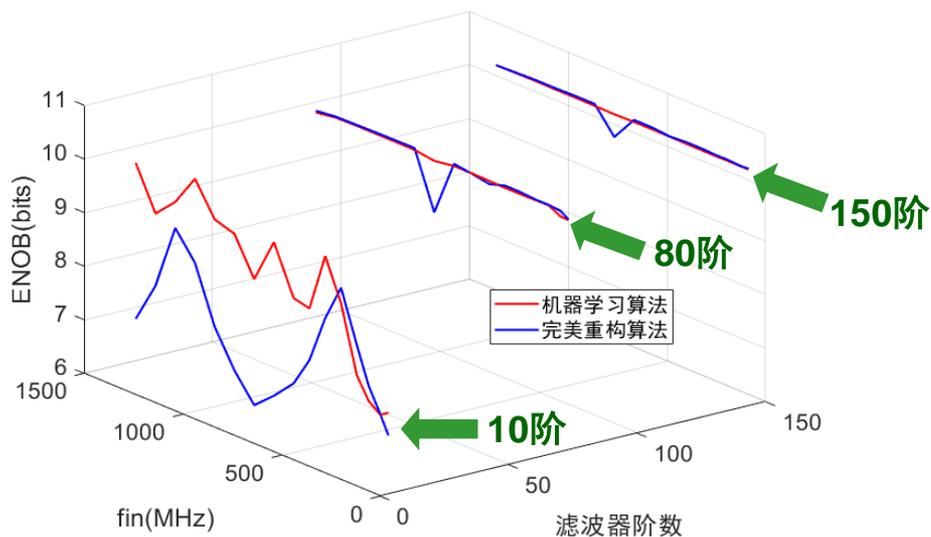


TIADC基本原理

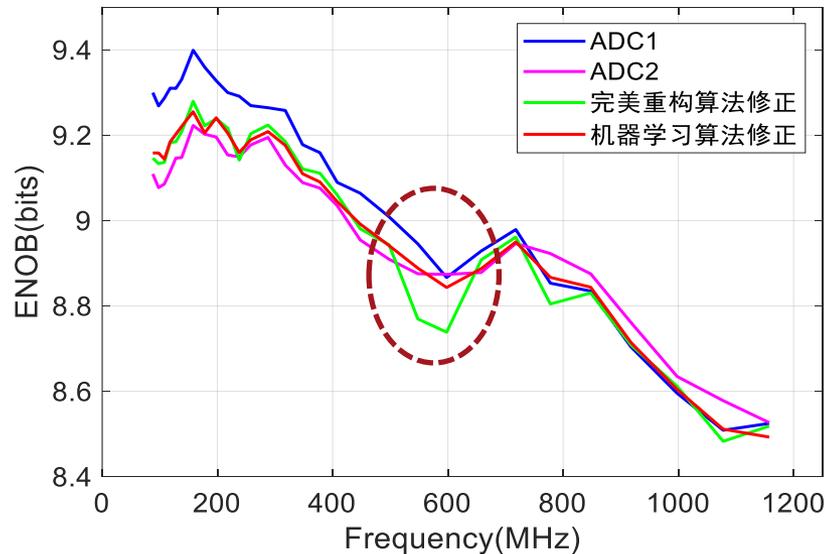


4、基于机器学习的TIADC失配误差修正方法研究（中科大）

- 基于机器学习的失配误差修正方法，达到相同修正效果时滤波器阶数要求降低、资源消耗减少
- 采用2片2.5 Gsps国产ADC，修正系数在线计算、实时修正
- 在80阶滤波器条件下，采样率提升2倍，同时总体精度好于同阶数传统计算得到的滤波器



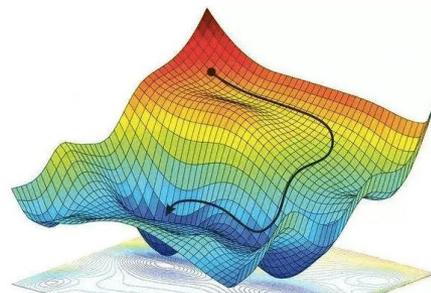
不同频率与阶数下的精度对比



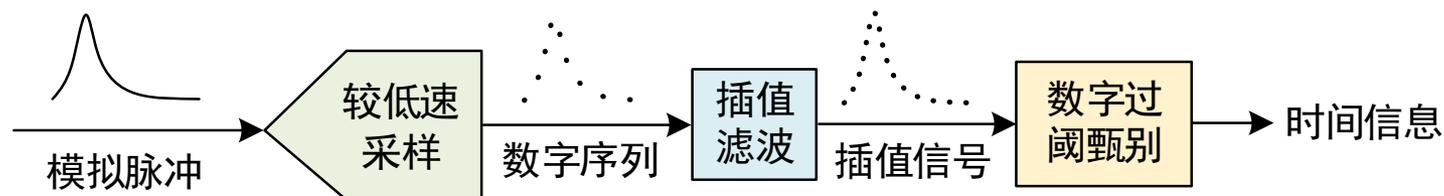
修正效果对比

5、基于遗传算法的低采样率高精度定时方法研究（中科大）

- 类型：遗传算法
- 采用插值定时方法可降低采样率需求
- 要保证插值后数字序列尽可能接近原始脉冲形状，需进行多参数优化



先后进行插值滤波和过阈甄别



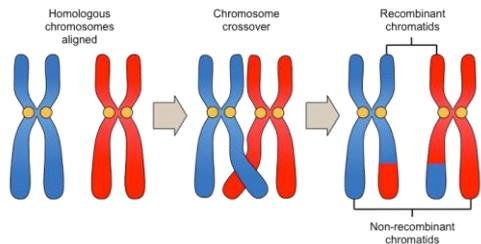
采样率为百兆量级

- 插值倍数 M
- 截止频率 f_c
- 滤波器阶数 N
- 窗函数参数 β
- 阈值位置 c_f

插值定时参数的选取

5、基于遗传算法的低采样率高精度定时方法研究（中科大）

- 该方法模拟自然界生物进化过程与机制求解极值问题
- 遗传过程中随机选择一个变量重新随机赋值，并保留“精英”个体



亲代1:

(M_1, N_1) $(f_{c1}, \beta_1, c_{f1})$

亲代2:

(M_2, N_2) $(f_{c2}, \beta_2, c_{f2})$

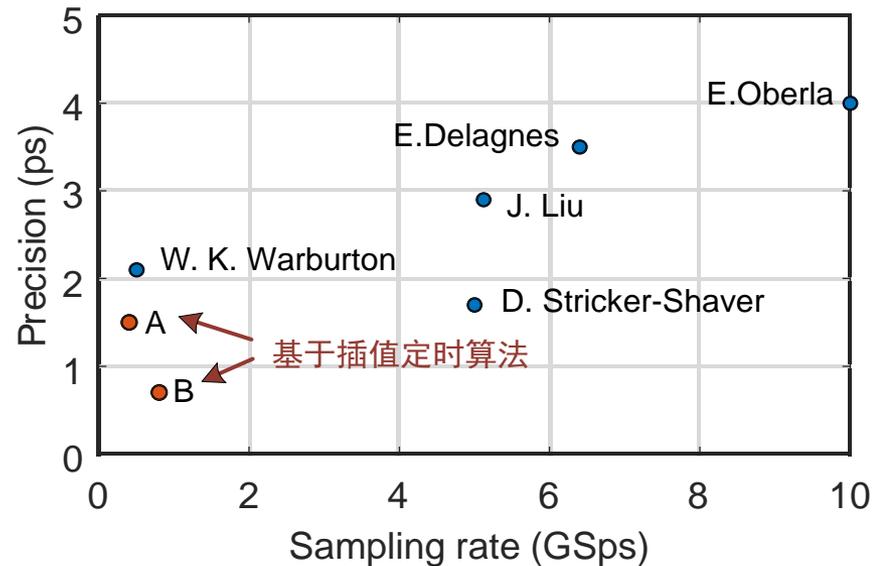
离散型

连续型

(M_1, N_1) 或 (M_1, N_2) $\lambda s_{p1_continuous} + (1-\lambda)s_{p2_continuous}$
 (M_2, N_2) 或 (M_2, N_1) $(1-\lambda)s_{p1_continuous} + \lambda s_{p2_continuous}$

组合成两个子代个体

参数优化流程



A: 脉冲上升时间约5 ns;
B: 脉冲上升时间约2.5 ns;

修正效果对比

研究方向分类

特征识别与提取

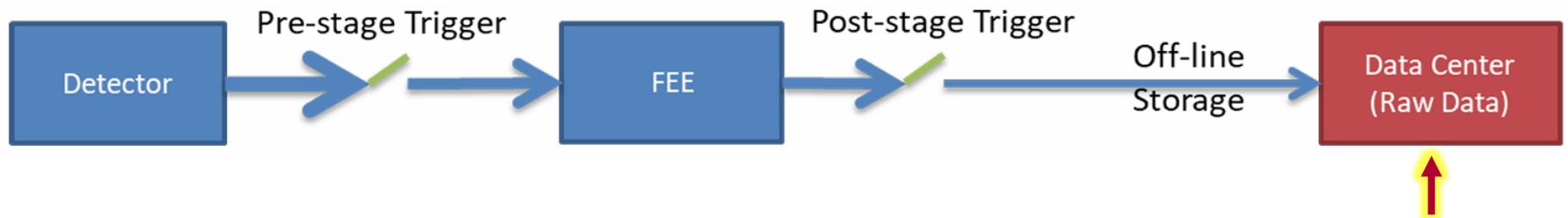
- 提高信号探测效率
- 增加初始有效信息比例、提高信息提取的精度

数据压缩

- 前端电子学
- 后端电子学

粒子物理实验前端电子学的数据带宽瓶颈

- 在高能物理实验中，探测器输出波形包含了感兴趣的信息，可将波形传输至后端存储，并通过离线方法进一步分析



人工智能与机器学习分析方法
大量面向离线数据，基于GPU等器件实现

- 随着高能物理实验亮度的提高以及探测器颗粒度的增加，上述数据处理方式对数据传输带宽提出了很高的挑战，并导致了可观的功耗
 - NICA-MPD量能器前端电子学单板面积344 cm²，数据传输带宽48 Gb/s，功耗16 W
 - HL-LHC内层像素探测器计数率高达3 GHz/cm²
 - HL-LHC量能器参与触发的总数据量高达300 TB/s

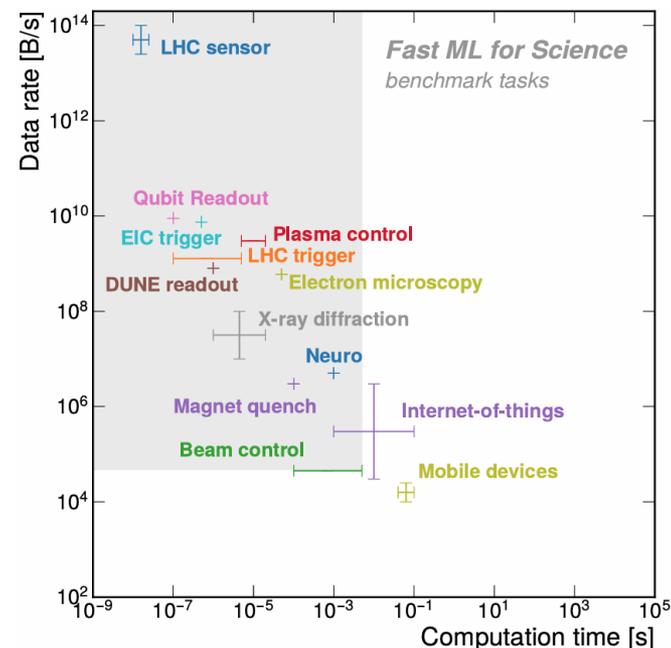
粒子物理实验上的边缘计算

- 一种有效的方法是，在靠近数据产生源处部署算力，对产生的数据进行实时处理，降低所需的上行数据带宽，称之为边缘计算。



- 在高能物理领域，人工智能已被应用于粒子鉴别，并采用GPU等器件进行加速。但将其应用于前端电子学将面临独特的挑战：

- 低延时
- 低功耗
- 抗辐照
-



前端电子学中的数据压缩

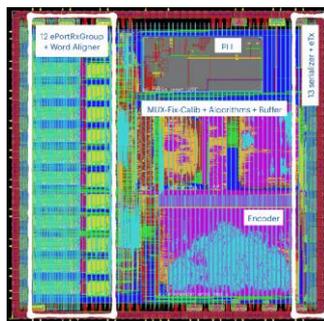
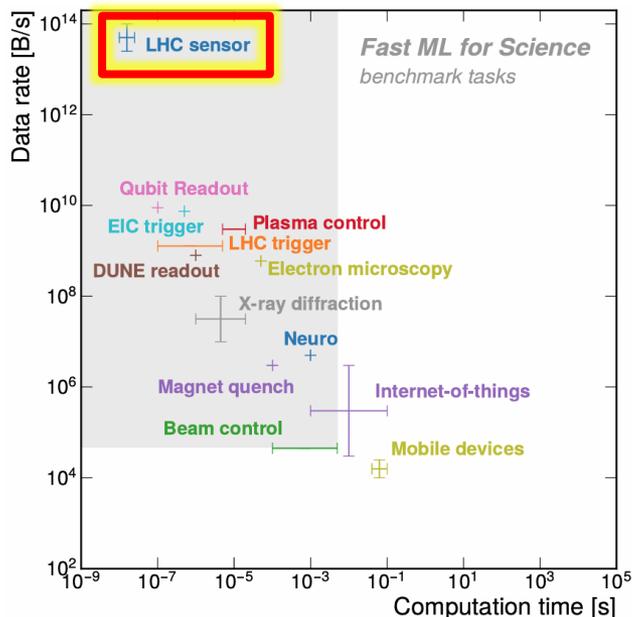
- 有损压缩算法
 1. HL-LHC上HGCal数据有损编码压缩（费米实验室）
- 像素芯片中背景事例筛除
 1. Hybrid型像素芯片低动量事例筛除（西北工业大学）
 2. MAPS型像素芯片低动量事例筛除（费米实验室）

前端电子学中的数据压缩

- 有损压缩算法
 1. HL-LHC上HGCal数据有损编码压缩（费米实验室）
- 像素芯片中背景事例剔除
 1. Hybrid型像素芯片低动量事例剔除（西北工业大学）
 2. MAPS型像素芯片低动量事例剔除（费米实验室）

用于HL-LHC上HGICAL的有损压缩算法

- 类型：自编码器
- HL-LHC计划使用高粒度端盖量能器（HGICAL），六百万读出通道产生的数据对传输带宽提出了巨大挑战
- HGICAL对前端HGCROC ASIC输出的数据进行压缩：
 - ECON-D 对数据进行零压缩
 - ECON-T 采用抗辐射神经网络进行数据压缩



芯片版图



芯片照片

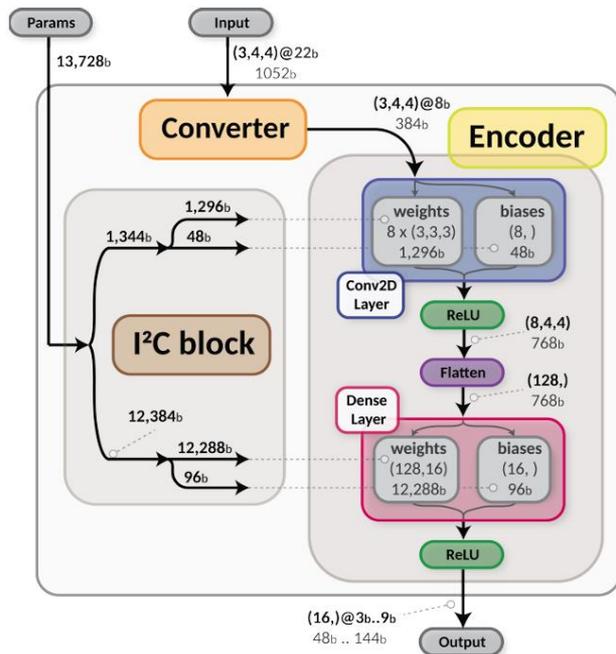
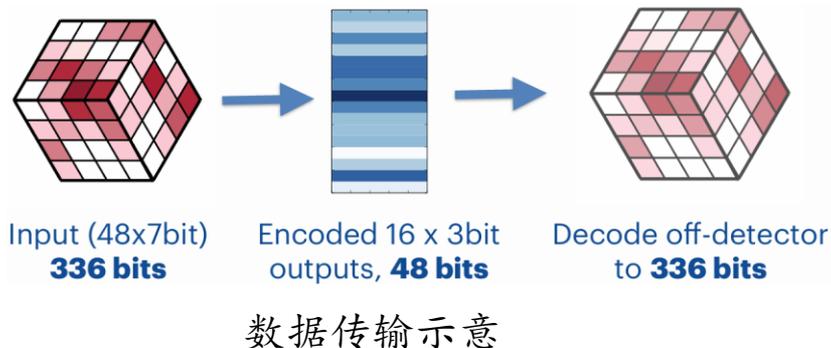


不同应用下的数据率与计算延时需求

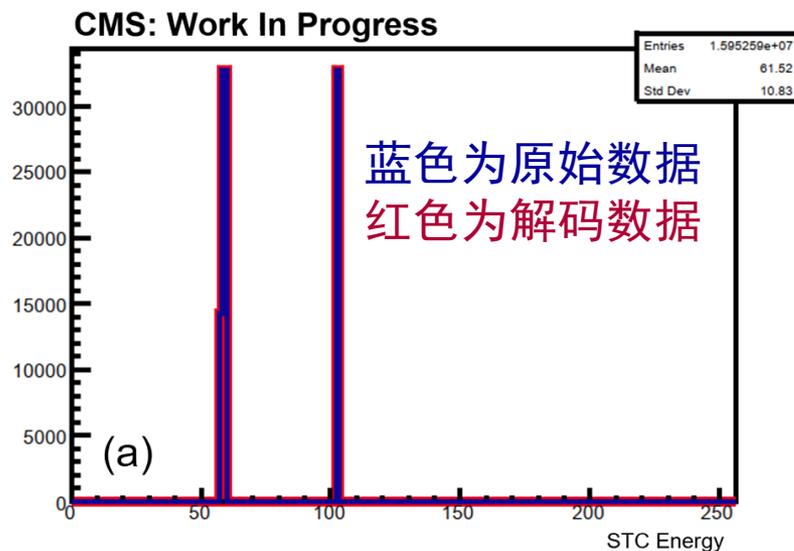
读出电子学各节点的数据率

用于HL-LHC上HGICAL的有损压缩算法

- ECON-T保留总能量信息，但采用有损编码方式对能量沉积图像进行压缩，降低传输带宽需求
- 束流测试结果满足设计需求



数据压缩处理流程



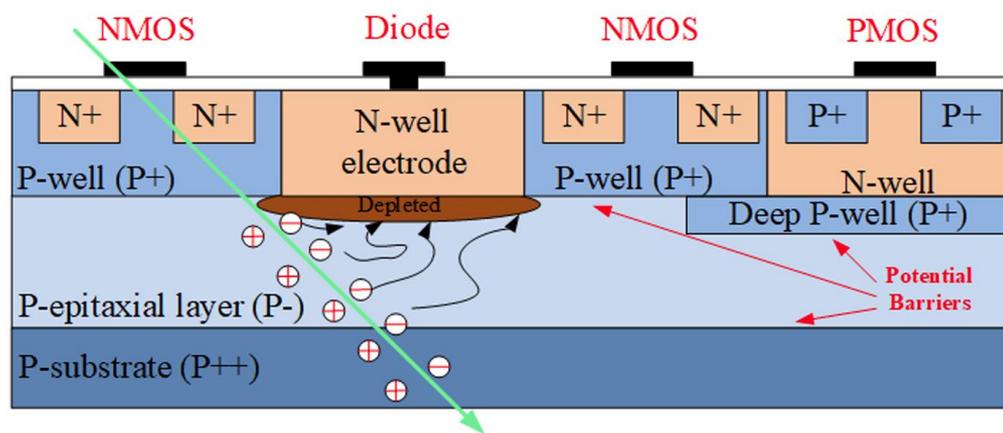
解码数据与原始数据对比

前端电子学中的数据压缩

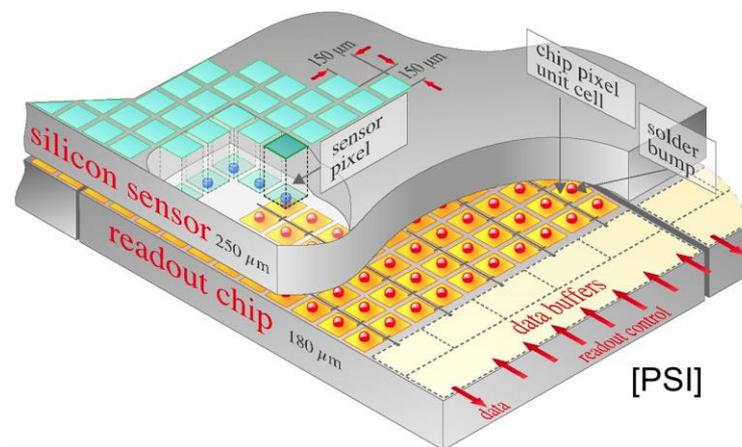
- 有损压缩算法
 1. HL-LHC上HGCal数据有损编码压缩（费米实验室）
- 像素芯片中背景事例剔除
 1. Hybrid型像素芯片低动量事例剔除（西北工业大学）
 2. MAPS型像素芯片低动量事例剔除（费米实验室）

像素芯片

- CMOS像素探测器将传感器和读出电路堆叠放置（hybrid pixel detector）或集成在同一块衬底（Monolithic active pixel sensors, MAPS）
- CMOS像素探测器在大型高能物理实验中已经得到了较为广泛应用，包括ALICE、ATLAS、CMS、LHCb和NA60等。



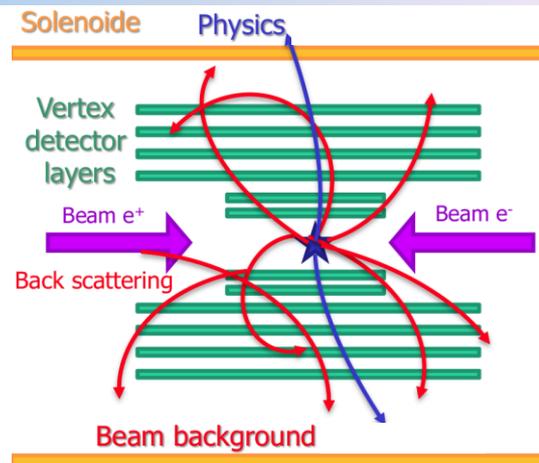
MAPS型像素探测器



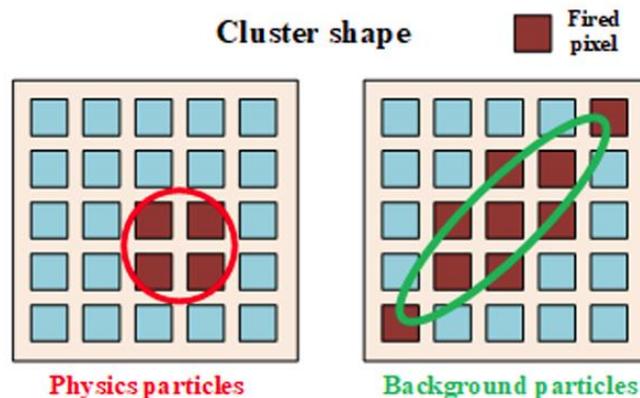
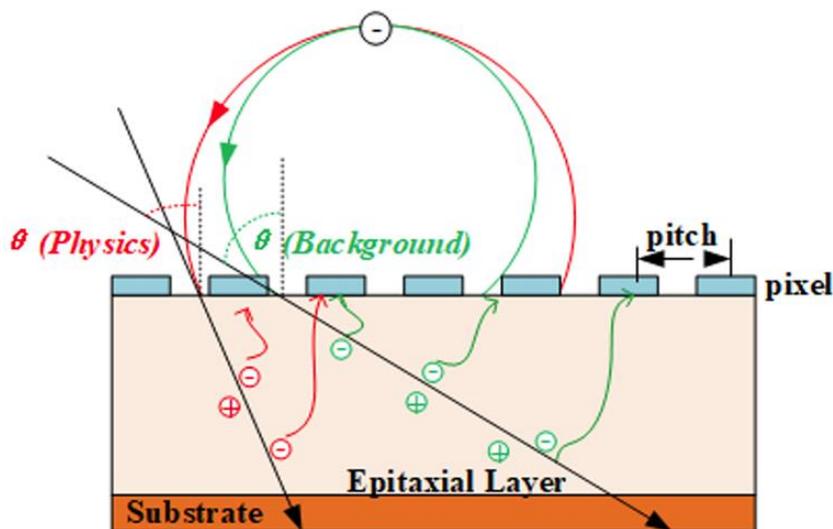
Hybrid型像素探测器

像素芯片中背景事例筛除

- 高亮度实验中不可避免存在大量来自背景
的击中信息，这些带电粒子通常具有较小动量
在探测器表面形成具有细长形状的簇。
- 通过神经网络在线提取径迹特征，剔除背景
事例，可降低读出数据率需求。



物理事例与背景径迹区别

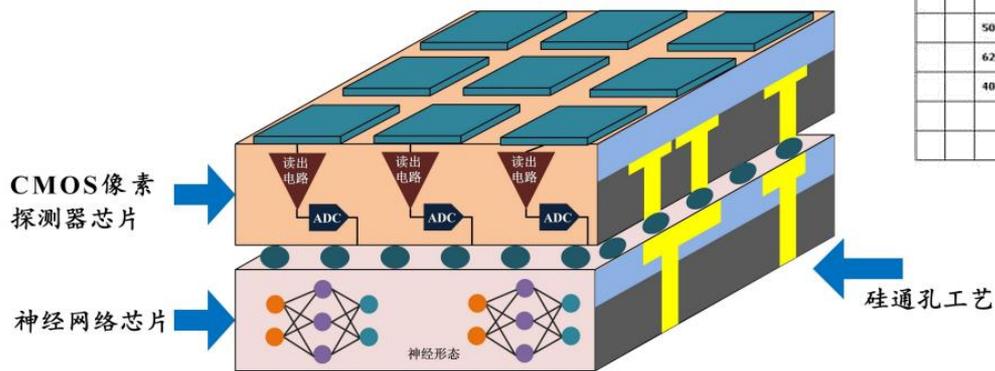


物理事例与背景击中像素形状区别

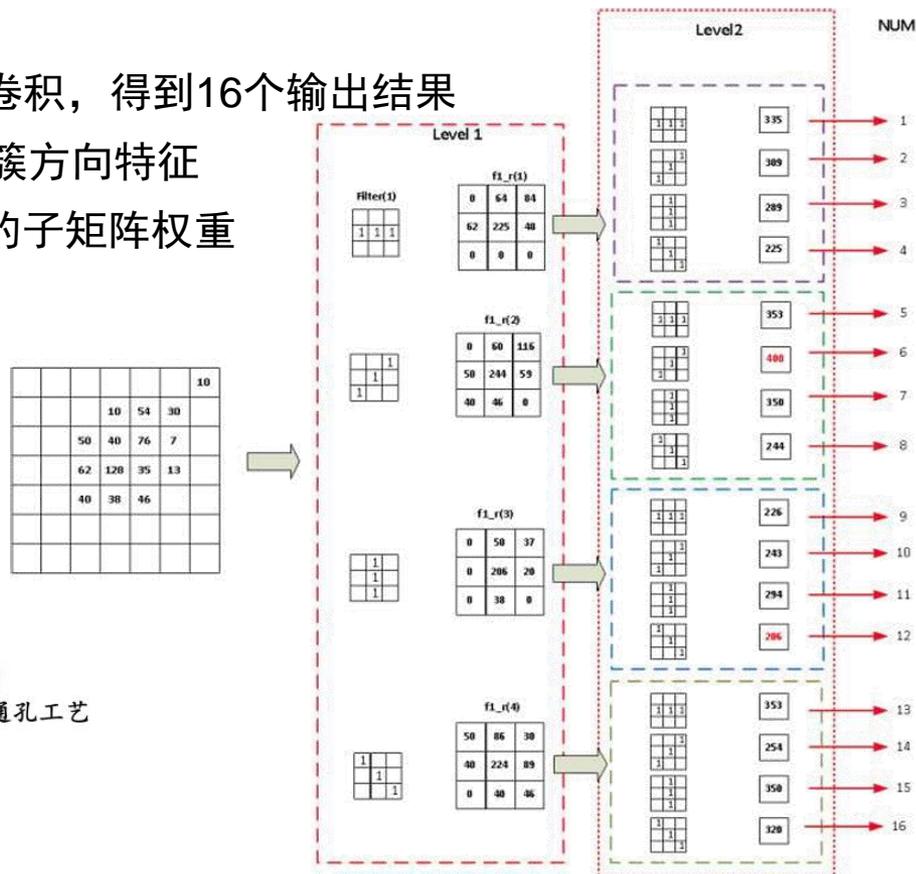
1、Hybrid型像素芯片低动量事例筛选（西北工业大学）

- 类型：主成分分析
- 在此方向上，西北工业大学展开了相应的研究工作
- 其中关键算法包括簇识别和簇特征提取，一种简单的簇特征提取方法如下：

- 将像素内的电荷信息与子矩阵进行两次卷积，得到16个输出结果
- 根据16个结果中的最大值与最小值得到簇方向特征
- 应用时需根据背景信号特征选择最合适的子矩阵权重



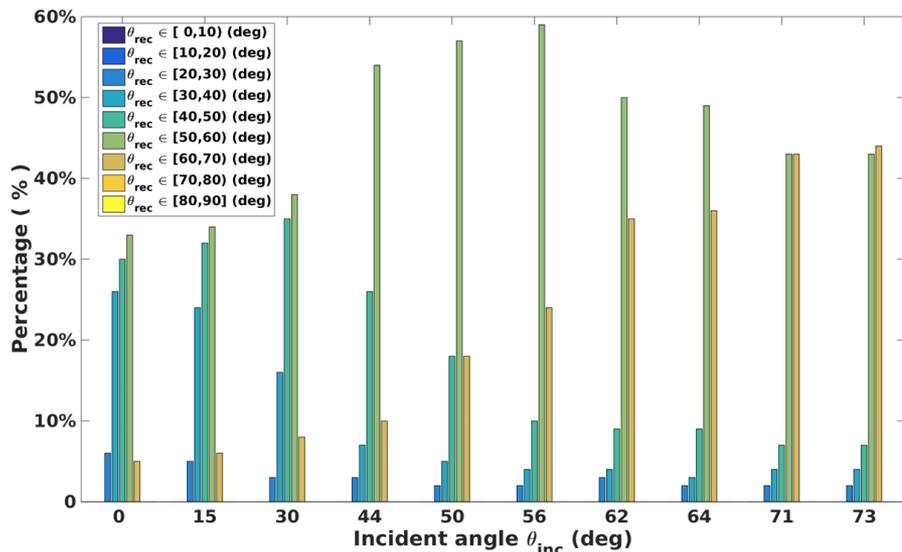
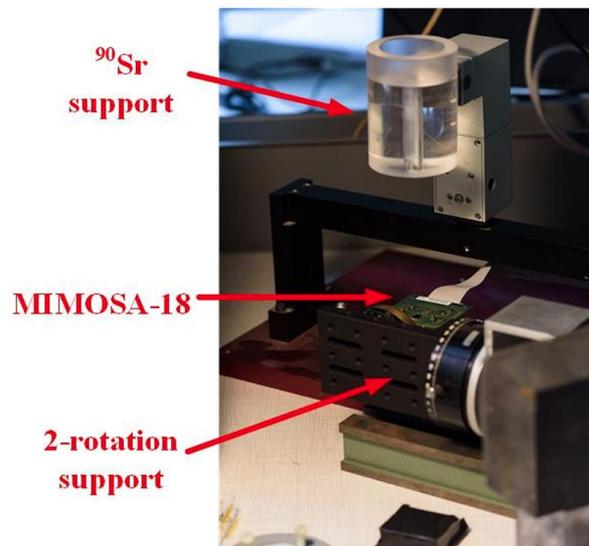
Hybrid型像素芯片结构示意图



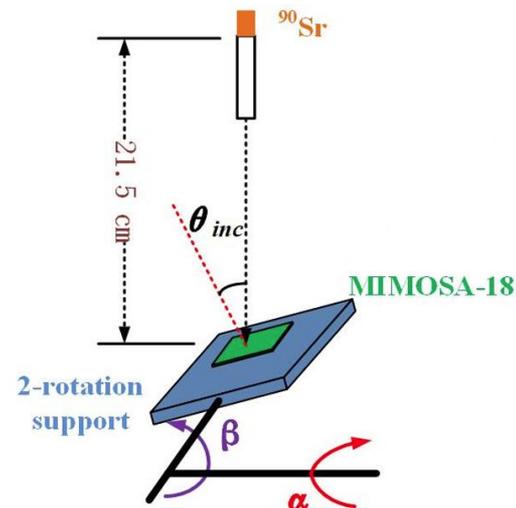
通过卷积判断簇方向的流程

1、Hybrid型像素芯片低动量事例筛选（西北工业大学）

- 基于MIMOSA-18像素探测芯片，在FPGA中进行了可行性验证，包括簇找寻模块、主成份分析模块、人工神经网络模块。
- 测试结果与离线神经网络重建结果一致，基本符合入射角度的变化规律。



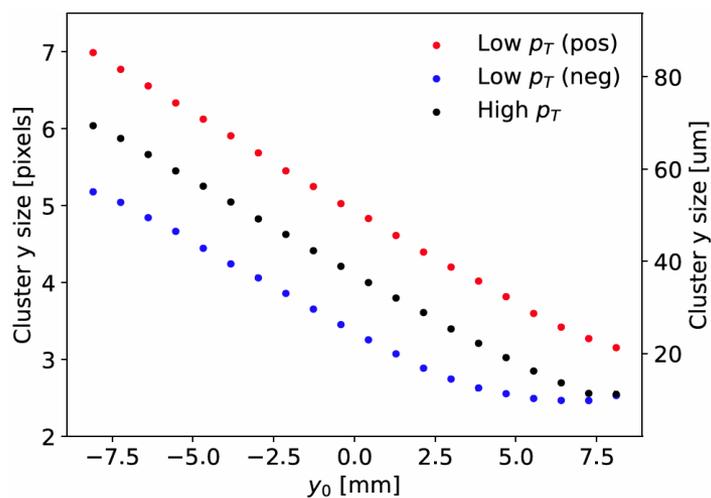
入射粒子角度重建结果



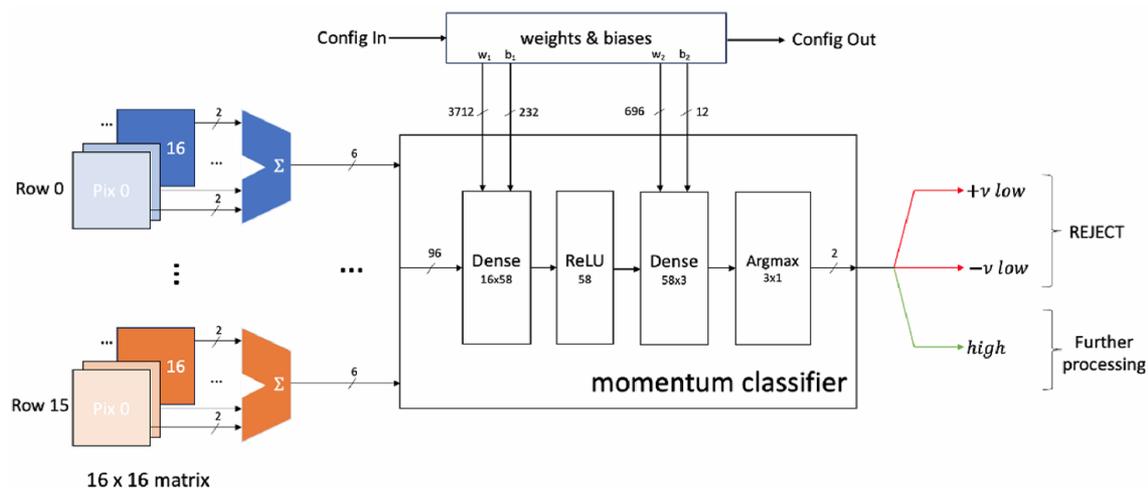
可改变入射粒子角度的测试台

2、MAPS型像素芯片低动量事例筛除（费米实验室）

- 类型：感知机
- Hybrid型像素芯片物质量相对偏高，费米实验室针对CMS内径迹探测器读出需求展开了MAPS型像素芯片像素内人工智能电路的研究
- 在HL-LHC升级后，CMS内径迹探测器将参与CMS Level-1 trigger产生，将面临严峻的实时数据传输带宽压力
- CMS不会重建动量 <100 MeV的粒子径迹，可在片上进行剔除



不同动量粒子的cluster size

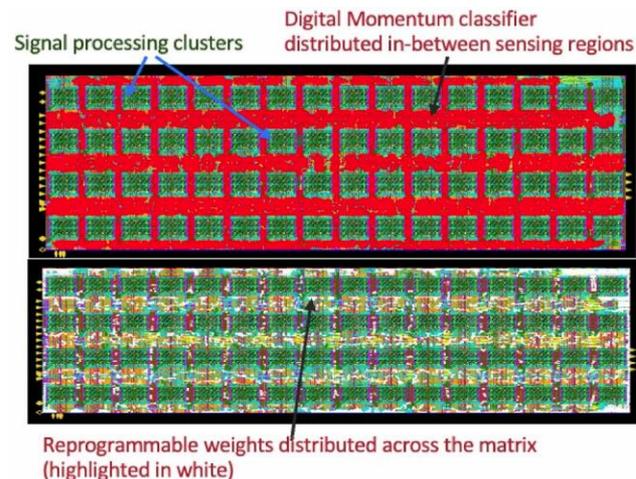


权重可重配置的动量分类智能电路框图

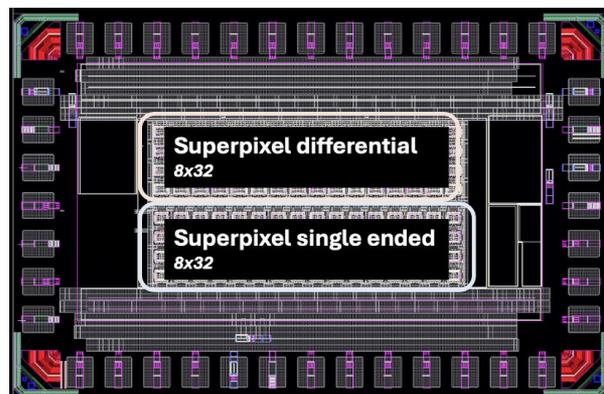
2、MAPS型像素芯片低动量事例筛除（费米实验室）

- 基于开源的CMS已有实验数据，采用QKeras进行包含截断误差分析的模型训练，然后通过hls4ml工具包转换为RTL实现。

工艺	28 nm
像素尺寸	$50\ \mu\text{m} \times 12.5\ \mu\text{m}$
等效输入电荷噪声	$\sim 30\ e^-$
信号效率	$> 90\%$
读出数据带宽	压缩至原来的24%~45%
总功耗	$\sim 6\ \mu\text{W}/\text{pixel}$
智能识别电路面积	$\sim 350\ \mu\text{m}^2/\text{pixel}$



布局（绿色为原本像素电路、红色为新加入的智能识别电路）



版图

研究方向分类

- 国内外已对人工智能在前端电子学中的应用展开了相关研究
- 受限于前端电子学的尺寸、功耗等限制，需在网络层数与参数数量均有限的条件下进行设计与实现

研究机构	用途	采用的人工智能类型	网络层数	参数数量	应用类型
CERN	小信号提取	卷积神经网络	3	21217	具体应用
BNL	幅度提取	感知机	4	377 ~ 9309	具体应用
华中师范大学	时间和能量测量	自编码器+感知机	8+3	/	具体应用
中科大	脉冲定时	感知机	1	26	电路结构确定， 与已有算法有机结合
中科大	TIADC误差修正	感知机	1	80	电路结构与算法确定， 参数最佳化
中科大	脉冲定时	遗传算法	/	/	电路结构与算法确定， 参数最佳化
费米实验室	数据压缩	自编码器	2	13728	具体应用
IPHC & 西北工业大学	低动量事例筛选	主成分分析	/	~72	电路结构确定， 与已有算法有机结合
费米实验室	低动量事例筛选	感知机	2	4652	电路结构确定， 与已有算法有机结合

总结

- 随着高能物理的发展，粒子物理实验领域已经展开一系列人工智能方法的应用研究。
- 通过引入人工智能，可提升前端电子学处理能力：
 - 直接应用人工智能进行海量数据处理
 - 在已有的电子学架构基础上，将人工智能与现有算法有机结合
 - 信号处理算法明确但参数难以计算，采用人工智能进行最优化设计
- 由于粒子物理实验前端电子学独有的特点，人工智能在此领域应用需要进行定向优化和相应的独特设计，以适应短延迟需求、抗辐照环境等。
- 随着未来人工智能技术的进一步发展，可以期望相应技术在本领域发挥更大的作用。

谢 谢