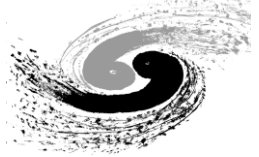


研究生8月考核报告

专业: **计算机应用技术**

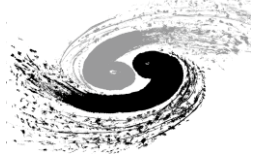
课题组: **触发与数据获取组**

考核人: **张叙**



- **CEPC在线业务算法研究**
 - 面向事例间的数据筛选算法: 霍夫变换
 - 面向事例内的数据筛选算法: CNN
- **CEPC在线异构软件框架**

CEPC背景:

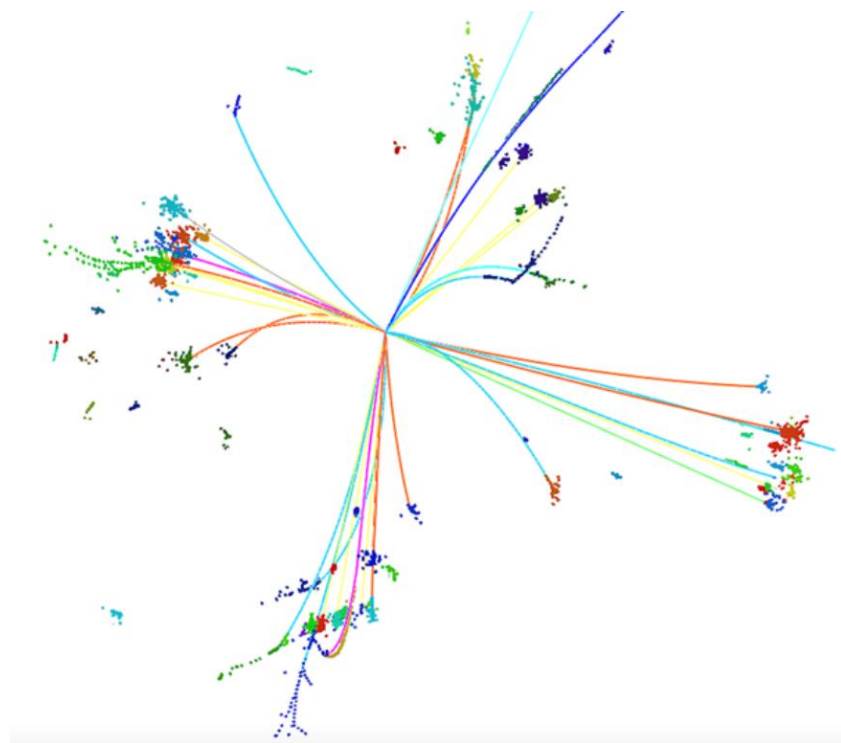
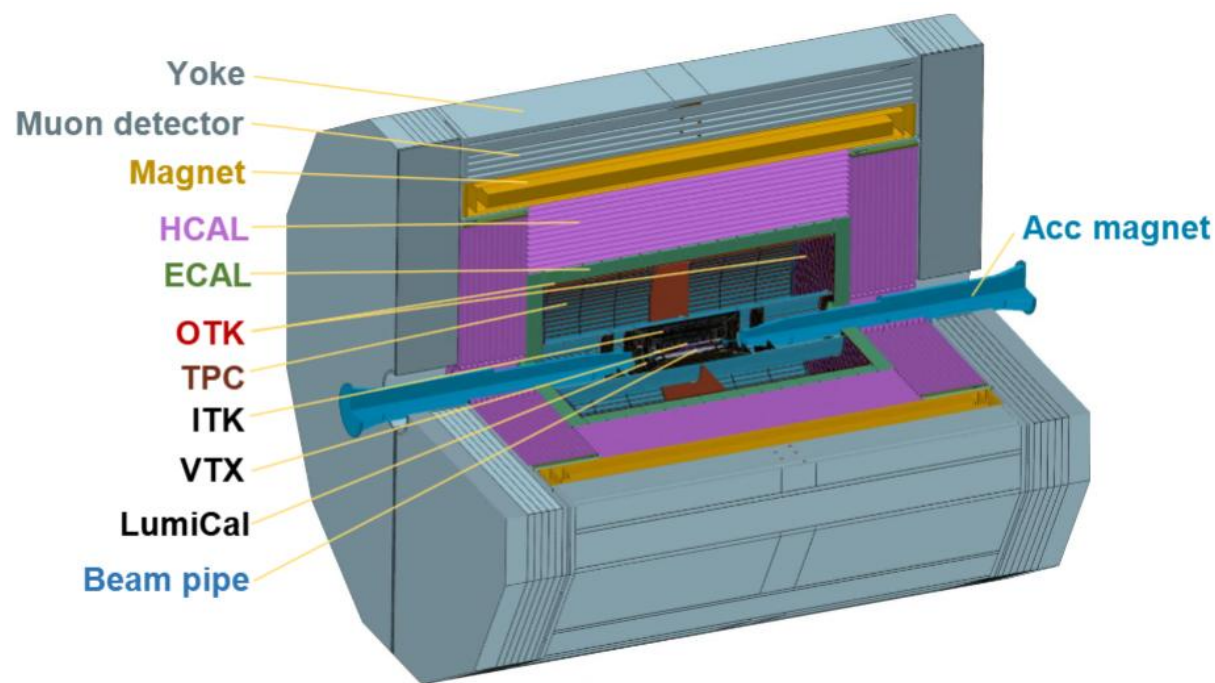


研究背景: CEPC将产生海量探测器数据，需要在在线阶段完成快速筛选与压缩。

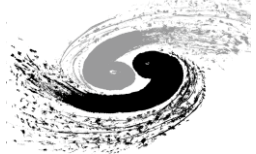
物理依据: 带电粒子在探测器中形成可识别的径迹，为区分物理事例与本底提供关键特征。

算法目标: 研究基于**径迹特征**的在线筛选与压缩算法，在有效保留物理信号的同时降低数据规模。

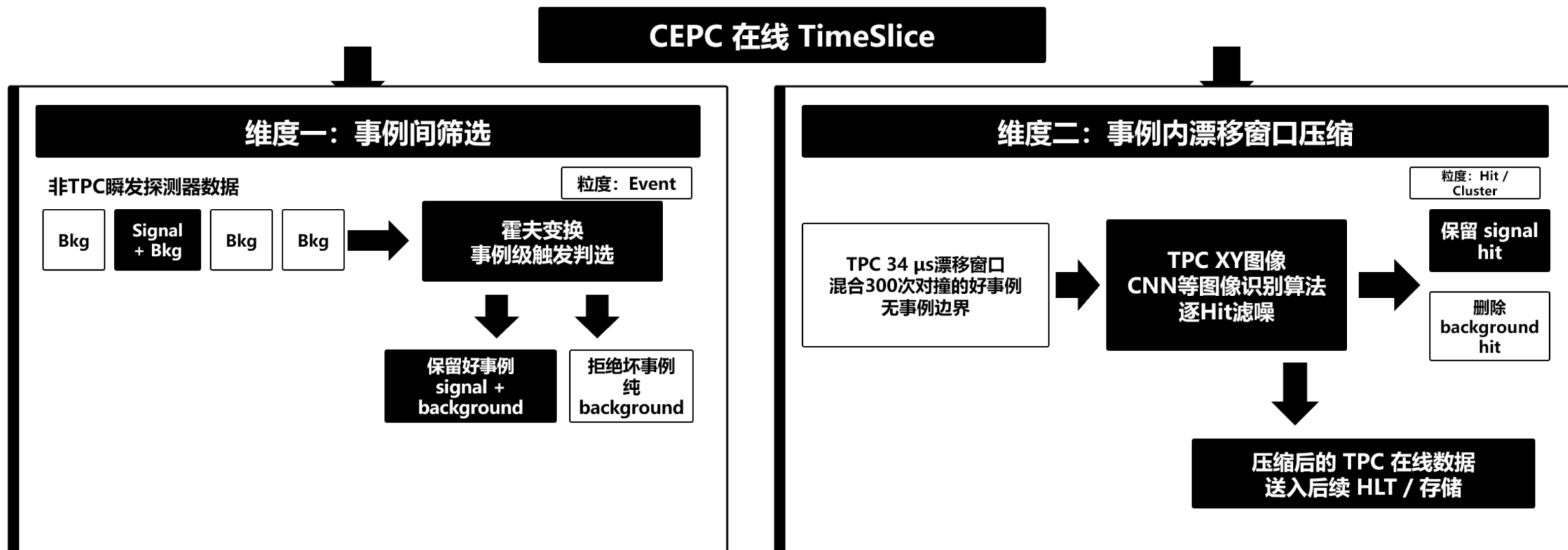
系统支撑: 开发**CPU+GPU**异构在线框架，为径迹识别等**高并行**算法提供统一、高效的部署与执行环境。



CEPC在线业务算法

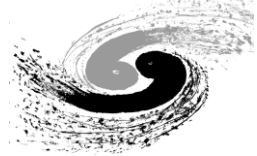


事例间筛选：普通物理事例具有清晰的事例时间边界，可利用**径迹**等事例特征区分物理事例与本底事例。
事例内筛选：TPC存在**漂移**时间，单个时间窗口内叠加大量Hit，需要利用Hit间的空间关联进行筛选与压缩。



两条独立路线：霍夫变换进行Event级事例筛选，CNN等图像识别算法进行Hit级漂移窗口压缩

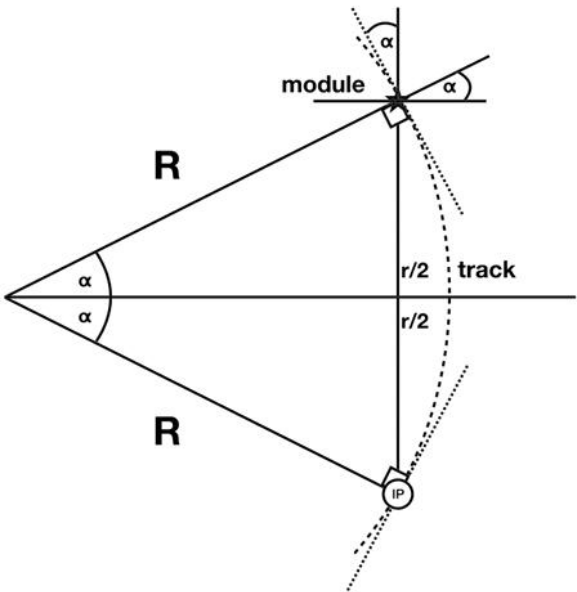
CEPC在线业务算法：事例的筛选



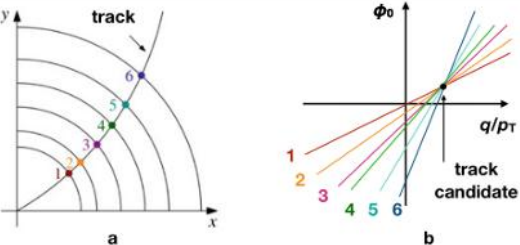
筛选原理： 利用**霍夫变换**提取典型径迹，根据是否存在有效径迹判断事例是否保留。

存在问题： 固定网格与硬阈值难以覆盖不同事例中的多样化径迹。

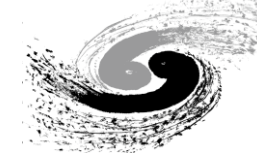
改进方法： 采用 “**霍夫变换 + 神经网络**”，提高径迹识别能力，减少信号漏选和本底误选。



方法	Signal efficiency	Background rejection	AUC
纯三维霍夫变换	98.40%	94.93%	约0.9953
霍夫候选 + DeepSets	98.67%	99.73%	约0.9975

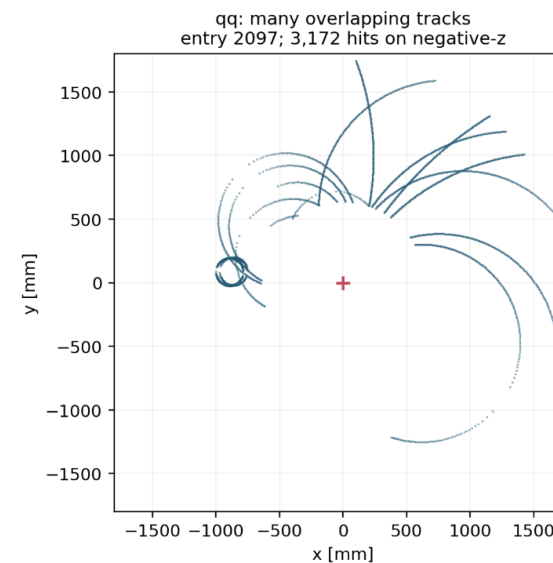
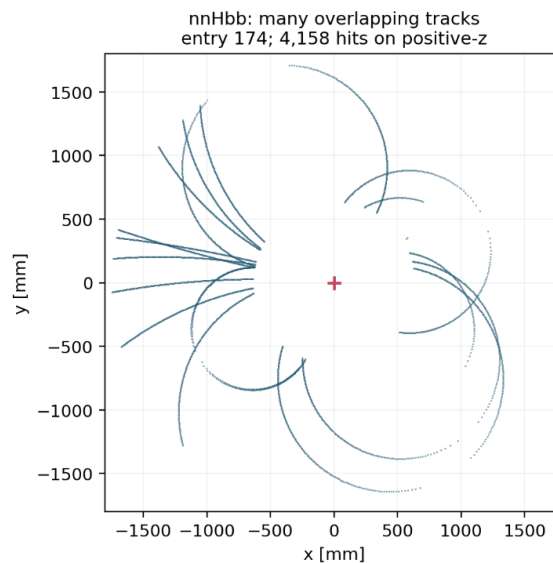
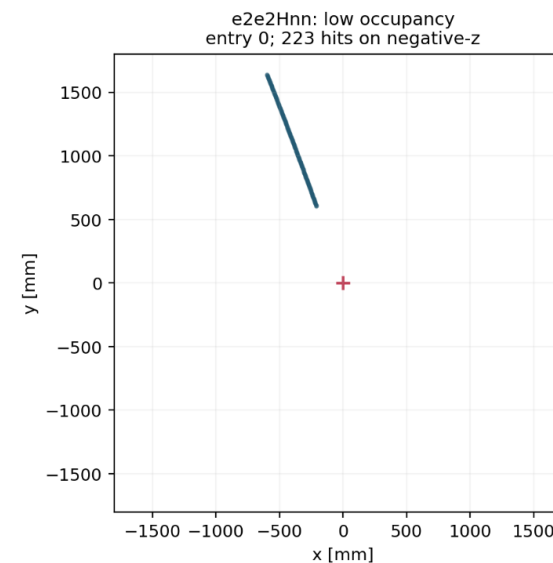
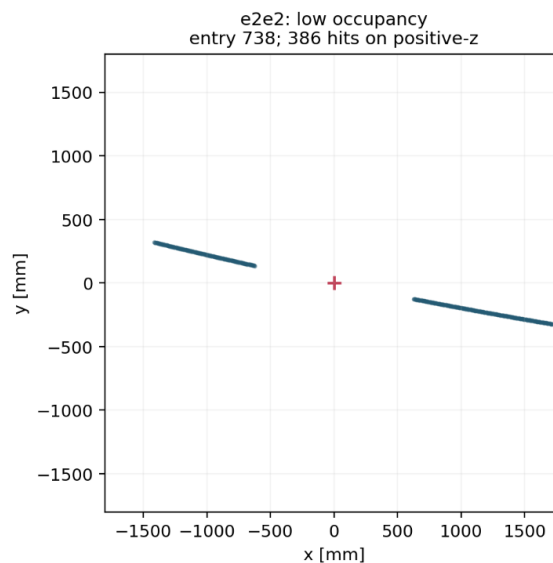


CEPC在线业务算法：事例内数据的筛选

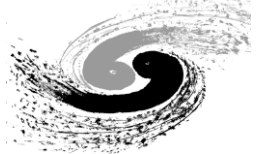


典型TPC事例研究：

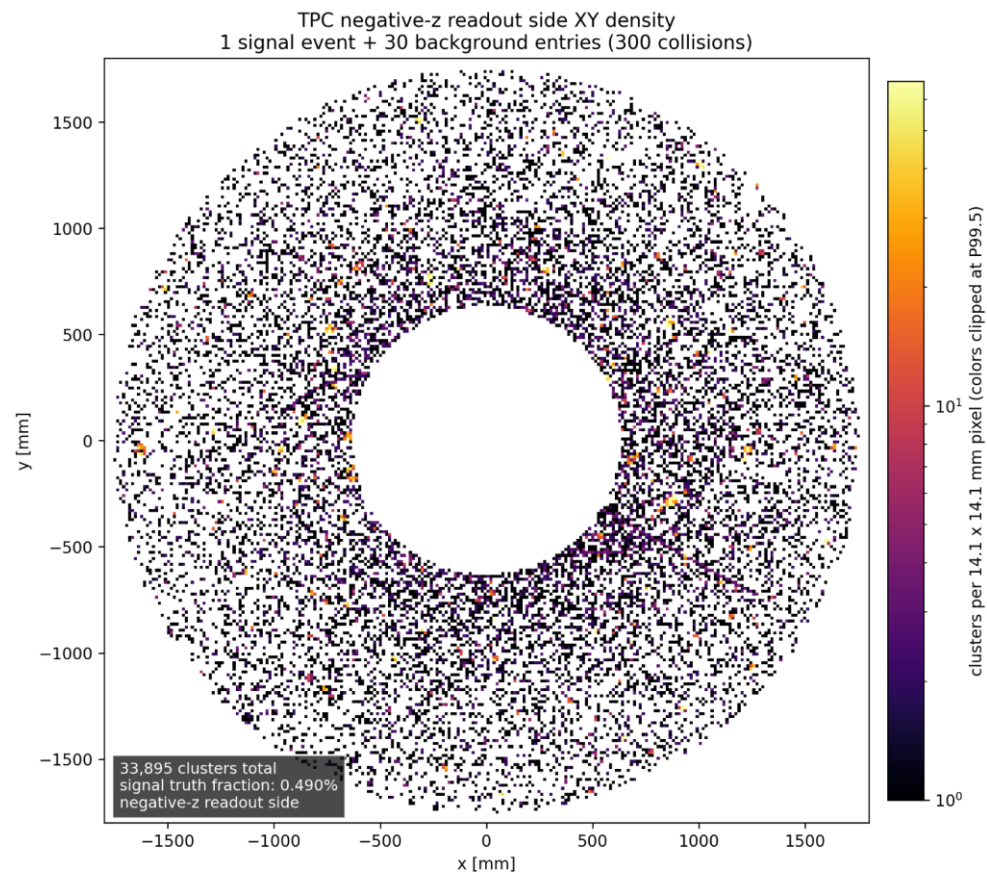
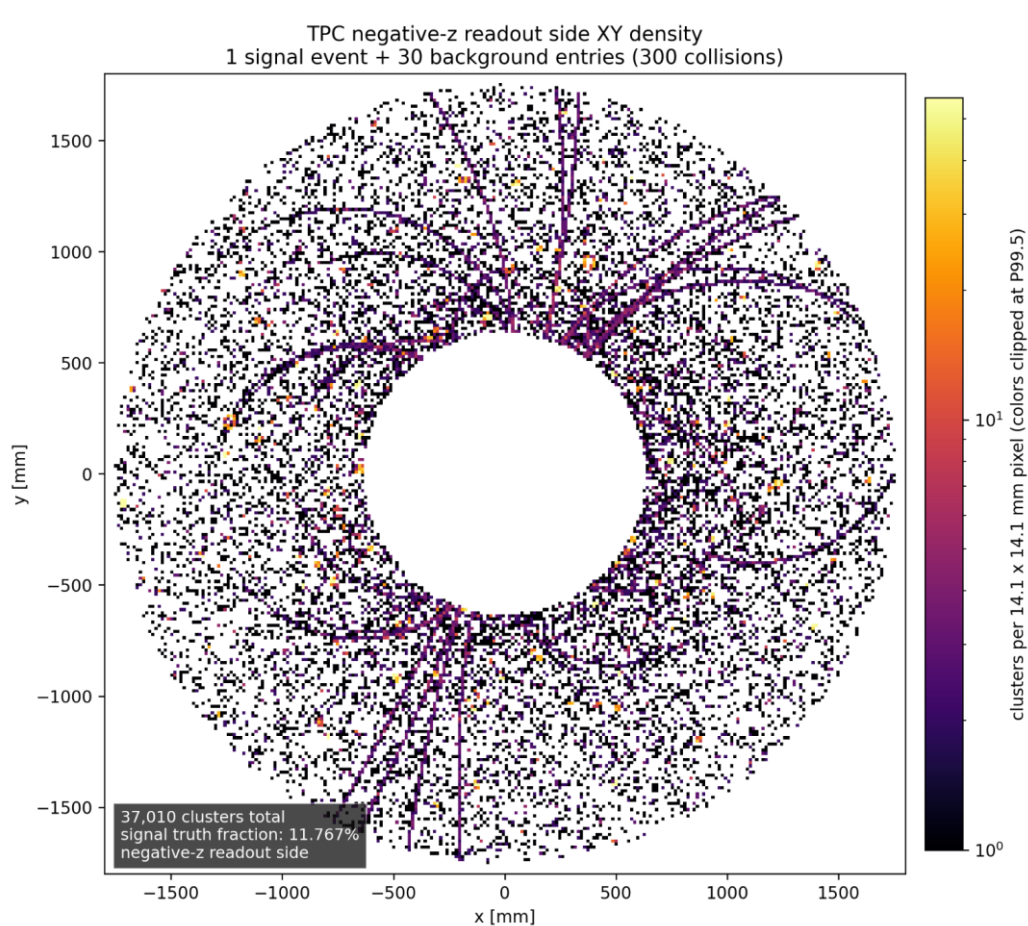
- 相对少径迹事例： e2e2、 e2e2Hnn
- 相对多径迹事例： nnHbb、 qq



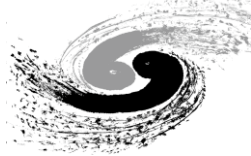
CEPC在线业务算法：事例内数据的筛选



事例内数据筛选原理：将300次本底碰撞与物理事例的Hit混合，分别研究少径迹和多径迹两类典型场景。从Hit的空间分布图中可以观察到局部径迹结构，因此可采用CNN等**图像识别**方法进行信号Hit识别与本底过滤。

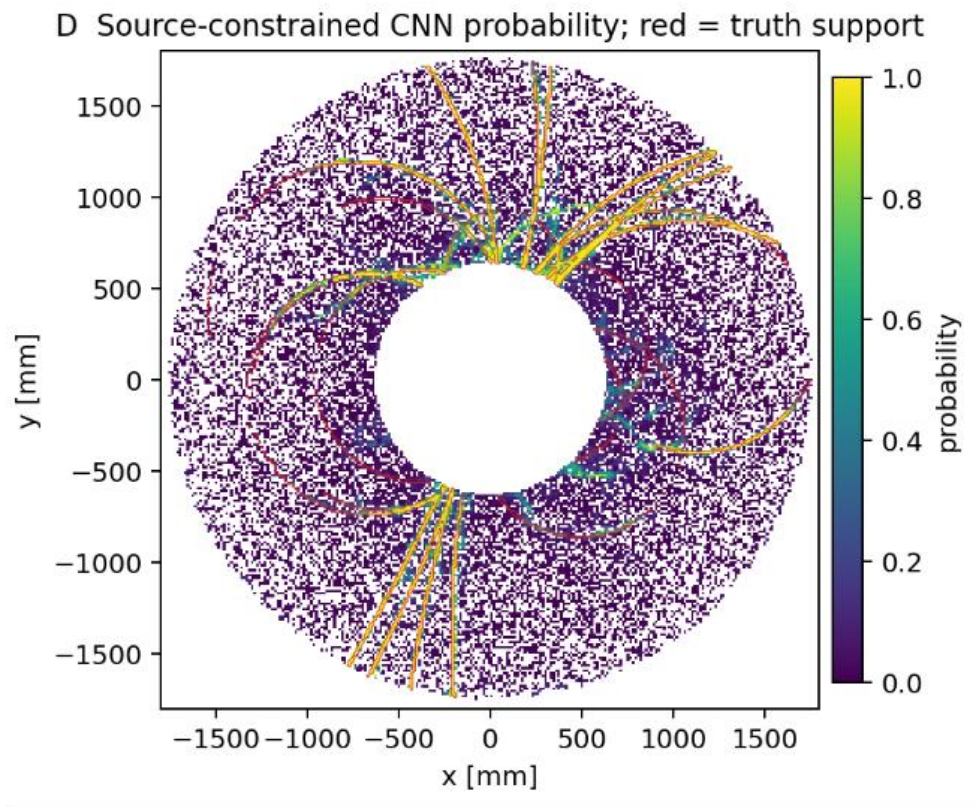


CEPC在线业务算法：事例内数据的筛选



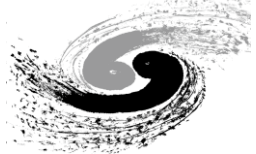
算法原理： CNN通过卷积核提取Hit的局部空间特征，并对Hit进行评分，判断其是否保留。

算法结论： 初步结果表明CNN具有应用潜力，本底拒绝率可达约80%；但仍需进一步理解，例如e2e2nn事例的信号保留率相对较低。



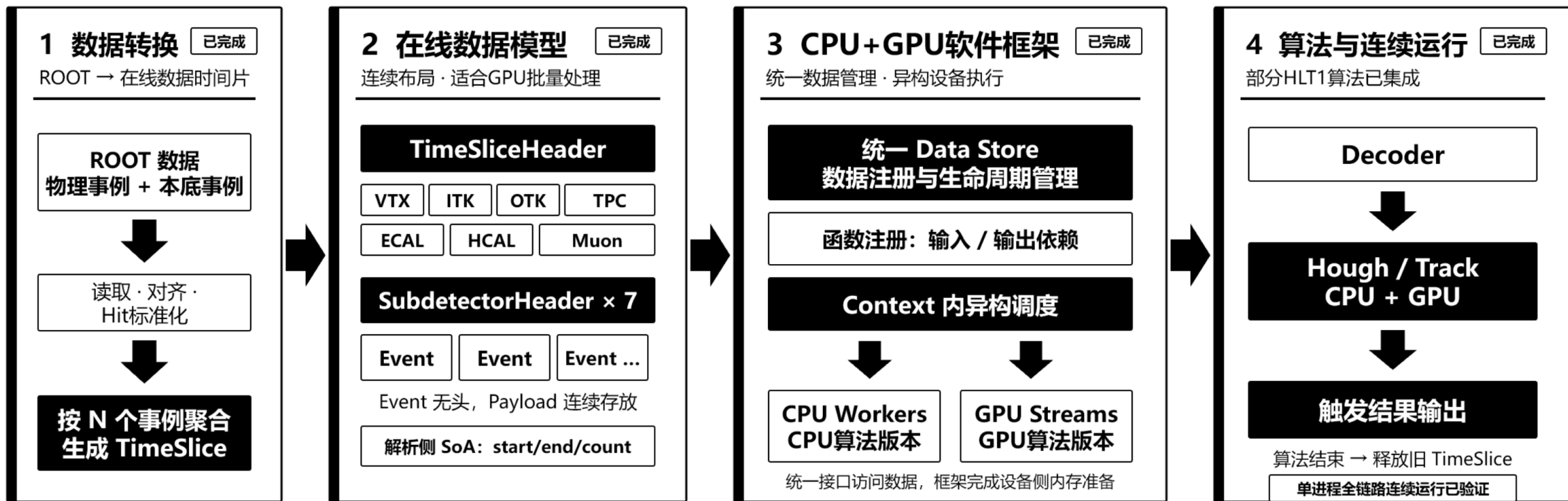
目标	总体 S/R	e2e2 S/R	e2e2nn S/R	nnbb S/R	qq S/R
T99	99.52 / 64.02	99.74 / 83.97	99.33 / 83.01	99.70 / 39.11	99.23 / 49.40
T98	98.19 / 75.43	99.33 / 91.03	85.54 / 90.77	98.70 / 55.20	98.52 / 64.23
T97	97.81 / 80.08	98.59 / 95.10	87.82 / 94.81	98.32 / 59.47	97.92 / 70.50
T96	96.64 / 78.93	98.98 / 85.96	97.01 / 86.44	96.48 / 69.00	96.73 / 74.12
T95	96.36 / 84.83	99.16 / 95.49	82.07 / 95.07	96.95 / 70.66	96.62 / 77.81

CEPC在线异构软件框架进展

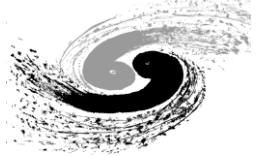


开发面向CEPC在线处理的CPU+GPU异构软件框架。

- 完成了从**ROOT数据**到**在线二进制数据**的转换
- 设计并实现了适合GPU并行处理的**在线数据模型**。
- 完成了部分算法集成及单进程**全链路**连续运行测试。



时间片级连续流: ROOT输入 → 时间片构建 → CPU/GPU处理 → 结果输出 → 释放旧数据 → 下一时间片



总结:

- 已形成初步的在线算法方案，包括面向事例间筛选的霍夫变换和面向事例内筛选的CNN。
- 已完成在线异构框架原型demo，实现时间片数据组织、算法调度和CPU/GPU执行。
- 已完成部分算法集成，并通过功能测试和单进程连续运行测试。

下一步:

- 继续优化在线筛选算法，提高信号保留率与本底拒绝率。
- 优化算法与框架的计算性能，提升系统处理吞吐。
- 增强框架的稳定性与可维护性，推动其面向生产环境运行。