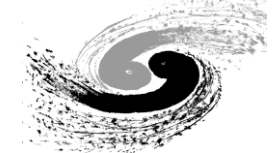


研究生4月考核报告

专业: **计算机应用技术**

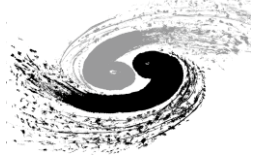
课题组: **触发与数据获取组**

考核人: **张叙**



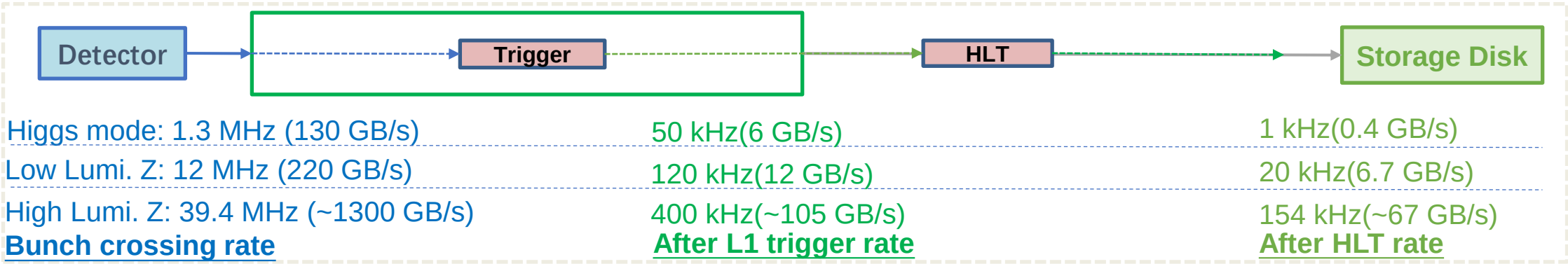
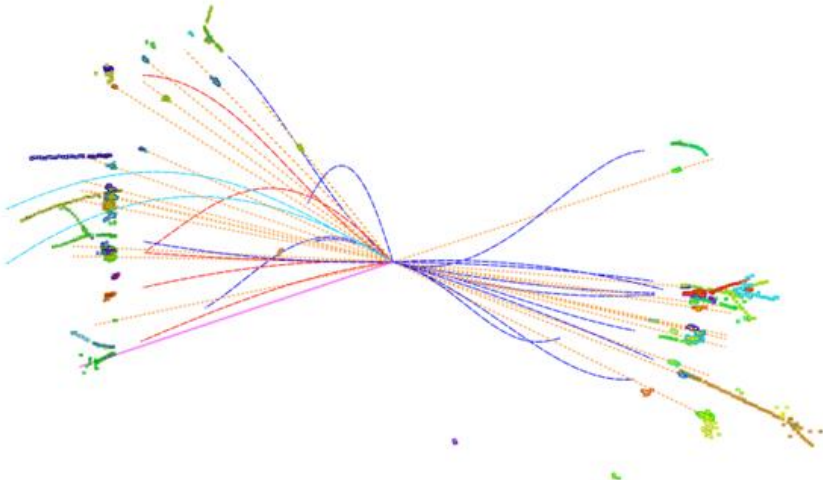
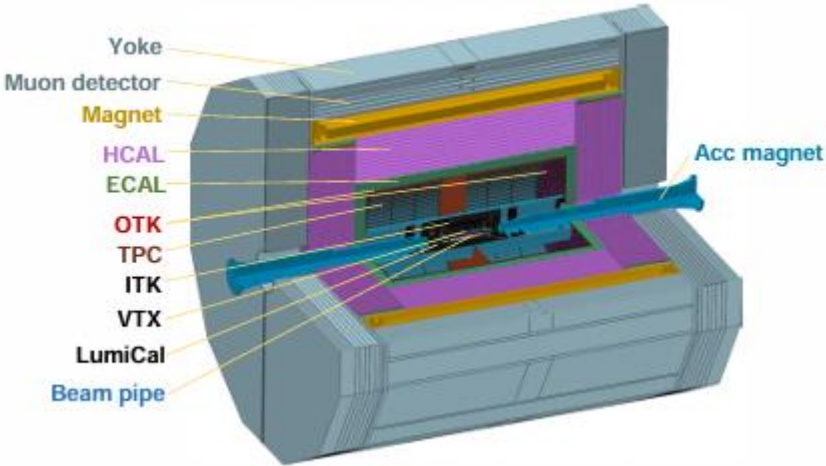
- **CEPC**
 - 基于Track的软件触发算法原型研究
 - 传统算法
 - 霍夫变换(GPU版本)
 - AI算法
 - Transformer
- **JUNO**
 - 小论文: JUNO Online Data Processing Software 的投递(TNS)与修改

基于Track的软件触发算法方案

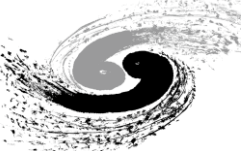


CEPC（环形正负电子对撞机）：旨在通过对撞产生大量Higgs粒子, Z玻色子以解决前沿物理问题。

- 正负电子在对撞中产生大量带电粒子经过探测器形成track(在磁场作用下形成螺旋线)。
- 通过快速重建track, 可以挑选感兴趣的事例，去除掉噪声事例。



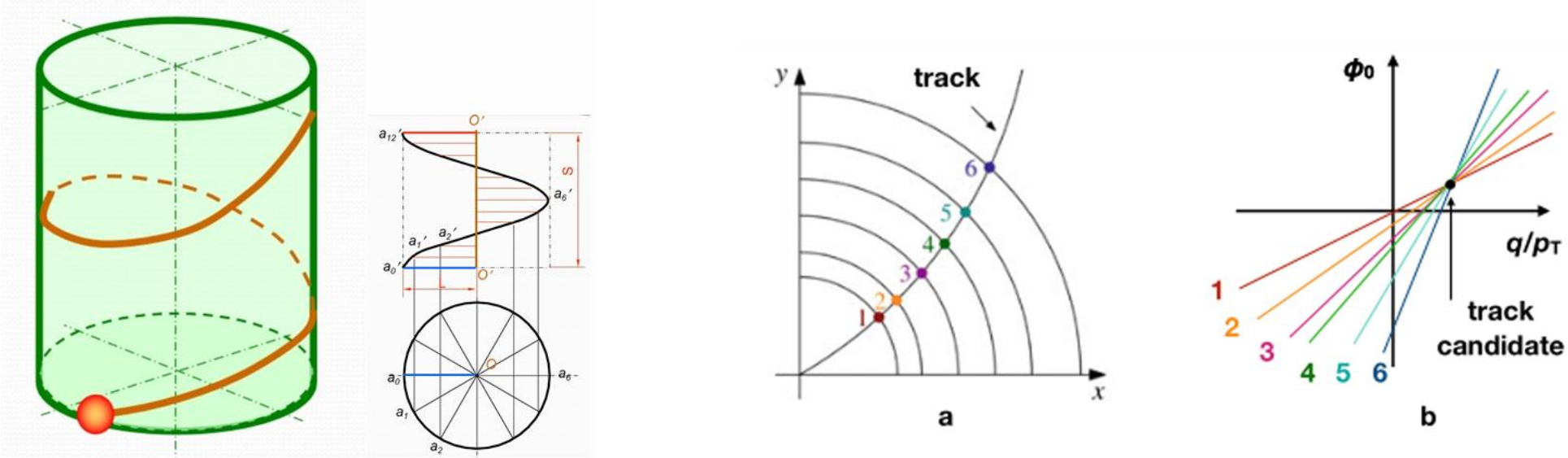
传统方案: 霍夫变换(TDR中的方案)



- **算法原理:** 将复杂的hits曲线拟合问题，转化为**参数空间**的峰值搜索问题。

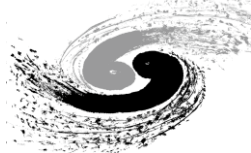
螺旋线方程: $\frac{rk}{2} = \sin(\phi - \phi_0), z = \tan\lambda \cdot s$

- **算法优点:** 算法通用性高，不依赖于业务，使用数学手段寻找是否有若干个hits位于一个**螺旋线**。



- 主要物理过程效率(Provided by BoPing):

Process	vvbb	vvuu	qq	uu	background
Efficiency	99.5%	93.5%	98%	77%	1.2%



- 霍夫变换CPU伪代码：

```
for 每一个点 (x, y, z) in points:
  for 这个点可能对应的每一种参数组合 (a, b, c):
    票箱索引 = 计算参数索引(a, b, c)
    //此accumulator为全局的累加
    accumulator[票箱索引] += 1
```

- 霍夫变换GPU伪代码：

```
一个点 (x, y, z) in points:
  这个点可能对应的一种参数组合 (a, b, c):
    票箱索引 = 计算参数索引(a, b, c)
    //此accumulator为全局的累加
    accumulator[票箱索引] += 1
```

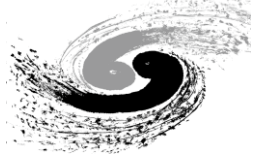
- 数据选取：
VTX ITK OTK hits的坐标

- 将两层for循环在GPU上展开：

	1600 hits	16000 hits	160000 hits	1600000 hits
CPU Intel(R) Xeon(R) Gold 6442Y	2ms	10ms	40ms	300ms
GPU NVIDIA RTX4090	7us	20us	135us	1300us

~7us per bunch crossing

Transformer track可行性讨论:



- 原理契合:

- 自注意力机制的功能之一是特征映射。
- 一条径迹上的点，在**高维空间**中，会被**聚类**成一类。

- 先例验证:

- Transformer在结构更复杂的**喷注分类**任务中已成为主流，证明了其处理HEP数据的卓越能力。
- 相比于喷注的复杂结构，轨迹是更**简单**的曲线模式。

- Transformer的优势:

目前许多Track方案属于局部序列递推

- **卡尔曼滤波**
- **RNN**

Transformer在原理上可以实现全局关联推理(**速度快, 精度高**)。

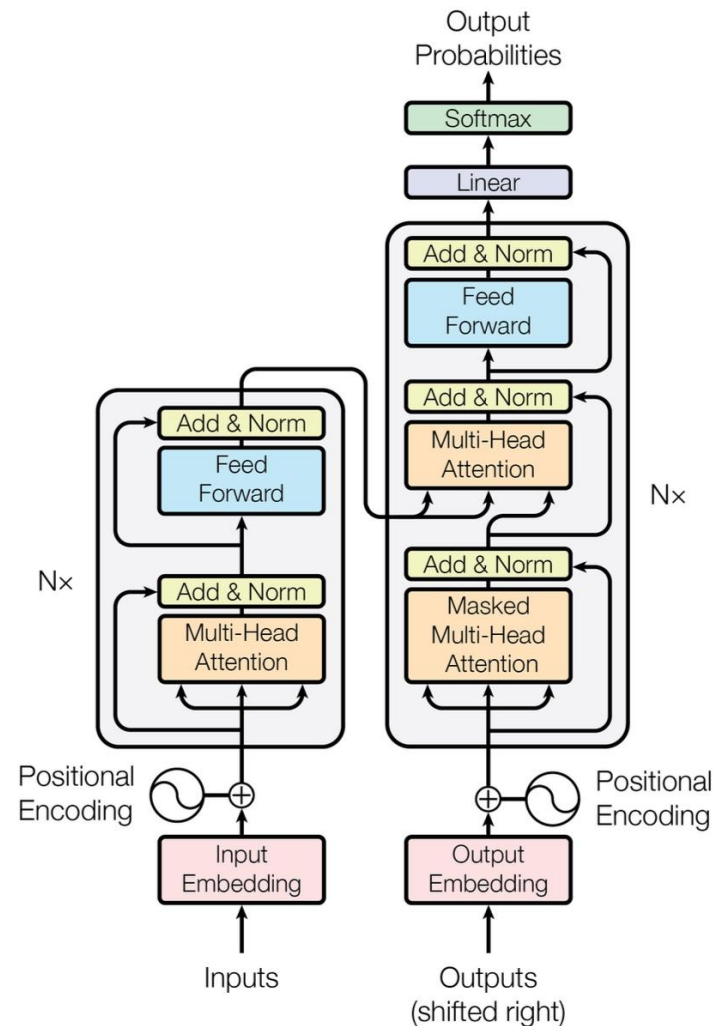
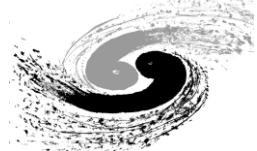
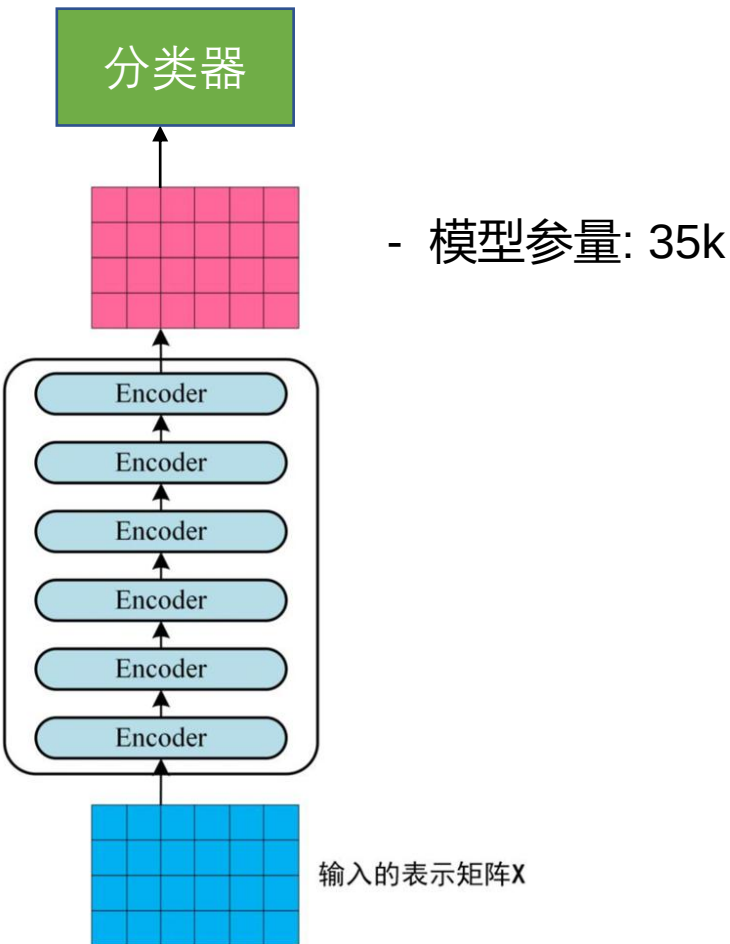


Figure 1: The Transformer - model architecture.

Transformer部署:



- 使用了**encoder**做特征映射后接**分类器**:
- **输入矩阵**的每一行是一个hit的表示(**64维**), 行数是单个event中hit的数量。
- **输出矩阵**的每一行是一个hit经过全局关联映射后的高维向量(**64维**)。

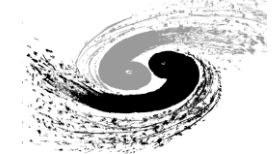


Dataset: VTX ITK OTK中hits的坐标

Process	vvbb + 混合本底	vvuu + 混合本底	qq + 混合本底	uu + 混合本底	background
总事例数:	5000	5000	5000	5000	2000
训练集:	1000	1000	1000	1000	1000
测试集:	5000	5000	5000	5000	2000

只用了20%数据训练 所有数据进行验证
以看泛化性能否足够好

Transformer vs Hough:



- 基于霍夫变换与transformer的触发原型算法物理效率对比

	vbbb	vvuu	qq	uu	background
Hough	99.5%	93.5%	98%	77%	1.2%
Transformer	99.6%	99.2%	98.8%	97.4%	1%

- 基于霍夫变换与transformer的触发原型算法计算效率对比

4090	GPU占用率	Time per event
Hough	17%	7us
Transformer	43%	73us

Transformer在基于track的软件触发方案上有潜力

More testing and validation are needed!!