



中国科学院高能物理研究所
Institute of High Energy Physics Chinese Academy of Sciences



中国科学院大学
University of Chinese Academy of Sciences

BESIII

ML-based PID at BESIII

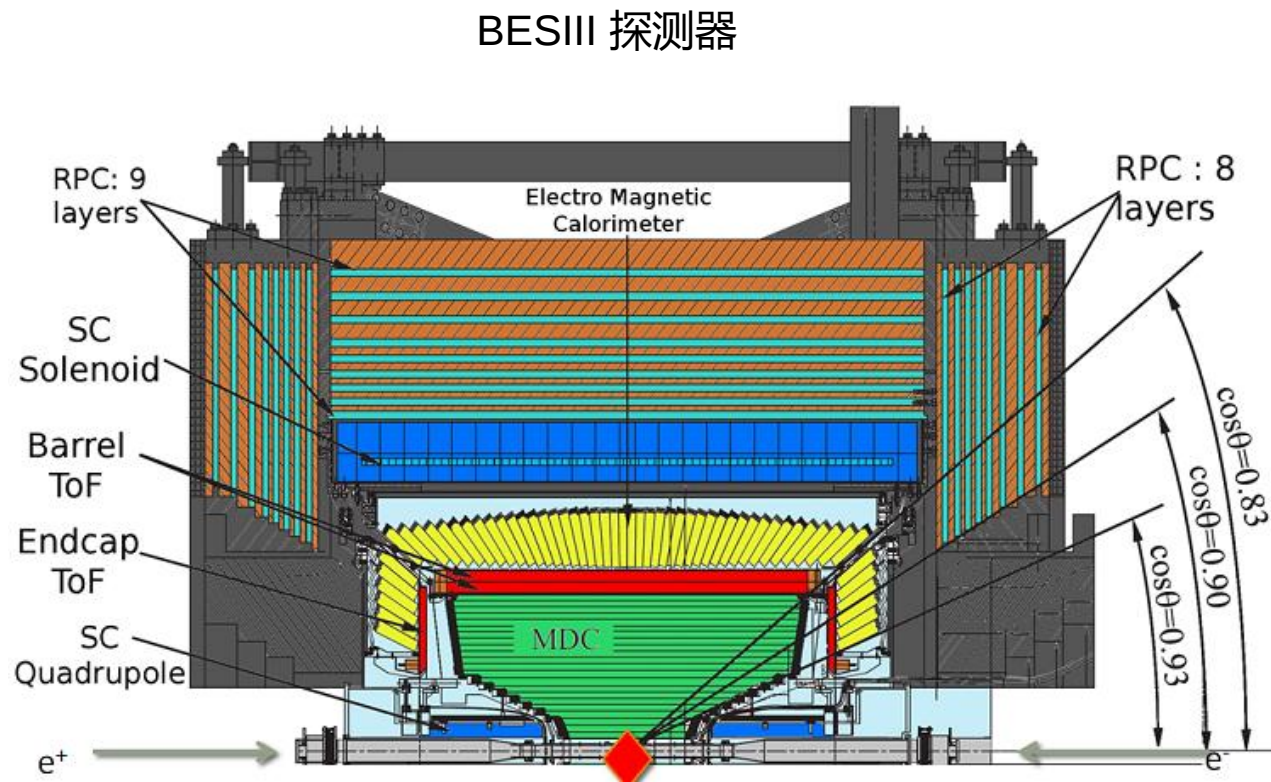
报告人：袁昊
2025超级陶粲装置大会
湘潭 2025.7.4

目录

- BESIII 简介
- 机器学习模型
- 数据获取与预处理
- 模型训练
- 模型性能检查
- 模型部署与物理中的应用
- 总结

BESIII 简介——探测器

- 北京谱仪 (BESIII) 实验是目前国际上唯一运行在 τ – 粲能区的大型粒子物理实验装置，作为高亮度前沿的重要代表，为 τ – 粲能区的物理课题提供了重要的研究平台
- BESIII探测器包含四个子探测器，由对撞点向外依次是主漂移室 (MDC)，飞行时间探测器 (TOF)，量能器 (EMC) 和 μ 子探测器 (MUC)



BESIII 简介——粒子鉴别

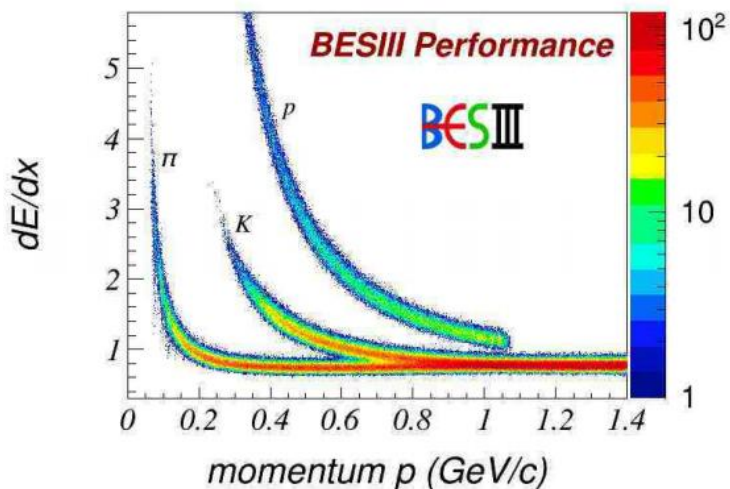
- 粒子鉴别在北京谱仪（BESIII）实验中非常重要
- BESIII实验通过联合dE/dx和飞行时间实现粒子鉴别，没有安装切伦科夫探测器，高动量区域粒子鉴别效率不能充分满足物理需求
- 如何最大程度获取粒子鉴别能力是一个关键科学问题

$$\text{粒子鉴别效率} = \frac{n}{N}$$

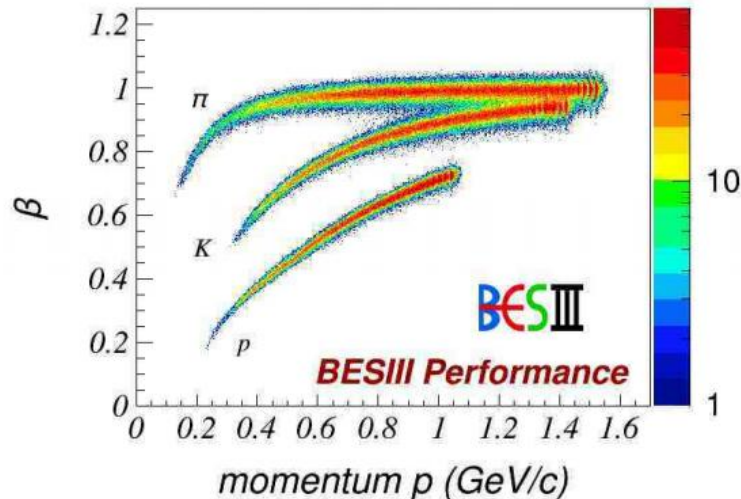
n: 被鉴别正确的强子数目

N: π 介子, K 介子, 质子和反质子的总数目

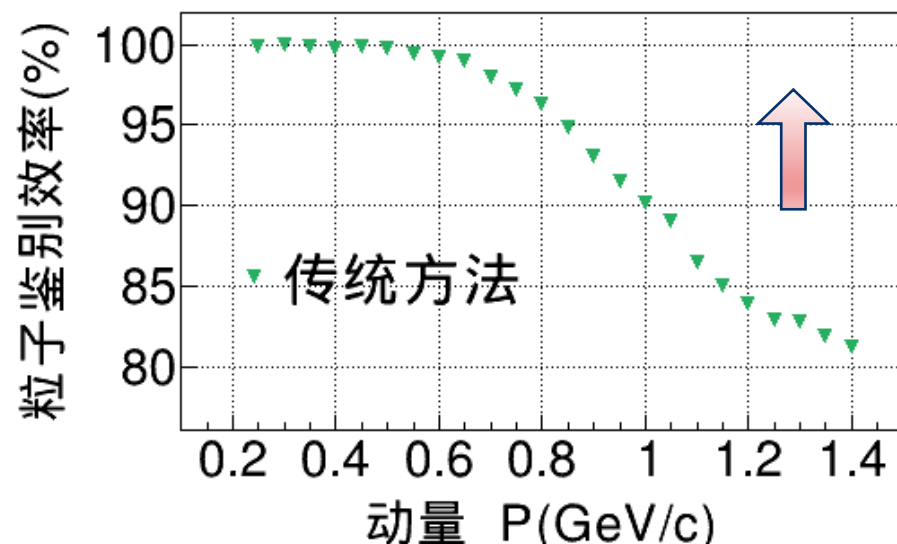
BESIII 主漂移室带电粒子
dedx与动量的分布



BESIII TOF测量的粒子飞行时间
与动量的分布



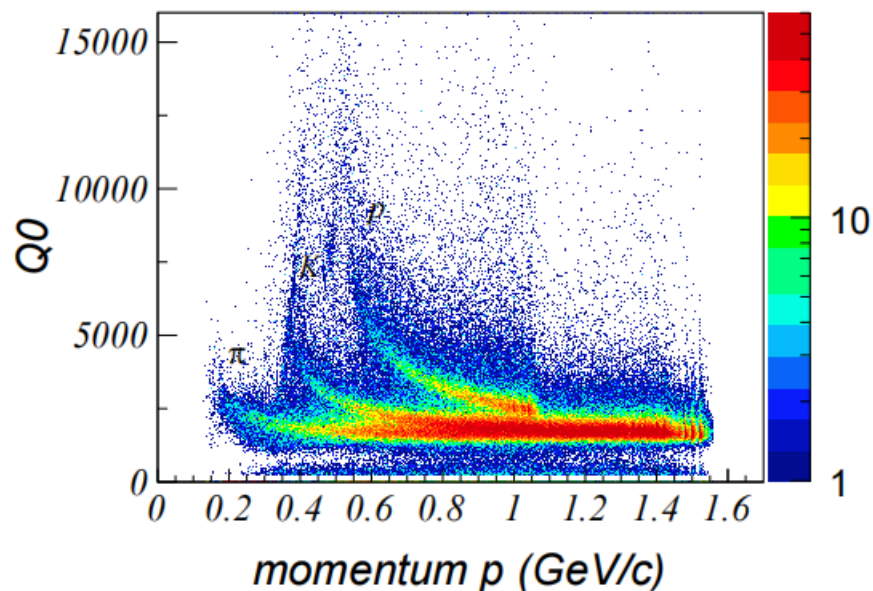
传统方法 π 介子粒子鉴别效率



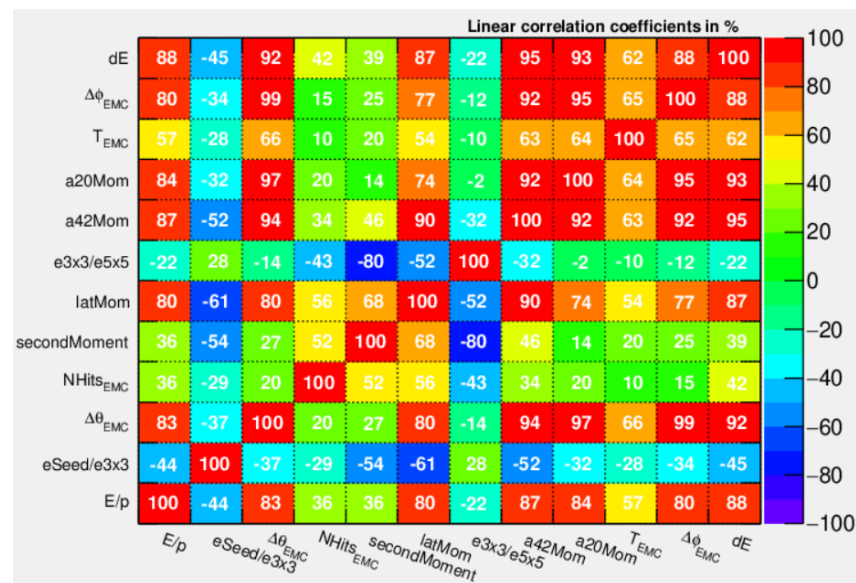
BESIII 简介——粒子鉴别

- BESIII各个子探测器都具有不同程度的粒子鉴别能力
- 机器学习方法适用于解决复杂多变量的关联问题
- 利用机器学习方法进行BESIII实验的快速精准的粒子鉴别方法的研究，发挥探测器最大性能

BESIII TOF脉冲幅度信息随动量的变化



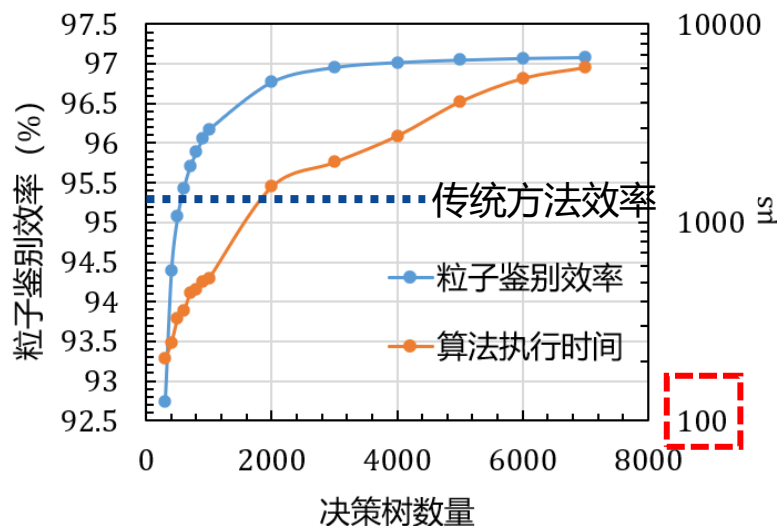
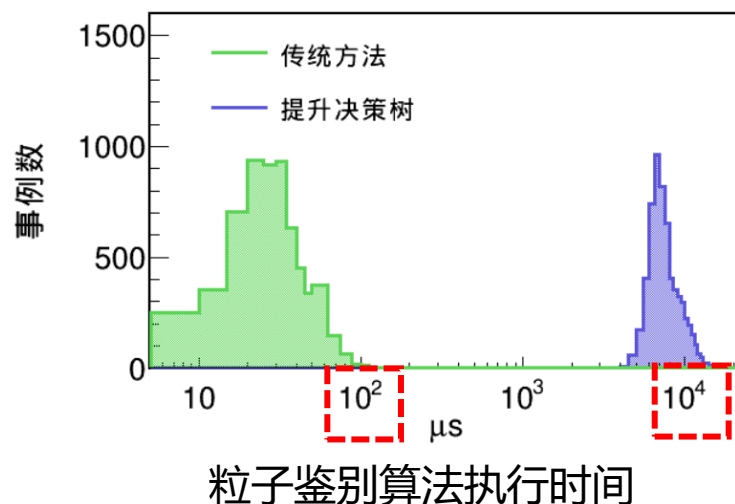
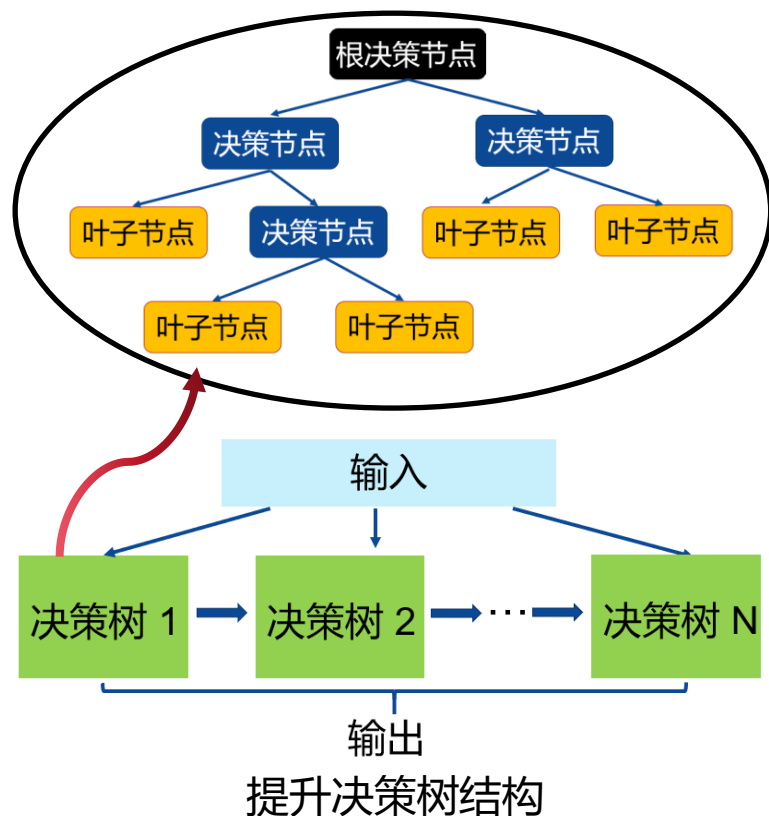
BESIII部分测量信息的关联矩阵



机器学习模型

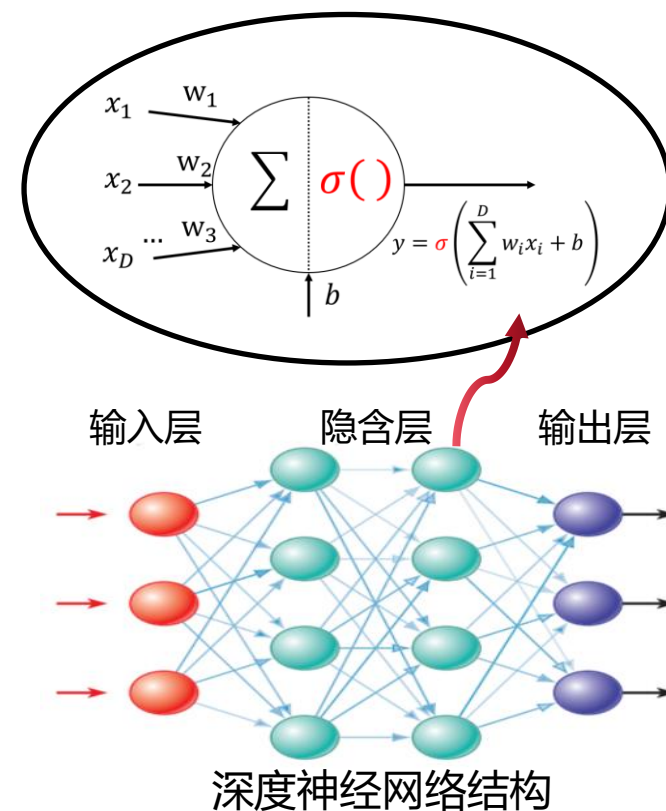
提升决策树 (BDTG) :

- ①粒子鉴别效率得到提升
- ②算法执行时间变长两个数量级



深度神经网络 (DNN) :

- ①缩短算法的执行时间
- ②提升粒子鉴别效率



数据获取与预处理

数据获取:

- 数据来源: J/ψ 真实数据和模拟样本
- Boss版本: 7.0.8
- 物理过程:
 1. $J/\psi \rightarrow \pi^+ \pi^- \pi^0, \pi^0 \rightarrow \gamma\gamma$
 2. $J/\psi \rightarrow K_s^0 K^\pm \pi^\mp, K_s^0 \rightarrow \pi^+ \pi^-$
 3. $J/\psi \rightarrow p \bar{p} \pi^+ \pi^-$

数据预处理:

1. 数据清洗

- 移除异常样本
- 将丢失测量值设为0
- 把动量低于0.4GeV/c粒子的TOF, EMC和MUC的测量信息设为0

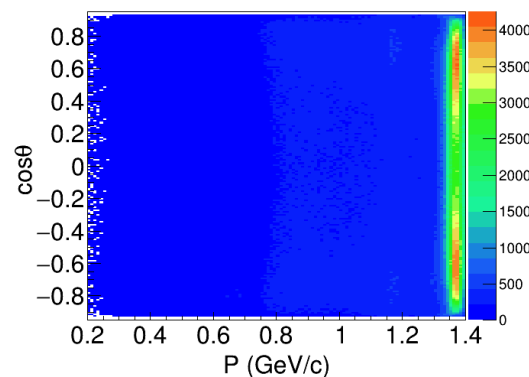
2. 数据归一

- 将动量和 $\cos\theta$ 分布转换成均匀分布

3. 类别平衡

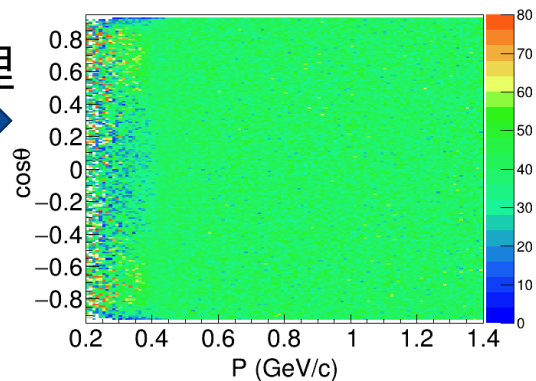
- 平衡三种强子样本的训练数目

预处理前K介子样本动量和 $\cos\theta$ 分布



预处理后K介子样本动量和 $\cos\theta$ 分布

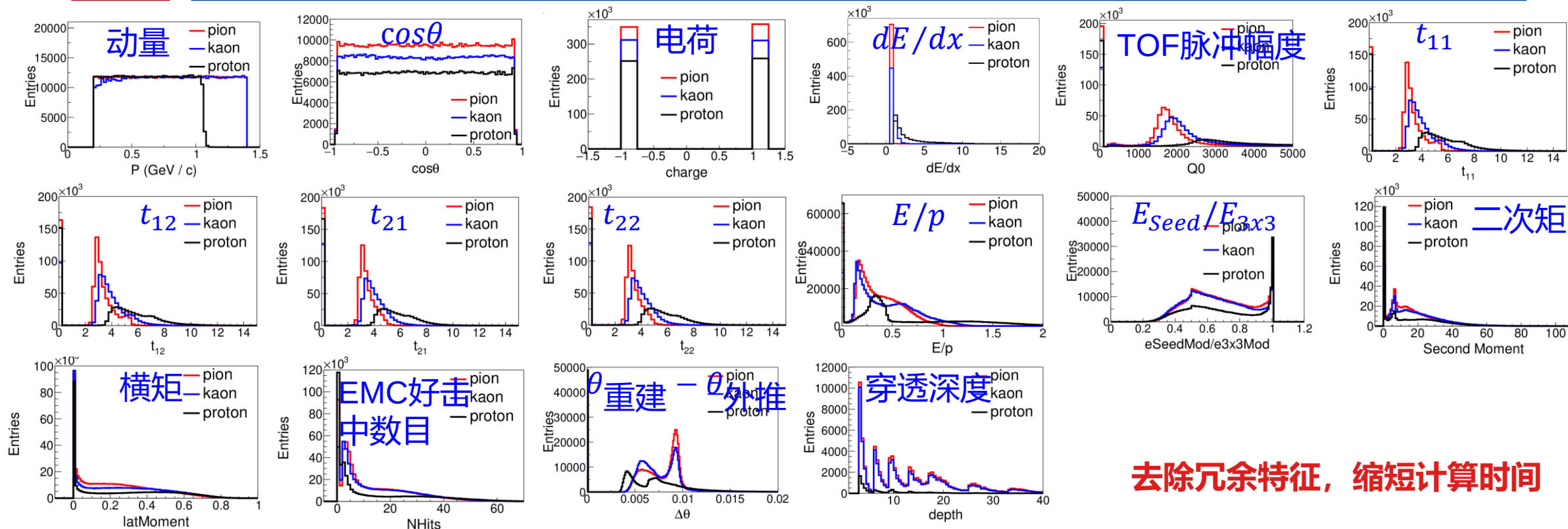
预处理



数据集划分

真实(模拟)	训练集	验证集	测试集
pion	0.57M(0.56M)	0.14M(0.14M)	0.18M(0.17M)
kaon	0.50M(0.51M)	0.12M(0.13M)	0.16M(0.16M)
proton	0.41M(0.39M)	0.10M(0.10M)	0.13M(0.12M)

模型训练——特征挑选



去除冗余特征，缩短计算时间

综合使用三种常用方法：

方法	原理	缺陷
过滤法	根据数据的通用表现去选择，如方差，相关性	割裂了模型与特征之间的联系
包裹法	遍历不同的特征子集组合	计算开销巨大
嵌入法	根据模型训练给出的重要性大小来选择	忽略了特征的物理意义

模型训练——模型超参数

深度神经网络模型

➤ 损失函数：交叉熵 (CrossEntropy)

$$\text{Loss} = -Y * \ln y_i - (1 - Y) \ln(1 - y_i)$$

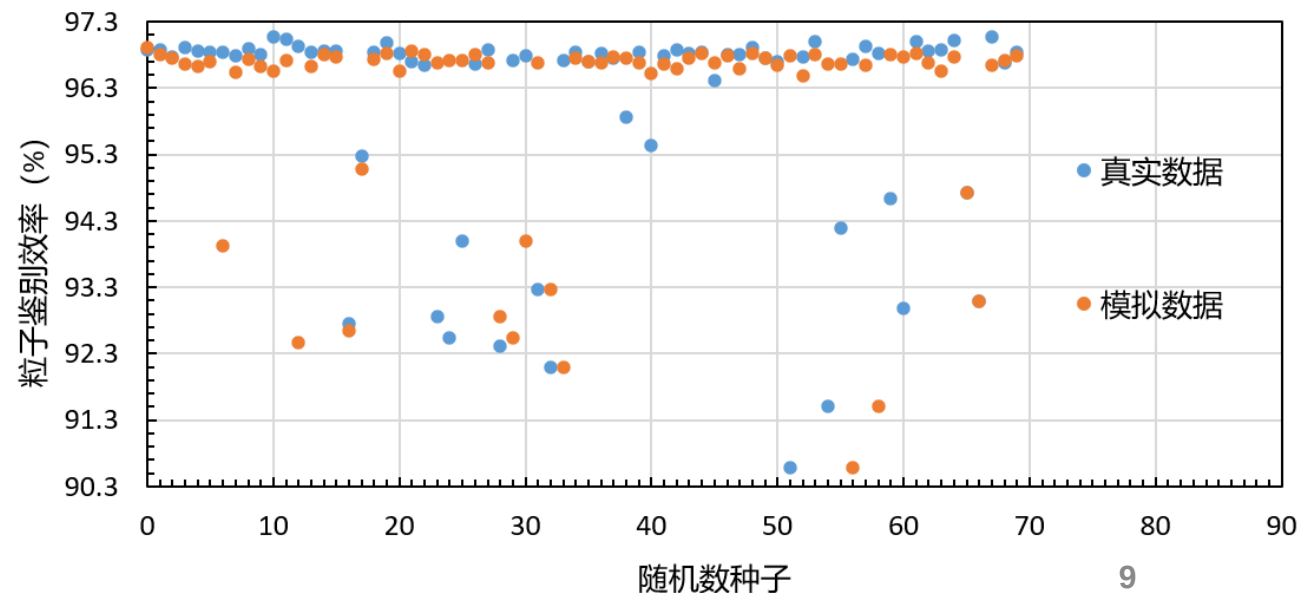
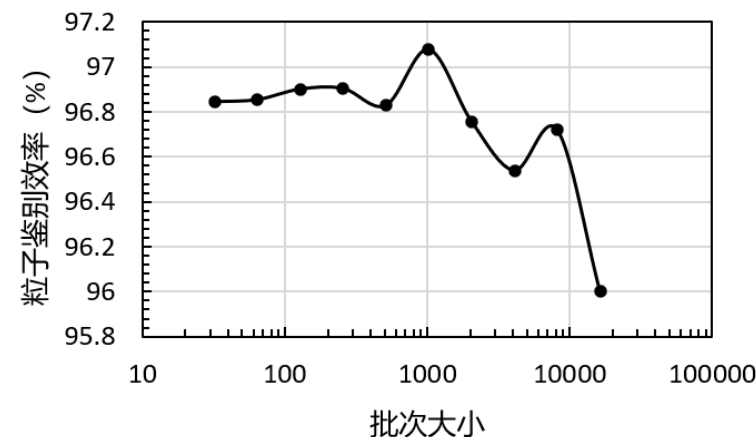
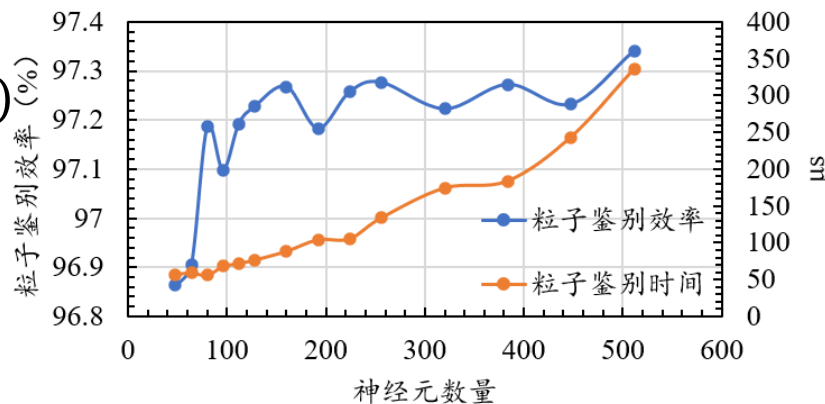
➤ 优化器：自适应矩估计 (ADAM)

➤ 网络架构

- 输入层：16个变量
- 隐含层：TANH|20,TANH|40,TANH|20
- 输出层：3 ($\text{Softmax}(y_i) = e^{y_i} / \sum_j e^{y_j}$)

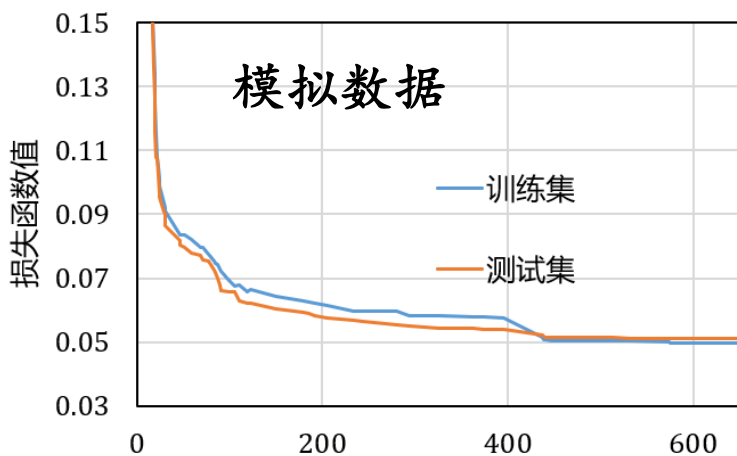
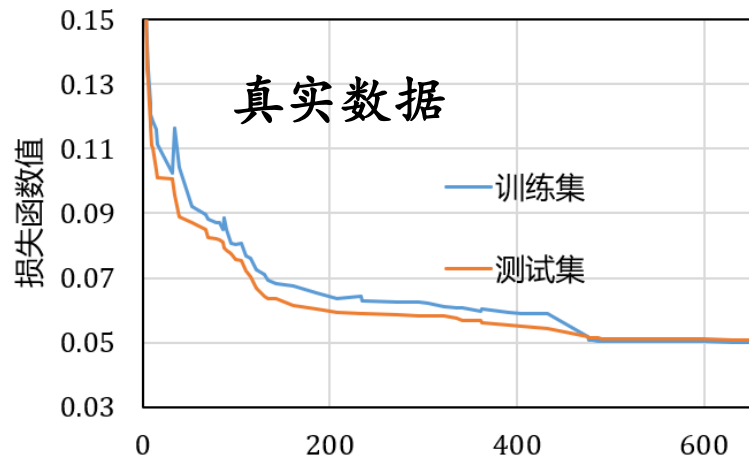
➤ 训练策略:

- 第一阶段：学习率= 10^{-2} ，批次大小=1024
- 第二阶段：学习率= 10^{-3} ，批次大小=1024
- 第三阶段：学习率= 10^{-4} ，批次大小=1024



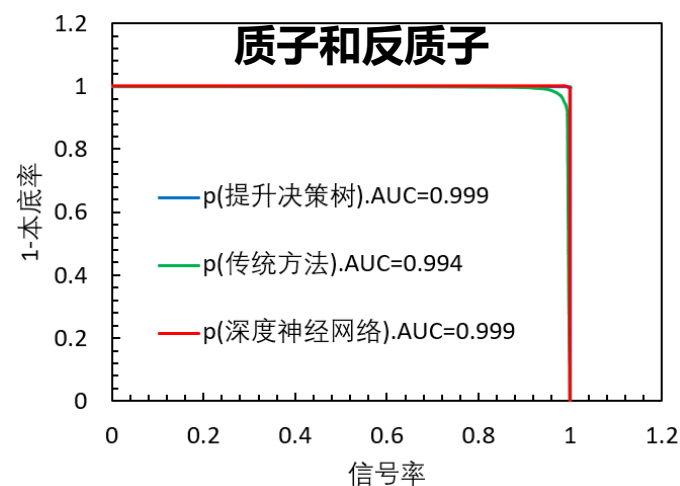
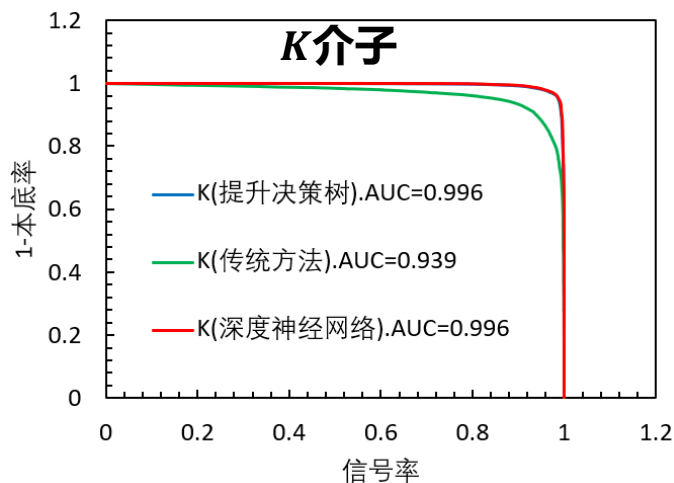
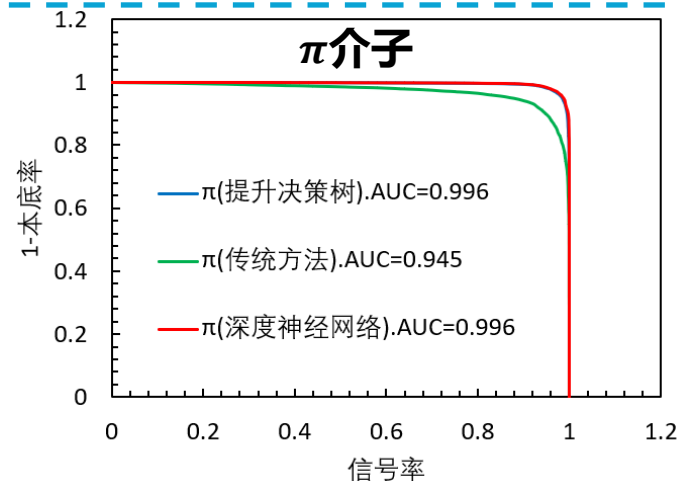
模型训练——损失函数与ROC曲线

训练过程中训练集与验证集损失函数值随训练轮次变换曲线



受试者工作特征 (ROC) 曲线

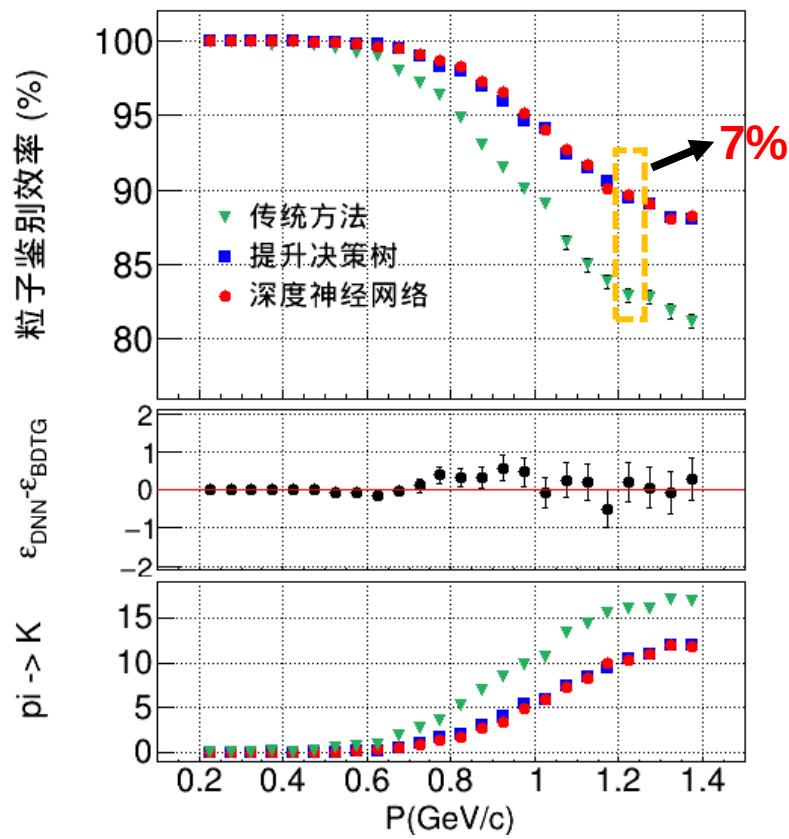
➤ 深度神经网络和提升决策树方法的性能大幅优于传统方法



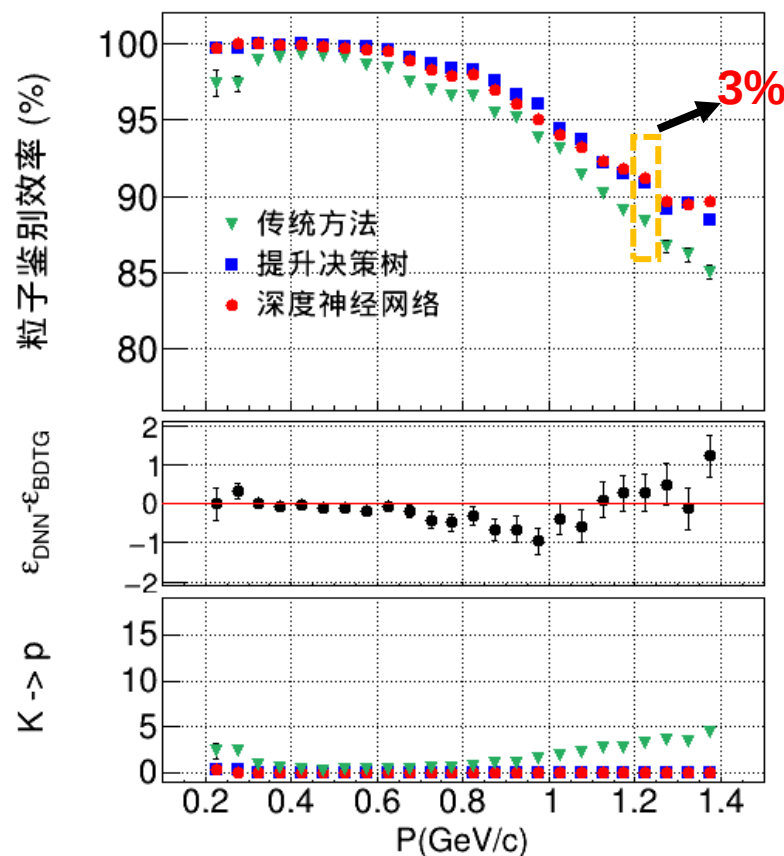
模型性能检查——粒子鉴别效率

- 深度神经网络模型的性能大幅优于传统方法，高动量区域粒子鉴别效率显著提升
- 深度神经网络模型的性能与提升决策树模型基本一致

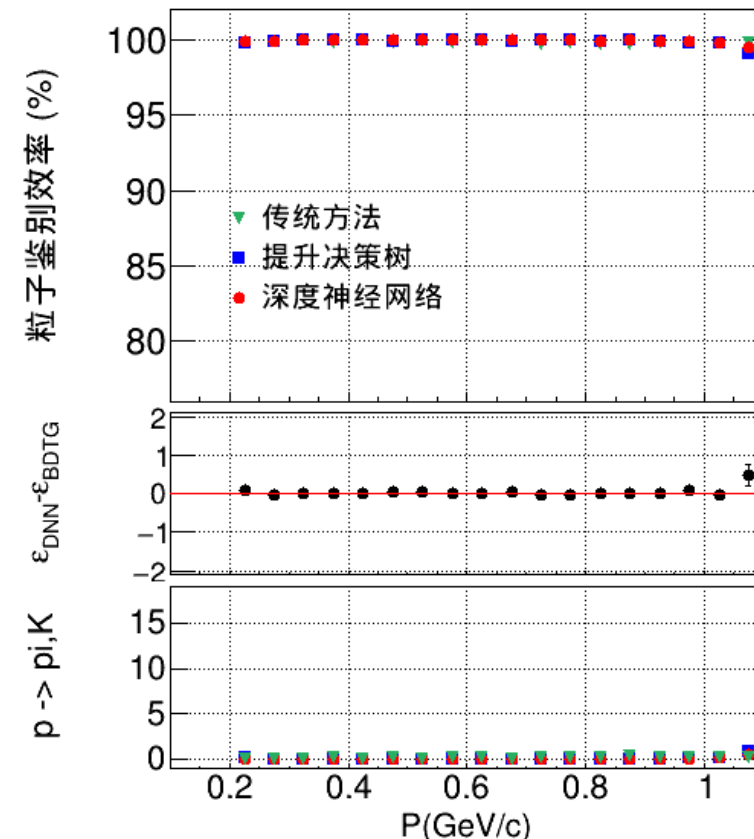
真实数据



π 介子



K 介子

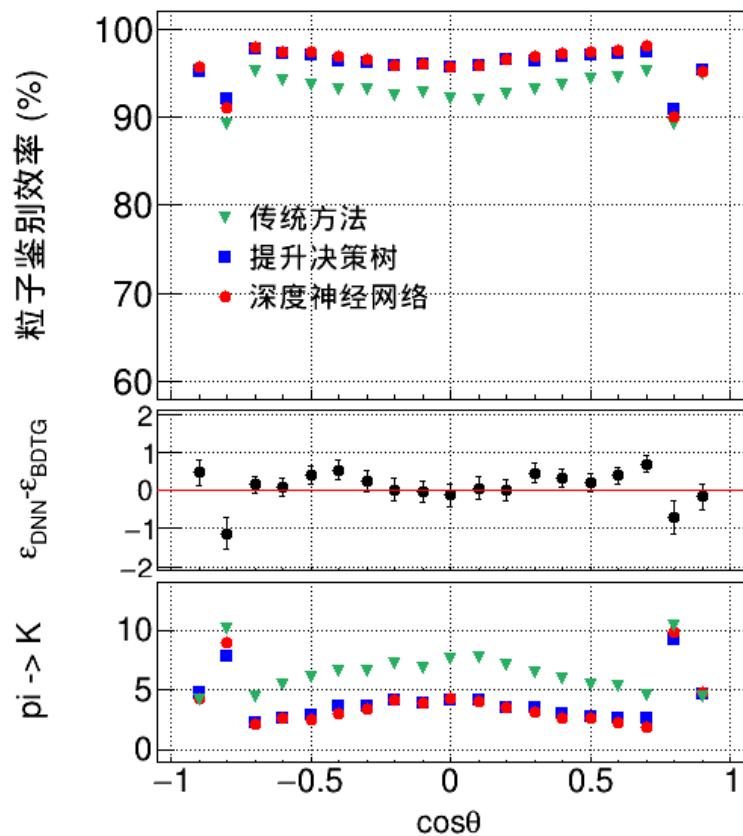


质子和反质子

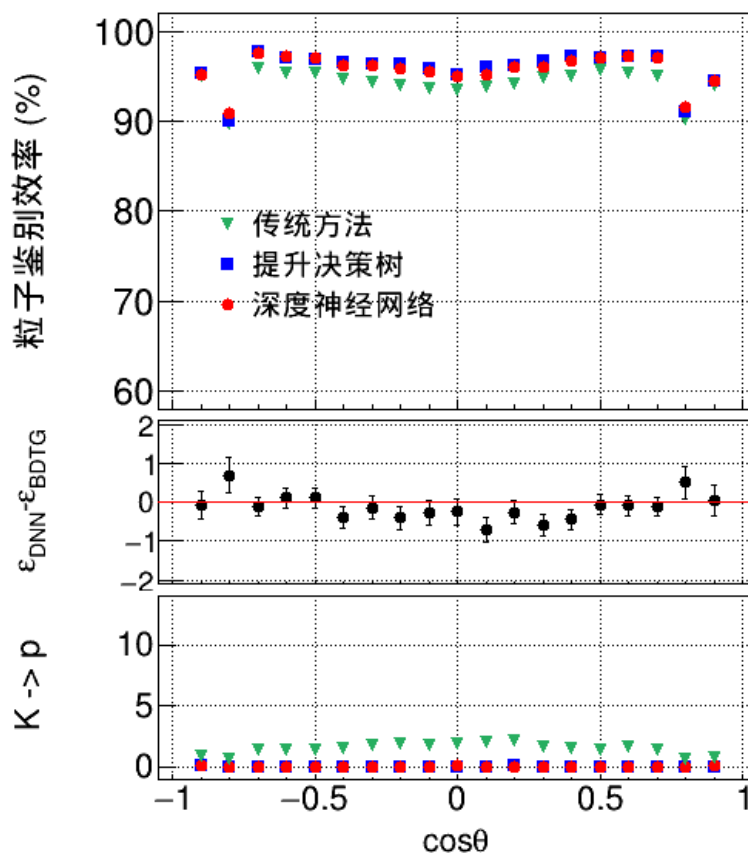
模型性能检查——粒子鉴别效率

- 深度神经网络模型的性能大幅优于传统方法，高动量区域粒子鉴别效率显著提升
- 深度神经网络模型的性能与提升决策树模型基本一致

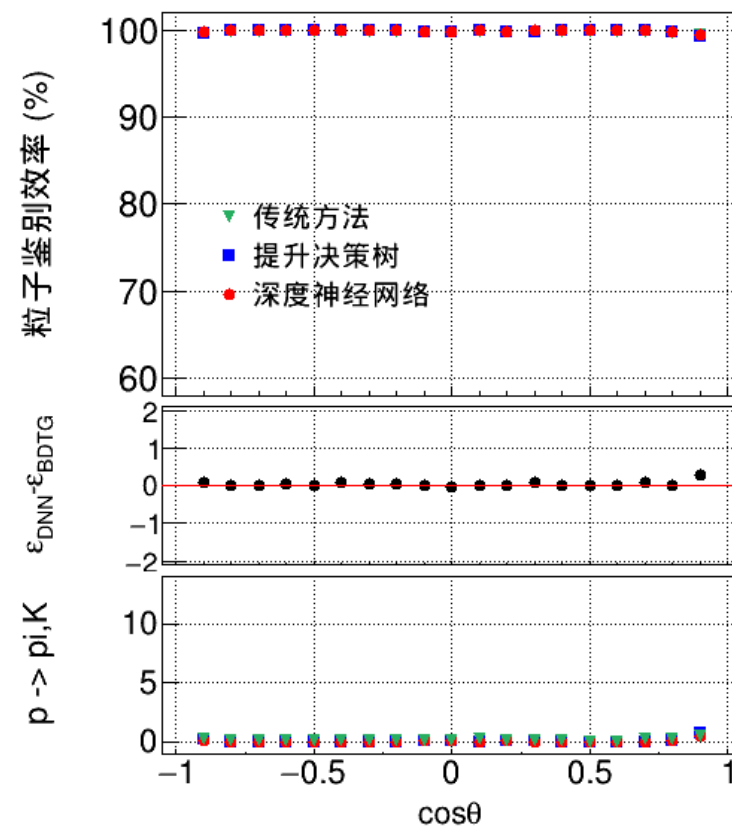
真实数据



π 介子



K 介子



质子和反质子

模型性能检查——粒子鉴别效率

真实数据

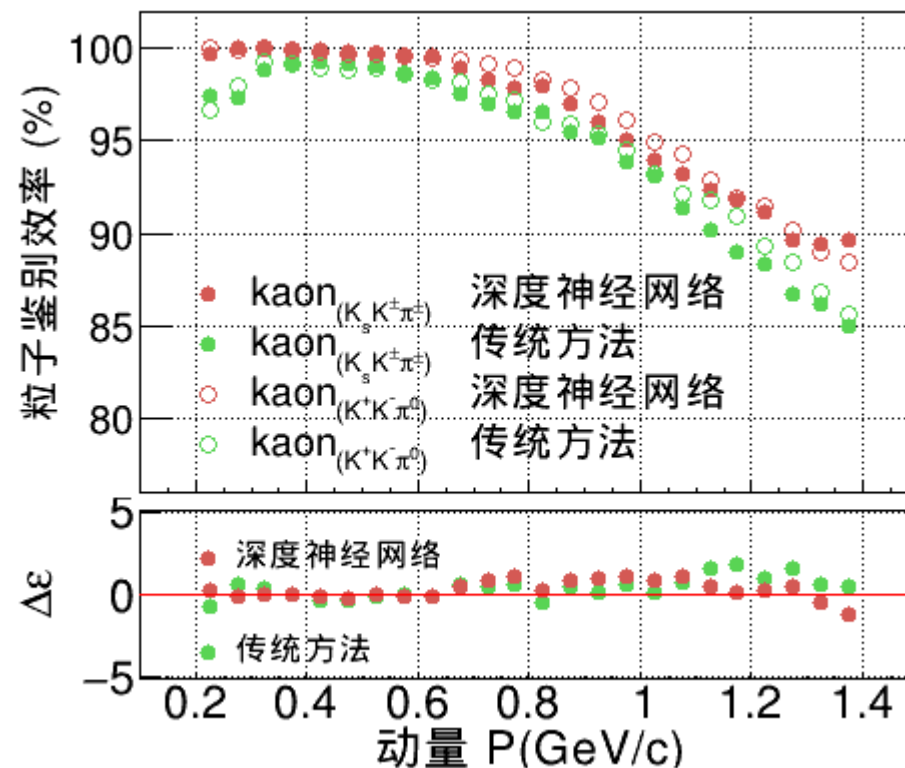
交叉验证:

Round11

BOSS708

物理过程: $J/\psi \rightarrow K^+ K^- \pi^0$

样本数: 0.82M



模型的性能对物理
过程没有依赖

模型性能检查——粒子鉴别算法执行时间

优化深度神经网络模型计算算法:

- 深度神经网络中包含大量矩阵的乘法运算
- 使用TMatrix类代替for循环来处理矩阵运算

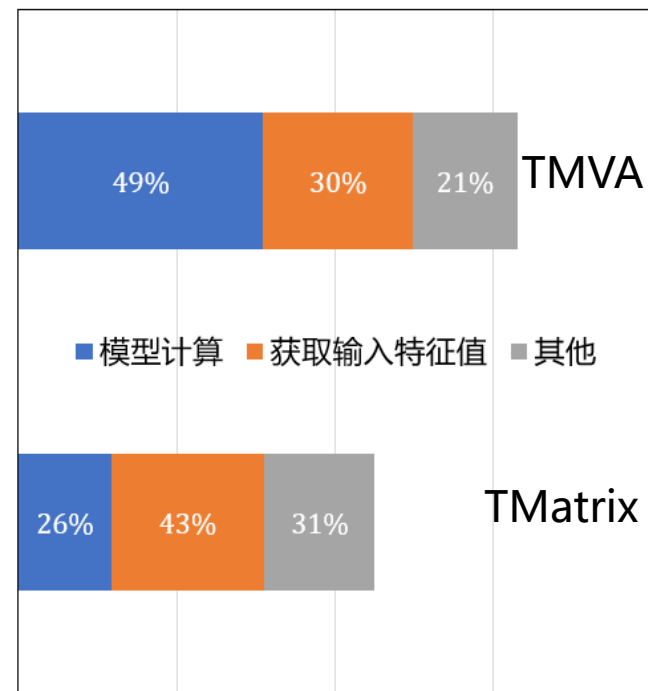
- 模型计算时间减少~50%
- 粒子鉴别算法执行时间减少~30%

TMVA

```
for i = 1:I
  for j = 1:J do
     $a_i^l = w_{i,j}^l * a_j^{l-1}$ 
     $a_i^l = \sigma(a_i^l + b_i^l)$ 
```

TMatrix

```
TMatrixD  $a^{l-1}, a^l, w^l, b^l$ 
 $a^l = \sigma(w^l * a^{l-1} + b^l)$ 
```



算法示意

深度神经网络算法各部分时间占比

模型性能检查——粒子鉴别算法执行时间

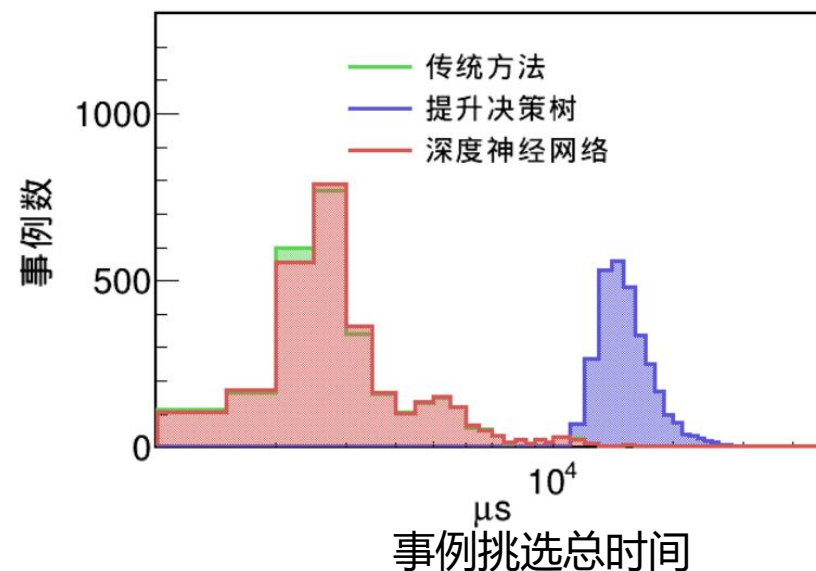
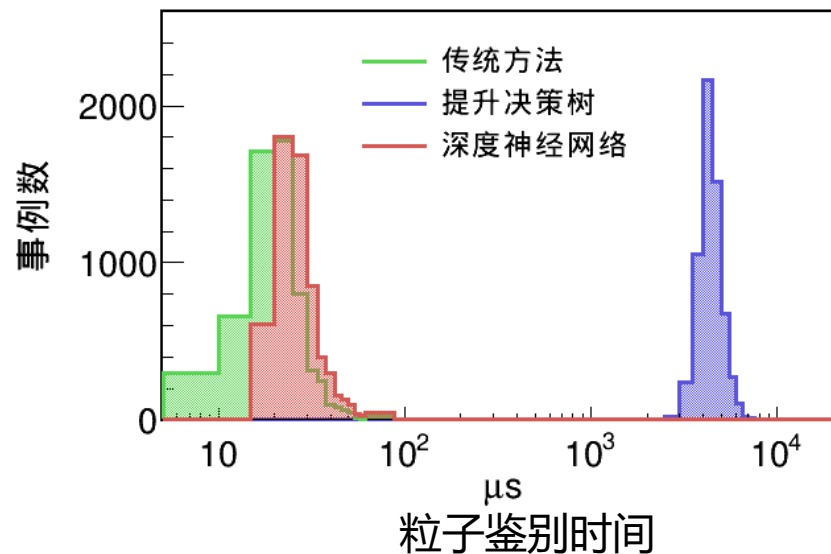
物理过程: $J/\psi \rightarrow \pi^+ \pi^- \pi^0$

事例挑选条件:

- 好带电径迹挑选
- 好光子挑选
- **粒子鉴别**
- 顶点拟合和运动学拟合

深度神经网络方法

- 粒子鉴别算法时间:
 - 仅是传统算法的~1.3倍
 - 比提升决策树快两个数量级
- 事例挑选总时间:
 - 与传统算法基本一致
 - 比提升决策树快~5倍

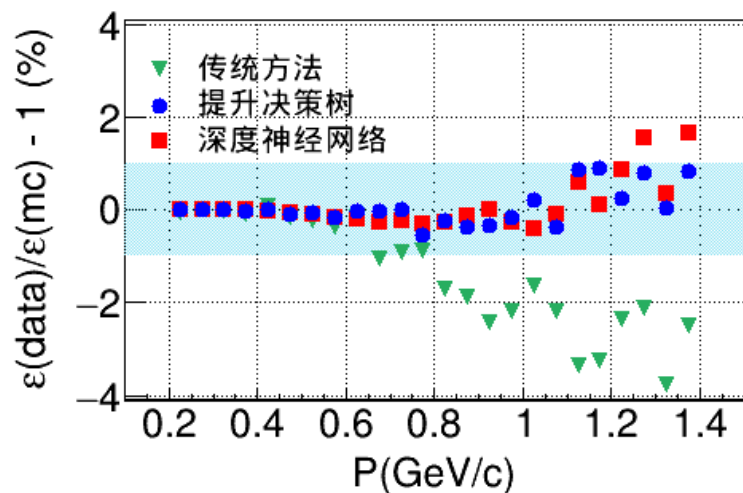


模型性能检查——系统误差

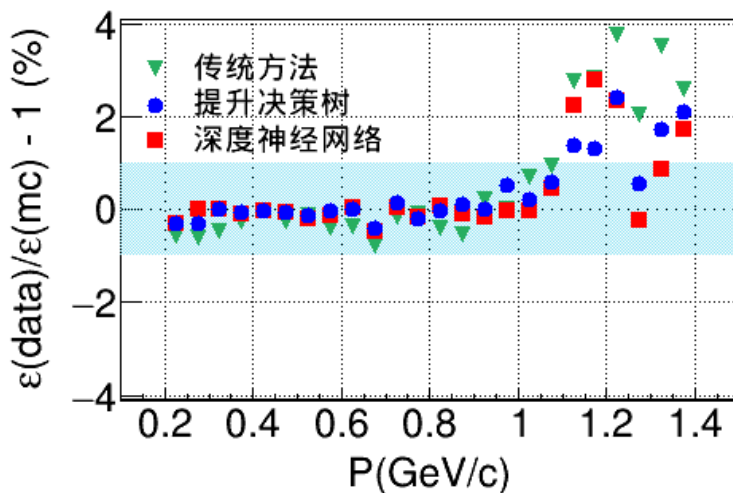
➤ 系统误差: $\Delta\varepsilon = \frac{\varepsilon(data) - \varepsilon(MC)}{\varepsilon(MC)}$

• ε : 粒子鉴别效率

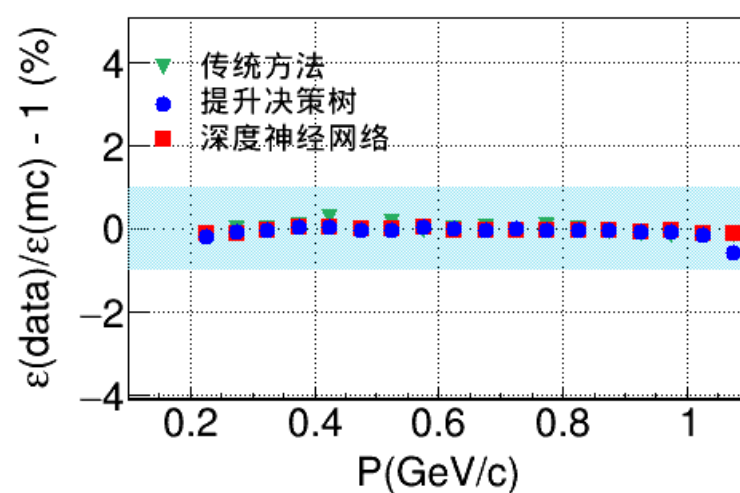
π 介子



K介子



质子和反质子

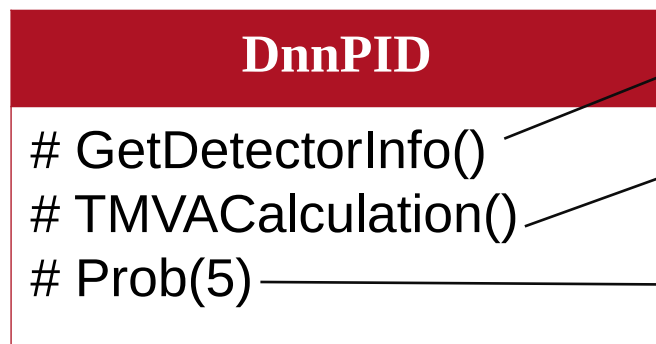


➤ 深度神经网络的系统误差:

- π 介子和K介子鉴别的系统误差总体小于1%
- 质子鉴别的系统误差总体小于0.2%

模型部署与物理中的应用——模型部署

深度神经网络模型在BESIII软件算法中的实现



读取重建信息

深度神经网络模型计算三种粒子
(π , K, p) 假设的概率

返回 (π , K, p) 假设的概率

深度神经网络模型的使用

```
ParticleID *pid = ParticleID::instance();  
for(int i = 0; i < nGood; i++) {  
    EvtRecTrackIterator itTrk = evtRecTrkCol->begin() + iGood[i];  
    pid->init();  
    ...  
    pid->usePidSys(pid->useDedx() | pid->useTofCorr()); // use PID sub-system  
    // pid->usePidSys(pid->useDnn()); // use Deep Neural Networks model  
    pid->identity(pid->onlyPionKaonProton()); // seperater Pion/Kaon/Proton  
    ...  
    pid->calculate();  
    ...  
    if((pid->probPion() < pid->probProton()) || (pid->probPion() < pid->probKaon())) continue;  
    ...  
}
```

Rhopi.cxx

模型部署与物理中的应用——应用测试

round11 J/ψ incMC

BOSS710

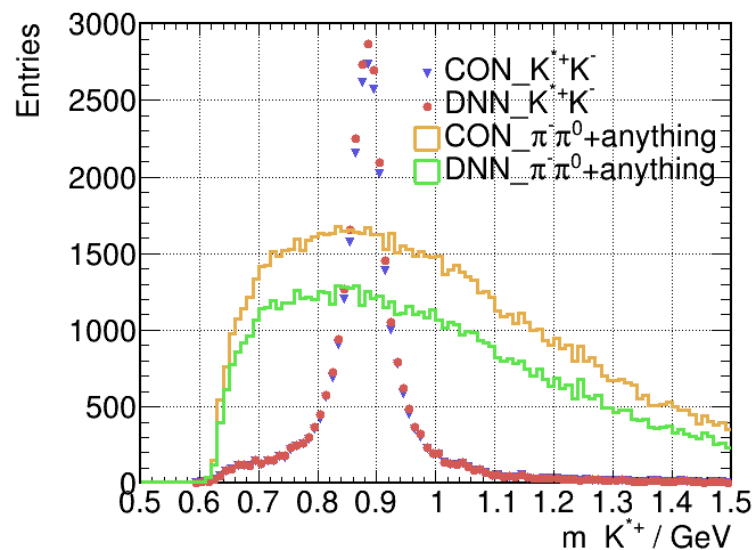
信号: $K^- K^{*+}$

本底: $\pi^- \pi^0 + anything$

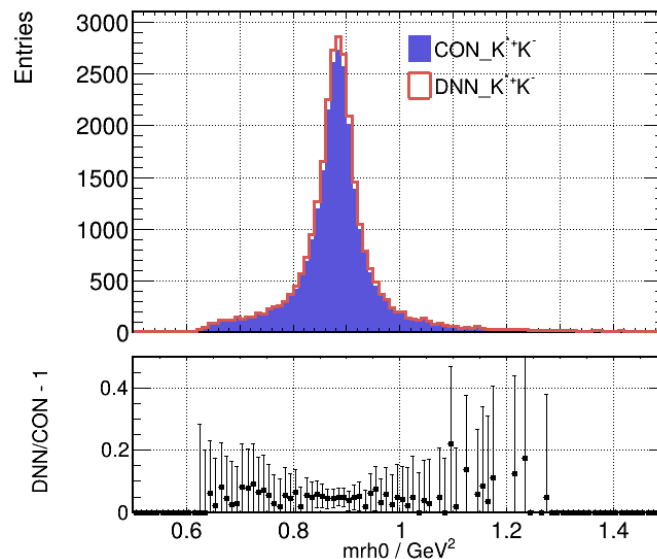
PID: $prob_{K^-} > prob_{\pi, P}$

方法	PID前事例总数	PID ($0.5 < m_{dcp} < 1.5$) π_0 Cut	显著度($s/\sqrt{s+b}$) ($0.5 < m_{\rho^0} < 1.2$)
传统方法	$K^{*+} K^-$: 37361	27429 (73.42%)	77.98 ↓ ~16% 90.54
	$\pi^- \pi^0 + anything$: 3458150	96284 (2.78%)	
深度神经网络	$K^{*+} K^-$: 37361	28704 (76.83%)	90.54
	$\pi^- \pi^0 + anything$: 3458150	71802 (2.08%)	

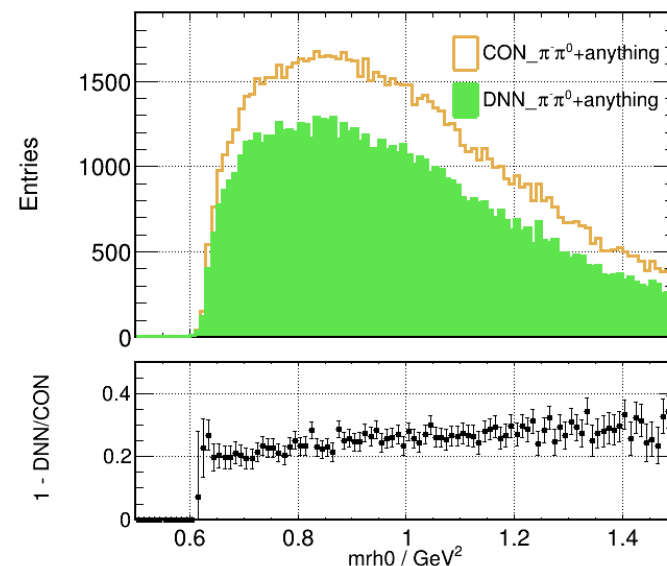
PID 后 $m_{K^{*+}}$ 分布



PID 后信号 $m_{K^{*+}}$ 分布



PID 后本底 $m_{K^{*+}}$ 分布



总结

- 利用机器学习方法高效联合BESIII实验四个子探测器的强子鉴别信息，开发了新的**快速精确**的粒子鉴别算法，可有效提升信号显著度和物理精度
- 基于**深度神经网络**的粒子鉴别算法与传统方法相比：

- ❑ **粒子鉴别效率：显著提高**

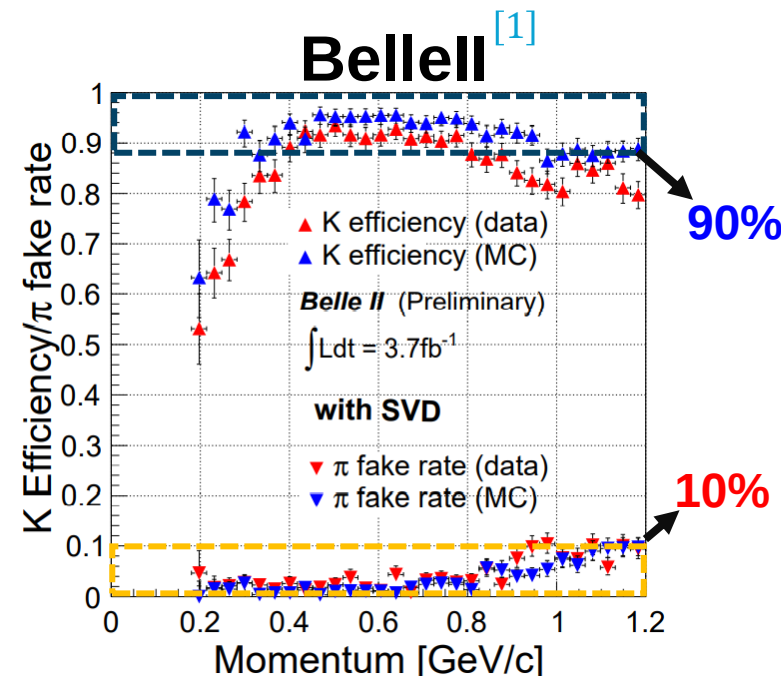
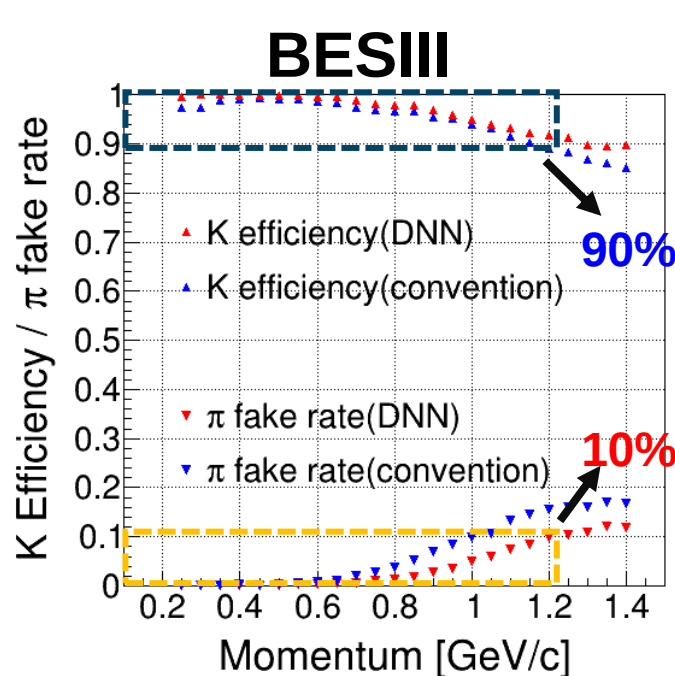
- π 介子提升 $\sim 7\%$ @1.2GeV/c
- K 介子提升 $\sim 3\%$ @1.2GeV/c

- ❑ **算法执行时间：非常接近**

- 粒子鉴别算法时间 ~ 1.3 倍
- 事例挑选总时间**基本一致**

- ❑ **系统误差：较大改善**

- π 介子和 K 介子总体小于 1%

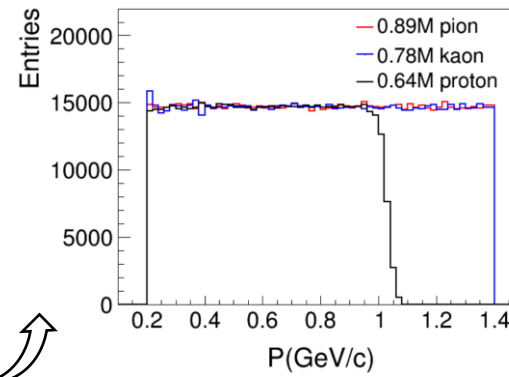
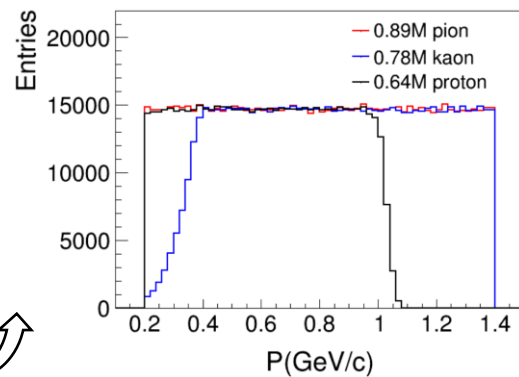
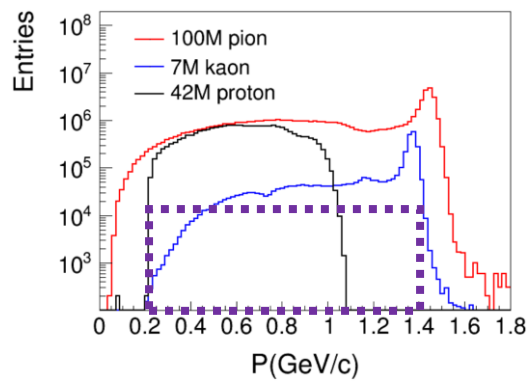


谢谢大家!

参考文献

- [1] S. Hazra, et al., Particle Identification in Belle II Silicon Vertex Detector, JPS Conf. Proc. 34 (2021) 01001

数据预处理



- K 介子的特殊处理

