



STCF fast simulation framework

Xiaoxia Liang, Xiaoshuai Qin
(On behalf of the fast simulation working group)
Shandong University

2026年超级陶粲装置研讨会, 西安
July.1st. 2026

目录

CONTENTS

1. 超级陶浆装置 (STCF) 简介
2. 快速模拟框架介绍
3. 快速模拟与全模拟性能对比
4. 讨论
5. 总结

超级陶粲装置 (STCF)

◆海量数据统计:

- 数据样本量大约是BESIII的100倍
- 原始数据处理链 (~110 PB/year)
- MC数据 (与原始数据规模相当)

◆多种物理课题:

- 粲强子, τ 物理
- CP 破坏: 超子, K^0 系统等
- 轻强子谱学, XYZ 粒子
- 基本物理量精确测量
- Collins效应

◆本底:

- 对于 1ab^{-1} 的数据量, 将有 $O(10^{11})$ 的本底事件

◆巨大的 CPU/存储消耗

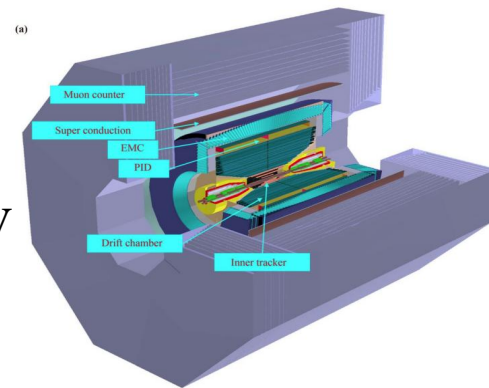
基本设计:

峰值亮度:

$$0.5 \times 10^{35} \text{cm}^{-2} \text{s}^{-1} @ E_{\text{cm}} = 4 \text{GeV}$$

年积分亮度: 1.4ab^{-1}

能量范围: $E_{\text{cm}} = 2-7 \text{GeV}$



超级陶粲装置中的快模拟

◆物理模拟的目标

- 物理可行性研究（物理灵敏度、本底等）
- 优化探测器设计（效率、分辨率等）
- 与过去/现有/未来实验的比较

◆模拟工具

- 全模拟（用于信号模拟以及作为快速模拟的标准）
- 快速模拟（用于本底研究、灵敏度估计等。）

◆快速模拟的特点：

- 速度快且存储容量小（inclusive MC 样本）
- 灵活应对探测器响应/参数变化
- 对用户方便友好

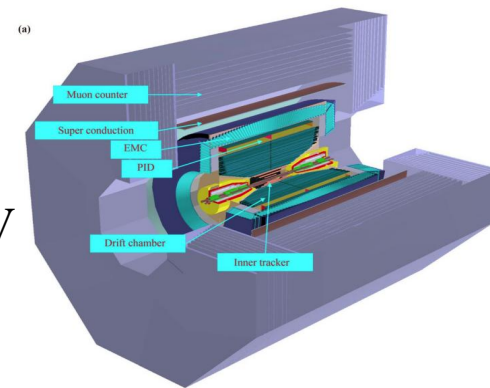
基本设计：

峰值亮度：

$$0.5 \times 10^{35} \text{cm}^{-2} \text{s}^{-1} @ E_{\text{cm}} = 4 \text{GeV}$$

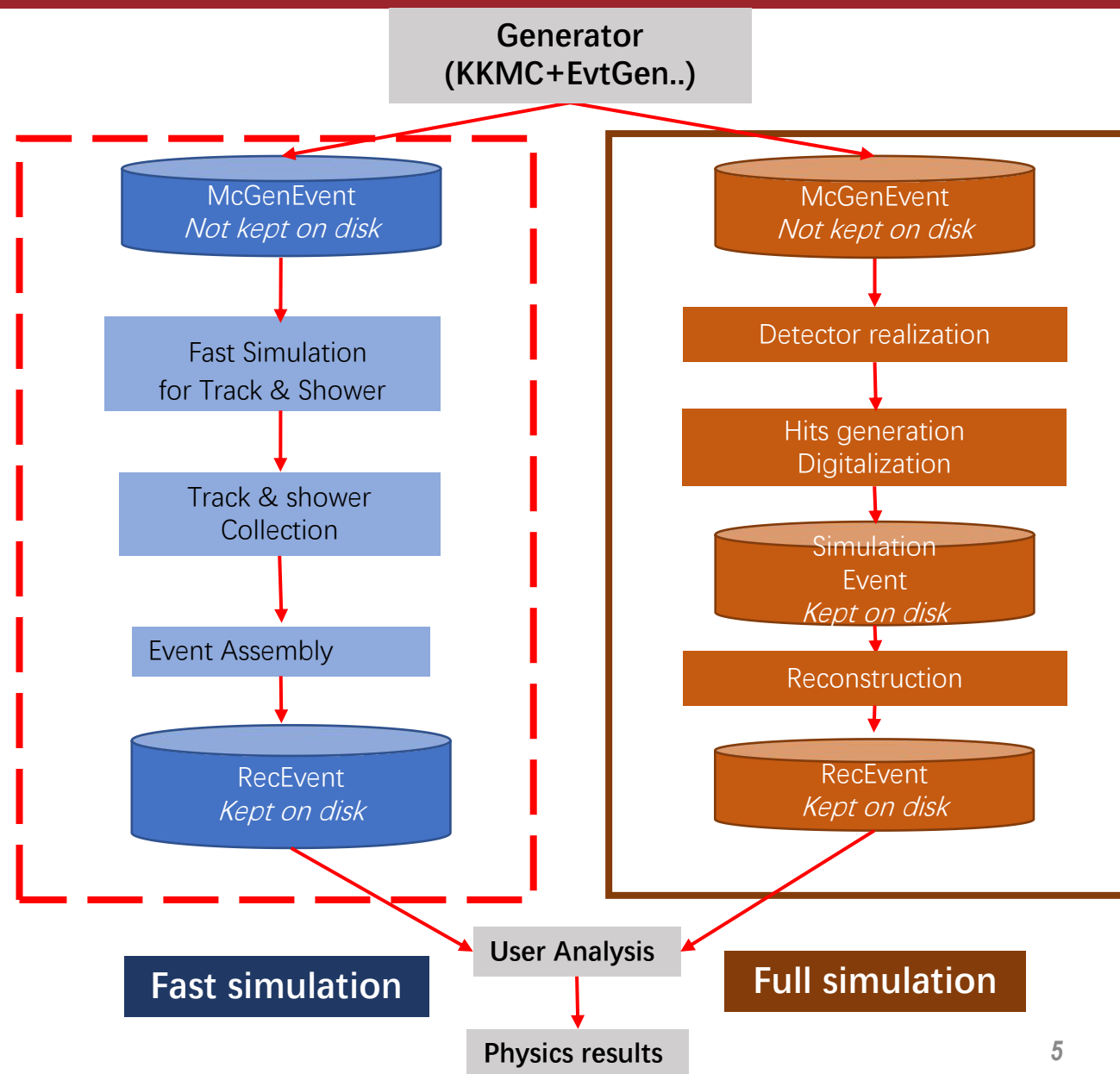
年积分亮度： 1.4ab^{-1}

能量范围： $E_{\text{cm}} = 2 - 7 \text{GeV}$



快模拟方案

- ◆产生器：将 McGenEvent 保存在内存中
- ◆带电和中性粒子的快速模拟
 - 分辨率、效率等
- ◆随机种子管理，用于重复分析
- ◆分析流程与全重建相同
- ◆应用此方案以减少 CPU 和存储消耗



快模拟与全重建部分信息处理的差别

- ◆实现了简单的PID 判定
 - 没有来自子探测器的详细测量值，如 dE/dx 、RICH 似然值、MUD 似然值
- ◆实现了 ECAL 沉积簇射能量以及MUD部分信息
- ◆没有详细的探测器几何结构，没有基于 Geant4 的粒子输运
 - 没有hits信息
 - 没有decaying in flight的信息
 - 没有gamma conversion
 - 没有精确的外推或 ExtTrack

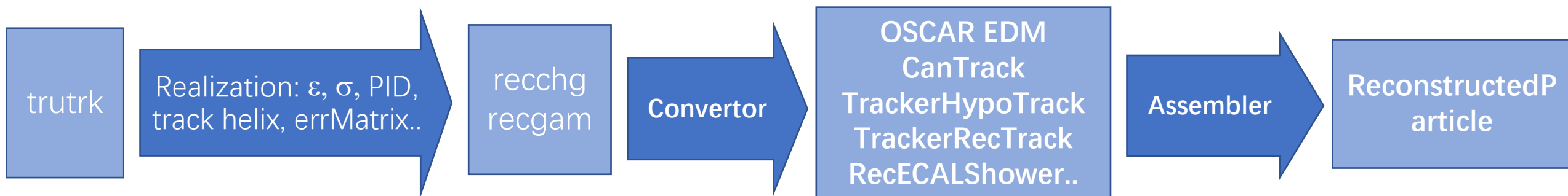
快模拟中的数据流

◆ 真实 (Truth) 信息:

- 粒子类型、动量、顶点、衰变链
- 数据类型: **struct** (结构体)

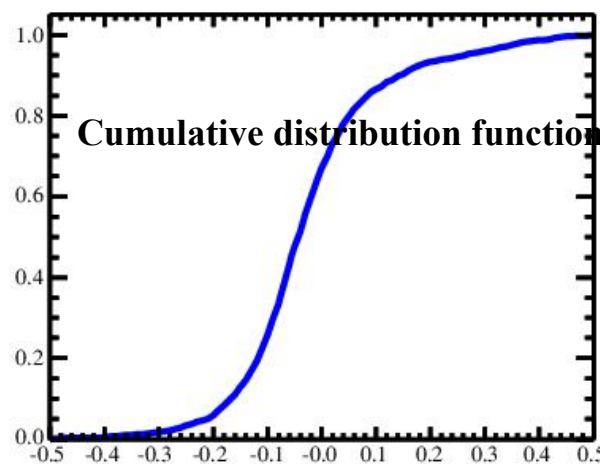
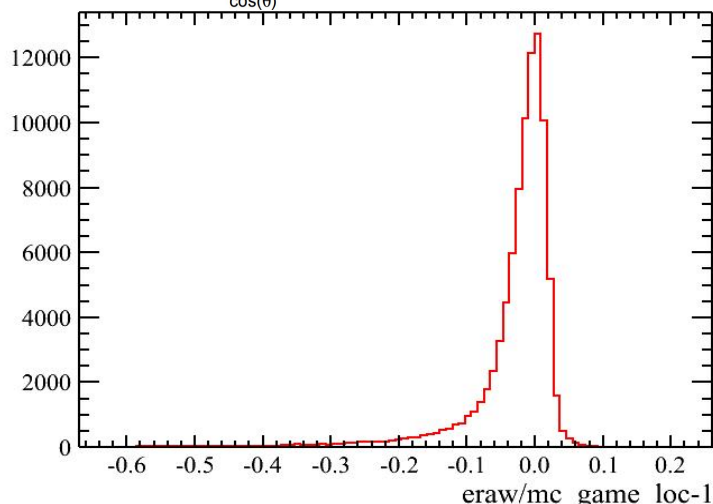
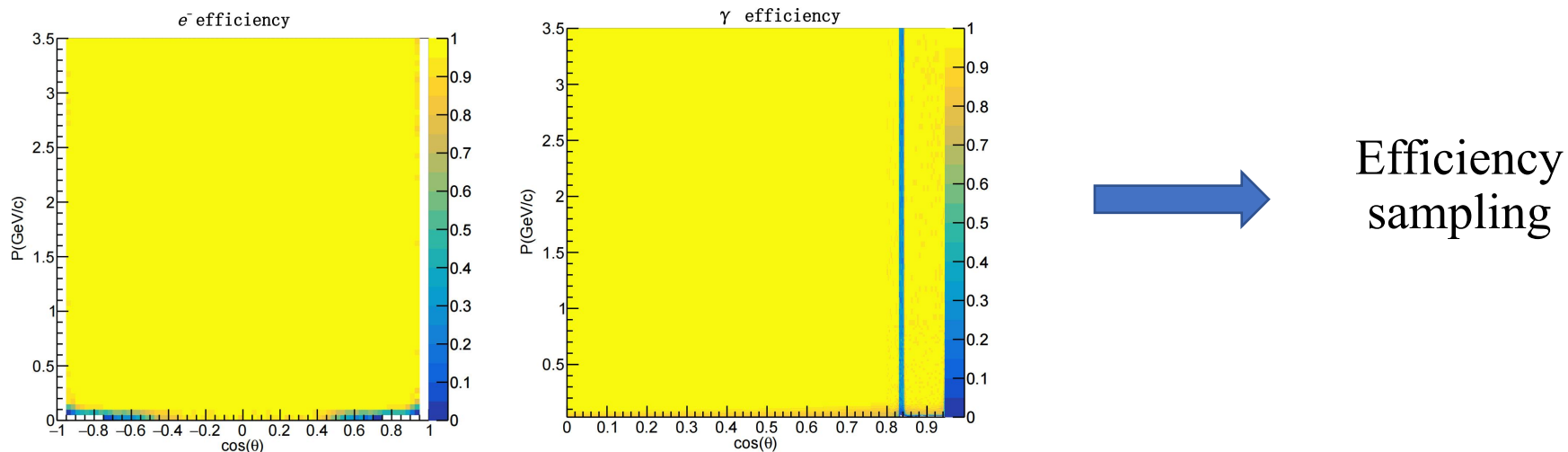
◆ 重建后的信息:

- 径迹效率、子探测器效率的实现, 动量/能量和空间分辨率、衰变顶点或产生点
- 径迹的螺旋线参数和误差矩阵
- 关于 p , $\cos(\theta)$ 的分区间 (**Binned**)
- 径迹和簇射对象构建
- 事件组装



径迹和簇射：效率和分辨

- ◆从全模拟单粒子 MC 中提取的探测器响应:能动量分辨、位置分辨、重建效率、粒子鉴别效率等
- ◆获得探测器响应的CDF曲线、效率直方图。
 - 通过产生子拿到粒子四动量信息利用探测器响应的CDF曲线、效率直方图进行抽样
 - 效率通过采样获得，动量/能量和空间分辨率通过smearing获得



Residual via
smearing(CDF)

径迹的误差矩阵处理

◆ 协方差矩阵元素

- $d_0, \phi_0, \kappa, d_z, \text{tg}\lambda$
- 按 $p, \cos(\theta)$ 分区间进行参数化
- 需要考虑相关性

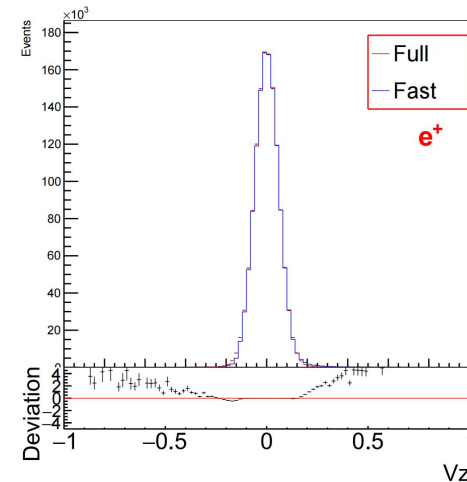
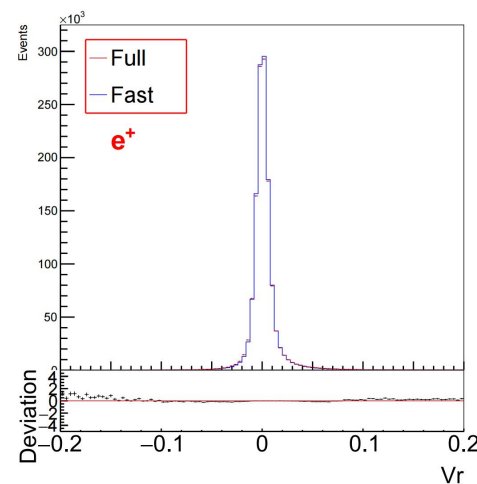
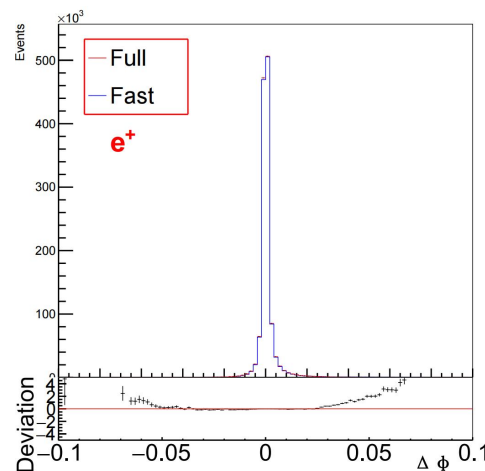
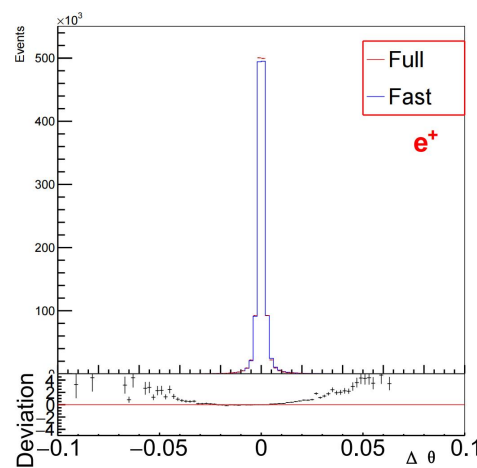
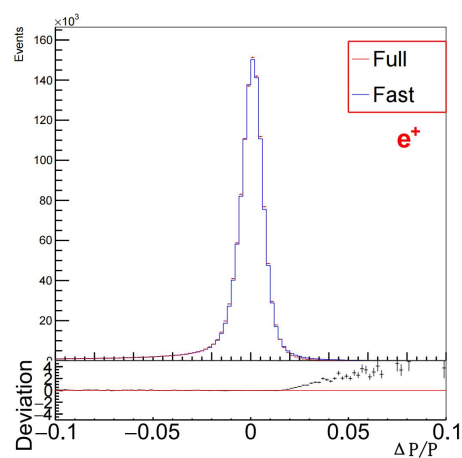
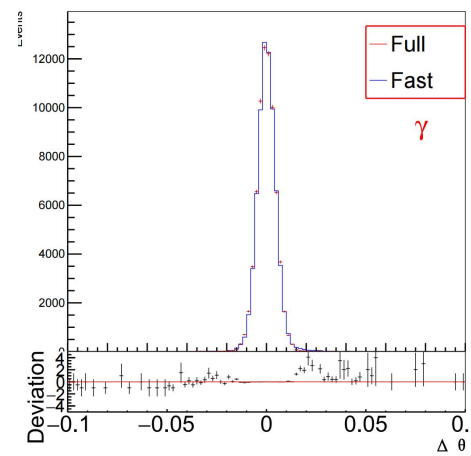
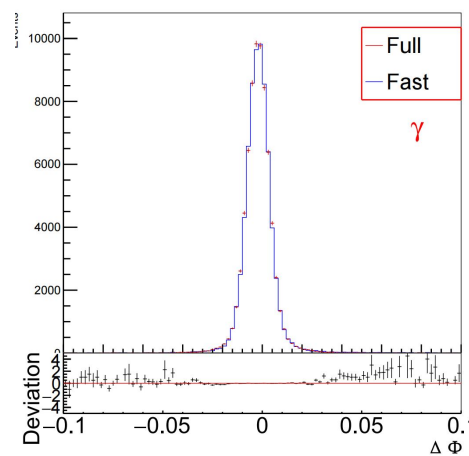
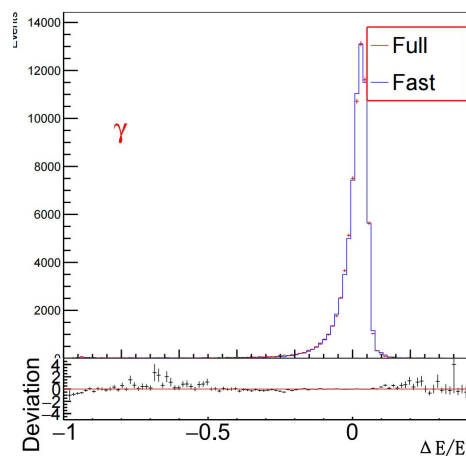
◆ 对角元素：通过累积分布函数抽取

◆ 非对角元素：

- 相关性部分使用修正形式来考虑：
$$\text{cov}'_{03} = \frac{\text{cov}(d_0, d_z)}{\sqrt{\text{cov}(d_0, d_0)\text{cov}(d_z, d_z)}}$$
- d_0 : 在 x - y 平面上，POCA（最近距离点）到 IP（对撞点）的距离
- d_z : 在 z 轴上，POCA 到 IP 的距离

快模拟性能验证: 单光子/带电径迹

- ◆ 能量、动量、顶点重建分辨率
- ◆ 残差分布与全重建吻合

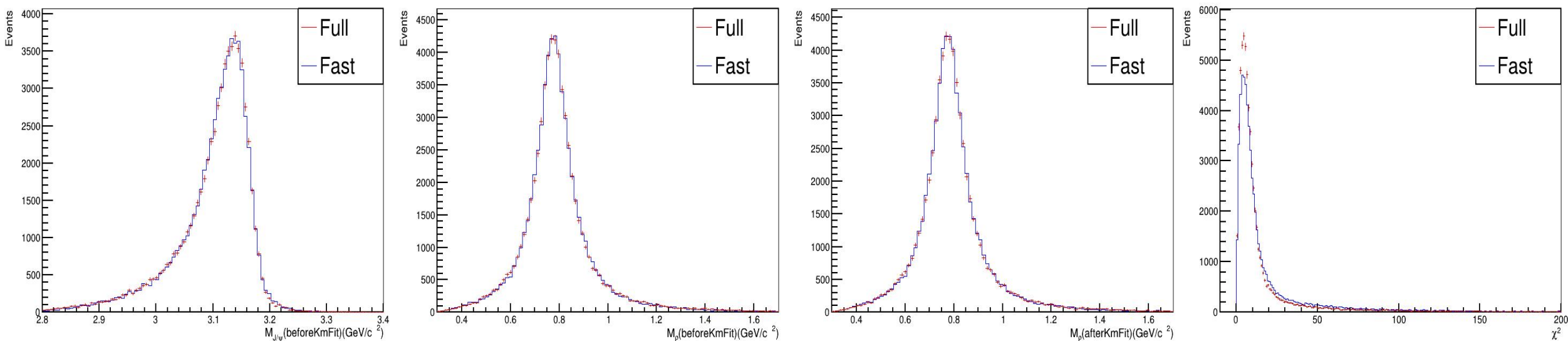


快模拟性能验证：物理过程 $J/\psi \rightarrow \rho\pi$

◆此物理过程包含：

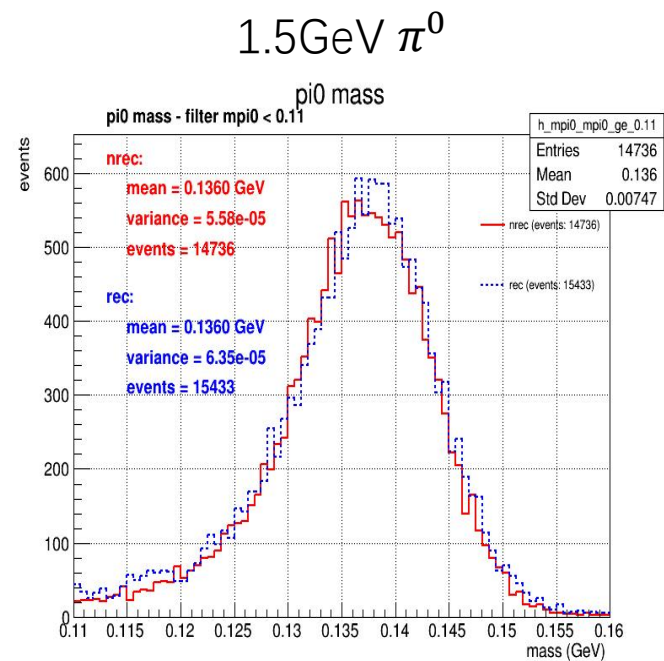
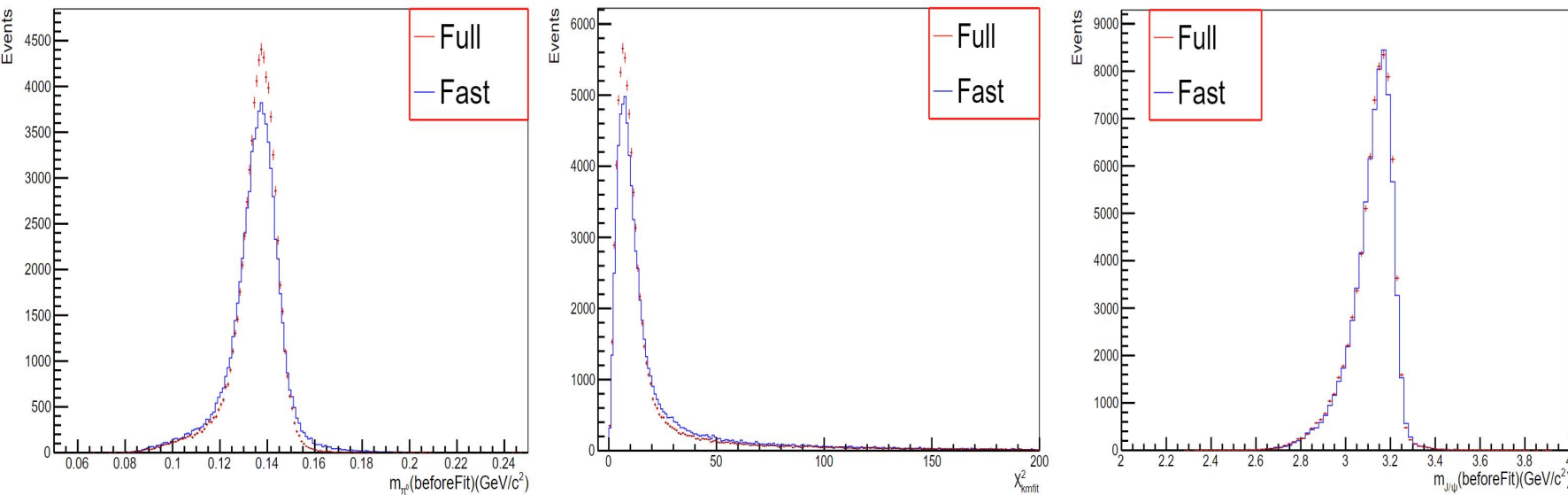
- 带电径迹、光子快速模拟
- 运动学拟合

◆与全重建分布吻合



快模拟性能验证：物理过程 $J/\psi \rightarrow \gamma\pi^0$

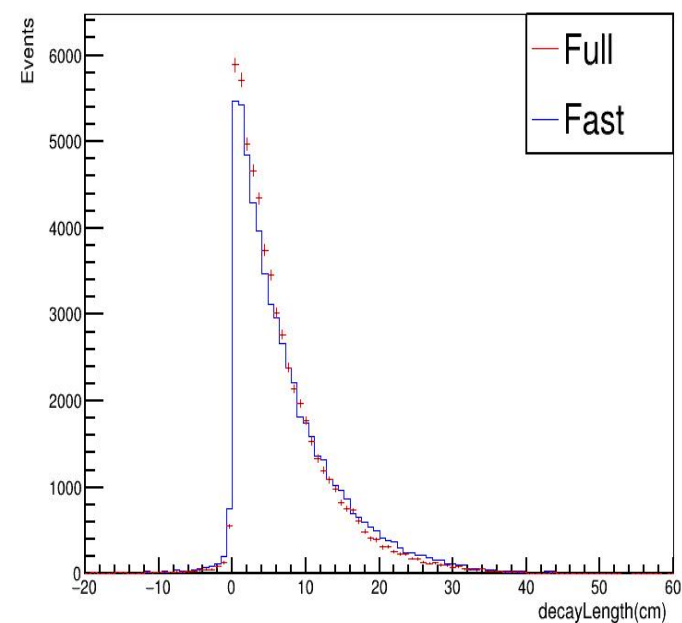
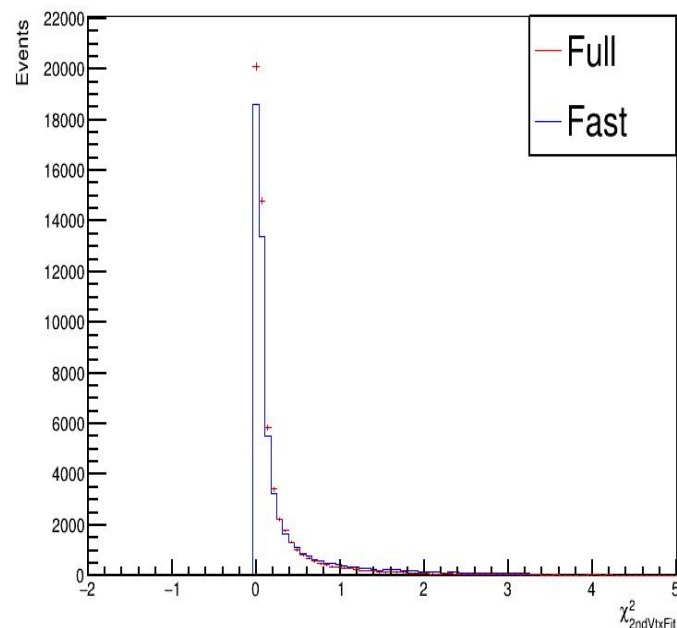
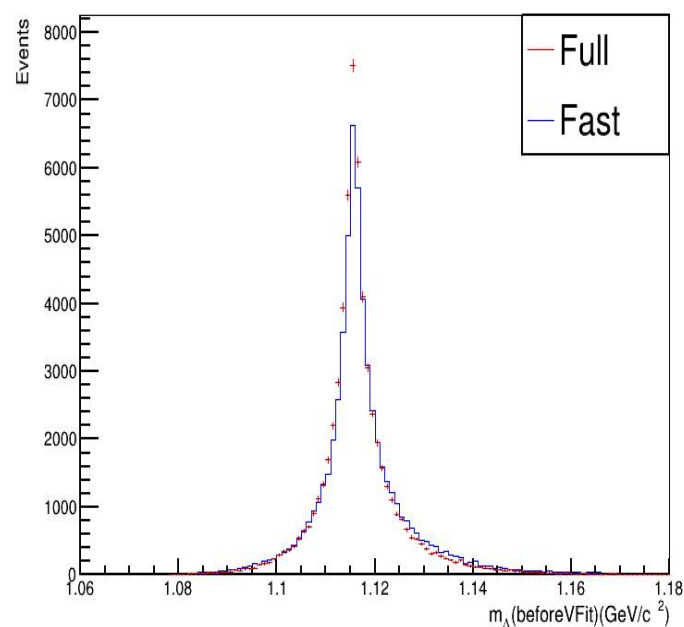
- ◆ 光子快速模拟, π^0 快速模拟
- ◆ 运动学拟合
- ◆ 通过多维抽样方法改善共振态等的质量谱与全重建样本的一致性



$J/\psi \rightarrow \gamma\pi^0$	快模拟	全重建
初选效率	73.85%	77.63%

快模拟性能验证：物理过程 $J/\psi \rightarrow \Lambda\bar{\Lambda}$

- ◆ 次级带电粒子（远离对撞点产生）
- ◆ 次级顶点与顶点拟合
- ◆ 分布较好吻合

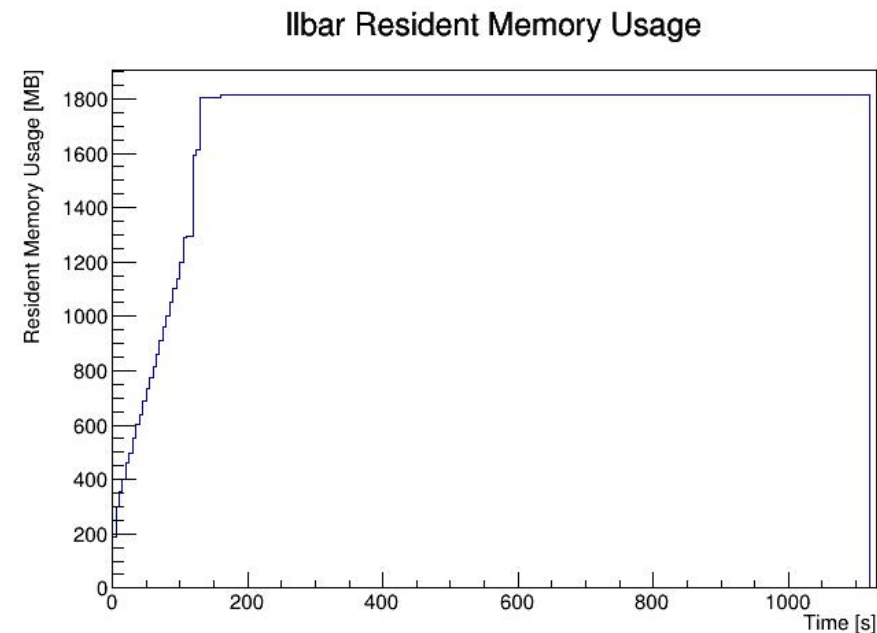


速度内存表现

◆资源消耗大幅减小:

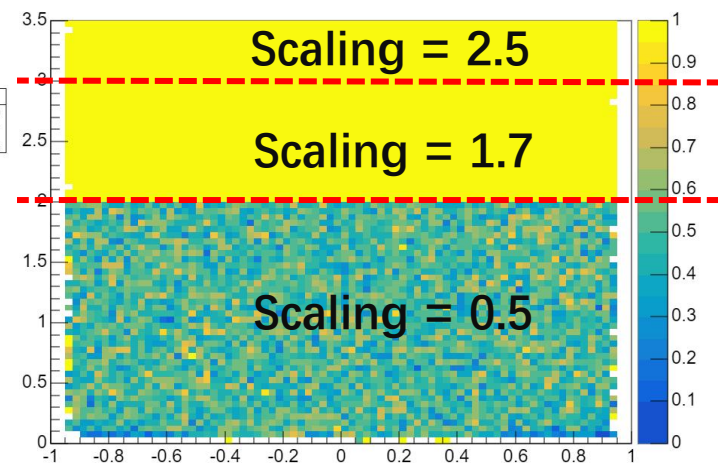
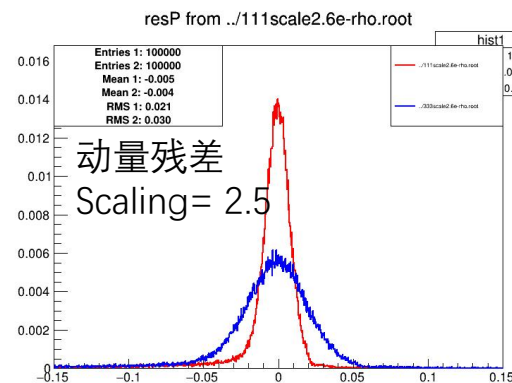
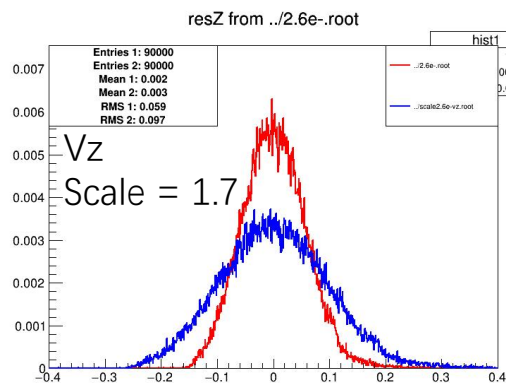
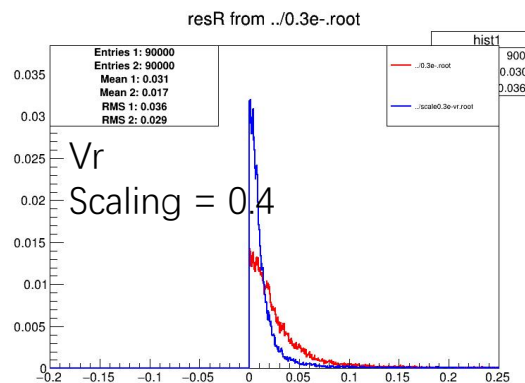
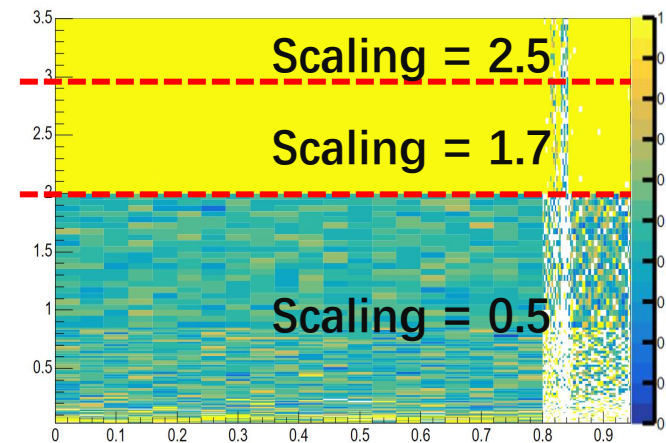
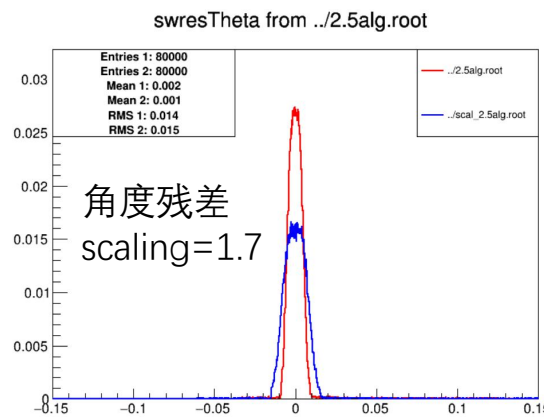
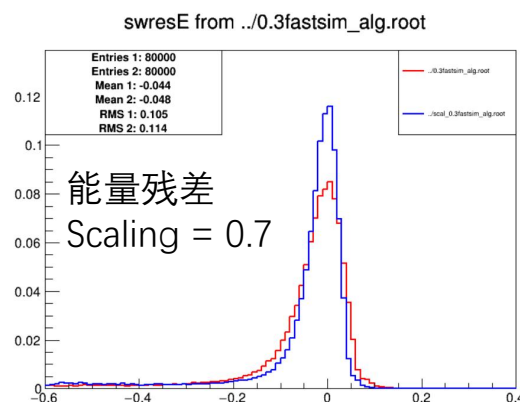
- 时间: 快约 $O(10^2)$ 倍
- 存储: 小10 倍以上
- 运行内存占用: 低于 2G

$\pi^+\pi^-\mu^+\mu^-$ (4 tracks)	Full Sim+Rec	Fast Sim
CPU time	$\sim 300\text{ms}(\text{sim}) + 800\text{ms}(\text{rec}) / \text{event}$	$\sim 2\text{ms} / \text{event}$
Event size	$\sim 77\text{kb}(\text{sim}) + 26\text{kb}(\text{rec}) / \text{event}$	$\sim 2\text{kb} / \text{event}$



对探测器性能的二维scaling

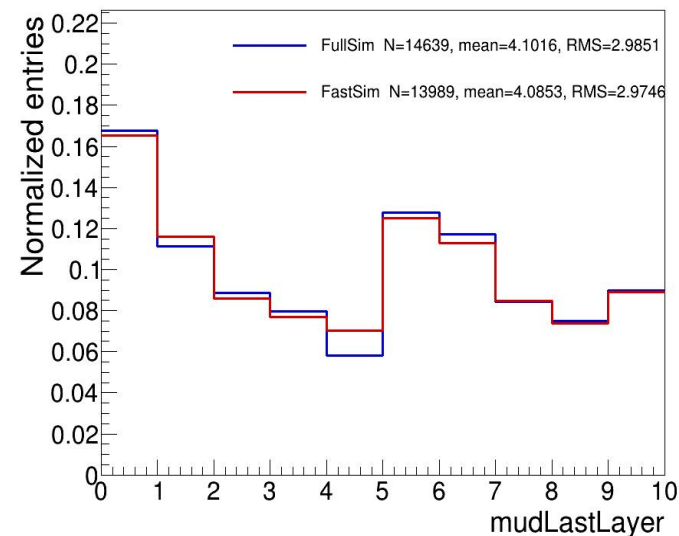
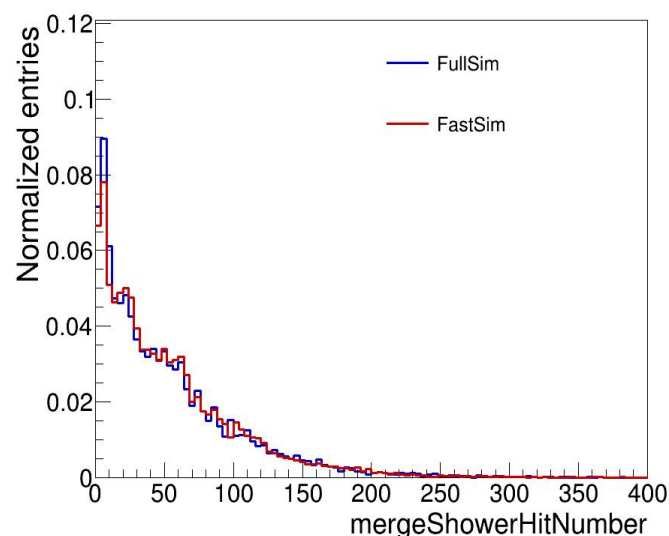
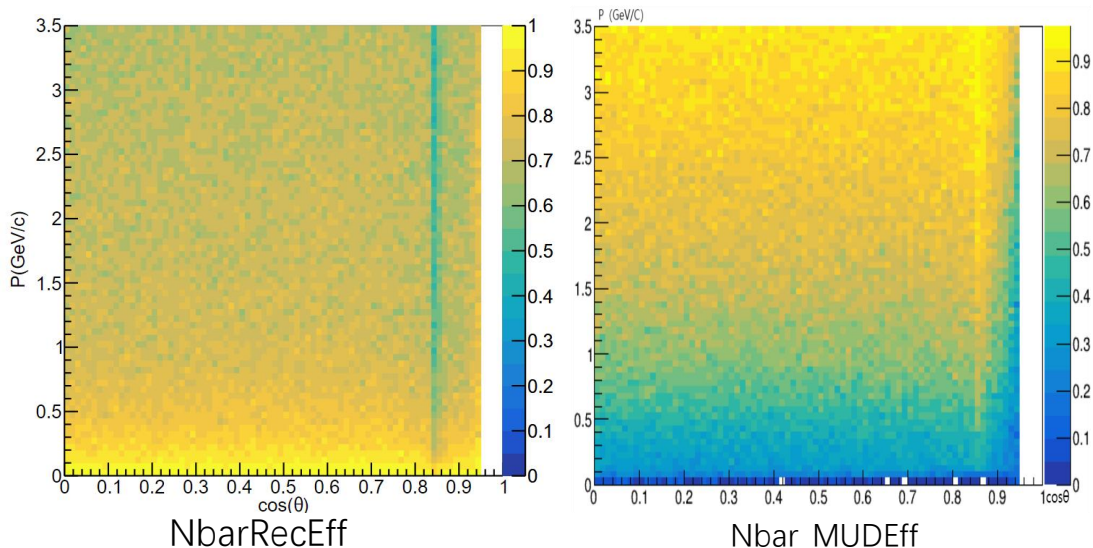
- ◆ 中性粒子: 实现对能量分辨, θ 角度分辨, Φ 角度分辨, 重建效率的scaling
- ◆ 带电粒子: 实现对动量分辨, 顶点 V_r , V_z , 重建效率的scaling
- ◆ 功能添加: 根据不同探测器方案可以切换到对应scaling参数



快模拟中对于中性粒子在ECAL与MUD信息处理

◆以1GeV Nbar为例

- 将MaxShower附近（按照 θ 和 Φ 向角度差）的簇射与其合并，计算合并后的HitNumber数量
- MUD中符合条件的cluster（MUD中cluster角度与ECAL中最大能量簇射角度在一定的 θ 和 Φ 角度范围内），找到这些cluster里的击中最深层
- 设置ECAL重建效率：NbarRecEff($|\cos\theta|$, p)
 - 实现角度残差和能量残差，即重建 θ , φ , E；并实现ShowerHitNumber
- ECAL有重建信息后，设置MUD重建效率：Nbar_MUDEff($|\cos\theta|$, p)
 - 实现MUDLastLayer

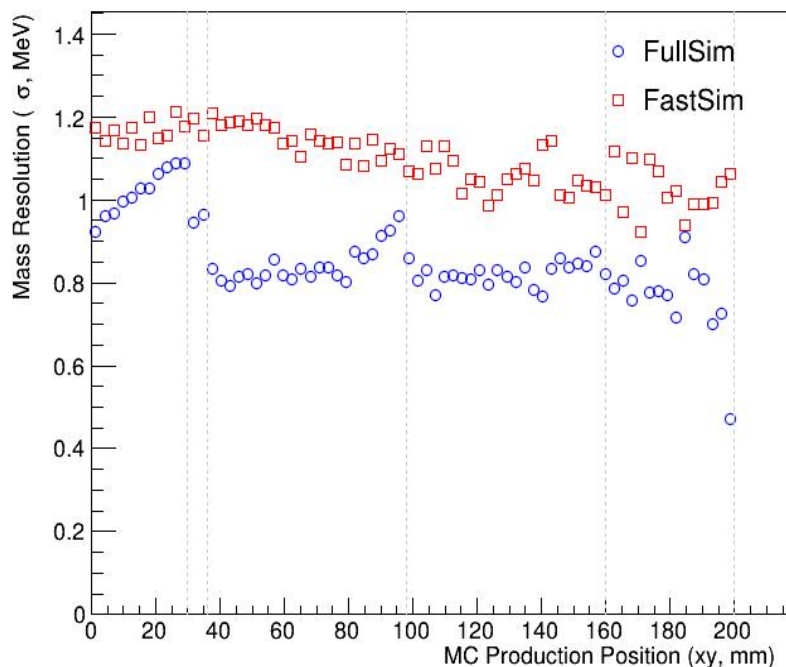


讨论：快模拟中次级粒子的表现

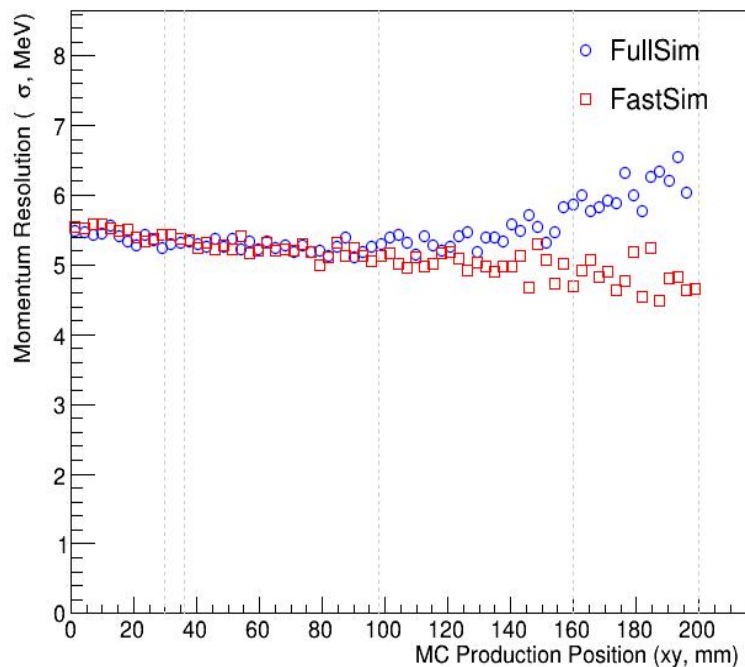
- ◆快模拟中误差矩阵的参数化是基于对撞点产生的样本
- ◆次级粒子末态径迹的产生位置远离 IP 点
- ◆动量残差和质量分辨率与径迹的产生点相关

$$J/\psi \rightarrow \Lambda \bar{\Lambda}$$

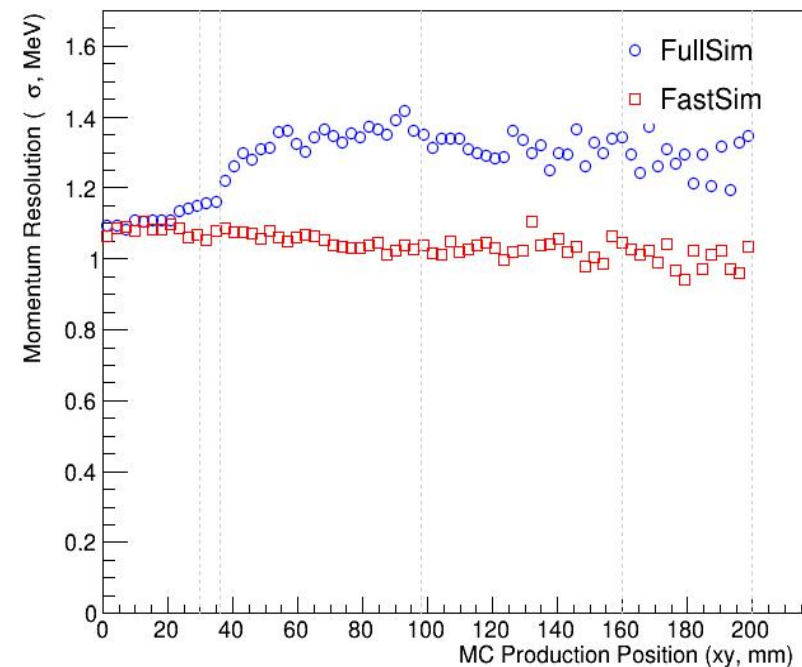
Lambda Mass Resolution vs Production Position (xy)



Proton Momentum Resolution vs Production Position (xy)

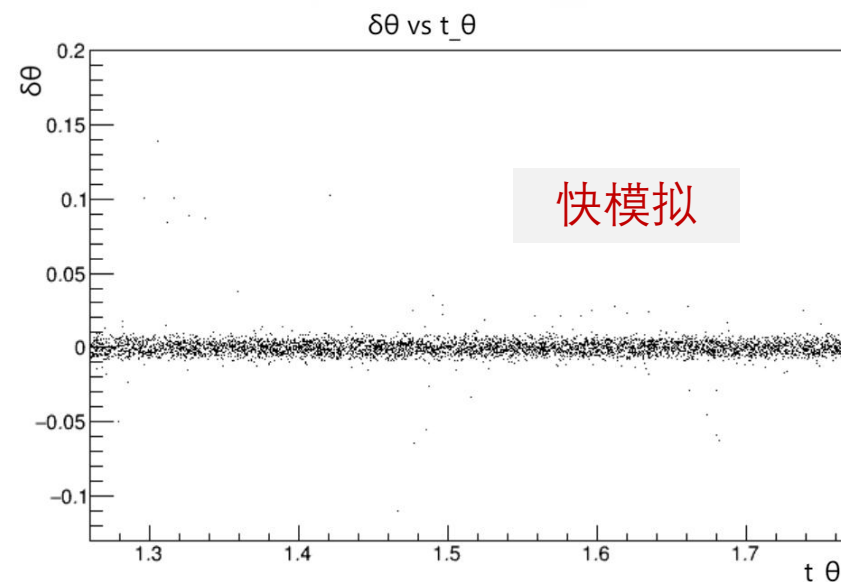
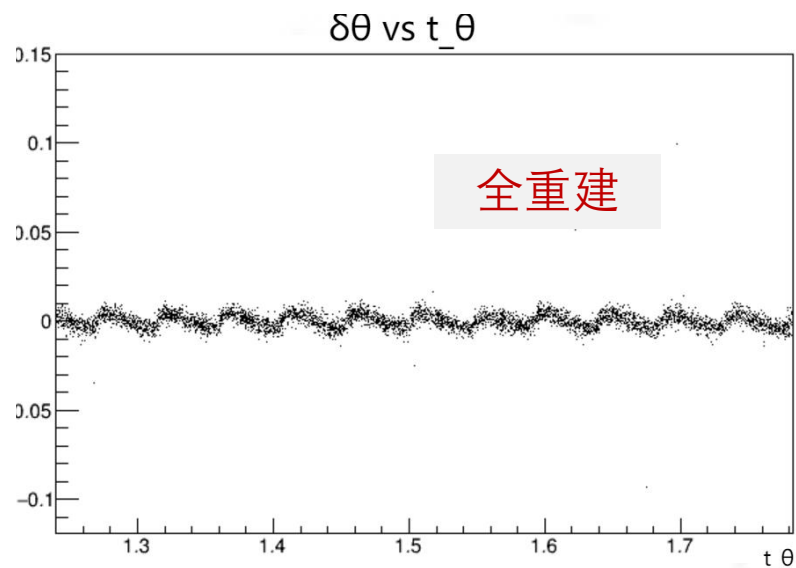
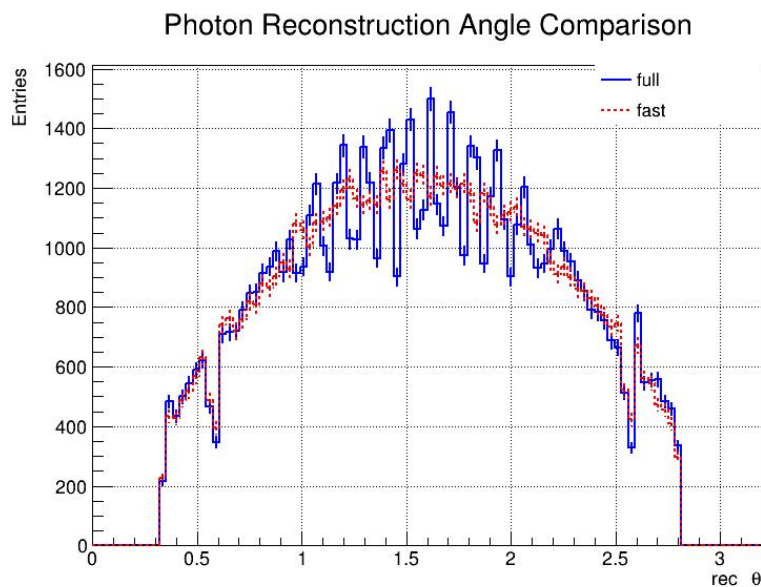


Pi- Momentum Resolution vs Production Position (xy)



讨论：快模拟中光子角度重建

- ◆光子重建角度残差随光子在晶体表面的不同入射角而变化
- ◆快速模拟中使用了分区间（**binned**）参数化，这种精细结构丢失了
- ◆由于这种依赖关系的颗粒度很小，可能不会影响数据分析



总结

- ◆快速模拟已实现基本功能，可用于物理分析
- ◆内存和时间消耗相对全重建大幅压缩
- ◆对不同的探测器响应可灵活设置
- ◆更多改进正在研究中



Thanks!



Back up

快模拟的更新

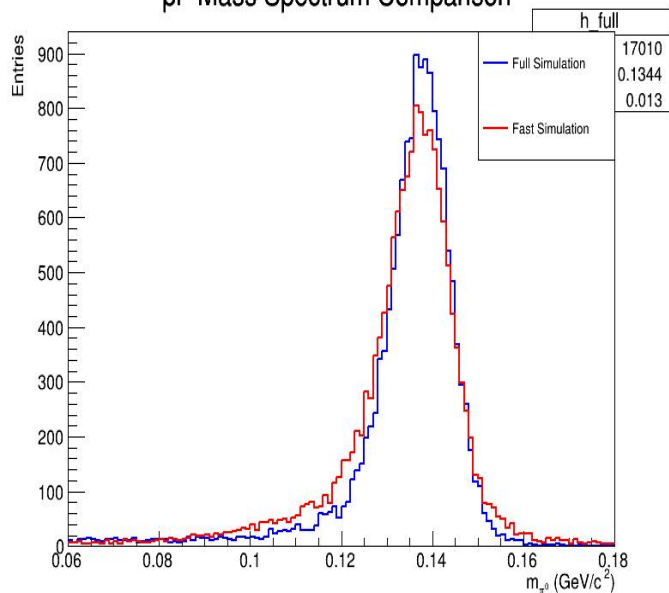
- ◆1.在OSCAR2.6.2下，产生新的全模拟样本并进行PDF的更新。
- ◆2.对快模拟进行降维处理，误差矩阵由6*6转换为5*5。
- ◆3.对快模拟进行了二维scaling。
- ◆4.在快模拟中加入了中性强子的击中数以及mud击中层数信息。

单粒子1.5GeV π^0 多维抽样

	多维抽样	全重建	相对效率	绝对效率
初始	20,000	20,000	多维抽样 vs 全重建	多维抽样 vs 全重建
trackID<0 的数量大于等于2	15472	17996	77.36% vs 89.98%	77.36% vs 89.98%
MC match cut角度和能量	15245	16556	98.53% vs 92%	76.23% vs 82.78%
m_pi0>0.05	15218	16376	99.82% vs 98.91%	76.09% vs 81.88%
m_pi0>0.11	14736	15433	96.83% vs 94.24%	73.68% vs 77.17%

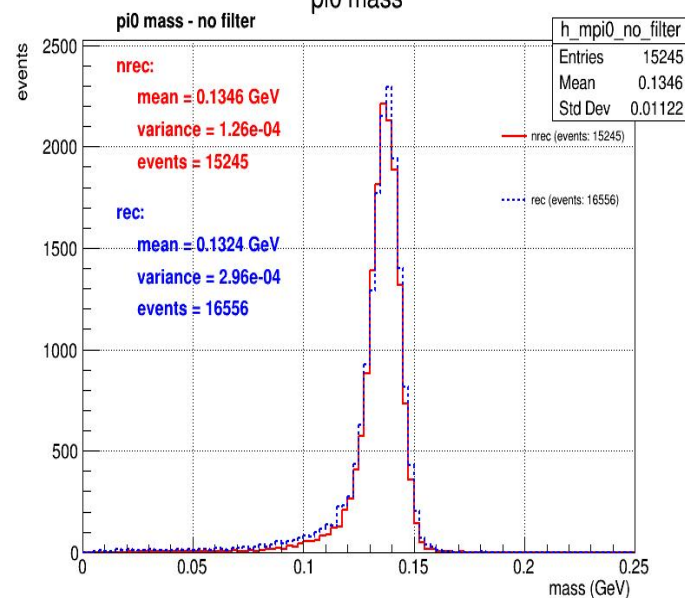
非多维抽样

π^0 Mass Spectrum Comparison



多维抽样

π^0 mass



π^0 mass

π^0 mass - filter mpi0 < 0.11

