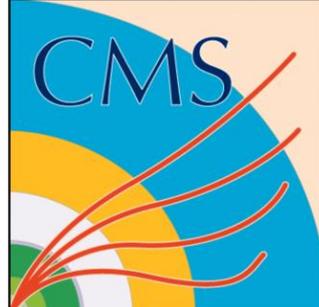




北京航空航天大学
BEIHANG UNIVERSITY

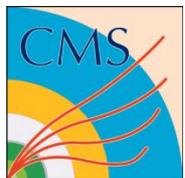


第三届CMS冬令营 喷注和 τ 子

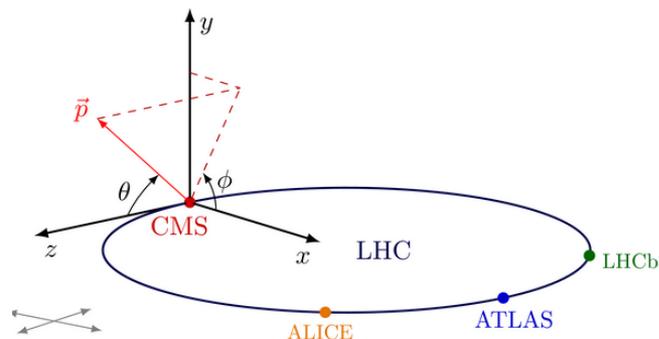
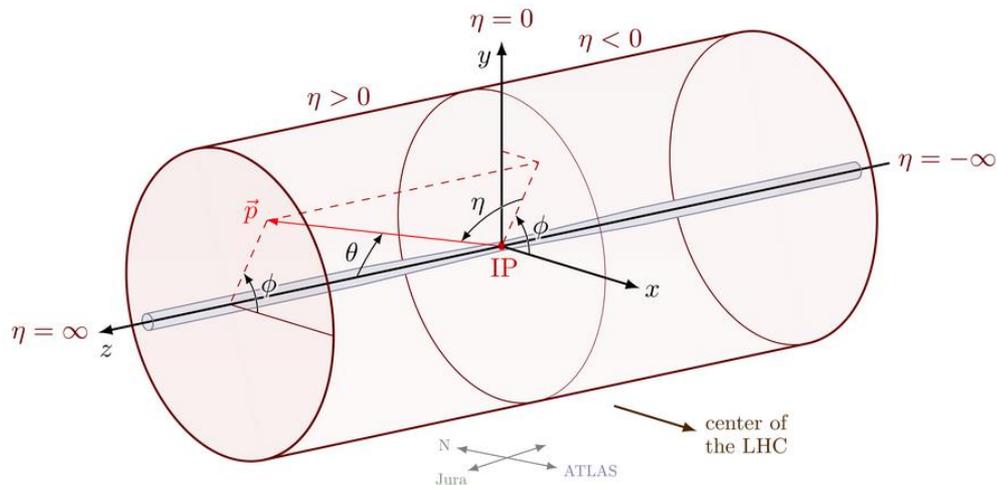
余涛哲

2025.01.17

北京航空航天大学



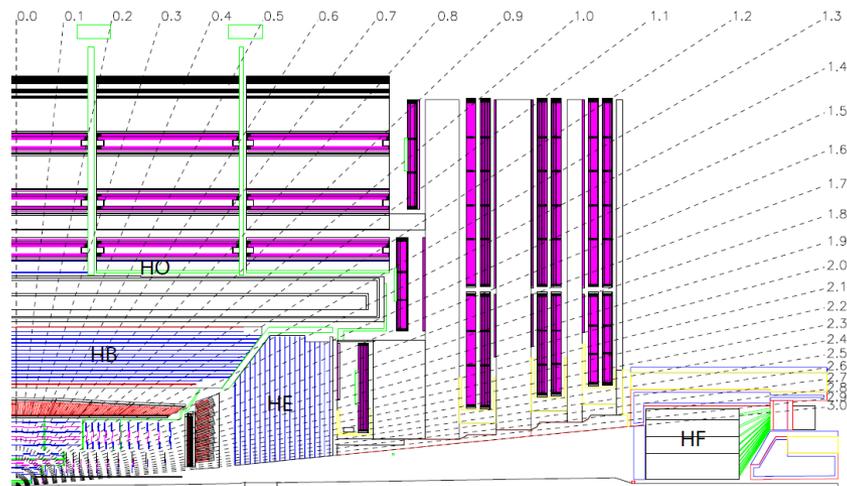
预知识:CMS探测器坐标系

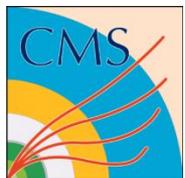


CMS的实验室坐标系见上图

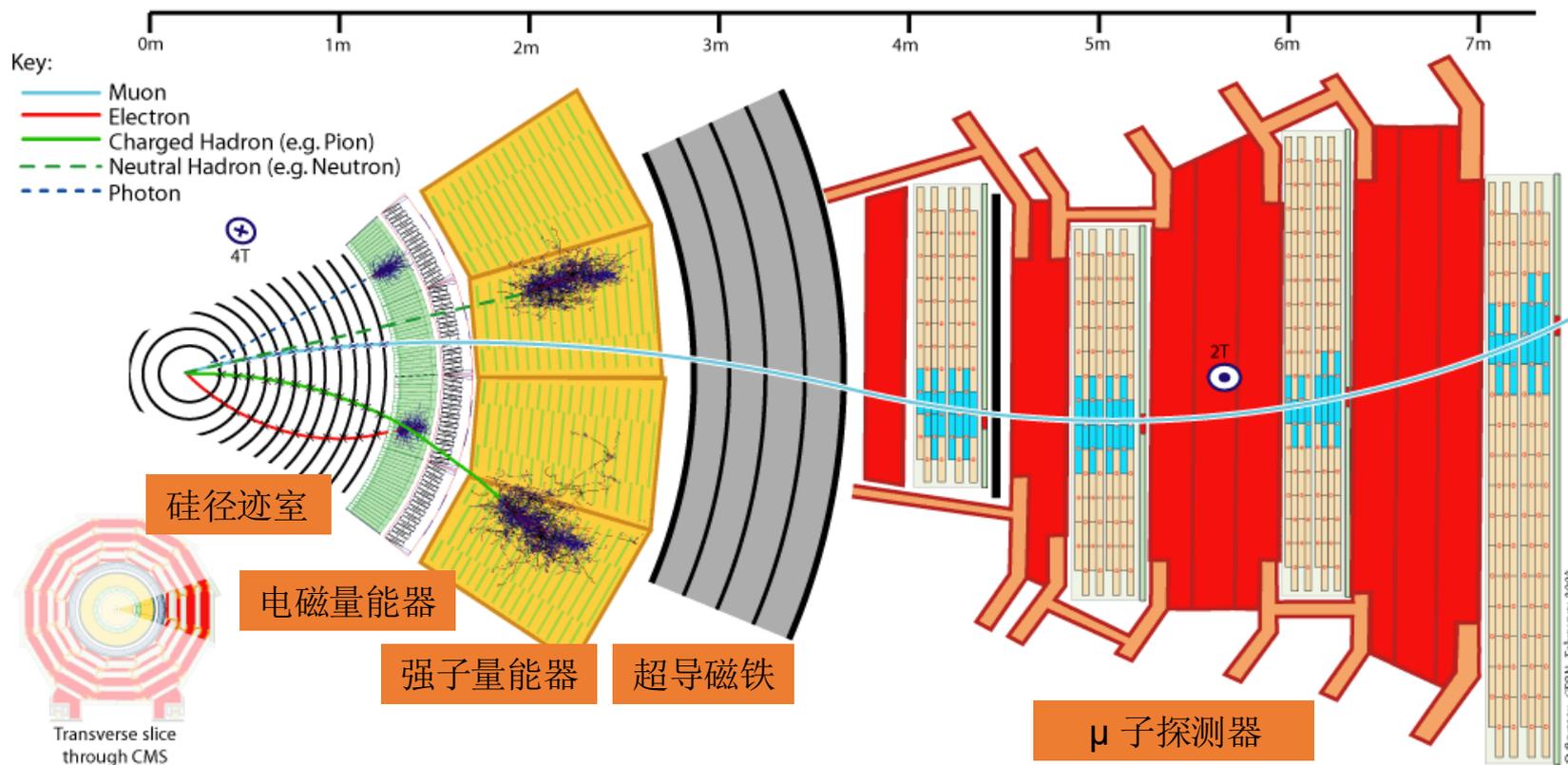
$$\eta = -\ln \tan \frac{\theta}{2} \quad \text{赝快度}$$

用 η 表示的 θ 的方位见右图





预知识: CMS 探测器介绍

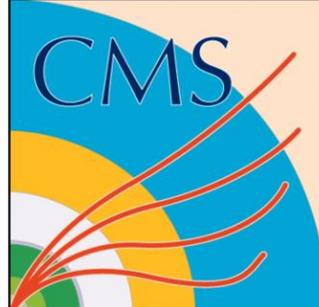


硅径迹室—重建径迹和顶点 电磁量能器—探测电子，光子 强子量能器—探测中子，质子
 超导磁铁—偏转带电粒子 μ子探测器—探测μ子

- CMS会结合所有的子探测器信息，重建出粒子流的候选者：电子，μ子，光子，带电强子，中性强子
- 这些粒子流候选者会进一步重建出更高一级的物理对象：喷注，τ子，损失横动量



北京航空航天大学
BEIHANG UNIVERSITY

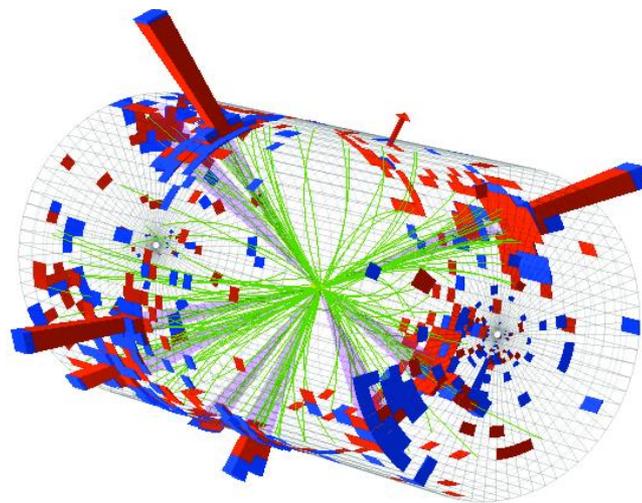
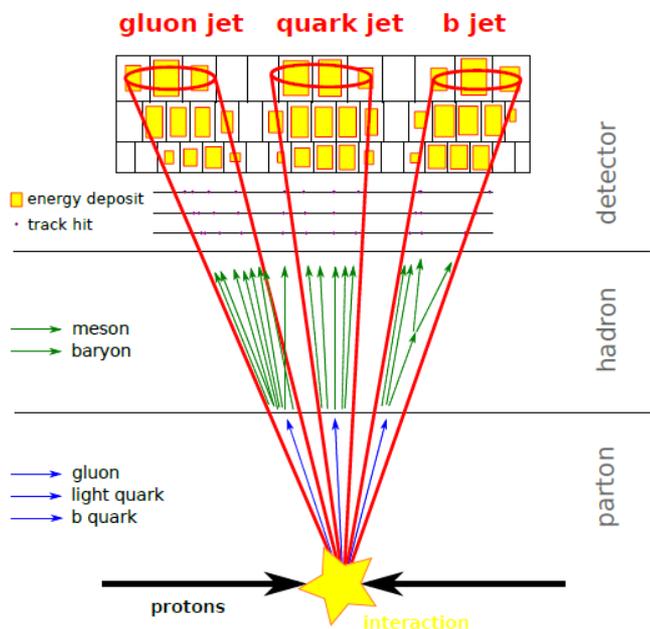


喷注(jet)

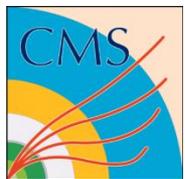


喷注是什么？

- LHC是强子对撞机，对撞后会产生大量的胶子和夸克
- 夸克和胶子由于色禁闭，不能独立存在
- 夸克和胶子强子化后形成喷注(jets)



CMS Experiment at LHC, CERN
 Data recorded: Mon May 23 21:46:26 2011 EDT
 Run/Event: 165567 / 347495624
 Lumi section: 280
 Orbit/Crossing: 73255853 / 3161



CMS探测器怎么重建喷注



- CMS通过聚类算法来将粒子流(particle flow)候选对象重建成喷注

$$d_{iB} = p_{\perp i}^{2a}$$

$$d_{ij} = \min(p_{\perp i}^{2a}, p_{\perp j}^{2a}) \times \frac{\Delta R_{ij}^2}{R^2}$$

1. $a = 1$ k_T 算法
2. $a = 0$ CA
3. $a = -1$ anti- k_T 算法



- d_{iB} 是对象i到原点的动量距离
- R 是描述喷注圆锥半径大小的参数;
- $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ 描述喷注候选对象i和j在探测器内的空间距离;

聚类算法寻找喷注遵循以下步骤:

选定一群粒子流作为喷注的候选对象，并从中找到最小的 d_{iB} 和对应的 d_{ij}

当 $d_{ij} < d_{iB}$ 时，喷注候选对象i和j合并，更新喷注候选对象i信息，
当 $d_{ij} > d_{iB}$ 时，喷注候选对象i就标定为一个喷注。

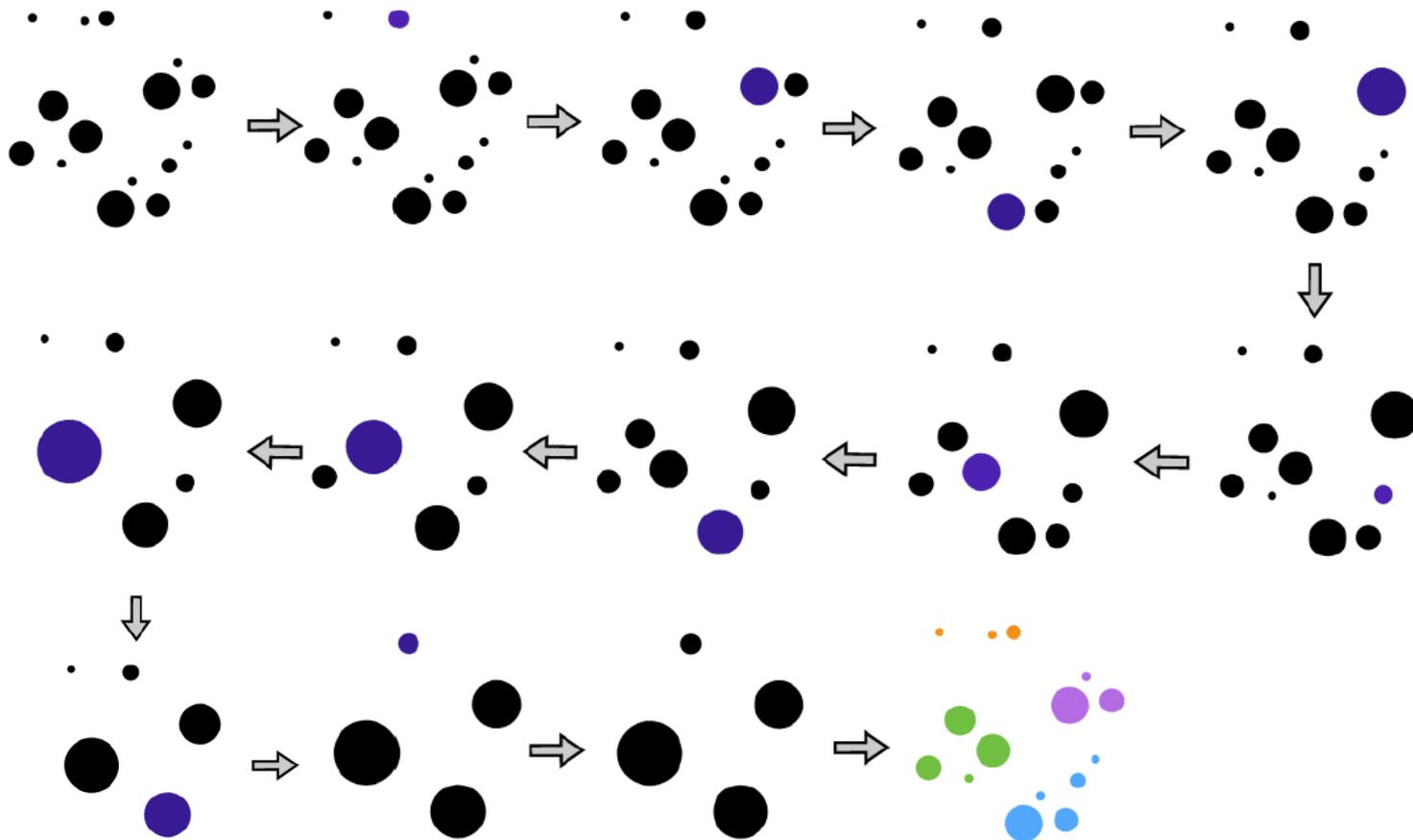
如果还有喷注候选对象未被标定，就继续第一步，直到所有喷注候选者标定为喷注



找到粒子流，并
把它们作为喷注
的候选对象

找到最小的 d_{iB} ，并找到对
应的最小 d_{ij} ，通过比较合
并为一个喷注候选对象

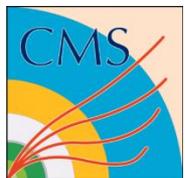
反复迭代，在每一步都合并
最小 d_{ij} 的两个喷注候选对象



第一次 $d_{ij} > d_{iB}$
把喷注候选对象
标定为喷注

所有的喷注候选
对象标定为喷注

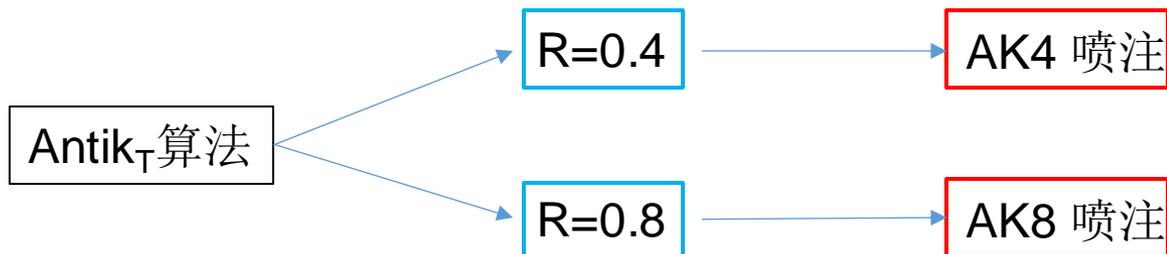
重建了四个喷注，
每一个都有N个成分



Anti-k_T 喷注



- 目前CMS上用的最多的聚类算法就是anti-k_T算法，根据R参数数值有两种算法最常用到



$$d_{iB} = p_{\perp i}^{2a} \quad (a=-1)$$

$$d_{ij} = \min(p_{\perp i}^{2a}, p_{\perp j}^{2a}) \times \frac{\Delta R_{ij}^2}{R^2}$$

注意！记住这两个喷注类型

目前CMS最常用的两种数据类型MiniAOD 和NanoAOD 都包含了这两种数据类型：

MiniAOD:

slimmedJets = ak4PFJets
slimmedJetsAK8 = ak8PFJets

思考：R的数值代表了什么

NanoAOD:

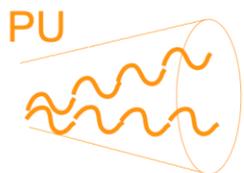
Jet = ak4PFJets
FatJet = ak8PFJets

也许你们以后分析工作中只会用到AK4喷注和AK8喷注，但也不要忘了另外两种算法 8

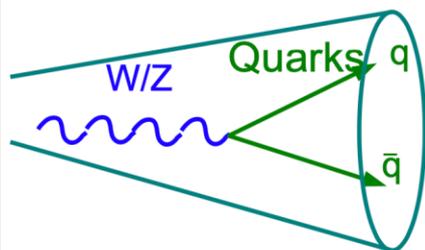


喷注圆锥半径的大小可以一定程度反应喷注的类型

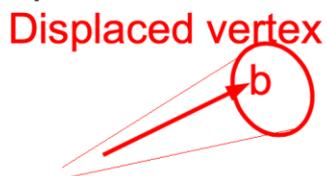
Pileup Jet



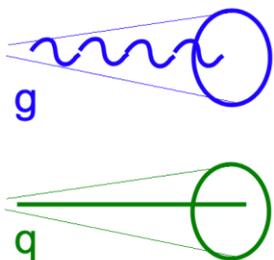
V Boson Jet



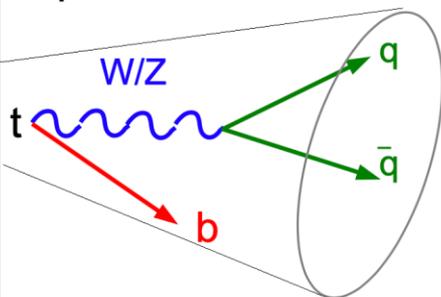
B-quark Jet



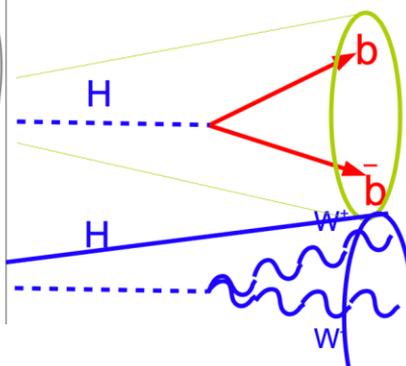
Quark/Gluon Jet



Top Jet



Higgs Jet



比较大的圆锥(AK8):
Top 喷注, W喷注

比较小的圆锥(AK4):
bottom喷注
Quark, gluon喷注

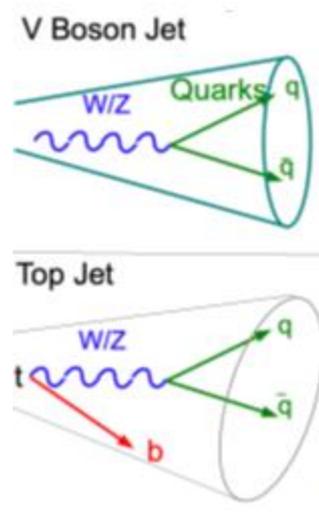
思考: Top喷注和W喷注都可以用AK8喷注来表示, 如何进一步去区分它们?



研究喷注的子结构有助于我们进一步来确定喷注的类型

思考：什么样的喷注会有子结构？

例子：具有高动量(boost)的top夸克强子化时，衰变产物的喷注会发生重叠，这时这种top喷注的子结构就会非常复杂



通过研究喷注的子结构来确定喷注的类型，给不同的喷注“贴上标签” (tagging)

思考：如果给喷注贴错标签(mistagging)了怎么办？



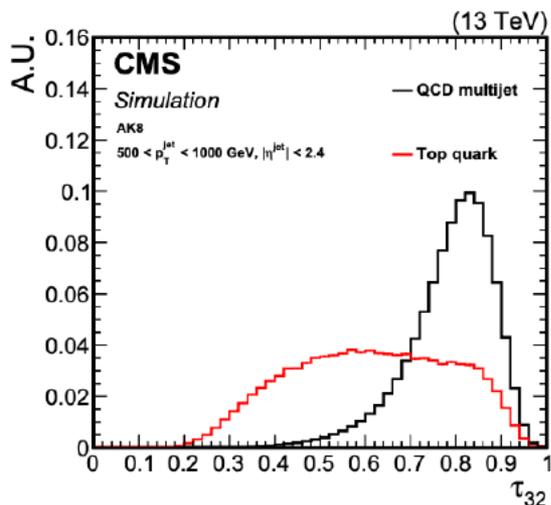
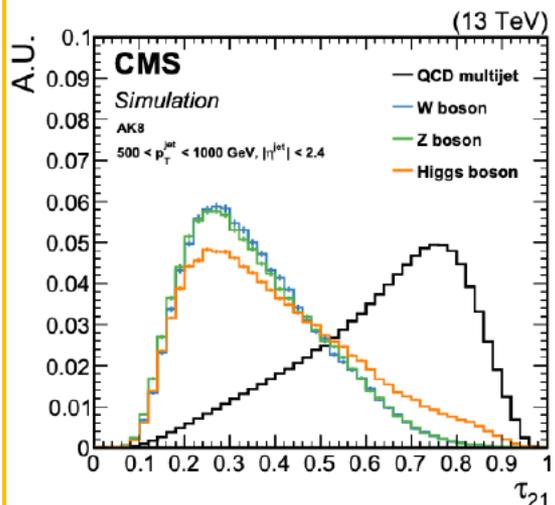


- N-subjettiness 是可以表征子喷注的物理量，定义为

$$\tau_N = \frac{1}{d_0} \sum_k [p_{Tk} \cdot \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k})]$$

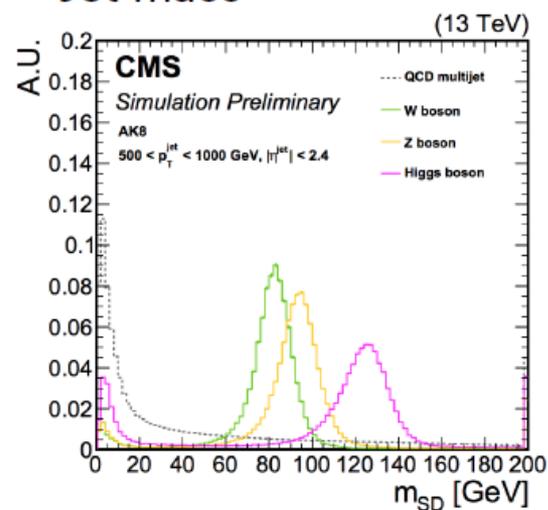
k是组成喷注粒子流的数量， P_{Tk} 是第k个组分的横动量，N是子喷注的数量， $\Delta R_{N,k}$ 是第k个组分和第n个子喷注的距离， $d_0 = \sum_k p_{Tk} R_0$ 是归一化因子。

用来分辨W喷注的变量是 $\tau_{21} = \tau_2/\tau_1$
 用来分辨Top喷注的变量是 $\tau_{32} = \tau_3/\tau_2$



对质量，动量进行选择进一步标定喷注的类型

Jet mass



目前CMS也使用机器学习技术来标定(tagging)不同类型的喷注

- 深度神经网络(DNN)→DeepAK(X)
- 卷积神经网络(CNN)→ParticleNet

思考：完成对喷注的标定后，我们还需要做什么？

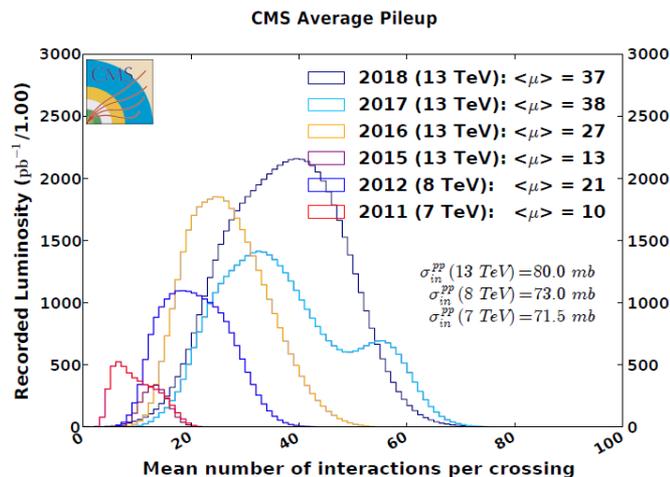
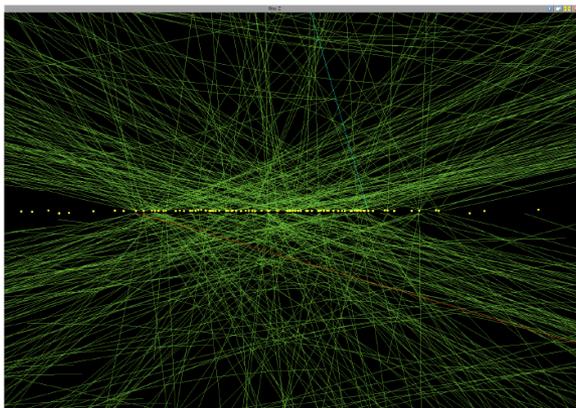


➤ 什么是粒子堆积?

LHC是高亮度的粒子对撞机, 平均每次束团对撞约有**26**个质子-质子对撞, 而我们关心的是最“硬”(hard)的那一个对撞, 这意味着平均有**25**个堆积(pile-up)事例

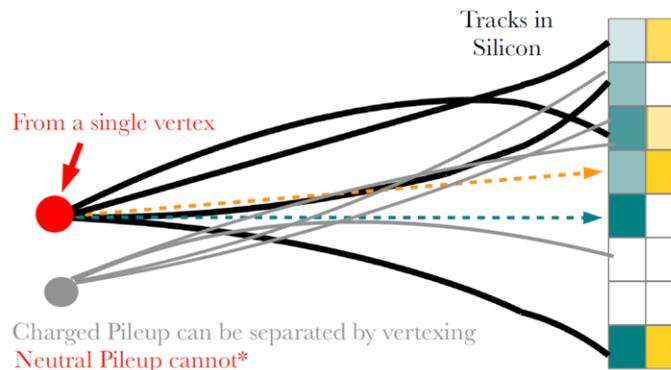
小知识: 什么是最“硬”的对撞

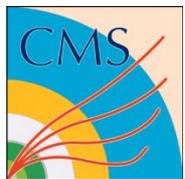
指在质心系中某一入射粒子在碰撞后其动能有较大幅度的损失的情况, 这样粒子损失的动能会转化为新粒子的质量



- 粒子堆积对我们分析的影响:
 - 会影响硬散射的喷注能量

我们需要压低粒子堆积对于喷注的影响!





粒子堆积效应的压低



➤ 带电强子扣除法(Charge-Hadron Subtraction, CHS)

- 应用在喷注的聚类算法之前，对于粒子流进行操作

效果：这种方法可以把粒子堆积产生的带电强子去除

局限性：1.只能去除带电强子 2.需要结合径迹探测器($|\eta|<2.5$)

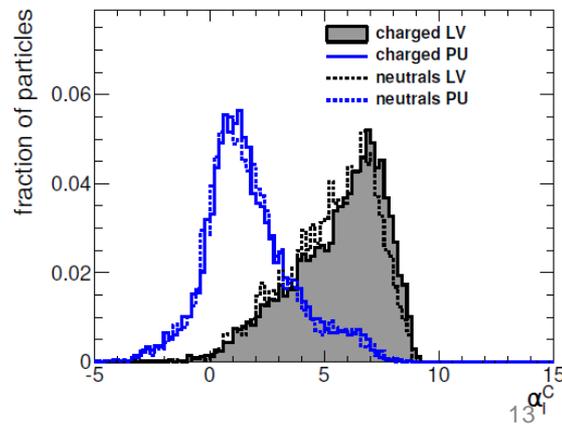
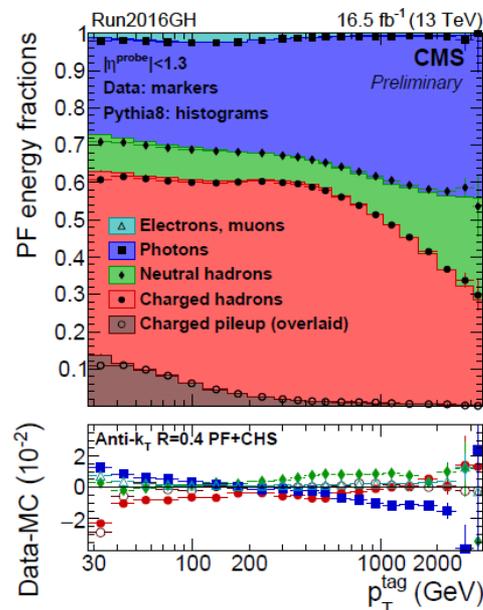
→中性强子的能量需要使用能量修正

➤ 粒子堆积鉴别法(Pileup per Particle Identification, PUPPI)

- 这种方法是建立在CHS方法基础之上，去除来自粒子堆积的中性强子和 $|\eta|>2.5$ 区域的粒子堆积效应产生的强子

- 方法：对于每个粒子流候选者分配一个权重，这个权重表明这个粒子流是有多大概率来自于粒子堆积

喷注的动力学特性受到来自粒子堆积效应的影响较小

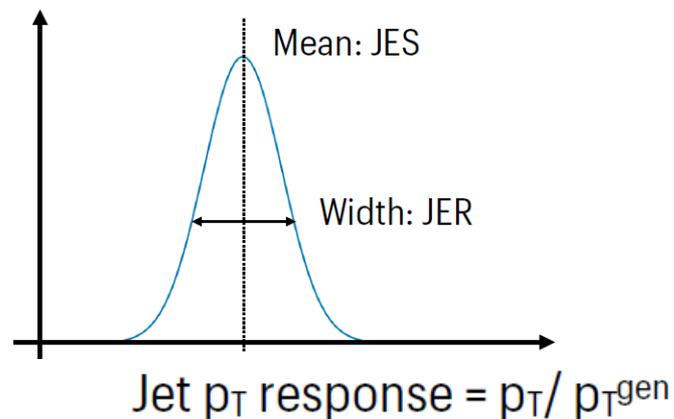




➤ 能量修正的原因:

- 探测器的噪声、非线性响应等因素
- 粒子堆积效应会给喷注带来额外的能量

➤ 喷注能量修正方法:

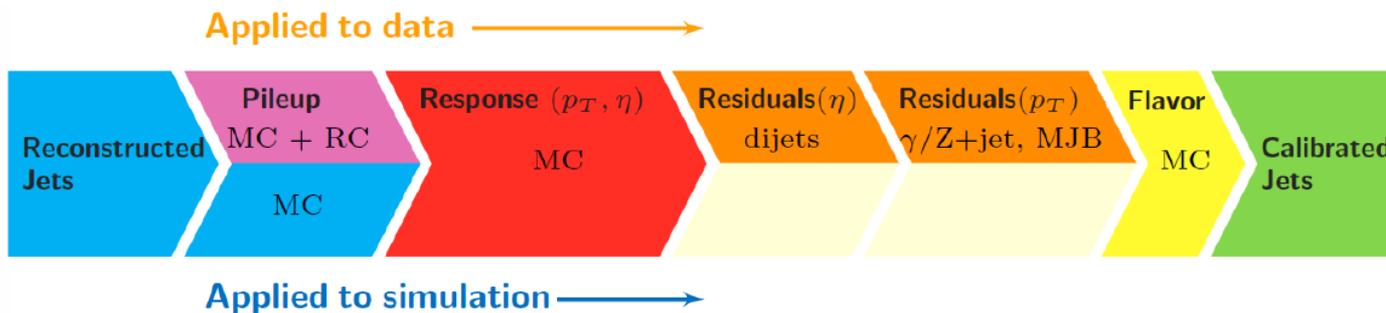


L1是对粒子堆积的修正, 进一步去除粒子堆积的能量
 L2, L3修正基于QCD-dijet模拟事例修正喷注四动量

$$p_{L1}^\mu = C_{L1}(\rho_{ev}, A_{jet}, p_{raw}^\mu) \times p_{raw}^\mu \quad (\text{pile-up})$$

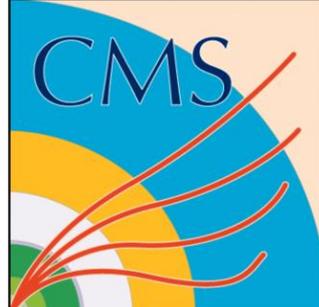
$$p_{L23MC}^\mu = C_{L23MC}(p_{L1}^\mu) \times p_{L1}^\mu \quad (\text{MC truth})$$

$$p_{L23Res}^\mu = C_{L23Res}(p_{L23MC}^\mu) \times p_{L23MC}^\mu \quad (\text{residual - data only})$$





北京航空航天大学
BEIHANG UNIVERSITY



子



tau(τ)的历史:

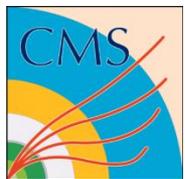
马丁·佩尔实验团队于**1975**年做实验探测到 τ 子,这实验主要使用到斯坦福直线加速器中心的斯坦福正负电子非对称圈 (**SPEAR**) 与劳伦斯伯克利国家实验室的磁性探测器。马丁·佩尔因为发现 τ 子而荣获**1995**诺贝尔物理学奖。

$$e^- + e^+ \rightarrow \tau^- + \tau^+ \rightarrow e^\mp + \mu^\pm + 4\nu$$

tau(τ)的性质:

1. 标准模型第三代带电轻子
2. 带电量: $-1e$
3. 质量: **1.78 GeV** 是电子质量的大约 **3478**倍
4. 自旋: $1/2$, 费米子
5. 寿命: 10^{-13} 秒
6. 动量为**40GeV**的 τ 在衰变前会在**CMS**中留下**2mm**的径迹





τ 子的衰变道

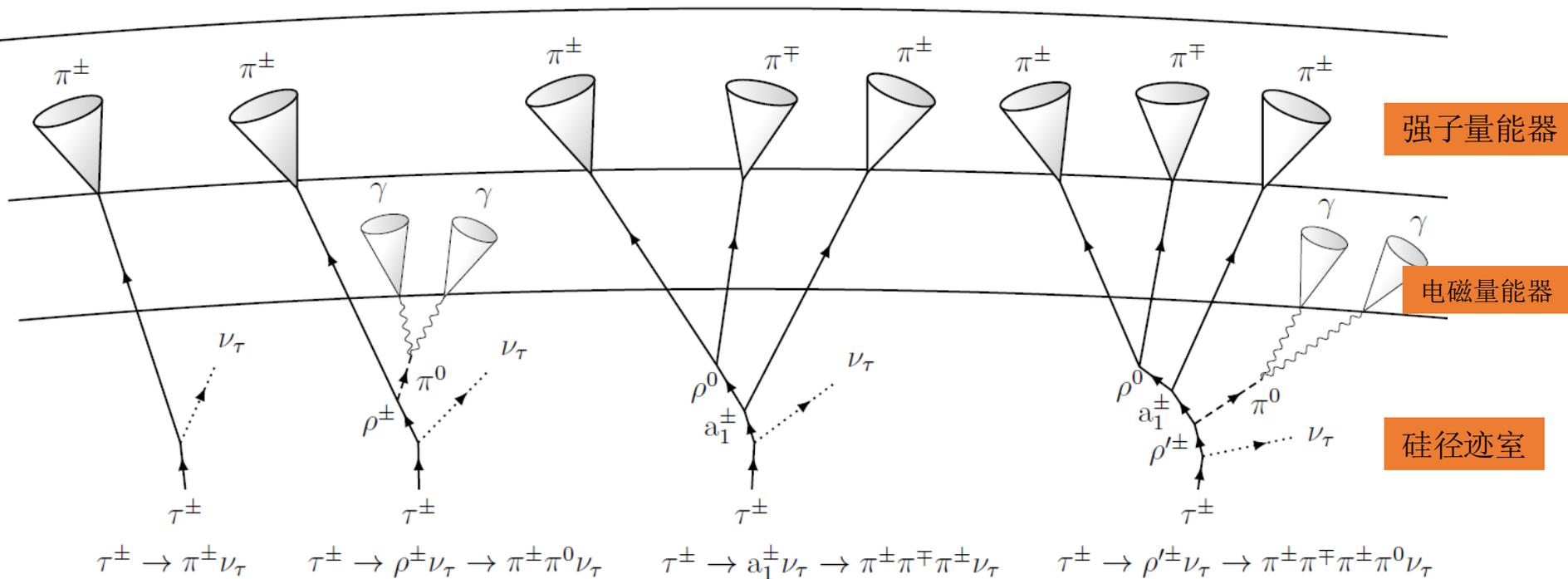


► τ 轻子是唯一能衰变为强子的轻子，其衰变道见下表

	Decay mode	Resonance	\mathcal{B} (%)
Leptonic decays			
≈ 1/3 of the decays Covered by the standard e/μ algorithms	$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$		17.8
	$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$		17.4
Hadronic decays			
≈ 2/3 of the decays τ_h, or sometimes just “τ” • dedicated L1 and HLT triggers • dedicated reconstruction algorithms • dedicated analysis methods	$\tau^- \rightarrow h^- \nu_\tau$		11.5
	$\tau^- \rightarrow h^- \pi^0 \nu_\tau$	$\rho(770)$	25.9
	$\tau^- \rightarrow h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	9.5
	$\tau^- \rightarrow h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
	$\tau^- \rightarrow h^- h^+ h^- \pi^0 \nu_\tau$		4.8
	Other		3.3

- τ 子的强子衰变分支比较大
- τ 子的强衰变包含一个带电粒子衰变道，三个带电粒子衰变道，这里 h^\pm 代表一个带电 π 介子或者是K介子

重建 τ 子的强衰变道对于CMS实验是一个巨大的挑战

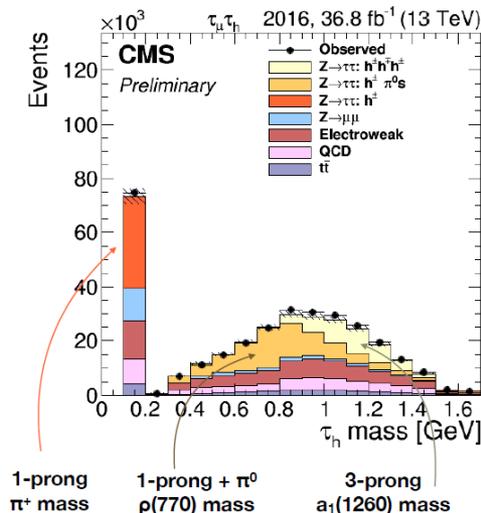
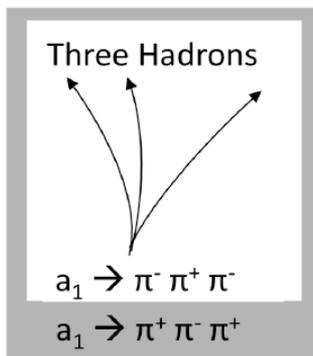
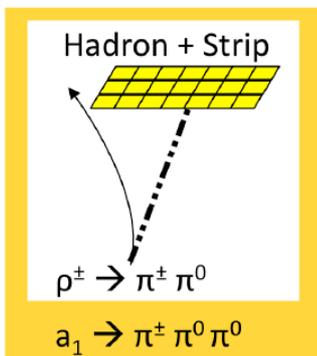
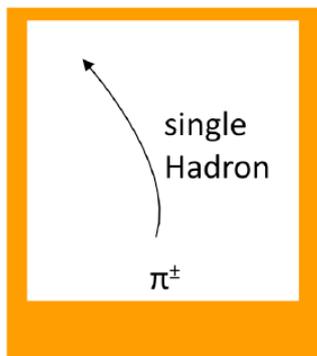


τ 子的强衰变的衰变产物与探测器相互作用的示意图

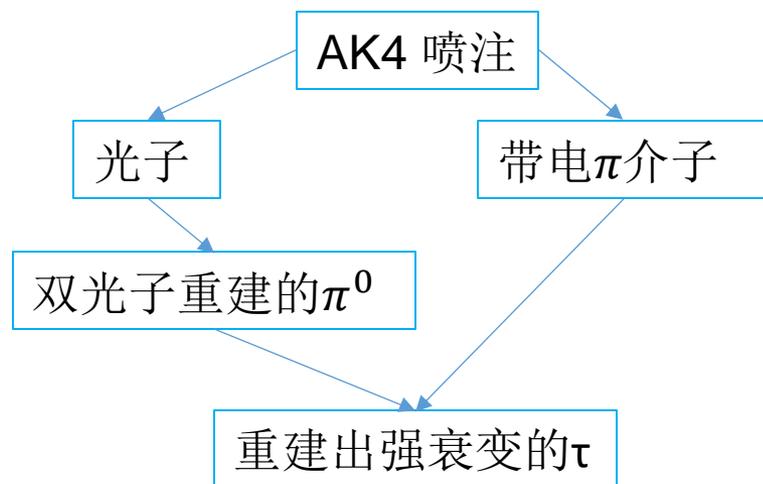
问题：如何重建强衰变的 τ 子



在CMS探测器上对于强衰变的 τ 的重建使用的是叫做hadrons-plus-strips(HPS)算法



- 1.调查一个AK4喷注的粒子流成分
- 2.通过电磁量能器通过粒子流的候选光子重建出 π^0
- 3.把 π^0 和粒子流算法重建的带电强子组合在一起,推断出 τ_h 的衰变模式

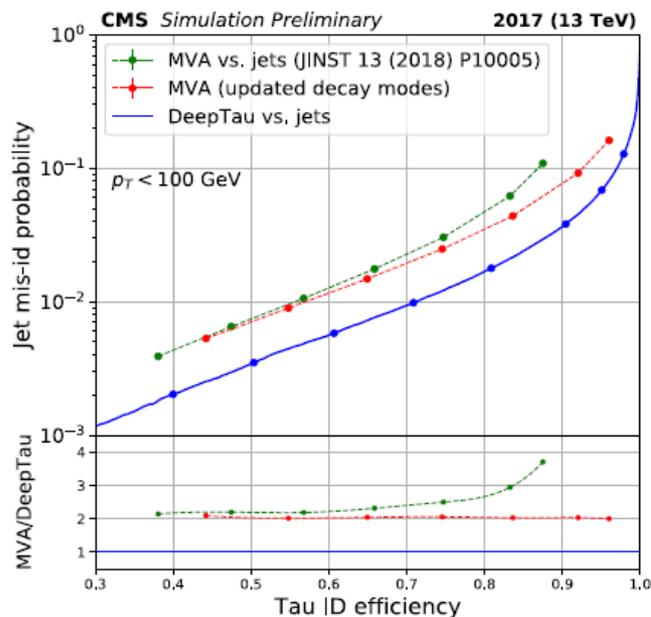
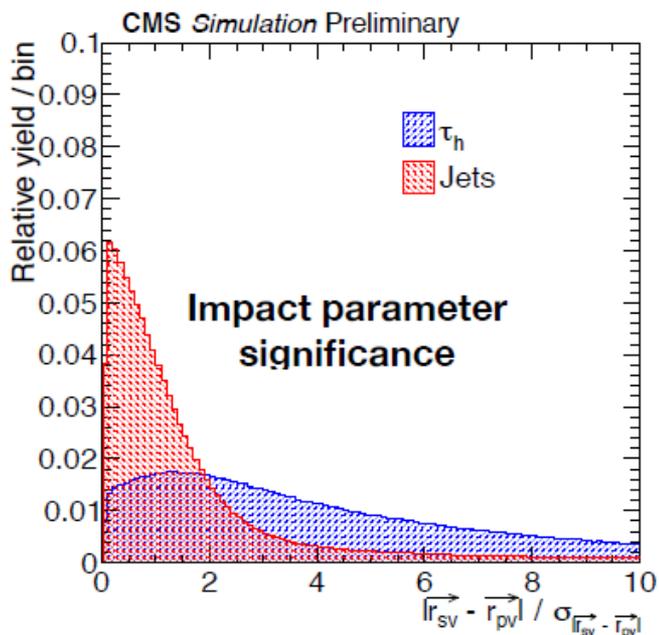
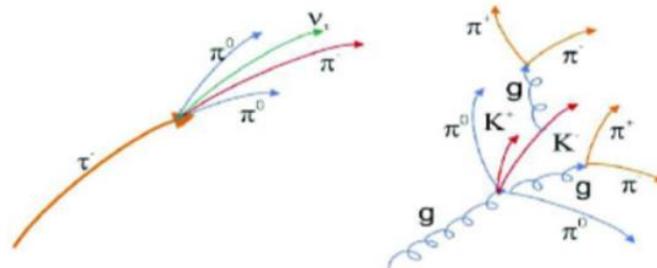


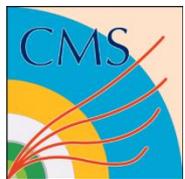


问题：如何区分 τ_h 和喷注

➤ CMS Run2开发了两种鉴别算法：

1. 基于决策树的多变量分析
2. 基于卷积深度神经网络(convolutional deep neural network)





总结



➤你学到了什么：

- 喷注的来源，喷注重建的方法和**CMS**最常用的喷注类型
- 根据喷注的大小，子结构，质量信息来标定不同类型的喷注
- 粒子堆积效应对于喷注能量的影响，及如何来压低这种影响
- 喷注能量的修正

- τ 子的基本性质和衰变道
- τ 子强衰变的重建和鉴别

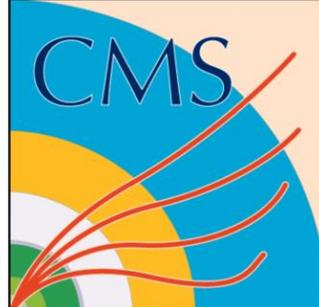
请大家进行hand-on的相关练习：

https://github.com/TaozheYu/CMS_winter_school

谢谢北京航空航天大学提供这样一个学习平台，感谢各位学员和老师的聆听



北京航空航天大学
BEIHANG UNIVERSITY

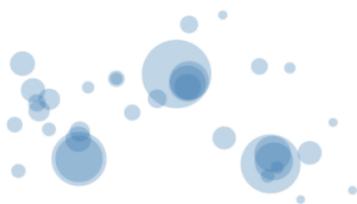


backup



➤ Particle Net

- 一种近些年在CMS分析中流行的用机器学习的方法对AK8 jet进行标定的新方法
- 把jet当做二维平面的“粒子云”，用卷积神经网络对其进行处理



- 目前CMS的NanoAOD数据类型已经包含了Particle Net:
FatJet_particleNet*
- 我们可以用它来区分:QCD 和top, W, Z,

Mistagging Rate	WP for Top	WP for W	WP for W (MD)
5%	-	0.68	-
2.5%	-	-	0.64
1%	0.49	0.94	0.85
0.5%	0.74	0.97	0.91
0.1%	0.96	-	-



➤ 粒子堆积修正

原因:

MC的粒子堆积分布



泊松分布

数据的粒子堆积分布



方法:

对比数据和MC的粒子堆积分布→修正因子

