



# ATLAS 上通过双轻子道寻找HH 和漂移室的粒子鉴别算法研究

---

学生 : 辛水艇  
指导老师: 娄辛丑  
学位类别: 理学博士  
学科专业: 粒子物理与原子核物理  
培养单位: 中国科学院高能物理研究所

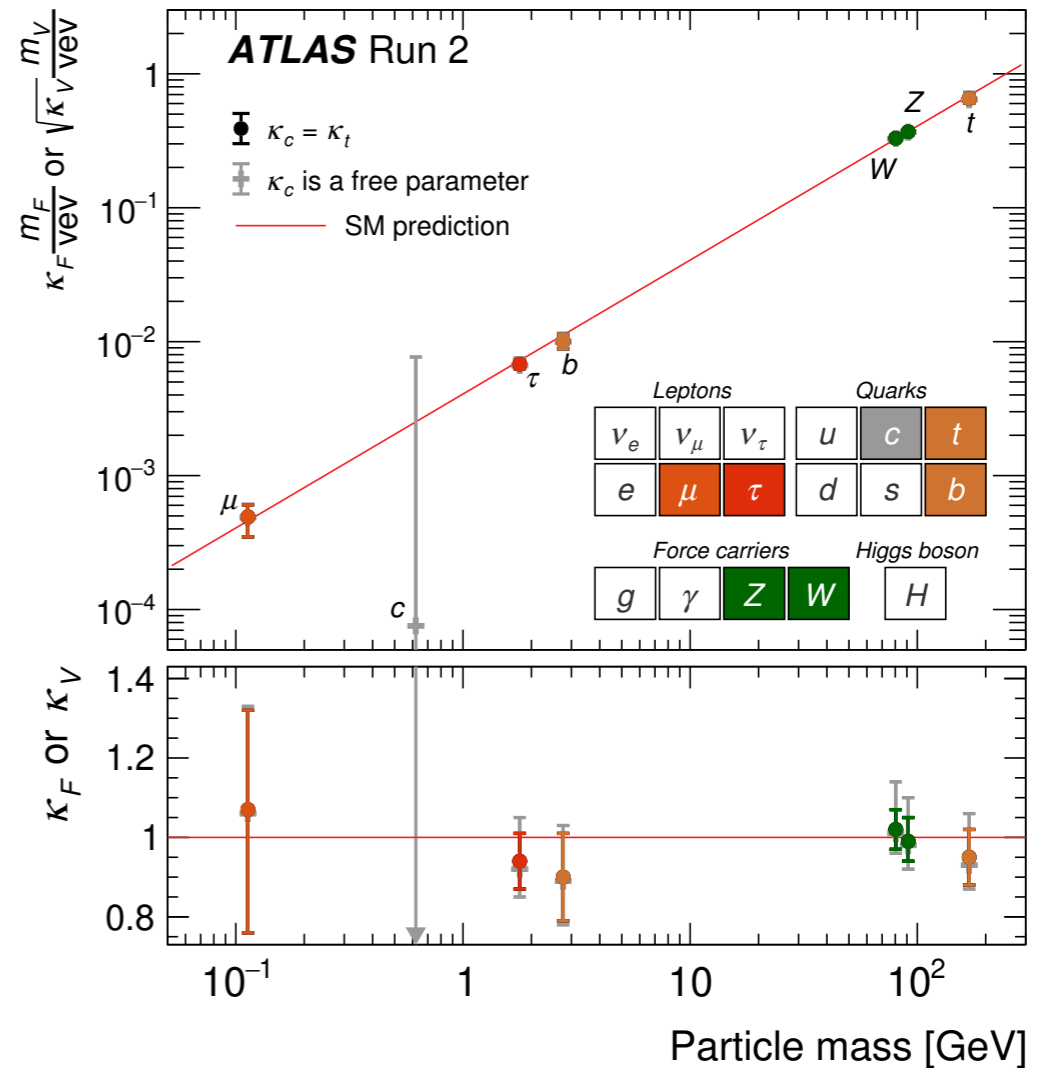
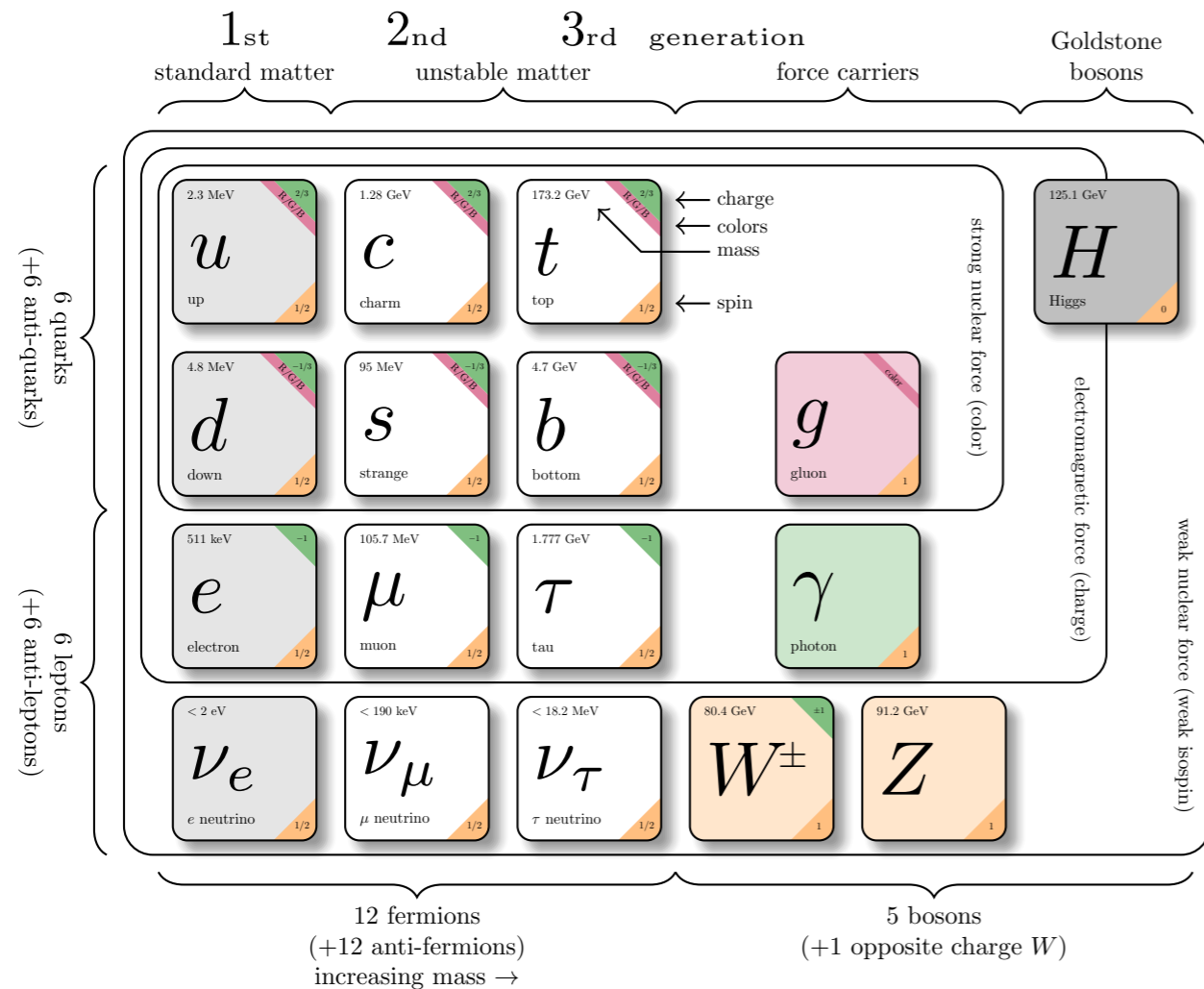
# 目录

---

- ▶ 研究背景
  - ▶ 标准模型和di-Higgs物理
  - ▶ ATLAS 探测器
- ▶ ATLAS 上di-Higgs 研究
  - ▶ 通过同电荷双轻子寻找标准模型非共振双希格斯粒子
  - ▶ 多轻子联合测量
- ▶ 漂移室的粒子鉴别算法研究
  - ▶ 原初电离计数（Cluster counting）研究背景
  - ▶ 基于BESIII 漂移室的粒子鉴别研究
  - ▶ Cluster counting 技术用于CEPC 漂移室
  - ▶ 束流测试
- ▶ 总结与展望
- ▶ 附文
  - ▶ 已取得研究成果
  - ▶ 主要贡献
  - ▶ 论文审阅建议

# 标准模型和希格斯物理

- ▶ 标准模型 (SM) 是准确涵盖基本粒子及其相互作用的成功理论。
  - ▶ 统一地描述了电弱理论和量子色动力学。
  - ▶ 希格斯场触发电弱对称性自发破缺 (EWSB)，是基本粒子质量的来源。



- ▶ 自2012年发现Higgs以来，对它的理解取得了巨大进展。产生过程 ( $> 5\sigma$ )、规范玻色子 ( $> 5\sigma$ )、费米子耦合模式 ( $H \rightarrow cc$  首次测量,  $H \rightarrow \mu\mu$   $3\sigma$ )、Higgs质量 ( $\sim 0.1\%$ )、 $J^{CP}$  .....

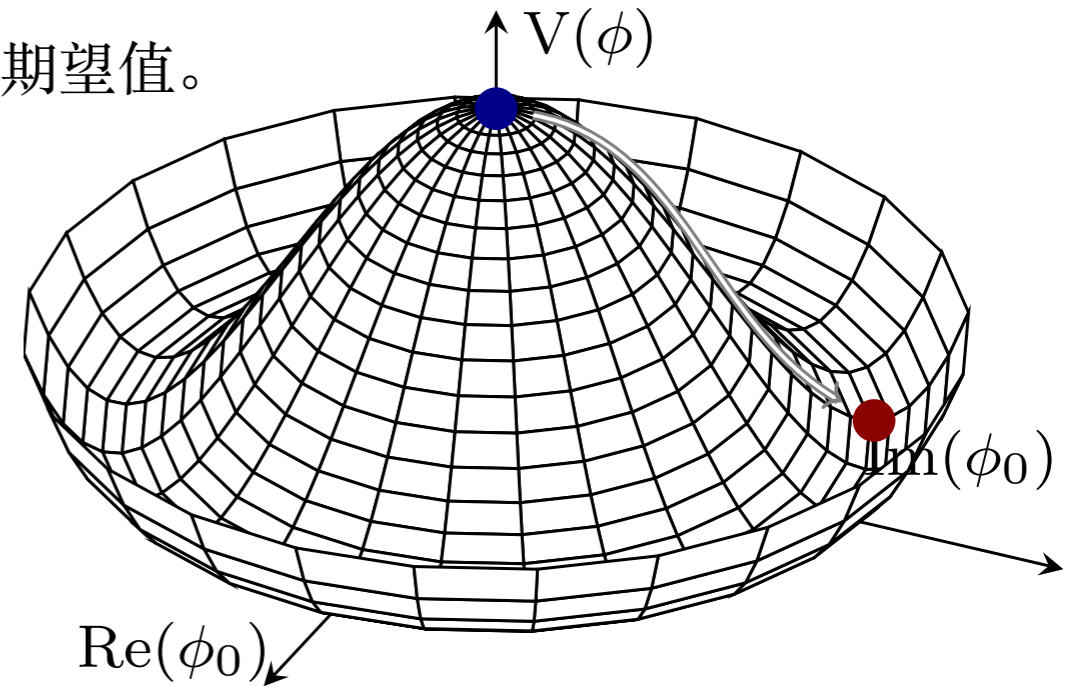
# 研究动机

- 电弱对称性破缺后，Higgs 势能移动到最小值，即真空期望值。

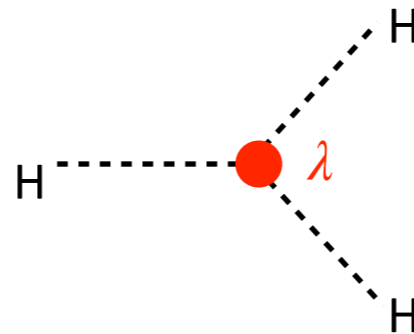
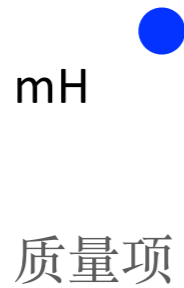
$$V(\Phi) = -\mu^2(\Phi^\dagger\Phi) + \lambda(\Phi^\dagger\Phi)^2$$

↓

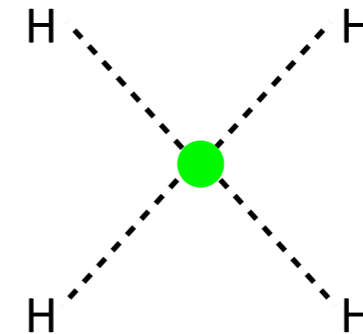
$$|\Phi|_{\min} = \sqrt{\frac{\mu^2}{2\lambda}}, \mu^2 = \lambda v^2$$



$$V(H) \stackrel{\Phi \rightarrow H+v}{\supset} \frac{1}{2} m_H^2 H^2 + \lambda v H^3 + \frac{1}{4} \lambda H^4 \dots$$



SM: 自耦合被  $m_H, v$  确定，  
需要实验测量和证实。



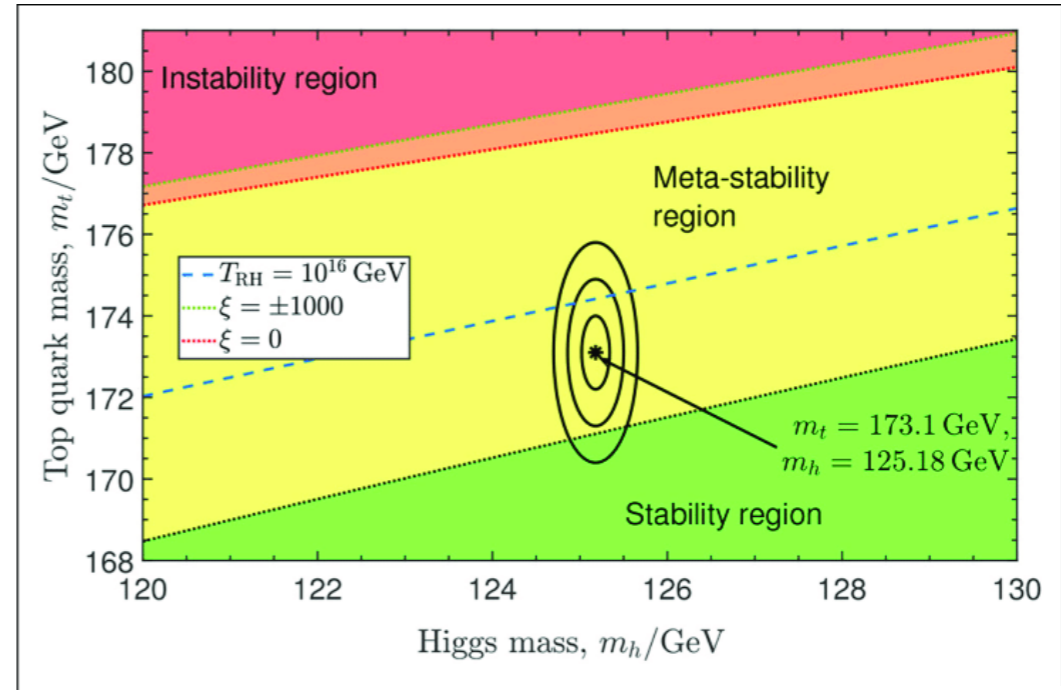
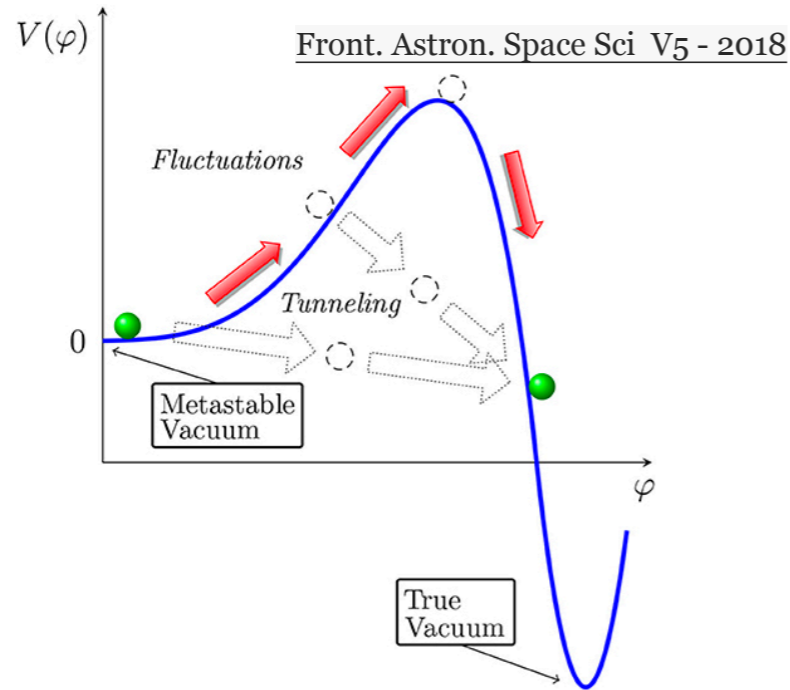
四线耦合

- Higgs 势只被测量了一部分，它的形状完全决定了标量场的性质。
- 测量 Higgs 自耦合是确定 Higgs 势能形状的关键。



# 研究动机: Higgs 势

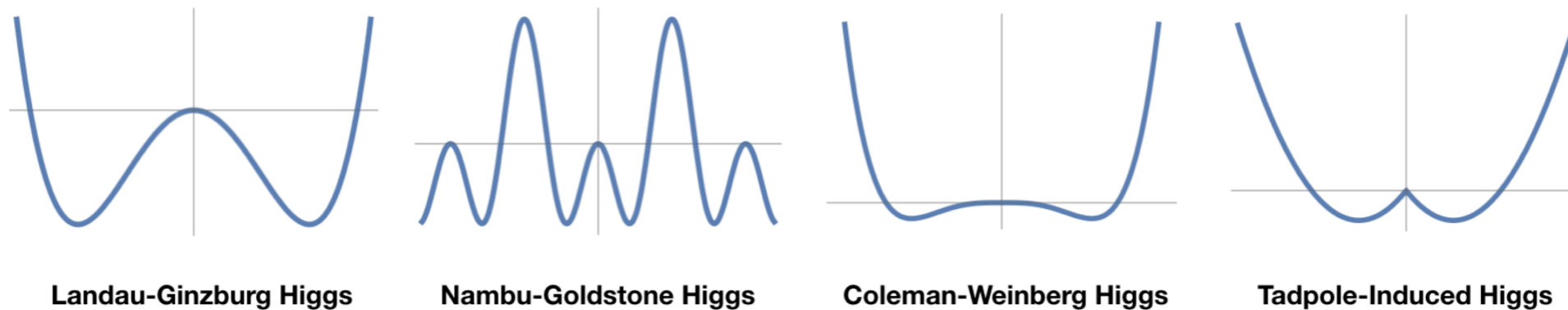
- ▶ 真空稳定性问题。当前宇宙的真空期望值是局部最小还是全局最小?



10.3389/fspas.2018.00040

- ▶ 电弱相变的本质仍然未知。
  - ▶ 在早期宇宙演化中，重子的产生要求一级电弱相变，这意味着与Higgs相互作用的新物理，可引起  $O(1)$  的Higgs自耦合改变。
- ▶ 希格斯玻色子对 (HH) 的测量可以区分不同的Higgs势能模型。

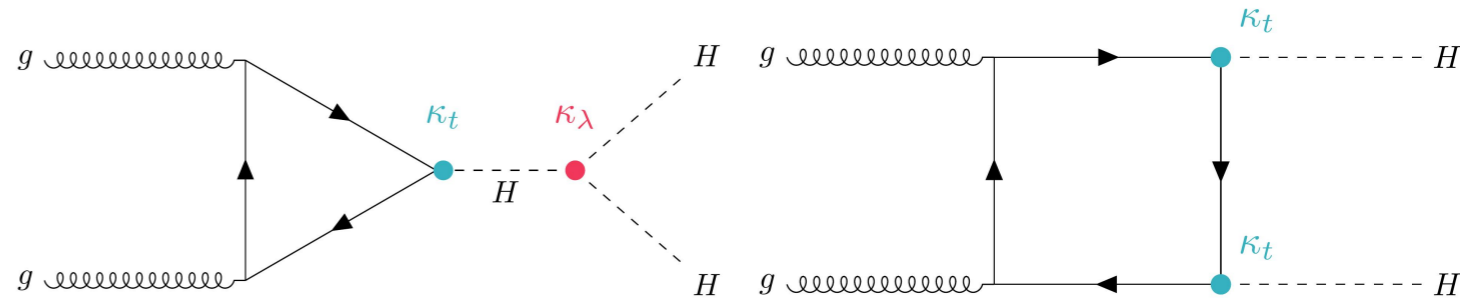
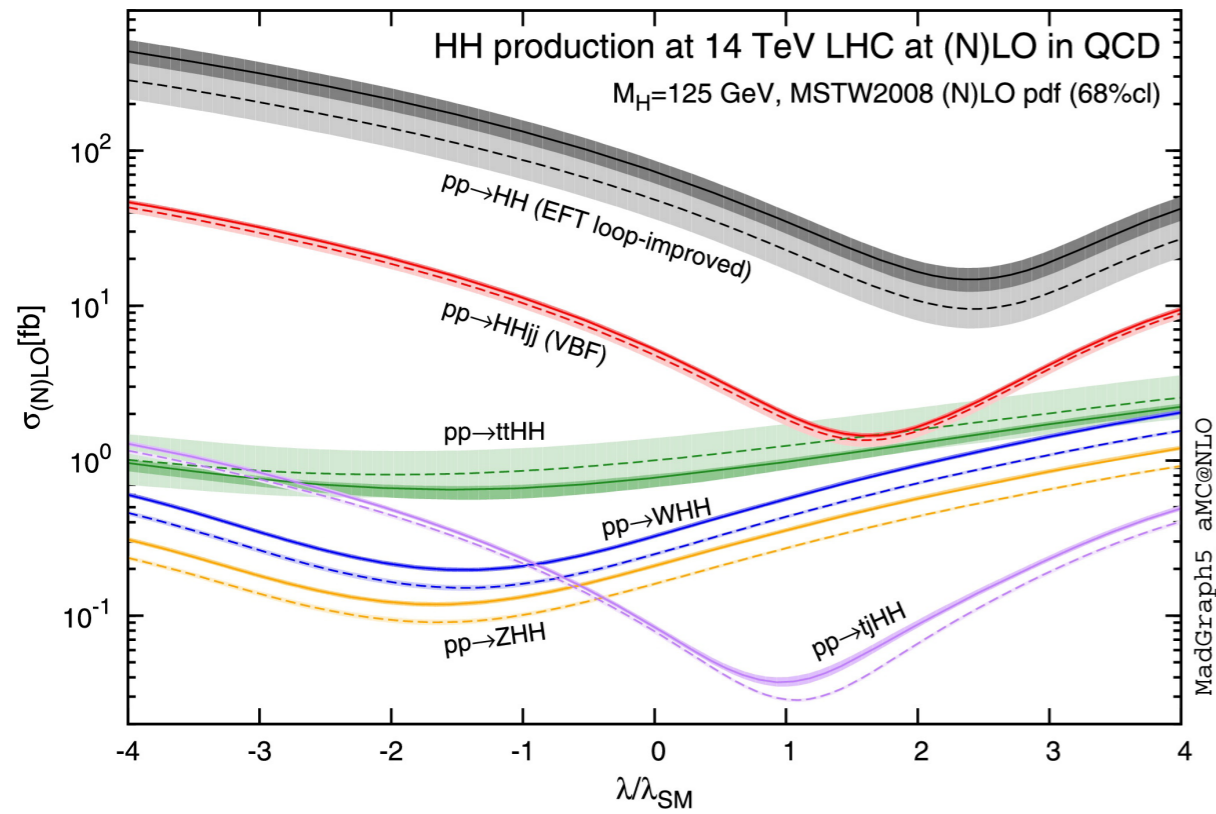
Phys. Rev. D 78, 063518



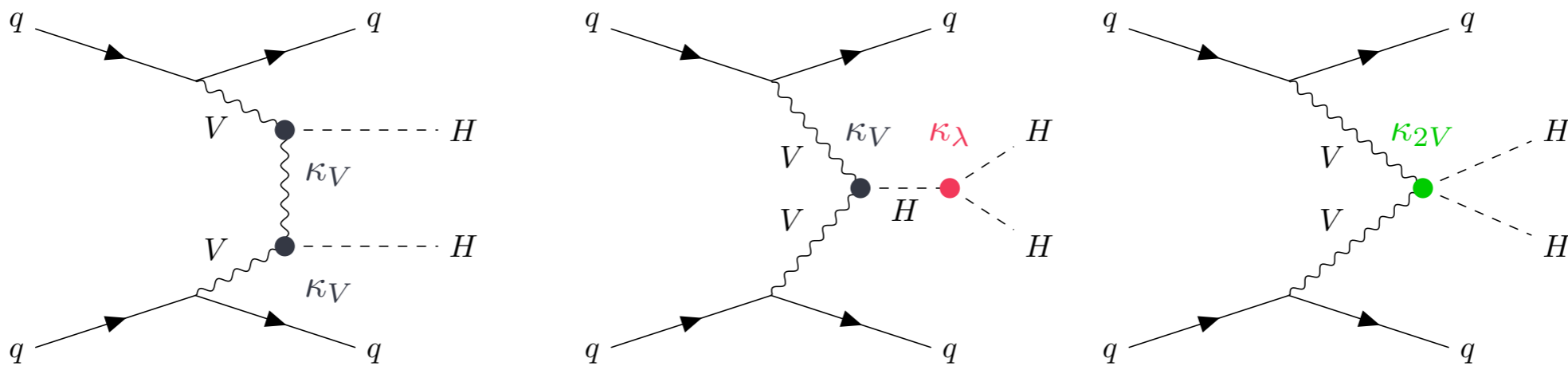
# LHC上HH产生过程

► 本文研究LHC上产生HH的主要模式ggF和VBF。  
VHH, ttHH过程更加稀有

► 胶子胶子融合 (ggF) :  $\sigma_{ggF}^{SM} \approx 31.05 \text{ fb [13TeV]}$



► 矢量玻色子融合 (VBF) :  $\sigma_{VBF}^{SM} \approx 1.7 \text{ fb [13TeV]}$

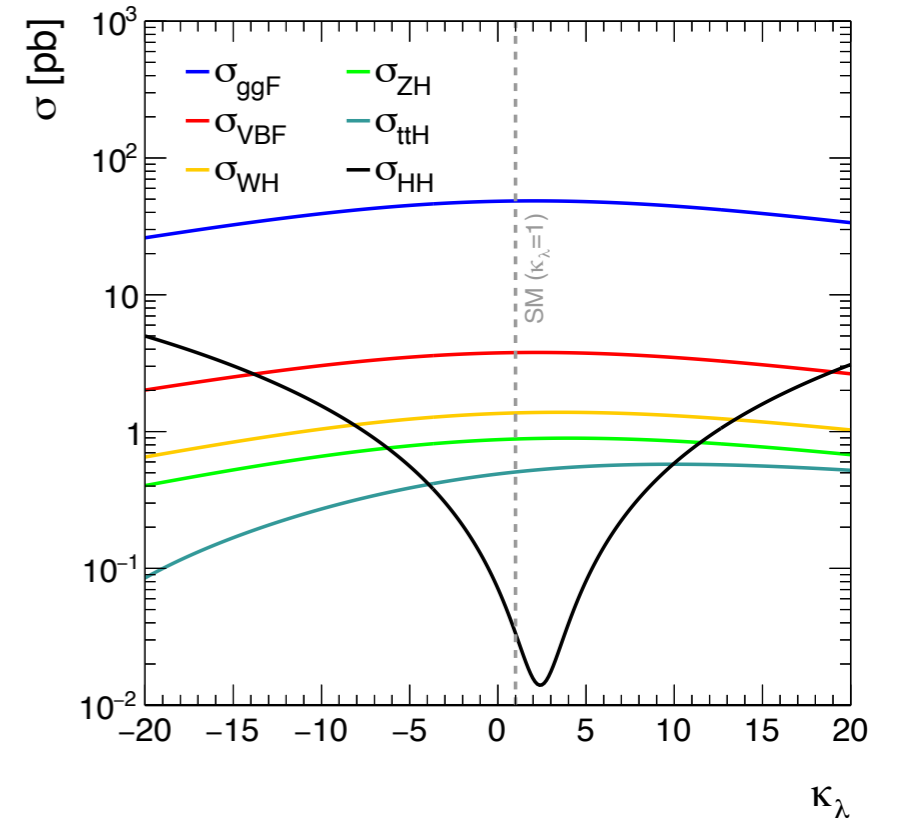
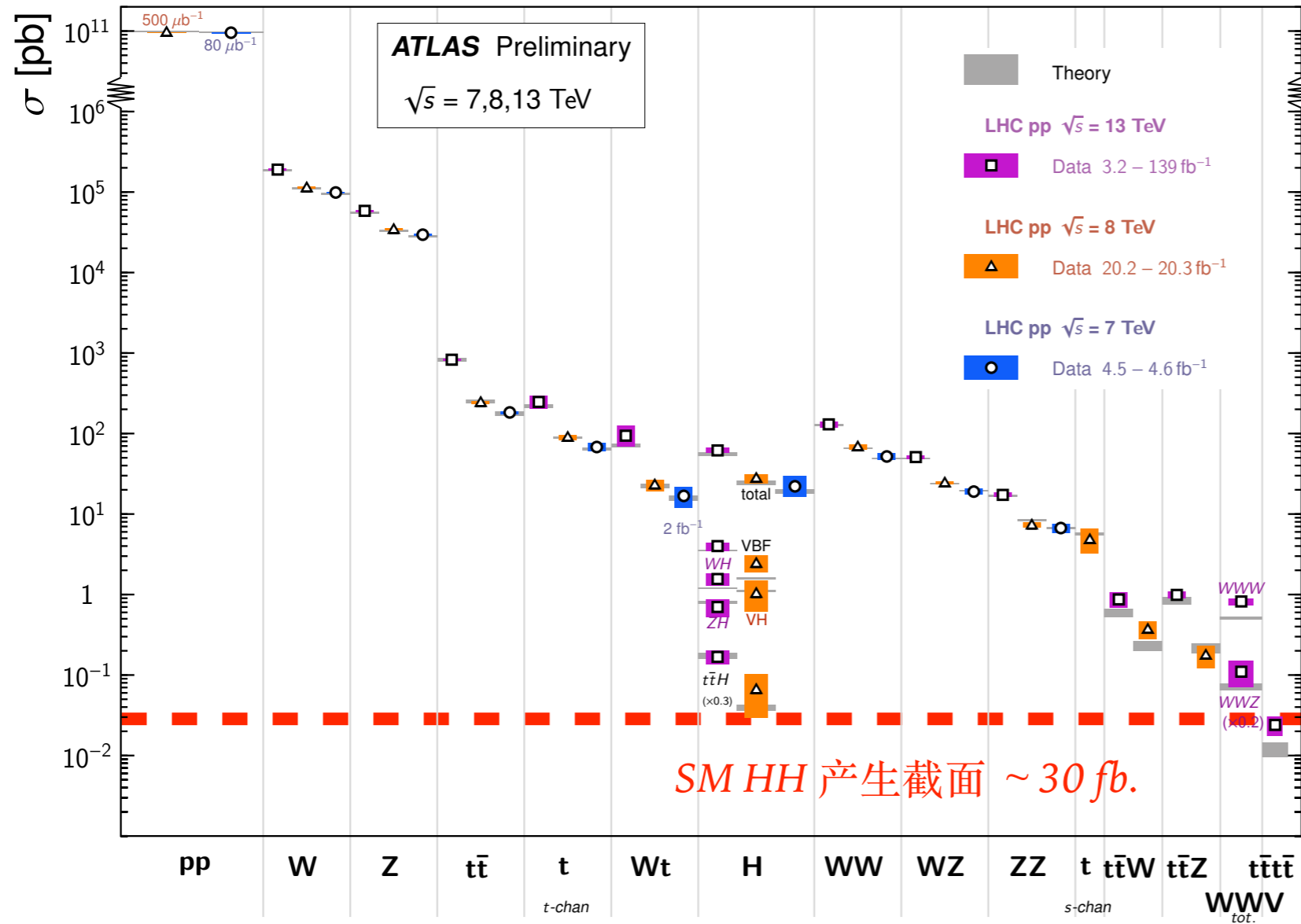


# HH产生截面

- ▶ SM HH的产生截面要比Higgs 小  $\sim 1000$  倍。

Standard Model Total Production Cross Section Measurements

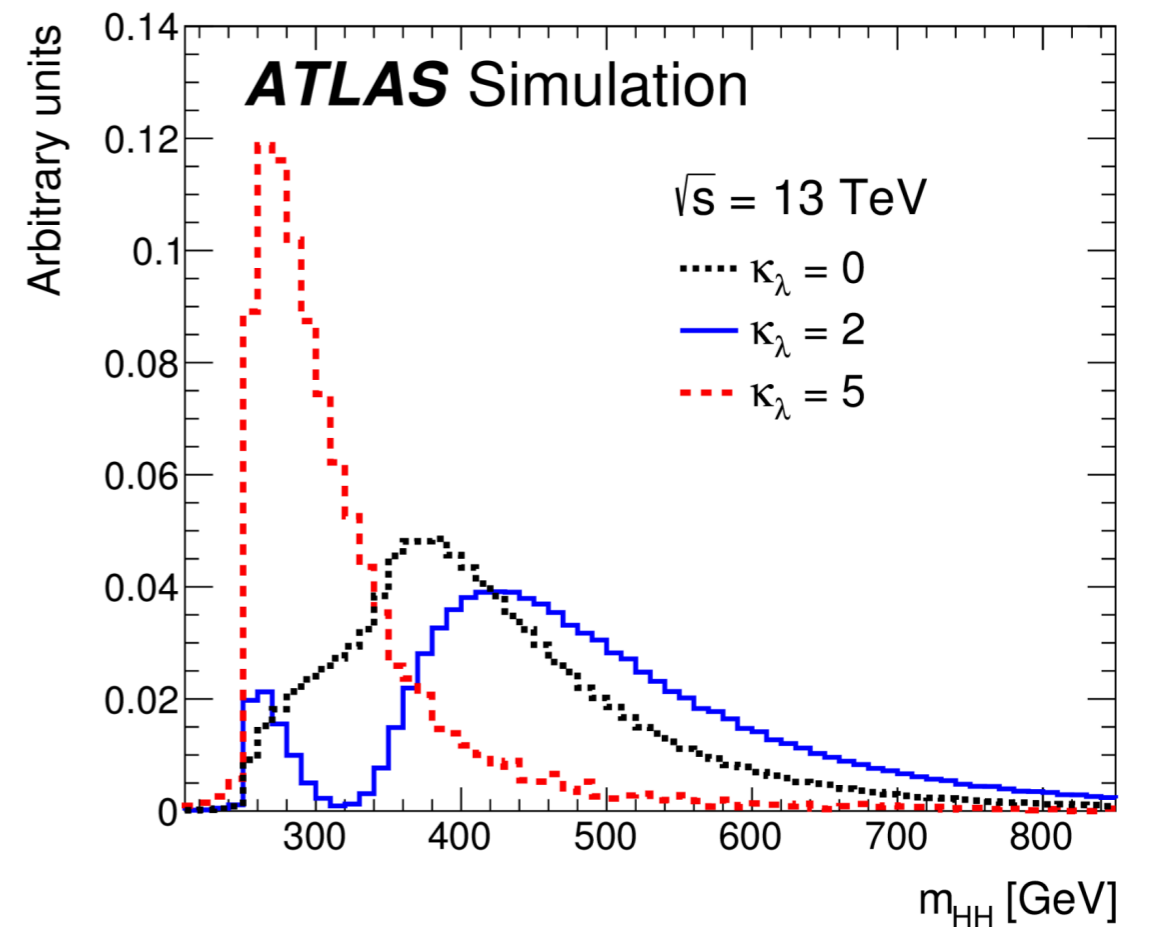
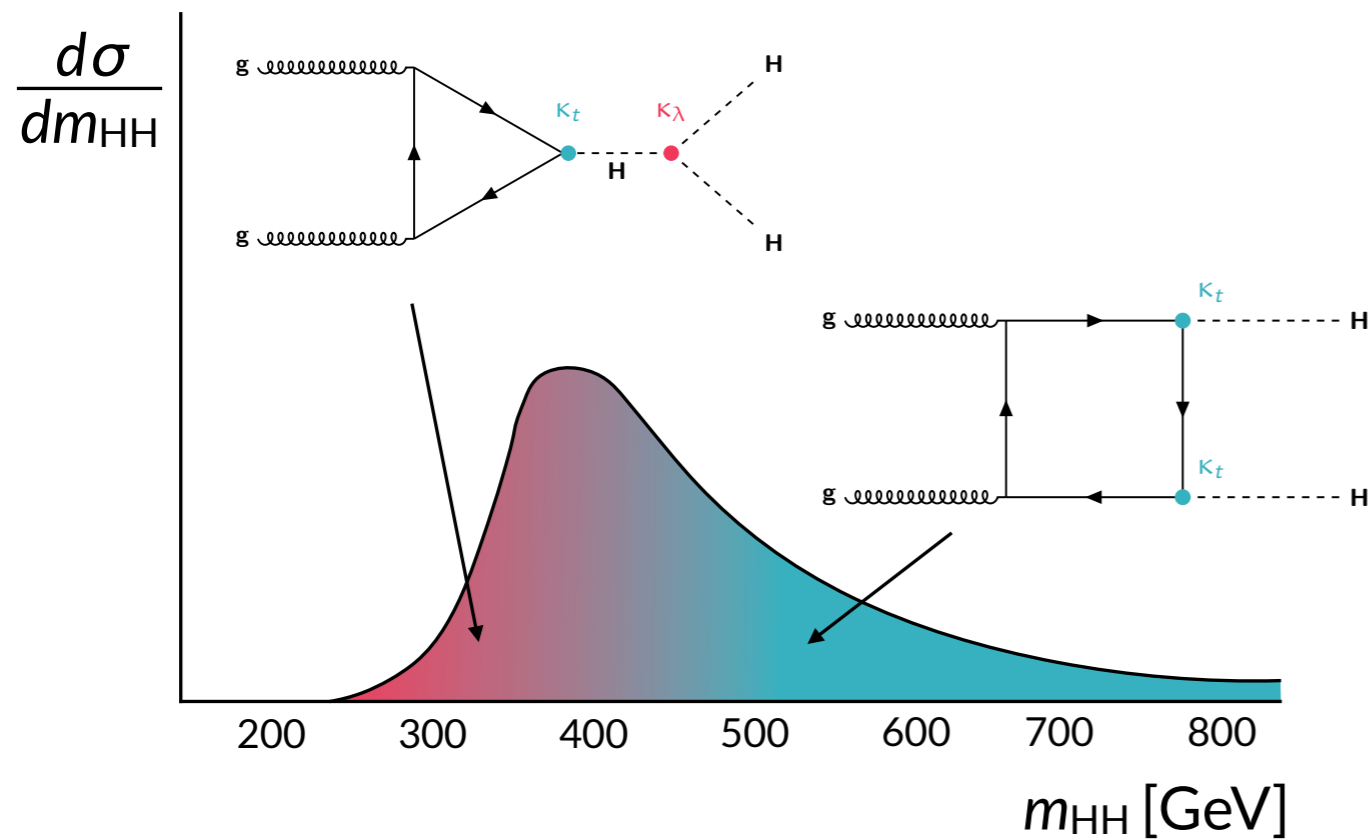
Status: February 2022



- ▶ 单Higgs 过程在LO不依赖于 $\kappa_\lambda$ ，然而NLO电弱圈图可引入 $\kappa_\lambda$  对Higgs 产生模式的影响。
- ▶  $\kappa_\lambda$ 的测量以研究HH过程为主导。

# HH产生截面特征

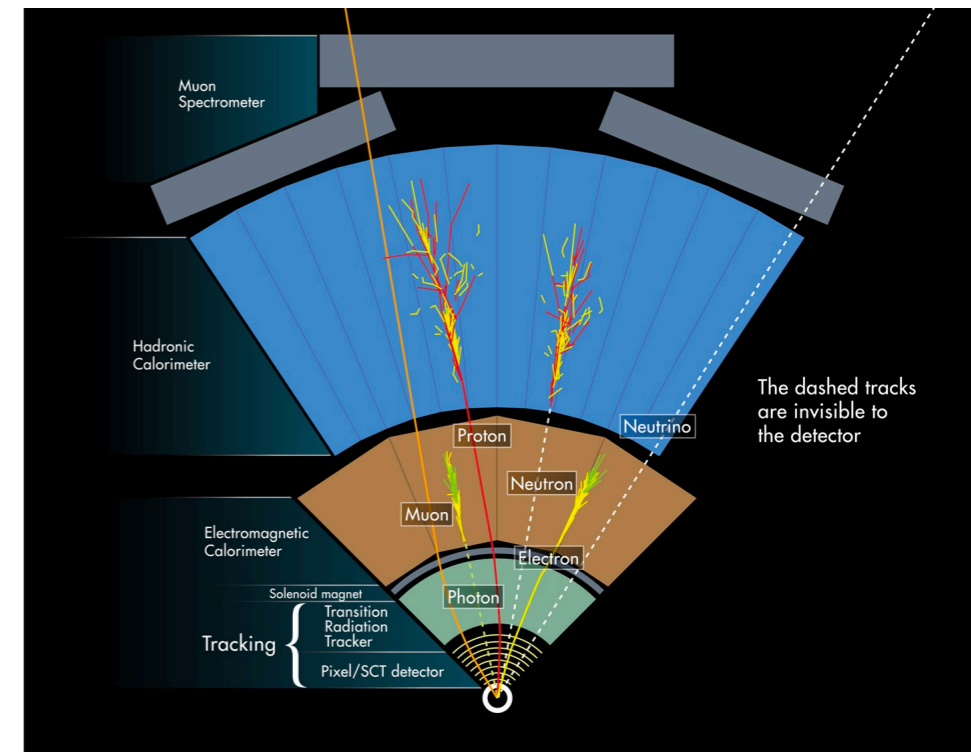
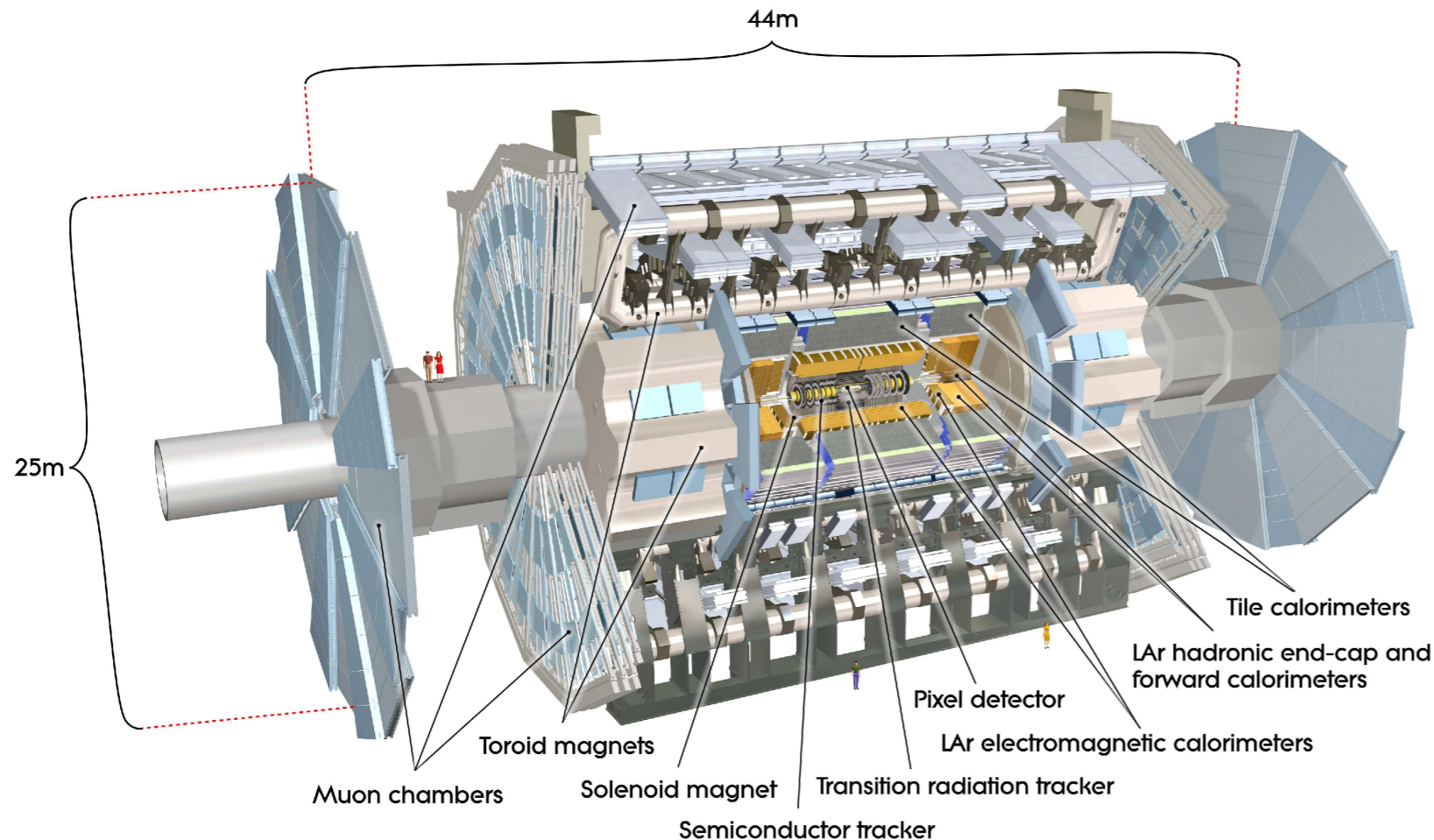
- ▶ 三角图和盒子图振幅的破坏性干涉导致了很小的截面，在 $\kappa_\lambda = 2.4$  干涉效应最显著。
- ▶  $m_{HH}$  谱随着 $\kappa_\lambda$  有较大变化。大的 $\kappa_\lambda$ 所对应的运动学很“平缓”。
- ▶ 测量HH产生截面限制 $\kappa_\lambda$  参数空间。测量若存在偏移可以是新物理的表现。





# ATLAS 探测器

- ▶ ATLAS (A Toroidal LHC apparatus)是大型强子对撞机 (LHC) 上通用型探测器。前后对称的圆筒形。
- ▶ ATLAS由子探测器负责不同测量。



- ▶ 内部径迹探测器：精确测量带电径迹、动量、顶点。
- ▶ 量能器系统：通过能量吸收测量电子，光子，*jet*能量。
- ▶ 缪子谱仪：在磁铁外部精确测量和鉴别缪子。

- 
- ▶ 研究背景
    - ▶ 标准模型和di-Higgs物理
    - ▶ ATLAS 探测器
  - ▶ ATLAS 上di-Higgs 研究
    - ▶ 通过同电荷双轻子寻找标准模型非共振双希格斯粒子
    - ▶ 多轻子联合测量
  - ▶ 漂移室的粒子鉴别算法研究
    - ▶ 原初电离计数 (Cluster counting) 研究背景
    - ▶ 基于BESIII 漂移室的粒子鉴别研究
    - ▶ Cluster counting 技术用于CEPC 漂移室
    - ▶ 束流测试
  - ▶ 总结与展望
  - ▶ 附文
    - ▶ 已取得研究成果
    - ▶ 主要贡献
    - ▶ 论文审阅建议



# 多轻子分析的HH 衰变模式

- ▶ 较低的截面，但受益于RUN2统计量，开拓额外的衰变道以寻找HH。
  - ▶ 衰变过程包含  $WWWW$ ,  $WWZZ$ ,  $WW\tau\tau\dots$  占HH衰变的  $\sim 12\%$
- ▶ 丰富的现象学，具有多种末态。
  - ▶ 按照末态划分分析策略，不关心中间产生过程。信号复杂的组合方式适合利用MVA方法分析。
- ▶ 多轻子末态：
  - ▶ 轻子在ATLAS上是较好的重建对象，在排除  $H \rightarrow bb$  信号中最佳的选择之一。
  - ▶ 双光子加上多轻子末态作为多轻子分析重要贡献。
  - ▶  $bbZZ$  (4l) ，预期有不错的信号区分能力。

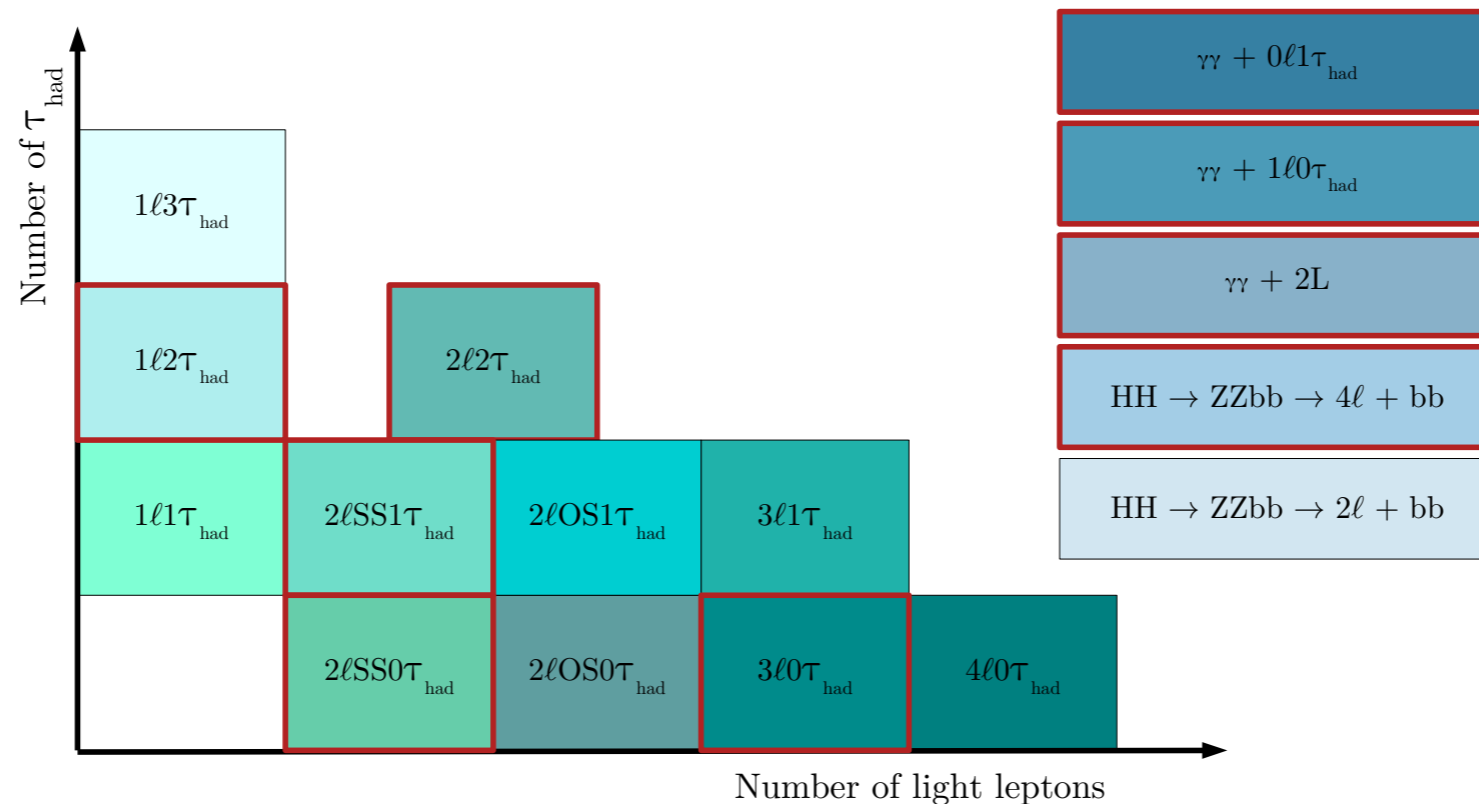
	bb	WW	$\tau\tau$	ZZ	$\gamma\gamma$
bb	34%				
WW	25%	4.6%			
$\tau\tau$	7.3%	2.7%	0.39%		
ZZ	3.1%	1.1%	0.33%	0.069%	
$\gamma\gamma$	0.26%	0.10%	0.028%	0.012%	0.0005%

- ATLAS**
- $bbVV$  (2l)
  - $WWWW$  (36.1 -1fb)
  - $WW\gamma\gamma$  (36.1 -1fb)

- CMS**
- $bbWW$  (2l2v, 36.1-1fb)
  - $bbZZ$  (4l)
  - $WWWW$ ,  $WW\tau\tau$ ,  $\tau\tau\tau\tau$
  - $WW\gamma\gamma$

# HH 多轻子分析

- ▶ 根据末态粒子个数，划分为不同物理道。包括6个轻子道和3个双光子道。
  - ▶  $\gamma\gamma + ML$  : 2 光子和轻子 ( $e, \mu, \tau_h$ ) ,  $\gamma\gamma + 1\ell 0\tau$ ,  $\gamma\gamma + 0\ell 1\tau$ ,  $\gamma\gamma + 2L$ .
  - ▶ 同电荷双轻子 有/无  $\tau_h$ :  $2\ell SS0\tau$  和  $2\ell SS1\tau$
  - ▶ 3个轻子无 $\tau_h$ :  $3\ell$
  - ▶ 一个/两个轻子 以及两个 $\tau_h$ :  $1\ell + 2\tau_h$  and  $2\ell + 2\tau_h$
  - ▶  $HH \rightarrow ZZb\bar{b}$ : 从  $H \rightarrow ZZ$  衰变而来的4个轻子及 2个 b-jets:  $b\bar{b}4\ell$



- ▶  $2\ell SS0\tau$  (2LSS) 是本文阐述的物理分析，是分析复杂度最高，对截面测量主要敏感的衰变道。

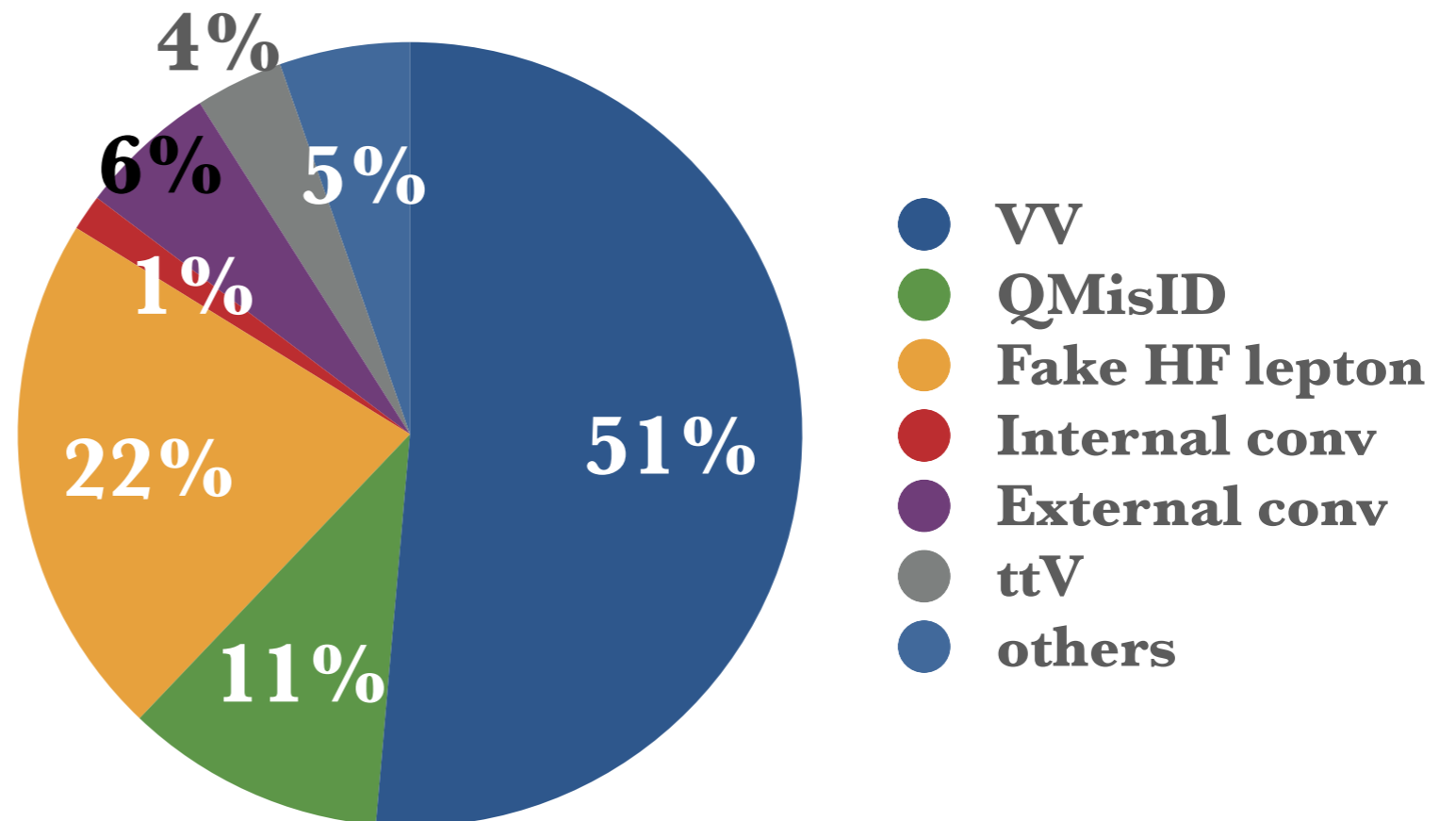
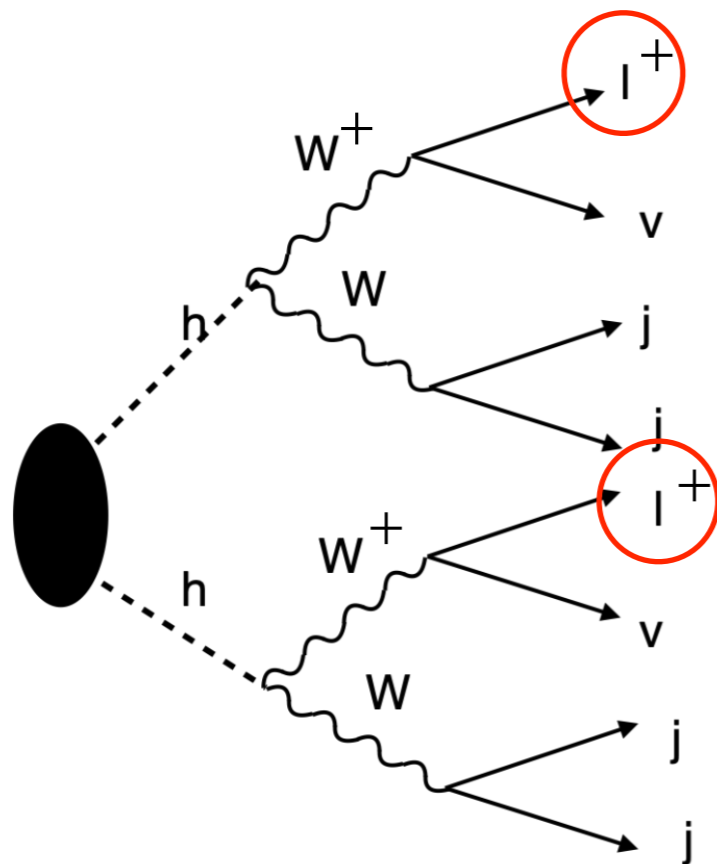
# 多轻子分析总览

- ▶ 轻子道尽可能采用同种物理对象定义和分析策略。
  - ▶ 在通过事例初选后，为每个物理道定义多变量（BDTG），作为最后的拟合变量。
  - ▶ 本底建模：
    - ▶ 不可约本底：MC估计，在CR中贡献程度大的利用数据驱动修正。
    - ▶ 可约本底：利用数据驱动方法估计
      - ▶ 模版拟合（TF）方法：假轻子本底。Fake factor方法作为交叉验证。
      - ▶ Fake factor 及 scale factor方法：假 $\tau$ 本底。
- ▶ yyML 道：利用 myy 的边带区做连续本底估计。不同BDT区间的myy 作为最后的拟合变量。
- ▶ 同时拟合所有的信号区间和各个物理道的控制区间（binned fit）。

	$2\ell SS$	$2\ell SS+1\tau_{had}$	$3\ell$	$b\bar{b}4\ell$	$1\ell+2\tau_{had}$	$2\ell+2\tau_{had}$	$\gamma\gamma+1\ell 0\tau_{had}$	$\gamma\gamma+0\ell 1\tau_{had}$	$\gamma\gamma+2L$
BDT trained against	VV, $t\bar{t}$ and V+jets separately	VV	total background	total background	VV	VV	MC total background		
Discriminant	Combined BDT	BDT	BDT	BDT	BDT	BDT	$m_{\gamma\gamma}$ in BDT regions		
Control regions	5	3	4	4	1	1	$m_{\gamma\gamma}$ in sideband region		
Non-prompt lepton strategy	TF	TF	TF	TF	-	-	-	-	-
Fake tau strategy	-	SF	-	-	FF	FF	-	-	-

# 2LSS分析总览

- ▶ 末态特征：同电荷双轻子，无b-jet，多个普通jet，SM过程本底被有效压低。
  - ▶ 以WWWW衰变模式为主导，占衰变产额的60%。
  - ▶ 分析以ggF信号特征为优化目标。
- ▶ 信号提取：利用事例的运动学，采用了基于决策树（BDT）多变量分析。
  - ▶ 联合多维BDT作为最终区分变量。
- ▶ 本底分析：
  - ▶ 对双玻色子（VV），假轻子（fake）和电荷误判（QmisID）做特定的本底建模。



# 蒙卡和数据样本

- ▶ ATLAS 2015-2018 期间完整RUN2阶段的139-1fb 收集数据。
- ▶ 蒙卡:
  - ▶ SM di-Higgs (ggF 及 VBF)  $\kappa_\lambda = 1, 10$ 。
  - ▶ 本底: VV, ttW, ttZ, ttbar, Zjets, Wjets, single Higgs 等。

Process	Generator	ME order	Process	Precision order	Cross section central value	Cross section uncertainty	Modelling uncertainty	Normalized to data
$t\bar{t}W$	SHERPA 2.2.10 (MG5_AMC)	NLO (NLO)	MC samples contributing to fake lepton templates					
$t\bar{t}t\bar{t}$	MG5_AMC (SHERPA 2.2.10)	NLO (NLO)	$t\bar{t}$	NNLO+NNLL	832 pb	-	alternative MC	Yes
$t\bar{t}H$	POWHEG-BOX [54] (Powheg-BOX) (MG5_AMC)	NLO (NLO) (NLO)	$s$ -, $t$ -channel single top	NLO	227 pb	-	-	Yes
$t\bar{t}(Z/\gamma^* \rightarrow l^+l^-)$	SHERPA 2.2.11 (MG5_AMC)	NLO (NLO)	$Wt$ single top	NNLO approx	71.7 pb	-	-	Yes
$t\bar{t} \rightarrow W^+bW^- \bar{b}l^+l^-$	MG5_AMC	LO	$Z/\gamma^* \rightarrow l^+l^-$	NNLO	0.9751×SHERPA	-	-	-
$t(Z/\gamma^*)$	MG5_AMC	NLO	$W \rightarrow \ell\nu$	NNLO	0.9751×SHERPA	-	-	-
$tW(Z/\gamma^*)$	MG5_AMC	NLO	MC samples of irreducible background processes					
$t\bar{t}W^+W^-$	MG5_AMC	LO	$t\bar{t}W$	NLO	601 fb	-	alternative MC scale variations	Yes
$t\bar{t}$	POWHEG-BOX (POWHEG-BOX)	NLO (NLO)	$t\bar{t}t\bar{t}$	NLO	12 fb	20%	alternative MC	No
$t\bar{t}t$	MG5_AMC	LO	$t\bar{t}(Z/\gamma^* \rightarrow l^+l^-)$	NLO	839 fb	-	alternative MC scale variations	Yes
$s$ -, $t$ -channel, $Wt$ single top	POWHEG-BOX [56, 57]	NLO	$t\bar{t}H$	NLO	507 fb	11%	alternative MC scale variations	No
VV, qqVV, $lowm\ell\ell, VVV$	SHERPA 2.2.2	NLO	$WH$	NLO	1102 fb	-	scale variations	No
$Z \rightarrow l^+l^-$	SHERPA 2.2.1	NLO	$ZH$	NLO	601 fb	-	scale variations	No
$Z \rightarrow l^+l^-$ (matCO)	POWHEG-BOX	NLO	VV, qqVV	NLO	SHERPA	-	10% (+LF jets), scale variations	Yes (+HF jets)
$Z \rightarrow l^+l^-+(\gamma^*)$	POWHEG-BOX	NLO	$t(Z/\gamma^*)$	LO	240 fb	5%	-	No
W+jets	SHERPA 2.2.1	NLO	$t\bar{t}t$	LO	1.6 fb	50%	-	No
VH	POWHEG-BOX	NLO	$tW(Z/\gamma^*)$	NLO	16 fb	50%	-	No
			$t\bar{t}W^+W^-$	NLO	9.9 fb	50%	-	No
			VVV	NLO	SHERPA	50%	-	No

# 对象选择和事例筛选

- ▶ 触发：单轻子或者双轻子触发（ST or DT）。
- ▶ 多轻子分析定义3种轻子用于不同末态过程。2LSS 物理道轻子使用Tight（T）定义。

孤立化：假轻子压低 BDT

电荷误判压低 BDT

$e/\gamma$ 模糊判别要求高质量电子

	$e$			$\mu$		
	B	L	T	B	L	T
Isolation	No	PLVLoose	PLVTight	No	PLVLoose	PLVTight
Identification	LooseLH		TightLH	Loose		Medium
Charge MisID BDT	No	Yes		N/A		
Ambiguity type	No	Yes		N/A		
$ d_0 /\sigma_{d_0}$	< 5			< 3		
$ z_0 \sin \theta $	< 0.5 mm					

- ▶ Jet 要求：  $p_T > 25 \text{ GeV}$ ,  $|\eta| < 2.5$ 。  $p_T < 60 \text{ GeV}$  和 要求 Tight JVT。
- ▶ 事例预选择：
  - ▶ 轻子横动量  $p_{T0}, p_{T1} > 20 \text{ GeV}$ 。对符合离线触发的事例，还要求  $p_T$  大于触发器的阈值。
  - ▶ b-jet 去除：DL1r 77% 工作点。
  - ▶ 至少2个 jet。
  - ▶  $\tau_h$  移除：RNN medium 重建  $\tau_h$ 。
  - ▶ 双轻子不变质量  $m_{\ell\ell} > 12 \text{ GeV}$ 。



# 多变量分析

- ▶ 使用BDT方法做信号和本底的二分类，提高信号显著度。
- ▶ 将信号（HH）与VV， ttbar， Vjet 样本分别训练BDT模型。

▶ 使用到的输入变量：18个。包含轻子、轻子-jet系统

▶ 质量信息： $M_{ll}, M_{all}, M_{\ell 0j}, M_{\ell 0j}, M_{W0}^T, M_{W1}^T$

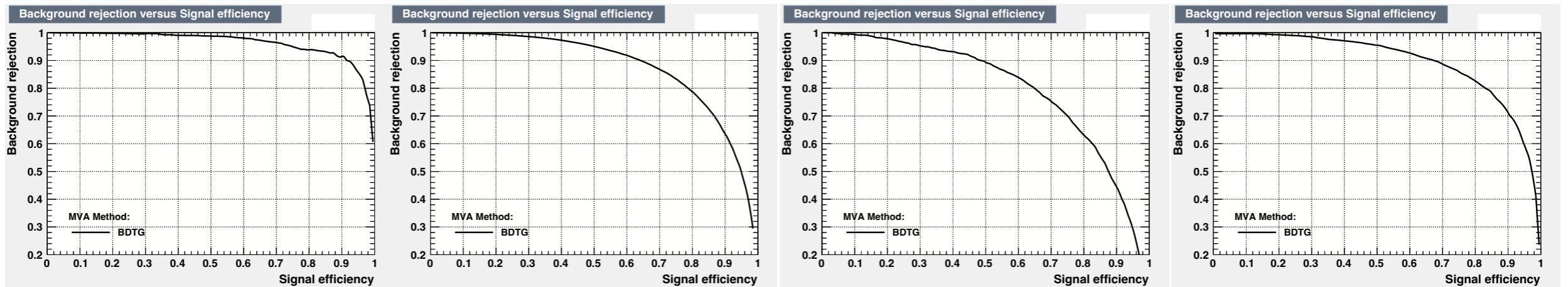
▶ 轻子信息： $\eta_0, \eta_1, \Delta_\eta, N_{jets}, H_T, H_T(Lep)$

▶ 角度信息： $\Delta R_{min(l0jets)}, \Delta R_{min(l1jets)}, \Delta R_{min(ljets)}, \Delta_{\ell\ell}$

▶ 事例信息： $\ell_{flavour}, E_T^{miss}, Q_{sum}$

$$Z = \sqrt{2 \times [(B + S) \times (\ln \frac{S + B}{B} - S)]}$$

- ▶ 3个独立BDT的输出作为最终 BDT（Combined BDT）的输入变量。
- ▶ 信号区间：最大化显著性。要求 $BDT_{All} > -0.4$ 。低端部分作为验证区间（VR）。
- ▶ 多分类神经网络（MLP）的提升不超过5%：受限于ttbar/Vjets 统计量。



Vjets BDT

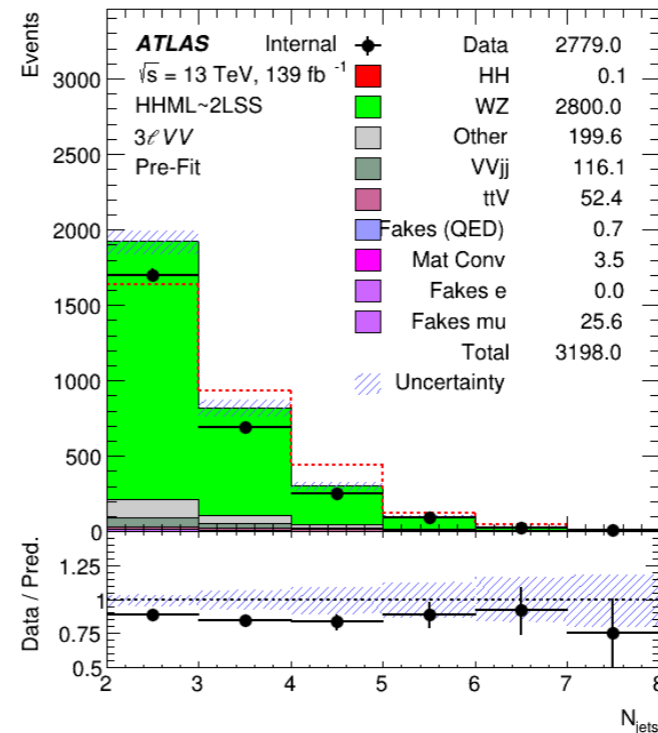
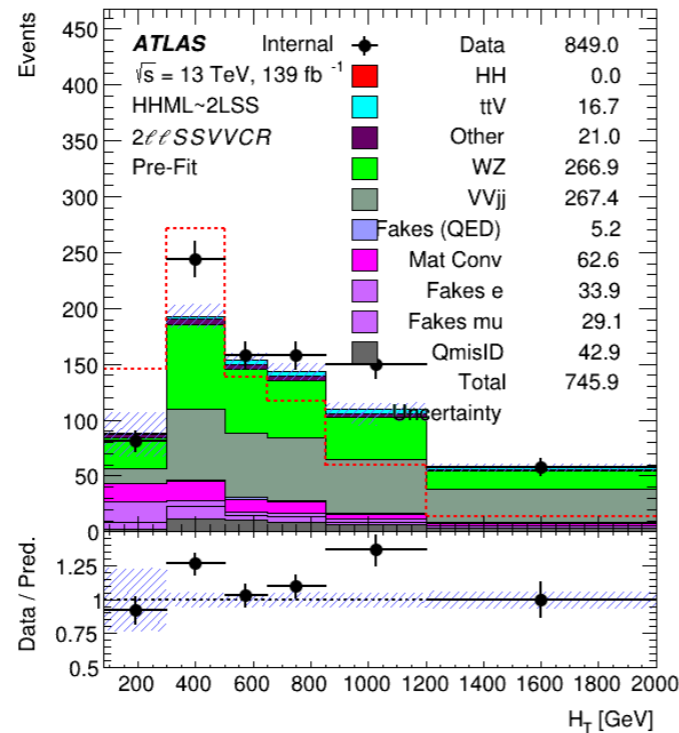
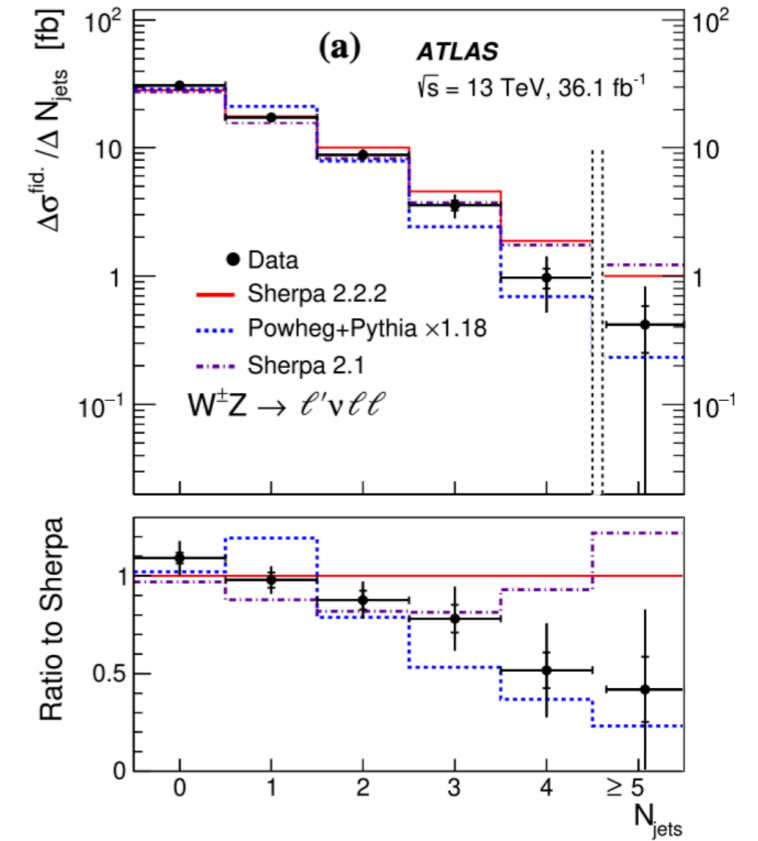
VV BDT

ttbar BDT

Combined BDT

# 双玻色子 (VV) 过程本底

- ▶ 主要有  $W^\pm W^\pm$ ,  $W^\pm Z$  的 QCD 及 EW VBS 过程。其中的可约本底部分使用特定fake/QMisID方法的估计。
- ▶ 在SR中, WZ占35%, VVjj (尤其是ssWWjj) 占比 12.5%。
- ▶ 由于这两个物理过程不同, 因此需要导出独立的归一化因子。
  - ▶ 在ATLAS实验上, 产生子对VV过程存在误建模是已知的。WZ: Sherpa 在  $N_{\text{jets}} > 2$  区域; VVjj 信号强度约为1.4
- ▶ 使用两个CR: WZ主导的CR和VVjj富集的CR。
  - ▶ WZ CR 要求一个额外的轻子, 模拟丢失1个轻子的过程。
  - ▶ VVjj CR要求  $m_{jj} > 300 \text{ GeV}$  以从WZ中区分。



# 假轻子本底

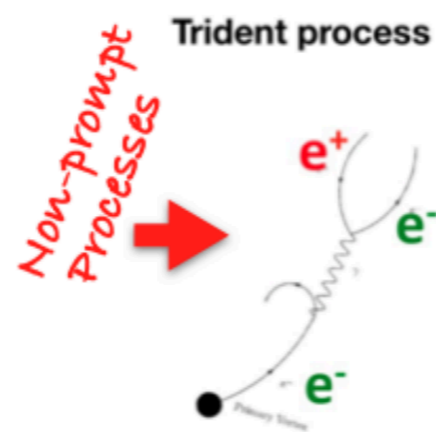
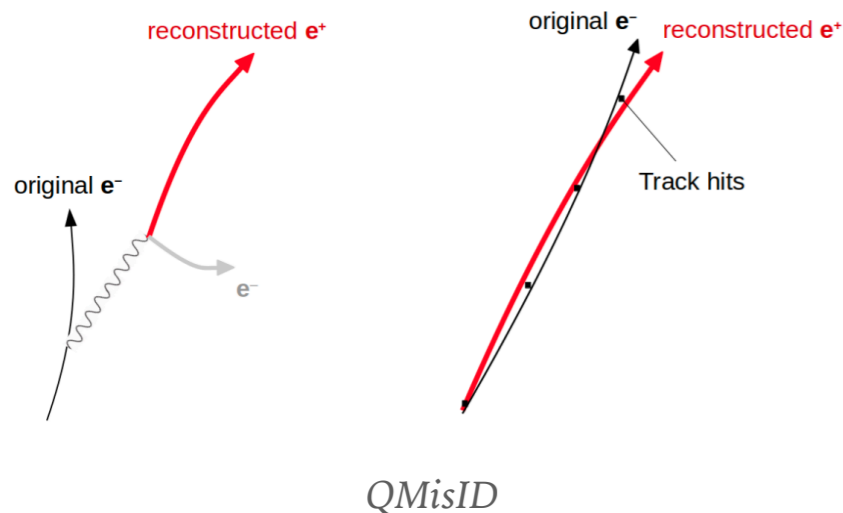
► 根据假轻子的来源，定义4种需要修正的假轻子过程。

## 电子

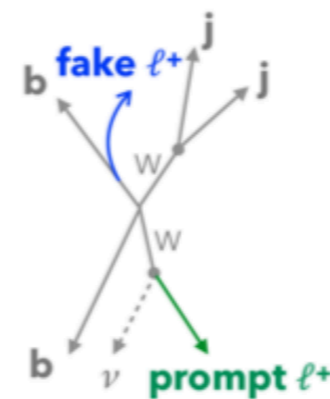
- 重味衰变 ( $b/c \rightarrow \ell$ )
- 光子转换: External ( $\gamma \rightarrow ee$ ) ; Internal ( $\gamma^* \rightarrow ee$ ) , 利用径迹不变质量和转换半径压低
- 电荷误判 (QMisID) : 数据驱动方法

## 缪子

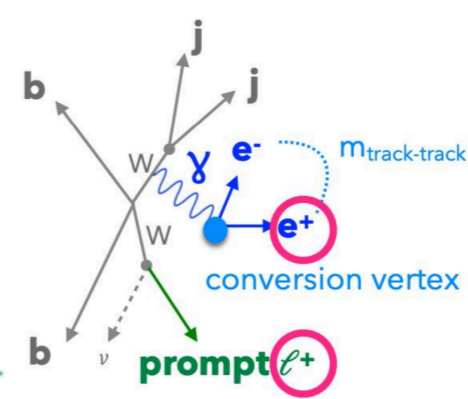
- 重味衰变 ( $b/c \rightarrow \ell$ )
- $\pi, K$  强子衰变



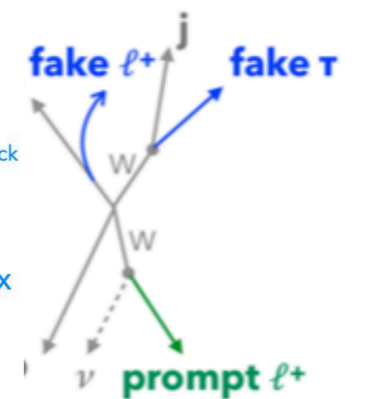
Semileptonic b-decay



Photon conversions

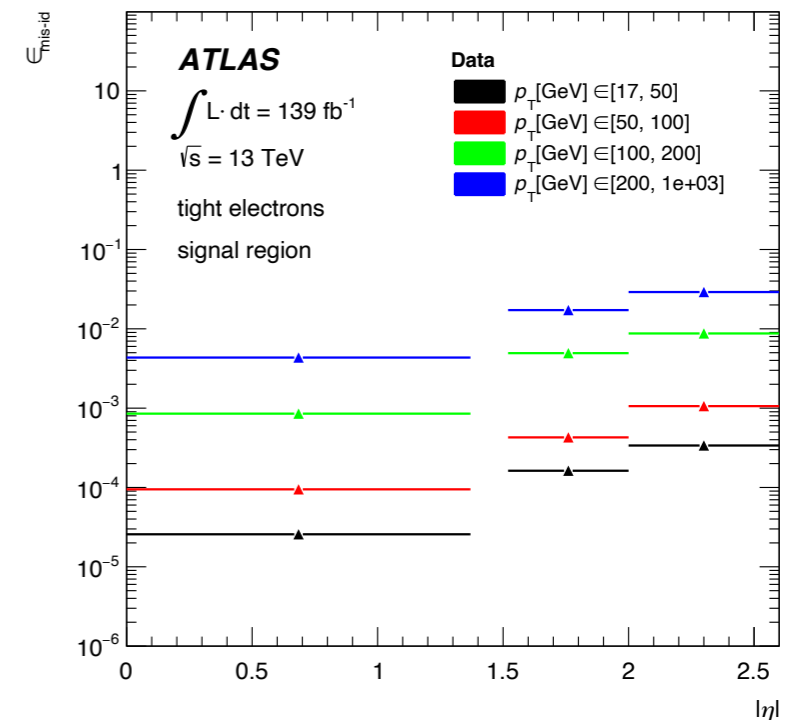
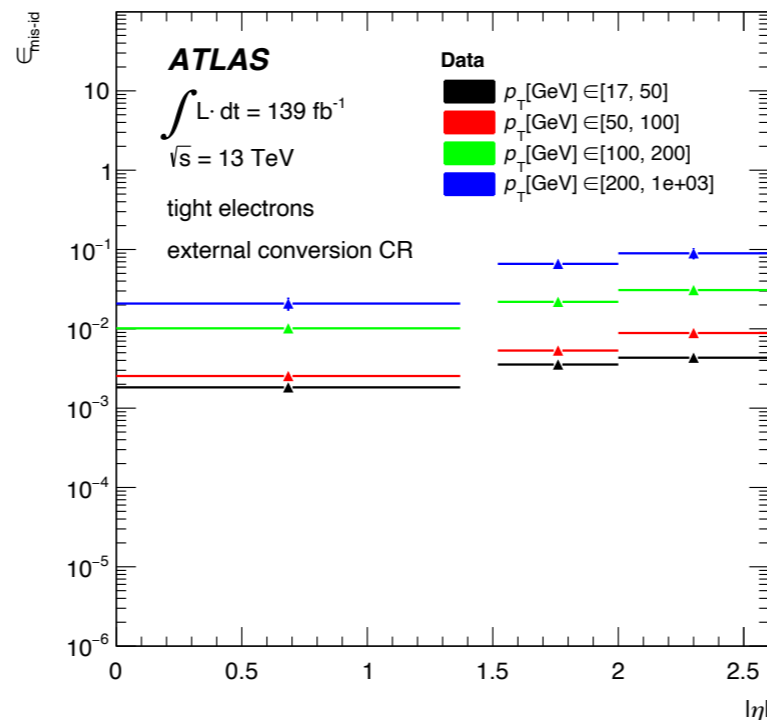
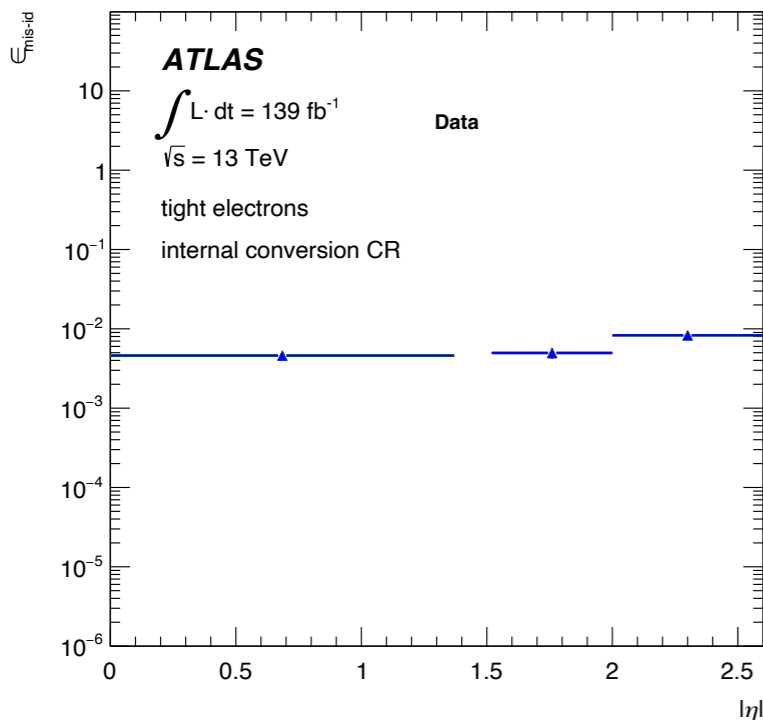
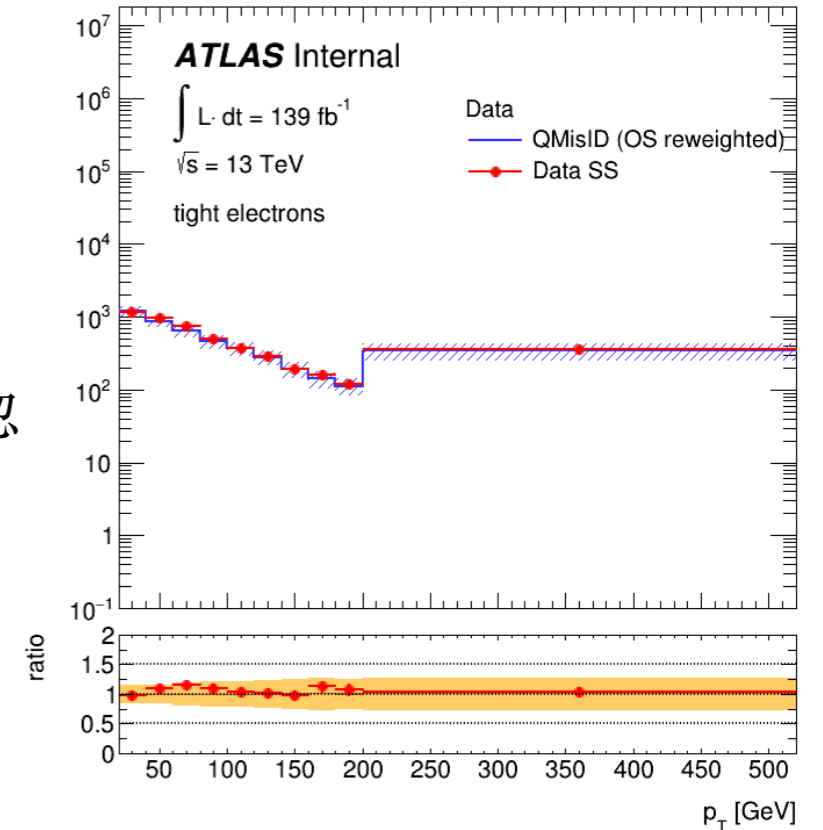


Non-prompt lepton & fake  $\tau$



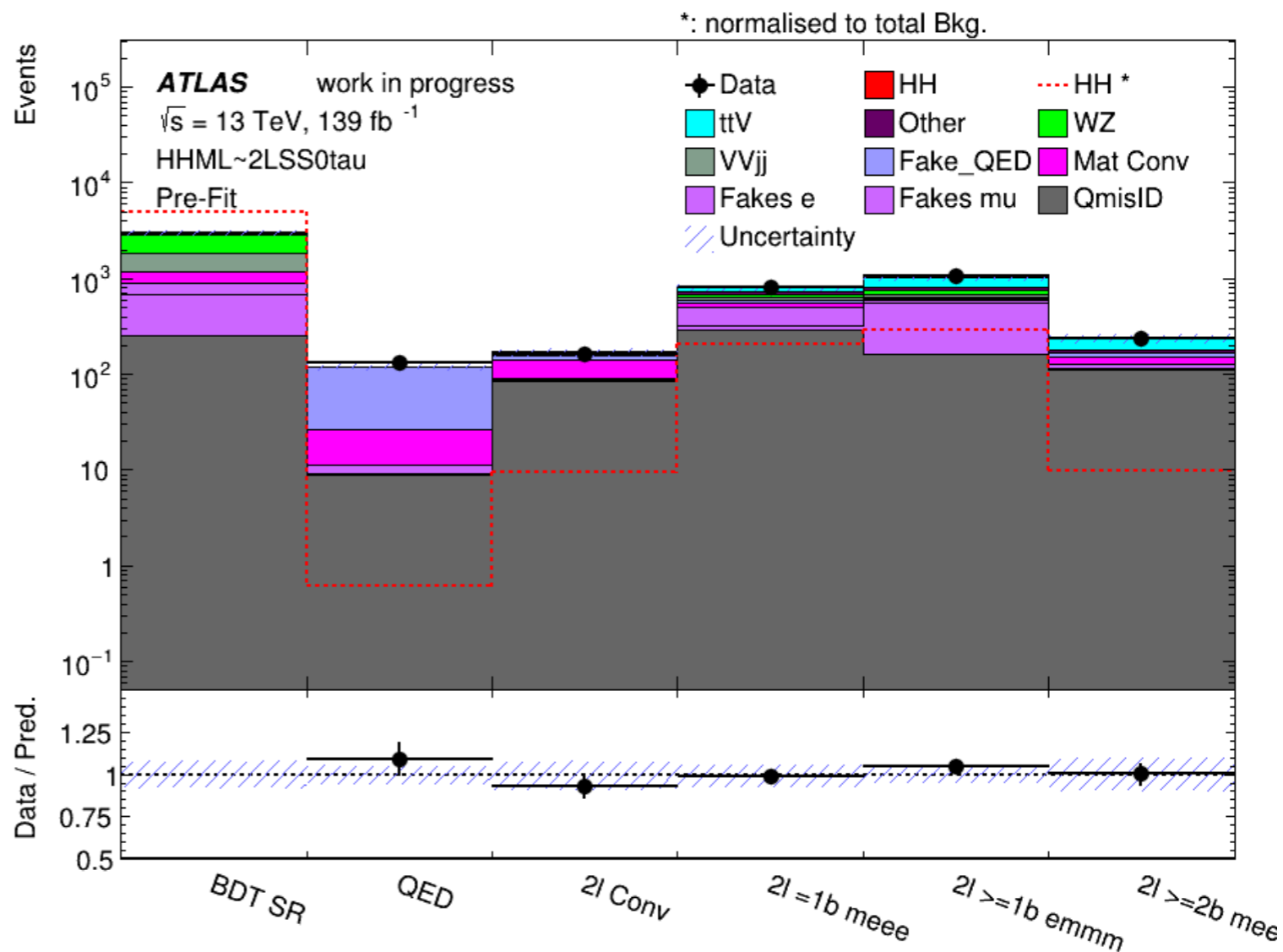
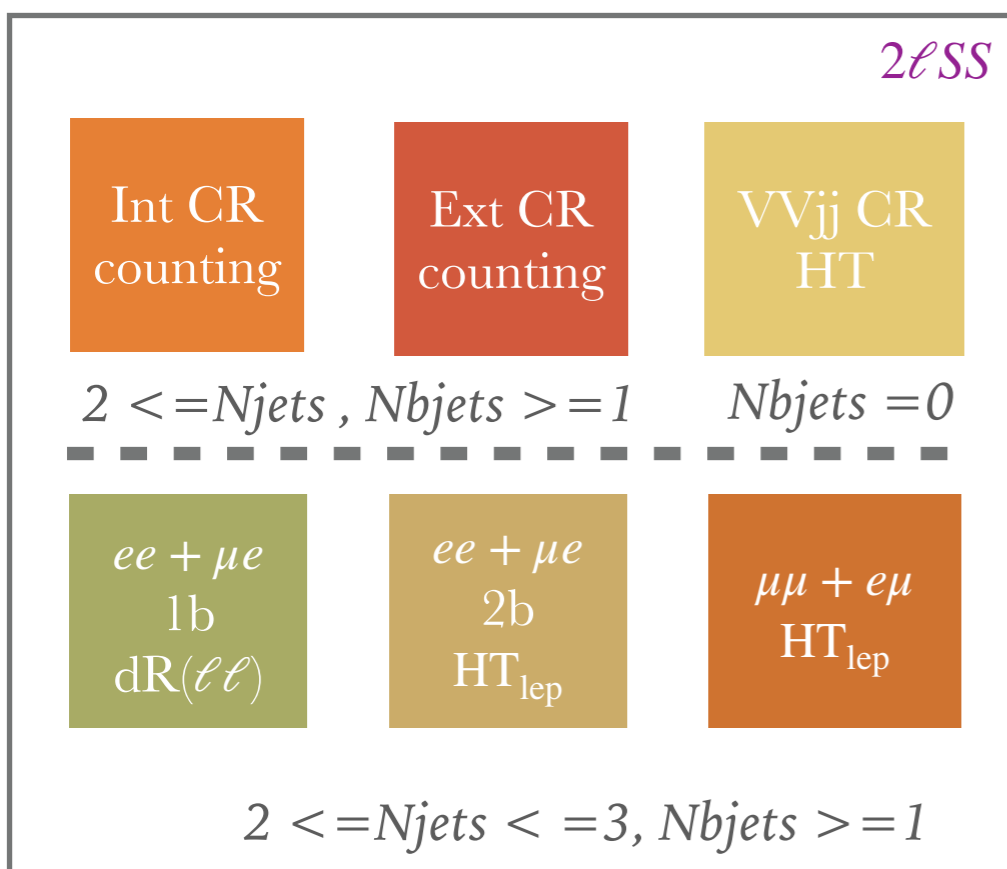
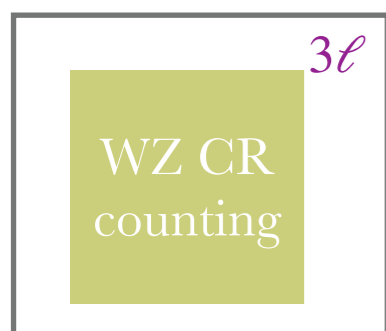
# 电荷误判本底

- ▶ 电荷误判 (QMisID) 本底主要存在于 $2\ell SS$ 过程, 贡献自 $ttbar/Zjets$ 过程。
  - ▶ 电子韧致辐射及不对称的转换 ( $e^\pm \rightarrow e^\pm \gamma^* \rightarrow e^\pm e^+ e^-$ )
  - ▶ 径迹曲率的错误测量
  - ▶ 只与 $ee$ 和 $e\mu$ 末态有关, 在分析的 $p_T$ 范围内的缪子的误判率忽略不计。
- ▶ 在数据 ( $Z \rightarrow e^+e^-$ ) 中利用极大似然估计误判率, 本底使用边带区域 ( $Z$ -mass 拟合的 $8\sigma$ 以外) 估计扣除。
- ▶ 根据电子的 $|\eta|, p_T$ , 类型建立3维的估计区域。
- ▶ 对 $P_t$ 分布做内插得到光滑的分布。



# 模版拟合方法

- ▶ 使用蒙特卡洛 (MC) 的形状估计假轻子本底及VV 本底，与信号强度做联合拟合。
  - ▶ 定义5个富含各类型本底的控制区间，基于MC truth 信息。
  - ▶ MC形状来自 ttbar, W+jets, Z+jets, singleTop, Vy
  - ▶ 运动学和事例产额作为拟合变量。
  - ▶ 浮动6个归一化参数: Internal/External 转化电子、重味衰变电子/缪子、WZ、VVjj。





# 本底建模误差

## ▶ 模版拟合过程中的本底误差

▶ MC模版形状的系统误差：通过放宽轻子孤立化和 $e/\gamma$ 模糊性条件计算。

$$\frac{Data - NonFakes}{Fake}$$

▶ MC自身建模误差，包括与实验关联的误差（物理对象重建、校准等）以及 $ttbar$ 的QCD scale, PDF等理论误差。

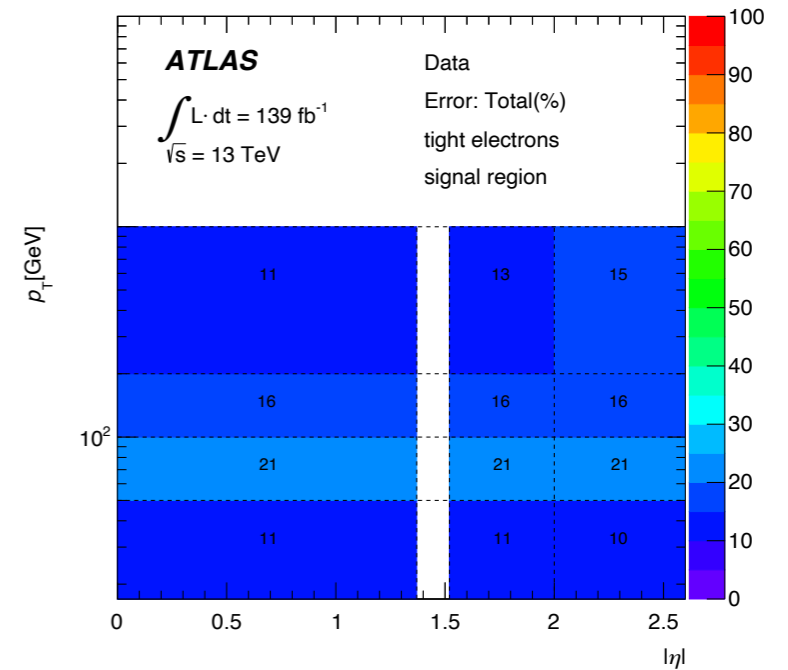
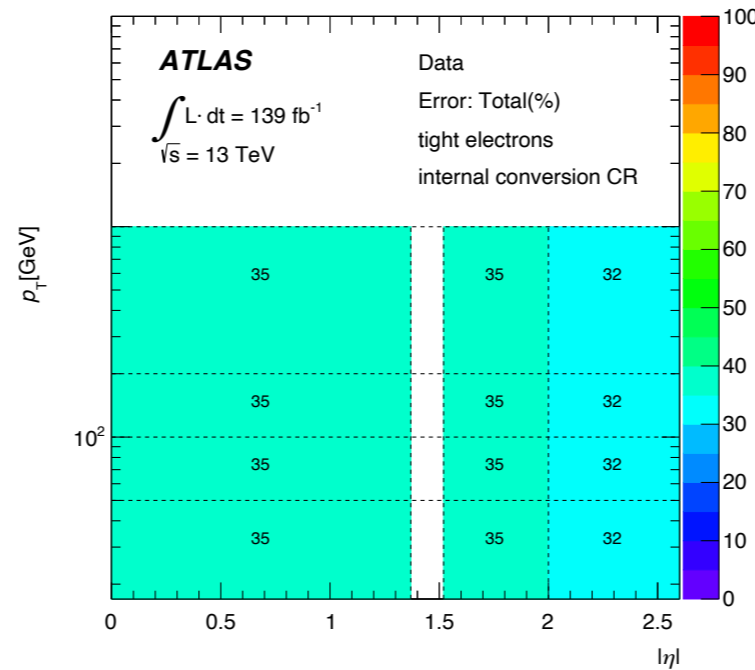
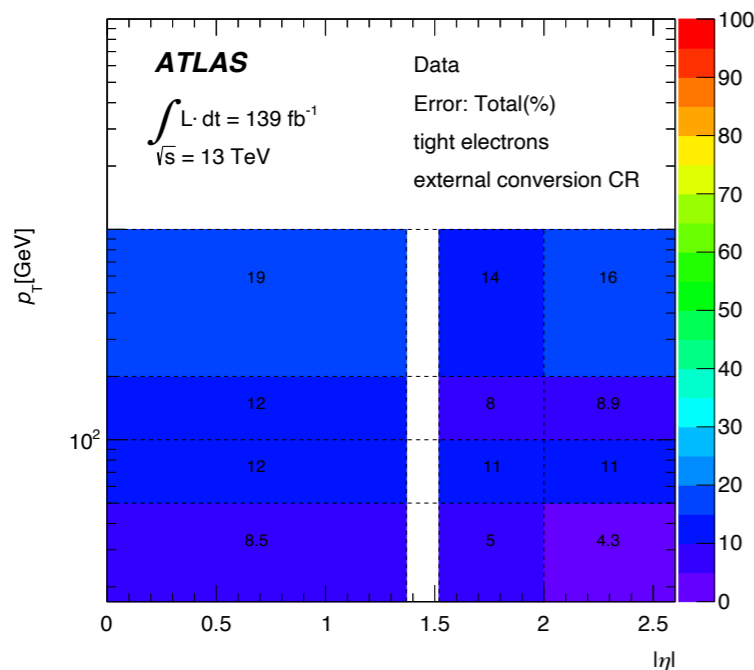
Uncertainties	BDT [-1,0.6]	BDT [-0.6,-0.2]	BDT [-0.2,0.2]	BDT [0.2,0.6]	BDT [0.6,1]
Heavy flavor muons	30%	15%	-15%	10%	5%
Heavy flavor electrons	100%	30%	-15%	10%	5%
Material Conversions	200%	100%	50%	5%	-10%

## ▶ QMisID 本底误差为三项来源的平方和

▶ 拟合过程中的统计误差（引入anti-tight 电子增大统计量）。

▶ 估计值与MC 真实信息差别的 unclosure（主导）。

▶ Z质量窗口的误差，移动中心值比较误判率的差别。





# 系统误差

## 实验误差:

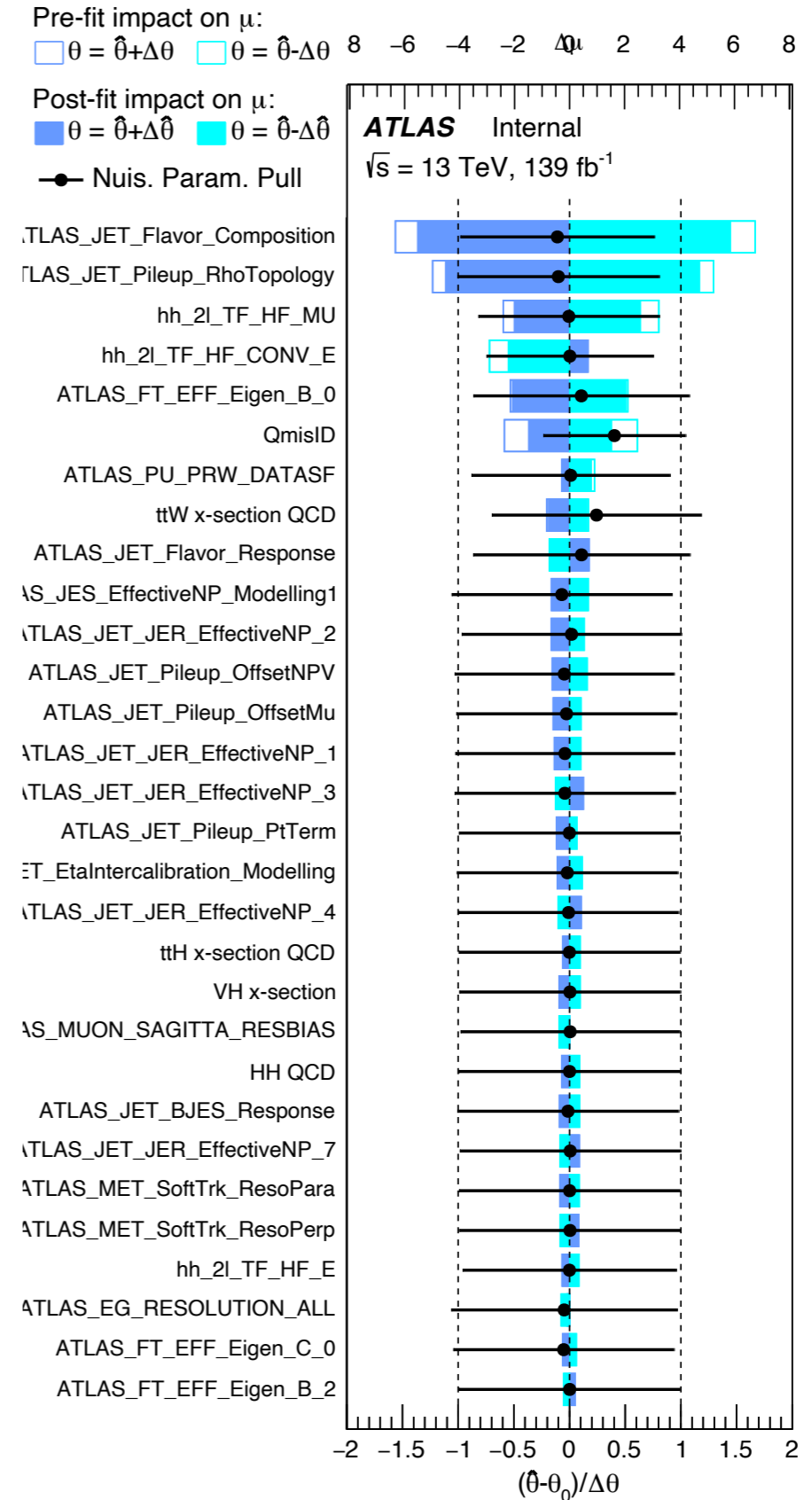
- 积分亮度, Pileup校准、轻子, jet, b-tagging, 丢失横动量 (132项)。

## 信号的建模误差

- QCD 尺度, 部分子分布函数 (PDF), 强子化模型 (Herwig vs Pythia) (3项)

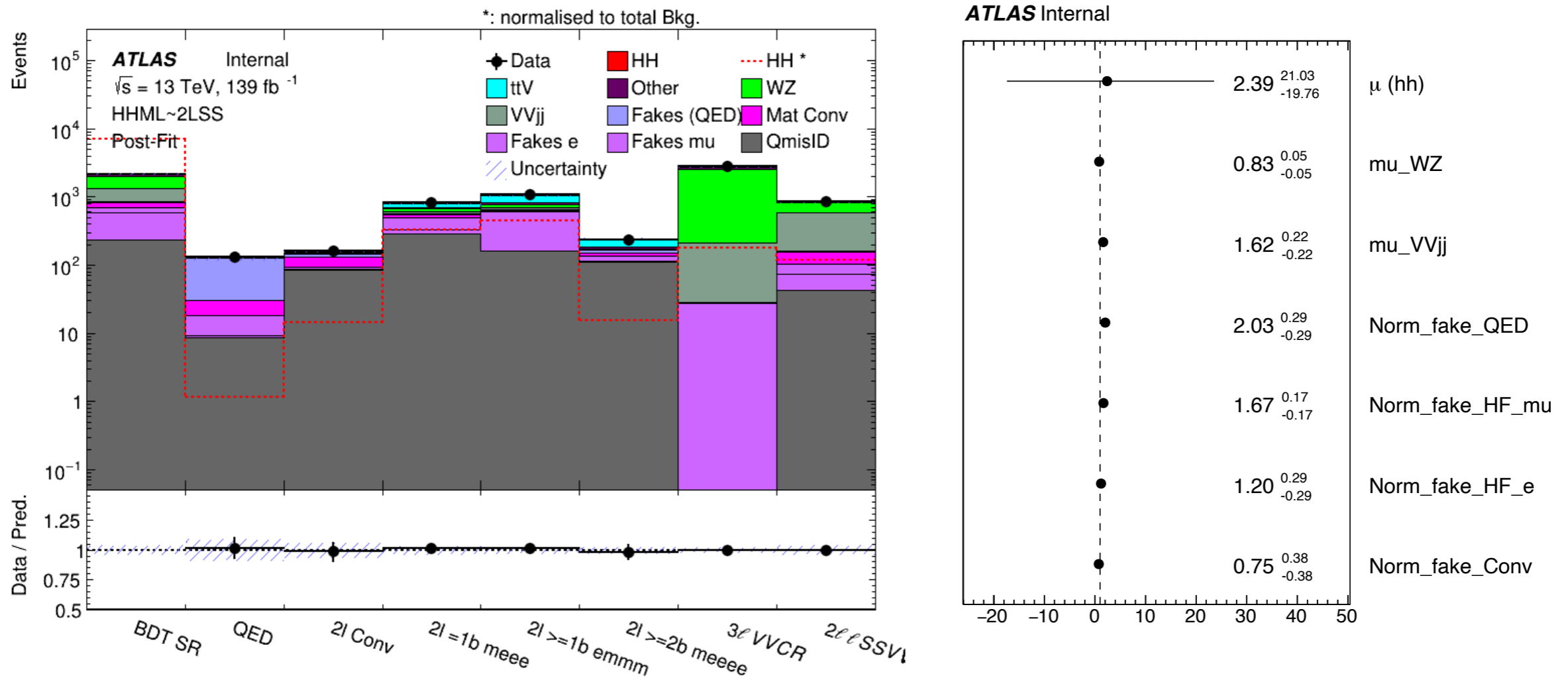
## 本底建模误差

- Prompt 本底的理论误差, 包括QCD scale, PDF: VV, ttV, rare top 等 (12项)
- 假轻子本底系统误差 (4项)
- 电荷误判本底系统误差 (1项)



# 拟合结果: CR

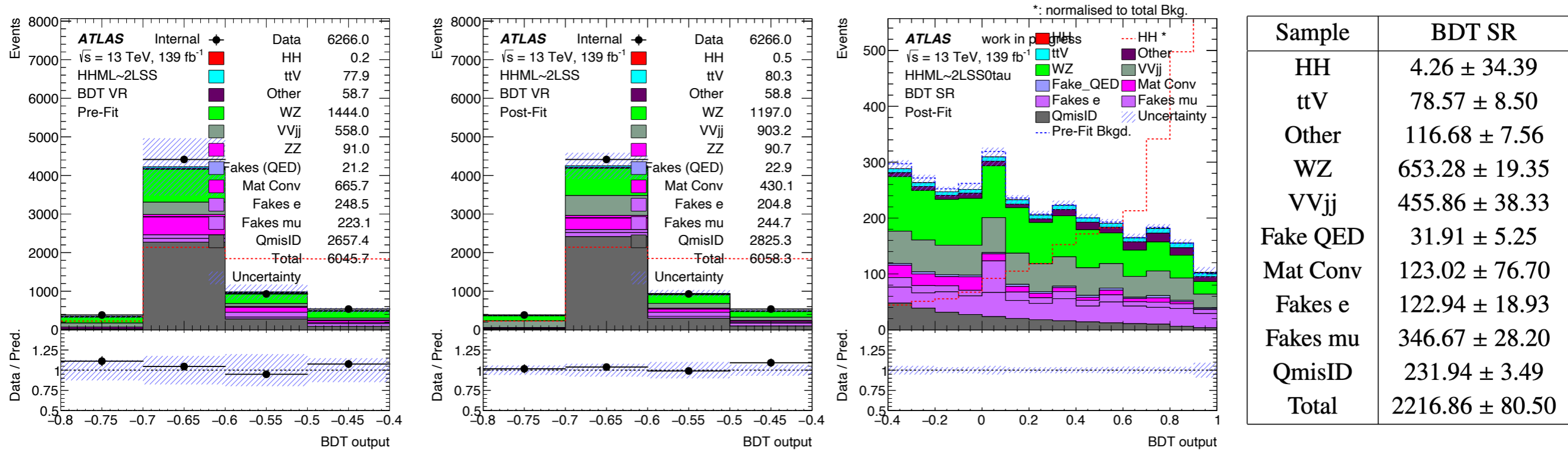
- ▶ 信号+本底拟合: CR (7) 中使用数据+ SR (1) 中使用Asimov数据。
- ▶ CR中的信号污染小于 0.01%
- ▶ 信号强度 ( $\mu_{hh}$ ) 作为POI, 其余归一化参数自由浮动, 系统误差项作为冗余参数 (NPs)。



- QmisID为数据驱动估计
- 2lss/3l 区间共享一组假轻子norm factors

# 拟合结果：VR及SR

- ▶ Combined BDT < -0.4 作为验证区间。
- ▶ Combined BDT > -0.4: 信号区间，数据被隐藏。

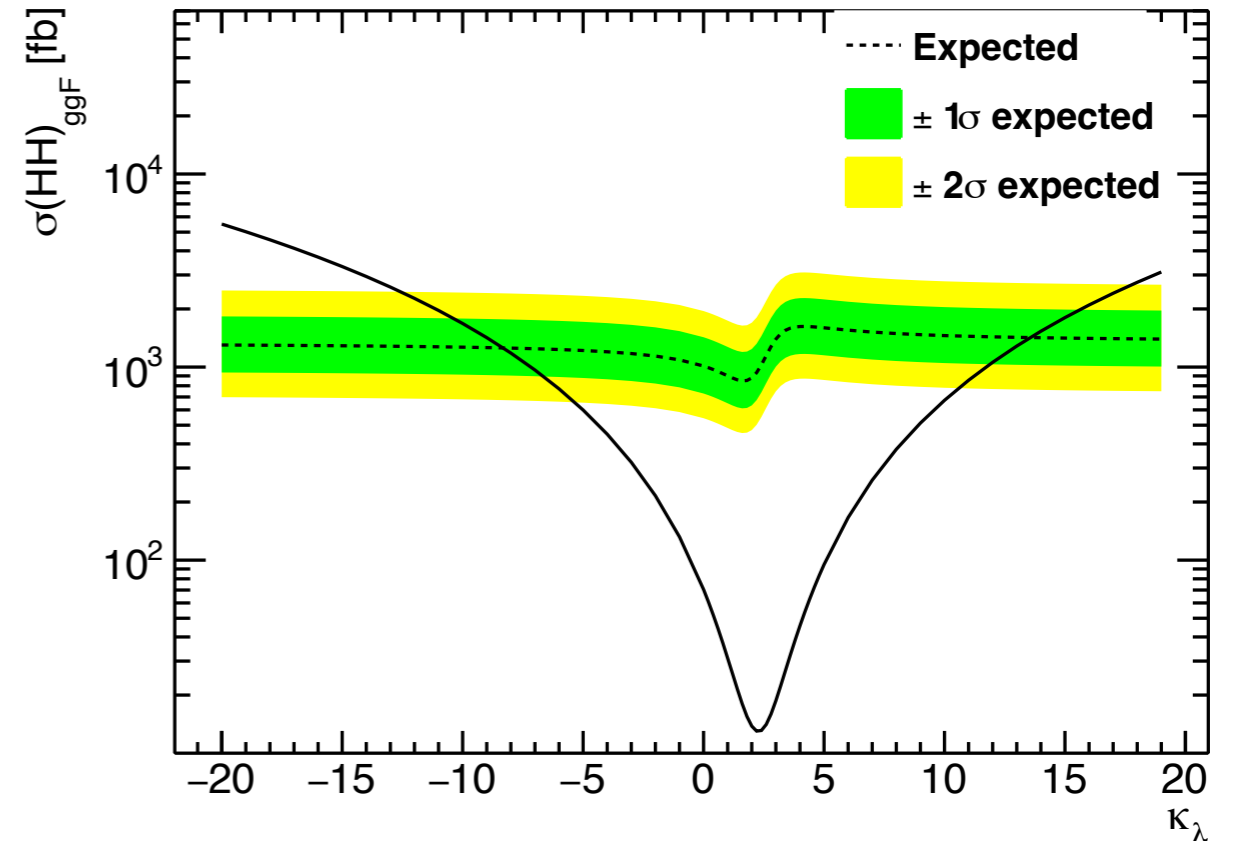
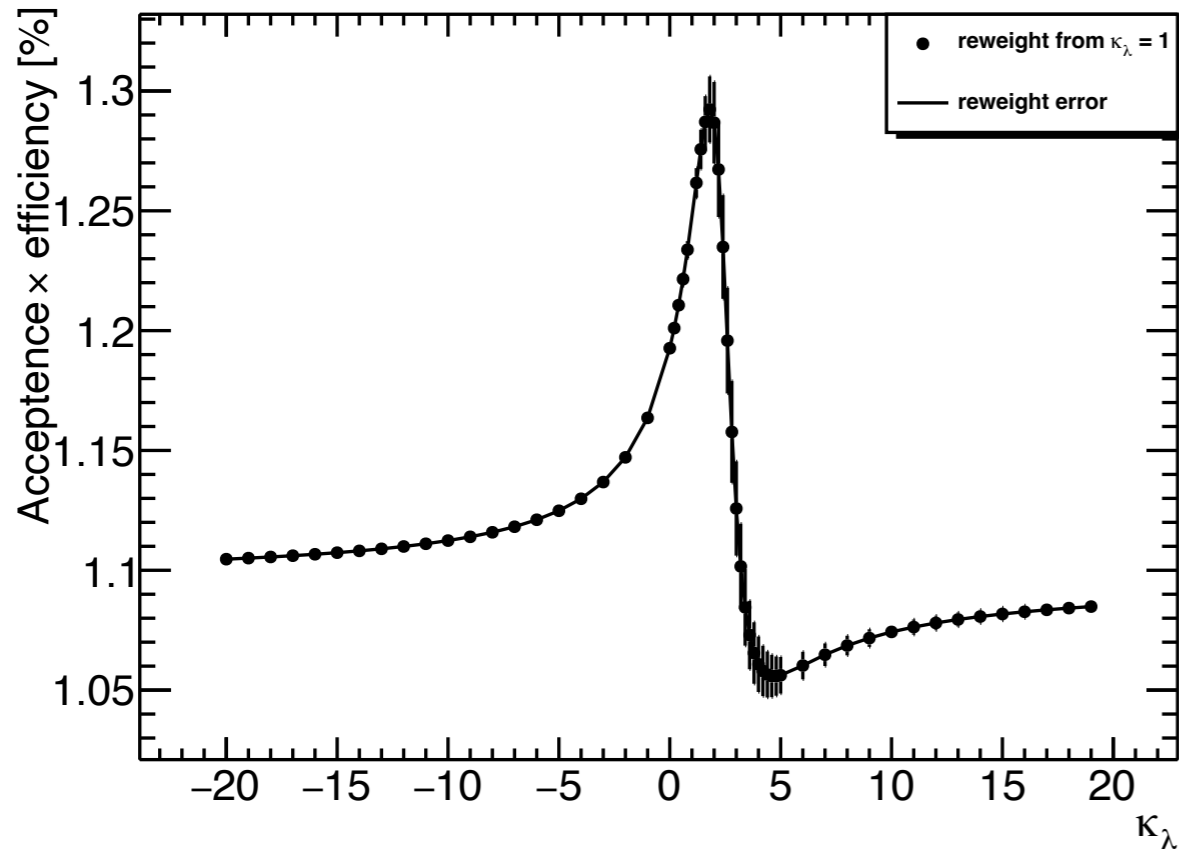


- ▶ CLs 方法计算95%置信度下的HH信号截面上限

	$-2\sigma$	$-1\sigma$	Expected	$+1\sigma$	$+2\sigma$	Observed
$\sigma_{HH}/\sigma_{HH}^{SM}$ Stats.	16.39	22.01	30.55	43.15	59.21	blinded
$\sigma_{HH}/\sigma_{HH}^{SM}$ Sys.	19.06	25.59	35.52	50.49	69.91	blinded

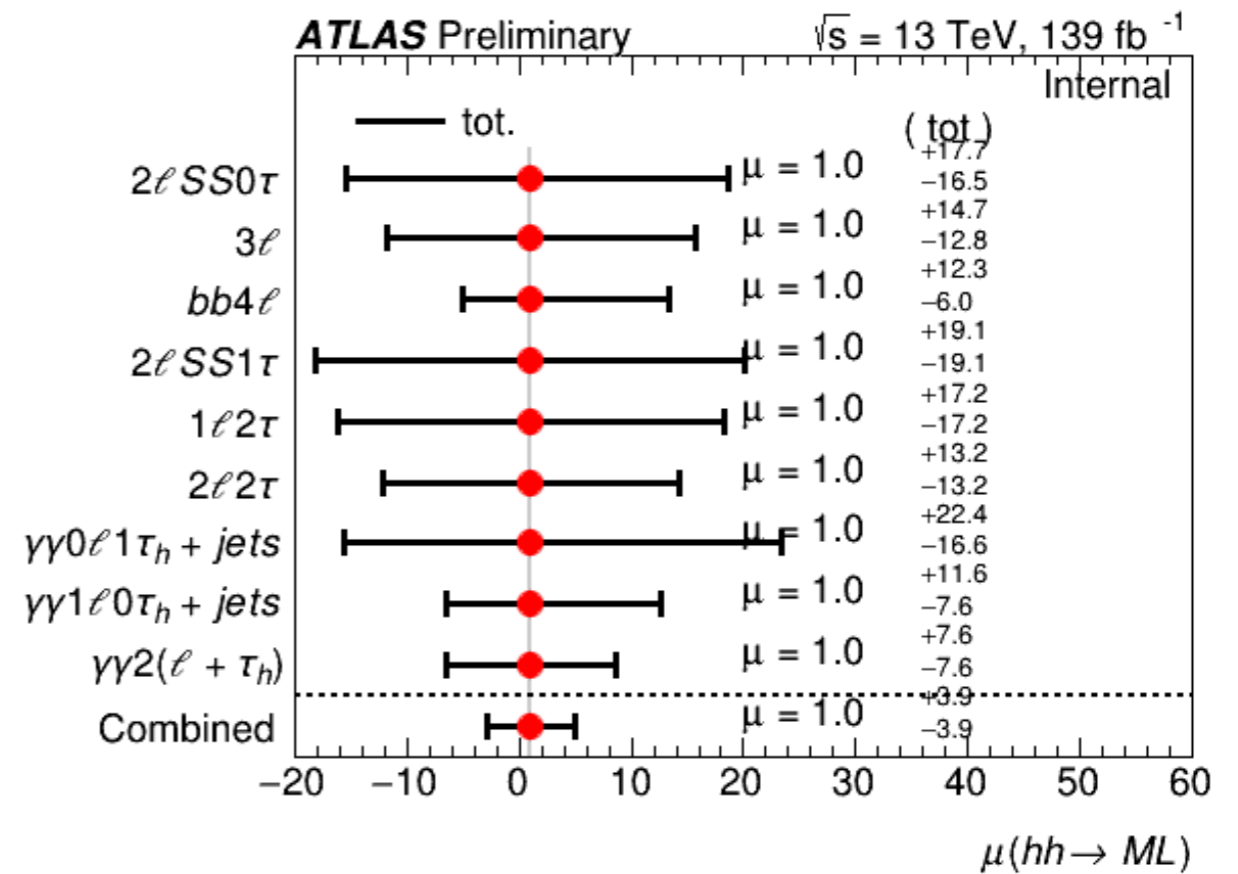
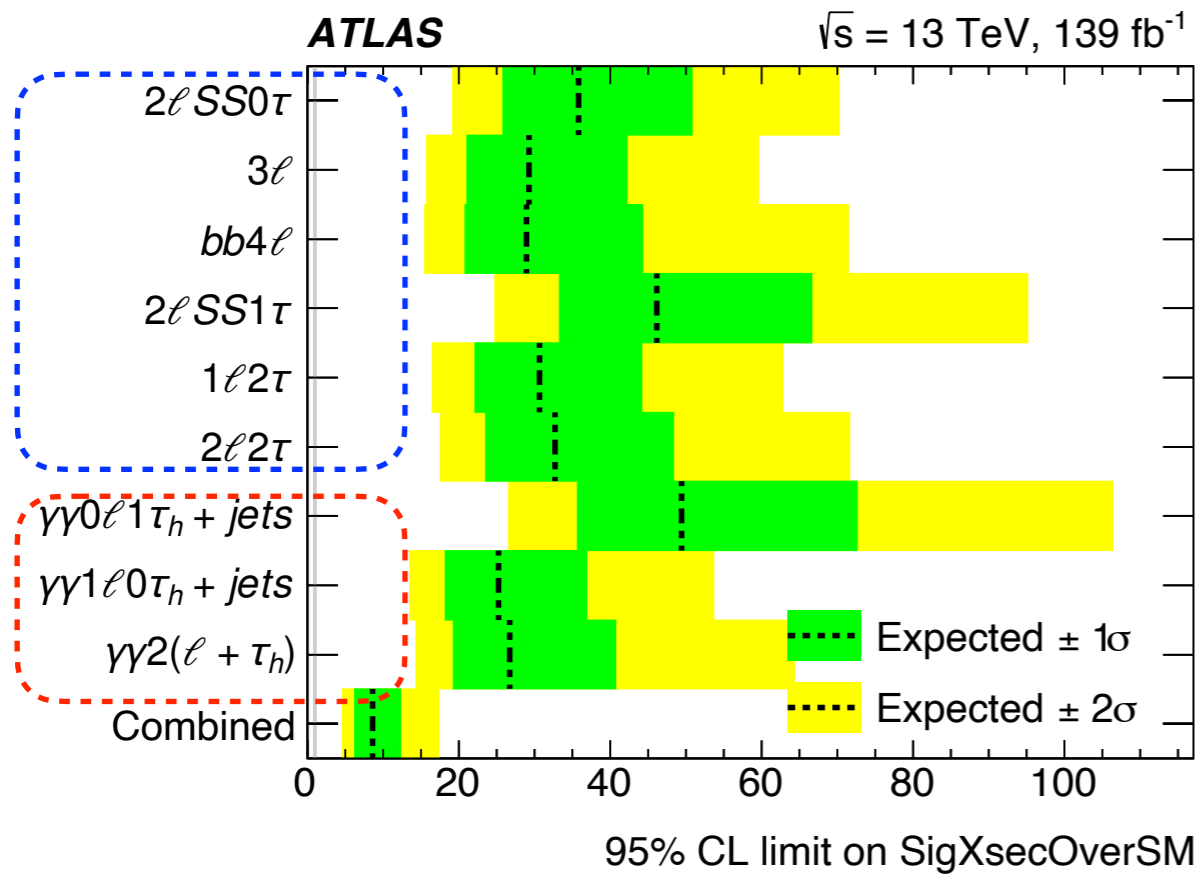
# $\kappa_\lambda$ 测量

- ▶  $\kappa_\lambda$  扫描在ATLAS HH 多轻子道是首次测量。
- ▶ 连续的 $\kappa_\lambda$ 样本通过 $\kappa_\lambda = (0, 1, 20)$ 的线性组合得到。
- ▶ 测量方法：应用SM HH的BDT决策变量，计算95% CLs 截面上限。



# HH多轻子道联合测量

- ▶ 9个多轻子物理道联合测量以达到最佳灵敏度。
- ▶ 36.1 fb<sup>-1</sup>分析预期结果: **WWWW** 分析为120倍 $\sigma_{HH}^{SM}$ ; **WWyy**分析为170倍 $\sigma_{HH}^{SM}$ 。
- ▶ 139 fb<sup>-1</sup>**纯多轻子物理道**: 12.18倍 $\sigma_{HH}^{SM}$ ; **yy+轻子物理道**14.98倍 $\sigma_{HH}^{SM}$  \*。
- ▶ 相比ATLAS 36.1 fb<sup>-1</sup>分析, 有约10倍的提升。
- ▶ 竞争力的结果。CMS 已投稿: 多轻子分析: 21 (19) 倍 $\sigma_{HH}^{SM}$ ; WWyy分析为 97 (52) 。



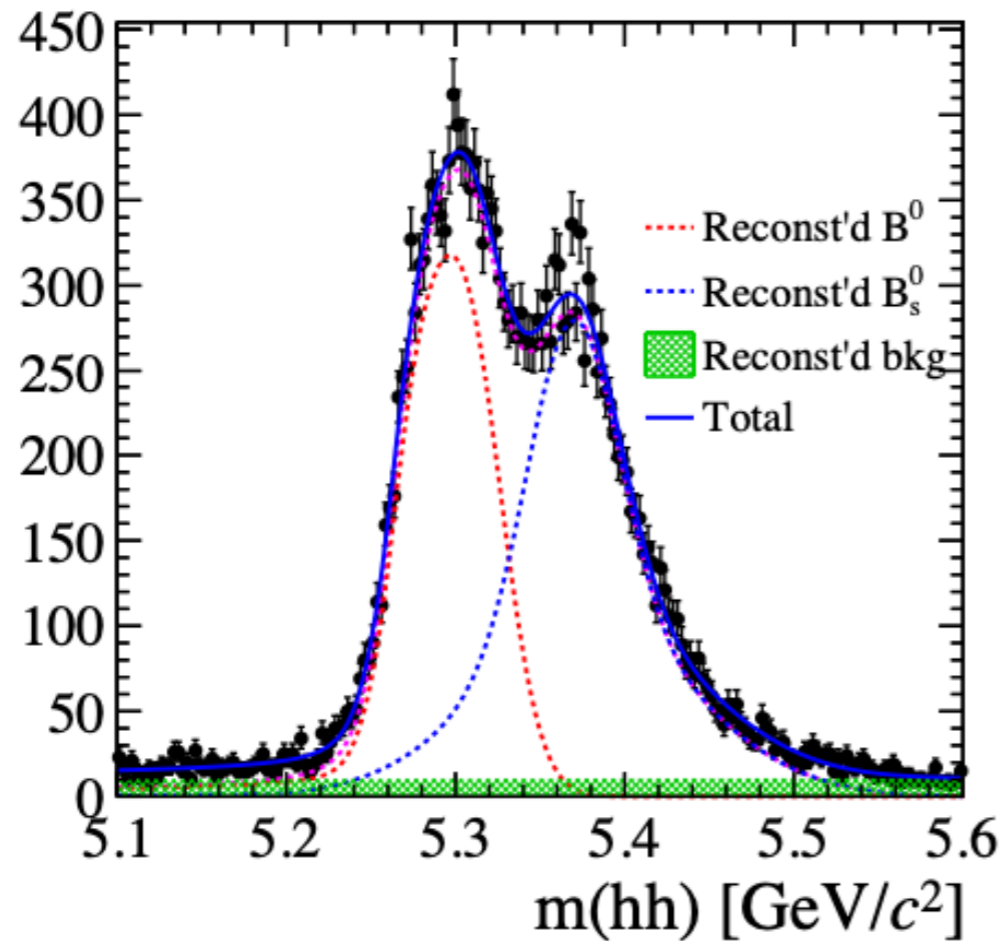
\* yy+轻子物理道的揭盲结果为 40 (18)

- 
- ▶ 研究背景
    - ▶ 标准模型和di-Higgs物理
    - ▶ ATLAS 探测器
  - ▶ ATLAS 上di-Higgs 研究
    - ▶ 通过同电荷双轻子寻找标准模型非共振双希格斯粒子
    - ▶ 多轻子联合测量
  - ▶ 漂移室的粒子鉴别算法研究
    - ▶ 原初电离计数（Cluster counting）研究背景
    - ▶ 基于BESIII 漂移室的粒子鉴别研究
    - ▶ Cluster counting 技术用于CEPC 漂移室
    - ▶ 束流测试
  - ▶ 总结与展望
  - ▶ 附文
    - ▶ 已取得研究成果
    - ▶ 主要贡献
    - ▶ 论文审阅建议

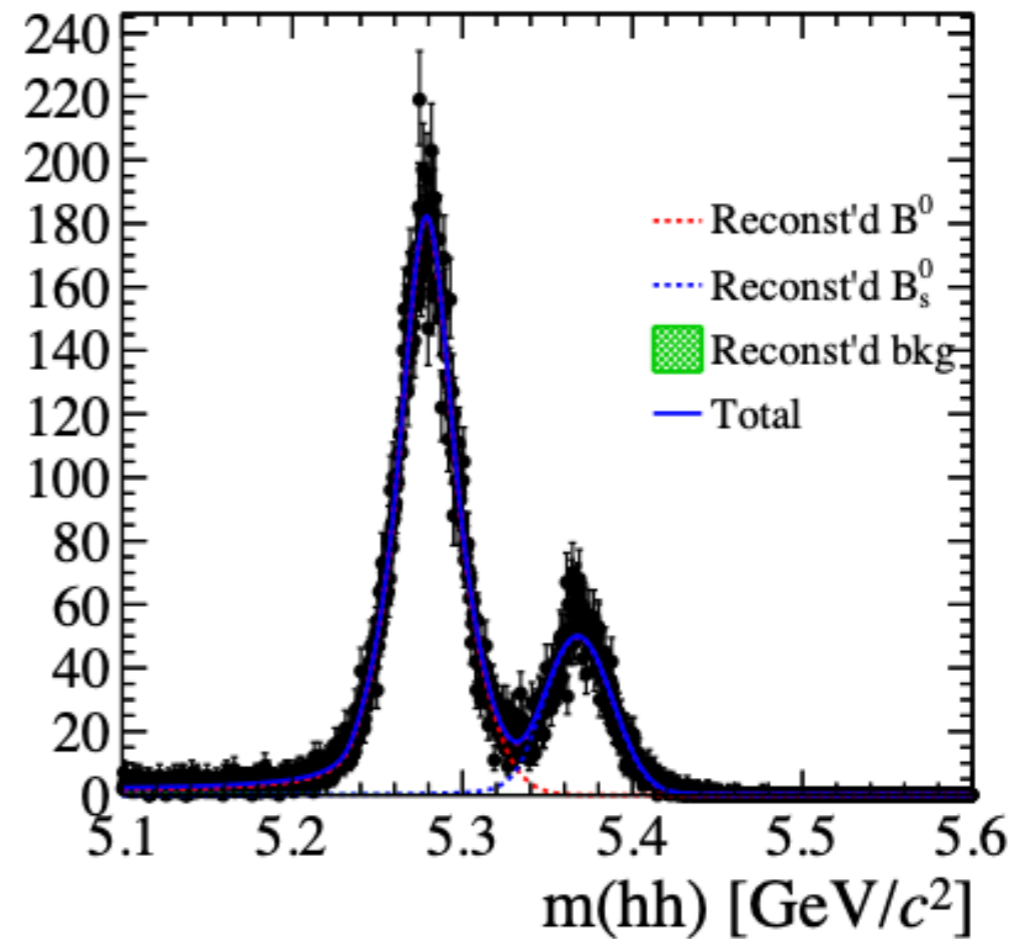


# 漂移室的粒子鉴别算法

- ▶ 粒子鉴别 (PID) 在  $e^+e^-$  对撞机上有着重要作用。
  - ▶ 重味夸克鉴别: b/c-tagging。
  - ▶ B物理,  $\tau$ 物理, 强子能谱, 尤其是在 Z-pole 产生模式: 降低组合本底, 提高质量分辨。
- ▶ 例如PID对重味衰变末态重建的影响: 有效区分  $B_s^0(B^0) \rightarrow h^-h^+$ 。



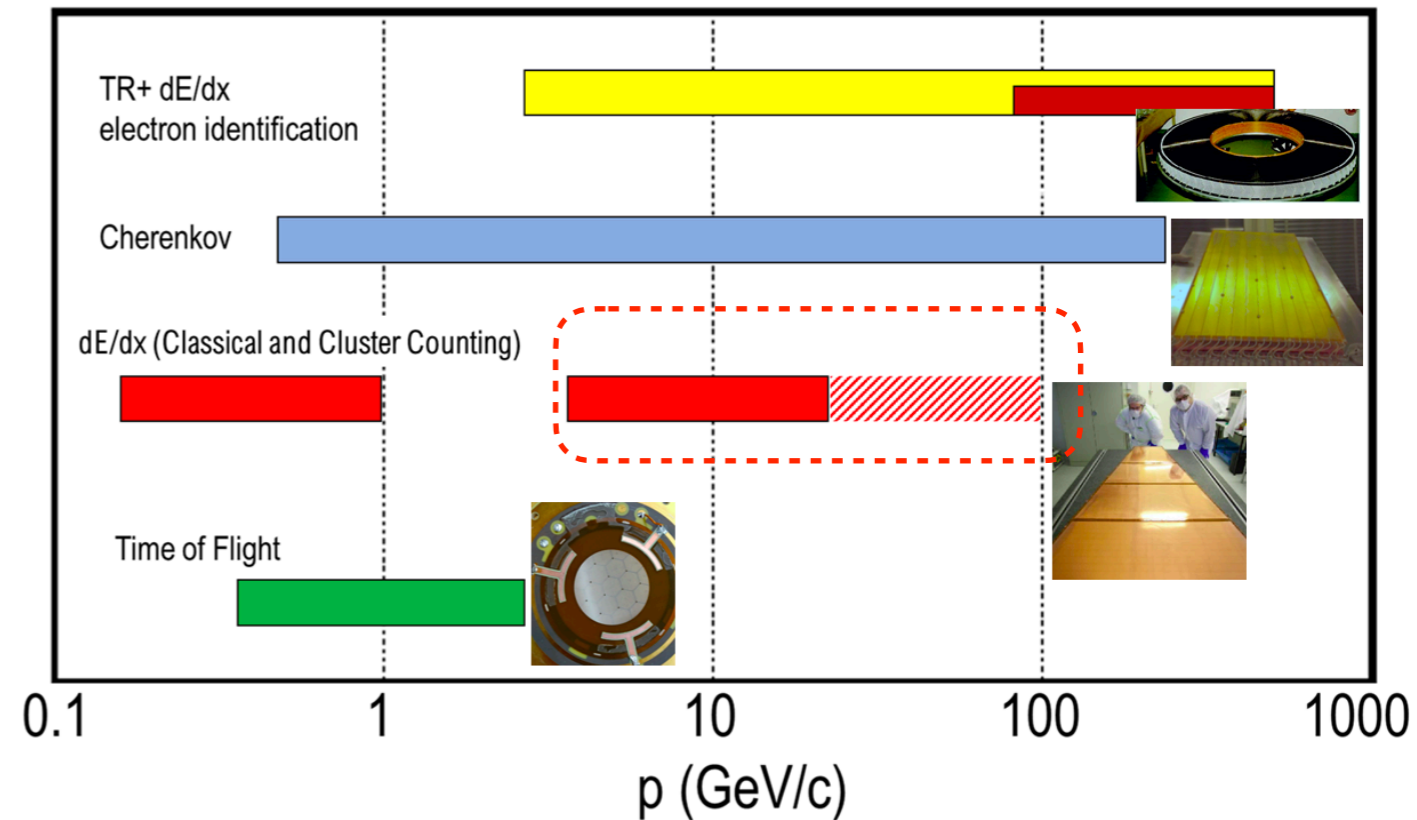
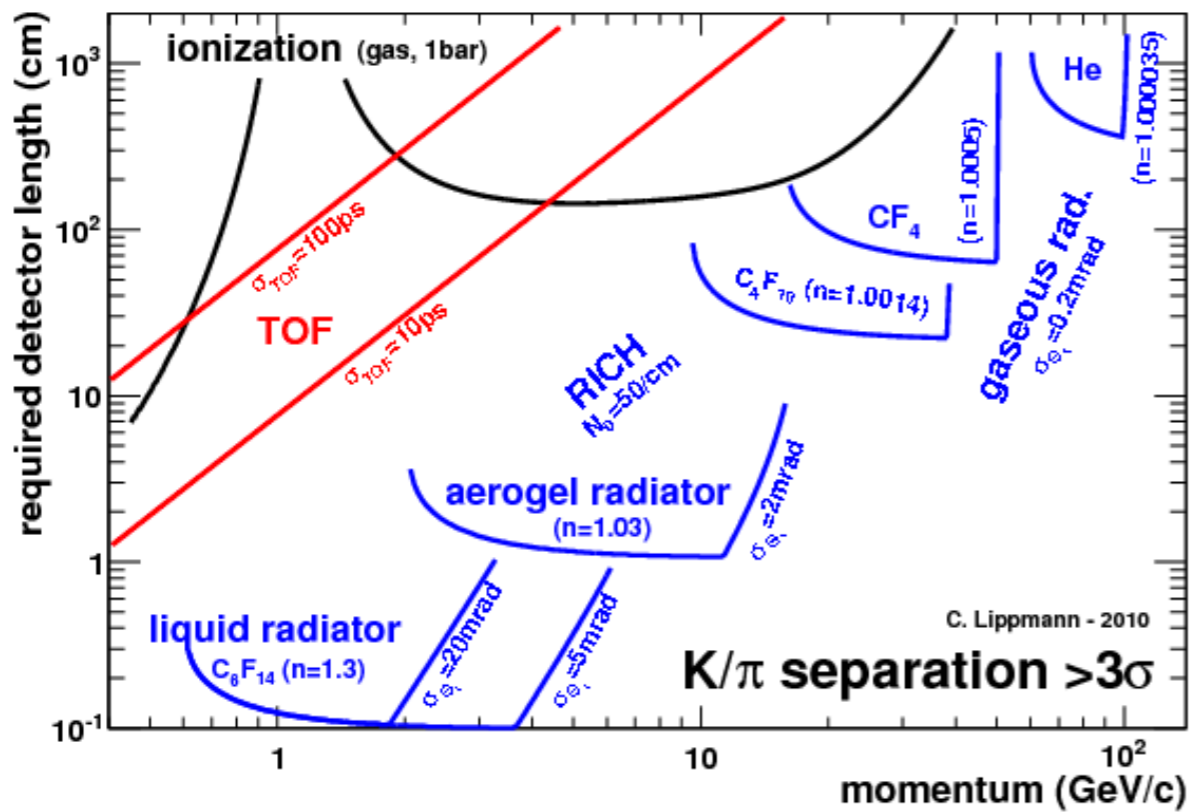
50% PID效率



100% PID效率

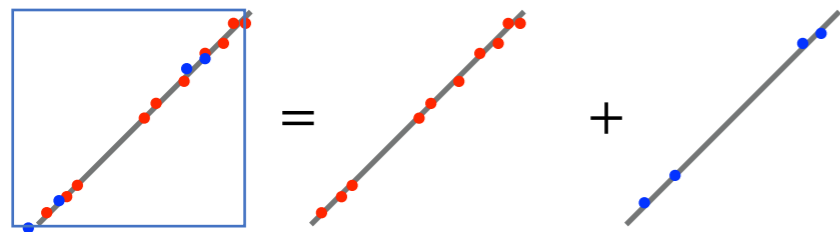
# 粒子鉴别方法

- ▶ 穿越辐射 (TR)
- ▶ 切伦科夫
- ▶ 电离能损 ( $dE/dx$ , 及原初电离计数, cluster counting)
  - ▶  $dE/dx$  的局限性: 要求探测器有足够的采样长度。另外, 在高动量区间无区分能力。
- ▶ 飞行时间 (TOF)



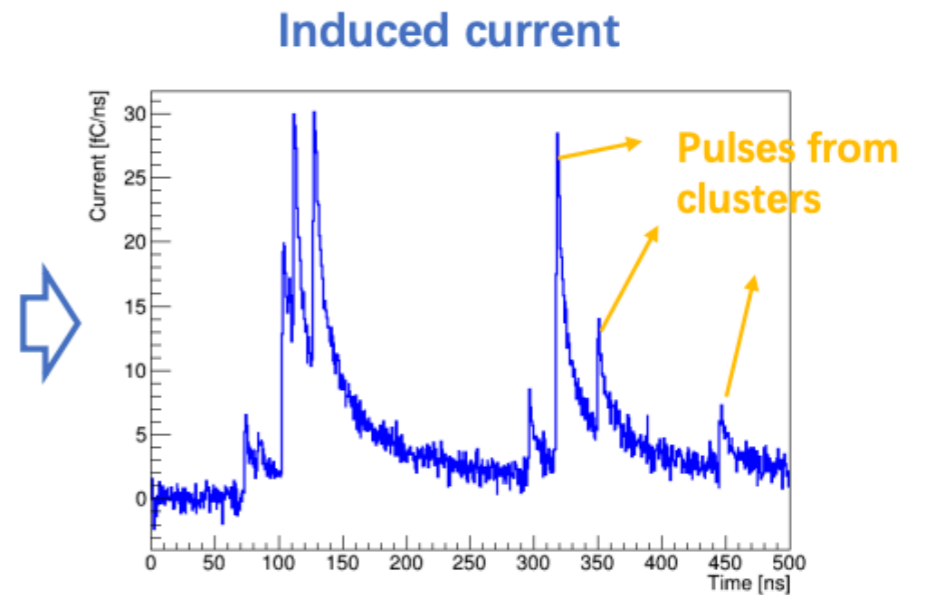
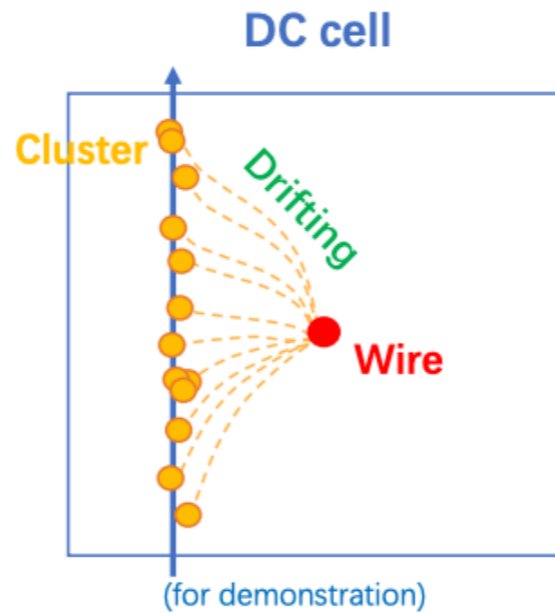
# 电离测量： $dE/dx$ 及 $dN/dx$

- ▶ 电离能损 ( $dE/dx$ )：测量总电离的积分
- ▶ 原初电离计数 ( $dN/dx$ )：测量原初电离 (cluster) 个数。



总电离 = 原初电离 + 次级电离

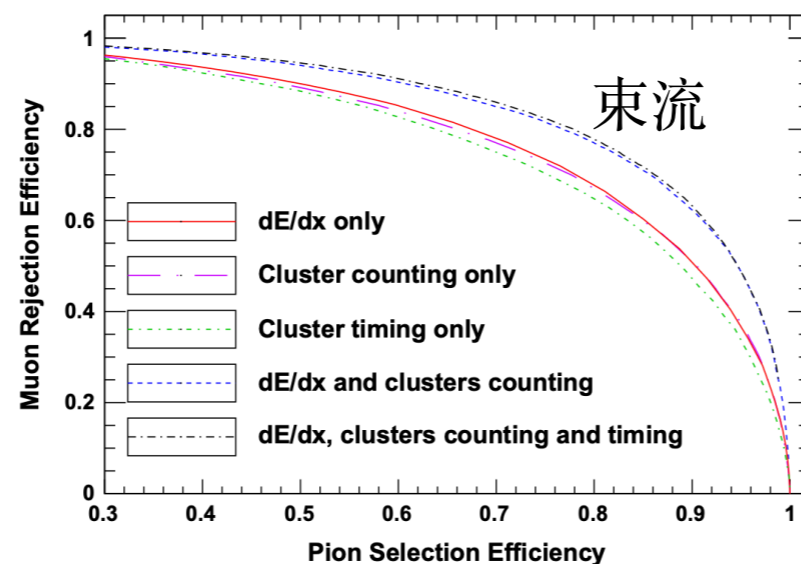
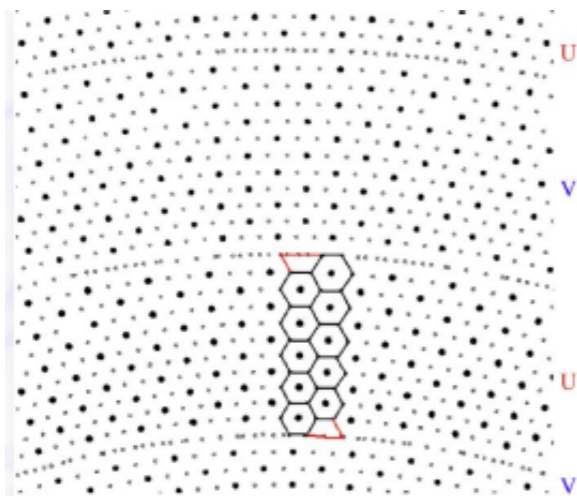
Cluster size分布：产生N个电子的概率



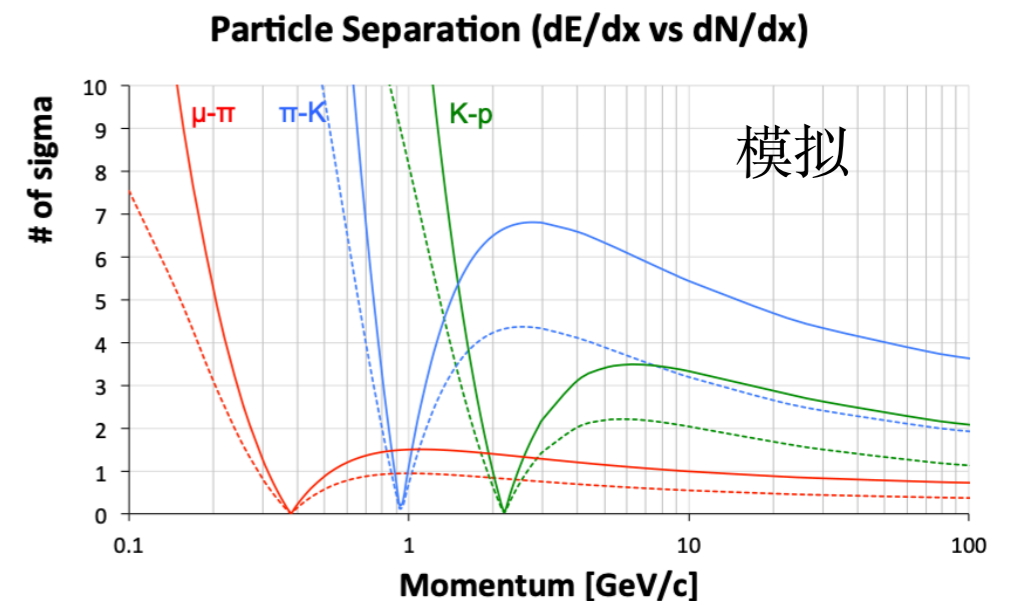
- ▶ Cluster size 的涨落导致了测量对电离取样的本征误差
  - ▶  $dE/dx$  为Landau分布：峰有较大的宽度和较长的尾巴（能力转移大的 $\delta$ 电子）
- ▶  $dN/dx$  的两大思路：靠时间/空间分辨。Ar气体：(20-30) cluster/cm,  $v=50 \mu\text{m}/\text{ns}$ , 意味着需要高时间分辨率 ( $\sim$ 几个ns) 或者高颗粒度 (几百 $\mu\text{m}$ ) 的探测器来分辨。
- ▶ 现状：被广泛使用的方法是测量电荷的积分，即“受限”能量损失， $dE/dx$ ，采用截断平均等方法估计其平均值。

# Cluster counting 研究概况

- ▶ 并不是一个新方法。概念在1930年代由E. J. Williams提出，70年代起出现详细的模拟和实验工作。然而，只有dE/dx方法真正用于实际的探测器上。
- ▶ 近二十年来，在 $e^+e^-$ 对撞实验利用cluster counting 开展的研究工作
  - ▶ ILC 第四代概念探测器（已停止）
  - ▶ Super-B 实验：束流测试结果，90% He 10% iC4H10 气体。
  - ▶ IDEA 探测器方案：CEPC 及 FCC 项目，模拟+束流实验
- ▶ 本文所展示的工作：
  - ▶ BESIII 高亮度升级方案
  - ▶ CEPC 第四代概念设计探测器



[J-F Caron, https://doi.org/10.1016/j.nima.2013.09.028](https://doi.org/10.1016/j.nima.2013.09.028)

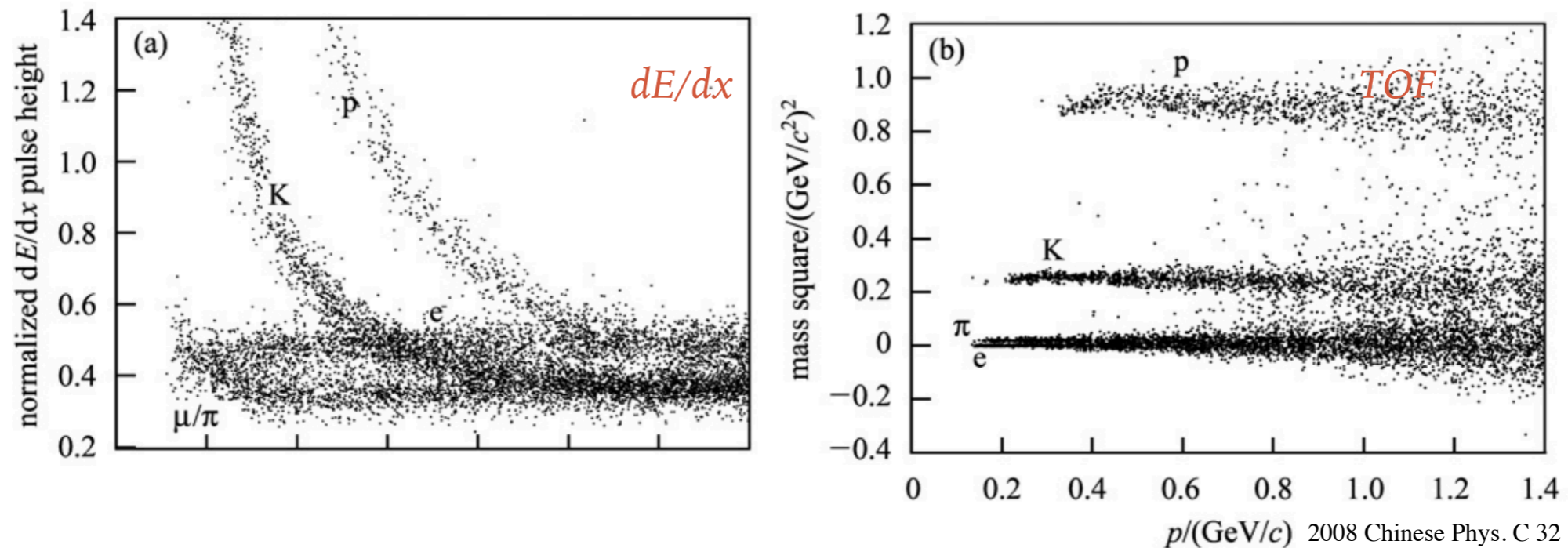


G. Chiarello et. al, NIM A 936 (2019) 503-504



# 基于BESIII漂移室的粒子鉴别研究动机

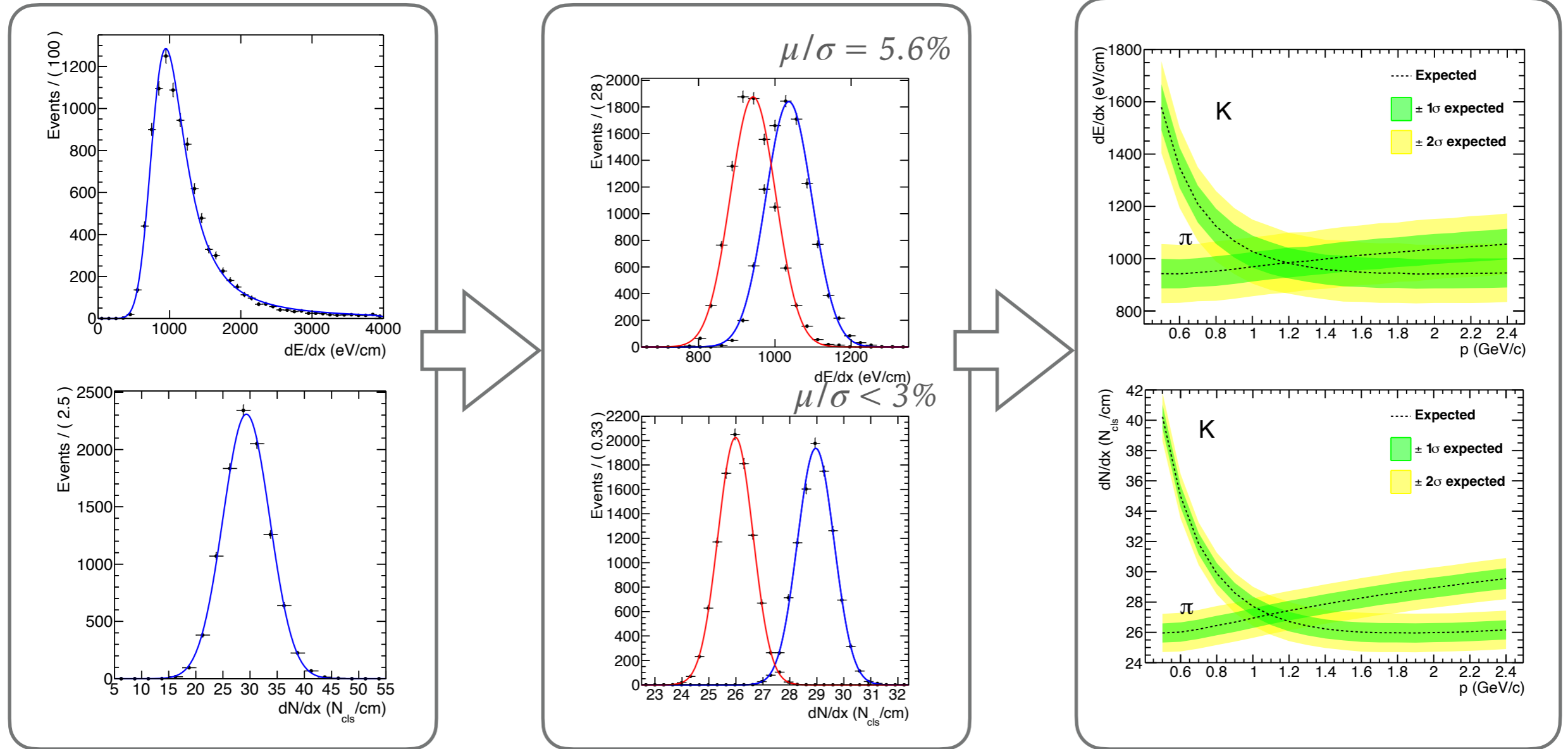
- ▶ 除了径迹探测器和量能器之外，PID在物理分析中是BESIII实验的一个关键方面。
- ▶ BESIII的PID系统由MDC的 $dE/dx$ （6%分辨率）测量和TOF（桶部68 ps，端盖60 ps）组成。
  - ▶ 相比CLEO-c、Belle 实验，使用切伦科夫探测器作为PID系统的补充，BESIII探测器的PID能力没有优势。



- ▶ BEPCII升级方案：最大束流能量  $2.3GeV \rightarrow 2.45GeV$ 
  - ▶ 出色的信号显著性需要优秀的PID能力：更高的PID效率和更低的误判率。
  - ▶ 随着MDC电子学的升级，cluster counting 方法提供了一种可能选择。

# 电离级别模拟

- ▶ 模拟带电粒子在 MDC 中的能量损失和电离数。
- ▶ 模拟环境：工作气体 He – C<sub>3</sub>H<sub>8</sub>60 : 40；磁场1T；MDC单元几何及信号丝/场丝。



单个单元模拟 (1.8 cm)

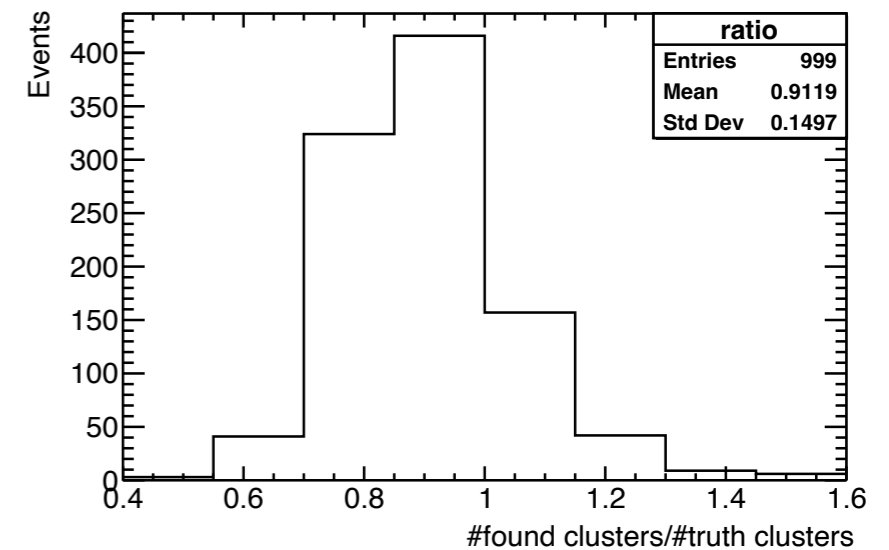
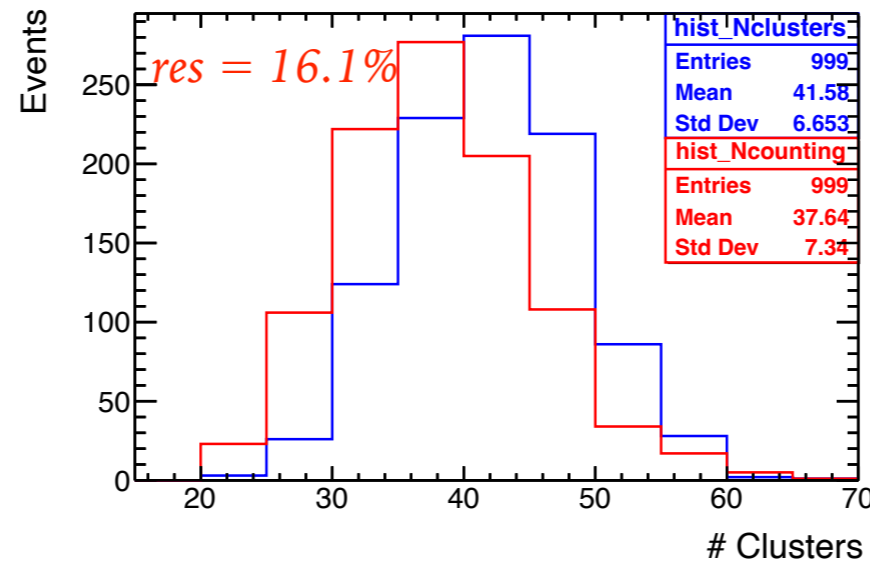
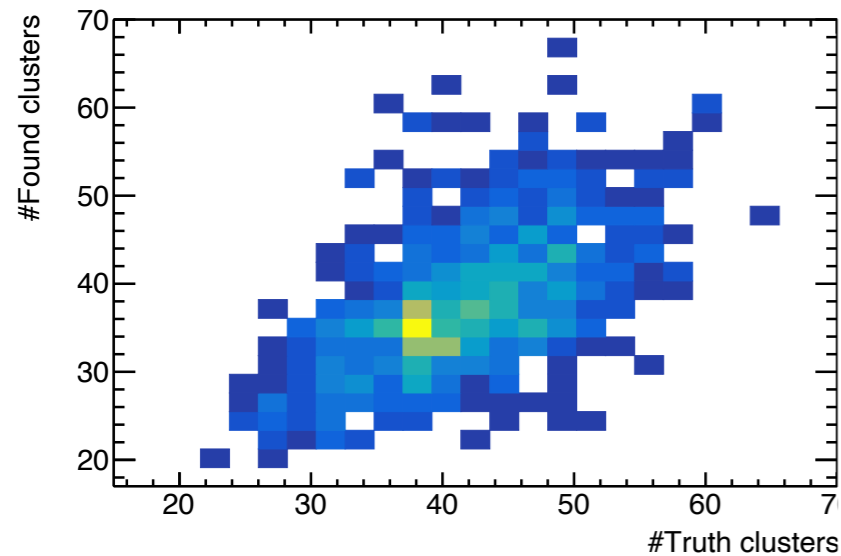
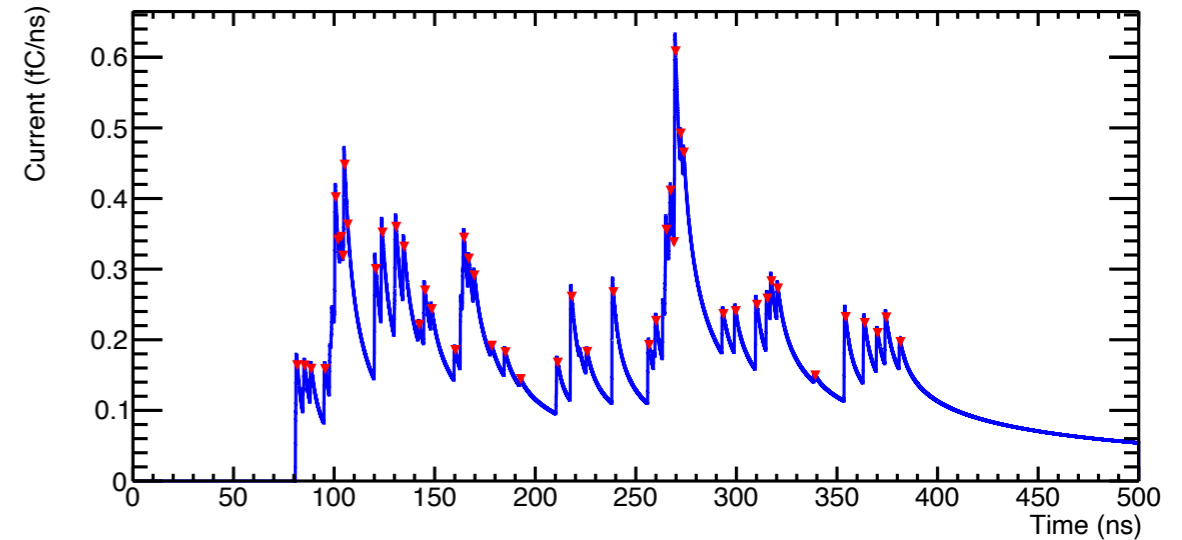
43 层模拟 (66 cm)  
dE/dx 截断：保留 5%-75%能损  
的击中

K/pi 分离趋势



# 基于波形的dN/dx模拟

- ▶ 改写了TSpectrum 寻峰算法，用来识别峰值。
- ▶ 理想条件的寻峰性能：eff = 91% @2GHz 采样率。
- ▶ 导致实验中分辨率下降的原因：
  - ▶ 较高的簇团密度，形成事例堆积（气体相关）
  - ▶ 二次电离（气体相关）
  - ▶ 电子学采样率、传输函数
  - ▶ 噪声使得峰值被误算

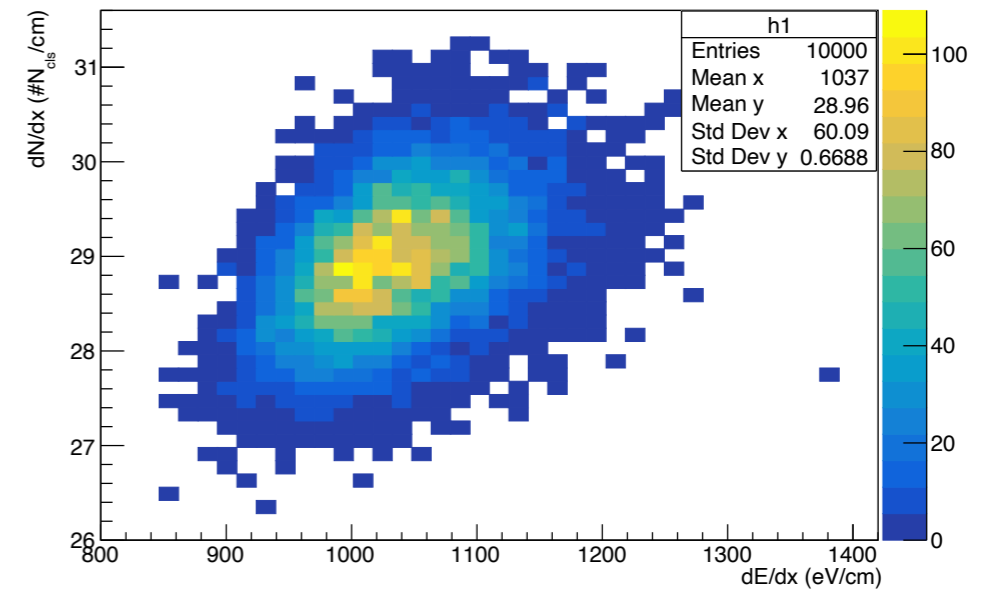
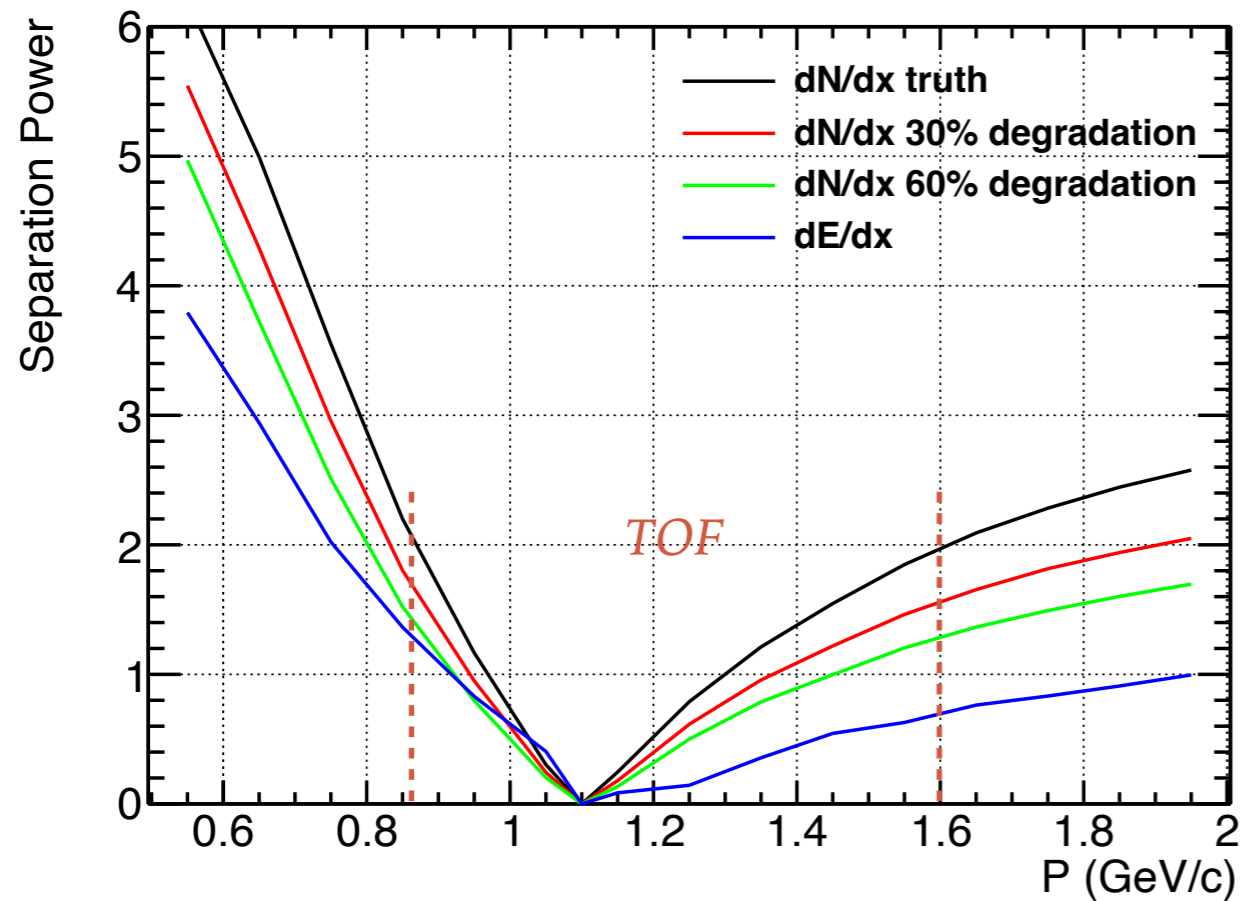


- ▶ 考虑所有影响做性能的定性分析，假设dN/dx分辨率变差30%，60%的情况。

# 波形模拟的dN/dx性能

- ▶ M是测量 (dE/dx或dN/dx) ，定义分离度

$$S_{K,\pi}^M = \frac{|\overline{M}_K - \overline{M}_\pi|}{\sqrt{\sigma(M_K)^2 + \sigma(M_\pi)^2}}$$



- ▶ 理想的dN/dx: 好2.5倍; dN/dx 60%退化: 好1.7倍。
- ▶ dN/dx 和 dE/dx 中度相关 ( $\rho \sim 0.4$ ) , 联合的PID性能提升不明显。

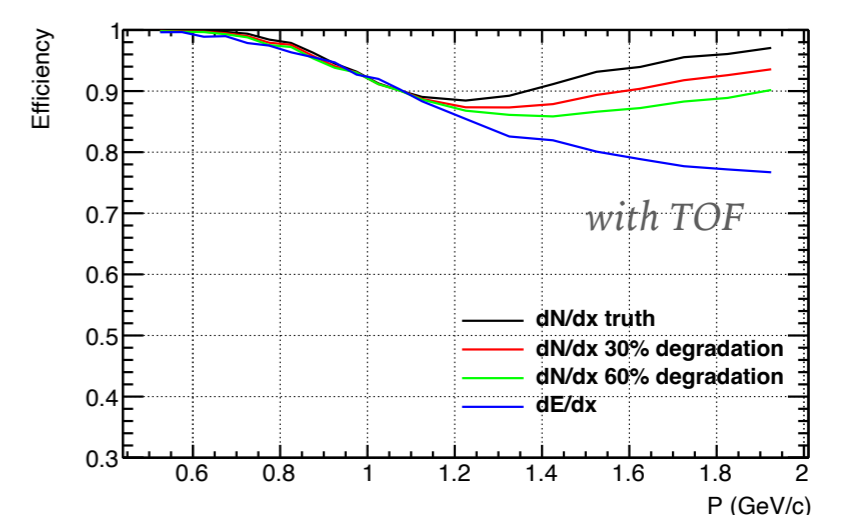
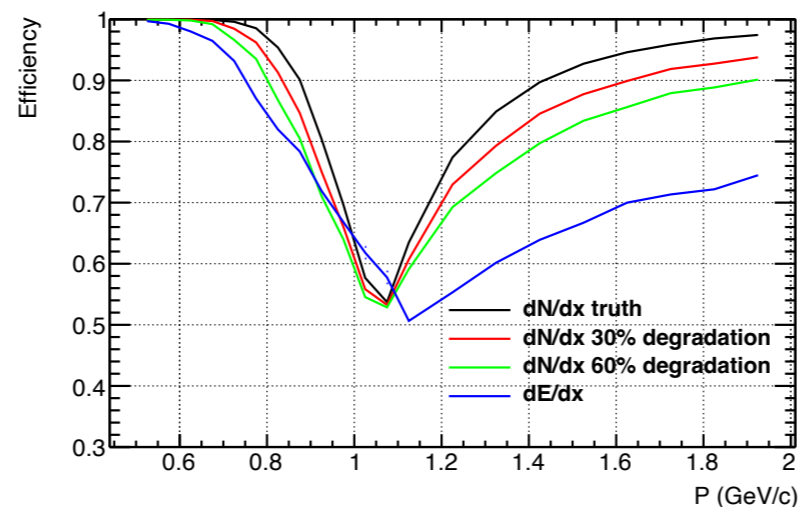
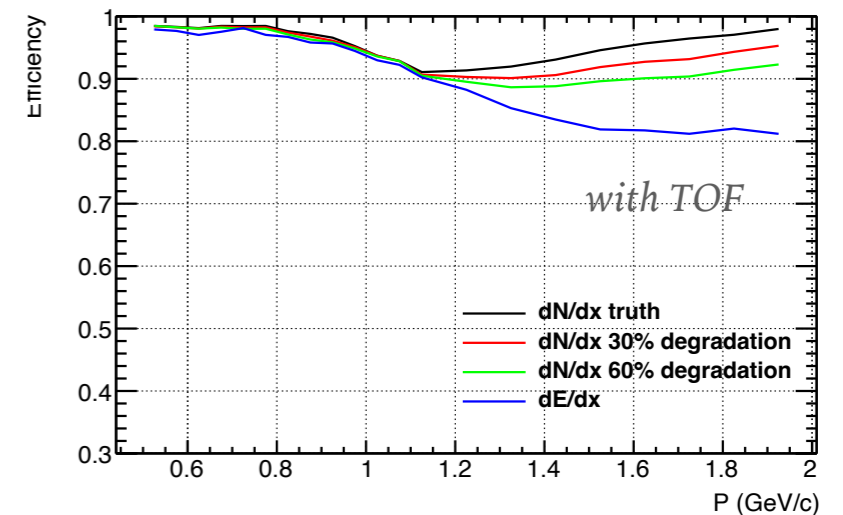
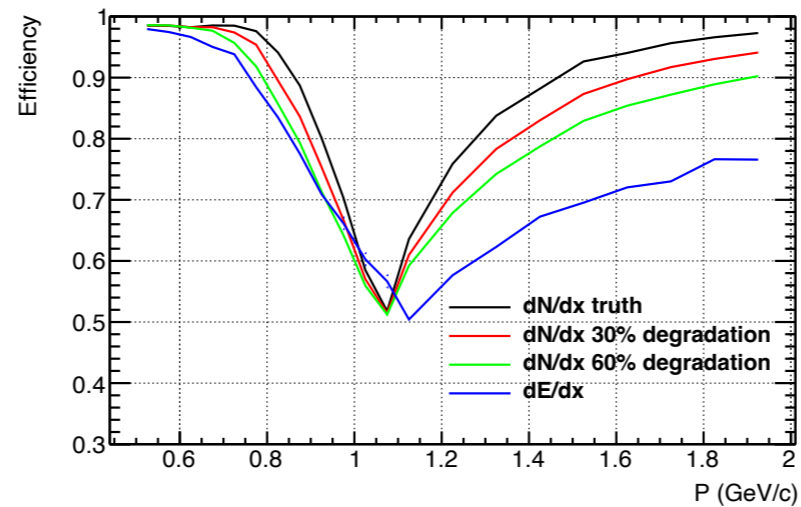
# BOSS下PID性能研究：PID 效率

- ▶ 模拟：利用BOSS框架产生的模拟数据研究。单粒子样本，筛选出无二次顶点的好径迹事例。
- ▶ 重建：dE/dx、TOF：BESIII 探测器重建级别的信息；dN/dx：径迹级别，从粒子的 $\beta\gamma$ 按照Garfield参数化结果抽样。
- ▶ PID 分析：Chi-squared 方法计算PID 效率：比较不同粒子假设的Chi-squared

$$\chi_{total}^2 = \chi_{dN/dx}^2 + \chi_{TOF}^2, \quad \chi_{dN/dx} = \frac{dN/dx_{meas} - dN/dx_{exp}}{\sigma_{dN/dx}}, \quad \chi_{TOF} = \frac{TOF_{meas} - TOF_{exp}}{\sigma_{TOF}}$$

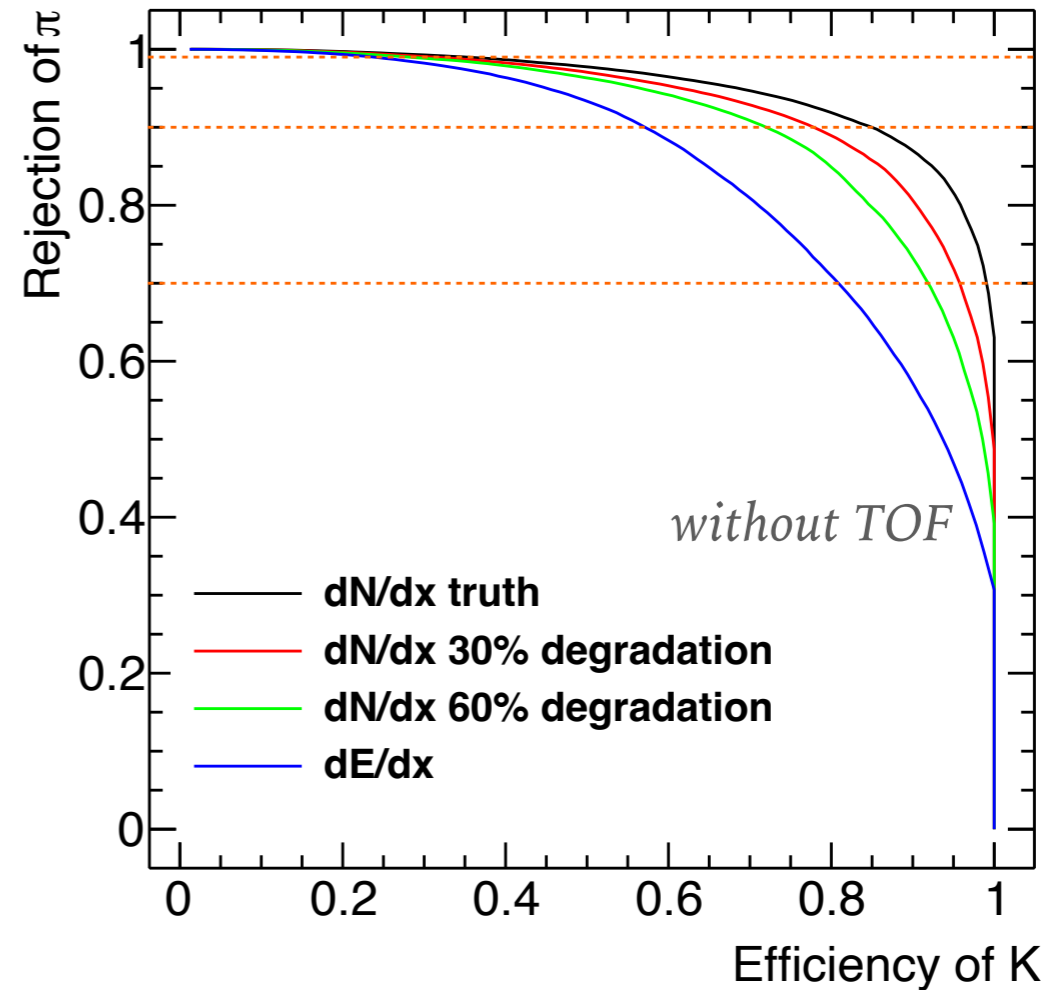
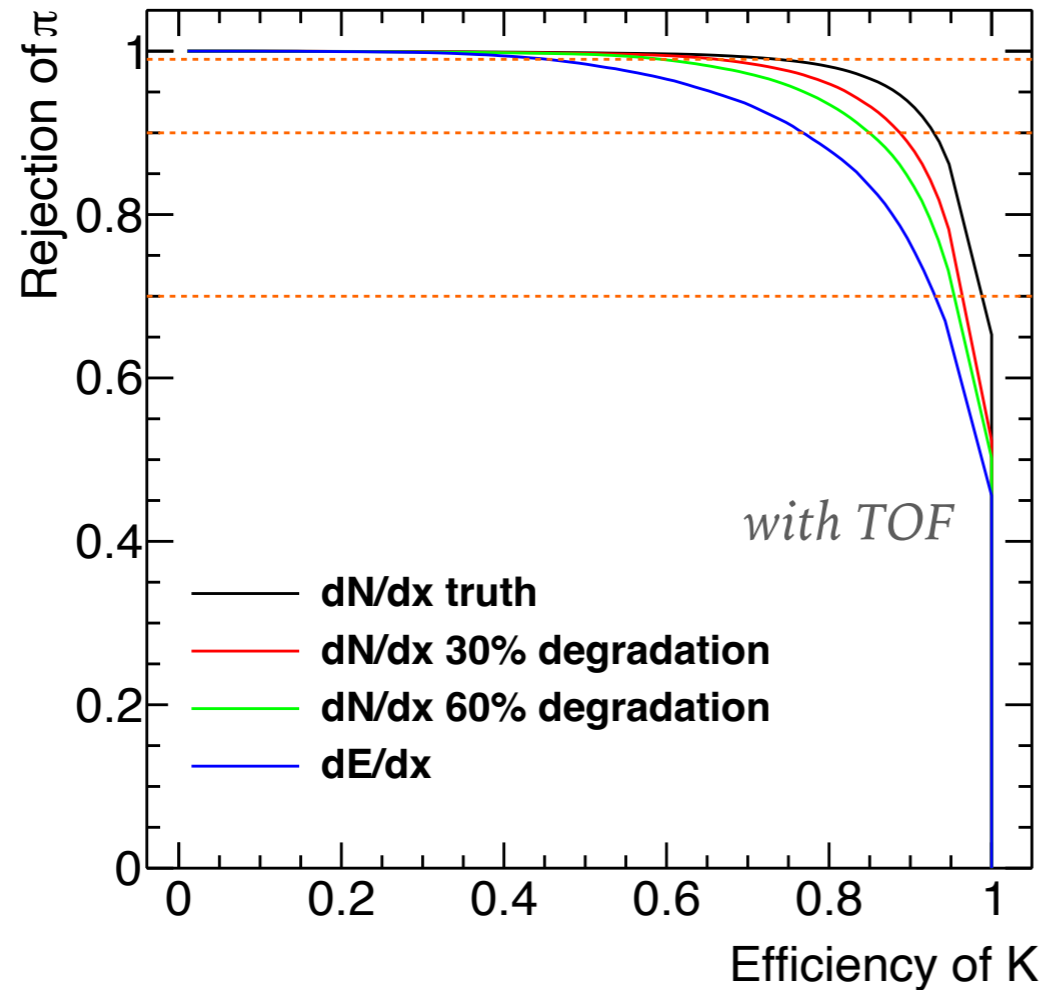
对于2GeV pion/kaon 径迹：

- Truth dN/dx : ~22%提升
- 30% dN/dx : ~18%提升
- 60% dN/dx : ~10%提升



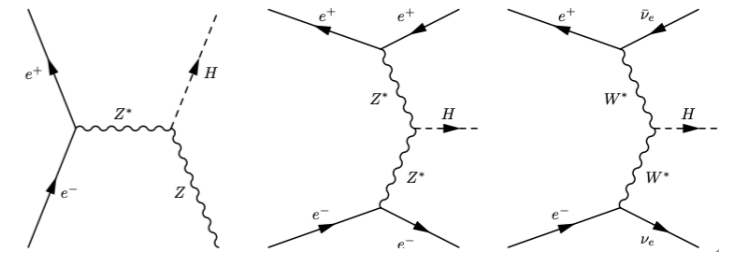
# BOSS下PID性能研究：ROC

► 以 $\pi$ 为本底，K的选择效率。



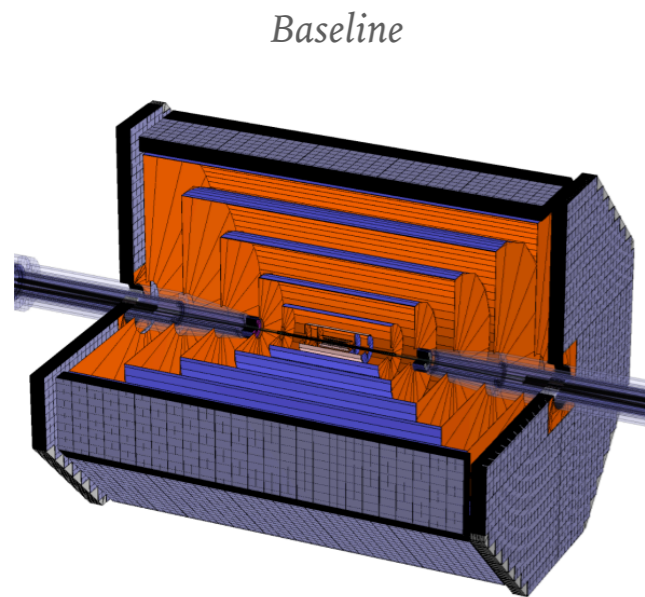
# CEPC 项目简介

- ▶ 核心任务：精确测量希格斯玻色子的性质（宽度、耦合、质量等）
- ▶ 精确测试标准模型（电弱物理、味物理、量子色动力学）
- ▶ 搜寻Z玻色子、B介子等的罕见衰变以及新物理。
- ▶ 新的运行策略提升了积分亮度、事例数及可能的 $t\bar{t}$ 运行。
- ▶ CDR中已有许多探测器设计方案

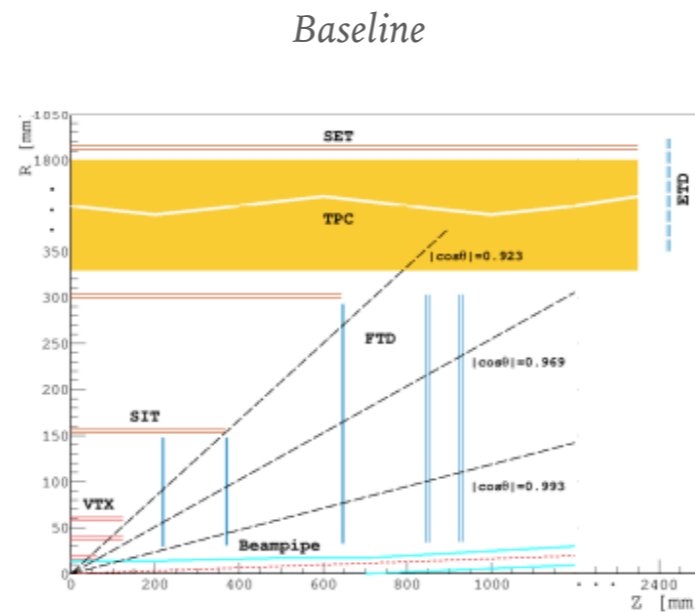


	Operation mode	Z factory	WW	Higgs factory	$t\bar{t}$
	$\sqrt{s}$ (GeV)	91.2	160	240	360
	Run time (year)	2	1	10	5
The CDR	Instantaneous luminosity ( $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , per IP)	32	10	3	-
	Integrated luminosity ( $\text{ab}^{-1}$ , 2 IPs)	16	2.6	5.6	-
	Event yields	$7 \times 10^{11}$	$2 \times 10^7$	$1 \times 10^6$	-
New nominal	Instantaneous luminosity ( $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , per IP)	191.7	26.6	8.3	0.83
	Integrated luminosity ( $\text{ab}^{-1}$ , 2 IPs)	100	6	20	1
	Event yields	$3 \times 10^{12}$	$1 \times 10^8$	$4 \times 10^6$	$5 \times 10^5$

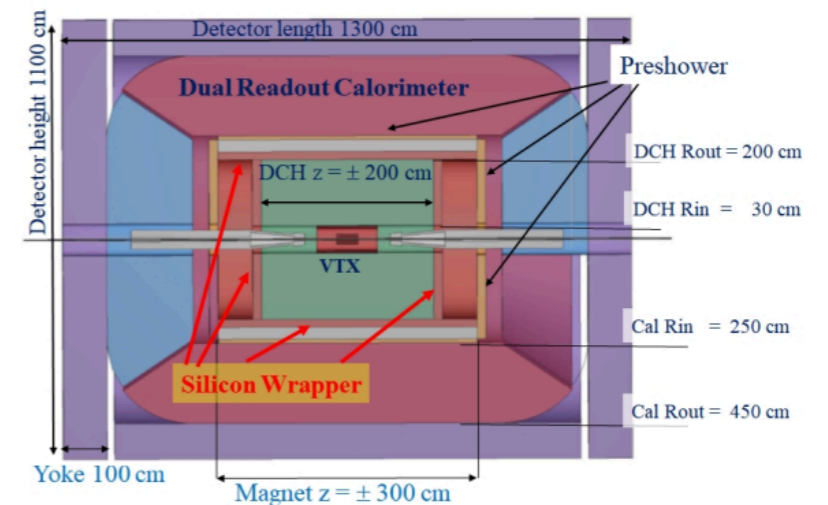
Alternative: IDEA



全硅, ILD设计



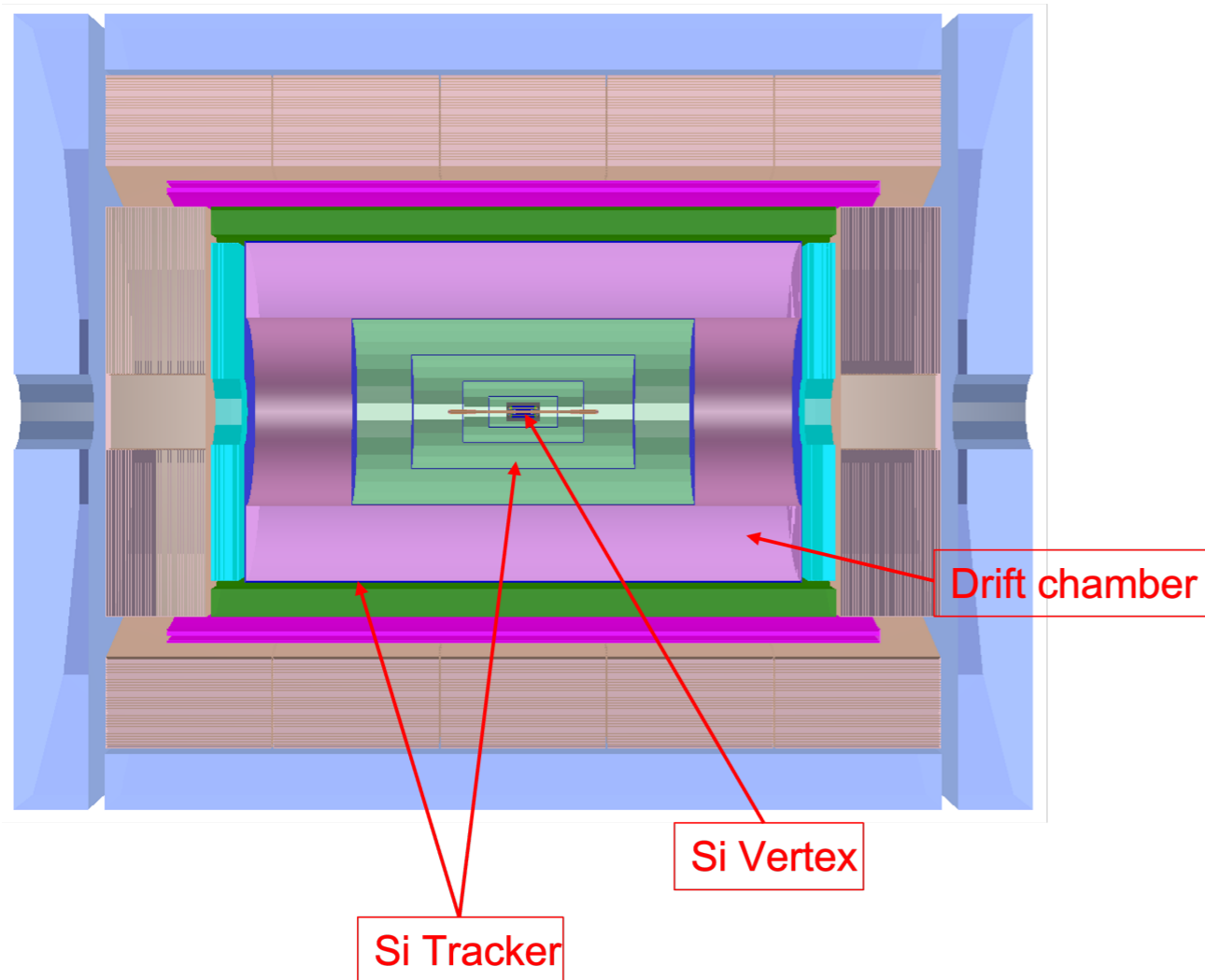
硅径迹+TPC+PFA-E/HCAL



硅径迹+漂移室

# 第四代概念设计探测器

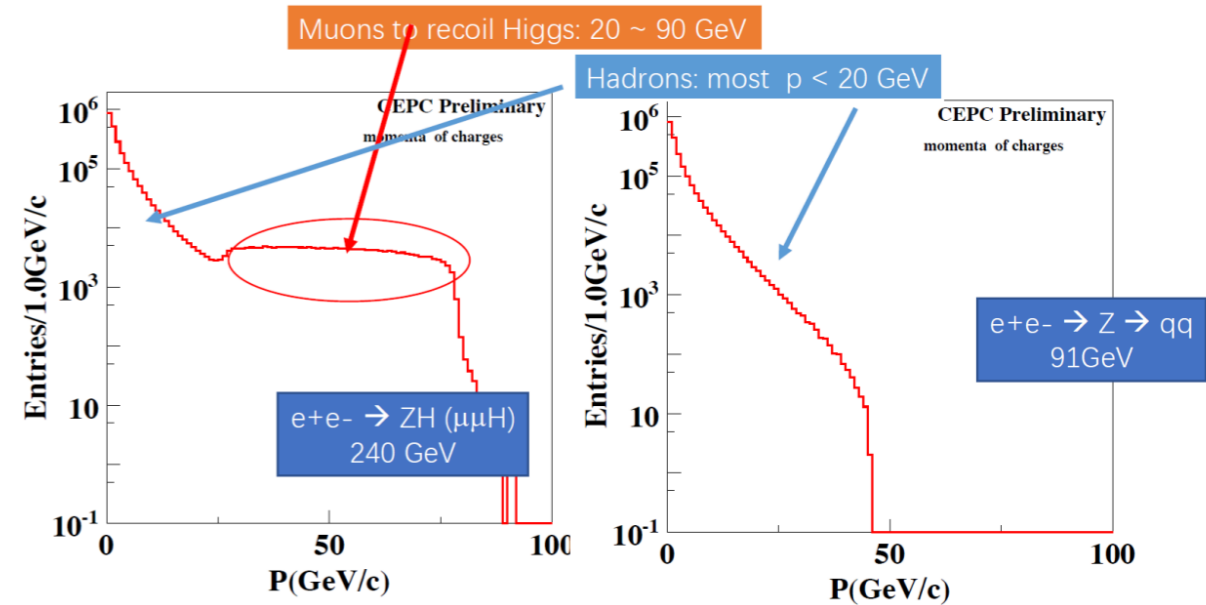
- ▶ 漂移室是CEPC第四代概念探测器中的一种设计
  - ▶ 位于全硅径迹层（0-0.8 m）之间，用于粒子识别（PID）
  - ▶ 通过应用原初电离计数技术进行PID优化。
- ▶ Z-pole运行物理目标：在 20GeV 达到 $2\sigma$  p/K 区分能力。



对探测器需求

Physics process	Measurands	Detector subsystem	Performance requirement
$ZH, Z \rightarrow e^+e^-, \mu^+\mu^-$ $H \rightarrow \mu^+\mu^-$	$m_H, \sigma(ZH)$ $BR(H \rightarrow \mu^+\mu^-)$	Tracker	$\Delta(1/p_T) = 2 \times 10^{-5} \oplus \frac{0.001}{p(\text{GeV}) \sin^{3/2} \theta}$
$H \rightarrow b\bar{b}/c\bar{c}/gg$	$BR(H \rightarrow b\bar{b}/c\bar{c}/gg)$	Vertex	$\sigma_{r\phi} = 5 \oplus \frac{10}{p(\text{GeV}) \times \sin^{3/2} \theta} (\mu\text{m})$
$H \rightarrow q\bar{q}, WW^*, ZZ^*$	$BR(H \rightarrow q\bar{q}, WW^*, ZZ^*)$	ECAL HCAL	$\sigma_E^{\text{jet}}/E = 3 \sim 4\%$ at 100 GeV
$H \rightarrow \gamma\gamma$	$BR(H \rightarrow \gamma\gamma)$	ECAL	$\Delta E/E = \frac{0.20}{\sqrt{E(\text{GeV})}} \oplus 0.01$

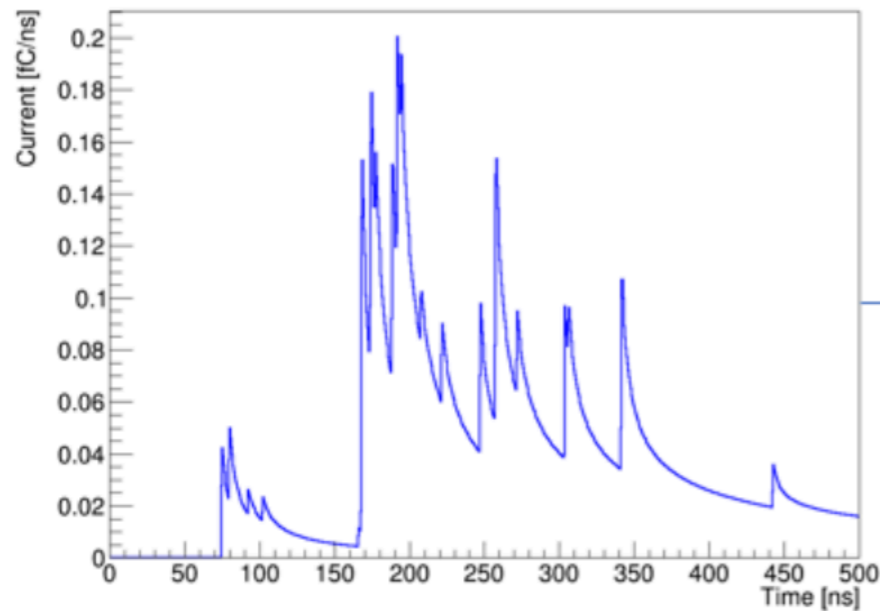
PID: 在 20GeV 达到 $2\sigma$  p/K 区分能力



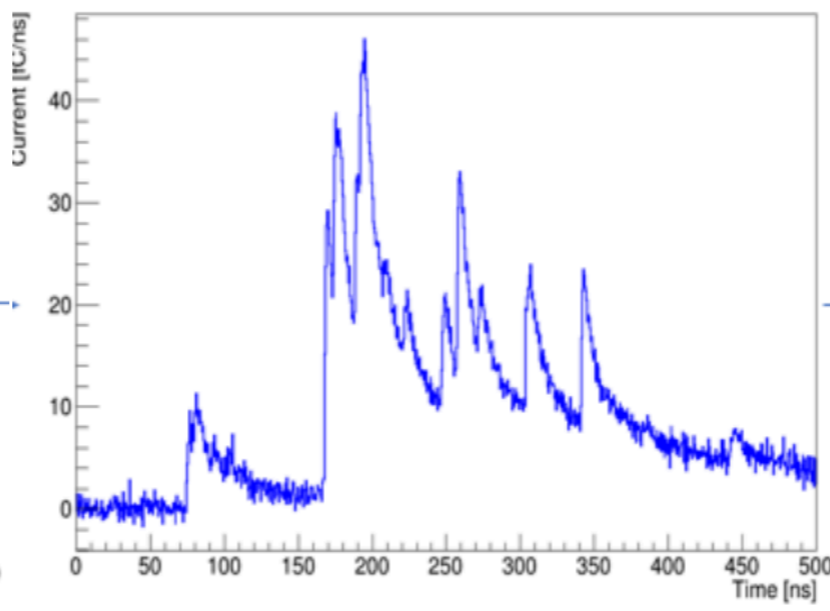


# 基于波形的模拟流程

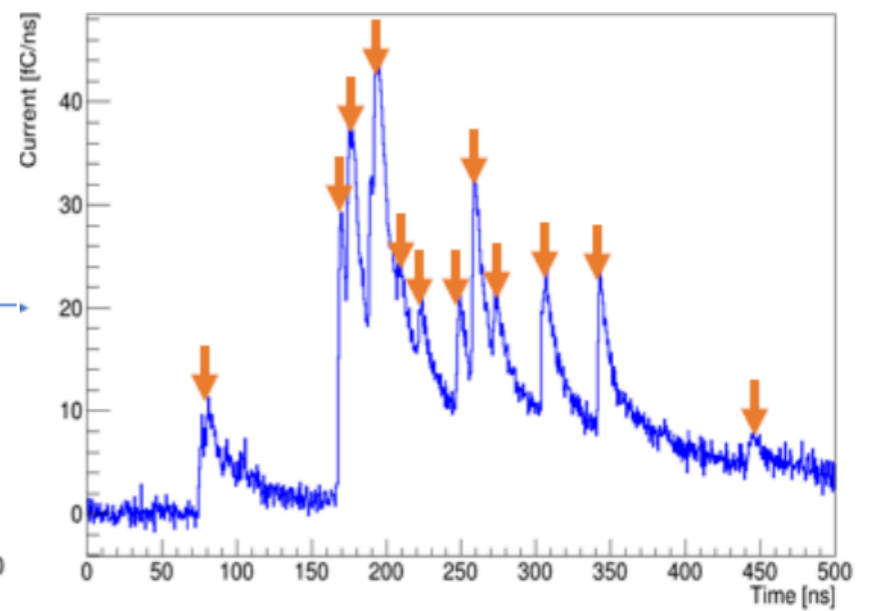
- ▶ Cluster counting 的关键内容是波形处理，端到端的工作流程包括：
  - ▶ 信号模拟、数字化、重建三个步骤。



模拟



数字化



重建

信号产生（基于Garfield++）

- ▶ *Heed*: 电离过程模拟
- ▶ *Magboltz*: 气体相关性质（漂移/扩散）

电子学

- ▶ 前置放大器
- ▶ 噪声
- ▶ ADC

重建算法：寻峰、聚类

- ▶ 微分算法
- ▶ 基于机器学习算法

# 数字化

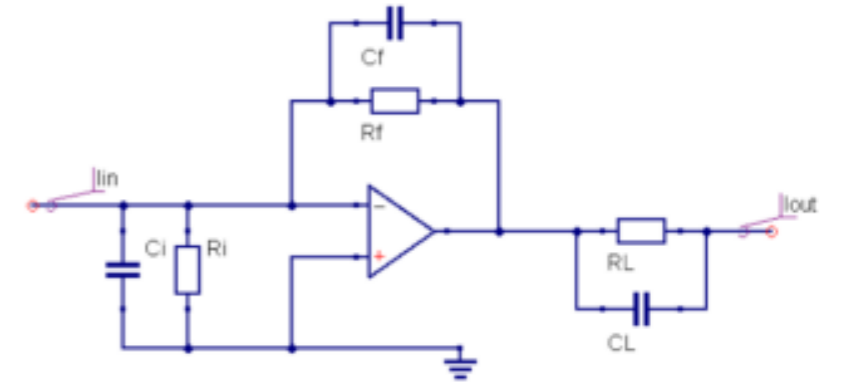
► 假设一个简单的电荷灵敏前置放大器。

► 将电子学传输函数卷积到原始信号上。

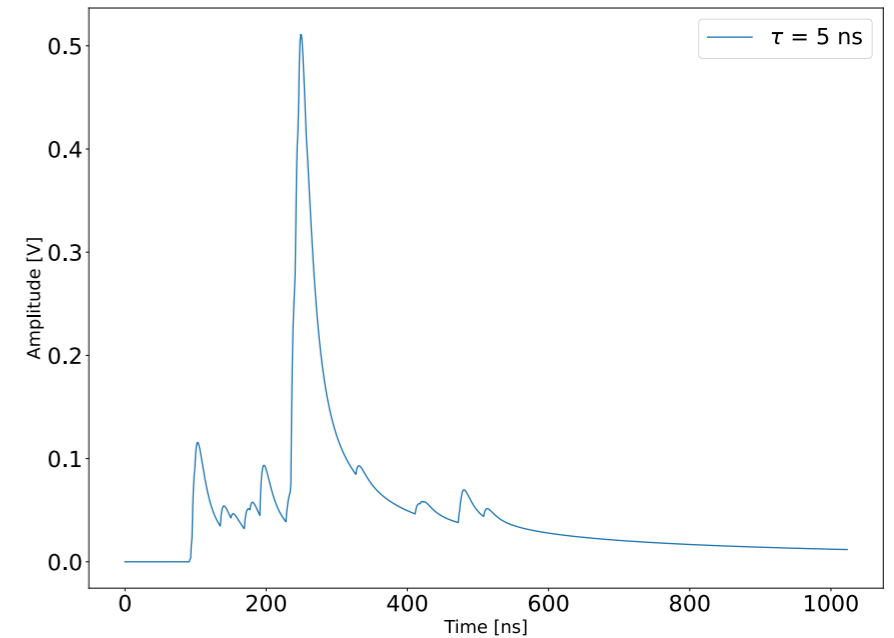
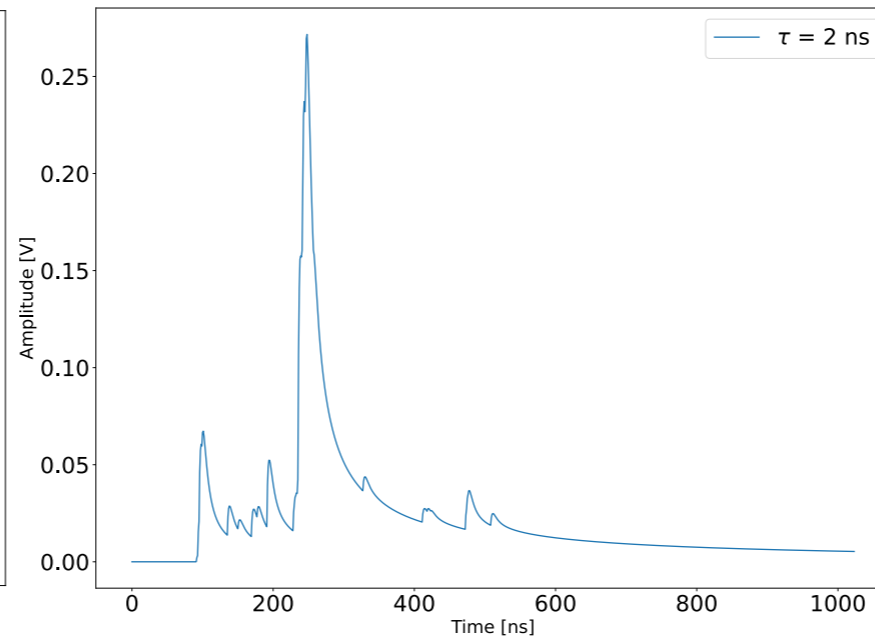
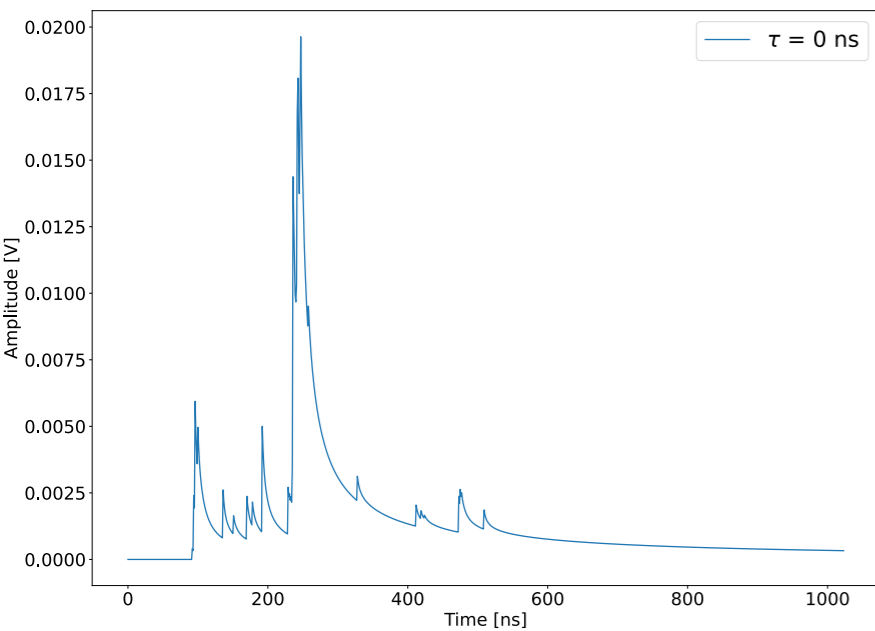
► 脉冲在时域空间的响应表示为： $h(t) = 1/\sqrt{R_f C_f} \times e^{-t/R_f C_f}$ ,  
 $R_f C_f = \tau$  (时间常数)

►  $\tau$  不同程度的平滑原始信号。

► 定义“上升时间”为脉冲达峰的90%幅度所对应的时间。



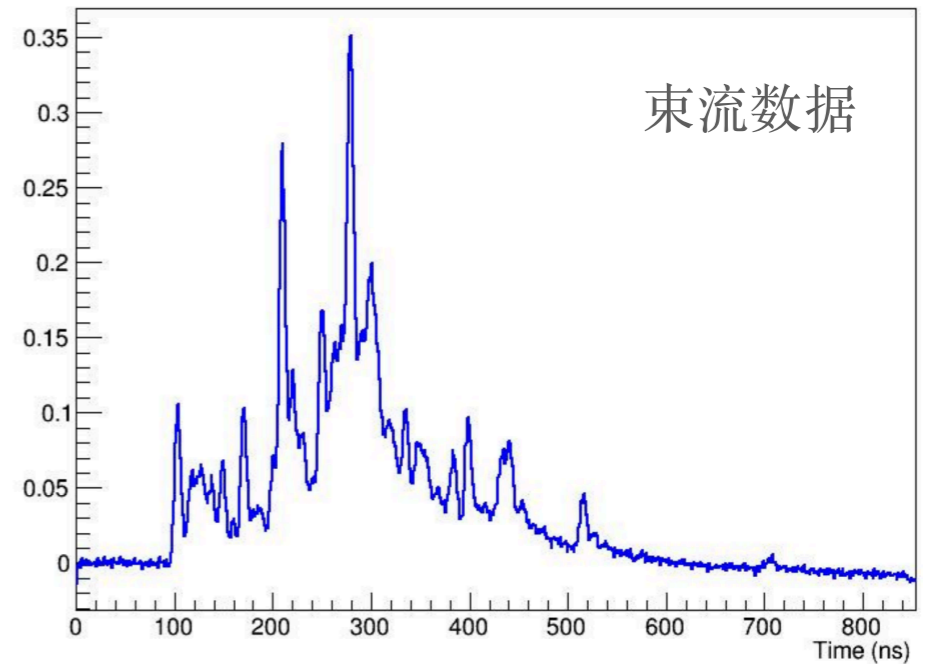
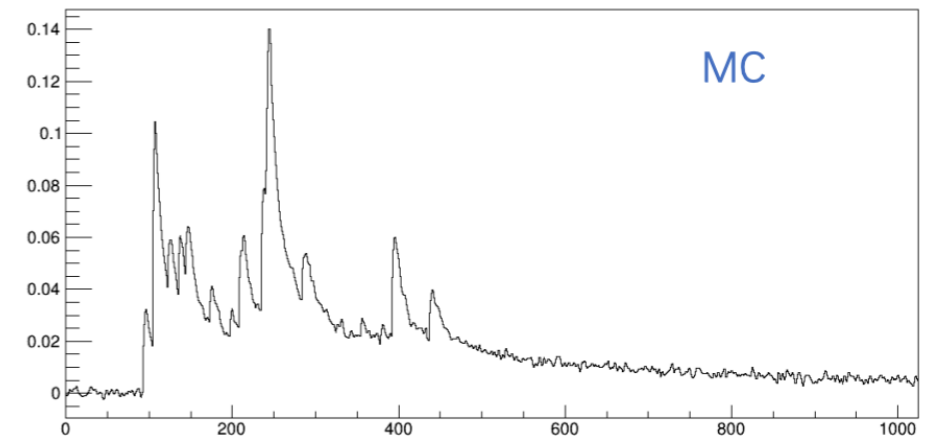
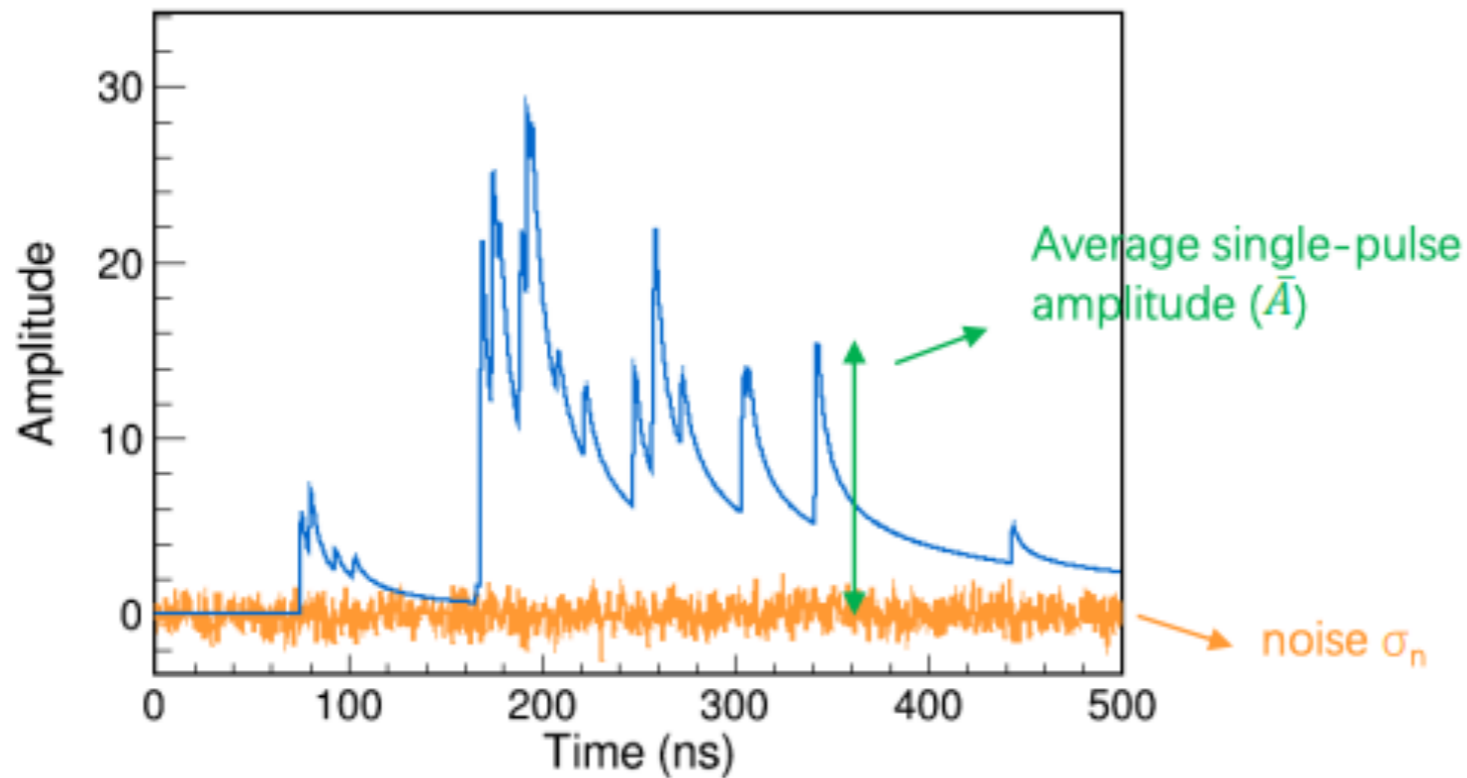
$$R_i \parallel C_i \gg R_f \parallel C_f \text{ and } C_L = 0$$



# 数字化

## ► 噪声模拟

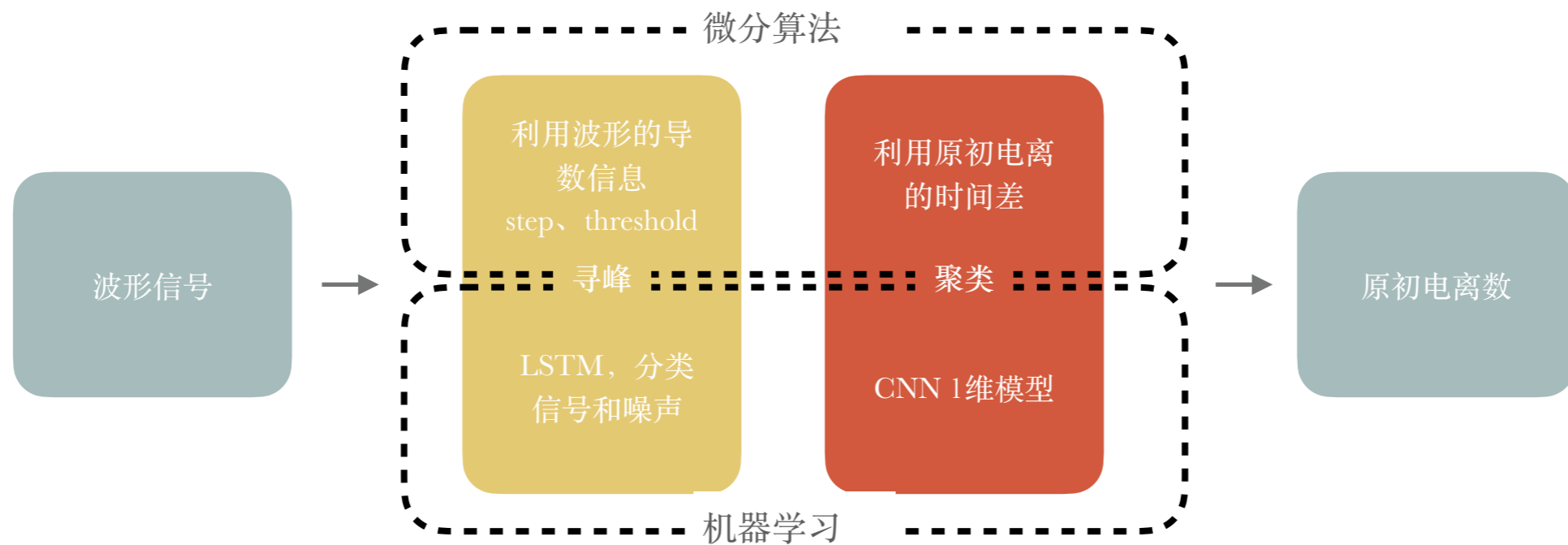
- 利用FFT从实验数据中的纯噪声样本提取频域信息。
- 假设均匀的相位分布，利用FFT拟变换产生随机噪声。
- 噪声水平定义： $\sigma_{Noise}/A_{signal}$



(比较噪声水平)

# 重建

- ▶ 重建原初电离实际上有两步工作：
  - ▶ 寻峰 (peak finding)：寻找波形上与噪声区分出来的电子（原初电离+二次电离）脉冲。
  - ▶ 聚类 (clusterization)：合并归属于同一原初电离的电子。
- ▶ 开发了两条重建算法路线：传统算法（基于微分）和机器学习（基于RNN，CNN模型）算法。



## ▶ LSTM

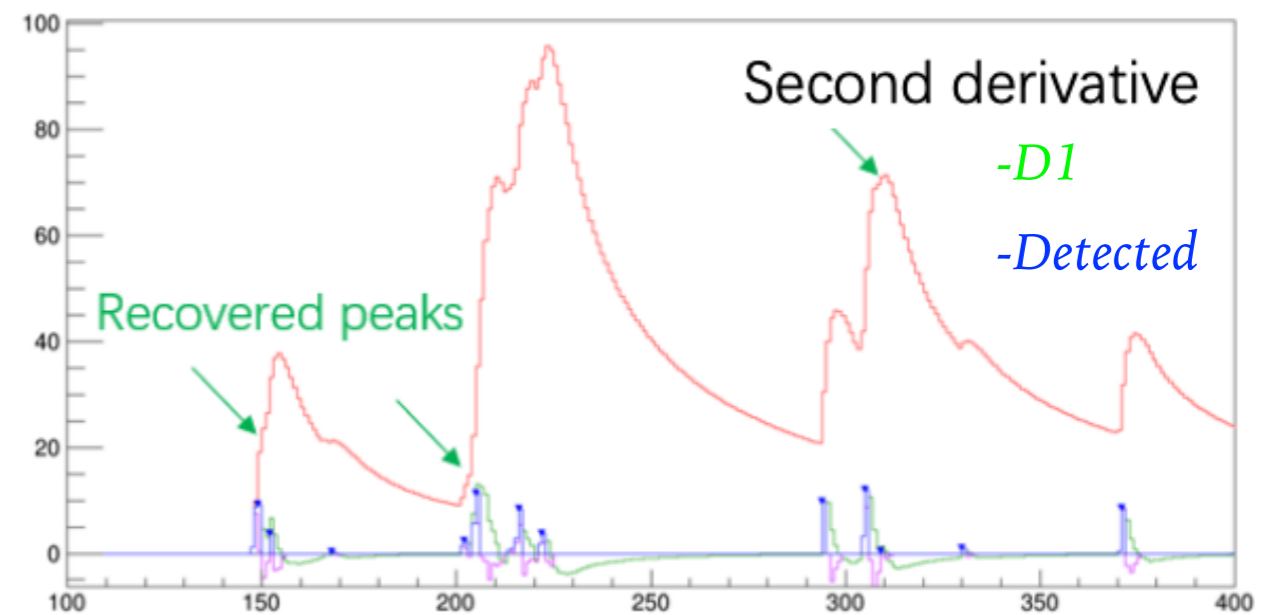
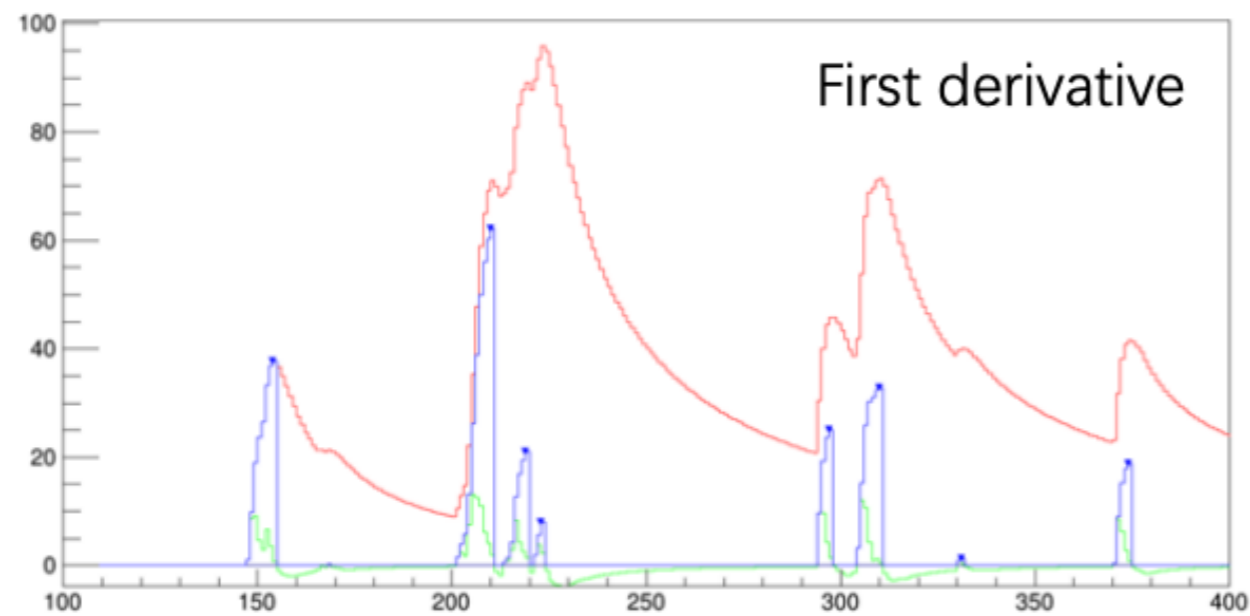
- ▶ 保留先前样本的“上下文”信息
- ▶ 适用于对事件顺序重要的数据集
- ▶ 寻峰

## ▶ CNN

- ▶ 无人干预下检测明显特征
- ▶ 提供高效的网络模型
- ▶ 聚类

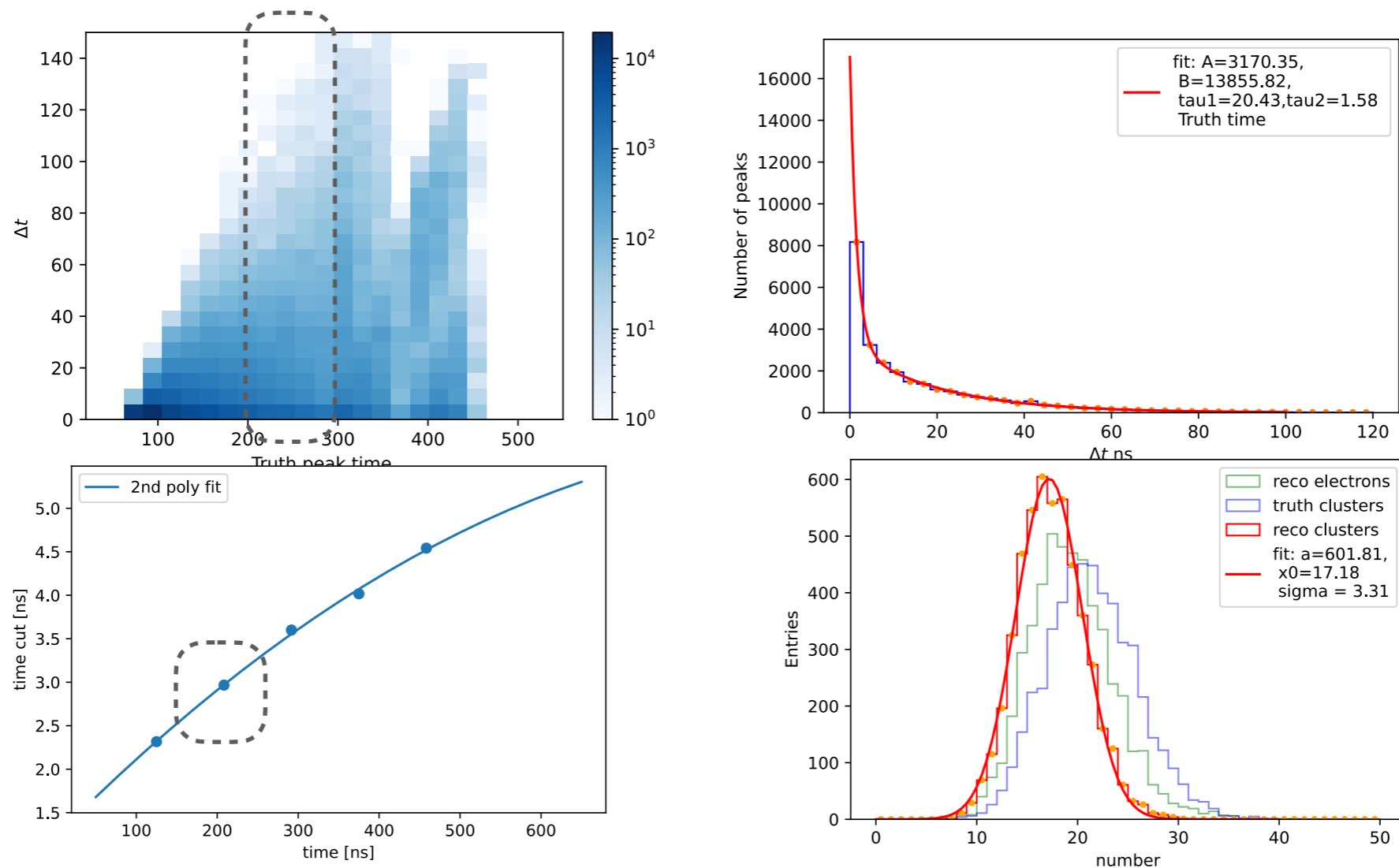
# 微分重建方法-寻峰

- ▶ 算法基于二阶导数，导数从正转变为负的“零点”对应峰的位置。
  - ▶ 一阶导数方法容易丢失位于上升沿的脉冲。
- ▶ 微分方法的优势在于
  - ▶ 简单，高效，对基线不敏感
  - ▶ 易于在硬件设备上实现（从前端存储完整波形所需资源）
- ▶ 缺陷
  - ▶ 只利用到边缘变化的信息
  - ▶ 在严重事例堆积和高噪声的情况下效率下降明显。



# 微分重建方法-聚类

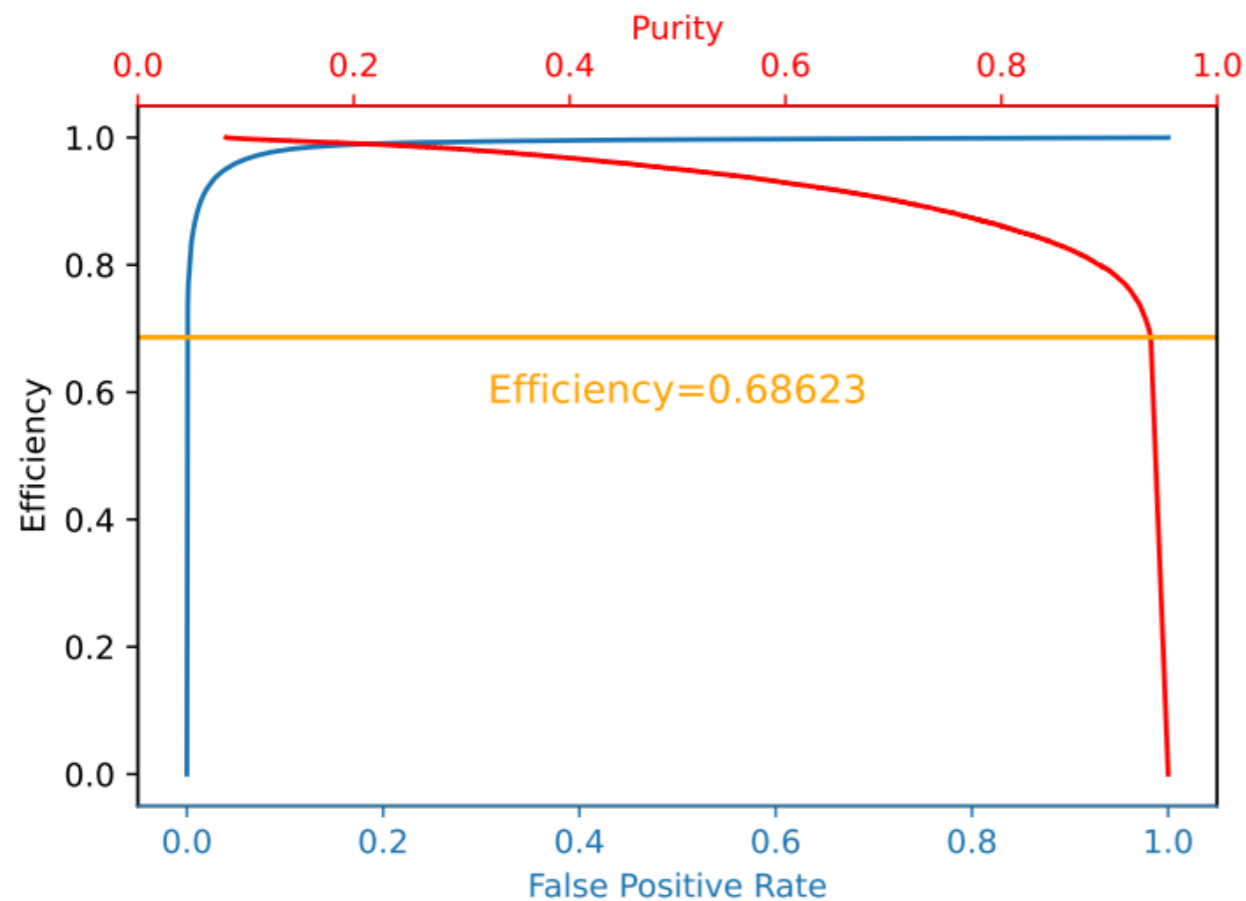
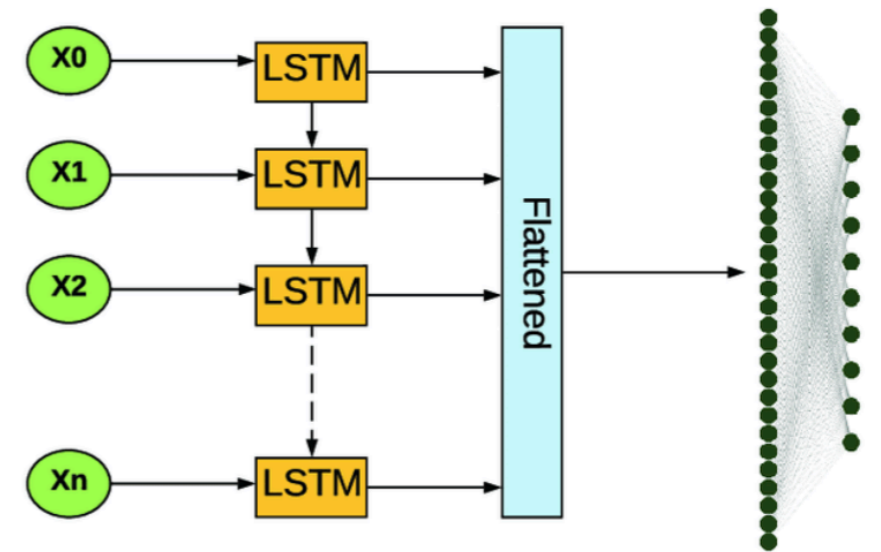
- ▶ 聚类策略：利用时间截断将连续的电子峰合并，时间截断取决于漂移时间。
- ▶ 认为连续电子的时间差分为两部分： $f(t) = Ae^{-t/\tau_1} + Be^{-t/\tau_2}$ 
  - ▶  $\tau_1$ 描述来自不同cluster的电子的漂移时间差， $\tau_2$ 描述来自相同cluster的电子的漂移时间差
  - ▶  $\tau_2$ 与电子的纵向扩散时间（2~4 ns）相当。
- ▶ 时间截断：当t满足 $\text{prob}(Ae^{-t/\tau_1}) > \text{prob}(Be^{-t/\tau_2})$ ，小于t的电子被合并。





# 机器学习重建算法

- ▶ 全监督的二分类学习
  - ▶ 标签：信号或噪声
  - ▶ 输入特征：长度为 (15, 1) 的连续波形
- ▶ LSTM 模型：1维的LSTM序列，连接平滑层和全连接层
- ▶ 模型性能：

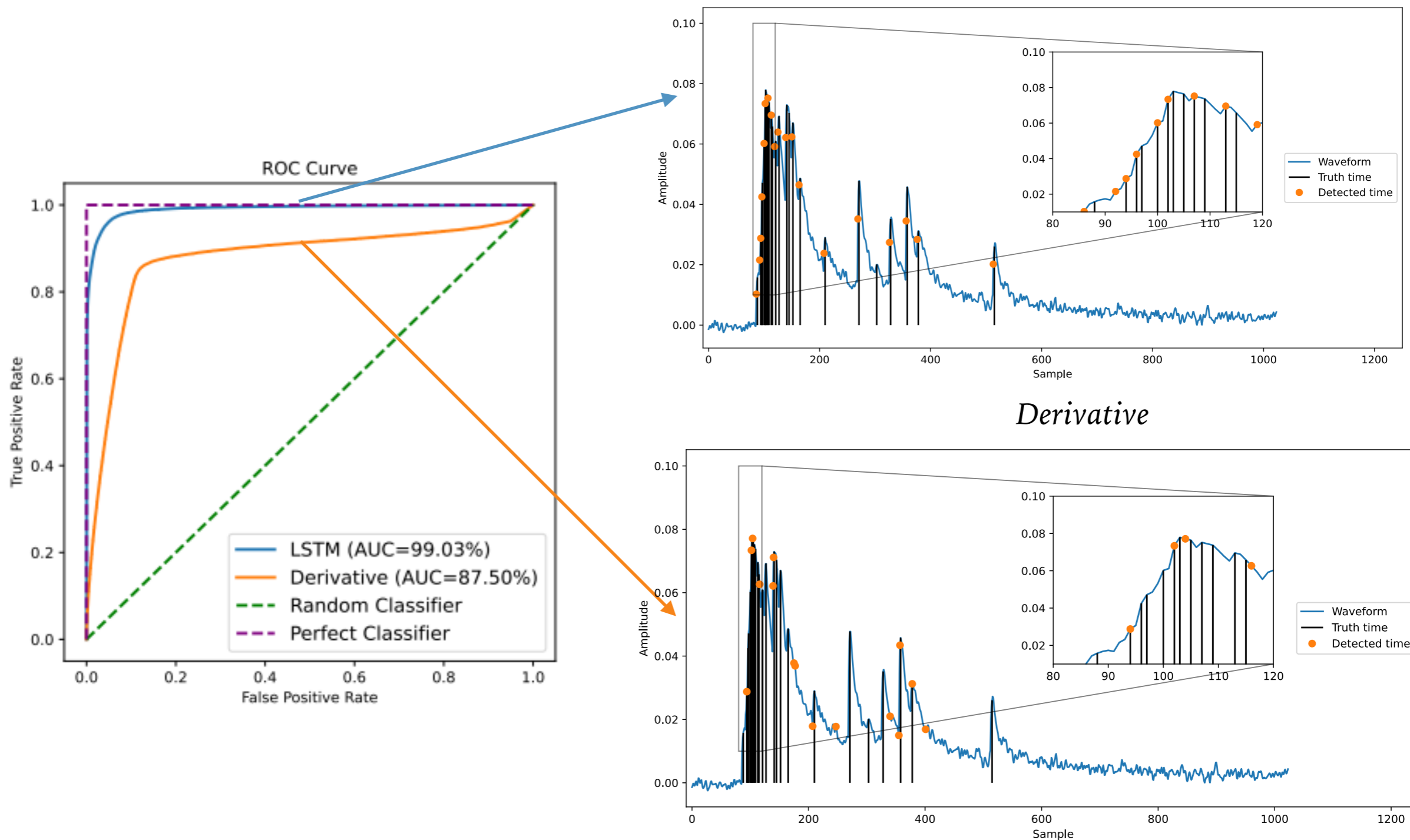


		Prediction	
		Signal	Noise
Truth	Signal	TP	FN
	Noise	FP	TN

- 纯度 =  $TP / (TP + FP) = 0.892$
- 效率 =  $TP / (TP + FN) = 0.686$
- 误判率 =  $FP / (FP + TN) = 0.0005$

# Peak finding 算法性能比较

► LSTM在事例堆积区间有更强的计数能力 (AUC ~ 99%)



# 机器学习重建算法-聚类

► 解释为回归问题，利用寻峰步骤的结果作为输入，预测波形的cluster数量。

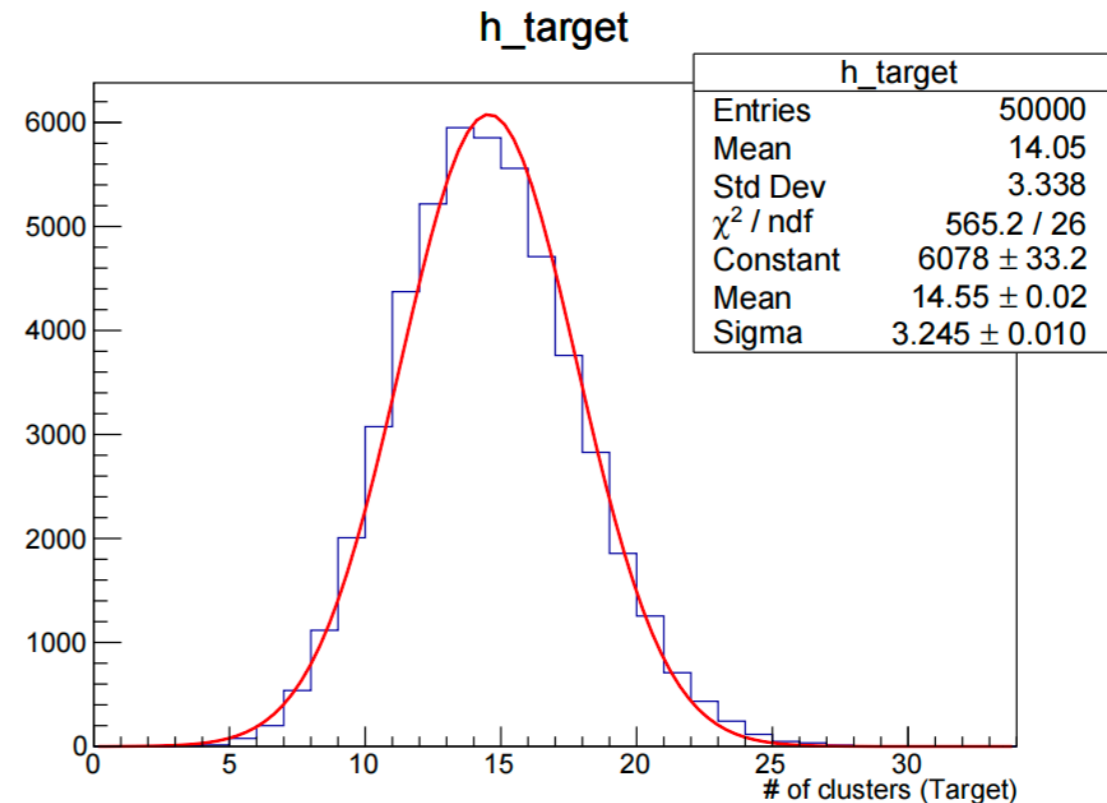
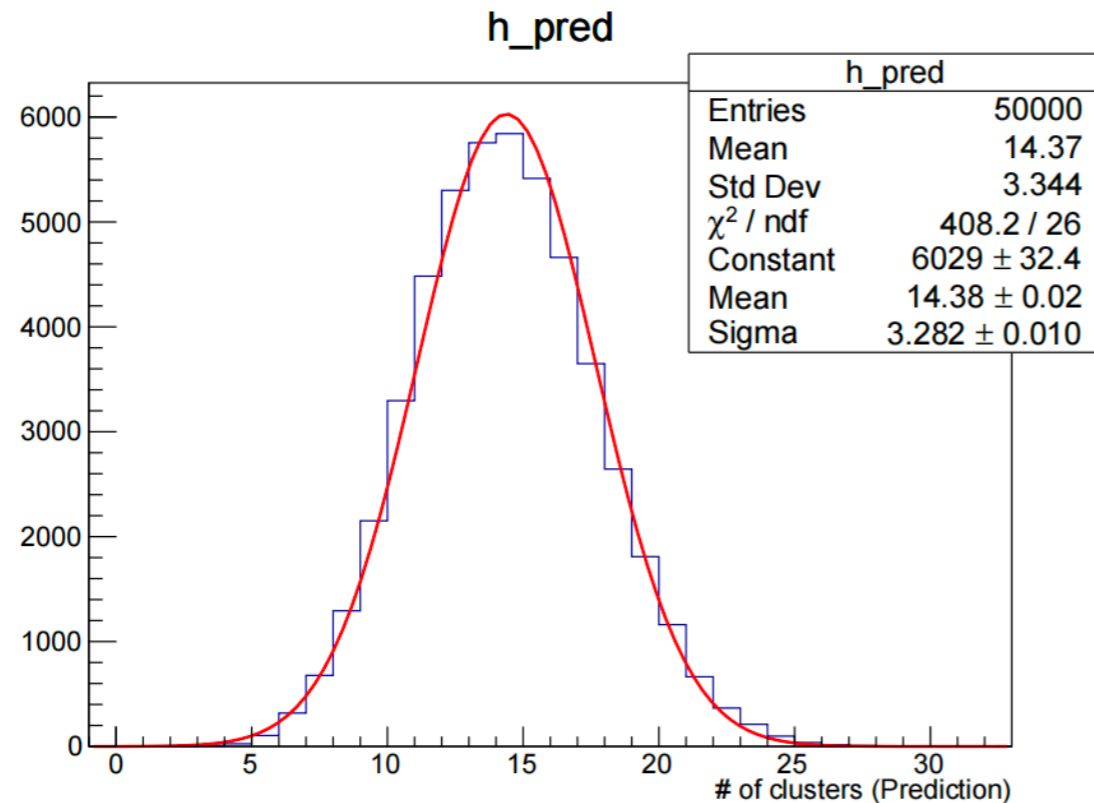
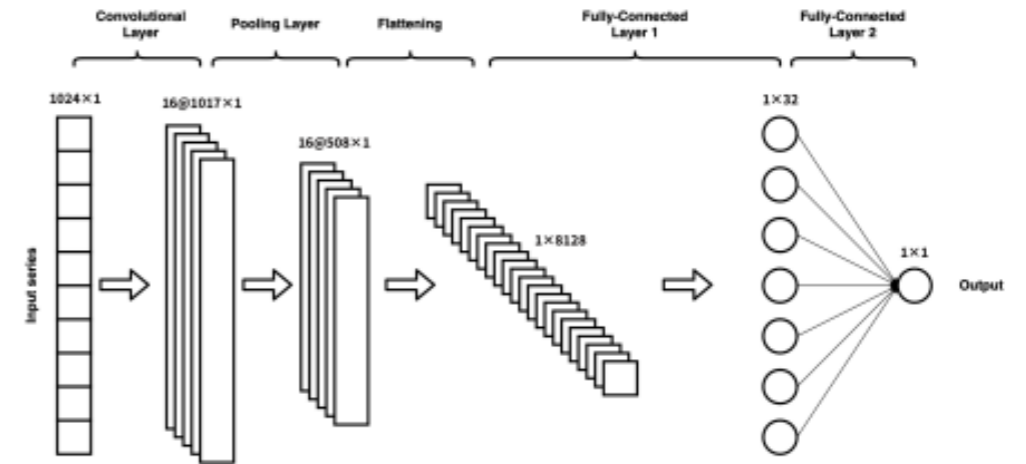
► 标签：cluster数量

► 特征：（第一步）标记的电子峰的时间

► 重建性能：

► 单个单元的分辨率 ( $\sigma/\mu$ ) : 22.8% (真实值: 22.3%)

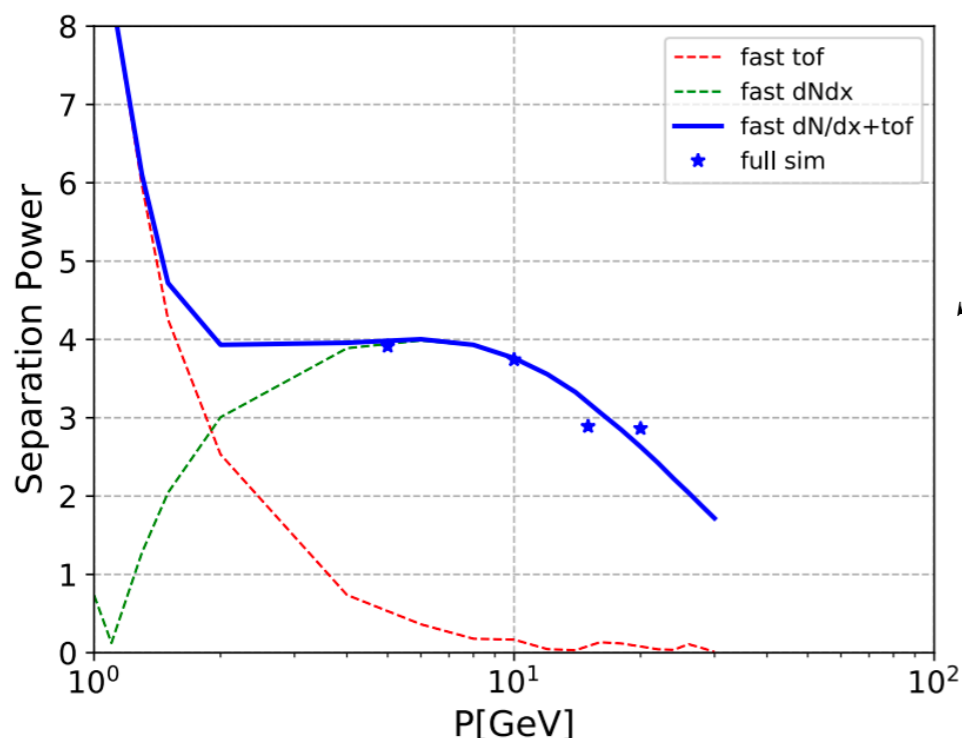
► 很好的柏松统计特征



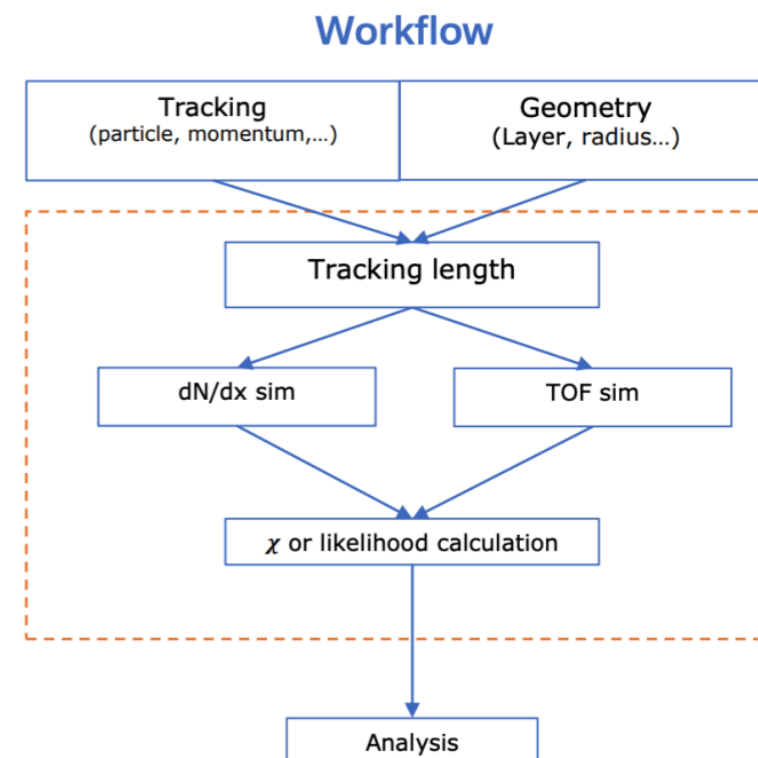
# K/π 粒子鉴别预期性能

## 快模拟流程

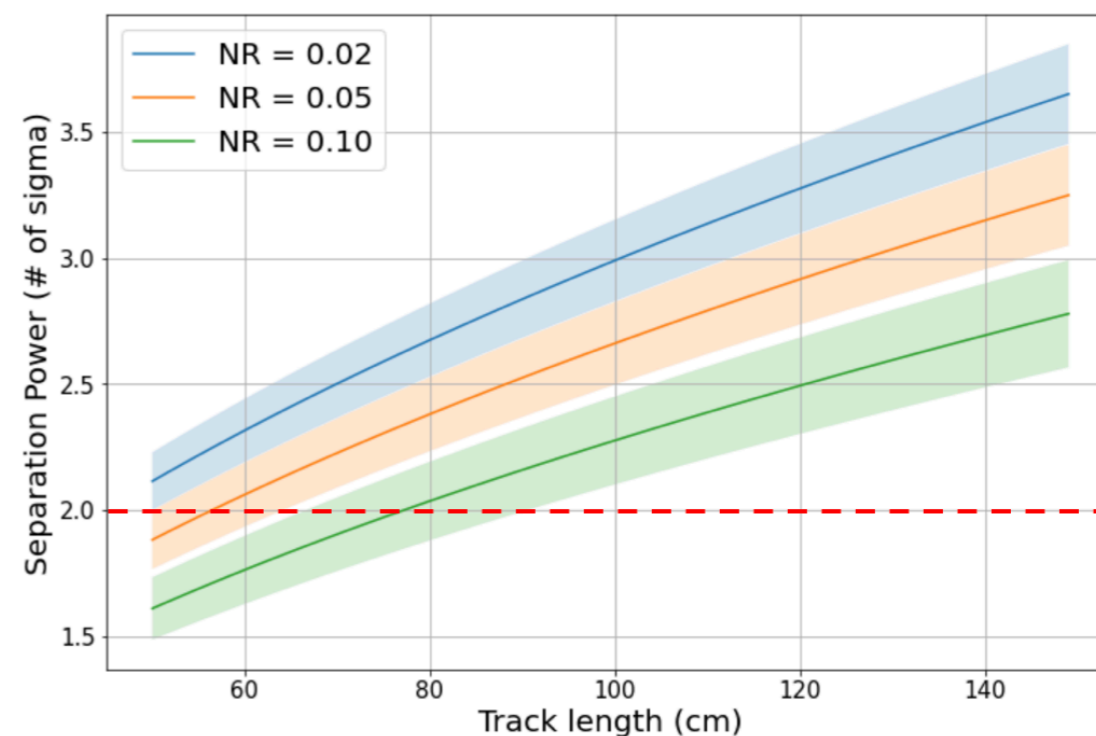
- 从tracking信息中联合dN/dx\_truth × ε 和飞行时间。



$$S = \frac{(dN/dx)_\pi - (dN/dx)_K}{(\sigma_\pi + \sigma_K)/2}$$

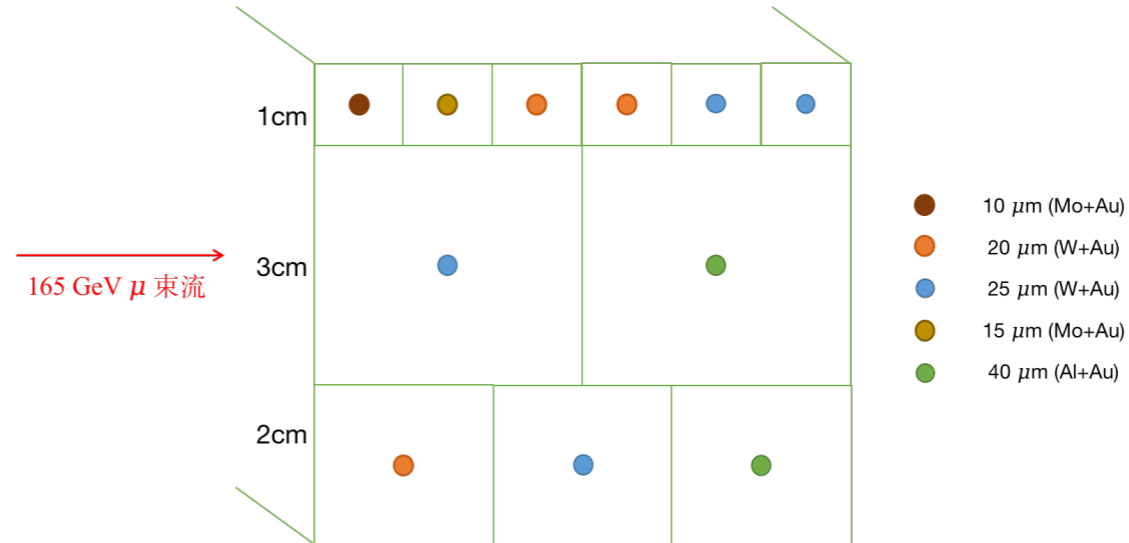
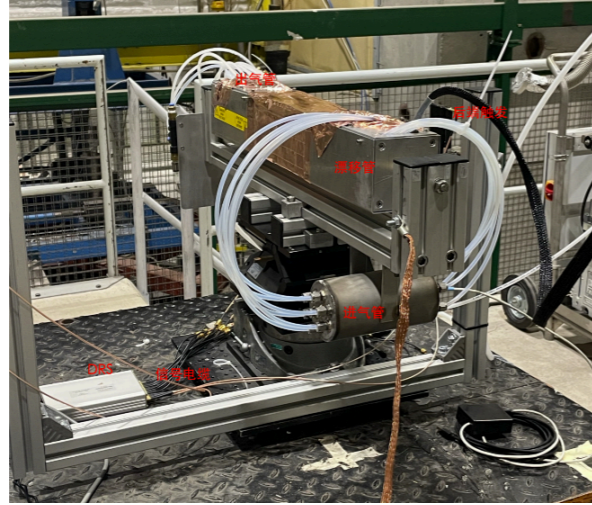


- 较小的径向厚度降低其他探测器（例如 SiTrk、量能器）的成本。
- 在1米的径迹长度内，可实现20 GeV/c下2σ K/π分离。

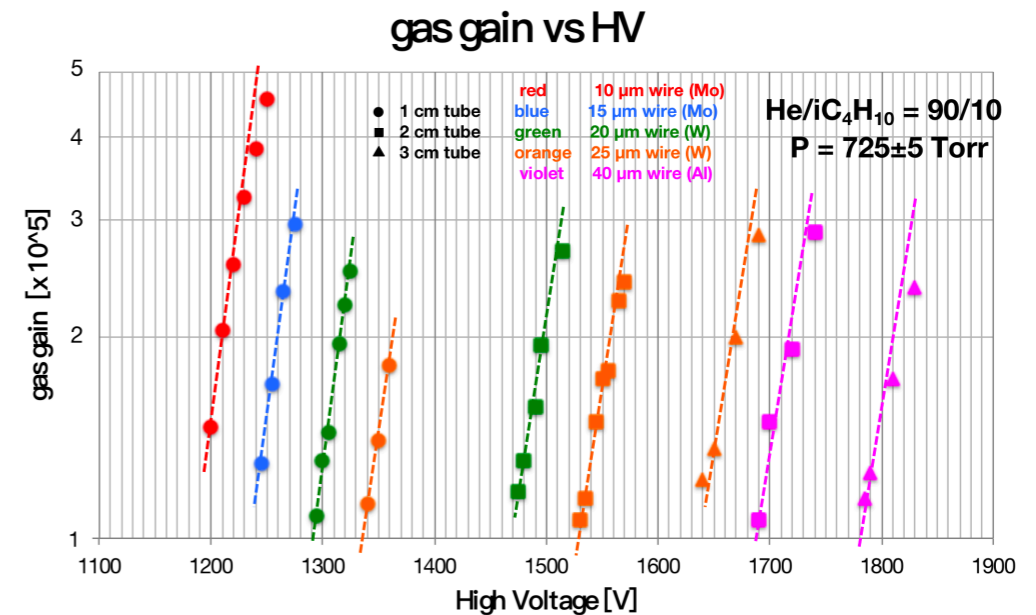


# 束流测试

- ▶ 硬件实验是检验原初电离测量方法有效性的唯一方法。
- ▶ 束流实验设置。 CERN SPS 束线, 165 GeV  $\mu$  束流, 11组漂移管。



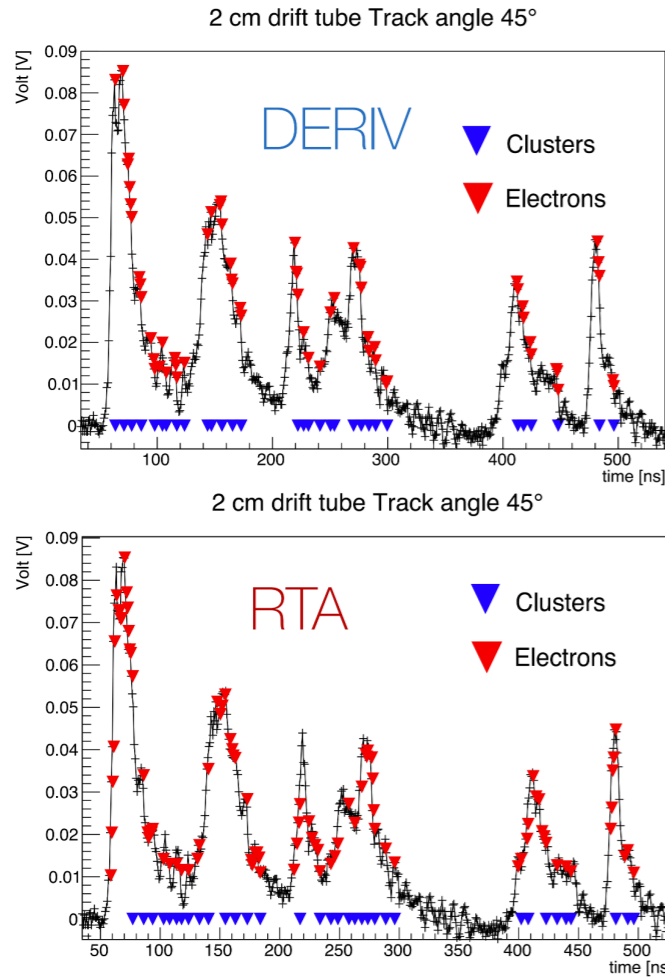
- ▶ 气体电压设置: 初值由Garfield 模拟提供, 测量最大信号幅度以校正至 $O(10^5)$ 。50V差别。
- ▶ 取得数据: Nov. 2021. 1.2 GSa/s, He:iC<sub>4</sub>H<sub>10</sub> (90/10, 80,20), 15°,30°,45°,60° ~ 100 k 事例; July 2022: 1.2, 1.5. 2.0 GSa/s, He:iC<sub>4</sub>H<sub>10</sub> (90/10, 85/15, 80/20) 15°,30°,45°,60° ~ 700k 事例





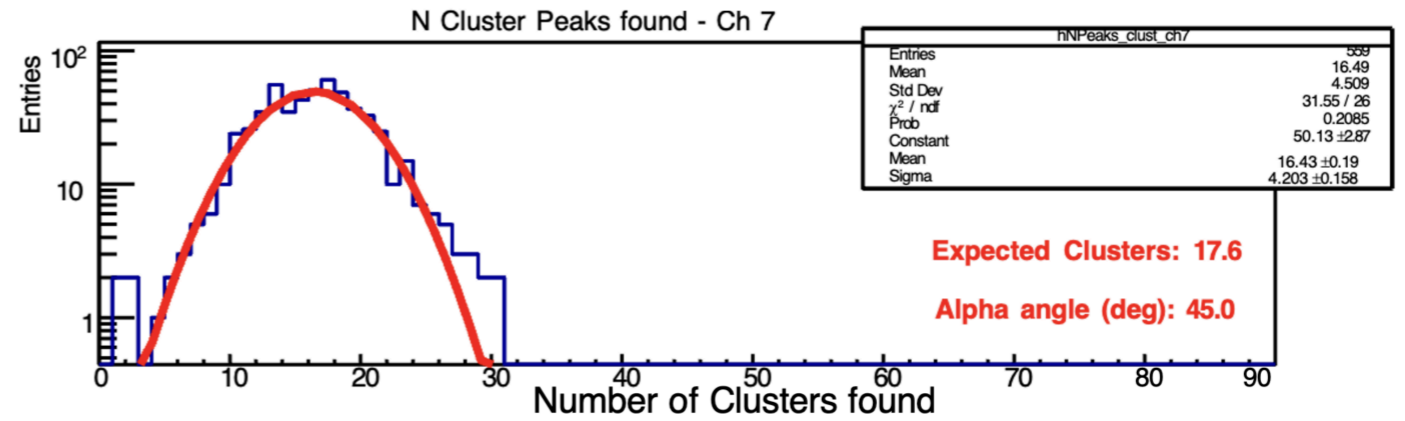
# 束流测试

## ► INFN DERIV/RTA 算法及结果



合并算法:

1. 连续的1个bin被合并成一个电子
2. 与2.5ns相符的相邻电子峰被认为是同一cluster
3. 最后一个电子被标记为cluster 位置

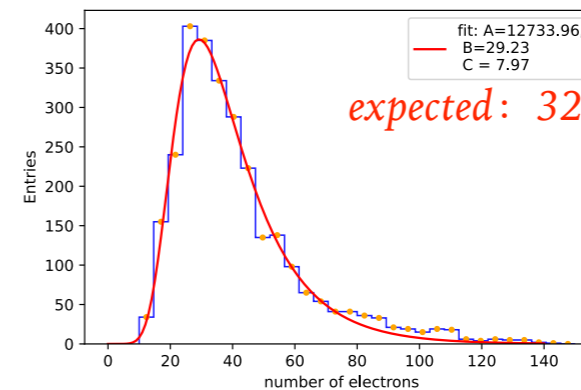


趋于泊松分布

$$N_{cluster(peak)} = (1.6) \times 1.3 \times \rho [cluster / cm] \times L [cm] \times \frac{1}{\cos \theta}$$

## ► 与CEPC 导数算法的比较:

- 不采取光滑 (moving average)
- 要求一阶导数





- 
- ▶ 研究背景
    - ▶ 标准模型和di-Higgs物理
    - ▶ ATLAS 探测器
  - ▶ ATLAS 上di-Higgs 研究
    - ▶ 通过同电荷双轻子寻找标准模型非共振双希格斯粒子
    - ▶ 多轻子联合测量
  - ▶ 漂移室的粒子鉴别算法研究
    - ▶ 原初电离计数（Cluster counting）研究背景
    - ▶ 基于BESIII 漂移室的粒子鉴别研究
    - ▶ Cluster counting 技术用于CEPC 漂移室
    - ▶ 束流测试
  - ▶ 总结与展望
  - ▶ 附文
    - ▶ 已取得研究成果
    - ▶ 主要贡献
    - ▶ 论文审阅建议

# 总结与展望

---

- ▶ 利用2015 -2018 年LHC Run-2 运行阶段ATLAS 收集的 $139\text{fb}^{-1}$ 的pp 对撞数据，在多轻子末态寻找标准模型希格斯粒子对。
  - ▶ 同电荷双轻子道
    - ▶ 利用信号区外数据测量了截面上限预期值，首次测量希格斯自耦合常数范围。
    - ▶ 相比WWWW  $36.1\text{fb}^{-1}$ 结果，截面上限提升了4.3倍。提升来自统计量、多变量分析方法、本底估计方法、系统误差控制。
  - ▶ 多轻子联合测量：合并9个物理道的截面上限初步结果：8.91 倍  $\sigma_{\text{HH}}^{\text{SM}}$ ，首次结果，对比CMS有竞争力。
  - ▶ 轻子物理道分析正在揭盲审议阶段。
- ▶ 展望
  - ▶ 利用多分类的神经网络提升事例分类能力：避免不同物理道训练的复杂度（大模型+大数据量），同时可加入额外物理道（CMS 实验中测量过）的贡献， $3\ell 1\tau, 1\ell 3\tau$  等。
  - ▶ 分析方法上进一步优化轻子重建方法、压低假轻子本底估计误差。
  - ▶ 在WW/ZZ轻子衰变过程寻找类希格斯标量X/S， 如2HDM+S模型。

# 总结与展望

---

## 漂移室的粒子鉴别算法研究

Cluster counting 方法利用电离过程中的柏松性质，用于在未来探测器的设计中克服dE/dx测量在PID能力上的限制。

- ▶ 基于漂移室波形分析，利用BOSS框架的模拟事例，在BESIII上研究了K/pi 分辨的预期性能和粒子鉴别效率。
  - ▶ 理想情况下，cluster counting相比电荷积分测量的分辨率好2倍，K/pi 分辨能力增强2.5倍。
  - ▶  $P_t > 2 \text{ GeV}/c$  动量区间，PID效率在理想情况（分辨60%下降）对dE/dx的提升 22%（10%）
- ▶ 展望
  - ▶ 气体选择：提高氦气比分，降低cluster 密度
  - ▶ 研究低噪声、高性能的前端电子学装置（快响应的读出

# 总结与展望

---

- ▶ 在CEPC上全面研究cluster counting技术的PID能力。
  - ▶ 针对PID技术优化的需求，我们开展了波形的模拟，数字化，重建算法。
  - ▶ 传统重建方法满足在 $P_t > 20 \text{ GeV}/c$  条件， $K/\pi$  的分辨达到  $3\sigma$ ，漂移室的尺寸可缩短至1m。
  - ▶ 利用LSTM、CNN 机器学习模型对cluster做了初步研究，其性能好于传统重建算法。该工作未来准备文章撰写。
  - ▶ 展望：嵌入物理道（如B-介子衰变过程）的物理分析。
- ▶ 与 INFN 等团队合作，在 CERN 进行了两次束流测试实验。利用第一次实验数据，1原初电离数的柏松性质被初步验证。
  - ▶ 展望：利用大样本数据，系统性的研究数据中的疑点：空间电荷效应、吸附效应
  - ▶ 优化模拟样本和机器学习算法，应用于实验数据的研究。

# 已取得研究成果

---

## ► 已发表SCI文章

- (一作) Simulation study of particle identification using cluster counting technique for the BESIII drift chamber, *JINST* 18 (2023) 01, T01006
- Particle identification with the cluster counting technique for the IDEA drift chamber. *Nucl.Instrum.Meth.A* 1048 (2023) 167969
- Leakage current simulations of Low Gain Avalanche Diode with improved Radiation Damage Modeling, *Nucl.Instrum.Meth.A* 1040 (2022) 167111
- Performance in beam tests of carbon-enriched irradiated Low Gain Avalanche Detectors for the ATLAS High Granularity Timing Detector *JINST* 18 (2023) 05, P05005
- The anomalous production of multi-lepton and its impact on the measurement of Wh production at the LHC *Eur.Phys.J.C* 81 (2021) 4, 365

# 主要贡献

---

- ▶ ATLAS diHiggs multilepton
  - ▶ 在CLHCP会议上做海报。
  - ▶ 分析主要贡献者，负责2lss道的物理分析工作，多轻子联合统计测量。
  - ▶ 统计联系人，内部文档撰写人。
  - ▶ 代表分析组做三次内部审核报告。
- ▶ ATLAS HGTD
  - ▶ Production database 最早的开发者，发布第一版数据库，内部文档编辑，取得ATLAS实验组Qualification资格。
  - ▶ TGAD辐射损伤泄漏电流模拟和碳辐照LGAD束流测试工作，测试人员。测量了IHEP-IME LGAD芯片的电荷收集能力和漏电流特性。性能结果。



# 主要贡献

---

- ▶ 漂移室电离计数方法应用于粒子鉴别研究
  - ▶ 在BESIII上利用模拟方法研究cluster counting技术的可行性。主要研究人员，发表了一篇SCI文章。
  - ▶ CEPC上的cluster counting研究。负责快速模拟软件、粒子鉴别算法预期性能研究。在APS会议上做口头报告，ICHEP会议上做口头报告，CEPC会议上做口头报告。
  - ▶ 参加IDEA漂移管模型束流测试。进行了电压计算、装置搭建、离线数据分析工作。作为合作者发表了一篇SCI文章。

# 论文审阅建议

---

- ▶ 论文对主要成果的描述中，关于95%置信度下预期上限期望值数值表述为“mean+/-sigma”，由于这一期望值不是常用的物理测量量，表述方式也不是常规读者可以理解的，因此在摘要中表述时应该在上下标数值前使用“+、-”符号或对上下标做一定的说明。
  - ▶ 修改情况：在论文的摘要和结论中，表述预期上限为其中心值（Medium），不包括1倍sigma，符合常见表述方式。
- ▶ 第一段最后一句“对希格斯粒子全方位的测量是标准模型物理和新物理寻找的重要方向之一”，很令人莫名。
  - ▶ 修改情况：对希格斯粒子进一步测量是新物理寻找的重要方向之一。
- ▶ 第二段“电荷误判本底和假轻子本底主要的非标准模型过程”，不严谨。
  - ▶ 修改情况：电荷误判本底和假轻子本底是主要的不可约本底过程。
- ▶ 论文中希格斯粒子还是Higgs应当统一。
  - ▶ 修改情况：统一为Higgs。

# 论文审阅建议

---

- ▶ cluster counting这种关键且常规名词，应当翻译。
  - ▶ 在摘要中补充“原初电离计数”。
- ▶ 建议将摘要最后一段提前。
- ▶ 结构上讲，第六章占比太小，建议与第7章合并。3.3节，HGTD的工作分量小且与论文主体不协调，建议放在附录中。
  - ▶ 修改情况：第六章是第七章的研究背景，保持不变。HGTD部分已按照意见修改。
- ▶ 第八页第一段有错别字。
  - ▶ 修改情况：去掉多余文字。
- ▶ 公式2.21符号A没有定义。
  - ▶ 修改情况：补充了振幅A的定义。
- ▶ 公式3.21应用数学体，否则  $\ln \tan$  不好识别。
  - ▶ 修改情况：已经是数学字符
- ▶ 第六章部分图标“internal”，部分标“work in progress”。请查阅ATLAS合作组关于博士论文的规定。
  - ▶ 修改情况：统一为work in process。
- ▶ 图7.1，注明是模拟还是实验数据。
  - ▶ 修改情况：注明是模拟数据。

图7.6及以后，Separation Power没有定义。是S吗？

修改情况：在定义S时补充说明了分离度（Separation Power）。

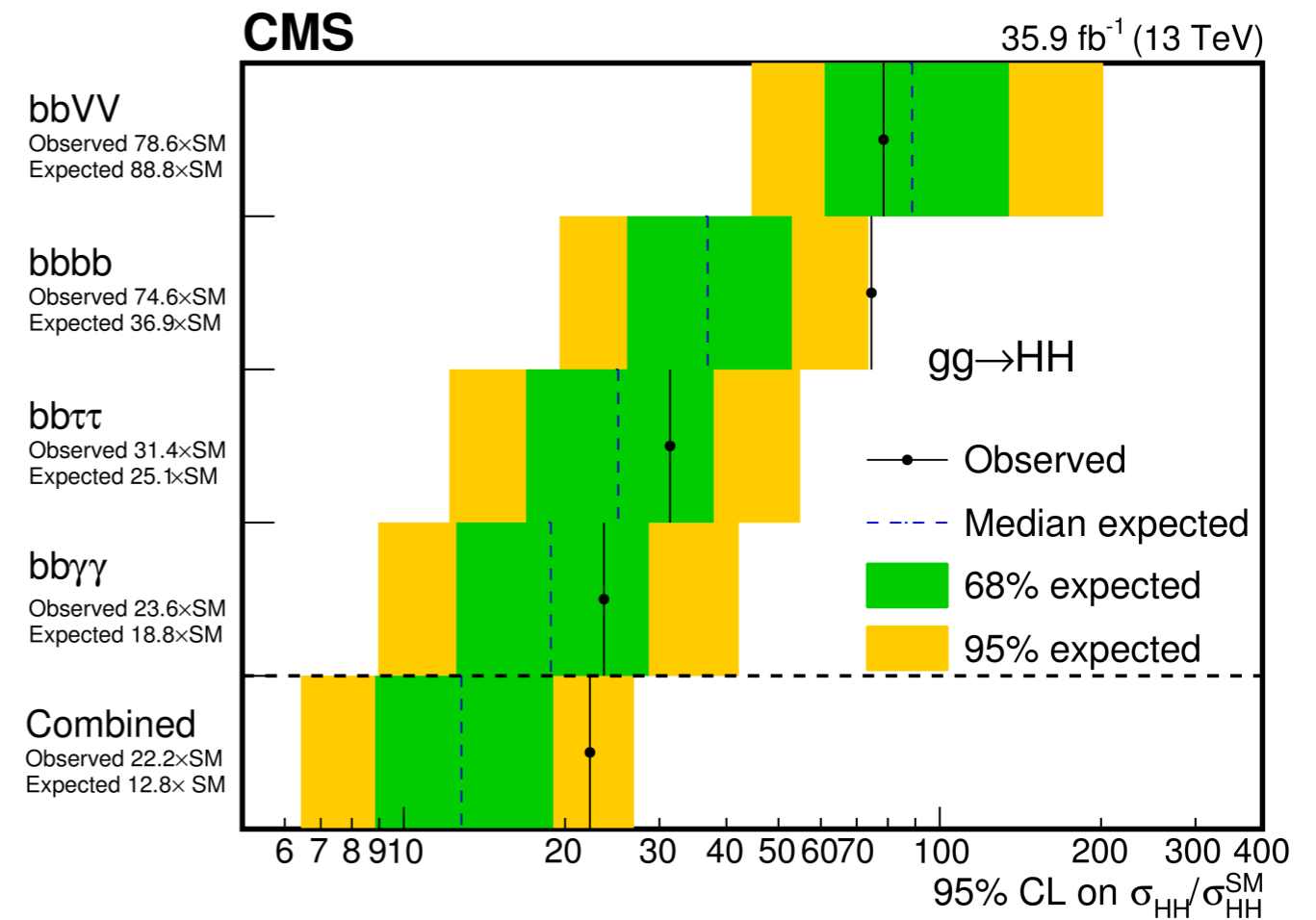
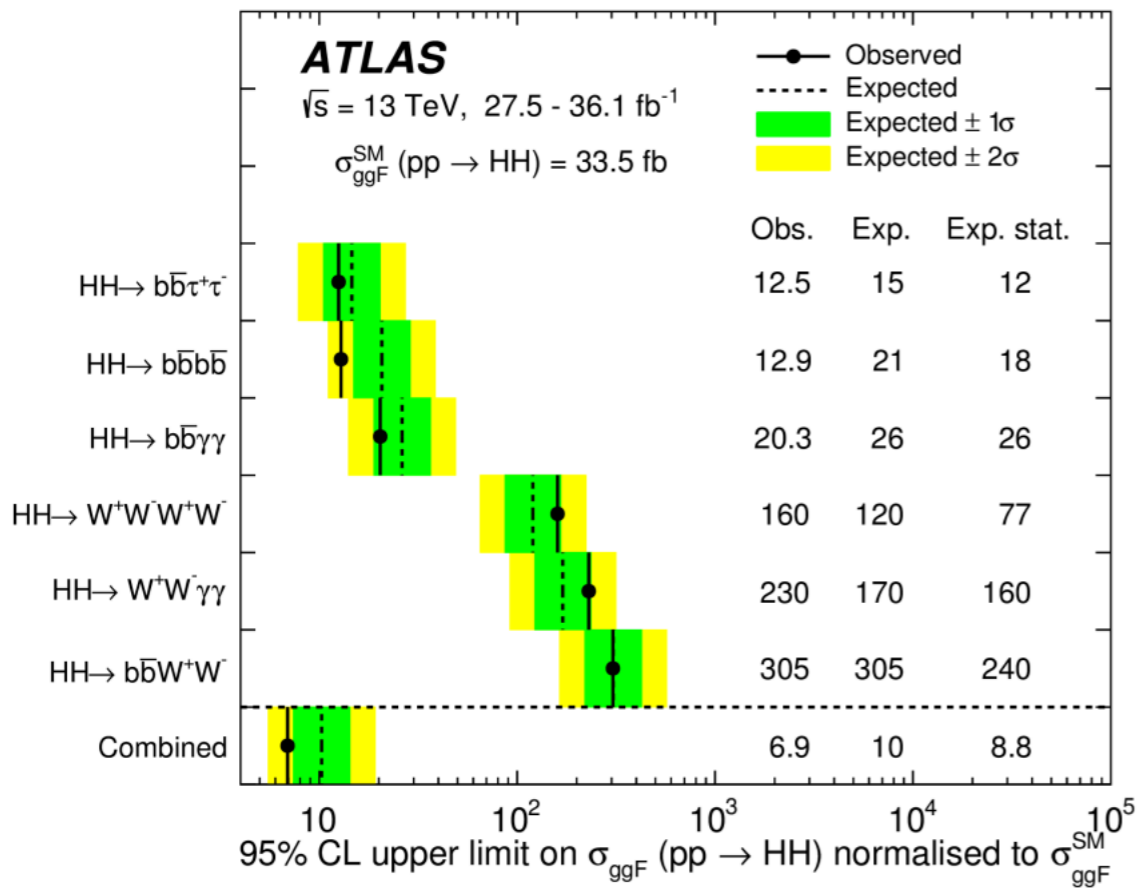
# 论文审阅建议

---

- ▶ 摘要中第二段第二行13 TeV处表达不对，第二段倒数第四行，正负误差没有+-符号
  - ▶ 修改情况：已修改。
- ▶ 图2.1，“图修改自[9]”，“Modified from [9]”表达不妥。同样图2.2
  - ▶ 修改情况：“图修改自网络资源[9]”。
- ▶ Sec4.2.2.1: “RUN 2 期间，自从 2017 年开始，ATLAS 采用拓扑簇射团(topo-cluster)算法”，表达不准确，topo-cluster不是从RUN2才开始的，而是从ATLAS一开始就在用
  - ▶ 修改情况：RUN2是取代了固定窗算法。修改为“RUN 2 期间，自从 2017 年开始，ATLAS 在原有拓扑簇射团（topo-cluster）算法基础上，用可变化的簇团大小对韧致辐射光子的重建，替代了原有的固定窗算法。”
- ▶ 图5.12至5.17，“work in progress”覆盖了“data”
  - ▶ 修改情况：已修改。
- ▶ 图7.29，7.30，正文没有引用，也没有解释图代表的结果
  - ▶ 修改情况：已在相应位置增加引用。
- ▶ P157,倒数第六行，“在[170]”表达不妥
  - ▶ 修改情况：修改成这一工作发表在SCI期刊[170]。

# HH 联合测量

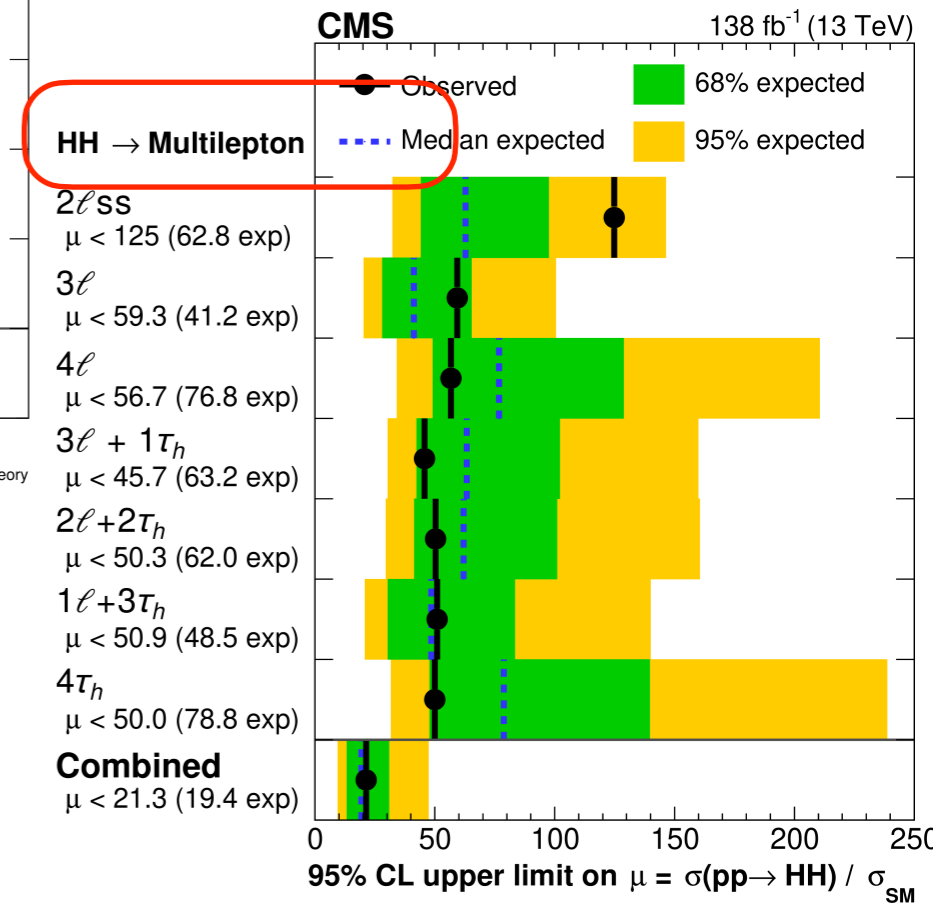
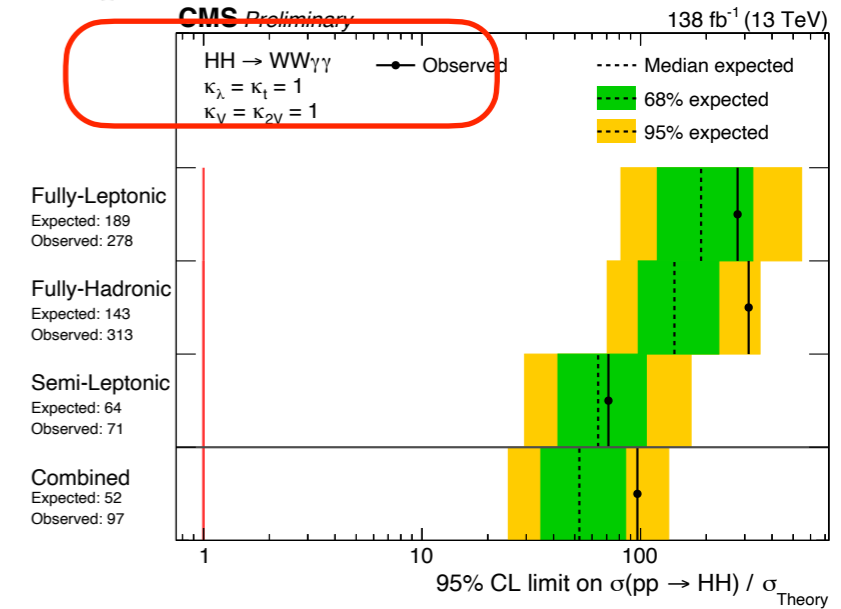
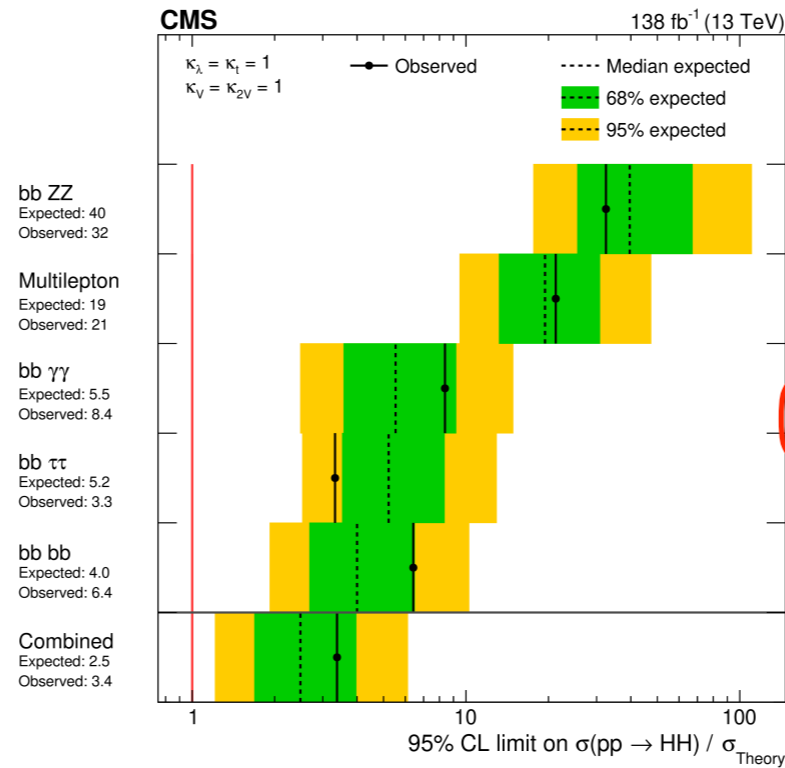
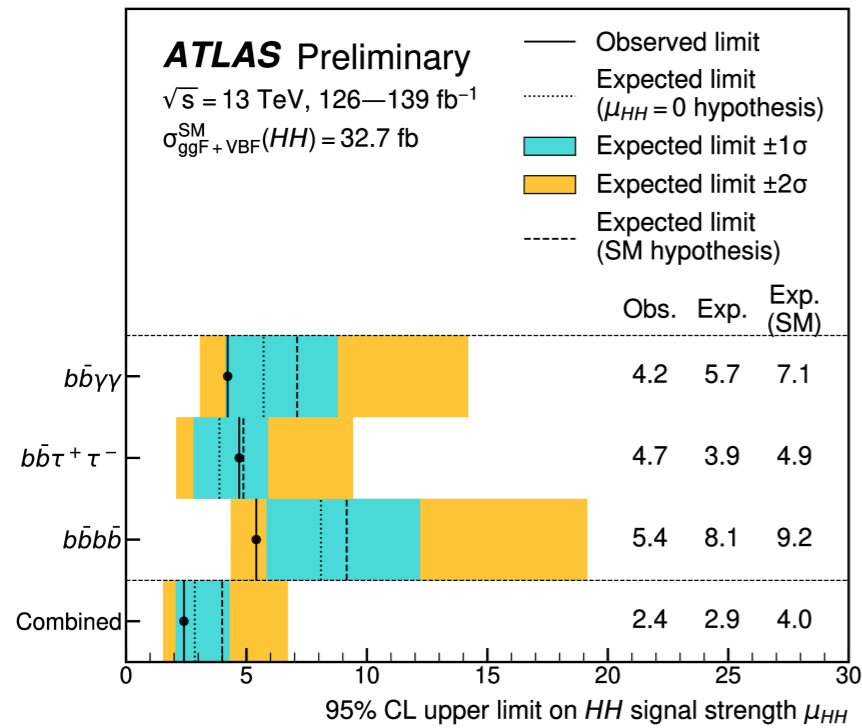
- ▶ ATLAS RUN2 36.1 fb<sup>-1</sup>结果：联合6个道，结果为6.9（10）倍截面标准模型上限。
- ▶ CMS RUN2 35.9 fb<sup>-1</sup>结果：结果为22.8（12.8）倍



# HH 联合测量

- ATLAS 联合最敏感的3个道: 2.4 (2.9)
- CMS 联合5个道: 2.5 (3.4)
- CMS WW $\gamma\gamma$ : 97 (52)

<https://cds.cern.ch/record/2840773>





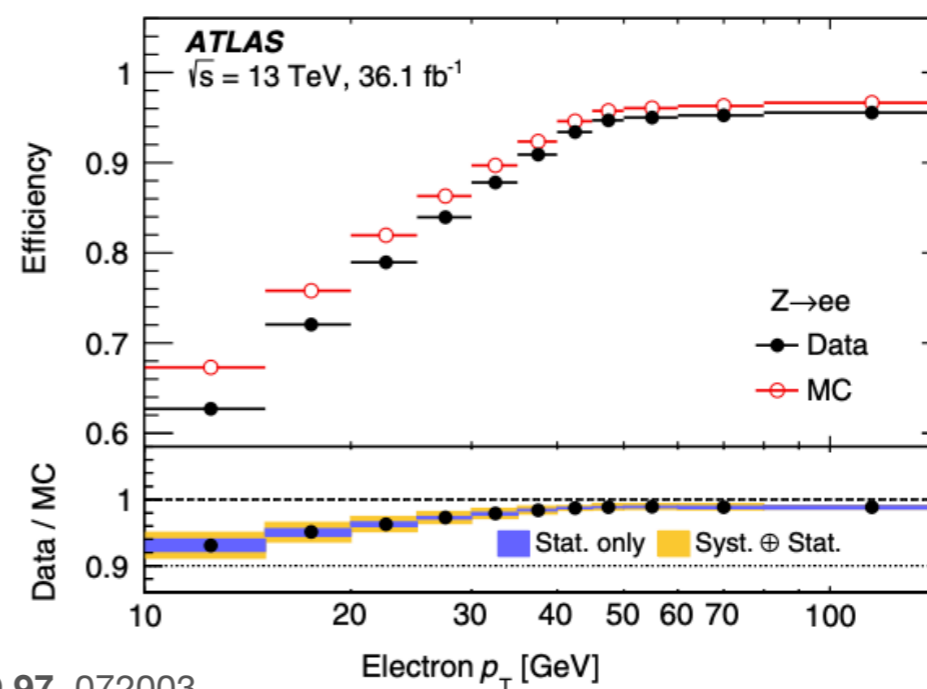
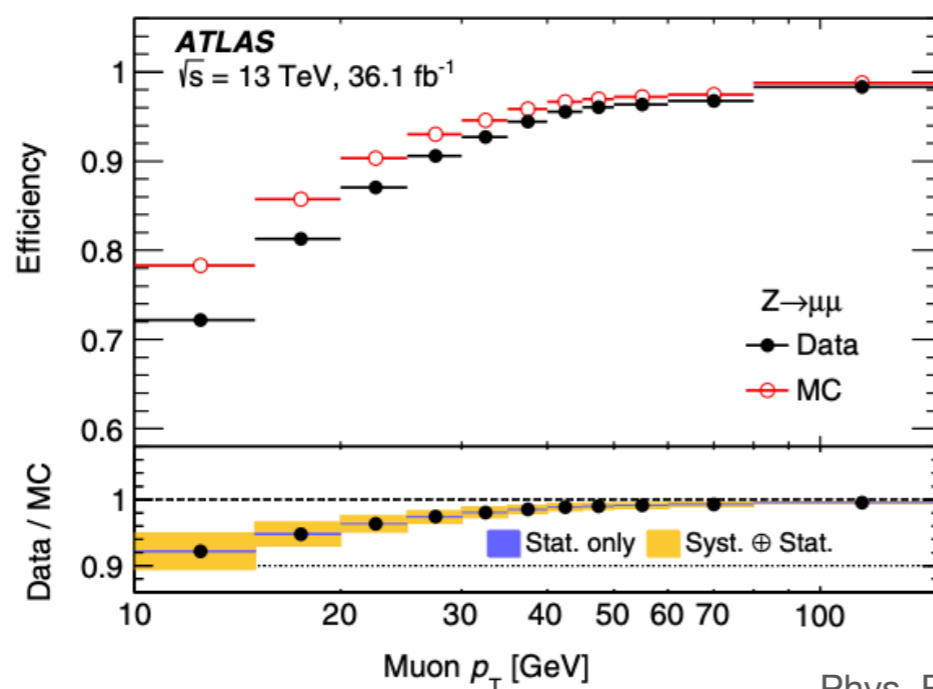
# Trigger 列表

Single lepton triggers (2015)	
$\mu$	HLT_mu20_iloose_L1MU15, HLT_mu50
$e$	HLT_e24_lhmedium_L1EM20VH, HLT_e60_lhmedium, HLT_e120_lhloose
Dilepton triggers (2015)	
$\mu\mu$ (asymm.)	HLT_mu18_mu8noL1
$ee$ (symm.)	HLT_2e12_lhloose_L12EM10VH
$e\mu, \mu e$ (~symm.)	HLT_e17_lhloose_mu14
Single lepton triggers (2016)	
$\mu$	HLT_mu26_ivarmedium, HLT_mu50
$e$	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0, HLT_e140_lhloose_nod0
Dilepton triggers (2016)	
$\mu\mu$ (asymm.)	HLT_mu22_mu8noL1
$ee$ (symm.)	HLT_2e17_lhvloose_nod0
$e\mu, \mu e$ (~symm.)	HLT_e17_lhloose_nod0_mu14
Single lepton triggers (2017 / 2018)	
$\mu$	HLT_mu26_ivarmedium, HLT_mu50
$e$	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0, HLT_e140_lhloose_nod0
Dilepton triggers (2017 / 2018)	
$\mu\mu$ (asymm.)	HLT_mu22_mu8noL1
$ee$ (symm.)	HLT_2e24_lhvloose_nod0
$e\mu, \mu e$ (~symm.)	HLT_e17_lhloose_nod0_mu14

# PLV working point

- ▶ PLV BDT 使用  $t\bar{t}$  样本训练，W 为强子衰变 ( $t\bar{t} \rightarrow bbqqqq$ ) plv score 接近 -1 属于 prompt 轻子。
- ▶ 电子（缪子）在选择的工作点上，PLV 的效率约为 65%（80%）
- ▶ 在  $p_T$  约为 20 GeV，45 GeV 时达到了 90%（95%）的平台。
- ▶ 对于 Fake 电子（或缪子）的拒绝率约为 10（3.5）。

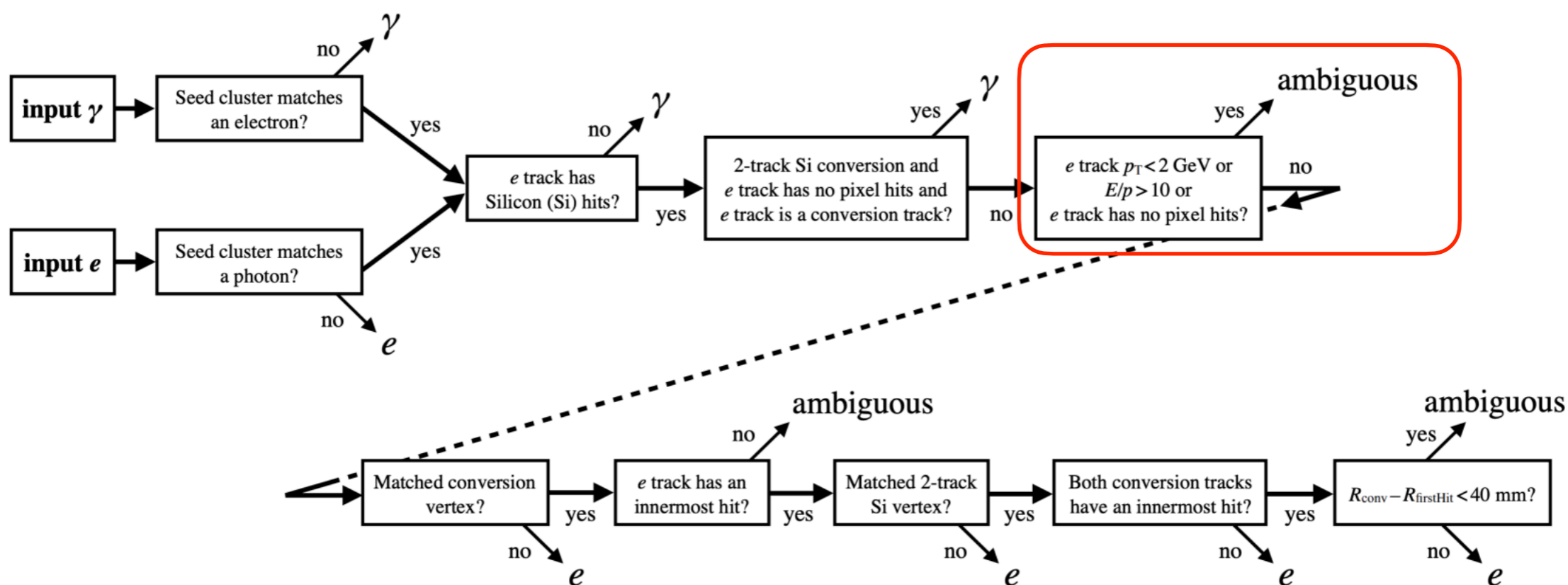
PLV 变量	描述
$N_{track}$ in track jet	Number of tracks collected by the track jet
RNNIP	Recurrent Neural Network with additional impact parameterised information of tracks inside the track-jet
DL1mu	DL1 (deep learning tagger) extended with Soft Muon Tagging information
$p_T^{rel}$	lepton $p_T$ in matched track jet direction, $p \cdot \sin(\langle \text{lepton, track jet} \rangle)$
$p^f$	The ratio of the lepton track $p_T$ and the track jet $p_T$
$\Delta_R(\text{lepton, track jet})$	$\Delta_R$ between the lepton and the track jet axis
$p_T VarCone30/p_T$	Lepton track isolation, with track collecting radius of $\Delta_R < 0.3$
$E_T VarCone30/p_T$	Lepton calorimeter isolation, with topological cluster collecting radius of $\Delta_R < 0.3$



Phys. Rev. D 97, 072003

# E/gamma ambiguity

- ▶ 要求e/gamma 模糊性以进一步提高候选电子的鉴别。
- ▶ 使用轨迹和量能器信息来评估某个电子/光子候选是否来自重建的转换顶点。
- ▶ 采用了最严格的e/ $\gamma$ -ambiguity resolution标准。该电子的 EM 簇射与单一径迹匹配，不与任何重建的光子转换顶点相关。

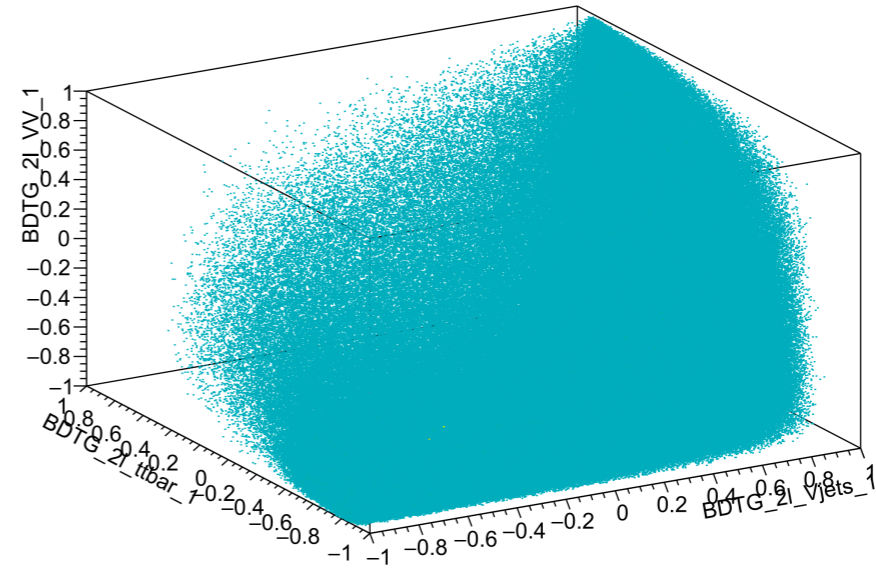


# 多变量分析

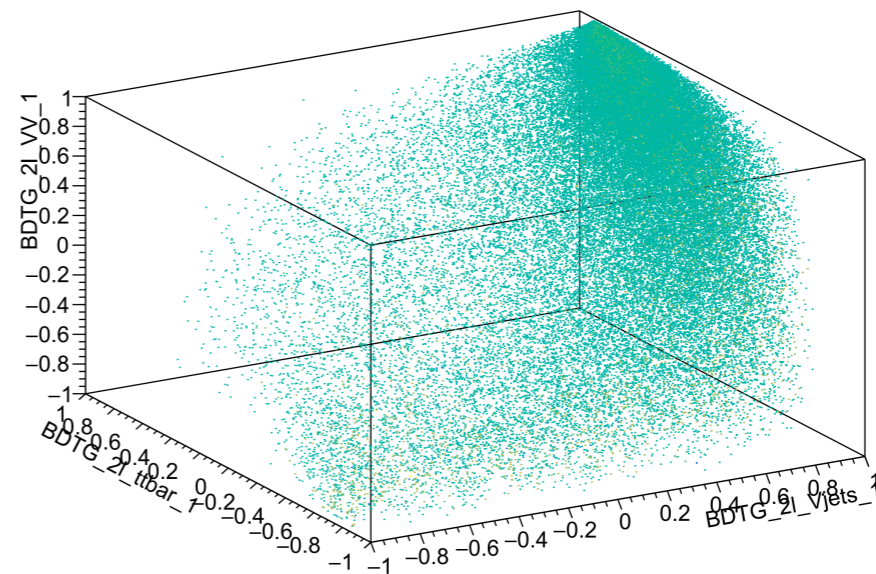
变量	Vjets	ttbar	VV
$M_{l_1j}$	3.5%	7.2%	5.6%
$M_{l_0j}$	3.9%	6.6%	6.3%
$M_{\ell\ell}$	4.2%	5.9%	4.6%
$M_{all}$	/	4.9%	5.9%
$M_{W0}^T$	2.6%	5.1%	4.5%
$M_{W1}^T$	3.2%	4.7%	5.2%
$HT_{lep}$	3.5%	4.8%	4.2%
HT	7.7%	4.8%	5.8%
$N_{jets}$	4.7%	5.4%	6.0%
$E_T^{miss}$	2.7%	4.4%	4.6%
Dilep_type	6.3%	/	4.4%
$\Delta R_{min\ell 1jets}$	10.1%	7.5%	8.2%
$\Delta R_{min\ell 0jets}$	9.1%	6.7%	7.8%
$\Delta R_{min\ell jets}$	11.7%	8.9%	10.7%
$ \eta_0 $	2.6%	3.9%	5.6%
$ \eta_1 $	3.1%	2.9%	4.6%
$\Delta R_{\ell\ell}$	11.7%	7.3%	8.4%
Total_Charge	/	/	4.0%

## 变量的区分能力

VV 样本



signal 样本



# CR 定义

表 5.6 控制区间定义。保证相对于相应信号区域的正交性的要求以蓝色突出显示。

Table 5.6 Definition of control regions. Orthogonality requirements with respect to signal reigon are highlighted in blue.

Channel	Region	$N_{\text{lep}}^{\sum Q_i^{\ell}}$	$N_{\tau_{\text{had}}}^{\sum Q_i^{\ell}}$	$N_{\text{jet}}$	$N_{b\text{-jet}}$	BDT cut	$m_{\ell\ell\ell}$ [GeV]	$m_{\ell^+\ell^-}$ [GeV]	$m_{\ell^\pm\ell^\pm}$ [GeV]	$E_T^{\text{miss}}$ [GeV]	$m_{jj}$ [GeV]
2 $\ell$ SS	WZ CR	3 $^{\pm 1}$	0	$\geq 2$	0	$BDT_{\text{All}} < -0.4$ $BDT_{V\text{jets}} > -0.8$	$ m_{\ell\ell\ell} - m_Z  > 10$	$ m_{\ell\ell} - m_Z  < 10$		$> 30$	
	W $^\pm$ W $^\pm$ jj CR	2 $^{\pm 2}$	0	$\geq 2$	0	$BDT_{\text{All}} < -0.4$ $BDT_{V\text{jets}} > -0.8$			$ m_{\ell\ell} - m_Z  > 10$		$> 300$
	QmisID	$e^\pm e^\pm$ $e^\pm e^\mp$	0	$< 2$	0			[78.5, 102.3]	[76.5, 101.3]		
	Conv CR	2 $^{\pm 2}$	0	$\geq 2$	$\geq 1$						
	QED CR	2 $^{\pm 2}$	0	$\geq 2$	$\geq 1$						
	HF-e CR	$\ell^\pm e^\pm$	0	2,3	1 2						
	HF- $\mu$ CR	$\ell^\pm \mu^\pm$	0	2,3	$\geq 1$						

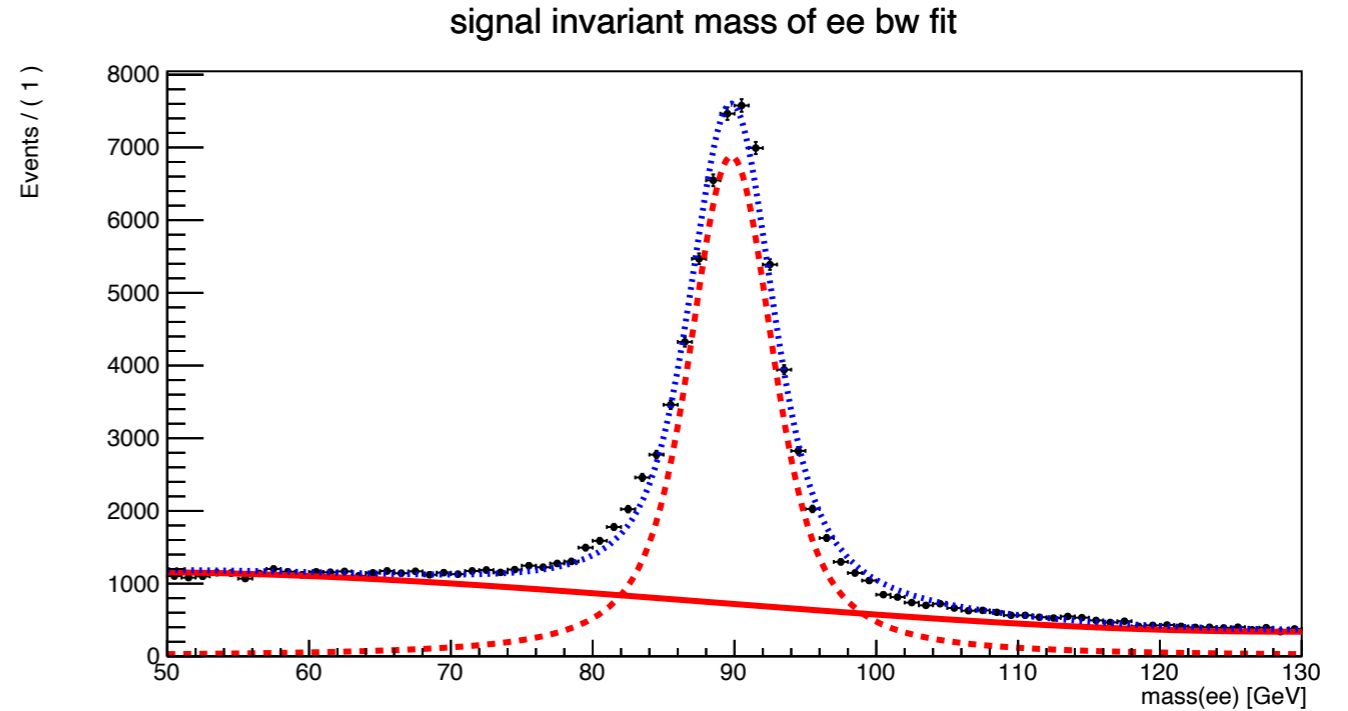


# QMisID 估计

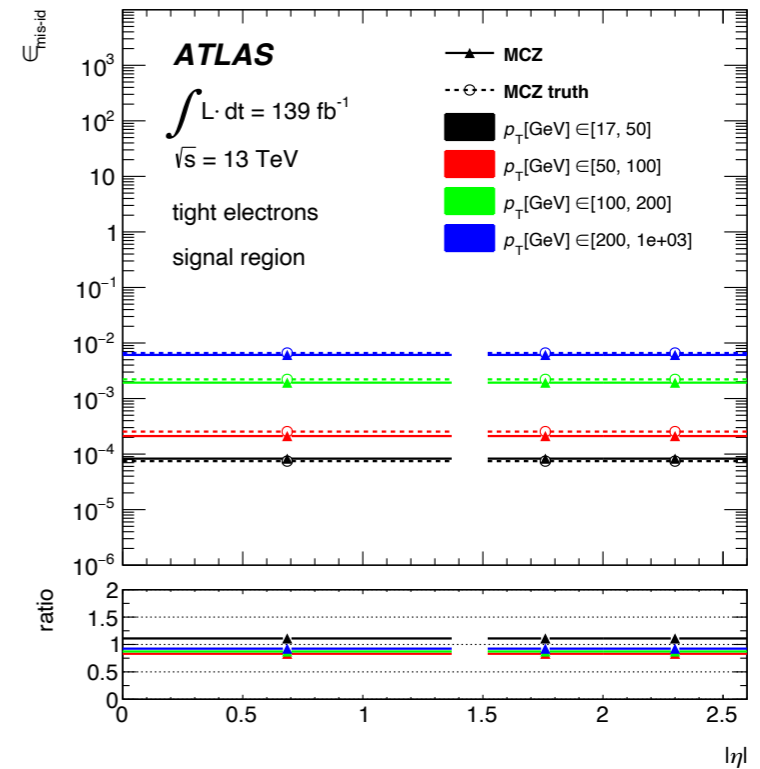
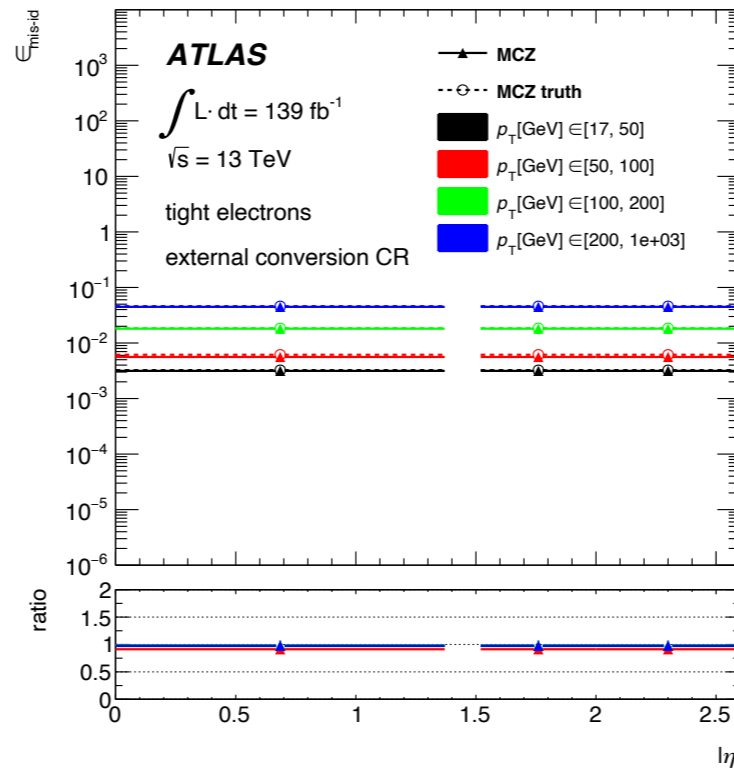
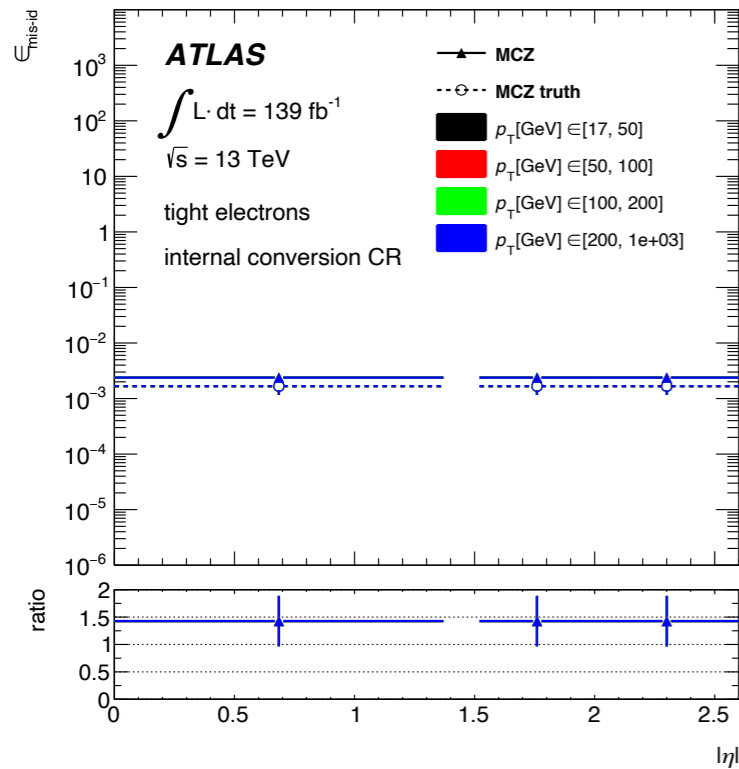
## ► Z-mass 窗口拟合:

- signal: Breit-Wigner 卷积高斯。
- background: 多项式。

$$\bar{N}_{SS} = \frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - (\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j)} N_{OS} \quad \bar{N}_{SS} = \frac{\epsilon_i}{1 - \epsilon_i} N_{OS}$$



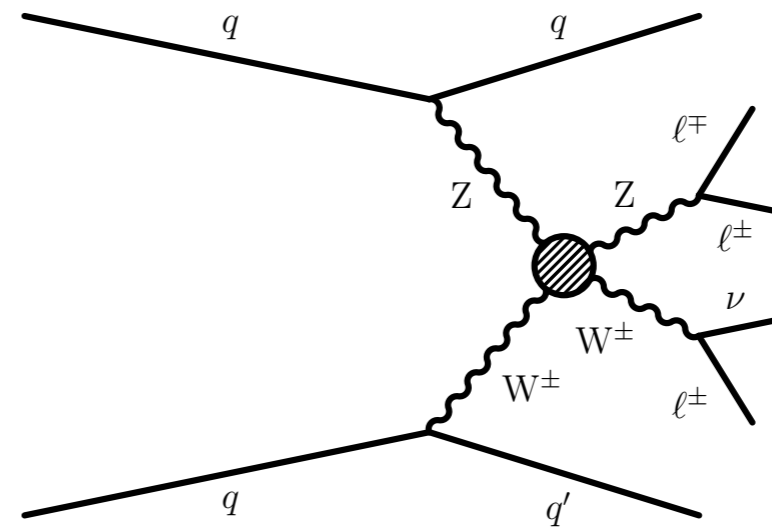
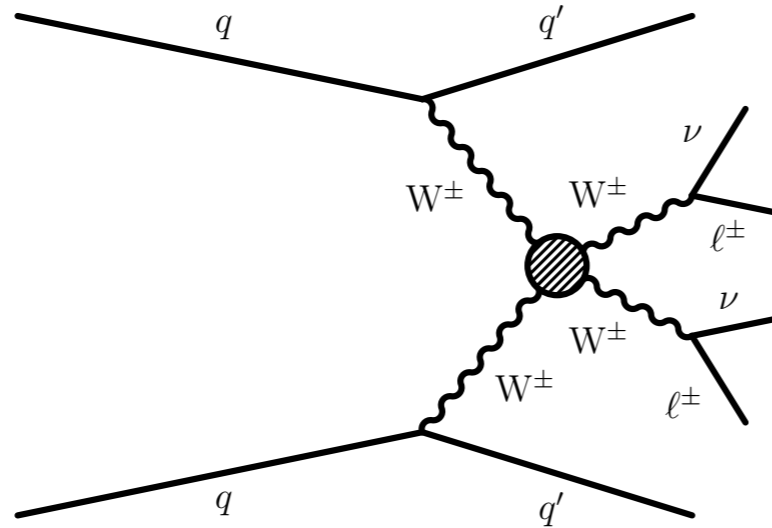
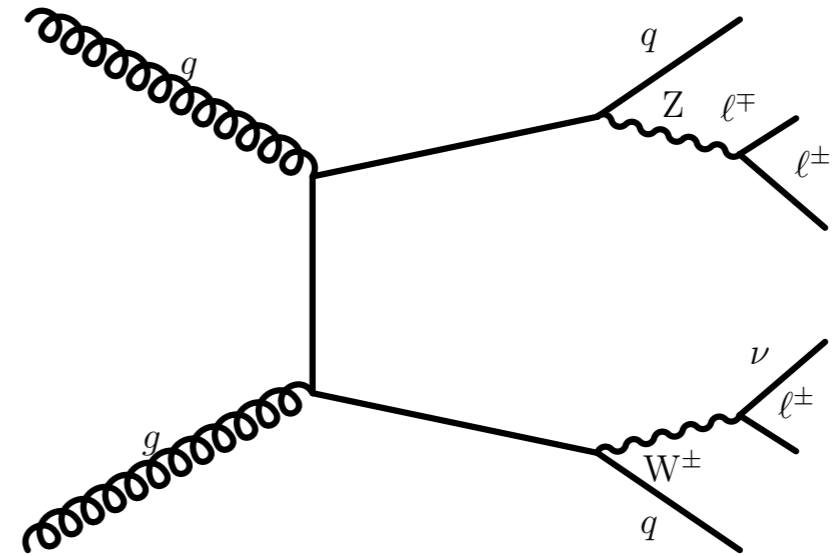
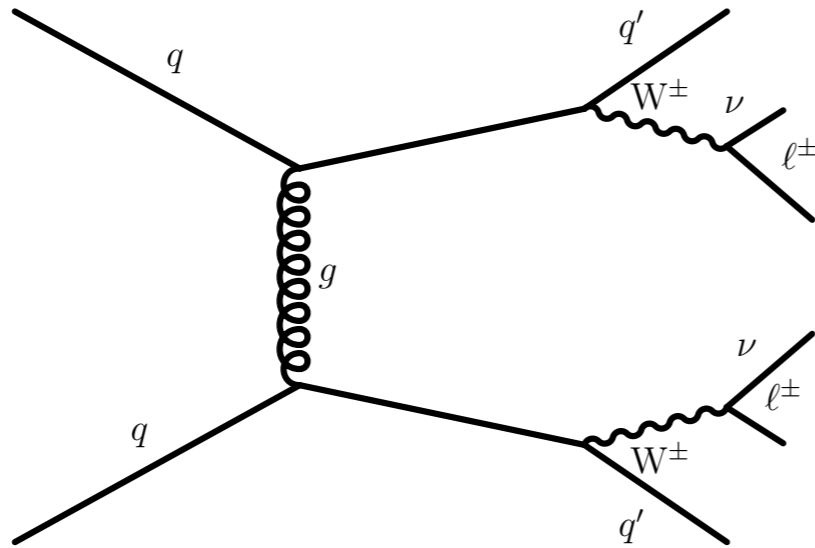
## ► 估计值与MC真实值之间的unclosure。





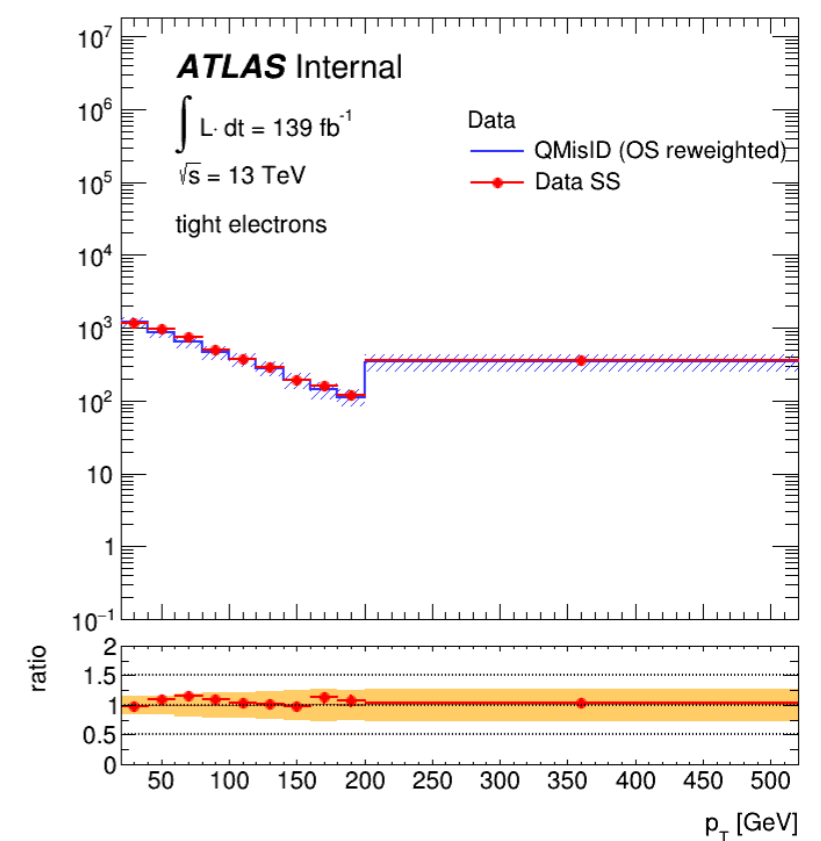
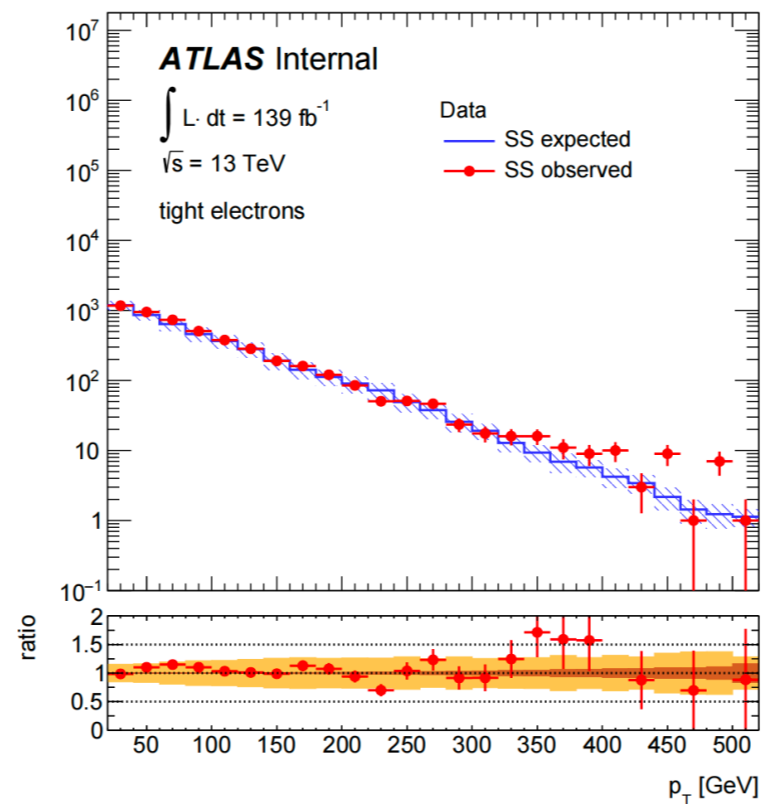
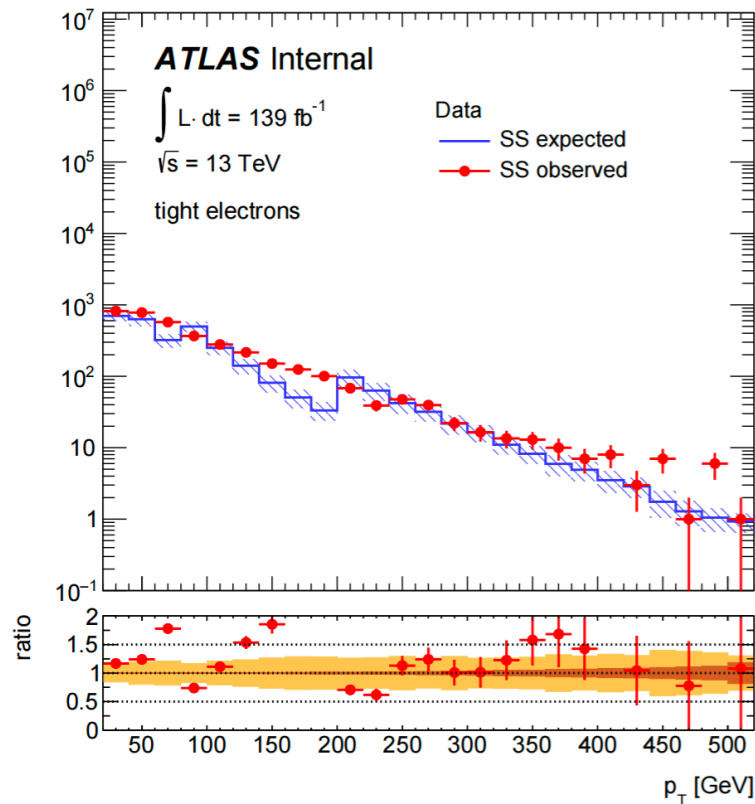
# WZ 过程费曼图

## ► EW (VBS) 和 QCD



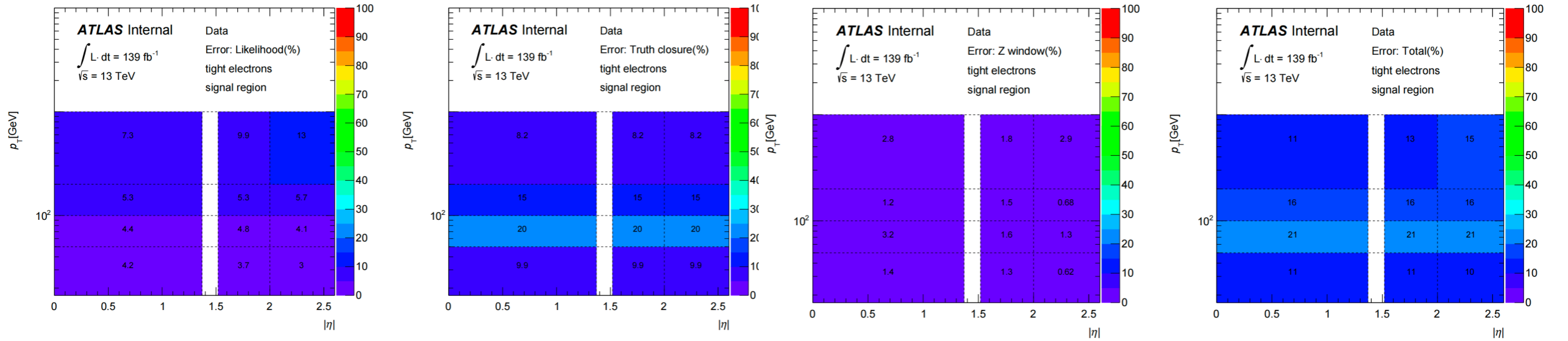
# QMisID 估计

- ▶ 在 $p_T$ 的边缘存在显著的non-closure，这些差别（只）可以被平均的non-closure 误差覆盖。
- ▶ 对 $p_T$ 做内插：给定某个 $p_T$ 值时的rates由相邻 $p_T$ 区间的rates加权求和得到
- ▶ 预测的QMisID（蓝色）和数据（红色）事件的比较。数据SS包括所有背景（QMisID + 假粒子 + 正确重建的事例），而QMisID是从OS加权的事例。

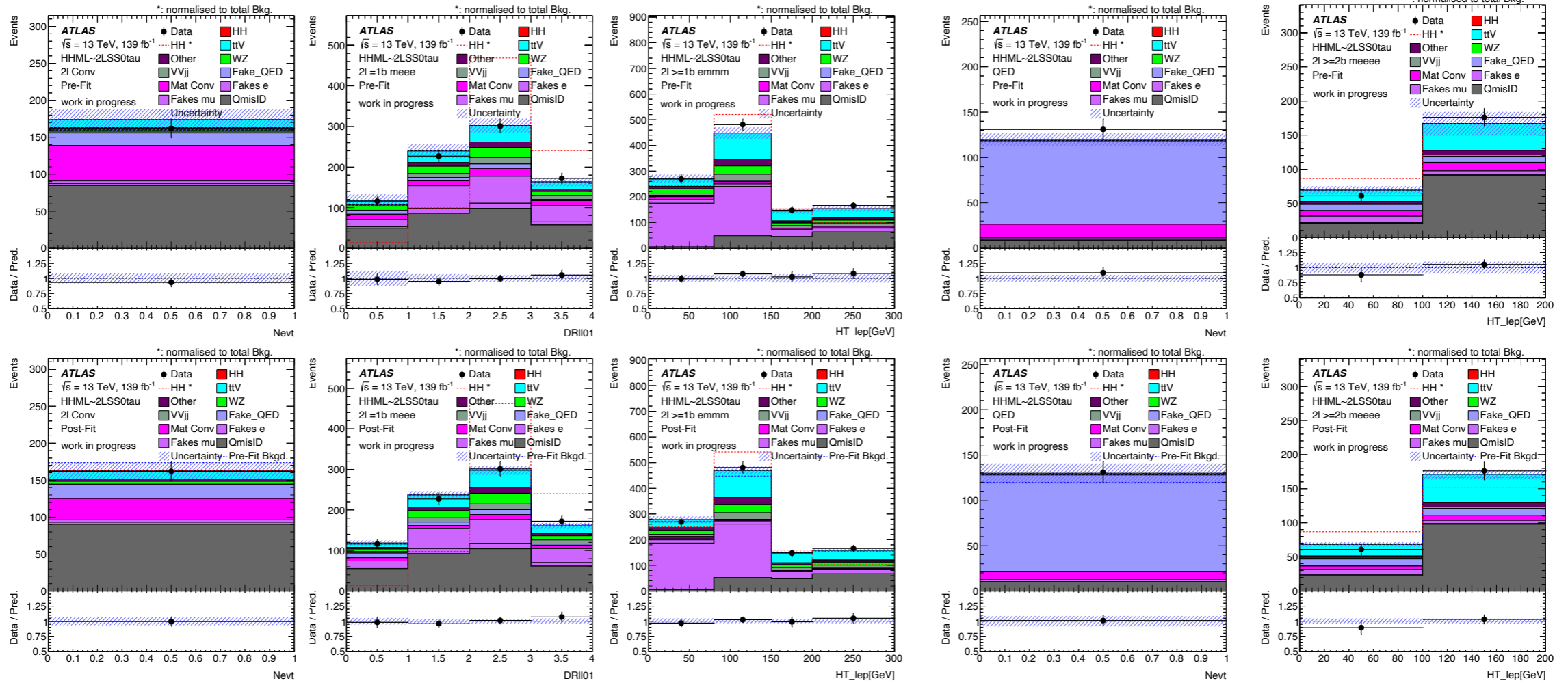


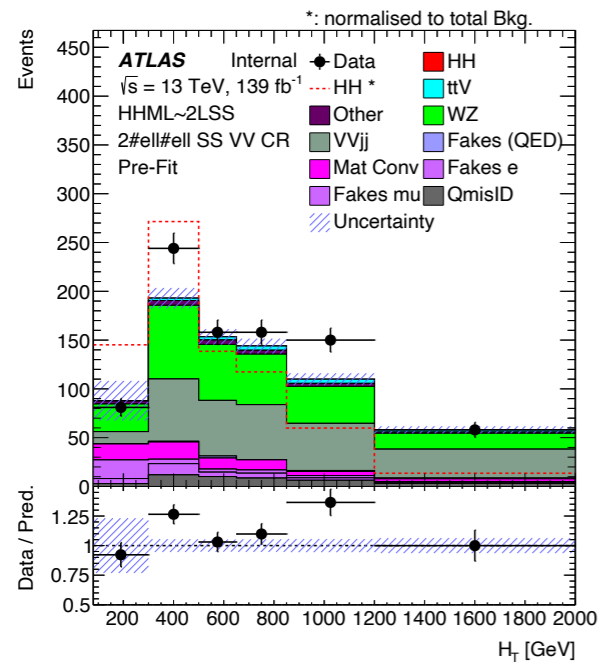
# QMisID 误差

## ► Tight 电子的误差来源



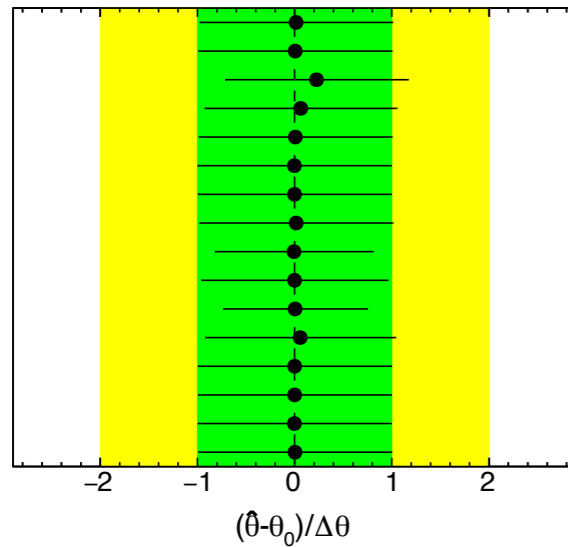
# CR分布





# 系统误差pull

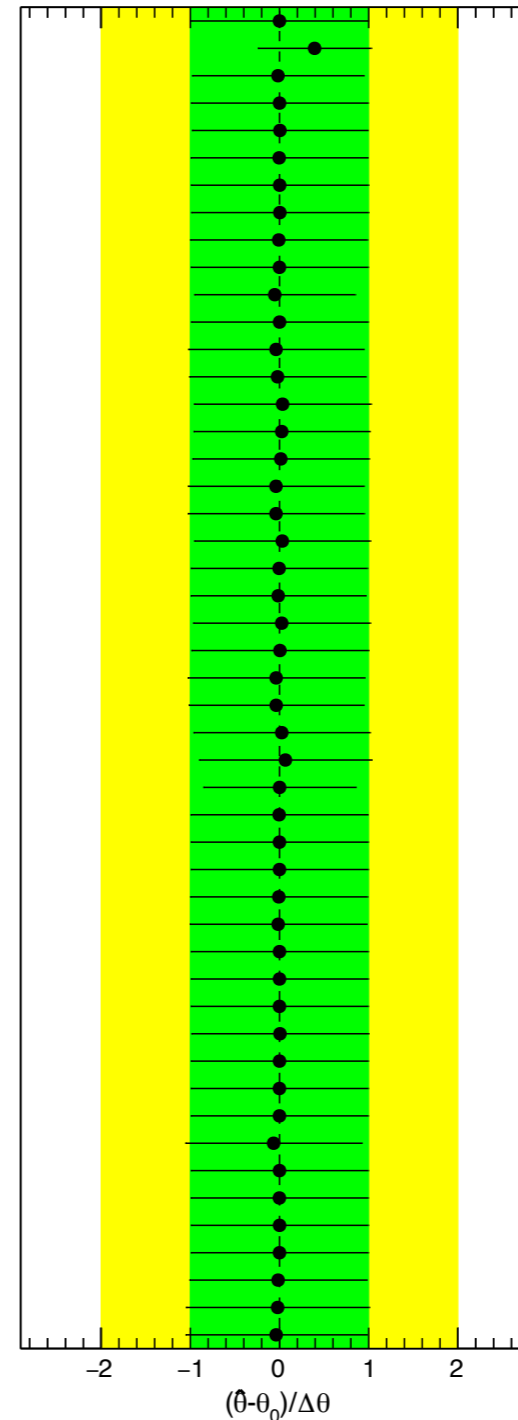
ATLAS Internal



Theory

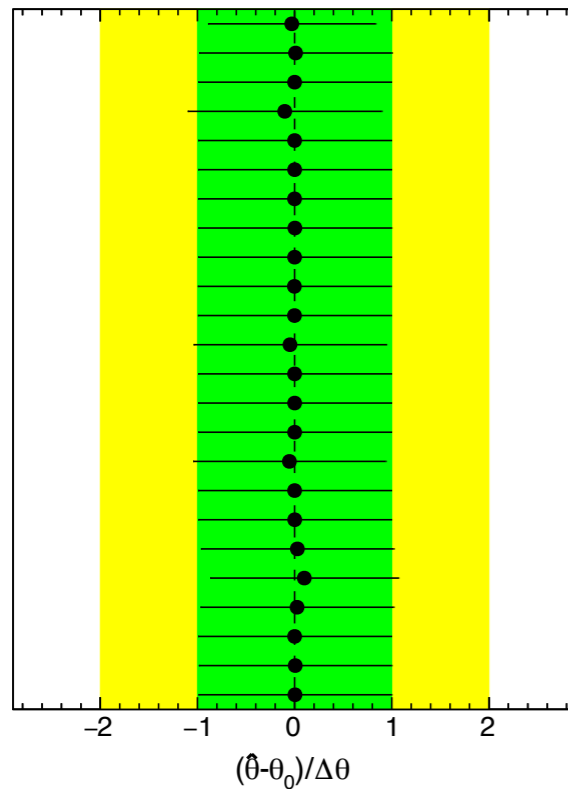
- ttZ x-section QCD
- ttZ x-section PDF
- ttW x-section QCD
- ttW x-section PDF
- ttWW x-section
- ttH x-section QCD
- ttH x-section PDF
- tZ x-section
- hh\_2l\_TF\_HF\_MU
- hh\_2l\_TF\_HF\_E
- hh\_2l\_TF\_HF\_CONV\_E
- WtZ x-section
- VVjj x-section QCD
- VVjj x-section PDF
- VVV x-section
- VH x-section

ATLAS Internal



- Norm\_VV
- QmisID
- ATLAS\_TAUS\_TRUEHADTAU\_SME\_TES\_MODEL\_C
- ATLAS\_MUON\_SCALE
- ATLAS\_MUON\_SAGITTA\_RESBIAS
- ATLAS\_MUON\_CB
- ATLAS\_MET\_SoftTrk\_Scale
- ATLAS\_MET\_SoftTrk\_ResoPerp
- ATLAS\_MET\_SoftTrk\_ResoPara
- ATLAS\_JET\_PunchThrough\_MC16
- ATLAS\_JET\_Pileup\_RhoTopology
- ATLAS\_JET\_Pileup\_PtTerm
- ATLAS\_JET\_Pileup\_OffsetNPV
- ATLAS\_JET\_Pileup\_OffsetMu
- ATLAS\_JET\_JER\_EffectiveNP\_9
- ATLAS\_JET\_JER\_EffectiveNP\_8
- ATLAS\_JET\_JER\_EffectiveNP\_7
- ATLAS\_JET\_JER\_EffectiveNP\_6
- ATLAS\_JET\_JER\_EffectiveNP\_5
- ATLAS\_JET\_JER\_EffectiveNP\_4
- ATLAS\_JET\_JER\_EffectiveNP\_3
- ATLAS\_JET\_JER\_EffectiveNP\_2
- ATLAS\_JET\_JER\_EffectiveNP\_12restTerm
- ATLAS\_JET\_JER\_EffectiveNP\_11
- ATLAS\_JET\_JER\_EffectiveNP\_10
- ATLAS\_JET\_JER\_EffectiveNP\_1
- ATLAS\_JET\_JER\_DataVsMC\_MC16
- ATLAS\_JET\_Flavor\_Response
- ATLAS\_JET\_Flavor\_Composition
- ATLAS\_JET\_EtaIntercalibration\_TotalStat
- ATLAS\_JET\_EtaIntercalibration\_NonClosure\_posEta
- ATLAS\_JET\_EtaIntercalibration\_NonClosure\_negEta
- ATLAS\_JET\_EtaIntercalibration\_NonClosure\_2018data
- ATLAS\_JET\_EtaIntercalibration\_Modelling
- ATLAS\_JES\_EffectiveNP\_Statistical6
- ATLAS\_JES\_EffectiveNP\_Statistical5
- ATLAS\_JES\_EffectiveNP\_Statistical4
- ATLAS\_JES\_EffectiveNP\_Statistical2
- ATLAS\_JES\_EffectiveNP\_Statistical1
- ATLAS\_JES\_EffectiveNP\_Modelling3
- ATLAS\_JES\_EffectiveNP\_Modelling2
- ATLAS\_JES\_EffectiveNP\_Modelling1
- ATLAS\_JES\_EffectiveNP\_Mixed3
- ATLAS\_JES\_EffectiveNP\_Mixed2
- ATLAS\_JES\_EffectiveNP\_Mixed1
- ATLAS\_JES\_EffectiveNP\_Detector1
- ATLAS\_JET\_BJES\_Response
- ATLAS\_EG\_SCALE\_ALL
- ATLAS\_EG\_RESOLUTION\_ALL

ATLAS Internal



Instrumental

- ATLAS\_PU\_PRW\_DATASF
- ATLAS\_MUON\_EFF\_TrigSystUncertainty
- ATLAS\_MUON\_EFF\_TrigStatUncertainty
- ATLAS\_MUON\_EFF\_PLVTIGHT\_ISO
- ATLAS\_MUON\_EFF\_ID\_SYS
- ATLAS\_JET\_JvtEfficiency
- ATLAS\_FT\_EFF\_extrapolation\_from\_charm
- ATLAS\_FT\_EFF\_extrapolation
- ATLAS\_FT\_EFF\_Eigen\_Light\_3
- ATLAS\_FT\_EFF\_Eigen\_Light\_2
- ATLAS\_FT\_EFF\_Eigen\_Light\_1
- ATLAS\_FT\_EFF\_Eigen\_Light\_0
- ATLAS\_FT\_EFF\_Eigen\_C\_3
- ATLAS\_FT\_EFF\_Eigen\_C\_2
- ATLAS\_FT\_EFF\_Eigen\_C\_1
- ATLAS\_FT\_EFF\_Eigen\_C\_0
- ATLAS\_FT\_EFF\_Eigen\_B\_3
- ATLAS\_FT\_EFF\_Eigen\_B\_2
- ATLAS\_FT\_EFF\_Eigen\_B\_1
- ATLAS\_FT\_EFF\_Eigen\_B\_0
- ATLAS\_EL\_SF\_Trigger\_TOTAL\_1NPCOR\_PLUS\_UNC
- ATLAS\_EL\_EFF\_RECO
- ATLAS\_EL\_EFF\_PLVTIGHT\_ISO
- ATLAS\_EL\_EFF\_ID



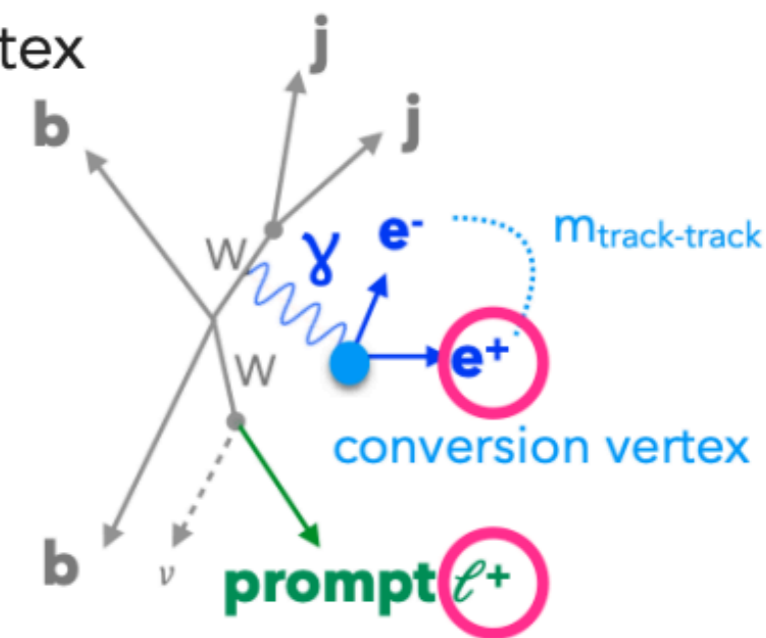
# HHML 联合拟合

Channels	Stats. Only (Asimov)	+ MC stats.(Asimov)	+ Systematics (Asimov)
$2\ell SS$	$32.34^{51.69}_{26.25}$	$33.55^{47.38}_{24.17}$	$35.8^{50.77}_{28.85}$
$3\ell$	$26.11^{37.12}_{18.82}$	$28.34^{40.77}_{20.43}$	$29.29^{42.2}_{15.7}$
$b\bar{b}4\ell$	$27.77^{41.59}_{20.02}$	$28.85^{44.01}_{20.79}$	$28.97^{44.28}_{20.87}$
$1\ell+2\tau_{\text{had}}$	$32.97^{47.30}_{23.74}$	$36.34^{51.76}_{26.18}$	$37.56^{53.65}_{27.06}$
$2\ell SStwotau$	$33.10^{48.77}_{23.88}$	$33.73^{49.55}_{24.30}$	$34.10^{50.20}_{24.55}$
$2\ell SStau$	$42.60^{61.10}_{30.70}$		$44.70^{64.50}_{32.2}$
Multilepton	$11.18^{15.89}_{8.05}$	$11.76^{16.76}_{8.47}$	$12.18^{17.40}_{8.78}$
$\gamma\gamma+1\ell 0\tau_{\text{had}}$	$25.46^{36.99}_{18.35}$	$25.50^{37.09}_{18.37}$	$25.64^{37.29}_{18.22}$
$\gamma\gamma+0\ell 1\tau_{\text{had}}$	$52.70^{76.71}_{37.98}$	$52.79^{76.91}_{38.04}$	$53.30^{77.60}_{38.40}$
$\gamma\gamma+2L$	$25.13^{38.23}_{18.11}$	$25.16^{38.32}_{18.13}$	$25.18^{38.34}_{18.15}$
$\gamma\gamma + ML$	$14.85^{21.65}_{10.70}$	$14.87^{21.70}_{10.71}$	$14.98^{21.86}_{10.80}$
Combined	$8.46^{12.06}_{6.09}$	$8.73^{12.49}_{6.29}$	$8.91^{12.77}_{6.42}$

# 光子转化电子定义

- \* Internal and material conversion electrons can be reconstructed as isolated electrons due to:
  - ▶ Reconstruction not considering conversions at  $r < 20$  mm (beam pipe: 24 mm)
  - ▶ Some conversion vertices are rejected due to track quality
  - ▶ Isolation not considering tracks associated with the electron candidate
- \* Variables are calculated in the gn1, new derivations have the information
  - ▶ the conversion radius
  - ▶ the invariant mass of the track associated to the electron and its closest track (originating from the conversion) calculated at conversion vertex
  - ▶ and the same invariant mass calculated at the primary vertex

electron CO selection	CO radius	$m_{\text{track-track}}$
(1) material CO	$> 20$ mm	$< 100$ MeV (wrt. CV)
(2) internal CO	<b>not (1)</b>	$< 100$ MeV (wrt. PV)
(3) very tight*	<b>not (1) and not (2)</b>	



# Fake factor 方法估计

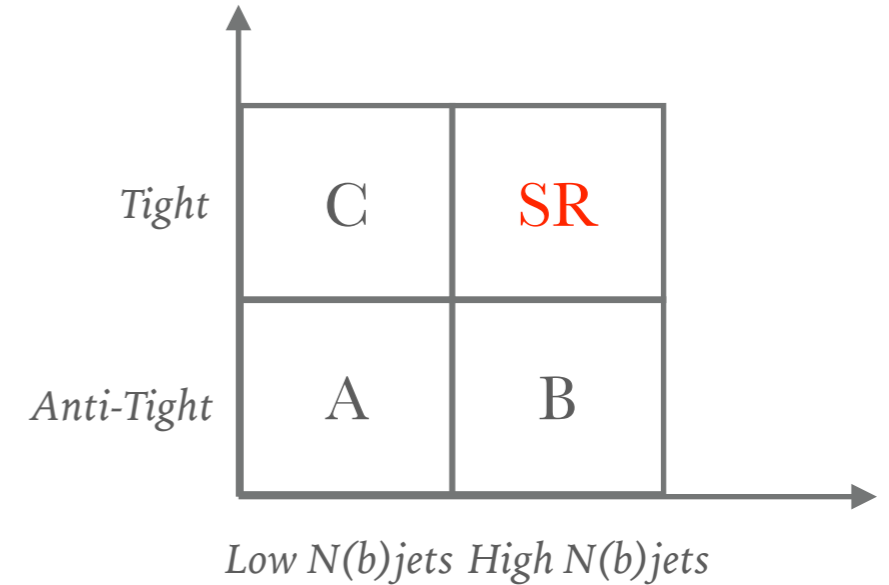
➤ 信号区的事例数估计:  $N_{SR} = \left(\frac{C}{A}\right) \times B$

$$N_{ee}^{\text{fakes}} (N_{\text{jet}} \geq 2) = \left( N_{e\phi}^{\text{data}} - N_{e\phi}^{\text{promptSS}} - N_{e\phi}^{V\gamma} - N_{e\phi}^{\text{QMisID MC}} \right) (N_{\text{jet}} \geq 2) \times \theta_e$$

$$N_{\mu\mu}^{\text{fakes}} (N_{\text{jet}} \geq 2) = \left( N_{\mu\mu}^{\text{data}} - N_{\mu\mu}^{\text{promptSS}} - N_{\mu\mu}^{V\gamma} \right) (N_{\text{b-jet}} \geq 1) \times \theta_\mu$$

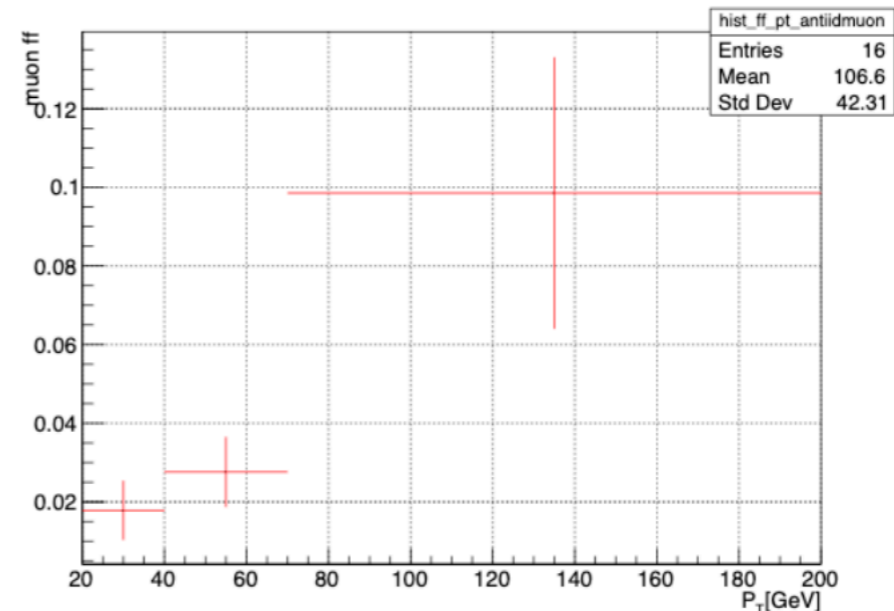
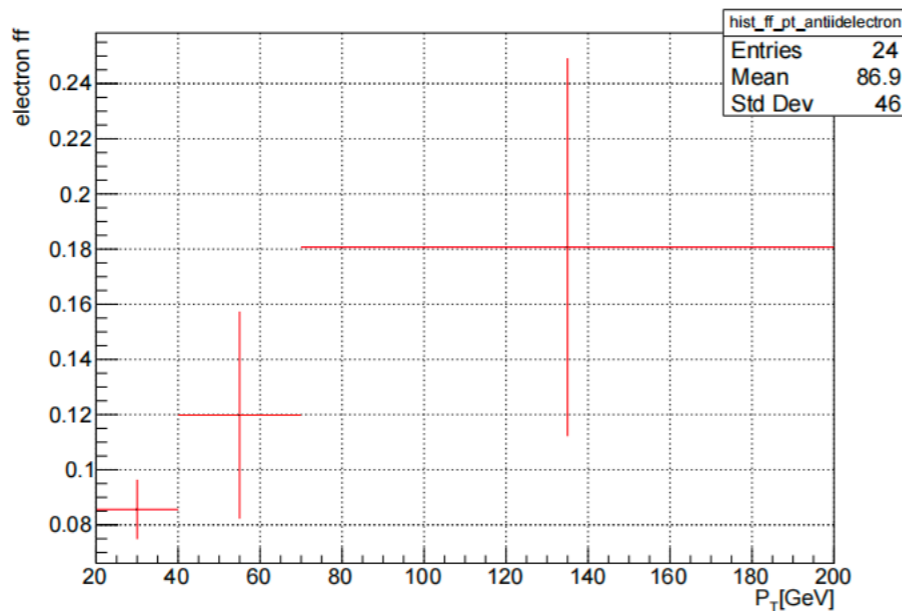
$$N_{e\mu}^{\text{fakes}} (N_{\text{jet}} \geq 2) = \left( N_{e\mu} - N_{e\mu}^{\text{promptSS}} - N_{e\mu}^{V\gamma} - N_{e\mu}^{\text{QMisID}} \right) (N_{\text{b-jet}} \geq 1) \times \theta_\mu$$

$$+ \left( N_{\phi\mu} - N_{\phi\mu}^{\text{promptSS}} - N_{\phi\mu}^{V\gamma} - N_{\phi\mu}^{\text{QMisID MC}} \right) (N_{\text{jet}} \geq 2) \times \theta_e$$

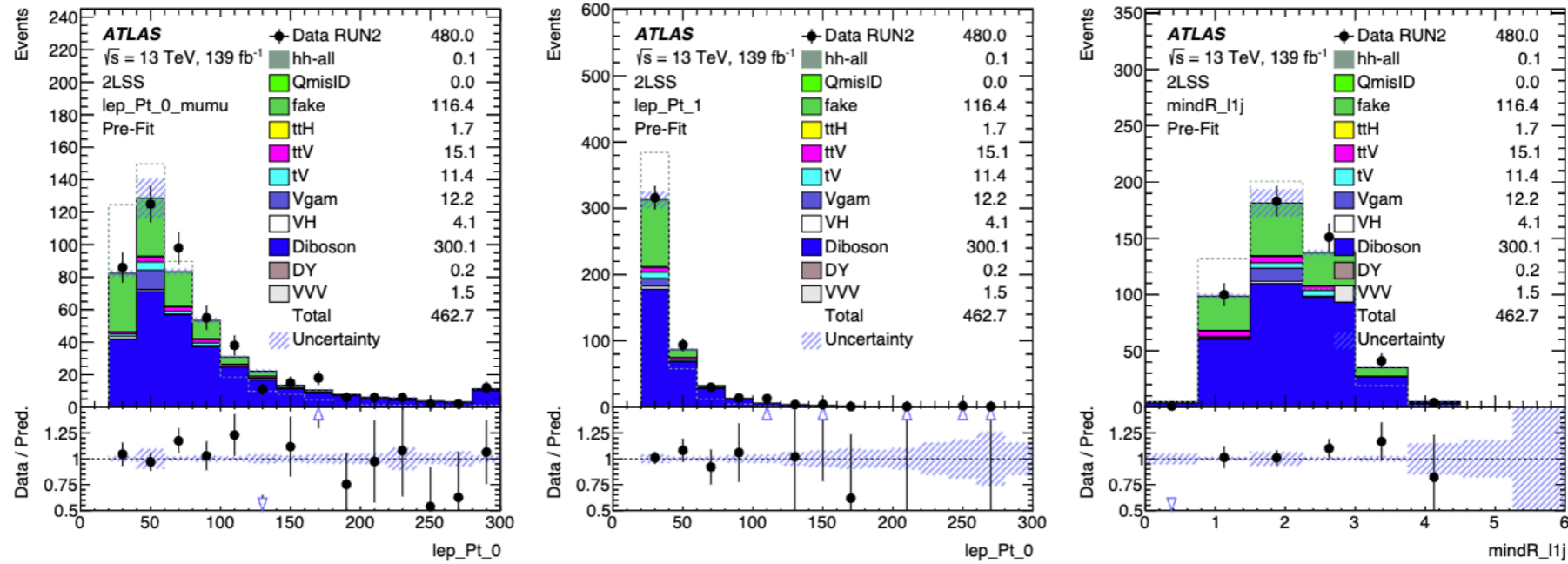


➤ fake factor 测量:

- Anti-Tight(A,B): 反转PLV 孤立化条件, 使得subleading轻子为non-prompt like 亲子。
- 测量区间(A,C): 电子: 较低的喷注数; 缪子: 较低的b-jet数

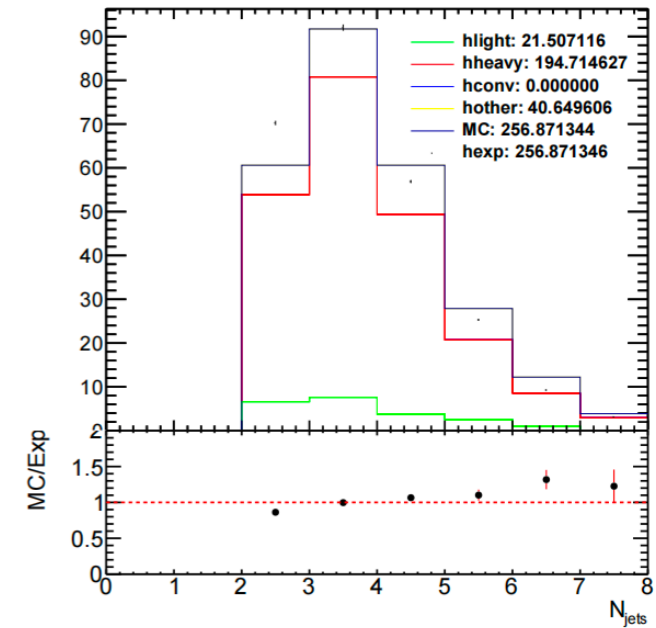
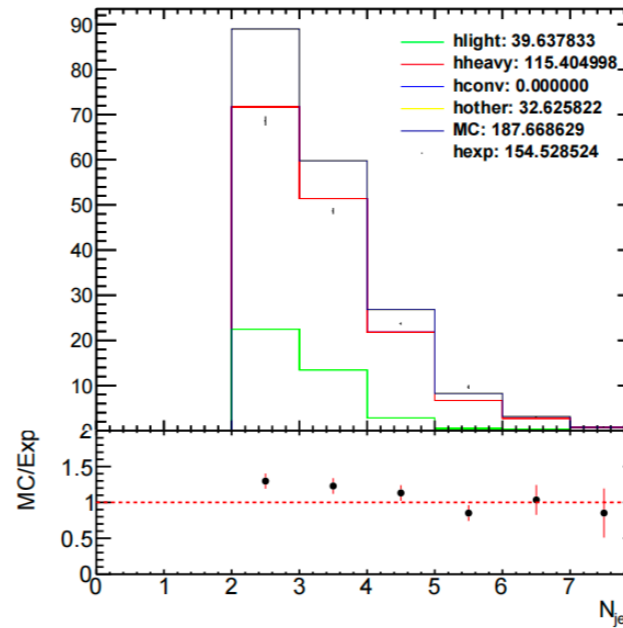


# Fake factor估计



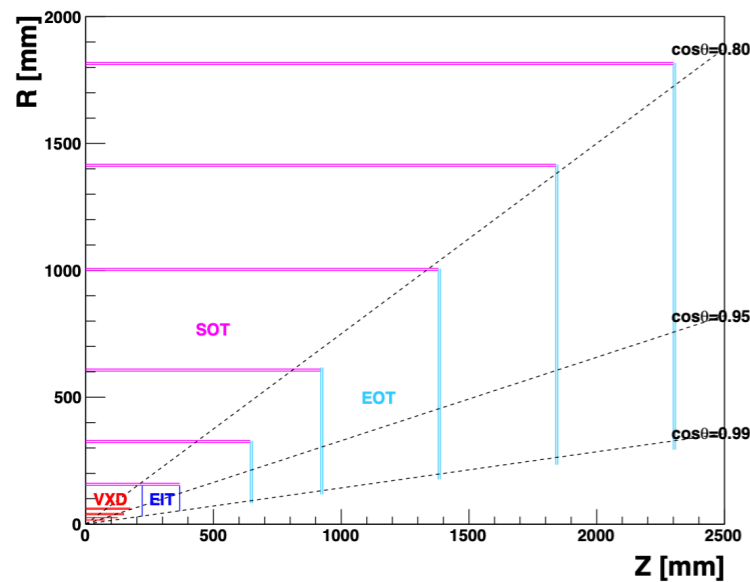
## ► CR to SR unclosure:

- CR和SR之间的组成差异。fake factor上的差异被考虑为系统性不确定性(对称处理)。
- 整体的yields误差: 17.6%。Jet 组分的变化: 重味 75.5%  $\rightarrow$  61.4%, 轻味 8.4%  $\rightarrow$  21.1%

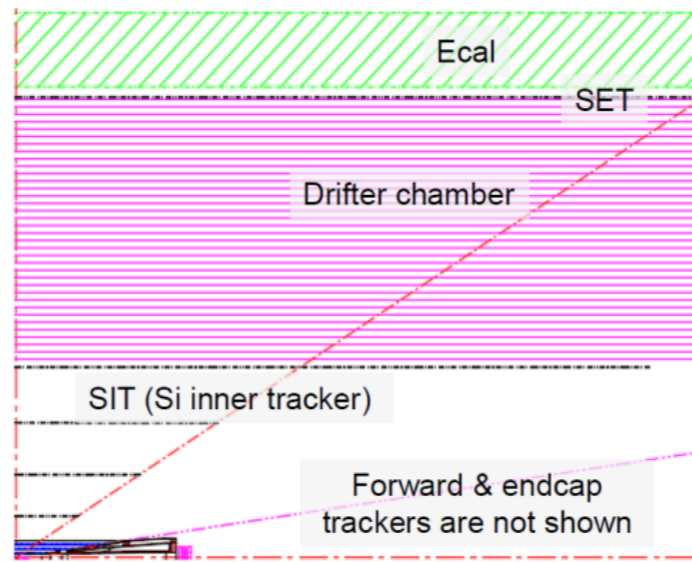


# Momentum performance

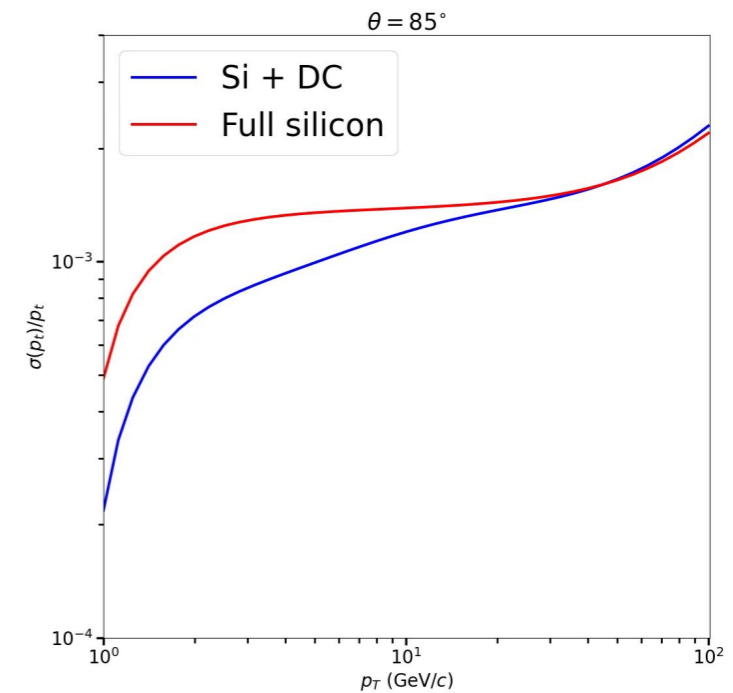
- High resolution and low budget Si system is required, such as silicon inner tracker, silicon external tracker for precise measurement.



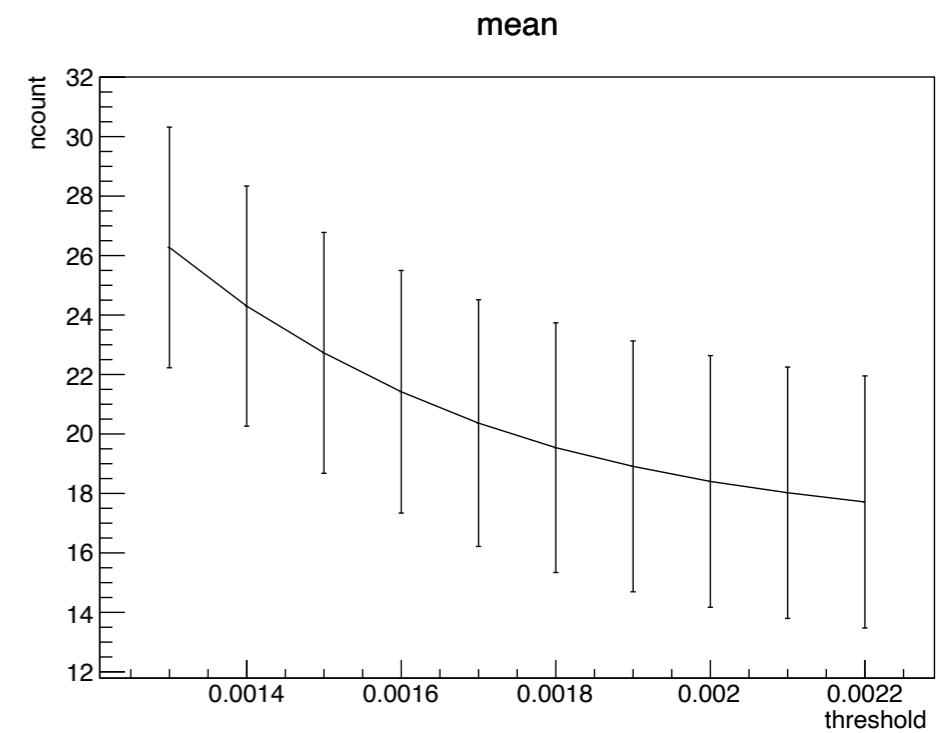
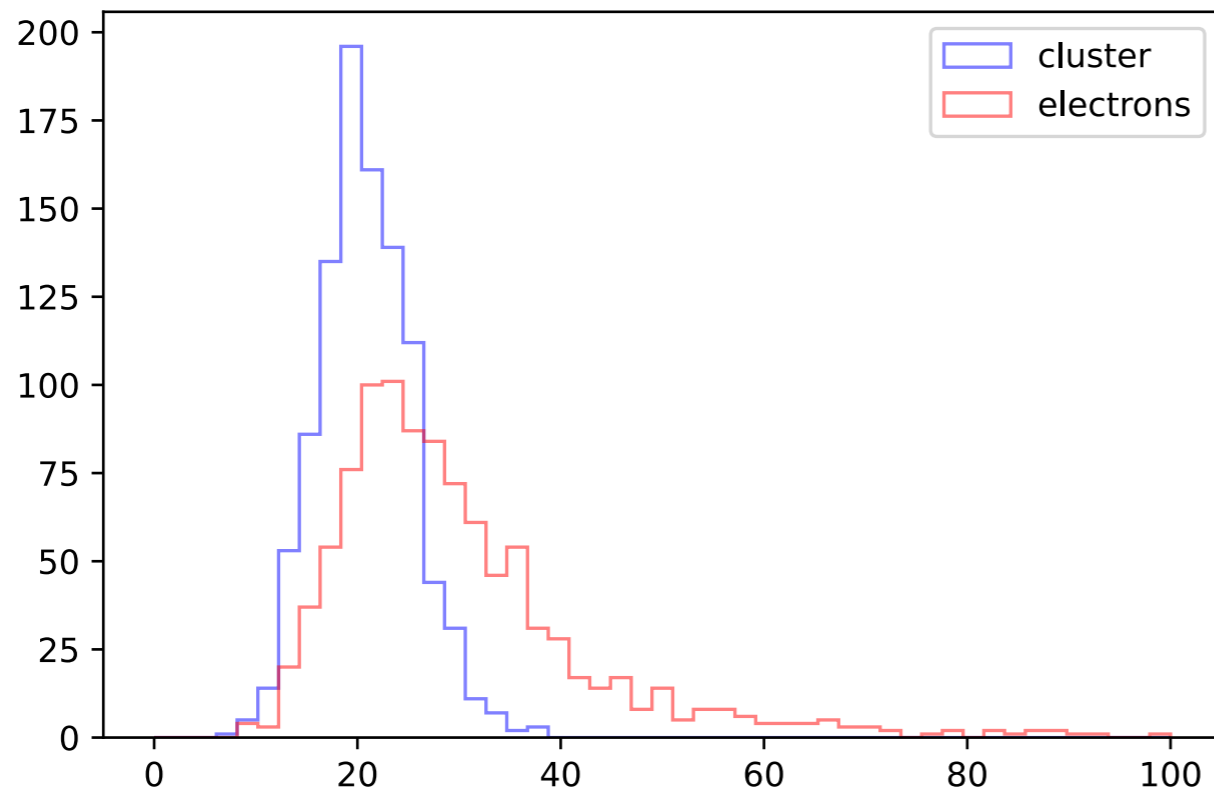
Full Silicon Tracker (FST)  
from CDR



Si + DC



- The change of drift chamber is significant to low momentum measurement.
- Si + DC design gives better momentum resolution than FST design  $\sim p_T < 50 \text{ GeV}/c$ .





.....

