



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

## 博士学位论文

ATLAS 上双光子末态希格斯 CP  
测量及 CEPC 电磁量能器重建算法

作者姓名: 郭方毅

指导教师: 姜辛丑 研究员 中国科学院高能物理研究所

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

2023 年 6 月



**The Higgs boson CP study in diphoton channel in ATLAS**  
**detector and the CEPC electromagnetic calorimeter**  
**reconstruction algorithm design**

**A dissertation submitted to**  
**University of Chinese Academy of Sciences**  
**in partial fulfillment of the requirement**  
**for the degree of**  
**Doctor of Natural Science**  
**in Particle Physics and Nuclear Physics**

**By**

**Guo Fangyi**

**Supervisor: Professor Lou Xinchou**

**Institute of High Energy Physics, Chinese Academy of Sciences**

**June, 2023**





## 中国科学院大学 学位论文原创性声明

本人郑重声明：所提交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。承诺除文中已经注明引用的内容外，本论文不包含任何其他个人或集体享有著作权的研究成果，未在以往任何学位申请中全部或部分提交。对本论文所涉及的研究工作做出贡献的其他个人或集体，均已在文中以明确方式标明或致谢。本人完全意识到本声明的法律结果由本人承担。

作者签名：

日 期：

## 中国科学院大学 学位论文授权使用声明

本人完全了解并同意遵守中国科学院大学有关收集、保存和使用学位论文的规定，即中国科学院大学有权按照学术研究公开原则和保护知识产权的原则，保留并向国家指定或中国科学院指定机构送交学位论文的电子版和印刷版文件，且电子版与印刷版内容应完全相同，允许该论文被检索、查阅和借阅，公布本学位论文的全部或部分内 容，可以采用扫描、影印、缩印等复制手段以及其他法律许可的方式保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延迟期后适用本声明。

作者签名：

日 期：

导师签名：

日 期：



## 摘要

粒子物理标准模型描述了基本粒子和相互作用，并在 TeV 量级以下粒子物理实验上取得了巨大成功。2012 年，ATLAS 和 CMS 实验宣布发现了希格斯玻色子，至此标准模型中预言的所有粒子都被找到。但标准模型并不是一个完美的理论，它无法描述实验中观测到的暗物质、中微子振荡、重子不对称等现象。寻找超出标准模型的新物理是如今粒子物理领域最重要的问题之一，新发现的希格斯玻色子可以成为这一研究的重要窗口。

ATLAS 是大型强子对撞机 LHC 上的通用探测器之一，它从 2008 年 LHC 运行之初七就在收集 LHC 上的质子-质子对撞数据，从中分析物理现象。液氙电磁量能器是 ATLAS 上的一个子探测器系统，用来测量电子和光子产生的电磁簇射。然而，模拟得到的电磁簇射无法精确描述实验数据是 ATLAS 上存在的一个长期问题，可能包括多方面原因，如探测器几何描述、电荷收集效应、探测器数字化模型等。本文介绍了对探测器电子学中的串扰和探测器几何模型对液氙量能器中电磁簇射形状影响的研究，当前结果显示电子学中的串扰效应可以显著影响电磁簇射  $\eta$  方向的形状，这证实了串扰可以是这一问题原因之一，但不是唯一原因。而探测器几何模型的调整显示出对簇射形状几乎没有影响，可以排除。这一工作是一个深入了解粒子物理实验的机会，但完全解决 ATLAS 上的这一问题还需要更多的研究。

在 Run 2 期间，ATLAS 实验上光子的重建采用了拓扑簇团的方法，而光子鉴别仍然基于电磁簇射形状变量。合作组期望拓扑簇团相关变量可以提供更多的信息，将有助于提升光子鉴别性能。因此本文中利用  $Z \rightarrow ll\gamma$  和  $Z \rightarrow ll + \text{jet}$  过程的蒙特卡洛模拟检查了拓扑簇团变量的光子鉴别能力。采用提升决策树方法，在加入 7 个拓扑簇团相关变量后信号和本底区分的 ROC 曲线有了显著提升。但与 ATLAS Run 2 部分数据的对比显示模拟样本对这些变量的描述并不完全准确，需要以后更深入的研究。

2015-2018 年 ATLAS 实验共采集了  $139 \text{ fb}^{-1} \sqrt{s} = 13 \text{ TeV}$  的质子-质子对撞数据，利用这些对撞数据我们可以进行各种物理问题的分析，如精确研究希格斯粒子的  $CP$  性质，这有助于进一步理解希格斯机制，同时寻找新物理的迹象。本文选用 ATLAS Run 2 全部数据在 VBF  $H \rightarrow \gamma\gamma$  过程测量希格斯粒子与矢量规范玻色子 (HVV) 相互作用过程中的  $CP$  破缺，引入了优化观测量方法，并训练

了两个提升决策树以得到更高纯度的 VBF 事例。我们在有效场论框架中用  $\tilde{d}$  和  $c_{H\tilde{W}}$  表示 HVV 相互作用中的 CP 破缺程度，目前结果没有发现存在 CP 破缺效应，并对这两个参数给出了 95% 置信度的约束，是当前对其最严格的约束。

在希格斯粒子发现后，世界各高能物理实验机构都在计划用做希格斯工厂的未来粒子对撞机，环形正负电子对撞机 (CEPC) 是提出的项目之一。CEPC 预计利用 10 年时间产生  $20 \text{ ab}^{-1} \sqrt{s} = 240 \text{ GeV}$  的  $e^+e^-$  对撞事例，用于希格斯物理的精确测量。在当前阶段探测器设计和优化是重点任务之一，CEPC 最新提出了一个由长晶体条组成的全吸收型电磁量能器设计方案，以期得到更好的电磁能量分辨和更低的造价。针对这一设计方案，本文提出了一套其需要的重建算法，将  $1 \times 1 \times \sim 40 \text{ cm}^3$  的长晶体条为单元的探测器重建为  $1 \times 1 \times 2 \text{ cm}^3$  的高颗粒度量能器，以满足粒子流重建 (PFA) 的要求。模拟研究证实了这套算法和这种量能器设计的可行性，相邻入射光子可以在 3 cm 以上的距离下被算法区分开并成功重建。但目前探测器的整体性能指标，玻色子质量分辨还没有得到。一套基于此算法的完整的新 PFA 软件框架也正在开发中。

CEPC 上光子相关的物理分析都可以从这一晶体电磁量能器中获益，因此我们选择  $H \rightarrow \gamma\gamma$  过程研究了在 CEPC 上  $\sigma(e^+e^- \rightarrow ZH) \times Br(H \rightarrow \gamma\gamma)$  预期的测量精度。这一分析根据 Z 玻色子的衰变末态分成了三个道，每个道中单独训练了一个提升决策树以区分信号和本底事例。在对 CEPC 系统误差进行了一些初步估计后，我们最终得到在  $5.6 \text{ ab}^{-1}$  数据量下预期精度为 7.9%，在  $20 \text{ ab}^{-1}$  数据量下为 4.6%。本文中进一步探究了电磁量能器能量分辨对  $\sigma(e^+e^- \rightarrow ZH) \times Br(H \rightarrow \gamma\gamma)$  预期精度的影响，这一结果对后续探测器设计和优化具有参考意义。

**关键词：** ATLAS 电磁簇射形状，光子鉴别，Higgs CP， $H \rightarrow \gamma\gamma$ ，CEPC 电磁量能器，粒子流算法

## Abstract

The Standard Model (SM) describes the fundamental interactions particles in our world. After its development in the last decades, the SM now reaches great success in sub-TeV level particle physics. The discovery of Higgs boson in 2012 by ATLAS and CMS groups marks all particles in the SM have been founded. However, the SM is not a complete theory. Several experimental phenomenons are our of its explanation, such as the dark matter, the neutrino oscillation, and the baryon asymmetry in the universe. Searching for the new physics beyond the SM is one of the most important topic in the field. And the precise measurement of the properties of recent discovered Higgs boson provides a good opportunity for this.

ATLAS is one of the general-purpose detectors at the Large Hadron Collider (LHC). It collects the proton-proton collision data delivered by LHC from the beginning of LHC run, at 2008. The Liquid Argon electromagnetic calorimeter (LAr EM calorimeter) is a sub-system of ATLAS, measuring the EM showers from photons and electrons. While there is a known long-term issue in ATLAS that the simulated EM shower development can not describe the data perfectly, especially in the lateral profile. The reason may come from many aspects, e.g. the detector geometry description, the charge collection effects, detector digitization, etc. In this paper we checked the impact from the cross talk effect in the detector electronics and the geometry description with the simulated single photon process. The results shows the cross talk can have notable influence on the  $\eta$  direction EM shower shapes, meaning the cross talk can be one of the reasons of shower shape mis-modeling issue. The geometry description can be excluded because it shows negligible influence on the shower shapes. This study is a good opportunity to deeply understand the particle experiment, while fully solving this issue still needs much effort from the collaboration.

The photon reconstruction in ATLAS experiment is based on the topo-cluster method in ATLAS Run-2, while the photon identification only uses the shower shapes in the LAr EM calorimeter. The additional information from the topo-clusters is expected to show some potential in improving the photon identification performance. This is checked with  $Z \rightarrow ll\gamma$  and  $Z \rightarrow ll + \text{jet}$  processes Monte Carlo simulation. With a boosted de-

cision tree, the ROC shows improvement after adding 7 topo-cluster related variables. These variables are further modeled with partial ATLAS Run-2 data. There is some mis-modeling in the variables awaiting dedicated studies in the future.

During the LHC Run-2 period from 2015 to 2018 ATLAS collects  $139 \text{ fb}^{-1} p-p$  collision data at  $\sqrt{s} = 13 \text{ TeV}$ , with which people can search and analyze physics objects like Higgs boson. Precise measurement of the CP properties of the Higgs boson can help us further understand the Higgs mechanism in the Standard model, and search the new CP violation source needed for the baryon asymmetry. We choose the VBF  $H \rightarrow \gamma\gamma$  channel to study the CP property in Higgs-gauge vector boson (HVV) interactions with ATLAS detector full Run-2 data. The Optimal Observable method is used to probe the CP structure in the interactions. Two boosted decision trees are trained to have high purity VBF events for the analysis. The results are presented within the effective field theory framework by  $\tilde{d}$  (HISZ basis) and  $c_{H\tilde{W}}$  (Warsaw basis). No sign of CP violation is observed in the data. The 95% confident level (CL) constraint intervals are obtained in different scenarios, which is the most stringent constraints on the CP violation in the HVV coupling.

After the discovery of Higgs boson people are planning the future collider to operate as a Higgs factory. The Circular Electron Positron Collider (CEPC) is one of the proposed projects. It plans to have  $20 \text{ ab}^{-1} e^+e^-$  collision data at  $\sqrt{s} = 240 \text{ GeV}$  after a 10-year running period for the Higgs physics study. At the first stage the design of detector is one of the main tasks. A new idea of homogeneous crystal EM calorimeter is under consideration, for better EM performance and lower cost. We proposed a set of reconstruction algorithms for this crystal bar EM calorimeter. The  $1 \times 1 \times \sim 40 \text{ cm}^3$  long crystal bar units in this EM calorimeter design are reconstructed to  $1 \times 1 \times 2 \text{ cm}^3$  high granularity small size hits for the Particle Flow Approach (PFA) reconstruction. The simulation proves the feasibility of this algorithm. The nearby photons can be separated with a distance of larger than 3 cm. While the detector benchmark, the boson mass resolution (BMR) is not derived yet. A complete new framework of the PFA reconstruction in CEPC is under development.

The photon-related studies in the CEPC can benefit from this crystal EM calorimeter. So we studied the expected measurement precision of  $\sigma(e^+e^- \rightarrow ZH) \times Br(H \rightarrow \gamma\gamma)$  at the CEPC. The analysis is departed into 3 sub-channels depending on the  $Z$  bo-

son decay. 3 boosted decision trees are developed in 3 sub-channels to separate the signal and background. After a preliminary consideration of the possible systematic uncertainties the precision can reach 7.9% with  $5.6 \text{ ab}^{-1}$  data and 4.6% with  $20 \text{ ab}^{-1}$  data. Furthermore, the impact of the EM calorimeter energy resolution on the expected  $\sigma(e^+e^- \rightarrow ZH) \times Br(H \rightarrow \gamma\gamma)$  precision is studied by smearing the photon energy in this analysis. The results can provide reference for the future detector optimization.

**Key Words:** ATLAS EM Shower Shape, Photon Identification, Higgs CP,  $H \rightarrow \gamma\gamma$ , CEPC Electromagnetic Calorimeter, Particle Flow.





## 目 录

第 1 章 引言	1
第 2 章 理论模型与研究背景	3
2.1 标准模型	3
2.1.1 规范对称性	3
2.1.2 强相互作用理论	5
2.1.3 电弱理论	6
2.1.4 自发对称性破缺机制与希格斯物理	9
2.1.5 Yukawa 耦合、费米子代混合和 CP 宇称	11
2.1.6 超出标准模型的新物理	12
2.2 希格斯物理相关实验	13
2.2.1 强子对撞中的希格斯物理	13
2.2.2 轻子对撞中的希格斯物理	15
第 3 章 大型强子对撞机和 ATLAS 探测器	19
3.1 大型强子对撞机	19
3.2 ATLAS 探测器	21
3.2.1 物理目标	21
3.2.2 磁铁	23
3.2.3 内部径迹探测器	23
3.2.4 量能器	24
3.2.5 缪子探测器	27
3.2.6 触发、读出、数据获取系统 [1, 2]	27
3.3 ATLAS 探测器上的粒子重建和鉴别	29
3.3.1 光子和电子 [3]	29
3.3.2 缪子 [4]	36
3.3.3 喷注 [5]	37
3.3.4 丢失横动量	38
第 4 章 ATLAS 上电磁簇射形状	39
4.1 电磁簇射形状的研究	39
4.1.1 串扰效应对簇射形状的影响	40
4.1.2 液氙量能器物质对簇射形状的影响	45
4.1.3 结论	45

4.2 基于簇射形状和拓扑簇团的光子鉴别 .....	46
4.2.1 数据和蒙特卡洛样本, 事例选择 .....	47
4.2.2 拓扑簇团变量 .....	48
4.2.3 多变量分析 .....	53
4.2.4 模拟与数据对照 .....	53
4.2.5 结论 .....	56
<b>第 5 章 ATLAS 探测器上通过 VBF <math>H \rightarrow \gamma\gamma</math> 过程测量希格斯粒子 CP 性质 .....</b>	<b>59</b>
5.1 研究背景 .....	59
5.2 理论模型与优化观测量 .....	60
5.2.1 HISZ 基下的 EFT .....	60
5.2.2 优化观测量 $\mathcal{O}\mathcal{O}$ .....	63
5.2.3 备选理论模型: SMEFT .....	65
5.3 数据与蒙特卡洛样本 .....	66
5.3.1 数据样本 .....	66
5.3.2 蒙特卡洛样本 .....	66
5.4 事例选择 .....	69
5.5 事例分类优化 .....	72
5.5.1 基于多变量分析的事例分类 .....	72
5.5.2 $\mathcal{O}\mathcal{O}$ 分割方式优化 .....	78
5.6 信号和本底参数化 .....	83
5.6.1 信号参数化 .....	83
5.6.2 本底参数化 .....	83
5.7 系统误差分析 .....	89
5.7.1 理论误差 .....	89
5.7.2 实验误差 .....	90
5.8 统计模型 .....	93
5.9 结果与总结 .....	95
<b>第 6 章 环形正负电子对撞机 (CEPC) 与粒子流重建 .....</b>	<b>101</b>
6.1 未来对撞机项目概况 .....	101
6.2 环形正负电子对撞机 (CEPC) 项目 .....	103
6.2.1 物理目标 .....	103
6.2.2 CEPC 基准探测器 .....	104
6.2.3 CEPC 软件环境 .....	106
6.2.4 CEPC 第四代探测器概念设计 .....	107
6.3 粒子流重建方法 .....	108

---

第 7 章 晶体电磁量能器重建算法的开发 .....	111
7.1 探测器几何搭建 .....	112
7.2 数字化 .....	113
7.3 重建算法软件设计 .....	115
7.3.1 晶体条能量劈裂算法 .....	115
7.3.2 能量-时间匹配算法 .....	118
7.3.3 软件实现 .....	120
7.3.4 初步性能 .....	121
7.4 总结与展望 .....	125
第 8 章 CEPC 上 $\sigma(ZH) \times Br(H \rightarrow \gamma\gamma)$ 测量精度的研究 .....	127
8.1 研究背景 .....	127
8.2 模拟样本 .....	127
8.3 物理对象重建与事例选择 .....	128
8.4 多变量分析 .....	131
8.5 信号与背景建模 .....	135
8.6 CEPC 上初步系统误差分析 .....	136
8.6.1 理论误差 .....	136
8.6.2 实验误差 .....	136
8.7 结果 .....	138
8.8 电磁量能器能量分辨的影响 .....	140
8.9 总结 .....	143
第 9 章 总结与展望 .....	145
参考文献 .....	149
致谢 .....	161
作者简历及攻读学位期间发表的学术论文与其他相关学术成果 ..	163



## 图目录

图 2-1 标准模型中的基本粒子 .....	4
图 2-2 $\alpha_s$ 随能标 $Q$ 跑动的测量结果, 总结截止至 2022 年 [6]。 .....	7
图 2-3 希格斯机制中引入的势能场 $V(\Phi)$ 。 .....	11
图 2-4 LHC 上产生希格斯玻色子过程的费曼图 [6]。其中: (a) 胶子熔合过程 ggF。 (b) 矢量玻色子熔合过程 VBF。 (c) (d) 矢量玻色子伴随过程 VH。 (e)(f)(g) 顶夸克伴随产生过程 ttH、tH。 .....	14
图 2-5 标准模型希格斯粒子的各过程产生截面随 $\sqrt{s}$ 变化 (左) 和衰变分支比随 $m_H$ 变化 (右) [6]。 .....	14
图 2-6 ATLAS (左) [7] 和 CMS (右) [8] 观测到的不同 $m_H$ 假设下的 local $p_0$ 值, 右图 CMS 的结果中显示了不同衰变道的贡献, 其中 $H \rightarrow \gamma\gamma$ 和 $H \rightarrow ZZ$ 占主导。 .....	15
图 2-7 ATLAS 实验上 Run 2 数据在 $H \rightarrow \gamma\gamma$ 道中对 STXS 参数的最佳拟合结果及误差 [9], 在误差范围内与标准模型相符。 .....	16
图 2-8 ATLAS 实验上 Run 2 数据在 $H \rightarrow \gamma\gamma$ 道中对 SMEFT 中耦合参数 $c_i$ 的测量结果及 68% 和 95% 置信区间, 不在图中的 SMEFT 参数被置为 0。图中蓝线和绿线分别表示对 SMEFT 进行线性项和线性项 + 二次项的参数化结果 [9]。 .....	17
图 2-9 $e^+e^-$ 对撞下希格斯粒子产生过程的截面 [10] .....	17
图 2-10 轻子对撞过程中希格斯粒子产生模式的费曼图, 从左到右依次为: $e^+e^- \rightarrow ZH$ , $e^+e^- \rightarrow \nu\bar{\nu}H$ , $e^+e^- \rightarrow e^+e^-H$ 。 .....	18
图 3-1 大型强子对撞机的加速器和储存环总揽图。 .....	20
图 3-2 ATLAS 在 Run 2 (左) 和 2022 Run 3 (右) 期间收集到的对撞数据量 (积分亮度) [11]。 .....	20
图 3-3 ATLAS 探测器示意图。 .....	21
图 3-4 ATLAS 磁铁系统示意图。 .....	23
图 3-5 ATLAS 内部径迹探测器。 .....	24
图 3-6 ATLAS 量能器结构示意图。 .....	26
图 3-7 ATLAS 触发系统流程图。 .....	29
图 3-8 电子和光子的重建流程。 .....	30
图 3-9 电子鉴别中使用的似然函数值分布图 [12]。 .....	33
图 3-10 不同工作点下 $Z \rightarrow e^+e^-$ 事例中的电子鉴别效率 [12]。 .....	34
图 3-11 Tight 工作点下光子鉴别效率和本底去除效率 [3]。 .....	34
图 3-12 ATLAS 数据和蒙特卡洛模拟中的 $R_\phi$ 和 $f_3$ 分布图 [3]。 .....	35
图 3-13 喷注能量分辨随横动量和 $\eta$ 的关系 [13]。 .....	38

图 4-1 从 $Z \rightarrow \mu^+ \mu^-$ 过程中得到的 LAr 电磁量能器桶部第一层（左）和第二层（右）串扰效应 [14]。其中红线为数字化中放入的串扰幅度，黑线为根据数字化模型从模拟样本中得到的串扰幅度，蓝线为数据中得到的串扰幅度。 .....	40
图 4-2 不同配置中簇射形状变量分布。 .....	42
图 4-3 不同 $\eta$ 下簇射形状变量的平均值，以 ATLAS 默认模拟样本中的值为参考。图中误差棒表示串扰修正上下限的影响。 .....	43
图 4-4 四个 $ \eta $ 区间内从簇射形状中拟合得到的串扰水平，左为 strip 层，右为 middle 层。 .....	43
图 4-5 Mean values of shower shape variables $R_\eta$ , $R_\phi$ , $w_{\eta_1}$ and $w_{\eta_2}$ in $ \eta  \in [0.2, 0.25]$ . .....	46
图 4-6 拓扑簇团几何性质的示意图。 .....	50
图 4-7 $Z \rightarrow ee\gamma$ 样本中 7 个拓扑簇团相关变量和 $m_{ll\gamma}$ 。 .....	51
图 4-8 $Z \rightarrow \mu\mu\gamma$ 样本中 7 个拓扑簇团相关变量和 $m_{ll\gamma}$ 。 .....	52
图 4-9 $Z \rightarrow e^+e^-\gamma$ 过程中 17 个输入变量的线性相关性矩阵。左：信号样本，右：本底样本。光子鉴别要求为 Loose ID。 .....	53
图 4-10 $Z \rightarrow e^+e^-\gamma$ 过程 17 个变量（左）和 10 个变量（右）训练出的 BDTG 分布。光子鉴别要求为 Loose ID。 .....	54
图 4-11 $Z \rightarrow e^+e^-\gamma$ 过程 17 个变量（左）和 10 个变量（右）训练出的 ROC 曲线。光子鉴别要求为 Loose ID。 .....	54
图 4-12 $Z \rightarrow e^+e^-\gamma$ 过程中 17 个输入变量的线性相关性矩阵。左：信号样本，右：本底样本。光子鉴别要求为 Tight ID。 .....	54
图 4-13 $Z \rightarrow e^+e^-\gamma$ 过程 17 个变量（左）和 10 个变量（右）训练出的 BDTG 分布。光子鉴别要求为 Tight ID。 .....	55
图 4-14 $Z \rightarrow e^+e^-\gamma$ 过程 17 个变量（左）和 10 个变量（右）训练出的 ROC 曲线。光子鉴别要求为 Tight ID。 .....	55
图 4-15 在 $Z \rightarrow e^+e^-\gamma$ （左）和 $Z \rightarrow \mu^+ \mu^- \gamma$ （右）过程中拟合 $m_{ll\gamma}$ 以得到数据中 $Z \rightarrow ll\gamma$ 和 $Z \rightarrow ll + \text{jets}$ 过程的比例。 .....	56
图 4-16 $Z \rightarrow e^+e^-\gamma$ 过程中 7 个拓扑簇团变量和 $m_{ll\gamma}$ 的数据-模拟对比。模拟样本中两部分按比例进行了叠加。 .....	56
图 4-17 $Z \rightarrow \mu^+ \mu^- \gamma$ 过程中 7 个拓扑簇团变量和 $m_{ll\gamma}$ 的数据-模拟对比。模拟样本中两部分按比例进行了叠加。 .....	57
图 5-1 EFT 模型中 $H \rightarrow \gamma\gamma$ 分支比随 $\tilde{d}$ 的变化。 .....	63
图 5-2 $\tilde{d} = 0.1$ （左上）、0.2（右上）、0.5（左下）和 0.9（右下）下 $\mathcal{O}\mathcal{O}$ 分布与 SM 过程的对比。BSM 的贡献被拆分成了干涉项（红色虚线）和二次项（绿色虚线）。 .....	64
图 5-3 真实级别和重建级别 VBF 过程 $\mathcal{O}\mathcal{O}$ 分布的对比。重建过程会由于能量分辨和效率问题对 $\mathcal{O}\mathcal{O}$ 谱型宽度有微小影响。 .....	65

图 5-4 VBF 过程在 $c_{H\bar{W}} = -0.5, 0$ (SM) 和 $0.5p_T^H$ (左) 和 $\mathcal{O}\mathcal{O}$ (右) 的对比。其他 Wilson 参数固定到 SM 值。 .....	66
图 5-5 MADGRAPH5_aMC@NLO 直接产生的 BSM 样本和经过 k-factor 方法加权得到的 BSM 样本的 $\mathcal{O}\mathcal{O}$ 分布。比值显示两种方法结果一致。 .....	68
图 5-6 VBF 事例分类中所用变量的分布图。可以看出 $\eta^{Zep}$ , $\Delta R_{\gamma,j}^{min}$ 和 $\Delta\eta_{jj}$ 的分布模拟样本无法精确描述边带数据。 .....	74
图 5-7 VBF (左)、ggF (中) 和连续本底 (右) 的变量线性相关性矩阵, $\mathcal{O}\mathcal{O}$ 和 $m_{\gamma\gamma}$ 也一同放入比较。图中无数据代表相关性小于 1%。 .....	75
图 5-8 在 $\eta^{Zep}$ (左) 和 $\Delta R_{\gamma,j}^{min}$ (右) 的分布中拟合 MC 与数据的比值进行加权修正。 .....	75
图 5-9 经过修正的事例分类变量分布图。 .....	76
图 5-10 BDT distribution (left) and ROC curve (right) in VBF-ggF BDT training. ....	77
图 5-11 BDT distribution (left) and ROC curve (right) in VBF-MC $\gamma\gamma$ BDT training. ....	77
图 5-12 BDT_VBF_ggH_tight (左) 和 BDT_VBF_ggH_loose (右) 区域中 VBF 信号综合显著性水平随 BDT 筛选点的变化。如果在一个筛选点下某一过程预期事例数为 0 则将显著性置为 0。 .....	78
图 5-13 分析中四类区域的定义。 .....	79
图 5-14 所有过程的两个 BDT 分布: $BDT_{VBF/ggF}$ (左) 和 $BDT_{VBF/\gamma\gamma}$ (右)。 .....	80
图 5-15 联合 VBF 显著性水平随 $\mathcal{O}\mathcal{O}$ 分割方式的变化, 扫描步长 0.5。 ..	80
图 5-16 本底样本构造过程, 以 TL $\mathcal{O}\mathcal{O}$ bin 3 为例。从左至右: 控制区数据得到的 $\gamma + jets$ 过程 (红色显示其中混入的 $\gamma\gamma$ 过程); 将高统计量 $\gamma\gamma$ 蒙特卡洛样本加权得到控制区数据 $\gamma + jets$ 过程 $m_{\gamma\gamma}$ 分布; 按照比例 $f_{\gamma\gamma}$ 将两部分混合后得到的最终本底样本, 与边带数据对比。 .....	85
图 5-17 TL $\mathcal{O}\mathcal{O}$ bin 3 中经过平滑后的 $m_{\gamma\gamma}$ 分布 (蓝线)、原本底样本分布 (红线) 和边带数据分布。 .....	86
图 5-18 TL $\mathcal{O}\mathcal{O}$ bin 3 中的 Wald 测试结果。 .....	89
图 5-19 VBF 过程中采用不同 PDF 计算 $\mathcal{O}\mathcal{O}$ 得到的分布。从图中可认为 PDF 的选择对 $\mathcal{O}\mathcal{O}$ 计算没有显著影响。 .....	90
图 5-20 ggF 过程不同区域和 $\mathcal{O}\mathcal{O}$ bin 中的 PDF、QCD、 $\alpha_S$ 误差 (上) 和分子簇射误差 (下)。 .....	91
图 5-21 VBF 过程不同区域和 $\mathcal{O}\mathcal{O}$ bin 中的 PDF、QCD、 $\alpha_S$ 误差 (上) 和分子簇射误差 (下)。 .....	91
图 5-22 揭盲后数据在三个分类 (左上、右上、左下) 和联合后 (右下) 的 $m_{\gamma\gamma}$ 分布。事例按照 $\ln(1 + S/B)$ 进行了加权。 .....	96

图 5-23 揭盲后数据在三个分类 (左上、右上、左下) 和联合后 (右下) 的 $\mathcal{O}\mathcal{O}$ 分布。统计 $\mathcal{O}\mathcal{O}$ 时仅计算 $m_{\gamma\gamma} \in [118, 132]$ GeV 区间中的事例数。图中一同展示了两个 BSM 样本中的 $\mathcal{O}\mathcal{O}$ 分布作为对比。 .....	97
图 5-24 通过似然函数模型得到的对 $\tilde{d}$ 和 $c_{H\tilde{W}}$ 的 NLL 曲线, 展示来自于不同分类的贡献。 .....	97
图 5-25 通过似然函数模型得到的对 $\tilde{d}$ 和 $c_{H\tilde{W}}$ 的 NLL 曲线, 展示考虑系统误差项的贡献。 .....	98
图 5-26 NLL curve .....	98
图 5-27 与 $H \rightarrow \tau^+\tau^-$ 联合测量后的 NLL 曲线 (预期结果与观测到结果)。 .....	99
图 6-1 ILC [15] (左) 和 CLIC [16] (右) 项目计划示意图。 .....	102
图 6-2 CEPC [17] (左) 和 FCC [18] (右) 项目计划示意图。 .....	103
图 6-3 CEPC 基准探测器在 $r-z$ (左) 和 $r-\phi$ (右) 方向上的结构示意图 [10]。 .....	105
图 6-4 iLCsoft 软件框架结构 (左) 和 CEPCsoft 工作流 (右)。 .....	107
图 6-5 CEPC 第四代探测器概念设计结构图。 .....	108
图 7-1 晶体 PFA 电磁量能器的两种设计思路: 小颗粒晶体块单端读出方案 (上) 和交叉晶体条两端读出方案 (下)。 [19, 20] .....	111
图 7-2 CEPC 第四代探测器概念设计中的八边形桶部电磁量能器总览 (左) 和其中一个梯形块 (右)。 .....	113
图 7-3 晶体条被击中时的数字化过程示意图。晶体中激发出的闪烁光子被两端 SiPM 接收并转化成电信号, 输出到读出系统中。 .....	114
图 7-4 重建算法流程示意。重建以数字化中给出的晶体条作为输入, 经历能量劈裂、超层内匹配、纵向连接的过程形成簇团, 进行两次迭代后形成粒子流对象 PFO。 .....	116
图 7-5 通过模拟光子样本拟合得到的电磁簇射横向发展分布。 .....	117
图 7-6 能量劈裂算法流程图。 .....	117
图 7-7 $M = N$ 情况下的模式识别示意图。首先计算每个交叉点的 $\chi_{ij}^2$ , 根据 $\chi_{ij}^2$ 计算组合出图案对应的 $\chi^2 comb$ , 最终选定 $\chi^2 comb$ 的图案。 ..	119
图 7-8 $M > N$ 情况下的模式识别示意图。此处我们假设 $\chi_{22}^2$ 是所有 $chi^2$ 中最小的, 点 (2, 2) 即被固定, 对应的一组 $U/V$ 簇射也被移除, 图案被简化为 $2 \times 1$ , 可直接确定。 .....	120
图 7-9 重建软件代码框架 (右)。左为参考的 PandoraSDK 代码框架 [21]。 ..	122
图 7-10 晶体 ECAL 中光子的相对能量分辨。红线为使用 $\sigma_E/E = a \oplus \frac{b}{\sqrt{E}} \oplus \frac{c}{E}$ 的拟合结果, 从拟合中得到 $a = 0.$ , $b = 1.45\%$ , $c = 3.80\%$ 。 ..	122
图 7-11 重建位置偏移随入射位置的变化, 图中误差棒对应重建位置分辨。 .....	124



图 7-12 入射光子示意图和重建效率随间距变化。重建效率定义为满足以下 3 个条件的重建事例的比例：重建出至少 2 个簇团，簇团能量正确 ( $\frac{2}{3}E_{\text{truth}} < E_{\text{rec}} < \frac{4}{3}E_{\text{truth}}$ ) 且簇团位置正确 ( $|Pos_{\text{truth}} - Pos_{\text{rec}}| < 10 \text{ mm}$ )。 ..... 124

图 8-1  $q\bar{q}\gamma\gamma$  道中训练变量分布 (已归一化)。 ..... 133

图 8-2  $\mu^+\mu^-\gamma\gamma$  道中训练变量分布 (已归一化)。 ..... 133

图 8-3  $\nu\bar{\nu}\gamma\gamma$  道中训练变量分布 (已归一化)。 ..... 134

图 8-4 三个道中信号样本双边水晶球函数拟合。从左至右： $q\bar{q}\gamma\gamma$  道， $\mu^+\mu^-\gamma\gamma$  道， $\nu\bar{\nu}\gamma\gamma$  道。 ..... 135

图 8-5 本底模型中测试的函数和拟合情况。从左至右： $q\bar{q}\gamma\gamma$  道， $\mu^+\mu^-\gamma\gamma$  道， $\nu\bar{\nu}\gamma\gamma$  道。 ..... 135

图 8-6 三个道中对 Asimov 数据的联合拟合。 ..... 139

图 8-7 全模拟  $H \rightarrow \gamma\gamma$  样本 (蓝线) 和两组 smearing 样本 (3% 红线, 16% 绿线) 中的  $m_{\gamma\gamma}$  分布。拟合得到的信号宽度分别为 2.81 GeV, 0.94 GeV 和 1.96 GeV。 ..... 141

图 8-8  $ZH \rightarrow q\bar{q}\gamma\gamma$  道中信号强度测量精度随 ECAL 能量分辨统计项的变化。数据点用式 8-2 中的关系进行了拟合。 ..... 142

图 8-9 不同 ECAL 能量分辨常数项情况下对应的统计项参考点。 ..... 143

表目录

表 3-1 ATLAS 探测器各部分的设计指标。表中  $E$  和  $p_T$  单位为 GeV。 .. 22

表 3-2 ATLAS 内部径迹探测器各部分的尺寸。 ..... 25

表 3-3 ATLAS 缪子探测器参数和功能。 ..... 28

表 3-4 电子与光子鉴别用到的变量。 ..... 32

表 3-5 电子孤立度工作点定义 ..... 35

表 3-6 光子孤立度工作点定义 ..... 36

表 3-7 不同工作点下缪子和轻强子的重建效率  $\epsilon_{\mu}$ ,  $\epsilon_{had}$ 。 Low- $p_T$  工作点下有两套方法, 分别基于筛选条件 (Low- $p_T$  cut-based) 和多变量分析 (Low- $p_T$  MVA)。Tight 工作点不重建  $p_T < 4 \text{ GeV}$  的, 缪子, 因此在  $3 < p_T < 5 \text{ GeV}$  区间缪子效率很低。 ..... 37

表 4-1 研究使用的模拟样本。 ..... 41

表 4-2 从 [14] 中得出的串扰幅度修正值。 ..... 41

表 4-3 数据中的 fudge factor、从数据中拟合得到的串扰水平和线性相关水平  $r^2$ 。前 5 行中的簇射形状变量对应于 strip 层, 后两行对应于 middle 层。仅在  $r^2 > 0.999$  情况下尝试从数据中拟合串扰水平。 ..... 44

表 4-4 本工作中使用的蒙特卡洛模拟样本。 ..... 47

表 4-5 拓扑簇团变量和簇射形状变量在两个道中的分辨能力。光子要求为满足 LooseID 和 LooseIso。 ..... 49

表 4-6 $Z \rightarrow e^+e^-\gamma$ 过程中用 10 个簇射形状变量和 17 个簇射形状 + 拓扑簇团变量训练出的 AUC 和 70% 本底去除率下的信号效率。 . . . . .	57
表 4-7 $Z \rightarrow \mu^+\mu^-\gamma$ 过程中用 10 个簇射形状变量和 17 个簇射形状 + 拓扑簇团变量训练出的 AUC 和 70% 本底去除率下的信号效率。 . . . . .	57
表 5-1 ATLAS Run 2 期间积分亮度和误差 [22, 23]。 . . . . .	66
表 5-2 分析中使用的蒙特卡洛样本。在本 VBF 分析中 ggF 过程产生的希格斯事例室主要共振本底，其他希格斯产生过程在 5.4 中验证过可忽略。 . . . . .	69
表 5-3 VBF、ggF 和 sideband data 的事例数和筛选效率。 . . . . .	71
表 5-4 其他希格斯产生模式经过选择条件后的事例数。经过事例分类的最终预期事例数在表 5-8 中。 . . . . .	72
表 5-5 VBF 事例分类中用到的变量和其分辨能力。 . . . . .	73
表 5-6 四个分类中各过程的预期事例数和 VBF 信号显著性。连续本底数归一到边带数据后从中抽取信号区 ( $m_{\gamma\gamma} \in [120, 130]$ GeV) 预期本底数，VBF 信号显著性水平则根据信号区本底进行计算。 . . . . .	79
表 5-7 每个分类和 $\mathcal{O}\mathcal{O}$ bin 中 VBF、ggF 预期事例数和边带数据事例数。 . . . . .	81
表 5-8 所有 bin 中各希格斯过程的预期事例数。最后一列为除 VBF 和 ggF 以外其他希格斯过程事例数占比。 . . . . .	82
表 5-9 二维边带区法测量得到的 $\gamma\gamma$ 组分分数 $f_{\gamma\gamma}$ ， $\gamma + jets$ 组分分数为 $1 - f_{\gamma\gamma}$ 。 . . . . .	84
表 5-10 最终决定的本底模型和对应的赝信号误差。在 TL $\mathcal{O}\mathcal{O}$ bin 5 和 TL $\mathcal{O}\mathcal{O}$ bin 6 中 $P(\chi^2) < 1\%$ ，在经过 toy 样本检查后确认为统计涨落，选定函数可以描述本底分布。 . . . . .	88
表 5-11 分析中系统误差项的总结。作用方式中 $N_{tot}$ 表示信号事例数误差值对所有希格斯过程 (VBF, ggF) 相同， $N_i$ 表示不同希格斯过程采用不同误差值。 . . . . .	94
表 5-12 预期和观测到的 $\tilde{d}$ 和 $c_{H\tilde{W}}$ 68% 和 95% 置信区间，以及与 $H \rightarrow \tau^+\tau^-$ 联合测量结果。表中 $H \rightarrow \tau^+\tau^-$ 预期结果与 [24] 有稍许差别，是由于其对信号区和控制区相关性处理方式对区别造成的。 . . . . .	96
表 6-1 CEPC 运行计划及各模式下预期的信号事例数 [25]。 . . . . .	103
表 6-2 喷注中各成分的能量占比及在探测器中的预期能量分辨。在使用各部分的最优测量结果综合全探测器信息后，可预期喷注总能量分辨为 $\sim 20\%/\sqrt{E_j}$ 。 . . . . .	109
表 7-1 数字化算法的可输入参数。 . . . . .	115
表 8-1 各物理过程的产生截面和使用的蒙特卡洛样本统计量。在 $q\bar{q}\gamma\gamma$ 和 $\mu^+\mu^-\gamma\gamma$ 道中仅考虑 ZH 过程为信号，在 $\nu\bar{\nu}\gamma\gamma$ 中 ZH $Z \rightarrow inv.$ 和 $W/Z$ fusion 过程均被考虑为信号。 . . . . .	128

表 8-2  $q\bar{q}\gamma\gamma$  道中的筛选条件和效率。 $\gamma_1(\gamma_2)$  定义为相对低能量（高能量）的光子； $\cos\theta_{\gamma\gamma}(\cos\theta_{jj})$  是双光子（双喷注）系统的极角； $\min|\cos\theta_{\gamma j}|$  是光子-喷注对中最小的  $\cos\theta$ 。 ..... 129

表 8-3  $\mu^+\mu^-\gamma\gamma$  道中筛选条件和事例数。 $\gamma_1(\gamma_2)$  定义与  $q\bar{q}\gamma\gamma$  道中相同； $M_{\gamma\gamma}^{recoil}$  是根据 CEPC 质心能量 240 GeV 计算的双光子系统反冲质量： $(M_{\gamma\gamma}^{recoil})^2 = (\sqrt{s} - E_{\gamma\gamma})^2 - p_{\gamma\gamma}^2 = s - 2E_{\gamma\gamma}\sqrt{s} + m_{\gamma\gamma}^2$ 。 ..... 130

表 8-4  $\nu\bar{\nu}\gamma\gamma$  道中筛选条件和事例数。 $M_{missing}$  是根据所有可见粒子计算得到的丢失质量。 ..... 130

表 8-5  $q\bar{q}\gamma\gamma$  道中 BDTG 训练输入变量。 ..... 131

表 8-6  $\mu^+\mu^-\gamma\gamma$  道中 BDTG 训练输入变量。 ..... 132

表 8-7  $\nu\bar{\nu}\gamma\gamma$  道中 BDTG 训练输入变量。 ..... 132

表 8-8 三个道中测试的 6 种本地函数拟合  $\chi^2/Ndof$  值。 ..... 136

表 8-9 对 Asimov 数据拟合得到的三个道中和联合后的  $\sigma(ZH) \times Br(H \rightarrow \gamma\gamma)$  预期精度。 $20 \text{ ab}^{-1}$  结果为对信号和本底样本重新加权后在 workspace 中重新拟合得到。统计误差中已包括了本底模型的贡献。 ..... 140

表 8-10 三个道中各系统误差项对  $(\sigma \times Br)/(\sigma \times Br)_{SM}$  结果的贡献。0.5% 理论误差仅加在  $q\bar{q}\gamma\gamma$  道中。 ..... 140



## 第 1 章 引言

自文明之初，人类一直在探寻物质世界的组成结构。从阴阳五行说到原子论，东西方哲人均对此有过大量思考。从 19 世纪末开始，现代物理学用实验的手段逐渐揭示物质的微观结构，用数学模型进行定量描述。此后一个世纪中，大量物理学家们不断进行深入探索，最终得到了标准模型 (Standard Model, SM) 这一迄今为止对物质结构和相互作用描述最为精确的科学理论。20 世纪中期以后，标准模型中预言的基本粒子逐一在大型加速器实验中被证实，2012 年欧洲核子中心 (CERN) 的大型强子对撞机实验 (LHC) 宣布发现了希格斯玻色子 [7, 8]，补全了标准模型的最后一块拼图。

然而标准模型并不是一个可以解释一切的终极理论，大量实验观测结果无法用标准模型去解释，如中微子振荡、暗物质、宇宙中正反物质不对称性等，我们需要一个能够兼容所有实验结果的新理论。目前 LHC 仍是研究这一问题最合适的高能物理实验装置之一。2015-2018 年间 LHC 通过质心能量 13 TeV 的质子-质子对撞产生高能粒子如希格斯玻色子，对其性质（质量、自旋宇称、耦合强度等）的精确测量可以用来寻找超出标准模型的新物理的迹象。除 LHC 以外，环形正负电子对撞机 (Circular Electron Positron Collider, CEPC) 是正在筹划中的下一代对撞机。相比于 LHC 上的质子对撞，CEPC 上的轻子对撞将提供非常干净的对撞环境，能够以更高精度实现对标准模型粒子（以希格斯粒子为主）性质的测量。这些实验装置将有希望推进人类对微观世界的认知。

本文中主要包含两个部分的内容：利用 ATLAS 探测器 Run 2 数据在双光子道中对希格斯粒子  $CP$  性质进行测量，及 CEPC 电磁量能器重建算法的开发。其中：

- 第二章简述了粒子物理标准模型，以及用于寻找超出标准模型新物理的有效场论框架。
- 第三章介绍了大型强子对撞机、ATLAS 探测器的组成结构和 ATLAS 实验中物理对象的重建方法。
- 第四章为对 ATLAS 液氩型电磁量能器上簇射形状的相关研究，包括实验数据与模拟数据间的对比，和利用簇射形状进行光子鉴别。
- 第五章描述了 ATLAS 实验上通过 VBF  $H \rightarrow \gamma\gamma$  过程对希格斯粒子与规范玻色子作用过程中  $CP$  性质的测量，相关结果发表在 [26]。

- 第六章简述了 CEPC 项目，包括物理目标、目前探测器设计方案和 CEPC 计划采用的粒子流重建。
- 第七章描述了 CEPC 第四代探测器概念设计中晶体条型电磁量能器重建算法的开发工作。
- 第八章为对 CEPC 上  $\sigma(\text{ZH}) \times \text{Br}(H \rightarrow \gamma\gamma)$  测量精度的估计，相关结果发表在 [27]。
- 第九章为总结和展望。

## 第 2 章 理论模型与研究背景

### 2.1 标准模型

标准模型是从量子场论发展而来，经过半个世纪的发展和补充逐渐形成的理论。其将组成世界的基本粒子分为两类：自旋半整数、组成物质的费米子（Fermion）和自旋整数、传播相互作用的玻色子（Boson）。费米子又根据参加的相互作用不同分为轻子（lepton）和夸克（quark），轻子和夸克各有 3 代，并且每个粒子还有其对应的反粒子。夸克带有分数电荷，由于夸克禁闭效应无法单独存在，通常由 3 个夸克组成重子如质子和中子，在核力的作用下结合成原子核，再吸引轻子中的电子组成原子，最终构成我们的世界。标准模型中的相互作用类型分为 3 类：电磁相互作用、弱相互作用和强相互作用（在理论发展中电磁和弱作用得到了统一），光子（ $\gamma$ ）、 $W^\pm$  玻色子和  $Z$  玻色子、胶子（gluon）分别为其传播子。另外还有希格斯玻色子通过自发对称性破缺机制为其他粒子赋予质量。图 2-1 展示了标准模型中的基本粒子种类。在场论的框架下标准模型的拉格朗日量可拆分为：

$$\mathcal{L}_{SM} = \mathcal{L}_{strong} + \mathcal{L}_{EW} + \mathcal{L}_{Yukawa} + \mathcal{L}_{Higgs}, \quad (2-1)$$

分别表示强相互作用项（QCD）、电弱相互作用项、赋予费米子质量的汤川耦合项和自发对称性破缺引入的希格斯项。其整体的规范群为  $SU(3)_C \times SU(2)_L \times U(1)_Y$ ， $C$  表示强相互作用过程的守恒量色荷， $L$  和  $Y$  分别表示电弱过程中的守恒量左手同位旋和超荷，群的对称性是标准模型的基础。下文对标准模型的规范对称性框架和各种相互作用有更为详细的介绍。

#### 2.1.1 规范对称性

物理学是对实验现象的归纳和总结。20 世纪中期以来，随着物理技术的发展，人们在实验中不断发现新的粒子和新的反应过程。其中对之后标准模型的建立最关键的就是总结出的守恒定律。而根据诺特定律，一个守恒量对应一种连续对称性。通过寻找满足相应对称性的规范场，我们可以构建起描述相应相互作用的规范理论，这是标准模型发展的基石。

此处以电磁相互作用为例简述这一过程。考虑自由狄拉克场  $\psi(x)$ ，其拉氏量为：

$$\mathcal{L}_{Dirac} = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \quad (2-2)$$

## Standard Model of Elementary Particles

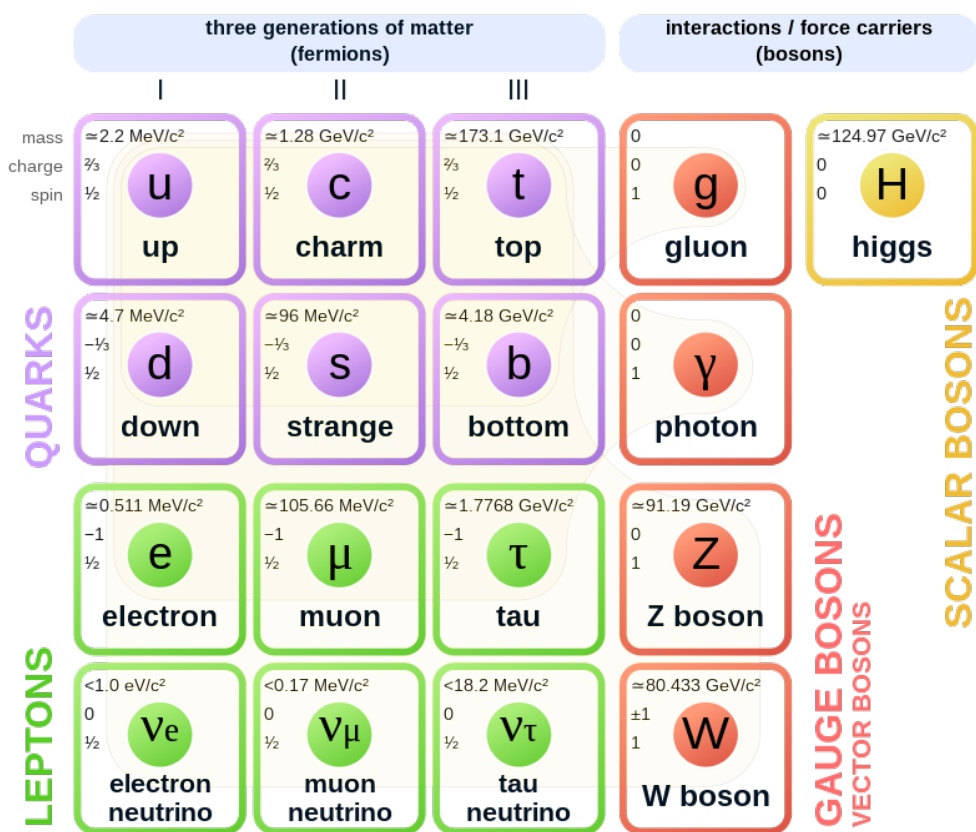


图 2-1 标准模型中的基本粒子



由电磁作用中的电荷守恒，此拉氏量应满足定域  $U(1)$  规范对称性：在变换

$$\psi(x) \rightarrow \psi(x)' = e^{ie\alpha(x)}\psi(x) \quad (2-3)$$

中保持不变。将式 2-3 带入式 2-2 中易发现， $\mathcal{L}'_{Dirac} = \mathcal{L}_{Dirac} - e\bar{\psi}(x)\gamma^\mu\psi(x)(\partial_\mu\alpha(x))$  并非不变，因此考虑将  $\partial_\mu$  替换为协变导数  $D_\mu$ ：

$$D_\mu = \partial_\mu + ieA_\mu(x), \quad (2-4)$$

其中  $A_\mu(x)$  为新引入的矢量场（即电磁场），在定域  $U(1)$  变换下满足变换关系：

$$A_\mu(x) \rightarrow A_\mu(x)' = A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x), \quad (2-5)$$

将其带入后可以发现，拉氏量  $\mathcal{L} = \mathcal{L}_{Dirac} - e\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu$  可以满足我们需要的规范对称性，后一项表示电子场（狄拉克场）与电磁场  $A_\mu(x)$  的相互作用  $\mathcal{L}_{int}$ 。在加入  $A_\mu(x)$  自身满足规范不变性的项  $\mathcal{L} = -\frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x)$ ， $F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$  后，我们就得到了 QED 的拉氏量， $A_\mu(x)$  的量子激发正是光子，且根据拉氏量光子是无质量的规范玻色子，也意味着电磁相互作用是远程相互作用。

### 2.1.2 强相互作用理论

在得到 QED 之后，对规范不变性的研究也推广到非阿贝尔群中 [28]，因此用于描述强相互作用的  $SU(3)$  规范群也可以通过规范不变性建立相互作用理论，即量子色动力学 QCD。考虑强相互作用中的不变量色荷，在色空间  $\alpha = (Red, Yellow, Blue)$  中自由拉氏量：

$$\mathcal{L}_0 = \Sigma_f \bar{q}_f(i\partial - m_f)q_f, \quad (2-6)$$

其中  $q_f$  表示夸克场， $f$  为夸克的味。类似 QED，考虑  $SU(3)$  规范变换的不变性后会引入协变导数  $D^\mu$ 。与 QCD 不同的是， $SU(3)$  群有 8 个生成元，因此在协变导数中有 8 个独立的规范玻色子场  $G^\mu(x)$ （即胶子场）：

$$D^\mu q_f \equiv (\partial^\mu - ig_s \frac{\lambda^a}{2} G_a^\mu(x))q_f \quad (2-7)$$

$g_s$  表示强相互作用的耦合常数，通常定义  $\alpha_s \equiv g_s^2/(4\pi)$ ， $\frac{\lambda^a}{2}$  ( $a = 1, 2, \dots, 8$ ) 是  $SU(3)$  群的生成元。规定  $G^\mu(x)$  的规范变换后，可以得到 QCD 拉氏量：

$$\begin{aligned} \mathcal{L}_{QCD} &= -\frac{1}{4}G_a^{\mu\nu}G_{\mu\nu}^a + \sum_f \bar{q}_f(i\gamma^\mu D_\mu - m_f)q_f. \\ &= -\frac{1}{4}(\partial^\mu G_a^\nu - \partial^\nu G_a^\mu)(\partial_\mu G_\mu^a - \partial_\nu G_\mu^a) + \sum_f \bar{q}_f^\alpha(i\gamma^\mu \partial_\mu - m_f)q_f^\alpha \\ &\quad + g_s G_a^\mu \sum_f \bar{q}_f^\alpha \gamma_\mu \left(\frac{\lambda^a}{2}\right)_{\alpha\beta} q_f^\beta \\ &\quad - \frac{g_s}{2} f^{abc}(\partial^\mu G_a^\nu - \partial^\nu G_a^\mu)G_\mu^b G_\nu^c - \frac{g_s^2}{4} f^{abc} f_{ade} G_b^\mu G_c^\nu G_\mu^d G_\nu^e. \end{aligned} \quad (2-8)$$

上式被拆分成了三部分，前两项表示胶子场和夸克场的动能项，第三项表示夸克-胶子相互作用，第四、五项表示三胶子和四胶子的自耦合。

$SU(3)$  理论的可重整化要求  $\alpha_s$  不能是定值，而是依赖于重整化能标  $\mu$ ，满足关系重整化群方程：

$$\mu^2 \frac{\partial \alpha_s}{\partial \mu^2} = \beta(\alpha_s) = -\frac{\beta_0}{4\pi} \alpha_s^2 - \frac{\beta_1}{(4\pi)^2} \alpha_s^3 - \dots \quad (2-9)$$

即能量越大 QCD 相互作用强度越弱，这与强作用为短程作用一致。同时这还使得 QCD 具有两个特殊性质：

**渐近自由** 在大动量转移、小尺度的作用过程，如强子内部或高能对撞反应中，强相互作用变弱，夸克可近似看作自由粒子，此时微扰论近似成立 ( $\alpha_s \ll 1$ )。

**夸克禁闭** 在低动量转移、长距离作用过程中，夸克之间相互作用变得非常大，其中势能使得可以凭空产生正反夸克对，形成新的强子。因此夸克只能存在于强子内部。同时在低能反应中，微扰论不再成立，必须通过非微扰 QCD 研究相关问题，这也使得强相互作用的计算到目前仍有一定困难。

图 2-2 展示了目前实验中对  $\alpha_s$  随能标  $Q$  跑动的测量结果 [6]。

### 2.1.3 电弱理论

20 世纪 30 年代，人们首先从原子核的  $\beta$  衰变过程中发现了弱作用力的存在，并逐渐发展出描述弱作用的费米有效理论。但在这一理论中，四费米子过程的反应截面  $\sigma \sim G_F^2 p_{cm}^2$  在高能量时发散。在类似 QED 引入中间矢量玻色子形成 V-A 弱流结构后这一问题得到解决，同时确定了弱作用满足  $SU(2) \times U(1)$  规范群，且可以与电磁作用完成统一（后续研究发现若要解决弱作用中的矢量玻色子质量问题，必须将电弱作用框架进行统一）。我们以第一代轻子为例讨论电弱模型：

$$\psi_1(x) = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \psi_2(x) = \nu_{eR}, \psi_3(x) = e_{R}^- \quad (2-10)$$

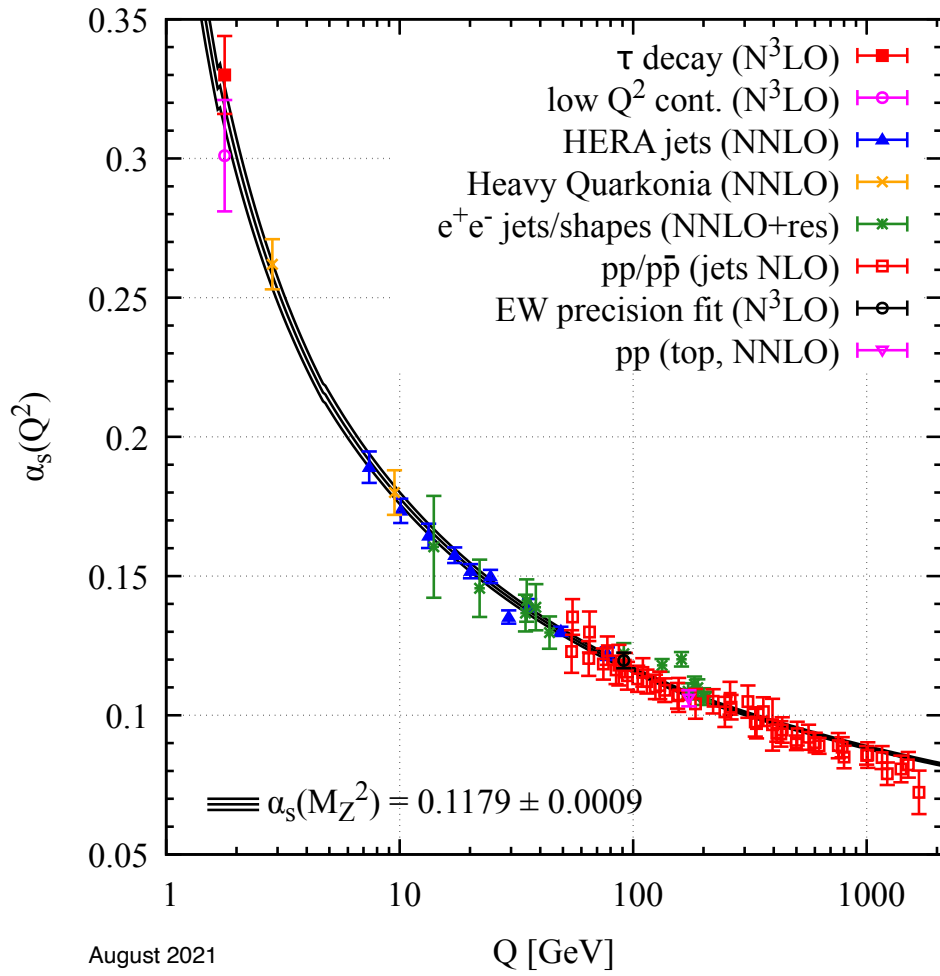


图 2-2  $\alpha_s$  随能标  $Q$  跑动的测量结果，总结截止至 2022 年 [6]。

自由拉氏量为:

$$\mathcal{L}_0 = \sum_{j=1}^3 i\bar{\psi}_j(x)\gamma^\mu\partial_\mu\psi_j(x). \quad (2-11)$$

左手费米子满足  $SU(2)$  规范变换不变性, 右手费米子满足  $U(1)$  规范变换不变性, 需要在协变导数中引入 4 个规范玻色子场:

$$\begin{aligned} D_\mu\psi_1(x) &\equiv [\partial_\mu - ig\tilde{W}_\mu(x) - ig'y_1B_\mu(x)]\psi_1(x), \\ D_\mu\psi_2(x) &\equiv [\partial_\mu - ig'y_2B_\mu(x)]\psi_2(x), \\ D_\mu\psi_3(x) &\equiv [\partial_\mu - ig'y_3B_\mu(x)]\psi_3(x), \end{aligned} \quad (2-12)$$

其中  $\tilde{W}_\mu(x) \equiv \frac{\sigma_i}{2}W_\mu^i(x)$ , 场强张量为:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk}W_\mu^jW_\nu^k. \quad (2-13)$$

电弱作用的拉氏量写为:

$$\begin{aligned} \mathcal{L}_{EW} &= \mathcal{L}_0 + \mathcal{L}_{int} \\ &= \sum_{j=1}^3 i\bar{\psi}_j(x)\gamma^\mu D_\mu\psi_j(x) - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^iW_i^{\mu\nu}. \end{aligned} \quad (2-14)$$

结合实验结果我们可以对此拉氏量进行一定变换:

**弱作用带电流** 式2-14中包含费米子与规范玻色子之间的带电流相互作用项:

$$\mathcal{L} = g\bar{\psi}_1\gamma^\mu\tilde{W}_\mu\psi_1 + g'B_\mu\sum_{j=1}^3 y_j\bar{\psi}_j\gamma^\mu\psi_j. \quad (2-15)$$

将  $\tilde{W}_\mu$  写为  $SU(2)$  矩阵形式:

$$\tilde{W}_\mu = \frac{\sigma_i}{2}W_\mu^i = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{2}W_\mu^3 & W_\mu^+ \\ W_\mu^- & -\sqrt{2}W_\mu^3 \end{pmatrix} \quad (2-16)$$

其中  $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2)$ , 带入拉氏量后有

$$\mathcal{L}_{CC} = \frac{g}{2\sqrt{2}} [W^+ [\bar{\nu}_e\gamma^\mu(1 - \gamma^5)e] + \dots] \quad (2-17)$$

将左右手轻子统一, 表示了带电流过程。但这与之前根据实验发展出的中间玻色子理论模型仍有一定差别, 在于此拉氏量中不包含玻色子质量项, 所描述的相互作用也应为远程作用。

**弱作用中性流** 式2-14中还包括中性规范场  $W_\mu^3$  和  $B_\mu$ ，我们对其做一定变换：

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \equiv \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \quad (2-18)$$

则中性流相互作用的拉氏量为：

$$\mathcal{L}_{NC} = \sigma_j \bar{\psi}_j \gamma^\mu \left[ A_\mu \left[ g \frac{\sigma_3}{2} \sin \theta_W + g' y_j \cos \theta_W \right] + Z_\mu \left[ g \frac{\sigma_3}{2} \sin \theta_W - g' y_j \cos \theta_W \right] \right] \psi_j. \quad (2-19)$$

当  $g \sin \theta_W = g' \cos \theta_W = e$  时， $A_\mu$  即 QED 中的电磁场，而  $Z_\mu$  对应于  $Z$  玻色子（除缺少质量项）， $\theta_W$  即为温伯格角， $\tan \theta_W = g'/g$ 。实验上  $Z$  玻色子的中性弱流于 1973 年在 CERN 的  $\nu_\mu e$  和  $\nu_\mu N$  散射实验中首次被观测到 [29, 30]，晚于电弱统一理论的形成，因此这也成为了对电弱统一理论的一次有效验证。 $W^\pm$  和  $Z$  玻色子的直接发现则要到 1983 年 CERN 质子-反质子对撞机 SPS 上的 UA1 和 UA2 实验组 [31–34]。

**规范玻色子的自相互作用** 在电弱模型的规范场动能项中，还包含 3-玻色子和 4-玻色子相互作用顶点，其中均包含一对正反  $W$  玻色子  $W^\pm$ ，但同样没有质量项，因为质量项会破坏  $SU(2) \times U(1)$  规范对称性。

#### 2.1.4 自发对称性破缺机制与希格斯物理

为解决之前反复提到的玻色子质量问题，1964 年 Francois Englert、Robert Brout [35] 和 Peter Higgs [36, 37] 分别独立提出了一种规范对称性的自发破缺机制，可以使规范玻色子获得质量。这一机制在原本的  $SU(2) \times U(1)$  电弱模型中加入了一个复标量场二重态  $\Phi(x)$  和其自身的势能函数  $V(\Phi)$ ：

$$\Phi(x) \equiv \begin{pmatrix} \phi^{(+)}(x) \\ \phi^{(0)}(x) \end{pmatrix}, \quad (2-20)$$

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad \mu^2 < 0, \lambda > 0.$$

复标量场自身的拉氏量和协变导数定义为：

$$\mathcal{L}_\Phi = \partial_\mu \Phi^\dagger \partial^\mu \Phi - V(\Phi), \quad (2-21)$$

$$D_\mu \Phi = \left( \partial^\mu - ig \tilde{W}^\mu - \frac{1}{2} ig' B^\mu \right) \Phi.$$

拉氏量本身在  $SU(2) \times U$  变换下是保持不变的。在确保  $\mu^2 < 0$  时  $V(\Phi)$  函数形式如图 2-3，有多个极小值，满足  $|\langle 0 | \phi^{(0)} | 0 \rangle| = \sqrt{\frac{-\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}}$ ， $v = \sqrt{\frac{-\mu^2}{\lambda}}$ 。当我

们选定一个特定的基态，如规定：

$$\langle \Phi \rangle_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}. \quad (2-22)$$

此时  $SU(2) \times U(1)$  对称性自发破缺到  $U(1)_{QED}$ 。我们将  $\Phi(x)$  写出其实分量  $\theta_j(x), j = 1, 2, 3, 4$ ，参数化为：

$$\Phi(x) = \exp\left(i\frac{\sigma_i}{2}\theta_i(x)\right) \begin{pmatrix} 0 \\ \frac{v+H(x)}{\sqrt{2}} \end{pmatrix} \quad (2-23)$$

对其进行规范变换重写为：

$$\Phi = \frac{1}{2} \begin{pmatrix} 0 \\ \frac{v+H(x)}{\sqrt{2}} \end{pmatrix} \quad (2-24)$$

$H(x)$  为物理的希格斯场。带入协变导数中，可得到

$$(D_\mu \Phi)^\dagger (D^\mu \Phi) = \frac{1}{2} (\partial_\mu H)^2 + (v + H)^2 \left[ \frac{g^2}{4} W_\mu^\dagger W^\mu + \frac{g^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu \right]. \quad (2-25)$$

即自发对称性破缺产生的 3 个无质量 Goldstone 粒子 [38] 在电弱拉氏量中被中间玻色子吸收，变为有质量的玻色子： $M_W = \frac{vg}{2}$ ， $M_Z = \frac{v}{2} \sqrt{g^2 + g'^2} = M_W / \cos \theta_W$ 。光子则由于保持了  $U(1)_{QED}$  对称性而仍旧没有质量。同时， $\mathcal{L}_\Phi$  中的动能项表示了希格斯场自相互作用：

$$\mathcal{L}_{\Phi,int} = \mu^2 H^2 - \lambda v H^3 - \frac{\lambda}{4} H^4 \quad (2-26)$$

给出了希格斯粒子的质量  $m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda}v$  和 3-希格斯、4-希格斯自耦合顶点。

以上自发对称性破缺的电弱理论中，各参数都可以做独立测量，进行交叉验证，并对实验尚未发现的新粒子质量进行预言。在此模型下对  $\beta$  衰变过程宽度的计算可以得到  $\frac{g^2}{M_W^2} = 4\sqrt{2}G_F$ ，即根据费米常数预言  $W$  玻色子质量，再结合  $\sin^2 \theta_W$  给出  $Z$  玻色子质量。从  $W$ 、 $Z$  玻色子的质量中可以得到真空期望值  $v \sim 246\text{GeV}$ ，又能够对希格斯粒子质量进行一定限制。2012 年，CERN 的大型强子对撞机 LHC 上两个合作组 ATLAS 和 CMS 同时宣布独立发现了希格斯玻色子，质量  $m_H \sim 125\text{GeV}$  [7, 8]。到目前为止，标准模型电弱理论显示出其对描述实验结果的巨大成功。但另一方面，如此复杂的模型能够相互嵌合让人不免怀疑其中的巧合，因此近年来人们一直希望通过对标准模型中各物理参数的精确测量来验证标准模型是否完全自洽，是否存在新物理的迹象。近年最重要的结果之

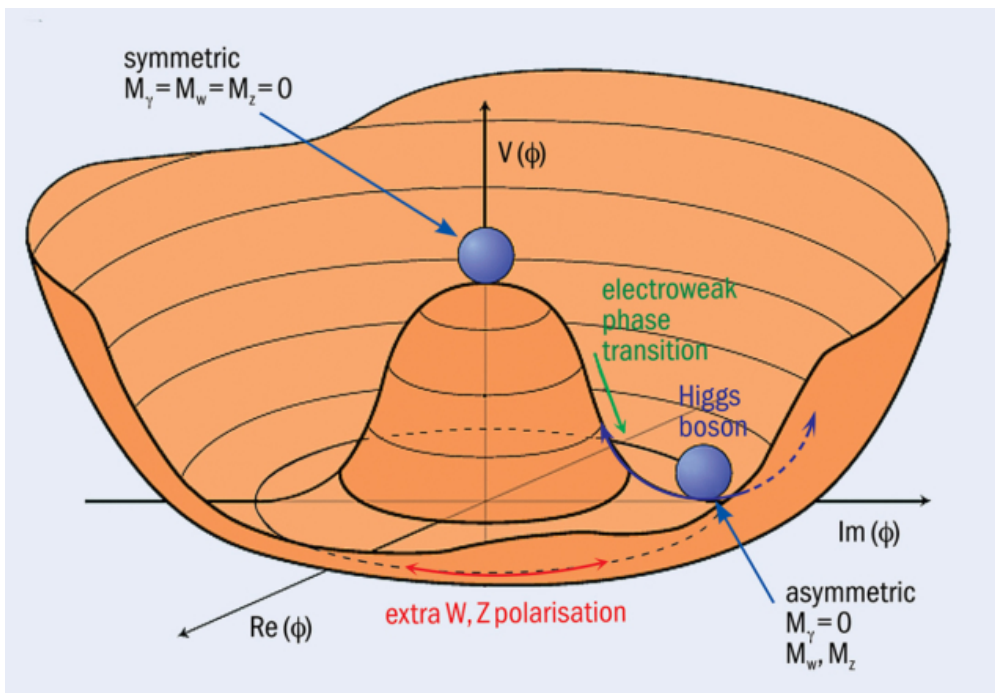


图 2-3 希格斯机制中引入的势能场  $V(\Phi)$ 。

一即 CDF 合作组利用 Tevatron 质子-反质子在 1.96 TeV 对撞数据中对  $W$  玻色子质量的测量结果  $M_W = 80433.5 \pm 9.4 \text{ MeV}/c^2$  [39]，与原标准模型预期值之间存在  $7\sigma$  的偏差。但考虑到电弱模型的复杂性，这一结果需要更多对电弱理论全局拟合的检验。

### 2.1.5 Yukawa 耦合、费米子代混合和 CP 宇称

在对称性自发破缺之前，由于电弱理论中左手费米子是弱同位旋二重态，右手费米子是弱同位旋单态，其间的耦合会破坏  $SU(2) \times U(1)$  对称性，因此所有费米子必须是 0 质量的，这与实验现象不符。在考虑自发对称性破缺后，希格斯场与费米子场之间存在耦合即 Yukawa 耦合，相互作用拉氏量为（仅以第一代费米子为例）：

$$\begin{aligned} \mathcal{L}_{Yukawa} &= y_d (\bar{u}, \bar{d})_L \Phi d_R + y_u (\bar{u}, \bar{d})_L \tilde{\Phi} u_R + y_e (\bar{\nu}_e, \bar{e})_L \Phi e_R + h.c. \\ &= \frac{v + H}{\sqrt{2}} (y_d d \bar{d} + y_u u \bar{u} + y_e e \bar{e}) + \dots \end{aligned} \quad (2-27)$$

从中可得到费米子质量  $m_f = \frac{y_f v}{\sqrt{2}}$ ，每种费米子具有其对应的 Yukawa 耦合常数  $y_f$ 。

在从第一代费米子扩充到 3 代费米子时，将拉氏量写为：

$$\mathcal{L}_{Yukawa} = - \left( \frac{v + H}{v} \right) \left[ \bar{d}'_L \mathbf{M}'_d d'_R + \bar{u}'_L \mathbf{M}'_u u'_R + \bar{l}'_L \mathbf{M}'_l l'_R + h.c. \right]. \quad (2-28)$$



$\mathbf{d}'$ 、 $\mathbf{u}'$  和  $\mathbf{l}$  为三代费米子味空间本征态， $\mathbf{M}'$  为质量混合矩阵  $(\mathbf{M}'_f)_{ij} \equiv -c_{ij}^f \frac{v}{\sqrt{2}}$ 。将质量矩阵对角化可以得到费米子的质量本征态和每个费米子的质量。在不考虑中微子质量时，中微子与希格斯场耦合常数  $y_\nu = 0$ ，则可以很容易对轻子的质量矩阵进行对角化。夸克项则必须引入一定的耦合矩阵，说明夸克的味本征态并不是质量本征态，两者之间存在混合，这一现象早已在强作用中得到证实。1973 年 Kobayashi 和 Maskawa 提出用一个  $3 \times 3$  矩阵描述 3 代夸克之间的混合，即 CKM 矩阵 (Cabibbo-Kobayashi-Maskawa Matrix) [40, 41]。这一矩阵中包含 3 个混合角和一个 CP 破坏相角  $\delta$ ，可以解释实验上观测到的 K 介子系统中的 CP 破坏效应 [42]。 $\delta$  是标准模型中唯一包含的 CP 破缺效应来源，但目前在实验中测得的  $\delta$  却如此之小，无法解释宇宙中观测到的正反物质不对称现象。

另外，20 世纪中后期开始的对中微子的研究表明 3 代中微子之间存在振荡 [43, 44]，表明中微子可能是有质量的，同时在 B 物理研究中也观测到了轻子均一性破坏 (lepton universality violation) 的迹象 ( $3.1\sigma$ ) [45]。这意味着在 Yukawa 耦合中对中微子 0 质量的假设不再适用，需要引入新的机制解释中微子质量问题 and 味混合问题。我们期待着一个更完善的新物理模型。

### 2.1.6 超出标准模型的新物理

寻找新物理需要在一定的理论框架中进行，目前在粒子物理中的常用做法是以标准模型为基础，用有效理论 (Effective theory) 来描述可能存在的高阶效应，对相关作用强度常数进行测量。在构造有效场论时，主要需要考虑三个问题：

自由度 模型自由度的选择通常是独立的，可以根据需要以及计算和实验精度进行增减。但在构造模型时，人们倾向于使用更少的自由度来描述更多的实验现象，使得模型尽量简洁。

对称性 与标准模型相同，构造的有效理论也应当满足一定对称性，以构造动力学系统。同样，对称性允许的作用项也应当被加入到有效理论模型中。

能标 考虑到现有模型应当是有效理论的近似，新理论需要规定一个能标，在此能标以下可以退化到标准模型，在能标以上新物理项能够有可观测效应。

如今发展的比较成熟的有效场论模型主要有标准模型有效场论 (SMEFT) [46] 和希格斯有效场论 (HEFT) [47, 48]。在 ATLAS 合作组中希格斯粒子相关的新物理寻找主要采用 SMEFT 模型，在考虑 SM 中的规范对称性和重子数守恒后，SMEFT 共有 59 个独立参数，包括 15 个玻色子项，19 个单费米子流项和 25 个四费米子项。对其实验测量仍是目前的重要研究方向。其他领域如中微子、QCD



物理等中对发现的超出标准模型现象另有理论模型进行描述，由于与本文和本人研究领域无关，在此不做描述。

## 2.2 希格斯物理相关实验

### 2.2.1 强子对撞中的希格斯物理

在电弱统一规范理论形成，并得到了  $W$  和  $Z$  玻色子的验证后，寻找希格斯粒子就成为了标准模型中最关键的问题。强子对撞机上希格斯粒子的产生主要有 4 种模式：胶子熔合 (Gluon Fusion, ggF)、矢量玻色子熔合 (Vector Boson Fusion, VBF)、矢量玻色子伴随产生 (VH) 和顶夸克伴随产生 (ttH、bbH、tH)。图 2-4 展示了这四种过程的费曼图，图 2-5 (左) 为各过程产生截面随  $p$ - $p$  对撞能量的变化。四种过程中 ggF 的产生截面最大，因为在质子中胶子携带了绝大多数动量。胶子本身与希格斯粒子没有耦合，这一过程主要通过重夸克 (顶夸克为主) 圈进行，是寻找希格斯粒子和测量顶夸克 Yukawa 耦合的重要过程。VBF 过程的产生截面其次，此过程中两夸克交换  $W/Z$  玻色子，并产生一个希格斯玻色子。两个夸克在反应后仍携带较大动量，在探测器中产生两个前向喷注，易于对此过程进行标记。这也是测量希格斯粒子与矢量玻色子相互作用的重要过程。

希格斯粒子可以衰变到多种末态。在理论模型中希格斯粒子与费米子之间的 Yukawa 耦合强度与费米子质量成正比，希格斯粒子与玻色子耦合强度与质量平方成正比，因此在允许的相空间内越重的粒子占据的分支比越大。图 2-5 (右) 展示了不同希格斯粒子质量假设下各衰变末态的分支比。在实验探测上，考虑到质子对撞中大量的 QCD 本底， $H \rightarrow b\bar{b}$  虽然有最大的分支比但并不容易进行测量。 $H \rightarrow \gamma\gamma$  和  $H \rightarrow ZZ^* \rightarrow 4l$  过程由于末态有非常清楚且易于测量的光子或轻子，因此虽然其分支比很小，却依然是 LHC 实验上希格斯粒子测量的黄金道。

2012 年 7 月，LHC 上的两个大型实验组 ATLAS [7] 和 CMS [8] 独立宣布在寻找标准模型希格斯粒子的过程中观测到了一个新粒子。ATLAS 使用了 LHC 2011 年  $4.8 \text{ fb}^{-1} \sqrt{s} = 7 \text{ TeV}$  和 2012 年  $5.8 \text{ fb}^{-1} \sqrt{s} = 8 \text{ TeV}$  数据，将  $H \rightarrow ZZ^* \rightarrow 4l$ 、 $H \rightarrow \gamma\gamma$ 、 $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  等多个道的观测结果联合后得到了相对于纯本底  $5.9\sigma$  的超出 (如图 2-6 (左)，local p-value)，新粒子质量为  $126.0 \pm 0.4(\text{stat.}) \pm 0.4(\text{sys.}) \text{ GeV}$ 。CMS 与 ATLAS 类似，使用了其收集到的  $5.1 \text{ fb}^{-1} 7 \text{ TeV}$  和  $5.3 \text{ fb}^{-1} 8 \text{ TeV}$  的 LHC 数据，在联合  $\gamma\gamma$ 、 $ZZ$ 、 $W^+W^-$ 、 $\tau^+\tau^-$  和  $b\bar{b}$  5 个衰变道的结果后得到了  $5.0\sigma$  显著性的超出 (如图 2-6 (右)，local p-value)，并在  $H \rightarrow \gamma\gamma$  和  $H \rightarrow ZZ$  道中拟合得到粒子质量为  $125.3 \pm 0.4(\text{stat.}) \pm 0.5(\text{sys.}) \text{ GeV}$ 。

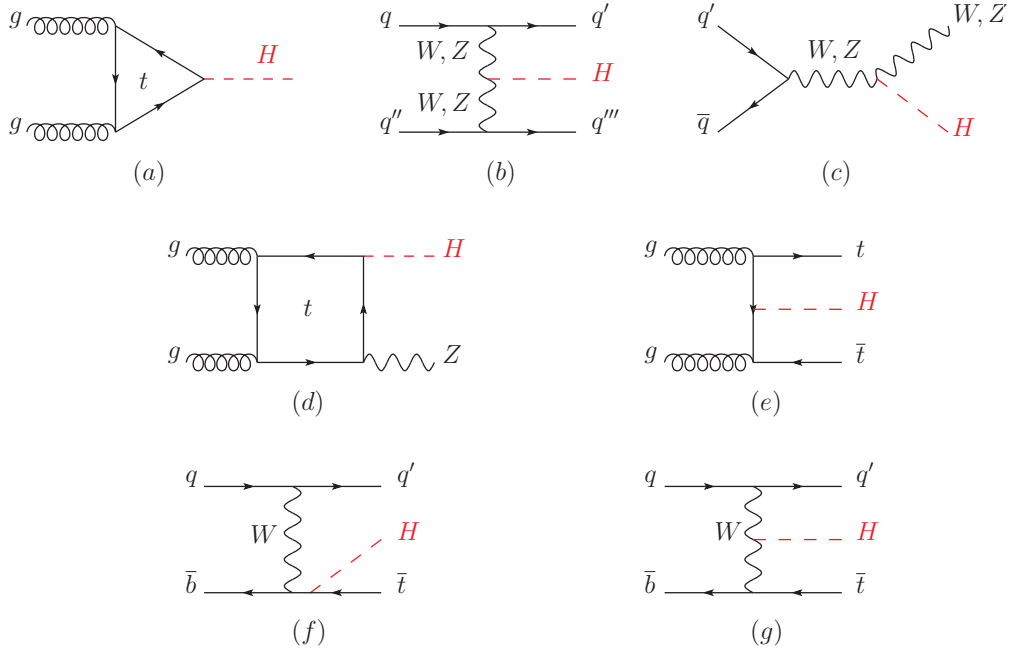


图 2-4 LHC 上产生希格斯玻色子过程的费曼图 [6]。其中：(a) 胶子熔合过程  $ggF$ 。(b) 矢量玻色子熔合过程  $VBF$ 。(c) (d) 矢量玻色子伴随过程  $VH$ 。(e)(f)(g) 顶夸克伴随产生过程  $ttH$ 、 $tH$ 。

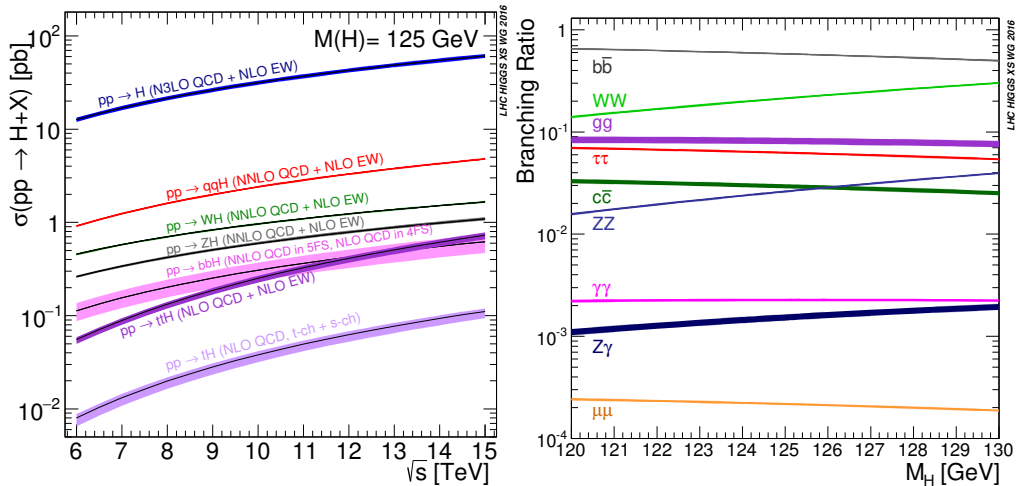


图 2-5 标准模型希格斯粒子的各过程产生截面随  $\sqrt{s}$  变化 (左) 和衰变分支比随  $m_H$  变化 (右) [6]。

两个实验上的观测结果均显示这一新粒子与标准模型中对希格斯粒子的描述相符，至此标准模型得到了最终的验证。

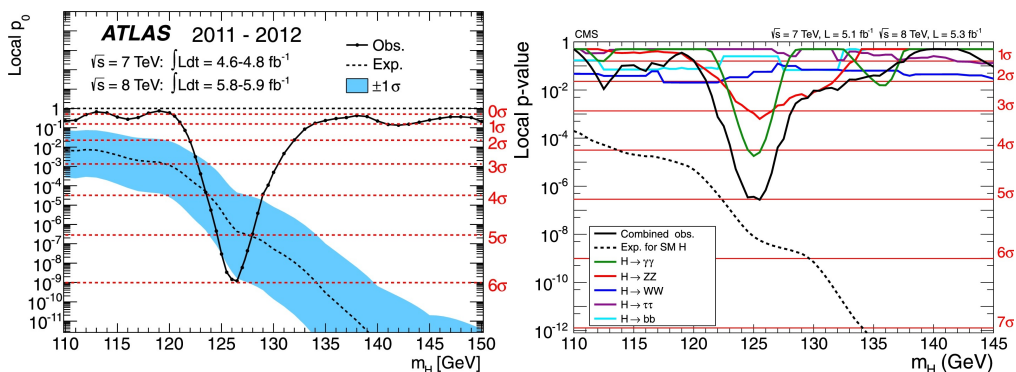


图 2-6 ATLAS (左) [7] 和 CMS (右) [8] 观测到的不同  $m_H$  假设下的 local  $p_0$  值，右图 CMS 的结果中显示了不同衰变道的贡献，其中  $H \rightarrow \gamma\gamma$  和  $H \rightarrow ZZ$  占主导。

在此之后到 2018 年 LHC 完成 Run 2 全部取数进入长时停机(Long shutdown), 与希格斯粒子相关的测量一直是 ATLAS 和 CMS 实验的工作重点。以 ATLAS 的结果为例，希格斯粒子的四个主要产生模式都分别被独立发现 [49, 50]，主要衰变道如  $H \rightarrow b\bar{b}$  [51]、 $H \rightarrow \tau^+\tau^-$  [52] 等均得到  $5\sigma$  显著性水平的证实，稀有衰变道如  $H \rightarrow c\bar{c}$  [53]、 $H \rightarrow \mu^+\mu^-$  [54]、 $H \rightarrow Z\gamma$  [55] 受限于统计量目前只能给出  $\sigma \times Br$  的上限。在完成 Run 3 取数后有望得到确切证实。希格斯粒子性质，如与费米子之间 Yukawa 耦合常数、自耦合常数  $\lambda$  的测量在简化模板微分截面 (Simplified Template Cross Section, STXS) 框架下进行 (如图 2-7)，同时通过测得截面推得有效场论 SMEFT 中超出标准模型项的耦合参数  $c_i$  如图 2-8 [9, 56]。目前的结果在误差范围内与标准模型相符，没有观测到明显超出。

### 2.2.2 轻子对撞中的希格斯物理

在 LHC 宣布发现希格斯粒子后，粒子物理的下一步计划建造轻子对撞机对希格斯粒子进行精确测量。轻子如正负电子对撞机上希格斯粒子的产生过程依赖于质心能量，在产生阈  $\sim 250$  GeV 附近以  $e^+e^- \rightarrow ZH$  过程为主，伴随有  $e^+e^- \rightarrow \nu\bar{\nu}H$  和  $e^+e^- \rightarrow e^+e^-H$  过程，其费曼图如图 2-10 所示，产生截面如图 2-9。LEP 实验中正负电子对撞的最高能量为 209 GeV [57]，仅仅稍低于这一阈值，错过了发现希格斯粒子的机会。这些过程可以提供对希格斯粒子和矢量玻色子相互作用，即电弱模型的精确测量，与 LHC 结果进行相互补充。希格斯粒子的衰变与产生相互独立，与强子对撞机上相同，但在轻子对撞机上 QCD 本底非常低，分支比最大的  $H \rightarrow b\bar{b}$ 、 $H \rightarrow gg$  等强子末态过程可以得到很好的测量。

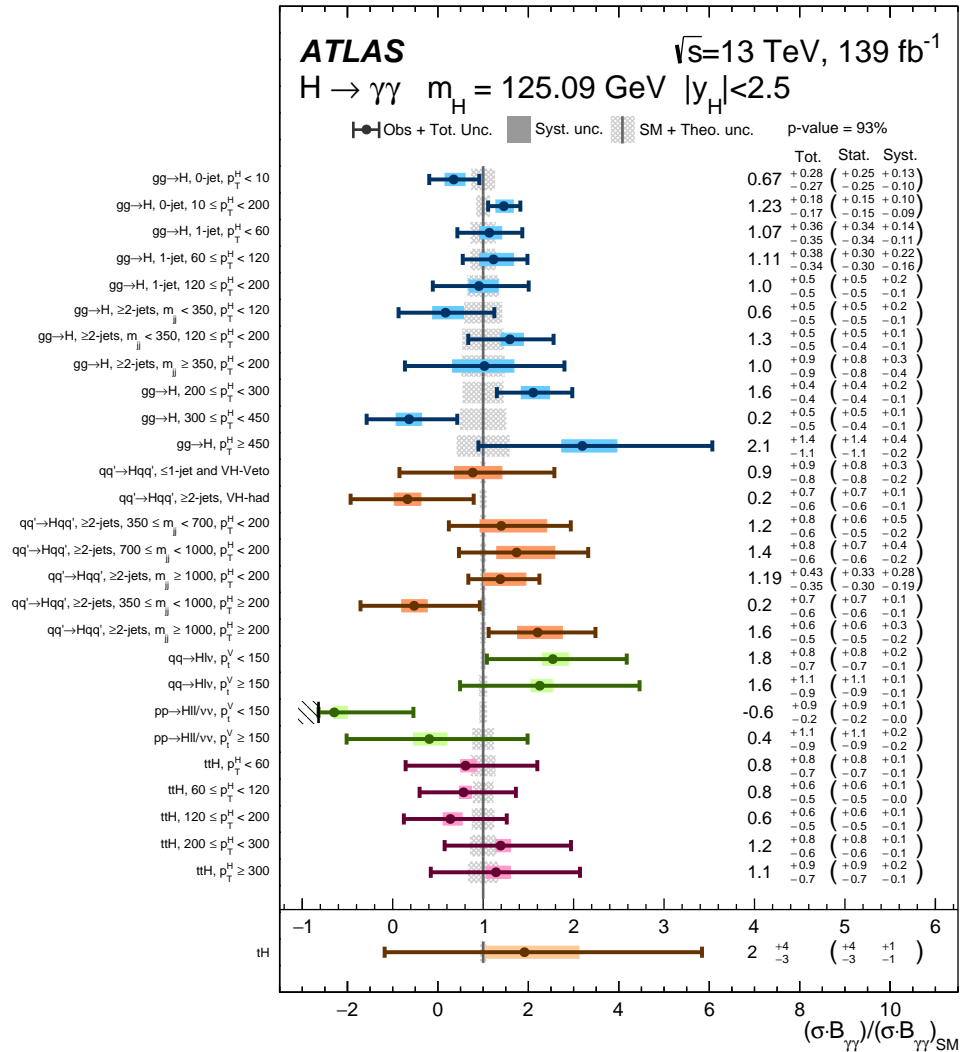


图 2-7 ATLAS 实验上 Run 2 数据在  $H \rightarrow \gamma\gamma$  道中对 STXS 参数的最佳拟合结果及误差 [9], 在误差范围内与标准模型相符。

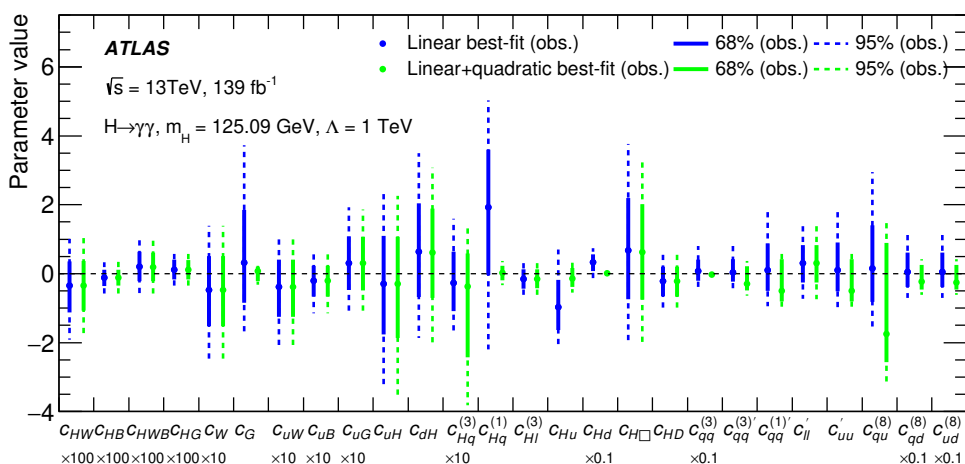


图 2-8 ATLAS 实验上 Run 2 数据在  $H \rightarrow \gamma\gamma$  道中对 SMEFT 中耦合参数  $c_i$  的测量结果及 68% 和 95% 置信区间，不在图中的 SMEFT 参数被置为 0。图中蓝线和绿线分别表示对 SMEFT 进行线性项和线性项 + 二次项的参数化结果 [9]。

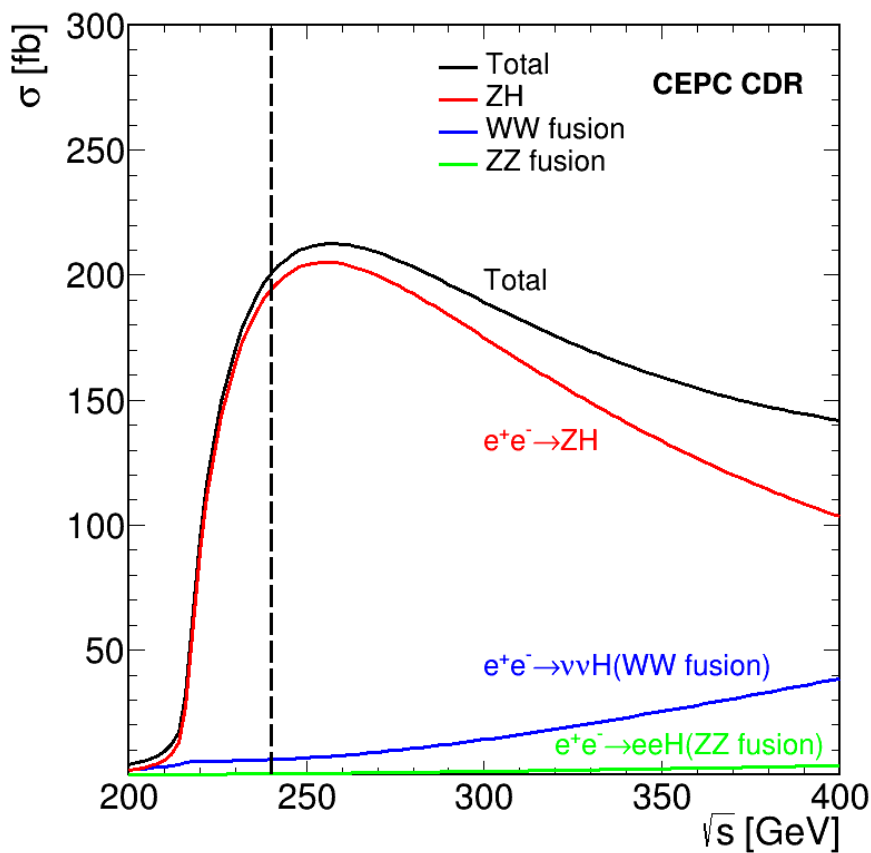


图 2-9  $e^+e^-$  对撞下希格斯粒子产生过程的截面 [10]

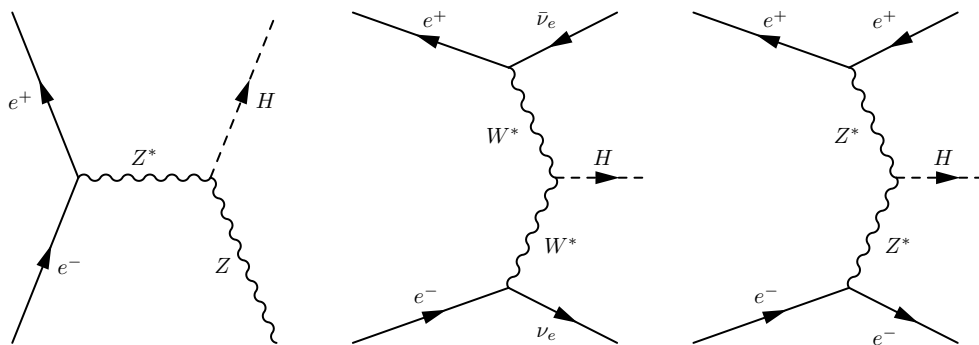


图 2-10 轻子对撞过程中希格斯粒子产生模式的费曼图,从左到右依次为: $e^+e^- \rightarrow ZH$ ,  $e^+e^- \rightarrow \bar{\nu}_e H$ ,  $e^+e^- \rightarrow e^+e^-H$ .

## 第 3 章 大型强子对撞机和 ATLAS 探测器

### 3.1 大型强子对撞机

大型强子对撞机 (Large Hadron Collider, LHC) 位于欧洲瑞士与法国边界地下 100 米深处, 是目前世界上最大的强子对撞机。其周长约 27 千米, 可进行质心能量 14 TeV 的质子-质子对撞或 2.3 TeV 的重离子 (铅核) 对撞。质子对撞的实际反应参与者为夸克或胶子对, 其携带的能量可用部分子分布函数 (Parton Distribution Function, PDF) 来进行描述, 覆盖从接近 0 到质子总能量的全部区间。因此强子对撞机中实际发生的物理过程的质心能量并不确定, 这有助于寻找对撞能标以下可发生的各种物理过程。

LHC 中的质子由氢气通过电磁场剥离电子得到。质子的加速过程如图 3-1 [58] 所示, 质子先通过直线加速器 Linac 2 加速到 50 MeV, 再被注入到周长 157 米的质子同步加速腔 (Proton Synchrotron Booster, PSB) 中加速至 1.4 TeV, 然后进入周长 628 米的质子同步加速环 (Proton Synchrotron, PS) 中, 加速至 25 GeV。接下来是周长 6.9 千米的超级质子同步加速器 (Super Proton Synchrotron, SPS), 在这里质子被加速到 450 GeV, 最后注入到 LHC 的储存环中。LHC 27 千米长的储存环上, 约 5 千米为加速单元, 包括 8 个 400 MHz 高频加速腔, 将质子加速至指定对撞能量。另外 22 千米为偏转单元, 包含 1232 个二级磁铁使束流约束在隧道中, 以及 392 个四级磁铁用于束流聚焦。储存环的两条束流管在 4 处对撞点 (Interaction Points, IPs) 进行交汇, 对应位置放置了 4 个探测器, ATLAS (A Toroidal LHC ApparatuS) [59], CMS (Compact Muon Solenoid) [60], LHCb (Large Hadron Collider beauty) [61] 和 ALICE (A Large Ion Collider Experiment) [62]。其中 ATLAS 和 CMS 为通用探测器, 覆盖对撞点周围  $4\pi$  空间, 可进行多种物理过程的测量, 主要目标为验证标准模型、寻找和测量希格斯玻色子、寻找新物理。LHCb 为前向探测器, 仅覆盖对撞点前向的部分空间, 主要进行 b 物理的研究。ALICE 是为重离子物理设计的探测器, 主要研究目标为夸克-胶子等离子体 (quark-gluon plasma)。在 ATLAS 和 CMS 的对撞点处束流被压缩至  $\beta^* = 0.30$  m 以增加对撞几率。

自 2010 年正式开机以来, LHC 已完成了两期取数工作。2010-2012 年为一期 (Run 1), LHC 对撞能量为 7/8 TeV, ATLAS 和 CMS 均收集到约  $25 \text{ fb}^{-1}$  的数据, 并基于此宣布发现了希格斯玻色子 [7, 8]。2015-2018 年为二期取数 (Run 2),

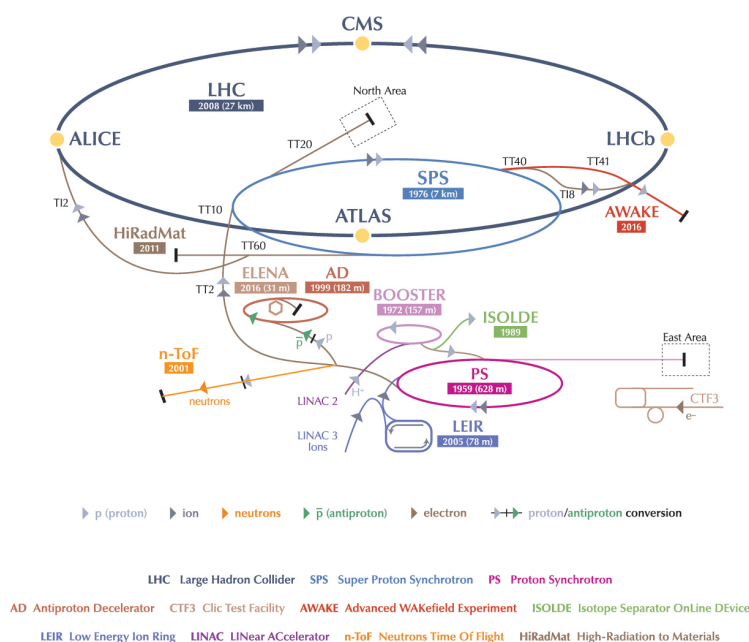


图 3-1 大型强子对撞机的加速器和储存环总揽图。

LHC 对撞能量提升至 13 TeV，ATLAS 取得  $139 \text{ fb}^{-1}$  可用于物理分析的数据，如图 3-2（左）所示。本论文中的研究基于此数据完成。2022 年 7 月，在经过 4 年的停机升级后，LHC 开始了第三期（Run 3）运行，对撞能量提升至 13.6 TeV，截至 2022 年冬季年度停机时已获取  $37.8 \text{ fb}^{-1}$  数据（如图 3-2 右）。初步的物理分析正在进行中。

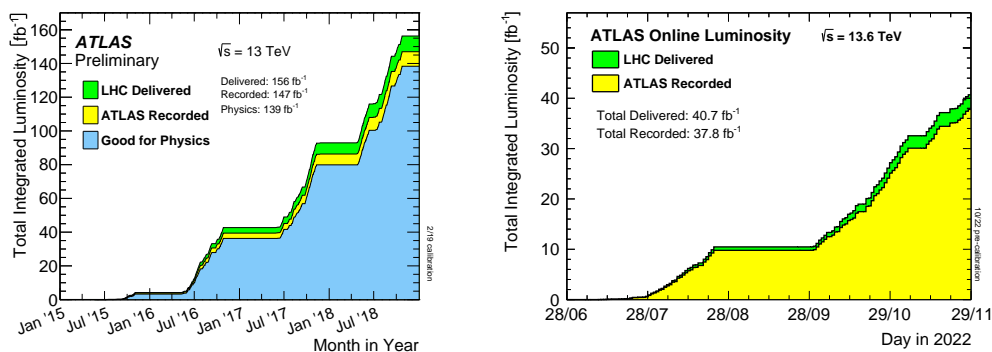


图 3-2 ATLAS 在 Run 2（左）和 2022 Run 3（右）期间收集到的对撞数据量（积分亮度） [11]。



## 3.2 ATLAS 探测器

ATLAS 是包围 LHC 束流管的圆柱形探测器，直径 25 米，长 46 米，总重约 7000 吨，主要组成部分如图 3-3 所示，有磁铁、内部径迹探测器、量能器、缪子探测器、前端探测器、触发和数据获取系统。

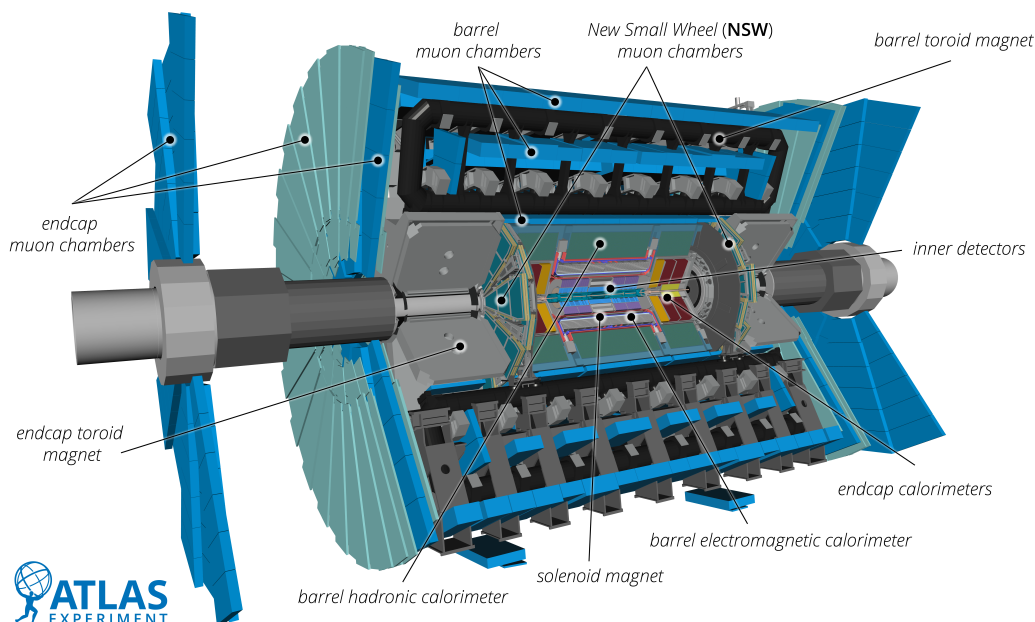


图 3-3 ATLAS 探测器示意图。

### 3.2.1 物理目标

ATLAS 的主要物理目标为寻找标准模型的希格斯玻色子，寻找超出标准模型的新物理，对 QCD、电弱、味物理和顶夸克进行精确测量。在设计之初人们并不确定希格斯玻色子的质量，根据理论预测，低质量 ( $m_H < 2m_Z$ ) 时希格斯粒子质量宽度在 MeV 量级，其强子衰变末态的喷注将难以在 QCD 本底中分辨出来，而  $H \rightarrow \gamma\gamma$  衰变道中两个干净的高能光子可以很好地重建出希格斯粒子，重建出的信号峰宽度依赖于探测器能量分辨，可以成为寻找希格斯粒子的黄金道。若  $m_H > 130 \text{ GeV}$ ， $H \rightarrow ZZ^{(*)} \rightarrow 4l$  道可以提供一个实验上非常干净的末态进行希格斯玻色子性质测量。若  $m_H > 600 \text{ GeV}$ ，则需要通过  $H \rightarrow WW/ZZ \rightarrow \text{jets}$  道寻找希格斯粒子，此时对从  $WW/ZZ$  衰变而来的前向喷注的测量将非常重要。而对于超出标准模型希格斯粒子，如最小超对称模型中的  $A$ ， $H^\pm$  的寻找，需要对包含  $\tau$  轻子和 b-jet 过程的高敏感度。

除希格斯粒子以外，新的重规范玻色子  $W'$ ， $Z'$  的质量可能达到  $\sim 6 \text{ TeV}$ ，其轻子衰变产物的动量或强衰变喷注的横动量可达到 TeV 量级。另一方面，寻

找味改变中性流 (Flavor-changing neutral currents, FCNC)、轻子数破缺 (lepton flavor violation, LFV) 也可能是开启新物理的窗口。因此探测器需要对大动量范围的带电粒子有很高的敏感度。

另外, 根据相关理论模型, 超对称粒子、引力子等新物理模型等产生与衰变通常会伴随产生不可见粒子, 实验上即丢失横动量  $E_T^{\text{miss}}$ 。这就需要探测器能做到尽可能大的空间覆盖率, 以实现对其  $E_T^{\text{miss}}$  的重建。

LHC 的设计瞬时亮度为  $10^3 4\text{cm}^{-2}\text{s}^{-1}$ , 这意味着每秒钟有  $10^9$  个非弹性散射事例, 同时  $p-p$  对撞中也会产生大量 QCD 喷注。对于物理分析来说这些过程是重要本底, 需要有高效的事例重建和鉴别能力来将其去除。

基于这些目标, LHC 上的探测器应满足如下要求 (如表 3-1):

- 响应时间快、抗辐照性能强, 同时探测器颗粒度高;
- 全空间覆盖, 并对前向 (大  $\eta$ ) 事例有很高接受度和灵敏度;
- 内部径迹室有很好的带电粒子动量分辨和重建效率, 顶点探测器需要能够分辨出次级顶点以满足对  $\tau$  轻子、b-jet 等事例的离线鉴别;
- 电磁量能器需要对电子和光子有很好的测量和鉴别能力, 强子量能器对喷注和丢失横动量也应有精确测量;
- 对大动量范围的缪子有很好的鉴别能力;
- 对低动量对象有高效的触发系统, 保证能够去除足够的本底, 以实现可接受的触发事例率。

子探测器	分辨率要求	$\eta$ 覆盖范围	
		测量	触发
径迹探测器	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	$\pm 2.5$	
电磁量能器	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$\pm 3.2$	$\pm 2.5$
强子量能器			
桶部和端盖	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\pm 3.2$	$\pm 3.2$
前端	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 <  \eta  < 4.9$	$3.1 <  \eta  < 4.9$
缪子探测器	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1 \text{ TeV}$	$\pm 2.7$	$\pm 2.4$

表 3-1 ATLAS 探测器各部分的设计指标。表中  $E$  和  $p_T$  单位为 GeV。

### 3.2.2 磁铁

ATLAS 的磁铁系统由一个螺线管型超导磁铁和 3 个环形线圈磁铁组成，其中一个桶部线圈两个端盖线圈，如图 3-4 所示。中央螺线管内直径 2.46 m，外直径 2.56 m，轴向长 5.8 m，为内部的径迹探测器提供 2 T 的轴向磁场，并把辐射长度控制在  $0.66 X_0$ 。桶部线圈磁铁由 8 个线圈组成，封装在单独的跑道型不锈钢真空容器中。其内直径 9.4 m，外直径 20.1 m，长 25.3 m，为内部缪子探测器提供约 0.5 T 的环形磁场。两个端盖线圈磁铁放置在探测器前端，同样各由 8 个线圈组成，内直径 1.65 m，外直径 10.7 m，长 5 m，为端盖缪子探测器提供 1 T 的偏转磁场。

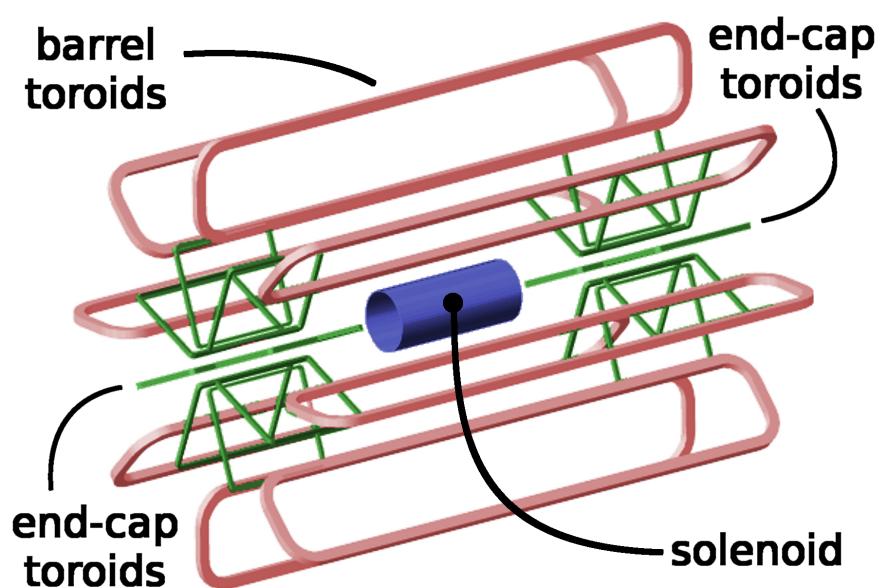


图 3-4 ATLAS 磁铁系统示意图。

### 3.2.3 内部径迹探测器

ATLAS 的内部径迹探测器（Inner Detector, ID）位于整个探测器最内层，由硅像素探测器（Pixel）、硅微条探测器（silicon microstrip layer, SCT）和穿越辐射探测器（transition radiation tracker, TRT）三部分组成，提供高分辨率的带电粒子动量测量、事例的初级和次级顶点的重建。整体如图 3-5 所示，参数如表 3-2。

最内层的硅像素探测器是最接近束流管的探测器，覆盖  $|\eta| < 2.58$  的区域，共有 4 层：一个插入式 B 层（Insertable B-layer, IBL）[63] 和 3 个硅像素层 [64]。探测器灵敏单元的尺寸为  $50 \times 400 \mu\text{m}^2$ ，位置分辨为桶部（端盖） $R - \phi$  方向  $10 \mu\text{m}$ ， $z$  ( $R$ ) 方向  $115 \mu\text{m}$ 。IBL 层是在 2013-2014 年 LHC 第一次长停机（Long Shutdown 1）时安装进 ATLAS 的，放置在原先硅像素层的内部，距束流管 33.2

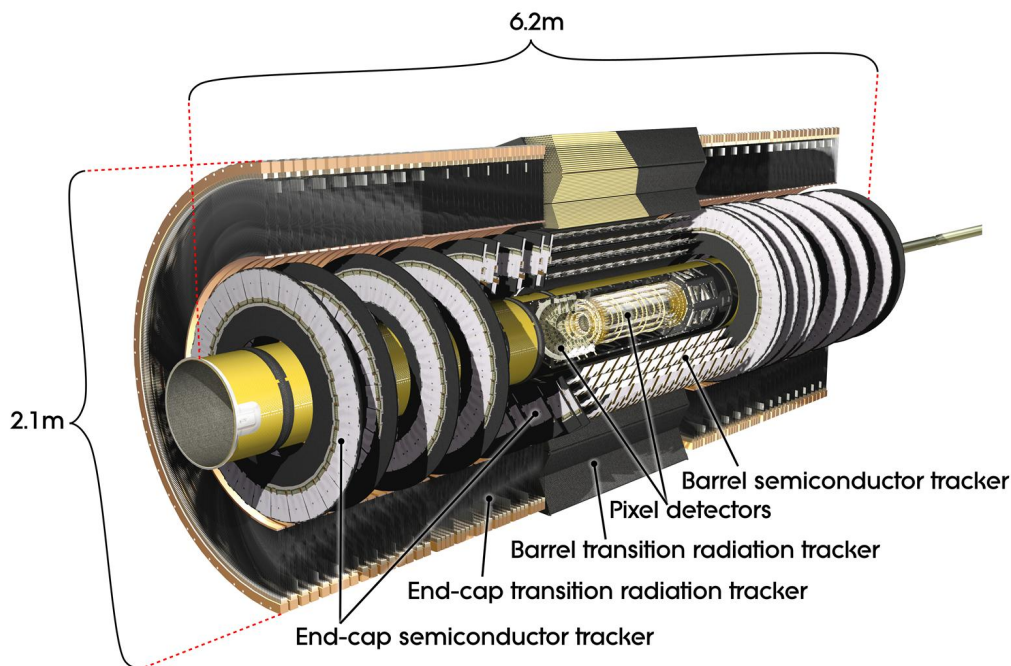


图 3-5 ATLAS 内部径迹探测器。

mm，像素尺寸  $250 \times 50 \mu\text{m}^2$ 。这一层的加入有助于 ATLAS 去除 LHC 亮度提升带来的堆积效应 (pile-up)，减缓随运行时间而来的内层径迹室坏道问题，提高径迹和顶点测量的精度，从而提升对 b-jet 相关物理的灵敏度。

硅微条探测器在硅像素探测器外围，桶部有 8 层，两层间小角度 ( $40 \text{ mrad}$ ) 交错排列，提供 4 个空间点的测量；两侧端盖各 9 层，同样以  $40 \text{ mrad}$  交错排列提供空间点定位。像素条带尺寸均为  $80 \mu\text{m} \times 6.4 \text{ mm}$ ，本征位置分辨  $R - \phi$  方向  $10 \mu\text{m}$ ， $z$  ( $R$ ) 方向  $580 \mu\text{m}$ 。SCT 总读出通道数约 630 万个。

带电粒子主要径迹测量由穿越辐射探测器 TRT 完成，平均每个带电径迹会留下 30 个测量点，最多 36 个。穿越辐射探测器由 37 万根稻草管 (漂移管) 组成，桶部每根稻草管长 1.44 m，直径 4 mm，管内填充 70% 氙气、27% 二氧化碳和 3% 氧气，管中心由一根直径  $30 \mu\text{m}$  的钨丝作为阳极丝。桶部 TRT 仅提供  $R - \phi$  方向的定位，精度  $130 \mu\text{m}$ 。端盖稻草管长 37 cm，共 160 层，沿径向排列，组成与桶部相同。电子在穿过包含氙气的稻草管时会发生穿越辐射，TRT 可探测到这一信号，从而具有电子鉴别能力。

### 3.2.4 量能器

对撞产生的粒子在穿过内部径迹探测器后将进入量能器，沉积绝大部分能量。ATLAS 的量能器分为电磁量能器和强子量能器两部分，如图 3-6 所示。整个

组分		径向尺寸 [mm]	轴向长度 [mm]
<b>ID 总体</b>		$0 < R < 1150$	$0 <  z  < 3512$
<b>IBL</b>		$\langle R \rangle = 25.7$	$ z  < 332$
<b>硅像素</b>			
3 层圆柱层	桶部	$50.5 < R < 122.5$	$ z  < 400.5$
2×3 圆盘层	端盖	$88.8 < R < 149.6$	$495 <  z  < 650$
<b>SCT</b>			
4 层圆柱层	桶部	$299 < R < 514$	$ z  < 749$
2×9 圆盘层	端盖	$275 < R < 560$	$839 <  z  < 2735$
<b>TRT</b>			
73 稻草管	桶部	$563 < R < 1066$	$ z  < 712$
160 稻草管	端盖	$644 < R < 1004$	$848 <  z  < 2710$

表 3-2 ATLAS 内部径迹探测器各部分的尺寸。

量能器覆盖区间为  $|\eta| < 4.9$ ，提供精确的可见粒子的能量测量以及  $E_T^{\text{miss}}$  测量。

#### 3.2.4.1 电磁量能器

ATLAS 电磁量能器由液氙作为灵敏层、铅为吸收体的取样型量能器，吸收体铅板呈手风琴型，由蜂窝状铜结构分隔开，中间灌注液氙。这一结构可以提供纵向颗粒度信息，并确保探测器在  $\phi$  向对称。几何上电磁量能器可分为桶部和端盖两个部分。桶部 ( $|\eta| < 1.475$ ) 厚度约 22 倍辐射长度 ( $X_0$ )，分成两个相同的部分，中间在  $z = 0$  处有一个 4 mm 的缝隙。纵向上分为 3 层 (strip layer, middle layer, back layer)，第一层厚度约  $4 X_0$ ，颗粒度  $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$  ( $|\eta| < 1.4$ ) 或  $0.025 \times 0.025$  ( $1.4 < |\eta| < 1.475$ )，可提高  $\eta$  向粒子分辨能力，如  $\gamma/\pi^0$ ， $e/\pi$  分辨，而  $\phi$  向由于磁场的存在，带电粒子可被较好地分开；第二层厚度约  $16 X_0$ ，颗粒度  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$  ( $|\eta| < 1.4$ ) 或  $0.075 \times 0.025$  ( $1.4 < |\eta| < 1.475$ ) 第三层厚度约  $2 X_0$ ，颗粒度  $\Delta\eta \times \Delta\phi = 0.050 \times 0.025$ 。每个探测单元都指向对撞点。

端盖量能器分为两个同轴轮，外部轮覆盖  $1.375 < |\eta| < 2.5$ ，内部轮覆盖范围  $2.5 < |\eta| < 3.2$ ，中间有一个 3 mm 宽的间隙。外部轮与桶部相同，在纵向上分为 3 层，内部轮为 2 层，颗粒度随  $\eta$  变化。整个探测器的铅吸收体厚度随  $\eta$  变化，以实现最优的能量分辨。



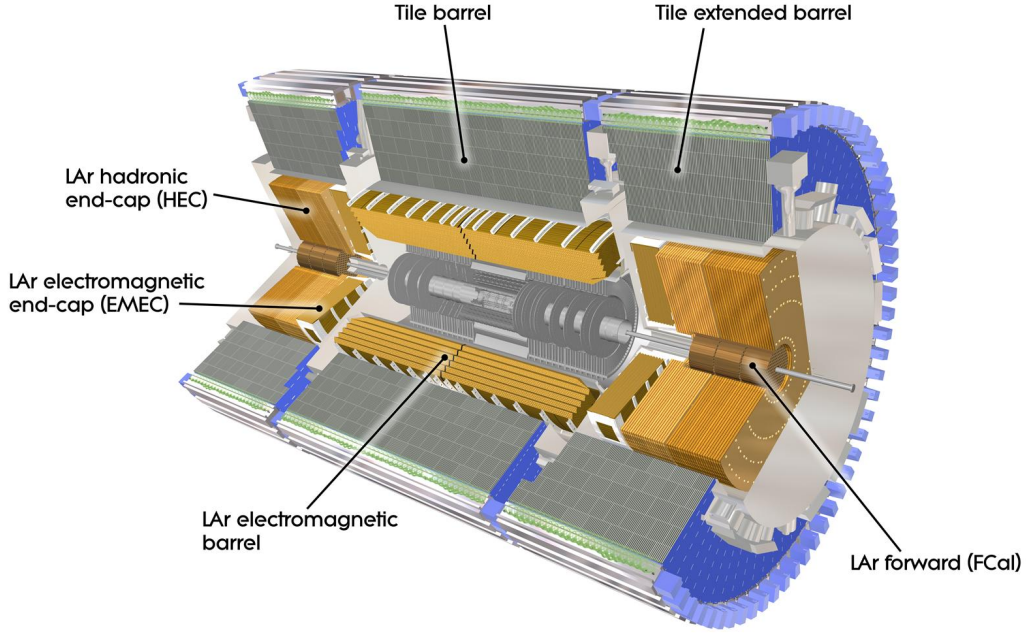


图 3-6 ATLAS 量能器结构示意图。

在电磁量能器前端还有一个预采样液氩探测器 (presampler LAr detector, PS), 覆盖范围  $|\eta| < 1.8$ , 厚度 0.5 mm (端盖) - 1.1 mm (桶部), 使得  $e/\gamma$  电磁簇射可以提前发生, 并完全沉积在量能器中。同时也可以用来修正量能器前泄漏的能量。

在安装前的束流测试中 [65, 66], ATLAS 电磁量能器的能量分辨约为  $\sigma_E/E = \frac{10\%}{\sqrt{E [\text{GeV}]} } \oplus 0.6\%$ 。

#### 3.2.4.2 强子量能器

穿过电磁量能器的强子会在强子量能器中发生簇射沉积能量。ATLAS 的强子量能器采用两种不同的结构, 中间 ( $|\eta| < 1.7$ ) 为瓦片型 (Tile) 量能器, 采用塑料闪烁体作为灵敏物质, 不锈钢作为吸收体, 径向厚度约 7.4 倍辐射长度。它又被分为桶部 (barrel,  $|\eta| < 1.0$ ) 和外缘 (extended barrel,  $0.8 < |\eta| < 1.7$ ) 两部分, 两者之间有一个 68 cm 的间隙用于引出内部探测器和电磁量能器的信号和供电线缆等, 同时也覆盖了一块三明治结构瓦片量能器来减少能量泄漏。瓦片量能器纵向上同样分为 3 层, 前两层颗粒度  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ , 最后一层为  $0.2 \times 0.1$ 。

强子量能器端盖部分 (Hadronic end-cap calorimeter, HEC,  $1.5 < |\eta| < 3.2$ ) 则采用液氩结构, 以铜为吸收体, 液氩为灵敏物质。每侧端盖由前后两个同轴轮组成, 轮外半径 2.03 m, 包含 32 个楔形模块。前轮的吸收体是 24 块 25 mm 厚的铜板, 以及一块 12.5 mm 厚的前端版, 后轮为 16 块 50 mm 厚的铜板和一块 25

mm 厚的后端版。在  $|\eta| < 2.5$  区间探测器的颗粒度为  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ ，更外部则为  $0.2 \times 0.2$ ，探测器单元被安排成指向对撞点的排列。

更外部是前向量能器 (Forward Calorimeter, FCal)，同样为液氩 + 吸收体结构，距离对撞点 4.7 m，覆盖  $3.1 < |\eta| < 4.9$  范围。第一层用铜做吸收体，主要用来探测电磁簇射成分；后两层用钨做吸收体，主要探测强子簇射。

强子量能器各部分在安装前的束流测试中给出的  $\pi$  能量分辨分别为：瓦片量能器  $\sigma_E/E = \frac{56.4\%}{\sqrt{E [\text{GeV}]} } \oplus 5.5\%$  [67]，端盖量能器  $\sigma_E/E = \frac{84.6\%}{\sqrt{E [\text{GeV}]} } [68]$ ，前向量能器  $\sigma_E/E = \frac{94.2\%}{\sqrt{E [\text{GeV}]} } \oplus 7.5\%$  [69]。

### 3.2.5 缪子探测器

缪子探测器 (Muon Spectrometer, MS) 位于 ATLAS 的最外层，用于探测逃逸出量能器的带电粒子，以缪子为主。MS 由 3.2.2 中提到的两套线圈提供磁场使带电粒子偏转，从而进行动量测量。 $|\eta| < 1.4$  范围的磁场由桶部线圈产生， $1.6 < |\eta| < 2.7$  的磁场由两侧的端盖线圈产生，中间  $1.4 < |\eta| < 1.6$  被称为过渡区 (transition region)，其磁场由桶部和端盖线圈一同产生。这套磁铁系统保证了磁场方向与缪子径迹方向正交，同时最小化多次散射带来的动量分辨下降。根据用途不同，缪子探测器被大致分为两个部分：漂移管室 (Monitored Drift tubes, MDT) 和阴极条室 (Cathode strip chambers, CSC) 提供中央和大  $\eta$  区域径迹的精确测量；阻性版室 (Resistive plate chambers, RPC) 和窄间隙室 (Thin gap chambers, TGC) 分别在桶部和端盖区作为 ATLAS 触发系统的一部分 (L1Muon)。各部分参数在表 3-3 中。

### 3.2.6 触发、读出、数据获取系统 [1, 2]

LHC 的设计亮度下，平均事例产生率为 40 MHz，需要通过触发系统将这一事例率降低至可接受水平再进行写出。在 Run 2 中 ATLAS 的触发系统分为两级：一级触发 (Level-1 trigger, L1)，和高级触发 (High Level Trigger, HLT)。一级触发通过电子学实现，HLT 几乎完全通过软件算法实现。L1 主要靠高  $p_T$  缪子、高能电子、光子、 $E_T^{\text{miss}}$ 、喷注和  $\tau$  强衰变产物的特征来鉴别事例，使用电磁量能器、RPC、TGC 的粗略位置信息 (Region of interest, RoI)，前者被称为 L1Calo，后者称为 L1Muon，另有 L1Topo 综合这两者对信息进行判断。一级触发在对撞发生后  $2.5 \mu\text{s}$  内判断是否读取事例，并将事例率压缩到 100 kHz。通过选择的事例信息被储存在读出系统 (Readout System, ROS) 中。HLT 用更全面的探测器信息对选出的事例进行组装和筛选，以 L1Calo 中的单元作为种子在  $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$

<b>监控漂移管</b>	<b>MDT</b>
- 覆盖范围	$ \eta  < 2.7$ (innermost layer: $ \eta  < 2.0$ )
- 室数目	1150
- 读出通道数	354 k
- 功能	径迹测量
<b>阴极条室</b>	<b>CSC</b>
- 覆盖范围	$2.0 <  \eta  < 2.7$
- 室数目	32
- 读出通道数	31 k
- 功能	径迹测量
<b>阻性板室</b>	<b>RPC</b>
- 覆盖范围	$ \eta  < 1.05$
- 室数目	606
- 读出通道数	359 k
- 功能	触发, 符合
<b>薄间隙室</b>	<b>TGC</b>
- 覆盖范围	$1.05 <  \eta  < 2.7$ (触发: $ \eta  < 2.4$ )
- 室数目	3588
- 读出通道数	318 k
- 功能	触发, 符合

表 3-3 ATLAS 缪子探测器参数和功能。



范围内构建簇团，读取 ID 中的击中信息进行快寻迹得到带电粒子径迹，综合全量能器信息重建和鉴别电子、光子、缪子、喷注、b-jet 等物理对象，以或逻辑进行事例的筛选。最终经过 HLT 存盘的事例率约为 1 kHz。流程总结如图 3-7。

在事例通过 L1 被传输给 HLT 的同时，数据获取系统（Data acquisition system, DAQ）也将探测器前端电子学系统中的信号通过读出驱动（Readout Drivers, RODs）传送到了读出系统（Readout System, ROS）中。这些数据被暂时放在 ROS 中供 HLT 使用，通过 HLT 事例的事例即被存盘。

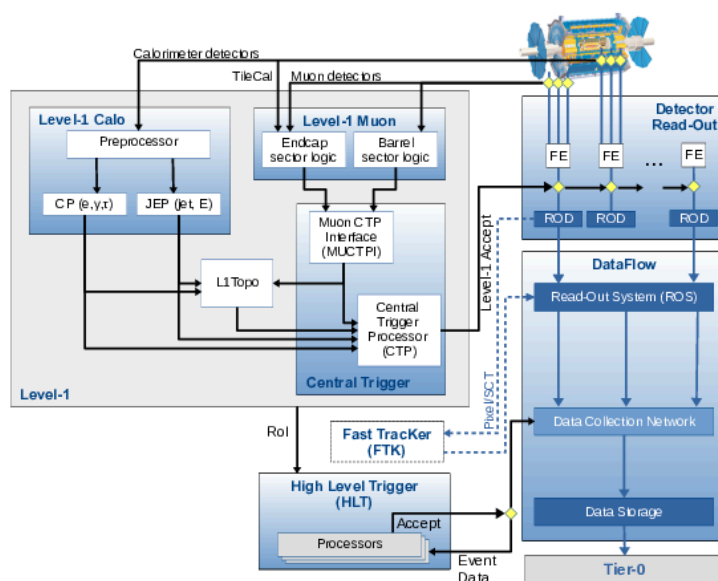


图 3-7 ATLAS 触发系统流程图。

### 3.3 ATLAS 探测器上的粒子重建和鉴别

LHC 上质子-质子对撞产生的末态粒子会在穿过 ATLAS 探测器时与探测器物质发生反应，留下电子学信号。这些信号在离线软件中被重建为不同物理对象以供物理分析。

#### 3.3.1 光子和电子 [3]

ATLAS 上光子和电子的重建类似，主要使用电磁量能器的沉积能量信息，电子的重建和鉴别需要额外的内部径迹室中的径迹信息。光子和电子在进入电磁量能器后通过电磁簇射过程在灵敏层（液氩）中沉积能量形成簇团（cluster），

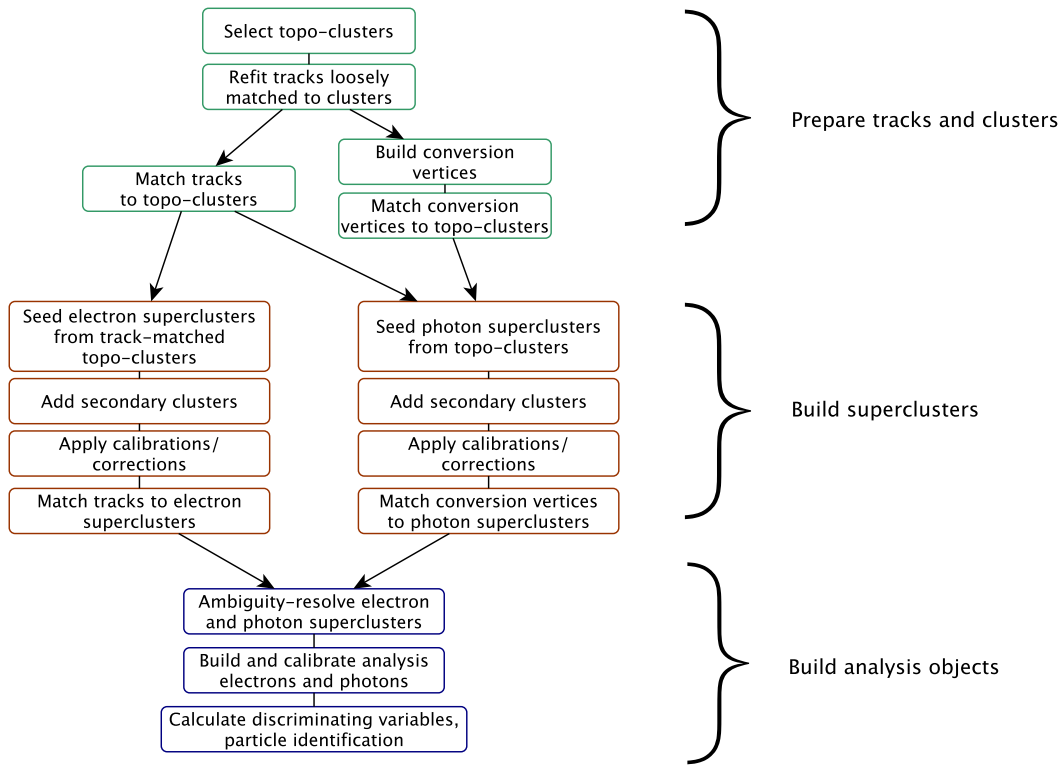


图 3-8 电子和光子的重建流程。

被探测器电子学转化为电荷输出，后期软件对电荷进行一系列刻度和修正后可以得到入射粒子的能量。同时探测器单元位置可以给出簇射形状，用于对重建出的粒子进行鉴别。电子在进入量能器前会通过电离过程在径迹室留下击中点，并在磁场的作用下有一定偏转，因此需要有来自对撞点的径迹与量能器中的簇团匹配。光子在进入量能器前有一定几率与内部物质发生作用转化成正负电子对，留下不从对撞点产生的一到两条径迹，因此光子的重建分为转化光子（converted photon）和非转化光子（unconverted photon）两部分。具体重建流程如图 3-8。

### 3.3.1.1 光子和电子的重建

重建算法首先构造量能器中的拓扑簇团（topo-cluster）。探测器中信号显著度  $|\zeta_{cell}^{EM}| \equiv \left| \frac{E_{cell}^{EM}}{\sigma_{noise,cell}^{EM}} \right| \geq 4$  的单元被选作簇团的种子，以此为中心向外围延伸，直到单元的  $|\zeta_{cell}^{EM}| < 2$ 。其中  $E_{cell}^{EM}$  是单元中沉积能量， $\sigma_{noise,cell}^{EM}$  是预期噪声水平，包含了电子学噪声和根据对装亮度对 pile-up 噪声的估计。如果两个簇团包含了同一个  $|\zeta_{cell}^{EM}| \geq 2$  的单元，那这两个簇团会被合并为一个。这一套重建流程通常被称为“4-2-0”拓扑簇团重建。

电子和转化光子重建过程所需的径迹由内部径迹室中的击中重建而成。硅像素探测器探测器中的每个击中点提供一个独立的三维空间信息，SCT 则将两层硅条的信息进行合并后得到一个空间点坐标。硅探测器中三个空间点组成一个径迹种子 (seed)，以此为基础采取  $\pi^\pm$  假设用模式识别算法向外层扩展，将扩展得到的径迹进行  $\chi^2$  拟合 (ATLAS Global  $\chi^2$  Track Fitter [70])，去除共享击中导致的重复径迹。对于没有通过  $\chi^2$  拟合但方位角与电磁量能器中簇团接近的径迹，再考虑韧致辐射效应使用正负电子假设来进行拟合，以提高效率。得到初步径迹后再利用卡尔曼滤波方法 ([71], [72]) 对径迹进行拟合和拓展，修正带电粒子在穿过探测器时物质效应导致的能量损失。最终综合利用硅像素探测器、SCT 和 TRT 信息重建出径迹参数。

如果两条电荷相反的径迹可以形成一个顶点，且对应于一个无质量粒子，则这个顶点可以作为转化光子顶点 (conversion vertex) 的备选。没有最内层击中的径迹也可以作为单转化径迹，最靠近束流的击中点作为转化顶点。为提高转化光子的纯度，选择的径迹应当被 TRT 或硅探测器判定为是电子径迹。

在得到拓扑簇团和径迹后，可以利用方位角条件将簇团和径迹进行匹配。未匹配上径迹且横向能量  $E_T$  高于 1.5 GeV 的簇团作为非转化光子的超簇团 (supercluster) 种子，把周围  $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$  的范围内的拓扑簇团作为卫星簇团合并其中；匹配上来自对撞点的径迹且  $E_T > 1$  GeV 的拓扑簇团被用作为电子的超簇团种子，在合并周围卫星簇团时范围扩展到  $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ ，以考虑电子韧致辐射产生的光子；如果簇团能够与转化顶点相匹配，则作为转化光子超簇团种子，匹配到同一转化顶点的簇团都被合并到一起。通过这些方法得到的超簇团被重建为电子或光子。为了降低超簇团对 pile-up 噪声的灵敏度，超簇团中的每个拓扑簇团  $\eta$  方向的尺寸被限制在 0.075 (桶部) 或 0.125 (端盖)。内部探测器中的磁场方向沿束流管，电子和光子与探测器物质的作用通常只会导致电磁簇射在  $\phi$  方向上的延展，因此这一  $\eta$  方向上的限制不会造成能量损失。

### 3.3.1.2 光子和电子的鉴别

在完成超簇团的重建后，我们可以构造一些变量来描述超簇团的形状，这些簇射形状变量 (shower shape variables) 可以用来进行光子和电子的鉴别。这可以有效去除对其他粒子的误判，如喷注中的  $\pi^0$ 。构造出的变量具体如表 3-4。

电子的鉴别使用似然函数法，根据使用的  $n$  个区分变量的信号与本底概率

Category	Description	Name	Usage
Hadronic leakage	Ratio of $E_T$ in the first layer of the hadronic calorimeter to $E_T$ of the EM cluster (used over the ranges $ \eta  < 0.8$ and $ \eta  > 1.37$ )	$R_{\text{had}_1}$	$e/\gamma$
	Ratio of $E_T$ in the hadronic calorimeter to $E_T$ of the EM cluster (used over the range $0.8 <  \eta  < 1.37$ )	$R_{\text{had}}$	$e/\gamma$
EM third layer	Ratio of the energy in the third layer to the total energy in the EM calorimeter	$f_3$	$e$
EM second layer	Ratio of the sum of the energies of the cells contained in a $3 \times 7 \eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a $7 \times 7$ rectangle, both centred around the most energetic cell	$R_\eta$	$e/\gamma$
	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$ , where $E_i$ is the energy and $\eta_i$ is the pseudorapidity of cell $i$ and the sum is calculated within a window of $3 \times 5$ cells	$w_{\eta_2}$	$e/\gamma$
	Ratio of the sum of the energies of the cells contained in a $3 \times 3 \eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a $3 \times 7$ rectangle, both centred around the most energetic cell	$R_\phi$	$e/\gamma$
EM first layer	Total lateral shower width, $\sqrt{(\sum E_i (i - i_{\text{max}})^2)/(\sum E_i)}$ , where $i$ runs over all cells in a window of $\Delta\eta \approx 0.0625$ and $i_{\text{max}}$ is the index of the highest-energy cell	$w_{s, \text{tot}}$	$e/\gamma$
	Lateral shower width, $\sqrt{(\sum E_i (i - i_{\text{max}})^2)/(\sum E_i)}$ , where $i$ runs over all cells in a window of 3 cells around the highest-energy cell	$w_{s, 3}$	$\gamma$
	Energy fraction outside core of three central cells, within seven cells	$f_{\text{side}}$	$\gamma$
	Difference between the energy of the cell associated with the second maximum, and the energy reconstructed in the cell with the smallest value found between the first and second maxima	$\Delta E_s$	$\gamma$
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies	$E_{\text{ratio}}$	$e/\gamma$
Track conditions	Ratio of the energy measured in the first layer of the electromagnetic calorimeter to the total energy of the EM cluster	$f_1$	$e/\gamma$
	Number of hits in the innermost pixel layer	$n_{\text{innermost}}$	$e$
	Number of hits in the pixel detector	$n_{\text{Pixel}}$	$e$
	Total number of hits in the pixel and SCT detectors	$n_{\text{Si}}$	$e$
	Transverse impact parameter relative to the beam-line	$d_0$	$e$
	Significance of transverse impact parameter defined as the ratio of $d_0$ to its uncertainty	$ d_0/\sigma(d_0) $	$e$
	Momentum lost by the track between the perigee and the last measurement point divided by the momentum at perigee	$\Delta p/p$	$e$
Likelihood probability based on transition radiation in the TRT	eProbabilityHT	$e$	
Track-cluster matching	$\Delta\eta$ between the cluster position in the first layer of the EM calorimeter and the extrapolated track	$\Delta\eta_1$	$e$
	$\Delta\phi$ between the cluster position in the second layer of the EM calorimeter and the momentum-rescaled track, extrapolated from the perigee, times the charge $q$	$\Delta\phi_{\text{res}}$	$e$
	Ratio of the cluster energy to the measured track momentum	$E/p$	$e$

表 3-4 电子与光子鉴别用到的变量。

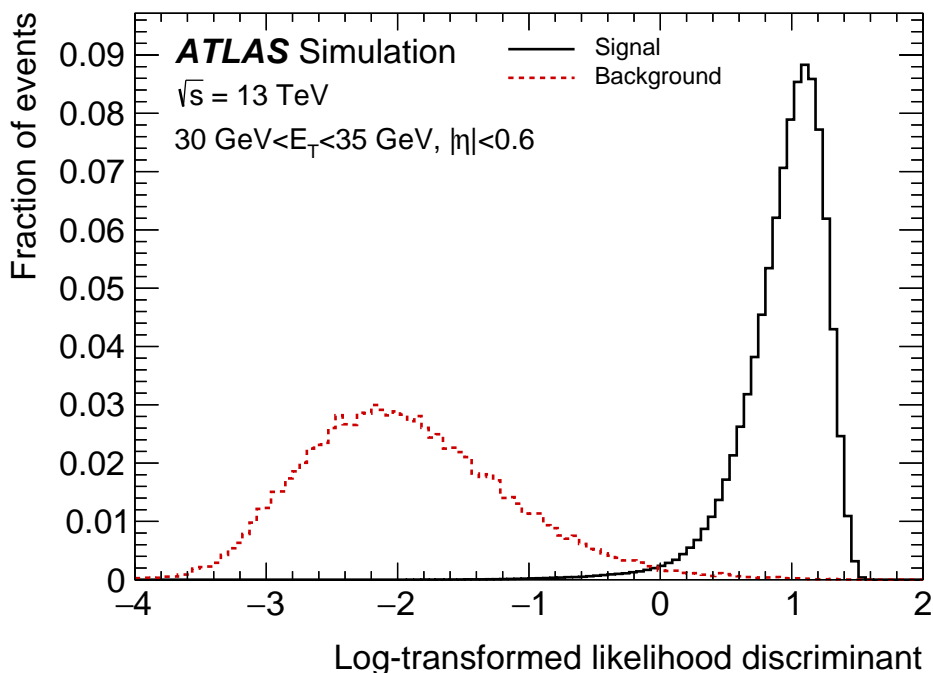


图 3-9 电子鉴别中使用的似然函数值分布图 [12]。

分布  $P_{S(B)}(\mathbf{x})$  构造似然函数值  $L_{S(B)}(\mathbf{x})$ :

$$L_{S(B)}(\mathbf{x}) = \prod_i^n P_{S(B),i}(x_i), \quad (3-1)$$

其中  $\mathbf{x}$  是使用的区分变量，变量之间的相关性不考虑。得到  $L_{S(B)}$  后可构造  $d_L = \ln(L_B/L_S)$  用于判定待选粒子，其分布如图 3-9。电子鉴别根据对信号和本底的区分能力分成了三个工作点：宽松，中等和严格。三个条件下的效率随  $E_T$  和  $\eta$  如图 3-10。

光子的鉴别则直接使用簇射形状变量进行筛选，并同样分为宽松、中等、严格三个工作点。其中宽松和中等工作点通常用于光子和双光子事例的触发，且不区分转化光子和非转化光子。严格工作点的条件用 TMVA [73] 进行了优化，且对于转化光子和非转化光子、不同  $E_T$  和  $\eta$  区间分别进行了研究。ATLAS Run 2 中使用的严格光子鉴别效率如图 3-11。

ATLAS 探测器实际取数时观测到数据和蒙特卡洛模拟中这些用于光子和电子鉴别的簇射形状变量并不完全一致，图 3-12 以  $R_\phi$  和  $f_3$  为例展示了其分布。这一误差可通过修正因子 (fudge factor) 进行描述，对模拟样本修正后再进行相关研究，以减小在实际分析中的影响。

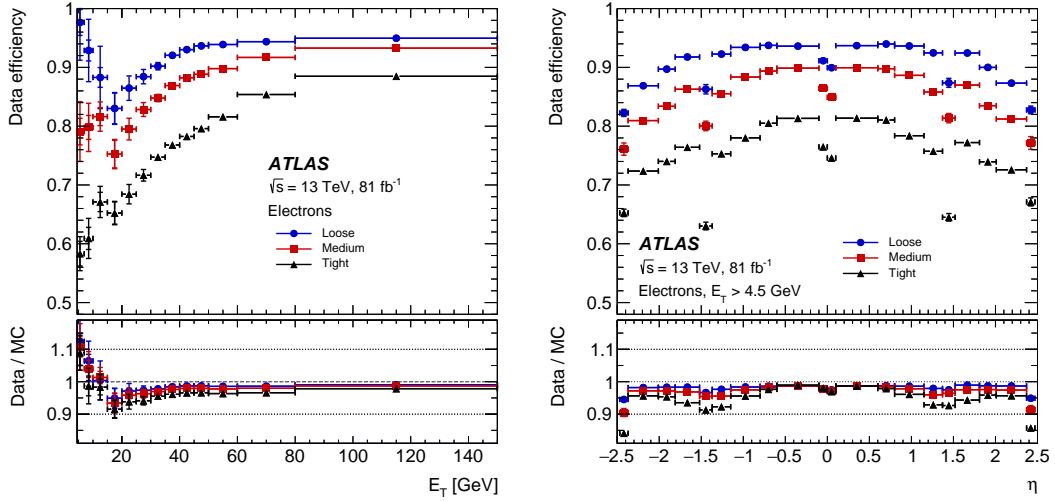


图 3-10 不同工作点下  $Z \rightarrow e^+e^-$  事例中的电子鉴别效率 [12]。

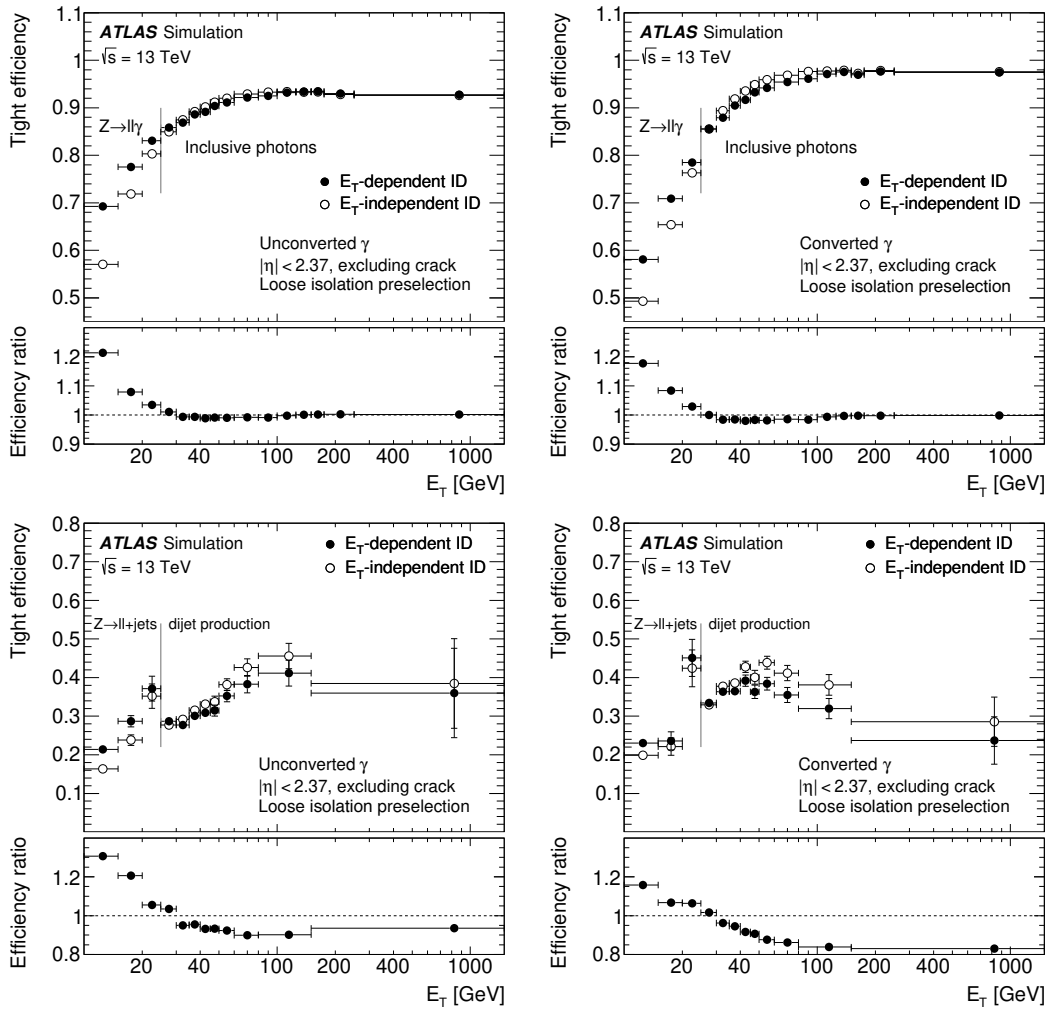
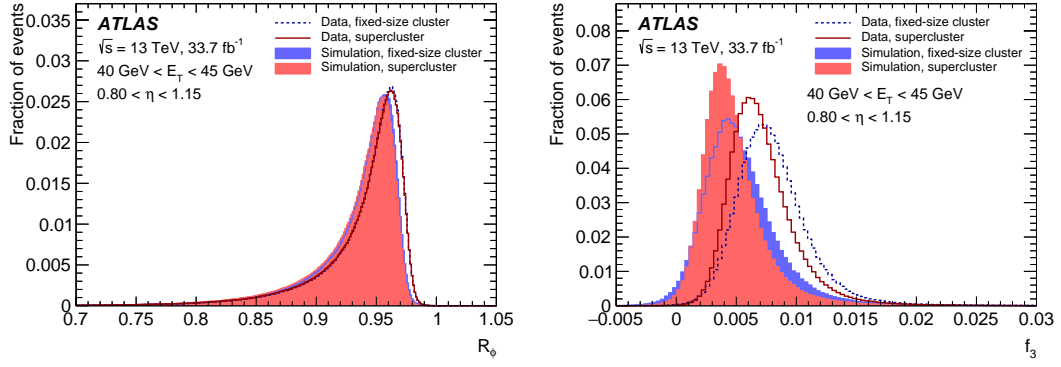


图 3-11 Tight 工作点下光子鉴别效率和本底去除效率 [3]。


 图 3-12 ATLAS 数据和蒙特卡洛模拟中的  $R_\phi$  和  $f_3$  分布图 [3]。

### 3.3.1.3 光子和电子的孤立度

光子和电子的孤立度 (isolation) 是为了排除对 QCD 喷注的误判, 选择出瞬发光子和电子 (prompt photon/electron)。喷注在量能器中产生的簇团通常较大, 在一个光子或电子周围还会有一定能量沉积, 而一个瞬发光子或电子通常是孤立的, 因此我们可以用量能器和径迹信息来定义这个孤立度:

$$\begin{aligned}
 E_T^{coneXX} &= E_{T,raw}^{isolXX} - E_{T,core} - E_{T,leakage}(E_T, \eta, \Delta R) - E_{T,pile-up}(\eta, \Delta R), \\
 p_T^{coneXX} &= \sum p_{T,track}(\Delta R, PV).
 \end{aligned} \tag{3-2}$$

式中  $XX$  为采用的圆锥角大小,  $\Delta R = XX/100$ , 通常采用  $\Delta R = 0.2$  或  $0.4$  ( $XX = 20/40$ )。  $E_{T,raw}^{isolXX}$  是电子/光子周围  $XX$  圆锥角内的横能量和;  $E_{T,core}$  是电子/光子产生的电磁簇射中心  $\Delta\eta \times \Delta\phi = 5 \times 7$  范围内的能量;  $E_{T,leakage}(E_T, \eta, \Delta R)$  是横向泄漏能量,  $E_{T,pile-up}(\eta, \Delta R)$  是 pile-up 导致的横向能量; 径迹孤立度  $p_T^{coneXX}$  即为圆锥角所有径迹的横动量之和。对电子的孤立度来说, 由于电子可能来自于高动量重粒子的衰变, 与其他衰变产物之间的距离可能很近,  $\Delta R$  采用可变尺寸  $\Delta R = \min(\frac{10}{p_T[GeV]}, \Delta R_{max})$ 。  $\Delta R_{max}$  通常为  $0.2$ 。光子和电子的孤立度分别有三个工作点, 具体条件在表 3-5 3-6 中。

表 3-5 电子孤立度工作点定义

Working point	Calorimeter isolation	Track isolation
Gradient	$\epsilon = 0.1143 \times p_T + 92.14\%$ (with $E_T^{cone20}$ )	$\epsilon = 0.1143 \times p_T + 92.14\%$ (with $p_T^{varcone20}$ )
HighPtCaloOnly	$E_T^{cone20} < \max(0.015 \times p_T, 3.5 \text{ GeV})$	-
Loose	$E_T^{cone20}/p_T < 0.20$	$p_T^{varcone20}/p_T < 0.15$
Tight	$E_T^{cone20}/p_T < 0.06$	$p_T^{varcone20}/p_T < 0.06$



表 3-6 光子孤立度工作点定义

Working point	Calorimeter isolation	Track isolation
Loose	$E_T^{\text{cone20}} < 0.065 \times E_T$	$p_T^{\text{cone20}}/E_T < 0.05$
Tight	$E_T^{\text{cone40}} < 0.022 \times E_T + 2.45 \text{ GeV}$	$p_T^{\text{cone20}}/E_T < 0.05$
TightCaloOnly	$E_T^{\text{cone40}} < 0.022 \times E_T + 2.45 \text{ GeV}$	-

### 3.3.2 缪子 [4]

#### 3.3.2.1 缪子重建

ATLAS 上产生的缪子可以稳定地穿过整个探测器，在内部径迹室和缪子探测器中留下击中点，在量能器中通过电离沉积一定能量，因此缪子的重建会综合全探测器的信息。内部探测器中的径迹重建与 3.3.1.1 中电子径迹的重建方法相同，缪子探测器中的径迹重建则从利用霍夫变换找出短直线的径迹段开始，再将这些径迹段连接成一条备选径迹，最后利用全局  $\chi^2$  拟合得到径迹参数。这一  $\chi^2$  拟合中考虑了缪子探测器中的磁场偏转、缪子与探测器物质的相互作用和可能的不同探测器部分之间的非准直（misalignment）效应。

在得到 ID 和 MS 中的径迹后，我们可以进行全局的缪子重建。根据重建策略我们可以把得到的缪子分为 5 种类型：

- 联合缪子（combined muon, CB muon）：如果 MS 和 ID 中的径迹能够匹配，这两段径迹被连接在一起，并在考虑量能器中能量损失后进行重新拟合，得到更精确的缪子重建。对于  $|\eta| > 2.5$  区域，MS 径迹可能仅能匹配到一个 pixel 和 SCT 中的 ID 径迹片段，这样重建得到的 CB 缪子被称为前向硅相关（silicon-associated forward, SiF）缪子。

- 内部-外部联合缪子（inside-out combined muon, IO muon）：IO 径迹由内部径迹外推到 MS 中得到，并要求匹配 MS 中至少三个击中点。相比于 CB 缪子，这一方法不依赖于 MS 中的径迹重建，可以提升低动量缪子的重建效率。

- 外推缪子（muon-spectrometer extrapolated muon, ME muon）：如果一条 MS 中的径迹无法与任何 ID 径迹相匹配，那这条径迹将被外推到束流管得到径迹顶点，并定义一个 ME 缪子。这有助于重建出不在 ID 接受范围内的缪子。

- 片段标记缪子（segment-tagged muon, ST muon）：当 ID 径迹外推到 MS 中时能与 MS 中至少一个击中点相匹配，这条 ID 径迹则被定义为一个 ST 缪子径迹，其径迹参数直接从 ID 径迹得到。



• 量能器标记缪子 (calorimeter-tagged muon, CT muon): 如果 ID 径迹外推到量能器中时可以匹配上一个最小电离粒子 (minimum-ionising particle, MIP) 的能量沉积, 这条径迹则被标记为 CT 缪子径迹, 参数由 ID 径迹决定。对于 CT 缪子, 横动量要求为  $p_T > 5 \text{ GeV}$  以去除大量的低动量本底。

### 3.3.2.2 缪子鉴别

为得到适合物理分析用的高质量缪子, ATLAS 采用了一套基于 ID 和 MS 中击中数、拟合优度和描述两部分间联系程度变量的缪子鉴别方案。同时 ATLAS 定义了多个工作点以适应不同物理分析的需求, 包括宽松 (Loose), 中等 (Medium), 严格 (Tight)、高动量 (High  $p_T$ ) 和低动量 (Low  $p_T$ )。表 3-7 总结了不同工作点下缪子和轻强子的重建效率。

Selection WP	3 < pT < 5		3 < pT < 5		3 < pT < 5		pT > 100	
	$eff_{mu}$	$eff_h$	$eff_{mu}$	$eff_h$	$eff_{mu}$	$eff_h$	$eff_{mu}$	$eff_h$
Loose	90	1.17	98	1.06	99	0.25	98	0.12
Medium	70	0.63	97	0.85	97	0.17	97	0.07
Tight	36	0.15	90	0.38	93	0.12	93	0.04
Low-pT (cut-based)	86	0.82	95	0.1	97	0.17	97	0.07
Low-pT (MVA)	88	0.73	96	0.66	97	0.17	97	0.07
High-PT	45	0.34	79	0.6	80	0.13	80	0.05

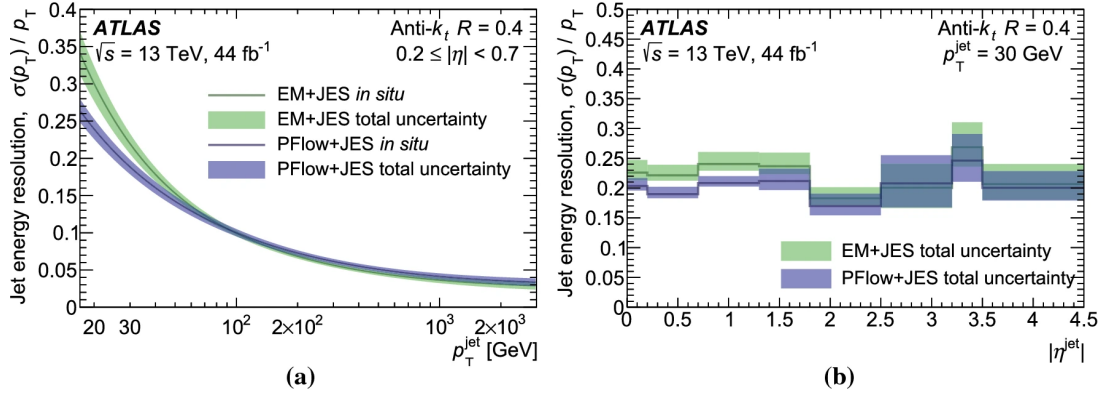
表 3-7 不同工作点下缪子和轻强子的重建效率  $\epsilon_\mu, \epsilon_{had}$ 。Low- $p_T$  工作点下有两套方法, 分别基于筛选条件 (Low- $p_T$  cut-based) 和多变量分析 (Low- $p_T$  MVA)。Tight 工作点不重建  $p_T < 4 \text{ GeV}$  的, 缪子, 因此在  $3 < p_T < 5 \text{ GeV}$  区间缪子效率很低。

### 3.3.3 喷注 [5]

质子-质子对撞中产生的大量夸克和胶子在经过自发强子化过程后会变成大量稳定的粒子, 这些粒子通常聚集在一定范围内, 在探测器上形成一个喷注 (jet)。ATLAS 中大量物理过程的发生都会伴随喷注的产生, 因此其重建和味道标记 (flavor tagging) 会对物理结果产生重要影响。喷注重建的方法为 anti- $k_t$  算法 [74], 在这一算法中, 两粒子间或粒子与喷注间的距离定义为  $d_{ij}$ , 粒子与束流间距定义为  $d_{iB}$ :

$$d_{ij} = \min(k_{ii}^{2p}, k_{ij}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \quad (3-3)$$

$$d_{iB} = k_{ii}^{2p}.$$


 图 3-13 喷注能量分辨随横动量和  $\eta$  的关系 [13]。

其中  $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ ,  $k_{ti}$ 、 $y_i$  和  $\phi_i$  为粒子  $i$  的横动量、赝快度和方位角。anti- $k_t$  算法中  $p$  固定为-1,  $R$  在 Run 2 ATLAS 实验上通常取 0.4。

根据 anti- $k_t$  算法中使用的粒子重建方式不同, ATLAS 上的喷注被分为两种: EMTopo jet 使用量能器中的拓扑簇团 (topocluster) 修正后的能量和位置作为输入; PFlow jet [75] 则使用粒子流算法得到的粒子, 其综合了量能器和径迹探测器中的信息, 因此相比于 EMTopo jet 拥有更好的能量分辨。图 3-13展示了两种喷注重建方法的能量分辨。

喷注味道标记方法的选择根据物理分析的目标而定, 通常针对需研究的末态粒子开发相应的标记算法以提高事例筛选效率, 如 b/c 夸克喷注标记 (b-tagging, c-tagging), 夸克-胶子喷注标记 (quark-gluon tagging) 等。

### 3.3.4 丢失横动量

在 LHC 对撞产物中有一些与探测器相互作用极小的粒子, 它们无法被直接观测到, 只能通过反应的横向能动量守恒根据可见粒子将其重建为丢失横动量 (missing transverse momentum,  $E_T^{\text{miss}}$ )。这些粒子通常包括中微子, 和一些超出标准模型的新粒子, 后者对于寻找新物理非常重要。 $E_T^{\text{miss}}$  的重建通过对可见粒子横向动量 ( $x, y$ ) 求和后求负得到, 且被分为硬成分 (hard objects: 包括电子、光子、缪子、 $\tau$  子和喷注) 和软成分 (soft objects: 重建出的来自于硬散射顶点的径迹和簇团):

$$\begin{aligned} E_{x(y)}^{\text{missing}} &= -(\sum_{\text{hard}} p_{x(y)} + \sum_{\text{soft}} p_{x(y)}), \\ E_T^{\text{miss}} &= \sqrt{(E_x^{\text{missing}})^2 + (E_y^{\text{missing}})^2}. \end{aligned} \quad (3-4)$$

## 第 4 章 ATLAS 上电磁簇射形状

### 4.1 电磁簇射形状的研究

如 3.3.1.1 中所述, ATLAS 的模拟并不能完美地描述实际数据中的簇射形状, 现用的 **fudge factor** 会作为光子和电子鉴别的系统误差项进入到物理分析中。长期以来 ATLAS 组一直试图找出造成这一非精确描述的原因, 可能包含以下 3 部分:

- ATLAS 探测器的几何描述与实际探测器有偏差。液氙电磁量能器之前的物质可能提高光子提前发生簇射几率从而改变在量能器中观测到的簇射形状; 液氙量能器组成成分的偏差, 如铅吸收板厚度、物质结构的组成、液氙灵敏区体积等会直接影响到探测器的等效莫里哀半径, 从而改变簇射形状。同时, 在液氙量能器安装并运行后的这 10 年里可能会由于重力、温度、老化等问题产生一定改变。

- Geant 4 模拟对于电磁簇射过程的描述存在偏差, 可能因素包括: 在进行模拟时为平衡精度和计算资源, 会人为选取模拟步长和能量截断; 模拟采用的 Geant 4 中的物理列表 (**physics list**) 对电磁簇射过程有不同精度的处理方式; 量能器中灵敏区内电场的模拟也会影响簇射发展的过程; 等等。

- 数字化中对于噪声的处理。ATLAS 量能器噪声水平的估计来自于模块的束流测试和刻度运行数据 (**calibration run**), 其中可能存在一定偏差; 不同信号之间的串扰 (**cross talk**) 会影响到读出信号幅度, 从而改变观测到的簇射形状; 对 **pile-up** 造成的噪声的模拟也可能有偏差。

这些因素中, 对电磁簇射过程的模拟一直由 Geant 4 合作组跟进, 多年来随着对物理过程理解的加深, 模拟软件经过了多次更新, 现使用的版本是目前已知最精确的电磁过程描述。ATLAS 探测器的几何描述同样在各子探测器组内有一定研究, 目前未发现有能够明显影响电磁簇射形状的因素。本论文中主要探讨串扰幅度和液氙量能器本身物质对于簇射形状的影响。需注意到这是一项复杂的研究, 需要考虑各个部分可能的影响并不断迭代, 目前结果无法形成准确结论, 但对于理解物理对象的实验探测很有帮助。

#### 4.1.1 串扰效应对簇射形状的影响

在探测器组装和建造中，一个读出单元或电缆中的电子学信号可能会引发临近单元中的信号，这一效应被称为电子学串扰。当粒子进入量能器产生信号时，在临近单元上产生的串扰信号会叠加在真实信号上，改变单元读出电荷分布，从而影响重建出的簇射形状。在 ATLAS 液氙量能器中串扰主要有三种：capacitive, resistive and inductive，分别为 strip layer 中各单元间串扰、middle-layer 与 strip-layer 之间串扰、middle-layer 单元间/middle-back 层间串扰的主要成分。为研究这一效应，ATLAS 组在 calibration run 中人为给单个探测单元注入脉冲信号，观察临近单元产生的扰动信号，并将结果存为数据库放入模拟中。在完成探测器安装并开始取数后，人们利用  $Z \rightarrow \mu^+ \mu^-$  过程对串扰水平进行了重新研究 [14]，观察  $\mu$  附近单元的信号幅度，发现模拟与实际数据存在一定差别，如图 4-1。以此为出发点，我们可以检查这一差距会对簇射形状造成多大影响。

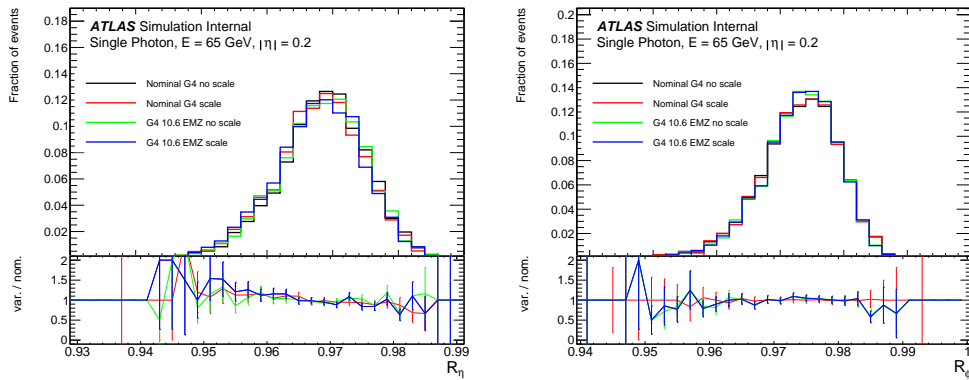


图 4-1 从  $Z \rightarrow \mu^+ \mu^-$  过程中得到的 LAr 电磁量能器桶部第一层（左）和第二层（右）串扰效应 [14]。其中红线为数字化中放入的串扰幅度，黑线为根据数字化模型从模拟样本中得到的串扰幅度，蓝线为数据中得到的串扰幅度。

这一研究使用粒子枪（Particle Gun）样本进行模拟，给定入射方向  $|\eta|$ ，从  $Z \rightarrow \mu^+ \mu^-$  结果中提取数据与模拟间串扰信号幅度的比值加入到液氙量能器数字化中。考虑到之前研究的统计误差，一共提取了 3 个修正值作为串扰幅度的误差范围。同时，已有研究表明相比于 ATLAS 模拟默认使用的 Geant 4 版本 Geant 4 10.1+physics list *FTFP\_BERT\_ATL*，10.6 + physics list *FTFP\_BERT\_EMZ* 对电磁过程的模拟更加精确，因此我们同样用两个版本进行了模拟，以与之前结果交叉验证。具体配置见表 4-1 4-2。

在完成样本的模拟和数字化后，这些单光子按照标准 ATLAS 流程进行了重建。我们选取其中的非转化光子，计算簇射形状变量  $w_{\eta_1}$ ， $w_{\eta_2}$ ， $R_\eta$  和  $R_\phi$  用于

Process	Single photon
DSID	431004, 431010, 431019, 431024
Energy	65.536 GeV
$ \eta $	[0.20, 0.25], [0.50, 0.55], [0.95, 1.00], [1.20, 1.25]
$\phi$	[0, $2\pi$ ]
Geometry	ATLAS-R2-2016-01-00-01
Geant 4 version and physics list	10.1 + <i>FTFP_BERT_ATL</i> 10.6 + <i>FTFP_BERT_EMZ</i>
Statistics	1000

表 4-1 研究使用的模拟样本。

		$ \eta  \in [0.20, 0.25]$	$ \eta  \in [0.50, 0.55]$	$ \eta  \in [0.95, 1.00]$	$ \eta  \in [1.20, 1.25]$
Strip layer	Central	1.174	1.222	1.737	1.780
	Lower	1.000	1.000	1.160	1.224
	Upper	1.289	1.477	1.974	2.361
Middle layer	Central	1.600	1.939	2.015	1.256
	Lower	1.102	1.027	1.241	1.056
	Upper	3.267	2.533	2.800	2.217

表 4-2 从 [14] 中得出的串扰幅度修正值。

研究串扰幅度的影响。图 4-2 展示了这四个变量的分布和在不同配置下与默认配置分布的比值。分布的平均值被提取出来并画在了图 4-3 中。目前结果显示串扰信号的幅度可以显著影响  $\eta$  方向的簇射形状，但对  $\phi$  方向簇射形状几乎无影响。这一结论与预期相符，因为目前的数字化并没有考虑任何  $\phi$  向串扰。

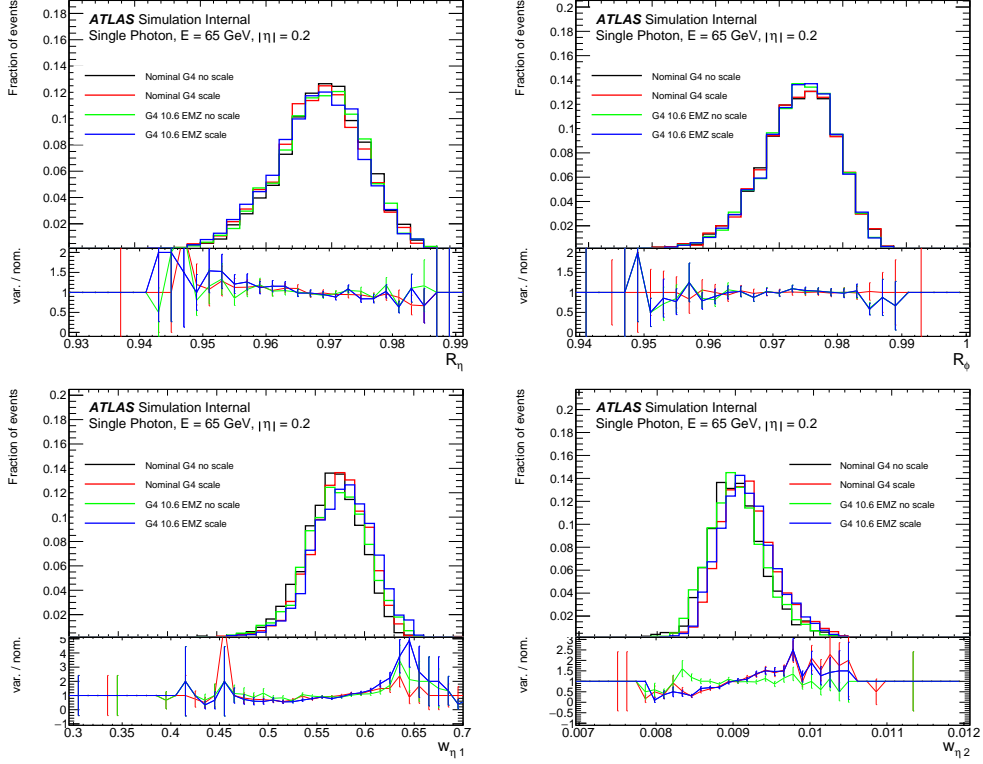


图 4-2 不同配置中簇射形状变量分布。

从这一结果中，我们还可以得到串扰幅度和簇射形状变量之间的关系，并且考虑使用电磁簇射形状来测量探测器中实际串扰效应的可能性。排除掉  $\phi$  方向的簇射形状变量后，我们选取了  $w_{\eta 1}$ ,  $w_{stot}$ ,  $f_{side}$ ,  $\Delta E$ ,  $E_{ratio}$ ,  $R_{\eta}$ ,  $w_{\eta 2}$  七个变量研究其与串扰幅度的关系，其中前 5 个为液氩电磁量能器 **strip layer** 形状变量，后 2 个为 **middle layer** 形状变量。我们使用线性函数来拟合串扰幅度修正值和簇射形状变量平均值之间的关系，并用  $r^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2}$  描述两者之间的线性相关水平。当  $r^2 > 0.999$  时我们认为此簇射形状变量和串扰幅度之间有良好的线性关系，可以尝试从数据的簇射形状（即 **data fudge factor**）反推探测器中串扰水平。四个  $|\eta|$  区间内的 **fudge factor**、拟合得到的串扰水平和线性拟合的  $r^2$  列在表 4-3 和图 4-4 中。可以看出，从不同变量中拟合得到的串扰水平并不一致，且对 **strip** 层的一些变量如  $w_{stot}$ 、 $f_{side}$  来说其拟合值已超出认为的串扰幅度误差范围，说明串扰效应不是簇射形状问题的唯一原因。

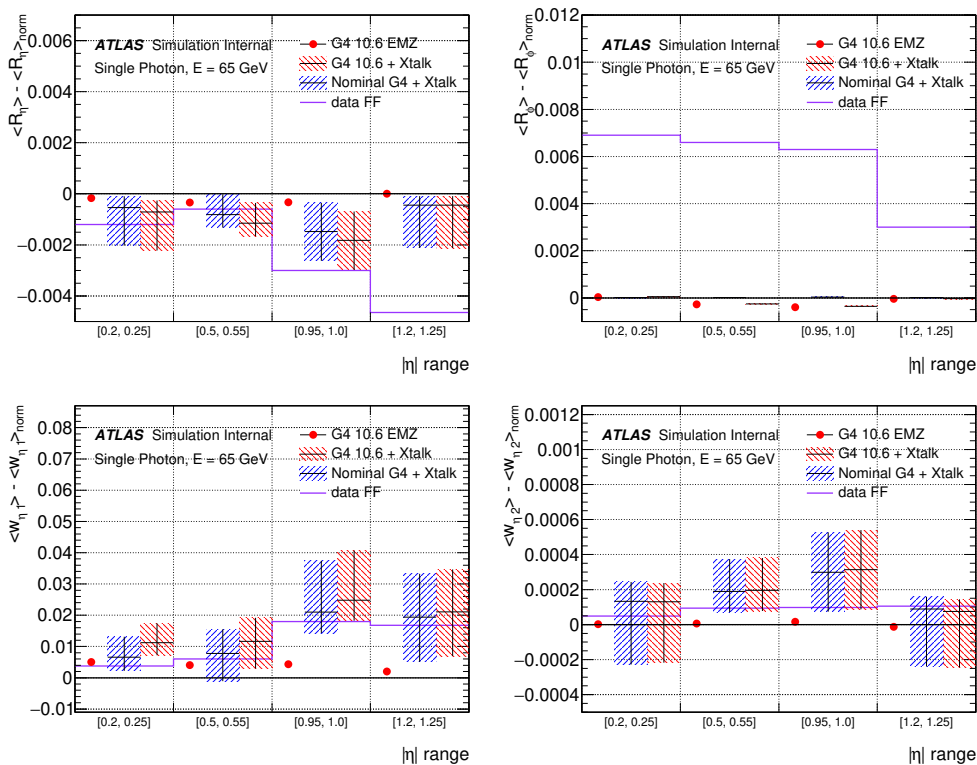


图 4-3 不同  $\eta$  下簇射形状变量的平均值，以 ATLAS 默认模拟样本中的值为参考。图中误差棒表示串扰修正上下限的影响。

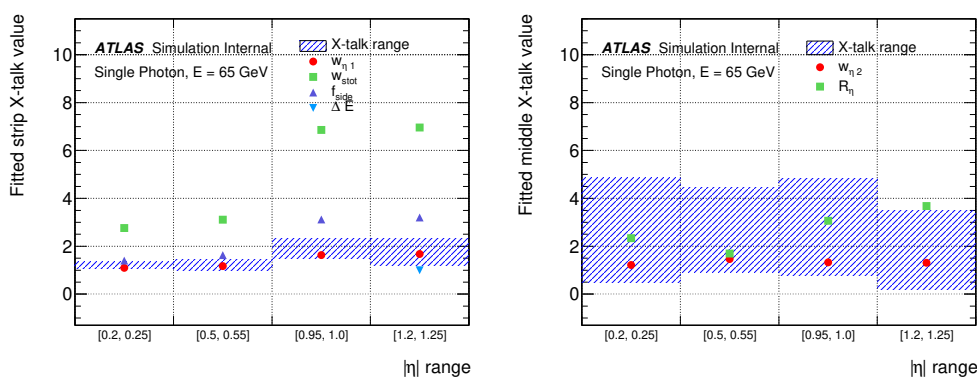


图 4-4 四个  $|\eta|$  区间内从簇射形状中拟合得到的串扰水平，左为 strip 层，右为 middle 层。





### 4.1.2 液氙量能器质量对簇射形状的影响

在串扰幅度之外，另一个考虑的因素是模拟中液氙量能器的几何描述不精确。这一因素有可能解决串扰无法解释的  $\phi$  方向簇射形状问题。现考虑的可能的几何描述偏差和处理方法为：

- 铅吸收体总质量。在对比了几何模型中吸收体总质量和 ATLAS TDR 中记录的质量后发现，几何模型可能对吸收体质量有 4% 的低估。对其修正通过增加模拟中铅的密度值 4% 实现。

- 铅吸收板厚度。同样，几何模型中吸收体厚度为 2.2 mm，而对储存在 CERN 的量能器吸收体样品的测量显示其厚度为 2.4 mm。如果确实存在这一偏差，可能导致灌注的液氙质量少于预估值，从而影响能量相应。对其修正通过减小液氙密度 3% 实现。

- 手风琴型铅吸收板拐角处的模型。在几何模型中，液氙量能器的手风琴型吸收体板由矩形板和圆筒型拐角组成，其密度均匀。而实际建造过程是将整块版弯曲构成，这将导致拐角处吸收板密度不均。因此在模拟中我们将吸收体平板和拐角处的铅分别定义，再降低 40% 拐角处铅的密度。40% 为根据几何尺寸（厚度、圆筒张角）计算得到。

除此之外，我们调用了三个已有的 ATLAS 几何模型，来验证量能器前物质对于簇射形状的影响，包括：

- 对内部探测器的总物质量进行 5% 修正；
- 对 IBL 的物质量进行 10% 修正；
- 对 PP0 的物质量进行 20% 修正；

以上几何模型同样使用单光子粒子枪事例进行模拟，固定能量为 65.536 GeV， $|\eta| \in [0.2, 0.25]$ 。数字化和重建部分使用 ATLAS 默认设置。重建出的非转化光子的簇射形状变量的平均值总结如图 4-5。可以看出，无论哪种几何描述都对簇射形状没有显著影响。

### 4.1.3 结论

在这一工作中我们研究了 ATLAS 液氙电磁量能器中串扰效应和探测器几何模型对电磁簇射形状的影响。在已知 ATLAS 量能器数字化模型与探测器中的串扰有一定偏差的情况下，我们尝试在模拟样本中对串扰幅度进行一个统一的修正，观察样本中簇射形状的变化。当前结果显示串扰效应可以对  $\eta$  方向的簇射形状产生可观的影响，但从数据的簇射形状变量中反推串扰水平的尝试没有成功，

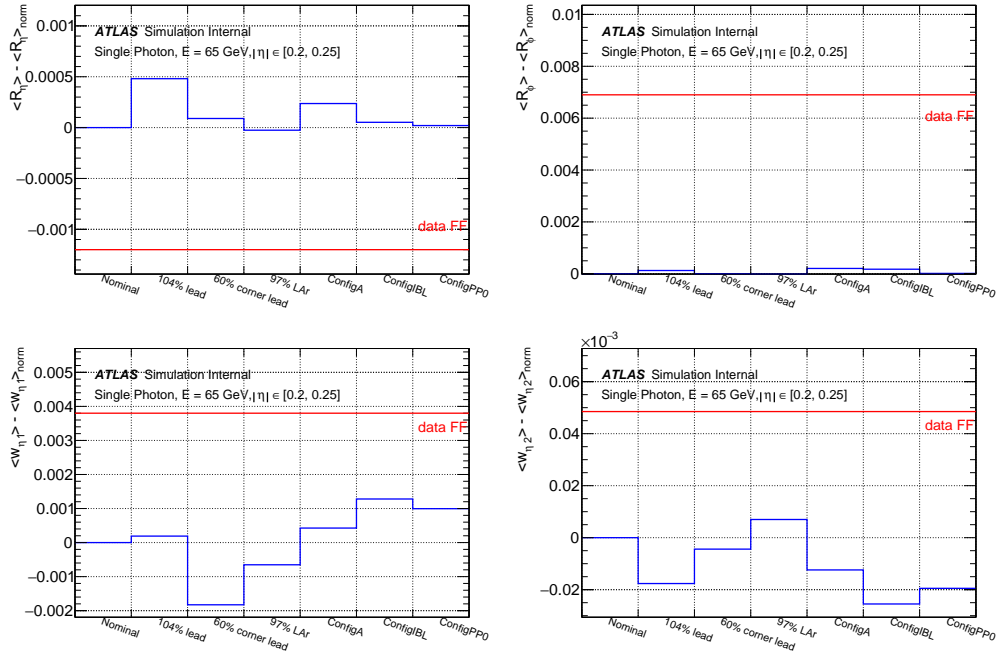


图 4-5 Mean values of shower shape variables  $R_\eta$ ,  $R_\phi$ ,  $w_{\eta_1}$  and  $w_{\eta_2}$  in  $|\eta| \in [0.2, 0.25]$ .

不同簇射形状变量给出的串扰水平并不一致，这说明数字化中串扰的偏差可能是簇射形状问题的原因之一。而调整探测器几何模型无法在簇射形状上观察到显著效应，即我们可以排除掉这一因素。

在其他因素上，目前已有一些最新的研究发现 Geant 4 模拟中对电磁簇射中  $e^+$  散射角正负电子多次库伦散射效应的修正可以使模拟中  $\eta$  方向簇射形状更接近数据，但对于  $\phi$  方向上的簇射形状  $R_\phi$  没有目前已知的因素可以进行可观修正。这一工作相当复杂，但对于深入理解粒子物理实验、ATLAS 探测器相应很有帮助。

## 4.2 基于簇射形状和拓扑簇团的光子鉴别

在 ATLAS Run 2 时，电子和光子的重建由 Run 1 中的固定探测器单元更新到 3.3.1.1 中描述的基于拓扑簇团的动态单元，但光子鉴别仍旧基于固定窗口定义的簇射形状变量。因此我们可以考虑在 Run 3 中加入拓扑簇团信息，更新光子鉴别算法，以得到更好的物理对象重建。本节介绍的工作即利用多变量分析 (MVA) 方法研究拓扑簇团用作光子鉴别的潜力，定量给出相对于现有光子鉴别方法的提升。

## 4.2.1 数据和蒙特卡洛样本，事例选择

这一工作需要干净的末态光子作为信号，被误判为光子的喷注作为本底。选用的物理过程为  $Z \rightarrow ll\gamma$  和  $Z \rightarrow ll$  并伴随一个被误判为光子的喷注（仅考虑  $l$  为电子和缪子）。其中  $Z$  玻色子共振峰可以用作数据和蒙卡模拟的标定。在蒙卡中具体产生的过程如表 4-4，对应选择 ATLAS Run 2 的  $Z \rightarrow ll\gamma$  数据。考虑到仅需要数据做谱型对比，这一研究中我们没有对数据的统计量进行严格要求。

信号过程	产生子	PDF	DSID
$Z \rightarrow ee\gamma$	SHERPA	NNPDF30NNLO	366140-366144
$Z \rightarrow \mu\mu\gamma$	SHERPA	NNPDF30NNLO	366145-366149
本底过程	产生子	PDF	DSID
$Z \rightarrow ee$	PowhegPythia8EvtGen	AZNLO CTEQ6L1	361106
$Z \rightarrow \mu\mu$	PowhegPythia8EvtGen	AZNLO CTEQ6L1	361107

表 4-4 本工作中使用的蒙特卡洛模拟样本。

为得到所需过程，我们对产生的蒙卡事例和数据进行了如下的筛选：

- **Good Run List:** 保证选取的数据是在探测器运行良好条件下取的数；
- 触发：通过单光子和轻子的触发条件；
- 拥有初级顶点；
- 两个轻子要求满足：
  - 至少包含两条径迹，
  - $p_T > 10\text{GeV}$ ，
  - 电子  $|\eta| < 2.47$ ，排除量能器缝隙；缪子  $|\eta| < 2.5$ ，
  - 电子鉴别 LHMedium 工作点要求，缪子鉴别 Medium 工作点，
  - FCLoose 孤立化条件，
  - 顶点的  $d_0/\sigma_{d_0} < 10$ ， $z_0 < 10\text{mm}$ 。
- 一个光子要求满足：
  - 光子鉴别 Loose 工作点要求，
  - $|\eta| < 2.5$ ，排除量能器缝隙区间  $[1.37, 1.52]$ ，
  - $10\text{GeV} < p_T < 140\text{GeV}$ 。
- $Zll\gamma$  事例要求：
  - 轻子对的电荷相反，

- $40\text{GeV} < m_{ll} < 120\text{GeV}$ ,
- $45\text{GeV} < m_{ll\gamma} < 125\text{GeV}$ ,
- $\Delta R_{l-\gamma} > 0.2$ .

同时对于蒙卡事例，还利用真实产生信息（粒子 PDGID，类型 ID，来源等）确保信号和背景中挑选出的光子为真实光子/来自喷注的假光子。

#### 4.2.2 拓扑簇团变量

如 3.3.1.1 所述，拓扑簇团是根据量能器单元中信号的显著度，通过“4-2-0”方式重建而来。光子是以拓扑簇团为种子，考虑径迹、临近簇团等信息后重建而来。图 4-6 展示了一个拓扑簇团大致的几何形状，以及因此定义的一些簇团变量。在本研究中，主要考虑的变量如下：

- $N_{topo-cluster}$  (nTopoCluster): 光子中拓扑簇团数量。
- $E_{topo}$  (topoCluster0\_Epos): 种子拓扑簇团的能量。
- $\sqrt{\langle r^2 \rangle}$  (topoCluster0\_secondR): 种子拓扑簇团到簇射轴径向距离的二次矩。
- $\sqrt{\langle \lambda^2 \rangle}$  (topoCluster0\_secondLambda): 种子拓扑簇团到簇射重心位置的纵向距离。
- $|r|$  (topoCluster0\_centerMag): 种子拓扑簇团重心到对撞点的距离  $\sqrt{x^2 + y^2 + z^2}$ 。
- $\lambda_{cent}$  (topoCluster0\_centerLambda): 种子拓扑簇团重心处的簇射深度。
- $I_{isolation}$  (topoCluster0\_isolation): 种子拓扑簇团周围未被吸收的单元能量分数。

图 4-7 4-8 中展示了信号和本底蒙卡中这 7 个拓扑簇团变量的分布图，以及  $m_{ll\gamma}$  的分布。从图中可看出这些变量的确具有一定的真假光子区分能力。我们使用式 4-1 定义的分辨能力  $\langle S^2 \rangle$  来量化这一区分能力。为了与现有的基于簇射形状的光子鉴别方案做对比，10 个现光子鉴别使用的簇射形状变量的分辨能力也一并计算并列在表 4-5 中。

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_s(y) - \hat{y}_b(y))^2}{\hat{y}_s(y) + \hat{y}_b(y)} dy. \quad (4-1)$$

排序	$Z \rightarrow e^+e^-\gamma$		$Z \rightarrow \mu^+\mu^-\gamma$	
	变量	分辨能力	变量	分辨能力
1	$E_{topo}$	0.32	$E_{ratio}$	0.31
2	$E_{ratio}$	0.22	$E_{topo}$	0.28
3	$\Delta E$	0.21	$\Delta E$	0.17
4	$f_{side}$	0.10	$R_\phi$	0.11
5	$R_\phi$	0.085	$f_{side}$	0.099
6	$w_{\eta_2}$	0.078	$R_\eta$	0.084
7	$\sqrt{\langle r^2 \rangle}$	0.073	$\sqrt{\langle r^2 \rangle}$	0.081
8	$R_\eta$	0.058	$w_{\eta_2}$	0.076
9	$R_{had1}$	0.041	$R_{had1}$	0.066
10	$R_{had}$	0.029	$R_{had}$	0.053
11	$\sqrt{\langle \lambda^2 \rangle}$	0.028	$\sqrt{\langle \lambda^2 \rangle}$	0.039
12	$\lambda_{cent}$	0.024	$\lambda_{cent}$	0.039
13	$ r $	0.014	$ r $	0.023
14	$I_{solation}$	0.013	$I_{solation}$	0.020
15	$w_1$	0.0068	$N_{topo-cluster}$	0.0091
16	$N_{topo-cluster}$	0.0054	$w_{stot}$	0.0028
17	$w_{stot}$	0.0008	$w_1$	0.0012

表 4-5 拓扑簇团变量和簇射形状变量在两个道中的分辨能力。光子要求为满足 LooseID 和 LooseIso。

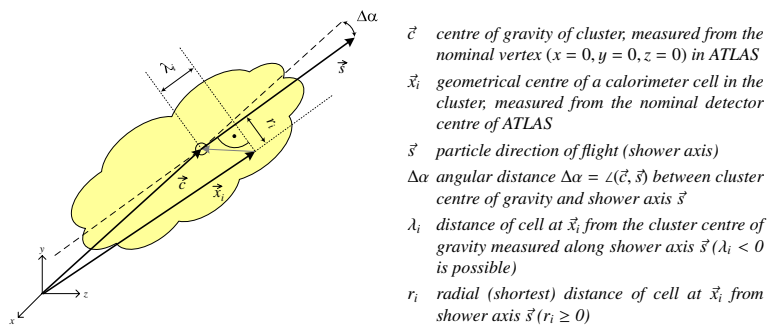


图 4-6 拓扑簇团几何性质的示意图。

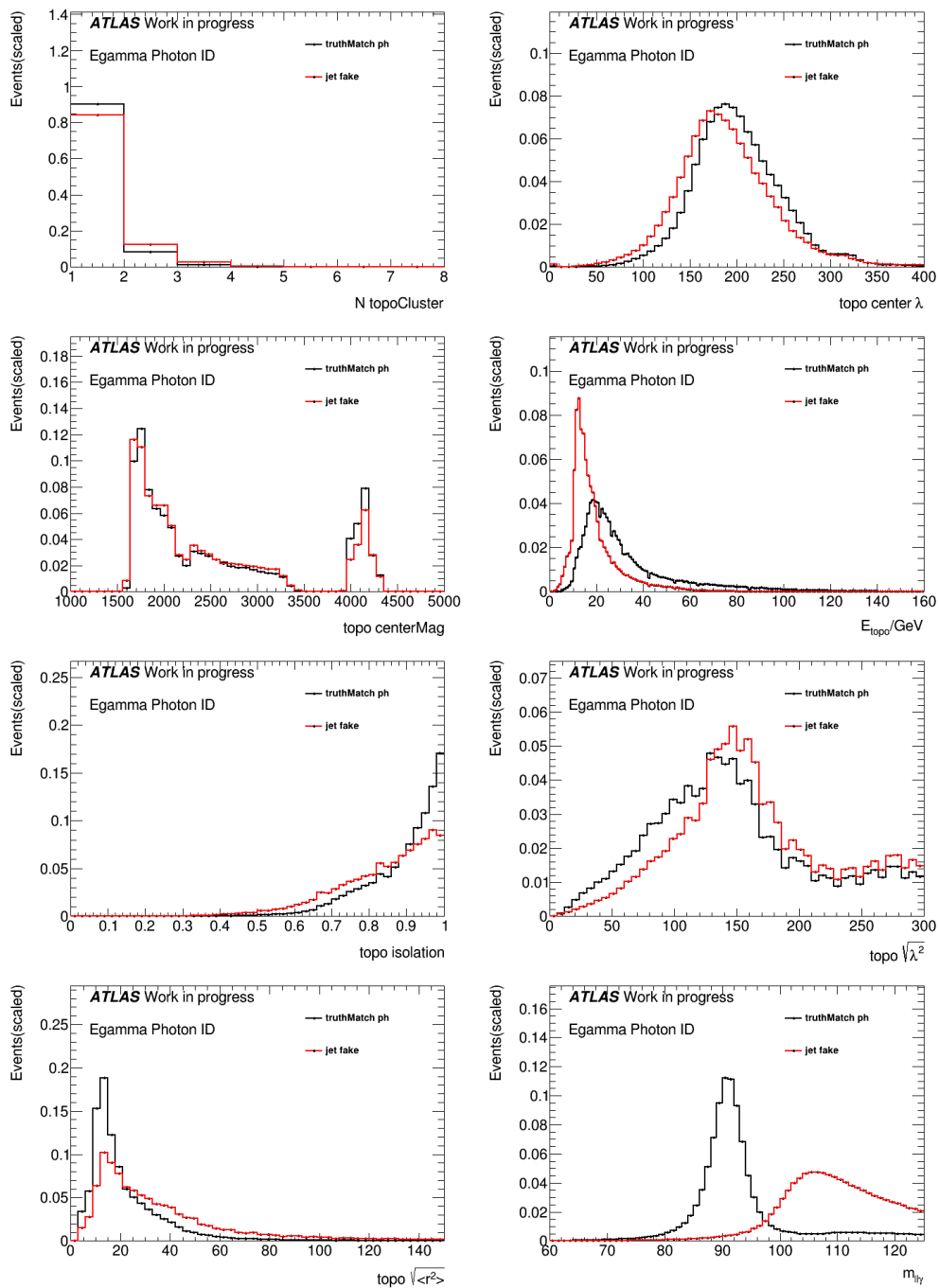
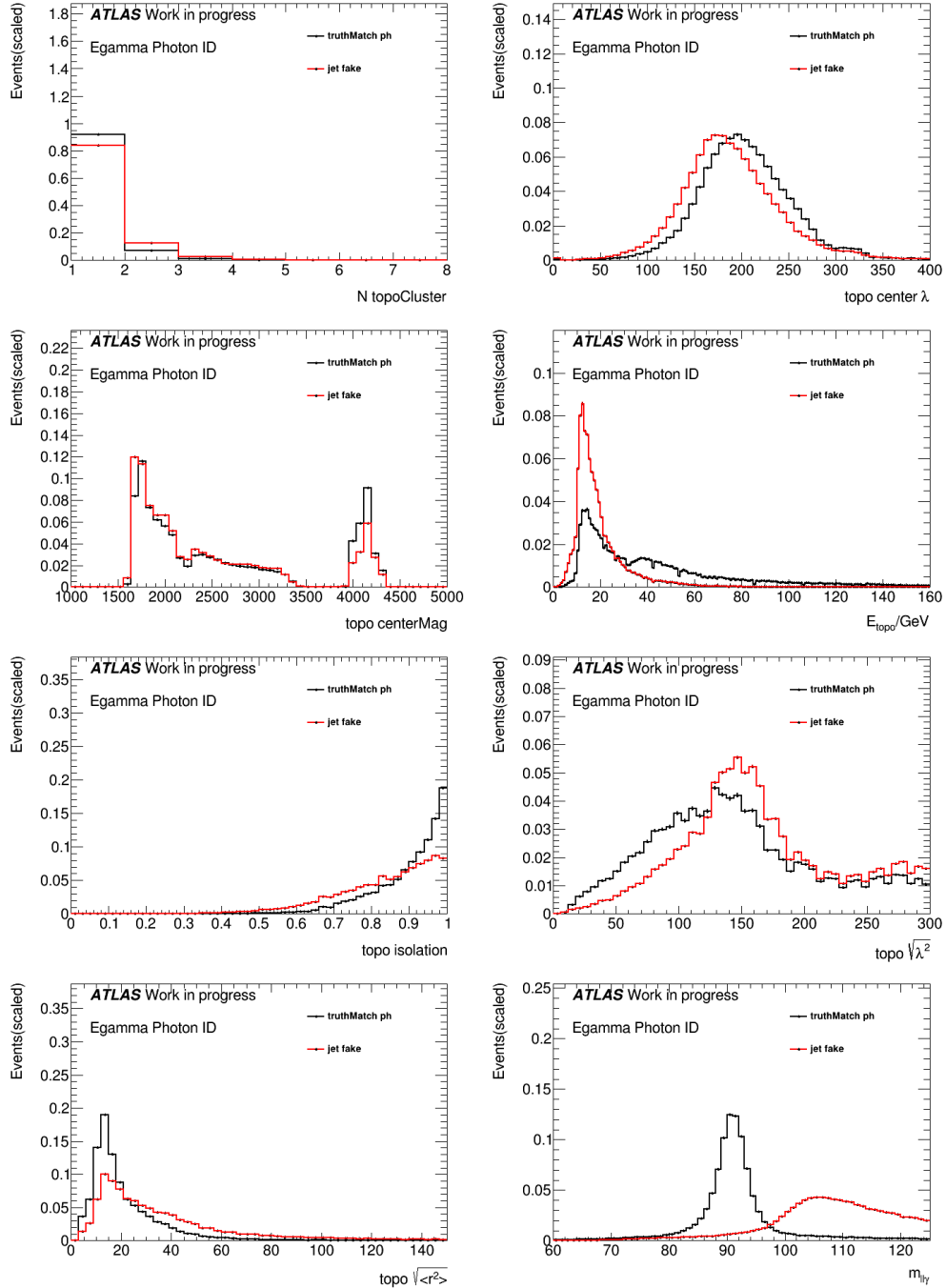


图 4-7  $Z \rightarrow ee\gamma$  样本中 7 个拓扑簇团相关变量和  $m_{ll\gamma}$ 。


 图 4-8  $Z \rightarrow \mu\mu\gamma$  样本中 7 个拓扑簇团相关变量和  $m_{ll\gamma}$ 。



### 4.2.3 多变量分析

相比于单一变量判选，多变量分析可以从多个维度对一个事例进行判断，以达到更好的效果。本研究选取的多变量分析方法为梯度提升决策树（Gradient Boosted Decision Tree, BDTG）。这是基于决策树发展而来的一种事例二分类方法，对输入的变量进行扫描，确定最优分割点，从而构建出一颗决策树，再通过多次迭代使得算法稳定。这种方法需要标记好的信号与本底事例进行训练，在本研究中输入的信号为  $Z \rightarrow ll\gamma$  的蒙特卡洛事例，其中光子经过真实信息认定为真实光子，输入的本底为  $Z \rightarrow ll + jets$  事例，其中光子来自于对喷注的误判。训练事例统计量为信号本底各 10000 个，并另有 10000 个事例被用作训练完成后的测试。由于完全重复现有的光子鉴别算法较为复杂，本研究通过比较采用 10 个簇射形状变量和 10 个簇射形状 +7 个拓扑簇团变量的训练结果来表征拓扑簇团变量带来的提升。17 个变量的线性相关矩阵和相应的两组训练结果（BDTG 响应值的分布和 ROC 曲线）如图 4-9 4-10 4-11。另外为排除现有光子鉴别的影响，确认这一结果，我们将对光子的要求提升至 Tight ID，并重复了这一训练，相关结果如图 4-12 4-13 4-14。同样研究在  $Z \rightarrow \mu^+ \mu^- \gamma$  过程中进行了重复。

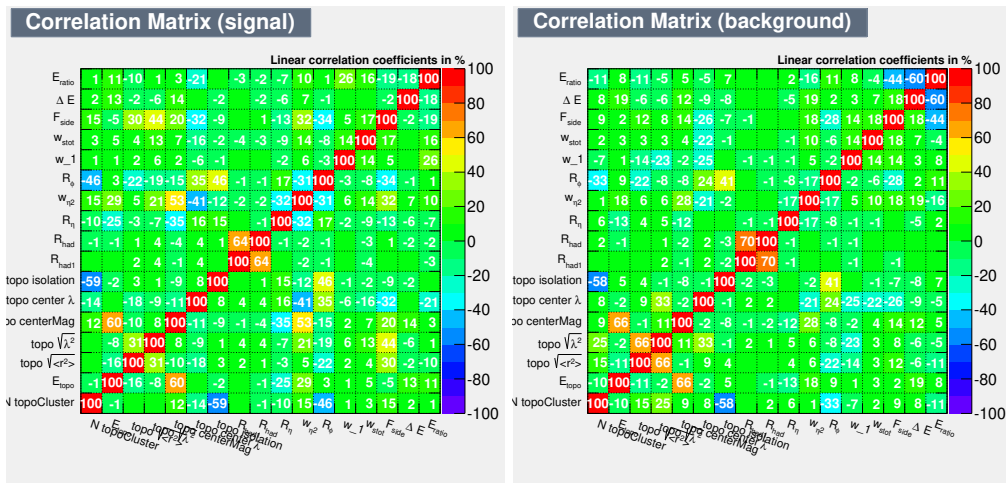


图 4-9  $Z \rightarrow e^+e^-\gamma$  过程中 17 个输入变量的线性相关性矩阵。左：信号样本，右：本底样本。光子鉴别要求为 Loose ID。

### 4.2.4 模拟与数据对照

想要使用以上拓扑簇团变量做光子鉴别需要保证蒙特卡洛对这些变量的模拟能够很好地描述实际数据在探测器中的表现，否则我们对光子鉴别的估计将有一定偏差，从而在物理分析中引入较大的系统误差。因此本研究进行了初步的数据与蒙特卡洛模拟的对比。在这里我们要求选择的光子满足 Tight ID 和 Loose

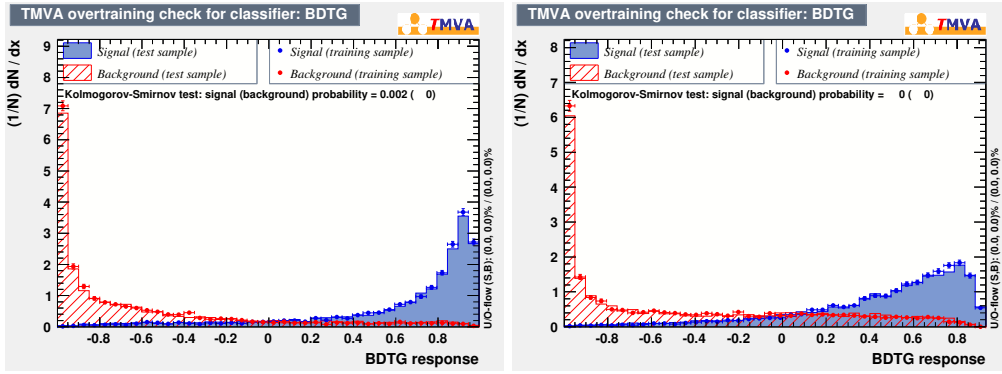


图 4-10  $Z \rightarrow e^+e^-\gamma$  过程 17 个变量 (左) 和 10 个变量 (右) 训练出的 BDTG 分布。光子鉴别要求为 Loose ID。

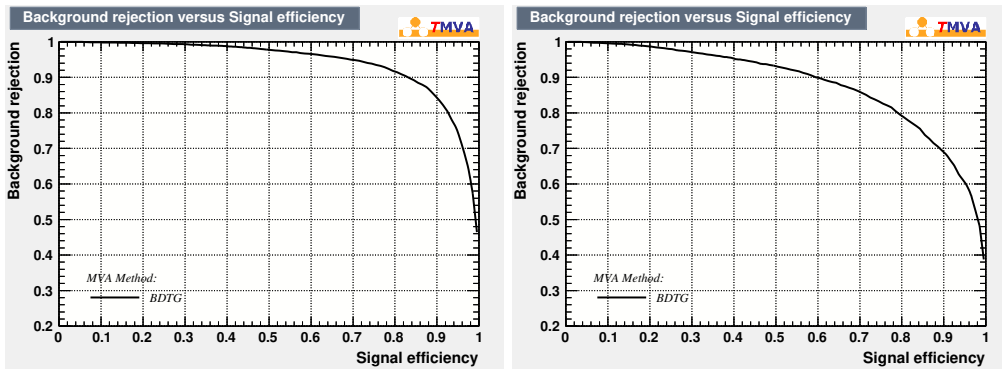


图 4-11  $Z \rightarrow e^+e^-\gamma$  过程 17 个变量 (左) 和 10 个变量 (右) 训练出的 ROC 曲线。光子鉴别要求为 Loose ID。

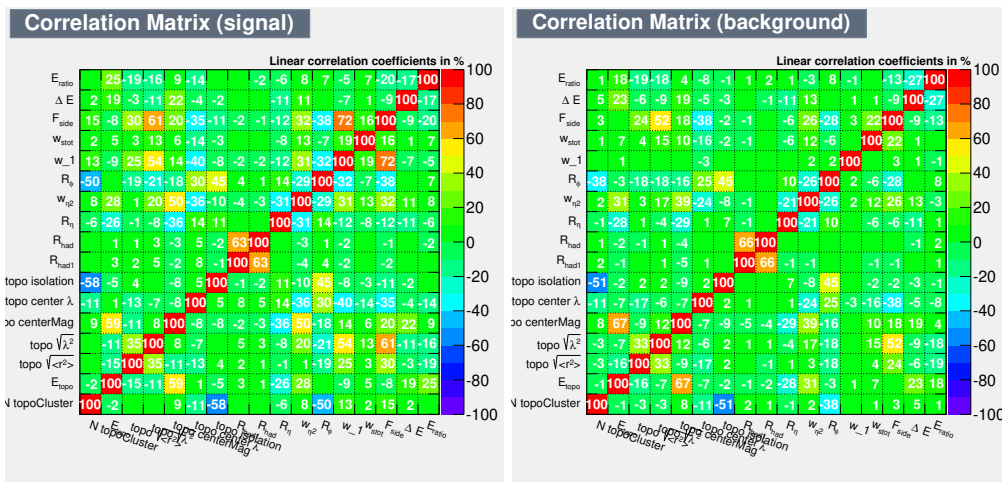


图 4-12  $Z \rightarrow e^+e^-\gamma$  过程中 17 个输入变量的线性相关性矩阵。左：信号样本，右：本底样本。光子鉴别要求为 Tight ID。

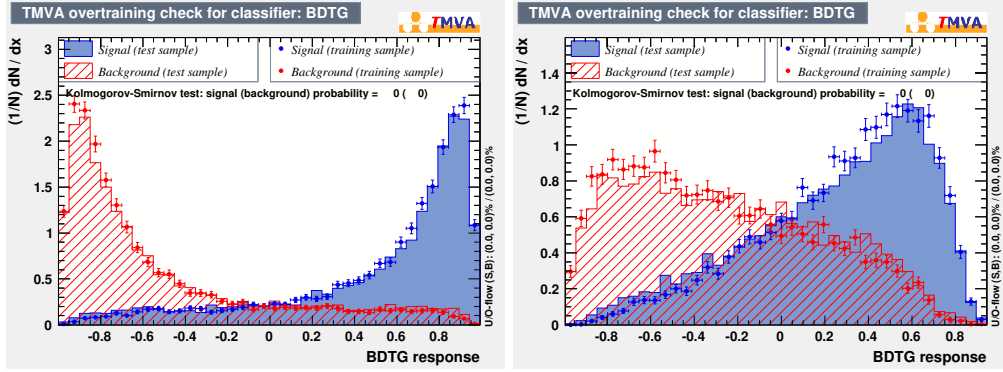


图 4-13  $Z \rightarrow e^+e^-\gamma$  过程 17 个变量 (左) 和 10 个变量 (右) 训练出的 BDTG 分布。光子鉴别要求为 Tight ID。

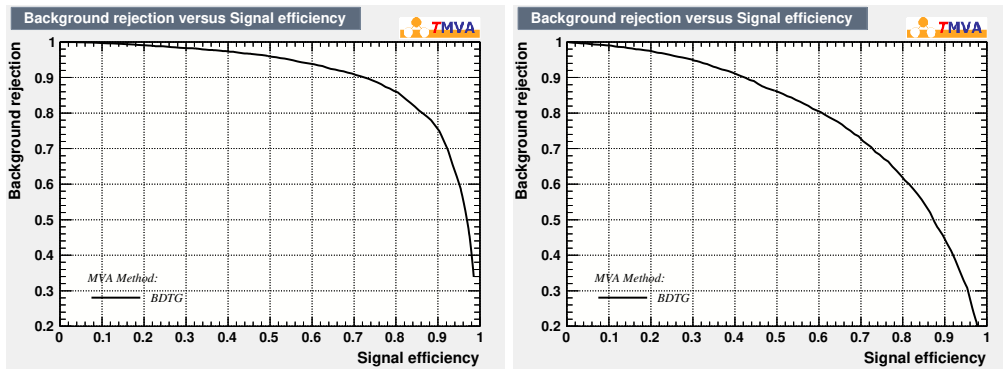


图 4-14  $Z \rightarrow e^+e^-\gamma$  过程 17 个变量 (左) 和 10 个变量 (右) 训练出的 ROC 曲线。光子鉴别要求为 Tight ID。

Isolation, 并假设选出的  $Z \rightarrow ll\gamma$  数据中仅包含真实  $Z \rightarrow ll\gamma$  事例和假光子的  $Z \rightarrow ll + \text{jets}$  事例, 即蒙特卡罗中的信号与本底过程。我们可以通过拟合  $m_{ll\gamma}$  谱来得到这两部分的比例, 如图 4-15。接着我们将信号与本底事例按这一比例进行组装, 得到拓扑簇团变量的分布, 并将其与数据进行对比, 如图 4-16 4-17。现有结果的确显示出数据与模拟之间存在一定不匹配, 以后我们可以进行更细致的研究, 如区分光子横动量区间, 采用更大统计量对其进行估计, 引入修正因子进行修正等。

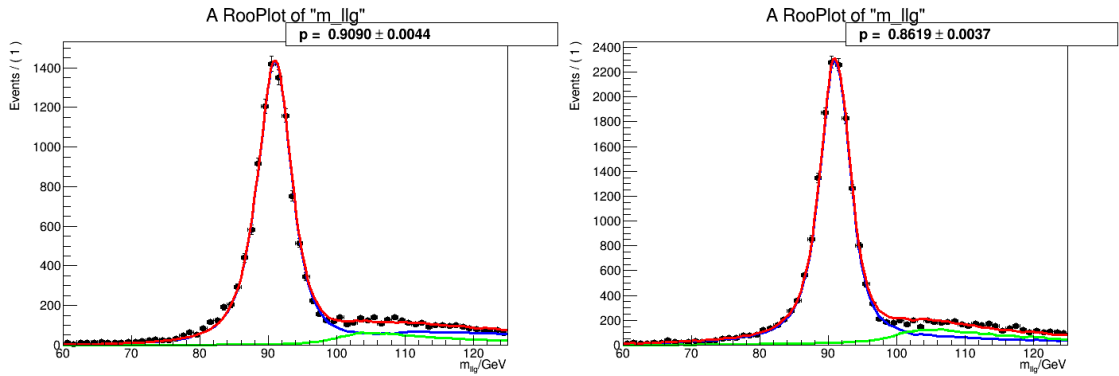


图 4-15 在  $Z \rightarrow e^+e^-\gamma$  (左) 和  $Z \rightarrow \mu^+\mu^-\gamma$  (右) 过程中拟合  $m_{ll\gamma}$  以得到数据中  $Z \rightarrow ll\gamma$  和  $Z \rightarrow ll + \text{jets}$  过程的比例。

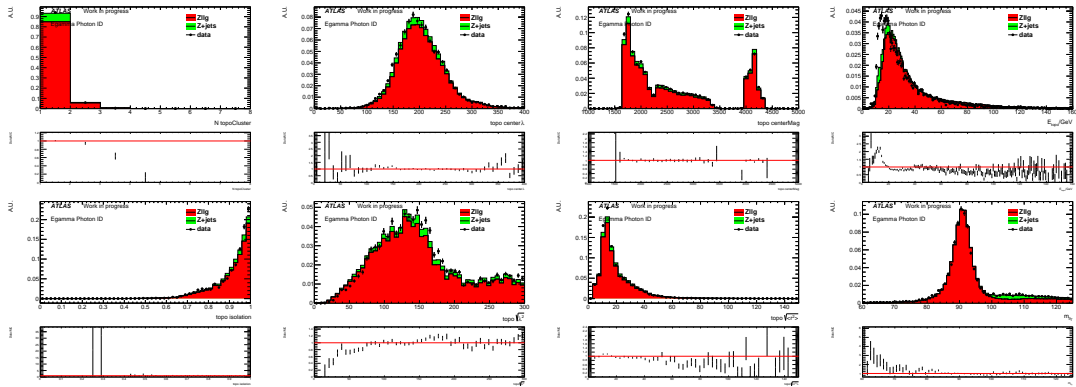


图 4-16  $Z \rightarrow e^+e^-\gamma$  过程中 7 个拓扑簇团变量和  $m_{ll\gamma}$  的数据-模拟对比。模拟样本中两部分按比例进行了叠加。

#### 4.2.5 结论

我们采用 ROC 曲线下面积 AUC (Area Under ROC Curve) 和本底排除率 70% 时的信号效率这两个指标来量化拓扑簇团的影响, 并通过比较 10 个训练变量与 17 个训练变量之间的差距来表示相对于现有的光子鉴别方法, 拓扑簇团变量能够带来的提高。相关结果在表 4-6 4-7 中。可见在 Loose ID 条件下, 引入拓扑簇团

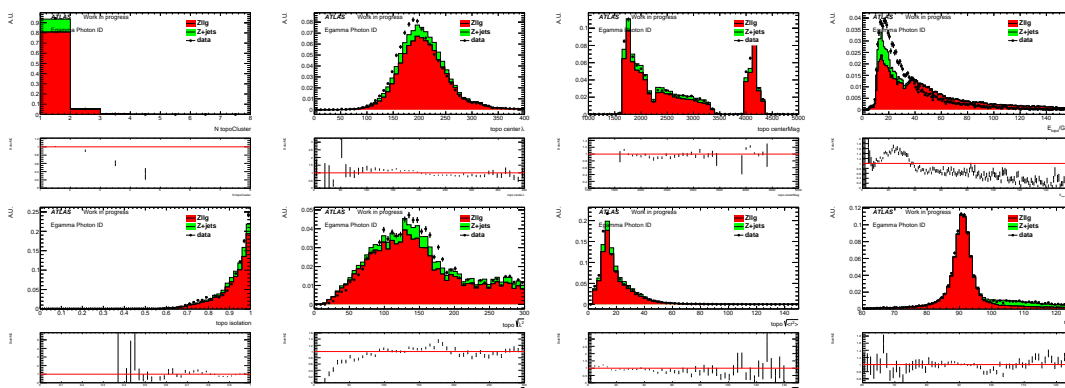


图 4-17  $Z \rightarrow \mu^+\mu^-\gamma$  过程中 7 个拓扑簇团变量和  $m_{ll\gamma}$  的数据-模拟对比。模拟样本中两部分按比例进行了叠加。

变量可以对光子鉴别带来显著提高，而且即使在现有的基于簇射形状的 Tight ID 的基础上使用 BDTG+ 拓扑簇团变量，相关表现同样能够得到提升。在 ATLAS Run 3 及后续运行中，我们可以期待更好的物理对象重建表现。

	AUC			$\epsilon_{sig} @ 1 - \epsilon_{bkg} = 70\%$		
	10 变量	17 变量	相对提升	10 变量	17 变量	相对提升
tight ID	0.787	0.905	15.0%	72.8%	92.2%	26.6%
loose ID	0.882	0.942	6.8%	89.0%	96.2%	8.1%

表 4-6  $Z \rightarrow e^+e^-\gamma$  过程中用 10 个簇射形状变量和 17 个簇射形状 + 拓扑簇团变量训练出的 AUC 和 70% 本底去除率下的信号效率。

	AUC			$\epsilon_{sig} @ 1 - \epsilon_{bkg} = 70\%$		
	10 变量	17 变量	相对提升	10 变量	17 变量	相对提升
tight ID	0.795	0.865	8.7%	74.4%	83.1%	11.7%
loose ID	0.882	0.921	4.4%	88.8%	92.1%	3.7%

表 4-7  $Z \rightarrow \mu^+\mu^-\gamma$  过程中用 10 个簇射形状变量和 17 个簇射形状 + 拓扑簇团变量训练出的 AUC 和 70% 本底去除率下的信号效率。



## 第 5 章 ATLAS 探测器上通过 VBF $H \rightarrow \gamma\gamma$ 过程测量希格斯粒子 $CP$ 性质

### 5.1 研究背景

自 2012 年 LHC 宣布发现了希格斯玻色子 [7, 8] 以来, 对其基本性质的研究就一直是粒子物理领域的热门问题。这些研究有助于确认发现的这一粒子是否是标准模型中预言的希格斯玻色子, 对标准模型进行进一步的检验, 同时寻找新物理的迹象。同时,  $CP$  宇称破缺 ( $CP$  violation,  $CPV$ ) 可以解释现如今观测到的宇宙正反物质不对称问题, 研究希格斯物理中是否同样存在  $CP$  破缺对于寻找新物理非常重要。

在 LHC Run 1 期间, ATLAS 和 CMS 就以 95% 以上的置信度排除了纯奇的希格斯粒子  $CP$  假设 [76, 77]。此后对希格斯粒子  $CP$  的研究以  $CP$  混合效应为主, 测量在希格斯粒子与费米子/玻色子耦合中  $CP$  偶与  $CP$  奇过程的相对强度, 有时也采用耦合强度直接表示。在本研究开始前, 希格斯粒子与费米子的汤川耦合过程中的  $CP$  性质是通过  $t\bar{t}H$  过程测量的, 其中希格斯衰变到双光子末态, ATLAS [78] 和 CMS [79] 合作组都对  $CP$  奇耦合强度给出了一定限制。而希格斯粒子与矢量玻色子作用过程中的  $CP$  性质测量是通过选取 VBF 过程产生的希格斯粒子, 在  $di\text{-tau}$  末态中完成的 [24, 80]。选取这一末态可以确保 HVV 过程中的  $CP$  破缺效应不会影响末态  $\tau^+\tau^-$  的测量, 同时拥有较清楚的本底和可观的统计量。另外, 在  $H \rightarrow \gamma\gamma$  微分截面测量的分析中, 也可以得到相关耦合参数的测量结果 [81]。

理论上讲, 希格斯物理中可能的  $CP$  破缺效应将影响反应末态粒子的动力学分布, 因此传统 HVV 过程中的  $CP$  测量通常采用两喷注方位角之差  $\Delta\phi_{jj}^{sign}$  作为研究变量。在 ATLAS VBF  $H \rightarrow \tau^+\tau^-$  道的希格斯  $CP$  测量 [24] 中, 研究人员引入了优化观测量 (Optimal Observable,  $\mathcal{OO}$ ) 来替代  $\Delta\phi_{jj}^{sign}$ , 并证明了相对于  $\Delta\phi_{jj}^{sign}$  其对  $CP$  的敏感度更高。优化观测量的计算与希格斯粒子衰变无关, 因此我们可以将同样的方法引入  $H \rightarrow \gamma\gamma$  分析中。在双光子衰变道末态更加干净, 本底控制更简单, 同时相关分析工具也更完备, 可以期望得到更精确的测量结果。



## 5.2 理论模型与优化观测量

在标准模型中，希格斯粒子为纯  $CP$  偶标量粒子，与其他粒子的相互作用也不包含  $CP$  破缺效应，因此我们需要在扩展的有效场论（Effective Field Theory, EFT）框架中研究希格斯粒子的  $CP$  破缺。为简化研究，此分析采用的 EFT 只拓展到质量量纲为 6 的作用项，并选取其能标  $\Lambda = 1 \text{ TeV}$  作为限制。质量量纲为 5 的算子将引入轻子数和重子数破缺，在此分析中不考虑，更高阶项则会被  $1/\Lambda$  压制。

### 5.2.1 HISZ 基下的 EFT

本研究考虑的是 HVV 作用过程中的  $CP$  破缺效应，因此在 EFT 中仅考虑标准模型和  $CP$  破缺的 HVV 作用项，这些项可以用希格斯二重态  $\Phi$  和电弱规范场  $B^\mu$  和  $W^{a,\mu}$  ( $a=1, 2, 3$ ) 来表示 [82]:

$$L_{\text{eff}} = L_{\text{SM}} + \frac{c_{\tilde{B}B}}{\Lambda^2} O_{\tilde{B}B} + \frac{c_{\tilde{W}W}}{\Lambda^2} O_{\tilde{W}W} + \frac{c_{\tilde{B}}}{\Lambda^2} O_{\tilde{B}}, \quad (5-1)$$

$$O_{\tilde{B}B} = \Phi^\dagger \hat{B}_{\mu\nu} \hat{B}^{\mu\nu} \Phi, \quad O_{\tilde{W}W} = \Phi^\dagger \hat{W}_{\mu\nu} \hat{W}^{\mu\nu} \Phi, \quad O_{\tilde{B}} = (D_\mu \Phi)^\dagger \hat{B}_{\mu\nu} D_\nu \Phi,$$

其中  $D_\mu$  表示协变导数  $D_\mu = \partial_\mu + \frac{i}{2}g' B_\mu + ig \frac{\sigma_a}{2} W_\mu^a \sigma_a$ ,  $\hat{V}_{\mu\nu}$  ( $V = B, W^a$ ),  $\tilde{V}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma}$  是场强张量，且有  $\hat{B}_{\mu\nu} + \hat{W}_{\mu\nu} = i\frac{g'}{2} B_{\mu\nu} + i\frac{g}{2}\sigma^a W_{\mu\nu}^a$ 。拉氏量的最后一项可通过  $\tilde{\kappa}_\gamma = \cot^2 \theta_W \tilde{\kappa}_Z = -\frac{m_W^2}{2\Lambda^2} c_{\tilde{B}}$  与  $\tilde{\kappa}_\gamma$  和  $\tilde{\kappa}_Z$  相联系，贡献  $CP$  破缺效应，但这一参数已被 ALEPH [83] 和 OPAL [84] 实验结果所限制，因此在这一分析中我们不考虑  $O_{\tilde{B}}$  项的贡献。

原拉氏量在考虑了电弱对称性破缺后，在 HISZ 基 [85, 86] 下可以写成 [57]:

$$L_{\text{eff}} = L_{\text{SM}} + \tilde{g}_{H\gamma\gamma} H \tilde{A}_{\mu\nu} A^{\mu\nu} + \tilde{g}_{H\gamma Z} H \tilde{A}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HZZ} H \tilde{Z}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HWW} H \tilde{W}_{\mu\nu}^+ W^{-\mu\nu}. \quad (5-2)$$

由于  $U(1)_Y \times SU(2)_{I_{W,L}}$  不变性的要求，4 个耦合常数  $\tilde{g}_{HVV}$  中仅有两个是独立的，可以用两个无量纲常数  $\tilde{d}$  和  $\tilde{d}_B$  来表示：

$$\tilde{g}_{H\gamma\gamma} = \frac{g}{2m_W} (\tilde{d} \sin^2 \theta_W + \tilde{d}_B \cos^2 \theta_W), \quad \tilde{g}_{H\gamma Z} = \frac{g}{2m_W} \sin 2\theta_W (\tilde{d} - \tilde{d}_B), \quad (5-3)$$

$$\tilde{g}_{HZZ} = \frac{g}{2m_W} (\tilde{d} \cos^2 \theta_W + \tilde{d}_B \sin^2 \theta_W), \quad \tilde{g}_{HWW} = \frac{g}{m_W} \tilde{d}.$$

这两个常数同样可以用 EFT 中的威尔逊常量  $c$  来表示：

$$\tilde{d} = -\frac{m_W^2}{\Lambda^2} c_{\tilde{W}W}, \quad \tilde{d}_B = -\frac{m_W^2}{\Lambda^2} \tan^2 \theta_W c_{\tilde{B}B}. \quad (5-4)$$



在实验中, 我们无法从 VBF 过程中区分各个矢量规范玻色子的贡献, 因此我们人为选取  $\tilde{d} = \tilde{d}_B$ 。这一规定也与 ATLAS Run 1 在  $H \rightarrow WW^*$  和  $H \rightarrow ZZ^*$  联合测量希格斯粒子  $CP$  分析中采用的  $\tilde{\kappa}_W = \tilde{\kappa}_Z$  假设一致 [87]。因此原先的耦合常数  $\tilde{g}_{HVV}$  可以写成:

$$\tilde{g}_{H\gamma\gamma} = \tilde{g}_{HZZ} = \frac{1}{2}\tilde{g}_{HWW} = \frac{g}{2m_W}\tilde{d} \quad \text{and} \quad \tilde{g}_{H\gamma Z} = 0. \quad (5-5)$$

这也意味着在  $\tilde{d} = \tilde{d}_B$  假设下,  $HWW$ ,  $HZZ$  和  $H\gamma\gamma$  顶点的  $CP$  奇结构强度相同, 而  $H\gamma Z$  顶点的贡献被忽略不计。此时  $\tilde{d}$  成为了在本 VBF 过程的  $CP$  分析中唯一表征  $CP$  破缺效应强度的无量纲耦合系数, 这在后续分析中将带来一定方便。VBF 过程的矩阵元也可以被解耦合为两个  $CP$  偶和  $CP$  奇的两部分:

$$\mathcal{M} = \mathcal{M}_{SM} + \tilde{d} \cdot \mathcal{M}_{CP-odd}. \quad (5-6)$$

反应截面则正比于矩阵元的平方:

$$|\mathcal{M}|^2 = |\mathcal{M}_{SM}|^2 + \tilde{d} \cdot 2\text{Re}(\mathcal{M}_{SM}^* \mathcal{M}_{CP-odd}) + \tilde{d}^2 \cdot |\mathcal{M}_{CP-odd}|^2. \quad (5-7)$$

式中第一项  $|\mathcal{M}_{SM}|^2$  为标准模型贡献, 在此不做讨论。第二项  $2\text{Re}(\mathcal{M}_{SM}^* \mathcal{M}_{CP-odd})$  来自于两个矩阵元的干涉, 为  $CP$  奇, 代表了在此 EFT 模型中希格斯场的  $CP$  破缺来源。在  $CP$  对称的相空间中对截面积分后此  $CP$  奇项会消失, 因此在一些采取了  $CP$  对称筛选条件的分析中并不会观测到总截面或 VBF 事例数的变化。第三项  $|\mathcal{M}_{CP-odd}|^2$  是  $CP$  偶, 不贡献  $CP$  破缺, 但会按  $\tilde{d}^2$  尺度增加 VBF 过程总截面。

以上讨论了 VBF 产生过程中此 EFT 模型的影响, 在  $H \rightarrow \gamma\gamma$  衰变过程中, 新引入的作用项会引入新的衰变作用顶点, 改变希格斯粒子的衰变宽度。对此我们做了特别计算, 在计算中 EFT 下末态粒子的螺旋度与 SM 情况下的正交, 因此没有干涉项。希格斯作为标量粒子, 其  $CP$  性质的改变不会影响末态光子的角分布, 仅考虑分之比的影响。在式 5-2 的拉氏量下,  $H \rightarrow \gamma Z$  和  $H \rightarrow \gamma\gamma$  过程的矩阵元在自旋求和后为:

$$|\mathcal{M}|_{H \rightarrow \gamma\gamma}^2 = 8\tilde{g}_{H\gamma\gamma}^2 \frac{m_H^4}{2}, \quad |\mathcal{M}|_{H \rightarrow \gamma Z}^2 = 2\tilde{g}_{H\gamma Z}^2 (m_H^2 - m_Z^2)^2. \quad (5-8)$$

EFT 下希格斯粒子宽度的变化  $\delta\Gamma$  为:

$$\delta\Gamma(H \rightarrow \gamma\gamma) = \frac{\tilde{g}_{H\gamma\gamma}^2 m_H^3}{4\pi}, \quad \delta\Gamma(H \rightarrow \gamma Z) = \frac{\tilde{g}_{H\gamma Z}^2 m_H^3}{8\pi} \left(1 - \frac{m_Z^2}{m_H^2}\right)^3. \quad (5-9)$$

在  $H \rightarrow ZZ^*$  和  $H \rightarrow WW^*$  道中离壳  $W(Z)$  衰变到费米子对  $f\bar{f}$ ，因此这一过程其实为三体衰变  $H \rightarrow V(\epsilon_1, k_1)f(p_1)\bar{f}(p_2)$ ,  $V = Z, W^+, W^-$ ，其矩阵元平方为：

$$\begin{aligned} |\mathcal{M}|_{H \rightarrow ZZ^*}^2 &= \frac{16N_f \tilde{g}_{HZZ}^2 (v_f^2 + a_f^2)}{(m_H^2 - m_{12}^2 - m_{23}^2)^2} (m_H^2 + m_Z^2 - m_{12}^2 - m_{23}^2) \times (m_{12}^4 + m_{23}^4 - 2m_H^2 m_Z^2), \\ |\mathcal{M}|_{H \rightarrow WW^*}^2 &= \frac{N_f \tilde{g}_{HWW}^2 (g^2)}{(m_H^2 - m_{12}^2 - m_{23}^2)^2} (m_H^2 + m_W^2 - m_{12}^2 - m_{23}^2) \times (m_{12}^4 + m_{23}^4 - 2m_H^2 m_W^2). \end{aligned} \quad (5-10)$$

考虑到  $m_H - m_Z(m_W)$  与 Z-pole 能标相距较远，我们忽略  $W$  和  $Z$  玻色子宽度的贡献。式中  $m_{12} \equiv (k_1 + p_1)^2$ ,  $m_{23} \equiv (k_1 + p_2)^2$ ,  $N_f$  是味参数， $v_f a_f$  是电弱过程中费米子的耦合常数：对带电轻子，有：

$$v_l = -\frac{e(1 - 4\sin^2\theta_W)}{4\sin\theta_W \cos\theta_W}, \quad a_l = -\frac{e}{4\sin\theta_W \cos\theta_W}, \quad N_f = 1. \quad (5-11)$$

对中微子，有：

$$v_\nu = a_\nu = \frac{e}{4\sin\theta_W \cos\theta_W}, \quad N_f = 1. \quad (5-12)$$

对 down-type 夸克 (d, s, b quark)，有：

$$v_d = -\frac{e}{4\sin\theta_W \cos\theta_W} \left(1 - \frac{4}{3}\sin^2\theta_W\right), \quad a_d = -\frac{e}{4\sin\theta_W \cos\theta_W}, \quad N_f = 3. \quad (5-13)$$

对 up-type 夸克 (u, c quark)，有：

$$v_u = \frac{e}{4\sin\theta_W \cos\theta_W} \left(1 - \frac{8}{3}\sin^2\theta_W\right), \quad a_u = \frac{e}{4\sin\theta_W \cos\theta_W}, \quad N_f = 3. \quad (5-14)$$

带入计算后可得两个道中希格斯粒子衰变宽度变化  $\delta\Gamma$ ：

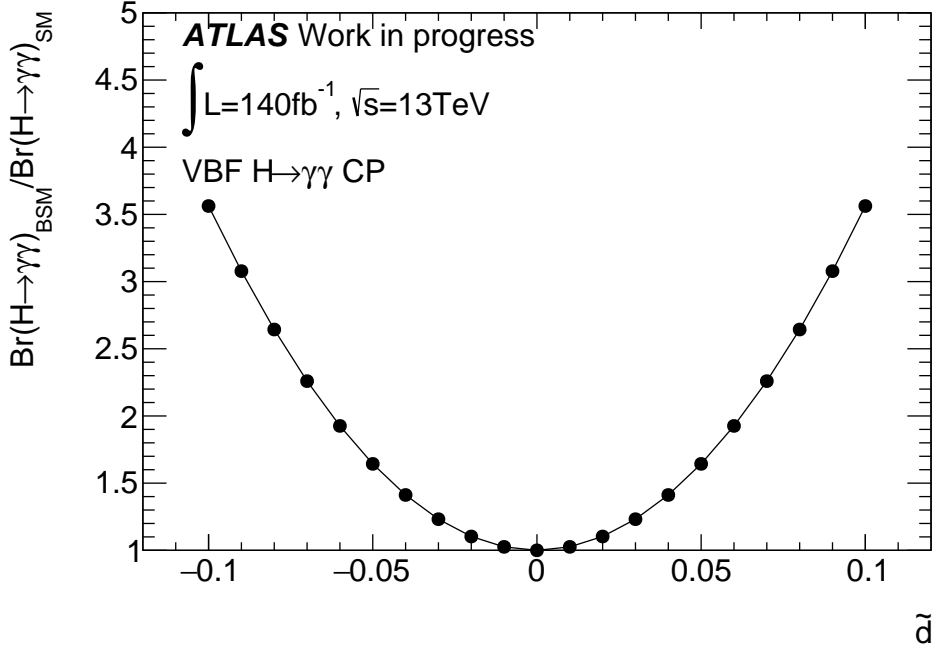
$$\begin{aligned} \delta\Gamma(H \rightarrow Zf\bar{f}) &\approx 0.09503 \frac{N_f \tilde{g}_{HZZ}^2 (v_f^2 + a_f^2) m_H^2}{9216\pi^3}, \\ \delta\Gamma(H \rightarrow W^\pm f\bar{f}) &\approx 2 \times 1.11053 \frac{N_c \tilde{g}_{HWW}^2 g^2 m_H^3}{147456\pi^2}. \end{aligned} \quad (5-15)$$

因此在此 EFT 模型下  $H \rightarrow \gamma\gamma$  的分之比为：

$$Br(H \rightarrow \gamma\gamma) = \frac{\Gamma(H \rightarrow \gamma\gamma)_{SM} + \delta\Gamma(H \rightarrow \gamma\gamma)}{\Gamma(H)_{SM} + \delta\Gamma(H \rightarrow \gamma\gamma) + \delta\Gamma(H \rightarrow \gamma Z) + \delta\Gamma(H \rightarrow Z^* Z) + \delta\Gamma(H \rightarrow W^* W)} \quad (5-16)$$

图 5-1 展示了此分之比随  $\tilde{d}$  的变化。

式 5-7 的第三项和  $Br(H \rightarrow \gamma\gamma)$  的计算结果都表明，VBF  $H \rightarrow \gamma\gamma$  过程的事例数可以直接对此模型中 CP 破缺效应的  $\tilde{d}$  做出一定约束。但考虑到总事例数可能被其他 BSM 效应影响，本分析不对观测到的事例数进行限制，即不考虑这两项的贡献，仅通过 CP 敏感变量的谱型对  $\tilde{d}$  进行限制。


 图 5-1 EFT 模型中  $H \rightarrow \gamma\gamma$  分支比随  $\tilde{d}$  的变化。

### 5.2.2 优化观测量 $\mathcal{OO}$

优化观测量的概念最初由  $\tau$  极化 [88] 和  $Z \rightarrow \tau^+\tau^-$   $CP$  分析 [89] 引入，当反应截面可以被解耦合如式 5-7 时，要测量  $\tilde{d}$  我们可以定义一个观测量：

$$\mathcal{OO} = \frac{2\text{Re}(\mathcal{M}_{SM}^* \mathcal{M}_{CP-odd})}{|\mathcal{M}_{SM}|^2} \quad (5-17)$$

这一观测量将对于  $\tilde{d}$  有很高的敏感性，因此被称为优化观测量。之前  $H \rightarrow \tau^+\tau^-$   $CP$  分析中已验证过，相对于传统变量  $\Delta\Phi_{jj}^{sign}$ ， $\mathcal{OO}$  具有更高的敏感度，本分析将不再重复这一研究，直接使用  $\mathcal{OO}$  作为分析的观测量。

在这一定义中，如果没有  $CPV$  效应， $\mathcal{OO}$  的期望值应为 0，其分布也应以 0 为中心的对称结构。 $CPV$  效应将改变  $\mathcal{OO}$  的平均值和谱型，使其不再对称，偏离程度与  $\tilde{d}$  相关。图 5-2 展示了在模拟的 VBF 过程中不同  $\tilde{d}$  下的  $\mathcal{OO}$  分布。

$\mathcal{OO}$  的计算需要 VBF 过程的矩阵元  $\mathcal{M}_{SM}$  和  $\mathcal{M}_{CP-odd}$ ，这需要对初末态粒子的四动量有精确描述：

$$\mathcal{M}^{ij \rightarrow klH} = \mathcal{M}(p_i^\mu, p_j^\mu, p_k^\mu, p_l^\mu, p_H^\mu) \quad (5-18)$$

其中  $i, j$  指初态夸克， $k, l$  为 VBF 过程的末态夸克， $H$  为希格斯粒子。对于真实模拟样本而言，矩阵元计算所需的初末态信息都可以从产生子的真实信息中得到。对重建后的样本和真实数据而言，这些信息需要一定的处理：

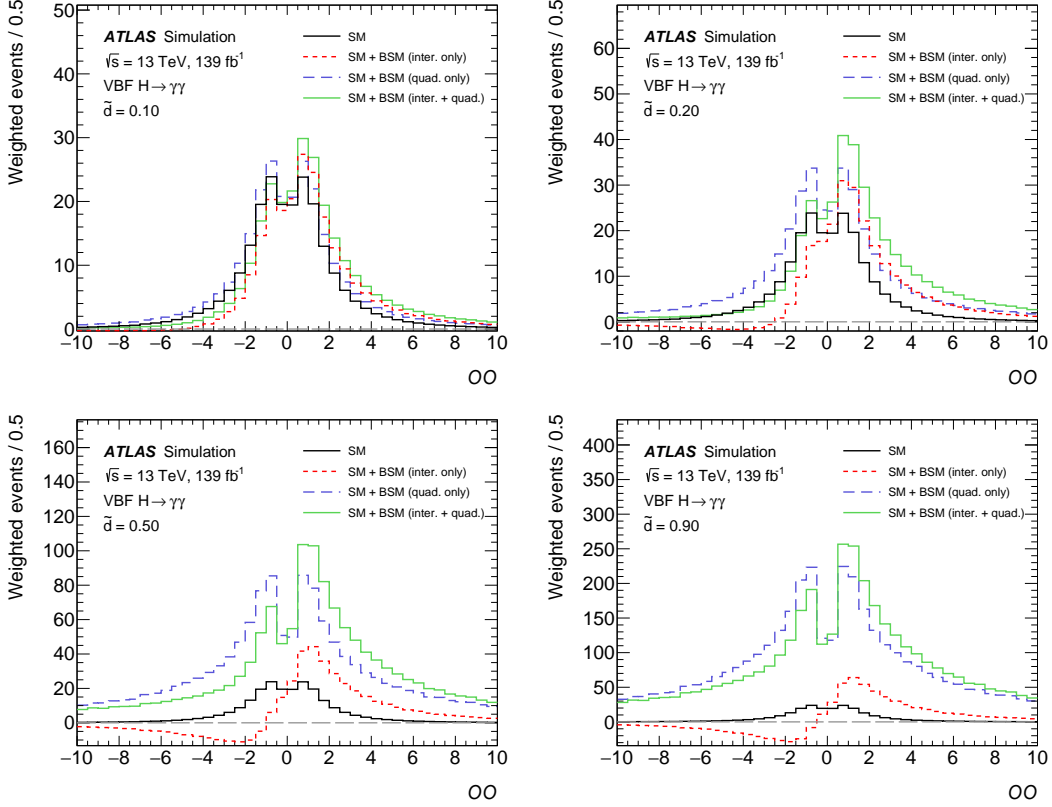


图 5-2  $\tilde{d} = 0.1$  (左上)、 $0.2$  (右上)、 $0.5$  (左下) 和  $0.9$  (右下) 下  $\mathcal{O}\mathcal{O}$  分布与 SM 过程的对比。BSM 的贡献被拆分成了干涉项 (红色虚线) 和二次项 (绿色虚线)。

• 初态两个夸克的动量可通过系统的初末态动量守恒进行估计，计算其占对撞质子的动量分数 Bjorken  $x$ :  $x_{1,2}^{reco} = \frac{m_{Hjj}}{\sqrt{s}} e^{\pm y_{Hjj}}$ ，其中 1 (2) 指沿正 (负)  $z$  方向的部分子， $m_{Hjj}$  与  $y_{Hjj}$  是末态希格斯粒子 + 两喷注系统的不变质量和赝快度。初末态夸克的味信息无法精确地从实验上获取，因此考虑根据 PDF 计算  $ij \rightarrow klH$  过程各种味组合的总截面：

$$|\mathcal{M}_{SM}|^2 = \sum_{i,j,k,l} f_i(x_1) f_j(x_2) |\mathcal{M}_{SM}^{ij \rightarrow klH}|^2, \quad (5-19)$$

$$2\text{Re}(\mathcal{M}_{SM}^* \mathcal{M}_{CP-odd}) = \sum_{i,j,k,l} f_i(x_1) f_j(x_2) 2\text{Re}((\mathcal{M}_{SM}^{ij \rightarrow klH})^* \mathcal{M}_{CP-odd}^{ij \rightarrow klH})$$

- 希格斯粒子由末态选择出的双光子重建而来，对光子的选择见 5.4。
- 末态夸克四动量信息由选择出的两个 VBF 喷注得到，喷注的重建方法见 3.3.3，选择条件见 5.4。

每个事例的以上信息被交给 HAWK [90]，进行计算后得出此事例的  $\mathcal{O}\mathcal{O}$ 。我们对 VBF 过程里真实信息得到的  $\mathcal{O}\mathcal{O}$  分布和重建后样本得到的  $\mathcal{O}\mathcal{O}$  分布进行了对比，见图 5-3

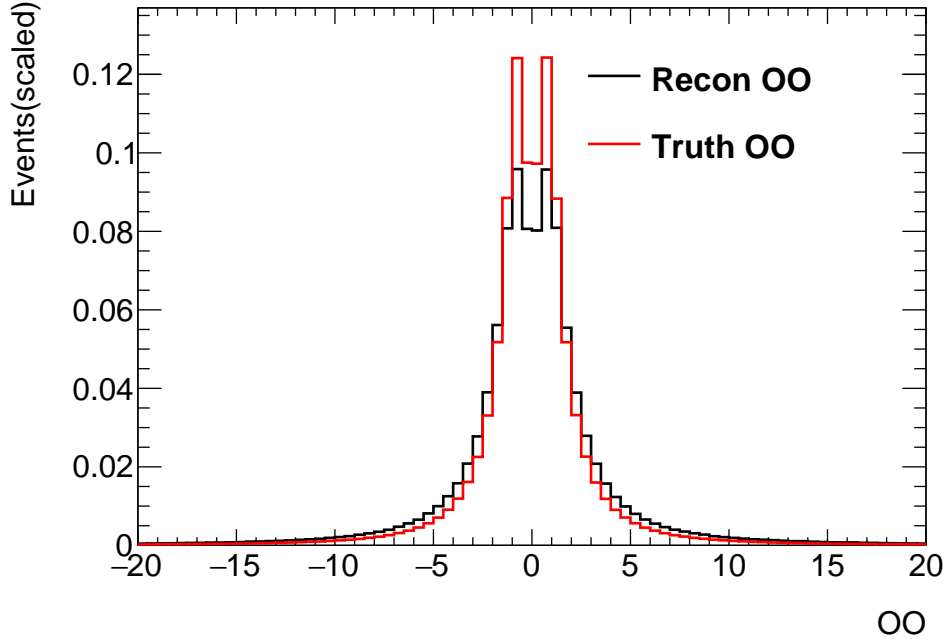


图 5-3 真实级别和重建级别 VBF 过程  $OO$  分布的对比。重建过程会由于能量分辨和效率问题对  $OO$  谱型宽度有微小影响。

### 5.2.3 备选理论模型：SMEFT

为方便与已有希格斯粒子性质测量结果相比较，本分析同时采用了另一套理论模型：在 Warsaw 基 [91] 下的 SMEFT [46]，这一理论描述已被采用为 ATLAS EFT 联合测量的通用模型，其拉氏量可写为：

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \frac{c_i^6}{\Lambda^2} \mathcal{O}_i^6. \quad (5-20)$$

在考虑了 SM 规范对称性和重子数守恒后，SMEFT 中共包含 59 个独立的算子和对应的威尔逊系数  $c_i$  (Wilson coefficients)。在此 VBF 过程的 HVV 顶点研究中，仅考虑以下项：

$$\mathcal{L}_{SMEFT}^{\text{CP-odd}} \supset \frac{c_{H\tilde{W}}}{\Lambda^2} H^\dagger H W_{\mu\nu}^I W^{\mu\nu I} + \frac{c_{H\tilde{B}}}{\Lambda^2} H^\dagger H B_{\mu\nu}^A B^{\mu\nu} + \frac{c_{H\tilde{W}B}}{\Lambda^2} H^\dagger \sigma^I H W_{\mu\nu}^I B^{\mu\nu}, \quad (5-21)$$

由于 VBF 过程实际以  $HW W$  为主导，本分析对  $c_{H\tilde{B}}$  和  $c_{H\tilde{W}B}$  几乎没有敏感度，仅考虑对  $c_{H\tilde{W}}$  的测量。 $OO$  同样可以展示出对  $c_{H\tilde{W}}$  的敏感度，如图 5-4 所示。需注意根据定义此时按之前方法计算得到的  $OO$  并非 SMEFT 下的最优化观测量，但同样可以作为  $CP$  敏感变量进行  $CP$  性质的测量。

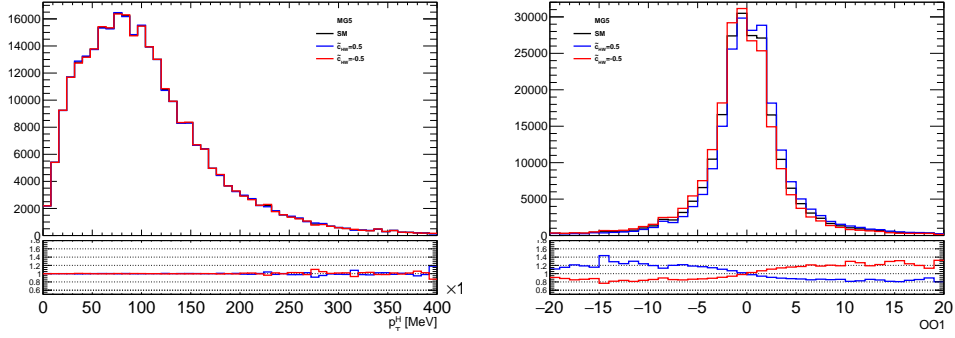


图 5-4 VBF 过程在  $c_{H\bar{W}} = -0.5$ ,  $0$  (SM) 和  $0.5p_T^H$  (左) 和  $\mathcal{O}\mathcal{O}$  (右) 的对比。其他 Wilson 参数固定到 SM 值。

### 5.3 数据与蒙特卡洛样本

#### 5.3.1 数据样本

本分析使用了 2015 - 2018 年 LHC Run 2 期间 ATLAS 收集到的全部  $\sqrt{s} = 13$  TeV 质子-质子对撞数据，在对数据质量进行了要求 [92] 后总积分亮度为  $139.0 \pm 2.4\text{fb}^{-1}$  [22, 23]，如表 reftab:dataLumi 所示。对撞期间平均每次对撞发生反应次数  $\mu = 34$ 。数据采集选用了单光子或双光子触发，其中 2015 - 2016 年的数据里单光子触发要求光子能量大于 120 GeV，2017 年这一阈值被提高至 140 GeV 以提高对高  $p_T^H$  事例的接受度。双光子触发要求两个光子能量分别大于 35 GeV 和 25 GeV，且要求光子满足 Loose 孤立度（2015 - 2016 年）或 Medium 孤立度（2017-2018 年）的要求，具体见 5.4。

运行时间	积分亮度 [ $\text{fb}^{-1}$ ]	绝对误差 [ $\text{fb}^{-1}$ ]	相对误差 [%]
2015+2016	36.2	0.8	2.1
2017	44.3	1.1	2.4
2018	58.5	1.2	2.0
总计	139.0	2.4	1.7

表 5-1 ATLAS Run 2 期间积分亮度和误差 [22, 23]。

#### 5.3.2 蒙特卡洛样本

在高能物理实验中我们采用蒙特卡洛方法得到物理过程的模拟样本，过程包括采用如 MADGRAPH5\_aMC@NLO [93] 的产生子计算物理过程的矩阵元，利用如 POWHEG-BOX [v2] [94-97] 计算 PDF 权重、QCD 能标、高阶修正等效应，

PYTHIA [98] 计算末态夸克和胶子的强子化过程, 对最终的稳定粒子用 Geant4 [99] 进行探测器模拟和数字化, 最终重建为物理对象。本分析中考虑的物理过程被分为信号和本底两部分, 信号为 SM 和 BSM 的 VBF 过程, 本底为其他希格斯粒子产生过程 (也被称为共振本底) 和双光子连续本底。EFT 模型中引入的 BSM HVV 顶点同样会对其他希格斯产生过程, 如 VH 有一定影响, 但本分析针对 VBF 过程开发, 在完成事例选择和分类后这些过程可忽略, 具体见 5.4 和 5.5。

### 5.3.2.1 信号样本

SM VBF 过程的模拟采用了 POWHEG-Box [v2] 产生子和 PDF4LHC15+AZNLO PDF, 用 PYTHIA [v8] 进行部分子簇射 (Parton Shower, PS) 的计算。POWHEG 是一个次领头阶 (Next-to-Leading Order, NLO) QCD 精度的产生子, 但总截面的计算达到了 NLO+ 近似 NNLO 的精度。模拟中的希格斯粒子质量被设定到 125 GeV, 质量宽度  $\Gamma_H = 4.07$  MeV。  $H \rightarrow \gamma\gamma$  分之比采用的是分析开始时 HDECAY [100–102] 和 PROPHECY4F [103–105] 的计算结果  $2.270 \times 10^{-3}$ 。事例数被归一到产生截面  $\times$  分之比上。其他共振本底的希格斯过程在希格斯粒子质量、质量宽度和分之比上使用了相同的设置。

BSM VBF 过程通过对 SM VBF 进行重新加权 (reweighting) 得到, 对应两种理论模型下的 BSM VBF, 我们采用了两种不同的方法。

**HISZ 基下的 BSM VBF** 根据式 5-7, BSM VBF 事例可以通过 SM VBF 事例乘一个权重得到, 权重值的定义为两者矩阵元的比值:

$$w \equiv \frac{|\mathcal{M}_{BSM}|^2}{|\mathcal{M}_{SM}|^2} = 1 + \tilde{d}w_{lin} + \tilde{d}^2w_{quad} \quad (5-1)$$

其中  $w_{lin} \equiv 2\text{Re}(\mathcal{M}_{SM}^* \mathcal{M}_{CP-odd})$  为线性项,  $w_{quad} \equiv |\mathcal{M}_{CP-odd}|^2$  为二次项。这是一个基于每个事例的权重,  $w_{lin}, w_{quad}$  值由 HAWK 根据 SM VBF 样本中的真实信息计算得到, 代入不同  $\tilde{d}$  值后可以得到一组 BSM VBF 样本, 可以节省单独产生 BSM 样本的计算资源。

**SMEFT Warsaw 基下的 BSM VBF** 在 SMEFT Warsaw 基下, BSM VBF 样本由 MADGRAPH5\_aMC@NLO 的重新加权方式产生, 每个事例包含一个 LO SM 截面的权重和一组对应于特定 Wilson 参数的截面的权重。从这组 MADGRAPH5\_aMC@NLO 样本中我们可以计算出特定相空间中权重的比值, 定义为 k-factor:

$$k_{c_i}|_{\Omega} = \frac{\sigma_{BSM c_i}}{\sigma_{SM}}|_{\Omega} \quad (5-2)$$

MADGRAPH5\_aMC@NLO 样本统计量不足以直接进行分析，因此我们将这组  $k$ -factor 应用到前述 SM VBF 样本中即可得到大统计量的 BSM VBF 样本。图 5-5 展示了 MADGRAPH5\_aMC@NLO 产生的 SM VBF 样本、直接产生 BSM VBF 样本和  $k$ -factor 加权得到的 BSM VBF 样本的  $\mathcal{O}\mathcal{O}$  分布，可确认此加权方法的可靠性。

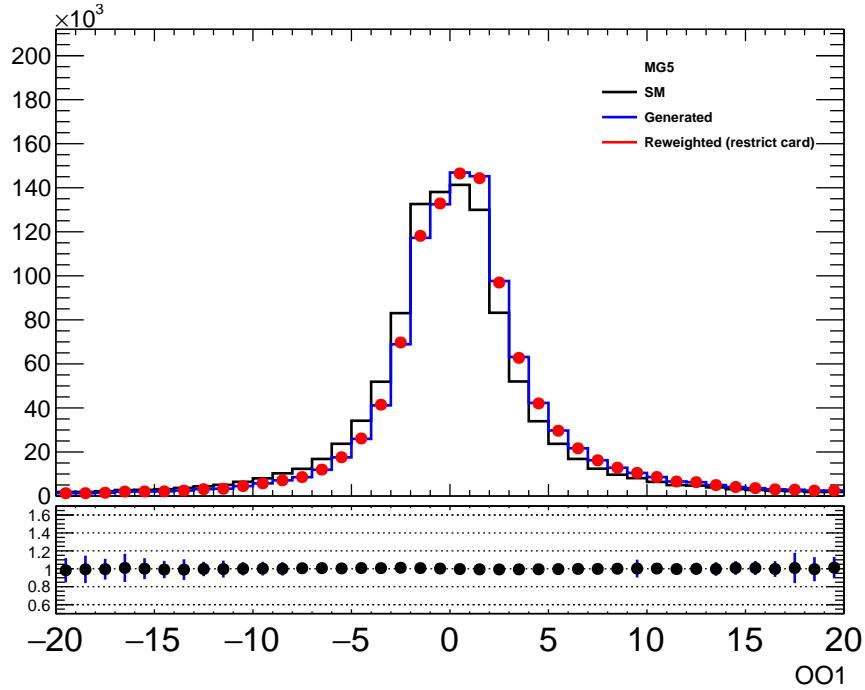


图 5-5 MadGraph5\_aMC@NLO 直接产生的 BSM 样本和经过  $k$ -factor 方法加权得到的 BSM 样本的  $\mathcal{O}\mathcal{O}$  分布。比值显示两种方法结果一致。

### 5.3.2.2 本底样本

共振本底部分为除 VBF 以外的其他希格斯产生过程，其中  $ggF$  过程因其在  $p-p$  对撞中产生截面最大而成为主导共振本底，其他过程均产生了对应的蒙特卡洛样本纳入考虑。对于 SM VBF 和 SM  $ggF$  过程，还产生了一组备用样本用于进行 PS 相关系统误差的计算，具体见 5.7。非共振本底为  $\gamma\gamma + jets$  过程，使用 SHERPA 产生子产生，其中 0、1 喷注过程的精度为 NLO，2、3 喷注过程精度为 LO。ATLAS 探测器模拟采用参数化探测器的快速模拟 (AF2 [106]) 以节省计算资源。另外还有因重建过程中将喷注误判为光子导致的可去除本底 (reducible background)，这部分样本通过数据驱动 (data driven) 方式获得，具体见 5.6.2。全部蒙特卡洛样本总结如表 5-2。



Process	DSID	Generator+PS	QCD+EW order	PDF ME + PS tune	Events in AOD
ggF	343981	POWHEG NNLOPS + PYTHIA	$N^3$ LO + NLO	PDF4LHC15 + AZNLO	18.3M
VBF	346214	POWHEG + PYTHIA	approx. NNLO + NLO	PDF4LHC15 + AZNLO	18.3M
$W^+H$	345318	POWHEG + PYTHIA	NNLO + NLO	PDF4LHC15 + AZNLO	0.6M
$W^-H$	345317	POWHEG + PYTHIA	NNLO + NLO	PDF4LHC15 + AZNLO	0.6M
$qq \rightarrow ZH$	345319	POWHEG + PYTHIA	NNLO + NLO	PDF4LHC15 + AZNLO	1.5M
$gg \rightarrow ZH$	345361	POWHEG + PYTHIA	NLO + NLO	PDF4LHC15 + AZNLO	0.15M
$t\bar{t}H$	346525	POWHEG + PYTHIA	NLO + NLO	PDF4LHC15 + A14	7.8M
$b\bar{b}H$	345315	POWHEG + PYTHIA	(5FS NNLO, 4FS NLO) + NLO	PDF4LHC15 + A14	0.3M
$tHbj$	346486	AMC@NLO + PYTHIA	5FS(NLO)	NNPDF3.0 + A14	0.4M
$WtH$	346188	AMC@NLO + PYTHIA	5FS(NLO)	NNPDF3.0 + A14	0.21M
$\gamma\gamma + \leq 3jets$	364352	SHERPA	NLO(0,1 jet)/LO(2,3jets)	NNPDF3.0NNLO	506M(AF2)

表 5-2 分析中使用的蒙特卡洛样本。在本 VBF 分析中 ggF 过程产生的希格斯事例室主要共振本底，其他希格斯产生过程在 5.4 中验证过可忽略。

## 5.4 事例选择

为选出高显著性和纯度的 VBF 事例，本分析对光子、喷注及整事例进行了一系列要求，具体如下：

**光子** 光子的重建、刻度和鉴别如 3.3.1 中所述，在分析中首先要求备选光子应满足 Loose 鉴别条件并且  $p_T > 25 \text{ GeV}$ ， $|\eta| < 1.37$  或  $1.52 < |\eta| < 2.37$ 。在所有备选光子中挑选出动量最大的两个光子（leading and subleading photon），其在量能器中的纵向信息（对非转化光子）或径迹信息（对转化光子）可以推到束流管得到其指向的顶点坐标  $z_{common}$ 。ATLAS HGam 工作组使用这一信息开发了一套基于神经网络的双光子顶点选择算法 [107]，并根据选出的顶点重新计算了双光子的方向，以减小 pile-up 的影响，提高能量分辨和双光子质量分辨。

**喷注** 分析中选用的喷注为 3.3.3 中描述的粒子流 (PFlow) 喷注，并要求其不是来自于探测器噪声、读出问题或其他非对撞信号，且  $p_T > 25 \text{ GeV}$ ， $|y| < 4.4$ 。为压低来自于事例堆积的喷注，分析中还采用了喷注顶点标记 (Jet Vertex Tagger, JVT) 和前向喷注顶点标记 (forward JVT, fJVT) 的多变量判别方法 [108]，其对硬散射喷注 (hard-scattering jet) 的效率能够达到约 90%，对 pile-up 喷注的误判为 5% 以下。

**去除重叠目标** 为避免对物理对象的重复计数，我们需要加入一些对临近对象的去除工作：

- 去除任何光子周围  $\Delta R < 0.4$  范围内的电子；

- 去除任何光子周围  $\Delta R < 0.4$  范围内的缪子；
- 去除任何光子周围  $\Delta R < 0.4$  范围内的喷注；
- 去除任何电子周围  $\Delta R < 0.2$  范围内的喷注；
- 去除任何喷注周围  $\Delta R < 0.4$  范围内的电子；
- 去除任何喷注周围  $\Delta R < 0.4$  范围内的缪子；

**事例终选** 在以上物理对象选择的基础上，最终事例选择的条件包括：

**事例触发** 事例需首先通过以下触发条件之一：

- **双光子触发 (2015+2016)**

2015 - 2016 年的数据采用的是 HLT\_g35\_loose\_g25\_loose 双光子触发条件。这一 HLT 触发使用了 L1 触发中两个能量大于 15 GeV 的电磁量能器簇团作为种子，要求其满足基于簇射形状的 Loose 判定条件，并且两个光子的能量应分别大于 35 GeV 和 25 GeV。

- **双光子触发 (2017+2018)**

相比于 2015 - 2016 年使用的触发，在 2017 - 2018 年运行中双光子触发采用的是 HLT\_g35\_medium\_g25\_medium\_L12EM20VH，其中 L1 触发中簇射团能量阈值被提高到 20 GeV，基于簇射形状的判定被提升到 Medium。这一更新是为应对新对撞环境下 pile-up 的增加，且造成了选择效率的一定降低 [109]。

- **单光子触发**

在 Run 2 中 ATLAS HGam 工作组统一加入了单光子触发 HLT\_g120\_loose (2015 - 2016) 和 HLT\_g140\_loose (2017 - 2018) [109, 110]，以提高对一些 BSM 寻找很重要的高能光子，对本分析几乎没有影响。

**好事例列表** 根据 ATLAS 记录的探测器运行状态，选择的事例应在如下列表中：

data15\_13TeV.periodAllYear\_DetStatus-v89-pro21-02  
\_Unknown\_PHYS\_StandardGRL\_All\_Good\_25ns.xml

data16\_13TeV.periodAllYear\_DetStatus-v89-pro21-01  
\_DQDefects-00-02-04\_PHYS\_StandardGRL\_All\_Good\_25ns.xml

data17\_13TeV.periodAllYear\_DetStatus-v99-pro22-01  
\_Unknown\_PHYS\_StandardGRL\_All\_Good\_25ns\_TriggerNo17e33prim.xml

data18\_13TeV.periodAllYear\_DetStatus-v102-pro22-04  
 \_Unknown\_PHYS\_StandardGRL\_All\_Good\_25ns\_TriggerNo17e33prim.xml

**好事例质量** 分析使用了标准的基于探测器质量的事例清理，这去除了有量能器报错和丢失部分探测器信息的不完整事例。

**初级顶点** 事例应有至少一个初级顶点。

**触发条件匹配** 最终选出的两个光子应与 HLT 中的光子一致。

**双光子事例** 事例应有至少两个满足初选条件的光子，并通过 Tight 鉴别 FixedCutLoose 孤立度条件。双光子事例还需要满足相对的能量要求  $p_T^{leading} > 0.35, p_T^{sub-leading} > 0.25$  和不变质量窗  $105\text{GeV} < m_{\gamma\gamma} < 160\text{GeV}$ 。

**VBF 事例** 为提高对 VBF 过程的选择效率，分析中加入了针对 VBF 的初选条件。事例中应包含至少两个喷注，且两喷注之间应有较大的赝快度差  $|\Delta\eta_{jj}| < 2.0$ ，Zeppenfeld  $\eta$  应小于  $5|\eta^{Zep}| < 5$ 。这些条件继承自 HGam 组在  $80\text{fb}^{-1}$  希格斯分析中针对 VBF 过程开发出的筛选条件。

以上条件中从事例触发至双光子事例的筛选条件已被 HGam 工作组整合进统一分析框架中，具体筛选效率见 [111]，本论文中将其整体记为“双光子事例”。VBF、ggF 和边带数据（sideband data，信号区  $120\text{GeV} < m_{\gamma\gamma} < 130\text{GeV}$  以外的全部数据）的筛选效率、剩余事例数及蒙特卡洛统计量总结如表 5-3，除 ggF 以外的其他希格斯过程的事例数总结如表 5-4。

	VBF			ggF			Sideband data	
	事例数	筛选效率	MC 统计量	事例数	筛选效率	MC 统计量	事例数	筛选效率
总计	1303.33			16563.30			22173720	
双光子事例	435.67	33.86%	2331753	5470.49	33.03%	6056327	251379	1.13%
$N_j \geq 2$	229.328	17.82%	1228478	710.82	4.29%	953783	137732	0.62%
$ \Delta\eta_{jj}  > 2$	192.227	14.94%	1027557	299.18	1.81%	394260	49970	0.23%
$ \eta^{zep}  < 5$	192.053	14.93%	1026602	296.19	1.79%	390817	49299	0.22%

表 5-3 VBF、ggF 和 sideband data 的事例数和筛选效率。

	$qq \rightarrow ZH$	$gg \rightarrow ZH$	$W^+H$	$W^-H$	$t\bar{t}H$	$b\bar{b}H$	$tHbj$	$WtH$
总计	260.22	41.88	287.27	181.88	173.38	166.05	25.60	5.18
双光子事例	76.45	15.47	79.38	55.94	59.17	59.08	7.84	1.92
$N_j \geq 2$	35.77	10.43	35.43	24.41	57.37	6.18	6.88	1.76
$ \Delta\eta_{jj}  > 2$	5.85	2.37	6.10	4.25	9.40	2.71	3.43	0.33
$ \eta^{Z\text{epp}}  < 5$	5.83	2.37	6.08	4.23	9.40	2.68	3.43	0.33

表 5-4 其他希格斯产生模式经过选择条件后的事例数。经过事例分类的最终预期事例数在表 5-8 中。

## 5.5 事例分类优化

### 5.5.1 基于多变量分析的事例分类

在之前的 VBF  $H \rightarrow \gamma\gamma$  相关分析，如希格斯的微分基准截面测量 [49] 中，研究人员曾开发过 VBF 事例的分类方法。这些研究在优化中使用 VBF 信号显著性水平作为标准，对其他希格斯过程的排除没有做特别要求，并不适合于本分析。因此我们参考之前的研究开发了一套基于多变量分析的分类方案，以期得到更高的 VBF 信号纯度，同时排除共振本底和连续本底。本分析选用的方法是提升决策树 BDT，选用的训练变量为 [49] 中使用的 VBF 相关变量： $p_T^{Hjj}$ ， $p_{Tt}$ ， $m_{jj}$ ， $\Delta\eta_{jj}$ ， $\Delta\Phi_{\gamma\gamma,jj}$ ， $\eta^{Z\text{epp}}$  和  $\Delta R_{\gamma,j}^{min}$ 。为了达到分类目标，我们单独训练了两个 BDT，以 VBF 事例为信号，分别以 ggF 和连续本底为背景。其中 VBF 和 ggF 事例的训练采用蒙特卡洛样本，连续本底分为双光子末态的不可去除本底 (irreducible background，记为  $\gamma\gamma$  过程) 和因将喷注误判为光子产生的可去除本底 (记为  $\gamma + \text{jets}$  过程)，不可去除本底使用蒙卡样本描述，可去除本底从控制区 (control region) 数据得到，二者相对比例可以用二维数据边带 (2x2D sideband) 方法得到，具体描述见 5.6.2。在做 BDT 训练时这一比例选用全部数据样本计算，得到结果为  $\gamma\gamma$  83.36%， $\gamma + \text{jets}$  16.64%。在完成所有分类后这一比例在每一类中进行了重新计算。为了在 CP 测量中不产生模型依赖，BDT 训练选用的变量保证尽量与 CPV 无关，在变量分布图 5-6 中检查了这一假设。

与 4.2 中一样，我们采用  $\langle S^2 \rangle$  来描述一个变量对信号和本底的区分度，表 5-5 列出了 7 个变量的定义和分辨能力，图 5-6 展示了 SM VBF、SM ggF、连续本底、边带数据和两个 BSM VBF 过程中这些变量的分布，图 5-7 展示了 SM VBF、SM ggF 和连续本底样本中这些变量之间以及其与  $m_{\gamma\gamma}$ ， $\mathcal{O}\mathcal{O}$  的线性相关系数。

Variable	Definition	VBF-ggF separation	VBF-yy separation
$m_{jj}$	两喷注系统不变质量	0.218	0.241
$\Delta\eta_{jj}$	两喷注赝快度差	0.152	0.219
$p_T^{Hjj}$	希格斯 + 两喷注系统的横动量	0.127	0.230
$\Delta\Phi_{\gamma\gamma,jj}$	两光子和两喷注系统的方位角差	0.120	0.186
$\Delta R_{\gamma,j}^{min}$	领头两光子和两喷注之间最小的 $\Delta R$	0.108	0.204
$\eta^{ZepP}$	$ \eta_{\gamma\gamma} - (\eta_{j1} + \eta_{j2})/2 $	0.060	0.078
$p_{T\perp}$	双光子横动量在与两个光子横动量之差垂直的方向上的投影	0.011	0.040

表 5-5 VBF 事例分类中用到的变量和其分辨能力。

从图 5-6 中，我们发现连续本底的样本和边带数据的  $\eta^{ZepP}$ 、 $\Delta R_{\gamma,j}^{min}$ 、 $\Delta\eta_{jj}$  分布存在一定差别。考虑到 SHERPA 对于软喷注的描述并不完美，这一差别是可以理解的。为修正这一差别我们对连续本底样本进行了重新加权以符合边带数据的分布，权重从  $\eta^{ZepP}$  和  $\Delta R_{\gamma,j}^{min}$  两个变量的分布中得到：对  $\Delta R_{\gamma,j}^{min}$  采用了线性加权，对  $\eta^{ZepP}$  采用二阶多项式加权。

$$w_{\eta^{ZepP}} = p_0 + p_1 |\eta^{ZepP}| + p_2 |\eta^{ZepP}|^2, \quad (5-1)$$

$$w_{\Delta R} = p_0 + p_1 \Delta R_{\gamma,j}^{min},$$

其中  $p_i$  为多项式参数，通过拟合两个变量数据与模拟样本分布的比值得到，如图 5-8。我们并没有对  $\Delta\eta_{jj}$  进行任何处理，因为在三个样本中  $\Delta\eta_{jj}$  与  $m_{jj}$  存在高度线性相关性，而  $m_{jj}$  拥有最高的信号-本底分辨能力。对  $\Delta\eta_{jj}$  进行加权修正可能导致  $m_{jj}$  描述出现偏差，从而影响到 BDT 训练结果。加权修正后 7 个变量的分布如图 5-9。

在之前的研究中，人们发现在大  $\Delta\Phi_{\gamma\gamma,jj}$  区域内 ggF + jets 过程微分截面的误差会显著增加，蒙卡样本并不完全可靠，因此在训练中为避免模型学习到这些特征， $\Delta\Phi_{\gamma\gamma,jj} > 2.94$  的事例被合并到一起（令事例  $\Delta\Phi_{\gamma\gamma,jj} = 2.94$ ），BDT 仅能读取到这一区间的总事例数。非用于训练的事例没有这一要求。

考虑到输入变量与  $m_{\gamma\gamma}$  和  $\mathcal{O}\mathcal{O}$  的低相关性，BDT 的训练并没有细分  $\mathcal{O}\mathcal{O}$  区间。两个 BDT 采用了相同的超参数进行训练：“Ntrees=850, minNodSize=2.5%, MaxDepth=3, BoostType=AdaBoost, BoostBeta=0.5”。

两个训练之间相互独立，训练出的 BDT 响应值与 ROC 曲线如图 5-10 5-11。

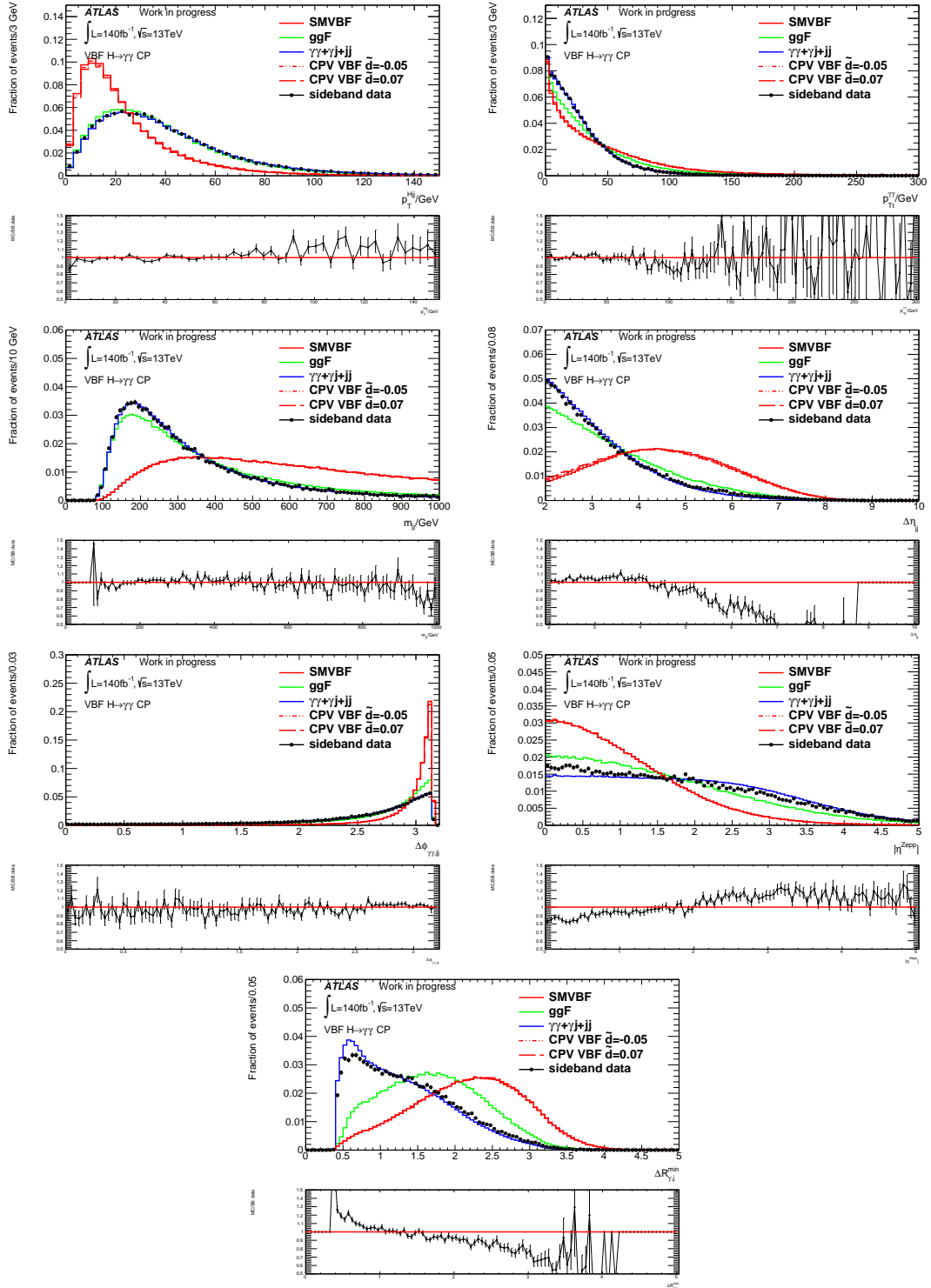


图 5-6 VBF 事例分类中所用变量的分布图。可以看出  $\eta^{Z_{epp}}$ ,  $\Delta R_{\gamma,j}^{min}$  和  $\Delta\eta_{jj}$  的分布模拟样本无法精确描述边带数据。

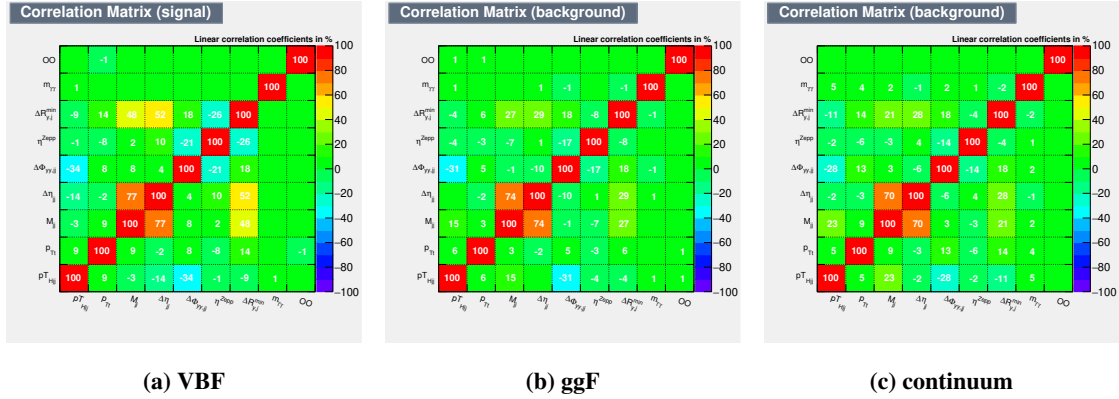


图 5-7 VBF (左)、ggF (中) 和连续本底 (右) 的变量线性相关性矩阵,  $O_0$  和  $m_{\gamma\gamma}$  也一同放入比较。图中无数据代表相关性小于 1%。

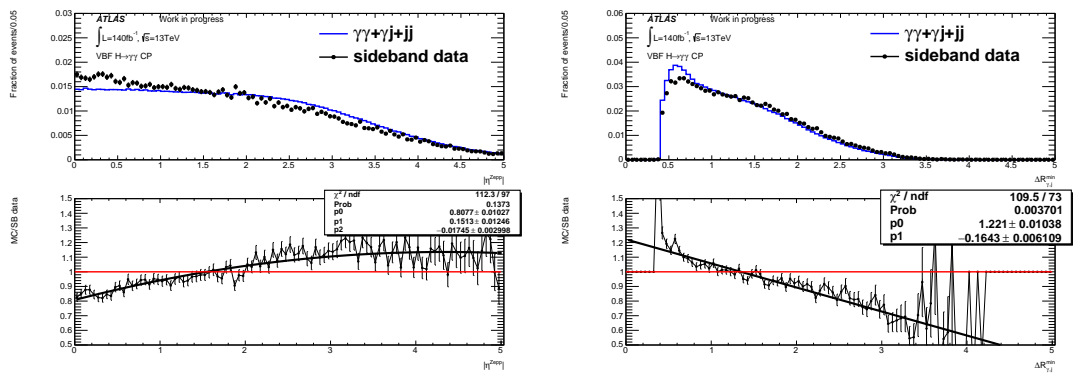


图 5-8 在  $\eta^{Zpp}$  (左) 和  $\Delta R_{\gamma\gamma}^{min}$  (右) 的分布中拟合 MC 与数据的比值进行加权修正。



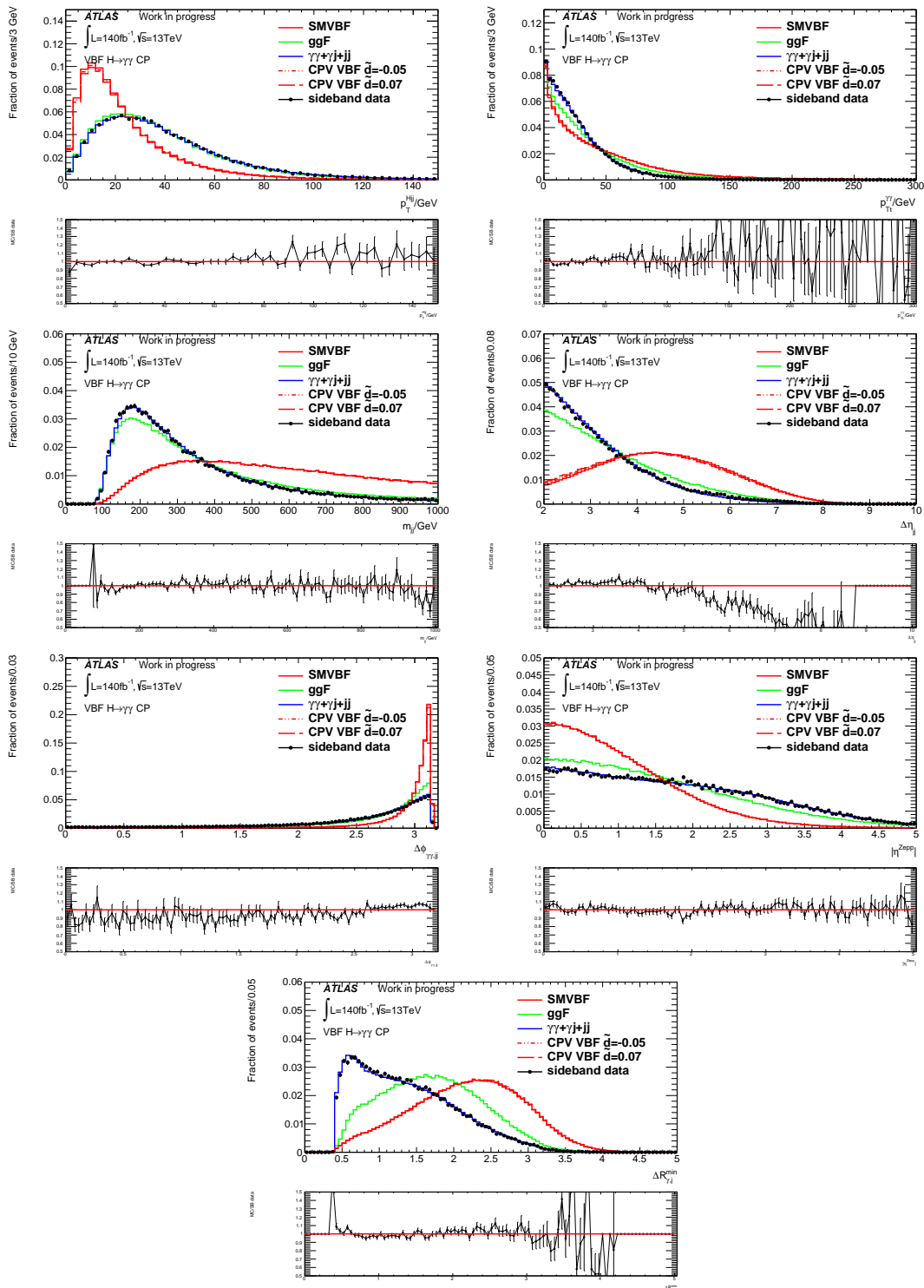


图 5-9 经过修正的事例分类变量分布图。



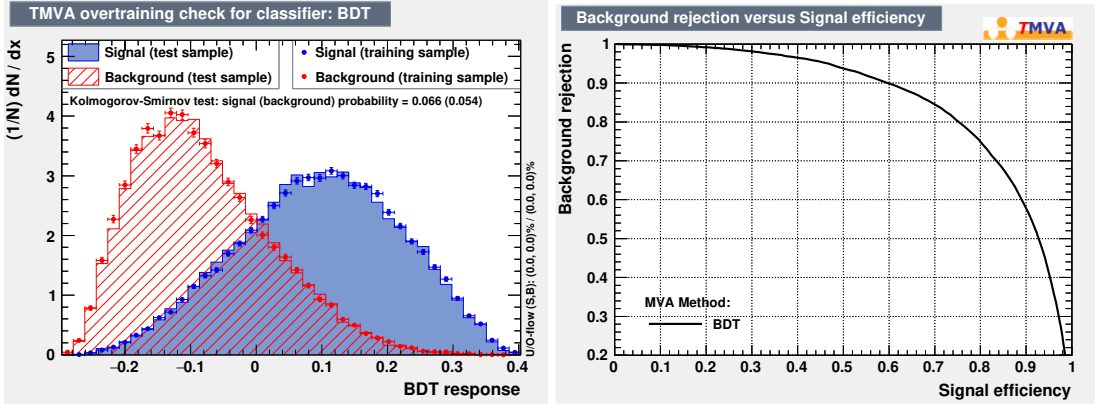
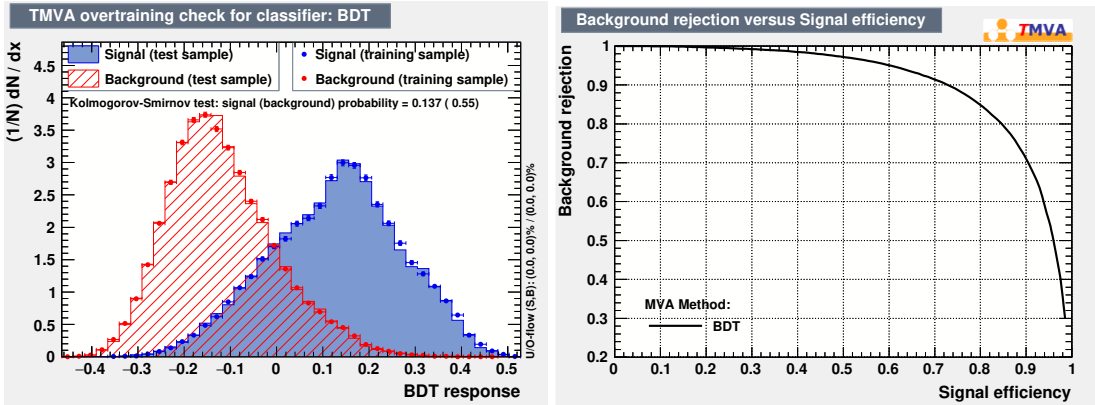


图 5-10 BDT distribution (left) and ROC curve (right) in VBF-ggF BDT training.


 图 5-11 BDT distribution (left) and ROC curve (right) in VBF-MC  $\gamma\gamma$  BDT training.

事例分类依据这两个 BDT 进行。其中  $BDT_{VBF/ggF}$  用于去除共振本底  $ggF$ ，以 VBF 过程纯度  $p = \frac{N_{VBF}}{N_{VBF} + N_{ggF}}$  为标准。从  $BDT_{VBF/ggF}$  分布不难看出，随着筛选变严 VBF 纯度会增加，但剩余事例数会不断减少，因此我们参照之前 VBF  $H \rightarrow \gamma\gamma$  分析中 Tight 区 VBF 效率 34% 选取工作点，将全部事例分为  $BDT_{VBF\_ggH\_tight}$  和  $BDT_{VBF\_ggH\_loose}$  两类，保证  $BDT_{VBF\_ggH\_tight}$  区的 VBF 效率同样为 34%。在此工作点下，相比于之前 VBF  $H \rightarrow \gamma\gamma$  分析中的 Tight 分类， $BDT_{VBF\_ggH\_tight}$  类中的 VBF 纯度和  $ggF$  去除都有了一定提升。

此后在  $BDT_{VBF\_ggH\_tight}$  和  $BDT_{VBF\_ggH\_loose}$  区中，我们使用  $BDT_{VBF/ggF}$  去除双光子连续本底。这一过程采用区域中联合 VBF 显著度为标准，其定义如 5-2，其中  $N_{vbf}^{tight}/N_{ggF}^{tight}/N_{nonRes}^{tight}$  ( $N_{vbf}^{loose}/N_{ggF}^{loose}/N_{nonRes}^{loose}$ ) 表示满足  $BDT_{VBF/\gamma\gamma} > Scanpoint$  ( $BDT_{VBF/\gamma\gamma} < Scanpoint$ ) 条件的 VBF/ $ggF$ /连续本底事例数。图 5-12 显示了在两个区域中联合显著度随  $BDT_{VBF/ggF}$  筛选标准点的变化，最终选定显著度最高处为标准点。

基于以上策略，所有事例被分为 4 类，记为 TT, TL, LT, LL。分类标准总结如表 5-13b, SM VBF 过程和边带数据的两个 BDT 的二维分布以及分类展示如图 5-13。SM VBF、SM ggF、连续本底和边带数据的两个 BDT 分布综合展示如图 5-14，四个区域中的事例数和 VBF 显著性水平以及总体联合显著性水平如表 5-6。

$$\begin{aligned}
 \sigma_{VBF}^{tight} &= \sqrt{2 \times ((N_{vbf}^{tight} + N_{ggF}^{tight} + N_{nonRes}^{tight}) \times \ln(1 + \frac{N_{vbf}^{tight}}{N_{ggF}^{tight} + N_{nonRes}^{tight}}) - N_{vbf}^{tight})}, \\
 \sigma_{VBF}^{loose} &= \sqrt{2 \times ((N_{vbf}^{loose} + N_{ggF}^{loose} + N_{nonRes}^{loose}) \times \ln(1 + \frac{N_{vbf}^{loose}}{N_{ggF}^{loose} + N_{nonRes}^{loose}}) - N_{vbf}^{loose})}, \\
 \sigma_{VBF}^{comb} &= \sqrt{(\sigma_{VBF}^{tight})^2 + (\sigma_{VBF}^{loose})^2}.
 \end{aligned} \tag{5-2}$$

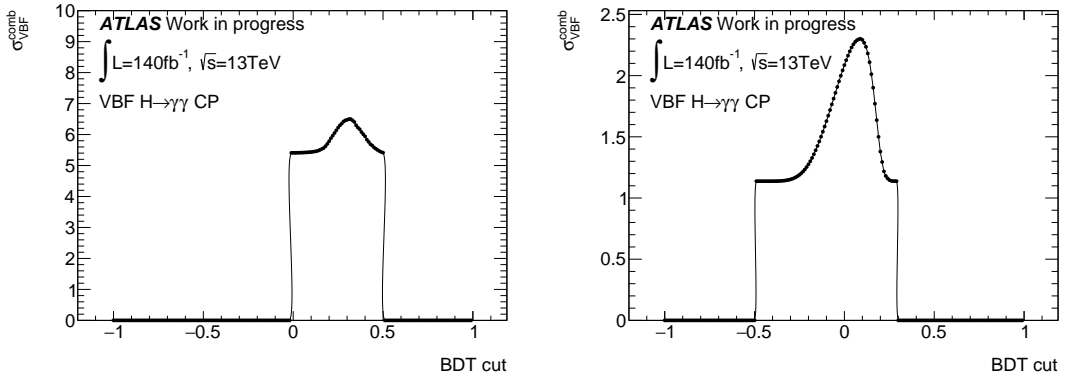
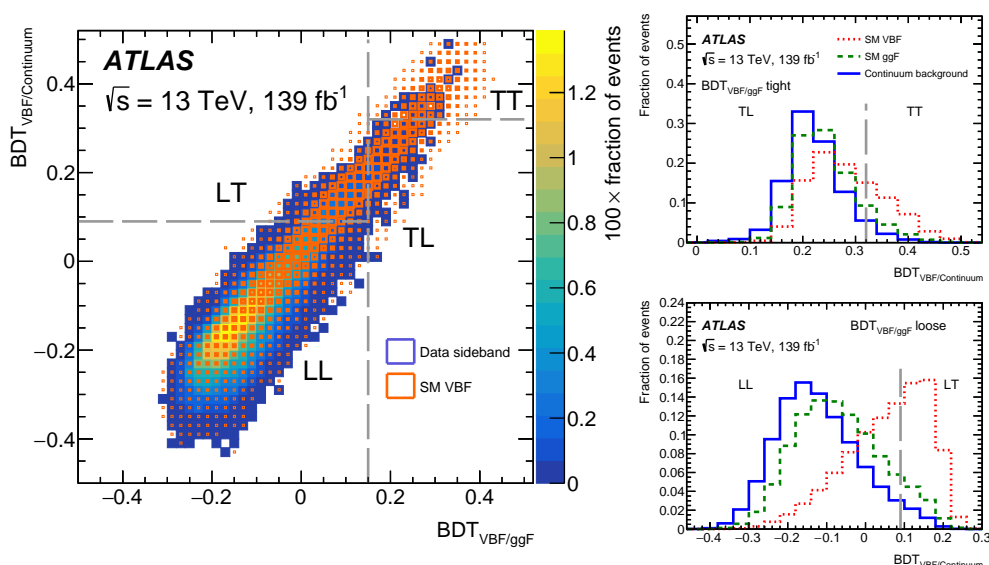


图 5-12 BDT\_VBF\_ggH\_tight (左) 和 BDT\_VBF\_ggH\_loose (右) 区域中 VBF 信号综合显著性水平随 BDT 筛选点的变化。如果在一个筛选点下某一过程预期事例数为 0 则将显著性置为 0。

### 5.5.2 $\mathcal{O}\mathcal{O}$ 分割方式优化

本分析中 CP 的测量基于  $\mathcal{O}\mathcal{O}$  谱型，由于没有可靠的理论模型来对  $\mathcal{O}\mathcal{O}$  谱型进行参数化处理，我们选择采用直方图来进行描述，而直方图分 bin 方式和精细程度将影响到 CP 测量的精度，因此我们对这一问题进行了一定研究。为平衡分析对 CP 敏感性和数据统计量，以及考虑到 SM 过程  $\mathcal{O}\mathcal{O}$  正负对称分布，我们选择采用正负对称的 6 bins。10 bins 和 20 bins 的区分方式同样进行了测试，没有观测到敏感性的显著提高。在此条件下， $\mathcal{O}\mathcal{O}$  全域的区分可以用两个参数  $p_1, p_2$



(a)

Category	Description
TT	$BDT_{VBF/ggF} \geq 0.15, BDT_{VBF/\gamma\gamma} \geq 0.32$
TL	$BDT_{VBF/ggF} \geq 0.15, BDT_{VBF/\gamma\gamma} < 0.32$
LT	$BDT_{VBF/ggF} < 0.15, BDT_{VBF/\gamma\gamma} \geq 0.09$
LL	$BDT_{VBF/ggF} < 0.15, BDT_{VBF/\gamma\gamma} < 0.09$

(b)

图 5-13 分析中四类区域的定义。

	TT	TL	LT	LL	总和
VBF	19.47	46.81	55.92	69.89	192.10
ggF	0.96	7.30	29.23	258.76	296.24
sideband	50	538	2352	46356	49296
信号区本底	12.81	137.1	608	7562	8320
显著性水平	5.25	3.89	2.21	0.79	1.76
联合显著性水平	6.94				

表 5-6 四个分类中各过程的预期事例数和 VBF 信号显著性。连续本底数归一到边带数据后从中抽取信号区 ( $m_{\gamma\gamma} \in [120, 130]$  GeV) 预期本底数, VBF 信号显著性水平则根据信号区本底进行计算。

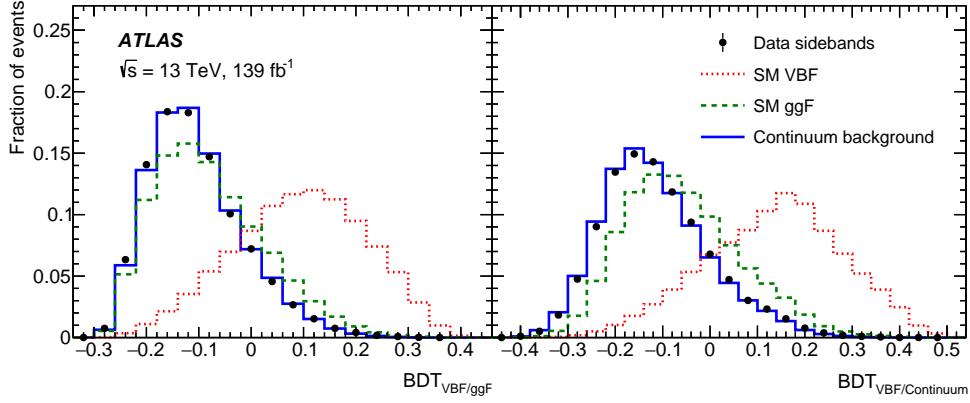


图 5-14 所有过程的两个 BDT 分布:  $BDT_{VBF/ggF}$  (左) 和  $BDT_{VBF/\gamma\gamma}$  (右)。

表示:  $(-\infty, -p_2], (-p_2, -p_1], (-p_1, 0], (0, p_1], (p_1, p_2], (p_2, \infty)$ 。我们选取 0.5 为步长扫描区分点  $p_1, p_2$ , 计算在此区分方式下 6 个 bin 中联合 VBF 显著性  $Z$ , 定义如式 ??。由于  $\mathcal{O}\mathcal{O}$  谱型会随  $CPV$  的  $\tilde{d}$  变化, 为实现在不同  $CPV$  情况下选取的区分方式都有较高 VBF 显著性, 我们使用多组  $\tilde{d} = -0.1$  到  $\tilde{d} = 0.1$  的 VBF 样本计算显著性  $Z_{\tilde{d}}$  并取平均值  $\bar{Z} = \frac{1}{N_{\tilde{d}}} \sum_{\tilde{d}=-0.1}^{0.1} Z_{\tilde{d}}$ 。各个  $p_1 p_2$  扫描点下  $\bar{Z}$  值如图 5-15。从这一扫描中我们注意到  $(p_1, p_2) = (1, 2)$  和  $(p_1, p_2) = (1.5, 5.5)$  两种组合都有最高的  $\bar{Z}$ , 但前者使得数据分割更为均匀, 在大  $\mathcal{O}\mathcal{O}$  区域能够保证较多统计量, 因此被选为最终方案。

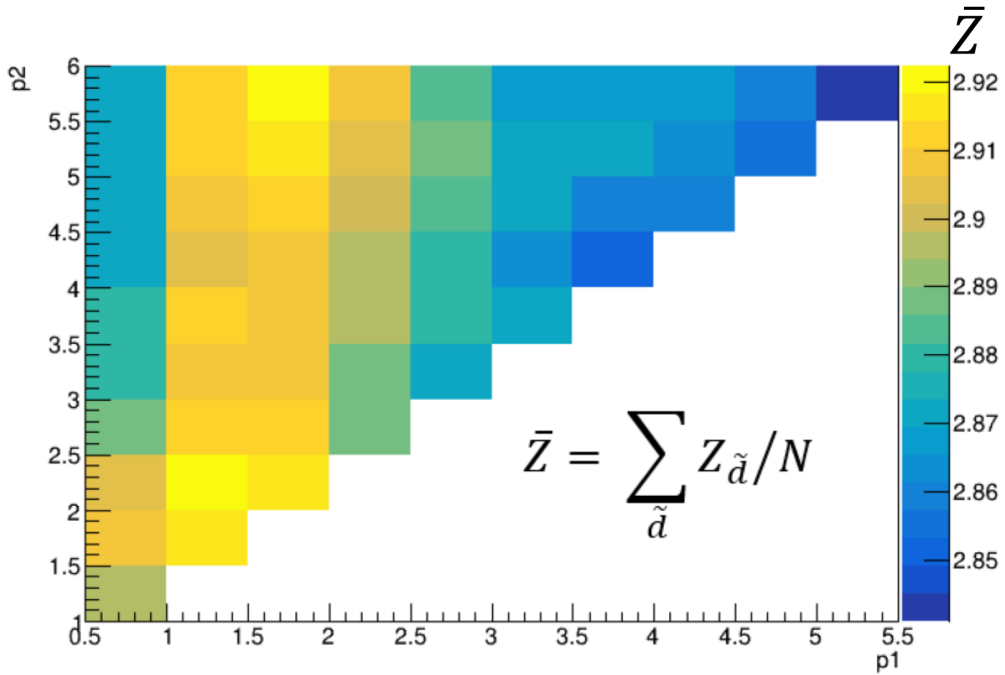


图 5-15 联合 VBF 显著性水平随  $\mathcal{O}\mathcal{O}$  分割方式的变化, 扫描步长 0.5。

最终所有事例根据两个 BDT 和  $\mathcal{O}\mathcal{O}$  分成了  $4 \times 6 = 24$  个区域, 表 ?? 中展示

了 SM VBF、SM ggF 和边带数据在这 24 个区域中的事例数，表 ?? 展示了在 24 个区域中其他希格斯过程的事例数及其所占的比例。注意到在 TT、TL 区域中这些比例小于 1%，我们认为这些过程的贡献可以忽略不计，LT 区域中这一比例作为一项统计误差进入分析中，相应过程不作考虑。LL 区域中这一比例为 5% - 10% 不可忽略，因此在后续分析中将不选用 LL 区域。

VBF	$[-\infty, -2]$	$[-2, -1]$	$[-1, 0]$	$[0, 1]$	$[1, 2]$	$[2, \infty]$	总和
TT	4.36	2.84	2.55	2.52	2.84	4.37	19.47
TL	8.27	7.24	7.89	7.87	7.27	8.25	46.80
LT	9.83	8.43	9.76	9.71	8.43	9.75	55.91
LL	11.28	10.18	13.55	13.52	10.10	11.24	69.88

(a)

ggF	$[-\infty, -2]$	$[-2, -1]$	$[-1, 0]$	$[0, 1]$	$[1, 2]$	$[2, \infty]$	总和
TT	0.13	0.12	0.21	0.22	0.15	0.13	0.96
TL	0.82	0.98	1.82	1.88	0.98	0.82	7.30
LT	4.07	3.67	6.77	6.91	3.68	4.12	29.22
LL	42.36	33.98	52.97	53.16	33.94	42.29	258.69

(b)

Side-band data	$[-\infty, -2]$	$[-2, -1]$	$[-1, 0]$	$[0, 1]$	$[1, 2]$	$[2, \infty]$	总和
TT	8	6	5	9	11	11	50
TL	70	89	115	110	88	66	538
LT	343	332	497	515	360	305	2352
LL	7418	6794	8944	8956	6746	7498	46356

(c)

表 5-7 每个分类和  $\mathcal{O}\mathcal{O}$  bin 中 VBF、ggF 预期事例数和边带数据事例数。

	SM VBF	ggF	$W^+H$	$W^-H$	$qq \rightarrow ZH$	$gg \rightarrow ZH$	$t\bar{t}H$	$b\bar{b}H$	$tHbj$	$WtH$	total Signal	Others frac
TTbin0	4.357	0.132	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0.000	4.491	0.058%
TTbin1	2.839	0.118	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	2.959	0.063%
TTbin2	2.546	0.207	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	2.756	0.091%
TTbin3	2.520	0.222	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	2.743	0.043%
TTbin4	2.835	0.149	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	2.986	0.086%
TTbin5	4.371	0.130	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.000	4.507	0.113%
TLbin0	8.265	0.818	0.006	0.004	0.004	0.004	0.003	0.005	0.021	0.000	9.131	0.523%
TLbin1	7.244	0.978	0.003	0.002	0.003	0.002	0.001	0.002	0.013	0.000	8.246	0.303%
TLbin2	7.894	1.821	0.002	0.001	0.002	0.002	0.001	0.008	0.008	0.000	9.740	0.256%
TLbin3	7.872	1.877	0.003	0.002	0.002	0.002	0.001	0.007	0.011	0.000	9.777	0.285%
TLbin4	7.268	0.982	0.002	0.002	0.002	0.002	0.001	0.002	0.018	0.000	8.279	0.348%
TLbin5	8.252	0.820	0.007	0.003	0.004	0.005	0.003	0.005	0.031	0.000	9.130	0.634%
LTbin0	9.827	4.075	0.050	0.033	0.048	0.041	0.041	0.026	0.076	0.003	14.219	2.235%
LTbin1	8.428	3.667	0.024	0.017	0.024	0.012	0.009	0.030	0.033	0.000	12.244	1.217%
LTbin2	9.761	6.767	0.025	0.017	0.023	0.013	0.007	0.035	0.030	0.000	16.678	0.896%
LTbin3	9.706	6.909	0.026	0.017	0.023	0.009	0.007	0.034	0.041	0.001	16.773	0.945%
LTbin4	8.431	3.677	0.023	0.018	0.023	0.012	0.008	0.025	0.045	0.000	12.262	1.260%
LTbin5	9.750	4.124	0.052	0.035	0.044	0.041	0.041	0.020	0.094	0.002	14.204	2.319%
LLbin0	11.269	42.375	1.281	0.837	1.187	0.680	3.294	0.366	0.768	0.105	62.161	13.702%
LLbin1	10.180	33.976	0.818	0.601	0.800	0.216	0.697	0.356	0.383	0.029	48.056	8.114%
LLbin2	13.545	52.968	0.846	0.612	0.829	0.211	0.640	0.515	0.342	0.023	70.532	5.697%
LLbin3	13.516	53.159	0.835	0.610	0.827	0.209	0.641	0.517	0.372	0.024	70.710	5.707%
LLbin4	10.101	33.936	0.814	0.592	0.803	0.223	0.684	0.360	0.369	0.030	47.912	8.088%
LLbin5	11.238	42.302	1.258	0.831	1.182	0.681	3.316	0.361	0.762	0.108	62.039	13.70%
total	192.015	296.189	6.078	4.234	5.831	2.365	9.396	2.676	3.426	0.325	522.536	

表 5-8 所有 bin 中各希格斯过程的预期事例数。最后一列为除 VBF 和 ggF 以外其他希格斯过程事例数占比。

## 5.6 信号和本底参数化

本分析通过构造  $\mathcal{O}\mathcal{O}$  分布的直方图来实现  $CP$  的测量，直方图中每个 bin 中的事例数通过拟合数据的  $m_{\gamma\gamma}$  分布的到。相比于直接对数据进行计数，拟合  $m_{\gamma\gamma}$  可以通过边带区数据估计连续本底数，从而实现更好的本底控制。这一工作需要每个区域和  $\mathcal{O}\mathcal{O}$  bin 中信号和本底的  $m_{\gamma\gamma}$  分布进行参数化。考虑到希格斯粒子自身宽度在 MeV 量级，在目前探测器分辨率条件下对  $m_{\gamma\gamma}$  分布的影响忽略不计。VBF 和 ggF 过程产生的希格斯粒子衰变出的双光子质量谱没有区别，两个过程可以用同样模型描述，具体见信号参数化部分 5.6.1。连续本底采用连续下降函数进行描述，具体函数选择采用赝信号 (spurious signal) 方法，具体见本底参数化部分 5.6.2。

### 5.6.1 信号参数化

$H \rightarrow \gamma\gamma$  过程产生会在  $m_{\gamma\gamma}$  分布上产生一个共振峰。在不考虑希格斯粒子自身宽度情况下，这一分布由光子的探测器响应主导。在 ATLAS  $H \rightarrow \gamma\gamma$  分析中普遍选用双边水晶球函数 (Double-side crystal ball function) 进行描述，该函数包括一个高斯核心和两个指数函数边缘部分，具体函数形式如下：

$$f_{\text{DSCB}}(m_{\gamma\gamma}) = N \times \begin{cases} e^{-t^2/2} & \text{if } -\alpha_{\text{low}} \leq t \leq \alpha_{\text{high}} \\ \frac{e^{-\frac{1}{2}\alpha_{\text{low}}^2}}{\left[\frac{1}{R_{\text{low}}}(R_{\text{low}} - \alpha_{\text{low}} - t)\right]^{n_{\text{low}}}} & \text{if } t < -\alpha_{\text{low}} \\ \frac{e^{-\frac{1}{2}\alpha_{\text{high}}^2}}{\left[\frac{1}{R_{\text{high}}}(R_{\text{high}} - \alpha_{\text{high}} + t)\right]^{n_{\text{high}}}} & \text{if } t > \alpha_{\text{high}} \end{cases} \quad (5-1)$$

其中  $N$  是归一化常数， $t = (m_{\gamma\gamma} - \mu_{\text{CB}}) / \sigma_{\text{CB}}$ ， $\mu_{\text{CB}}$  和  $\sigma_{\text{CB}}$  是高斯核心的中心值和宽度； $\alpha_{\text{high}}$ ， $\alpha_{\text{low}}$  是函数高低两端核心和边缘部分的边界， $n_{\text{high}}$ ， $n_{\text{low}}$  是两侧指数函数的指数部分， $R_{\text{low}} = \frac{n_{\text{low}}}{\alpha_{\text{low}}}$ ， $R_{\text{high}} = \frac{n_{\text{high}}}{\alpha_{\text{high}}}$ 。

在构造统计模型时，每个区域的共振  $m_{\gamma\gamma}$  分布均采用此函数描述，但函数参数固定到此区域中对 VBF+ggF 蒙卡样本的最佳拟合值。实验误差对信号分布的影响主要考虑光子能量刻度和能量分辨的影响，通过滋扰参数 (nuisance parameter) 方法贡献到  $\mu_{\text{CB}}$  和  $\sigma_{\text{CB}}$  中，具体见 5.7.2.2。

### 5.6.2 本底参数化

在双光子分析中，连续本底过程以标准模型的光子对过程 ( $\gamma\gamma$ ) 为主导，同时包含由于将其他物理对象，通常由很多中性成分的喷注，误判为光子的  $\gamma + \text{jets}$

过程。因此本底参数化的工作包含以下几步：

- 计算  $\gamma\gamma$  和  $\gamma + jets$  两部分的相对比例。
- 根据比例混合  $\gamma\gamma$  蒙卡和用于描述  $\gamma + jets$  分布的控制区数据，构造完整本底样本。
- 使用赝信号方法或 Wald 测试法选择每个区域内的本底模型函数，并计算模型系统误差。

**本底拆解** 背景拆解采用的是二维边带区法(2×2D sideband decomposition method) [112]，选用的数据相比于 5.4 中分析用数据去除了光子孤立度要求，并将光子鉴别要求进行了一定放松，即在 Tight 鉴别中去除了对  $w_{s3}$ 、 $f_{side}$  和  $\Delta E_s$  的要求，记为 Loose'-3 或  $L'$ 。每个光子根据是否通过 Tight 孤立度和鉴别要求可分成 4 类，双光子事例则可分为 16 个相互正交类，每个区域中的事例数可以用光子鉴别和孤立度的效率、误判率 and 对应各成分的事例数表示。联合求解 16 个区域可得各组分比例。本分析中这一计算在每个 BDT 分类和  $\mathcal{O}\mathcal{O}$  bin 中进行，表 5-9 列出了  $\gamma\gamma$  组分比例  $f_{\gamma\gamma}$ ， $\gamma + jets$  组分比例为  $1 - f_{\gamma\gamma}$ 。

	TT	TL	LT
$\mathcal{O}\mathcal{O}$ bin1	$0.83 \pm 0.05$	$0.78 \pm 0.02$	$0.86 \pm 0.01$
$\mathcal{O}\mathcal{O}$ bin2	$0.86 \pm 0.45$	$0.80 \pm 0.07$	$0.78 \pm 0.04$
$\mathcal{O}\mathcal{O}$ bin3	$0.78 \pm 0.25$	$0.84 \pm 0.06$	$0.83 \pm 0.03$
$\mathcal{O}\mathcal{O}$ bin4	$0.84 \pm 0.11$	$0.84 \pm 0.06$	$0.83 \pm 0.04$
$\mathcal{O}\mathcal{O}$ bin5	$0.53 \pm 0.37$	$0.87 \pm 0.05$	$0.86 \pm 0.04$
$\mathcal{O}\mathcal{O}$ bin6	$0.88 \pm 0.15$	$0.86 \pm 0.04$	$0.88 \pm 0.02$

表 5-9 二维边带区法测量得到的  $\gamma\gamma$  组分分数  $f_{\gamma\gamma}$ ， $\gamma + jets$  组分分数为  $1 - f_{\gamma\gamma}$ 。

**本底样本构造**  $\gamma + jets$  过程的  $m_{\gamma\gamma}$  分布从控制区 (CR) 数据得到，这个控制区的定义为至少一个光子没有通过 Tight ID 和 Isolation 要求，其中混入的  $\gamma\gamma$  过程事例可忽略不计。得到的控制区数据统计量较低，为减小统计误差，我们将  $\gamma\gamma$  蒙卡样本以线性权重加权到  $\gamma + jets$  样本上，得到高统计量  $\gamma + jets$  样本的  $m_{\gamma\gamma}$  分布。最后两部分本底样本按比例进行混合。与边带数据的比较显示这种本底样本构造方法可以很好地描述边带数据。图 5-16 以一个区域 (TL  $\mathcal{O}\mathcal{O}$  bin 3) 为例展



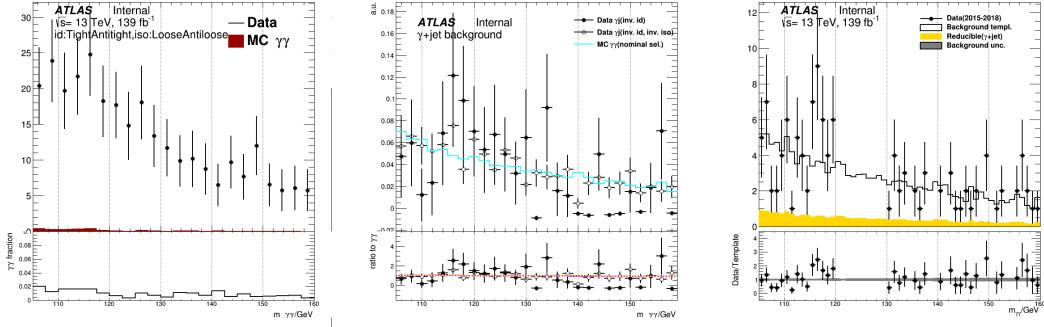


图 5-16 本底样本构造过程，以 TL  $\mathcal{O}\mathcal{O}$  bin 3 为例。从左至右：控制区数据得到的  $\gamma$  + jets 过程（红色显示其中混入的  $\gamma\gamma$  过程）；将高统计量  $\gamma\gamma$  蒙卡样本加权得到控制区数据  $\gamma$  + jets 过程  $m_{\gamma\gamma}$  分布；按照比例  $f_{\gamma\gamma}$  将两部分混合后得到的最终本底样本，与带边数据对比。

示了本底构造过程中每一步的  $m_{\gamma\gamma}$  分布。

图中的本底样本仍然有一些统计涨落，为进一步去除其影响我们采用了高斯回归方法（Gaussian Process Regression, GPR）[113] 进行平滑处理。本底过程中每个  $m_{\gamma\gamma}$  bin 中的数据量可以看作一个高斯过程，如果我们认为本底  $m_{\gamma\gamma}$  分布应该是平滑下降的，那就可以从整个数据集推测每个 bin 中的数据数，从而使得分布变得平滑。这一方法在  $H \rightarrow \gamma\gamma$  分析中使用的可行性已在之前分析中验证过。但注意到若样本统计量过低，经过平滑后的样本相对于原样本会有一些偏差，因此分析中仅蒙卡样本中每个  $m_{\gamma\gamma}$  bin 统计量大于 10 的分类进行 GPR 平滑操作，即 TL、LT 区域的所有  $\mathcal{O}\mathcal{O}$  bin。TT 区域的本底样本由于统计量过低不进行平滑处理。图 5-17 同样以 TL  $\mathcal{O}\mathcal{O}$  bin 3 为例展示了平滑后的  $m_{\gamma\gamma}$  分布。

**赝信号方法进行本底模型选择** 连续下降的本底可以用多种函数来描述， $H \rightarrow \gamma\gamma$  分析中通常采用赝信号（spurious signal）方法来决定，即在一个连续本底上进行信号共振峰 + 本底的拟合。选用不同本底函数拟合会得到不同的赝信号数  $S$  的拟合结果，在通过拟合优度和针对  $S$  的条件后拥有最小自由度的函数即被选为本底模型， $S$  则被视为选择这一模型引入到信号数上的模型误差。在这一方法中考虑的函数包括：

- 一阶指数函数 (Exp):  $e^{c \cdot m_{\gamma\gamma}}$
- 二阶指数多项式 (ExpPoly2):  $e^{c_1 \cdot m_{\gamma\gamma}^2 + c_2 \cdot m_{\gamma\gamma}}$
- 三阶指数多项式 (ExpPoly3):  $e^{c_1 \cdot m_{\gamma\gamma}^3 + c_2 \cdot m_{\gamma\gamma}^2 + c_3 \cdot m_{\gamma\gamma}}$
- 3 - 5 阶 Bernstein 多项式:  $B_N(m_{\gamma\gamma}) = \sum_{i=0}^N c_i \cdot b_{i,N}$  with  $b_{i,N} = \binom{N}{i} m_{\gamma\gamma}^i (1 - m_{\gamma\gamma})^{N-i}$

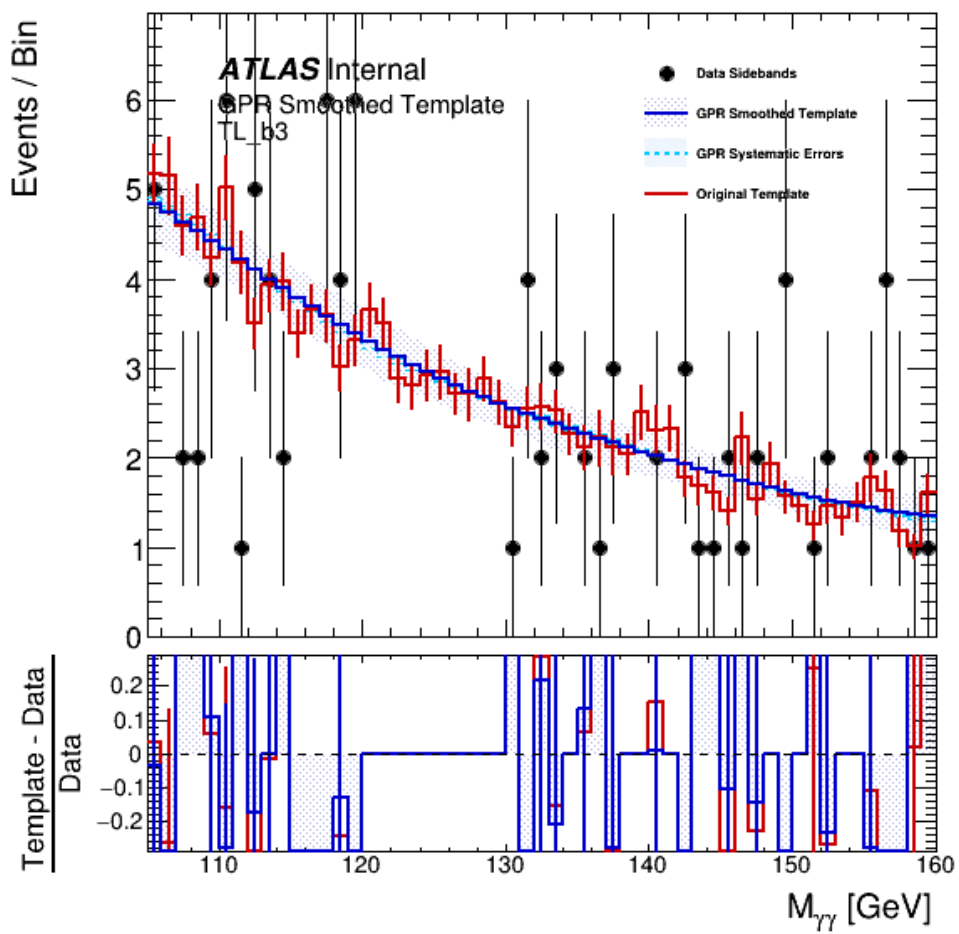


图 5-17 TL  $00$  bin 3 中经过平滑后的  $m_{\gamma\gamma}$  分布 (蓝线)、原本底样本分布 (红线) 和边带数据分布。

$m_{\gamma\gamma})^{N-i}$ ,  $N=3, 4, 5$

- 一阶幂函数:  $m_{\gamma\gamma}^c$

在拟合中, 信号模型的峰值以 0.5 GeV 为步长从 121 GeV 扫到 129 GeV,  $S$  取这组拟合中得到的最大的赝信号数。对  $S$  的要求为至少满足以下之一:

- $S < 10\%N_{s,exp}$ ,  $N_{s,exp}$  是在此区域中通过信号蒙卡得到的预期希格斯信号事例数,  $\mu_{sp} = S_{spur}/N_{s,exp}$ ;

- $S < 20\%\sigma_{bkg}$ ,  $\sigma_{bkg}$  是拟合中  $S$  的统计误差,  $Z_{sp} = S_{spur}/\sigma_{bkg}$

另外, 选中函数对样本的纯本底模拟应给出合理的  $\chi^2$  值, 要求按  $\chi^2$  分布, 得到此  $\chi^2$  值的概率  $P(\chi^2)$  大于 1%。

以上赝信号方法在有充足数据统计量时有较好表现, 而当数据统计量较低时, 数据的统计涨落会很明显, 赝信号方法会倾向于选择更高自由度的函数, 但其实这并无必要, 也可能引入一定模型偏差。基于这一考虑, 我们人为选定边带数据数量大于 100 的区域 (LT 所有  $\mathcal{O}\mathcal{O}$  bin) 采用上述赝信号方法, 边带数据量小于 100 的区域 (TT, TL 的所有  $\mathcal{O}\mathcal{O}$  bin) 采用 Wald 测试法。在这一方法中, 函数模型的选择被限制在 1, 2, 3 阶指数多项式中:

$$f_i(x) = e^{\sum_{j=0}^i a_j x^{j+1}} \quad (5-2)$$

对三个函数的拟合计算统计量  $\lambda_{(i,i+1)} = -2\ln(L_i/L_{i+1})$ ;  $i = 0, 1$ , 其中  $L_i$  为用  $i+1$  阶指数函数拟合边带数据的似然函数值, 统计量  $\lambda$  应服从自由度为 1 的  $\chi^2$  分布。Wald 测试从  $i = 0$  开始, 若  $\lambda_{i,i+1}$  的 P 值  $p(\lambda_{i,i+1})$  小于 5%, 证明  $i$  阶指数函数不足以描述边带数据分布, 则考虑  $i+1$  阶函数, 直到  $p(\lambda_{i,i+1}) > 5\%$ 。图 5-18 展示了 TL  $\mathcal{O}\mathcal{O}$  bin 3 的 Wald 测试结果。在利用 Wald 测试方法完成本底模型选择后, 选中函数同样进行了赝信号拟合, 以得到作为本底模型误差的赝信号数。最终结果总结如表 5-10。

Category	function	maxS	$\mu_{sp}[\%]$	$Z_{sp}[\%]$	$P(\chi^2)[\%]$	Wald p-value
TT cate. $\mathcal{O}\mathcal{O}$ bin1	Exp.	-0.264	-5.95	-26.4	7.9	62.7%
TT cate. $\mathcal{O}\mathcal{O}$ bin2	Exp.	0.261	8.88	22.5	9.05	45.1%
TT cate. $\mathcal{O}\mathcal{O}$ bin3	Exp.	0.135	5.14	14	93.4	81.0%
TT cate. $\mathcal{O}\mathcal{O}$ bin4	Exp.	0.254	9.3	19.2	7.01	65.0%
TT cate. $\mathcal{O}\mathcal{O}$ bin5	Exp.	0.537	18.2	32.6	34.2	97.9%
TT cate. $\mathcal{O}\mathcal{O}$ bin6	Exp.	0.547	12.3	35.1	35.1	64.4%
TL cate. $\mathcal{O}\mathcal{O}$ bin1	Exp.	0.362	3.96	10.1	1.62	40.2%
TL cate. $\mathcal{O}\mathcal{O}$ bin2	Exp.	-0.386	-4.74	-10.1	4.2	6.40%
TL cate. $\mathcal{O}\mathcal{O}$ bin3	Exp.	-0.403	-4.18	-9.01	23.7	45.3%
TL cate. $\mathcal{O}\mathcal{O}$ bin4	Exp.	-0.374	-3.88	-8.7	4.07	31.5%
TL cate. $\mathcal{O}\mathcal{O}$ bin5	Exp.	0.162	2.33	3.79	0.94	20.8%
TL cate. $\mathcal{O}\mathcal{O}$ bin6	Exp.	0.17	2.2	5.04	0.39	65.6%
LT cate. $\mathcal{O}\mathcal{O}$ bin1	Exp.	0.217	1.74	3	46.1	
LT cate. $\mathcal{O}\mathcal{O}$ bin2	Exp.	-0.397	-3.63	-5.34	7.78	
LT cate. $\mathcal{O}\mathcal{O}$ bin3	Pow.	1.02	6.17	11.1	11.5	
LT cate. $\mathcal{O}\mathcal{O}$ bin4	Exp.	-1.26	-7.61	-11.8	79.3	
LT cate. $\mathcal{O}\mathcal{O}$ bin5	Exp.	1.38	11.3	15.8	82.5	
LT cate. $\mathcal{O}\mathcal{O}$ bin6	Exp.	0.476	3.88	6.84	23.5	

表 5-10 最终决定的本底模型和对应的赝信号误差。在 TL  $\mathcal{O}\mathcal{O}$  bin 5 和 TL  $\mathcal{O}\mathcal{O}$  bin 6 中  $P(\chi^2) < 1\%$ ，在经过 toy 样本检查后确认为统计涨落，选定函数可以描述本底分布。

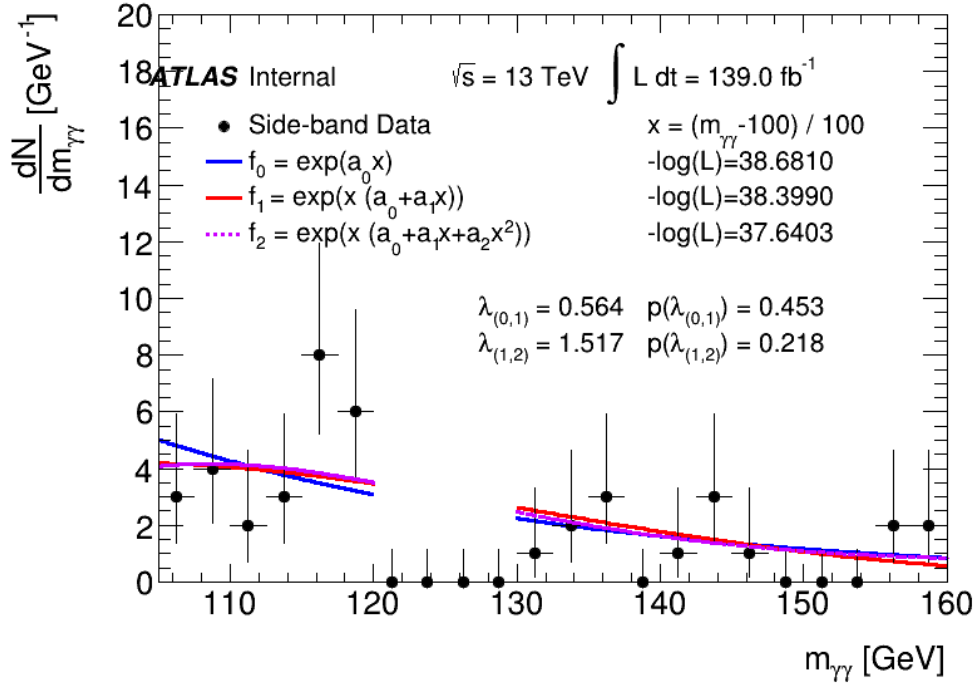


图 5-18 TL 00 bin 3 中的 Wald 测试结果。

## 5.7 系统误差分析

分析中需要考虑的误差主要分为理论误差和实验误差两部分，通过影响信号事例数或信号模型参与到统计模型中，影响最终结果的测量。本节讨论这些系统误差来源，统计模型的搭建见 5.8。

### 5.7.1 理论误差

由于分析中对 VBF 和 ggF 过程事例数的估计来自于蒙卡模拟，对相关截面的理论计算会影响这一结果，其中主要包括产生子中未考虑的高阶修正、PDF 选取和  $\alpha_S$  的不确定性。同时，对末态部分子簇射的描述会影响到末态动力学分布，从而影响特定相空间中的事例数。理论误差在统计模型中作用到 VBF 和 ggF 过程的信号事例数上。

**QCD 微扰误差** 微扰 QCD 计算中缺失高阶项引起的误差可以通过改变归一化因子和形状因子获得，对应权重储存在产生子文件中。本分析中微扰 QCD 误差的计算沿用  $H \rightarrow \gamma\gamma$  简化模板微分截面 (Simplified Template Cross Section, STXS) 分析中的框架，因此在除了截面的全局误差外，还包括不同相空间中事例迁移的误差。因此对 ggF 过程，共考虑了 17 (迁移) +1 (全局) 个相关误差项；VBF 过程则有 9+1 个误差项。

**PDF 和  $\alpha_S$  误差** PDF 和  $\alpha_S$  的误差从产生子中的权重估计。PDF4LHC15 中包含 30 个对应于 PDF 中参数的误差项，调用不同的权重可以得到相应误差项的影响。分析中还有两组使用了不同  $\alpha_S$  值的 PDF，对应于  $\alpha_S = 0.118 \pm 0.0015$ ，即  $\alpha_S$  值 68% 置信度的误差范围，这两项作为  $\alpha_S$  误差参与到分析中。另外在计算  $\mathcal{O}\mathcal{O}$  时曾使用了 CT10 PDF 来考虑初末态部分子的味组合，图 5-19 展示了 SM VBF 过程中用 CT10 和 PDF4LHC 两组 PDF 计算  $\mathcal{O}\mathcal{O}$  的差别，基于此我们认为这一误差相比于其他 PDF 误差可忽略，没有进行特别处理。

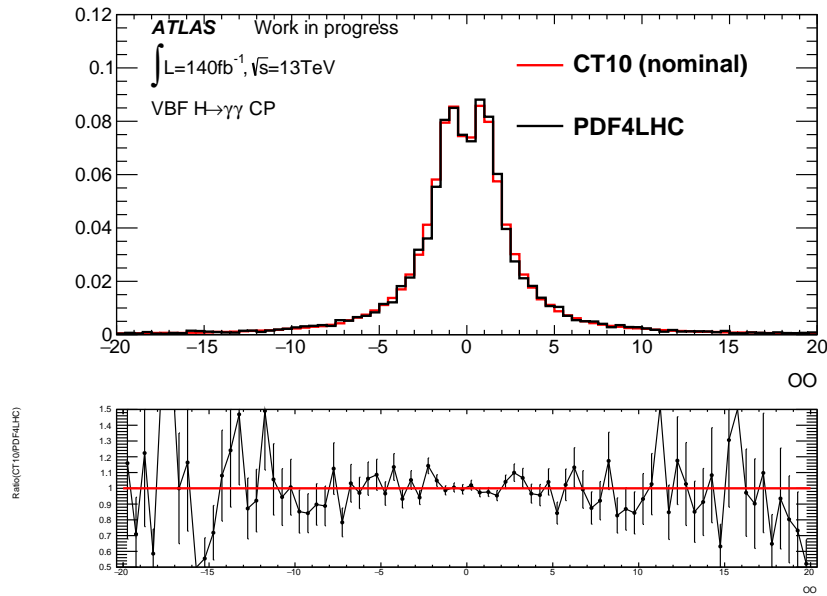


图 5-19 VBF 过程中采用不同 PDF 计算  $\mathcal{O}\mathcal{O}$  得到的分布。从图中可认为 PDF 的选择对  $\mathcal{O}\mathcal{O}$  计算没有显著影响。

**部分子簇射误差** 部分子簇射的误差通过选用不同模型进行考虑，如 5.3 所述，我们分别使用 PYTHIA 和 HERWIG 产生了 SM VBF 和 SM ggF 过程的蒙卡样本，两组样本之间的差别被视为部分子簇射误差。

图 5-20 5-21 中展示了 3 个 BDT 分类 6 个  $\mathcal{O}\mathcal{O}$  bin 中的理论误差，ggF 过程总理论误差为 10% - 23%，VBF 过程总理论误差为 2% - 5%。

### 5.7.2 实验误差

实验误差根据作用到统计模型中方式的不同，可以分为两类：影响信号事例数的误差和影响信号模型的误差。

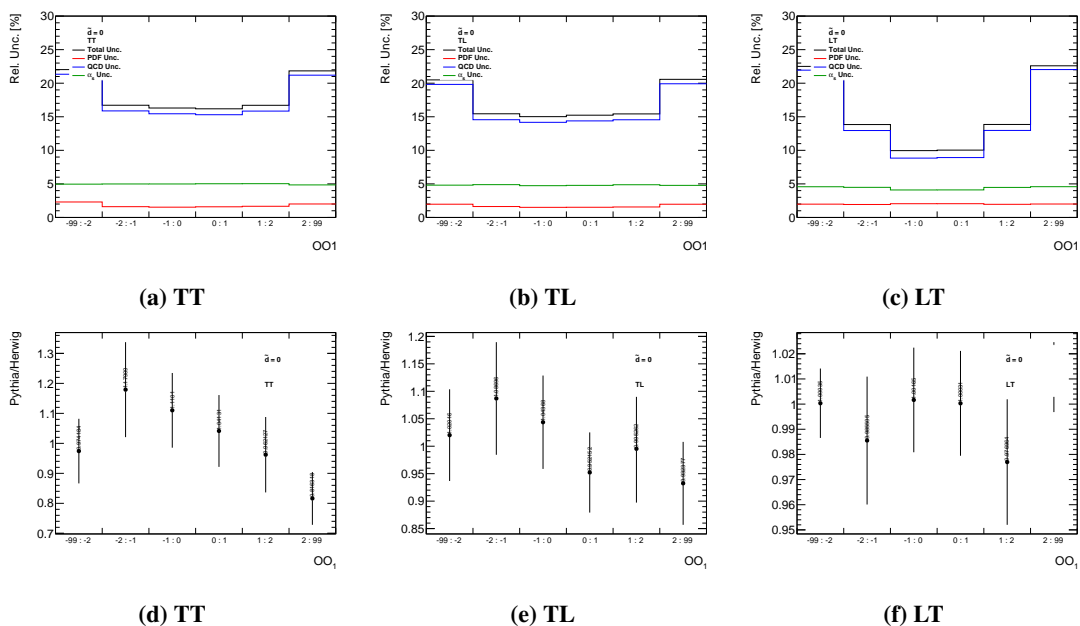


图 5-20 ggF 过程不同区域和  $\mathcal{OO}$  bin 中的 PDF、QCD、 $\alpha_s$  误差 (上) 和分子簇射误差 (下)。

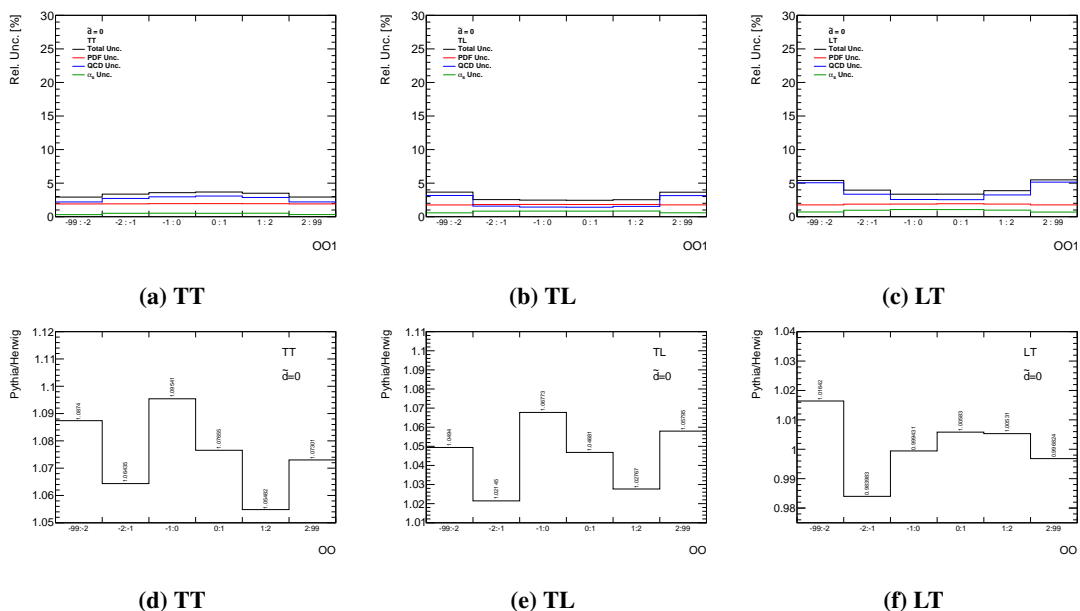


图 5-21 VBF 过程不同区域和  $\mathcal{OO}$  bin 中的 PDF、QCD、 $\alpha_s$  误差 (上) 和分子簇射误差 (下)。

### 5.7.2.1 信号事例数误差

**积分亮度误差** ATLAS 实验的积分亮度由 LUCID-2 探测器 [23] 测量，2015 - 2018 年期间积分亮度的误差为 1.7%。

**事例堆积误差** 蒙特卡洛样本中对事例堆积效应的模拟是通过对非弹性散射过程进行重新加权估计的，这会对事例数的估计产生一定影响，在不同过程不同区域中不等。

**选择效率误差** 包括双光子触发效率的误差、光子鉴别和孤立度的选择效率的误差、喷注重建效率的误差。这些误差由 ATLAS 相关工作组提供，在分析中将蒙卡样本中的物理对象按正负一倍标准差修正，比较修正前后特定区域内的事例数即可得到对应项的误差。在光子鉴别效率误差的估计中，另加入了一项 2% 的误差用于弥补快模拟样本和全模拟样本之间的差别。

**光子和喷注能量刻度和分辨的误差** 光子和喷注能量刻度和分辨主要影响其能量的重建，通过筛选条件间接作用到特定区域的事例数上。其中光子能量刻度和分辨的误差相比于其他系统误差项来说较小，因此分别被整和到一个系统误差项中 (EG\_RESOLUTION\_ALL, EG\_SCALE\_ALL)。

**本底模型误差** 如 5.6.2 中所述，本底模型相关的误差统一吸收到赝信号数中，在每个  $\mathcal{O}\mathcal{O}$  bin 中加入一项对称的作用于信号数的误差。

**TT 区域中 ggF 蒙卡样本统计误差** 由于 TT 区域对 VBF 相关筛选非常严格，ggF 蒙卡样本仅有少部分进入此区域，蒙卡自身的统计误差不可忽略。因此在此区域的 ggF 样本上加入了蒙卡统计误差项，其他区域和样本中此类误差忽略不计。

**LT 区域中来自于其他希格斯产生过程的误差** 如表 5-8 所示，在 LT 区域中其他希格斯过程占比为 1% - 3%，相比于 TT、TL 无法直接忽略，因此这一比例作为 VBF 过程事例数的误差考虑进统计模型中。

**SMEFT Warsaw 基中 k-factor 方法引入的误差** 在测量  $c_{H\bar{W}}$  的模型中，k-factor 来自于另外产生的 MADGRAPH5\_aMC@NLO 样本中 BSM 和 SM 权重的比值，其计算会在  $\mathcal{O}\mathcal{O}$  bin 事例数上引入一项误差，大致在 O(1%) 量级。



### 5.7.2.2 信号模型误差

如 5.6.1 中所述，信号模型采用双边水晶球描述，其峰值  $\mu_{CB}$  和高斯核宽度  $\sigma_{CB}$  会受光子能量刻度、能量分辨和希格斯粒子质量测量误差的影响，而函数中的其他参数被认为不受影响。ATLAS EGamma 工作组提供了 40 项光子能量刻度误差和 9 项能量分辨误差，通过比较考虑误差后双光子谱平均值的变化  $\delta\mu_{CB}^{\pm 1\sigma} = \frac{\langle m_{\gamma\gamma}^{\pm 1\sigma} \rangle}{\langle m_{\gamma\gamma}^{nom} \rangle} - 1$  和宽度的变化  $\delta\sigma_{CB}^{\pm 1\sigma} = \frac{S^{\pm 1\sigma}}{S^{nom}} - 1$  来考虑其对信号模型的影响。

分析中采用的希格斯粒子质量测量结果为  $125.09 \pm 0.24$  GeV [114]，因此每个分类中会在质量峰  $\mu_{CB}$  上引入此项 0.19% 的误差。其他因素，如喷注相关误差已证明对信号模型的影响可忽略不计。

## 5.8 统计模型

分析采用极大似然法测量  $CP$  相关参数  $\tilde{d}$  和  $c_{H\tilde{W}}$ 。似然函数构建如下：

$$\mathcal{L}(\mu_{vbf}, N_{bkg}, \xi_{bkg}, \theta | m_{\gamma\gamma}) = \prod_i^{N_{bins}} Pois(n_i | v_i(\mu_{vbf}, \theta)) \prod_{j=1}^{n_i} f_i(m_{\gamma\gamma}^j; \theta) \prod_k G(\theta_k)$$

$$f_i(m_{\gamma\gamma}) = \frac{1}{v_i} \left[ (\mu_{vbf} \times N_{vbf}^i(\theta) + N_{ggF}^i(\theta) + N_{spur}^i \times \theta_{spur}^i) \times f_{sig}(m_{\gamma\gamma}; \mu_{CB}^i(\theta), \sigma_{CB}^i(\theta), \xi_{CB}^i) + N_{bkg}^i \times f_{bkg}(m_{\gamma\gamma}; \xi_{bkg}^i) \right]$$

$$v_i = \mu_{vbf} \times N_{vbf}^i(\theta) + N_{ggF}^i(\theta) + N_{spur}^i \times \theta_{spur}^i + N_{bkg}^i$$

其中：

- $\mu_{vbf}$  是 VBF 过程信号强度，是统计模型中的感兴趣参数 (Parameter of interest, POI)，在 18 个区域中共享 (不是分析的 POI)。

- $N_{bins}$  是所有区域数量 (共计 18 个)， $n_i$  是每个区域中观测到的事例数。 $\theta$  是滋扰参数，每个系统误差项对应一个  $\theta$ ，在拟合中浮动。

- $f_i(m_{\gamma\gamma})$  是双光子不变质量谱的模型， $f_{sig}$  为 5.6.1 中描述的双边水晶球函数，其中参数  $\alpha_{low}$ ,  $\alpha_{high}$ ,  $n_{low}$ ,  $n_{high}$  固定到对蒙卡样本拟合值， $\mu_{CB}^i(\theta)$  和  $\sigma_{CB}^i(\theta)$  受  $\theta$  影响。 $f_{bkg}$  为在 5.6.2 中选定的函数，其中参数  $\xi_{bkg}$  均保持浮动。

- $v_i$  是信号和本底的总期望事例数， $N_{VBF}^i(\theta)$  和  $N_{ggF}^i(\theta)$  是考虑滋扰参数的 VBF 和 ggF 信号事例数， $N_{spur}^i$  是赝信号事例数， $N_{bkg}$  是连续本底数。其中  $N_{bkg}$  在拟合中浮动。

滋扰参数通过以下方式进入到  $N_{VBF}^i(\theta)$ ,  $N_{ggF}^i(\theta)$ ,  $\mu_{CB}^{nom}$  和  $\sigma_{CB}^{nom}$ ,  $\xi_{CB}$  中：

$$N_{VBF}^i(\theta) = N_{VBF}^{i,nom} \times \prod_k Resp_k(\theta_k) \quad , \quad N_{ggF}^i(\theta) = N_{ggF}^{i,nom} \times \prod_k Resp_k(\theta_k)$$

$$\mu_{CB}^i(\theta) = \mu_{CB}^{i,nom} \times \prod_k Resp_k(\theta_k) \quad , \quad \sigma_{CB}^i(\theta) = \sigma_{CB}^{i,nom} \times \prod_k Resp_k(\theta_k)$$

其中  $N_{VBF}^{nom}$  和  $N_{ggF}^{nom}$  为蒙卡预测事例数， $\mu_{CB}^{nom}$  和  $\sigma_{CB}^{nom}$  为对信号样本的最佳拟合值。系统误差的滋扰参数通过相应函数  $Resp(\theta)$  作用到统计模型的特定项上， $G(\theta)$  为对滋扰参数的高斯约束。若认为误差服从高斯分布，响应函数  $Resp(\theta) = F_G(\sigma, \theta) = (1 + \sigma\theta)$ ， $\sigma$  为误差幅度。若认为误差满足 log-normal 分布，响应函数选用  $Resp(\theta) = F_{LN}(\sigma, \theta) = e^{\theta\sqrt{\ln(1+\sigma^2)}}$ 。各误差项采用的响应方式如表 5-11。

系统误差来源		作用方式
实验误差	亮度	$N_{tot}$ log-normal
	喷注	$N_t$ asymmetric
	光子孤立度	$N_t$ asymmetric
	光子鉴别	$N_t$ asymmetric
	光子能量刻度	$N_t$ asymmetric
	光子能量分辨	$N_t$ asymmetric
	事例堆积	$N_t$ asymmetric
理论误差	PDF	$N_t$ log-normal
	$\alpha_s$	$N_t$ log-normal
	QCD	$N_t$ log-normal
	部分子簇射	$N_t$ Gaussian
希格斯质量	ATLAS-CMS $m_H$	$\mu_{CB}$ Gaussian
	光子刻度	$\mu_{CB}$ Gaussian
	光子能量分辨	$\sigma_{CB}$ log-normal
本底模型	赝信号	$N_{spur}$ Gaussian

表 5-11 分析中系统误差项的总结。作用方式中  $N_{tot}$  表示信号事例数误差值对所有希格斯过程 (VBF, ggF) 相同， $N_t$  表示不同希格斯过程采用不同误差值。

$\tilde{d}(c_{H\tilde{W}})$  是统计模型中的隐含参数，通过影响  $N_{vbf}^i$  和  $f_{sig}(m_{\gamma\gamma})$  参与到似然函数中。我们通过换用一系列对应于不同  $\tilde{d}(c_{H\tilde{W}})$  的 BSM VBF 样本构造 BSM 似然函数，利用这组似然函数对数据进行样本拟合 (template fit)，最终得到对  $\tilde{d}(c_{H\tilde{W}})$  的极大似然估计值。在拟合中，VBF 信号强度  $\mu_{VBF}$ 、连续本底

数和本底模型参数保持浮动，这意味着没有任何模型依赖的截面计算的限制，同时可能的本底模型描述误差也被考虑进测量误差中。拟合结果使用似然函数负对数 (negative log-likelihood, NLL) 曲线表示，NLL 最小值  $NLL_{min}$  处的  $\tilde{d}$  ( $c_{H\tilde{W}}$ ) 即极大似然拟合值， $\Delta NLL = NLL - NLL_{min} = 0.5$  (1.92) 处的  $\tilde{d}$  ( $c_{H\tilde{W}}$ ) 值即其 68% (95%) 置信度 (confidence level, CL.) 的置信区间。

## 5.9 结果与总结

图 5-22 中展示了 TT、TL、LT 和三个区域联合的  $m_{\gamma\gamma}$  分布。通过对其拟合我们可以得到数据的  $\mathcal{O}\mathcal{O}$  分布，如图 5-23。将数据与 5.8 中描述的统计模型计算似然函数后， $\tilde{d}$  和  $c_{H\tilde{W}}$  得到的 NLL 曲线 (3 个区域的贡献和联合测量结果) 如图 5-24。在构造统计模型时  $\tilde{d}$  ( $c_{H\tilde{W}}$ ) 扫描步长为 0.01 (0.05)，连接扫描点得到平滑曲线，最佳拟合值为使用最低点附近的 7 个点进行二次函数拟合得到。这一结果中， $\tilde{d}$  考虑式 5-7 中干涉项 (interference term) 和二次项 (quadratic term) 的影响， $c_{H\tilde{W}}$  仅考虑理论中干涉项的影响。而为了完整性和与其他分析相比较，我们通过调整 BSM VBF 样本中的权重也得到了仅考虑干涉项的  $\tilde{d}$  结果和考虑干涉项和二次项的  $c_{H\tilde{W}}$  结果，总结在表 5-12。

在考虑的系统误差中，占主导地位的是本底模型引起的赝信号误差、VBF 过程部分子簇射误差和喷注重建中味道成分引起的误差。其中后两项也是 VBF 过程信号强度测量中主导的系统误差 [9]，而赝信号误差为每个  $\mathcal{O}\mathcal{O}$  bin 中  $m_{\gamma\gamma}$  拟合时本底模型和数据统计量引起的信号数误差，直接作用于  $\mathcal{O}\mathcal{O}$  谱型，因此对基于谱型的测量有显著影响。本分析的误差依然为统计误差主导，不同系统误差项对  $\tilde{d}$  和  $c_{H\tilde{W}}$  测量精度的影响如图 5-25 所示。各系统误差项对应的滋扰参数排序如图 5-26

在双光子道中完成本分析后，我们将  $\tilde{d}$  的测量结果与之前 VBF  $H \rightarrow \tau^+\tau^-$  道中的结果进行了联合，结果如图 5-27 和表 5-12 中所示。由于  $H \rightarrow \tau^+\tau^-$  分析中仅使用了 ATLAS Run 2 部分数据 ( $36 \text{ fb}^{-1}$ )，联合的结果中  $H \rightarrow \gamma\gamma$  的贡献占主导。

本分析使用 ATLAS Run 2 全部数据完成了在 VBF 过程中对 HVV 相互作用的  $CP$  性质的测量，得到了  $\tilde{d}$  和  $c_{H\tilde{W}}$  的 95% 置信区间。目前测量结果与 SM 预言相吻合，没有在 HVV 过程中发现  $CP$  破缺效应。在 HISZ 基下本分析是首次给出  $\tilde{d}$  的 95% 置信区间，在 Warsaw 基下本分析对  $c_{H\tilde{W}}$  的限制比之前 ATLAS 和 CMS 的分析严格约两倍。到目前为止本分析给出了对 HVV 过程中  $CP$  性质

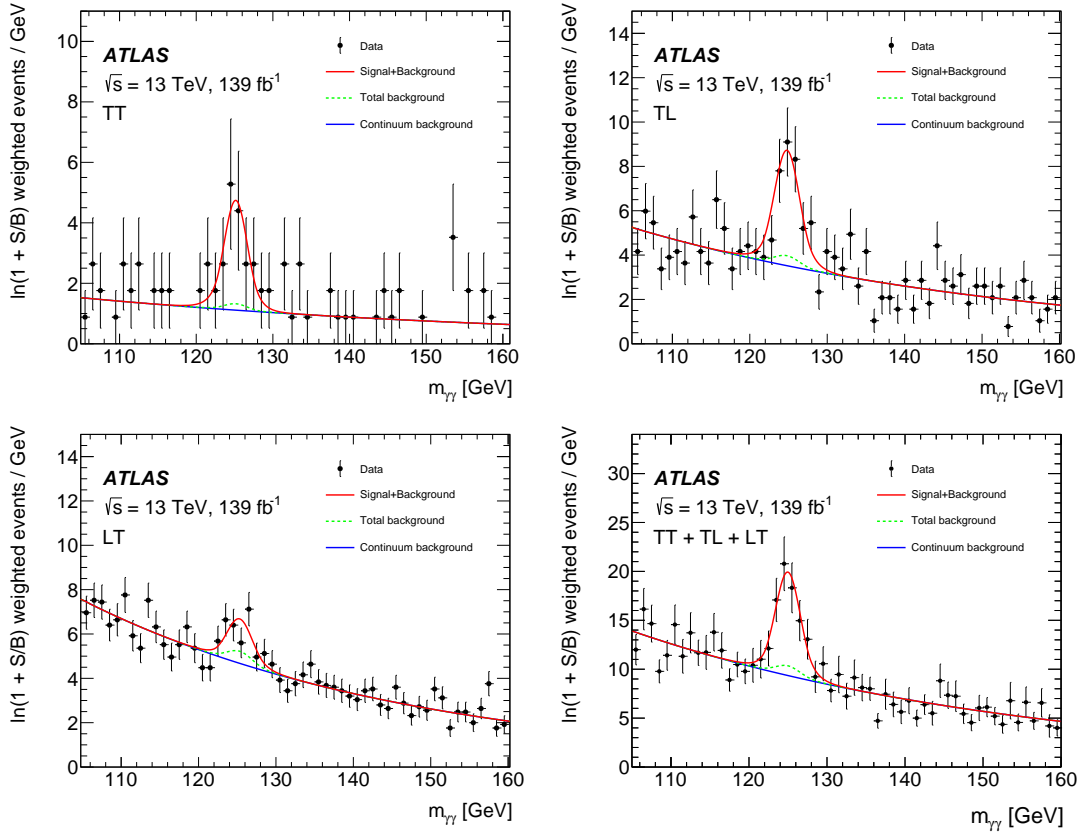


图 5-22 揭盲后数据在三个分类（左上、右上、左下）和联合后（右下）的  $m_{\gamma\gamma}$  分布。事例按照  $\ln(1 + S/B)$  进行了加权。

表 5-12 预期和观测到的  $\tilde{d}$  和  $c_{H\tilde{W}}$  68% 和 95% 置信区间，以及与  $H \rightarrow \tau^+\tau^-$  联合测量结果。表中  $H \rightarrow \tau^+\tau^-$  预期结果与 [24] 有稍许差别，是由于其对信号区和控制区相关性处理方式对区别造成的。

	68% (exp.)	95% (exp.)	68% (obs.)	95% (obs.)
$\tilde{d}$ (inter. only)	[-0.027, 0.027]	[-0.055, 0.055]	[-0.011, 0.036]	[-0.032, 0.059]
$\tilde{d}$ (inter.+quad.)	[-0.028, 0.028]	[-0.061, 0.060]	[-0.010, 0.040]	[-0.034, 0.071]
$\tilde{d}$ from $H \rightarrow \tau\tau$	[-0.038, 0.036]	—	[-0.090, 0.035]	-
Combined $\tilde{d}$	[-0.022, 0.021]	[-0.046, 0.045]	[-0.012, 0.030]	[-0.034, 0.057]
$c_{H\tilde{W}}$ (inter. only)	[-0.48, 0.48]	[-0.94, 0.94]	[-0.16, 0.64]	[-0.53, 1.02]
$c_{H\tilde{W}}$ (inter.+quad.)	[-0.48, 0.48]	[-0.95, 0.95]	[-0.15, 0.67]	[-0.55, 1.07]

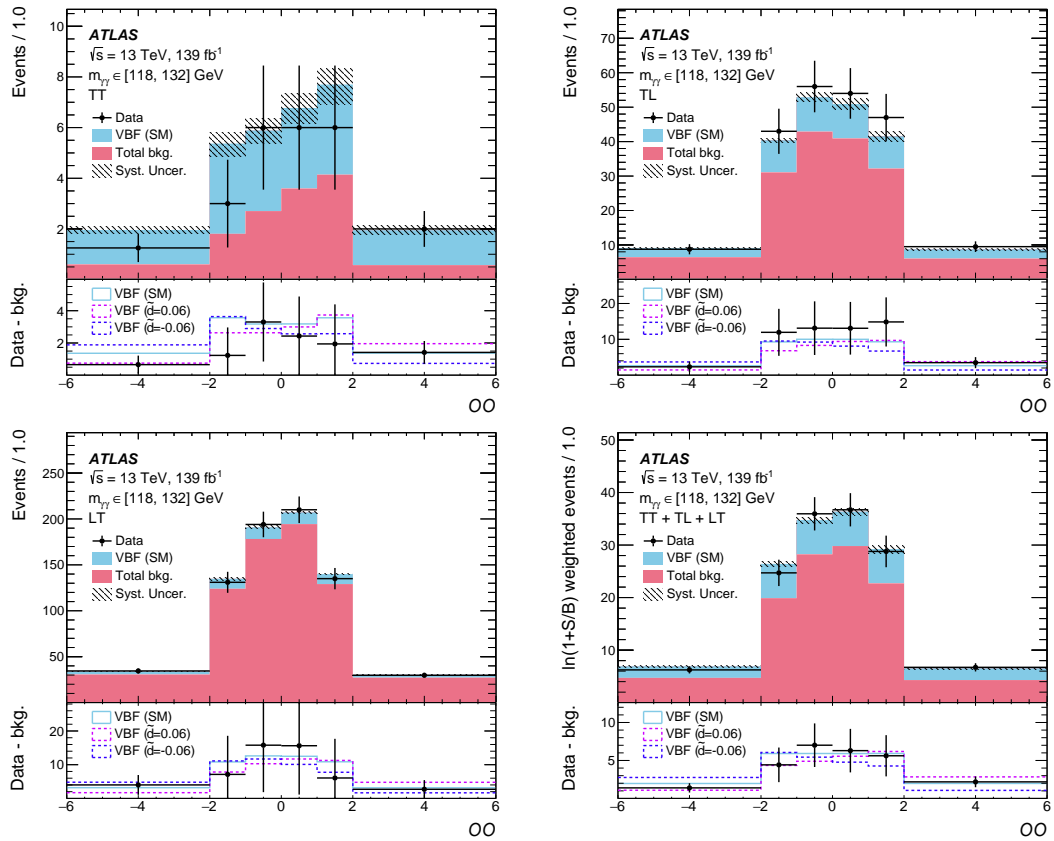


图 5-23 揭盲后数据在三个分类（左上、右上、左下）和联合后（右下）的  $\mathcal{O}\mathcal{O}$  分布。统计  $\mathcal{O}\mathcal{O}$  时仅计算  $m_{\gamma\gamma} \in [118, 132]$  GeV 区间中的事例数。图中一同展示了两个 BSM 样本中的  $\mathcal{O}\mathcal{O}$  分布作为对比。

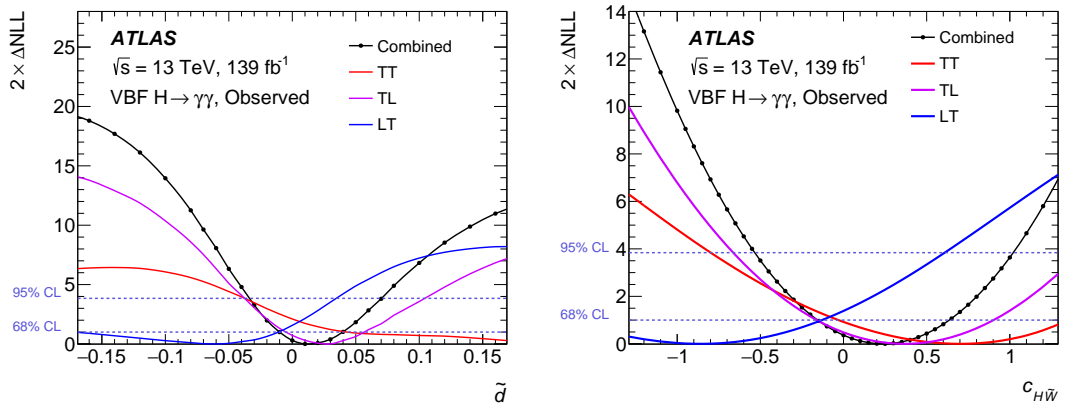


图 5-24 通过似然函数模型得到的对  $\tilde{d}$  和  $c_{H\tilde{W}}$  的 NLL 曲线，展示来自于不同分类的贡献。

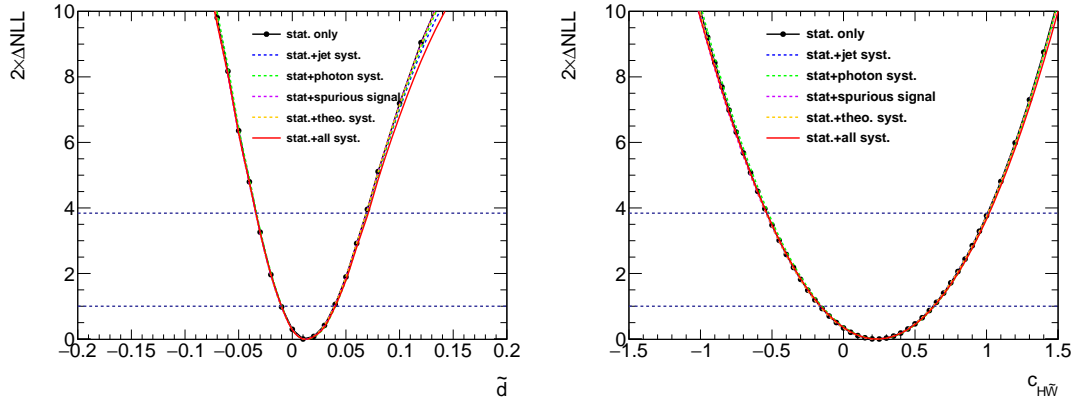


图 5-25 通过似然函数模型得到的对  $\tilde{d}$  和  $c_{H\tilde{W}}$  的 NLL 曲线，展示考虑系统误差项的贡献。

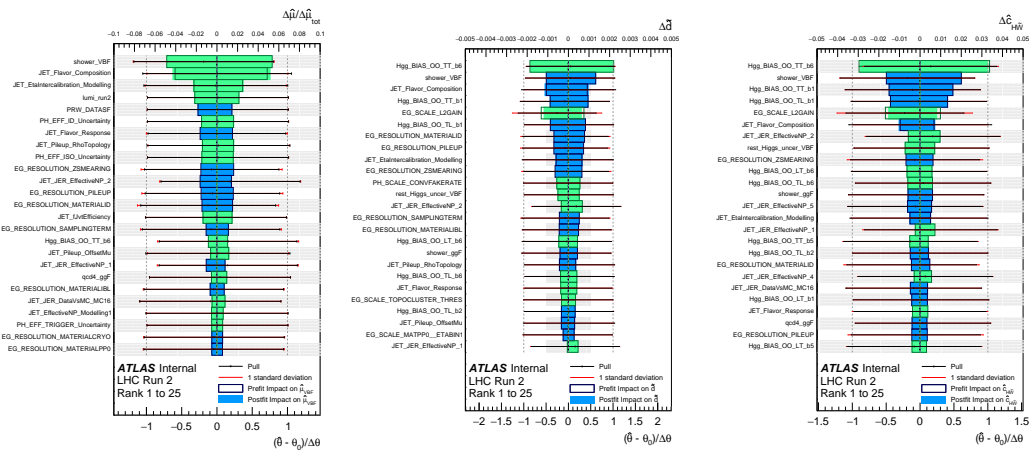


图 5-26 NLL curve

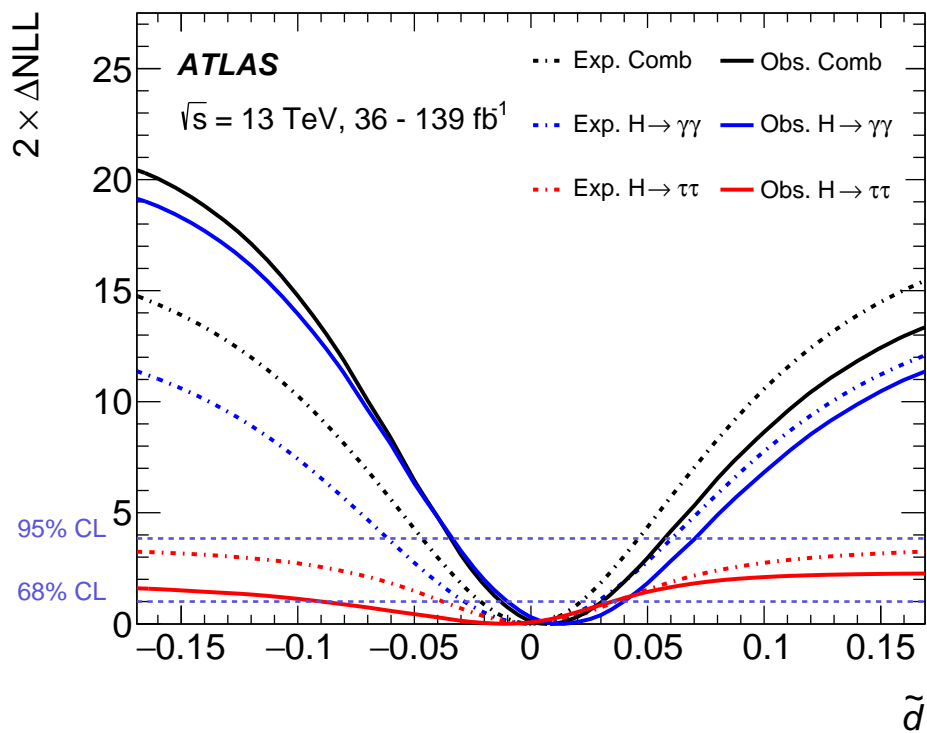


图 5-27 与  $H \rightarrow \tau^+\tau^-$  联合测量后的 NLL 曲线 (预期结果与观测到结果)。

最精确的限制。





## 第6章 环形正负电子对撞机 (CEPC) 与粒子流重建

### 6.1 未来对撞机项目概况

在 2012 年 ATLAS 和 CMS 宣布发现了希格斯玻色子后, 世界多家粒子物理研究机构都在讨论下一代对撞机项目的计划。LHC 的质子-质子对撞中产生的希格斯玻色子通常伴随有大量 QCD 本底, 不利于其性质的精确测量, 因此对未来对撞机的计划主要聚焦于正负电子对撞机。正负电子作为基本粒子, 直接参与对撞反应, 对撞能量可以由粒子束能量进行精确控制, 相对于质子对撞有非常干净的反应末态, 本底控制也更加容易, 适合大规模产生干净的希格斯粒子进行研究, 因此也被视为下一代希格斯工厂 (Higgs factory)。目前主流的对撞机计划中, 按照加速器类型可分为直线对撞机如 ILC (International Linear Collider) [15, 115, 116] 和 CLIC (Compact Linear Collider) [16], 与环形对撞机如 CEPC (Circular Electron Positron Collider) [10, 17] 和 FCC (Future Circular Collider) [18]。

直线对撞机将大量超导高频腔沿数十千米长的直线排布, 正负电子在经过时会受到腔中电场作用加速, 到达目标能量后在对撞点发生对撞。这种结构的优点在于无需使用磁场使带电粒子偏转, 也就不会产生同步辐射造成能量损失, 可以将粒子加速到非常高的能量。并且加速器可拓展性强, 通过延长加速器、增加高频腔数量可以继续提高对撞能量, 探索高能量前沿物理。但缺点在于粒子束经过一次对撞后即丢弃, 能量利用率较低, 而且大量的高频腔使得造假较为昂贵。ILC 和 CLIC 项目的早在发现希格斯粒子之前就已提出, 目的在于在轻子对撞机上直接寻找希格斯粒子, 同时配合 LHC 研究 TeV 级新物理。在初期计划中 ILC 加速轨道长度 20.5 km, 质心能量  $\sqrt{s} = 250$  GeV, 瞬时亮度  $1.35 \sim 2.7 \times 10^{34} \text{cm}^{-1} \text{s}^{-1}$ , 在 4 年的取数中累计得到  $400 \text{fb}^{-1}$  的数据。这一能量达到了  $e^+e^- \rightarrow \text{ZH}$  过程的产生阈值, 适合用于单希格斯粒子的研究。二期阶段 ILC 计划将加速器延长到 31 km, 对撞质心能量 500 GeV, 研究希格斯自耦合、顶夸克、 $W$  玻色子、电弱理论和  $f\bar{f}$  过程等。三期阶段加速器轨道可升级至 40 km, 对撞质心能量达到 1 TeV, 用于直接寻找新物理。目前 ILC 的前期研究和设计已经完成, 正在等待批准建造。CLIC 相比于 ILC 计划采用双束流技术加速正负电子, 加速梯度可达 70 - 100 MV/m, 目标对撞能量为 380 GeV, 1.5 TeV 和 3 TeV, 是目前唯一成熟的数 TeV 级轻子对撞机方案, 其主要物理目标在于直接寻找 TeV 级新物理。ILC 和 CLIC 的项目示意图如 6-1 所示。

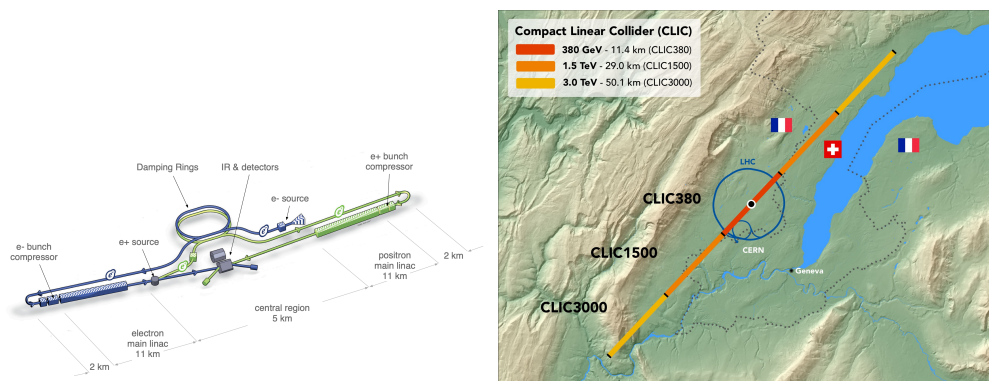


图 6-1 ILC [15] (左) 和 CLIC [16] (右) 项目计划示意图。

环形对撞机如前对 LHC 介绍中所述，利用磁场将带电粒子轨道限制在圆环中，在少数位置放置加速单元用于粒子加速。这种方案对粒子束的利用效率高，多次聚焦使得对撞亮度较高。缺点在于带电粒子会有同步辐射能量损失，每圈损失能量  $\Delta E$  与偏转半径成反比，与粒子能量的 4 次方成正比，与粒子质量的 4 次方成反比，因此在环形正负电子对撞机中对撞能量很难提高，且需要较大半径的储存环。考虑到希格斯玻色子质量仅 125 GeV，在环形正负电子对撞机中产生希格斯粒子仍然可行。CEPC 项目设计对撞环周长 100 km，对撞质心能量从 91 GeV 到 360 GeV，分 4 个运行周期，主要研究  $W$  玻色子、 $Z$  玻色子、希格斯玻色子和顶夸克的物理性质，从精确测量结果中寻找新物理的迹象。在 CEPC 运行结束后，其隧道经过改造可用于质子对撞 (Super Proton-Proton Collider, SPPC)，对撞能量计划达到 100 TeV，作为再下一代大型对撞机继续探索新物理。中国科学院高能物理研究所在 2012 年初步提出了 CEPC 的计划，此后一直稳步推进，在 2018 年发布了 CEPC 概念设计报告 (Concept Design Report) [10, 17]，并计划争取在 2027 年之前获得项目批准，2035 年完成建造开始取数。FCC 项目与 CEPC 类似，同样选用 100 km 隧道运行正负电子对撞计划 (FCC-ee)，并在结束后升级为电子-质子 (FCC-eh) 和质子-质子 (FCC-hh) 对撞，目前仍在设计和研究阶段。CEPC 和 FCC 项目示意图如图 6-2 所示。

除此以外，目前学界仍在研究的对撞机项目还有  $C^3$  (Cool Copper Collider) [117, 118]、缪子对撞机 [119] 等，各项目之间有竞争也有合作。但探索未知是全世界粒子物理学家共同的目标，无论哪个项目成功运行都将成为粒子物理学新的里程碑，拓展人类认知边界，推动科学发展。

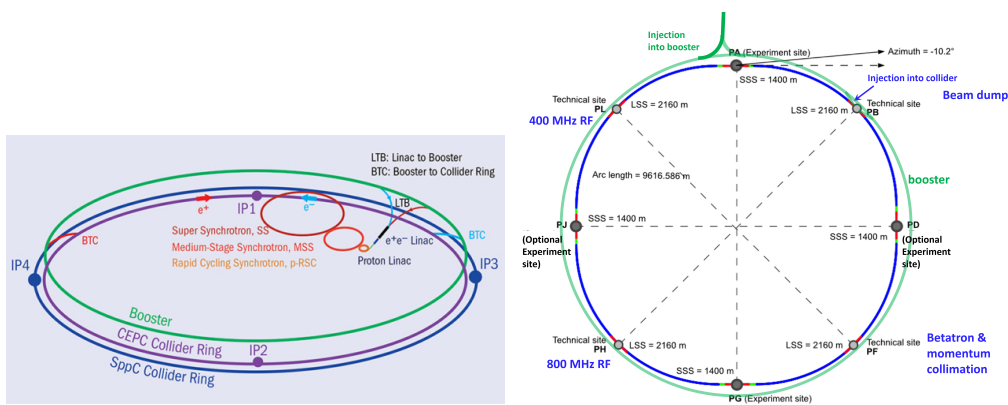


图 6-2 CEPC [17] (左) 和 FCC [18] (右) 项目计划示意图。

## 6.2 环形正负电子对撞机 (CEPC) 项目

### 6.2.1 物理目标

CEPC 项目主要作为希格斯工厂，进行希格斯物理的精确测量，同时也承担对标准模型中其他基础物理性质的研究。在目前最新的计划中 [25]，CEPC 将进行 10 年的  $\sqrt{s} = 240\text{GeV}$  希格斯工厂模式取数，累计数据量  $20 \text{ ab}^{-1}$ ，预期可产生  $4 \times 10^6$  个希格斯玻色子。随后计划进行至少 2 年的 Z 工厂模式 (Z-factory,  $\sqrt{s} = 91.2 \text{ GeV}$ ) 和 1 年的 WW 阈值模式 (WW threshold,  $\sqrt{s} = 160 \text{ GeV}$ ) 取数。此后 CEPC 将进行能量升级，进行至少 5 年的  $t\bar{t}$  模式 ( $e^+e^- \rightarrow t\bar{t}$ ,  $\sqrt{s} = 360\text{GeV}$ ) 取数。具体计划如表 6-1。

运行模式	Z 工厂	WW 阈	Higgs 工厂	$t\bar{t}$
$\sqrt{s}$ (GeV)	91.2	160	240	360
运行时间 (年)	2	1	10	5
积分亮度 ( $\text{ab}^{-1}$ , 2 IPs)	100	6	20	1
预期事例数	$3 \times 10^{12}$	$1 \times 10^8$	$4 \times 10^6$	$5 \times 10^5$

表 6-1 CEPC 运行计划及各模式下预期的信号事例数 [25]。

CEPC 上首先计划进行的希格斯物理研究是希格斯粒子耦合性质的测量，通过测量希格斯粒子产生截面  $\sigma(\text{ZH})$  和各衰变道分支比  $Br(H \rightarrow X)$  完成。相比于强子对撞机，CEPC 物理过程简单，产生过程中基本不受 PDF 影响和高阶 QCD 修正影响，截面的理论计算可以达到很高精度，同时对撞环境本底干净，几乎没有事例堆积影响。在目前的预研工作中各分支比的测量分别进行，在采用了一些分析方法如多变量分析、喷注味鉴别 (flavor tagging) 等方法后预期精度可

达到  $O(1\%)$  量级 [25, 120]。其他性质如希格斯粒子的自旋、 $CP$  宇称和相关耦合参数的测量也有初步结果，预期精度相比于 HL-LHC 有显著提升 [121]。得到这些结果后，我们可以在理论框架如  $\kappa$ -framework、SMEFT 中进行总体拟合得到希格斯粒子与其他标准模型粒子的耦合参数，以此判断在标准模型的希格斯场部分中是否存在新物理。另外，由于正负电子对撞机上满足能动量守恒，我们可以利用  $Z$  玻色子的反冲信息对  $\sigma(ZH)$ 、 $\Gamma_H$  进行模型无关的测量，也可以利用反冲质量  $M_{recoil}$  直接希格斯粒子质量  $m_H$  和希格斯粒子不可见末态的分支比  $Br(H \rightarrow inv.)$ 。在选择特定  $Z$  衰变道如  $Z \rightarrow e^+e^-/Z \rightarrow \mu^+\mu^-$  后，这些分析也能够达到相当高的精度。

$Z$  工厂和  $WW$ -threshold 模式下运行的 CEPC 可以产生数百亿的  $Z$  玻色子和上亿  $W$  玻色子用于电弱测量和味物理研究。对电弱过程基本参数如  $W$ 、 $Z$  玻色子质量和质量宽度、电弱耦合角  $\sin^2 \theta_W$ 、前后电荷不对称度  $A_{FB}^f$  的测量可验证标准模型的一致性， $W/Z$  玻色子衰变过程也可用于寻找轻子数破缺 (Lepton Flavor Violation, LFV) 等稀有物理过程，从中寻找新物理。2022 年 CDF 合作组公布的  $W$  玻色子质量  $m_W$  测量结果与标准模型预测值之间存在  $7\sigma$  的偏差 [39]，CEPC 有机会对这一结果进行验证。在得到相关结果后，对希格斯物理和电弱的联合拟合是对电弱对称性破缺及标准模型自洽性很关键的检验，有助于加深对标准模型的理解，回答如阶层问题 (hierarchy problem) 等粒子物理中最基础的问题。 $Z$  工厂模式中产生的大量  $Z$  玻色子可衰变到  $b\bar{b}$ 、 $c\bar{c}$ 、 $\tau^+\tau^-$  末态，通过强子化产生重味强子，这可以为我们提供在 CEPC 上研究味物理的机会，如测量介子质量、CMK 矩阵元、 $CP$  相角等。同时，味物理也是目前研究粒子物理中的基本问题，如费米子代数、轻子混合 (lepton mixing / lepton universality) 等问题的最佳窗口。

### 6.2.2 CEPC 基准探测器

为达到以上物理测量的目标，CEPC 对探测器和物理对象的测量提出了一系列条件，简略描述如下<sup>1</sup>：

- 径迹测量：对 1 GeV 以上的径迹，重建效率应达到 99% 以上，动量分辨应达到 0.001 级别。
- 轻子鉴别：对于 5 GeV 以上的孤立轻子，鉴别效率应达到 99% 以上，误判率 2% 以下。对喷注中的轻子也应有较高的鉴别效率。

<sup>1</sup>以下要求取自 CEPC CDR Volume 2 [10]，此后对探测器指标的更新目前未考虑。



- 带电强子鉴别：带电强子的鉴别效率和纯度都应达到 90% 以上。
- 光子测量和鉴别：光子能量分辨应达到  $20\%/\sqrt{E}$  以内，对喷注中  $\pi^0$  的重建效率应达到 95% 以上。
- 喷注和丢失能量：应能够分辨出  $W \rightarrow q\bar{q}$  和  $Z \rightarrow q\bar{q}$  事例，对应于重建出的玻色子质量分辨 (Boson Mass Resolution, BMR) 4%。
- 味鉴别：b-jet 标记效率应大于 80%，c-jet 标记效率大于 60%，

在此标准下，CEPC 的基准探测器 (baseline detector) 沿用了 ILC 探测器 ILD 的设计方案，根据 CEPC 对撞环境进行了优化。探测器从内到外包括硅像素顶点探测器、硅内部径迹探测器 (Silicon Internal Tracker, SIT)、时间投影室 (Time Projection Chamber, TPC)、外部硅探测器 (Silicon External Tracker, SET)、硅-钨取样型电磁量能器 (Si-W ECAL)、基于不锈钢-玻璃阻性版室 (steel-Glass Resistive Plate Chamber, GRPC) 的取样型强子量能器 (HCAL)、3 Tesla 超导磁铁线圈和缪子探测器。在探测器前端还有 5 对硅径迹版，用于前向径迹的测量。探测器整体示意图如图 6-3。

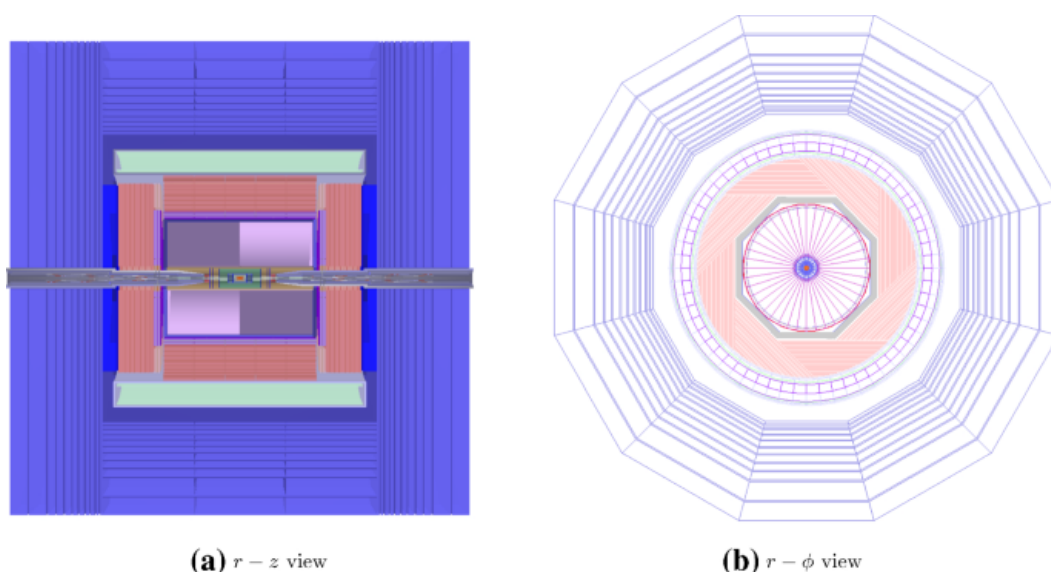


图 6-3 CEPC 基准探测器在  $r-z$  (左) 和  $r-\phi$  (右) 方向上的结构示意图 [10]。

顶点探测器由 3 组共 6 个同心圆桶组成，围绕在对装点周围 1.6 ~ 6 cm 范围内， $z$  向长度 62 ~ 125 mm，可提供  $5 \mu\text{m}$  级别的顶点分辨。SIT 和 SET 为硅微条层，SIT 的两层位于顶点探测器和 TPC 之间，距束流管 15 cm 和 30 cm 处，SET 位于 TPC 和 ECAL 之间，距束流管 180 cm。两者联合提供了 TPC 前后带电粒子击中点的定位，有助于提高整个径迹探测器的探测效率，同时可以将径迹信息外推至量能器中。TPC 内外半径分别为 30 cm 和 180 cm，总长 4.7 m，沿径向分为

220 层,  $\phi$  向分割为 1 mm 宽的探测单元, 在 3 Tesla 磁场中可达到对单个点  $r-\phi$  向 ( $z$  向) 位置分辨  $100 \mu\text{m}$  ( $500 \mu\text{m}$ )。在结合所有径迹探测器信息后, 径迹动量分辨  $\Delta(1/p_T) \sim 2 \times 10^{-5} \text{ GeV}^{-1}$ , 且总物质量控制在  $O(1\%)X_0$  级别。

为实现  $\text{BMR} \sim 4\%$  的物理要求, 量能器系统是关键。ECAL 和 HCAL 的设计按照粒子流 (Particle Flow, 具体见 6.3) 重建的要求, 采用高粒度取样型量能器。ECAL 以硅为灵敏层, 钨板为吸收体, 沿探测器径向分为 30 层, 总厚度  $24X_0$ 。横向上探测单元被分割为  $1 \times 1\text{cm}^2$ , 以实现簇射三维结构的探测。HCAL 以不锈钢为吸收体, RPC 为灵敏单元, 尺寸  $1 \times 1\text{cm}^2$ , 同样具有三维探测能力。这组量能器可达到  $16\%/\sqrt{E/\text{GeV}}$  的光子能量分辨和  $60\%/\sqrt{E/\text{GeV}}$  的中性强子能量分辨, 在采用粒子流算法进行重建后, 对 100 GeV 以下喷注的能量分辨约为  $\sim 30\%/\sqrt{E/\text{GeV}}$ , 可达到 BMR 要求。

在基准探测器设计以外, CEPC 工作组还研究了采用全硅径迹室代替 TPC 的方案, 以及一个备选探测器方案 IDEA (Innovative Detector for Electron-positron Accelerator, FCC-ee 的参考探测器方案)。在 CDR 发布后依然在对此方案进行优化, 并提出第四代探测器概念设计方案。

### 6.2.3 CEPC 软件环境

CEPC 的软件环境 CEPCsoft [122] 沿用自 iLCsoft [123]。这套框架由 ILC 团队设计, 主要分为 4 个部分: 事例产生、模拟、重建和分析, 以 C++、Java 和 Fortran 77 为编程语言, 采用通用的数据模型 (Event Data Model, EDM) LCIO [124] 和探测器几何描述 DD4hep [125]。在事例产生框架中 iLCsoft 内嵌了 Whizard [126] 和 PYTHIA 产生子用于产生正负电子对撞的物理过程, 事例的全模拟则使用基于 Geant 4 的 Mokka [127] 完成。重建和分析软件则在 Marlin (Modular Analysis & Reconstruction for the LINear collider) [128] 框架下进行开发, 这是一套 C++ 语言的应用框架, 将每个算法包装成一个 Processor, 采用固定流程处理每个实例, 而算法参数可以通过 XML 文件灵活调配。iLCsoft 的设计目标为一个用于轻子对撞机的简洁、灵活、模块化的软件流, 使用粒子物理实验通用工具和数据模型, 利于国际合作和共同开发。iLCsoft 的结构和 CEPCsoft 的工作流如图 6-4。

在发现希格斯粒子后, 未来对撞机实验经讨论决定使用 Key4HEP [129] 作为未来粒子物理的公用软件环境, 目的同样在于最大化现有软件工具和模型的利用率。相比于 iLCsoft, Key4HEP 中跟随时代发展以 C++ 和 Python 语言为主, 采用 EDM4HEP [130] 作为公用数据模型, 用 Gaudi [131] 代替 Marlin 作为事例处理

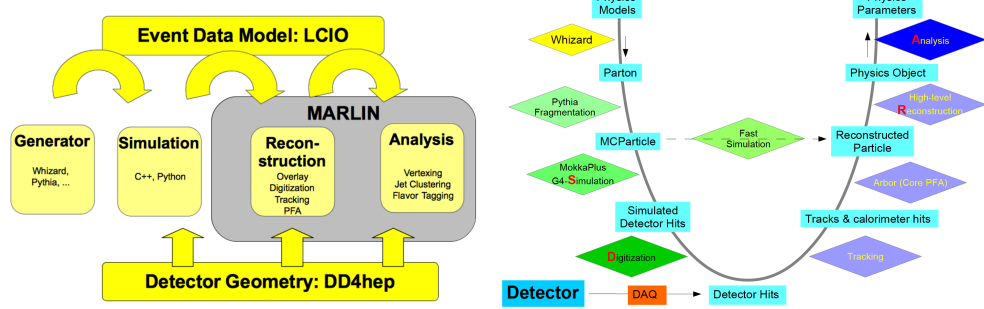


图 6-4 iLCsoft 软件框架结构 (左) 和 CEPCsoft 工作流 (右)。

的软件框架，几何描述则继续使用 DD4hep。同时，Key4HEP 计划提供更多近年出现的软件工具的接口，以及对原 Marlin 框架下算法的移植工具。CEPC 项目也跟随粒子物理主流方向，正在将软件环境转移到基于 Key4HEP 的 CEPCSW [132] 中。

#### 6.2.4 CEPC 第四代探测器概念设计

第四代探测器的概念设计是在对基准探测器的性能、工程结构、造价等进行多方面考虑后提出的新型设计方案，如图 6-5，相比于基准探测器，其主要更新包括：

- 针对粒子鉴别 (PID) 能力优化的漂移室 (Drift Chamber, DC) 代替原 TPC。漂移室对高事例率对撞的接受程度更高，使得探测器可以应对高亮度的  $Z$  工厂模式，其挑战在于保持动量分辨和粒子鉴别能力。
- 以长晶体条为探测单元的全吸收型电磁量能器代替原硅-钨取样型量能器。这可以将光子能量分辨提高至  $\sim 3\%/\sqrt{E/\text{GeV}}$ ，并拥有更好的  $\gamma/\pi^0$  分辨能力，有助于味物理的研究，但缺点在于不完全符合粒子流量能器要求，需要单独的重建算法。
- 以闪烁玻璃为灵敏单元的强子量能器，其优点在于更低的造价和更高的密度，可以减小外部缪子探测器和超导线圈的体积从而进一步减少造价，缺点在于光产额不足，需对其优化。

这一探测器具体的设计工作仍在进行中，还没有得到完整的物理性能。本人参与的对晶体电磁量能器重建算法的开发主要在第 7 章中讨论。

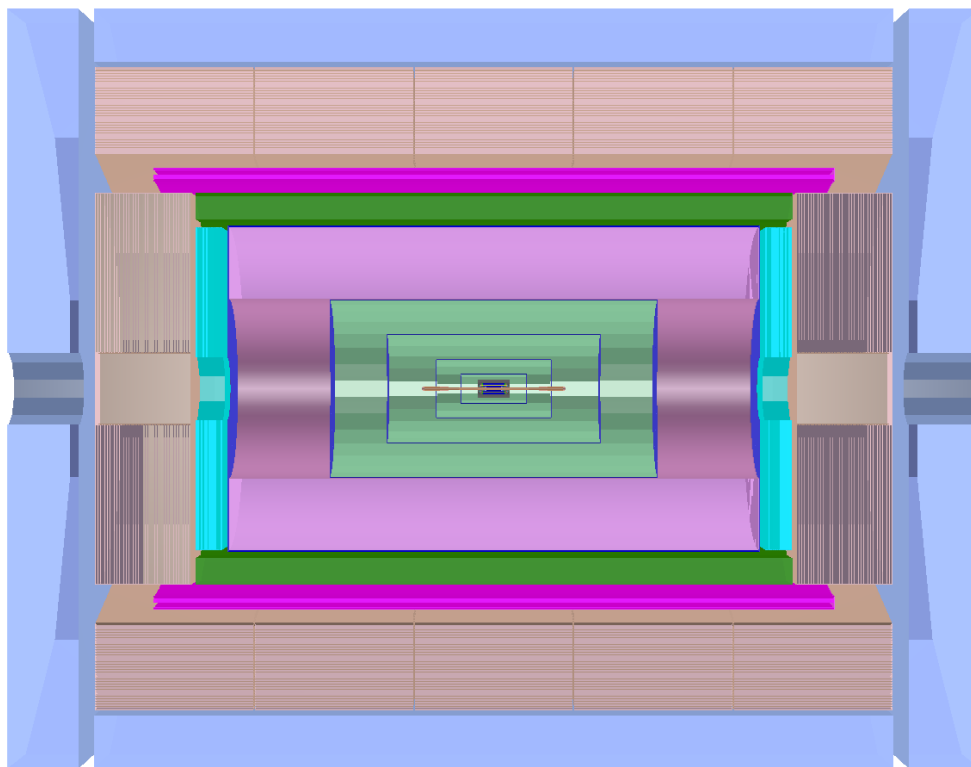


图 6-5 CEPC 第四代探测器概念设计结构图。

### 6.3 粒子流重建方法

在传统的喷注重建方法中，喷注能量的重建来自于对 ECAL 和 HCAL 中分别的沉积能量，这会使得总能量分辨  $> 60\%/\sqrt{E/\text{GeV}}$ ，无法满足在未来轻子对撞机中区分  $W \rightarrow q\bar{q}/Z \rightarrow q\bar{q}$  的要求。粒子流的思想就是为解决这一问题而提出。LEP 的实验结果提供了对喷注成分的认识 [133, 134]，平均来讲在喷注中的短寿命粒子完成衰变后，喷注能量的 62% 被带电粒子（带电强子为主）带走，27% 被光子带走，10% 被稳定中性强子带走，1.5% 被中微子带走。粒子流方法指出，可以采用不同子探测器对喷注成分进行分别测量，组合后整体能量分辨可以得到很大提高，如表 6-2 总结 [135]。

在粒子流重建中，首先通过径迹探测器重建出径迹，再在量能器中根据能量沉积重建出簇团，要求能够将不同粒子产生的簇团区分开。然后在算法中进行径迹和簇团的匹配，在量能器中扣除掉认为是带电粒子产生的能量沉积，将其他簇团重建为光子或中性强子。这些重建好的粒子即被定义为粒子流对象（Particle Flow Object, PFO），可以进行一定粒子鉴别。此后即可使用其他喷注重建算法，如  $ee-k_T$  [136]、 $\text{anti-}k_T$  [74] 将 PFO 组合成喷注，得到最终重建结果。

从以上流程中可看出，粒子流重建是对探测器和重建软件的双重要求。对探



组分	探测器	能量占比	测量分辨	喷注能量分辨
带电粒子 ( $X^\pm$ )	Tracker	$\sim 0.6 E_j$	$10^{-4} E_{X^\pm}^2$	$< 3.6 \times 10^{-5} E_j^2$
光子 ( $\gamma$ )	ECAL	$\sim 0.3 E_j$	$0.15/\sqrt{E_\gamma}$	$0.08/\sqrt{E_j}$
中性强子 ( $h^0$ )	HCAL	$\sim 0.1 E_j$	$0.55/\sqrt{E_{h^0}}$	$0.17/\sqrt{E_j}$
总和	全探测器	1		$0.19/\sqrt{E_j}$

表 6-2 喷注中各成分的能量占比及在探测器中的预期能量分辨。在使用各部分的最优测量结果综合全探测器信息后，可预期喷注总能量分辨为  $\sim 20\%/\sqrt{E_j}$ 。

测器，特别是量能器来说，应满足 [137]:

- 高颗粒度的探测单元，对应较小的探测单元尺寸，以实现簇团的精确识别和区分；
- 选用物质有较小莫里哀半径  $R_M$ ，以限制电磁簇射的横向发展尺寸；
- 较大的核作用长度与辐射长度比值  $\lambda_I/X_0$ ，使得电磁簇射与强子簇射在纵向上有很好的区分；

同时，这一高颗粒度量能器也对读出系统有一定要求，包括：

- 前端电子学应能够对信号进行一定预处理，如设计触发条件，采用零压制 (zero-suppression) 方法，以减少读出通道数；
- 电子学元件完全嵌入在探测器读出板上，要求有很好的质量和抗辐照性能；
- 探测器冷却系统能够带走如此大量读出信号产生的热量；
- 探测器大量区域被 PCB 覆盖，将提升整体探测器造价。

对重建软件来说，主要挑战来自于对簇团的识别和分割。对沉积能量的错误聚类会导致对中性粒子能量的错误重建，从而影响到喷注重建。过程中还需引入径迹的位置和动量与重建出的簇团进行匹配，调整簇团聚类，多次迭代后得到稳定分割结果。多年以来 CALICE 合作组提出了多种粒子流量能器设计方案 [138, 139]，其中包括 CEPC 基准探测器中采用的 Si-W ECAL 和 GRPC HCAL，并针对每个方案进行了大量的研究和测试。软件方面目前主流的粒子流重建算法 (Particle Flow Algorithm, PFA) 有 PandoraPFA [135]，ArborPFA [140, 141]，后者为 CEPC 研究中采用的重建算法，在 CEPC CDR 中已证明可以达到目标能量分辨。



## 第7章 晶体电磁量能器重建算法的开发

如 6.2.4 中所述，CEPC 第四代探测器中计划采用全吸收型晶体电磁量能器，以提高光子能量分辨。为满足 PFA 重建方法的要求，这一探测器的结构有两种思路，如图 7-1 所示：第一种思路为采用晶体立方体填充整个电磁量能器，以满足高颗粒度的要求，但其对读出电子学、支撑、冷却等系统的设计有相当大的挑战。第二种思路中，探测器由  $1 \times 1 \times \sim 40\text{cm}^3$  的长晶体条组成，相邻两层中的晶体条正交排放组成一个超层（super layer），以一个超层为基础单元通过后期重建提供三维空间点的定位。每个晶体条中的信号由两端的硅基光倍增器（Silicon Photon Multiplier, SiPM）进行双端读出，同时输出电荷和时间信息。14 个超层（对应 24 倍辐射长度）在探测器径向上叠放成一个约  $40 \times 40 \times 28\text{cm}^3$  的超单元（super cell）作为探测器的基础模块。

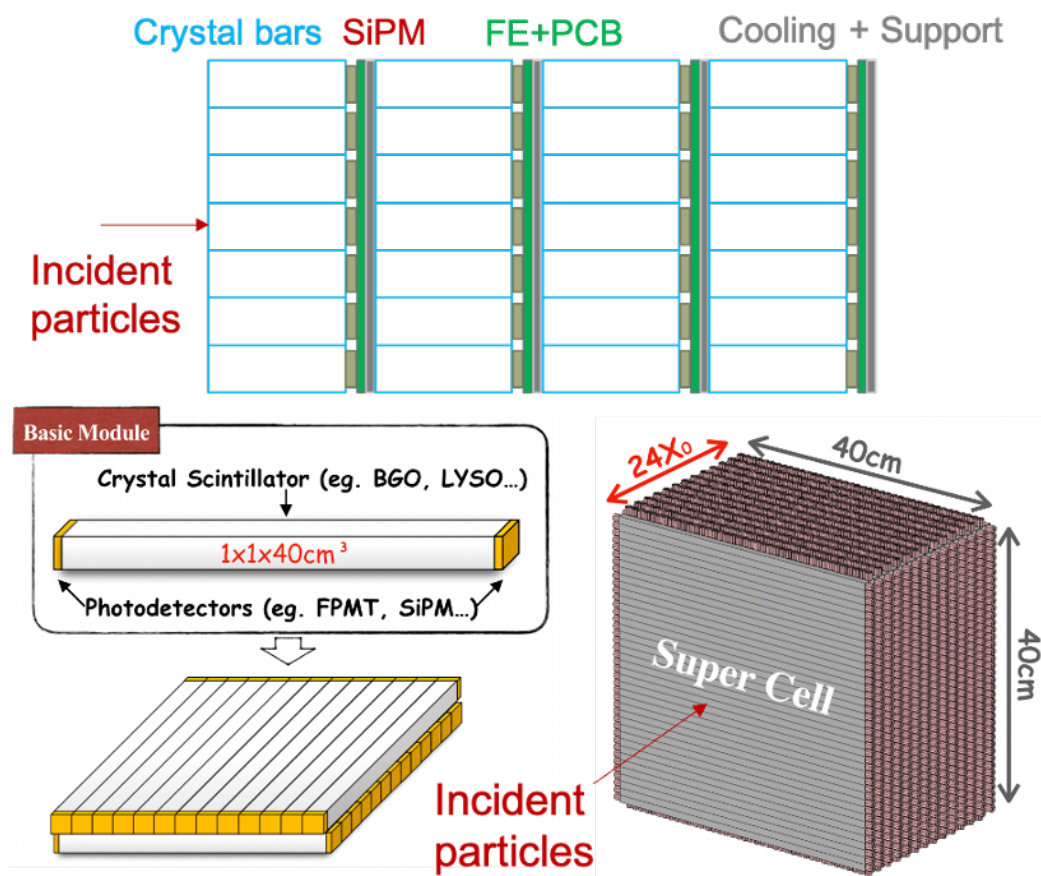


图 7-1 晶体 PFA 电磁量能器的两种设计思路：小颗粒晶体块单端读出方案（上）和交叉晶体条两端读出方案（下）。[19, 20]

与第一种设计思路相比,第二种方案中读出和支撑系统可以放置在 super cell 四周,不会成为灵敏单元中间的死物质影响探测效率,同时信号通道数大幅减少,降低了数据传输和冷却系统的压力,也能够节省造价。但也有其缺点:一是需要额外的重建算法将晶体条信息重建为高颗粒度击中点;二是当多个粒子同时进入一个 super cell 时,可能会由于误组合重建出“幽灵击中”(ghost hits)。这些幽灵击中将对后期能量簇团的识别和物理对象的重建带来很大的影响。

另外,这两种晶体型量能器的设计方案都将面临晶体大  $R_M$  带来的电磁簇射横向展宽的问题。如仍旧采用 PandoraPFA 和 ArborPFA 中根据击中点区分簇团的方法,在面对喷注中的多粒子同时入射时混淆问题会更加严重,对喷注能量分辨有一定影响。

考虑到以上问题,本章的主要工作就是开发出能够通过长条探测单元得到高颗粒度击中点并去除幽灵击中的重建算法,验证此长条形晶体电磁量能器设计方案的可行性。

## 7.1 探测器几何搭建

重建算法的开发首先需要有一个探测器的几何模型。我们从桶部开始,为适配 CEPC 探测器整体结构,新的电磁量能器采用与 CEPC 基准探测器中相似的八边形结构,如图 7-2 (左)。8 个全同的梯形块覆盖了  $\phi$  向  $2\pi$  范围的全部空间,其叠放方式保证了没有指向对撞点的缝隙。桶部  $z$  向总长度 4.6 m,内径 1.8 m,覆盖角度  $|\cos\theta| < 0.78$ 。每个梯形块由  $4 \times 11$  个 super cell 组成,随层数增加 super cell 尺寸逐渐缩小,因此晶体条实际长度从 33 cm 到 46 cm 不等,横截面固定为  $1 \times 1\text{cm}^2$ ,如图 7-2 (右)。为便于重建时纵横方向晶体条的匹配,所有晶体条长度取整数,保证每个 super layer 为规则长方体。这将导致在梯形块斜边,即两个梯形块交界处有  $2 \times 2\text{cm}^2$  的直角三角型空隙,可作为电子学芯片或线缆等的位置。

为满足未来 CEPC 相关软件开发的要求,这一几何模型的构建在 DD4Hep 中完成。在研究的初级阶段我们暂时不考虑部分实际探测器结构,如吸收体、读出电子学、包装材料等,仅用晶体作为灵敏物质填充整个理想化的探测器。CEPCSW 框架对此提供了必要的软件库支持,其模拟模块将几何信息读入,调用 Geant4 选取指定步长模拟粒子与探测器物质之间的相互作用,并记录每一步中的位置和沉积能量,采用标准数据类型 EDM4HEP 输出为 PODIO 文件。

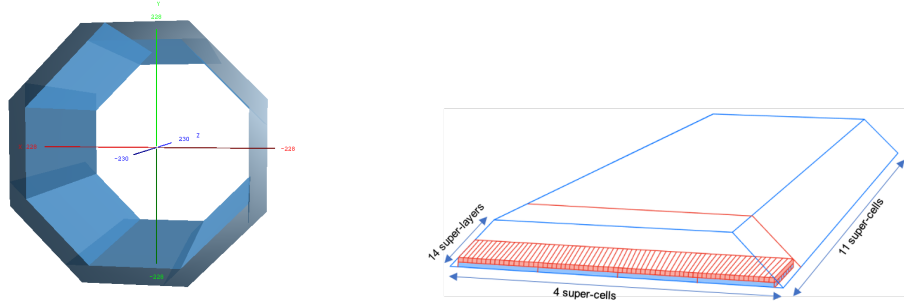


图 7-2 CEPC 第四代探测器概念设计中的八边形桶部电磁量能器总览 (左) 和其中一个梯形块 (右)。

## 7.2 数字化

数字化提供了对电子学读出过程的模拟。在构造一个探测器模型时，需要对电子学读出系统有精确的了解，如 SiPM 对光的接收效率、接收到晶体中的光信号后将其转化为电信号的幅度、整体噪声水平等。如此得到的探测器模拟数据才能够与实际测量相对比。在目前的研究阶段我们同样采用理想化的数字化模型：认为 SiPM 对晶体中闪烁光的接收效率为 100%；忽略闪烁光在晶体中的衰减；不考虑元器件响应曲线，认为在晶体灵敏体积内沉积的能量和即 SiPM 读出信号幅度；忽略噪声和芯片台基信号等电子学效应；SiPM 记录的时间记为达到一定阈值后入射粒子激发出的闪烁光子到达 SiPM 的最快传播时间，以高斯项考虑时间测量的分辨率。如此，每根晶体条输出两端读出的电荷和时间信号  $(Q_-, T_-)$ ,  $(Q_+, T_+)$ ：

$$Q_{\pm}^i = c \cdot E_i \cdot e^{-\frac{L_{bar}/2 \pm z_i}{L_{Att}}}; \quad (7-1)$$

$$Q_{\pm} = \sum_i Q_{\pm}^i. \quad (7-2)$$

其中  $c$  是刻度常数，在我们的理想模型中固定为 1； $E_i$  为 Geant4 step 的沉积能量； $L_{bar}/2 \pm z_i$  为其到晶体两端的距离； $L_{Att}$  为衰减长度，理想模型中设为无穷大。读出时间有：

$$T_{\pm}^i = T_0 + \text{Gaus}(z_{\pm}^i/v, \sigma_T); \quad (7-3)$$

$$T_{\pm} = T_{\pm}^k | (\sum_{i=1}^k Q_{\pm}^i > \text{threshold}) \quad (7-4)$$

其中  $T_0$  是给定的初始时间； $\text{Gaus}(z_{\pm}^i/v, \sigma_T)$  表示对传播时间进行高斯 smearing 以考虑探测中的分辨，研究中设为 500 ps；阈值 threshold 设为 5%。图 7-3 为此过程的示意图。

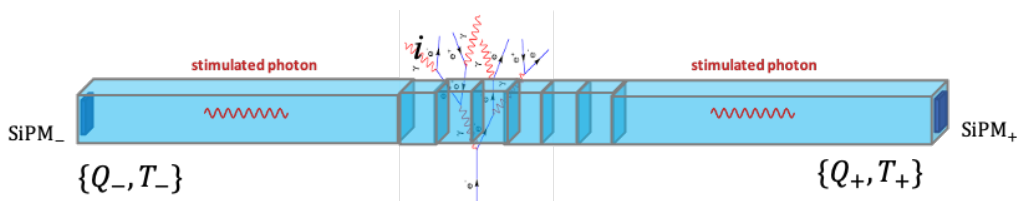


图 7-3 晶体条被击中时的数字化过程示意图。晶体中激发出的闪烁光子被两端 SiPM 接收并转化成电信号，输出到读出系统中。

晶体条的读出电荷和时间信息将对重建有重要影响。通过电荷刻度得到准确的对应沉积能量是实现高分辨率能量测量的关键；在后续重建算法中时间信息被用来判定和去除幽灵击中，时间测量的分辨率对去除能力有显著影响。受限于实验条件，本文没有对其进行深入研究，而将重点放在算法开发上。

算法编写在 Gaudi 框架中完成，以模拟文件作为输入，读取其中的 *edm4hep::SimCalorimeterHitCollection*，在按 1 所示的工作流对每个事例中每个晶体条进行处理后得到数字化后的信号并写为 *edm4hep::CalorimeterHitCollection*。算法的可输入参数如表 7-1，使用 Gaudi 框架在执行脚本文件中进行控制，方便优化过程的参数扫描。

---

#### Algorithm 1: 数字化算法 EcalDigiAlg

---

**Input:** 模拟中晶体条 *edm4hep::SimCalorimeterHitCollection*;

**Output:** 数字化后晶体条 *edm4hep::CalorimeterHitCollection*;

- 1 初始化 GeoService，探测器解码器 Decoder;
  - 2 **foreach** 读入晶体条 **do**
    - 3 初始化 CalorimeterHit;
    - 4 对内部所有 G4 step 能量求和，考虑向两端衰减效应得到双端读出信号幅度;
    - 5 根据 G4 step 在晶体中相对位置，计算传播时间得到双端读出时间;
    - 6 写入 CalorimeterHit;
  - 7 将数字化后的晶体条写出;
-

输入参数	默认值
晶体中总能量能量阈值	0.1 MeV
刻度常数	1
晶体中光衰减长度	$\infty$
测量时间分辨	500 ps
介质折射率	2.15 (BGO)
选取最快时间对应的能量分数阈值	5%

表 7-1 数字化算法的可输入参数。

### 7.3 重建算法软件设计

这一电磁量能器的重建主要分为三步：

(1) 将临近着火晶体条进行聚类，如果识别认为存在电磁簇射的重叠就对晶体条能量进行劈裂；

(2) 在一个超层中匹配两个方向的晶体条，得到（伪）高颗粒度击中点的信息，同时利用时间信息去除 ghost hit，得到平面内的二维簇团；

(3) 将每一层中的二维簇团进行纵向连接，形成簇团。

我们为前两步开发了特定的子算法，并在下文中进行了详细描述。第三步相对而言较为普通，目前使用了圆锥角聚类算法，从探测器内向外逐层推进，将在圆锥角内部的二维簇团连接起来。为去除晶体量能器中的杂散簇团这一重建流程还进行了多次迭代，以上一步迭代形成的三维簇团作为参照，将附近杂散簇团吸收进主簇团中。迭代流程如图 7-4。注意到此重建软件还在开发中，目前设计的流程为阶段性结果，后续还计划进行调整和优化。

在软件设计方面，我们参考了 Pandora 的软件开发工具包 PandoraSDK [21]，希望搭建一个模块化、可扩展、灵活调整的算法框架。考虑到算法内置于 CEPCSW 框架中，可能会作为粒子物理公用软件进行推广，数据模型的接口采用 EDM4HEP 的公共模型。具体设计在下文中进行了详细介绍。

#### 7.3.1 晶体条能量劈裂算法

高能光子在进入介质时，会通过电子对效应产生一对正负电子，正负电子再辐射产生次级光子，这一过程经过级联发展形成电磁簇射，直到光子能量低于电子对效应的阈值，通过康普顿效应或光电效应损失能量，次级粒子数目逐渐减少，簇射结束。电磁簇射过程能量沉积的纵向分布可以使用  $\Gamma$  函数进行描



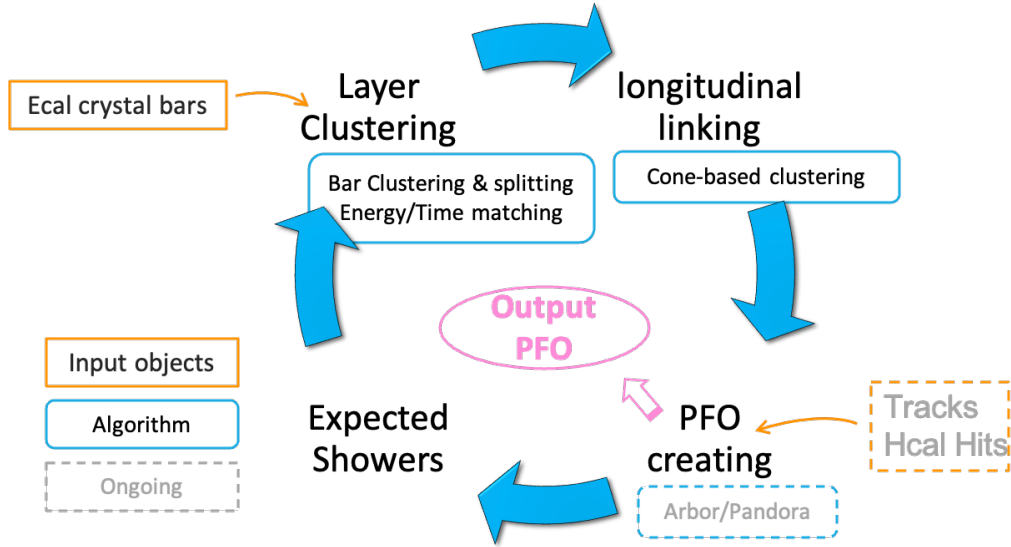


图 7-4 重建算法流程示意。重建以数字化中给出的晶体条作为输入，经历能量劈裂、超层内匹配、纵向连接的过程形成簇团，进行两次迭代后形成粒子流对象 PFO。

述 [142]，但对于横向发展仍缺乏统一结论。本文中我们参照 BESIII 实验的光子重建 [143]，使用二阶指数函数描述电磁簇射的横向能量沉积：

$$f(x) = p_0 \exp\left(-\frac{p_1 x}{R_M}\right) + p_3 \exp\left(-\frac{p_4 x}{R_M}\right), \quad (7-1)$$

其中  $x$  是与簇射中心的距离， $R_M$  是介质的莫里哀半径， $p_i$  是模型参数，可以通过对模拟样本的拟合得到，拟合结果如图 7-5。簇射横向形状应当与纵向深度有关，即  $p_i = p_i(\text{depth})$ ，但目前暂不做考虑。

在算法中，同一层的相邻晶体首先被聚类为簇团 (cluster)，其中能量超出相邻两块晶体的被定义为种子 (seed)。如果一个簇团中包含多个种子，并且计算得到的簇团二次矩  $S = \sum_i \frac{E_i(x_i - \bar{x})^2}{E_{tot}}$  大于阈值，则认为这一簇团由多个簇射 (shower) 组成，每个种子对应一个簇射。此时簇团中的每个晶体条  $i$  中的能量都被认为包含来自所有簇射  $\alpha$  的贡献，并据此计算出一组权重值：

$$E_{i\alpha}^{exp} = E_{seed\alpha} \times f(x_i), \quad (7-2)$$

$$w_{i\alpha} = E_{i\alpha}^{exp} / \sum_{\alpha} E_{i\alpha}^{exp}. \quad (7-3)$$

其中  $E_{seed\alpha}$  是判定  $\alpha$  簇射种子的能量， $x_i$  是晶体条  $i$  与簇射  $\alpha$  中心之间的间距， $E_{i\alpha}^{exp}$  即簇射  $\alpha$  预期在晶体条  $i$  处沉积的能量。然后对所有晶体条复制为  $\alpha$  个，对每个晶体重新赋予能量： $E_{i\alpha} = w_{i\alpha} \times E_i^{measured}$ ，并将其完全归于簇射  $\alpha$ 。这一计算过程经过了多次迭代，采用修正后的簇射重新计算簇射中心位置，再进行能量分配，直至迭代收敛。图 7-6 展示了这一算法的流程图。



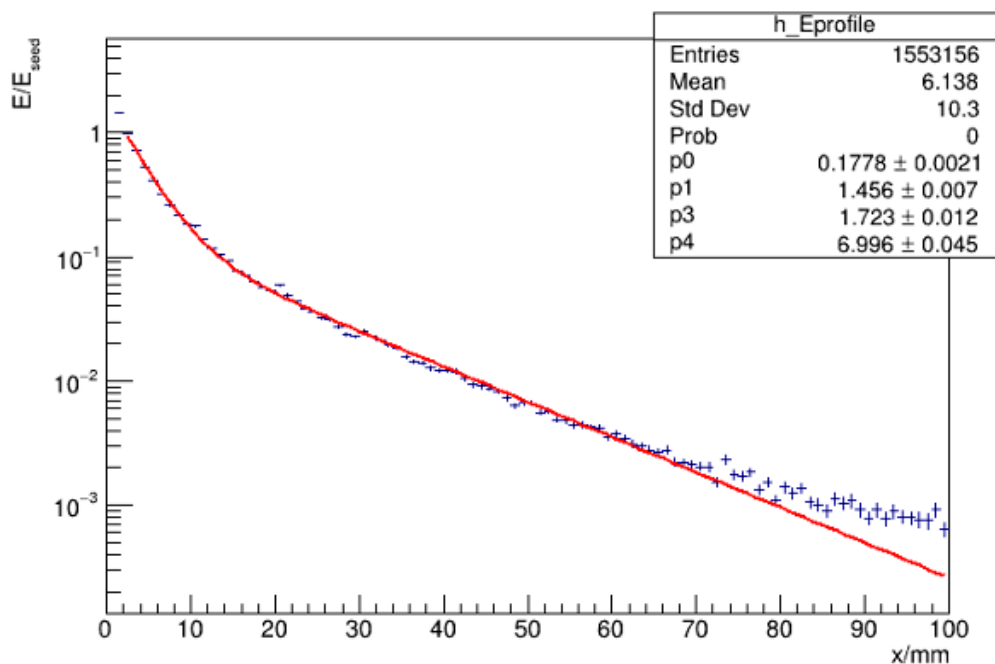


图 7-5 通过模拟光子样本拟合得到的电磁簇射横向发展分布。

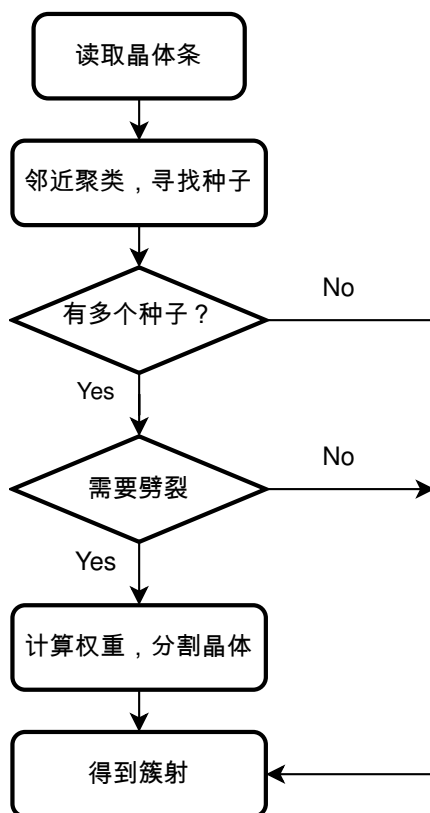


图 7-6 能量劈裂算法流程图。

### 7.3.2 能量-时间匹配算法

这一算法目的在于将晶体条结构重建为高颗粒度结构，并去除匹配过程中的 ghost hit。重建过程参考了 CALICE 合作组为闪烁体量能器 ScECAL 开发的条分割算法 (Strip Splitting Algorithm, SSA) [144]: 考虑同一超层中交叉排列的两组晶体条  $U$  条和  $V$  条，将每根条在纵向分割成  $1 \times 1 \times 1 \text{ cm}^3$  的小块，每一块中的能量分数由相邻层得到的能量分布决定。通过这种方式，原本  $1 \times 1 \times 40 \text{ cm}^3$  结构的探测单元可以等效重建为  $1 \times 1 \times 2 \text{ cm}^3$  颗粒度，能量为：

$$E_{\text{cell } k}^U = E_{\text{bar}}^U \times (E_{\text{bar } k}^V / \sum_i E_{\text{bar } i}^V), \quad (7-4)$$

$$E_{\text{cell } k}^V = E_{\text{bar}}^V \times (E_{\text{bar } k}^U / \sum_i E_{\text{bar } i}^U) \quad (7-5)$$

$k$  是  $U/V$  方向条中被分割成的小块的序号， $E_{\text{bar } k}^{V/U}$  是相邻层中与  $k$  单元交叉处的晶体条中的能量。

在这种思路下，如果一个 super-cell 被多个粒子同时击中，算法会在所有交叉点处重建出击中点，即 ghost-hit。之前有研究采用 CEPC 上典型多喷注过程  $e^+e^- \rightarrow \text{ZH} \rightarrow 4\text{jets}$  研究了多粒子击中一个  $40 \times 40 \text{ cm}^2$  super-cell 区域的情况 [145]，其中 1、2 和多个粒子同时击中的概率分别为 74%、15% 和 11%。而当对撞质心能量提升至 360 GeV ( $t\bar{t}$  mode) 时，喷注能量增加，多重数预期会进一步上升。因此在使用上述条分割算法前需要一个算法能够对此情况进行识别，选择特定  $U/V$  方向簇射进行分割，去除 ghost-hit。

我们考虑以下情形：在  $U$  和  $V$  方向上分别识别出  $M$  和  $N$  个簇射 ( $M > 1, N > 1$ )，此时若直接使用条分割算法，可重建出的总组合数为  $M \times N$ ，其中至少有  $\max(M, N)$  个为真实入射粒子产生的簇射，其他为 ghost-hit。而考虑到在 4 jet 末态的模拟结果中，多重数并不那么严重，我们选择在重建时忽略多于  $\max(M, N)$  个粒子同时入射的情况。考虑使用以下信息：

- 晶体条 SiPM 双端读出时的时间信息可以给出簇射的粗略定位，并且应当与相邻层给出的定位一致；
- 相邻两层中沉积能量应当一致，或能够根据电磁簇射纵向发展给出预测；

基于此两点考虑，我们对每个组合点定义一个  $\chi^2_{point}$ ：

$$\chi_E^2 = \frac{(E_U - E_V)^2}{\sigma_E^2}; \quad (7-6)$$

$$\chi_{Ti}^2 = \frac{(z_T - z_i)^2}{\sigma_s^2 + \sigma_{z(t)}^2}, \quad i = U/V; \quad (7-7)$$

$$\chi_{point}^2 = w_E \cdot \chi_E^2 + w_T \cdot (\chi_{Tu}^2 + \chi_{Tv}^2)/2; \quad (7-8)$$

其中  $z_T$  为根据时间计算出的簇射位置， $\sigma_s$  是探测器单元的本征位置分辨， $\sigma_{z(t)}$  是根据时间得到的位置的分辨。 $w_T$  和  $w_E$  是对于时间项和能量项的权重，通过调整其相对值可以调整算法对于时间和能量分辨的灵敏度。这一参数设置可以在后期进行优化。考虑以下情况利用  $\chi^2$  去除 ghost-hit：

- 如果  $M = N > 1$ ：在忽略额外多重数的情况下，此时全部  $N^2$  个组合点可以给出  $N!$  种不同入射图案的组合，每种组合可定义其  $\chi_{comb}^2 \equiv \sum_i^N \chi_{point}^2$ ，我们选定  $\chi_{comb}^2$  最小的图案为重建的组合方式，如图 7-7 所示。

- 如果  $M \neq N$ ，假定  $M > N$ ：所有组合点中拥有最小  $\chi_{point}^2$  的先被固定下来，对应的  $U/V$  方向簇射也被移除，再考虑剩余  $(M - 1) \times (N - 1)$  的图案；重复这一过程，直至  $U/V$  方向组合图案变为  $(M - N + 1) \times 1$ ，这些簇射再根据条分割算法重建为  $(M - N + 1)$  个簇射。这一过程中整体簇射图案被逐步确定，如图 7-8 所示。

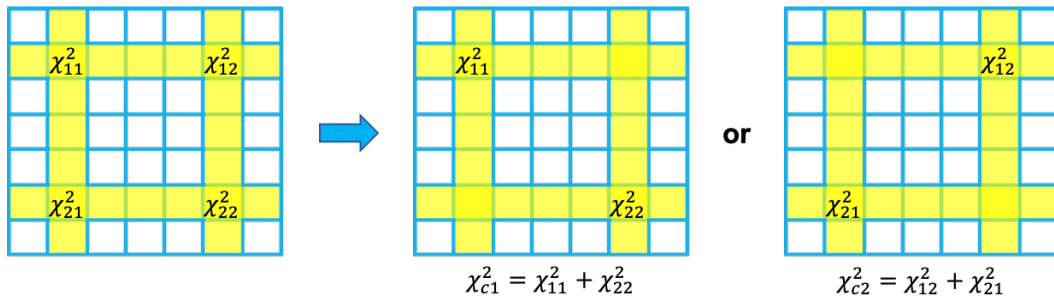


图 7-7  $M = N$  情况下的模式识别示意图。首先计算每个交叉点的  $\chi_{ij}^2$ ，根据  $\chi_{ij}^2$  计算组合出图案对应的  $\chi^2_{comb}$ ，最终选定  $\chi^2_{comb}$  的图案。

这一模式识别过程在每一个超层中独立进行，完成后所有重建出的二维簇团作为算法结果进行输出，在下一步中使用圆锥角聚类的方法组合成三维簇团。我们正在重新组织算法结构，综合纵向上所有超层的信息判定入射粒子模式，去除 ghost-hit，预期可以进一步压低错误识别产生的假簇团。

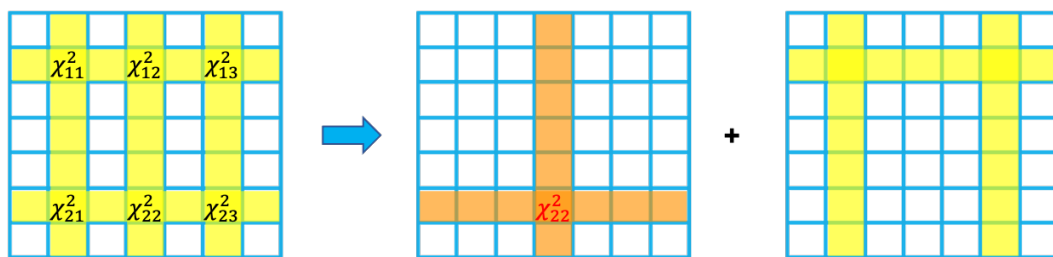


图 7-8  $M > N$  情况下的模式识别示意图。此处我们假设  $\chi_{22}^2$  是所有  $\chi^2$  中最小的，点 (2,2) 即被固定，对应的一组  $UV$  簇射也被移除，图案被简化为  $2 \times 1$ ，可直接确定。

### 7.3.3 软件实现

重建算法的代码实现考虑了模块化、可扩展的要求，参考 PandoraSDK [21] 进行整体架构的设计，包括：

- 内部事例模型 Self-EDM：定义算法中需要用到事例模型，提供一定功能函数方便算法使用，现有数据模型包括：

- CaloUnit：描述着火晶体条，包括单元编号 cellID、位置、双端读出能量和时间；

- Calo1DCluster：由一组 CaloUnit 组成，描述一层中相邻着火晶体条组成的簇团或劈裂后产生的簇射，包含簇团种子信息；

- Calo2DCluster：由两组 Calo1DCluster 组成，描述一个超层中一组交叉晶体条或经过二维重建后的二维簇团；

- Calo3DCluster：由多组 Calo2DCluster 组成，描述经过纵向连接后在纵向多层内的三维簇团；

- CaloHit：对应于 edm4hep::CalorimeterHit 的内部数据模型，用于存储经过重建后得到的  $1 \times 1 \times 2\text{cm}^3$  颗粒度的击中点和 HCAL 中的击中点；

- CaloCluster：对应于 edm4hep::Cluster 的内部数据模型，用于存储重建后的簇团、HCAL 簇团，作为接口转换到 edm4hep::Cluster；

内部事例模型随算法设计一同开发，目前仍在调整和修改中。同时算法中包含一个 DataCol 类管理事例中的所有数据，并管理内存。

- 接口类 Client App：提供 EDM4HEP 的事例模型 (Event Data Model, EDM) 和算法自身事例模型 (Self-EDM) 之间的转换，作为算法输入输出的接口，目前包括：

- MCParticleCreator：从输入文件中读入 MCParticle；

- TrackCreator：从输入文件中读入径迹，计算径迹外推到量能器中的击中

点;

- CaloHitsCreator: 从输入文件中读入各种类型的 CalorimeterHit, 根据探测器解码器对读入 Hit 编号进行解码得到位置;
- OutputCreator: 算法运行结束后读取 DataCol, 从中提取数据构造输出的粒子流对象 ReconstructedParticleCollection 和簇团 ClusterCollection;
  - 算法管理函数 AlgorithmManager: 统一注册调用的算法模块和各自参数, 按照运行脚本中传入的配置信息调用子算法, 其中包含算法数据库 AlgorithmMap, 注册后的算法按调用顺序储存在 AlgorithmMap 中, 并配置好算法执行参数;
    - 子算法类 Sub-Algorithm: 重建算法的核心, 统一调用 DataCol 读取输入数据、存储输出数据, 实现特定功能, 目前包括:
      - GlobalClusteringAlg: 在全探测器上进行邻近聚类;
      - LocalMaxFindingAlg: 在聚类后的一维簇团中寻找局域最大, 将其标记为种子;
      - EnergySplittingAlg: 根据找到的种子和簇团性质按照能量劈裂算法对重叠簇团进行能量劈裂;
      - EnergyTimeMatchingAlg: 按照能量-时间匹配算法处理 ghost-hit, 重建出二维簇团和高颗粒击中点;
      - ConeClusteringAlg: 将每层中的二维簇团进行纵向连接得到三维簇团;
      - ClusterMergingAlg: 对簇团质量进行初步判断, 将杂散小簇团合并到邻近簇团中;

所有子算法继承虚基类 virtual Algorithm, 定义了虚函数 ReadSettings(), Initialize(), RunAlgorithm(), ClearAlgorithm() 以供管理函数 AlgorithmManager 统一调用。子算法模块仍在开发和优化中。

架构模式如图 7-9。算法的整体控制由 EcalRecAlg 完成, 遵从 Gaudi 框架下的运行流程 (算法 2)。

### 7.3.4 初步性能

由于全探测器重建尚未完成, 目前仅能给出初步的性能结果。全吸收型晶体量能器相比于 Si-W 取样型量能器最大的优势在于对光子优秀的能量分辨, 因此我们首先对这一结果进行了测试。我们产生了 1 GeV~ 30 GeV 的单光子枪样本, 正入射进一个模块中央, 重建能量的分辨率如图 7-10所示。对数据点采用  $\sigma_E/E = a \oplus \frac{b}{\sqrt{E}} \oplus \frac{c}{E}$  进行拟合, 得到结果  $b = 1.45\%$ ,  $c = 3.80\%$ 。c 项来自于对晶体量能器中小簇射团处理不完美而产生的能量泄漏, 这计划在以后进行优

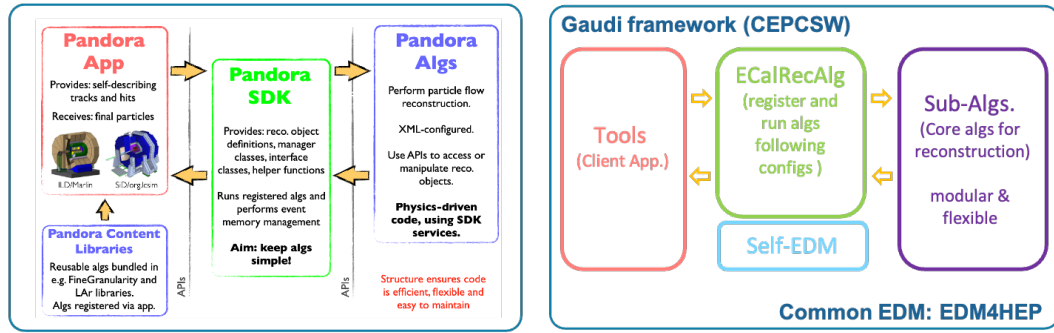


图 7-9 重建软件代码框架（右）。左为参考的 PandoraSDK 代码框架 [21]。

化和补偿。

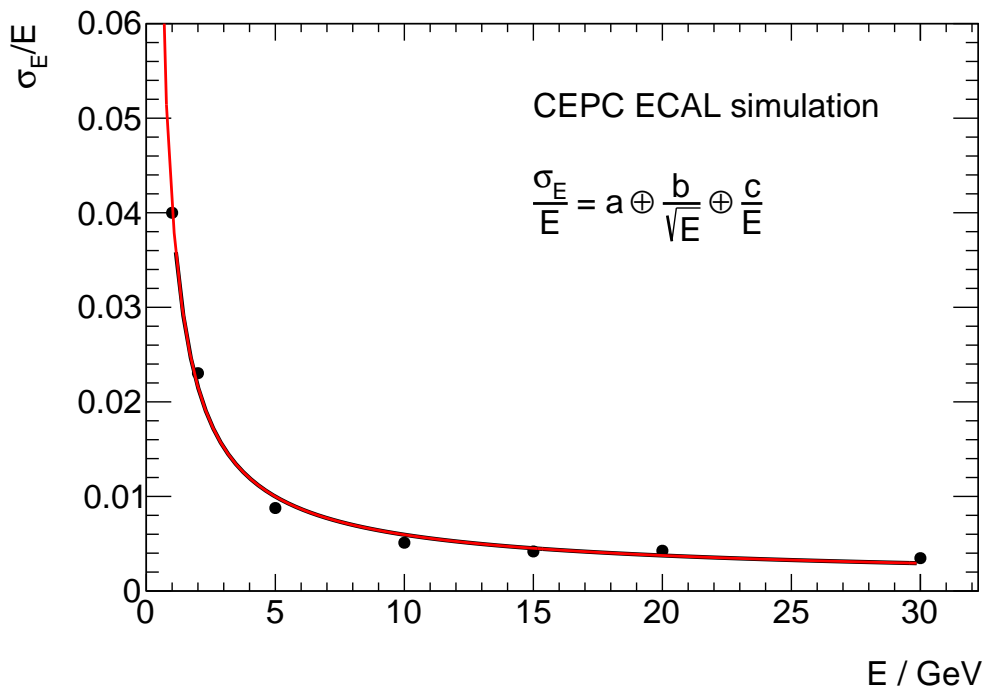


图 7-10 晶体 ECAL 中光子的相对能量分辨。红线为使用  $\sigma_E/E = a \oplus \frac{b}{\sqrt{E}} \oplus \frac{c}{E}$  的拟合结果，从拟合中得到  $a = 0.$ ,  $b = 1.45\%$ ,  $c = 3.80\%$ 。

对光子精确的位置重建对于 PFA 也非常重要，因此我们对位置分辨也进行了相应测试。我们另产生了一组 5 GeV 的单光子样本垂直打入探测器，产生位置沿束流方向（z 向）平移，图 7-11 展示了重建出的位置偏移、位置分辨随产生点变化。所有点点位置分辨均在 1 mm 范围内，与 1 cm 晶体块对应的本征分辨相符，考虑到晶体的  $R_M$  这一结果已符合 PFA 的要求。随入射点变化的位置偏移来自于晶体尺寸，这一效应可以通过对重建位置加权进行修正，这已在 BESIII 实验上应用 [143]。

Ghost-hit 问题的解决通过同时入射在正方形对角处的两个 5 GeV 光子检验。

**Algorithm 2:** 重建算法 EcalRecAlg**Input:** 数字化后晶体条 edm4hep::CalorimeterHitCollection;**Output:** 重建后的粒子流对象 PFO

(edm4hep::ReconstructedParticleCollection)

- 1 初始化：设置全局参数 GlobalSettings;
- 2 初始化：初始化接口函数 ObjectCreators，从配置文件中读取对象名,;
- 3 初始化：算法管理函数 AlgorithmManager 注册算法;
- 4 初始化：AlgorithmManager 配置算法调用顺序，从配置文件中为每个算法配置执行参数;
- 5 初始化：初始化几何服务 GeoService、探测器解码器 Decoder;
- 6 初始化：注册分析文件 ntuple;
- 7 **foreach** 事例 **do**
- 8     ObjectCreators 读取对象数据，暂存在 DataCol;
- 9     AlgorithmManager 按配置调用算法执行计算，结果储存进 DataCol;
- 10    OutputCreator 读取 DataCol 中数据，写出重建后对象  
      ReconstructedParticleCollection;
- 11    写入 ntuple;
- 12    DataCol 清空内存;
- 13 写出所有数据，清空内存;

两光子间距（正方形边长）从 5 mm 扫描至 70 mm，在正确位置重建出正确能量的两个光子的效率如图 7-12 所示。可看出 ghost-hit 问题已完全解决，并且当两光子间距达到  $30 \text{ mm} \times \sqrt{2}$  时重建效率可达 90%。考虑到在  $e^+e^- \rightarrow ZH \rightarrow 4 \text{ jets}$  事例中，超过一般的多粒子入射情景仅包含两个粒子，其中双带电粒子 (C+C)、一带电一中性 (C+N) 和双中性 (N+N) 的比例分别为 28%、40% 和 32%。有带电粒子时径迹信息能够辅助进行 ECAL 中簇射定位，N+N 情形是对此设计的主要挑战。目前的结果证明这一问题有希望通过后期重建算法进行解决。

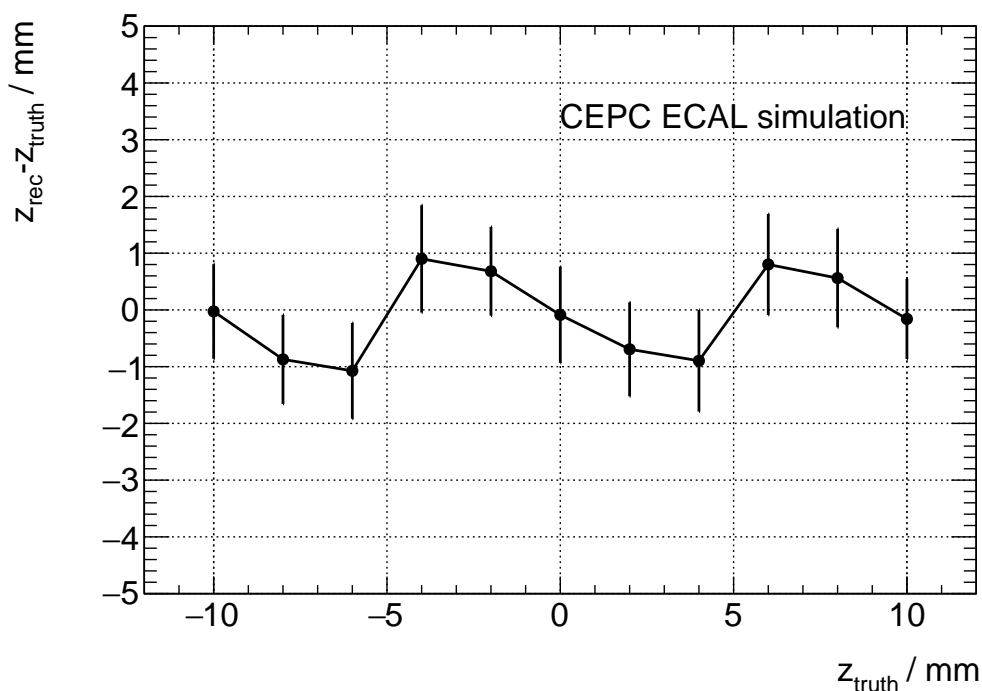


图 7-11 重建位置偏移随入射位置的变化，图中误差棒对应重建位置分辨。

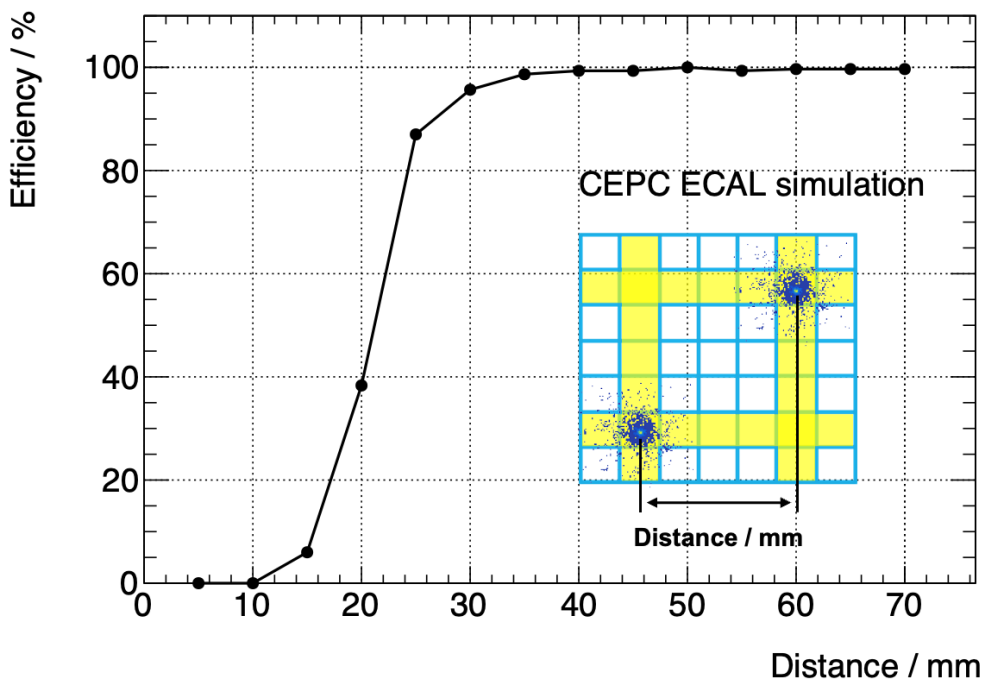


图 7-12 入射光子示意图和重建效率随间距变化。重建效率定义为满足以下 3 个条件的重建事例的比例：重建出至少 2 个簇团，簇团能量正确 ( $\frac{2}{3}E_{\text{truth}} < E_{\text{rec}} < \frac{4}{3}E_{\text{truth}}$ ) 且簇团位置正确 ( $|Pos_{\text{truth}} - Pos_{\text{rec}}| < 10 \text{ mm}$ )。



## 7.4 总结与展望

CEPC 第四代探测器的设计工作仍在进行中，其可行性验证和性能分析是前期工作重点。本工作着重于这一问题，构造了探测器几何，设计数字化模型，开发了初步算法对交叉晶体条 ECAL 方案的可行性进行了研究。本人开发的基于能量-时间的匹配算法解决了交叉条结构的 ghost-hit 问题，引入电磁簇射形状的能量劈裂算法也使得在晶体量能器中簇射能量可以得到正确重建，拓展了 PFA 对硬件的要求。相关文章已撰写完成，正在内部审阅中。目前的工作仍非常初步，有很多细节需要补充，如：

- 没有考虑量能器中除灵敏晶体外的其他物质，如读出电子学、包装材料、线缆、冷却系统、支撑系统等；
- 数字化方案过于简化，能量衰减没有考虑，时间分辨需要更细致的研究；
- 簇团重建采用圆锥角重建，可以进行优化以求更好的能量分辨；
- 没有完成全探测器重建，暂时没有得出关键指标 BMR；
- 不考虑更多粒子同时入射的复杂度问题，缺乏这种简化对喷注能量分辨的定量研究；

与此同时，我们还在考虑优化算法，引入新的方法进行簇射的模式识别和重建，如引入霍夫变化识别光子主轴、接入 ArborPFA 实现全探测器重建等。而在代码框架方面，我们也在进行不断优化，使其结构更加合理、易读易开发，以期望以后进行推广和协作开发。



## 第 8 章 CEPC 上 $\sigma(ZH) \times Br(H \rightarrow \gamma\gamma)$ 测量精度的研究

### 8.1 研究背景

$H \rightarrow \gamma\gamma$  衰变是研究希格斯物理的黄金道之一，其优点在于实验上从未态两个高能光子中可以清楚地重建出希格斯粒子，理论上希格斯粒子与双光子作用需通过顶夸克圈或有质量玻色子圈进行，BSM 新物理很有可能参与到这一过程中来，相关耦合强度的测量对寻找新物理非常重要。CEPC 作为希格斯工厂， $H \rightarrow \gamma\gamma$  是必不可少的研究过程。在目前的预研阶段， $H \rightarrow \gamma\gamma$  道的测量精度也可以作为 ECAL 设计和优化的重要参考。目前 LHC Run 2 对希格斯粒子双光子道信号强度的测量结果为  $1.04_{-0.09}^{+0.10}$  (ATLAS [9]) 和  $1.03_{-0.09}^{+0.11}$  (CMS [146])，在完成 HL-LHC 取数后期望的最佳测量精度为 2.5% [121]。

本分析为基于蒙特卡洛模拟的对 CEPC 上  $\sigma(ZH) \times Br(H \rightarrow \gamma\gamma)$  测量精度的研究，模拟样本采用对 CEPC 基准探测器的全模拟，根据  $Z$  衰变模式分为  $q\bar{q}\gamma\gamma$ 、 $\mu^+\mu^-\gamma\gamma$  和  $\nu\bar{\nu}\gamma\gamma$  三个子道，在每个道中用多变量分析区分信号和本底事例。同时我们引入了 ATLAS 分析中的滋扰参数方法对系统误差进行了初步估计，最终给出了在  $5.6 \text{ fb}^{-1}$  (CEPC CDR 中预期数据量) 和  $20 \text{ fb}^{-1}$  (Snowmass 2021 中预期数据量) 下  $\sigma(ZH) \times Br(H \rightarrow \gamma\gamma)$  的测量精度，包括统计误差和系统误差。考虑到正在开发中的晶体 ECAL，我们还在  $q\bar{q}\gamma\gamma$  道中以简化流程研究了 ECAL 能量分辨对于这一精度的相对影响。

### 8.2 模拟样本

在  $\sqrt{s} \sim 240 \text{ GeV}$  的正负电子对撞机如 CEPC 上，希格斯粒子主要通过  $e^+e^- \rightarrow ZH$ 、 $e^+e^- \rightarrow \nu\bar{\nu}H$  和  $e^+e^- \rightarrow e^+e^-H$  过程产生。考虑到三个过程的产生截面中  $ZH$  为主导，本分析中将  $e^+e^- \rightarrow ZH \rightarrow f\bar{f}\gamma\gamma$  过程作为主要信号，并根据  $Z$  玻色子衰变将事例分为 3 个子道： $Z \rightarrow q\bar{q}$ ， $Z \rightarrow \mu^+\mu^-$ ， $Z \rightarrow \nu\bar{\nu}$ 。我们放弃了  $Z \rightarrow e^+e^-$  道的分析，因为在正负电子对撞机上伴随有初末态辐射 (ISR / FSR) 光子的 Bhabha 过程事例将非常多，这一分析道的信噪比将很低； $Z \rightarrow \tau^+\tau^-$  衰变道也由于  $\tau$  喷注鉴别的复杂性和现阶段的低效率而不做考虑。因此本底过程主要为伴随 ISR / FSR 光子两费米子过程  $e^+e^- \rightarrow f\bar{f} + \gamma\gamma$ ，其在 CEPC 上占据主导的产生截面，其他 4 费米子本底和希格斯共振本底被认为可以忽略。

这些标准模型事例由 Whizard [126] 在 LO 精度下产生，并采用 PYTHIA 6 描

述  $q\bar{q}$  末态的部分子簇射，其中强子化过程的参数基于 LEP 实验的结果 [147]。事例产生时考虑了 ISR 和 FSR 过程，以及在 CEPC 中由于束流同步辐射导致的 0.16% 电子束能量弥散 [17]。表 8-1 中总结了三个分析道使用的物理过程、对应 240 GeV 下的产生截面，以及蒙特卡洛样本数量。样本详情可见 [148]。

CEPC 基准探测器结构如 6.2.2 中介绍，对探测器响应的模拟由 MokkaPlus [127] 完成。为节省计算资源，仅信号事例进行了探测器全模拟，本底事例则对产生的真实样本中末态稳定粒子进行了基于探测器效率和分辨的参数化的涂抹 (smearing)。事例产生和模拟在 CEPCsoft 软件框架下进行。

物理过程	$\sigma$	样本统计量
$q\bar{q}\gamma\gamma$ sub-channel		
$e^+e^- \rightarrow ZH \rightarrow q\bar{q}\gamma\gamma$	0.31 fb	100 k
$e^+e^- \rightarrow q\bar{q}$	54.1 pb	20 M
$\mu^+\mu^-\gamma\gamma$ sub-channel		
$e^+e^- \rightarrow ZH \rightarrow \mu^+\mu^-\gamma\gamma$	0.15 fb	100 k
$e^+e^- \rightarrow \mu^+\mu^-$	5.3 pb	20 M
$\nu\bar{\nu}\gamma\gamma$ sub-channel		
$e^+e^- \rightarrow ZH \rightarrow \nu\bar{\nu}\gamma\gamma$	0.11 fb	100 k
$e^+e^- \rightarrow \nu\bar{\nu}H \rightarrow \nu\bar{\nu}\gamma\gamma$		
$e^+e^- \rightarrow \nu\bar{\nu}$	54.1 pb	20 M

表 8-1 各物理过程的产生截面和使用的蒙特卡洛样本统计量。在  $q\bar{q}\gamma\gamma$  和  $\mu^+\mu^-\gamma\gamma$  道中仅考虑 ZH 过程为信号，在  $\nu\bar{\nu}\gamma\gamma$  中 ZH  $Z \rightarrow inv.$  和  $W/Z$  fusion 过程均被考虑为信号。

### 8.3 物理对象重建与事例选择

如 6.3 中所述，CEPC 使用粒子流算法 ArborPFA 进行事例重建。径迹重建采用 Clupatra 算法 [149] 进行，得到的径迹与量能器中的击中点一起输入给 ArborPFA，ArborPFA 在进行聚类 and 径迹-簇团匹配后将其组装成粒子流对象 PFO。这些 PFO 可初步分为带电粒子、备选光子、备选中性强子和能量碎片 (unassociated fragments)，其中通过一些簇射形状要求并且没有匹配径迹的备选光子即被认为是重建出的光子。目前在 CEPC 中没有考虑转化光子的贡献，这会导致在桶部丢失 5 - 10% 的光子，在前端丢失 25% 的光子 [10]。带电轻子 ( $e^\pm, \mu^\pm$ ) 从带电 PFO 中挑选，使用内置于 ArborPFA 中的基于似然函数的粒子鉴别算法 LICH [150] 区

分电子、缪子和强子。喷注的重建采用 Durham 算法进行 [136]，在排除感兴趣的对象后将其余 PFO 聚类为喷注。此分析中没有对喷注味道进行鉴别。目前重建粒子的能量刻度基于蒙卡样本，相信在未来 CEPC 正式取数后可以使用如  $Z \rightarrow q\bar{q}$  等过程实现更准确的刻度。

分析中的事例选择根据分析道中物理过程的拓扑结构进行，主要为初步去除本底，更精细的信号-本底区分工作则由多变量分析完成。 $ZH \rightarrow \nu\bar{\nu}\gamma\gamma$  道中要求末态仅有 2 个光子，且事例的丢失能量  $E_{missing}$  大于 10 GeV (inclusive 2 photons);  $ZH \rightarrow \mu^+\mu^-\gamma\gamma$  道中则要求能够挑选出 2 个光子和 2 个缪子，并且没有其他粒子，丢失能量和丢失质量均小于 10 GeV，且缪子对的不变质量应接近 Z 玻色子质量 (exclusive 2 photons and 2 muons); 在  $ZH \rightarrow q\bar{q}\gamma\gamma$  道中，两个能量最高的光子先被挑出，其他粒子按照 Durham 方法重建为 2 个喷注 (exclusive 2 photons and 2 jets)。除此以外，我们还对一些动力学变量进行了要求，具体条件和筛选效率见表 8-2, 8-3, 8-4。

筛选条件	Higgs 信号	$q\bar{q}\gamma\gamma$ 本底
Exclusive 2 jets and 2 photons	85.56%	69.57%
$E_{\gamma 1} > 25$ GeV	100.00%	2.35 %
$E_{\gamma 2} \in [35, 95]$ GeV	98.37%	35.33%
$\cos \theta_{\gamma\gamma} > -0.95$	95.20%	68.01%
$\cos \theta_{jj} > -0.95$	90.86%	85.54%
$pT_{\gamma 1} > 20$ GeV	93.42%	56.94%
$pT_{\gamma 2} > 30$ GeV	93.25%	54.54%
$m_{\gamma\gamma} \in [110, 140]$ GeV	97.50%	21.14%
$E_{\gamma\gamma} > 120$ GeV	99.47%	98.41%
$\min  \cos \theta_{\gamma j}  < 0.9$	71.67%	48.05%
总筛选效率	44.08%	0.01%
5.6 ab <sup>-1</sup> 下事例数	766.64	26849.38

表 8-2  $q\bar{q}\gamma\gamma$  道中的筛选条件和效率。 $\gamma 1(\gamma 2)$  定义为相对低能量 (高能量) 的光子;  $\cos \theta_{\gamma\gamma}(\cos \theta_{jj})$  是双光子 (双喷注) 系统的极角;  $\min |\cos \theta_{\gamma j}|$  是光子-喷注对中最小的  $\cos \theta$ 。

筛选条件	Higgs 信号	$\mu^+\mu^-\gamma\gamma$ 本底
Exclusive 2 muons and 2 photons	70.18%	5.18%
$E_\gamma > 35$ GeV	99.21%	8.39%
$ \cos\theta_\gamma  < 0.9$	83.79%	38.14%
$pT_{\gamma 1} \in [10, 70]$ GeV	99.84%	86.30%
$pT_{\gamma 2} \in [30, 100]$ GeV	99.96%	95.59%
$m_{\gamma\gamma} \in [110, 140]$ GeV	98.08%	37.62%
$M_{\gamma\gamma}^{recoil} \in [85, 105]$ GeV	80.12%	21.29%
$E_{\gamma\gamma} \in [125, 145]$ GeV	99.88%	95.86%
总筛选效率	45.69%	0.01%
$5.6 \text{ ab}^{-1}$ 下事例数	39.32	2662.77

表 8-3  $\mu^+\mu^-\gamma\gamma$  道中筛选条件和事例数。 $\gamma 1(\gamma 2)$  定义与  $q\bar{q}\gamma\gamma$  道中相同； $M_{\gamma\gamma}^{recoil}$  是根据 CEPC 质心能量 240 GeV 计算的双光子系统反冲质量： $(M_{\gamma\gamma}^{recoil})^2 = (\sqrt{s} - E_{\gamma\gamma})^2 - p_{\gamma\gamma}^2 = s - 2E_{\gamma\gamma}\sqrt{s} + m_{\gamma\gamma}^2$ 。

筛选条件	Higgs 信号	$\nu\bar{\nu}\gamma\gamma$ 本底
Inclusive 2 photons	85.51%	0.34%
$E_\gamma > 30$ GeV	99.81%	20.13%
$ \cos\theta_\gamma  < 0.8$	70.48%	11.56%
$pT_\gamma > 20$ GeV	99.97%	99.26%
$M_{missing} > 60$ GeV	98.17%	99.71%
$m_{\gamma\gamma} \in [110, 140]$ GeV	97.51%	22.86%
$E_{\gamma\gamma} \in [120, 150]$ GeV	99.16%	99.58%
总筛选效率	57.08%	0.002%
$5.6 \text{ ab}^{-1}$ 下事例数	335.89	3640.20

表 8-4  $\nu\bar{\nu}\gamma\gamma$  道中筛选条件和事例数。 $M_{missing}$  是根据所有可见粒子计算得到的丢失质量。

## 8.4 多变量分析

本分析采用梯度提升决策树 BDTG [73] 来进一步区分信号和本底，在每个道中 ZH 过程和 2 费米子过程分别作为训练的信号和背景，且蒙卡样本中的所有事例被随机分成 2 份，用于做 2-fold 检验 [151] 以避免过训练。由于没有相关研究提供参考，我们考虑按以下方式构造训练变量：

- 参与构造变量的信息为所有末态粒子的四动量，从中构造出光子、费米子和多粒子系统的动量  $p$ 、横动量  $p_T$ 、能量  $E$ 、角度  $\cos\theta$ 、反冲质量  $M_{recoil}$ ，两个对象之间的  $\Delta P$ 、 $\Delta E$ 、 $\Delta\Phi$ 、 $\Delta\cos\theta$ 、 $\Delta R$ ，以及事例的丢失质量  $M_{missing}$ 。
- 采用式 4-1 中的  $\langle S^2 \rangle$  表征变量的区分能力，对其进行排序。
- 为构造 8.5 中描述的二维统计模型，要求训练变量与  $m_{\gamma\gamma}$  之间的线性相关系数  $|\text{Corr}_{v-m_{\gamma\gamma}}| < 30\%$ ，以此保证  $m_{\gamma\gamma}$  与 BDTG 响应值之间的低相关性。
- 为避免训练冗余变量，要求两个训练变量之间的线性相关系数  $|\text{Corr}_{v_1-v_2}| < 40\%$ 。两个变量中  $\langle S^2 \rangle$  更小的被去除。

经过以上条件，我们选出了用于 BDTG 训练的变量，其定义和分辨能力在表 8-5~8-7 中，变量分布如图 8-1~8-3。

变量	定义	分辨能力 $\langle S^2 \rangle$
$pT_{\gamma 1}$	次领头光子横动量	0.209
$\cos\theta_{\gamma 2}$	领头光子极角	0.197
$\Delta\Phi_{\gamma\gamma}$	两光子之间的方位角	0.147
$\min\Delta R_{\gamma,j}$	光子和喷注间 $\Delta R$ 的最小值	0.054
$E_{j1}$	次领头喷注能量	0.041
$\Delta\Phi_{\gamma\gamma,jj}$	两光子和两喷注系统之间的方位角	0.033
$pT_{j2}$	领头喷注的横动量	0.032
$\cos\theta_{j1}$	次领头喷注极角	0.032
$\cos\theta_{\gamma\gamma,jj}$	两光子和两喷注极角差 $\cos(\theta_{\gamma\gamma} - \theta_{jj})$	0.024
$\cos\theta_{\gamma 1,j1}$	$\gamma 1$ 和 $j1$ 之间极角差 $\cos(\theta_{\gamma 1} - \theta_{j1})$	0.023

表 8-5  $q\bar{q}\gamma\gamma$  道中 BDTG 训练输入变量。

变量	定义	分辨能力 $\langle S^2 \rangle$
$\min \Delta R_{\gamma,\mu}$	光子和缪子间 $\Delta R$ 的最小值	0.335
$E_{\mu\mu}$	双缪子系统能量	0.259
$\cos \theta_{\gamma 1,\mu 1}$	$\gamma 1$ 和 $\mu 1$ 之间极角差	0.189
$E_{\gamma 2}$	领头光子能量	0.160
$\Delta \Phi_{\gamma\gamma}$	两光子之间方位角	0.090
$\cos \theta_{\gamma 2}$	领头光子极角	0.072
$\Delta \Phi_{\gamma\gamma,\mu\mu}$	两光子系统和两缪子系统之间方位角	0.034
$\cos \theta_{\mu 1}$	次领头缪子极角	0.014

表 8-6  $\mu^+\mu^-\gamma\gamma$  道中 BDTG 训练输入变量。

变量	定义	分辨能力 $\langle S^2 \rangle$
$pT_{\gamma 1}$	次领头光子横动量	0.089
$\cos \theta_{\gamma 2}$	领头光子极角	0.079
$\Delta \Phi_{\gamma\gamma}$	两光子之间方位角	0.054
$pT_{t_{\gamma\gamma}}$	双光子横动量在与两个光子横动量之差垂直的方向上的投影	0.042
$pT_{\gamma 2}$	领头光子横动量	0.037

表 8-7  $\nu\bar{\nu}\gamma\gamma$  道中 BDTG 训练输入变量。



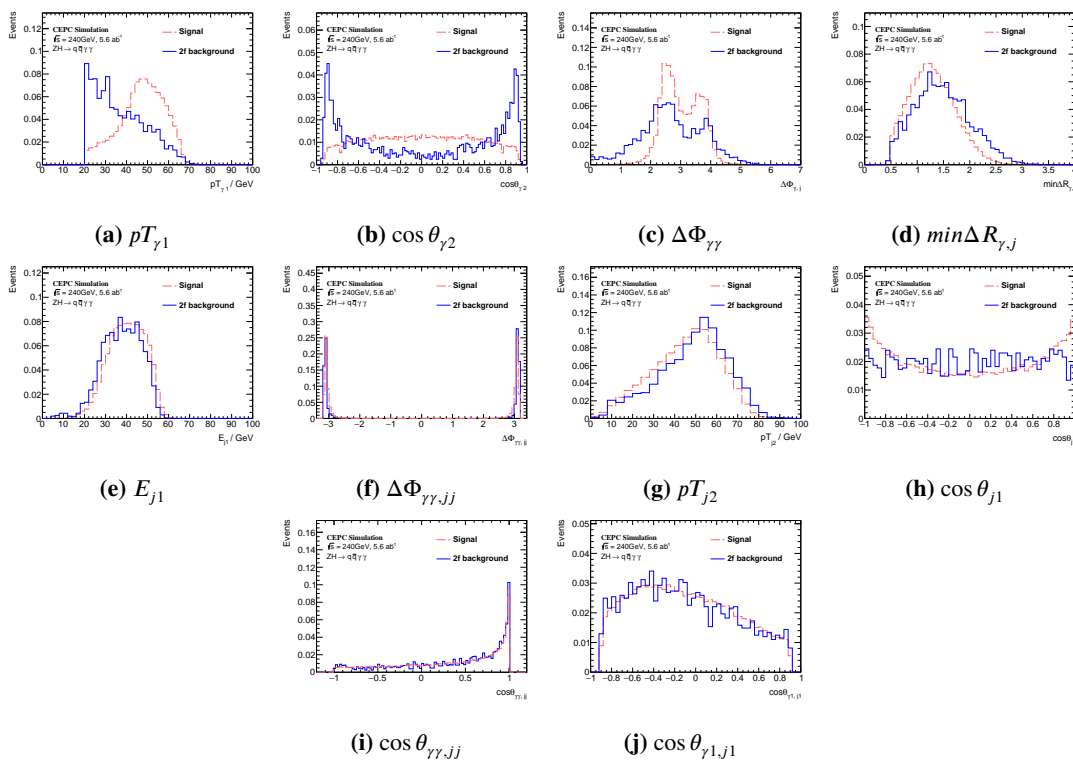


图 8-1  $q\bar{q}\gamma\gamma$  道中训练变量分布 (已归一化)。

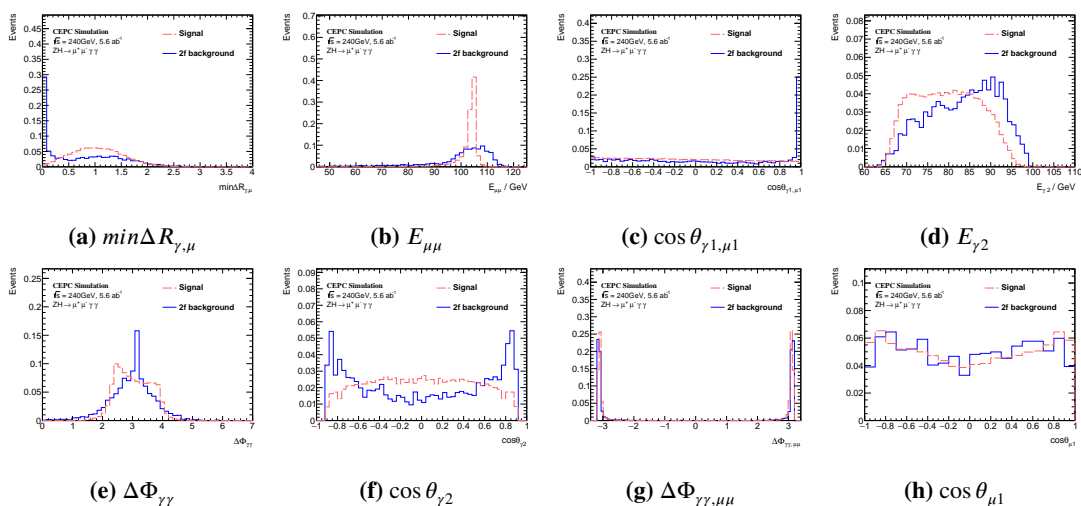


图 8-2  $\mu^+\mu^-\gamma\gamma$  道中训练变量分布 (已归一化)。

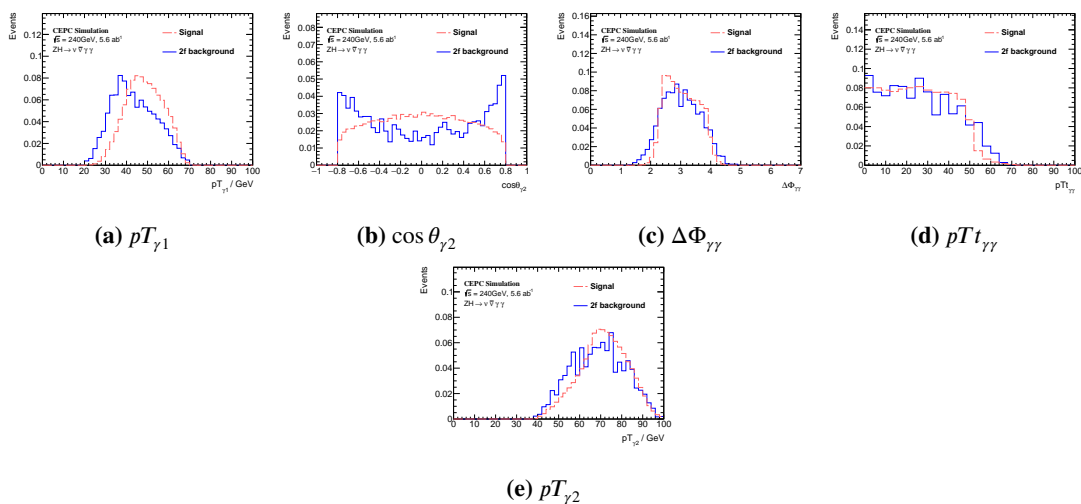


图 8-3  $\nu\bar{\nu}\gamma\gamma$  道中训练变量分布（已归一化）。

## 8.5 信号与背景建模

为提高信号和本底区分能力，本分析构建了  $m_{\gamma\gamma} \times \text{BDTG}$  的二维模型。其中  $m_{\gamma\gamma}$  的信号模型采用与 5.6.1 相同的双边水晶球函数描述，对信号蒙卡样本拟合得到的希格斯共振峰宽度在  $q\bar{q}\gamma\gamma / \mu^+\mu^-\gamma\gamma / \nu\bar{\nu}\gamma\gamma$  道中分别为 2.81 GeV, 2.68 GeV 和 2.74 GeV，如图 8-4。对  $m_{\gamma\gamma}$  连续本底来说，我们测试了多个平滑函数，包括 1、2 阶切比雪夫多项式，1、2 阶指数函数和 1、2 阶多项式，比较其  $\chi^2/\text{N dof}$ ，选取  $\chi^2/\text{N dof}$  最小的作为本底模型。三个分析道本底拟合的  $\chi^2/\text{N dof}$  值如表 8-8，拟合情况如图 8-5。最终在  $q\bar{q}\gamma\gamma$  和  $\mu^+\mu^-\gamma\gamma$  道中二阶切比雪夫多项式被选为本底模型， $\nu\bar{\nu}\gamma\gamma$  道中为一阶切比雪夫多项式。本分析中不考虑其他连续本底，如因对光子误判造成的可去除本底。

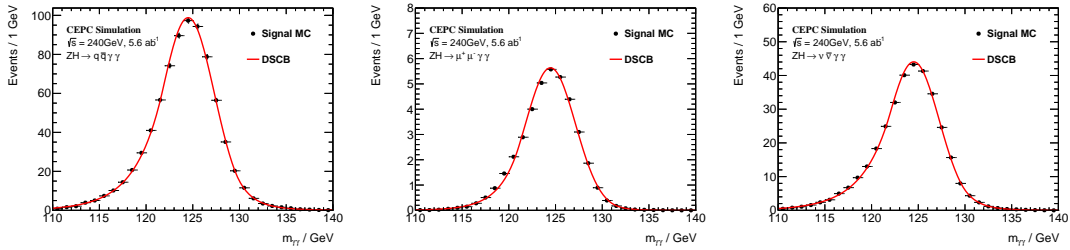


图 8-4 三个道中信号样本双边水晶球函数拟合。从左至右： $q\bar{q}\gamma\gamma$  道， $\mu^+\mu^-\gamma\gamma$  道， $\nu\bar{\nu}\gamma\gamma$  道。

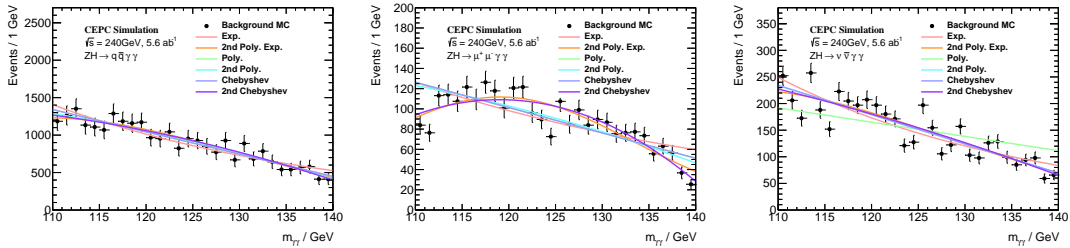


图 8-5 本底模型中测试的函数和拟合情况。从左至右： $q\bar{q}\gamma\gamma$  道， $\mu^+\mu^-\gamma\gamma$  道， $\nu\bar{\nu}\gamma\gamma$  道。

BDTG 的分布没有可供参考的函数理论模型，因此我们使用蒙卡样本的直方图构建分 bin 概率密度函数 (binned PDF, 后续也记为 HistPDF) 来作为信号和本底模型。在选择 BDTG 训练变量时我们已要求其应与  $m_{\gamma\gamma}$  无关，因此可认为训练出的 BDTG 也与  $m_{\gamma\gamma}$  无关，并在此假设下通过将  $m_{\gamma\gamma}$  模型与 BDTG 模型简单相乘构建二维模型  $f(m_{\gamma\gamma}, \text{BDTG}) = f(m_{\gamma\gamma}) \times \text{HistPDF}(\text{BDTG})$ ，忽略两者之间的相关性。作为检查，三个分析道信号样本中  $m_{\gamma\gamma}$  和 BDTG 的线性相关系数分别为 -3.45% ( $q\bar{q}\gamma\gamma$ ), -11.6% ( $\mu^+\mu^-\gamma\gamma$ ), 8.33% ( $\nu\bar{\nu}\gamma\gamma$ ), 背景样本的为 11.6% ( $q\bar{q}\gamma\gamma$ ), 28.2% ( $\mu^+\mu^-\gamma\gamma$ ), 28.4% ( $\nu\bar{\nu}\gamma\gamma$ )。较大的相关性可能会导致信号或和本底模型的偏差，从而影响测量结果。

	$q\bar{q}\gamma\gamma$	$\mu^+\mu^-\gamma\gamma$	$\nu\bar{\nu}\gamma\gamma$
1st order Exp.	0.941	5.423	3.786
2nd order Exp.	0.610	2.035	3.435
1st order Poly.	0.644	4.321	7.399
2nd order Poly.	0.600	3.758	3.439
1st order Chebyshev	0.644	4.321	3.320
2nd order Chebyshev	0.596	1.789	3.411

表 8-8 三个道中测试的 6 种本地函数拟合  $\chi^2/\text{Ndof}$  值。

## 8.6 CEPC 上初步系统误差分析

目前在 CEPC 上已有的研究还不支持进行完整的系统误差分析，但我们可以先人为估计系统误差项，引入相关分析方法，来研究系统误差对分析精度的影响。具体量化的系统误差研究仍需等待后续的理论计算、探测器表现和真实数据的结果。参考 5.7 中 ATLAS 分析上的系统误差分析方法，我们同样考虑将系统误差分为影响信号事例数的误差和影响信号  $m_{\gamma\gamma}$  模型的误差两类。本底事例数和本底模型误差通过在拟合时浮动响应参数考虑到测量结果中，BDTG 模型相关误差被整合为一项误差，应用到事例数误差中。考虑的误差来源包括  $q\bar{q}\gamma\gamma$  道中的理论计算、积分亮度、光子探测效率、光子能量刻度和重建、希格斯质量和 BDTG 模型。

### 8.6.1 理论误差

与强子对撞机不同，轻子对撞机上的反应过程很少受理论误差影响。高阶 QCD 修正、 $\alpha_s$  和 PDF 对信号产生截面计算的影响都不会进入到  $e^+e^- \rightarrow ZH \rightarrow f\bar{f}\gamma\gamma$  过程中，并且在此  $\sigma \times Br$  直接测量中，信号和本底事例数直接来自于对数据的拟合， $Br(H \rightarrow \gamma\gamma)$  的误差也被抵消掉。因此，在三个分析道中仅  $q\bar{q}\gamma\gamma$  道中末态部分子簇射和强子化过程的模拟会引入系统误差，与 ATLAS 分析类似，这可以采用多组蒙卡样本比较得到。受限于目前计算资源，我们没有进行响应事例产生，仅在  $q\bar{q}\gamma\gamma$  道的信号事例数上人为加入一项 0.5% 的误差表示此效应。其他理论误差均认为可以忽略。

### 8.6.2 实验误差

对实验误差的讨论考虑以下几项：

**积分亮度** CEPC 上的积分亮度由亮度量能器(Lumi-Cal)采用高统计量的 Bhabha 过程进行检测, CEPC CDR 中认为对积分亮度测量的误差可达到 0.1%, 直接作用到整体信号事例数上。

**事例堆积、探测器模拟** 相对于质子对撞, 轻子是基本粒子, 其通过对撞发生反应的截面要小几个数量级, 因此在轻子对撞机上不考虑事例堆积效应和隐藏事例 (underlying event)。目前我们认为对探测器几何的正确建模可以提供精确的探测器接受度和响应的模拟, 从而在蒙卡模拟事例中得到正确的估计, 相关系统误差暂不考虑。

**光子探测效率、能量刻度和分辨** 光子重建、鉴别和能量刻度的情况依赖于具体的探测器结构和相关算法, 在 CEPC CDR 的预研中认为这些误差可以控制在 1% 以下。当得到真实数据后, 我们也可以利用  $\pi^0 \rightarrow \gamma\gamma$ 、 $Z \rightarrow e^+e^- + \gamma$  等物理过程作为标准烛光对光子进行标定, 提高测量精度。目前我们假设光子重建效率误差为 1%, 光子能量刻度 (PES) 和能量分辨 (PER) 误差为 0.05%。缪子和喷注的相应误差暂不考虑, 在此光子道分析中认为其效应是次要的。我们通过随机去除 1% 光子, 或对光子能量向上/向下刻度 0.05%, 或对光子能量加入 0.05% smear 产生另外的模拟样本, 并采用相同的筛选条件进行筛选, 将其中事例数与原样本的事例数对比得到光子探测相关误差对信号事例数的影响  $\delta n^i = \frac{|n_{var}^i - n_{nom}^i|}{n_{nom}^i}$ 。得到的结果中, 光子效率对信号事例数影响约为 2%, 能量分辨和刻度对事例数的影响为  $O(0.01\%)$ 。同时, 我们考虑光子刻度和分辨对信号模型的影响, 这将作用于信号峰位  $\mu_{CB}$  和宽度  $\sigma_{CB}$ 。这一估计来自于对两组样本  $m_{\gamma\gamma}$  谱的重新拟合, 得到  $\delta\mu_{CB} = \frac{\mu_{CB,var} - \mu_{CB,nom}}{\mu_{CB,nom}}$ ,  $\delta\sigma_{CB} = \frac{\sigma_{CB,var} - \sigma_{CB,nom}}{\sigma_{CB,nom}}$  作为其误差。在三个分析道中, 光子能量刻度对  $\mu_{CB}$  的影响在 0.04% 到 0.10% 之间, 光子能量分辨对  $\sigma_{CB}$  的影响为 0.004% ~ 0.02%。

**希格斯质量  $m_H$**  CEPC 预期可以将希格斯质量测量的误差控制在 5.9 MeV, 因此我们在信号质量峰上考虑此 5.9 MeV 的误差。

**BDTG 模型** 前述提到的探测器效应同样会影响到蒙卡样本 BDTG 的分布, 从而影响二维模型。我们考虑将各系统误差样本与原样本的 BDTG 分布进行对比, 计算每个 bin 中相对变化值  $\delta n^{bin} = \frac{|n_{var} - n_{nom}|}{n_{nom}}$ , 将其中最大的作为 BDTG 模型误差作用到信号事例数上, 其中去掉了低统计量 BDTG bin (bin 中事例数小于此

样本总事例数的 5%)。而 BDTG 方法本身造成的误差被认为包括在这一误差中。在 3 个分析道中这一项为 0.5% ~ 0.7%。

## 8.7 结果

此分析中我们构建类似于 5.8 中的似然函数，用极大似然法在三个分析道中联合拟合信号数，并表示为分析 POI  $\mu = \frac{N(e^+e^- \rightarrow ZH \rightarrow f\bar{f}\gamma\gamma)}{N_{SM}(e^+e^- \rightarrow ZH \rightarrow f\bar{f}\gamma\gamma)}$ 。似然函数定义如下：

$$\mathcal{L}(\mu, \theta; (m_{\gamma\gamma}, \text{BDT})) = \prod_c \text{Pois}(n_c | N_c(\mu, \theta)) \cdot \prod_i^n f_c((m_{\gamma\gamma}, \text{BDT})^i; \theta) \cdot \prod_j G(\theta_j), \quad (8-1)$$

其中：

- $\theta$  是系统误差对应的滋扰参数；
- $n_c$  是数据中  $c$  道中的事例数；
- $N_c(\mu, \theta) = \mu S_{SM,c}(\theta_{yield}) + B_c \cdot S_{SM,c}(\theta_{yield})$  是  $c$  道中的预期信号事例数，包括相应滋扰参数， $B_c$  是  $c$  道中本底事例数；

- $f_c((m_{\gamma\gamma}, \text{BDT})^i; \theta)$  是 ?? 中构造的信号和本底模型的概率密度函数：

$$f_c((m_{\gamma\gamma}, \text{BDT})^i; \theta) = \frac{1}{N_c} \times [\mu S_{SM,c}(\theta_{yield}) f_{c,sig}((m_{\gamma\gamma}, \text{BDT})^i; \theta) + B_c f_{c,bkg}((m_{\gamma\gamma}, \text{BDT})^i; \theta)]. \quad (8-2)$$

- 信号数  $S_{SM,c}$ 、信号峰  $\mu_{CB}$  和宽度  $\sigma_{CB}$  受系统误差影响，考虑其响应函数：

$$\begin{aligned} S_{SM,c}(\theta_{yield}) &= S_{SM,c} \prod_j e^{\theta_j \sqrt{\ln(1+\delta_j^2)}}, \\ \mu_{CB}(\theta_{peak}) &= \mu_{CB}^{nom} \prod_j (1 + \delta_j \theta_j), \\ \sigma_{CB}(\theta_{width}) &= \sigma_{CB}^{nom} \prod_j e^{\theta_j \sqrt{\ln(1+\delta_j^2)}} \end{aligned} \quad (8-3)$$

- $G(\theta_j)$  是对滋扰参数  $\theta_j$  的高斯约束项，中心值为 0，宽度为 1。

在拟合中信号模型参数固定道对信号蒙卡样本的最佳拟合值，本底事例数、模型参数和所有滋扰参数保持浮动。

目前我们并没有真实数据，为研究此分析的测量精度我们根据信号 + 本底模型产生了一组 Asimov 数据 [152]。图 8-6 展示了三个分析道中  $m_{\gamma\gamma}$  和 BDTG 分别的拟合情况。最终拟合得到的  $\mu$  的精度总结如表 8-9，其中  $20 \text{ fb}^{-1}$  结果为对

事例重新加权外推得到。同时我们对不同系统误差项的贡献进行了拆解，结果如表 8-10。本底模型误差的贡献为分别固定和浮动本底模型参数得到，已包括在表 8-9 中的统计误差中。

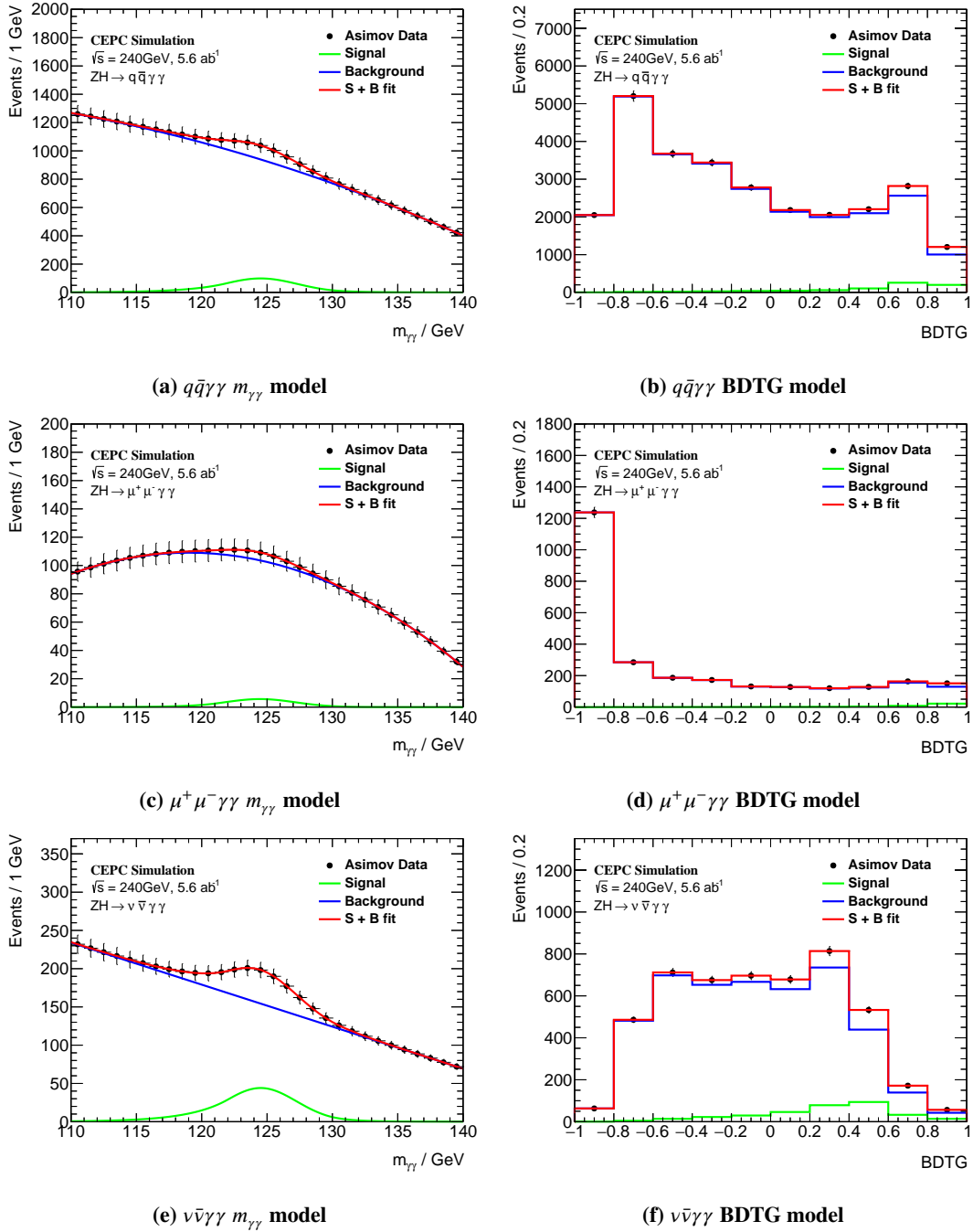


图 8-6 三个道中对 Asimov 数据的联合拟合。

	5.6 ab <sup>-1</sup>		20 ab <sup>-1</sup>	
	$\frac{\Delta_{tot}}{(\sigma \times Br)_{SM}}$	$\frac{\Delta_{stat}}{(\sigma \times Br)_{SM}}$	$\frac{\Delta_{tot}}{(\sigma \times Br)_{SM}}$	$\frac{\Delta_{stat}}{(\sigma \times Br)_{SM}}$
$q\bar{q}\gamma\gamma$	0.101	0.098	0.056	0.052
$\mu^+\mu^-\gamma\gamma$	0.373	0.371	0.202	0.200
$\nu\bar{\nu}\gamma\gamma$	0.130	0.127	0.071	0.067
Combined	0.079	0.077	0.046	0.040

表 8-9 对 Asimov 数据拟合得到的三个道中和联合后的  $\sigma(ZH) \times Br(H \rightarrow \gamma\gamma)$  预期精度。20 ab<sup>-1</sup> 结果为对信号和本底样本重新加权后在 workspace 中重新拟合得到。统计误差中已包括了本底模型的贡献。

	$q\bar{q}\gamma\gamma$	$\mu^+\mu^-\gamma\gamma$	$\nu\bar{\nu}\gamma\gamma$
理论误差 0.5%	0.005	-	-
亮度 0.1%	0.001	0.001	0.001
光子效率 1%	0.019	0.020	0.020
光子能量刻度 0.05%	0.001	<0.001	0.001
光子能量分辨 0.05%	<0.001	<0.001	<0.001
$m_H$ 5.9 MeV	<0.001	<0.001	<0.001
BDTG 模型	0.006	0.006	0.007
本底模型	0.029	0.062	0.006

表 8-10 三个道中各系统误差项对  $(\sigma \times Br)/(\sigma \times Br)_{SM}$  结果的贡献。0.5% 理论误差仅加在  $q\bar{q}\gamma\gamma$  道中。

## 8.8 电磁量能器能量分辨的影响

在构建模型、拟合  $m_{\gamma\gamma}$  时，信号峰宽度是  $H \rightarrow \gamma\gamma$  道中测量精度和 ECAL 能量分辨的直接联系。如 6.2.4 中所述，目前 CEPC 正在进行第四代探测器的概念设计，ECAL 能量分辨计划从基准探测器的  $16\%/\sqrt{E} \oplus 1\%$  提升到  $3\%/\sqrt{E} \oplus 1\%$ 。这将有助于光子的探测，以及中性介子  $\pi^0$  的重建，从而提升  $H \rightarrow \gamma\gamma$  分析和一些  $\pi^0 \rightarrow \gamma\gamma$  末态的味物理分析，如  $B_{(s)}^0 \rightarrow \pi^0\pi^0$  [153]。喷注的重建和能量分辨被认为没有显著变化，因为在 PFA 喷注重建中探测器颗粒度是主要因素。因此我们在  $q\bar{q}\gamma\gamma$  道中采用本分析中的方法，进行一定简化后研究了 ECAL 能量分辨对  $H \rightarrow \gamma\gamma$  测量的影响。

此研究中我们将分析中重建的光子替换为产生子中的真实光子，对其能量



进行一系列 smearing 以模拟不同的 ECAL 分辨。通常来讲 ECAL 能量分辨可以用下式表示：

$$\frac{\sigma_E}{E} = A \oplus \frac{B}{\sqrt{E}} \oplus \frac{C}{E}, \quad (8-1)$$

其中  $A$  为常数项，来自于探测器能量泄漏、电子学读出阈值等效应； $B$  为统计项，来自于探测器灵敏物质中产生的光电子的统计涨落； $C$  来自于电子学噪声。目前我们认为噪声项  $C = 0$ ， $A$  可以控制在 1% 量级。光子能量的 smearing 即基于此公式，调整  $B$  从 1% 变化至 35%。图 8-7 即展示了两个不同  $B$  值下 smearing 得到的  $m_{\gamma\gamma}$  分布，以及分析中探测器全模拟得到的  $m_{\gamma\gamma}$  分布。 $q\bar{q}\gamma\gamma$  道中喷注不做调整，认为探测器的替换不会影响到喷注重建。

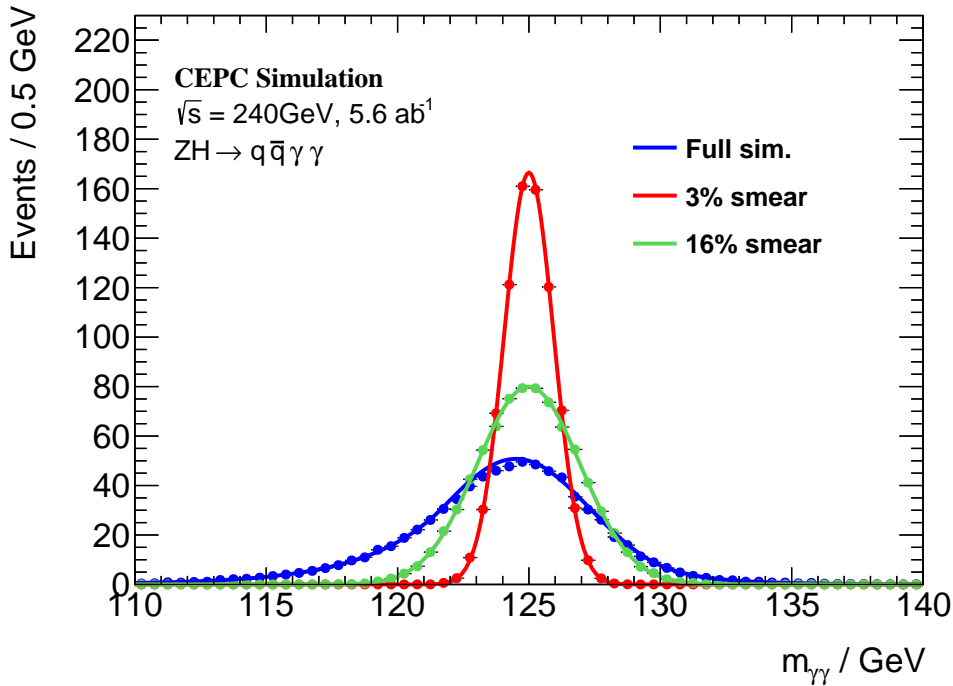


图 8-7 全模拟  $H \rightarrow \gamma\gamma$  样本（蓝线）和两组 smearing 样本（3% 红线，16% 绿线）中的  $m_{\gamma\gamma}$  分布。拟合得到的信号宽度分别为 2.81 GeV，0.94 GeV 和 1.96 GeV。

在此简化分析中，我们采用了 8.3 中的事例筛选条件，但没有加入 BDTG，以仅关注光子表现。这预计会导致分析精度有 30% 的降低，但考虑到  $m_{\gamma\gamma}$  和 BDTG 无相关性，不同 ECAL 能量分辨下分析精度的相对变化不会收到影响。统计模型的构建类似 8.7，构造似然函数，采用极大似然法拟合预期信号数，仅将信号的双边水晶球函数替换为高斯函数，将二维模型简化为  $m_{\gamma\gamma}$  的一维模型，且忽略系统误差（去除所有滋扰参数）。最终可以得到  $q\bar{q}\gamma\gamma$  道中预期分析精度  $\Delta\mu/\mu$ ，这一结果随  $B$  变化显示如图 8-8。

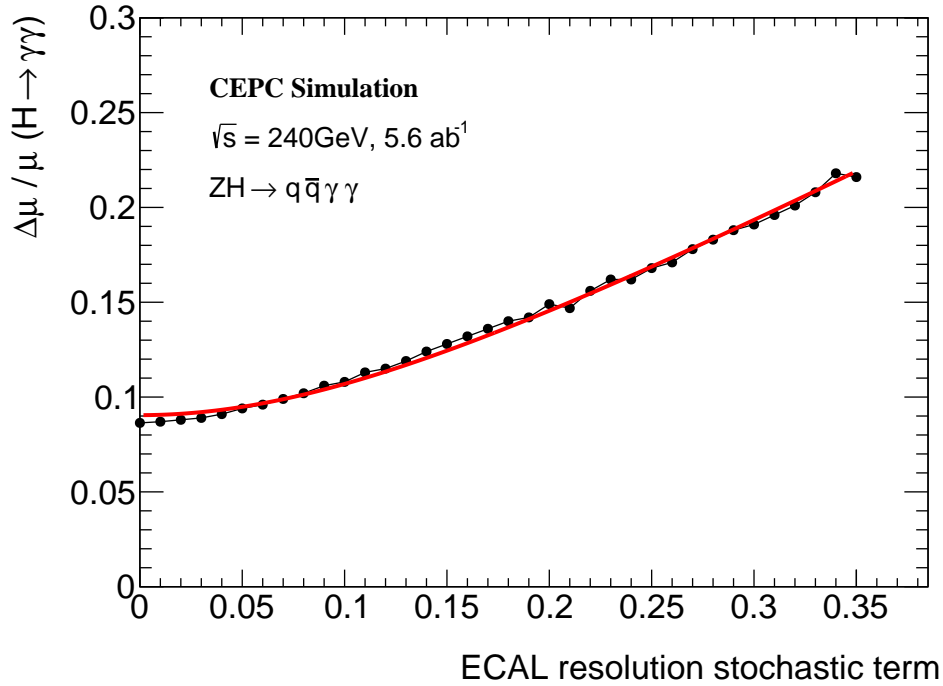


图 8-8  $ZH \rightarrow q\bar{q}\gamma\gamma$  道中信号强度测量精度随 ECAL 能量分辨统计项的变化。数据点用式 8-2 中的关系进行了拟合。

我们可以对这一关系进行一定分析。类似于光子能量分辨，我们采用以下函数来拟合数据点：

$$\frac{\delta\mu}{\mu} = p_0 \oplus (p_1 \times B), \quad (8-2)$$

其中  $p_0$  和  $p_1 \times B$  分别表示常数项  $A$  和统计项  $B$  在  $\Delta\mu/\mu$  中的贡献。根据此关系新的全吸收型晶体 ECAL 可以对分析精度带来 28% 的提高。另外我们可以定义一个“参考点”：统计项与常数项对  $\Delta\mu/\mu$  的贡献相等时的  $A$  和  $B$ ，即  $p_0 = p_1 \times B$ 。当常数项固定为 1% 时，这一定义下  $B$  的参考点为 14%，这意味着在  $3\%/\sqrt{E} \oplus 1\%$  的 ECAL 上常数项将贡献主要的能量分辨，在探测器设计时需考虑如何减小能量泄漏等效应，具有重要参考意义。不同  $A$  值对应的统计项  $B$  如图 8-9 所示。

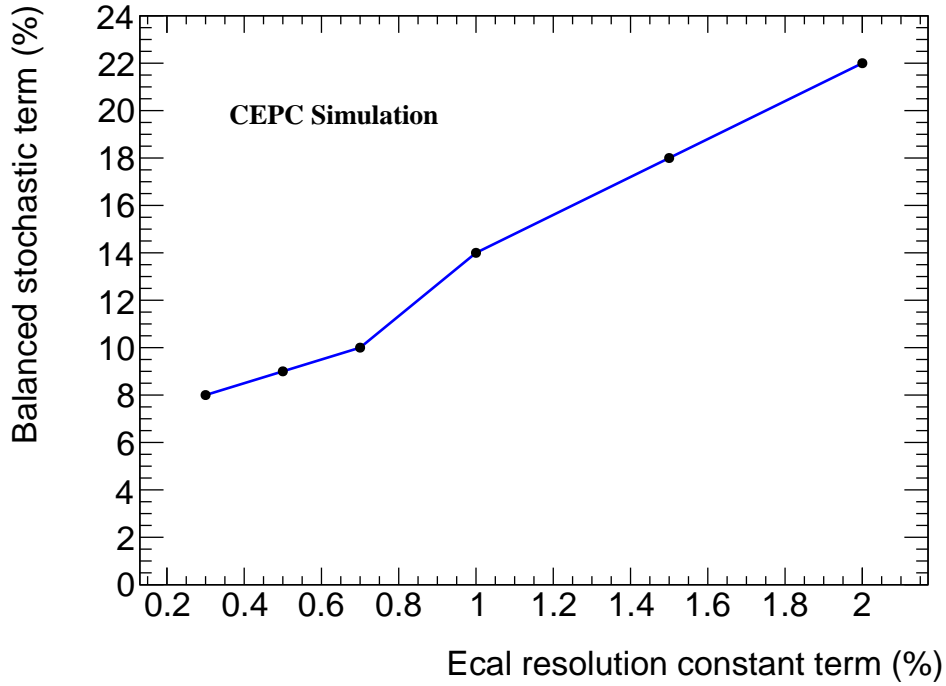


图 8-9 不同 ECAL 能量分辨常数项情况下对应的统计项参考点。

## 8.9 总结

$H \rightarrow \gamma\gamma$  道的分析对展现 CEPC 测量希格斯物理、寻找新物理的能力和前期探测器设计均具有重要意义。本分析研究了在 CEPC  $\sqrt{s} = 240\text{GeV}$  环境下  $ZH \rightarrow q\bar{q}\gamma\gamma$ 、 $ZH \rightarrow \mu^+\mu^-\gamma\gamma h$  和  $ZH \rightarrow \nu\bar{\nu}\gamma\gamma$  道中  $\sigma(e^+e^- \rightarrow ZH) \times Br(H \rightarrow \gamma\gamma)$  的预期测量精度，采用 CEPC 基准探测器进行事例模拟，利用 BDTG 区分信号和本底事例，并用其与  $m_{\gamma\gamma}$  一起构造统计模型，进行极大似然拟合。同时，本分析还搭建了滋扰参数法的系统误差分析框架，并初步估计了部分系统误差项。在  $5.6 \text{ ab}^{-1}$  统计量下，CEPC 中此分析的预期精度为 7.9%（统计误差 7.7%）；而在新的  $20 \text{ ab}^{-1}$  取数计划下这一精度可以达到 4.6%（统计精度 4.0%），意味着在此情况下系统误差将会有可见的贡献。本分析还在简化的  $q\bar{q}\gamma\gamma$  道中给出了分析精度与 ECAL 能量分辨之间的关系，对 ECAL 设计和优化具有一定参考意义。



## 第9章 总结与展望

标准模型是目前最成功的粒子物理理论模型，2012年发现的希格斯玻色子补完了标准模型的最后一块拼图。自此以后寻找超出标准模型的新物理是粒子物理领域最重要的问题之一，而希格斯玻色子可以成为这一研究的重要窗口。目前希格斯玻色子仅在LHC上可以通过p-p对撞大量产生，因此LHC上的ATLAS实验是现在研究希格斯粒子性质和寻找新物理的很好的机会。而在未来，CEPC计划在正负电子对撞中同样大量产生非常干净的希格斯粒子事例，以更高的精度进行相关测量。本文内容主要分为三部分，ATLAS探测器上电磁簇射相关研究，利用ATLAS探测器Run 2期间收集到的 $139 \text{ fb}^{-1}$  13 TeVp-p对撞数据在VBF  $H \rightarrow \gamma\gamma$ 道中测量HVV过程中的CP破缺效应，和对CEPC晶体电磁量能器重建算法的开发。

液氙电磁量能器中电磁簇射形状模拟问题是ATLAS上长期存在的一个问题，现阶段通过数据驱动的fudge factor进行修正，这会在物理分析中引入一项相关的系统误差。可能引起这一问题的原因包括多方面，目前已有研究显示Geant 4对电磁簇射发展过程的模拟对簇射形状有一定影响，本文则探究了电子学中的串扰效应和探测器几何描述对簇射形状的影响。ATLAS液氙量能器对串扰效应的模拟在数字化过程中实现，通过一个数据库载入探测器不同 $\eta$ 区域探测器单元中的串扰水平， $\phi$ 向串扰效应没有考虑，液氙量能器strip层和middle层由于结构不同也有不同的串扰水平。本文中使用的单光子事例进行模拟，调整数字化中的串扰幅度，研究其对簇射形状的影响。目前结果显示串扰效应可以显著影响簇射形状，在给予合适的串扰水平时可以将一些簇射形状变量修正到与数据中一致，但不同簇射形状变量需要的修正值并不一致，说明串扰效应可能是簇射形状问题的原因之一，但不是唯一原因，想要彻底解决这一问题还需要更多更细致的研究，如寻找机会重新测量ATLAS探测器电子学串扰信号的形状和幅度、考虑温度、压力、探测器老化等因素、在模拟中考虑远距离单元串扰等。而对探测器几何描述的调整没有显示出能够对簇射形状产生显著影响，这一因素可以被排除。注意到这一研究已是在 $O(0.1\%)$ 精度水平研究ATLAS探测器中的电磁簇射形状，是一个难得的研究粒子与探测器介质相互作用过程、探测器响应等效应的机会。这部分工作是本人在CERN联合培养期间的主要工作，相关结果计划以ATLAS公开文档的形式发表，目前正在内部审议中。

ATLAS 上光子的重建在 Run 2 期间升级到了动态的拓扑簇团方法，但光子鉴别仍基于 Run 1 重建时采用的簇射形状方法，合作组希望在 Run 3 期间能够对光子鉴别也进行升级，加入拓扑簇团相关变量，得到更好的性能表现。因此本文研究了拓扑簇团变量在光子鉴别工作中的潜力。我们选用  $Z \rightarrow ll\gamma$  和  $Z \rightarrow ll + \text{jet}$  过程分别作为真光子信号和假光子本底，选择了 7 个拓扑簇团相关变量，采用 BDT 训练方法研究仅采用簇射形状变量和簇射形状加拓扑簇团变量进行光子鉴别的性能。初步结果显示在加入拓扑簇团变量后，BDT 模型的 ROC 曲线有了显著改善，在  $Z \rightarrow e^+e^-\gamma$  和  $Z \rightarrow \mu^+\mu^-\gamma$  过程中有一致的表现。这一结果证实了拓扑簇团变量在光子鉴别工作中的能力，后续工作可以考虑实际开发基于拓扑簇团的光子鉴别方法，细致研究可选取的拓扑簇团变量，优化分类方法，根据光子能量、 $\eta$  区间、是否为转化光子设定不同工作点以得到最佳性能。另外也需要考虑在 Run 3 及后续高亮度 (HL-LHC) 运行期间事例堆积效应相对于 Run 2 有一定提升，可能对光子重建和鉴别产生一定影响。这部分工作是本人在 ATLAS 组内的 Qualification task，相关结果记录在 [154] 中。

希格斯粒子 CP 性质测量是在发现希格斯粒子后的重要后续研究之一，同时也是在希格斯机制中寻找超出标准模型新物理的重要机会。本文在 VBF  $H \rightarrow \gamma\gamma$  过程中基于 ATLAS Run 2 期间  $139 \text{ fb}^{-1}$  数据分析了 HVV 相互作用中的 CP 破缺性质。我们参考 ATLAS VBF  $H \rightarrow \tau^+\tau^-$  分析引入了优化观测量  $\mathcal{O}\mathcal{O}$  作为 CP 敏感观测量进行分析，在 EFT 的两个基下分别给出了对 CP 破缺参数  $\tilde{d}$  和  $c_{H\tilde{W}}$  的 95% 置信度约束，并与  $H \rightarrow \tau^+\tau^-$  的  $\tilde{d}$  测量进行了联合，当前结果没有显示出 HVV 作用中存在 CP 破缺效应。这是在 HISZ 基下首次给出  $\tilde{d}$  的 95% 置信度限制，而对 Warsaw 基下的参数  $c_{H\tilde{W}}$  的限制相比于之前 ATLAS 和 CMS 的分析有两倍的提升，相关结果发表在 [26]，本人是本分析的主要研究人员，负责引入优化观测量方法、确定事例分类策略、信号建模、计算本底组分分数、用鹰信号方法估计本底模型误差，并在 ATLAS Higgs 组内给了揭盲审议报告和两次结果审议报告。分析中采取的双 BDT 事例分类方法可以进行进一步优化，考虑采用如 DNN 的多分类器，直接区分多种过程的事例。当前结果仍是统计误差主导，因此相信在 ATLAS 于 2022 年 7 月开始 Run 3 取数后本分析的精度能够有进一步的提升。而优化观测量的方法是一个模型参数测量的通用方法，相信可以引入到其他分析中，如 CEPC 实验上对希格斯粒子的精确测量等。

CEPC 作为未来对撞机项目之一，有相当重要的科学潜力。CEPC 探测器相关设计和优化则是前期预研阶段的工作重点。为达到电弱、希格斯和顶夸克相关

物理的精确测量的要求，CEPC 探测器在设计上选用了粒子流设计思路，而为平衡性能与造价其电磁量能器设计上提出了由晶体条构成的全吸收型量能器方案。本文基于这一设计方案，开发了一套与其适配的重建算法，利用晶体条双端读出的电荷和时间信号将晶体条结构重建为高颗粒度结构，并解决因结构引起的 ghost hit 问题，证实了这一设计的可行性。全探测器重建工作则仍在进行中，这部分需要加入其他子探测器部分如径迹室和强子量能器，对粒子在探测器中发生的簇射进行更全面研究和识别，加入粒子鉴别方法，最终完成整个物理事例的粒子流重建。在粒子流重建方面，本文将电磁簇射形状信息引入重建算法中，希望借此扩展粒子流重建对探测器硬件结构的要求，后续希望将相关结果在高能物理实验领域进行推广。这一晶体电磁量能器和之前 CEPC 基准探测器方案对物理分析的影响则通过 CEPC 上  $H \rightarrow \gamma\gamma$  相关分析进行了研究，将事例按 Z 玻色子衰变模式分为 3 个子道，采用 BDTG 模型和  $m_{\gamma\gamma}$  二维拟合区分信号和本底，在考虑一定系统误差后给出了  $\sigma(e^+e^- \rightarrow ZH) \times Br(H \rightarrow \gamma\gamma)$  的预期测量精度。同时通过对光子能量进行 smear 的方式研究了电磁量能器能量分辨对这一测量精度的影响，联系起了探测器性能与物理指标。目前软件部分工作仍在进行中，本人完成了探测器几何模型的搭建、数字化算法设计、重建算法的初步可行性验证和软件框架的设计和编写。后续计划在重建中以子算法模块方式加入更多功能，如利用霍夫变换进行光子簇射的识别、电磁簇射和强子簇射的鉴别、接入 ArborPFA 完成全探测器重建等。CEPC 物理性能的分析则已发表在 [27]，本人为第一作者，也是分析的主要研究人员。在 CEPC 第四代探测器概念设计完成后，可以考虑在新探测器中重复这一研究。另外，如今工业界在计算机视觉领域关于点云的相关研究与粒子流重建的思路具有共通性，有希望将其引入开发出新的粒子流重建方法，得到对各种探测器通用的粒子流重建模型。





## 参考文献

- [1] Ruiz-Martinez, Aranzazu and ATLAS Collaboration. The Run-2 ATLAS Trigger System [R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2133909>. DOI: 10.1088/1742-6596/762/1/012003.
- [2] Panduro Vazquez, William and ATLAS Collaboration. The ATLAS Data Acquisition system in LHC Run 2: 3 [R/OL]. Geneva: CERN, 2017. <https://cds.cern.ch/record/2244345>. DOI: 10.1088/1742-6596/898/3/032017.
- [3] Collaboration A. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data [J/OL]. JINST, 2019, 14(12): P12006. DOI: 10.1088/1748-0221/14/12/P12006.
- [4] Collaboration A. Muon reconstruction and identification efficiency in ATLAS using the full Run 2  $pp$  collision data set at  $\sqrt{s} = 13$  TeV [J/OL]. Eur. Phys. J. C, 2021, 81(7): 578. DOI: 10.1140/epjc/s10052-021-09233-2.
- [5] ATLAS Collaboration. Atlas jet and missing energy reconstruction, calibration and performance in lhc run-2 [J/OL]. Journal of Instrumentation, 2017, 12(06): C06038. <https://dx.doi.org/10.1088/1748-0221/12/06/C06038>.
- [6] Workman R L, Others. Review of Particle Physics [J/OL]. PTEP, 2022, 2022: 083C01. DOI: 10.1093/ptep/ptac097.
- [7] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC [J/OL]. Phys. Lett. B, 2012, 716: 1. DOI: 10.1016/j.physletb.2012.08.020.
- [8] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC [J/OL]. Phys. Lett. B, 2012, 716: 30. DOI: 10.1016/j.physletb.2012.08.021.
- [9] ATLAS Collaboration. Measurement of the properties of Higgs boson production at  $\sqrt{s} = 13$  TeV in the  $H \rightarrow \gamma\gamma$  channel using  $139 \text{ fb}^{-1}$  of  $pp$  collision data with the ATLAS experiment [J]. 2022.
- [10] Dong M, et al. CEPC Conceptual Design Report: Volume 2 - Physics & Detector [J]. 2018.
- [11] The ATLAS Collaboration. Public atlas luminosity results for run-2 of the lhc [EB/OL]. 2022. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- [12] ATLAS Collaboration. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at  $\sqrt{s} = 13$  tev [J/OL]. The European Physical Journal C, 2019, 79(8). <https://doi.org/10.1140/epjc/s10052-019-7140-6>. DOI: 10.1140/epjc/s10052-019-7140-6.
- [13] ATLAS Collaboration. Jet energy scale and resolution measured in proton-proton collisions

- at  $\sqrt{s}=13$  TeV with the ATLAS detector [J/OL]. The European Physical Journal C, 2021, 81(8): 689. <https://doi.org/10.1140/epjc/s10052-021-09402-3>.
- [14] Kivernyk O. Measurement of the W boson mass with the ATLAS detector [EB/OL]. 2017. <https://cds.cern.ch/record/2262454>.
- [15] The International Linear Collider Technical Design Report - Volume 1: Executive Summary [J]. 2013.
- [16] Robson A, Burrows P N, Catalan Lasheras N, et al. The Compact Linear  $e^+e^-$  Collider (CLIC): Accelerator and Detector [J]. 2018.
- [17] CEPC Conceptual Design Report: Volume 1 - Accelerator [J]. 2018.
- [18] Fcc-ee: The lepton collider [J/OL]. The European Physical Journal Special Topics, 2019, 228(2): 261-623. <https://doi.org/10.1140/epjst/e2019-900045-4>.
- [19] Qi B, Liu Y. R & d of a novel high granularity crystal electromagnetic calorimeter [J/OL]. Instruments, 2022, 6(3). <https://www.mdpi.com/2410-390X/6/3/40>. DOI: 10.3390/instruments6030040.
- [20] Liu Y, Jiang J, Wang Y. High-granularity crystal calorimetry: conceptual designs and first studies [J/OL]. Journal of Instrumentation, 2020, 15(04): C04056. <https://dx.doi.org/10.1088/1748-0221/15/04/C04056>.
- [21] Marshall J S, Thomson M A. The pandora software development kit for particle flow calorimetry [J/OL]. Journal of Physics: Conference Series, 2012, 396(2): 022034. <https://dx.doi.org/10.1088/1742-6596/396/2/022034>.
- [22] ATLAS Collaboration. Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC [M/OL]. 2019. <https://cds.cern.ch/record/2677054>.
- [23] Avoni G, Bruschi M, Cabras G, et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS [J/OL]. Journal of Instrumentation, 2018, 13(07): P07017-P07017. <https://doi.org/10.1088/1748-0221/13/07/p07017>.
- [24] ATLAS Collaboration. Test of CP invariance in vector-boson fusion production of the Higgs boson in the  $h \rightarrow \tau\tau$  channel in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector [J/OL]. Physics Letters B, 2020, 805: 135426. <https://www.sciencedirect.com/science/article/pii/S0370269320302306>. DOI: <https://doi.org/10.1016/j.physletb.2020.135426>.
- [25] Cheng H, et al. The Physics potential of the CEPC. Prepared for the US Snowmass Community Planning Exercise (Snowmass 2021) [C]//2022 Snowmass Summer Study. 2022.
- [26] Study of the CP property of the Higgs boson to electroweak boson coupling in the VBF  $H \rightarrow \gamma\gamma$  channel with the ATLAS detector [J]. 2022.
- [27] Guo F, Fang Y, Li G, et al. The expected measurement precision of the branching ratio of the Higgs decaying to the di-photon at the CEPC [J/OL]. 2022. DOI: 10.1088/1674-1137/aca22.
- [28] Yang C N, Mills R L. Conservation of Isotopic Spin and Isotopic Gauge Invariance [J/OL]. Phys. Rev., 1954, 96: 191-195. DOI: 10.1103/PhysRev.96.191.

- [29] Hasert F J, et al. Search for Elastic  $\nu_{\mu}$  Electron Scattering [J/OL]. Phys. Lett. B, 1973, 46: 121-124. DOI: [10.1016/0370-2693\(73\)90494-2](https://doi.org/10.1016/0370-2693(73)90494-2).
- [30] Hasert F J, et al. Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment [J/OL]. Phys. Lett. B, 1973, 46: 138-140. DOI: [10.1016/0370-2693\(73\)90499-1](https://doi.org/10.1016/0370-2693(73)90499-1).
- [31] The UA1 Collaboration. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540$  GeV [J/OL]. Phys. Lett. B, 1983, 122: 103-116. DOI: [10.1016/0370-2693\(83\)91177-2](https://doi.org/10.1016/0370-2693(83)91177-2).
- [32] The UA1 Collaboration. Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c\*\*2 at the CERN SPS Collider [J/OL]. Phys. Lett. B, 1983, 126: 398-410. DOI: [10.1016/0370-2693\(83\)90188-0](https://doi.org/10.1016/0370-2693(83)90188-0).
- [33] The UA2 Collaboration. Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider [J/OL]. Phys. Lett. B, 1983, 122: 476-485. DOI: [10.1016/0370-2693\(83\)91605-2](https://doi.org/10.1016/0370-2693(83)91605-2).
- [34] The UA2 Collaboration. Evidence for  $z^0 \rightarrow e^+e^-$  at the cern pp collider [J/OL]. Physics Letters B, 1983, 129(1): 130-140. <https://www.sciencedirect.com/science/article/pii/037026938390744X>. DOI: [https://doi.org/10.1016/0370-2693\(83\)90744-X](https://doi.org/10.1016/0370-2693(83)90744-X).
- [35] Englert F, Brout R. Broken Symmetry and the Mass of Gauge Vector Mesons [J/OL]. Phys. Rev. Lett., 1964, 13: 321-323. DOI: [10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321).
- [36] Higgs P W. Broken symmetries, massless particles and gauge fields [J/OL]. Phys. Lett., 1964, 12: 132-133. DOI: [10.1016/0031-9163\(64\)91136-9](https://doi.org/10.1016/0031-9163(64)91136-9).
- [37] Higgs P W. Broken Symmetries and the Masses of Gauge Bosons [J/OL]. Phys. Rev. Lett., 1964, 13: 508-509. DOI: [10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508).
- [38] Goldstone J, Salam A, Weinberg S. Broken Symmetries [J/OL]. Phys. Rev., 1962, 127: 965-970. DOI: [10.1103/PhysRev.127.965](https://doi.org/10.1103/PhysRev.127.965).
- [39] CDF COLLABORATION. High-precision measurement of the  $\langle w \rangle$  boson mass with the cdf ii detector [J/OL]. Science, 2022, 376(6589): 170-176. <https://www.science.org/doi/abs/10.1126/science.abk1781>.
- [40] Cabibbo N. Unitary Symmetry and Leptonic Decays [J/OL]. Phys. Rev. Lett., 1963, 10: 531-533. DOI: [10.1103/PhysRevLett.10.531](https://doi.org/10.1103/PhysRevLett.10.531).
- [41] Kobayashi M, Maskawa T. CP Violation in the Renormalizable Theory of Weak Interaction [J/OL]. Prog. Theor. Phys., 1973, 49: 652-657. DOI: [10.1143/PTP.49.652](https://doi.org/10.1143/PTP.49.652).
- [42] Christenson J H, Cronin J W, Fitch V L, et al. Evidence for the  $2\pi$  Decay of the  $K_2^0$  Meson [J/OL]. Phys. Rev. Lett., 1964, 13: 138-140. DOI: [10.1103/PhysRevLett.13.138](https://doi.org/10.1103/PhysRevLett.13.138).
- [43] Fukuda Y, et al. Evidence for oscillation of atmospheric neutrinos [J/OL]. Phys. Rev. Lett., 1998, 81: 1562-1567. DOI: [10.1103/PhysRevLett.81.1562](https://doi.org/10.1103/PhysRevLett.81.1562).
- [44] Ahmad Q R, et al. Measurement of the rate of  $\nu_e + d \rightarrow p + p + e^-$  interactions produced by

- $^8\text{B}$  solar neutrinos at the Sudbury Neutrino Observatory [J/OL]. Phys. Rev. Lett., 2001, 87: 071301. DOI: [10.1103/PhysRevLett.87.071301](https://doi.org/10.1103/PhysRevLett.87.071301).
- [45] Aaij R, et al. Test of lepton universality in beauty-quark decays [J/OL]. Nature Phys., 2022, 18(3): 277-282. DOI: [10.1038/s41567-021-01478-8](https://doi.org/10.1038/s41567-021-01478-8).
- [46] Grzadkowski B, Iskrzynski M, Misiak M, et al. Dimension-Six Terms in the Standard Model Lagrangian [J/OL]. JHEP, 2010, 10: 085. DOI: [10.1007/JHEP10\(2010\)085](https://doi.org/10.1007/JHEP10(2010)085).
- [47] Feruglio F. The Chiral approach to the electroweak interactions [J/OL]. Int. J. Mod. Phys. A, 1993, 8: 4937-4972. DOI: [10.1142/S0217751X93001946](https://doi.org/10.1142/S0217751X93001946).
- [48] Alonso R, Gavela M B, Merlo L, et al. The Effective Chiral Lagrangian for a Light Dynamical "Higgs Particle" [J/OL]. Phys. Lett. B, 2013, 722: 330-335. DOI: [10.1016/j.physletb.2013.04.037](https://doi.org/10.1016/j.physletb.2013.04.037).
- [49] Measurements of Higgs boson properties in the diphoton decay channel using  $80\text{ fb}^{-1}$  of  $pp$  collision data at  $\sqrt{s} = 13\text{ TeV}$  with the ATLAS detector [R/OL]. Geneva: CERN, 2018. <https://cds.cern.ch/record/2628771>.
- [50] ATLAS Collaboration. Observation of higgs boson production in association with a top quark pair at the lhc with the atlas detector [J/OL]. Physics Letters B, 2018, 784: 173-191. <https://www.sciencedirect.com/science/article/pii/S0370269318305732>. DOI: <https://doi.org/10.1016/j.physletb.2018.07.035>.
- [51] ATLAS Collaboration. Observation of  $h \rightarrow b\bar{b}$  decays and  $vh$  production with the atlas detector [J/OL]. Physics Letters B, 2018, 786: 59-86. <https://www.sciencedirect.com/science/article/pii/S0370269318307056>. DOI: <https://doi.org/10.1016/j.physletb.2018.09.013>.
- [52] ATLAS Collaboration. Cross-section measurements of the higgs boson decaying into a pair of  $\tau$ -leptons in proton-proton collisions at  $\sqrt{s} = 13\text{ TeV}$  with the atlas detector [J/OL]. Phys. Rev. D, 2019, 99: 072001. <https://link.aps.org/doi/10.1103/PhysRevD.99.072001>.
- [53] ATLAS Collaboration. Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector [J/OL]. Eur. Phys. J. C, 2022, 82: 717. DOI: [10.1140/epjc/s10052-022-10588-3](https://doi.org/10.1140/epjc/s10052-022-10588-3).
- [54] ATLAS Collaboration. A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector [J/OL]. Phys. Lett. B, 2021, 812: 135980. DOI: [10.1016/j.physletb.2020.135980](https://doi.org/10.1016/j.physletb.2020.135980).
- [55] ATLAS Collaboration. A search for the  $z\gamma$  decay mode of the higgs boson in  $pp$  collisions at  $s=13\text{ tev}$  with the atlas detector [J/OL]. Physics Letters B, 2020, 809: 135754. <https://www.sciencedirect.com/science/article/pii/S0370269320305578>. DOI: <https://doi.org/10.1016/j.physletb.2020.135754>.
- [56] ATLAS Collaboration. Measurement of the total and differential Higgs boson production cross-sections at  $\sqrt{s} = 13\text{ TeV}$  with the ATLAS detector by combining the  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $H \rightarrow \gamma\gamma$  decay channels [J]. 2022.

- [57] Achard P, Adriani O, Aguilar-Benitez M, et al. Search for anomalous couplings in the higgs sector at lep [J/OL]. *Physics Letters B*, 2004, 589(3-4): 89–102. <http://dx.doi.org/10.1016/j.physletb.2004.03.048>.
- [58] De Melis C. The CERN accelerator complex. *Complexe des accélérateurs du CERN* [J/OL]. 2016. <https://cds.cern.ch/record/2119882>.
- [59] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider [J/OL]. *JINST*, 2008, 3: S08003. DOI: [10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- [60] CMS Collaboration. The CMS Experiment at the CERN LHC [J/OL]. *JINST*, 2008, 3: S08004. DOI: [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- [61] LHCb Collaboration. The LHCb Detector at the LHC [J/OL]. *JINST*, 2008, 3: S08005. DOI: [10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005).
- [62] ALICE Collaboration. The ALICE experiment at the CERN LHC [J/OL]. *JINST*, 2008, 3: S08002. DOI: [10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002).
- [63] Capeans M, Darbo G, Einsweiler K, et al. ATLAS Insertable B-Layer Technical Design Report [R/OL]. 2010. <https://cds.cern.ch/record/1291633>.
- [64] Technical Design Report for the ATLAS Inner Tracker Pixel Detector [R/OL]. Geneva: CERN, 2017. <https://cds.cern.ch/record/2285585>. DOI: [10.17181/CERN.FOZZ.ZP3Q](https://doi.org/10.17181/CERN.FOZZ.ZP3Q).
- [65] Aharrouche M, Colas J, Di Ciaccio L, et al. Energy linearity and resolution of the atlas electromagnetic barrel calorimeter in an electron test-beam [J/OL]. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2006, 568(2): 601-623. <https://www.sciencedirect.com/science/article/pii/S0168900206013222>. DOI: <https://doi.org/10.1016/j.nima.2006.07.053>.
- [66] Aubert B, Ballansat J, Bazan A, et al. Performance of the atlas electromagnetic calorimeter end-cap module 0 [J/OL]. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2003, 500(1): 178-201. <https://www.sciencedirect.com/science/article/pii/S0168900203003449>. DOI: [https://doi.org/10.1016/S0168-9002\(03\)00344-9](https://doi.org/10.1016/S0168-9002(03)00344-9).
- [67] Hurwitz M. Module-to-module uniformity at 180 GeV in 2002-2003 TileCal calibration test-beams [J]. 2006.
- [68] Dowler B, Pinfold J, Soukup J, et al. Performance of the atlas hadronic end-cap calorimeter in beam tests [J/OL]. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2002, 482(1): 94-124. <https://www.sciencedirect.com/science/article/pii/S0168900201013389>. DOI: [https://doi.org/10.1016/S0168-9002\(01\)01338-9](https://doi.org/10.1016/S0168-9002(01)01338-9).
- [69] Archambault J P, Artamonov A, Cadabeschi M, et al. Energy calibration of the atlas liquid argon forward calorimeter [J/OL]. *Journal of Instrumentation*, 2008, 3(02): P02002. <https://dx.doi.org/10.1088/1748-0221/3/02/P02002>.

- [70] Cornelissen T G, Elsing M, Gavrilenko I, et al. The global  $\chi^2$  track fitter in atlas [J/OL]. Journal of Physics: Conference Series, 2008, 119(3): 032013. <https://dx.doi.org/10.1088/1742-6596/119/3/032013>.
- [71] Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung [R/OL]. Geneva: CERN, 2012. <https://cds.cern.ch/record/1449796>.
- [72] Frühwirth R. Application of kalman filtering to track and vertex fitting [J/OL]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1987, 262(2): 444-450. <https://www.sciencedirect.com/science/article/pii/0168900287908874>. DOI: [https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4).
- [73] Hoecker A, Speckmayer P, Stelzer J, et al. Tmva - toolkit for multivariate data analysis [M/OL]. arXiv, 2007. <https://arxiv.org/abs/physics/0703039>. DOI: 10.48550/ARXIV.PHYSICS/0703039.
- [74] Cacciari M, Salam G P, Soyez G. The anti-kt jet clustering algorithm [J/OL]. Journal of High Energy Physics, 2008, 2008(04): 063. <https://dx.doi.org/10.1088/1126-6708/2008/04/063>.
- [75] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the atlas detector [J/OL]. The European Physical Journal C, 2017, 77(7): 466. <https://doi.org/10.1140/epjc/s10052-017-5031-2>.
- [76] ATLAS Collaboration. Study of the spin and parity of the higgs boson in diboson decays with the atlas detector [J/OL]. The European Physical Journal C, 2015, 75(10): 476. <https://doi.org/10.1140/epjc/s10052-015-3685-1>.
- [77] CMS Collaboration. Constraints on the spin-parity and anomalous  $h\nu\nu$  couplings of the higgs boson in proton collisions at 7 and 8 tev [J/OL]. Phys. Rev. D, 2015, 92: 012004. <https://link.aps.org/doi/10.1103/PhysRevD.92.012004>.
- [78] ATLAS Collaboration.  $cp$  properties of higgs boson interactions with top quarks in the  $t\bar{t}h$  and  $th$  processes using  $h \rightarrow \gamma\gamma$  with the atlas detector [J/OL]. Phys. Rev. Lett., 2020, 125: 061802. <https://link.aps.org/doi/10.1103/PhysRevLett.125.061802>.
- [79] CMS Collaboration. Measurements of  $t\bar{t}h$  production and the  $cp$  structure of the yukawa interaction between the higgs boson and top quark in the diphoton decay channel [J/OL]. Phys. Rev. Lett., 2020, 125: 061801. <https://link.aps.org/doi/10.1103/PhysRevLett.125.061801>.
- [80] CMS Collaboration. Constraints on anomalous  $h\nu\nu$  couplings from the production of higgs bosons decaying to  $\tau$  lepton pairs [J/OL]. Phys. Rev. D, 2019, 100: 112002. <https://link.aps.org/doi/10.1103/PhysRevD.100.112002>.
- [81] ATLAS Collaboration. Measurements of the Higgs boson inclusive and differential fiducial cross-sections in the diphoton decay channel with pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector [J/OL]. JHEP, 2022, 08: 027. DOI: 10.1007/JHEP08(2022)027.
- [82] Hankele V, Klämke G, Zeppenfeld D, et al. Anomalous higgs boson couplings in vector boson

- fusion at the cern lhc [J/OL]. Phys. Rev. D, 2006, 74: 095001. <https://link.aps.org/doi/10.1103/PhysRevD.74.095001>.
- [83] ALEPH Collaboration. Improved measurement of the triple gauge-boson couplings  $\gamma ww$  and  $zww$  in  $e^+e^-$  collisions [J/OL]. Physics Letters B, 2005, 614(1): 7-26. <https://www.sciencedirect.com/science/article/pii/S0370269305004120>. DOI: <https://doi.org/10.1016/j.physletb.2005.03.058>.
- [84] Abbiendi et al. G. Measurement of w boson polarisations and cp-violating triple gauge couplings from  $w^+ w^-$  production at lep [J/OL]. The European Physical Journal C, 2001, 19(2): 229–240. <http://dx.doi.org/10.1007/s100520100602>.
- [85] Hagiwara K, Ishihara S, Szalapski R, et al. Low energy effects of new interactions in the electroweak boson sector [J/OL]. Phys. Rev. D, 1993, 48: 2182-2203. <https://link.aps.org/doi/10.1103/PhysRevD.48.2182>.
- [86] Falkowski A. Higgs Basis: Proposal for an EFT basis choice for LHC HXSWG [J/OL]. 2015. <https://cds.cern.ch/record/2001958>.
- [87] Aad G, Abbott B, Abdallah J, et al. Study of the spin and parity of the higgs boson in diboson decays with the atlas detector [J/OL]. The European Physical Journal C, 2015, 75(10). <http://dx.doi.org/10.1140/epjc/s10052-015-3685-1>.
- [88] Davier M, Duflot L, Le Diberder F, et al. The optimal method for the measurement of tau polarization [J/OL]. Physics Letters B, 1993, 306(3): 411-417. <https://www.sciencedirect.com/science/article/pii/037026939390101M>. DOI: [https://doi.org/10.1016/0370-2693\(93\)90101-M](https://doi.org/10.1016/0370-2693(93)90101-M).
- [89] Diehl M, Nachtmann O. Optimal observables for the measurement of three gauge boson couplings in  $e^+e^- \rightarrow w^+w^-$  [J/OL]. Zeitschrift für Physik C Particles and Fields, 1994, 62 (3): 397-411. <https://doi.org/10.1007/BF01555899>.
- [90] Denner A, Dittmaier S, Kallweit S, et al. Hawk 2.0: A monte carlo program for higgs production in vector-boson fusion and higgs strahlung at hadron colliders [J/OL]. Computer Physics Communications, 2015, 195: 161–171. <http://dx.doi.org/10.1016/j.cpc.2015.04.021>.
- [91] Brivio I, Trott M. The Standard Model as an Effective Field Theory [J/OL]. Phys. Rept., 2019, 793: 1-98. DOI: [10.1016/j.physrep.2018.11.002](https://doi.org/10.1016/j.physrep.2018.11.002).
- [92] ATLAS Collaboration. ATLAS data quality operations and performance for 2015–2018 data-taking [J]. 2019.
- [93] Alwall J, Frederix R, Frixione S, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations [J/OL]. JHEP, 2014, 07: 079. DOI: [10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079).
- [94] P. Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms [J]. JHEP, 2004, 11: 040.
- [95] Alioli S, Nason P, Oleari C, et al. NLO Higgs boson production via gluon fusion matched with shower in POWHEG [J/OL]. JHEP, 2009, 0904: 002. DOI: [10.1088/1126-6708/2009/04/002](https://doi.org/10.1088/1126-6708/2009/04/002).



- [96] Alioli S, Nason P, Oleari C, et al. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX [J/OL]. JHEP, 2010, 06: 043. DOI: [10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043).
- [97] Frixione S, Nason P, Oleari C. Matching NLO QCD computations with parton shower simulations: the POWHEG method [J/OL]. JHEP, 2007, 11: 070. DOI: [10.1088/1126-6708/2007/11/070](https://doi.org/10.1088/1126-6708/2007/11/070).
- [98] Sjöstrand T, Mrenna S, Skands P Z. A brief introduction to PYTHIA 8.1 [J/OL]. Comput. Phys. Commun., 2008, 178: 852. DOI: [10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036).
- [99] Agostinelli S, et al. GEANT4 – a simulation toolkit [J/OL]. Nucl. Instrum. Meth. A, 2003, 506: 250. DOI: [10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [100] Djouadi A, Kalinowski J, Spira M. HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension [J/OL]. Comput. Phys. Commun., 1998, 108: 56. DOI: [10.1016/S0010-4655\(97\)00123-9](https://doi.org/10.1016/S0010-4655(97)00123-9).
- [101] Spira M. QCD Effects in Higgs physics [J/OL]. Fortsch. Phys., 1998, 46: 203-284. DOI: [10.1002/\(SICI\)1521-3978\(199804\)46:3<203::AID-PROP203>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1521-3978(199804)46:3<203::AID-PROP203>3.0.CO;2-4).
- [102] Djouadi A, Mühlleitner M M, Spira M. Decays of supersymmetric particles: The Program SUSY-HIT (SUSpect-SdecaY-Hdecay-InTerface) [J]. Acta Phys. Polon. B, 2007, 38: 635-644.
- [103] Bredenstein A, Denner A, Dittmaier S, et al. Radiative corrections to the semileptonic and hadronic Higgs-boson decays  $H \rightarrow WW/ZZ \rightarrow 4$  fermions [J/OL]. JHEP, 2007, 02: 080. DOI: [10.1088/1126-6708/2007/02/080](https://doi.org/10.1088/1126-6708/2007/02/080).
- [104] Bredenstein A, Denner A, Dittmaier S, et al. Precise predictions for the Higgs-boson decay  $H \rightarrow WW/ZZ \rightarrow 4$  leptons [J/OL]. Phys. Rev. D, 2006, 74: 013004. DOI: [10.1103/PhysRevD.74.013004](https://doi.org/10.1103/PhysRevD.74.013004).
- [105] Bredenstein A, Denner A, Dittmaier S, et al. Precision calculations for the Higgs-boson decay  $H \rightarrow ZZ/WW \rightarrow 4$  leptons [J/OL]. Nucl. Phys. Proc. Suppl., 2006, 160: 131-135. DOI: [10.1016/j.nuclphysbps.2006.09.104](https://doi.org/10.1016/j.nuclphysbps.2006.09.104).
- [106] Collaboration A. The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim [R/OL]. Geneva: CERN, 2010. <https://cds.cern.ch/record/1300517>.
- [107] ATLAS Collaboration. Measurement of higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 tev with the atlas detector [J/OL]. Physical Review D, 2014, 90(11). <http://dx.doi.org/10.1103/PhysRevD.90.112015>. DOI: [10.1103/physrevd.90.112015](https://doi.org/10.1103/physrevd.90.112015).
- [108] ATLAS Collaboration. Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector [J/OL]. Eur. Phys. J. C, 2017, 77: 580. DOI: [10.1140/epjc/s10052-017-5081-5](https://doi.org/10.1140/epjc/s10052-017-5081-5).
- [109] Aad G, Abbott B, Abbott D C, et al. Performance of electron and photon triggers in atlas



- during LHC Run 2 [J/OL]. The European Physical Journal C, 2020, 80(1). <http://dx.doi.org/10.1140/epjc/s10052-019-7500-2>.
- [110] Monticelli F, Orellana G E, Hoya J, et al. Performance of Photon triggers in ATLAS during LHC Run-2 [R/OL]. Geneva: CERN, 2018. <https://cds.cern.ch/record/2652154>.
- [111] Nomidis I, De Vivie De Regie J B, Proklova N, et al. Event selection, performance and background estimation in the  $H \rightarrow \gamma\gamma$  channel with Run-2 data [R/OL]. Geneva: CERN, 2020. <https://cds.cern.ch/record/2718255>.
- [112] Carminati L, Delmastro M, Kataoka M, et al. Measurement of the isolated di-photon cross section in 4.9 fb<sup>-1</sup> of pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector [R/OL]. Geneva: CERN, 2012. <https://cds.cern.ch/record/1450063>.
- [113] Frate M, Cranmer K, Kalia S, et al. Modeling Smooth Backgrounds and Generic Localized Signals with Gaussian Processes [J]. 2017.
- [114] ATLAS, Collaborations C. Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments [J/OL]. Phys. Rev. Lett., 2015, 114: 191803. DOI: [10.1103/PhysRevLett.114.191803](https://doi.org/10.1103/PhysRevLett.114.191803).
- [115] The International Linear Collider Technical Design Report - Volume 2: Physics [J]. 2013.
- [116] Abramowicz H, et al. The International Linear Collider Technical Design Report - Volume 4: Detectors [J]. 2013.
- [117] Bai M, et al. C<sup>3</sup>: A "Cool" Route to the Higgs Boson and Beyond [C]//2022 Snowmass Summer Study. 2021.
- [118] Dasu S, et al. Strategy for Understanding the Higgs Physics: The Cool Copper Collider [C]//2022 Snowmass Summer Study. 2022.
- [119] Delahaye J P, Diemoz M, Long K, et al. Muon Colliders [J]. 2019.
- [120] An F, et al. Precision Higgs physics at the CEPC [J/OL]. Chin. Phys. C, 2019, 43(4): 043002. DOI: [10.1088/1674-1137/43/4/043002](https://doi.org/10.1088/1674-1137/43/4/043002).
- [121] Cepeda M, et al. Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC [J/OL]. CERN Yellow Rep. Monogr., 2019, 7: 221-584. DOI: [10.23731/CYRM-2019-007.221](https://doi.org/10.23731/CYRM-2019-007.221).
- [122] CEPC. Cepec software home page [EB/OL]. 2018. <http://cepcsoft.ihep.ac.cn/>.
- [123] iLC. Linear collider software [EB/OL]. 2023. <https://github.com/iLCSoft>.
- [124] Gaede F, Behnke T, Graf N, et al. LCIO: A Persistency framework for linear collider simulation studies [J]. eConf, 2003, C0303241: TUKT001.
- [125] Frank M, Gaede F, Grefe C, et al. Dd4hep: A detector description toolkit for high energy physics experiments [J/OL]. Journal of Physics: Conference Series, 2014, 513(2): 022010. <http://dx.doi.org/10.1088/1742-6596/513/2/022010>.
- [126] Kilian W, Ohl T, Reuter J. WHIZARD: Simulating Multi-Particle Processes at LHC and ILC [J/OL]. Eur. Phys. J. C, 2011, 71: 1742. DOI: [10.1140/epjc/s10052-011-1742-y](https://doi.org/10.1140/epjc/s10052-011-1742-y).

- [127] Mora de Freitas P, Videau H. Detector simulation with MOKKA / GEANT4: Present and future [J]. 2002: 623-627.
- [128] Gaede F. Marlin and LCCD: Software tools for the ILC [J/OL]. Nucl. Instrum. Meth. A, 2006, 559: 177-180. DOI: [10.1016/j.nima.2005.11.138](https://doi.org/10.1016/j.nima.2005.11.138).
- [129] Ganis G, Helsen C, Völkl V. Key4hep, a framework for future HEP experiments and its use in FCC [J/OL]. Eur. Phys. J. Plus, 2022, 137(1): 149. DOI: [10.1140/epjp/s13360-021-02213-1](https://doi.org/10.1140/epjp/s13360-021-02213-1).
- [130] Gaede F, Ganis G, Hegner B, et al. EDM4hep and podio - The event data model of the Key4hep project and its implementation [J/OL]. EPJ Web Conf., 2021, 251: 03026. DOI: [10.1051/epjconf/202125103026](https://doi.org/10.1051/epjconf/202125103026).
- [131] Clemencic M, Hegner B, Leggett C. Gaudi evolution for future challenges [J/OL]. J. Phys. Conf. Ser., 2017, 898(4): 042044. DOI: [10.1088/1742-6596/898/4/042044](https://doi.org/10.1088/1742-6596/898/4/042044).
- [132] CEPCSW. Cepc offline software prototype based on key4hep [EB/OL]. 2023. <https://github.com/cepc/CEPCSW>.
- [133] Knowles I G, Lafferty G D. Hadronization in  $Z^0$  decay [J/OL]. J. Phys. G, 1997, 23: 731-789. DOI: [10.1088/0954-3899/23/7/003](https://doi.org/10.1088/0954-3899/23/7/003).
- [134] M. G. Green and others. Electron-positron physics at the z [M/OL]. Bristol, UK ; Philadelphia : Institute of Physics Pub., 1998. <https://search.library.wisc.edu/catalog/999834167302121>.
- [135] Thomson M. Particle flow calorimetry and the PandoraPFA algorithm [J/OL]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2009, 611(1): 25-40. DOI: [10.1016/j.nima.2009.09.009](https://doi.org/10.1016/j.nima.2009.09.009).
- [136] Catani S, Dokshitzer Y, Olsson M, et al. New clustering algorithm for multijet cross sections in  $e^+e^-$  annihilation [J/OL]. Physics Letters B, 1991, 269(3): 432-438. DOI: [https://doi.org/10.1016/0370-2693\(91\)90196-W](https://doi.org/10.1016/0370-2693(91)90196-W).
- [137] Ruchti R, Krüger K. Particle Flow Calorimetry [C]//2022 Snowmass Summer Study. 2022.
- [138] The CALICE collaboration. Design and electronics commissioning of the physics prototype of a si-w electromagnetic calorimeter for the international linear collider [J/OL]. Journal of Instrumentation, 2008, 3(08): P08001-P08001. <https://doi.org/10.1088/1748-0221/3/08/p08001>. DOI: [10.1088/1748-0221/3/08/p08001](https://doi.org/10.1088/1748-0221/3/08/p08001).
- [139] The CALICE collaboration. Construction and response of a highly granular scintillator-based electromagnetic calorimeter [J/OL]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2018, 887: 150-168. <https://www.sciencedirect.com/science/article/pii/S0168900218300305>. DOI: <https://doi.org/10.1016/j.nima.2018.01.016>.
- [140] Ruan M. Arbor, a new approach of the particle flow algorithm [M/OL]. arXiv, 2014. DOI: [10.48550/ARXIV.1403.4784](https://doi.org/10.48550/ARXIV.1403.4784).
- [141] Ruan M, et al. Reconstruction of physics objects at the Circular Electron Positron Collider with Arbor [J/OL]. Eur. Phys. J. C, 2018, 78(5): 426. DOI: [10.1140/epjc/s10052-018-5876-z](https://doi.org/10.1140/epjc/s10052-018-5876-z).

- [142] Grindhammer G, Peters S. The Parameterized simulation of electromagnetic showers in homogeneous and sampling calorimeters [C]//International Conference on Monte Carlo Simulation in High-Energy and Nuclear Physics - MC 93. 1993.
- [143] He M. Simulation and reconstruction of the BESIII EMC [J/OL]. Journal of Physics: Conference Series, 2011, 293: 012025. <https://doi.org/10.1088/1742-6596/293/1/012025>.
- [144] Kotera K, Jeans D, Miyamoto A, et al. A novel strip energy splitting algorithm for the fine granular readout of a scintillator strip electromagnetic calorimeter [J/OL]. Nucl. Instrum. Meth. A, 2015, 789: 158-164. DOI: [10.1016/j.nima.2015.04.001](https://doi.org/10.1016/j.nima.2015.04.001).
- [145] Wang Y. Crystal ecal design for cepe. online mini-workshop on a detector concept with a crystal ecal, 22–23 july 2020 [EB/OL]. 2022. [https://indico.ihep.ac.cn/event/11938/contributions/14786/attachments/6967/7875/Yuexin\\_CrystalECALWS\\_20200722.pdf](https://indico.ihep.ac.cn/event/11938/contributions/14786/attachments/6967/7875/Yuexin_CrystalECALWS_20200722.pdf).
- [146] Measurements of Higgs boson properties in the diphoton decay channel at  $\sqrt{s} = 13$  TeV [J]. 2020.
- [147] Taylor T, Treille D. The Large Electron Positron Collider (LEP): Probing the Standard Model [J/OL]. Adv. Ser. Direct. High Energy Phys., 2017, 27: 217-261. 45 p. <https://cds.cern.ch/record/2312570>. DOI: [10.1142/9789814749145\\_0007](https://doi.org/10.1142/9789814749145_0007).
- [148] Mo X, Li G, Ruan M Q, et al. Physics cross sections and event generation of  $e^+e^-$  annihilations at the CEPC [J/OL]. Chin. Phys. C, 2016, 40(3): 033001. DOI: [10.1088/1674-1137/40/3/033001](https://doi.org/10.1088/1674-1137/40/3/033001).
- [149] Gaede F, Aplin S, Glattauer R, et al. Track reconstruction at the ILC: the ILD tracking software [J/OL]. Journal of Physics: Conference Series, 2014, 513(2): 022011. DOI: [10.1088/1742-6596/513/2/022011](https://doi.org/10.1088/1742-6596/513/2/022011).
- [150] Yu D, Ruan M, Boudry V, et al. Lepton identification at particle flow oriented detector for the future  $e^+e^-$  higgs factories [J/OL]. The European Physical Journal C, 2017, 77(9). DOI: [10.1140/epjc/s10052-017-5146-5](https://doi.org/10.1140/epjc/s10052-017-5146-5).
- [151] Stone M. Cross-validatory choice and assessment of statistical predictions [J/OL]. Journal of the Royal Statistical Society. Series B (Methodological), 1974, 36(2): 111-147[2022-08-12]. <http://www.jstor.org/stable/2984809>.
- [152] Cowan G, Cranmer K, Gross E, et al. Asymptotic formulae for likelihood-based tests of new physics [J/OL]. The European Physical Journal C, 2011, 71(2). DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0).
- [153] Wang Y, Descotes-Genon S, Deschamps O, et al. Prospects for  $B_{(s)}^0 \rightarrow \pi^0 \pi^0$  and  $B_{(s)}^0 \rightarrow \eta \eta$  modes and corresponding  $CP$  asymmetries at Tera- $Z$  [J]. 2022.
- [154] Guo F, Fang Y, Lou X. Topo-cluster variable performance study with Radiative  $Z$  events in 13TeV ATLAS data [R/OL]. Geneva: CERN, 2021. <https://cds.cern.ch/record/2774734>.



致 谢

---

致 谢

2023 年 6 月



## 作者简历及攻读学位期间发表的学术论文与其他相关学术成果

### 作者简历:

郭方毅, 新疆维吾尔自治区博尔塔拉蒙古自治州博乐市人, 中国科学院高能物理研究所博士研究生。

2014年09月——2018年06月, 在中国科学技术大学物理学院近代物理系获得学士学位。

2018年09月——2023年06月, 在中国科学院高能物理研究所攻读博士学位。

2021年11月——2023年1月, 在欧洲核子研究中心(CERN)联合培养博士生。

### 工作经历:

2017年——至今, 参加位于欧洲核子研究中心的 ATLAS 实验, 主要参与希格斯物理在双光子道中的物理分析, 2021年获得合作组作者身份。

2018年——至今, 参加环形正负电子对撞机实验的前期预研工作, 包括  $H \rightarrow \gamma\gamma$  道物理性能研究和电磁量能器重建算法开发工作。

2021年11月——2023年1月, 被中国留学基金委国家建设高水平大学公派研究项目选为公派联合培养博士生, 前往 CERN 参与 ATLAS 实验相关工作。

### 已发表(或正式接受)的学术论文:

- (1) An F, et al., "Precision Higgs physics at the CEPC", Chin.Phys.C,2019,43(4):043002.
- (2) Q. Sha et al., "Probing Higgs CP properties at the CEPC", Eur. Phys. J. C 82, 981 (2022).
- (3) Fangyi Guo, Yaquan Fang, Gang Li and Xinchou Lou. The expected measurement precision of the branching ratio of the Higgs decaying to the di-photon at the CEPC[J]. Chinese Physics C.(Accepted)
- (4) ATLAS Collaboration, "Study of the CP property of the Higgs boson to electroweak boson coupling in the VBF  $H \rightarrow \gamma\gamma$  channel with the ATLAS detector", arxiv:2208.02338, accepted by Phys.Rev.Lett.

获奖情况：

2022 年 11 月，中国科学院高能物理研究所第二十二届所长奖学金优秀奖。