

國 立 中 央 大 學

物 理 學 系  
碩 士 論 文

Search for the rare decays of Z and Higgs bosons to  
J/ $\psi$  plus photon at  $\sqrt{s} = 13$  TeV

研 究 生：鄭皓仁

指 導 教 授：郭 家 銘

中 華 民 國 一〇八 年 七 月



# 國立中央大學圖書館 碩博士論文電子檔授權書

(104年5月最新修正版)

本授權書授權本人撰寫之碩/博士學位論文全文電子檔(不包含紙本 詳備註1說明)

明) 在 國立中央大學圖書館博碩士論文系統。(以下請擇一勾選)

()同意 (立即開放)

()同意 (請於西元 \_\_\_\_\_ 年 \_\_\_\_\_ 月 \_\_\_\_\_ 日開放)

()不同意,原因是 \_\_\_\_\_

在國家圖書館 臺灣博碩士論文知識加值系統」

()同意 (立即開放)

()同意 (請於西元 \_\_\_\_\_ 年 \_\_\_\_\_ 月 \_\_\_\_\_ 日開放)

()不同意 原因是 \_\_\_\_\_

以非專屬 無償授權國立中央大學 台灣聯合大學系統圖書館與國家圖書館，基於推動「資源共享 互惠合作」之理念 於回饋社會與學術研究之目的 得不限地域 時間與次數 以紙本 微縮、光碟及其它各種方法將上列論文收錄 重製 與利用，並得將數位化之上列論文與論文電子檔以上載網路方式 提供讀者基於個人非營利性質之線上檢索 閱覽 下載或列印。

研究生簽名 鄭伯仁 學號 106222038

論文名稱: Search for the rare decays of Z and Higgs bosons to

指導教授姓名 鄭家金 Z/γ plus photon

系所 物理 所 博士班 碩士班 at √S = 13 TeV

填單日期 108.06.21

## 備註

1. 本授權書之授權範圍僅限電子檔 紙本論文部分依著作權法第15條第3款之規定 採推定原則即預設同意圖書館得公開上架閱覽，如您有申請專利或投稿等考量，不同意紙本上架陳列 須另行加填申請書 詳細說明與紙本申請書下載請至本館數位博碩論文網頁
2. 本授權書請填寫並親筆簽名後 裝訂於各紙本論文封面後之次頁 全文電子檔內之授權書簽名 可用電腦打字代替)
3. 讀者基於個人非營利性質之線上檢索 閱覽 下載或列印上列論文 應遵守著作權法規定

# 國立中央大學碩士班研究生 論文指導教授推薦書

物理 學系/研究所 鄭皓仁 研究生所提之論

文 Search for the rare decays of  $Z$  and Higgs bosons to  $J/\psi$  plus photon  
(題目)  
at  $\sqrt{S} = 13 \text{ TeV}$

係由本人指導撰述，同意提付審查。

指導教授 郭家銘 (簽章)

107年9月5日

國立中央大學碩士班研究生  
論文口試委員審定書

物理 學系/研究所 鄭皓仁 研究生  
所提之論文

Search for the rare decays of Z and Higgs bosons to J/ψ plus photon at  $\sqrt{s} = 13$  TeV

經本委員會審議，認定符合碩士資格標準。

學位考試委員會召集人

林貴林

委

員

張元翰

郭家銘

中華民國 107 年 9 月 14 日

**Search for the rare decays of Z and Higgs bosons to  
J/ $\psi$  plus photon at  $\sqrt{s} = 13$  TeV**

by

**Hao-Ren Jheng**

at the

NATIONAL CENTRAL UNIVERSITY

June 2018

© National Central University 2018. All rights reserved.

Author..... Hao-Ren Jheng

Department of Physics and Center of high energy of high field  
physics, National Central University, Taiwan

Monday 24<sup>th</sup> June, 2019

Certified by..... Chia-Ming Kuo

Associate Professor  
Thesis Supervisor

**Search for the rare decays of Z and Higgs bosons to  $J/\psi$  plus  
photon at  $\sqrt{s} = 13 \text{ TeV}$**   
by  
**Hao-Ren Jheng**

## **Abstract**

A search is presented for decays of Z and Higgs bosons to a  $J/\psi$  meson and a photon, with the subsequent decay of the  $J/\psi$  to  $\mu^+ \mu^-$ . The analysis uses data from proton-proton collisions with an integrated luminosity of  $35.9 \text{ fb}^{-1}$  at  $\sqrt{s} = 13 \text{ TeV}$  collected with the CMS detector at the LHC. The observed limit on the  $Z \rightarrow J/\psi \gamma$  decay branching fraction, assuming that the  $J/\psi$  meson is produced unpolarized, is  $1.4 \times 10^{-6}$  at 95% confidence level, which corresponds to a rate higher than expected in the standard model by a factor of 15. For extreme-polarization scenarios, the observed limit changes from -13.6 to +8.6% with respect to the unpolarized scenario. The observed upper limit on the branching fraction for  $H \rightarrow J/\psi \gamma$  where the  $J/\psi$  meson is assumed to be transversely polarized is  $7.6 \times 10^{-4}$ , a factor of 260 larger than the standard model prediction. The results for the Higgs boson are combined with previous data from proton-proton collisions at  $\sqrt{s} = 8 \text{ TeV}$  to produce an observed upper limit on the branching fraction for  $H \rightarrow J/\psi \gamma$  that is a factor of 220 larger than the standard model value.

Thesis Supervisor: Chia-Ming Kuo  
Title: Associate Professor

# 中文摘要

本篇論文目的為尋找Z玻色子和希格斯玻色子衰變至一個J/ $\psi$ 介子和一個光子的頻道，其中J/ $\psi$ 介子進一步衰變至渺子對 ( $\mu^+\mu^-$ )。此分析使用於2016年由大型強子對撞機 (LHC) 產生的質子對撞，質心能量為13兆電子伏特 (TeV)，並由緊湊渺子線圈 (CMS) 所收集，對應於總光度35.9飛靶 (inverse femto-barn) 之事件。若J/ $\psi$ 介子不帶極性，在95%信心水準下，Z玻色子衰變頻道之衰變分支比例的觀測上限為 $1.4 \times 10^{-6}$ ，約對應至15倍的標準模型預測值。假設J/ $\psi$ 介子帶有縱向或橫向極性，衰變分支比例的觀測上限對應於不帶極性之假設有-13.6至+8.6%的差異。在J/ $\psi$ 介子帶有橫向極性且在95%信心水準下，希格斯玻色子衰變頻道之衰變分支比例的觀測上限為 $7.6 \times 10^{-4}$ ，約對應至260倍的標準模型預測值。在統計上，此希格斯玻色子衰變頻道之結果與由質心能量8兆電子伏特所得到之結果合併計算，得到衰變分支比例的觀測上限約對應至220倍的標準模型預測值。

# Acknowledgement

從剛開始對高能物理的認知僅限於科學人上的文章，到實際參與分析工作、完成這篇論文、甚至將研究結果發表到期刊上，首先要感謝的是指導老師郭家銘教授。記得大三上剛加入實驗室，總覺得自己是不是能力不足，group meeting裡完全聽不懂學長姐們討論的問題，不僅聽不懂，連該問怎樣的問題讓自己能了解都不知道。老師這句話「剛開始有很多不懂沒關係，每次meeting或每個禮拜強迫自己學到一件事情，那麼過了一年，你就能知道比一年前多幾十個概念。」讓我一直記著。有時我無法短時間內理解某些概念，老師也是耐心地一再解釋，直到我理解為止。除了這些學術上的教導，老師也提醒我如何正向看待每件事情，除了學術能力外，心態與壓力調適也是做研究中不可輕視及忽略的因素。不僅如此，我也從與老師相處過程裡知道，我必須要更open-minded、更謙虛，也要更能接受別人的意見。最後，也謝謝老師給予很多機會，不管是暑假去CERN做研究參加beamtest、參加大大小小會議、或是讓我在碩班當自己論文的contact person，即便有時候感到很辛苦(尤其是當contact person...)，但這些都是寶貴的經驗，也讓我更確信自己想走的方向。實驗室秘書張媽和曾小姐，任何文書、報帳、科技部計畫、獎學金、出國開會等等，我不管處理過多少次這種事情還是常常搞不清楚狀況，你們總是耐心地幫助我完成各種繁瑣的程序，謝謝你們不求回報的幫助。

Andrey，如果沒有你的幫助和教導，我覺得我無法順利建構和理解分析步驟。你的思考敏銳與深度，和如何有效率處理事情的方法，一直是我學習的目標。希望未來我也能成為像你一樣能獨立做研究和獨立思考。Shilpi，謝謝你教會我bias study及background method裡的各種技術細節。Hien，作為除了Andrey和Shilpi之外我們實驗室最資深的成員，在你待在台灣的時間裡，你每天都是最早到實驗室，也都很晚才離開，你的認真與認份一直是我的榜樣。本璿、政威和為尹學長，謝謝你們在我剛進實驗室時毫不保留地教我你們會的事情，一直以來我都期許自己在當學長時，也能像你們一樣將我會的事情交給學弟妹。你們所建立的實驗室文化，希望我也有好好地傳遞下去。而不僅僅是在學術方面，常常也會跟你們聊到各式各樣的話題，透過跟你們的討論，也讓我認識到不同的聲音與思考方向，每次都覺得很充實。Victoria，謝謝你幫助我熟悉latex，和有形無形上的英文指導。有時候透過跟你討論，也讓我對一些事情有更深一步的思考。又或是有時候跟你聊天的時候，常常會讓我哭笑不得，你

的腦袋迴路也是很神奇。培築，謝謝你平常都會跟我討論好看的日劇和劇情，讓我等不及就在週末兩天看完。平時你問我問題的時候，也讓我重新想過那些我也曾經問過學長或老師的，讓我更加深印象。明晏，謝謝你在我們都在CERN的時候常常討論一些實驗上的問題或是各種大小事，甚至不知道吃了幾餐你煮的飯。跟培築一樣，跟你討論時也讓我對過去已經聽過的概念更加熟悉。用威，謝謝你在我之前做HGCAL實驗時遇到硬體上的問題，都會替我講解。做為你的直屬學長，每次跟你吃家聚的時候都會聊到很多有趣的事情，讓我不會尷尬無聊。承璋(=爽哥、瑋爺、最成功的男人)，第一次聽到有學弟說從“欸，你待會吃完飯要幫我地毯式檢查我的年會海報嗎？”光是這句能吐槽的點就太多了，不一一贅述。祝你的memes (/mi:m/) 和研究都能加油。昀修a.k.a修哥、博學的男人、像風一樣的男人，你也是常常會讓我笑出來，尤其是博學的部分。也透過在教你不會的事情的時候，強烈感到我很沒有耐心，希望我後來漸漸有比較耐心一點。

我也要謝謝我大學的室友們，小白、槍槍、JEEG豪，和同學鳥哥、翊仁、翊哲。小白，雖然在花蓮的時候我們讀同一所國中、高中，但一直到上大學時才認識你，還一起當室友整個四年。謝謝你在兩年實驗班跟我一起當隊友，忍受我的脾氣和很多事情的固執與堅持。你真的是我認識的人裡面個性最好的人。槍槍，從大一的時候刷牙時遇到打個招呼的點頭之交，到大三後住在一起，知道你常常會突然爆出一句很好笑的話，或是時不時就開始玩奇怪的小遊戲，我記得寢室梗JEEG也是從你開始，從那時起無論什麼東西都可以拿來跟JEEG扯上關係。JEEG豪，你曾經在我很低潮的時候聽我講話，或是玩遊戲破不了關的時候幫我過關（?）。不管玩遊戲或是桌遊，你總是有辦法馬上理解遊戲規則，沒有人知道你到底為什麼不把才能分一點給你的學業，延續數 + 自30的傳奇。晚上聊天聊到凌晨三點還因為笑到肚子痛到睡不著、玩世紀帝國衝康槍槍把槍槍家擰爆、因為颱風學校整整停電三天兩夜、只好每天跑去中壢夜市亂晃打保齡球、搬宿舍晚出來幾分鐘被舍監逼寫悔過書只好發揮文采寫藏頭詩怒噴一波...能跟你們當上室友真是太好了。

最後我要謝謝我的家人，謝謝你們一直在背後支持鼓勵我，讓我可以專注在我有興趣的事情上。即便很多事情你們不能幫到我什麼，但是每次和你們講完後，我都能更安心和更加堅定。

# Contents

<b>1 Introduction</b>	<b>1</b>
1.1 The standard model of particle physics . . . . .	1
1.1.1 Gauge invariance . . . . .	2
1.1.2 Weak interaction and the electroweak unification . . . . .	6
1.1.3 The Higgs mechanism . . . . .	13
1.1.4 The production of the Higgs boson and its decays . . . . .	21
1.1.5 The measurement of the Higgs couplings . . . . .	28
1.2 The rare decays $Z/H \rightarrow J/\psi \gamma$ . . . . .	35
1.2.1 Overview . . . . .	35
1.2.2 Features of the decays . . . . .	37
1.2.3 Previous results from the ATLAS and CMS Collaborations . .	39
<b>2 Experimental apparatus</b>	<b>43</b>
2.1 Large Hadron Collider . . . . .	43
2.2 Compact Muon Solenoid . . . . .	45
2.3 Object reconstruction . . . . .	52
2.3.1 Particle-Flow algorithm . . . . .	52
2.3.2 Pile-up & Primary vertex . . . . .	58
<b>3 Analysis procedures</b>	<b>60</b>
3.1 Data and simulated samples . . . . .	60
3.1.1 Data sample . . . . .	60
3.1.2 Simulated samples . . . . .	61

<b>3.2 Trigger</b>	69
<b>3.3 Object identification</b>	80
<b>3.3.1 Muon identification</b>	80
<b>3.3.2 Photon identification</b>	88
<b>3.4 Event Selection</b>	89
<b>3.5 Background modeling</b>	108
<b>3.5.1 F-test</b>	109
<b>3.5.2 Bias study</b>	111
<b>3.6 Signal modeling</b>	124
<b>3.7 Systematic uncertainties</b>	124
<b>3.8 Statistical method</b>	133
 <b>4 Results and conclusion</b>	 140
<b>4.1 Limits on decay branching fraction</b>	140
<b>4.2 Conclusion</b>	142
<b>4.3 Outlook</b>	143
 <b>A Additional materials for the bias study</b>	 145
<b>A.1 Linearity</b>	145
<b>A.1.1 <math>H \rightarrow J/\psi \gamma</math></b>	146
<b>A.1.2 <math>Z \rightarrow J/\psi \gamma</math> Cat1</b>	148
<b>A.1.3 <math>Z \rightarrow J/\psi \gamma</math> Cat2</b>	150
<b>A.1.4 <math>Z \rightarrow J/\psi \gamma</math> Cat3</b>	152
<b>A.2 Pseudo-event</b>	154
<b>A.2.1 Pseudo-events for <math>H \rightarrow J/\psi \gamma</math></b>	155
<b>A.2.2 Pseudo-events for Cat1 of <math>Z \rightarrow J/\psi \gamma</math></b>	156
<b>A.2.3 Pseudo-events for Cat2 of <math>Z \rightarrow J/\psi \gamma</math></b>	157
<b>A.2.4 Pseudo-events for Cat3 of <math>Z \rightarrow J/\psi \gamma</math></b>	158
 <b>B Discussion on the systematic uncertainties</b>	 159

<b>C Beam test for the CMS high granularity endcap calorimeter in 2018</b>	<b>165</b>
C.1 BDT method for energy reconstruction . . . . .	167
C.2 Electron and pion separation . . . . .	177
C.3 Machine learning technique for Electron and pion separation . . . . .	182
<b>Bibliography</b>	<b>191</b>

# Chapter 1

## Introduction

The thesis is organized as follows. The theoretical background, from the concept of the gauge invariance, the electroweak theory to the Higgs mechanism, will be introduced. The experimental perspective and an overview of the searched decays  $Z/H \rightarrow J/\psi \gamma$  are followed. Chapter 2 will briefly mention the experiment apparatus, with the object reconstruction. In Chapter 3, the analysis procedure and methods, including data and simulated samples, the object identification, background and signal models construction, systematic uncertainties estimation, and the statistical methods, are described in detail. Chapter 4 represents the results of this analysis, as well as the possible improvements.

### 1.1 The standard model of particle physics

The standard model (SM) of particle physics provides so far the most effective and appropriate theory framework to describe the fundamental constituents of the Universe, and the interactions between them, the force<sup>1</sup>, which are carried by the gauge boson. The last piece of the SM is the Higgs boson, which is the manifestation of the mechanism by which particles acquire masses.

There are twelve fundamental fermions in the SM, and are categorized into

---

<sup>1</sup>The interactions here do not include the gravitation. In the following text, "the interactions in the SM" will simply refer to the electromagnetic, weak, and strong interactions.

quarks and leptons by the types of interactions they experience. All the fermions involve in the weak interaction, which is mediated by the  $W^\pm$  and  $Z$  bosons. Except for the electrically neutral neutrinos, the remaining nine fermions participate in the electromagnetic interaction, which is mediated by the photon  $\gamma$ . The theory of the electromagnetic interaction is the Quantum Electrodynamics (QED), which is the most accurately tested physics theory. The above two interactions can be unified into the Electro-Weak theory (EW), and will be described later in the text. Only the quarks carry the color charge and undergo the strong interaction, which is mediated by the gluons  $g$ . The theory for the strong interaction is the Quantum Chromodynamics (QCD). The color is a label for the three orthogonal states in the  $SU(3)$  symmetry group of the QCD. Quarks are always bound together to form hadrons, which can either be mesons (consist of a quark and a anti-quark) or baryons (consist of three quarks). This is the nature of the QCD, called color confinement – quarks are always observed to be confined to bound colorless states. An overview of QCD can be found in the lecture [1] and will not be discussed in this thesis. The elementary particles, and their basic properties, are summarized in Fig. 1-1.

### 1.1.1 Gauge invariance

In the context of Quantum Field Theory (QFT), particles are described by excitations of a quantum field which satisfies the quantum field equation. In a continuous system, the *field* represents the generalized coordinates at each point in space-time, and therefore is written in the form of a continuous function. The dynamics of the field is often expressed by the Lagrangian density  $\mathcal{L}(\phi_i, \partial_\mu \phi_i)$  where  $\phi_i$  is the field. Later in the text a simplified term "the Lagrangian" will be used to replace the Lagrangian density. The equation of motion describing the dynamics of the field can be derived from the Euler-Lagrange equation

$$\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_i} = 0. \quad (1.1)$$

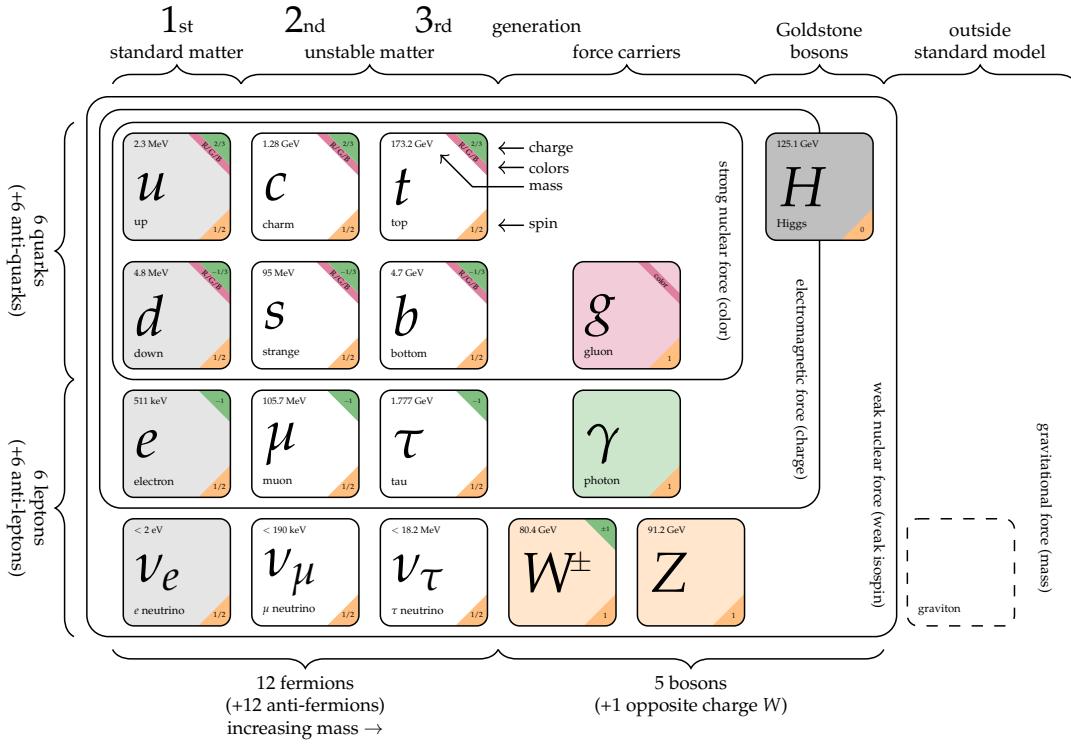


Figure 1-1: The elementary particles of SM, with the three generations of fermions, four gauge bosons, and the Higgs boson.

The three interactions, QED, weak, and QCD, can be derived by requiring the *local gauge invariance*: the Lagrangian is invariant under the *local phase transformation* of the fields,

$$\psi(x) \rightarrow \psi'(x) = U(x)\psi(x) = e^{iq\chi(x)}\psi(x). \quad (1.2)$$

The Lagrangian for a free spin- $\frac{1}{2}$  particle (referred to as free Lagrangian)

$$\mathcal{L}_{\text{free}} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi. \quad (1.3)$$

With the U(1) local gauge transformation, Eq. 1.3 becomes

$$\mathcal{L}_{\text{free}} \rightarrow \mathcal{L}'_{\text{free}} = \mathcal{L}_{\text{free}} - q\bar{\psi}\gamma_\mu(\partial_\mu\chi)\psi. \quad (1.4)$$

The free Lagrangian is obviously not invariant under U(1) local gauge transformation. The solution to deal with the extra term in Eq. 1.4 is to replace the derivative  $\partial_\mu$  in the free Lagrangian with the *covariant derivative*  $D_\mu$ ,

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + iqA_\mu, \quad (1.5)$$

with the introduction of a new field  $A_\mu$ . After the replacement, the new field  $A_\mu$  transforms in coordination with the local phase transformation of the  $\psi$  as

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu \chi, \quad (1.6)$$

The invariance of the Lagrangian can be preserved. It is worth noting that Eq. 1.6 is actually the concept of gauge transformation of the electromagnetic vector potential  $A_\mu$  in the classical electromagnetism. The requirement of the U(1) local invariance of the Lagrangian takes price, which is to introduce a vector field that couples to the spin- $\frac{1}{2}$  particles. The full Lagrangian should include this newly introduced vector field. The corresponding terms in the Lagrangian is known as the Proca Lagrangian

$$\mathcal{L}_{\text{Proca}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m_A^2 A^\mu A_\mu. \quad (1.7)$$

where the  $F_{\mu\nu} \equiv (\partial_\mu A_\nu - \partial_\nu A_\mu)$  is the field-strength tensor. However, the  $F_{\mu\nu}$  is invariant under Eq. 1.6 while the  $A^\mu A_\mu$  term transforms as

$$\frac{1}{2}m_A^2 A^\mu A_\mu \rightarrow \frac{1}{2}m_A^2 (A_\mu - \partial_\mu \chi)(A^\mu - \partial^\mu \chi) \neq \frac{1}{2}m_A^2 A^\mu A_\mu, \quad (1.8)$$

which is certainly not invariant. A conclusion can be drawn that the U(1) local gauge symmetry can only be satisfied with the *massless* gauge boson of the interaction. The Lagrangian describing the QED takes the form

$$\mathcal{L}_{\text{QED}} = (i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi) - (q\bar{\psi}\gamma_\mu\psi)A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \quad (1.9)$$

The introduction of the new field not only exhibits the observed gauge invari-

ance of classical electromagnetism, but also corresponds to a wave equation with an interaction term of the form

$$q\gamma^\mu A_\mu \psi. \quad (1.10)$$

This is the QED interaction potential, and its vertex is shown in Fig. 1-2. The re-

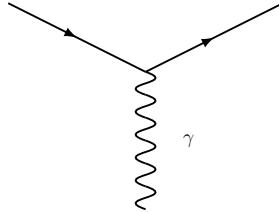


Figure 1-2: The Feynman diagram of the QED vertex.

quirement of the physics to be invariant under local U(1) phase transformations implies that a gauge field must exist, and the excitation of this field is now commonly identified as the massless gauge boson – the photon.

The same construction can be applied to the weak and the strong interactions (QCD, quantum chromodynamics), of which the underlying symmetry is the invariance under SU(2) and SU(3) local phase transformations respectively,

$$\psi(x) \rightarrow \psi'(x) = \exp \left[ ig_{W(S)} \boldsymbol{\alpha}(x) \cdot \boldsymbol{M} \right] \psi(x), \quad (1.11)$$

with the corresponding replacements of the partial derivatives to covariant derivatives,

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ig_{W(S)} \boldsymbol{M} \cdot \boldsymbol{G}_\mu(x), \quad (1.12)$$

where  $g_{W(S)}$  is the coupling constant of weak (strong) interaction,  $\boldsymbol{M}$  are the generators of SU(2) (SU(3)) symmetry group, and  $\boldsymbol{G}$  are the three (eight) new gauge fields of weak (strong) interaction. The well-known representations of the SU(2) group are the Pauli matrices and of the SU(3) are the Gell-Mann matrices.

In the following paragraphs, the weak interaction will be introduced a bit deeper.

### 1.1.2 Weak interaction and the electroweak unification

The weak interaction at first was proposed to explain the beta decay. Fermi (1933) treated the process as a contact interaction, which takes place at a single space-time point and does not require mediating particles. Nowadays, it is widely known that the Fermi's model is the low energy approximation and will fail at high energy regime.

At the beginning, this theory only includes the charged-current weak interaction which can be associated with invariance under SU(2) local phase transformation

$$\psi(x) \rightarrow \psi'(x) = \exp \left[ ig_W \chi(x) \cdot \mathbf{M} \right] \psi(x), \quad (1.13)$$

where  $\mathbf{M}$  are the three generators of the SU(2) symmetry group, of which the representation is the Pauli matrix,

$$\mathbf{M} = \frac{1}{2} \boldsymbol{\sigma}. \quad (1.14)$$

The local gauge invariance is satisfied with the three introduced fields,  $W_\mu^k$  with  $k = 1, 2, 3$ , corresponding to three gauge bosons  $W^{(1)}$ ,  $W^{(2)}$ , and  $W^{(3)}$ . Since the SU(2) generators are represented by  $2 \times 2$  matrices, the wavefunction must have two additional degrees of freedom. Furthermore, only left-handed (LH) chiral particles and right-handed (RH) chiral antiparticles couple to the weak charged-current interaction, LH particles and RH antiparticles are placed in weak isospin doublets. On the other hand, RH particles and LH antiparticles are put into weak isospin singlets and hence will not be affected by the transformation of Eq. [1.13]. Consequently, the wave functions can be interpreted as

$$\psi(x) = \begin{pmatrix} \mu_i \\ \ell_i \end{pmatrix}_L, \begin{pmatrix} u_i \\ d_i \end{pmatrix}_L, (u_i)_R, (d_i)_R, (\ell_i)_R, \quad (1.15)$$

where  $i = 1, 2, 3$  for the three families of fermions. Again, the requirement of the local gauge invariance necessitates the modification of the Dirac equation to

include a new interaction term

$$ig_W T_k \gamma^\mu W_\mu^k \psi_L = ig_W \frac{1}{2} \sigma_k \gamma^\mu W_\mu^k \psi_L, \quad (1.16)$$

where  $\psi_L$  stands for the weak isospin doublet of LH particles. From this form of interaction, three weak currents can be associated with Pauli matrices,

$$j_1^\mu = \frac{g_W}{2} \bar{\psi}_L \gamma^\mu \sigma_i \psi_L, \quad (1.17)$$

where  $i = 1, 2, 3$ . The actual charged-currents relate to the isospin raising the lowering operators,  $\sigma_\pm = \frac{1}{2}(\sigma_1 \pm i\sigma_2)$ , and read as

$$j_\pm^\mu = \frac{1}{\sqrt{2}} \left( j_1^\mu \pm ij_2^\mu \right) = \frac{g_W}{\sqrt{2}} \bar{\psi}_L \gamma^\mu \sigma_\pm \psi_L. \quad (1.18)$$

In the case of the doublet formed by the LH electron and electron neutrino, the currents  $j_\pm^\mu$ , corresponding to the exchange of the physical  $W^\pm$  bosons, are

$$j_+^\mu = \frac{g_W}{\sqrt{2}} (\bar{\nu}_L \bar{e}_L) \gamma^\mu \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \nu_l \\ e_L \end{pmatrix} = \frac{g_W}{\sqrt{2}} \bar{\nu} \gamma^\mu \frac{1}{2} (1 - \gamma^5) e, \quad (1.19)$$

$$j_-^\mu = \frac{g_W}{\sqrt{2}} (\bar{\nu}_L \bar{e}_L) \gamma^\mu \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \nu_l \\ e_L \end{pmatrix} = \frac{g_W}{\sqrt{2}} \bar{e} \gamma^\mu \frac{1}{2} (1 - \gamma^5) \nu, \quad (1.20)$$

consistent with the experimental observation of the vector minus axial vector (V-A) structure. The physical  $W$  bosons are identified as

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left( W_\mu^{(1)} \mp i W_\mu^{(2)} \right). \quad (1.21)$$

The  $SU(2)_L$  does not only give two weak charged-currents, but also implies the existence of a weak neutral-current

$$j_3^\mu = g_W \bar{\psi}_L \gamma^\mu \frac{1}{2} \sigma_3 \psi_L. \quad (1.22)$$

In the case of the fermion doublet (again, the LH electron and electron neutrino are used as an example), it reads as

$$j_3^\mu = g_W \frac{1}{2} (\bar{\nu}_L \bar{e}_L) \gamma^\mu \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \nu_l \\ e_L \end{pmatrix} = g_W \frac{1}{2} \bar{\nu}_L \gamma_L^\nu - g_W \frac{1}{2} \bar{e}_L \gamma^\mu e_L \quad (1.23)$$

or in a more compact form

$$j_3^\mu = I_W^{(3)} g_W \bar{f} \gamma^\mu \frac{1}{2} (1 - \gamma^5) f, \quad (1.24)$$

where  $f$  represents the fermion doublet and  $I_W^{(3)}$  is the third component of the weak isospin. The property that RH particles and LH antiparticles do not couple to the weak interaction is preserved, as they possess  $I_W^{(3)} = 0$ . (One should not mix this weak neutral-current with the SM Z boson that currently known, as the reason will be stated in the following paragraphs.)

There is another evidence and argument that the weak neutral-current must exist: the cross-section of the W boson pair production in the electron-positron collisions do not converge if there is no neutral-current interaction. Fig. [1-3] shows the leading order diagrams of the  $e^+e^- \rightarrow W^+W^-$  process. The left most diagram is the charged-current process. The middle one is the electromagnetic process as it is mediated by the photon, and there is also a  $\gamma WW$  vertex indicating that the  $\gamma$  can couple with W boson since they carry electric charge. In the right most diagram, a neutral boson, which is now known as the Z boson, acts as the mediator. Fig. [1-4] shows the predicted  $e^+e^- \rightarrow W^+W^-$  cross-sections of three cases: only the  $\nu_e$  diagram included; only  $\nu_e$  and  $\gamma$  diagrams included; all diagrams included [2]. With only the first two diagrams, the cross-section will increase without limit. The inclusion of the neutral-current interaction makes the calculated cross-section converge and consistent with the experimental observation.

The cancellation that preserves the unitary of  $e^+e^- \rightarrow W^+W^-$  indicates that the coupling of the  $\gamma$ , charged- and neutral-currents are related. A unification of the electromagnetic and weak interaction was proposed, and a unified electroweak

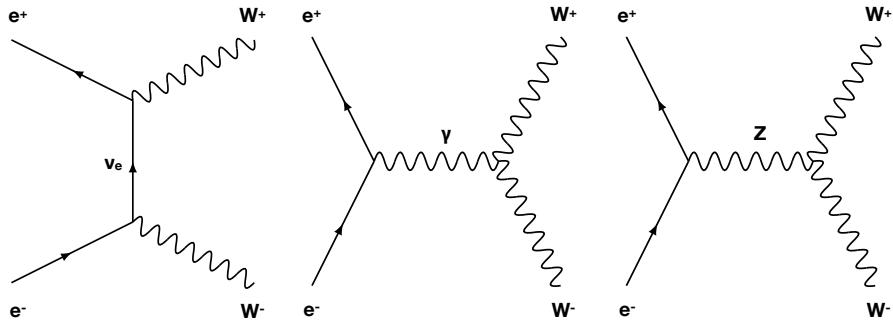


Figure 1-3: The leading order diagrams of the  $e^+e^- \rightarrow W^+W^-$  process.

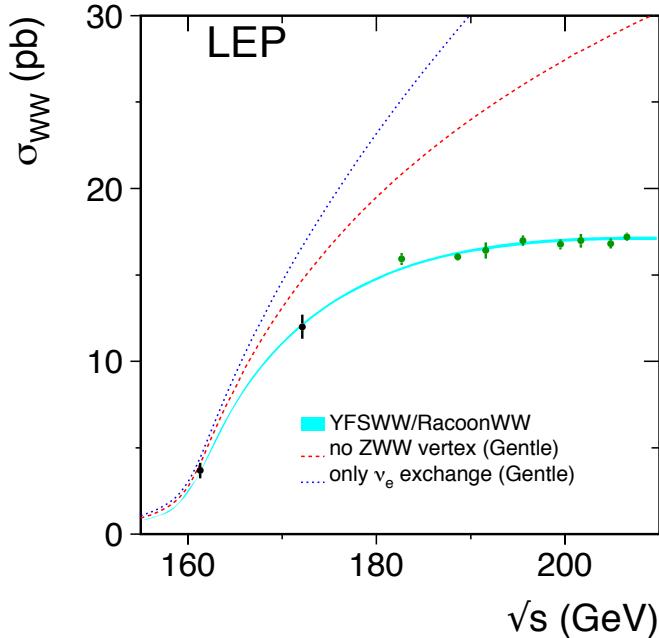


Figure 1-4: Measurements of the W-pair production cross-section, compared to the different predictions. The shaded area represents the uncertainty on the theoretical predictions [2].

model was completed by Sheldon Glashow, Abdus Salam, and Steven Weinberg, and now it is called GSW model.

One thing that must be incorporated in the unification is the correspondence between the weak neutral-current and the physical Z boson. The neutral-current

previously stated does not couple to RH particles/LH antiparticles, which is in contrast to the experimental evidence that the neutral Z boson couples, not equally, to both LH and RH particles. At the first step, a  $U(1)_Y$  local gauge symmetry is introduced to replace the  $U(1)$  gauge group of the electromagnetism with the transformation

$$\psi(x) \rightarrow \psi'(x) = U(x)\psi(x) = \exp\left[ig'\frac{Y}{2}\chi'(x)\right]\psi(x), \quad (1.25)$$

with a new field  $B_\mu$  and a new weak hypercharge  $Y$ . This new symmetry yields the same interaction term as the  $U(1)$  symmetry of the QED in Eq. [1.10],

$$g'\frac{Y}{2}\gamma^\mu B_\mu \psi. \quad (1.26)$$

The physical photon  $\gamma$  and Z boson are expressed as,

$$A_\mu = +B_\mu \cos \theta_W + W_\mu^{(3)} \sin \theta_W, \quad (1.27)$$

$$Z_\mu = -B_\mu \sin \theta_W + W_\mu^{(3)} \cos \theta_W, \quad (1.28)$$

where the  $\theta_W$  is the weak mixing angle. The physical QED and weak neutral-current are therefore,

$$j_{em}^\mu = j_Y^\mu \cos \theta_W + j_3^\mu \sin \theta_W, \quad (1.29)$$

$$j_Z^\mu = -j_Y^\mu \sin \theta_W + j_3^\mu \cos \theta_W, \quad (1.30)$$

with the weak neutral-current  $j_3$  of Eq. [1.23] and the current associated with the interaction term  $j_Y$  of Eq. [1.26]

$$j_Y^\mu = \frac{1}{2}g'Y_{e_L}\bar{\mathbf{e}}_L\gamma^\mu\mathbf{e}_L + \frac{1}{2}g'Y_{e_R}\bar{\mathbf{e}}_R\gamma^\mu\mathbf{e}_R + \frac{1}{2}g'Y_{\nu_L}\bar{\nu}_L\gamma^\mu\nu_L + \frac{1}{2}g'Y_{\nu_R}\bar{\nu}_R\gamma^\mu\nu_R \quad (1.31)$$

On the other hand, the electromagnetic current (of the electron doublet) is simply

$$j_{em}^\mu = Q_e e \bar{\mathbf{e}}_L \gamma^\mu \mathbf{e}_L + Q_e e \bar{\mathbf{e}}_R \gamma^\mu \mathbf{e}_R. \quad (1.32)$$

The underlying symmetry group of the electroweak sector, as described in GSW model, is  $U(1)_Y \times SU(2)_L$ . In order to preserve the invariance under  $U(1)_Y$  and  $SU(2)_U$  local gauge transformation, the hypercharges of particles in a weak isospin doublet should be the same. Having this argument and equating each component of the Eq. [1.29] with  $j_3^\mu$  and  $j_Y^\mu$  substituted and Eq. [1.32], the weak hypercharge can be expressed as a linear combination of the electromagnetic charge  $Q$  and the third component of weak isospin  $I_W^{(3)}$

$$Y = 2(Q - I_W^{(3)}), \quad (1.33)$$

Relations between the weak coupling  $g_W$ , the hypercharge coupling  $g'$  and the electric charge can be derived

$$e = g_W \sin \theta_W = g' \cos \theta_W. \quad (1.34)$$

The GSW model successfully bridges the couplings of QED, weak, and the hypercharge with the simple relation. The measurement of the weak mixing angle, in convention, provides the value of  $\sin^2 \theta_W$ , which is also the ratio of the weak to electromagnetic coupling constant

$$\sin^2 \theta_W = \frac{\alpha}{\alpha_W} = \frac{e^2}{g_W^2} \sim 0.23. \quad (1.35)$$

The coupling of the physical Z boson can be determined similarly. From Eq. [1.30], the current of the interaction between the Z boson and a fermion (with flavor  $f$ ) can be written as

$$\begin{aligned} j_Z^\mu &= g_Z(I_W^{(3)} - Q_f \sin^2 \theta_W) \bar{u}_L \gamma^\mu u_L - g_Z(Q_f \sin^2 \theta_W) \bar{u}_R \gamma^\mu u_R \\ &\equiv g_Z(c_L \bar{u}_L \gamma^\mu u_L + c_R \bar{u}_R \gamma^\mu u_R) \end{aligned} \quad (1.36)$$

where  $u_{L(R)}$  is the spinor of LH (RH) states,  $c_L = I_W^{(3)} - Q_f \sin^2 \theta_W$  and  $c_R = -Q_f \sin^2 \theta_W$  indicating the strengths of the coupling, and the coupling of the phys-

ical Z boson defined as

$$g_Z = \frac{g_W}{\cos \theta_W} = \frac{e}{\sin \theta_W \cos \theta_W}. \quad (1.37)$$

As stated previously, the physical Z boson does couple to LH and RH particles, however, unequally. This is intuitively reasonable, as the current associated with the Z boson is the mixture of the weak and  $U(1)_Y$  interactions, where the former one couples only to LH particles but the latter one equally couples to LH and RH particles.

In 1967, Steven Weinberg obtained the formula for the W and Z boson masses [3], with the  $\theta_W$  which had not yet been determined then. In the following years, the  $\theta_W$  was measured in various experiments, and in 1982 the masses of the W and Z bosons were predicted to be  $m_W = 82 \pm 2 \text{ GeV}/c^2$  and  $m_Z = 92 \pm 2 \text{ GeV}/c^2$ . In 1983, Carlos Rubbia and his group discovered the W and the Z boson [4, 5] with measured masses  $m_W = 80.403 \pm 0.029 \text{ GeV}/c^2$  and  $m_Z = 91.188 \pm 0.002 \text{ GeV}/c^2$ . Experiments later on also confirmed the couplings. The GSW model is now considered as one of the most important successes in the SM.

Despite the triumph of the electroweak unification, it did have some questions regarding the whole mechanism. First of all, Eq. [1.27] and [1.28] demonstrate that the fields of  $U(1)_Y$  and  $SU(2)_L$  are mixed to give physical bosons. The underlying nature of this mixture was unclear. Secondly, four electroweak gauge bosons have different masses, especially when comparing the photon with other three massive particles. This fact seems to contradict the physical picture that both electromagnetic and weak interactions are manifestations of a more fundamental electroweak interaction. The problem with the masses happens also on the fermions. In Eq. [1.9]

the mass term in the QED Lagrangian can be expressed in the chiral states

$$\begin{aligned}
-m\bar{\psi}\psi &= -m\bar{\psi}\left[\frac{1}{2}(1-\gamma^5) + \frac{1}{2}(1+\gamma^5)\right]\psi \\
&= -m\bar{\psi}\left[\frac{1}{2}(1-\gamma^5)\psi_L + \frac{1}{2}(1+\gamma^5)\psi_R\right] \\
&= -m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R).
\end{aligned} \tag{1.38}$$

In the  $SU(2)_L$  gauge transformation of the weak interaction, LH particles transform as doublets while RH particles as singlets. Eq. 1.38 obviously does not follow the required gauge invariance. Thirdly, a problem was found: the unitarity violation of the scattering process  $W^+W^- \rightarrow W^+W^-$ . An overview of the WW scattering process can be found in Ref. [6]. The original calculation for the amplitude included the diagrams, shown in Fig. 1-5. The unitarity violation results from the longitudinal polarized states of W boson and the process  $W_LW_L \rightarrow W_LW_L$ . The issue is solved by introducing a new scalar particle to mediate the WW process. The diagrams are shown in Fig. 1-6. All the above three problems necessitate a new mechanism, which is now called the Higgs mechanism, with its manifestation, the Higgs boson.

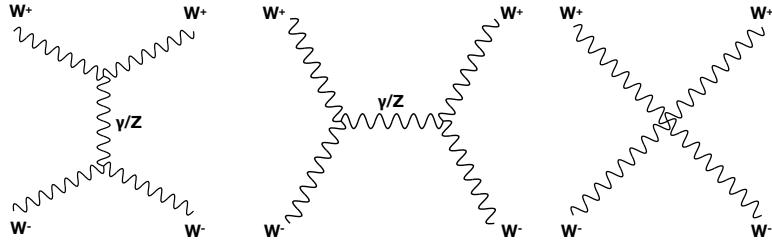


Figure 1-5: The leading order diagrams for  $W^+W^- \rightarrow W^+W^-$  scattering process.

### 1.1.3 The Higgs mechanism

The Higgs mechanism was proposed back to 1964 by Robert Brout and François Englert, Peter Higgs, and Gerald Guralnik, C. R. Hagen, and Tom Kibble [7, 8, 9].

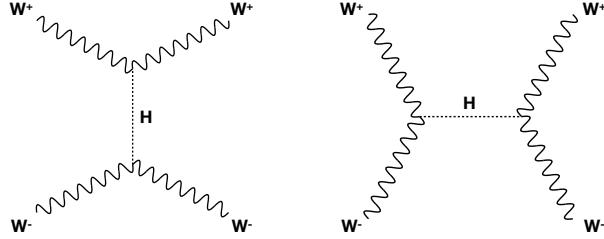


Figure 1-6: The diagrams for  $W^+W^- \rightarrow W^+W^-$  scattering process with a scalar boson as mediator.

Before formally introducing the Higgs mechanism in the SM, a single scalar field  $\phi$  is used as an example to illustrate the concept. Consider the potential of the form

$$V(\phi) = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4. \quad (1.39)$$

The corresponding Lagrangian is given by

$$\begin{aligned} \mathcal{L}_{ex} &= \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - V(\phi) \\ &= \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}\mu^2\phi^\mu - \frac{1}{4}\lambda\phi^4. \end{aligned} \quad (1.40)$$

In this example Lagrangian, the term of  $(\partial_\mu\phi)(\partial^\mu\phi)$  can be associated with the kinematic energy of the scalar particle. The term of  $\phi^2$  can be read as the mass of the particle (strictly to say, when  $\mu^2 > 0$ , it is the coefficient of the  $\phi^2$  term that associates to the mass). The  $\phi^4$  term is identified as self-interactions of the scalar field.

The vacuum state is the lowest energy state of the field. In the field theory, the particles state (or the excitations of the field) can be obtained by applying perturbations of the field around the vacuum state. In order to have minima for the potential, the  $\lambda$  must be positive. When  $\mu^2 > 0$ , the minimum of the potential happens to be at  $\phi = 0$ . When  $\mu^2 < 0$ , the term can no longer be interpreted as mass, and the potential now has two degenerate minima at  $\phi = \pm v = \pm |\sqrt{\frac{-\mu^2}{\lambda}}|$ . One needs to arbitrarily select one of the degenerate states as the ground state, then the ground state no longer preserves the symmetry of the Lagrangian. This way to

obtain the asymmetric vacuum state is known as *spontaneous symmetry breaking*.

In the SM, the Higgs mechanism is embedded in the  $U(1)_Y \times SU(2)_L$  local gauge symmetry of the electroweak sector. As the Higgs mechanism is required to generate masses of the electroweak gauge bosons, one of the scalar fields must be neutral (therefore termed as  $\phi^0$ ), and the other must be charged ( $\phi^+$  and  $\phi^- = (\phi^+)^*$ ) to give the longitudinal polarization states of the W bosons<sup>2</sup>. The simplest Higgs model, which has four degrees of freedom and consists of two complex scalar fields, is placed in a weak isospin doublet,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (1.41)$$

The Lagrangian of this doublet of fields is

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi), \quad (1.42)$$

To preserve the invariance under the  $U(1)_Y \times SU(2)_L$  local gauge transformation, the derivative in the Lagrangian should be replaced by the covariant derivative of the form

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ig_W \mathbf{T} \cdot \mathbf{W}_\mu + ig' \frac{Y}{2} B_\mu, \quad (1.43)$$

where  $\mathbf{T} = \frac{1}{2}\sigma$  are the three generators of the  $SU(2)$  group. The Higgs potential is of the form

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad (1.44)$$

where  $\lambda$  is positive. The visualization of the Higgs field is shown in Fig. 1-7. The potential is spherically symmetric, and thus the original Lagrangian is spherically symmetric. For  $\mu^2 < 0$ , the potential has infinite degenerate minima

$$\phi^\dagger \phi = \frac{1}{2} (\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}. \quad (1.45)$$

---

<sup>2</sup>Before the Higgs mechanism, the gauge bosons do not have masses. Hence, they can only have transverse polarization states. After acquiring the masses, gauge bosons become massive particles, which can have longitudinal polarization state.

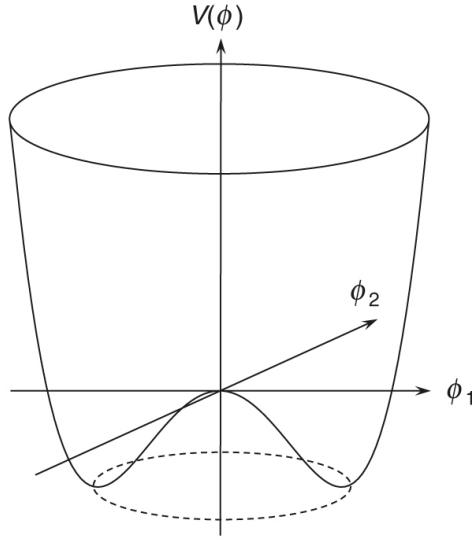


Figure 1-7: The Higgs potential for  $\mu^2 < 0$ .

For the neutral photon to be massless after the symmetry breaking, the vacuum state is chosen to be

$$\phi^{\text{vacuum}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (1.46)$$

The symmetry of the original Lagrangian is broken, given that a particular ground state is selected among the degenerate states. A field  $\eta$  is introduced when applying the perturbation around the vacuum state

$$\phi^{\text{vacuum}} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ v + \eta + i\phi_4 \end{pmatrix}. \quad (1.47)$$

By substituting Eq. 1.47 into the Lagrangian, however, will produce massless Goldstone bosons and terms associated with the couplings between the massive gauge fields and the Goldstone fields. An important fact is that every choice of the gauge transformation, as long as it follows correct form, will not break the symmetry of the Lagrangian. Therefore, a clever way to eliminate the Goldstone fields from the Lagrangian is to choose a gauge transformation called *Unitary gauge*, and after which the complex scalar fields will be entirely real. The Higgs doublet after the

Unitary gauge is written as

$$\phi^{\text{vacuum}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix}, \quad (1.48)$$

where  $\eta$  is replaced by  $h$ , which represents the physical field. After expanding all the terms of the Lagrangian, the masses of gauge bosons can be identified as the coefficients of the quadratic in the gauge fields.

In the Higgs doublet, the lower component is neutral ( $Q = 0$ ) and has  $I_W^{(3)} = -\frac{1}{2}$ , therefore the whole doublet has weak hypercharge  $Y = 1$ . Expanding the term  $(D_\mu \phi)^\dagger (D^\mu \phi)$

$$(D_\mu \phi)^\dagger (D^\mu \phi) = \frac{1}{2} (\partial_\mu h)(\partial^\mu h) + \frac{1}{8} g_W^2 (W_\mu^{(1)} + iW_\mu^{(2)}) (W^{(1)\mu} - iW^{(2)\mu}) (v + h)^2 + \frac{1}{8} (g_W W_\mu^{(3)} - g' B_\mu) (g_W W^{(3)\mu} - g' B^\mu) (v + h)^2. \quad (1.49)$$

one can identify the quadratic terms as

$$\frac{1}{8} v^2 g_W^2 (W_\mu^{(1)} W^{(1)\mu} + W_\mu^{(2)} W^{(2)\mu}) + \frac{1}{8} v^2 (g_W W_\mu^{(3)} - g' B_\mu) (g_W W^{(3)\mu} - g' B^\mu) \quad (1.50)$$

Identify the mass of the W boson by comparing

$$\frac{1}{2} m_W^2 W_\mu^{(1)} W^{(1)\mu} = \frac{1}{8} v^2 g_W^2 W_\mu^{(1)} W^{(1)\mu}, \quad (1.51)$$

therefore

$$m_W = \frac{1}{2} g_W v. \quad (1.52)$$

The mass the physical W boson is determined by the coupling constant of the  $SU(2)_L$  gauge interaction  $g_W$  and the vacuum expectation value of the Higgs field  $v$ .

The second term in Eq. 1.50 is associated with the neutral  $W^{(3)}$  and  $B$  fields,

and can be written as

$$\frac{1}{8}v^2 \left( g_W W_\mu^{(3)} - g' B_\mu \right) \left( g_W W^{(3)\mu} - g' B^\mu \right) = \frac{1}{8}v^2 \begin{pmatrix} W_\mu^{(3)} & B_\mu \end{pmatrix} \begin{pmatrix} g_W^2 & -g_W g' \\ -g_W g' & g'^2 \end{pmatrix} \begin{pmatrix} W^{(3)\mu} \\ B^\mu \end{pmatrix} \quad (1.53)$$

The matrix (referred to as mass matrix) appearing in the equation is non-diagonal, showing that the off-diagonal elements couple the  $W^{(3)}$  and  $B$  fields and allow them to mix. The physical boson fields (termed as  $Z_\mu$  and  $A_\mu$ ) correspond to the eigenstates of the mass matrix, which can be obtained by solving the characteristic equation

$$\det(\mathbf{M} - \lambda I) = (g_W^2 - \lambda)(g'^2 - \lambda) - g_W^2 g'^2 = 0. \quad (1.54)$$

As a result, the eigenvalues  $\lambda = 0$  or  $g_W^2 + g'^2$  with the eigenstates

$$\begin{aligned} A_\mu &= \frac{g' W_\mu^{(3)} + g_W B_\mu}{\sqrt{g_W^2 + g'^2}}, \quad m_A = 0 \text{ (photon)} \\ Z_\mu &= \frac{g_W W_\mu^{(3)} - g' B_\mu}{\sqrt{g_W^2 + g'^2}}, \quad m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2} \text{ (Z boson)} \end{aligned} . \quad (1.55)$$

Now, by defining the ratio of the coupling as

$$\frac{g'}{g_W} = \tan \theta_W, \quad (1.56)$$

Eq. 1.55 can be expressed as

$$\begin{aligned} A_\mu &= +B_\mu \cos \theta_W + W_\mu^{(3)} \sin \theta_W \\ Z_\mu &= -B_\mu \sin \theta_W + W_\mu^{(3)} \cos \theta_W. \end{aligned}$$

Eq. 1.27 and 1.28 are retained. With Eq. 1.56, the mass of the physical Z boson is

$$m_Z = \frac{1}{2} \frac{g_W}{\cos \theta_W} v. \quad (1.57)$$

Combining with the W boson mass from Eq. 1.52, one would obtain

$$\frac{m_W}{m_Z} = \tan \theta_W. \quad (1.58)$$

The mass of the Higgs boson  $m_H$  can be identified as the quadratic term in the Higgs boson field which is generated by the potential  $V(\phi)$  in the Lagrangian,

$$m_H^2 = 2\lambda v^2. \quad (1.59)$$

In Eq. 1.49, the gauge boson fields appears in the form of  $VV(v + h)^2$ , where  $V$  stands for gauge fields. The  $VVv^2$  terms relate to the mass of the gauge bosons, and the  $VVvh$  and  $VVhh$  terms represent the triple and quartic couplings between the Higgs bosons and the gauge bosons. From the weak theory, the physical W bosons are constructed as linear combination of the  $W^{(1)}$  and  $W^{(2)}$ , as shown in Eq. 1.21. Hence, the second term in Eq. 1.49 associated with the  $W^{(1)}$  and  $W^{(2)}$  can be rewritten as

$$\frac{1}{4}g_W^2 W_\mu^- W^{+\mu} (v + h)^2 = \frac{1}{4}g_W^2 v^2 W_\mu^- W^{+\mu} + \frac{1}{2}g_W^2 v W_\mu^- W^{+\mu} h + \frac{1}{4}g_W^2 W_\mu^- W^{+\mu} hh. \quad (1.60)$$

The first terms gives the masses of W boson as stated previous, the second term represents the triple  $HW^+W^-$  coupling, and the third term gives rise to the quartic  $HHW^+W^-$  coupling. The coupling strength of the  $HW^+W^-$  vertex is

$$g_{HWW} = \frac{1}{2}g_W^2 v = g_W m_W. \quad (1.61)$$

Similarly, the coupling  $HZZ$  can be derived  $g_{HZZ} = \frac{g_W}{\cos \theta_W} m_Z \equiv g_Z m_{cPZ}$ . The couplings of the Higgs boson and the gauge bosons are proportional to the mass of the gauge bosons.

As mentioned previously, the fermion mass term  $-m\bar{\psi}\psi = -m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)$  is not invariant under  $SU(2)_L \times U(1)_Y$  transformation, since the RH and LH fermions

transform differently

$$\begin{aligned} \text{LH doublet fermions : } \psi_L &\rightarrow \psi'_L = \psi_L e^{ig_W T \cdot W + ig' \frac{Y}{2} B} \\ \text{RH singlet fermions : } \psi_R &\rightarrow \psi'_R = \psi_R e^{ig' \frac{Y}{2} B} \end{aligned} . \quad (1.62)$$

The solution is to construct a *singlet* under  $SU(2)_L \times U(1)_Y$  in the Lagrangian. Consider an infinitesimal  $SU(2)$  local transformation on the  $SU(2)$  doublet  $\phi$  of the Higgs fields,

$$\phi \rightarrow \phi' = (I + ig_W \epsilon(x) \cdot T)\phi, \quad (1.63)$$

where  $T$  are generators of the  $SU(2)$  group. The LH doublets  $L$  undergoes the same transformation

$$\begin{aligned} L &\rightarrow L' = (I + ig_W \epsilon(x) \cdot T)L \\ \bar{L} &= L^\dagger \gamma^0 \rightarrow \bar{L}' = \bar{L}(I - ig_W \epsilon(x) \cdot T) \end{aligned} \quad (1.64)$$

It is clear that a term of  $\bar{L}\phi$  is invariant under the  $SU(2)_L$  transformation, or in other word, a singlet under  $SU(2)_L \times U(1)_Y$ . The effects of the transformation on the  $\phi$  and  $\bar{L}$  compensate to each other. Combining the  $\bar{L}\phi$  with RH singlet  $R$  also results in a singlet under  $SU(2)_L \times U(1)_Y$  (The conjugate of the combination is also a singlet). Consequently, a term in the Lagrangian of the form  $-y_f(\bar{L}\phi R + \bar{R}\phi^\dagger L)$  possesses the  $SU(2)_L \times U(1)_Y$  gauge symmetry. The Lagrangian, after spontaneous symmetry breaking and in the unitary gauge, is now

$$\mathcal{L}_{\text{fermion mass}} = -\frac{y_f}{\sqrt{2}} v (\bar{\ell} f_R + \bar{f}_R \ell) - \frac{y_f}{\sqrt{2}} (\bar{\ell} f_R + \bar{f}_R \ell). \quad (1.65)$$

where  $y_f$  is a constant known as *Yukawa coupling*. The first term corresponds to the fermion masses,  $m_\ell = \frac{y_f v}{\sqrt{2}}$ , representing the coupling of the fermions to the Higgs field through the non-zero vacuum expectation value. The second term corresponds to the interaction between the fermions and the physical Higgs boson.

The non-zero vacuum expectation value appears only in the lower component of the Higgs doublet, thus only fermions in the lower component of the  $SU(2)$

doublet (charged fermions and down-type quarks) can acquire masses, which is obviously not the case. The way to give masses to up-type quarks is to construct the conjugate doublet of the Higgs field  $\phi_c$  which transforms in the same way as the doublet  $\phi$

$$\phi_c = -i\sigma_2\phi^* = \begin{pmatrix} -\phi^{0*} \\ \phi^- \end{pmatrix} = \begin{pmatrix} -\phi_3 + i\phi_4 \\ \phi_1 - i\phi_2 \end{pmatrix} \quad (1.66)$$

The Lagrangian of the up-type quark masses is the same as Eq. [1.65] except  $\phi$  now is replaced by  $\phi_c$ . Consequently, the Lagrangian, after the symmetry breaking, is

$$\mathcal{L}_{\text{up-type quark masses}} = -\frac{y_{f,\text{up}}}{\sqrt{2}}v(\bar{u}u_R + \bar{u}_Ru) - \frac{y_{f,\text{up}}}{\sqrt{2}}(\bar{u}u_R + \bar{u}_Ru). \quad (1.67)$$

where the up-type quark masses can be identified as  $m_{\text{up}} = \frac{y_{f,\text{up}}v}{\sqrt{2}}$ . The Yukawa coupling of the fermions to the Higgs field is jointly written as

$$y_f = \frac{\sqrt{2}m_f}{v}, \quad (1.68)$$

and its value is determined to be consistent with the observed fermion masses.

The neutrino masses are yet another story. The possible mechanism to account for the neutrino masses was first introduced in Ref. [10, 11], and is now known as the seesaw mechanism. This mechanism will not be discussed in this thesis.

A review of the Higgs boson production at the LHC will be introduced in the next sub-section.

### 1.1.4 The production of the Higgs boson and its decays

The main production processes at the hadron collider are gluon-gluon fusion (ggF), vector boson fusion (VBF, or qqH), associated vector boson production (VH), and associated top quark pair production (ttH). The diagrams for these production modes are shown in Fig. [1-8] and the Higgs boson production cross-sections at the center-of-mass frame energy  $\sqrt{s} = 13 \text{ TeV}$  are shown in Fig. [1-9] [12]. The profound results of the deep inelastic scattering experiments showed that the mo-

mentum of the proton is not only carried by its three valence quarks, but also by the gluons that mediate the strong interaction between quarks. In high energy collisions at the LHC, the majority of energy is carried by gluons, and hence the hard processes are dominantly produced by the gluon-gluon interactions.

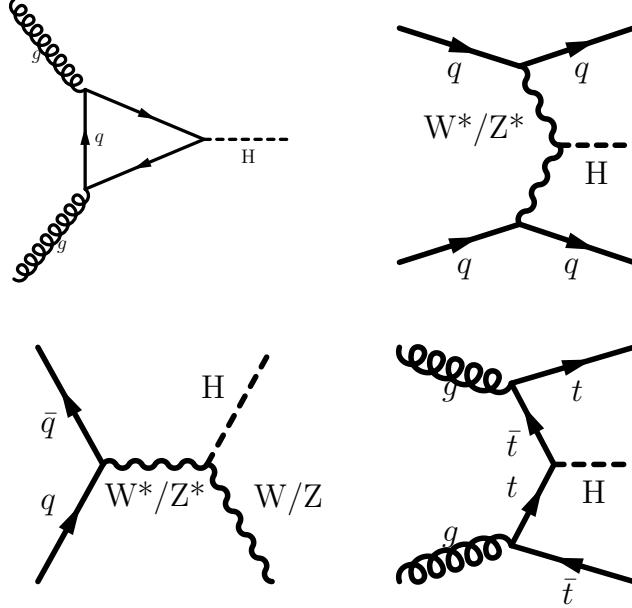


Figure 1-8: The diagrams for dominant production modes. (Top left) gluon-gluon fusion; (Top right) vector boson fusion; (Bottom left) associated vector boson production; (Bottom right) associated top quark pair production.

Since the Higgs boson is the manifestation of the Higgs mechanism which gives fundamental particles masses, in principle it can decay into all particles, if it is kinematically allowed. The decay probability is interpreted as branching ratio. The branching ratio of the most important decay channels as function of the Higgs boson mass are shown in Fig 1-10. In the following paragraphs, I will discuss main decay channels of the Higgs boson.

The Higgs boson cannot decay into top quarks as the top quark is too heavy [13]. The coupling between the Higgs boson and the top quark  $y_t$  is then realized in terms of the  $t\bar{t}H$  production and loops of virtual top quarks in the  $ggF$  production or in the decays to the massless particles, such as  $H \rightarrow \gamma\gamma$  and  $H \rightarrow gg$ . The combined measurement of the rate of Higgs boson production through gluon-

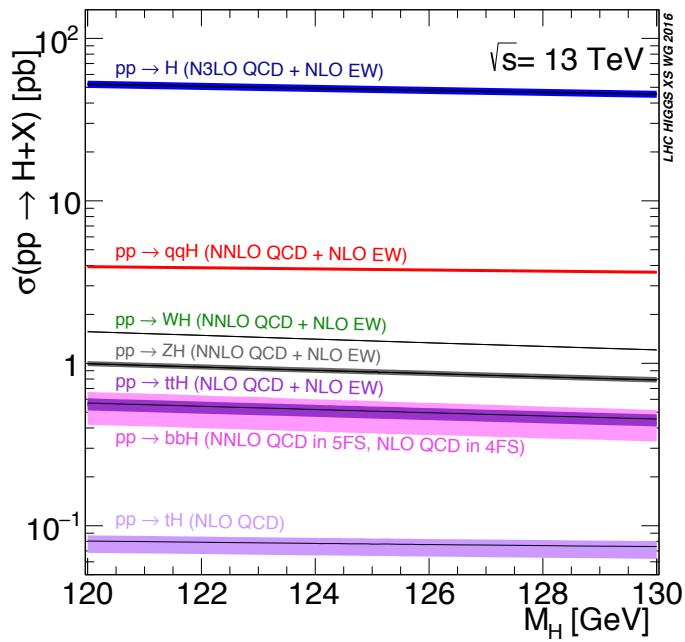


Figure 1-9: The SM Higgs boson production cross sections at  $\sqrt{s} = 13 \text{ TeV}$  as a function of the Higgs boson mass [12].

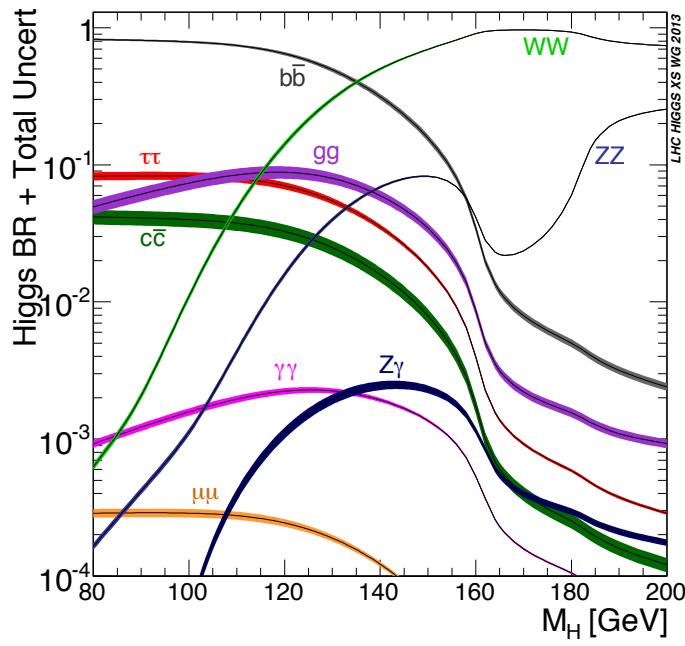


Figure 1-10: The SM Higgs boson decay branching ratios [12].

gluon fusion and of the  $H \rightarrow \gamma\gamma$  decay with LHC Run1 data suggested that the Higgs boson coupling to top quarks is consistent with SM prediction within uncertainties [14]. A measurement of the production rate of the tree-level  $t\bar{t}H$  process can provide further information as to whether there exists non-SM particles in the loops that introduce terms compensating for other deviations from the SM. The analysis is very difficult, as the top-quark decays to a W bosons and b-quark, and shortly afterwards the W decays hadronically to two jets or leptonically to a lepton and a neutrino. Both the ATLAS and CMS Collaboration recently observe this production channel, and establish the confirmation of the tree-level coupling of the Higgs boson to top quarks with the combined analyses of datasets collected at  $\sqrt{s} = 7, 8,$  and  $13\text{ TeV}$  [15, 16]. The best-fit signal strength  $\hat{\mu}$  from the ATLAS measurement is  $1.32^{+0.28}_{-0.26}(\text{Total}) \pm 0.18(\text{Stat.})^{+0.21}_{-0.19}(\text{Syst.})$ , and from the CMS is  $1.26^{+0.31}_{-0.26}(\text{Total}) \pm 0.16(\text{Stat.})^{+0.27}_{-0.22}(\text{Syst.})$ . The ATLAS obtained a significance of 6.3 standard deviations ( $\sigma$ ) relative to the background-only hypothesis, where the expected significance is  $5.1\sigma$ . The CMS also obtained the observed significance of  $5.2\sigma$  with the expected significance is  $4.2\sigma$ . The Higgs-top coupling can also be probed in the search for the production of Higgs boson in association with a single top quark. The production cross-section of this process is not only sensitive to the absolute values of the modifiers of the Higgs-top coupling,  $\kappa_t$ , and the coupling of vector bosons to the Higgs boson,  $\kappa_V$ , but also to their relative signs with respect to those predicted in the SM. Hence, it provides additional information toward the nature of the Higgs boson. The CMS Collaboration performs this search with data collected in 2016 [17], and the results show that the observed data favor positive sign of the coupling.

The largest branching ratio of the Higgs boson of mass  $m_H = 125\text{ GeV}$  is to bottom quarks, with  $\mathcal{BR}(H \rightarrow b\bar{b}) \approx 58.2\%$ . The measurement of the rate of the  $H \rightarrow b\bar{b}$  decay offers a direct test to the magnitude of Hbb coupling, while the relative sign of the coupling can be determined by the decay process  $H \rightarrow Y + \gamma$ , where the Y meson is the bound state of the b and anti-b quarks [18]. In order to suppress the QCD backgrounds, the analysis is designed to search for the VH

production where a W or Z boson decays leptonically, corresponding to five independent channels:  $Z(\ell\ell)H$ ,  $W(\ell\nu)H$ , and  $Z(\nu\nu)H$  where  $\ell = e, \mu$ . A multivariate regression technique [19, 20, 21] is applied to calibrate the measured energy of the b-tagged jets to improve the dijet mass resolution, after which the mass resolution is approximately 10–15%. Both the ATLAS and CMS Collaborations recently observe this decay channel. The CMS Collaboration reaches an observed (expected) significance of  $5.6$  ( $5.5$ )  $\sigma$  with the signal strength of  $\hat{\mu} = 1.04 \pm 0.20$  [22]. The ATLAS Collaboration announces an observed (expected) significance of  $5.4$  ( $5.5$ )  $\sigma$  with the signal strength  $\hat{\mu} = 1.01 \pm 0.20$  [23].

The  $H \rightarrow \tau^+\tau^-$  decay mode has been considered as the only accessible leptonic decay mode that probes the coupling of the Higgs boson to the fermionic sector. It can also be used to constrain CP violation in the VBF production [24] and provide sensitivity to CP violation in the Higgs boson coupling to leptons [25]. This decay benefits from a favorable signal-to-background conditions than the  $H \rightarrow b\bar{b}$  decay, however, slightly worse mass resolution of  $\approx 10 - 20\%$ , resulting from the inaccuracy of the momentum reconstruction of the  $\tau$  lepton. The  $\tau$  lepton can decay leptonically as  $\tau \rightarrow \nu_\tau \ell \bar{\nu}_l$  where  $\ell = e, \mu$ , and hadronically to charged or neutral pions. The analyses from both the ATLAS and CMS utilize the four most sensitive  $\tau\tau$  final states:  $e\mu$ ,  $e\tau_h$ ,  $\mu\tau_h$ , and  $\tau_h\tau_h$ , where  $\tau_h$  denotes the hadronic decay. The ATLAS Collaboration reports the signal strength  $\hat{\mu} = 1.09^{+0.36}_{-0.30}$  with an observed (expected) significance of  $6.4$  ( $5.4$ )  $\sigma$  with a combined analysis with  $\sqrt{s} = 7$ ,  $8$ , and  $13$  TeV data [26]. The CMS Collaboration also obtains the signal strength  $\hat{\mu} = 1.09^{+0.27}_{-0.26}$  with an observed (expected) significance of  $5.9$  ( $5.9$ )  $\sigma$  in combination with Run1 data [27].

Prior to the discovery of the Higgs boson, the decay mode  $H \rightarrow WW$  was considered the most sensitive channel in the mass range around the WW threshold of  $160$  GeV, and thus was important to the exclusion in such range. The  $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$  analysis profits from the fact that it has large branching fraction and has a relatively low-background final state. As a result, this decay channel has very good sensitivity to most production processes, in particular ggF and VBF. However, the

presence of neutrinos in the final state prevents the full reconstruction of the Higgs boson mass, and hence worse mass resolution of  $\approx 20\%$ . The different-flavor leptonic decay mode  $e\mu$  has the largest branching fraction, is the least affected by background processes, and therefore is the most sensitive channel of the analysis. The ATLAS Collaboration provides results of ggF and VBF production with 2016 data separately [28]. For the ggF production the signal strength  $\hat{\mu} = 1.21^{+0.22}_{-0.21}$  with an observed (expected) significance of  $6.3$  ( $5.2$ )  $\sigma$ , while for the VBF the signal strength  $\hat{\mu} = 0.62^{+0.37}_{-0.36}$  with an observed (expected) significance of  $1.9$  ( $2.7$ )  $\sigma$ . The CMS Collaboration reports the signal strength  $\hat{\mu} = 1.28^{+0.18}_{-0.17}$  with an observed (expected) significance of  $9.1$  ( $7.1$ )  $\sigma$ , combining all considered channels [29].

The  $H \rightarrow ZZ^* \rightarrow 4\ell$  ( $\ell = e$  or  $\mu$ ) decay has low branching fraction, but fortunately has the lowest background contamination, resulting in very good sensitivity. It provides the direct probe in constraining the  $HZZ$  coupling. The precise reconstruction of the final state products allows the complete determination of the kinematics of the reconstructed Higgs boson with mass resolution of  $\approx 1 - 2\%$ , which makes it one of the most important channels to measure the properties of the Higgs boson. The ATLAS and CMS Collaborations have both performed analyses for this channel with the Run1 data to determine the mass and spin-parity of the boson [30, 31, 32, 33, 34], its width [35, 36, 37], the fiducial cross sections [22, 23], and the tensor structure of its interaction with a pair of neutral gauge bosons [32, 34, 36]. These measurements provided results that are so far consistent with the SM predictions. The CMS Collaboration provides results, based on the combined data collected in 2016 and 2017, of the signal strength  $\hat{\mu} = 1.06^{+0.15}_{-0.13}$  [38]. The ATLAS Collaboration reports the signal strength  $\hat{\mu} = 1.18^{+0.13}_{-0.13}$  [39]. A model-independent measurement of the Higgs boson width is performed by the CMS Collaboration with 2016 data using the  $m_{4\ell}$  distribution in the range  $105 < m_{4\ell} < 140$  GeV, and is able to constrain the width to be  $\Gamma_H < 1.10$  ( $1.60$ ) GeV at 95% confidence level (CL) for observed (expected) value [40].

Despite the small branching fraction predicted by the SM, the  $H \rightarrow \gamma\gamma$  decay provides a clean final state, two energetic photons, with an invariant mass peak

that can be reconstructed with high precision with mass resolution of  $\approx 1 - 2\%$ . Consequently, this channel was one of the most important channels for the Higgs boson discovery and first measurements of its properties [41, 42]. Since the  $H \rightarrow \gamma\gamma$  decay proceeds mainly through W- and top-loop processes, interference effects make its branching fraction sensitive to the relative sign of the fermion and vector boson couplings. The differential cross sections enables us to test the perturbative QCD predictions for Higgs boson production, and can be used to probe the spin and CP properties of the Higgs boson. The CMS Collaboration provides the results using 2016 data of the signal strength  $\hat{\mu} = 1.18^{+0.17}_{-0.14}$  [43], while the ATLAS Collaboration obtains  $\hat{\mu} = 0.99^{+0.14}_{-0.14}$ . The interpretation of the coupling measurements from both collaborations shows that the observed data favors the positive sign of the coupling [44, 45]. The ATLAS Collaboration also tries to investigate the strength and tensor structure of the Higgs boson interactions using an effective Lagrangian, which introduces additional CP-even and CP-odd interactions [45], but no significant new physics contributions are observed.

The decay of  $H \rightarrow Z/\gamma^* + \gamma$  shares similar diagrams to the  $H \rightarrow \gamma\gamma$  decay, where in the former one a Z boson or a virtual photon  $\gamma^*$  is radiated from the loop. Measurement of this rare decay can enhance the current understanding of the nature of the Higgs boson, and can also provide an alternative way to test if there is any beyond standard model (BSM) couplings induced in the loop diagrams. A brief summary of these extension of SM can be found in Ref. [46, 47]. If there exists BSM that is manifested through CP violation, one can also observe the anomaly though a measurement of the forward-backward asymmetry. The ATLAS Collaboration sets an observed (expected) exclusion upper limit on the production cross section times the branching ratio of the  $H \rightarrow Z\gamma$  decay of 6.6 (5.2) times the SM prediction at 95% CL for a Higgs boson mass  $m_H = 125.09$  GeV, while the upper limits from the CMS Collaboration varies between 6.1 and 11.4 (3.9 and 9.1) times the SM value in the mass range of  $120 < m_H < 130$  GeV [46, 47]. The CMS Collaboration also provides so far the most stringent limit on the  $H \rightarrow \gamma^*\gamma$  decay, varying between 1.4 and 4.0 (2.1 and 2.3) times the SM prediction in the range of

$120 < m_H < 130$  GeV [46].

The rare decay  $H \rightarrow \mu\mu$  offers the best possibility to measure the Higgs coupling to second-generation fermions at the LHC. The expected branching fraction for a Higgs boson mass  $m_H = 125.09$  GeV is  $\mathcal{BR}(H \rightarrow \mu\mu) \approx 2.2 \times 10^{-4}$  [48] which is roughly one order of magnitude smaller than the  $H \rightarrow Z/\gamma^* + \gamma$  decay, owing to the small Yukawa coupling of the muon to the Higgs field. The CMS Collaboration sets the observed (expected) upper limit on the signal strength of 2.92 (2.16) times the SM prediction, with combination of 7, 8, and 13 TeV data [49], while the ATLAS Collaboration reports an upper limit of 2.1 (2.0) times the SM values [50].

The other decay of the Higgs boson to second-generation fermions that was searched for is the  $H \rightarrow c\bar{c}$  process. It is commonly considered impossible to discover this channel even in high luminosity run of the LHC (HL-LHC) due to the small branching fraction, large background in hadron collider, and jet flavor identification inefficiency [51, 52]. Nevertheless, direct search for the  $H \rightarrow c\bar{c}$  decay is important in the long-term perspective, as the development of the charm-tagging technique and the direct constraint of the Higgs-charm coupling would be valuable inputs to the next generation of particle colliders. The ATLAS Collaboration presents the first search for this process with data collected in 2016, utilizing the  $ZH$  production with the subsequent decay of the  $Z$  boson to dilepton. The observed (expected) upper limit on the production cross-section  $\sigma(pp \rightarrow ZH) \times \mathcal{BR}(H \rightarrow c\bar{c})$  is found to be  $2.7$  ( $3.9^{+2.1}_{-1.1}$ ) pb at the 95% CL, corresponding to an observed (expected) upper limit on the signal strength  $\hat{\mu} < 110$  ( $150^{+80}_{-40}$ ) [53].

### 1.1.5 The measurement of the Higgs couplings

The ATLAS and CMS Collaborations both reported the observation of a new boson with a mass of  $m_H = 125.09 \pm 0.21(\text{stst.}) \pm 0.11(\text{syst.})$  GeV [54] in 2012, and subsequent measurements revealed its Higgs-boson-like properties [31, 32, 55, 56, 57, 58, 59, 60]. One of the important analyses, and most related to this thesis, is the measurement of the Higgs coupling. A combined measurement were performed

by ATLAS and CMS with data collected at 7 and 8 TeV [14], and the CMS Collaboration provides the latest results with 13 TeV data [61]. The results from CMS with 13 TeV data will be shown in the following paragraphs.

The inputs of the analysis are the four main production processes introduced previously, decay channels to bosons  $H \rightarrow ZZ$ ,  $WW$ ,  $\gamma\gamma$ , and to fermions  $H \rightarrow \tau\tau$ ,  $b\bar{b}$ ,  $\mu\mu$ . In this work, a so-called  $\kappa$ -framework [62] is used<sup>3</sup>. Within the framework, there are assumptions made such that the production and decay of the Higgs boson can be factorized and parametrized as

$$\sigma_i \cdot \mathcal{BR}^f = \frac{\sigma_i(\vec{\kappa}) \cdot \Gamma^f(\vec{\kappa})}{\Gamma_H}, \quad (1.69)$$

where  $\Gamma_H$  is the total width of the Higgs boson and  $\Gamma^f$  is the partial width for Higgs boson decay to the final state  $f$ . Coupling modifiers,  $\vec{\kappa}$ , are introduced in order to test deviations in the couplings of the Higgs boson to other particles, and are defined as

$$\kappa_j^2 = \frac{\sigma_j}{\sigma_j^{\text{SM}}} \text{ or } \kappa_j^2 = \frac{\Gamma^j}{\Gamma_{\text{SM}}^j}, \quad (1.70)$$

where all  $\kappa_j = 1$  in the SM and  $j$  denotes the tested production or decay mode. Tree-level Higgs boson couplings, such as the  $H - Z$ ,  $H - W$ ,  $H - t$ ,  $H - b$ ,  $H - \tau$ , and  $H - \mu$ , are introduced as individual coupling modifiers. For those processes that occur at leading-order (LO) involving box or triangular loop diagrams, the loops are resolved in terms of the corresponding coupling modifiers, weighted by their individual contribution. Interference effects between the different diagrams provide sensitivity to the relative signs of the Higgs boson couplings to different particles. The coupling modifiers  $\kappa_c$  and  $\kappa_s$  are allowed to vary as function of other modifiers, provided that current LHC data are insensitive to these couplings. The constraint on  $\kappa_c$  will be introduced separately later. Other coupling modifiers  $\kappa_u$ ,  $\kappa_d$ , and  $\kappa_e$  are not included in combination given that their magnitudes are marginal.

There are two parametrization schemes. One is defined such that two addi-

---

<sup>3</sup>It was referred to as Interim framework in the cited reference.

tional effective coupling modifiers,  $\kappa_g$  and  $\kappa_\gamma$ , which describe the loop processes for ggF production and  $H \rightarrow \gamma\gamma$  decay, are introduced to account for the situation that BSM particles may be present in these loops. The other one is to resolve the ggF and  $H \rightarrow \gamma\gamma$  processes as function of remaining coupling modifiers. Fig. [1-11] shows the summary plots for the  $\kappa$ -framework model with the resolved loop scheme and the assumption  $\mathcal{BR}_{\text{BSM}} = 0$ . The points indicate the best fit values while the thick and thin horizontal bars show the  $1\sigma$  and  $2\sigma$  CL intervals, respectively. Without loss of generality, the value of  $\kappa_t$  is restricted to be positive. For this model, both positive and negative values of  $\kappa_W$ ,  $\kappa_Z$ , and  $\kappa_b$  are considered. The result shows that negative values of  $\kappa_W$  are disfavored by more than  $2\sigma$ . The interference between diagrams of the ZH production leads to the break of the degeneracy between signs, and indicates that a positive value of  $\kappa_Z$  is favored. A negative value of  $\kappa_b$  is preferred in this model, however, the difference between the best-fit point and the minimum in the positive region is small. Fig. [1-12] shows the summary plots with effective couplings scheme. In the left figure the constraint  $\mathcal{BR}_{\text{BSM}} = \mathcal{BR}_{\text{inv}} + \mathcal{BR}_{\text{undet}} = 0$  is imposed, and both positive and negative values of  $\kappa_W$  and  $\kappa_Z$  are considered. In the right figure a constraint  $|\kappa_V| \leq 1$ , where  $\kappa_V$  denotes  $\kappa_Z$  or  $\kappa_W$ , is imposed (same sign of  $\kappa_Z$  and  $\kappa_W$ ), while  $\mathcal{BR}_{\text{inv}} > 0$  and  $\mathcal{BR}_{\text{undet}} > 0$  are free parameters. The preferred sign of the  $\kappa_W$ , opposite to the first scheme, is negative. In Fig. [1-13] left plot shows the scan of the test statistic as a function of  $\mathcal{BR}_{\text{inv}}$ , and the right plot shows the 68% and 95% CL contours for  $\mathcal{BR}_{\text{inv}}$  vs.  $\mathcal{BR}_{\text{undet}}$ , indicating the 95% CL upper limits of  $\mathcal{BR}_{\text{inv}} < 0.22$  and  $\mathcal{BR}_{\text{undet}} < 0.36$ .

Another fit is performed using a phenomenological parameterization relating the masses of the fermions and vector bosons to the corresponding modifiers with two parameters,  $M$  and  $\epsilon$  [63, 64]. In this parametrization, the coupling modifiers,  $M$  and  $\epsilon$  are related as  $\kappa_F = \frac{v \cdot m_f^\epsilon}{M^{1+\epsilon}}$  for fermions and  $\kappa_V = \frac{v \cdot m_V^{2\epsilon}}{M^{1+2\epsilon}}$  for vector bosons, where  $v = 246.22 \text{ GeV}$  is the vacuum expectation value [65]. The SM expectation of  $\kappa = 1$ , corresponds to  $(M, \epsilon) = (v, 0)$ . The left plot in Fig. [1-14] shows the  $1\sigma$  and  $2\sigma$  CL regions in the  $(M, \epsilon)$  fit, and the results of the fit using the six modifiers are

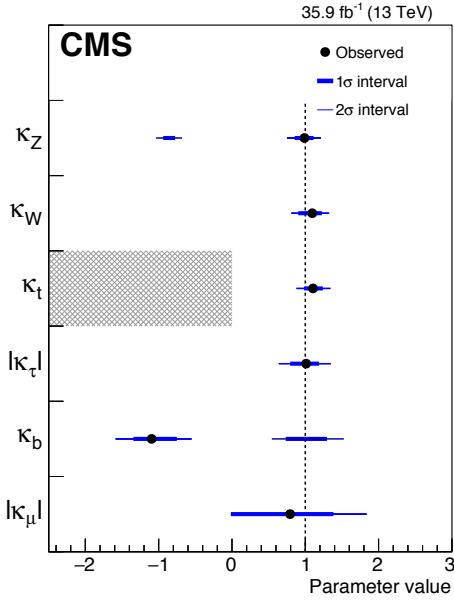


Figure 1-11: Summary for the  $\kappa$ -framework model with the resolved loop scheme [61].

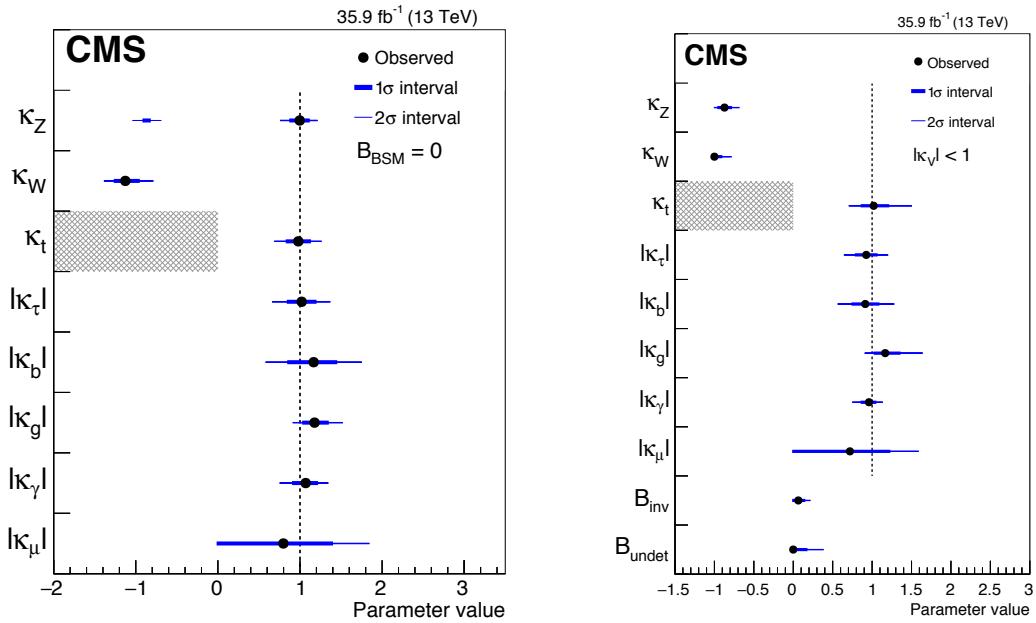


Figure 1-12: Summary for the  $\kappa$ -framework model with the effective couplings scheme [61].

plotted versus the particle masses on the right-hand side, as well as the result of the  $(M, \epsilon)$  fit. A "reduced" vector boson coupling  $\frac{\sqrt{\kappa_V \cdot m_V}}{v}$  is shown to represent the

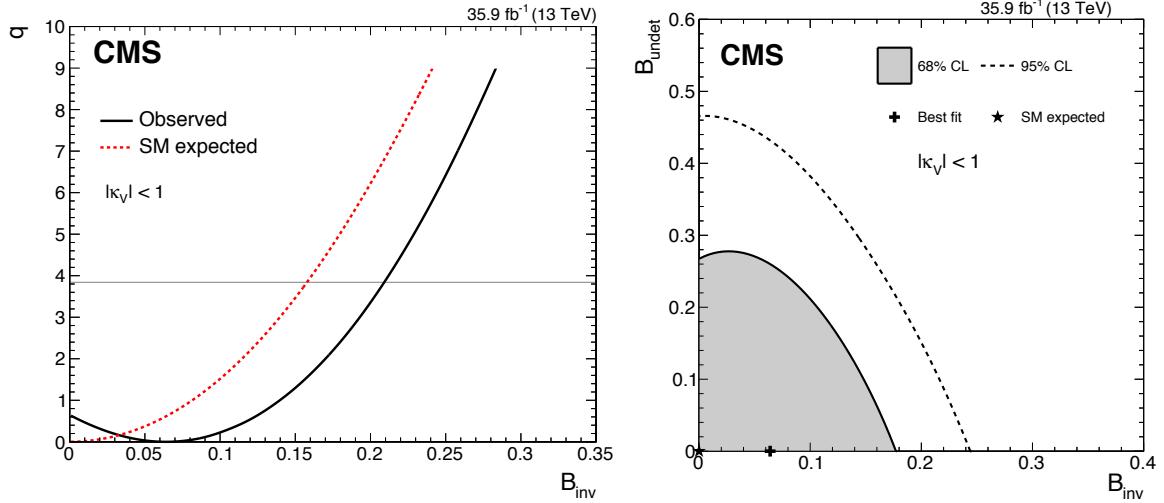


Figure 1-13: Scan of the test statistic as a function of  $\mathcal{B}\mathcal{R}_{\text{inv}}$  (left), and 68% and 95% CL regions for  $\mathcal{B}\mathcal{R}_{\text{inv}}$  vs.  $\mathcal{B}\mathcal{R}_{\text{undet}}$  [61].

couplings of the vector bosons in the same plot. As one can see, the couplings of these six particles to the Higgs boson are consistent within uncertainties with the SM predictions.

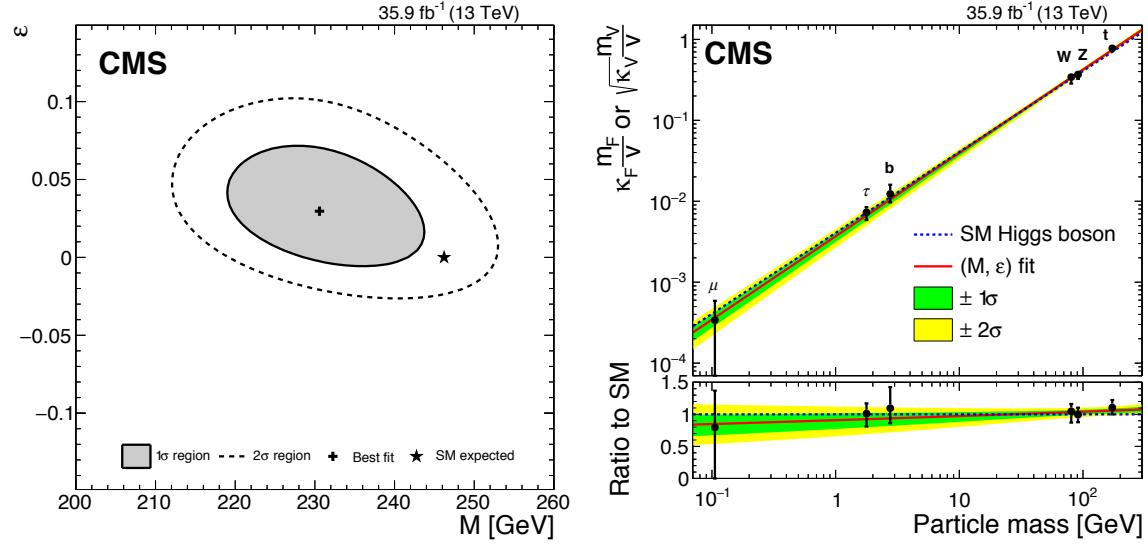


Figure 1-14: (Left) Likelihood scan in the  $M - \epsilon$  plane. The best fit point and the  $1\sigma$  and  $2\sigma$  CL regions are shown, along with the SM prediction. (Right) Result of the phenomenological  $(M, \epsilon)$  fit with the loop-resolved scheme of  $\kappa$ -framework model [61].

## The Higgs-charm coupling

As stated previously, a sensitive measurement of Higgs-charm coupling is not feasible in the environment of the LHC. There are still ways to constrain the size of the coupling. Since c- and b-jets share rough similarities, jets originating from charm quarks may be mistagged as b- jets. Hence, with the tagging efficiency of c- and b-jets, one can recast the existing analyses of  $H \rightarrow b\bar{b}$  to constrain the  $H \rightarrow c\bar{c}$  rate [51]. This results in a model-independent bound on the charm signal strength of  $\mu_c = 95^{+90}_{-95}$  with the results of the  $H \rightarrow b\bar{b}$  search in VH production from both ATLAS and CMS Collaborations. Both ATLAS and CMS Collaboration give a model-independent bound on the Higgs total width from the invariant-mass distribution of the  $H \rightarrow ZZ^*$  and  $H \rightarrow \gamma\gamma$  in the Run1 analyses. This bound on the total width can be used to constrain the Higgs-charm coupling by assuming the entire Higgs width is formed by  $Hc\bar{c}$ . With this method, the upper bounds at 95% CL with the CMS results is  $\kappa_c < 120$  and with the ATLAS results is  $\kappa_c < 150$ . A method that relies on the measurements of transverse momentum distributions of Higgs boson was proposed to determine the limit on the coupling modifier  $\kappa_c$  [66]. Fig. 1-15 shows the impact of the coupling modifier  $\kappa_c$  on the normalized  $p_T^H$  spectrum in inclusive Higgs production. This letter takes the  $p_T$  spectrum from the ATLAS combined measurement of  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^*$  decays with Run1  $\sqrt{s} = 8$  TeV data, and obtains the bounds on  $\kappa_c$  at 95% CL of  $\kappa_c \in [-16, 18]$ . The spectrum of the  $p_T^H$  at  $\sqrt{s} = 13$  TeV is expected to be slightly harder than that of  $\sqrt{s} = 8$  TeV, thus will enhance the sensitivity to  $\kappa_c$  at ongoing LHC runs as well as possible future hadron colliders at higher energies. The CMS Collaboration applies this method with the distributions from  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^*$  analyses using data collected in 2016 to set limit on the constrain of  $\kappa_c$  [67]. Fig. 1-16 shows the simultaneous fit results for  $\kappa_b$  and  $\kappa_c$ . On the left plot, 1 and  $2\sigma$  deviation contours for the combined ( $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^*$ ) fit to data and for  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^*$  separately, assuming coupling dependency of the branching fractions, while the right plot assumes freely floating branching fractions in the fit. The ob-

served (expected) constraints on  $\kappa_c$  are

$$-4.3 < \kappa_c < 4.3 \quad (-5.4 < \kappa_c < 5.3) \text{ (coupling dependent } \mathcal{BR}), \quad (1.71)$$

$$-18.0 < \kappa_c < 22.9 \quad (-15.7 < \kappa_c < 19.3) \text{ (freely floating } \mathcal{BR}). \quad (1.72)$$

If the branching fractions are fixed to the SM expectations, the expected constraint will be

$$-8.7 < \kappa_c < 10.6 \text{ (SM branching fractions).} \quad (1.73)$$

Rare exclusive decays of the Higgs boson to mesons in association with a photon can be used to explore these couplings. For example, the  $H \rightarrow J/\psi \gamma$  decay can probe the Higgs boson coupling to the charm quark [18]. This decay is the focus in the thesis, and will be discussed in the next section. Using Run1 results of the upper limit on  $H \rightarrow J/\psi \gamma$ , the bound at 95% CL is set at  $\kappa_c < 220$ .

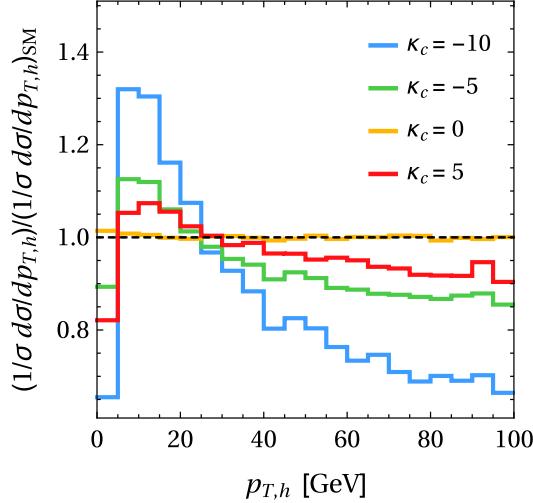


Figure 1-15: The normalized  $p_T^H$  spectrum of inclusive Higgs production at  $\sqrt{s} = 8$  TeV with different values of  $\kappa_c$  [18].

In some extensions to the SM, modified  $Hc\bar{c}$  couplings can arise [68]. For example, within the context of the effective field theory [69, 70, 71] the  $Hc\bar{c}$  coupling is modified in the presence of dimension-six operator, leading to only an enhancement of the coupling with respect to the SM at the cutoff scale  $\Lambda$  that can be as small

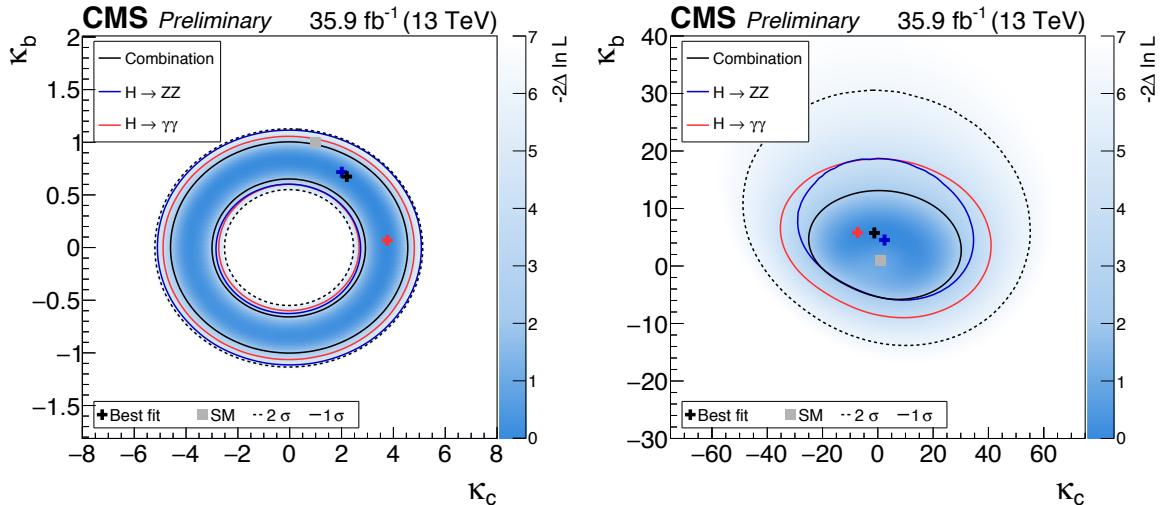


Figure 1-16: Simultaneous fit results for  $\kappa_b$  and  $\kappa_c$  [67].

as about 30 TeV, and leaving no other signature of new physics at the LHC. In the two Higgs doublet model with minimal flavor violation [72, 73], the  $Hc\bar{c}$  coupling can be significantly enhanced by breaking the flavor symmetry, while other couplings are not severely affected. The composite pseudo-Nambu-Goldstone boson model [74] parametrizes the coupling by the degree of compositeness and compositeness scale, which can be experimentally constrained by the direct search for the charm partner [75].

## 1.2 The rare decays $Z/H \rightarrow J/\psi \gamma$

### 1.2.1 Overview

The rare decay of  $H \rightarrow J/\psi \gamma$  is one of the proposed ways to probe the Higgs-charm coupling. The corresponding decay of the Z boson,  $Z \rightarrow J/\psi \gamma$ , can be used as an experimental benchmark for the  $H \rightarrow J/\psi \gamma$  search, given that the mass of the Z boson is not far from that of the Higgs boson, and to test various QCD factorization approaches that are being used in the estimation of branching fractions for hadronic radiative decays of bosons [76, 77, 78].

Both the Higgs and Z boson decays have contributions from direct and indirect processes. In the direct mechanism, Z and Higgs bosons couple to charm quarks, and charm quarks then hadronize to form  $J/\psi$  mesons. In the indirect mechanism, the Higgs and Z bosons decay through the quark and W boson loops to  $\gamma\gamma^*$ , and the  $\gamma^*$  then converts to a  $c\bar{c}$  resonant state. The Feynman diagrams for these decay modes are shown in Fig. 1-17. The widths of the decays are expected to be

$$\begin{aligned}\Gamma_{H \rightarrow J/\psi \gamma} &= \frac{1}{8\pi} \frac{m_H - m_{J/\psi}}{m_H} |\mathcal{A}_{\text{direct}} + \mathcal{A}_{\text{indirect}}|^2 \\ &= \left[ (11.71 \pm 0.17) - [(0.659^{+0.085}_{-0.085}) + i(0.073^{+0.035}_{-0.035})] \kappa_c \right] \times 10^{-10} \text{ GeV} \\ &= 1.221^{+0.042}_{-0.041} \times 10^{-8} \text{ GeV},\end{aligned}\tag{1.74}$$

$$\Gamma_{Z \rightarrow J/\psi \gamma} = \frac{m_Z^3}{96\pi m_{J/\psi}^2} |\mathcal{A}_{\text{direct}} + \mathcal{A}_{\text{indirect}}|^2 = 2.236^{+0.377}_{-0.344} \times 10^{-7} \text{ GeV},\tag{1.75}$$

where in Eq. 1.74 the equality and numerical results are taken from Ref. [18, 79], and those in Eq. 1.75 are from Ref. [80]. In these theoretical calculations, a framework of the nonrelativistic QCD (NRQCD) factorization [81] is used, where the nonperturbative effects are parametrized in terms of the quarkonium light-cone distribution amplitudes (LCDAs) [82, 83]. These computations will not be discussed in detail here. With the total widths of both the Higgs  $\Gamma_H = 4.20 \text{ MeV}$  and Z boson  $\Gamma_Z = 2.4952 \text{ GeV}$  and  $\kappa_c = 1$  in the SM, the branching fractions of both decays are then:

$$\mathcal{B}_{\text{SM}}(H \rightarrow J/\psi \gamma) = (3.0^{+0.2}_{-0.2}) \times 10^{-6}.\tag{1.76}$$

$$\mathcal{B}_{\text{SM}}(Z \rightarrow J/\psi \gamma) = (9.0^{+1.5}_{-1.4}) \times 10^{-8},\tag{1.77}$$

The direct and indirect amplitudes interfere destructively in both decays. In the Higgs decay, the contribution from the indirect process is larger. Including only direct process in the calculation leads to a branching fraction of  $5.28 \times 10^{-8}$ , while taking only indirect diagrams into account results in a branching fraction of  $3.25 \times 10^{-6}$ .

The branching fraction of the Z decay, compared to the Higgs decay, is smaller by 1-2 orders of magnitude. This results from the suppression of the indirect amplitude, which is less than 1% of the magnitude of direct amplitude. One qualitative explanation uses the Landau-Yang theorem [84], which states that the Z boson does not decay to two on-shell photon. This requires that the indirect amplitude tends to zero in the limit  $m_{J/\psi} \rightarrow 0$ .

With the branching fractions shown above, one obtains

$$\begin{aligned}\sigma(pp \rightarrow H) \times \mathcal{B}_{\text{SM}}(H \rightarrow J/\psi \gamma \rightarrow \mu\mu\gamma) = \\ 55 \text{ pb} \times 3.0 \times 10^{-6} \times 0.059 = 9.8 \times 10^{-3} \text{ fb},\end{aligned}\tag{1.78}$$

$$\begin{aligned}\sigma(pp \rightarrow Z) \times \mathcal{B}_{\text{SM}}(Z \rightarrow J/\psi \gamma \rightarrow \mu\mu\gamma) = \\ 5.7 \times 10^4 \text{ pb} \times 9.0 \times 10^{-8} \times 0.059 = 3.0 \times 10^{-1} \text{ fb}.\end{aligned}\tag{1.79}$$

where the cross-section of the Higgs boson are summed over the ggF, VBF, VH, and ttH productions, and taken from Ref. [48]. The cross-section of the Z boson are calculated using FEWZ 3.1.b2 program [85].

Deviations from the SM predictions for the couplings can affect the interference terms and may result in changes in the branching fractions. For example, the shift in the branching fraction for  $H \rightarrow J/\psi \gamma$  can be more than 100% if the  $Hc\bar{c}$  coupling deviates from its SM value by more than a factor of 2, as shown in Fig. 1-18. Measurements of the direct decay of  $H \rightarrow c\bar{c}$  leave the overall signs of the couplings undetermined. This ambiguity can be resolved by the interference terms in  $H \rightarrow J/\psi \gamma$ , providing us with additional information about the Higgs properties.

### 1.2.2 Features of the decays

Due to the relatively heavy Z and Higgs bosons, the  $J/\psi$  and  $\gamma$  from their decays will have high transverse momenta  $p_T$  and energy  $E_T$  (boosted). The high- $E_T$  photon will be produced back-to-back to the  $J/\psi$  particle, and hence can be distinguished from backgrounds easily and be identified as an isolated photon. Since

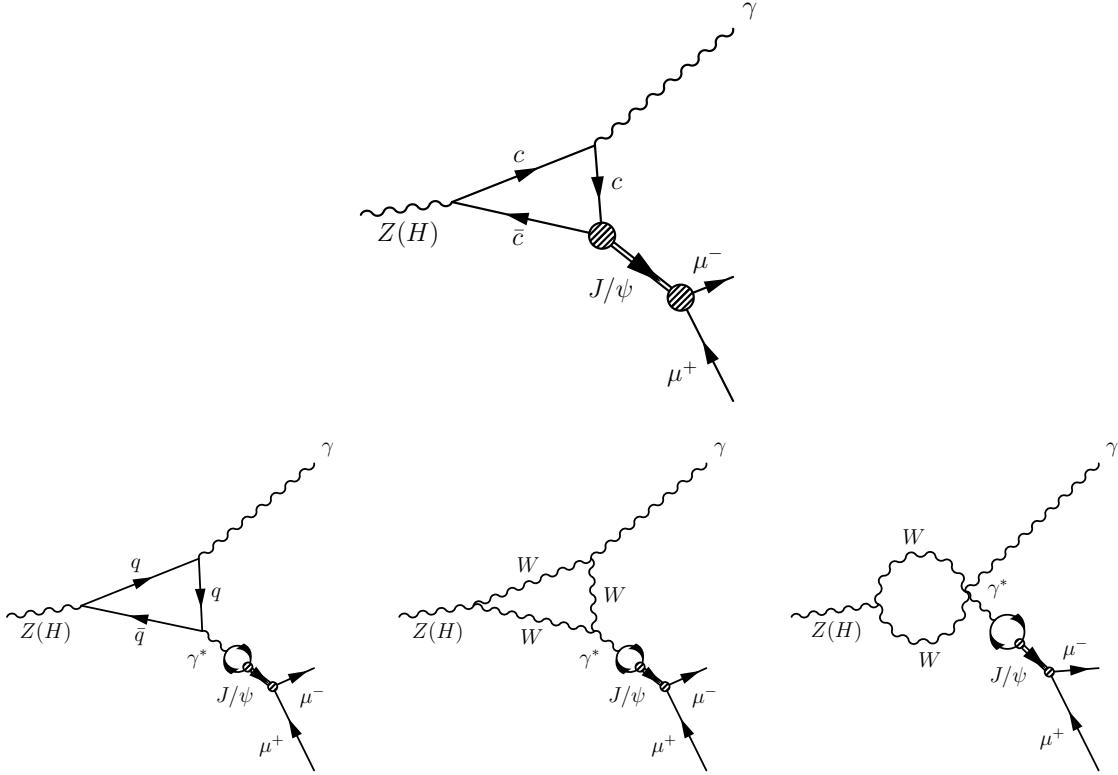


Figure 1-17: Feynman diagrams for  $Z(H) \rightarrow J/\psi \gamma$  decay. The top diagram shows the direct process and the remaining diagrams show the indirect processes.

the  $J/\psi$  meson from  $Z$  (Higgs) boson decay is boosted, the  $p_T$  of the two muons from its decay are anti-correlated. Further, these two muons are very close to each other spatially. Therefore, dedicated strategies for trigger algorithms and both off-line reconstruction are needed.

The photon should be well separated from each muon. This event signature can be utilized to design kinematic requirements such as the angular separation  $\Delta R^4$  to reject backgrounds.

Fig. 1-19 shows the distributions of key variables at the generator level. All the distributions shown in the figure are normalized to unity. One can see that, the momenta of muons cover a wide range: the transverse momentum  $p_T$  of trailing muon<sup>5</sup> can be less than 10 GeV, while that of leading muon can be greater than 40 (60) GeV in the  $Z$  (Higgs) boson decay respectively. The photon can have high

---

<sup>4</sup>The coordinate system will be introduced in the next chapter.

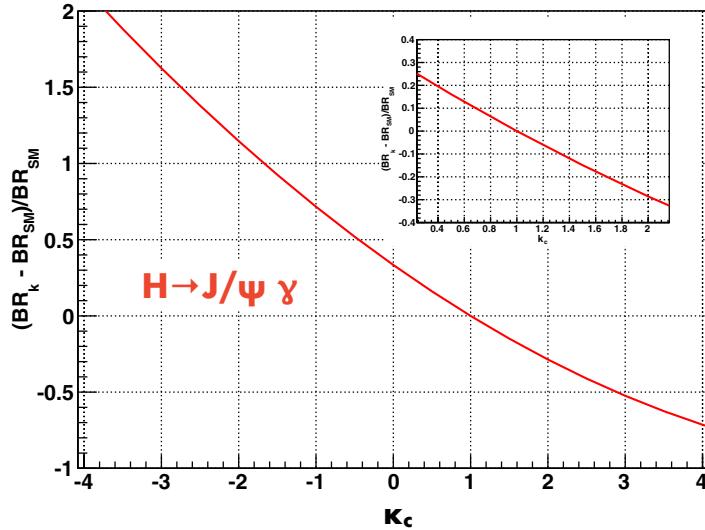


Figure 1-18: The relative deviations in the branching fraction for  $H \rightarrow J/\psi \gamma$  as function of  $\kappa_c$  [18].

transverse energy. The muons and the photon distribute mostly in the central region. The high- $E_T$  photon is back-to-back to the dimuon system, while the two muons are close to each other spatially.

### 1.2.3 Previous results from the ATLAS and CMS Collaborations

The  $Z \rightarrow J/\psi \gamma$  decay was searched for by the ATLAS Collaboration using the data set collected at  $\sqrt{s} = 8$  TeV [86]. An observed (expected) upper limit on the branching fraction of  $2.6 (2.0^{+1.0}_{-0.6}) \times 10^{-6}$  was reported. Searches for the  $H \rightarrow J/\psi \gamma$  decay have been performed by the ATLAS and CMS Collaborations using the data set collected at  $\sqrt{s} = 8$  TeV respectively [86, 87]. Observed (expected) limits on the branching fraction were  $1.5 (1.2^{+0.6}_{-0.3}) \times 10^{-3}$  from the ATLAS Collaboration and  $1.5 (1.6^{+0.8}_{-0.8}) \times 10^{-3}$  from the CMS Collaborations. Fig. 1-20 shows the three-body invariant mass  $m_{\mu\mu\gamma}$  and  $p_T^{\mu\mu\gamma}$  distributions, along with the signal-plus-background fit to observed data collected at  $\sqrt{s} = 8$  TeV from ATLAS results. Fig. 1-21 shows the non-resonant background fit to the  $m_{\mu\mu\gamma}$  distributions observed in data col-

---

<sup>5</sup>In the analysis, two muons will be selected in the final state. The one with higher  $p_T$  is referred to as leading muon, and the other one is then referred to as trailing or subleading muon.

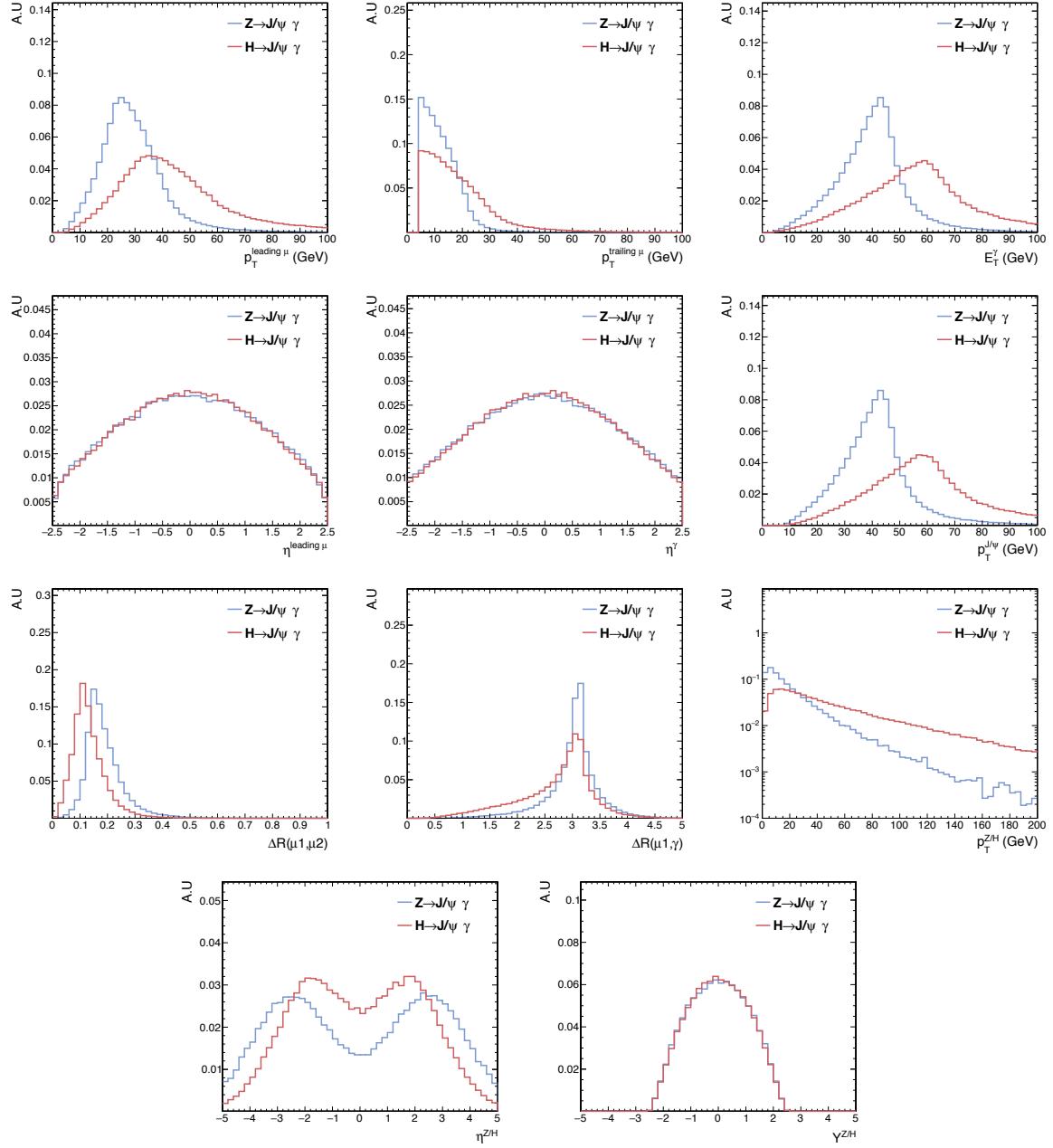


Figure 1-19: The distributions of key variables at generator level in both the  $Z$  and Higgs boson decays:  $p_T$  and  $E_T$  of the leading, trailing muon and the photon, pseudorapidity  $\eta$  of the leading muon and the photon,  $p_T$  of the  $J/\psi$  meson , angular separation  $\Delta R$  between muons,  $\Delta R$  between the leading muon and the photon,  $p_T$  of the  $Z$  and Higgs boson,  $\eta$  of the  $Z$  and Higgs boson, and the rapidity  $Y$  of the  $Z$  and Higgs boson

lected at  $\sqrt{s} = 8$  TeV with CMS search. Recently, ATLAS provides results with data collected in 2016 for both decays. An observed (expected) upper limit on the branching fraction of  $Z \rightarrow J/\psi \gamma$  decay is set at  $2.3 (1.1^{+0.5}_{-0.3}) \times 10^{-6}$ , and of the  $H \rightarrow J/\psi \gamma$  is at  $3.5 (3.0^{+1.4}_{-0.8}) \times 10^{-4}$  [88]. Fig. 1-22 shows the recent results from ATLAS Collaboration.

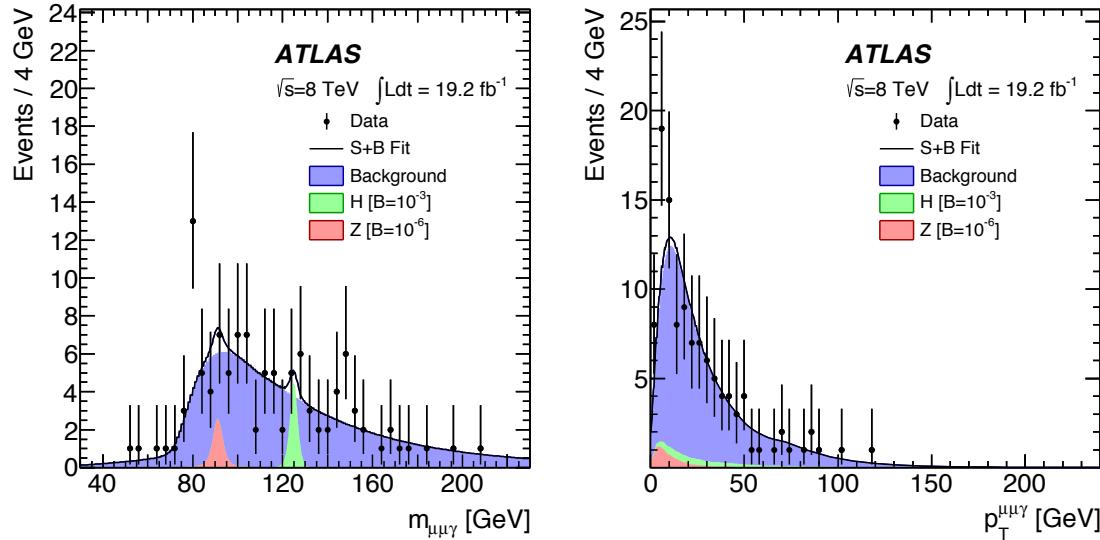


Figure 1-20: Previous result of  $Z$  ( $H$ )  $\rightarrow J/\psi \gamma$  decay search from the ATLAS Collaboration. The three-body invariant mass  $m_{\mu\mu\gamma}$  and  $p_T^{\mu\mu\gamma}$  distributions, along with the results of signal-plus-background fit to observed data collected at  $\sqrt{s} = 8$  TeV [86].

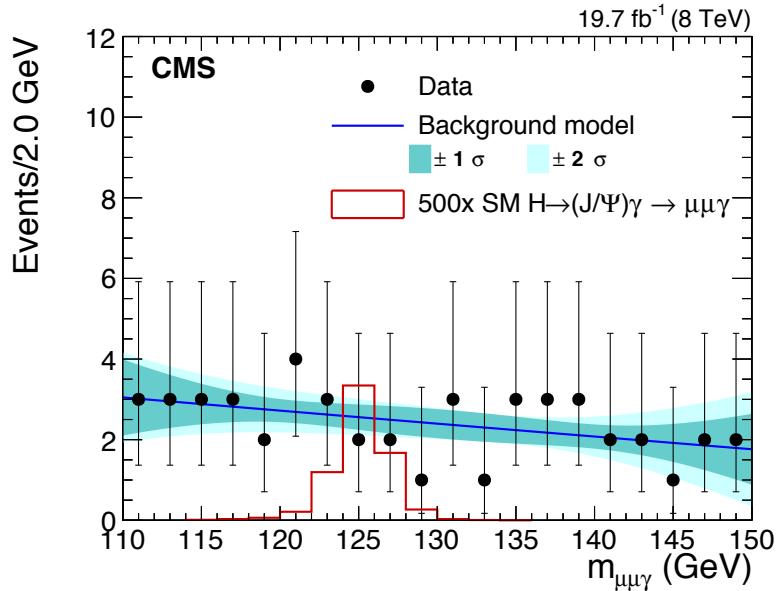


Figure 1-21: Previous result of  $H \rightarrow J/\psi \gamma$  decay search from the CMS Collaboration. Non-resonant background fit to the  $m_{\mu\mu\gamma}$  distributions observed in data collected at  $\sqrt{s} = 8$  TeV [86].

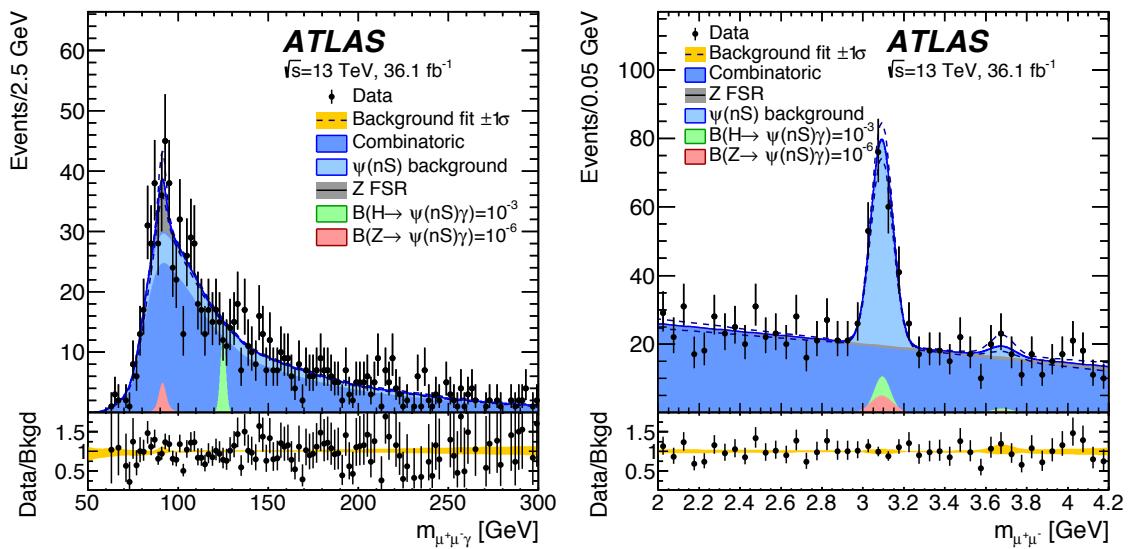


Figure 1-22: Result of  $Z$  ( $H$ )  $\rightarrow J/\psi \gamma$  decay search from the ATLAS Collaboration with data collected at  $\sqrt{s} = 13$  TeV in 2016 [88].

# Chapter 2

## Experimental apparatus

In this chapter, the overview of the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) will be introduced. The object reconstruction will be summarized in the last section, as it is closely related to the detectors.

### 2.1 Large Hadron Collider

The LHC is so far the largest particle accelerator that human have ever built, and currently hosted by the Europe Organization of Nuclear Research (CERN). It possesses a 26.7 km of ring and is placed more than 100 m deep beneath Geneva and France. Such large circumference makes it able to provide high energy collisions, and enables us to examine the validity of the SM and explore the physics such as the existence of the Higgs boson, supersymmetry particles (SUSY), extra-dimension, or even dark matter (DM). The ring consists of two individual and parallel beam pipes, in which protons (or heavy-ions) circulate in opposite directions.

The protons are grouped together into 2808 bunches, and each bunch contains  $1.15 \times 10^{11}$  protons. The time interval between two bunches is 25 ns, corresponding to a collision rate of 40 MHz. A series of machines then successively accelerate and bring proton beams to higher energy. Each beam is accelerated up to an energy of 6.5 TeV when it finally arrives at the LHC beam line.

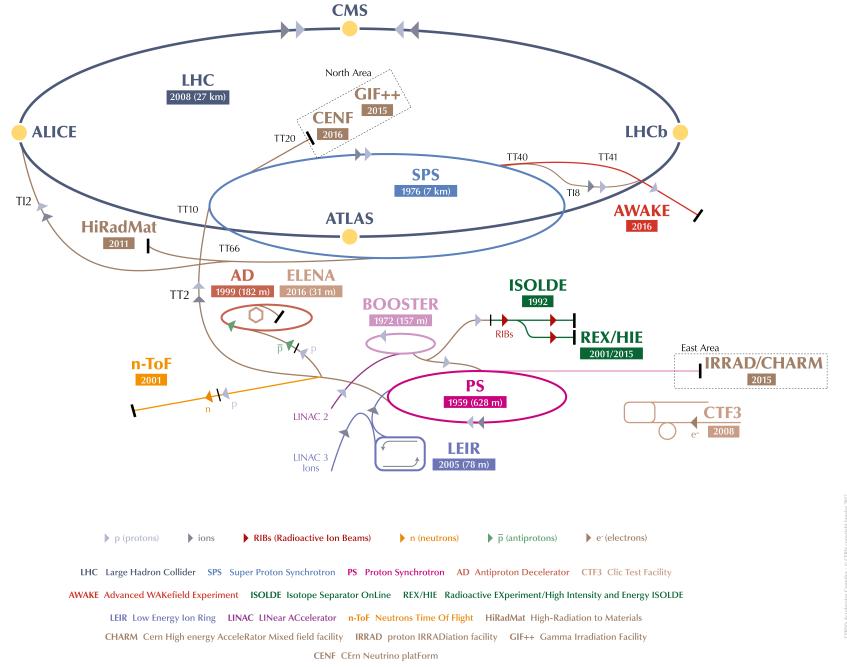


Figure 2-1: The CERN accelerator complex. The protons are accelerated from the LINAC2, PSB, PS, SPS, and finally to LHC [89].

Fig. 2-1 shows the whole system of the CERN complex [89].

An important quantity in the collider physics is the luminosity  $\mathcal{L}$ . The instantaneous luminosity is defined as:

$$\frac{dN}{dt} = \sigma_{\text{event}} \frac{d\mathcal{L}}{dt} \quad (2.1)$$

The  $\frac{dN}{dt}$  is the event production rate, and  $\sigma_{\text{event}}$  is the interaction cross section. The integrated luminosity  $\mathcal{L}_{\text{Tot}}$  is the integral of the instantaneous luminosity over a period of time and a measure of the amount of data. Fig. 2-2 shows the integrated luminosity that CMS recorded in each data-taking year [90].

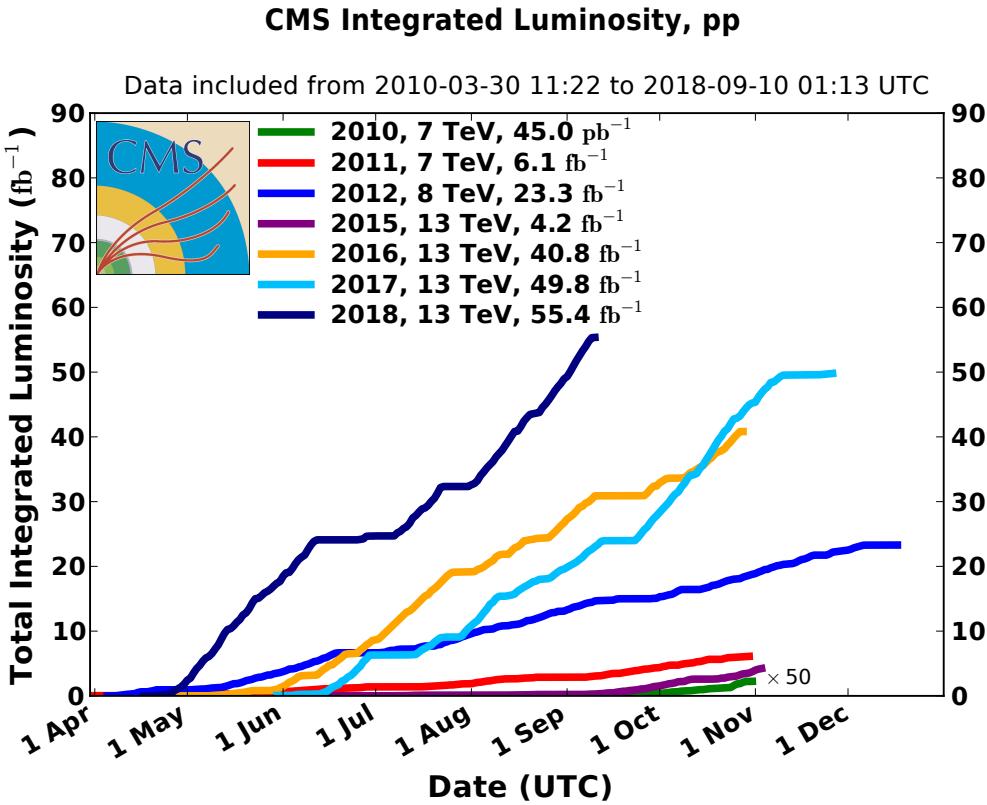


Figure 2-2: Cumulative luminosity versus day delivered to CMS during stable beams for pp collisions at nominal center-of-mass energy. This is shown for data-taking in 2010 (green), 2011 (red), 2012 (blue), 2015 (purple), 2016 (orange), 2017 (light blue), and 2018 (deep blue) [90].

## 2.2 Compact Muon Solenoid

Compact Muon Solenoid is one of the general purpose detectors located at the LHC ring. The central feature of the CMS apparatus is a superconducting solenoid of 13 m in length and 6 m in internal diameter, providing an axial magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity ( $\eta$ ) coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid.

The adopted coordinate system, as shown in Fig. [2-3], has the origin at the nominal collision point inside CMS detector, where the y-axis pointing vertically upward, the x-axis pointing radially inward toward the center of the LHC, and the z-axis pointing along the beam direction. The azimuthal angle  $\phi$  is measured from the x-axis in the x-y plane, while the polar angle  $\theta$  is measured from the z-axis. Rapidity,  $Y$ , is defined as  $Y \equiv \frac{1}{2} \ln \left( \frac{E+p_z c}{E-p_z c} \right)$ , where  $E$  is the energy of the particle and  $p_z$  is the momentum in the z direction. This Lorentz invariant quantity indicates the angle between the x-y plane and the direction of the measured particle. For the highly relativistic particles, the other quantity called pseudorapidity, defined as  $\eta = -\ln \tan(\theta/2)$ , is used, where  $\theta$  is the angle between the particle trajectory and the z-axis (beam pipe).

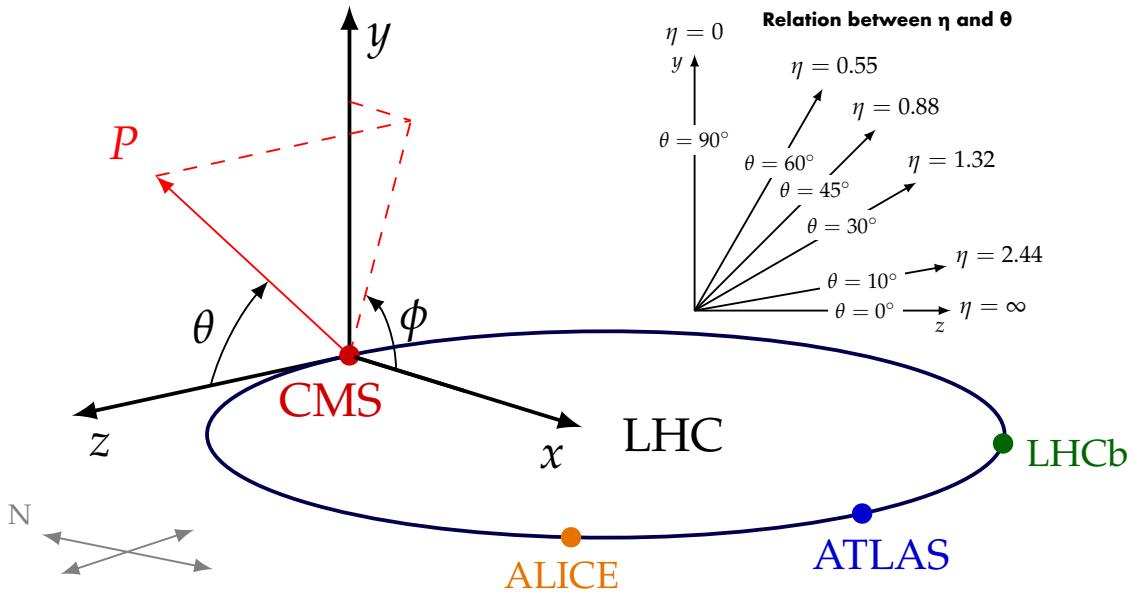


Figure 2-3: The adopted coordinate system in CMS.

The momentum and energy transverse to the beam direction, denoted by  $p_T$  and  $E_T$ , respectively, are computed from the x and y components.

Fig. [2-4] shows a global view of the CMS detector [91]. Brief description of each sub-detector is summarized as follows.

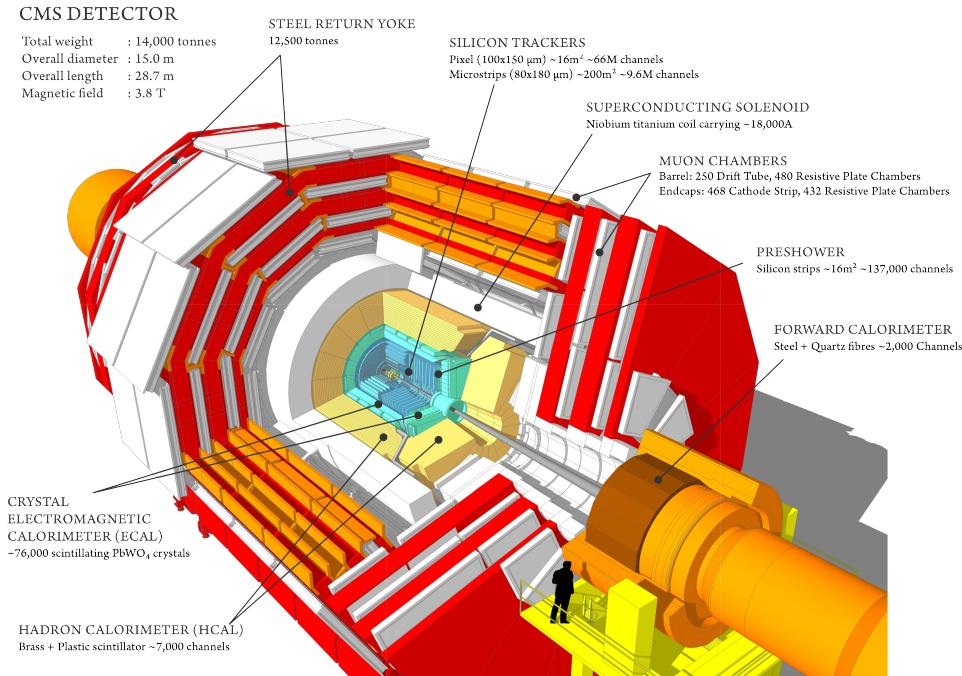


Figure 2-4: A cutaway view of the CMS detector [91].

## Superconduction magnet

The superconducting solenoid magnet, formed by a cylindrical coil of superconducting fibers, was originally designed to provide a magnetic field of 4 Tesla (T), while in the actual operation it produces a 3.8 T of field. Such large bending power enables us to measure the momentum of high energy charged particles precisely. The magnetic field is confined to the volume of the detector. This is done by the steel yoke, consisting of five layers for barrel part and three layers for each endcap.

## Silicon tracker

The CMS tracker is composed of two systems: a pixel detector (for a total of 1440 silicon pixel modules) with three barrel layers, and a silicon strip tracker (for a total of 15148 silicon strip modules) with ten barrel detection layers, four layers of tracker inner barrel (TIB) and six layers of tracker outer barrel (TOB), extending outwards. Each system is completed by endcaps, which consist of two disks in the

pixel detector, three tracker inner disks (TID) and nine disks of tracker endcaps (TEC) in the strip tracker on each side of the barrel. The acceptance of the whole tracker system extends up to a  $|\eta| < 2.5$ . Fig. 2-5 shows the schematic view of the silicon tracker in the r-z plane. The upper plot is the cross section through the tracker, and the lower one is one quarter of the tracker, where the paths of the laser rays (R), the alignment tubes (A) and the beam splitters (B) of the laser alignment system are illustrated.

For non-isolated particles with transverse momentum,  $p_T$ , between 1 and 10 GeV and  $|\eta| < 1.4$ , the track resolutions are typically 1.5% in  $p_T$  and 25–90 (45–150)  $\mu\text{m}$  in the transverse (longitudinal) impact parameter [92].

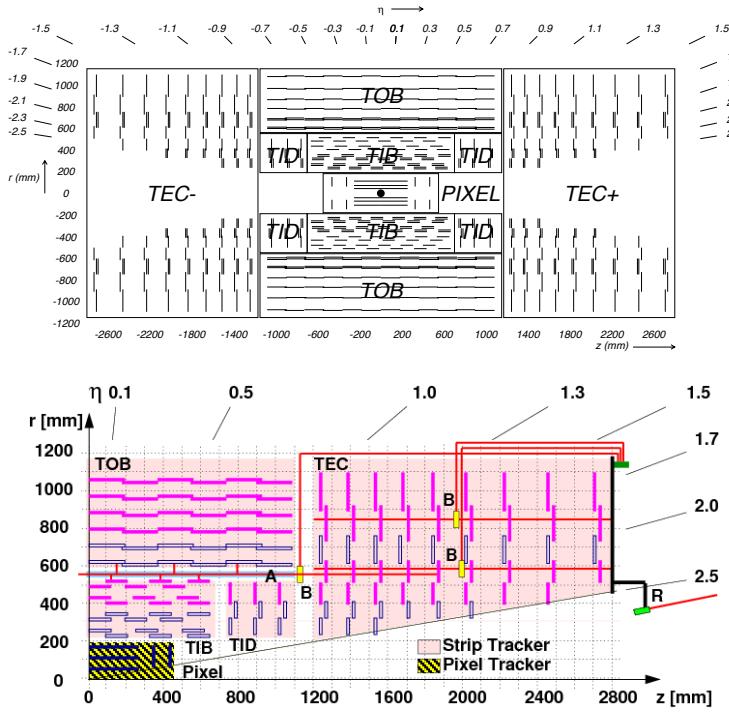


Figure 2-5: Schematic view of one quarter of the silicon tracker in the r-z plane [93].

## Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is a homogeneous calorimeter made of 61200 lead tungstate ( $\text{PbWO}_4$ ) crystals in the barrel part ( $0 < |\eta| < 1.48$ ) and 7324

crystals in each of endcaps ( $1.48 < |\eta| < 3.0$ ). The high density ( $8.28 \text{ g/cm}^3$ ) and short radiation length  $X_0^{\text{I}}$ <sup>1</sup> (0.89 cm) of the crystal result in a compact calorimeter with fast response, fine granularity, and strong resistance to the radiation. A sampling calorimeter, preshower detector (ES), is placed in front of the endcap crystals and covers the range of  $1.65 < |\eta| < 2.6$ . It consists of two planes of silicon sensors interleaved with a total of  $3X_0$  of lead. The main task of this detector is to help on distinguishing between single high-energy photons and the close pairs of low-energy photons, usually from the decay of neutral pion. The ES also improves the ability of identifying electrons against minimum ionizing particles and the position determination of electrons and photons. Fig. [2-6] shows layout of the CMS ECAL.

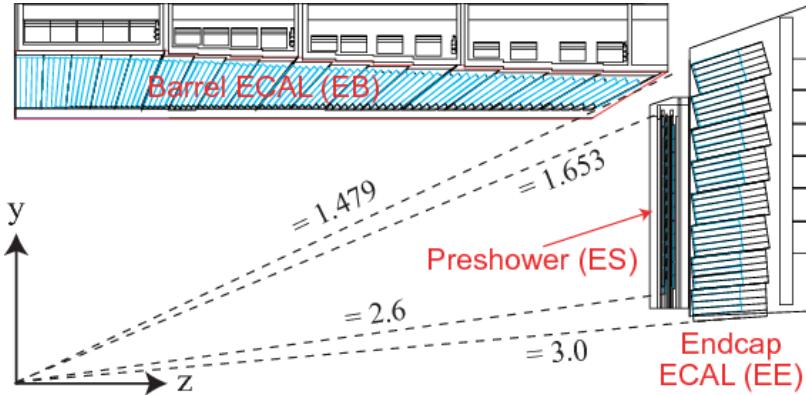


Figure 2-6: Layout of the ECAL [94].

## Hadron calorimeter

The hadron calorimeters (HCAL), a sampling calorimeter, measures the energy of hadron jets and provides indirect measurement of missing transverse energy, which can be neutrinos or exotic particles that do not interact with matters. Fig. [2-7] shows the longitudinal view of the CMS detector with the dashed lines representing fixed  $\eta$  values. The HCAL consists four parts: the HCAL barrel (HB), the

---

<sup>1</sup>One radiation length of a given material is defined as the distance after which the electron loses  $1/e$  of its original energy.

HCAL endcap (HE), the HCAL outer (HO), and the HCAL forward (HF). The HB, covering the range of  $|\eta| < 1.3$ , is placed radially between the outer extent of the ECAL and the inner extent of the magnet coil. The HO sits outside the solenoid complementing the barrel part, and ensure the leakage of the energy not detected by HB to be minimal. The HE covers the range of  $1.3 < |\eta| < 3.0$ , a region containing about 34% of the particles produced in the final state. The HF is place at the range of  $|\eta| > 3.0$ , where much higher energy will be deposited compared to other sub-detectors.

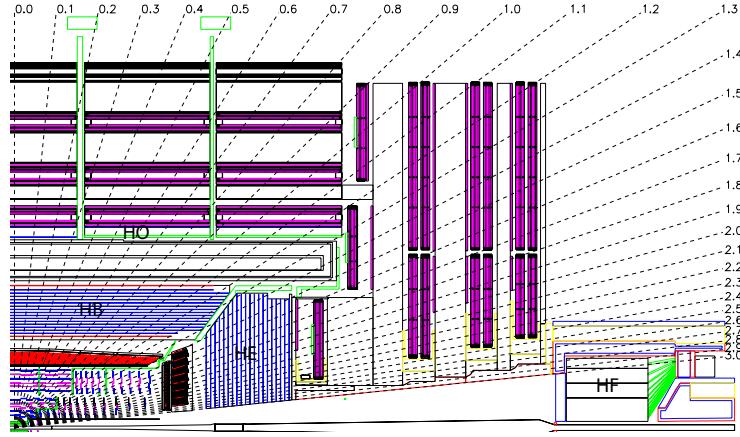


Figure 2-7: Longitudinal view of the CMS detector showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters [93].

## Muon system

The muon system is located outside the solenoid and covers the range  $|\eta| < 2.4$ . It is composed of three types of gaseous detectors, drift tubes (DTs), cathode strip chambers (CSCs), and resistive plate chambers (RPCs), sandwiched among the layers of the steel yoke. The DTs are segmented into drift cells; the position of the muon is determined by measuring the drift time to an anode wire of a cell with a shaped electric field. The CSCs operate as standard multi-wire proportional counters but with a finely segmented cathode strip readout, which yields an accurate

measurement of the position of the bending plane ( $R - \phi$ ) coordinate at which the muon crosses the gas volume. The DT and CSC chambers are located in the regions  $|\eta| < 1.2$  and  $0.9 < |\eta| < 2.4$ , respectively, and are complemented by RPCs in the range  $|\eta| < 1.9$ . Three regions are defined and referred to as the barrel ( $|\eta| < 0.9$ ), overlap ( $0.9 < |\eta| < 1.2$ ), and endcap ( $1.2 < |\eta| < 2.4$ ) regions [95]. Fig. 2-8 shows the arrangement of the muon system.

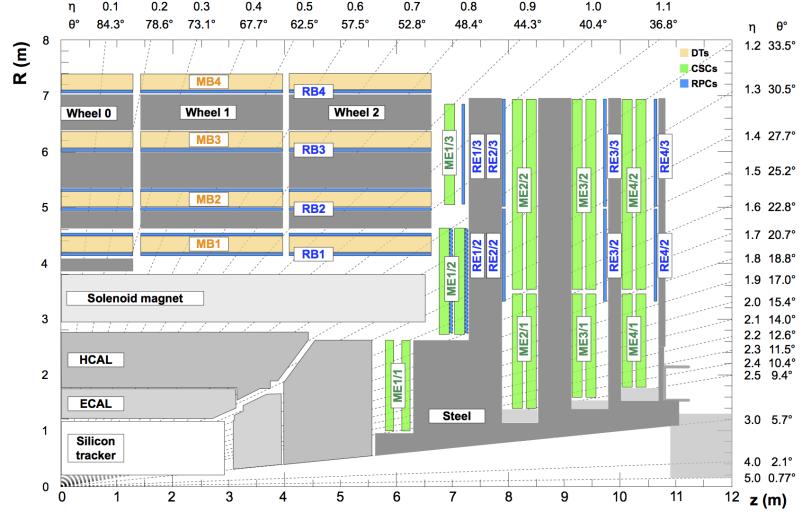


Figure 2-8: An R-z cross section of the muon station. The drift tube stations (DTs) are labeled MB ("Muon Barrel") and the cathode strip chambers (CSCs) are labeled ME ("Muon Endcap"). Resistive plate chambers (RPCs) are mounted in both the barrel and endcaps of CMS, where they are labeled RB and RE, respectively [95].

## Trigger and data acquisition system

The LHC provides pp and heavy-ion collisions at high interaction rate. This corresponds to an enormous amount of data that are currently not able to be completely stored. Furthermore, most of these interactions would be low-energy glancing collisions, rather than energetic and head-on interactions where processes of interest may occur. The trigger system is designed to reduce the rate and to start the physics event selection process. Fig. 2-9 shows the schematic diagram of the trigger architecture and data acquisition system. The level-1 trigger (L1) consists of custom-designed and programmable electronics. Information from muon system

(including DTs, CSCs, and RPCs), ECAL, HCAL, and HF is used to reconstruct candidate trigger objects, and these quantities are combined and forwarded to the Global Trigger (GT), which calculates the trigger decision and sends out the signal if it is "L1 Accept (L1S)". This step reduces the data rate from the 40 MHz of the LHC bunch crossing rate down to a maximum of 100 kHz. In case of a positive L1 decision all data for the corresponding bunch crossing time is read out from the CMS detector and transferred to the HLT, which consists of a software system implemented in a filter farm. The high level trigger algorithm (HLT) performs a full reconstruction of events using a faster version of offline software and writes data out to permanent storage at a typical rate of several hundred Hz.

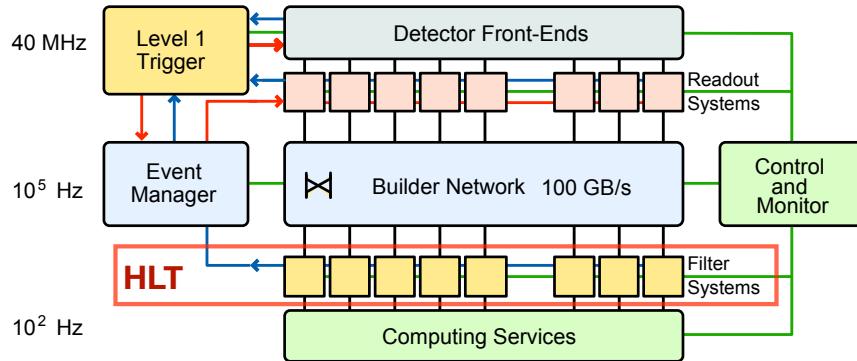


Figure 2-9: A schematic diagram of the trigger system [93].

## 2.3 Object reconstruction

### 2.3.1 Particle-Flow algorithm

The core concept of this algorithm is to optimally correlate tracks or clusters from all sub-detectors and combines the information to reconstruct final state particles. In order to have PF algorithm as efficient as possible, the magnetic field should be strong enough to maximize the separation between charged and neutral particles, and the detector should have fine spatial granularity layers that can distinguish

merged particles, especially those in jets. The CMS meets all of these advantages to use PF reconstruction as a global event description.

The ECAL energy clusters without being associated to extrapolated tracks from tracker are reconstructed as photons. Electrons are reconstructed by tracks in the tracker system with associated energy deposits in the ECAL. The bremsstrahlung emission and energy losses when traveling through tracker materials are properly accounted for. Muons tracks can be reconstructed in tracker, in muon system, or the combination of the two. Charged hadrons are reconstructed by the tracks not identified as electrons or muons with energy cluster in HCAL. The energy clusters that correspond to excesses of energy with respect to charged hadrons and not linked to charged particle trajectories are reconstructed as neutral hadrons.

In this analysis, photon and muons are selected as final states particles. Hence, their reconstructions are described in detail in the following paragraphs.

## Photon reconstruction

Photons are reconstructed from energy deposits in the ECAL. The algorithms, without any hypothesis as to whether the particle from the interaction point is a photon or an electron, identify the energy clusters and constrains them to the expected sizes and shapes, based on the study of simulation. The measurements of photon trigger, reconstruction, and identification efficiencies and energy scale and resolution can therefore utilize the electrons from  $Z \rightarrow e^+e^-$  events with a well defined invariant mass.

The clustering algorithms are used to sum over all energy deposits in crystals in the same electromagnetic shower. A basic cluster (BC) is chosen to be the local maximum among the energy deposits. Several BCs are combined to construct a supercluster (SC). The radiated energies, such as the conversions of photons or bremsstrahlung from electrons, are corrected and recovered for their corresponding SC. The energy of the photon is determined by summing the amplitude in channels  $A_i$  over the crystals  $i$  in the supercluster where the photon leaves energy,

corrected by the intercalibration  $c_i$  and light monitoring  $S_i(t)$  constants. The procedure can be summarized in a formula,

$$E_{fl} = \left[ \sum_i \left( S_i(t) \times c_i \times A_i \right) \times G(\eta) + E_{ES} \right] \times F_{fl} \quad (2.2)$$

where  $G(\eta)$  is the ADC to GeV factor.

Independent methods are used to calculate the intercalibration constants (ICs), and the combined factor is obtained from the mean of the individual IC at a fixed value of  $\eta$ , weighted by their respective precisions. A light monitoring system, consisting of a system of lasers that inject light to crystals, is used to monitor the time dependence of response in the ECAL resulting from the decreases in crystal transparency in radiation exposure. The difference between input and read laser amplitudes are then used to calculate correction factors  $S_i(t)$ . For photon in the region  $1.65 < |\eta| < 2.6$  the energy deposits in the preshower  $E_{ES}$  are also accounted for. The cluster corrections  $F_{fl}$  is applied to take the variation of shower containment in the clustered crystals and the shower losses of photons that convert before reaching the calorimeter into account. The correction factors are computed with a multivariate regression technique that estimates the energy of the photon and its uncertainty simultaneously. The resolution of photon energy is optimized after applying the factors.

The ECAL energy resolution was measured in beam tests, and found to be:

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E \text{ (GeV)}}} \oplus \frac{12\%}{E \text{ (GeV)}} \oplus 0.3\%. \quad (2.3)$$

The first contribution is the stochastic term, which represents the event to event fluctuations in the lateral shower containment. The second term comes from the electronic noise. The last one is the constant term, characterizing the resolution at high energy region.

The energy scale and resolution is further measured and calibrated using a high purity  $Z \rightarrow e^+e^-$  samples with 2% of background contamination, estimated from

simulation. An unbinned maximum likelihood fit to the invariant mass distribution is performed. A Breit-Wigner distribution convolved with a Crystal Ball (CB) function [?] is used.

$$CB(m - \Delta m) = \begin{cases} e^{-\frac{1}{2}(\frac{m-\Delta m}{\sigma_{CB}})^2}, & \frac{m-\Delta m}{\sigma_{CB}} > \alpha \\ (\frac{\gamma}{\alpha})^\gamma \cdot e^{-\frac{\alpha^2}{2}} \cdot \left( \frac{\gamma}{\alpha} - \alpha - \frac{m-\Delta m}{\sigma_{CB}} \right)^{-\gamma}, & \frac{m-\Delta m}{\sigma_{CB}} < \alpha \end{cases} \quad (2.4)$$

where the parameter  $\Delta m$  quantifies the displacement of the peak with respect to the nominal Z boson mass;  $\sigma_{CB}$  is the width of the Gaussian component of the CB function and serves as a measure of the energy resolution; the parameters  $\alpha$  and  $\gamma$  describe the tail part of CB, accounting for electrons of which energy is not fully retained after the clustering algorithms. In this step, the ADC to GeV factor  $G(\eta)$  is also adjusted and determined such that the peak value from the fit to  $Z \rightarrow e^+e^-$  distribution agrees with that of the simulation, independently for the barrel and endcap. There are still unknown effects that make the resolution of  $Z \rightarrow e^+e^-$  distribution in data worse than that in simulation. These residual discrepancies are corrected by adding a Gaussian smearing, where the parameters of smearing function are determined by a comparison between the lineshapes of  $Z \rightarrow e^+e^-$  in data and simulation. As a result, the corrections to the energy scale vary in time,  $|\eta|$  and  $R_9$  variable, which is defined as the energy sum of the  $3 \times 3$  crystals centered on the most energetic crystal in the candidate electromagnetic cluster divided by the energy of the candidate. The amount of smearing required changes from about 0.1% to about 2.7%, depending on the same categories as the energy scale corrections. The comparison of the dielectron invariant mass distributions in data and simulation after energy smearing are shown in Fig. 2-10.

In the barrel section of the ECAL, an energy resolution of about 1% is achieved for unconverted or late-converting photons in the tens of GeV energy range. The remaining barrel photons have a resolution of about 1.3% up to a pseudorapidity of  $|\eta| = 1$ , rising to about 2.5% at  $|\eta| = 1.4$ . In the endcaps, the resolution of unconverted or late-converting photons is about 2.5%, while the remaining endcap

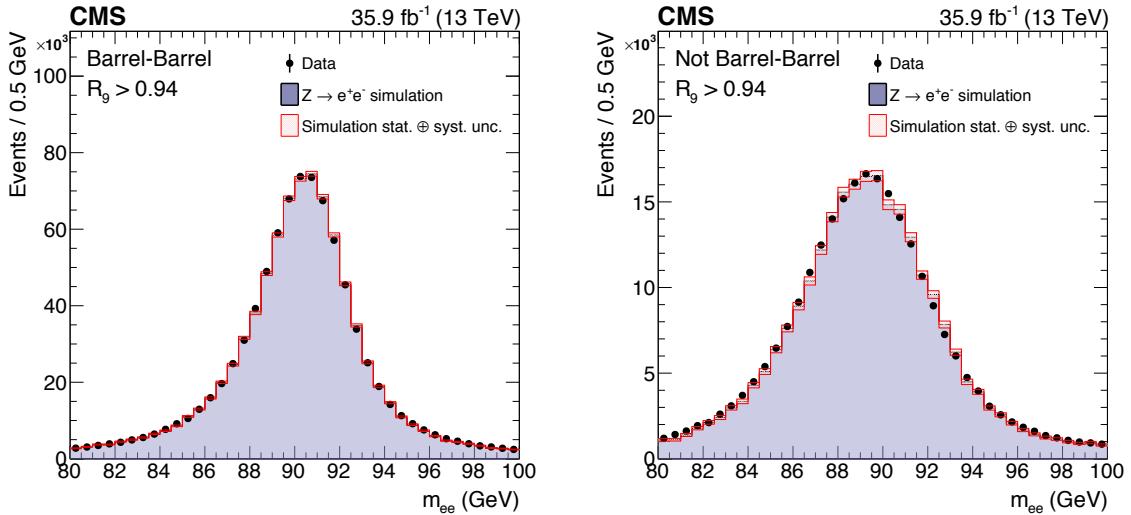


Figure 2-10: The comparison of the dielectron invariant mass distributions in data and simulation after energy smearing [43].

photons have a resolution between 3 and 4% [96].

## Muon reconstruction

The muon reconstruction starts with hits in DTs, CSCs and RPCs. Those hits are combined to form segments. This step is called local reconstruction. Three collections of muons reconstructed by different methods are described as follows:

- **Standalone Muon reconstruction.** The segments are used to generate the seeds including the information of positions, directions, and estimated muon  $p_T$ . The segments and hits from DTs, CSCs and RPCs with the seeds are then fitter by the Kalman-filter technique [97]. The resulting objects are referred to as standalone muon.
- **Global Muon reconstruction.** Each standalone muon track is matched 'outside-in' to a tracker track (also referred to as inner track or silicon track). This global muon track is then fitted by combining the hits from both tracker and standalone tracks using Kalman-filter technique.
- **Tracker Muon reconstruction.** Tracker tracks with  $p_T > 0.5 \text{ GeV}$  and total

momentum  $p > 2\text{ GeV}$  are matched ‘inside-out’ to the muon system, with bending effect from magnetic field, multiple scattering and expected energy losses while traveling through the detector materials taking into account. The extrapolated track will be considered as tracker muon if it matches to at least one muon segment, formed by hits within each DT and CSC.

In general, tracker muon reconstruction is more efficient than the global muon reconstruction at low momenta  $p \leq 5\text{ GeV}$ , as it merely requires a single muon segment in the muon system. The downside is that the hadron shower with high energy may “punch through” the calorimeter and reach the innermost muon station, which is then misidentified as a tracker muon. As for the global muon reconstruction, high efficiency is maintained for muons with higher  $p_T$ , which can traverse through more than one muon station. As a result, around 99% of muons within the acceptance of the muon system can be well reconstructed either as global muon or tracker muon, and usually as both. For those only reconstructed as standalone muon, they are usually not used in physics analyses as they have worse momentum resolution and are more probable from cosmic-ray.

The ensemble of reconstructed muons (abbreviated as reco muon) is used as ingredient by the PF event reconstruction. In the PF algorithm, some of identification requirements together with the measurement of energy in the calorimeter are optimized to identify muons with high efficiency and low fake rate, especially those in jets as fake or missed reconstructed (identified) muons can bias measurements of jets and missing transverse energy  $E_T^{\text{miss}}$ . Consequently, this selection is able to retain not only isolated muons but also non-isolated muons, and those from decay products of hadron that typically treated as background.

Three sets of requirements are imposed to label reco muons as “isolated”, “pf-tight”, and “pf-loose”, and are grouped as particle-flow muons. Reco muons are considered to be isolated if the sum of the  $p_T$  of the tracks and of the transverse energy of the calorimeter hits calculated in a cone of size  $\Delta R = 0.3$  centered on the muon is less than 10% of the muon  $p_T$ . The pf-tight and pf-loose selections, tuned to identify muons in jets, are applied to the remaining reco muons. The pf-tight

criteria requires the muon track to have a certain number of hits with compatibility with the muon segment and the energy deposited in calorimeter, defined by a template-based simulation. In the pf-loose selection the required number of hits are relaxed and the compatibility requirements are simply replaced to a matching of the track to hits in the muon stations.

Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution, for muons with  $p_T$  up to 100 GeV, of 1% in the barrel ( $|\eta| < 0.9$ ) and 3% in the endcaps ( $|\eta| > 0.9$ ). The  $p_T$  resolution in the barrel is better than 7% for muons with  $p_T$  up to 1 TeV [95]. The improvement compared to the 2010 results [98] is primarily due to the improvement to the tracker alignment [99].

### 2.3.2 Pile-up & Primary vertex

The high instantaneous luminosity of the LHC results in multiple proton-proton interactions per bunch crossing, which is often referred to as event pile-up. In 13 TeV collisions in 2016 data-taking period, there was on average 27 interactions per bunch crossing, as shown in Fig. 2-11 [90].

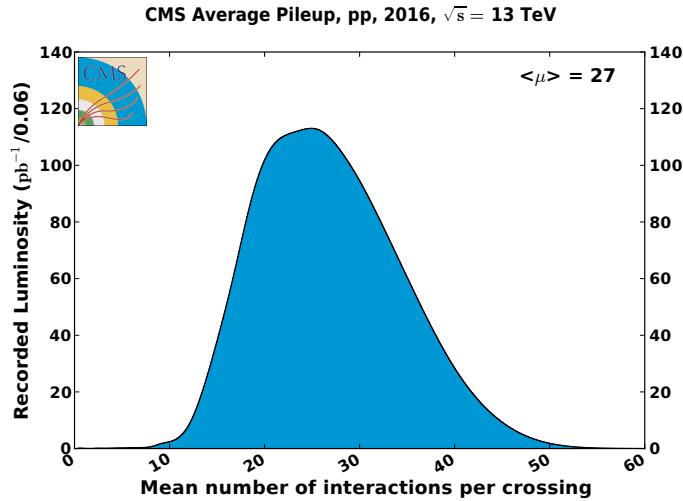


Figure 2-11: Mean number of interactions per bunch crossing for the 2016 pp run at 13 TeV [90].

The reconstructed vertex with the largest value of summed physics-object  $p_T^2$

is taken to be the primary pp interaction vertex. The physics objects are the jets, clustered using the jet finding algorithm [100, 101] with the tracks assigned to the vertex as inputs, and the associated missing transverse momentum, taken as the negative vector sum of the  $p_T$  of those jets. The simulated  $t\bar{t}$  events (inclusive decays) are used to validate the performance of the vertexing algorithm. Consequently, a resolution, defined as the difference between the position of the reconstructed vertex and the true vertex along the z direction, better than 1 mm can be achieved, and a harder  $p_T$  threshold ( $p_{T,\min}$ ) for a track to be taken into account does not result in a significantly degradation of the resolution. The efficiency of reconstructing the primary vertex within 5 mm of the true vertex is  $\sim 97\%$ . Restricting it to be within 1 mm of the generated vertex, the efficiency is about 90% for  $p_{T,\min} = 2 \text{ GeV}$ , and remains at  $\sim 86\%$  with  $p_{T,\min} = 5 \text{ GeV}$  [102].

# Chapter 3

## Analysis procedures

### 3.1 Data and simulated samples

#### 3.1.1 Data sample

The MuonEG dataset collected in 2016 at  $\sqrt{s} = 13$  TeV, corresponding to a total integrated luminosity of  $35.9 \text{ fb}^{-1}$ , is used. The data for each run period is summarized in Table 3.1. The data used are recorded with all sub-detectors running under good condition.

Dataset Name	Luminosity( $\text{fb}^{-1}$ )
/MuonEG/Run2016B-03Feb2017_ver2-v2/MINIAOD	5.8
/MuonEG/Run2016C-03Feb2017-v1/MINIAOD	2.6
/MuonEG/Run2016D-03Feb2017-v1/MINIAOD	4.2
/MuonEG/Run2016E-03Feb2017-v1/MINIAOD	4.0
/MuonEG/Run2016F-03Feb2017-v1/MINIAOD	2.7
/MuonEG/Run2016F-03Feb2017-v1/MINIAOD	0.4
/MuonEG/Run2016G-03Feb2017-v1/MINIAOD	7.5
/MuonEG/Run2016H-03Feb2017_ver2-v1/MINIAOD	8.4
/MuonEG/Run2016H-03Feb2017_ver3-v1/MINIAOD	0.2

Table 3.1: Summary of data sample used in the analysis.

### 3.1.2 Simulated samples

#### Signal samples

The  $H \rightarrow J/\psi \gamma \rightarrow \mu\mu\gamma$  sample, with  $m_H = 125$  GeV, is produced with POWHEG v2.0 [103, 104] for ggF, VBF, VH, and ttH productions. The generator is interfaced with PYTHIA 8.212 [105, 106] for hadronization and fragmentation with underlying event tune CUETP8M1 [107]. The parton distribution function PDF set used is NNPDF3.0 [108]. The samples used, with the cross-section for each production mode taken from Ref. [12], are summarized in the Table 3.2. The cross sections for all the productions are calculated with QCD and electroweak (EW) corrections. The EW correction for each mode includes the calculation up to next-to-leading order (NLO). The QCD correction for the ggF is calculated at next-to-next-to-next-to-leading order, at next-to-next-to-leading order (NNLO) for the VBF and VH , and at NLO for the ttH.

Dataset name	Production	Cross-section(pb)	Order
/ggH_HToJPsG*/RunII Summer16*/*	ggF	48.6	N3LO QCD & NLO EW
/VBFH_HToJPsG*/RunII Summer16*/*	VBF	3.78	NNLO QCD & NLO EW
/ZH_HToJPsG*/RunII Summer16*/*	ZH	0.884	NNLO QCD & NLO EW
/WpHJ_HToJPsG*/RunII Summer16*/*	$W^+H$	0.840	NNLO QCD & NLO EW
/WmHJ_HToJPsG*/RunII Summer16*/*	$W^-H$	0.538	NNLO QCD & NLO EW
/ttH_HToJPsG*/RunII Summer16*/*	ttH	0.507	NLO QCD & NLO EW
	<b>Total</b>	<b>55.1</b>	

Table 3.2: Summary of Higgs boson signal samples.

The  $Z \rightarrow J/\psi \gamma \rightarrow \mu\mu\gamma$  sample, with  $m_Z = 91.2$  GeV [65], is produced with the PYTHIA 8.226 generator for hadronization and fragmentation with tune CUETP8M1. The SM Z boson production cross section includes the NNLO contribution, QCD and electroweak corrections from FEWZ 3.1 using the NLO PDF set NNPDF3.0. To account for the potential mismodeling of the Z  $p_T$  distribution and the missing  $\gamma^*$  contribution in the sample, we apply the Z  $p_T$  reweighting. The Drell-Yan jets samples (with  $m_{ll} > 50$  GeV), one generated with MADGRAPH5\_aMC@NLO matrix-element generator and the other one with POWHEG generator, are used as references to calculate the reweighting factors. In both samples, the NLO contribution, the interference, and the contribution of the  $\gamma^*$  diagrams are included. The left plot

of Fig. 3-1 shows the  $Z$   $p_T$  distributions at generator level of the  $Z \rightarrow J/\psi \gamma$  and Drell-Yan jets samples. The interference between diagrams at NLO in *aMC@NLO* sample are properly handled. The ratio of the two  $p_T$  distributions "Drell-Yan jets (*aMC@NLO*)" to " $Z \rightarrow J/\psi \gamma$ " is used as binned weight, as shown in the right plot of Fig. 3-1.

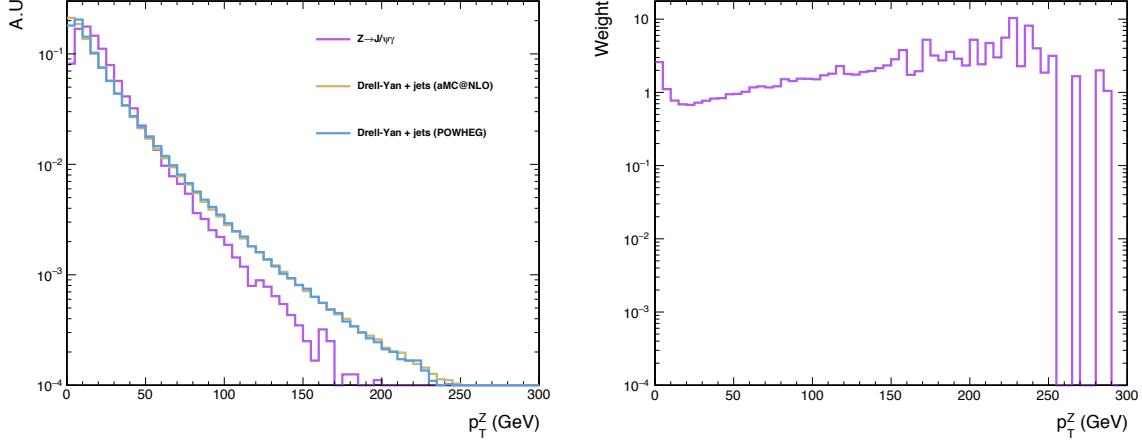


Figure 3-1: The right plot shows the  $Z$   $p_T$  distributions at generator level of the  $Z \rightarrow J/\psi \gamma$  and Drell-Yan jets samples. The left plot shows the ratio of the two  $p_T$  distributions "Drell-Yan jets(*aMC@NLO*)" to " $Z \rightarrow J/\psi \gamma$ ", as binned weight to be applied to the PYTHIA sample.

## J/ $\psi$ polarization

The Higgs boson is now commonly believed to be a spin-0 particle, and the  $J/\psi$  from its decay is therefore transversely polarized (with  $J_Z = \pm 1$ ). However, this polarization is not correctly simulated in the generation. The distribution of  $\cos\theta$  was checked, where  $\theta$  is the angle between the muon and the direction of  $J/\psi$ , and is derived at the generator level. The angle  $\theta$  is calculated without kinematic requirement and in the rest frame of  $J/\psi$ , where the direction of  $J/\psi$  is obtained from the center-of-mass (CM) frame of the Higgs boson (the helicity frame). The  $H \rightarrow J/\psi \gamma$  samples are therefore reweighted using weight  $w = 3/4 \times (1 + (\cos\theta)^2)$  per event [109]. This reweighting preserves the total number of events in the samples,

however, results in a decrease of the signal acceptance by 7.0%. No systematic uncertainty is assigned for this procedure since the reweighting is done via exact formula, and the angular distribution after reweighting is the one we expect. Fig. 3-2 shows the distributions of the  $H \rightarrow J/\psi \gamma$  samples before (green), after (blue) reweighting, and of the  $H \rightarrow \gamma^* \gamma$  sample (red) where the polarization of  $\gamma^*$  is correctly simulated.

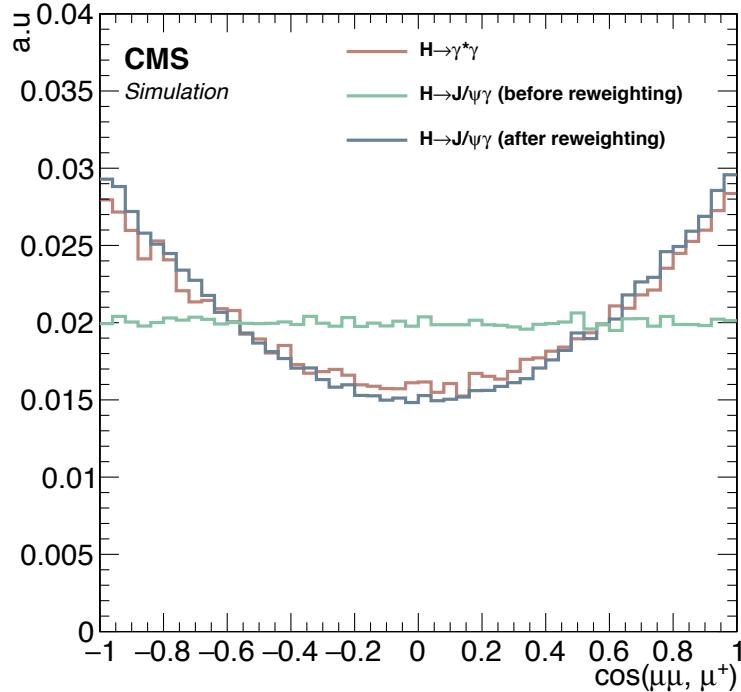


Figure 3-2: Distributions of  $\cos\theta$  of  $J/\psi \rightarrow \mu\mu$  and  $\gamma^* \rightarrow \mu\mu$ . The green distribution is the  $H \rightarrow J/\psi \gamma$  sample before reweighting; the red distribution is from  $H \rightarrow \gamma^* \gamma$ ; the blue distribution is  $H \rightarrow J/\psi \gamma$  sample after reweighting.

The Z boson is a spin-1 particle, the  $J/\psi$  from its decay can be transversely (with  $J_Z = \pm 1$ ) or longitudinally polarized (with  $J_Z = 0$ ), depending on the polarization of the Z boson. Fig. 3-3 shows the  $\cos\theta$  distributions with different polarization scenarios. Table 3.3 summarizes the reweight formulae and effects on acceptance from different polarization assumption.

The results from the Z boson polarization measurement are not used to con-

strain the helicity of the  $J/\psi$  meson in this analysis. The signal acceptance is therefore calculated for the unpolarized case and for two extreme scenarios, one in which the  $J/\psi$  is fully transversely polarized and the other in which it is fully longitudinally polarized

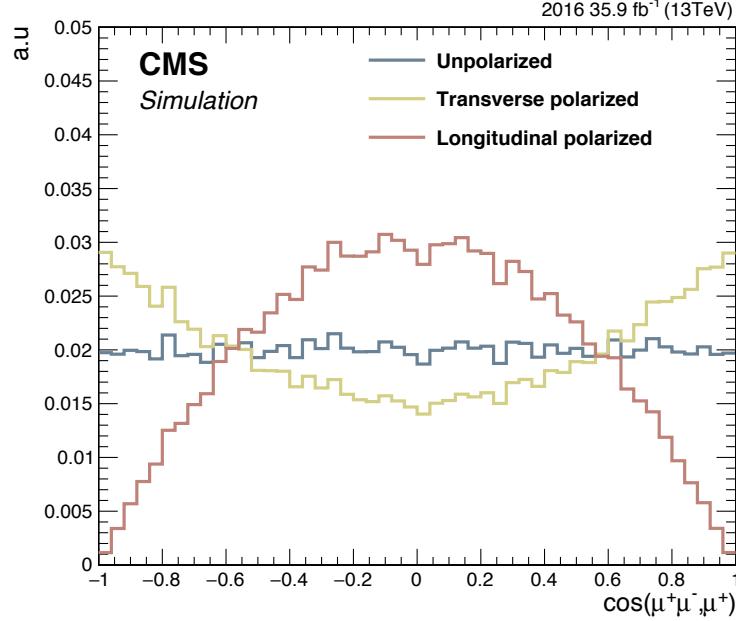


Figure 3-3: Distributions of  $\cos\theta$  of  $J/\psi \rightarrow \mu\mu$  decay resulting from different polarization scenarios. The blue distribution is the unpolarized scenario; the earthy yellow distribution is fully transversely polarized scenario; the red distribution is fully longitudinal polarized scenario.

$J_Z$	Polarization scenario	Formula	Effect on acceptance
$\pm 1$	Transverse	$3/4 \times (1 + (\cos\theta)^2)$	-7.8%
0	Longitudinal	$3/2 \times (1 - (\cos\theta)^2)$	+15.6%

Table 3.3: Summary of the reweight formulae and effects on acceptance from different polarization scenarios.

## Background

The Higgs boson Dalitz decay [110],  $H \rightarrow \gamma^*\gamma \rightarrow \mu\mu\gamma$ , results in the same final state as the signal. This process exhibits a peak in the three-body invariant mass

$m_{\mu\mu\gamma}$  at the Higgs boson mass, and is therefore referred to as a peaking, or resonant, background. It is taken into account when deriving the upper limit on the branching fraction for  $H \rightarrow J/\psi \gamma$ . The diagrams for  $H \rightarrow \gamma^*\gamma$  process are shown in Fig. [3-4]. Samples of Higgs boson Dalitz decays, produced in ggF, VBF, VH for  $m_H = 125$  GeV, are simulated at NLO using the MADGRAPH5\_aMC@NLO 2.6.0 matrix element generator [111], interfaced with PYTHIA 8.212 for parton showering and hadronization. The dimuon invariant mass  $m_{\mu\mu}$  in the ggF sample is restricted to be less than 50 GeV, while in VBF and VH samples it is less than 60 GeV. The contribution of the ttH is accounted for by scaling the VBF signal to the ttH production cross section. The branching fraction for  $H \rightarrow \gamma^*\gamma$  is obtained from MCFM 7.0.1 program [112]. The other source of peaking background comes from the decay of a Higgs boson into two muons, with a photon radiated from one of the muons. Fig. [3-5] shows the distributions of some kinematic variables for the  $H \rightarrow \mu\mu$  and the  $H \rightarrow J/\psi \gamma$  decays. As one can see, the event signatures of the decay are different from those of the  $H \rightarrow J/\psi \gamma$ . The contribution of this background is found to be negligible after the event selection.

Dataset name	$\mathcal{B}_{SM}(H \rightarrow \gamma^*\gamma \rightarrow \mu\mu\gamma)$
/GluGluHToMuMuG_M125_mll-0To50*/RunII Summer16*/MINIAODSIM	$3.83 \times 10^{-5}$
/VBFHToMuMuG_M125_MLL-0To60*/RunII Summer16*/MINIAODSIM	$3.92 \times 10^{-5}$
/ZHToMuMuG_M125_MLL-0To60*/RunII Summer16*/MINIAODSIM	$3.92 \times 10^{-5}$
/WHToMuMuG_M125_MLL-0To60*/RunII Summer16*/MINIAODSIM	$3.92 \times 10^{-5}$

Table 3.4: Summary of Higgs Dalitz decay samples.

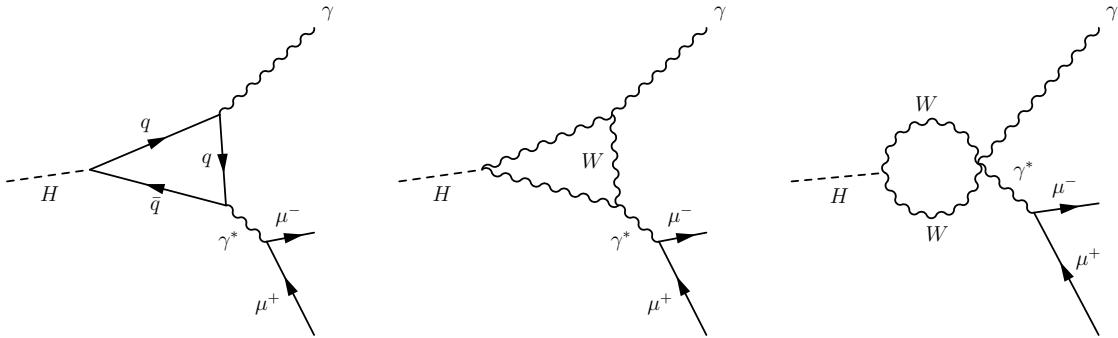


Figure 3-4: Main diagrams for the Higgs Dalitz decay,  $H \rightarrow \gamma^*\gamma \rightarrow \mu\mu\gamma$ .

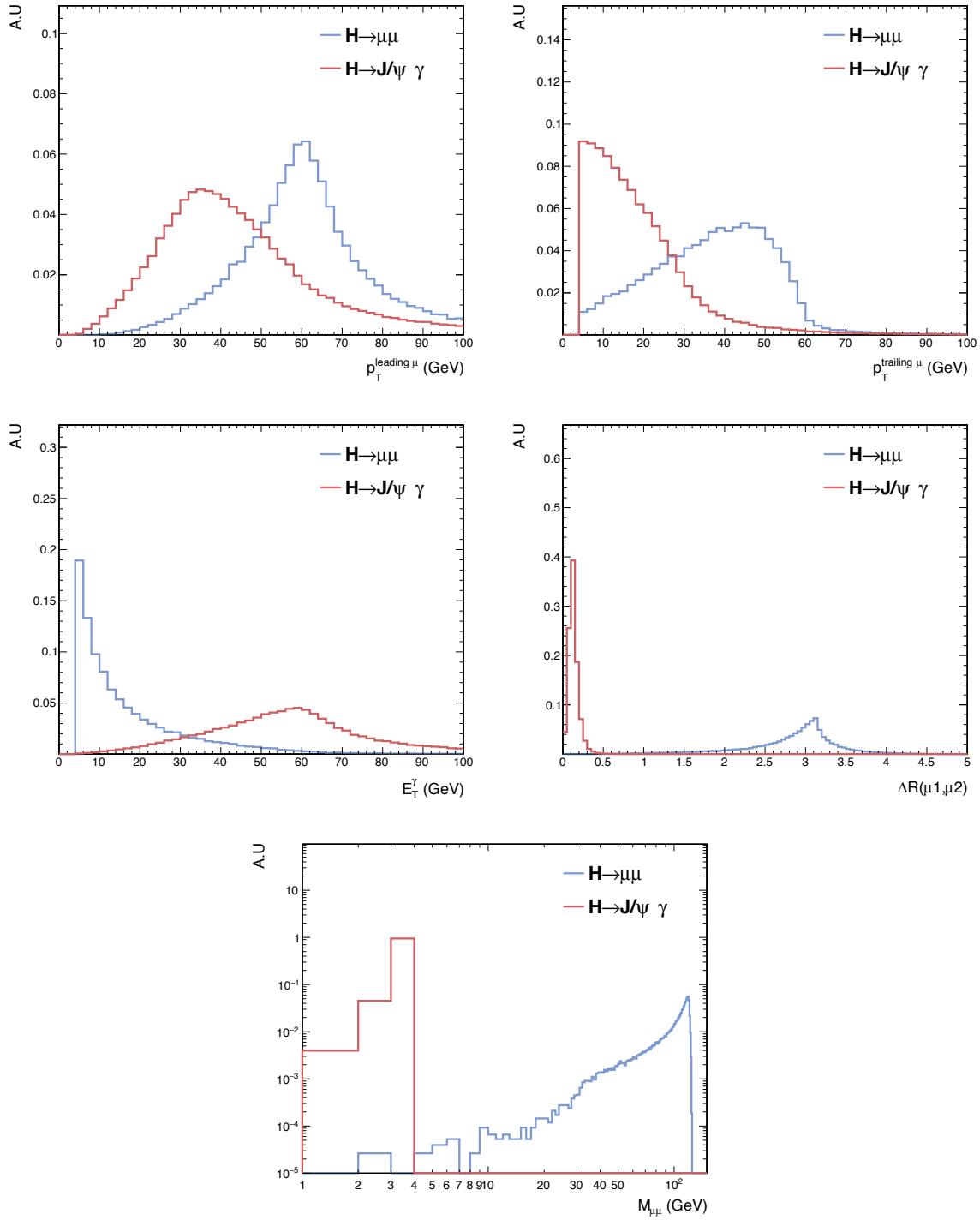


Figure 3-5: Distributions of kinematic variables for the  $H \rightarrow \mu\mu$  and the  $H \rightarrow J/\psi \gamma$  decays. (Top left)  $p_T$  of the leading muon; (Top middle)  $p_T$  of the trailing muon; (Top right)  $E_T$  of the photon; (Bottom left) angular separation  $\Delta R$  between muons; (Bottom right) dimuon mass  $M_{\mu\mu}$ .

Similarly, the Drell–Yan process,  $\text{pp} \rightarrow Z \rightarrow \mu\mu\gamma$  is a peaking background for  $Z \rightarrow J/\psi \gamma$ . The diagrams for the  $\text{pp} \rightarrow Z \rightarrow \mu\mu\gamma$  process are shown in Fig. 3-6. The MADGRAPH5\_aMC@NLO 2.6.0 generator at leading order with the NNPDF3.0 PDF set, interfaced with PYTHIA 8.226 for parton showering and hadronization with tune CUETP8M1, is used to generate a sample of these resonant background events. The photons in these events are all produced in final-state radiation from the  $Z \rightarrow \mu\mu$  decay and therefore the  $m_{\mu\mu\gamma}$  distribution peaks at the  $Z$  boson mass and there is no continuum contribution. Kinematic requirements, such as  $2 < m_{\mu\mu} < 15 \text{ GeV}$  and  $E_T^\gamma > 20 \text{ GeV}$ , are imposed when generating the sample, and results in an inclusive cross section of 93.0 pb. The additional photons added by

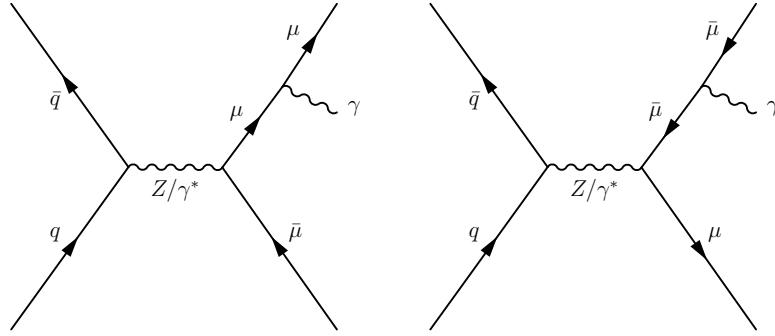


Figure 3-6: Main diagrams for the Drell-Yan process,  $\text{pp} \rightarrow Z \rightarrow \mu\mu\gamma$ .

PYTHIA may modify the photon  $E_T$  modeling in the sample. The effect is checked by using generator level information. Figs. 3-7 shows two distributions, one is the  $E_T$  of the photons which are prompt final states<sup>1</sup> (in blue) and the other one is the  $E_T$  of the photons added by the PYTHIA8 when it is interfaced with *aMC@NLO* (in red). The number of photons with  $E_T > 33 \text{ GeV}$  added by PYTHIA8 is only 0.3% of those from hard scattering. Therefore, the interface with PYTHIA has minimal effect on the overall photon  $E_T$  spectrum. No additional uncertainty is assigned.

There are also background processes that do not give resonance peaks in the three-body invariant mass spectrum. These are referred to as non-peaking (non-

---

<sup>1</sup>A particle is labeled as prompt if it is from the hard process in an interaction.

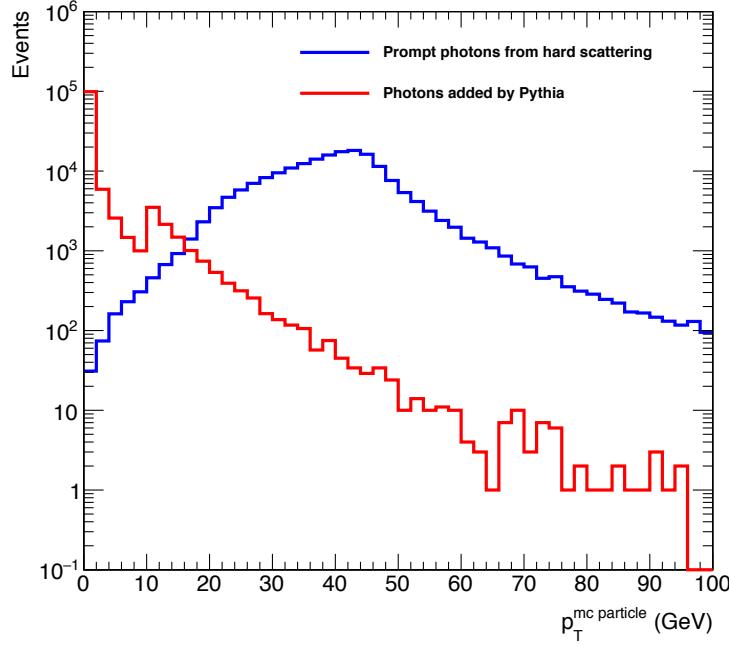


Figure 3-7: The  $E_T$  of the photons which are prompt final state (in blue) and the other one is the  $E_T$  of the photons added by the PYTHIA when it is interfaced with *aMC@NLO* (in red).

resonant) backgrounds. These processes include

- The Drell-Yan FSR process:  $pp \rightarrow Z + \gamma_{\text{FSR}} \rightarrow \mu\mu\gamma_{\text{FSR}}$ , where  $m_{\mu\mu\gamma}$  is within the Higgs ( $Z$ ) mass window.
- The Drell-Yan ISR process:  $pp \rightarrow Z/\gamma^* + \gamma_{\text{ISR}} \rightarrow \mu\mu\gamma_{\text{ISR}}$ , where  $m_{\mu\mu}$  is within the  $J/\psi$  mass window and  $m_{\mu\mu\gamma}$  is within the Higgs ( $Z$ ) mass window.
- $pp \rightarrow Z/\gamma^*(\rightarrow \mu\mu) + \text{jets}$ , where a jet is misidentified as an energetic photon which can fire the trigger and pass the event requirements.
- $pp \rightarrow \gamma + \text{jets}$ , where the muons can come from the jets.
- Inclusive quarkonium production with a jet reconstructed as a photon  $pp \rightarrow J/\psi + \text{jets}/\gamma$ , where the muons come from the quarkonium,  $J/\psi$ , in our cases.

Since currently no proper simulated samples for those processes are available, these non-resonant backgrounds are modeled using the fits to  $m_{\mu\mu\gamma}$  in data, which

will be introduced in Sec. 3.5.

### Pile-up reweighting

The simulated sample is reweighted in analysis level using minimum bias events with cross section of 69.2mb. The corresponding systematic uncertainties are described in Sec. 3.7, and are estimated to be less than 1.5% on the expected yields of the signal in both channels.

## 3.2 Trigger

The HLT\_Mu17\_Photon30\_CaloIdL\_L1ISO trigger is used in this analysis. At the L1 (L1\_Mu5IsoEG18), the trigger requires the presence of a muon with  $p_T$  greater than 5 GeV and an isolated electromagnetic object with  $p_T$  greater than 18 GeV. The main HLT requires a muon and a photon with  $p_T$  greater than 17 GeV and 30 GeV, respectively. No isolation requirement is imposed on the muon by the fact that the small angular separation between muons in the final state.

### The choice of the trigger

A study is made to compare the resulting signal efficiency with different triggers. In the single muon trigger, the  $p_T$  threshold on muon is high and there is isolation requirement calculated in the cone  $\Delta R = 0.3$ . For the double muon trigger, the  $p_T$  cut of 8 GeV is imposed on the subleading muon, and there are requirements on the isolation calculated using tracker information. Among the triggers used in analyses associated with heavy flavor or quarkonium physics, most of them are pre-scaled and target at different physics content. The only suitable choice is the HLT\_Dimuon20\_Jpsi\_v6. The L1 seed of this quarkonium trigger requires two muons of  $p_T$  greater than 13 and 6 GeV respectively.

Fig. 3-8 shows the trigger efficiency in the  $H \rightarrow J/\psi \gamma$  signal as function of  $p_T$  of the leading muon,  $p_T$  of the dimuon system, photon  $E_T$ , angular separation ( $\Delta R$ )

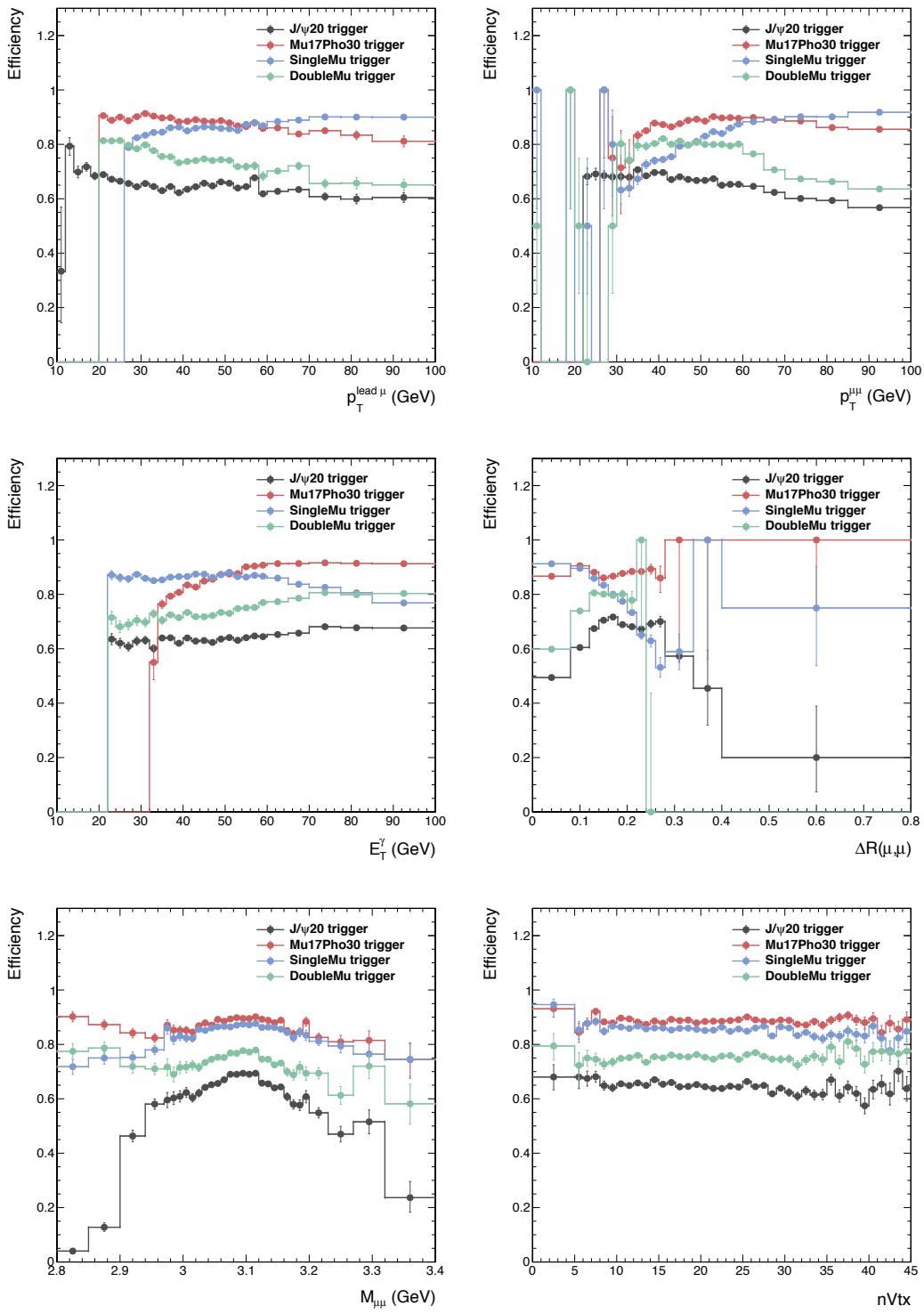


Figure 3-8: The trigger efficiency in the  $H \rightarrow J/\psi \gamma$  signal as function of leading muon  $p_T$  (top left),  $p_T$  of the dimuon system (top right), photon  $E_T$  (middle left), angular separation  $\Delta R$  between the muons (middle right), invariant mass of the dimuon system  $M_{\mu\mu}$  (bottom left), and the number of vertex (bottom right).

Trigger path
Single muon trigger
HLT_IsoMu24_v* OR
HLT_IsoTkMu24_v*
Double muon trigger
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v* OR
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v* OR
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* OR
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v
Muon-Photon trigger
HLT_Mu17_Photon30_CaloIdL_L1ISO_v*
Quarkonium trigger
HLT_Dimuon20_Jpsi_v*

Table 3.5: Triggers used in the signal efficiency study.

between the muons, invariant mass of the dimuon system  $m_{\mu\mu}$ , and the number of vertex. As one can see, the muon-photon trigger preserves the highest signal efficiency. The inefficiencies of double muon and quarkonium triggers are responsible for that both triggers are not specifically designed for the muons with small separation. The efficiency of single muon trigger is slightly lower than that of the muon-photon trigger, which may be due to the isolation requirement and high  $p_T$  threshold. Consequently, the muon-photon trigger is chosen.

In the actual signal events, the trigger efficiency is 89.2 (84.2)% in the Higgs (Z) boson decay. The trigger efficiency is measured in the control sample, and found to be 81.5 (83.3)% in data (simulation). The method of this measurement is described in the next section.

## Trigger efficiency measurement

Trigger efficiency in data is measured using  $Z \rightarrow \mu\mu\gamma$  control sample in the dataset collected by single muon trigger, while in the simulated events the Drell-Yan jets with  $m_{ll} > 50 \text{ GeV}$  sample is used. Events must have at least two muons and one photon in the final state, and are required to pass at least one of the two single muon triggers, HLT\_MuIso24 or HLT\_MuTkIso24. The muon that fires one or both

triggers is considered as the tag muon, and is further required to pass Tight Muon ID and relative isolation requirement [95]. One muon and one photon are then selected as probe objects, and are required to pass the kinematic selections listed below, which ensure that they come from the Z decay with a final-state-radiated (FSR) photon.

- $0.1 < \Delta R(\mu, \gamma) < 0.8$ , where the lower bound of 0.1 rejects events where the selected photon picks up the track from one of the muons, and the upper bound of 0.8 rejects events where neither muons emitted the photon
- $m_{\mu\mu} + m_{\mu\mu\gamma} < 180 \text{ GeV}$  to reject contribution from initial-state-radiated (ISR) photons
- $60 < m_{\mu\mu\gamma} < 120 \text{ GeV}$ , the mass window cut used to identify the Z boson.

If there are two muons passing tag selections simultaneously, we could choose between two possible tag muons. In this case, both choices are considered and tested, and the event is counted twice. This is to avoid underestimating the efficiency and the potential bias on the measurement.

The Z boson candidate mass distribution in data and MC obtained through this method are shown in Fig. 3-9. Offline selection requirements of the analysis are applied in order to factorize the selection efficiency. The events passing all these selections are counted as the denominator of the trigger efficiency. For the numerator, the probe muon (photon) is tested to see if it can fire the muon (photon) leg of the muon-photon trigger used in the analysis. The filters in the muon-photon trigger are listed in Table 3.6. (The filters checked for the muon and photon legs are different between runs B to E and F to H. The filters in the MC sample are the same as those in run F to H in data) The filters marked in red color are used for testing the muon leg, while those in blue are for the photon leg.

There is almost no Run-dependency in trigger efficiency (except for period B), as shown in the red points in Fig. 3-10 as well as the constant fits and the resulting  $\chi^2/\text{ndf}$ . The black points shown here, which serve as a reference, are the

efficiencies with the standard loose muon ID with additional  $d_Z$  and  $d_{xy}$  cuts used previously in this analysis.

Trigger efficiency as a function of probe photon  $E_T$ , probe muon  $p_T$ , probe muon pseudorapidity  $\eta^\mu$ , and probe photon supercluster pseudorapidity  $\eta_{SC}^\gamma$  are shown in Fig. [3-11]. The efficiency as function of probe muon  $p_T$  is made with the probe photon  $E_T > 33\text{ GeV}$ . Similarly, the plot as function of probe photon  $E_T$  is made with the probe muon  $p_T > 20\text{ GeV}$ .

The trigger efficiency scale factors – the ratio of Data/MC efficiencies – are to be applied to simulated samples. They are derived in bins of probe muon  $p_T$  and probe photon  $E_T$  in 2 photon supercluster eta  $\eta_{SC}$  regions : Ecal Barrel (EB) region ( $0 < \eta_{SC} < 1.4442$ ) and Ecal Endcap (EE) region ( $1.566 < \eta_{SC} < 2.5$ ). When applying the trigger efficiency scale factors to MC samples, it is assumed that the leading muon is the one that fires the muon leg of the trigger, so the leading muon  $p_T$  and photon  $E_T$  are used to determine which trigger efficiency bin to apply on an event. Results for the trigger efficiency measurement are shown in Fig. [3-12] and the scale factors are shown in Fig. [3-13]. The uncertainty of each bin on Fig. [3-12] only includes statistical uncertainty, while uncertainties shown in Fig. [3-13] are total systematic uncertainties, which will be detailed in Section [3.7].

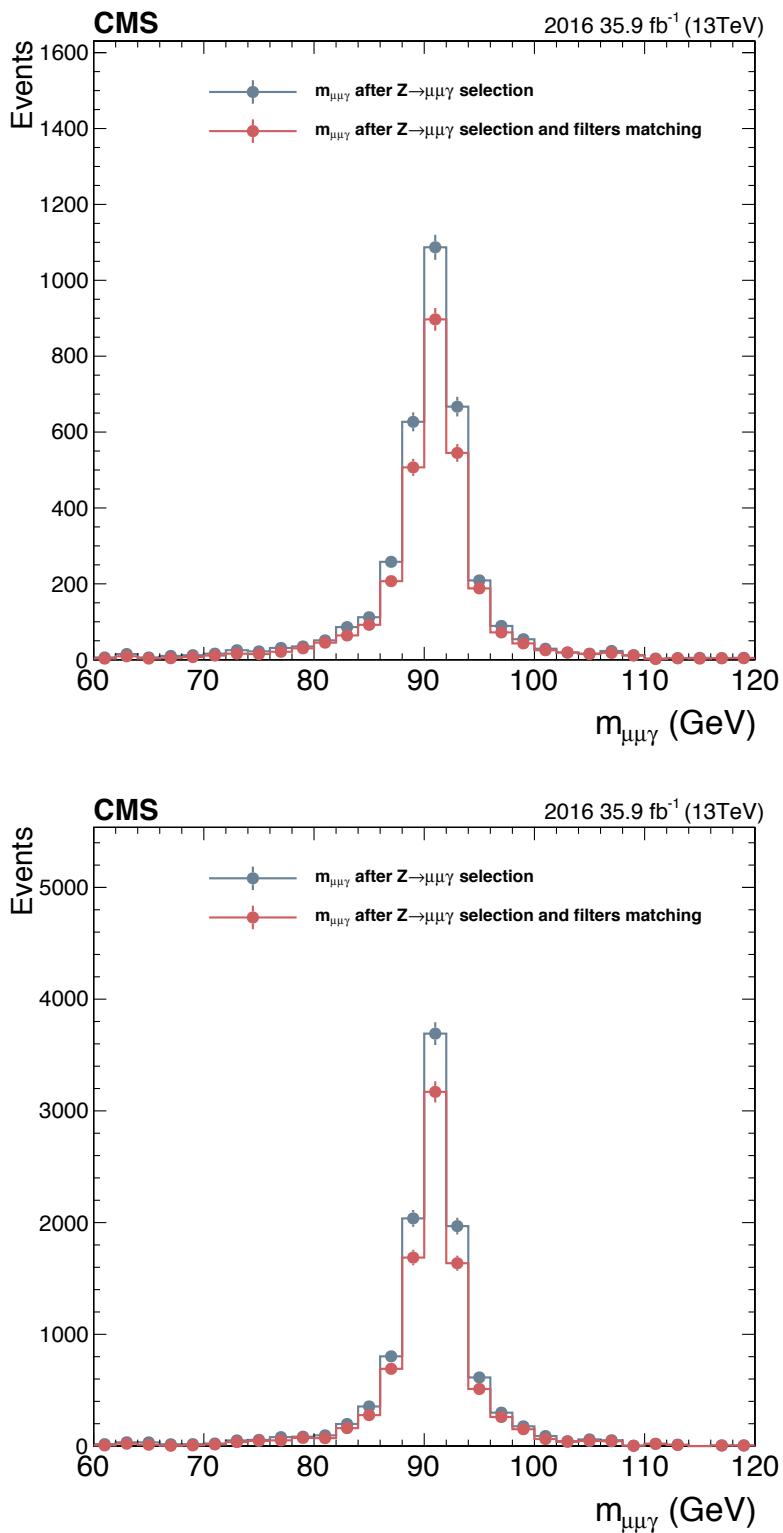


Figure 3-9: The  $Z$  boson candidate mass after selection in data (top) and MC (bottom).

<b>HLT_Mu17_Photon30_CaloIdL_L1ISO_v6</b>	
Run B~E	
Filters	hltL1sMu5IsoEG18 hltPreMu17Photon30CaloIdLL1ISO hltL1fL1sMu5IsoEG18L1Filtered5 hltL2fL1sL1Mu5IsoEG18L1f5L2Filtered7 <b>hltL3fL1sL1Mu5IsoEG18L1f5L2f7L3Filtered17</b> hltEgammaCandidates hltEGL1Mu5IsoEG18Filter hltMu17Photon30CaloIdLL1ISOEtFilter hltEgammaClusterShape hltMu17Photon30CaloIdLL1ISOClusterShapeFilter hltEgammaHoverE <b>hltMu17Photon30CaloIdLL1ISOHEFilter</b>
<b>HLT_Mu17_Photon30_CaloIdL_L1ISO_v9</b>	
RunF~H, MC samples	
Filters	hltL1sMu5IsoEG18IorMu5IsoEG20 hltPreMu17Photon30CaloIdLL1ISO hltL1fL1sMu5IsoEG18ORMu5IsoEG20L1Filtered5 hltL2fL1sL1Mu5IsoEG18ORL1Mu5IsoEG20L1f5L2Filtered7 <b>hltL3fL1sL1Mu5IsoEG18ORL1Mu5IsoEG20L1f5L2f7L3Filtered17</b> hltEgammaCandidates hltEGL1Mu5IsoEG18ORMu5IsoEG20Filter hltMu17Photon30CaloIdLL1ISOOREtFilter hltEgammaClusterShape hltMu17Photon30CaloIdLL1ISOORClusterShapeFilter hltEgammaHoverE <b>hltMu17Photon30CaloIdLL1ISOORHEFilter</b>

Table 3.6: Filters in the muon-photon trigger, listed in sequence. The filters marked in red color are used for testing the muon leg, while those in blue are for the photon leg.

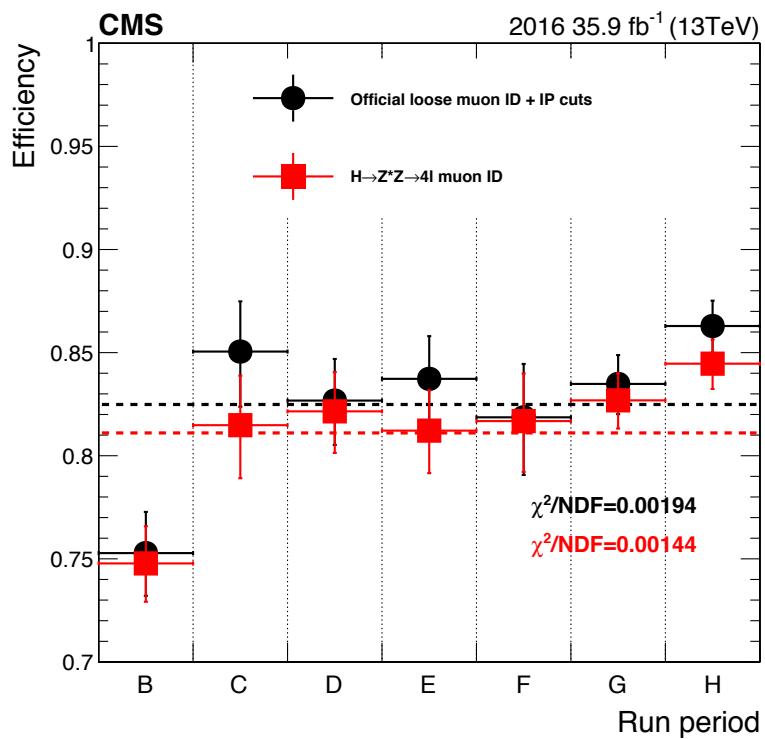


Figure 3-10: Trigger efficiency in each run period. Points in black correspond to the efficiencies measured when using the standard loose muon ID with additional  $d_Z$  and  $d_{xy}$  cuts used previously in this analysis, while red points correspond to the efficiencies measured using the muon ID optimized for  $H \rightarrow ZZ^* \rightarrow 4l$  analysis that is currently used.

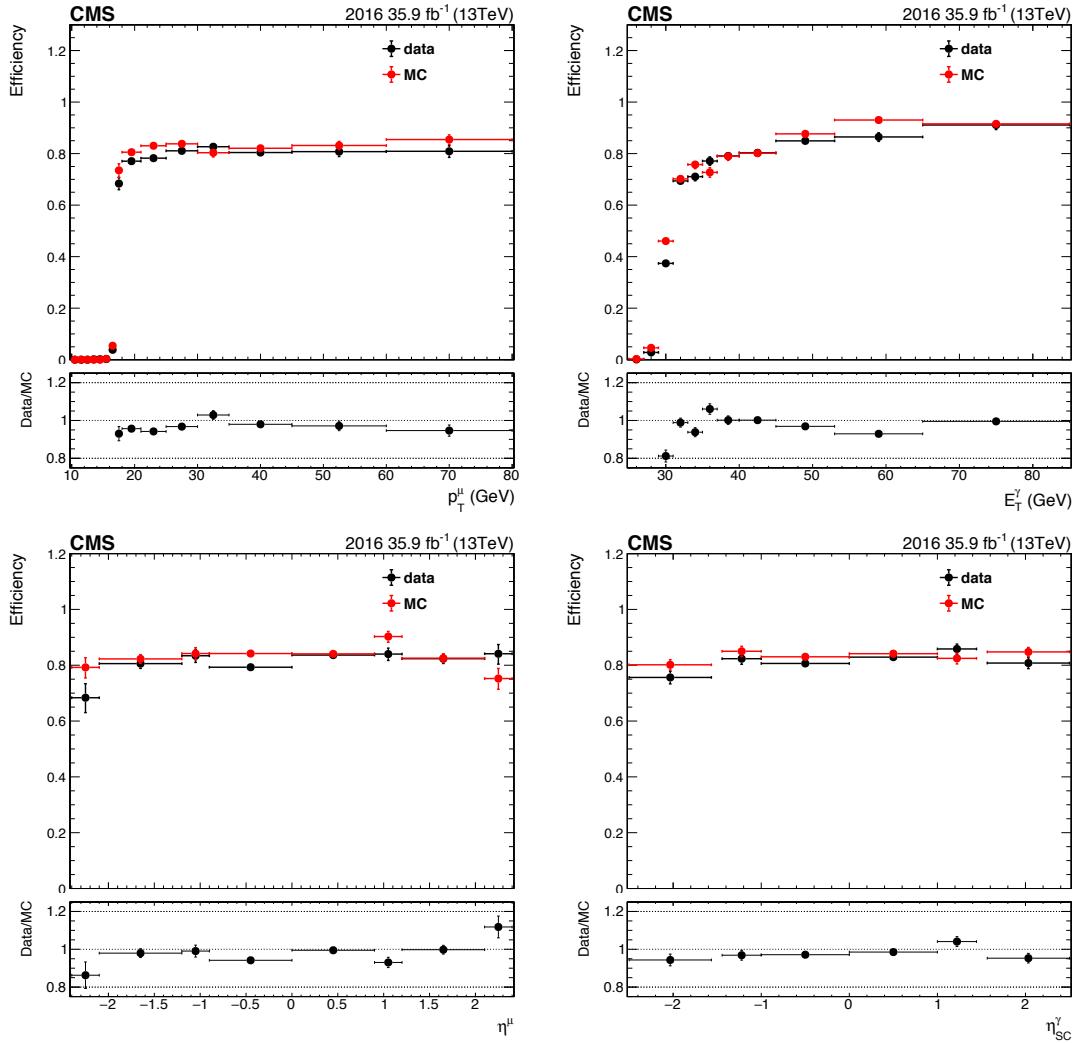


Figure 3-11: Trigger efficiency as a function of probe muon  $p_T$  (top left), probe photon  $E_T$  (top right), probe muon  $p_T$  (bottom left), and probe photon  $\eta_{SC}^{\gamma}$  (Bottom right).

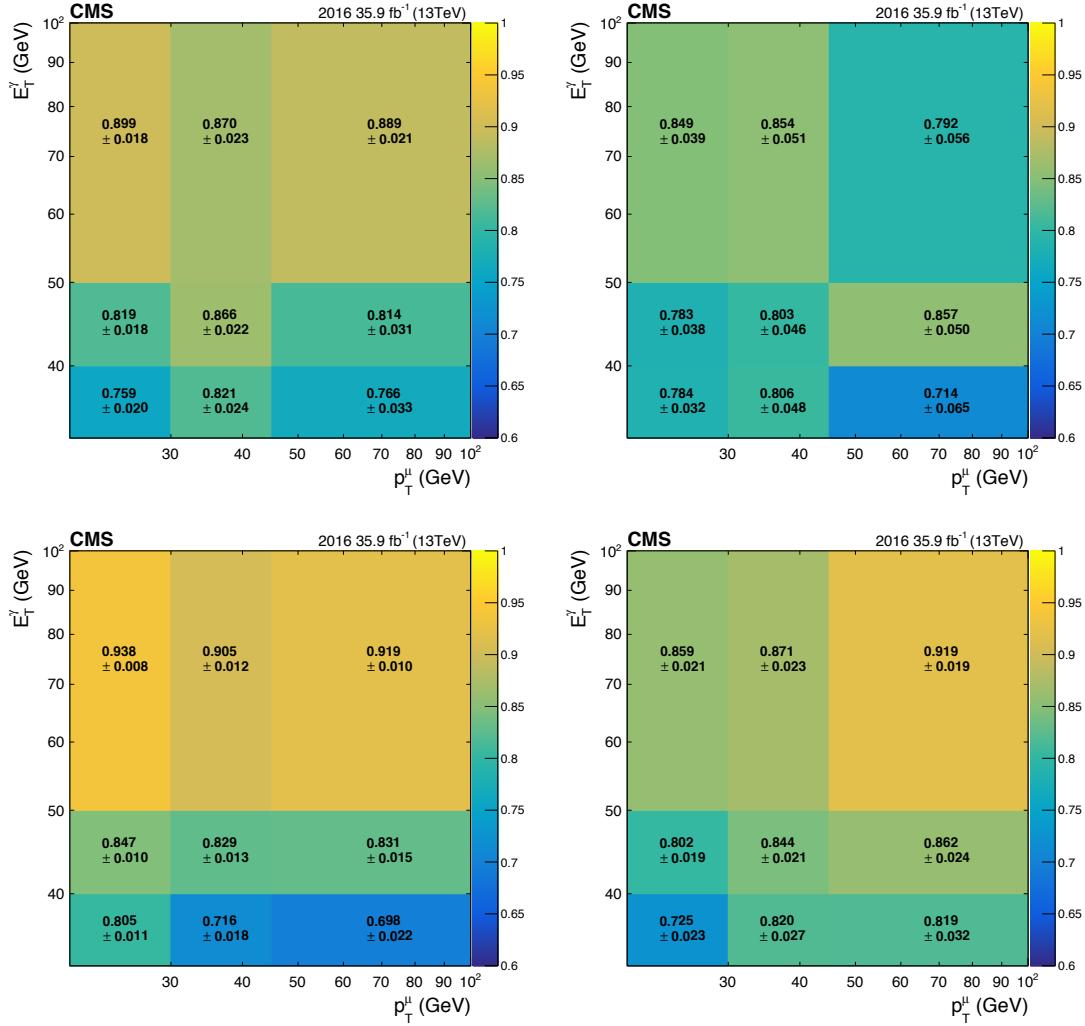


Figure 3-12: Trigger efficiency in bins of muon  $p_T$  vs photon  $E_T$  for data with the photon in EB region (top left) and in EE region (top right), and for MC with the photon in EB region (bottom left) and in EE region (bottom right).

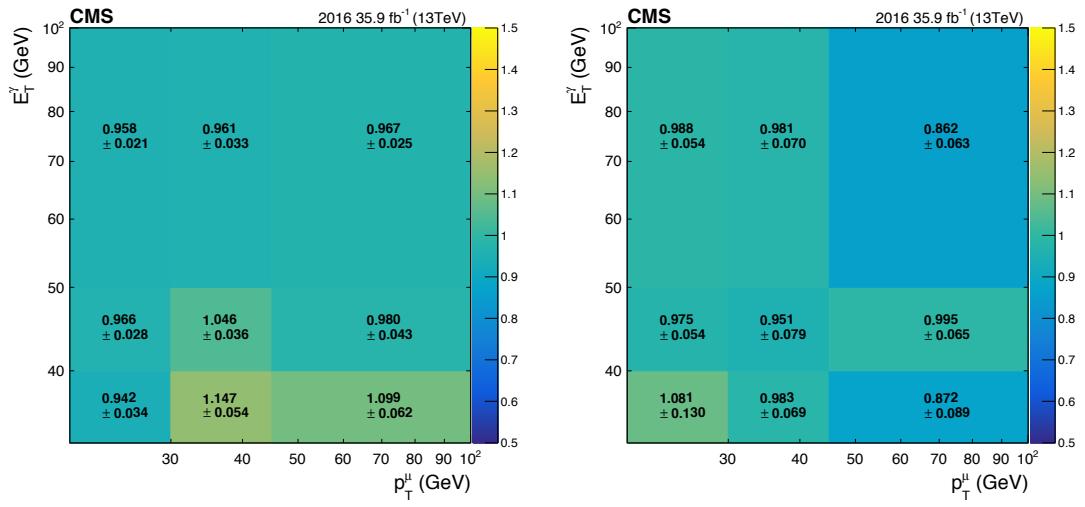


Figure 3-13: Trigger efficiency scale factors in bins of photon  $p_T$  vs muon  $p_T$  for the selected photon in ECAL EB region (left), for the selected photon in ECAL EE region (right).

## 3.3 Object identification

### 3.3.1 Muon identification

It was observed in 2016 data that a single muon may be incorrectly reconstructed as two or more muons. To deal with this situation, the “ghost cleaning” procedure is performed. Tracker muons matched to segments in at least two muons stations are retained. If there are two muons sharing more than 50% of their segments, the one with lower reconstruction quality is removed.

Two opposite-sign muons are selected with the identification requirements which are motivated by  $H \rightarrow ZZ^* \rightarrow 4\ell$  analysis [40] and are listed as follows:

- Muons must be reconstructed as particle-flow muons, and can either be global muons or tracker muons. Those only reconstructed as standalone muon are rejected.
- $p_T > 4, |\eta| < 2.4$
- Muons must have  $d_{xy} < 0.5 \text{ cm}$ ,  $d_z < 1 \text{ cm}$ , where  $d_{xy}$  and  $d_z$  are defined as the closest distance between the track of the muon and the PV in the  $\phi$  plane and the z direction respectively.
- Significance of the impact parameter in 3-dimensional space  $SIP_{3D} = |IP/\sigma_{IP}| < 4$ , where IP is the closest distance between the track of the muon and the event vertex,  $\sigma_{IP}$  is the uncertainty of the IP.

The usage of impact parameter cuts suppresses the muons from the decays of heavy-flavor hadrons or products of cosmic ray. If the muon  $p_T$  is greater than 200 GeV, it is selected if it passes Tracker High- $p_T$  ID. After the whole set of selection, there is no event with the muon  $p_T$  greater than 200 GeV in both Higgs and Z boson searches.

In order to discriminate prompt muons from Higgs (Z) boson decays from those from electroweak decays of hadrons within jets, the Particle-Flow isolation require-

ment is applied. In this analysis, the relative isolation is calculated for the leading muon.

$$\mathcal{I}^\mu \equiv \frac{\sum p_T^{\text{charged}} + \max \left[ 0, \sum E_T^{\text{neutral}} + \sum E_T^\gamma - p_T^{\text{PU}}(\mu) \right]}{p_T^\mu} \quad (3.1)$$

A cone of size  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.3$  is constructed around the direction of muon momentum. The  $\sum p_T^{\text{charged}}$  is the scalar sum of the transverse momenta of charged hadrons originating from the chosen primary vertex of the event. The  $\sum E_T^{\text{neutral}}$  and  $\sum E_T^\gamma$  are the scalar sums of the transverse energy for neutral hadrons and photons, respectively. Since the isolation variable is sensitive to energy deposits from pileup interactions, the  $p_T^{\text{PU}}(\mu)$  contribution is subtracted. The pileup contribution  $p_T^{\text{PU}}(\mu) \equiv 0.5 \sum_i p_T^{\text{PU},i}$ , where  $i$  runs over the momenta of the charged hadron PF candidates not originating from the primary vertex, and the factor of 0.5 corrects for the different fraction of charged and neutral particles in the cone. These momentum and energy sums do not include the contribution from the muon itself.  $\Delta\beta$  correction is applied, where  $\Delta\beta \equiv 0.5 \sum_{\text{PU}}^{\text{charged hadron}} p_T$  is the estimation of the energy deposit of neutral hadrons and photons from other pileup vertices. The isolation is required to be less than 0.35 for the leading muon, corresponding to  $\sim 96\%$  of signal efficiency and  $\sim 81\%$  of background rejection power.

The reason that the isolation is not calculated for the trailing muon is that the  $\Delta R$  for most of selected muon pairs are less than 0.3 (as can be seen from Fig. [3-26], [3-27] and [3-28]), which means that the trailing muon is within the isolation cone defined with the leading muon. The Isolation efficiencies as functions of  $p_T^{\text{leading } \mu}$ ,  $p_T^{\text{trailing } \mu}$ ,  $\eta^{\text{leading } \mu}$ ,  $\eta^{\text{trailing } \mu}$ , and  $p_T^{\mu\mu}$  are shown in Fig. [3-14]. Applying isolation on both muons is about 7% less efficient than applying it only on the leading muon, which is due to the fact that the trailing muon  $p_T$  is not significantly greater than other activities in the defined cone.

When the subleading muon is in the isolation cone of the leading muon, its  $p_T$  contribution is subtracted in the isolation sum of the leading muon, and vice versa. This can be verified by looking at the isolation of the leading muon divided by

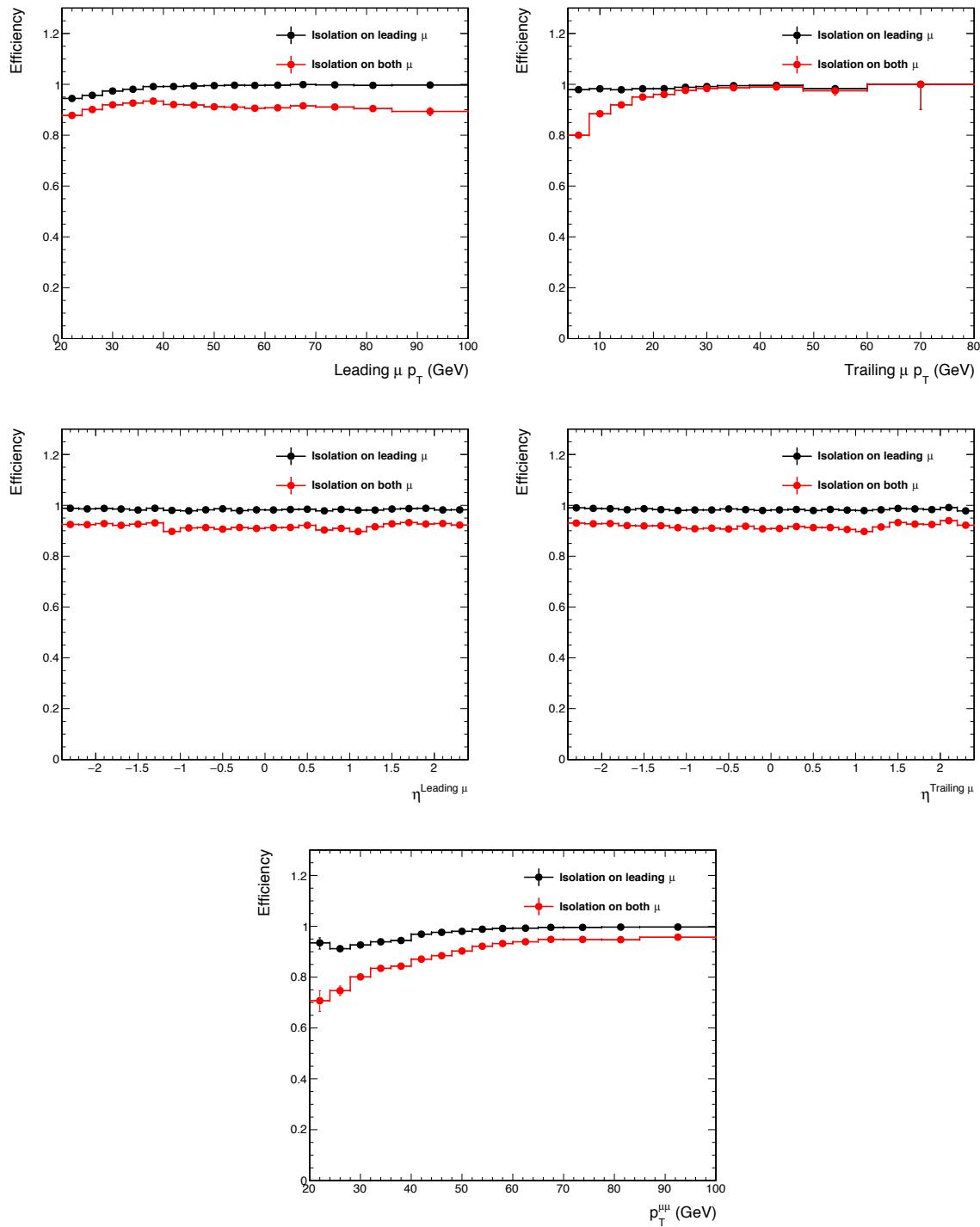


Figure 3-14: Relative isolation efficiency for muon as function of  $p_T^{\text{leading } \mu}$  (top left),  $p_T^{\text{trailing } \mu}$  (top right),  $\eta^{\text{leading } \mu}$  (bottom left),  $p_T^{\mu\mu}$  (bottom right).

the  $p_T$  of the trailing muon in each sample with isolation requirement relaxed, as shown in Fig. 3-15. All the distributions are normalized to unity. If the subleading muon is not excluded in the isolation of the leading muon, then it is expected that there will be a peak at  $\sim 1$  on the distribution, which is not seen.

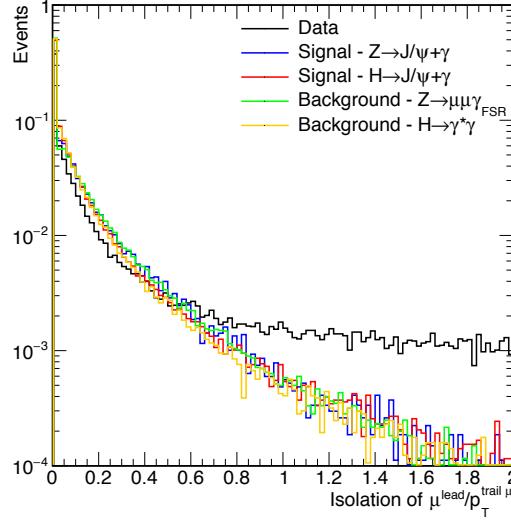


Figure 3-15: The isolation of the leading muon divided by the  $p_T$  of the trailing muon in each sample. All the distributions are normalized to unity.

Fig. 3-16 shows the  $m_{\mu\mu}$  distributions of the events selected with isolation requirement (left) and without isolation requirement (right). Muons from  $J/\psi$  decay must be true muons, so the fake muons should mostly fall in the continuum background but not form in  $J/\psi$  peak. Therefore, the numbers of background,  $N_{bkg}$ , from the fit can roughly tell us how many fake muons will be selected if no isolation requirement is imposed. By removing the isolation cut,  $N_{bkg}$  changes from  $\sim 492$  to  $\sim 756$ , meaning that fake muons roughly decrease by 34.9%.

The other information that can be extracted here is that, lots of events from QCD background can be removed by applying the isolation, based on the fact that the  $J/\psi$  in the distributions are from QCD events rather than from actual signal  $H(Z) \rightarrow J/\psi\gamma$ . Whether the isolation is applied or not has negligible impact on the expected signal yields (less than 1%).

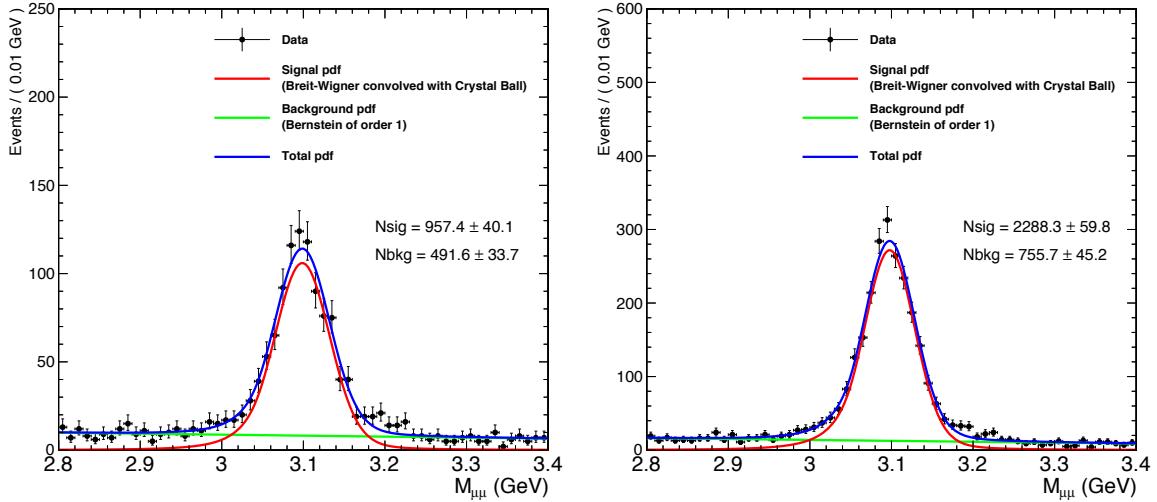


Figure 3-16: The  $m_{\mu\mu}$  distributions of the events selected with isolation requirement (left) and without isolation requirement (right). By removing the isolation cut, the fake muons roughly decrease by 34.9%.

### Muon momentum calibration

In this analysis, Rochester Muon Momentum Corrections [113] derived for 2016 dataset are applied. Biases in the measurement of muon momenta in hadron collider experiments can originate from several sources such as misalignment of the detectors, the deficiency in the software reconstruction, and uncertainties in magnetic field. Corrections are developed to remove such biases. The momentum scale corrections are extracted using the average of  $1/p_T$  ( $\langle 1/p_T \rangle$ ) spectra of muons from Z decay, while the resolution corrections and scale factors are derived by comparing the  $m_{\mu\mu}$  distributions between data and MC. The corrections are then applied to correct the momentum scale in data events and resolution in simulated events. We validate whether the Rochester correction would give consistent energy scale and resolution between data and MC for the muons from decay of J/ $\psi$  candidates in  $H \rightarrow J/\psi\gamma$  events. In this validation study, the events are required to satisfy the nominal selection requirements with relaxed dimuon and photon transverse momenta ( $p_T^{\mu\mu}, E_T^\gamma / m_{\mu\mu\gamma} > 0.16(20/125)$ ). To quantify the scale and resolution, a Breit-Wigner convolved with a Crystal Ball function (Eq. 3.2) is used to fit

the distribution for the signal events. For the data events, Breit-Wigner convolved with a Crystal Ball function in addition of the Bernstein 1<sub>st</sub> polynomial (Eq. 3.3) is used as model. As can be seen in Fig. 3-17, the  $m_{\mu\mu}$  distribution in MC is smeared, while the scale of the  $m_{\mu\mu}$  distribution in data is shifted.

$$f_{J/\psi-MC} = \text{BW}(m, \Delta) \otimes \text{CB}(0, \sigma_{CB}, \alpha, n) \quad (3.2)$$

$$f_{J/\psi-\text{data}} = N_{sig} \times f_{J/\psi-MC} + N_{bkg} \times \text{Bern.1st}(p1) \quad (3.3)$$

Associated systematic uncertainty is quoted and will be detailed in Sec ??.

### Muon efficiency measurements

A “tag-and-probe” method [114] based on samples of  $Z \rightarrow \mu\mu$  and  $J/\psi \rightarrow \mu\mu$  events in data and simulation is used to measure the efficiency, and is found to be between 94–98 (92–97)% in the barrel (endcap), depending on muon  $p_T$  and  $\eta$ . The isolation efficiency is measured with  $Z \rightarrow \mu\mu$  events, and found to be  $p_T$  dependent and between 90 (92) and 100% in the barrel (endcap).

The difference in the efficiencies measured in simulation and data, which on average is 1% per muon, is used to rescale the selection efficiency in the simulated samples. The products of all the data to simulation scale factors for muon tracking, reconstruction, identification, impact parameter and isolation requirements and corresponding uncertainties are shown in Fig. 3-18.

Reconstruction and identification	$p_T > 20 \text{ GeV}$	$Z \rightarrow \mu\mu$ events are used
	$p_T < 20 \text{ GeV}$	$J/\psi \rightarrow \mu\mu$ events are used
Impact parameter	$Z \rightarrow \mu\mu$ events are used for the whole $p_T$ range	
Isolation	$Z \rightarrow \mu\mu$ events are used for the whole $p_T$ range	
Tracking	$Z \rightarrow \mu\mu$ events are used for the whole $p_T$ range	

Table 3.7: The summary table of muon efficiencies and scale factors measurement.

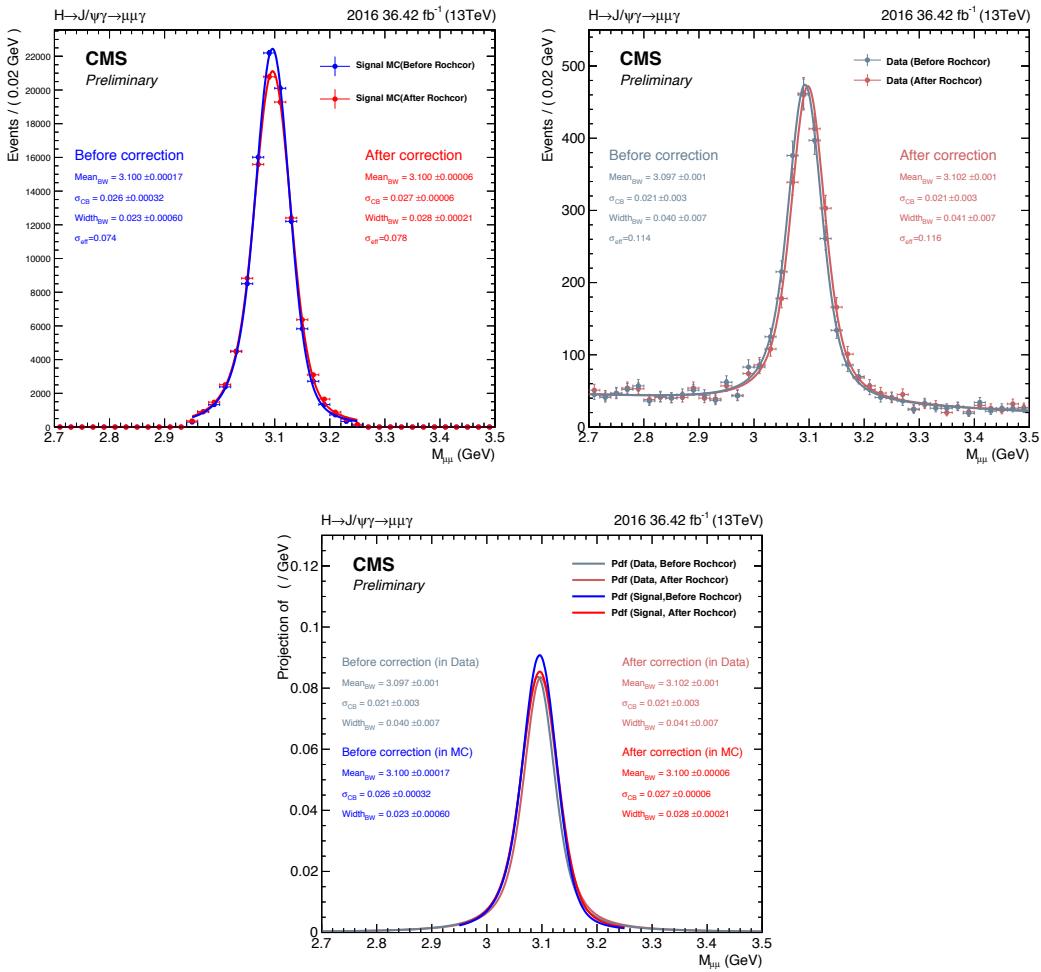


Figure 3-17: Comparisons between the dimuon mass  $m_{\mu\mu}$  distributions with and without the corrections in both data and signal MC.

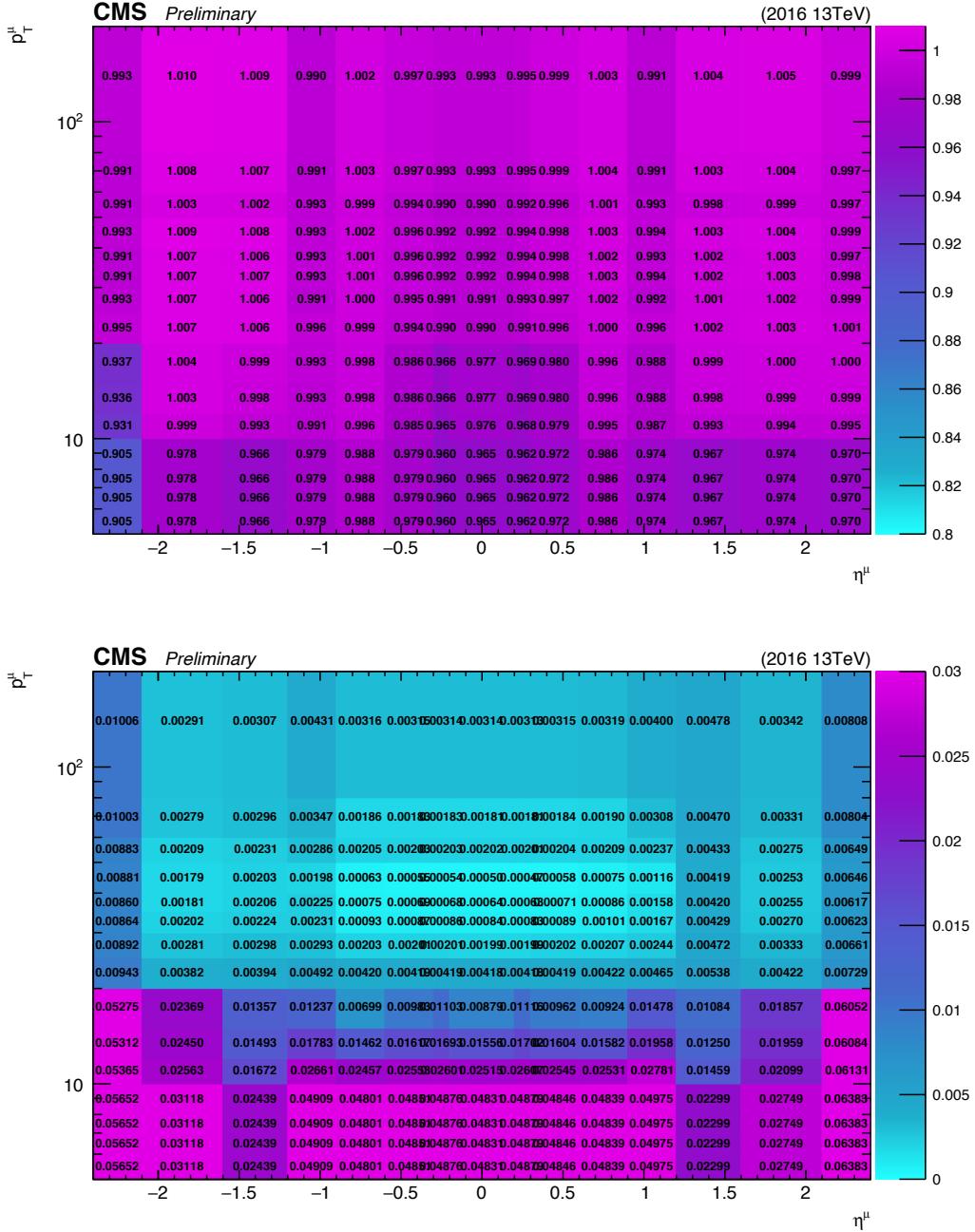


Figure 3-18: The histograms of overall data to simulation scale factors (reconstruction, identification, impact parameter and isolation requirements and tracking SF) and corresponding uncertainty.

### 3.3.2 Photon identification

MVA based ID with working point (W.P) 90% is used. This ID is trained on a sample of simulated  $\gamma + \text{jet}$  events, where the photon candidates matching the prompt photon are used as signal, and photon candidates not matching the prompt photon are identified as background. The input variables for the photon MVA training include the shower shapes variables, photon isolation, and charged hadron isolation. The general purpose MVA has two categories, one for photons in barrel (EB) region and the other for those in endcap (EE) region. The suggested cut values, 0.2 for both categories, result in 95.2% (93.9%) of signal efficiency for  $Z \rightarrow J/\psi \gamma$  events and 60.3% (67.3%) of background rejection power, defined as  $1 - \epsilon_{\text{Bkg}}$ , for the EB (EE) region. Here, the events selected in data are treated as background. Fig. [3-19] shows the ROC curves for photon MVA ID obtained from  $Z \rightarrow J/\psi \gamma$  signal events and data events (treated and labeled as background in the plots), the point corresponding to the 90% W.P for each category is shown as red solid star.

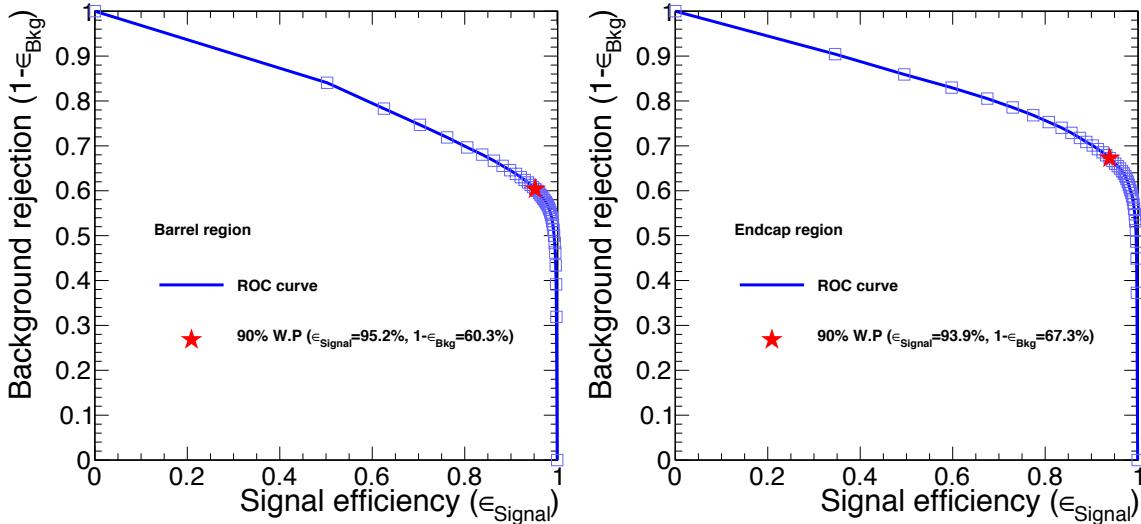


Figure 3-19: The ROC curves for photon MVA ID obtained from  $Z \rightarrow J/\psi \gamma$  signal events and data events for EB (left) and EE (right) category. The red solid star corresponds to the efficiency for 90% W.P.

The contamination of fake photons is estimated by checking the ratio of the  $Z + \text{jets}$  yields to the  $Z + \gamma$  yields. This gives a rough idea on the performance

of photon ID. It is found that the ratio of  $Z + \text{jets}/Z\gamma$  events is  $\sim 30\%$  for photon  $E_T$  between 33 and 40 GeV, and  $\sim 20\%$  for photon  $E_T$  between 60 and 80 GeV. The ratios of  $Z + \text{jets}/Z\gamma$  events in different photon  $E_T$  regions are summarized in Table 3.8.

photon $E_T$	$Z+\text{jets}/Z\gamma$ (in %)
$33 < E_T^\gamma < 40 \text{ GeV}$	30
$40 < E_T^\gamma < 50 \text{ GeV}$	28
$50 < E_T^\gamma < 60 \text{ GeV}$	22
$60 < E_T^\gamma < 80 \text{ GeV}$	21

Table 3.8: The ratios of  $Z+\text{jets}/Z\gamma$  events in different photon  $E_T$  regions.

Conversion safe electron veto (CSEV) is used to reject photons from electron conversions by requiring that there be no charged-particle track with a hit in the inner layer of the pixel detector associated to the photon cluster in the ECAL. The small number of inoperative sensors and possible cases where a track can pass between the first layer of sensors without leaving a hit are accounted for. The photon inefficiency is largely reduced and the residual comes from photons converting in the beam pipe. Up to 99.1% (97.8%) of photon in EB (EE) can pass CSEV, and 5.3% (19.6%) of electrons in EB (EE) can also satisfy this requirement. The efficiency of the photon identification is measured from  $Z \rightarrow ee$  events using tag-and-probe techniques, and found to be between 84 and 91% (77 and 94%), depending on the transverse energy  $E_T$ , in the barrel (endcap). The electron veto efficiencies are measured with  $Z \rightarrow \mu\mu\gamma$  events, where the photon is produced by final-state radiation, and found to be 98 (94%) in the barrel (endcap). The scale factors for the photon ID in bins of photon  $E_T$  and  $\eta_{SC}$  are shown in Fig. 3-20, and those for the CSEV are shown in Fig. 3-21.

## 3.4 Event Selection

In addition to the object identification and isolation, kinematic selections are applied to further discriminate the background.

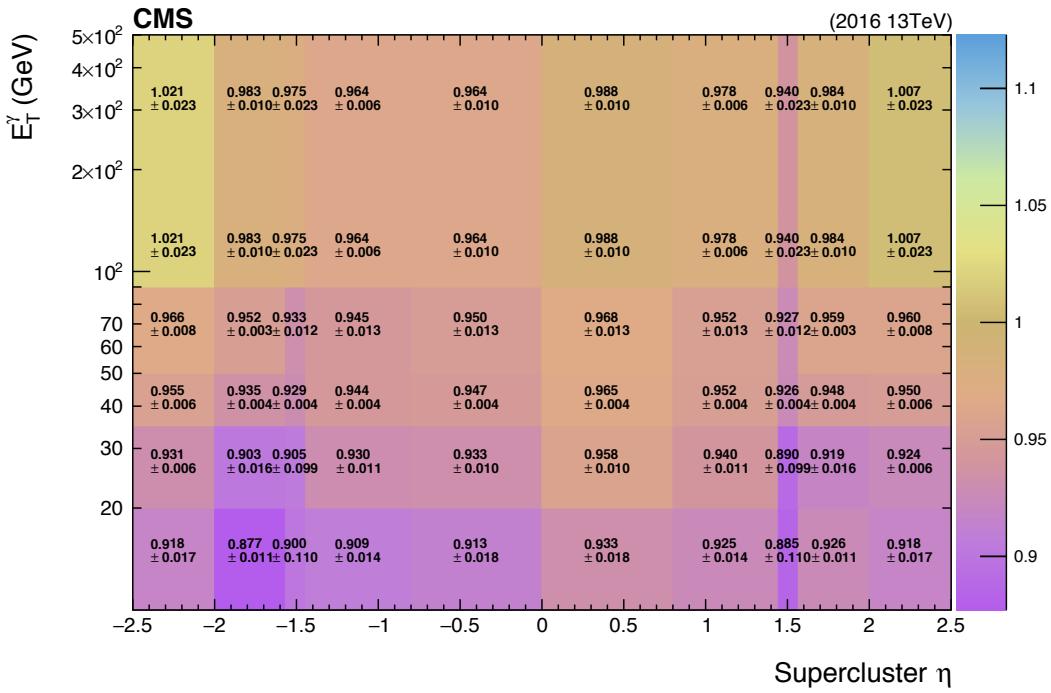


Figure 3-20: The scale factors in bins of photon  $E_T$  and  $\eta_{SC}$ .

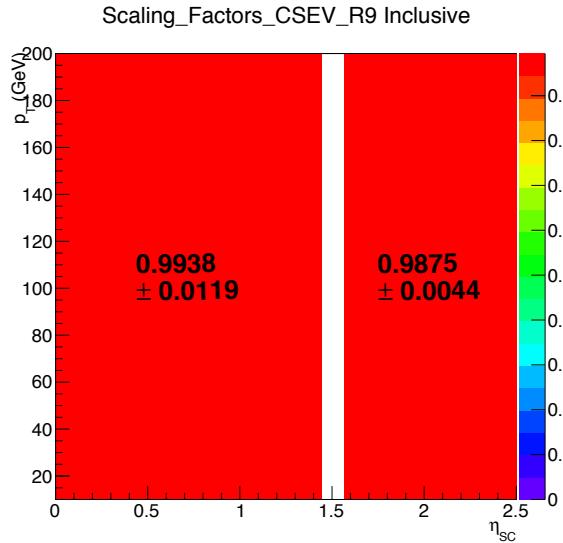


Figure 3-21: The scale factors of SCEV in bins of photon  $\eta_{SC}$ .

- Two opposite charged muons with  $p_T^{\mu_1} > 20$  GeV,  $p_T^{\mu_2} > 4$  GeV,  $|\eta^\mu| < 2.4$ .  
The  $p_T$  cut value on the leading muon is driven by the trigger threshold.
- $J/\psi$  candidate selection  $3.0 < m_{\mu\mu} < 3.2$  GeV.

- $E_T^\gamma > 33 \text{ GeV}$ ,  $|\eta_{SC}^\gamma| < 2.5$ , excluding the Barrel-Endcap transition region at  $1.4442 < |\eta_{SC}^\gamma| < 1.566$ . The  $E_T$  cut value on the photon is driven by the trigger threshold.
- $\Delta R(\mu_1, \gamma) > 1$ ,  $\Delta R(\mu_2, \gamma) > 1$ ,  $\Delta R(\mu\mu, \gamma) > 2$ , and  $|\Delta\phi(\mu\mu, \gamma)| > 1.5$ . The angular separation  $\Delta R$  cuts on each muon and the photon are imposed to suppress Drell-Yan process with FSR photon. As we do not have proper background MC samples, the cut values are determined such that a higher total signal efficiency is kept.
- $p_T^{\mu\mu}, E_T^\gamma / m_{\mu\mu\gamma} > 0.28$  (35/125) for  $H \rightarrow J/\psi \gamma$ , 0.384 (35/91.2) for  $Z \rightarrow J/\psi \gamma$ . If a hard cut on  $E_T$  or  $p_T^{\mu\mu}$  is imposed, there will be an obvious turn-on at the  $Z$  mass region, as shown in Fig 3-22, which will complicate the background model. This ratio cut also helps to reject the  $\gamma^* + \text{jet}$  and  $\gamma + \text{jet}$  backgrounds. As for the cut value, 91.2 and 125.0 GeV are the nominal mass of the  $Z$  and Higgs boson respectively.

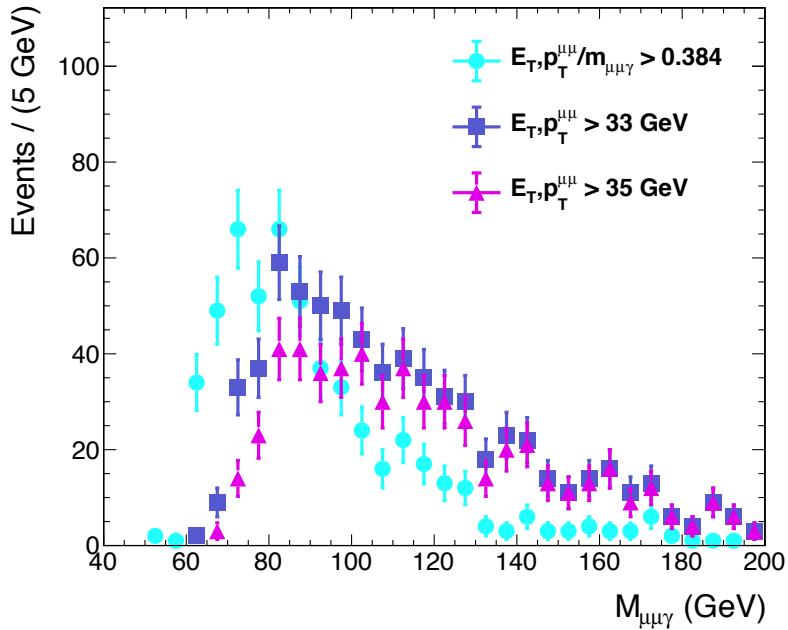


Figure 3-22:  $m_{\mu\mu\gamma}$  distributions with different forms of  $p_T$  or  $E_T$  cuts.

Table 3.9 summarizes event selections in this analysis

Trigger : HLT_Mu17_Photon30_CaloIdL_L1ISO_v
Muon identification, Particle Flow Isolation in cone 0.3 for $\mu_{\text{lead}} < 0.35$
$p_T^{\mu_{\text{lead}}} > 20 \text{ GeV}$ , $p_T^{\mu_{\text{trai}}} > 4 \text{ GeV}$ , $ \eta_\mu  < 2.4$
Photon MVA ID(90% WP), $E_T^\gamma > 33 \text{ GeV}$
$ \eta_{SC}^\mu  < 2.5$ , excluding those in Barrel-Endcap transition region of ECAL.
$\Delta R(\mu_1, \gamma) > 1$ , $\Delta R(\mu_2, \gamma) > 1$ , $\Delta R(\mu\mu, \gamma) > 2$ , and $ \Delta\phi(\mu\mu, \gamma)  > 1.5$
$3.0 < m_{\mu\mu} < 3.2 \text{ GeV}$
$p_T^{\mu\mu}/m_{\mu\mu\gamma} > 0.384(0.28)$ , $E_T^\gamma/m_{\mu\mu\gamma} > 0.384(0.28)$ for the Z (H) $\rightarrow J/\psi \gamma$ .

Table 3.9: The selection requirements in this analysis, including ID, isolation and kinematic selection.

In the  $Z \rightarrow J/\psi \gamma$  search, selected events are classified into mutually exclusive categories in order to enhance the sensitivity of the search. The categorization is based on the  $\eta$  of the photon and the photon  $R_9$  variable (defined as the energy sum of  $3 \times 3$  ECAL crystals centered on the most energetic crystal in the supercluster divided by the energy of the supercluster). Unconverted photons have high values of  $R_9$  and a threshold of 0.94 is used to classify reconstructed photons with high  $R_9$  (thus with a better resolution) and low  $R_9$  (worse resolution). The background is larger in the converted photon category. The three categories are: photon in the barrel region with a high  $R_9$  value (referred to as EB high  $R_9$ ); photon in the barrel region with low  $R_9$  value (referred to as EB low  $R_9$ ); photon in the endcap region (referred to as EE). The EE category is not divided into high/low  $R_9$  because there are few events in this category. By this categorization, this improvement on the search limit is  $\sim 2.0\%$ . In the  $H \rightarrow J/\psi \gamma$  search events are not divided into categories. The statistical sample is simply not large enough to overcome statistical fluctuations if it is further subdivided. As a result, we do not expect that the  $H \rightarrow J/\psi \gamma$  sample with limited statistics can define the shape of our discriminating variable,  $m_{\mu\mu\gamma}$ , used in the hypothesis test that differs between the categories. The possibility of splitting the EE category is investigated, but this indeed does not result in a significant improvement.

The exact definition of the three event categories in  $Z \rightarrow J/\psi \gamma$  search are shown in Table 3.10. The table includes the fractions of expected events in each category

for signal and of the observed events for data. The  $\sigma_{\text{eff}}$  of the  $m_{\mu\mu\gamma}$  distribution of each category is also included.

	Category 1 $0 <  \eta_\gamma^{\text{SC}}  < 1.4442$ $R_9 > 0.94$	Category 2 $0 <  \eta_\gamma^{\text{SC}}  < 1.4442$ $R_9 > 0.94$	Category 3 $1.566 <  \eta_\gamma^{\text{SC}}  < 2.5$ -
Data	40.3%	36.2%	23.5%
Signal	49.0%	30.6%	20.3%
$\sigma_{\text{eff}}$	3.58 GeV	3.86 GeV	4.08 GeV

Table 3.10: Definition of the three event classes in  $Z \rightarrow J/\psi \gamma$  and the fraction of selected events in signal and data. The expected mass resolution on the signal are also shown.

Table 3.11 summarizes the expected number of events from signals and observed yields in data in steps of event selection of both the Higgs and Z boson decays. For the Z boson decays, the numbers are with the unpolarized  $J/\psi$  assumption and  $p_T$  reweighting. Table 3.12 shows the impacts of different polarization scenarios and the Z  $p_T$  reweighting. The variations on the yields resulting from the extreme polarization assumption is -7.8% (transverse) to +16% (longitudinal), corresponding to the total signal efficiency varying from 13.1% to 16.4%. The Z  $p_T$  reweighting, with weights derived from the *aMC@NLO* sample, results in +2.3% of increase on the expected yields of the Z decay. The difference between the yield with weights derived from the *aMC@NLO* sample and that from *POWHEG* is only 0.13%, and no additional uncertainty is assigned. In both Z and Higgs decays the number of events coming from the peaking background  $H(Z) \rightarrow \mu\mu\gamma$  is large compared to signal processes. On the other hand, it is small compared to the total background. Hence, it has minimal effect on the upper limit on  $\mathcal{B}(H(Z) \rightarrow J/\psi \gamma)$ . With the constraint  $100$  (70)  $< m_{\mu\mu\gamma} < 150$  (120) GeV, the total signal efficiency, including kinematic acceptance, trigger and reconstruction efficiencies, and  $p_T$  reweighting for the Z boson decay, of about 22.6% and 14.2% in Higgs and Z boson decays. The difference in the total signal efficiency between the Higgs and the Z boson decay is mostly due to the kinematic acceptance, which comes from the difference in  $p_T$  distributions of muons and photon given that the Z boson is lighter

than the Higgs boson.

	$H \rightarrow J/\psi \gamma$			$Z \rightarrow J/\psi \gamma$		
	Data	$H \rightarrow J/\psi \gamma$ signal	$H \rightarrow \gamma^*\gamma$ background	Data	$Z \rightarrow J/\psi \gamma$ signal	$Z \rightarrow \mu\mu\gamma$ background
Total (Before selection)	170M	0.350	91.7	170M	10.8	3335
HLT	30.3M	0.190	51.3	30.3M	4.24	1932
Muon selection	650K	0.136	35.9	650K	2.67	1317
Photon selection	152K	0.116	30.7	152K	2.17	1066
$\Delta R, \Delta\phi$	59.4K	0.101	23.5	59.4K	2.09	1020
$m_{\mu\mu}$	1088	0.0929	0.274	1088	1.93	5.29
$m_{\mu\mu\gamma}$	363	0.0928	0.273	637	1.90	5.37
$p_T^{\mu\mu}, E_T^{\gamma}/m_{\mu\mu\gamma}$	279	0.0884	0.257	384	1.58	4.57
Expected signal yields (with the pileup weight, all the scale factors and efficiencies)						
All	279	0.0765	0.207	384	1.54	4.47
Cat1	-	-	-	148	0.770	2.14
Cat2	-	-	-	144	0.468	1.20
Cat3	-	-	-	92	0.299	1.12

Table 3.11: The expected signal yield and the number of selected events in data, for the integrated luminosity of  $35.9 \text{ fb}^{-1}$ .

	Inclusive Yield	Inclusive Difference (in %)
<b>unpolarized &amp; with <math>p_T</math> reweighting</b>	1.54	
<b>transversely polarized &amp; with <math>p_T</math> reweighting</b>	1.42	-7.86
<b>longitudinally polarized &amp; with <math>p_T</math> reweighting</b>	1.78	+15.7
<b>unpolarized &amp; without <math>p_T</math> reweighting</b>	1.50	-2.24
<b>transversely polarized &amp; without <math>p_T</math> reweighting</b>	1.38	-9.85
<b>longitudinally polarized &amp; without <math>p_T</math> reweighting</b>	1.74	+13.0

Table 3.12: Summary of the impacts of different polarization scenarios and the Z  $p_T$  reweighting.

Fig. 3-23 and 3-24 show the  $m_{\mu\mu}$  distributions in  $H \rightarrow J/\psi \gamma$  (top plots in Fig. 3-23), Cat1 of  $Z \rightarrow J/\psi \gamma$  (bottom plots in Fig. 3-23), Cat2 of  $Z \rightarrow J/\psi \gamma$  (top plots in Fig. 3-24), and Cat3 of  $Z \rightarrow J/\psi \gamma$  (bottom plots in Fig. 3-24). The black points with error bars are distributions in data, while the filled histograms are distributions in signal events. Plots on the left hand side are with the  $m_{\mu\mu}$  constraint, while those on the right hand side are not. The peak at the  $J/\psi$  mass in data shows that real  $J/\psi$  candidates are reconstructed and selected. These events come from inclusive quarkonium production, for which no simulation is available. The backgrounds from  $H \rightarrow \gamma^*\gamma$  and  $Z \rightarrow \mu\mu\gamma$  events, for which there is a simulation, are much smaller than that from inclusive quarkonium production and they are scaled to

	$H \rightarrow J/\psi \gamma$ signal					
	ggF	VBF	ZH	$W^+H$	$W^-H$	ttH
Total (Before selection)	0.307	0.0240	0.00596	0.00565	0.00360	0.00334
HLT	0.167	0.0132	0.00303	0.00279	0.00193	0.00226
Muon selection	0.119	0.00939	0.00216	0.00198	0.00139	0.00168
Photon selection	0.103	0.00803	0.00178	0.00161	0.00114	0.00125
$\Delta R, \Delta\phi$	0.0925	0.00480	0.00110	0.00100	0.000742	0.000510
$m_{\mu\mu}$	0.0858	0.00442	0.000938	0.000784	0.000594	0.000351
$m_{\mu\mu\gamma}$	0.0858	0.00442	0.000932	0.000776	0.000589	0.000330
$p_T^{\mu\mu}, E_T^\gamma/m_{\mu\mu\gamma}$	0.0820	0.00401	0.000855	0.000714	0.000541	0.000305
Expected signal yields (with the pileup weight, all the scale factors and efficiencies)	0.0710	0.00352	0.000711	0.000597	0.000454	0.000266

Table 3.13: The expected signal yield for each Higgs production mode.

make it visible. Figures 3-25, 3-26, 3-27, 3-28 show the distributions of kinematic variables in  $H \rightarrow J/\psi \gamma$ , Cat1, Cat2, and Cat3 of  $Z \rightarrow J/\psi \gamma$ . The variables shown are :  $p_T$  of leading muon,  $p_T$  of trailing muon,  $E_T$  of photon,  $\eta$  of leading muon,  $\eta$  of trailing muon,  $\eta_{SC}$  of photon,  $p_T$  of reconstructed dimuon system,  $\Delta R$  between two muons, and  $\Delta R$  between leading muon and photon.

The normalization of each distribution from data events is the number of events selected in the corresponding category. The number of events in distributions from signal simulated events are normalized to 750 (40) times the SM prediction for Higgs (Z) decays. The number of events in distributions from peaking background MC events are normalized to 150 (5) times their SM expectation for Higgs (Z) decays. These scale factors in the plots are chosen to give better visualization.

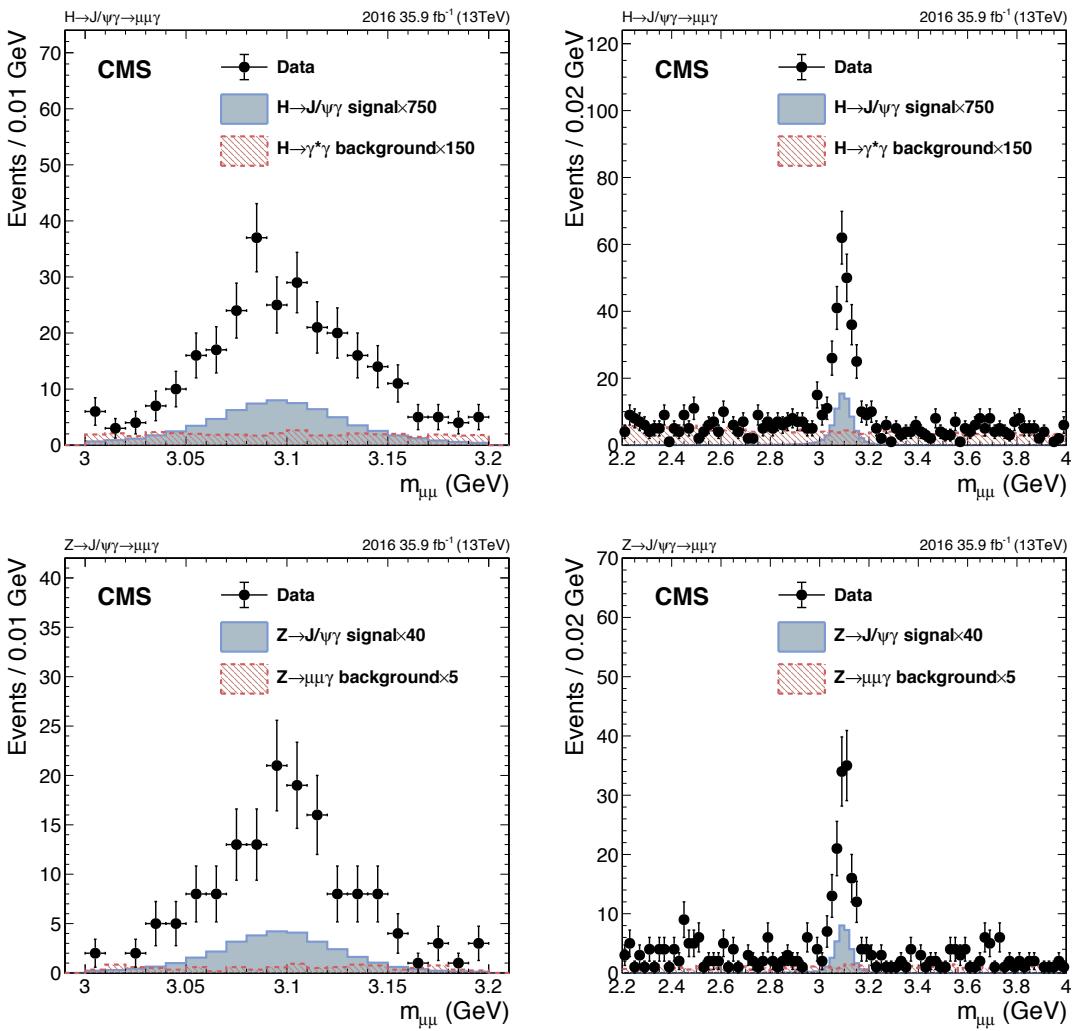


Figure 3-23: The  $m_{\mu\mu}$  distributions from data and signal events of:  $H \rightarrow J/\psi \gamma$  (top), Cat1 of  $Z \rightarrow J/\psi \gamma$  decay (bottom).

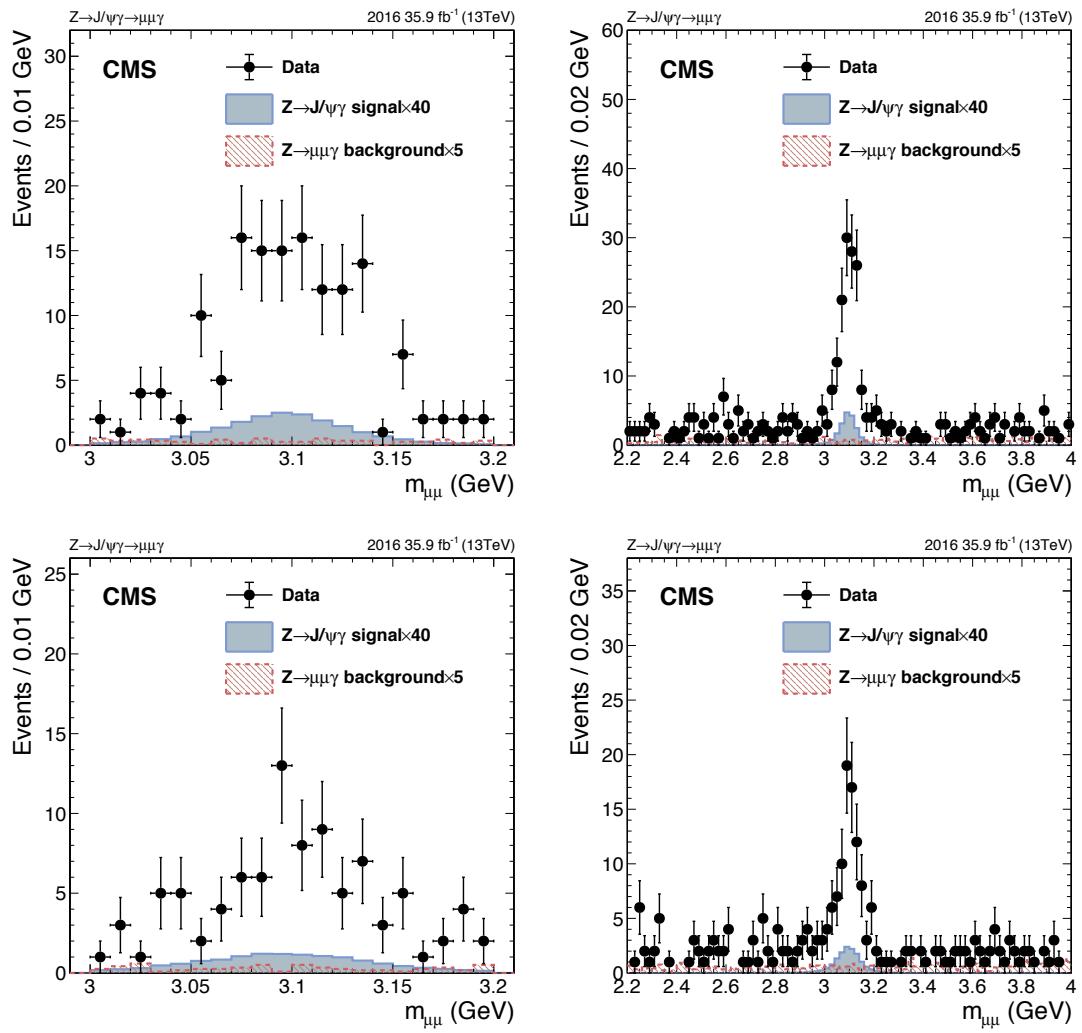


Figure 3-24: The  $m_{\mu\mu}$  distributions from data and signal events of: Cat2 of  $Z \rightarrow J/\psi \gamma$  decay (top), and Cat3 of  $Z \rightarrow J/\psi \gamma$  decay (bottom).

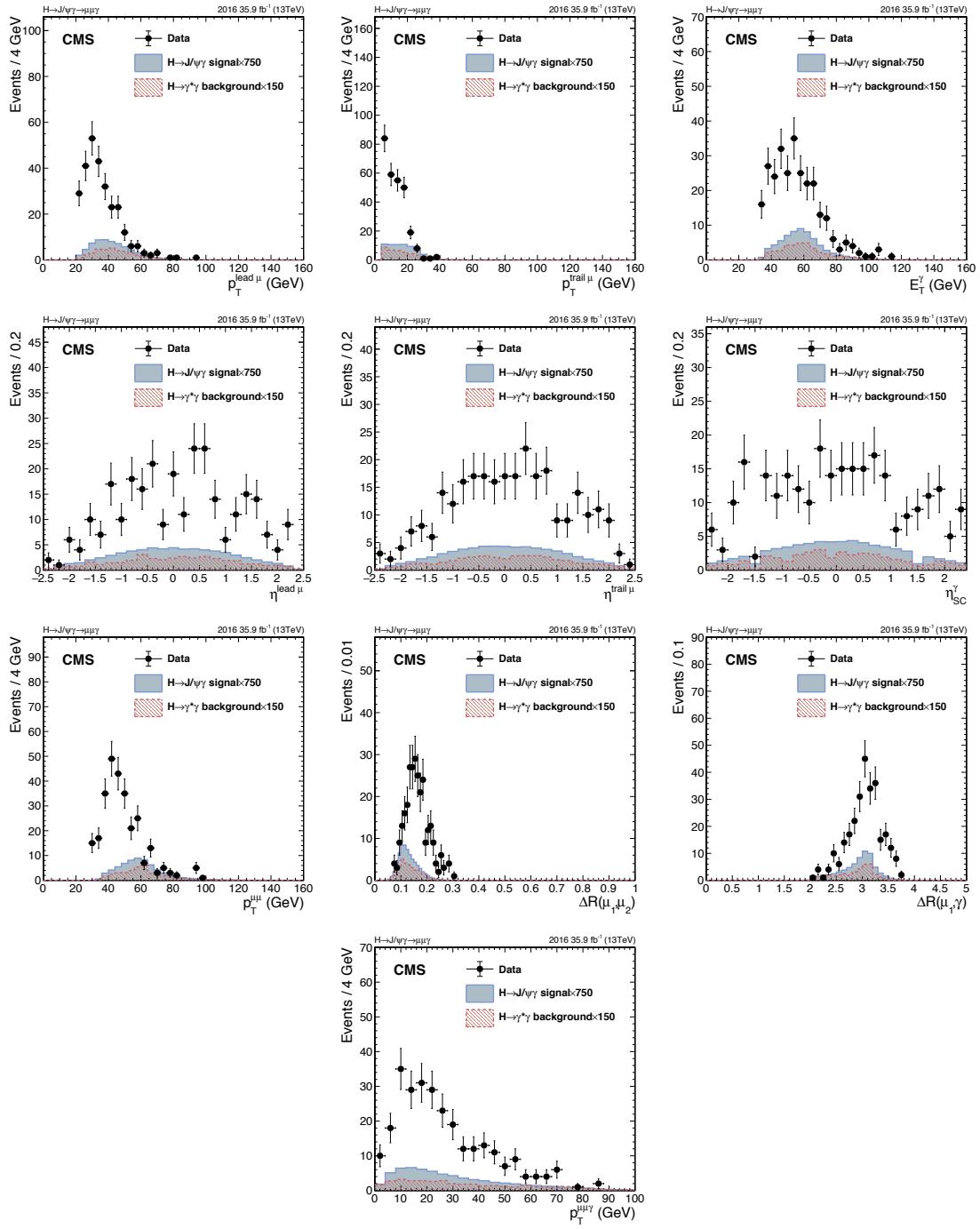


Figure 3-25: Distributions of the key variables from data and signal events in  $H \rightarrow J/\psi \gamma$  decay. Transverse momenta of the muons and the photon; pseudorapidity of the muons and the photon; transverse momenta of the dimuon system; distances  $\Delta R$  between the two muons and between the leading muon and the photon; the transverse momenta of the three-body system,  $p_T^{\mu\mu\gamma}$

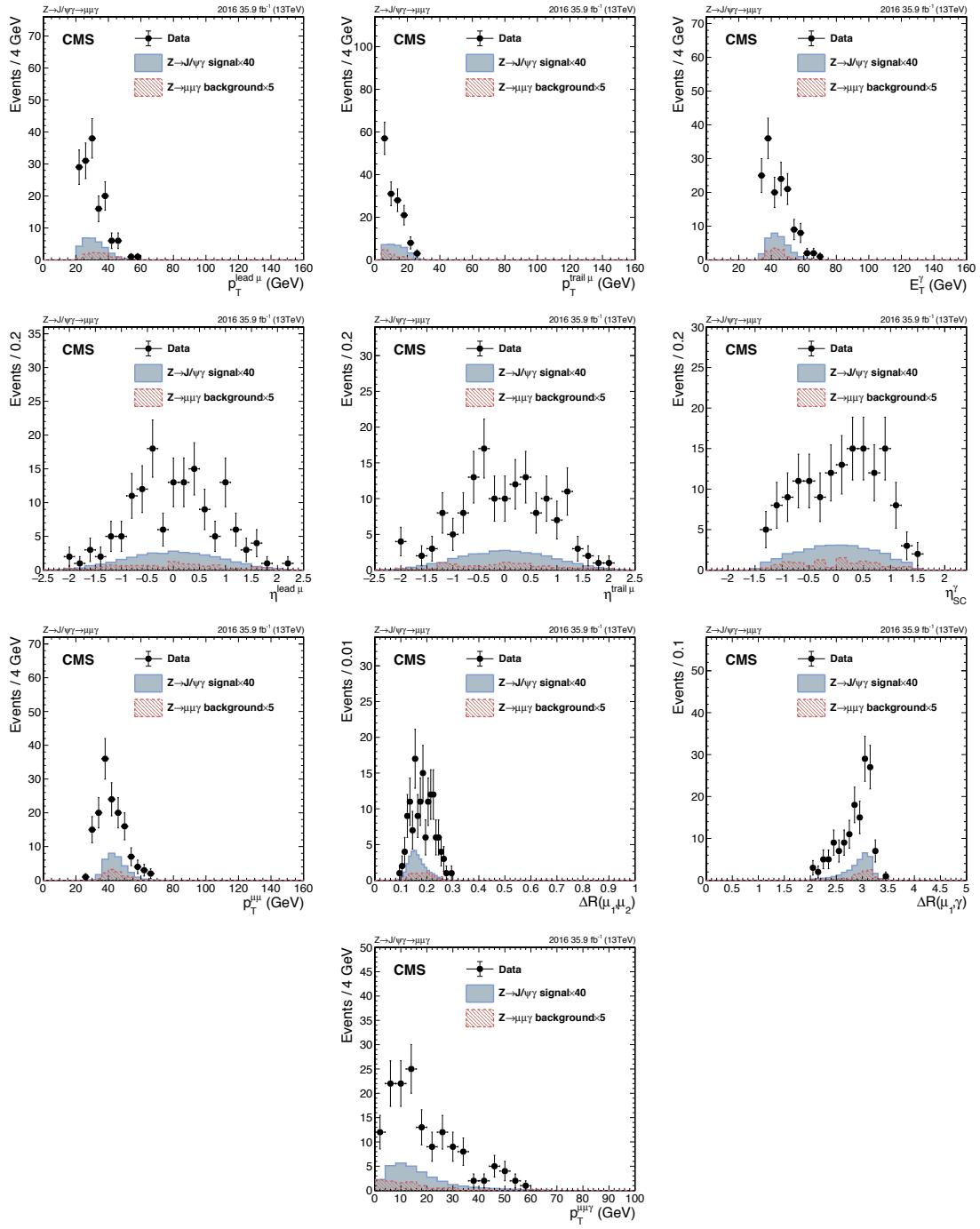


Figure 3-26: Distributions of the key variables from data and signal events of Cat1 in  $Z \rightarrow J/\psi \gamma$  decay. Transverse momenta of the muons and the photon; pseudorapidity of the muons and the photon; Transverse momenta of the dimuon system; distances  $\Delta R$  between the two muons and between the leading muon and the photon; the transverse momenta of the three-body system,  $p_T^{\mu\mu\gamma}$ .

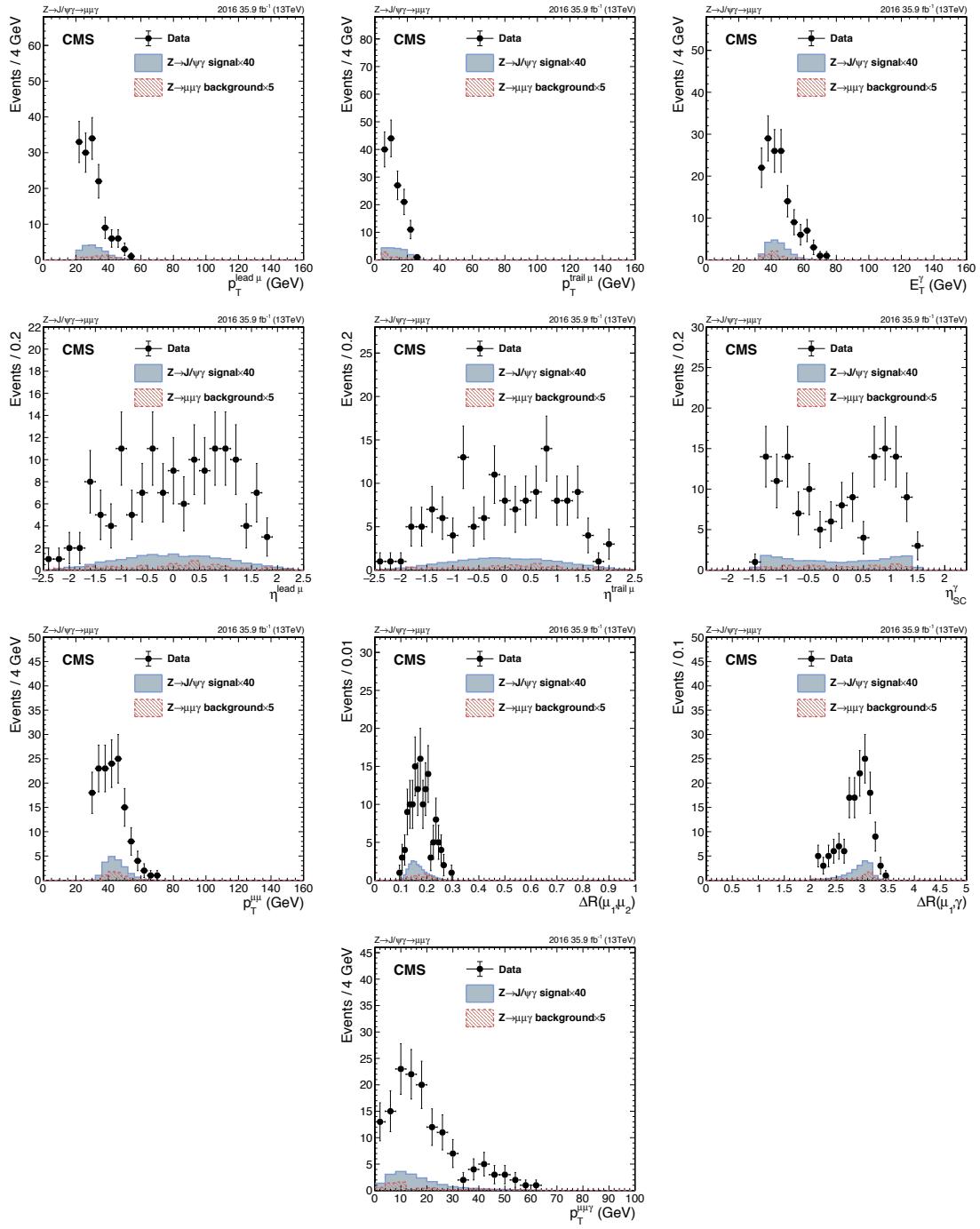


Figure 3-27: Distributions of the key variables from data and signal events of Cat2 in  $Z \rightarrow J/\psi \gamma$  decay. Transverse momenta of the muons and the photon; pseudorapidity of the muons and the photon; Transverse momenta of the dimuon system; distances  $\Delta R$  between the two muons and between the leading muon and the photon; the transverse momenta of the three-body system,  $p_T^{\mu\mu\gamma}$ .

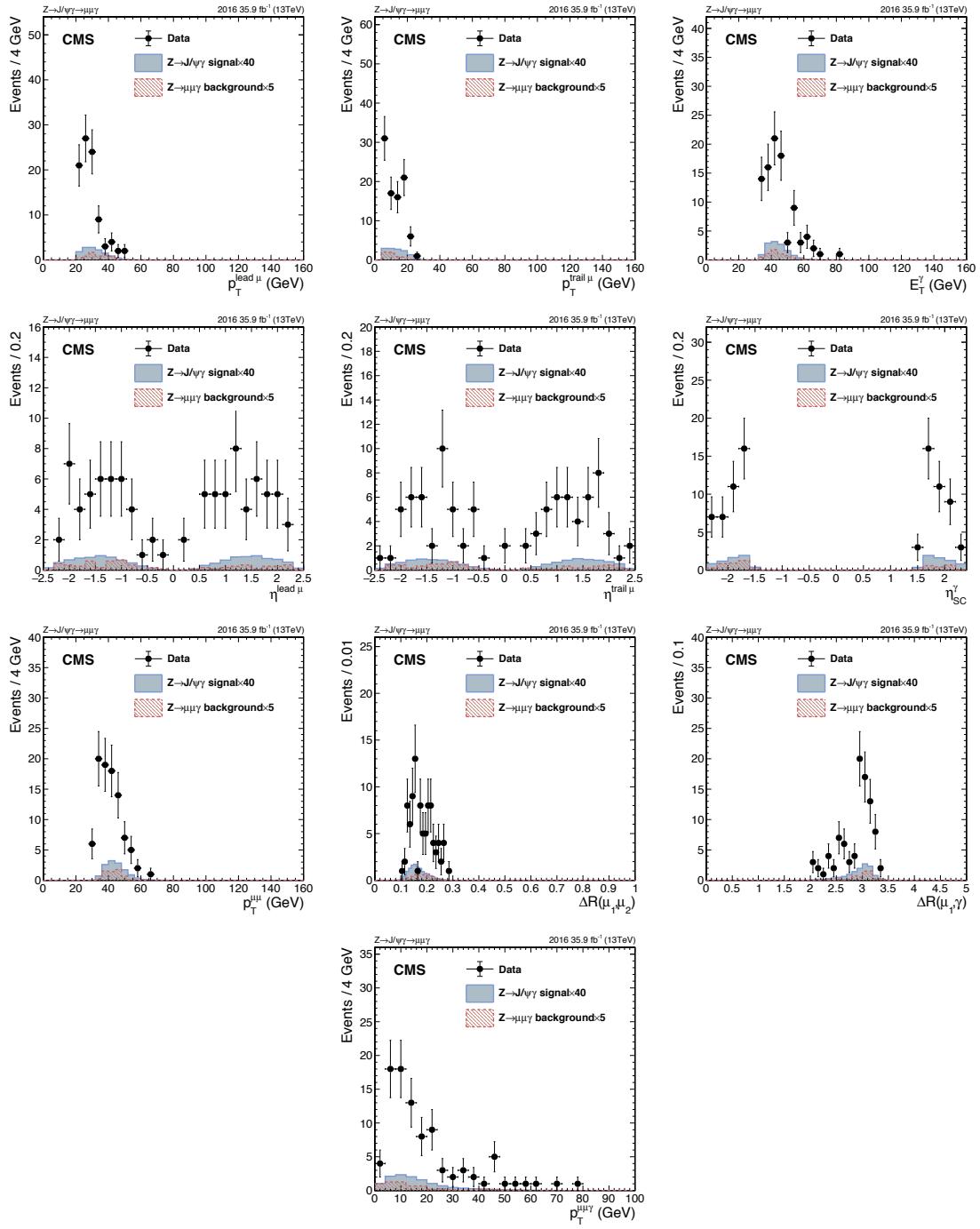


Figure 3-28: Distributions of the key variables from data and signal events of Cat3 in  $Z \rightarrow J/\psi \gamma$  decay. Transverse momenta of the muons and the photon; pseudorapidity of the muons and the photon; Transverse momenta of the dimuon system; distances  $\Delta R$  between the two muons and between the leading muon and the photon. the transverse momenta of the three-body system,  $p_T^{\mu\mu\gamma}$ .

A study of the muon vertex is done to ensure whether the reconstructed  $J/\psi$  after the full selection are promptly produced at the pp interaction point, referred to as “prompt  $J/\psi$ ”, and not from the displaced heavy hadron decays, referred to as “non-prompt  $J/\psi$ ”. It is expected that in signal events the  $J/\psi$  are produced promptly since the lifetimes of the Z and Higgs boson are very short.

Vertex-related variables examined in this study are:

- Dimuon vertex position ( $x$ ,  $y$  and  $z$  coordinates)
- The transverse decay length  $L_{xy} = \frac{\vec{r}_T \cdot \vec{p}_T^{J/\psi}}{|\vec{p}_T^{J/\psi}|}$ , where  $\vec{r}_T$  is the vector from PV to the dimuon vertex in transverse plane.
- $R_{xy} = |L_{xy}|$
- $SL_{xy} = |L_{xy}|/\sigma(L_{xy})$ . The significance of the  $L_{xy}$  is defined as the absolute value of  $L_{xy}$  divided by its error  $\sigma(L_{xy})$ .
- $Cos(\alpha)$ , where  $\alpha$  is defined as the angle between the reconstructed momentum vector of the dimuon system and the vector from the PV to the dimuon vertex.
- Dimuon vertex  $\chi^2$ , one of the indicators of the goodness of the fit
- Dimuon vertex probability, which is the chi-square probability given the dimuon vertex  $\chi^2$  and the number of degree of freedom in the fit.
- Validity of the dimuon vertex. The vertex returned may not be valid in some cases. The status of the vertex will be invalid when the maximum number of iterations is exceeded or the fitted position is out of the tracker bounds.
- Proper decay time  $t = \frac{m_{J/\psi}}{p_T^{J/\psi}} \cdot L_{xy}$ , where the  $m_{J/\psi}$  is the mass of the reconstructed  $J/\psi$  candidate. The proper time can be negative by the fact that the  $L_{xy}$  is defined either to be positive or negative. The positive (negative) value indicates that the angle between the  $L_{xy}$  vector and the vector of  $p_T^{J/\psi}$  is smaller (larger) than  $\pi/2$ .

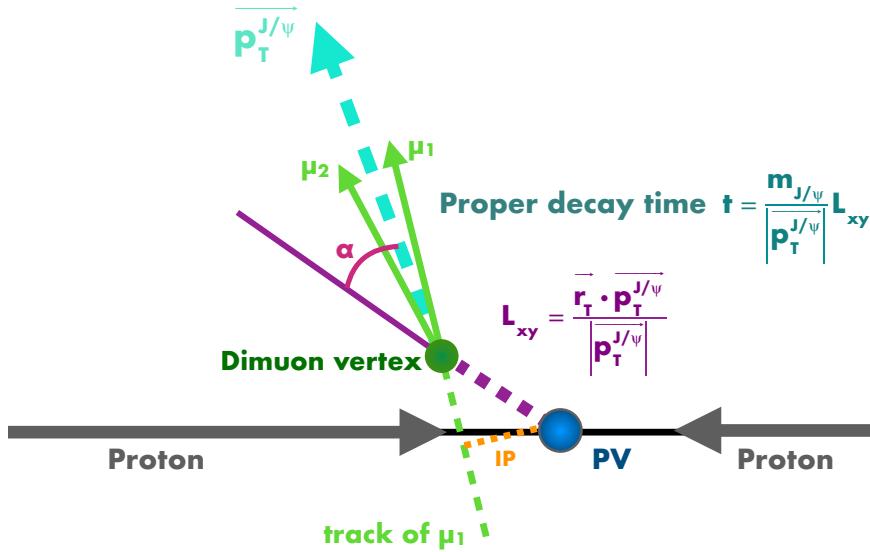


Figure 3-29: Schematic figure for vertex variables.

The distributions of the vertex-related variables from data (in black points with error bars) and signal (filled histograms) for the Higgs and Z boson searches are shown in Figs. 3-30, 3-31, 3-32, and 3-33. These distributions are normalized to the number of selected events in data. The distributions suggest that the  $J/\psi$  candidates reconstructed in data, like the signal events, are produced promptly at the  $p\bar{p}$  interaction point, rather than coming from displaced heavy hadron decays. Based on the above-mentioned argument, no additional requirement associated with these vertex variables is imposed any, since the  $d_{xy}$ ,  $d_z$ , and the  $SIP_{3D}$  cuts already reject non-prompt  $J/\psi$ .

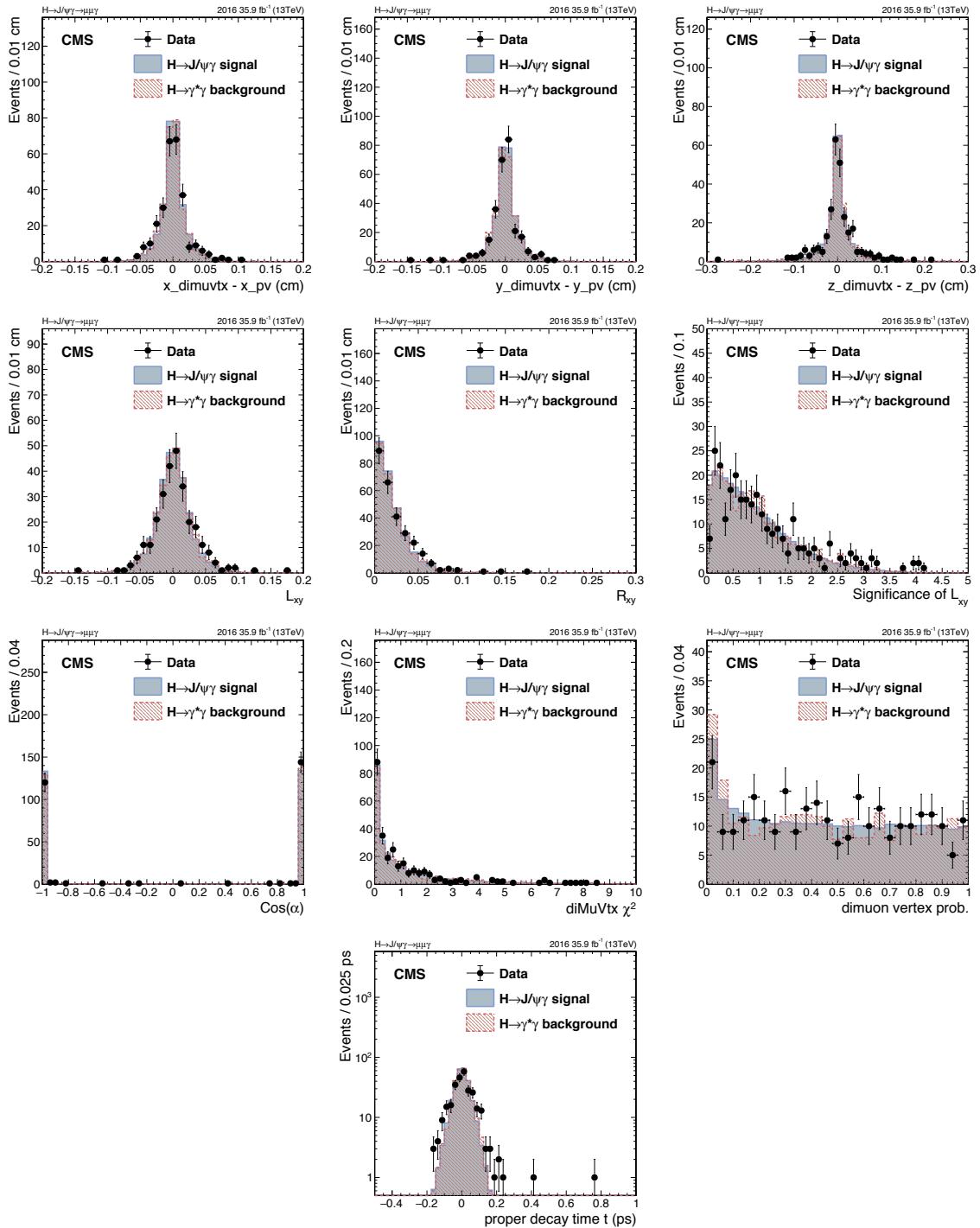


Figure 3-30: Distributions of the vertex-related variables from data and signal events in  $H \rightarrow J/\psi \gamma$  decay.

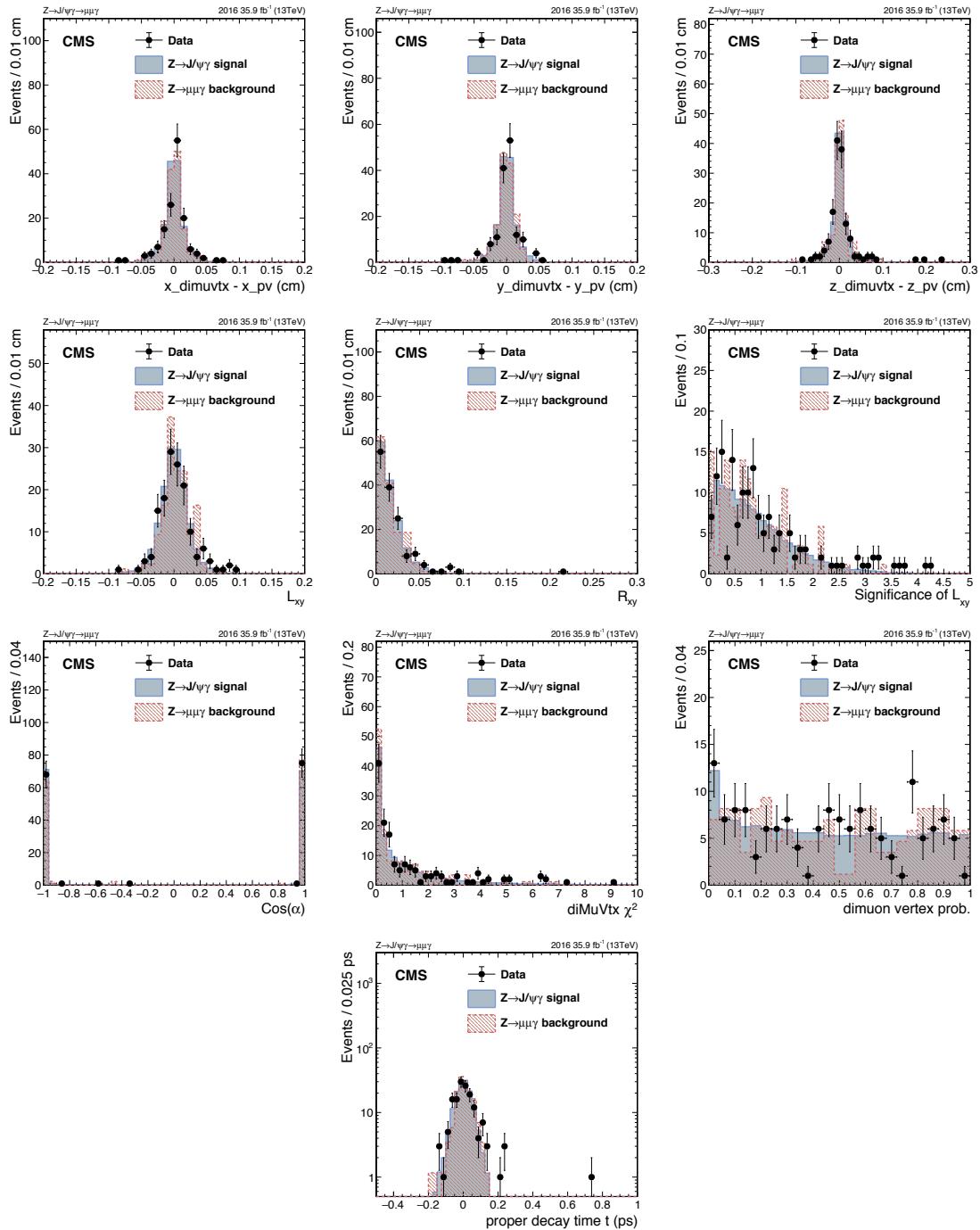


Figure 3-31: Distributions of the vertex-related variables from data and signal events of Cat1 in  $Z \rightarrow J/\psi \gamma$  decay.

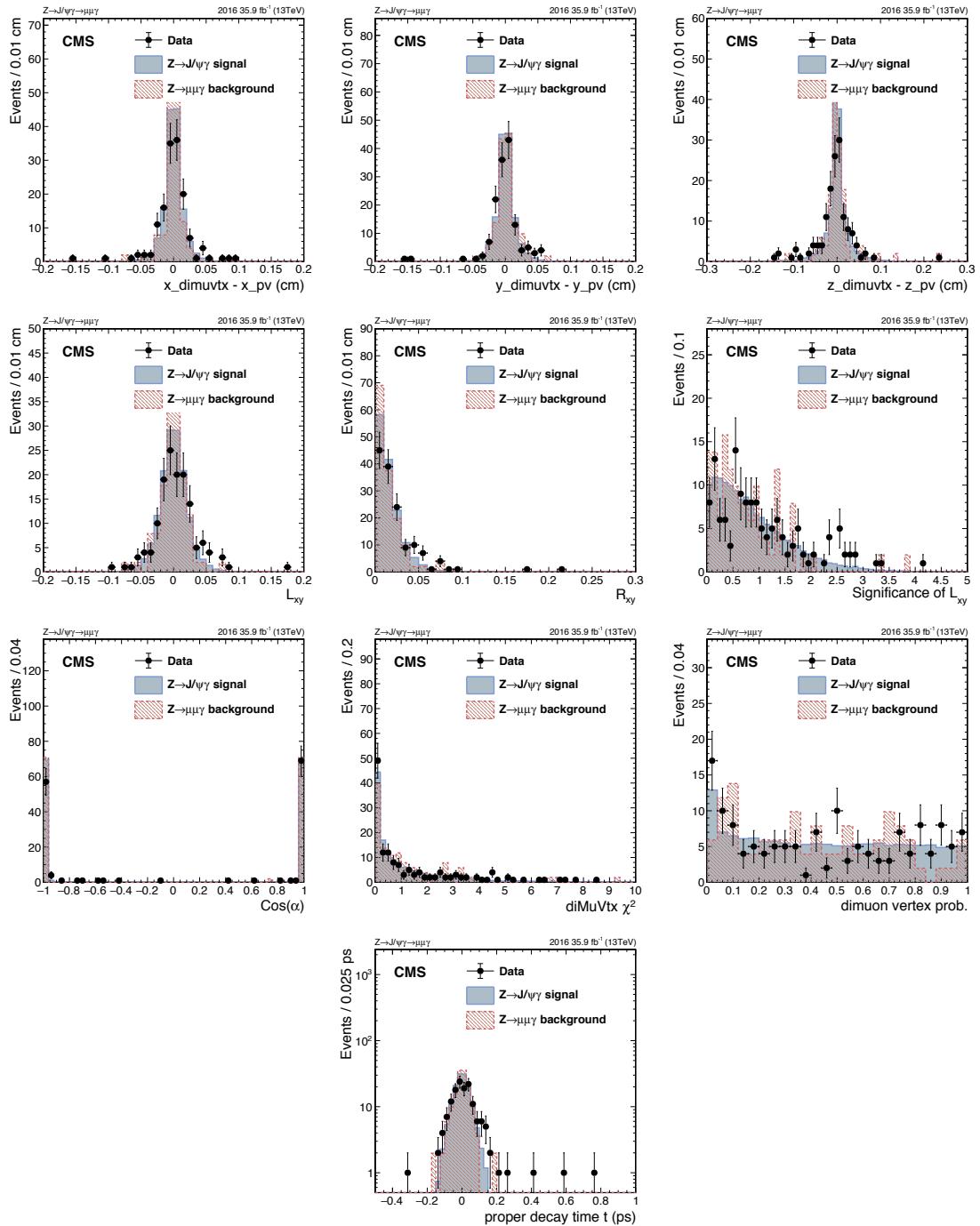


Figure 3-32: Distributions of the vertex-related variables from data and signal events of Cat2 in  $Z \rightarrow J/\psi \gamma$  decay.

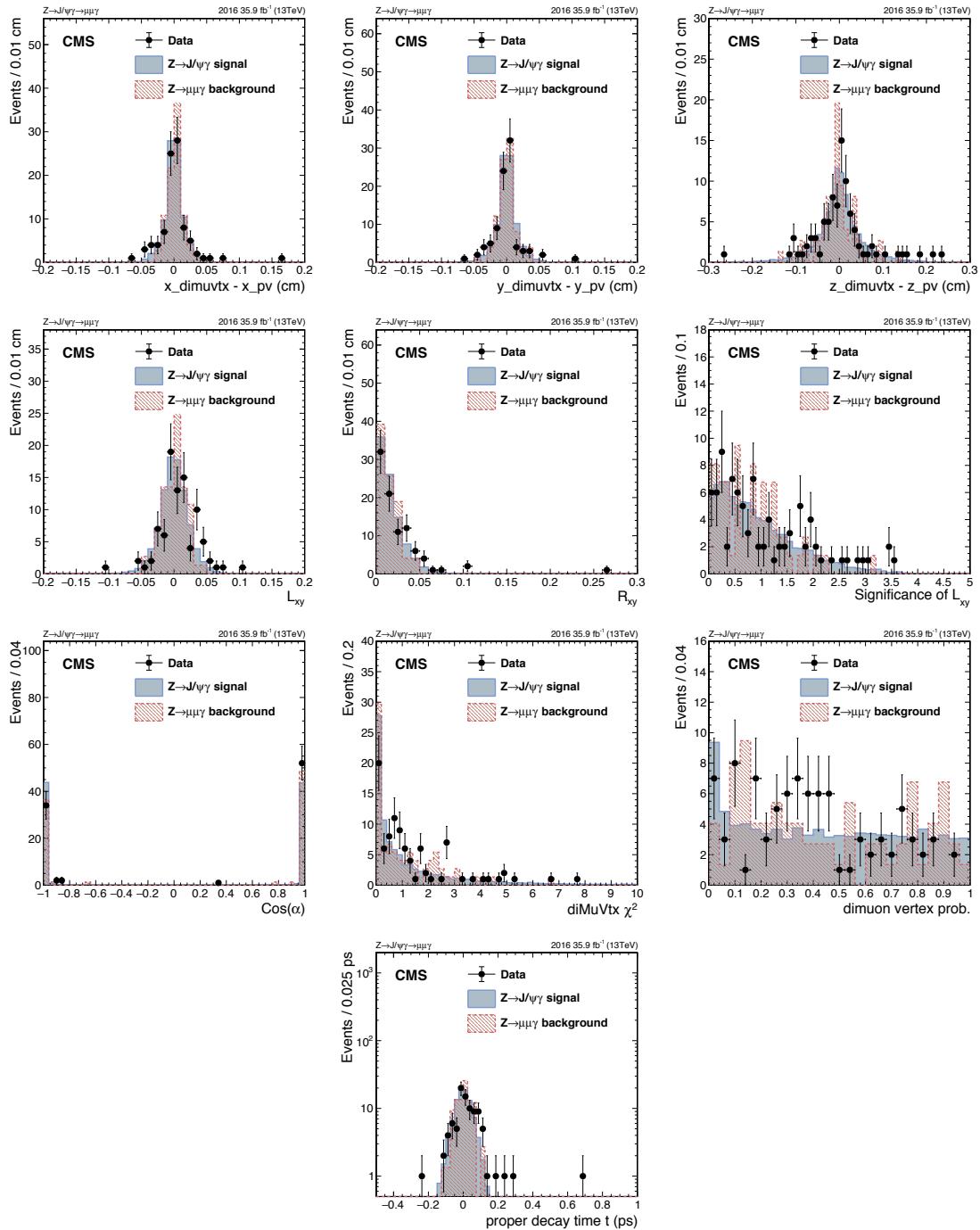


Figure 3-33: Distributions of the vertex-related variables from data and signal events of Cat3 in  $Z \rightarrow J/\psi \gamma$  decay.

### 3.5 Background modeling

While the sub-dominant, peaking, backgrounds are estimated from the simulated samples, the dominant continuum background of each category of both the Z and Higgs boson decays is estimated and modeled from data by fitting parameteric functions to the  $m_{\mu\mu\gamma}$  distributions. An un-binned maximum likelihood fit is performed over the range  $70$  ( $100$ )  $< m_{\mu\mu\gamma} < 120$  ( $150$ ) GeV for the  $Z$  ( $H$ )  $\rightarrow J/\psi \gamma$  search.

The following functions are considered:

- Bernstein polynomials of order  $N$  (NPol)

$$\text{Bern}_N(m_{\mu\mu\gamma}) = \sum_{i=1}^N f_i^2 \binom{N}{i} m_{\mu\mu\gamma}^i (1 - m_{\mu\mu\gamma})^{N-i} \quad (3.4)$$

with  $N$  free parameters.

- A sum of  $N$  exponential functions

$$N\text{Exp}(m_{\mu\mu\gamma}) = \sum_{i=1}^N f_i e^{p_i (m_{\mu\mu\gamma})} \quad (3.5)$$

with  $2N - 1$  free parameters:  $p_i < 0$  and  $f_i$ . The lowest order considered has  $N = 1$ , i.e. one term.

- The sum of  $N$  power-functions

$$N\text{Pow}(m_{\mu\mu\gamma}) = \sum_{i=1}^N f_i (m_{\mu\mu\gamma})^{p_i}, \quad (3.6)$$

with  $2N - 1$  free parameters  $p_i < 0$  and  $f_i$ . The lowest order considered has  $N = 1$ , i.e. one term.

- Laurent series with 2, 3 and 4 terms

$$2\text{Lau}(m_{\mu\mu\gamma}) = f_2 (m_{\mu\mu\gamma})^{-4} + f_3 (m_{\mu\mu\gamma})^{-5}, \quad (3.7)$$

$$3\text{Lau}(m_{\mu\mu\gamma}) = f_1(m_{\mu\mu\gamma})^{-3} + f_2(m_{\mu\mu\gamma})^{-4} + f_3(m_{\mu\mu\gamma})^{-5}, \quad (3.8)$$

and

$$4\text{Lau}(m_{\mu\mu\gamma}) = f_1(m_{\mu\mu\gamma})^{-3} + f_2(m_{\mu\mu\gamma})^{-4} + f_3(m_{\mu\mu\gamma})^{-5} + f_4(m_{\mu\mu\gamma})^{-6}, \quad (3.9)$$

with  $N$  free parameters  $f_1 \dots f_4$ .

Fits to the  $m_{\mu\mu\gamma}$  distributions in data from the Higgs and Z boson decays using different functions are shown on Fig 3-34. To choose the best fit function out of the above-mentioned families of functions, a F-test is performed and follows with the bias study. F-test is performed for all the functions except for Bernstein polynomials. For Bernstein family, the bias study is performed all the orders up to order 6.

### 3.5.1 F-test

To choose the best fit order from a family of functions, a F-test on data is performed. First, for a given family, the lowest order function in that family is fit to a single category. Then, the next order function is fit to the data in the same category. The difference of twice the negative log-likelihood(NLL) between the two fits,  $2\Delta NLL_{N+1} = 2(NLL_{N+1} - NLL_N)$ , indicates the improvement of the fit and whether or not the data support the hypothesis of the higher order function. This argument is made by the fact that the  $2\Delta NLL_{N+1}$  should be distributed as a  $\chi^2$  distribution of  $M$  degrees of freedom, where  $M$  is the difference in the number of free parameters in the  $(N+1)_{th}$ -order function and  $N_{th}$ -order function. For example, for exponential family,  $M = [2(N+1) - 1] - [2(N) - 1] = 2$ , while for the Bernstein polynomials  $M = (N+1) - (N) = 1$ . A p-value is defined and calculated as

$$\text{p-value} = p(2\Delta NLL > 2\Delta NLL_{N+1} | \chi^2(M)). \quad (3.10)$$

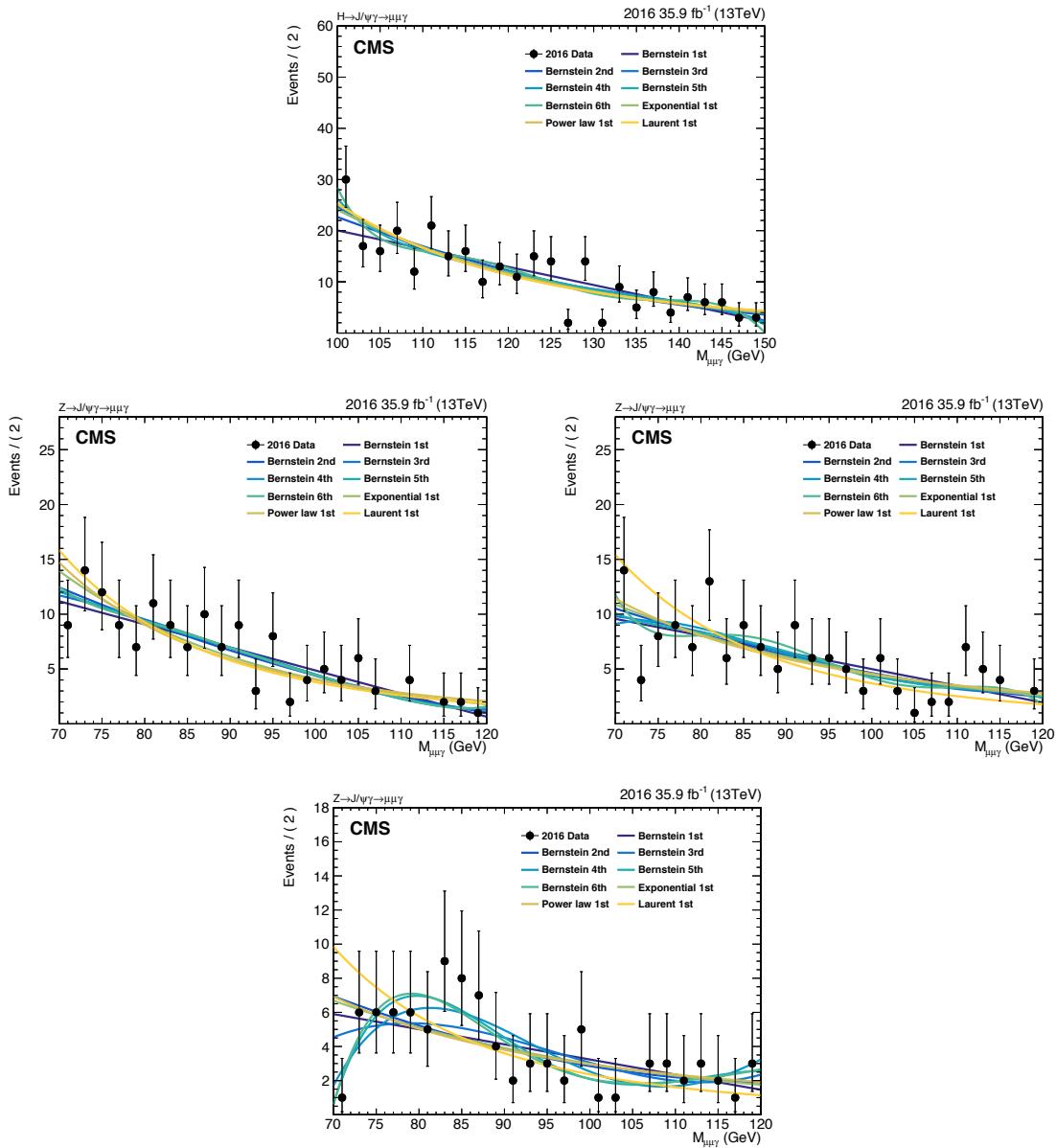


Figure 3-34: Fits on the three-body invariant mass  $m_{\mu\mu\gamma}$  distributions of data for  $H \rightarrow J/\psi \gamma$  (top),  $Z \rightarrow J/\psi \gamma$  Cat1 (middle left),  $Z \rightarrow J/\psi \gamma$  Cat2 (middle right), and  $Z \rightarrow J/\psi \gamma$  Cat3 (bottom).

If the p-value is less than 0.05, the higher order function is supported by the data since the probability of obtaining a NLL with  $(N + 1)_{th}$  order function being greater or equal to a NLL with  $N_{th}$  function is small. The procedure then continues to test the next higher order function in the family. If the p-value is more than 0.05, meaning that an additional increase of parameters does not result in a significant improvement of the fit. Therefore the higher order function is considered to be too flexible for the given  $m_{\mu\mu\gamma}$  distribution in data. The procedure terminates, and the highest order of function in a family is found.

As a result, the functions with 1 degree of freedom of exponential, power law, and Laurent form are picked up by the F-test. These 3 functions with Bernstein polynomials from 1st to 6th order will be tested in the bias study.

Table 3.14 shows the functions to be used in the bias study.

Category	Bernstein polynomial	Exponential	Power-law	Laurent
$H \rightarrow J/\psi \gamma$				
Inclusive	1st - 6th	1Exp	1Pow	1Lau
$Z \rightarrow J/\psi \gamma$				
Cat1, EB_HR9	1st - 6th	1Exp	1Pow	1Lau
Cat2, EB_LR9	1st - 6th	1Exp	1Pow	1Lau
Cat3, EE	1st - 6th	1Exp	1Pow	1Lau

Table 3.14: The functions to be used in the bias study for both Higgs and Z decays.

### 3.5.2 Bias study

Bias study is performed to determine the best function out of those resulting from the F-test. The procedures of bias study are as follows. One of the functions listed in Table 3.14 is chosen to fit to  $m_{\mu\mu\gamma}$  distribution from data events. Pseudo-events are randomly generated by using the resulting fit (referred to as the true function) as background model to simulate possible experiment results. Signal events with signal strength  $\mu_{\text{True}}$  are introduced when generating the pseudo-events. We should note that  $\mu_{\text{True}} = 1$  corresponds to injecting  $1 \times (\text{expected signal yield})$  events on top of the background. A fit is made to the distribution using one of the

functions in the four families combined with a signal model, where the normalization of the signal in this step is allowed to be negative. This procedure is repeated many times, and it's expected that ideally on average the signal strength predicted by the fit  $\mu_{\text{Fit}}$  will be equal to  $\mu_{\text{True}}$ . A pull value, defined as  $(\mu_{\text{Fit}} - \mu_{\text{True}}) / \sigma_{\text{Fit}}$ , where  $\sigma_{\text{Fit}}$  is the error on  $\mu_{\text{Fit}}$ , is calculated for each pseudo-event. The criteria used to determine the unbiased fit is that, the distribution of the pull value  $(\mu_{\text{Fit}} - \mu_{\text{True}}) / \sigma_{\text{Fit}}$  from all pseudo-events with a given combination of true and fit function should be a Gaussian with a mean value less than 0.20 and width around 1. The criteria of 0.20 ensures that a possible bias is at least 20% times smaller than the statistical fluctuation, hence can be neglected. This also implies that the error on the frequentist coverage of the quoted measurement in the analysis is less than 1%, where the coverage is defined as the fraction of experiments in which the true value is contained within the confidence interval. Since the bias introduced by the unbiased fit is negligible, no additional uncertainty is assigned for the background modeling.

The 2-D bias maps of the study with true function (used to generate the toys) on the X-axis and the fitted function (used to fit the toys) on the Y-axis of  $H \rightarrow J/\psi \gamma$  (Fig. 3-35), Cat1 in  $Z \rightarrow J/\psi \gamma$  (Fig. 3-36), Cat2 in  $Z \rightarrow J/\psi \gamma$  (Fig. 3-37), and Cat3 in  $Z \rightarrow J/\psi \gamma$  (Fig. 3-38) are shown. For the  $H \rightarrow J/\psi \gamma$ , the table with  $\mu_{\text{True}} = 300$  is shown. For all the three categories of  $Z \rightarrow J/\psi \gamma$ , the tables with  $\mu_{\text{True}} = 200$  are shown.

The pull-value distributions are shown in Fig. 3-39, 3-40, 3-41, and 3-42. Some of the pseudo-events generated in this study are shown in Appendix A.

For the  $H \rightarrow J/\psi \gamma$  channel, the lowest order satisfying the criteria of bias 20% is Bernstein polynomial of 2nd order. For the  $Z \rightarrow J/\psi \gamma$  channel, the lowest order satisfying the criteria for all three categories are Bernstein polynomial of 3rd order.

The background fits with the best fit functions for both Higgs and Z boson are shown in Fig 3-43 (Top:  $H \rightarrow J/\psi \gamma$ ; Middle left: Cat1 of  $Z \rightarrow J/\psi \gamma$ ); Middle right: Cat2 of  $Z \rightarrow J/\psi \gamma$ ; Bottom: Cat3 of  $Z \rightarrow J/\psi \gamma$ ).

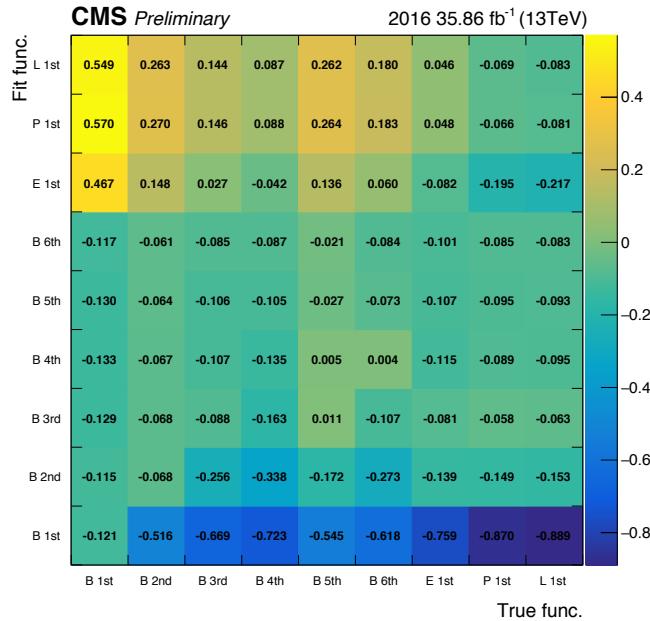


Figure 3-35: The 2-D bias maps of the study for  $\mu_{\text{True}} = 300$  with true function (used to generate the toys) on the X-axis and the fitted function (used to fit the toys) on the Y-axis of  $H \rightarrow J/\psi \gamma$

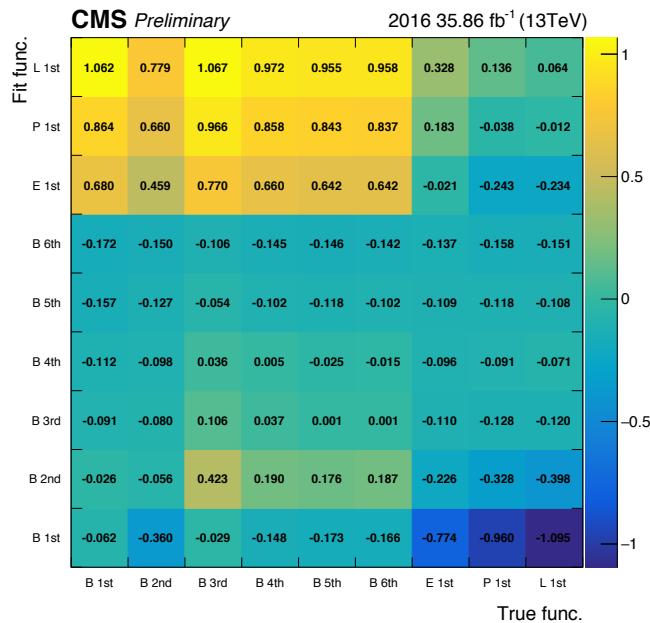


Figure 3-36: The 2-D bias maps of the study for  $\mu_{\text{True}} = 200$  with true function (used to generate the toys) on the X-axis and the fitted function (used to fit the toys) on the Y-axis of Cat1 in  $Z \rightarrow J/\psi \gamma$

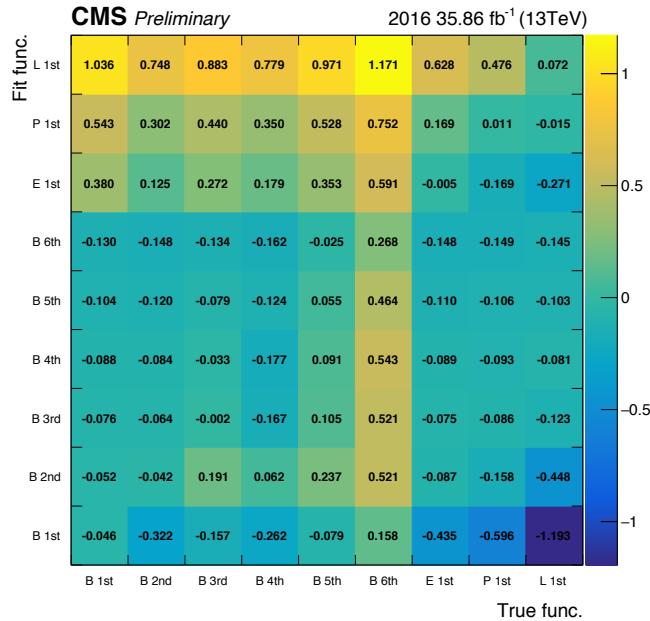


Figure 3-37: The 2-D bias maps of the study for  $\mu_{\text{True}} = 200$  with true function (used to generate the toys) on the X-axis and the fitted function (used to fit the toys) on the Y-axis of Cat2 in  $Z \rightarrow J/\psi \gamma$

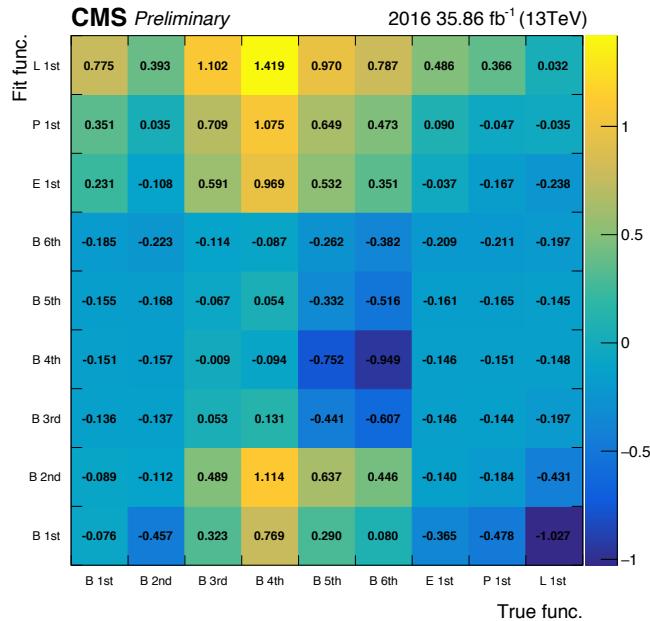


Figure 3-38: The 2-D bias maps of the study for  $\mu_{\text{True}} = 200$  with true function (used to generate the toys) on the X-axis and the fitted function (used to fit the toys) on the Y-axis of Cat3 in  $Z \rightarrow J/\psi \gamma$

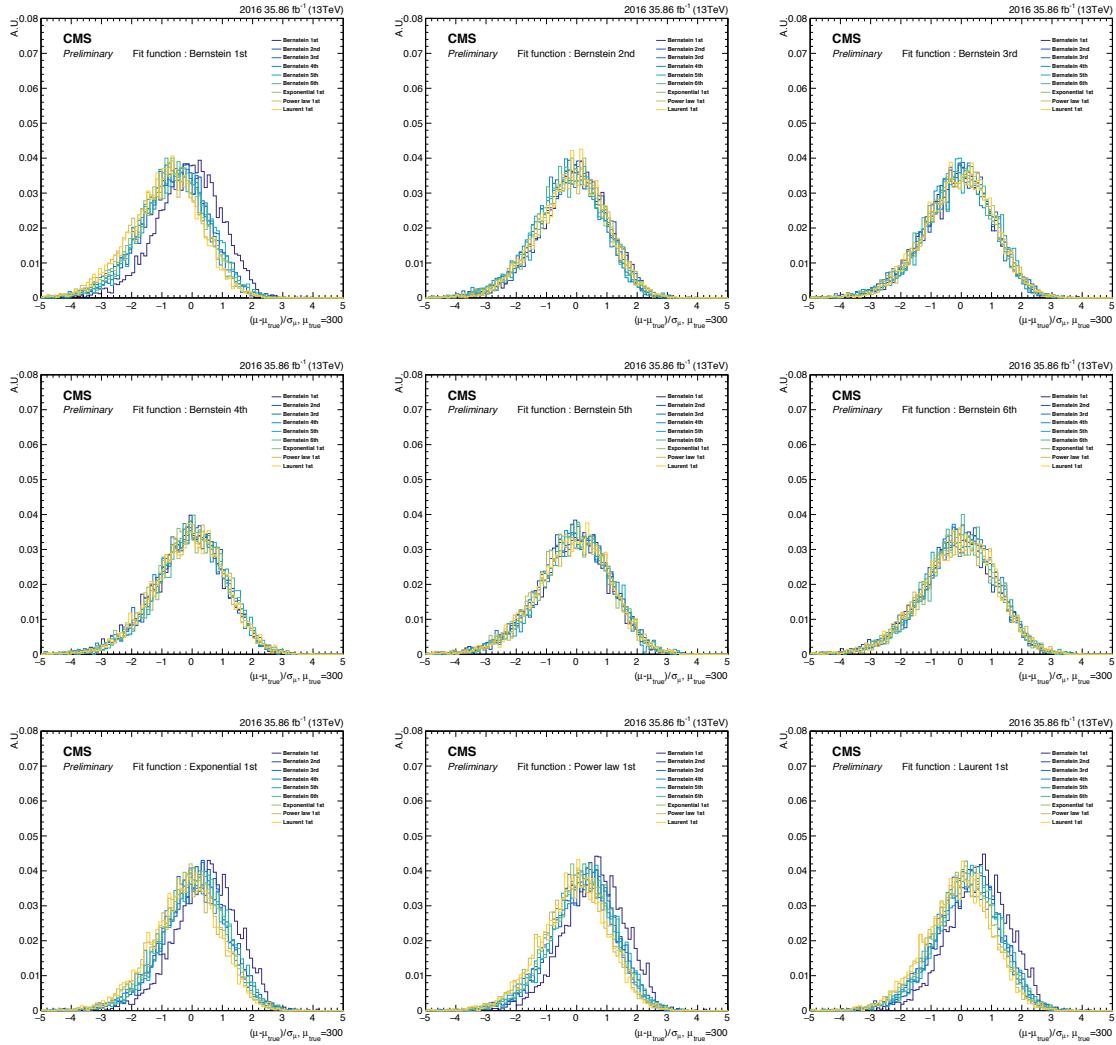


Figure 3-39: The pull-value distributions of bias study in the Higgs boson search. In these plots, the legend labels the distributions using different true functions. The fit function of each plots is: (Top left) Bernstein of 1st order; (Top middle) Bernstein of 2nd order; (Top right) Bernstein of 3rd order; (Middle left) Bernstein of 4th order; (Middle central) Bernstein of 5th order; (Middle right) Bernstein of 6th order; (Bottom left) Exponential with 1 d.o.f (1Exp); (Bottom middle) Power law with 1 d.o.f (1Pow); (Bottom right) Laurent series with 2 terms (1Lau).

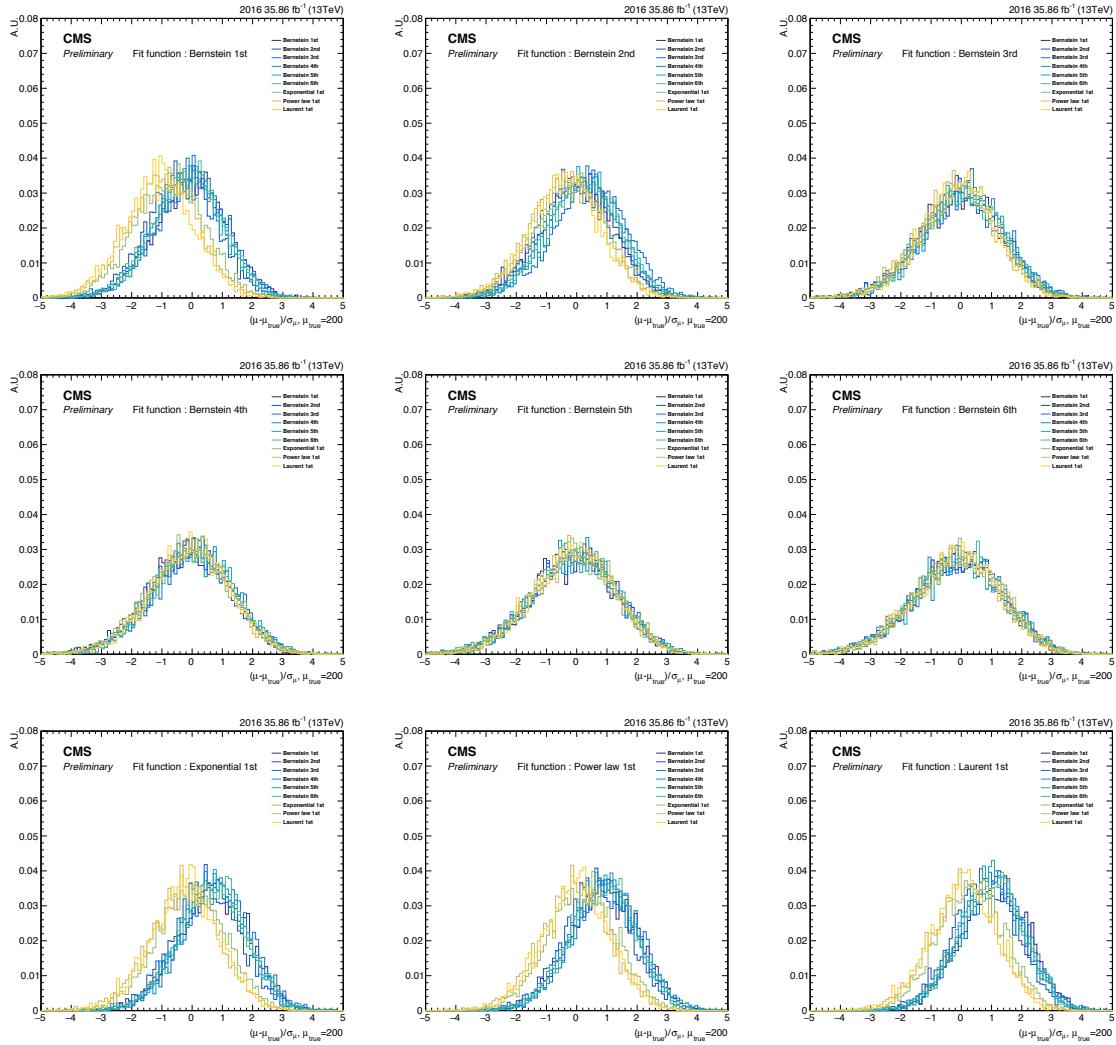


Figure 3-40: The pull-value distributions of bias study of Cat1 in the Z boson search. In these plots, the legend labels the distributions using different true functions. The fit function of each plots is: (Top left) Bernstein of 1st order; (Top middle) Bernstein of 2nd order; (Top right) Bernstein of 3rd order; (Middle left) Bernstein of 4th order; (Middle central) Bernstein of 5th order; (Middle right) Bernstein of 6th order; (Bottom left) Exponential with 1 d.o.f (1Exp); (Bottom middle) Power law with 1 d.o.f (1Pow); (Bottom right) Laurent series with 2 terms (1Lau).

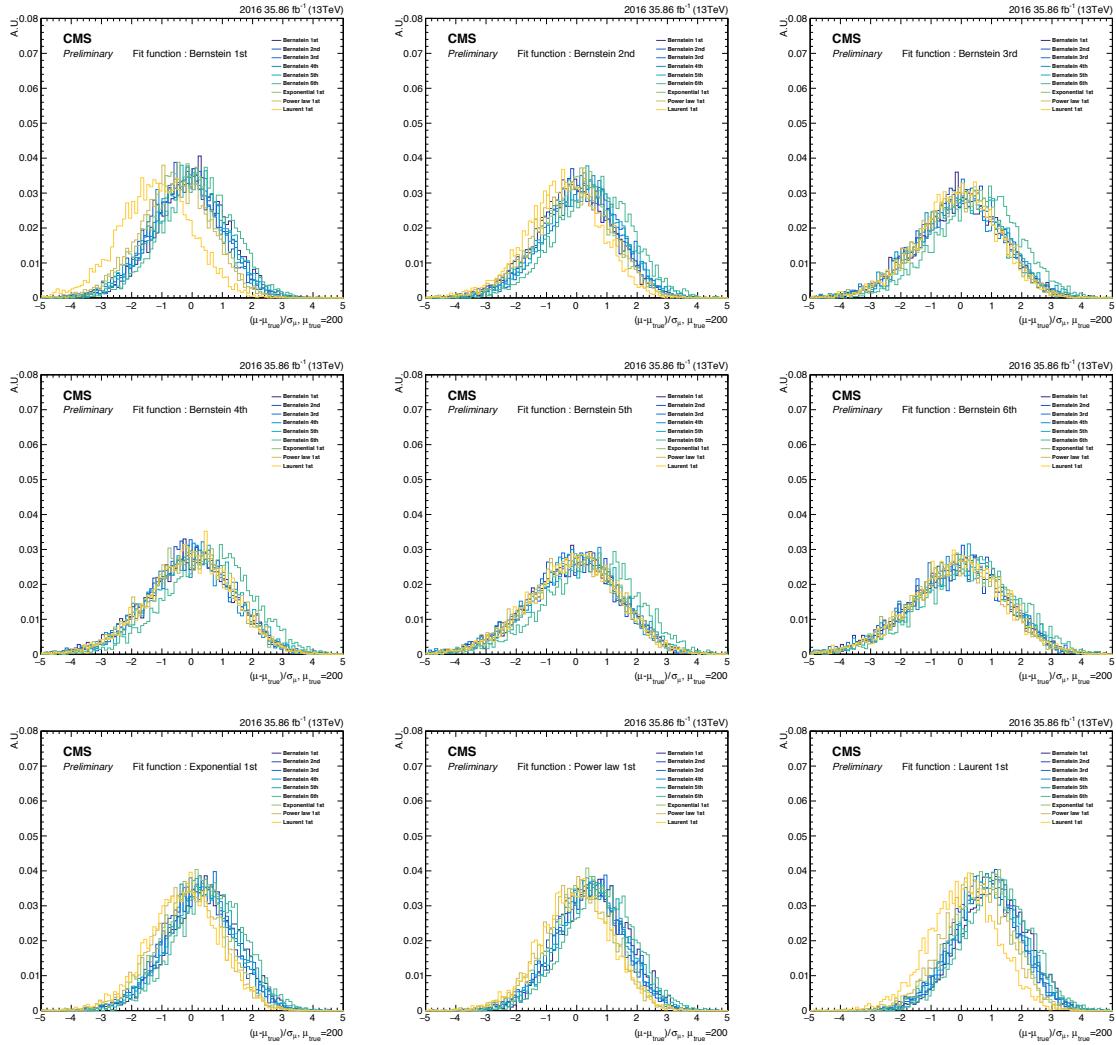


Figure 3-41: The pull-value distributions of bias study of Cat2 in the Z boson search. In these plots, the legend labels the distributions using different true functions. The fit function of each plots is: (Top left) Bernstein of 1st order; (Top middle) Bernstein of 2nd order; (Top right) Bernstein of 3rd order; (Middle left) Bernstein of 4th order; (Middle central) Bernstein of 5th order; (Middle right) Bernstein of 6th order; (Bottom left) Exponential with 1 d.o.f (1Exp); (Bottom middle) Power law with 1 d.o.f (1Pow); (Bottom right) Laurent series with 2 terms (1Lau).

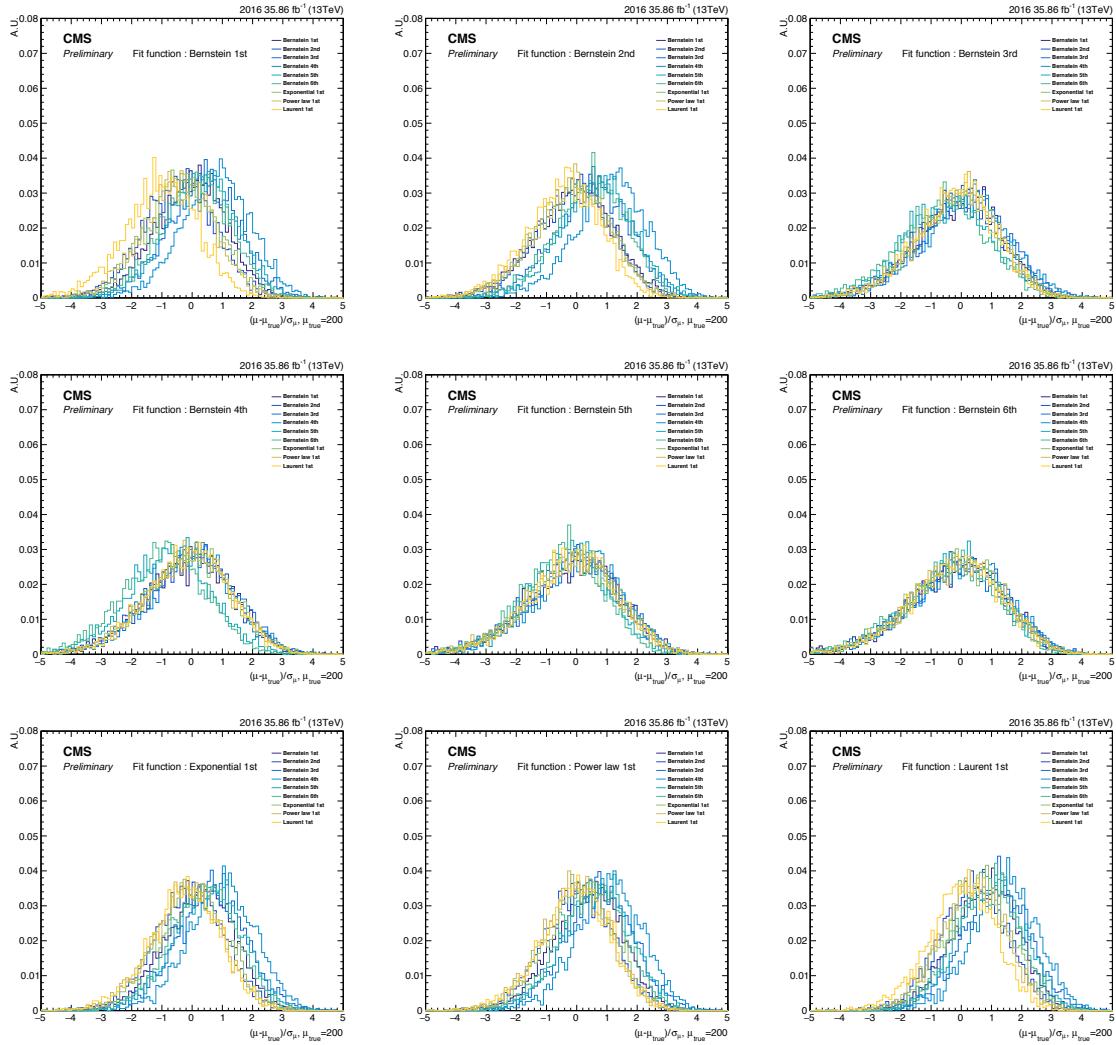


Figure 3-42: The pull-value distributions of bias study of Cat3 in the Z boson search. In these plots, the legend labels the distributions using different true functions. The fit function of each plots is: (Top left) Bernstein of 1st order; (Top middle) Bernstein of 2nd order; (Top right) Bernstein of 3rd order; (Middle left) Bernstein of 4th order; (Middle central) Bernstein of 5th order; (Middle right) Bernstein of 6th order; (Bottom left) Exponential with 1 d.o.f (1Exp); (Bottom middle) Power law with 1 d.o.f (1Pow); (Bottom right) Laurent series with 2 terms (1Lau).

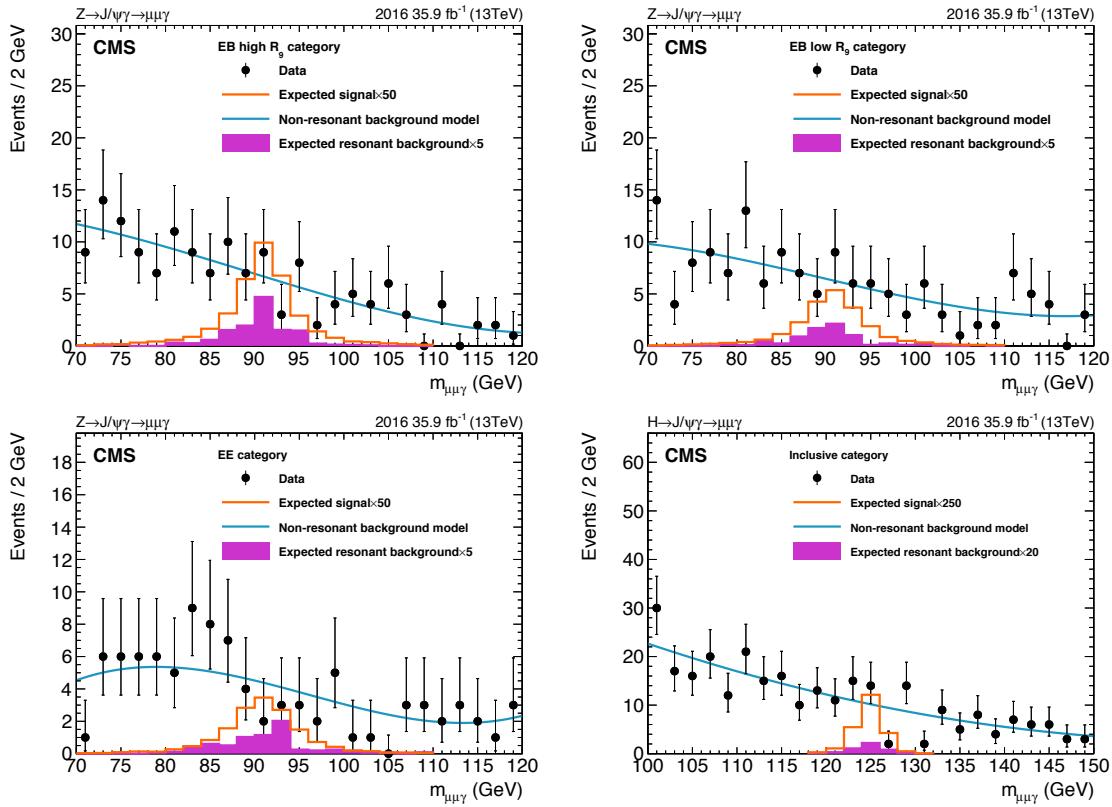


Figure 3-43: Non-resonant background fits with the lowest order unbiased functions to the three-body invariant mass  $m_{\mu\mu\gamma}$  distributions observed in data for the  $Z \rightarrow J/\psi \gamma$  channel in the EB high  $R_9$  category (top left), the EB low  $R_9$  category (top right), the EE category (bottom left), as well as the  $H \rightarrow J/\psi \gamma$  channel (bottom right).

## Motivation of B-only fit

Here we show the plots comparing background-only (B-only) with signal-plus-background (B+S) fit to motivate that including signal region in B-only fits does not change the background model significantly. 2 sets of comparisons are made. Fig. 3-44 shows the s+b fit where signal component is fixed to be the expected yield in each category. Fig. 3-45 shows the s+b fit where signal component is allowed to float when the fit is performed.

Here, an argument is made that the B+S fit in the “full mass” range is actually not too much different from B-only fit in sidebands in combination to signal shape, where the signal shape takes care the region, say,  $\pm 2\sigma$  of the signal distribution (that is, the range containing 95% of signal events). Then based on the plots attached previously, the conclusion can be drawn that the difference between the background model resulting from sideband region and that from the whole range is not significant at all.

Another study is made with binned fit. Fig. 3-46 shows the sideband-only fit (in red) and the sideband-plus-signal region fit (in blue) to the event in  $H \rightarrow J/\psi \gamma$  search. The  $\chi^2/NDF$  of each fit is also shown in the legend. The reasonable assumption in this study is that the resulting function forms from binned fit and un-binned fit are similar.

As one can see, neither including the signal component in the fit does not have significant impact on the overall shape. Whether a sideband-only fit or not will not affect the background model much.

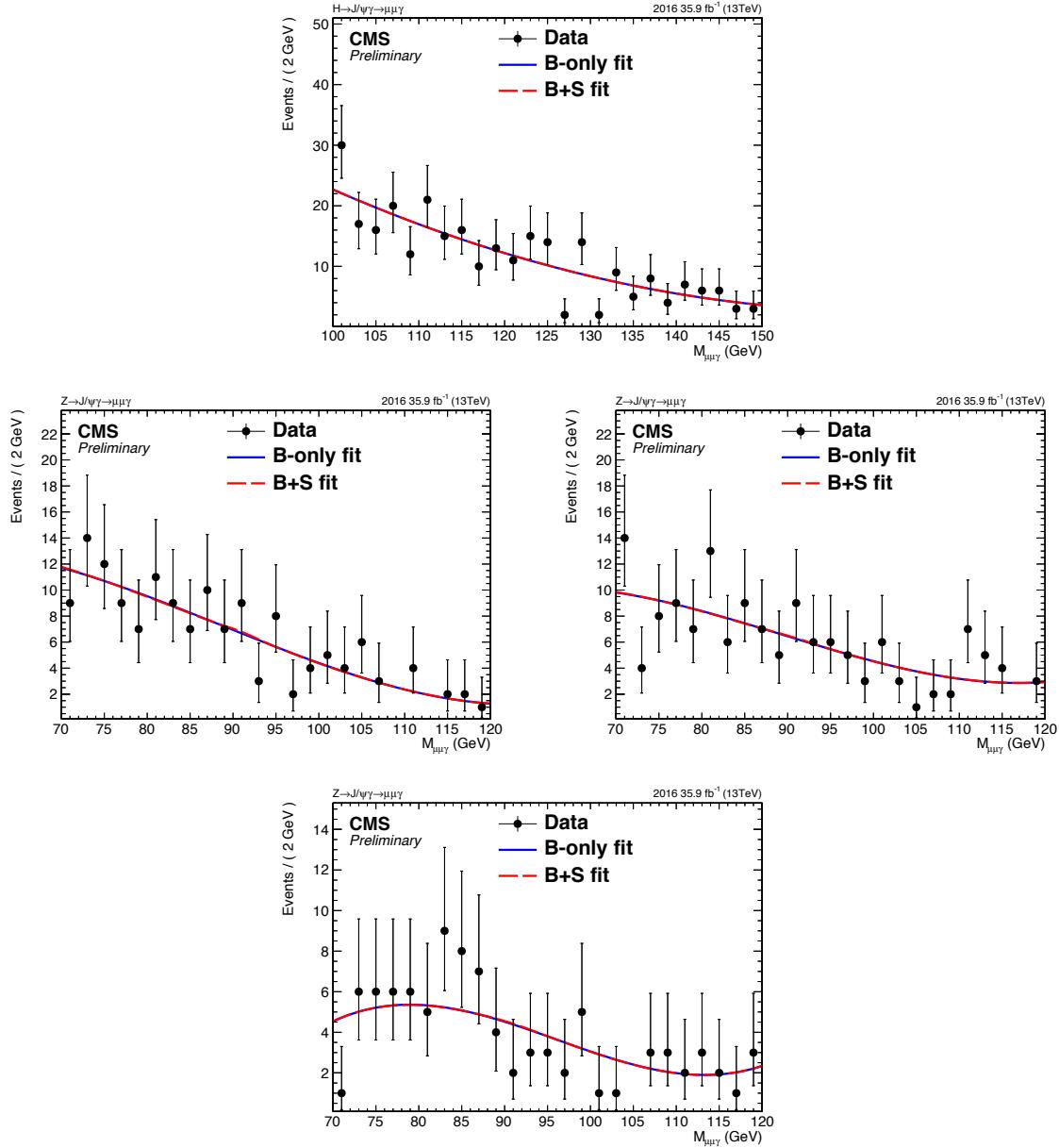


Figure 3-44: B+S fits, where signal component is fixed to be the expected yield in each category, with lowest order unbiased functions on the three-body invariant mass  $m_{\mu\mu\gamma}$  distributions of data for  $H \rightarrow J/\psi \gamma$  (top),  $Z \rightarrow J/\psi \gamma$  Cat1 (middle left),  $Z \rightarrow J/\psi \gamma$  Cat2 (middle right), and  $Z \rightarrow J/\psi \gamma$  Cat3 (bottom).

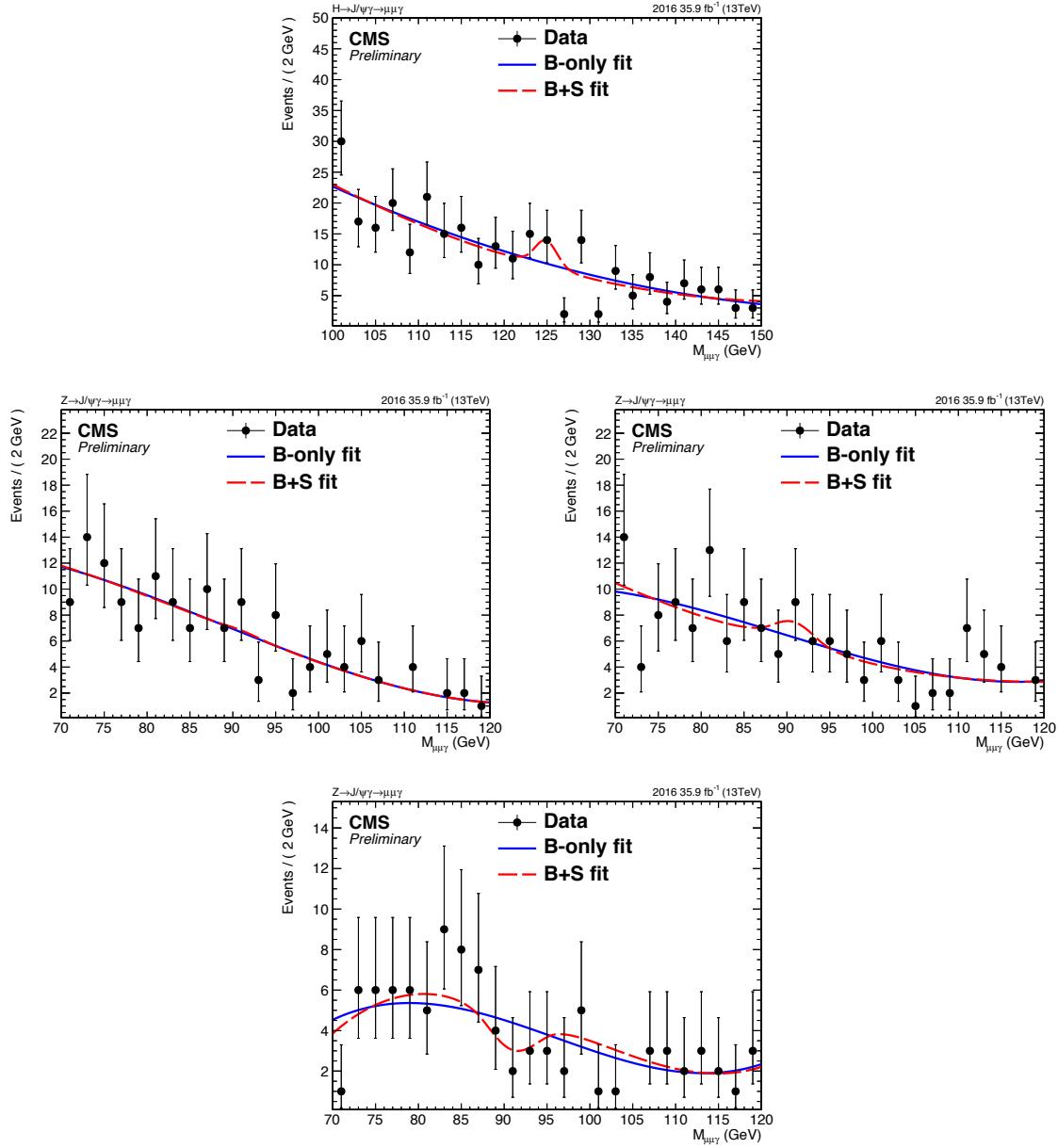


Figure 3-45: B+S fits, where signal component is allowed to float, with lowest order unbiased functions on the three-body invariant mass  $m_{\mu\mu\gamma}$  distributions of data for  $H \rightarrow J/\psi \gamma$  (top),  $Z \rightarrow J/\psi \gamma$  Cat1 (middle left),  $Z \rightarrow J/\psi \gamma$  Cat2 (middle right), and  $Z \rightarrow J/\psi \gamma$  Cat3 (bottom).

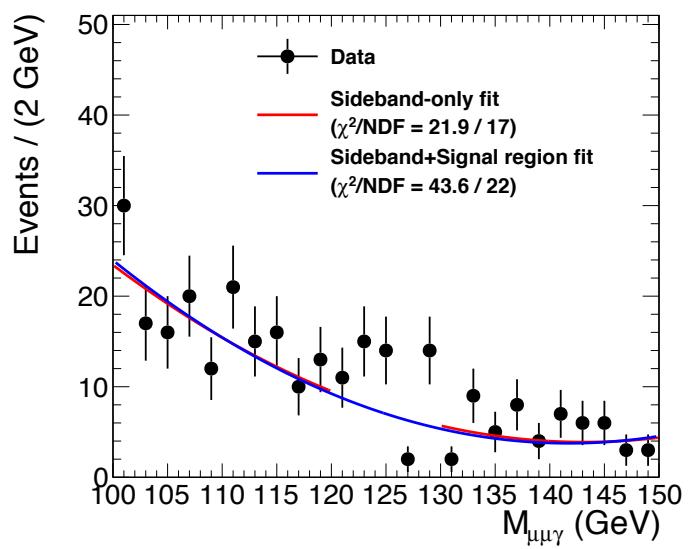


Figure 3-46: The binned fit with sideband-only fit (in red) and with sideband-plus-signal region fit (in blue).

## 3.6 Signal modeling

For the  $H \rightarrow J/\psi \gamma$  decay, a Gaussian function in addition to a Crystal Ball function with common mean value is used. It is a 6-parameter fit (CB: power,  $\alpha$ ,  $\sigma_{CB}$ ; Gaussian:  $\sigma_{Gau}$ ; mean value; fraction of the Gaussian and the Crystal Ball function). For the Higgs Dalitz background, a Crystal Ball function is used to model the shape. For the  $Z \rightarrow J/\psi \gamma$ , we use a double-sided Crystal Ball function. It has 6 parameters: mean,  $\sigma$ ,  $n_1$ ,  $n_2$ ,  $\alpha_1$ , and  $\alpha_2$ . For the  $Z \rightarrow \mu\mu\gamma$  background, we take the  $Z \rightarrow J/\psi \gamma$  signal shape, since the events after full selections are not enough to give reasonable fits. The signal fits for both Higgs and Z boson are shown in Fig. [3-47] and [3-49]. The Higgs Dalitz background shapes for the Higgs decay are shown in Fig. [3-48].

## 3.7 Systematic uncertainties

Systematic uncertainties arising from incomplete knowledge of the detector simulation and theoretical prediction on signal production mechanism may affect the results. Uncertainties for the simulated signal are evaluated by varying contributing sources within their corresponding uncertainties and propagating to the signal yield or shape.

The background modeling and prediction is purely derived from data, so only statistical uncertainties are considered, which are translated into uncertainty on each parameter of the fit function. Besides, the bias study mentioned in previous section is performed to ensure the bias on the choice of the background function is negligible. Hence, no additional systematic uncertainty is assigned.

In both Higgs and Z boson decays, the uncertainties can be classified into two classes, one affecting the predicted signal yields and the other affecting the shape of the signal model.. They are described separately in the following subsections.

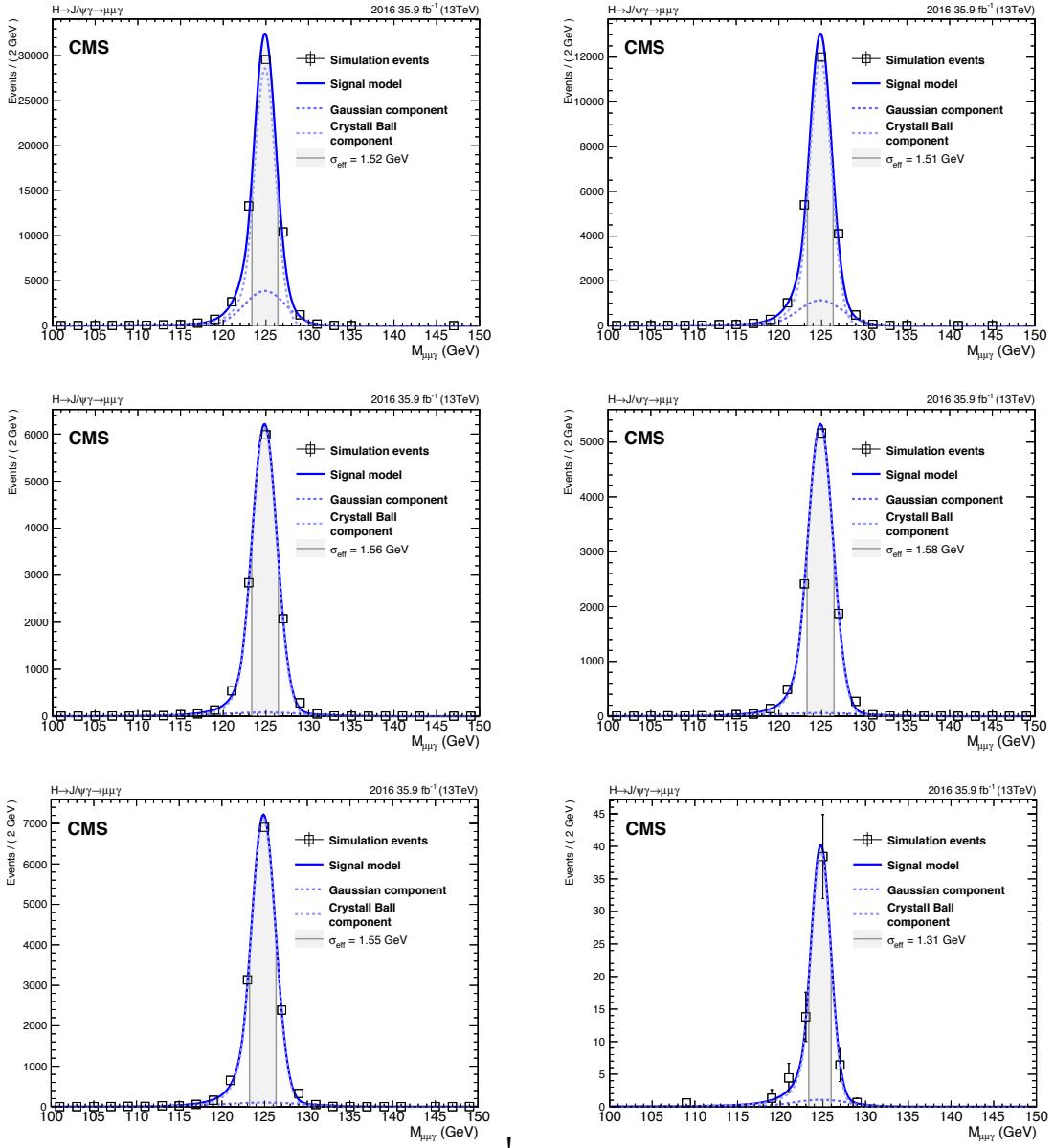


Figure 3-47: Signal model of  $H \rightarrow J/\psi \gamma$  for each production mode. (Top left) gluon fusion; (Top right) vector-boson fusion; (Middle left) ZH production; (Middle right)  $W^+H$  production; (Bottom left)  $W^-H$  production; (Bottom right) associated top quark production ttH.

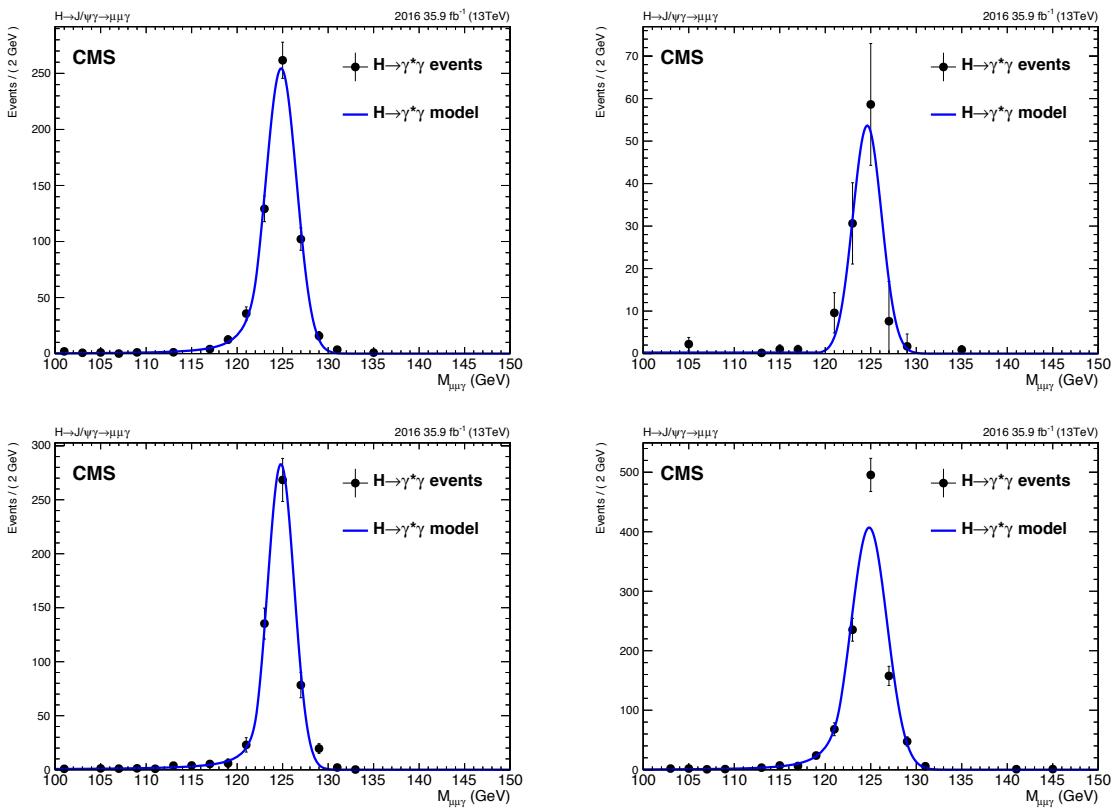


Figure 3-48: The shape of peaking background  $H \rightarrow \gamma^*\gamma$  for ggF (top left), VBF (top right), ZH (bottom left), and  $W^+H$  (bottom right).

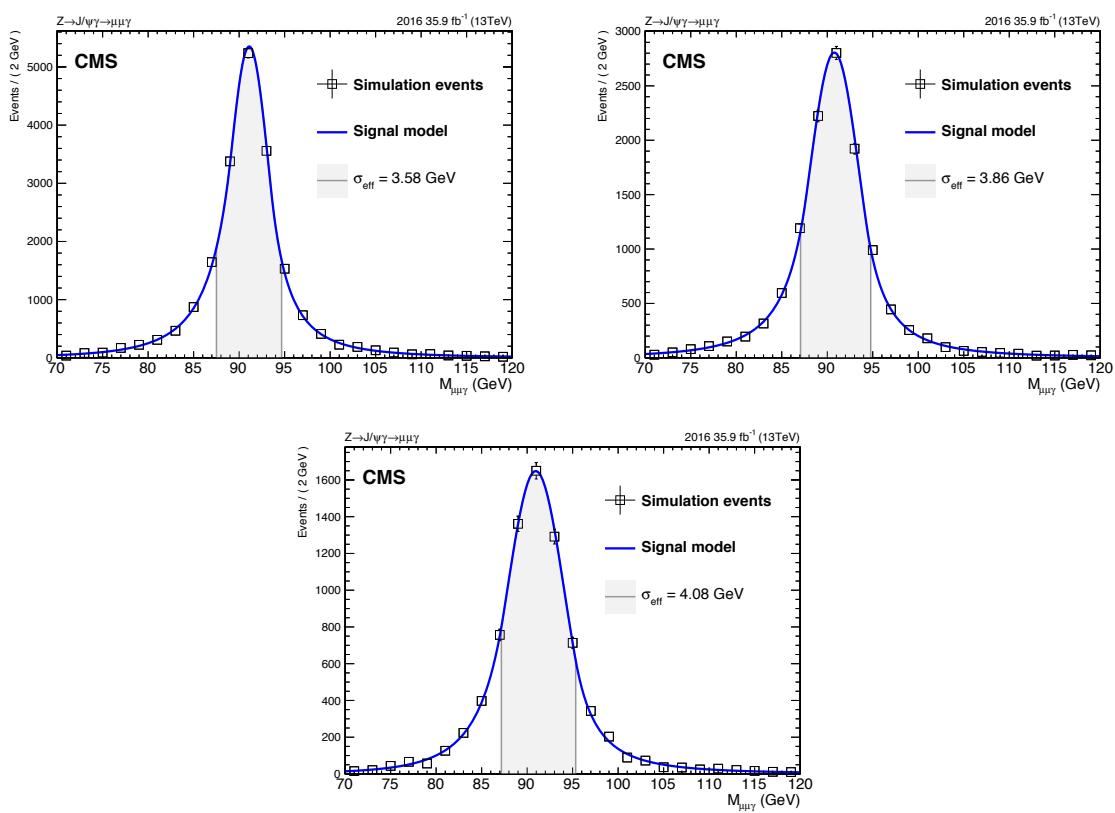


Figure 3-49: Signal model of each category of  $Z \rightarrow J/\psi \gamma$ .

## Uncertainties affecting the predicted signal yields

- **Luminosity measurement** [115]. The recommended value of 2.5% is used.
- **ID and isolation of the objects.** For the muons and photon MVA ID, the uncertainties are derived from three sources:
  1. Different signal and background functions used in tag-and-probe method to obtain the scale factors
  2. Different tag requirements
  3. Compute the passing and failing probes by simply counting in the simulated events

For the photon CSEV, the uncertainties come from

1. Different pileup reweighting references
  2. Adding background simulated events
  3. Different generators used to generate the signal events
- **Trigger.** Uncertainties in the measurement of trigger efficiency scale factors are derived by adding background simulated samples for the computation of the scale factors and by varying the pileup weight references. The systematic uncertainty from each source is taken to be the difference between the nominal value of scale factors and values obtained after varying these parameters. These two were estimated separately and added in quadrature along with the statistical uncertainty to give total uncertainties.
  - **Pileup.** The minimum bias cross-section of 69.2 mb for pileup reweighting is used in the analysis. The analysis is run with varied weights,  $\pm 4.6\%$  with respect to the nominal one. The largest difference in the yields is quoted as the uncertainty.
  - **Theoretical sources.** These include

1. The effects of the parton density function (PDF) choice on the signal cross-section and strong coupling constant  $\alpha_s$  [48, 108, 116]
2. The lack of higher-order calculations for the cross-section and renormalization scale [117, 118, 119, 120, 121]
3. The prediction of the decay branching fraction [122]

## Uncertainties affecting the shape of the signal model

Since the energy resolution of simulated events is better than that in real data, smearing corrections are applied on simulated events. The energy scale in real data is corrected to match the simulated event.

- **Muon momentum scale and resolution.** Rochester correction derived for full 2016 dataset is used in the analysis. There are several sources contributing to the uncertainties, including statistical uncertainty, the effect of correction without reweighting reference to data, varied profile mass windows, and varied fitting mass windows. For each source, the analysis is run many times, varying the members given in the package. Different corrections on the  $p_T$  are applied to muons, and the differences on muon  $p_T$  are then propagated to  $m_{\mu\mu\gamma}$ . Fits to the resulting  $m_{\mu\mu\gamma}$  distributions are done using previously mentioned signal model to obtain the mean and width ( $\sigma$ ) of the Gaussian component of the signal model, which are measures of the scale and resolution uncertainties. When the mean values are to be obtained, the parameters of the signal model are fixed except for the mean value. Similarly, when the  $\sigma$  values are to be obtained, all other parameters than  $\sigma$  are fixed. The largest variation on the mean/ $\sigma$  among the members in each source is quoted. The uncertainties from these four sources are added in quadrature and assigned as total systematic uncertainty for the scale and resolution.
- Photon energy scale. The uncertainty in the photon energy scale is estimated by varying the energy correction. Three sources are considered: statistical

uncertainties, systematic uncertainties (cut-based selection,  $R_9$  categorization, etc.), and gain switch uncertainties. Each of the sources contains up and down corrections. The analysis is run with these six variations and varied  $m_{\mu\mu\gamma}$  distributions are obtained. The signal model with all parameters fixed except for the mean is fitted to varied distributions. The largest variation on the mean value of the fit with respect to the nominal one is taken as systematic uncertainty on the photon energy scale.

- Photon energy resolution. The smearing of the photon energy is done with two parameters, rho and phi, corresponding to constant term and  $E_T$  dependent term. Each of them contains up and down corrections. Similar to what has been done for photon energy scale, the analysis is run with these four variations and varied  $m_{\mu\mu\gamma}$  distributions are obtained. Alternatively, the signal model with all parameters fixed except the  $\sigma$  is fitted to varied distributions. The largest variation on the  $\sigma$  value of the fit with respect to the nominal one is taken as systematic uncertainty on the photon energy resolution.

Table 3.15, 3.16, and 3.17 show the sources of the all systematic uncertainties in both Higgs and Z analyses and the pre-fit value of each source.

Source	Pre-fit value (in %)					
	ggF	VBF	$H \rightarrow J/\psi \gamma$ signal			
			ZH	$W^+H$	$W^-H$	ttH
<b>Integrated luminosity</b>	2.5					
<b>Theoretical uncertainties</b>						
Cross section (scale)	+4.6 -6.7					
Cross section (PDF + $\alpha_s$ )	3.2					
<b>Detector simulation, reconstruction:</b>						
Pileup reweighting	0.686	0.684	0.927	0.606	0.907	1.509
Trigger (per event)	3.92	4.05	4.12	4.23	4.12	4.05
Muon ID/Isolation	2.08	2.04	2.05	2.06	2.06	2.16
Photon ID	1.21	1.18	1.18	1.22	1.17	1.13
Electron veto	1.05	1.05	1.04	1.02	1.04	1.07
<b>Signal model fits:</b>						
Mean (scale)	0.0966	0.0884	0.0804	0.0927	0.0953	0.112
Sigma (resolution)	4.95	4.30	3.35	4.61	3.79	14.1

Table 3.15: Systematic uncertainties for the  $H \rightarrow J/\psi \gamma$  signal.

The statistical uncertainty dominates in the present analysis. The largest systematic uncertainties come from theory and these affect the predicted signal yields

Source	Pre-fit value (in %)			
	$H \rightarrow \gamma^*\gamma$ background			
	ggF	VBF	ZH	WH
<b>Integrated luminosity</b>	2.5			
<b>Theoretical uncertainties</b>				
SM H boson cross section (scale)	+4.6 -6.7			
SM H boson cross section (PDF + $\alpha_s$ )	3.2			
SM BR( $H \rightarrow \gamma^*\gamma$ )	6.0			
<b>Detector simulation, reconstruction:</b>				
Pileup reweighting	1.71	0.103	1.80	1.39
Trigger (per event)	4.10	4.09	4.09	4.29
Muon ID/Isolation	2.50	2.63	2.49	2.20
Photon ID	1.18	1.10	1.17	1.19
Electron veto	1.04	1.11	1.04	1.01

Table 3.16: Systematic uncertainties for the  $H \rightarrow \gamma^*\gamma$  background.

Source	Pre-fit value (in %)					
	$Z \rightarrow J/\psi\gamma$			$Z\gamma \rightarrow \mu\mu\gamma$		
	Cat1	Cat2	Cat3	Cat1	Cat2	Cat3
<b>Integrated luminosity</b>	2.5					
<b>Theoretical uncertainties</b>						
SM Z boson XS (scale)	3.5			5.0		
SM Z boson XS (PDF + $\alpha_s$ )	1.73			5.0		
<b>Detector simulation, reconstruction</b>						
Pileup reweighting	0.990	0.200	1.34	0.940	1.45	4.38
Trigger (per event)	3.30	3.30	6.50	3.41	3.40	6.52
Muon ID/Isolation	2.92	2.95	3.01	3.31	3.42	3.58
Photon ID	1.12	1.11	1.11	1.08	1.08	1.14
Electron veto	1.20	1.20	0.450	1.20	1.92	0.446
<b>Signal model</b>						
Mean (scale)	0.0495	0.0767	0.0685	—		
Sigma (resolution)	0.990	0.690	1.45	—		

Table 3.17: Systematic uncertainties in the Z boson decay.

most. The next two largest are from trigger and muon identification and isolation where the efficiency measurements come from a limited sample. Smaller systematic uncertainties are likely in future analyses since more improved and robust methods are being developed and planned for the future. Nevertheless, as a summary, the overall impact from the uncertainties in the final observed and expected limits are small.

## $\psi(2S)$ feed-down

The decay  $\psi(2S) \rightarrow J/\psi(\rightarrow \mu\mu) + X$ , where  $X$  can be anything, contributes as a background source. Currently there is no theoretical reference on the branching ratio of the  $Z \rightarrow \psi(2S)\gamma$ , so here an assumption is made,

$$\frac{N(Z \rightarrow J/\psi \gamma)}{N(Z \rightarrow \psi(2S)\gamma)} \simeq \frac{N(Z \rightarrow J/\psi + ll)}{N(Z \rightarrow \psi(2S) + ll)} \simeq 3.5 \quad (3.11)$$

By taking the branching ratio of the  $\psi(2S) \rightarrow J/\psi(\rightarrow \mu\mu) + X$  into account,

$$\frac{N(Z \rightarrow J/\psi \gamma)}{N(Z \rightarrow \psi(2S)\gamma[\rightarrow J/\psi(\rightarrow \mu\mu) + X])} \simeq \frac{N(Z \rightarrow J/\psi + ll)}{N(Z \rightarrow \psi(2S) + ll[\rightarrow J/\psi(\rightarrow \mu\mu) + X])} \simeq 5.7 \quad (3.12)$$

we then expect to have  $1.54/5.7 \sim 0.270$  events from the  $\psi(2S)$  decay, where 1.54 is the expected yield of  $Z \rightarrow J/\psi \gamma$ . This is negligible amount compared to the total background, 384.

The mass shapes of this background at the generator level are shown in Fig. [3-50], where  $m_{\mu\mu\gamma}$  distributions from the  $Z \rightarrow (\psi(2S) \rightarrow J/\psi + X) \gamma$  are in blue and from the  $Z \rightarrow J/\psi \gamma$  are in red. The distribution in solid line is without the kinematic cuts used in the selection, while the filled distribution in dashed line is after imposing the kinematic cuts. The distribution without kinematic cuts is normalized to 1, while the one with kinematic cuts is normalized to the fraction of the events passing kinematic cuts. As one can see, the  $Z \rightarrow (\psi(2S) \rightarrow J/\psi + X) \gamma$  actually contributes as peaking background, with the peak shifts around 10 GeV toward lower value. Since it is estimated to be 1/6 of signal and small compared to total background, it will be taken care by the background fit. Further more, from the red dashed distribution the range containing  $\sim 68\%$  of events is of 87.4 to 94.6 GeV, which corresponds to 17.3% of events of  $Z \rightarrow \psi(2S)\gamma$  after kinematic selection. It is less than 2.9% of the  $Z \rightarrow J/\psi \gamma$  yield for which are relevant at limit calculation.

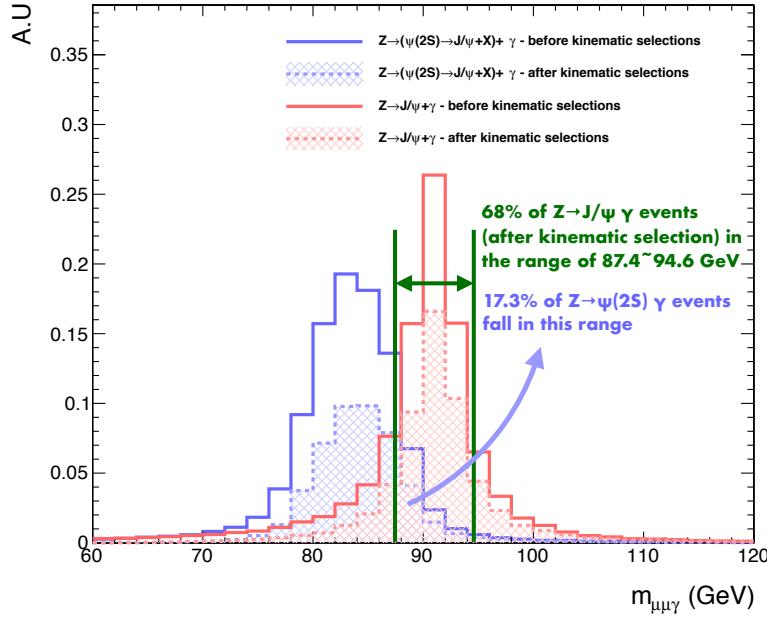


Figure 3-50: The comparison of the  $m_{\mu\mu\gamma}$  distributions between the  $Z \rightarrow (\psi(2S) \rightarrow J/\psi + X) \gamma$  (blue) and the  $Z \rightarrow J/\psi \gamma$  (red) at the generator level.

### 3.8 Statistical method

The model-independent limit is set on the signal cross section times branching ratio ( $\sigma \times \mathcal{BR}$ ) with procedures followed from Ref. [123, 124, 125, 126].

First, a likelihood function is constructed as:

$$\mathcal{L}(\text{data} \mid \mu, \theta) = \text{Poisson}(\text{data} \mid \mu \cdot s(\theta_{\text{sig}}) + b(\theta_{\text{bkg}})) \cdot p(\tilde{\theta} \mid \theta), \quad (3.13)$$

where “data” can either be actual experiment observation or pseudo-events;  $\mu$  is the signal strength modifier, defined as  $\sigma \times \mathcal{BR} / (\sigma \times \mathcal{BR})_{\text{SM}}$ ;  $\theta_{\text{sig(bkg)}}$  represents the set of nuisance parameters associated with the signal and background model. Systematic uncertainties are treated as nuisance parameters which are of uninterested in the analysis but can affect results, and handle by introducing probability distribution functions (pdfs)  $p(\tilde{\theta} \mid \theta)$ . Here  $\theta$  stands for the whole set of nuisance in the analysis, and  $\tilde{\theta}$  represents the set of default values of the nuisance parameter, which reflecting our knowledge or belief on what values of these parameters can

be. There are different choices of pdfs for nuisance parameters.

- **Gaussian** pdf is used for parameters that can be either positive or negative.

$$p(\tilde{\theta}|\theta) = \frac{1}{\sqrt{2}} \exp \left( -\frac{(\tilde{\theta} - \theta)^2}{2\sigma^2} \right), \quad (3.14)$$

The uncertainties in the parameter of the signal model belong to this class. Two multiplicative factors  $\kappa_m \equiv 1 \pm \delta\kappa_m$  and  $\kappa_\sigma \equiv 1 \pm \delta\kappa_\sigma$  are introduced such that the mean and width of the signal model are modified as

$$m' = \kappa_m \cdot m, \quad \sigma' = \kappa_\sigma \cdot \sigma, \quad (3.15)$$

where  $m$  and  $\sigma$  are original parameters.

- **Log-normal** pdf is an alternative pdf for positively defined parameters.

$$p(\tilde{\theta}|\theta) = \frac{1}{\sqrt{2\pi \ln(w)}} \exp \left( -\frac{(\ln(\tilde{\theta}/\theta))^2}{2(\ln(w))^2} \right) \frac{1}{\tilde{\theta}}, \quad (3.16)$$

where  $w$  characterizes the width of the log-normal pdf. This distribution has a longer tail than the Gaussian and goes to zero at  $\theta = 0$ . This class includes uncertainties in luminosity, cross-section, efficiency measurements.

- The parameters for background model are allowed to freely float across their ranges and not Gaussian constrained.

The unbinned likelihood is computed as,

$$k^{-1} \prod_i (\mu S f_s(x_i) + B f_b(x_i)) \cdot e^{-(\mu S + B)}. \quad (3.17)$$

$f_{s(b)}(x_i)$  are pdfs (models) of signal and background of observable(s)  $x_i$ , and  $S$  and  $B$  are event yields for signal and background. The observable used in this analysis is the three-body invariant mass  $m_{\mu\mu\gamma}$ .

The likelihood function can be used to represent *background-only* (*b-only*) hypothesis,  $\mathcal{L}_b = \mathcal{L}(\mu = 0)$ , and *signal plus background* (*s+b*) hypothesis,  $\mathcal{L}_{s+b} = \mathcal{L}(\mu)$ .

For the nominal SM hypothesis  $\mu = 1$ .

Based on the Neyman & Pearson lemma [127], the likelihood ratio  $\frac{\mathcal{L}_{s+b}}{\mathcal{L}_b}$  provides the most powerful test for hypothesis test. Hence, it is used as *test statistic*,

$$t(\mu) = \frac{\mathcal{L}_{s+b}}{\mathcal{L}_b} = \frac{\mathcal{L}(\mu, \theta)}{\mathcal{L}(0, \theta)}. \quad (3.18)$$

However, since the expected signal yields from the SM prediction are small,  $\mathcal{L}_{s+b}$  and  $\mathcal{L}_b$  are not well separated. In other words, we are not sensitive to determine the presence of the signal yet. Instead, an upper limit on the  $\mu$  is set, and a different test statistic is used.

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} & \text{if } \hat{\mu} \geq 0 \\ \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(0, \hat{\theta}_{\mu=0})} & \text{if } \hat{\mu} < 0 \end{cases}$$

where  $\hat{\theta}_\mu$  is the value of  $\theta$  that maximizes  $\mathcal{L}$  for a specific  $\mu$ ; the  $\mathcal{L}(\hat{\mu}, \hat{\theta})$  is the global (unconditional) maximum of the likelihood function, where  $\hat{\mu}$  and  $\hat{\theta}$  are values such that the likelihood function is maximized. In the second part where  $\hat{\mu} < 0$ , the definition of  $\tilde{\lambda}(\mu)$  is determined to constrain the signal yield to be positive.

Apart from the negative signal rate constraint, upward fluctuations of the data such that  $\hat{\mu} > \mu$  are not considered as evidence against the signal hypothesis  $\mu$ . Based on this argument, the test statistic is modified as,

$$\tilde{q}_\mu = \begin{cases} -2 \ln \tilde{\lambda}(\mu) & \text{if } 0 \leq \hat{\mu} \leq \mu \\ 0 & \text{if } \hat{\mu} \geq \mu \end{cases}$$

The *observed* value of the test statistic  $\tilde{q}_\mu$  for a given signal strength  $\mu$  under test  $\tilde{q}_\mu^{obs}$ , as well as the value of nuisance parameters  $\hat{\theta}_0^{obs}$  and  $\hat{\theta}_\mu^{obs}$  that maximize the likelihood for *b-only* and *s+b* hypotheses respectively, can be found. Next, pseudo-

events are generated, based on the pdfs for signal and background, to construct pdfs for  $\tilde{q}_\mu$  for *b-only* and *s+b* hypotheses,  $f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{obs})$  and  $f(\tilde{q}_\mu|0, \hat{\theta}_0^{obs})$ . Example distributions are shown in Fig. [3-51].

Having  $f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{obs})$  and  $f(\tilde{q}_\mu|0, \hat{\theta}_0^{obs})$  distributions, two p-values are defined to be associated with the actual observation for *s+b* and *b-only* hypotheses,  $p_\mu$  and  $p_b$ ,

$$p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} | s + b) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{obs}) d\tilde{q}_\mu^{obs} \quad (3.19)$$

$$p_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} | b - only) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|0, \hat{\theta}_0^{obs}) d\tilde{q}_\mu^{obs} \quad (3.20)$$

and  $CL_s(\mu)$  is defined as a ratio of these two p-values

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b}. \quad (3.21)$$

To quote the 95% confidence level (C.L) upper limit on  $\mu$ , denoted as  $\mu^{95\%CL}$ , the  $\mu$  value is adjusted until the  $CL_s(\mu) = 0.05$ . The derived limit is called *observed limit*<sup>2</sup>.

The traditional way to compute the *expected* limit for *b-only* hypothesis is to generate a large number of pseudo-events based on the pdfs of the signal and background, without using the true data, treat them as real data, and calculate the  $CL_s$  and  $\mu^{95\%CL}$  for each of them. A pdf for the  $\mu^{95\%CL}$  and corresponding cumulative probability distribution (or cumulative distribution function, CDF) can be obtained. An example is shown in Fig. [3-52]. The point where the CDF crosses 50% of entries is the median expected value. The  $\pm 1\sigma$  (68%) band is defined as points crossings of the 16% and 84% entries. Points crossings at 2.5% and 97.5% define the  $\pm 2\sigma$  (95%) band.

Instead, in this analysis *expected* limits are set with the *asymptotic* method. The detail discussion of the method is described in Ref. [126], here a brief summary is

---

<sup>2</sup>If, for example,  $\mu = 1$  and  $CL_s \leq \alpha$ , we would state that “The SM Higgs/Z boson decay is excluded with  $(1 - \alpha)CL_s$  confidence level (C.L).”.

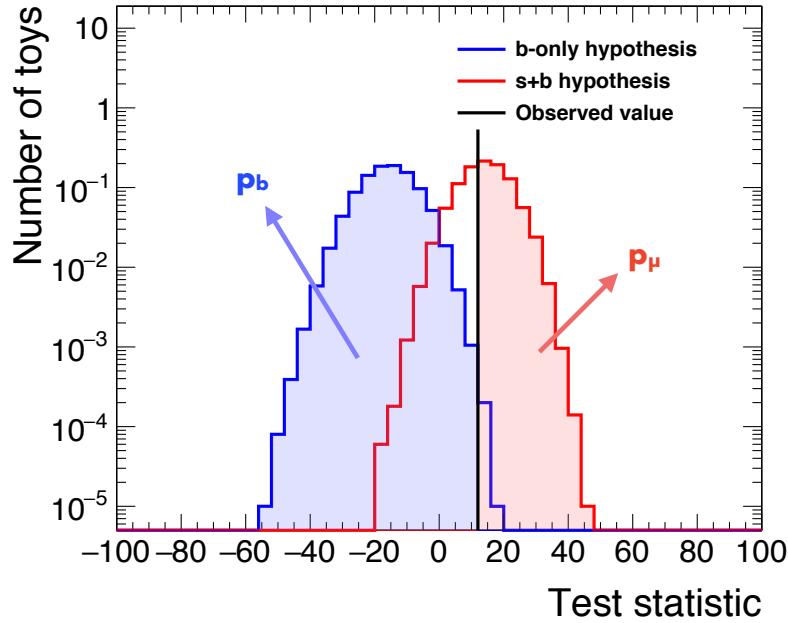


Figure 3-51: Test statistic distributions for pseudo-events generated with signal+background and background-only hypotheses.

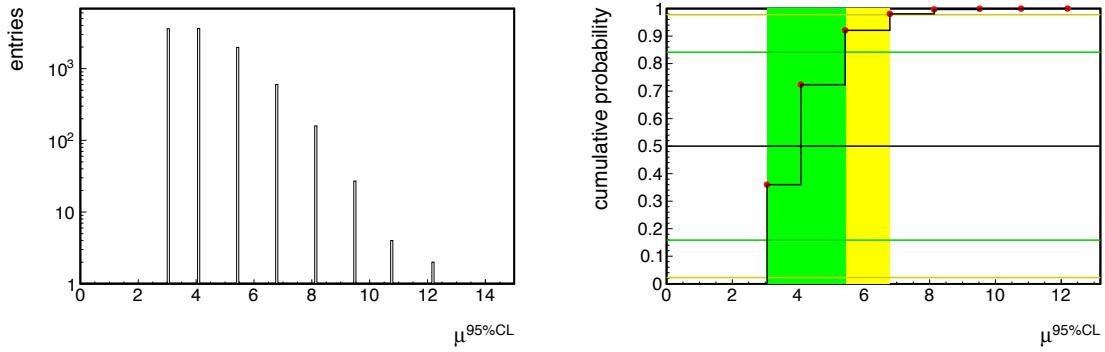


Figure 3-52: An example of distribution of  $\mu^{95\%CL}$  for the *b-only* hypothesis (left), and the corresponding CDF with horizontal lines indicating the 2.5%, 16%, 50%, 84%, and 97.5% quantiles, and vertical green and yellow bands show the  $\pm 1\sigma$  and  $\pm 2\sigma$  ranges of  $\mu^{95\%CL}$  [125].

shown. It is found that with the large data sample size (asymptotic regime)<sup>[8]</sup>, the modified test statistic  $\tilde{q}_\mu$  is in the form,

$$\tilde{q}_\mu = \begin{cases} \frac{\mu^2}{\sigma^2} - \frac{2\mu\hat{\mu}}{\sigma^2} & \text{if } \hat{\mu} < 0 \\ \frac{(\mu - \hat{\mu})^2}{\sigma^2} & \text{if } 0 \leq \hat{\mu} \leq \mu \\ 0 & \text{if } \hat{\mu} > \mu \end{cases}$$

where  $\sigma$  is a factor that characterizes effects from all nuisance parameters. The pdf  $f(\tilde{q}_\mu|\mu)$  is found to follow a well defined formula (here the  $\theta$  is drop as the  $\sigma$  takes care of their effects)

$$f(\tilde{q}_\mu|\mu) = \frac{1}{2}\delta(\tilde{q}_\mu) + \begin{cases} \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{q}_\mu}}e^{-\tilde{q}_\mu/2} & \text{if } 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2 \\ \frac{1}{\sqrt{2\pi}(2\mu/\sigma)}\exp\left(-\frac{1}{2}\frac{(\tilde{q}_\mu + \mu^2/\sigma^2)^2}{(2\mu/\sigma)^2}\right) & \text{if } \tilde{q}_\mu > \mu^2/\sigma^2 \end{cases}$$

where

$$\sigma^2 = \frac{\mu^2}{q_{\mu,A}}. \quad (3.22)$$

$q_{\mu,A}$  is the test statistic evaluated with the expected background and nominal nuisance parameters.  $A$  in the  $q_{\mu,A}$  stands for the Asimov data set<sup>4</sup>. The same construction can also be used for  $f(\tilde{q}_\mu|b - \text{only})$ . A novel result states that, by assuming the large sample size, one can obtain the exact formulae for  $f(\tilde{q}_\mu|\mu)$  and  $f(\tilde{q}_\mu|0)$ , whose parameter  $\sigma$  can be extracted from a single representative Asimov data set. The median expected limits and their bands are therefore easily obtained using this data set, without performing any generation of pseudo experiments. The median expected CL<sub>s</sub> limit,  $\mu_{\text{up}}^{\text{med}}$  is expressed as

$$\mu_{\text{up}}^{\text{med}} = \hat{\mu} + \sigma\Phi^{-1}(1 - 0.5\alpha), \quad (3.23)$$

---

<sup>3</sup>This is a critical assumption, which enables us to factorize the test statistic into Gaussian and non-Gaussian part. In the large data sample limit, the contribution from the non-Gaussian component is negligible.

<sup>4</sup>The Asimov data set is defined such that when maximizing the likelihood associated to this data set, one would get the maximum likelihood estimators of the parameters to be the assumed (true) values of the parameters.

and the  $\pm n\sigma$  band is given by

$$\text{Band}_{n\sigma} = \hat{\mu} + \sigma(\Phi^{-1}(1 - \alpha) \pm n), \quad (3.24)$$

where  $\Phi^{-1}$  is the inverse of the cumulative distribution of the standard Gaussian.

The  $\alpha = 0.05$  is chosen corresponding to the 95% CL.

# Chapter 4

## Results and conclusion

### 4.1 Limits on decay branching fraction

The distributions in  $m_{\mu\mu\gamma}$  observed in the data are in agreement with the SM expectation of the background-only hypothesis. The results are used to derive upper limits on the branching fractions,  $\mathcal{B}(Z \rightarrow J/\psi \gamma)$  and  $\mathcal{B}(H \rightarrow J/\psi \gamma)$  for the Z and Higgs boson.

The observed (expected) exclusion upper limit on the cross-section times the branching fraction at 95% CL for the  $H \rightarrow J/\psi \gamma$ , where the  $J/\psi$  meson is fully transversely polarized, is,

$$\sigma(pp \rightarrow H) \times \mathcal{B}(H \rightarrow J/\psi \gamma \rightarrow \mu\mu\gamma) < 2.5 (1.7^{+0.8}_{-0.5}) \text{ fb}, \quad (4.1)$$

where upper and lower bounds for 68% of interval of expected limits are shown as superscript and subscript. With the known values of  $\sigma(pp \rightarrow H) = 55.1 \text{ pb}$  and  $\mathcal{B}(J/\psi \rightarrow \mu\mu) = 0.059$ , the above result can be interpreted in terms of limit on the branching fraction,

$$\mathcal{B}(H \rightarrow J/\psi \gamma) < 7.6 (5.2^{+2.4}_{-1.6}) \times 10^{-4}. \quad (4.2)$$

which corresponds to 260 (170) times the SM prediction.

For the Z boson decay, with the unpolarized J/ $\psi$  meson assumption ,the observed (expected) upper limit on the cross-section times the branching fraction is,

$$\sigma(pp \rightarrow Z) \times \mathcal{B}(Z \rightarrow J/\psi \gamma \rightarrow \mu\mu\gamma) < 4.6 (5.3^{+2.3}_{-1.6}) \text{ fb}, \quad (4.3)$$

With the known value  $\sigma(pp \rightarrow Z) = 5.71 \times 10^4 \text{ pb}$ , the observed (expected) upper limit in branching fraction is

$$\mathcal{B}(Z \rightarrow J/\psi \gamma) < 1.4 (1.6^{+0.7}_{-0.5}) \times 10^{-6}. \quad (4.4)$$

, corresponding to 15 (18) times the SM prediction.

Extreme polarization scenarios give rise to variations from  $-13.6$  ( $-13.5$ )%, for a fully longitudinally polarized J/ $\psi$ , to  $+8.6$  ( $+8.2$ )%, for a fully transversely polarized J/ $\psi$  meson, in the observed (expected) branching fraction. The observed (expected) exclusion limits on the cross sections and branching fractions at 95% confidence level for the Z and Higgs boson decays are summarized in Table 4.1.

Channel	Polarization scenario	$\sigma$ (fb)	$\mathcal{B}(Z \text{ (H)} \rightarrow J/\psi \gamma)$	$\frac{\mathcal{B}(Z \text{ (H)} \rightarrow J/\psi \gamma)}{\mathcal{B}_{\text{SM}}(Z \text{ (H)} \rightarrow J/\psi \gamma)}$
$Z \rightarrow J/\psi \gamma$	Unpolarized	$4.6 (5.3^{+2.3}_{-1.6})$	$1.4 (1.6^{+0.7}_{-0.5}) \times 10^{-6}$	15 (18)
	Transverse	$5.0 (5.9^{+2.5}_{-1.7})$	$1.5 (1.7^{+0.7}_{-0.5}) \times 10^{-6}$	16 (19)
	Longitudinal	$3.9 (4.6^{+2.0}_{-1.4})$	$1.2 (1.4^{+0.6}_{-0.4}) \times 10^{-6}$	13 (15)
$H \rightarrow J/\psi \gamma$	Transverse	$2.5 (1.7^{+0.8}_{-0.5})$	$7.6 (5.2^{+2.4}_{-1.6}) \times 10^{-4}$	260 (170)

Table 4.1: Upper observed (expected) limits on cross sectiona  $\sigma(pp \rightarrow Z \text{ (H)} \rightarrow (J/\psi \rightarrow \mu\mu)\gamma)$  (fb) and branching fractions of  $Z \text{ (H)} \rightarrow J/\psi \gamma$  decays, where the latter are computed assuming SM cross section of the Z (H) boson. Variations of the branching fractions of the Z decay for complete transverse and longitudinal polarizations for J/ $\psi$  are also shown. The upper and lower bounds of the expected 68% confidence level interval for the expected limits are shown as superscripts and subscripts respectively.

## Combination with 8 TeV result

The results of the  $H \rightarrow J/\psi \gamma$  are combined with the results of a similar search performed by the CMS Collaboration using data recorded with pp collisions at  $\sqrt{s} = 8$  TeV, corresponding to an integrated luminosity of  $19.7 \text{ fb}^{-1}$  [87]. The combination results in an upper limit corresponding to 220 (160) times the SM prediction. All systematic uncertainties are assumed to be uncorrelated in the combination, apart from the theoretical calculations for the cross section and branching fractions. The two experiments were conducted at different energies, the triggers were substantially different to deal with the increased luminosity, and the reconstruction algorithms were different and more sophisticated to deal with the larger backgrounds. Besides, independent calibrations were performed for different years of data taking and thus it is unlikely that there is much correlation. Nonetheless, a calculation assuming the uncertainties are fully correlated, as the extreme situation, is performed and the difference in the final result is 0.3%, which is negligible.

## 4.2 Conclusion

A search is performed for decays of the standard model (SM) Z and Higgs bosons into a  $J/\psi$  meson and a photon with the  $J/\psi$  meson subsequently decaying into  $\mu\mu$ . Data from pp collisions at  $\sqrt{s} = 13$  TeV, corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$  is used. No excess has been observed above the predicted background. The observed (expected) exclusion limit at 95% CL on the branching fraction of the Higgs boson is set at  $\mathcal{B}(H \rightarrow J/\psi \gamma) < 7.6 (5.2) \times 10^{-4}$ , corresponding to 260 (170) times the SM value. The 68% confidence level interval ranges from 3.6 to  $7.6 \times 10^{-4}$ . The limit on the branching fraction of the Z boson decay in the unpolarized scenario is set at  $\mathcal{B}(Z \rightarrow J/\psi \gamma) < 1.4 (1.6)$ , corresponding to 15 (18) times the SM prediction. The 68% confidence level interval ranges from 1.1 to  $2.3 \times 10^{-6}$ . Extreme polarization scenarios give rise to variations from  $-13.6$  ( $-13.5$ )%, for a

fully longitudinally polarized  $J/\psi$  meson, to  $+8.6$  ( $+8.2\%$ ), for a fully transversely polarized  $J/\psi$  meson, in the observed (expected) branching fraction. The results for the Higgs boson channel are combined with the results obtained by a similar search performed at  $\sqrt{s} = 8$  TeV by the CMS Collaboration, yielding an observed (expected) upper limit on the branching fraction for the decay  $H \rightarrow J/\psi \gamma$  of 220 (160) times the SM prediction.

### 4.3 Outlook

Improvements can be done in order to make the analysis more advanced. The proper simulation of the background processes is of the first priority. The difficulty is mainly due to the large cross sections of the low mass dimuon system in the final states, and therefore efficient ways to produce such samples should be developed. The background samples will enable us to have better understanding of the background composition and make the optimization of the event selection feasible. Furthermore, the multivariate analysis (MVA) or the matrix element method<sup>1</sup>(MEM) can be exploited to better discriminate the signal and background.

The analysis can be extended to include the decay of  $Z(H) \rightarrow Y(nS)\gamma$ , where the  $Y(nS)$  mesons decay to a muon pair. The one dimension fit in the  $m_{\mu\mu\gamma}$  space to estimate the background in this analysis will need to be modified to cope with the non-negligible contribution of the peaking background  $Z \rightarrow \mu\mu\gamma$ . A two dimension (in the  $m_{\mu\mu}$  and  $m_{\mu\mu\gamma}$  space) or multi-dimension fit is suggested. The background composition can also be estimated by this data-driven method, and in turn can be used to validate the background simulation samples. The development of the identification and reconstruction of merged electrons can be used in the electron channel. The projection study is performed, and the expected distributions of  $m_{\mu\mu\gamma}$  with  $3000 \text{ fb}^{-1}$  of data from both decay channels are shown in Fig. 4-1. The upper limit on  $\mathcal{B}(Z \rightarrow J/\psi \gamma)$  is around 2 times its SM value, while that on the

---

<sup>1</sup>For example, the Matrix Element Likelihood Analysis (MELA) used in the  $H \rightarrow ZZ^* \rightarrow 4l$  analysis.

$\mathcal{B}(H \rightarrow J/\psi \gamma)$  is expected to be less than 20 times its SM prediction. With the addition of the electron channel and foreseeable improvements, the  $Z \rightarrow J/\psi \gamma$  would be sensitive to its current SM predicted rate after the high luminosity run of the LHC, possibly leading to the first observation of this rare decay of the  $Z$  boson.

Apart from the analysis techniques, projects of detector upgrade to improve the capability of the CMS are planned. One of the important projects is the high granularity endcap calorimeter (HGCal). I involved in the beamtest for the HGCal and worked on the energy reconstruction and basic particle identification in the testbeam data. The study is described in Appendix C.

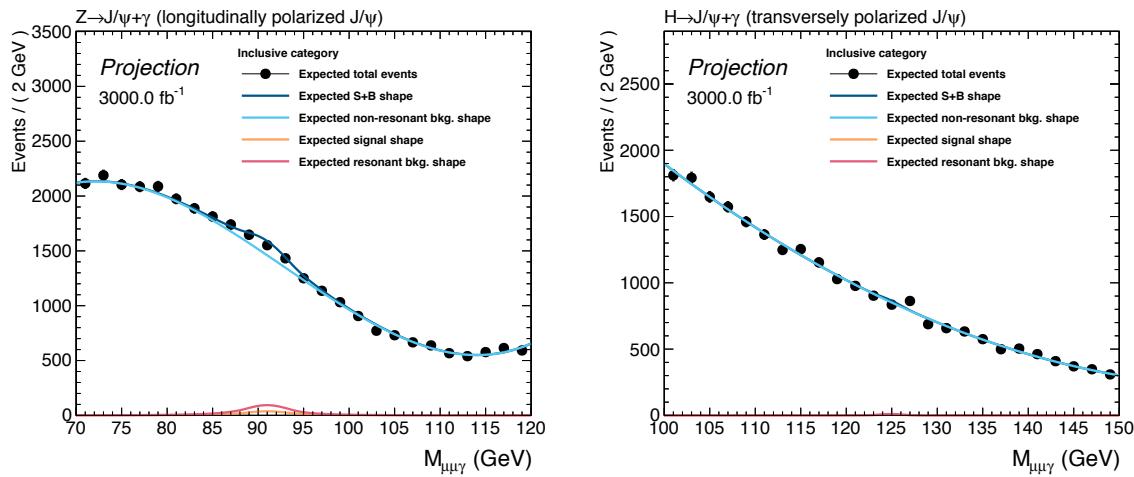


Figure 4-1: The expected distributions of  $m_{\mu\mu\gamma}$  at  $3000 \text{ fb}^{-1}$  of data from both decay channels.

# **Appendix A**

## **Additional materials for the bias study**

### **A.1 Linearity**

It was suggested to do the bias study with more signal events introduced when generating the pseudo-event. Following plots show how the mean and width of the pull distribution evolve as more signal events are introduced.

### A.1.1 $H \rightarrow J/\psi \gamma$

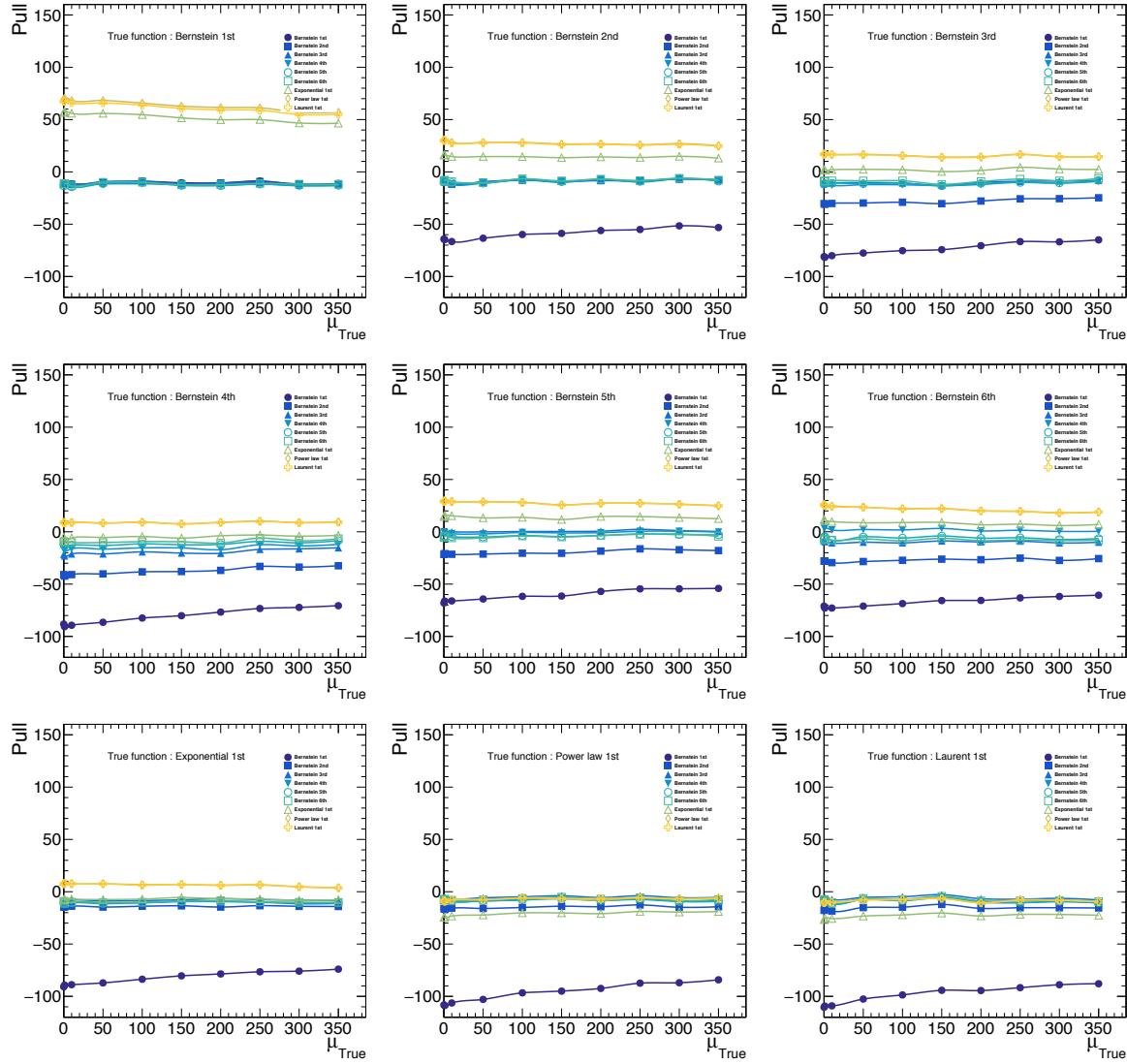


Figure A-1: The evolution of the mean of the pull value distribution as more signal events are introduced in the Higgs decay.

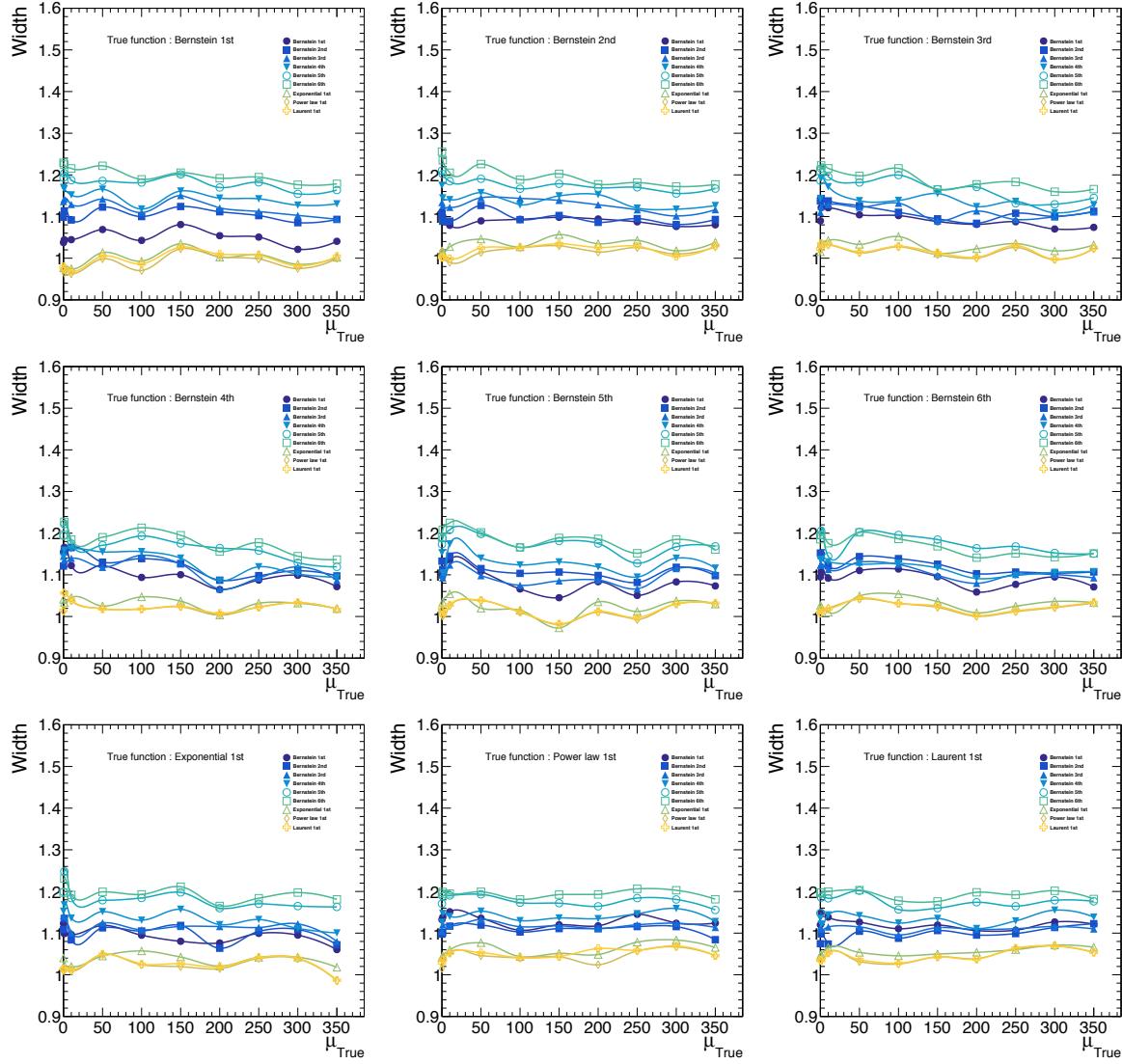


Figure A-2: The evolution of the width of the pull value distribution as more signal events are introduced in the Higgs decay.

### A.1.2 $Z \rightarrow J/\psi \gamma$ Cat1

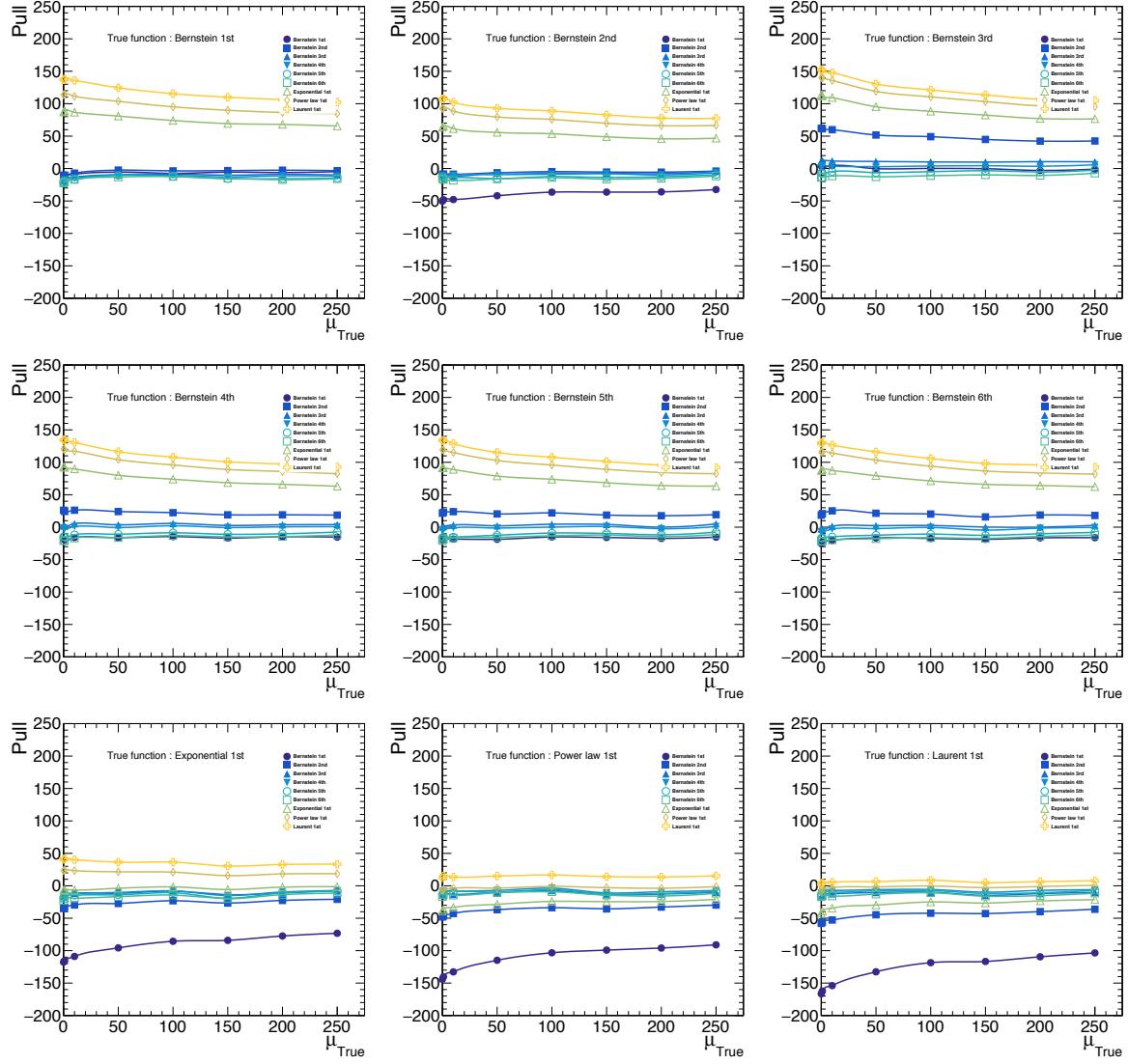


Figure A-3: The evolution of the mean of the pull value distribution as more signal events are introduced in the Cat1 of the  $Z$  decay.

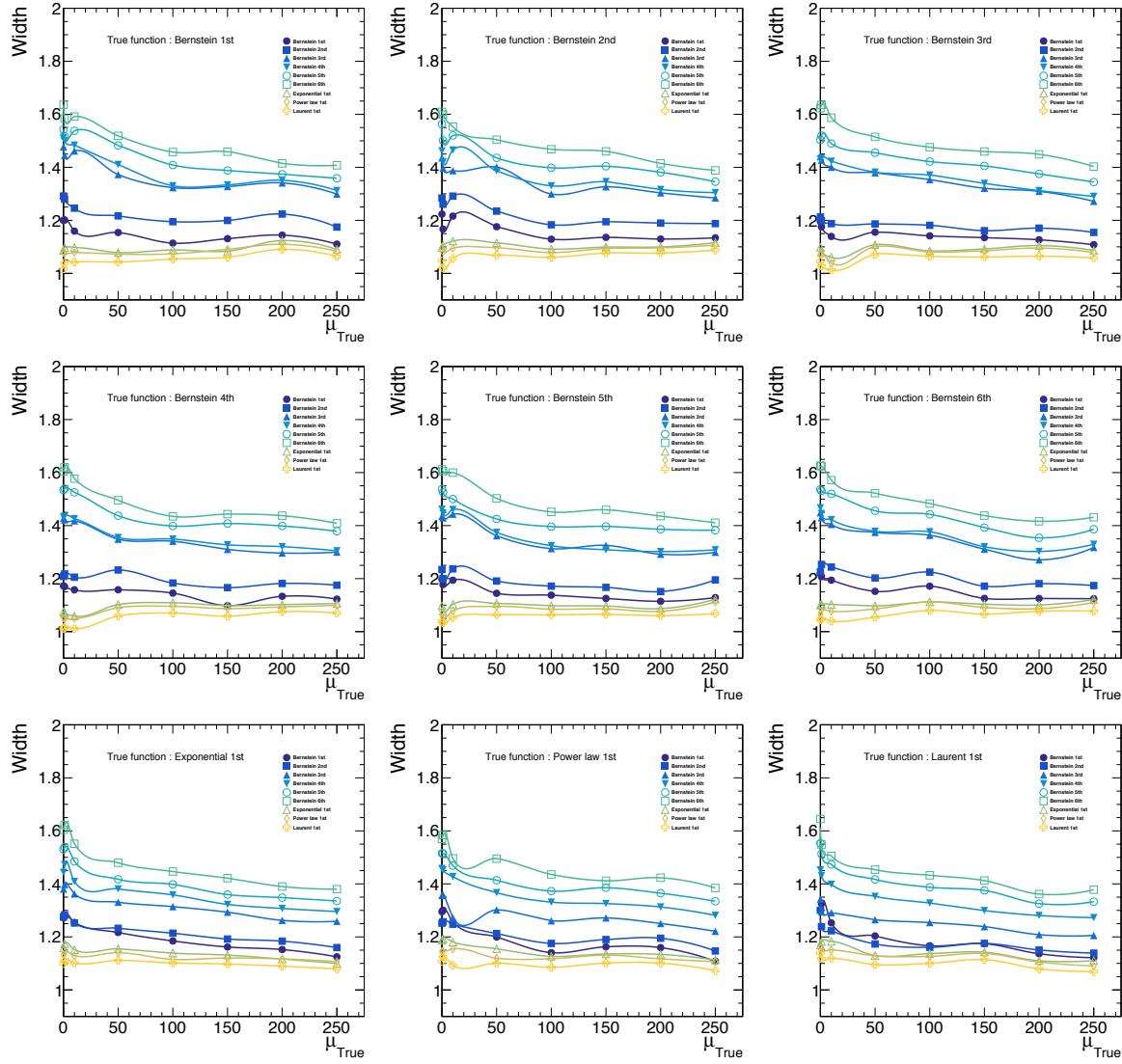


Figure A-4: The evolution of the width of the pull value distribution as more signal events are introduced in the Cat1 of the Z decay.

### A.1.3 $Z \rightarrow J/\psi \gamma$ Cat2

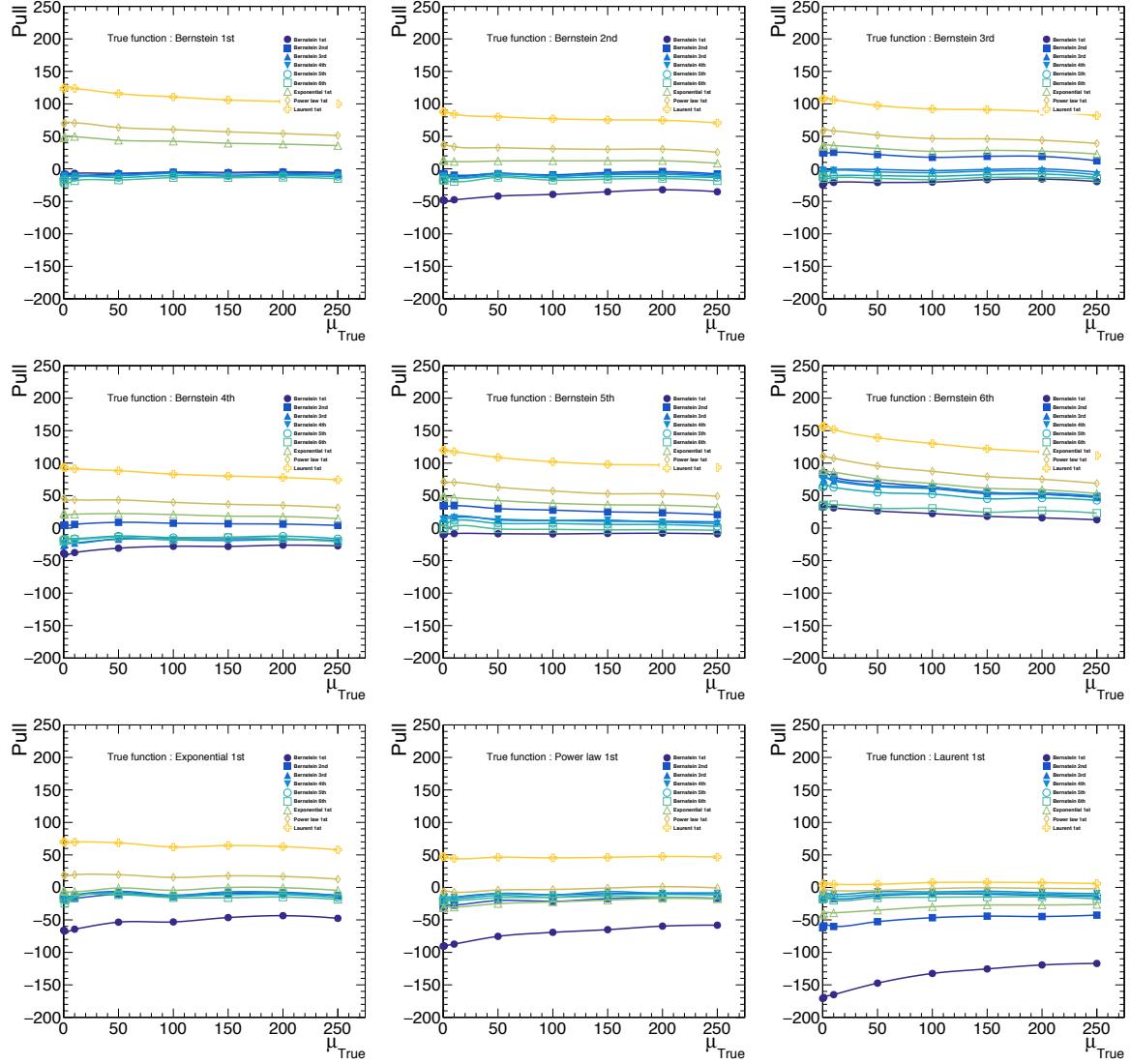


Figure A-5: The evolution of the mean of the pull value distribution as more signal events are introduced in the Cat2 of the  $Z$  decay.

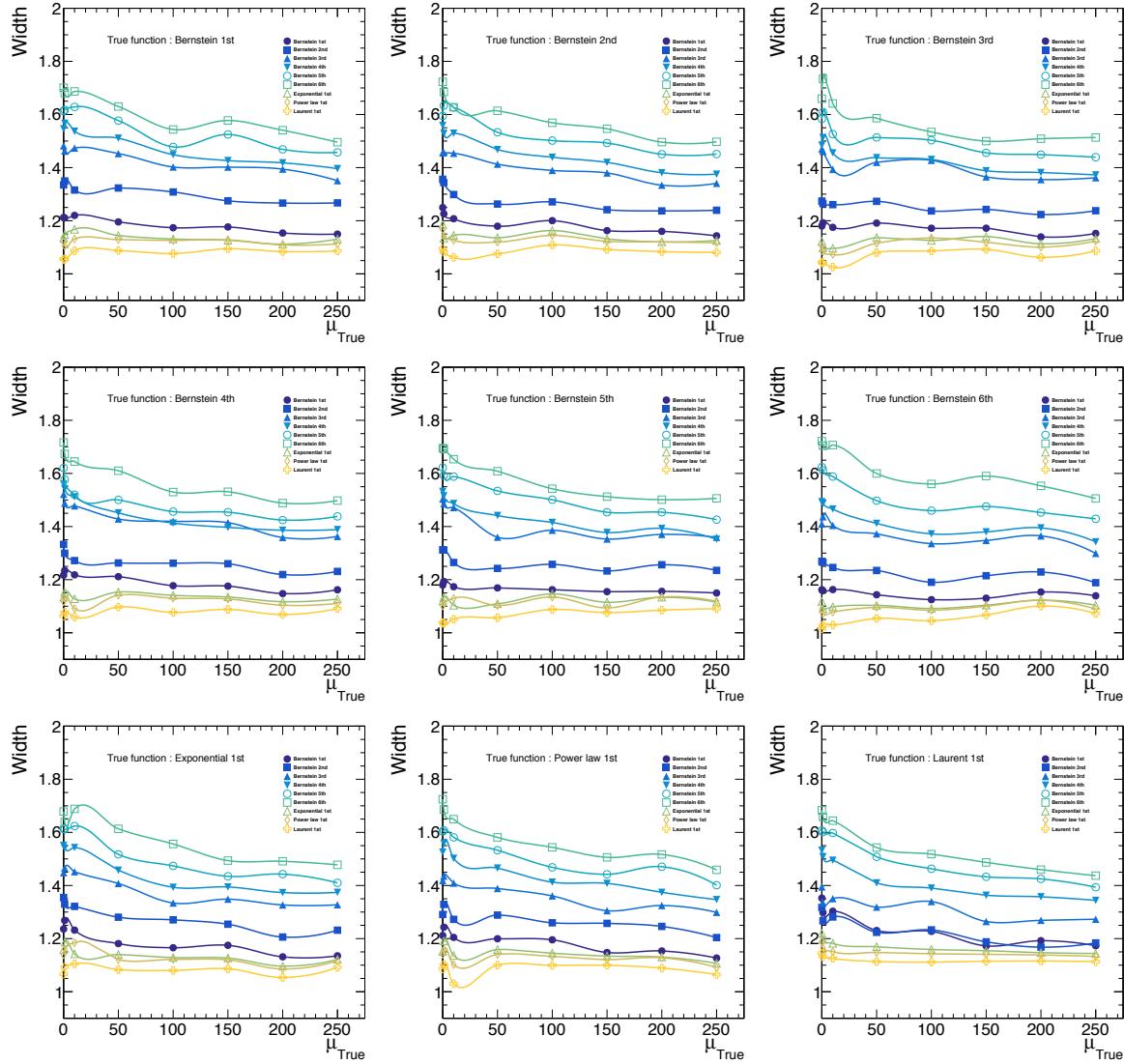


Figure A-6: The evolution of the width of the pull value distribution as more signal events are introduced in the Cat2 of the Z decay.

### A.1.4 $Z \rightarrow J/\psi \gamma$ Cat3

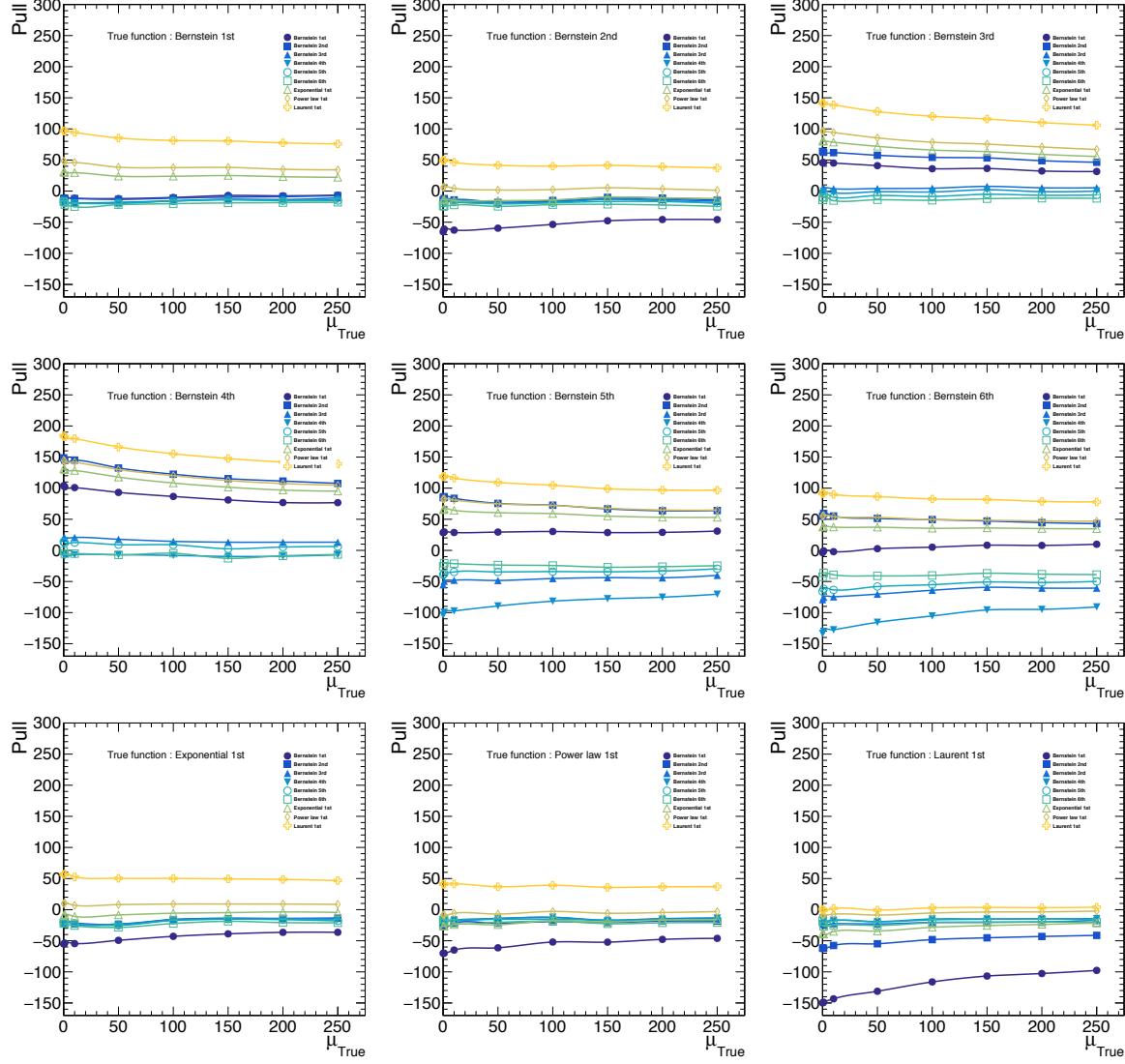


Figure A-7: The evolution of the mean of the pull value distribution as more signal events are introduced in the Cat3 of the  $Z$  decay.

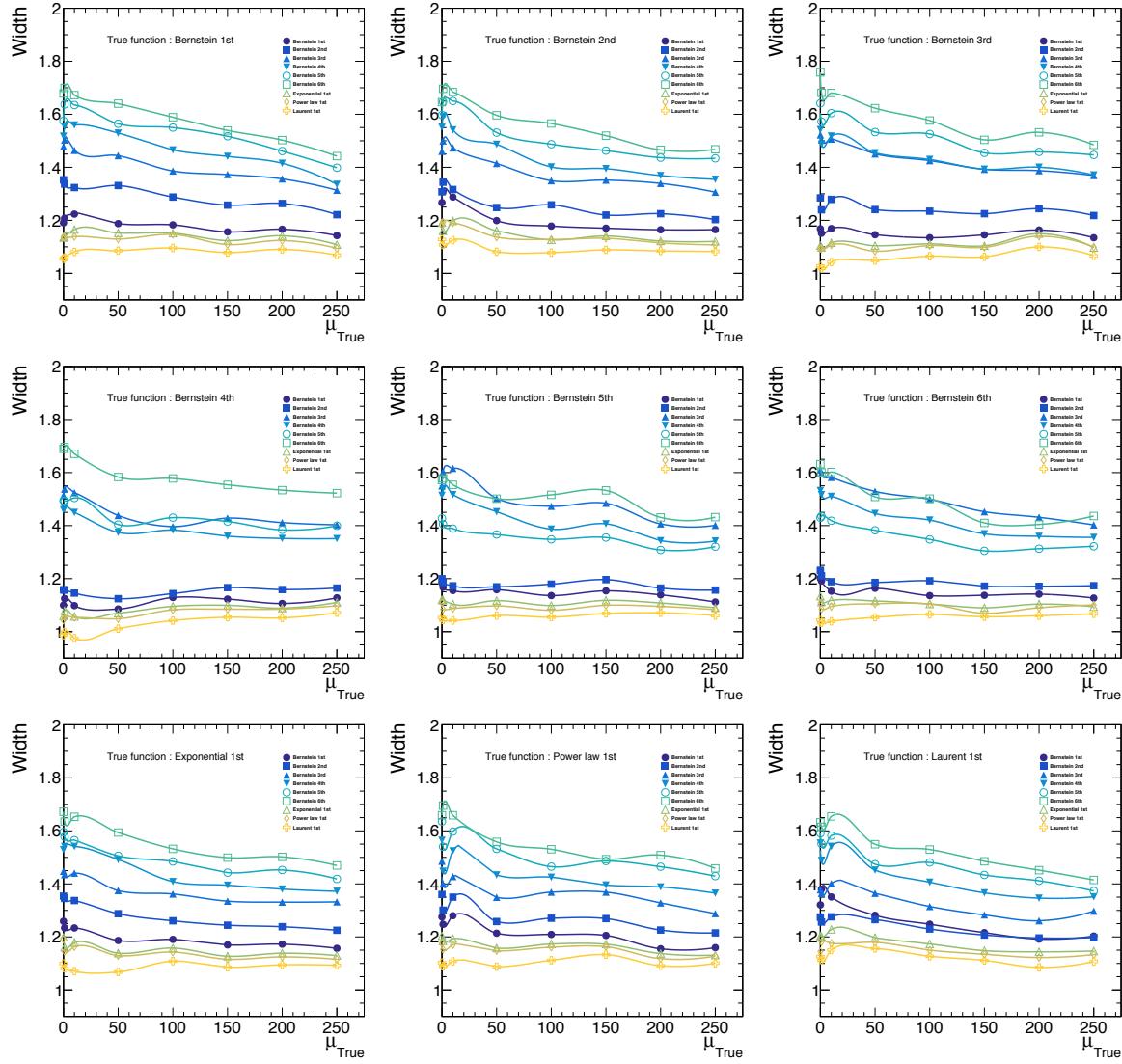


Figure A-8: The evolution of the width of the pull value distribution as more signal events are introduced in the Cat3 of the Z decay.

## A.2 Pseudo-event

Examples of pseudo-events for the Higgs and all the three categories of Z boson searches are shown in this section. The pseudo-events are generated from the least-bias functions for each category. The fits using the least-bias functions are also shown in the plots, where the green one is the signal component of the resulting fit, red one is the background component, and the blue one is the combination of the signal and background component.

### A.2.1 Pseudo-events for $H \rightarrow J/\psi \gamma$

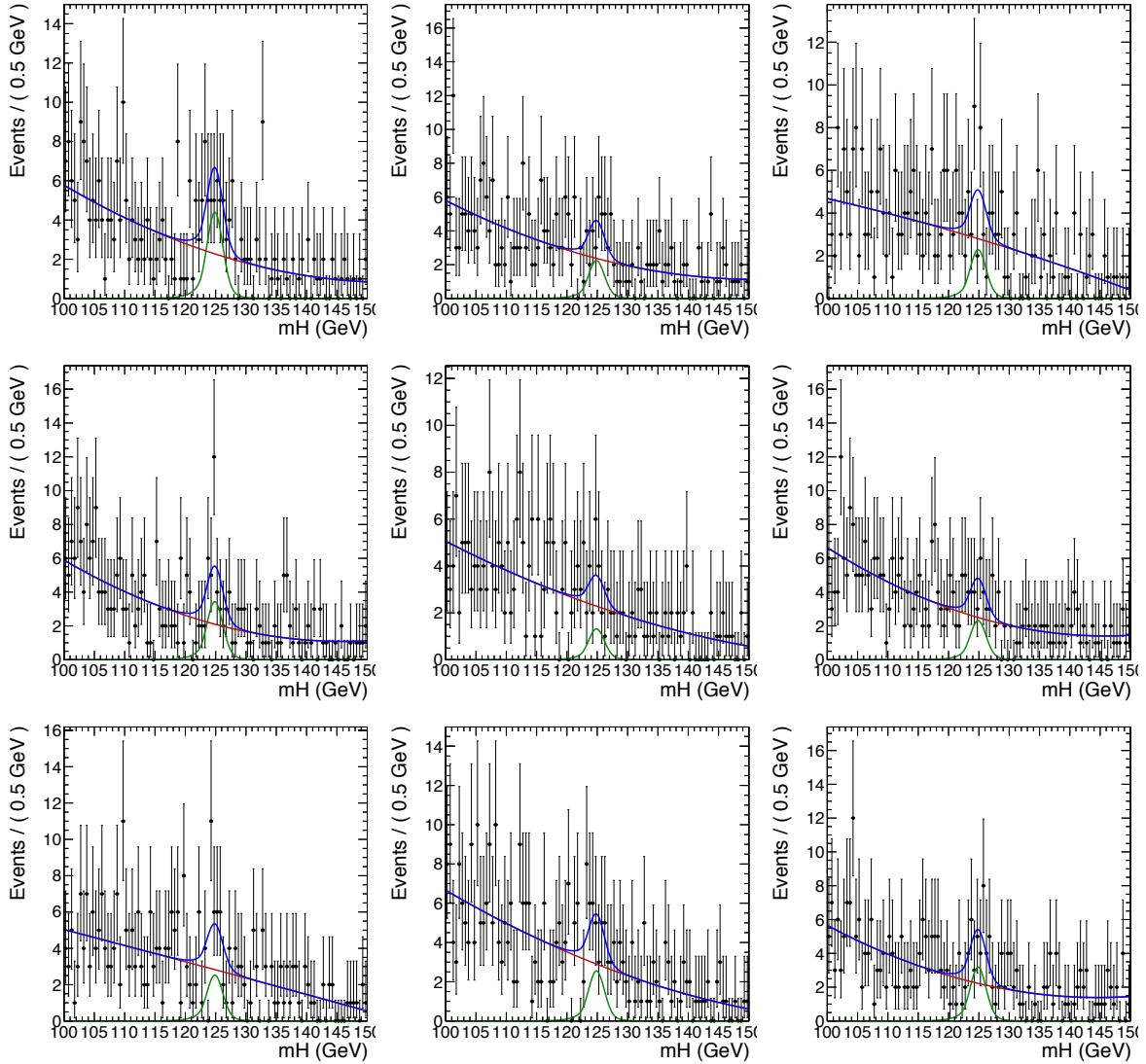


Figure A-9: Examples of the pseudo-events for bias study in Higgs search. The toys are generated from Bernstein polynomial of 2nd order, and the background fit (red line) is the Bernstein polynomial of 2nd order.

### A.2.2 Pseudo-events for Cat1 of $Z \rightarrow J/\psi \gamma$

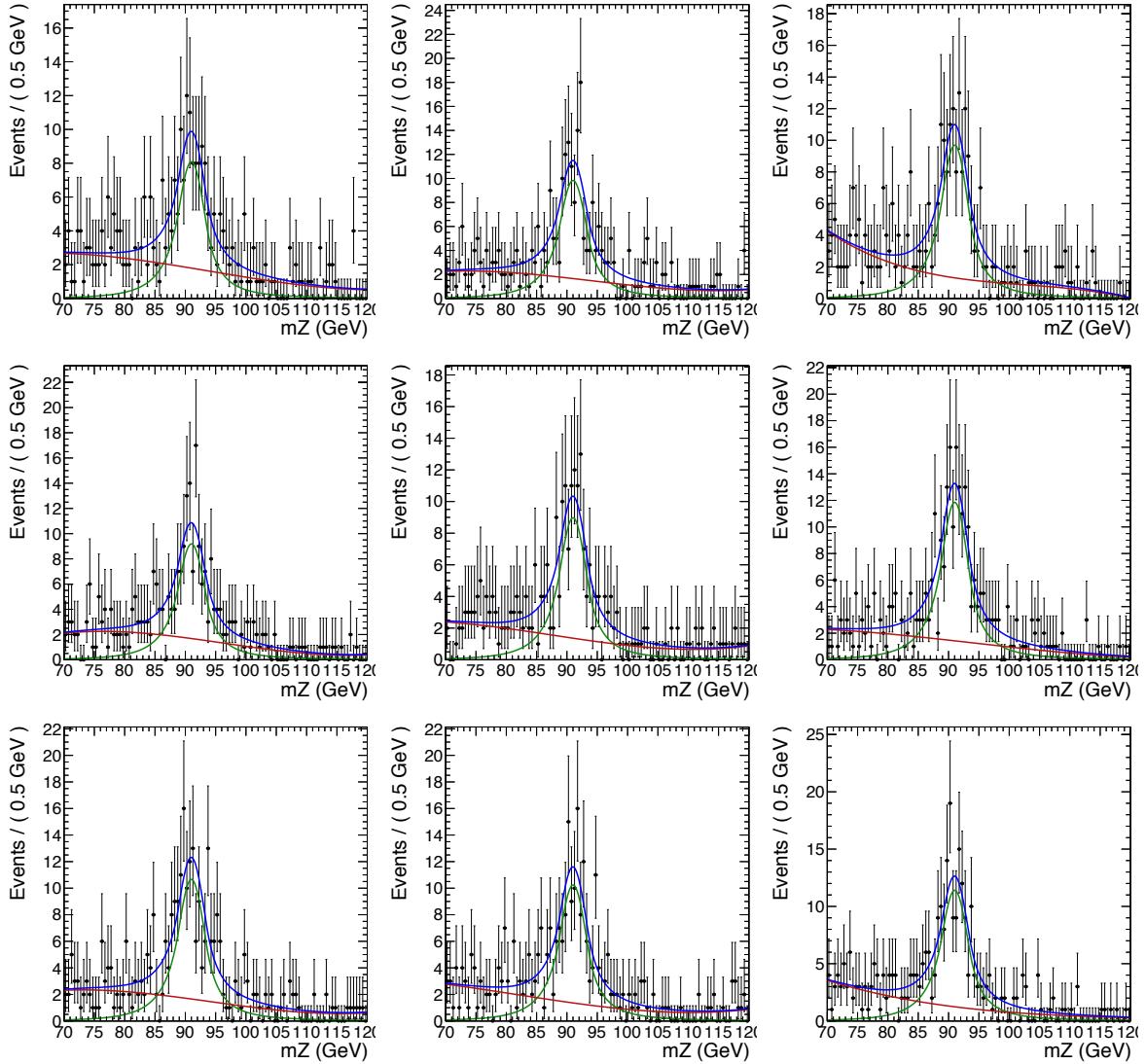


Figure A-10: Examples of the pseudo-events for bias study in Cat1 of  $Z$  search. The toys are generated from Bernstein polynomial of 3rd order, and the background fit (red line) is the Bernstein polynomial of 3rd order.

### A.2.3 Pseudo-events for Cat2 of $Z \rightarrow J/\psi \gamma$

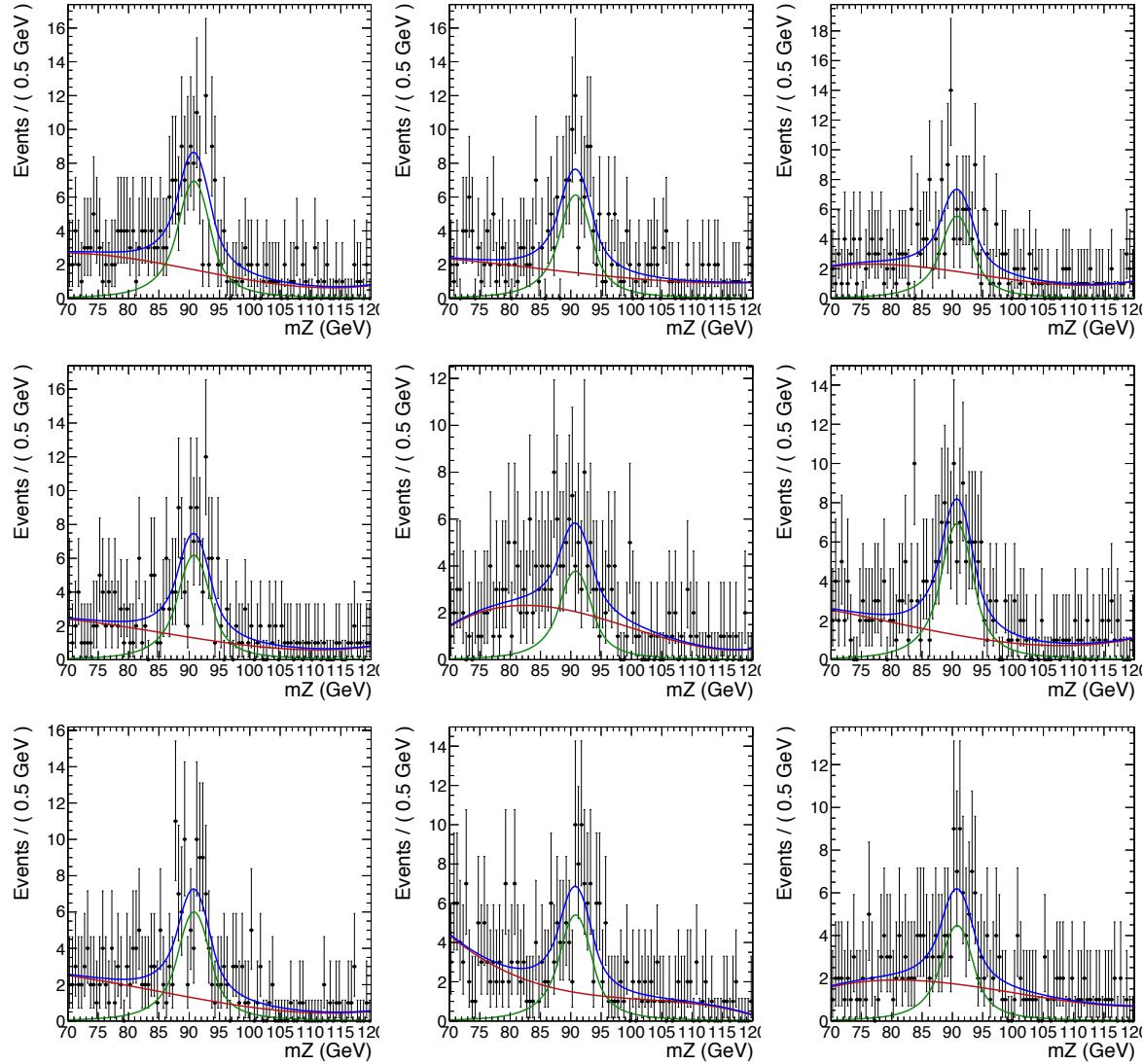


Figure A-11: Examples of the pseudo-events for bias study in Cat2 of  $Z$  search. The toys are generated from Bernstein polynomial of 3rd order, and the background fit (red line) is the Bernstein polynomial of 3rd order.

#### A.2.4 Pseudo-events for Cat3 of $Z \rightarrow J/\psi \gamma$

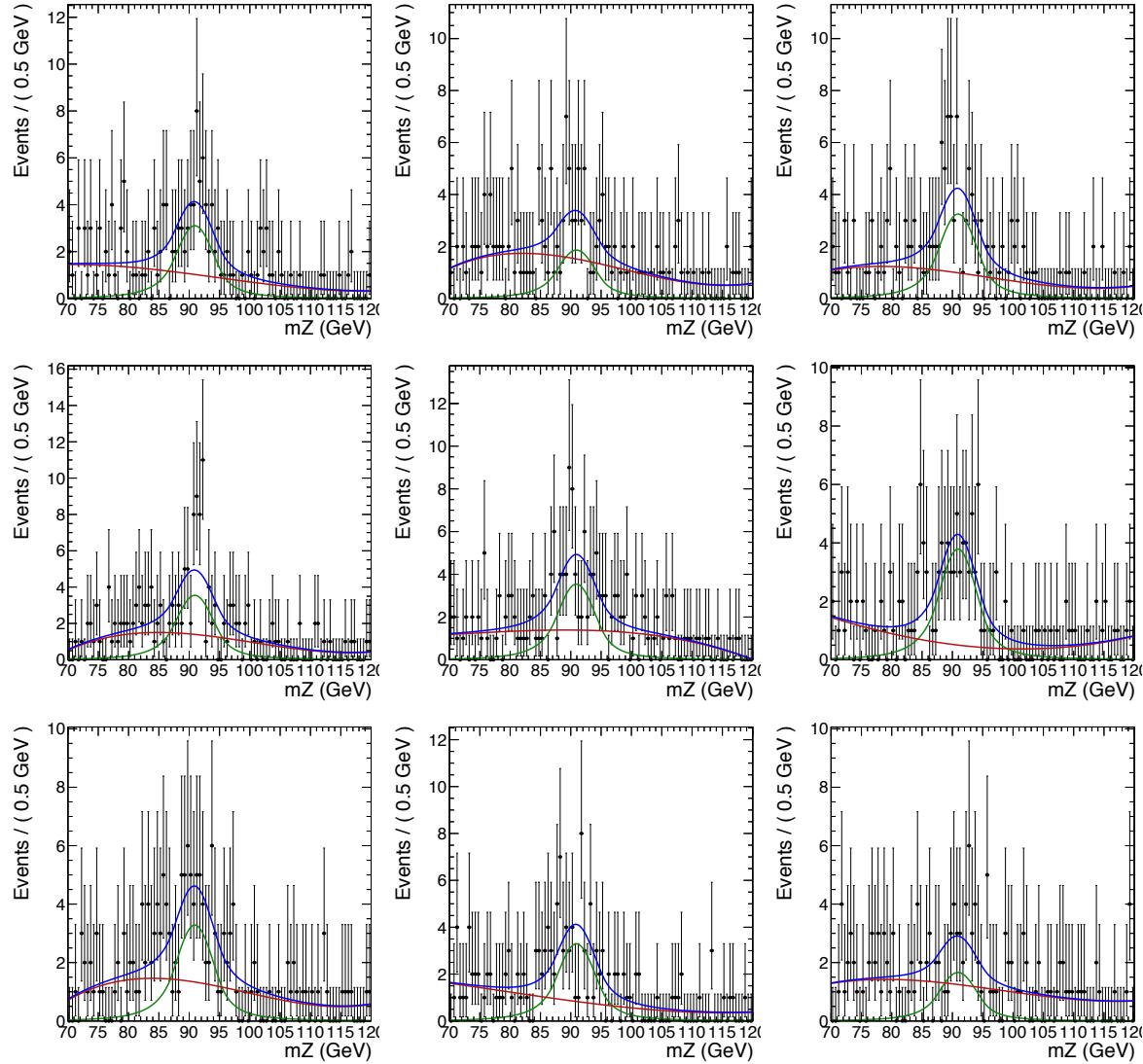


Figure A-12: Examples of the pseudo-events for bias study in Cat3 of  $Z$  search. The toys are generated from Bernstein polynomial of 3rd order, and the background fit (red line) is the Bernstein polynomial of 3rd order.

# Appendix B

## Discussion on the systematic uncertainties

Comments and discussions on systematic uncertainties are summarized as follows.

- **Pileup.** The uncertainty in the Cat2 in the Z decay is small compared to other categories. No weird behavior in the pileup weights of all three categories is found, no mistake is made when the pileup weights are evaluated and applied. Table B.1 shows the detail numbers that give the final uncertainties in all the categories. Fig. B-1 shows the distributions of the difference between the up (down) variation and the nominal pileup weight of all the three categories in the Z decay. Fig. B-2 shows the 2D distributions of the difference between the up (down) variation and the nominal pile-up weight versus the photon  $R_9$  value. In Fig. B-3, the x-axis is the event number while the y-axis is the difference with respect to the sum of nominal pile-up weight over all events. This plot clearly shows how the difference evolves with the events in each category. As one can see, such small uncertainty in EBLR9 category is due to the cancellation of positive and negative weights.

	[1]	[2]	fraction [1]/[2] (in %)	[3]	[4]	[5]	Uncertainty (in %)
EBHR9	4423	5447	44.8	589.0	-687.3	10050.1	-0.98
EBLR9	2898	3257	47.1	387.2	-399.7	6213.0	-0.20
EE	1800	2287	44.0	234.7	-290.9	4196.7	-1.34
	[1]	[2]	fraction [1]/[2] (in %)	[3]	[4]	[5]	Uncertainty (in %)
EBHR9	4956	4914	50.2	728.8	-629.1	10050.1	0.99
EBLR9	2910	3245	47.3	418.6	-413.0	6213.0	0.091
EE	2074	2013	50.7	307.9	-254.1	4196.7	1.28

[1]: number of events where ( $\text{puwei\_up}/\text{down} - \text{puwei}$ )> 0  
[2]: number of events where ( $\text{puwei\_up}/\text{down} - \text{puwei}$ )< 0  
[3]: sum over positive value of ( $\text{puwei\_up}/\text{down} - \text{puwei}$ )  
[4]: sum over negative value of ( $\text{puwei\_up}/\text{down} - \text{puwei}$ )  
[5]: sum over all puwei

Table B.1: The uncertainties in pile-up weight of each category.

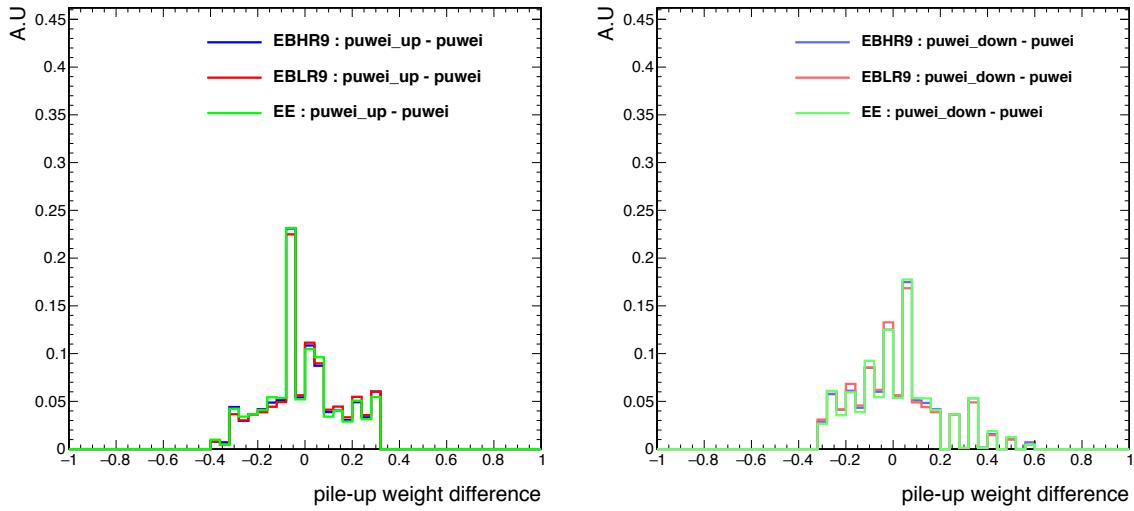


Figure B-1: The 1D distributions of the difference between the up(down) variation and the nominal pile-up weight of all the 3 categorizes in the Z decay.

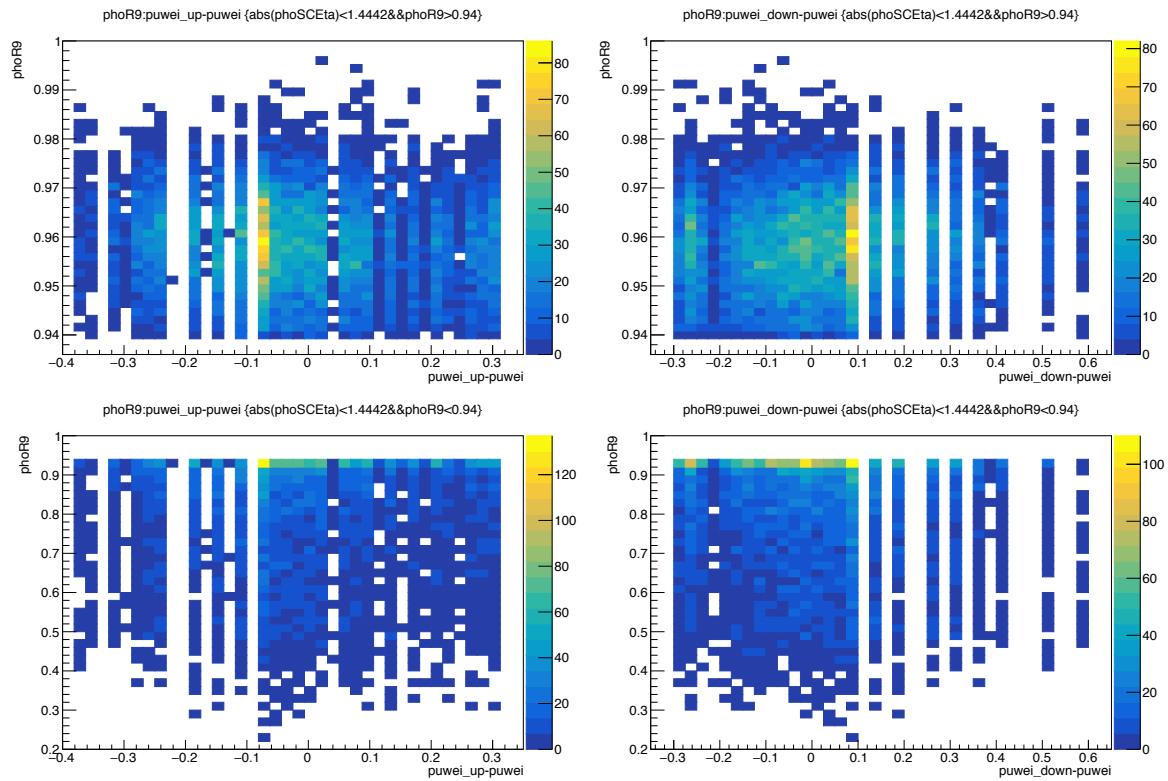


Figure B-2: The 2D distributions of the difference between the up(down) variation and the nominal pile-up weight versus the photon  $R_9$  value. (Top left) ( $puwei_{up} - puwei$ ) v.s photon  $R_9$  in EBHR9; (Top right) ( $puwei_{down} - puwei$ ) v.s photon  $R_9$  in EBHR9; (Bottom left) ( $puwei_{up} - puwei$ ) v.s photon  $R_9$  in EBLR9; (Bottom right) ( $puwei_{down} - puwei$ ) v.s photon  $R_9$  in EBLR9

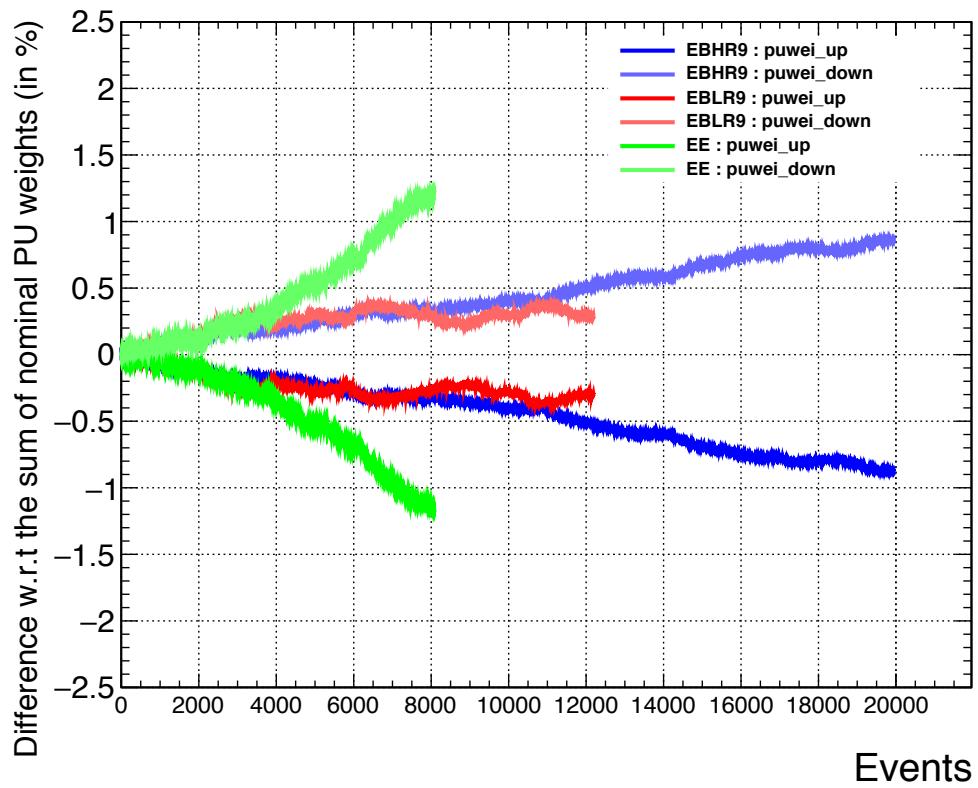


Figure B-3: The evolution of the difference with respect to the sum of nominal pile-up weights of all the 3 categories in the Z decay.

- **Muon ID/Isolation.** Muons from Z boson decay are softer than those from the Higgs boson decay, which can be seen in Fig. [3-25], [3-26], [3-27], [3-28]. The  $p_T$  of muons from the Z boson decay distribute mostly in the range of 20~30 GeV, while those from the Higgs boson decay are mostly in the range of 30~40 GeV, and from the Fig. [3-18] one can see that uncertainties in the range of 20~30 GeV are slightly higher than those in 30~40 GeV. Consequently, uncertainties in muon ID and isolation in the Z boson decay are higher than those in the Higgs boson decay.
- **Electron veto.** As shown in the Fig. [3-21], the uncertainty of photons in endcap region is smaller than that of the photon in barrel region by a factor of  $0.0044/0.0119 = 0.37(37\%)$ . The ratio of the uncertainty on the yields in categories of barrel and endcap should be comparable to this number,  $0.450/1.200.375 (37.5\%)$ . Therefore, the difference of uncertainties between barrel and endcap region is reasonable.
- **Scale uncertainty in the signal modeling.** The individual uncertainty from each source in each category of the Z boson decay is shown in Table [B.2]. There are four sets of variation in the muon momentum correction and three sets in the photon energy correction. The final uncertainty in each category are summed in quadrature over the muon and photon part.
- **Resolution uncertainty in the signal modeling.** The uncertainties in the  $\sigma$  of the signal model are larger in the Higgs boson decays than in the Z decay. No unusual behaviors in the distributions of  $m_{\mu\mu\gamma}$  resulting from different sets of correction is found, and fits are all reasonable. The difference may come from the correction itself, for which individual analysis cannot do much. The natural width of the Z boson itself is larger, and so relative uncertainty becomes smaller compared to the Higgs boson case. In addition, for the Z decay the first two categories for barrel photons where the uncertainties are smaller, while in Higgs all events are combined and uncertainties from different kinematic regime are averaged. Uncertainties in the muon and

photon correction separately are summarized in Table B.3 and B.4. The total uncertainty is derived by summing the uncertainties in the muon and photon parts in quadrature.

Table B.2: The uncertainties in the mean of the signal model from muon and photon correction.

	Cat1 EBHR9		Cat2 EBLR9		Cat3 EE	
	Scale	Uncertainty (in %)	Scale	Uncertainty (in %)	Scale	Uncertainty (in %)
Nominal	91.002		90.768		90.950	
Muon - Set1	91.004	0.00220	90.785	0.0187	90.966	0.0176
Muon - Set2	90.997	0.00549	90.782	0.0154	90.961	0.0121
Muon - Set4	90.992	0.0110	90.785	0.0187	90.956	0.00660
Muon - Set5	90.997	0.00549	90.782	0.0154	90.957	0.00770
Muon - Total		0.0136		0.0343		0.0236
Photon - gain up	90.995	0.00769	90.772	0.00441	90.995	0.00769
Photon - gain down	90.995	0.00769	90.772	0.00441	90.995	0.00769
Photon - stat. up	90.996	0.00659	90.772	0.00441	91.000	0.00220
Photon - stat. down	90.994	0.00879	90.772	0.00441	90.991	0.0121
Photon - syst. up	91.030	0.0308	90.830	0.0683	91.046	0.0484
Photon - syst. down	90.960	0.0462	90.713	0.0606	90.945	0.0626
Photon - Total		0.0476		0.0686		0.0643
Total uncertainty		0.0495		0.0767		0.0685

Table B.3: The uncertainties in the sigma of the signal model from muon and photon correction in the H decay. The total uncertainty is derived by summing the uncertainties in the muon and photon parts in quadrature. The numbers in the table are in percentage.

	$H \rightarrow (J/\psi)\gamma$					
	ggF	VBF	ZH	$W^+H$	$W^-H$	$t\bar{t}H$
muon	1.69	1.27	1.60	1.38	2.00	2.97
photon	4.65	4.15	2.95	4.40	3.22	13.8
Total	4.94	4.30	3.35	4.61	3.79	14.1

Table B.4: The uncertainties in the sigma of the signal model from muon and photon correction in the Z decay. The total uncertainty is derived by summing the uncertainties in the muon and photon parts in quadrature. The numbers in the table are in percentage.

	$Z \rightarrow (J/\psi)\gamma$		
	Cat1	Cat2	Cat3
muon	0.44	0.38	0.49
photon	0.89	0.57	1.37
Total	0.99	0.69	1.45

# Appendix C

## Beam test for the CMS high granularity endcap calorimeter in 2018

The Run-2 of the Large Hadron Collider (LHC), starting in 2015 at a center-of-mass energy  $\sqrt{s} = 13$  TeV, has successfully came to an end in 2018. Despite the fact that plenty of significant results were obtained since the Run-1, an increase of integrated luminosity enables us to test the standard model (SM) in detail, measure its parameters more precisely, and probably open up a window to new physics. In the Run-2 period, the highest instantaneous luminosity was  $1.7 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , exceeding its original design, while the planned instantaneous luminosity is up to  $5.0 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  after the third long shutdown (LS3), which is scheduled from 2023 to late 2026. The operational phase after the LS3 is often referred to as High Luminosity LHC (HL-LHC). The projected LHC performance is shown in Fig. [C-1]. It is foreseeable that the high luminosity operation will impose great challenges for either radiation tolerance for detectors or event pileup for particle reconstructions and identifications<sup>[1]</sup>. During the LS3, extensive upgrades for different sub-detectors of the Compact Muon Solenoid (CMS) will be carried out. The existing endcap calorimeters will be replaced by the high granularity calorimeter (HGCAL). It includes two sections: electromagnetic (CE-E) and hadronic (CE-H) compartments. In the latest design, the former uses lead as the main absorber and hexagonal silicon sensors as the active detector. Fig. [C-2] shows the longitudinal

cross section of the upper half of one endcap calorimeter [128].

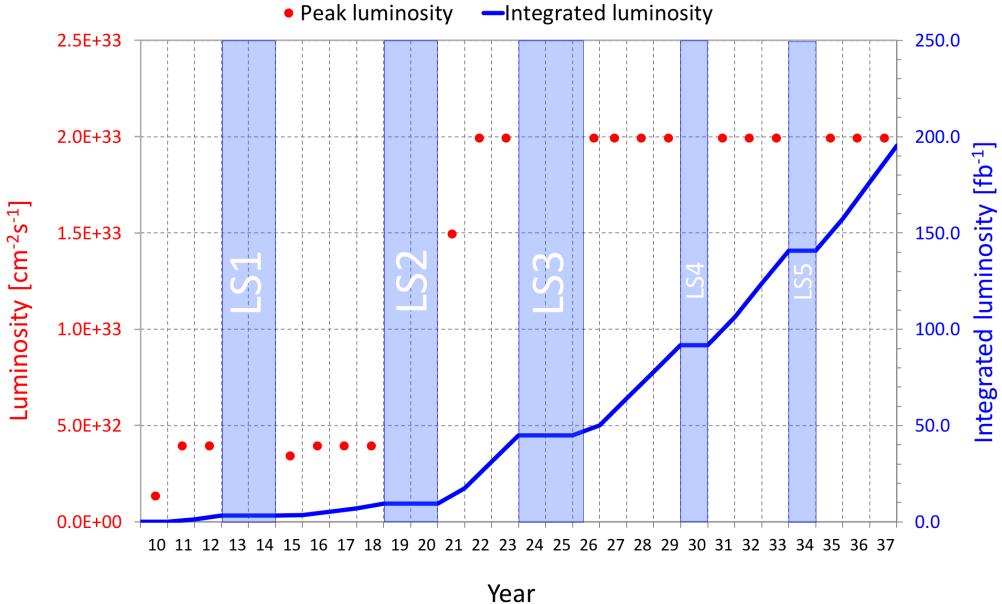


Figure C-1: Projected LHC performance through 2035, with dates of long shutdowns of LHC, periods of data-taking, and projected instantaneous and integrated luminosities.

In this beamtest, the HGCAL, including electromagnetic and hadronic compartments, and the analogue hadronic calorimeter (AHCAL) [130] were tested jointly. The actual setup for the beamtest is shown in Fig. C-3.

The simplified analysis flow is shown in Fig. C-4. My analysis focuses on the energy reconstruction using boosted decision tree (BDT) method and electron identification, and mainly uses the reconstructed-level (RECO) objects. The basic element of the RECO object is the reconstructed hit (RecHit), which is energy deposit in the sensor with pedestal and common-mode (CM) noises subtracted. The definitions of the pedestal and CM noises can be found in Ref. [131] and will not be described in detail in this report. Although I participated in the October beamtest, the used datasets for the studies were taken (or simulated) from beamtest in June, where 28 layers of modules were tested.

<sup>1</sup>The expected mean number of interaction per bunch crossing (pileup) in HL-LHC is approximately 140, more than 3 times of that in 2018 [129].

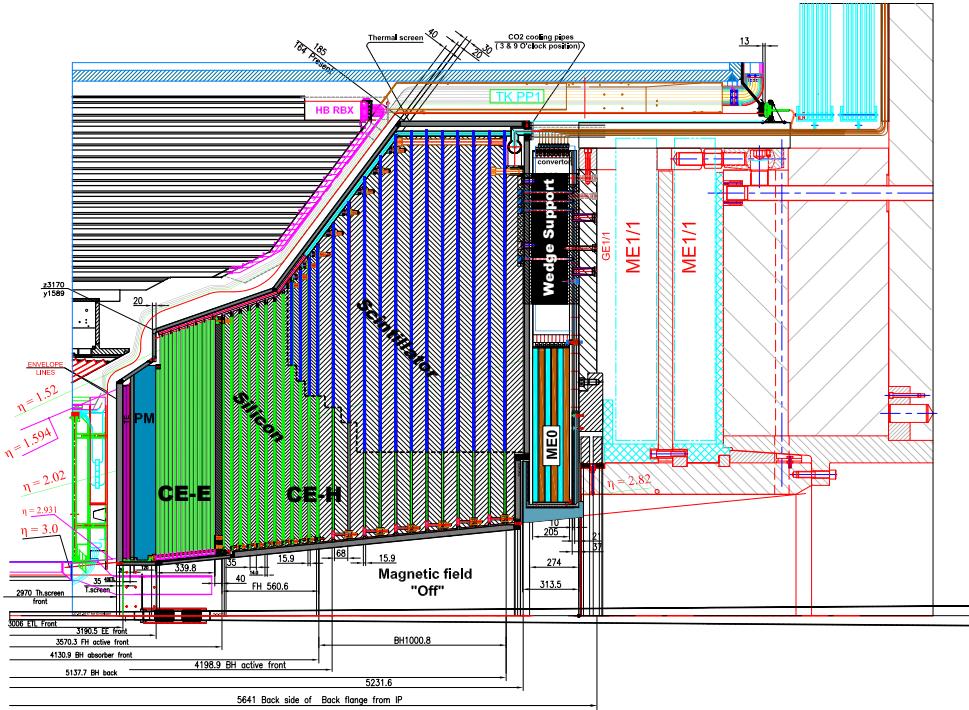


Figure C-2: Longitudinal cross section of the upper half of one endcap calorimeter [128].

## C.1 BDT method for energy reconstruction

The study aims at comparing various methods for energy reconstruction. My study focuses on the electron. The strategy is to use the simulation samples for regression, and the final goal is to apply the training result on beamtest data. In previous beamtests, the energy of the shower was calculated as the sum of the energy deposits in the active silicon sensors and in the passive absorbers over all tested layers, where the energy deposits in the absorbers were estimated using the stopping power<sup>2</sup> of the absorber materials which can be obtained from PDG [132] and simulation. In the following text, I will refer to this method as dEdx method. The full description of this method can be found in Ref. [131].

This study uses the XGBRegressor in XGBoost library [133], where the Gradient Boosting algorithm is implemented. The hyper-parameters adopted are listed in

---

<sup>2</sup>The energy losses per unit length ( $dE/dx$ ) of certain particle in the given material.

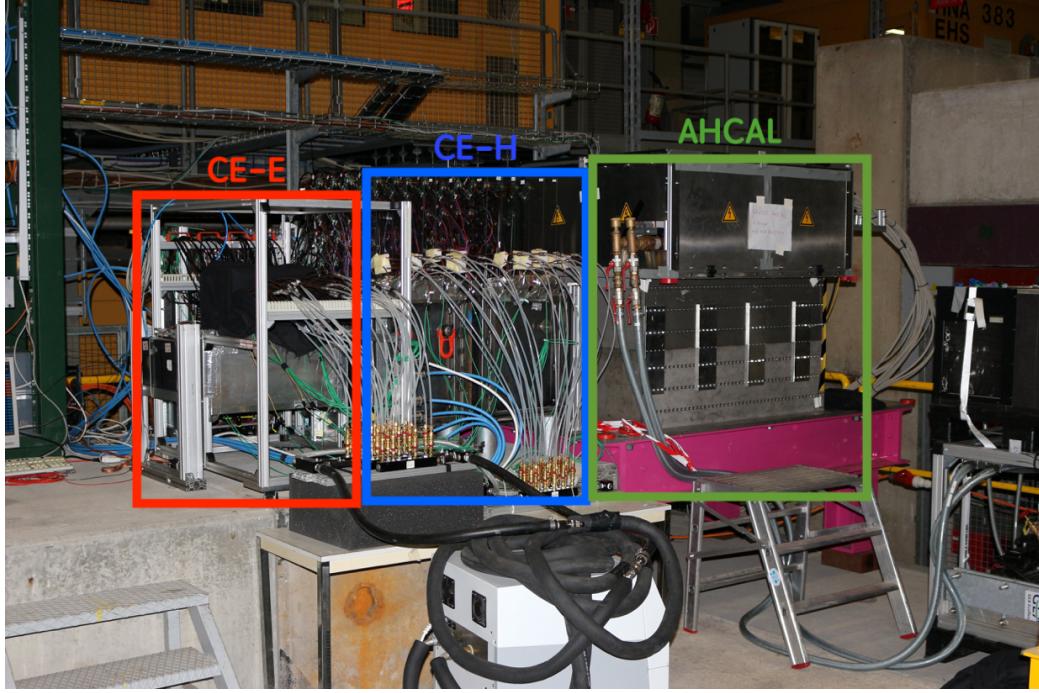


Figure C-3: Beamtest setup.

Table. C.1

hyperparameter	Parameters value
n_estimators	500
learning_rate	0.08
gamma	0
subsample	0.75
colsample_bytree	1
max_depth	7

Table C.1: The hyper-parameters adopted in the XGBRegressor for the regression.

The definitions of variables used to construct training features are listed in Table. C.2

Each layer has individual value for EAll, sum1, sum7, and sum19. From these basic variables, one can further construct simple lateral shower shape variables, such as sum1/sum7 (referred to as E1/E7), sum7/sum19 (E7/E19), and sum1/sum19 (E1/E19).

In the very first test [134], only the variables EAll and EAll/Etot were used as

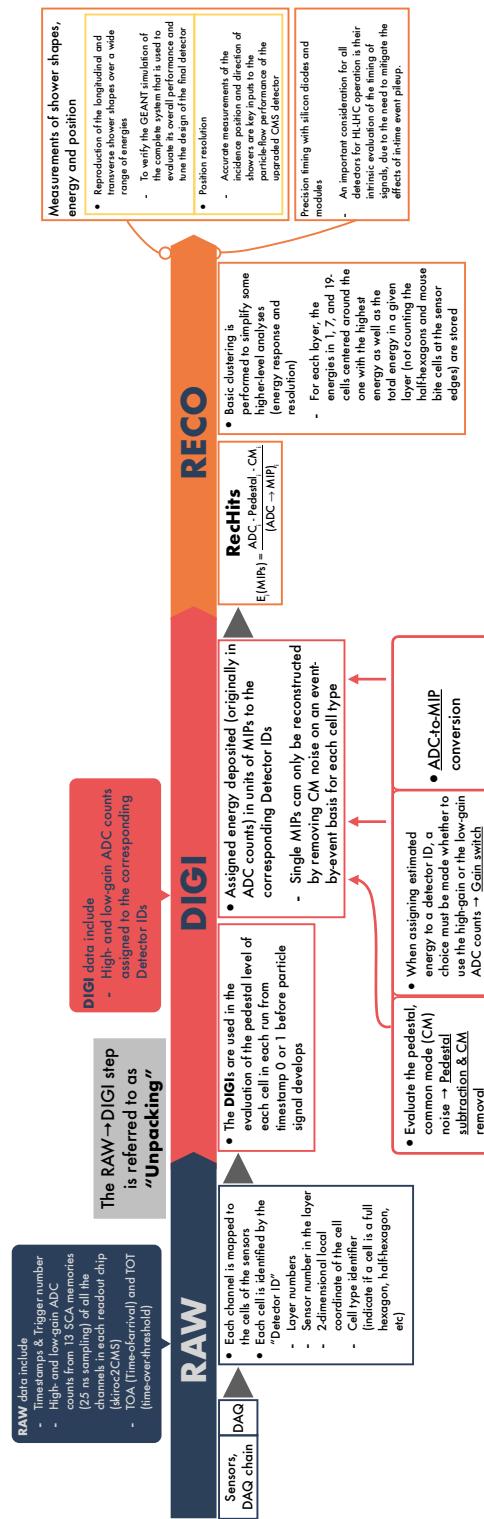


Figure C-4: The workflow of the data analysis framework, data preparation and reconstruction.

Variable	Definition
EAll	Energy deposits in each layer
Etot	Energy deposits in all layers
sum1	Energy deposit of the most energetic cell in each layer
sum7	Energy deposits in 7 cells centered at the most energetic cell in each layer
sum19	Energy deposits in 19 cells centered at the most energetic cell in each layer

Table C.2: The variables used to construct training features for the regression.

training features (for total of 56 variables) with the beam energy as target. Two sets of training were studied, one with the trained dataset having a uniform energy profile, and the other one with certain energy range ( $\pm 10$  GeV) for each energy point. Fig. [C-5] shows the results of the relative resolution as a function of predicted beam energy from regression (left) and the energy response as a function of predicted beam energy (right). The relative resolution is defined as the width of the reconstructed energy distribution divided by the predicted beam energy, while the energy response is defined as the ratio between the mean value of the reconstructed energy distribution and the true beam energy. Examples of the reconstructed energy distributions are shown in Fig. [C-6]. The observations from Fig. [C-5] and [C-6] are

- On the left plot, one clearly sees that results of relative resolution obtained from both sets of training are poorer than that from dEdx method, and the differences are larger at low energy points than at high energy points.
- The differences can be reduced with the training with energy range, indicating that at low energy points the predictions from regression are less precise, and the precision can be improved with the training with energy window.
- Although BDT regression gives worse resolutions than dEdx method, the scale of the reconstructed energy is more precise and gives more linear response, as can be seen on right plot.
- From the reconstructed energy distributions, one can see that the distributions are non-Gaussian and there are low energy outliers. By looking at the

event displays, as shown in Fig. C-7, the showers from events with low predicted energy (bottom plot, where the predicted energy is 15.8 GeV) do not develop as deep as those from normal predicted energy events (top plot, where the predicted energy is 20.1 GeV). Apart from this observation, there are few events where the electron hits the edge of the hexagon. However, rejecting events where the hit with the maximum energy deposit in the first layer is outside the  $2\text{ cm} \times 2\text{ cm}$  area around the hexagon center has marginal impact on the reconstructed energy distribution.

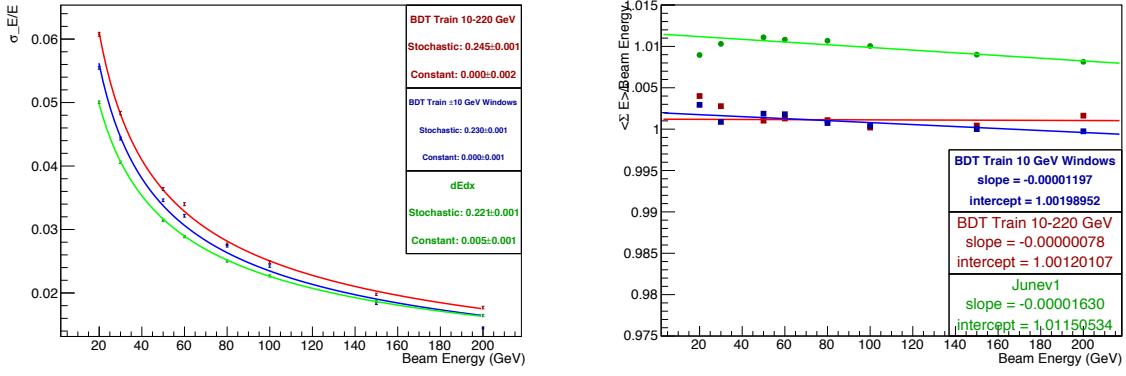


Figure C-5: Results of the relative resolution as a function of predicted beam energy from regression (left) and the energy response as a function of predicted beam energy (right).

New sets of regression are tested with (1) adding in other variables, such as lateral shower shape variables, as training features, and (2) dynamic energy window (i.e., narrower range for low energy points). Two sets of dynamic energy window are used, and are summarized in Table C.3. The feature importances of the regression for all energy points are shown in Fig. C-8. Some interesting observations are summarized in the following list.

- The pattern of the feature importance for each energy point are different from each other.
- EAll is the most important one among the five types of variables.

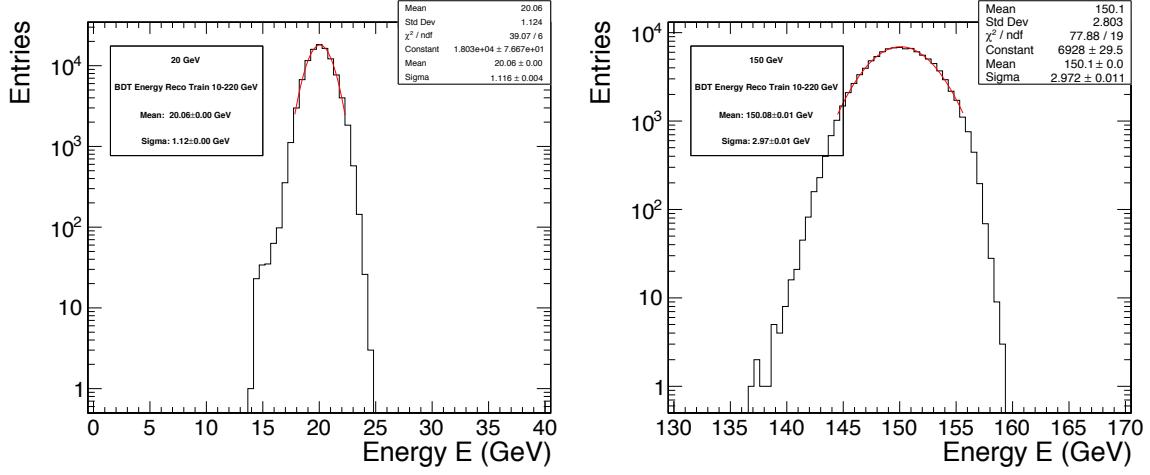


Figure C-6: Reconstructed energy distributions for 20 GeV electrons (left) and 150 GeV electrons (right) predicted from regression.

- E7/E19 is the most important lateral shower shape variable, compared to other two.
- For the low energy points, the contributions of the features from variables in last few layers are marginal.
- For the high energy points, the importances of EAll and EAll/Etot increase between layer 7 and 12, while the importances of E1/E19 and E1/E19 decrease between layer 4 and 10 and then increase onward.

The full correlation matrix of the training features is shown in Fig. C-9.

The latest results are shown in Fig. C-10, where resolutions can be better than those from dEdx method when beam energy is greater than 50 GeV, yet for 20 and 30 GeV energy points the resolutions and energy responses are still worse than from dEdx method. The possible improvement is to add the shower depth information as the training features and see if the precision of the regression can be improved.

Since this regression result will be applied on the beamtest data eventually, it is important to ensure the agreement between data and simulation, for which one need to select electron samples in beamtest data as pure as possible.

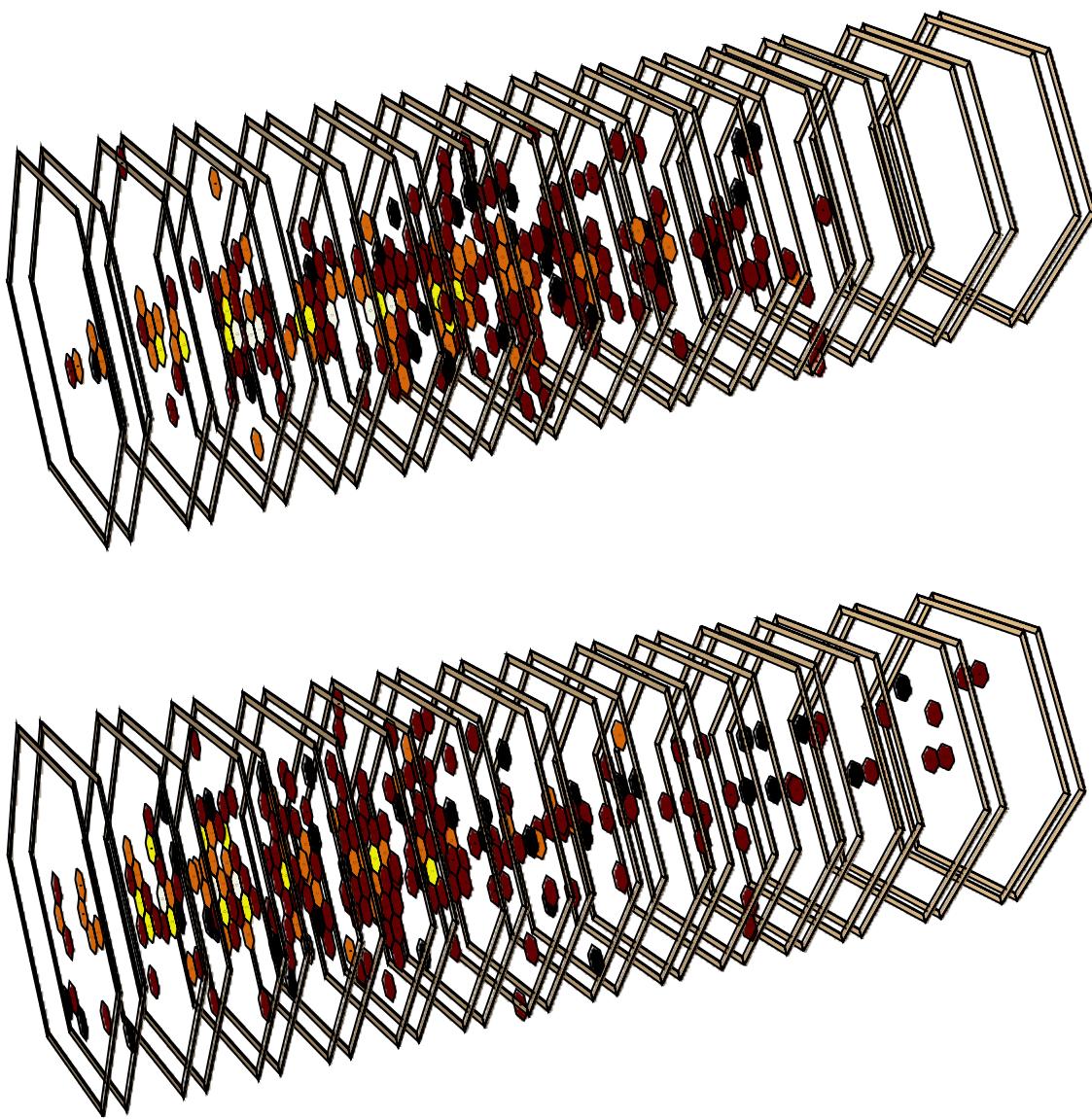


Figure C-7: Examples of event display of the electron shower. The predicted energy for the shower in top plot is 20.1 GeV, while that in bottom plot is 15.8 GeV.

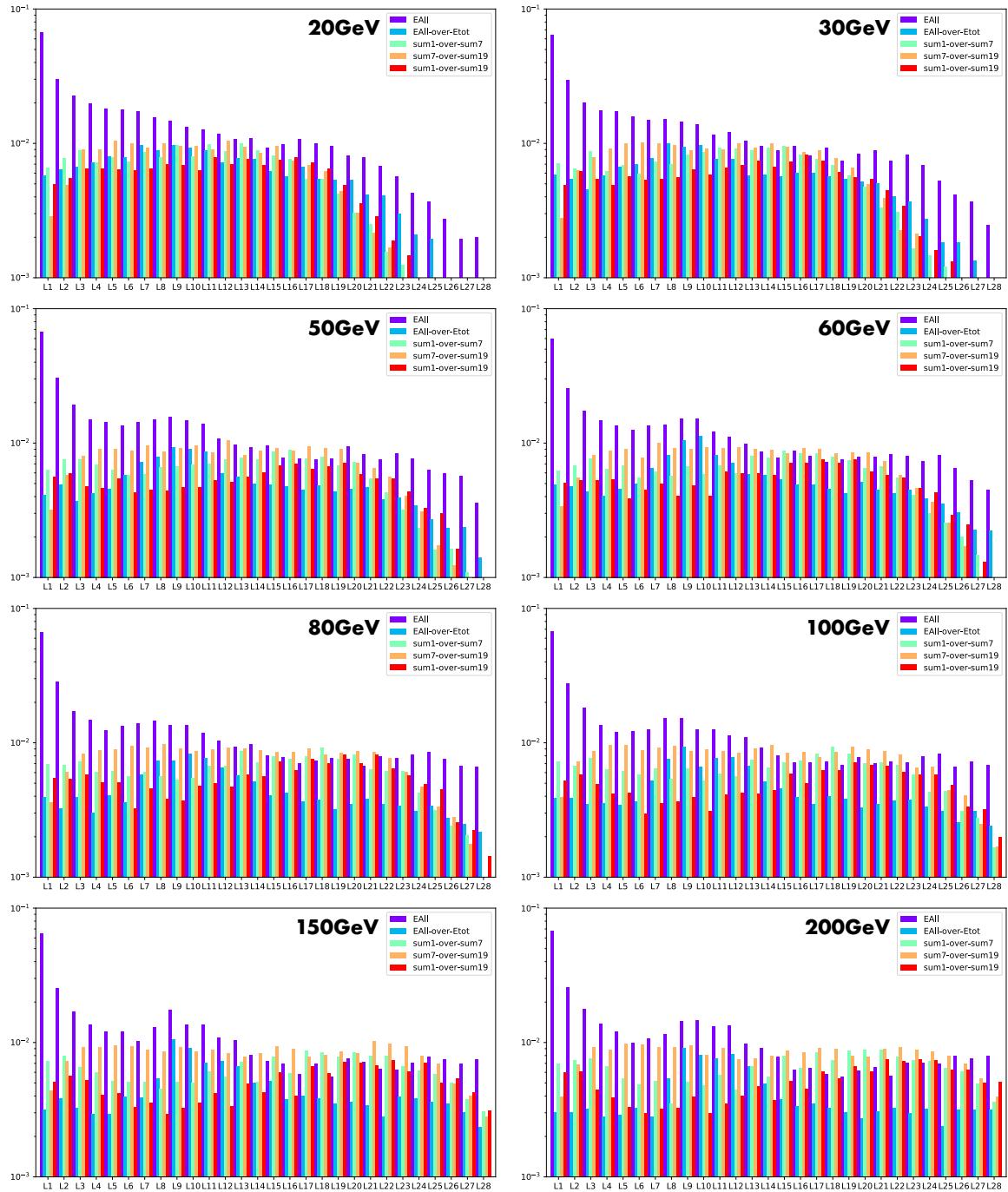


Figure C-8: Feature importances of the regression for all energy points.

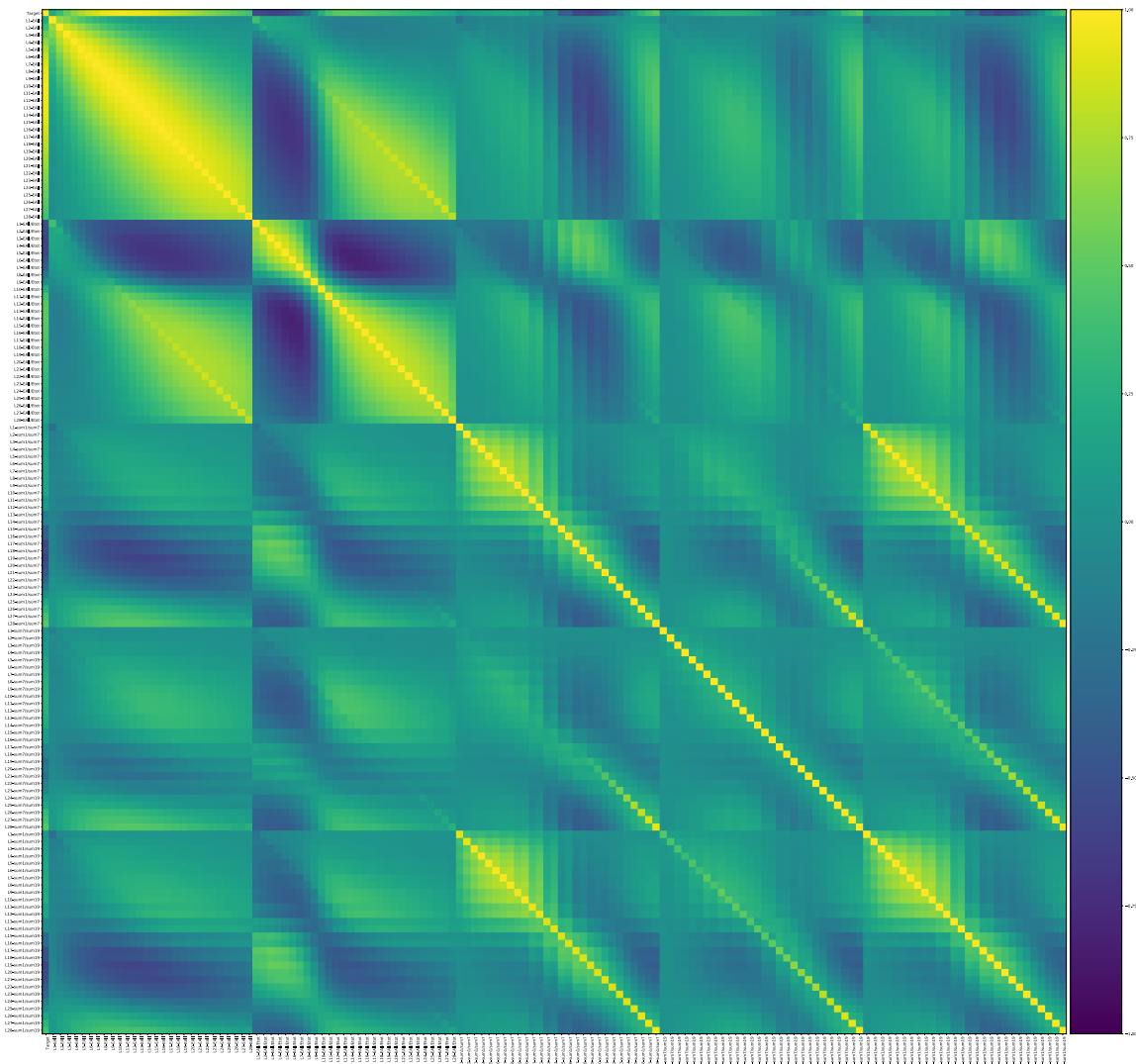


Figure C-9: The correlation matrix of the training features.

Set	(20, 30, 50, 60, 80, 100, 150, 200) GeV
1	( $\pm 8$ GeV, $\pm 10$ GeV, $\pm 10$ GeV)
2	( $\pm 6$ GeV, $\pm 6$ GeV, $\pm 6$ GeV, $\pm 6$ GeV, $\pm 8$ GeV, $\pm 10$ GeV, $\pm 10$ GeV)

Table C.3: The sizes of dynamic energy window used in two different sets of regression.

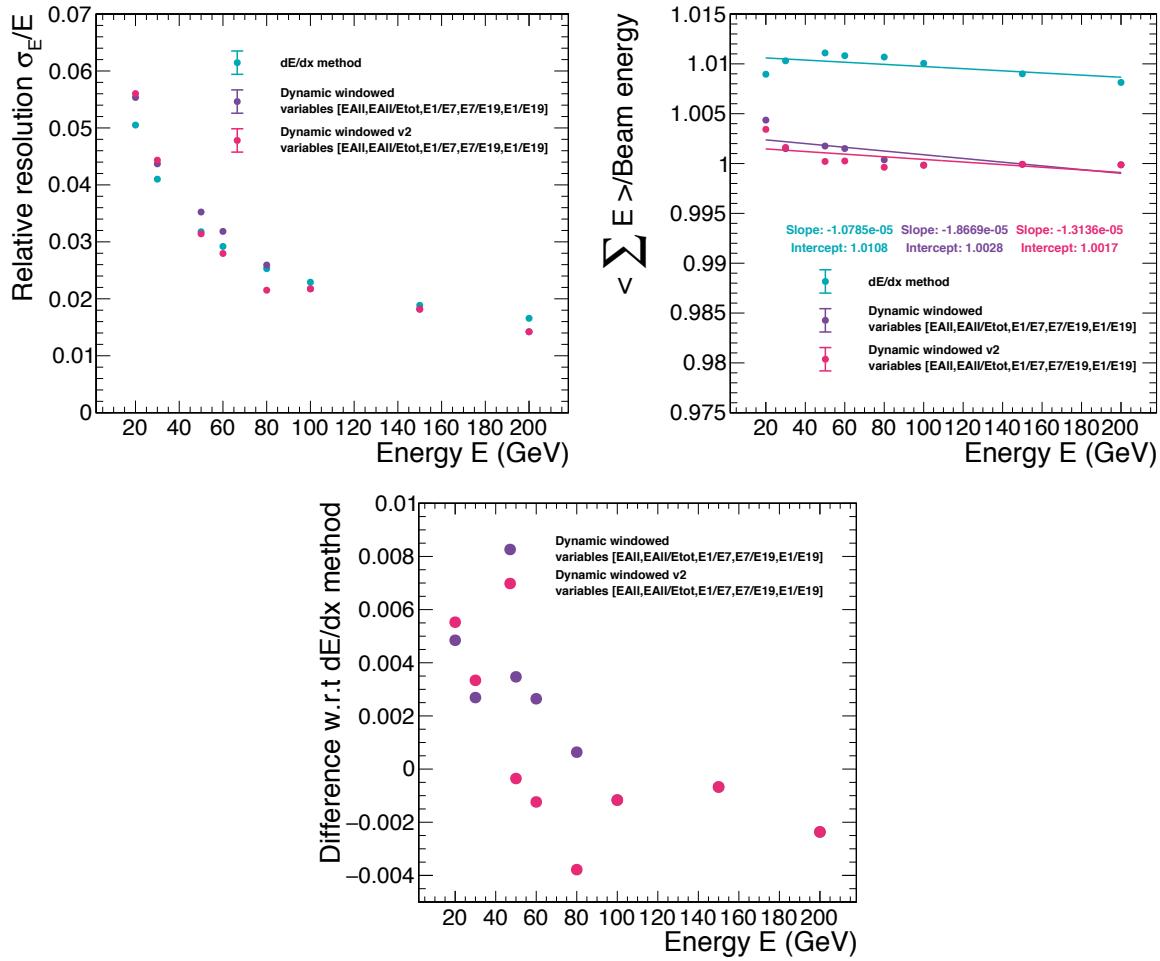


Figure C-10: Results of the relative resolution as a function of predicted beam energy from regression (top left), the energy response as a function of predicted beam energy (top right), and the differences in relative resolutions with respect to those from dEdx method, with lateral shower shape variables being used as training features and dynamic window.

In the next section, the first systematical way to separate electron and pion for the beamtest data will be introduced.

## C.2 Electron and pion separation

One of the issues in beamtest data is the pion contamination. From previous experience a suggested way to discriminate electron and pion is to look at the median value of the RecHit energy distribution and the energy-weighted longitudinal shower depth, defined as

$$\text{Longitudinal shower depth} = \frac{\sum_{i=1}^{28} (\text{EAll}_i \times X_{0,i})}{\sum_{i=1}^{28} \text{EAll}_i}. \quad (\text{C.1})$$

Fig. C-11 shows an example of the scatter plot of the median value of the RecHit energy distribution (will be abbreviated as mdn of RecHit in the following text) and the energy-weighted longitudinal shower depth (will be abbreviated as depthX0 in the following text) from 100 GeV beamtest data and simulation samples, and for better visualization, 2-dimension histograms for all the three samples are also shown. It is obvious that the pure electron events distribute differently from pure pion events, and in the beamtest data there is pion contamination.

The simplest way to reject pion events is to impose a straight line, and identify the events on the right hand side of it as electrons. However, it is difficult to choose a proper slope and intersection systematically. Alternatively, a "2-dimension window" cut is proposed. The basic ideas are:

- To construct a 2-D window that contains a fraction  $N\%$  of electron events,  $\sqrt{N}\%$  of events should be contained in each dimension (i.e., mdn of RecHit or depthX0).
- From the 1-dimension distributions of mdn of RecHit (depthX0) in certain range of depthX0 (mdn of RecHit), one can obtain  $\sqrt{N}\%$  quantile of the distribution. Examples of the 1-dimension distributions can be found in Fig. C-12, where the lines with different colors indicate the starting and ending points of certain quantile.
- By scanning over the range where there are enough statistics, one can obtain all the starting and ending points of the quantiles of the distributions over

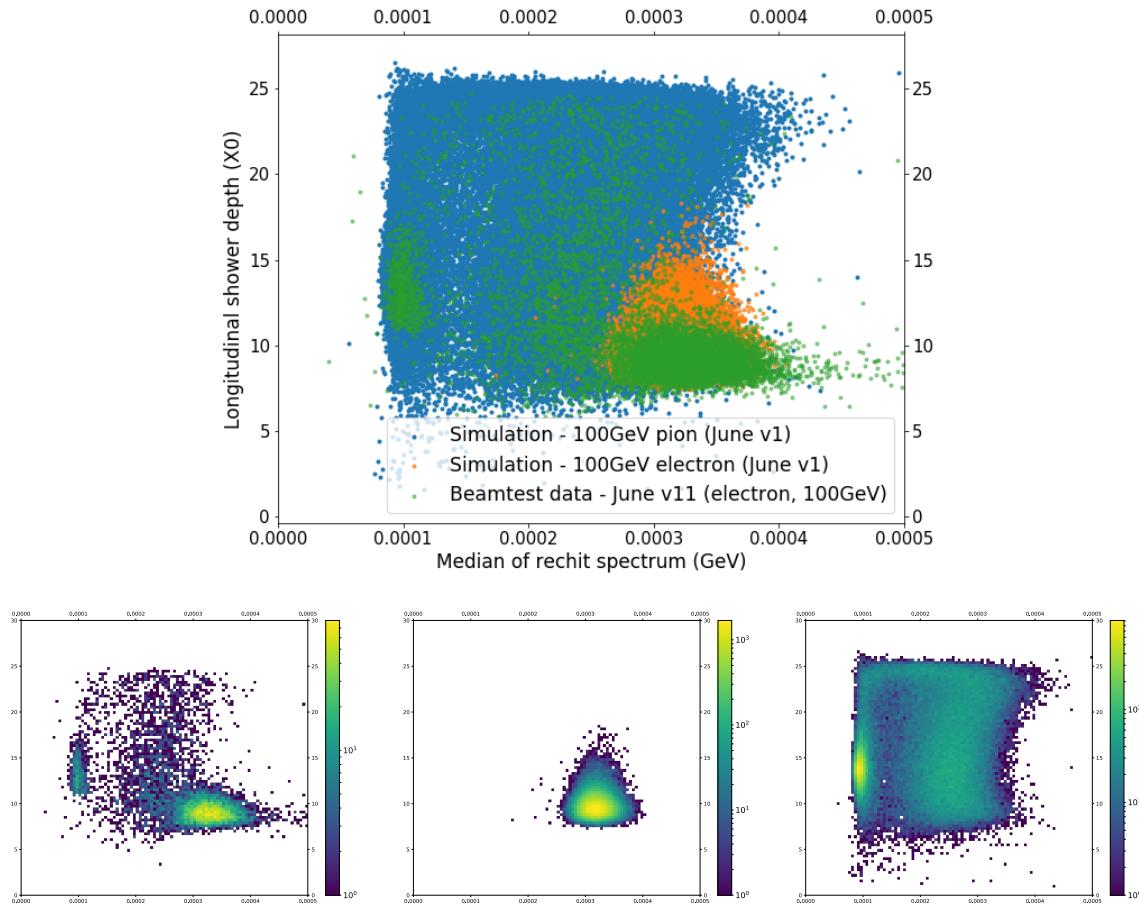


Figure C-11: The scatter plot of the median value of the RecHit energy distribution and the energy-weighted shower depth from 100 GeV beamtest data and simulation samples (top); 2-dimension histograms for beamtest data (bottom left), electron simulation (bottom middle), and pion simulation (bottom right).

the scanned range. The left plot of Fig. C-13 shows an example. Here one can already see the outline of the window.

- By fitting all the starting and ending points for given quantile with straight line (polynomial of order one), one obtains functions that roughly describe the relation between the mdn of RecHit and depthX0. The right plot of Fig. C-13 shows the fit results. The events inside the quadrangles are then identified as electron events, where different sizes represent the "2-dimension window" cut with different signal efficiencies (working points).

Fig. [C-14] shows the "2-dimension window" cut applied on beamtest data and pion simulation samples. The electron (signal) efficiency as a function of the pion (background) efficiency is shown in the left plot of Fig. [C-15], while the electron (signal) efficiencies as a function of background rejection power, defined as the reciprocal of background efficiency, is in the right plot. One can see that the actual signal efficiencies are close to the desired working points. The efficiencies in beamtest data, defined as the ratio of the number of events retained in the window cut over total number of events in the sample, are listed in Table. [C.4]

WP (%)	Efficiency in beamtest data (%)
68.3	19.4
80.0	25.0
90.0	30.9
95.0	35.7
99.0	42.5

Table C.4: The efficiencies in beamtest data with different working points.

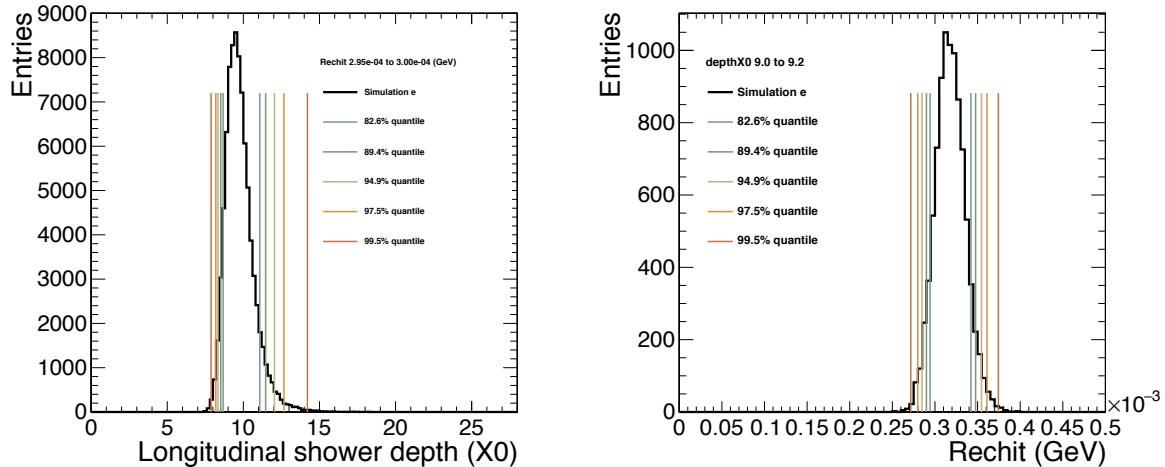


Figure C-12: 1-dimension distributions of mdn of RecHit (depthX0) in certain range of depthX0 (mdn of RecHit), with different quantiles (labeled in legend).

The constructed window cut do not result in bias in the reconstructed energy, which can be seen in Fig. [C-16] showing the distributions of total energy deposits in all layers with different working points, where all the distributions are normal-

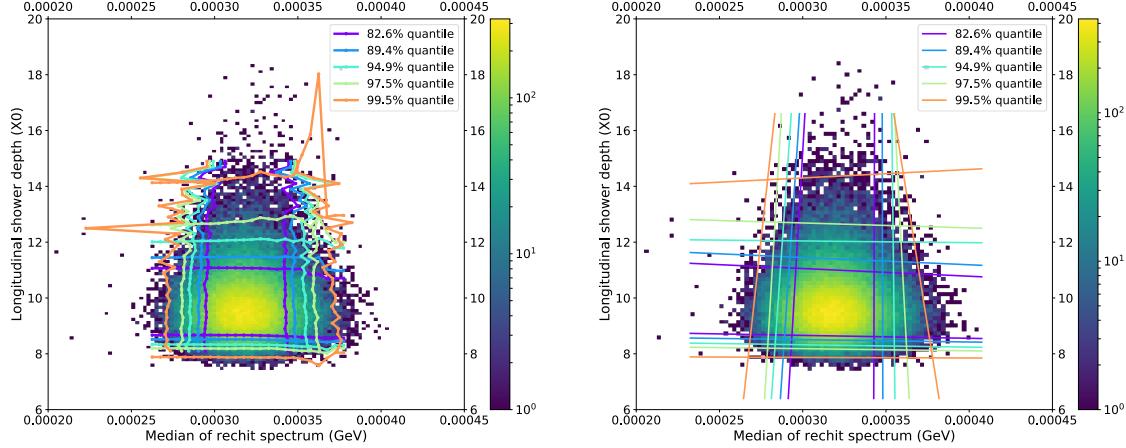


Figure C-13: The starting and ending points of the quantiles over the range where there are enough statistics.

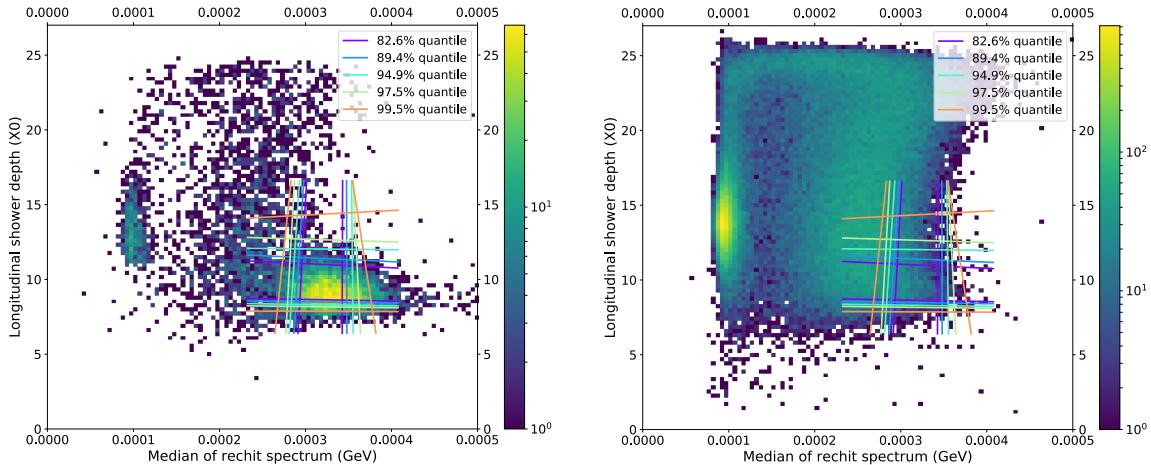


Figure C-14: The "2-dimension window" cut applied on beamtest data and pion simulation samples.

ized to unity. In the electron simulation sample, the tightest cut gives a 0.11% of difference in the median of the distribution with respect to that without window cut, while in beamtest data the difference is 3.0%.

Fig. C-17 shows the comparisons of EAll of layer 1, E7/E19 of layer 1, and sum of EAll/Etot over layer 1 to 10 between the beamtest data and simulation samples for both electron and pion events, the plots in left column are without applying the window cut and the plots on the right are with the window cuts of 68.3% working

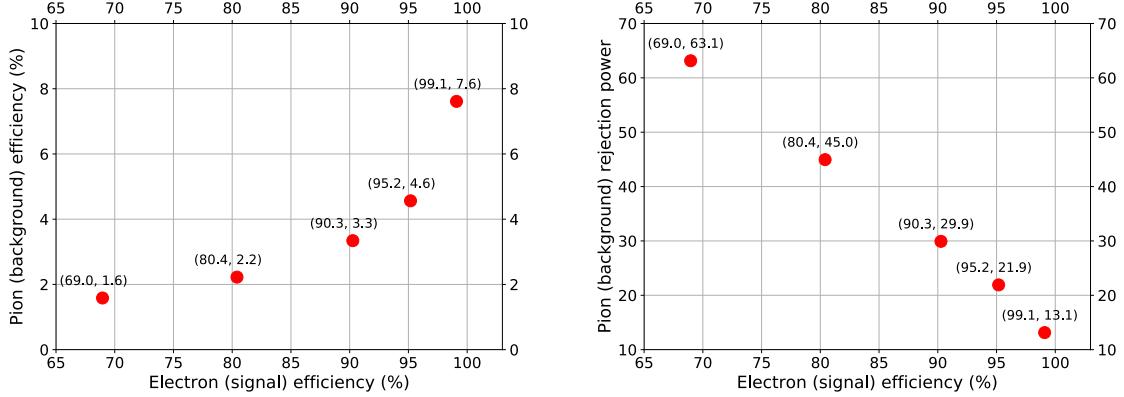


Figure C-15: The electron (signal) efficiency as a function of the pion (background) efficiency (left) and The electron (signal) efficiencies as a function of background rejection power, defined as the reciprocal of background efficiency (right).

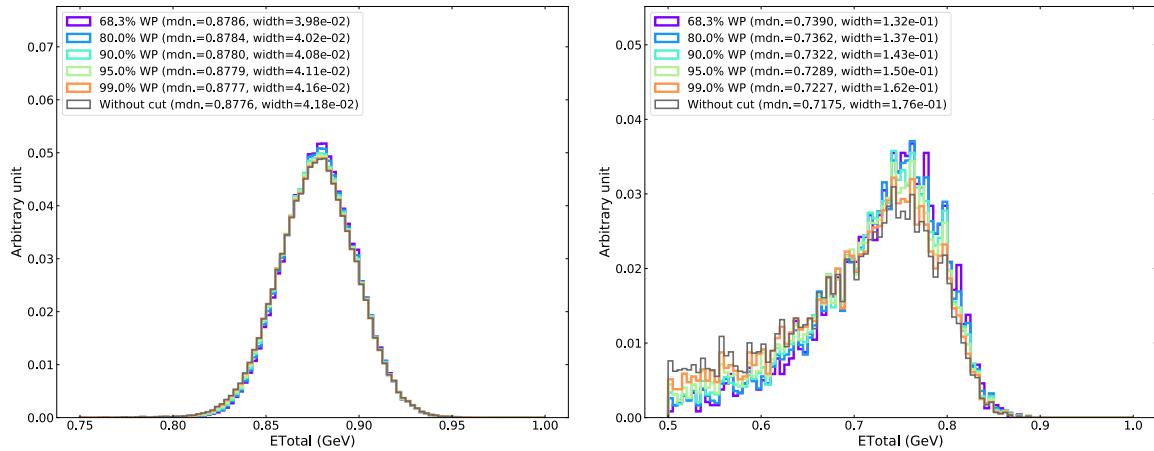


Figure C-16: The distributions of total energy deposits from simulation (left) and beamtest data (right) in all layers with different working points, where all the distributions are normalized to unity.

point. The agreement of the distributions between beamtest data and simulations of electron events improves after applying the tightest window cut. There are still residual pion events mimicking electrons after applying the tightest window cut, meaning that those pion events cannot be distinguished by this window cut. This can be seen from the event displays, Fig. C-18, of the pion simulation events that contained in the window cut. A more powerful identification is needed if one

wants purer electron samples from beamtest data. The machine learning technique is proposed and tested, and will be discussed in the next subsection.

### C.3 Machine learning technique for Electron and pion separation

In the first step, I tried using mdn of RecHit and depthX0 as training features for the classifier. This should give the baseline performance for the multivariate identification. Four commonly used classifiers are tested

- Linear support vector machine (linear SVM)
- XGBClassifier, based on gradient boosting algorithm
- Adaptive boosting classifier (AdaBoostClassifier)
- Random forest classifier (RandomForestClassifier)

The details of the classifiers and the corresponding algorithms will not be described in this report. The classifier outputs<sup>3</sup> from tested classifiers are shown in Fig. C-19, and the receiver operating characteristic (ROC) curves obtained from the classifier outputs are in Fig. C-20. The hyper-parameters used in the classifiers are the default setting of XGBClassifier. Table. C.5 summarizes the background (pion) efficiency with 99.0% of signal (electron) efficiency from each tested classifier, which quantifies the performance on electron and pion separation. Among the four classifiers, XGBClassifier gives the best performance on discriminating the electron and pion events, resulting an 2.2% of improvement with respect to the window cut. Therefore, XGBClassifier will be used as the classifier in the following study.

---

<sup>3</sup>The value of classifier output is defined as the probability of each event being predicted as electron.

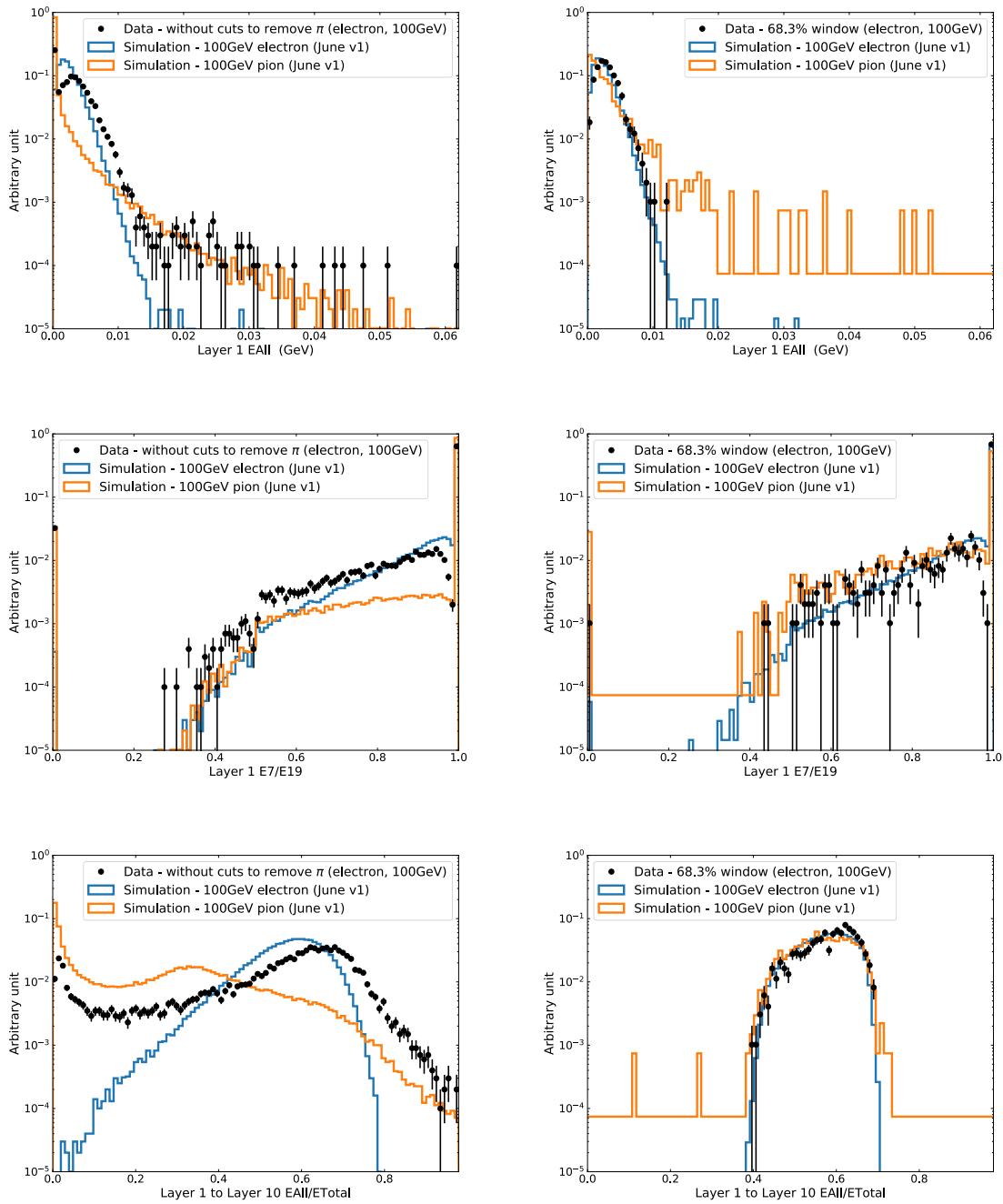


Figure C-17: The comparison between the beamtest data and simulation samples for both electron and pion events, the plots in left row are without applying the window cut and the plots on the right are with the window cuts of 68.3% working point.

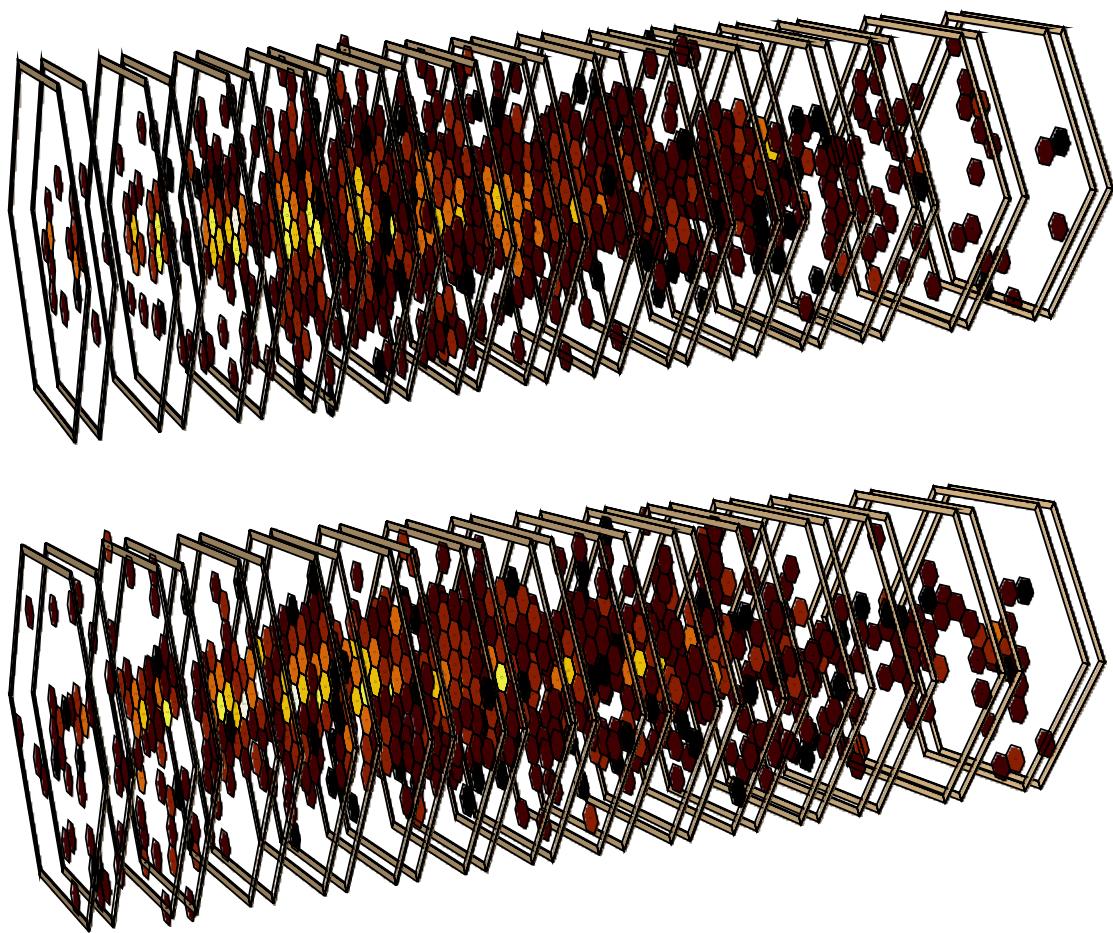
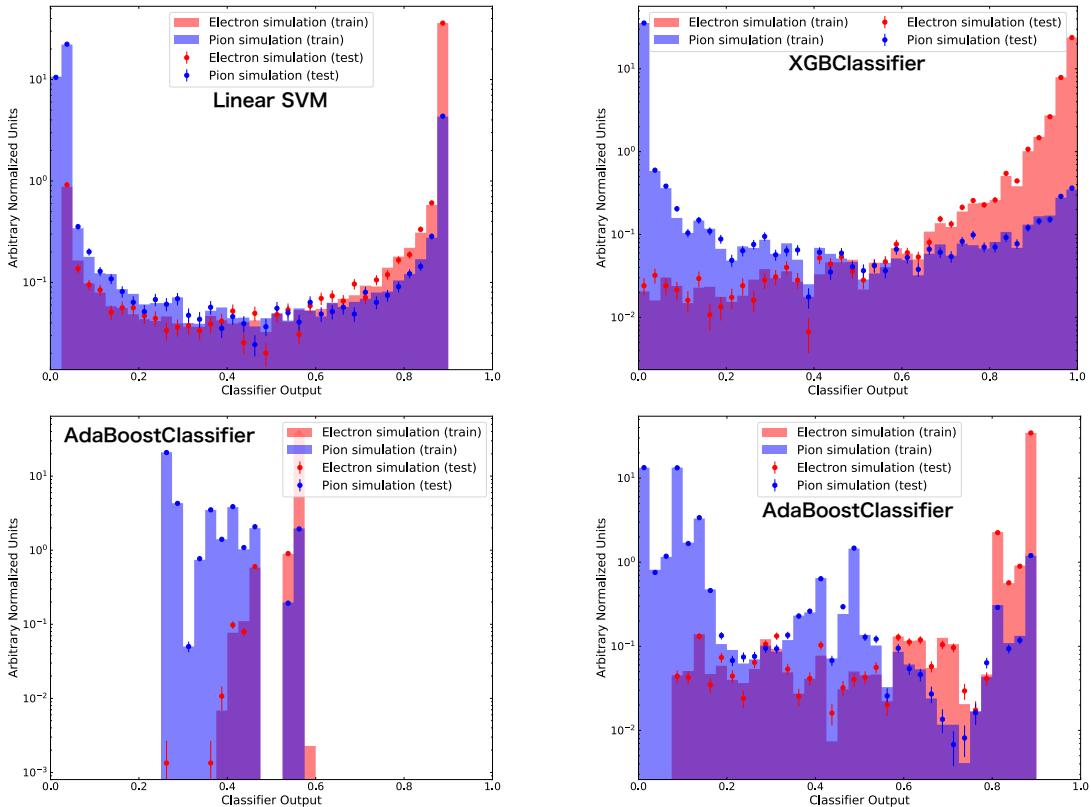


Figure C-18: Examples of event display of the pion events that contained in the window cut.



1

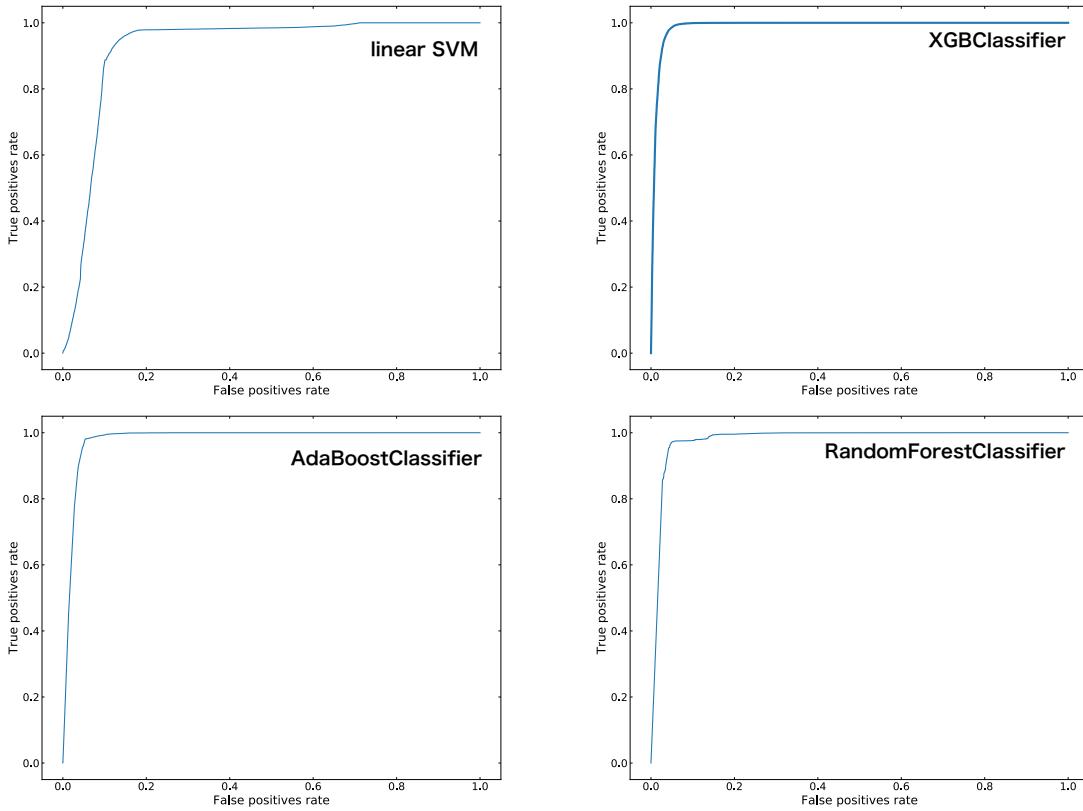
Figure C-19: Classifier outputs from linear SVM (top left), XGBClassifier (top right), AdaBoostClassifier (bottom left), RandomForestClassifier (bottom right).

Classifier	Background efficiency (with 99.0% of signal efficiency)
XGBClassifier	5.43
AdaBoostClassifier	8.40
RandomForestClassifier	14.2
linear SVM	64.9

Table C.5: The background (pion) efficiencies from tested classifiers with 99.0% of signal (electron) efficiency.

The next study is to add in more variables as training features. Different sets of training features are used, which are summarized as below.

1. Baseline (mdn of RecHit and depthX0) + Number of hits in all layers, referred to as nhits, or sum over all energy deposits in all layers



1

Figure C-20: ROC curves from linear SVM (top left), XGBClassifier (top right), AdaBoostClassifier (bottom left), RandomForestClassifier (bottom right).

2. 4 variables - Baseline + nhits + Etot

3. 7 variables - Baseline + nhits + Etot + 3 variables, where the 3 variables are

- The ratio of the energy sum in first two layers over Etot, referred to as L1\_L2\_EAll\_over\_ETotal in the following figures.
- The ratio of the energy sum in first three layers over Etot, referred to as L1\_L3\_EAll\_over\_ETotal in the following figures.
- The ratio of the energy sum in first ten layers over Etot, referred to as L1\_L4\_EAll\_over\_ETotal in the following figures.

4. 28 variables listed in Table. C.2

5. 35 variables - Baseline + nhits + Etot + 3 variables + 28 variables listed in Table. C.2

Fig. C-21 shows the respect feature importances from different trainings listed above. For the trainings with 7, 28, and 35 variables, only the results from the combination of variables that give the best performance are shown. Interestingly, Etot is always the most important variable in the trainings where it is used. The correlation matrices of 35 variables from electron and pion events are shown in Fig. C-22.

Fig. C-23 shows the background efficiencies with different sets of 28 variables listed in Table. C.2. From the left plot, the combination [Baseline + nhits + Etot + 3 variables + 28 E1 variables] so far gives the best performance among all tested combinations. From both plots, one can see that including lateral shower shape variables does not seem helpful on discriminating the electron and pion events.

Table. C.6 summarizes the performances from different sets of training features listed previously in the text. With the machine learning technique, the baseline identification gives 28.6% improvement with respect to the window cut. From the fact that adding only the Etot brings approximately 90% of improvement and the feature importances, one can conclude that Etot is a critical variables in the electron and pion discrimination. However, when adding 28 variables, the improvement seems marginal. The next step is to perform hyper-parameter optimization and see if the improvement is actually limited by the default setting of the hyper-parameter. Another issue is that there are huge number of variables that can be used in the training, and dumping all the variables into the training seems redundant. How to choose training features to give optimal performance should be dedicated.

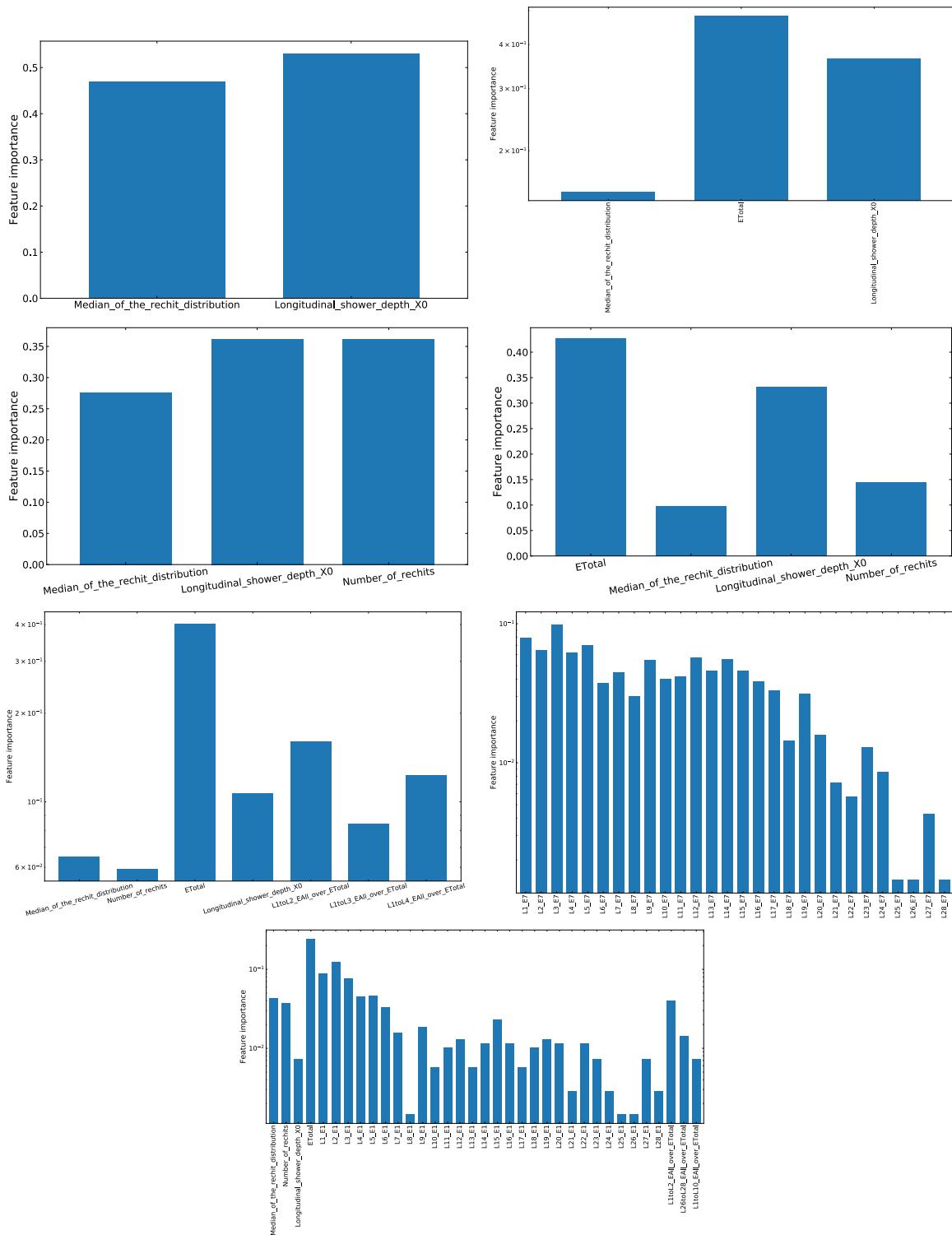


Figure C-21: The respect feature importances from different trainings, as listed previously in the text.

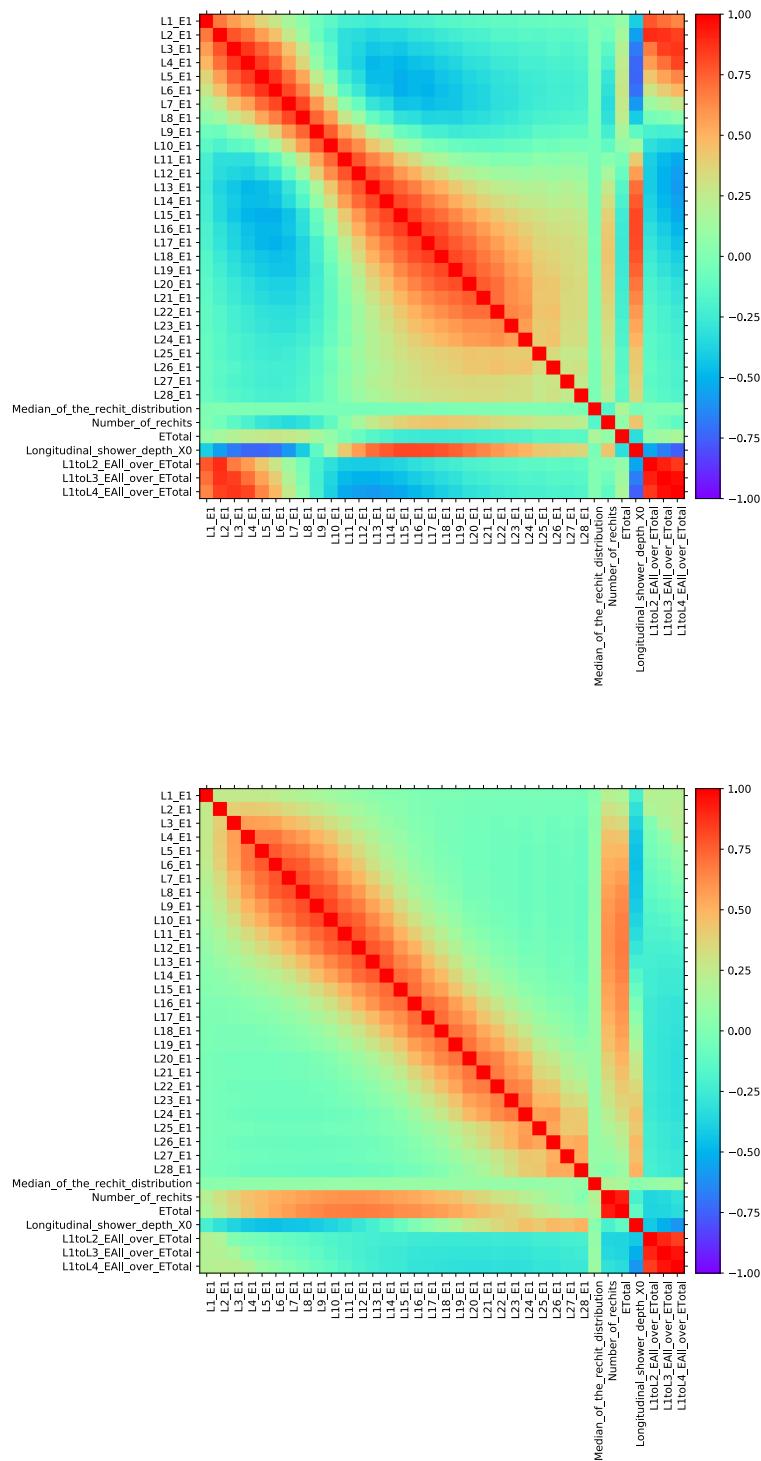


Figure C-22: The correlation matrix of the training features for 35 variables.

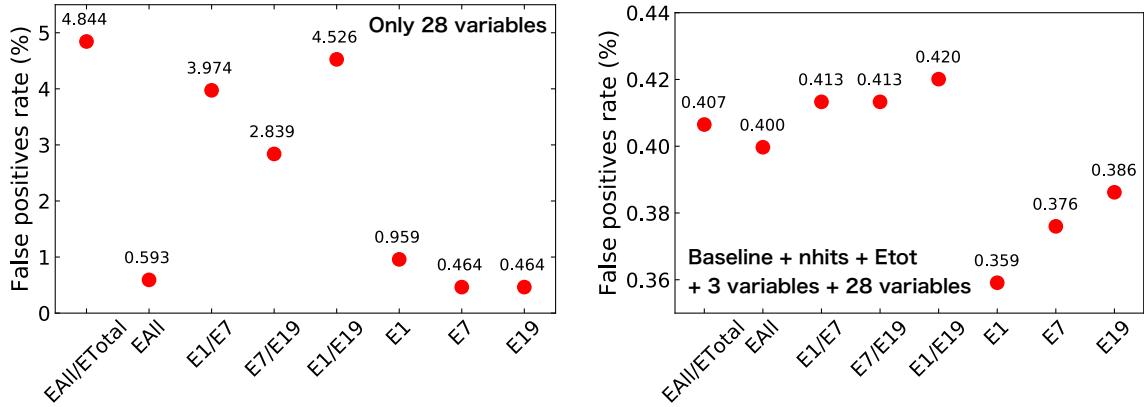


Figure C-23: The background efficiencies with different sets of 28 variables listed in Table C.2.

Features	Number of training features	Background efficiency (%) (at 99% of signal efficiency)	Improvement (%) w.r.t window cut	Improvement (%) w.r.t baseline
Window cut	2	7.60	—	—
Baseline	2	5.43	28.6	—
Baseline + nhits	3	2.81	63.0	48.3
Baseline + Etot	3	0.613	91.9	88.7
4 variables	4	0.562	92.6	89.7
7 variables	7	0.427	94.4	92.1
28 variables	28	0.464	93.9	91.5
35 variables	35	0.359	95.3	93.4

Table C.6: Summary of the performances from different sets of training features.

# Bibliography

- [1] Peter Skands. [Introduction to QCD](#). 2013.
- [2] The ALEPH, DELPHI, L3, OPAL Collaborations, the LEP Electroweak Working Group. [Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP](#). *Phys. Rept.*, 532:119, 2013.
- [3] Steven Weinberg. [A Model of Leptons](#). *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.
- [4] G. Arnison et al. [Experimental observation of isolated large transverse energy electrons with associated missing energy at  \$\sqrt{s}=540\$  GeV](#). *Phys. Lett. B*, 122(1):103 – 116, 1983.
- [5] G. Arnison et al. [Observation of muonic  \$Z\_0\$ -decay at the pp collider](#). *Phys. Lett. B*, 147(1):241 – 248, 1984.
- [6] Michał Szleper. [The Higgs boson and the physics of WW scattering before and after Higgs discovery](#). 2014.
- [7] F. Englert and R. Brout. [Broken Symmetry and the Mass of Gauge Vector Mesons](#). *Phys. Rev. Lett.*, 13:321–323, Aug 1964.
- [8] Peter W. Higgs. [Broken Symmetries and the Masses of Gauge Bosons](#). *Phys. Rev. Lett.*, 13:508–509, Oct 1964.
- [9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. [Global Conservation Laws and Massless Particles](#). *Phys. Rev. Lett.*, 13:585–587, Nov 1964.
- [10] Murray Gell-Mann, Pierre Ramond, and Richard Slansky. [Complex Spinors and Unified Theories](#). *Conf. Proc.*, C790927:315–321, 1979.
- [11] Mohapatra, Rabindra N. and Senjanović, Goran. [Neutrino Mass and Spontaneous Parity Nonconservation](#). *Phys. Rev. Lett.*, 44:912–915, Apr 1980.
- [12] LHC Higgs Cross Section Working Group. History: r294.
- [13] LHCTopWG Summary Plots. History: r52.

- [14] Georges Aad et al. Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at  $\sqrt{s} = 7$  and 8 TeV. *JHEP*, 08:045, 2016.
- [15] Morad Aaboud et al. Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector. 2018.
- [16] Albert M Sirunyan et al. Observation of  $t\bar{t}H$  production. *Phys. Rev. Lett.*, 120(23):231801, 2018.
- [17] Search for the associated production of a Higgs boson and a single top quark in pp collisions at  $\sqrt{s} = 13$  TeV. Technical Report CMS-PAS-HIG-18-009, CERN, Geneva, 2018.
- [18] Geoffrey T. Bodwin, Frank Petriello, Stoyan Stoynev, and Mayda Velasco. Higgs boson decays to quarkonia and the  $H\bar{c}c$  coupling. *Phys. Rev. D*, 88:053003, 2013.
- [19] T. Aaltonen, A. Buzatu, B. Kilminster, Y. Nagai, and W. Yao. Improved  $b$ -jet energy correction for  $H \rightarrow b\bar{b}$  searches at CDF. 2011.
- [20] Vardan Khachatryan et al. Search for the associated production of a Higgs boson with a single top quark in proton-proton collisions at  $\sqrt{s} = 8$  TeV. *JHEP*, 06:177, 2016.
- [21] Georges Aad et al. Search for the  $b\bar{b}$  decay of the Standard Model Higgs boson in associated  $(W/Z)H$  production with the ATLAS detector. *JHEP*, 01:069, 2015.
- [22] Albert M Sirunyan et al. Observation of Higgs boson decay to bottom quarks. *Submitted to: Phys. Rev. Lett.*, 2018.
- [23] Morad Aaboud et al. Observation of  $H \rightarrow b\bar{b}$  decays and  $VH$  production with the ATLAS detector. 2018.
- [24] Georges Aad et al. Test of CP Invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector. *Eur. Phys. J.*, C76(12):658, 2016.
- [25] Stefan Berge, Werner Bernreuther, and Sebastian Kirchner. Prospects of constraining the Higgs boson's CP nature in the tau decay channel at the LHC. *Phys. Rev. D*, 92:096012, 2015.
- [26] Cross-section measurements of the Higgs boson decaying to a pair of tau leptons in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. Technical Report ATLAS-CONF-2018-021, CERN, Geneva, Jun 2018.

- [27] Albert M Sirunyan et al. Observation of the Higgs boson decay to a pair of  $\tau$  leptons with the CMS detector. *Phys. Lett. B*, 779:283–316, 2018.
- [28] Measurement of gluon fusion and vector boson fusion Higgs boson production cross-sections in the  $H \rightarrow WW^* \rightarrow e\nu\mu\nu$  decay channel in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. Technical Report ATLAS-CONF-2018-004, CERN, Geneva, Mar 2018.
- [29] Albert M Sirunyan et al. Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at  $\sqrt{s} = 13$  TeV. Submitted to: *Phys. Lett. B*, 2018.
- [30] Serguei Chatrchyan et al. Measurement of the properties of a Higgs boson in the four-lepton final state. *Phys. Rev. D*, 89(9):092007, 2014.
- [31] Serguei Chatrchyan et al. Study of the mass and spin-parity of the Higgs boson candidate via its decays to Z Boson pairs. *Phys. Rev. Lett.*, 110(8):081803, 2013.
- [32] Vardan Khachatryan et al. Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV. *Phys. Rev. D*, 92(1):012004, 2015.
- [33] Georges Aad et al. Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Phys. Rev. D*, 91(1):012006, 2015.
- [34] Georges Aad et al. Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector. *Eur. Phys. J.*, C75(10):476, 2015. [Erratum: *Eur. Phys. J.* C76,no.3,152(2016)].
- [35] Vardan Khachatryan et al. Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs. *Phys. Lett. B*, B736:64–85, 2014.
- [36] Vardan Khachatryan et al. Limits on the Higgs boson lifetime and width from its decay to four charged leptons. *Phys. Rev. D*, 92(7):072010, 2015.
- [37] Georges Aad et al. Constraints on the off-shell Higgs boson signal strength in the high-mass ZZ and WW final states with the ATLAS detector. *Eur. Phys. J.*, C75(7):335, 2015.
- [38] Measurements of properties of the Higgs boson in the four-lepton final state at  $\sqrt{s} = 13$  TeV. Technical Report CMS-PAS-HIG-18-001, CERN, Geneva, 2018.
- [39] Measurements of the Higgs boson production, fiducial and differential cross sections in the  $4\ell$  decay channel at  $\sqrt{s} = 13$  TeV with the ATLAS detector. Technical Report ATLAS-CONF-2018-018, CERN, Geneva, Jun 2018.

- [40] Albert M Sirunyan et al. Measurements of properties of the Higgs boson decaying into the four-lepton final state in  $pp$  collisions at  $\sqrt{s} = 13$  TeV. *JHEP*, 11:047, 2017.
- [41] Vardan Khachatryan et al. Observation of the diphoton decay of the Higgs boson and measurement of its properties. *Eur. Phys. J.*, C74(10):3076, 2014.
- [42] Georges Aad et al. 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. *Phys. Rev. D*, 90(11):112015, 2014.
- [43] Albert M Sirunyan et al. Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at  $\sqrt{s} = 13$  TeV. 2018.
- [44] Albert M Sirunyan et al. Measurement of inclusive and differential Higgs boson production cross sections in the diphoton decay channel in proton-proton collisions at  $\sqrt{s} = 13$  TeV. 2018.
- [45] Morad Aaboud et al. Measurements of Higgs boson properties in the diphoton decay channel with  $36\text{ fb}^{-1}$  of  $pp$  collision data at  $\sqrt{s} = 13$  TeV with the ATLAS detector. 2018.
- [46] Albert M Sirunyan et al. Search for the decay of a Higgs boson in the  $\ell\ell\gamma$  channel in proton-proton collisions at  $\sqrt{s} = 13$  TeV. Submitted to: *JHEP*, 2018.
- [47] M. Aaboud et al. Searches for the  $Z\gamma$  decay mode of the Higgs boson and for new high-mass resonances in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. *JHEP*, 10:112, 2017.
- [48] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016.
- [49] Albert M Sirunyan et al. Search for the Higgs boson decaying to two muons in proton-proton collisions at  $\sqrt{s} = 13$  TeV. 2018.
- [50] A search for the rare decay of the Standard Model Higgs boson to dimuons in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS Detector. Technical Report ATLAS-CONF-2018-026, CERN, Geneva, Jul 2018.
- [51] Gilad Perez, Yotam Soreq, Emmanuel Stamou, and Kohsaku Tobioka. Constraining the charm Yukawa and Higgs-quark coupling universality. *Phys. Rev. D*, 92(3):033016, 2015.
- [52] Gilad Perez, Yotam Soreq, Emmanuel Stamou, and Kohsaku Tobioka. Prospects for measuring the Higgs boson coupling to light quarks. *Phys. Rev. D*, 93(1):013001, 2016.

- [53] M. Aaboud et al. [Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment](#). *Phys. Rev. Lett.*, 120(21):211802, 2018.
- [54] Georges Aad et al. [Combined measurement of the Higgs boson mass in  \$pp\$  collisions at  \$\sqrt{s} = 7\$  and 8 TeV with the ATLAS and CMS experiments](#). *Phys. Rev. Lett.*, 114:191803, 2015.
- [55] Georges Aad et al. [Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC](#). *Phys. Lett. B*, 716:1–29, 2012.
- [56] Serguei Chatrchyan et al. [Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC](#). *Phys. Lett. B*, 716:30–61, 2012.
- [57] Serguei Chatrchyan et al. [Observation of a new boson with mass near 125 GeV in  \$pp\$  collisions at  \$\sqrt{s} = 7\$  and 8 TeV](#). *JHEP*, 06:081, 2013.
- [58] Georges Aad et al. [Measurements of the Higgs boson production and decay rates and coupling strengths using  \$pp\$  collision data at  \$\sqrt{s} = 7\$  and 8 TeV in the ATLAS experiment](#). *Eur. Phys. J.*, C76(1):6, 2016.
- [59] Vardan Khachatryan et al. [Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV](#). *Eur. Phys. J.*, C75(5):212, 2015.
- [60] Georges Aad et al. [Evidence for the spin-0 nature of the Higgs boson using ATLAS data](#). *Phys. Lett. B*, 726:120–144, 2013.
- [61] [Combined measurements of the Higgs boson’s couplings at  \$\sqrt{s} = 13\$  TeV](#). Technical Report CMS-PAS-HIG-17-031, CERN, Geneva, 2018.
- [62] J R Andersen et al. [Handbook of LHC Higgs Cross Sections: 3. Higgs Properties](#). 2013.
- [63] John Ellis and Tevong You. [Global analysis of the Higgs candidate with mass 125 GeV](#). *JHEP*, 09:123, 2012.
- [64] John Ellis and Tevong You. [Updated Global Analysis of Higgs Couplings](#). *JHEP*, 06:103, 2013.
- [65] C et al. Patrignani. [Review of Particle Physics, 2016-2017](#). *Chin. Phys. C*, 40(10):100001. 1808 p, 2016.
- [66] Fady Bishara, Ulrich Haisch, Pier Francesco Monni, and Emanuele Re. [Constraining Light-Quark Yukawa Couplings from Higgs Distributions](#). *Phys. Rev. Lett.*, 118(12):121801, 2017.

- [67] Combined measurement and interpretation of differential Higgs boson production cross sections at  $\sqrt{s}=13$  TeV. Technical Report CMS-PAS-HIG-17-028, CERN, Geneva, 2018.
- [68] Cédric Delaunay, Tobias Golling, Gilad Perez, and Yotam Soreq. Enhanced Higgs boson coupling to charm pairs. *Phys. Rev. D*, 89(3):033014, 2014.
- [69] W. Buchmuller and D. Wyler. Effective Lagrangian analysis of new interactions and flavor conservation. *Nucl. Phys. B*, 268:621, 1986.
- [70] Steven Weinberg. Effective gauge theories. *Phys. Lett. B*, 91:51, 1980.
- [71] Roberto Contino, Margherita Ghezzi, Christophe Grojean, Margarete Muhlleitner, and Michael Spira. Effective Lagrangian for a light Higgs-like scalar. *JHEP*, 07:035, 2013.
- [72] Michael Trott and Mark B. Wise. On theories of enhanced CP violation in  $B_{s,d}$  meson mixing. *JHEP*, 11:157, 2010.
- [73] Martin Jung, Antonio Pich, and Paula Tuzon. Charged-Higgs phenomenology in the Aligned two-Higgs-doublet model. *JHEP*, 11:003, 2010.
- [74] G. F. Giudice, C. Grojean, A. Pomarol, and R. Rattazzi. The strongly-interacting light Higgs. *JHEP*, 06:045, 2007.
- [75] Cédric Delaunay, Thomas Flacke, J. Gonzalez-Fraile, Seung J. Lee, Giuliano Panico, and Gilad Perez. Light non-degenerate composite partners at the LHC. *JHEP*, 02:055, 2014.
- [76] B. Guberina and J.H. Kühn and R.D. Peccei and R. Rückl. Rare decays of the  $Z^0$ . *Nuclear Physics B*, 174:317, 1980.
- [77] Ting-Chung Huang and Frank Petriello. Rare exclusive decays of the  $Z$ -boson revisited. *Phys. Rev. D*, 92:014007, 2015.
- [78] Yuval Grossman, Matthias König, and Matthias Neubert. Exclusive radiative decays of  $W$  and  $Z$  bosons in QCD factorization. *JHEP*, 04:101, 2015.
- [79] Geoffrey T. Bodwin, Hee Sok Chung, June-Haak Ee, and Jungil Lee. Addendum: New approach to the resummation of logarithms in Higgs-boson decays to a vector quarkonium plus a photon [Phys. Rev. D 95, 054018 (2017)]. *Phys. Rev. D*, 96:116014, 2017.
- [80] Geoffrey T. Bodwin, Hee Sok Chung, June-Haak Ee, and Jungil Lee.  $Z$ -boson decays to a vector quarkonium plus a photon. *Phys. Rev. D*, 97(1):016009, 2018.
- [81] Geoffrey T. Bodwin, Eric Braaten, and G. Peter Lepage. Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium. *Phys. Rev. D*, 51:1125–1171, Feb 1995.

- [82] G. Peter Lepage and Stanley J. Brodsky. Exclusive processes in perturbative quantum chromodynamics. *Phys. Rev. D*, 22:2157–2198, Nov 1980.
- [83] Chernyak, V. L. and Zhitnitsky, A. R. Asymptotic Behavior of Exclusive Processes in QCD. *Phys. Rept.*, 112:173, 1984.
- [84] C. N. Yang. Selection Rules for the Dematerialization of a Particle into Two Photons. *Phys. Rev.*, 77:242–245, Jan 1950.
- [85] Ye Li and Frank Petriello. Combining QCD and electroweak corrections to dilepton production in FEWZ. *Phys. Rev. D*, 86:094034, 2012.
- [86] Georges Aad et al. Search for Higgs and Z Boson decays to  $J/\psi\gamma$  and  $\Upsilon(nS)\gamma$  with the ATLAS detector. *Phys. Rev. Lett.*, 114:121801, 2015.
- [87] Vardan Khachatryan et al. Search for a Higgs boson decaying into  $\gamma^*\gamma \rightarrow \ell\ell\gamma$  with low dilepton mass in pp collisions at  $\sqrt{s} = 8$  TeV. *Phys. Lett. B*, 753:341, 2016.
- [88] Morad Aaboud et al. Searches for exclusive Higgs and Z boson decays into  $J/\psi\gamma$ ,  $\psi(2S)\gamma$ , and  $\Upsilon(nS)\gamma$  at  $\sqrt{s} = 13$  TeV with the ATLAS detector. 2018.
- [89] Esma Mobs. The CERN accelerator complex. Complexe des accélérateurs du CERN. Jul 2016. General Photo.
- [90] CMS Luminosity - Public Results. History: r140.
- [91] Tai Sakuma and Thomas McCauley. Detector and event visualization with SketchUp at the CMS experiment. Technical Report CMS-CR-2013-379. arXiv:1311.4942, CERN, Geneva, Oct 2013. Comments: 5 pages, 6 figures, Proceedings for CHEP 2013, 20th International Conference on Computing in High Energy and Nuclear Physics.
- [92] Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9:P10009, 2014.
- [93] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.
- [94] A. Benaglia. The CMS ECAL performance with examples. *JINST*, 9:C02008, 2014.
- [95] A. M. Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV. *JINST*, 13(06):P06015, 2018.
- [96] Vardan Khachatryan et al. Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV. *JINST*, 10:P08010, 2015.

- [97] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A*, 262:444–450, 1987.
- [98] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in  $pp$  collision events at  $\sqrt{s} = 7 \text{ TeV}$ . *JINST*, 7:P10002, 2012.
- [99] Serguei Chatrchyan et al. Alignment of the CMS tracker with LHC and cosmic ray data. *JINST*, 9:P06009, 2014.
- [100] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- $k_t$  jet clustering algorithm. *JHEP*, 04:063, 2008.
- [101] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur. Phys. J. C*, 72:1896, 2012.
- [102] D Contardo, M Klute, J Mans, L Silvestris, and J Butler. Technical proposal for the Phase-II upgrade of the CMS detector. Technical Report CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, Geneva, Jun 2015. Upgrade Project Leader Deputies: Lucia Silvestris (INFN-Bari), Jeremy Mans (University of Minnesota) Additional contacts: Lucia.Silvestris@cern.ch, Jeremy.Mans@cern.ch.
- [103] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO Higgs boson production via gluon fusion matched with shower in POWHEG. *JHEP*, 04:002, 2009.
- [104] Paolo Nason and Carlo Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 02:037, 2010.
- [105] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. A brief introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852, 2008.
- [106] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159, 2015.
- [107] Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J. C*, 76:155, 2016.
- [108] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [109] Sandro Palestini. Angular distribution and rotations of frame in vector meson decays into lepton pairs. *Phys. Rev. D*, 83:031503, 2011.
- [110] Ali Abbasabadi, David Bowser-Chao, Duane A. Dicus, and Wayne W. Repko. Radiative Higgs boson decays  $H \rightarrow ff\gamma$ . *Phys. Rev. D*, 55:5647, 1997.

- [111] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. [The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations](#). *JHEP*, 07:079, 2014.
- [112] John M. Campbell and R. Ellis. [MCFM for the Tevatron and the LHC](#). *Nucl. Phys. Proc. Suppl.*, 205:10, 2010.
- [113] A. Bodek, A. van Dyne, J. Y. Han, W. Sakumoto, and A. Strelnikov. [Extracting muon momentum xcale corrections for hadron collider experiments](#). *Eur. Phys. J.*, C72:2194, 2012.
- [114] Serguei Chatrchyan et al. [Measurement of the inclusive W and Z production cross sections in pp collisions at  \$\sqrt{s} = 7\$  TeV](#). *JHEP*, 10:132, 2011.
- [115] CMS Luminosity measurements for the 2016 data taking period. Technical Report CMS-PAS-LUM-17-001, 2017.
- [116] Jon Butterworth et al. [PDF4LHC recommendations for LHC Run II](#). *J. Phys. G*, 43:023001, 2016.
- [117] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. [Parton distributions for the LHC](#). *Eur. Phys. J. C*, 63:189, 2009.
- [118] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C.-P. Yuan. [New parton distributions for collider physics](#). *Phys. Rev. D*, 82:74024, 2010.
- [119] Sergey Alekhin et al. [The PDF4LHC Working Group Interim Report](#). 2011.
- [120] Michiel Botje, Jon Butterworth, Amanda Cooper-Sarkar, Albert de Roeck, Joel Feltesse, Stephano Forte, Alexander Glazov, Joey Huston, Ronan McNulty, and Sjöstrand. [The PDF4LHC Working Group Interim Recommendations](#). 2011.
- [121] Richard D. Ball, Valerio Bertone, Francesco Cerutti, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, Juan Rojo, and Maria Ubiali. [Impact of heavy quark masses on parton distributions and LHC phenomenology](#). *Nucl. Phys. B*, 849:296, 2011.
- [122] Giampiero Passarino. [Higgs boson production and decay: Dalitz sector](#). *Phys. Lett. B*, 727:424, 2013.
- [123] Thomas Junk. [Confidence level computation for combining searches with small statistics](#). *Nucl. Instrum. Meth. A*, 434:435, 1999.
- [124] Alexander L. Read. [Presentation of search results: The  \$CL\_s\$  technique](#). *J. Phys. G*, 28:2693, 2002.

- [125] ATLAS and CMS Collaborations, The LHC Higgs combination group. [Procedure for the LHC Higgs boson search combination in Summer 2011](#). Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, 2011.
- [126] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. [Asymptotic formulae for likelihood-based tests of new physics](#). *Eur. Phys. J. C*, 71:1554, 2011. [Erratum: Eur. Phys. J.C 73,2501(2013)].
- [127] J. Neyman and E. S. Pearson. *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, pages 73–108. Springer New York, New York, NY, 1992.
- [128] CMS Collaboration. The Phase-2 Upgrade of the CMS Endcap Calorimeter. Technical Report CERN-LHCC-2017-023. CMS-TDR-019, CERN, Geneva, Nov 2017. Technical Design Report of the endcap calorimeter for the Phase-2 upgrade of the CMS experiment, in view of the HL-LHC run.
- [129] CMS Luminosity - Public Results. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>
- [130] C. Adloff et al. Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype. *JINST*, 5:P05004, 2010.
- [131] N. Akchurin et al. First beam tests of prototype silicon modules for the cms high granularity endcap calorimeter. *Journal of Instrumentation*, 13(10):P10023, 2018.
- [132] M. Tanabashi et al. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.
- [133] XGBoost Documentation. <https://xgboost.readthedocs.io/en/latest/>
- [134] Energy Reconstruction in TB Simulation. <https://indico.cern.ch/event/770743/contributions/3205328/attachments/1747440/2829827/sim-energy-reco-11-06-18.pdf>.