

國 立 中 央 大 學

物 理 學 系
碩 士 論 文

**Search for the production of two Higgs
bosons in the final state with two photons
and two b quarks in proton-proton collision
at $\sqrt{s} = 13 \text{ TeV}$**

研 究 生：葉政威

指 導 教 授：郭 家 銘

中 華 民 國 一 ○ 七 年 六 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)，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: 葉政威 學號: 105222031

論文名稱: Search for the production of two Higgs bosons in the final states with

指導教授姓名: 郭家金 two photons and two b quarks in
proton-proton collision at $\sqrt{s}=13$

系所: 物理 所 博士班 碩士班 TeV.

填單日期: 2018/6/26

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 葉政威 研究生所提之論文

Search for the production of two Higgs bosons in the final state

with two photons and two b quarks in proton-proton collision 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 6 月 19 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 葉政威 研究
生所提之論文 Search for the production of two Higgs bosons
in the final state with two photons and two b quarks in
proton-proton collision at $\sqrt{s} = 13 \text{ TeV}$ 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
委 員

高崇文

張惠庭

郭家銘

中華民國 107 年 6 月 22 日

**Search for the production of two Higgs bosons in the
final state with two photons and two b quarks in
proton-proton collision at $\sqrt{s} = 13$ TeV**

by

Yeh, Cheng-Wei

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Master of Physics

at the

NATIONAL CENTRAL UNIVERSITY

June 2018

© National Central University 2018. All rights reserved.

Author

Department of Physics
July 18, 2018

Certified by

Kuo, Chia-Ming
Associate Professor
Thesis Supervisor

Search for the production of two Higgs bosons in the final state with two photons and two b quarks in proton-proton collision at $\sqrt{s} = 13$ TeV

by

Yeh, Cheng-Wei

Submitted to the Department of Physics
on July 18, 2018, in partial fulfillment of the
requirements for the degree of
Master of Physics

Abstract

The search is presented for the production of a pair of Higgs bosons in the final state with two photons and two b quarks by full 2016 data, which corresponds to an integrated luminosity of 35.9 fb^{-1} recorded by the CMS detector. Both resonant and non-resonant processes are investigated for the Standard Model (SM) and the Beyond the Standard Model (BSM) theories. The non-resonant production helps us to understand the Higgs field structure in the SM and other possible effects from BSMs. The resonant production is predicted by many BSMs. In this thesis, the hypothesis with the spin-0 and spin-2 new heavy particles which can decay into two Higgs bosons is searched and compared with the prediction from the warped extra dimension theory. The b-jet energy regression specifically developed for this analysis are employed to improve the sensitivity about 10%. The observed results agree with the standard model prediction, and the limits on the exclusion of the BSM productions are also set.

Thesis Supervisor: Kuo, Chia-Ming

Title: Associate Professor

摘要

本篇論文旨在尋找於質子-質子對撞下產生的雙希格斯粒子，分別衰變至一對光子及一對底夸克。本分析使用了於2016年由大強子對撞機(LHC)產生的質子-質子對撞，總能量為 $\sqrt{s} = 13 \text{ TeV}$ ，並由緊湊渺子線圈(CMS)所記錄，總亮度達到 35.9 fb^{-1} 。此研究基於標準模型以及超越標準模型的理論，同時尋找非共振衰變的雙希格斯粒子以及由新粒子衰變的雙希格斯粒子。非共振衰變可用於驗證希格斯機制以及探索其他可能的希格斯粒子與其他粒子的交互作用。多維度模型預測了兩種與重力相關的新粒子，且這些新粒子可衰變到雙希格斯粒子。本研究使用機器學習來輔助重建來自底夸克的強子噴流能量，使預測的生產截面的信心水準上限下降百分之十，達到更佳的結果。本研究沒有觀察到顯著的訊號事件，並提供了實驗上對理論參數以及新粒子重量的限制區間。

致謝

感謝指導教授郭家銘老師帶領我走過這四年的高能物理研究生活，教導我研究所需的技能以及知識，給予我出國參與國際團隊的機會，讓我得以一窺實驗高能物理研究的面貌。同時也感謝口試委員，讓這本論文得以完成。

感謝系上的老師們教導眾多的物理知識，讓我學會更多物理工具以及開拓高能物理領域之外的眼界。

感謝系辦的工作人員們以及高能實驗室助理張惠虹小姐協助處理繁瑣的手續，讓各項流程完美運轉。

感謝這段期間在高能物理實驗研究團隊的成員：Maxim、Andrey、Shilpi、Hien、Hoa、Victoria、Pocholo、Adam、張佑祥、張秋萍、曾師彥、杜仲倫、劉爲尹，以及好勝的八加九們：汪本璿、鄭皓仁、賴培築、李明晏、張用威、童宇軒、劉書孝、徐郁瑜、葉智翔、張峻堯。謝謝你們在研究中提供種種助力，以及在上課、研究之餘陪我玩。

感謝好同學們：蔡伯罕、蘇丰彥、王奕惟、藍亭奕、曾柏穎、馮思賢，以及熱音社的夥伴們，陪我玩、吵架、交換各種資訊、玩音樂，衝表演等，讓我的生活多采多姿。

最後感謝我的家人，提供我經濟後盾以及精神支持，讓我得以完成學業。

一百萬個感謝！

Contents

1	Introduction	1
1.1	Introduction	1
1.2	Theoretical overview	2
1.2.1	Higgs mechanism	2
1.2.2	Resonant pair production	5
1.2.3	Non-resonance pair production	8
1.2.4	Decay channels	12
1.2.5	Previous result	14
ATLAS Run I search	14
CMS Run I search	15
2	The LHC Machine and the CMS detector	19
2.1	The Large Hadron Collider	19
2.2	The Compact Muon Solenoid	21
2.2.1	Magnetic system	22
2.2.2	Tracking system	23
2.2.3	Electromagnetic calorimeter (ECAL)	24
2.2.4	Hadronic calorimeter (HCAL)	27
2.2.5	Muon system	27
2.2.6	Trigger and data acquisition system	29
3	Multivariate data analysis	33
3.1	Toolkit for Multivariate Data Analysis (TMVA)	33
3.1.1	Boosted decision trees (BDT)	33
3.1.2	Gradient Boosting and pruning	34
3.2	Photon identification	37
3.2.1	Di-photon vertex identification	37
3.2.2	Photon identification	38
3.3	Photon energy regression	39
3.4	B-jet identification	40
3.5	B-jet energy regression	46
3.5.1	Training samples	46

3.5.2	Training method	47
3.5.3	Performance	48
	Mass resolution	55
3.5.4	The improvement from the three additional input variables and the training with leading and trailing jets separately	57
3.5.5	Data validation	60
3.5.6	Impact on the results	65
3.6	$HH \rightarrow b\bar{b}\gamma\gamma$ categorization MVA	65
4	Event selection	69
4.1	Datasets	69
4.1.1	Data	69
4.1.2	Resonant Monte Carlo signal samples	69
4.1.3	Non-resonant Monte Carlo signal samples	72
4.1.4	Background simulation	74
4.2	Physical objects	74
4.2.1	The $H \rightarrow \gamma\gamma$ candidate	74
	Trigger mimic pre-selection	74
	$H \rightarrow \gamma\gamma$ selection	75
4.2.2	The $H \rightarrow b\bar{b}$ candidate	76
4.2.3	The di-Higgs system	76
4.2.4	Signal selection efficiency	78
5	Modeling	79
5.1	Signal modeling	79
5.1.1	Correlation Studies	80
5.2	Background modeling	81
5.2.1	Bias Studies	83
5.2.2	Correlation Studies	86
5.2.3	Single Higgs background modeling	87
5.3	Systematic uncertainties	89
6	Results	91
6.1	Resonance results	91
6.2	Non-resonance results	91
7	Conclusion	95
Bibliography		97

Chapter 1

Introduction

1.1 Introduction

The discovery of new boson with mass around 125 GeV at the Large Hadron Collider (LHC) by the CMS [1] and ATLAS [2] opened a new era of particle physics. The measured properties of this new particle are consistent with the Higgs boson in the Standard Model within the uncertainties. [3] In the Standard Model (SM), the Higgs mechanism involves Higgs potential field to solve the mass problem of the gauge bosons and the fermions. With the determined Higgs mass value, the structure of the Higgs potential field and the Higgs self-couplings are predicted in the SM. To confirm independently whether the new boson is Higgs boson in the SM, the measurement of the self-coupling constant is required. In LHC experiment, the double Higgs production is the only possible probe of the self-coupling constant. The di-Higgs production is mainly via gluon-gluon fusion in proton-proton collisions, which is a rare process. The cross section predicted by the LHC Higgs Cross Section Group is about 33.49 fb [4] in 13 TeV pp collisions, which is not expected to be sensitive at LHC. Nevertheless, there are many theories beyond the Standard Model (BSM) predicting other processes, which may increase the cross section of the di-Higgs boson production. Several BSM models extend the Standard Model, such as the Higgs singlet model, the two-Higgs-doublet model (2HDM), the minimal supersymmetric standard model (MSSM) and the warped extra dimension (WED) theory. They predict that there are new heavy particles, which mass can be up to TeV scale, decaying into a pair of Higgs bosons. Although new physics may not be observable in the low energy region, the effect of the BSMs in low energy scale can be parameterized as several anomalous couplings by the effective field theory (EFT). Both kinds of resonance and non-resonance production are presented in this thesis.

The search for the diHiggs production in this thesis explores the decay channel where one Higgs decays to a pair of b quarks and one decays to a pair of

photons. The data are collected by the CMS detector in full 2016, which corresponds to an integrated luminosity of 35.9 fb^{-1} .

This thesis is organized as follows: The motivation and the theoretical overview are described briefly in this chapter. In Chapter 2, an overview of the LHC and the CMS detector is presented. Chapter 3 describes the MVA techniques used in this analysis. Chapter 4 mentions the dataset and the event selection. In Chapter 5, the modeling for the signal and background estimation is introduced. The results are presented in Chapter 6, both for the resonant and non-resonant processes.

1.2 Theoretical overview

1.2.1 Higgs mechanism

The Higgs mechanism was published in 1964 independently by Robert Brout and François Englert [5], by Peter Higgs [6], and by Gerald Guralnik, C. R. Hagen, and Tom Kibble [7]. The concept of the Higgs mechanism is based on spontaneous symmetry breaking. The phenomenon of spontaneous symmetry breaking often happens in Nature. For example, a ball locates at the top of the dome and then rolls down in a specific direction. The rotational symmetry of this system was broken after rolling.

The Higgs potential is involved and assumed to be

$$V(\phi) = \mu^2 \phi^* \phi + \lambda (\phi^* \phi)^2, \quad (1.1)$$

where ϕ is the complex scalar field (particle) which interacts with the potential. The corresponding Lagrangian density can be written as

$$\mathcal{L} = (\partial_\mu \phi)^* (\partial^\mu \phi) - \mu^2 \phi^* \phi - \lambda (\phi^* \phi)^2. \quad (1.2)$$

The vacuum state should be the lowest potential energy state. To have a finite minimum, the parameter λ should be positive. The parameter μ^2 is required being negative to form a symmetry minimum ring, like Fig 1.1. The physical vacuum state will be at a particular point of the minimum potential ring, which breaks the global $U(1)$ symmetry.

In the Salam-Weinberg model, the Higgs mechanism is embedded to the $U(1)_Y \times SU(2)_L$ local gauge symmetry of electroweak sector. The electroweak scalar field should be a complex doublet, which contains one neutral field ϕ^0 and

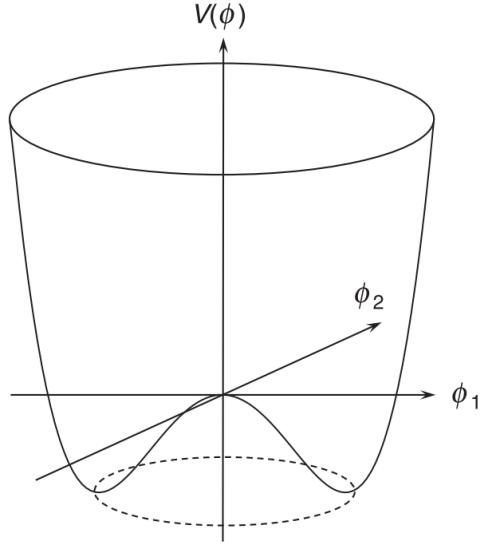


FIGURE 1.1: The structure of the Higgs potential for a complex scalar field $\phi = \phi_1 + i\phi_2$. [8]

one charged field ϕ^+ .

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (1.3)$$

According to the structure of the potential, the vacuum state with lowest potential satisfies

$$\phi^* \phi = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}, \quad (1.4)$$

where v is the vacuum expectation value. The vacuum expectation value v should not be zero to remain photon to be massless. The constraint is

$$|\langle 0 | \phi^0 | 0 \rangle| = \frac{v}{\sqrt{2}}. \quad (1.5)$$

Once the particular ground state is chosen, the symmetry is broken. The excited state of the field (particle) can be obtained from the perturbation of the vacuum state by introducing a real field $\eta(x)$.

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ v + \eta(x) + i\phi_4 \end{pmatrix} \quad (1.6)$$

According to the Goldstone theorem, this field contains one massive scalar field and three massless Goldstone bosons. These three Goldstone fields can be absorbed by doing appropriate gauge transformation with a local phase. The field

can be rewritten:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (1.7)$$

where Higgs field $h(x)$ is introduced. To respect Eq. 1.2 to be $SU(2)_L \times U(1)_Y$ local symmetry, the derivatives are replaced by the covariant derivative with two gauge fields \mathbf{W} and \mathbf{B} :

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + \frac{1}{2} i g_W \sigma^i \cdot \mathbf{W}_\mu^i + i g' \frac{Y}{2} \mathbf{B}_\mu, \quad (1.8)$$

where σ^i with ($i = 1, 2, 3$) are the three generators of $SU(2)$ symmetry and the weak hypercharge $Y = 2(Q - I_w^{(3)}) = 2(0 - (-\frac{1}{2})) = 1$. Combines Eq. 1.2, Eq. 1.7 and Eq. 1.8, the three massive gauge bosons and also one massless boson (photon) are determined by the terms of $(D_\mu \phi)^*(D^\mu \phi)$ in the Lagrangian. The masses of gauge bosons are found to be

$$m_W = \frac{1}{2} g_W v, \quad m_Z = \frac{1}{2} v \sqrt{g_W^2 + g'^2}, \quad m_h = \sqrt{2\lambda} v. \quad (1.9)$$

By the measurement of m_W and g_W , the vacuum expectation value is found to be $v = 246$ GeV.

The Lagrangian also describes the couplings between Higgs field and weak bosons, which include hWW , hZZ , $hhWW$, $hhZZ$, hhh and $hhhh$. The Higgs-self coupling can be described by the terms of the Higgs potential:

$$V(h) = \frac{1}{2} m_h^2 h^2 + \lambda v h^3 + \frac{\lambda}{4} h^4, \quad (1.10)$$

where the trilinear self-coupling constant is $\lambda_{hhh} = \lambda v = \frac{m_h^2}{2v}$ and also the quadrilinear self-coupling constant is $\lambda_{hhhh} = \frac{\lambda}{4} = \frac{m_h^2}{8v^2}$, which can be predicted by the measurement of Higgs boson mass m_h .

For fermions, the left-handed chiral fermions are described by $SU(2)$ doublets and right-handed fermions are described by $SU(2)$ singlets. The familiar process can be done to derive the fermion masses through the Yukawa coupling Lagrangian and the Higgs doublet.

Although the Higgs mass is already measured as about 125 GeV, the direct measurement of the self-coupling is necessary to provide an independent validation. The cross-section of the Higgs boson pair production and the triple Higgs boson production are known as the possible probes of the coupling constant. The triple Higgs boson production associated to the quadrilinear self-coupling is highly suppressed by v^2 . The possible triple Higgs boson production processes

are shown in Fig. 1.2, which are too complicated to explore in the experiment. The total cross section is in the order of 0.1 fb , which is too rare to search in the experiment in a foreseen future. [9] Therefore, the possible way to measure the Higgs self-coupling in experiment is through the Higgs boson pair production. In proton-proton collisions, it is mainly via gluon-gluon fusion, as Fig. 1.4 shown. [10, 11] There are two processes in the same order of magnitude, shown in Fig 1.3. One involves the trilinear Higgs boson self-coupling corresponding to λ_{hhh} . Another involves a heavy quark loop corresponding to the top quark Yukawa coupling $y_t = \sqrt{2}m_t/v$, which is called box diagram.

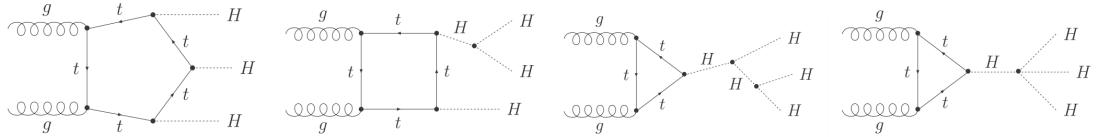


FIGURE 1.2: The Feynman diagrams of triple Higgs boson production via gluon-gluon fusion. From left to right: production through pentagon top loop, rectangle top loop with subsequent decay through trilinear self-coupling, triangle top loop with subsequent decay through two trilinear self-coupling and triangle top loop with one quartic self-coupling.

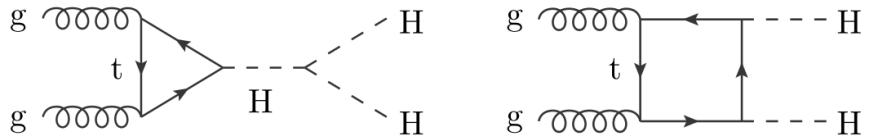


FIGURE 1.3: The Feynman diagrams of Higgs boson pair production.

Because these two diagrams are interfered destructively in leading order, the cross-section of the pair production becomes small. The cross section predicted by the LHC Higgs Cross Section Group is about 33.49 fb [4] in 13 TeV pp collisions, which is not sensitive at the LHC. However, it can be a probe of BSM physics. Several phenomena in Nature implicit that the SM is incomplete. For example, the SM doesn't include the gravity. The BSM physics associated Higgs pair production may increase the cross-section. In next section, two types of BSMs are introduced.

1.2.2 Resonant pair production

There are many models extending the SM and predicting that the new particles exist and can decay into a pair of Higgs bosons. For example, the Higgs singlet model involves a Higgs singlet field and mixes it with the original Higgs

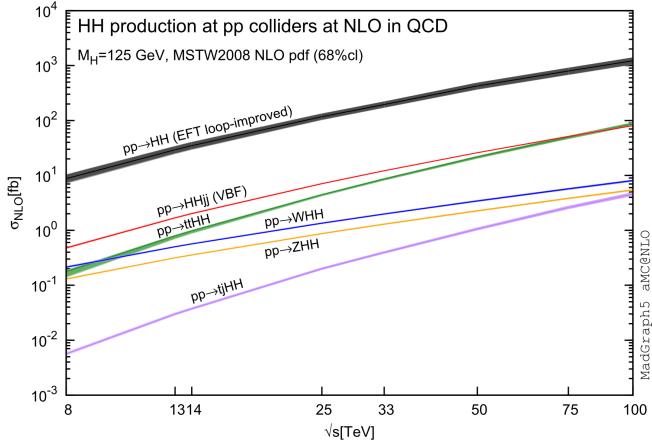


FIGURE 1.4: The cross-section of Higgs pair production through gluon-gluon fusion (black), vector boson fusion (red), and the pair production associated with top quark pair (green), vector boson (blue and yellow) and single top quark (purple). [11]

doublet. It predicts that there are a light SM-like Higgs boson h and a heavy Higgs boson H , which can couple to each other. Other extensions, such as two Higgs doublet model (2DHM) and minimal supersymmetric model (MSSM), involve two Higgs doublet fields and predict the existence of five new particles: two heavy charged Higgs bosons H^\pm , one heavy neutral Higgs bosons H , one pseudoscalar boson A and one light SM-like Higgs boson h .

On the other hand, models with extra dimension are developed to solve the problem of gravity. In the 1920's, the theory built by Kaluza and Klein involves the compact 5th spacetime dimension, which can be compactified as a circle, and try to unify the electroweak force and gravity force. Moreover, other theories try to explain the hierarchy problem between electroweak scale and the Planck scale. [12] These theories allow the compact 5-D space to be large. The new hierarchy problem about the compactification scale comes. To solve new hierarchy problem, the model proposed by Randall and Sundrum (RS model) introduces the concept of warped extra dimension (WED). [13] They consider that there are two 3-branes (three dimensional space with time), one allows SM field to propagate through (weak brane) and another is for gravity force (graviton brane). The metric of the 5-D spacetime is described as a 4-D spacetime multiplied by a "warp" factor and is referred to as the solution to 5-D Einstein's equations:

$$ds^2 = e^{-2kr_c\phi} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\phi^2, \quad (1.11)$$

where r_c is the compactification radius, k is the curvature to control the warped factor e^{-2kr_c} and ϕ is the 5th warped dimension. To solve the Einstein's equation, the value of k is equal to $\sqrt{\frac{-\Lambda}{24M^3}}$, where Λ is the cosmological constant, which

is also treated as vacuum energy density, and M is the Planck mass in 5-D dimension. The 5th dimension is constraint as a circle, which links two 3-branes (3-D space and time). The gravity brane is located at $\phi = 0$ and the weak brane is located at $\phi = \pi$, as Fig. 1.5 [14] shown. The interval between two branes is called "bulk". The two new particles are coming from the fluctuations of the 5-D

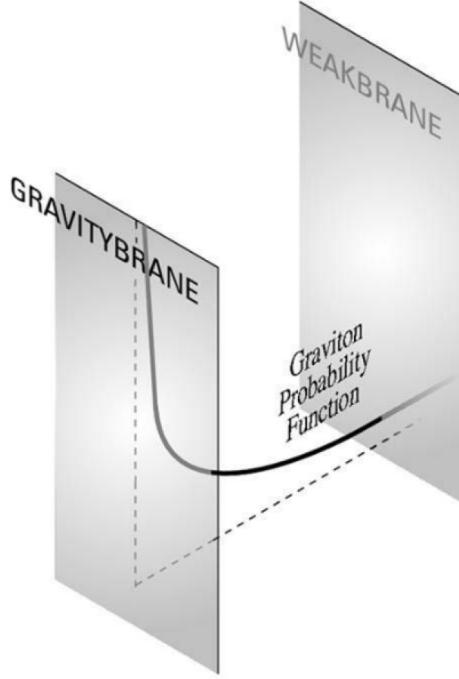


FIGURE 1.5: Scheme of dimensions on RS theory. The extra dimension ϕ compactified on the interval $(0, \pi)$ links two branes. The probability function is exponential decay curve.

metric, which are the excited fields of the metric. The tensor fluctuation from the original 4-D parts in the metric Eq. 1.11 generates the 4D effective massive particle, which is called Graviton. On the other hand, the scalar fluctuation of the 5th dimension predicts so-called Radion. To explore the different phase space of kinematics, these two particles are assumed arbitrarily that radion is spin-0 and graviton is spin-2. The angular distribution of spin-0 is distinct from the spin-2 particles by different kinematics. There are two scenarios of the RS model. One is that the SM fields only can propagate on the weak brane (RS1 model), another is that the SM fields can propagate in the bulk (bulk RS model). In the latter case, the coupling strength between gravity and the SM fields depends on the position in the bulk of the SM fields, which gives us more interesting phenomenology. For example, the light quarks can be localized closely to the Planck brane, while the top quark is near the weak brane and has a large mass. [15] The

properties of radion are similar in these two scenarios. It can be parameterized $\Lambda_R = \sqrt{6}e^{-\pi kr_c}k\sqrt{\frac{M_5^3}{k^3}}$ for radion. The properties of bulk graviton can be parameterized by the curvature $\tilde{k} = k/\bar{M}_{Pl}$, where $\bar{M}_{Pl} = M_{Pl}/\sqrt{8\pi}$. The Planck mass here is the effective 4-D Planck mass obtained by multiplying the volume of the compact space, which becomes $M_{Pl}^2 = M_{4+n}^{n+2}V_n$, where $n = 1$ is the number of the extra dimension and M_{4+n} is the (4+n)-D Planck mass. In this thesis, the theoretical interpretation is based on the bulk scenario and described in Ref. [14] and Ref. [16]. The details of setup are shown in Sec. 4.1.2.

1.2.3 Non-resonance pair production

The value of λ_{hhh} is already predicted by the SM once the mass of Higgs bosons m_h and vacuum expectation value v are measured. However, λ_{hhh} can be modified by considering the effect of the BSM physics. [17] In some BSM models, the mass of new particles can reach several TeVs and hard to achieve. Their effects in the low energy scale are possible to influence the contribution of di-Higgs boson production, which become the indirect way to study BSMs. For example, the effect on the di-Higgs production can be parameterized and described as the deviation from the SM prediction, which is defined as $\kappa_\lambda = \lambda_{hhh}/\lambda_{hhh}^{SM}$. In different models, the constraints of κ_λ can be different, which is the hint of the discrimination to different BSMs. In general, the large range of κ_λ is allowed and κ_λ can be referred to as a free parameter.

A more general approach can be derived by the effective field theory (EFT) with top-down perspective, which provides a model-independent way to search new physics. The concept is that the physics which are well understood in unobservable high energy scale can be reduced and integrated to match the phenomena in low energy scale. The effective terms of new physics are described by the high dimensional operators and suppressed by the powers of their high energy scale Λ . The BSM effects in the SM measurement can be considered as possible uncertainties, and increase the Higgs pair production rate. The effective Lagrangian relative to the Higgs pair production can be written as:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \sum_i \frac{c_i^{(5)}}{\Lambda} \mathcal{O}_i^{(5)} + \sum_i \frac{c_i^{(6)}}{\Lambda^2} \mathcal{O}_i^{(6)} + \sum_i \frac{c_i^{(7)}}{\Lambda^3} \mathcal{O}_i^{(7)} + \dots, \quad (1.12)$$

where $\mathcal{O}^{(n)}$ is the n-dimensional operators with corresponding *Wilson coefficients* $c_i^{(n)}$. At the LHC, the EFT only with 6-D operators are focused [4, 18, 19], because 5-D operator violates the lepton number [20] and higher dimensional operators are highly suppressed.

After spontaneous symmetry breaking, the modified coupling terms are generated by these high dimension operators. For the di-Higgs production through gluon-gluon fusion, the effective terms can be rewritten with the anomalous coupling constants, which provides a simple physics interpretation: [21]

$$\begin{aligned}\Delta\mathcal{L} = & \frac{1}{2}\partial_\mu H\partial^\mu H - \frac{m_H^2}{2}H^2 - \kappa_\lambda\lambda^{SM}vH^3 \\ & - \frac{m_t}{v}(v + \kappa_t H + \frac{c_2}{v}HH)(\bar{t}_L t_R + h.c.) \\ & + \frac{\alpha_s}{12\pi v}(c_g H - \frac{c_{2g}}{2v}HH)G_{\mu\nu}^A G^{A,\mu\nu}.\end{aligned}\quad (1.13)$$

It contains five anomalous Higgs couplings. The two modified couplings which already exist in the SM are described as the deviation from the SM Higgs self-coupling κ_λ , and the deviation from the SM top Yukawa coupling κ_t . Three pure BSM contact interaction couplings are the coupling of a gluon pair with a Higgs boson pair c_{2g} , the coupling and of a gluon pair with a single Higgs boson c_g , the coupling of a top quark pair with a Higgs boson pair c_2 . Corresponding Feynman diagrams are shown in Fig. 1.6.

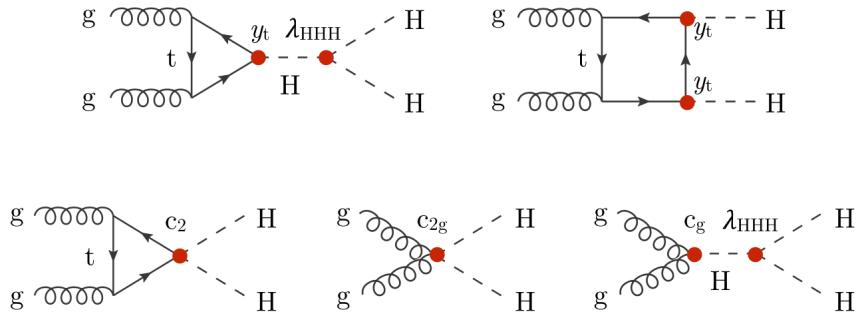


FIGURE 1.6: The Feynman diagrams of anomalous coupling in HH production.

The effect on the cross section of Higgs pair production is written as the function of five anomalous coupling constants, which is shown in Eq. 1.14. [21, 22]

$$\begin{aligned}R_{HH} = \frac{\sigma_{HH}}{\sigma_{HH}^{SM}} = & A_1\kappa_t^4 + A_2c_2^2 + (A_3k_t^2 + A_4c_g^2)\kappa_\lambda^2 + A_5c_{2g}^2 \\ & + (A_6c_2 + A_7\kappa_t\kappa_\lambda)\kappa_t^2 + (A_8\kappa_t\kappa_\lambda + A_9c_g\kappa_\lambda)c_2 \\ & + A_{10}c_2c_{2g} + (A_{11}c_g\kappa_\lambda + A_{12}c_{2g})\kappa_t^2 \\ & + (A_{13}\kappa_\lambda c_g + A_{14}c_{2g})\kappa_t\kappa_\lambda + A_{15}c_g c_{2g}\kappa_\lambda.\end{aligned}\quad (1.14)$$

The coefficients A_i can be extracted from a fit to the cross section estimated by the simulation in different BSM parameter points [22]. The fit results are shown in Table. 1.1.

TABLE 1.1: Coefficient of Eq. 1.14 in 13 TeV proton-proton collisions. [22]

$A_1=2.09$	$A_2=10.15$	$A_3=0.28$	$A_4=0.10$	$A_5=1.33$
$A_6=-8.51$	$A_7=-1.37$	$A_8=2.83$	$A_9=1.46$	$A_{10}=-4.92$
$A_{11}=-0.68$	$A_{12}=1.86$	$A_{13}=0.32$	$A_{14}=-0.84$	$A_{15}=-0.57$

Not only the cross-section, but the signal topology is influenced by the coupling constants. The small modification of the coupling may change the kinematic of the final states drastically. Because exploring all possible combination of five coupling constants is time consuming and not possible for experiment search, the shape benchmark models are built and described in Ref. [21]. The 5-D parameter space is scanned to understand their kinematic properties. The events which have similar kinematic distribution shape are grouped as a cluster. The latest recommended 12 benchmark points are shown in Tab. 1.2, and the kinematic distribution shapes are demonstrated in Fig. 1.7.

 TABLE 1.2: The benchmark points in 5D parameter space of non-resonance HH production with $N_{cluster} = 12$.

Benchmark	κ_λ	κ_t	c_2	c_g	c_{2g}
1	7.5	1.0	-1.0	0.0	0.0
2	1.0	1.0	0.5	-0.8	0.6
3	1.0	1.0	-1.5	0.0	-0.8
4	-3.5	1.5	-3.0	0.0	0.0
5	1.0	1.0	0.0	0.8	-1
6	2.4	1.0	0.0	0.2	-0.2
7	5.0	1.0	0.0	0.2	-0.2
8	15.0	1.0	0.0	-1	1
9	1.0	1.0	1.0	-0.6	0.6
10	10.0	1.5	-1.0	0.0	0.0
11	2.4	1.0	0.0	1	-1
12	15.0	1.0	1.0	0.0	0.0
SM	1.0	1.0	0	0	0

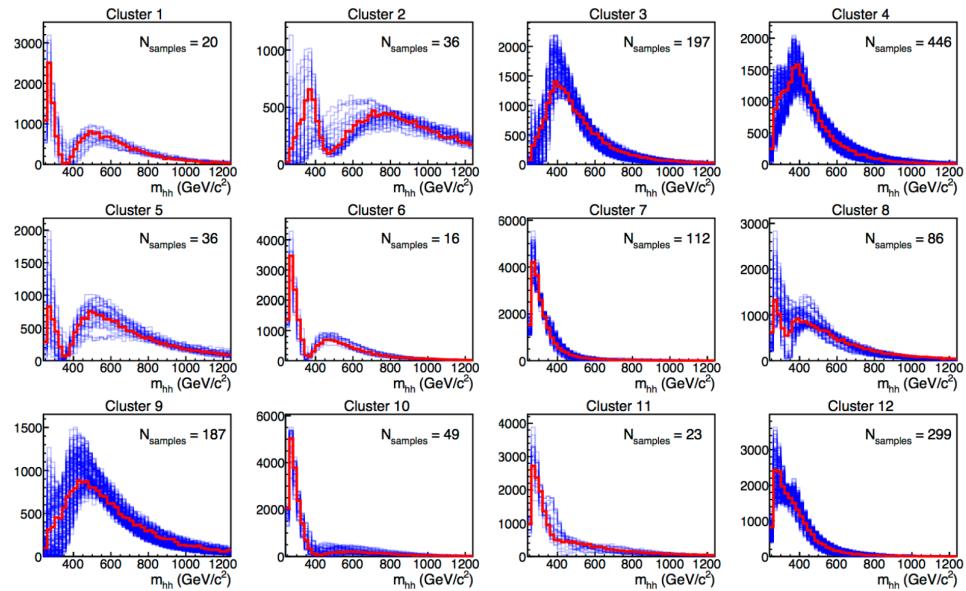


FIGURE 1.7: Generator-level distribution of di-Higgs boson mass m_{hh} in different clusters. [21] Total 1507 points are scanned and split into 12 clusters.

1.2.4 Decay channels

The SM expected Higgs boson decay channels and branching ratio as a function of the Higgs mass near 125 GeV are shown in Fig. 1.8. The decay branching ratio of Higgs with mass $m_H = 125.09$ GeV is summarized in Tab. 1.3 and the di-Higgs decay branching ratio of some channels is shown in Fig. 1.9. Because the Higgs pair production is quite rare, the preferred decay channels are limited. To be sensitive in experiment, it is requested that at least one Higgs which decays into a pair of b quarks or a pair of W bosons due to the large branching ratios. However, the $H \rightarrow bb$ channel suffers from QCD background contamination. The $H \rightarrow WW$ channel usually requests that one of the W bosons should decay leptonically to suppressed multi-jets background, which reduces the branching ratio. The $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$, $b\bar{b}\tau\tau$ and $b\bar{b}WW$ channels are expected to be the most sensitive channels, where the first three are explored in CMS Run I (see Sec. 1.2.5). The advantages and challenges strongly depend on different final states. The $b\bar{b}b\bar{b}$ channel is benefited by the highest branching ratio and fully reconstructed by the b-tagging technique. (see Sec. 3.4) Additionally, this channel is possible to be explored in high mass region due to the special topology of two overlapped b jets from a highly boosted objects. [23] The $b\bar{b}WW$ channel profits the second largest branching ratio, but is contaminated by $t\bar{t} \rightarrow b\bar{b}WW$. The $b\bar{b}\tau\tau$ channel is trade-off between the branching ratio and the signal purity. Although it is hard to distinguish the τ lepton which decays into electron and muon (labeled as τ_{lep}), the τ decaying into hadronic jets (labeled as τ_{had}) is well tagged due to the characteristic of small jet cone with a low particle multiplicity. [24] The $b\bar{b}\gamma\gamma$ channel is a clean channel due to high selection efficiency and good energy resolution in $H \rightarrow \gamma\gamma$ decay. This channel is suffer from the low branching ratio.

TABLE 1.3: The summarized table of Higgs decay branching ratio with Higgs mass $H_m = 125.09$ GeV. [4]

Decay mode	Branching ratio (%)
$H \rightarrow bb$	58.09
$H \rightarrow WW$	21.52
$H \rightarrow gg$	8.180
$H \rightarrow \tau\tau$	6.256
$H \rightarrow cc$	2.884
$H \rightarrow ZZ$	2.641
$H \rightarrow \gamma\gamma$	0.227
$H \rightarrow Z\gamma$	0.154
$H \rightarrow \mu\mu$	0.022

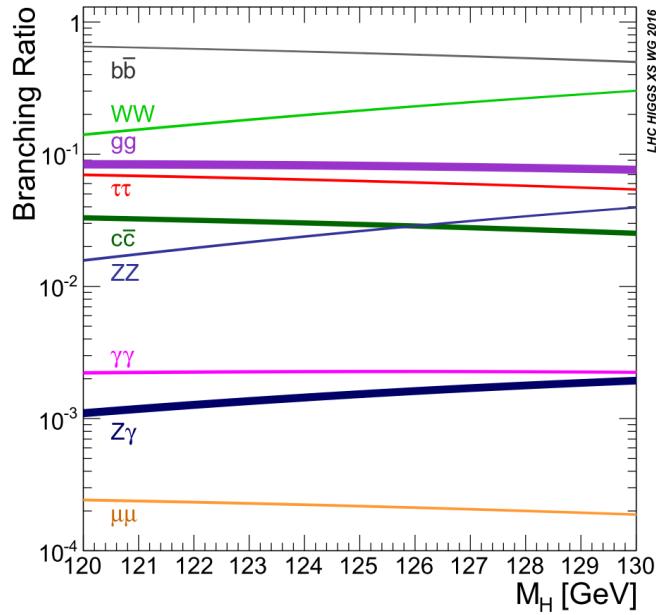


FIGURE 1.8: The Higgs decay branching ratio as a function of the Higgs mass near 125 GeV. [4]

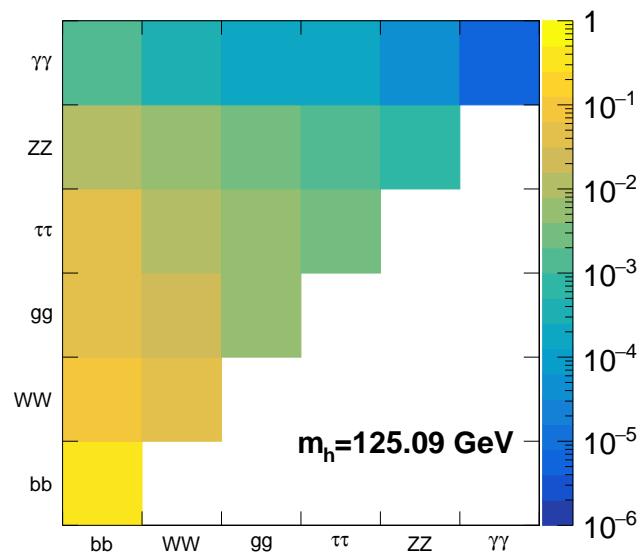


FIGURE 1.9: The Higgs boson pair decay branching ratio in five channels.

1.2.5 Previous result

ATLAS Run I search

Both searches of the resonance and the non-resonance di-Higgs boson productions are performed by the ALTAS collaboration, using LHC run I data of pp collision at $\sqrt{s} = 8 \text{ TeV}$ corresponding to an integrated luminosity of 20.3 fb^{-1} . The final states with $b\bar{b}b\bar{b}$ [25], $b\bar{b}\gamma\gamma$ [26], $b\bar{b}\tau_{had}\tau_{lep}$ and $\gamma\gamma WW$ [27] are searched and evaluated the combined sensitivity assuming the SM Higgs decay branching ratios.

In case of resonant di-Higgs production, the spin-0, neutral and heavy resonance Higgs H is searched. The upper limits at 95% confidence level (CL) upper limits are set for the product of the production cross-section $\sigma(gg \rightarrow H \rightarrow hh)$ as function of the resonance mass m_H . The searched resonance mass range depends on different channels. The observed and expected limits of searched channels and combined result are shown in Fig. 1.10. The combined limit varies from 2.1 pb at 260 GeV to 0.011 pb at 1000 GeV. The most significant excess is at 300 GeV with 2.5 standard deviations σ . The results are interpreted to exclude the resonance mass ranges in different BSM models.

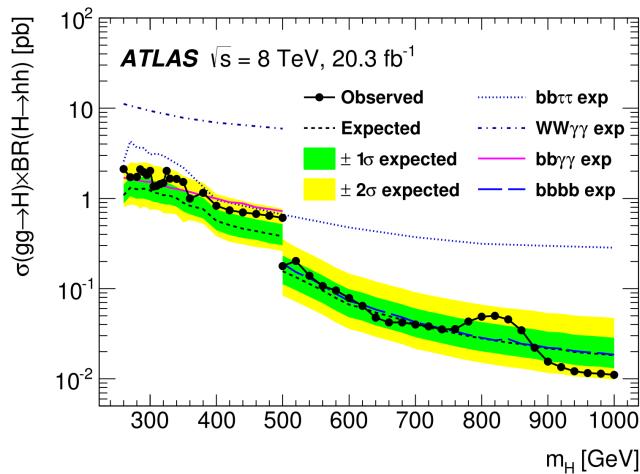


FIGURE 1.10: The observed and expected 95% CL upper limits on $\sigma(gg \rightarrow H \rightarrow hh)$ at $\sqrt{s} = 8 \text{ TeV}$ as function of the resonance mass m_H , which combine all of the searched channels. [27]

For the non-resonance di-Higgs production, the SM di-Higgs production is searched and the results are compared with the SM predicted cross-section $\sigma_{SM}(gg \rightarrow hh) = 9.9 \pm 1.7 \text{ fb}$ with Higgs mass $m_H = 125.4 \text{ GeV}$ at 8 TeV pp collisions [28]. The summary of non-resonance result of Higgs pair production is shown in Tab. 1.4. The observed (expected) combined upper limit is 0.69 (0.47)

pb, which is about 70 (48) times larger than SM prediction. The combined observed result presents 1.7 standard deviations.

All results are no significant excess in the data beyond the background expectation.

TABLE 1.4: The observed and expected 95% CL upper limits on the cross-section of non-resonance process $\sigma(gg \rightarrow hh)$ at $\sqrt{s} = 8\text{ TeV}$.

Channel	Observed (expected) upper limit on $\sigma(gg \rightarrow H \rightarrow hh)$	Relative to σ_{SM}
$\gamma\gamma bb$	2.2 (1.0)	220 (100)
$\gamma\gamma WW$	11.4 (6.7)	1150 (680)
$bb\tau\tau$	1.6 (1.3)	160 (130)
$bbbb$	0.62 (0.62)	63 (63)
Combined	0.69 (0.47)	70 (48)

CMS Run I search

By the CMS collaboration, the resonance searches are performed by $b\bar{b}b\bar{b}$ [29], $b\bar{b}\gamma\gamma$ [30] and $b\bar{b}\tau_{had}\tau_{had}$ channels [31], using LHC run I data of pp collision at $\sqrt{s} = 8\text{ TeV}$ corresponding to an integrated luminosity of about 19 fb^{-1} . Both spin-0 and spin-2 resonance HH productions are searched and combined with three searched channels. The combined observed (expected) limits on $\sigma(pp \rightarrow X \rightarrow HH)$, which are shown in Fig. 1.11, range from 1134 (776) and 1088 (760) fb at $m_x = 300\text{ GeV}$ to 21 (31) and 18 (26) fb at $m_x = 1000\text{ GeV}$ for spin-0 and spin-2 resonances respectively. The results are compared with the theoretical prediction based on the bulk and RS1 models and perform the exclusions. The resonance in very high mass range from 1 TeV to 3 TeV is explored in $b\bar{b}b\bar{b}$ channel [32] and from 800 to 2500 GeV in $b\bar{b}\tau_{had}\tau_{lep}$ channel [33], which is benefited by the special b-tagging technique for boosted b jets. [23]

The non-resonance of SM-like di-Higgs production is searched in $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau\tau$ channels and combined together. [31] The observed (expected) limit of 0.43 (0.47) fb is set for the SM Higgs pair production cross-section with decay branching ratio, which is 43 (47) times larger than the SM prediction. The deviation κ_λ from SM Higgs self-coupling are scanned in $b\bar{b}\gamma\gamma$ channel, and the result is shown in Fig. 1.12. The range of κ_λ in $\kappa_\lambda < -17$ and $\kappa_\lambda < -22.5$ is excluded. [30]

All results are no significant excess in the data beyond the background expectation.

The comparison of the limits on spin-0 resonance HH production cross-section as function as the resonance mass m_X are shown in Fig. 1.13, which includes the

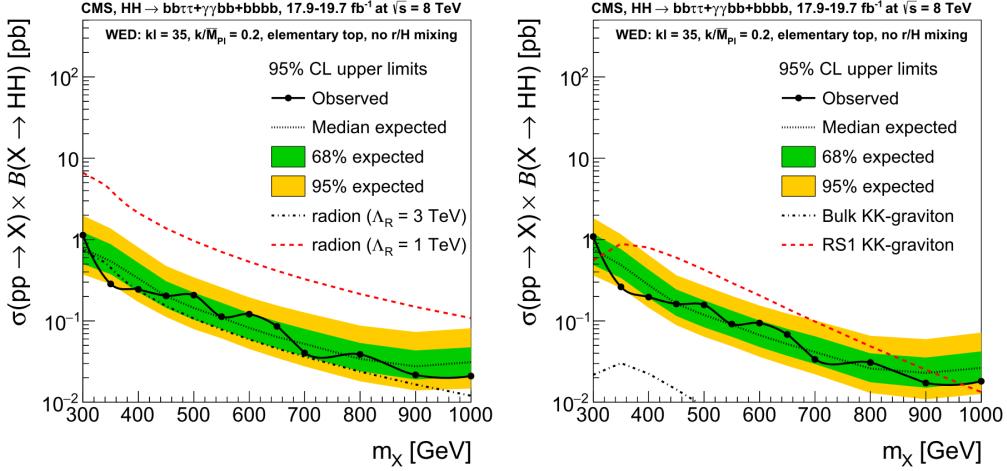


FIGURE 1.11: The combined observed and expected limits on resonance Higgs pair production for spin-0 (left) and spin2 (right). The theoretical curves are predicted by WED models with different parameters: Radion in different mass scale $\Lambda_R = 1, 3 \text{ TeV}$, graviton in different scenarios. The other WED parameters are $kl = kr_c = 35$ and $k/\overline{M}_{Pl} = 0.2$. [31]

searches in different channels by the CMS and the combination results by the ATLAS. The different channels complement each other in the different resonance mass range. For example, the $b\bar{b}\gamma\gamma$ channel dominates the sensitivity of low mass region $m_X \leq 400 \text{ GeV}$, the $b\bar{b}\tau\tau$ channel is good at intermediate mass region $400 \text{ GeV} < m_X \leq 700 \text{ GeV}$ and the $b\bar{b}b\bar{b}$ channel dominates the high mass region $m_X > 700 \text{ GeV}$. Therefore, the exploration in different channels is important to increase the sensitivity.

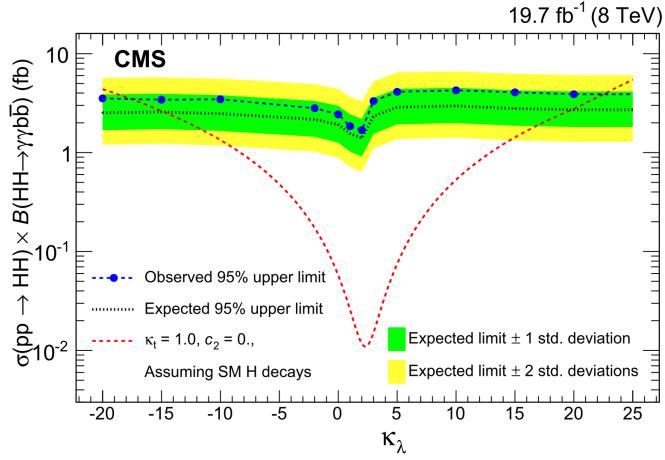


FIGURE 1.12: The limits on the cross-section of non-resonance di-Higgs production in $b\bar{b}\gamma\gamma$ channel in scanned anomalous coupling κ_λ . The other coupling constants are the same as SM prediction. [30]

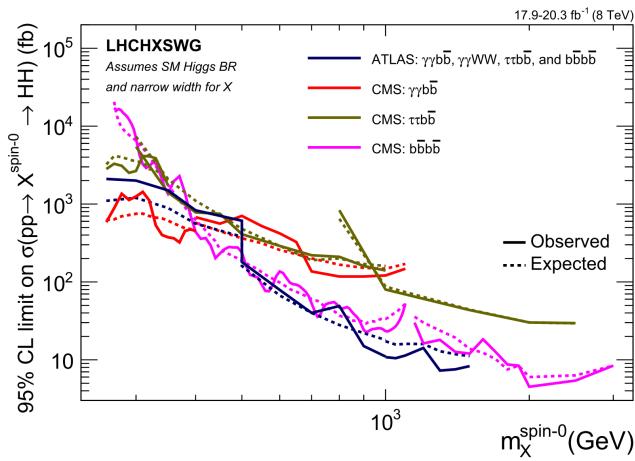


FIGURE 1.13: The comparison of the limits on $pp \rightarrow X^{spin-0} \rightarrow HH$ in different HH decay channels by the CMS collaboration, also the combined result from the ATLAS collaboration. [4]

Chapter 2

The LHC Machine and the CMS detector

This thesis is based on the data of proton-proton collisions recorded by the Compact Muon Solenoid (CMS) at the Large Hadron Collider at CERN. The protons are accelerated at the energy of 6.5 TeV by the LHC and collided. This chapter presents the overview of the LHC and the CMS.

2.1 The Large Hadron Collider

The Large Hadron Collider is the largest and most powerful particle accelerator in the world. It consists of a 26.7 km superconducting ring with a number of accelerating structures, underground about 100 meters. Fig. 2.1 shows the location and the accelerator complex. The LHC is designed to accelerate and collide proton beams in the centre-of-mass energy of protons at 14 TeV with luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. [34]

The protons are born from the hydrogen atoms which are ionized by an electric field. The journey of protons begins in the linear accelerator, Linac2. The protons are accelerated to 50 MeV. Then the protons beam is injected into the Proton Synchrotron Booster (PSB), accelerated to 1.4 GeV, and fed to the Proton Synchrotron to be accelerated to 25 GeV. Next, they are sent to the Super Proton Synchrotron to be accelerated to 450 GeV. Finally, they are transferred to the two separate beam pipes of the Large Hadron Collider (LHC), split up into clockwise and anti-clockwise direction. The protons are accelerated to 6.5 TeV in the LHC, travel in the speed about 99.99% times the speed of light. The superconducting dipole magnets providing about 8 T magnetic field are used to keep the particles in their circular orbits. The collisions occur at the intersection points in the LHC rings. There are four intersection points used for the experiments: CMS, ATLAS, LHCb and ALICE. The quadrupole magnets focus the two beams to perform a



CERN's Accelerator Complex

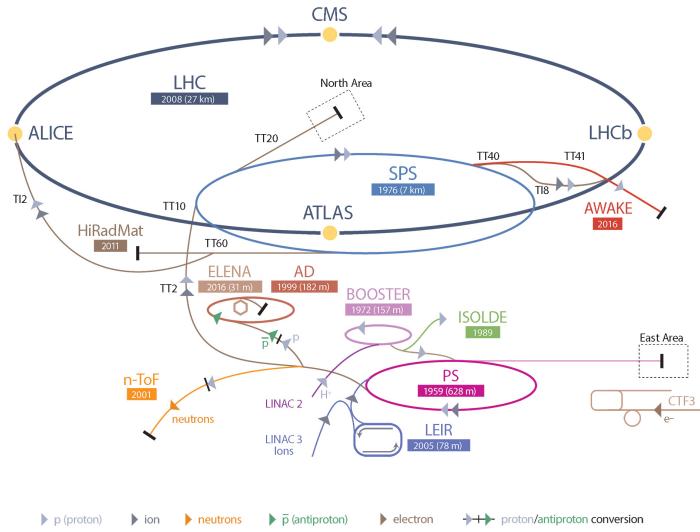


FIGURE 2.1: (Top) The LHC is located beneath the border between Switzerland and France. (Bottom) The accelerator complex of the LHC.

bunch crossing, enhance the probability of the collisions. Under normal operating conditions, the LHC provides a proton beam containing 2808 bunches. The bunch length is about 7.5 cm and the spacing is 25 ns (about 7.5 m). Each bunch is expected to contain about 1.15×10^{11} protons.

In 2016, the centre-of-mass energy of protons reached 13 TeV. The maximum number of injected bunches are 2208. The beam intensity is up to 2.5×10^{14} , about 1.25×10^{11} protons in each bunch. The peak instantaneous luminosity is $1.4 \times 10^{34} \text{ cm}^{-2} \text{s}^{-1}$, 40% above the design value. [35] The delivered integrated luminosity is 40.82 fb^{-1} and 37.76 fb^{-1} are recorded by the CMS. (Fig. 2.2) [36]

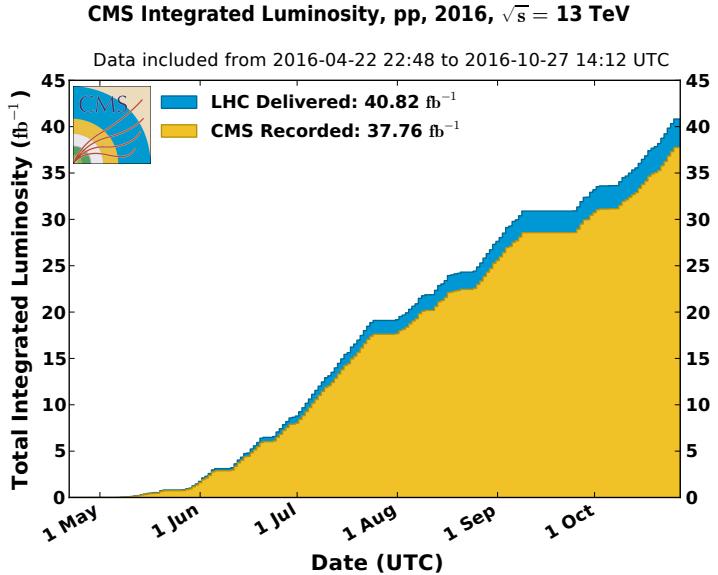


FIGURE 2.2: The integrated luminosity delivered by the LHC and recorded by the CMS in 2016.

2.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is a general-purpose detector at the LHC [37], designed to record the proton-proton collisions with high luminosity. To increase the luminosity, the high intensity of protons released by the LHC can generate more inelastic interaction and produce about 1000 particles every 25 ns . In 2016, the average number of interaction vertices produced per bunch crossing, called pile-up, is up to 27 as shown in Fig 2.3, which is about 25% higher than 2012. The high pile-up effect is a grand challenge for reconstruction. Therefore, the detectors have to feature high granularity and fast response. The prime goals of the CMS are to verify the Standard Model and also to explore other BSM (beyond Standard Model) physics at TeV energy scale. The CMS detector is cylindrical symmetry around the beam axis, composed of a barrel and two end-caps with a diameter 15 m and a length 28.7 m . It is quite compact with a weight about 14000 tonnes. The main components of the CMS are a superconducting solenoid electromagnet providing 3.8 Tesla magnetic field inside the solenoid, silicon tracking system, electromagnetic calorimeter, hadronic calorimeter, and also a muon detecting system outside of the solenoid. The geometry of sub-detectors can be found in Fig. 2.4.

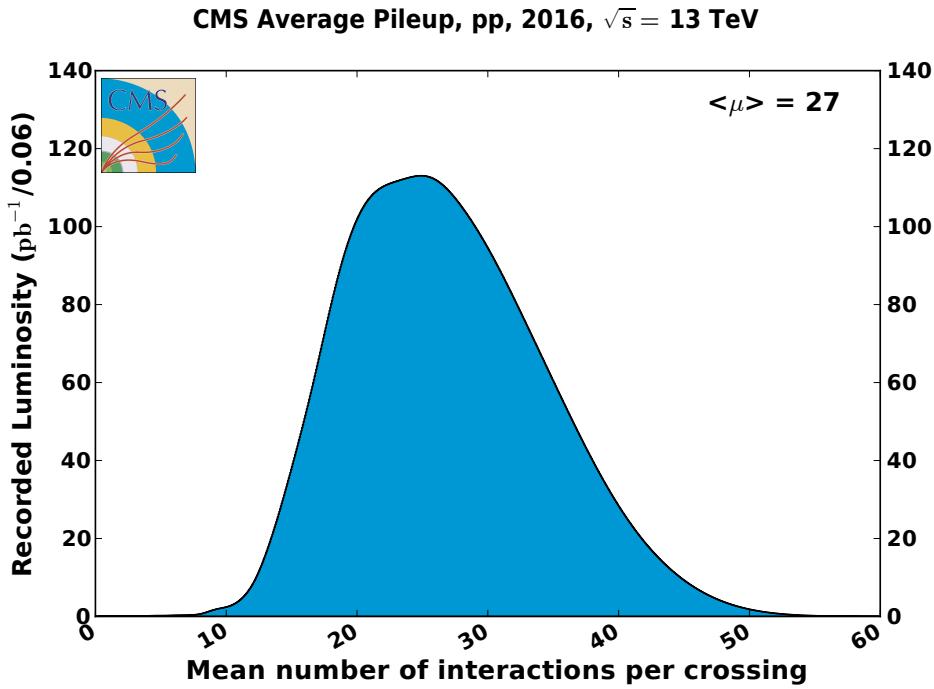


FIGURE 2.3: The distribution of number of pile-up per bunch crossing in the CMS in 2016. The maximum number is up to 53 and the average is 27.

2.2.1 Magnetic system

The magnetic system are a superconducting cylindrical coil with a length of 12.5 m and a diameter of 6 m , designed to reach a uniform 4 T field in a free bore. The superconducting coils are Rutherford-type cables made by NbTi conductor, form as the 4-layer winding, total 2168 turns operating current 19.5 kA , storing energy 2.69 GJ . [38] The coils are in the cold box and cooled by the helium to 4.45 K and insulated by a vacuum vessel. There is a steel yoke outside the solenoid, where the muon system is housed, to return the magnetic flux. The yoke comprises 5 wheels and 2 endcaps, composed of three layers each. The magnetic field in the barrel and outer endcap return yoke is about 1.7 T and points in a direction opposite to the direction of the field inside the coil. Inside the solenoid, the tracking system, electromagnetic calorimeter, and the hadronic calorimeter are installed. The muon system is outside the solenoid and interspersed in the return yoke. This system used to bend charged particles flying outward from the collision point helps to identify the charge and improves the measurement of momentum.

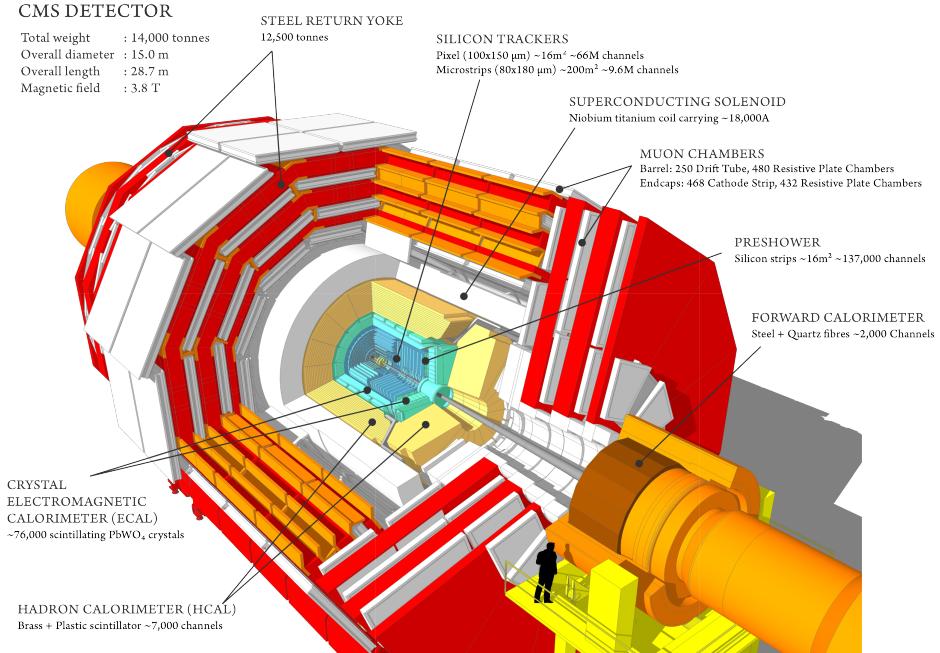


FIGURE 2.4: A perspective view of the CMS detector.

2.2.2 Tracking system

The inner tracking system (tracker) surrounds the collision points and occupies a cylindrical volume with a length of 5.8 m and a diameter of 2.5 m . The solenoid provides a homogeneous magnetic field of 4 T over the full tracker, and is expected to reconstruct the trajectories of charged particles with transverse momentum p_T around 0.8 GeV . The reconstruction of trajectories is also used to reconstruct the primary vertex as well as secondary vertex, which is the most important part of the b-tagging technique. The tracker is based on the silicon detector technology. The charged particles cause small ionization current which can be measured when they pass through the silicon. It is composed of a pixel detector with 3 barrel layers and a silicon strip tracker with 10 barrel layers. For the two endcaps, each comprises 2 disks of pixel detector and 3 inner disk plus 9 outer disks of the strip trackers. The acceptance of the tracker is up to a pseudorapidity of $|\eta| < 2.5$. The full layout is shown in Fig. 2.5 and described in Tab. 2.1. [39] The size of each pixel cell is $100 \times 150\text{ }\mu\text{m}^2$, and provides the hit position resolution about $10\text{ }\mu\text{m}$ in the transverse coordinate and $20\text{-}40\text{ }\mu\text{m}$ in the longitudinal coordinate. Total 1440 pixel modules are used with 66 million readout channels. On the other hand, the strips tracker has 15148 modules with total 9.3 million strips, and provides position resolution in $r\phi$ approximately $13\text{-}38\text{ }\mu\text{m}$ (TIB, TID), $18\text{-}47\text{ }\mu\text{m}$ (TOB, TEC). More details can be found in Ref. [40].

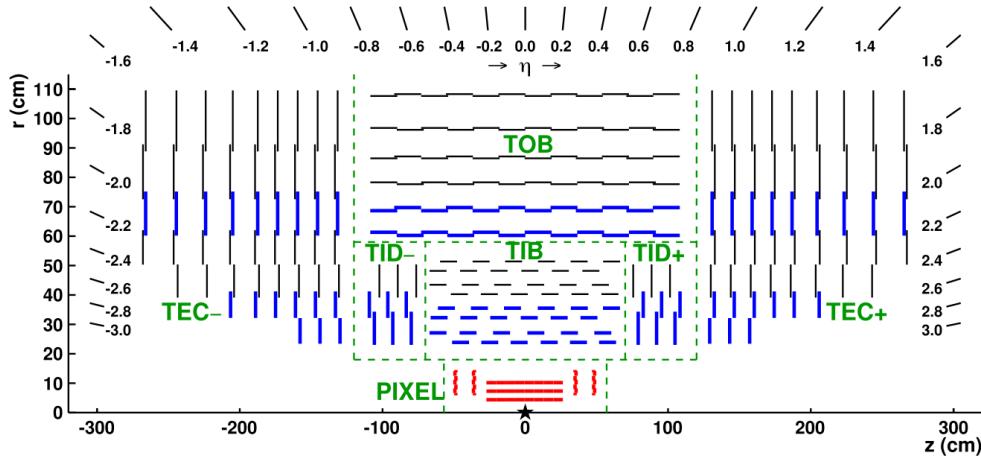


FIGURE 2.5: The layout of the trackers in r - z plane. The start point is the interaction point. The other components are Pixel Detector (PIXEL), Tracker Inner Barrel (TIB), Tracker Outer Barrel (TOB), Tracker Inner Disk (TID), and Track endcap (TEC).

TABLE 2.1: A summary table of the components of the tracker. Pitch in the barrel is the distance between neighbouring strips. The barrel strips run parallel to the beam axis.

Tracker subsystem	Layers	Pitch	Location
Pixel tracker barrel	3 cylindrical	$100 \times 150 \mu\text{m}^2$	$4.4 < r < 10.2 \text{ cm}$
Strip tracker inner barrel (TIB)	4 cylindrical	$80 - 120 \mu\text{m}$	$20 < r < 55 \text{ cm}$
Strip tracker outer barrel (TOB)	6 cylindrical	$122 - 183 \mu\text{m}$	$55 < r < 116 \text{ cm}$
Pixel tracker endcap	2 disks	$100 \times 150 \mu\text{m}^2$	$34.5 < z < 46.5 \text{ cm}$
Strip tracker inner disks (TID)	3 disks	$100 - 141 \mu\text{m}$	$58 < z < 124 \text{ cm}$
Strip tracker endcap (TEC)	9 disks	$97 - 184 \mu\text{m}$	$124 < z < 282 \text{ cm}$

2.2.3 Electromagnetic calorimeter (ECAL)

The electromagnetic calorimeter (ECAL) in CMS is a hermetic homogeneous calorimeter, able to obstruct the electromagnetic (EM) particles and measures their energy deposits and positions. One of the important tasks of ECAL is to explore the $H \rightarrow \gamma\gamma$ decay channel. The lead tungstate crystals PbWO_4 are used to be scintillators, which have short radiation length ($X_0 = 0.89 \text{ cm}$) to stop the EM particles, small Moliere radius (2.2 cm) to have small electromagnetic showers. When EM particles strike a scintillator material, the material is excited and then emits light. The scintillation emission spectrum of the ECAL crystals peaks at around 440 nm . The emitted light is received by the photo diodes to measure the energy deposits of the incoming EM particles. About 80% scintillation light is emitted in 25 ns . The time-response is fast to cope with the bunch crossing time of the LHC. Although the crystal is radiation hard, the radiation damage still affects the transparency of the crystals. The correction for the transparency

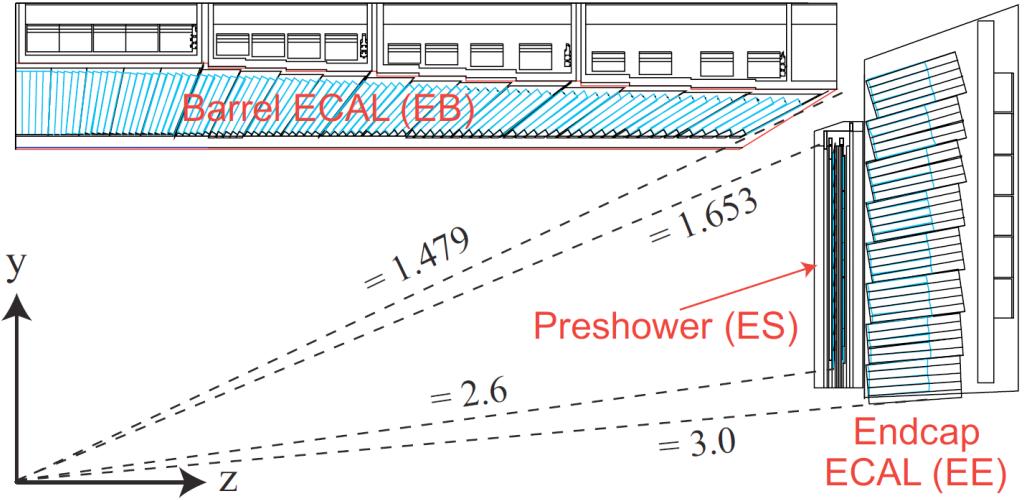


FIGURE 2.6: The layout of ECAL.

loss is needed. [41]

ECAL is constructed with a barrel and two endcaps. The ECAL barrel (EB) covers the pseudorapidity range $|\eta| < 1.479$. It is constructed by 61200 crystals, and each of them is $22 \times 22 \text{ mm}^2$ at the front face and $26 \times 26 \text{ mm}^2$ at the rear face (0.0174×0.0174 in $\phi\text{-}\eta$), with a length 230 mm ($25.8 X_0$). The ECAL endcap (EE) covers the pseudorapidity range $1.479 < |\eta| < 3$. The crystal size is $28.62 \times 28.62 \text{ mm}^2$ at the front face and $30 \times 30 \text{ mm}^2$ at the rear face with a length 220 mm ($24.7 X_0$). To suppress the background $\pi^0 \rightarrow \gamma\gamma$ for $H \rightarrow \gamma\gamma$, the spatial resolution should be improved. A sampling detector, Preshower (ES), installed in front of EE, covers the pseudorapidity range $1.65 < |\eta| < 2.61$. Each endcap contains two lead absorbers and two planes of silicon strips which are orthogonal to each other. The plane of silicon strips are behind the absorber to measure the EM showers from the incoming particles striking the absorber. The radiation length of the absorber is $2 X_0$ for the front plane and $1 X_0$ for the rear plane. The whole layout is shown in Fig. 2.6 and Fig. 2.7. More details can be found in Ref. [42].

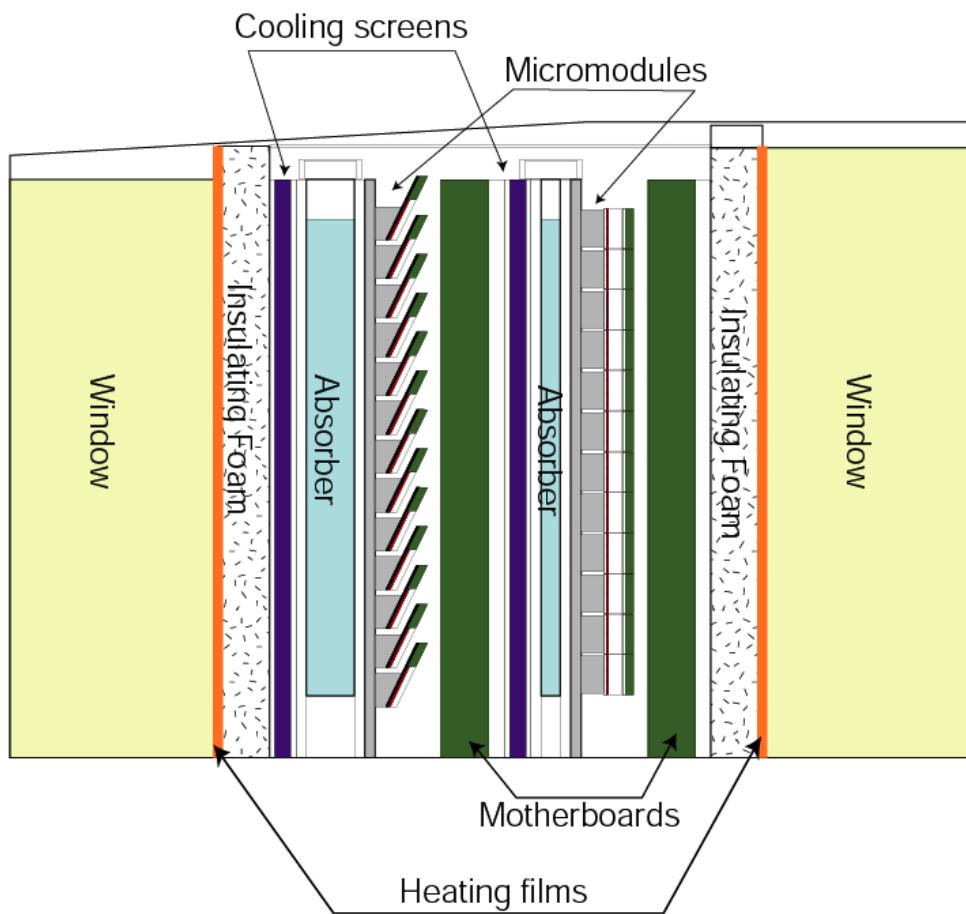


FIGURE 2.7: The layout of Preshower.

2.2.4 Hadronic calorimeter (HCAL)

The Hadron calorimeter is a sampling calorimeter surrounding the ECAL, designed to detect the strongly interacting particles and help to determine the missing energy. The HCAL is composed of four parts: barrel (HB), endcap (HE), outer (HO) and forward (HF). The longitudinal scheme is shown in Fig. 2.8. The barrel covers pseudorapidity range $|\eta| < 1.3$, between the outer extent of the ECAL (1.77 m) and inner extent of the magnetic coil (2.95 m). The absorbers in HCAL barrel consist a front steel plate with a thickness 40 mm, 8 layers brass plates with a thickness 50.5 mm, 6 layers brass plates with thickness 56.5 mm and a back steel plate with a thickness 75 mm. Total thickness is 5.82 interaction lengths at 90°. The brass materials are non-ferromagnetic and high nuclear interaction length, which are easy to be simulated and make the detector be more compact. The plastic scintillator behind the absorber is segmented in $\eta - \phi$ towers granularity 0.087×0.087 , which is equivalent to the area of the 5×5 ECAL crystals. 70000 tiles in the scintillators are connected to the wavelength shifting fibres which collect the emitted light and shift the blue-violet light to the green light. The endcaps cover the region $1.3 < |\eta| < 3$. The absorbers are 79-mm-thick brass plates. The granularity in $\eta - \phi$ is 0.087×0.087 for $|\eta| < 1.6$ and 0.17×0.17 for $|\eta| \geq 1.6$. The outer calorimeter is located outside the magnetic coil, is designed to complete the containment for hadron showers of HB and HE. The HO is placed as the first layer of the return yoke. The magnetic coil is used as an additional absorber equal to $1.4/\sin\theta$ interaction length. At $\eta = 0$, the interaction length of the HB is minimum, there are two layers of scintillators with 19.5 cm thick piece of iron. Other wheels have a single layer. The granularity of HO is same as HB. Lastly, the two forward calorimeters (HF) are positioned at $3 < |\eta| < 5.2$, outside the muon chambers. The HF will experience extremely high particles fluxes, about 10 MGy in the LHC with an integrated luminosity $5 \times 10^5 pb^{-1}$. This calorimeter is sensitive to the EM component of showers, because the shower particles can achieve the energy above the Cherenkov threshold. The absorber in the HF is a cylindrical steel with an inner radius 12.5 cm, an outer radius 130 cm and a length 165 cm along z-axis. The granularity of HF is 0.175×0.175 in $\eta - \phi$ towers. More details can be found in Ref. [43].

2.2.5 Muon system

The particles produced from the collision are absorbed by the inner detectors with high probability. The most particles escape outside are muons, which provide an unmistakable signature. The muon system is able to identify muons,

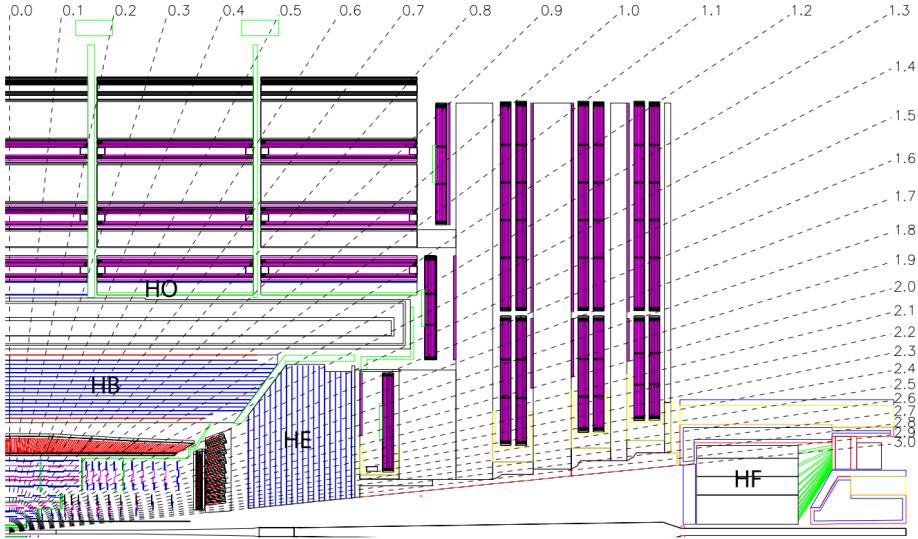


FIGURE 2.8: Longitudinal view of the HCAL, showing the location of the barrel (HB), endcap (HE), outer (HO) and forward (HF) of HCAL.

measure their momentum, and deal with the muon triggering. It is composed of a barrel and two endcaps with three kinds of gaseous detectors. In the barrel region at $|\eta| < 1.2$, the muon rate is low and the 4-T magnetic field is quite uniform in the return yoke. The drift tube (DT) chambers with 400 ns drifting time are used, which is interspersed among the layers of the return yoke, total 4 stations. 3 stations are consisted of 8 chambers to measure in $r - \phi$ bending plane, and 4 chambers measure in z direction. The other one station doesn't include the z -direction measurement, but provides the best angular resolution. The drift cells of each chamber are shifted by a half-cell to eliminate the gaps. In both endcaps at $0.9 < |\eta| < 2.4$, the muon rates are high and the magnetic field is non-uniform. The Cathode strip Chambers (CSC) are used for the endcaps, which have fast response and high radiation resistance. There are 4 stations in each endcap interspersed in the return yoke. Each strip is run radially outward and provides $r - \phi$ plane measurement. The anode wires are also used to measure the η position and the beam-crossing time. They are perpendicular to the cathode strips. To have a better muon triggering, the resistive plate chambers (RPC) are added in both the barrel and endcaps at $|\eta| < 1.6$. The RPCs provide a fast response with good time resolution but coarse spatial resolution. It also helps to resolve the tracks from the multi-hits in a chamber. Full layout is shown in Fig. 2.9. More details can be found in Ref. [44].

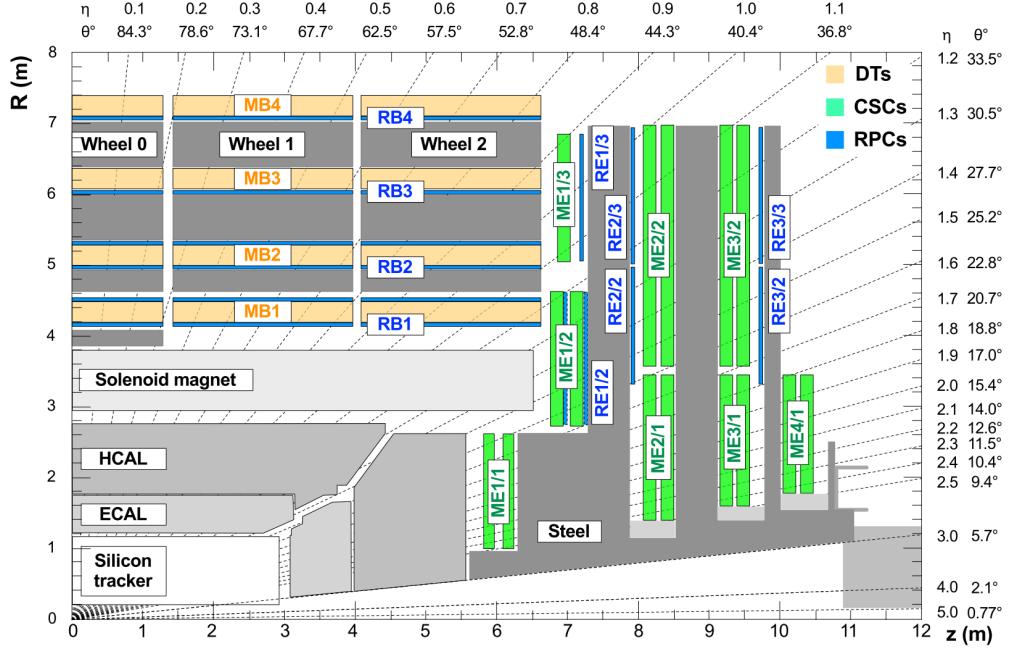


FIGURE 2.9: Layout of one quadrant of the muon system.

2.2.6 Trigger and data acquisition system

At the designed luminosity in the LHC, there are about 10^9 events generated per second, which is impossible to process and store every event. Therefore, it is important to reduce the data rate and select the events we interest. The trigger system is the start of the event selection process in two steps – level-1 trigger (L1) and high level trigger (HLT).

The level-1 trigger is composed by custom-designed, large programmable electronics based on FPGA with maximum output rate 100 kHz, which are installed partly on the detectors. (Other parts are in the control room.) The L1 trigger has to analyze each bunch crossing every 25 ns without any dead-time, but the trigger latency is 3.2 μ s. The data are stored in the pipelined memories first and wait for the L1 trigger decision transmitted to the detector electronics within 3.2 μ s. Due to the 3.2 μ s restriction, the event selection by L1 should rely on the coarser data and simple algorithm, from the calorimeters and muon system.

There are three components of the L1: local, regional and global as shown in Fig. 2.10. The local trigger is based on the calorimeters and the muon trigger, respectively. The calorimeter local triggers are done by the Trigger Primitive Generator circuit in each calorimeter cell and begins with the sum of the tower energy from ECAL, and HCAL. The regional calorimeter trigger (RCT) combines the information of all the local triggers of ECAL and HCAL to determine

the trigger objects, such as photons, electrons, taus and jets. It also finds the isolated/non-isolated photons/electrons separately. Then the global calorimeter trigger (GCT) sorts trigger objects by the energy and quality, selects top 4 of each type of the trigger objects. It also calculates the total transverse energy and missing energy vector. The muon local triggers are composed of three different algorithms corresponding to the different kinds of muon chambers, which are described in Sec. 2.2.5 and Ref. [45]. Then the Global Muon Trigger combines three local triggers and validates the muon sign, sorts muons by the energy and quality, selects top 4 of muons. Finally, the global trigger combines all information and makes a decision.

Next, the data acquisition (DAQ) system receives the events which pass the L1 triggers and provide the computing power for the filter systems (HLT). The structure of the DAQ is shown in Fig. 2.11. The information will be transmitted to the DAQ, which are used, calculated by the global calorimeter trigger, the global muon trigger and the global trigger. One zero-suppressed event occupies about 1 MByte, the data flow to the DAQ is about 100 GByte/s. The computer farm of DAQ with thousands of commercial CPUs is designed to process such kind of huge data with more complex algorithms and performs the High Level Trigger. The High Level Triggers is similar to the off-line reconstruction, but more efficient and lower CPU-time-consuming. The CMS software framework are used for the HLT and off-line reconstruction, which is open source and available in Ref. [46]. Finally, the rate of data recorded for off-line analysis is on the order of 10^2 Hz. The more details about the HLT and DAQ are in Ref. [47]

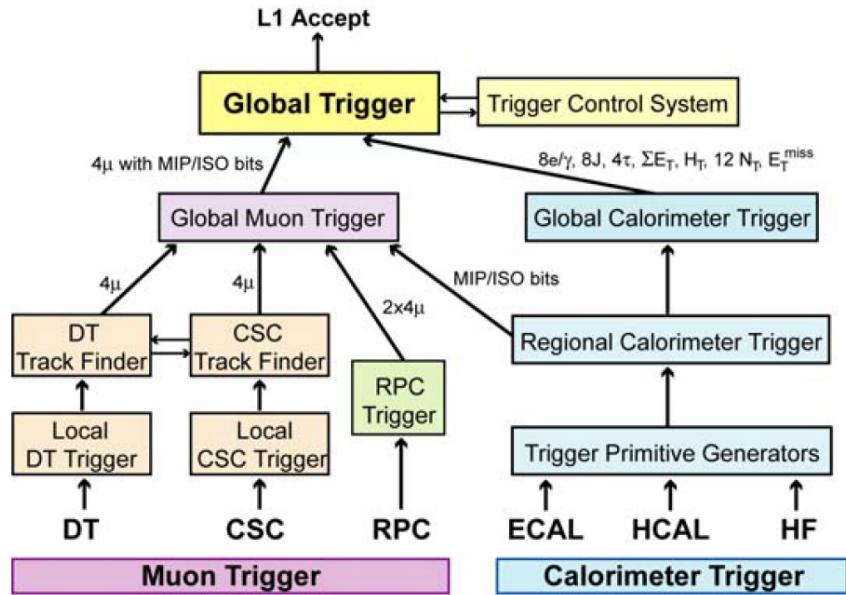


FIGURE 2.10: Architecture of the CMS L1 trigger system.

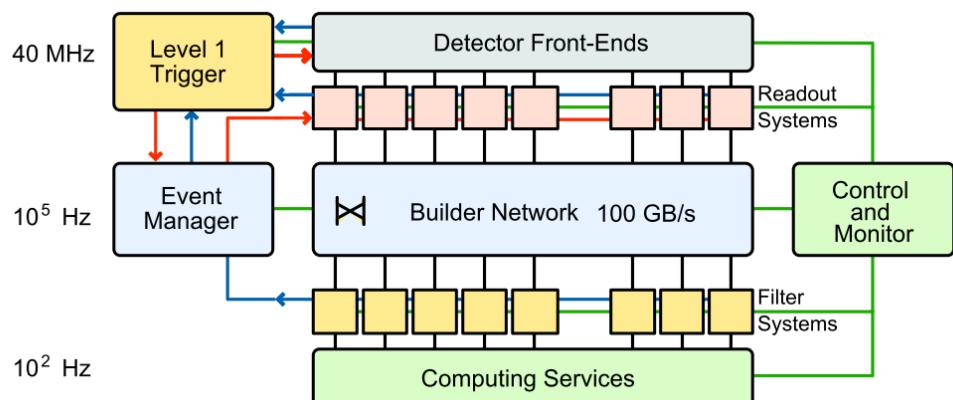


FIGURE 2.11: Architecture of the CMS DAQ system.

Chapter 3

Multivariate data analysis

3.1 Toolkit for Multivariate Data Analysis (TMVA)

To extract small signals from large data, the multivariate analysis (MVA) technique is used for signal/background classification and energy correction (regression). In this analysis, the Toolkit for Multivariate Data Analysis with ROOT [48] is used for photon identification, b-tag identification, b-jet energy regression, and the categorization. It provides a machine-learning environment to process multivariate classification and regression techniques, which is designed for high energy physics. The algorithms in TMVA belong to “supervised learning”. The events used for training contain known output target to build the prediction function that describes the decision boundary (classification) or the target value (regression). In this section, the MVA algorithm used in this analysis is introduced.

3.1.1 Boosted decision trees (BDT)

The MVA method often used in CMS to build classification models or regression functions is boosted decision trees (BDT). [48] A decision tree is a binary tree structured like Fig. 3.1. The tree starts from a root node containing the full training sample, and grows by creating splitting nodes repeatedly. Each branch node splits training data to left or right subsets by a discrimination of one variable. The discrimination is optimized to have the best separation power by scanning all variables. The range of each variable is divided into several grid points for scanning, which is set by `nCut` in the program. There are several methods provided by TMVA to find the criteria. For classification, the default one is according to Gini index. The Gini index of a subset is defined as

$$Gini = p \cdot (1 - p), \quad (3.1)$$

where p is the signal purity $p = \text{signal}/(\text{background} + \text{signal})$. The sum of the Gini index of two separated subsets weighted by their fraction of events should be minimum, which means that the purity is increased. To find the next node, the events in the subsets should be more than the setting of minimum number of events. The parameter given to the program `MinNodeSize` defines the minimum percentage of training events in each node. The tree will grow until reaching the maximum depth, which is set by `MaxDepth`. Finally, the terminal nodes (leaves) are identified as signal-like or background-like according to the signal purity of the events in the nodes.

For regression, the scheme is similar to classification. The criteria are found by average square error, which is defined by

$$\text{Average square error} = \frac{\sum_i^N (y_i - \hat{y})^2}{N}, \quad (3.2)$$

where N is the number of events in the node, y is the value of the target variable of event i , \hat{y} is the mean of all events in the node. The criteria with the smallest weighted sum of error value will be chosen. The leaves are labeled as the average value of the events in the nodes.

In principle, a decision tree can grow until the signal and background are fully separated or predicts every target value in the training samples, like Fig. 3.2. However, this kind of model takes the random fluctuation and extreme events into account. It is too sensitive to the known training events and hard to give a general prediction for the test data. That is so-called overtraining (overfitting). To avoid overtraining, the decision tree needs to be pruned. Also, a decision tree is altered easily by the statistical fluctuation of the training sample. The so-called “boosting” technique is introduced to increase the stability. The same training events are used to grow different decision trees and form a forest. The parameter `NTrees` defines the number of trees in the forest in the program. The forest is combined into one single BDT discriminator for classification problems or one target value for the regression. The detail of the boosting and pruning method used in this analysis is described in Sec. 3.1.2.

3.1.2 Gradient Boosting and pruning

Boosting is a kind of mathematical optimization to reduce bias in machine-learning. It describes that a strong learner (the built model which can predict the target precisely and is close to the truth function) can be formed by some

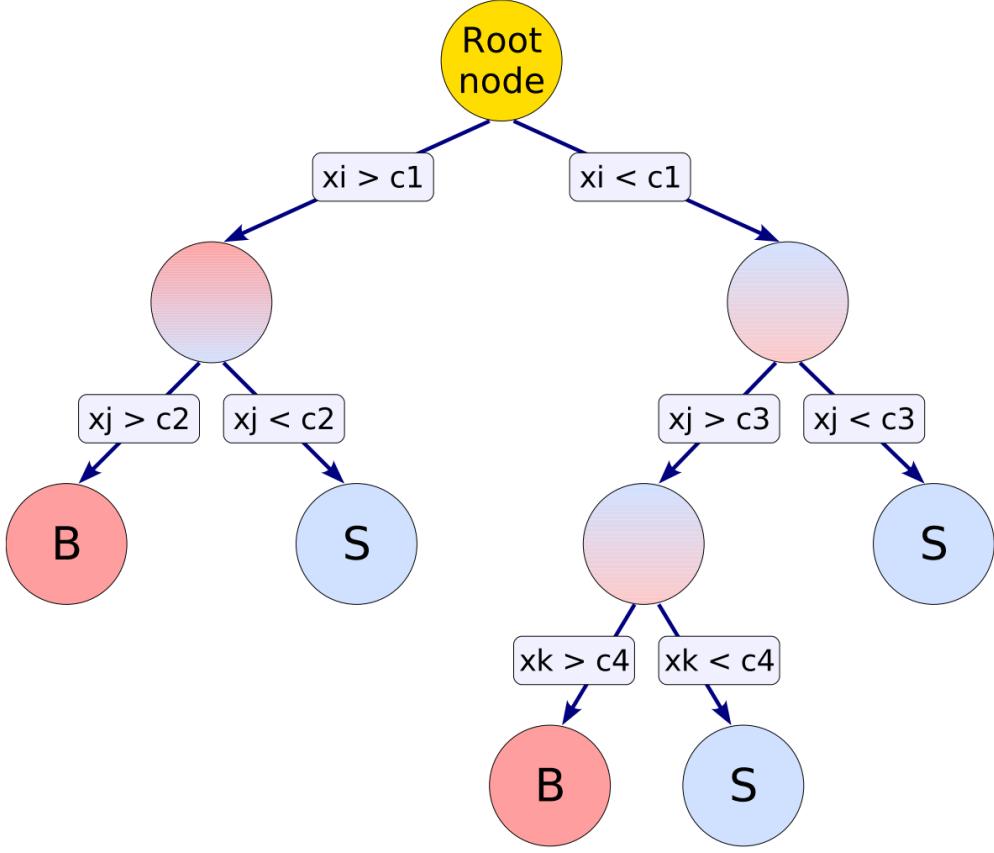


FIGURE 3.1: The schematic of a decision tree. The tree grows from the root node and is split by the scanned criterion " $x_i (> or <) c_a$ " repeatedly until reaching the maximum depth. Finally, the terminal nodes (leaves) are identified as signal-like or background-like.

weak learners (the models which have low classification power). In the gradient boosting algorithm, it starts from a weak learner F_0 function, then finds next stronger learner step by step:

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + h(\mathbf{x}), \quad 1 \leq m \leq M, \quad (3.3)$$

where \mathbf{x} is a vector which includes all input variables, M is the number of trees we want to grow.

In each step, the deviation of the function $F_m(\mathbf{x})$ from the true function should be reduced. The deviation can be described by the so-called *loss function*. For binary classification (signal or background), the chosen loss function is the binomial log-likelihood loss

$$L(F, y) = \ln(1 + e^{-2F(\mathbf{x})y}), \quad (3.4)$$

where y is the output value of the true function (classification score) from the

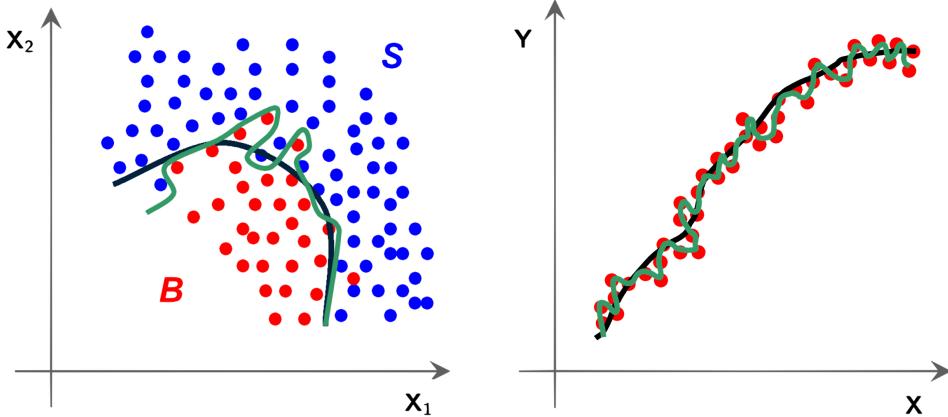


FIGURE 3.2: The sketch of overtraining for classification (left) and regression (right). The green curves are the model which is overtrained, and black curves are the expected general models.

training sample. From Eq. 3.3, the learner $h(\mathbf{x})$ is equal to $F_m(\mathbf{x}) - F_{m-1}(\mathbf{x})$, which is the deviation between a stronger learner and a weaker leaner. Like the gradient descent, $h(\mathbf{x})$ can be derived by the first order gradient of the *loss function*:

$$h(\mathbf{x}) = -\frac{\partial L(F_{m-1}(\mathbf{x}), y)}{\partial F_{m-1}(\mathbf{x})}. \quad (3.5)$$

Then the tree of the function $h(\mathbf{x})$ grows and updates the $F_m(\mathbf{x})$ accordingly:

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \nu \cdot \gamma \cdot h(\mathbf{x}), \quad (3.6)$$

where ν is the shrinkage parameter used to control the learning rate. Empirically, the small shrinkage (0.1-0.3) gives the better prediction. The distance of descent γ is optimized to minimize the sum of deviation $\sum_{i=1}^n L(F_m, y_i)$ when growing the tree. After M iterative steps, the final model F_M is obtained as output model.

For regression, the true value y can be the target variable, and the Huber loose function is used:

$$L(F, y) = \begin{cases} \frac{1}{2}(y - F)^2 & |y - F| \leq \delta \\ \delta(|y - F| - \frac{\delta}{2}) & |y - F| > \delta \end{cases}. \quad (3.7)$$

The BDT method works best with the weak learners, which are the small trees with limited growth depth. It almost eliminates the tendency of overtraining, which typically grows huge trees.

Pruning is the process to remove the statistically insignificant nodes from bottom up after the tree is built. Because the BDT method always grows small

trees (tree depth about 2 to 6), the pruning is not needed, or only applied when the depth > 3 . The cost-complexity pruning is used in this analysis. To decide which subtrees grown from the node t in the tree T should be pruned from T , the parameter is defined:

$$g(t) = \frac{R(t) - R(T)}{f(t) - 1}, \quad (3.8)$$

where $R(t)$ is the error rate of this node and $R(T)$ is the sum of error rate of the leaves in the tree T . The function $f(t)$ is the number of leaves of the subtree grown from the node t , so $f(t) - 1$ means the number of the pruned nodes (the leaves is not counted) when we want to prune this subtree. This parameter considers the discrimination power of the subtree. The algorithm starts from the whole tree T_0 . The subtree t_0 with the minimum $g(t_0)$ is pruned and we get the pruned tree T_1 . This process is executed iteratively until $g(t_n)$ is larger than the threshold which is set by `PruneStrength` in the program.

3.2 Photon identification

3.2.1 Di-photon vertex identification

The specific identification for the vertex of $H \rightarrow \gamma\gamma$ is developed for the discovery of $H \rightarrow \gamma\gamma$ channel in the LHC Run I. [49] and retuned for the LHC Run II. The vertex choice for $H \rightarrow \gamma\gamma$ decay has a impact on the di-photon mass resolution. For example, the di-photon mass resolution as a function of the distance between the truth vertex and selected vertex is shown in Fig. 3.3. Because photons are not detected by the tracker, there is no tracks to reconstruct the vertex of $H \rightarrow \gamma\gamma$ decay. Moreover, the homogeneous ECAL of the CMS has no longitudinal space resolution. To choose the correct vertex, a BDT algorithm is built to estimate the probability of the right vertex of the di-photon. It exploits the tracks recoiling and also the tracks of electrons from photon conversions.

For no converted events, three discriminating variables are defined:

$$\begin{aligned} sumpt2 &= \sum_i |\vec{p}_T^i|^2 \\ ptbal &= - \sum_i (\vec{p}_T^i \cdot \frac{\vec{p}_T^{\gamma\gamma}}{|\vec{p}_T^{\gamma\gamma}|}) \\ ptasym &= (|\sum_i \vec{p}_T^i| - \vec{p}_T^{\gamma\gamma}) / (|\sum_i \vec{p}_T^i| + \vec{p}_T^{\gamma\gamma}), \end{aligned} \quad (3.9)$$

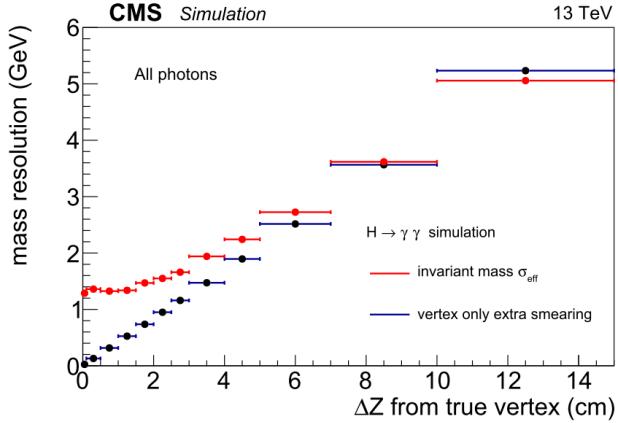


FIGURE 3.3: The impact of the vertex reconstruction on the di-photon invariant mass resolution in gluon-gluon fusion $H \rightarrow \gamma\gamma$ simulation. The additional energy smearing is applied in blue as a function of the Z difference. The mass resolution is calculated as effective sigma σ_{eff} , which is defined as the narrowest region which contains 68% statistics of the full distribution.

where \vec{p}_T^i is the momentum of the i-th track associated with a given vertex and $\vec{p}_T^{\gamma\gamma}$ is the momentum of the photon pair. Their distributions are shown in Fig. 3.4, which shows good discriminating power from the tracks recoiling the photon pair. These three variables are combined and built a first vertex identification MVA by BDT. The vertex with best MVA score is selected to be $H \rightarrow \gamma\gamma$ vertex.

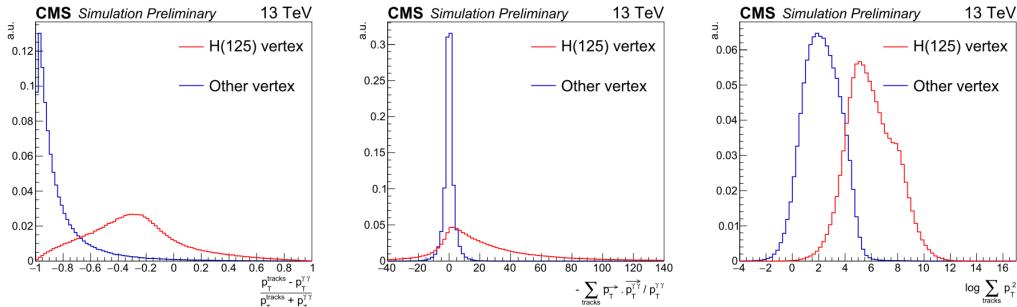


FIGURE 3.4: The variables of track recoiling for the right vertices matched to MC truth vertices (red) and for the wrong vertices failing the matching.

3.2.2 Photon identification

The goal of the photon identification is to distinguish the prompt photon through the non-prompt photon coming from the neutral mesons (π^0) with high momentum decaying into two photons, merged and detected as one photon. In the ECAL barrel, to separate of the two photons from π^0 , the transverse momentum p_T of π^0 should smaller than 15 GeV, where the distance between two photon

is the same as the crystal size. Additionally, the photons converted into e^+e^- in the tracker material are separated by the 3.8 T magnetic field, which lead the wilder shower in ϕ direction. Therefore, the photon isolation and shower shape variables are relied heavily in the identification.

In the first step, the photons should be checked to be "conversion-safe electron veto" to reject the electrons. It requires that there is no charged-particle track with hit in the inner layer of the pixel detector matched the conversion vertex pointing to the clusters of photon candidates. After passing the veto, 99% (97%) photons survive with 5% (20%) electrons in the ECAL barrel (endcap).

The multivariate photon identification is developed in the LHC Run I [50] and updated in Run II. It is trained by the EM-enriched $\gamma + \text{jets}$ simulated samples with BDT algorithm. The prompt photons matched to the photons in the generator level are used as signal, and the others are used as background. The p_T and η_{SC} distribution are weighted to the same between the signal and background to be kinematics independent. The classifiers for ECAL barrel and ECAL endcap are separated. The input variables are listed in Tab. 3.1. The detailed TMVA modified options are as follows:

```
BoostType=Grad:Shrinkage=0.1:MinNodeSize=5%
NTrees=2000:MaxDepth=6:nCuts=2000
PruneMethod=CostComplexity:PruneStrength=5
```

The BDT values of two working points are shown in Tab. 3.2 with their corresponding efficiency.

3.3 Photon energy regression

To have better energy resolution in the photon energy measurement, the photon energy regression is developed. [50] This regression is used to recover the effects from the local containment of the photon shower, such as the energy losses in the gaps and cracks of the ECAL, and the global containment, such as the converted photon showers and the pile-up. The training is performed for ECAL barrel and endcap separately by the double photon-gun simulated samples with about 5 million events. They are trained by the semi-parametric likelihood technique, which combines the parametric likelihood fit on the target variable with the double-sided Crystal-Ball function (see Eq. 5.1) and the non-parametric fit by BDT technique. The source code can be found at Ref. [51]. The input variables rely on the energy ratios in different selected crystals, H/E (the ratio between the hadronic energy and electromagnetic energy), ρ , number of primary vertex,

the shower shape variables and the preshower energies (only for endcap region). Some definitions are already mentioned in Tab. 3.1.

An example of the performance of the regression is given in Fig. 3.5, which show the comparison of the ratios between reconstructed energy and true energy before/after regression.

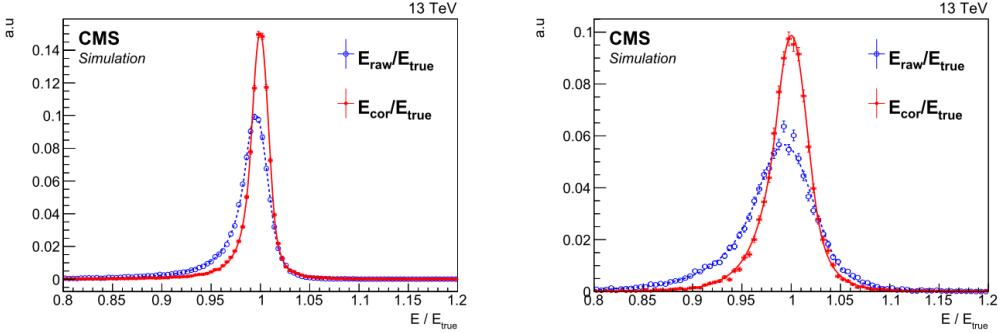


FIGURE 3.5: The distribution of ratio between photon reconstructed energy and true energy in the SM-like $H \rightarrow \gamma\gamma$ samples before (blue line) and after (ref line) regression in the ECAL barrel (left) and endcap (right).

3.4 B-jet identification

Many researches for QCD, Standard Model and new physics rely on the b-quark final states as probes, because their production is perturbatively calculable. [52] In experiment, extraction b jets from other light jets, which are originated from gluons, up and down quarks, is the key-point to reject the background. The b-tag algorithm exploits the specific properties of b hadrons: the long lifetime (~ 1.5 ps) and large mass (~ 5 GeV). The long lifetime leads to have a large impact parameter (IP) and form a secondary vertex (SV) displaced from primary vertex (PV), like Fig. 3.6. In addition, the b hadronization process has high track multiplicity, hard fragmentation function and high semiletonic branching ratio. Due to the hard fragmentation function, the b hadron in the jets carries large momentum from the original b quark.

In CMS, the well-designed track system is able to identify b jets. The b-tagging algorithm used in this analysis is Combined Secondary Vertex Version 2 (CSVv2) [53], which combines the information of displaced tracks and secondary vertices. The anti- k_T jet algorithm with a distance parameter $R = 0.4$ is used to reconstruct the jets. [54] Tracks associated to the jets passing the criteria, which is shown in Tab. 3.3, are selected to derive a series of track-based variables. If there is no selected track, the negative value is sent to the algorithm.

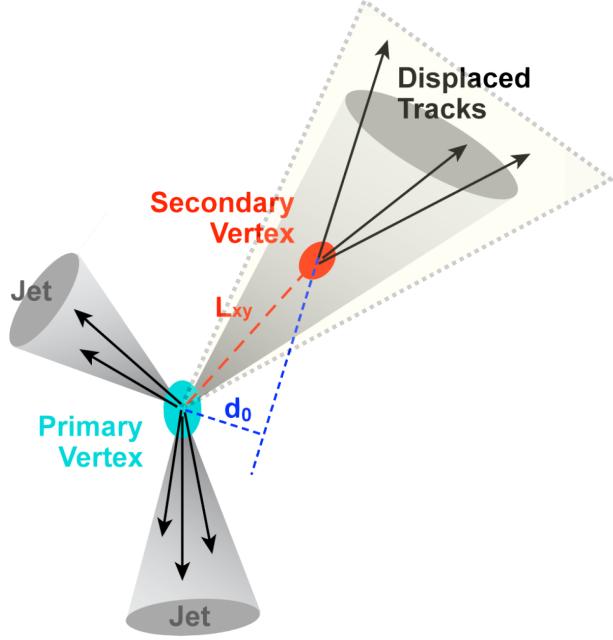


FIGURE 3.6: The sketch of the secondary vertex with large impact parameter d_0 and 2D flight distance from the primary vertex L_{xy} .

In version 2, Inclusive Vertex Finder algorithm is used to reconstruct the secondary vertices. [55] To increase efficiency, this algorithm is independent from the jet algorithm. All reconstructed tracks in the collision event passing a looser requirements are input. The tracks are required to have at least 8 hits at the tracker, p_T greater than 0.8 GeV and the longitudinal impact parameter (IP) smaller than 0.3 cm. The tracks with large impact parameter (at least $50 \mu m$) and impact parameter significance¹ (at least 1.2) are referred to as seeds. Other tracks near the seed are selected to form a cluster by the distance and angle between them. The clusters are fitted by the adaptive vertex fitter algorithm to identify the vertices. The vertices are merged when they share 70% of the tracks and have small distance significance (less than 2). At this stage, both the primary vertex and the secondary vertex are found. One track can be included by the PV and the SV simultaneously. The tracks are removed from the SV when they have less than 1 hit in the pixel detector or when they are more compatible with the PV. If there are still at least two tracks existing, the vertex is refitted. The vertex is removed when they share 20% of the tracks and have small distance significance (less than 10). The remaining SV candidates are selected by the criteria:

- the fraction of shared tracks between the secondary vertex candidate and the primary vertex should not exceed 79% ;

¹the impact parameter significance is defined as the impact parameter divided by its uncertainty. The same definition is used for the 2D (3D) distance.

- the angular distance between the jet axis and the secondary flight direction ΔR is required to be smaller than 0.3 ;
- the 2D flight distance significance should be at least 2 ;

There are three categories of SV reconstruction: the event contains real SVs, pseudo vertex or no vertex. In the case of pseudo vertex, there is no real secondary vertex reconstructed. The jet contains at least two tracks which are not in the K^0 mass window of 50 MeV and signed impact parameter significance exceeding 2. In the case of no vertex, only the information about the displaced track is employed. Finally, the track-based variables and secondary vertex variables are combined by MVA technique. A special tuned MVA based on CMSSW [46] for b-tagging are developed to evaluate the discriminator. The multi-layer perception technique, which is so-called neural network, is used. The tools and the details can be found at [56]. The summary of input variables can be found in Tab. 3.4.

Three kinds of samples with different selections are studied to realize the performance of algorithm in the data and the Monte-Carlo (MC) simulation: the multijet sample, the muon-enriched sample and the dilepton $t\bar{t}$ sample. For example, the combined b-tag discriminator in muon-enriched sample is shown in Fig. 3.7. [57] The good agreement between the data and MC is observed. In addition, the efficiency and the misidentification probability between data and MC is studied to determine the working points (WPs) and derive the scaling factors (SFs). The three WPs corresponding to the misidentification rate are shown in Tab. 3.5. It provides a flexible b jet selection in different signal purity levels. The SFs depending on p_T , η and the b-tagging score are applied to the MC simulation to correct the simulation predictions.

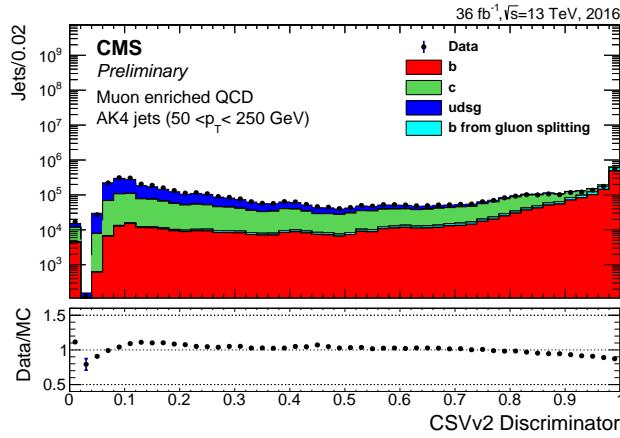


FIGURE 3.7: The distribution of CSVv2 discriminator in the muon-enriched sample compared with full 2016 data.

TABLE 3.1: The input variables and their definition for photon MVA identification.

Type	Variable	Definition
Shower shape	R_9 ($E_{3\times 3}/E_{SC}$)	The energy ratio between the energy sum of the 3 by 3 crystals surrounding the seed crystal of the super cluster and the energy sum of the super cluster.
	$S4$ ($E_{2\times 2}/E_{5\times 5}$)	The energy ratio between the energy sum of the 2 by 2 crystals and the the energy sum of the 5 by 5 crystals surrounding the seed crystal of the super cluster.
	$\sigma_{i\eta i\eta}$	Lateral extension of the electromagnetic shower measured in terms of crystal cells.
	η_{SC} width	The logarithmic energy weighted standard deviation of single crystal η within the super cluster.
	ϕ_{SC} width	The logarithmic energy weighted standard deviation of single crystal ϕ within the super cluster.
	$cov_{i\eta i\eta}$	The covariance of the single crystal η and ϕ in terms of crystal cells within the 5x5 crystals centered on the super cluster seed crystal.
	ES σ_{RR} (endcap only)	The standard deviation of the shower spread in the x and y planes of the preshower detector.
Isolation	Photon isolation	Transverse energy sum associated with all particles identified as photons by the particle-flow algorithm falling inside a cone size $\Delta R=0.3$ around the photon candidate direction.
	Charged isolation	Transverse energy sum associate with all particles identified as charged hadrons by the particle-flow algorithm falling inside a cone size $\Delta R=0.3$ around the photon candidate direction. Measured with respect to the selected vertex.
	Charged isolation (worst vertex)	Same as charged isolation but measured with respect to the worst vertex.
Other	η_{SC}	The η in the detector-based coordinate.
	ϕ_{SC}	The ϕ in the detector-based coordinate.
	Raw E_{SC}	The energy sum of super cluster without correction.
	ρ	The energy median density per unit area in the event.
	E_{ES}/E_{Raw} (endcap only)	The ratio between the sum of the energy deposits in the preshower detector and the energy of photon without correction.

TABLE 3.2: The working points of photon MVA identification.

Category	90% efficiency working point (WP90)	80% efficiency working point (WP80)
ECAL Barrel	>0.20	>0.68
ECAL Endcap	>0.20	>0.60

TABLE 3.3: The selection of tracks for b-tagging

Track variable	Requirement
number of pixel hit	≥ 1
track transverse impact parameter $ IP_{2D} $	$<0.2 \text{ cm}$
track transverse momentum p_T	$>1 \text{ GeV}/c$
track normalized χ^2	<5
track longitudinal impact parameter $ IP_z $	$<17 \text{ cm}$
track distance to jet axis	$<0.07 \text{ cm}$
track decay length	$<5 \text{ cm}$

TABLE 3.4: The input variables for CSVv2 training.

Variable	Reco Vertex	Pseudo Vertex	No Vertex
the vertex category	0	1	2
2D flight distance significance	used	n/a	n/a
vertex mass	used	used	n/a
number of tracks at the vertex	used	used	n/a
ratio of the energy carried by tracks at the vertex with respect to all tracks in the jet	used	used	n/a
the pseudo-rapidity of the tracks at the vertex with respect to the jet axis	used	used	available
2D IP significance of the first track	used	used	available
3D signed IP significances for all tracks in the jet	used	used	used
number of tracks in the jet	used	available	available
ΔR between the secondary vertex flight direction and the jet axis	used	available	n/a
number of secondary vertices	used	n/a	n/a

TABLE 3.5: The working points of CSVv2 b-tag and corresponding selection efficiency in $t\bar{t}$ MC samples.

Working Point	CSVv2 Threshold	Misidentification rate for light jets	Efficiency
Loose	0.5426	10%	$\approx 83\%$
Medium	0.8484	1%	$\approx 69\%$
Tight	0.9535	0.1%	$\approx 49\%$

3.5 B-jet energy regression

The reconstruction of Higgs decaying into $b\bar{b}$ final state is a basic component for this analysis. The semiletonic decay of b hadrons makes the b jet energy resolution worse, due to the undetectable neutrinos. To improve the signal sensitivity in this analysis, the di-jet invariant mass resolution should be optimized. The specific b jet energy regression for $HH \rightarrow b\bar{b}\gamma\gamma$ analysis is implemented by MVA techniques with BDT algorithm. The regression provides the correction function for b jet energy that is applied after the standard CMS jet energy correction. It is trained by the jet kinematics, constituents, and also event-wide information in MC samples. The validation via real data is also checked. The details are introduced in this section.

3.5.1 Training samples

To avoid biases to the signal, the regression is trained on $X \rightarrow HH \rightarrow b\bar{b}b\bar{b}$ MC sample. It combines the samples with all mass points of the resonant m_X , which are listed in Tab. 3.6. Input variables to the training include jet kinematics, energy deposited in the HCAL and vertex information (total of 15 variables), similar to the regression performed for $H \rightarrow b\bar{b}$ analysis. [58] In addition to those, we include in the training the two variables related to missing transverse energy of the event (MET) and the distance between the two jets, $\Delta R(j, j)$. A summary of all input variables is shown in Tab. 3.7.

The target for the regression training is p_T^{gen}/p_T^{reco} , where the generated level jet labeled as “gen” and the reconstructed jet labeled as “reco”. Standard gen-jet collection in CMS does not include the neutrinos, so we add them manually from gen-particle collection, using ΔR cone of 0.4 for matching.

Adding neutrinos to the gen-jet brings the energy of the jet closer to the energy of the original b-quark, which is illustrated in Fig. 3.8. The relative p_T difference between b-quark and gen-jet is peak at zero and has narrower width after adding neutrinos. Due to the limit of the jet algorithm with ΔR cone of 0.4, the jet can’t be exactly the same as the b-quark in the generator level. Fig. 3.9 shows additional distributions, also comparing the gen-jets with and without neutrino additions: invariant mass of the Higgs boson candidates and the m_{jjjj} . Events from all mass samples are combined, which explains the shape of m_{jjjj} distribution. From these figures one can see the effects on the mass resolution of adding neutrinos to gen-jets.

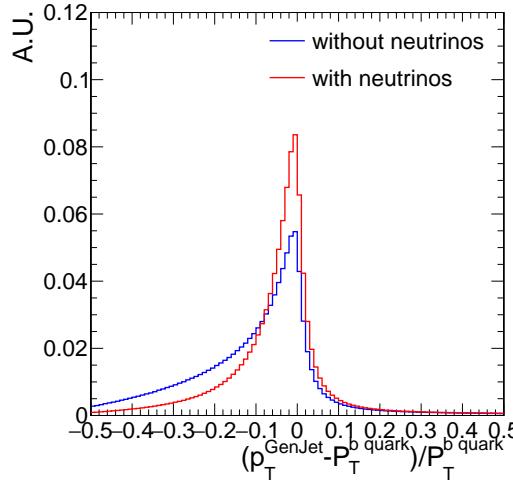


FIGURE 3.8: Relative p_T difference of the b-quark and the corresponding gen-jet, obtained from $HH \rightarrow b\bar{b}b\bar{b}$ samples. Red histogram is for gen-jets containing neutrinos, blue is for jets without neutrinos.

3.5.2 Training method

For the training we select jets that satisfy the following criteria:

- Jets with $p_T > 20\text{GeV}$, $|\eta| < 2.4$;
- Jets matched to the generated level jet within a cone $\Delta R < 0.4$ (this matching is done as part of MiniAOD reconstruction);
- The events with two generated level $H \rightarrow b\bar{b}$ b-quarks pairs;
- Jets matched to a b-quark within a cone $\Delta R < 0.4$;

After selection, about 5 million jets are used for training. The jet with the highest p_T is defined as leading jet and the other one is called trailing jet. We perform six different trainings to check the impact of additional variables:

- Using 15 variables based on $H \rightarrow b\bar{b}$ training, listed in Tab. 3.7. It is denoted as **Baseline** on the figures below.
- Using 15 variables plus MET and $\Delta\phi(\text{Jet}, \text{MET})$.
- Using 15 variables plus MET, $\Delta\phi(\text{Jet}, \text{MET})$, and also $\Delta R(\text{Leading jet}, \text{Trailing jet})$. This training is denoted as **full 15+3var** in the text.
- Using 15 variables as above but in addition, for each pair of jets from the Higgs boson, the training is performed separately for the leading and trailing jets. That is, two BDT functions are derived, one for the leading and one for the trailing jet in the event. This method is denoted as **js** on the figures and in the text.

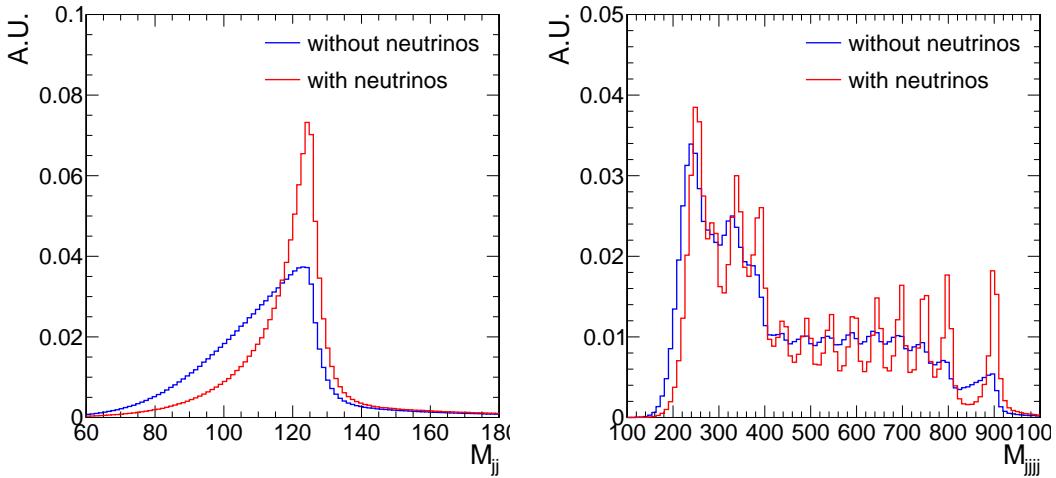


FIGURE 3.9: m_{jj} and m_{jjjj} distributions using jets in all samples with various m_G , obtained from $HH \rightarrow b\bar{b}bb$ MC samples. Red histograms for gen-jets containing neutrinos, blue for jets without neutrinos.

- Using 15 plus MET and $\Delta\phi(Jet, MET)$ variables, and separating the training for leading and trailing jets as above.
- Using 15 plus MET, $\Delta\phi(Jet, MET)$ and ΔR variables, and separating the training for leading and trailing jets as above.

The regression is performed with the TMVA package, using BDT technique with gradient boosting. 500 decision trees with depth 5 are created to estimate the target. The pruning technique is applied. The detailed TMVA modified options are as follows:

```
BoostType=Grad:Shrinkage=0.1:MinNodeSize=1%
NTrees=500:MaxDepth=5:nCuts=500
PruneMethod=CostComplexity:PruneStrength=5
```

Each parameter is already mentioned in Sec. 3.1.

3.5.3 Performance

After the training is done, its performance is checked in the signal samples, $X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$, at all mass points of m_X . The selection is the same as the training samples. All of our trainings are compared with the one done by $H \rightarrow b\bar{b}$ analysis, which is denoted as **Hbb** on the figures. The distribution of relative p_T difference between the gen-jets and the reconstructed jets is shown in Fig. 3.10 for the leading and trailing jets. From distributions such as in Fig. 3.10 we obtain the mean value (μ) for the **scale** and sigma (σ) for the **resolution**. The mean value and sigma are from the fit of Bukin function, which is the convolution of

Gaussian + exponential function, defined in Eq. 3.10.

$$\mathcal{P}(x; x_p, \sigma_p, \xi, \rho_1, \rho_2) = A_p \exp \left[\frac{\xi \sqrt{\xi^2 + 1} (x - x_1) \sqrt{2 \ln 2}}{\sigma_p (\sqrt{\xi^2 + 1} - \xi)^2 \ln(\sqrt{\xi^2 + 1} + \xi)} + \rho \left(\frac{x - x_i}{x_p - x_i} \right)^2 - \ln 2 \right]$$

where $\begin{cases} \rho = \rho_1, x_i = x_1, & \text{for } x < x_1 \\ \rho = \rho_2, x_i = x_2, & \text{for } x \geq x_2 \end{cases}, \quad x_{1,2} = x_p + \sigma_p \sqrt{2 \ln 2} \left(\frac{\xi}{\sqrt{\xi^2 + 1}} \mp 1 \right)$

(3.10)

The scale and resolution of the leading and trailing jets versus their p_T are shown in Fig. 3.11. From this figure we can conclude that adding MET variables into the training improves the resolution significantly. As expected the **15+3var js** training gives the best per-jet resolution across the whole p_T range.

Fig. 3.12 and 3.13 show the response of each input variable for leading and trailing jets. We check the average deviation of p_T^{reco} from p_T^{gen} as a function of each input variable before (in blue) and after (in pink) regression. The $H \rightarrow b\bar{b}$ training (in red) and our best training are compared. After regression, the p_T deviation is reduced to zero and becomes more stable as a function of each input variable. With the **15+3var js** training, the strongly dependence of the three additional variables is removed. For example, without regression, the p_T deviation of the jets with high MET is large because a part of energy of the jet is not detected and is included into MET. After our regression, these jets are corrected to become closer to the gen-jet.

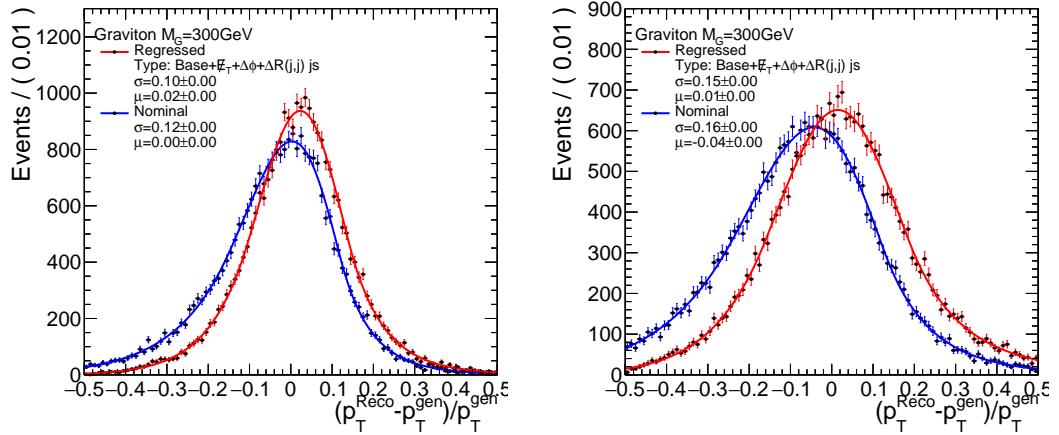


FIGURE 3.10: Relative p_T difference of the reconstructed and generated level jets after regression (red histograms) and without the regression (blue histograms).

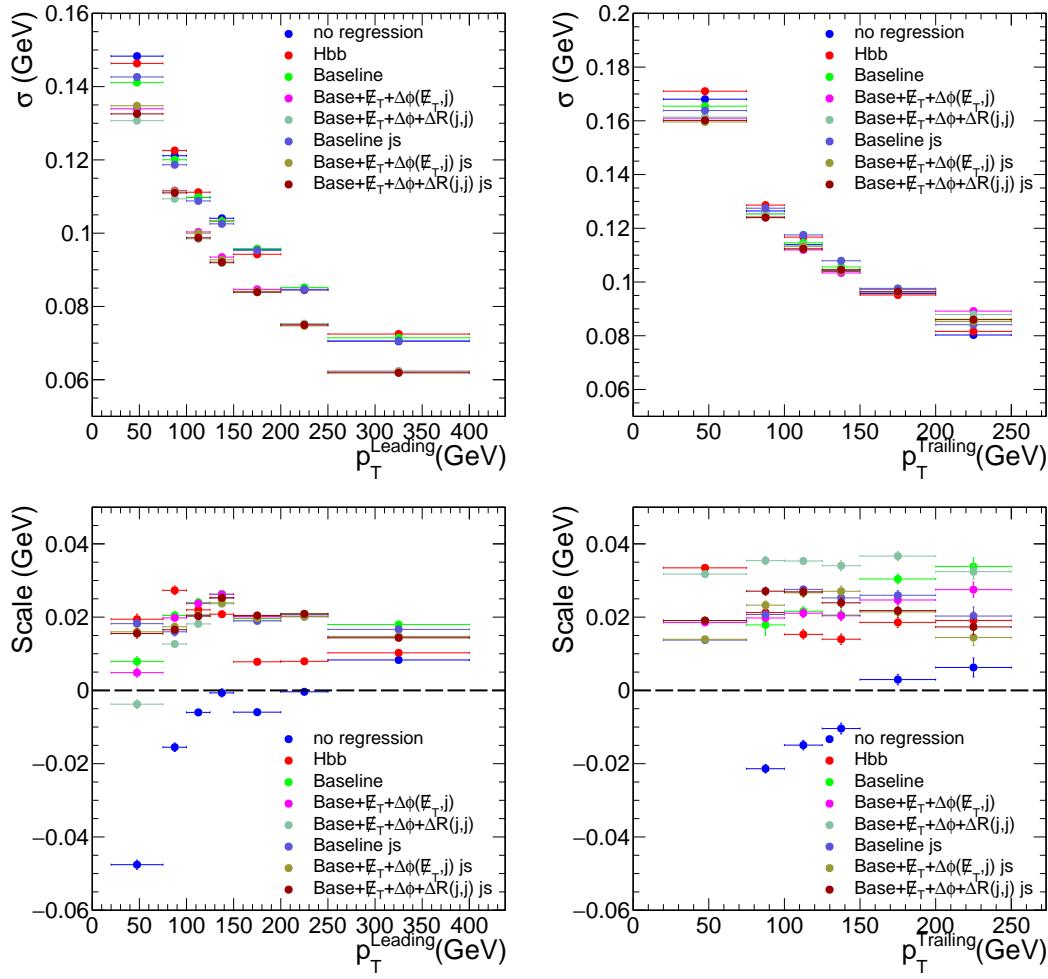


FIGURE 3.11: The resolution of the jet p_T (top) and the scale (bottom), for leading (left) and trailing (right) jets from the signal sample $G \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$.

TABLE 3.6: The samples used for b jet energy regression training. The label “M-X” indicates the resonant mass points.

Samples	Number of events
GluGluToRadionToHHTo4B_M-260_narrow_13TeV	299996
GluGluToRadionToHHTo4B_M-270_narrow_13TeV	299997
GluGluToRadionToHHTo4B_M-300_narrow_13TeV	294392
GluGluToRadionToHHTo4B_M-350_narrow_13TeV	299991
GluGluToRadionToHHTo4B_M-400_narrow_13TeV	299991
GluGluToRadionToHHTo4B_M-450_narrow_13TeV	99193
GluGluToRadionToHHTo4B_M-500_narrow_13TeV	99995
GluGluToRadionToHHTo4B_M-550_narrow_13TeV	99997
GluGluToRadionToHHTo4B_M-600_narrow_13TeV	99596
GluGluToRadionToHHTo4B_M-650_narrow_13TeV	99995
GluGluToRadionToHHTo4B_M-750_narrow_13TeV	99785
GluGluToRadionToHHTo4B_M-800_narrow_13TeV	99993
GluGluToRadionToHHTo4B_M-900_narrow_13TeV	99980
GluGluToBulkGravitonToHHTo4B_M-260_narrow_13TeV	299997
GluGluToBulkGravitonToHHTo4B_M-270_narrow_13TeV	299989
GluGluToBulkGravitonToHHTo4B_M-300_narrow_13TeV	292395
GluGluToBulkGravitonToHHTo4B_M-350_narrow_13TeV	299994
GluGluToBulkGravitonToHHTo4B_M-400_narrow_13TeV	299794
GluGluToBulkGravitonToHHTo4B_M-450_narrow_13TeV	99995
GluGluToBulkGravitonToHHTo4B_M-500_narrow_13TeV	99993
GluGluToBulkGravitonToHHTo4B_M-550_narrow_13TeV	99994
GluGluToBulkGravitonToHHTo4B_M-600_narrow_13TeV	99991
GluGluToBulkGravitonToHHTo4B_M-650_narrow_13TeV	99188
GluGluToBulkGravitonToHHTo4B_M-750_narrow_13TeV	99183
GluGluToBulkGravitonToHHTo4B_M-800_narrow_13TeV	99980
GluGluToBulkGravitonToHHTo4B_M-900_narrow_13TeV	99981

TABLE 3.7: Input variables used in BDT regression. Upper part lists the variables that are also used in $H \rightarrow b\bar{b}$ analysis, and the lower part lists additional variables.

Input variables as in $H \rightarrow b\bar{b}$ analysis	
nPVs	Number of primary vertex of the event
Jet p_T	Jet transverse momentum
Jet η	Jet pseudorapidity
Jet m_T	Jet transverse mass
Jet lead track p_T	Transverse momentum of the leading track in the jet
Jet lepton p_T	Transverse momentum of the leading lepton candidate in the jet
Jet relative lepton p_T	The projection of 4-momentum of the leading lepton candidate in the jet on the jet axis.
$\Delta R(\text{Jet, lepton})$	The distance in the η - ϕ space of the lepton in the jet and the jet
Jet neHEF	Neutral hadron energy fraction of the jet
Jet neEmF	Photon energy fraction of the jet
Jet vtx p_T	p_T of the jet secondary vertex
Jet vtx Mass	Invariant mass of the jet secondary vertex
Jet vtx 3d L	The 3-d flight length of the jet secondary vertex
Jet vtx Ntrk	Track multiplicity of the reconstructed secondary vertex
Jet vtx 3d eL	Error on the 3-d flight length of the jet secondary vertex
Additional variables for this analysis	
MET	Missing energy of the event
$\Delta\phi(\text{Jet}, MET)$	The difference ϕ of MET and the jet
$\Delta R(\text{Jet}_1, \text{Jet}_2)$	The distance in the η - ϕ space of two jets

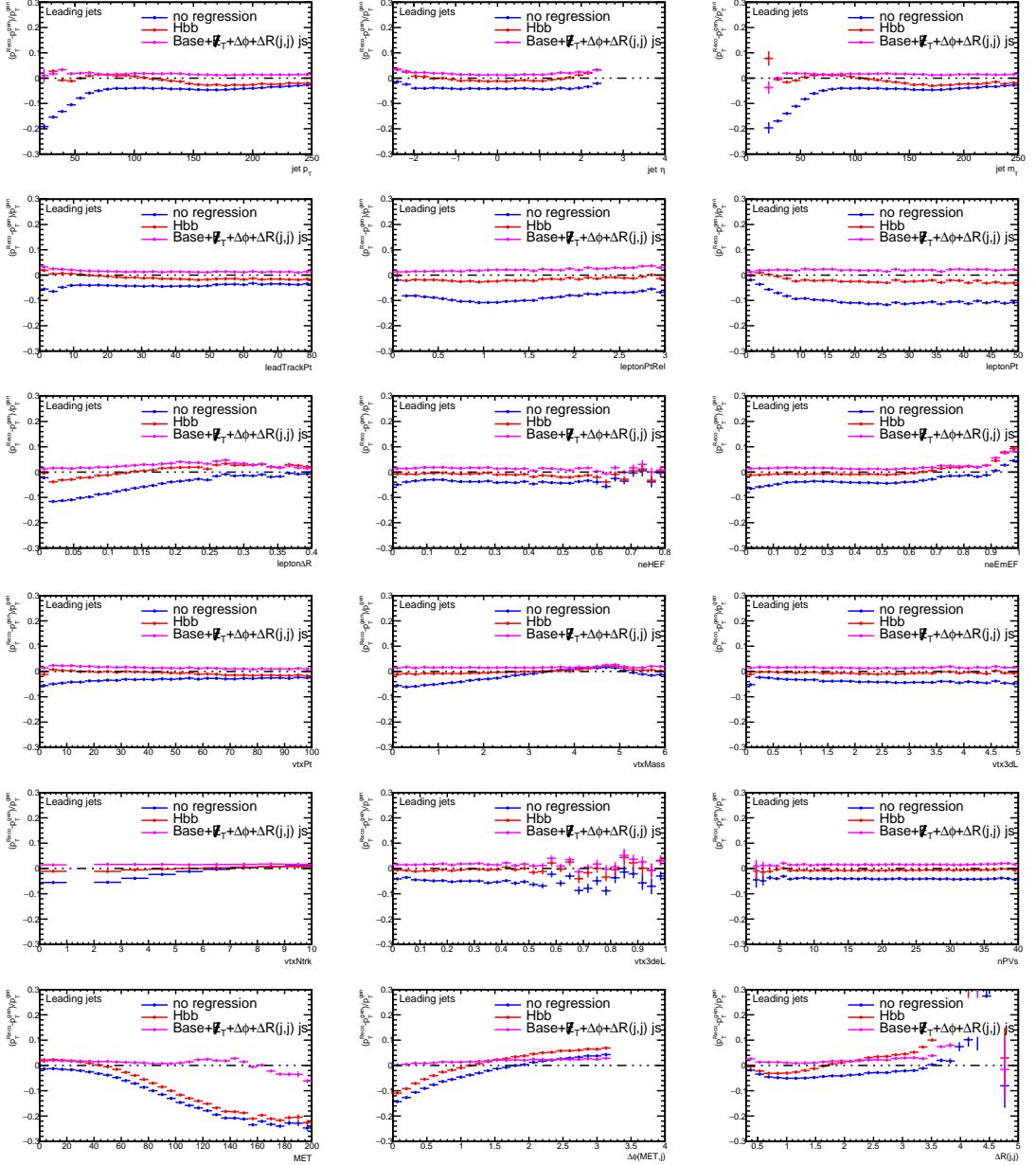


FIGURE 3.12: The correlation between $(p_T^{reco} - p_T^{gen}) / p_T^{gen}$ and input variables in leading jets, $X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ samples.

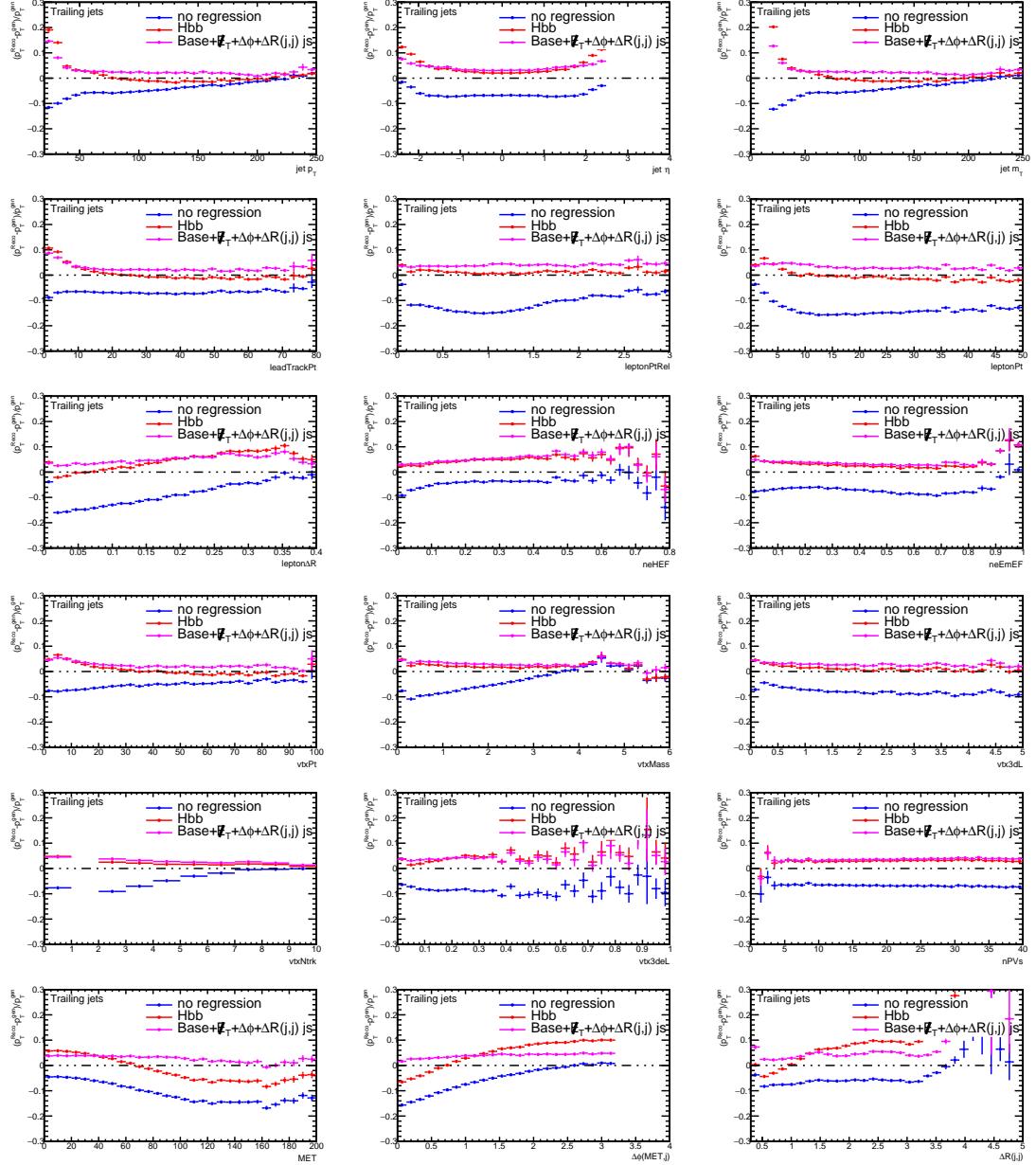


FIGURE 3.13: The correlation between $(p_T^{reco} - p_T^{gen}) / p_T^{gen}$ and input variables in trailing jets, $X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ samples.

Mass resolution

To obtain better signal modeling, the purpose of the regression is to improve the Higgs boson mass resolution from the $H \rightarrow b\bar{b}$ decay in $X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ signal. The distributions together with the fit by Bukin function are shown in Fig. 3.14, where the reconstructed mass is shown before and after applying the **full 15+3var js** regression training, for $m_G = 300\text{GeV}$ and $m_G = 900\text{GeV}$.

After the fit to the corresponding function is done, we obtain the mean and the width parameter of the fit. The width and mean (both in GeV) are shown on Fig. 3.15 versus the mass of the Graviton particle. The regression also tested in the non-resonant samples, which is shown in Fig. 3.16. From these figures we arrive at the same conclusion as for single-jet plots of Fig. 3.11: the MET variables improve the resolution and the **full 15+3var js** training gives the best mass resolution. The training with 15 variables gives similar results to the **Hbb** training. In all trainings the scale does not match the nominal value of the Higgs boson mass. This is expected because the jets (both at reco and gen levels) do not contain the whole energy of the Higgs boson decay.

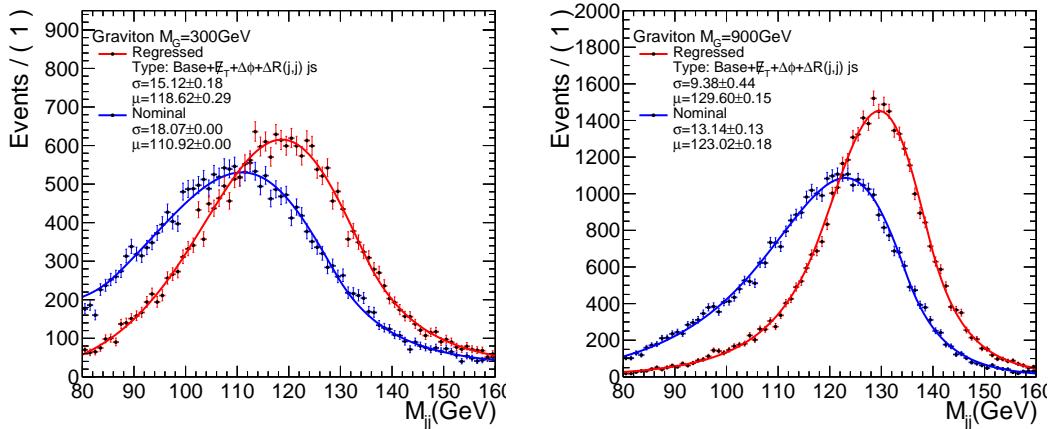


FIGURE 3.14: M_{jj} distributions from the reco-jets before and after the **full 15 variables with js** regression for $m_G = 300\text{GeV}$ signal sample (left) and $m_G = 900\text{GeV}$ signal sample (right).

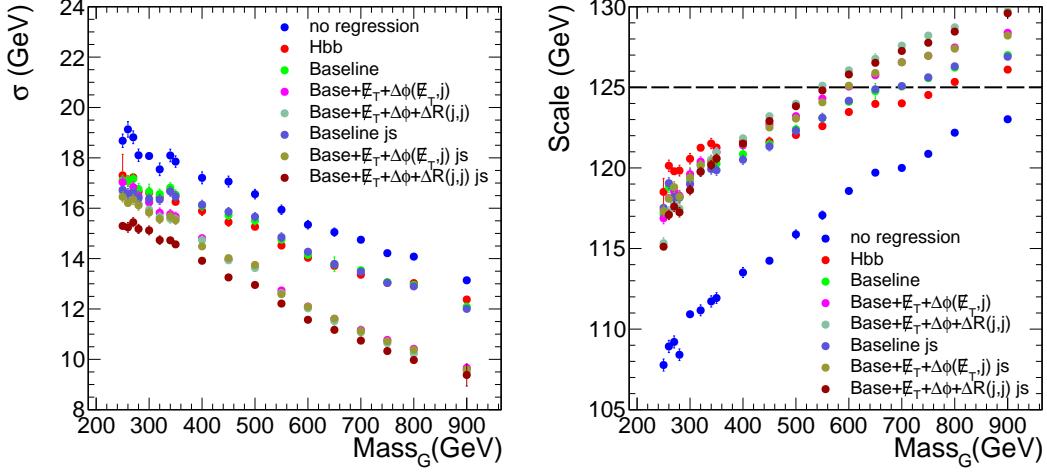


FIGURE 3.15: Performance plot comparing different regression trainings in resonant signal samples.

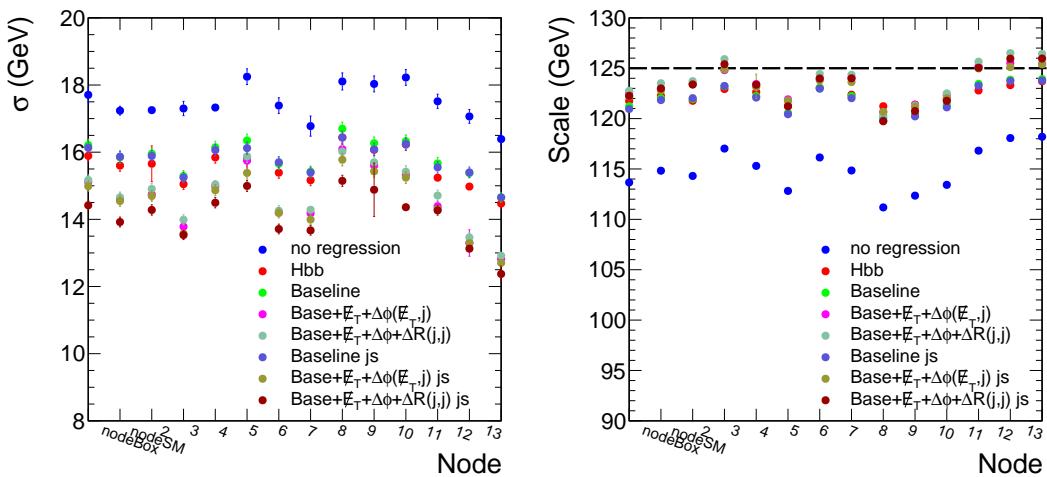


FIGURE 3.16: Performance plot from M_{jj} distributions comparing different regression trainings in non-resonant signal samples.

3.5.4 The improvement from the three additional input variables and the training with leading and trailing jets separately

From previous sections, we know that the improvement from MET, $\Delta\phi(Jet, MET)$ is significant. For the $\Delta R(Jet_1, Jet_2)$, the training with separating leading and trailing jets is needed.

Fig. 3.17 shows the correlation between $\Delta R(\text{Leading jet}, \text{Trailing jet})$ and other input variables. Some variables such as $\Delta R(\text{Jet, lepton})$, invariant mass of the jet secondary vertex and track multiplicity of the secondary vertex have interesting shape and are different between leading and trailing jets. With $\Delta R(Jet_1, Jet_2)$ and the training with leading and trailing jets separately, this correlation is taken into account. If we put the leading and trailing jets together, the correlation becomes smaller. Just adding $\Delta R(Jet_1, Jet_2)$ does not give us the better result.

Although $\Delta R(Jet_1, Jet_2)$ is highly jet- p_T dependent, Fig. 3.18 shows that the special shapes of these variables do not appear as a function of p_T of the jets. This special correlation just happens for $\Delta R(Jet_1, Jet_2)$. In addition, the other input variables have similar correlation with jet- p_T between leading and trailing jets. Without $\Delta R(Jet_1, Jet_2)$, the difference between leading and trailing jets depends on the p_T of jets in one event. That is why training leading and trailing jets separately without $\Delta R(Jet_1, Jet_2)$ doesn't lead the result to be better compared to the non-separate training.

Adding $\Delta R(Jet_1, Jet_2)$ and training with leading and trailing jets separately gives BDT the strong information of the correlation between other input variables and $\Delta R(Jet_1, Jet_2)$. That's where the improvement comes.

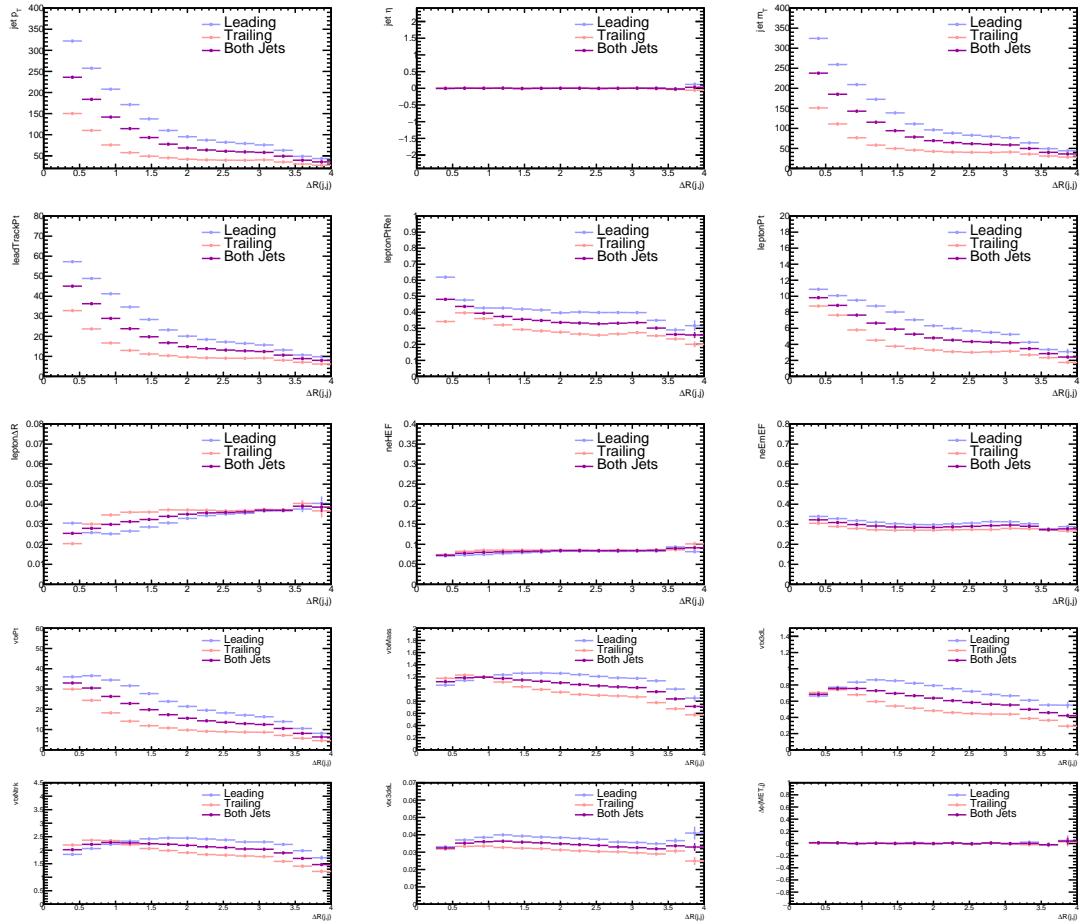


FIGURE 3.17: The correlation between $\Delta R(\text{Leading jet, Trailing jet})$ and other input variables comparing leading jets (blue), trailing jets (red) and both jets (purple) in $G \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ samples. The y-axis is the mean value of the input variables in different $\Delta R(\text{Leading jet, Trailing jet})$ bins.

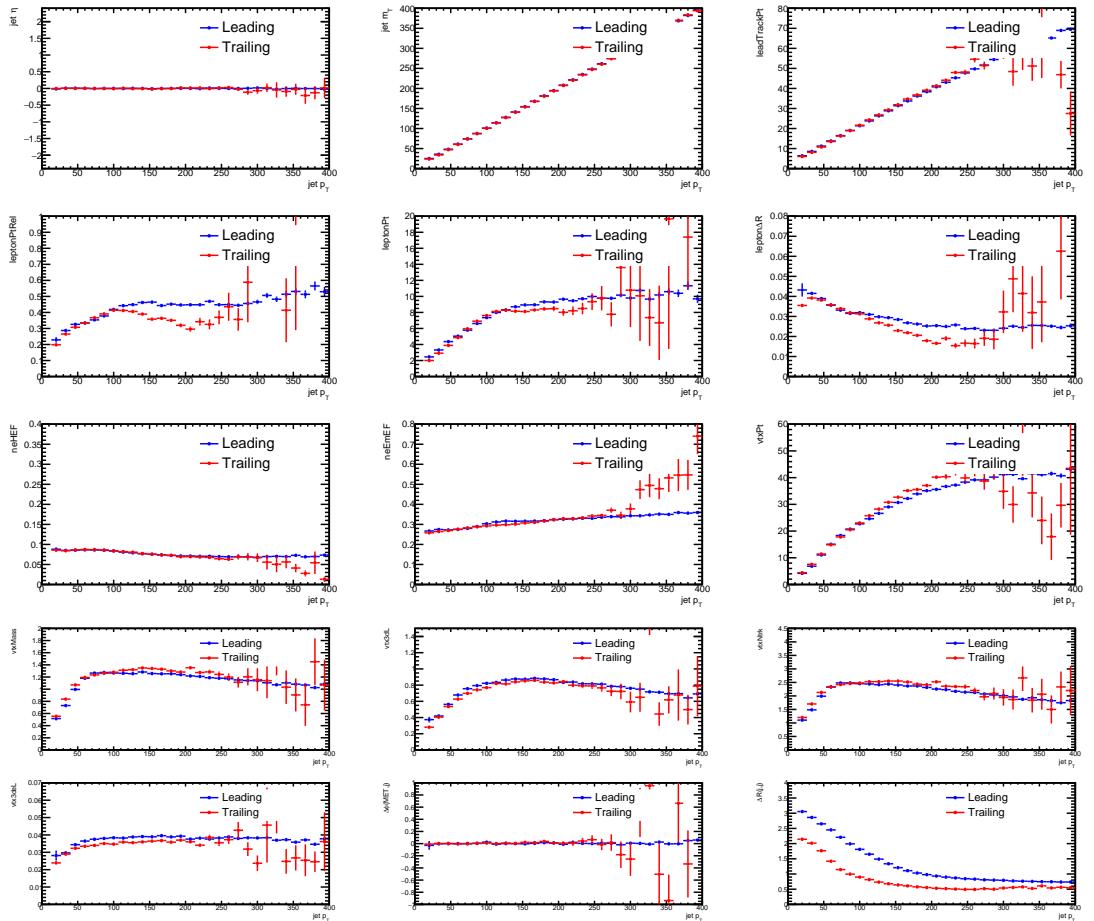


FIGURE 3.18: The correlation between p_T and other input variables comparing leading jets, trailing jets in $G \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ samples.

3.5.5 Data validation

In order to validate the developed regression in data we select events with $Z \rightarrow \ell\ell$ decay which also contain two b-tagged jets. [59] It is assumed that a di-jet is recoiled against Z boson, and therefore the $p_T(jj)$ must balance the $p_T(\ell\ell)$. This check was done both in muon and electron channels of Z boson decay, analyzing DoubleMuon and DoubleEG full 2016 datasets, which collect the events passing di-muon and di-electron/di-photon triggers correspondingly.

TABLE 3.8: The lepton selection corresponding to electron and muon channel

Electron channel from DoubleElectron dataset
pass trigger HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ
pass electron MVA ID WP90
Muon channel from DoubleMuon dataset
pass trigger HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ
or HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ
pass Tight muon ID and Loose particle flow isolation

The leptons should pass the requirement corresponding to the electron and muon channel, which is shown in Tab. 3.8. Further event selection requirements in both channels are as follows:

- $p_T(\ell_1) > 25 \text{ GeV}$, $p_T(\ell_2) > 15 \text{ GeV}$;
- $p_T(\ell\ell) > 50 \text{ GeV}$;
- $75 < m_{\ell\ell} < 105 \text{ GeV}$;
- $\Delta R_{\ell,jet}^{min} > 0.4$;
- jet $p_T > 20 \text{ GeV}$, $|\eta| < 2.4$;
- Both two jets pass Loose particle flow identification [60], medium pile up jet identification [61], tagged as b-jets with CSVv2 Medium WP ;

Fig. 3.19 shows the input variables in data and MC samples. Good agreement between data and MC is observed. Fig. 3.20 shows the mentioned p_T -balance distributions, $p_T^{jj}/p_T^{\ell\ell}$. The data is compared to the MC predictions. It can be seen that before any regression is applied the peak of the ratio distribution is below one. With the regression applied the peak moves to 1 for both our **full 15 variables js** training (center) and the one from **Hbb** (right). The response of the regression is the same in data and MC. This indicates that the regression does indeed brings the p_T of the b-jets closer to their true values.

Similarly, Fig. 3.21 shows the mass distributions, m_{jj} , before and after regression. These figures indicate that m_{jj} is not distorted in any bad way, and no artificial peaks are created. This ensures us that the background distributions in our analysis signal region will not be distorted either.

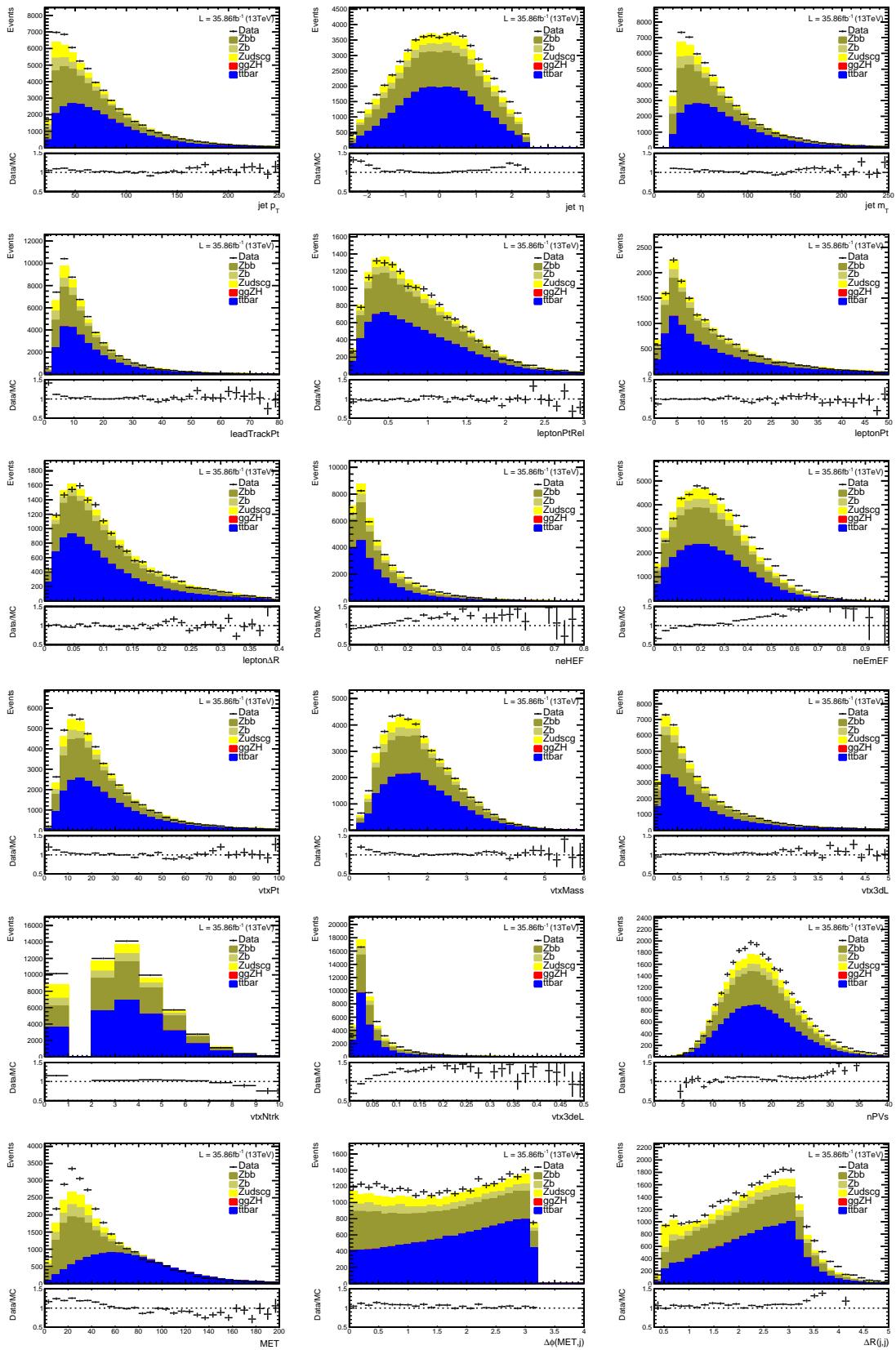


FIGURE 3.19: The distribution of input variables comparing data and MC. Electron and muon channels are combined.

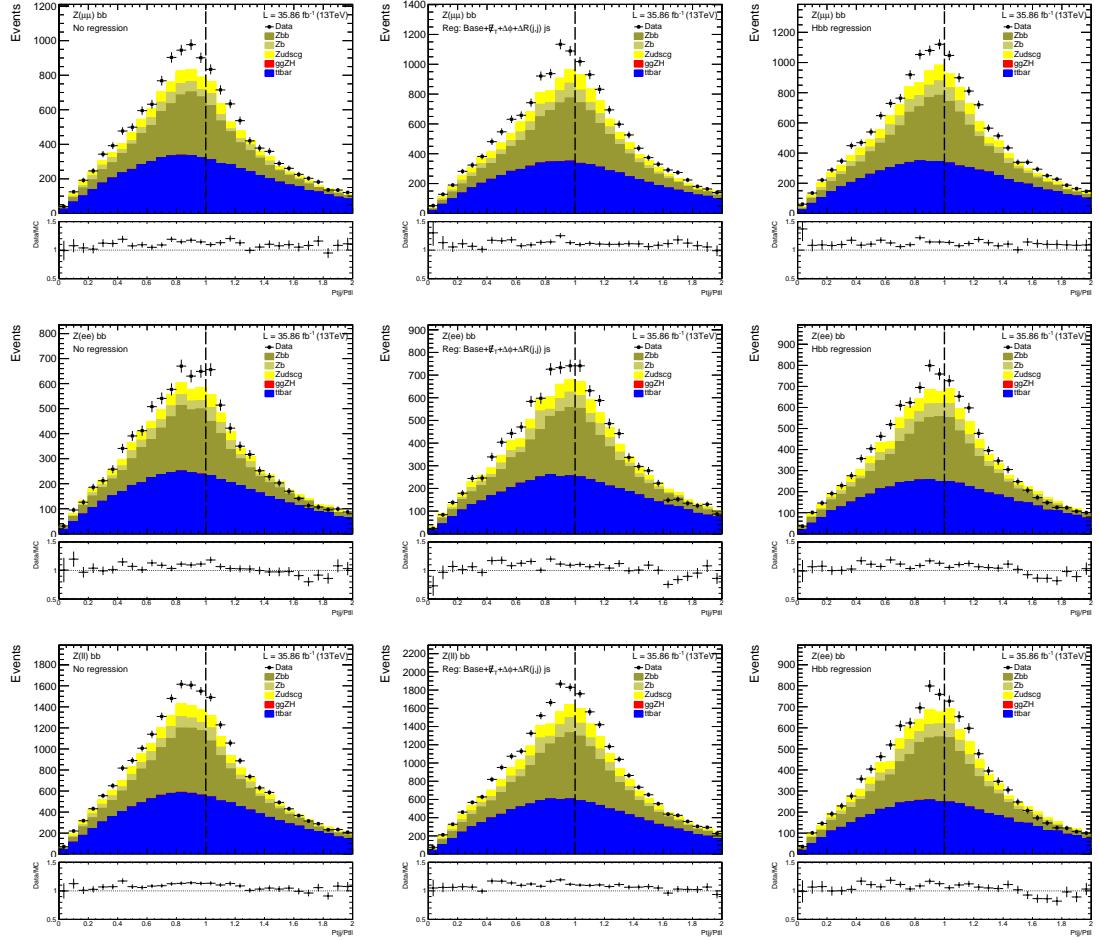


FIGURE 3.20: Pt balance (ratio) of the di-jet and di-lepton. On the left are the distributions without regression, in the center - using **full 15+3var js** training and on the right - using **Hbb** regression. Top plots are for muon channel, middle for electron channel and bottom is the combination (sum) of the two.

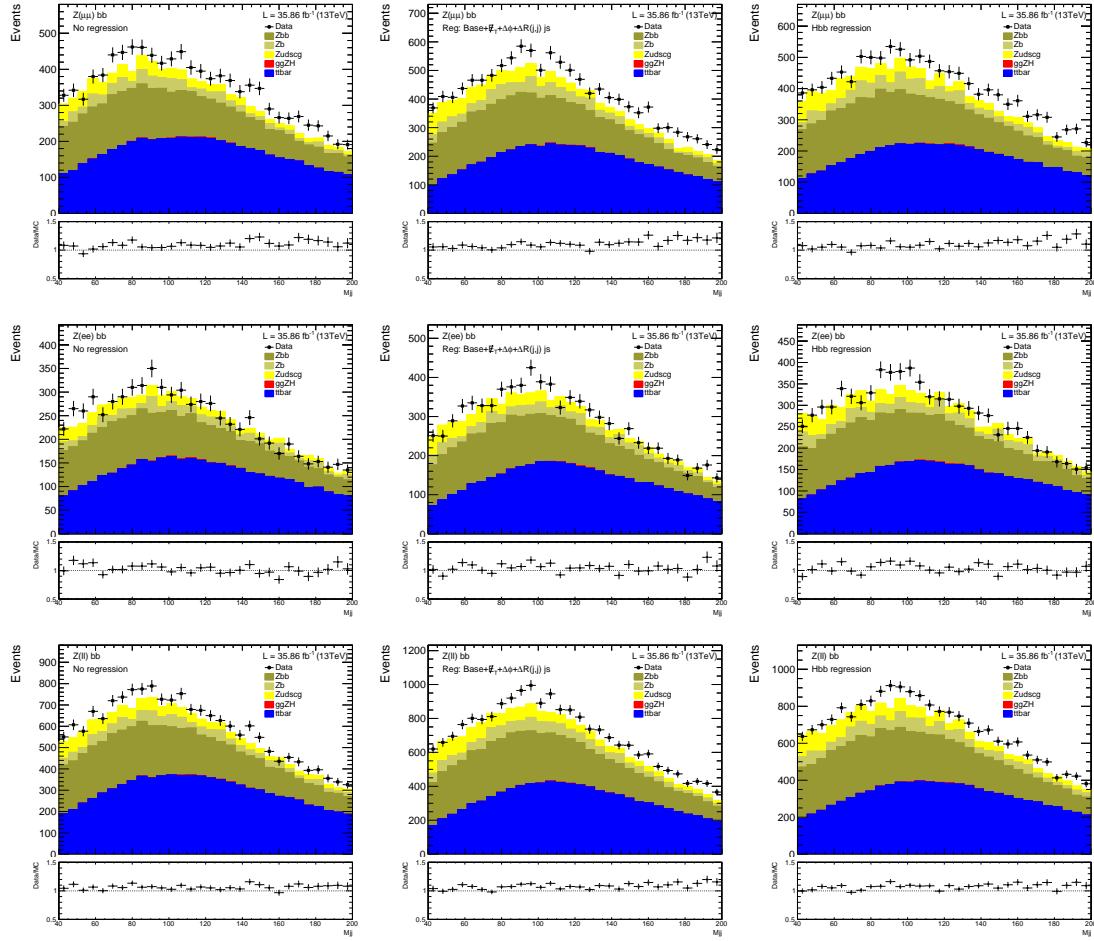


FIGURE 3.21: Distributions of the m_{jj} . On the left are plots with no regression, in the center - using **full 15+3var js** training and on the right - using **Hbb** regression. Top plots for muon channel, middle for electron channel and bottom is the combination (sum) of the two.

3.5.6 Impact on the results

In the SM-like HH search, the expected limit combining all categories (see next section) set on the cross-section of production is improved from 1.77 fb to 1.6 fb (10%). The signal-over-background ratios in different categories are shown in Tab. 3.9. Fig. 3.22 shows the expected limits for the non-resonance search with the recommended points in the anomalous couplings space. The improvements are from 10% to 13%, which depend on the benchmark points.

TABLE 3.9: The signal-over-background ratios in the search of SM-like HH production in different categories. The categories are labeled as low mass region (LM), high mass region (HM), high signal purity (HPC) and low signal purity (MPC). The details about categorization are described in Sec. 3.6.

	LM HPC	LM MPC	HM HPC	HM MPC
Nominal	0.2556	0.4425	3.9813	3.5879
Regressed	0.2675	0.5319	4.2661	4.1079
Improvement	4.65%	20.02%	7.15%	14.49%

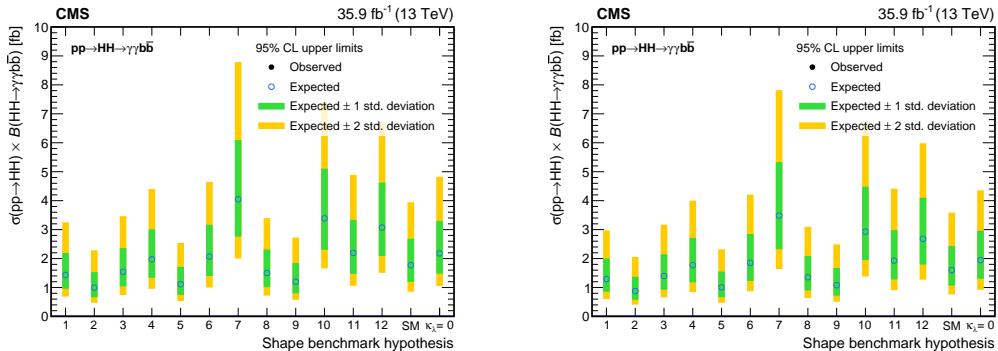


FIGURE 3.22: The upper limits at 95% confidence level on the cross-section of non-resonance HH production in benchmark points before (left) and after (right) regression.

3.6 $HH \rightarrow b\bar{b}\gamma\gamma$ categorization MVA

There are some kinematics variables found able to classify the background-like and signal-like events. These variables are combined to one BDT classification discriminator. To increase the sensitivity, the events are divided into two signal purity categories depending on the ratio of signal to background.

The signal hypotheses are all signal samples described in Sec. 4.1.2 and Sec. 4.1.3. Both classifiers take the data control region as background, which is described in Sec. 4.2.1. Furthermore, we split the training into high mass region and low

mass region by \tilde{M}_X , the definition is mentioned in Sec. 4.2.3. In the resonance search, the high mass region is $\tilde{M}_X > 350$ GeV. This region is optimized for the SM di-Higgs production because there is a smooth peak around 400 GeV in \tilde{M}_X spectrum of SM-like process, shown in Fig. 3.23. The low mass region $\tilde{M}_X < 350$ GeV is also searched, since the anomalous coupling constant may distort the distribution and form a peak in low mass region. In the resonance search, the events are segmented by $\tilde{M}_X = 600$ GeV. The low mass region contains higher background yields and lower signal efficiency, and the high mass region is inverse. This cut value is also checked to ensure both ensembles have enough statistics for training. The BDT classifiers are trained for resonance and non-resonance searches in difference mass region separately. There are total four trainings.

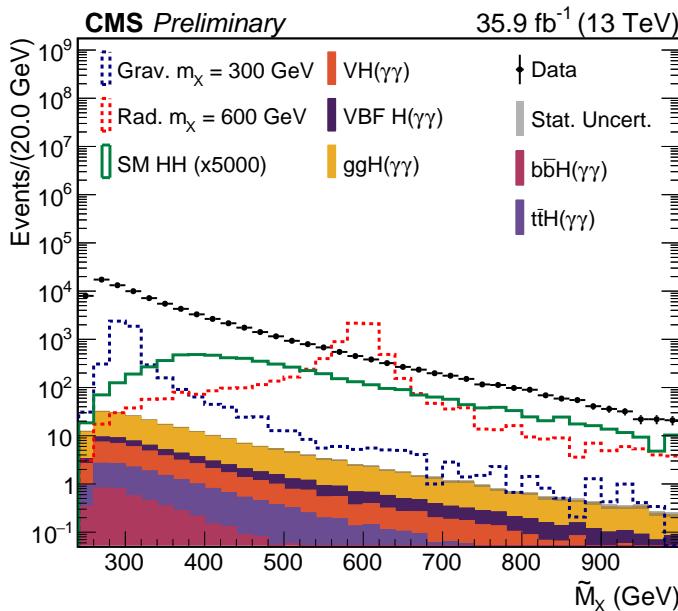


FIGURE 3.23: The \tilde{M}_X spectrum of data and signal/background MC simulation after $HH \rightarrow b\bar{b}\gamma\gamma$ selection, which is described in Ch. 4. The resonant signal is normalized to a cross section of 500 fb and SM-like signal is 5000 times higher than the SM prediction.

The training variables are described in Tab. 3.10, and the distribution of BDT discriminator of non-resonant high mass region is shown in Fig. 3.24. The values of the BDT discriminator used for categorization are optimized to reach the best sensitivity. A summary of categorization is shown in Tab. 3.11. In the non-resonance low mass region, an additional requirement on the b-tag score corresponding to 80% efficiency is applied.

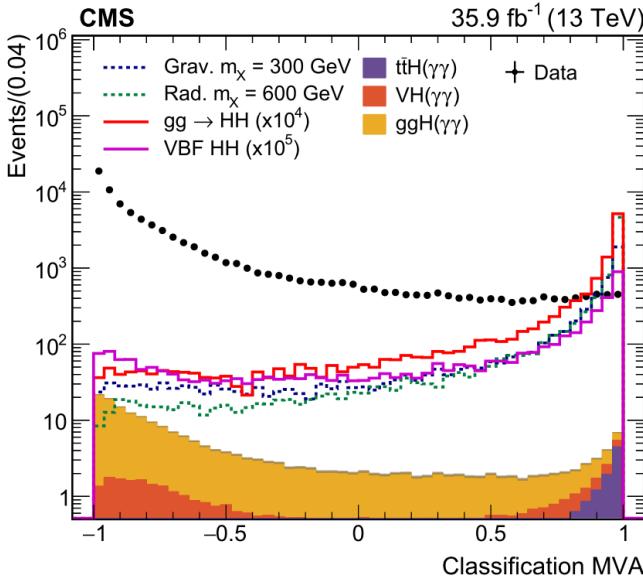


FIGURE 3.24: The categorization MVA output distribution obtained from the non-resonance high mass region training. Data are compared with the SM-like signal samples through gluon-gluon fusion and vector boson fusion (VBF), the resonant signal samples, and the single-Higgs samples after full selections.

TABLE 3.10: The training variables of categorization MVA.

Input variables for categorization MVA	
Jet_1 b-tag score	The b-tag score of leading jet
Jet_2 b-tag score	The b-tag score of trailing jet
$p_{T\gamma\gamma}/p_{THH}$	The transverse momentum ratio between di-photon and di-Higgs boson
$p_{T\gamma\gamma}/p_{Tbb}$	The transverse momentum ratio between di-photon and di-jet
$ \cos(\theta_{CS}^*) $	The angle between the direction of the $H \rightarrow \gamma\gamma$ candidate and the Colin-Sopper reference frame
$ \cos(\theta_{bb}^*) $	The angle between the direction of the $H \rightarrow b\bar{b}$ candidate and the leading jet
$ \cos(\theta_{\gamma\gamma}^*) $	The angle between the direction of the $H \rightarrow \gamma\gamma$ candidate and the leading photon

TABLE 3.11: The summary of the categorization by MVA. The high signal purity category (HPC) and medium signal purity category (MPC) are defined for both non-resonant and resonant analyses

Analysis	Mass region	Categorization MVA
Non-resonant	$\tilde{M}_X > 350 \text{ GeV}$	HPC: MVA > 0.97 MPC: $0.6 < \text{MVA} < 0.97$
	$\tilde{M}_X < 350 \text{ GeV}$	HPC: MVA > 0.985 MPC: $0.6 < \text{MVA} < 0.985$
Resonant	$\tilde{M}_X > 600 \text{ GeV}$	HPC: MVA > 0.5 MPC: $0 < \text{MVA} < 0.5$
	$\tilde{M}_X < 600 \text{ GeV}$	HPC: MVA > 0.96 MPC: $0.7 < \text{MVA} < 0.96$

Chapter 4

Event selection

4.1 Datasets

4.1.1 Data

The data used in this analysis are collected by the CMS in full 2016, corresponding to an integrated luminosity of 35.9 /fb recorded by the CMS detector. The “DoubleEG” dataset is used in this analysis, which is off-line reconstructed in 2017 and collects the events firing the specific di-electron/di-photon triggers. In this analysis, we use di-photon HLT trigger which is the same used in $H \rightarrow \gamma\gamma$ study. [62] This HLT is seeded by the hardware L1 trigger electromagnetic candidates. There should be a single candidate with transverse momentum higher than 25 GeV or a pair of candidates higher than 15 and 10 GeV respectively. The HLT algorithm extends the L1 candidates (L1 seed) from the rectangle by $\Delta\eta \times \Delta\phi = 0.14 \times 0.4$ to be clusters. There should exist the clusters passing the isolation plus calorimeter identification or R_9 selection, which is shown in Tab. 4.1. Furthermore, the events should contain two clusters within $|\eta| < 2.5$ having $E_T > 30$ GeV and $E_T > 18$ GeV, $H/E < 0.12$ and invariant mass of two objects above 90 GeV. The rate of this HLT is 14 Hz during 2016 data taking.

4.1.2 Resonant Monte Carlo signal samples

The signal samples is simulated at leading order (LO) by MADGRAPH5_aMC@NLO 2.3.2 [63, 64] interfaced with LHAPDF6 [65]. The next-to-leading-order (NLO) parton distribution function set PDF4LHC15_NLO_MC [66–70] is used. The events are processed with PYTHIA 8.212 [71] with the tune CUETP8M1 [72] for showering, hadronization, underlying event and pile-up. The the CMS detector is simulated by GEANT 4 [73].

TABLE 4.1: The list of filters used in di-photon high level trigger (HLT).

	$R_9(5 \times 5)$	$\sigma_{in\eta}(5 \times 5)$	ECAL PF cluster isolation	Track isolation
Barrel region $R_9 > 0.85$	>0.5	-	-	-
Barrel region $R_9 \leq 0.85$	>0.5	<0.015	$<6.0 + 0.012E_T$	$<6.0 + 0.002E_T$
Endcap region $R_9 > 0.90$	>0.8	-	-	-
Endcap region $R_9 \leq 0.90$	>0.8	<0.035	$<6.0 + 0.012E_T$	$<6.0 + 0.002E_T$

The Warped Extra Dimension model in bulk RS scenario with narrow-width resonance is chosen as the benchmark model. The new heavy particles both spin-0 Radion and spin-2 Graviton with the resonance masses from 250 GeV to 900 GeV, which are produced through gluon-gluon fusion and decay into two Higgs bosons with the mass $m_H = 125$ GeV, are simulated. The particles are assumed having narrow decay width which is much smaller than the detector resolution and negligible. The used resonance mass points are summarized in Tab. 4.2.

The cross-section for theoretical interpretation can be found at Ref. [16]. The configuration cards of MADGRAPH5_aMC@NLO can be found at Ref. [74].

TABLE 4.2: The resonant signal samples used in this analysis. The label “M-X” indicates the resonant mass points.

Samples	Number of events
GluGluToRadionToHHTo2B2G_M-250_narrow_13TeV	50000
GluGluToRadionToHHTo2B2G_M-260_narrow_13TeV	50000
GluGluToRadionToHHTo2B2G_M-270_narrow_13TeV	50000
GluGluToRadionToHHTo2B2G_M-280_narrow_13TeV	49600
GluGluToRadionToHHTo2B2G_M-300_narrow_13TeV	50000
GluGluToRadionToHHTo2B2G_M-320_narrow_13TeV	49998
GluGluToRadionToHHTo2B2G_M-340_narrow_13TeV	50000
GluGluToRadionToHHTo2B2G_M-350_narrow_13TeV	49999
GluGluToRadionToHHTo2B2G_M-400_narrow_13TeV	49996
GluGluToRadionToHHTo2B2G_M-450_narrow_13TeV	49998
GluGluToRadionToHHTo2B2G_M-500_narrow_13TeV	49997
GluGluToRadionToHHTo2B2G_M-550_narrow_13TeV	49998
GluGluToRadionToHHTo2B2G_M-600_narrow_13TeV	49998
GluGluToRadionToHHTo2B2G_M-650_narrow_13TeV	49999
GluGluToRadionToHHTo2B2G_M-700_narrow_13TeV	49198
GluGluToRadionToHHTo2B2G_M-750_narrow_13TeV	49997
GluGluToRadionToHHTo2B2G_M-800_narrow_13TeV	49792
GluGluToRadionToHHTo2B2G_M-900_narrow_13TeV	49994
GluGluToBulkGravitonToHHTo2B2G_M-250_narrow_13TeV	49799
GluGluToBulkGravitonToHHTo2B2G_M-260_narrow_13TeV	49998
GluGluToBulkGravitonToHHTo2B2G_M-270_narrow_13TeV	48400
GluGluToBulkGravitonToHHTo2B2G_M-280_narrow_13TeV	50000
GluGluToBulkGravitonToHHTo2B2G_M-300_narrow_13TeV	49200
GluGluToBulkGravitonToHHTo2B2G_M-320_narrow_13TeV	50000
GluGluToBulkGravitonToHHTo2B2G_M-340_narrow_13TeV	49998
GluGluToBulkGravitonToHHTo2B2G_M-350_narrow_13TeV	50000
GluGluToBulkGravitonToHHTo2B2G_M-400_narrow_13TeV	49999
GluGluToBulkGravitonToHHTo2B2G_M-450_narrow_13TeV	49999
GluGluToBulkGravitonToHHTo2B2G_M-500_narrow_13TeV	49198
GluGluToBulkGravitonToHHTo2B2G_M-550_narrow_13TeV	49995
GluGluToBulkGravitonToHHTo2B2G_M-600_narrow_13TeV	49998
GluGluToBulkGravitonToHHTo2B2G_M-650_narrow_13TeV	50000
GluGluToBulkGravitonToHHTo2B2G_M-700_narrow_13TeV	49998
GluGluToBulkGravitonToHHTo2B2G_M-750_narrow_13TeV	49999
GluGluToBulkGravitonToHHTo2B2G_M-800_narrow_13TeV	49999
GluGluToBulkGravitonToHHTo2B2G_M-900_narrow_13TeV	49993

4.1.3 Non-resonant Monte Carlo signal samples

The MC generator set used for non-resonance analysis is the same as the resonance signal MC samples. Both SM-like HH production and BSM HH production through gluon-gluon fusion are simulated based on the effective field theory (EFT) Lagrangian, which contains five parameters related to the Higgs boson coupling strength and is already described in Sec. 1.2.3. It is not feasible to generate the simulation samples with all possible combination of the five parameters. The mentioned shape benchmark technique is used to produce MC samples. The events which are simulated by different value of parameters and have similar shape of kinematics distribution are grouped into a cluster. The benchmark points represent these clusters to reduce the complexity of the analysis. The first version of recommended 13 benchmark points are used to generate our signal samples. They are listed in Tab. 4.3, contain the SM-like, box-diagram only and other BSM values. The configuration cards of `MADGRAPH5_aMC@NLO` can also be found at Ref. [74].

To explore the phase space beyond the simulated benchmark points, the event weighting technique is employed to approach other points. [75] It starts from the assumption that the HH production is a $2 \rightarrow 2$ scattering. The two Higgs bosons are produced back-to-back in the azimuthal direction and have the same momentum. The azimuthal angle of the two Higgs boson is isotropic. Consequently, the kinematics can just be determined by two variables: invariant mass of Higgs boson pair m_{HH} and the absolute value of the cosine of the polar angle of one Higgs boson with respect to the beam axis $|cos\theta^*|$. The phase space in these two variables is sliced optimally into 59 bins in m_{hh} and 4 bins in $|cos\theta^*|$.

The event weights are derived from the differential cross-section of each bin, which is extended from the total cross-section described in Ch. 1, Eq. 1.14. It has the same form of Eq. 1.14 with coefficients A_i^j in different bin j . The differential cross-section becomes

$$R_{HH}^j = \frac{\sigma_{HH}}{\sigma_{HH}^{SM}} \frac{Frac^j}{Frac_{SM}^j} = Poly(A_i^j). \quad (4.1)$$

The $Frac^j$ is the fraction of events of the simulated BSM samples in bin j , and $Frac_{SM}^j$ is for the simulated sample of SM-like HH production. The coefficients A^j are extracted by the likelihood fit and scanning the space of the five parameters. This procedure is similar to the procedure to extract the total cross-section coefficients. To obtain the events of the arbitrary point in the space of five parameters, the all simulated events are summed together. The events in bin j are

applied the weight:

$$W_j = R_{HH}^j \frac{\sigma_{HH}^{SM}}{\sigma_{HH}} \frac{Frac_{SM}^j}{C_{norm}}, \quad (4.2)$$

where $C_{norm} = \sum_j R_{HH}^j \frac{\sigma_{HH}^{SM}}{\sigma_{HH}} Frac_{SM}^j$ is the factor to limit the errors from the fit and simulated samples. In the ideal case, C_{norm} is equal to one.

Fig. 4.1 shows the comparison in different kinematics variables between two SM-like sample, one is from MC simulation and the other one is from re-weighting. The good agreement is observed.

TABLE 4.3: The list of non-resonant signal samples and the values of five parameters in 13 clusters. The box-diagram only sample is also produced.

Samples	Number of events	κ_λ	κ_t	c_2	c_g	c_{2g}
Node1 (SM)	49998	1.0	1.0	0.0	0.0	0.0
Node2	49598	7.5	2.5	-0.5	0.0	0.0
Node3	50000	15.0	1.5	-3.0	-0.0816	0.3010
Node4	49996	5.0	2.25	3.0	0.0	0.0
Node5	49999	10.0	1.5	-1.0	-0.0956	0.1240
Node6	49998	1.0	0.5	4.0	-1.0	-0.3780
Node7	49998	2.4	1.25	2.0	-0.2560	-0.1480
Node8	49998	7.5	2.0	0.5	0.0	0.0
Node9	49600	10.0	2.25	2.0	-0.2130	-0.0893
Node10	49799	15.0	0.5	1.0	-0.0743	-0.0668
Node11	49998	-15.0	2.0	6.0	-0.1680	-0.5180
Node12	49996	2.4	2.25	2.0	-0.0616	-0.1200
Node13	49997	-15.0	1.25	6.0	-0.0467	-0.5150
NodeBox	49999	0.0	1.0	0.0	0.0	0.0

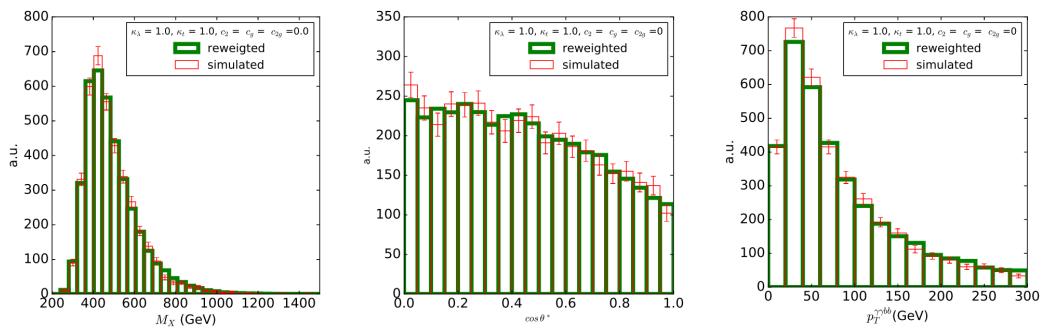


FIGURE 4.1: The comparison between SM-like simulated and re-weighting samples in three kinematics variables: the four body invariant mass, the angle between two Higgs bosons and the four body transverse momentum. [75]

4.1.4 Background simulation

The background is contributed mainly by $n\gamma + \text{jets}$ events in the phase space of our analysis. They mainly come from the non-resonance QCD production of two photons and two jets. Meanwhile, the jets misidentified as photons with multi-jets and other photons can be a source. It is challenging to simulate the QCD events accurately due to large effects in high order. [76] In this analysis, these contributions are estimated from the data.

On the other hand, the contribution of single Higgs production in non-resonance search is not negligible. The SM-like single Higgs boson production with two additional jets, where the Higgs boson decays into two photons, through gluon-gluon fusion (ggH), vector boson fusion (VBF H) and associated production with $t\bar{t}$ ($t\bar{t}\text{H}$), $b\bar{b}$ ($b\bar{b}\text{H}$), and vector boson (VH) is considered as a source of background. They are simulated by `MADGRAPH5_aMC@NLO` 2.2.2 for VH , 2.3.3 for $b\bar{b}\text{H}$, and `POWHEG` 2.0 [77–80] at NLO for ggH , VBF H and $t\bar{t}\text{H}$. They are also interfaced with `PYTHIA` 8.212 and `GEANT 4` as the signal samples.

4.2 Physical objects

All of the final-state physics objects are reconstructed by the particle flow algorithm. [81] This algorithm links all of the reconstructed ingredients together from the sub-detectors in the CMS and identify the particle type. Each reconstructed object is mutually exclusion, where the ingredients are only used once for one objects. The identification is benefited by the structure of the CMS detector, as shown in Fig. 4.2. Starting from the interaction region, the charged particles are detected by the tracker first. The electromagnetic particles are absorbed by the ECAL, and the hadronic particles are obstructed by the HCAL. Finally, the escaping muons cause the hit in the muon system. The particles can be identified according to the linked hits in each sub-detector. The correlations between each sub-detector are taken into account to identify the properties of particles.

4.2.1 The $H \rightarrow \gamma\gamma$ candidate

Trigger mimic pre-selection

To achieve a good agreement between data and simulation, a pre-selection which is tighter than the trigger selection is applied in order to mimic the trigger. This pre-selection is developed for the $H \rightarrow \gamma\gamma$ search [62] and is employed in

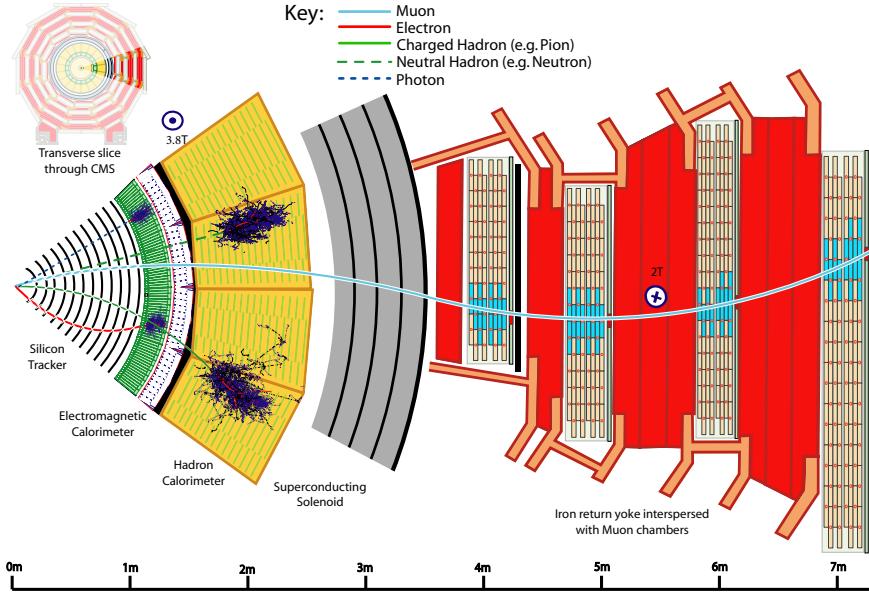


FIGURE 4.2: A sketch of the specific particle interactions in a transverse slice of the CMS detector. [81]

$HH \rightarrow b\bar{b}\gamma\gamma$ analysis. The scale factors and uncertainties related to this selection are also applied in this analysis. The selection is summarized in Tab. 4.4.

TABLE 4.4: The trigger mimic selection for the photon trigger.

Barrel region $ \eta < 1.4442$	$R_9(5 \times 5) > 0.85$	-
	$R_9(5 \times 5) \leq 0.85$	$R_9(5 \times 5) > 0.5$ $\text{pfPhoIso} < 4 \text{ GeV}$ $\sigma_{in\eta}(5 \times 5) < 0.015$
Endcap region $1.556 < \eta < 2.5$	$R_9(5 \times 5) > 0.9$	-
	$R_9(5 \times 5) \leq 0.9$	$R_9(5 \times 5) > 0.8$ $\text{pfPhoIso} < 4 \text{ GeV}$ $\sigma_{in\eta}(5 \times 5) < 0.035$
Pass electron-veto to reject the electron; $R_9(5 \times 5) > 0.8$ or $\text{chargeIso} < 20 \text{ GeV}$ or $\text{chargeIso}/E_T < 0.3$; $H/E < 0.08$;		

$H \rightarrow \gamma\gamma$ selection

The photon pair selection is based on the kinematics variables and the photon MVA identification which is described in Sec 3.2. The selection is as follows:

- Leading photon divided by the di-photon invariant mass $p_T^{\gamma 1}/M_{\gamma\gamma} > 1/3$, trailing photon $p_T^{\gamma 2}/M_{\gamma\gamma} > 1/4$;

- $100 < M_{\gamma\gamma} < 180 \text{ GeV};$
- Both photons pass photon MVA identification with 90 % efficiency working point;

The trigger efficiency of the events passing the selection above is 100 %.

In addition, the control region is obtained by selecting the photon pair with only one photon passing the photon MVA identification. The other procedures of selection are the same as the normal selection. This control region is used by the training of categorization MVA, which is described in Sec. 3.6, and also the background modeling validation, which is mentioned in Sec. 5.2.1.

4.2.2 The $H \rightarrow b\bar{b}$ candidate

The jets in LHC Run-II at CMS are reconstructed by clustering the particle-flow objects with the anti- k_t algorithm using cone size $\Delta R = 0.4$. [54] The cone size is smaller than LHC Run I analysis, which is $\Delta R = 0.5$, due to the higher luminosity and number of pile-up. This change leads jets to be worse resolution, because there is less energy clustered. The energy of jets are corrected by the flavor blind standard regression. [82] In addition, the specific energy regression for b-jet is applied, which is described in Sec. 3.5. The jets are selected by the criteria as follows:

- Pass loose particle flow jet identification [60];
- $p_T > 25 \text{ GeV};$
- $|\eta| < 2.4$, which is within the tracker of the CMS and can be tagged as b-jet;
- Outside the cone of selected photons with $\Delta R(jet, \gamma) > 0.4$;
- $70 < M_{jj} < 90$;
- The jet pair which has the highest sum of b-tag score is selected.

4.2.3 The di-Higgs system

After the di-photon and di-jet selection, the selected objects are combined to become a di-Higgs boson candidate. To minimize the dependence on the resolution of photons and jets, the effective four body invariant mass [83] is defined as:

$$\tilde{M}_X = M_{jj\gamma\gamma} - M_{jj} - M_{\gamma\gamma} + 250 \text{ GeV}. \quad (4.3)$$

This variable scales the di-photon and di-jet invariant mass to 125 GeV. Fig. 4.3 shows the comparison between the distributions of the four body invariant mass and the effective mass \tilde{M}_X . The impact on the low mass resonance is very huge.

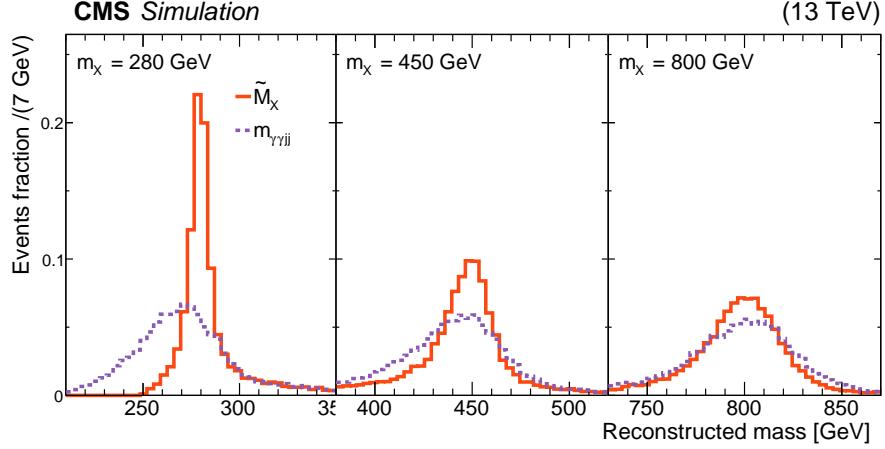


FIGURE 4.3: Comparison between \tilde{M}_X (red) and $m_{jj\gamma\gamma}$ (purple) distribution of spin-2 resonance with different masses. All of events pass the di-photon and di-jet selections.

The \tilde{M}_X is used to restrict the signal phase space for the signal modeling in the resonance analysis. The size of effective mass window around \tilde{M}_X peak is optimized to reach the best sensitivity for each of searched resonance point. The mass window is chosen to cover 60 % signal shape in \tilde{M}_X with the minimum range. The exactly mass window range is a function of the resonance mass, which is fit by the polynomial and shown in Fig. 4.4.

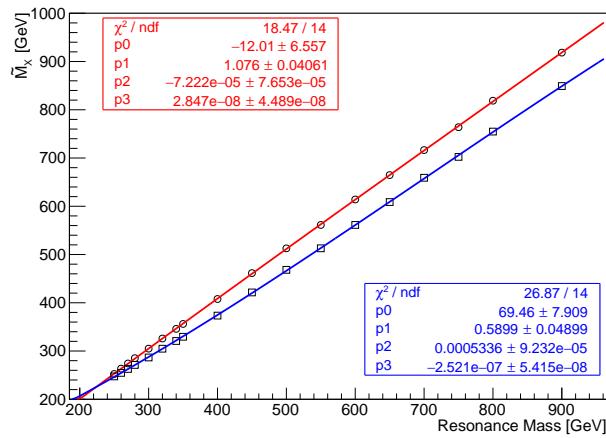


FIGURE 4.4: The upper bounds and lower bounds of the window in \tilde{M}_X as a function of searched resonance mass. The values of bound are fit by a 3rd polynomial.

4.2.4 Signal selection efficiency

The signal step-by-step selection acceptance times efficiency as a function of resonance mass is shown in Fig. 4.5, which includes the online (trigger mimic) selection, the di-photon selection, the di-jet selection and the MVA categorization in different regions. The final efficiencies range from approximately 20% (low mass) to 50% (high mass) for both spin-0 and spin-2 hypotheses. In the non-resonance case, the efficiency is 30% for the SM-like HH production, with 25% in the high mass region and 5% in the low mass region.

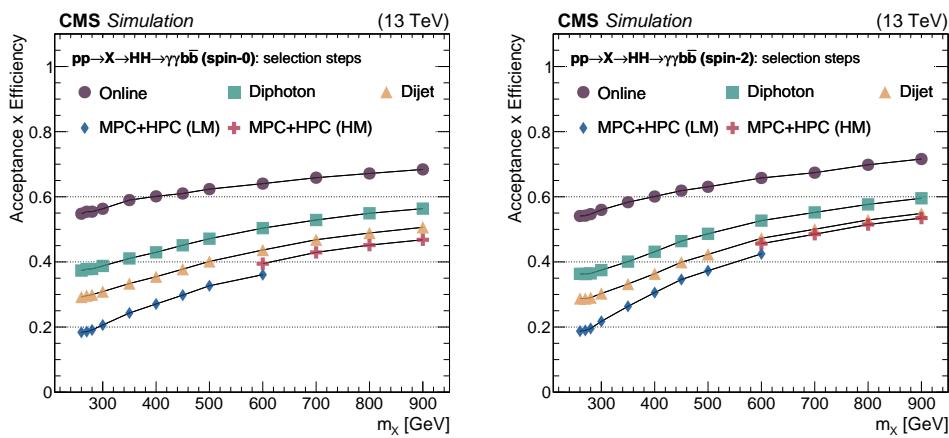


FIGURE 4.5: Consecutive selection efficiencies step-by-step for two resonance hypotheses: spin-0 (left) and spin-2 (right).

Chapter 5

Modeling

An accurate modeling is an important part to estimate the signal strength and excluded limits. In this analysis, the modeling is derived via both the invariant mass of the di-photon ($m_{\gamma\gamma}$) and the di-jet (m_{jj}) candidates. The 2D fit on the $m_{\gamma\gamma} \times m_{jj}$ plane is employed to model the expected signal and background shapes. The function of fit is the product of the two independent functions for $m_{\gamma\gamma}$ and m_{jj} with assumption that these two variables are uncorrelated. This assumption is checked for both signal and background modelings. The procedure of modeling is introduced in this chapter.

5.1 Signal modeling

The function used to model both resonant and non-resonant signal shapes is double-sided Crystal Ball function. The double-sided Crystal Ball function is composed of a Gaussian core and two independent exponential tails:

$$f(x; \mu, \sigma, \alpha_L, p_L, \alpha_R, p_R) = N \cdot \begin{cases} A_L \cdot (B_L - \frac{x-\mu}{\sigma})^{-p_L}, & \text{for } \frac{x-\mu}{\sigma} > -\alpha_L \\ A_R \cdot (B_R + \frac{x-\mu}{\sigma})^{-p_R}, & \text{for } \frac{x-\mu}{\sigma} > \alpha_R \\ e^{\frac{(x-\mu)^2}{\sigma^2}}, & \text{for } \frac{x-\mu}{\sigma} < -\alpha_L \text{ and } \frac{x-\mu}{\sigma} > \alpha_R \end{cases}, \quad (5.1)$$

where the A_L, A_R, B_L, B_R constants are defined by:

$$A_k = \left(\frac{p_k}{|\alpha_k|} \right)^{p_k} \cdot e^{-\frac{\alpha_k^2}{2}}, \quad (5.2)$$

$$B_k = \frac{p_k}{|\alpha_k|} - |\alpha_k|, \quad (5.3)$$

where k is either L or R . The additional exponentials provide the good description for the events in the lower and higher energy tail which may be mismeasured, and the Gaussian core extracts the peak position and resolution. This function is able to model the mass distribution well both for the high resolution $m_{\gamma\gamma}$ and low resolution m_{jj} in whole analysis.

The parameters of function are obtained by fits to the simulated signal samples after selection and categorization. Each fit is independently done for each analyzed point. Some results of fit projected to one-dimension are demonstrated in Fig. 5.1, Fig. 5.2, and Fig. 5.3. In the figures, the effective sigma σ_{eff} is extracted by a half of the width of the narrowest region containing 68.3% of the signal shape. The mean value μ is obtained from the fitted Gaussian core of the Crystal Ball function.

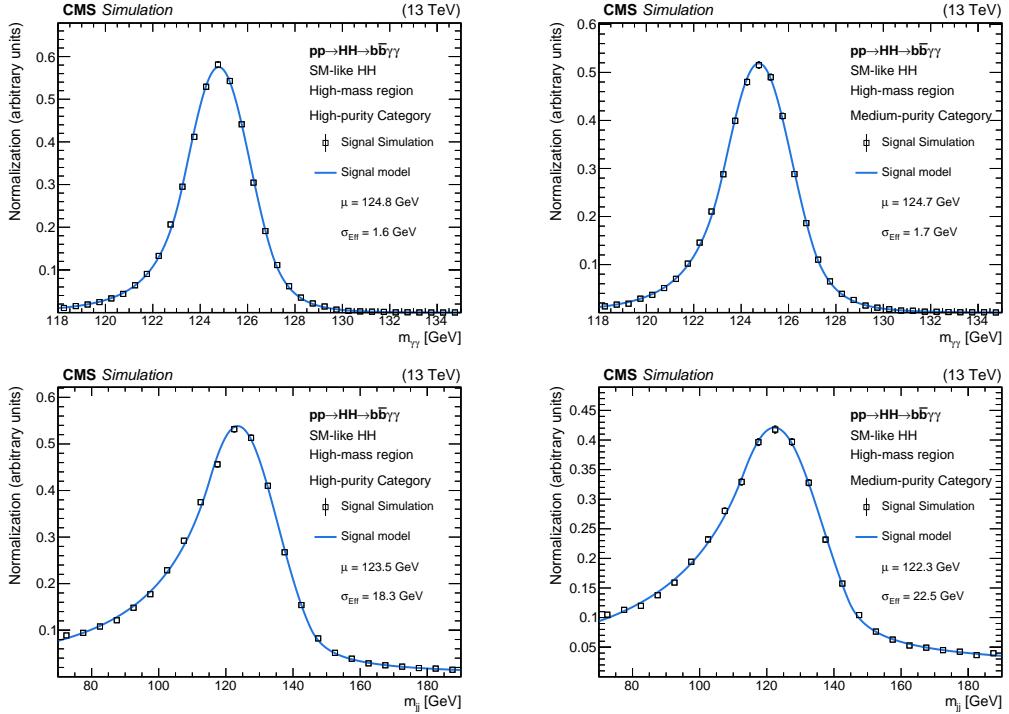


FIGURE 5.1: Signal fits for the SM HH non-resonant sample after full analysis selection, in high mass region, high (left) and medium (right) purity categories. Top plots: $m_{\gamma\gamma}$. Bottom plots: m_{jj} .

5.1.1 Correlation Studies

The chosen 2D fit function is the direct product of two independent functions for two variables which doesn't take the correlation between two variables into account. Therefore, whether the correlation is sensitive to our signal samples should be checked. This check is performed by the comparison with the simulated signal samples and the 2D fit functions. The different between the distribution of the simulated samples and the function is defined as residual R_{ij} :

$$R_{ij} = \frac{N_{ij}^{\text{PDF}} - N_{ij}^{\text{MC}}}{\sigma_{N_{ij}^{\text{PDF}}}^{\text{Poisson}}}, \quad (5.4)$$

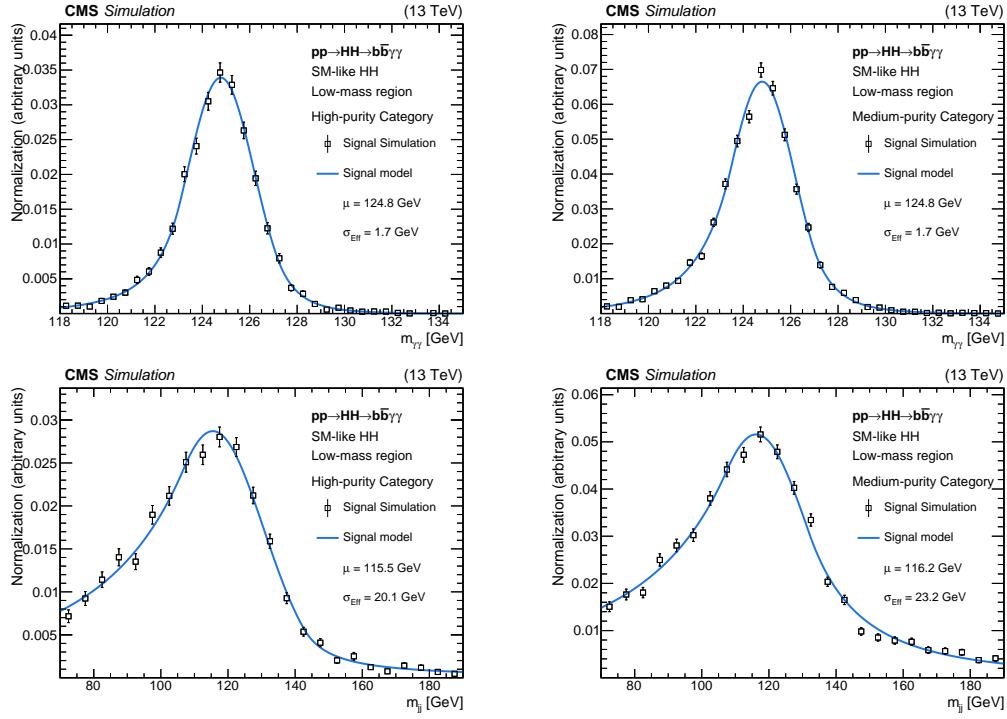


FIGURE 5.2: Signal fits for the SM HH non-resonant sample after full analysis selection, in low mass region, high (left) and medium (right) purity categories. Top plots: $m_{\gamma\gamma}$. Bottom plots: m_{jj} .

where ij refers to bin i in $m_{\gamma\gamma}$ and bin j in m_{jj} , and σ is the Poisson error of the function (PDF) and the simulated events (MC). The results are shown in Fig. 5.4 - 5.7. There is no specific structure in the region where the signal is expected. Therefore, we assume that there is no correlation between these two modeling variables.

5.2 Background modeling

In our analysis, the signal is not expected to have large contribution and doesn't form a huge peak in the spectrum of $m_{\gamma\gamma}$ and m_{jj} in the data. The background model can be derived by fitting the selected data with a smooth falling function, which is so-called data-driven method. The probable signal can be referred as the fluctuation and not fit by the falling function. For the non-resonant background, the contribution of single Higgs is considered as background and may become a peak in the mass spectrum. In the resonance search, the contribution from single Higgs production is already excluded by the mass window requirement. The fits of simulated single Higgs production are added into the background models, which are described in Sec. 5.2.3.

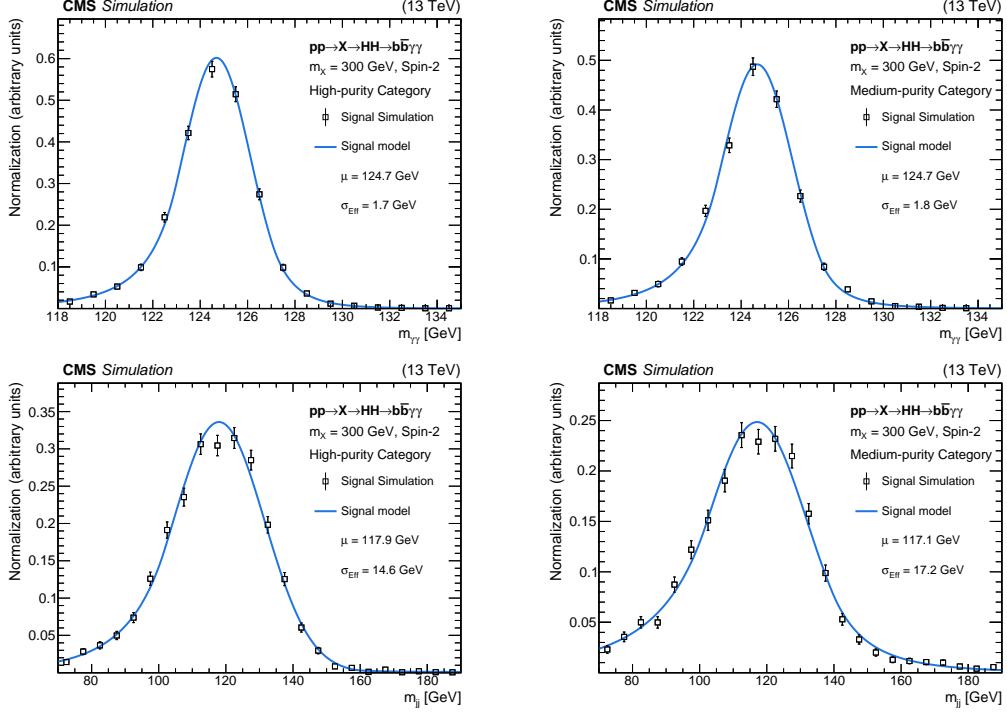


FIGURE 5.3: Signal fits for the Graviton 600 GeV mass sample after full analysis selection, in high (left) and medium (right) purity categories. Top plots: $m_{\gamma\gamma}$. Bottom plots: m_{jj} .

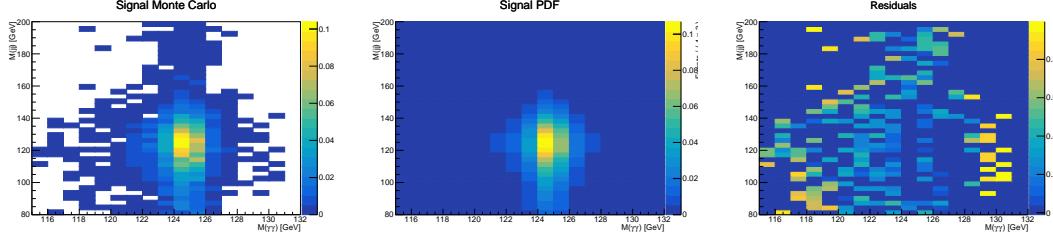


FIGURE 5.4: 2D distributions of the signal MC (left), fitted PDF model (center) and 2D residuals (right) for the High Mass-High Purity Category non-resonant selection.

The smooth falling functions chosen in this analysis are the family of Bernstein polynomials $\sum_{i=0}^n p_i B_i^n$ with order n , for example:

$$\begin{aligned} 1st \text{ order} &= p_0(1-x) + p_1x \\ 2nd \text{ order} &= p_0(1-x)^2 + p_12x(1-x) + p_2x^2 \\ 3rd \text{ order} &= p_0(1-x)^3 + p_13x(1-x)^2 + p_23x^2(1-x) + p_3x^3. \end{aligned} \quad (5.5)$$

The higher order polynomials have more degrees of freedom, and the precision of the fit is relative to the number of events. The chosen orders of polynomials are the same in both variables. Due to the different requirements of each signal region, the expected background yields can be very different. The chosen order

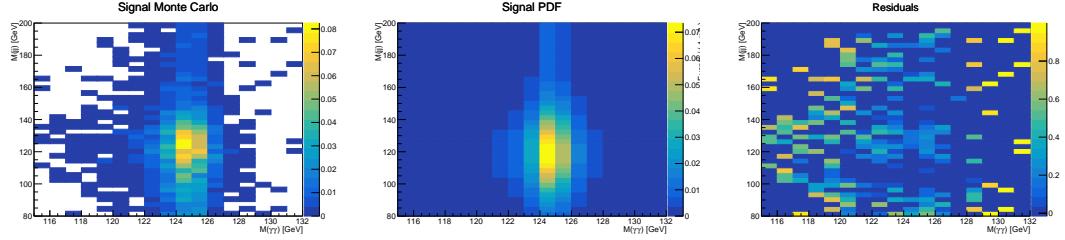


FIGURE 5.5: 2D distributions of the signal MC (left), fitted PDF model (center) and 2D residuals (right) for the High Mass-Medium Purity Category non-resonant selection.

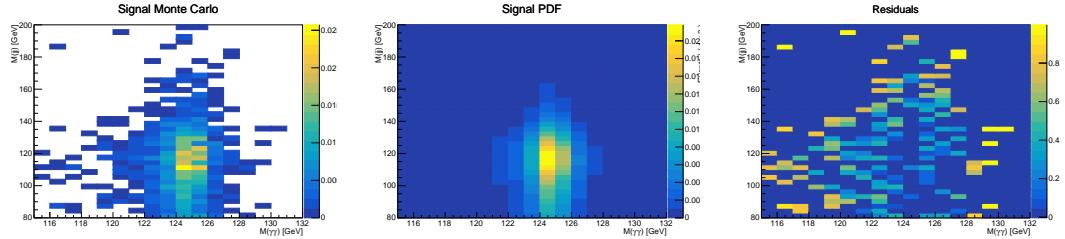


FIGURE 5.6: 2D distributions of the signal MC (left), fitted PDF model (center) and 2D residuals (right) for the Low Mass-High Purity Category non-resonant selection.

of Bernstein depends on the number of selected yields to obtain the good fitness. The categories with selected yields less than 15 are fit by first order, otherwise are fit by second order.

The fit results projected to one of variables in resonance case are shown in Fig. 5.8 and Fig. 5.9. The background only (green dash line) is from the fit of Bernstein polynomials. The full background (blue line) includes the single Higgs production background which are normalized according to their cross-section. The background modeling for resonance search is illustrated in Fig. 5.10.

5.2.1 Bias Studies

It is impossible to define an exact function to model the background. The chosen function for background modeling must be checked that have no bias toward our signal. The bias is measured by extracting the signal strength μ with the chosen background function fit to MC toys. The MC toys are generated by other background shape hypotheses, which are obtained by the fit on the data photon control region. The fit results should not appear a statistically significant bias in the signal strength with all background shape hypotheses. The bias is defined as

$$B = (\mu_{fit} - \mu_{true})/\sigma_\mu, \quad (5.6)$$

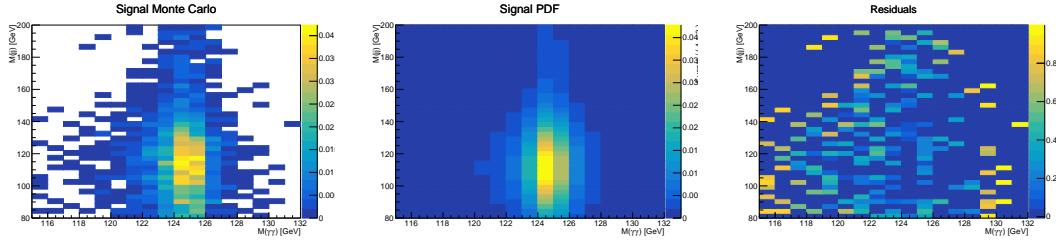


FIGURE 5.7: 2D distributions of the signal MC (left), fitted PDF model (center) and 2D residuals (right) for the Low Mass-Medium Purity Category non-resonant selection.

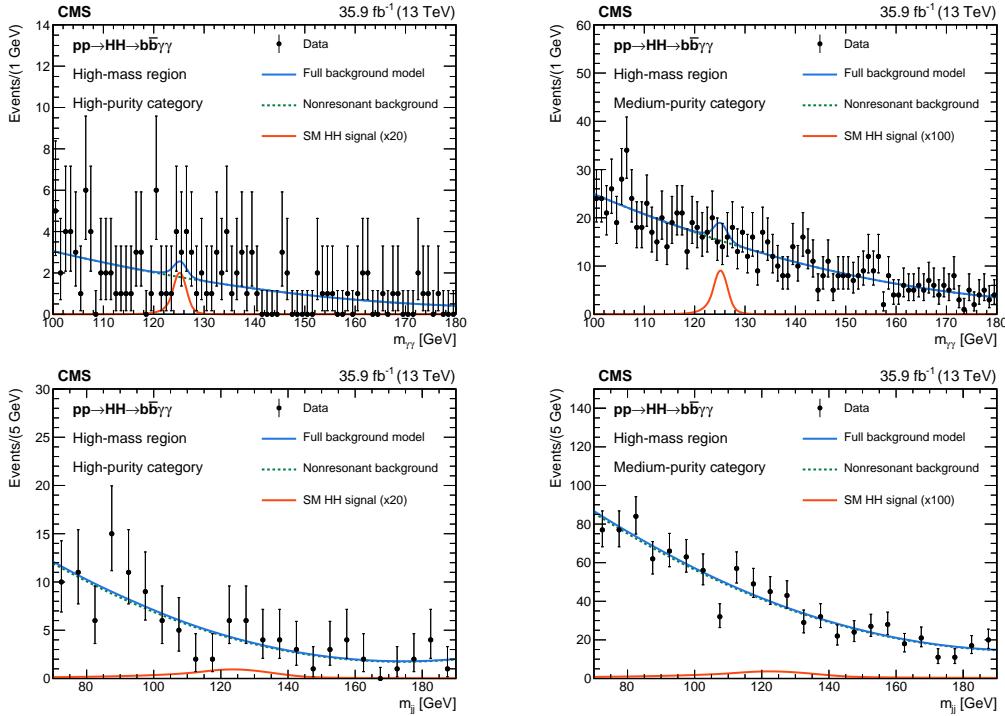


FIGURE 5.8: Background fits for the SM HH non-resonant analysis in the high mass region. The plots on the left (right) show the distributions in the HPC (MPC) region.

where σ_μ is the uncertainty from the fit procedure, and the signal strength μ describes the composition of the signal model in the fit:

$$\text{Fit Result} = \mu \times (\text{Signal model}) + (\text{Background model}). \quad (5.7)$$

The bias of chosen function fit to other background shape hypotheses is required to be smaller than 14%, which is justified by the effect of uncertainties on the signal strength which would be smaller than 1%.

The test background hypotheses are Laurent series and sums of n exponentials for both $m_{\gamma\gamma}$ and m_{jj} . Each hypothesis is used to generate 2000 toys for the different categories. These toys are injected the expected signal yields according to the different categories with assuming signal cross-section of 1 fb. ($\mu_{\text{true}} = 1$)

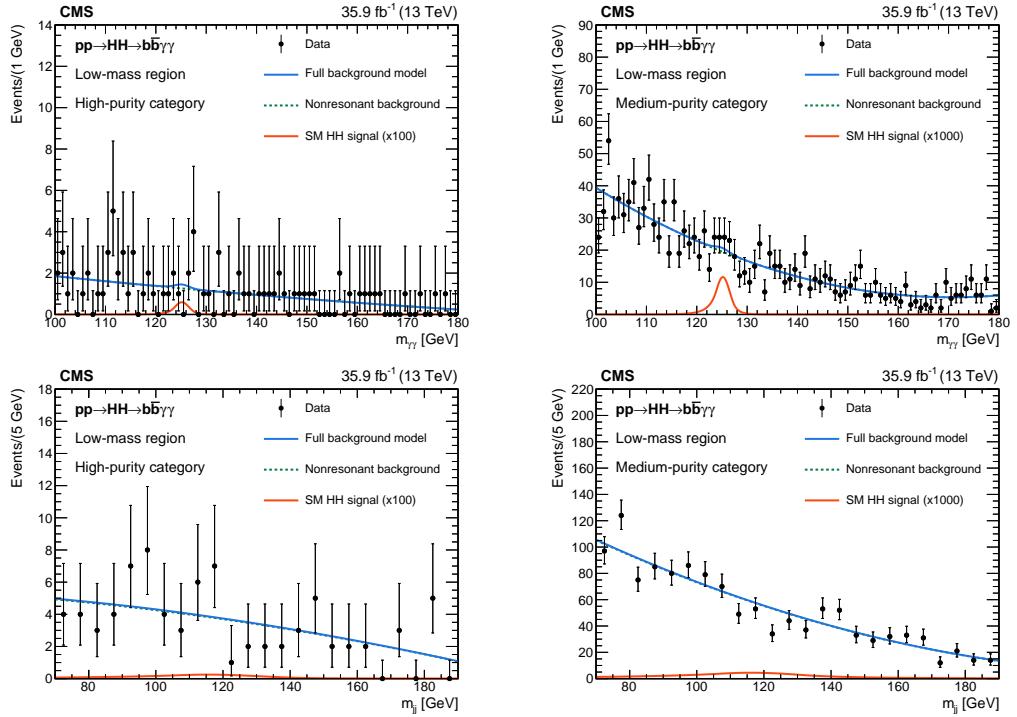


FIGURE 5.9: Background fits for the SM HH non-resonant analysis in the low mass region. The plots on the left (right) show the distributions in the HPC (MPC) region.

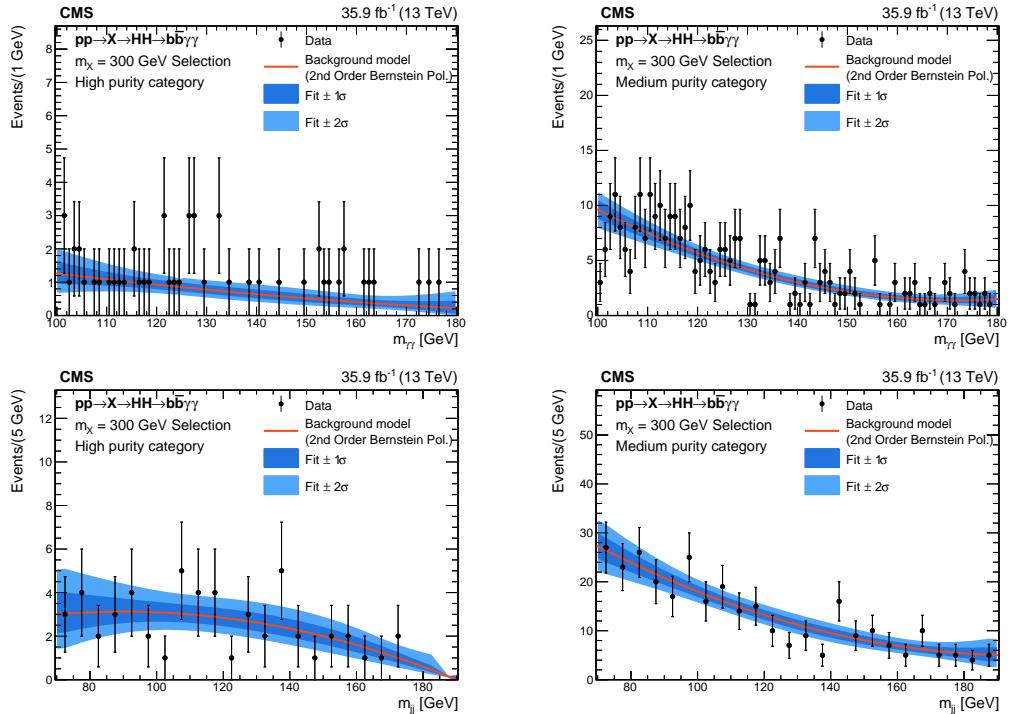


FIGURE 5.10: Background fits for the resonant analysis selection, assuming a spin-2 (Graviton) resonance with $m_X = 300$ GeV. The plots on the left (right) show the distributions in the HPC (MPC) region.

Some examples of measured biases are shown in Fig. 5.11, Fig. 5.12. The measured bias of our chosen function is acceptable.

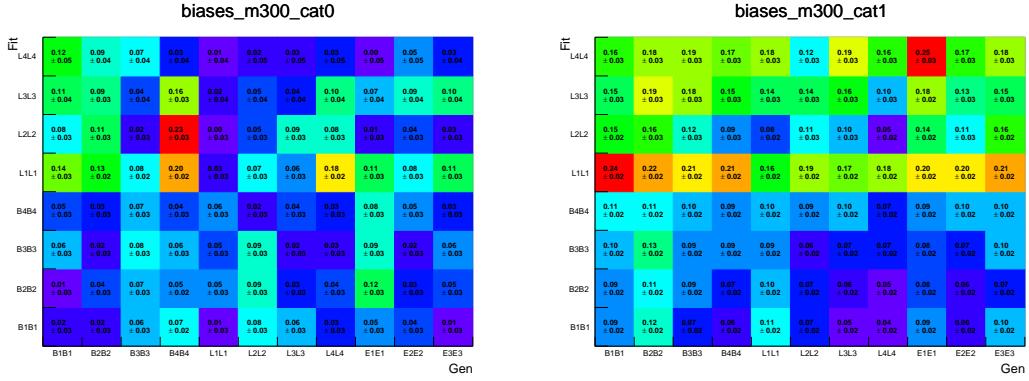


FIGURE 5.11: Biases measured in the 300 GeV resonant selection in the HPC (left) and MPC (right).

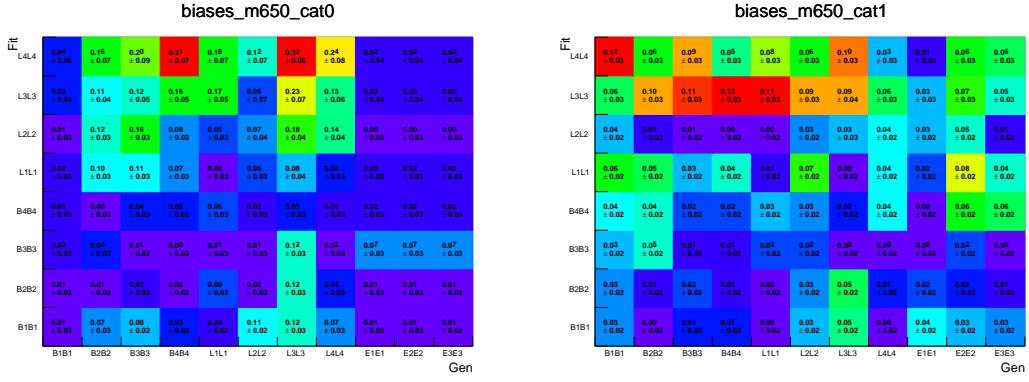


FIGURE 5.12: Biases measured in the 650 GeV resonant selection in the HPC (left) and MPC (right).

5.2.2 Correlation Studies

Since there is no peak structure in the background shape, it is hard to check the correlation just by the structure of residues. We generate the toy dataset according to the fit function with additional terms to simulate the correlation between $m_{\gamma\gamma}$ and m_{jj} . The original fit function $g(x, y)$ assuming that two variables are independent can be expressed as:

$$g(x, y) = \left(\sum_i a_i x^i \right) \left(\sum_k a_k y^k \right), \quad (5.8)$$

where x and y is $m_{\gamma\gamma}$ and m_{jj} in our case.

In general, the correlation can happen in the coefficient a_i and a_k . We just assume the polynomials fit function with more degree of freedom for fitting to approach this condition:

$$f(x, y) = \sum_i \sum_k c_{ik} x^i y^k. \quad (5.9)$$

Now we add back more terms to $g(x, y)$ to increase the degree of freedom. With the second order polynomials, the additional terms can be written as:

$$g_{corr}(x, y) = g(x, y) + \alpha \cdot m_{\gamma\gamma} \cdot m_{jj} + \beta \cdot m_{\gamma\gamma}^2 \cdot m_{jj} + \omega \cdot m_{\gamma\gamma} \cdot m_{jj}^2. \quad (5.10)$$

The toys are generated by $g_{corr}(x, y)$ with different coefficient (α, β, ω) and different number of events, and injected the signal events. These toys are fit by $g(x, y)$ to extract the signal strength, and calculate the bias.

Fig. 5.13 shows the bias as a function of α . Because there is no bias larger than 14%, the impact on the correlation terms can be ignored.

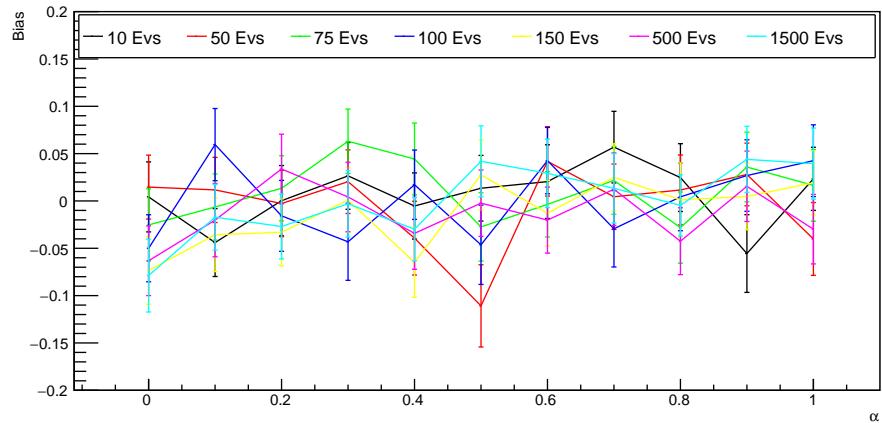


FIGURE 5.13: The average bias on measuring the signal with $g(m_{\gamma\gamma}, m_{jj})$ on toys created by $g_{corr}(m_{\gamma\gamma}, m_{jj})$ with α from 0 to 1.

5.2.3 Single Higgs background modeling

The single Higgs modeling relies on the simulated samples described in Sec. 4.1.4. There are five mentioned production modes considered in this analysis: ggH, VBF H, $t\bar{t}$ H, $b\bar{b}$ H, and VH. Different production modes cause the different shapes in the distribution of $m_{\gamma\gamma}$ and m_{jj} and rely on different fit functions. For the gluon-gluon fusion production (ggH) and vector boson fusion (VBF H), two selected background jets constitute a smooth falling spectrum. They are modeled by double-sided Crystal Ball function \times second order Bernstein polynomials for

$m_{\gamma\gamma} \times m_{jj}$. For the other single Higgs production associated with $b\bar{b}$, $t\bar{t}$, and one vector boson, the resonance is expected in the m_{jj} spectrum due to the kinematics turn on and vector boson resonance. They are modeled by two double-sided Crystal Ball functions both for $m_{\gamma\gamma}$ and m_{jj} .

The predicted cross-section of these processes and selection efficiencies are listed in Tab. 5.1. The examples of fit model are shown in Fig. 5.14 for ggH process and Fig. 5.15 for ttH process.

TABLE 5.1: SM single Higgs cross sections at 13 TeV with their respective selection efficiencies for the four different non-resonant analysis categories.

	Cross section (pb)	HM-HPC (%)	HM-MPC (%)	LM-HPC (%)	LM-MPC (%)
ggH	44.14	0.029 ± 0.0017	0.148 ± 0.0038	0.033 ± 0.0018	0.151 ± 0.0039
VBF	3.7820	0.038 ± 0.001	0.239 ± 0.0025	0.048 ± 0.0011	0.242 ± 0.0025
VH	2.257	0.271 ± 0.0038	0.748 ± 0.0063	0.367 ± 0.0044	0.962 ± 0.0071
$b\bar{b}H$	0.488	0.0297 ± 0.0035	0.262 ± 0.010	1.02 ± 0.020	2.59 ± 0.032
$t\bar{t}H$	0.5071	3.41 ± 0.027	3.69 ± 0.029	8.38 ± 0.042	8.17 ± 0.042

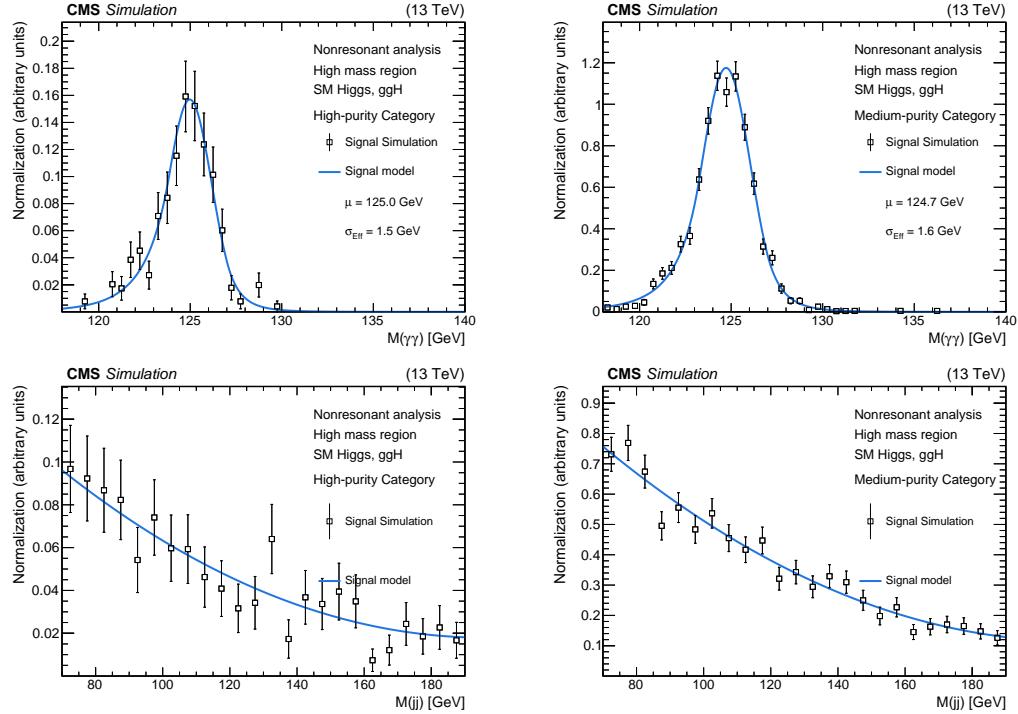


FIGURE 5.14: The modeling for single Higgs production through gluon-gluon fusion in high mass, high purity (left) and medium purity (right) region. Top plots are the projections in $m_{\gamma\gamma}$ and bottom plots are the projections in m_{jj} .

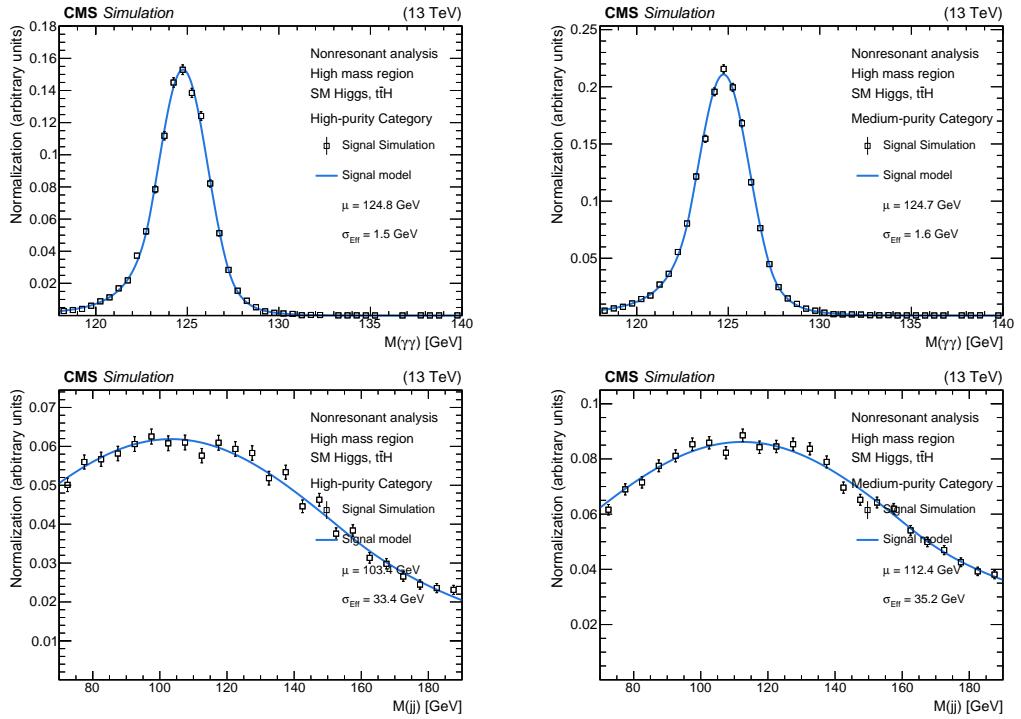


FIGURE 5.15: The modeling for single Higgs production associated with a pair of top quarks in high mass, high purity (left) and medium purity (right) region. Top plots are the projections in $m_{\gamma\gamma}$ and bottom plots are the projections in m_{jj} .

5.3 Systematic uncertainties

In this analysis, the systematics uncertainties come from two possible categories: normalization and shape.

First, the uncertainty of estimation of the integrated luminosity modifies the normalization of expected signal. It is taken as 2.5%. [84] Other uncertainties influence the selection efficiency and impact on the modeling and signal extraction.

The photon relative uncertainties are provided by SM $H \rightarrow \gamma\gamma$ analysis, since we follow the same selection. [85] The uncertainties from the measurement of photon energy scale (PES) and resolution (PER) are transferred into $\Delta m_{\gamma\gamma}/m_{\gamma\gamma}$ and $\Delta\sigma m_{\gamma\gamma}/\sigma m_{\gamma\gamma}$, and are estimated about 0.5 % and 5 % respectively.

In the case of jets, the similar uncertainties as photons from jet energy scale (JES) and resolution (JER) are taken into account. [82] They impacts on JER about 5% and on JES about 1%. Their impacts on our selection procedure are about 0.5%.

Additionally, the uncertainty related to b-tag is considered due to the categorization MVA. The scale factors are applied to the simulated training samples to match the shape of data in the b-tag score distribution. The uncertainty of the

classification MVA is estimated by varying the scale factor within one standard deviation of its uncertainty. The impacts from the other source of uncertainties are found to be negligible in the classification procedure.

The theoretical uncertainties of the single-Higgs production and double-Higgs production are considered according to Ref. [4]. The uncertainties from BSM are taken into account in the procedure of signal extraction.

All of the systematic uncertainties are summarized in Tab. 5.2.

TABLE 5.2: Summary of systematic uncertainties.

Source of systematic uncertainties	Type	Value
General uncertainties		
Integrated luminosity	Normalization	2.5%
Photon related uncertainties		
PES $\frac{\Delta m_{\gamma\gamma}}{m_{\gamma\gamma}}$	Shape	2.0%
PER $\frac{\Delta \sigma m_{\gamma\gamma}}{\sigma m_{\gamma\gamma}}$	Shape	1.0%
Di-photon selection (with trigger uncertainties and PES)	Normalization	0.5%
Photon identification	Normalization	5.0%
Jet related uncertainties		
JES $\frac{\Delta m_{jj}}{m_{jj}}$	Shape	0.5%
JER $\frac{\Delta \sigma m_{jj}}{\sigma m_{jj}}$	Shape	1.0%
Di-jet selection (JES+JER)	Normalization	5.0%
Resonant analysis specific uncertainties		
Mass window selection (JES+JER)	Normalization	3.0%
Classification MVA (HPC)	Normalization	11-19%
Classification MVA (MPC)	Normalization	3-9%
Non-resonant analysis specific uncertainties		
M_x Classification	Normalization	0.5%
Classification MVA (HPC)	Normalization	11-19%
Classification MVA (MPC)	Normalization	3-9%
Theoretical uncertainties of SM single-Higgs boson production		
QCD missing orders (ggH, VBF H, VH, ttH)	Normalization	0.4-5.8%
PDF and uncertainties (ggH, VBF H, VH, ttH)	Normalization	1.6-3.6%
Theory uncertainty bbH	Normalization	20%
Theoretical uncertainties of SM di-Higgs boson production		
QCD missing orders	Normalization	4.3-6%
PDF and α_s uncertainties	Normalization	3.1%
m_T effects	Normalization	5%

Chapter 6

Results

The procedure of limits extraction follows Ref. [86], with Higgs Analysis Combination tools [87]. The high purity and medium purity categories are combined for all results.

6.1 Resonance results

The observed and expected upper limits at 95% confidence level (CL) are shown in Fig. 6.1, for the cross-section of $pp \rightarrow X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ process assuming spin-0 and spin-2 heavy resonances. The observed data are from 0.23 fb to 4.2 fb, depending on the resonance mass hypotheses.

The results are compared with the cross sections for the spin-0 radion and spin-2 KK-graviton production in the bulk RS WED models. In radion case, the cross-section is proportional to the parameter $1/\Lambda_R^2$. In the Run I results, $\Lambda_R = 1$ TeV is excluded below 980 GeV in the radion search and have no sensitivity at $\Lambda_R = 3$ TeV. In this analysis, we compare the results of the radion search with $\Lambda_R = 2$ TeV and $\Lambda_R = 3$ TeV. The observed limits exclude the points of m_x below 840 GeV in $\Lambda_R = 2$ TeV and of m_x below 540 GeV in $\Lambda_R = 3$ TeV. In the graviton case, the parameter for the properties of graviton k/\overline{M}_{Pl} equals to 1.0 and 0.5 are compared. With assuming $k/\overline{M}_{Pl} = 1.0$, the range $290 < m_x < 810$ GeV is excluded. With assuming $k/\overline{M}_{Pl} = 0.5$, the range $350 < m_x < 530$ GeV is excluded.

6.2 Non-resonance results

The observed (expected) 95% CL upper limits of SM-like di-Higgs production on the cross-section of $pp \rightarrow HH \rightarrow b\bar{b}\gamma\gamma$ are 2.0 (1.6) fb, which can be translated into 0.79 (0.63) pb for the cross-section of $pp \rightarrow HH$ corresponding to about 24 (19) times the SM prediction.

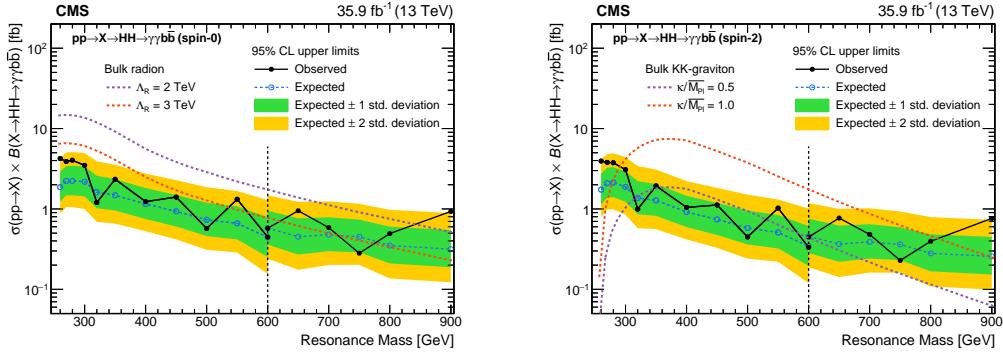


FIGURE 6.1: Observed and expected 95% CL upper limits on the cross-section $\sigma(pp \rightarrow X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma)$ combining two categories in high mass region and low mass region. The limits of 600 GeV are set in both high mass and low mass methods. The green and yellow bands present one and two standard deviation, respectively. The theoretical predictions in WED with different parameters are compared.

The results are also performed in the anomalous couplings study. The results of limits on the BSM benchmark points in 5-D BSM coupling space are shown in Fig. 6.2, which provide the constraint in different regions. The tightest constraint happens on the benchmark 2, which has a wide peak around 800 GeV in m_{HH} spectrum and has a lot of events in the high mass category. The loosest constraint is placed on the benchmark 7, which has most events locating below $m_{HH} = 300$ GeV and in low mass category. It is observed that the high mass category has good sensitivity.

In addition, the scanning of κ_λ with the other parameters keeping in the SM value ($\kappa_t = 1$, other BSMs = 0) is also searched. The 95% CL limits are set on the cross-section as a function of κ_λ , and are shown in Fig. 6.3. The κ_λ is constrained in the range $-11 < \kappa_\lambda < 17$ by the observed data.

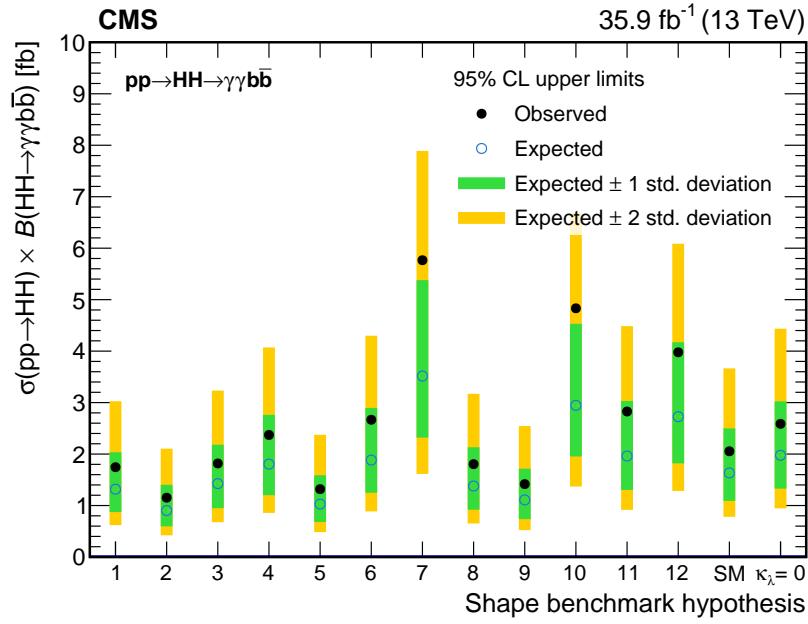


FIGURE 6.2: Observed and expected 95% CL upper limits on the cross-section $\sigma(pp \rightarrow X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma)$ in recommended benchmark points.

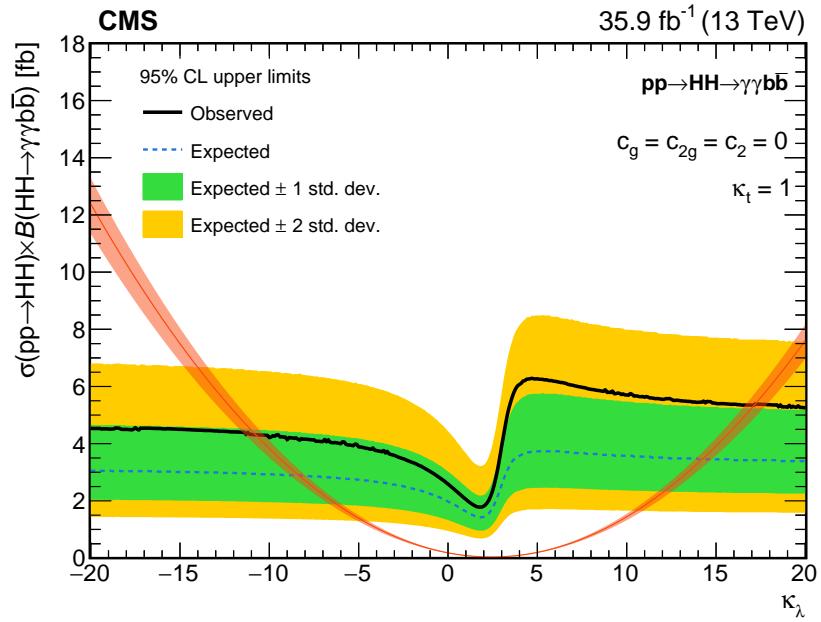


FIGURE 6.3: Observed and expected 95% CL upper limits on the cross-section $\sigma(pp \rightarrow X \rightarrow HH \rightarrow b\bar{b}\gamma\gamma)$ as a function of κ_λ . Other parameters are the same as the SM.

Chapter 7

Conclusion

This thesis is devoted for the search of the di-Higgs boson (HH) production decaying into the $b\bar{b}\gamma\gamma$ in the proton-proton collisions at $\sqrt{s} = 13\text{TeV}$ during the LHC Run II. The data are collected by the CMS detector, which correspond to an integrated luminosity of 35.9 fb^{-1} . Both the Standard Model and BSMs are searched for the resonant and the non-resonant HH productions. The new developed techniques are updated in this analysis, and the new energy regression for the b jets improves about 10% in the sensitivity. Upper limits at a 95% confidence level are set on the cross sections for the HH production decaying to two photons and two b jets. No statistically significant deviations from the expectations are found in all searches. The new particles in WED model with spin-0 (radion) and spin-2 (graviton) decaying into HH boson are searched in the mass range between 250 GeV and 900 GeV. The observed limits are from 0.23 fb to 4.2 fb, which exclude the radion (spin-0) signal hypothesis, assuming the scale parameter $\Lambda_R = 3 \text{ TeV}$, for all masses below $m_X = 550 \text{ GeV}$, and the KK-graviton (spin-2) hypothesis for the mass range $280 < m_X < 800 \text{ GeV}$, assuming $\kappa/\overline{M}_{Pl} = 1$. The observed limit on the cross-section of the SM-like HH production is 2.0 fb, which is about 24 times the SM prediction. The anomalous couplings for the BSMs approach are also searched. The deviation of the Higgs self-coupling constant in the SM is scanned and constrained in the range $-11 < \kappa_\lambda < 17$. The recommended points in the anomalous couplings space with specific kinematics shape are also observed. There is no significant data excess in all of points.

Bibliography

- [1] The CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: (2012). DOI: 10.1016/j.physletb.2012.08.021. eprint: arXiv:1207.7235.
- [2] The ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: (2012). DOI: 10.1016/j.physletb.2012.08.020. eprint: arXiv: 1207.7214.
- [3] ATLAS and CMS Collaborations. “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”. In: (2016). DOI: 10.1007/JHEP08(2016)045. eprint: arXiv: 1606.02266.
- [4] D. de Florian et al. “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector”. In: (2016). DOI: 10.23731/CYRM-2017-002. eprint: arXiv:1610.07922.
- [5] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Physical Review Letters* 13.9 (1964), pp. 321–323. DOI: 10.1103/physrevlett.13.321. URL: <https://doi.org/10.1103/physrevlett.13.321>.
- [6] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Physical Review Letters* 13.16 (1964), pp. 508–509. DOI: 10.1103/physrevlett.13.508. URL: <https://doi.org/10.1103/physrevlett.13.508>.
- [7] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Physical Review Letters* 13.20 (1964), pp. 585–587. DOI: 10.1103/physrevlett.13.585. URL: <https://doi.org/10.1103/physrevlett.13.585>.
- [8] Mark Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. ISBN: 9781107034266. URL: <http://www-spires.fnal.gov/spires/find/books/www?cl=QC793.2.T46::2013>.
- [9] F. Maltoni, E. Vryonidou, and M. Zaro. “Top-quark mass effects in double and triple Higgs production in gluon-gluon fusion at NLO”. In: (2014). DOI: 10.1007/JHEP11(2014)079. eprint: arXiv:1408.6542.

- [10] J. Baglio et al. “The measurement of the Higgs self-coupling at the LHC: theoretical status”. In: (2012). DOI: 10.1007/JHEP04(2013)151. eprint: arXiv:1212.5581.
- [11] R. Frederix et al. “Higgs pair production at the LHC with NLO and parton-shower effects”. In: (2014). DOI: 10.1016/j.physletb.2014.03.026. eprint: arXiv:1401.7340.
- [12] Nima Arkani-Hamed, Savas Dimopoulos, and Gia Dvali. “The Hierarchy Problem and New Dimensions at a Millimeter”. In: (1998). DOI: 10.1016/S0370-2693(98)00466-3. eprint: arXiv:hep-ph/9803315.
- [13] Lisa Randall and Raman Sundrum. “A Large Mass Hierarchy from a Small Extra Dimension”. In: (1999). DOI: 10.1103/PhysRevLett.83.3370. eprint: arXiv:hep-ph/9905221.
- [14] Alexandra Carvalho. *Gravity particles from Warped Extra Dimensions, predictions for LHC*. 2014. eprint: arXiv:1404.0102.
- [15] A. Liam Fitzpatrick et al. “Searching for the Kaluza-Klein Graviton in Bulk RS Models”. In: (2007). DOI: 10.1088/1126-6708/2007/09/013. eprint: arXiv:hep-ph/0701150.
- [16] <https://github.com/CrossSectionsLHC/WED>.
- [17] Aielet Efrati and Yosef Nir. *What if $\lambda_{hhh} \neq 3m_h^2/v$* . 2014. eprint: arXiv: 1401.0935.
- [18] B. Grzadkowski et al. “Dimension-Six Terms in the Standard Model Lagrangian”. In: (2010). DOI: 10.1007/JHEP10(2010)085. eprint: arXiv: 1008.4884.
- [19] Florian Goertz et al. “Higgs boson pair production in the D=6 extension of the SM”. In: (2014). DOI: 10.1007/JHEP04(2015)167. eprint: arXiv: 1410.3471.
- [20] W. Buchmüller and D. Wyler. “Effective lagrangian analysis of new interactions and flavour conservation”. In: *Nuclear Physics B* 268.3-4 (1986), pp. 621–653. DOI: 10.1016/0550-3213(86)90262-2. URL: [https://doi.org/10.1016/0550-3213\(86\)90262-2](https://doi.org/10.1016/0550-3213(86)90262-2).
- [21] Alexandra Carvalho et al. “Higgs Pair Production: Choosing Benchmarks With Cluster Analysis”. In: (2015). DOI: 10.1007/JHEP04(2016)126. eprint: arXiv:1507.02245.
- [22] Alexandra Carvalho et al. *Analytical parametrization and shape classification of anomalous HH production in the EFT approach*. 2016. eprint: arXiv:1608.06578.

- [23] *Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events*. Tech. rep. CMS-PAS-BTV-13-001. Geneva: CERN, 2013. URL: <http://cds.cern.ch/record/1581306>.
- [24] CMS Collaboration. “Reconstruction and identification of tau lepton decays to hadrons and tau neutrino at CMS”. In: (2015). DOI: 10.1088/1748-0221/11/01/P01019. eprint: arXiv:1510.07488.
- [25] ATLAS Collaboration. “Search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state from pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: (2015). DOI: 10.1140/epjc/s10052-015-3628-x. eprint: arXiv:1506.00285.
- [26] ATLAS Collaboration. “Search For Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State using pp Collision Data at $\sqrt{s} = 8$ TeV from the ATLAS Detector”. In: (2014). DOI: 10.1103/PhysRevLett.114.081802. eprint: arXiv:1406.5053.
- [27] ATLAS Collaboration. “Searches for Higgs boson pair production in the $hh \rightarrow b\bar{b}\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector”. In: (2015). DOI: 10.1103/PhysRevD.92.092004. eprint: arXiv:1509.04670.
- [28] Daniel de Florian and Javier Mazzitelli. “Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD”. In: (2013). DOI: 10.1103/PhysRevLett.111.201801. eprint: arXiv:1309.6594.
- [29] CMS Collaboration. “Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs in proton-proton collisions at 8 TeV”. In: (2015). DOI: 10.1016/j.physletb.2015.08.047. eprint: arXiv:1503.04114.
- [30] CMS Collaboration. “Search for two Higgs bosons in final states containing two photons and two bottom quarks in proton-proton collisions at 8 TeV”. In: (2016). DOI: 10.1103/PhysRevD.94.052012. eprint: arXiv:1603.06896.
- [31] CMS Collaboration. “A search for Higgs boson pair production in the bb tau tau final state in proton-proton collisions at $\sqrt{s} = 8$ TeV”. In: (2017). DOI: 10.1103/PhysRevD.96.072004. eprint: arXiv:1707.00350.
- [32] CMS Collaboration. “Search for heavy resonances decaying to two Higgs bosons in final states containing four b quarks”. In: (2016). DOI: 10.1140/epjc/s10052-016-4206-6. eprint: arXiv:1602.08762.
- [33] *Search for resonant pair production of Higgs bosons decaying to $b\bar{b}$ and $\tau^+\tau^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*. Tech. rep. CMS-PAS-EXO-15-008. Geneva: CERN, 2015. URL: <https://cds.cern.ch/record/2125293>.

- [34] Lyndon Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.
- [35] CERN. *Accelerator Performance and Statistics*. <http://acc-stats.web.cern.ch/acc-stats/>.
- [36] CMS Collaboration. *CMS Luminosity Public Results*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [37] S. Chatrchyan et al. "The CMS experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.
- [38] *The CMS magnet project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: <http://cds.cern.ch/record/331056>.
- [39] The CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *JINST* 9.10 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009.
- [40] V Karimäki et al. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/368412>.
- [41] Cristina Biino. "The CMS Electromagnetic Calorimeter: overview, lessons learned during Run 1 and future projections". In: *Journal of Physics: Conference Series* 587.1 (2015), p. 012001. URL: <http://stacks.iop.org/1742-6596/587/i=1/a=012001>.
- [42] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: <http://cds.cern.ch/record/349375>.
- [43] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: <http://cds.cern.ch/record/357153>.
- [44] *The CMS muon project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: <http://cds.cern.ch/record/343814>.
- [45] *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical Design Report CMS. URL: <http://cds.cern.ch/record/706847>.
- [46] CMS Collaboration. *CMS software*. <https://github.com/cms-sw>.
- [47] Sergio Cittolin, Attila Rácz, and Paris Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. Geneva: CERN, 2002. URL: <http://cds.cern.ch/record/578006>.

- [48] Andreas Hoecker et al. "TMVA: Toolkit for Multivariate Data Analysis". In: *PoS ACAT* (2007), p. 040. arXiv: physics/0703039.
- [49] CMS Collaboration. "Search for the standard model Higgs boson decaying into two photons in pp collisions at $\sqrt{s}=7$ TeV". In: (2012). DOI: 10.1016/j.physletb.2012.03.003. eprint: arXiv:1202.1487.
- [50] CMS Collaboration. "Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV". In: (2015). DOI: 10.1088/1748-0221/10/08/P08010. eprint: arXiv:1502.02702.
- [51] CMS Collaboration. *GBRlikelihood package*. <https://github.com/cms-egamma/HiggsAnalysis/tree/master/GBRLikelihood>.
- [52] E. Norrbin and T. Sjöstrand. "Production and Hadronization of Heavy Quarks". In: (2000). DOI: 10.1007/s100520000460. eprint: arXiv: hep-ph/0005110.
- [53] *Identification of b quark jets at the CMS Experiment in the LHC Run 2*. Tech. rep. CMS-PAS-BTV-15-001. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2138504>.
- [54] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti- k t jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), p. 063. URL: <http://stacks.iop.org/1126-6708/2008/i=04/a=063>.
- [55] The CMS Collaboration. "Measurement of B anti-B Angular Correlations based on Secondary Vertex Reconstruction at $\sqrt{s}=7$ TeV". In: (2011). DOI: 10.1007/JHEP03(2011)136. eprint: arXiv:1102.3194.
- [56] CMS Collaboration. *MVA Framework Offline Guide*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMVAFramework>.
- [57] "Heavy flavor identification at CMS with deep neural networks". In: (2017). URL: <https://cds.cern.ch/record/2255736>.
- [58] *Search for the standard model Higgs boson produced through vector boson fusion and decaying to bb with proton-proton collisions at $\sqrt{s} = 13$ TeV*. Tech. rep. CMS-PAS-HIG-16-003. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2160154>.
- [59] CMS Collaboration. "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks". In: (2013). DOI: 10.1103/PhysRevD.89.012003. eprint: arXiv:1310.3687.
- [60] *Jet algorithms performance in 13 TeV data*. Tech. rep. CMS-PAS-JME-16-003. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2256875>.

- [61] *Pileup Jet Identification*. Tech. rep. CMS-PAS-JME-13-005. Geneva: CERN, 2013. URL: <https://cds.cern.ch/record/1581583>.
- [62] CMS Collaboration. *Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$* . 2018. eprint: arXiv: 1804.02716.
- [63] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: (2014). DOI: 10.1007/JHEP07(2014)079. eprint: arXiv:1405.0301.
- [64] Benoit Hespel, David Lopez-Val, and Eleni Vryonidou. “Higgs pair production via gluon fusion in the Two-Higgs-Doublet Model”. In: (2014). DOI: 10.1007/JHEP09(2014)124. eprint: arXiv:1407.0281.
- [65] Andy Buckley et al. “LHAPDF6: parton density access in the LHC precision era”. In: (2014). DOI: 10.1140/epjc/s10052-015-3318-8. eprint: arXiv:1412.7420.
- [66] Stefano Carrazza et al. *A compression algorithm for the combination of PDF sets*. 2015. eprint: arXiv:1504.06469.
- [67] Jon Butterworth et al. “PDF4LHC recommendations for LHC Run II”. In: (2015). DOI: 10.1088/0954-3899/43/2/023001. eprint: arXiv: 1510.03865.
- [68] Sayipjamal Dulat et al. “New parton distribution functions from a global analysis of quantum chromodynamics”. In: (2015). DOI: 10.1103/PhysRevD.93.033006. eprint: arXiv:1506.07443.
- [69] L. A. Harland-Lang et al. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: (2014). DOI: 10.1140/epjc/s10052-015-3397-6. eprint: arXiv:1412.3989.
- [70] The NNPDF Collaboration et al. “Parton distributions for the LHC Run II”. In: (2014). DOI: 10.1007/JHEP04(2015)040. eprint: arXiv:1410.8849.
- [71] Torbjörn Sjöstrand et al. “An Introduction to PYTHIA 8.2”. In: (2014). DOI: 10.1016/j.cpc.2015.01.024. eprint: arXiv:1410.3012.
- [72] CMS Collaboration. “Event generator tunes obtained from underlying event and multiparton scattering measurements”. In: (2015). DOI: 10.1140/epjc/s10052-016-3988-x. eprint: arXiv:1512.00815.
- [73] S. Agostinelli et al. “Geant4—a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), 250–303. DOI: 10.1016/s0168-9002(03)01368-8.

- [74] https://github.com/syuvivida/DibosonBSMSignal_13TeV.
- [75] Alexandra Carvalho et al. *On the reinterpretation of non-resonant searches for Higgs boson pairs*. 2017. eprint: arXiv:1710.08261.
- [76] Daniel Faeh and Nicolas Greiner. “Diphoton production in association with two bottom jets”. In: (2017). DOI: 10.1140/epjc/s10052-017-5296-5. eprint: arXiv:1706.08309.
- [77] Paolo Nason. “A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms”. In: (2004). DOI: 10.1088/1126-6708/2004/11/040. eprint: arXiv:hep-ph/0409146.
- [78] Stefano Frixione, Paolo Nason, and Carlo Oleari. “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”. In: (2007). DOI: 10.1088/1126-6708/2007/11/070. eprint: arXiv:0709.2092.
- [79] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: (2010). DOI: 10.1007/JHEP06(2010)043. eprint: arXiv:1002.2581.
- [80] E. Bagnaschi et al. “Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM”. In: (2011). DOI: 10.1007/JHEP02(2012)088. eprint: arXiv:1111.2854.
- [81] CMS Collaboration. “Particle-flow reconstruction and global event description with the CMS detector”. In: (2017). DOI: 10.1088/1748-0221/12/10/P10003. eprint: arXiv:1706.04965.
- [82] CMS Collaboration. “Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS”. In: (2011). DOI: 10.1088/1748-0221/6/11/P11002. eprint: arXiv:1107.4277.
- [83] Nilanjana Kumar and Stephen P. Martin. “LHC search for di-Higgs decays of stoponium and other scalars in events with two photons and two bottom jets”. In: (2014). DOI: 10.1103/PhysRevD.90.055007. eprint: arXiv:1404.0996.
- [84] CMS Luminosity Measurements for the 2016 Data Taking Period. Tech. rep. CMS-PAS-LUM-17-001. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2257069>.
- [85] CMS Collaboration. “Observation of the diphoton decay of the Higgs boson and measurement of its properties”. In: (2014). DOI: 10.1140/epjc/s10052-014-3076-z. eprint: arXiv:1407.0558.
- [86] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: (2010). DOI: 10.1140/epjc/s10052-011-1554-0. eprint: arXiv:1007.1727.

- [87] CMS Collaboration. *Higgs Analysis Combination tools*. <https://github.com/cms-analysis/HiggsAnalysis-CombinedLimit>.