Evidence for the Higgs boson production in association with top-quark pair with $\sqrt{s} = 13$ TeV of proton-proton collisions at LHC with the ATLAS detector

DISSERTATION
Presented in Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Graduate University for Advanced Studies, SOKENDAI

Satoshi Higashino

The Graduate University for Advanced Studies, SOKENDAI

December 10, 2019

Abstract

In the Standard Model of the particle physics, the mass of fermions is generated dynamically through the Yukawa interaction with Higgs field. However, the Yukawa interaction for top-quark is not directly observed yet, and could be a portal to the new physics. Therefore we conduct the search for the Yukawa interaction for top-quark, and report on the study of the Higgs boson production in association with top-quark pair using di-photon decay channel with 79.8 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector. In this analysis, the deep understanding of the photon selection efficiency is crucial. We studied the photon efficiency in detail using $Z \to ll\gamma$ and $Z \to ee$ events. Furthermore, we introduced the multivariate analysis with machine learning to achieve better signal separation from the background.

After the top-quark pair selection, we found the di-photon invariant mass peak at 125 GeV, which represents the existence of the Higgs boson. The statistical significance is estimated to be 4.2 standard deviations relative to the background-only hypothesis while the expected significance is 3.6 standard deviations. This result provides the evidence of the coupling between top-quark and Higgs boson. The cross section of the Higgs boson production is measured to be

$$\sigma_{ttH} = \left[694.9\ ^{+198.0}_{-179.3}\ (\text{stat.})\ ^{+141.7}_{-108.2}\ (\text{syst.})\right]\ \text{fb} = \left[694.9\ ^{+243.1}_{-207.7}\right]\ \text{fb},$$

which is slightly larger than the Standard Model prediction with the NLO calculation, $506.5\ ^{+34.6}_{-50.5}$ fb, although the uncertainty in the measurement is not small enough to judge whether the deviation is significant or not. The further study with better precision is important program in the high energy physics in future.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Standard Model (SM) is the particle physics theory describing three fundamental interactions; the electromagnetic, weak and strong forces. The SM predicts that there are 12 fermions, four gauge bosons and one scalar boson called Higgs boson which was discovered [1, 2] by the ATLAS [3] and CMS [4] experiments in 2012. Figure 1.1 shows the table of all the SM particles.

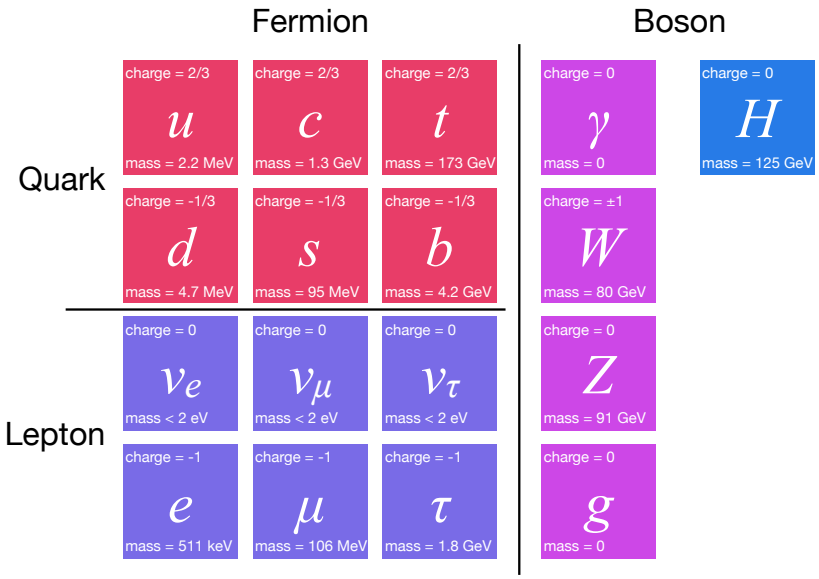


Figure 1.1: Standard model particles. Parameters are referred from [5].

Although the SM proofs its integrity, it cannot explain everything in the Universe. One of the mysteries in the Universe is the dark matter issue. The dark matter cannot be explained as any SM particles, thus new physics model(s) are required in addition to the SM.

The properties of Higgs boson are not precisely measured yet. One of the interesting measurable parameters is the coupling constant between Higgs boson and fermions, introduced

as the "Yukawa coupling" in the SM. The existence of the Yukawa coupling allows to give mass of each fermion. However, since the Yukawa coupling is introduced without fundamental principles, the detailed structure of the coupling is still unknown although the existence of Higgs boson has been proved. Thus, the measurement of the Yukawa coupling plays an important role to explore new physics model(s). This thesis focuses on the coupling between Higgs boson and top-quark (called Top-Yukawa coupling). In the SM, the fermion mass is linear to the Yukawa coupling and thus top-quark has the largest Yukawa coupling among all fermions. The top-quark mass is measured to be 173.34 $\pm$ 0.76 GeV by the ATLAS, CDF, CMS and DØ experiments [6–8]. Therefore, the measurement of the Top-Yukawa coupling allows to check the relation between the fermion mass and the Yukawa coupling.

However, the Top-Yukawa coupling is not observed at any experiments. In this thesis, we present the search for the Top-Yukawa coupling with the ATLAS detector at the LHC.

1.1 Higgs mechanism and Yukawa coupling

The SM is based on the gauge theory [9–12]. The difficulty of the gauge theory without the Higgs mechanism [13–18] is explained in Section 1.1.1. The solution of this problem is to introduce the Higgs mechanism which is presented in Section 1.1.2. The Yukawa coupling is described in Section 1.1.3.

1.1.1 Gauge theory

We start to discuss from the Quantum Electrodynamics (QED) which explains the electromagnetic interaction because it is the simplest gauge theory. The Lagrangian density $\mathcal{L}_{\rm QED}$ is described with a spin $\frac{1}{2}$ field ψ with mass m and an electromagnetic field A_μ with mass M as

$$\mathcal{L}_{\rm QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\psi}(i\gamma^\mu D_\mu - m)\psi + \frac{1}{2}M^2 A_\mu A^\mu, \tag{1.1}$$

where $F_{\mu\nu}$ is the electromagnetic field tensor defined by $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and D_μ is the covariant derivative $D_\mu = \partial_\mu + iqA_\mu$, where q is the electric charge. Under the gauge symmetry, $\mathcal{L}_{\rm QED}$ is invariant under the gauge transformation with arbitrary $\alpha(x)$ denoted as

$$\psi(x) \to e^{-iq\alpha(x)}\psi(x), \qquad A_\mu(x) \to A_\mu(x) + \partial_\mu\alpha(x). \tag{1.2}$$

In order to satisfy the gauge invariance, the mass of the electromagnetic field (equivalent to photon) must be zero, which is consistent with the experimental results.

On contrary with the electromagnetic interaction, the weak interaction is based on $SU(2)$ symmetry. The Lagrangian density is written as

$$\mathcal{L}_{weak} = -\frac{1}{4}\boldsymbol{W}_{\mu\nu}\boldsymbol{W}^{\mu\nu} + \overline{\psi}(i\gamma^\mu D_\mu - m)\psi \tag{1.3}$$

$$\boldsymbol{W}_{\mu\nu} = \partial_\mu \boldsymbol{W}_\nu - \partial_\nu \boldsymbol{W}_\mu - g\boldsymbol{W}_\mu \times \boldsymbol{W}_\nu, \tag{1.4}$$

where the $\boldsymbol{W}_\mu = (W^1_\mu, W^2_\mu, W^3_\mu)$ is an isospin triplet of $SU(2)$ field. The covariant derivative D_μ is defined as $D_\mu = \partial_\mu + ig\boldsymbol{W}_\mu \cdot \boldsymbol{T}$, where g is the coupling constant. $\boldsymbol{T}$ is the isospin operator represented by $T^i = \frac{1}{2}\sigma^i$ (i = 1, 2, 3), where σ^i is the i-th element of the Pauli matrices.

In this case, the gauge transformation with arbitrary $\boldsymbol{\alpha}(x)$ is through the special unitary group ($SU(2)$),

$$\psi(x) \rightarrow e^{ig\boldsymbol{\alpha}(x)\cdot\boldsymbol{T}}\psi, \tag{1.5}$$

$$\boldsymbol{W}_\mu(x) \rightarrow \boldsymbol{W}_\mu(x) + \partial_\mu\boldsymbol{\alpha}(x) + \boldsymbol{\alpha}(x) \times \boldsymbol{W}_\mu. \tag{1.6}$$

Under the transformation, the mass of the gauge bosons must be zero to conserve the gauge symmetry. However, this is inconsistent with the experimental results, i.e. massive gauge bosons, $W^\pm$ and Z^0 exist. It implies that some mechanisms are needed to explain the mass of the gauge bosons.

1.1.2 Higgs mechanism

In order to give mass to $W^\pm$ and Z^0 with keeping the theory gauge symmetric, addition of a scalar field is proposed in the model developed by Brout, Englert and Higgs. This scalar field is called as "Higgs field" and the spontaneous symmetry breaking of the field results in generation of gauge boson masses. This mechanism is called as "Higgs mechanism". In this mechanism, the Lagrangian density with the $SU(2)$ scalar doublet field is introduced as

$$\mathcal{L}_\phi = |D_\mu\phi|^2 - V(|\phi|^2), \tag{1.7}$$

where the scalar field ϕ and the scalar potential $V(|\phi|^2)$ are defined as

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^+_R + i\phi^+_I \\ \phi^0_R + i\phi^0_I \end{pmatrix}, \tag{1.8}$$

$$V(|\phi|^2) = \mu^2|\phi|^2 + \lambda|\phi|^4, \tag{1.9}$$

where ϕ^+_R and ϕ^0_R are the real parts of the scalar doublet, ϕ^+_I and ϕ^0_I are the imaginary parts, and $|\phi|^2$ denotes $\phi^\dagger\phi$. μ and λ imply the Higgs mass parameter and the Higgs self-coupling constant, respectively. In order for the field to be stable, λ is greater than zero. The Lagrangian is gauge invariant under the $\phi \rightarrow e^{i\alpha(x)}\phi$ gauge transformation.

Figure 1.2 shows the Higgs potential shapes for $\mu^2 > 0$, and $\mu^2 < 0$. The transition from $\mu^2 > 0$ to $\mu^2 < 0$ is believed to be occurred as the universe gets cold. In the situation of $\mu^2 < 0$, the potential minimum is described as

$$|\phi| = \frac{v}{\sqrt{2}} = \sqrt{-\frac{\mu^2}{2\lambda}}, \tag{1.10}$$

where v is called as vacuum expectation value. There are infinite number of states with minimum energy, as illustrated by the red dashed circle in Figure 1.2. Without breaking the gauge invariance we can arbitrarily select

$$\phi_{vacuum} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{1.11}$$

which is called as "the spontaneously symmetry breaking".

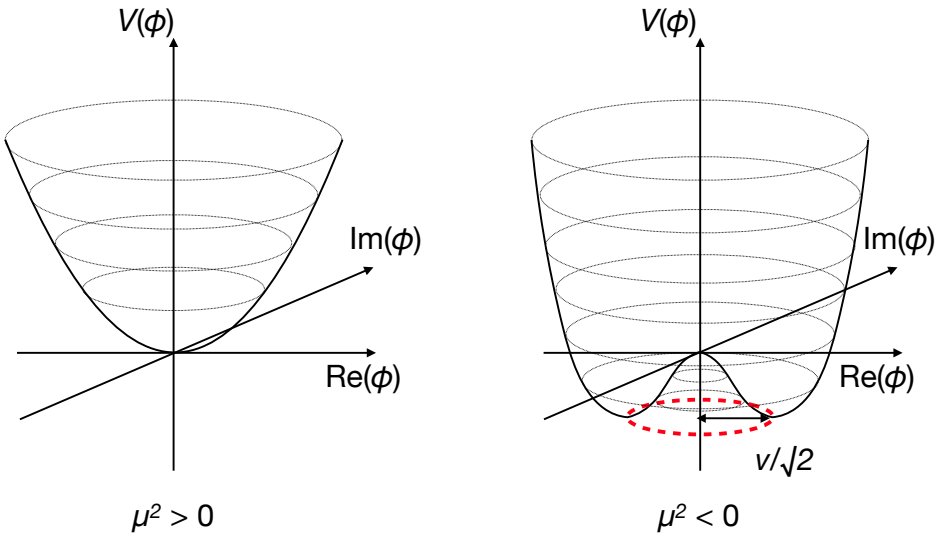


Figure 1.2: Illustration of the Higgs potential for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).

Under this particular vacuum state, the scalar field ϕ is represented by

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1.12}$$

where $H(x)$ is the scalar field corresponding to the Higgs boson. The Higgs potential is described as

$$V = -\frac{\mu^4}{4\lambda} + \frac{1}{2}(-2\mu^2)H^2 + \lambda v H^3 + \frac{\lambda}{4}H^4. \tag{1.13}$$

The second term describes Higgs mass with $M_H = \sqrt{-2\mu^2}$.

The gauge boson masses are expressed in the kinetic term of $\mathcal{L}_\phi$, assuming the simple combination of $SU(2)$ and $U(1)$, which is expressed as $SU(2) \times U(1)$. The covariant derivative D_μ is represented as

$$D_\mu = \partial_\mu + \frac{ig}{2}\boldsymbol{W}_\mu \cdot \boldsymbol{\sigma} + \frac{ig'}{2}A_\mu Y, \tag{1.14}$$

where g and g' are the gauge couplings of $SU(2)$ and $U(1)$, respectively. Y is hyper charge defined as

$$Q = T^3 + \frac{1}{2}Y. \tag{1.15}$$

Then the kinetic term of the Higgs Lagrangian is

$$\begin{aligned} |D_\mu \phi|^2 &= \frac{1}{2}(0, v+H)(\frac{g}{2}\boldsymbol{W}_\mu \cdot \boldsymbol{\sigma} + \frac{g'}{2}A_\mu)^2 \begin{pmatrix} 0 \\ v+H \end{pmatrix} \\ &= \frac{1}{2}M_W^2(W^{+\mu}W_\mu^- + W^{-\mu}W_\mu^+) + \frac{1}{2}M_Z^2 Z^\mu Z_\mu + (terms\ of\ H), \end{aligned} \tag{1.16}$$

with

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \quad Z_\mu = \frac{gW_\mu^3 - g'A_\mu}{\sqrt{g^2+g'^2}}, \tag{1.17}$$

$$M_W = \frac{g}{2}v, \quad M_Z = \frac{\sqrt{g^2+g'^2}}{2}v. \tag{1.18}$$

The gauge field A_μ which is actually a photon is still massless. The mass terms for the $W^\pm$ and Z^0 appear through the symmetry breaking. The coupling terms between Higgs and weak bosons (called as gauge coupling in the following) also appear in the terms of the Higgs field in Equation 1.16;

$$(terms\ of\ H) = \left(\frac{M_W^2}{v^2}(W^{+\mu}W_\mu^- + W^{-\mu}W_\mu^+) + \frac{M_Z^2}{v^2}Z^\mu Z_\mu\right)(2vH + H^2). \tag{1.19}$$

Thus the gauge coupling Y_{HVV} and Y_{HHVV} (V means vector boson; $W^\pm$ or Z^0) are written as

$$Y_{HVV} \propto \frac{M_V^2}{v}, \quad Y_{HHVV} \propto \frac{M_V^2}{v^2}. \tag{1.20}$$

1.1.3 Yukawa coupling

Although the mass of the weak bosons can be explained with the Higgs mechanism, the origin of the fermion mass is not clear. In the SM, the Lagrangian density related to Yukawa coupling is introduced as

$$\mathcal{L}_{Yukawa} = -y\overline{\psi}\phi\psi = -y(\overline{\psi}_L + \overline{\psi}_R)\phi(\psi_L + \psi_R) = -y\overline{\psi}_L\phi\psi_R - y\overline{\psi}_R\phi\psi_L, \tag{1.21}$$

where y is the Yukawa coupling constant. Left-handed fermion fields are expressed by $SU(2)$ doublet and right-handed fields by singlet. Fermions for the i-th generation are written as

$$\psi_{iL} = \begin{pmatrix} u_{iL} \\ d_{iL} \end{pmatrix}, \begin{pmatrix} \nu_{iL} \\ e_{iL} \end{pmatrix}, \quad \psi_{iR} = (u_{iR}), (d_{iR}), (e_{iR}). \tag{1.22}$$

Then Equation 1.21 for leptons (ones for quarks are not explained for its simplicity) are represented as

$$\mathcal{L}^{lep}_{Yukawa} = -y_i\overline{\psi}_{iL}\phi e_{iR} + h.c., \tag{1.23}$$

where y_i is the Yukawa coupling constant for a charged lepton. Again, vacuum state is selected (= spontaneously symmetry breaking) to satisfy $\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+H \end{pmatrix}$ and then the Lagrangian density is represented as

$$\mathcal{L}^{lep}_{Yukawa} = -\frac{1}{\sqrt{2}}y_i v(\overline{e}_{iL}e_{iR} + \overline{e}_{iR}e_{iL}) - \frac{1}{\sqrt{2}}y_i(\overline{e}_{iL}He_{iR} + \overline{e}_{iR}He_{iL}). \tag{1.24}$$

The mass terms are appeared through the Yukawa coupling Y_{Hff} which has the following relation

$$Y_{Hff} \propto \frac{M_f}{v}. \tag{1.25}$$

However, there are no fundamental principles for the introduction of the Yukawa coupling in the SM. Therefore, it is important to confirm that fermions and Higgs boson are truely coupled to understand the origin of the fermion masses.

1.2 Search for the Top-Yukawa interaction at the LHC

The Large Hadron Collider (LHC) is the proton-proton collider with the center-of-mass energy $\sqrt{s} = 13$ TeV, which is the only facility to produce and study Higgs boson in the world. In this section we present how to search for the Top-Yukawa coupling by the ATLAS experiment at the LHC.

1.2.1 ttH process

In order to study the Yukawa coupling, use of the decay of Higgs boson into fermion pair is the most straightforward. However, the Higgs mass $m_H = 125$ GeV is insufficient to decay into top-quark pair. For this reason, we use the Higgs boson production process via Higgs boson and top-quark interaction (ttH process) to measure Top-Yukawa coupling. The diagrams of this process are shown in Figure 1.3. The cross section of the ttH process is estimated by the SM to be 506.5 fb with the NLO QCD calculation [19] at the LHC with $\sqrt{s} = 13$ TeV.

The Top-Yukawa coupling appears also in the other productions and decays of Higgs boson via loop diagrams shown in Figure 1.4.(a)-(d). In the loop diagrams, not only top-quark but also other particles may contribute to the interaction. Thus extraction of the Top-Yukawa coupling through decay processes requires some assumptions, such as SM, to estimate and subtract the contribution from other particles. The measurement of ttH process is the clearest approach to access the Top-Yukawa coupling.

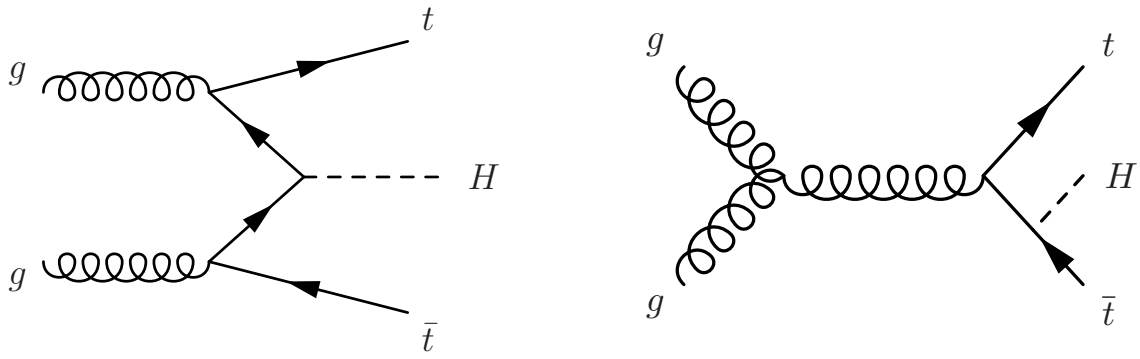


Figure 1.3: Feynman diagrams of the Higgs boson production in association with a top-quark pair.

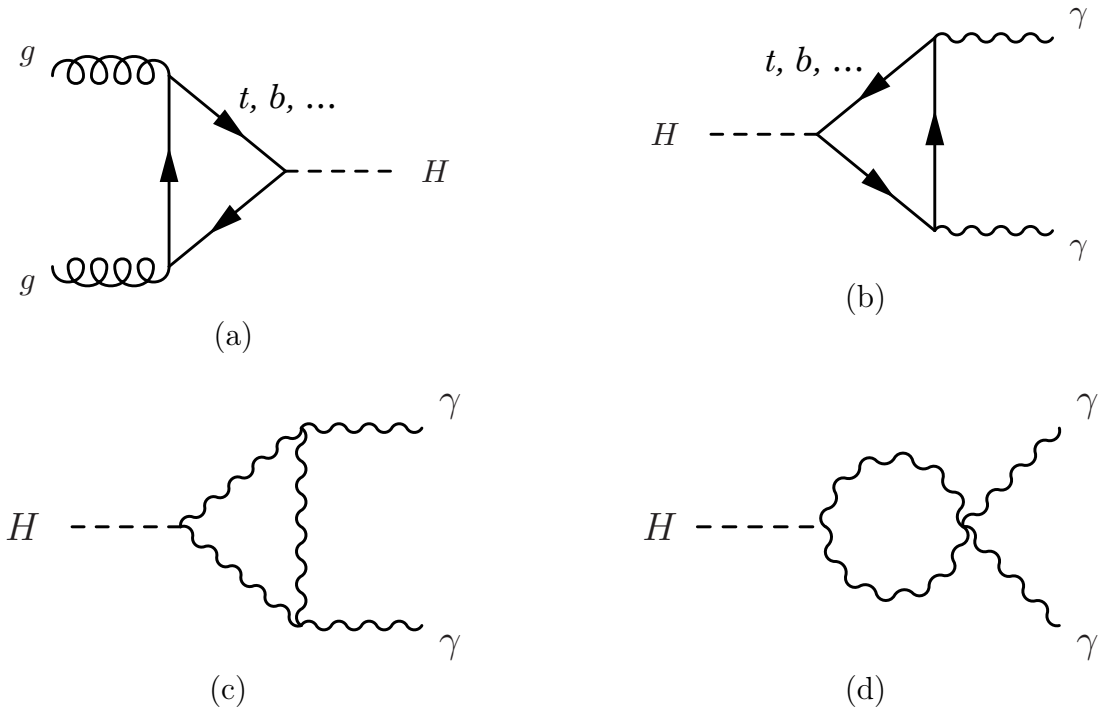


Figure 1.4: Feynman diagrams including the Top-Yukawa coupling. (a) The gluon fusion process contains the top-quark loop while other quarks are also contributed. (b, c, d) The $H \to \gamma\gamma$ decay occurs through the top-quark loop as well as the other fermions and the gauge bosons.

1.2.2 Previous result for each Higgs decay mode

The major Higgs decay channels are $H \to bb$, $H \to WW$, $H \to \tau\tau$, $H \to ZZ$ and $H \to \gamma\gamma$ whose branching ratios are 58 %, 22 %, 6.3 %, 2.6 % and 0.2 %, respectively. The ATLAS experiment published the result of the search for ttH process on December 2017 using 36.1 fb^{-1} of data collected in 2015 and 2016 with $\sqrt{s}$ = 13 TeV [20–23] as shown in Figure 1.5. The signal strength μ_{ttH} is defined as

$$\mu_{ttH} = \frac{\sigma_{ttH}^{obs} \cdot BR^{obs}}{\sigma_{ttH}^{SM} \cdot BR^{SM}}, \tag{1.26}$$

where σ_{ttH}^{obs} is the observed cross section and σ_{ttH}^{SM} is the cross section predicted by SM. The BR^{obs} and BR^{SM} are the branching ratios of each Higgs decay from the measurement and the SM expectation, respectively. For the measurement of μ_{ttH}, BR^{obs} is assumed to be the SM prediction ($BR^{obs} = BR^{SM}$). The μ_{ttH} is measured to be 1.2 ± 0.3 by combining various Higgs decay channels. The background-only hypothesis is excluded at 4.2 standard deviations, while the expectation is 3.8 standard deviations.

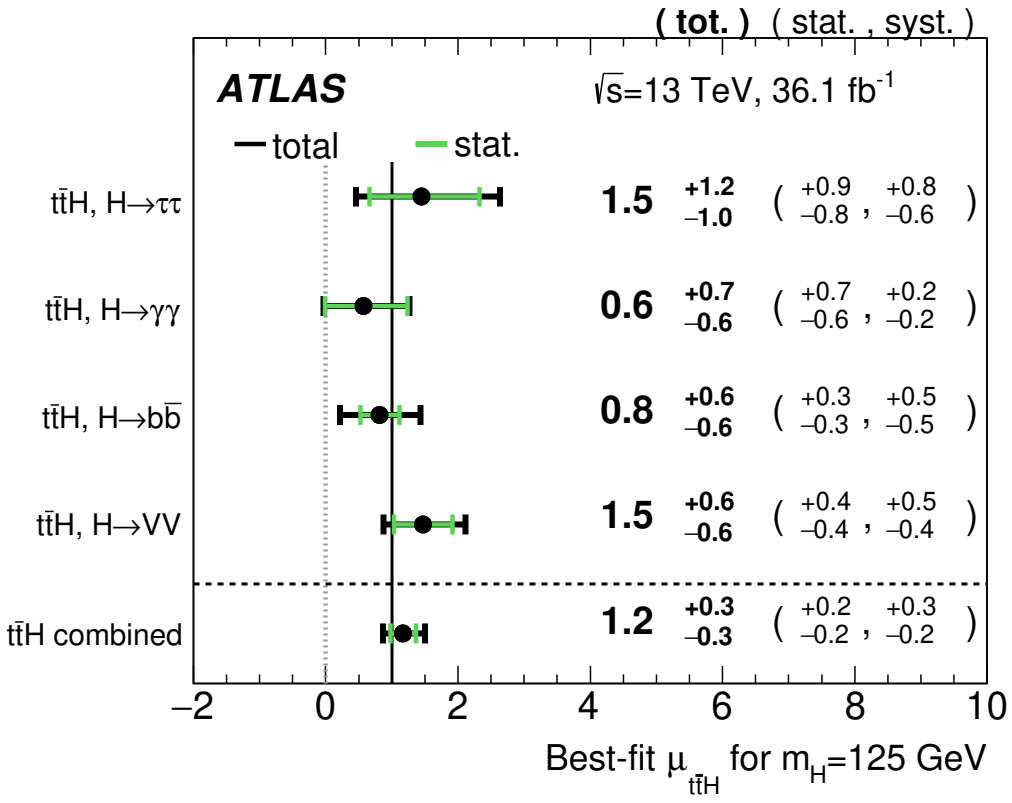


Figure 1.5: The measured value of the signal strength of the ttH process [20–23].

The size of the statistical uncertainty depends on the branching ratio of the Higgs decays. Thus the uncertainty for the $H \to bb$ channel is relatively small while the one for $H \to \gamma\gamma$

is large. In the SM, the number of $H \to \gamma\gamma$ decays in the ttH process is estimated to be 42 events with 36.1 fb^{-1} data. On the other hand, the systematic uncertainty in the $H \to \gamma\gamma$ channel is much smaller than that in the other channels because the clear invariant mass peak in $H \to \gamma\gamma$ mode enables to reduce the uncertainty of the background estimation. Therefore the $H \to \gamma\gamma$ decay mode will be the most sensitive channel once the statistics is increased.

Furthermore, the mass peak from two photons allows to claim the existence of Higgs boson, while one can only claim that the number of events significantly excesses from SM background hypothesis in the other channels. Such excess could be caused by new particles coupled with a top-quark.

For the above reasons, $H \to \gamma\gamma$ decay channel is used in this analysis. In the past search the significance measured by the single $H \to \gamma\gamma$ channel was 0.9 standard deviations with an expectation of 1.7 standard deviations [22].

1.3 Analysis overview

We search for the ttH process where Higgs boson decays into two photons ($ttH(H \to \gamma\gamma)$ channel) using 79.8 fb^{-1} data collected by the ATLAS detector at the LHC. The analysis overview is presented in this section.

1.3.1 Signal final state

The final state of the $ttH(H \to \gamma\gamma)$ channel consists of two photons decayed from a Higgs boson and decay products from a top-quark pair. Since a top-quark decays into b-quark and W boson with $BR(t \to bW) \sim 100\%$, the final state depends on the W decay, which can be classified into two types depending on whether it contains at least one charged lepton or not;

- $t\bar{t}H \to b\bar{b}W^+W^- + \gamma\gamma \to [b\bar{b}q\bar{q}l^+\overline{\nu} + \gamma\gamma]$ or $[b\bar{b}q\bar{q}l^-\nu + \gamma\gamma]$ or $[b\bar{b}l^+l^-\overline{\nu}\nu + \gamma\gamma]$ (56 %)
- $t\bar{t}H \to b\bar{b}W^+W^- + \gamma\gamma \to b\bar{b}q\bar{q}q\bar{q} + \gamma\gamma$ (44 %)

where $q = u, d, c, s$, and the numbers in parentheses are branching fractions. The first decay mode is called Leptonic channel, and the second Hadronic channel in the following.

1.3.2 Background

The characteristics of the final state is to have two photons and multi-jets, in addition to one or two lepton(s) only for the Leptonic channel. The possible background events are:

- $\gamma\gamma$: non-resonant di-photon production together with multi-jets
- $t\bar{t} + \gamma\gamma$: non-resonant di-photon production together with a top-quark pair
- Non-ttH Higgs production: decay of Higgs boson to two photons where the Higgs boson is produced via non-ttH processes

The $\gamma\gamma$ and $t\bar{t}+\gamma\gamma$ events have continuum di-photon invariant mass distributions as shown in Figure 1.6. This continuum background includes events with jets mis-identified as photons. Only for the Leptonic channel, the $\gamma\gamma$ events can be suppressed by requiring the existence of a charged lepton. On contrary, the non-ttH Higgs production events make a peak in the di-photon invariant mass distribution as shown in Figure 1.6. Figure 1.7 shows the diagrams of the non-ttH Higgs productions.

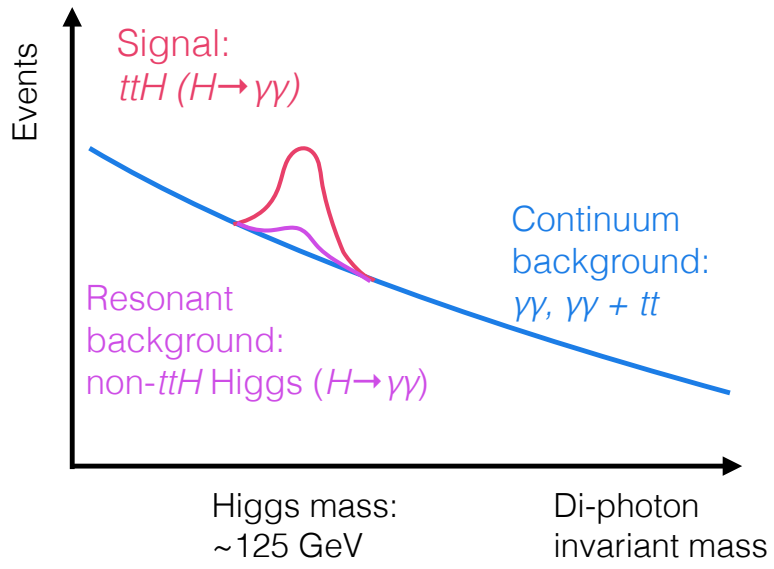


Figure 1.6: Illustration of the di-photon invariant mass distribution for the signal and the resonant and continuum backgrounds.

1.3.3 Event selection

The basic idea of the event selection is to look for a peak by Higgs bosons in di-photon invariant mass distribution in events containing a top-quark pair.

Di-photon selection

Photons are reconstructed from the energy deposit measured by the ATLAS calorimeter. To make the di-photon mass peak clear, good resolution of photon energy, position and decay vertex is required. Furthermore, there are a lot of background photons in the reconstructed objects due to mis-identification of hadronic jets as photons. In order to reject these fake photons, the photon identification is important for this analysis.

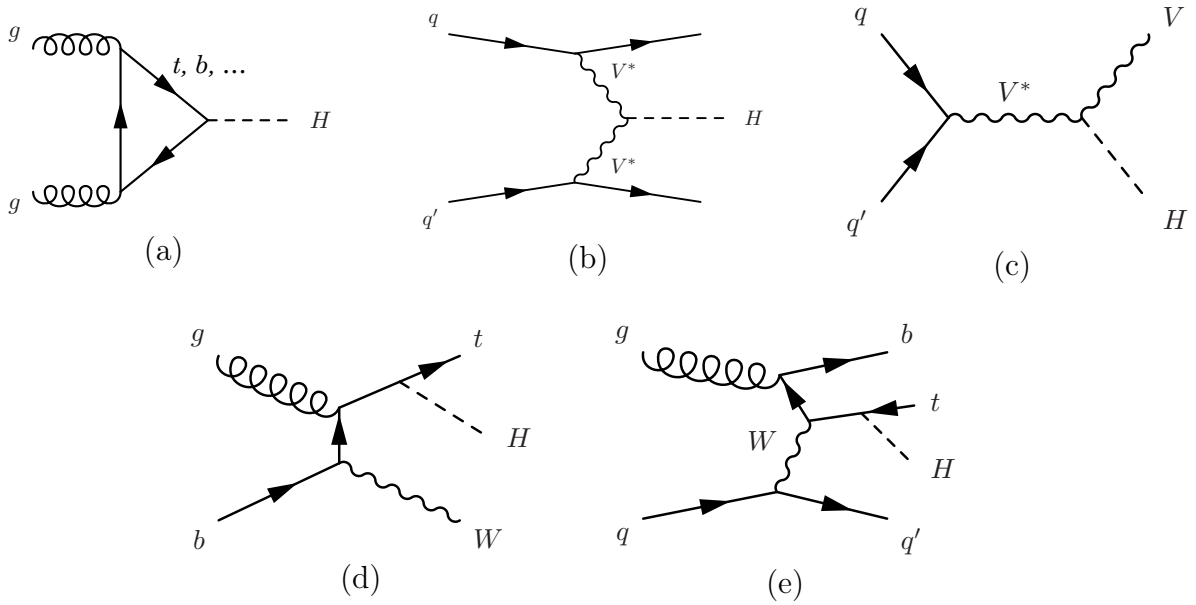


Figure 1.7: Feynman diagrams of the non-ttH Higgs production processes. (a) Gluon fusion process (ggF). (b) Vector boson fusion process (VBF). (d) Vector boson associated production process (VH, $V = W, Z$). (d, e) single top associated production processes (tWH, $tHjb$).

Top-quark pair selection

Additional selection criteria are applied to select events containing a top-quark pair. One of the characteristics of top-quark pair events is to have b-hadrons in the final state. To identify events containing a top-quark pair, b-jet (jet originating from b-hadron) is required to exist. Thus it is important to isolate b-jets from jets originating from other light hadrons.

Event categorization

Events passing both the di-photon and top-quark pair selection criteria are still suffered from $\gamma\gamma$ backgrounds, especially for the Hadronic channel. In order to distinguish signal events from $\gamma\gamma$ backgrounds, kinematic variables (energy, momentum and position) for all objects in the final states are used. These variables are used as inputs of the multivariate analysis (MVA).

In order to improve the signal sensitivity, events are categorized into some groups with different signal to background ratio based on the MVA output. In this analysis, we made three categories for the Leptonic channel and four categories for the Hadronic channel. After the categorization, the numbers of signals and backgrounds are extracted by fitting the di-photon invariant mass distribution.

The remaining of this thesis is organized as follows. The overviews of the LHC and the ATLAS detector are described in Chapter 2. The dataset and the Monte Carlo simulation samples used in this analysis are presented in Chapter 3. The event reconstruction is given in Chapter 4. The event selection and categorization for $ttH(H \to \gamma\gamma)$ using the MVA technique is described in Chapter 5. After the categorization, the extraction of $ttH(H \to \gamma\gamma)$ signals is performed using a statistical procedure. The detail of the signal extraction and its result are presented in Chapter 6. Finally, we discuss our result in Chapter 7 and conclude in Chapter 8.

Chapter 2

LHC-ATLAS experiment

2.1 Large Hadron er (LHC)

The Large Hadron Collider (LHC) [24] is built at the European Organization for Nuclear Research (CERN). The LHC consists of a 27-kilometer ring inside the underground tunnel on France-Switzerland border near Geneva.

Inside the LHC tunnel there are two beam pipes where proton beams are circulated clockwise and anti-clockwise as shown in Figure 2.1. The two beams cross each other at four points which are surrounded by the ATLAS, CMS, ALICE and LHCb detectors, respectively. The center-of-mass energy of proton-proton collisions reaches 13 TeV in 2015, which is the highest collision energy in the world so far.

2.1.1 Design

Figure 2.2 shows the overview of the LHC, including the beam injection systems. The proton injector chain consists of the Linac 2, the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super ProtonSynchrotron (SPS). Table 2.1 lists up the beam energy of each injector component and the LHC.

Table 2.1: Maximum beam energy of the proton injectors and the LHC.

Injector	Maximum beam energy [GeV]
Linac 2	0.05
PSB	1.4
PS	26
SPS	450
LHC	7000

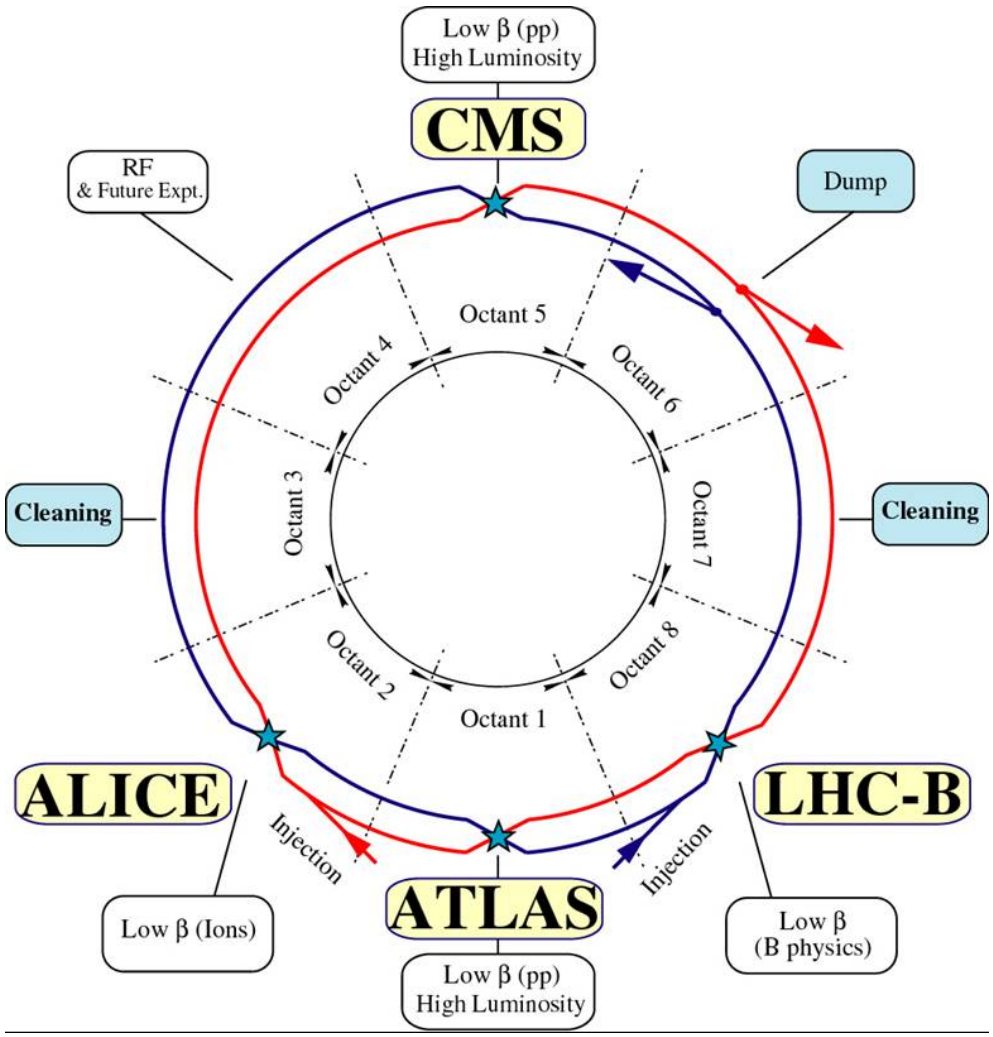


Figure 2.1: Overview LHC schematic [25]. It shows the four main experiments and the two-ring structure.

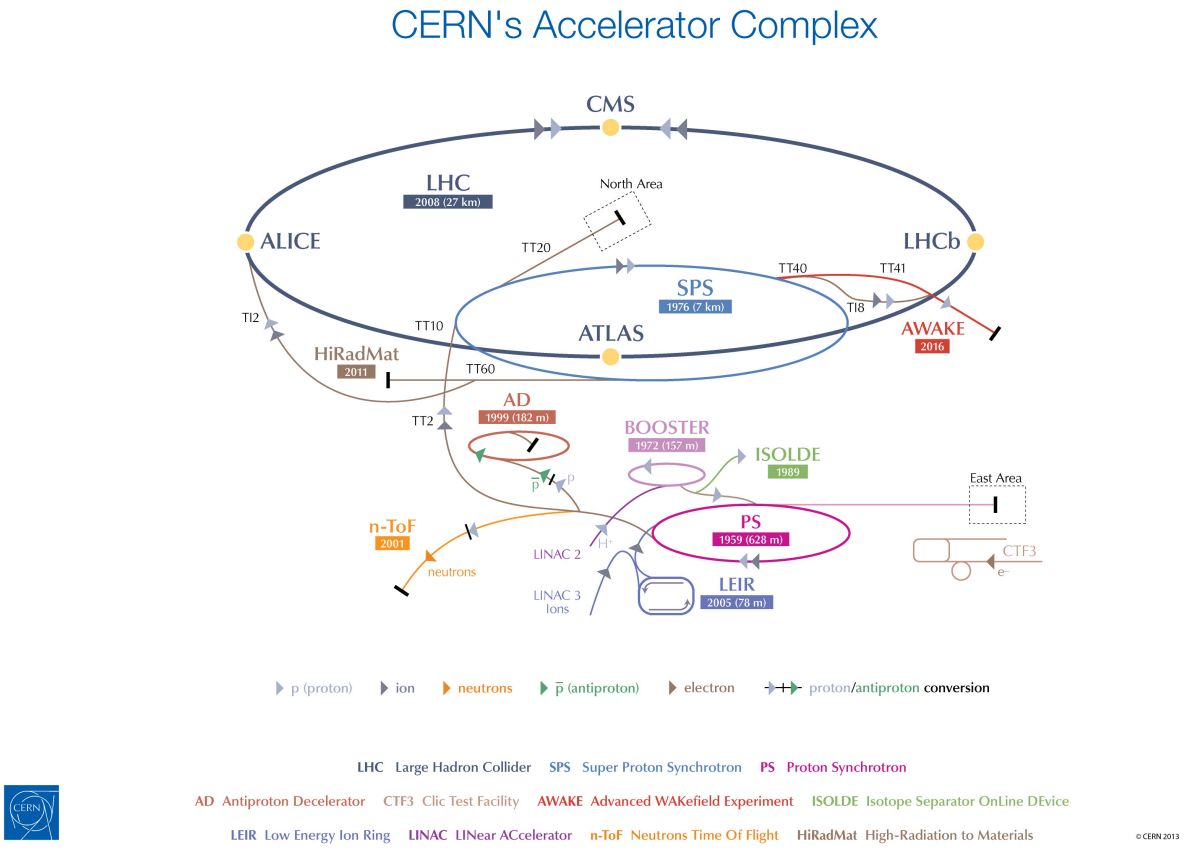


Figure 2.2: Accelerator complex in CERN [26].

The luminosity of a collider with a Gaussian beam distribution is expressed as

$$L = \frac{N_b^2 n_b f_{rev} \gamma}{4\pi\epsilon\beta^*} \times F, \tag{2.1}$$

$$F = \frac{1}{\sqrt{1 + \left(\frac{\theta_c \sigma_z}{2\sigma_{xy}}\right)^2}}, \tag{2.2}$$

where the parameters in this formula are explained in Table 2.2. With the designed LHC beam parameter, the luminosity is 1×10^{34} cm^{-2}s^{-1} with the bunch interval of 25 ns.

Table 2.2: Summary of the LHC beam design and delivered values in 2017.

Parameter	Design value	Value in 2017	Description
N_b	1.15×10^{11}	1.35×10^{11}	the number of protons per bunch
n_b	2808	2544	the number of bunches per beam
f_{rev}	11.245 kHz	11.245 kHz	revolution frequency
γ	7460	6930	Lorentz factor
ϵ	3.75 μm rad	2.9 - 1.6 μm rad	beam emittance
β^*	0.55 m	0.42 - 0.3 m	beta function at the interaction point
θ_c	285 μrad	400 - 180 μrad	beam crossing angle
σ_{xy}	16.6 μm	9 - 6.2 μm	transverse beam size
σ_z	7.55 cm	8.2 - 7.9 cm	longitudinal beam size
E_{beam}	7000 GeV	6500 GeV	beam energy

2.1.2 Performance

Table 2.2 also shows the achieved values in 2017. Thanks to the reduced beta function resulting in smaller beam emittance, the instantaneous luminosity exceeds the design value. Figure 2.3 shows the peak and integrated luminosities recorded in 2015, 2016 and 2017. The peak luminosity reaches 2×10^{34} cm^{-2}s^{-1} in 2017.

Since the ATLAS detector has a limitation on the number of proton-proton interactions per bunch crossing from the view point of data taking speed, the luminosity leveling was introduced in the 2017 run to control peak luminosity. There are three types of leveling techniques:

- Separation leveling
- Crossing angle leveling
- β^* leveling

Figure 2.4 shows one of the proton-proton collision fills in 2017 where the luminosity leveling was taken place for about three hours at the beginning of the fill. The peak luminosities are leveled to 1.5×10^{34} cm^{-2}s^{-1} during 2017.

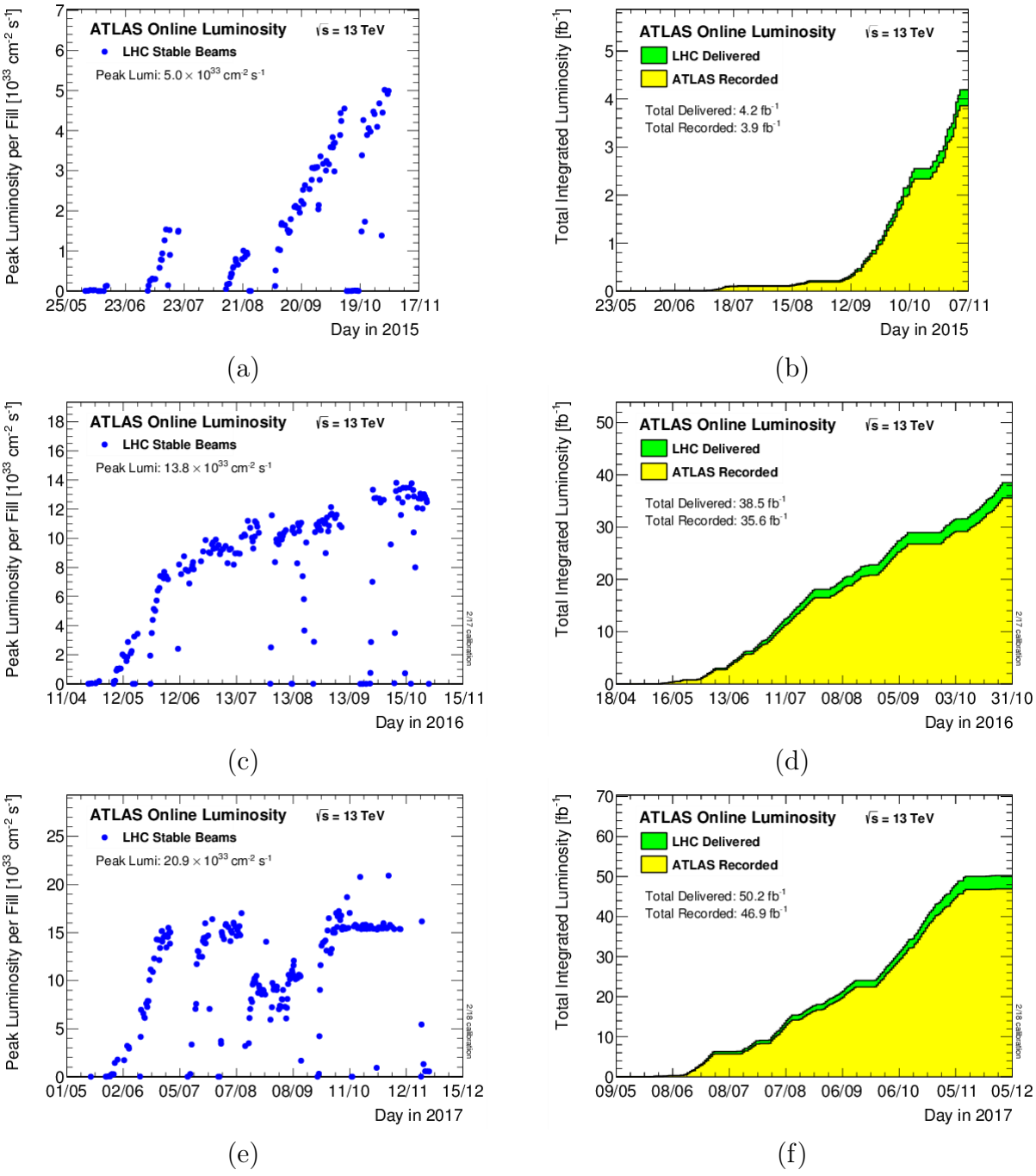


Figure 2.3: Peak and integrated luminosities recorded in 2015, 2016 and 2017 [27].

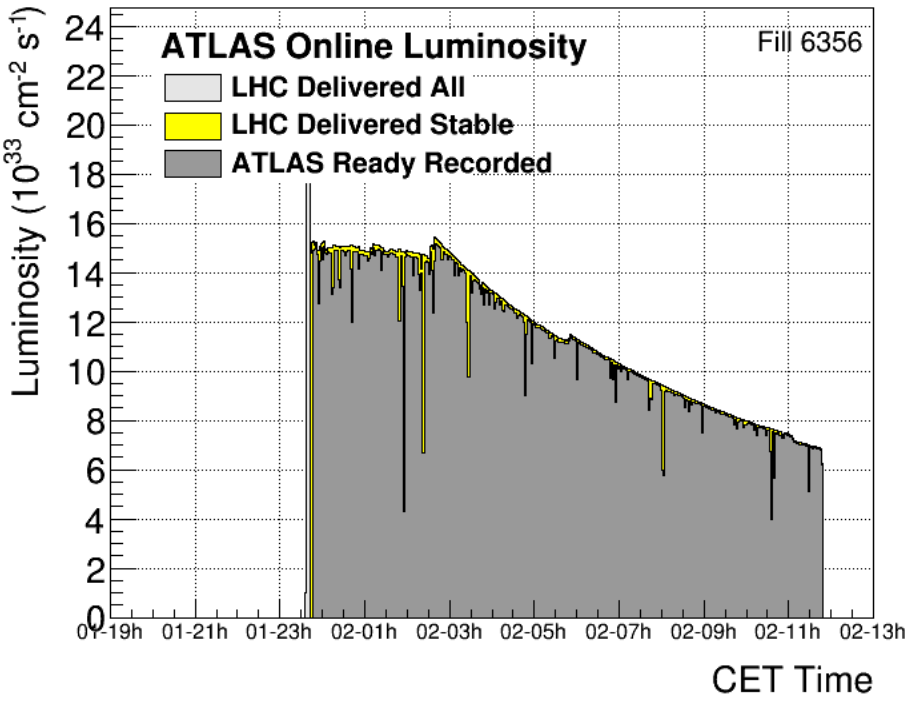


Figure 2.4: One of the proton-proton collision fills in 2017 [28]. The luminosity leveling was taken place from around 02-00h to 02-03h.

The high instantaneous luminosity of the LHC leads to an overlap of multiple proton-proton interactions in a single bunch crossing. This kind of multiple interaction is called "pile-up". The average pile-up or the number of interactions per bunch crossing, μ is expressed as

$$\mu = \frac{L \times \sigma_{inel}}{n_b \times f_{rev}}, \tag{2.3}$$

where L is the instantaneous luminosity, σ_{inel} is the inelastic cross section of the proton-proton collision for $\sqrt{s} = 13$ TeV, n_b is the number of bunches per beam, and f_{rev} is the revolution frequency. Figure 2.5 shows the luminosity-weighted distribution of the mean number of interactions per crossing ($< \mu >$) in 2015, 2016 and 2017 run.

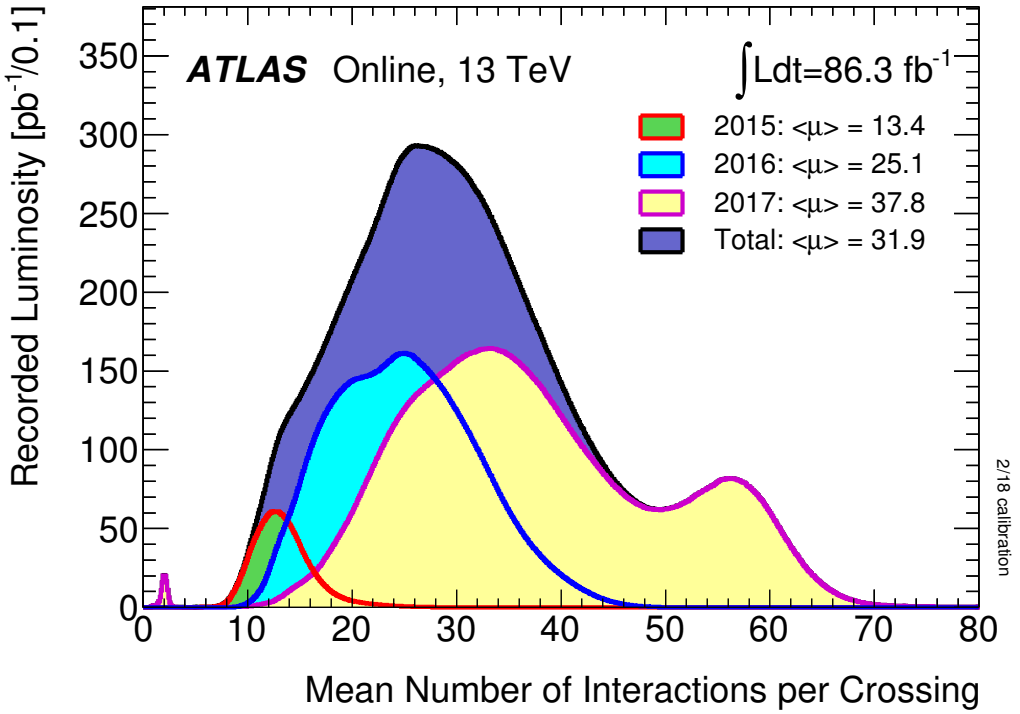


Figure 2.5: Luminosity-weighted distribution of the mean number of interactions per crossing in 2015, 2016 and 2017. The average number of interactions per crossing ($< \mu >$) corresponds to the mean of the poission distribution of the number of interactions per crossing for each bunch [27].

2.2 ATLAS detector overview

The ATLAS (A Toroidal LHC ApparatuS) [3] is the general purpose detector built in one of the collision points of the LHC. The ATLAS detector is forward-backward symmetric with respect to the interaction point. The overall ATLAS detector layout is shown in Figure 2.6.

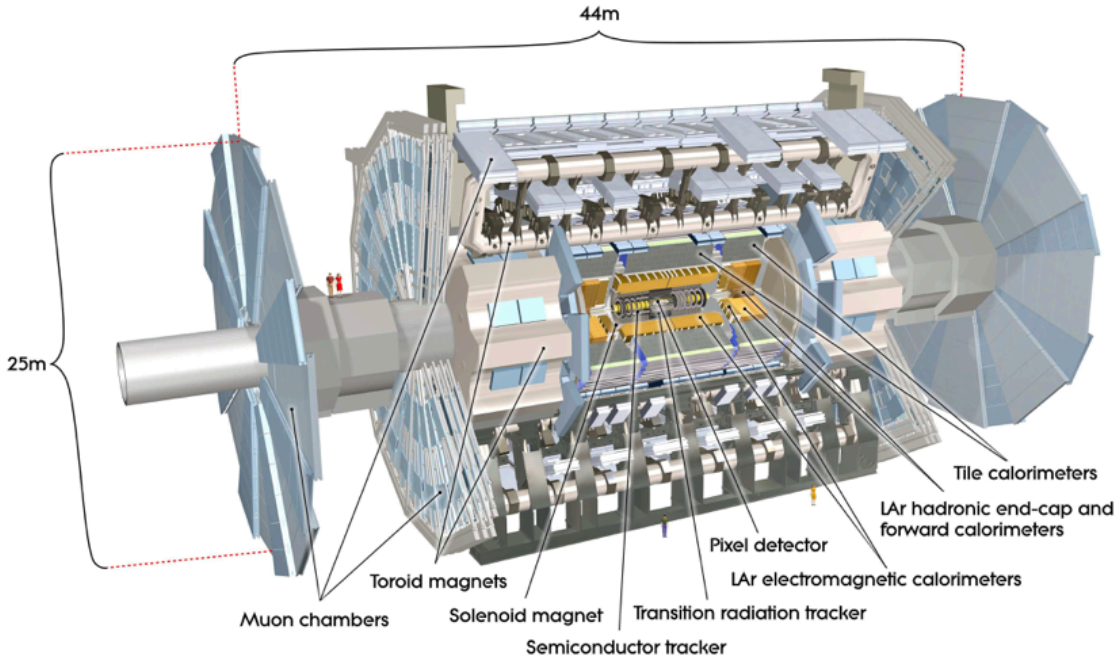


Figure 2.6: Overall ATLAS detector layout [3]. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tonnes.

The Inner Detector consists of the silicon pixel, silicon strip and straw-tube tracking detectors locating just outside of the proton-proton interaction point. A thin superconductive solenoid magnet covers the Inner Detector. The combination of the Inner Detector and the solenoid magnet allows to measure momentum of charged particles. Liquid-argon (LAr) electromagnetic sampling calorimeters are placed outside the solenoid magnet to measure energy of electromagnetic showers. A scintillator-tile (Tile) calorimeter surrounding the LAr calorimeters is used to measure energy of hadron showers. The Inner Detector and all the calorimeters are covered by the muon spectrometers which are composed by four types of chambers and large toroid magnets.

2.2.1 ATLAS coordinate system

The ATLAS group adopts a right-handed coordinate system defining the interaction point as the origin as shown in Figure 2.7. The positive x-axis points to the center of the LHC ring and the positive y-axis points upwards. Instead of the $x - y$ coordinate system, the $r - \phi$

system is also defined as shown in Figure 2.7, where r is the radial distance and ϕ is the azimuthal angle. Furthermore, the rapidity y and the pseudorapidity η are defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \tag{2.4}$$

$$\eta = -\ln \tan \frac{\theta}{2}, \tag{2.5}$$

where E is the particle energy, p_z is the component of the particle momentum along the beam direction and θ is the polar angle.

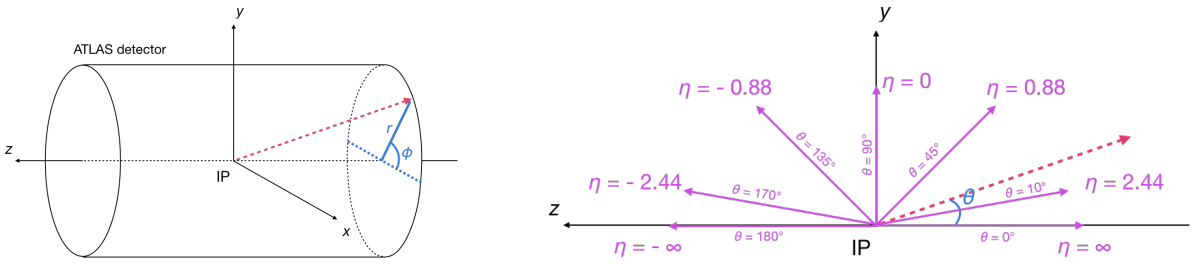


Figure 2.7: ATLAS coordinate system.

Since the phase space of a particle generated by hadron collisions is uniform in the rapidity, we often use the rapidity instead of the polar angle. In the case of a high energy limit (the particle energy is much greater than its mass), the rapidity can be approximated as the pseudorapidity. Since we handle only relativistic particles in the ATLAS experiment, the pseudorapidity is used for its simplicity.

In hadron colliders such as the LHC, the parton-parton collision energy is different event by event. Moreover the center of mass of the parton-parton system is boosted to the beam axis by unknown amount. As a result, the longitudinal momentum of the initial state is completely unknown, which complicates the final state kinematics. On the other hand, the transverse momentum of the initial state is zero and conserves in the final state. For this reason, the transverse momentum p_T and transverse energy E_T defined in Eq. 2.6 are useful and used in this analysis. These are defined as

$$p_T = \sqrt{p_x^2 + p_y^2} = p \sin\theta, \qquad E_T = E \sin\theta. \tag{2.6}$$

Furthermore, the missing transverse energy E_T^{miss} is defined as the momentum imbalance on the $x - y$ plane. More detail is discussed in Section 4.7.

2.3 Inner Detector

The Inner Detector is located at the center of the ATLAS detector, surrounding the beam pipe within $|\eta| < 2.5$. The Inner Detector is composed by the silicon pixel detector, the

silicon strip detector and the transition radiation tracker. The layout of these three detectors is shown in Figure 2.8. These detectors have high granularity to precisely measure the position of charged tracks and its decay vertices. Furthermore, the solenoid magnet is placed on the outside of these three detectors to make curved pathes for charged particles. More description of each detector and magnet is given in this section.

2.3.1 Silicon pixel detector

The silicon pixel detector (Pixel) is located innermost of all the detector subsystems. It is composed by four layers in the barrel and three disks in both the end-cap regions as shown in Figure 2.8. In the original ATLAS detector, the barrel part consisted of three layers. In 2015, the innermost layer, called Insertable B-Layer (IBL) [30], is newly installed. Each layer is composed of modules consisting of the silicon sensors, the readout ASICs and the printed circuit as shown in Figure 2.9.

Table. 2.3 summarizes the numbers of modules and pixels, as well as the pixel size and their depth.

Table 2.3: Summary of silicon pixel modules.

Layer	#Modules	#Pixels / Module	Pixel size [μm^2]	Sensor depth [μm]
1st barrel layer	280	80 × 320 (Planar sensor)	250 × 50	200 (Planar sensor)
(IBL)		80 × 160 (3D sensor)	250 × 50	230 (3D sensor)
2nd barrel layer	286	144 × 328	400 × 50	250
3rd barrel layer	494	144 × 328	400 × 50	250
4th barrel layer	676	144 × 328	400 × 50	250
End-caps (both sides)	288	144 × 328	400 × 50	250
Total	2,024	$\sim$ 88 M		

2.3.2 Silicon strip detector

The silicon strip detector (Semi-Conductor Tracker; SCT) is composed of four layers for the barrel and nine disks for the end-cap regions. Both the barrel and end-caps consist of modules. Each of them has the silicon strip sensors, the readout ASICs and the printed circuit. Figure 2.10 shows the schematic of the SCT module. The SCT uses the single-sided 80 μm pitch micro-strip sensor with the p-in-n technology. Four sensors are mounted on each SCT module. Two 6 cm-long daisy-chained strip sensors are located on each the top and bottom side. They are rotated with each other by 40 mrad to measure two dimensional position of charged particles.

2.3.3 Transition radiation tracker

The Transition Radiation Tracker (TRT) surrounds the SCT detector consisting of one barrel and two end-caps. Both the barrel and the end-caps are composed of polyimide straw drift tubes whose diameter is 4 mm. There is a tungsten gold-plated anode wire in the center

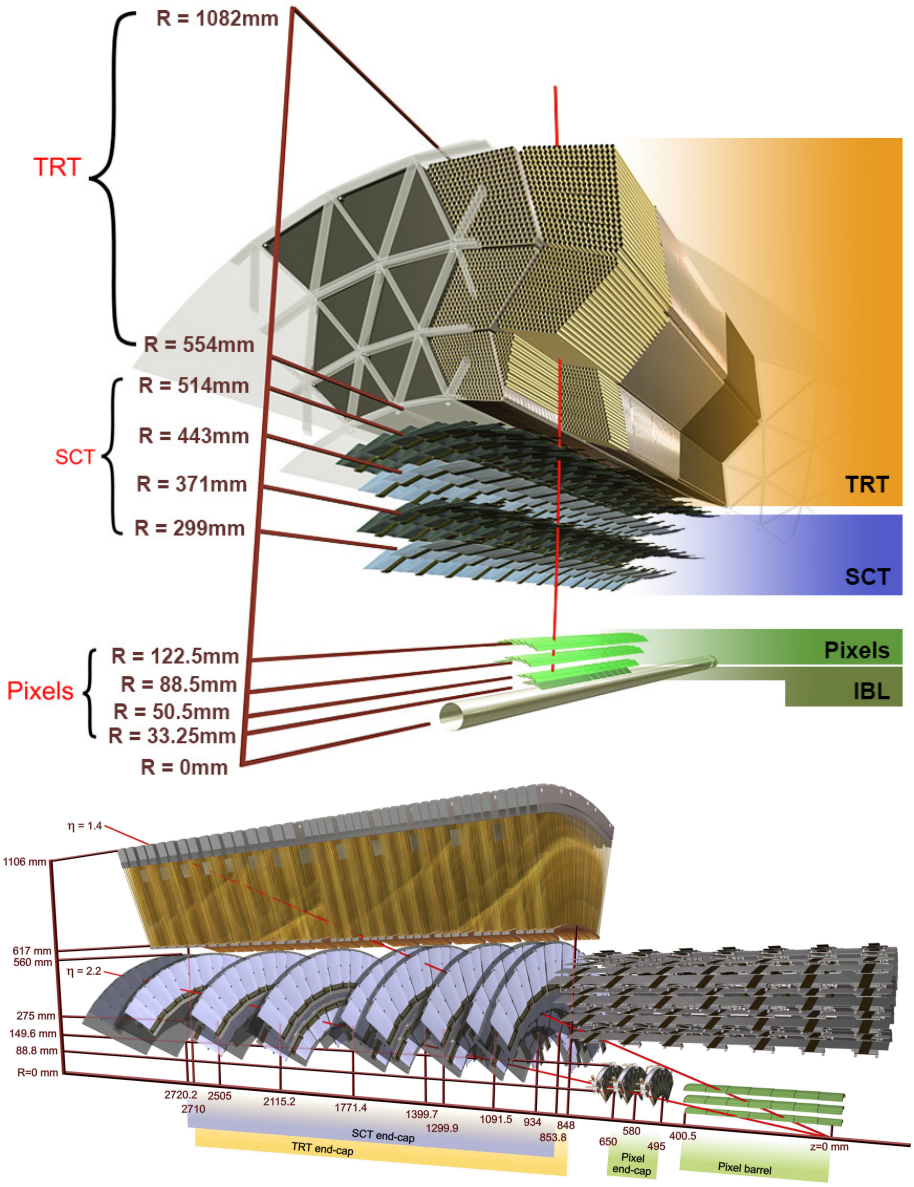


Figure 2.8: Overview of the barrel and end-cap parts of the Inner Detector consisting of the silicon pixel (IBL and Pixel), the silicon strip (SCT) and the transition radiation tube (TRT) detectors [29].

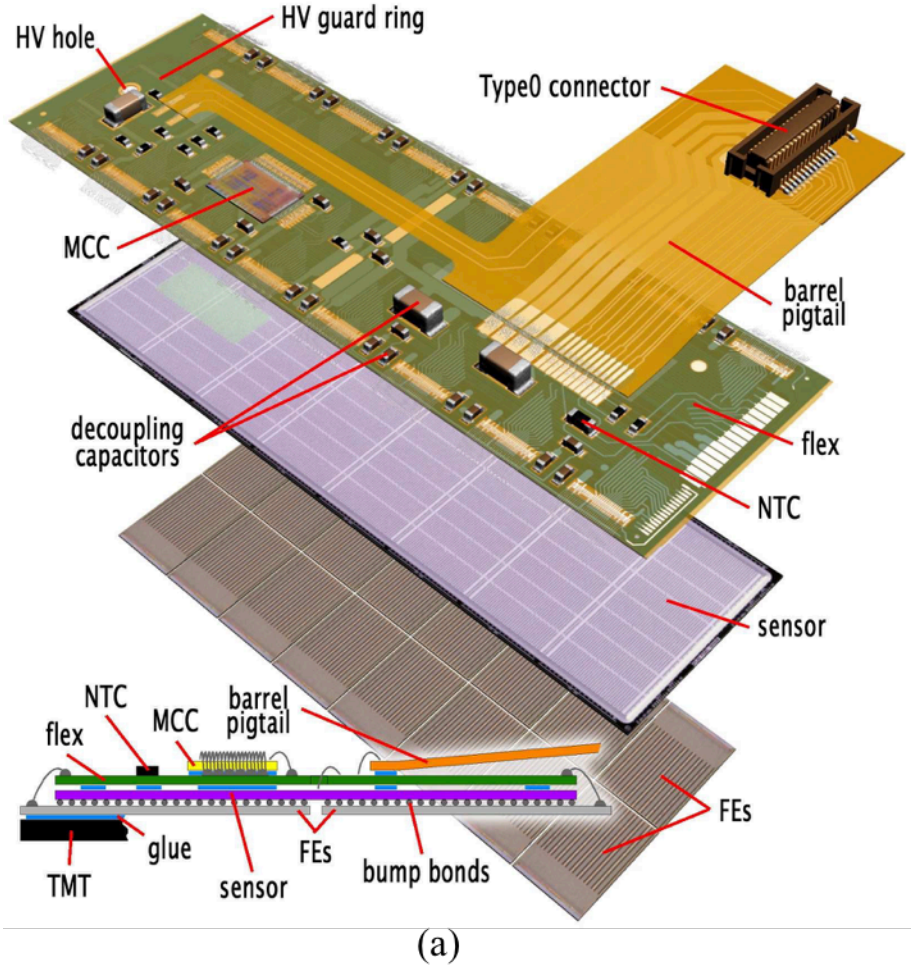


Figure 2.9: Schematic of the silicon pixel module [3] consisting of the silicon sensor, the readout ASICs and the printed circuit.

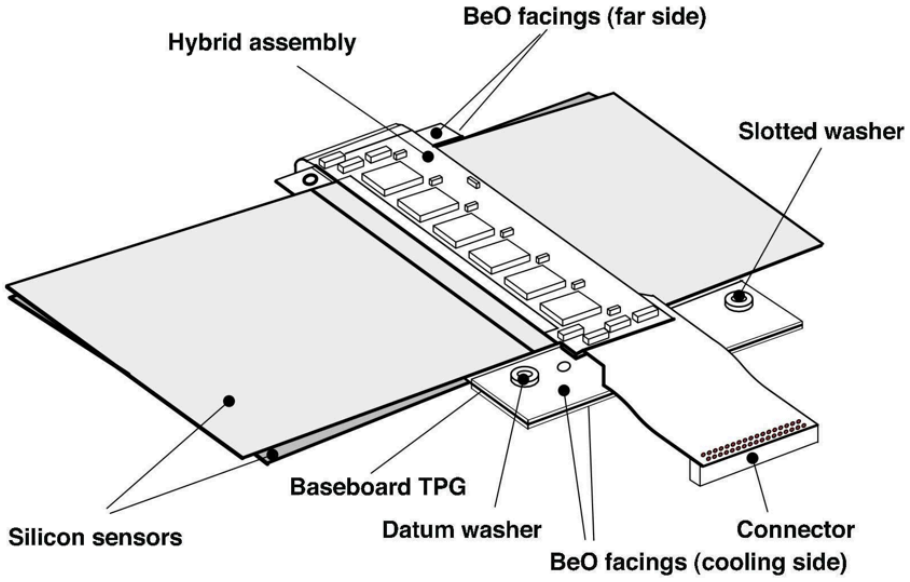


Figure 2.10: Schematic of the SCT module [3].

of the straw as shown in Figure 2.11. These straws are filled with a gas mixture of 70 % Xe, 27 % CO_2 and 3 % O_2. The TRT covers roughly up to 1100 mm in the r direction and 2700 mm in the z direction. The TRT provides only $r-\phi$ information of charged particle tracks. However, the combination of precise tracking by the silicon pixel and strip detectors with the TRT gives robust pattern recognition in both $r-\phi$ and z coordinates due to TRT's large volume.

In addition, the TRT provides a discrimination between electrons and pions using the transition radiation effect. Transition radiations from electrons are emitted in the radiator filled with the space between straws. The radiated photons are detected by the straws.

2.3.4 Solenoid magnet

The silicon detectors and TRT are surrounded by a solenoid magnet whose photograph is displayed in Figure 2.12. The solenoid magnet generates 2 T magnetic field. It is 5.3 m long, 2.4 m diameter, 4.5 cm thick and 5 tonne weight. In order to achieve such a strong magnetic field, the superconducting electromagnet which is composed by Al-stabilized NbTi conductor is used.

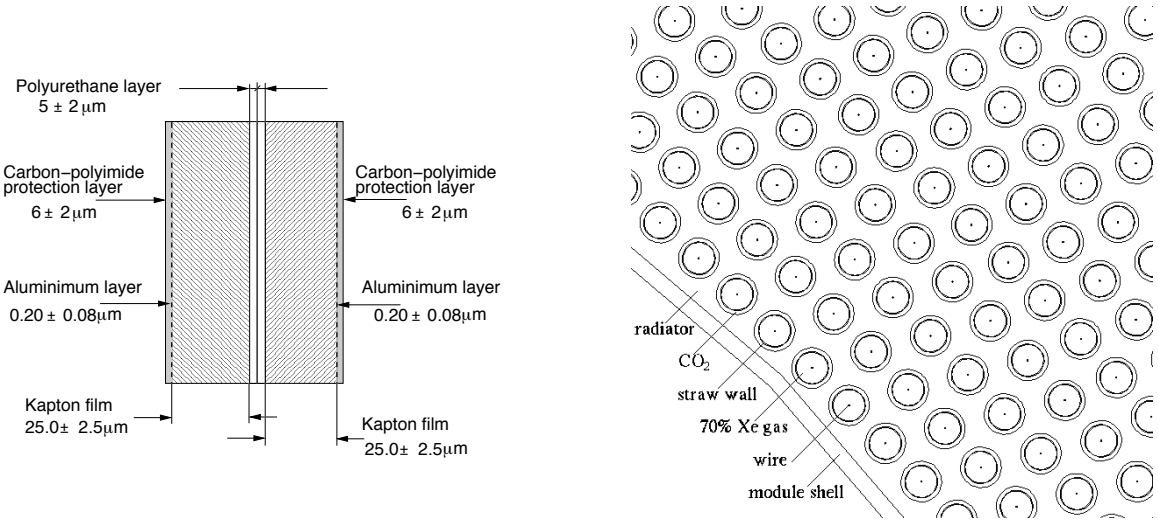


Figure 2.11: TRT tube [31] and cross section of the TRT barrel part [32].

Figure 2.12: ATLAS central solenoid magnet [33].

2.4 Calorimeters

There are two types of calorimeter surrounding the Inner Detector. One is the electromagnetic (EM) calorimeter placed just outside of the solenoid magnet. The other is the hadronic calorimeter outside the EM calorimeter. Figure 2.13 shows the cut-away view of the calorimeter system. Table 2.4 shows the summary of the parameters of the calorimeter system.

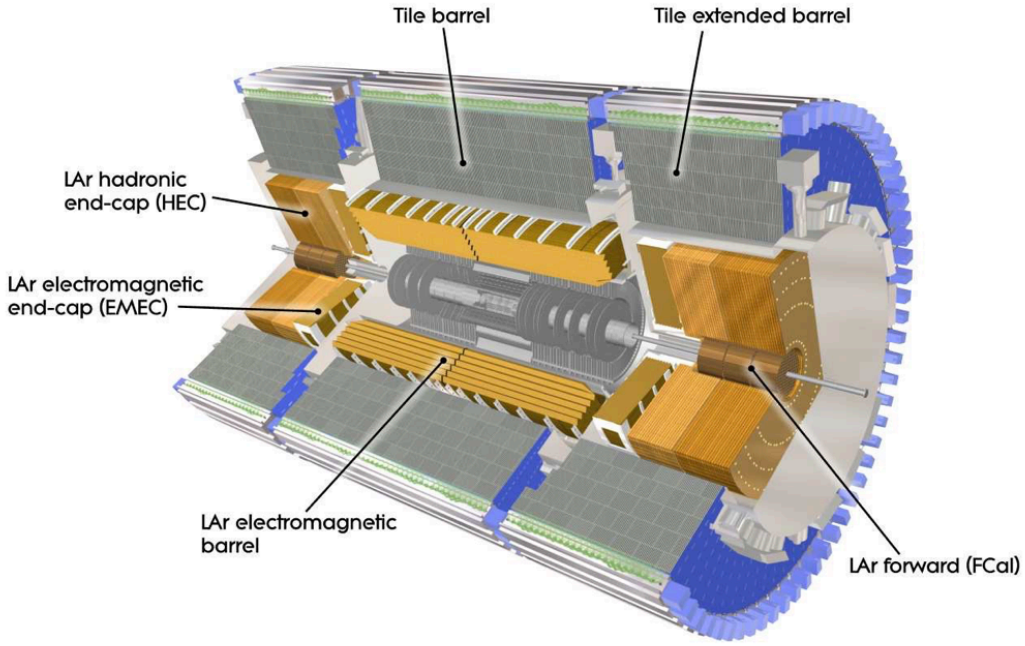


Figure 2.13: Cut-away view of the ATLAS calorimeter system [3].

2.4.1 EM calorimeter

The EM calorimeter consists of lead absorbers, liquid-argon (LAr), and electrodes. This detector has an accordion shape as shown in Figure 2.14. This shape allows full coverage in ϕ without any insensible cracks and fine segmentation along η direction.

As shown in Figure 2.14, the EM calorimeter is segmented into three layers in depth. The first layer is finely segmented along η. The second layer, on the other hand, is coarser than the first layer but it has large volume in the longitudinal direction, which allows to collect the large fraction of the energy deposit. The third layer is placed to detect only the tail of the EM shower, thus longitudinal length is short and the segmentation along η is rough. This high granularity arrangement in the η direction allows precise shower shape analysis especially for the photon and electron reconstruction. In addition, the presampler detector, which consists of liquid-argon, is placed in front of the first layer only for the barrel part.

Table 2.4: Summary of the parameters of the calorimeter system [3].

	Barrel		**End-cap**	
EM calorimeter				
Number of layers and $\|\eta\|$ coverage				
Presampler	1	$\|\eta\| < 1.52$	1	$1.5 < \|\eta\| < 1.8$
Calorimeter	3	$\|\eta\| < 1.35$	2	$1.375 < \|\eta\| < 1.5$
	2	$1.35 < \|\eta\| < 1.475$	3	$1.5 < \|\eta\| < 2.5$
			2	$2.5 < \|\eta\| < 3.2$
Granularity $\Delta\eta \times \Delta\phi$ versus $\|\eta\|$				
Presampler	0.025×0.1	$\|\eta\| < 1.52$	0.025×0.1	$1.5 < \|\eta\| < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$\|\eta\| < 1.40$	0.050×0.1	$1.375 < \|\eta\| < 1.425$
	0.025×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.1	$1.425 < \|\eta\| < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \|\eta\| < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \|\eta\| < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \|\eta\| < 2.4$
			0.025×0.1	$2.4 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 2nd layer	0.025×0.025	$\|\eta\| < 1.40$	0.050×0.025	$1.375 < \|\eta\| < 1.425$
	0.075×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.025	$1.425 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 3rd layer	0.050×0.025	$\|\eta\| < 1.35$	0.050×0.025	$1.5 < \|\eta\| < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	
LAr hadronic end-cap				
$\|\eta\|$ coverage			$1.5 < \|\eta\| < 3.2$	
Number of layers			4	
Granularity $\Delta\eta \times \Delta\phi$			0.1×0.1	$1.5 < \|\eta\| < 2.5$
			0.2×0.2	$2.5 < \|\eta\| < 3.2$
Readout channels			5632 (both sides)	
LAr forward calorimeter				
$\|\eta\|$ coverage			$3.1 < \|\eta\| < 4.9$	
Number of layers			3	
Granularity $\Delta x \times \Delta y$ (cm)			FCal1: 3.0×2.6	$3.15 < \|\eta\| < 4.30$
			FCal1: $\sim$ four times finer	$3.10 < \|\eta\| < 3.15$,
				$4.30 < \|\eta\| < 4.83$
			FCal2: 3.3×4.2	$3.24 < \|\eta\| < 4.50$
			FCal2: $\sim$ four times finer	$3.20 < \|\eta\| < 3.24$,
				$4.50 < \|\eta\| < 4.81$
			FCal3: 5.4×4.7	$3.32 < \|\eta\| < 4.60$
			FCal3: $\sim$ four times finer	$3.29 < \|\eta\| < 3.32$,
				$4.60 < \|\eta\| < 4.75$
Readout channels			3524 (both sides)	
Scintillator tile calorimeter				
	Barrel		Extended barrel	
$\|\eta\|$ coverage	$\|\eta\| < 1.0$		$0.8 < \|\eta\| < 1.7$	
Number of layers	3		3	
Granularity $\Delta\eta \times \Delta\phi$	0.1×0.1		0.1×0.1	
Last layer	0.2×0.1		0.2×0.1	
Readout channels	5760		4092 (both sides)	

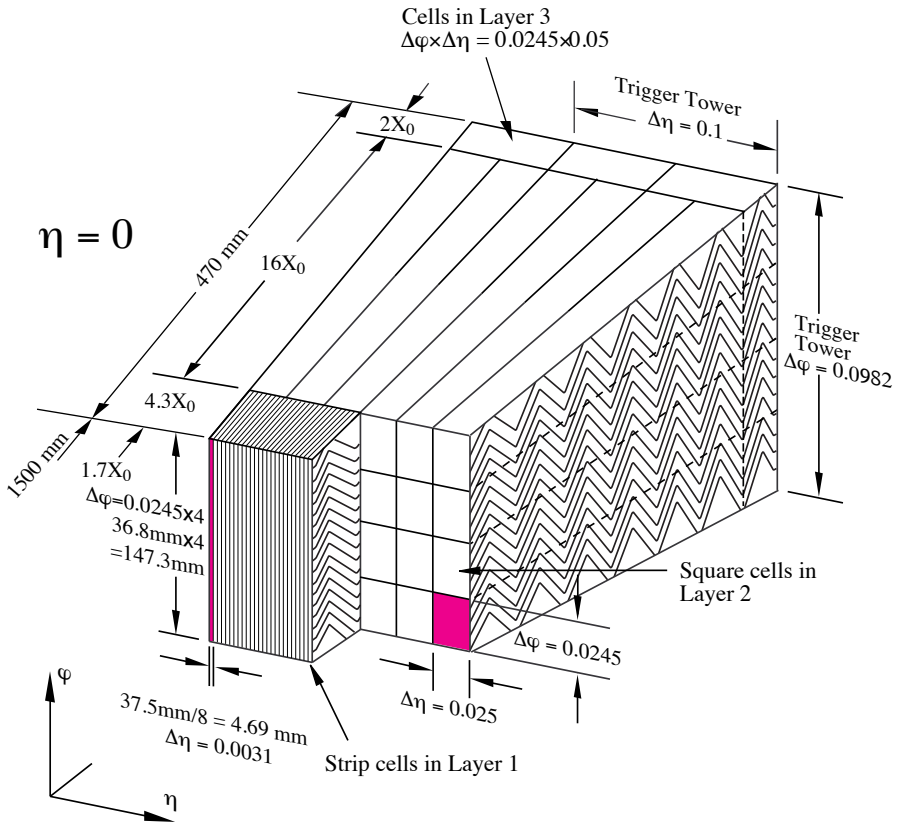
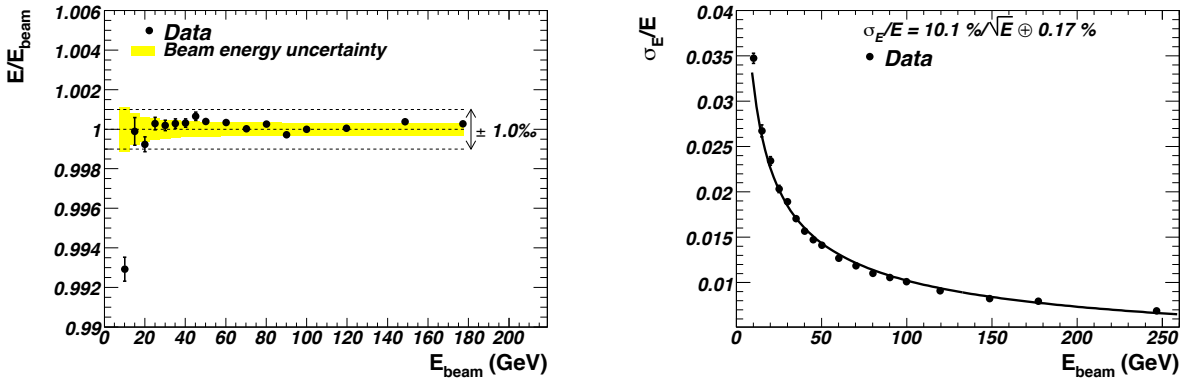


Figure 2.14: Segmentation of the LAr calorimeter [3].

The total thickness of the EM calorimeter is > 22 radiation length (X_0) in the barrel, and > 24 X_0 in the end-caps.

The linearity of response and the fractional energy resolution of the EM calorimeter to electron is measured by beam tests [3], whose result is shown in Figure 2.15. In the energy range of 15 - 180 GeV, the reconstructed energy response is linear within ± 0.1 %. The energy resolution was measured to be $\sigma_E/E = 10/\sqrt{E(\mathrm{GeV})} \oplus 0.17$ %.



(a) (b)

Figure 2.15: Performance of the EM calorimeter using electron and positron beams [3]. (a) Linearity of response as a function of the electron beam energy, E_{beam}. (b) Fractional energy resolution as a function of the electron beam energy, E_{beam}.

2.4.2 Hadronic calorimeter

There are three types of hadronic calorimeters in the ATLAS. One is the LAr hadronic end-cap calorimeter consisting of copper/LAr placed in the region of $1.5 < |\eta| < 3.2$. Next is the LAr forward calorimeter which consists of copper-tangsten/LAr covering $3.1 < |\eta| < 4.9$. The last one is the Tile calorimeter placed in the barrel region. It is a sampling calorimeter with a steel-scintillator sandwich structure. Figure 2.16 shows the schematic of the Tile calorimeter.

The Tile calorimeter has approximately 7.4 interaction length in depth. The signal from the scintillator is read out by photomultiplier tubes through wave length shifting fibers.

2.5 Muon spectrometer

The muon spectrometer is the outermost detector of the ATLAS as shown in Figure 2.17. Basically only muons can reach this region because most particles are absorbed at the calorimeter except for punch-through particles and neutrinos. The spectrometer is composed of four types of gas chambers: Monitored drift tubes (MDT), Cathode strip chambers (CSC), Resistive plate chambers (RPC) and Thin gap chambers (TGC). The MDT and CSC are used for

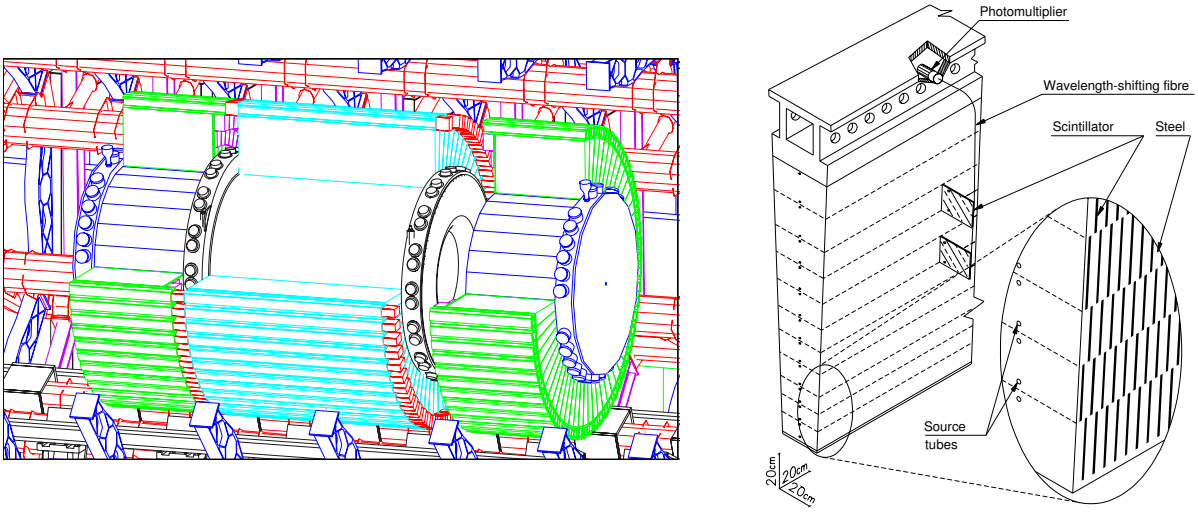


Figure 2.16: Overview of the Tile calorimeter [34] and its component [3].

the precise tracking. The MDT covers $|\eta| < 2.7$ except for the innermost layer which covers $|\eta| < 2.0$. The CSC is located at the region of $2.0 < |\eta| < 2.7$. The RPC and TGC are used for the triggering. The RPC covers $|\eta| < 1.05$ while the TGC covers $1.05 < |\eta| < 2.7$ region.

Three large air-core toroids generate the magnetic field to measure the muon momentum as shown in Figure 2.17. The toroid system is composed of one barrel and two end-caps. Each of them has eight segmented toroidal coils.

2.6 Luminosity detectors

There are two types of detectors in the ATLAS to measure the luminosity [35]; the LUminosity measurement using a Cherenkov Integrating Detector (LUCID) and the Beam Conditions Monitor (BCM).

2.6.1 LUCID

The main purpose of the LUCID is to measure the luminosity by detecting inelastic proton-proton scatterings. It also provides the online monitoring of the instantaneous luminosity and beam conditions. The current version of the LUCID was installed just before 2015 run. It consists of quartz as the Cherenkov medium and photon multipliers to detect charged tracks from the collision point. Figure 2.18 shows the LUCID detector system.

2.6.2 BCM

The BCM is composed of four 8×8 mm^2 diamond sensors, placed around the beam pipe at $|z| = \pm$ 1.84 m on each side of the ATLAS interaction point. The diamond sensor is selected for its radiation hardness.

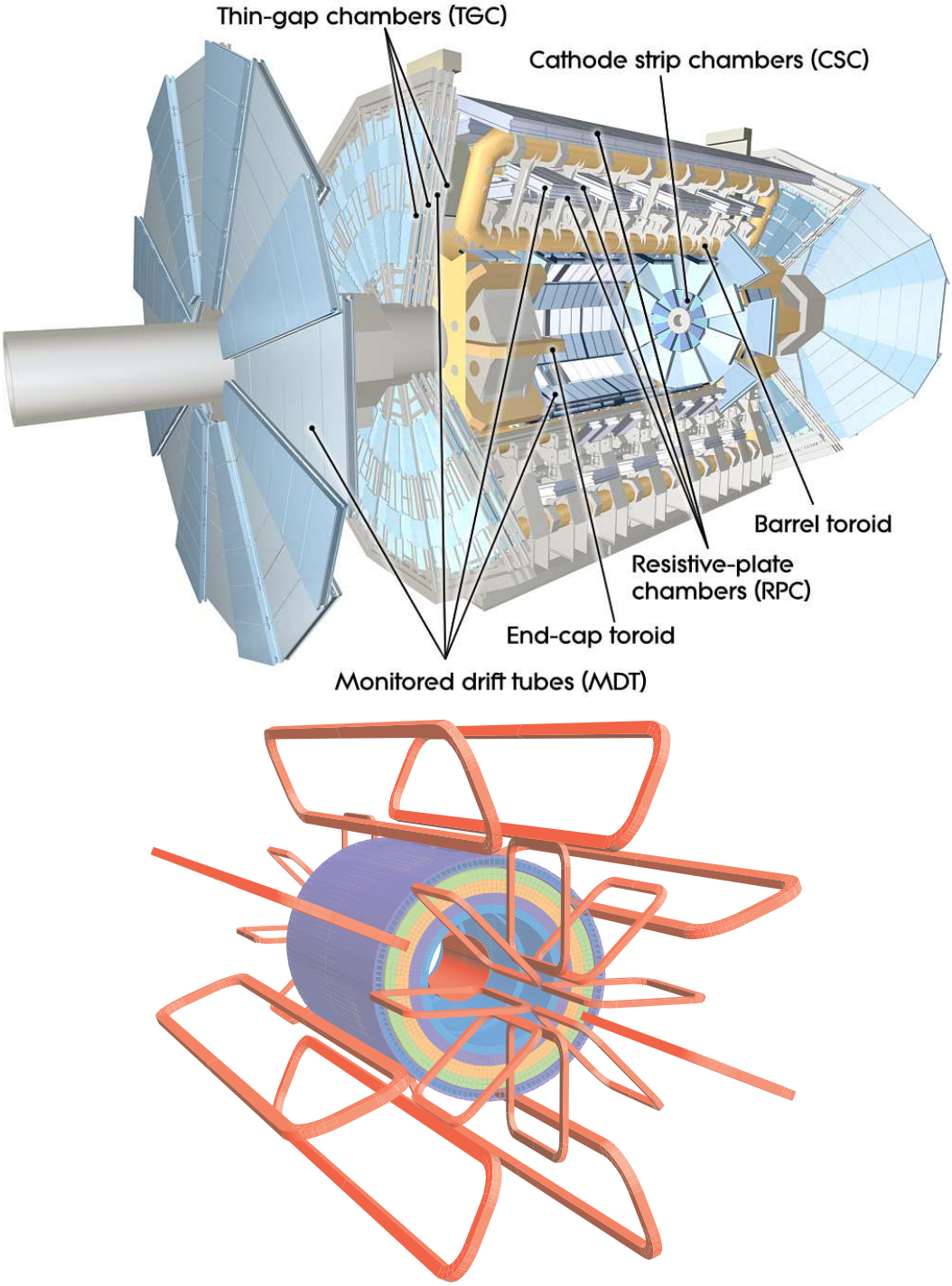


Figure 2.17: Muon spectrometer and toroid systems [3].

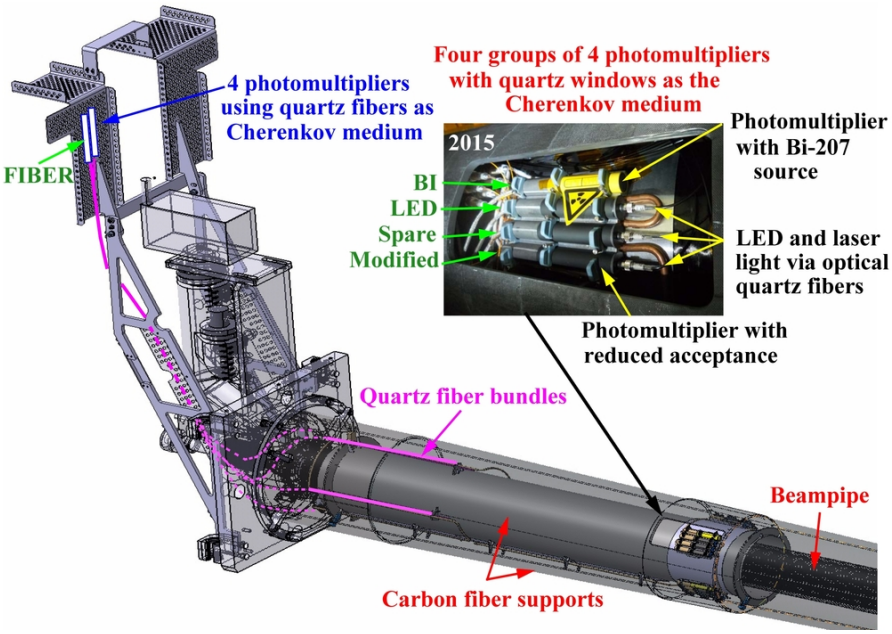


Figure 2.18: LUCID detector system [36].

2.7 Trigger system

All the collision data cannot be recorded because the event rate of LHC collisions is too high. Only interesting collision data are selected and stored using the Trigger and Data Acquisition (TDAQ) system [37, 38]. Figure 2.19 shows the schematics of the TDAQ system. The trigger system consists of a hardware-based first level trigger (Level-1) and a software-based second level trigger (High Level Trigger: HLT). A Level-1 trigger signal is generated when a particular event is occurred in a collision (e.g. there are two high p_T photons in case of the analysis in this thesis), based on either the calorimeter or the muon spectrometer. After that the trigger signal is sent to the central trigger processor (CTP). Then the CTP provides a trigger signal for all the ATLAS sub-detectors. A trigger signal is generated beam bunch by bunch. Since it is impossible to make a trigger decision within a bunch crossing interval (25 ns), bunch by bunch collision data for each detector must be retained in the pipeline memories on the front-end detector electronics until the trigger decision is made by the CTP. Once each subsystem receives a trigger signal, data taken in the corresponding collision bunch is sent to the HLT. Only events passing the HLT selection are sent to the data storage.

2.7.1 Level-1 trigger

The Level-1 trigger decision is made by the calorimeter and the muon spectrometer. From the calorimeter, triggers related to photons, electrons, taus, jets and missing transverse energy are generated while triggers related to muons are generated by the muon spectrometer. In the

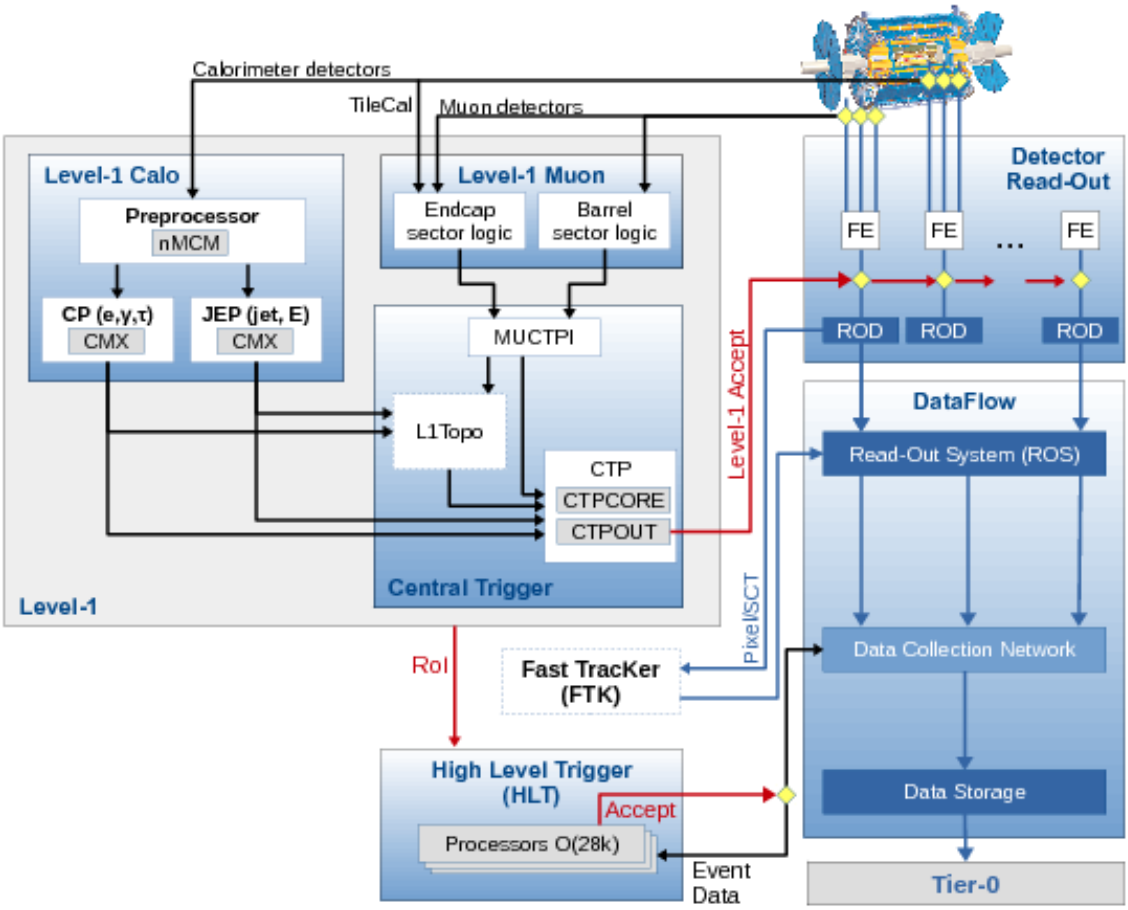


Figure 2.19: The ATLAS TDAQ system [38] consisting of a hardware-based first level trigger (Level-1) and a software-based second level trigger (HLT). The Level-1 trigger decision is formed by the Central Trigger Processor (CTP), which receives inputs from Level-1 calorimeter and Level-1 muon triggers. After the Level-1 trigger acceptance, the events are processed by the HLT using the Region-of-Interest (RoI) from the Level-1.

$ttH(H \to \gamma\gamma)$ analysis, events with at least two photons are selected by the trigger system (Di-photon trigger). Since the ATLAS is the general purpose detector, not only two photons but also the combinations of high p_T electrons, muons, jets and so on are used to trigger events for the other analyses. The Level-1 trigger output rate is up to 100 kHz. The Level-1 trigger latency which is the time from the collision until an event is accepted by the Level-1 trigger system is set at 2.5 μs. In order to achieve the latency of 2.5 μs, the Level-1 trigger system is implemented in the fast custom electronics.

Furthermore, The Level-1 trigger defines Regions of Interest (RoI's). Only the data associated with the RoI's are used in the HLT analysis.

Photon (Electron) trigger

Data collected by the Di-photon trigger are used in this analysis. The trigger algorithm identifies a cluster consisting of 2×2 EM trigger towers as shown in Figure 2.20. Here one trigger tower is defined as the size in $\eta - \phi$ plane of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in most parts. In each cluster, at least one vertical or horizontal energy sums of the neighboring towers in the cluster (see Figure 2.20) must exceed a certain threshold. In order to require isolation, the energy sum of 12 EM towers surrounding 2×2 EM seed towers must be below a threshold. Moreover, 4×4 hadronic calorimeter trigger towers behind EM towers must be below a threshold to suppress hadron contaminations.

The 2×2 EM seed tower is defined as the coordinates of the electron/photon RoI. At the stage of the Level-1 trigger, photons and electrons cannot be distinguished because there are no charged track information which should be provided by the Inner Detector. The photon-electron separation is responsible for the HLT.

2.7.2 High Level Trigger (HLT)

RoI's defined by the Level-1 trigger are sent to the HLT in which precise selection algorithms run with the full granularity detector information for well-separated objects such as photon, electron, muon and tau. Only the case of jets and global event quantities (e.g. E_T^{miss}) reconstruction, the full calorimeter information is used. In the HLT algorithm, the object reconstruction is performed within a processing time of about 200 ms using large amount of computing resources. In the Di-photon trigger, photons are reconstructed using energy deposits in the calorimeter and charged particle tracks provided by the Inner Detector. The reconstructed objects at the stage of the HLT are used only for the triggering. The offline object reconstruction used for the physics analysis is performed independently on the HLT. The HLT reduces the event rate from 100 kHz to approximately 1 kHz in average.

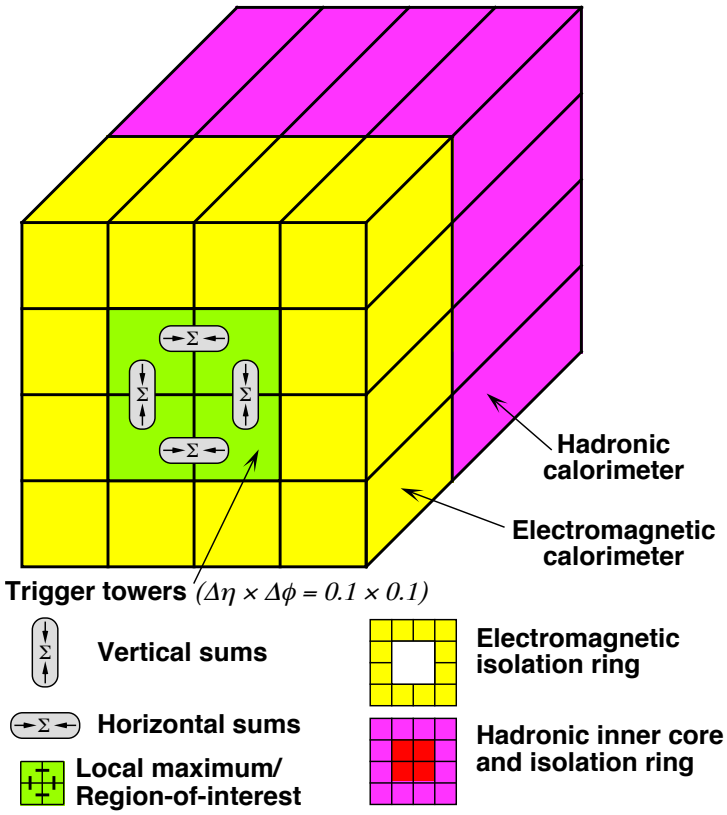


Figure 2.20: EM calorimeter cluster used in the Level-1 electron/photon trigger [3].

Chapter 3

Data and Simulation samples

The dataset used for this analysis is described in this chapter. This analysis uses the proton-proton collision data provided by the LHC and recorded by the ATLAS detector from 2015 to 2017. A detailed description of the dataset is presented in Section 3.1. In order to understand the behavior of the data and the detector, simulated samples of the proton-proton collision are used. Section 3.2 describes the simulation samples in detail.

3.1 Data sample

Table 3.1 shows the integrated luminosity collected in 2015, 2016 and 2017. The integrated luminosity delivered by the LHC reaches 92.9 fb^{-1} while the ATLAS detector successfully records 86.4 fb^{-1}.

During data taking, while detector shifters monitor the status of each detector, data quality shifters additionally check reconstructed objects such as tracks, calorimeter clusters, muons and jets in online. In addition, the data quality is checked again after the offline reconstruction. In the data quality check, some data taking periods whose data is improper to use for physics analysis are removed by various criteria. Some of the sources of such bad events are severe coverage losses, timing shifts and data corruption. After removing these bad events, the integrated luminosity used for the analysis is reduced to 79.8 fb^{-1}. This condition is called "Good for physics" in Table 3.1.

Table 3.1: The integrated luminosity collected in 2015, 2016 and 2017.

Year	LHC delivered [fb^{-1}]	ATLAS recorded [fb^{-1}]	Good for physics [fb^{-1}]
2015	4.2	3.9	3.2
2016	38.5	35.6	33.0
2017	50.2	46.9	43.6
Total	92.9	86.4	79.8

3.2 Simulation samples

The acceptance and efficiency of signal and resonant background events are estimated by Monte Carlo (MC) simulation. The resonant background is composed by the non-ttH Higgs production processes: ggF, VBF, $VH(V = W, Z)$, tH and others. In addition, MC samples which make continuum background are also used to understand its behavior. The continuum background mainly comes from direct di-photon productions. Two types of the direct di-photon simulation sample are prepared; one is that two photons are generated together with a top-quark pair ($t\bar{t}+\gamma\gamma$), and the other does not contain any top-quarks ($\gamma\gamma$). The method of the MC event generation and the samples used for this analysis are described in Section 3.2.1 and 3.2.2, respectively.

3.2.1 Event simulation

In order to simulate each event produced in a proton-proton collision, there are three steps in the event simulation process: event generation, detector simulation and digitization [39].

First, particular physics events are generated with a MC generator. The generator handles hard processes, initial and final state radiation (ISR, FSR), parton showering, underlying events, decays of unstable particle and hadronization processes. To model data, it is useful to combine two generator packages which handle the matrix element calculation and parton shower modeling, respectively. In addition, we need to select various types of parameter tunings such as the parton distribution function (PDF) sets and the renormalization and factorization scales for the QCD calculation.

After the event generation and hadronization, all MC samples are simulated with GEANT4 toolkit [40] to take into account the detector geometry. At this stage, the energy deposition, position and timing information corresponding to each sub-detector are recorded as a "hit" information.

Finally, these hit information are converted into detector responses. This process is called digitization. A digital information is produced when the voltage or current on a particular readout channel has exceeded a detector's threshold within its timing window. The charge collection on each sub-detector is also modeled, which includes cross-talk, electronic noise and channel-dependent variations. Furthermore, the pile-up is overlaid to each event before the digitization.

3.2.2 Signal and background samples

Table 3.2 shows all MC samples for the signal and backgrounds.

Simulated $ttH(H \to \gamma\gamma)$ events are generated using $\mathrm{P}_{\mathrm{OWHEG}}$ [41, 42] with the NLO PDF set by NNPDF3.0 [43]. The parton showering, hadronization and underlying event for this sample are modeled by $\mathrm{P}_{\mathrm{YTHIA}}8$ [44, 45]. In order to reduce the systematic uncertainty from the MC statistics, 4.5 million events are produced (about 45,000 times more than the number of expected events).

Non-ttH Higgs background samples listed on Table 3.2 are also generated using $\mathrm{P}_{\mathrm{OWHEG}}+\mathrm{P}_{\mathrm{YTHIA}}8$ generators except for the tH (tWH and $tHjb$) samples. The tH samples are gen-

Table 3.2: MC sample list used for the signal and background estimation

	Process	Generator	Parton shower	PDF	$\sigma \times BR$ [fb]
Signal	ttH $(H \to \gamma\gamma)$	PowHeg	Pythia8	NNPDF3.0	1.150
Non-ttH Higgs background	ggF $(H \to \gamma\gamma)$	PowHeg	Pythia8	NNLOPS	110.1
	VBF $(H \to \gamma\gamma)$	PowHeg	Pythia8	NNPDF3.0	8.578
	W^+H $(H \to \gamma\gamma)$	PowHeg	Pythia8	NNPDF3.0	1.902
	W^-H $(H \to \gamma\gamma)$	PowHeg	Pythia8	NNPDF3.0	1.206
	ZH $(H \to \gamma\gamma)$	PowHeg	Pythia8	NNPDF3.0	1.725
	bbH $(H \to \gamma\gamma)$	PowHeg	Pythia8	NNPDF2.3	1.104
	$ggZH$ $(H \to \gamma\gamma)$	PowHeg	Pythia8	NNPDF3.0	0.279
	tWH $(H \to \gamma\gamma)$	MadGraph5_aMC@NLO	Herwig++	CT10	0.034
	$tHjb$ $(H \to \gamma\gamma)$	MadGraph5_aMC@NLO	Pythia8	CT10	0.169
Continuum background	$t\bar{t}+\gamma\gamma$ (0-lepton)	MadGraph5_aMC@NLO	Pythia8	NNPDF2.3	6.482
	$t\bar{t}+\gamma\gamma$ ($\geq$1-lepton)	MadGraph5_aMC@NLO	Pythia8	NNPDF2.3	4.076
	$\gamma\gamma$ ($m_{\gamma\gamma} \in [50, 90]$ GeV)	Sherpa 2.2.4	Sherpa 2.2.4	NNPDF3.0	1.391×10^5
	$\gamma\gamma$ ($m_{\gamma\gamma} \in [90, 175]$ GeV)	Sherpa 2.2.4	Sherpa 2.2.4	NNPDF3.0	5.182×10^5

erated with MadGraph5_aMC@NLO [46]. The parton showering is modeled by the Herwig++ [47] for the tWH sample and Pythia8 for the $tHjb$ sample.

For the continuum background, $t\bar{t}+\gamma\gamma$ samples are separated into two top-quark decay modes. The 0-lepton sample indicates that both W bosons decayed from a top-quark pair decay into quarks while $\geq$1-lepton sample denotes that at least one W boson decays into leptons. Both $t\bar{t}+\gamma\gamma$ samples are generated with MadGraph5_aMC@NLO. The parton showering is modeled by Pythia8. The $\gamma\gamma$ samples are prepared with each $m_{\gamma\gamma}$ slices; $m_{\gamma\gamma} \in [50, 90]$ GeV and $m_{\gamma\gamma} \in [90, 175]$ GeV, where $m_{\gamma\gamma}$ is the di-photon invariant mass. Sherpa 2.2.4 [48] is used as the generator and the parton shower model for both $\gamma\gamma$ samples.

In order to evaluate the systematic uncertainties which arise from the generator and the parton shower modeling, hadronization, underlying event and PDF, alternative signal MC samples are produced [49, 50]. These samples are listed in Table 3.3. To estimate the generator uncertainty, the MC sample generated with the MadGraph5_aMC@NLO + Pythia8 is prepared. For the parton shower modeling, hadronization, underlying event and PDF uncertainty estimate, Herwig++ generator is used instead of Pythia8. This uncertainty is evaluated by the difference between MadGraph5_aMC@NLO + Pythia8 and MadGraph5_aMC@NLO + Herwig++.

Table 3.3: ttH $(H \to \gamma\gamma)$ MC samples used to estimate uncertainties related to the generator, parton shower, hadronization, underlying event and PDF.

Generator	Parton shower	PDF	Description
PowHeg	Pythia8	NNPDF3.0	Reference sample
MadGraph5_aMC@NLO	Pythia8	NNPDF3.0	For the uncertainty by the generator
MadGraph5_aMC@NLO	Herwig++	CT10	For the uncertainty by the parton shower, hadronization, underlying event and PDF

Chapter 4

Event reconstruction

In the $ttH(H \to \gamma\gamma)$ analysis, photons, electrons, muons and jets exist in the final state. The reconstruction and the selection criteria of each object are described in this chapter. A jet originated from a b-quark is identified using the "b-tagging" algorithm which is explained in Section 4.6. In addition to the objects above, the missing transverse energy (E_T^{miss}), as the signature of invisible neutrinos, is reconstructed. The E_T^{miss} reconstruction is described in Section 4.7.

4.1 Tracking and vertexing

4.1.1 Tracking

Charged tracks are used in the reconstruction of photons, electrons, muons and jets. Tracks are reconstructed by hit positions recorded by the Inner Detector [51]. The reconstruction procedure has three steps.

The first step is the creation of hit points in the silicon detectors and TRT. The silicon pixel and strip hits form a "cluster" with neighboring pixels and strips, respectively. From these clusters, three dimensional hit points referred to as "space-points" are defined. In the silicon pixel detector, each cluster corresponds to one space-point, while in the silicon strip detector, clusters from both sides of a strip layer are combined to obtain a space-point. Only two dimensional position is provided by the TRT because it does not have the stereo measurement.

The second step is the creation of a track seed from sets of three space-points by finding a straight line on the $r - z$ plane identified by the silicon detectors. Once a seed is found, the candidate track is formed from the seed by incorporating additional space-points using a Kalman filter [52] fitting. A track score is assigned to each candidate to reduce fake tracks. The score is determined by the χ^2 and the number of space-points used for the track fitting. Tracks with the score smaller than a certain threshold are removed from the candidates.

The final step is the extrapolation of the remaining track candidates further to the TRT, adding hits to the candiates. All surviving candidates are re-fitted using all space-points and TRT hits. After the re-fitting, track candidates passing the criteria listed in Table 4.1 are

used as charged particles.

Table 4.1: Summary of the charged track selection criteria.

Parameter	Cut value
Number of silicon hits	≥ 9 (if $\|\eta\| \leq 1.65$) ≥ 11 (if $\|\eta\| \geq 1.65$)
Number of innermost + second innermost pixel layer hits	> 0
Number of missing hits in the pixel detector	$== 0$
p_T	> 10 GeV
Longitudinal impact parameter, $\|z_0\|$	< 1.5 mm

4.1.2 Vertexing

Multiple particles sometimes appear at a certain point by, for example, a proton-proton collision or a particle decay. This point is called as vertex. In this section, the reconstruction procedure of vertices from proton-proton collisions is described. Since there are many proton-proton interactions per bunch crossing, many vertices are produced. The vertex that has the largest sum of squared p_T of the associated tracks is defined as the primary vertex (PV), while the remaining vertices are regarded as the "pile-up" vertices.

The vertex reconstruction [53] begins with searching for the position of a vertex seed where tracks are the most densely populated. For the seed, all tracks are examined if they are compatible with forming a vertex by the χ^2 fitting [54]. A weight, w, based on the χ^2 of the fit is assigned to each track. The weight is low as χ^2 is large. Tracks with low weight have less impact on the calculation of the vertex position. The weighted χ^2 (χ^2_w) is defined as

$$\chi^2_w(\boldsymbol{x}) = \sum_i w_i(\boldsymbol{x})\chi^2_i(\boldsymbol{x}), \tag{4.1}$$

where $\boldsymbol{x}$ is the vertex position, w_i and χ_i are the weight and χ^2 for i-th track, respectively. The position $\boldsymbol{x}$ where the χ^2_w is minimized is selected as the vertex position. After the vertex position is determined, tracks displaced by more than 7σ, where σ is the error of the vertex position, from the vertex are used in the determination of other vertices. This procedure is repeated until no unassociated tracks are left in the event or no additional vertices can be found.

4.2 Photon

Reconstruction

Photons (as well as electrons and jets) are reconstructed from three dimensional clusters built by calorimeter cells which have large energy deposit. Such a cluster is called "topo-cluster" [55]. The topo-cluster is defined by connecting calorimeter cells which have large

signal. The cell unit is listed in Table 2.4. The cell which has energy of $|E_{\rm cell}/\sigma_{\rm cell}| \geq 4$ is considered to be the cluster seed, where $|E_{\rm cell}|$ is the measured energy at the cell and $\sigma_{\rm cell}$ is the cell noise. The seed cell and neighboring cells with $|E_{\rm cell}/\sigma_{\rm cell}| \geq 2$ are grouped to form a cluster, where the neighboring cells are defined as four cells around a seed cell in a given layer, and cells in adjacent layers having partial overlap in the (η, ϕ) plane. Each neighbor cell becomes a seed cell, and then new neighboring cells are searched. This iteration is repeated until no new neighboring cells are found. Finally, neighboring cells satisfying $|E_{\rm cell}/\sigma_{\rm cell}| \geq 0$ are added to the cluster.

Both the hadronic calorimeter and the EM calorimeter are used to form the topo-cluster. The EM energy fraction f_{EM} is defined as

$$f_{EM} = \frac{E_{EM}^{cluster}}{E_{EM}^{cluster} + E_{Had}^{cluster}}, \tag{4.2}$$

where $E_{EM}^{cluster}$ and $E_{Had}^{cluster}$ are the cluster energies in the EM and hadronic calorimeter, respectively. In the photon reconstruction, $f_{EM} > 0.5$ is required to reduce hadrons. This requirement can reject $\sim$60 % of hadrons while the photon efficiency is kept to be > 99 %.

To discriminate photons from electrons, photons are required not to have any tracks matching with a calorimeter cluster. Photon candidates with $|\eta_{track} - \eta_{cluster}| < 0.05$ and $-0.10 < q \cdot (\phi_{track} - \phi_{cluster}) < 0.05$ are removed, where $\eta_{cluster}$ and $\phi_{cluster}$ are the positions of a cluster barycenter, η_{track} and ϕ_{track} are the positions of a track extrapolated to the second layer of the EM calorimeter, and q is the charge of the track. The remaining cluster after the above requirements is called "unconverted photon".

In addition, two tracks matching to one cluster are checked if they form a conversion vertex. In case the pair of tracks form a conversion vertex, the cluster with such a pair of tracks is regarded as "converted photon". Figure 4.1 shows the concept of the unconverted photon, converted photon and electron.

In order to recover converted photons with only one track matching with a topo-cluster, the supercluster [56] is introduced as shown in Figure 4.2. There are two types of the supercluster. First, two adjacent topo-clusters in the distance $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ are considered to be a supercluster. Second, two topo-clusters with matching tracks forming a common conversion vertex are considered to be a supercluster, regardless of its distance.

Selection

The photon p_T is defined as $p_T = E_T = E/\cosh(\eta)$, where E is the energy of photon cluster and η is the pseudorapidity of the barycenter of the cluster in the second layer of the EM calorimeter. The photons are required to have $p_T > 22$ GeV and $|\eta| < 2.37$ (excluding $1.37 < |\eta| < 1.52$), and to pass the selection criteria to reject backgrounds of hadronic jet which contains neutral hadrons or photons. The photons typically produce narrower shower in the EM calorimeter and have smaller leakage to the hadronic calorimeter compared to jets. The cuts on the following variables are used to discriminate photons from hadronic jets [57]:

- Energy leakage to the hadronic calorimeter
- Energy deposit in the middle layer of the EM calorimeter

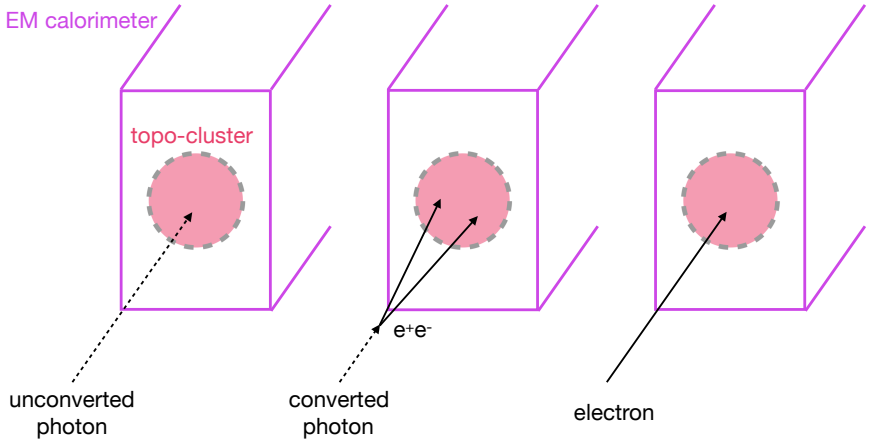


Figure 4.1: Track-cluster matching in the photon (electron) reconstruction. The unconverted photon has no matched tracks. The converted photon has a charged track pair matched to the cluster. These tracks form the vertex apart from a collision point. The electron candidate has associated charged track coming from a collision point.

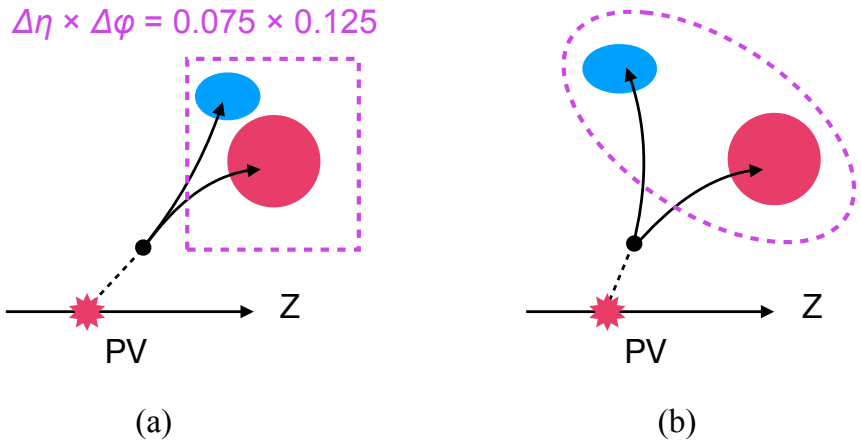


Figure 4.2: Diagrams of the superclustering algorithm for converted photons. (a) two adjacent topo-clusters in the distance $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ are considered to be a supercluster. (b) two topo-clusters whose matched tracks have a common conversion vertex are considered to be a supercluster.

- Shower width measured by the finely segmented layer (called strip layer) of the EM calorimeter

The list of all the discriminating variables is shown in Appendix A. The cuts are tuned separately for unconverted and converted photons in several η regions. There are three cut points called *Loose*, $Medium$ and $Tight$. The *Loose* selection uses the energy leakage to the hadronic calorimeter and the energy deposit in the middle layer of the EM calorimeter. The *Loose* cut point has the highest photon efficiency in the three cut points, which is 98.9 % for unconverted photons and 96.3 % for converted photons. The $Medium$ selection uses part of variables for the strip layer of the EM calorimeter in addition to the *Loose* selection, to reduce fake photons more. The $Medium$ cut point is used only for the photon trigger as discussed in Section 4.2.2. The $Tight$ selection uses all variables for the strip layer of the EM calorimeter in addition to the *Loose* selection.

Performance

- Energy calibration

The energy deposited in a topo-cluster is estimated based on the test-beam studies [58]. The three types of corrections are applied; the energy loss in the materials in front of the EM calorimeter, the energy leakage to the cells adjacent in η and ϕ, and the energy leakage to back of the EM calorimeter. The size of these corrections is determined by MC. Both data and MC use the same correction.

After the energy calibration of the topo-cluster, the energy scale is calibrated with electrons coming from $Z \to ee$ decays [59] because both photon and electron energies are determined by the topo-cluster. The electron energy in data is adjusted so that the di-electron invariant mass is matched to the world average of the Z boson mass. The difference in the energy resolution between data and MC is also measured to correct MC. The systematic uncertainty in the energy scale calibration varies between 0.25 % and 1 %, depending on η, for photons with p_T around 60 GeV.

4.2.1 $Tight$ selection efficiency measurement

The efficiency of the $Tight$ photon selection with respect to reconstructed photons is measured in data. The efficiency is calculated by counting the number of photons passing the $Tight$ selection, with respect to the number of reconstructed photons, and therefore highly pure photon sample is needed. The photon reconstruction efficiency is estimated to be > 95 % for photons with $p_T > 25$ GeV using the $Z \to ll\gamma$ ($l = e, \mu$) MC sample. The difficulty of the measurement is the absence of such pure photon samples over a large p_T range. To deal with the problem, two different measurements are used.

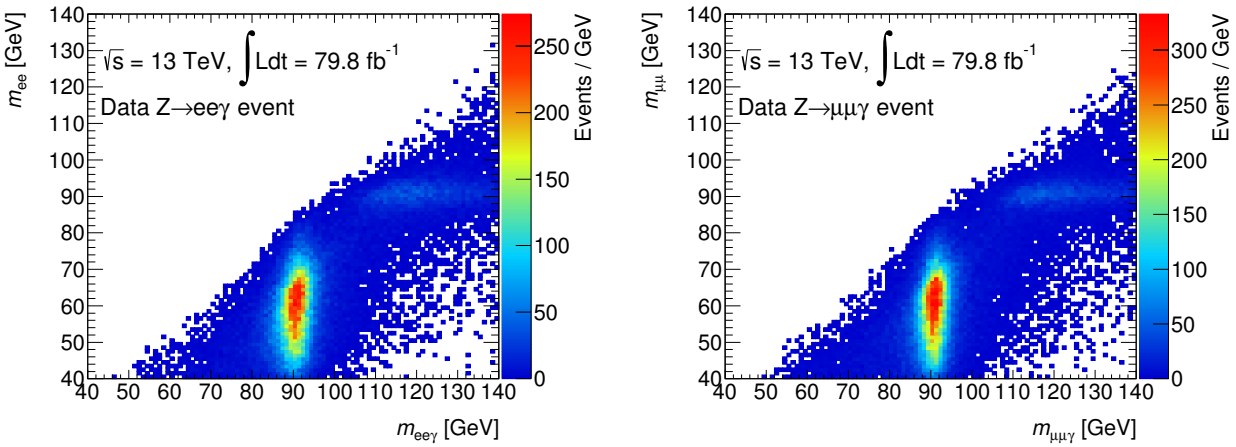
- *Radiative Z decays* : This method uses $Z \to ll\gamma$ ($l = e, \mu$) events where photon comes from the final state radiation (FSR). High purity photons can be obtained but the statistics of high p_T photon is not sufficient. For this reason, it allows to measure the efficiency only in the region of $20 < p_T < 100$ GeV.

- *Extrapolation of electron to photon* : By making use of the similarities of the shower shape between electrons and photons in the EM calorimeter, electrons are used to measure the photon efficiency with some corrections. This method uses a pure sample of electrons obtained from $Z \to ee$ events. This allows to measure the efficiency in the region of $30 < p_T < 1000$ GeV.

Radiative Z decays

Events are required to pass logical *OR* of the single electron (muon) trigger and di-electron (di-muon) trigger to select $Z \to ee\gamma$ ($Z \to \mu\mu\gamma$) events. An opposite charge pair of the same flavor leptons and at least one photon are required to exist. Photons are required to have $p_T > 20$ GeV and pseudorapidity in the range of $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. The isolation selection to photons described in Section 5.2 is applied.

The main backgrounds are initial state radiation (ISR) events, where the photon is radiated from initial quarks, and Z + jets events where a jet fakes a photon. In order to reduce them, events are required to have 80 GeV $< m_{ll\gamma} <$ 100 GeV, where $m_{ll\gamma}$ is the invariant mass of two leptons and a photon. Figure 4.3 shows the two dimensional invariant mass distribution of ll and $ll\gamma$. The signal purity is estimated in data by fitting $m_{\gamma\gamma}$ distribution with the shape of $Z \to ll\gamma$ and $Z \to ll$ (+ ISR or jet) expected by the MC. The signal purity is > 99 % for both electron and muon channels as shown in Figure 4.4.

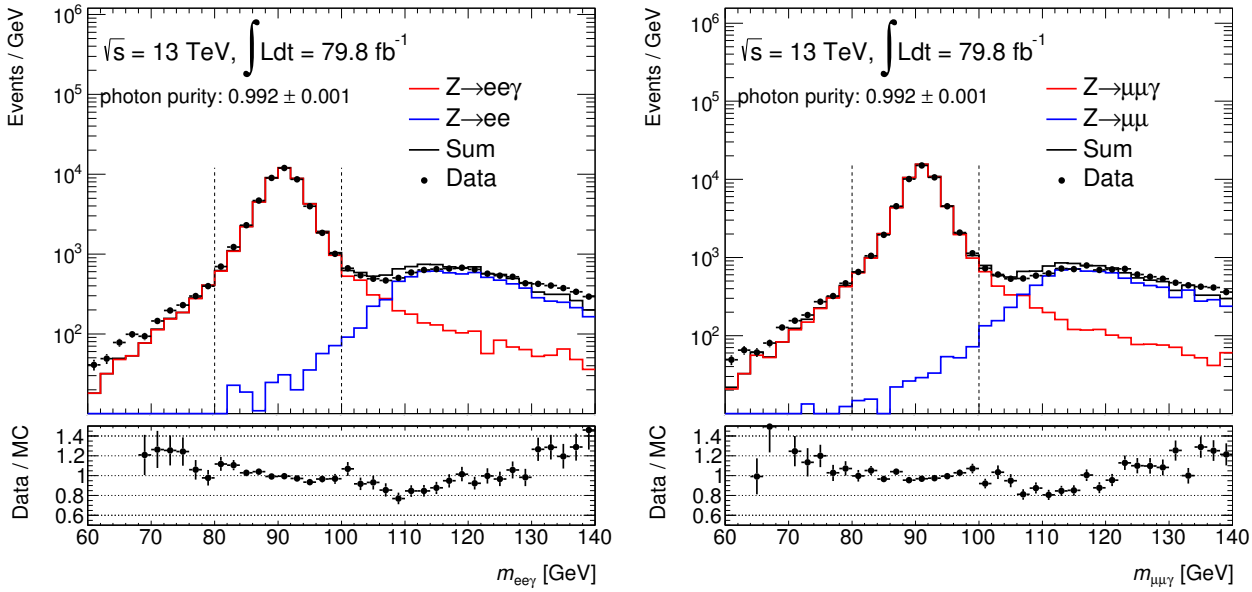


(a) (b)

Figure 4.3: Two dimenaional invariant mass distributions of (a) ee vs $ee\gamma$ and (b) $\mu\mu$ vs $\mu\mu\gamma$.

Extrapolation of electron to photon

In this method, the efficiency is calculated by counting the number of electrons passing the *Tight* photon selection with some corrections to the shower shape to compensate the difference between electrons and photons. To collect pure electrons without biasing the electron which is used to evaluate the efficiency, the $Z \to ee$ decays are selected by the



(a) (b)

Figure 4.4: Mass distribution of $Z \to ll\gamma$ events for (a) electron and (b) muon channel.

tag-and-probe method, where an unbiased electron, 'probe', is obtained by requiring strict selections to another electron, 'tag'. Event selection starts from to pass the single electron trigger. The trigger object must match the tag electron. An opposite charged pair of the same flavor electrons is required to exist. Both tag and probe electrons are required to have $p_T > 30$ GeV and pseudorapidity in the range of $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. The isolation selection is applied for both electrons. Only the tag electron is required to pass the electron identification selection as described in Section 4.3 to reduce fakes. The events are required to have the invariant mass of di-electron in the range of $70 < m_{ee} < 110$ GeV.

The backgrounds are mainly multi-jet and W+jets events, where a jet fakes an electron. Figure 4.5 shows the di-electron invariant mass distribution for $Z \to ee$ events. The signal purity is estimated by fitting the mass distribution with the signal and background template shapes. The signal template is prepared by the $Z \to ee$ MC while the background template is obtained from the data where the probe electron fails both the identification and isolation selections. The signal purity is estimated to be > 99 % by the fitting.

The difference of the shower shape between photons and electrons is evaluated using MC. The photon (electron) shower shape is obtained by γ + jet ($Z \to ee$) MC. There are eleven variables to represent the shower shape as listed in Appendix A. Each electron variable, x_e, is transformed to photon variable, x_γ, obeying the Smirnov transformation [60] with f defined as

$$x_\gamma = f(x_e) \equiv F_\gamma^{-1}(F_e(x_e)), \tag{4.3}$$

where $F_\gamma(x)$ and $F_e(x)$ are the cumulative distribution functions (CDF's) of x for simulated

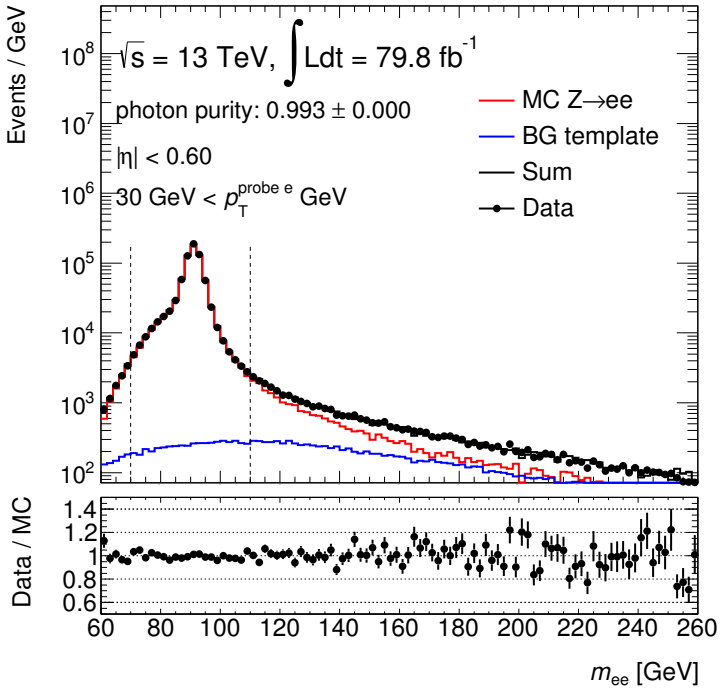


Figure 4.5: Distribution of the di-electron invariant mass for $Z \to ee$ events.

photons and electrons, respectively. Electrons have relatively broader shower shape because they tend to make a shower in front of the calorimeter than photons. To demonstrate the transformation the variable R_ϕ is selected as the example, where R_ϕ is the ratio of the energy deposit in 3×7 ($\eta\times\phi$) cells over the energy deposit in 7×7 cells of the middle EM calorimeter layer, centered at the photon cluster position. The R_ϕ distribution for electrons and photons, as shown in Figure 4.6 (a), are used to get the corresponding CDF's (Figure 4.6 (b)). From these CDF's, a Smirnov transformation is derived, which is a mapping of R_ϕ between electron and photon as shown in Figure 4.6 (c). Based on this mapping, the electron distribution is transformed to the pseudo-photon distribution as shown in Figure 4.6 (d).

After deriving x_γ's, the number of electrons passing the photon requirements with the x_γ's is simply counted to obtain the efficiency.

Result of the efficiency measurement

The *Tight* photon selection efficiency ε^{Tight} is defined as

$$\varepsilon^{Tight} = \frac{N^{Tight}_{probe\ \gamma(e)} - N^{Tight}_{BG\ \gamma(e)}}{N_{probe\ \gamma(e)} - N_{BG\ \gamma(e)}}, \tag{4.4}$$

where $N_{probe\ \gamma(e)}$ and $N^{Tight}_{probe\ \gamma(e)}$ are the numbers of probe photons (electrons), and the ones passing the *Tight* selection, respectively. The $N_{BG\ \gamma(e)}$ and $N^{Tight}_{BG\ \gamma(e)}$ are the numbers of

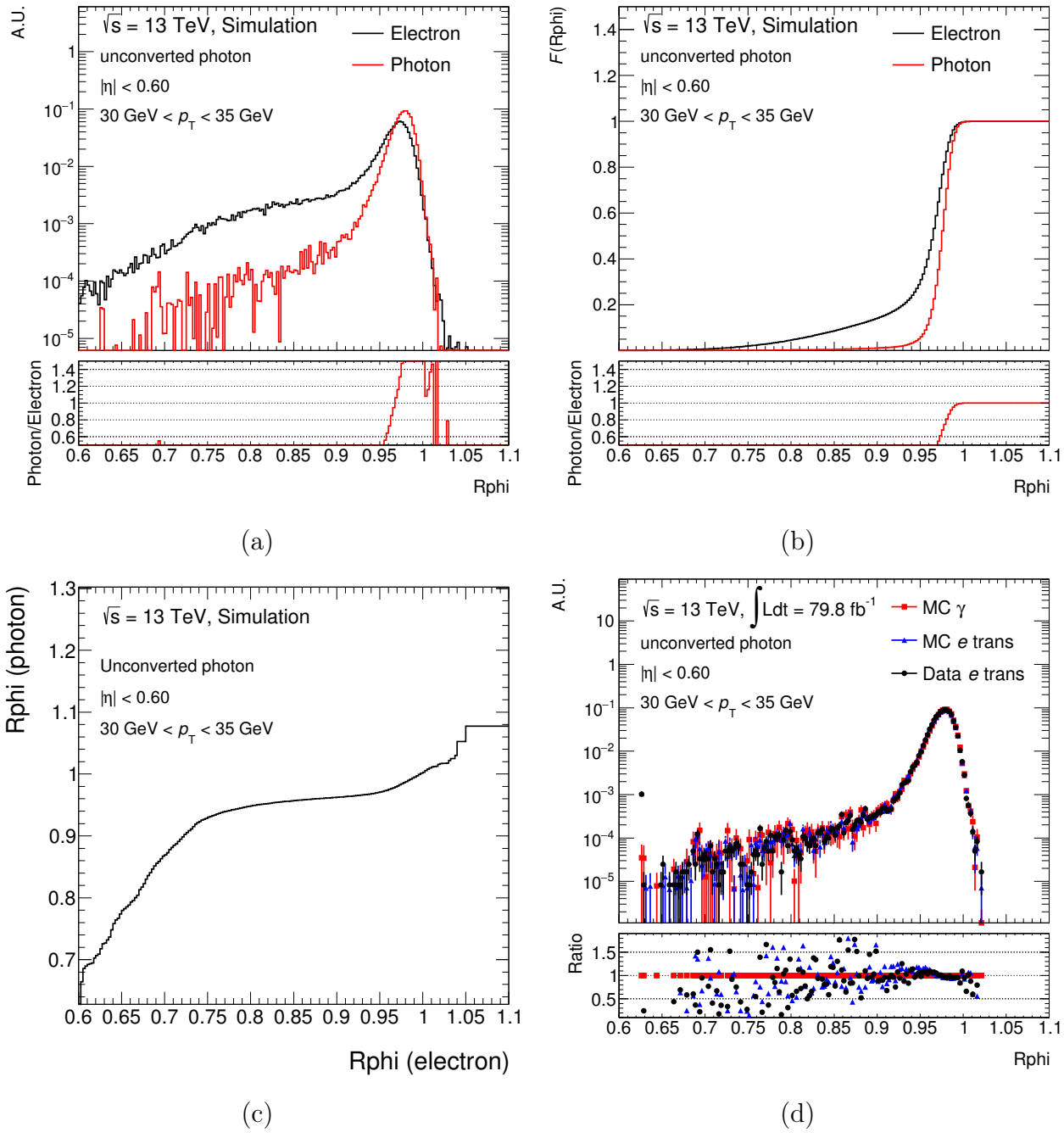


Figure 4.6: Demonstration of the extrapolation of electron to photon method: (a) the initial R_ϕ for electrons and photons, (b) the CDF of R_ϕ for electrons and photons, (c) the transformation map for the R_ϕ from electrons to photons, (d) the R_ϕ distribution after the transformation.

backgrounds before and after the $Tight$ selection. The ε^{Tight} is measured as a function of p_T and $|\eta|$ of the reconstructed photon or electron, and separately for the converted and unconverted photons. Figure 4.7 and 4.8 show the results. The efficiency varies from 95 % to 98 % and from 90 % to 94 %, depending on η, for the converted and unconverted photons with p_T around 50 GeV, respectively. The two methods give the consistent results.

4.2.2 Di-photon trigger efficiency measurement

The trigger system is separated into the Level-1 and HLT as described in Section 2.7. After the decision of the Level-1, the Regions of Interest (RoI's) defined by the Level-1 trigger are sent to the HLT. The decision in the HLT is performed by the online photon reconstruction. The difference of the reconstruction algorithm between the online and offline described above is the definition of the calorimeter cluster. In the online reconstruction, simpler clustering algorithm is used instead of the topo-cluster finding to save the CPU time, as explained below [61].

A cluster is formed by energy deposits only in EM calorimeter cells. The EM calorimeter is divided by the elements of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. Inside each of these elements, the energy of all cells in all longitudinal layers is summed. Each element is called "tower" in the following. Then, a window with fixed size $N_\eta \times N_\phi = 3 \times 7$ and 5×5, in units of the tower size, is defined for the barrel and end-caps of EM calorimeter, respectively. The window is moved across each element of the tower grid in the RoI. If the sum of E_T in the window is a local maximum, the group of towers in the window is defined as the cluster. Photons are reconstructed by the clusters with the same method as the offline reconstruction while the supercluster is not formed in the online reconstruction. The decision of photon triggers is based on the photon p_T and $Loose$ or $Medium$ selection criteria.

Table 4.2 shows the requirements of the di-photon trigger at the Level-1 and HLT, respectively. Since the data taking period in 2017 is in relatively higher pile-up environment than that in 2015 and 2016, the requirement in 2017 is tighter than the one in 2015 and 2016. The p_T threshold of both photons is increased in the Level-1 while the $Medium$ selection is used instead of $Loose$ in the HLT.

Table 4.2: The di-photon trigger requirement in 2015, 2016 and 2017

Year	Level-1	HLT
2015 , 2016	$E_T > 15$ GeV for both photons	$p_T > 35(25)$ GeV for (sub-)leading photon Photon selection: "$Loose$"
2017	$E_T > 20$ GeV for both photons	$p_T > 35(25)$ GeV for (sub-)leading photon Photon selection: "$Medium$"

Since the di-photon trigger efficiency cannot be directly measured because of lack of pure data sample, the efficiency is obtained by multiplying the single photon trigger efficiencies. The di-photon trigger efficiency with respect to $ttH(H \to \gamma\gamma)$ events is defined as

$$\varepsilon^{trig}_{ttH(H\to\gamma\gamma)} = \int \varepsilon^{trig(35\mathrm{GeV})} P^{lead}(p_T) \times \varepsilon^{trig(25\mathrm{GeV})} P^{sub-lead}(p_T) dp_T, \tag{4.5}$$

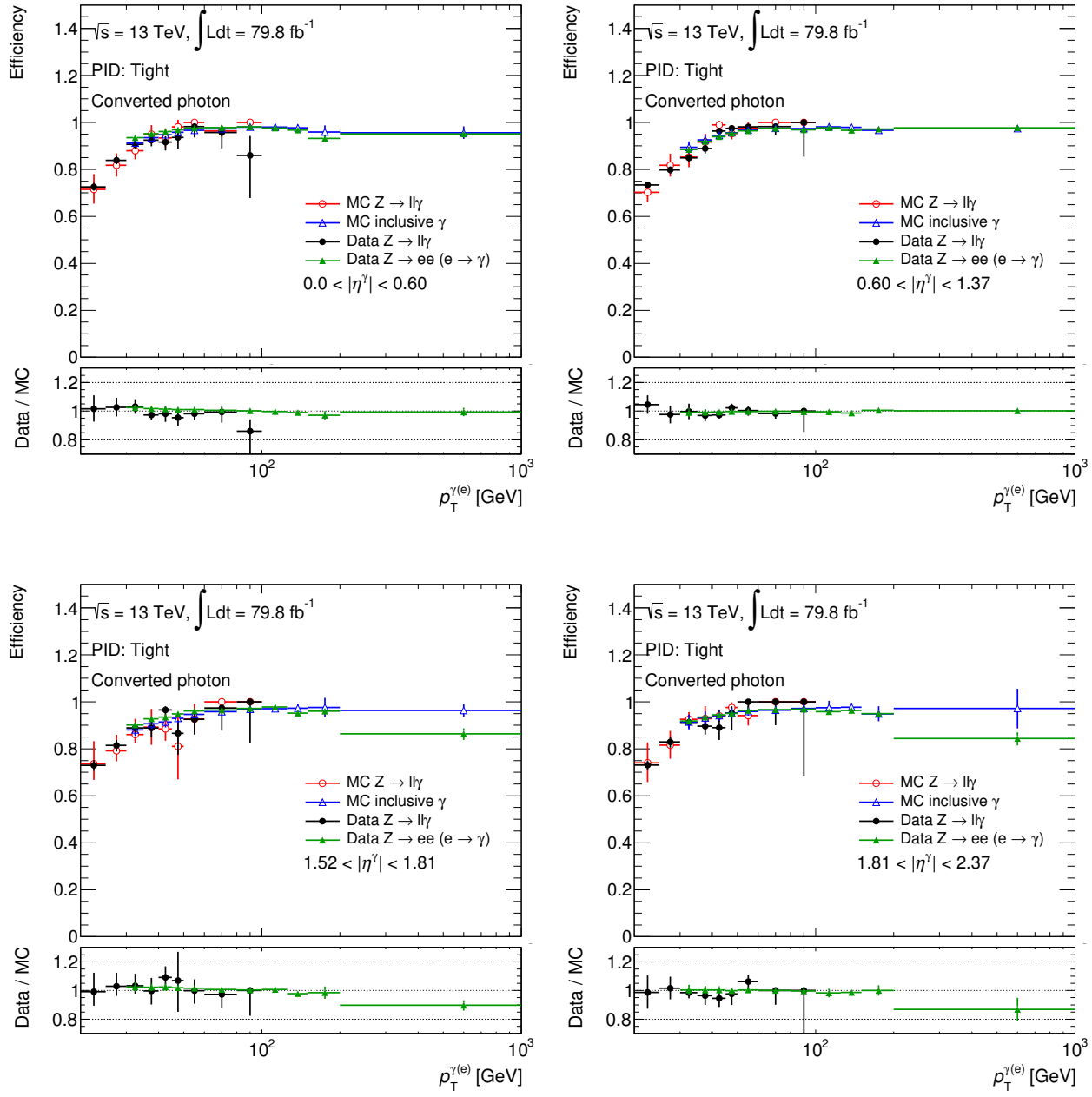


Figure 4.7: Photon efficiency for converted photon, measured by the *Radiative Z decays* method (red and black) and the *extrapolation of electron to photon* method (blue and green), respectively. The error bars correspond to the sum of the systematic and statistical uncertainties.

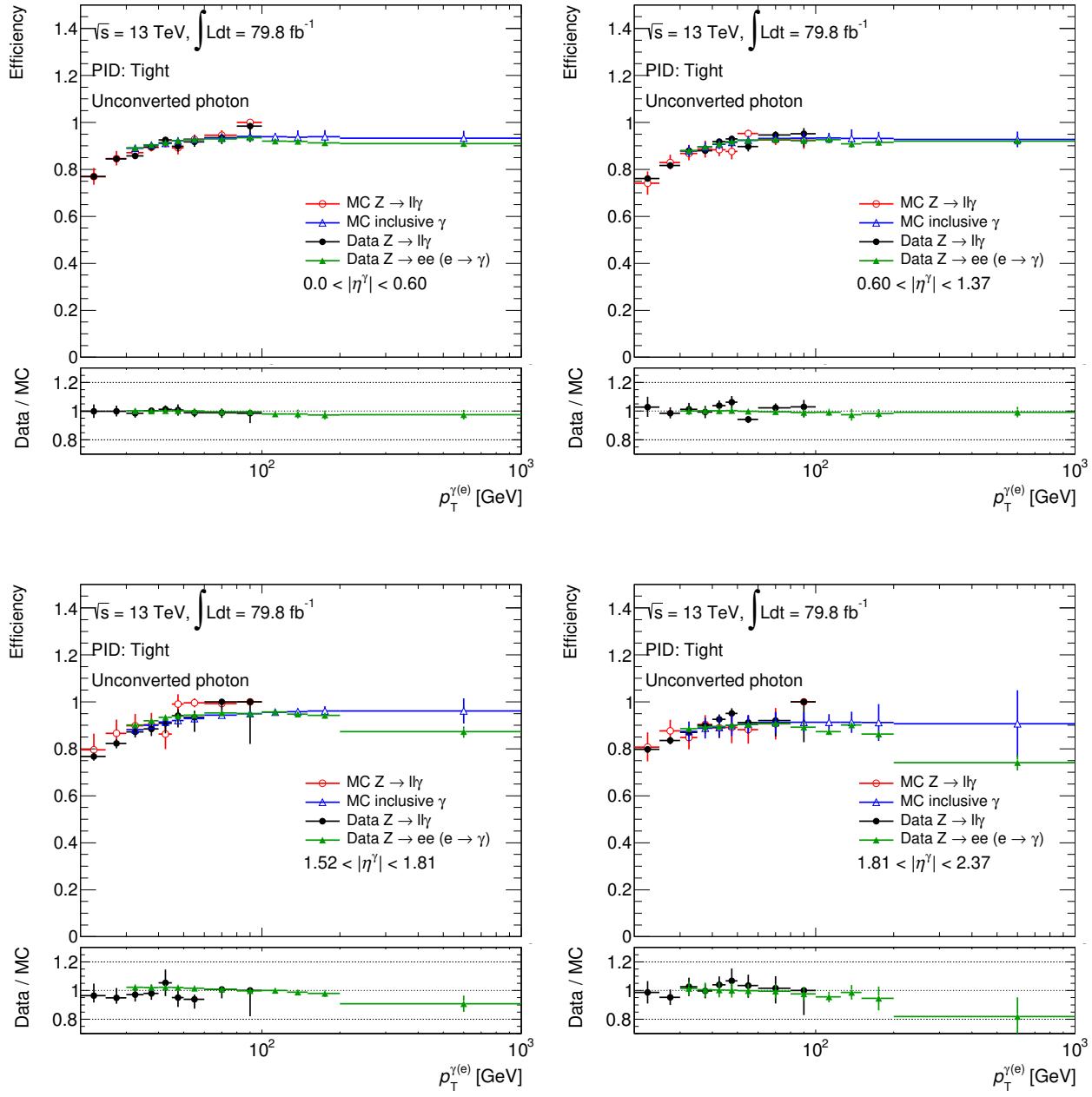


Figure 4.8: Photon efficiency for unconverted photon, measured by the *Radiative Z decays* method (red and black) and the *extrapolation of electron to photon* method (blue and green), respectively. The error bars correspond to the sum of the systematic and statistical uncertainties.

where $P^{lead}(p_T)$ and $P^{sub-lead}(p_T)$ are the probability density functions of the leading and sub-leading photon p_T in the $ttH(H \to \gamma\gamma)$ events made by MC, respectively. $\varepsilon^{trig(35\mathrm{GeV})}$ and $\varepsilon^{trig(25\mathrm{GeV})}$ are the single photon trigger efficiencies for the leading and sub-leading photons, respectively. Each single photon trigger efficiency is measured in data, by counting the number of photons passing the trigger requirement with respect to the number of reconstructed photons. In order to obtain pure photon samples, $Z \to ee\gamma$ and $Z \to \mu\mu\gamma$ events are used. The detail of the event selection is explained in Section 4.2.1. Figure 4.9 shows the measured single photon trigger efficiencies. The difference of the efficiencies between the di-photon trigger and the multiplication of the single photon triggers is appeared when two photons overlap. This difference is estimated to be $0.1 \sim 25$ %, depending on photon p_T, by using $ttH(H \to \gamma\gamma)$ MC sample. It is considered as a correction factor. The efficiencies of the triggers are $99.9\ ^{+0.1}_{-1.0}$ % in 2015 and 2016, and $99.3\ ^{+0.3}_{-1.1}$ % in 2017.

4.3 Electron

The electron reconstruction is similar to the photon reconstruction. Only the difference is that electron is required to have an associated charged track coming from primary vertices.

The electrons are required to have $p_T > 10$ GeV and $|\eta| < 2.47$ (excluding $1.37 < |\eta| < 1.52$). The transverse and longitudinal impact parameters, d_0 and z_0, are required to be $|d_0/\sigma(d_0)| < 5.0$ and $|z_0 \sin\theta| < 3.0$ mm, respectively.

Additional requirement is needed to reduce backgrounds such as hadronic jets, electrons from heavy flavor decays, and photons. In order to remove these backgrounds, electrons are required to satisfy a set of selection criteria [62]. Jets produce broader shower in the EM calorimeter and have larger leakage to the hadronic calorimeter compared to electrons. Electrons from photon conversion have large impact parameter and have less space-points in their tracks, and broader shower in the EM calorimeter. Based on the above, the following kinds of variables are used to discriminate electrons from hadronic jets and photons:

- Energy leakage to the hadronic calorimeter
- Energy deposit in the middle layer of the EM calorimeter
- Shower width measured by the strip layer of the EM calorimeter
- Track conditions: the number of hits in the silicon detectors and the transverse impact parameter
- Transition radiation in the TRT: the number of TRT hits by transition radiation
- Track-cluster matching: distance between the cluster and track positions extrapolated to the EM calorimeter

The electron selection is based on the likelihood, where the signal and background probability density functions (PDF's) are the discriminating variables listed above. The signal PDF's are obtained from $Z \to ee$ MC while the background PDF's from multi-jet, $qg \to q\gamma$, $qq \to g\gamma$,

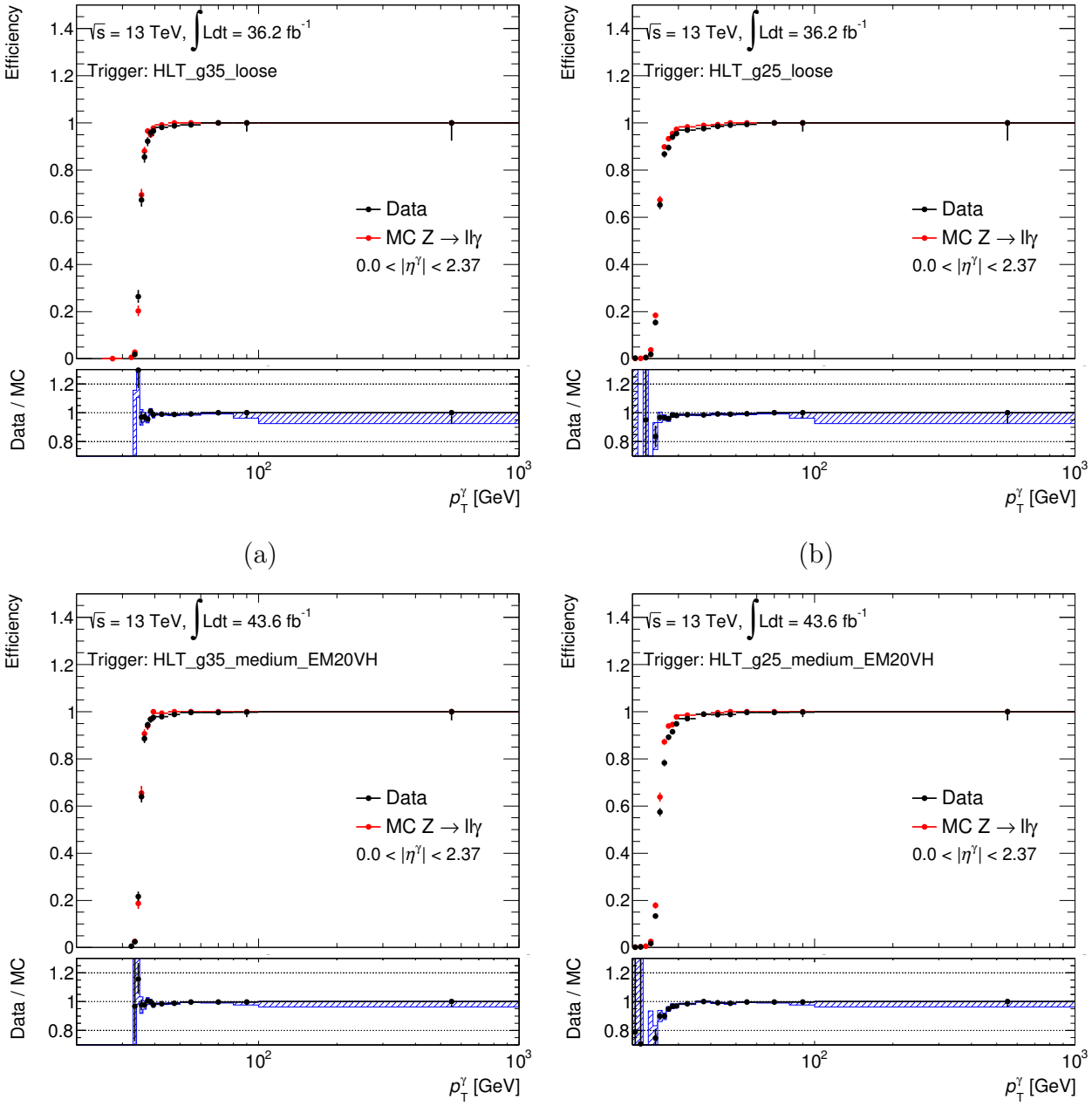


Figure 4.9: The single photon trigger efficiencies measured using the *Radiative Z decays* events. (a) $p_T > 25$ GeV and (b) $p_T > 35$ GeV, and "*Loose*" identification is required in the 2015 and 2016 measurement. (c) $p_T > 25$ GeV and (d) $p_T > 35$ GeV, and "*Medium*" identification is required in the 2017 measurement.

W, Z, and top-quark production MC. The selection cut is applied on the output of the likelihood. The selection efficiency is measured with the $Z \to ee$ events. It varies from 75 % to 93 % depending on p_T and $|\eta|$.

The electron energy scale and resolution are measured with the $Z \to ee$ events. The systematic uncertainty in the energy scale calibration varies between 0.05 % and 0.26 %, depending on η, for electrons with p_T around 45 GeV.

4.4 Muon

The muon reconstruction starts with forming a muon track in the muon spectrometer [63]. The track in the muon spectrometer is then combined with A track reconstructed in the Inner Detector to form a global muon track.

The track reconstruction in the muon spectrometer starts with a search for hit patterns in each layer of muon chambers to form segments. Muon track candidates are then built by fitting the segments. At least two segments are required to exist for the track candidates. Segments providing large contribution of the χ^2 are removed and the fitting is repeated.

Once the track in the muon spectrometer is formed, it is combined with the charged track reconstructed by the Inner Detector, where fitting is performed using all the hits in both the Inner Detector and the muon spectrometer. During the fitting, hits in the muon spectrometer may be added to or removed from the track if the fit quality improves. The global tracks with $p_T > 10$ GeV and $|\eta| < 2.7$ are selected as muons.

Further selections are applied to reduce muons from decays in flight of charged hadrons and punch through charged hadrons. Muons originating from in-flight decays of charged hadrons can be characterized by the presence of a kink topology in their track. Thus, the impact parameter and the track quality requirements are helpful to reject such backgrounds. The impact parameters are required to be $|d_0/\sigma(d_0)| < 3.0$ and $|z_0 \sin\theta| < 0.5$ mm. In addition, a set of track quality cuts is applied for the following variables:

- The number of hits in the muon spectrometer:
  ≥ 3 hits in at least two layers are required

- The number of hits in the Inner Detector:
  At least one silicon pixel hit, at least 5 silicon strip hits and less than three missing hits in the silicon pixel and strip detectors. In addition, the number of TRT hits nearby the muon track but not included in the track fit is required to be less than 90 % of the number of TRT hits on the muon track in the range of $0.1 < |\eta| < 1.9$.

- *q/p significance*:
  $|q/p_{ID} - q/p_{MS}|/\sqrt{\sigma_{ID}^2 + \sigma_{MS}^2} < 7$ is required, where q/p_{ID} and q/p_{MS} are the ratio of the charge and momentum of the muon measured in the Inner Detector and muon spectrometer, and σ_{ID} and σ_{MS} are the corresponding uncertainties, respectively.

The efficiency is measured using the $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events to be close to 99 %.

The muon momentum scale and resolution are measured using the $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events [63]. The muon momentum is scaled and smeared in simulation to match to data. The

muon momentum uncertainty varies from a minimum of 0.05 % for $|\eta| < 1$ to a maximum of 0.3 % for $|\eta| \sim 2.5$.

4.5 Jet

Reconstruction

High energy quarks and gluons originated from proton-proton collisions make collimated hadron bundles via the fragmentation and hadronization processes. This bunch of hadrons is called jet that is reconstructed from the energy cluster of the calorimeter. Jets are reconstructed by forming a bunch of topo-clusters defined in Section 4.2.

The boundary of jets may be affected by soft gluon radiation and splitting (so called infrared and collinear unsafe). In order to minimize the fluctuation in energy and position measurements due to such gluon radiation and splitting, the anti-k_t algorithm [64] is used. In this algorithm, the distance of two-cluster pair d_{ij} is defined as

$$d_{ij} = \min\{k_{t_i}^{-2}, k_{t_j}^{-2}\}\frac{\Delta R_{ij}}{R}, \tag{4.6}$$

where i and j are the indices of the clusters, k_t is the transverse momentum of the topo-cluster, and ΔR_{ij} is the distance between two clusters defined as $\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$. R is called radius parameter, and $R = 0.4$ is used in this analysis. If a two-cluster pair that minimizes d_{ij} is found, they are merged as a new cluster. After merging, if a cluster i, which satisfies $d_{ij} < k_{t_i}^{-2}$ for any other cluster j is found, the cluster i is removed from the cluster list and regarded as a jet. This merging procedure is repeated until no clusters are left. Figure 4.10 shows the jet shapes in the $y - \phi$ plane for the anti-k_t algorithm, where y is rapidity defined as Equation 2.4. The soft particles tend to be bundled with the energetic clusters without modifying the shape of the jet themselves.

Selection

In order to select jets from the primary vertex (PV), tracks matched to the topo-clusters are checked if they come from the PV or pile-up vertex. The track information is parameterized as the "Jet-Vertex-Tagger (JVT)" [65]. The JVT is constructed using two variables, $corrJVF$ and R_{p_T}, defined as

$$corrJVF = \frac{\sum_k p_T^{trk_k}(PV)}{\sum_l p_T^{trk_l}(PV) + \frac{\sum_n \sum_l p_T^{trk_l}(PU_n)}{0.01 \times N_{trk}^{PU}}}, \tag{4.7}$$

$$R_{p_T} = \frac{\sum_k p_T^{trk_k}(PV)}{p_T^{jet}}, \tag{4.8}$$

where $\sum_k p_T^{trk_k}(PV)$ is the scalar p_T sum of the tracks associated with the jet originating from PV, $\sum_n \sum_l p_T^{trk_l}(PU_n)$ is the scalar p_T sum of the tracks originating from any of pile-up vertices, N_{trk}^{PU} is the number of pile-up vertices, and p_T^{jet} is the jet p_T. Both the $corrJVF$

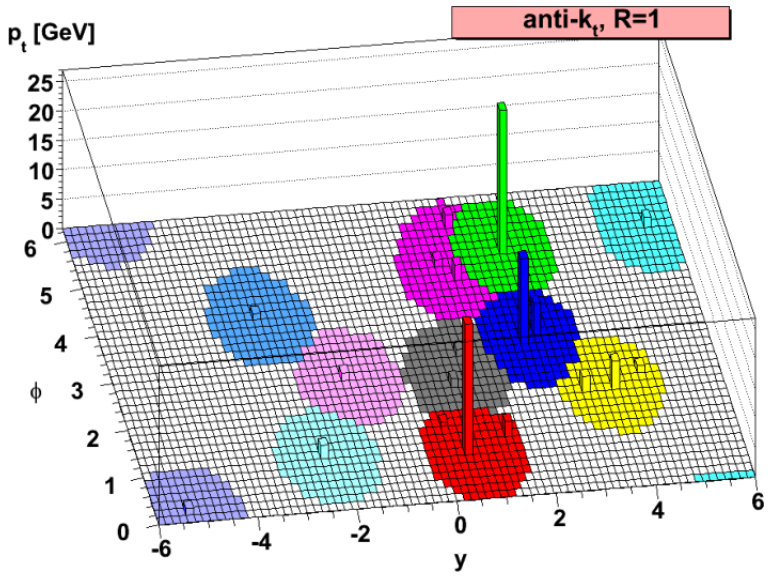


Figure 4.10: Jet shapes in the $y - \phi$ plane (y is rapidity) for the anti-k_t algorithm with $R = 1.0$ [64].

and R_{p_T} are close to 1 when the jet arises from PV. The JVT is the likelihood discriminant defined as

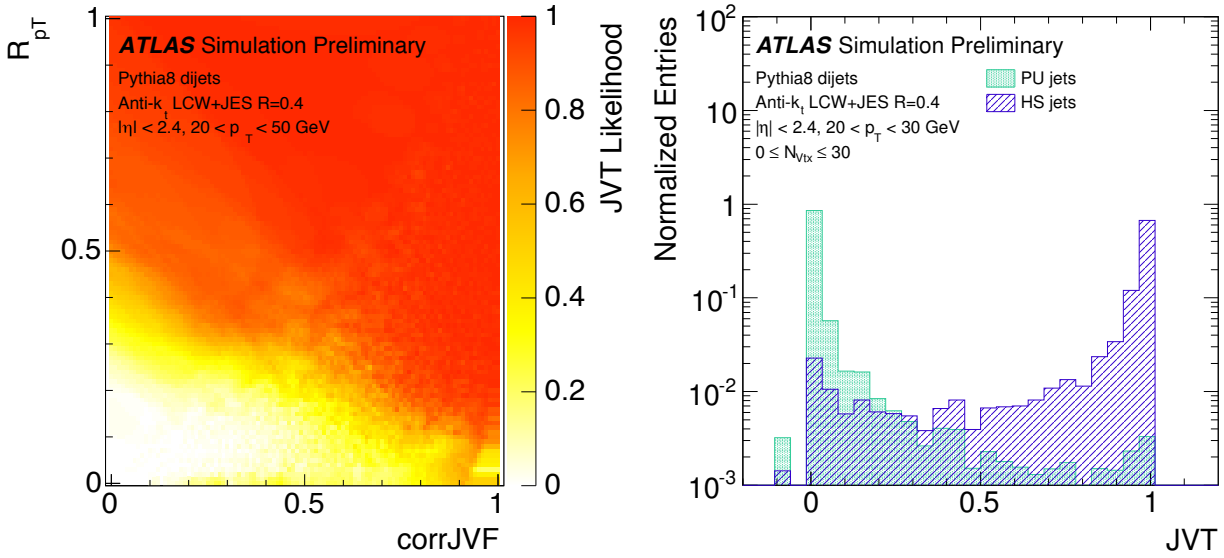
$$\text{JVT} = \frac{L_S(corrJVF, R_{p_T})}{L_S(corrJVF, R_{p_T}) + L_B(corrJVF, R_{p_T})}, \tag{4.9}$$

where L_S and L_B are the likelihoods for the signal (jet from PV) and the background (jet from pile-up vertices) defined in two dimensional distribution of $corrJVF$ and R_{p_T}. The probability density functions for these likelihoods are obtained from the di-jet MC. Figure 4.11 shows JVT as functions of $corrJVF$ and R_{p_T}, and JVT for jets from PV and pile-up, respectively. In this analysis, jets with JVT > 0.59 are selected.

Performance

- JVT efficiency

The JVT selection efficiency is evaluated using $Z \to \mu\mu + jets$ events with the following requirements. Z boson is identified by the requirement of $81 < m_{\mu\mu} < 101$ GeV, where $m_{\mu\mu}$ is the invariant mass of two muons. The leading jet with $p_T > 20$ GeV and $|\eta| < 2.4$ is required to be $|\Delta\phi(Z, \text{jet})| > 2.8$. Figure 4.12 shows the JVT selection efficiency with a cut on 0.59. The difference in efficiency between data and MC is corrected by a scale-factor as a function of p_T.



(a) (b)

Figure 4.11: (a) The two dimensional JVT likelihood distribution and (b) the JVT distribution for jets originating from PV (HS jets) and pile-up vertex (PU jets) [65].

- Jet energy calibration

The energy of topo-clusters used to reconstruct jets is calibrated with the assumption that the energy is deposited by electromagnetic showers while the following effects are not taken into accounts;

- energy loss by particles coming from pile-up interactions.
- hadronic interaction in the calorimeter.
- energy deposit at the inactive materials in the calorimeter.

Since what we handle here is hadronic jets, the energy scale is further calibrated to hadronic jet energy scale. This correction is performed by MC, and confirmed by data with a technique that exploits the p_T balance between a jet and a reference object in back-to-back topology [67]. Figure 4.13 shows the ratio of the jet energy scale response defined as $R = p_T^{jet}/p_T^{ref}$ in data and MC, where p_T^{jet} is p_T of the jet and p_T^{ref} is p_T of the reference object. As a reference object, Z, γ, or jet is used. The uncertainty of the jet energy scale is at a level of 2 % at 25 GeV and 1 % at 100 GeV. The ratio R_{data}/R_{MC} is used to correct the jet energy scale in data.

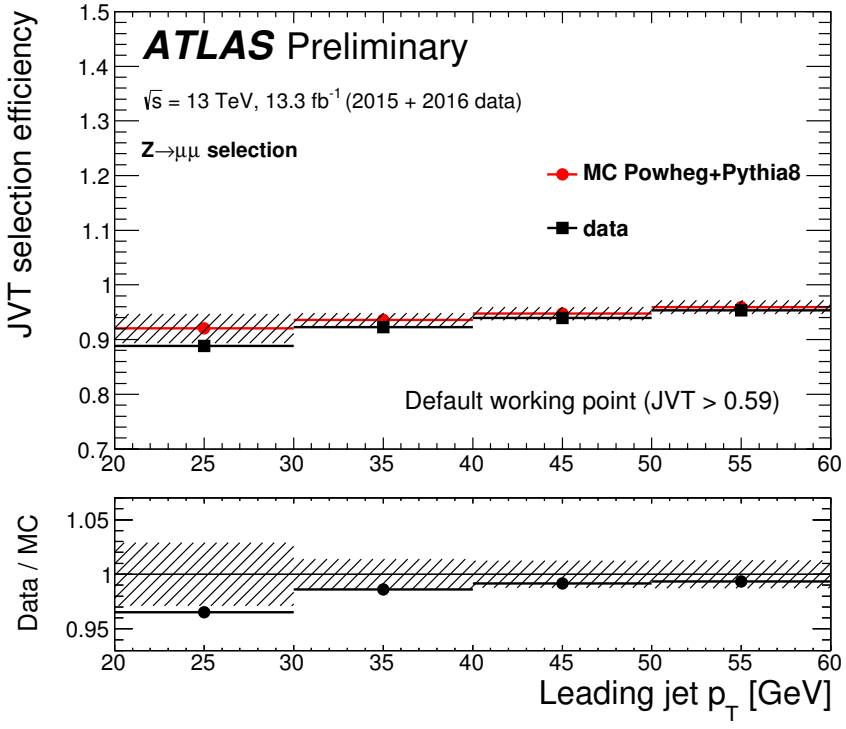


Figure 4.12: Jet selection efficiency of a JVT > 0.59 requirement as a function of leading jet p_T [66].

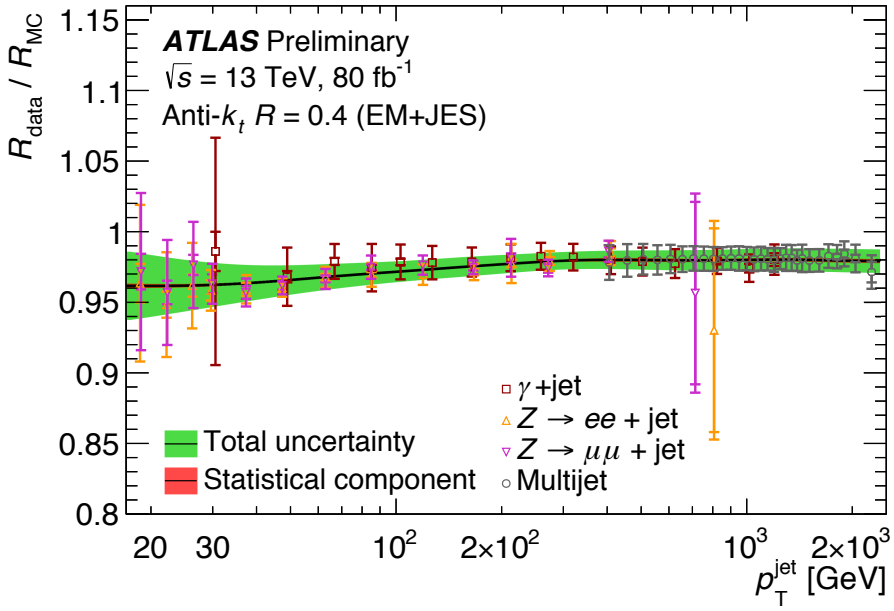


Figure 4.13: Ratio of the jet energy scale response in data to that in MC as a function of jet p_T measured with Z+jet, γ+jet and multi-jet events [68].

4.6 b-tagging

b-jet tagging algorithm

Since b-hadrons tend to fly $O(1)$ mm from the proton-proton interaction point before the decay, jets originated from b-quark (b-jet) can be distinguished from others (light-jet originated from u-, d-, s-quark and gluon, and c-jet originated from c-quark) by finding out the signature of b-hadron decays [69]. Figure 4.14 illustrates the feature of a b-jet and a light-jet (including c-jet). There are three concepts to identify b-jets, as described below.

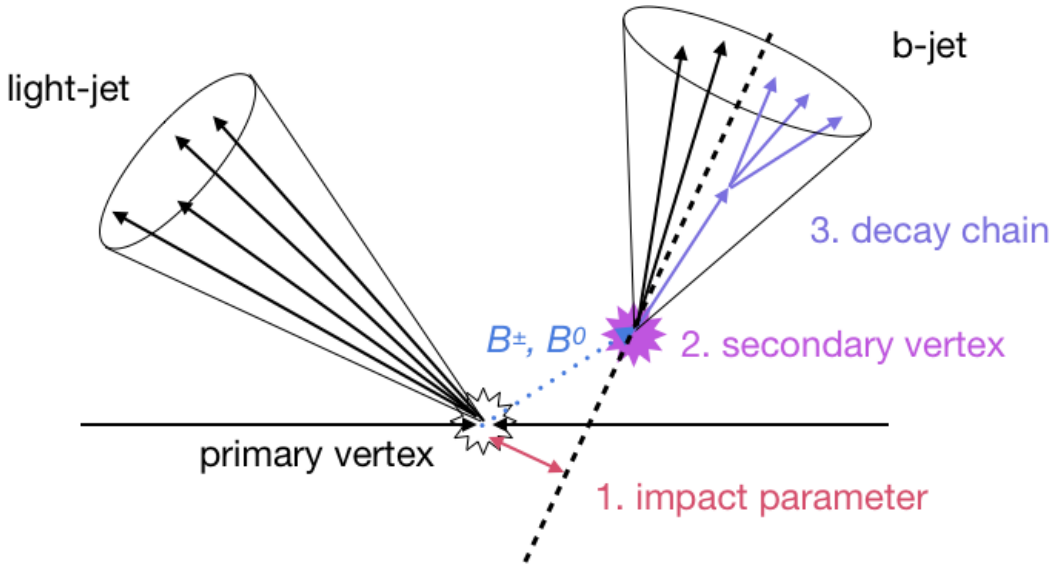


Figure 4.14: Feature of a b-jet. The b-tagging decision is made by impact parameter, secondary vertex and decay chain related variables.

1. Impact parameter.

The impact parameter of tracks decayed from a b-hadron tends to be larger than the ones from light hadrons. Therefore the impact parameter of tracks in a jet can be a discriminant between b-jets and other jets. In the actual implementation, likelihood of the impact parameters are introduced with the probability distribution functions generated by MC. The log likelihood ratio of b- and c- or light-jet is used as one of the b-tagging discriminants.

2. Secondary vertex reconstruction based variables

The existence of secondary vertex (SV) is a strong indication of b-jets. The SV reconstruction begins with forming all possible two-track vertices using tracks matched to topo-clusters in a jet. Vertices formed by heavy particle decays such as K and Λ, or photon conversion are removed by checking the invariant mass or the position. All two-track vertices are then combined into a single vertex using an iterative fitting procedure to remove the track that gives the worst χ^2. The vertex fit is repeated until an acceptable χ^2 is obtained. Some of the properties related to SV have discriminating power against light-jets. The variables used as the discriminants are listed below.

- Invariant mass of tracks from the SV: b-jets have larger mass than that of light-jets or c-jets.
- Energy fraction of tracks from the SV with respect to the jet: b-jets have larger energy fraction.
- The number of tracks from the SV: it is larger for b-jets.

- The number of two-track vertex candidates: b-jets have more candidates.
- Distance between the PV and the SV: b-jets have larger distance.
- ΔR between the jet axis and the direction of the SV relative to the PV: that of b-jets is smaller than others.

3. Multi-vertex reconstruction to follow decay chain.

Not only the inclusive secondary vertices but also the b-hadron's full decay chain (i.e. PV $\rightarrow$ b-hadron decay $\rightarrow$ c-hadron decay) can be reconstructed [70]. The reconstruction algorithm assumes that all tracks from the b- and c-hadron decays are produced on the b-hadron flight axis. A Kalman filter is used to find a common direction connecting the PV to both the b and c decay vertices. With this approach, the decay vertices can be resolved even when only a single track can be found on the b-hadron flight axis. The variables related to the b- and c-hadron decay vertices (called displaced vertices in the following) used as the discriminants are listed below.

- Invariant mass of tracks from displaced vertices
- Energy fraction of tracks from the displaced vertices with respect to the jet
- The number of displaced vertices
- The number of tracks from displaced vertices
- Significance of the average distance between the PV and displaced vertices
- ΔR between the jet axis and vectorial sum of momenta of all tracks from displaced vertices

All three types of discriminants described above are used as inputs of a multivariate analysis (MVA) to identify b-jets. In this analysis, the MV2c10 algorithm [71] is used as a MVA, which is based on the Boosted Decision Tree (BDT) machine learning. The MV2c10 is trained by MC $t\bar{t}$ events. The signal is b-jets originating from top-quark decays while the background is the combination of inclusive c- and light-jets. Figure 4.15 shows the performance of the b-tagging with the MV2c10. As shown in the left plot, the MV2c10 makes strong discrimination between b-jets and light-jets while the separation from c-jets is worse than that from light-jets because c-hadron also flies from the proton-proton interaction point before the decay. In this analysis, jets with the BDT score, which is the output of the MV2c10, larger than 0.64 are classified as b-jets, which gives 77 % selection efficiency to true b-jets.

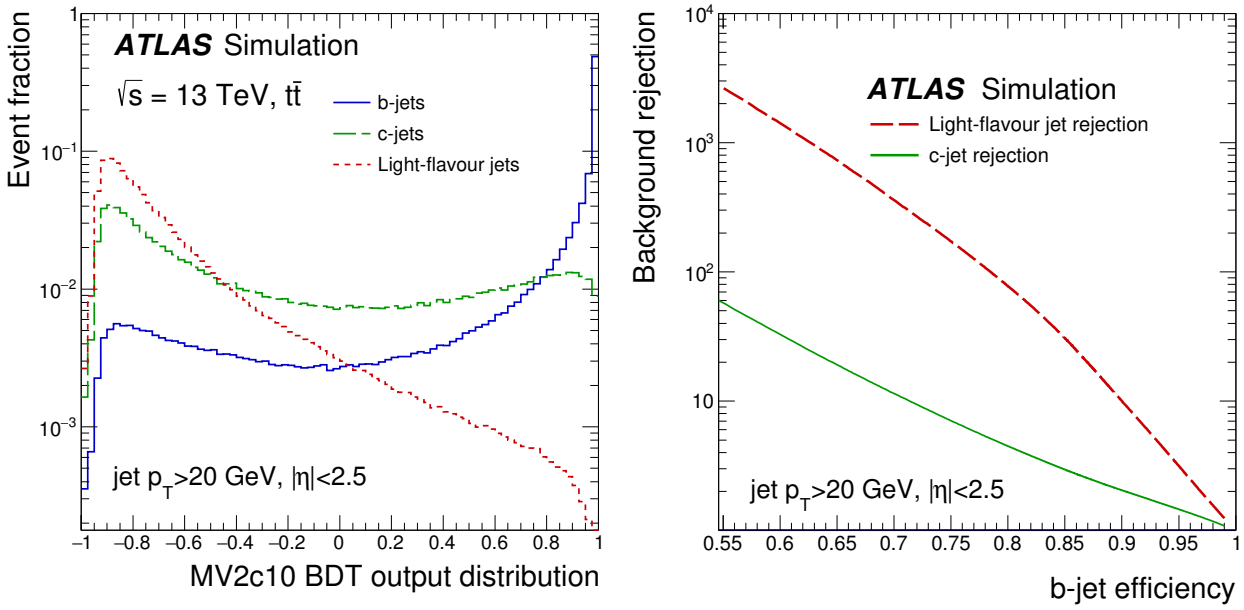


Figure 4.15: The MV2c10 BDT output (left) and the light-jet and c-jet rejection factor as a function of the b-jet efficiency (right). The performance for b-, c-, and light-jet is evaluated using $t\bar{t}$ MC events [72].

b-jet tagging efficiency

The b-jet tagging efficiency is measured with data using $t\bar{t}$ di-lepton channel ($t\bar{t} \to bWbW \to bbll\nu\nu$) [73]. Opposite charged, and different flavor lepton pair ($e + \mu$) is required to exist in each event to reduce Z boson decays. In addition, exactly two jets are required to exist to reduce additional jets coming from gluon radiation in the initial and final states. This high purity $t\bar{t}$ sample provides the b-jet without using the requirement of the b-tagging. The b-jet tagging efficiency is calculated by counting jets passing the b-tagging selection. Figure 4.16 shows the measured b-jet tagging efficiency and its scale factor (ratio of the efficiency measured by data to that by MC) as a function of jet p_T. The total uncertainty of the scale factor ranges from 1.0% to 7.6% depending on the jet p_T.

4.7 Missing transverse energy

The missing transverse energy E_T^{miss} is defined as the momentum imbalance on the $x - y$ plane, formulated as

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2}, \tag{4.10}$$

$$E_{x(y)}^{miss} = -\Sigma E_{x(y)}^{\gamma} - \Sigma E_{x(y)}^{e} - \Sigma E_{x(y)}^{\mu} - \Sigma E_{x(y)}^{jet} - \Sigma E_{x(y)}^{soft}, \tag{4.11}$$

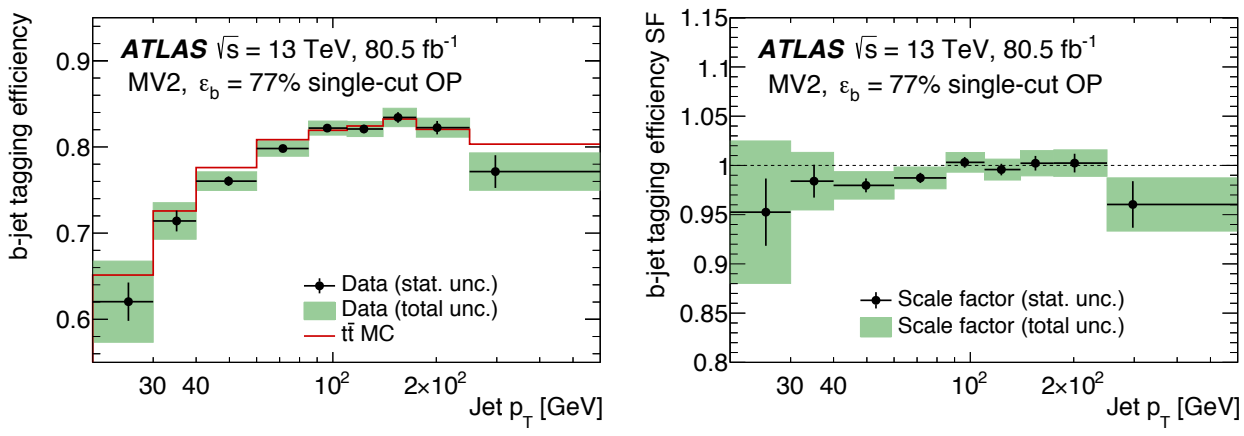


Figure 4.16: The b-jet tagging efficiency (left) and the scale factor (right) evaluated by data and $t\bar{t}$ MC sample [73].

where $\Sigma E_{x(y)}^{\gamma,e,\mu,jet}$ is the vectorial sum of $x(y)$ momentum of each object. These contributions are referred to as hard terms in the following, while $E_{x(y)}^{soft}$ is called as soft term which is the vectorial sum of $x(y)$ momentum of particles that are not reconstructed as either photon, electron, muon nor jet. The soft term is calculated by tracks associated to the primary vertex, except for tracks matched to the hard term objects, because the calorimeter-based soft term suffers from objects coming from pile-up vertices [74], while the track based soft term is more robust against the pile-up.

4.8 Overlap removal

Since the reconstruction of each object is proceeded independently, the same topo-cluster or track could belong to the different reconstructed objects. In order to avoid such a double-counting, the following "overlap removal" procedure is applied, where objects overlapping with photons are preferentially removed to keep the photon efficiency high. If the distance of two objects $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is closer than 0.4, one of them is removed with the following order;

1. jet
2. electron or muon
3. photon

except if an electron is overlapped with any jets within $0.2 < \Delta R < 0.4$, the electron is removed.

Chapter 5

Event selection and categorization

In this chapter, the event selection and categorization criteria are described. The first basic event selection is called pre-selection described in Section 5.1. The signal event, $ttH(H \to \gamma\gamma)$, has two components in the final state; two photons coming from a Higgs boson, and other objects coming from a top-quark pair. The di-photon and top-quark pair selection criteria are presented in Section 5.2 and 5.3, respectively. Events are further selected to reduce backgrounds using kinematic variables of all final state objects. The selected events are finally separated into categories which have different signal to background ratios to gain the sensitivity. The multivariate analysis (MVA) is used for the categorization and final event selection, as discussed in Section 5.4.

5.1 Pre-selection

The following selections are applied as the basic event selection:

- Require events to pass the di-photon trigger.
- Require events to be in "Good for physics" condition as discussed in Section 3.1 to assure all the sub-detectors work properly.
- Require events to have a primary vertex to ensure the event quality.

5.2 Di-photon selection

The characteristics of two photons coming from a Higgs boson decay is to have relatively higher p_T than continuum backgrounds in addition to have an invariant mass peak around 125 GeV. There are many hadronic jets in the reconstructed photons due to mis-identification. In order to reduce them, further photon selection criteria are required to pass, as discussed below.

At least two photons, which pass the *Loose* selection criteria, with $p_T > 25$ GeV and $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$, are required to exist at the first stage of the di-photon

selection. The leading and sub-leading p_T photons are considered as the $H \to \gamma\gamma$ candidate and required the following selection to pass.

Trigger matching

The leading and sub-leading photons must be the objects firing the di-photon trigger.

Tight selection

The *Tight* selection is implemented to reduce jets mis-identified as photons using the shower shape in the EM calorimeter, as described in Section 4.2. Both photons are required to pass the *Tight* selection.

Isolation selection

In addition to the *Tight* selection, the isolation selection is implemented to reduce misidentified jets. Signal photons are isolated from any other objects because they are the decay objects from a heavy object, i.e. Higgs, while fake photons tend to be close to hadrons. To exploit this difference, two types of isolation requirements are implemented.

- Calorimeter isolation: $E_T^{iso}/p_T < 0.065$,
  where E_T^{iso} is the sum of the transverse energy of topo-clusters within an opening $\Delta R < 0.2$ [1] with respect to the photon direction after subtracting the energy deposited by the photon itself.

- Track isolation: $p_T^{iso}/p_T < 0.05$,
  where p_T^{iso} is the sum of the transverse momentum of all the tracks with $p_T > 1$ GeV and having a distance of closest approach to the primary vertex along the beam axis of less than 3 mm within an opening $\Delta R < 0.2$ with respect to the photon direction.

Figure 5.1 and 5.2 show the distributions of the calorimeter and track isolation variables, respectively. Distributions for the data and the signal MC sample are drawn in both plots to compare them. Since the photon samples in data contains jets, both calorimeter and track isolation variables for data are relatively larger than that for MC (note that the track isolation distribution is in a log scale). Photons with $p_T^{iso}/p_T = 0$ mean that there are no tracks within $\Delta R < 0.2$ with respect to the photon direction.

Relative p_T cut

The p_T requirement relative to $m_{\gamma\gamma}$ is applied instead of the absolute value of p_T. Figure 5.3 shows the $p_T/m_{\gamma\gamma}$ distribution for the data and the signal MC sample. The $p_T/m_{\gamma\gamma}$ for signal is larger than that for background for both leading and sub-leading photons. In this analysis, events are required to satisfy $p_T/m_{\gamma\gamma} > 0.35$ for the leading photon and $p_T/m_{\gamma\gamma} > 0.25$ for the sub-leading photon.

[1] ΔR is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, where $\Delta\eta$ and $\Delta\phi$ are the differences between two particles in η and ϕ, respectively

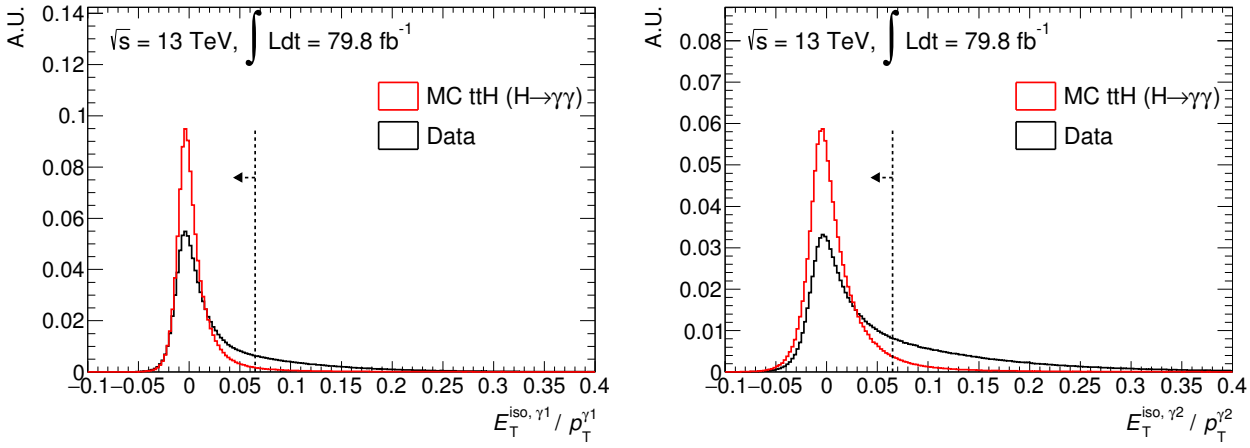


Figure 5.1: Distributions of the calorimeter isolation for the leading (left) and sub-leading (right) photons. All the photon selections are applied except for the isolation requirement. The region shown by an arrow is selected. Events are normalized to unity.

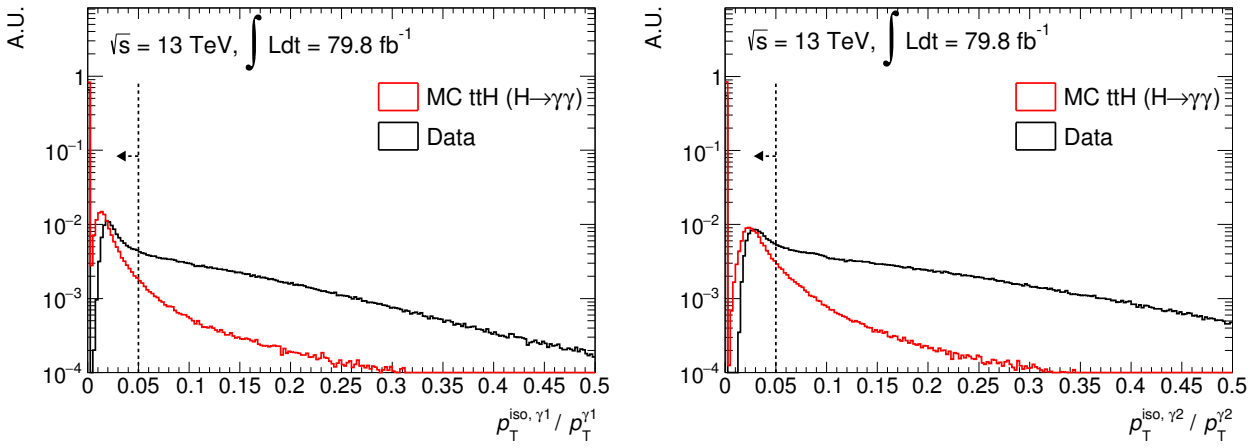


Figure 5.2: Distributions of the track isolation for the leading (left) and sub-leading (right) photons. All the photon selections are applied except for the isolation requirement. The region shown by an arrow is selected. Events are normalized to unity.

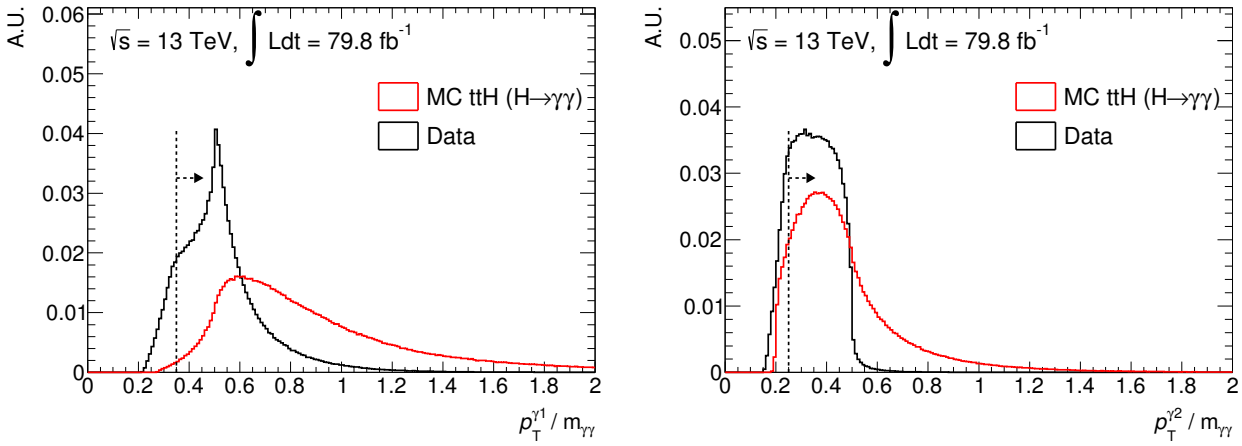


Figure 5.3: Distributions of the $p_T/m_{\gamma\gamma}$ for the leading and sub-leading photons. All the photon selections are applied except for the $p_T/m_{\gamma\gamma}$ requirement. The region shown by an arrow is selected. Events are normalized to unity.

After the photon requirements, the di-photon invariant mass is required to be in the range of $105 < m_{\gamma\gamma} < 160$ GeV.

5.3 Top-quark pair selection

Top-quarks decay into W boson and b-quark with the 100 % branching ratio. In this analysis, events are separated into two signal regions depending on the W decay mode; the "Leptonic" and "Hadronic" regions. The Leptonic region requires that at least one charged lepton exists in the final state. In this region, $t\bar{t} \to b\bar{b}q\bar{q}l^+\overline{\nu}$, $t\bar{t} \to b\bar{b}q\bar{q}l^-\nu$ and $t\bar{t} \to b\bar{b}l^+l^-\overline{\nu}\nu$ are inclusively considered. The Hadronic region requires no charged leptons exist in the final state, assuming both W bosons decay into quarks.

Leptonic region

In the Leptonic region, at least one electron or muon is required to exist in addition to pass all di-photon selection criteria. Either electron or muon is required to pass the isolation selection criteria in addition to the selection explained in Section 4.3 for electron and 4.4 for muon, respectively. There are non-resonant $\gamma\gamma$ events and small contribution of $W + \gamma\gamma$ events in the Leptonic region. In order to reduce them, at least one jet is b-tagged. Figure 5.4 shows the two dimensional distribution of the number of jets and the number of b-jets for data and $ttH(H \to \gamma\gamma)$ MC, where the di-photon selection is passed. The large amount of background events are reduced by the requirement of b-jet.

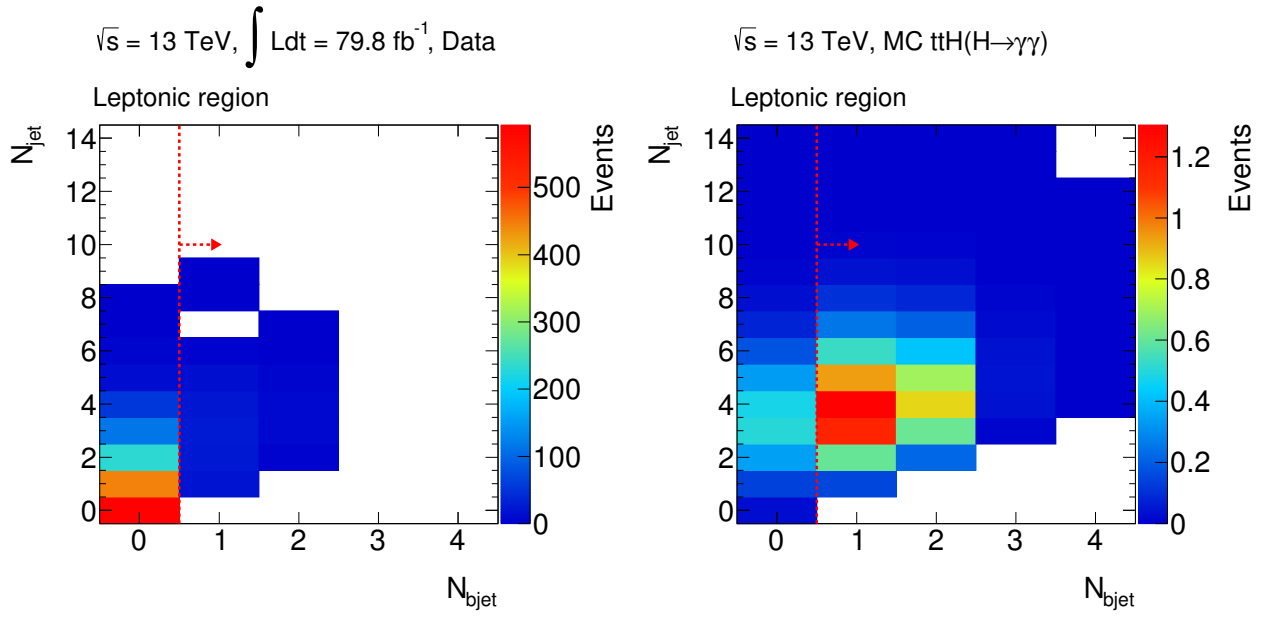


Figure 5.4: Two dimensional distributions of the number of jets and the number of b-jets for data (left) and $ttH(H \to \gamma\gamma)$ MC (right) in the Leptonic region. At least one electron or muon is required to exist, in addition to pass all the di-photon selection criteria.

Hadronic region

Events failing the selection for Leptonic region are classified as the Hadronic region. The main background source is non-resonant $\gamma\gamma$ event as the Leptonic region. The final state of a top-quark pair in the Hadronic region has six jets, while $\gamma\gamma$ events tend to have less number of jets. Thus the requirement of the number of jets is helpful to reduce $\gamma\gamma$ events, as can be seen in Figure 5.5, which shows the two dimensional distribution of the number of jets and the number of b-jets for data and MC. At least three jets and at least one b-tagged jet are required to exist for the Hadronic region.

Table 5.1 shows the cut flow of the pre-selection, the photon selection and the top-quark pair selection, separating into the Leptonic and Hadronic regions. The number of events in the Leptonic region is 139 in total while ones in the Hadronic region is 4,869 in total. Assuming the SM, the signal yield for the Leptonic and Hadronic regions are 7.9 and 17.7, respectively. There are still poor signal to background ratio in both regions (especially for the Hadronic region) due to non-resonant $\gamma\gamma$ background at this stage.

Table 5.1: Cut flow of the di-photon and top-quark pair selection. The MC denotes the $ttH(H \to \gamma\gamma)$ samples. The Data column represents the number of events while the MC column represents the acceptance with respect to all MC events. The cut flow is separated into 2015, 2016 and 2017 since the trigger requirement is changed in the 2017 data taking.

Selection	Data 2015	Data 2016	Data 2017	MC for 2015+2016	MC for 2017
				100 %	100 %
Pre-selection	12,199,177	122,360,256	62,504,764	77.6 %	77.2 %
Two photons	3,544,701	36,090,940	16,360,562	60.9 %	60.0 %
Trigger match	2,597,327	26,987,092	8,062,623	57.8 %	51.7 %
Tight	429,883	4,246,802	4,750,250	48.7 %	46.8 %
Isolation	156,537	1,523,797	1,834,630	40.3 %	37.7 %
Relative p_T	136,038	1,324,491	1,595,147	36.7 %	34.6 %
$m_{\gamma\gamma}$	30,162	298,882	374,981	36.4 %	34.2 %
Leptonic region					
$N_{e,\mu} \geq 1$	79	669	855	11.3 %	10.4 %
$N_{b-jet} \geq 1$	11	61	67	9.0 %	8.3 %
Hadronic region					
$N_{e,\mu} = 0$	30,083	298,213	374,126	25.1 %	23.8 %
$N_{jet} \geq 3$	1,899	22,594	36,837	24.4 %	23.1 %
$N_{b-jet} \geq 1$	172	1,927	2,770	19.9 %	18.8 %

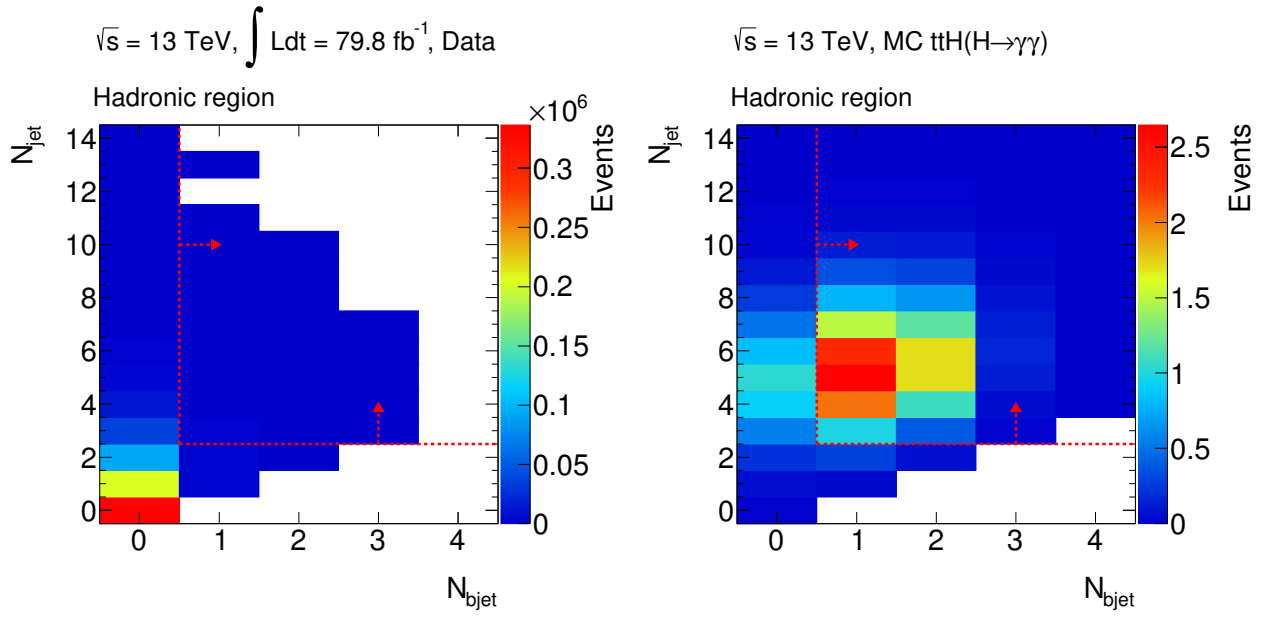


Figure 5.5: Two dimensional distributions of the number of jets and the number of b-jets for data (left) and $ttH(H \to \gamma\gamma)$ MC (right) in the Hadronic region. Neither electron nor muon is required to exist, in addition to pass all the di-photon selection criteria.

5.4 Event categorization using MVA technique

5.4.1 Concept

After the event selection, events in the Leptonic (Hadronic) region are further categorized into three (four) groups with the different signal to background ratio to improve the signal sensitivity. Below we describe why such a categorization improves the sensitivity.

The sensitivity is evaluated by the statistical significance of observing signal. In order to introduce the significance, the likelihood function $L(\mu)$ is defined as

$$L(\mu) = \frac{(\mu \cdot s + b)^n}{n!} e^{-(\mu \cdot s + b)}, \tag{5.1}$$

where s and b are the numbers of signals and backgrounds expected by SM, respectively, n is the number of observed events, and μ is called signal strength, with $\mu = 1$ corresponding that the signal yield is consistent with the SM prediction and $\mu = 0$ corresponding to the background-only hypothesis. The expected significance Z_0, concerning $\mu = 1$, is defined as [75]

$$Z_0 = \sqrt{-2 \ln \frac{L(0)}{L(1)}} = \sqrt{2 \left((s+b) \ln \left(1 + \frac{s}{b} \right) - s \right)}. \tag{5.2}$$

Next, assuming that events are classified into N categories, the likelihood function $L^{cat}(\mu)$ is expanded from Equation 5.1 to

$$L^{cat}(\mu) = \prod_{i}^{N} L_i(\mu) = \prod_{i}^{N} \frac{(\mu \cdot s_i + b_i)^{n_i}}{n_i!} e^{-(\mu \cdot s_i + b_i)}, \tag{5.3}$$

where s_i, b_i and n_i are the number of expected signals, backgrounds and observed events in the i-th category, respectively. The significance in case with the categorization, Z_0^{cat}, is then obtained as

$$Z_0^{cat} = \sqrt{\sum_{i}^{N} (Z_{0,i}^{cat})^2}, \quad Z_{0,i}^{cat} = \sqrt{2\left((s_i + b_i)\ln\left(1 + \frac{s_i}{b_i}\right) - s_i\right)}. \tag{5.4}$$

Then, the relation between Z_0^{cat} and Z_0 becomes

$$Z_0^{cat} > Z_0 \quad (\text{if } \frac{s}{b} \neq \frac{s_i}{b_i} \text{ for any } i) \tag{5.5}$$

or

$$Z_0^{cat} = Z_0 \quad (\text{if } \frac{s}{b} = \frac{s_i}{b_i} \text{ for all } i), \tag{5.6}$$

implying that the categorization which have the different signal to background ratio improves the significance. The key point of the categorization is that events failing the selection contribute to gain the significance.

The event categorization is based on the MVA. Four-vector-like variables (p_T, η, ϕ, E) of all objects in the final state are used as the inputs of the MVA which is actually the XGBoost (eXtreme Gradient Boosting) package [76]. The XGBoost outputs a discriminant showing a compatibility with signal (called XGBoost score in the following). Based on the XGBoost score, events are categorized into three groups for the Leptonic region and four groups for the Hadronic region.

Later in this section, the input variables for the XGBoost are described in Section 5.4.2. The technical description of the XGBoost, and its training are presented in Section 5.4.3. The procedure of the event categorization is described in Section 5.4.4.

5.4.2 Input variables

In order to reduce backgrounds, events are further selected using the difference of kinematics between the signal and background. However, since there are many objects in the final state, it is hard to find small differences by human eyes. For this reason, we use four-vector-like variables of all objects in the final state as the inputs of the MVA. Table 5.2 and 5.3 summarize the input variables for the Leptonic and Hadronic regions, respectively. The distribution of each input variable can be found in Appendix B.

Table 5.2: Input variables of the XGBoost for the Leptonic region.

Object	The number of objects	Variable
Photon	two leading photons	$p_T/m_{\gamma\gamma}$, η, ϕ, $E/m_{\gamma\gamma}$
Jet	up to four leading jets	p_T, η, ϕ, E, b-tagging decision
Lepton (e or μ)	up to two leading leptons	p_T, η, ϕ, E
E_T^{miss}	-	$\lvert E_T^{miss}\rvert$, ϕ

Table 5.3: Input variables of the XGBoost for the Hadronic region.

Object	The number of objects	Variable
Photon	two leading photons	$p_T/m_{\gamma\gamma}$, η, ϕ, $E/m_{\gamma\gamma}$
Jet	up to six leading jets	p_T, η, ϕ, E, b-tagging decision
E_T^{miss}	-	$\lvert E_T^{miss}\rvert$, ϕ

5.4.3 Technical description of the XGBoost training

XGBoost

The XGBoost uses the Boosted Decision Tree (BDT) procedure as shown in Figure 5.6. In the stage of the machine learning (called "training" in the following), many different decision trees are created iteratively, tuning parameters at each step. A score corresponding to "signal-like" parameter is assigned in each branch of trees. The sum of scores obtained by all decision trees is the output of the machine learning algorithm. In the XGBoost, the "Gradient Boosting [77]" algorithm is adopted, which uses the gradient descent to optimize parameters at the iteration. The machine learning process of the XGBoost is faster than that of the nominal Gradient Boosting method because of the improvements of the optimization method.

XGBoost training

Two dedicated types of decision trees are trained to discriminate signal from backgrounds for the Leptonic and Hadronic regions, respectively. The signal and background events used for the development of decision trees are separated into two samples as following:

- *Training* sample: events used for the training.
- *Validation* sample: events used for the validation of the training result.

For the signal samples, $ttH(H \to \gamma\gamma)$ MC is used for the training and validation of both the Leptonic and Hadronic regions. Events are selected by the di-photon and top-quark pair selection criteria as described in Section 5.2 and 5.3. In order to avoid to use the same dataset for the training and validation, the dataset is separated into two for each sample. Since a lot of MC events are generated for this analysis as mentioned in Section 3.2.2, the number of events is enough to develop decision trees.

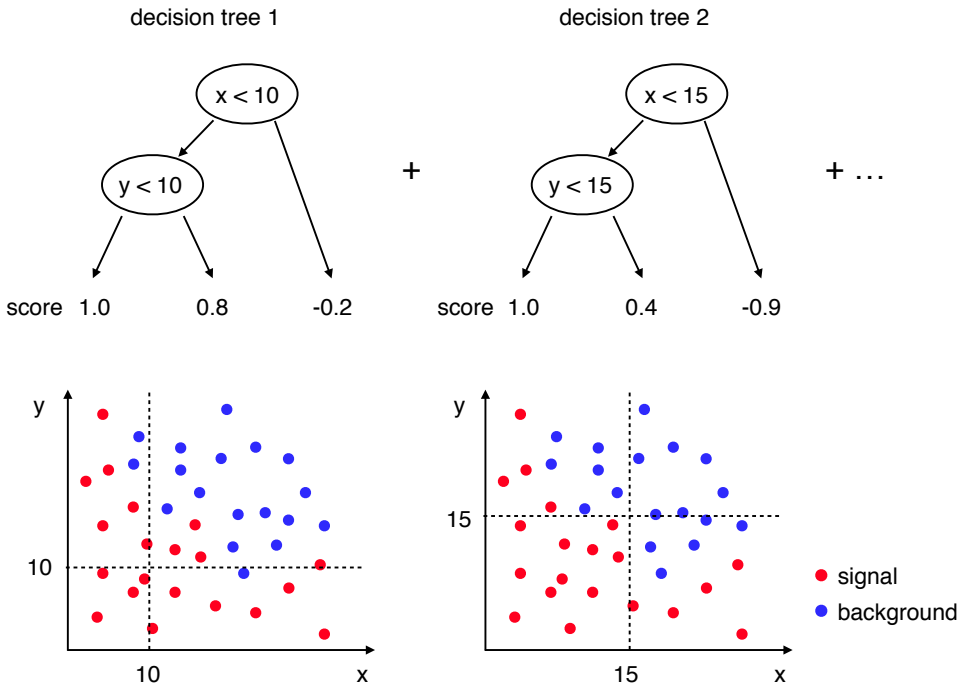


Figure 5.6: Illustration of the Boosted Decision Tree procedure.

Table 5.4: Summary of the signal and background samples used for the training and validation. Samples are prepared separately for the Leptonic and Hadronic regions.

	Leptonic region		Hadronic region	
	Sample	Selection	Sample	Selection
Signal training	MC $ttH(H \to \gamma\gamma)$	di-photon selection $N_{b-jet} \geq 1$	MC $ttH(H \to \gamma\gamma)$	di-photon selection $N_{jet} \geq 3$ $N_{b-jet} \geq 1$
Signal validation		di-photon selection $N_{b-jet} \geq 1$		di-photon selection $N_{jet} \geq 3$ $N_{b-jet} \geq 1$
Background training	Data	di-photon NTI selection $N_{b-jet} = 0$	Data	di-photon NTI selection $N_{jet} \geq 3$ $N_{b-jet} \geq 1$
Background validation		di-photon NTI selection $N_{b-jet} \geq 1$		di-photon NTI selection $N_{jet} \geq 3$ $N_{b-jet} \geq 1$

For the background samples, data is used for the training and validation of both the Leptonic and Hadronic regions. The ideal samples for both the training and validation are those after the di-photon and top-quark pair selections. However, due to lack of statistics, events from a data control region is used as the background sample. Such region in data is defined with the same selection as each Leptonic and Hadronic region, except that at least one photon has to fail either *Tight* identification or isolation requirement. This selection is referred to as "Non-Tight-Isolated (NTI)" in the following. Furthermore, only for the training sample for the Leptonic region, another control region, where no b-tagged jets are required in addition to the NTI requirement, is defined. The sets of the signal and background samples for the training and validation are summarized in Table 5.4.

The machine learning may excessively fit to the training sample. This phenomenon is called "over training". In order to check the over training, the validation sample is used by comparing the XGBoost score distribution with that made from the training sample. Figure 5.7 shows the distribution of the XGBoost score for the Leptonic and Hadronic regions. The higher XGBoost score represents the more signal-like events. For the Leptonic region, the disagreement of the background distributions can be seen due to the difference of the event selection criteria between the training and validation samples. However, the distribution of the signal training sample is in good agreement with that of the signal validation sample. The over training is not observed. For the Hadronic region, the distribution is in good agreement between the training and validation samples for both the signal and background.

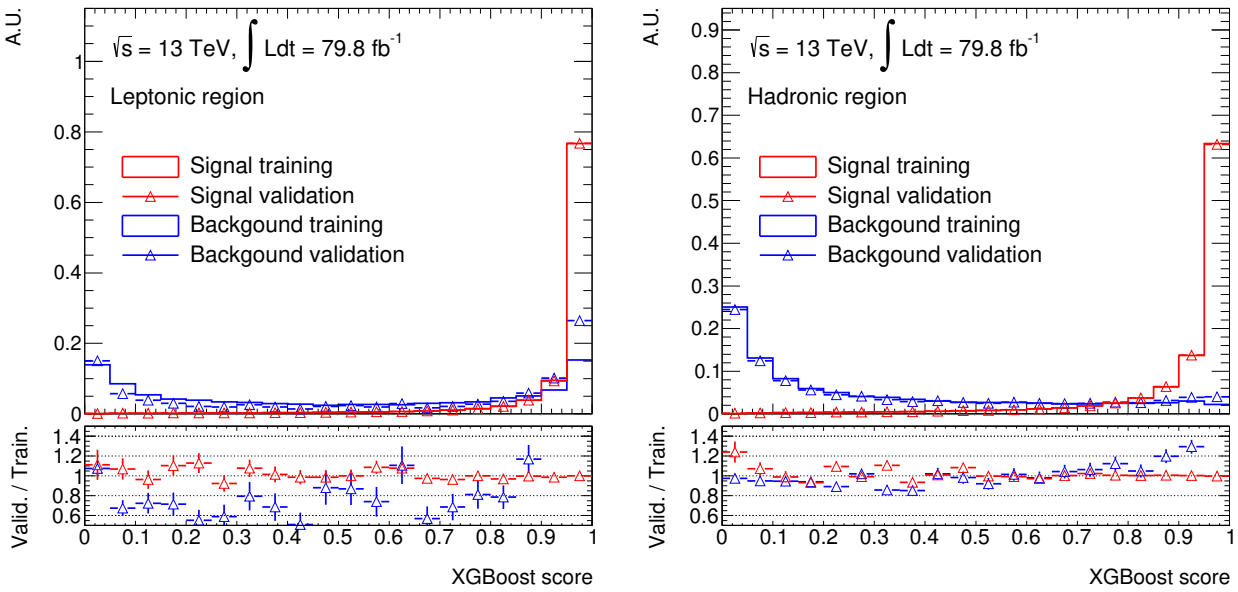


Figure 5.7: The XGBoost score distribution for the Leptonic (left) and the Hadronic (right) regions. Events whose di-photon invariant mass in a range $123 < m_{\gamma\gamma} < 127$ GeV are removed from the validation sample. The higher XGBoost score represents the more signal-like events. Events are normalized to unity.

Since the XGBoost score is calculated to discriminate $ttH(H \to \gamma\gamma)$ events from inclusive continuum backgrounds using data, it is important to know which types of background components are rejected or selected in the training. Figure 5.8 shows the XGBoost score distribution for each signal MC, background MC and data. It indicates that events including objects coming from $t\bar{t}$ have higher XGBoost score, and thus the $t\bar{t} + \gamma\gamma$ events are selected as well as the $ttH(H \to \gamma\gamma)$ events.

The most effective set of inputs is the (p_T, η, ϕ, E) of jets coming from $t\bar{t}$ for this XGBoost training. On the other hand, the discrimination of $ttH(H \to \gamma\gamma)$ events from $t\bar{t} + \gamma\gamma$ events is difficult since the kinematic topology is very similar. It is discussed in chapter 7.

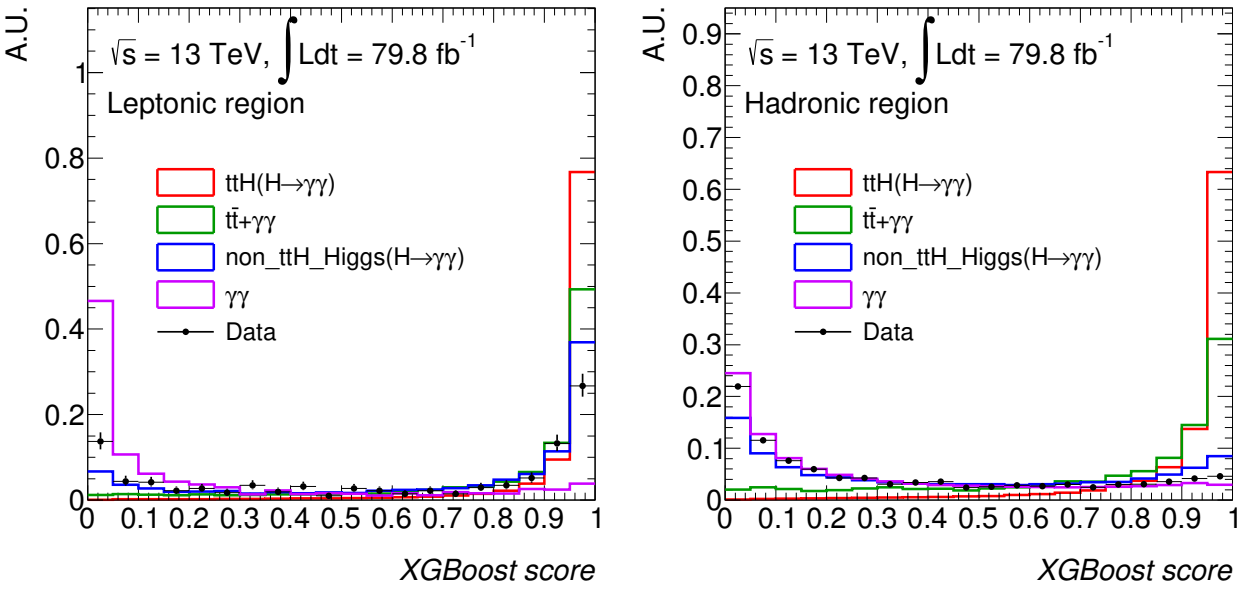


Figure 5.8: The XGBoost score distribution for the Leptonic (left) and Hadronic (right) regions. The non-ttH Higgs process denotes all Higgs boson production except for the $ttH(H \to \gamma\gamma)$ process as listed in Table 3.2. Events are normalized to unity.

5.4.4 Event categorization

After the di-photon and top-quark pair selections, events are categorized based on the XGBoost score. The categorization is proceeded as the following:

1. Define categories with certain boundaries of the XGBoost score.

2. Estimate the numbers of expected signals and backgrounds using the $m_{\gamma\gamma}$ distribution for each category. Both the signal and background events must be in the $m_{\gamma\gamma}$ selection; $123 < m_{\gamma\gamma} < 127$ GeV (called "signal window").

3. Calculate the combined significance Z_0^{cat} using Equation 5.4.

4. Repeat this process from 1, iteratively scanning boundaries.

5. Find the set of boundaries maximizing the combined significance.

Figure 5.9 shows the categorization procedure. The detail for each process is described below.

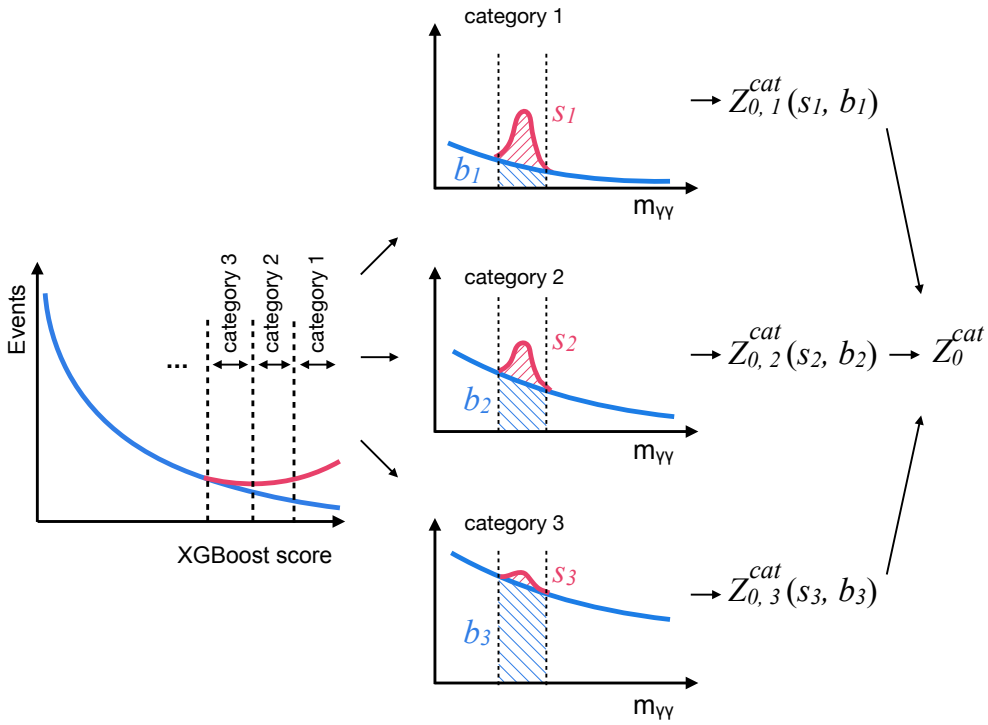


Figure 5.9: Illustration of the categorization procedure.

Signal estimation

The number of expected signals s, assuming the SM, is calculated as;

$$s = L \times \sigma_{ttH} \times BR(H \rightarrow \gamma\gamma) \times \varepsilon_{acc}. \tag{5.7}$$

The integrated luminosity L is measured to be 79.8 fb^{-1} while the cross section σ_{ttH} and branching ratio $BR(H \rightarrow \gamma\gamma)$ are the SM predictions. The acceptance ε_{acc} is estimated from MC.

Background estimation

The background events are classified into two types; resonant and continuum backgrounds as discussed in Section 1.3.2. The sources of the resonant backgrounds are $H \rightarrow \gamma\gamma$ decays

where Higgs bosons are produced by non-ttH processes. The number of expected resonant backgrounds, N^{exp}_{res-BG}, is estimated from the MC as well as the signal, assuming the SM.

The number of expected continuum backgrounds, $N^{exp}_{cont-BG}$, is estimated using events with $m_{\gamma\gamma} < 123$ GeV or $m_{\gamma\gamma} > 127$ GeV in real data (called "sideband"). Under the assumption that the shape of the $m_{\gamma\gamma}$ distribution with the NTI selection is the same as that of the nominal selection, $N^{exp}_{cont-BG}$ can be calculated as

$$N^{exp}_{cont-BG} = N^{sb}_{cont-BG} \times \frac{N^{sw}_{NTI}}{N^{sb}_{NTI}}, \tag{5.8}$$

where $N^{sb}_{cont-BG}$ is the number of events in the sideband, N^{sw}_{NTI} and N^{sb}_{NTI} are the numbers of events passing the NTI selection criteria in the signal window and the sideband, respectively. Figure 5.10 shows the $m_{\gamma\gamma}$ distribution with a requirement of XGBoost score > 0.5 for the Leptonic and Hadronic regions, respectively. The shape of the $m_{\gamma\gamma}$ distribution with NTI requirement is in good agreement with the one in the sideband. The number of total backgrounds is simply the sum of N^{exp}_{res-BG} and $N^{exp}_{cont-BG}$.

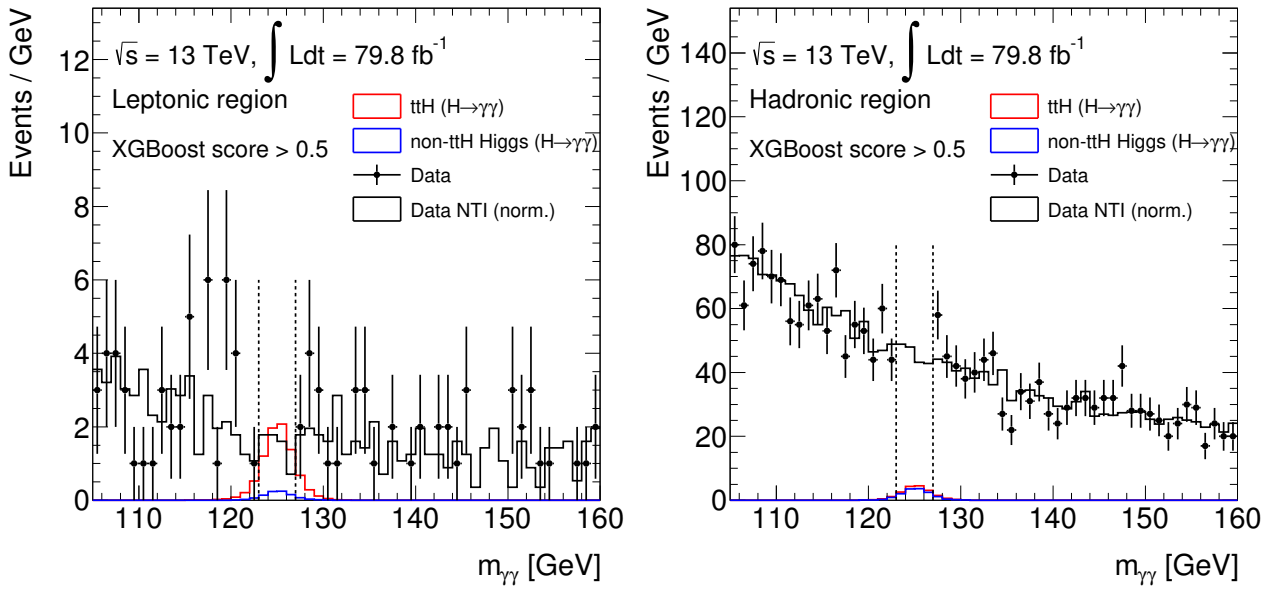


Figure 5.10: Di-photon invariant mass distributions for the Leptonic (left) and Hadronic (right) regions. The mass region $123 < m_{\gamma\gamma} < 127$ GeV is referred to as "signal window", while the outside of the signal window is called "sideband". Events in the signal window are not plotted. Events with NTI selection are also plotted by normalizing the number of events to be the same as the one in the sideband in data.

Boundary scan

The categories with the lowest XGBoost score in both the Leptonic and Hadronic regions are not considered in this analysis because these categories have the least signal to background ratio, giving the negligible significance, in addition to the large systematic uncertainty of the number of backgrounds. The boundaries are scanned with 0.001 step to maximize Z_0^{cat}.

Table 5.5 and 5.6 show the XGBoost boundaries and Z_0 of each category for the Leptonic and Hadronic regions, respectively. The combined significance is calculated from Equation 5.4. Table 5.7 and 5.8 show the numbers of expected signals and backgrounds.

Table 5.5: Expected significance for each category in the Leptonic region, using a mass window of $123 < m_{\gamma\gamma} < 127$ GeV.

Category	XGBoost score boundaries	Significance Z_0
Leptonic Cat. 1	[0.987, 1.000]	2.5 ± 0.2
Leptonic Cat. 2	[0.942, 0.987]	1.0 ± 0.1
Leptonic Cat. 3	[0.705, 0.942]	0.4 ± 0.0
Combined		2.7 ± 0.2

Table 5.6: Expected significance for each category in the Hadronic region, using a mass window of $123 < m_{\gamma\gamma} < 127$ GeV.

Category	XGBoost score boundaries	Significance Z_0
Hadronic Cat. 1	[0.996, 1.000]	2.5 ± 0.3
Hadronic Cat. 2	[0.991, 0.996]	1.1 ± 0.1
Hadronic Cat. 3	[0.971, 0.991]	0.8 ± 0.0
Hadronic Cat. 4	[0.911, 0.971]	0.4 ± 0.0
Combined		2.9 ± 0.3

Table 5.7: The number of expected signals and backgrounds for each category in the Leptonic region, using a mass window of $123 < m_{\gamma\gamma} < 127$ GeV.

Category	ttH	$ggF + bbH$	VBF	WH	ZH	$tHjb$	tWH	cont. BG
Leptonic Cat. 1	4.1	0.0018	0	0.024	0.0081	0.070	0.13	1.4
Leptonic Cat. 2	1.9	0.00050	0	0.041	0.011	0.13	0.054	2.7
Leptonic Cat. 3	0.66	0.010	0.0016	0.048	0.024	0.15	0.023	2.3

5.4.5 Mass bias check

Since the number of continuum backgrounds is estimated by fitting the $m_{\gamma\gamma}$ distribution, the mass shape must be smooth, for example, like an exponential function. The XGBoost uses kinematic variables related to di-photon system, therefore it may affect the shape of $m_{\gamma\gamma}$

Table 5.8: The number of expected signals and backgrounds for each category in the Hadronic region, using a mass window of $123 < m_{\gamma\gamma} < 127$ GeV.

Category	ttH	$ggF + bbH$	VBF	WH	ZH	$tHjb$	tWH	cont. BG
Hadronic Cat. 1	3.9	0.11	0.017	0.036	0.079	0.066	0.13	1.4
Hadronic Cat. 2	3.0	0.23	0.019	0.048	0.11	0.10	0.082	6.7
Hadronic Cat. 3	3.9	0.61	0.083	0.14	0.31	0.24	0.10	22
Hadronic Cat. 4	2.4	1.11	0.18	0.25	0.48	0.34	0.066	37

distribution. This mass bias is checked with three samples; events in data passing the NTI selection, $t\bar{t} + \gamma\gamma$ MC, and $\gamma\gamma$ MC. Figure 5.11, 5.12 and 5.13 show the $m_{\gamma\gamma}$ distributions for these three samples for the Leptonic and Hadronic regions, respectively. The shapes with or without XGBoost score cut (> 0.9) are compared, and turned out to be in good agreement. No bias for the mass distribution is found in the XGBoost score based selection.

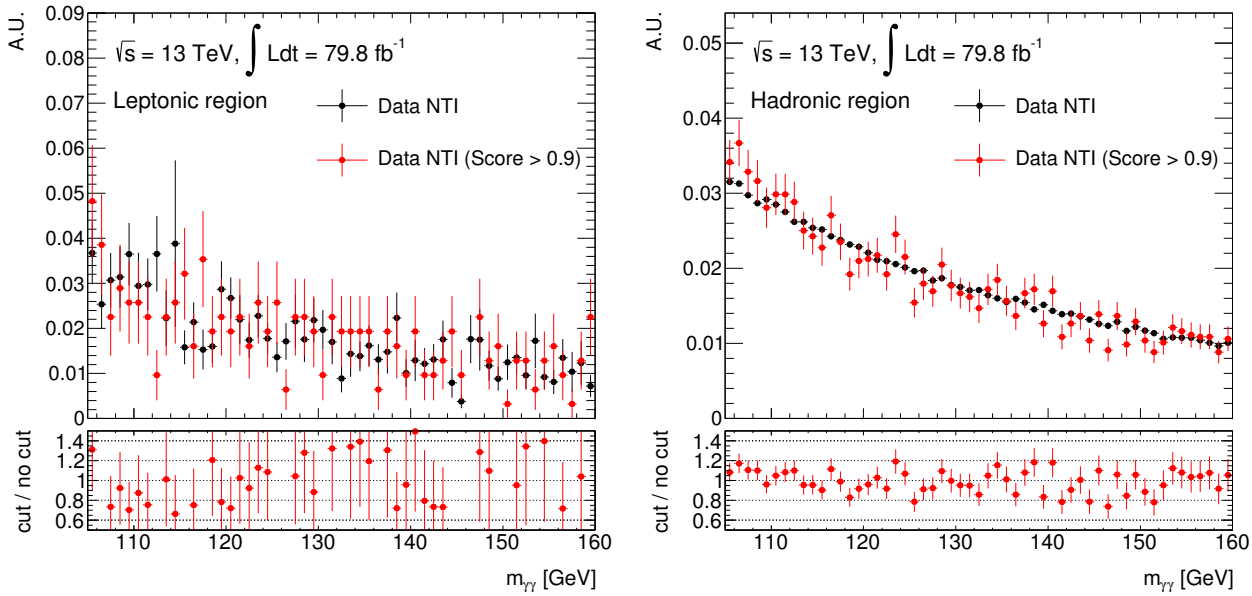


Figure 5.11: Di-photon invariant mass distribution for events from data with NTI selection for the Leptonic (left) and Hadronic (right) regions, respectively. The black dots denote all events after passing the top-quark pair selection while the red dots denote events passing XGBoost score > 0.9 in addition. Events are normalized to unity.

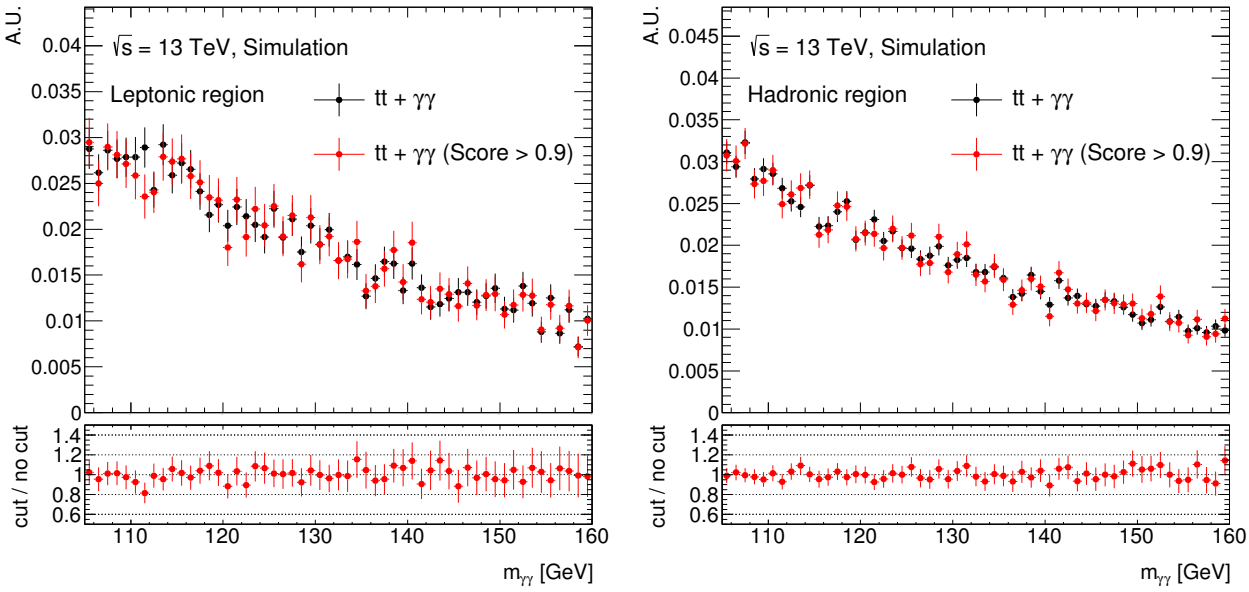


Figure 5.12: Di-photon invariant mass distribution for $t\bar{t} + \gamma\gamma$ MC sample for the Leptonic (left) and Hadronic (reght) regions, respectively. The black dots denote all events after passing the top-quark pair selection while the red dots denote events passing XGBoost score > 0.9 in addition. Events are normalized to unity.

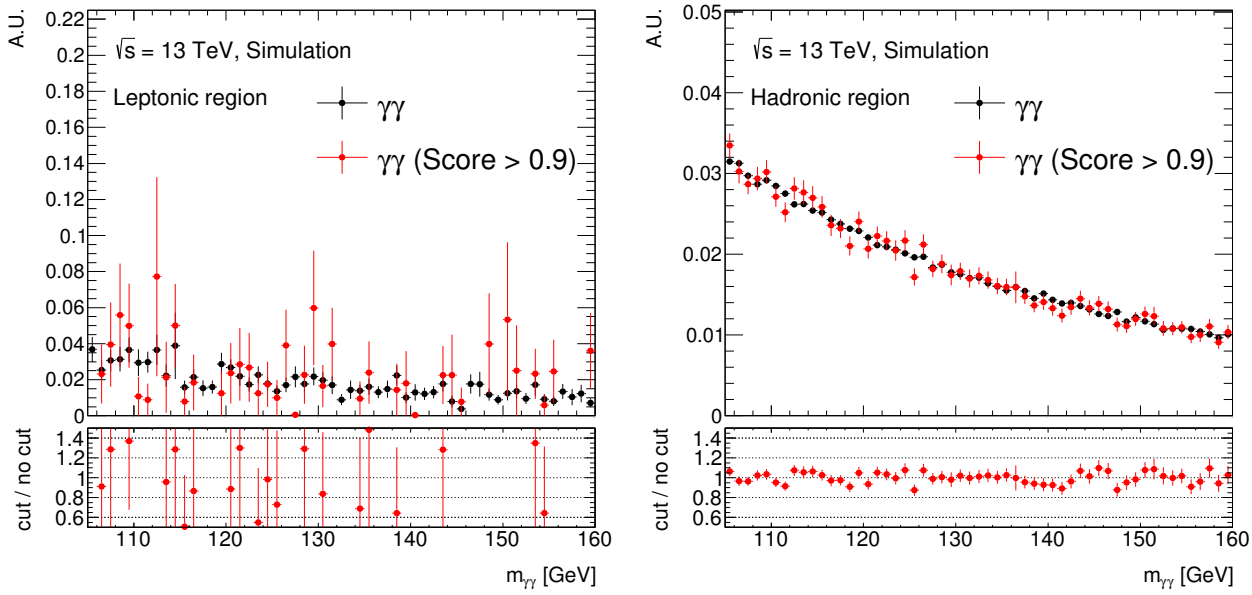


Figure 5.13: Di-photon invariant mass distribution for events from $\gamma\gamma$ MC sample for the Leptonic (left) and Hadronic (right) regions, respectively. The black dots denote all events after passing the top-quark pair selection while the red dots denote events passing XGBoost score > 0.9 in addition. Events are normalized to unity.

Chapter 6

Extraction of $ttH(H \to \gamma\gamma)$ signal

This chapter presents the $ttH(H \to \gamma\gamma)$ signal extraction using the $m_{\gamma\gamma}$ distribution. Section 6.1 presents the $m_{\gamma\gamma}$ distribution for each category in the Leptonic and Hadronic regions. Since the numbers of signals and backgrounds in each category are small, the unbinned maximum likelihood fit is used to evaluate the signal yield. Section 6.2 describes how to estimate the signal yield. The various systematic uncertainties are concerned in the fitting, which is discussed in Section 6.3. The validation of the signal extraction method is described in Section 6.4. Finally, the fitting result is presented in Section 6.5.

6.1 Mass distribution

Figure 6.1 and 6.2 show the $m_{\gamma\gamma}$ distribution for each category in the Leptonic and Hadronic regions, respectively. Figure 6.3 shows the $m_{\gamma\gamma}$ distribution for all events passing the selection. The "Higgs-like" peak is seen around $m_{\gamma\gamma} = 125$ GeV although the statistics is not enough. We introduce the statistical procedure to evaluate the significance of the $ttH(H \to \gamma\gamma)$ signal as described in the next section.

6.2 Yield estimation

This section describes how to estimate the number of signals by the unbinned maximum likelihood fit to the $m_{\gamma\gamma}$ distribution. Section 6.2.1 explains the detail of the unbinned maximum likelihood fit. The $m_{\gamma\gamma}$ shape is modeled by analytic functions. The function for $H \to \gamma\gamma$ peak is discussed in Section 6.2.2, and the one for the continuum background is described in Section 6.2.3. The expectation of the signal yield and significance are discussed in Section 6.2.4.

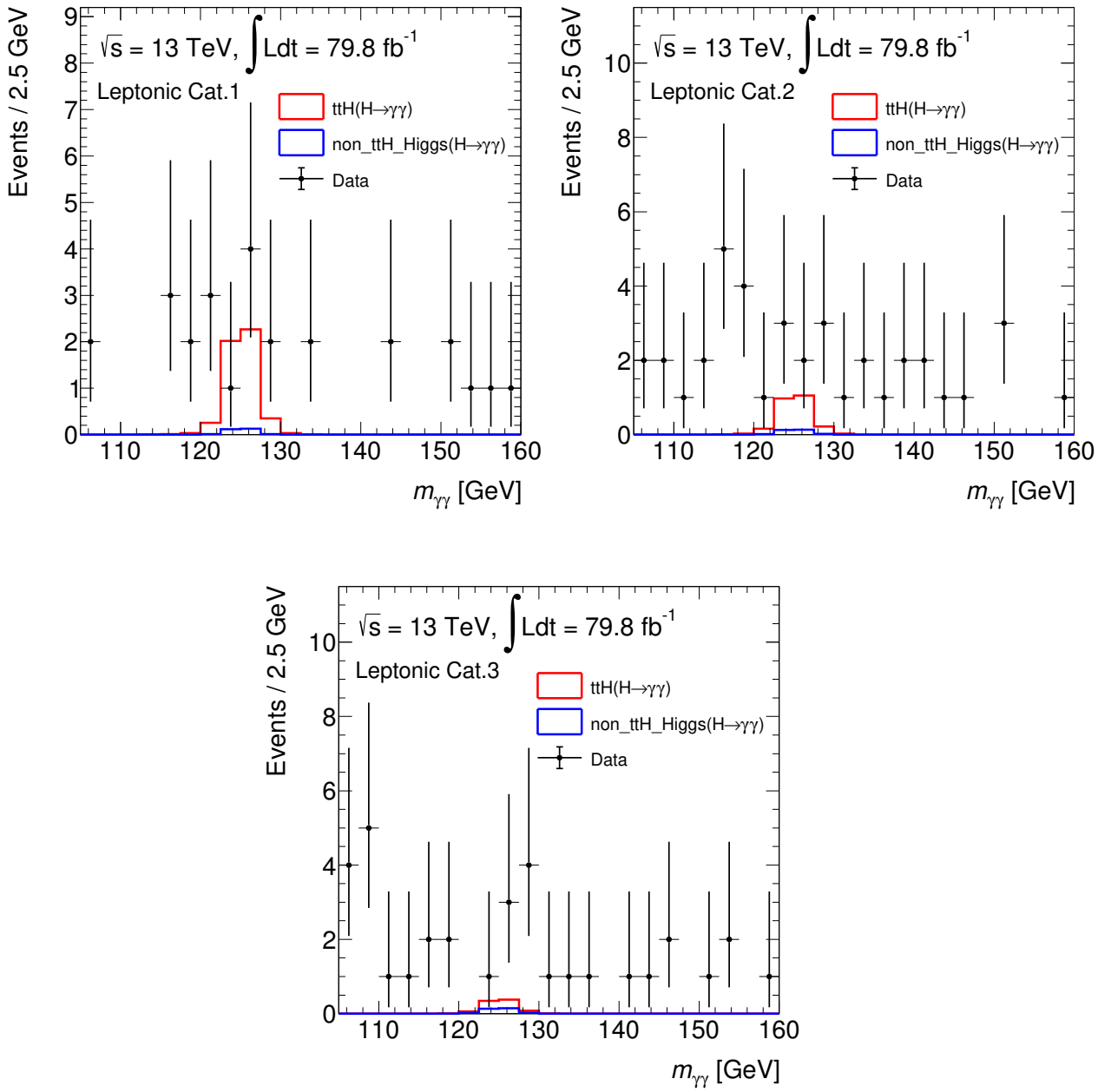


Figure 6.1: Di-photon invariant mass distribution of events for each category in the Leptonic region.

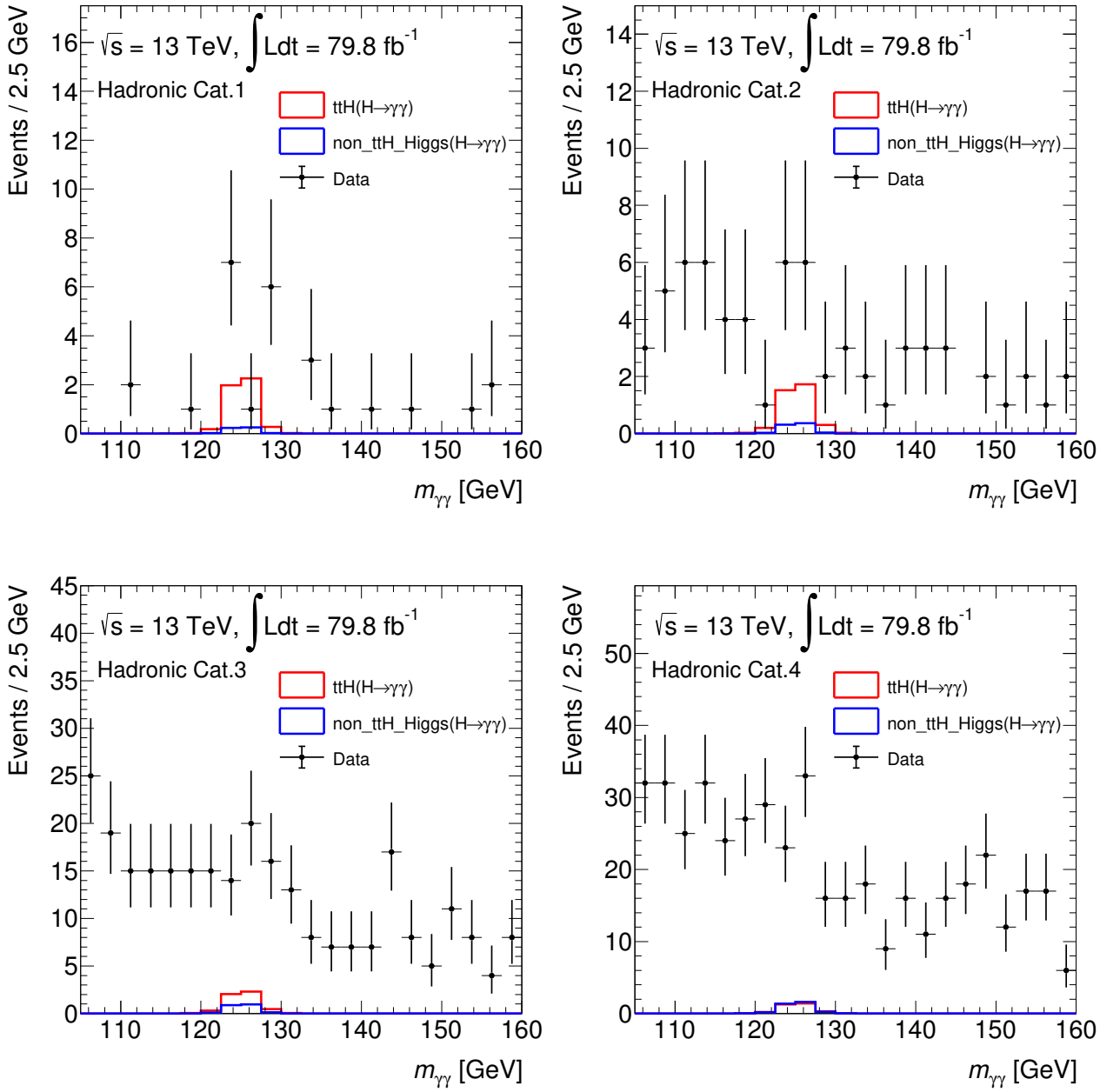


Figure 6.2: Di-photon invariant mass distribution of events for each category in the Hadronic region.

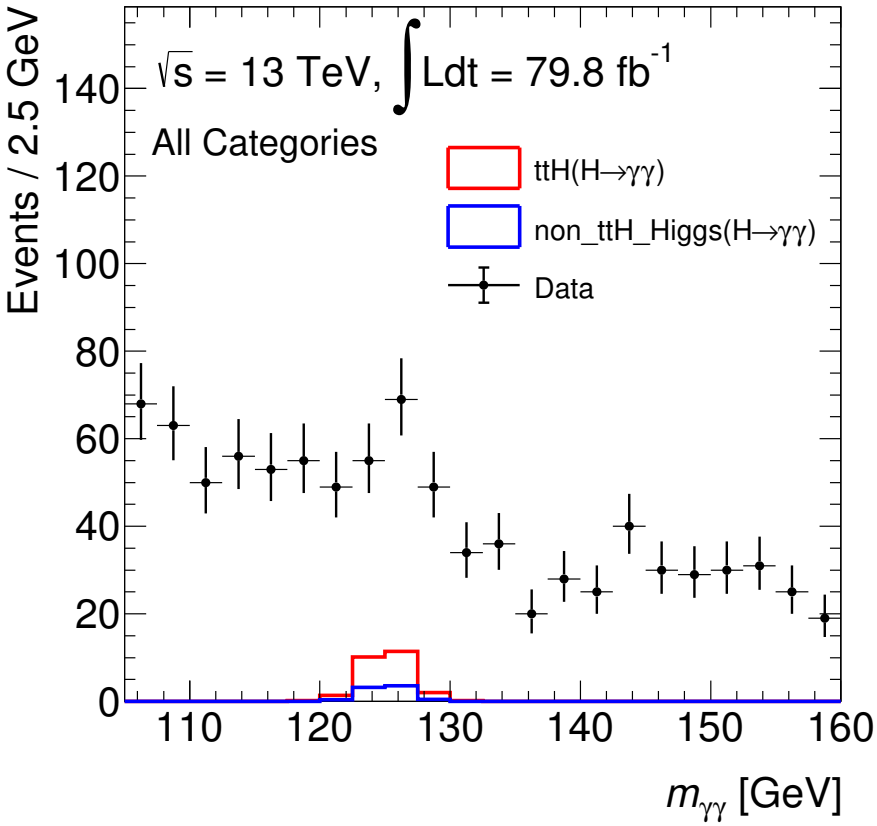


Figure 6.3: Di-photon invariant mass distribution of all events.

6.2.1 Statistical model

The likelihood function, L [75], is defined as

$$L = \prod_{c}^{N_{cat}} \left[\frac{(N_c(\boldsymbol{\theta}))^{n_c}}{n_c!} e^{-N_c(\boldsymbol{\theta})} \cdot \prod_{i}^{n_c} f_c(m^i_{\gamma\gamma}, \boldsymbol{\theta}) \right] \cdot L_{\boldsymbol{\theta}} \ , \tag{6.1}$$

where N_{cat} is the number of categories ($N_{cat} = 7$ in this analysis), n_c is the number of observed events in c-th category, $\boldsymbol{\theta}$ are nuisance parameters, $f_c(m^i_{\gamma\gamma})$ is the probability density function (PDF) of i-th event in terms of the $m_{\gamma\gamma}$ distribution, and $L_{\boldsymbol{\theta}}$ is the factor of Gaussian constraints by certain nuisance parameters. N_c is defined as

$$N_c(\boldsymbol{\theta}) = \sum_{p}^{prod} \mu_p s_{p,c}(\boldsymbol{\theta}) + b_c(\boldsymbol{\theta}) \ , \tag{6.2}$$

where μ_p and $s_{p,c}$ are the signal strength and the number of expected events of each Higgs production mode ($p \in ttH, ggF, VBF, ...$), respectively, and b_c is the number of continuum backgrounds. The PDF for c-th category can be written as

$$f_c\left(m^i_{\gamma\gamma}, \boldsymbol{\theta}\right) = \left[\sum_{p}^{prod} \mu_p s_{p,c}\left(\boldsymbol{\theta}\right) f_c^{H\to\gamma\gamma}\left(m^i_{\gamma\gamma}, \boldsymbol{\theta}\right) + b_c\left(\boldsymbol{\theta}\right) f_c^{cont}\left(m^i_{\gamma\gamma}, \boldsymbol{\theta}\right) \right] / N_c\left(\boldsymbol{\theta}\right), \tag{6.3}$$

where $f_c^{H\to\gamma\gamma}$ and f_c^{cont} are the PDF's of $H \to \gamma\gamma$ peak and continuum background, respectively. Note that the PDF's of signal and resonant backgrounds are considered to be the same, thus the resonant PDF ($f_c^{H\to\gamma\gamma}$) is common for all Higgs production modes.

The floated and fixed parameters of the fitting functions are summarized in Table 6.1. Since the parameter of interests is μ_{ttH}, the signal strengths for other Higgs production processes are fixed to $\mu_p = 1$ while μ_{ttH} is floated. The parameters related to $f_c^{H\to\gamma\gamma}$ are fixed while the parameters related to f_c^{cont} are floated. The PDF's $f_c^{H\to\gamma\gamma}$ and f_c^{cont} have various parameters which are discussed in Section 6.2.2 and 6.2.3, respectively.

Table 6.1: The parameters of the fitting functions. Parameters with the check mark (✓) denote floated while the ones with the minus mark (-) denotes fixed.

Normalization parameters				Shape parameters		
μ_p		s_p	b	$f_c^{H\to\gamma\gamma}(m_{\gamma\gamma})$		$f_c^{cont}(m_{\gamma\gamma})$
$p = ttH$	$p \neq ttH$			mean, sigma	others	
✓	-	-	✓	-	-	✓

The signal strength and its uncertainty are determined by minimizing the following test statistics q_μ;

$$q_\mu \equiv -2\ln\lambda(\mu_{ttH}) \equiv -2\ln\frac{L(\mu_{ttH}, \hat{\hat{\boldsymbol{\theta}}}_\mu)}{L(\hat{\mu}_{ttH}, \hat{\boldsymbol{\theta}})}, \tag{6.4}$$

where $\hat{\mu}_{ttH}$ and $\hat{\boldsymbol{\theta}}$ are the values of the signal strength and nuisance parameters that maximize the likelihood, while $\hat{\boldsymbol{\theta}}_{\mu}$ are the nuisance parameters that maximize the likelihood on the condition that μ_{ttH} is fixed to a given value.

The significance for discovery Z_0 is defined to be the square root of the test statistics with respect to background-only hypothesis as

$$Z_0 = \sqrt{q_0} = \sqrt{-2\ln\frac{L(0)}{L(\hat{\mu}_{ttH})}} \quad . \tag{6.5}$$

6.2.2 Modeling of the $H \to \gamma\gamma$ peak

The PDF of the $m_{\gamma\gamma}$ distribution for $H \to \gamma\gamma$ events, $f^{H\to\gamma\gamma}(m_{\gamma\gamma})$, is discussed in this section. The ideal shape of the $H \to \gamma\gamma$ mass peak is a Breit-Wigner curve, which peaks at the Higgs mass $m_H = 125$ GeV and has a narrow width of 4 MeV. However, the distribution observed in data is smeared by the finite detector resolution which is much larger than 4 MeV. Thus, the $m_{\gamma\gamma}$ distribution is modeled by a Gaussian-based function. The simple Gaussian is insufficient to model the tail due to the detector response. Furthermore the low mass tail is expected to be wider than high mass tail because it is possible that one of the electrons converted from photons is mis-detected. For these reasons, the $m_{\gamma\gamma}$ peak is modeled by a double-sided Crystal Ball function which is a combination of a Gaussian and power law tails defined as:

$$f^{H\to\gamma\gamma}(m_{\gamma\gamma}) = N \times \begin{cases} e^{-\frac{t^2}{2}} & (-\alpha_L \le t \le \alpha_H), \\ e^{-\frac{\alpha_L^2}{2}}\left[\frac{\alpha_L}{n_L}\left(\frac{n_L}{\alpha_L} - \alpha_L - t\right)\right]^{-n_L} & (t < -\alpha_L), \\ e^{-\frac{\alpha_H^2}{2}}\left[\frac{\alpha_H}{n_H}\left(\frac{n_H}{\alpha_H} - \alpha_H + t\right)\right]^{-n_H} & (t > \alpha_H), \end{cases} \tag{6.6}$$

$$t = \frac{\mu_{DSCB} - m_{\gamma\gamma}}{\sigma_{DSCB}}. \tag{6.7}$$

Here μ_{DSCB} and σ_{DSCB} are the mean and width of the Gaussian, respectively. α_L and α_H are the positions of the transition from the Gaussian to the power low tails on the low and high mass sides, respectively. n_L and n_H are the exponents of the low and high mass tails. N is the normalization factor. All parameters, except for N, are determined by the inclusive $H \to \gamma\gamma$ MC samples. Categories with high significance tend to have higher photon p_T and small $|\eta|$, resulting in better photon energy resolution, and hence better di-photon mass resolution. For this reason, the peak shape is decided category by category.

Figure 6.4 and 6.5 show the di-photon invariant mass distributions of MC for all Higgs productions including $ttH(H \to \gamma\gamma)$, for three categories in the Leptonic region and four categories in the Hadronic region. The parameters of $f^{H\to\gamma\gamma}(m_{\gamma\gamma})$ are retrieved by the unbinned maximum likelihood fit to these distributions.

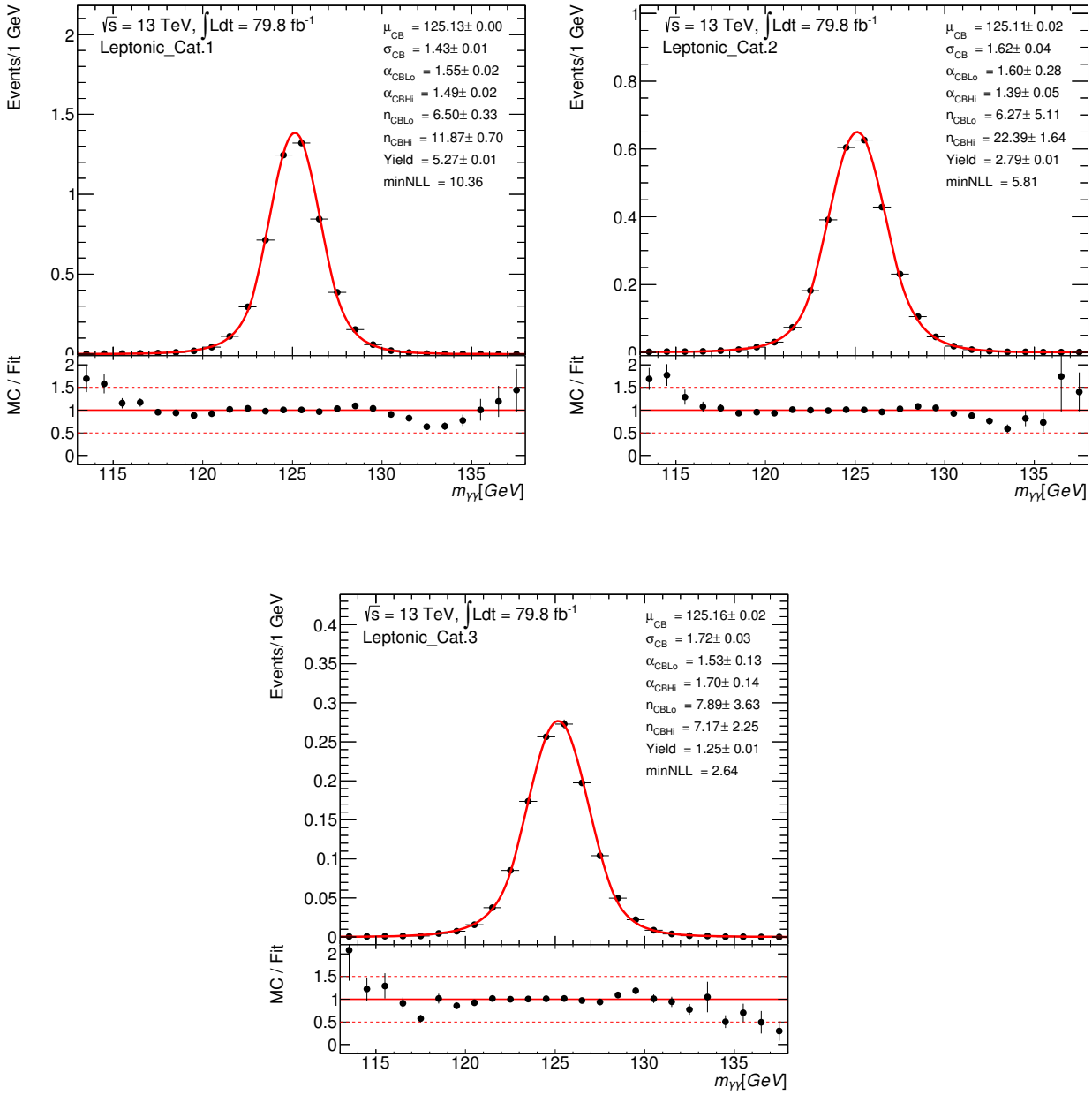


Figure 6.4: Di-photon invariant mass distribution fitted by the double-sided Crystal Ball function for three categories in the Leptonic region.

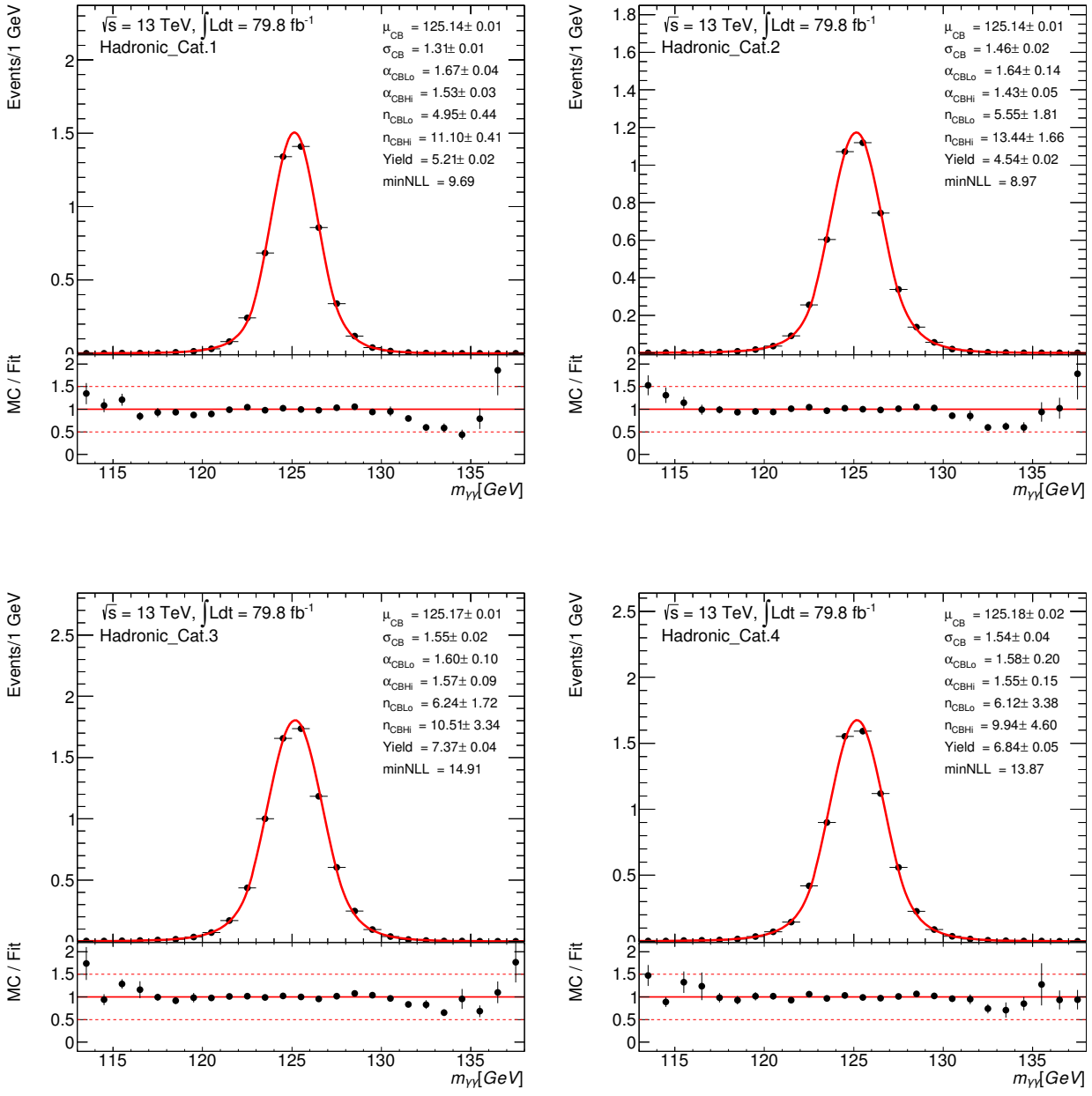


Figure 6.5: Di-photon invariant mass distribution fitted by the double-sided Crystal Ball function for four categories in the Hadronic region.

6.2.3 Modeling of the continuum background

The PDF of the $m_{\gamma\gamma}$ distribution for the continuum background, $f^{cont}(m_{\gamma\gamma})$, is discussed in this section. The sources of the continuum background are mainly $t\bar{t}+\gamma\gamma$ process in the Leptonic region, and $\gamma\gamma$ and $t\bar{t}+\gamma\gamma$ processes in the Hadronic region. The shape of these background distributions is smoothly falling like an exponential. Thus the continuum background is modeled by an exponential function.

In order to model the background shape, it is appropriate to use the sideband in data. However, because of lack of the statistics in the sideband, the templates made from $t\bar{t}+\gamma\gamma$ MC for the Leptonic region, and from data with NTI selection (defined in Section 5.4.3) for the Hadronic region are used. Figure 6.6 shows the comparison of the $m_{\gamma\gamma}$ distributions between data and the template. Data in the region of $121 < m_{\gamma\gamma} < 129$ GeV is removed to avoid the contamination of $H \to \gamma\gamma$ events. The templates are normalized so that the number of events in the sideband is the same as data. The detailed study about the comparison of the shape of the sideband and template is discussed in Appendix C.

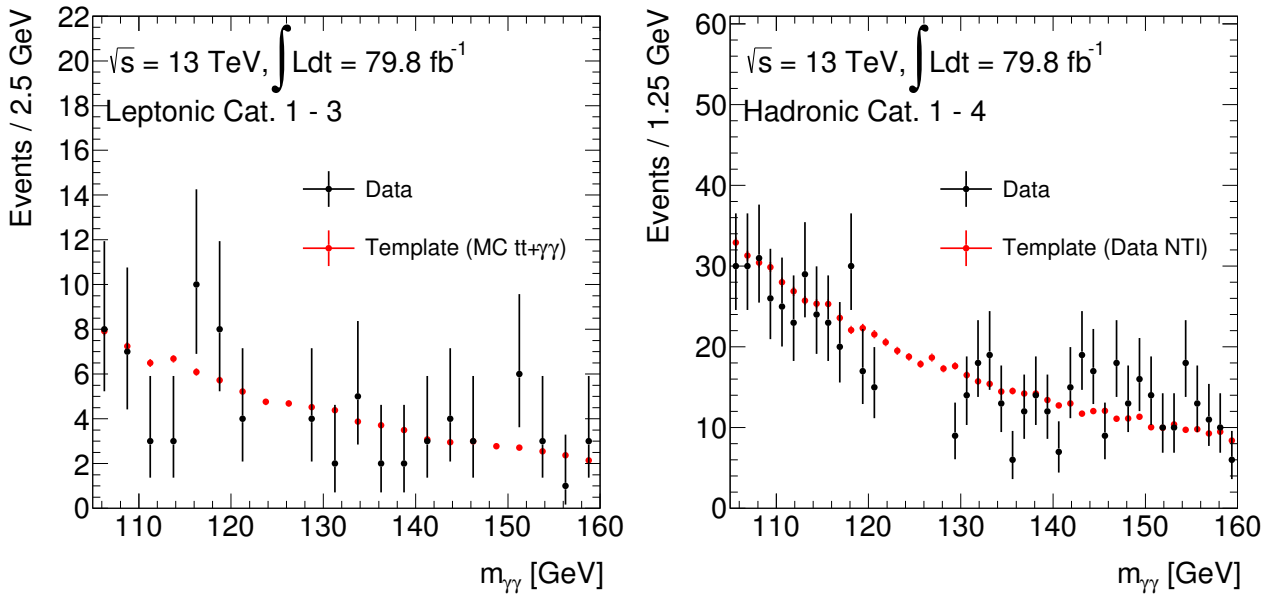


Figure 6.6: Template distribution used for the continuum background modeling for the Leptonic (left) and Hadronic (right) regions. Data in the region $121 < m_{\gamma\gamma} < 129$ GeV are removed to avoid the contamination of the $H \to \gamma\gamma$ events.

Figure 6.7 and 6.8 show the template distributions fitted by the exponential function for each category. The χ^2's divided by the number of degrees of freedom (NDF) are close to one for all categories, indicating the validity of the modeling.

To evaluate the systematic uncertainty due to the background modeling, alternative candidate functions are prepared. The difference of the number of backgrounds between the

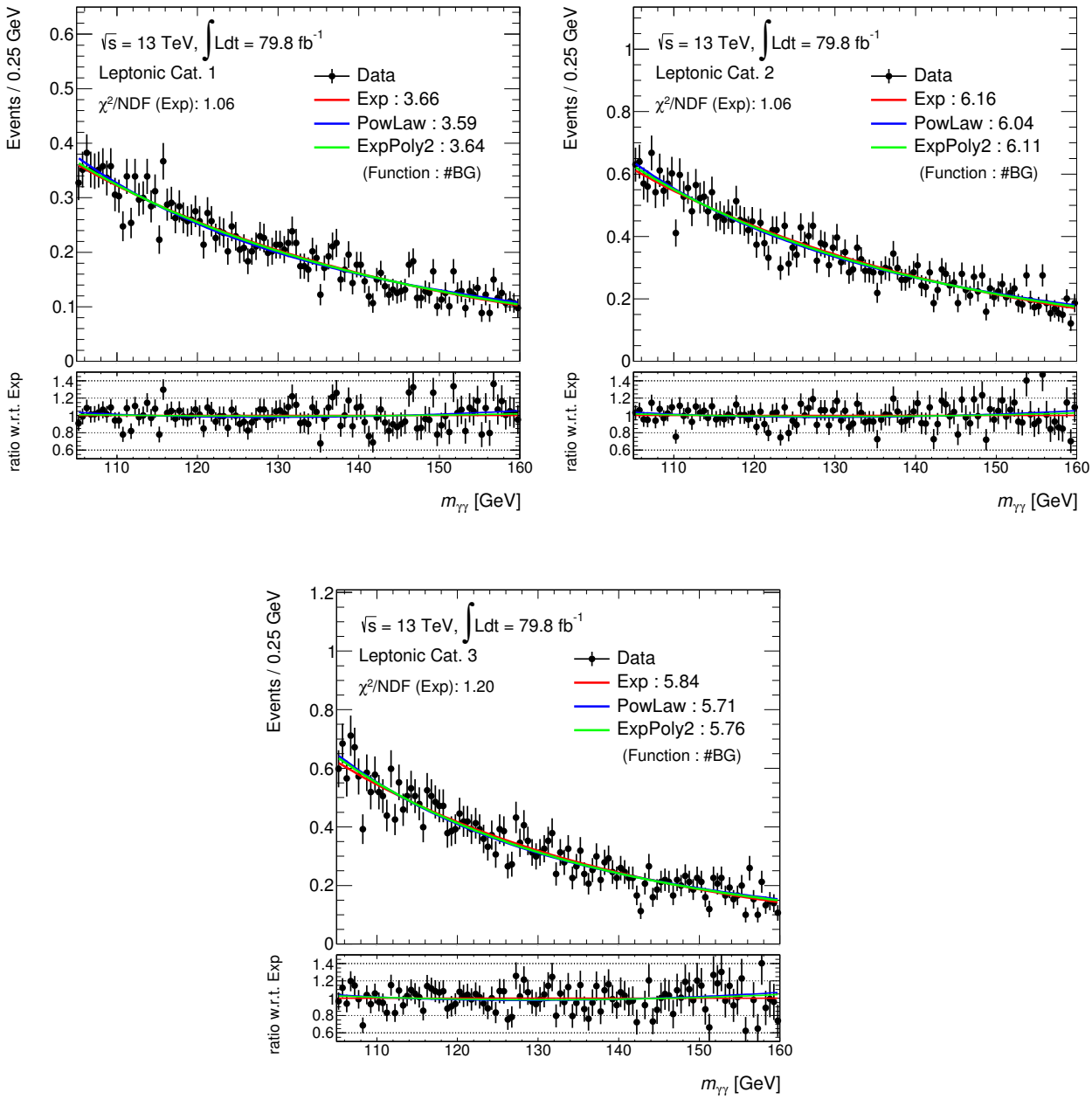


Figure 6.7: Template distribution for continuum background for each category in the Leptonic region. Templates are fitted by the Exponential (Exp), Power law (PowLaw) and Exponential of second-order Polynomial (ExpPoly2) functions. The χ^2 divided by the number of degrees of freedom (NDF) for the exponential fitting is shown. The number of backgrounds is calculated by integral in $121 < m_{\gamma\gamma} < 129$ GeV.

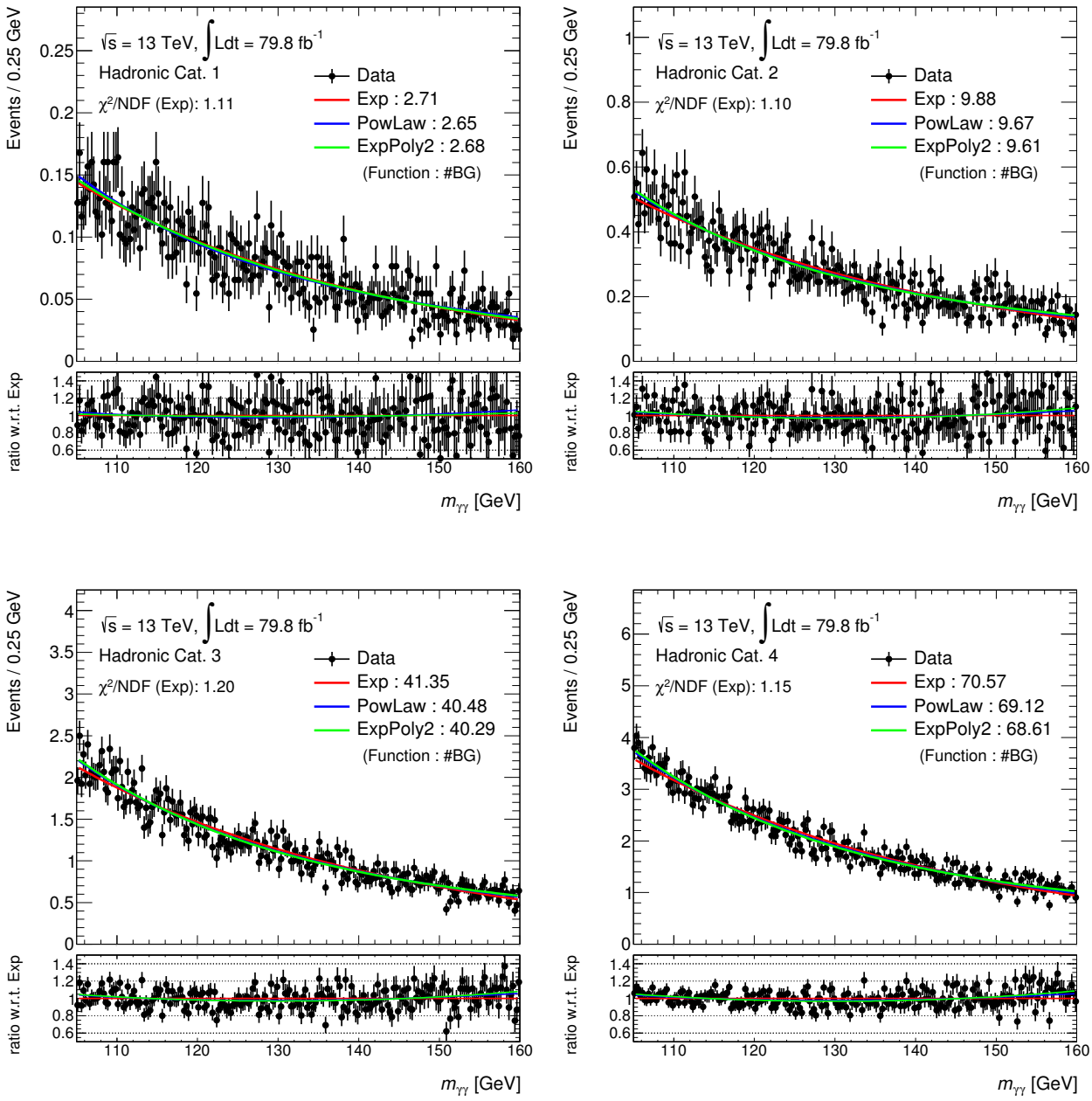


Figure 6.8: Template distribution for continuum background for each category in the Hadronic region. Templates are fitted by the Exponential (Exp), Power law (PowLaw) and Exponential of second-order Polynomial (ExpPoly2) functions. The χ^2 divided by the number of degrees of freedom (NDF) for the exponential fitting is shown. The number of backgrounds is calculated by integral in $121 < m_{\gamma\gamma} < 129$ GeV.

different functions is assigned as the systematic uncertainty. The following functions are used.

- Exponential: $f^{cont}(m_{\gamma\gamma}) = e^{c_1 m_{\gamma\gamma}}$
- Power Law: $f^{cont}(m_{\gamma\gamma}) = (m_{\gamma\gamma})^{c_1}$
- Exponential of second-order Polynomial: $f^{cont}(m_{\gamma\gamma}) = e^{c_1 m_{\gamma\gamma}^2 + c_2 m_{\gamma\gamma}}$

The fitting results of the remaining two functions are also shown in Figure 6.7 and 6.8. The number of backgrounds shown in the plots is calculated by the integral in $121 < m_{\gamma\gamma} < 129$ GeV. The uncertainty is relatively small compared to the number of expected signals for each category.

6.2.4 Expectation

Table 6.2 shows the summary of the number of expected signals, non-ttH Higgs backgrounds and continuum backgrounds in $121 < m_{\gamma\gamma} < 129$ GeV for each category. The number of expected signals and non-ttH Higgs backgrounds are estimated using MC, while the number of continuum backgrounds is estimated by the integral of the function determined in Section 6.2.3.

Table 6.2: The number of expected signals and backgrounds in $121 < m_{\gamma\gamma} < 129$ GeV for each category in the Leptonic and Hadronic regions.

Category	Signal	non-ttH Higgs BG	continuum BG
Leptonic Cat. 1	4.70 ± 0.79	0.25 ± 0.08	3.66 ± 0.07
Leptonic Cat. 2	2.30 ± 0.31	0.28 ± 0.11	6.16 ± 0.12
Leptonic Cat. 3	0.82 ± 0.14	0.31 ± 0.15	5.84 ± 0.13
Hadronic Cat. 1	4.54 ± 0.86	0.50 ± 0.33	2.71 ± 0.06
Hadronic Cat. 2	3.58 ± 0.57	0.70 ± 0.48	9.88 ± 0.29
Hadronic Cat. 3	4.91 ± 0.65	1.94 ± 1.63	41.4 ± 1.06
Hadronic Cat. 4	3.05 ± 0.39	3.18 ± 2.81	70.6 ± 1.96

The expected significance Z_0 is estimated by fitting $m_{\gamma\gamma}$ of the ensemble of the simulated dataset, called Asimov dataset [75]. The PDF's of the signal and resonant background are obtained from MC while that of the continuum background are obtained from sideband data. The Asimov dataset is generated separately for each category, obeying the PDF's. Figure 6.9 and 6.10 show the $m_{\gamma\gamma}$ distribution of the Asimov dataset for each category in the Leptonic and Hadronic regions, respectively, overlaid with the fitting results. The expected significance only considering the statistical uncertainty is 4.0 σ. The impact of the systematic uncertainty on the significance is evaluated in the next section.

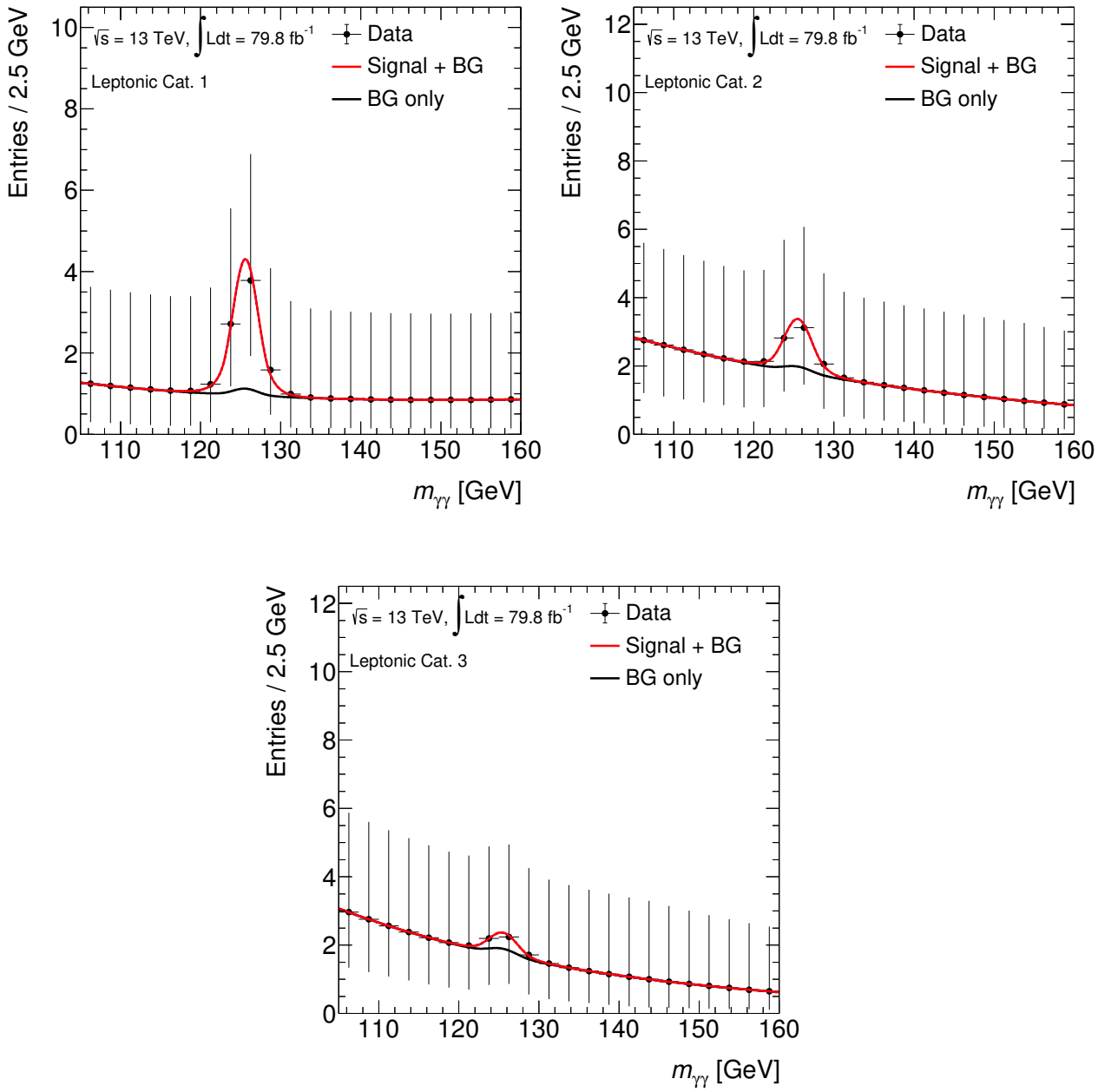


Figure 6.9: Di-photon invariant mass spectrum of the Asimov dataset for the three Leptonic categories. The black line is the background-only spectrum and the red line is the signal + background spectrum from the SM prediction.

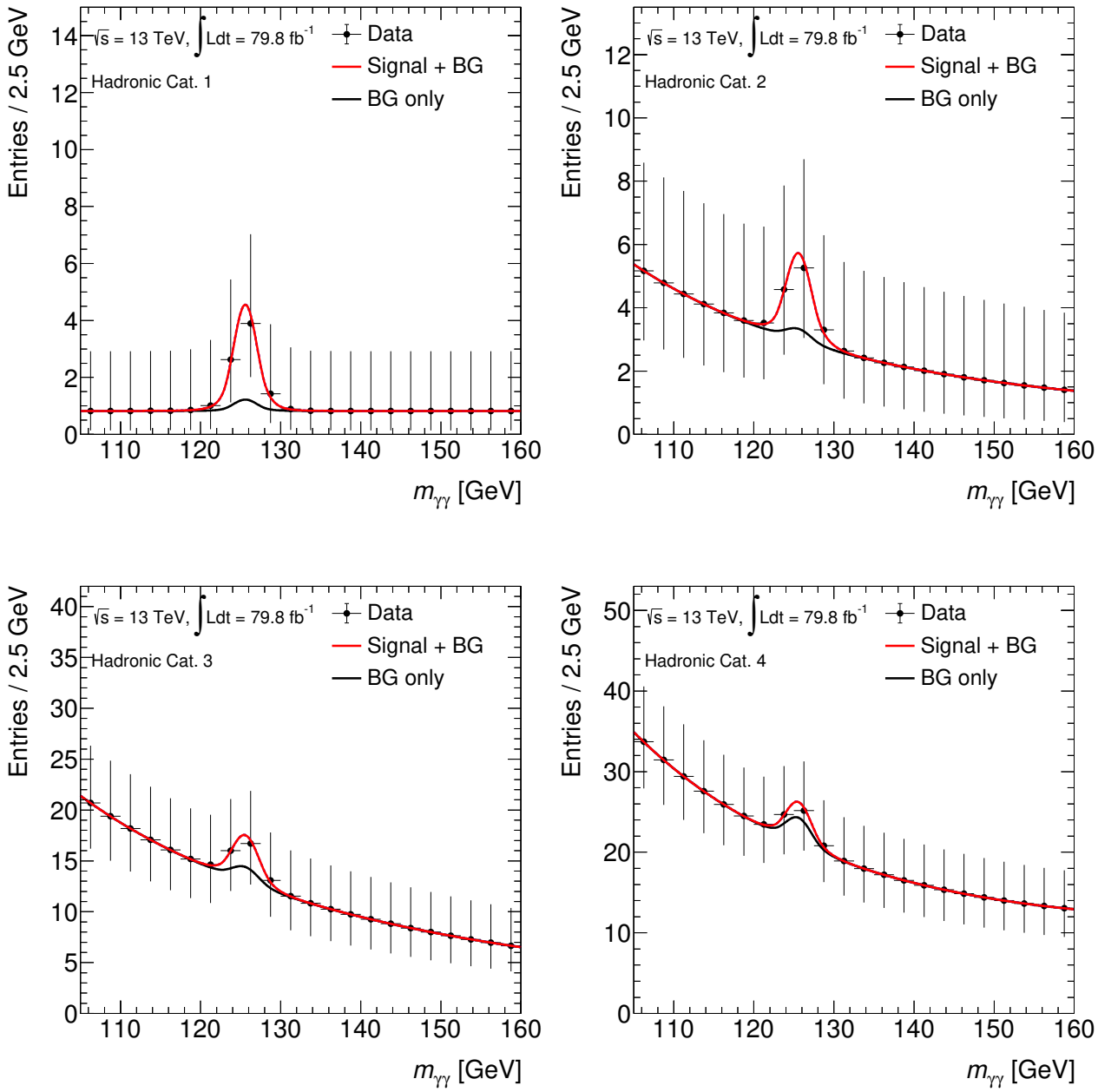


Figure 6.10: Di-photon invariant mass spectrum of the Asimov dataset for the four Hadronic categories. The black line is the background-only spectrum and the red line is the signal + background spectrum from the SM prediction.

6.3 Systematic uncertainty

The sources of the systematic uncertainty can be classified into three types;

- Uncertainty related to the yield.
- Uncertainty related to the di-photon mass scale.
- Uncertainty related to the di-photon mass resolution.

Figure 6.11 shows the illustration of these three types of uncertainties. Table 6.3 lists all the systematic uncertainties considered in this analysis. The check mark (✓) indicates if each source affects the signal, non-ttH Higgs background or continuum background.

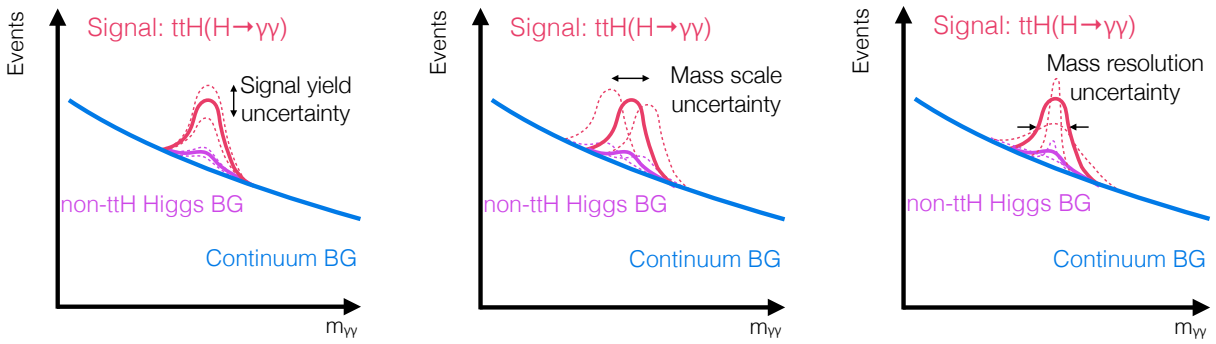


Figure 6.11: Illustration of the systematic uncertainties affecting to the mass fit: (a) the uncertainty related to the yield, (b) the uncertainty related to the di-photon mass scale and (c) the uncertainty related to the di-photon mass resolution.

In order to estimate the systematic uncertainties, the event selection and categorization are repeated using MC samples, varying each uncertainty source by $\pm$ 1 σ. The size of each uncertainty is evaluated by taking a ratio of the result obtained with the modified MC to the nominal result. The uncertainties are evaluated category by category. The detail of each uncertainty source is described in the following.

6.3.1 Uncertainty related to the yield

The number of observed events for i-th category N_i is defined as

$$N_i = N_{\text{cont.BG}} + N_{H\to\gamma\gamma} \tag{6.8}$$

$$= N_{\text{cont.BG}} + L \times BR(H \to \gamma\gamma) \times \sum_{p}^{prod} \left[\sigma_{prod} \times \varepsilon^{acc}_{prod,i}\right], \tag{6.9}$$

where $N_{\text{cont.BG}}$ is the number of continuum backgrounds, $N_{H\to\gamma\gamma}$ is the number of $H \to \gamma\gamma$ events for all Higgs production, L is the integrated luminosity, σ_{prod} is the cross section for

Table 6.3: List of all systematic uncertainties. The check mark (✓) represents which uncertainty affects the signal, non-ttH Higgs background or continuum background.

		Signal	Non-ttH BG	Cont. BG
Yield	Luminosity	✓	✓	-
	$BR(H \to \gamma\gamma)$	✓	✓	-
	Cross section			
	μ_r and μ_f in the QCD	-	✓	-
	PDF + α_S	-	✓	-
	Acceptance			
	μ_r and μ_f in the QCD	✓	✓	-
	PDF + α_S	✓	✓	-
	MC generator modeling	✓	-	-
	Photon trigger	✓	✓	-
	Photon $Tight$ selection	✓	✓	-
	Photon isolation selection	✓	✓	-
	Jet	✓	✓	-
	Lepton	✓	✓	-
	b-tagging	✓	✓	-
	Missing transverse energy	✓	✓	-
	Heavy flavor jet modeling	-	✓ (ggF, VBF, VH)	-
	Background modeling	-	-	✓
Mass scale	Photon energy scale	✓	✓	-
	Higgs mass	✓	✓	-
Mass resolution	Photon energy resolution	✓	✓	-

each Higgs production process, the $BR(H \to \gamma\gamma)$ is the branching ratio of the $H \to \gamma\gamma$ decay and $\varepsilon^{acc}_{prod,i}$ is the acceptance for the i-th category estimated by MC for each Higgs production.

Since $N_{\text{cont.BG}}$ is obtained from data, its systematic uncertainty is not considered except for the background modeling as discussed in Section 6.2.3. Table 6.4 shows the magnitude of the background modeling uncertainty for each category and the impact on the ttH cross section.

Table 6.4: Magnitude of the background modeling uncertainty and the impact on the ttH cross section.

Source of uncertainty	$\Delta N_{\text{cont.BG}}/N_{\text{cont.BG}}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat.1	Cat.2	Cat.3	Cat.1	Cat.2	Cat.3	Cat.4	
Background modeling	1.9	2.0	2.2	2.2	3.0	2.6	2.8	1.7

The luminosity is measured using the forward detectors, LUCID and BCM as mentioned in Section 2.6. The measured integrated luminosity in the period of 2015-2017 satisfying the "Good for physics" condition is 79.8 $\pm$ 1.6 fb^{-1}, and consequently the uncertainty due to the luminosity measurement is $\pm$ 2.0 %. The branching ratio of the $H \to \gamma\gamma$ decay is given to be $2.27\ ^{+0.07}_{-0.06} \times 10^{-3}$ from the theoretical prediction assuming the SM [19]. Table 6.5 shows their impact on the ttH cross section.

Table 6.5: Magnitude of the he integrated luminosity and $BR(H \to \gamma\gamma)$ uncertainties and their impact on the ttH cross section.

Source of uncertainty	$\Delta N_{H\to\gamma\gamma}/N_{H\to\gamma\gamma}$ [%]	$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
Luminosity	2.0	2.3
$BR(H \to \gamma\gamma)$	2.9	3.2

To measure the ttH cross section, the non-ttH Higgs production processes are assumed as the SM. In the calculation of the predicted cross section, we introduced two artificial parameters in the perturbative expansion of the QCD process; the renormalization scale μ_r and the factorization scale μ_f. The renormalization scale μ_r is introduced for the fixed order QCD calculation. The factorization scale μ_f is applied to separate the calculation of the cross section into so called long- and short-distance parts. The long-distance physics is included in the parton distribution functions (PDF's) while the short-distance physics is included in the hard-scattering cross section. The cross section is calculated at next-to-next-to-next-to-leading order (N3LO), next-to-next-to-leading order (NNLO) or next-to-leading order (NLO) QCD, depending on the production processes. Due to the missing higher-order terms of the QCD calculation, the cross section depends on μ_r and μ_f. In this analysis, both are fixed to be $\mu_r = \mu_f = M_t + M_H/2$, where M_t and M_H are the top-quark and Higgs boson masses, respectively. The uncertainty of the missing higher-order terms in the QCD calculation is obtained by varying μ_r and μ_f by factors of 0.5 and 2.0. Table 6.6 lists the magnitude of the cross section uncertainties by the renormalization and factorization scales in the QCD calculation for each Higgs production process [19].

The cross section also depends on the PDF for incoming protons. Since the PDF cannot be calculated by the perturbative QCD, it is empirically determined using data. The uncertainty from the variation of PDF's is estimated by the NLO PDF sets of NNPDF3.0 [43]. The uncertainty of the strong coupling constant α_S is added in quadrature. Table 6.6 also lists the magnitude of the cross section uncertainties by the PDF+α_S for each Higgs production process [19].

Table 6.6: Magnitude of the cross section uncertainty for each Higgs production process [19] and their impact on the ttH cross section.

Process	$\Delta\sigma_{prod}/\sigma_{prod}$ [%]		$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	μ_r and μ_f in the QCD	PDF + α_S	
ggF (N3LO)	+3.9 / -3.9	+3.2 / -3.2	0.4
VBF (NNLO)	+0.4 / -0.3	+2.1 / -2.1	0.0
WH (NNLO)	+0.5 / -0.7	+1.9 / -1.9	0.0
ZH (NNLO)	+3.8 / -3.0	+1.6 / -1.6	0.1
$tHjb$ (NLO)	+6.5 / -14.7	+3.7 / -3.7	0.0
tWH (NLO)	+4.9 / -6.7	+6.3 / -6.3	0.0
ttH (NLO)	+5.8 / -9.2	+3.6 / -3.6	-

Uncertainty related to the signal acceptance

There are various sources of the acceptance uncertainty. Table 6.7 lists the sources of the signal acceptance uncertainty for each category. The detail of the main signal acceptance uncertainties is explained below. The acceptance uncertainties for non-ttH Higgs backgrounds are described in Appendix D.

Table 6.7: Magnitude of the acceptance uncertainties for $ttH(H \to \gamma\gamma)$ process evaluated for each category and their impact on the ttH cross section.

Source of uncertainty	$\Delta\epsilon_{ttH}/\epsilon_{ttH}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat.1	Cat.2	Cat.3	Cat.1	Cat.2	Cat.3	Cat.4	
MC generator modeling	9.1	4.5	7.3	12	7.1	3.9	3.0	7.9
μ_r and μ_f in the QCD	0.4	0.6	1.2	0.5	0.3	0.5	0.1	0.5
PDF+α_S	1.3	1.3	1.3	1.4	0.9	1.2	0.9	1.3
Photon trigger	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Photon $Tight$ selection	3.7	3.5	3.5	4.0	3.6	3.5	3.5	3.7
Photon isolation selection	4.5	4.3	4.2	4.7	4.5	4.3	4.3	5.1
Jet	5.1	2.8	10	7.9	5.9	2.9	2.3	5.7
Lepton	1.0	1.1	1.3	-	-	-	-	0.2
b-tagging	1.3	1.9	2.2	2.2	1.6	1.0	0.8	1.6
Missing transverse energy	0.2	0.3	0.3	0.5	0.3	0.2	0.5	0.4
Total	12	8.1	14	16	11	7.6	6.9	12

- Uncertainty due to μ_r and μ_f in the QCD calculation and the PDF+α_S

The uncertainty of the renormalization and factorization scales in the QCD calculation and the PDF+α_S affect not only to the cross section but also to the acceptance. Table 6.7 shows the size of the signal acceptance uncertainty caused by them. The impact on the signal acceptance is relatively small.

- Uncertainty related to the MC modeling

The signal MC modeling uncertainty arises from the generator, parton shower, hadronization, underlying event and PDF. To evaluate the size of the uncertainties, the alternative MC samples are used as discussed in Section 3.2.2. The generator uncertainty is estimated by comparing the nominal $\mathrm{P_{OWHEG}}$ + $\mathrm{P_{YTHIA}8}$ sample with the MadGraph5_aMC@NLO + $\mathrm{P_{YTHIA}8}$ sample, while the other uncertainties are inclusively estimated by comparing the MadGraph5_aMC@NLO + $\mathrm{P_{YTHIA}8}$ sample with the MadGraph5_aMC@NLO + Herwig++ sample. It is the largest source of the cross section uncertainty as shown in Table 6.7.

- Uncertainty of the photon *Tight* selection efficiency

The *Tight* photon selection efficiency is measured using the $Z \to ll\gamma$ events and the $Z \to ee$ events as discussed in Section 4.2.1 for converted and unconverted photons. It varies from 0.3 % to 12 % for converted photons and from 1.1 % to 13 % for unconverted photons. The dominant source of the uncertainty is the one associated with the disagreement of shower shape distributions between data and MC, which comes from the material mis-modeling.

- Uncertainty of the photon isolation selection efficiency

The photon isolation efficiency is measured using $Z \to ll\gamma$ events and inclusive photon events. The uncertainty varies from 1.9 % to 5.0 % in the range of $20 < p_T < 300$ GeV. The impact on the signal yield is from 4.2 % to 4.7 %, depending on the categories.

6.3.2 Uncertainty related to the di-photon mass scale

Because the position of the signal Higgs peak is fixed to the world average in the fitting, the di-photon mass scale contributes to the systematic uncertainty.

The uncertainty of the di-photon invariant mass scale comes from the uncertainty of the photon energy scale. In principle, measurement of the opening angle of two photons could cause the uncertainty. However, the fine lateral and longitudinal granularity of the EM calorimeter gives precise measurement of the position of photon resulting in negligible effect to the uncertainty.

The uncertainty also comes from the current world average of the Higgs mass, which is $m_H = 125.09 \pm 0.24$ GeV [78]. Thus, $\pm$ 0.2 % of the mass scale uncertainty is considered.

The mass scale uncertainty is evaluated by varying each uncertainty source from the fixed peak position (μ_{DSCB}). Table 6.8 lists the source of uncertainty and the magnitude of the mass scale uncertainty in each category.

Table 6.8: List of the sources of the di-photon mass scale uncertainties for each category.

Source of uncertainty	$\Delta\mu_{DSCB}/\mu_{DSCB}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat. 1	Cat. 2	Cat. 3	Cat. 1	Cat. 2	Cat. 3	Cat. 4	
Photon energy scale	0.4	0.3	0.3	0.4	0.4	0.3	0.3	1.5
Higgs mass	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.9

6.3.3 Uncertainty related to the di-photon mass resolution

Because again the width of the signal Higgs peak is fixed in the fitting, the mass resolution contributes to the systematic uncertainty. The mass resolution uncertainty is evaluated by varying each uncertainty source from the fixed peak width (σ_{DSCB}). Table 6.9 shows the magnitude of the mass resolution uncertainty in each category, which only attributes the photon energy resolution uncertainty, and the magnitude of the cross section uncertainty propagated by the mass resolution uncertainty.

Table 6.9: List of the sources of the di-photon mass resolution uncertainty for each category.

Source of uncertainty	$\Delta\sigma_{DSCB}/\sigma_{DSCB}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat. 1	Cat. 2	Cat. 3	Cat. 1	Cat. 2	Cat. 3	Cat. 4	
Photon energy resolution	12	8.4	9.3	15	13	11	9.5	5.1

6.3.4 Summary of the systematic uncertainty

Table 6.10 shows the main systematic uncertainties to the ttH cross section.

Table 6.10: Top ten sources of uncertainty obtained from the fit to the observed dataset.

Uncertainty	$-\Delta\sigma_{ttH}/\sigma_{ttH}$[%]	$+\Delta\sigma_{ttH}/\sigma_{ttH}$[%]
Parton shower / hadronization / underlying event / PDF	7.7	8.1
Jet energy scale / energy resolution / JVT	5.7	5.7
Photon isolation selection efficiency	5.0	5.2
Photon energy resolution	5.2	4.8
Photon $Tight$ selection efficiency	3.1	4.2
Modeling of heavy flavor jet in ggF process	3.6	3.8
$BR(H \to \gamma\gamma)$	3.2	3.3
Background modeling	2.4	2.3
Luminosity	2.2	2.3
B-tagging	2.2	2.2

6.4 Validation of the signal extraction method

Pseudo experiments are performed to confirm the linearity and the validity of the size of uncertainty of the fitting. The $m_{\gamma\gamma}$ distribution is prepared using toy MC, where the shape of the signal and the resonant background are obtained by MC while the continuum background by the sideband data. The number of signals is first assumed to be the SM ($\mu^{input} = 1$) when creating pseudo datasets. The number of events in each pseudo dataset is fluctuated to follow the Poisson statistics. The number of the pseudo datasets is 10,000. The fitting is performed to the 10,000 pseudo datasets.

Figure 6.12 shows the distribution of the signal strength μ^{output} and its pull defined as $(\mu^{output} - \mu^{input})/\delta\mu^{output}$, where μ^{output} and its error $\delta\sigma^{output}$ are obtained by the fitting. The mean and sigma of the distributions are obtained by the gaussian fitting. The mean of the μ^{output} distribution is consistent with the input value. The sigma of the pull distribution is also consistent with unity implying the validity of the size of statistical error.

In order to test the linearity, μ^{input} is varied from 0 to 2 with 0.5 step and repeated the pseudo experiments. Figure 6.13 shows μ^{output} and pull as a function of μ^{input}. The response of the fit shows no bias. The distributions of the μ^{output} and its pull for various μ^{input} are shown in Appendix E.

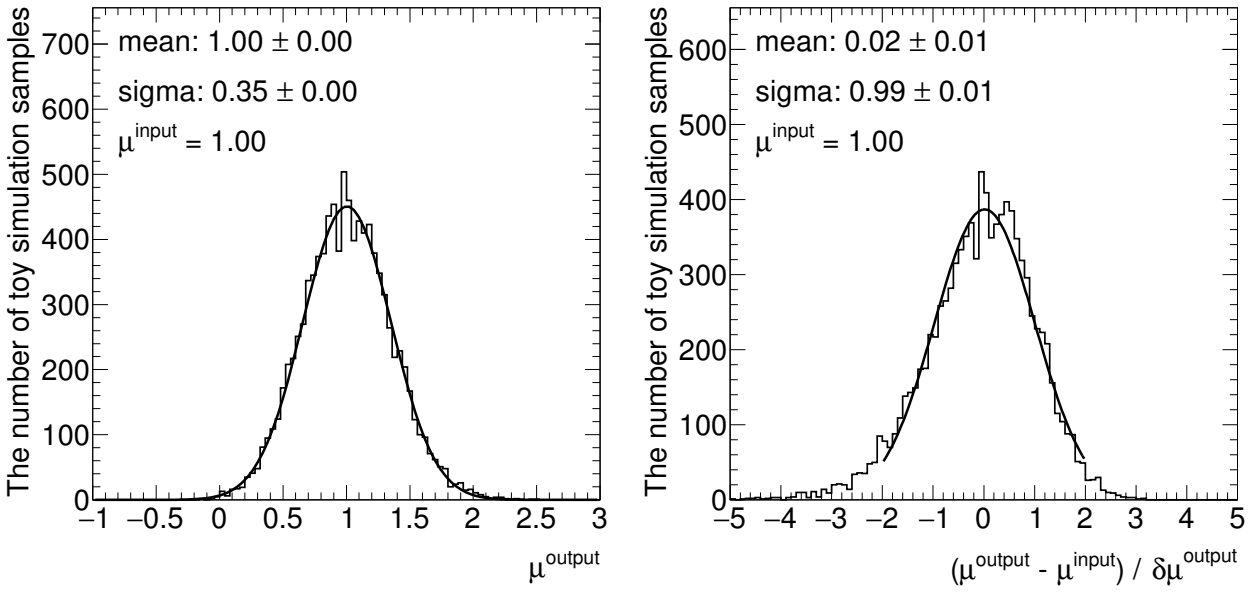


Figure 6.12: Distribution of the signal strength μ^{output} (left) and its pull (right) obtained by fitting pseudo datasets generated by toy MC. The input signal strength is assumed to be $\mu^{input} = 1$ in the generation of pseudo data.

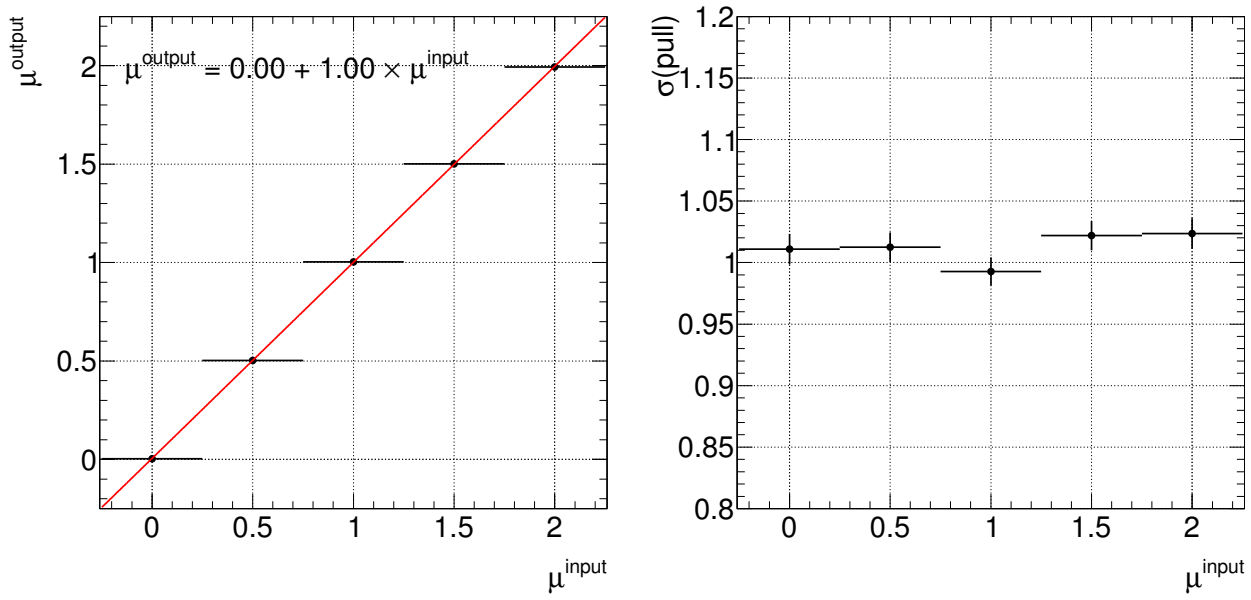


Figure 6.13: Response of the signals strength μ^{output} (left) and its pull (right) for various μ^{input}.

6.5 Result

The unbinned maximum likelihood fit is performed to all categories in both the Leptonic and Hadronic regions simultaneously. Figure 6.14 and 6.15 show the $m_{\gamma\gamma}$ distributions of the data with fitting functions for each category. The combined plot of all categories is shown in Figure 6.16. From the fitting, the statistical significance is determined to be 4.2 σ, while expected significance is 3.6 σ assuming the SM. Not only a visual inspection of the $H \to \gamma\gamma$ peak in $m_{\gamma\gamma}$ but also the statistical interpretation provides evidence for the ttH production using $H \to \gamma\gamma$ decay mode.

The numbers of signals and backgrounds are calculated by integrating the fitted results in the range of $121 < m_{\gamma\gamma} < 129$ GeV. Table 6.11 shows the numbers of signals S and backgrounds B, the signal to background ratio S/B, and the significance Z_0. The most sensitive categories in each Leptonic and Hadronic region (Leptonic Cat.1 and Hadronic Cat.1) dominate the significance, while the remaining categories contribute to 13 % of the total significance. The Hadronic region has slightly better sensitivity in total than the Leptonic region. In general, since at least one lepton (e, μ) is required to exist in the Leptonic region, it has better S/B than the Hadronic region. However, the S/B in the sensitive categories in the Hadronic region is comparable with the one in the Leptonic region. One of the reason is the tight requirement for the XGBoost score in the Hadronic region to reduce the number of backgrounds. In addition, the number of signals in the Hadronic region is roughly twice than that in the Leptonic region as shown in Table 5.1, because of the smaller branching ratio of W in the Leptonic region and some inefficiency for electron and muon. With these reasons,

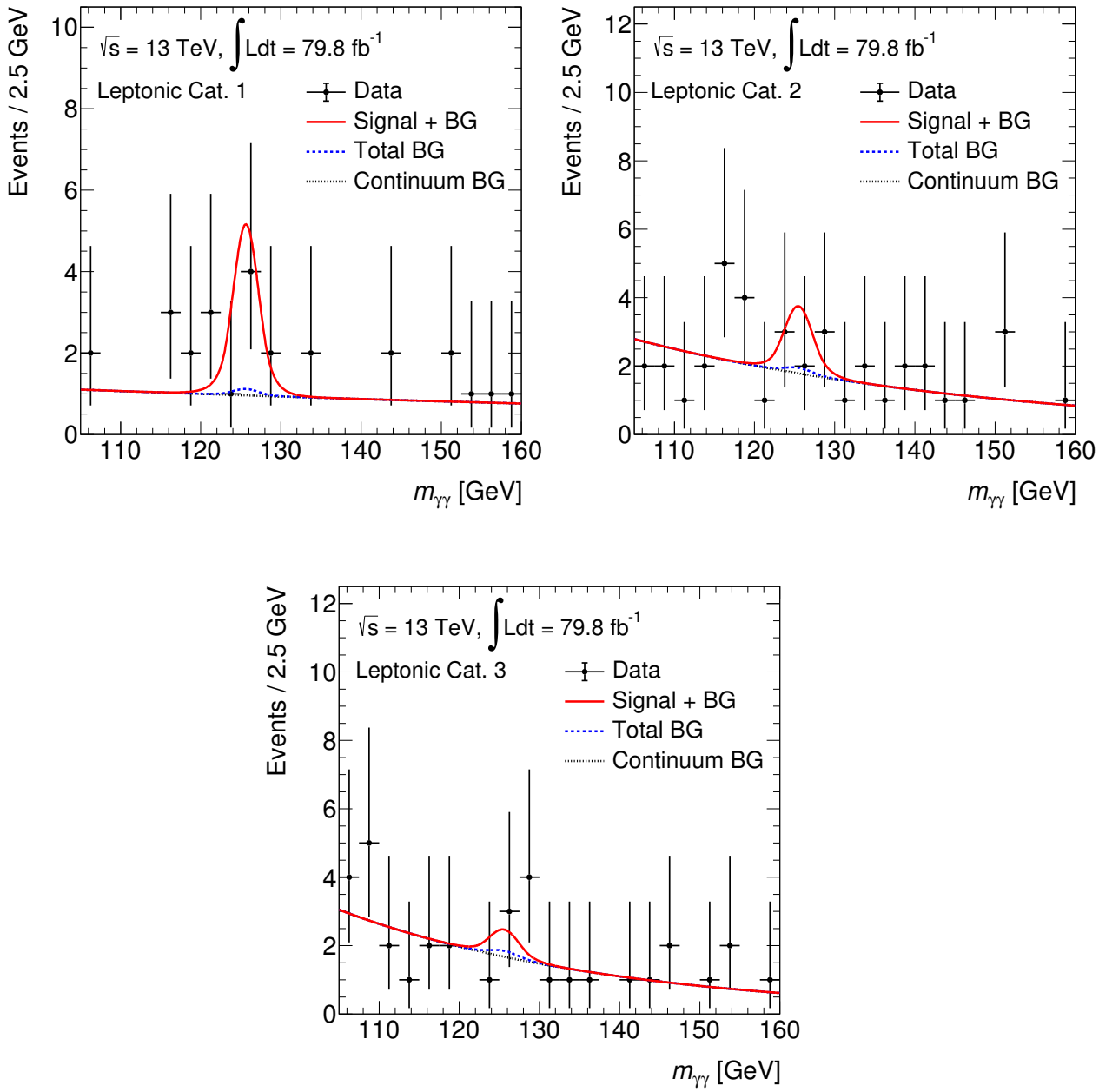


Figure 6.14: Di-photon invariant mass distribution of the observed events for three categories in the Leptonic region. The solid red curve shows the fitted signal + background model. The dashed blue and dotted black curves show the total and continuum background components of the fit, respectively.

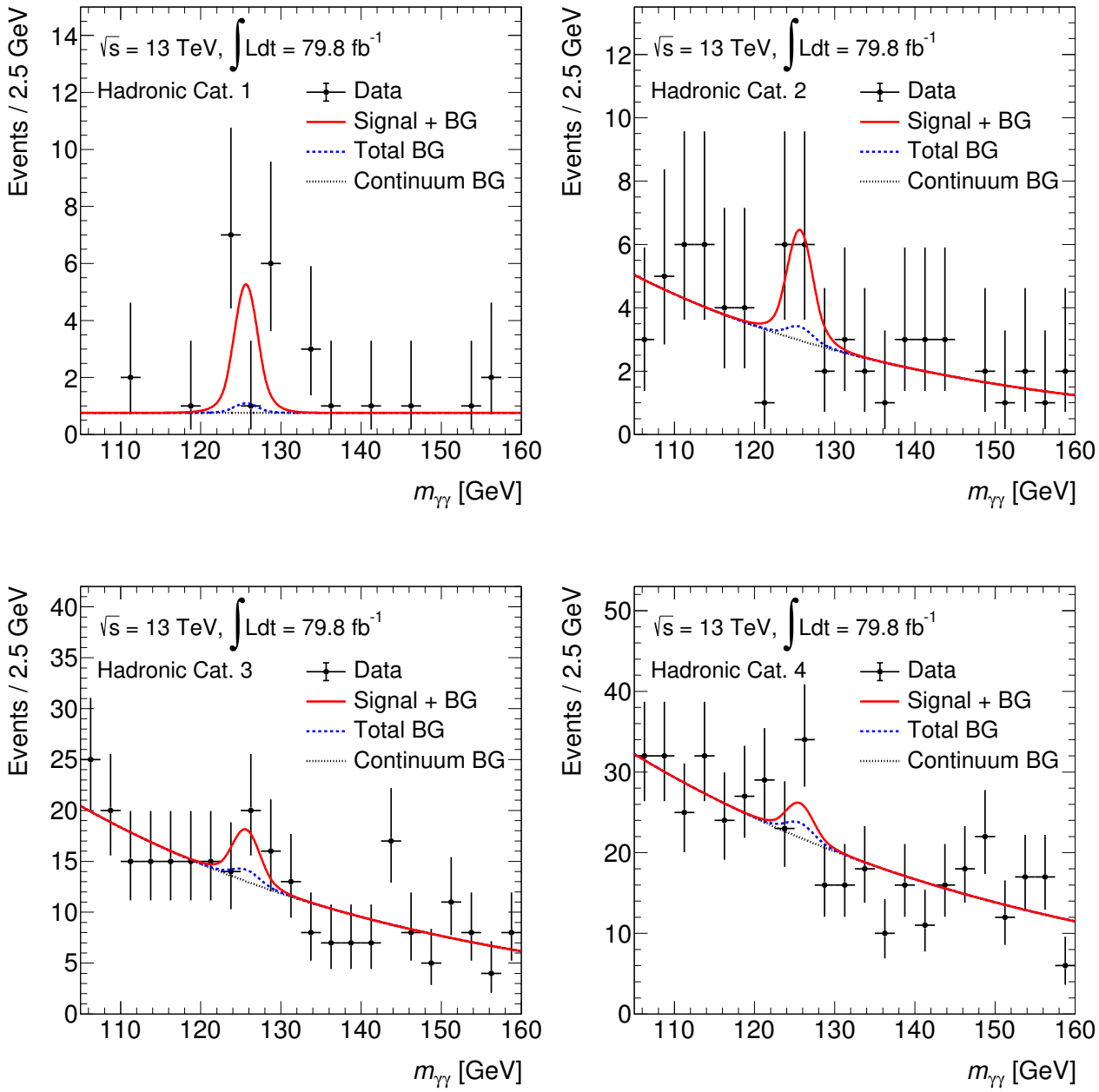


Figure 6.15: Di-photon invariant mass distribution of the observed events for four categories in the Hadronic region. The solid red curve shows the fitted signal + background model. The dashed blue and dotted black curves show the total and continuum background components of the fit, respectively.

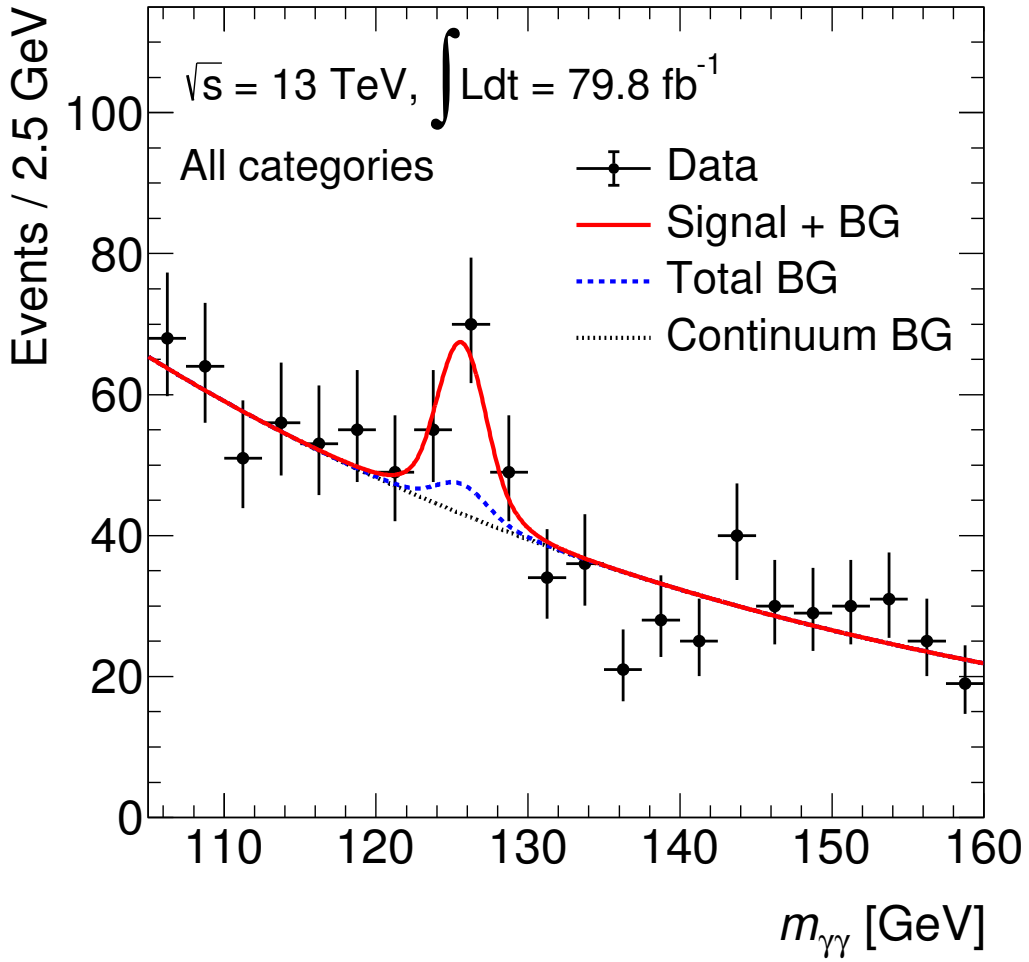


Figure 6.16: Di-photon invariant mass distribution of all the observed events including both the Leptonic and Hadronic regions. The solid red curve shows the fitted signal + background model. The dashed blue and dotted black curves show the total and continuum background components of the fit, respectively.

the Hadronic region has similar or better Z_0 compared to the Leptonic region.

Table 6.11: The numbers of signals S and backgrounds B in the interval of $121 < m_{\gamma\gamma} < 129$ GeV, the signal to background ratio S/B, and the significance Z_0.

Category	S	B	S/B	Z_0
Leptonic Cat. 1	6.4	3.3	1.9	2.9
Leptonic Cat. 2	3.2	6.1	0.5	1.2
Leptonic Cat. 3	1.1	5.8	0.2	0.5
Hadronic Cat. 1	6.2	2.9	2.1	2.9
Hadronic Cat. 2	4.9	10	0.5	1.4
Hadronic Cat. 3	6.7	44	0.2	1.0
Hadronic Cat. 4	4.2	74	0.1	0.5

Now we know the existence of the ttH process, and evaluate the cross section from the signal strength μ_{ttH} assuming $\sigma_{ttH}^{SM} = 506.5$ fb which is calculated with the NLO QCD. The profile of the negative log-likelihood ratio $q_\mu = -2\ln\lambda(\mu_{ttH})$ is shown in Figure 6.17. The best fit value of μ_{ttH} is found to be

$$\mu_{ttH} = 1.37\ ^{+0.39}_{-0.35}\ (\text{stat.})\ ^{+0.28}_{-0.21}\ (\text{syst.}) = 1.37\ ^{+0.48}_{-0.41}. \tag{6.10}$$

The statistical uncertainty is still dominant, which is about 2.1 times larger than the systematic uncertainty.

The cross section of the ttH process is calculated from the Equation 1.26, assuming $BR(H \to \gamma\gamma)^{obs} = BR(H \to \gamma\gamma)^{SM} = 2.27\ ^{+0.07}_{-0.06} \times 10^{-3}$. The cross section σ_{ttH}^{obs} is found to be

$$\sigma_{ttH}^{obs} = \left[694.9\ ^{+198.0}_{-179.3}\ (\text{stat.})\ ^{+141.7}_{-108.2}\ (\text{syst.})\right]\ \text{fb} = \left[694.9\ ^{+243.1}_{-207.7}\right]\ \text{fb}. \tag{6.11}$$

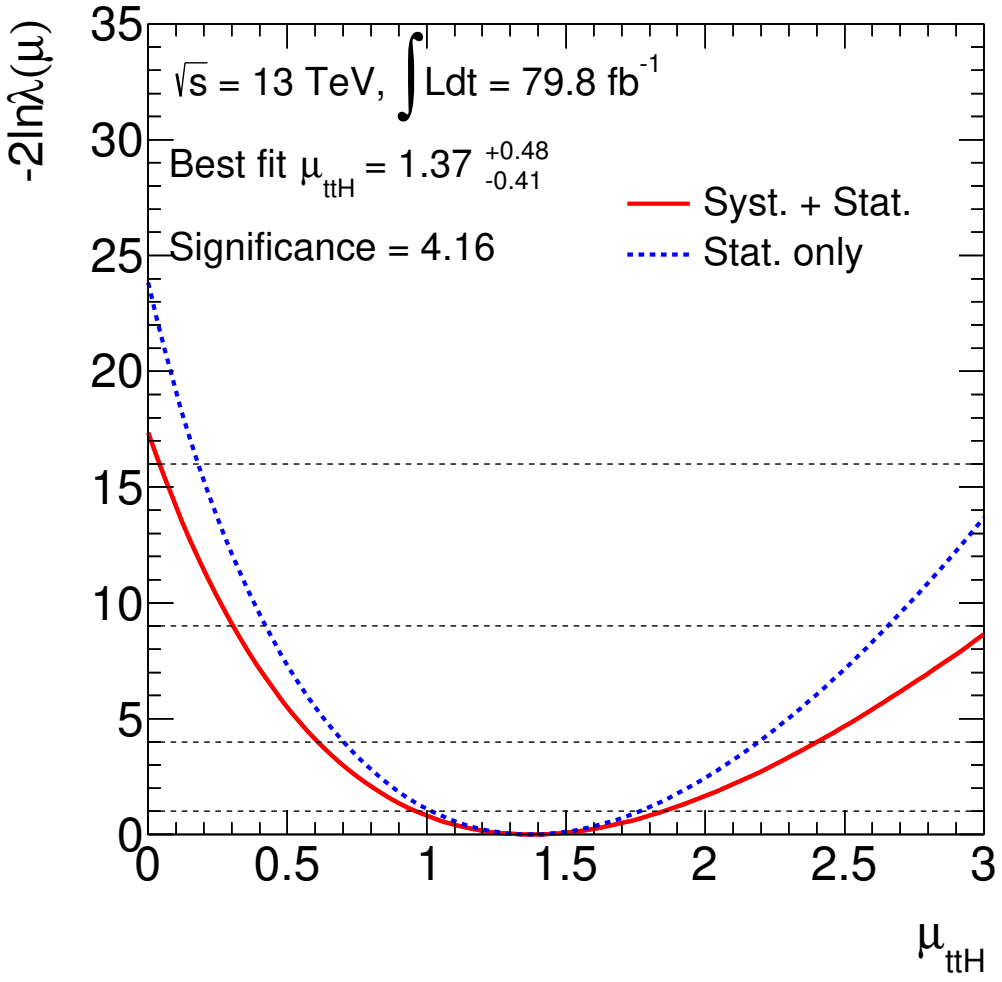


Figure 6.17: The profile of the negative log-likelihood ratio $-2\ln\lambda(\mu_{ttH})$ of the signal strength μ_{ttH} of data in case of the SM. The intersections of the solid and dashed curves with the horizontal dashed line at $-2\ln\lambda(\mu_{ttH}) = 1$ indicate that the 68 % confidence intervals of the expected results.

Chapter 7

Discussion

We compare the observed ttH cross section with the SM prediction in Section 7.1. Due to the large uncertainty, we cannot conclude that there are signatures of the new physics. We discuss the possibility how to reduce the uncertainty in Section 7.2. We also discuss the idea to further separate $t\bar{t} + \gamma\gamma$ events from the signal in Section 7.3.

7.1 Comparison with the theoretical prediction

The observed cross section is about 1.4 times larger than the SM prediction $\sigma_{ttH}^{SM} = 506.5\ ^{+34.6}_{-50.5}$ fb calculated with the NLO QCD although the uncertainty in the measurement is large.

The CMS experiment also published the result of the search for $ttH(H \to \gamma\gamma)$ process with 77.4 fb^{-1} of proton-proton collision data [79]. Table 7.1 shows the results by the ATLAS (this thesis) and CMS. The statistical significance in the CMS is also found to be $Z_0 > 4\ \sigma$, which supports the existence of the $ttH(H \to \gamma\gamma)$. The signal strength μ_{ttH} is slightly higher than the SM prediction as the ATLAS result. Although it is a wild trial, the simple weighted average of μ_{ttH} by the ATLAS and CMS is $\mu_{ttH}^{comb} = 1.5\ ^{+0.4}_{-0.3}$.

Table 7.1: The results of the $ttH(H \to \gamma\gamma)$ signal strength in the ATLAS (this thesis) and CMS [79]. Z_0 is the statistical significance of the $ttH(H \to \gamma\gamma)$ process. The combined values are the simple minded weighted average of the ATLAS and CMS results.

	Luminosity	Z_0 (observed)	Z_0 (expected)	μ_{ttH}
ATLAS (this thesis)	79.8 fb^{-1}	4.2 σ	3.6 σ	$1.4\ ^{+0.5}_{-0.4}$
CMS	77.4 fb^{-1}	4.1 σ	2.7 σ	$1.7\ ^{+0.6}_{-0.5}$
Combined		5.9 σ	4.5 σ	$1.5\ ^{+0.4}_{-0.3}$

One of the new physics models that possibly predicts the deviation of the Top-Yukawa coupling from the SM is the Two Higgs Doublet Model (2HDM) [80, 81]. However, the precision of the measurement needs to be improved to discriminate such new physics models from the SM.

7.2 Possible reduction of uncertainty

The uncertainty of the measured cross section $\sigma_{ttH}^{obs} = [694.9\ ^{+198.0}_{-179.3}\ (\text{stat.})\ ^{+141.7}_{-108.2}\ (\text{syst.})]$ fb is dominated by the statistical uncertainty while the systematic uncertainty is not negligible. Assuming the statistical uncertainty will be improved in future at the High-Luminosity LHC, we discuss the possibility to reduce the systematic uncertainty. Furthermore, we also discuss about the theoretical uncertainty which is needed in the comparison with the measurement.

7.2.1 Systematic uncertainty

The systematic uncertainty of the cross section is assigned to be 18 % in the current result. Since this will be dominant in future, it is important to discuss how to reduce the systematic uncertainty even now.

Systematic uncertainty related to the parton shower, hadronization, underlying event and PDF

The largest contribution of the systematic uncertainty comes from the modeling of the parton shower, hadronization, underlying event and parton distribution function (PDF). Figure 7.1 shows the p_T distributions for the leading and sub-leading jets for the signal MC samples generated with the nominal and alternative models. As shown in the plots, the different models predict the different jet p_T distributions. This comparison does not allow to identify the most important systematic effect, because the modeling of parton shower, hadronization, underlying event and PDF are changed at the same time from the nominal MC sample. To understand what causes the difference, these model have to be changed one by one.

Systematic uncertainty related to the measurement of each object

There are two main objects which have relatively large uncertainty; jet and photon. The source of uncertainties related to jet is dominated by the jet energy scale. This uncertainty mainly affects the XGBoost score. The uncertainty of the jet energy scale is already suppressed to be 1 % at 100 GeV as discussed in Section 4.5. However, it still affects the signal acceptance because there are many jets in the final state. One of the relatively large contributions of the jet energy scale uncertainty comes from the hadronization modeling. The uncertainty of the jet energy scale is assigned as the difference of the response R defined in Section 4.5, which is estimated by the comparison between PYTHIA8 and Herwig++. However, there are not only the hadronozation modeling but many other uncertainty sources in the jet energy scale. To reduce the jet energy scale uncertainty, sub percent level high precision calibration is needed for each source, which is challenging. As shown in Figure 4.13, low p_T jets have relatively large jet energy scale uncertainty. Thus it is possible to reduce the uncertainty if lower p_T ordered jets are removed from the MVA inputs. Since fourth (sixth) jet could not be the one originated from top-quarks in the Leptonic (Hadronic) region, their removal may make a small change of the XGBoost distribution although further studies are needed to make it clear.

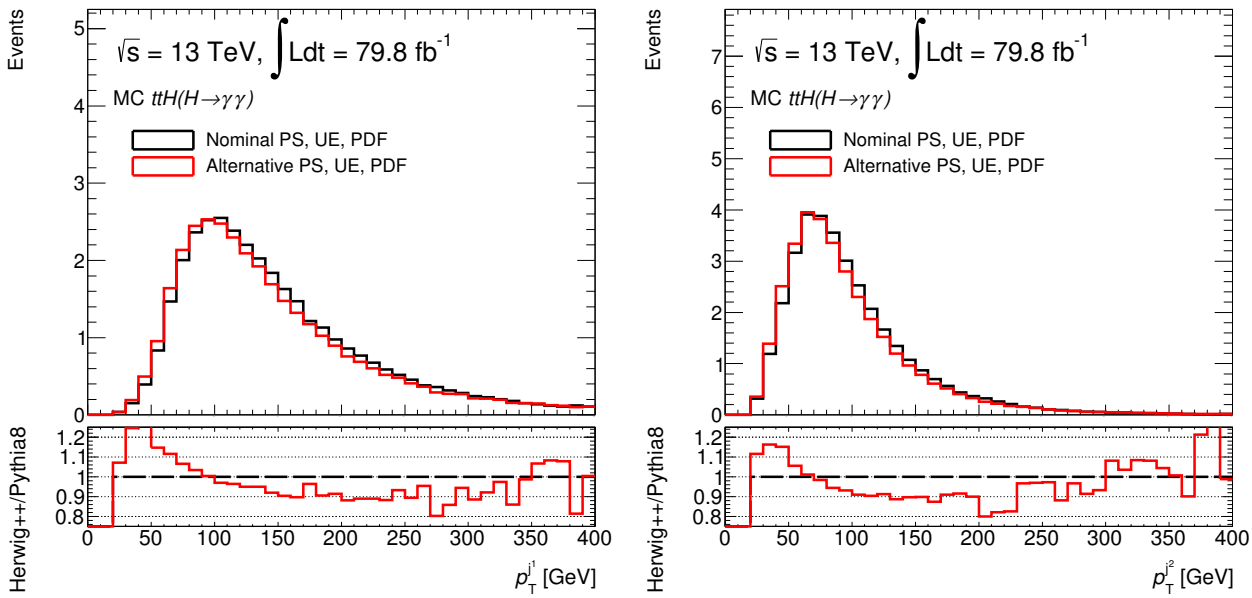


Figure 7.1: The p_T distribution of the leading (left) and sub-leading (right) jets for the $ttH(H \to \gamma\gamma)$ MC generated with the nominal and alternative parton shower and hadronization (PS), underlying event (UE) and PDF model. Events are required to pass the di-photon selection.

Because there are various sources of the uncertainty related to photon, we need to make various efforts to reduce photon systematics. One of the relatively large uncertainty sources comes from the modeling of material inside the EM calorimeter, which mainly affects the photon energy resolution and the photon selection efficiency. Currently, the material of the Inner Detector is estimated by counting the number of converted photons using data taken in 2015 [85]. The precision of the material estimation is limited by the understanding of the reconstruction efficiency of converted photons. Because we currently cannot measure the reconstruction efficiency, its uncertainty is estimated by varying the selection criteria where the variation is artificially large to be conservative. In order to estimate the material uncertainty precisely, we need to establish the method to measure the reconstruction efficiency of converted photons.

7.2.2 Theoretical uncertainty for the SM ttH cross section

The ttH cross section in the SM prediction is $\sigma_{ttH}^{SM} = 506.5\ ^{+34.6}_{-50.5}$ fb with the NLO QCD calculation. This theoretical uncertainty is relatively smaller than the experimental uncertainty. However, it will be important when the experimental uncertainties are reduced in the future to look for a deviation from the SM prediction.

The variation of the renormalization and factorization scales in the QCD calculation largely contributes to the cross section uncertainty: $\Delta\sigma_{ttH}^{SM}/\sigma_{ttH}^{SM} = +5.8/-9.2$ %. To reduce the uncertainty, it is the most straightforward to perform the higher order QCD calculation (e.g. NNLO, N3LO, ...). Figure 7.2 shows the renormalization scale μ_r and factorization scale μ_f dependence of the ttH cross section calculated with LO and NLO [19]. The cross section is obtained by simultaneously varying μ_r and μ_f, with $\mu = \mu_r = \mu_f$ in this plot. The scale dependence becomes small as the cross section is calculated at higher order. Another possibility to improve the uncertainty is to include the next-to-leading logarithmic (NLL) corrections to the NLO QCD calculation. Table 7.2 shows the cross section uncertainties at the various order calculations. The NLO+NLL QCD calculation reduces the uncertainty by 70 % from the NLO QCD calculation.

Table 7.2: Cross section uncertainties for the ttH production process at the LHC [19]. The uncertainties are estimated varying μ_r and μ_f independently by considering the minimum and maximum values obtained with $(\mu_r/\mu,\ \mu_f/\mu) = (0.5, 0.5), (0.5, 1), (1, 0.5), (1, 1), (1, 2), (2, 1), (2, 2)$.

Process	$\sqrt{s}$	NLO	NLO+NLL	NLO+NLL (with C)
ttH	13 TeV	$^{+5.8}_{-9.2}$ [%]	$^{+4.2}_{-6.5}$ [%]	$^{+7.8}_{-5.5}$ [%]
	14 TeV	$^{+6.0}_{-9.2}$ [%]	$^{+4.5}_{-6.5}$ [%]	$^{+7.9}_{-5.5}$ [%]

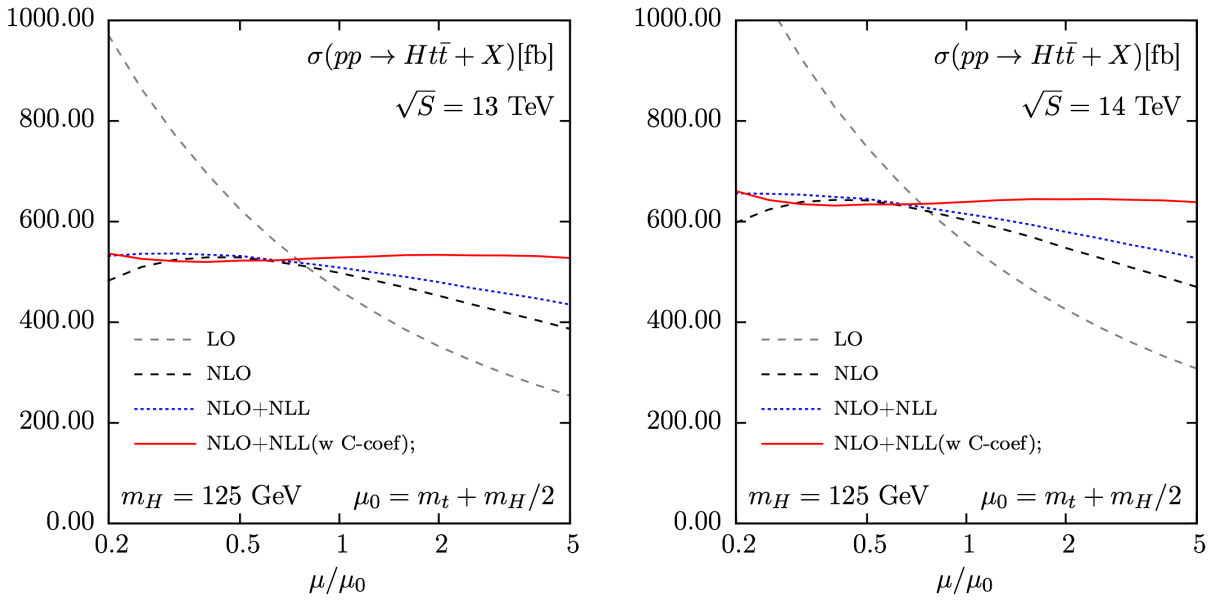


Figure 7.2: Renormalization scale μ_r and factorization scale μ_f dependence of the ttH cross section calculated with LO, NLO and NLO+NLL at $\sqrt{s} = 13$ TeV (left) and 14 TeV (right) [19]. The cross section is obtained by simultaneously varying μ_r and μ_f, with $\mu = \mu_r = \mu_f$.

7.3 Separation from the $t\bar{t}+\gamma\gamma$ background

Since $ttH(H \to \gamma\gamma)$ and $t\bar{t}+\gamma\gamma$ events have the same final state, it is difficult to reduce $t\bar{t}+\gamma\gamma$ backgrounds even with the MVA. In this section, we discuss about the idea to further reduce $t\bar{t}+\gamma\gamma$ backgrounds.

The training samples of the MVA are the signal MC and real data passing NTI selection as the inclusive background containing both $\gamma\gamma$ and $t\bar{t}+\gamma\gamma$. However, this background sample is dominated by $\gamma\gamma$ events and thus the characteristics of $t\bar{t}+\gamma\gamma$ events are not shown up.

To tune the MVA so that the reduction of $t\bar{t}+\gamma\gamma$ is enhanced, we develop two independent XGBoosts; one is for $ttH(H \to \gamma\gamma)$ and $\gamma\gamma$ separation and the other for $ttH(H \to \gamma\gamma)$ and $t\bar{t}+\gamma\gamma$ separation. For the training sample, each $\gamma\gamma$ and $t\bar{t}+\gamma\gamma$ MC is used instead of data. Since the number of $\gamma\gamma$ MC events are very limited in the Leptonic region, this study covers only for the Hadronic region. The input variables to the XGBoost are the same as the nominal analysis (Table 5.3 shows the list of variables). In the following, the XGBoost trained by $ttH(H \to \gamma\gamma)$ and $\gamma\gamma$ samples are called as XGBoost($\gamma\gamma$), and the one trained by $ttH(H \to \gamma\gamma)$ and $t\bar{t}+\gamma\gamma$ as XGBoost($t\bar{t}+\gamma\gamma$), respectively. Figure 7.3 shows the distributions of the XGBoost($\gamma\gamma$) and XGBoost($t\bar{t}+\gamma\gamma$) scores. $t\bar{t}+\gamma\gamma$ events are better separated from $ttH(H \to \gamma\gamma)$ by the XGBoost($t\bar{t}+\gamma\gamma$).

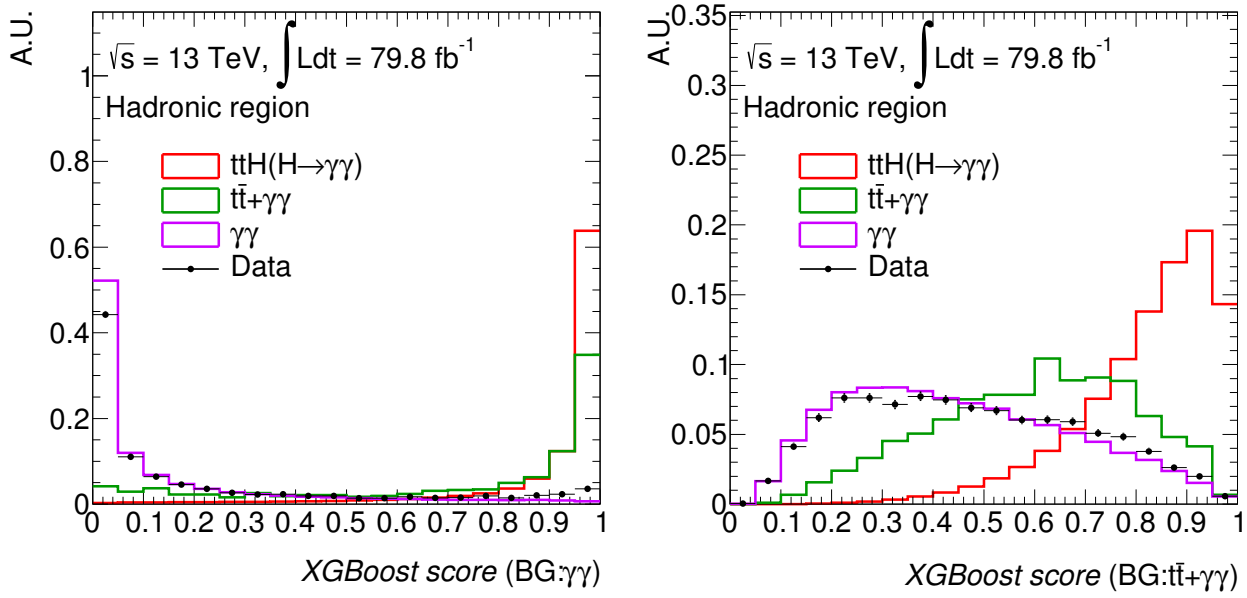


Figure 7.3: XGBoost score distributions. $\gamma\gamma$ MC (left) and $t\bar{t}+\gamma\gamma$ MC (right) are used as the training samples, instead of data. Events are normalized to unity.

To reduce not only $t\bar{t}+\gamma\gamma$ but also $\gamma\gamma$ backgrounds, we need to check the XGBoost($t\bar{t}+\gamma\gamma$) distribution with the high XGBoost($\gamma\gamma$) score region. Figure 7.4 shows the two dimensional distributions of the XGBoost($\gamma\gamma$) and XGBoost($t\bar{t}+\gamma\gamma$) for $ttH(H \to \gamma\gamma)$, $\gamma\gamma$, $t\bar{t}+\gamma\gamma$ MC,

and data. It turned out that events with higher XGBoost($\gamma\gamma$) score tend to have higher XGBoost($t\bar{t}+\gamma\gamma$) score. Figure 7.5 shows the XGBoost($t\bar{t}+\gamma\gamma$) score distribution with XGBoost($\gamma\gamma$) score > 0.9. Signal events can be separated from $t\bar{t}+\gamma\gamma$ events although they have high XGBoost($\gamma\gamma$) score.

Events are categorized into four groups by two dimensional boundaries of the XGBoost($\gamma\gamma$) and XGBoost($t\bar{t}+\gamma\gamma$) scores. Table 7.3 and 7.4 show the summary of the nominal and new categorization results, respectively, where N_{sig}^{exp} is the number of expected signals assuming the SM and $N_{cont-BG}^{exp}$ is the number of continuum backgrounds estimated from sideband data. The signal to background ratio is about 1.8 times improved by the new categorization, although this categorization is not introduced in this analysis because the expected significance is not improved. However, the systematic uncertainty related to background could be reduced by the improvement of the signal to background ratio. The expected significance Z_0 has large uncertainty because of the lack of events in the sideband. The benefit of the new categorization could be appeared when the statistics is more abundant.

Table 7.3: The numbers of signals and backgrounds, expected significance and the signal to background ratio for each category in the Hadronic region, using the nominal XGBoost as used in Section 5.4.4. Only the statistical uncertainty is considered.

Category	N_{sig}^{exp}	$N_{cont-BG}^{exp}$	Significance Z_0	$N_{sig}^{exp}/N_{cont-BG}^{exp}$
Cat. 1	3.9 ± 0.0	1.4 ± 0.5	2.5 ± 0.3	2.9 ± 1.1
Cat. 2	2.9 ± 0.0	6.7 ± 1.5	1.1 ± 0.1	0.4 ± 0.1
Cat. 3	3.9 ± 0.0	22.2 ± 2.5	0.8 ± 0.0	0.2 ± 0.0
Cat. 4	2.4 ± 0.0	36.6 ± 2.8	0.4 ± 0.0	0.1 ± 0.0
Combined	12.5 ± 0.0	66.9 ± 4.1	2.9 ± 0.3	0.2 ± 0.0

Table 7.4: The numbers of signals and backgrounds, expected significance and the signal to background ratio for each category in the Hadronic region, using the XGBoost($\gamma\gamma$) and XGBoost($t\bar{t}+\gamma\gamma$). Only the statistical uncertainty is considered.

Category	N_{sig}^{exp}	$N_{cont-BG}^{exp}$	Significance Z_0	$N_{sig}^{exp}/N_{cont-BG}^{exp}$
Cat. 1	3.0 ± 0.0	1.3 ± 0.5	2.1 ± 0.2	2.3 ± 0.8
Cat. 2	2.2 ± 0.0	2.0 ± 0.6	1.4 ± 0.2	1.1 ± 0.3
Cat. 3	1.5 ± 0.0	3.5 ± 0.8	0.8 ± 0.1	0.4 ± 0.1
Cat. 4	4.7 ± 0.0	26.8 ± 2.3	0.9 ± 0.0	0.2 ± 0.0
Combined	11.4 ± 0.0	33.7 ± 2.6	2.7 ± 0.3	0.3 ± 0.0

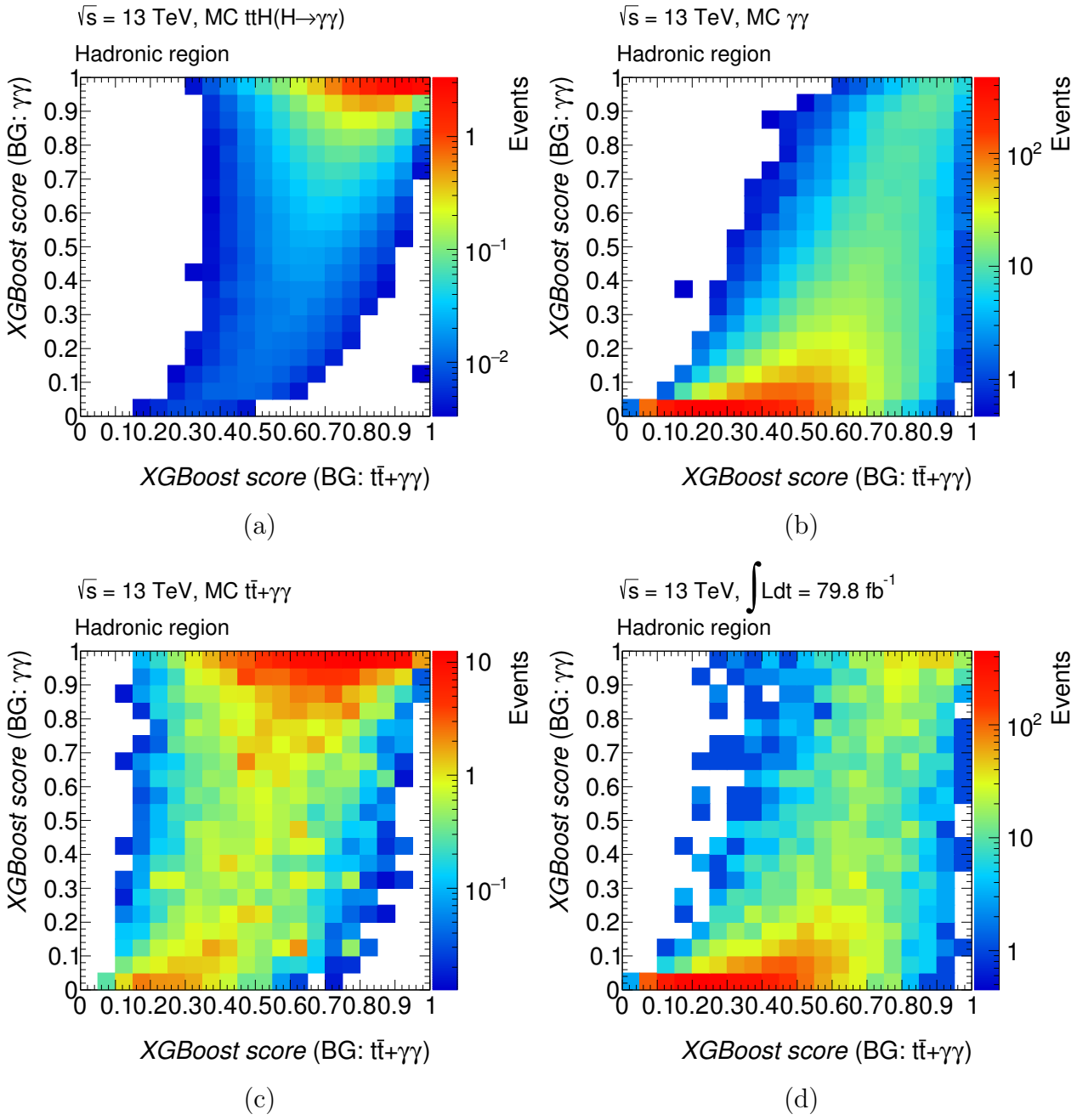


Figure 7.4: Two dimensional distributions of the XGBoost($\gamma\gamma$) and XGBoost($t\bar{t}+\gamma\gamma$) for $ttH(H \to \gamma\gamma)$ MC (a), $\gamma\gamma$ MC (b), $t\bar{t}+\gamma\gamma$ MC (c), and data (d).

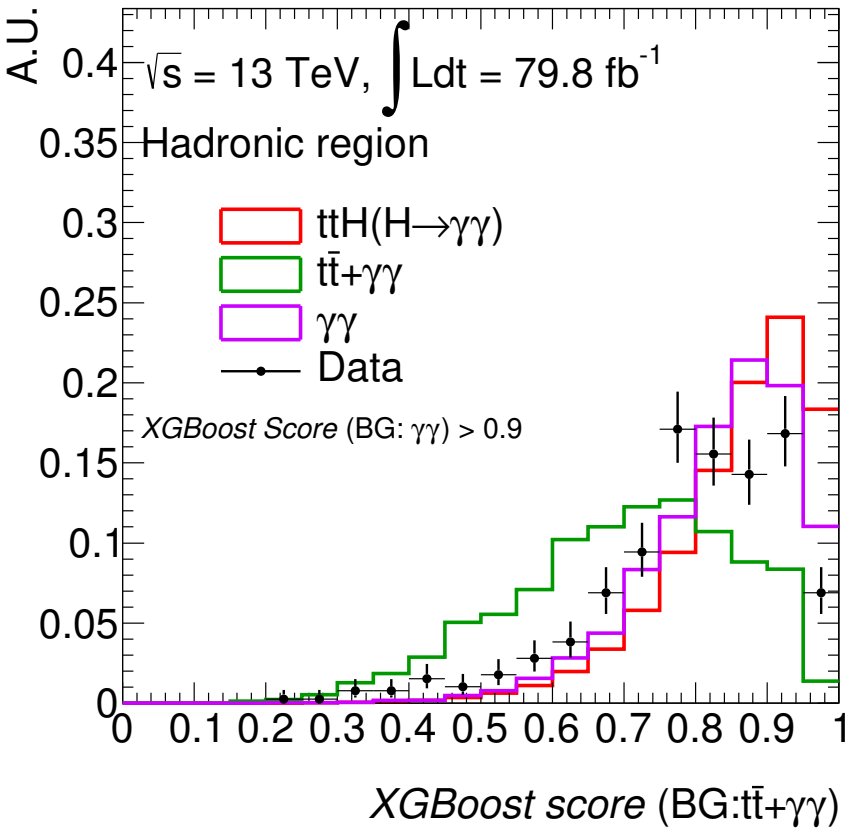


Figure 7.5: XGBoost($t\bar{t}+\gamma\gamma$) score distribution requiring XGBoost($\gamma\gamma$) score > 0.9 selection criterion. Events are normalized to unity.

Chapter 8

Conclusion

We performed the search for the production of the Higgs boson in association with top-park pair with the $H \to \gamma\gamma$ decay channel using proton-proton 13 TeV collision data accumulated by the ATLAS detector.

Signal events are selected by the multivariate analysis using the kinematic variables (p_T, η, ϕ, E) of the final state objects. The selected events are categorized into some groups with different signal to background ratio to improve signal sensitivity.

We found the "Higgs-like" events around 125 GeV in the di-photon invariant mass distribution although the statistics is not enough to make clear peak. The unbinned maximum likelihood fit is performed to extract the number of signals to all categories simultaneously, indicating the statistical significance is 4.2 standard deviations away from the background-only hypothesis, while the expected significance is 3.6 standard deviations. This is the evidence for the Higgs boson production in association with top-quark pair with $H \to \gamma\gamma$ decay channel only.

The cross section of the ttH process σ_{ttH} is found to be

$$\sigma_{ttH} = [694.9\ ^{+198.0}_{-179.3}\ (\text{stat.})\ ^{+141.7}_{-108.2}\ (\text{syst.})]\ \text{fb}. \tag{8.1}$$

It is slightly higher than the Standard Model expectation, $\sigma_{ttH}^{SM} = 506.5^{+34.6}_{-50.0}$ fb, calculating with the NLO QCD, although it is consistent within the uncertainty. It indicates that some new physics models can be involved in the Top-Yukawa coupling. More precise measurement is needed to discover new physics.

Appendix A

Discriminating variables used for the photon selection

Table A.1 shows the list of discriminating variables used for the *Loose*, *Medium* and *Tight* photon selection.

Table A.1: Discriminating variables used for *Loose*, *Medium* and *Tight* selection [57].

Category	Description	*Loose*	*Medium*	*Tight*
Energy leakage to the hadronic calo.	Ratio of E_T in the first sampling layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ of $\|\eta\| > 1.52$)	✓	✓	✓
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ of $\|\eta\| > 1.52$)	✓	✓	✓
Middle layer of the EM calo.	Ratio of the energy in 3×7 ($\eta \times \phi$) cells over the energy in 7×7 cells centered around the photon cluster position	✓	✓	✓
	Lateral shower width, $\sqrt{(\Sigma E_i \eta_i^2)/(\Sigma E_i) - ((\Sigma E_i \eta_i)/(\Sigma E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	✓	✓	✓
	Ratio of the energy in 3×3 ($\eta \times \phi$) cells over the energy of 3×7 cells centered around the photon cluster position			✓
Strip layer of the EM calo.	Lateral shower width, $\sqrt{(\Sigma E_i (i - i_{max})^2)/(\Sigma E_i)}$, where i runs over all strips in a window of $20 \times (\eta \times \phi)$ strips, and i_{max} is the index of the highest energy strip measured in the strip layer			✓
	Total lateral shower width $\sqrt{(\Sigma E_i (i - i_{max})^2)/(\Sigma E_i)}$, where i runs over all strips in a window of 20×2 ($\eta \times \phi$) strips, and i_{max} is the index of the highest energy strip measured in the strip layer			✓
	Energy outside the core of the three central strips but within seven strips divided by energy within the three central strips			✓
	Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with the minimum value found between the first and second maxima			✓
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in the secondary maximum in the cluster to the sum of these energies		✓	✓
	Ratio of the energy in the first layer to the total energy of the EM cluster			✓

Appendix B

Distribution of the XGBoost Input variables

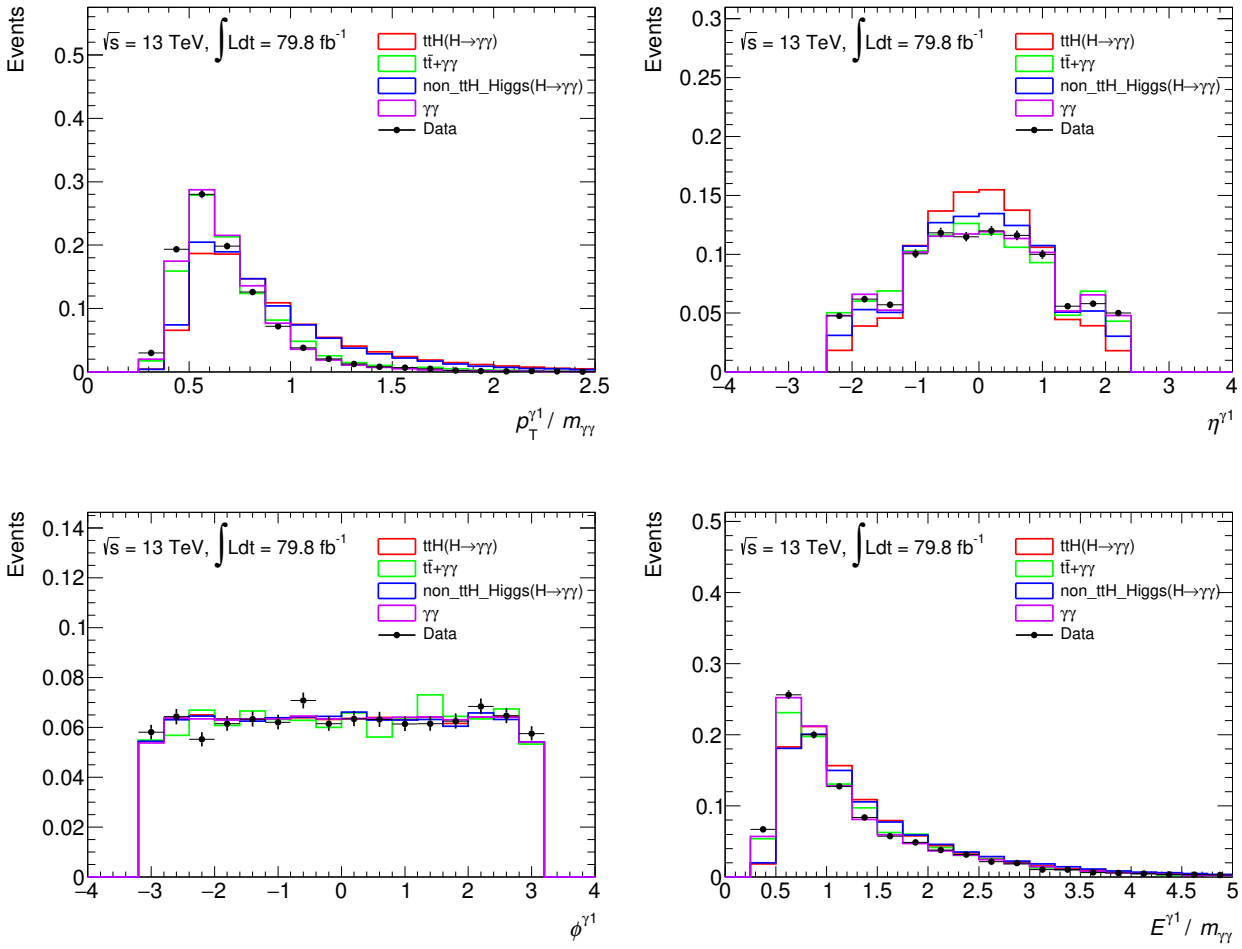


Figure B.1: Input variable distributions of the XGBoost for the Hadronic region related to the leading photon after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

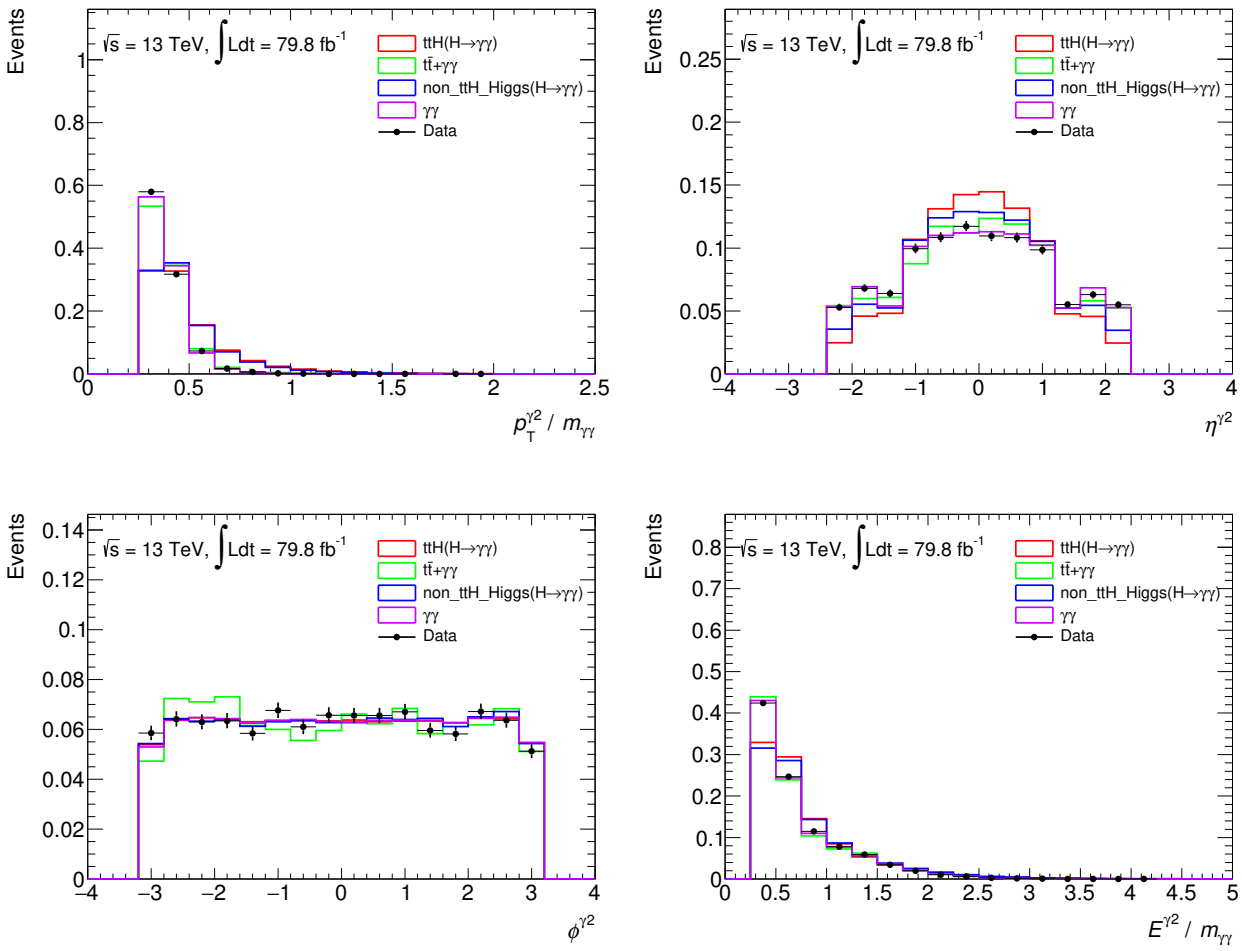


Figure B.2: Input variable distributions of the XGBoost for the Hadronic region related to the sub-leading photon after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

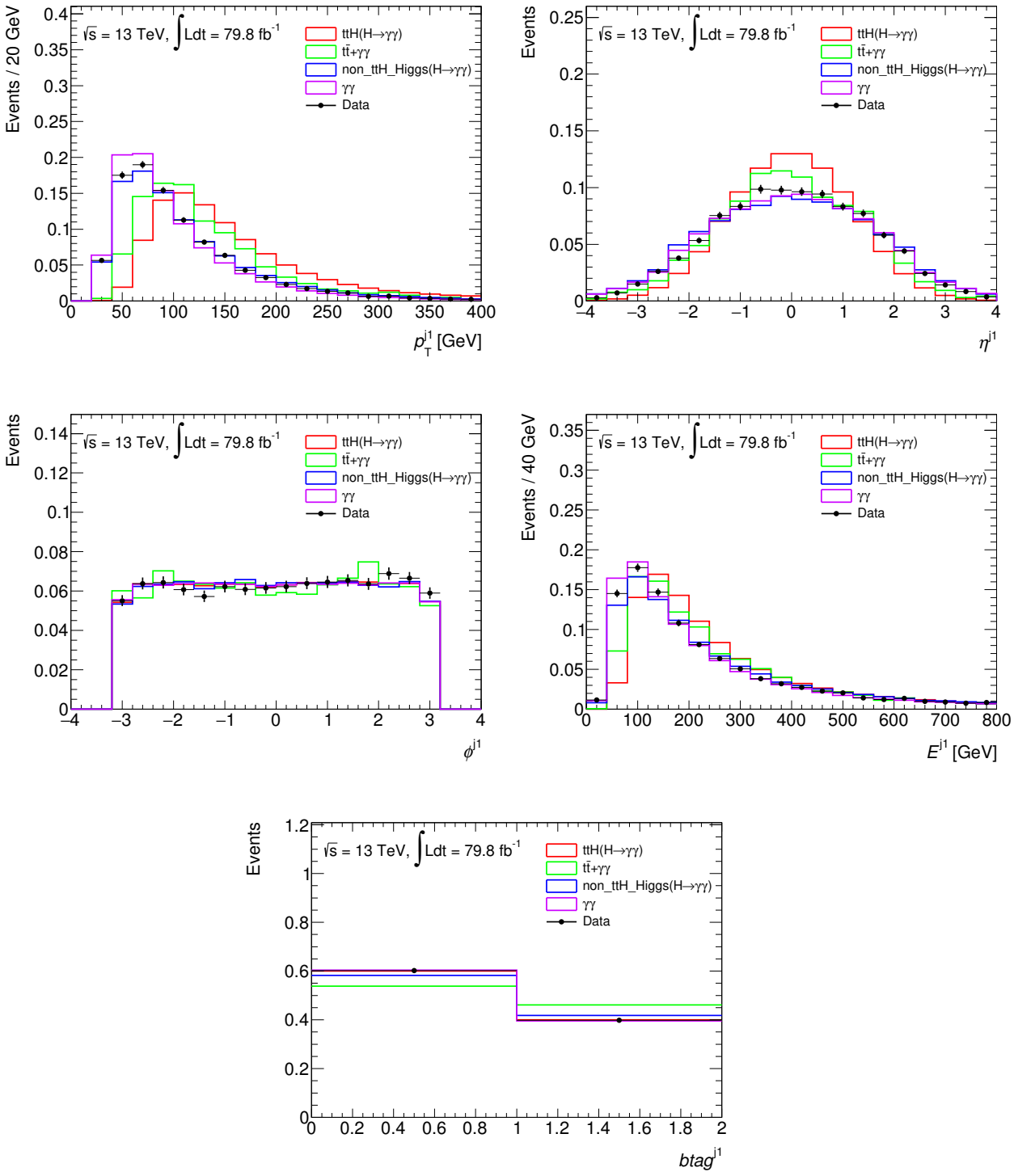


Figure B.3: Input variable distributions of the XGBoost for the Hadronic region related to the leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

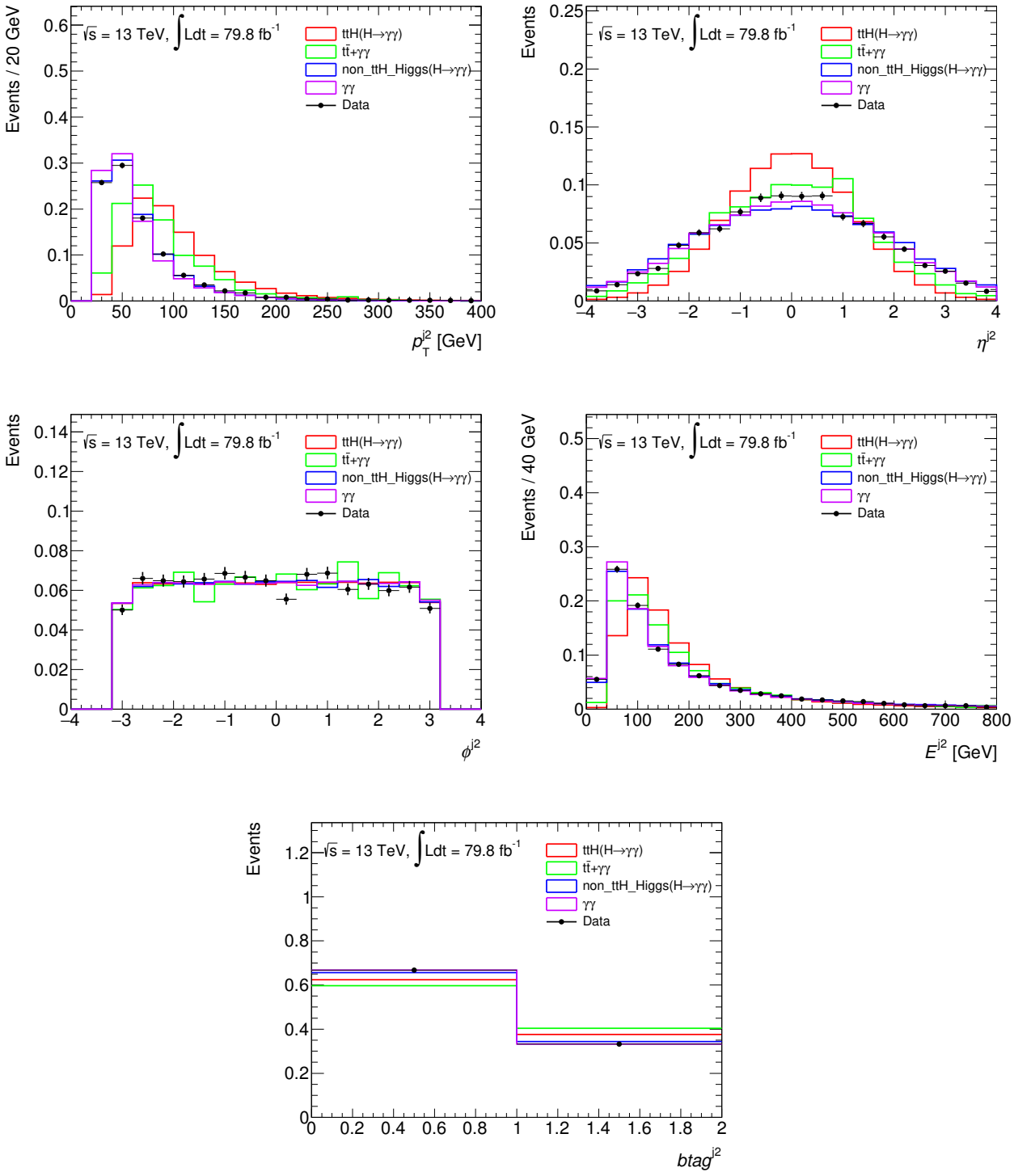


Figure B.4: Input variable distributions of the XGBoost for the Hadronic region related to the second leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

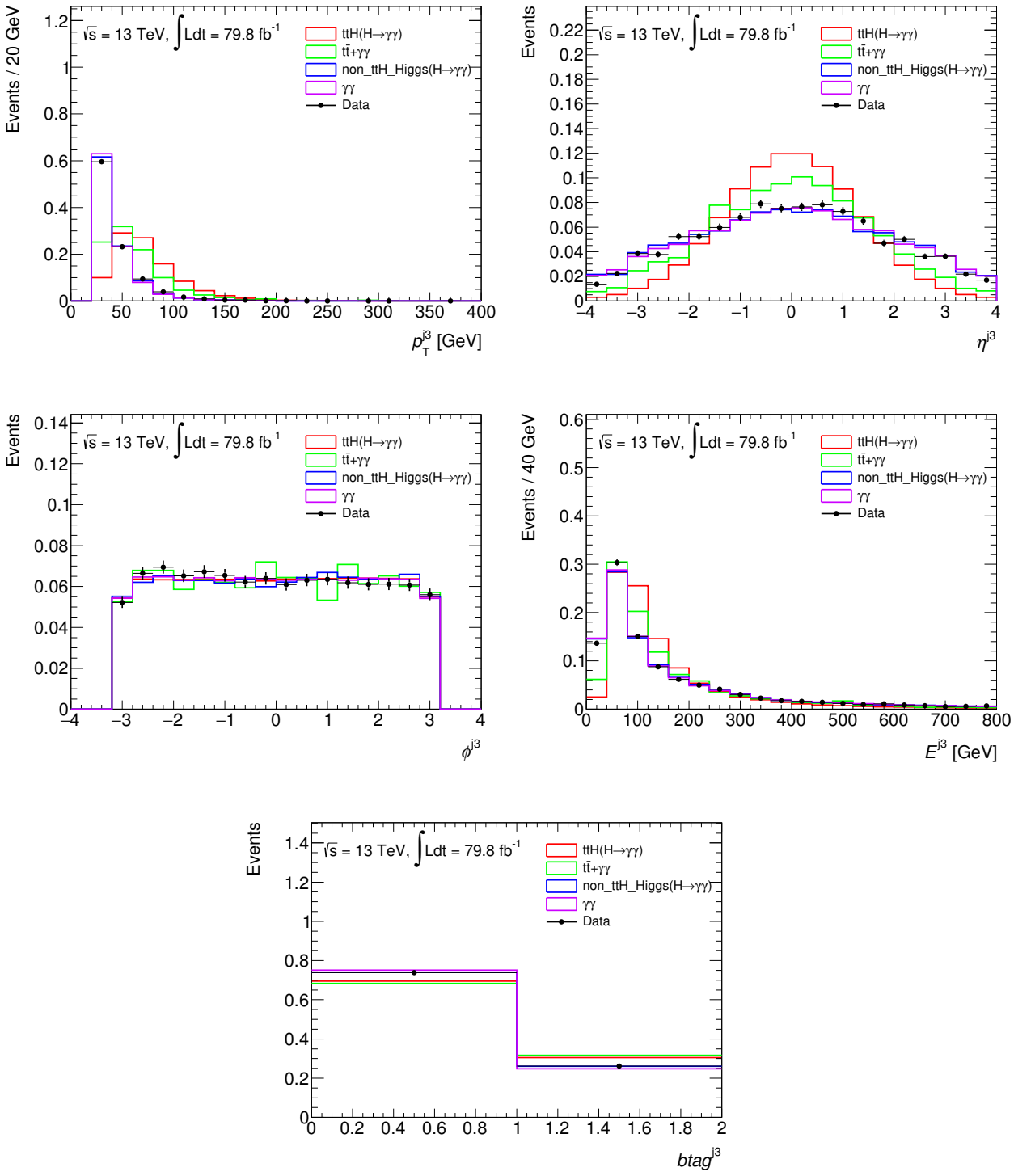


Figure B.5: Input variable distributions of the XGBoost for the Hadronic region related to the third leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

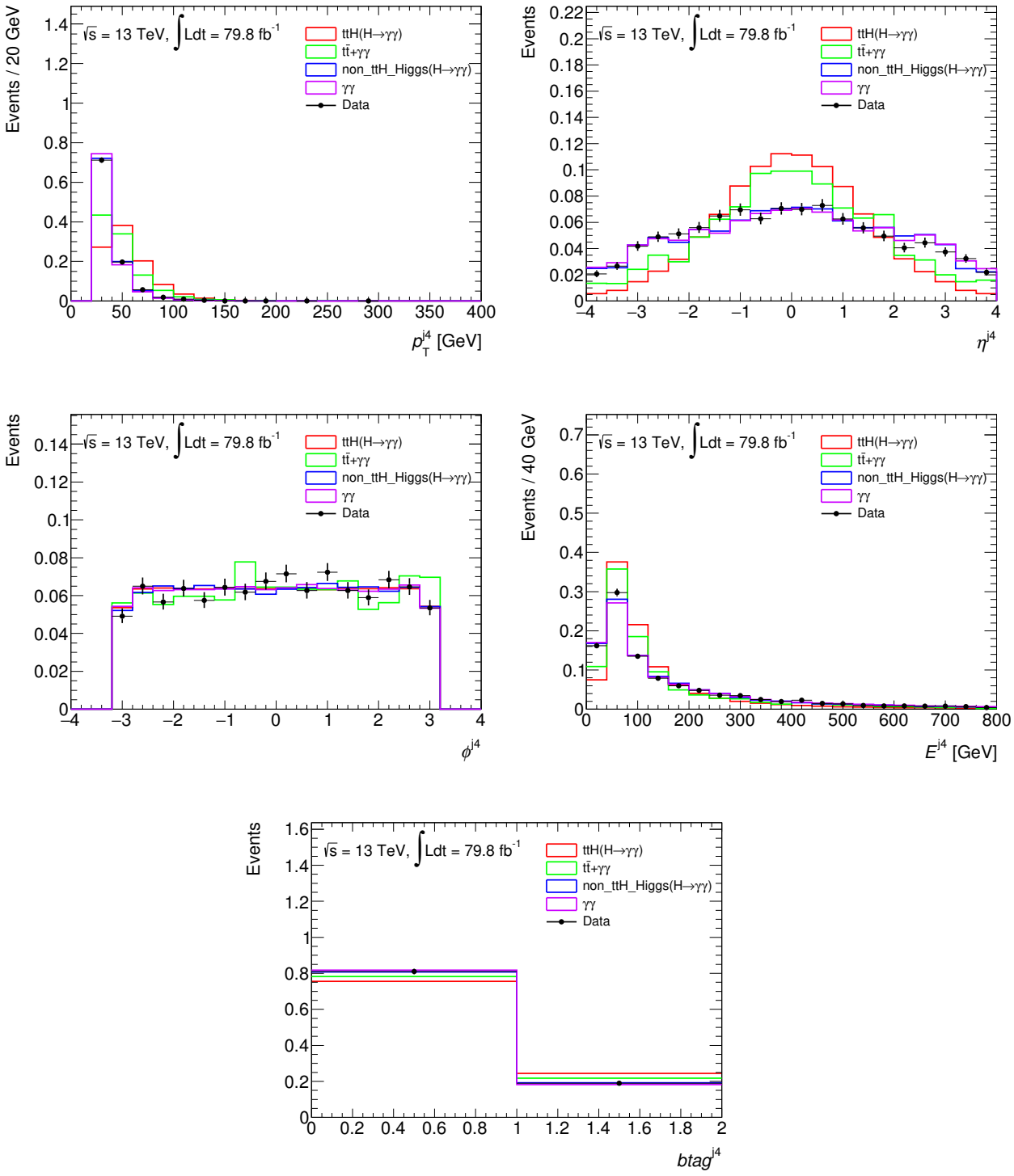


Figure B.6: Input variable distributions of the XGBoost for the Hadronic region related to the fourth leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

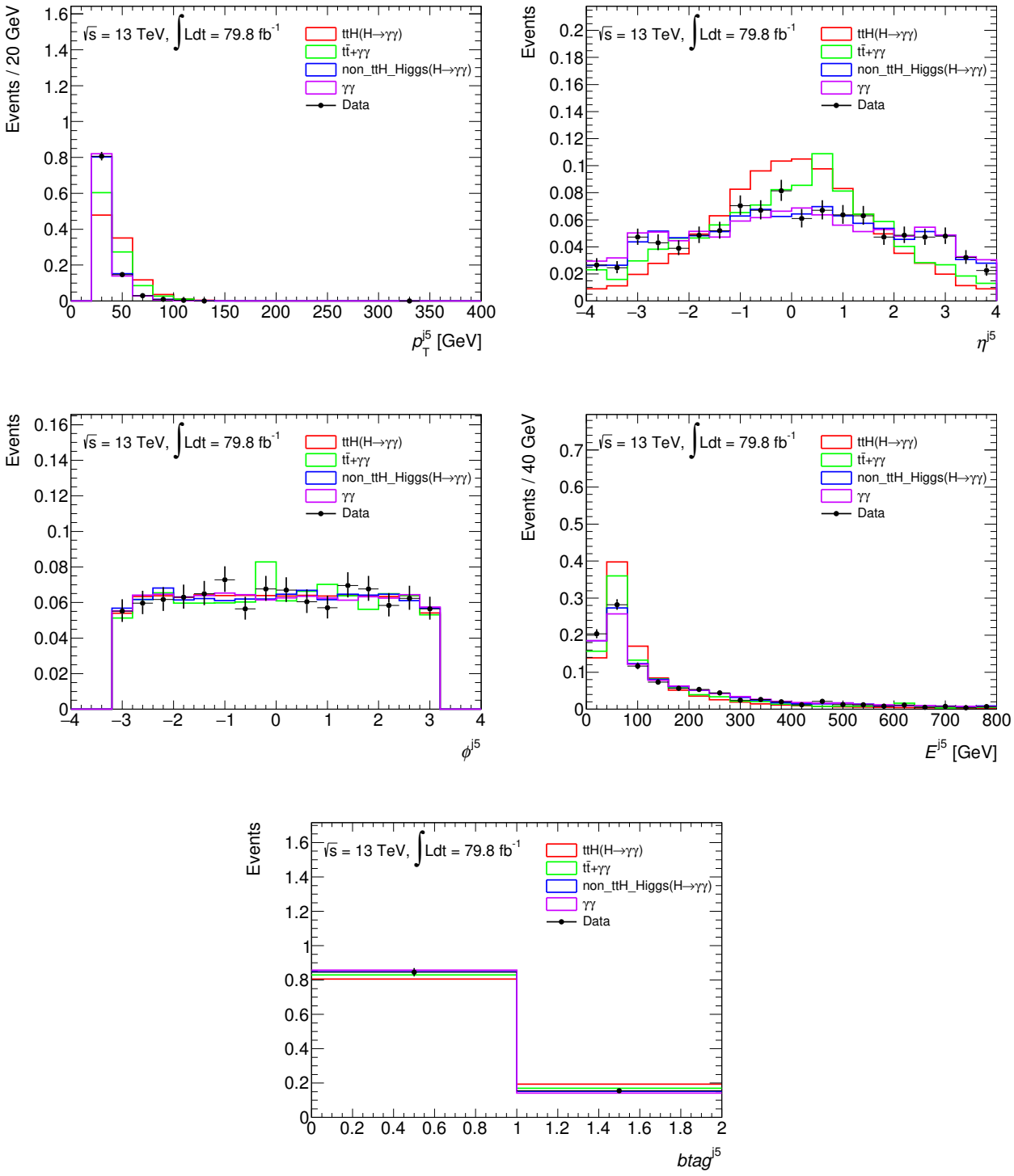


Figure B.7: Input variable distributions of the XGBoost for the Hadronic region related to the fifth leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

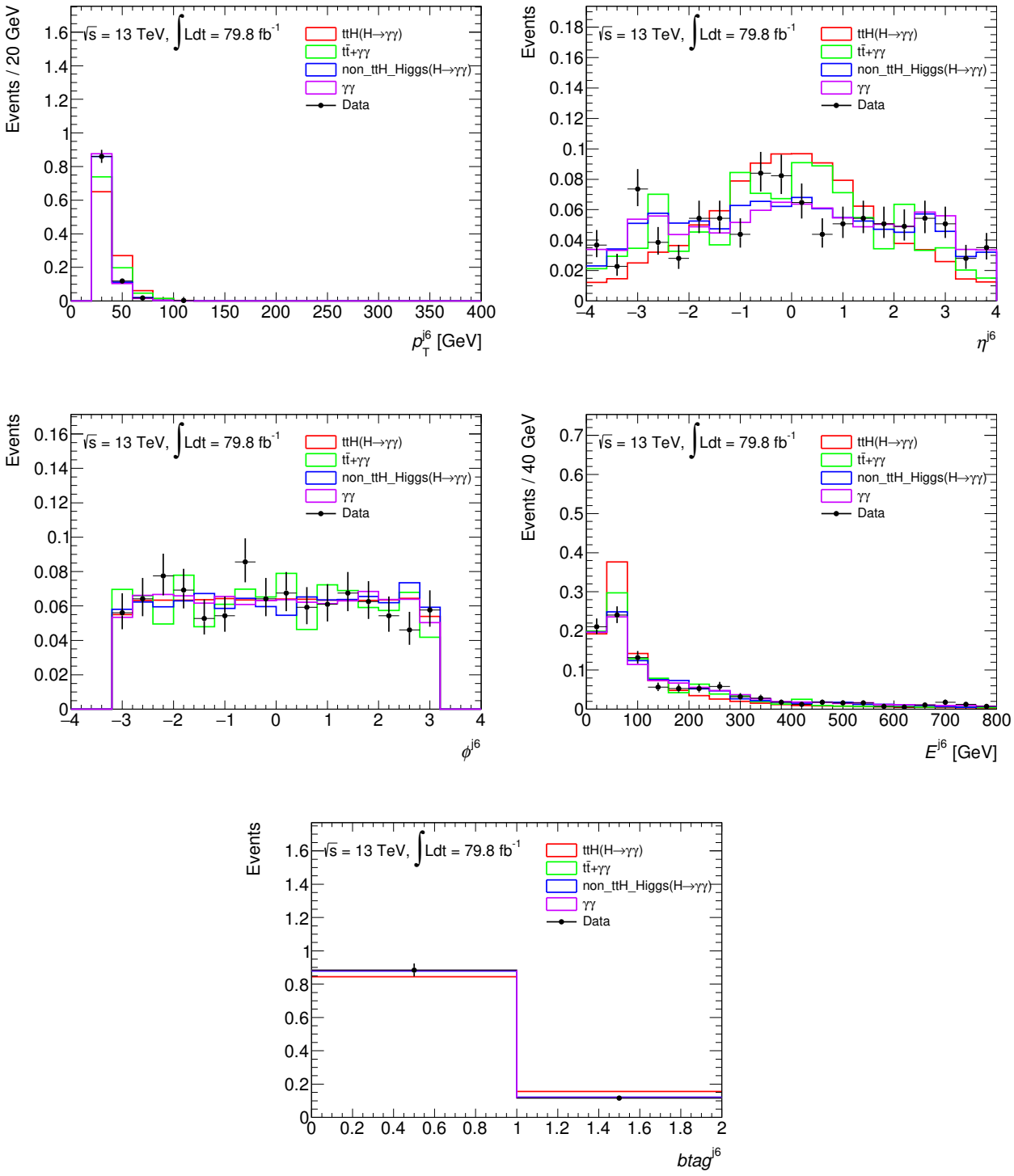


Figure B.8: Input variable distributions of the XGBoost for the Hadronic region related to the sixth leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

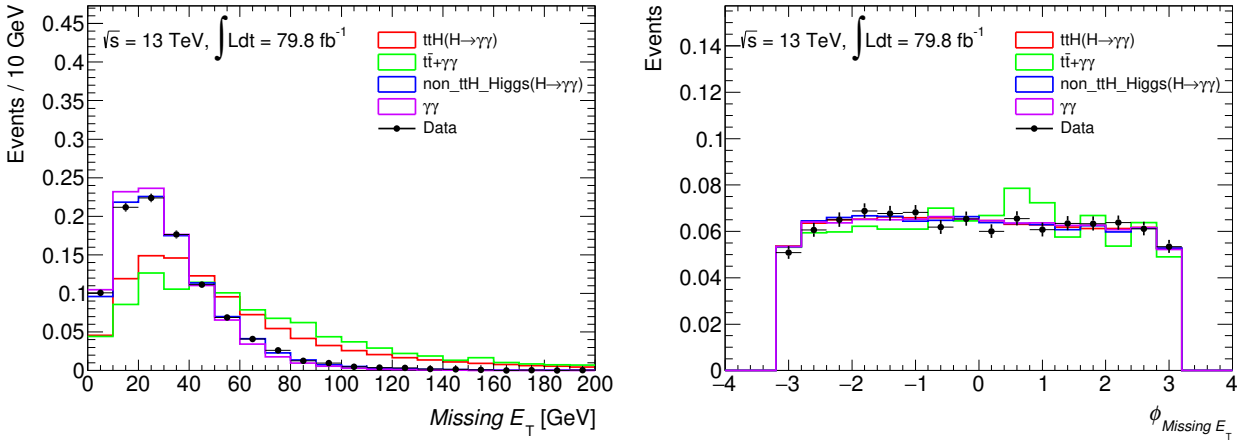


Figure B.9: Input variable distributions of the XGBoost for the Hadronic region related to the missing transverse energy E_T^{miss} after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

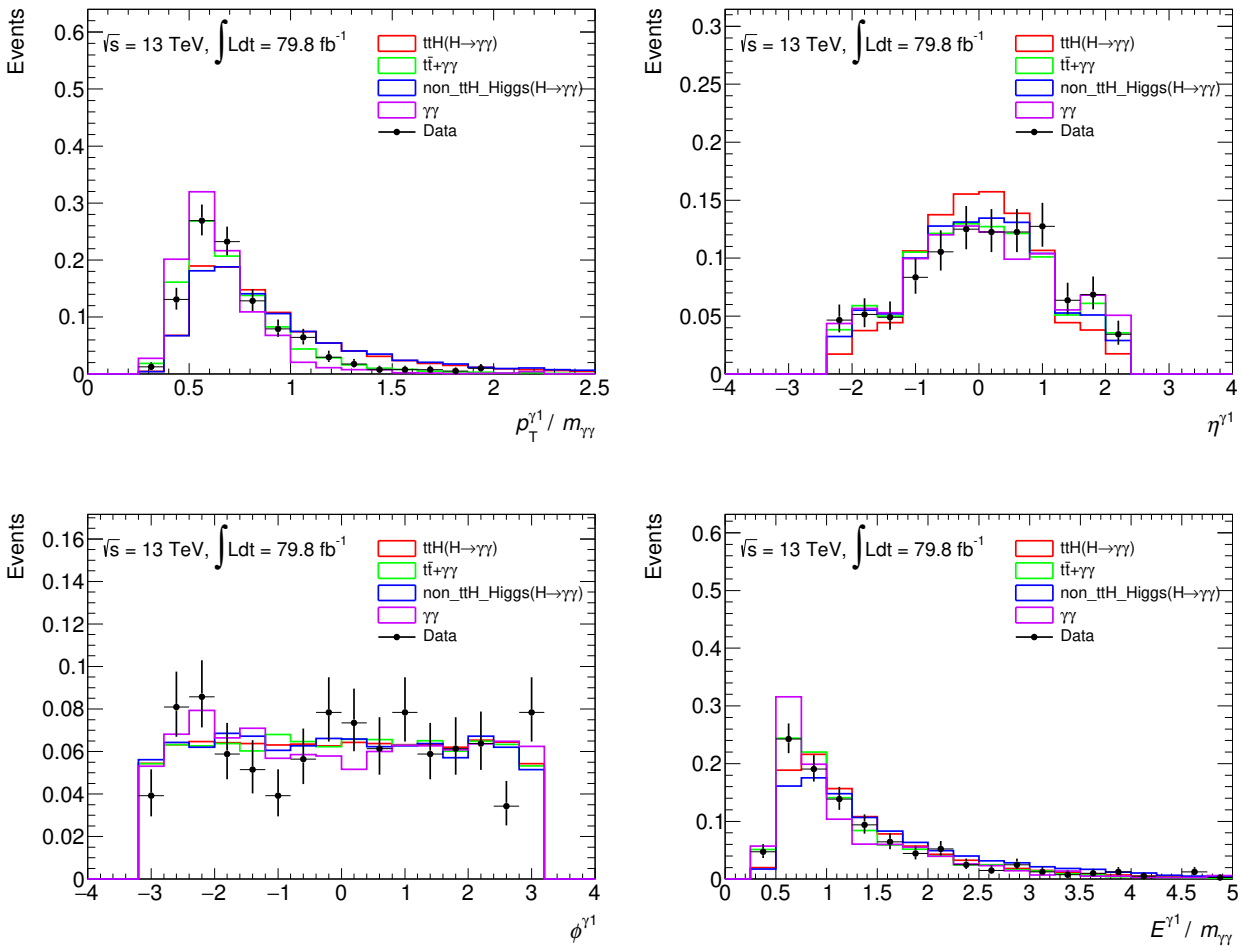


Figure B.10: Input variable distributions of the XGBoost for the Leptonic region related to the leading photon after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

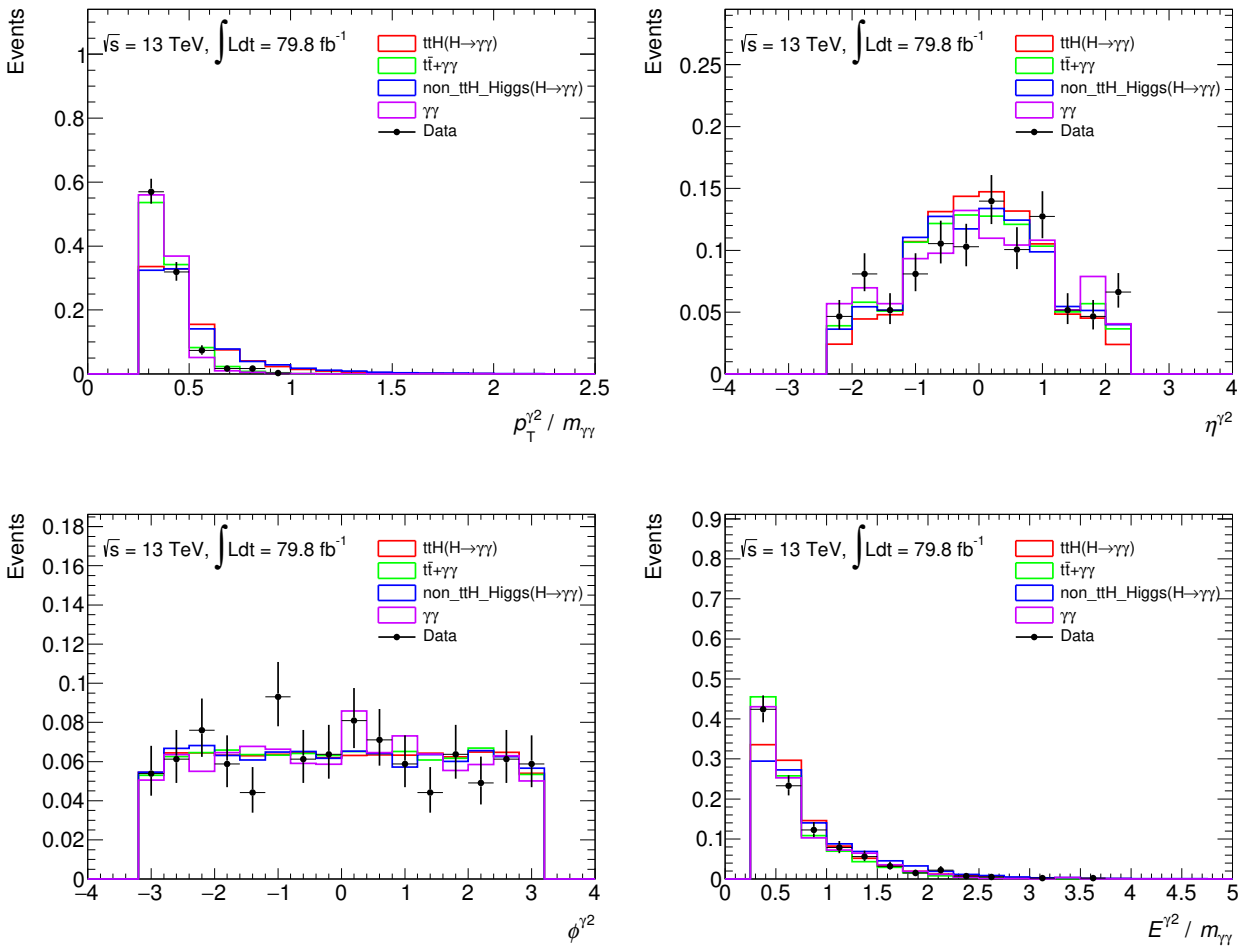


Figure B.11: Input variable distributions of the XGBoost for the Leptonic region related to the sub-leading photon after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

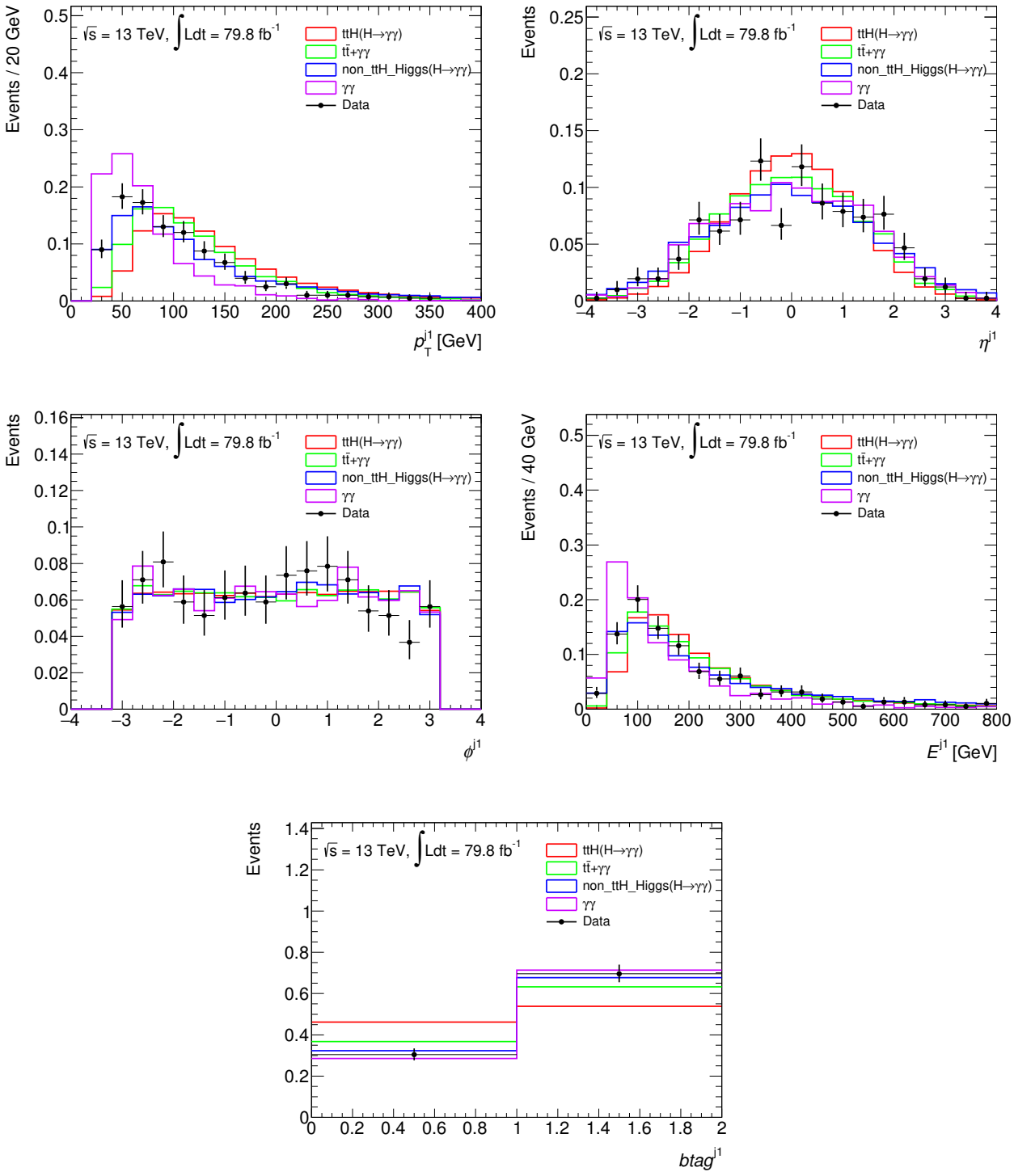


Figure B.12: Input variable distributions of the XGBoost for the Leptonic region related to the leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

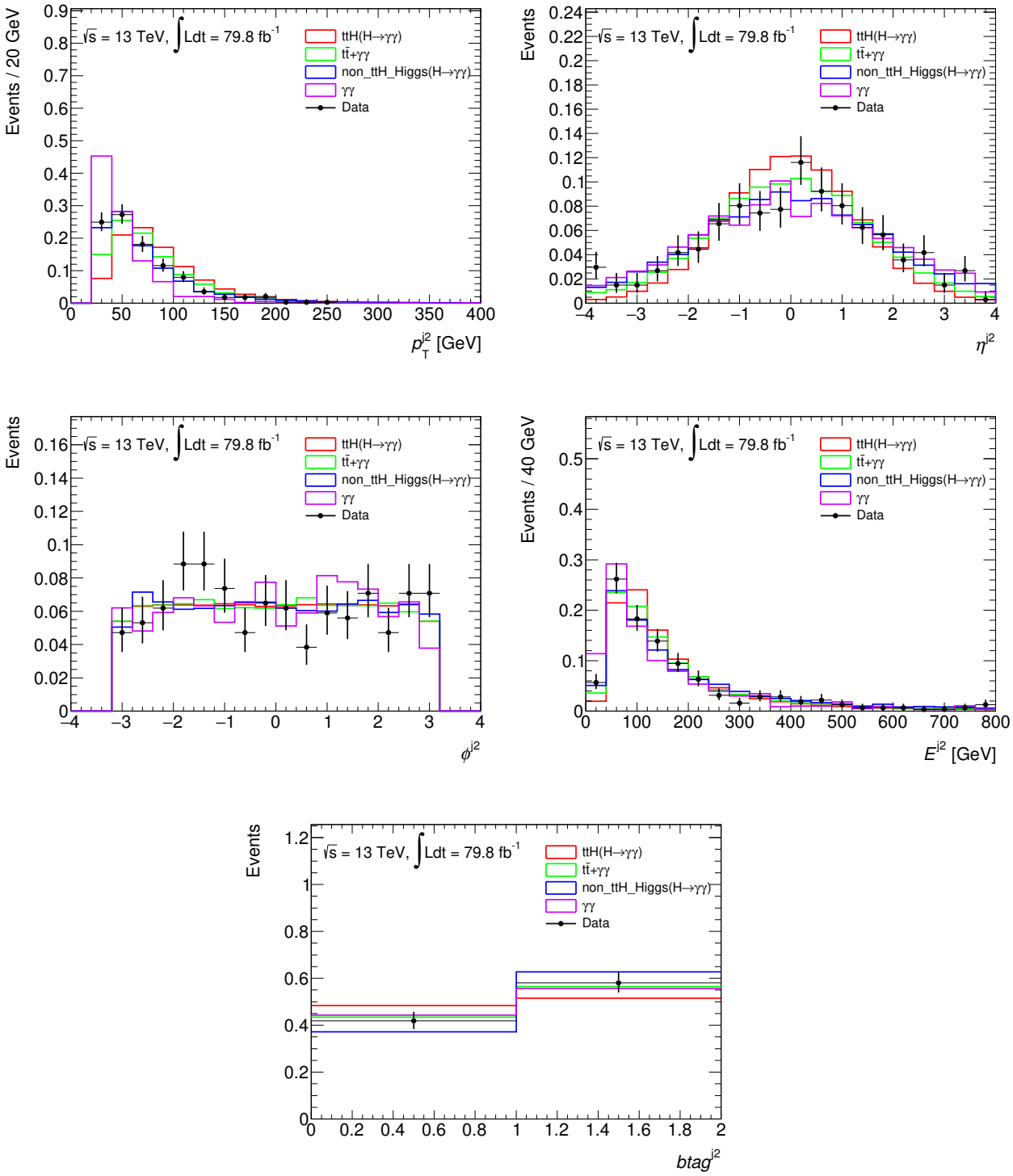


Figure B.13: Input variable distributions of the XGBoost for the Leptonic region related to the second leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

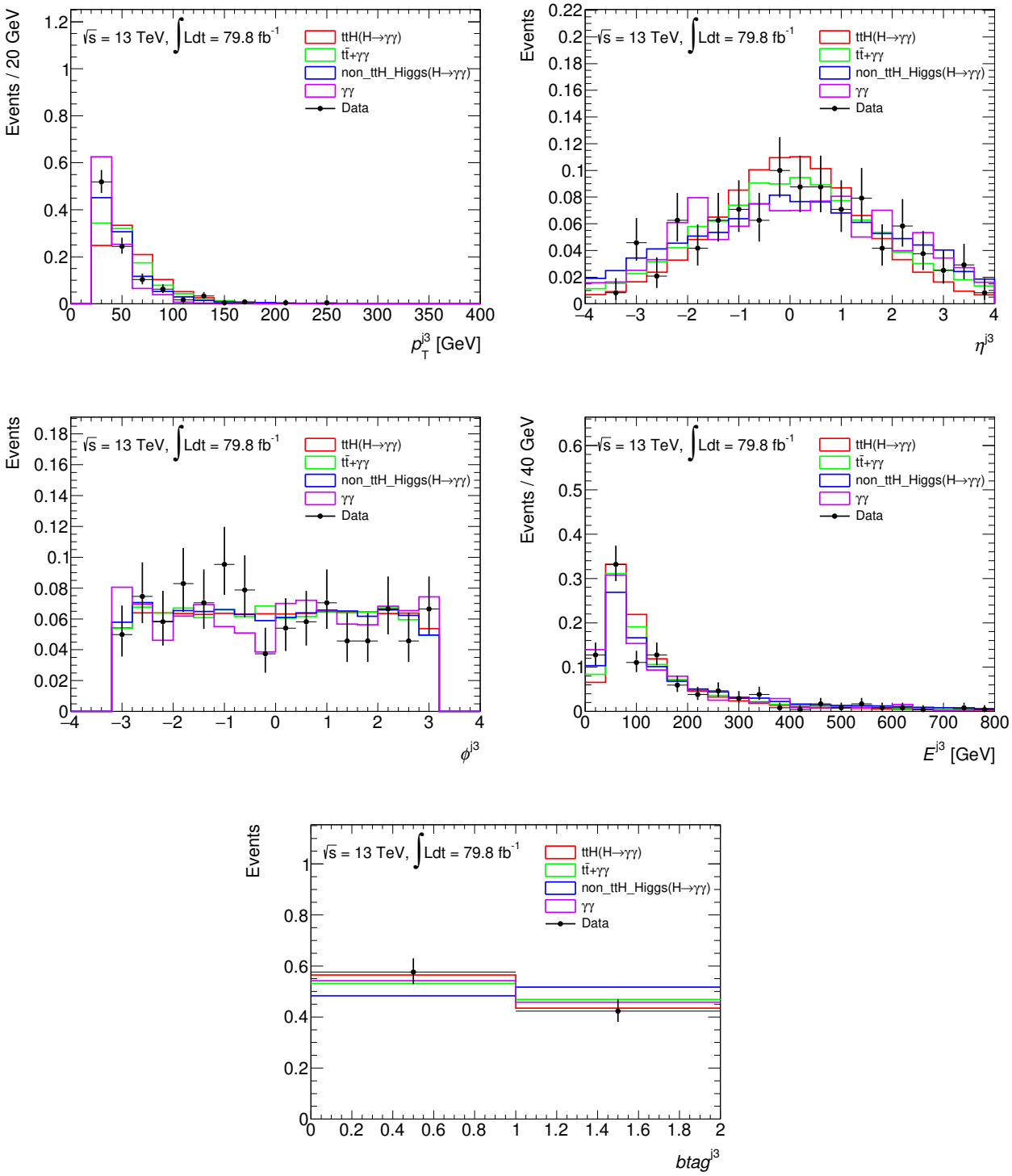


Figure B.14: Input variable distributions of the XGBoost for the Leptonic region related to the third leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

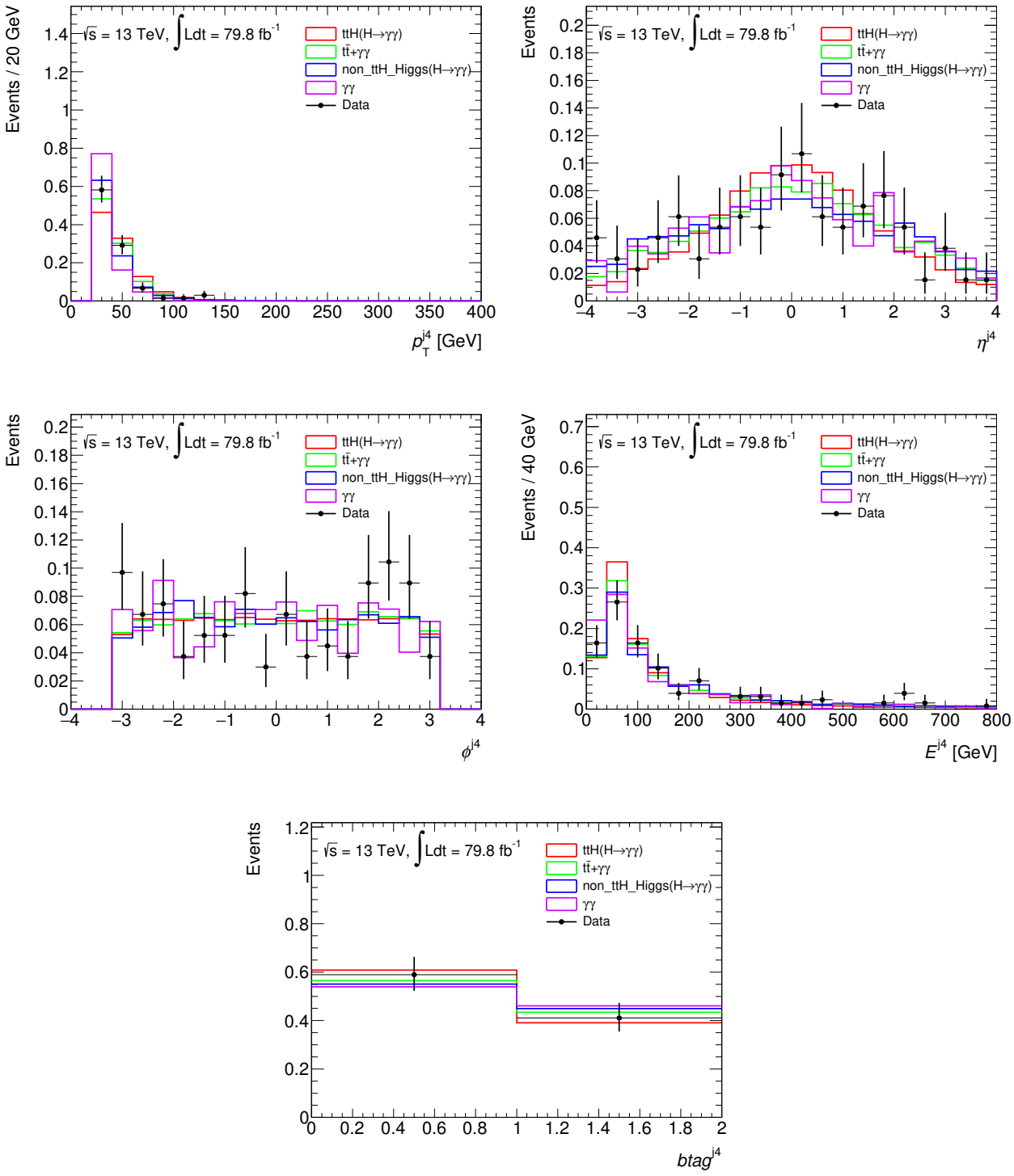


Figure B.15: Input variable distributions of the XGBoost for the Leptonic region related to the fourth leading jet after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

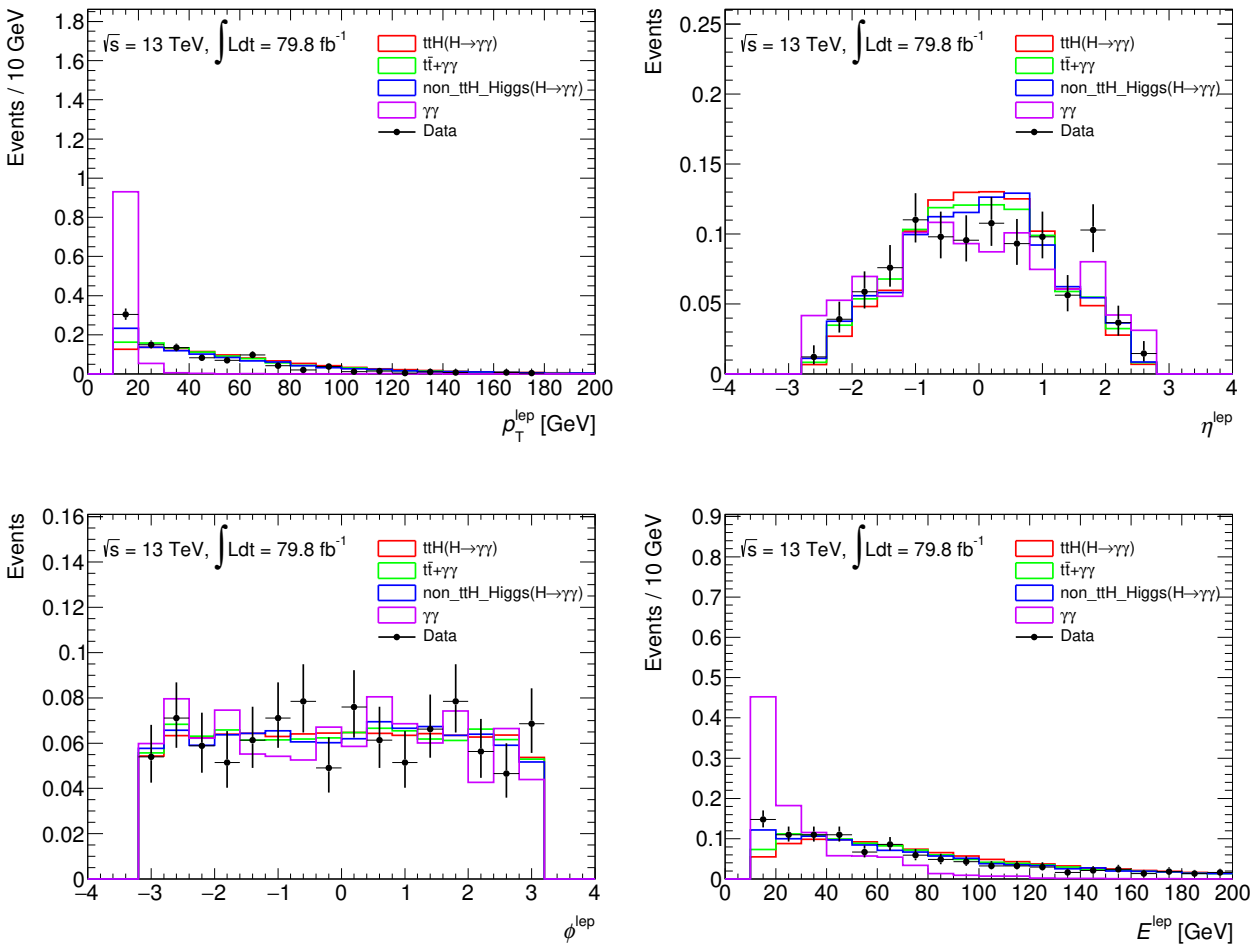


Figure B.16: Input variable distributions of the XGBoost for the Leptonic region related to the leading lepton after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

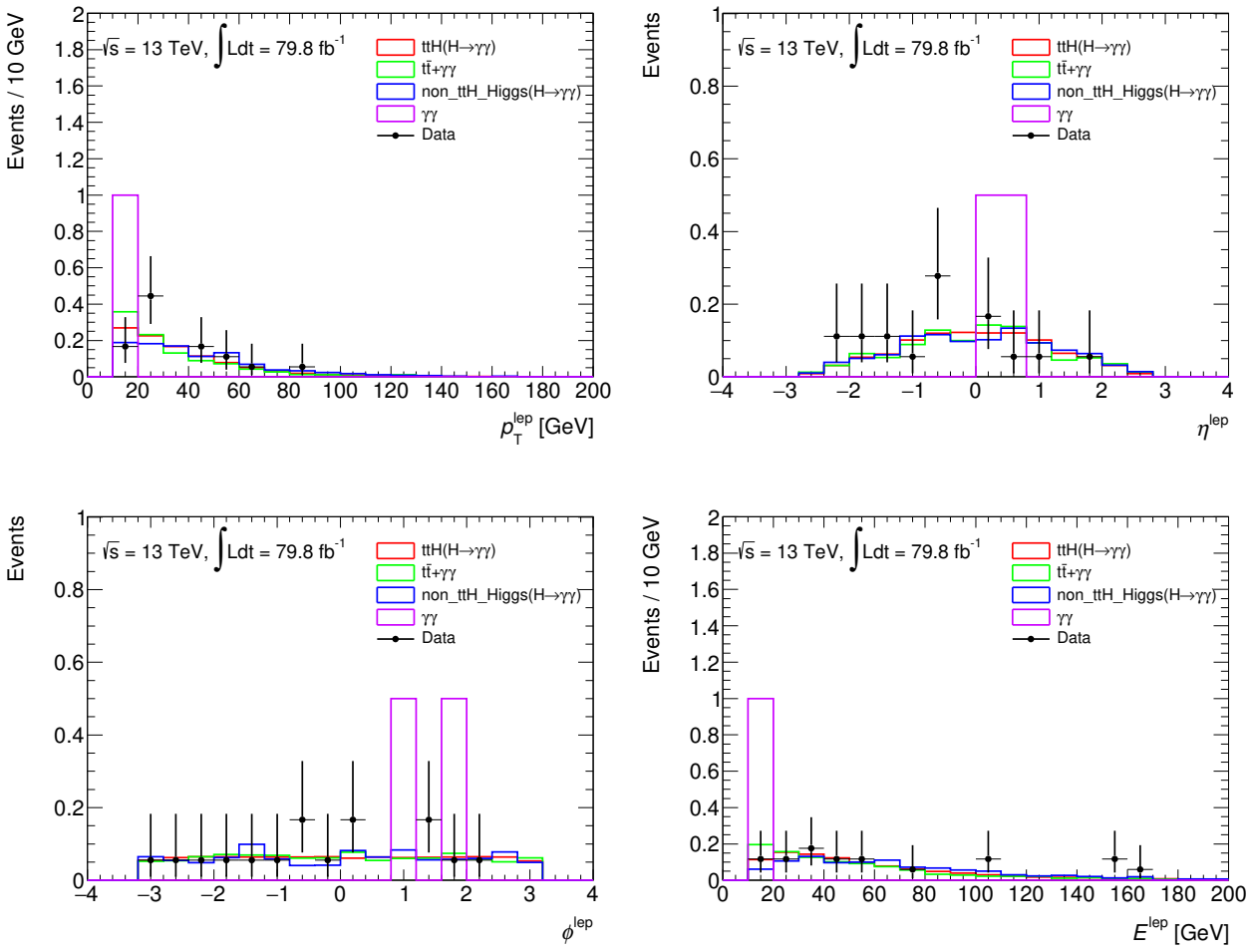


Figure B.17: Input variable distributions of the XGBoost for the Leptonic region related to the leading lepton after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

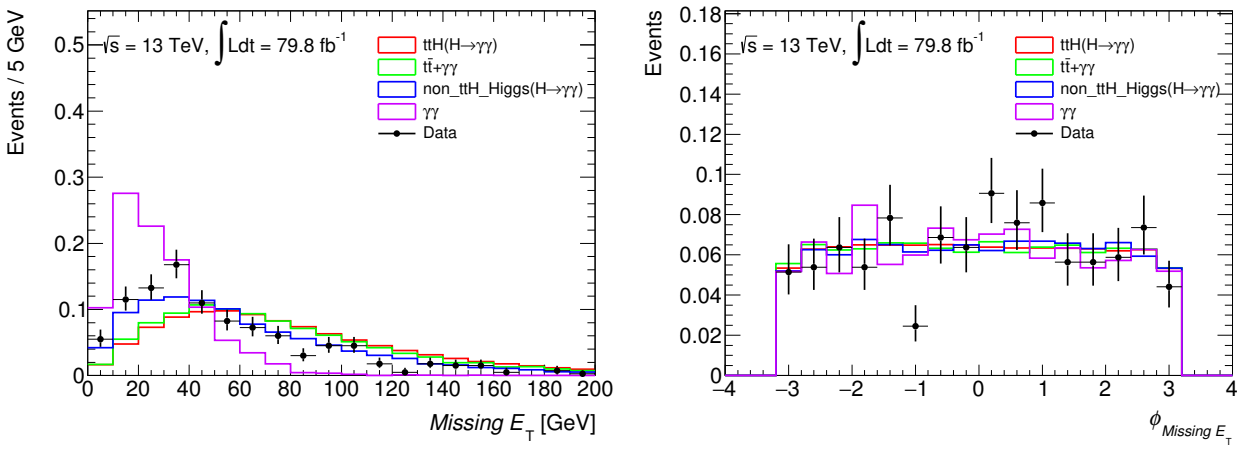


Figure B.18: Input variable distributions of the XGBoost for the Leptonic region related to the missing transverse energy E_T^{miss} after the di-photon and top-quark pair selections. The non-ttH Higgs denotes the mixture of the ggF, VBF, VH ($V = W^{\pm}, Z$), bbH, $ggZH$, tWH and $tHjb$ MC samples. Events are normalized to unity.

Appendix C

Continuum background modeling using data sideband

In this analysis, the continuum background shape is modeled using certain templates as discussed Section 6.2.3. The study with sideband data is also available although the statistics is very limited. The continuum background modeling using sideband data is performed as shown in Figure C.1 and C.2 for the Leptonic and Hadronic regions, respectively.

Figure C.3 shows the fitting parameter c_1 or c_2 for the three functions used to evaluate the background modeling uncertainty as described in Section 6.2.3. The background shapes retrieved from the templates are slightly different from that from the sideband data. However, it is hard to discuss the difference of the shapes between the templates and sideband data due to the large statistical uncertainty. For this reason, the shape of the templates and sideband is regarded as the same, and the templates are used for the continuum background modeling in this analysis.

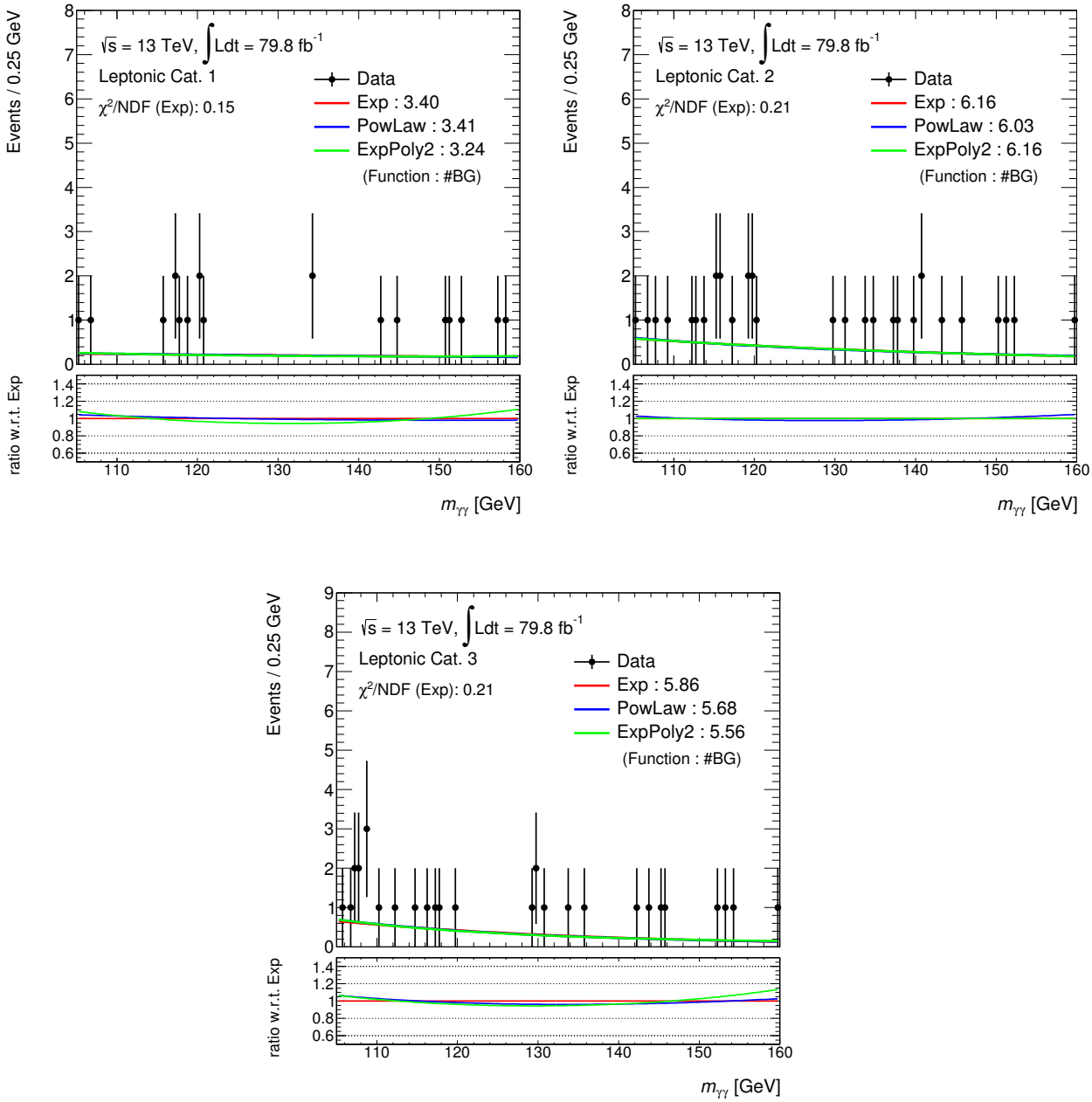


Figure C.1: Sideband data distribution for each category in the Leptonic region. The region $121 < m_{\gamma\gamma} < 129$ GeV is removed to avoid the contamination of the $H \to \gamma\gamma$ events. Data are fitted by the Exponential (Exp), Power law (PowLaw) and Exponential of second-order Polynomial (ExpPoly2) functions. The χ^2 divided by the number of degrees of freedom (NDF) for the exponential fitting is shown. The number of backgrounds is calculated by integral in $121 < m_{\gamma\gamma} < 129$ GeV.

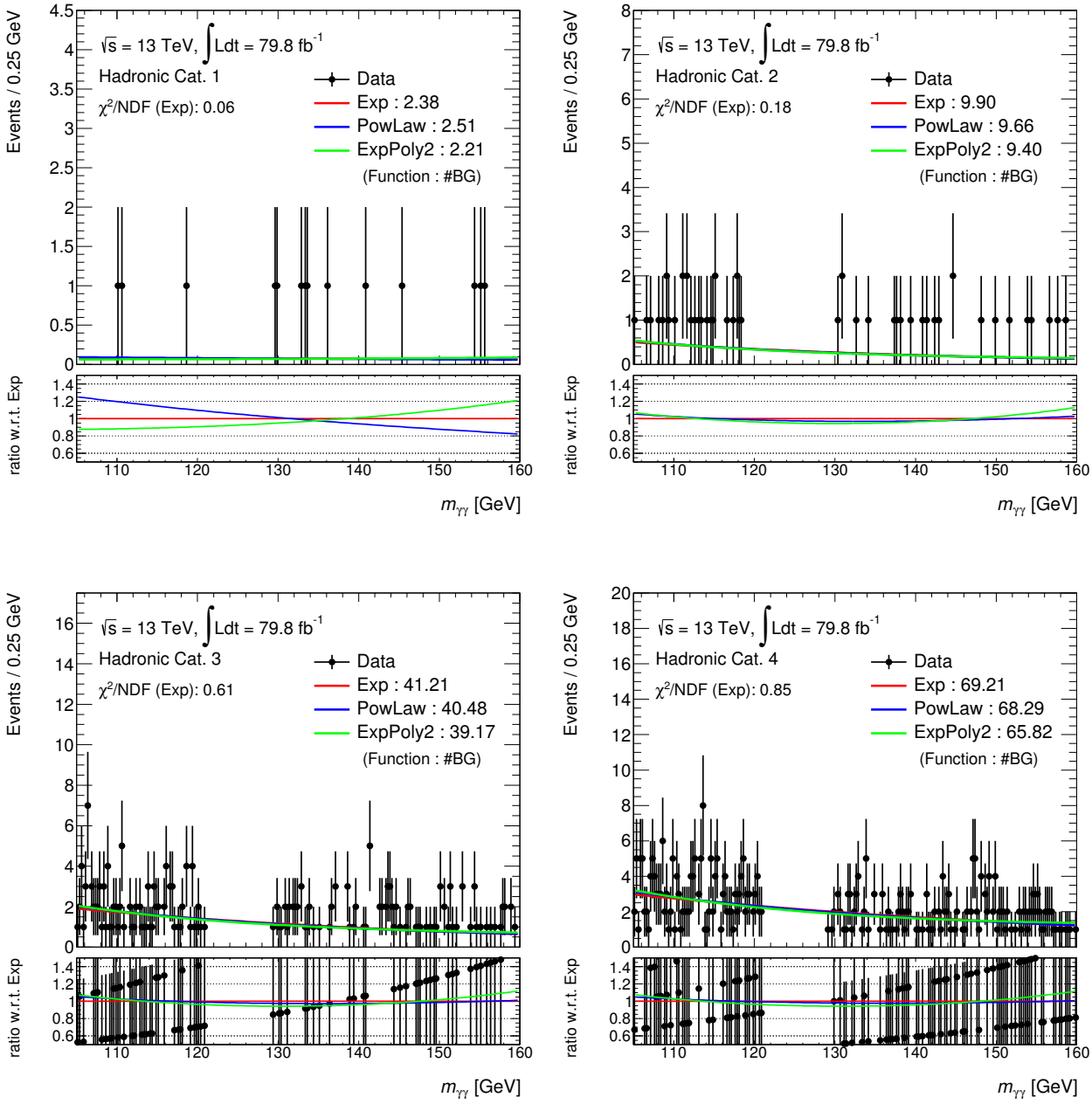


Figure C.2: Sideband data distribution for each category in the Hadronic region. The region $121 < m_{\gamma\gamma} < 129$ GeV is removed to avoid the contamination of the $H \to \gamma\gamma$ events. Data are fitted by the Exponential (Exp), Power law (PowLaw) and Exponential of second-order Polynomial (ExpPoly2) functions. The χ^2 divided by the number of degrees of freedom (NDF) for the exponential fitting is shown. The number of backgrounds is calculated by integral in $121 < m_{\gamma\gamma} < 129$ GeV.

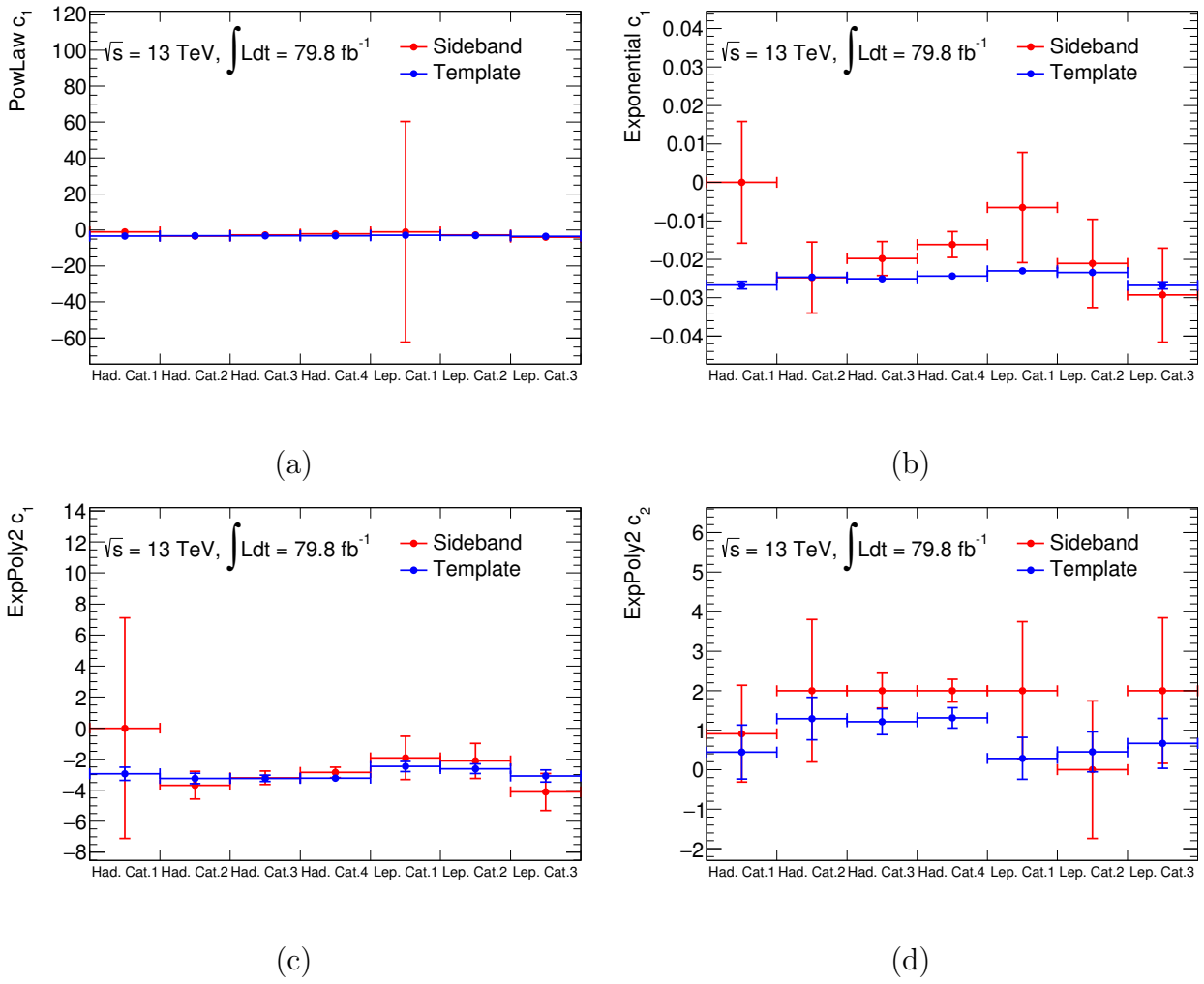


Figure C.3: Fitting parameters for the Power Law (a), the Exponential (b) and the Exponential of second-order Polynomial (c, d). The detail of these parameters are described in Section 6.2.3. The fitting is performed to the sideband data and templates.

Appendix D

Detail of the systematic uncertainties related to acceptance

The main sources of the acceptance uncertainty is discussed in Section 6.3.1. We discuss about the remaining uncertainties although there are small contributions to the ttH cross section measurement. Table D.1, D.2, D.3, D.4, D.5 and D.6 show the lists of the sources of acceptance uncertainty for the non-ttH Higgs backgrounds.

Uncertainty of the photon trigger efficiency

The di-photon trigger efficiency is measured by the multiplication of the single photon trigger efficiency as discussed in Section 4.2.2. The uncertainty of $^{+0.2}_{-1.1}$ % is considered.

Uncertainties related to leptons

The acceptance uncertainty related to electron and muon mainly comes from the selection and isolation efficiencies. The small contribution is appeared only in the Leptonic region.

Uncertainties related to jets

The uncertainties of jet energy scale, jet energy resolution and JVT efficiencies also affect the acceptance uncertainty. These measurements are mentioned in Section 4.5.

Uncertainties of the b-tagging efficiency

The b-tagging efficiency and its uncertainty are measured with data using $t\bar{t}$ di-lepton channel as discussed in Section 4.6. The dominant source of the uncertainty is modeling of $t\bar{t}$ MC, such as the generator and parton shower modeling.

Uncertainties of the missing transverse energy

The uncertainty of the missing transverse energy calculated as Equation 4.10 comes from the energy scale and resolution of hard and soft terms. The uncertainty of the hard term, consisting of the energy of photons, electrons, muons and jets, is already considered above.

Heavy flavor jet modeling uncertainties

The theoretical uncertainty of the additional production of b-jets via gluon radiation and splitting into b-quark pair is assigned to the acceptance of non-ttH Higgs background. This uncertainty is only assigned to the ggF, VBF, WH and ZH processes. The predicted acceptance of these processes is conservatively assigned with 100 % uncertainty. Since there are small contaminations by the non-ttH Higgs background in each category as shown in Table 5.8 and 5.7, the impact on the ttH cross section measurement is small.

Table D.1: Magnitude of the acceptance uncertainties for ggF process evaluated for each category and their impact on the ttH cross section.

Source of uncertainty	$\Delta N_{ggF}/N_{ggF}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat.1	Cat.2	Cat.3	Cat.1	Cat.2	Cat.3	Cat.4	
μ_r and μ_f in the QCD	18	17	24	36	8.9	21	19	1.2
PDF+α_S	10	5.3	17	5.9	6.3	8.6	11	0.5
Heavy flavor modeling	100	100	100	100	100	100	100	3.6
Photon trigger	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Photon $Tight$ selection	5.1	3.5	3.9	4.7	4.3	4.4	4.0	3.7
Photon isolation selection	4.3	4.3	4.1	4.6	5.5	5.2	4.6	5.1
Jet	44	167	16	15	22	15	14	5.7
Lepton	0.6	0.0	0.0	-	-	-	-	0.2
b-tagging	6.5	3.7	7.1	9.2	8.4	7.1	8.8	1.6
Missing transverse energy	7.1	0.0	0.0	2.0	1.4	2.2	2.2	0.4

Table D.2: Magnitude of the acceptance uncertainties for VBF process evaluated for each category and their impact on the ttH cross section.

Source of uncertainty	$\Delta N_{VBF}/N_{VBF}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat.1	Cat.2	Cat.3	Cat.1	Cat.2	Cat.3	Cat.4	
μ_r and μ_f in the QCD	-	-	6.7	2.3	1.7	1.8	1.5	0.0
PDF+α_S	-	-	4.5	3.4	1.9	2.1	3.3	0.0
Heavy flavor modeling	-	-	100	100	100	100	100	0.5
Photon trigger	-	-	0.7	0.7	0.7	0.7	0.7	0.7
Photon $Tight$ selection	-	-	2.9	4.9	4.5	4.2	4.1	3.7
Photon isolation selection	-	-	4.2	9.1	4.4	4.5	4.5	5.1
Jet	-	-	18	22	19	20	18	5.7
Lepton	-	-	0.0	-	-	-	-	0.2
b-tagging	-	-	0.0	0.0	0.0	0.0	0.0	1.6
Missing transverse energy	-	-	0.0	0.7	1.9	3.5	2.2	0.4

Table D.3: Magnitude of the acceptance uncertainties for WH process evaluated for each category and their impact on the ttH cross section.

Source of uncertainty	$\Delta N_{WH}/N_{WH}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat.1	Cat.2	Cat.3	Cat.1	Cat.2	Cat.3	Cat.4	
μ_r and μ_f in the QCD	4.2	4.5	3.5	5.0	4.2	3.8	3.3	0.1
PDF+α_S	3.0	4.1	3.8	3.2	1.5	1.8	0.7	0.0
Heavy flavor modeling	100	100	100	100	100	100	100	1.2
Photon trigger	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Photon *Tight* selection	4.0	4.0	4.1	4.6	4.1	4.2	4.0	3.7
Photon isolation selection	4.3	4.4	4.2	4.5	4.4	4.1	4.5	5.1
Jet	16	8.3	10	18	23	16	20	5.7
Lepton	0.4	0.4	0.3	-	-	-	-	0.2
b-tagging	10	11	13	8.2	7.8	5.6	5.6	1.6
Missing transverse energy	1.1	0.7	1.2	1.6	1.1	2.8	2.3	0.4

Table D.4: Magnitude of the acceptance uncertainties for ZH process evaluated for each category and their impact on the ttH cross section.

Source of uncertainty	$\Delta N_{ZH}/N_{ZH}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat.1	Cat.2	Cat.3	Cat.1	Cat.2	Cat.3	Cat.4	
μ_r and μ_f in the QCD	5.1	4.6	4.0	3.7	4.0	3.4	2.0	0.1
PDF+α_S	3.8	3.0	3.4	2.8	1.9	1.5	0.8	0.0
Heavy flavor modeling	100	100	100	100	100	100	100	1.2
Photon trigger	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Photon *Tight* selection	4.4	4.0	3.8	4.6	4.0	4.1	4.0	3.7
Photon isolation selection	4.4	4.8	4.2	4.7	4.3	4.7	4.4	5.1
Jet	19	27	12	17	16	16	15	5.7
Lepton	1.0	2.0	3.1	-	-	-	-	0.2
b-tagging	6.4	7.6	9.0	3.4	3.7	3.5	5.4	1.6
Missing transverse energy	2.6	7.8	13	1.0	0.9	1.8	2.2	0.4

Table D.5: Magnitude of the acceptance uncertainties for $tHjb$ process evaluated for each category and their impact on the ttH cross section.

Source of uncertainty	$\Delta N_{tHjb}/N_{tHjb}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat.1	Cat.2	Cat.3	Cat.1	Cat.2	Cat.3	Cat.4	
Photon trigger	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Photon *Tight* selection	4.4	3.9	3.9	4.8	4.3	4.1	3.9	3.7
Photon isolation selection	4.9	4.4	4.2	5.0	4.8	4.7	4.6	5.1
Jet	11	4.6	2.4	13	11	13	9.1	5.7
Lepton	0.4	0.3	0.4	-	-	-	-	0.2
b-tagging	1.2	0.8	0.6	0.9	0.9	1.1	1.2	1.6
Missing transverse energy	0.5	0.3	0.6	0.7	0.9	0.8	1.0	0.4

Table D.6: Magnitude of the acceptance uncertainties for tWH process evaluated for each category and their impact on the ttH cross section.

Source of uncertainty	$\Delta N_{tWH}/N_{tWH}$ [%]							$\Delta\sigma_{ttH}/\sigma_{ttH}$ [%]
	Leptonic region			Hadronic region				
	Cat.1	Cat.2	Cat.3	Cat.1	Cat.2	Cat.3	Cat.4	
Photon trigger	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Photon *Tight* selection	4.5	3.8	3.5	5.0	4.6	4.2	3.9	3.7
Photon isolation selection	8.7	6.0	4.9	9.2	9.8	7.1	4.3	5.1
Jet	4.6	2.8	11	6.8	6.4	4.3	4.7	5.7
Lepton	0.8	0.6	0.4	-	-	-	-	0.2
b-tagging	2.7	2.2	1.5	1.6	1.4	1.2	1.9	1.6
Missing transverse energy	1.8	0.9	0.2	0.8	0.3	1.2	0.6	0.4

Appendix E

Detail of the validation of the fitting

Figure E.1, E.2, E.3 and E.4 show the distributions of the signal strength μ^{output} and its pull obtained by fitting pseudo datasets generated by toy MC for various μ^{input}. There are tails at the lower side in the distribution of μ^{output} and its pull, especially for the ones for $\mu^{input} < 1$. It is caused by the lack of events around the Higgs mass due to the statistical fluctuation and consequently μ^{output} can be much small value ($\mu^{output} < 0$) although it is not allowed by physics. To avoid such an influence, the range of gaussian fitting of the pull distribution is limited from -2 to 2 in this analysis. We need to consider how to treat the case of $\mu^{output} < 0$.

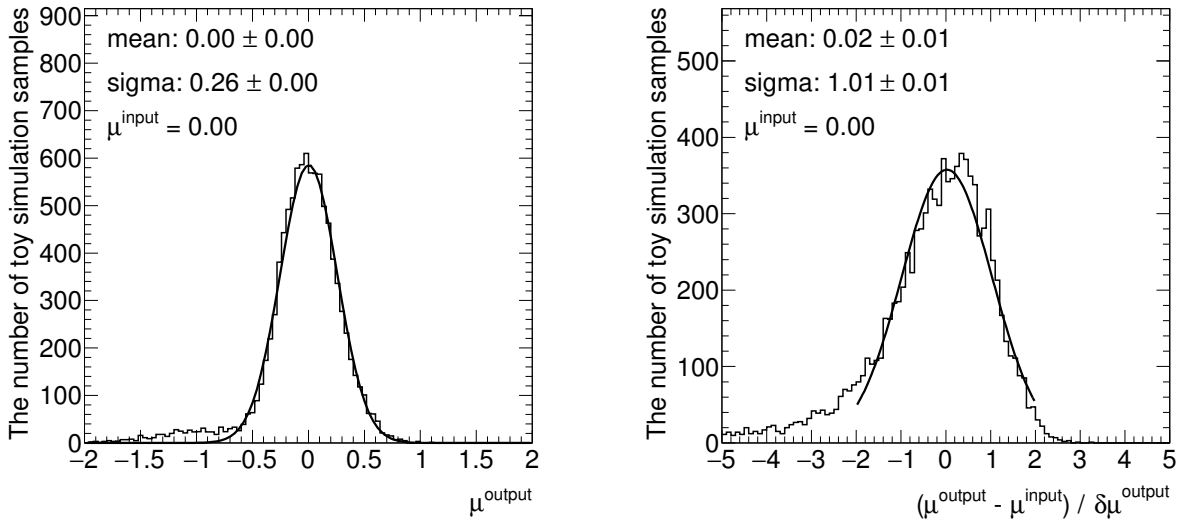


Figure E.1: Distribution of the signal strength μ^{output} (left) and its pull (right) obtained by fitting pseudo datasets generated by toy MC. The input signal strength is assumed to be $\mu^{input} = 0$ in the generation of pseudo data.

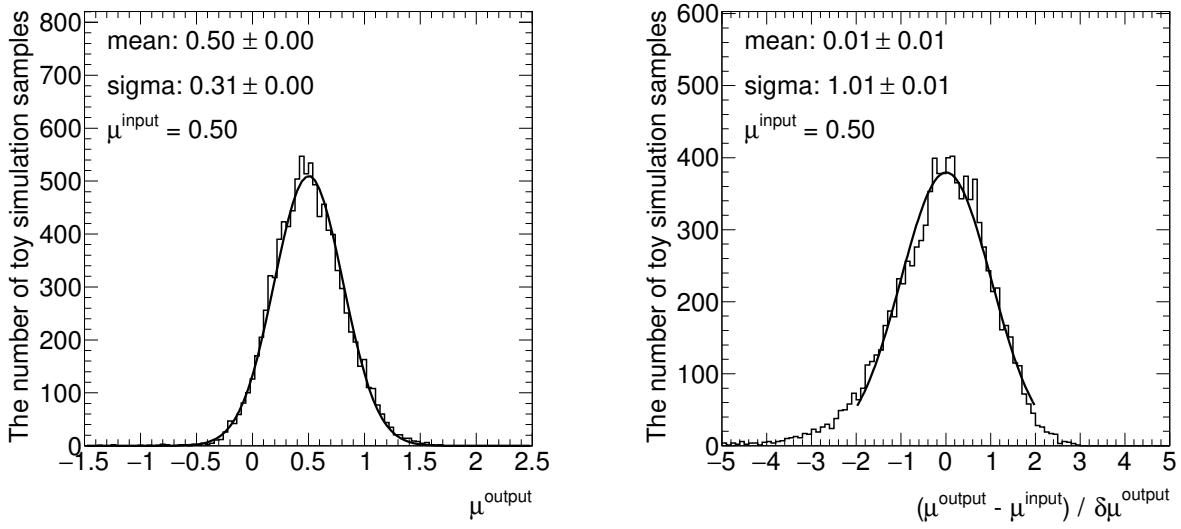


Figure E.2: Distribution of the signal strength μ^{output} (left) and its pull (right) obtained by fitting pseudo datasets generated by toy MC. The input signal strength is assumed to be $\mu^{input} = 0.5$ in the generation of pseudo data.

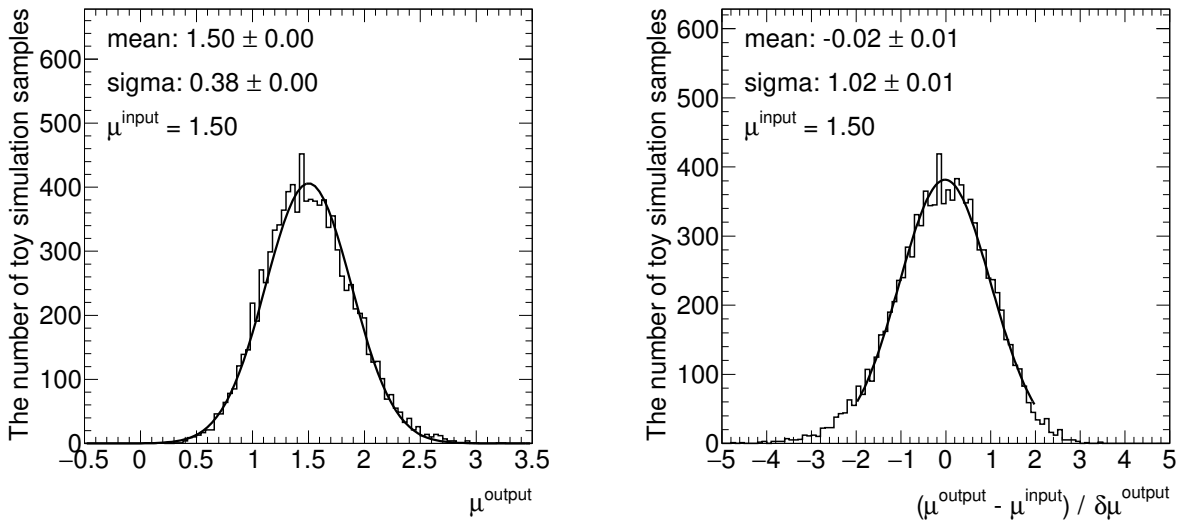


Figure E.3: Distribution of the signal strength μ^{output} (left) and its pull (right) obtained by fitting pseudo datasets generated by toy MC. The input signal strength is assumed to be $\mu^{input} = 1.5$ in the generation of pseudo data.

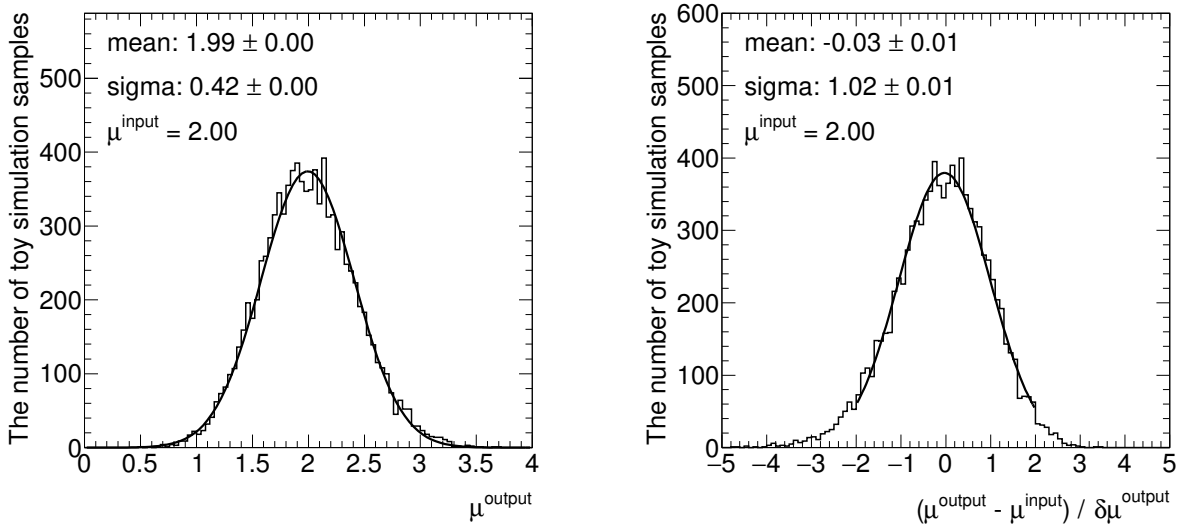


Figure E.4: Distribution of the signal strength μ^{output} (left) and its pull (right) obtained by fitting pseudo datasets generated by toy MC. The input signal strength is assumed to be $\mu^{input} = 2$ in the generation of pseudo data.

Acknowledgement

First of all, I would like to express the deepest appreciation to my supervisor, Prof. Kazunori Hanagaki, for giving me a great opportunity to study the particle physics again in SOKENDAI. His enthusiastic teaching and support always encourage me. I could not make such a productive research life without his supports. I also would like to show my appreciation to Prof. Taku Yamanaka, who encourage me to come back to study physics.

I am deeply grateful to Y. Enari, K. Nakamura, Y. Okumura and Y. Yamaguchi for their productive advices to my research. I would have had a lot of difficulties without their helps. I also really enjoyed discussing about many physics topics with them.

I appreciate the collaborators of the $H \to \gamma\gamma$ analysis group. Especially, the conveners of the group, M. Delmastro and H. Wang, and co-analyzer, C. Zhou give me a lot of knowledges about the technical procedure of the physics analysis.

My deepest appreciation goes to the ATLAS Pixel working group; D. Ferrer, M. Kocian, Y. Takubo, S. Tsuno, H. Oide and K. Lantzsch. They gave me many chances to study about the Pixel detector on the front line. Especially, I really would like to thank Y. Takubo, S. Tsuno and H. Oide for their warm support for my first year after moving to CERN.

I received generous support from the KEK ATLAS group, M. Togawa, K. Nagano, M. Aoki, S. Suzuki, T. Honda, J. Usui, A. Mizukami, A. Furuichi and S. Midorikawa. I also would like to offer my special thanks to Y. Honda, who kindly supported me to make comfortable research environment.

I really enjoyed my research life at CERN in the past four years thanks to a lot of my colleagues, H. Otono, M. Hirose, K. Mochizuki, S. Hirose, J.J. Teoh, N. Ishijima, K. Todome, S. Honda, S. Shirabe, D. Yamaguchi and K. Yajima. There are a lot of great experiences at CERN with them, in my memories. I am really happy person to have best buddies.

Finally, I would like to express my gratitude to my family for their support.

Reference

[1] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[2] Serguei Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[3] The ATLAS Collaboration et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (2008), S08003. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[4] The CMS Collaboration et al. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (2008), S08004. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[5] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[6] "First combination of Tevatron and LHC measurements of the top-quark mass". In: (2014). arXiv: 1403.4427 [hep-ex].

[7] Timo A. Aaltonen. "Combination of CDF and DO Results on the Mass of the Top Quark Using up to 8.7 fb^{-1} at the Tevatron". In: (2013). arXiv: 1305.3929 [hep-ex].

[8] The ATLAS collaboration. "Combination of ATLAS and CMS results on the mass of the top-quark using up to 4.9 fb^{-1} of $\sqrt{s} = 7$ TeV LHC data". In: (2013).

[9] S. L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[10] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[11] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc.* C680519 (1968), pp. 367–377.

[12] Gerard 't Hooft and M. J. G. Veltman. "Regularization and Renormalization of Gauge Fields". In: *Nucl. Phys.* B44 (1972), pp. 189–213. DOI: 10.1016/0550-3213(72)90279-9.

[13] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). [,157(1964)], pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[14] Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[15] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). [,160(1964)], pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[16] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (1964). [,162(1964)], pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[17] Peter W. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons". In: *Phys. Rev.* 145 (1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.

[18] T. W. B. Kibble. "Symmetry breaking in nonAbelian gauge theories". In: *Phys. Rev.* 155 (1967). [,165(1967)], pp. 1554–1561. DOI: 10.1103/PhysRev.155.1554.

[19] 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. arXiv: 1610.07922 [hep-ph].

[20] Morad Aaboud et al. "Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector". In: *Phys. Rev.* D97.7 (2018), p. 072003. DOI: 10.1103/PhysRevD.97.072003. arXiv: 1712.08891 [hep-ex].

[21] Morad Aaboud et al. "Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* D97.7 (2018), p. 072016. DOI: 10.1103/PhysRevD.97.072016. arXiv: 1712.08895 [hep-ex].

[22] Morad Aaboud et al. "Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* D98 (2018), p. 052005. DOI: 10.1103/PhysRevD.98.052005. arXiv: 1802.04146 [hep-ex].

[23] Morad Aaboud et al. "Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 03 (2018), p. 095. DOI: 10.1007/JHEP03(2018)095. arXiv: 1712.02304 [hep-ex].

[24] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (2008), S08001. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[25] Public web site. "LHC Images." In: (). General Photo. URL: https://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/lhc_in_pictures.htm.

[26] Fabienne Marcastel. "CERN's Accelerator Complex." In: (Oct. 2013). General Photo. URL: https://cds.cern.ch/record/1621583.

[27] The ATLAS Collaboration. *ATLAS Luminosity Public Results*. Tech. rep. Geneva: CERN. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsR

[28] Nikolina Ilic on behalf of the ATLAS Collaboration. "ATLAS Status Report". In: 132nd LHCC Meeting, 2017. URL: https://indico.cern.ch/event/679087/.

[29] Karolos Potamianos. "The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider". In: *PoS* EPS-HEP2015 (2015), p. 261. arXiv: 1608.07850 [physics.ins-det].

[30] M. Capeans et al. "ATLAS Insertable B-Layer Technical Design Report". In: (2010).

[31] E. Abat et al. "The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance". In: *JINST* 3 (2008), P02013. DOI: 10.1088/1748-0221/3/02/P02013.

[32] *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331063.

[33] *ATLAS Photos*. URL: http://atlasexperiment.org/photos/magnets-solenoid.html.

[34] *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: http://cds.cern.ch/record/331062.

[35] The ATLAS collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC". In: (2019).

[36] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *Journal of Instrumentation* 13.07 (2018), P07017. URL: http://stacks.iop.org/1748-0221/13/i=07/a=P07017.

[37] Georges Aad et al. "Performance of the ATLAS Trigger System in 2010". In: *Eur. Phys. J.* C72 (2012), p. 1849. DOI: 10.1140/epjc/s10052-011-1849-1. arXiv: 1110.1530 [hep-ex].

[38] Morad Aaboud et al. "Performance of the ATLAS Trigger System in 2015". In: *Eur. Phys. J.* C77.5 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex].

[39] G. Aad et al. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J.* C70 (2010), pp. 823–874. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[40] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[41] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [hep-ph].

[42] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[43] Richard D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[44] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[45] 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: *Journal of High Energy Physics* 2014.7 (July 17, 2014), p. 79. ISSN: 1029-8479. DOI: 10.1007/JHEP07(2014)079. URL: https://doi.org/10.1007/JHEP07(2014)079.

[46] 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: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[47] M. Bahr et al. "Herwig++ Physics and Manual". In: *Eur. Phys. J.* C58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[48] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[49] *Simulation of top quark production for the ATLAS experiment at sqrt(s) = 13 TeV*. Tech. rep. ATL-PHYS-PUB-2016-004. Geneva: CERN, Jan. 2016. URL: https://cds.cern.ch/record/2120417.

[50] *Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run2*. Tech. rep. ATL-PHYS-PUB-2018-009. Geneva: CERN, July 2018. URL: http://cds.cern.ch/record/2630327.

[51] T. Cornelissen et al. "Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)". In: (2007). Ed. by A. Salzburger.

[52] R. Fruhwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nucl. Instrum. Meth.* A262 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.

[53] Morad Aaboud et al. "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC". In: *Eur. Phys. J.* C77.5 (2017), p. 332. DOI: 10.1140/epjc/s10052-017-4887-5. arXiv: 1611.10235 [physics.ins-det].

[54] R. Fruhwirth, W. Waltenberger, and P. Vanlaer. "Adaptive vertex fitting". In: *J. Phys.* G34 (2007), N343. DOI: 10.1088/0954-3899/34/12/N01.

[55] Georges Aad et al. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J.* C77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[56] *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*. Tech. rep. ATL-PHYS-PUB-2017-022. Geneva: CERN, Dec. 2017. URL: http://cds.cern.ch/record/2298955.

[57] Morad Aaboud et al. "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016". In: *Submitted to: Eur. Phys. J.* (2018). arXiv: 1810.05087 [hep-ex].

[58] M. Aharrouche et al. "Response uniformity of the ATLAS liquid argon electromagnetic calorimeter". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 582.2 (2007), pp. 429–455. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2007.08.157. URL: http://www.sciencedirect.com/science/article/pii/S0168900207018591.

[59] Georges Aad et al. "Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data". In: (2019). arXiv: 1908.00005 [hep-ex].

[60] Luc Devroye. *Non-Uniform Random Variate Generation*. New York, NY, USA: Springer-Verlag, 1986.

[61] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, Apr. 2008. URL: https://cds.cern.ch/record/1099735.

[62] Morad Aaboud et al. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV". In: *Submitted to: Eur. Phys. J.* (2019). arXiv: 1902.04655 [physics.ins-det].

[63] Georges Aad et al. "Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[64] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[65] *Tagging and suppression of pileup jets with the ATLAS detector*. Tech. rep. ATLAS-CONF-2014-018. Geneva: CERN, May 2014. URL: https://cds.cern.ch/record/1700870.

[66] ATLAS Collaboration 2016. *JVT Public Plots for ICHEP 2016*. Tech. rep. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2016-011/.

[67] M. Aaboud et al. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* D96.7 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[68] ATLAS Collaboration 2018. *Jet energy scale and uncertainties in 2015-2017 data and simulation*. Tech. rep. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2018-006/.

[69] *Expected performance of the ATLAS b-tagging algorithms in Run-2*. Tech. rep. ATL-PHYS-PUB-2015-022. Geneva: CERN, July 2015. URL: http://cds.cern.ch/record/2037697.

[70] G. Aad et al. "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics". In: (2009). arXiv: 0901.0512 [hep-ex].

[71] *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. ATL-PHYS-PUB-2017-013. Geneva: CERN, July 2017. URL: http://cds.cern.ch/record/2273281.

[72] Morad Aaboud et al. "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV". In: *JHEP* 08 (2018), p. 089. DOI: 10.1007/JHEP08(2018)089. arXiv: 1805.01845 [hep-ex].

[73] Georges Aad et al. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV". In: (2019). arXiv: 1907.05120 [hep-ex].

[74] Morad Aaboud et al. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J.* C78.11 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[75] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an].

[76] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. DOI: 10.1145/2939672.2939785. URL: http://doi.acm.org/10.1145/2939672.2939785.

[77] Jerome H. Friedman. "Greedy function approximation: A gradient boosting machine." In: *Ann. Statist.* 29.5 (Oct. 2001), pp. 1189–1232. DOI: 10.1214/aos/1013203451. URL: https://doi.org/10.1214/aos/1013203451.

[78] Georges Aad et al. "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments". In: *Phys. Rev. Lett.* 114 (2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex].

[79] CMS Collaboration. "Measurement of the associated production of a Higgs boson and a pair of top-antitop quarks with the Higgs boson decaying to two photons in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: (2018).

[80] T. D. Lee. "A Theory of Spontaneous T Violation". In: *Phys. Rev.* D8 (1973). [,516(1973)], pp. 1226–1239. DOI: 10.1103/PhysRevD.8.1226.

[81] G. C. Branco et al. "Theory and phenomenology of two-Higgs-doublet models". In: *Phys. Rept.* 516 (2012), pp. 1–102. DOI: 10.1016/j.physrep.2012.02.002. arXiv: 1106.0034 [hep-ph].

[82] *Further studies on simulation of top-quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV*. Tech. rep. ATL-PHYS-PUB-2016-016. Geneva: CERN, Aug. 2016. URL: https://cds.cern.ch/record/2205262.

[83] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J.* C76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[84] Georges Aad et al. "Measurement of the $t\bar{t}$ production cross-section as a function of jet multiplicity and jet transverse momentum in 7 TeV proton-proton collisions with the ATLAS detector". In: *JHEP* 01 (2015), p. 020. DOI: 10.1007/JHEP01(2015)020. arXiv: 1407.0891 [hep-ex].

[85] Morad Aaboud et al. "Study of the material of the ATLAS inner detector for Run 2 of the LHC". In: *JINST* 12.12 (2017), P12009. DOI: 10.1088/1748-0221/12/12/P12009. arXiv: 1707.02826 [hep-ex].