Search for the dimuon decay of the Higgs boson in 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

ATLAS 実験における重心系エネルギー 13 TeV の陽子陽子衝突データ 139 fb^{-1} を用いた μ 粒子対に崩壊するヒッグス粒子の探索

Tomomi Kawaguchi

High Energy Physics Laboratory

Graduate School of Science, Nagoya University

February 21, 2021

Abstract

The Standard Model (SM) of particle physics explains various phenomena observed so far. However, there are phenomena of the Higgs boson which are predicted in the SM but not yet confirmed experimentally. One of the Higgs boson's phenomena not yet well verified is the Yukawa interaction. It is assumed that "charged fermions in all generations obtain mass proportional to each coupling strength through Yukawa interaction with the Higgs field". Couplings of the Higgs boson to third-generation charged fermions have been observed. On the other hand, the couplings to first- and second-generation fermions have not been observed yet.

This thesis presents a search for a dimuon decay of the Higgs boson. This decay contains the Yukawa interaction of the Higgs boson and the muon. The muon is one of the second-generation fermions in the SM. The dimuon decay of the Higgs boson provides a unique probe of the coupling between the Higgs boson and the second-generation fermions thanks to a more clean signature of the muons compared to the charm and strange quarks. Search for this decay is very challenging due to the huge $Z \to \mu\mu$ background events, and this decay has not been observed in previous studies using $\sqrt{s} = 13$ TeV pp collisions data collected in 2015 and 2016 at the Large Hadron Collider (LHC).

This analysis uses $\sqrt{s} = 13$ TeV pp collisions data collected with the ATLAS detector at LHC in 2015–2018. The integrated luminosity is 139 fb^{-1}. The dimuon invariant mass is reconstructed and a peak around 125 GeV (the Higgs boson mass) above the irreducible background is searched for. In order to increase the sensitivity of the search, the improvements in the invariant mass resolution and the separation of the signal events from the background events are crucial. To improve the mass resolution, the photon from the QED final state radiation (FSR) of a muon is taken into account. In the events with the FSR, the photon carries away the energy of the muon, resulting in a smaller dimuon invariant mass. By reconstructing FSR photons and including the four-momentum in the invariant mass calculation, the number of signal events in the invariant mass region 120–130 GeV was increased by 1.4% and the mass resolution was improved by 2.8%. To separate signal from background, the events are divided into 20 categories, based on the different final state particles in each Higgs production process, using multivariate analysis techniques. The invariant mass distribution of the background events is modeled using the Leading Order Drell-Yan analytic line-shape multiplied by empirical functions to account for imperfect modeling by the line shape.

The observed signal significance is 2.0 standard deviations with respect to the background-only hypothesis. The signal strength, defined as the ratio of the observed signal yield to the expected one in the SM, is 1.2 ± 0.6. The result is consistent with the SM within the uncertainty. The CMS experiment reported an observed signal significance of 3 standard deviations. It is compatible with the result described in this thesis. Both results indicate that the origin of the muon mass arises from the Higgs mechanism. The results constitute the first evidence of the coupling of the Higgs boson to second-generation fermions. At present, the measurement of the signal strength is dominated by the statistical uncertainty. I expect that the Yukawa coupling to the muon is verified using higher statistics at LHC in the near future.

Contents

Chapter 1

Introduction

In 2012, the Higgs boson was discovered [1, 2] at the ATLAS [3] and the CMS experiments [4] and all the particles predicted in the Standard Model (SM) are observed. The SM predicts that "the fermions in all generations obtain mass proportional to each coupling strength through Yukawa interaction with the Higgs field". In the SM, the Higgs boson has a unique property, and the difference in fermion mass, which differs by six orders depending on the generation, is explained only by the difference in the Yukawa coupling constant. In order to elucidate this unnaturalness, the measurement of coupling strength of the Higgs boson and the fermions is crucial. After the Higgs boson discovery, we have verified the Higgs boson properties with further accumulating the datasets. As a result, the Higgs boson couplings to the gauge bosons (W, Z) and the third-generation charged fermions (t, b, τ) have been observed. The results are consistent with the SM. However, the Higgs boson couplings to the first- and second- generation fermions have not been observed yet. The universality of the coupling mechanism between the generations of fermions is not confirmed yet. For the confirmation, a search for the Higgs boson coupling to the second-generation fermions is important.

In this thesis, I report the search for the decay of the Higgs boson to muons. The Higgs boson decay to muons has larger sensitivity than those to other second-generation fermions at the LHC. The search allows us to verify the mass origin of second-generation fermions for the first time. A new physics model which can explain the muon $g-2$ anomaly [5] could enhance and say that the Higgs coupling strength to the muon. The difficulty of this search is a small $\mathrm{H} \to \mu\mu$ signal yield above a huge Drell-Yan (DY) background. The cross section of the $pp \to H \to \mu\mu$ at $\sqrt{s} = 13$ TeV is 0.012 pb while the cross section of the DY in the dimuon invariant mass of 120–130 GeV is approximately 6 pb. In the $\mathrm{H} \to \mu\mu$ search using $\sqrt{s} = 13$ TeV pp collisions data collected with the ATLAS detector at LHC in 2015 and 2016 corresponding to 36.1 fb^{-1}, no significant excess was observed. An upper limit on the cross section of the Higgs boson production times the $\mathrm{H} \to \mu\mu$ branching ratio was 3.0 times the SM prediction at the 95% confidence level. Total uncertainty of the previous result is dominated by data statistical uncertainty.

I performed the $\mathrm{H} \to \mu\mu$ search using $\sqrt{s} = 13$ TeV pp collisions data collected with the ATLAS detector at LHC in 2015–2018 corresponding to 139 fb^{-1}. The dimuon invariant mass is reconstructed and a peak around the Higgs boson mass 125 GeV is searched for above the

irreducible background. In order to improve the mass resolution, I focused on the events in which a muon emits a photon by the QED final state radiation (FSR). I developed a method to reconstruct the photons from FSR. In addition, in order to separate the signal from the background, the events are divided by the difference in the final state particles and further classified into 20 categories in total using the multivariate analysis. The invariant mass distribution of the background events is modeled by the Leading Order Drell-Yan analytic line-shape multiplied by empirical functions to account for imperfect modeling by the line shape.

This thesis is organized as follows.

Chapter 2 The Higgs mechanism and the Yukawa coupling are introduced. The results on the already observed Yukawa couplings are summarized.

Chapter 3 An overview of the LHC accelerator and the ALTAS detector is given.

Chapter 4 The analysis of the $H \to \mu\mu$ search is presented.

Chapter 5 The results are discussed with reference to the results of the CMS experiment. Future prospects of the search are also described.

Chapter 6 Conclusion is given.

Chapter 2

Motivation

2.1 Theoretical motivation

2.1.1 The Standard Model

The Standard Model (SM) is a theory of particle physics that represents our current best understanding of the elementary particles and their interactions. In the SM, there are 12 fermions that compose matters, four gauge bosons that mediate force, and a Higgs boson that is the quantum manifestation of the Higgs field, which gives mass to elementary particles. Fermions are further divided into three generations, a pair of quarks and a pair of leptons belong to each generation. For quarks, up (u) and down (d) quarks, charm (c) and strange (s) quarks, and top (t) and bottom (b) quarks are in the first, second, and third generations, respectively. For leptons, electron (e) and electron neutrino (ν_e), muon (μ) and muon neutrino (ν_μ), and tau (τ) and tau neutrino (ν_τ) are in the first, second, and third generations, respectively. The larger generation corresponds to the larger particle mass. These particles interact by three forces: the electromagnetic, weak, and strong forces. The electromagnetic force is mediated by the photon (γ). The weak force is mediated by the weak bosons (W and Z bosons). The strong force is mediated by the gluon (g). In 2012, the Higgs boson was discovered at the ATLAS and CMS experiments, and the observations of the SM particles were completed. The SM particles are summarized in Figure 2.1. The spin, the electric charge, the weak hyper charge, and the third component of the weak isospin for each particle are summarized in Table 2.1.

The interactions and the motions of the SM particles are described by the SM Lagrangian:

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{Gauge}} + \mathcal{L}_{\mathrm{Fermion}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}} \tag{2.1}$$

Here $\mathcal{L}_{\mathrm{Gauge}}$, $\mathcal{L}_{\mathrm{Fermion}}$, $\mathcal{L}_{\mathrm{Higgs}}$, and $\mathcal{L}_{\mathrm{Yukawa}}$ are terms related to gauge bosons, fermions, the Higgs boson, and Yukawa interaction, respectively.

The $\mathcal{L}_{\mathrm{Gauge}}$ describes the kinematics and self-interaction of gauge bosons. It consists of three terms related to gauge fields (B_μ, $\boldsymbol{W}_\mu$, and $\boldsymbol{G}_\mu$) required by $U(1)$, $SU(2)$ and $SU(3)$

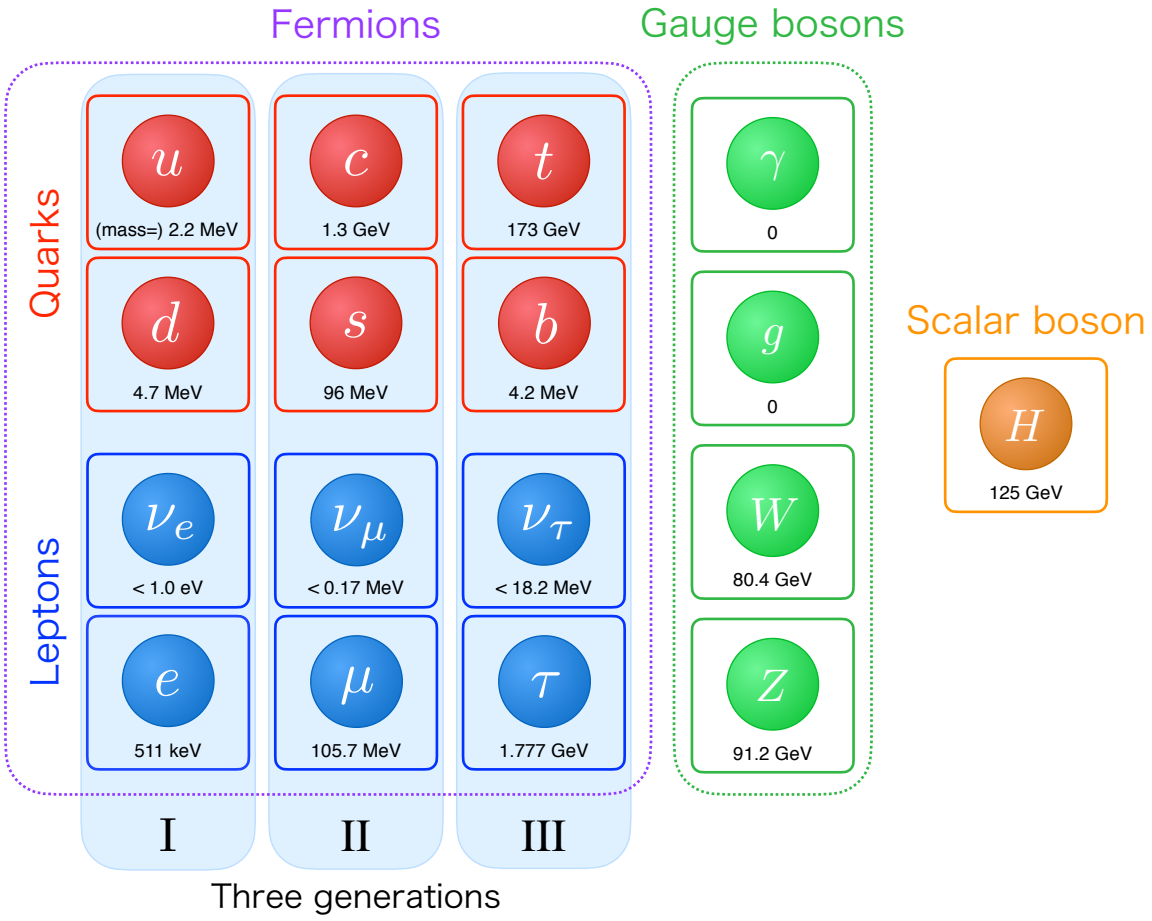


Figure 2.1 The SM particles.

gauge symmetries:

$$\mathcal{L}_{\text{Gauge}} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}\boldsymbol{W}_{\mu\nu}\boldsymbol{W}^{\mu\nu} - \frac{1}{4}\boldsymbol{G}_{\mu\nu}\boldsymbol{G}^{\mu\nu}, \tag{2.2}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.3}$$

$$\boldsymbol{W}_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - \frac{g_w}{2}(W_\mu W_\nu - W_\nu W_\mu), \tag{2.4}$$

$$\boldsymbol{G}_{\mu\nu} = \partial_\mu G_\nu - \partial_\nu G_\mu - \frac{g_s}{2}(G_\mu G_\nu - G_\nu G_\mu). \tag{2.5}$$

Here $B_{\mu\nu}$, $\boldsymbol{W}_{\mu\nu}$, and $\boldsymbol{G}_{\mu\nu}$ are the gauge field strength tensors. The factors g_w and g_s are coupling strength parameters.

The $\mathcal{L}_{\text{Fermion}}$ describes the kinematics of fermions and interaction between fermions and gauge fields as

$$\mathcal{L}_{\text{Fermion}} = \bar{\psi}_L\gamma^\mu iD_\mu\psi_L + \bar{\psi}_R\gamma^\mu iD_\mu\psi_R, \tag{2.6}$$

where $\psi_{L(R)}$ is the fermion fields of left-handed and right-handed, and γ^μ is gamma matrix. The

Table 2.1 The spin S, the electric charge Q, the weak hyper charge Y, and the third component of the weak isospin T^3. The subscripts of L and R correspond to the left-handed and right-handed, respectively.

				S	Q	Y	T^3
Quark	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$	$\frac{1}{2}$	$\begin{pmatrix} +\frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$	$+\frac{1}{6}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$
	u_R	c_R	t_R	$\frac{1}{2}$	$+\frac{2}{3}$	$+\frac{2}{3}$	0
	d_R	s_R	b_R	$\frac{1}{2}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0
Lepton	$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}$	$\frac{1}{2}$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$-\frac{1}{2}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$
	ν_{eR}	$\nu_{\mu R}$	$\nu_{\tau R}$	$\frac{1}{2}$	0	0	0
	e_R	μ_R	τ_R	$\frac{1}{2}$	-1	-1	0
Photon	γ			1	0	0	0
Weak boson	$\begin{pmatrix} W^+ \\ W^- \end{pmatrix}$			1	$\begin{pmatrix} +1 \\ -1 \end{pmatrix}$	0	$\begin{pmatrix} +1 \\ -1 \end{pmatrix}$
	Z			1	0	0	0
Gluon	g			1	0	0	0
Higgs boson	$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$			0	$\begin{pmatrix} +1 \\ 0 \end{pmatrix}$	$\frac{1}{2}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$

global covariant derivative D_μ is represented using local covariant derivative ∂_μ as

$$D_\mu = \partial_\mu - ig_w^{'}\frac{Y}{2}B_\mu - ig_w \boldsymbol{T}_w \boldsymbol{W}_\mu - ig_s \boldsymbol{T}_s \boldsymbol{G}_\mu, \tag{2.7}$$

where Y is the hyper charge, $g_w^{'}$ is the corresponding coupling strength parameter, and $\boldsymbol{T}_w$ and $\boldsymbol{T}_s$ are the isospin operators.

The $\mathcal{L}_{\text{Higgs}}$ and $\mathcal{L}_{\text{Yukawa}}$ describe the kinematics of the Higgs boson and the masses of the Higgs boson, gauge boson, and fermion. These terms are described in the following sections.

2.1.2 Higgs mechanism

In the SM, the Englert-Brout-Higgs mechanism spontaneously breaks $U(1)$ and $SU(2)$ gauge symmetries and generates mass terms for W and Z bosons and also for fermions.

The Higgs field ϕ is introduced by complex scalar fields as:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{2.8}$$

The Lagrangian related to the Higgs field and the Higgs potential $V(\phi)$ are described as

$$\mathcal{L}_{\text{Higgs}} = |D_\mu \phi|^2 - V(|\phi|^2), \tag{2.9}$$

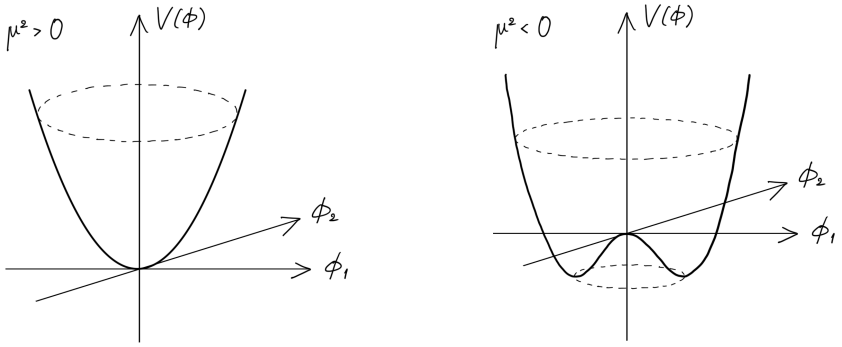
$$V(|\phi|^2) = \mu^2|\phi|^2 + \lambda|\phi|^4, \tag{2.10}$$

where μ and $\lambda(>0)$ are arbitrary constants. Since the complex scalar field is color-singlet, D_μ is described as:

$$D_\mu = \partial_\mu - ig_w^{'}\frac{Y}{2}B_\mu - ig_w \boldsymbol{T}_w \boldsymbol{W}_\mu. \tag{2.11}$$

The Higgs potential shapes for $\mu^2 > 0$ and $\mu^2 < 0$ are shown in Figure 2.2. It is believed that the potential before a phase transition had $\mu^2 > 0$ and then it changed to $\mu^2 < 0$ due to the transition caused by the cooling of the universe. If $\mu^2 > 0$, the vacuum which corresponds to the minimum potential is $\phi = 0$. In this case, the $U(1)$ and $SU(2)$ gauge symmetries are conserved. On the other hand, the vacuum for $\mu^2 < 0$ has the non-zero expectation value v:

$$\phi = \pm\sqrt{\frac{-\mu^2}{\lambda}} \equiv v. \tag{2.12}$$



(a) (b)

Figure 2.2 Shapes of the Higgs potential for $\mu^2 > 0$ (a) and $\mu^2 < 0$ (b). The axes for ϕ_3 and ϕ_4 are omitted for simplicity.

The Higgs field at the minimum for $\mu^2 < 0$ can be described as:

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix} \qquad (\phi_1, \phi_2, \phi_3, \phi_4) = (0, 0, v, 0). \tag{2.13}$$

The $U(1)$ and $SU(2)$ gauge symmetries are broken spontaneously.

The field is expanded around the vacuum as

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{2.14}$$

where $h(x)$ is the real scalar field corresponding to the physical Higgs boson. In the expansion, the following local gauge transformations are considered

$$U(1) \quad : \qquad \phi(x) \to e^{i\theta(x)/v}\phi(x), \tag{2.15}$$

$$SU(2) \quad : \qquad \phi(x) \to e^{i\boldsymbol{\tau}\boldsymbol{\theta}(x)/v}\phi(x). \tag{2.16}$$

The scalar kinetic energy term is described as

$$\begin{aligned} |D_\mu\phi|^2 &= \frac{1}{2}|\partial_\mu h|^2 \\ &+\frac{1}{8}g_w^2(v+h)^2(W_\mu^1 - iW_\mu^2)(W^{1\mu} - iW^{2\mu}) \\ &+\frac{1}{8}(v+h)^2(-g_w^{'}B_\mu + g_w W_\mu^3)^2. \end{aligned} \tag{2.17}$$

The second and third terms are written using the gauge fields of $W^\pm$ and Z as:

$$|D_\mu\phi|^2_{\text{2nd}-\text{term}} = \frac{g_w^2 v^2}{4}W_\mu^+ W^{-\mu} + \frac{g_w^2 v}{2}hW_\mu^+ W^{-\mu} + \frac{g_w^2}{4}h^2 W_\mu^+ W^{-\mu}, \tag{2.18}$$

$$|D_\mu\phi|^2_{\text{3rd}-\text{term}} = \frac{(g_w^2 + g_w^{'2})v^2}{8}Z_\mu Z^\mu + \frac{(g_w^2 + g_w^{'2})v}{4}hZ_\mu Z^\mu + \frac{g_w^2 + g_w^{'2}}{8}h^2 Z_\mu Z^\mu. \tag{2.19}$$

The first terms of Eqs. (2.18) and (2.19) are the mass terms for the W and Z bosons provided by the spontaneous symmetry breaking. The masses of the W and Z bosons are proportional to the vacuum expectation value:

$$m_{W^+} = m_{W^-} = \frac{g_w v}{2}, \tag{2.20}$$

$$m_Z = \frac{1}{2}v\sqrt{g_w^2 + g_w^{'2}}. \tag{2.21}$$

The second and third terms describe the interactions of the Higgs boson to W and Z bosons: triple coupling (W^+W^-H and ZZH) and quadratic coupling (W^+W^-HH and $ZZHH$), respectively.

Similarly, the potential is described as

$$\begin{aligned} V(|\phi|^2) &= \frac{1}{2}\mu^2(v+h)^2 + \frac{1}{4}\lambda(v+h)^4 \\ &= \frac{1}{4}\lambda h^4 + \lambda v h^3 + \lambda v^2 h^2 - \frac{1}{4}\lambda v^4. \end{aligned} \tag{2.22}$$

Here the first and second terms describe the Higgs trilinear and quartic couplings. The third term is the mass term for the Higgs boson. The mass is

$$m_H = \sqrt{2\lambda}v = \sqrt{-2\mu^2}. \tag{2.23}$$

2.1.3 Yukawa coupling

The Higgs field can also provide mass to fermions by mixing right and left chiralities via the Yukawa couplings. The Lagrangian of the Yukawa coupling is described as:

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= -y\bar{\psi}\phi\psi \\ &= -y\bar{\psi}_L\phi\psi_R - y\bar{\psi}_R\phi\psi_L. \end{aligned} \tag{2.24}$$

For leptons, the Lagrangian is written as

$$\begin{aligned}\mathcal{L}^{\ell}_{\text{Yukawa}} &= -y_\ell(\bar{\nu}_\ell, \bar{\ell})_L \phi \ell_R - y_\ell \bar{\ell}_R \phi \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L \\ &= -\frac{y_\ell(v+h)}{\sqrt{2}}(\bar{\ell}_L \ell_R + \bar{\ell}_R \ell_L) \\ &= -\frac{y_\ell v}{\sqrt{2}}\bar{\ell}\ell - \frac{y_\ell}{\sqrt{2}}\bar{\ell}\ell h \end{aligned} \tag{2.25}$$

where $\ell = (e, \mu, \tau)$. The first term is the mass term for the lepton. The lepton mass is

$$m_\ell = \frac{y_\ell v}{\sqrt{2}}. \tag{2.26}$$

Similarly, the quark mass is provided as $m_q = \frac{y_q v}{\sqrt{2}}$, where $q = (u, d, c, s, t, b)$. The second term of Eq. (2.25) is the interaction term for the lepton and the Higgs boson. Since the Yukawa coupling is proportional to the mass of the particle, a heavier particle has a larger Yukawa coupling.

The masses of the gauge bosons and fermions are provided by the vacuum expectation value. The strength of the couplings of the Higgs boson to the gauge bosons and fermions are represented by g and y, respectively, which are proportional to the mass of the particles. The Higgs self-coupling is provided with the coupling strength λ. Strangely, only the Higgs field distinguishes the fermion generations, although the fields in the QED and QCD do not distinguish the fermion generations. In addition, the SM does not predict the structure of the Higgs potential. In order to determine the structure of the Higgs potential for fundamental understanding of the Higgs mechanism, the Higgs couplings (g, y, and λ) to all particles need to be measured.

2.2 Higgs boson studies at LHC

Currently, the SM Higgs boson can be produced directly only at the LHC. In this section, the phenomena in the pp collisions at LHC are explained, and then the Higgs boson productions and decays are described.

The cross section of a process $pp \to X$ is described using the factorization theory [6] as

$$\sigma_{pp\to X} = \sum_{ij} \int dx_1 \int dx_2 \; f_i(x_1, \mu_{\text{F}}^2) f_j(x_2, \mu_{\text{F}}^2) \; \hat{\sigma}_{ij\to X}\left(x_1 p_1, x_2 p_2, \alpha_s^2(\mu_{\text{R}}), \mu_{\text{R}}, \mu_{\text{F}}\right), \tag{2.27}$$

which consists of the non-perturbative part $f_i(x_1, \mu_{\text{F}}^2) f_j(x_2, \mu_{\text{F}}^2)$ and perturbative part $\hat{\sigma}_{ij\to X}$. The $f_{i(j)}(x_{1(2)}, \mu_{\text{F}}^2)$ and $\hat{\sigma}_{ij\to X}$ are the parton distribution function (PDF) for partons $i(j)$ and the cross section of the elementary process, respectively. The PDF provides the parton density carrying a fraction $x_{1(2)}$ of the first (second) proton momentum at an energy scale μ_{F}. The μ_{F}, referred to as the QCD factorization scale, represents the energy scale at the boundary between perturbative QCD and non-perturbative QCD. The PDFs are measured by deep inelastic scattering experiments since they cannot be derived from calculation due to the non-perturbativity of the QCD. The PDF distributions of NNPDF3.1NNLO are shown in Figure 2.3.

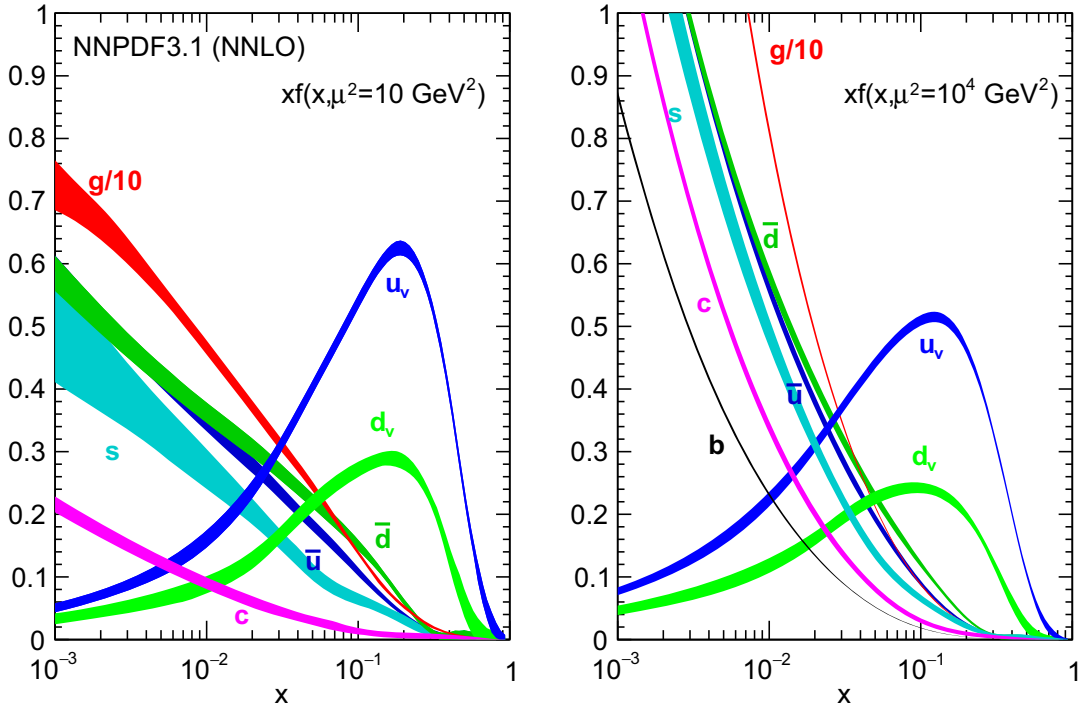


Figure 2.3 The NNPDF3.1NNLO PDFs, evaluated at $\mu^2 = 10$ GeV^2 (left) and $\mu^2 = 10$ GeV^4 (right). Here, $\mu = \mu_F = \mu_R$. [7]

On the other hand, the cross section $\hat{\sigma}_{ij\to X}$ is calculable by the QCD and EW theories. The cross section can be represented by a function of the parton momentum $x_1p_1(x_2p_2)$ of parton $i(j)$, the QCD renormalization scale μ_R, the QCD coupling constant α_s, and μ_F. The μ_R is the scale on which the renormalization is performed to avoid the divergence of the perturbation calculation. The QCD coupling constant α_s depends on μ_R.

Typical products from a pp collision are shown in Figure 2.4. There are many products from the processes in a pp collision not only from hard scattering.

Parton Shower (PS) Parton involved in a hard interaction causes gluon radiation, photon radiation, and quark anti-quark splitting. Especially, the radiation of the parton in the parton-level initial and final state is called initial state radiation (ISR) and final state radiation (FSR), respectively.

Hadronization Parton transits into colorless hadrons at parton-level final state. This is described by the fragmentation function, which uses QCD-inspired phenomenological models. The most common model is "Lund-String-Model" [8]. The parameters of the model are derived from experimental results since hadronization is a non-perturbative process.

Underlying event This is from the interaction between the remnants of protons at lower energy.

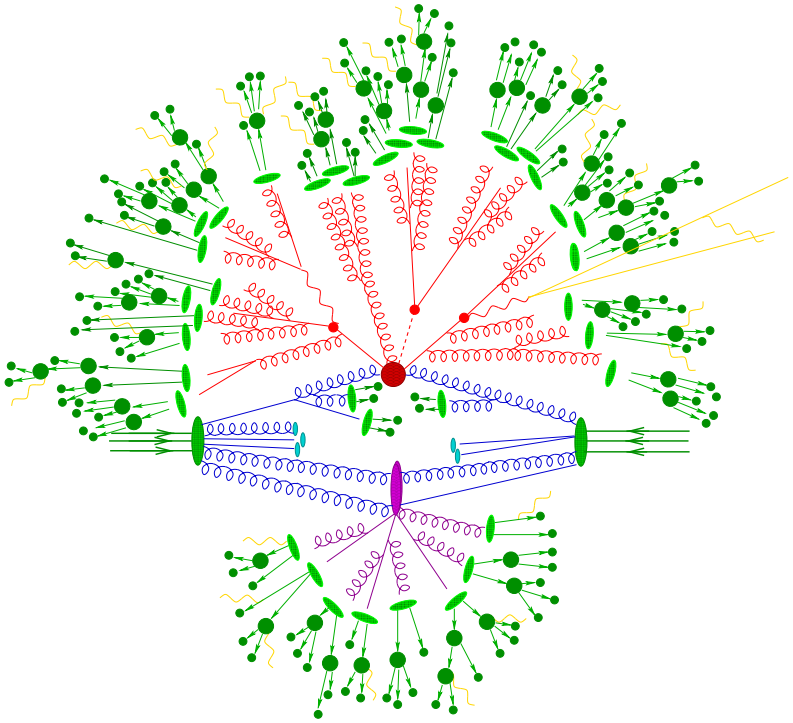


Figure 2.4 Sketch of a pp collision (adapted from [9]).

The main production processes of the SM Higgs boson at LHC are the following:

- gluon-gluon fusion production (ggF): $gg \to H$
- vector-boson fusion production (VBF): $q\bar{q} \to q\bar{q}H$
- associated production with vector bosons (VH): $q\bar{q} \to VH$ ($V = W$ or Z)
- associated production with pair of top quarks (ttH): $q\bar{q}/gg \to t\bar{t}H$

Feynman diagrams are shown in Figure 2.5. The SM Higgs boson production cross sections at $\sqrt{s} = 13$ TeV are summarized in Table 2.2. The ggF process is the dominant Higgs boson production process. The interaction between the gluon and the Higgs boson is mediated predominantly by a heavy quark loop since the gluon is massless. The second dominant production process is the VBF process. The VBF process has two forward jets in the final state. The VH process is the third dominant production process. It has a different final state (number of leptons and quarks, and missing transverse momentum) between WH and ZH process. The production cross section of the ttH process is smaller than other processes. It contains the Yukawa coupling to top quark. The final state of the ttH process has two jets from bottom quarks.

The branching ratios of the main decays of the Higgs boson are summarized in Table 2.3. Since the Higgs boson has a larger coupling to a particle with larger mass as explained in Section 2.1, the branching ratio for fermions is larger in descending order of mass, with top

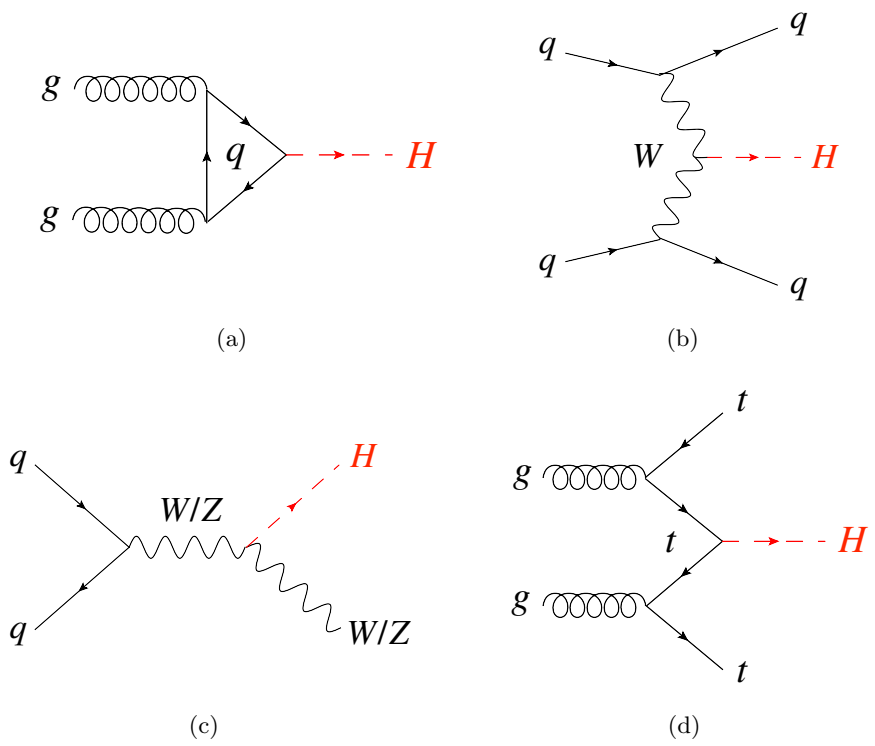


Figure 2.5 Feynman diagrams of the SM Higgs boson production. (a): gluon-gluon fusion, (b): vector-boson fusion, (c) associated production with vector boson and (d) associated production with pair of top quarks.

Table 2.2 The SM Higgs boson production modes and the cross section at $m_H = 125.09$ GeV [10].

Production mode	cross section [pb]
ggF	48.517
VBF	3.779
WH	1.369
ZH	0.8824
ttH	0.5065

quarks excluded ($b\bar{b} > \tau\tau > c\bar{c} > \mu\mu$). For gauge bosons, the branching ratio is relatively small since it occurs only virtually. The ggF, VBF, VH, ttH production processes and $H \to WW^*, H \to ZZ^*, H \to \gamma\gamma, H \to b\bar{b}, H \to \tau\tau$ decay processes have been observed with a significance of greater than 5σ. Of the processes including the Higgs coupling to fermions, only the couplings to third-generation charged fermions have been observed. The couplings to the first- and second- generation fermions have not been observed yet. Of the second-generation fermions, the charm quark has the largest mass, and the branching ratio for $H \to c\bar{c}$ is the largest in the SM. However, the discrimination of charm quark from other quarks is difficult at hadron collider since charm quark produces jets inside a detector. Although the branching ratio of H $\to \mu\mu$ is small, 0.02% in the SM, it is easier to reconstruct the muons. Therefore, $H \to \mu\mu$ decay is the next target for measurements of the Higgs coupling at LHC.

Table 2.3 The main decay modes of the SM Higgs boson and the branching ratios at $m_H = 125.09$ GeV [10].

Dacay mode	Branching ratio (%)
$H \to b\bar{b}$	58.09
$H \to WW$	21.52
$H \to gg$	8.180
$H \to \tau\tau$	6.256
$H \to c\bar{c}$	2.884
$H \to ZZ$	2.641
$H \to \gamma\gamma$	0.2270
$H \to Z\gamma$	0.1541
$H \to \mu\mu$	0.02171

The ATLAS experiment published the result of the H $\to \mu\mu$ search using $\sqrt{s} = 13$ TeV pp collisions data collected in 2015 and 2016 corresponding to 36.1 fb^{-1} at LHC [11]. No significant excess was observed, and an upper limit was set on the cross section times H $\to \mu\mu$ branching ratio to be 3.0 times the SM prediction at the 95% confidence level.

Figure 2.6 shows the result of the Higgs coupling measurements so far [12]. The Higgs couplings to gauge bosons (W, Z) and the third-generation charged fermions (t, b, τ) are consistent with the SM. For the Higgs coupling to the muon, a preliminary result of H $\to \mu\mu$ using partial Run-2 data of 79.8 fb^{-1} [13] is considered. The Higgs coupling to the muon has still large uncertainty, and the universality of the coupling mechanism between the fermion generations is not confirmed yet.

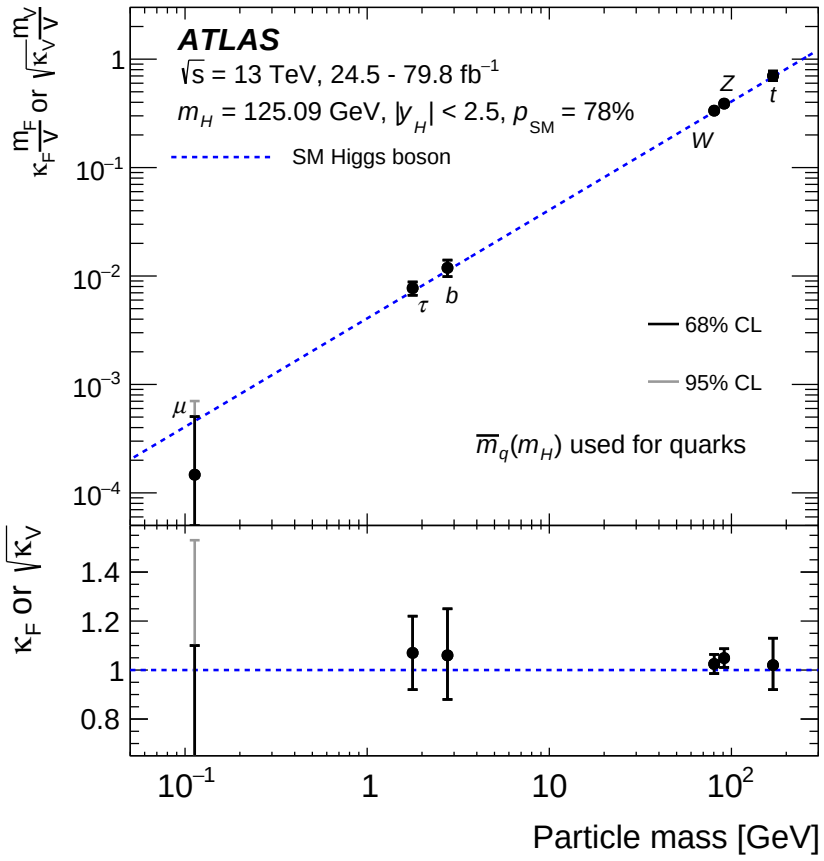


Figure 2.6 Coupling-strength scale factors as a function of the particle mass. [12]

Chapter 3

The LHC-ATLAS experiment

3.1 The LHC accelerator

The Large Hadron Collider (LHC) [14] is located 100 m underground at the European Organization for Nuclear Research (CERN) across the Swiss-French border. It is the largest energy circular proton-proton accelerator in the world with the design center of mass energy up to 14 TeV.

Protons extracted from hydrogen atoms are accelerated in stages by one linear accelerator and four circular accelerators. An overview of LHC accelerator system is shown in Figure 3.1. The protons are first accelerated by the linear accelerator LINAC2 to the energy of 50 MeV, then to 1.4 GeV by the BOOSTER, to 25 GeV by the Proton Synchrotron (PS), and to 450 GeV by the Super Proton Synchrotron (SPS). The accelerated protons are split into two portions and inserted into each of the two beam pipes in the LHC ring with a circumference of 27 km. In the LHC ring, there are 1232 superconducting dipole magnets, which provide a magnetic field of 8.3 T for bending the beams and 392 quadrupole magnets for focusing the beams. The beams are accelerated in opposite directions up to 7 TeV maximum in the LHC ring. The beams have a bunch structure. Each bunch contains about 10^{11} protons. The bunch spacing is 25 ns.

During the data taking periods in 2015–2018, the LHC accelerated the proton beams to 6.5 TeV and operated with a center of mass energy of 13 TeV. The design values of the LHC beam and machine parameters and of the values used in the operation in 2015–2018 are listed in Table 3.1.

The LHC beam intensity is described using the instantaneous luminosity defined as:

$$L = \frac{N_b^2 n_b f_{\mathrm{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F, \tag{3.1}$$

$$F = \frac{1}{\sqrt{1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*_{xy}}\right)^2}}, \tag{3.2}$$

where γ_r, σ_z, and σ^*_{xy} are the Lorentz factor, the longitudinal beam length, and the transverse beam size, respectively. The instantaneous luminosity corresponds to the interaction frequency of particles per unit time and unit area. Therefore, the expected number of generated events in unit time for a process with cross section $\sigma_{\mathrm{process}}$ is $L \times \sigma_{\mathrm{process}}$.

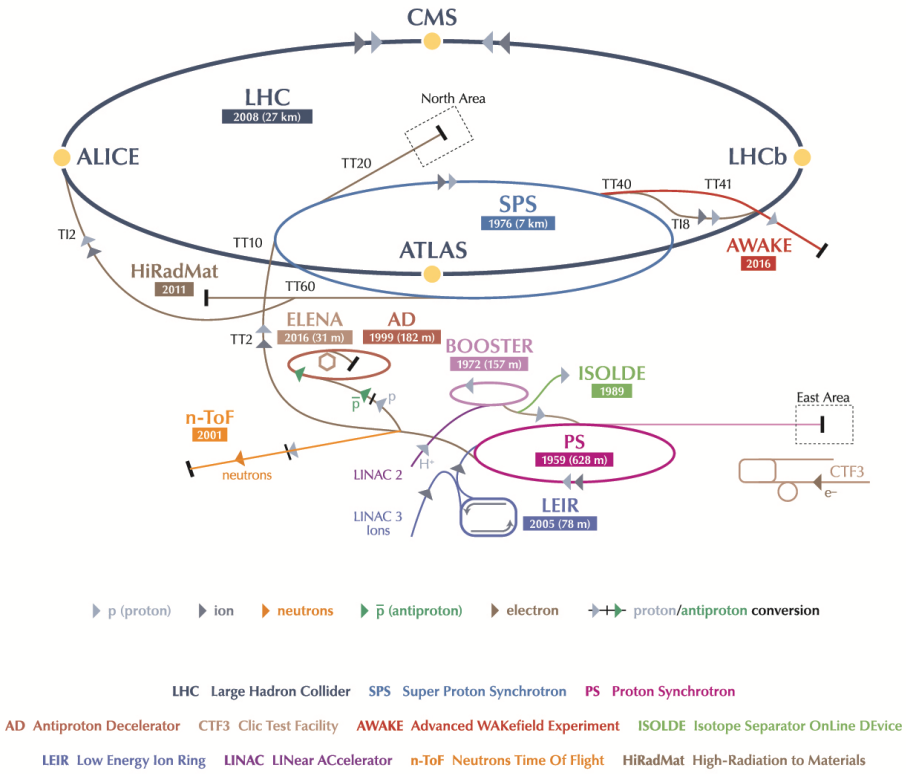


Figure 3.1 Overview of the LHC accelerator system. [15] The protons accelerated by LINAC2, BOOSTER, PS, and SPS are inserted in the LHC. There are four experiments at the four beam collision points on the ring: ATLAS, CMS, ALICE, and LHCb.

Multiple pp interactions arise per bunch crossing due to the high instantaneous luminosity. The additional interactions are called "pile-up" and typically arise from the inelastic QCD interaction. The mean number of interactions per bunch crossing, referred to as average pile-up $< \mu >$, is calculated by

$$< \mu >= \frac{L \times \sigma_{\text{inel}}}{N_b \times f_{\text{rev}}}, \tag{3.3}$$

where σ_{inel} is the inelastic pp cross section. The pile-up is subdivided into in-time and out-of-time pile-up. The in-time pile-up is additional pp interactions in the same bunch crossing as the hard interaction of interest. The out-of-time pile-up is produced by pp interactions in the bunch crossing before or after the collision of interest. The pile-up affects the performance of the trigger decision and object reconstruction.

The amount of data taken in a period is described by the integrated luminosity, which is defined by integrating the instantaneous luminosity during the time of period. The integrated luminosity of the collisions delivered by the LHC during 2015–2018 is 156 fb^{-1}.

Table 3.1 Beam and machine parameters of the LHC in 2018 compared to the design [16].

Parameters	Design	2018
N_b: Number of protons per bunch [$\times 10^{11}$]	1.15	1.15
n_b: Number of bunches per LHC ring	2808	2556
$1/f_{\mathrm{rev}}$: Bunch spacing [ns]	25	25
β^*: Beta function at the interaction point [cm]	55	30–25
θ_c: Beam crossing angle [μrad]	285	320–260
ϵ_n: Normalized transverse emittance at the start of collisions [μm]	3.75	2.0
E_{beam}: Beam energy [TeV]	7	6.5

3.2 The ATLAS detector

The ATLAS detector is a huge cylindrical general-purpose detector installed at one of the four LHC collision points. An overview of the ATLAS detector is shown in Figure 3.2. The size is 25 m in diameter and 44 m in total length. The weight is 7,000 tons. It covers nearly 4π solid-angle around the collision point. The ATLAS detector consists of inner detectors, electromagnetic calorimeters, hadron calorimeters, and muon spectrometers in that order from the inside. In addition, three superconducting magnets are installed in the ATLAS detector. The design of sub-detectors differs between the top and bottom parts, referred to as "end-cap", and the side part, referred to as "barrel", of the cylinder.

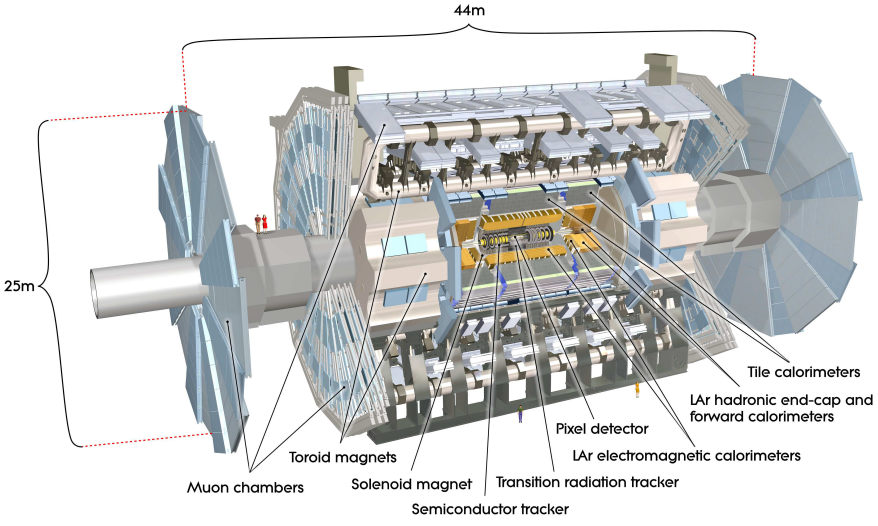


Figure 3.2 Overview of the ATLAS detector. [3]

3.2.1 Coordinate system

Figure 3.3 shows the coordinate system used in the ATLAS experiment. The center of the detector is defined as the origin and two coordinate axes are used: Cartesian coordinate system and cylindrical coordinate system. In the Cartesian coordinate system, the x-axis, y-axis, and z-axis are defined as pointing to the center of the LHC ring, as perpendicular to the ground and upward, and as along the beam line, respectively. In the cylindrical coordinate system, the radial distance $R(=\sqrt{x^2+y^2})$, the azimuthal angle ϕ $(-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2})$, and z are used. In the ATLAS experiment, pseudorapidity η is also used to express the θ $(-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2})$ direction

$$\eta = -\ln \tan \frac{\theta}{2}. \tag{3.4}$$

In the high energy limit ($E \sim p$), pseudorapidity is an approximation of the rapidity y of a particle with energy E and momentum p as described by:

$$y = \frac{1}{2} \ln \left(\frac{E + p\cos\theta}{E - p\cos\theta} \right) \sim \frac{1}{2} \ln \left(\frac{1+\cos\theta}{1-\cos\theta} \right) = \eta. \tag{3.5}$$

Similarly, transverse momentum p_{T} (energy E_{T}) defined below is often used since the transverse momentum (energy) of the particles from the collisions of interest is relatively large.

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2} = p \sin\theta \tag{3.6}$$

$$E_{\mathrm{T}} = \sqrt{E_x^2 + E_y^2} = E \sin\theta \tag{3.7}$$

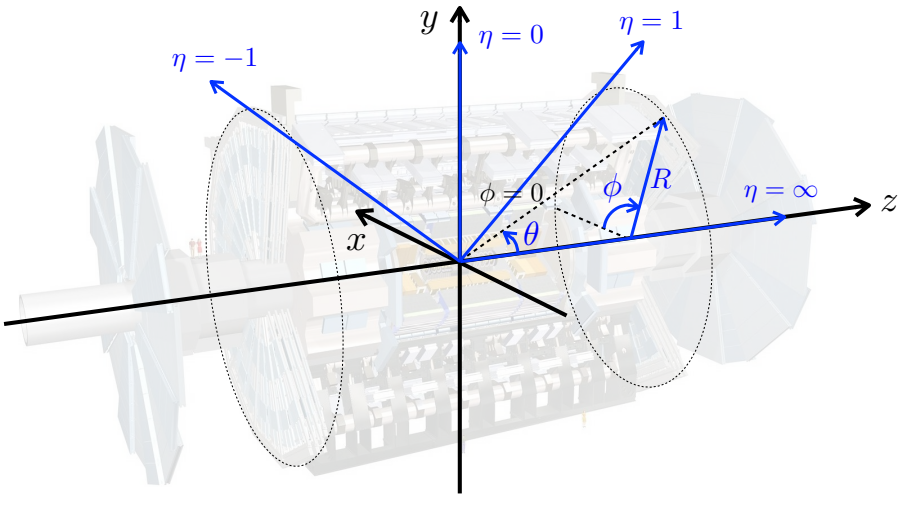


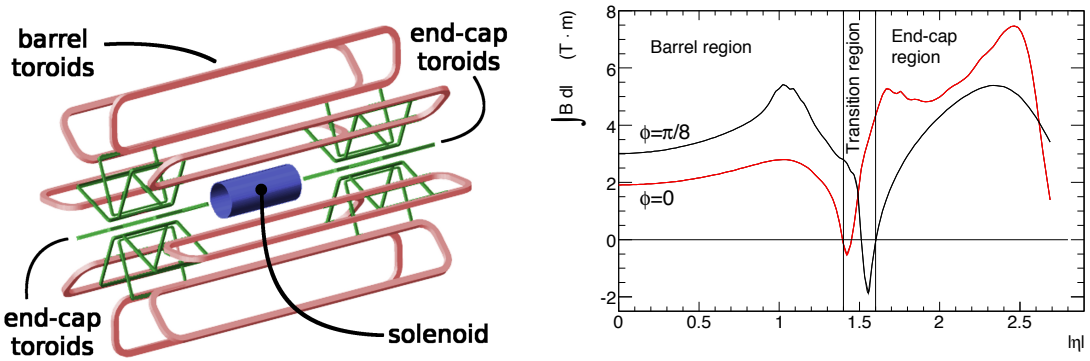
Figure 3.3 The ATLAS coordinate system. The origin of the system is the center of the detector. There are the cartesian coordinate system (x, y, z) and the cylindrical coordinate system (R, ϕ, θ). Pseudorapidity $\eta = -\ln \tan \frac{\theta}{2}$ is also used.

3.2.2 Magnet system

Since the momentum of charged particles is measured using the curvature of the trajectory, magnetic fields are essential in the ATLAS experiment. The ATLAS magnet system consists of three superconducting magnets as shown in Figure 3.4(a): the central solenoid, the barrel toroid, and the end-cap toroid.

The central solenoid with 2.5 m diameter and 5.8 m length surrounds the inner detector. It provides a uniform magnetic field along the beam line with 2 T field strength inside the coil. The charged particles from the interaction point are bent in the ϕ direction by the magnetic field.

Each of the barrel toroid and the end-cap toroid consists of eight toroidal magnets located outside the calorimeter. These magnets provide toroidal magnetic fields along the ϕ direction inside the coils. The magnetic field integral shown in Figure 3.4(b) is 8-fold symmetric due to the structure of the magnets. In the transition region ($1.4 < |\eta| < 1.6$), the magnetic field integral is small. In the barrel region ($0 < |\eta| < 1.4$), the magnetic field varies 0.15–2.5 T depending on $|\eta|$. In the end-cap region ($1.6 < |\eta| < 2.7$), it is 0.2–3.5 T. The charged particles from the interaction point are bent in the η direction by the toroidal magnetic field.

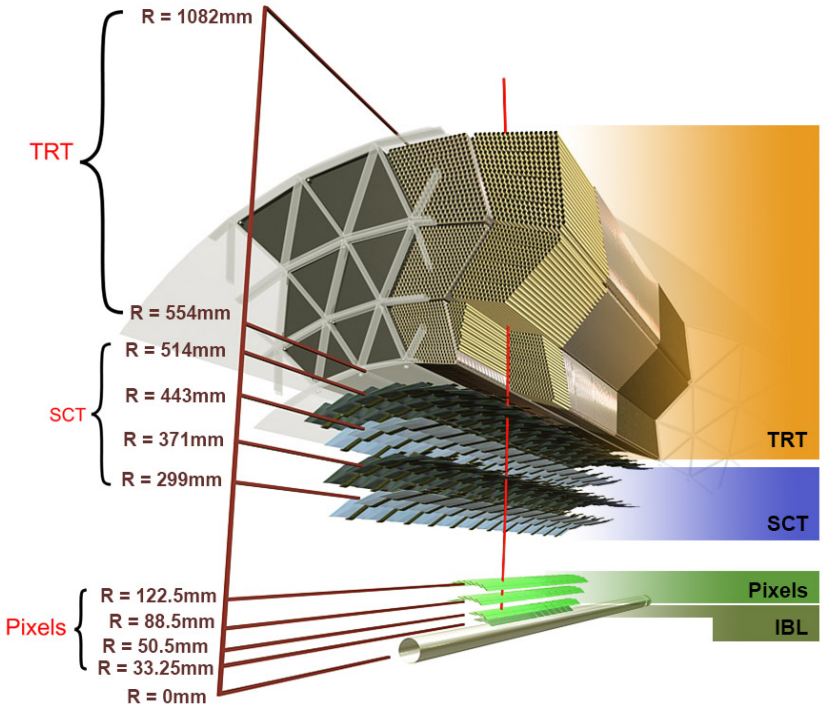


(a) (b)

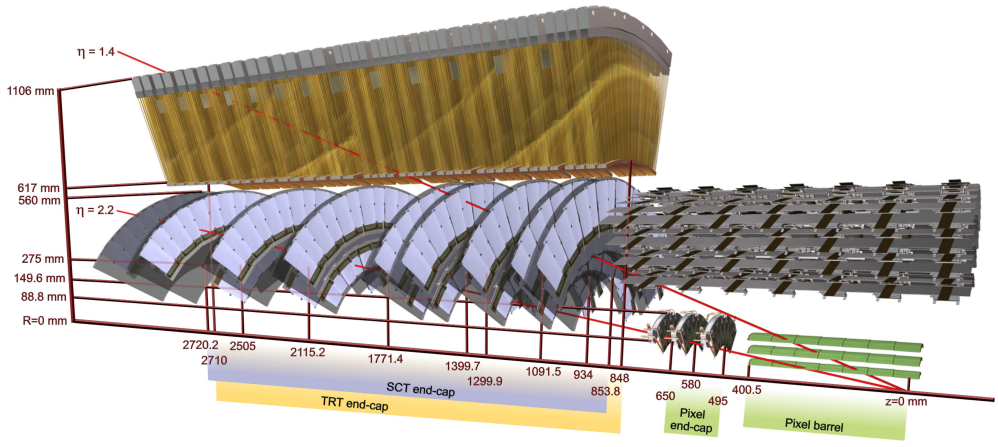
Figure 3.4 The ATLAS magnet system (a) and the integrated toroidal magnetic fields as a function of $|\eta|$ (b).

3.2.3 Inner detector

The inner detector (ID) is installed in the innermost part of the ATLAS detector and surrounded by the solenoid magnet. It measures the momenta and directions of charged particles from the interaction point by reconstructing the trajectories. The ID consists of three sub-detectors as shown in Figure 3.5: the silicon pixel detector, the semiconductor tracker (SCT), and the transition radiation tracker (TRT) in that order from the inside.



(a)



(b)

Figure 3.5 Cutaway view of the inner detector in the barrel (a) and end-cap (b) regions. [17] The ID consists of the silicon pixel detector (IBL and Pixel), the SCT, and the TRT in that order from the inside.

Silicon pixel detector

The silicon pixel detector consists of three Pixel layers and the Insertable B-Layer (IBL), which was newly installed in 2015 [18], in the barrel and three Pixel disks in each of the end-caps. The silicon pixel detector covers the region $|\eta| < 2.5$. The Pixels and IBL have high-granularity pixel sensors with the size of 50 μm $\times$ 400 μm and 50 μm $\times$ 250 μm in the R–$\phi \times z$ plane, respectively. The total number of readout channels for the Pixels and IBL are 80 M and 12 M, respectively.

Semiconductor tracker

The SCT is a silicon detector with single-sided silicon strip sensors of 80 μm pitch and 128 mm length. It consists of four SCT layers in the barrel and nine SCT disks in each of the end-caps. The SCT covers the region $|\eta| < 2.5$. In each module, the sensors are glued with a stereo angle of 40 mrad for two-dimensional measurement of hit position. The total number of readout channels is 6.3 M.

Transition Radiation Tracker

The TRT is installed outside the SCT and consists of polyimide straw tubes with a diameter of 4 mm. The straws are filled with a gas mixture of Xenon (70%), CO_2 (27%), and O_2 (3%). The maximum drift time is 45 ns at the operating voltage of the cathode wire of 1.5 kV. The intrinsic position resolution per tube is about 130 μm. The tubes are arranged in parallel along the beam pipe with a 7 mm interval in the barrel. In the end-cap, the tubes are arranged along the R direction. Since the TRT provides only the information in the R–ϕ plane, the information from the pixel and SCT are combined for three-dimensional reconstruction of the trajectory. The total number of readout channels is 350 k.

3.2.4 Calorimeters

The calorimeter system consists of the electromagnetic (EM) calorimeters and hadronic calorimeters. The EM calorimeters surround the solenoid magnet and the hadronic calorimeters surround the EM calorimeters as shown in Figure 3.6. Both calorimeters are sampling calorimeters, composed of absorber material and active material. The calorimeters have crack regions between barrel and end-caps of $1.37 < |\eta| < 1.52$.

Electromagnetic calorimeters

The EM calorimeters measure the energy deposit and position of electromagnetic shower from electrons and photons. The calorimeters are divided into the barrel region ($|\eta| < 1.475$) and the end-cap region ($1.375 < |\eta| < 3.2$). For both of the calorimeters, lead and liquid argon (LAr) are used as the absorber and the active material, respectively. These have accordion structure as shown in Figure 3.7 to be uniform in ϕ with keeping the distance between readout electrodes. The EM calorimeters are divided into two or three layers in the depth direction. The first layer (the innermost) has high granularity in the η direction with approximately 4.3 radiation lengths (X_0). The second layer has large depth of about 16 X_0 so that most of the energy deposit occurs

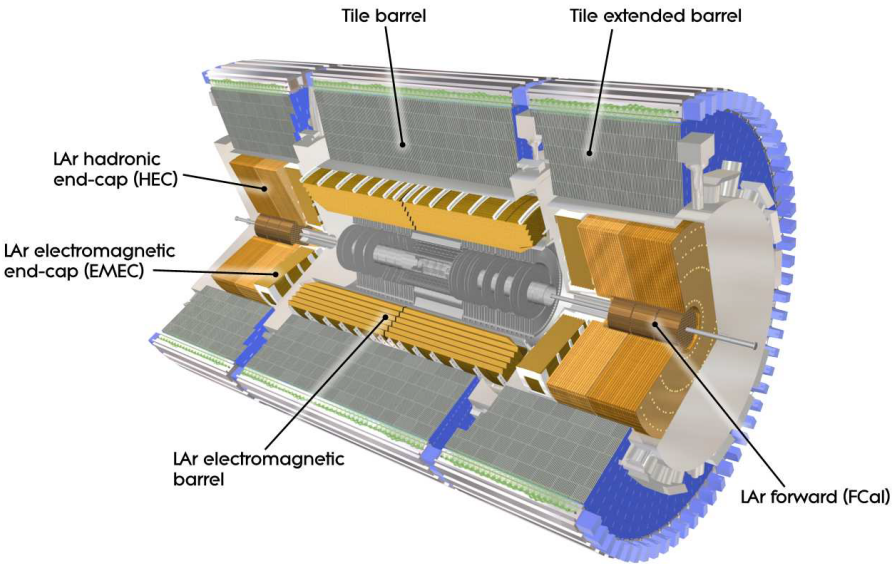


Figure 3.6 Cutaway view of the calorimeter system. The calorimeter system consists of the electromagnetic calorimeters and hadronic calorimeters.

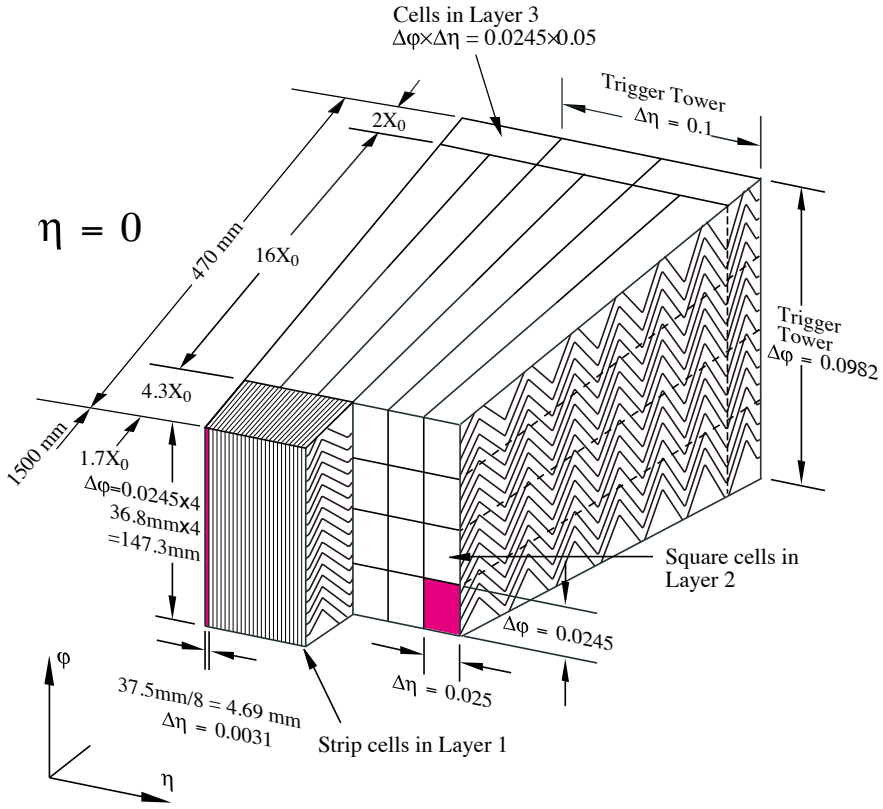


Figure 3.7 Illustration of calorimeter layers.

in this layer. The third layer (the outermost) measures the shower tail to separate electrons from $\pi^{\pm}$. In the region $|\eta| < 1.8$, the presampler consisting of one LAr layer measures the energy loss in front of the EM calorimeters. The total thickness of the EM calorimeters is $> 22\ X_0$ ($> 24\ X_0$) in the barrel (end-caps) region. The calorimeters are designed with energy resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$. The total number of readout channels is 170 k.

Hadronic calorimeters

The hadronic calorimeters measure the energy deposit and position of hadron shower. It consists of three calorimeters; tile ($|\eta| < 1.7$), end-cap ($1.5 < |\eta| < 3.2$), and forward calorimeters ($3.1 < |\eta| < 4.9$).

The tile calorimeters use iron and scintillator as the absorber and the active material, respectively. The tile calorimeters consist of three layers. The $\Delta\eta \times \Delta\phi$ granularity is 0.1×0.1 in the first and second layers and 0.2×0.1 in the last layer.

For the end-cap calorimeter, copper and LAr are used as the absorber and the active material, respectively. The end-cap calorimeter consists of four layers. The $\Delta\eta \times \Delta\phi$ granularity is 0.1×0.1 in $1.5 < |\eta| < 2.5$ and 0.2×0.2 in $2.5 < |\eta| < 3.2$.

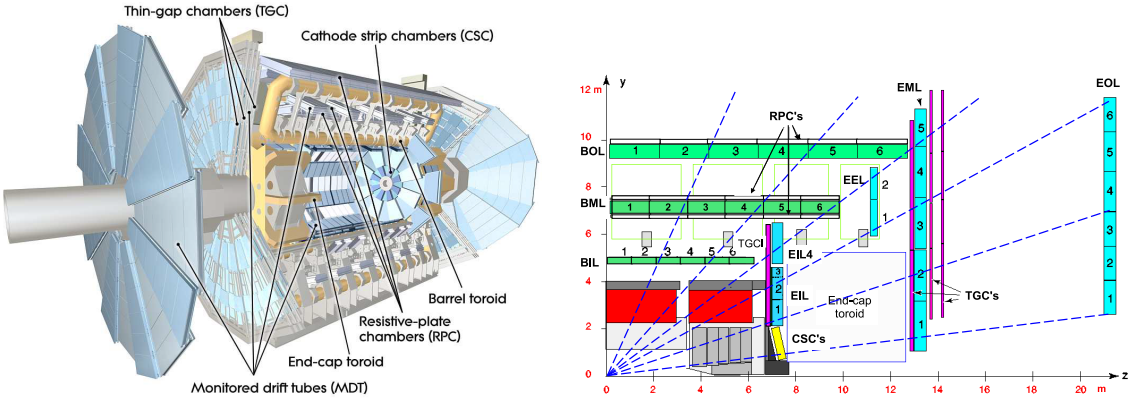
The forward calorimeters consist of one EM calorimeter layer and two hadronic calorimeter layers. They measure both EM and hadronic energy. For the EM layer, which is the first layer, copper and LAr are used as the absorber and the active material, respectively. For the hadronic layers, which are the second and third layers, tungsten and LAr are used as the absorber and the active material, respectively.

The typical energy resolution is $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$. The energy resolution for the forward calorimeter is $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$. The total number of readout channels is 14.3 k (Tile: 5.2 k, End-cap: 5.6 k, and Forward: 3.5 k).

3.2.5 Muon Spectrometer

The muon spectrometer (MS) is located at the outermost of the ATLAS detector [19]. Muons passing through the calorimeter are bent by the toroid magnetic field and then enter the MS. The MS measures the curvature of the muon trajectories and derives the transverse momentum. The MS consists of four sub-detectors as shown in Figure 3.8(a). The MS is divided into two types depending on the purpose: precision chambers and trigger chambers. The precision chambers, the Monitored Drift Tubes (MDT) and the Cathode Strip Chambers (CSC), are used to measure the momentum with high precision. The trigger chambers, the Resistive Plate Chambers (RPC) and the Thin Gap Chambers (TGC), are used to trigger events that contain muons with high transverse momentum. The trigger chambers are required to have a fast response to provide the muon trajectory within 25 ns.

Figure 3.8(b) shows the layout of the MS. There are multiple stations in each of the barrel and end-cap regions. The muon trajectory is derived from the position of the muon in each station.



(a) (b)

Figure 3.8 Cutaway view of the muon spectrometers (a). The muon spectrometers layout at $\phi = \pi/2$ in the R–z plane (b).

Monitored Drift Tubes

The MDT is the main precision chamber. It consists of multiple cylindrical drift tubes with a diameter of 30 mm. The tubes are filled with a gas mixture of Ar (93%) and CO_2 (7%) at 3 bar absolute pressure. The tungsten-rhenium gold-plated wires with a diameter of 50 μm are used for the anode wires at a potential of 3080 V. In this condition, each drift tube has about 700 ns maximum drift time and 2×10^4 amplification factor so that the average resolution is 80 μm. The MDT subsystem consists of three stations, inner, middle, and outer, in concentric cylinder shapes around the beam axis in the barrel region ($|\eta| < 1$). In the end-cap region ($1 < |\eta| < 2.7$), it consists of three or four stations, inner, extra, middle, and outer, in perpendicular wheels. In each station, there are six or eight tube layers. In the end-cap inner region, the MDT covers only $|\eta| < 2.0$ due to the higher hit rate of > 150 Hz/cm^2 in $|\eta| > 2.0$. The total number of readout channels is 357 k.

Cathode Strip Chambers

The CSC is the precision chamber that covers the inner $2.0 < |\eta| < 2.7$ region. It is a multi-wire proportional chamber with the anode wires of 2.5 mm pitch parallel to the R direction and the cathode strips of 5 mm pitch parallel to the ϕ direction. The anode wires have a diameter of 30 μm and are operated at a potential of 1900 V. The gap is filled with a gas mixture of Ar (80%) and CO_2 (20%). The gas gain is 6×10^4. The CSC has a relatively short drift time of < 40 ns and thus that can be used in the region with the higher hit rate. The CSC consists of four layers in a perpendicular wheel-shaped station. The position resolution is 60 μm in the R direction and 5 mm in the ϕ direction. The total number of readout channels is 31 k.

Resistive Plate Chambers

The RPC is the trigger chamber for the barrel region ($|\eta| < 1.05$). It is a gaseous parallel electrode-plate detector with a gas gap of 2 mm. The gap is filled with a gas mixture of $C_2H_2F_4$ (94.7%), Iso-C_4H_{10} (5%), and SF_6 (0.3%). A uniform electric field between the plates of about 4.9 kV/mm produces the avalanche multiplication of the ionization electrons created by a charged particle. The signal is read out via strips mounted on the outer faces of the plates. The rise time of the avalanche signal is as fast as several ns. The typical space-time resolution is of the order of 1 cm $\times$ 1 ns. The total number of readout channels is 383 k.

Thin Gap Chambers

The TGC is the trigger chamber for the end-cap region ($1.05 < |\eta| < 2.4$). It is a multi-wire proportional chamber with a thinner distance between the anode wire and cathode strip (1.4 mm) than that between the anode wires (1.8 mm). The thin gap contributes to the fast response of the chamber. The strips are arranged parallel to the R direction. The wires with a diameter of 50 μm are arranged orthogonally to the strips. The TGC provides two dimensional hit position using the strips and wires. The operating voltage of the anode wires is 2800 V. In the gap, a highly quenching gas mixture of CO_2 (55%) and n-C_5H_{12} (45%) is filled. The gas amplification is 3×10^5.

There are three stations outside the toroidal magnet. Each station consists of two or three gas gaps and two or three wire-layers. There are seven gas gaps and seven wire-layers in total. Inside the toroidal magnet, there is one station with two gas gaps and two wire-layers. Due to the high granularity of the TGC wires, less than 35 wires corresponding to the η pitch of less than 60 mm are readout as one channel depending on $|\eta|$. The total number of readout channels is 320 k.

3.3 Trigger and data acquisition

Since the ATLAS detector has about 100 M readout channels, the size of the raw data per pp collisions is very large. The ATLAS experiment cannot record this huge amount of data for all pp collisions that occur with a frequency of 40 MHz due to limitations on data transferring and recording rate. In addition, most of the pp collisions arise from the QCD interactions with small momentum transfer, and the event rate is small for the processes of interest, such as the Higgs boson production. The cross section of the total Higgs boson production is $O(10)$ pb, while the cross section of the total inelastic pp scattering is $O(10^{10})$ pb. The events are examined and selected for storage during the data taking ('triggering"). The ATLAS trigger system typically selects the events with high-p_{T} leptons, photons, and jets since the processes of interest may have high-p_{T} particles in the final state.

The ATLAS trigger and data acquisition system is shown in Figure 3.9 [20]. The events are triggered in two stages, the Level-1 trigger and the High-Level trigger. The Level-1 trigger is a hardware-based trigger that uses ASICs and logic integrated circuits in order to perform high-speed processing. It roughly selects the events and reduces the event rate to 100 kHz within 2.5 μs. In the Level-1 trigger, there are the Level-1 Muon and Level-1 Calo triggers which

make decisions based on the MS and the calorimeters, respectively. The Level-1 trigger decision is formed by the Central Trigger Processor, which receives inputs from the Level-1 Muon and Level-1 Calo triggers. The data of the events triggered by the Level-1 trigger are transfered to the High-Level trigger. The High-Level trigger is a software-based trigger that reconstructs events more precisely with more detailed detector information. It reduces the event rate to approximately 1 kHz within a few seconds.

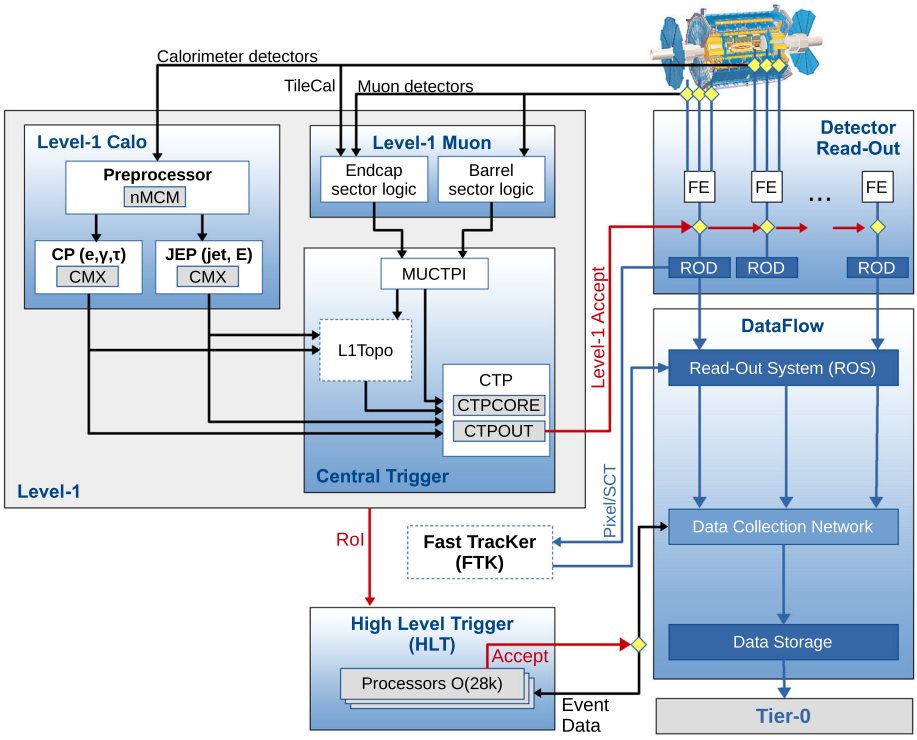
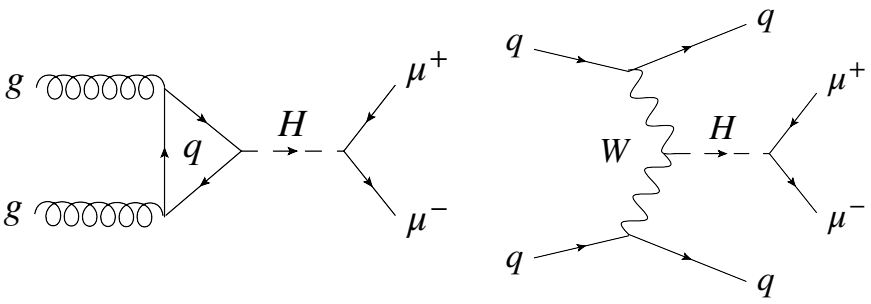


Figure 3.9 The ATLAS trigger and data acquisition system for the data taking in 2015–2018.

Chapter 4

Search for $\mathrm{H} \to \mu\mu$

We search for the $\mathrm{H} \to \mu\mu$ decay using $\sqrt{s} = 13$ TeV pp collisions data collected in 2015–2018 corresponding to 139 fb^{-1}. In this search, the ggF, VBF, VH, and ttH production processes (their Feynman diagrams are shown in Figure 4.1–4.3) are considered for signal. In the final state for all processes, there is an isolated opposite-charge muon pair from the Higgs boson decay. The p_{T} distribution of the muons has a peak at around 60 GeV. For the VH and ttH processes, only the processes including the leptonic decay of the boson and $t\bar{t}$ dileptonic and semileptonic decays are considered, respectively, thanks to the better background separation. The only charged leptons considered for leptonic decays are electrons and muons.



(a) (b)

Figure 4.1 Feynman diagrams of the $\mathrm{H} \to \mu\mu$ decay in the ggF (a) and VBF (b) processes. In the VBF process, there are two jets in addition to the opposite-charge muon pair in the final state.

In each signal process, constituent particles of the final state and main background process are different. The expected signal yield in the $\sqrt{s} = 13$ TeV 139 fb^{-1}, the constituent particles of the final state, and the main background process in each signal process are summarized in Table 4.1.

The dominant background process is the Drell-Yan (DY) process shown in Figure 4.4. The cross section of the DY process in the dimuon invariant mass of 120–130 GeV is approximately 6 pb [21], while the signal cross section is 0.012 pb. The DY background is an irreducible background due to the same final state particles as a signal. Besides the DY process, the

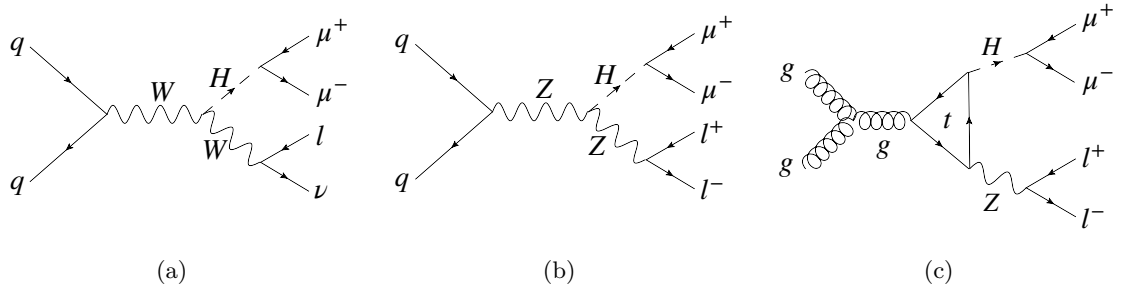


Figure 4.2 Feynman diagrams of the H → $\mu\mu$ decay in the VH processes. The $qq \to WH$ (a), $qq \to ZH$ (b), and $gg \to ZH$ (c) are main processes. Only the processes including the leptonic decay of the boson are considered ($\ell = e$ or μ). In the final state of the WH process, there is one charged lepton from $W \to l\nu$ decay in addition to the opposite-charge muon pair from the Higgs boson decay. The final state of ZH process has the muon pair and two additional charged leptons from $Z \to ll$ decay.

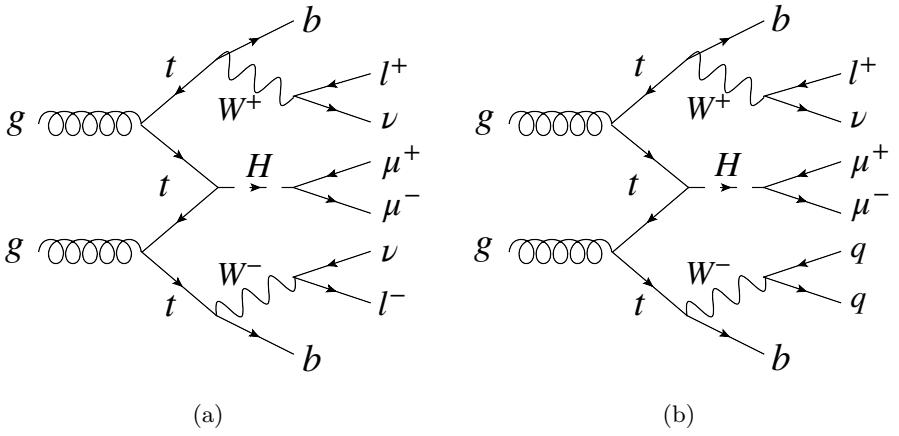


Figure 4.3 Feynman diagrams of the H → $\mu\mu$ decay in the ttH processes. The $t\bar{t}$ dileptonic (a) and semileptonic (b) decays are considered ($\ell = e$ or μ). In the final state for the dileptonic process, there are two b-jets, additional two charged leptons, and two neutrinos in addition to the opposite-charge muon pair from the Higgs boson decay. The final state for the semileptonic process has the muon pair, two b-jets, an additional charged lepton, one neutrino, and two jets.

Table 4.1 The expected signal yield, the final state particles considered in this analysis, and the main background in each production mode ($\ell = e$ or μ). The expected signal yields are calculated by the cross section times the branching ratio in the $\sqrt{s} = 13$ TeV 139 fb^{-1}.

Process		Expected signal yield	Final state particles	Main background
ggF		1464	μ^+, μ^-	Drell-Yan
VBF		114	μ^+, μ^-, 2 jets	Drell-Yan, Z + 2 jets
$W(\to \ell\nu)H$		8.8	μ^+, μ^-, ℓ, ν	WZ
$Z(\to \ell\ell)H$		1.8	$\mu^+, \mu^-, 2\ell$	ZZ
ttH	$t\bar{t}$ semileptonic	5.1	$\mu^+, \mu^-, b, \bar{b}, \ell, \nu$, 2 jets	$t\bar{t}, t\bar{t}Z$
	$t\bar{t}$ dileptonic	0.69	$\mu^+, \mu^-, b, \bar{b}, 2\ell, 2\nu$	

$Z \to \mu\mu$ decay together with two jets (referred to as "Z + 2 jets") is also a background in the VBF process. For the VH process (WH or ZH), there are three or four charged leptons in the final state. The dominant background for the WH and ZH processes is $WZ \to \ell\nu\mu\mu$ and $ZZ \to \mu\mu\ell\ell$, respectively. The ttH process is characterized by the b-jets and additional leptons. The $t\bar{t}$ dimuon decay ($t\bar{t} \to b\bar{b}W^+W^- \to b\bar{b}\mu^+\mu^-\nu\nu$) and the $t\bar{t}Z$ process, which consists of $t\bar{t}$ dimuon decay with $Z \to \ell\ell$ decay and $t\bar{t}$ dileptonic decay with $Z \to \mu\mu$ decay, are dominant background for the ttH process.

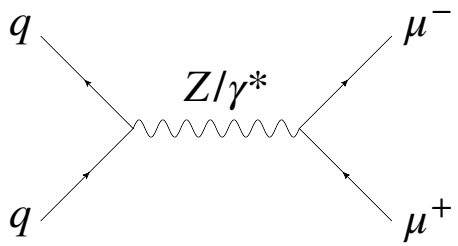


Figure 4.4 Feynman diagram of the Drell-Yan process.

In addition to the final state particles described above, photons are used in this analysis to reconstruct the photon from the QED FSR. In order to improve the dimuon invariant mass resolution, the FSR photons are reconstructed and their four-momentum is included in the dimuon invariant mass calculation. This correction is referred to as "FSR recovery" and described in Section 4.4.

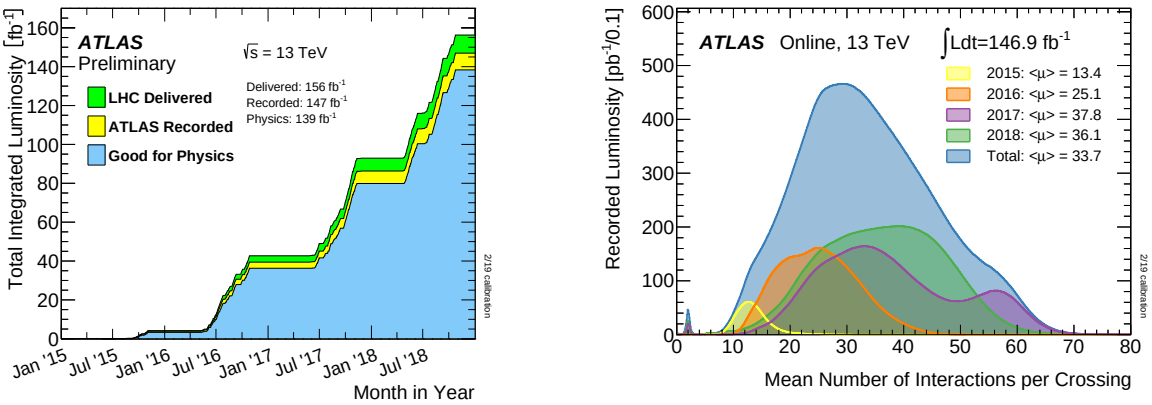
4.1 Data and simulation samples

4.1.1 Collision data

The dataset used in this analysis was recorded during the data taking periods in 2015–2018. In all data taking periods, the correct performance of all detectors is checked and the data whose quality is determined to be bad are removed. The integrated luminosity before and after removing the bad quality data is shown in Table 4.2. Total integrated luminosity of good quality data used for this analysis is 139 fb^{-1}. Figure 4.5 shows total integrated luminosity versus time and the average pile-up during stable beams for pp collisions [22].

Table 4.2 The integrated luminosity recorded by ATLAS and certified to be good quality data in 2015–2018.

Year	ATLAS recorded [fb^{-1}]	Good quality [fb^{-1}]
2015	3.9	3.2
2016	35.6	33.0
2017	46.9	44.3
2018	60.6	58.5
Total	147.0	139.0



(a) (b)

Figure 4.5 Total integrated luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at $\sqrt{s} = 13$ TeV in 2015–2018 (a). The average pile-up for the data (b). [22]

In the $H \to \mu\mu$ process, the final state has two muons, and the leading muon has relatively high momentum ($p_T \sim 60$ GeV). The events for this analysis are collected using two single muon triggers which require to have at least one high-p_T muon [23]. The single muon triggers used for this analysis are summarized in Table 4.3. For the data in 2015, the "OR" combination of "HLT_mu20_iloose_L1MU15" and "HLT_mu50" is used. The "HLT_mu50" trigger requires to satisfy the requirement for p_T greater than 20 GeV for the L1 muon trigger system ("L1_MU20") and to satisfy p_T greater than 50 GeV for the HLT muon trigger. For muons with $p_T < 50$ GeV, the "HLT_mu20_iloose_L1MU15" trigger which has an isolation requirement from other ID tracks is used to reduce the event rate. The "HLT_mu20_iloose_L1MU15" trigger requires that a muon with $p_T > 15$ GeV is selected by the L1 muon trigger system ("L1_MU15") and to pass the "HLT_mu20_iloose" trigger at the HLT system. For the events that pass the "L1_MU15" trigger, the "HLT_mu20_iloose" trigger requires that the muon at the HLT level has $p_T > 20$ GeV and satisfies loose isolation selection criteria using online ID track reconstruction by the HLT. For the data in 2016–2018, the "OR" combination of "HLT_mu26_ivarmedium" and "HLT_mu50" is used to cope with higher luminosity than 2015. The "HLT_mu26_ivarmedium" trigger requires to pass "L1_MU20" and to satisfy p_T threshold of 26 GeV in the HLT muon trigger and to pass a medium isolation selection criteria using online ID track reconstruction by the HLT.

Table 4.3 The single muon triggers and the requirements. The "OR" combination of two single muon triggers is used in each data taking period.

Year	Name	Requirements	
		L1	HLT
2015	HLT_mu20_iloose_L1MU15	$p_T > 15$ GeV	$p_T > 20$ GeV and isolation
	HLT_mu50	$p_T > 20$ GeV	$p_T > 50$ GeV
2016–2018	HLT_mu26_ivarmedium	$p_T > 20$ GeV	$p_T > 26$ GeV and isolation
	HLT_mu50	$p_T > 20$ GeV	$p_T > 50$ GeV

The trigger efficiencies of the "OR" combination of "HLT_mu20_iloose_L1MU15" and "HLT_mu50", and the "OR" combination of "HLT_mu26_ivarmedium" and "HLT_mu50" are summarized in Figure 4.6 and Figure 4.7, respectively. The efficiencies are measured using the tag-and-probe method for the events with the $Z \to \mu\mu$ candidate [23]. In the tag-and-probe method, one muon of the decay, referred to as the "tag", requires to satisfy stringent identification criteria and to fires the trigger. The second muon candidate, referred to as the "probe", is used to study the efficiency of a certain reconstruction algorithm or selection criteria. The probes are usually required to be reconstructed by a system independent of the one being studied.

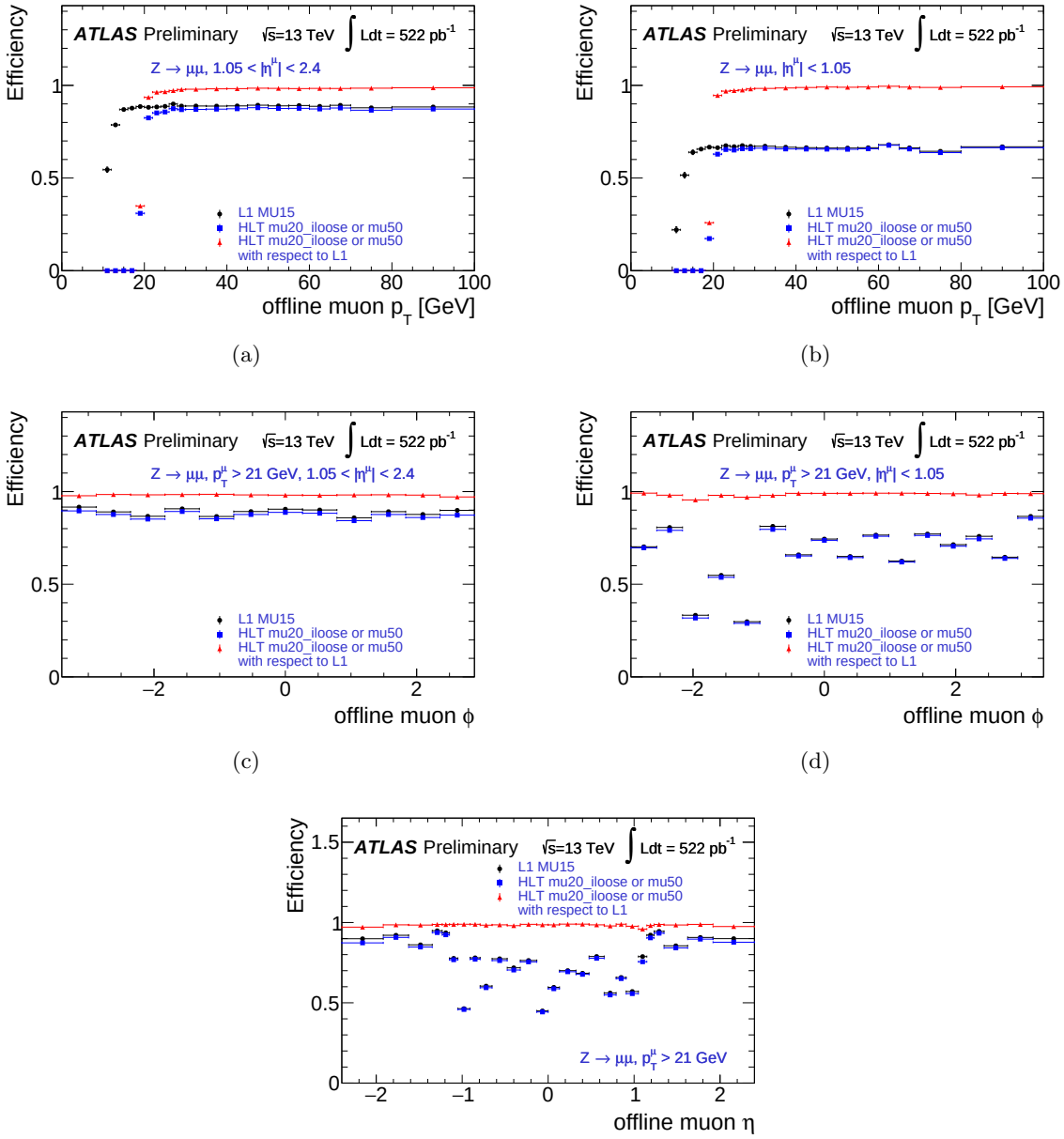


Figure 4.6 Efficiencies of the muon triggers used for the data taking in 2015 in this analysis. The efficiencies as a function of p_{T} in end-cap region (a), p_{T} in barrel region (b), ϕ in end-cap region (c), ϕ in barrel region (d), and η (e) are measured using events with $Z \to \mu\mu$. Black circles are for the absolute efficiencies of "L1MU15". Blue squares are for the absolute efficiencies of the "OR" combination of "HLT_mu20_iloose" and "HLT_mu50". Red triangles are for the relative efficiencies of the "OR" combination of "HLT_mu20_iloose" and "HLT_mu50". The error bars show only statistical uncertainties. [23]

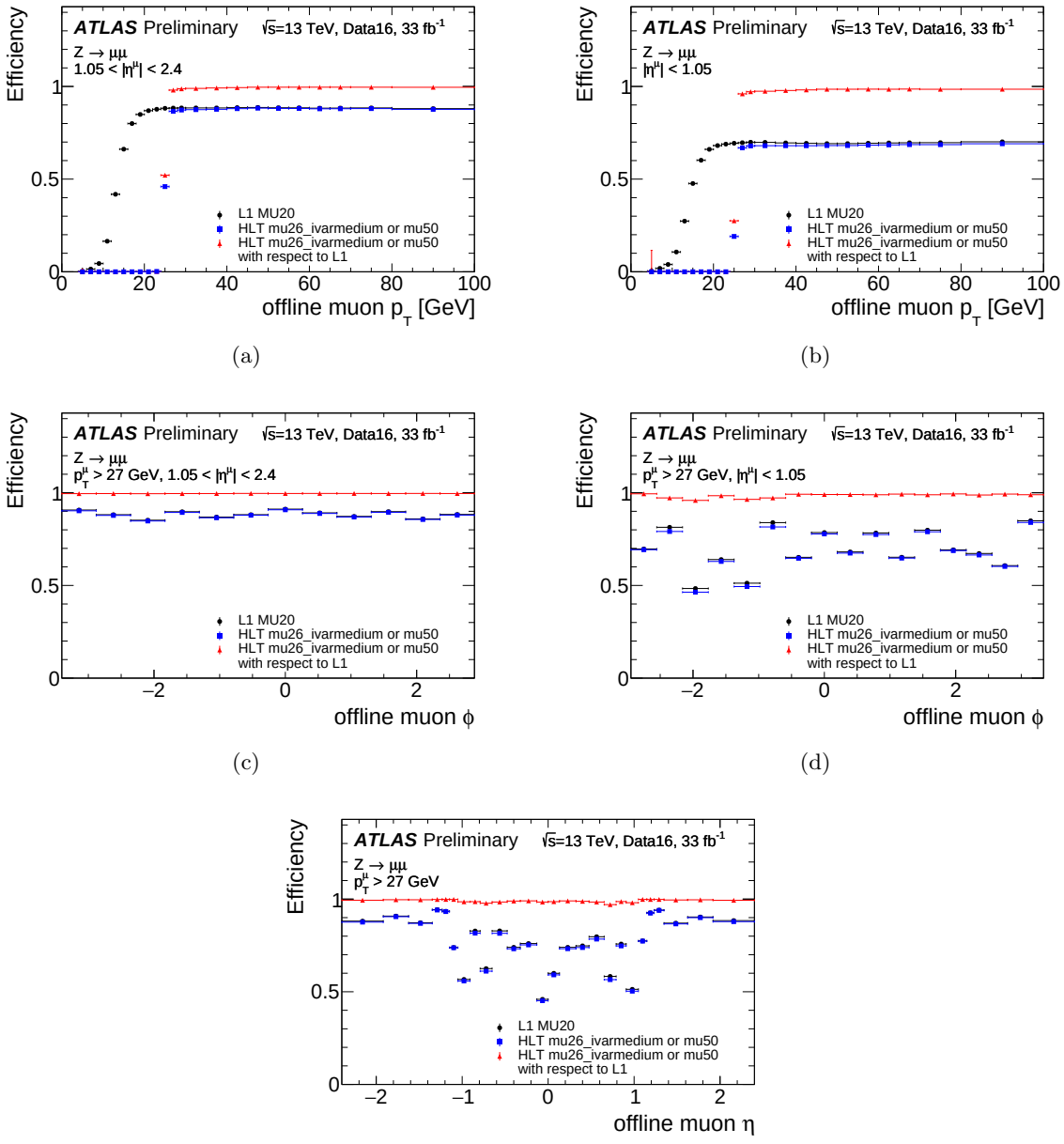


Figure 4.7 Efficiencies of the muon triggers used for the data taking in 2016–2018 in this analysis. The efficiencies as a function of p_{T} in end-cap region (a), p_{T} in barrel region (b), ϕ in end-cap region (c), ϕ in barrel region (d), and η (e) are measured using events with $Z \to \mu\mu$ for the data taken in 2016. Black circles are for the absolute efficiencies of "L1_MU20". Blue squares are for the absolute efficiencies of the "OR" combination of "HLT_mu26_ivarmedium" and "HLT_mu50". Red triangles are for the relative efficiencies of the "OR" combination of "HLT_mu26_ivarmedium" and "HLT_mu50". The error bars show only statistical uncertainties. [23]

4.1.2 Monte Carlo simulation

For optimization of event selection and categorization, signal and background modeling, and estimation of systematic uncertainties, simulated samples are used in this analysis. Physics processes are simulated by event generators using a Monte-Carlo (MC) method. The samples are generated by three steps: event generation, pile-up simulation, and detector simulation. All steps are performed within the ATLAS software framework ATHENA [24].

The simulated samples have three levels of information: truth, detector, and reconstruction. These correspond to the information before detector simulation, after detector simulation, and after object reconstruction, respectively.

Event generation

The event involved in pp collision is generated at parton-level by MC event generators with PDF sets. In the event generator, the fixed-order predictions which consider the perturbative expansion for additional real-emission and virtual corrections at the lower order are used. The parton-level events are processed to consider the higher-order effects from parton shower, hadronization and underlying event. In the process, the parton shower approximation [6] is applied to the event to provide parton showers. The hadronization and underlying event are simulated using phenomenological models (see Section 2.2) tuned by experimental results.

The main MC samples and the production setup in this analysis are summarized in Table 4.4. The signal samples are generated with the Higgs boson mass $m_H = 125$ GeV and the width $\Gamma = 4.07$ MeV. These signal samples are normalized with the latest available theoretical calculations corresponding to the SM production cross sections.

Table 4.4 Setup for MC sample production. Event generator, PDF set, accuracy in QCD perturbation theory, interface of PS simulation, and tuned parameters set used for PS are summarized for each sample.

Process	Generator	PDF	Accuracy	PS	Tune
Signal					
ggF	Powheg NNLOPS [25,26]	PDF4LHC15 [27]	NNLO	Pythia 8 [28]	AZNLO [29]
VBF	Powheg-Box [30–32]	PDF4LHC15	NLO	Pythia 8	AZNLO
$q\bar{q}/qg \to VH$	Powheg-Box	PDF4LHC15	NLO	Pythia 8	AZNLO
$gg \to ZH$	Powheg-Box	PDF4LHC15	LO	Pythia 8	AZNLO
ttH	MadGraph5_aMC [33]	NNPDF3.0NLO [34]	NLO	Pythia 8	A14 [35]
Background					
Drell-Yan	Serpa v2.2.1 [36]	NNPDF3.0NNLO	NLO / LO [a]	Serpa PS [37]	MEPS@NLO
Diboson (semi-leptonic)	Serpa v2.2.1	NNPDF3.0NNLO	NLO / LO [b]	Serpa PS	MEPS@NLO [38,39]
Diboson (dileptonic)	Serpa v2.2.2	NNPDF3.0NNLO	NLO / LO [b]	Serpa PS	MEPS@NLO
$Z + jj$	Serpa v2.2.1	NNPDF3.0NNLO	NLO / LO [a]	Serpa PS	MEPS@NLO
$t\bar{t}$	Powheg-Box	NNPDF3.0NLO	NLO	Pythia 8	A14
single-t	Powheg-Box	NNPDF3.0NLO	NLO	Pythia 8	A14
$t\bar{t}V$	MadGraph5_aMC	NNPDF3.0NLO	NLO	Pythia 8	A14
$H \to Z(\to \ell\ell)\gamma$	Powheg-Box	PDF4LHC15	NNLO	Pythia 8	AZNLO

[a] NLO (LO) accuracy for up to two (four) partons calculated with the Comix [40] and OpenLoops [41,42]
[b] NLO (LO) accuracy for up to one (two or three) additional partons

Pile-up simulation

In order to simulate the effects of the pile-up, soft QCD events are generated using Pythia 8 with the NNPDF2.3LO set [43] and the A3 set of tuned parameters [44]. They are overlaid to each hard-interaction event generated. The overlaid events are reweighted to data to correct the pile-up distribution.

Detector simulation

The generated events are processed through the full ATLAS detector simulation [45] based on a detailed description of the detector geometry and of the particle interaction simulation in the detector material with GEANT4 [46].

High statistics DY samples

The samples above provide an equivalent integrated luminosity that is typically 5–20 times higher than that of data. However, the statistical uncertainties in the dominant DY background limit

the classification and background modeling study for a small signal yield. High statistics DY samples are produced using approximate detector simulation since the full detector simulation needs too large CPU time. Two event generation setups were developed for high statistics DY samples.

The primary fast simulation DY sample is parton-level generated by Sherpa 2.2.4 at LO accuracy for up to three additional partons with the CT14 NNLO PDF set [47]. The parton-level generated events are processed with Pythia 8 to provide parton shower and hadronization. The CKKW-L merging algorithm [48] with a merging scale of 20 GeV is used to remove double counting between this sample and the main DY sample.

For further cross-checks, an additional fast simulation DY sample is prepared. In the additional sample, events for inclusive $Z/\gamma^* + 0, 1$ parton and for $Z/\gamma^* + 2$ partons are separately simulated. The former events are parton-level generated using Powheg-Box with the CT10 PDF set [49]. For the latter ones, Alpgen [50] with the CTEQ6L1 PDF set [51] is used. These events are processed with an approximate QCD parton shower algorithm using Pythia 8 and the overlaps between these samples are removed. The QED FSR is provided by Photos [52].

For high statistics DY samples, detector responses are emulated in an approximate way by applying parameterized "smearing functions" to truth-level objects. The parametrization is applied to the detector responses to muons, FSR photons, jets, and the missing transverse momentum. The smearing functions for each objects are derived using full simulation sample. In total, two samples were prepared with significantly large statistics corresponding to an integrated luminosity of at least 50 ab^{-1} in the kinematic phase space relevant for the analysis.

4.2 Reconstruction of final state particles

4.2.1 Track

In the ATLAS detector, muons are reconstructed by two independent detector systems. The inner detector (ID) track is the charged-particle trajectory reconstructed using the hits in the Pixel, SCT and TRT detectors. The muon spectrometer (MS) track is used for muon identification, which is described in Section 4.2.3. The details of the reconstruction procedure of the ID track can be found in Ref. [53]. The procedure consists of the following steps. Pixels and strips with detected hits are grouped into clusters by connected component analysis [54]. From these clusters, three-dimensional hit points referred to as "space-points" are created. After cluster creation, the track seeds are formed from three space-points. The track seeds need to pass the requirements specified to maximize purity, which is defined as a fraction of seeds that provide good quality tracks. Then, the track candidates are reconstructed from the track seeds using the Kalman filter [55]. For each track candidate, the track score is calculated based on the track quality such as the resolution of subdetectors, multiplicity of clusters and holes on the track trajectory, and the χ^2 from the track fit. Here a hole is defined as an active sensor that registers no-hit despite it on the track trajectory. If a cluster is shared among multiple track candidates, a neural network classifier [56] is used to identify whether the cluster was created by multiple tracks or not, and then the track score of the candidate is re-calculated if needed. Track candidates with a bad score are rejected and the remaining candidates that don't fulfill additional basic quality criteria based on p_{T}, η, number of clusters and holes, and distance of the track from the interaction point are also rejected. For the track candidates fulfilling all the requirements mentioned above, high-precision fit using all available information is performed. In the end, the tracks are extended into the TRT detector, and refit is performed only in the case in which compatible hits on the TRT are found. The details of the extension can be found in Ref. [53].

4.2.2 Primary vertex

There are multiple pp collisions per bunch crossing. The positions of pp collisions are referred to as "vertex". The vertex that seems to be from hard scattering, referred to as "primary vertex", is selected from reconstructed vertices. The vertex reconstruction and definition of the primary vertex are described below.

The vertices are reconstructed by collecting and combining reconstructed charged tracks in the ID [57]. The ID tracks which satisfy the following quality requirements are considered in vertex reconstruction:

- $p_{\mathrm{T}} > 400$ MeV
- $|\eta| < 2.5$
- Number of silicon (SCT or Pixel) hits $\geq \begin{cases} 9 & \text{if } |\eta| \leq 1.65 \\ 11 & \text{if } |\eta| > 1.65 \end{cases}$

- IBL hits + B-layer hits ≥ 1
- A maximum of 1 shared module (1 shared pixel hit or 2 shared SCT hits)
- No Pixel holes
- SCT holes ≤ 1

A hole is defined as an active sensor, which registers no-hit, on the trajectory.

After the selection of ID tracks, the combination of vertex and tracks are determined by the following steps.

- Definition of seed position: A vertex seed position is found by searching for the global maximum in the distribution of z coordinates of the tracks.
- Vertex fitting: The vertex position is found using the adaptive vertex fitting algorithm [58], which is an iterative χ^2 minimization based fitting. The seed position and the tracks around the seed are used as inputs of the algorithm.
- Compatibility check: The compatibility of track with the vertex is checked for each track. The weights for the fitting algorithm of incompatible tracks are reduced.
- Compatible track finding: By repeating the vertex fitting and compatibility check, all compatible tracks with the vertex are searched for. Incompatible tracks are removed in the vertex to use for other vertices.
- Repeat above steps until no unassociated tracks are left in the corresponding event or no additional vertex can be found. All vertices require to have at least two associated tracks.

The vertex positions are calculated using only the compatible tracks. There are approximately 10–30 reconstructed vertices per bunch crossing.

The vertex with the largest sum of squared p_{T} of contributing tracks (Σp_{T}^2), referred to as "primary vertex", provides significant information and is used for physics analysis.

4.2.3 Muon

Muons used for this analysis are required to pass three steps, reconstruction, identification, and isolation. In the reconstruction step, the MS track is reconstructed first and the combined track is reconstructed using the MS track and the information from other subdetectors. In the isolation step, the isolated muons are selected to suppress the muons associated with jets from $b\bar{b}$ and $c\bar{c}$ processes, which are a large cross section. In addition, the momentum calibration of the reconstructed muons is performed.

Track reconstruction in the muon spectrometer

Muons are reconstructed using the information of the ID tracks and the MS tracks, which are reconstructed independently. The MS tracks are reconstructed using the information of the muon spectrometer, and the reconstruction procedure consists of the following two steps [59, 60].

The first step is a segment finding. In each station of the MDT and CSC, the segments are formed using the hits inside the detector. For the MDT segments, a Hough transform [61] is used to search for hits on a straight trajectory in the bending plane of the detector. The segment position in the coordinate orthogonal to the bending plane is provided using hits in the RPC or TGC. For the CSC segments, a separate combinatorial search in the η and ϕ detector planes is used.

After the segment finding, track candidates are built by fitting hits associated with segments on the different stations. The segments used for the candidate build need to fulfill the requirements based on hit multiplicity on the detector and fit quality. To build the track candidates, at least two selected segments are required, except for the endcap–barrel transition region. The track candidates are refitted using all hits associated with the candidate by a global χ^2 fit. The track candidates for which the χ^2 satisfies the selection criteria are accepted, and the candidates with large χ^2 are removed.

Combined reconstruction

The reconstruction of muons is performed using the information from the ID tracks, the MS tracks and the energy deposit in the calorimeter. The following four muon types are defined depending on which subdetector information is used for reconstruction:

- Combined (CB) muons: Based on the ID and MS tracks, the combined track is formed by a global refitting to hits from both the ID and MS subdetectors. In the global fit, hits on the muon spectrometer associated with the track may be updated and the track fit is repeated to improve the fit quality. Most muons are reconstructed using the MS tracks and the ID tracks matched to the MS track extrapolated inward, referred to as "outside-in algorithm". In order to recover efficiency for muons such as the low-p_{T} muons which may not reach the middle MS station, the complementary "inside-out algorithm", in which the ID tracks are extrapolated outward and MS hits matched to the ID tracks are searched for, is used.

- Segment-tagged (ST) muons: ST muons are reconstructed using the ID tracks and the MDT or CSC local segments matched to the ID track extrapolated outward. In order to efficiently reconstruct muons, ST muons are used when muons cross only one MS station due to the limited MS coverage. The parameters of ST muons are taken from the ID track fit.

- Calorimeter-tagged (CT) muons: CT muons are reconstructed using the information from the ID tracks and the calorimeter. CT muons are identified when the calorimeter energy deposits consistent with minimum-ionizing particles are matched to the outward extrapolated ID track. The CT muon reconstruction type has the lowest purity than other types, but it recovers acceptance in the region $|\eta| < 0.1$ where the MS is installed partially. Whereas other reconstruction algorithms use the ID track with $p_{\mathrm{T}} > 2$ GeV, the CT algorithm uses the ID track with $p_{\mathrm{T}} > 5$ GeV since the purity gets worse at the low-p_{T}. The parameters of CT muons are taken from the ID track fit.

- Extrapolated (ME) muons: ME muons are reconstructed based on only the MS tracks which are compatible with originating from the interaction point when the MS tracks cannot be matched to ID tracks. The ME muons are mainly used in the forward region $2.5 < |\eta| < 2.7$, where the inner detector does not cover. Whereas the MS track requires to traverse two MS stations in general, the MS tracks used for the ME reconstruction in the forward region are required to traverse three MS stations. The parameters of ME muons are defined at the interaction point, taking into account the estimated muon energy loss in the calorimeters.

The transverse momentum resolution of the CB muons is shown in Figure 4.8. All reconstruction types of muons are used for this analysis.

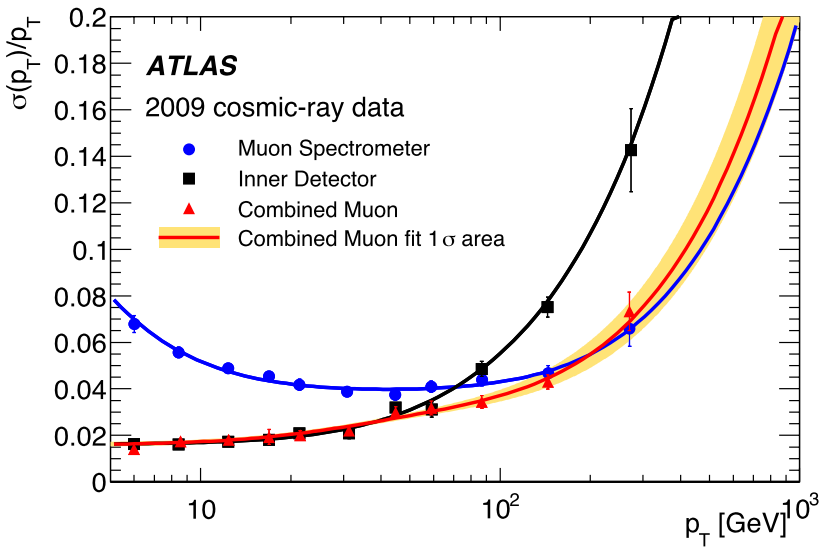


Figure 4.8 The transverse momentum resolution $\sigma(p_{\mathrm{T}})$ on relative p_{T} as a function of p_{T} for ID tracks (black squares), MS tracks (blue circles), and CB tracks (red triangles). [62]

Identification

In order to select high-quality muon candidates used for physics analysis, the muon identification is performed by applying quality requirements. The requirements are decided so as to suppress background, which is mainly in-flight decays of charged light hadrons such as pion and kaon. The background muons are characterized by the poor fit quality of the combined track and the difference in the momentum measurement between ID and MS. For the CB, ST and CT muons, the ID tracks are required to satisfy the following requirements to guarantee the quality of tracks used for identification.

- at least one hit in the Pixel,

- at least five hits in the SCT,
- less than three holes in total in the Pixel or SCT.

Three muon identification working points (WPs) are provided for different physics analyses: Medium, Loose, and Tight.

The Loose, Medium and Tight WPs are standard WPs and have a higher purity in the order. These are inclusive categories (Loose $\subset$ Medium $\subset$ Tight). For the definition of these WPs, the following variables are used.

- number of precision stations defined as the number of different MS stations which have at least three hits
- q/p significance defined for CB muons:

$$q/p \text{ significance} = \frac{|(q/p)_{\mathrm{ID}} - (q/p)_{\mathrm{MS}}|}{\sqrt{\sigma^2(q/p)_{\mathrm{ID}} + \sigma^2(q/p)_{\mathrm{MS}}}} \tag{4.1}$$

  where $(q/p)_{\mathrm{ID}}$ and $(q/p)_{\mathrm{MS}}$ are the ratio of the charge q and momentum p of muons measured in the ID and MS, while $\sigma^2(q/p)_{\mathrm{ID}}$ and $\sigma^2(q/p)_{\mathrm{MS}}$ are the corresponding uncertainties.

- ρ' defined for CB muons:

$$\rho' = \frac{|p_{\mathrm{T,ID}} - p_{\mathrm{T,MS}}|}{p_{\mathrm{T,CB}}} \tag{4.2}$$

  where $p_{\mathrm{T,ID}}$ and $p_{\mathrm{T,MS}}$ are muon p_{T} measured in the ID and MS, while $p_{\mathrm{T,CB}}$ is muon p_{T} of the combined track.

- normalized χ^2 of the combined track fit

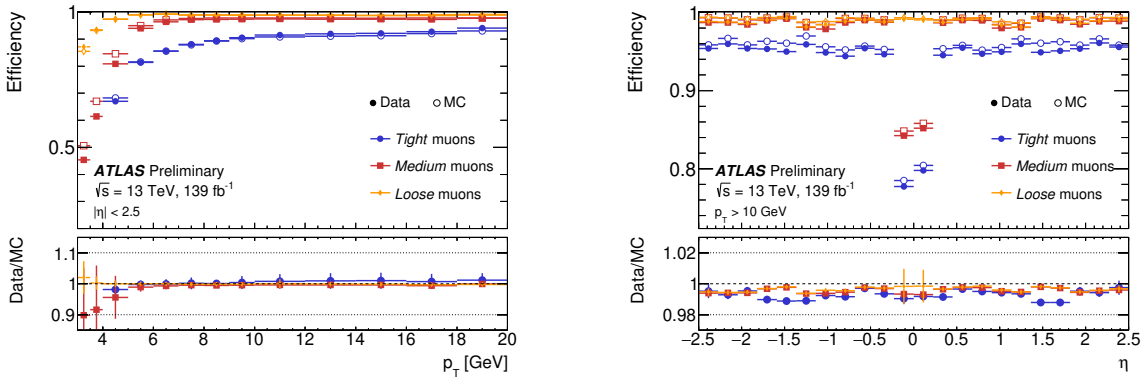
The criteria of the Loose, Medium and Tight WPs are described in the following:

Medium muons The medium WP minimizes the systematic uncertainty on muon reconstruction. Only CB tracks within the ID acceptance $|\eta| < 2.5$ and ME tracks in the $2.5 < |\eta| < 2.7$ region are used. In order to ensure the track quality, there are requirements on the number of precision stations for CB and ME tracks. CB tracks are required to have at least two precision stations for the $|\eta| > 0.1$ region, and to have at least one precision station but no more than one precision hole station for the $|\eta| < 0.1$ region. For ME tracks, at least three precision stations are required. In addition, q/p significance is to be less than 7 to suppress the background.

Loose muons The loose WP is designed to maximize the reconstruction efficiency keeping good-quality muon tracks. It is optimized for the reconstruction of the Higgs boson candidates decaying in the four-muon final state. For higher reconstruction efficiency, CB, ST, CT and ME tracks are used. For CB and ME tracks, the same criteria as medium WP is used. For ST and CT tracks, muons are restricted to the $|\eta| < 0.1$ region.

Tight muons The tight WP provides the highest purity at the cost of a few percent efficiency loss. Only CB tracks with at least two precision stations and satisfying the medium criteria are used. The normalized χ^2 of the combined track fit is to be less than 8 to remove pathological tracks. In order to provide better background rejection for lower-p_{T} muon, in which higher background is expected, the two-dimensional cut in the q/p significance and ρ' is performed depending on the muon p_{T} and $|\eta|$.

The reconstruction and identification efficiencies as a function of p_{T} and $|\eta|$ for the Loose, Medium and Tight WPs are shown in Figure 4.9. The hadron misidentification rates are summarized in Table 4.5. The Loose WP is used for this analysis to secure higher signal efficiency.



(a) (b)

Figure 4.9 The muon reconstruction and identification efficiencies as a function of p_{T} (a) and η (b) for the Loose, Medium, and Tight WPs. The p_{T} dependence is measured in $J/\psi \to \mu\mu$ events and the η dependence is measured in $Z \to \mu\mu$ events for muons with $p_{\mathrm{T}} > 10$ GeV.

Table 4.5 The muon efficiency ϵ_μ and the hadron misidentification rate ϵ_{had}. These are evaluated in a $t\bar{t}$ MC sample in different p_{T} regions for muons in $|\eta| < 2.5$.

Working point	5 GeV $< p_{\mathrm{T}} <$ 20 GeV		20 GeV $< p_{\mathrm{T}} <$ 100 GeV		$p_{\mathrm{T}} >$ 100 GeV	
	ϵ_μ [%]	ϵ_{had} [%]	ϵ_μ [%]	ϵ_{had} [%]	ϵ_μ [%]	ϵ_{had} [%]
Loose	98	1.06	99	0.25	98	0.12
Medium	97	0.85	97	0.17	97	0.07
Tight	90	0.38	93	0.12	93	0.04

Isolation

A track-based isolation variable and a calorimeter based isolation variable are used to select tracks isolated from other tracks or clusters [60]. The track-based isolation variables are defined as the scalar sum of p_{T} of the ID tracks in a cone around the muon p_{T} (p_{T}^{μ}) excluding the muon track itself. The cone is defined by $\Delta R = 0.2$ and $\Delta R = \min(10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu}, 0.3)$, labeled as $p_{\mathrm{T}}^{\mathrm{cone20}}$ and $p_{\mathrm{T}}^{\mathrm{varcone30}}$, respectively. For the calorimeter based isolation variable, $E_{\mathrm{T}}^{\mathrm{topocone20}}$ is used. It is defined as the sum of E_{T} of the topological cluster described in Section 4.2.4 in a cone of size $\Delta R = 0.2$ around the muon position after subtracting the muon energy deposit itself and correcting for pile-up effects. The track-based isolation has better resolution and lower pile-up dependence. On the other hand, the calorimeter-based isolation takes into account neutral particles and charged particles with p_{T} below the threshold in the ID track reconstruction. Since both isolation variables have the contribution from charged particles, the particle flow algorithm [63] is used for removing the overlap. In the particle flow algorithm, the ID tracks are extrapolated to the calorimeter regions, and the corresponding energy is removed from the measured calorimeter energy. The variable $E_{\mathrm{T}}^{\mathrm{neflow20}}$, defined as the E_{T} of neutral particle-flow objects in a cone of size $\Delta R = 0.2$ around the muon, is used as particle-flow-based isolation variable. In the particle-flow-based isolation, $p_{\mathrm{T}}^{\mathrm{cone20}}$ and $p_{\mathrm{T}}^{\mathrm{varcone30}}$ are also used for the muon with $p_{\mathrm{T}}^{\mu} > 50$ GeV and $p_{\mathrm{T}}^{\mu} < 50$ GeV, respectively. To correct the contribution from the muon energy deposit itself and pile-up effects, a weighting factor $w = 0.4$ is applied to $E_{\mathrm{T}}^{\mathrm{neflow20}}$.

There are several isolation WPs defined using the above isolation variables [60]. To ensure high efficiency for muon suppressing the background of muons from light-hadrons decay, muons are required to satisfy the following isolation criteria in this analysis. The track p_{T} needs to be $p_{\mathrm{T}} > 500$ MeV. The muons with $p_{\mathrm{T}}^{\mu} < 50$ GeV are required to satisfy:

$$p_{\mathrm{T}}^{\mathrm{varcone30}} + 0.4 \cdot E_{\mathrm{T}}^{\mathrm{neflow20}} < 0.16 \cdot p_{\mathrm{T}}^{\mu}. \tag{4.3}$$

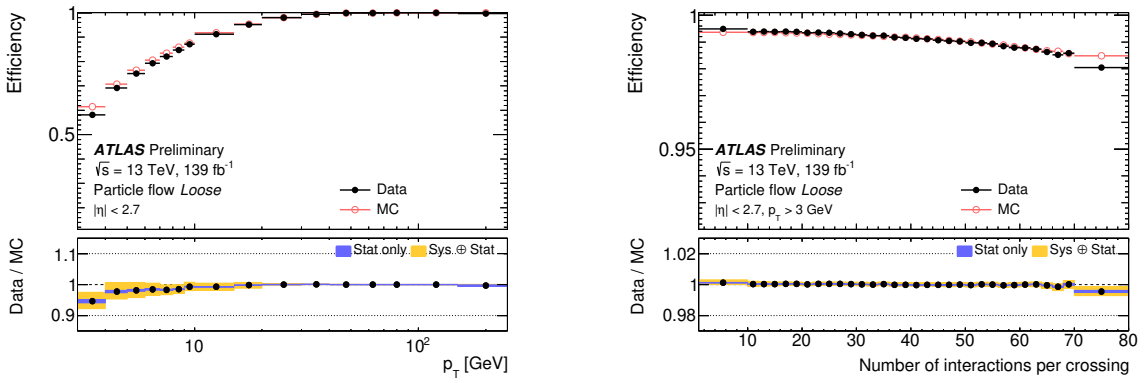
On the other hand, muons with $p_{\mathrm{T}}^{\mu} > 50$ GeV need to satisfy:

$$p_{\mathrm{T}}^{\mathrm{cone20}} + 0.4 \cdot E_{\mathrm{T}}^{\mathrm{neflow20}} < 0.16 \cdot p_{\mathrm{T}}^{\mu}. \tag{4.4}$$

The isolation efficiency for this WP is shown in Figure 4.10.

Calibration

The muon momentum scale and resolution are calibrated to take into account the fluctuation of the energy loss in the traversed material and the fluctuation of the energy loss, multiple scattering, and local magnetic field inhomogeneities [59]. Since the simulation cannot explain data perfectly, the momentum scale predicted by the simulation differs at the per mille level from that of the data. The momentum resolution is different at the per cent level between the simulation and the data. The simulated muon momentum scale and resolution are corrected to agree with data [59]. The correction factors are evaluated with the $J/\psi \to \mu\mu$ events for low-p_{T} muon and the $Z \to \mu\mu$ events for high-p_{T} muon. The correction factors are extracted using the dimuon invariant mass distributions in each η, ϕ, and p_{T} region. The dimuon invariant mass distributions before and after the correction are shown in Figure 4.11. After the correction, a good agreement between data and simulation is obtained.



(a) (b)

Figure 4.10 The muon isolation efficiencies for the WP used in this analysis. These are measured in $Z \to \mu\mu$ events for muons with $p_{\mathrm{T}} > 3$ GeV as a function of p_{T} (a) and the number of interactions per bunch crossing (b).

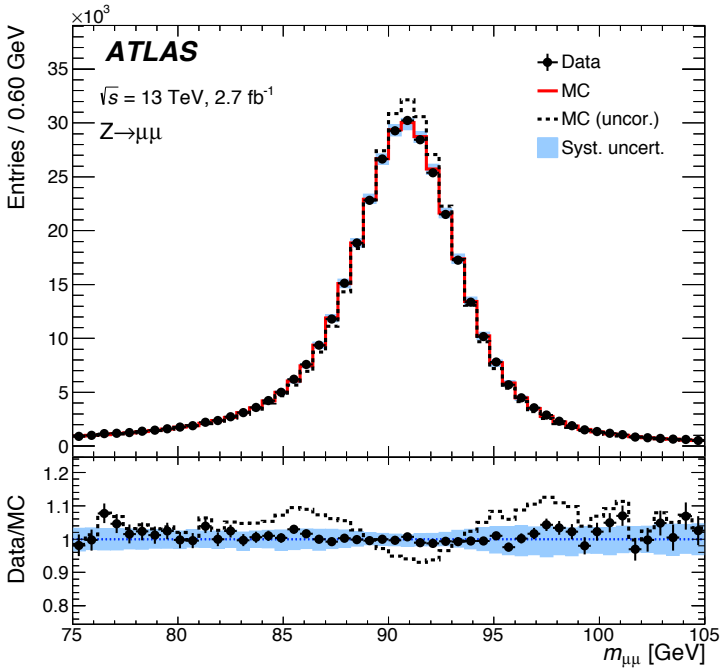


Figure 4.11 The invariant mass of $\mu\mu$ of $Z \to \mu\mu$ candidate events for data and MC simulation before and after the correction. In the upper panel, the black points, red histogram, and black dotted histogram show the data, simulation after the correction, and simulation before the correction, respectively. The lower panel shows the ratio of data to the simulation. The error band shows the effect of the systematic uncertainty on the momentum correction.

4.2.4 Electron

Electrons used for this analysis are required to pass three steps, reconstruction, identification, and isolation. In the reconstruction step, the "topo-cluster" is reconstructed first and used with tracks [64]. In addition, the energy calibration of the reconstructed electrons is performed [65].

Topo-cluster

Topo-clusters are formed from calorimeter cell signals based on seeding and collecting [64]. The clustering algorithm is performed based on the spatial signal significance for each cell $\zeta_{\text{cell}}^{\text{EM}}$, which is defined as the ratio of the cell signal $E_{\text{cell}}^{\text{EM}}$ to the expected noise $\sigma_{\text{noise,cell}}^{\text{EM}}$ in the cell:

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}} \tag{4.5}$$

The procedure of clustering is the following.

- Look for the cell with $\zeta_{\text{cell}}^{\text{EM}} > 4$. The cell is identified as a seed of a cluster.
- The neighboring cells with $\zeta_{\text{cell}}^{\text{EM}} > 2$ are added to the cluster.

The above steps are initiated with the cell with the largest $\zeta_{\text{cell}}^{\text{EM}}$ and repeated until there are no more cells to be added. After the growing is stopped, all neighboring cells of the cluster are added to the cluster. This step allows the cells with signal significance close to the noise levels to be added to the clusters. If there are multiple local signal maxima defined by $E_{\text{cell}}^{\text{EM}} >$ 500 MeV in a cluster, the cluster is split into two clusters so that the two highest-energy cells are separated. These algorithms suppress from electronic noise and fluctuations such as pile-up, keeping signal efficiency high. Topo-clusters are used for not only electron reconstruction but also muon identification, muon isolation, and jet reconstruction.

Reconstruction

The electron reconstruction is performed using the information of the topo-clusters and the track seeds in the inner detectors [65]. The procedure of reconstruction consists of mainly four steps: cluster selection, track reconstruction, track-cluster matching, and supercluster reconstruction.

First, the topo-clusters used for the reconstruction are selected using two requirements. One is that the energy from the cells in the EM calorimeter, referred to as EM energy, is larger than 400 MeV. The other is that the EM fraction f_{EM}, defined as the ratio of the EM energy to total energy of the cluster, is larger than 0.5 to reject pile-up clusters. The clusters which satisfy the above requirements are referred to as "EM topo-clusters".

Tracks of the electron are reconstructed by two steps: pattern recognition and track fit. The pattern recognition algorithm builds the track candidates using the pion and electron hypothesis for the model of energy loss due to the interaction with the detector material. The track candidates are then fitted using the global χ^2 track fitter [66].

After the cluster selection and the track reconstruction, a loose matching of the track to the cluster is required. The tracks with $|\eta_{\text{track}} - \eta_{\text{cluster}}| < 0.05$ and $-0.10 < q \cdot (\phi_{\text{track}} - \phi_{\text{cluster}}) < 0.05$

are considered loosely matched. Here η_{track} and ϕ_{track} are the η and ϕ coordinates of the extrapolated track, and q is the reconstructed charge of the track. η_{cluster} and ϕ_{cluster} are the η and ϕ coordinates of cluster barycentre. The tracks loosely matched to clusters are re-fitted using the optimized Gaussian Sum Filter [67] to take into account the effect of non-linear bremsstrahlung.

In order to consider the energy deposit from bremsstrahlung and the topo-cluster splitting, supercluster is defined. The EM topo-clusters which have EM energy larger than 1 GeV and match to tracks with at least 4 silicon (Pixel and SCT) hits are selected as the supercluster seeds. The satellite clusters are searched within a window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ around the barycenter of seed clusters. The superclusters are formed using seed clusters and the satellite clusters associated with the seeds. Finally, the track matching is performed again using the superclusters by the same method as the matching to EM topo-clusters.

Identification

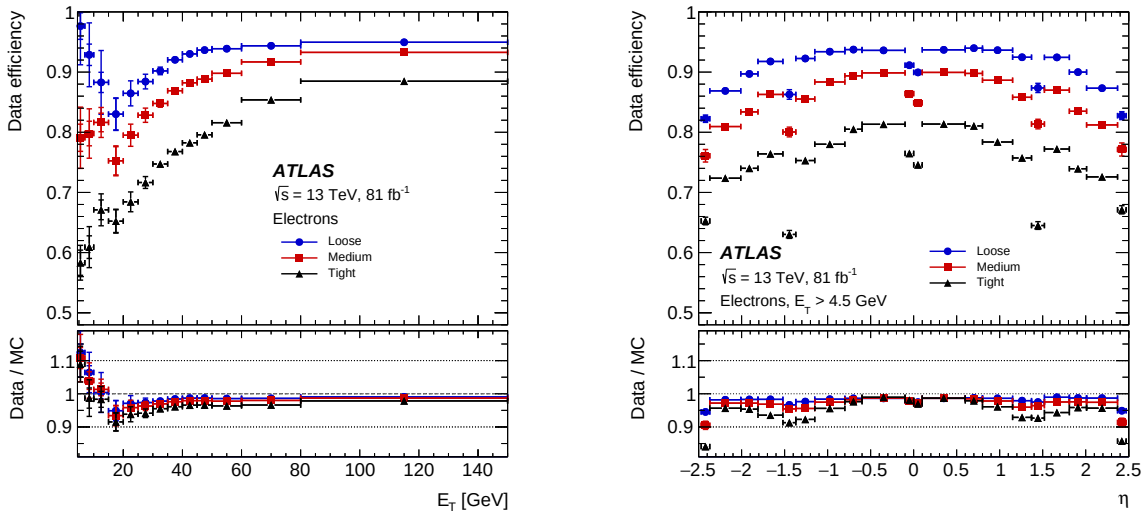
In order to improve the purity of the reconstructed electrons, the electrons are identified by a likelihood discriminant. The likelihoods are calculated from probability density functions (PDFs), which are created using histograms of the quantities measured in the inner detector, the calorimeter and the combination of them. Typically 13 variables, related to energy deposits in the different EM layers and in the hadronic calorimeter, track quality, and track-cluster matching, are used for the PDF creation. The list of all variables is found in Ref. [65]. Four working points are defined on the LH discriminant so that the requirement can be optimized depending on the analyses. The working points are referred to as VeryLoose, Loose, Medium, and Tight. The reconstruction and identification efficiencies as a function of E_{T} and $|\eta|$ for the Loose, Medium and Tight WPs are shown in Figure 4.12. In this analysis, the electrons identified by the Medium WP are used.

Isolation

The electron isolation criteria use a calorimeter-based isolation variable and a track-based isolation variable [65]. The calorimeter-based isolation variable $E_{\text{T}}^{\text{cone20}}$ is the transverse energy of the topo-cluster in a cone of size $\Delta R = 0.2$ around the cluster barycenter after the subtraction of the transverse energy of the energy deposit of electron candidate $E_{\text{T,core}}$. The calculation of $E_{\text{T}}^{\text{cone20}}$ uses the total transverse energy of topo-cluster within the cone, $E_{\text{T,raw}}^{\text{iso20}}$, and $E_{\text{T,core}}$. The $E_{\text{T,core}}$ is derived by the total transverse energy of cells in a window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ around the direction of the electron candidate. The isolation variable $E_{\text{T}}^{\text{cone20}}$ is defined by the following taking into account the leakage of $E_{\text{T,core}}$ to outside of the cone, $E_{\text{T,leakage}}$, and the effects of pile-up, $E_{\text{T,pileup}}$:

$$E_{\text{T}}^{\text{cone20}} = E_{\text{T,raw}}^{\text{iso20}} - E_{\text{T,core}} - E_{\text{T,leakage}} - E_{\text{T,pileup}}. \tag{4.6}$$

The track-based isolation variable, $p_{\text{T}}^{\text{varcone20}}$, is the scalar sum of p_{T} of the ID tracks in a cone of size $\Delta R = \min(10\ \text{GeV}/p_{\text{T}^e}, 0.2)$ around the electron candidate excluding the electron track itself.



(a) (b)

Figure 4.12 Electron reconstruction and identification efficiencies measured in $Z \to ee$ events as a function of E_{T} (a) and η (b) for the Loose, Medium, and Tight criteria. The region of $1.37 < |\eta| < 1.52$ is the crack region.

There are several isolation WPs defined using the above isolation variables. In the Loose WP used for this analysis, electrons are required to satisfy the following requirements:

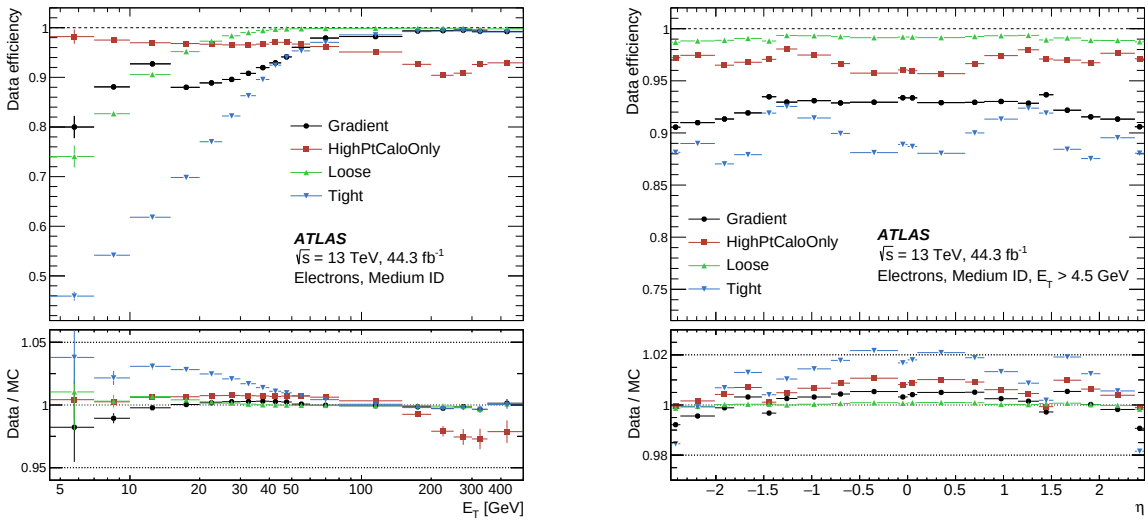
$$E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}}^{e} < 0.20, \tag{4.7}$$

$$p_{\mathrm{T}}^{\mathrm{varcone20}}/p_{\mathrm{T}}^{e} < 0.15. \tag{4.8}$$

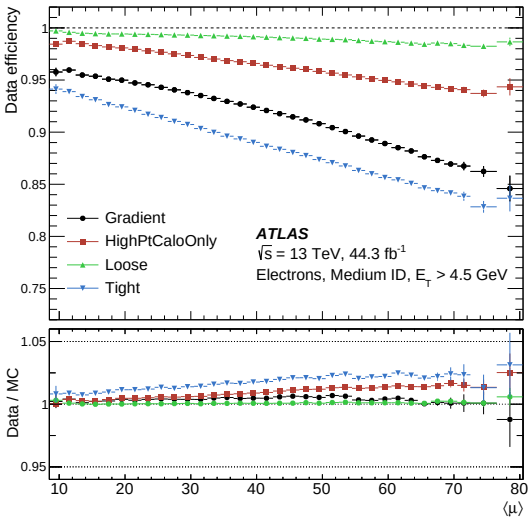
The isolation efficiencies for different WPs are shown in Figure 4.13.

Energy calibration

The energy of electrons is estimated from the energy of clusters by the multivariate regression algorithm tuned on MC. The estimation is performed taking into account the energy loss of the interaction of the material in front of the calorimeter. The energy in data is adjusted to correct the difference in the relative energy scale between different layers of the EM calorimeter. In addition, the non-uniformity of the energy response of the calorimeter due to the geometric effects of the boundary of the calorimeter modules and the different HV setting is corrected. After the above corrections, there are still some differences between the simulated energy and the energy obtained in data. The simulated energy scale and resolution are corrected to agree with data. These energy corrections are validated using $Z \to ee$ events. The invariant mass distribution of two electrons in the $Z \to ee$ events after corrections is shown in Figure 4.14 [65]. After the corrections, good agreement between data and simulation is obtained.



(a) (b)



(c)

Figure 4.13 Electron isolation efficiencies for different criteria measured in $Z \to ee$ events as a function of E_{T} (a), η (b) and the number of interactions per bunch crossing (c). The electrons are identified by Medium criteria.

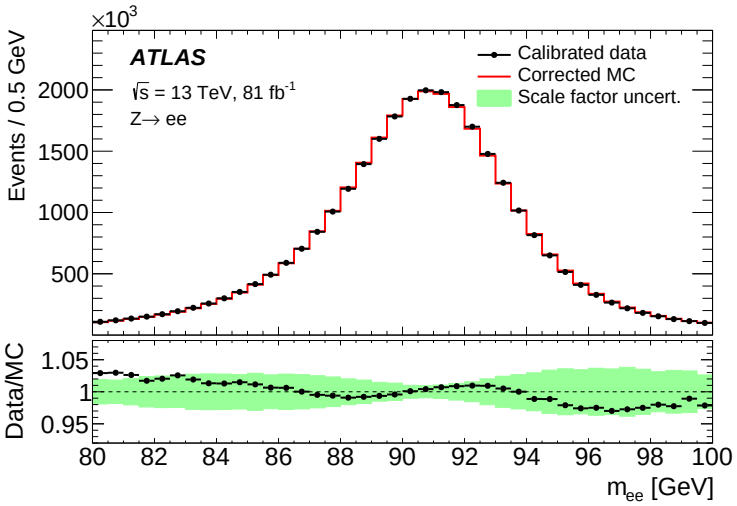


Figure 4.14 The invariant mass of ee of $Z \to ee$ candidate for data (black) and MC (red) aftter the energy calibration.

4.2.5 Photon

In this analysis, photons are used for the reconstruction of FSR photons as described in Section 4.4. The photons are required to pass the three steps: reconstruction, identification, and isolation [65]. The reconstruction is performed by a procedure similar to that of electrons. The energy calibration is also performed using the $Z \to ll\gamma$ decay by a procedure similar to that of electrons.

Reconstruction

Photons are reconstructed by a procedure similar to that of electrons using the information of EM topo-clusters and reconstructed tracks. In the track loose matching step, the cluster matched to tracks with $|\eta_{\text{track}} - \eta_{\text{cluster}}| < 0.05$ and $-0.10 < q \cdot (\phi_{\text{track}} - \phi_{\text{cluster}}) < 0.05$ is removed to discriminate from electrons. The remaining clusters are reconstructed as "unconverted photons". In addition, "converted photons" are reconstructed if there is a conversion vertex matched to the cluster. There are two types of conversion vertices: double-track conversion and single-track conversion. The double-track conversion vertices are reconstructed from two opposite-charge tracks in the case that the tracks can form a vertex consistent with a massless particle. The single-track conversion vertices are reconstructed from tracks with no hits in the innermost sensitive layers.

The superclusters for photons are built independently of that of electrons. In the supercluster seed selection, the cluster is required to have the EM energy larger than 1.5 GeV. The clusters within a window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ around the barycenter of the seed cluster are considered as satellite clusters. If there is a cluster selected as seed for both photon and electron, the assignment is determined from track information and the distance between track and cluster.

Identification

The photon identification criteria are designed to keep efficiency for isolated photons high and reject backgrounds from hadronic jets. There are three identification working points: Loose, Medium, and Tight. The Loose and Medium working points are constructed from cut-based selections of the shower shape variables and used for triggers. For the physics analysis, Tight photons are used. The Tight photons are selected from Medium photons by requirement optimized using TMVA. The Tight identification is also optimized in separate bins of E_{T}.

Isolation

The photon isolation criteria use a calorimeter-based isolation variable and a track-based isolation variable [65]. For the calorimeter-based isolation variable, the same variable $E_{\mathrm{T}}^{\mathrm{cone20}}$ as the electron isolation is used. For the track-based isolation variable, $p_{\mathrm{T}}^{\mathrm{cone20}}$ defined as the scalar sum of p_{T} of the ID tracks in a cone of size $\Delta R = 0.2$ around the photon candidate excluding the photon track itself is used. There are several isolation WPs defined using the above isolation variables. In the Loose WP used for this analysis, photons are required to satisfy the following requirements:

$$E_{\mathrm{T}}^{\mathrm{cone20}} < 0.065 \times E_{\mathrm{T}}^{\gamma}, \tag{4.9}$$

$$p_{\mathrm{T}}^{\mathrm{cone20}}/E_{\mathrm{T}}^{\gamma} < 0.05. \tag{4.10}$$

The efficiency of the Tight identification for the photons passing the Loose isolation WP is shown in Figures 4.15 and 4.16.

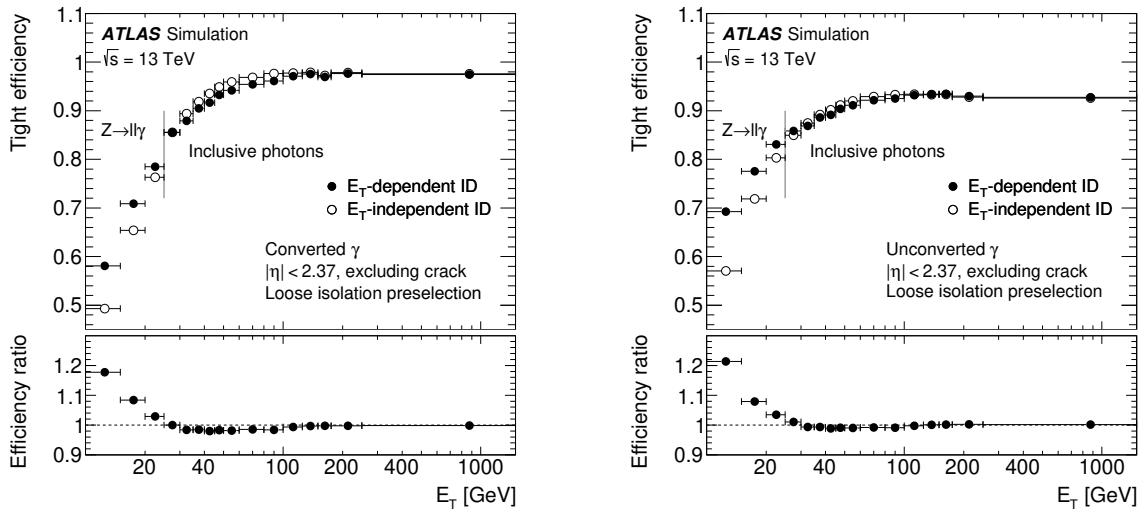
4.2.6 Jet

Jets are reconstructed using the information of the ID tracks and the topo-clusters [63]. Since the ID tracks and the topo-clusters are reconstructed independently, there may be overlaps of contribution from charged particles. First, the particle flow algorithm extrapolates the ID tracks to the calorimeter regions, and the corresponding energy deposits are subtracted from topo-clusters. The energy after subtraction and the ID tracks are used as input for the anti-kt algorithm to reconstruct jets. After the reconstruction, the energy calibration of the reconstructed jets is performed using several algorithms [63]. In addition, the jets originating from the primary vertex are selected by the Jet Vertex Tagger algorithm [68].

Particle flow algorithm

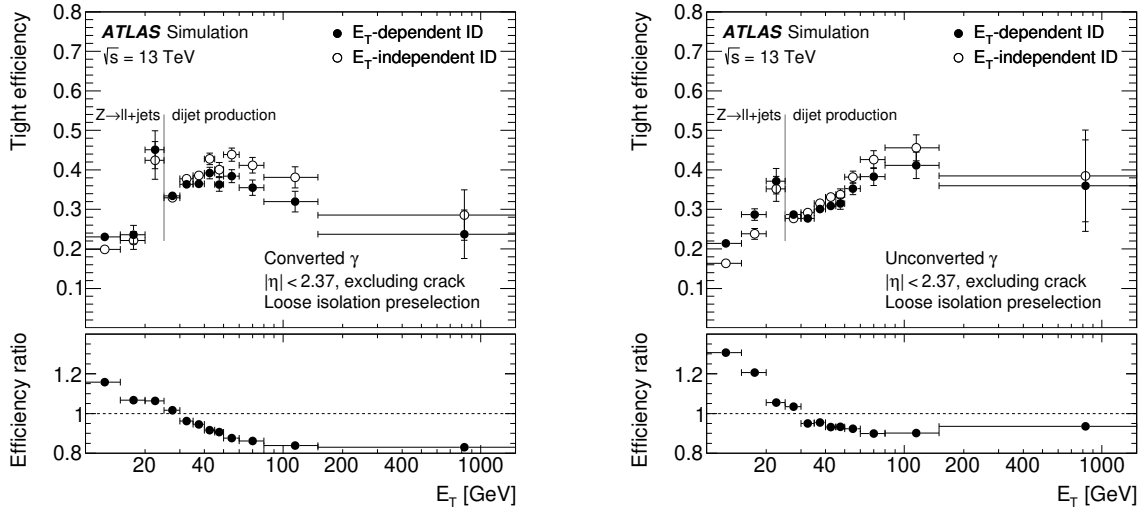
The particle flow algorithm [63] is a cell-based energy subtraction algorithm using the information of ID tracks and topo-clusters. The procedure of the algorithm shown in Figure 4.17 has five main steps described in the following.

- Track selection:
  The tracks used for the subtraction algorithm are required to pass stringent quality criteria. In addition, tracks matched to electron or muon candidates are not used because the algorithm is optimized for the hadronic shower subtraction.



(a) (b)

Figure 4.15 The Tight identification efficiencies for converted (left) and unconverted (right) photons which pass Loose isolation WP. The events in $Z \to ll\gamma$ MC and inclusive photon MC [65] are used for photons with $E_{\mathrm{T}} < 25$ GeV and $E_{\mathrm{T}} > 25$ GeV, respectively.



(a) (b)

Figure 4.16 The Tight identification efficiencies for converted (left) and unconverted (right) photons which pass Loose isolation WP. The events in $Z \to ll$+jets MC and dijets MC [65] are used for background photons from jets with $E_{\mathrm{T}} < 25$ GeV and $E_{\mathrm{T}} > 25$ GeV, respectively.

- Looking for the clusters matched to tracks:
  The clusters considered for matching with the selected tracks need to satisfy $E^{\text{clus}}/p^{\text{trk}} > 0.1$, where E^{clus} is the cluster energy and p^{trk} is the selected track momentum. The cluster closest to the extrapolated track is selected as matching the track.

- Computing $< E_{\text{dep}} >$:
  After the matching between tracks and clusters, the energy deposit is estimated for the particle which produced the track. The average of energy deposit ($< E_{\text{dep}} >$) by a particle with momentum p^{trk} measured in the ID is estimated by $< E_{\text{dep}} >= p^{\text{trk}} < E^{\text{clus}}_{\text{ref}}/p^{\text{trk}}_{\text{ref}} >$, where $< E^{\text{clus}}_{\text{ref}}/p^{\text{trk}}_{\text{ref}} >$ is the expectation value determined by a simulation. The expectation value $< E^{\text{clus}}_{\text{ref}}/p^{\text{trk}}_{\text{ref}} >$ varies depending on the p_{T}, η of the track and the calorimeter layer of the highest energy density.

- Recovering split showers:
  The number of clusters needed to capture at least 90% of the energy of the particle is typically 1–5. In this step, the algorithm determines whether the shower is split into several clusters by comparing the energy in the cluster with $< E_{\text{dep}} >$. If the shower is split into several clusters, further clusters within a cone of $\Delta R_{\text{track,cluster}} = 0.2$ are added for consideration.

- Cell-by-cell subtraction:
  Once a set of clusters corresponding to the track has been selected, energy subtraction is executed. The subtraction is performed in order in rings, which have several radii defined to contain at least one cell around the extrapolated track. The algorithm computes the energy density in each ring and starts a subtraction step from the ring with the highest energy density. The cells in the current ring are removed if the energy in the cells is less than the remaining energy required to reach $< E_{\text{dep}} >$. If the energy in the cells is larger than the remaining energy, the energy in the cells is scaled down to reach $< E_{\text{dep}} >$. The subtraction step is repeated in different rings and layers until subtracted energy reaches $< E_{\text{dep}} >$.

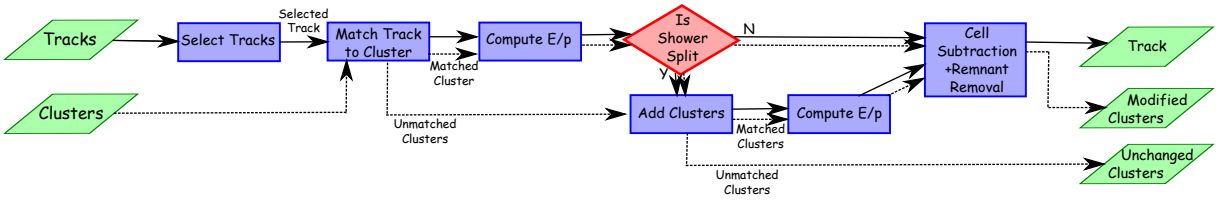


Figure 4.17 A flow chart of the particle flow algorithm procedure. [63]

Anti-kt algorithm

Jets are reconstructed by the anti-kt algorithm [69]. It uses as inputs the ensembles of the topo-clusters surviving the subtraction step of the particle flow algorithm and the selected tracks matched to the primary vertex. The algorithm is based on the distance between two topo-clusters d_{ij} defined as

$$d_{ij} = \min(k_{t_i}^{-2}, k_{t_j}^{-2}) \frac{\Delta R_{ij}^2}{R^2}, \tag{4.11}$$

where i and j are the indices of clusters, k_t is the transverse momentum of the cluster, and ΔR_{ij} is the angular distance between the clusters defined by $\Delta R = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$. In addition, R is the radius parameter set to $R = 0.4$ in this analysis. The algorithm merges the two clusters with the smallest d_{ij} and repeats the merging until $d_{ij} > k_{t_i}^{-2}$. If d_{ij} of cluster i and any other cluster j is $d_{ij} > k_{t_i}^{-2}$, the cluster i is identified as a jet. Using the remaining clusters, merging and identification are repeated until all clusters are connected to jets. A feature of this algorithm is that the reconstructed hard-jets have a circular boundary with radius R as shown in Figure 4.18.

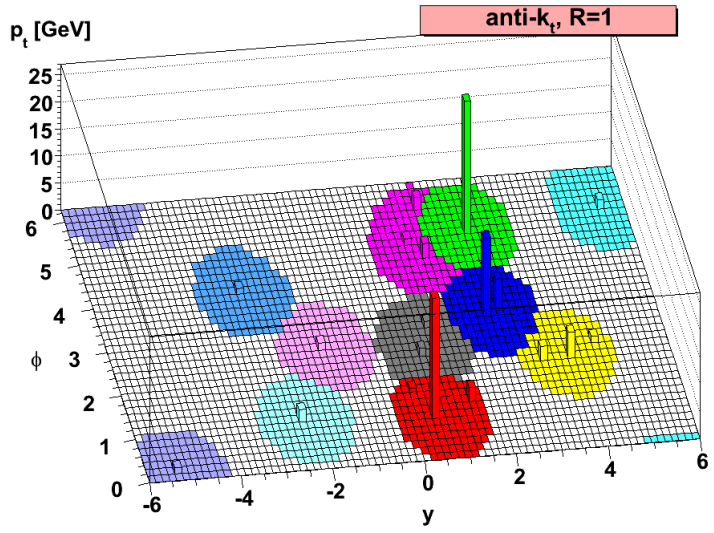
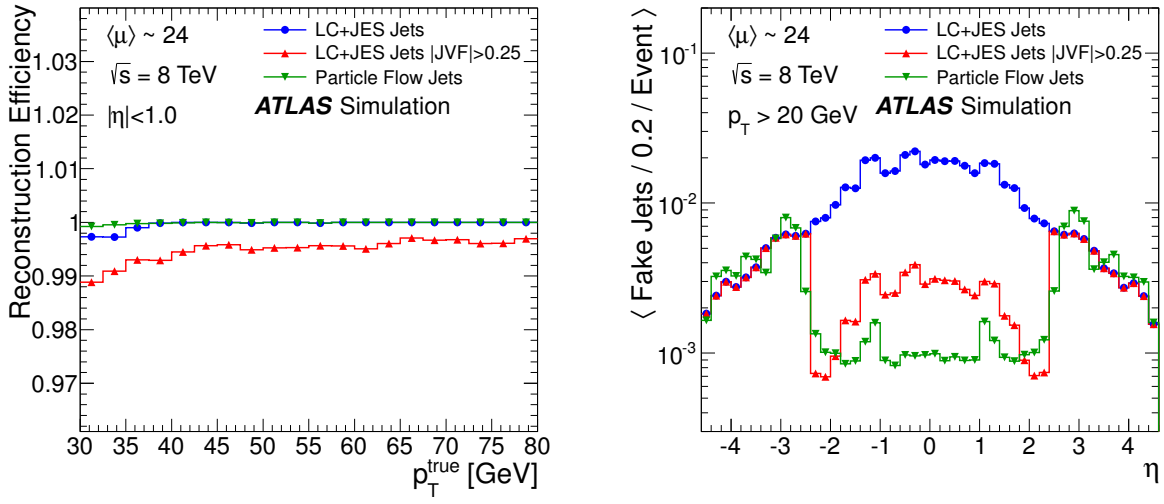


Figure 4.18 The shapes of reconstructed jets in the y–ϕ plane using anti-kt algorithm with the radius parameter $R = 1$.

Jet reconstruction efficiency and pile-up rejection are shown in Figure 4.19.

Jet calibration

The calibration of the reconstructed jets is executed for the range 20 GeV $< p_{\mathrm{T}} <$ 1500 GeV. There are four steps of the calibration: area-based pile-up correction, MC-based numerical inversion, global sequential correction, and residual *in-situ* calibration [63, 70].



(a) (b)

Figure 4.19 The jet reconstruction performance in the average pile-up $\langle\mu\rangle \sim 24$ by different reconstruction methods. The jet reconstruction efficiencies as a function of truth p_{T} (a) and the rates of fake jets from pile-up as a function of η (b) are shown. The red distributions correspond to the particle flow and the anti-kt with $R = 0.4$ algorithms.

First, the area-based pile-up correction is performed using the jet ghost-area subtraction method [71]. This method uses the transverse energy density ρ calculated from clusters and the jet area A in the η–ϕ. Due to in-time and out-of-time pile-up effects, ρ depends on the number of vertex, N_{vertex}, of pp collisions and the average number of interaction points per bunch crossing, $< \mu >$. Therefore, the effect on p_{T} of pile-up contamination is corrected by

$$p_{\mathrm{T}}^{\mathrm{corr.}} = p_{\mathrm{T}}^{\mathrm{reco.}} - \rho \times A - \alpha \times (N_{\mathrm{vertex}} - 1) - \beta \times < \mu >, \tag{4.12}$$

where $p_{\mathrm{T}}^{\mathrm{corr.}}$ and $p_{\mathrm{T}}^{\mathrm{reco.}}$ are the jet transverse momentum after and before the correction. In addition, α and β are constants derived from MC.

After the area-based pile-up correction, MC-based numerical inversion [72] is applied to correct jet energy scale from detector level to particle level. The energy responses $R = E_{\mathrm{reco.}}/E_{\mathrm{truth}}$ in various η and p_{T} regions are derived using MC and are used for the correction.

The global sequential correction [73] corrects the response differences depending on the flavor of the origin of jets and the composition of the hadrons created in jet fragmentation. This correction uses the following three additional variables. One is the fraction of the energy derived from tracks associated to the jet to consider the degree of under-calibrated signal due to the lower energy deposit of hadrons. The others are the fractions of energy deposited in the first tile calorimeter layer and in the third electromagnetic calorimeter layer to consider the energy deposit at the inactive region in the calorimeter. The corrections are applied consecutively.

Finally, the residual *in-situ* calibration is performed to account for the differences in jet response between data and MC. The jet responses in data and MC are measured separately

using the p_{T} balance in $Z \to \mu\mu$ + jet events. The ratio of the response in data and MC is used for correction in data.

Jet Vertex Tagger

Jet Vertex Tagger (JVT) [68] is a discriminant based on a 2D-likelihood to effectively select the jets from the primary vertex. Jet vertex fraction (JVF) defined by the following is a variable that can efficiently distinguish between jets from pile-up vertex and the jets from the primary vertex. The JVF is defined as the ratio of the scalar sum of p_{T} of the tracks associated with the jet and originating from the primary vertex to the scalar sum of p_{T} of all the associated tracks

$$\mathrm{JVF} = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk}_k}(\mathrm{PV}_0)}{\sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_0) + \sum_{n\geq 1}\sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_n)}, \tag{4.13}$$

where PV_0 is the primary vertex and PV_j ($j \geq 1$) are pile-up vertices. For further selection, the corrected JVF, corrJVF, is calculated by the following taking into account that the JVF depends on the total number of pile-up tracks per event $n_{\mathrm{trk}}^{\mathrm{PU}}$:

$$\mathrm{corrJVF} = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk}_k}(\mathrm{PV}_0)}{\sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_0) + \frac{\sum_{n\geq 1}\sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_n)}{0.01\times n_{\mathrm{trk}}^{\mathrm{PU}}}}. \tag{4.14}$$

In addition, the ratio $R_{p_{\mathrm{T}}}$ of the scalar sum of p_{T} of the tracks associated with the jet and originating from the primary vertex to the fully calibrated jet p_{T}, defined as the following, is used:

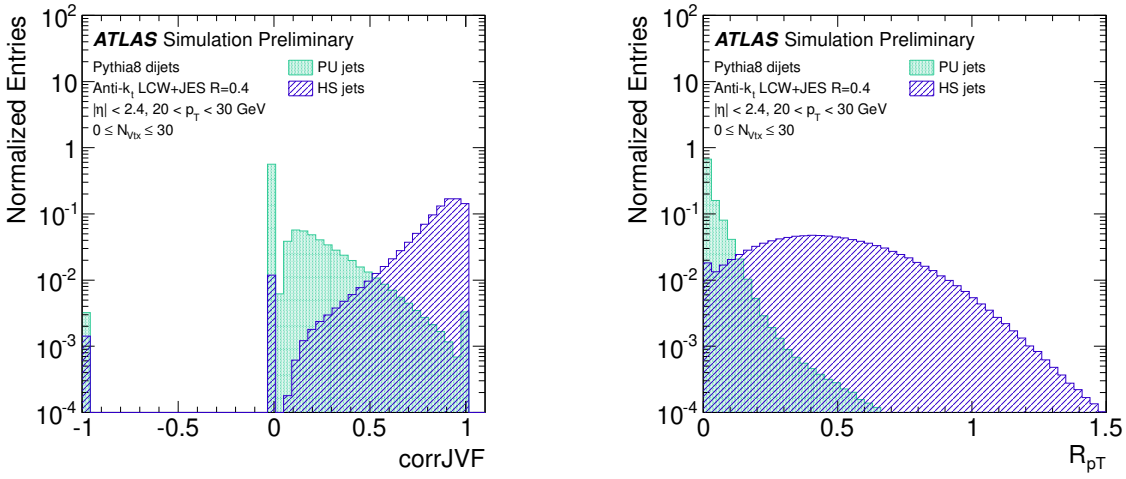
$$R_{p_{\mathrm{T}}} = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk}_k}(\mathrm{PV}_0)}{p_{\mathrm{T}}^{jet}}. \tag{4.15}$$

The jets from the primary vertex can be selected effectively by using corrJVF and $R_{p_{\mathrm{T}}}$ as inputs of the JVT. The distributions of corrJVF, $R_{p_{\mathrm{T}}}$, and JVT score are shown in Figure 4.20. The efficiency of the jet from the primary vertex for JVT score > 0.2 is around 98% depending on p_{T} in the range 20 GeV $< p_{\mathrm{T}} <$ 60 GeV.

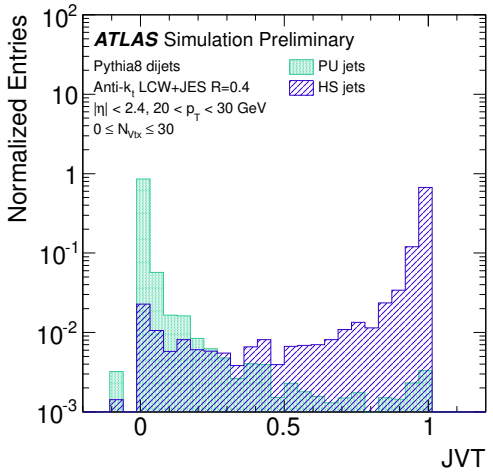
4.2.7 b-tagging

The b-hadrons decaying via weak interaction have relatively large flight length ($c\tau \sim 450$ μm). In addition, b-hadrons have relatively high mass and decay multiplicity. Jets originating from b-quarks are identified by exploiting these features. The b-tagging algorithm has two stages: low-level algorithm and high-level algorithm [74]. In the low-level algorithm, features of b-jets are reconstructed using two different approaches.

One is to use the impact parameters of all tracks associated with jets. For this approach, the IP2D and IP3D algorithms [75], which are likelihood-based classifiers, are used. The IP2D and IP3D algorithms generate PDFs for the transverse and longitudinal impact parameters using MC. The log-likelihood ratio of the b-jet, c-jet, and light-flavor jet is calculated for each track using the PDFs.



(a) (b)

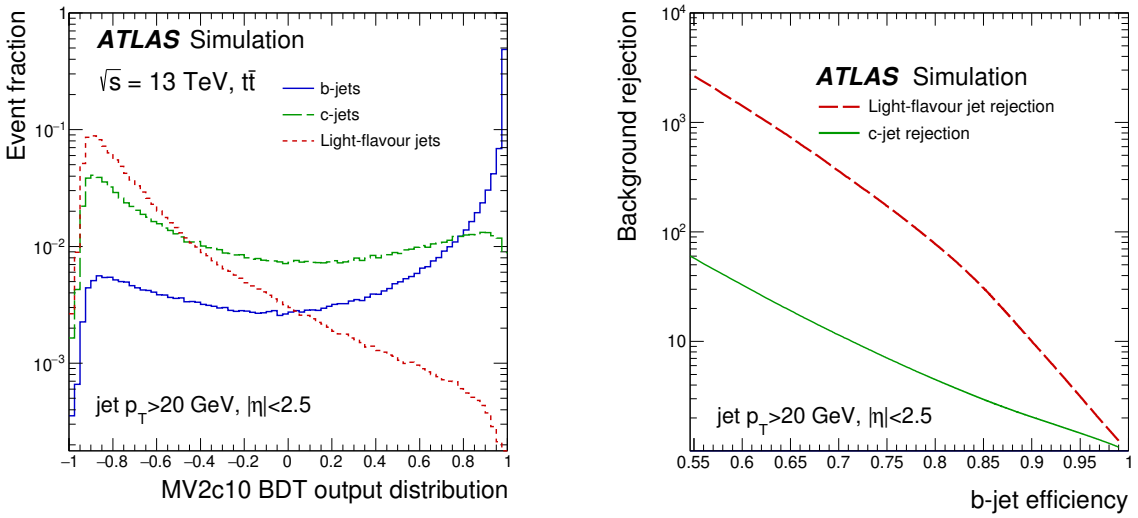


(c)

Figure 4.20 The distributions of corrJVF (a), $R_{p_{\mathrm{T}}}$ (b), and JVT score (c) for the jets from the primary vertex (blue) and pile-up jets (green). The entries with corrJVF $= -1$ correspond to jets with no associated tracks.

The other approach is to reconstruct the displaced vertices. For this approach, the secondary vertex tagging algorithm (SV1) [76] and the topological multi-vertex algorithm (JetFitter) [77] are used. Both algorithms use the information of all tracks associated with the jet as input. The SV1 algorithm reconstructs the secondary vertex using a χ^2 fit. The JetFitter algorithm reconstructs the full decay chain inside the jet using a modified Kalman filter.

In the high-level algorithm, the MV2c10 algorithm [78] is used to maximize the b-tagging performance. The MV2c10 algorithm is formed by a boosted decision tree (BDT) algorithm that combines the results from the low-level algorithm and the kinematic properties of the jets (p_{T} and $|\eta|$). The MV2c10 output score and the background rejection as a function of b-jet efficiency are shown in Figure 4.21.



(a) (b)

Figure 4.21 The b-tagging performance by the MV2c10 algorithm [78]. The MV2c10 output scores for b-jets (blue), c-jets (green) and light-flavor jets (red) are shown in (a). The rejections of c-jet (green) and light-flavor jet (red) as a function of the b-jet efficiency of the MV2c10 algorithm are shown in (b). These are evaluated using $t\bar{t}$ simulation events.

4.2.8 Missing transverse momentum

The missing transverse momentum is a general parameter connected to kinematics of invisible particles in the event. The missing transverse momentum is calculated using transverse momenta of reconstructed particles and the ID tracks ($p_{\mathrm{T}} > 400$ MeV) associated with the primary vertex but not associated with reconstructed particles, referred to as "soft-signals" [79]. The vector $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$ given by

$$\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum \boldsymbol{p}_{\mathrm{T}}^{e} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mu} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mathrm{jet}} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mathrm{soft}}, \tag{4.16}$$

where each term corresponds to the vector sum of transverse momenta of all reconstructed electrons, muons, jets, and soft-signals, respectively. In addition, the magnitude and the direction of the missing transverse momentum is derived using the components of missing transverse momentum $E^{\text{miss}}_{x(y)}$ defined by $\boldsymbol{E}^{\text{miss}}_{\text{T}} = (E^{\text{miss}}_x, E^{\text{miss}}_y)$:

$$\text{magnitude} \quad : \quad E^{\text{miss}}_{\text{T}} = |\boldsymbol{E}^{\text{miss}}_{\text{T}}| = \sqrt{(E^{\text{miss}}_x)^2 + (E^{\text{miss}}_y)^2}, \tag{4.17}$$

$$\text{direction} \quad : \quad \phi^{\text{miss}} = \tan^{-1}(E^{\text{miss}}_y / E^{\text{miss}}_x). \tag{4.18}$$

4.2.9 Overlap removal

Since the muons, the electrons and the jets are reconstructed independently using different algorithms, a single particle may be reconstructed or identified as two or more objects at the same time. In order to avoid the overlap and select the correct object, the "overlap removal" is applied. First the overlap removal between leptons is performed when leptons share the same ID tracks. Then overlap removal between leptons and jets is performed based on $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. The detailed criteria are summarized in Table 4.6.

Table 4.6 The overlap removal criteria of each object. When the "objects to be removed" and the "reference objects" satisfy at the same time the "applicable condition", the "objects to be removed" are removed. These criteria are executed in order from the top.

Objects to be removed	Reference objects	Applicable condition
electrons (lower p_{T})	electrons (higher p_{T})	objects share the same track
muons	electrons	muon is identified as CT muon and objects share the same track
electrons	muons	objects share the same track
jets	electrons	$\Delta R_{e,jet} < 0.2$
electrons	jets	$0.2 < \Delta R_{e,jet} < 0.4$
jets	muons	$\Delta R_{\mu,jet} < 0.2$ and $N^{jet}_{\text{trk}} < 3$ or $(p^{jet}_{\text{T}}/p^{\mu}_{\text{T}} < 2$ and $p^{\mu}_{\text{T}}/\sum p^{\text{trk}}_{\text{T}} > 0.7)$
muons	jets	$0.2 < \Delta R_{\mu,jet} < 0.4$ or ($\Delta R_{\mu,jet} < 0.2$ if converse case of the bottom line in the above row)

4.2.10 Selection for the analysis

For each reconstructed particle to be used in this analysis, further selection criteria are applied.

The muons with $p_{\mathrm{T}} > 6$ GeV in the region $|\eta| < 2.7$ are selected to reject muons from hadron decays. In addition, requirements related to the impact parameters, $|d_0^{\mathrm{BL}}/\sigma_{d_0^{\mathrm{BL}}}| < 3$ and $|z_0^{\mathrm{PV}} \cdot \sin\theta| < 0.5$ mm, are applied, to reject cosmic muons. In addition, the muons are required to the requirements related to impact parameters $|d_0^{\mathrm{BL}}/\sigma_{d_0^{\mathrm{BL}}}| < 3$ and $|z_0^{\mathrm{PV}} \cdot \sin\theta| < 0.5$ mm, to reject the effect of cosmic muons. Here d_0^{BL} and z_0^{PV} are the transverse and longitudinal impact parameters.

The electrons are selected with $p_{\mathrm{T}} > 7$ GeV in the region $|\eta| < 2.47$ excluding the crack region $1.37 < |\eta| < 1.52$. Electrons are rejected in the case that a faulty cell is used for the reconstruction. Similarly to muons, the transverse and longitudinal impact parameters are required to satisfy $|d_0^{\mathrm{BL}}/\sigma_{d_0^{\mathrm{BL}}}| < 5$ and $|z_0^{\mathrm{PV}} \cdot \sin\theta| < 0.5$ mm.

The jets are selected with $p_{\mathrm{T}} > 25$ GeV ($p_{\mathrm{T}} > 30$ GeV) in the region $|\eta| < 2.4$ ($2.4 < |\eta| < 4.5$). In order to suppress pile-up , the cut of JVT score > 0.2 is applied to jets with $|\eta| < 2.4$ and $p_{\mathrm{T}} < 60$ GeV.

For jets containing b-hadrons, the MV2c10 b-tagging algorithm is used for $p_{\mathrm{T}} > 20$ GeV in the region $|\eta| < 2.5$. In this analysis, two different WPs are used. For the ttH category defined in Section 4.3, the WP, which is designed to provide 85% b-tagging efficiency in $t\bar{t}$ events, is used. The rejection factors for c-jets and light-flavor jets are 2.7 and 25, respectively. For other categories, the WP, which are designed to provide 60% b-tagging efficiency in $t\bar{t}$ events, is used. The rejection factors for c-jets and light-flavor jets are 23 and 1200, respectively.

4.3 Event selection

The events used for this analysis are pre-selected using common features in the target signal production modes. In order to exploit the difference of the signal kinematics for each process, the pre-selected events are further selected depending on the production modes and analyzed independently. The events for the ttH, VH, VBF, and ggF production modes are selected exclusively. Since the signal yields of ttH and VH are relatively small, the events for ttH and VH are selected first. The common pre-selection for all processes is described in Section 4.3.1. The ttH, VH and ggF/VBF selection criteria are presented in Sections 4.3.2–4.3.4.

4.3.1 Pre-selection

Events are required to have at least one opposite-charge muon pair and at least one vertex of pp interaction. The leading and subleading muons are required to have $p_{\mathrm{T}} > 27$ GeV and $p_{\mathrm{T}} > 15$ GeV (except for VH3L channel described in Section 4.5.2), respectively. For VH3L, a lower threshold of p_{T} for the subleading muon (10 GeV) is used due to the lower signal efficiency. In addition, the events are required that either the leading or subleading muon is matched to the muon used for the single muon trigger. For the main signal samples (ggF and VBF), the main background sample ($Z \to \mu\mu$), and the data sample, the number of events passing each cut is summarized in Table 4.7.

Table 4.7 The number of events passing each cut (# events) and the efficiencies (Eff.) for the ggF, VBF, and $Z \to \mu\mu$ samples and data.

Requirements	ggF		VBF		$Z \to \mu\mu$		Data	
	# events	Eff. [%]	# events	Eff. [%]	# events	Eff. [%]	# events	Eff. [%]
Before pre-selection	1596043	-	374085	-	95441274	-	3709769	-
$N_{\mathrm{muons}} \geq 2$	1323314	83	310700	83	73105594	77	368096	10
Trigger matching	1313884	99	308642	99	70935778	97	351231	95
$p_{\mathrm{T}}^{\mathrm{lead(sub)}} > 27$ (15) GeV	1292217	98	300210	97	69249801	98	330264	94
Opposite-charge muons	1292183	100	300199	100	69249801	100	330169	100

4.3.2 ttH selection

The ttH process is characterized by two features: the existence of two b-jets and the high multiplicity of the final state particles. In this analysis, only the $t\bar{t}$ dileptonic decay and the $t\bar{t}$ semileptonic decay shown in Figure 4.3 are considered. In the final state, there are at least one electron or one muon and the missing transverse energy from neutrino in addition to the b-jets and the muon pair.

In order to select the events of the ttH process keeping higher signal efficiency, the events are loosely selected using the requirements on the basic features described above. The events are required to have at least one jet passing 85% WP of the MV2c10 b-tagging algorithm described in Section 4.2.7. Although there are 2 b-jets in the ttH process, the requirements "at least one" is applied for higher sensitivity of the search. In addition, the events are required to have at least one electron or one muon with $p_{\mathrm{T}} > 15$ GeV in addition to the muon pair. If there are more than three muons in an event, the opposite charge muon pair with maximum p_{T} is regarded as muons from the Higgs boson decay. The efficiency of the correct pairing by this method is 80%.

4.3.3 VH selection

In the VH selection, the events selected for ttH are excluded. The VH process consists of two possibilities of WH and ZH processes. In this analysis, both WH and ZH processes including the leptonic decay of the boson, as shown in Figure 4.2, are considered. In the WH process, there are three leptons, which are one lepton from $W \to l\nu$ decay and two muons from the Higgs boson decay. On the other hand, the final state of ZH process has four leptons, which are two leptons from $Z \to ll$ decay and two muons from the Higgs boson decay. Since the target processes have different number of leptons in the final state, the events are separated into two channels by different requirements. The first channel, referred to as "VH3L", contains the events with three leptons, targeting the WH process. The second channel, referred to as "VH4L", contains the events with four leptons, targeting the ZH process.

For the VH3L, the events are selected by requiring exactly three leptons and no b-tagged jets at 85% WP. The additional lepton among the three leptons is to be an electron or a muon. In order to suppress the $Z \to \mu\mu$ background, the events with an opposite charge muon pair in the range 80 GeV $< m_{\mu\mu} <$ 105 GeV are rejected, if there are three muons in the event.

For the VH4L, the events are selected by requiring at least four leptons and no b-tagged jets at 85% WP. The additional leptons are required to include at least one opposite charge electron pair with $p_{\mathrm{T}} > 8$ GeV or muon pair with $p_{\mathrm{T}} > 6$ GeV. Up to one Z candidate is allowed in the event.

If there are more than three muons, the assignments of the Higgs boson and W or Z boson candidates are determined based on the charges of muons and the χ^2 for the masses of the candidates. The χ^2 for VH3L and VH4L is defined as Eqs. (4.19) and (4.20), respectively.

$$\text{VH3L} \quad : \quad \chi^2 = \frac{\left(m(\mu_i^+\mu_i^-) - 125\ \text{GeV}\right)^2}{(3.0\ \text{GeV})^2} + \frac{\left(m_{\mathrm{T}}(\mu_j^{+(-)} E_{\mathrm{T}}^{\mathrm{miss}}) - 70\ \text{GeV}\right)^2}{(20\ \text{GeV})^2} \tag{4.19}$$

$$\text{VH4L} \quad : \quad \chi^2 = \frac{\left(m(\mu_i^+\mu_i^-) - 125\ \text{GeV}\right)^2}{(3.0\ \text{GeV})^2} + \frac{\left(m(\mu_j^+\mu_j^-) - 91.1\ \text{GeV}\right)^2}{(3\ \text{GeV})^2} \tag{4.20}$$

The combination of (i, j) with the smallest χ^2 is regarded as the signal candidate. The efficiencies of the correct pairing for the VH3L and VH4L by this method are 97% and 93%, respectively.

4.3.4 ggF/VBF selection

For the ggF and VBF processes, the events selected by the ttH and VH selections are excluded. In the VBF process, there are two jets in the final state. The ggF process may contain jets from the initial state radiation. Since there is no clear difference in the final states between the ggF and VBF processes, a common loose selection is first used for these processes.

The events are required to have no additional muons and no b-jets at 60% WP. For the ggF and VBF processes, the jets are used for categorization. In order to keep efficiency of signal events with jets high, the 60% b-tagging WP with higher rejection factor for light-flavor jets is chosen for no b-jet requirement.

The number of events passing each selection is summarized in Table 4.8. In each selection, the dimuon invariant mass distributions are shown in Figure 4.22. There are about 500 times as many background events as the signal events in 120 GeV $< m_{\mu\mu} <$ 130 GeV. The invariant mass distribution for signal events has a tail, which is caused by the FSR photon taking away the energy of the muon.

Table 4.8 The number of events passing each selection in 120 GeV $< m_{\mu\mu} <$ 130 GeV.

Selection	Signal					Background		
	ggF	VBF	WH	ZH	ttH	$Z \to \mu\mu$	Diboson	Top
ttH	0.01	0.003	0.1	0.05	1.7	9.2	43	144
VH3L	0.1	0.03	2.7	0.2	0.1	76	135	58
VH4L	0.001	0.0001	0.0004	0.3	0.01	0.4	19	0.6
ggF/VBF	740	57	17	10	1.3	394517	4760	11911

In order to extract the signal more sensitively in the invariant mass distribution, we increase the signal yield (A), decrease the background yield (B), and improve the invariant mass resolution (C). The concept of this strategy is shown in Figure 4.23. By reconstructing and including the FSR photon in the four-momentum in the dimuon invariant mass calculation, (A) and (C) are performed. For (B), the events passing these selection are further classified into 20 categories in total using the multivariate analysis.

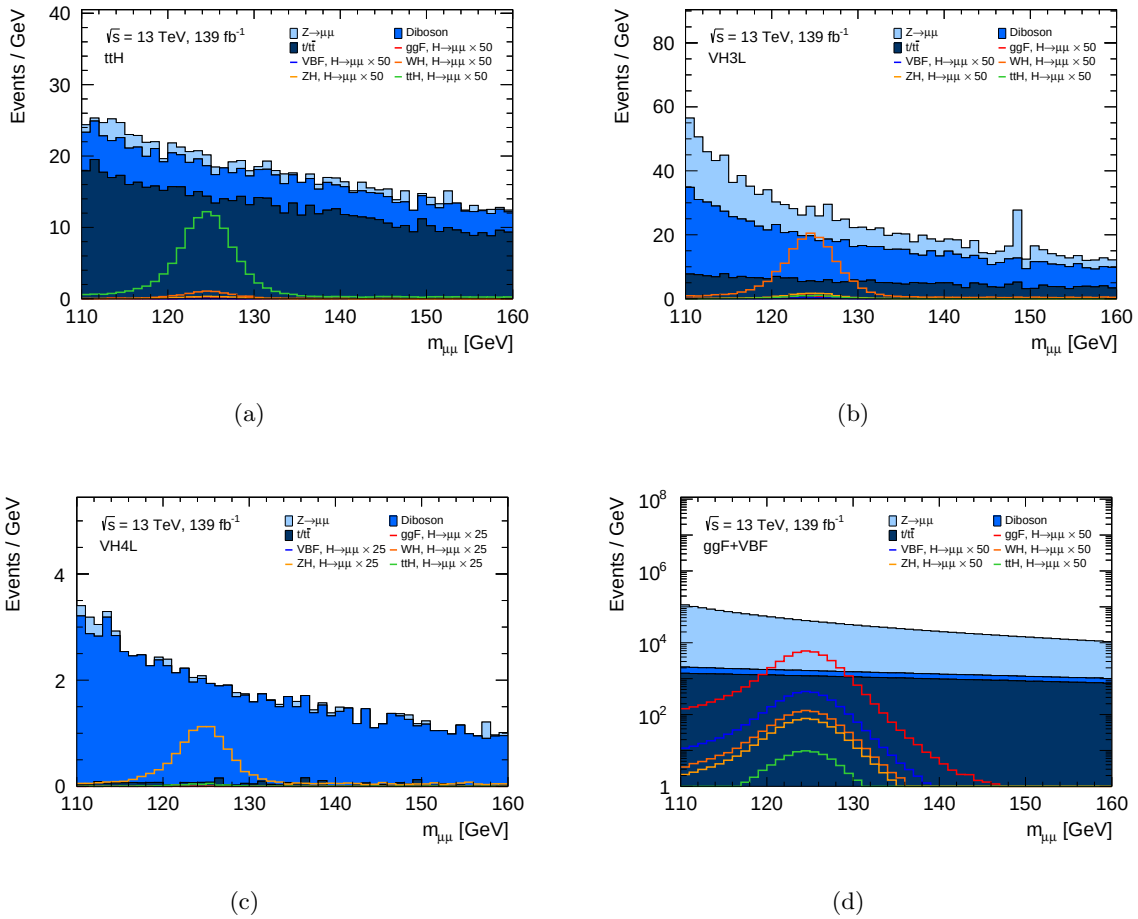


Figure 4.22 The dimuon invariant mass distributions for the events passing the ttH (a), VH3L (b), VH4L (c), and ggF/VBF (d) selection.

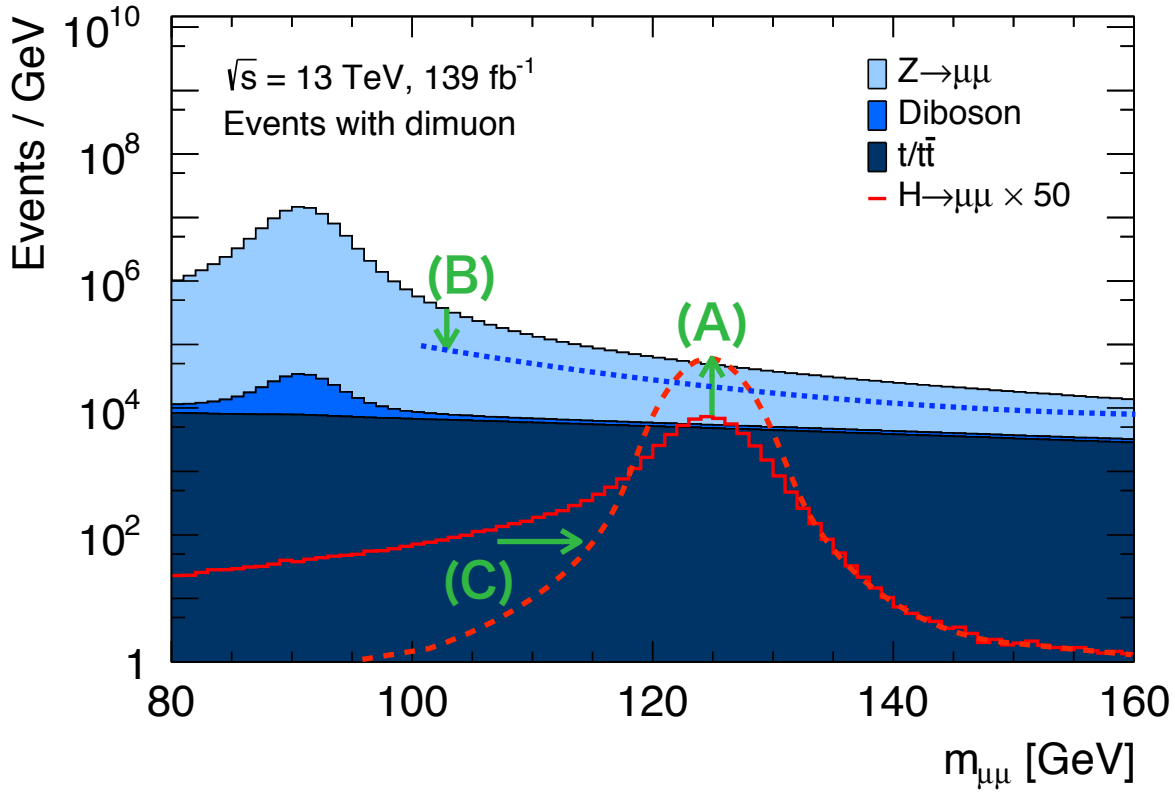


Figure 4.23 The strategies of the H → $\mu\mu$ sensitivity improvement; we increase the signal yield (A), decrease the background yield (B), and improve the invariant mass resolution (C).

4.4 FSR recovery

Muons from H → μμ may lose a significant fraction of their energy by QED Final State Radiation (FSR). The diagram is shown in Figure 4.24. In order to improve the sensitivity of the H → μμ search, we select the photon from the QED FSR, referred to as the FSR photon. The four-momentum is included in the calculation of the invariant mass of muons.

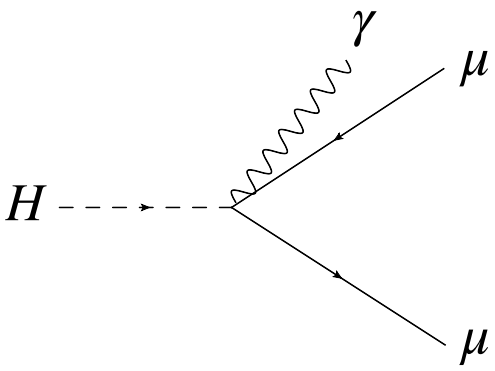
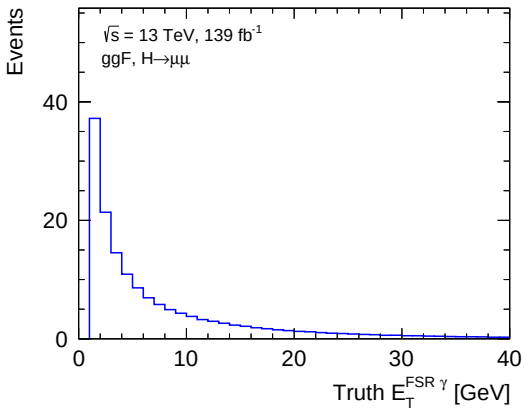


Figure 4.24 Feynman diagram of the final state radiation in the H → μμ process.

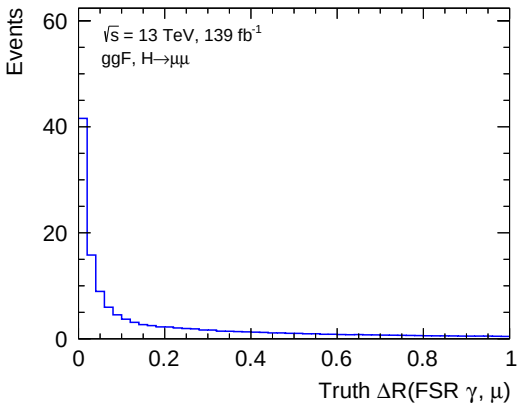
4.4.1 FSR photon in the H → μμ event

The kinematics of the FSR photons in the H → μμ signal events are studied using the truth-level and reconstruction-level information of $gg\mathrm{F}, H \to \mu\mu$ MC. The events used for this study are required to have one opposite-charge muon pair.
Sixteen percent of all events has truth-level FSR photon with $E_{\mathrm{T}} > 1$ GeV. The FSR photons tend to have very small E_{T} and small ΔR between the photon and the muon as shown in Figure 4.25. The distributions of the invariant mass of two reconstructed muons before and after the truth-level FSR recovery are shown in Figure 4.26. In the truth-level recovery, the mean value of the invariant mass for the events with FSR photon is improved from 123.3 GeV to 124.7 GeV. The standard deviation improves from 5.7 GeV to 3.1 GeV. The number of events in the range 120 GeV $< m_{\mu\mu} <$ 130 GeV is increased by 10%.

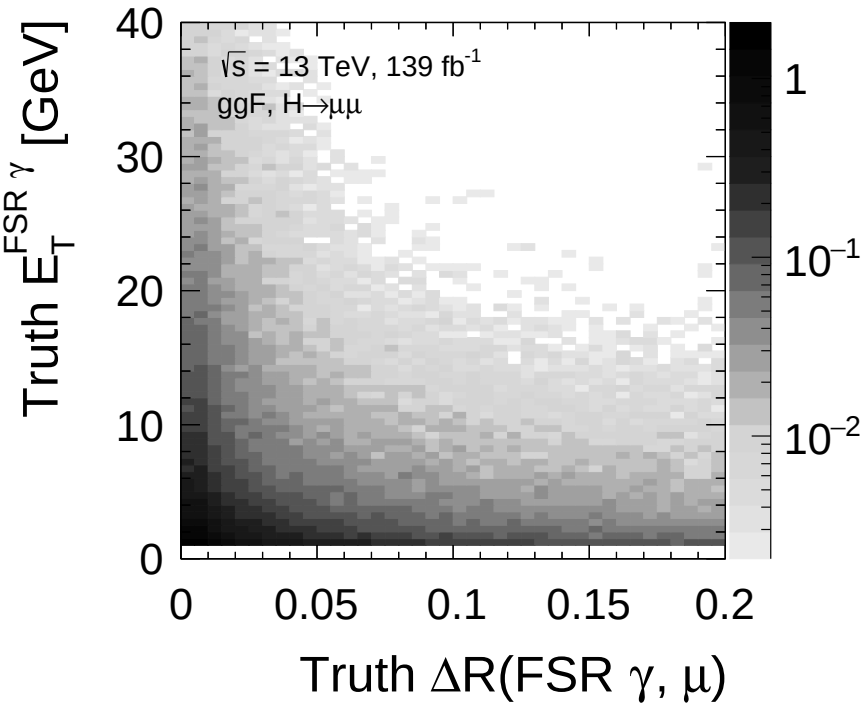
The H → μμ decay with FSR photon has the same final state as the $H \to Z(\to \mu\mu)\gamma$ decay. The branching ratios of the $H \to Z\gamma$ decay and the $Z \to \mu\mu$ decay are 2×10^{-3} and 0.03, respectively. The yield of $H \to Z(\to \mu\mu)\gamma$ events is smaller than the Drell-Yan, but the invariant mass of $\mu\mu\gamma$ for the $H \to Z(\to \mu\mu)\gamma$ events peaks at around 125 GeV. In contrast to the H → μμ decay, the FSR candidates in the $H \to Z(\to \mu\mu)\gamma$ MC have large ΔR and typically $E_{\mathrm{T}} \sim 30$ GeV as shown in Figure 4.27. In order to minimize the contamination of the $H \to Z(\to \mu\mu)\gamma$ events around 125 GeV, photons are required to satisfy $\Delta R < 0.2$.



(a)

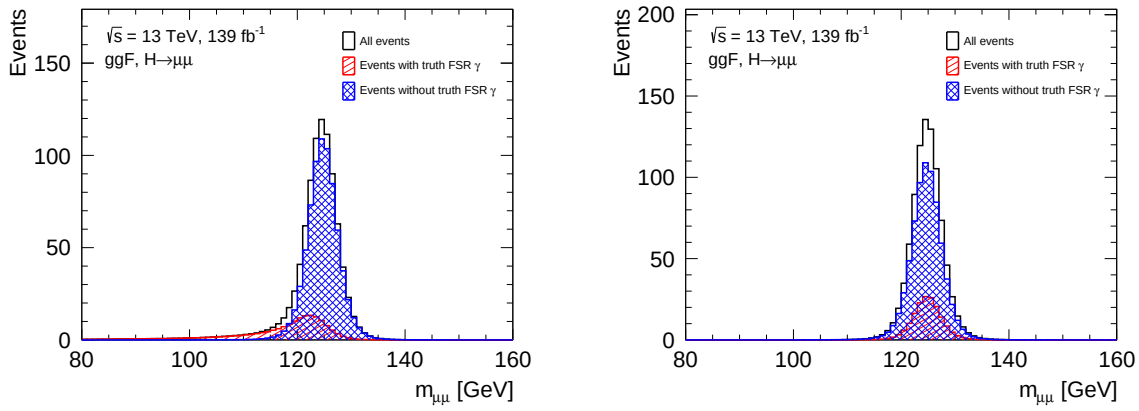


(b)



(c)

Figure 4.25 The distributions of E_{T} and $\Delta R_{\gamma,\mu}$ for truth-level FSR photons.



(a) (b)

Figure 4.26 The dimuon invariant mass distributions before (a) and after (b) the truth-level FSR recovery. The total events, the events with FSR photons and the events without FSR photons are shown by black, red and blue, respectively.

4.4.2 Reconstruction

The FSR photon candidates are reconstructed using the reconstructed photons and electrons by the ATLAS standard method described in Section 4.2 excluding the isolation requirement. In this reconstruction, the clusters are searched with a size of 3×5 in units of 0.025×0.025 in the $\eta \times \phi$ space, which are seeded by topo-clusters. The clusters satisfying the following requirements are reconstructed as FSR candidates:

- the transverse energy of the cluster: $E_{\mathrm{T}} > 3.5$ GeV,
- the distance between the cluster and the muon: $\Delta R_{\mathrm{cluster},\mu} (= \sqrt{\Delta\eta^2 + \Delta\phi^2}) < 0.2$, and
- the fraction of the energy deposit in the front sampling of the calorimeter over the total energy deposit of the cluster: $f_1 > 0.1$.

The requirement of f_1 suppresses the cluster of the energy deposit by the muon ionization. The energy of the candidates with $\Delta R_{\mathrm{cluster},\mu} < 0.05$ is corrected by subtracting the energy deposit of 400 MeV$/\cosh\eta_\mu$ due to the muon ionization, where η_μ is the pseudorapidity of the muon. In addition to the above FSR candidates, the clusters specialized for smaller E_{T} are also reconstructed. The criteria are

- the transverse energy of the cluster: 1.5 GeV$< E_{\mathrm{T}} < 3.5$ GeV,
- the distance between the cluster and the muon: $\Delta R_{\mathrm{cluster},\mu} (= \sqrt{\Delta\eta^2 + \Delta\phi^2}) < 0.08$, and
- the fraction of the energy deposited in the first sampling layer of the EM calorimeter over the total energy deposit of the cluster: $f_1 > 0.2$.

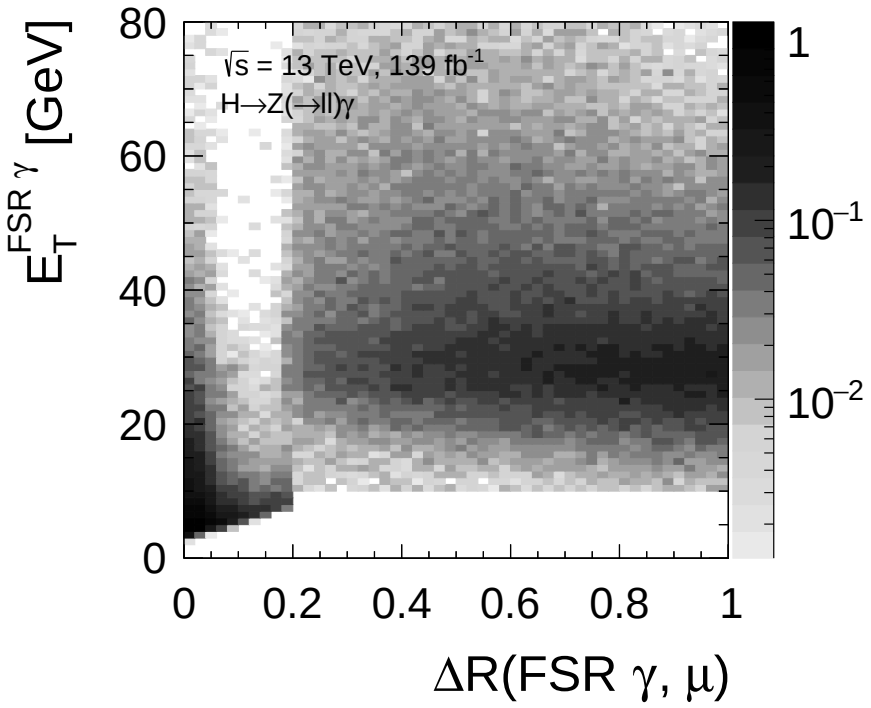


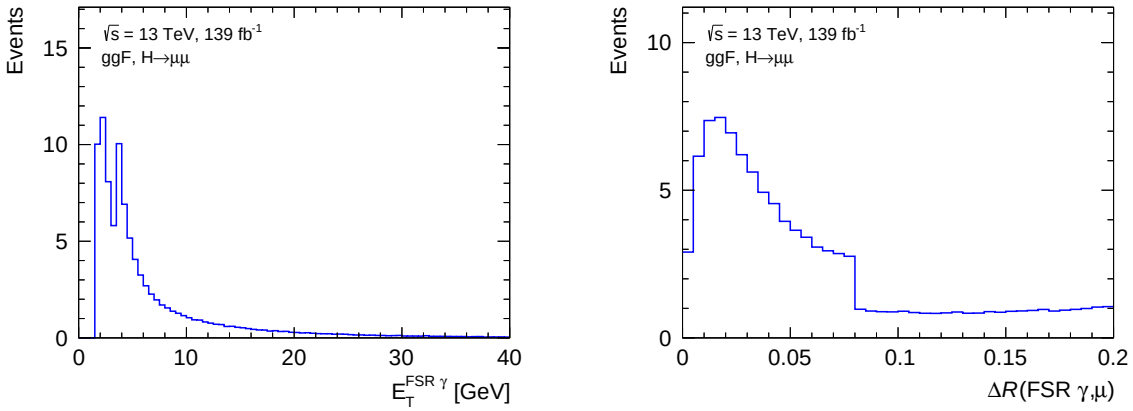
Figure 4.27 The E_{T}–ΔR distribution for the FSR photon candidates in the $H \to Z(\to \mu\mu)\gamma$ events. The FSR photon candidates in $\Delta R < 0.2$ are reconstructed by the optimized selection described in this thesis. The candidates in $\Delta R > 0.2$ are satisfied $E_{\mathrm{T}} > 10$ GeV and loose isolation requirement.

This reconstruction procedure is also introduced in the $H \to ZZ \to 4l$ analysis using data taken in 2011–2012 [80].

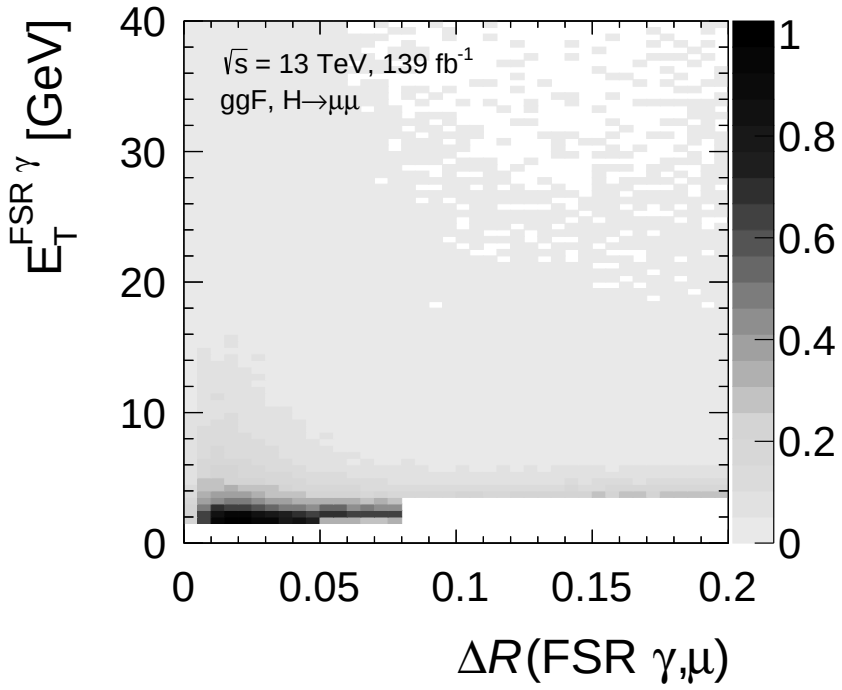
In this analysis, up to one FSR photon candidate is reconstructed per event. If there are multiple candidates, the candidate with higher E_{T} is selected. The FSR candidate is assigned to the muon with smaller ΔR.

In the reconstruction-level, the events with FSR photon candidate are found for 11% of total events. The distributions of $E_{\mathrm{T}}^{\mathrm{cand}}$ and $\Delta R_{\mathrm{cand},\mu}$ for the reconstructed FSR photon candidates are shown in Figure 4.28.

The FSR photon candidates do not match truth-level FSR photon in around 39% of the reconstructed FSR photon candidates. The FSR candidates in such events are called “fake” FSR candidates. The matching between the reconstructed candidate and truth-level FSR photon is performed by requiring ΔR between the candidate and the truth-level FSR photon to be less than 0.2. Most of the fake FSR candidates arise from pile-up interaction. There are about 17 reconstructed photons per event in average, which have low-E_{T} and smaller ΔR than FSR photon. The distribution of the number of reconstructed photons and the two-dimensional



(a) (b)



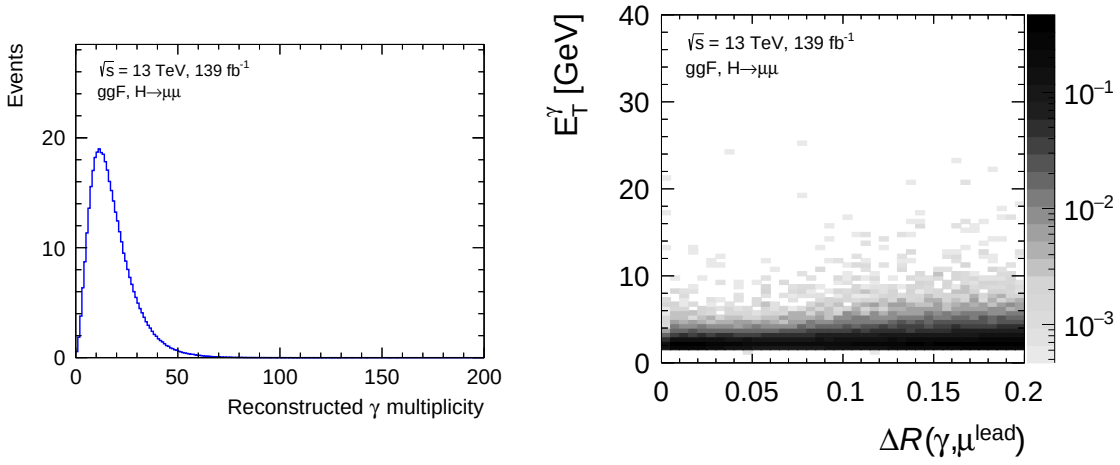
(c)

Figure 4.28 The distributions of E_{T} and $\Delta R_{\gamma,\mu}$ for reconstruction-level FSR photon candidates. The gaps of $E_{\mathrm{T}} = 3.5$ GeV and $\Delta R = 0.08$ are caused by different selection of the FSR photon reconstruction.

distribution on E_{T} and ΔR are shown in Figure 4.29. The fake FSR rate defined in Eq. (4.21) increases as the number of pile-up collisions increases as shown in Figure 4.30.

$$\text{Fake FSR rate} = \frac{\text{Number of events with the fake FSR candidate}}{\text{Number of events with reconstructed FSR candidate}} \tag{4.21}$$

The distributions of $E_{\mathrm{T}}^{\mathrm{fake}}$ and $\Delta R_{\mathrm{fake},\mu}$ for the fake FSR candidates are shown in Figure 4.31. The fake FSR candidates tend to have relatively small E_{T}. In addition, the fake FSR candidate increases in the region of larger dRΔR for the region $\Delta R > 0.08$. The fake FSR candidate increases the dimuon invariant mass as shown in Figure 4.32. As a result, the peak position and the width of the $m_{\mu\mu}$ distribution are higher than the original distribution.



(a) (b)

Figure 4.29 Number of reconstructed photons in an event (a). The reconstructed photon distribution in the E_{T}–ΔR plane (b).

4.4.3 Optimization of selection

The reconstruction procedure is optimized based on the data taken in 2011–2012. To cope with the higher pile-up, the reconstructed FSR candidates are further selected based on the fake FSR rate. The fake FSR rate depends on E_{T} and ΔR as shown in Figure 4.33. In order to suppress the fake FSR candidates, E_{T} threshold depending on ΔR is scanned and the expected significance is estimated by the following method:

- invariant mass distribution after FSR recovery using the FSR candidates with E_{T} larger than the threshold is obtained,
- invariant mass ranges are defined by $m_{\mu\mu} = \mu \pm \Delta$ in each category[1], where μ indicates

[1] In this optimization study, the categorization for the H $\to \mu\mu$ analysis using 79.8 fb^{-1} in ATLAS [13] is used. The performance of FSR recovery have no significant difference between categories.

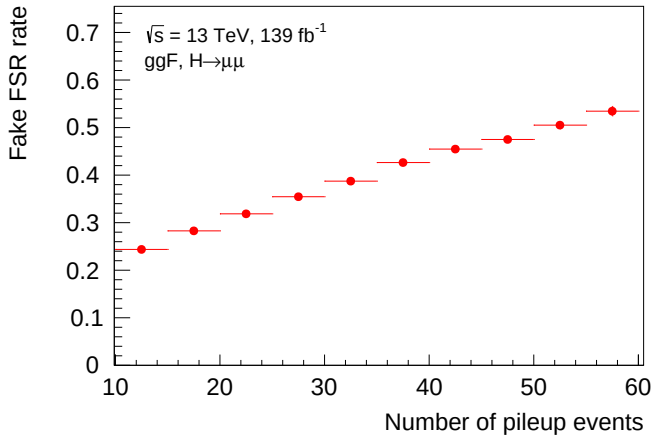


Figure 4.30 Fake FSR rate depending on pile-up.

the mean of the invariant mass distribution for $ggF, H \to \mu\mu$ MC and Δ is scanned from 2.5 GeV to 10 GeV with a step size of 0.1 GeV,

- the significance $S/\sqrt{B}$, where S and B is the numbers of $ggF, H \to \mu\mu$ signal and $Z \to \mu\mu$ background, is derived in the invariant mass range for each category,
- the maximum value of $S/\sqrt{B}$ is selected as an expected significance.

The E_{T} requirement which maximizes the expected significance is selected for this analysis.

In the E_{T}–ΔR threshold scan, the E_{T} threshold is defined by the following linear function with a slope of 25 GeV corresponding to the change of the fake FSR rate

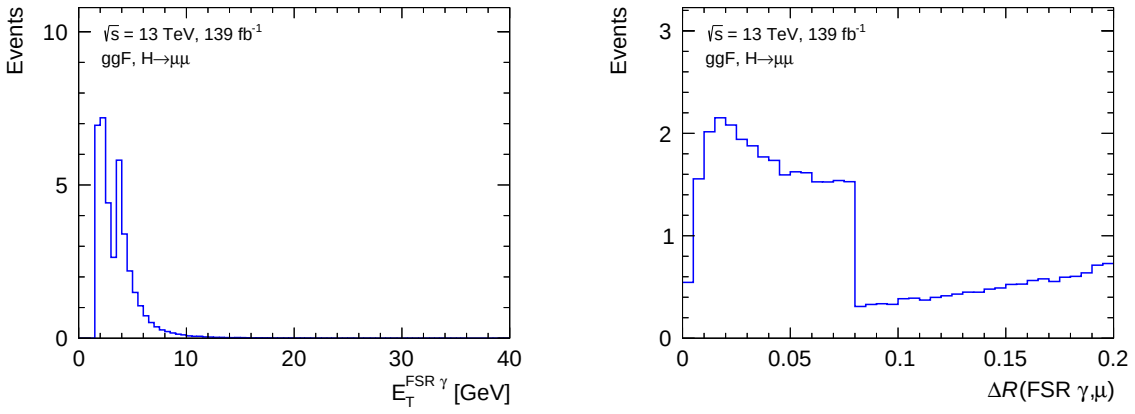
$$E_{\mathrm{T}}^{\mathrm{threshold}}\ [\mathrm{GeV}] = 25\ [\mathrm{GeV}] \cdot \Delta R_{\mathrm{cand},\mu} + A\ [\mathrm{GeV}], \tag{4.22}$$

where A is the intercept of the function. The A is scanned from 2 GeV to 6 GeV with a step size of 1 GeV. The ratio of the number of signal events in 120 GeV $< m_{\mu\mu} <$ 130 GeV over that in 110 GeV $< m_{\mu\mu} <$ 160 GeV, the mean value and the standard deviation of the invariant mass, the truth FSR efficiency defined as Eq. (4.23), and the fake FSR rate are summarized in Table 4.9.

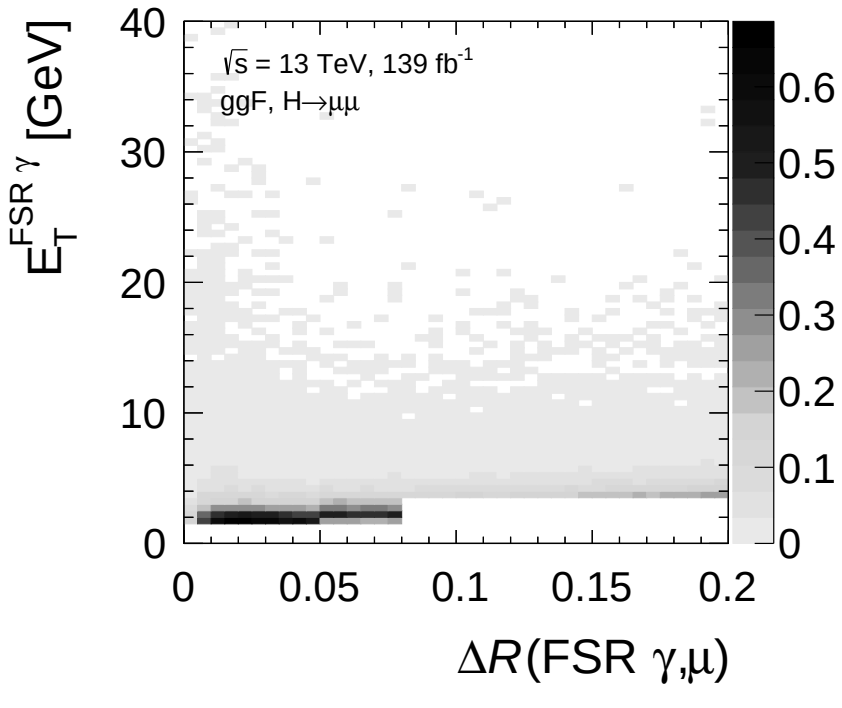
$$\text{Truth FSR Eff.} = \frac{\text{Number of events with the FSR candidatematchedtothetruthFSR}}{\text{Number of events with truth FSR}} \tag{4.23}$$

The expected significance is maximized at $A = 3$ GeV as shown in Figure 4.34. The improvement of expected significance by the cut E_{T} [GeV] > 25 [GeV] $\cdot\, \Delta R_{\mathrm{cand},\mu} + 3$ [GeV] with respect to that without E_{T} threshold is around 2.0%.

The cut E_{T} [GeV] > 25 [GeV] $\cdot\, \Delta R_{\mathrm{cand},\mu} + 3$ [GeV] is applied to the FSR candidates. By this cut, pile-up dependence of the fake FSR rate is significantly reduced as shown in Figure 4.35.



(a) (b)



(c)

Figure 4.31 The distributions of E_{T} and $\Delta R_{\gamma,\mu}$ for fake FSR photon candidates.

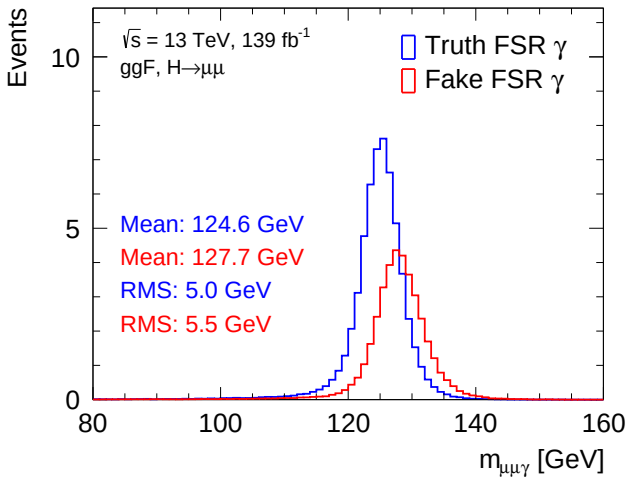


Figure 4.32 Invariant mass distributions after the FSR recovery for the events with the FSR candidate matched to the truth-level FSR (blue) and for the events with the fake FSR candidate (red).

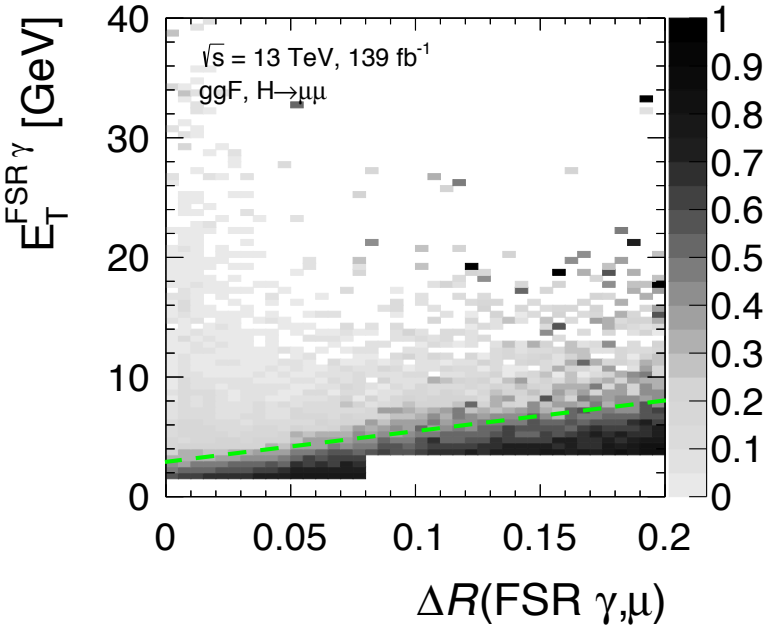


Figure 4.33 The distribution of the fake FSR rate on the E_{T}–ΔR plane. The green dotted line shows the optimized E_{T} threshold depending on ΔR.

Table 4.9 The ratio of the number of signal events in 120 GeV $< m_{\mu\mu} <$ 130 GeV over that in 110 GeV $< m_{\mu\mu} <$ 160 GeV, the mean value and the standard deviation of the invariant mass for the events with the FSR candidate, the truth FSR efficiency, and the fake FSR rate. The values are shown for the default selection, A = 2, 3, 4, 5, and 6 GeV.

A	Event ratio	Mean [GeV]	Std. Dev. [GeV]	Truth FSR Eff.	Fake FSR rate
Default	87%	126.3	4.2	41%	0.39
2 GeV	88%	124.9	4.1	31%	0.20
3 GeV	88%	124.4	4.1	26%	0.12
4 GeV	87%	124	4.1	22%	0.08
5 GeV	87%	123.6	4.1	19%	0.06
6 GeV	87%	123.4	4.2	17%	0.05

Since there are ambiguities in the matching between the FSR candidate and muons when there are three or more muons, the FSR recovery is applied only to the events selected by the ggF and VBF selection.

4.4.4 Performance of FSR recovery

The FSR candidates are found in 4.8% of total ggF signal events. The effect on the invariant mass distribution for ggF signal MC of the FSR recovery is shown in Figure 4.36. The standard deviation of the invariant mass for all events is reduced by around 3%. The mean of the invariant mass for all events approaches 125 GeV. In addition, the number of signal events in 120 GeV $< m_{\mu\mu} <$ 130 GeV and in 110 GeV $< m_{\mu\mu} <$ 160 GeV increases by 3.0% and 1.4%, respectively.

The FSR candidates are found in 3.5% of total Drell-Yan background events. The effect on the invariant mass distribution of the FSR recovery for the Drell-Yan background MC is shown in Figure 4.37. The number of signal events in 120 GeV $< m_{\mu\mu} <$ 130 GeV increases by around 2%.

The expected significance calculated by $S/\sqrt{B}$ increases by around 4% with respect to the value before applying FSR recovery.

4.4.5 Validation using $Z \to \mu\mu$ data

The performance of the FSR recovery is validated using $Z \to \mu\mu$ data and the background MC around the Z-mass peak. The background MC includes $Z \to \mu\mu$, diboson, and top MC samples. In both data and MC, the events for the validation are selected by the same dimuon selection as for the final analysis except for the invariant mass.

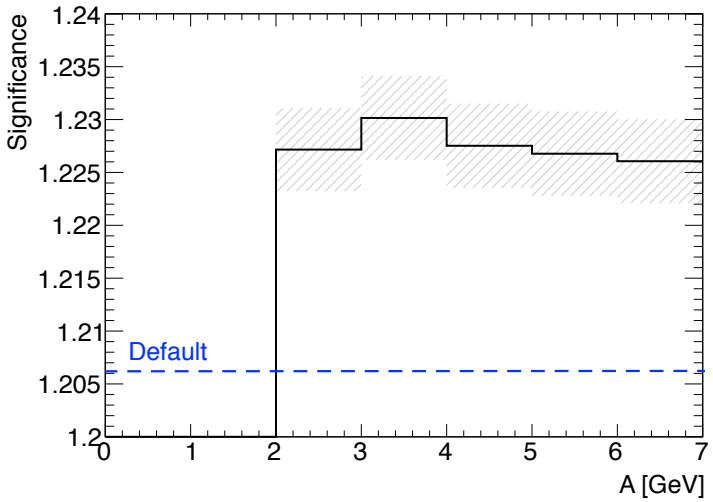


Figure 4.34 The expected significance depending on A.

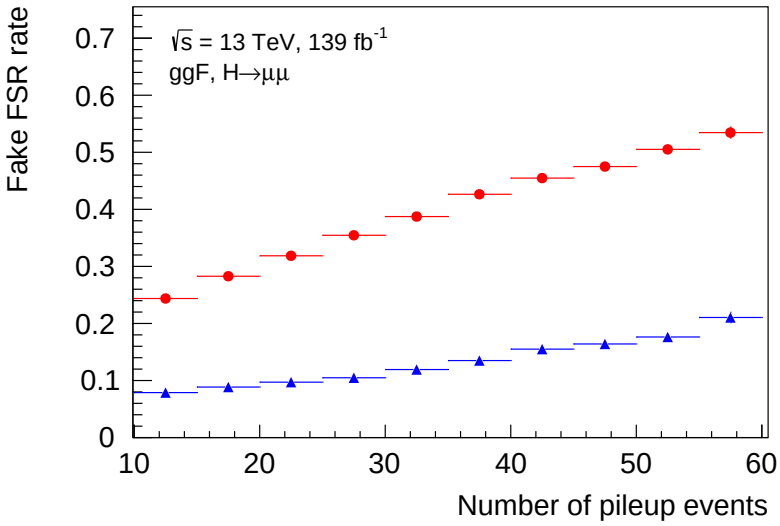


Figure 4.35 The fake FSR rate depending on pile-up before (red) and after (blue) applying the requirement of E_{T} [GeV] > 25 [GeV] $\cdot\, \Delta R_{\mathrm{cand},\mu} + 3$ [GeV].

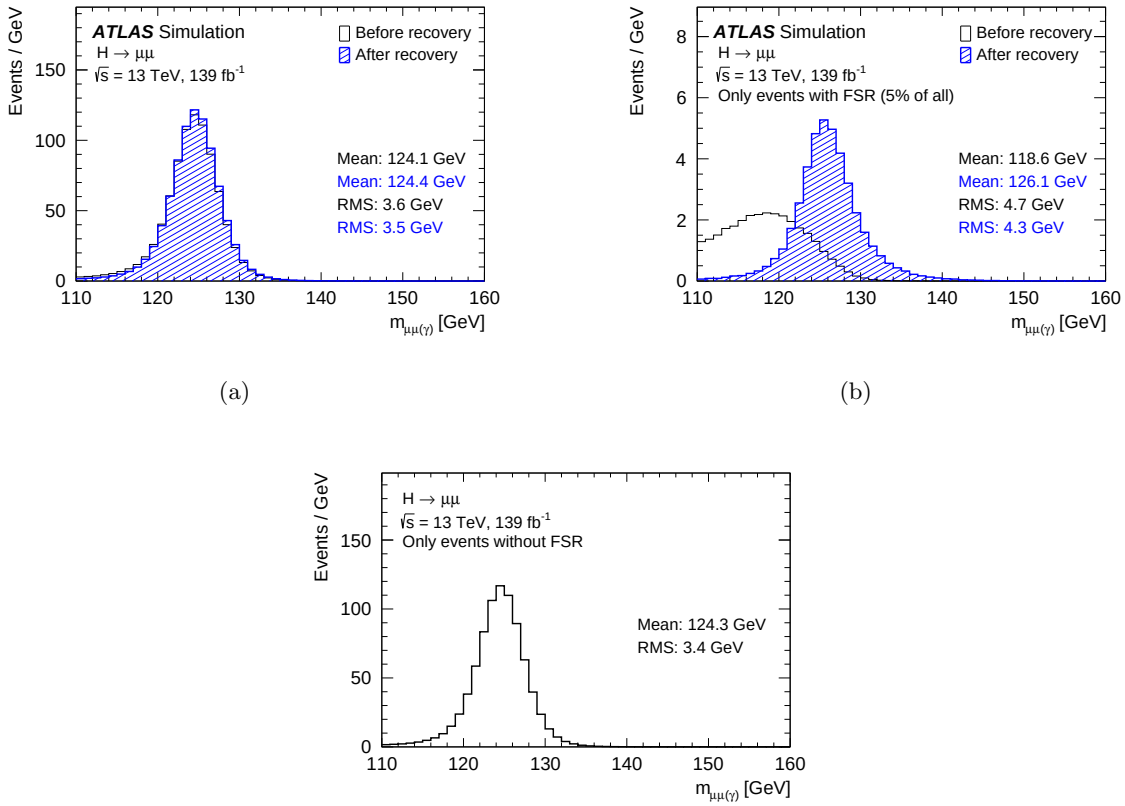
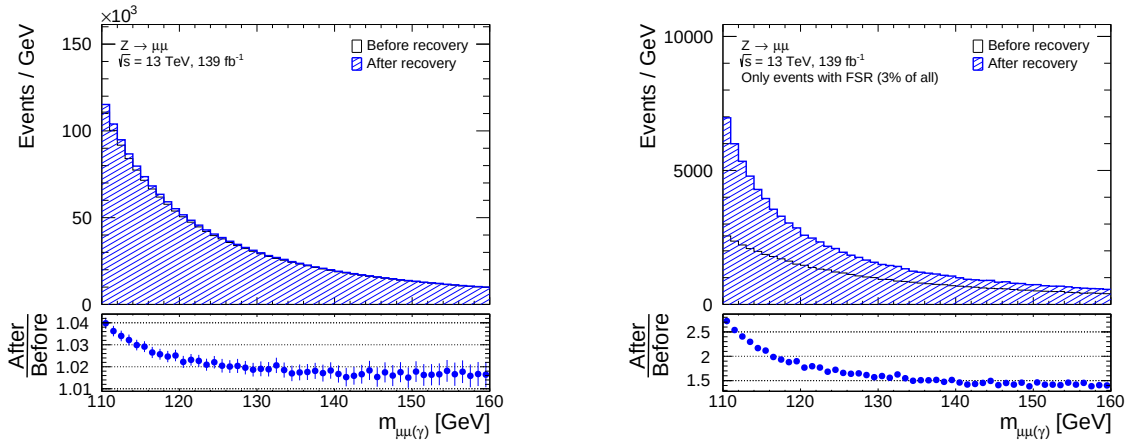


Figure 4.36 The invariant mass distributions for all events (a), the events with FSR candidates (b) and the events without FSR candidate (c). For (a) and (b), the black and blue histograms are before and after FSR recovery, respectively.



(a) (b)

Figure 4.37 The invariant mass distributions of the Drell-Yan background MC sample for all events (a) and for the events with FSR candidates (b) before (black) and after (blue) FSR recovery.

The event with the FSR candidate is found in 2.9% of the total events for both data and MC in 70 GeV $< m_{\mu\mu} <$ 110 GeV. In the $m_{\mu\mu}$ of 110–120 GeV or 130–160 GeV, the events with the FSR candidate are found in 2.3% and 2.2% of the total events for data and MC, respectively. The comparison of the invariant mass distributions between data and MC before and after the FSR recovery is shown in Figure 4.38. These show good agreement between data and MC.

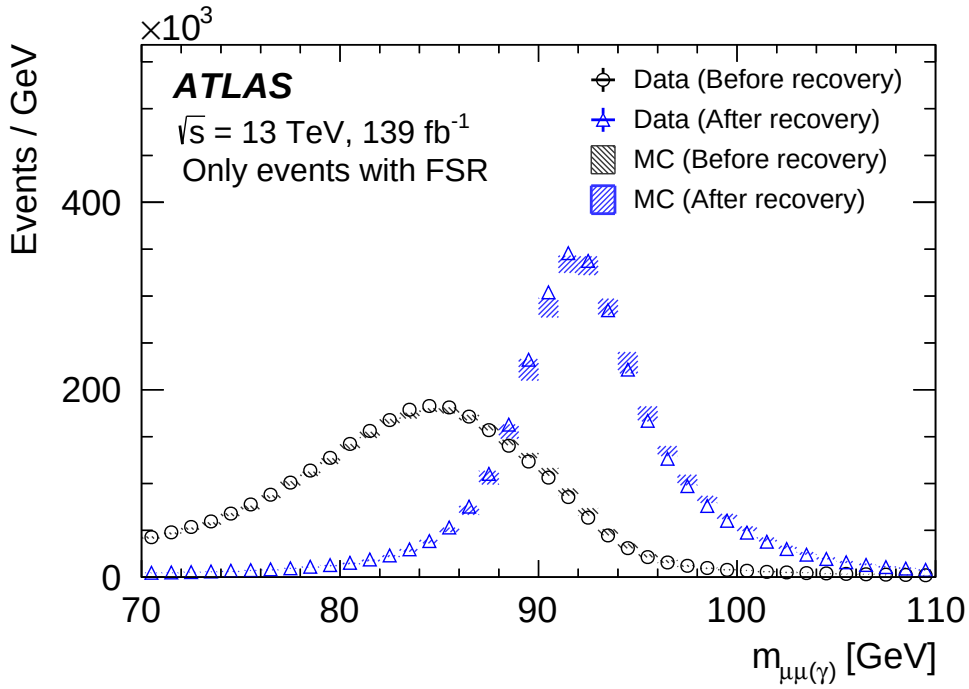


Figure 4.38 The invariant mass distributions for the $Z \to \mu\mu$ events with a reconstructed FSR photon candidate around the Z-mass peak. The black circles (black hatched bands) represent the distribution before the FSR recovery in data (MC). The blue triangles (blue hatched bands) represent the distribution after the FSR recovery in data (MC). The MC bands are scaled to 139 fb^{-1} and include the uncertainties due to the MC statistics and the muon reconstruction.

4.5 Event categorization

To improve the sensitivity of the search by isolating regions of higher signal-to-background ratio, the events selected for each production described in Section 4.3 are categorized into 20 mutually exclusive categories. There are sixteen, three, and one categories targeting the ggF/VBF, VH, and ttH production processes, respectively. The categorization is based on the presence of additional leptons, the number of jets and b-tagged jets, and classifiers of boosted decision trees (BDT) [81, 82] trained using the XGBoost package [83]. The XGBoost package is a scalable machine learning system for tree boosting, which allows heavy training to be performed efficiently and is used widely in the elementary particle experiment [84]. For each event selected by different selections, the different BDT classifiers are used.

4.5.1 ttH categorization

The events selected by the ttH selection are further selected by an XGBoost BDT classifier. The classifier takes advantage of the complex final state to separate from background. The classifier is trained with the $ttH, H \to \mu\mu$ signal MC and background MC which includes Drell-Yan, $t\bar{t}$, diboson, and single top-quark events. The events for training samples are required to pass the ttH selection and to be in the range 100 GeV $< m_{\mu\mu} <$ 200 GeV.

The 4-fold training is performed to avoid the bias from overtraining on the test samples. Both the signal and background samples are split into four samples based on the value of the event number divided by 4. In each fold, two samples corresponding to 50% of the full sample are used for training. One of the rest samples corresponding to 25% of the full sample is used for validation where the hyperparameters and the input variables are determined. The remaining sample is used as the test sample for the decision of output score. In total four folds can be formed by changing the combination of the samples as shown in Table 4.10. The training, the validation, and the test are repeated using four folds to avoid the bias due to the over training.

Table 4.10 The 4-fold training method.

	Sample 0	Sample 1	Sample 2	Sample 3
Fold-0	Training	Training	Validation	Test
Fold-1	Test	Training	Training	Validation
Fold-2	Validation	Test	Training	Training
Fold-3	Training	Validation	Test	Training

The set of input variables is chosen with the fewest training variables while keeping a similar performance. The following 12 variables in total are used as inputs of the classifier.

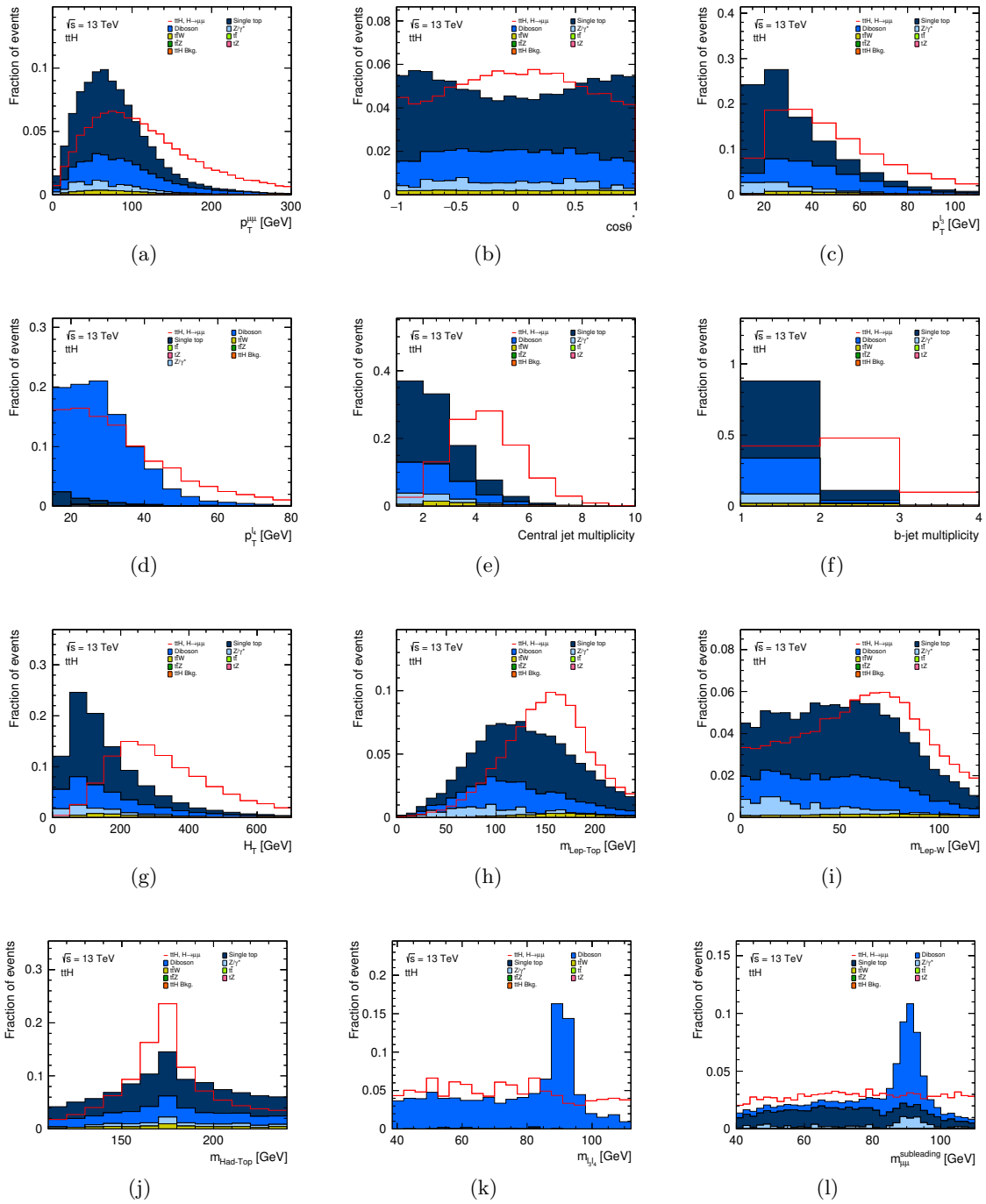
- $p_{\mathrm{T}}^{\mu\mu}$: transverse momentum of the Higgs boson candidate
- $\cos\theta^*$: cosine of the lepton decay angle in the Collins-Soper frame [85]

- $p_{\mathrm{T}}^{l_3}$: transverse momentum of the additional lepton (l_3) with the highest p_{T}
- $p_{\mathrm{T}}^{l_4}$: transverse momentum of the additional lepton (l_4) with the second highest p_{T} (only the case there are two additional leptons in the event)
- central jet multiplicity: number of jets in $|\eta| < 2.5$ in the event
- b-jet multiplicity: number of b-tagged jets selected with 85% WP in the event
- H_{T}: scalar sum of the transverse momenta of all jets in the event
- $m_{\text{Lep-Top}}$: transverse mass of the top-quark candidate with associated the W boson candidate decaying leptonically. The system is composed of l_3, the missing transverse momentum and the b-tagged jet. If there are more than one b-tagged jet in the event, the b-jet giving the closest value to 173 GeV is used.
- $m_{\text{Lep-W}}$: transverse mass of the W boson candidate decaying leptonically. The system is composed of l_3 and the missing transverse momentum.
- $m_{\text{Had-Top}}$: mass of the top-quark candidate with associated the W boson candidate decaying hadronically. The system is composed of three jets, where at least one jet is b-tagged. If there is only one b-tagged jet in the event, the same b-tagged jet is used as $m_{\text{Lep-Top}}$. If there are more than three jets in the event, the jet pair is used that maximizes the probability of compatibility of $m_{\text{Had-Top}}$ and $m_{\text{Had-W}}$ with top-quark mass and W boson mass, respectively. Here, $m_{\text{Had-W}}$ is the mass of the W boson candidate decaying hadronically and is calculated using two non-b-tagged jets.
- $m_{l_3 l_4}$: mass of l_3 and l_4 if they are both muons or electrons with opposite charge (only the case there are two additional leptons in the event)
- $m_{\mu\mu}^{\text{subleading}}$: if there is an additional opposite charge muon pair in the event, the mass of the muon pair is used.

The distributions of the input variables for the events passing the basic selection are shown in Figure 4.39.

In the BDT selection, the cut based on the output score of the classifier O_{ttH} is applied. Figure 4.40 shows the O_{ttH} distributions for MC samples. The cut value is determined so that the value is as loose as possible and is close to the maximum significance. The events with $O_{\text{ttH}} > 0.35$ pass the selection. They are categorized as the events in the ttH category. The signal efficiency of the $O_{\text{ttH}} > 0.35$ cut is 65%. The dimuon invariant mass distribution for the ttH category is shown in Figure 4.41.

The signal-to-background ratio in the region 120 GeV $< m_{\mu\mu} <$ 130 GeV is 0.08. Assuming the SM Higgs boson, 1.2 signal events are expected in the ttH category. The expected number of produced events of $ttH, H \to \mu\mu$ is 5.8 in the $\sqrt{s} = 13$ TeV 139 fb^{-1} data. Therefore, the selection efficiency for $ttH, H \to \mu\mu$ events is $\sim 21\%$. The purity of $ttH, H \to \mu\mu$ events against the signal events of all Higgs boson production is 98%.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

Figure 4.39 The BDT input variables for ttH classifier. The distributions are shown for MC samples.

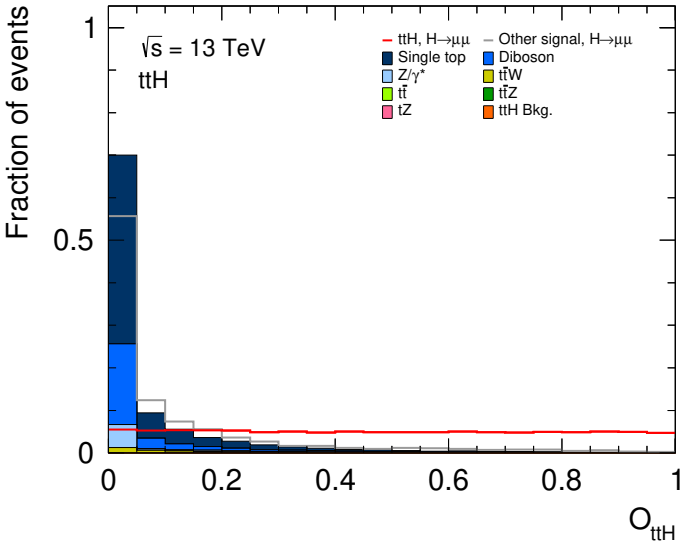


Figure 4.40 The BDT output score of ttH classifier O_{ttH} for MC samples.

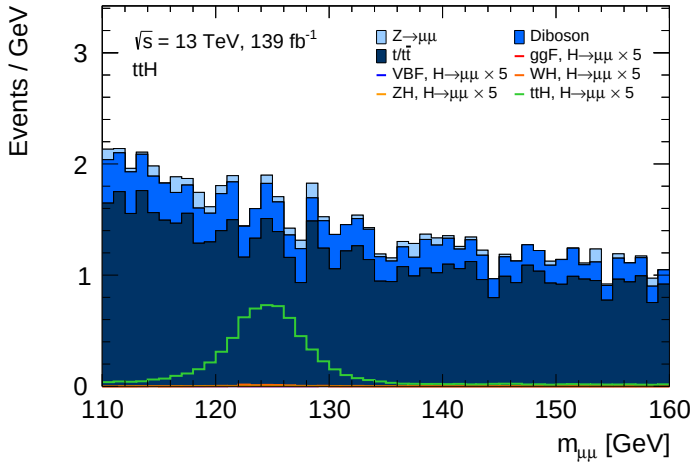


Figure 4.41 The dimuon invariant mass distribution for the ttH category.

4.5.2 VH categorization

In order to better separate signal from background using the difference in the final state between VH3L and VH4L, different classifiers are used for VH3L and VH4L. The VH3L classifier and the VH4L classifier use the same setup but are separately trained. The $W(\to l\nu)H(\to \mu\mu)$ and $Z(\to ll)H(\to \mu\mu)$ MC samples are used as signal sample for the VH3L and VH4L classifier, respectively. For both classifiers, background MC samples of Drell-Yan, $t\bar{t}$, diboson, and single top-quark events are used as background sample. The events for training are required to pass the pre-selection and basic selection and to have $m_{\mu\mu}$ in the range 110 GeV $< m_{\mu\mu} <$ 160 GeV. A 5-fold training method is used to avoid the bias from overtraining on the test samples. Both the signal and background samples are split into five samples based on the value of the event number divided by 5. In each fold, three samples are used for training. One of the rest samples is used for the validation and the other is used as test sample.

For VH3L, the following eight variables are used as the inputs of the VH3L classifier. Figure 4.42 shows the distributions of the input variables.

- $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mu\mu)$: difference of the azimuthal angle between the Higgs boson candidate and the missing transverse momentum
- $p_{\mathrm{T}}^{l_3}$: transverse momentum of the additional lepton
- $m_{\text{Lep-W}}$: transverse mass of the candidate of the W boson decaying leptonically. The system is composed of the additional lepton and the missing transverse momentum.
- $\Delta\phi(l_3, \mu\mu)$: difference of the azimuthal angle between the Higgs boson candidate and the additional lepton
- $\Delta\eta(l_3, \mu\mu)$: difference of the pseudorapidity between the Higgs boson candidate and the additional lepton
- $E_{\mathrm{T}}^{\mathrm{miss}}$: the missing transverse momentum
- $p_{\mathrm{T}}^{j_1}$: transverse momentum of the jet with the highest p_{T}
- jet multiplicity: number of jets in the event

For VH4L, the following seven variables are used as the inputs of the VH4L classifier. Figure 4.43 shows the distributions of the input variables.

- $\Delta\phi(l_3, l_4)$: difference of the azimuthal angle between the additional two leptons with the highest p_{T}
- $\Delta\eta(Z, \mu\mu)$: difference of the pseudorapidity between the Higgs boson candidate and the Z candidate
- $\Delta\phi(Z, \mu\mu)$: difference of the azimuthal angle between the Higgs boson candidate and the Z candidate
- m_Z: mass of the additional electron pair or muon pair

- jet multiplicity: number of jets in the event
- $p_{\mathrm{T}}^{j_1}$: transverse momentum of the jet with the highest p_{T}
- $p_{\mathrm{T}}^{j_2}$: transverse momentum of the jet with the second highest p_{T}

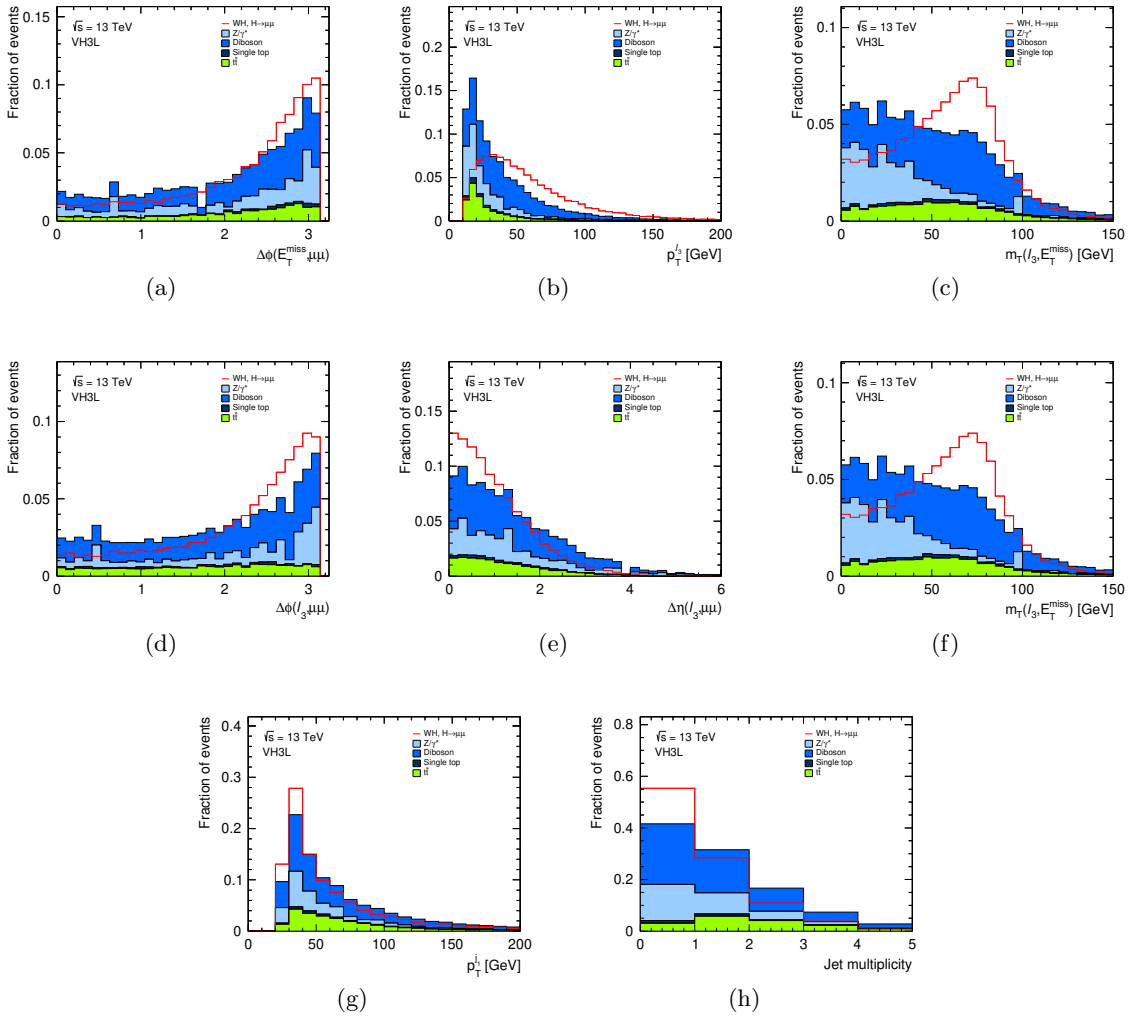


Figure 4.42 The BDT input variables for VH3L classifier. The distributions are shown for MC samples.

The output scores of the classifiers for VH3L and VH4L, referred to as O_{VH3L} and O_{VH4L}, are obtained as shown in Figure 4.44. The cut value for categorization is determined so that the value is as loose as possible and is close to the maximum significance. In addition, the VH3L classifier categorizes the events into two categories to isolate regions of higher signal-to-background ratio.

For the VH3L, the events with $O_{\mathrm{VH3L}} > 0.1$ pass the BDT selection. In addition, the events with $O_{\mathrm{VH3L}} > 0.7$, corresponding to the region with higher signal-to-background ratio, are categorized as the events in the VH3L-High category. The remaining events in the VH3L

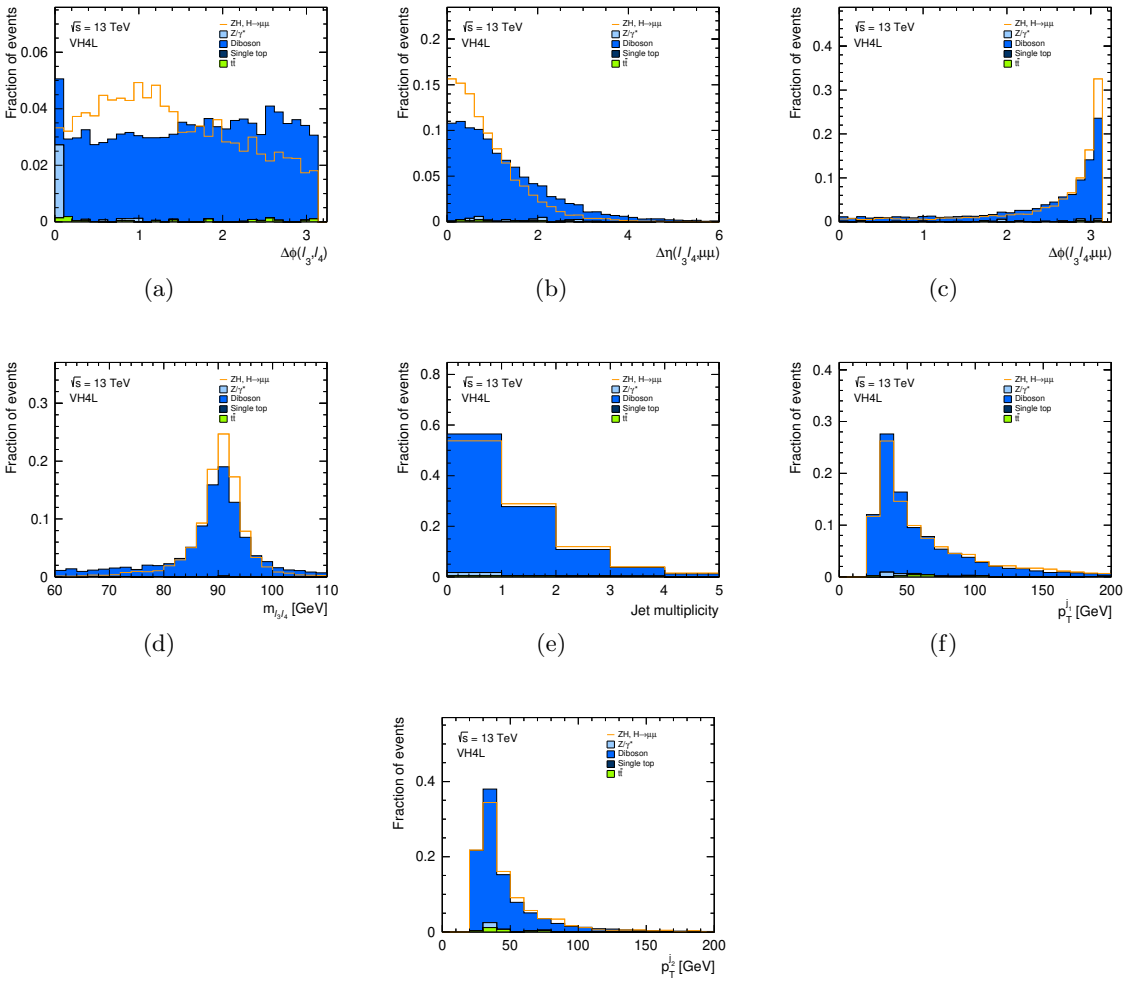
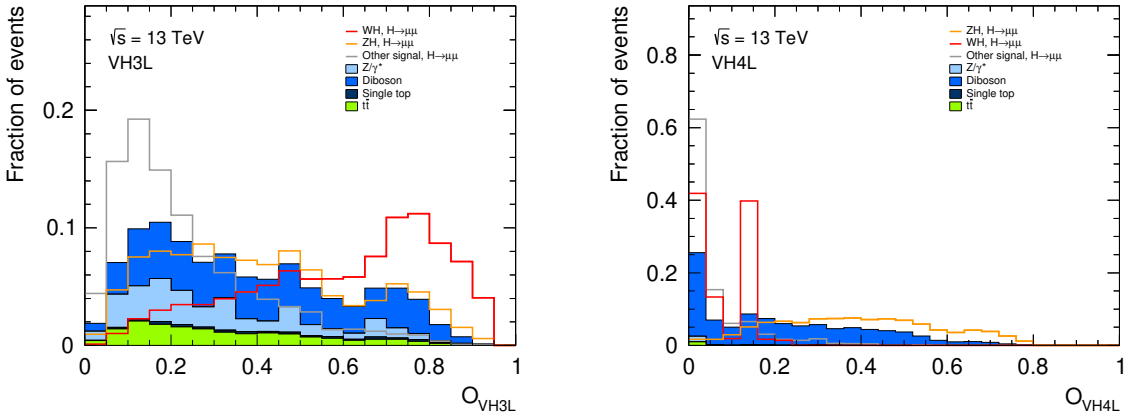


Figure 4.43 The BDT input variables for VH4L classifier. The distributions are shown for MC samples.

($0.1 < O_{\text{VH3L}} < 0.7$) are categorized as the events in the VH3L-Medium category. For the VH4L, the events with $O_{\text{VH4L}} > 0.12$ pass the BDT selection and categorized as the events in the VH4L category.

The signal-to-background ratios for VH3L-High, VH3L-Medium and VH4L categories in the region 120 GeV $< m_{\mu\mu} <$ 130 GeV are 0.037, 0.008 and 0.026, respectively. The purities, defined as the ratio between $H \to \mu\mu$ events in WH production and signal events in all Higgs boson productions, for VH3L-Medium and VH3L-High categories are 89% and greater than 99%, respectively. The purity of the $H \to \mu\mu$ events in ZH production to signal events in all Higgs boson productions for VH4L category is greater than 99%.



(a) (b)

Figure 4.44 The BDT output scores of VH3L and VH4L classifiers for MC samples.

4.5.3 ggF and VBF categorization

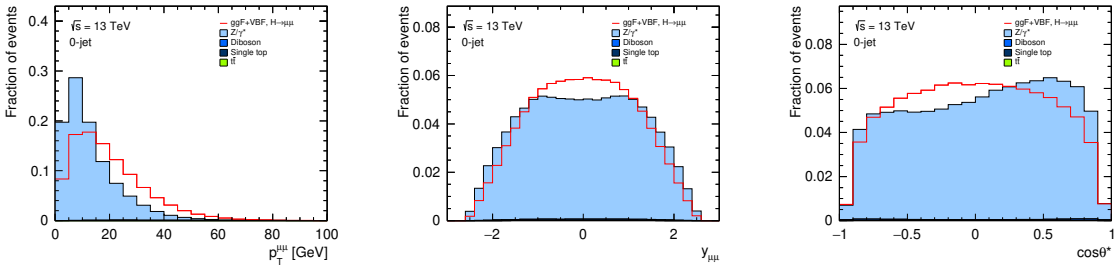
The ggF and VBF categorization is based on the number of jets and BDT classifiers. The events passing the ggF/VBF selection are first divided into three channels based on jet multiplicity n_j: 0-Jet, 1-Jet and 2-Jet channels. The 0-Jet, 1-Jet and 2-Jet channels include the events with $n_j = 0$, $n_j = 1$ and $n_j \geq 2$, respectively. The events in the 0-Jet and 1-Jet channels are classified into four categories each based on BDT classifiers. In the 2-Jet channels, the events are classified into four VBF-enriched categories and four ggF-enriched categories based on BDT classifiers. Different classifiers are used and trained for the three channels. In the 2-Jet channel, the classifier specialized for VBF process, referred to as the VBF classifier, selects VBF-like events first. The remaining events in the 2-Jet channel and the events in the 0-Jet and 1-Jet are classified using the classifiers trained to separate ggF+VBF signal from background. They are referred to as Higgs-2Jet, Higgs-1Jet and Higgs-0Jet classifiers, respectively. The Higgs-0Jet, Higgs-1Jet, and Higgs-2Jet classifiers are trained using the MC samples of the $H \to \mu\mu$ decay in ggF and VBF production as a signal sample. The background MC samples including Drell-Yan, $t\bar{t}$, diboson and single top-quark are used as background sample. The VBF classifier uses the MC

samples of the $H \to \mu\mu$ decay in VBF production as a signal sample and the same background MC as other classifiers as background. For all classifiers, the events for training are required to pass the pre-selection and the common ggF/VBF selection and to have $m_{\mu\mu}$ in the range 120 GeV $< m_{\mu\mu} <$ 130 GeV. These classifiers are trained using the 4-fold training method. The input variables are different for jet multiplicity channels. The variables are listed in Table 4.11. The distributions of the input variables for the events are shown in Figures 4.45–4.47.

Table 4.11 The input variables of the classifiers for the VBF and ggF channels.

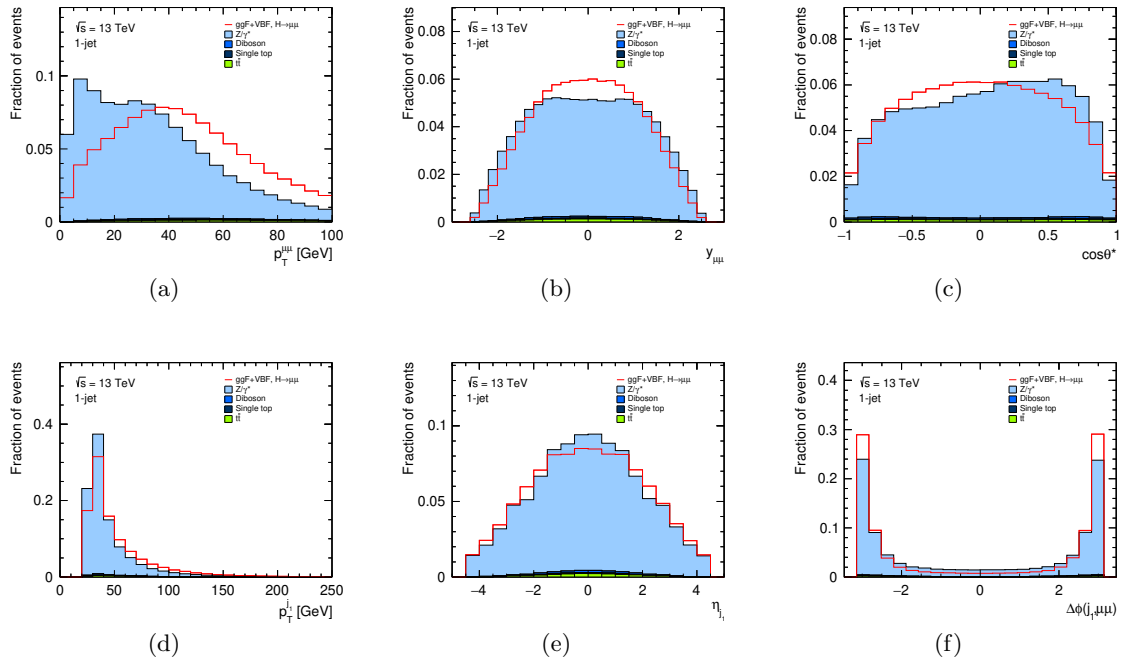
Variables	Description	Channel
$p_{\mathrm{T}}^{\mu\mu}$	transverse momentum of the dimuon system	0/1/2-Jet
$Y_{\mu\mu}$	rapidity of the dimuon system	0/1/2-Jet
$\cos\theta^*$	cosine of the dimuon decay angle in the Collins-Soper frame	0/1/2-Jet
$p_{\mathrm{T}}^{j_1}$	transverse momentum of the jet (j_1) with the highest p_{T}	1/2-Jet
η_{j_1}	pseudorapidity of the jet j_1	1/2-Jet
$\Delta\phi_{j_1,\mu\mu}$	azimuthal angle difference between the jet j_1 and the dimuon system	1/2-Jet
$N_{\mathrm{track}}^{j_1}$	multiplicity of the ID tracks with $p_{\mathrm{T}} > 0.5$ GeV in the jet j_1 (available if $p_{\mathrm{T}}^{j_1} > 50$ GeV and $\lvert\eta_{j_1}\rvert < 2.1$)	1/2-Jet
$p_{\mathrm{T}}^{j_2}$	transverse momentum of the jet (j_2) with the second highest p_{T}	2-Jet
η_{j_2}	pseudorapidity of the jet j_2	2-Jet
$\Delta\phi_{j_2,\mu\mu}$	azimuthal angle difference between the jet j_2 and the dimuon system	2-Jet
$N_{\mathrm{track}}^{j_2}$	multiplicity of the ID tracks with $p_{\mathrm{T}} > 0.5$ GeV in the jet j_2 (available if $p_{\mathrm{T}}^{j_2} > 50$ GeV and $\lvert\eta_{j_2}\rvert < 2.1$)	2-Jet
p_{T}^{jj}	transverse momentum of the dijet system (using j_1 and j_2)	2-Jet
Y_{jj}	rapidity of the dijet system	2-Jet
$\Delta\phi_{jj,\mu\mu}$	azimuthal angle difference between the dijet and the dimuon systems	2-Jet
m_{jj}	invariant mass of the dijet system	2-Jet
$E_{\mathrm{T}}^{\mathrm{miss}}$	missing transverse mass in the event	2-Jet
H_T	scalar sum of the transverse momentum of the jets	2-Jet

In the 0-Jet and 1-Jet channels, the events are classified into four categories each based on the output scores, $O_{ggF}^{(0)}$ and $O_{ggF}^{(1)}$, of the Higgs-0Jet and Higgs-1Jet classifiers, respectively. In the 2-Jet channel, the events with high output score, O_{VBF}, of the VBF classifier are selected and



(a) (b) (c)

Figure 4.45 The input variables for the classifier in the 0-Jet channel. The distributions are shown for MC samples.



(a) (b) (c)

(d) (e) (f)

(g)

Figure 4.46 The input variables for the classifier in the 1-Jet channel. The distributions are shown for MC samples.

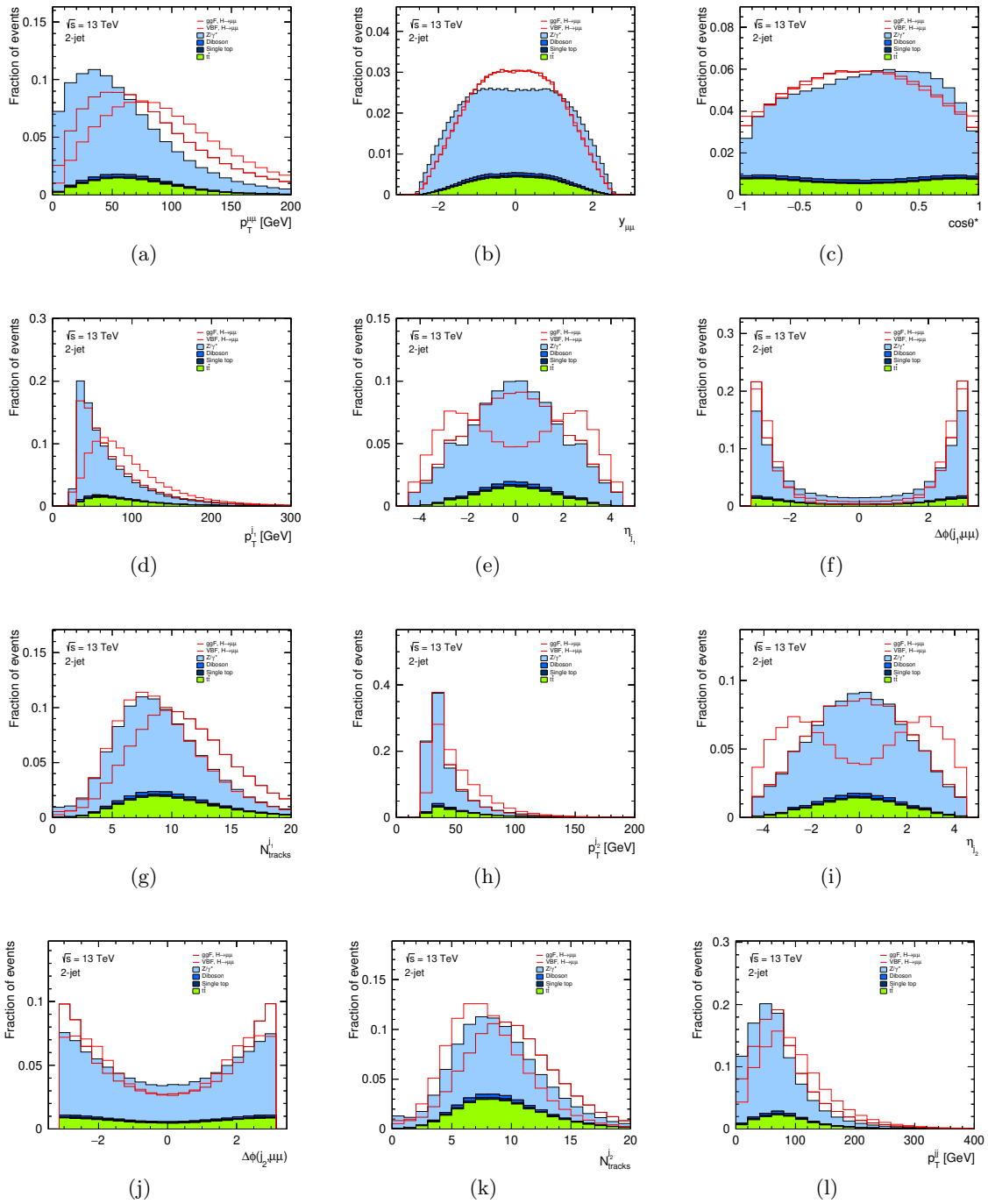


Figure 4.47 The input variables for the classifiers in the 2-Jet channel. The distributions are shown for MC samples.

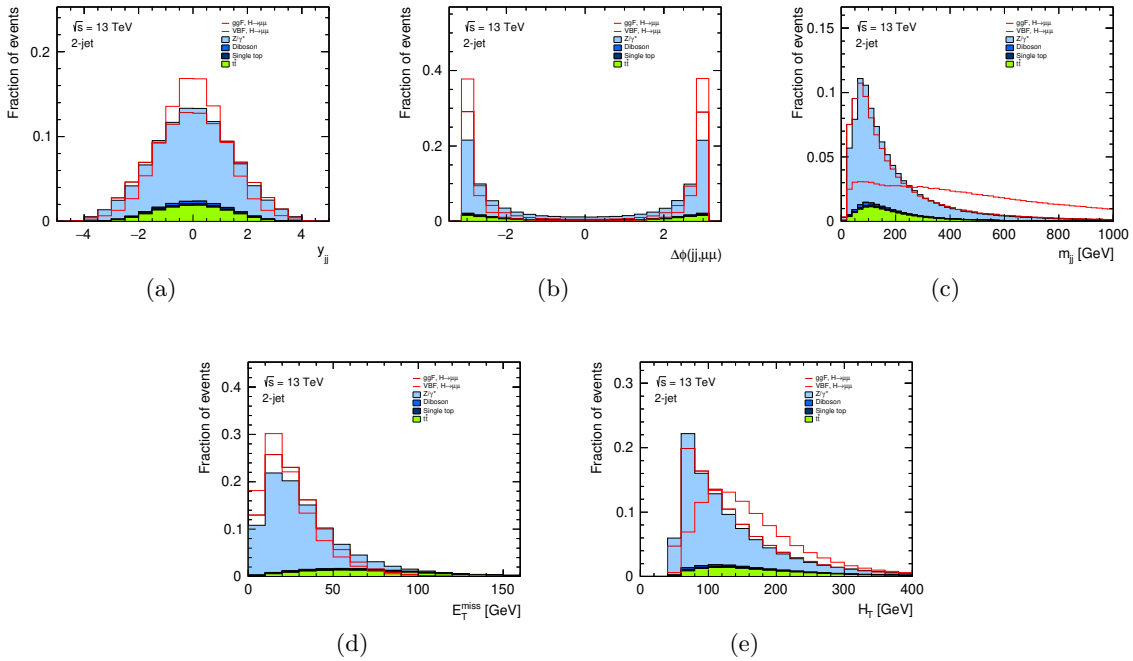


Figure 4.47 The input variables for the classifiers in the 2-Jet channel (continued). The distributions are shown for MC samples.

classified into four categories based on the score. The remaining events in the 2-Jet channel are classified into four categories based on the output score, $O^{(2)}_{ggF}$, of the Higgs-2Jet classifier. The four categories classified by Higgs-NJet classifier ($N = 0, 1, 2$) are referred to as NJet-VeryHigh, NJet-High, NJet-Medium, and NJet-Low in descending order of purity of the $ggF, H \to \mu\mu$ events. The four categories classified by VBF classifier are referred to as VBF-VeryHigh, VBF-High, VBF-Medium, and VBF-Low in descending order of purity of the VBF, $H \to \mu\mu$ events. The output score distributions are shown in Figure 4.48. The criteria of the categorization based on $O^{(N)}_{ggF}$ and O_{VBF} are summarized in Table 4.12. In each classifier, the boundaries of score for categorization are simultaneously optimized to maximize the total expected significance.

4.5.4 Categorization summary

In total, the events are classified into 20 categories: 4 VBF, 12 ggF+VBF, 3 VH, and 1 ttH enriched categories. Figure 4.49 shows the signal and background compositions for the 20 categories in 120 GeV $< m_{\mu\mu} <$ 130 GeV. The signal and background compositions depend on the categories. The signal-to-background ratio varies between 0.001 and 0.20 across the categories.

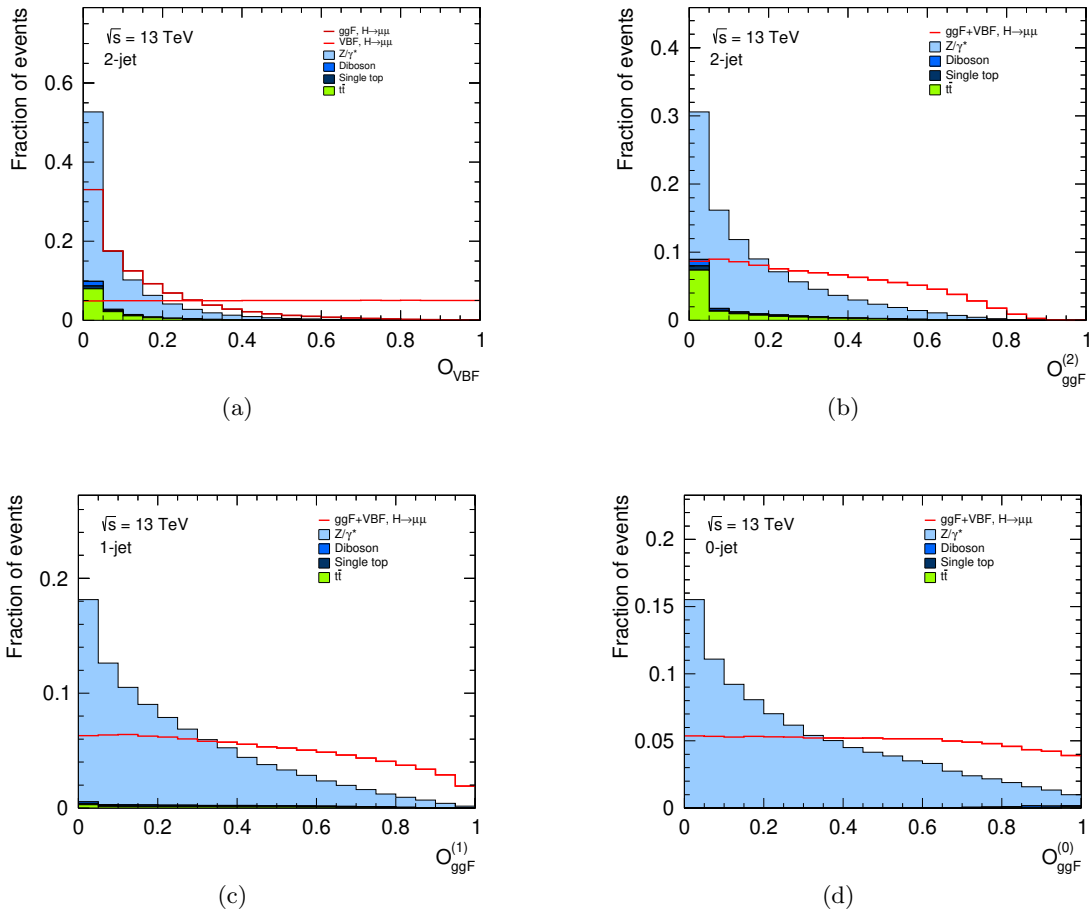


Figure 4.48 The output scores of the VBF (a), the Higgs-2Jet (b), the Higgs-1Jet (c), and the Higgs-0Jet (d) classifiers for MC samples.

Table 4.12 The output score boundaries in each category.

Category	Classifier			
	$N_j = 2$		$N_j = 1$	$N_j = 0$
	VBF	Higgs-2Jet ($O_{\mathrm{VBF}} < 0.62$)	Higgs-1Jet	Higgs-0Jet
VeryHigh	$O_{\mathrm{VBF}} \geq 0.93$	$O^{(2)}_{ggF} \geq 0.65$	$O^{(1)}_{ggF} \geq 0.88$	$O^{(0)}_{ggF} \geq 0.81$
High	$0.85 \leq O_{\mathrm{VBF}} < 0.93$	$0.42 \leq O^{(2)}_{ggF} < 0.65$	$0.67 \leq O^{(1)}_{ggF} < 0.88$	$0.53 \leq O^{(0)}_{ggF} < 0.81$
Medium	$0.75 \leq O_{\mathrm{VBF}} < 0.85$	$0.16 \leq O^{(2)}_{ggF} < 0.42$	$0.36 \leq O^{(1)}_{ggF} < 0.67$	$0.21 \leq O^{(0)}_{ggF} < 0.53$
Low	$0.62 \leq O_{\mathrm{VBF}} < 0.75$	$O^{(2)}_{ggF} < 0.16$	$O^{(1)}_{ggF} < 0.36$	$O^{(0)}_{ggF} < 0.21$

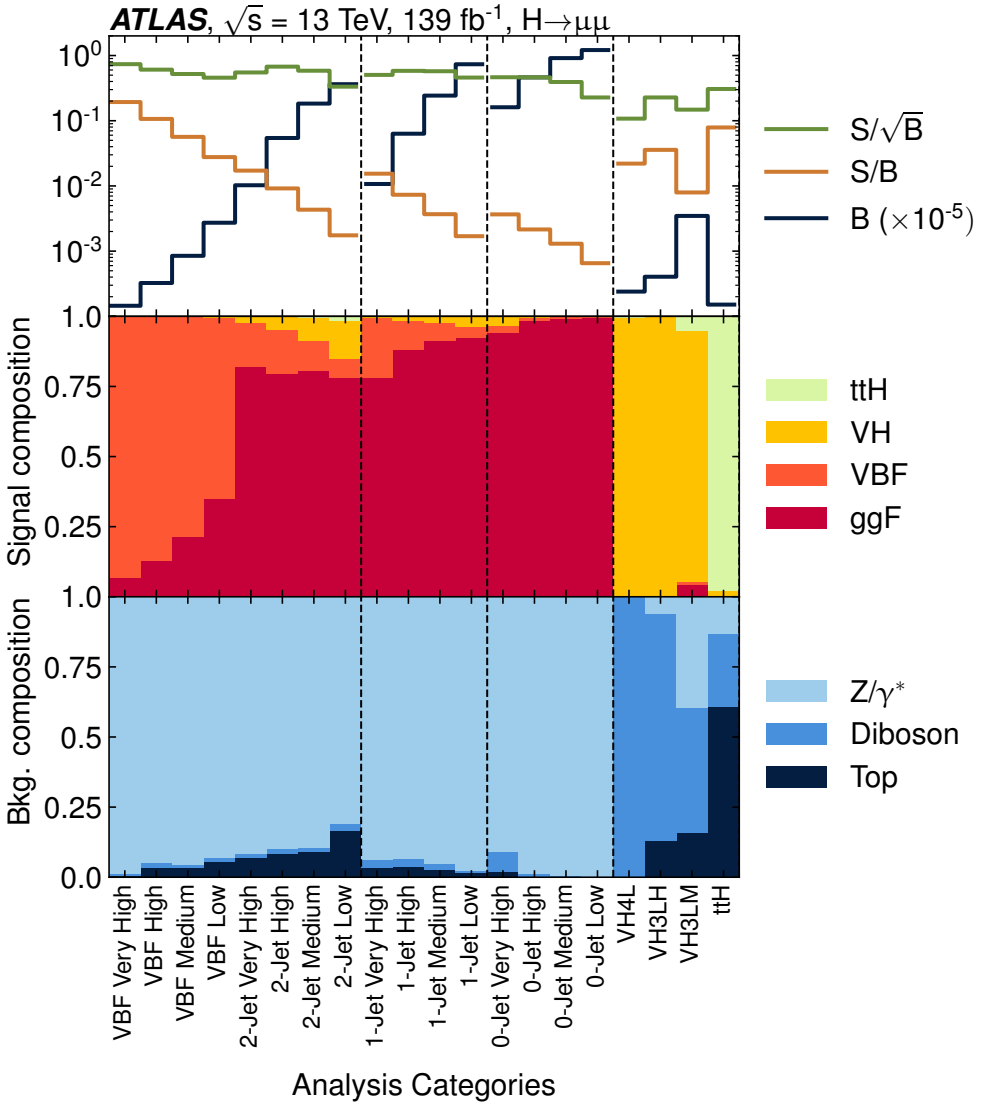


Figure 4.49 The signal and background compositions for the 20 categories in 120 GeV < $m_{\mu\mu}$ < 130 GeV. The event ratios $S/\sqrt{B}$, S/B, and B are also shown. [86]

4.6 Signal extraction

The signal yield is extracted by the simultaneous signal+background fit to the dimuon invariant mass spectrum in all categories. The fit is performed in the 110 GeV $< m_{\mu\mu} <$ 160 GeV region. The fit function for signal is modeled using high statistics signal MC samples. The background is modeled by high statistics background MC sample reweighted to data. The signal and background modeling is described in Section 4.6.1 and Section 4.6.2, respectively. In the Section 4.6.3, systematic uncertainties considered in this analysis are presented. Finally, the statistical procedure is explained in Section 4.6.4.

4.6.1 Signal modeling

The dimuon invariant mass distribution for the signal MC sample is shown in Figure 4.50. The main peak can be expressed by the Gaussian distribution since the width of the Higgs boson (4.1×10^{-3} GeV) is much smaller than the invariant mass resolution ($\sim$ 2.7 GeV). There are tails on both sides since the resolution depends on the region of the ATLAS detector where the muon passes through. On the lower-mass side, there is an additional tail component due to the $H \to \mu\mu$ events with FSR photon.

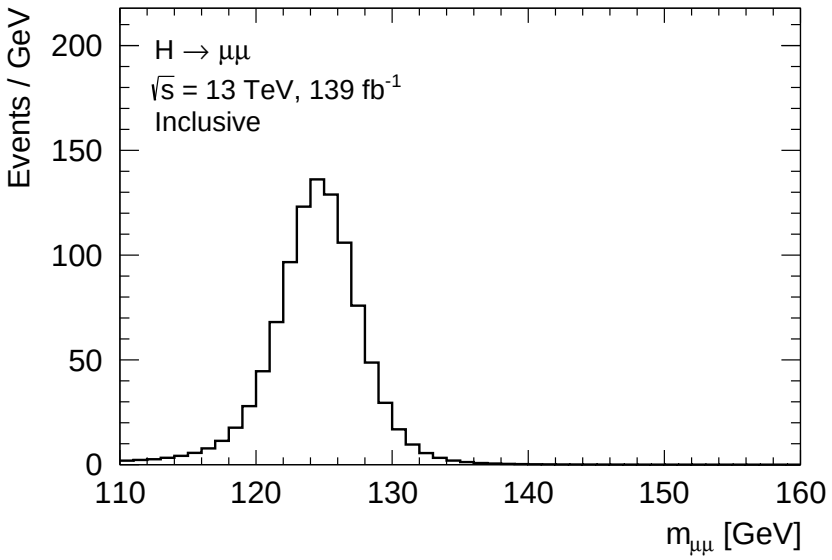


Figure 4.50 The dimuon invariant mass distribution for the $H \to \mu\mu$ MC events in inclusive category.

The function for signal is selected with considering the asymmetric mass peak and the effects from detector resolution. A Double-Sided Crystal Ball function, which is a modification of the Crystal Ball function and consists of a Gaussian with a power-law tail on each side, is used for

signal modeling. The Double-Sided Crystal Ball function has the following parameters and is defined by Eqs. (4.24) and (4.25).

- M_{CB}: mean of the Gaussian core
- σ_{CB}: width of the Gaussian core,
- α_{L}: threshold for the lower-end tail of power-law
- α_{H}: threshold for the higher-end tail of power-law
- n_{L}: power in the lower-end tail of power-law
- n_{H}: power in the higher-end tail of power-law
- N: normalization factor corresponding to signal yield

$$f_{\mathrm{sig}}(m_{\mu\mu}) = N \cdot \begin{cases} e^{-\frac{t^2}{2}} & \text{for } -\alpha_{\mathrm{L}} \leq t \leq \alpha_{\mathrm{H}} \\ e^{-\frac{\alpha_{\mathrm{L}}^2}{2}} \left\{ \frac{\alpha_{\mathrm{L}}}{n_{\mathrm{L}}} \left(\frac{\alpha_{\mathrm{L}}}{n_{\mathrm{L}}} - \alpha_{\mathrm{L}} - t \right) \right\}^{-n_{\mathrm{L}}} & \text{for } t < -\alpha_{\mathrm{L}} \\ e^{-\frac{\alpha_{\mathrm{H}}^2}{2}} \left\{ \frac{\alpha_{\mathrm{H}}}{n_{\mathrm{H}}} \left(\frac{\alpha_{\mathrm{H}}}{n_{\mathrm{H}}} - \alpha_{\mathrm{H}} + t \right) \right\}^{-n_{\mathrm{H}}} & \text{for } t > \alpha_{\mathrm{H}} \end{cases} \tag{4.24}$$

$$t = \frac{M_{\mathrm{CB}} - m_{\mu\mu}}{\sigma_{\mathrm{CB}}} \tag{4.25}$$

The fits are performed for the signal MC samples in the range 110 GeV $< m_{\mu\mu} <$ 160 GeV with all shape parameters as free parameters. The fits are independently performed in each category due to differing the shapes by the event selection and the categorization. The normalization factor (N) in each category is calculated from the total number of events. The shape parameters are determined from the inclusive MC sample, which includes all signal production modes. The same shape parameters are used in each production process in the same category since the shapes are similar. The results for all categories are shown in Figures 4.51–4.53. The σ_{CB} roughly corresponding to the invariant mass resolution varies from 2.56 GeV to 3.23 GeV. The normalization factors and the shape parameters are summarized in Tables 4.13 and 4.14.

4.6.2 Background modeling

In this analysis, the signal-to-background ratio is very small (0.2% in 120 GeV $< m_{\mu\mu} <$ 130 GeV). In order to avoid a significant bias in the extracted signal yields, the development of the function that accurately describes the background is very important. For reliable background modeling, both full and fast simulation background samples are reweighted in each category to the data sidebands. The events in the ggF categories are reweighted using second-order polynomial functions in $m_{\mu\mu}$. For the VBF, VH, and ttH categories, first-order polynomial functions are used due to limited data statistics.

The invariant mass distribution for total background MC has a steep slope in the lower mass region (110 GeV $< m_{\mu\mu} <$ 120 GeV) due to the tail of the Z-boson mass distribution with a

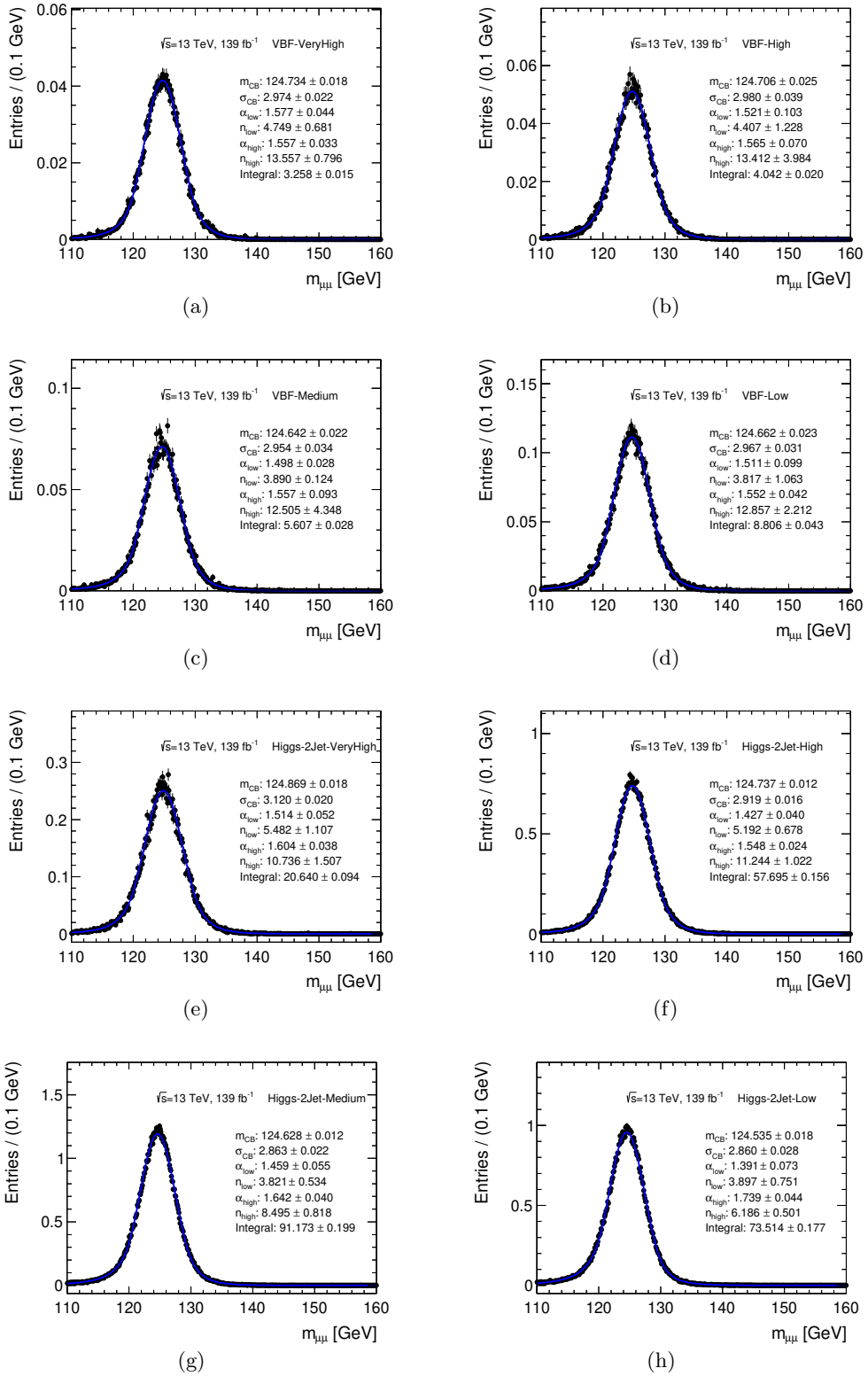


Figure 4.51 The results of signal modeling in VBF and Higgs-2Jet categories.

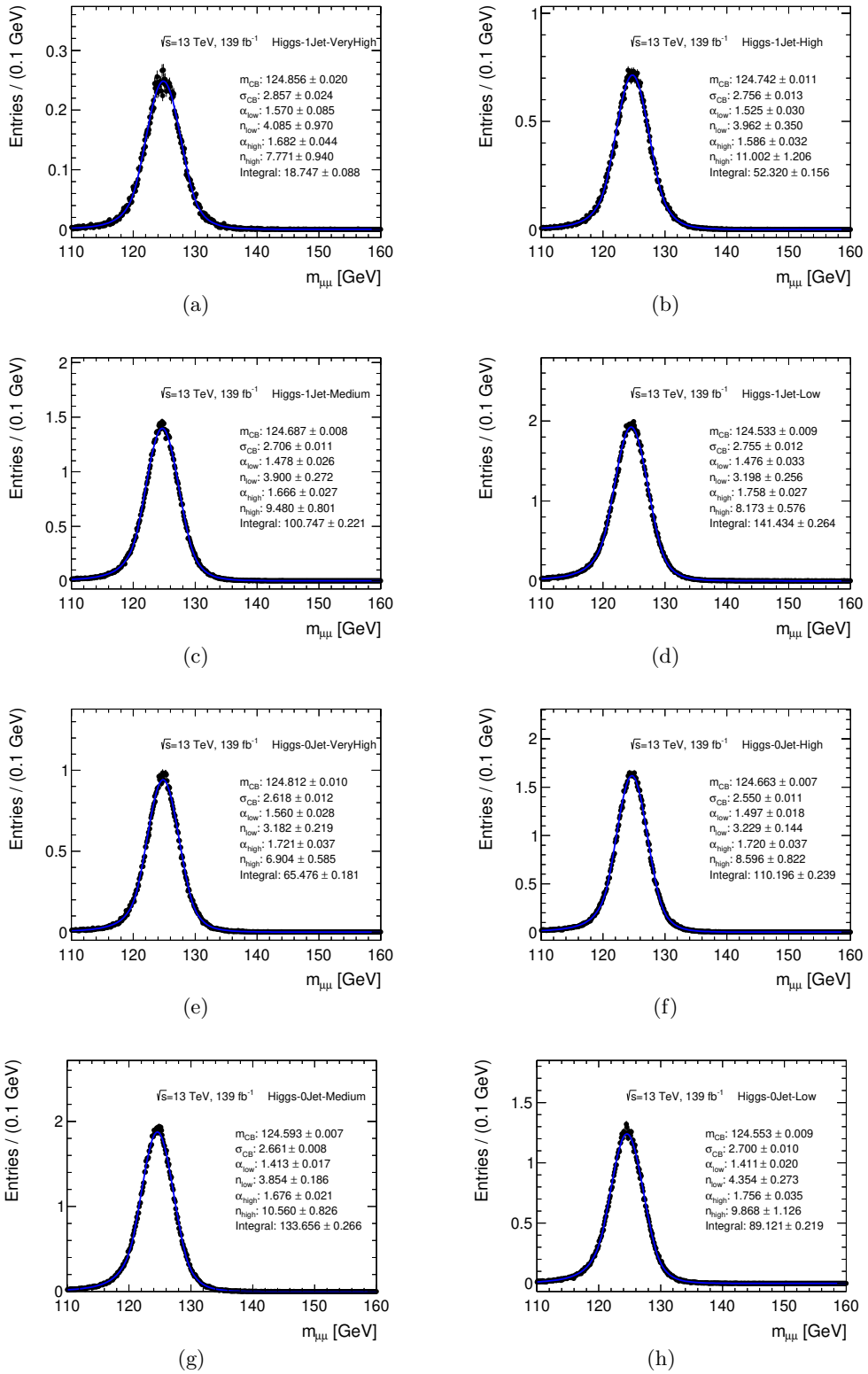


Figure 4.52 The results of signal modeling in Higgs-1Jet and Higgs-0Jet categories.

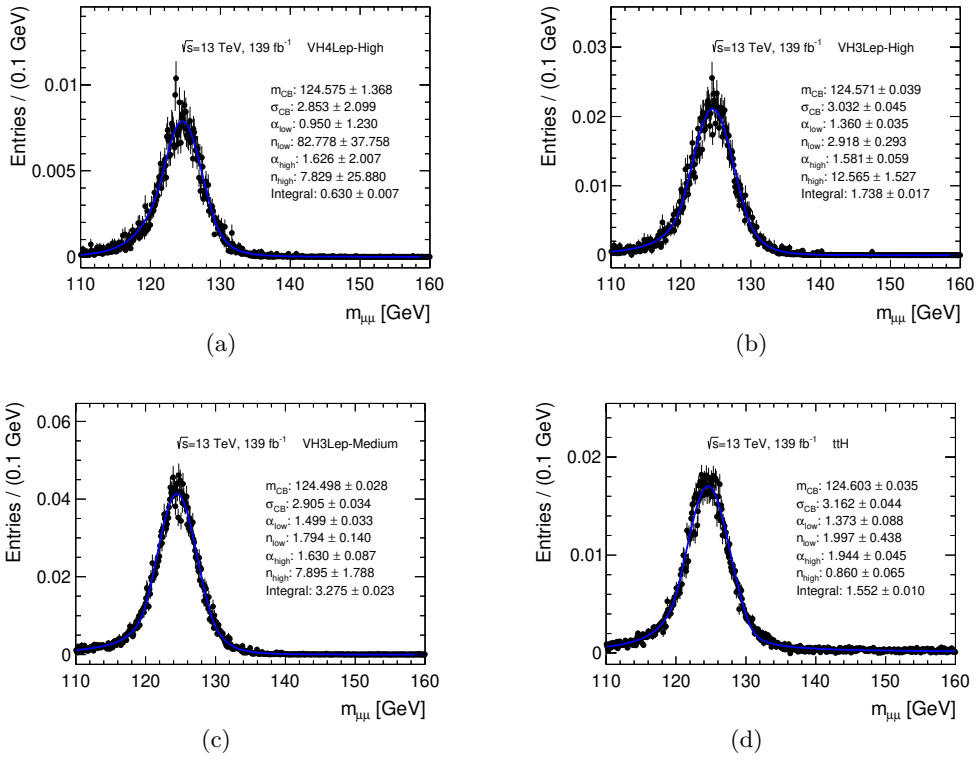


Figure 4.53 The results of signal modeling in VH and ttH categories.

Table 4.13 The normalization factors and the shape parameters for signal modeling in the VBF, Higgs-2Jet, and Higgs-1Jet categories.

Category	VBF				Higgs-2Jet				Higgs-1Jet			
	VeryHigh	High	Medium	Low	VeryHigh	High	Medium	Low	VeryHigh	High	Medium	Low
N	3.26	4.04	5.61	8.81	20.64	57.70	91.17	73.51	18.75	52.32	100.75	141.43
M_{CB}	124.73	124.71	124.64	124.66	124.87	124.74	124.63	124.54	124.86	124.74	124.69	124.53
σ_{CB}	2.97	2.98	2.95	2.97	3.12	2.92	2.86	2.86	2.86	2.76	2.71	2.76
α_{L}	1.58	1.52	1.50	1.51	1.51	1.43	1.46	1.39	1.57	1.53	1.48	1.48
α_{H}	1.56	1.57	1.56	1.55	1.60	1.55	1.64	1.74	1.68	1.59	1.67	1.76
n_{L}	4.75	4.41	3.89	3.82	5.48	5.19	3.82	3.90	4.09	3.96	3.90	3.20
n_{H}	13.56	13.41	12.51	12.86	10.74	11.24	8.50	6.19	7.77	11.00	9.48	8.17

Table 4.14 The normalization factors and the shape parameters for signal modeling in the Higgs-0Jet, VH, and ttH categories.

Category	Higgs-0Jet				VH			ttH
	VeryHigh	High	Medium	Low	4L-High	3L-High	3L-Medium	
N	65.48	110.20	133.66	89.12	0.63	1.74	3.27	1.55
M_{CB}	124.81	124.66	124.59	124.55	124.58	124.57	124.50	124.60
σ_{CB}	2.62	2.55	2.66	2.70	2.85	3.03	2.91	3.16
α_{L}	1.56	1.50	1.41	1.41	0.95	1.36	1.50	1.37
α_{H}	1.72	1.72	1.68	1.76	1.63	1.58	1.63	1.94
n_{L}	3.18	3.23	3.85	4.35	82.78	2.92	1.79	2.00
n_{H}	6.90	8.60	10.56	9.87	7.83	12.57	7.89	0.86

the peak at around 90 GeV. In order to describe the steep slope accurately, the product of core and empirical functions are used.

For the core function, LO DY analytic line-shape [87] smeared with Gaussian for mass resolution is used. The LO DY analytic line-shape is a function based on the physics for the Drell-Yan process:

$$DY(m_{\mu\mu}) = \sum_{q} \mathcal{L}_{q\bar{q}}(m_{\mu\mu}) \cdot \sigma_{q\bar{q}}(m_{\mu\mu}), \tag{4.26}$$

$$q = u, s, d. \tag{4.27}$$

The parton luminosity contribution $\mathcal{L}_{q\bar{q}}$ is derived from PDF4LHC15 set as a function of $m_{\mu\mu}$ and parameterized using a sixth-order polynomial. The cross section contribution $\sigma_{q\bar{q}}$ is expressed

as:

$$\sigma_{q\bar{q}}(m^2_{\mu\mu}) = \frac{4\pi\alpha^2}{3m^2_{\mu\mu}N_c}\left[\mathcal{Q}_q^2 - 2\mathcal{Q}_q V_l V_q \chi_{Z\gamma}(m^2_{\mu\mu}) + (A_l^2+V_l^2)(A_q^2+V_q^2)\chi_Z(m^2_{\mu\mu})\right], \quad (4.28)$$

$$\chi_{Z\gamma}(m^2_{\mu\mu}) = \kappa\frac{m^2_{\mu\mu}(m^2_{\mu\mu}-m^2_Z)}{(m^2_{\mu\mu}-m^2_Z)^2+\Gamma_Z^2 m_Z^2}, \quad (4.29)$$

$$\chi_Z(m^2_{\mu\mu}) = \kappa^2\frac{(m^2_{\mu\mu})^2}{(m^2_{\mu\mu}-m^2_Z)^2+\Gamma_Z^2 m_Z^2}, \quad (4.30)$$

$$\kappa = \frac{\sqrt{2}G_F m_Z^2}{4\pi\alpha}. \quad (4.31)$$

Here $\mathcal{Q}$, V, and A denote the electric charges, vector and axial-vector couplings of the fermions. The factors α and G_F are the electromagnetic coupling constant and the Fermi coupling constant, respectively. The parameters m_Z and Γ_Z are the mass and the width of the Z boson. $N_c = 3$ denotes the number of QCD color charges. Although multiple background processes are considered in this analysis, the Drell-Yan process is the dominant background in the most sensitive categories. For this reason, the function describing the Drell-Yan process is selected as the core function. The mass resolution is estimated by the RMS of the residual distribution between the truth-level and the reconstruction-level using the full-simulation Drell-Yan samples described in Section 4.1.2. The core function has no free parameters. The same function is used for all categories.

In order to correct for distortions of the mass shape due to event selection, categorization, higher order theory correction and other small background contributions, flexible functions are used as empirical functions. In each category, an empirical function is selected from four types of the power laws (Power) or four types of the exponentials with polynomials on the index (Epoly), as listed in Table 4.15. The number of free parameters varies from one to four depending on the functions.

The empirical function for each category is independently determined by the following criteria.

- The χ^2 probability of the fit needs to be greater than 1% when the background-only fits are performed to the following samples:
  - data sidebands (110 GeV $< m_{\mu\mu} <$ 120 GeV and 130 GeV $< m_{\mu\mu} <$ 160 GeV).
  - full-simulation background MC reweighted to data sidebands[2]:
    top + diboson + Drell-Yan Sherpa MC samples are used for the ggF and VBF categories, and
    top + diboson MC samples are used for the VH and ttH categories.
  - fast-simulation Drell-Yan Sherpa MC reweighted to data sidebands (only for the ggF and VBF categories)

[2]The ratio of the dimuon invariant mass distributions for the data and the MC samples is fitted with polynomial functions, and the MC sample is reweighted bin by bin using the fitted function. The second-order polynomial function is used for the ggF categories and the first-order polynomial function is used for the VBF, VH, and ttH categories.

Table 4.15 Candidates of the empirical functions.

Function	Expression	Number of free parameters
Power0	$m_{\mu\mu}^{a_0}$	1
Power1	$m_{\mu\mu}^{a_0+a_1 m_{\mu\mu}}$	2
Power2	$m_{\mu\mu}^{a_0+a_1 m_{\mu\mu}+a_2 m_{\mu\mu}^2}$	3
Power3	$m_{\mu\mu}^{a_0+a_1 m_{\mu\mu}+a_2 m_{\mu\mu}^2+a_3 m_{\mu\mu}^3}$	4
Epoly1	$\exp\left(a_1 m_{\mu\mu}\right)$	1
Epoly2	$\exp\left(a_1 m_{\mu\mu}+a_2 m_{\mu\mu}^2\right)$	2
Epoly3	$\exp\left(a_1 m_{\mu\mu}+a_2 m_{\mu\mu}^2+a_3 m_{\mu\mu}^3\right)$	3
Epoly4	$\exp\left(a_1 m_{\mu\mu}+a_2 m_{\mu\mu}^2+a_3 m_{\mu\mu}^3+a_4 m_{\mu\mu}^4\right)$	4

- The spurious signal SS after the subtraction (see below) needs to be smaller than 20% of the expected statistical uncertainty δS of the signal yield in 120 GeV $< m_{\mu\mu} <$ 130 GeV. The spurious signal SS is estimated as the signal yield in 120 GeV $< m_{\mu\mu} <$ 130 GeV obtained by fitting the signal+background functions to the background-only MC sample. In the estimation, the mean of the signal function m_H is scanned from 120 GeV to 130 GeV with a step size of 1 GeV. The SS is derived for each mean as shown in Figure 4.54. The maximum value of SS in that signal mass range is conservatively used. When applying the requirement, the MC statistical error is subtracted from the absolute value of the SS. The SS after the subtraction is denoted by $SS_{\pm 1\sigma}$ in Figure 4.54. The reweighted fast-simulation Drell-Yan Sherpa MC is used as the MC sample for the ggF and VBF categories. The reweighted full-simulation sample including top and diboson processes is used as the MC sample for the VH and ttH categories.

- If there are multiple functions satisfying the above requirements, first the function with the smallest number of free parameters is selected, then the function with the smallest maximum value of SS. This priority decreases statistical uncertainty and maximizes the sensitivity.

The maximum value of SS, denoted as max(SS), is considered as the systematic uncertainty due to the mismodeling in each category. Since all SS values are compatible with 0 within 2σ of MC statistical uncertainty at signal mass 125 GeV, no statistically significant mismodelling is found. The background-only fits to the fast-simulation Drell-Yan Sherpa MC in the ggF and VBF categories are shown in Figure 4.55. The selected empirical function, the χ^2 probability, the maximum value of $SS/\delta S$ and the maximum value of SS in each category are summarized in Table 4.16.

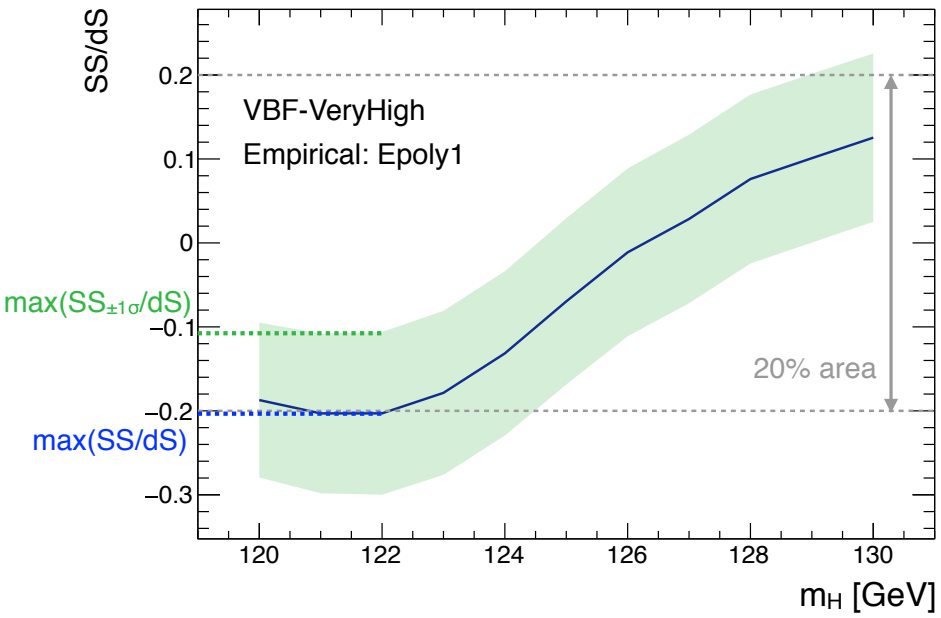


Figure 4.54 The spurious signal relative to its uncertainty $SS/\delta S$ as a function of the dimuon invariant mass. The figure is shown for the VBF-VeryHigh category. The value of $|\max(SS_{\pm 1\sigma}/\delta S)|$ is required to be smaller than 20%.

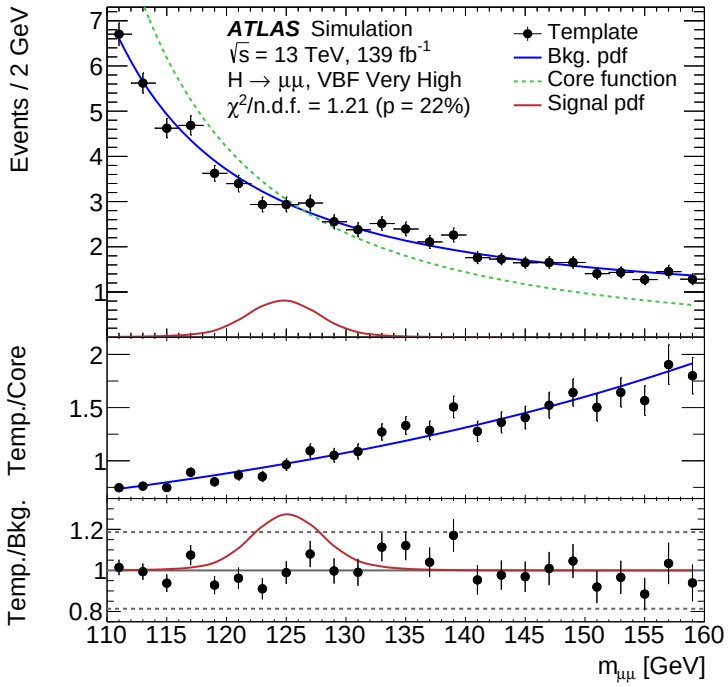


Figure 4.55 Illustration of the background modeling procedure using the fast-simulation Drell-Yan Sherpa MC. [86] The figure is shown for the VBF-VeryHigh category. The top panel shows the dimuon invariant mass distribution (black dots) overlaid with background model function(blue curve), core function (green dashed curve), and signal model function (red curve). The middle panel shows the MC template and the background model, both divided by the core component of the background model (black dots), thus showing the empirical part of the background model. The bottom panel shows the MC template divided by the background model (black dots) with the signal model (red curve).

Table 4.16 The selected empirical function, the χ^2 probabilities, the maximum value of SS relative to its uncertainty δS, and the maximum value of SS are also shown.

Category	Empirical	$P(\chi^2)_{\text{data}}$	$P(\chi^2)_{\text{fullMC}}$	$P(\chi^2)_{\text{fastMC}}$	$\max(SS/\delta S)$ [%]	$\max(SS_{\pm 1\sigma}/\delta S)$ [%]	$\max(SS)$
VBF-VeryHigh	Epoly1	0.56	0.40	0.22	−20.3	−10.8	−1.25
VBF-High	Power0	0.07	0.89	0.51	11.7	0.8	0.82
VBF-Medium	Power0	0.87	0.29	0.76	8.5	0.0	0.92
VBF-Low	Power0	0.34	0.22	0.96	11.2	3.0	2.14
Higgs-2Jet-VeryHigh	Power1	0.17	0.06	0.19	−13.3	4.8	−7.24
Higgs-2Jet-High	Epoly2	0.94	0.63	0.12	−19.8	−8.1	−23.80
Higgs-2Jet-Medium	Power1	0.83	0.89	0.48	19.8	10.0	37.30
Higgs-2Jet-Low	Epoly3	0.08	0.17	0.50	2.1	0.0	5.91
Higgs-1Jet-VeryHigh	Epoly2	0.57	0.02	0.38	21.9	14.3	−10.00
Higgs-1Jet-High	Epoly2	0.53	0.69	0.82	−7.8	0.0	−9.70
Higgs-1Jet-Medium	Power1	0.07	0.72	0.85	4.2	0.0	7.94
Higgs-1Jet-Low	Power1	0.13	0.59	0.12	17.3	9.0	72.60
Higgs-0Jet-VeryHigh	Power1	0.12	0.08	0.27	19.2	11.0	33.30
Higgs-0Jet-High	Power1	0.34	0.36	0.73	−19.4	−10.2	47.80
Higgs-0Jet-Medium	Power1	0.19	0.62	0.02	25.8	18.6	93.00
Higgs-0Jet-Low	Epoly3	0.16	0.93	0.03	−20.8	−16.8	−89.50
VH-4L-High	Power1	0.23	0.11	-	20.7	9.3	1.38
VH-3L-High	Epoly2	0.03	0.45	-	36.9	16.5	3.58
VH-3L-Medium	Epoly3	0.61	0.90	-	33.6	13.8	9.15
ttH	Power0	0.74	0.19	-	32.2	14.2	1.87

As a cross-check, the spurious signal is also evaluated using the fast-simulation Drell-Yan Powheg MC and the full-simulation MC including Drell-Yan process for the ggF and VBF categories. In addition, further cross-checks are performed using the fast-simulation Drell-Yan Sherpa MC applying several systematics variations. In all of these checks, no significant increase in the SS is found.

4.6.3 Systematic uncertainties

The following systematic uncertainties on the H $\to \mu\mu$ signal strength are considered:

- theoretical uncertainties on the expected signal yield,
- experimental uncertainties on the expected signal yield, the dimuon invariant mass scale and the dimuon invariant mass resolution, and
- spurious signal uncertainty due to the mismodeling of background.

The spurious signal uncertainty is described in the previous subsection. The theoretical uncertainty for the signal process and the experimental uncertainties are summarized in this section.

Theoretical uncertainty

The uncertainties of higher order QCD correction, PDF, underlying event and parton shower affect the Higgs boson production cross section, the H $\to \mu\mu$ branching ratio, and the signal yield in each category. The following uncertainties are considered in this analysis.

- **Uncertainties on Higgs boson cross section and** H $\to \mu\mu$ **branching ratio**:
  Uncertainties due to the QCD renormalization μ_{R} and factorization μ_{F} scale variation, choice of PDF set, and strong coupling constant α_s are considered as described in Ref. [10]. The μ_{R} and μ_{F} variations and the choice of PDF set depend on the production. For the main production of ggF, μ_{R} and μ_{F} are varied in the range from $m_H/4$ to m_H and the PDF set of PDF4LHC15_nnlo_100 [10] is used. The impacts are summarized in Table 4.17.
- **Uncertainties on signal yield due to higher order QCD correction**:
  For the ggF signal yield, the QCD correction uncertainty is divided into the effects of the QCD scale, QCD resummation, jet-multiplicity migration, the Higgs boson p_{T} migration, VBF/VH migration, and treatment of the top-quark mass in the loop corrections. The impacts on the ggF signal yield in each category are estimated to be from 0% to 10%. The QCD correction uncertainty for the VBF and VH signal yield take the effects of the QCD scale into account. The effects on the VBF and VH signal yield in each category are less than 3% and 3–25%, respectively.
- **Uncertainties on signal yield due to PDF**:
  The impacts on the ggF, VBF and VH signal yields due to the choice of PDF set are estimated by comparing the result with the HessianPDF set. All of these impacts are a few %. The effects on the ggF, VBF and VH signal yields due to the α_s uncertainty are estimated by varying α_s by ± 0.001. The impacts on the ggF, VBF and VH signal yields are a few %, less than 1%, and around 1%, respectively.

- **Uncertainties due to underlying event and parton shower**:
  For the ggF and VBF signal yields in the ggF/VBF categories, the uncertainties due to the underlying event and parton shower are considered. The impacts are estimated by comparing the events generated with different parton shower program: Pythia 8 and Herwig 7. The impacts range from 0% to 40%.

- **Uncertainty due to the ggF signal yield in the ttH category**: For the ggF signal yield in the ttH category, we assigned additional 100% uncertainty to cover the possible mismodeling of heavy-flavor production associated with the Higgs boson.

Table 4.17 The impacts on the Higgs boson production cross sections and the H → μμ branching ratio due to the QCD scale, PDF, and strong coupling constant α_s.

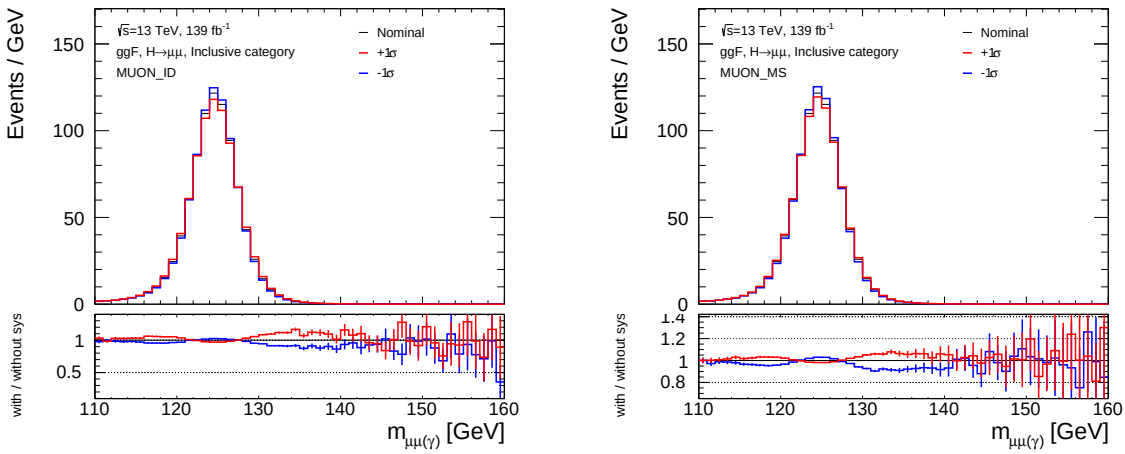
		Uncertainties [%]		
		QCD scale	PDF	α_s
Cross section	ggF	± 3.9	± 1.9	± 2.6
	VBF	$^{+0.4}_{-0.3}$	± 2.1	± 0.5
	WH	$^{+0.5}_{-0.7}$	± 1.7	± 0.9
	ZH	$^{+3.8}_{-3.1}$	± 1.3	± 0.9
	ttH	$^{+5.8}_{-9.2}$	± 3.0	± 2.0
BR(H → μμ)		± 1.23	$^{+0.97}_{-0.99}$	$^{+0.59}_{-0.64}$

Experimental uncertainty

The following experimental uncertainties are considered in this analysis.

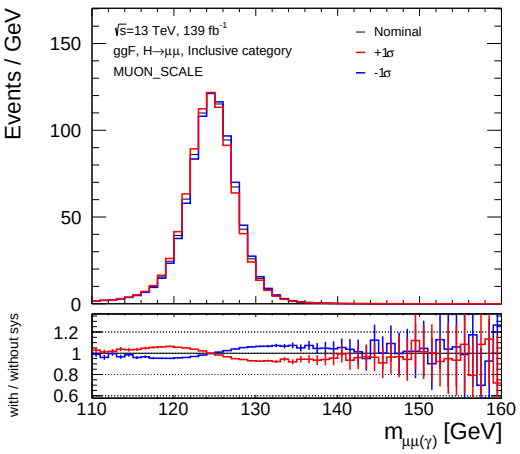
- **Muons**: The uncertainties due to the efficiencies of trigger, reconstruction, identification, and isolation affect the signal yield. In addition, the uncertainties due to the calibration of muon momentum scale and resolution (as described in Section 4.2.3) affect not only the signal yield but also the dimuon invariant mass scale and resolution. These effects on the invariant mass distribution are shown in Figure 4.56. The impacts on the signal yield, the $m_{\mu\mu}$ resolution, and the $m_{\mu\mu}$ scale in each category are less than 1%, 0–4%, and $\sim 0.1\%$, respectively.

- **Jets**: The uncertainties due to the calibration of jet energy scale and resolution and the efficiencies of JVT and b-tagging (as described in Sections 4.2.6 and 4.2.7) are considered.



(a)

(b)



(c)

Figure 4.56 The effects of uncertainties due to the calibration of muon momentum resolution in the ID (a), resolution in the MS (b), and scale (c) on the dimuon invariant mass.

These affect the signal yield especially in the jet-enriched categories since the jets and b-jets are used for categorization and event selection. The uncertainties on the signal yields due to jets and b-tagging are 0–8% and 0–12%, respectively.

- **Missing transverse mass**: The uncertainties due to the momentum scale and resolution of the soft-term tracks in the missing transverse mass determinations (as described in Section 4.2.8) are considered. These affect the signal yield since missing transverse mass is used as input of the VH and ttH classifiers. The impacts on the signal yield in each category are from 0% to ∼ 1%.

- **Electrons and photons**: The uncertainties due to the scale and the resolution of the energy of the electrons and the photons (as described in Section 4.2.4) are considered. These affect the signal yield, especially in the VH and ttH categories. The impacts on the signal yield are around 0.1%. The uncertainties dedicated for the FSR recovery are not considered since no significant difference is found between MC and data in the validation of FSR recovery.

- **Luminosity**: The uncertainty is derived from the luminosity measurement using the forward detectors [88]. This affect the cross section. The total uncertainty in 2015–2018 data is 1.7%.

- **Pile-up**: The difference in pile-up distribution between MC and data is corrected using a scale factor. The uncertainty of the scale factor is considered. This is below a few %.

4.6.4 Statistical analysis

For the signal extraction, the extended maximum likelihood fit is used in this analysis. The likelihood function depending on the signal strength μ and nuisance parameters $\boldsymbol{\theta}$ is defined as

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{k=1}^{N_{\text{cate}}} \prod_{j=1}^{N_{\text{bin}}} \frac{(\mu \cdot s_{k,j} + b_{k,j})^{n_{k,j}}}{n_{k,j}!} \mathrm{e}^{-(\mu \cdot s_{k,j} + b_{k,j})}, \tag{4.32}$$

where $N_{\text{cate}} = 20$ and $N_{\text{bin}} = 500$ are the number of categories and the number of bins of $m_{\mu\mu}$ distribution. $n_{k,j}$ is the number of events in the jth bin in the kth category. $s_{k,j}$ and $b_{k,j}$ are signal and background yields defined by

$$s_{k,j} = s_{\text{tot},k} \int_{\text{bin } j} f_{s,k}(m_{\mu\mu}; \boldsymbol{\theta_{s,k}})\, dm_{\mu\mu}, \tag{4.33}$$

$$b_{k,j} = b_{\text{tot},k} \int_{\text{bin } j} f_{b,k}(m_{\mu\mu}; \boldsymbol{\theta_{b,k}})\, dm_{\mu\mu}. \tag{4.34}$$

Here $f_{s(b),k}$ and $\boldsymbol{\theta_{s(b),k}}$ are PDF and nuisance parameters for signal (background) in the kth category. $s_{\text{tot},k}$ is the sum of signal yield for each process in the kth category. $s_{\text{tot},k}$ is fixed to the value expected by the SM. $b_{\text{tot},k}$ is total background yield in the kth category and one of the nuisance parameters. The integral $\int_{\text{bin } j}$ represents the integral inside the bin j.

The profile likelihood ratio is defined using the likelihood function:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})}. \tag{4.35}$$

The numerator is the conditional maximum-likelihood (ML) function of θ, which is a function of μ. The $\hat{\hat{\boldsymbol{\theta}}}(\mu)$ denotes the value of θ that maximizes $\mathcal{L}$ in the specified μ. The denominator is the unconditional ML function, where $\hat{\mu}$ and $\hat{\theta}$ maximize their $\mathcal{L}$. From the $\lambda(\mu)$ definition, the possible values are $0 \leq \lambda(\mu) \leq 1$, with λ near 1 implying good agreement between the data and the hypothesized μ.

The convenient statistic t_μ defined as below as the basis of a statistical test, referred to as the test statistics, is often used to test a hypothesized value of the signal strength μ:

$$t_\mu = -2\ln\lambda(\mu) = \begin{cases} -2\ln\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\theta})} & (\text{if } \hat{\mu} \geq 0) \\ -2\ln\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & (\text{if } \hat{\mu} < 0) \end{cases} \tag{4.36}$$

Higher t_μ values correspond to larger discrepancy between the data and the hypothesis. To quantify the level of disagreement, the p-value is computed as

$$p_\mu = \int_{t_\mu,\text{obs}}^{\infty} f(t_\mu|\mu)\ dt_\mu, \tag{4.37}$$

where $t_{\mu,\text{obs}}$ is the observed test statistics from the data and $f(t_\mu|\mu)$ is the value of PDF of t_μ under the assumption of the signal strength μ. In the particle physics, the p-value is usually converted into an equivalent significance Z defined as

$$Z = \Phi^{-1}(1-p), \tag{4.38}$$

where Φ^{-1} is the inverse of the cumulative distribution for the standard Gaussian. The $f(t_\mu|\mu)$ distribution and the relation of the p-value to the observed $t_{\mu,\text{obs}}$ and to the significance Z are illustrated in Figure 4.57.

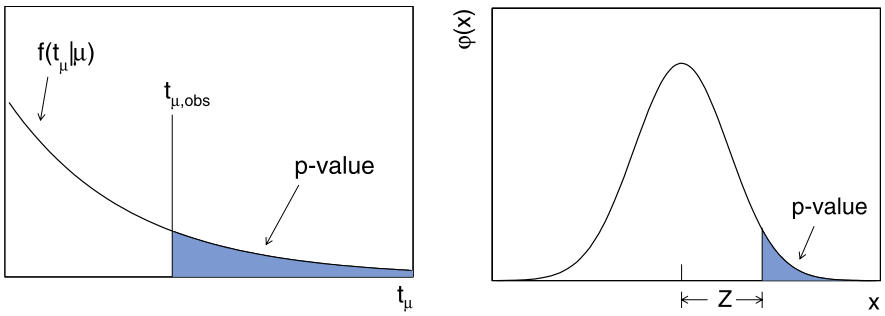
For discovery testing, the inconsistency with the background-only ($\mu = 0$) hypothesis is examined. In the test, the test statistic, q_0, and the p-value, p_0, defined as below are used:

$$q_0 = \begin{cases} -2\ln\frac{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}(0))}{\mathcal{L}(\hat{\mu},\hat{\theta})} & (\text{if } \hat{\mu} \geq 0) \\ 0 & (\text{if } \hat{\mu} < 0) \end{cases} \tag{4.39}$$

$$p_0 = \int_{q_0,\text{obs}}^{\infty} f(q_0|0)\ dq_0. \tag{4.40}$$

We regard as the "evidence" when the p_0 is less than 0.0013, corresponding to $Z_0 = 3$. If p_0 less than 2.87×10^{-7}, corresponding to $Z_0 = 5$, we regard as the "discovery".

If the result is consistent with background-only hypothesis, an upper limit on the signal strength μ is established by upper limit testing. In the test, the following test statistic q_μ and



(a) (b)

Figure 4.57 Illustration of the relation of the p-value to the observed $t_{\mu,\text{obs}}$ (a) and to the significance Z (b). [89]

p-value p_μ are used:

$$q_\mu = \begin{cases} -2\ln\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & (\text{if } \hat{\mu} \leq 0) \\ -2\ln\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\theta})} & (\text{if } 0 \leq \hat{\mu} \leq \mu) \\ 0 & (\text{if } \hat{\mu} > \mu) \end{cases} \tag{4.41}$$

$$p_\mu = \int_{q_\mu,\text{obs}}^{\infty} f(q_\mu|\mu)\, dq_\mu. \tag{4.42}$$

For establishing the limit, we look for a value of μ where the model can be excluded with 95% confidence level. The confidence level is defined based on the ratio of p-values for the background-only hypothesis and signal plus background hypothesis:

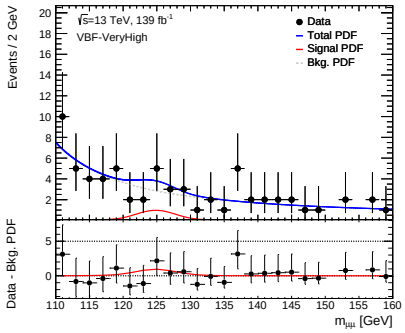
$$\text{CL}_s = \frac{p_\mu}{p_0}. \tag{4.43}$$

4.7 Results

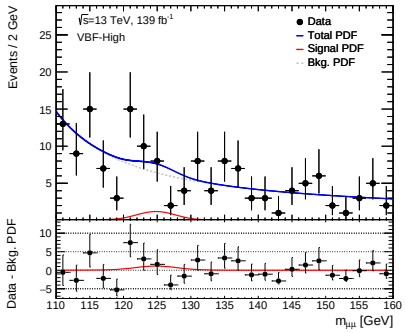
The maximum likelihood signal+background fit is performed to data in 110 GeV $< m_{\mu\mu} <$ 160 GeV as shown in Figure 4.58–4.61. The signal and background yields obtained by the fit are summarized in Table 4.18.

Table 4.18 Number of events observed in the 120 GeV $< m_{\mu\mu} <$ 130 GeV window in data, the number of signal (S) and background (B) as derived from the fit to the data, the observed $S/\sqrt{B}$, and the observed signal-to-background ratio (S/B).

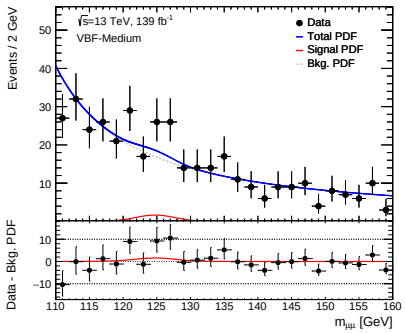
Category	Data	S	B	$S/\sqrt{B}$	S/B [%]
VBF-VeryHigh	15	3.29	14.4	0.86	22.8
VBF-High	39	4.04	32.4	0.71	12.5
VBF-Medium	112	5.60	84.9	0.61	6.6
VBF-Low	284	8.77	272.6	0.53	3.2
Higgs-2Jet-VeryHigh	1030	20.18	1024.2	0.63	2.0
Higgs-2Jet-High	5433	56.74	5434.7	0.77	1.0
Higgs-2Jet-Medium	18311	89.25	18310.4	0.66	0.5
Higgs-2Jet-Low	36409	70.99	36342.7	0.37	0.2
Higgs-1Jet-VeryHigh	1097	19.31	1071.2	0.59	1.8
Higgs-1Jet-High	6413	54.27	6323.5	0.68	0.9
Higgs-1Jet-Medium	24576	104.22	24286.3	0.67	0.4
Higgs-1Jet-Low	73459	142.79	73473.8	0.53	0.2
Higgs-0Jet-VeryHigh	15986	68.50	16079.4	0.54	0.4
Higgs-0Jet-High	46523	115.54	46200.9	0.54	0.3
Higgs-0Jet-Medium	91392	137.72	91290.9	0.46	0.2
Higgs-0Jet-Low	121354	91.07	121314.0	0.26	0.1
VH4Lep-High	34	0.61	24.4	0.12	2.5
VH3Lep-High	41	1.70	40.7	0.27	4.2
VH3Lep-Medium	358	3.22	348.1	0.17	0.9
ttH	17	1.40	15.1	0.36	9.2



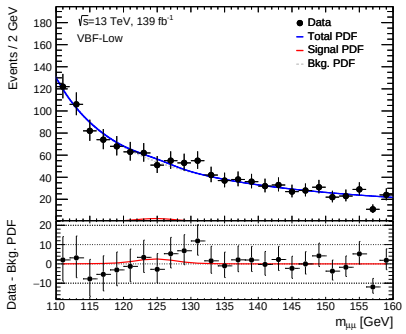
(a)



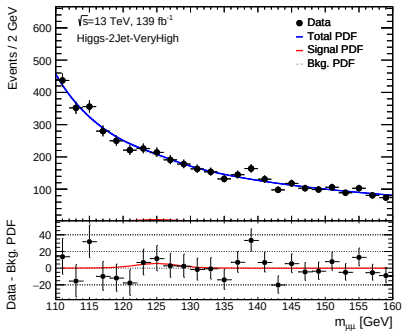
(b)



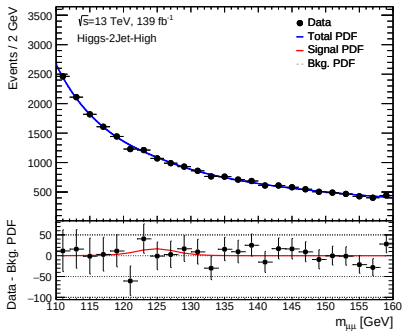
(c)



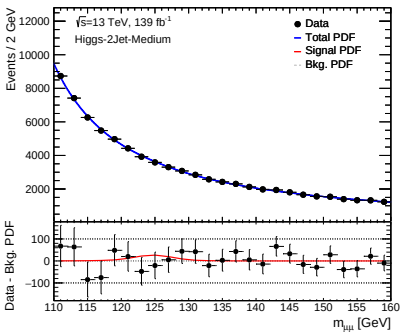
(d)



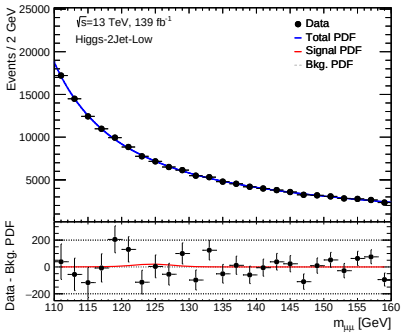
(e)



(f)

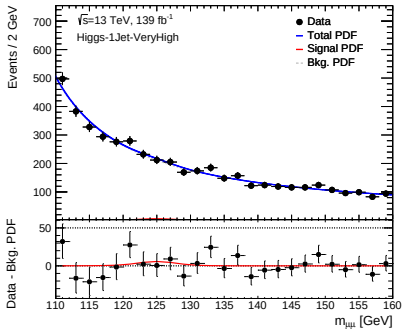


(g)

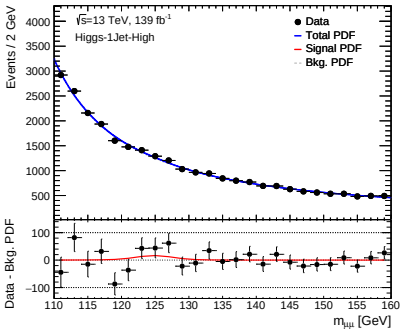


(h)

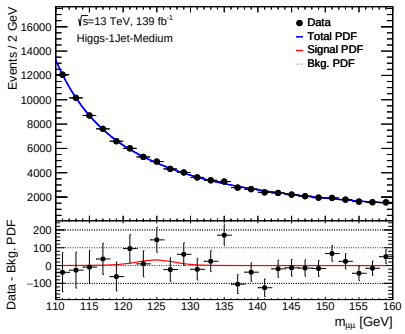
Figure 4.58 Invariant mass distribution observed in data for each category in VBF and Higgs-2Jet categories.



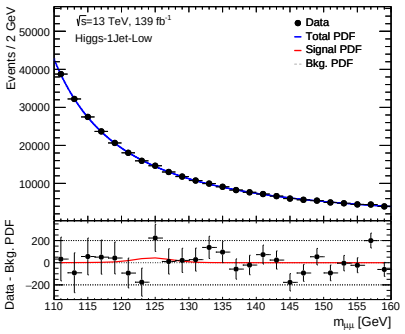
(a)



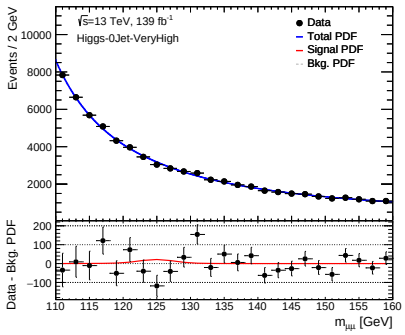
(b)



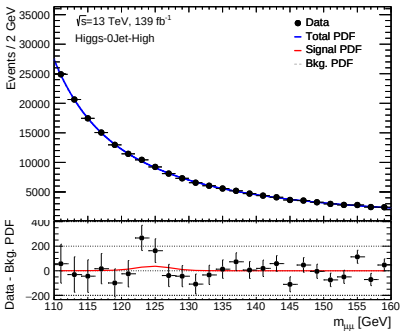
(c)



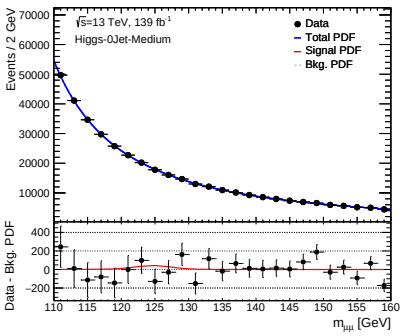
(d)



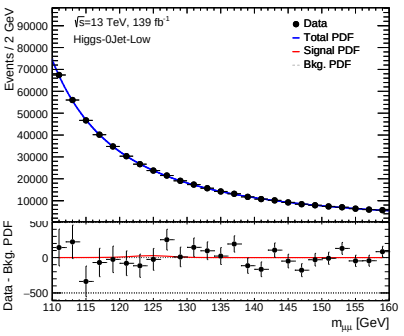
(e)



(f)



(g)



(h)

Figure 4.59 Invariant mass distribution observed in data for each category in Higgs-1Jet and Higgs-0Jet categories.

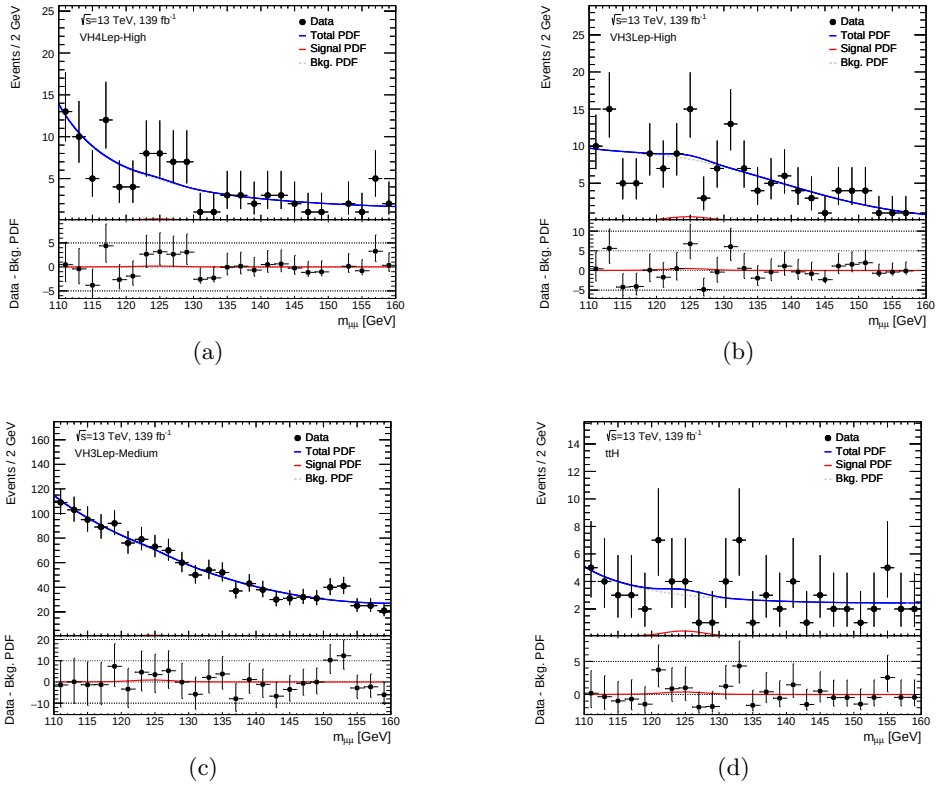


Figure 4.60 Invariant mass distribution observed in data for each category in VH and ttH categories.

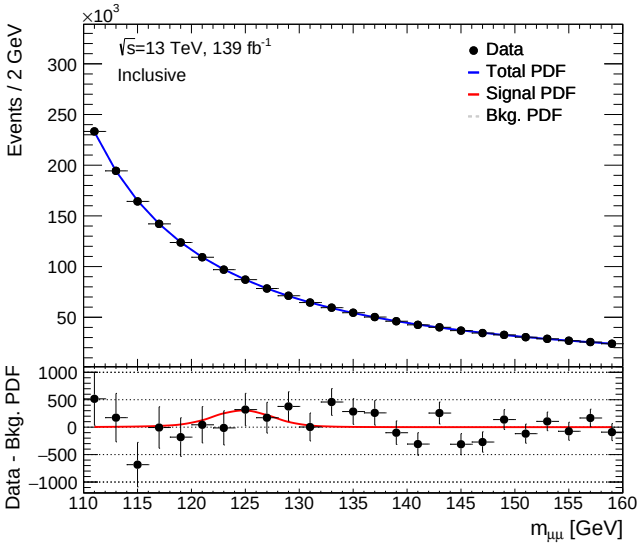


Figure 4.61 Invariant mass distribution observed in data for the inclusive of all category.

The likelihood as a function of μ is shown in Figure 4.62. No significant excess is observed in data. The observed (expected) significance is 2.02σ (1.73σ) relative to the background-only hypothesis for the SM Higgs boson with a mass of 125.09 GeV. The signal strength μ is found to be $\mu = 1.17^{+0.58}_{-0.57}(\text{stat})\ ^{+0.18}_{-0.13}(\text{syst})$. This shows the good compatibility with the SM value.

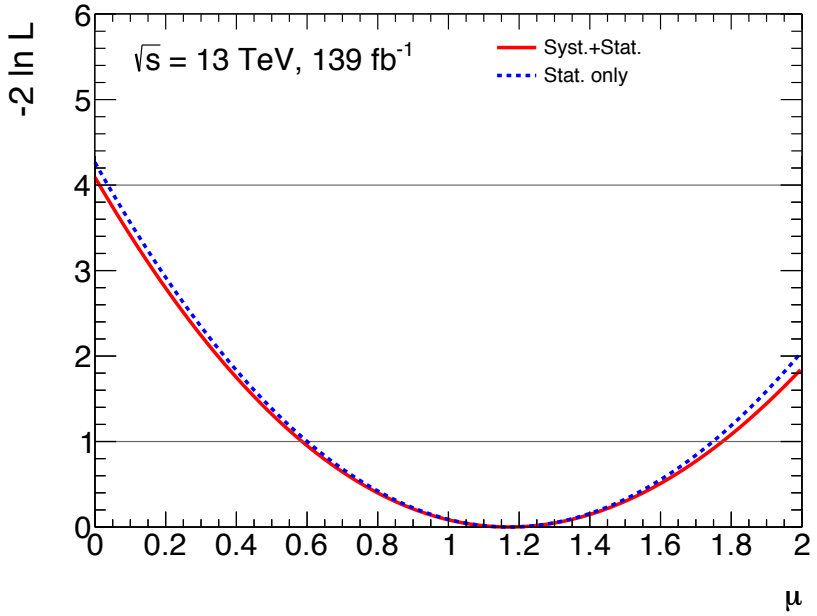


Figure 4.62 The profile of the negative log-likelihood ratio $-2\ln L$ as a function of the signal strength μ.

The independent signal strength μ in each category is defined and a simultaneous fit to all categories is performed. The signal strength μ and the significance for each category are summarized in Table 4.19.

Total uncertainty is dominated by statistical uncertainty of data. The ranking plot of the impacts for the top 30 systematic uncertainties is shown in Figure 4.63. The contributions of the systematic uncertainties are $^{+0.13}_{-0.08}$ from the theoretical uncertainties, ± 0.10 from the spurious signal, and $^{+0.07}_{-0.03}$ from the experimental uncertainties.

The observed upper limit on the signal strength at 95% confidence level is obtained to be 2.26. The expected upper limit is 1.08 for background only scenario and 1.97 for the SM H → μμ scenario. The corresponding value of the observed upper limit on the branching ratio is 4.7×10^{-4} assuming the SM cross section. These results show a 2.5 times better expected sensitivity compared to the results of the previous ATLAS publication [11]. The improvement of a factor of about two is from a larger dataset and the additional 25% improvement is from advanced analysis techniques: event selection, FSR recovery, classification, and background modeling. This result provides further constraints to the coupling of the Higgs boson to the second-generation fermions.

Table 4.19 The expected and observed signal strength μ and the significance.

Category	Expected			Observed		
	Fitted μ	$\delta\mu$	Significance	Fitted μ	$\delta\mu$	Significance
VBF-VeryHigh	1.00	1.67	0.62	0.33	1.29	0.26
VBF-High	1.00	1.86	0.55	0.29	1.65	0.18
VBF-Medium	1.00	2.07	0.49	5.97	2.14	3.11
VBF-Low	1.00	2.41	0.41	0.74	2.10	0.35
Higgs-2Jet-VeryHigh	1.00	2.10	0.41	0.65	2.09	0.31
Higgs-2Jet-High	1.00	1.89	0.52	-0.03	1.64	0.00
Higgs-2Jet-Medium	1.00	1.92	0.45	-1.56	1.92	0.00
Higgs-2Jet-Low	1.00	4.38	0.22	-3.36	3.98	0.00
Higgs-1Jet-VeryHigh	1.00	2.46	0.40	0.74	2.13	0.35
Higgs-1Jet-High	1.00	2.07	0.48	3.21	1.85	1.79
Higgs-1Jet-Medium	1.00	1.86	0.47	3.25	1.87	1.77
Higgs-1Jet-Low	1.00	2.40	0.36	-0.49	2.40	0.00
Higgs-0Jet-VeryHigh	1.00	2.33	0.38	-4.94	2.31	0.00
Higgs-0Jet-High	1.00	2.27	0.38	5.45	2.37	2.37
Higgs-0Jet-Medium	1.00	2.84	0.31	-2.43	2.74	0.00
Higgs-0Jet-Low	1.00	6.28	0.16	1.97	5.45	0.36
VH4Lep-High	1.00	8.90	0.10	28.37	12.20	2.70
VH3Lep-High	1.00	6.41	0.15	2.64	5.17	0.51
VH3Lep-Medium	1.00	10.60	0.09	9.41	9.46	1.00
ttH	1.00	4.06	0.25	0.80	3.60	0.23

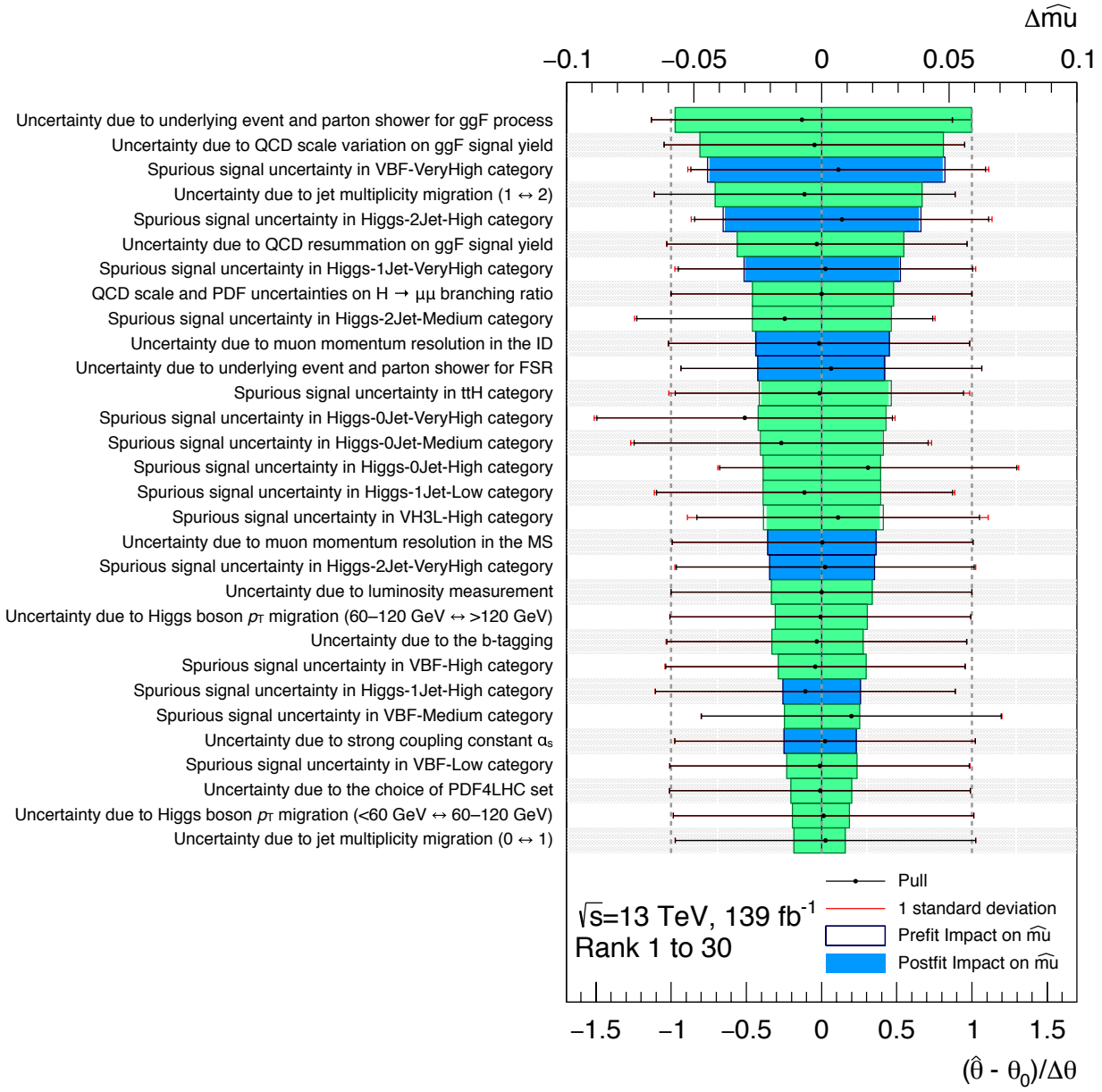


Figure 4.63 The ranking of the impacts for the top 30 systematic uncertainties on the prefit and postfit pulls. The black points show the pulls. The red points show the 1σ region. The white and blue bands are the prefit and postfit uncertainties, respectively. They show the expected and observed impact on μ.

Chapter 5

Discussion

In this chapter, the result obtained by this analysis are discussed. The comparison with the result of the CMS experiment is presented in Section 5.1. Future prospects of the H $\rightarrow \mu\mu$ search at the LHC are presented in Section 5.2. For the High-Luminosity LHC (HL-LHC) which is a future project at the LHC (see Section 5.2), the upgrade of the trigger and readout system is indispensable. For the first level end-cap muon trigger, an improved trigger algorithm that reduces the trigger rate will be introduced to take full advantage of HL-LHC. To suppress the total trigger rate, it is important to reduce the rate in the end-cap region since the rate in the end-cap region has been dominant. In this upgrade, I developed a sophisticated trigger algorithm based on the TGC track reconstruction. The study of the muon trigger upgrade is presented in Section 5.3.

5.1 Comparison with CMS results

The CMS experiment published the following results in Reference [90].

- Data set: $\sqrt{s} = 13$ TeV pp collision data corresponding to 137 fb^{-1}
- Significance: 3.0σ (2.5σ) observed (expected)
- Signal strength μ: $1.19^{+0.41}_{-0.39}$(stat) $^{+0.17}_{-0.16}$(syst)

These results are compatible with the results shown in this thesis. Both results indicate that the mass origin of muons is from the Higgs mechanism. In addition, the observed significance of 3.0σ in CMS represents the evidence of the H $\rightarrow \mu\mu$ decay. The larger significance at CMS is predominantly due to higher dimuon invariant mass resolution. The dimuon invariant mass resolution for the ggF signal is 1.47–2.12 GeV in the CMS analysis, while that in ATLAS is 2.55–3.12 GeV. The difference originates from the better muon momentum resolution of CMS due to the stronger magnetic field of 3.8 T of the solenoid magnet (2 T for ATLAS).

The observed (expected) significance is 3.6σ (3.0σ) by statistically combining these results without considering the correlation in the systematic uncertainties. This result more strongly suggests the relationship between the mass of the second-generation fermions and the Higgs mechanism.

5.2 Future prospect

The future schedule of the LHC is shown in Figure 5.1. After the data-taking period in 2015–2018, referred to as "Run 2", the injectors will be upgraded in the Phase 1 upgrade period. Another data-taking period "Run 3" is planned to deliver the data up to the total integrated luminosity of 350 fb^{-1} [91]. In order to access broader physics programs which include a detailed exploration of the Higgs mechanism, new physics searches through more precise study of rare SM processes, and searches for new heavy states, the High-Luminosity LHC (HL-LHC) project is planned after Run 3. During the long shutdown 3 (LS3), new elements including focusing magnets and possibly crab cavities are installed in the LHC interaction regions to provide the peak instantaneous luminosity of 5–7.5$\times 10^{34}$ cm^{-2}s^{-1}, corresponding to an average pile-up of 200. In order to cope with the severe radiation environment and higher trigger rate at the HL-LHC, some detectors and the trigger and data acquisition system will also be upgraded during the LS3. This large scale upgrade is called the Phase 2 upgrade. The data-taking of the HL-LHC is planned to accumulate a total integrated luminosity of 3000–4000 fb^{-1} in about 10 years.

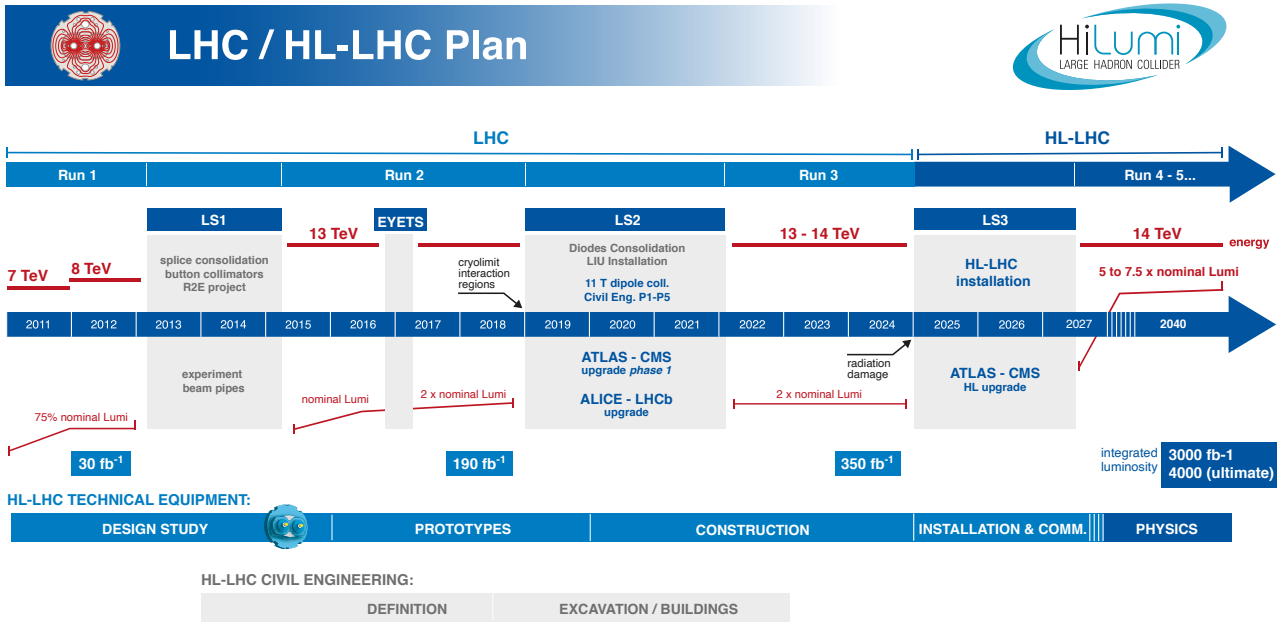


Figure 5.1 The LHC / HL-LHC Plan. After the data-taking period "Run 2", the Long shutdown 2 (LS2), another data-taking period "Run 3", and the HL-LHC are planned.

The expected significance of the H $\to \mu\mu$ search will be improved by using such a large dataset, as shown in Table 5.1. It is expected to exceed 7σ in the HL-LHC. The expected

significances for different data statistics are estimated by scaling the statistical uncertainty by the integrated luminosity. This estimation was performed assuming the SM signal strength ($\mu = 1$) and ignoring any changes in the effects of the systematic uncertainties since the statistical uncertainty is dominant in the search.

Table 5.1 The integrated luminosity delivered by the LHC and the expected significance of the $H \to \mu\mu$ search. The integrated luminosity of the Run 3 is the value assuming that the LHC delivers a further 164 fb^{-1} by the end of Run 3.

Periods	Run 2	Run 3	HL-LHC		
Integrated luminosity [fb^{-1}]	156	320	470	1300	3000 (4000)
Expected significance [σ]	1.7	2.4	3.0 (evidence)	5.0 (discovery)	7.5 (8.7)

In addition to the increased statistics, the resolution of the dimuon invariant mass and the muon trigger efficiency will be improved by the Phase 2 upgrade. In the Phase 2 upgrade, the inner detectors are replaced with a new inner tracker made of silicon sensors with higher granularity [92]. As a result, the resolution of the dimuon invariant mass is expected to be improved by 25% as shown in Figure 5.2. The expected significance is estimated to be improved

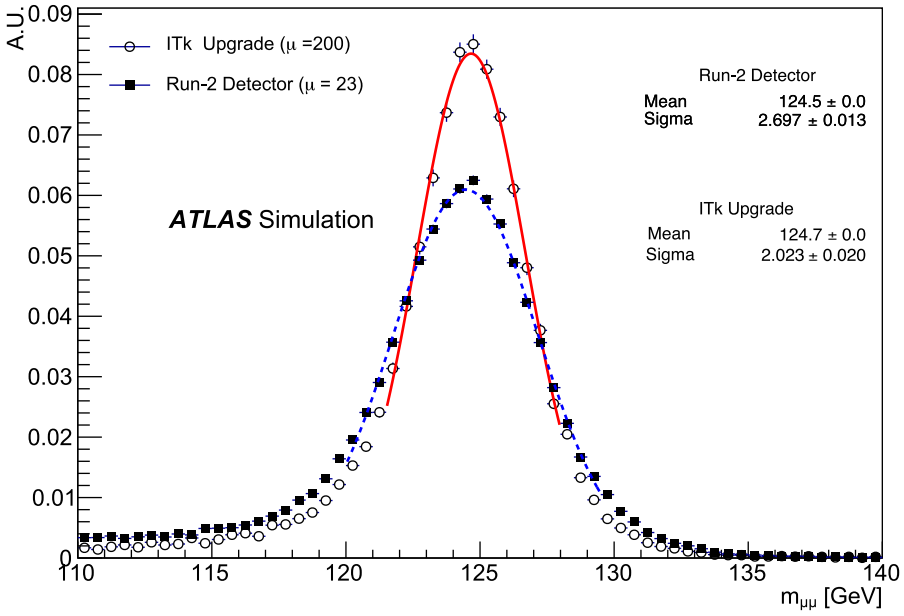


Figure 5.2 The expected dimuon invariant mass distribution in the HL-LHC condition corresponding to the average pile-up of 200, compared with the Run 2 condition.

by 15% by this resolution improvement.

Furthermore, the RPCs will be added in the inner Barrel region [93]. In the Run 2, the Level-1 single muon trigger efficiency for the muons with $p_{\mathrm{T}} > 20$ GeV was only 70% in the barrel (see Figure 4.6(b)) due to the RPC layout constrained by the mechanical structure of the barrel toroid magnet. In the HL-LHC, the acceptance will be increased thanks to the installation of the additional RPCs, and the trigger efficiency is expected to be improved to $\sim 96\%$ as shown in Figure 5.3. The expected significance is estimated to be improved by about 5% by this improvement.

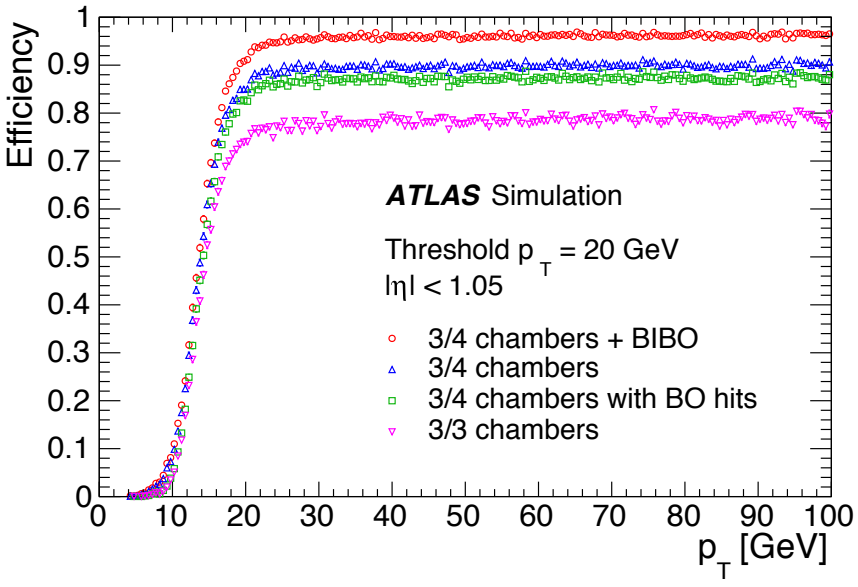


Figure 5.3 The efficiency of the muon trigger targeting the muons with $p_{\mathrm{T}} > 20$ GeV in the Barrel region ($|\eta| < 1.05$).

Taking into account both improvements, the integrated luminosity required to reach the significance of 3σ (5σ) is estimated to be about 320 (890) fb^{-1}.

5.3 TGC tracking trigger

In the Run 2, the first level end-cap muon trigger was based on the coincidence of hits among the TGC layers (described in Section 3.2.5). A coincidence is taken by requiring 2 out of 3 layers of the wires of the most inner station. Another independent coincidence taken by requiring 3 out of 4 layers from the remaining stations. Combining the information obtained, the trigger makes a decision. The first level muon trigger rate of the p_{T} threshold 20 GeV was up to 20 kHz.

In the HL-LHC, the event rate increases due to the higher luminosity. To cope with the higher event rate, the threshold of the first level muon trigger needs to raise to 50 GeV if the trigger and readout systems are not upgraded [93]. In the no upgrade scenario, the integrated acceptance for $W \to \ell\nu$ decay decreases by about 40% as shown in Figure 5.4. The acceptance

for the H $\rightarrow \mu\mu$ decay is also expected to decrease by about 60%. In order to cope with the higher event rate keeping the threshold, the maximum allowable value of the first level trigger rate will be increased from 100 kHz to 1 MHz. To extend the first level trigger rate, the trigger and readout electronics at Run 3 needs to be replaced by new one. Furthermore, for the end-cap muon trigger, an improved trigger algorithm will be introduced to take full advantage of HL-LHC. In this upgrade, I developed a sophisticated trigger algorithm based on the muon track reconstruction using TGC hits. Before the HL-LHC, a new muon detector that can reconstruct muon segments with around 1 mrad resolution, referred to as New Small Wheel (NSW), will be introduced in the inner part of the magnetic field [94]. For the HL-LHC, in addition to NSW, TGC provides the muon segment in the outer region of the magnetic field using the hit information of all TGC layers. The trigger decision is made with the deflection angle between the NSW segment and the TGC segment (referred to as NSW+TGC trigger). Thanks to an improved p_{T} resolution for muons with the deflection angle, the new trigger suppresses the events including muons with p_{T} smaller than the p_{T} threshold.

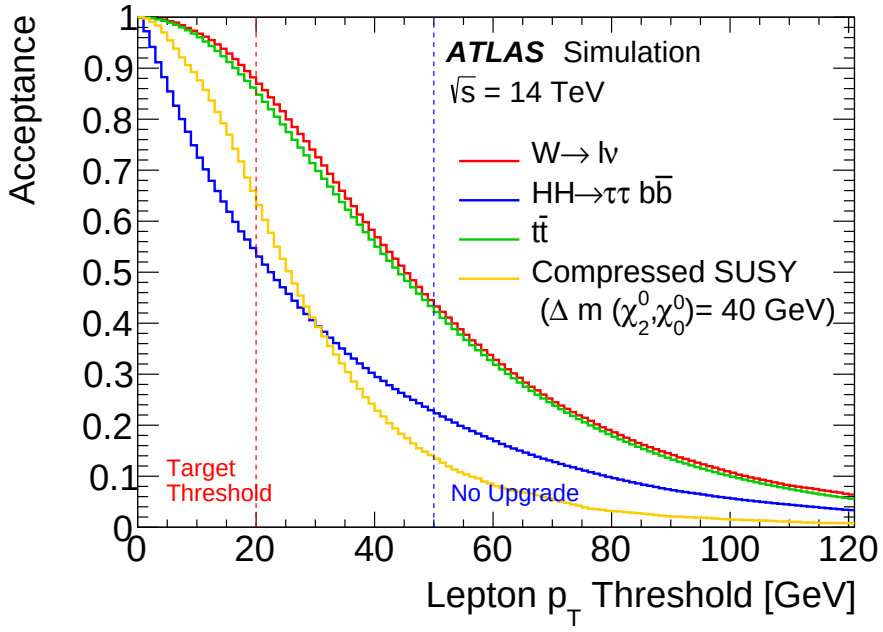


Figure 5.4 The integrated acceptance as a function of the single lepton p_{T} threshold for $W \rightarrow \ell\nu$ decay (red line). The target p_{T} threshold which is the same value as that of Run 2 primary is shown by the red dotted line. The threshold for the no upgrade scenario is shown by the blue dotted line. [93]

In this study, I developed the reconstruction algorithm of muon segments using TGC hits and evaluated the performance of the NSW+TGC trigger. Since the TGC segments should be reconstructed within around 1 μs from the bunch crossing, pattern matching method is used for the segment reconstruction. In the pattern matching method, the patterns of TGC hit channels for each layer are converted to the segment information, including the position, the

angle, and the quality of the hit pattern. The pattern list is prepared by scanning the muon straight trajectory. Each predefined hit pattern has associated segment information. By using this algorithm, TGC segments are reconstructed with around 4 mrad resolution as shown in Figure 5.5. The product of the acceptance and segment reconstruction efficiency is greater than 90%. The trigger efficiency for the NSW+TGC trigger as a function of the muon p_{T} reconstructed by offline analysis is shown in Figure 5.6. The NSW+TGC trigger provides higher efficiency for muons with p_{T} greater than p_{T} threshold and better rejection for muons with p_{T} lower the p_{T} threshold. The high efficiency at the plateau arises from a loose requirement for the coincidence of the TGC layers in the pattern matching algorithm, which requires 5 hits in 7 layers. The better p_{T} resolution in the turn-on curve is from the improved algorithm using the deflection angle. Thanks to the improved algorithm, the first level end-cap muon trigger rate of the p_{T} threshold 20 GeV is expected to be smaller than 30 kHz at HL-LHC. The threshold of the first level muon trigger is kept to 20 GeV. Therefore, the acceptance for the $\mathrm{H} \to \mu\mu$ decay is also kept to the same level as that of Run 2.

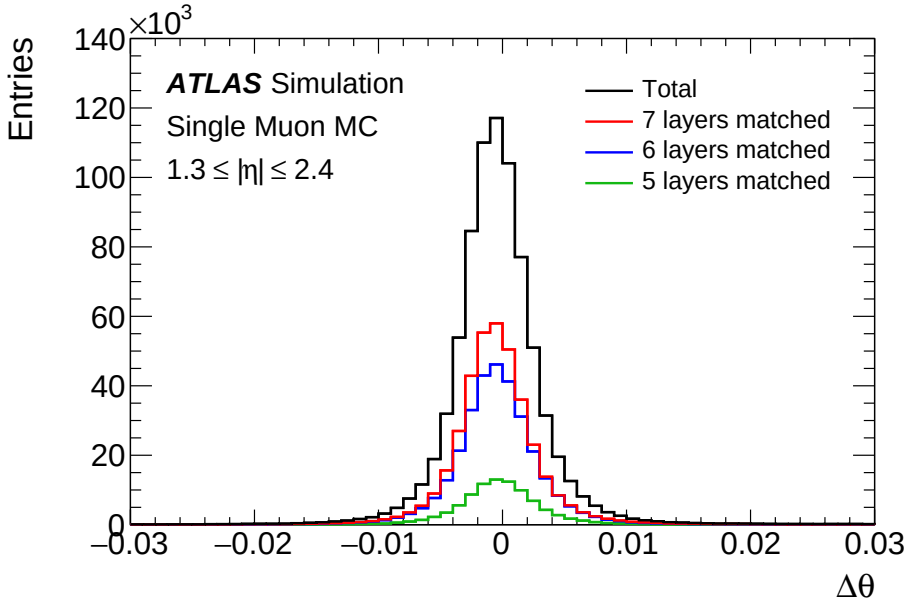


Figure 5.5 Distributions of the difference of the polar angle θ between the reconstructed TGC segment and the segment reconstructed by the ATLAS full offline analysis. The red, blue, and green histograms are for the TGC segments reconstructed with seven, six, and five hits, respectively, over the seven layers. The black histogram shows the sum of them.

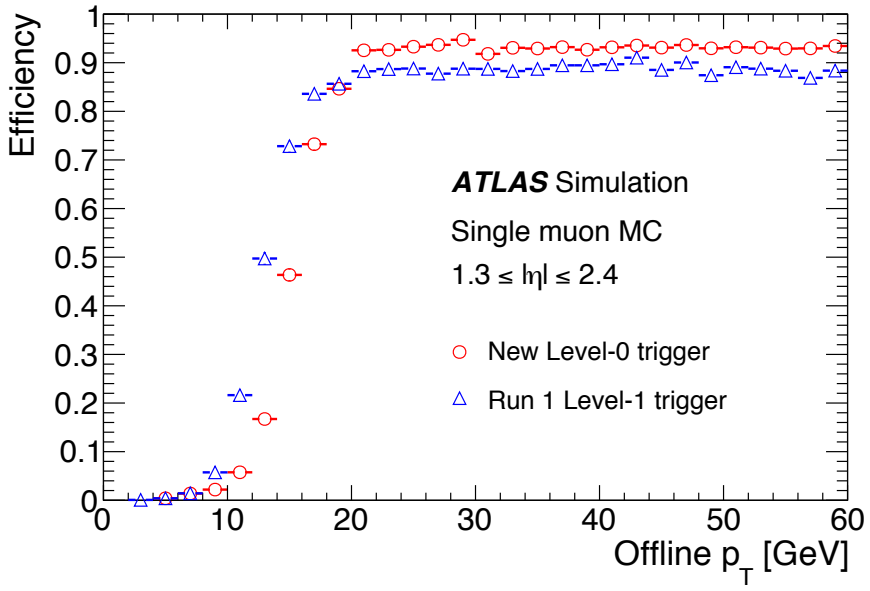


Figure 5.6 The muon trigger efficiency for p_{T} threshold 20 GeV of NSW+TGC trigger (red) and of the trigger for Run 1 which has similar performance to Run 2 (blue).

Chapter 6

Conclusion

The SM well explains various phenomena observed so far. However, it is unnatural that the difference in fermion mass, which differs by six orders depending on the generation, is explained only by the difference in the Yukawa coupling constant. The Yukawa couplings between the Higgs boson and the third-generation charged fermions (t, b, τ) have been observed and they are consistent with the SM. However, the Yukawa couplings of the first- and second-generation fermions have not been observed yet. In order to confirm whether there is any difference in the coupling between the generations of fermions, I performed the $H \to \mu\mu$ search using $\sqrt{s} =$ 13 TeV pp collisions data collected with the ATLAS detector at LHC in 2015–2018 corresponding to 139 fb^{-1}.

In this analysis, I developed a method to reconstruct the photon in the events in which a muon emits a photon by the QED FSR. By including the four-momentum of the photon in the invariant mass calculation, the mass resolution was improved by 2.8%. In addition, in order to separate the signal events from the background events, the events are classified into 20 categories in total using the multivariate analysis. In order to extract the signal yields without significant bias, the invariant mass distribution of the background events was modeled by the function based on the LO Drell-Yan analytic line-shape.

The observed (expected) signal significance is 2.0 (1.7) standard deviations with respect to the background-only hypothesis for the Higgs boson with a mass of 125.09 GeV. The signal strength, defined as the ratio of the observed signal yield to the expected one in the SM, is 1.2 ± 0.6. The CMS experiment also reported the result, yielding the observed signal significance of 3 standard deviations. It is compatible with the result described in this thesis. These results are consistent with the SM, representing that the coupling strength of the Higgs boson and the muon is much smaller than that of the third-generation fermions. In addition, these results indicate for the first time the mass origin of the second-generation fermions seems to arise from the unnaturalness of the Higgs mechanism. The results provide important ingredients in unraveling the mass origin of the fermions and the mystery of the generations of the fermions.

At present, the uncertainty of the signal strength is dominated by the statistical uncertainty. There are further data taking plans at (HL-)LHC. I expect that the Yukawa coupling to second-generation fermion is verified with significance greater than 5 standard deviations at HL-LHC. In addition, the FSR recovery method established in this study will be useful in the future

experiment under a higher pile-up environment since the optimized FSR photon selection is robust against the number of pile-ups.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor Makoto Tomoto for giving me the opportunity to research in a supportive environment where I could fully devote to the research. Thanks to his encouragement, I was able to obtain remarkable results and psychological strength as well. I would also like to express my thanks to Yasuyuki Horii and Yuya Kano for providing me with technical advice and useful discussions. In addition, I was supported in my research and life by a number of people including Toru Iijima, all the other N-lab members, the ATLAS H $\rightarrow \mu\mu$ members, the ATLAS Japan members, my friends, and my family. I sincerely appreciate all those people.

This work was supported by MEXT KAKENHI Grant Number JP16H06493 and JSPS KAKENHI Grant Number JP18J21699.

References

[1] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[2] CMS Collaboration, S. Chatrchyan et al., *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[3] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[4] CMS Collaboration, S. Chatrchyan, G. Hmayakyan, and V. Khachatryan, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08004–S08004, https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08004.

[5] T. Abe, R. Sato, and K. Yagyu, *Muon specific two-Higgs-doublet model*, JHEP **07** (2017) 012, arXiv:1705.01469 [hep-ph].

[6] G. Salam, *Elements of QCD for hadron colliders*, CERN Yellow Rep. School Proc. **5** (2020) 1–56.

[7] NNPDF Collaboration, R. D. U. Ball, *Parton distributions from high-precision collider data*, Eur. Phys. J. C **77** (2017) 663, arXiv:1706.00428 [hep-ph].

[8] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton fragmentation and string dynamics*, Physics Reports **97** (1983) 31 – 145, http://www.sciencedirect.com/science/article/pii/0370157383900807.

[9] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, *Event generation with SHERPA 1.1*, Journal of High Energy Physics **2009** (2009) 007–007, https://doi.org/10.1088%2F1126-6708%2F2009%2F02%2F007.

[10] LHC Higgs Cross Section Working Group Collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, arXiv:1610.07922 [hep-ph].

[11] ATLAS Collaboration, M. Aaboud et al., *Search for the dimuon decay of the Higgs boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. Lett. **119** (2017) 051802, arXiv:1705.04582 [hep-ex].

[12] ATLAS Collaboration, G. Aad et al., *Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*, Phys. Rev. D **101** (2020) 012002, arXiv:1909.02845 [hep-ex].

[13] ATLAS Collaboration, *A search for the rare decay of the Standard Model Higgs boson to dimuons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, ATLAS-CONF-2018-026, 2018, https://cds.cern.ch/record/2628763.

[14] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001–S08001, https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001.

[15] C. De Melis, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, http://cds.cern.ch/record/2119882, General Photo.

[16] J. Wenninger, *LHC status and performance*, PoS **CHARGED2018** (2019) 001.

[17] K. 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*, PoS **EPS-HEP2015** (2015) 261, arXiv:1608.07850 [physics.ins-det].

[18] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, cern-lhcc-2010-013; atlas-tdr-19, 2010. https://cds.cern.ch/record/1291633.

[19] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*, cern-lhcc-97-022; atlas-tdr-10, 1997. https://cds.cern.ch/record/331068.

[20] ATLAS Collaboration, M. Aaboud et al., *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv:1611.09661 [hep-ex].

[21] CMS Collaboration, *Measurement of the differential Drell-Yan cross section in proton-proton collisions at 13 TeV*, CMS-PAS-SMP-16-009, 2016, https://cds.cern.ch/record/2205152.

[22] ATLAS Collaboration, *Luminosity Public Results Run2*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[23] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015, arXiv:2004.13447 [hep-ex].

[24] ATLAS Collaboration, *ATLAS computing: Technical design report*, cern-lhcc-2005-022; atlas-tdr-017, 2005. https://cds.cern.ch/record/837738.

[25] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv:hep-ph/0409146.

[26] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[27] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv:1510.03865 [hep-ph].

[28] S. Catani and M. Grazzini, *An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC*, Phys. Rev. Lett. **98** (2007) 222002, arXiv:hep-ph/0703012.

[29] ATLAS Collaboration, G. Aad et al., *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09** (2014) 145, arXiv:1406.3660 [hep-ex].

[30] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **02** (2010) 037, arXiv:0911.5299 [hep-ph].

[31] G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, G. Ossola, T. Reiter, and F. Tramontano, *Automated One-Loop Calculations with GoSam*, Eur. Phys. J. C **72** (2012) 1889, arXiv:1111.2034 [hep-ph].

[32] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, *$HW^{\pm}/HZ$ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*, JHEP **10** (2013) 083, arXiv:1306.2542 [hep-ph].

[33] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[34] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph].

[35] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, https://cds.cern.ch/record/1966419.

[36] Sherpa Collaboration, E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv:1905.09127 [hep-ph].

[37] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv:0709.1027 [hep-ph].

[38] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv:1111.1220 [hep-ph].

[39] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, arXiv:1207.5030 [hep-ph].

[40] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv:0808.3674 [hep-ph].

[41] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv:1111.5206 [hep-ph].

[42] A. Denner, S. Dittmaier, and L. Hofer, *Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations*, Comput. Phys. Commun. **212** (2017) 220–238, arXiv:1604.06792 [hep-ph].

[43] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[44] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, 2016, https://cds.cern.ch/record/2206965.

[45] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[46] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, and S. Banerjee, *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) 250 – 303, http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[47] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. Yuan, *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93** (2016) 033006, arXiv:1506.07443 [hep-ph].

[48] L. Lonnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, JHEP **05** (2002) 046, arXiv:hep-ph/0112284.

[49] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[50] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:hep-ph/0206293.

[51] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, and W. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:hep-ph/0201195.

[52] P. Golonka and Z. Was, *Next to Leading Logarithms and the PHOTOS Monte Carlo*, Eur. Phys. J. C **50** (2007) 53–62, arXiv:hep-ph/0604232.

[53] ATLAS Collaboration, M. Aaboud et al., *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv:1704.07983 [hep-ex].

[54] A. Rosenfeld and J. L. Pfaltz, *Sequential Operations in Digital Picture Processing*, J. ACM **13** (1966) 471, https://doi.org/10.1145/321356.321357.

[55] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) 444–450, http://www.sciencedirect.com/science/article/pii/0168900287908874.

[56] ATLAS Collaboration, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, Journal of Instrumentation **9** (2014) P09009–P09009, http://dx.doi.org/10.1088/1748-0221/9/09/P09009.

[57] F. Meloni, *Primary vertex reconstruction with the ATLAS detector*, Journal of Instrumentation **11** (2016) C12060–C12060, https://doi.org/10.1088%2F1748-0221%2F11%2F12%2Fc12060.

[58] W. Waltenberger, R. Frühwirth, and P. Vanlaer, *Adaptive vertex fitting*, Journal of Physics G: Nuclear and Particle Physics **34** (2007) N343–N356, https://doi.org/10.1088%2F0954-3899%2F34%2F12%2Fn01.

[59] ATLAS Collaboration, G. Aad et al., *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv:1603.05598 [hep-ex].

[60] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2020-030, 2020, http://cds.cern.ch/record/2725736.

[61] J. Illingworth and J. Kittler, *A survey of the hough transform*, Computer Vision, Graphics, and Image Processing **44** (1988) 87 – 116, http://www.sciencedirect.com/science/article/pii/S0734189X88800331.

[62] ATLAS Collaboration, G. Aad et al., *Studies of the performance of the ATLAS detector using cosmic-ray muons*, Eur. Phys. J. C **71** (2011) 1593, arXiv:1011.6665 [physics.ins-det].

[63] M. Aaboud, G. Aad, B. Abbott, J. Abdallah, O. Abdinov, B. Abeloos, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, and et al., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, The European Physical Journal C **77** (2017), http://dx.doi.org/10.1140/epjc/s10052-017-5031-2.

[64] ATLAS Collaboration, G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv:1603.02934 [hep-ex].

[65] ATLAS Collaboration, G. Aad et al., *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006, arXiv:1908.00005 [hep-ex].

[66] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse, and A. Salzburger, *The new ATLAS track reconstruction (NEWT)*, Journal of Physics: Conference Series **119** (2008) 032014, https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032014.

[67] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, https://cds.cern.ch/record/1449796.

[68] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, https://cds.cern.ch/record/1700870.

[69] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[70] ATLAS Collaboration, G. Aad et al., *Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, arXiv:2007.02645 [hep-ex].

[71] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. B **659** (2008) 119–126, arXiv:0707.1378 [hep-ph].

[72] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, ATLAS-CONF-2015-037, 2015, https://cds.cern.ch/record/2044941.

[73] ATLAS Collaboration, *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at sqrt(s) = 8 TeV*, ATLAS-CONF-2015-002, 2015, https://cds.cern.ch/record/2001682.

[74] ATLAS Collaboration, G. Aad et al., *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv:1907.05120 [hep-ex].

[75] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, https://cds.cern.ch/record/2273281.

[76] ATLAS Collaboration, *Secondary vertex finding for jet flavour identification with the ATLAS detector*, ATL-PHYS-PUB-2017-011, 2017, https://cds.cern.ch/record/2270366.

[77] ATLAS Collaboration, *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*, ATL-PHYS-PUB-2018-025, 2018, https://cds.cern.ch/record/2645405.

[78] ATLAS Collaboration, M. Aaboud et al., *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, JHEP **08** (2018) 089, arXiv:1805.01845 [hep-ex].

[79] ATLAS Collaboration, M. Aaboud et al., *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv:1802.08168 [hep-ex].

[80] ATLAS Collaboration, G. Aad et al., *Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. D **91** (2015) 012006, arXiv:1408.5191 [hep-ex].

[81] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Wadsworth and Brooks, 1984.

[82] B. P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor, *Boosted decision trees, an alternative to artificial neural networks*, Nucl. Instrum. Meth. A **543** (2005) 577–584, arXiv:physics/0408124.

[83] T. Chen and C. Guestrin, *XGBoost*, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2016), http://dx.doi.org/10.1145/2939672.2939785.

[84] ATLAS Collaboration, M. Aaboud et al., *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173–191, arXiv:1806.00425 [hep-ex].

[85] J. C. Collins and D. E. Soper, *Angular distribution of dileptons in high-energy hadron collisions*, Phys. Rev. D **16** (1977) 2219–2225, https://link.aps.org/doi/10.1103/PhysRevD.16.2219.

[86] ATLAS Collaboration, G. Aad et al., *A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector*, Phys. Lett. B **812** (2021) 135980, arXiv:2007.07830 [hep-ex].

[87] ATLAS Collaboration, M. Aaboud et al., *Measurement of the Drell-Yan triple-differential cross section in pp collisions at* $\sqrt{s} = 8$ *TeV*, JHEP **12** (2017) 059, arXiv:1710.05167 [hep-ex].

[88] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, https://cds.cern.ch/record/2677054.

[89] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv:1007.1727 [physics.data-an], [Erratum: Eur.Phys.J.C 73, 2501 (2013)].

[90] CMS Collaboration, *Measurement of Higgs boson decay to a pair of muons in proton-proton collisions at* $\sqrt{s} = 13\,\text{TeV}$, CMS-PAS-HIG-19-006, 2020, https://cds.cern.ch/record/2725423.

[91] CERN, *HL-LHC Project schedule*, https://project-hl-lhc-industry.web.cern.ch/content/project-schedule.

[92] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, cern-lhcc-2017-005; atlas-tdr-025, 2017. https://cds.cern.ch/record/2257755.

[93] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System*, cern-lhcc-2017-020; atlas-tdr-029, 2017. https://cds.cern.ch/record/2285584.

[94] ATLAS Collaboration, *New Small Wheel Technical Design Report*, cern-lhcc-2013-006; atlas-tdr-020, 2013. https://cds.cern.ch/record/1552862.