ALBERT - LUDWIGS - UNIVERSITÄT
FREIBURG

MASTER'S THESIS

Search for CP - violation in gluon fusion production of the Higgs boson using the $H \rightarrow \tau_{\mathbf{lep}}\tau_{\mathbf{had}}$ decay at $\sqrt{s} = 13$ TeV with the ATLAS detector

Author:
Daariimaa BATTULGA

Supervisor:
Prof. Dr. Markus SCHUMACHER

A thesis submitted in fulfillment of the requirements
for the degree of Master of Science

in the

Physikalisches Institut
Fakultät für Mathematik und Physik

November 18, 2019

ALBERT - LUDWIGS - UNIVERSITÄT FREIBURG

Abstract

Fakultät für Mathematik und Physik
Physikalisches Institut

Master of Science

Search for CP - violation in gluon fusion production of the Higgs boson using the H → $\tau_{\textbf{lep}}\tau_{\textbf{had}}$ decay at $\sqrt{s} = 13$ TeV with the ATLAS detector

by Daariimaa BATTULGA

CP-violation in the Standard Model as described by the CKM matrix [1, 2] is not sufficient to explain the observed baryon asymmetry [3]. Thus, it is essential to find new sources of CP- violation in the Higgs-boson sector. Since the observation of the Higgs boson in 2012 [4, 5], the CP nature of the Higgs boson can be studied by its coupling to fermions and gauge bosons.

This analysis focuses on the effective anomalous coupling of the Higgs boson to gluons. Higgs boson production in gluon fusion is dominated by the top quark loop, which is assumed to be in the limit of infinite top quark mass. The mixing angle of $\tan\alpha$ between CP-even and CP-odd contribution parametrizes the vertex of the Higgs boson coupling to gluons. In order to measure the $\tan\alpha$ value, this thesis studies several CP-odd observables: the Optimal Observable and the azimuthal angle between two outgoing jets. A direct test of CP-invariance is performed in Higgs boson production via gluon fusion in association with two jets using the decay $H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$. Analysis of this thesis is based on proton-proton collision at the center of mass energy of $\sqrt{s} = 13$ TeV and a dataset of the integrated luminosity of 139 fb^{-1} collected by the ATLAS experiment.

Contents

Chapter 1

Introduction

The theoretical model describing the fundamental particles and their interactions were developed between 1960 and 1970 [6–8]. The Standard Model (SM) contains the electromagnetic, strong and weak forces, which are mathematically described by relativistic quantum field theories. The fourth force is gravity, which is not described in this model. In the SM, fundamental particles are grouped into fermions and bosons depending on their spin. Fermions have a half-integer spin particle which form the matter while bosons have an integer spin particle which carry the interaction forces.

Initially, quantum field theories of the SM contain no mass terms for the fundamental particles, which breaks local gauge invariance. However, fermions and bosons have a mass, as observed in the experiment. For example, gauge bosons $W^{\pm}$ with a mass of $m_W = 80.379 \pm 0.012$ GeV and the Z^0-boson with a mass of $m_Z = 91.1876 \pm 0.0021$ GeV have been found experimentally [9–11]. This phenomenon is explained by the scalar Higgs field with a vacuum expectation value by introducing the Brout-Englert-Higgs (BEH) mechanism [12–14]. The BEH mechanism is based on the spontaneous symmetry breaking of the electroweak theory. According to this mechanism, fermion, bosons get masses through the interaction with the vacuum expectation value of the Higgs field. The existence of a new particle called Higgs boson is postulated.

Observation of the Higgs boson was announced in 2012 by ATLAS and CMS collaborations at LHC. The Higgs boson has a mass of $m_H = 125.09 \pm 0.21(\text{stat}) \pm 0.11(\text{syst})$ from the combined measurement of the ATLAS and CMS [4, 5]. This scientific achievement allows the detailed investigation of the Higgs boson properties like its production and decay modes, and its spin (J).

In the SM, the Higgs boson has a positive eigenstate of the charge conjugation and parity transformation (CP), which is denoted as $J_{CP} = 0^+$. In particle physics, charge conjugation describes the symmetry between particle and antiparticle, while parity transformation describes the mirror inversion of the physics system. This C and P and CP symmetry are conserved in electromagnetic and strong interactions. Lev Landau [15] suggests the CP symmetry in order to reinstate the parity violation in the weak interaction, as observed in the 1950s [16–18]. Then, CP-violation in weak interaction is included in the SM via the CKM matrix [1, 2]. However, the observed CP-violation in the SM is not sufficient to explain the observed asymmetry of the baryons in the universe [3]. Therefore, an additional source of CP-violation is necessary. Hence, this thesis focuses on the study of the CP nature of the Higgs boson sector.

CP-even particle is an eigenstate of CP transformation with eigenvalue +1, and particle with eigenvalue -1 is said to be in CP-odd. CP-violation occurs when the particles are found to not correspond to a CP eigenstates. Also, CP-violation can be identified if a mixture of CP-even and CP-odd interactions is obtained. The CP

nature of the Higgs boson can be analyzed by studying the couplings of the Higgs boson to other particles. The CP properties are measured in the Higgs boson coupling to vector bosons. The analysis of the Higgs boson production in vector boson fusion and using the decay H→ $\tau\tau$ and H→ZZ, WW, $\gamma\gamma$ shows that the Higgs boson agrees with the SM prediction of spin-zero with dominant contribution from CP-even interaction [19–22]. In addition, direct tests of CP-invariance are also performed in the Higgs boson production in gluon fusion associated with two jets using decay channel of H→ $\tau_{lep}\tau_{had}$ at $\sqrt{s} = 8$ TeV with $\mathcal{L} = 20.3\,\text{fb}^{-1}$ [23]. Investigation of the Higgs boson production in the gluon fusion (ggF) is important as it can, in principle, contain CP-even and CP-odd contribution at the same level. In the vector boson fusion, CP-odd contribution is suppressed due to the loop level effect.

This thesis focuses on the anomalous coupling structure of the Higgs boson to gluons in the ggF production with two jets in the final state in order to test CP-invariance. The subsequent decay of the Higgs boson H→ $\tau_{lep}\tau_{had}$ is considered. The coupling of the Higgs boson to gluons is dominated by the top quark loop, which can be approximated by effective point-like coupling in the limit of $m_{top} \rightarrow \infty$. The effective field theory to describe the Higgs boson coupling to gluons allows to parametrize the mixing between CP-even and CP-odd states. The relative contribution of CP-even and CP-odd interaction is parametrized by the parameter $\tan\alpha$. In order to measure the relative contributions, this analysis uses the following CP-odd observables; the Optimal Observable and the azimuthal angle difference between two outgoing jets. This analysis focus on the data collected by the ATLAS experiment at the LHC in collisions of the proton-proton at a center of mass energy of $\sqrt{s} = 13$ TeV with an integrated luminosity of $\mathcal{L} = 139\,\text{fb}^{-1}$.

The necessary theoretical background is described in chapter 2. The signal and background processes are discussed in chapter 3. It is followed by the investigation of the Higgs boson coupling to gluons with the reweighting technique [23] in chapter 4. Also, this chapter describes the sensitivity study to the anomalous coupling at the generator level for different subprocesses of the gluon fusion production. The experimental setup is discussed in chapter 5 with the definition of physics object in chapter 6.

Furthermore, the selection of the signal events is described in chapter 7. The expected results of the analysis to the sensitivity of the Higgs boson coupling to gluons are presented in chapter 8. The results of the analysis and outlook for future studies are summarized in chapter 9.

Chapter 2

Theoretical Background

This chapter provides the theoretical background for this analysis. The fundamental particles and their interaction forces will be discussed. In particular, the phenomenology of the Higgs boson and its decay and production modes are reviewed. In this thesis, the Higgs boson interacting with gluons is of particular interest. Therefore, an overview of the effective field theory of the Higgs boson coupling with gluons is given. Furthermore, the method to test CP-invariance as well as, CP-odd observables will be presented.

2.1 The Standard Model

The Standard Model (SM) [6–8] is the theoretical model for describing the fundamental particles (elementary particles) and their interactions. This theoretical model has been developed between 1960 and 1970, and it contains the electromagnetic, strong, and weak forces of the elementary particle physics. The fourth force is gravity; however, this interaction is so far, not included in the SM. Most of the physics phenomena can be explained by these three forces. Fig. 2.1 illustrates particle content of the SM.
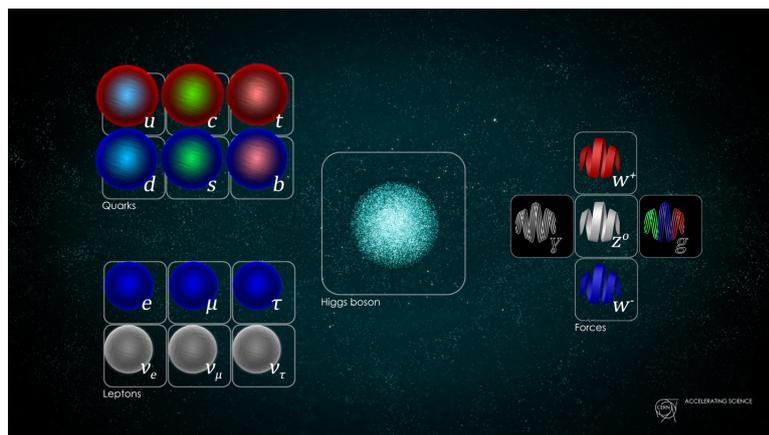
FIGURE 2.1: Standard Model of particle physics [24]. Fundamental particles, fermions as categorized as leptons and quarks with their interaction force carrier bosons, with the Higgs boson are illustrated.

The elementary particles are categorized as the fermions and the bosons. The fermions are called the matter particles, and they have a half-integer spin. Furthermore, fermions are classified into quarks and leptons based on their interactions. The quarks and the leptons contain six particles each, which they are corresponds

to three different generations. There are three generations of the elementary particles, which are divided by their stability and particle's mass and particle flavour and respective quantum numbers. The first generation particles are, in principle, have a small mass and it is stable which means does not decay further. For second and third-generation particles are more massive, and it becomes unstable. Also, it decays

	Leptons				Quarks			
Generation	Particle		Q[e]	mass [GeV]	Particle		Q[e]	mass [GeV]
I	electron	e^-	-1	0.0005	down	d	-1/3	0.005
	neutrino	ν_e	0	$< 10^{-9}$	up	u	+2/3	0.003
II	muon	μ^-	-1	0.106	strange	s	-1/3	0.1
	neutrino	ν_μ	0	$< 10^{-9}$	charm	c	+2/3	1.3
III	tau	τ^-	-1	1.78	bottom	b	-1/3	4.2
	neutrino	ν_τ	0	$< 10^{-9}$	top	t	+2/3	172.4

TABLE 2.1: The fermions are listed with their generation and electric charge Q with its mass [25, 26]. The Q is usually has a units of elementary electric charge $e = 1.6 \times 10^{-19}$ C which it is carried by an electron.

to stable elementary particles. In the Table 2.1 first generation quarks are up (u) and down (d) quarks, second-generation quarks are charm (c), and strange (s) quarks. Third generation quarks are top (t) and bottom (b) quarks. Each quark has an anti-quark that has an opposite charge and other additive quantum numbers. Quarks carry color charge: red, blue and, green of the strong force. In general, quarks are confined in composite particle called hadrons, which are color neutral.

Like quarks, the leptons are also divided into three different generations with their antiparticles. The first generation leptons are electron (e^-) with corresponding electron neutrino (ν_e). The second generations are muon (μ^-) and muon neutrino (ν_μ), and the third generation leptons are tau-lepton (τ^-) and tau neutrino (ν_τ). Except the lepton neutrinos, leptons have an electrical charge. From the neutrino oscillations neutrinos observed to carry very little mass.

Another class of elementary particles are the bosons which have an integer spin and it is called the force carrier particles. Each fundamental forces have their own force carrier particles and it acts in different range with different strength. The photons (γ) have no mass and they carry the electromagnetic force, while massless gluons (g) carry the strong interaction. The weak interactions is carried by the massive $W^\pm$ and Z-bosons.

Fundamental force	Strength	Boson		Spin	Q[e]	Mass [GeV]
Strong	1	Gluon	g	1	0	0
Electromagnetic	10^{-3}	Photon	γ	1	0	0
Weak	10^{-8}	W-boson	$W^\pm$	1	±1	80.4
		Z-boson	Z^0	1	0	91.2

TABLE 2.2: The fundamental forces in nature are listed with corresponding strength, force carrier and as well as spin, charge, and mass. Note that strength values are approximately defined between two particles with a distance of 10^{-15} m [25, 26].

All the force carrier particles have a spin of one and are called gauge bosons. These boson are listed with their corresponding fundamental forces in the Table 2.2.

The electromagnetic force has an infinite range and force carrier photon interacts with all particle with electric charge (e$^-$, μ^-, τ^-, and quarks). In the strong force, gluons have a eigth flavour states depending on their color charges. Those flavour states are defined by the Gell-Mann matrices. Massive gauge W$^\pm$ and Z-bosons interact with all leptons by weak isospin (I$_w$) in limited range.

In the SM, the Higgs boson is a scalar particle with spin of zero. The phenomenology of this scalar particle will be discussed thoroughly in the Sec. 2.2.

2.1.1 Gauge theories of fundamental interactions in the SM

In particle physics, all fundamental particles and their interacting forces are described mathematically by quantum field theory (QFT). QFT is the theoretical framework that uses the Lagrangian formalism in order to describe a model for elementary particles and their interaction.

The gauge theory describes the Lagrangian is invariant under the local transformation of any group in Lie algebra. The Lie groups are classified in general abelian and non-abelian groups. Non-abelian groups are usually referred to as non-commutative symmetry groups.

Furthermore, all particle interactions can be visualized by Feynman diagram. These Feynman diagrams are used as a computing tool in the relativistic quantum field theory.

In the following, relativistic field theories will be discussed in brief.

Quantum Electrodynamics

The electromagnetic interaction between elementary particles is mathematically described by Quantum Electrodynamics (QED) [27–29]. QED is the abelian gauge theory which is based on U(1)$_Q$ symmetry group, and its Lagrangian is written as

$$\mathcal{L}_{\mathrm{QED}} = \overline{\psi}(\imath\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.1}$$

where γ_μ are Dirac's matrices, ψ, $\overline{\psi} := \psi^\dagger\gamma^0$ are Dirac spinor fields for particles with half-integer spin. And $F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor. Whereas, $D_\mu := \partial_\mu + \imath Q A_\mu$ is the gauge covariant derivative. Here, A_μ is the covariant four potential of the electromagnetic field. In addition, e is the electric charge and m is the mass of the fermion, $\partial_{\mu/\nu} = \partial/\partial x^{\mu/\nu}$ is the partial derivatives. The Lagrangian of the QED is based on the Dirac equation which can be derived from Euler-Lagrangian equation of motion.

The $\mathcal{L}_{\mathrm{QED}}$ is invariant under local transformation U(1)$_Q$ symmetry group given by

$$\begin{aligned} \psi(x) &\to e^{-\imath Q\alpha(x)}\psi(x) \\ A_\mu &\to A_\mu + \partial_\mu\alpha(x), \end{aligned} \tag{2.2}$$

where $\alpha(x)$ is the local phase which depends on the space and the time. The mass $(-1/2)m^2 A^\mu A_\mu$ term would break the gauge invariance of the U(1)$_Q$, hence QED needs a massless γ photon.

Quantum Chromodynamics

The strong interaction between quarks and gluons is mathematically described by Quantum Chromodynamics (QCD) [30]. This field theory is a non-abelian gauge theory, which is invariant under local SU(3)_C transformations, where C denotes colour charge.

As discussed earlier, gluons are the force carrier of the strong interaction, and quarks gluons formed as hadrons. Also, quarks and gluons are found to carry the three different color charges: red, blue, and green. In nature, quarks are not observed as free particles due to the confinement phenomenon. A pair of quark and anti-quark is formed as mesons or three quark (anti-) quarks formed as baryons because the strength of the strong force increases when the distance between quarks is increased. Increasing strong force creates the another hadron when the spatial separation between quarks increased. Another QCD phenomena are the asymptotic freedom [30, 31], which describes the two particles' spatial separation decreases the strength of the interaction become weak. As a result, quarks, are freely moving in a certain distance.

Strong interaction formulated as

$$\mathcal{L}_{\text{QCD}} = \overline{\psi}_i(\imath(\gamma^\mu D_\mu)_{ij} - m\delta_{ij})\psi_j - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a, \tag{2.3}$$

where $\psi_{i,j}(x)$ is quark field where a function of space and time, the indices i, j illustrate the color indices in the SU(3) group, m is the quark mass, and $D_\mu := \partial_\mu - \imath g_s G^a_\mu \lambda_a/2$ is the gauge covariant derivative. And $G^a_{\mu\nu} := \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu$ is the gluon field tensor as same as electromagnetic field tensor $F_{\mu\nu}$. The gluon field denoted as $G^{a/b/c}_{\mu/\nu}(x)$ as a function of space and time, where index a, b, c represents the eight different gluons and λ_a are the Gell-Mann matrices. This eight Gell-Mann matrices descibe the colour symmetry representation of the SU(3) group. The coupling constant denoted as g_s, and f^{abc} are the structure constant of SU(3) group.

The Lagrangian of the QCD is invariant under following local SU(3) transformation

$$\psi(x) \rightarrow exp\left(\imath\frac{g_s}{2}\sum_a^8 \lambda_a\beta_a(x)\right)\psi(x), \tag{2.4}$$

$$G^\mu_a \rightarrow G^\mu_a - \partial^\mu\beta_a(x) - g_s f_{abc}\beta_b(x)G^\mu_c,$$

where $\beta_{a/b}(x)$ are the eight-component transformation functions as function of space and time.

Electroweak theory and broken symmetries

The weak interaction between elementary particles has a small range, approximately $m_W^{-1} \sim 10^{-15}$ cm. This interaction couples to the weak isospin through charged $W^\pm$ and neutral Z-bosons. Furthermore, the weak interaction is well established by the electroweak theory. The electroweak theory combines the two fundamentals of four interactions: electromagnetic and weak interactions. This theory has been developed by Sheldon Glashow [6], Abdus Salam [7], and Steven Weinberg [8] in 1967-1968, and it is known as the Glashow-Weinberg-Salam Model (GWS).

The gauge invariance always forces the mediating gauge bosons to be massless. However, the $W^\pm$ and Z-bosons have a large mass that experimentally proven.

Therefore any mass term appears in the Lagrangian, broke the gauge symmetry due to mass generation in the vector field. The spontaneous symmetry breaking can explain the heavy mass of the vector bosons.

Initially, gauge bosons and fermion introduced to be massless by a gauge symmetry. Then the scalar Higgs field introduces the non-vanishing vacuum expectation value by the Higgs mechanism. Formally, it is called as Brout-Englert-Higgs (BEH) mechanism [12–14, 32]. As a result, the theory of the spontaneous symmetry breaking gives the mass to electron, muon, tau leptons, quarks, and gauge bosons. The photons and gluons remain massless. In this analysis, neutrinos are considered to be massless.
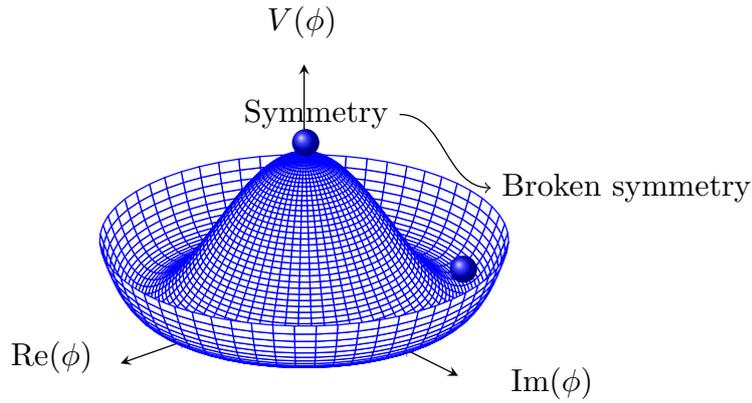
FIGURE 2.2: The Higgs potential field is illustrated in spontaneous symmetry breaking.

The lagrangian of the electroweak theory is based on the $SU(2)_{L, I_w} \times U(1)_Y$ symmetry group. The index L refers to the weak interaction couples to the left-handed fermions by weak isospin I_w and Y is the hypercharge. The relation between I_w and Y and Q is described by the Gell-Mann-Nishijima relation,

$$Q = I_w^3 + \frac{Y}{2}. \tag{2.5}$$

The SU(2) singlets describe fermions with the right-handed states with weak isospins of zero. In contrast, left-handed fermions are described by the SU(2) doublets with $I_w^3 = \pm 1/2$ weak isospin and same hypercharge.

The lagrangian formalism before symmetry breaking is given by

$$\mathcal{L}_{EW} = \mathcal{L}_g + \mathcal{L}_f + \mathcal{L}_h + \mathcal{L}_y, \tag{2.6}$$

where $\mathcal{L}_g$ describes the interactions of three W vector bosons and B vector boson. And this lagrangian is formulated as

$$\mathcal{L}_g = -\frac{1}{4} W_a^{\mu\nu} W_{\mu\nu}^a - \frac{1}{4} B^{\mu\nu} B_{\mu\nu}, \tag{2.7}$$

where $W_a^{\mu\nu}$, $B^{\mu\nu}$ are the field strength tensors to the gauge field of the weak isospin and weak hypercharge, and index $a = 1, 2, 3$ numbers of the vector bosons. The second term, $\mathcal{L}_f$, describes the interaction between gauge bosons and fermions. This

term is given by

$$\mathcal{L}_f = \overline{Q}_i \imath \slashed{D} Q_i + \overline{u}_i \imath \slashed{D} u_i + \overline{d}_i \imath \slashed{D} d_i + \overline{L}_i \imath \slashed{D} L_i + \overline{e}_i \imath \slashed{D} e_i, \tag{2.8}$$

where Q_i is the left-handed quark doublet, u_i is the right-handed singlet up, and d_i is the right-handed singlet down quark fields. Whereas, L_i is the left-handed lepton doublet, e_i is the right-handed singlet electron field. The index i represents the three generations of the fermions. The notation $\slashed{D} := \gamma^\mu D_\mu$ is the gauge covariant derivative.

The third term, $\mathcal{L}_h$, illustrates the Higgs field and its interaction of the Higgs field and to the gauge bosons. And this is written by

$$\mathcal{L}_h = |D_\mu h|^2 - \lambda \left(|h|^2 - \frac{v^2}{2} \right), \tag{2.9}$$

where h is the Higgs field, and v is the vacuum expectation value which is zero.

The last lagrangian term $\mathcal{L}_y$ defines the Yukawa interaction of the fermions by

$$\mathcal{L}_y = -y_{uij} \varepsilon^{ab} h_b^\dagger \overline{Q}_{ia} u_j^c - y_{dij} h \overline{Q}_i d_j^c - y_{eij} h \overline{L}_i e_j^c + h.c., \tag{2.10}$$

where y_{uij}, y_{dij} y_{eij} are the Yukawa coupling strength for fermions, and $h.c.$ denotes the hermitian conjugation of the corresponding term.

The lagrangian of the electroweak theory after the symmetry breaking is mathematically formulated by

$$\mathcal{L}_{EW} = \mathcal{L}_K + \mathcal{L}_N + \mathcal{L}_C + \mathcal{L}_H + \mathcal{L}_{HV} + \mathcal{L}_{WWV} + \mathcal{L}_{WWVV} + \mathcal{L}_Y, \tag{2.11}$$

where $\mathcal{L}_K$ includes the mass terms and self interaction of $W^\pm$, Z-boson. It defined as following

$$\begin{aligned}
\mathcal{L}_K &= \sum_f \overline{f}(\imath \slashed{\partial} - m_f)f - \frac{1}{4} A_{\mu\nu} A^{\mu\nu} - \frac{1}{2} W_{\mu\nu}^+ W^{-\mu\nu} + m_W^2 W_\mu^+ W^{-\mu} \\
&\quad - \frac{1}{4} Z_{\mu\nu} Z^{\mu\nu} + \frac{1}{2} m_Z^2 Z_\mu Z^\mu + \frac{1}{2} (\partial^\mu H)(\partial_\mu H) - \frac{1}{2} m_H^2 H^2,
\end{aligned} \tag{2.12}$$

where the sum contains all fermions.

$$\mathcal{L}_N = e J_\mu^{EM} A_\mu + \frac{g}{\cos \theta_W} (J_\mu^3 - \sin^2 \theta_W J_\mu^{EM}) Z^\mu, \tag{2.13}$$

where $e = g \sin \theta_W = g' \cos \theta_W$ charge, and $J_\mu^{EM} := \sum_f d_f \overline{f} \gamma_\mu f$ is the electromagnetic current while $J_\mu^3 := \sum_f I_f^3 \overline{f} \gamma_\mu \frac{1-\gamma^5}{2} f$ is neutral weak currents. Here, q_f are the electric charge of the fermion and I_f^3 are the weak isospin of the fermions.

The next term describes the charged current of the Lagrangian, and it is,

$$\mathcal{L}_C = -\frac{g}{\sqrt{2}} \left[\overline{u}_i \gamma^\mu \frac{1-\gamma^5}{2} M_{ij}^{CKM} d_j + \overline{v}_i \gamma^\mu \frac{1-\gamma^5}{2} e_i \right] W_\mu^+ + h.c., \tag{2.14}$$

where $\frac{1-\gamma^5}{2}$ is the factor of weak couplings. It projects to the left-handed components of the spinor fields. The term $\mathcal{L}_H$ describes the self-interaction term of the Higgs

boson as

$$\mathcal{L}_H = -\frac{gm_H^2}{4m_W}H^3 - \frac{g^2m_H^2}{32m_W^2}H^4, \tag{2.15}$$

where m_H, m_W is the mass of the Higgs boson and W-boson. And $\mathcal{L}_{HV}$ gives the interaction term of the gauge bosons with vector boson by

$$\mathcal{L}_{HV} = \left(gm_WH + \frac{g^2}{4}H^2\right)\left(W_\mu^+W^{-\mu} + \frac{1}{2\cos^2\theta_W}Z_\mu Z^\mu\right), \tag{2.16}$$

where θ_W is the weak mixing angle. A self three-point interactions of the gauge boson defined the following

$$\begin{aligned}
\mathcal{L}_{WWV} &= -\imath g\left[(W_{\mu\nu}^+W^{-\mu} - W^{+\mu}W_{\mu\nu}^-)(A^\nu\sin\theta_W - Z^\nu\cos\theta_W)\right.\\
&+ \left. W_\nu^-W_\mu^+(A_{\mu\nu}\sin\theta_W - Z^{\mu\nu}\cos\theta_W)\right].
\end{aligned} \tag{2.17}$$

And four-point interaction term of the gauge bosons are illustrated as

$$\begin{aligned}
\mathcal{L}_{WWVV} &= -\frac{g^2}{4}\left\{\left[2W_\mu^+W^{-\mu} + (A_\mu\sin\theta_W - Z_\mu\cos\theta_W)^2\right]^2\right.\\
&- \left[W_\mu^+W_\nu^- + W_\nu^+W_\mu^-\right.\\
&+ \left.\left.(A_\mu\sin\theta_W - Z_\mu\cos\theta_W)(A_\nu\sin\theta_W - Z_\nu\cos\theta_W)\right]^2\right\}.
\end{aligned} \tag{2.18}$$

The last term are the Yukawa interactions,

$$\mathcal{L}_Y = -\sum_f \frac{gm_f}{2m_W}\overline{f}fH, \tag{2.19}$$

of the Higgs field interacting with the fermions.

2.2 Phenomenology of the Higgs boson

The Higgs mechanism describes the spontaneous symmetry breaking in the electroweak theory. This broken symmetry explains how gauge bosons of the weak interaction acquire their heavy mass. The mass of the W-boson $m_W = 80.379 \pm 0.012$ GeV and mass of the Z-boson $m_Z = 91.1876 \pm 0.0021$ GeV are experimentally determined [9–11].

From the Higgs mechanism, the Higgs field has a non-zero vacuum expectation value everywhere. The existence of this field is proved by the existence of the scalar particle called Higgs boson, which is the excitation of the Higgs field.

With the Higgs boson, the Standard Model of particle physics is completed. The predicted Higgs boson has been observed observed on July 4, 2012, by ATLAS and CMS experiments in proton-proton collisions at the LHC at CERN. The Higgs boson has a mass of $m_H = 125.09 \pm 0.21$ (stat) ± 0.11 (syst), with zero spin and no electric or colour charge [4, 5]. Besides, Higgs boson has a positive eigenstate under charge and parity transformation (CP), and this is denoted as $J^{CP} = 0^+$. Different properties of the SM Higgs boson will suggest physics beyond the Standard Model (BSM).

The lifetime of the Higgs boson is short and it is approximately 10^{-22} s, therefore, it decays further into the fermions and bosons.

The following sections are written based on the reference [33].

2.2.1 Production and decay rates of Higgs boson at the LHC

In proton-proton (pp) collisions, the Higgs boson can be produced by in general, four different production modes: gluon fusion (ggF), vector boson fusion (VBFH), in association with weak vector boson (VH), and in association with a pair of top quark production (tt̄H). Example of Feynman diagrams are sketched in Fig. 2.3. The
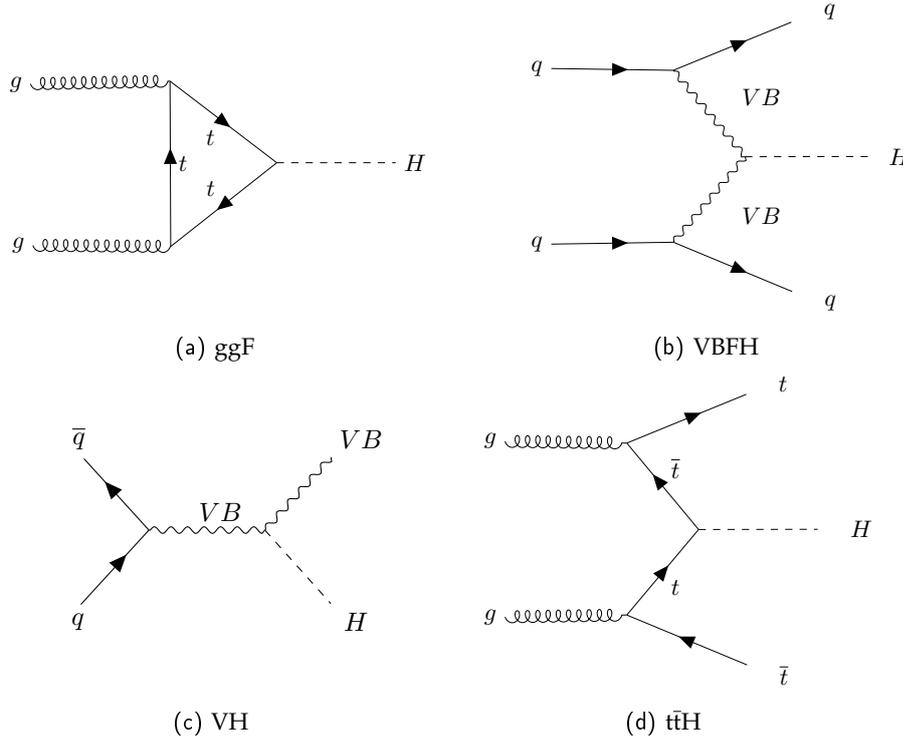
FIGURE 2.3: Example of Feynman diagrams to the productions of the Higgs boson in (a) gluon fusion, (b) vector boson fusion, (c) in association with weak vector boson, and (d) in association with a pair of top quarks at leading order (LO) are shown.

dominant production mode of the Higgs boson is the gluon fusion process, because the contribution of the gluons at the relevant energy scale is huge. This quark and gluon contribution is described by the parton distribution functions (PDF).

This PDF gives the probability to find parton i, which carries the momentum fraction x_i of the proton probed at momentum scale μ^2. The parton distributions are shown in Fig. 2.4.

The total cross section of a process ($pp \rightarrow X$) in pp collisions can be calculated by

$$\sigma_{pp \rightarrow X} = \int \int dx_i dx_j f(x_i, \mu^2) f(x_j, \mu^2) \hat{\sigma}_{ij} \tag{2.20}$$

$$\text{with } \hat{\sigma}_{ij} = \int \frac{|\mathcal{M}(ij \rightarrow X)|^2}{F} dQ$$

where index i, j represents the partons, and $\hat{\sigma}_{ij}$ is the cross section at parton level. And $\mathcal{M}$ is the matrix element of the processes, whereas $|\mathcal{M}|^2$ is the transition probability between initial and final states, F is the flux of the particle, and dQ is the phase space factor of the kinematics. The cross section measurements for SM processes in
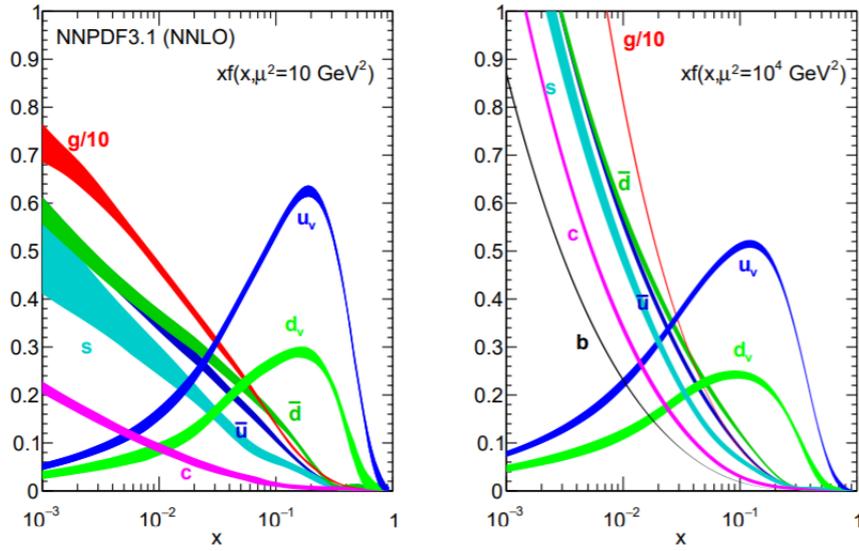
FIGURE 2.4: Parton distribution functions (PDF) at $\mu^2 = 10\,\text{GeV}^2$ (left) and at $\mu^2 = 10^4\,\text{GeV}^2$ (right) . The PDFs are calculated at next-to-next-to-leading (NNLO) order and it is evaluated by NNLOPDF3.1 [34].

pp collisions are shown in Fig. 2.5. Detailed information for relevant event processes in this analysis will be given in Chapter 3.
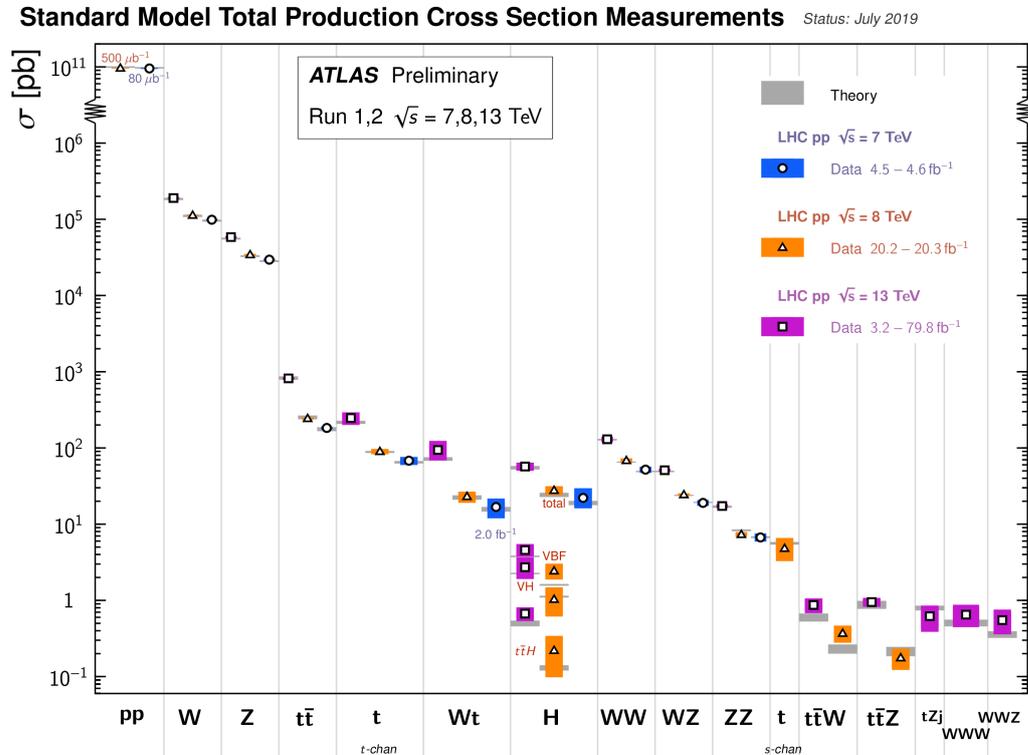
FIGURE 2.5: SM cross section in pp collisions in comparison to theoritical predictions [35].

As shown in Fig. 2.3a, the Higgs boson production in gluon fusion via top quark loop is the dominant production mode at the LHC. A b-quark loop contributes to the gluon fusion, but it is negligible.

Furthermore, Higgs boson can be produced by vector boson ($W^\pm$, Z^0) fusion, and example Feynman diagram at LO are illustrated in Fig. 2.3b. From the initial state quarks, vector bosons are radiated, and further fusion of the bosons produce the Higgs boson. Outgoing jets have a large difference of the pseudorapidity ($\Delta\eta_{jj}$) and have a large invariant mass m_{jj} in the vector boson production.

Fig. 2.3c shows an example of Feynman diagrams at LO of the Higgs production in association with a vector boson. Annihilation of a pair of quarks produces a vector boson and Higgs boson. Also, this production mode gives the Higgs boson coupling to a pair of W and Z-boson like VBFH.

Another production mode of the Higgs boson is in association with top quark pair, as sketched in Fig. 2.3d. The Higgs bson stems from a pair of the top quark. The production cross sections at $\sqrt{s} = 13$ TeV for a Higgs boson with mass of 125 GeV are shown in Fig. 2.6. Moreover, Table 2.3 summarizes the inclusive cross sec-
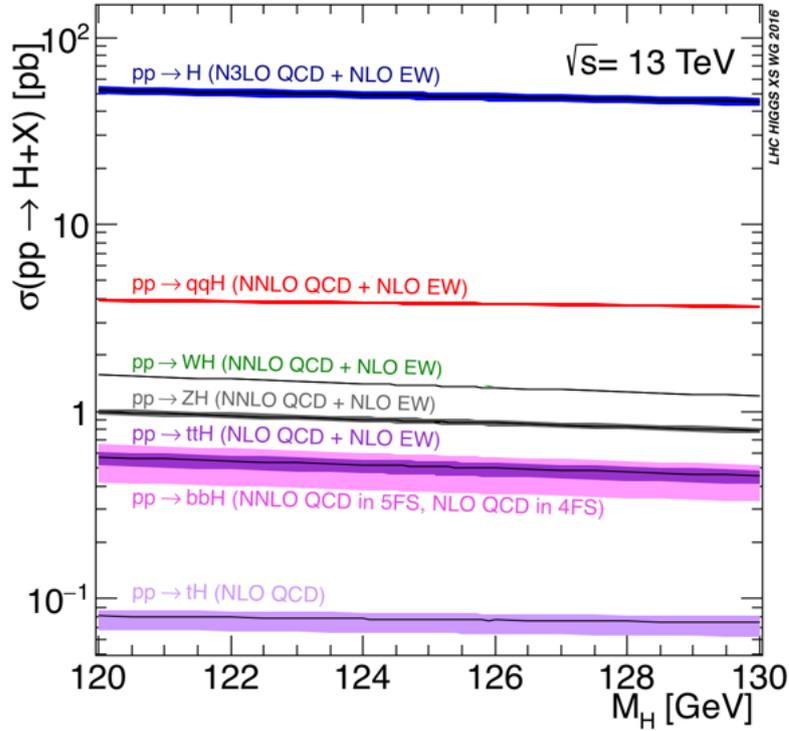
FIGURE 2.6: Cross section of the Higgs boson production as a function of the mass at $\sqrt{s} = 13$ TeV in Standard Model [36]. The cross section of the tH production includes only s- and t-channel contributions.

tion with uncertainties for the main production modes of the Higgs boson. Uncertainty of cross section of the gluon fusion production is significant because, beyond the leading order, additional jets are radiated in the final state in QCD corrections. As a result, it changes the kinematics in gluon fusion; therefore, it has large QCD corrections.

Branching fractions of various decay channels of the Higgs boson in the SM as a function of Higgs boson mass are shown in Fig. 2.7, and are summarized with their

Process	$\sigma_{\text{inc}}[pb]$	Δ_{scale} [%]	$\Delta_{\text{PDF}} + \alpha_S$ [%]
ggF	48.58	$\pm^{4.6}_{6.7}$	± 3.2
VBFH	3.78	$\pm^{0.43}_{0.33}$	± 2.1
WH	1.37	$\pm^{0.5}_{0.7}$	± 1.9
ZH	0.88	$\pm^{3.8}_{3.1}$	± 1.6
t$\bar{\text{t}}$H	0.49	$\pm^{5.8}_{9.2}$	± 3.6

TABLE 2.3: Inclusive cross section of the Higgs productions with their respective uncertainties: Δ_{scale} is the uncertainity of QCD scale [36].
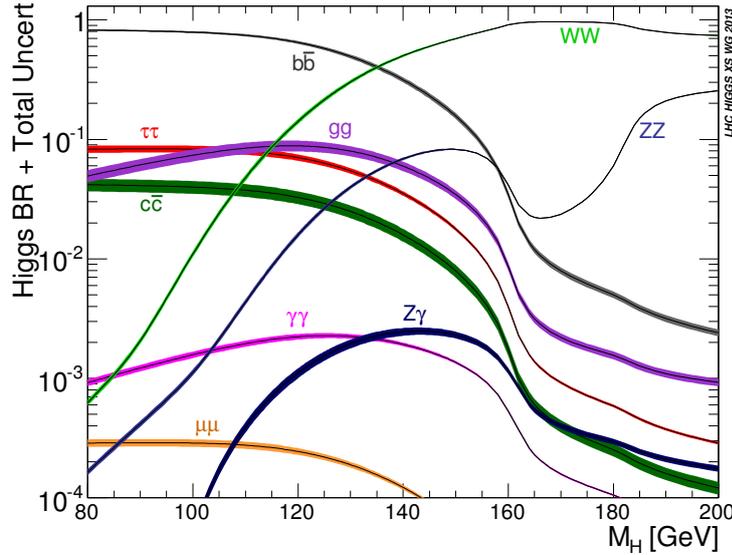
FIGURE 2.7: Predicted branching ratios of the Higgs boson with their respective uncertainties are depicted [36].

in Table 2.4.

Decay channels	Branching fraction [%]	THU [%]	PU(m_q) [%]	PU(α_S) [%]
H$\to$ b$\bar{\text{b}}$	58.24	± 0.65	$\pm^{0.72}_{0.74}$	$\pm^{0.78}_{0.80}$
H$\to$ WW	21.37	± 0.99	$\pm^{0.99}_{0.98}$	$\pm^{0.66}_{0.63}$
H$\to$ gg	8.18	$\pm^{3.40}_{3.41}$	$\pm^{1.12}_{1.13}$	$\pm^{3.69}_{3.61}$
H$\to$ $\tau\tau$	6.27	$\pm^{1.17}_{1.16}$	$\pm^{0.99}_{0.98}$	± 0.62
H$\to$ c$\bar{\text{c}}$	2.89	± 1.20	$\pm^{5.26}_{0.98}$	± 1.25
H$\to$ ZZ	2.62	± 0.99	$\pm^{0.99}_{0.98}$	$\pm^{0.66}_{0.63}$
H$\to$ $\gamma\gamma$	0.22	$\pm^{1.73}_{1.72}$	$\pm^{0.93}_{0.99}$	$\pm^{0.61}_{0.62}$
H$\to$ Zγ	0.15	± 5.71	$\pm^{0.98}_{1.01}$	$\pm^{0.58}_{0.65}$
H$\to$ $\mu\mu$	0.02	± 1.23	$\pm^{0.97}_{0.99}$	$\pm^{0.59}_{0.64}$

TABLE 2.4: Branching fractions of the decay channel of the Higgs boson with respective uncertainties [36]. Here, THU is the theory uncertainty, and PU(m_q) is the parametric uncertainty as a function of quark mass and PU(α_S) is the parametric uncertainty as a function of strong coupling.

The dominant decay of the Higgs boson is the H$\to$ b$\bar{\text{b}}$ and its contribution is roughly 58 %. In addition, Higgs boson decays into fermions: H$\to$ $\tau\tau$ (6.27 %),

H→ c$\bar{c}$, H→ $\mu\mu$. Decay into a pair of top quark is kinematically not allowed due to its high mass.

In this thesis, the Higgs boson decays into a pair of τ-leptons are of particular interest. Especially, semileptonic decay of the τ-leptons will be further studied in Higgs production mode in gluon fusion. It will be discussed in detail Chapter 3.

The Higgs boson also decays into gauge bosons: H→ WW, H→ gg, H→ $\gamma\gamma$, H→ Zγ. The second-largest decay channel of the Higgs boson is a pair of W-bosons (21 %).

The H→ $\gamma\gamma$, H→ ZZ → 4ℓ decay are the discovery channel of the Higgs boson. In 2012, ATLAS and CMS experiments at the LHC found the Higgs boson with a mass of 125.09 GeV [4, 5].

2.2.2 CP-properties of the Higgs boson

This section will motivate the test of CP-invariance in the anomalous coupling of Higgs boson to gluons, which is the main interest of this analysis. Therefore, first of all, the necessary understanding of CP-invariance will be given in the following.

CP transformation refers to the product of charge conjugation and parity symmetry. In particle physics, the charge conjugation illustrates the symmetry between particle and antiparticle. While transformation describes the mirror inversion of the physics system.

C and P and CP-symmetry is conserved in electromagnetic and strong interactions. This CP symmetry suggested by Lev Landau in 1957 [15] to reinstate parity symmetry since the observed parity violation in weak interaction in the 1950s.

In the SM, the CP-symmetry (CP-invariance) is violated in weak interactions and it has been firstly observed in kaon decay in 1964 [16–18]. This is explained by the Cabibbo-
Kobayashi-Maskawa (CKM) matrix [1, 2]. The CKM matrix describes the transition probability V_{ij} of quark i to the quark j and it is defined as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.21}$$

The magnitude of CP-violation that appears in the CKM matrix is not sufficient to explain the observed baryon asymmetry. This asymmetry can be explained by one of the Sakharov conditions, which is the C and CP-violation [3]. Therefore, a new source of CP-violation is required.

In the Standard Model, the spin-parity configuration of the Higgs boson is $J^{CP} = 0^+$. It has a positive eigenstate of the charge conjugation and parity transformation (CP). A positive eigenstate with eigenvalue of +1 is called CP-even, whereas a negative eigenstate with eigenvalue of -1 is referred to the CP-odd. CP-properties of the Higgs boson can be studied by investigating its coupling to the gauge bosons and fermions. Study of CP-properties of the Higgs boson is performed by ATLAS and CMS experiments in the decay H→ $\tau\tau$ and H→ZZ, WW, $\gamma\gamma$ shows that the Higgs boson agrees with the SM prediction of spin-zero with dominant contribution from CP-even interaction [19–22]. In the vector boson fusion, CP-odd contribution is suppressed due to the loop level effect. Therefore, investigation of the Higgs boson production in the gluon fusion (ggF) is important as it can, in principle, contain CP-even and CP-odd contribution at the same level.

This direct test of CP-invariance has been performed in Higgs production in gluon fusion with associated with two jets at $\sqrt{s} = 8$ TeV with $\mathcal{L} = 20.3\,\text{fb}^{-1}$ [23]. This analysis covers the fully leptonic decay of the Higgs boson. Furthermore, Higgs production in vector boson production in the ditau decay channel has been performed [21, 22].

This thesis will investigate the anomalous coupling structure of the Higgs boson to the gluons associated with two jets in order to test CP-invariance. Only the decay of H$\to \tau_{\text{lep}}\tau_{\text{had}}$ is considered. This direct CP-invariance investigation is performed based on effective field theory to parametrize the contribution of a CP-odd coupling in addition to the CP-even coupling in the SM.

2.3 Effective Field Theory of Higgs boson coupling to gluons

The Lagrangian of effective field theory which describes the Higgs characterization is constructed from an underlying $\text{SU}(2)_{I_w, L} \times \text{U}(1)_Y$ symmetry [37]. The advantages of this effective lagrangian are that it includes the dimension six operators for CP-odd contribution, which are invariant under the electroweak scale. Therefore, it provides a description of CP-even and CP-odd mixing which is used in this analysis. This effective Lagrangian describes the three-point Higgs boson interaction with fermions and bosons. However, the term of interest is only Higgs boson interaction with gluons. Hence, the Lagrangian can be written as

$$\mathcal{L}_0^f = - \sum_{f=t,b,\tau} \overline{\psi}_f (\cos\alpha k_{Hff} g_{Hff} + \iota \sin\alpha k_{Aff} g_{Aff}) \psi_f X_0, \qquad (2.22)$$

where $\psi_f, \overline{\psi}_f$ are the fermion spinor field, $g_{Hff} = m_f/v$ is the scalar coupling strength while $g_{Aff} = m_f/v$ denotes the pseudoscalar coupling strength. Here, $v \sim 246$ GeV. Angle α can describe the mixing between 0^+ ($\cos\alpha = 1$, $\sin\alpha = 0$) and 0^- ($\cos\alpha = 0$, $\sin\alpha = 1$) states and k_{Hff}, k_{Aff} are dimensionless real coupling parameters, and X_0 represents the scalar boson. The top-quark loop dominates in Higgs boson production in gluon fusion. In addition, the b-quark and τ-lepton loops are negligible as discussed in Sec. 3.1. In the limit of an infinite mass of the top quark the Lagrangian can be written as,

$$\mathcal{L}_{eff} = \cos(\alpha) g_{Hgg} k_{Hgg} G^a_{\mu\nu} G^{a,\,\mu\nu} H + \sin(\alpha) g_{Agg} k_{Agg} G^a_{\mu\nu} \tilde{G}^{a,\,\mu\nu} A, \qquad (2.23)$$

where $G^a_{\mu\nu} := \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^a_\nu$ is the gluon field tensor whereas $\tilde{G}^a_{\mu\nu} := \varepsilon_{\mu\nu\rho\sigma} G^{a,\,\rho\sigma}$ is the dual tensor. Here, $\varepsilon_{\mu\nu\rho\sigma}$ is the total asymmetric tensor with metric tensor $g^{\mu\nu} = diag(+1, -1, -1, -1)$. And H is the pure scalar ($\cos\alpha = 1$, $\sin\alpha = 0$), A is the pure pseudoscalar ($\cos\alpha = 0$, $\sin\alpha = 1$) component of X_0. As a results of infinite top quark mass, the effective couplings strength are obtained to be

$$g_{Hgg} = -\frac{\alpha_s}{3\pi v}; \quad g_{Agg} = \frac{\alpha_s}{2\pi v}, \qquad (2.24)$$

where α_s is the strong coupling strength. The dimensionless couplings strength ($k_{Hgg} = 1$, $k_{Agg} = -2/3$) is not varied. To obtain this

$$k_{Hgg} g_{Hgg} = k_{Agg} g_{Agg} = -\frac{\alpha_s}{3\pi v} \equiv \tilde{g}. \qquad (2.25)$$

Moreover, to measure the relative contribution of CP-even and odd states, parametrization by tangent is convenient. Therefore, the effective Higgs boson gluon interaction is expressed as,

$$\mathcal{L}_{eff} = \cos(\alpha)\tilde{g}\left[G^a_{\mu\nu}G^{a,\,\mu\nu}H + \tan(\alpha)G^a_{\mu\nu}\tilde{G}^{a,\,\mu\nu}A\right]. \tag{2.26}$$

Here, $\tan\alpha = 0$ describes the no CP-violation (SM) case and $\tan\alpha \neq 0$ yields the CP-violating scenarios.

2.4 Test of CP-invariance and CP-odd observables

Model-dependent or model-independent methods can be used to probe the invariance of the processes. In this thesis, the study of an anomalous coupling of the Higgs boson to the gluons uses a model-independent test of CP-invariance. It can be performed by measuring the mean value of a CP-odd observable $\mathcal{O}_{CP}$, and it defined as

$$\langle\mathcal{O}_{CP}\rangle = \int \frac{\mathcal{O}_{CP}d\Omega}{d\Omega}, \tag{2.27}$$

where Ω is a CP-symmetric phase space. The CP symmetry is conserved when $\langle\mathcal{O}_{CP}\rangle = 0$. And $\langle\mathcal{O}_{CP}\rangle \neq 0$ illustrates the CP-violation. This method refers to a model-independent test of CP-invariance.

The observables which are sensitive to the CP-odd contribution to Higgs boson in gluon fusion are the azimuthal angle between outgoing jets so called tagging jets ($\Delta\Phi_{jj}$) and the Optimal Observable ($\mathcal{OO}$).

2.4.1 Azimuthal angle between two tagging jets

This analysis is concentrated on structure of the Higgs boson production in gluon fusion with associated two jets, ggF H+ 2 jets. In order to probe the CP-invariance, the kinematics of the process is important. The azimuthal angle between the two jets are sensitive to the CP-mixing states, and it is defined as,

$$\Delta\Phi_{jj} = \Phi_{j_1} - \Phi_{j_2}, \tag{2.28}$$

where $j_{1,2}$ represents the outgoing jets. In this analysis, $\Delta\Phi_{jj}$ is ordered by the pseudorapidity η and further, it will be denoted as $\Delta\Phi_{jj}^{\text{ordered}}$. The distribution of $\Delta\Phi_{jj}^{\text{ordered}}$ for $\tan\alpha = -0.4$ and $\tan\alpha = 1$ are showin in Fig. 2.8. The $\Delta\Phi_{jj}^{\text{ordered}}$ event distributions in process ggF H+2 jets are generated by MadGraph at LO at $\sqrt{s} = 13$ TeV. From the figure, CP-even and CP-mixing states have a distinct distributions. It is quantified by the mean value of $\Delta\Phi_{jj}^{\text{ordered}}$. In the SM, $\langle\Delta\Phi_{jj}^{\text{ordered}}\rangle = 0$ consistent with the CP-invariance. However, for CP-mixing states, $\langle\Delta\Phi_{jj}^{\text{ordered}}\rangle = -0.066 \pm 0.005$ is obtained at $\tan\alpha = -0.4$ and $\langle\Delta\Phi_{jj}^{\text{ordered}}\rangle = 0.099 \pm 0.005$ is obtained at $\tan\alpha = 1$.

A different choice of defining a CP-odd angular observable is given by choosing two outgoing jets in different hemispheres. It is referred to as a signed azimuthal angle between two jets and denoted as $\Delta\Phi_{jj}^{\text{sign}}$. It can be expressed as [38],

$$\varepsilon_{\mu\nu\rho\sigma}b_+^\mu p_+^\nu b_-^\rho p_-^\sigma = 2p_{T+}p_{T-}\sin(\Phi_+ - \Phi_-) = 2p_{T+}p_{T-}\sin\Delta\Phi_{jj}, \tag{2.29}$$
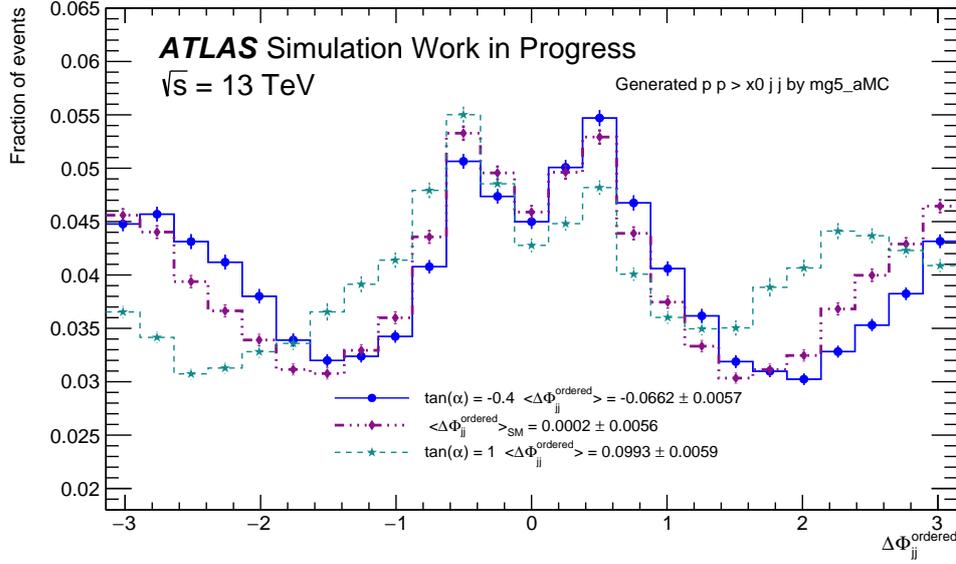
FIGURE 2.8: Normalized distibutions of the azimuthal angle between
jets $\Delta\Phi_{jj}$ ordered by pseudorapidity for different mixing angles $\tan\alpha$
values with the statistical uncertainty. The dashed violet line indi-
cates the SM case where $\tan\alpha = 0$, the solid blue line indicates the
CP-mixing state with $\tan\alpha = -0.4$, and dashed green line illustrates
the $\tan\alpha = 1$ distribution.

where b_+^μ (b_-^μ) are the clockwise(anti-clockwise) circulating proton beams normal-
ized four-momenta, p_+^μ (p_-^μ) are the four-momenta of two outgoing jets. And p_{T+} (p_{T-})
is the four momentum vector of jet point to same detector hemispheres as b_+^μ (b_-^μ).
Normalized distributions of $\Delta\Phi_{jj}^{\text{sign}}$ for SM and CP-mixing states are illustrated in
Fig. 2.9. Also, event distribution is generate by MadGraph5 [39].

2.4.2 Optimal Observable

Another CP-odd observable to investigate the Higgs boson coupling structure to
gluons is the Optimal Observable ($\mathcal{OO}$) [40–42].

This section introduces the parametrization of the CP-even and CP-odd states for
$\mathcal{OO}$ observable.

Effective Lagrangian for anomalous coupling of the Higgs boson to the gluons
is discussed in the Sec. 2.3. From Eq. 2.3, the matrix element that describes the
kinematics of the process can be written as,

$$\mathcal{M} = \mathcal{M}_{\text{SM}} + \tan(\alpha)\mathcal{M}_{\text{CP-odd}}, \tag{2.30}$$

where $\mathcal{M}_{\text{SM}}(\mathcal{M}_{\text{CP-odd}})$ is the matrix element of standard model (CP-odd) Higgs bo-
son coupling to gluons. Moreover, the relative contributions of CP-even and CP-odd
state are described by the mixing angle $\tan\alpha$. And the squared matrix element of the
process is given by

$$|\mathcal{M}|^2 = |\mathcal{M}_{SM}|^2 + \tan(\alpha)2Re\{\mathcal{M}_{SM}^*\mathcal{M}_{CP-odd}\} + \tan^2(\alpha)|\mathcal{M}_{CP-odd}|^2. \tag{2.31}$$
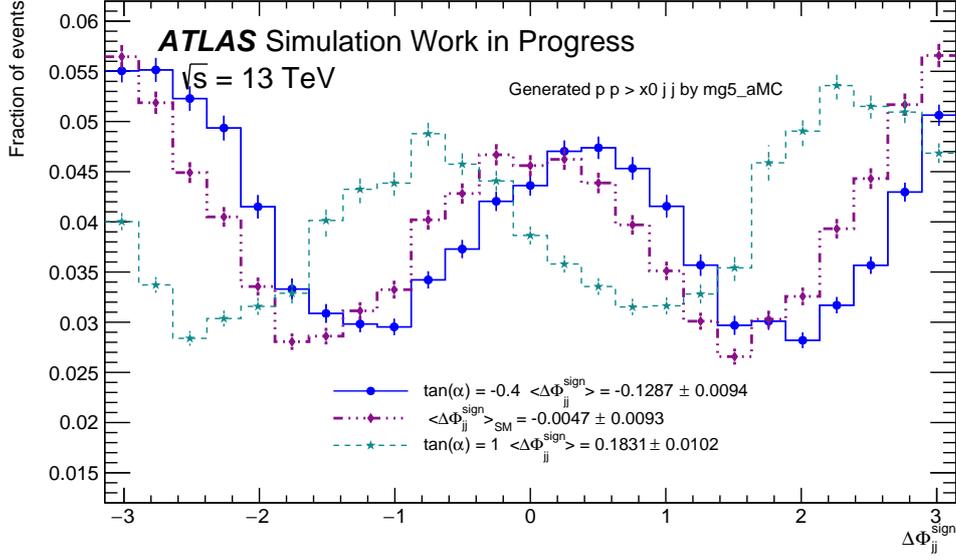
FIGURE 2.9: Normalized distributions of the signed azimuthal angle $\Delta\Phi_{jj}^{\text{sign}}$ for different mixing angles $\tan\alpha$ values with the statistical uncertainty. The dashed violet line indicates the SM case where $\tan\alpha = 0$. While the solid blue line indicates the CP-mixing state with $\tan\alpha = -0.4$, and the dashed green line illustrates the $\tan\alpha = 1$ distributions.

Here, $2Re\{\mathcal{M}_{SM}^*\mathcal{M}_{CP-odd}\}$ is the CP-odd interference term of CP-even and CP-odd couplings. This interference term is sensitive to CP-violation. However, the first and the third term is conserving the sign under the CP-transformation. Hence, $\mathcal{OO}$ is defined by the ratio of interference term and squared matrix element of SM by

$$\mathcal{OO} = \frac{2Re\{\mathcal{M}_{SM}^*\mathcal{M}_{CP-odd}\}}{|\mathcal{M}_{SM}|^2}. \tag{2.32}$$

The advantages of the $\mathcal{OO}$ are that it can be calculated from the four-momentum vectors of Higgs boson and jets. Also, it does not depend on the subsequent decay of the Higgs boson. Most importantly, this one-dimensional variable contains the full phase space information. Thus, it can provide a high sensitivity to the anomalous coupling than other variables. Normalized distributions of the $\mathcal{OO}$ for different CP-mixing scenario are shown in Fig. 2.10.

$\mathcal{OO}$ distribution in the SM is symmetric, and the mean value is compatible with zero. However, in the CP-mixing states, distributions are asymmetric, and for negative $\tan\alpha$ value distributions are negatively shifted, and for positive value is vice versa.

The matrix element of the process depends on the initial and final state particles, and it can be written as,

$$\mathcal{M}(p_{p_1}^\mu, p_{p_2}^\mu, p_{k_1}^\mu, p_{k_2}^\mu, p_H^\mu) = \sum_{f_1f_2 \to f_3f_4} \mathcal{M}(p_{p_1,f_1}^\mu, p_{p_2,f_2}^\mu, p_{k_1,f_3}^\mu, p_{k_2,f_4}^\mu, p_H^\mu)$$
$$\times \ F(x_1, f_1)F(x_2, f_2), \tag{2.33}$$
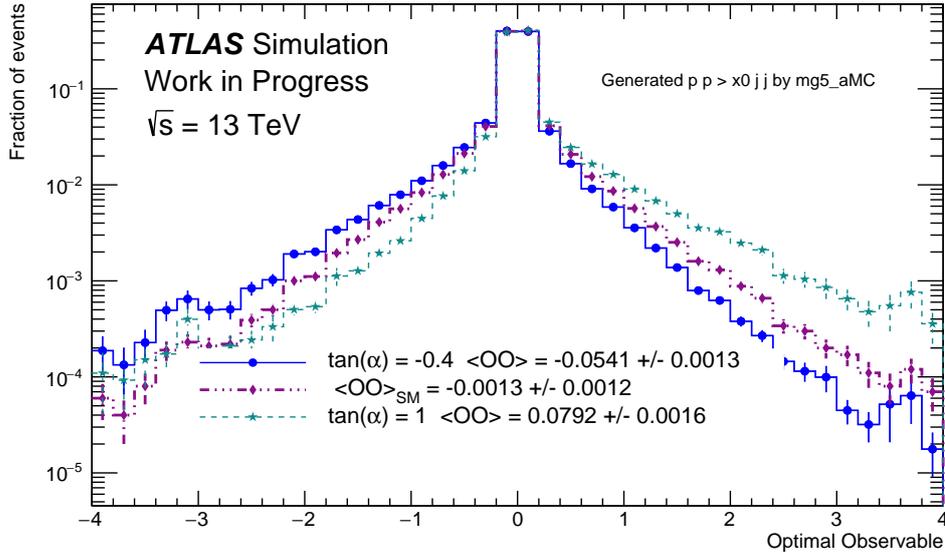
FIGURE 2.10: Normalized distibutions of the Optimal Observable $\mathcal{OO}$ for different mixing angles $\tan\alpha$ values with the statistical uncertainty. The dashed violet line indicates the SM case where $\tan\alpha = 0$, the solid blue line indicates the CP-mixing state with $\tan\alpha = -0.4$, and dashed green line illustrates the $\tan\alpha = 1$ distributions.

where sum runs over all combinations of the incoming and outgoing partons. The $p^{\mu}_{p_{1,2}}$, $p^{\mu}_{k_{1,2}}$ represents the four momentum of incoming and outgoing particles, and p^{μ}_{H} is the four momentum of the Higgs boson. Here, $F(x_i, f_i)$ is the parton distribution function, which i-th parton with the flavour f_i carrying the momentum fraction x_i of the proton. The matrix element calculation is extracted from MadGraph5 at LO using CT10 PDF set [43] with LHAPDF package [44].

Outgoing partons are measured as color neutral jets in the detector and kinematics of the incoming partons are not available. In order to calculate the matrix element of the measured events, initial state momentum vectors are build via

$$p^{\mu}_{p_{1/2}} = x_{1/2} \frac{\sqrt{s}}{2} (1, 0, 0 \pm 1), \tag{2.34}$$

where the momentum fractions $x_{1/2}$ can be calculated from the final state by

$$x_{1/2} = \frac{m_{\text{final}}}{\sqrt{s}} e^{\pm y_{\text{final}}}, \tag{2.35}$$

where $\sqrt{s}$ is the center of mass energy and m_{final} is the invariant mass of the final state of Higgs boson and two jets and y_{final} is the rapidity of the final state of Higgs boson and two jets.

Chapter 3

Signal and Background Processes

This analysis exploits the gluon fusion production mode (ggF) where the Higgs boson in produced in association with two jets. In particular, decay of the Higgs boson H$\to \tau_{\text{lep}}\tau_{\text{had}}$ is considered. The gluon fusion production (ggF) is the dominant production mode (see Section 2.2.1) at the LHC due to the contribution of the gluons to proton parton density function (PDF). The Higgs boson mass is considered to be $m = 125$ GeV with a lifetime of 1.56×10^{-22} s. The Higgs boson decays into a pair of τ-leptons 6.27% [25] of the time. The τ-leptons have a short lifetime of approximately 2.9×10^{-13} s and a mass of 1.78 GeV [25]. The τ-lepton decay is depicted in Fig. 3.1: 64.79 % [25] of tau-leptons decay hadronically and 35.21 % leptonically [25]. The τ-lepton decay is shown in Fig. 3.1. The probability for the fully leptonic di-tau decay is 12 %, for the semileptonic di-tau decay is 46 %, and for the fully hadronic di-tau decay is 42 % [25].

$$\nu_\tau$$

$$\tau^- \qquad \overline{\nu}_e/\overline{\nu}_\mu/\overline{u}$$

$$W^-$$

$$e^-/\mu^-/d$$
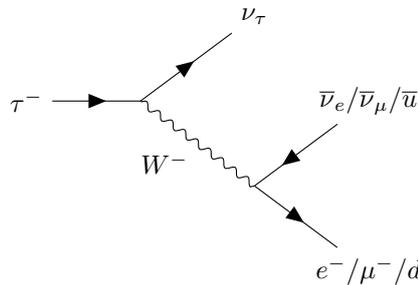
FIGURE 3.1: Feynman diagram for τ-lepton decay

When protons collide, numerous particles are produced. When investigating a specific process, the other physics processes need to be rejected in order to achieve a high signal purity. Therefore, this chapter discusses the signal and the background processes for this analysis, the test of CP-invariance of ggF H+2 jets production. The final state particles are one hadronically decaying τ-lepton, one light lepton (electron or muon), missing transverse energy and 2 jets. Finally, Monte Carlo generators for the simulation of signal and background samples, as well as the cross sections are briefly reviewed.

3.1 Signal process

The signal process considered in this analysis is the ggF Higgs boson production with two associated jets. The Higgs boson is dominantly produced via a top quark (a small contribution from a b-quark is found to be negligible) loop in the ggF. The top quark is massive, $m_{\text{top}} = 172.5$ GeV, but in this analysis, the top quark mass is approximated to infinite, $m_{\text{top}} \to \infty$, by an effective coupling vertex of Higgs boson to
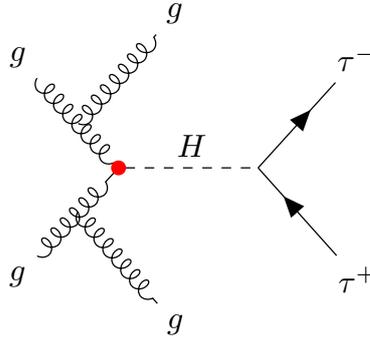
FIGURE 3.2: Example of Feynman diagram for the signal process, ggF
H+2 jets, are shown. The Higgs boson decays into τ-leptons. The red
dot symbolizes the effective coupling of Higgs boson to gluon vertex
in the limit of $m_{\text{top}} \to \infty$.

gluons. The related Feynman diagram for the signal process, ggF H+2 jets, is drawn
in Fig. 3.2. The final state signature is exactly one light lepton (e/μ), one hadroni-
cally decaying τ-lepton with an opposite charge, and two jets. Three neutrinos arise
from the decay of the τ-leptons. In the final state of the gluon fusion, there have
no additional partons produced at the leading order (LO). Although, additional par-
ton(s) are produced by next-to-leading order (NLO), higher-order QCD corrections
due to initial state partons. The signal process is classified into three different cate-
gories depending on the initial state partons. The categories are the gluon-gluon, the
gluon-quark, and the quark-quark initiated processes. These processes are discussed
in detail in Sec. 4.3.

3.2 Background processes

Background processes, in general, are classified into two different categories: re-
ducible background and irreducible background. The irreducible background has a
same final state particles as signal process whereas reducible background has at least
one final state object is misidentified e.g., jet as a hadronically decaying τ-leptons.
Selection criteria can suppress the reducible background. For the ggF H+2 jets analy-
sis, the dominant irreducible background is the Z-boson production associated with
jets, where the subsequent decay of Z-boson is a pair of τ-leptons. The following
sections introduce the main background processes.

3.2.1 Higgs boson background processes

There are four different production modes for the Higgs boson, as discussed earlier
in Sec. 2.2.1. In order to have a well-distinguished signal, other production modes
of the Higgs boson need to be suppressed because we would like to investigate the
anomalous coupling of the Higgs boson in the ggF production mode. The related
Feynman diagrams of these processes are shown in Fig. 3.3; 3.4. As shown in Fig.
3.3, Higgs boson is produced by vector bosons fusion (VBF). From the initial state
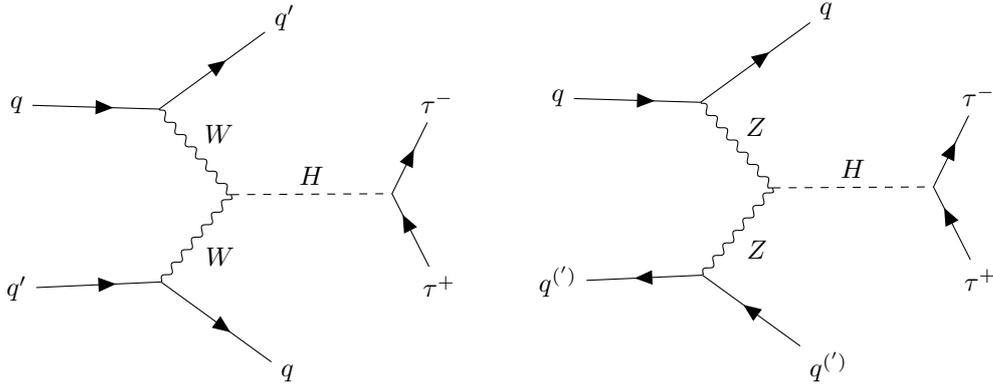
FIGURE 3.3: Example of Feynman diagrams of the Higgs boson in vector-boson fusion with subsequent Higgs boson decay into a pair of tau leptons. The left diagram shows the W-boson fusion, and the right one is the Z-boson fusion processes.

quarks, vector bosons are radiated, and further fusion of the bosons produce the Higgs boson. In the VBF production, outgoing jets have a large difference of the pseudorapidity ($\Delta \eta_{jj}$) and have a large invariant mass m_{jj} in comparison to ggF.

The Higgs boson produced in association with a weak vector boson (VH) is known as Higgs-strahlung. When weak vector boson decay into hadronically $V \rightarrow q\bar{q}^{(')}$, this VH background contributes to the irreducible backgrounds. Feynman diagrams of W/Z-boson associated production are shown in Fig. 3.4.
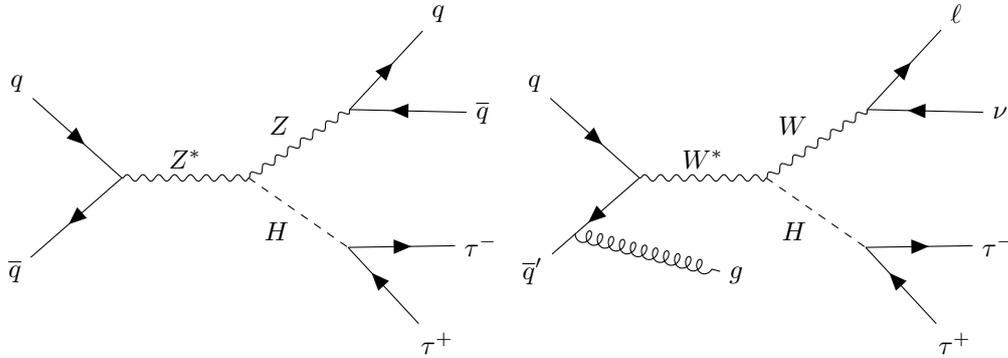
FIGURE 3.4: Example of Feynman diagrams of the associated W/Z-boson (VH) to the Higgs boson are illustrated. The W/Z-boson decays are in the left, hadronically, and in the right, leptonically.

Another production mode of the Higgs boson is the associated production with a pair of top quark ($t\bar{t}H$). The cross section of this process is small as discussed in Sec. 2.2.1; therefore, it is found to be negligible.

3.2.2 W/Z-boson production associated with jets

Vector boson production in association with jets gives the dominant background contribution to the ggF H+2 jets analysis. It produces both reducible and irreducible background contributions. The irreducible background is the Z-boson (or virtual photon γ^*) production decaying into a pair of τ-leptons ($Z/\gamma^* \rightarrow \tau^-\tau^+ \rightarrow \tau_{\text{had}} +$

$\ell + 3\nu$). Feynman diagrams of the Z boson production without jet, with one jet, and with two jets are shown in Fig. 3.5, left to the right, respectively. A small contribution

(a) Z-boson   (b) Z+jet   (c) Z+ 2jets
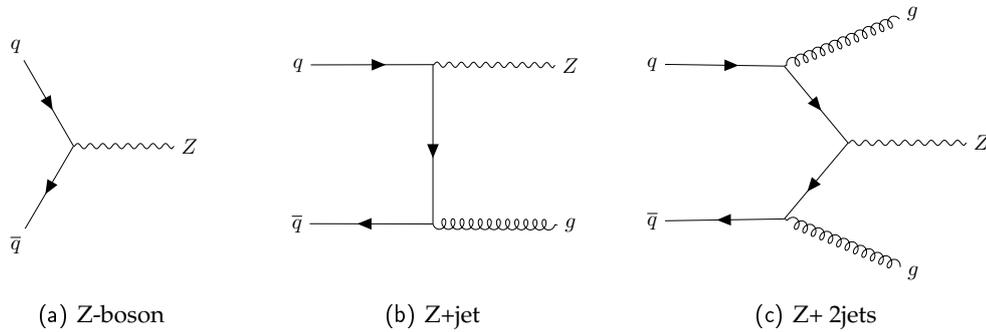
FIGURE 3.5: Example of Feynman diagrams that illustrate the Z-boson production without jets, with one jet, and with two jets, respectively, are shown.

to the background arises from $Z \to \ell\ell$ production. The Z-boson can be produced association with jets as shown in the Feynman diagrams Fig. 3.5. The jets can be misidentified as a hadronically decaying τ-lepton. Also, it could be one lepton is not detected and one jet is misidentified as a hadronically decaying τ-lepton. As same

(a) W-boson   (b) W+jet   (c) W+2 jets
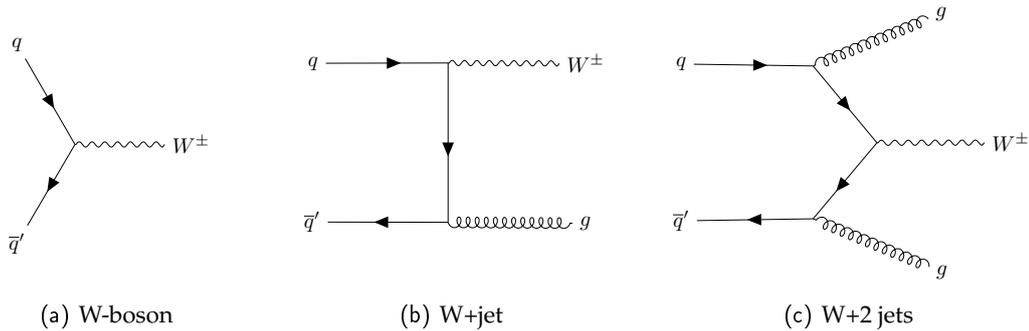
FIGURE 3.6: Example of Feynman diagrams of the W-boson production associated with jets.

as Z-boson and its decay, Feynman diagrams of W-boson production with jets are illustrated in Fig. 3.6. This production of W+jets contributes the background when W-boson decays leptonically, an associated jets can be misidentified as a hadronically decaying τ-lepton. This W+jets background gives the dominant contribution to the jets are misidentified as a hadronically decaying τ-leptons. This background is reducible, and it can be suppressed by the selection criteria like the transverse momentum of the jets.

3.2.3 Diboson production

Diboson background contributions are the production of a pair of W-boson ($W^{\pm}W^{\pm}$), a pair of Z-boson (ZZ), and a pair of W and Z-bosons.
In some cases, these vector bosons couple to each other. This coupling is called Trilinear Gauge Boson Coupling (TGC) [45, 46]. Feynman diagrams of this process are shown in Fig. 3.7. In the production of a pair of W-boson, $WW \to \ell\nu q\bar{q}'$, jets are misidentified as a hadronically decaying τ-leptons. Moreover, the Z-boson decays either leptonically or (two neutrinos, but this is not the case in this analysis)
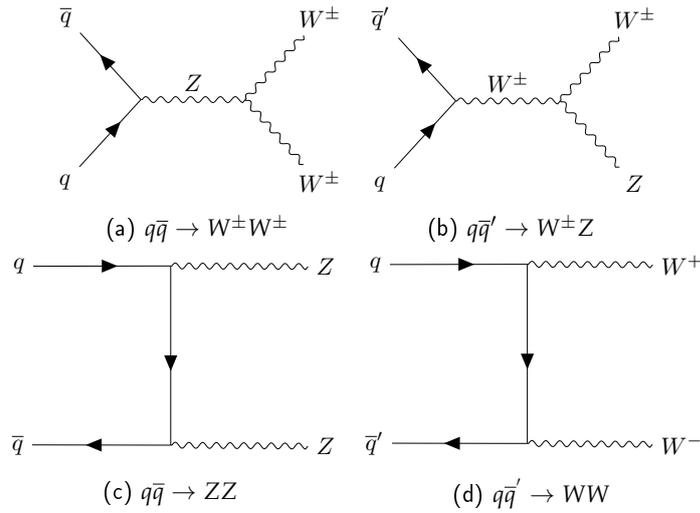
FIGURE 3.7: Example of Feynman diagrams for diboson productions.

or hadronically. Therefore, in the ZZ production, one Z boson decays into a pair of τ-leptons and another one decays into a pair of quarks. Thus, this production can have the same final state signature as the signal.

3.2.4 Top-quark production

The top-quark background includes the single top-quark production and production of a top quark pair. The top quark decays to a W-boson and a b-quark. The
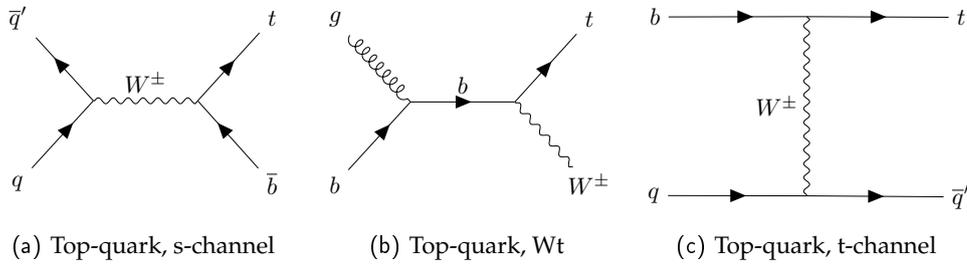
FIGURE 3.8: Example of Feynman diagrams of the single top-quark productions.
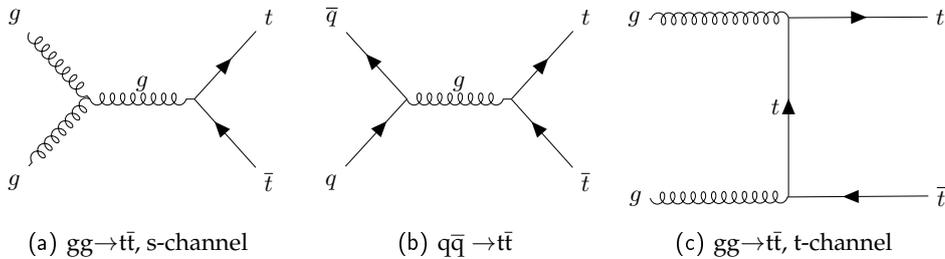
FIGURE 3.9: Example of Feynman diagrams of pair of top-quark production. From the left to the right, top-quark pair produced from gluon and quarks.

Feynman diagrams of the single top-quark production are depicted in Fig. 3.8. This

background is reducible, and it can be suppressed by the b-jet tagging method. W-boson decays can also be misidentified as a hadronically decayed τ-lepton. However, misidentified jets does not need to be originated from W-boson. Moreover, leptonic decay of the W-boson and misidentified jets have the same final state like the signal.

Most of the contribution arises from a pair of top-quark production. The Feynman diagrams of these processes are shown in Fig. 3.9.

3.2.5 QCD multi-jet production

The cross section for multi-jet production is huge at the LHC. If at least four jets are produced, and one identified as electron or muon and one as a hadronically decaying τ-lepton. It contributes to the reducible backgrounds.
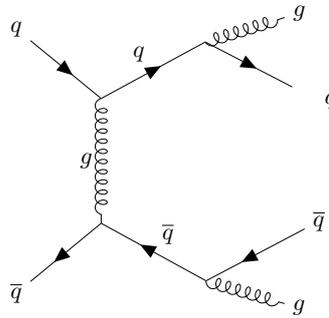
FIGURE 3.10: Example of Feynman diagram for the multi-jet production.

In Fig. 3.10 example of Feynman diagram is shown. This multi-jet background is suppressed by the selection criteria by identification of electron and muon and hadronically decaying τ-lepton identification. And missing transverse energy selection criteria can reduce this QCD multi-jet production.

3.3 Simulated event samples

Simulated event samples of the signal and background are generated by multiple Monte Carlo (MC) event generators. The signal ggF H+2 jets samples are simulated by MC generator using the Powheg method [47–50] with the NNLOPS program [51] at Quantum Chromodynamics (QCD) next-to-leading order (NLO) accuracy. The cross section of the events of the ggF H+1 jet is calculated at next-to-next-to-leading order (NNLO) in QCD. Whereas the cross section of the ggF H+2 jets is calculated at NLO in QCD using the MiNLO approach [52].

All Higgs boson production processes are simulated by Powheg interfaced with Pythia 8.212 [53] for the modeling of parton showering, hadronization, and underlying event (UEPS). Herwig 7.0.3 [54] is used to estimate the impact of uncertainties for the Higgs boson production processes on the UEPS. The parton distribution function (PDFs), PDF4LHC15 [55], is used for all Higgs boson production modes. The overall cross section for ggF process is calculated at N^3LO in QCD with NLO EW corrections [56–59]. Moreover, the cross section of VBFH is calculated at NNLO accuracy in QCD with NLO EW corrections [60–62], and the VH process cross section calculated at NNLO in QCD with NLO EW corrections [63–65] as well.

The diboson and the W/Z-boson background are simulated with the Sherpa 2.2.1 [66] generator. The matrix element of Z-boson production associated with up to 2 jets is calculated at NLO in QCD, whereas Z+3/4 jets is calculated at LO accuracy. The inclusive cross section of Z boson production associated with jets is calculated at NNLO using FEWZ [67, 68]. The NNPDF30NNLO PDF set [69] is used for the diboson and W/Z-boson+jets processes. Top-quark background is produced with Powheg interfaced with Pythia 6.428 [70] for UEPS. The cross section of this background is estimated at NNLO in QCD [71–75] for $t\bar{t}$ using the Top++2.0 program [76]. The cross section for $W\bar{t}$ production is calculated at NLO. CT10 [43] is used for top quark production PDF set.

The Monte Carlo generators and cross sections of the simulated samples are presented in Table 3.1 and Table 3.2. From Table 3.1, one can see the cross section of the signal process and background of Higgs boson production in different modes. Also, the cross section of diboson and top quark productions are summarized with the respective MC generators. In Table 3.2, inclusive cross sections of the dominant background $Z/\gamma^* \to \tau\tau$ in association with jets and $Z/\gamma^* \to \ell\ell$ are given.

Event Sample	MC generator	UEPS	Cross section [pb]
ggF H $\to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	Powheg	Pythia 8.212	0.461
VBF H$\to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	Powheg	Pythia 8.230	0.038
$W^{\pm}$H, H $\to \tau\tau$	Powheg	Pythia 8.212	0.085
ZH, H$\to \tau\tau$	Powheg	Pythia 8.212	0.047
Diboson background			
ZZ$\to$ qq$\nu\nu$	Sherpa 2.2.1	Sherpa 2.2.1	4.354
ZZ$\to$ qq$\ell\ell$	Sherpa 2.2.1	Sherpa 2.2.1	2.172
WZ$\to$ qq$\nu\nu$	Sherpa 2.2.1	Sherpa 2.2.1	6.797
WZ$\to$ qq$\ell\ell$	Sherpa 2.2.1	Sherpa 2.2.1	3.437
WZ$\to$ $\ell\nu$qq	Sherpa 2.2.1	Sherpa 2.2.1	11.413
$W^{\pm}W^{\pm} \to$ qq$\ell\nu$	Sherpa 2.2.1	Sherpa 2.2.1	24.717/24.734
Top-quark background			
$t\bar{t}$	Powheg	Pythia 6.428	452.36
Single t/$\bar{t}$, t-channel	Powheg	Pythia 6.428	44.154/26.275
Wt/W$\bar{t}$	Powheg	Pythia 6.428	34.849/35.846
Single t/$\bar{t}$, s-channel	Powheg	Pythia 6.428	2.061/1.288

TABLE 3.1: Signal and background Monte Carlo generators with underlying event, hadronization and parton showering model (UEPS) are described. As well as the inclusive cross section of the standard model process, H $\to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ at $\sqrt{s} = 13$ TeV.

The W+jets and QCD backgrounds are estimated using a data-driven method. The estimation of reducible background contributions, where a jet is misidentified as a hadronically decaying τ-lepton is discussed shortly in Sec. 7.2. The simulation of the ATLAS detector, which describes detector material and detector response is performed using the GEANT4 [77].

Event Sample	Cross section [pb]
Z+4 jets, $Z/\gamma^* \to \tau\tau$, Sherpa 2.2.1	
Z+jets EW	0.633
$0 < \max(HT, pTV) < 70$ GeV	103.012
$70 < \max(HT, pTV) < 140$ GeV	7.747
$140 < \max(HT, pTV) < 280$ GeV	3.709
$280 < \max(HT, pTV) < 500$ GeV, c&b veto	4.672
$280 < \max(HT, pTV) < 500$ GeV, b-veto, c-filter	2.219
$280 < \max(HT, pTV) < 500$ GeV, b-filter	1.465
$500 < \max(HT, pTV) < 1000$ GeV	1.764
$1 < \max(HT, pTV) < 13$ TeV	0.144
$10 < m_{\ell\ell} < 40$ GeV, $0 < \max(HT, pTV) < 70$ GeV b-veto	2275.811
$10 < m_{\ell\ell} < 40$ GeV, $0 < \max(HT, pTV) < 70$ GeV b-filter	79.083
$10 < m_{\ell\ell} < 40$ GeV, $70 < \max(HT, pTV) < 280$ GeV b-veto	43.659
$10 < m_{\ell\ell} < 40$ GeV, $70 < \max(HT, pTV) < 280$ GeV b-filter	5.4029
$10 < m_{\ell\ell} < 40$ GeV, $0.28 < \max(HT, pTV) < 13$ TeV b-veto	2.724
$10 < m_{\ell\ell} < 40$ GeV, $0.28 < \max(HT, pTV) < 13$ TeV b-filter	0.458

Z+4 jets, $Z/\gamma^* \to \ell\ell$, Sherpa 2.2.1		
	ee	$\mu\mu$
Z+2jets EW	0.632	0.634
$0 < \max(HT, pTV) < 70$ GeV, c&b veto	1586.660	1589.631
$0 < \max(HT, pTV) < 70$ GeV, c-filter, b-veto	218.160	218.146
$0 < \max(HT, pTV) < 70$ GeV, b-filter	123.301	124.013
$70 < \max(HT, pTV) < 140$ GeV, c&b veto	74.393	73.148
$70 < \max(HT, pTV) < 140$ GeV, c-filter, b-veto	19.829	19.841
$70 < \max(HT, pTV) < 140$ GeV, b-filter	12.308	12.080
$140 < \max(HT, pTV) < 280$ GeV, c&b veto	24.406	23.680
$140 < \max(HT, pTV) < 280$ GeV, c-filter, b-veto	9.138	9.044
$140 < \max(HT, pTV) < 280$ GeV, b-filter	5.931	5.688
$280 < \max(HT, pTV) < 500$ GeV, c&b veto	4.748	4.654
$280 < \max(HT, pTV) < 500$ GeV, c-filter, b-veto	2.223	2.209
$280 < \max(HT, pTV) < 500$ GeV, b-filter	1.457	1.454
$500 < \max(HT, pTV) < 1000$ GeV	1.763	1.743
$1 < \max(HT, pTV) < 13$ TeV	0.145	0.144
$10 < m_{\ell\ell} < 40$ GeV, $0 < \max(HT, pTV) < 70$ GeV b-veto	2273.175	2272.164
$10 < m_{\ell\ell} < 40$ GeV, $0 < \max(HT, pTV) < 70$ GeV b-filter	79.332	80.208
$10 < m_{\ell\ell} < 40$ GeV, $70 < \max(HT, pTV) < 280$ GeV b-veto	43.851	43.761
$10 < m_{\ell\ell} < 40$ GeV, $70 < \max(HT, pTV) < 280$ GeV b-filter	5.345	4.987
$10 < m_{\ell\ell} < 40$ GeV, $0.28 < \max(HT, pTV) < 13$ TeV b-veto	2.708	2.691
$10 < m_{\ell\ell} < 40$ GeV, $0.28 < \max(HT, pTV) < 13$ TeV b-filter	0.461	0.460

TABLE 3.2: Inclusive cross section of the Z/γ^*+jets background with Monte Carlo generator. The cross section of Z/γ^* with up to 2 jets calculated at NLO and 3, 4 jets calculated at LO at $\sqrt{s} = 13$ TeV.

Chapter 4

Study of anomalous couplings of the Higgs boson to gluons at generator level

It is essential to study the CP nature of the Higgs boson by looking for anomalous coupling structure of the Higgs boson in the gluon fusion production. In this chapter, the effective Higgs boson coupling to gluons is investigated in the Higgs boson production in gluon fusion in association with with two jets, ggF H+2 jets events, at generator level. The generator level refers to events produced with no parton shower nor hadronization, and without detector effects. Background processes and systematic uncertainties are not considered in this study.

The effective Lagrangian for Higgs boson interaction with gluons can be found in Sec. 2.3. In order to test the CP-invariance, the relative contributions of CP-even and CP-odd interactions are important. The mixing angle $\tan\alpha$ parametrizes the relative contributions of CP-even and CP-odd interactions.

For this, CP-odd observables as introduced in Sec. 2.4, are used to study the anomalous couplings structure of the Higgs boson to gluons. The mean values of the CP-odd observables for various $\tan\alpha$ coupling models can be used to illustrate the sensitivity of the observable to CP-violating effects. In the Standard Model (SM), the CP-symmetry is conserved in the Higgs boson interactions; the mean values of the CP-odd observables have to be compatible with zero. Hence, nonzero mean values indicate the violation of CP-invariance.

In order to study the $\tan\alpha$ dependence, a matrix element based reweighting technique is used in this analysis to generate signal prediction for non-vanishing $\tan\alpha$ values. This chapter introduces the reweighting technique and shows the sensitivity to anomalous couplings in ggF H+2 jets events at the generator level. Furthermore, the sensitivity of the subprocesses contributing ggF H+2 jets production is studied. The subprocesses are categorized depending on the flavor of incoming and outgoing particles in ggF H+2 jets.

4.1 Reweighting technique

Various coupling models corresponding to different $\tan\alpha$ values are desired in order to investigate the Higgs boson coupling structure to gluons. The simulation of these signal samples are not efficient. Instead, a reweighting technique can be applied. It uses the matrix element of a particular process to calculate the event weight to a specific $\tan\alpha$ value and it is expressed by

$$\mathcal{W}(\tan\alpha) = \frac{|\mathcal{M}(\tan\alpha)|^2}{|\mathcal{M}_{SM}|^2}. \tag{4.1}$$

Here $\mathcal{M}(\tan\alpha)$ is the matrix element of the process of interest, and $\mathcal{M}_{SM} = \mathcal{M}(\tan\alpha = 0)$ where corresponds to the SM interaction. The reweighting technique uses the truth level four-momentum vectors and flavours of the incoming and outgoing particles as input information [23].

Variable	Tau [Gev]	Jets [Gev]
p_T^{min}	10	20
$|\eta|$	2.5	5
$\Delta R_{min}(\tau\tau(jj))$	0.4	0.4

TABLE 4.1: Threshold of kinematic variables for jets and τ-leptons in the generator of leading order (LO) ggF H+2 jets events in Mad-Graph5 [39]. The τ-leptons are considered to be stable.

ggF H+2 jets signal events are generated at leading order (LO) with MadGraph5 [39] using the CT10 PDF set [43] within at the Higgs-Characterization model [37]. The matrix element of the processes at LO is extracted from the MadGraph5 [39] for calculating the weight used in reweighting technique. The center of mass energy is $\sqrt{s} = 13$ TeV, and the analysis considers a four-flavor scheme (up, down, strange, charm quarks) without parton shower. As discussed in Chapter 3, the subsequent decay of the Higgs boson is the semileptonic a pair of τ-leptons is considered. However, for this sensitivity study on the anomalous couplings of the Higgs boson to gluons, the decay channel of the Higgs boson is not relevant. Table 4.1 summarizes the threshold value of the kinematic variables for τ-lepton and jets applied in the event generator.

By applying the reweighting technique, reweighted distributions are compared to the directly generated events by MadGraph5 [39] for certain $\tan\alpha$-models.

For the validation, the distributions of CP-odd observables and their corresponding mean values are compared for $\tan\alpha = 1$ for LO reweighted and LO generated events. A comparison of the distributions of the CP-odd observables is shown in Fig. 4.1.

As illustrated in Fig. 4.1a; 4.1b; 4.1c, the distributions of $\Delta\Phi_{jj}^{\text{ordered}}$; $\Delta\Phi_{jj}^{\text{sign}}$; $\mathcal{OO}$ observables respectively have a good agreement between reweighted and generated events. In addition, the mean values are compared in Table 4.2 and they are compatible with each other within statistical uncertainties.

The signal sample ggF H+2 jets can be grouped into gluon-gluon (gg), gluon-quark (gq) and quark-quark (qq) initial states depending on their combination of incoming parton flavors. Therefore, it is also important to validate the reweighting

(a)

(b)

(c)
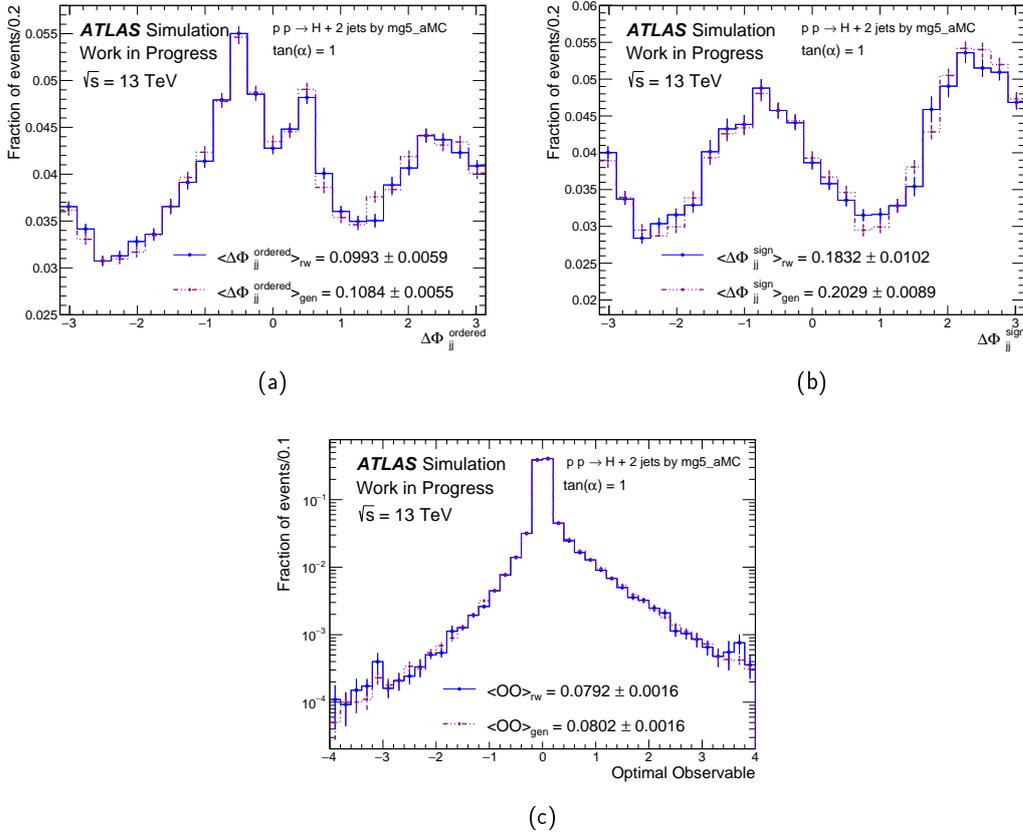
FIGURE 4.1: Comparison of LO reweighted (rw) and LO generated (gen) events for $\tan\alpha = 1$ for the normalized distributions of the following CP-odd observables: (a) $\Delta\Phi_{jj}^{\text{ordered}}$, (b) $\Delta\Phi_{jj}^{\text{sign}}$, and (c) $\mathcal{OO}$. Here, events are generated by MadGraph5 [39]. Only statistical uncertainties are shown. Reweighted events are drawn a solid blue line, and generated events are drawn as purple dashed lines in all figures.

method for the different subprocesses of the ggF H+2 jets seperately. The comparison for the different distributions of subprocesses at $\tan\alpha = 1$ between reweighted and generated events are shown in Fig. 4.2; 4.3; 4.4.

The $\Delta\Phi_{jj}^{\text{sign}}$ distributions in each subprocess are shown in Fig. 4.2 for reweighted and generated events at $\tan\alpha = 1$. As seen in Fig. 4.2a, the mean value of the $\Delta\Phi_{jj}^{\text{sign}}$ in the gg initial state shows some fluctuations. The events are reduced by 58.4 % when requiring the outgoing jets to be in opposite hemispheres. Hence, reducing the statistics, reweighted and generated events for $\Delta\Phi_{jj}^{\text{sign}}$ give some fluctuation. The gq initial, and qq initial subprocesses in Fig. 4.2b, Fig. 4.2c respectively, show good agreement between reweighted and generated events within the statistical uncertainties.

Since the jet selection in $\Delta\Phi_{jj}^{\text{sign}}$ reduces number of the events, another approach for the jets is studied. The outgoing jets are ordered by their pseudorapidity η-value. This $\Delta\Phi_{jj}^{\text{ordered}}$ distributions for subprocesses in reweighted and generated events are shown in Fig. 4.3. From the distributions, one can see the agreement between reweighted and generated events. The mean values are compatible within uncertainties.

(a)

(b)

(c)
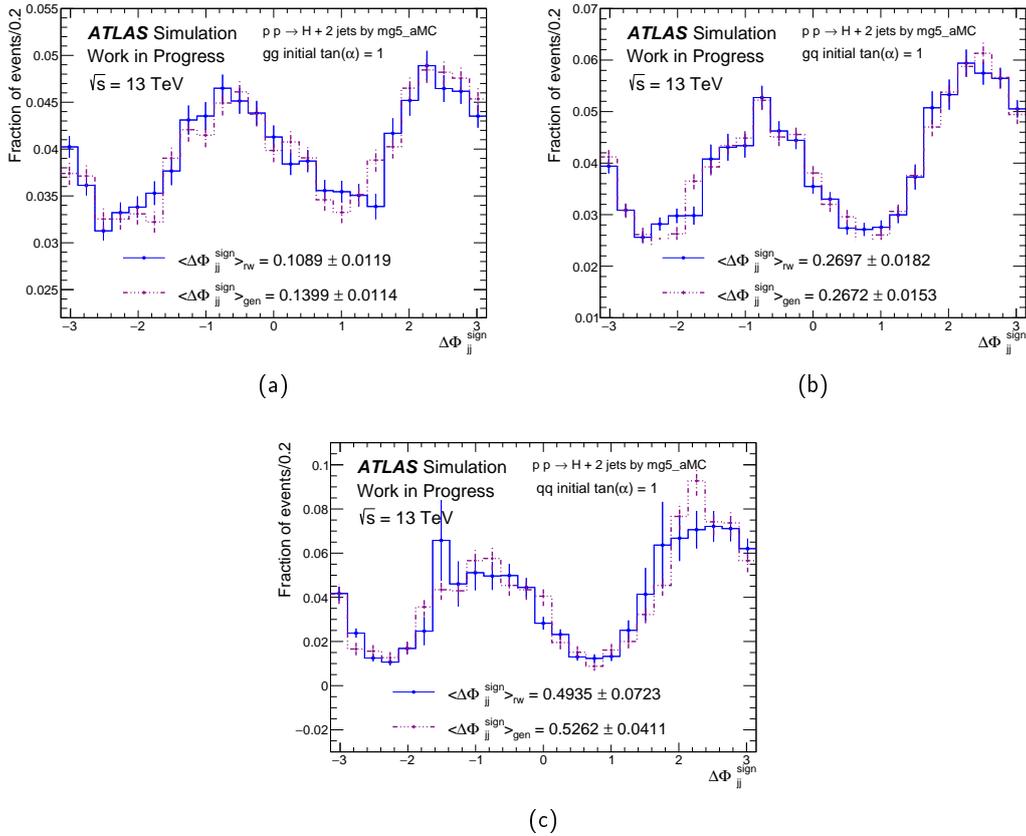
FIGURE 4.2: Comparison of reweighted (rw) and generated events (gen) at $\tan \alpha = 1$ for $\Delta \Phi_{jj}^{\text{sign}}$ in (a) gluon-gluon (gg), (b) gluon-quark (gq), and (c) quark-quark (qq) initiated processes for ggF H+2 jets events. The distributions are normalized and only statistical uncertainties are shown.

The $\mathcal{OO}$ distributions for different subprocesses as shown in Fig. 4.4, also show good agreement between reweighted and generated events. In addition, the corresponding mean values at $\tan \alpha = 1$ for reweighted and generated events are compatible with each other within uncertainties.

These results allow to validate the LO reweighting method for LO generated ggF H+2 jets events.

However, as discussed in Sec. 3.1, this analysis uses signal events, which are generated at next-to-leading order (NLO) with the Monte Carlo event generator Powheg [47–50]. Therefore, it is examined how well the LO reweighting method works, when it is applied to NLO events. The events of the NLO sample are produced with the same threshold value for kinematic variables as LO Madgraph event samples as summarized in Table 4.1. Hence, we can compare the mean values of the CP-odd observables, as shown in Table 4.2.

From Table 4.2, one can see the LO reweighting comparison for LO ggF H+2jets produced with MadGraph5 and NLO ggF H+2 jets sample produced with Powheg for different CP-odd observables. There are deviations between the CP-odd observable mean values for LO and NLO samples when the LO reweighting method is used. In principle, the reweighting method at NLO order is wanted, but it is not available. The LO order reweighting method does not consider the effect of NLO
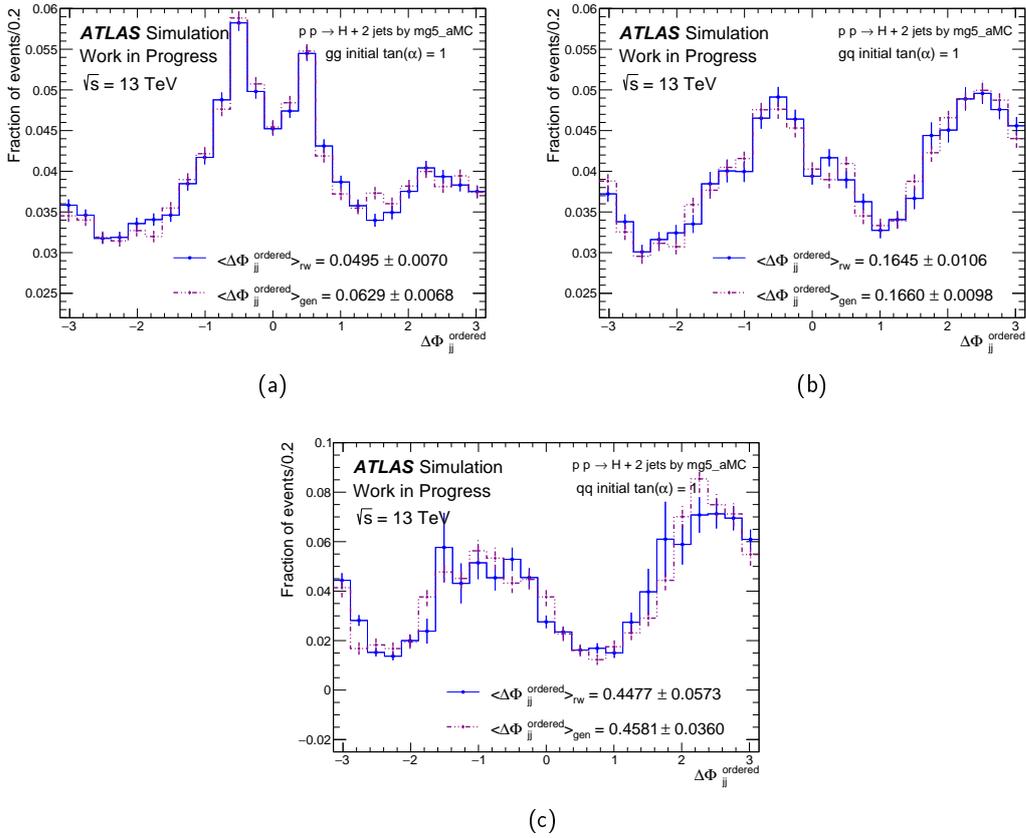
FIGURE 4.3: Comparison of reweighted (rw) and generated events (gen) at $\tan\alpha = 1$ for $\Delta\Phi_{jj}^{\text{ordered}}$ in (a) gluon-gluon (gg), (b) gluon-quark (gq), and (c) quark-quark (qq) initiated processes for ggF H+2 jets events. The distributions are normalized and only statistical uncertainties are shown.

corrections. However, as the application of NLO reweighting is not trivial, it could not be developed in the context of this thesis.

(a)

(b)

(c)
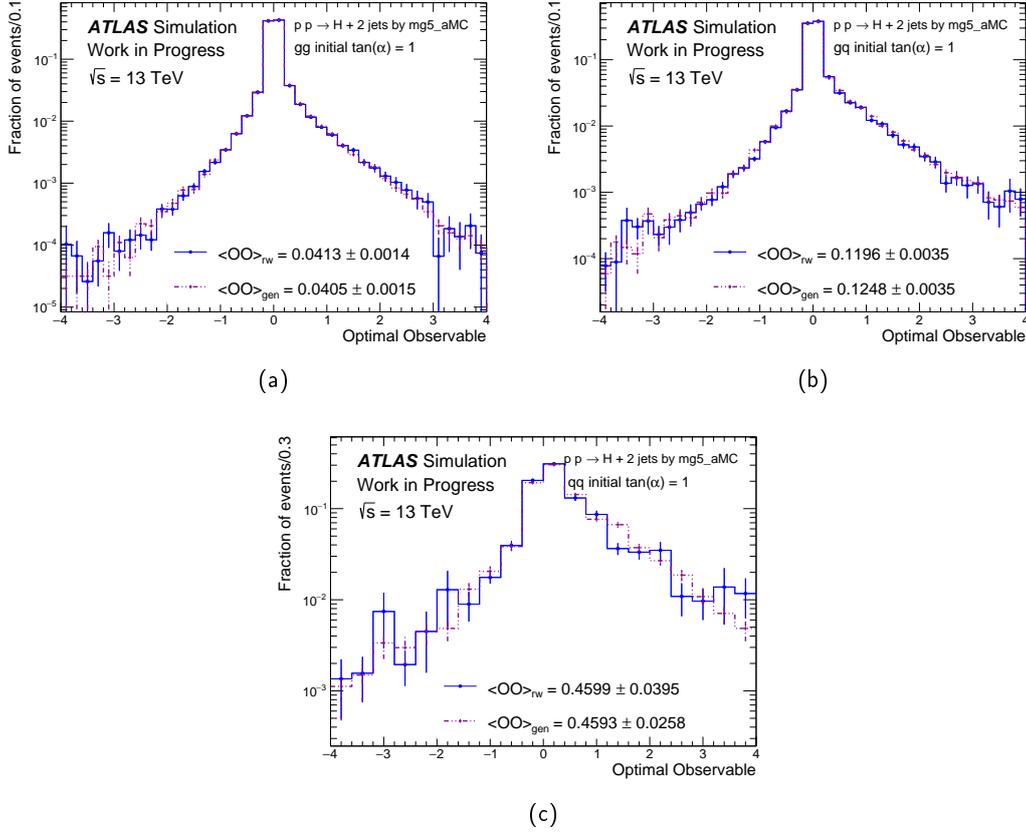
FIGURE 4.4: Comparison of reweighted (rw) and generated events (gen) for $\mathcal{OO}$ in (a) gluon-gluon (gg), (b) gluon-quark (gq), and (c) quark-quark (qq) initiated processes for ggF H+2 jets events. The distributions are normalized, and only statistical uncertainties are shown.

ggF H+2jets		MadGraph5 LO		Powheg NLO
$\tan\alpha$	$\mathcal{OO}$	Generated	LO reweighted	
1	all	0.0802 ± 0.0017	0.0792 ± 0.0016	0.0789 ± 0.0011
	gg Initial	0.0405 ± 0.0015	0.0413 ± 0.0014	0.0617 ± 0.0010
	gq Initial	0.1248 ± 0.0035	0.1196 ± 0.0035	0.0801 ± 0.0026
	qq Initial	0.4593 ± 0.0258	0.4599 ± 0.0395	0.4631 ± 0.0175
$\tan\alpha$	$\Delta\Phi_{jj}^{\mathbf{sign}}$	Generated	LO reweighted	
1	all	0.2029 ± 0.0089	0.1831 ± 0.0102	0.1584 ± 0.0056
	gg Initial	0.1399 ± 0.0114	0.1089 ± 0.0120	0.1499 ± 0.0060
	gq Initial	0.2672 ± 0.0153	0.2697 ± 0.0182	0.1314 ± 0.0117
	qq Initial	0.5262 ± 0.0411	0.4935 ± 0.0723	0.4788 ± 0.0326
$\tan\alpha$	$\Delta\Phi_{jj}^{\mathbf{ordered}}$	Generated	LO reweighted	
1	all	0.1084 ± 0.0055	0.0993 ± 0.0059	0.0899 ± 0.0034
	gg Initial	0.0629 ± 0.0068	0.0496 ± 0.0070	0.0829 ± 0.0037
	gq Initial	0.1660 ± 0.0098	0.1645 ± 0.0106	0.0785 ± 0.0072
	qq Initial	0.4581 ± 0.0360	0.4477 ± 0.0573	0.3853 ± 0.0258

TABLE 4.2: Comparison of the mean values for different CP-odd observables between generated and reweighted events with $\tan\alpha = 1$.

4.2 Sensitivity study

In order to study the anomalous couplings between Higgs boson and gluons, this analysis uses the CP-odd observables. $\Delta\Phi_{jj}^{\text{sign}}$, $\Delta\Phi_{jj}^{\text{ordered}}$, $\mathcal{OO}$, which are introduced in Sec. 2.4. Furthermore, the sensivity to CP-violation can be studied by determining the mean values for various CP mixing states corresponding to $\tan\alpha$ values which deviate from zero (SM). A mean value of zero corresponds to CP-invariance, and it should be obtained for $\tan\alpha = 0$ in each CP-odd observable.

(a)

(b)

(c)
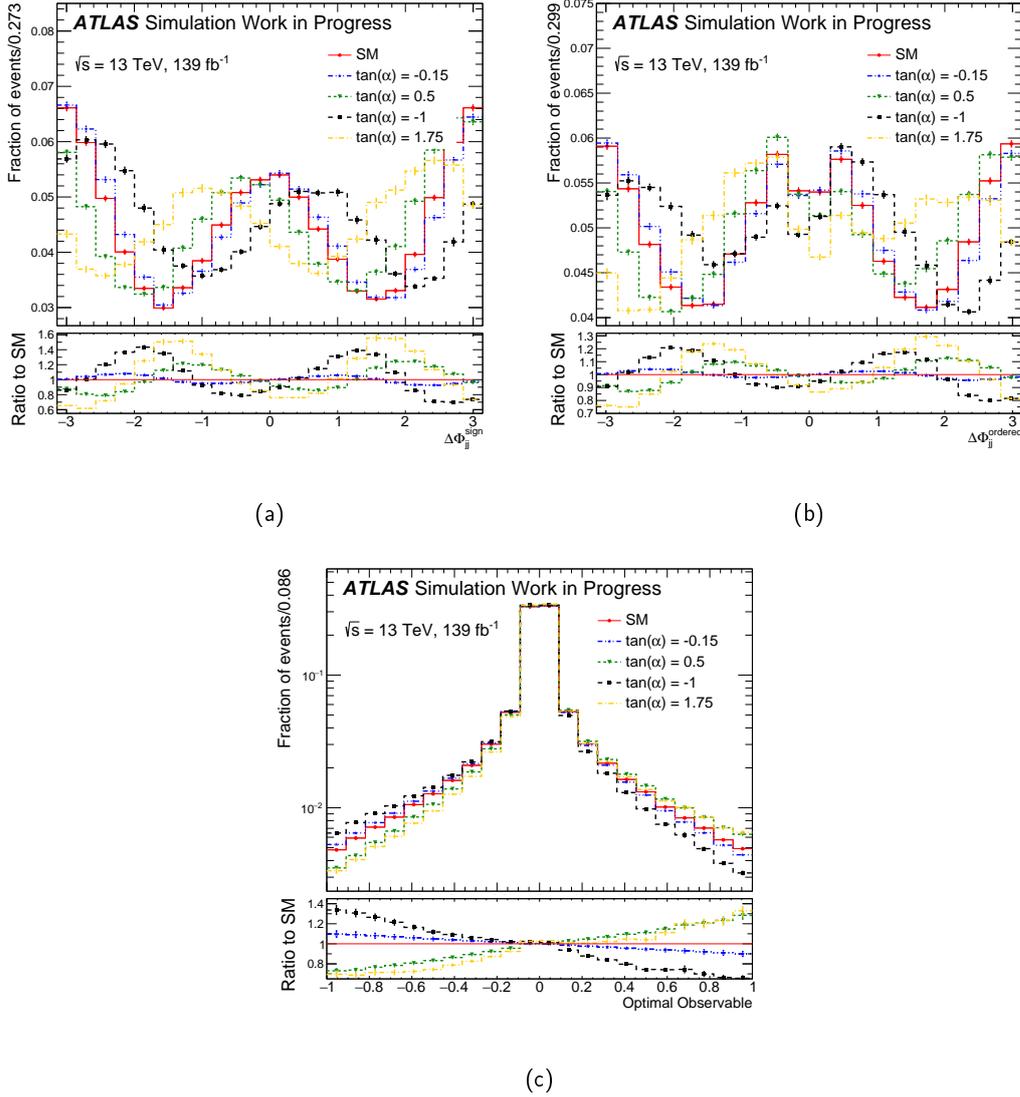
FIGURE 4.5: Normalized distributions of the CP-odd observables: (a) $\Delta\Phi_{jj}^{\text{sign}}$, (b) $\Delta\Phi_{jj}^{\text{ordered}}$, (c) Optimal Observable for various $\tan\alpha$ values. Here, NLO event distributions are reweighted at LO and only statistical uncertainties are shown. From the ratio plots, one can see that distributions are shifted to the left for negative $\tan\alpha$ values, while distributions are shifted to the right for positive $\tan\alpha$ values.

As discussed in the previous section, the reweighting technique allows to obtain signal events with various anomalous coupling models of the Higgs boson and gluons. The normalized distributions of the CP-odd observables for various $\tan\alpha$

values are shown in Fig. 4.5. The distributions of the observables are shifted to the left for negative $\tan\alpha$ values and shifted to the right for positive $\tan\alpha$ values. The mean value of the CP-odd observables as a function of various $\tan\alpha$ values are illustrated in Fig. 4.6. Here, the sensitivity to the effective couplings of the Higgs boson to gluons are compared for LO and NLO reweighted events.

Fig. 4.6a shows the $\langle\Delta\Phi_{jj}^{\text{sign}}\rangle$ dependency on $\tan\alpha$ for LO Madgraph5 [39] and NLO Powheg [47–50] events. Selecting the jets in opposite hemisphere reduces the number of events by 58.4 % and therefore enhances the statistical uncertainties. As shown in Fig. 4.6b, the sensitivity to anomalous couplings of Higgs boson to gluons is reduced due to the application of the LO reweighting method to the events generated at NLO.

In order to retain the full available statistics $\Delta\Phi_{jj}^{\text{ordered}}$ is used to study anomalous couplings as well, and it is shown in Fig. 4.6c; 4.6d for LO and NLO events, respectively.

Furthermore, the Optimal Observable mean value distributions for LO and NLO as a function of $\tan\alpha$ are shown in Fig. 4.6e; 4.6f.

All CP-odd observables allow to distinguish various models of $\tan\alpha$. For all CP-odd observables mean value distributions, it can be seen that negative mean values of the observables occur for negative $\tan\alpha$ values. For the positive mean values, it is vice versa. Moreover, the maximum mean values of the CP-odd observables are obtained for $\tan\alpha = 1$ corresponding to the same amount of CP-even and CP-odd contributions. To investigate which CP-odd observable is more sensitive to the anomalous coupling of the Higgs boson and gluons, the mean values of the observables are normalized to the root mean square (RMS) as shown in Fig. 4.7. Here, one can see that the Optimal Observable provides a better sensitivity to anomalous couplings ggF H+2 jets than the azimuthal angle difference between the outgoing jets. Furthermore, $\Delta\Phi_{jj}^{\text{sign}}$ shows larger deviations of the mean value from zero compared to the η-ordered $\Delta\Phi_{jj}^{\text{ordered}}$.
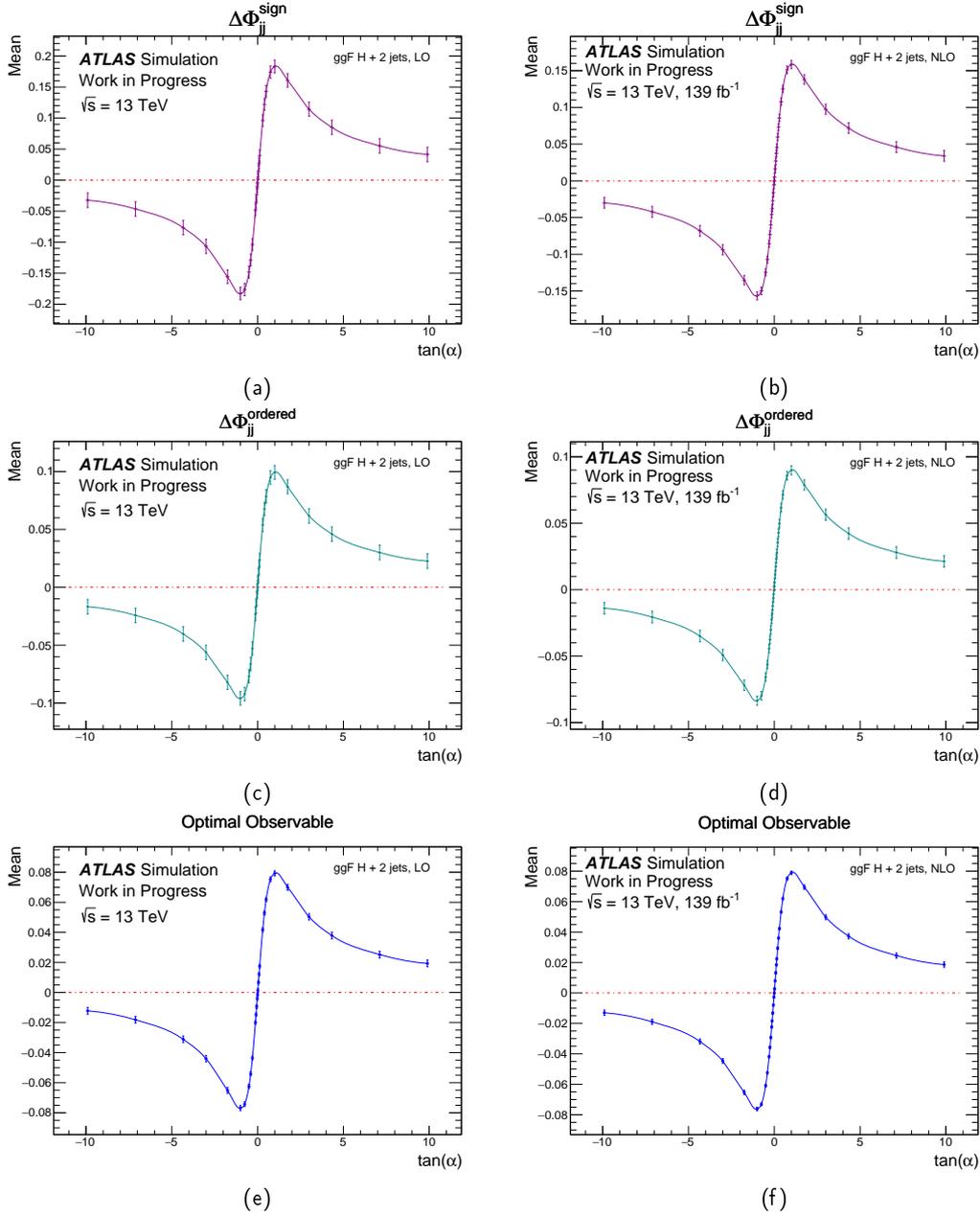
FIGURE 4.6: Mean values of CP-odd observables as a function of various $\tan\alpha$ values. (a) and (b) show the mean of $\Delta\Phi_{jj}^{\text{sign}}$ in the LO and NLO event sample, respectively. Also (c), (d) show the mean value of $\Delta\Phi_{jj}^{\text{ordered}}$ for LO and NLO events; (e), (f) show the Optimal Observable mean value distributions for LO and NLO with various $\tan\alpha$. The red dashed line indicates the zero mean value of the observables, which illustrates the CP-invariance.
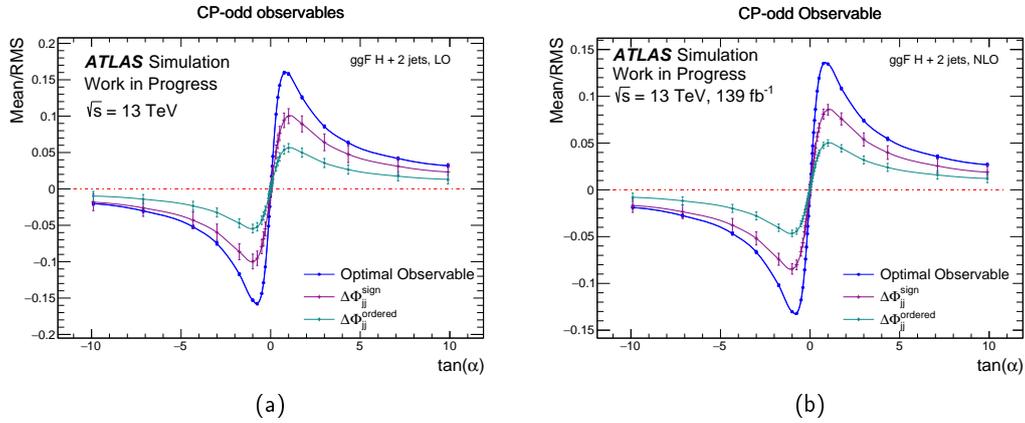
FIGURE 4.7: Comparison of the mean values of CP-odd observables normalized to the root mean square (RMS) as function of $\tan\alpha$ for (a) LO and (b) NLO samples. The red dashed line indicates the expectation of mean value zero for SM coupling. The solid blue line indicates the normalized mean value for Optimal Observable, and the purple line indicates the normalized mean values for $\Delta\Phi_{jj}^{\text{sign}}$ and green line indicates the $\Delta\Phi_{jj}^{\text{ordered}}$ observables respectively. MadGraph5 [39] produces the LO ggF H+2 jets events, whereas the Powheg Monte Carlo generator [47–50] produces the NLO event sample.

4.3 Sensitivity of different subprocesses

The ggF H+2 jets events can be classified into three different groups depending on their flavor combination of the initial state partons. This gluon-gluon (gg), gluon-quark (gq) and quark-quark (qq) subprocesses are described in Sec. 3.1. The validation of the reweighting method for each subprocess is discussed in Sec. 4.1. The relative contributions of the subprocesses to the total ggF H+2 jets event sample are summarized in Table 4.3 and examples of Feynman diagrams are shown in Fig. 4.8. The dominant contribution arises from gg initiated processes with 63.3 % followed

Initial partons	Subprocesses	Contribution
gluon - gluon	$gg \to Hgg$ $gg \to Hqq$	63.3.%
gluon - quark	$gq \to Hgq$ $g\bar{q} \to Hg\bar{q}$	33.9 %
quark - quark	$qq' \to Hqq'$ $q\bar{q}' \to Hq\bar{q}'$ $qq \to Hqq$ $q\bar{q} \to H\bar{q}q'$ $q\bar{q} \to Hgg$ $q\bar{q} \to Hq\bar{q}$ $\overline{qq}' \to H\overline{qq}'$ $\overline{qq} \to H\overline{qq}$	2.8 %

TABLE 4.3: Relative contribution of the different subprocesses for ggF H+2 jets production at LO.

by the gq initiated process with 33.9 % and the qq initiated process with 2.8 %.

Fig. 4.9 shows the $\Delta\Phi_{jj}^{\text{sign}}$ distributions for the different in subprocesses of ggF H+2 jets for various $\tan\alpha$ values. The $\Delta\Phi_{jj}^{\text{sign}}$ mean values for various $\tan\alpha$ coupling models are shown in Fig. 4.10. The highest deviations of the mean value is obtained with the quark-quark initiated subprocess, while gg, gq initiated processes give less sensitivity to the anomalous coupling between Higgs boson and gluons. However, gluon initiated subprocesses give the largest contribution to the total ggF sample.

Another CP-odd observable is $\Delta\Phi_{jj}^{\text{ordered}}$. Its distributions for the different subprocesses of ggF H+2 jets are shown in Fig. 4.11. The mean values as a function of the $\tan\alpha$ coupling are shown in Fig. 4.12. Similar to $\Delta\Phi_{jj}^{\text{sign}}$, qq initiated processes are more sensitive to the effective Higgs boson and gluon couplings than gluon initiated processes. Quark-quark initiated processes have larger statistical uncertainties due to their small contribution to ggF H+2 jets events.

The distributions of the Optimal observable for different $\tan\alpha$ values are shown in Fig. 4.13 for the different subprocess. Fig. 4.14 shows the sensitivity of the subprocesses to anomalous couplings of the Higgs boson to the gluons. The qq initiated process also here gives highest sensitivity to anomalous couplings of the Higgs boson and gluons.

Fig. 4.10, 4.12, 4.14 provide a comparison between the reweighted events at LO
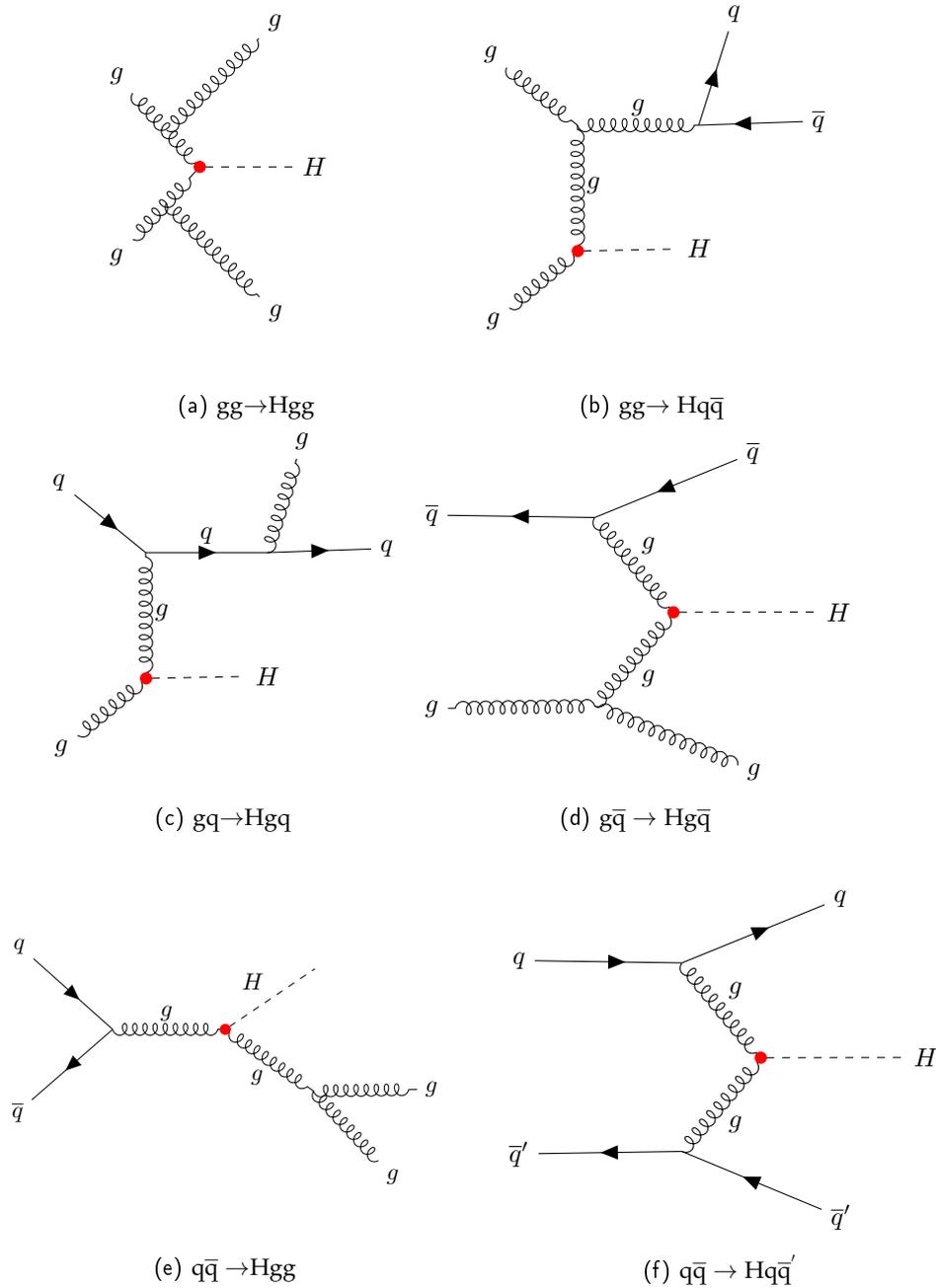
FIGURE 4.8: Example of Feynman diagrams at leading-order (LO) for different subprocesses of ggF H+2 jets. The event ggF H+2 jets can be classified depending on their incoming parton flavor combinations as a gluon-gluon (gg): (a) gg→Hgg, (b) gg→ Hq$\bar{q}$; gluon-quark (gq): (c) gq→Hgq, (d) g$\bar{q}$ → Hg$\bar{q}$; and quark-quark (qq): (e) q$\bar{q}$ →Hgg; (f) q$\bar{q}$ → Hq$\bar{q}'$. Red dot illustrates the top quark loop with assumption of infinite top quark mass $m_{\text{top}} \rightarrow \infty$. Feynman diagrams are from MadGraph5 [39].

and NLO. In the NLO sample, the sensitivity of subprocesses to the anomalous coupling of the Higgs boson and gluons are reduced due to the LO reweighting. Furthermore, the gluon-gluon initiated processes shown similar sensitivity compared

(a)

(b)

(c)
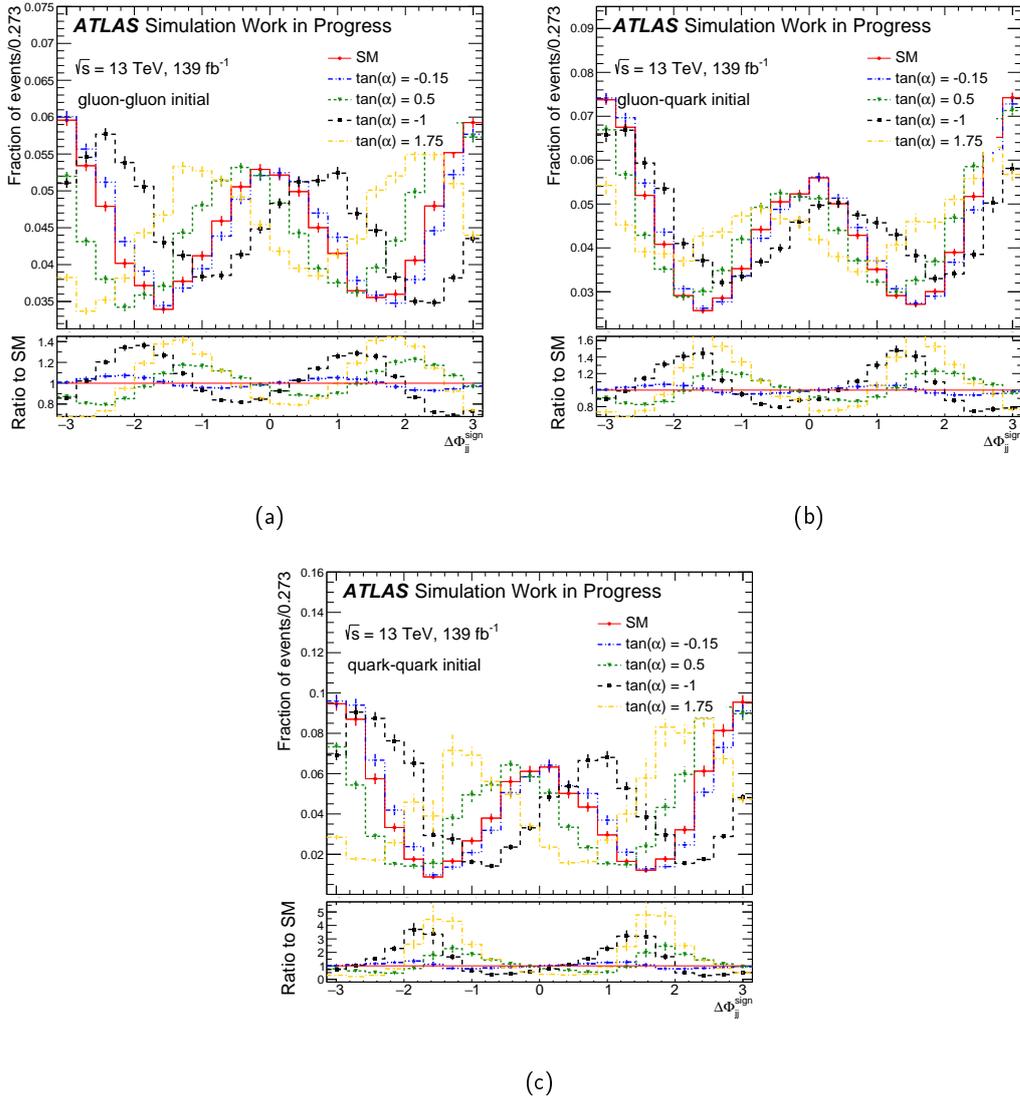
FIGURE 4.9: Normalized distributions of $\Delta\Phi_{jj}^{\text{sign}}$ for various $\tan\alpha$ values for the different (a) gg, (b) gq, and (c) qq subprocesses with statistical uncertainties. Here, ggF H+2 jets events are produced at NLO and reweighted at LO. Ratio plot shows the ratio for effective anomalous coupling to the SM prediction.

to the gluon-quark initiated processes. From the LO sample, one can see that the gluon-quark initiated process provide large deviations of the mean value from zero than the gg initiated process.
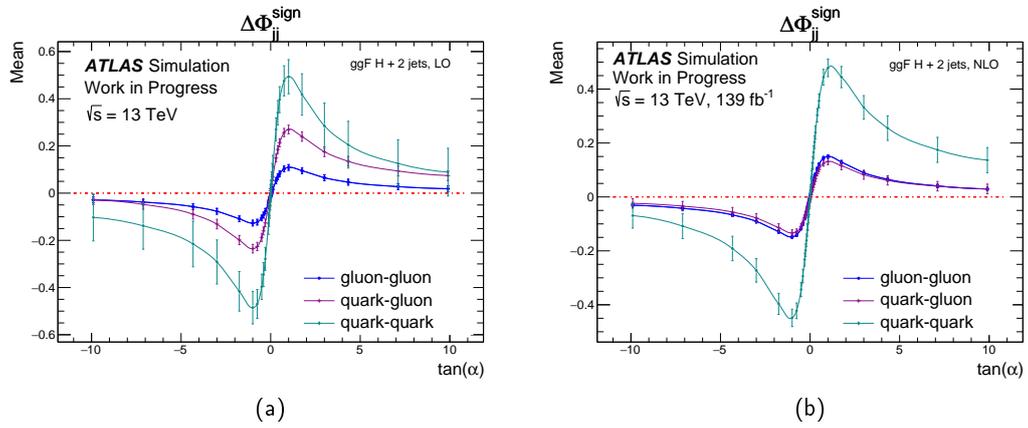
FIGURE 4.10: Distribution of $\Delta\Phi_{jj}^{\text{sign}}$ mean values as a function of $\tan\alpha$ for different initial states. (a) LO and (b) NLO reweighted samples are compared in gluon-gluon (blue), for gluon-quark (purple), and in quark-quark (green) initiated processes. Only statistical uncertainties are shown.

(a)

(b)

(c)
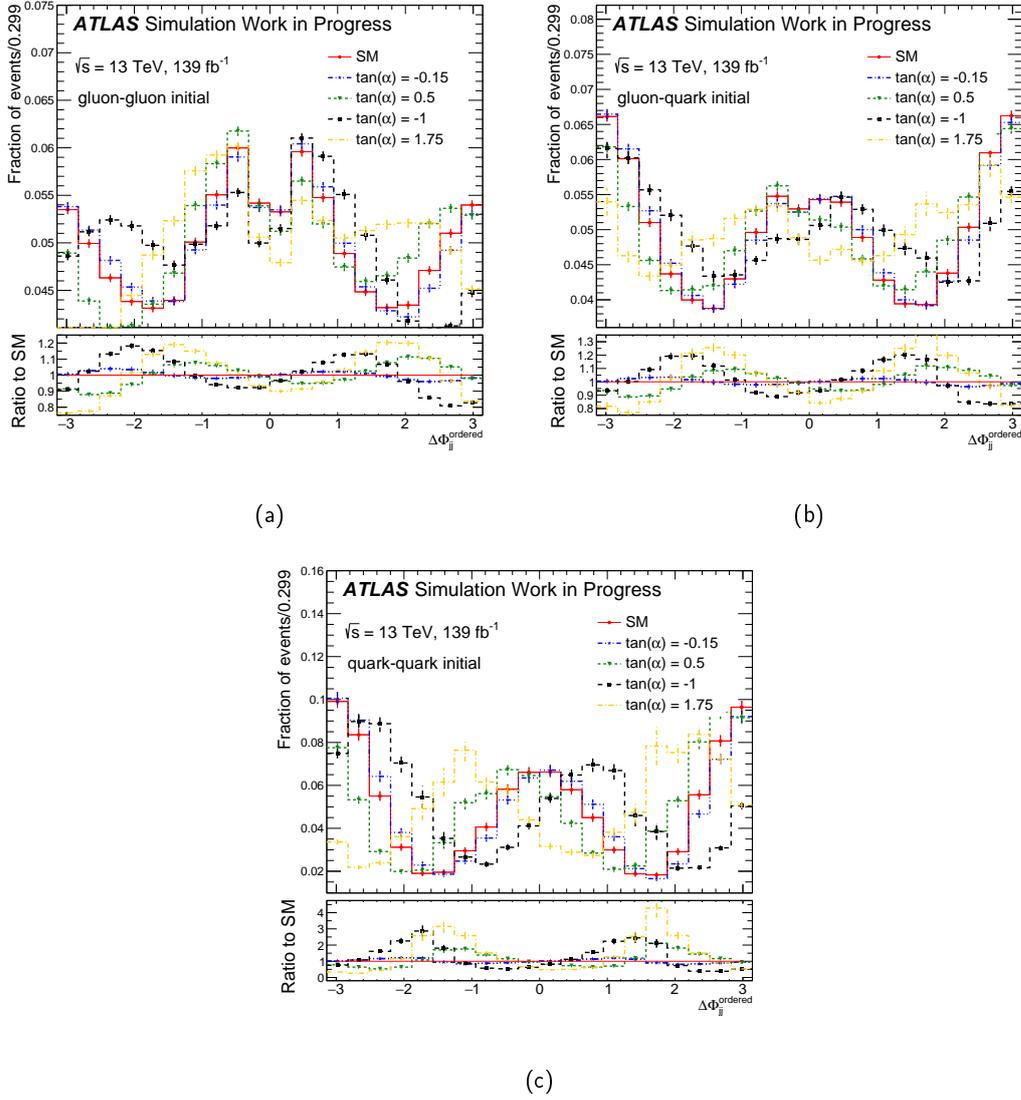
FIGURE 4.11: Normalized distributions of $\Delta\Phi_{jj}^{\text{ordered}}$ for various $\tan\alpha$ values for the different (a) gg, (b) gq, and (c) qq subprocesses with statistical uncertainties. Here, ggF H+2 jets events are produced at NLO and reweighted at LO. Ratio plot shows the ratio for effective anomalous coupling to the SM prediction.
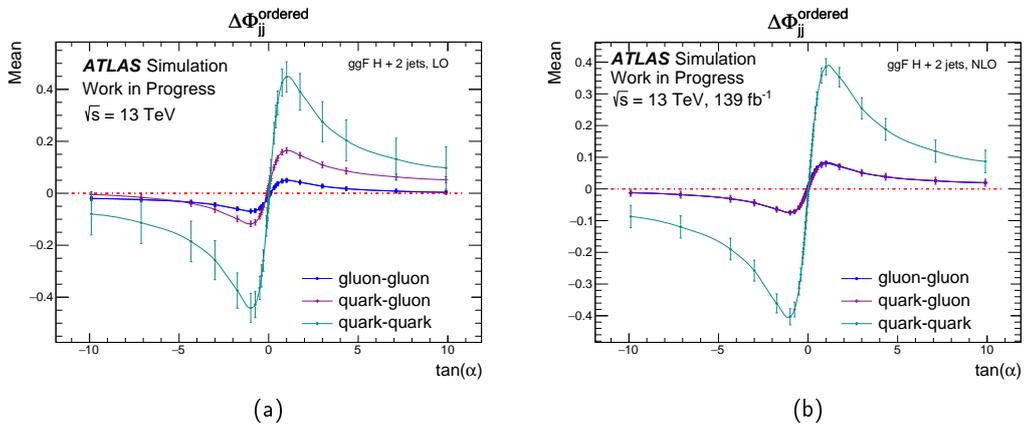
FIGURE 4.12: $\Delta\Phi_{jj}^{ordered}$ mean values distribution as a function of $\tan\alpha$ for different initial states. (a) LO and (b) NLO reweighted samples are compared in gluon-gluon (blue), for gluon-quark (purple), and in quark-quark (green) initiated processes. Only statistical uncertainties are shown.

(a)

(b)

(c)
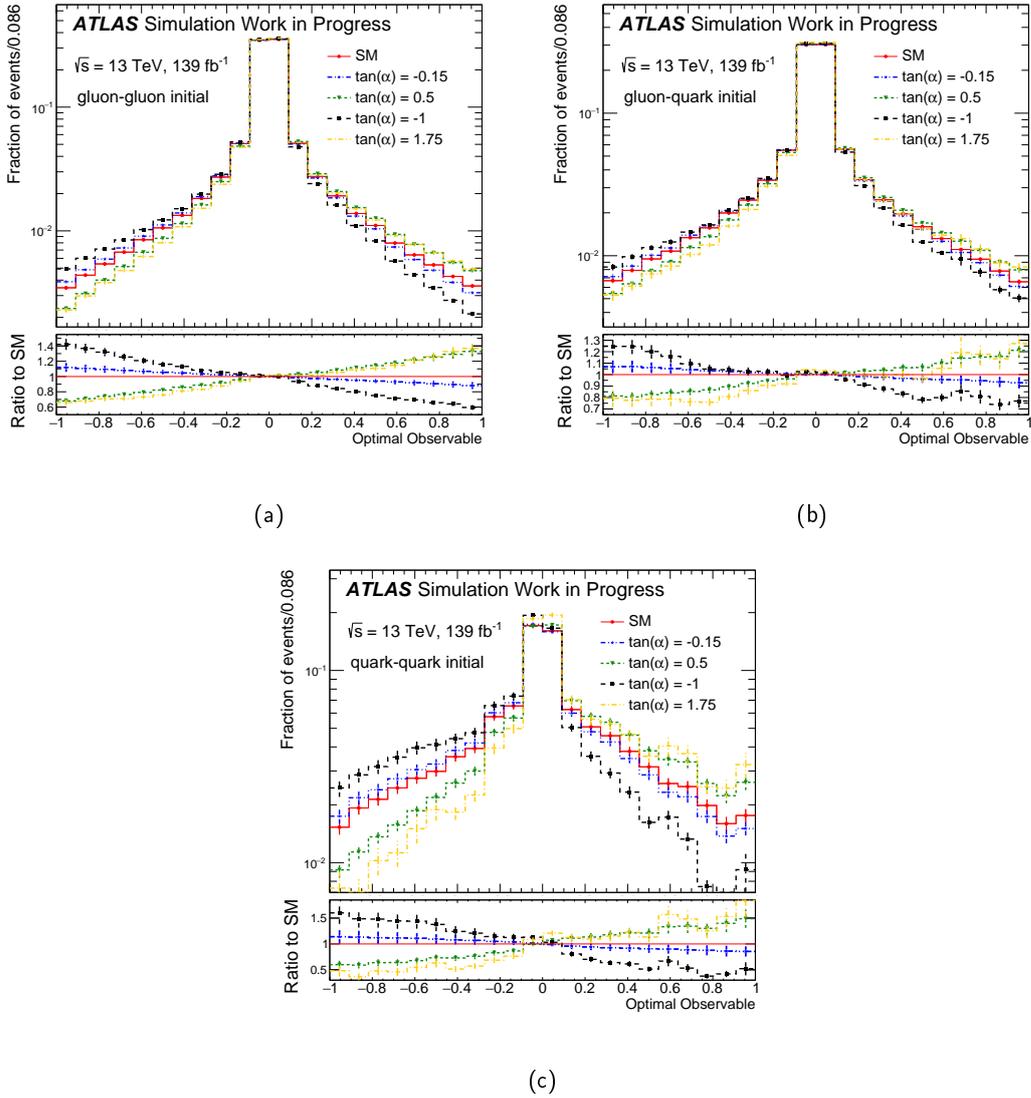
FIGURE 4.13: Normalized distributions of $\mathcal{OO}$ for various $\tan\alpha$ values for the different (a) gg, (b) gq, and (c) qq subprocesses with statistical uncertainties. Here, ggF H+2 jets events are produced at NLO and reweighted at LO. Ratio plot shows the ratio for effective anomalous coupling to the SM prediction.
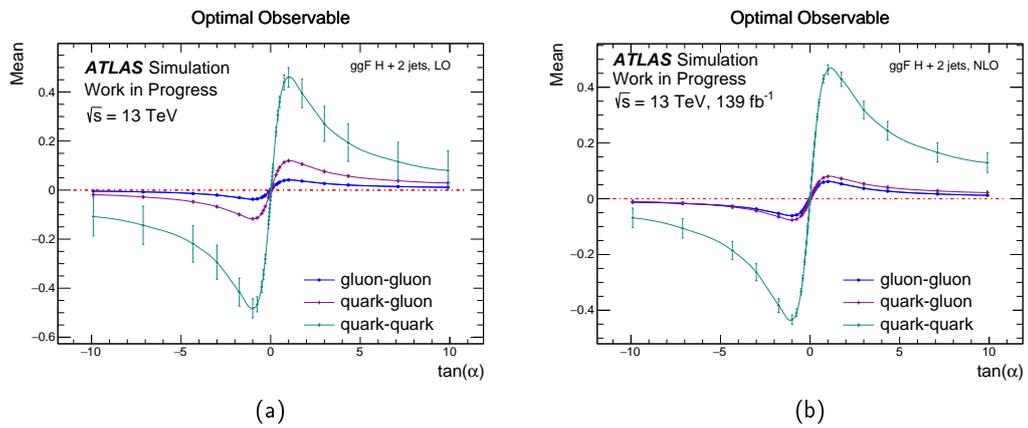
FIGURE 4.14: $\mathcal{OO}$ mean values distribution as a function of $\tan\alpha$ for different initial states. (a) LO and (b) NLO reweighted samples are compared for gluon-gluon (blue), in gluon-quark (purple), and in quark-quark (green) initiated processes. Only statistical uncertainties are shown.

Chapter 5

ATLAS Experiment at the LHC

This chapter provides information for a particle accelerator, in particular, LHC, and high luminosity experiment, the ATLAS detector. Furthermore, proton-proton (pp) collisions data collected in the ATLAS experiment will be presented.

In principle, particle accelerator constructed in two different ways, linear colliders (LINAC) and circular or synchrotron accelerators. Circular or synchrotron accelerators can provide higher energies than linear colliders [78]. The accelerators experiment can be performed in two different ways, fixed target experiments $\sqrt{s} = \sqrt{2E_{\text{beam}} m_{\text{target}}}$ and colliding beam experiments $\sqrt{s} = 2E_{\text{beam}}$. The collision of two particles has the advantage of getting a higher center of mass energies than the fixed target experiment. The LHC is a synchrotron accelerator.

5.1 Large Hadron Collider

The Large Hadron Collider, LHC, [79] is the machine designed to investigate Standard Model (SM) physics and beyond. This accelerator is located at the CERN laboratory near Geneva at the edge of France and Switzerland. This proton-proton collider installed up to 170 m below the ground within a 27 km circumference tunnel. The LHC is built in the former LEP tunnel, which was running in the period of 1989-2000. It designed to reach maximum $\sqrt{s} = 14$ TeV in pp collisions, also collides the heavy-ion collisions like lead Pb-Pb. Initially, LHC has been colliding proton beams in the center mass of energy 900 GeV and reached 7 TeV in 2010. Run 1 of the LHC is the period of data taking in 2010-2011(2012) with $\sqrt{s} = 7(8)$ TeV. Run 2 covers the time between 2015-2018 with $\sqrt{s} = 13$ TeV. This particle accelerator has two proton rings. The beam circle is bent by superconducting magnets, which are 15 m long 35-ton dipoles. The 1232 dipole magnets are called "2-in-1" because two proton beams are placed in opposite directions in each beam pipe. The dipoles create the 8.33 T magnetic field and are cooled by liquid helium to a temperature of 1.9^o K. Beams consists of 2808 proton bunches with bunch spacing 25 ns. And each bunch contains approximately 10^{11} protons. The energy stored in the beam is 360 MJ. The collision rates are proportional to the beam size. Therefore the beams are squeezed at the interaction points.

The LHC has four different experiment detectors, and the accelerator chain is sketched in Fig. 5.1. The LINAC produces 50 MeV protons and the Booster fed by those protons and further accelerates them up to 1.4 GeV. Then protons are injected to the Proton Synchrotron (PS) to get 26 GeV energy. From the PS, Super Proton Synchrotron (SPS) accelerate proton beams up to 450 GeV. Then, protons are injected into the two rings, where they reach the 6.5 TeV center of mass-energy in each beam. The collision of protons will be recorded at detectors.
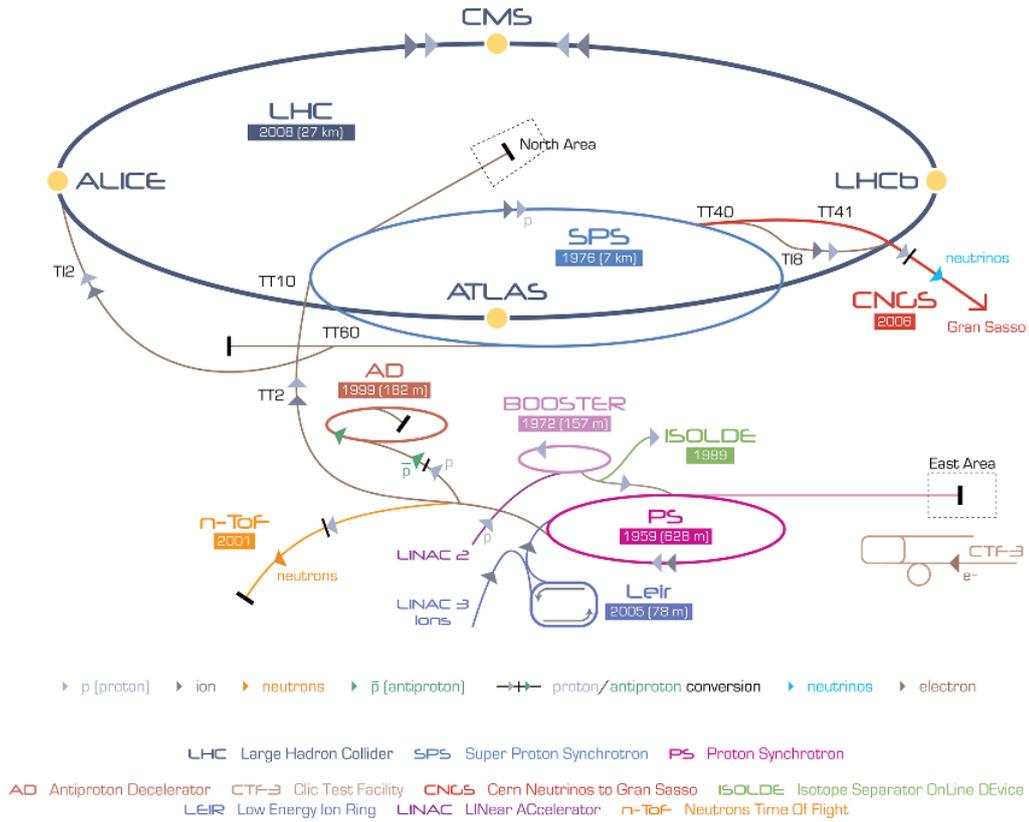
FIGURE 5.1: The LHC accelerator chain [78].

The experiments CMS and ATLAS are general-purpose detectors. LHCb studies rare decays, CP-violation, and b-quark physics. Another detector, ALICE, investigates quark-gluon plasma.

This thesis uses data collected at the ATLAS experiment.

5.1.1 Luminosity

In a particle collider, there are some important parameters. The number of collision events produced per second is defined by the collision rate and it is given by

$$R = \frac{dN}{dt} = \mathcal{L}\sigma, \tag{5.1}$$

where, σ is the cross section of the process with unit of area usually in a barn and $\mathcal{L}$ is the instantaneous luminosity with unit of [1/(cross section $\times$ time)]. The number of events produced N can be found by the integrating equation 5.1,

$$N_{\text{produced events}} = \sigma \times \int \mathcal{L}dt, \tag{5.2}$$

where $\int \mathcal{L}dt$ is the integrated luminosity. All events cannot be detected due to the inefficiency of the trigger. Therefore, these have to calculated experimentally,

$$N_{\text{observed events}} = \sigma \times \int \mathcal{L}dt \times \epsilon, \tag{5.3}$$

with the total efficiency ϵ of identifying collision events of interest. The difference between integrated and instantaneous luminosity is instantaneous luminosity drops when the protons collide. For two colliding proton bunches instantaneous luminosity defined as

$$\mathcal{L} = \frac{N_p^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F,$$ (5.4)

where particles number in a bunch denoted as N_p, bunches number in a beam symbolized as n_b. And the revolution frequency is denoted as f_{rev}, γ_r stands for the relativistic gamma factor. Beam emmitance to the transverse plane is represented by ϵ_n, the collsion point beta function is symbolized by β^*. And F is the reduction factor for geometric luminosity which depends on the crossing angle at the interaction point. More detailed information can be found on [80].
Integrated luminosity is alternatively,

$$\mathcal{L}_{\text{int}} = \mathcal{L}_0 \tau_{\mathcal{L}} [1 - e^{-T_{\text{run}}/\tau_{\mathcal{L}}}]$$ (5.5)

where T_{run} is the total length of the luminosity run and $\tau_{\mathcal{L}}$ which denotes the beam half-life. The maximum instantaneous luminosity of the LHC is 10^{33} - 10^{34} cm^{-2} s^{-1}, whereas integrated luminosity per year is the 10 - 100 fb^{-1}.

5.2 ATLAS Detector

The ATLAS detector [81] located at CERN, is a general-purpose detector, and it has symmetry in the forward-backward direction as well as to the interaction point. This large high luminosity detector has a length of 44 m and a height of 25 m as illustrated in Fig. 5.2, The detector consists of an inner tracking detector surrounded by the
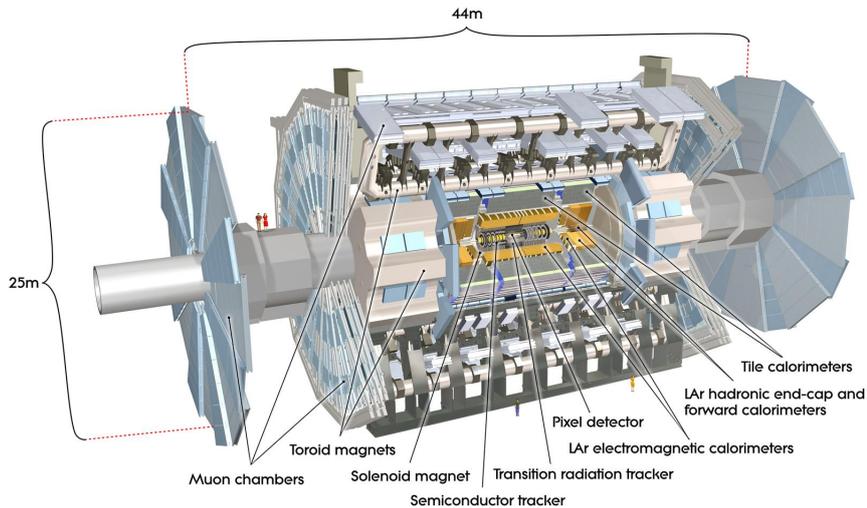
FIGURE 5.2: ATLAS detector structure [81].

superconducting solenoid, electromagnetic and hadronic calorimeters, and muon chambers. ATLAS detector has 2 magnets: the inner magnet is solenoid with 2 T magnetic field and the outer magnets are toroid which gives a 4 T magnetic field. The following sections will give more detailed information for each part of the ATLAS detector [81]. Before going further, there are necessary definitions will be mentioned

Detector part	Detector technology	Resolution	η coverage	
			Meas.	Trigger
Tracking	silicon/gas	$\sigma_{p_T}/p_T = 0.05\% \ p_T \oplus 1\%$	± 2.5	
EM calorimeter	liquid Argon	$\sigma_E/E = 10\% \ \sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimeter				
barrel and end-cap	steel/scintillator	$\sigma_E/E = 50\% \ \sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward		$\sigma_E/E = 100\% \ \sqrt{E} \oplus 10\%$	$3.1 < |\eta| < 4.9$	$3.1 < |\eta| < 4.9$
Muon spectrometer	RPC's/drift	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

TABLE 5.1: The performance and detector technology of the ATLAS
experiment [78, 81].

herein brief.

Rapidity and Pseudorapidity: A cylindrical coordinate system is used in the collider experiment. The z-axis is defined in direction of the beam. And θ is the polar angle and the azimuthal angle is noted ϕ. Here, the interaction point is $z = 0$. The rapidity of the particle is defined as,

$$y = \frac{1}{2} \log \left(\frac{E + p_z}{E - p_z} \right) = \tanh^{-1} \left(\frac{p_z}{E} \right),\tag{5.6}$$

where E is the energy and p_z is the particles' momentum in the z-direction. In this coordinate system, θ is not Lorentz invariant therefore the pseudorapidity is defined as function of θ,

$$\eta \equiv -\log \tan(\theta/2).\tag{5.7}$$

From Eq. 5.7, forward region is defined by $\eta \geq 1$, backward region $\eta \leq -1$. The rapidity difference Δy is Lorentz invariant under the boost along the beam direction. For massless particle, the rapidity and the pseudorapidity are equal.

Separation of particles, ΔR distance, is used to determine the particles' separation in $(\eta - \phi)$-plane. It is defined as

$$\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2},\tag{5.8}$$

where $\Delta \eta$ is the pseudorapidity difference and $\Delta \phi$ is the azimuthal angle difference.

Transverse quantities: In hadron collisions, one does not know the initial state partons of the event. Initial p_z is unknown, therefore, final p_z cannot be constrained like $\sum p_T = 0$. The transverse mass is defined as

$$m_T^2 = \sqrt{E_T^2 - p_T^2}.\tag{5.9}$$

where $E_T = E \sin \theta$ is transverse energy and $p_T = p \sin \theta$ is transverse momentum of a particle. Missing transverse energy is

$$E_T^{\text{miss}} = - \sum_{\text{all visible}} \overrightarrow{E_T}.\tag{5.10}$$

5.2.1   Inner Tracking Detector

The purpose of the tracking detector is to measure the trajectory and charge of a particle and its momentum. This sub-detector should be constructed from as little material as possible to suppress multiple scattering. Two main technologies are used

for inner detector gas and silicon detectors. The detector is surrounded by a magnetic field to cause the curvature of the particle's track. The inner detector of ATLAS
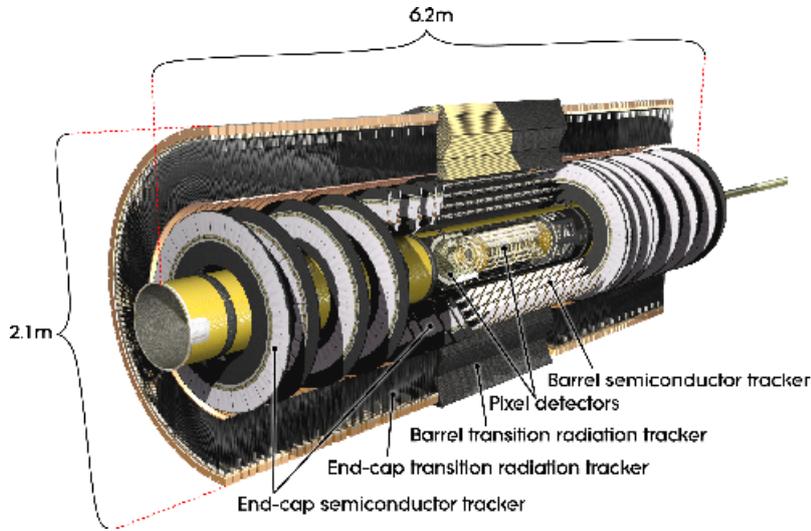
FIGURE 5.3: The layers of inner detector [81].

is shown in Fig. 5.3. Inner tracking detector has a length of 6.2 m and a diameter of 2.1 m. The solenoid generates the 2T magnetic fields. The tracking detector consists of pixel and silicon microstrip (SCT) trackers and Transition Radiation Tracker (TRT). The pixel and SCT enclose the $|\eta| < 2.5$ area. The pixel detector has a four layers, each layers $50 \times 400 \, \mu m^2$ size in $(R - \phi \times z)$ plane with 80.4 million readout channels. New insertable B-layer (IBL) was implemented in the deep-seated pixel detector layer. This IBL is close to beamline which is important for heavy flavor identification. The SCT has eight strip layers, coordinates measured by the 40mrad stereo strips in the barrel region in $(R - \phi \times z)$ plane with 6.3 million readout channels. And TRT has straw tubes with 4 mm, gives the coordinates in $R - \phi$ plane. The tracks are covering $|\eta| < 2.0$ with 351000 readout channels.

5.2.2 Calorimeters

The calorimeters measure the energy of the particles and their respective positions. There are 2 different calorimeters, electromagnetic (EM) and hadronic calorimeters. EM calorimeter is designed to detect particles with electromagnetic interaction such as photon and electron, whereas hadronic calorimeter is detecting hadrons like neutron. One of the important parameters of the calorimeters is the relative uncertainty (resolution of the calorimeters). It is defined as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b \oplus \frac{c}{E}, \tag{5.11}$$

where a is the stochastic term and b is the constant and c is primarily due to noise such as electronic. Here, $\oplus$ is quadrature adding, $a \oplus b = \sqrt{a^2 + b^2}$. The calorimeters can be two different types of detectors such as homogeneous calorimeter and sampling calorimeter. The homogeneous calorimeter is made of scintillating high Z heavy inorganic material like crystals and liquid Argon. It generates the electromagnetic signal. And sampling calorimeters made of the active and passive medium to create a signal and absorb the energy, respectively. Also, they have different energy resolutions which you can see from table 5.1.
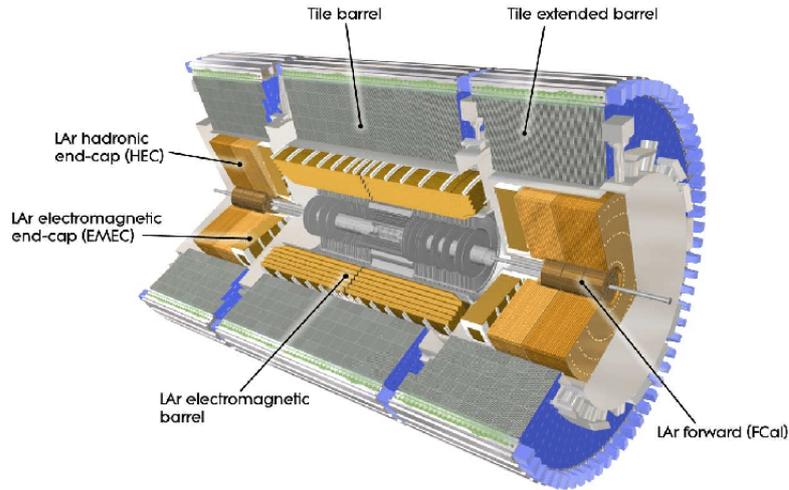
FIGURE 5.4: The layers of calorimeter [81].

As shown in the Fig. 5.4, **electromagnetic calorimeter** of the ATLAS experiment is made by liquid Argon (LAr) and has sublayers such as barrel part with $|\eta| < 1.475$ coverage and two end cap components with $1.375 < |\eta| < 3.2$ coverage. The geometry of this calorimeter has a full coverage, complete ϕ symmetry. The total thickness of the EM calorimeter's barrel is 22 - 30 radiation lengths (X_0) in respect to $0 < |\eta| < 0.8$, and end-caps thickness is 24 X_0 - 33 X_0 in $0.8 < |\eta| < 1.3$.

Hadronic calorimeters consist of tile calorimeter, LAr hadronic end-cap calorimeter (HEC) and LAr forward calorimeter (FCal). The tile calorimeter is inserted around to the EM calorimeter and it has a barrel with extended region $0.8 < |\eta| < 1.7$. This sublayer is the sampling calorimeter that is made of steel and scintillating active material. Thickness of this layer is 9.7 λ at $\eta = 0$. The HEC contains in the end-cap 2 independent EM calorimeter. It expands to $|\eta| = 3.2$ and it causes the overlap with a forward calorimeter. Also, it overlaps with the tile calorimeter with a range of $|\eta| = 1.5$.

5.2.3 Muon Spectrometer

The muon spectrometer is designed to measure the muons charge and positions. Tracks of the muon are reconstructed by inner tracking detector and by muon spectrometry. This detector is placed as outermost layer of the ATLAS detector that is outside of the hadronic calorimeters. Muon is minimum ionizing particles because muons penetration. From Fig. 5.5, the muon spectrometer consists of large superconducting air-core toroid magnets, trigger, and high precision measurement tracking chambers. The large barrel toroid covers the pseudorapidy of $|\eta| < 1.4$. The end-cap toroids cover the range of $1.6 < |\eta| < 2.7$. The region $1.4 < |\eta| < 1.6$ is a transition region supplied by both barrel and end-cap magnetic fields deflection. There are 3 different hight precision measurement tracking chambers: Monitored Drift Tubes (MDT's), Cathode Strip Chambers (CSC's) and Resistive Plate Chambers (RPC's). This drift tubes cover almost all of the pseudorapidity range. The CSC's are gas-filled with wires intersected with metallic strips which gives the muon 2D coordinates. These strip chambers cover the range of $2 < |\eta| < 2.7$. T The RPC's are located in the barrel and Thin Gap Chambers (TGC's) used in the end-cap regions, covers the $|\eta| < 2.4$ region.
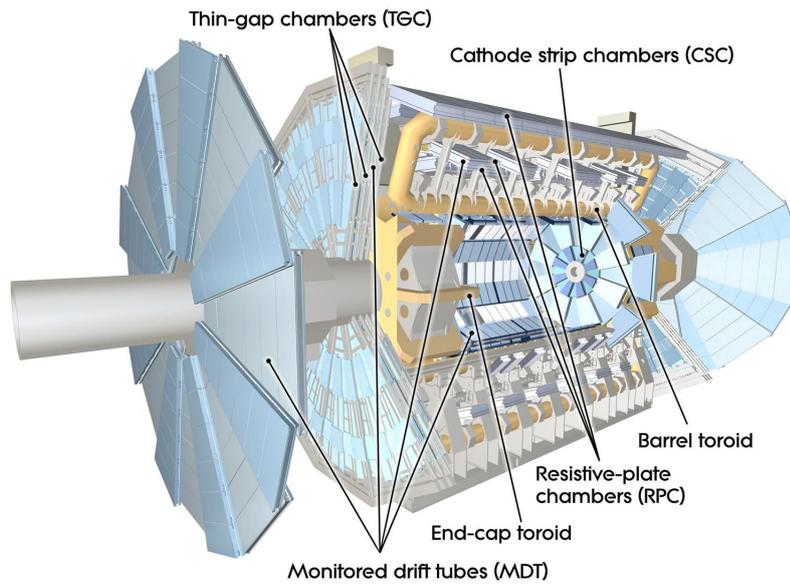
FIGURE 5.5: The Muon system [81].

5.2.4   Trigger system in ATLAS

Trigger system decides the which event can be kept or removed for the further analysis. In order to do, this trigger system uses the given information of the bunch-crossing interaction. The collision rate at the LHC is 40 MHz. The trigger system of the ATLAS experiment has two levels [81, 82]. The first level is hardware-based Level-1 trigger and second level is software based high-level trigger referred to as HLT. The Level-1 trigger uses the information from coarse granularity of calorimeter and muon spectrometer. This input information determines the region of interest (RoIs) in the detector. The RoI is the information of η and ϕ coordinates. The decision time of the first level trigger is approximately 2.5 μs and it reduces the trigger rate by 100 kHz. The next level, the HLT uses full granularity information. This level reduces the trigger rate by approximately 1 kHz with 200 ms decision time. The Fig.
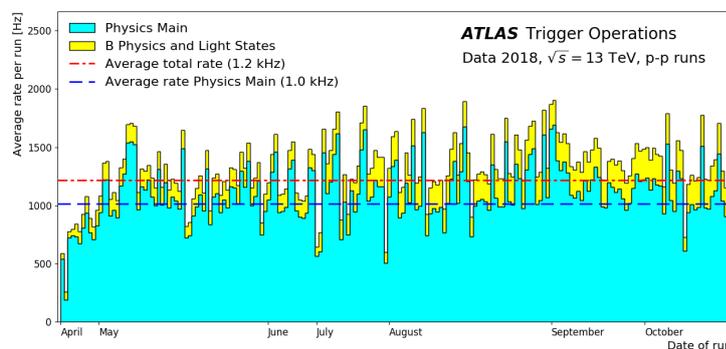
FIGURE 5.6: Trigger operation in 2018 at ATLAS experiment. Average *pp* data recording rate in September 2018 [83].

5.6 shows the average recording rate in 2018 of ATLAS trigger operation.

5.2.5   Data taking in Run 2

The ATLAS experiment collected data of proton-proton collision between the period of 2015-2018 referred as to LHC Run 2. During the data taking in Run 2, center of mass-energy was 13 TeV with integrated luminosity of 139 fb^{-1} with an uncertainty of 2.4 fb^{-1} [84]. Integrated luminosity during Run 2 summarized in table 5.2. In the

Data sample	2015 + 2016	2017	2018	Combined
$\mathcal{L}_{\text{int}}$ [fb^{-1}]	36.2	44.3	58.5	139.0
Total uncertainty [fb^{-1}]	0.8	1.0	1.2	2.4
Relative uncertainty [%]	2.1	2.4	2.0	1.7

TABLE 5.2: Integrated luminosity after data quality requirements of Run 2 with $\sqrt{s} = 13$ TeV in *pp* collisions [84]

left plot of Fig. 5.7 shows the total integrated luminosity versus years and right side
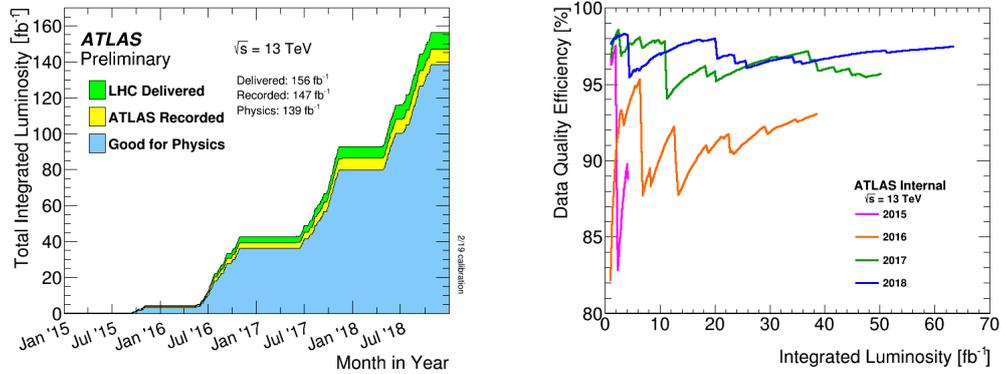
FIGURE 5.7: The integrated luminosity in 2015-2018 [85, 86].

plot indicates the data efficiency in function of luminosity in $\sqrt{s} = 13$ TeV at the ATLAS detector.

Chapter 6

Reconstruction and Identification of Physics Objects

A reconstruction and an identification of physics objects are necessary for particles detected by ATLAS detector during proton-proton collisions. This chapter will discuss the reconstruction and identification of electrons, muons, jets, hadronically decaying τ-leptons, and missing transverse energy. These definitions are important for selecting a H$\rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ event with two jets. In addition, the trigger selection is briefly discussed.

6.1 Tracks and Vertices

Trajectories of charged particles are measured by the inner detector (ID) of the ATLAS experiment [87, 88]. In the inner detector, the magnetic field curves the trajectories and the momentum is proportional to the curvature. The inner detector is responsible for measuring the momentum of charged particles and the determination of primary and secondary vertices of the charged particles. In order to reconstruct the tracks and vertices of the charged particles, the subdetector: Pixel Detector, Semiconductor Tracker (SCT), and Transition Radiation Tracker (TRT) are used. Tracks reconstruction is based on the "hit" in the sub-tracking layers in the inner detector. This hit refers to the three-dimensional space-point measurements in the detector. Track reconstruction uses a pattern recognition algorithm. Reconstructed tracks have to fulfill quality criteria in order to reject misidentified tracks or badly measured tracks.

Furthermore, vertices of the interactions are reconstructed from the tracks [89]. Multiple vertices are reconstructed per collision since multiple protons interact. Therefore, the primary vertex is reconstructed by tracks with the highest sum of squared transverse momenta.

6.2 Electron

Reconstruction of the electrons is based on the combined information of energy deposition in the electromagnetic calorimeter and charged track information in the inner detector. Electron candidates are reconstructed in the central region $|\eta| < 2.47$ of the ATLAS detector with transverse momentum $p_T > 15\,\text{GeV}$ and with the medium identification criterion. In principle, the electron candidates are loose, medium, and tight, which are based on a likelihood identification [90]. To avoid bad identification and bad reconstruction performance, the transition region between the calorimeters, $1.37 < |\eta| < 1.52$ is excluded. Electrons within a distance of ΔR between 0.2 and

0.4 of a jet are excluded. Reconstructing electrons with high efficiency and isolating from other particles are important. The efficiencies of the reconstruction and identification of the electron candidates can be calculated from the tag and probe method using $Z \rightarrow ee$, $J/\psi \rightarrow ee$. The efficiency of the reconstructed electrons is in the range of 96 % $-$ 99%.

6.3 Muon

The muons leave a clear distinctive signature in the detector with minimum energy loss [91, 92]. Tracks of the muon are measured in the inner detector, and in muon spectrometry (MS). Muons are initially reconstructed separately in the ID and MS, then later combined. The hit pattern is used for reconstruction in MS. A track candidates fitted to all formed segments in the muon chamber. Furthermore, tracks need to fulfill the selection criteria. Muons are selected with transverse momentum greater than 10 GeV and with pseudorapidity $|\eta| < 2.5$ with medium identification. The quality criteria mostly suppress the pion and kaon decays. In addition, muons are identified as a loose, medium, tight, and high-p_T muons. The identification of the loose muons has a high reconstruction efficiency with good quality tracks. The medium muons have minimum systematic uncertainties from its reconstruction and its calibration. Moreover, tight identification has a low reconstruction efficiency but with maximum purity. Another identification algorithm is the high-p_T muons, which has a maximum momentum resolution for $p_T > 100$ GeV. Efficiencies of the reconstruction and identification algorithm can be estimated from the tag-and-probe method using $Z \rightarrow \mu\mu$ and $J/\psi \rightarrow \mu\mu$ events. The reconstruction and identification efficiencies of the muon is almost 99 %. Muon candidates are neglected when muon and jets distance is $\Delta R < 0.4$.

6.4 Jets

According to color confinement of the quantum chromodynamics (QCD), quarks and gluons are confined in hadrons. In the detector, quarks and gluons are hadronized and create a spray of colorless hadrons, which are called jets. Reconstruction of the jets is performed using the anti-k_t algorithm based on a topological cluster of the electromagnetic and hadronic calorimeter. Jets are reconstructed with a radius of 0.4 and transverse momentum greater than 20 GeV with $|\eta| < 4.9$. The jets arising from the pileup with $p_T < 50$ GeV and $|\eta| < 2.4$ suppressed by an algorithm called Jet Vertex Tagger (JVT) [93, 94]. While forward JVT algorithm rejects the pileup jets in the forward region with $p_T < 50$ GeV and $|\eta| > 2.5$.

Jets initiated from bottom quark is called b-jets. This b-jets can be identified by flavor-tagging methods [95, 96]. The b-flavored hadrons have a long lifetime compared to other hadrons. Therefore its decay creates a secondary vertex separated from the primary vertex. This secondary vertex can be reconstructed using the track information of the particles within the jet. With a multivariate analysis, b-jets can be tagged in a range of $|\eta| < 2.5$. The efficiency of b-jets selection is 85 %. Tagging and vetoing the b-jets is essential for this analysis to reduce the background events from top quark, because top quark decay almost in 100 % of the cases to b-quarks. Also, b-jets are not taken into account when distance of $\Delta R = 0.2$ between jets and electons.

6.5 Hadronic Tau

The τ-lepton can decay leptonically and hadronically. The visible part of leptonically decaying τ-lepton can be reconstructed and identified as an electron or muon. The visible decay of the hadronically decaying τ-lepton ($\tau_{\text{had-vis}}$) is reconstructed from jets [97]. Hadronically decaying τ-leptons often called a 1-prong or 3-prong, which depend on the associated number of tracks. The visible decay of the hadronically decaying τ-lepton is reconstructed with $p_T > 20$ GeV with $|\eta| < 2.47$. The transition region between the barrel and end cap of the calorimeter with $1.37 < |\eta| < 1.51$ is excluded due to poor reconstruction and identification performance. In this analysis, $\tau_{\text{had-vis}}$ is selected with medium identification, which is based on the Boosted decision tree method. This method is used to distinguish jets which originates from quark and gluons. In addition, τ-leptons are ignored when they have angular distance of $\Delta R = 0.2$ to jets. Furthermore, the efficiency of the reconstruction and identification of the hadronically decaying τ-leptons are estimated from the tag-and-probe method, which is based on Z$\to \tau\tau$ events. Also, in order to reject electron misidentified as a hadronically decaying τ-leptons, 1-prong $\tau_{\text{had-vis}}$ is selected. The efficiencies are 55 % for the 1-prong, and 40 % for the 3-prong hadronically decaying τ-lepton candidates.

6.6 Missing Transverse Energy

Invisible particles in the detector from the proton-proton collisions are measured as a missing transverse energy [98]. This missing transverse momentum can be obtained from the vector sum of the transverse momentum of the reconstructed particles. And it defined by

$$
\begin{aligned}
E_T^{\text{miss}} &= E_T^{\text{miss},e} + E_T^{\text{miss},\gamma} + E_T^{\text{miss},\tau_{\text{had}}} + E_T^{\text{miss},\mu} + E_T^{\text{miss,jet}} + E_T^{\text{miss,soft}}, \quad (6.1) \\
E_T^{\text{miss}} &= \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2},
\end{aligned}
$$

where $E_T^{\text{miss,object}} = -\sum p_T^{\text{object}}$.

The energy and momentum of the initial state particles are not accessible along the beam axis. This transverse momentum of the final state particles is equal to zero from the energy and momentum conservation. Reconstruction of the missing transverse energy contains the hard objects term and soft term. The hard term is reconstructed from electrons, muons, photons, τ-leptons and jets. Moreover, the soft term defines the reconstruction of a charged particle along with hard scattering vertices.

6.7 Trigger Selection

In the ATLAS experiment, hardware-based first level trigger and software-based High-Level Trigger (HLT) are used as discussed in Sec. 5.2.4. This analysis uses the Higgs boson production in gluon fusion associated with two jets. The subsequent decay of the Higgs boson is one hadronically decaying τ-lepton and one leptonically decaying τ-lepton. Therefore, single lepton triggers are used [99]. For the events collected from pp collision in 2015, single-electron channel uses the trigger `HLT_e24_lhmedium_L1EM20VH` with transverse momentum of electron $p_T > 25$ GeV.

The single muon channel uses the trigger `HLT_mu20_iloose_L1MU15` with transverse momentum of muon $p_T > 21$ GeV.

For the 2016 events from pp collisions, single electron channel uses the trigger `HLT_e26_lhtight_nod0_ivarloose` with electron tranverse momentum $p_T > 27$ GeV.

The single muon channel uses the trigger `HLT_mu26_ivarmedium` with muon transverse momentum $p_T > 27.3$ GeV. Table 6.1 summarizes the requirements for transverse momentum of the light leptons at trigger level for Run 2.

2017 and 2018 events from pp collisions use the same trigger selection for light leptons as described as 2016 events.

Trigger	$p_T^{\mathbf{min}}$ **[GeV]**	
	2015	2016-2018
Single electron	25	27
Single muon	21	27.3

TABLE 6.1: One lepton trigger selection for light leptons [99].

Chapter 7

Event selection

To select the signal process, one needs to apply the selection criteria to the simulated and observed data. Therefore, this chapter will focus on the primary selection criteria, main techniques to reconstruct the invariant mass, and multivariate analysis for signal and background separation.

7.1 Reconstruction of di-tau invariant mass

To reconstruct the invariant mass of the di-tau system is the important method because this variable is useful to distinguish signal and background. The signal process is the Higgs production in gluon fusion associated with jets, ggF H+2 jets. And the subsequent decay of the Higgs boson is one hadronically decay and one leptonically decaying τ-leptons. It is difficult to reconstruct the mass of di-tau-system due to the three neutrinos arising from their decay products. Neutrinos are invisible in the detector, therefore missing transverse energy is used from the detector. Here, two techniques for the reconstruction of the invariant mass of the di-tau final states will be briefly presented.

7.1.1 Collinear mass approximation

The collinear approximation is proposed to reconstruct the invariant mass of the Higgs boson decay to a pair of τ-leptons [100]. This method assumes that the neutrinos are collinear to the visible decay products of the τ-leptons. Neutrinos are invisible to the detector; therefore, the invisible momentum in τ-decay can be calculated via

$$
\begin{aligned}
E_{T_x}^{\text{miss}} &= p_{\text{miss}_1} \sin \theta_{\text{vis}_1} \cos \phi_{\text{vis}_1} + p_{\text{miss}_2} \sin \theta_{\text{vis}_2} \cos \phi_{\text{vis}_2} \\
E_{T_y}^{\text{miss}} &= p_{\text{miss}_1} \sin \theta_{\text{vis}_1} \cos \phi_{\text{vis}_1} + p_{\text{miss}_2} \sin \theta_{\text{vis}_2} \cos \phi_{\text{vis}_2},
\end{aligned}
\tag{7.1}
$$

where $E_{T_x}^{\text{miss}}$, $E_{T_y}^{\text{miss}}$ are x, y component of the missing transverse energy and p_{miss_1}, p_{miss_2} are momentum of the neutrinos. Whereas θ_{vis}, ϕ_{vis} are the polar and azimuthal angle of the visible decay products of τ-lepton. The invariant mass of di-tau system, which neglecting the masses of τ-lepton, light leptons and pions, can be calculated as

$$
M_{\tau\tau}^{\text{coll}} = \frac{m_{\text{vis}}}{x_1 x_2},
\tag{7.2}
$$

where m_{vis} is the invariant mass of the visible τ-lepton decay products, i.e., electron or muon and hadronically decaying τ-lepton and $x_{1,2}$ is the momentum fractions are

carried by visible decay products in τ-lepton. Then momentum fractions defined as

$$x_{1,2} = \frac{p_{\text{vis}_{1,2}}}{p_{\text{vis}_{1,2}} + p_{\text{miss}_{1,2}}}. \tag{7.3}$$

This collinear approximation technique can reconstruct the mass of the di-tau system; however, it suffers from the back-to-back topology of the decay products of the τ-leptons. Most of the decay products of the τ-leptons are decaying back-to-back, and it causes the Eq. 7.1 tend to be degenerate. When $\phi_{\text{vis}_1} = \phi_{\text{vis}_2} + \pi \Rightarrow p_{\text{miss}_{1,2}} \sim \sin^{-1}(\phi_{\text{vis}_1} - \phi_{\text{vis}_2})$. And it diverges to $|\phi_{\text{vis}_1} - \phi_{\text{vis}_2}| \rightarrow \pi$.

Also, in this analysis, invariant mass of di-tau system is reconstructed by the Missing Mass Calculator (MMC). And the detailed information of this method can be found on [101].

7.2   Event Preselection

Event preselection is the first step to suppress the background and select the signal events. Firstly, there are data quality requirements applied. In the following, selection criteria [102] for the events are listed.

1. **Truth matching:** This criterion only applied to the Monte Carlo simulated event samples. Jets are misidentified as a hadronically decaying τ-leptons. Hadronically decaying τ-leptons are matched to the truth τ-leptons. Therefore, background of jets misidentified as τ_{had} is suppressed in the simulated events. Hence, this fake backgrounds are estimated using the data-driven method [103].

2. **Primary vertex:** It is selected to be a minimum one primary vertex with two associated tracks. Furthermore, associated tracks transverse momentum has to fulfill $p_T > 0.5$ GeV.

3. **Momentum fraction:** Momentum fractions carried by visible decay products of di-tau system are x_1, x_2. This momentum fractions have to fulfill $0.1 < x_1 < 1.4$ and $0.1 < x_2 < 1.2$. This selection criteria removes the event with invisible tau decay and suppress the event with mismatched missing transverse energy to the decay of τ-leptons. Momentum fractions are calculated via the collinear mass approximation method as described in 7.1.1

4. **Angular variables:** The angular distance between visible decay products of the leptonically decaying and hadronically decaying τ-leptons has to fulfill $\Delta R_{\tau\tau} < 2.5$ and $\Delta \eta_{\tau\tau} < 1.5$ GeV. This selection suppresses the non-resonant background.

5. **Missing transverse energy:** The missing transverse energy has to fulfill $E_T^{\text{miss}} > 20$ GeV. This selection criterion removes events without neutrino.

6. **Trigger selection:** Depending on the data taking period and flavor of the particles, one needs to apply different trigger selection. The trigger selection has already been discussed in 6.7. In brief, the single lepton trigger is used with light lepton (e/μ) with $p_T > 27$ GeV and $|\eta| < 2.47$. Additionally, hadronically decaying τ-lepton with $p_T > 30$ GeV and $|\eta| < 2.47$. Further, visible decay part of the hadronically decaying τ-lepton is denoted as $\tau_{\text{had}}^{\text{vis}}$.

7. **Tau selection and opposite charge:** Final state τ-leptons have to be the opposite charge, in order to conserve the charge of the Higgs boson. Moreover, τ_{lep} selected as gradient isolated and, both τ-leptons fulfills the medium identification criteria (above mentioned variables value in trigger selection).

8. **b-jet veto:** b-jets are identified by flavour tagging method. The b-jets are removed in order to suppress the backgrounds of a pair of top-quark and single top quark productions. Efficiency of rejecting the b-jets is 85 %.

9. **Two jets:** Events with a minimum of two jets in the final state are required. The leading (subleading) jet transverse momentum has to fulfill $p_T^{j_1} > 40(30)$ GeV.

10. **Transverse mass:** The transverse mass reconstructed from the light lepton and missing transverse momentum. Here, light lepton is the visible decay part of the leptonically decaying τ-lepton and it is denoted τ_{lep}^{vis}. This transverse mass $m_T < 70$ GeV is selected. This criterion reduces the background contribution from W+jets production.

11. **MMC mass:** Invariant mass of the final state particle reconstructed using the MMC method. Events are removed when algorithm does not converge.

The above-listed selection criteria suppress the backgrounds and increase the sensitivity of the signal for the test of CP-invariance. The distribution of some variables used in preselection are shown in Fig. 7.1. Fig. 7.1e is shown the invariant mass distribution of the di-tau system reconstructed by the MMC technique after all preselection criteria applied. Table. 7.1 shows the event yield of the respective processes after selection criteria (11) at $\sqrt{s} = 13$ TeV with $\mathcal{L} = 139\,\text{fb}^{-1}$. The CP-sensitive observables distributions are shown in Fig. 7.2 after preselection criteria.

Process	Event yield
ggF H+2 jets, H$\to \tau_{lep}\tau_{had}$	587.54 ± 3.62
VBFH, H$\to \tau_{lep}\tau_{had}$	199.59 ± 0.58
VH, H$\to \tau_{lep}\tau_{had}$	58.37 ± 0.67
Z$\to \tau_{lep}\tau_{had}$	$28\,108.07 \pm 123.71$
Z$\to \ell\ell$	2468.79 ± 117.44
Top	1219.82 ± 14.61
Diboson	920.01 ± 6.59
Fakes	$12\,165.50 \pm 41.64$
$\sum_{\text{bkg}}$	$45\,140.15 \pm 176.32$
Data	$45\,617.00 \pm 213.58$
s/b	0.01 ± 0.00
s/$\sqrt{\text{s+b}}$	2.76 ± 0.02

TABLE 7.1: Summary of the event yield after preselection criteria (11) for simulated samples and observed data at $\sqrt{s} = 13$ TeV with $\mathcal{L} = 139\,\text{fb}^{-1}$. Here are only statistical uncertainties are considered.

The dominant background is $Z/\gamma^* \to \tau\tau$, and it contributes 62 % of the total background. Furthermore, fakes background contributes 27 %, and Z$\to \ell\ell$ contributes 6% of the total background. The data and simulation model have a good agreement. The signal and background ratio is s/b = 0.013 and further separation

of signal and background are required. The next step to suppress the background is
the Boosted Decision Tree method which is multivariate analysis.
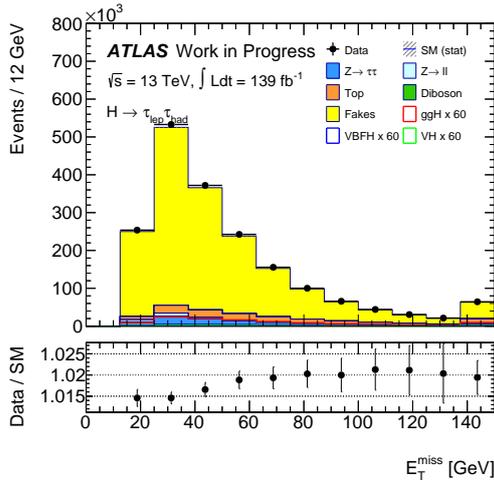
(a) Missing transverse energy distribution after selection criteria (6)
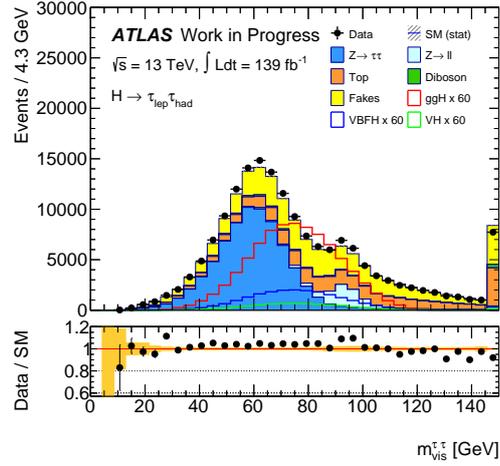
(b) The visible di-tau mass distribution after selection criteria (7)
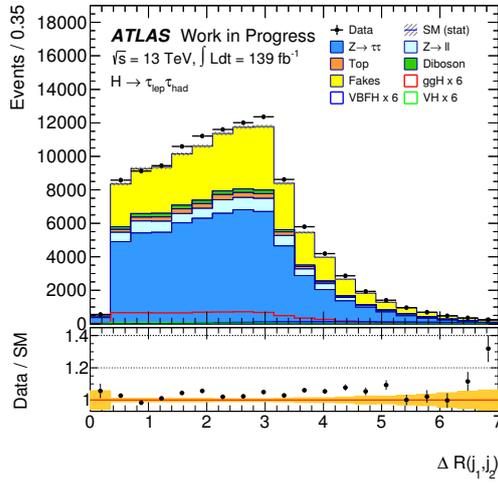
(c) The angular distance between jets distribution after selection criteria (8)
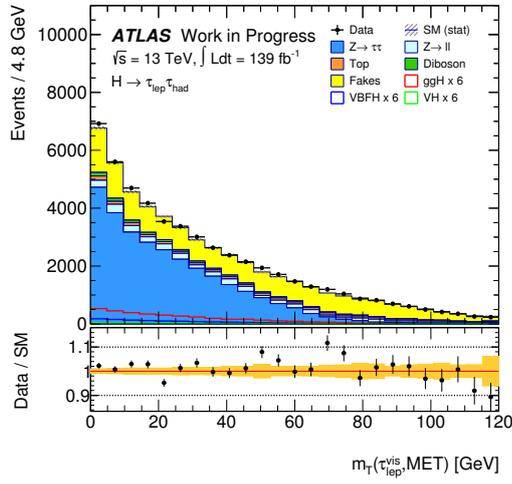
(d) The transverse mass distribution after selection criteria (9)
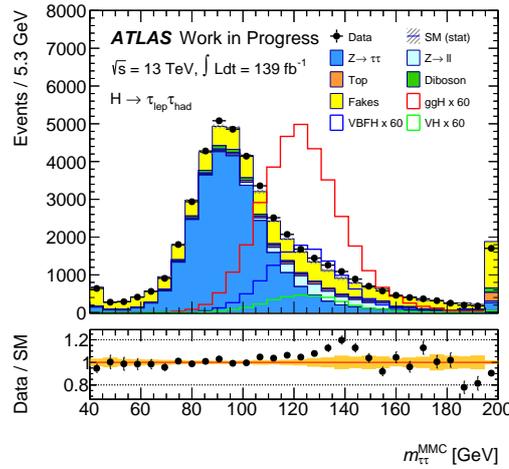
(e) The invariant mass of the Higgs boson after selection criteria (10)

FIGURE 7.1: Distribution of variables used in the preselection for signal and background processes. Only statistical uncertainty is considered. In all figures, notation SM (stat) represents the statistical uncertainties of the background.
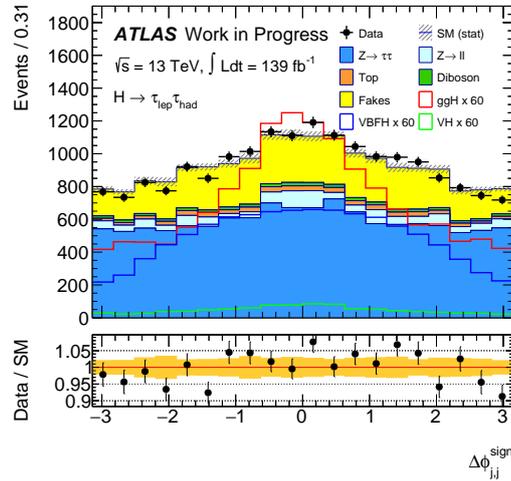
(a) The distribution of the signed azimuthal angle
between jets after selection criteria (11)
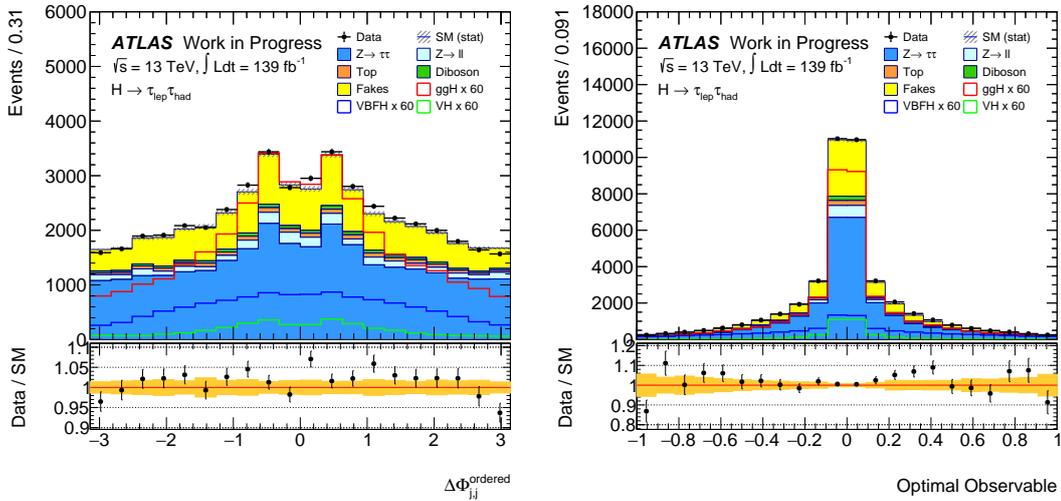
(b) The distribution of the azimuthal angle be-
tween jets ordered by η after selection criteria (11)

(c) The distribution of the Optimal Observable af-
ter selection criteria (11)

FIGURE 7.2: The distribution of CP-sensitive observables: (a) $\Delta\Phi_{jj}^{\text{sign}}$,
(b) $\Delta\Phi_{jj}^{\text{ordered}}$, (c) $\mathcal{OO}$ after selection criteria (11) are shown. Here
only statistical uncertainties are considered. In all figures, notation
SM (stat) represents the statistical uncertainties of the background.

7.3 Multivariate Analysis

The suppression of the background contribution based on the event preselection is not sufficient. In order to enhance the the signal-to-background ratio, a multivariate analysis technique is applied. The reasons to use multivariate analysis are correlations between observables in the signal and background are not visible by just defining the threshold values on the variables. Another reason is that the multivariate analysis combines several individual input variables into a single output for discrimination.

This analysis uses the method of Boosted Decision Trees (BDT) for separating signal and background. The BDT training is one of the analyzing tools based on the Toolkit for Multivariate Data Analysis (TMVA) package [104].

The following sections introduce the basic concept and performance of the BDT training. Finally, the signal region is defined based on the BDT output.

7.3.1 Boosted Decision Tree method

The Boosted Decision Tree (BDT) is an algorithm based on binary response for classifying signal and the background events. This machine learning technique uses diverse input variables to split the event sample to signal and background events. The schematic view of a single decision tree is shown in Fig. 7.3. The set of con-
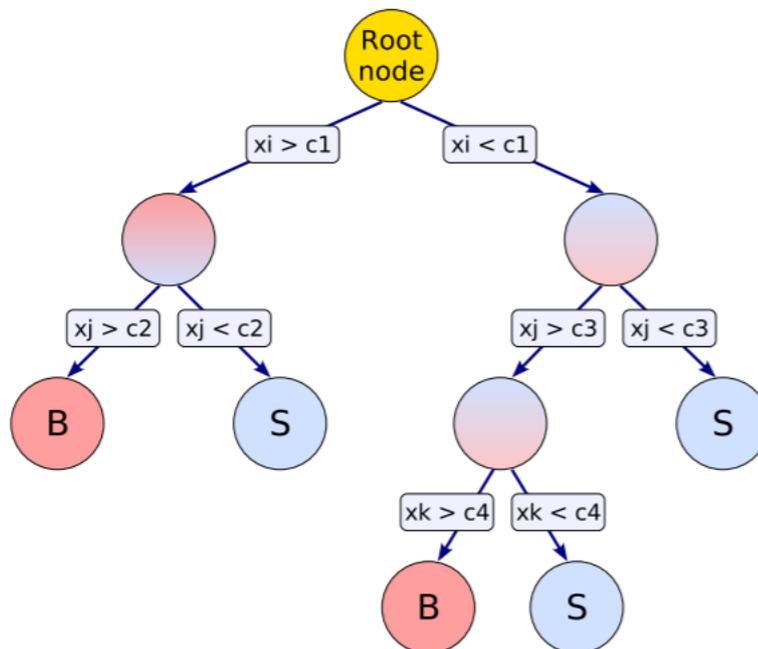
FIGURE 7.3: The schematic diagram of a Decision Tree [104] is illustrated. Defining the threshold value c_1 of the input variable x_i, to split the signal and background events. This splitting recursively continues until a consecutive set of nodes are fulfilled. The leaf nodes, B is the background, and S is the signal.

secutive nodes are asking the threshold values for each input variables to make a decision. The final verdict is called the leaf node, and splitting stops after fulfilling the criteria for the maximum number of nodes. This separation of the events decided

by so-called **Gini-index**, and it is defined as,

$$G = p(1 - p), \qquad\qquad (7.4)$$

where p is the purity of the node, which is the number of signal events divided by the sum of all events in the corresponding nodes. The values of the purity lie between zero and one, which indicates the mixedness of the signal and background samples. The BDT separation power depends on the combination of hundreds of simple decision tree called a **forest**. The forest can be achieved by boosting algorithm. The idea of the boosting method is applying an event weight for each training sample in the tree which depends on the decision of previous tree. The Boosting algorithm in the tree stabilizes the statistical fluctuations of the classifier and improves the separation power. In this analysis the Gradient Boost algorithm [105] is used.

The optimization of classification of signal and background events and the boosting technique are usually called the **BDT training**. During the BDT training, statistical fluctuations can occur in the decision tree. This is referred as to **overtraining**. In general, the BDT training tends to be overtrained. The overtraining can be detected by splitting the input samples in training and samples. However, this reduces the available statistics of the samples by a half. The solution of this problem is the cross validation method. This analysis uses five fold cross validation in order to increase the event statistics, which will be discussed in the Sec. 7.3.2.

Tuning hyper parameters of the BDT training can obtain optimization of the discrimination of the signal and background. The definitions of the parameters for configuration are summarised in the Table 7.2. This table shows the hyper parameter values used in this analysis. These hyper parameters are optimized in order to reduce the overtraining of the BDT and to obtain good separation between signal and backgrounds. The final result of the forest is called **BDT output** or BDT score.

Parameters	Description	Value in the BDT
NTrees	The number of decision tree	250
MinNodeSize	Minimum number of events in each final node (in percent of the total amount of events)	1%
Max Depth	Maximum allowed depth of the decision tree	5
nCut	The number of steps during the node threshold optimization	20
BoostType	Boosting type for trees in the forest	Gradient Boost
Shrinkage	Learning rate for boosting algorithm	0.1
Separation type	Defines the splitting of signal and background	GiniIndex
IgnoreNegWeights	Prevents the bias in the result	True

TABLE 7.2: Configuration of hyper parameters for the BDT training are described as well as values used for training the BDT [104].

The output of the BDT events is weighted by the normalized boosted weight in each decision tree. The BDT output distribution ranges between -1 and 1. Negative values are background-like and positive values are signal-like.

Input variables	Definitions	BDT$_{bkg}$	BDT$_{VBF}$
m_{MMC}	Invariant mass of the Higgs boson reconstructed by MMC technique	×	
$\Delta R_{\tau\tau}$	ΔR distance between visible decay products in hadronically and leptonically decaying τ-leptons	×	
$m_{\tau\tau}^{coll}$	Invariant mass of the Higgs boson reconstructed by collinear approximation	×	
$m_{\tau\tau}^{vis}$	Invariant mass of visible decay products mass in the τ-leptons	×	
$\Delta\eta_{\tau_{lep}^{vis},j_1}$	η-separation between light lepton and leading jets	×	×
m_{Hj_1}	Invariant mass of the Higgs boson with leading jet	×	×
$\Delta\eta_{\tau_{lep}^{vis},j_2}$	η-separation between light lepton and subleading jet	×	×
$\Delta R^{min}_{\tau_{lep}^{vis},j_1}$	Minimum ΔR distance between light lepton and leading jet	×	×
$\Delta R^{min}_{\tau_{had}^{vis},j_1}$	Minimum ΔR distance between hadronically decaying τ-lepton and leading jet	×	×
E_T^{miss}	The missing transverse energy	×	
p_T^H	The momenta of the Higgs boson reconstructed by visible and invisible decay products in τ-leptons	×	
$p_T^{\tau_{lep}^{vis}}/p_T^{\tau_{had}^{vis}}$	Transverse momentum ratio between light lepton and hadronically decaying τ-lepton	×	
m_T	Transverse mass reconstructed via light lepton and missing transverse energy	×	
$\sum_{scal} p_T$	Scalar sum of the momenta of visible decay of di-tau tau and jets	×	
m_{jj}	Invariant mass of the two leading jets		×
ΔR_{jj}	Distance ΔR between two leading jets		×
p_T^{total}	The vector sum of the transverse momenta of the visible tau decay products with two leading jets, and E_T^{miss}		×
$\eta_{j_1} \times \eta_{j_2}$	η-product of the two leading jets		×
$C_{\eta_1\eta_2}(\eta_{\tau_{lep}^{vis}})$	Lepton centrality quantifies the η position of light lepton with respect to the two leading jets in the event		×
$E_T^{miss}\phi$ centrality	This variable quantifies the relative angular position of the missing transverse momentum with respect to the visible tau decay products in the transverse plane.	×	

TABLE 7.3: Discriminating variables used in the BDT training. Definitions of the variables are given and the cross symbol (×) indicates which variables are used in the two BDTs.

7.3.2   Performance of the BDT training

As discussed in the previous section, event samples are divided internally in the MVA by training and test samples. It reduces the available event statistics, but it is useful to detect overtraining. The BDT training can increase the event statistics by using the method of k-fold cross validation. In this analysis, five-fold cross validation is used for BDT training.Further, it refers to BDT f1, BDT f2, BDT f3, BDT f4, and BDT f5. Each training classified as training and test samples and they are statistically independent.

For the BDT training, the signal sample ggF H+2 jets are produced at NLO in QCD, as discussed in Chapter 3. In the NLO order Monte Carlo simulation, some events have a negative event weight in particular phase space regions. The events with negative weight are ignored during the BDT training.

To separate the signal and backgrounds events, two different BDTs are trained and used in this analysis. The first BDT training is referred to as $\mathrm{BDT_{bkg}}$, which is trained to separate the signal processes ggF H+2 jets from all background processes. The BDT training uses the input variables for classification. The 15 input variables are used for $\mathrm{BDT_{bkg}}$. Variables are chosen based on their separation power of the background and signal. For the BDT training, there are no CP- sensitive variables are included. The definition of the input variables is described in the Table 7.3 and corresponding distributions are shown in the Fig. 7.4; 7.5; 7.6. The importance of the

Rank	Variable	Importance
1	m_{MMC}	0.14
2	$m_{\tau\tau}^{\mathrm{coll}}$	0.11
3	$m_{\tau\tau}^{\mathrm{vis}}$	0.09
4	$\Delta R_{\tau\tau}$	0.08
5	m_T	0.06
6	$\Delta\eta_{\tau_{\mathrm{lep}}^{\mathrm{vis}},j_1}$	0.06
7	p_T^H	0.06
8	$\Delta\eta_{\tau_{\mathrm{lep}}^{\mathrm{vis}},j_2}$	0.06
9	$\Delta R_{\ell,j_1}^{\min}$	0.06
10	E_T^{miss}	0.05
11	$E_T^{\mathrm{miss}}\phi$ centrality	0.05
12	$\Delta R_{\tau_{\mathrm{had}}^{\mathrm{vis}},j_1}^{\min}$	0.05
13	$\sum_{\mathrm{scal}} p_T$	0.05
14	$p_T^{\tau_{\mathrm{lep}}^{\mathrm{vis}}}/p_T^{\tau_{\mathrm{had}}^{\mathrm{vis}}}$	0.05
15	m_{Hj_1}	0.04

(a)

Rank	Variable	Importance
1	m_{jj}	0.15
2	ΔR_{jj}	0.14
3	p_T^{total}	0.12
4	$\eta_{j_1} \times \eta_{j_2}$	0.10
5	$C_{\eta_1\eta_2}(\eta_{\tau_{\mathrm{lep}}^{\mathrm{vis}}})$	0.09
6	$\Delta\eta_{\tau_{\mathrm{lep}}^{\mathrm{vis}},j_1}$	0.08
7	m_{Hj_1}	0.08
8	$\Delta\eta_{\tau_{\mathrm{lep}}^{\mathrm{vis}},j_2}$	0.08
9	$\Delta R_{\tau_{\mathrm{lep}}^{\mathrm{vis}},j_1}^{\min}$	0.07
10	$\Delta R_{\tau_{\mathrm{had}}^{\mathrm{vis}},j}^{\min}$	0.06

(b)

TABLE 7.4: Input variables for discrimination of signal and background in each BDT training. Table 7.4a shows the rank of the variables for $\mathrm{BDT_{bkg}}$ training. And Table 7.4b shows the rank of the input variables for $\mathrm{BDT_{VBF}}$ training.

variables is quantified by how often those variables are used to split the background and signal and they are listed in the Table 7.4. From the table, mass distributions: m_{MMC}, $m_{\tau\tau}^{\mathrm{coll}}$, and $m_{\tau\tau}^{\mathrm{vis}}$ ranked high but they are highly correlated to each other. Additionally, the variables $\Delta R_{\tau\tau}$ and m_T have a good separation power. As discussed

before, the performance of the BDT training can be improved by tuning the configuration parameters as listed in Table 7.2. The performance of the overall BDT training can be illustrated by the Receiver Operating Characteristic (ROC) curve. This curve shows the background rejection (1-signal efficiency) as a function of the signal efficiency. The ROC curve of the BDT_{bkg} performance are shown in Fig. 7.7. From the figure, it is visible that there are no significant deviations between BDT f1, BDT f2, BDT f3, BDT f4, and BDT f5. This means overtraining is not observed during the BDT training. The result of the BDT_{bkg} output is shown in Fig. 7.8a. A good separation is obtained between signal and non-Higgs boson backgrounds.

However, the Higgs boson productions via VBFH, VH, H$\rightarrow \tau_{lep}\tau_{had}$ are classified as signal events. Therefore, suppression of the VBFH background is crucial. Because VBFH contributed as a 34 % to the signal events (ggF H+2 jets) after preselection criteria. The event statistics of the VH background are small; therefore, in additional suppression of this background is not necessary. For this reason, the second BDT training is performed in order to discriminate ggF H+2 jets events from VBFH events. This will be referred as a BDT_{VBF} training. In this BDT_{VBF} training, 10 input variables are used. The definition of those variables are in the Table. 7.3. The importance of variables for BDT_{VBF} are shown in Table. 7.4b. The distributions of variables, m_{jj}, ΔR_{jj} are ranked high. They exploit the VBF event topology. Also, p_T^{total} and the product of two leading jets pseudorapidity have a good seperation power. The performance of the BDT_{VBF} is shown in the Fig. 7.7. No overtraining is observed according to the ROC curve. The BDT_{VBF} output is shown in Fig. 7.8a; good discrimination between signal and VBFH events is achieved.
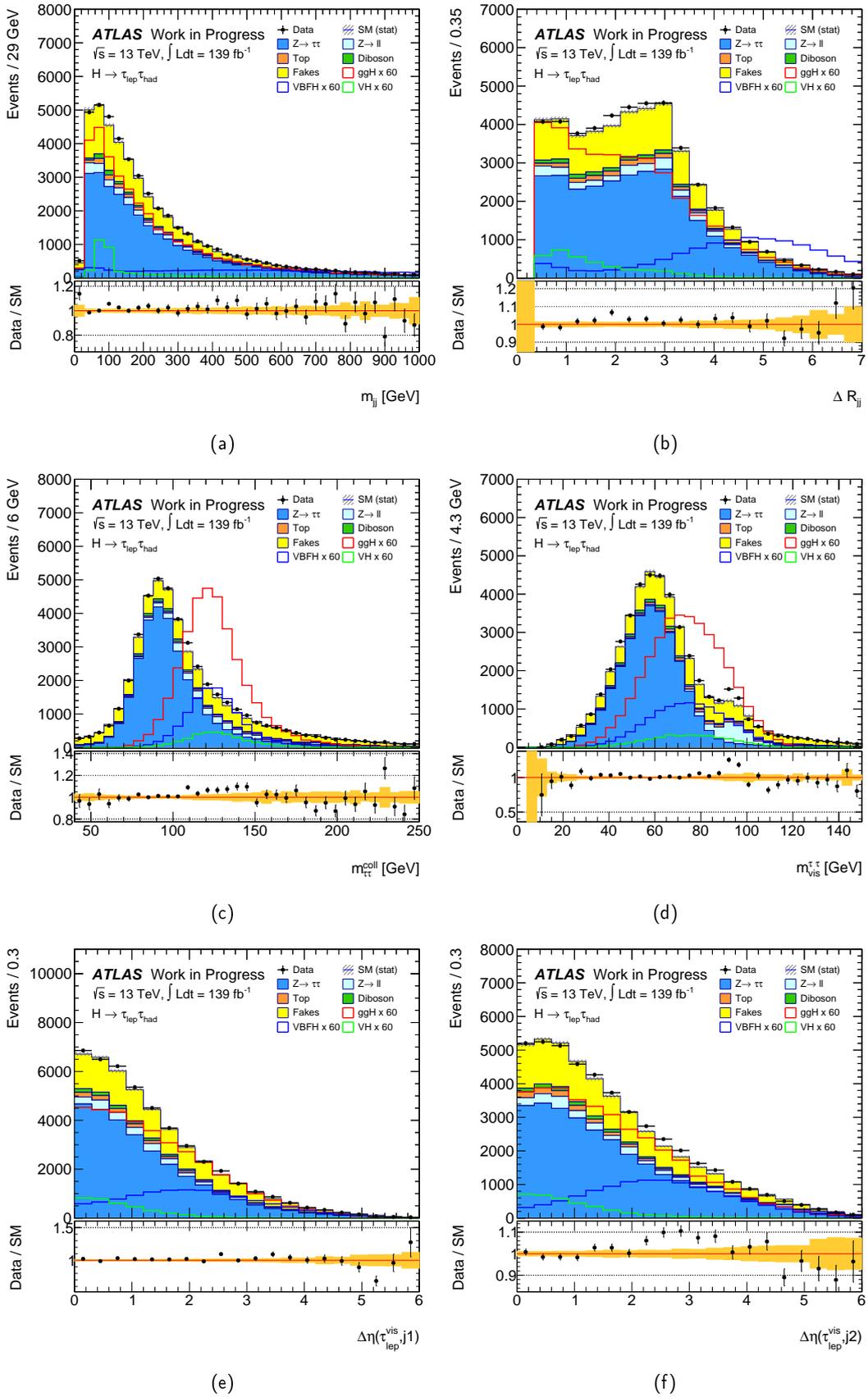
FIGURE 7.4: Distribution of discriminating variables used in the BDT
training process. Only statistical uncertainties are shown.

(a)

(b)

(c)

(d)

(e)

(f)
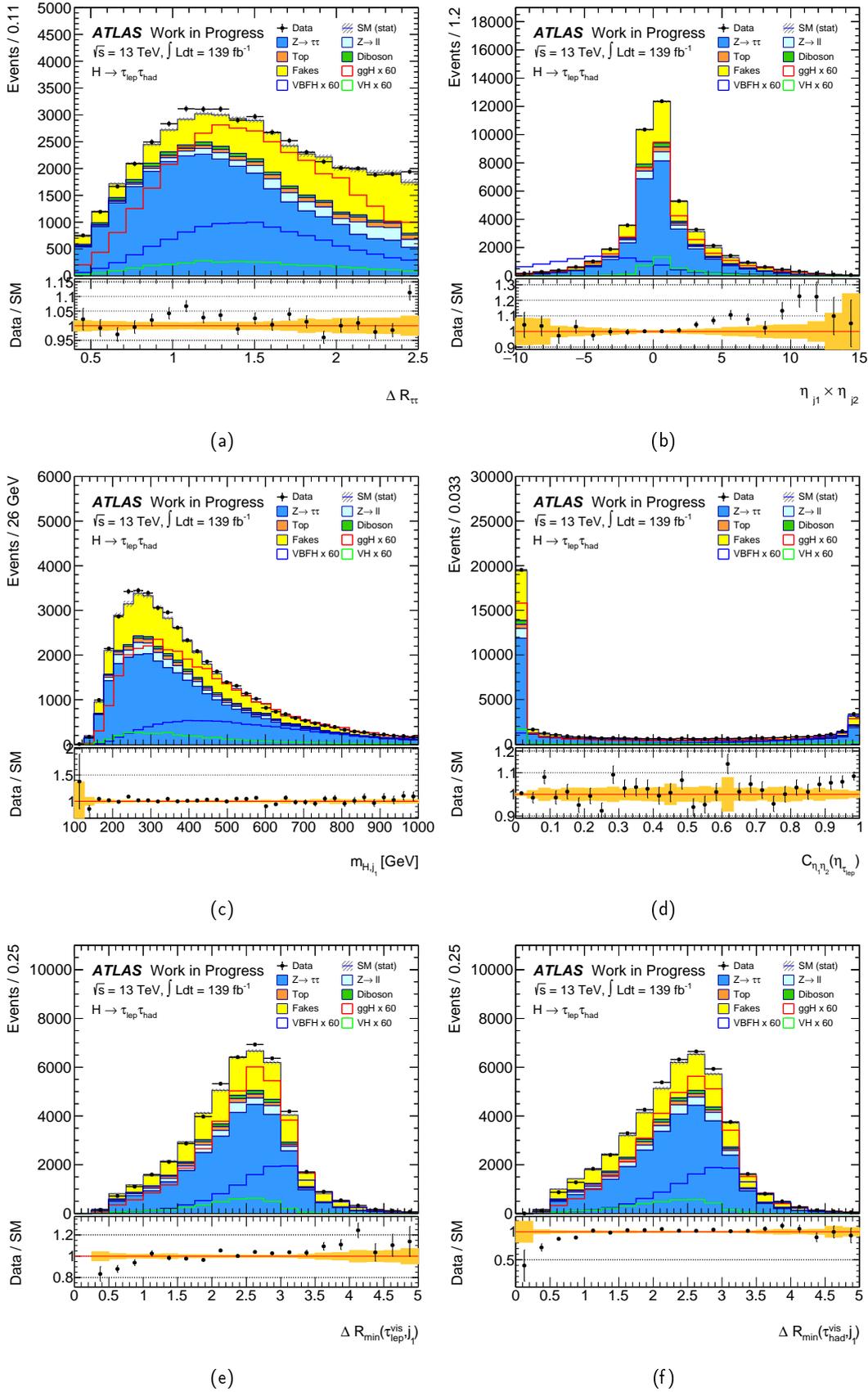
FIGURE 7.5: Distribution of discriminating variables used in the BDT training process. Only statistical uncertainties are shown.

(a)

(b)

(c)

(d)

(e)

(f)
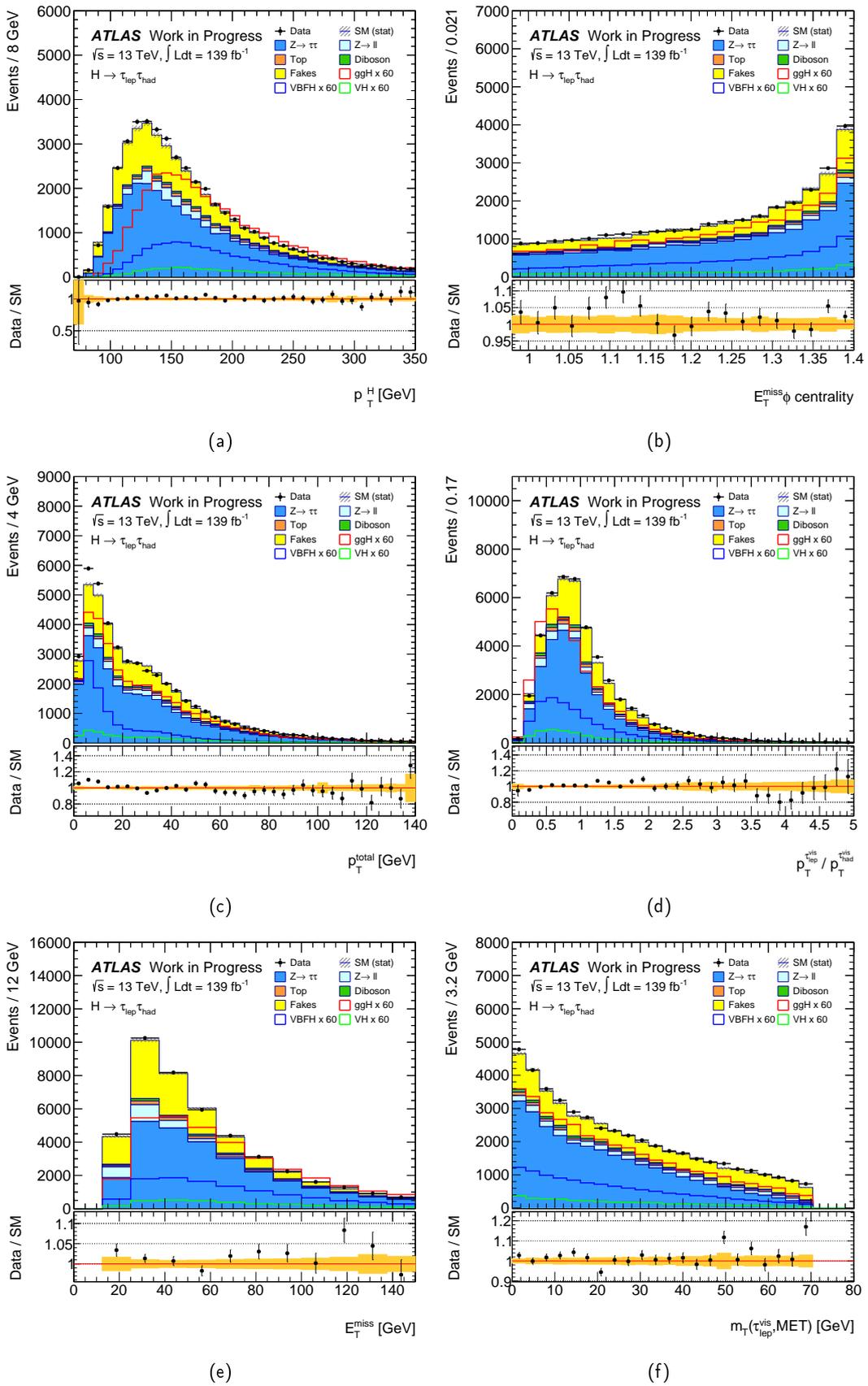
FIGURE 7.6: Distribution of discriminating variables used in the BDT training process. Only statistical uncertainties are shown.

(a)   (b)
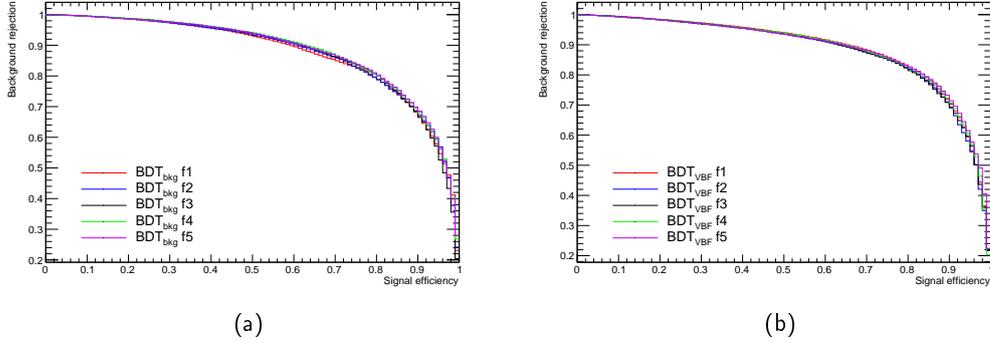
FIGURE 7.7: ROC curve for both BDT trainings on the left
side (a) for BDT$_{bkg}$ and on the right side (b) for BDT$_{VBF}$; and
BDT$_{bkg, VBF}$ f1, 2, 3, 4, 5 are described in the text.

(a)   (b)
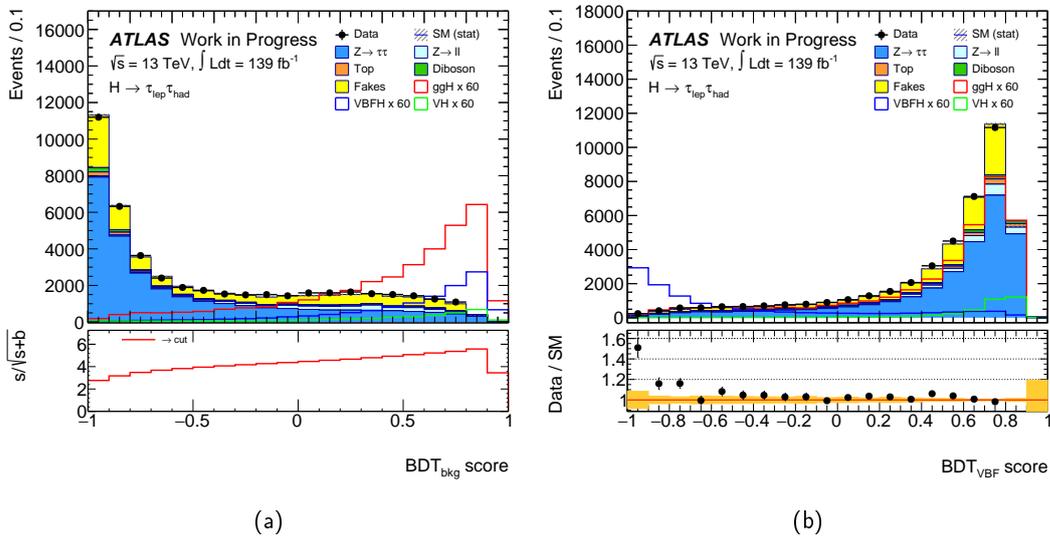
FIGURE 7.8: Distributions of the BDT output for signal and back-
ground processes: (a) BDT$_{bkg}$ and (b) BDT$_{VBF}$. Only statistical uncer-
tainties are shown. The lower panel shows in (a) signal significance
($s/\sqrt{s+b}$) after BDT training.

7.3.3 Signal Region Definition

After the preselection criteria, the signal to the background ratio is 0.013, and the significance defined as $s/\sqrt{s+b}$ corresponds to 2.76 ± 0.02. In order to achieve higher sensitivity to probe anomalous couplings in the signal process, the BDT method is used. Selecting signal events by cutting on the BDT scores, allows to increase the signal-to-background ratio.
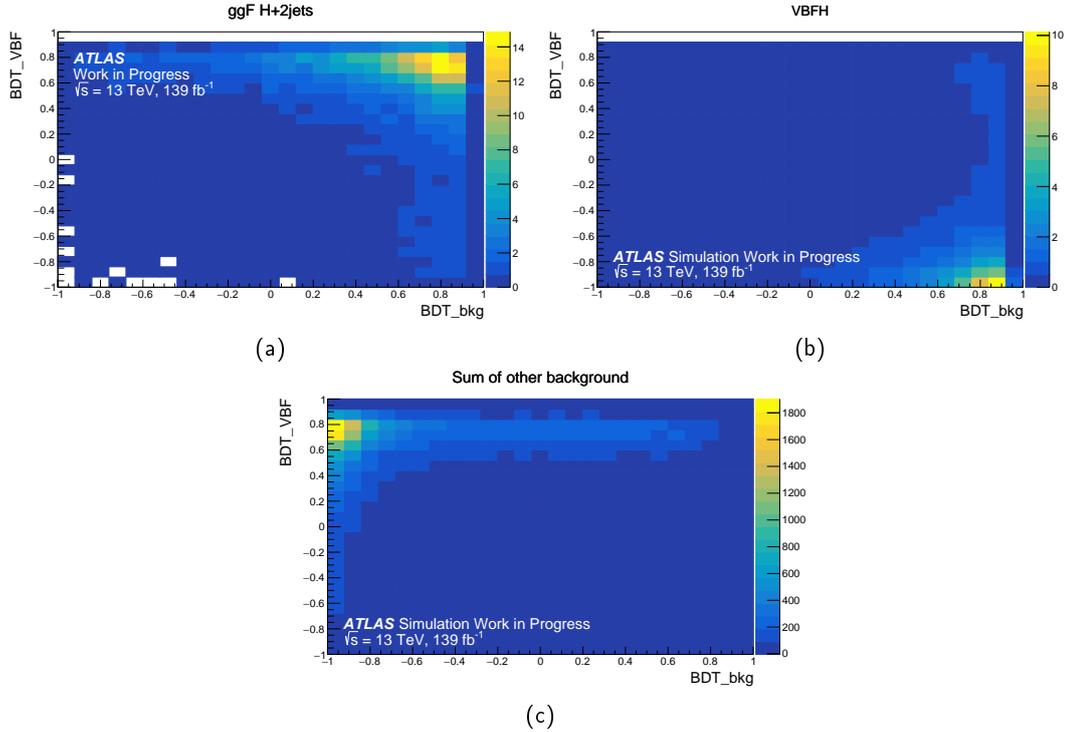
FIGURE 7.9: Correlations between the BDT_{bkg} and BDT_{VBF} output for (a) signal process, (b) VBF Higgs boson production, (c) all other backgrounds are shown.

Fig. 7.9 shows the correlation between two BDT trainings: BDT_{bkg} and BDT_{VBF}. The signal ggF H+2 jets is classified as a signal for both BDT training. The VBFH, H$\rightarrow \tau\tau$ background shows the correlations are depicted in the right bottom corner of Fig. 7.9b. Furthermore, the sum of backgrounds without VBFH is shown in Fig. 7.9c and correlation is illustrated in the left top corner of the plot. The optimal cut values on BDT_{bkg} and BDT_{VBF} are found by maximizing the signal-to-background ratio. The BDT_{bkg} output is scanned between 0.6 and 0.8 in step of 0.02. While, BDT_{bkg} is scanned between -0.5 and 0.0 in step of 0.1. The reason is the significance is dropped at 0.9 of the BDT_{bkg} as shown in Fig. 7.8a. The maximum signal to background ratio is obtained as 0.187 ± 0.005 with the significance of 4.52 ± 0.08 at the $BDT_{bkg} > 0.8$ while $BDT_{VBF} > -0.4$. Before the BDT cut, VBFH background events are 34 % of the whole signal. After the BDT cut, VBFH events are 13 % of the whole signal. The signal is reduced by 80 %, and VBFH is reduced by 92 % after BDT selection. Moreover, 98.6 % of the background of the whole is reduced by this BDT selection. The expected event yield for signal and background are shown in the Table. 7.5.

The dominant background contributions are $Z/\gamma^* \rightarrow \tau\tau$ with 48 %, and fake background with 32 %. The distributions of the CP-sensitive observables are shown

Process	Event yield
ggF H+2 jets, H$\to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	115.92 $\pm$ 1.61
VBFH, H$\to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	15.39 $\pm$ 0.16
VH, H$\to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	12.98 $\pm$ 0.32
Z$\to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$	296.32 $\pm$ 14.04
Z$\to \ell\ell$	33.50 $\pm$ 5.56
Top	37.69 $\pm$ 2.55
Diboson	25.26 $\pm$ 0.73
Fakes	199.74 $\pm$ 4.82
$\sum_{\mathrm{bkg}}$	620.87 $\pm$ 16.07
Data	723.00 $\pm$ 26.89
s/b	0.19 $\pm$ 0.01
s/$\sqrt{s+b}$	4.52 $\pm$ 0.08

TABLE 7.5: Summary of the event yield of the high BDT signal region for simulated samples and observed data at $\sqrt{s} = 13$ TeV with $\mathcal{L} = 139\,\mathrm{fb}^{-1}$. Only statistical uncertainties are shown.

in Fig. 7.10.

(a)

(b)

(c)
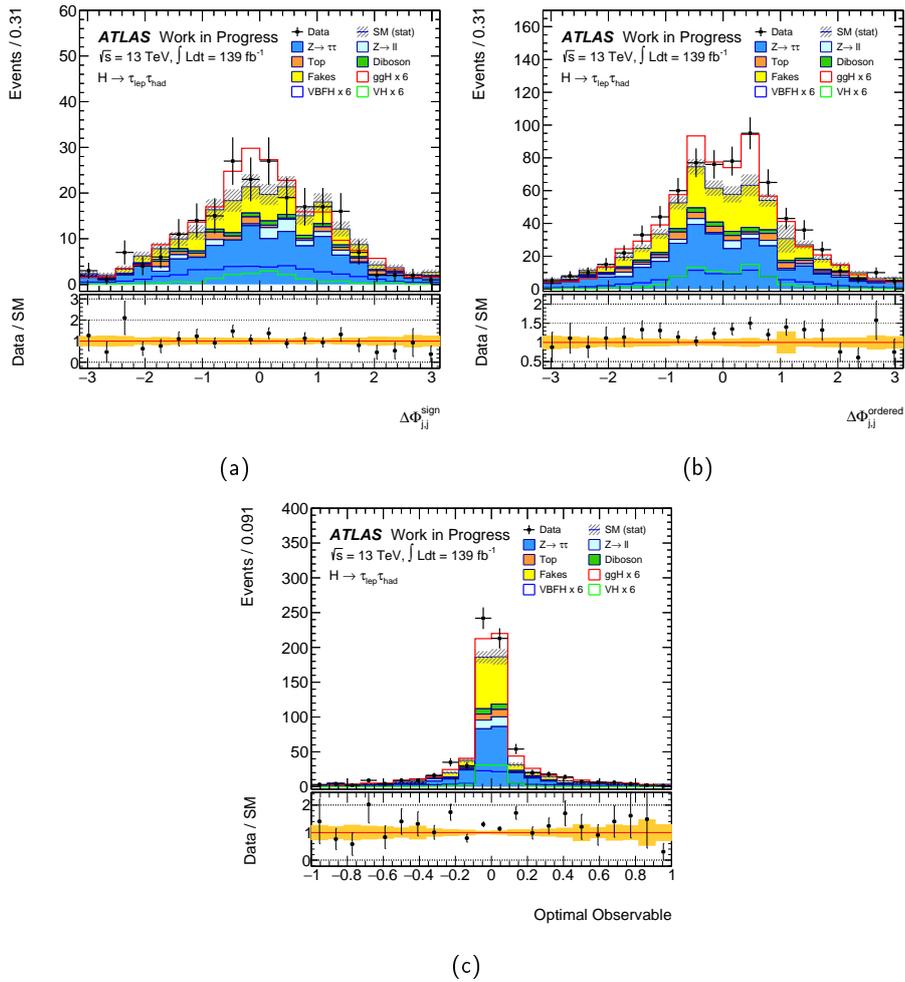
FIGURE 7.10: CP-sensitive observable distributions: (a) $\Delta\Phi_{jj}^{\text{sign}}$ (b) $\Delta\Phi_{jj}^{\text{ordered}}$ (c) Optimal Observable in the signal region with statistical uncertainties at $\sqrt{s} = 13$ TeV with $\mathcal{L} = 139\text{fb}^{-1}$.

Chapter 8

Statistical Analysis and Results

In order to determine the sensitivity of the analysis to effective anomalous couplings of ggF H+2 jets the decay H$\to \tau_{\text{lep}}\tau_{\text{had}}$ is used. As discussed in the previous chapters, the sensitivity to the effective coupling is exploited analyzing different CP-odd observables: the Optimal Observable, $\Delta\Phi_{jj}^{\text{sign}}$ and $\Delta\Phi_{jj}^{\text{ordered}}$. A profile likelihood fit is used to determinate the sensitivity and to derive expected and observed confidence intervals on $\tan\alpha$.

This chapter gives an overview of the profile likelihood fit function, which is used in this analysis. Furthermore, the fitting procedure is explained and the expected and observed results are presented.

8.1 Likelihood Function

A fit in the CP-odd observable distributions is performed using the profile likelihood method. This method includes signal strength and background normalization factors as free fit parameters. The signal strength μ_{sig} is the cross section ratio between the observed signal events and SM expectation. The likelihood fit is implemented in the software framework called TRExFitter. This software framework includes the HistFactory tool [106].

The likelihood function also includes so-called nuisance parameters ($\boldsymbol{\theta}$). These parameters are included to account for the impact of systematic uncertainties on the predicted signal and background yields. The likelihood function [107, 108], which is maximized to find the best parameter values, is given by,

$$
\begin{aligned}
L(\mathbf{n}, \boldsymbol{\theta^0}|\mu_{\text{sig}}, \tan\alpha, \mathbf{b}, \boldsymbol{\theta}) &= \prod_{i \in SR} P(n_i|\lambda_i(\mu_{\text{sig}}, \tan\alpha, \mathbf{b}, \boldsymbol{\theta})) \\
&\times \prod_{i \in CR} P(n_i|\lambda_i(\mu_{\text{sig}}, \tan\alpha, \mathbf{b}, \boldsymbol{\theta})) \\
&\times C_{\text{syst}}(\boldsymbol{\theta^0}, \boldsymbol{\theta}),
\end{aligned}
\tag{8.1}
$$

where $P(n_i|\lambda_i(\mu_{\text{sig}}, \tan\alpha, \mathbf{b}, \boldsymbol{\theta}))$ is the Poisson distribution, which gives the probability to find the number of observed events n_i. Index i runs over all bins for each observables distributions in the signal (SR) and control regions (CR). Here, $\lambda_i(\mu_{\text{sig}}, \tan\alpha, \mathbf{b}, \boldsymbol{\theta}) = \mu_{\text{sig}} \cdot s_i(\boldsymbol{\theta}, \tan\alpha) + b_i(\boldsymbol{\theta})$, where s and b is expected event yield of signal and background events in bin i, and parameter of interest (POI) is the $\tan\alpha$. The s_i depends on the $\tan\alpha$, and $\sum s_i$ is fixed to the SM value. And the $\mathbf{b}$ is the normalization factors for the background. The bold symbol indicates the several nuisance parameters are in the fit. In this analysis, normalization factor for $Z \to \tau\tau$ background is used. The last term of Eq. 8.1 describes systematic uncertainties

which are constrained by the Gaussian function. It can be written as,

$$C_{\text{syst}}(\boldsymbol{\theta^0}, \boldsymbol{\theta}) = \prod_{j \in S} G(\theta_j^0 - \theta_j), \tag{8.2}$$

where $G(\boldsymbol{\theta_j^0} - \boldsymbol{\theta_j})$ is the Gaussian function with in unit width and index j is runs over all systematic uncertainties (S). Here, θ_j^0 is the nominal value of the systematic uncertainty. And θ_j symbolizes the fitted parameter of the systematic uncertainties. The systematic uncertainties are initially uncorrelated to each other in the likelihood function but their correlations are determined in the process of the fit.

This analysis uses the negative logarithmic likelihood function and it can be written as

$$\text{NLL}_{\text{min}} = \min\{-\ln L(\mu_{\text{sig}}, \tan \alpha, \boldsymbol{\theta})\} = \ln L(\hat{\mu}_{\text{sig}}, \hat{\tan} \alpha, \hat{\boldsymbol{\theta}}), \tag{8.3}$$

where $\hat{\mu}_{\text{sig}}$, $\hat{\tan} \alpha$, $\hat{\boldsymbol{\theta}}$ are the best fit values, which are obtained from minimizing the negative logarithm of the likelihood function.

8.1.1 Nuisance parameters

In this analysis, the simplified systematic uncertainties are used based on [102]. As described earlier, expected event yields of signal and background depend on the nuisance parameters. In this analysis, following simplified systematic uncertainties are considered. All nuisance parameters are constrained by Gaussian function.

- **Luminosity:** The integrated luminosity uncertainty is ± 2.9 % for the data taking in Run 2.

- **Jet Energy Scale (JES):** This uncertainties applied to the signal and backgrounds. JES uncertainties have impact on signal events with up 0.4 % and downward 0.5% fluctuation. And impact on background events for $Z \rightarrow \tau\tau$ with 2.8 %/-3.4 %, for $Z \rightarrow \ell\ell$ with 4.3 %/-3.2 %, for top quark background with -4.1 %/4.5 % and for the other backgrounds with 0.9 %/-1.3 % in respective up/downward fluctuations.

- **Tau Energy Scale (TES):** The impact on this uncertainties to the signal ggF H+2 jets is up 1.3 % and downward 0.7 %. And the up/downwards fluctuations for the backgrounds are $Z \rightarrow \tau\tau$ with ± 2.2 %, top quark background with 2.3 %/-6.1 %, and for other backgrounds with ± 1.7 %, respectively.

- **Tau reconstruction and identification efficiency:** Impact of this systematic uncertainties for the signal is approximately ± 6.5 %, for backgrounds are less than 6.6 %.

- **Normalization for the background:** The uncertainty on the fake background is 5 % and $Z \rightarrow \ell\ell$ is less than 10 % variation.

- **MC statistical uncertainties:** Monte Carlo statistical uncertainties gammas for each bin of the signal and control region is taken into account to the nuisance parameters.

- **QCD scale uncertainty:** The cross section uncertainties for the signal is up 16.4 % and for $Z \rightarrow \tau\tau$ is up to 3.4 %.

8.2   Fitting procedure

The strategy of this analysis is to determine the sensitivity to the effective couplings of the Higgs boson to gluons by performing a profile likelihood fit to the CP-odd observables. The likelihood fit is performed to the $\Delta\Phi_{jj}^{\text{sign}}$, $\Delta\Phi_{jj}^{\text{ordered}}$, and Optimal Observable distributions in the signal region. This signal region distributions are shown in Fig. 8.1a; 8.1b; 8.1c, respectively. The signal region is defined by cutting on the two different BDT outputs as discussed in Sec. 7.3.3. And the signal region is defined as $\text{BDT}_{\text{bkg}} > 0.8$ and $\text{BDT}_{\text{VBF}} > -0.4$. The shape of CP-odd observables is used in fitting procedure. Therefore, the binning of the observables are chosen to contain as much information as possible as long as no bins have negative content due to negative Monte Carlo weights. Because bins with the negative content are unphysical.

And rest of the two BDT outputs ($\text{BDT}_{\text{bkg}} < 0.8$ and $\text{BDT}_{\text{VBF}} < -0.4$) are considered as a control region. The distribution of $m_{\text{H}}^{\text{MMC}}$ in control region is shown in Fig. 8.2. This distribution of invariant mass of the Higgs boson calculated by MMC method is used for the control region because it allows to discriminate between $Z \to \tau\tau$ and $H \to \tau\tau$. This control region is defined in order to constrain the background contribution from $Z \to \tau\tau$.

The SM ggF H+2 jets signal sample is reweighted for various $\tan\alpha$ values as described in Sec. 4.1. Likelihood fits for different $\tan\alpha$ points use the templates for all $\tan\alpha$ values and continuous parameters that representing tangent in the fit. These templates contain the histogram of the CP-odd observables for different $\tan\alpha$. Then fit takes two nearest $\tan\alpha$ histograms, which are correspond to the current value of $\tan\alpha$ being fitted. Furthermore, fit in a bin by bin does a linear interpolation of the histograms. All other histograms that are too far from the current $\tan\alpha$ value are scaled to zero. In another word, the $s_i(\tan\alpha)$ are obtained from the reweighting method and that then a fit in $\tan\alpha$ is performed. The $s_i(\tan\alpha)$ from the reweighting and fit is performed in $\tan\alpha$ values.

The confidence intervals of $[a, b]$ at 68.3 % confidence interval (CL) and $[c, d]$ at 95 % CL are obtained from the conditions:

$$\ln L(\hat{\tan}\alpha) - \ln L(a/b) = -0.5, \qquad (8.4)$$
$$\ln L(\hat{\tan}\alpha) - \ln L(c/d) = -1.92.$$

In the fitting procedure, the signal strength μ_{sig} is left free floating and only the shape of the observables are used. To determine the expected confidence intervals for $\tan\alpha$ an Asimov data is constructed. The Asimov dataset is constructed by the SM expectation for signal in addition with background in the signal and control regions. This CP analysis considers simplified systematic uncertainties according to [102], which are listed in Sec. 8.1.1. The results will be compared between statistic uncertainty only and including the effect of systematic uncertainties. This comparison shows by how much systematic uncertainties effect the final results. The event yields with statistical and inclusion of systematic uncertainties for each observables are summarized in Table 8.1 for the signal region, and in Table 8.2 for the control region. The event yield for $\Delta\Phi_{jj}^{\text{sign}}$ is decreased with respect to $\Delta\Phi_{jj}^{\text{ordered}}$, $\mathcal{OO}$due to requirement for the two tagging jets to be in different hemispheres.

The analysis uses the full Run 2 dataset to determine the observed sensitivity. A detailed discussion of the results is in the following sections.

(a)   (b)

(c)
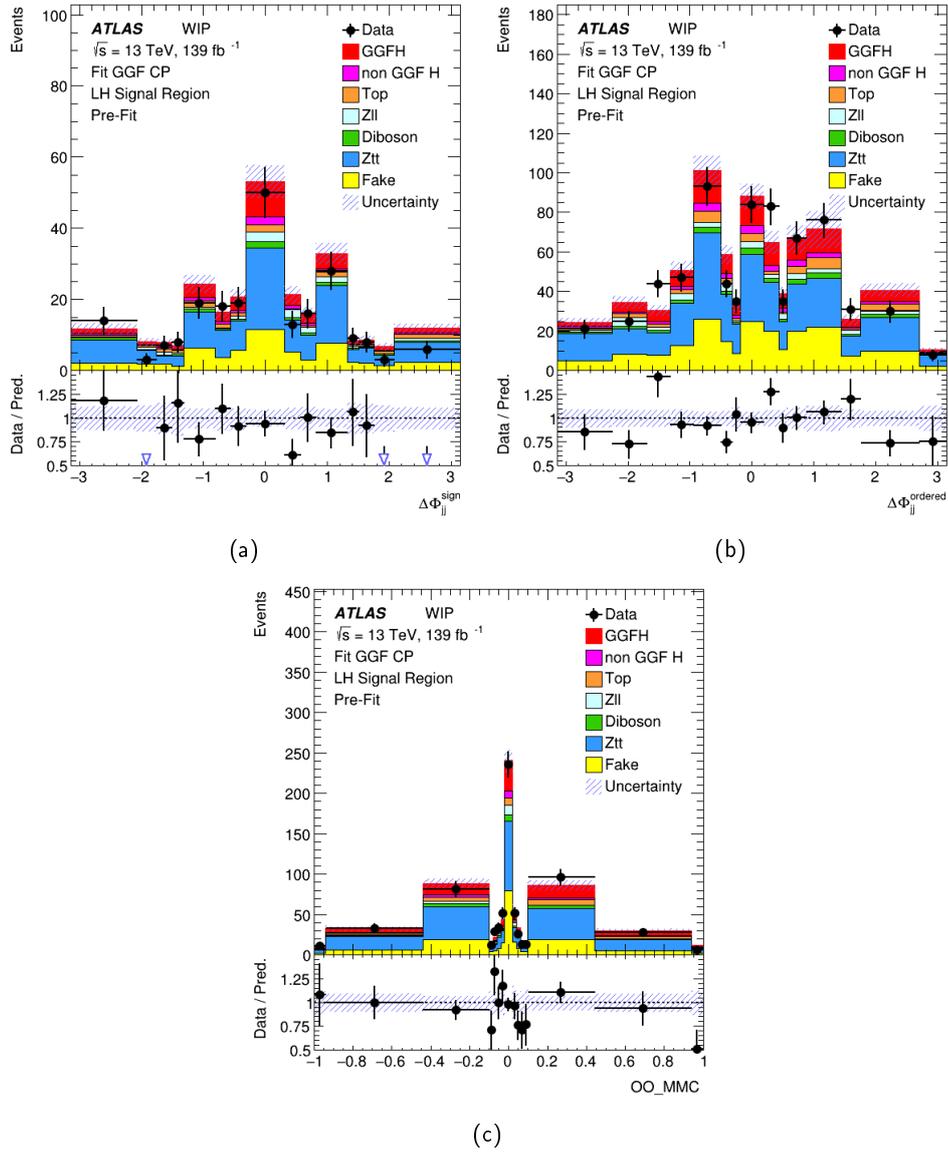
FIGURE 8.1: Prefit distributions of the (a) $\Delta\Phi_{jj}^{\text{sign}}$, (b) $\Delta\Phi_{jj}^{\text{ordered}}$, (c) Optimal Observable in the signal region with statistical and systematic uncertainties are shown.
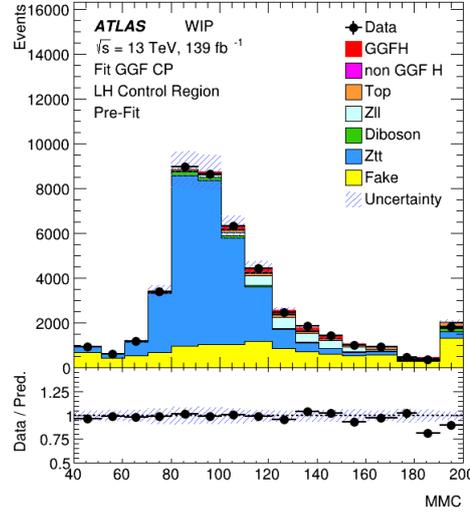
FIGURE 8.2: The prefit distribution of m_H^{MMC} is shown in the control region with statistical and systematic uncertainties.

	Event yield in Signal region					
	$\Delta\Phi_{jj}^{\mathrm{sign}}$	Uncertainty		$\mathcal{OO}(\Delta\Phi_{jj}^{\mathrm{ordered}})$	Uncertainty	
Process		stat.	stat.+syst.		stat.	stat.+syst.
VH	$2.70 \pm$	$0.11 \pm$	0.29	12.98	$\pm\ 0.32 \pm$	1.24
Top	$12.64 \pm$	$0.50 \pm$	1.93	37.69	$\pm\ 2.55 \pm$	4.58
$Z \to \ell\ell$	$13.54 \pm$	$0.54 \pm$	3.49	33.50	$\pm\ 5.56 \pm$	7.33
Diboson	$9.01 \pm$	$0.33 \pm$	0.94	25.26	$\pm\ 0.73 \pm$	2.50
$Z \to \tau\tau$	$112.85 \pm$	$3.70 \pm$	13.30	296.32	$\pm\ 14.04 \pm$	34.32
Fake	$58.82 \pm$	$2.15 \pm$	4.31	199.74	$\pm\ 4.82 \pm$	12.51
VBFH	$6.95 \pm$	$0.25 \pm$	0.65	15.39	$\pm\ 0.16 \pm$	1.43
ggF H+2 jets	$39.27 \pm$	$1.14 \pm$	1.14	121.30	$\pm\ 1.61 \pm$	3.52
Total	$255.78 \pm$	$10.68 \pm$	14.63	742.18	$\pm\ 16.16 \pm$	37.83

TABLE 8.1: Prefit event yields with statistic and with effect of systematic uncertainties of the analysis for CP-odd observables in the signal region. Here simplified compatible systematic uncertainties are considered [102].

| | Event yield in Control Region | | |
| | | | Uncertainty |
Process			stat only	syst + stat
VH	45.39 ±		0.77 ±	4.23
Top	1182.13 ±		17.71 ±	120.43
Z→ $\ell\ell$	2435.29 ±		41.82 ±	366.11
Diboson	894.76 ±		13.87 ±	85.11
Z→ $\tau\tau$	27 811.80 ±		245.72 ±	2941.68
Fake	11 965.80 ±		181.96 ±	692.87
VBFH	184.20 ±		3.26 ±	17.02
ggF H+2 jets	893.65 ±		221.96 ±	25.92
Total	45 413.02 ±		380.82 ±	3048.00

TABLE 8.2: Prefit event yields with statistic and simplified compatible systematic uncertainties [102] of the analysis for control region is summarized.

8.3 Results

The profile likelihood function as described in Eq. 8.4 is used to fit to the different CP-odd observables. The negative profile likelihood allows to derive an exclusion limits on the $\tan \alpha$ value parametrizing the anomalous coupling between Higgs boson and gluons. During the fit, the best fit value of signal strength μ_{sig} is also determined. These results are in general based on statistical uncertainty. As an additional step simplified systematic uncertainties are considered to determine the impact on the results. The expected results will be discussed in Sec. 8.3.1 using the Asimov data set

8.3.1 Expected confidence intervals

The likelihood fit is first performed to the Asimov dataset in order to get expected results for a SM signal hypothesis. Another purpose of this fit is to validate the fit setup. Thus, in this fit, SM signal and background prediction in signal and control regions is taken into account. With this, the signal strength and normalization factors are expected to be one in the Asimov fit.
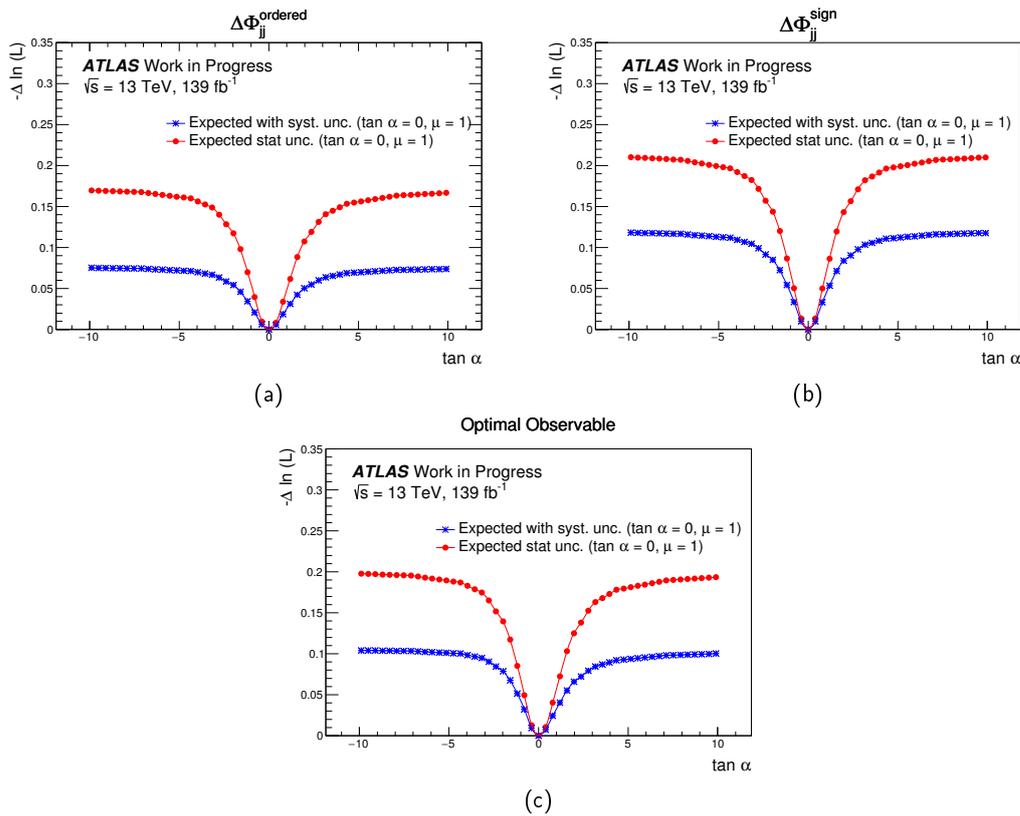
FIGURE 8.3: Expected negative $\Delta \ln(L)$ curves as function of $\tan \alpha$ values defining the underlying signal hypothesis with simplified systematic uncertainties (blue) [102] and statistical uncertainty only (red) for (a) $\Delta \Phi_{jj}^{\text{ordered}}$ (b) $\Delta \Phi_{jj}^{\text{sign}}$ (c) Optimal Observable . The Asimov data is constructed with signal strength of $\mu_{\text{sig}} = 1$ and $\tan \alpha = 0$.

Result for fits on Asimov dataset are shown in Fig. 8.3. Here, each $\Delta \ln(L)$ curves are compared for the inclusion of systematic uncertainties and the statistical uncertainties only. The minimum of $\Delta \ln(L)$ is obtained at $\tan \alpha = 0$ for all CP-odd observables as predicted in the SM.

The best fit signal strength in $\tan \alpha = 0$ value for CP-odd observables with impact of systematics is obtained. The signal strength $\hat{\mu}_{\mathrm{sig}} = 1.00 \pm 0.24$ and normalization factor $\hat{b}_{Z\tau\tau} = 1.00 \pm 0.11$ for $\Delta \Phi_{jj}^{\mathrm{ordered}}$ and for Optimal Observable are obtained, whereas $\hat{\mu}_{\mathrm{sig}} = 1.00 \pm 0.33$ and normalization factor $\hat{b}_{Z\tau\tau} = 1.00 \pm 0.10$ for $\Delta \Phi_{jj}^{\mathrm{sign}}$ is obtained.

As shown in Fig. 8.3a, fit to the $\Delta \Phi_{jj}^{\mathrm{ordered}}$ with systematic uncertainties, curve reaches a plateau of a $-\Delta \ln(L) = 0.07$. Whereas, fit result where considers only statistical uncertainties, $-\Delta \ln(L) = 0.17$ is obtained. The expected results of the fit to $\Delta \Phi_{jj}^{\mathrm{ordered}}$ does not reach $\pm 1\sigma$ confidence interval.

A negative $\Delta \ln(L)$ curve of the $\Delta \Phi_{jj}^{\mathrm{sign}}$ as illustrated in Fig. 8.3b. With inclusion of systematic uncertainties, the expected value of $-\Delta \ln(L) = 0.12$ is obtained, while expected results that only statistical uncertainties, curve reaches a plateau of $-\Delta \ln(L) = 0.21$. Therefore, no exclusion limit is found.

The $-\Delta \ln(L) = 0.10$ is obtained for fit to the Optimal Observable with effect of the systematic uncertainties as shown in Fig. 8.3c. On the other hand, fit with only statistical uncertainties, expected value of $-\Delta \ln(L) = 0.20$ is obtained. Also, due to limited sensitivity to effective coupling of Higgs boson to gluons, the exclusion limits can not be obtained.

These expected results show that $\Delta \Phi_{jj}^{\mathrm{sign}}$ observable is more sensitive than others. The reasons, why $\Delta \Phi_{jj}^{\mathrm{sign}}$ appears to be more sensitive in the fit could be caused by the binning choice or influence from the LO reweighting method to the NLO event sample.

The likelihood fits to the CP-odd observables do not allow to derive exclusion limit due to the limited sensitivity.

The postfit distributions with statistical and systematic uncertainties of observables in signal and control regions are shown in Fig. 8.4; 8.5; 8.6. And their respective event yields are summarized in Table 8.3; 8.4; 8.5.

	$\Delta \Phi_{jj}^{\mathrm{sign}}$ in SR	Control Region
VH	2.70 ± 0.26	45.39 ± 4.11
Top	12.64 ± 1.24	1182.13 ± 108.41
Zll	13.54 ± 1.61	2435.29 ± 257.04
Diboson	9.01 ± 0.82	894.76 ± 73.67
Ztt	112.85 ± 3.78	$27\,811.80 \pm 256.64$
Fake	58.82 ± 2.40	$11\,965.80 \pm 291.05$
VBFH	6.95 ± 0.67	184.20 ± 16.70
GGFH	39.27 ± 12.76	893.65 ± 290.43
Total	255.79 ± 11.40	$45\,412.90 \pm 212.25$
Data	221	44894

TABLE 8.3: Postfit event yields from the fit in the $\Delta \Phi_{jj}^{\mathrm{sign}}$ in signal region and control region. Statistical and systematic uncertainties are included. Data is shown for references and it is not used in the fit.
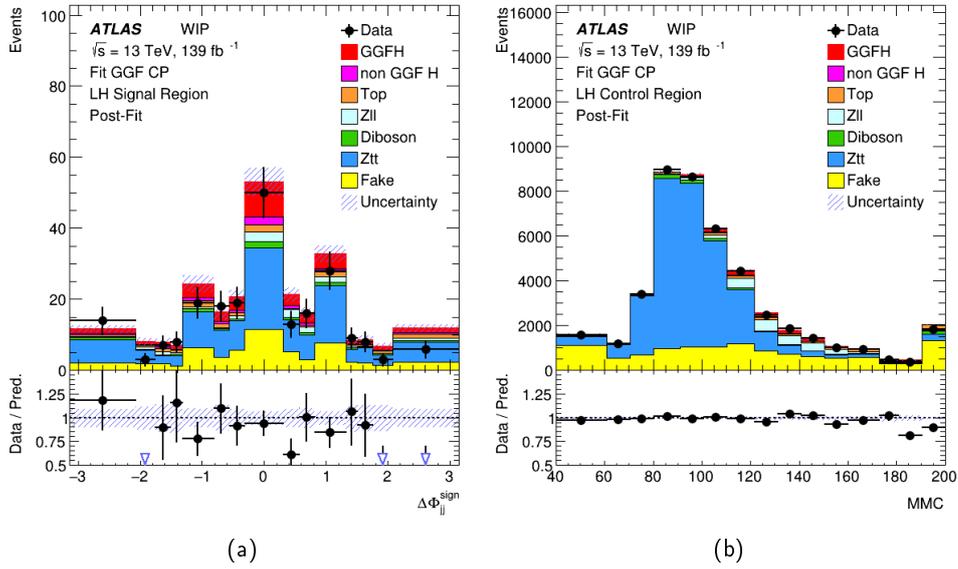
FIGURE 8.4: Postfit distributions of (a) $\Delta\Phi_{jj}^{\text{sign}}$ in the signal region and (b) m_H^{MMC} distribution with statistical and systematic uncertainties are shown. In all plots, WIP abbreviates the Work in Progress.
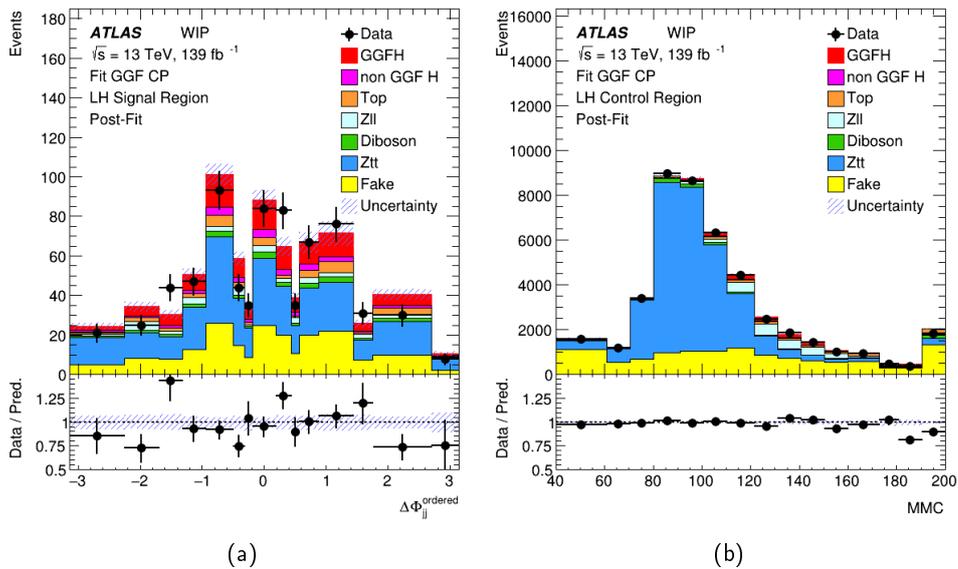
FIGURE 8.5: Postfit distributions of (a) $\Delta\Phi_{jj}^{\text{ordered}}$ in the control region, (b) m_H^{MMC} distribution with statistical and systematic uncertainties are shown. In all plots, WIP abbreviates the Work in Progress.

	$\Delta\Phi_{jj}^{\text{ordered}}$ in SR	Control Region
VH	12.98 ± 1.22	45.39 ± 4.11
Top	37.69 ± 3.59	1182.13 ± 108.48
Zll	33.50 ± 3.48	2435.29 ± 226.43
Diboson	25.26 ± 2.19	894.76 ± 71.64
Ztt	296.32 ± 7.59	$27\,811.80 \pm 254.04$
Fake	199.74 ± 7.05	$11\,965.80 \pm 291.73$
VBFH	15.39 ± 1.44	184.20 ± 16.70
GGFH	121.30 ± 29.26	839.65 ± 215.55
Total	742.17 ± 23.71	$45\,412.90 \pm 212.53$
Data	723	44894

TABLE 8.4: Postfit event yields from the fit in the $\Delta\Phi_{jj}^{\text{ordered}}$ in signal region and control region. Statistical and systematic uncertainties are included. Data is shown for references and it is not used in the fit.

(a)   (b)
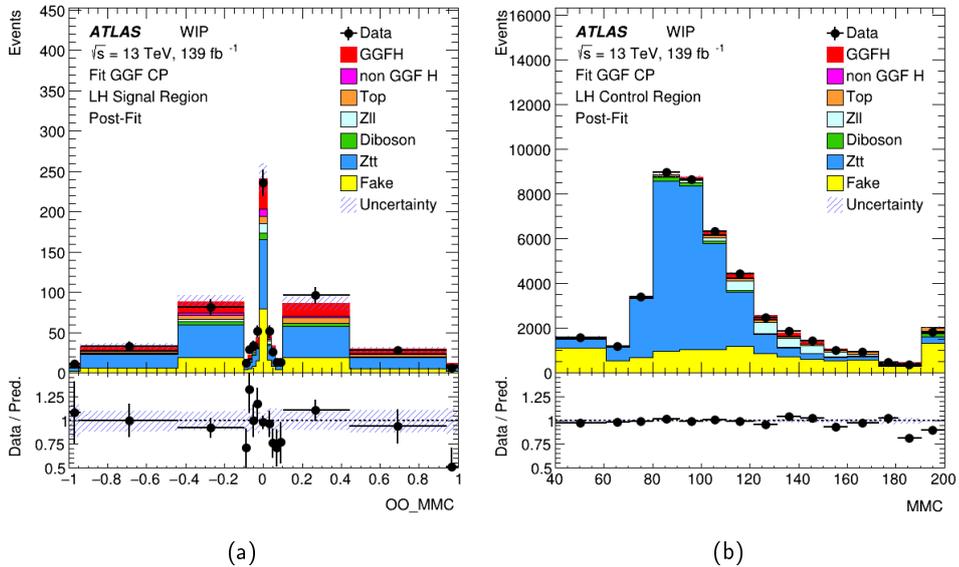
FIGURE 8.6: The postfit distribution of (a) Optimal Observable distribution in the signal region and (b) $m_{\text{H}}^{\text{MMC}}$ distribution in control region with statistic and systematic uncertainties. In the plot, WIP abbreviates the Work in Progress.

	$\mathcal{OO}$ in SR		Control Region	
VH	12.98 $\pm$	1.23	45.39 $\pm$	4.11
Top	37.69 $\pm$	3.56	1182.13 $\pm$	108.44
Zll	33.50 $\pm$	3.51	2435.29 $\pm$	225.64
Diboson	25.26 $\pm$	2.18	894.76 $\pm$	71.67
Ztt	296.32 $\pm$	7.47	27 811.80 $\pm$	253.72
Fake	199.74 $\pm$	7.00	11 965.80 $\pm$	291.29
VBFH	15.39 $\pm$	1.44	184.20 $\pm$	16.70
GGFH	121.30 $\pm$	29.13	893.65 $\pm$	214.64
Total	742.17 $\pm$	62.02	45 412.90 $\pm$	472.85
Data	723		44894	

TABLE 8.5: Postfit event yields from fit in the Optimal Observable in signal region and control region. Statistical and systematic uncertainties are included. Data is shown for references and it is not used in the fit.

Chapter 9

Conclusion and Outlook

This thesis presents the study of the CP property of the Higgs boson production in the gluon fusion in association with two jets, ggF H+2 jets. The sensitivity of the effective anomalous coupling is studied by using the decay channel of the Higgs boson H→ $\tau_{\text{lep}}\tau_{\text{had}}$. The interaction of the Higgs boson to gluons allows testing the CP-invariance directly. The relative contribution between CP-even and CP-odd interactions is described by the mixing angle of $\tan \alpha$, which has zero value in the Standard Model (SM). Non zero value will describe CP-violation in the coupling of the Higgs boson to gluons.

In order to test the CP-invariance, this analysis uses CP-odd observables, which are CP-odd under the transformation of the charge conjugation and parity transformation. Three different CP-odd observables are used to extract the signal ggF H+2 jets sensitivity. The Optimal Observable, which uses the full phase space information based on the matrix element of the physics processes, is used. In addition, the azimuthal angle between two tagging jets, which are selected in the opposite hemispheres or ordered by their pseudorapidity, is investigated. This thesis is based on the proton-proton collisions data at the center of mass energy of $\sqrt{s} = 13$ TeV with an integrated luminosity of $\mathcal{L} = 139\,\text{fb}^{-1}$ collected by the ATLAS experiment at the LHC, CERN.

By studying the CP-odd observables, sensitivity to the effective anomalous coupling of the signal ggF H+2 jets is investigated at the generator level, which is ignoring the background processes and systematic uncertainties. The mean value of the CP-odd observables for various $\tan \alpha$ coupling models can be used to illustrate the sensitivity of the observable to CP-violation effects. Various $\tan \alpha$ coupling models on the signal ggF H+2 jets are obtained using a leading order (LO) reweighting technique. The reweighting technique is validated at LO generated events with MadGraph5 [39] and used in the NLO events sample produced by Powheg [47–50].

A mean value which is compatible to zero is obtained in all CP-odd observables in the CP conserving coupling model $\tan \alpha = 0$. While mean value deviating from zero of the CP-odd observables are obtained in nonzero values of $\tan \alpha$. The Optimal Observable, $\Delta\Phi_{jj}^{\text{sign}}$ and $\Delta\Phi_{jj}^{\text{ordered}}$ show a good separation between various anomalous couplings. The Optimal Observable shows larger sensitivity than other observables. Furthermore, $\Delta\Phi_{jj}^{\text{sign}}$ shows broader deviations of the mean value from zero than the η-ordered $\Delta\Phi_{jj}^{\text{ordered}}$. However, $\Delta\Phi_{jj}^{\text{sign}}$ has large statistical uncertainty because of the jet selection in opposite hemispheres.

Also, a sensitivity study on different subprocesses of the gluon fusion is studied. The subprocesses are classified as gluon-gluon, gluon-quark, and quark-quark initial states depending on the flavor combination of initial state partons. The dominant

contribution comes from the gluon initiated process (97.2 %); however, the quark-quark initiated process (2.8 %) gives the largest deviations in the mean value of the CP-odd observables.

A comparison of LO and NLO events are studied when applying the LO reweighting technique. In the NLO sample, the sensitivity of different subprocesses to the anomalous coupling of the Higgs boson and gluons are reduced due to the LO reweighting. Probably a proper NLO reweighting method can increase the sensitivity to the anomalous coupling in further analysis.

The sensitivity to ggF H+2 jets is also investigated at detector level. In this analysis, the dominant background contribution comes from $Z \to \tau\tau$ production and from jets misidentified as hadronically decaying τ-leptons backgrounds. Higgs production in vector boson fusion and associated vector boson production modes are considered as a background. Measurement of the $\tan\alpha$ is obtained in the data by reweighting the simulated signal events for different coupling models.

The selection criteria are applied to the simulated events and the data. Furthermore, the Boosted Decision Tree (BDT) method, which is a multivariate analysis, is used. Two distinctive BDT are trained in order to achieve good separation between signal and background events. One BDT is trained for signal against all backgrounds; a second BDT is trained to separate ggF from VBF production. The signal region is defined by defining the threshold value for both BDT outputs. This threshold value has been chosen to optimize the signal-to-background ratio. It reduces the contribution of the VBFH background. As a result, the dominant contribution arises from $Z/\gamma \to \tau\tau$ with 48% and fake background with 32%. A signal-to-background ratio of 0.19 is obtained with a significance of 4.52.

To measure the relative contribution of the CP-even and CP-odd in Higgs boson coupling to gluons, a profile likelihood fit function is used. The negative logarithmic profile likelihood fit is performed to the CP-odd observables in order to test the signal hypothesis as a function of $\tan\alpha$. The distributions of CP-odd observables are used to define the signal region. The invariant mass of the Higgs boson distribution estimated by the MMC method is used for the control region in order to normalize the $Z \to \tau\tau$ background. In the fit procedure, the parameter of interest is $\tan\alpha$ with normalization factor signal strength μ_{sig} and normalization factor $b_{Z\tau\tau}$. During the fit procedure, the normalization factors are left for free-floating. This analysis uses simplified systematic uncertainties for both background and signal, which are based on [102].

For testing the SM signal hypothesis, expected results are performed using an Asimov dataset. As a result, the expected sensitivity is compared for fit results based on statistical uncertainties only and the effect of including the systematic uncertainties. Expected result of the fit with only statistical uncertainties, negative likelihood curve reaches a plateau of $-\Delta\ln(L) = 0.17$ for $\Delta\Phi_{jj}^{\text{ordered}}$, $-\Delta\ln(L) = 0.21$ for $\Delta\Phi_{jj}^{\text{sign}}$ and $-\Delta\ln(L) = 0.20$ for the Optimal Observable.

On the other hand, the expected result of the fit including the systematic uncertainties, curve reaches a plateau of $-\Delta\ln(L) = 0.07$ for $\Delta\Phi_{jj}^{\text{ordered}}$, $-\Delta\ln(L) = 0.12$ for $\Delta\Phi_{jj}^{\text{sign}}$ and $-\Delta\ln(L) = 0.10$ for the Optimal Observable.

From the expected result, $\Delta\Phi_{jj}^{\text{sign}}$ shows the highest sensitivity but yet $\pm 1\sigma$ confidence level for the signal hypothesis is not obtained. The reason for $\Delta\Phi_{jj}^{\text{sign}}$ appearing more sensitive could be caused by the binning choice and the use of the LO reweighting applied to the NLO events. These expected results show that an

additional study should be performed using a newly developed NLO reweighting technique with proper systematic uncertainties.

In future studies, the signal events for various $\tan \alpha$ coupling models can be improved by the usage of the NLO reweighting method. Also, using all decay channels of the Higgs boson can increase the sensitivity to anomalous coupling in ggF H+2 jets events.

List of Figures

List of Tables

Bibliography

[1] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*,
Phys. Rev. Lett. **10** (1963) 531,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[2] M. Kobayashi and T. Maskawa,
CP-Violation in the Renormalizable Theory of Weak Interaction,
Progress of Theoretical Physics **49** (1973) 652, ISSN: 0033-068X, eprint: http:
//oup.prod.sis.lan/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf,
URL: https://doi.org/10.1143/PTP.49.652.

[3] A. D. Sakharov,
Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe,
Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[4] A. Collaboration, *Observation of a new particle in the search for the Standard
Model Higgs boson with the ATLAS detector at the LHC*,
Phys. Lett. **B716** (2012) 1, arXiv: 1207.7214 [hep-ex].

[5] C. Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS
experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv: 1207.7235 [hep-ex].

[6] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics (1961).

[7] A. Salam, *Weak and Electromagnetic Interactions*,
Conf. Proc. **C680519** (1968) 367.

[8] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[9] U. Collaboration, *Experimental Observation of Isolated Large Transverse Energy
Electrons with Associated Missing Energy at $\sqrt{s} = 540\,GeV$*,
Phys. Lett. **122B** (1983) 103.

[10] U. Collaboration, *Observation of Single Isolated Electrons of High Transverse
Momentum in Events with Missing Transverse Energy at the CERN $p\overline{p}$ Collider*,
Phys. Lett. **122B** (1983) 476.

[11] U. Collaboration, *Evidence for $Z^0 \to e^+e^-$ at the CERN $p\overline{p}$ Collider*,
Phys. Lett. **129B** (1983) 130.

[12] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*,
Phys. Rev. Lett. **13** (1964) 508,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[13] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*,
Phys. Rev. Lett. **13** (1964) 321,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[14] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble,
Global Conservation Laws and Massless Particles, Phys. Rev. Lett. **13** (1964) 585,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[15]   L. Landau, *On the conservation laws for weak interactions*, **3** (1957) 127,
        URL: https://ui.adsabs.harvard.edu/abs/1957NucPh...3..127L.

[16]   J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay,
        Evidence for the 2π Decay of the K20 Meson, Phys. Rev. Lett. **13** (1964) 138.

[17]   T. D. Lee and C. N. Yang, *Question of Parity Conservation in Weak Interactions*,
        Phys. Rev. **104** (1956) 254,
        URL: https://link.aps.org/doi/10.1103/PhysRev.104.254.

[18]   C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson,
        Experimental Test of Parity Conservation in Beta Decay,
        Phys. Rev. **105** (1957) 1413,
        URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[19]   A. Collaboration, *Determination of spin and parity of the Higgs boson in the
        WW* → evμν decay channel with the ATLAS detector*,
        Eur. Phys. J. **C75** (2015) 231, arXiv: 1503.03643 [hep-ex].

[20]   A. Collaboration, *Study of the spin and parity of the Higgs boson in diboson
        decays with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 476,
        arXiv: 1506.05669 [hep-ex].

[21]   A. Collaboration, *Test of CP Invariance in vector-boson fusion production of the
        Higgs boson using the Optimal Observable method in the ditau decay channel with
        the ATLAS detector*, Eur. Phys. J. **76** (2016) 658, arXiv: 1602.04516 [hep-ex].

[22]   A. Collaboration, *Test of CP invariance in vector-boson fusion production of the
        Higgs boson in the H → ττ channel in proton–proton collisions at $\sqrt{s}$ = 13 TeV
        with the ATLAS detector*, (2019),
        URL: https://cds.cern.ch/record/2693960.

[23]   A. Lösle,
        *Study of CP-Properties of the Higgs-Boson produced in Gluon fusion with two jets
        in Proton-Proton Collisions at $\sqrt{s} = 8\,TeV$ with the ATLAS Experiment*, (2017),
        URL: https://terascale.physik.uni-
        freiburg.de/Publikationen/abschlussarbeiten/diplom-
        masterarbeiten/ALmaster/view.

[24]   D. Dominguez, *Particles of the Standard Model of particle physics*, CERN (),
        URL: https://home.cern/science/physics/standard-model.

[25]   B. Tanabashi et al., *Particle Data Group*,
        Phys. Rev. D **98** (2018 and 2019 update),
        URL: http://pdg.lbl.gov/2019/tables/contents_tables.html.

[26]   M. Thomson, *Modern particle physics*, Cambridge University Press, 2013,
        ISBN: 9781107034266, URL: http://www-
        spires.fnal.gov/spires/find/books/www?cl=QC793.2.T46::2013.

[27]   R. P. Feynman,
        Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction,
        Phys. Rev. **80** (1950) 440,
        URL: https://link.aps.org/doi/10.1103/PhysRev.80.440.

[28]   S. Tomonaga, *On a Relativistically Invariant Formulation of the Quantum Theory
        of Wave Fields**, Progress of Theoretical Physics **1** (1946) 27, ISSN: 0033-068X,
        eprint: http://oup.prod.sis.lan/ptp/article-pdf/1/2/27/24027031/1-
        2-27.pdf, URL: https://doi.org/10.1143/PTP.1.27.

[29] J. Schwinger, *Quantum Electrodynamics. I. A Covariant Formulation*,
Phys. Rev. **74** (1948) 1439,
URL: https://link.aps.org/doi/10.1103/PhysRev.74.1439.

[30] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*,
Phys. Rev. Lett. **30** (1973) 1343,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1343.

[31] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*,
Phys. Rev. Lett. **30** (26 1973) 1346,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1346.

[32] T. W. B. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*,
Phys. Rev. **155** (1967) 1554,
URL: https://link.aps.org/doi/10.1103/PhysRev.155.1554.

[33] H. E. Logan,
TASI 2013 lectures on Higgs physics within and beyond the Standard Model,
(2014), arXiv: 1406.1786 [hep-ph].

[34] R. D. Ball et al., *Parton distributions from high-precision collider data*,
Eur. Phys. J. **C77** (2017) 663, arXiv: 1706.00428 [hep-ph].

[35] A. Collaboration, *Standard Model Summary Plots Summer*, (2019),
URL: http://cds.cern.ch/record/2682186.

[36] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph].

[37] P. Artoisenet et al., *A framework for Higgs characterisation*, JHEP **11** (2013) 043,
arXiv: 1306.6464 [hep-ph].

[38] G. Klamke and D. Zeppenfeld,
Higgs plus two jet production via gluon fusion as a signal at the CERN LHC,
JHEP **04** (2007) 052, arXiv: hep-ph/0703202 [HEP-PH].

[39] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al.,
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].

[40] D. Atwood and A. Soni, *Analysis for magnetic moment and electric dipole moment form factors of the top quark via $e^+e- \rightarrow t\bar{t}$*,
Phys. Rev. D **45** (1992) 2405,
URL: https://link.aps.org/doi/10.1103/PhysRevD.45.2405.

[41] M. Davier, L. Duflot, F. L. Diberder, and A. Rougé,
The optimal method for the measurement of tau polarization,
Physics Letters B **306** (1993) 411, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/037026939390101M.

[42] M. Diehl and O. Nachtmann, *Optimal observables for the measurement of three gauge boson couplings in $e^+e^- \rightarrow W^+W^-$*,
Zeitschrift für Physik C Particles and Fields **62** (1994) 397, ISSN: 1431-5858,
URL: https://doi.org/10.1007/BF01555899.

[43] H.-L. Lai et al., *New parton distributions for collider physics*,
Phys. Rev. **D82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[44] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*,
Eur. Phys. J. **C75** (2015) 132, arXiv: 1412.7420 [hep-ph].

[45] K. Hagiwara, R. D. Peccei, D. Zeppenfeld, and K. Hikasa,
Probing the weak boson sector in $e^+e^- \to W^+W^-$, Nucl. Phys. **B282** (1987) 253.

[46] M. Bilenkii, J. L. Kneur, F. M. Renard, and D. Schildknecht, *Trilinear couplings among the electroweak vector bosons and their determination at LEP2*,
Nucl. Phys. **B408** (1993) 22, URL: http:
//www.sciencedirect.com/science/article/pii/055032139390445U.

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

[48] 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].

[49] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*,
JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph].

[50] G. Bagnaschi E.and Degrassi, P. Slavich, and A. Vicini, *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*,
Journal of High Energy Physics **2012** (2012) 88.

[51] K. Hamilton, P. Nason, E. Re, and G. Zanderighi,
NNLOPS simulation of Higgs boson production, JHEP **10** (2013) 222,
arXiv: 1309.0017 [hep-ph].

[52] K. Hamilton, P. Nason, and G. Zanderighi,
Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator,
JHEP **05** (2015) 140, arXiv: 1501.04637 [hep-ph].

[53] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*,
Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph].

[54] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*,
JHEP **01** (2001) 010, arXiv: hep-ph/0011363.

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

[56] S. Actis, G. Passarino, C. Sturm, and S. Uccirati,
NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders,
Phys. Lett. **B670** (2008) 12, arXiv: 0809.1301 [hep-ph].

[57] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger,
Higgs Boson Gluon-Fusion Production in QCD at Three Loops,
Phys. Rev. Lett. **114** (2015) 212001, arXiv: 1503.06056 [hep-ph].

[58] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, JHEP **05** (2016) 058,
arXiv: 1602.00695 [hep-ph].

[59] C. Anastasiou et al., *Mixed QCD-electroweak corrections to Higgs production via gluon fusion in the small mass approximation*,
Journal of High Energy Physics (2019) 162.

[60] M. Ciccolini, A. Denner, and S. Dittmaier,
Strong and Electroweak Corrections to the Production of a Higgs Boson + 2 Jets via Weak Interactions at the Large Hadron Collider,
Phys. Rev. Lett. **99** (2007) 161803,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.99.161803.

[61] M. Ciccolini, A. Denner, and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC,*
Phys. Rev. **D77** (2008) 013002, arXiv: 0710.4749 [hep-ph].

[62] P. Bolzoni, F. Maltoni, S. Moch, and M. Zaro,
Higgs production via vector-boson fusion at NNLO in QCD,
Phys. Rev. Lett. **105** (2010) 011801, arXiv: 1003.4451 [hep-ph].

[63] O. Brein, A. Djouadi, and R. Harlander,
NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders,
Phys. Lett. **B579** (2004) 149, arXiv: hep-ph/0307206 [hep-ph].

[64] A. Denner, S. Dittmaier, S. Kallweit, and A. Muck, *Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK,*
JHEP **03** (2012) 075, arXiv: 1112.5142 [hep-ph].

[65] L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak, and T. J. E. Zirke,
Gluon-induced Higgs-strahlung at next-to-leading order QCD,
JHEP **02** (2013) 078, arXiv: 1211.5015 [hep-ph].

[66] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, et al.,
Event generation with SHERPA 1.1, JHEP **02** (2009) 007,
arXiv: 0811.4622 [hep-ph].

[67] K. Melnikov and F. Petriello,
*Electroweak gauge boson production at hadron colliders through O(alpha(s)**2),*
Phys. Rev. D **74** (2006) 114017, arXiv: hep-ph/0609070.

[68] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush,
FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order,
Comput. Phys. Commun. **182** (2011) 2388, arXiv: 1011.3540 [hep-ph].

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

[70] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 physics and manual,*
JHEP **05** (2006) 026, arXiv: hep-ph/0603175.

[71] M. Beneke, P. Falgari, S. Klein, and C. Schwinn,
Hadronic top-quark pair production with NNLL threshold resummation,
Nucl. Phys. **B855** (2012) 695, arXiv: 1109.1536 [hep-ph].

[72] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels,* JHEP **12** (2012) 054,
arXiv: 1207.0236 [hep-ph].

[73] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason,
Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation, Phys. Lett. **B710** (2012) 612,
arXiv: 1111.5869 [hep-ph].

[74] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction,* JHEP **01** (2013) 080,
arXiv: 1210.6832 [hep-ph].

[75] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha \frac{4}{S})$*, Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph].

[76] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph].

[77] S. Agostinelli et al., GEANT4 – *a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[78] E. Halkiadakis, *Proceedings for TASI 2009 Summer School on "Physics of the Large and the Small": Introduction to the LHC experiments*, (2007), arXiv: 1004.5564 [hep-ex].

[79] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[80] M. Hostettler and G. Papotti, *Luminosity Lifetime at the LHC in 2012 Proton Physics*, (2013) 3 p, URL: http://cds.cern.ch/record/1567672.

[81] A. Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[82] A. Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. **C77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[83] A. Collaboration, ATLAS Twiki (2018), URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults.

[84] A. Collaboration, *Luminosity Determination in pp Collisions at $\sqrt{s} = 13\,TeV$ using the ATLAS Detector at the LHC*, (2019), URL: https://cdsweb.cern.ch/record/2677054.

[85] A. Collaboration, ATLAS Twiki (2015-2018), URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[86] A. Collaboration, ATLAS Twiki (2015-2018), URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010.

[87] A. Collaboration, *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].

[88] A. Collaboration, *Early Inner Detector Tracking Performance in the 2015 Data at $\sqrt{s} = 13\,TeV$*, (2015), URL: https://cds.cern.ch/record/2110140.

[89] A. Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13\,TeV$*, (2015), URL: https://cds.cern.ch/record/2037717.

[90] A. Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data*, (2019), arXiv: 1908.00005 [hep-ex].

[91] A. Collaboration, *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].

[92] A. Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} =13\,TeV$*, Eur. Phys. J. **C76** (2016) 292, arXiv: 1603.05598 [hep-ex].

[93] A. Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. **C76** (2016) 581, arXiv: 1510.03823 [hep-ex].

[94] A. Collaboration, *Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector*, Eur. Phys. J. **C77** (2017) 580, arXiv: 1705.02211 [hep-ex].

[95] A. Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** (2016) P04008, arXiv: 1512.01094 [hep-ex].

[96] A. Collaboration,
Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run, (2016),
URL: http://cds.cern.ch/record/2160731.

[97] A. Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, (2017), URL: http://cds.cern.ch/record/2261772.

[98] A. Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. **C78** (2018) 903, arXiv: 1802.08168 [hep-ex].

[99] A. Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. **C77** (2017) 317, URL: https://doi.org/10.1140/epjc/s10052-017-4852-3.

[100] R. K. Ellis, I. Hinchliffe, M. Soldate, and J. J. Van Der Bij, *Higgs decay to $\tau^+\tau^-$ A possible signature of intermediate mass Higgs bosons at high energy hadron colliders*, Nucl. Phys. **B297** (1988) 221, ISSN: 0550-3213.

[101] A. Elagin, P. Murat, A. Pranko, and A. Safanov, *A New Mass Reconstruction Technique for Resonances Decaying to $\tau\tau$*, Nucl. Instrum. and Meth. **A654** (2011) 481, arXiv: 1012.4686 [hep-ex].

[102] A. Collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. **D99** (2019) 072001, arXiv: 1811.08856 [hep-ex].

[103] F. Becherer, *Estimation of background processes with a jet misidentified as a hadronically decaying τ-lepton and measurements of Simplified Template Cross Sections in the $H \to \tau_{lep}\tau_{had}$ decay channel with the ATLAS experiment at $\sqrt{s} = 13$ TeV*, (2017) 70, URL: https://terascale.physik.uni-freiburg.de/Publikationen/abschlussarbeiten/diplom-masterarbeiten/masterFabianBecherer/view.

[104] A. Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, (2007), arXiv: physics/0703039 [physics.data-an].

[105] J. H. Friedman, *Greedy function approximation: A gradient boosting machine*, Ann. Statist. **29** (2001) 1189, URL: https://doi.org/10.1214/aos/1013203451.

[106] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, (2012), URL: https://cds.cern.ch/record/1456844.

[107] G. Cowan, *Statistical Data Analysis*, Clarendon Press (1998).

[108] S. Brandt, *Data Analysis: Statistical and Computational Methods for Scientists and Engineers*, Springer **4th ed.** (2014).

Declaration

Herewith I affirm that the submitted master thesis was written autonomously by myself and that I did not use any other sources and auxiliaries than declared in this work. Under the acknowledged rules of scientific work (lege artis) literal or analogous content borrowed from the work of others was appropriately identified. Furthermore, I insure that the submitted master thesis is not and was never part of any other examination procedures, either complete or in substantial parts.

Signed:

Date:

Acknowledgements

I would like to thank my Prof. Dr. Markus Schumacher, for giving me the opportunity to write a master thesis in his research group of particle physics. I am very thankful for my supervisor Dr. Kathrin Becker for all the guidance and responses to all my questions related to physics and programming. Also, I am very thankful to Dr. David Hohn for also supervising my thesis. Furthermore, I am appreciated for all group members for helping me every possible situation, especially, I am deeply appreciated to Alena Lösle for all the advice.

Also, I would like to thank my friend Dr. Christian Scheffzük, who supports me all the time and helps me to overcome the problematic situations. I am thankful to my parents, who believed in me all the time.

I am deeply grateful to all of you! Thank you!