Search for Higgs boson pair-production in the $bb\tau\tau$ final state using proton-proton collisions at $\sqrt{s}$ = 13 TeV data with the ATLAS detector

Zhiyuan Li

March 14, 2022

Contents

1 Introduction **1**

2 Theory and motivation **2**

2.1 The Standard Model of Particle Physics 2

2.1.1 Particle Content 3

2.1.2 Symmetries in the Standard Model 5

2.1.3 Quantum Chromodynamics 7

2.1.4 Electroweak theory 9

2.1.5 The Higgs mechanism 11

2.1.6 Yukawa coupling 15

2.1.7 Higgs boson production at the LHC 16

2.1.8 Higgs boson pair production at the LHC 19

2.1.8.1 The $b\bar{b}\tau^+\tau^-$ decay channel 22

2.2 Beyond the Standard Model 23

2.2.1 Big questions in the Standard Model 23

2.2.2 Two-Higgs-doublet Model 25

2.2.3 Effective field theory interpretation 27

2.3 Machine learning theory 29

2.4 Statistical Interpretation 32

2.4.1 Test statistics 33

2.4.2 Look elsewhere effect 36

Chapter 1

Introduction

Chapter 2

Theory and motivation

Particle physics is in the heart of our understanding of the laws of nature. The subject is concerned with the fundamental constituents of the Universe, the *elementary particles*, and the interactions between them, the *forces*. The Standard Model (SM) embodies the current understanding of particle physics, providing a unified picture where the forces between the particles are themselves described by the exchange of particles. It provides a successful description of all current experimental data and represents on of the triumphs of mordern physics.

The last missing piece of the SM, the Higgs boson has been observed by the ATLAS and CMS experimental at the Large Hadron Collider [1, 2]. The SM is therefore hailed by many as the most successful theory ever conceived. Despite its success, the SM is yet not the ultimate theory, as many unanswered questions remain. For example, why the SM has so many free parameters (26) that have to be input by hand; what is the particle content of the dark matter; what is the origin of the matter-antimatter assymmetry in the universe?

This chapter is structured as follows. Section 2.1 introduces the basic concepts of the SM, including the gauge theory and fundamental forces and the Higgs mechanism. Section 2.2 gives a hint of beyond the SM (BSM) theories that address some of the important and unanswered questions. Section 2.3 and section 2.4 outline the machine learning algorithms and the statistical methods used in this thesis, respectively.

2.1 The Standard Model of Particle Physics

The core of the SM was outlined by Steven Weinberg just over a half-century ago, when he published the short but revolutionary paper titled “A Model of Leptons” in the journal Physical Review Letters [3]. The SM was developed in stages throughout the latter half of the 20th century, through the work of many scientists around the world, with the current formulation being finalized in the mid–1970s upon experimental confirmation of

the existence of quarks [4, 5].

In the SM, most of the everyday phenomena we see in the physical world is just the low energy manifestation of the twelve elementary particles and three interactions: electromagnetic, weak and strong forces. For example, atoms, which were believed to be the most basic building blocks of the world, are in fact comprised of a negatively charged electrons and positively charged nucleus, bound by electromagnetic attraction. Such phenomenon is a low energy manifestation of the fundamental theory of electromagnetism, namely Quantum Electrodynamics (QED). While in the nucleus, the protons and neutrons are bound together by the strong nuclear force (where the protons and neutrons are comprised of quarks bound by the strong force), which is again the low energy manifestation of the fundamental theory of strong interactions, namely Quantum Chromodynamics (QCD). The fundamental interactions of the particles are completed by the weak force, which controls for example beta decay. Although gravity is oftenly ignored in the context of particle physics due to its relatively much smaller strength compared to the other three fundamental interactions (around 10^{34} times weaker than the electromagnetic force), gravity must be integrated into the theory. The SM is widely considered to be incompatible with the most successful theory of gravity to date, general relativity. To date no clear approach is available for combining the two behemoths, but huge effort has been put into physics beyond SM, such as the Supersymmetry (SUSY), Loop quantum gravity or String theory, which may shed some light on the ultimate theory of everything.

2.1.1 Particle Content

In the SM, the electron, the electron neutrino, the up-quark and the down-quark are known collectively as the first generation. As far as we know, they are elementary particles, instead of being composite, and represent the basic building blocks of the low-engery universe. For each of the first generation particles, there are exactly two copies which differ only in their masses. These additional eight particles are known as the second and the third generations. For example, the second generation muon is essentially a heavier version of the electron with mass $m_\mu \approx 200 m_e$, and the third generation τ-lepton is an even heavier copy with $m_\tau \approx 3500 m_e$. The three generations of particles, are collectively called as *fermions*. Fermions have intrinsic spin $s = \frac{1}{2}$ and obey the Pauli exclusion principle, which states that no two fermions can occupy the same quantum state.

The dynamics of each of the twelve fundamental fermions are governed by the Dirac equation of relativistic quantum mechanics [6],

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \tag{2.1}$$

where the γ^μ are the gamma matrices, ψ is the Dirac spinor and m is the mass of the particle. An important consequence is that for each of the twelve fermions there exists an antiparticle state with exactly the same mass, but opposite charge and intrinsic spin. The antiparticles are denoted either by their charge or by a bar over the corresponding particle symbol, for example, the anti-electron (positron) is denoted by e^+, and the anti-up-quark is written as $\bar{u}$.

In contrast to fermions, *bosons* are defined as particles that have integer intrinsic spin, and do not obey the Pauli exclusion principle. In particular, bosons with spin 1 are called *gauge bosons*. In modern particle physics, the three fundamental forces are described by a Quantum Field Theory (QFT), with each gauge boson can be seen as the excitations of the quantum field of each forces. For example, the familiar photon is the gauge boson of the Quantum Electrodynamics (QED), and for the strong interaction, the force-carrying partilce is called a *gluon*. While the photon and gluon are massless, the weak charged-current interaction, which is responsible for nuclear β-decay, is mediated by the W^- or the W^+ bosons with masses of 80.4 GeV. In addition, the neutral-current interaction is mediated by the chargeless Z boson, with a mass of 91.2 GeV.

Due to the large mass of the mediator, the weak force is, as its name suggests, much weaker than the electromagnetic force and the strong force: about 10^5 times weaker than the electromagnetic force; while the strong force is intrinsicly much stronger than the other two: about 1000 times stronger than the electromagnetic force(note that the strength of interaction depends greatly on the distance and energy scale being considered). Another consequence is, the weak force has an extremely short effective range of around 10^{-18} m, while the massless photon enables the electromagnetic to apply at infinite distance. The gluon is also massless, but has an effective range of around 10^{-15} m, due to *colour confinement* (which will be discussed in more detail in section 2.1.3). Lastly, all fermions can 'feel' the weak force, while the electromagnetic force only applies to particles with electric charge and the strong force only applies to quraks and gluons themselves.

The final element of the SM is the Higgs bosons. Unlike all other SM particles, which have either spin 1 or $\frac{1}{2}$, the Higgs boson is the only known fundamental scalar particle, having a spin of 0. The *Higgs mechanism* plays an importance role in the SM by providing mass for all known particles: without it all particles would be massless, making the universe a very different place! More specifically, unlike other fields associated with the fundamental fermions and bosons, the Higgs field has a non-zero vacuum expectation value; the interaction of the particles with the Higgs field is what provides them with mass. This mechanism is discussed in more detail in section 2.1.5.

The 12 elementary fermions and the 5 elementary bosons (6, if counting the hypothetical graviton) are illustrated in Figure 2.1. All particles in the SM are assumed to be point-like.

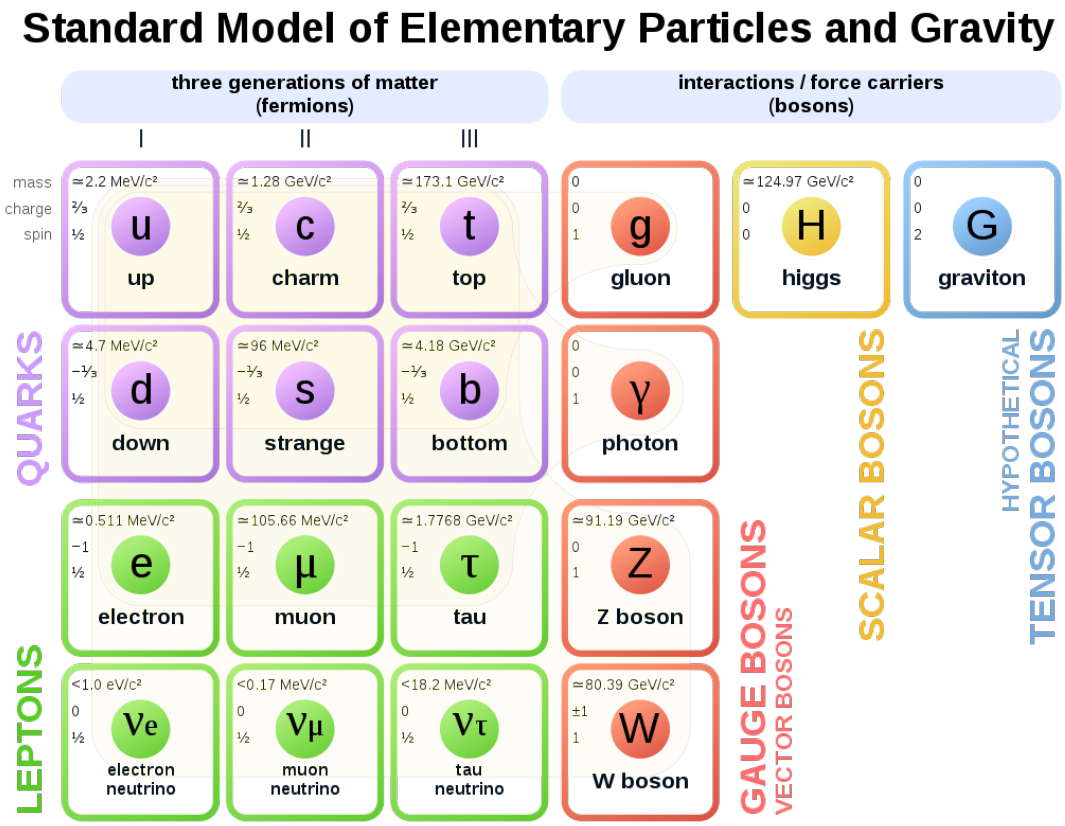


Figure 2.1: SM of elementary particles: the 12 fundamental fermions and 5 fundamental bosons (along with the hypothetical Graviton). The mass, charge and spin of each particle are given inside the particle boxes [7].

2.1.2 Symmetries in the Standard Model

Symmetry plays a crucial role in modern physics, particularly in the SM. SM is a relativistic quantum gauge theory containing the internal symmetries of the unitary product group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The $SU(3)_C$ is the symmetry group of the strong interaction which is non-abelian and the letter C refers to the color charge which is the corresponding conserved quantity. The $SU(2)_L \times U(1)_Y$ is the symmetry group of the electroweak interaction that unifies the weak and electromagnetic interactions, the letter L stands for left and indicates that the symmetry only involves left-handed particles, while the letter Y stands for the weak hypercharge which is the conserved quantity corresponding to the $U(1)_Y$ group. The weak hypercharge is related to the electric charge (Q) which is conserved due to the global gauge invariance of the electromagnetic field and the weak isospin (T) given as $Q = T_3 + Y/2$, with T_3 being the third component of the weak isospin and is

conserved due to the $SU(2)_L$ symmetry.

The SM is described by the Lagrangian formalism, and the Lagrangian density (or just Lagrangian) is constructed from four components:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{Electroweak}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}} \tag{2.2}$$

where $\mathcal{L}_{\text{QCD}}$ describes the dynamics of the strong force, $\mathcal{L}_{\text{Electroweak}}$ describes the dynamics of the electroweak force and $\mathcal{L}_{\text{Higgs}}$ and $\mathcal{L}_{\text{Yukawa}}$ are the terms which introduce mass to fermions and gauge bosons respectively. The explicit definition of each term will be introduced in the following sections.

To understand the relation between symmetries and conserved charges, one can consider this simple example: suppose the dynamics are determined by an action S written in terms of a Lagrangian density $\mathcal{L}(x)$ that contains the free Lagrangian of the fields ($\psi(x)$), which accounts for their free propagation, and additional terms that respect the above symmetries and account for their interactions:

$$S = \int d^4x \mathcal{L}(x). \tag{2.3}$$

The Euler-Lagrangian equations can be derived assuming that the action is stationary, i.e. $\delta S = 0$:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \right) - \frac{\partial \mathcal{L}}{\partial \psi} = 0. \tag{2.4}$$

In gauge theory, the Lagrangian is invariant under gauge transformation of Infinitesimal change $\delta\psi$:

$$\psi(x) \rightarrow \psi'(x) = \psi(x) + \delta\psi. \tag{2.5}$$

Noether's theorem states (informally) that if a system has a continuous symmetry property, then there are corresponding quantities whose values are conserved. In this simple example, Noether's theorem follows as:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \delta\psi \right) = \partial_\mu J^\mu = 0, \tag{2.6}$$

where J^μ is the conserved current and

$$Q = \int dx J^0 = \text{constant}, \tag{2.7}$$

is the conserved charge associated to the symmetry.

2.1.3 Quantum Chromodynamics

It is useful to introduce the concept of local gauge invariance, which is a familiar idea from electromagnetism. The physical electric file **E** and the magnetic field **B** is unchanged under this transformation:

$$\psi \rightarrow \psi' = \psi - \partial\chi/\partial t \text{ and } \mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla\chi, \tag{2.8}$$

and in covariance form, this can be written as

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu\chi, \tag{2.9}$$

where $A_\mu = (\psi, -\mathbf{A})$ and $\partial_\mu = (\partial_0, \nabla)$. For the U(1) transformation on the wave function ψ, $\psi(x) \rightarrow \psi'(x) = \hat{U}(x)\psi(x) = e^{iq\chi(x)}\psi(x)$ where $\hat{U}(x)$ is the *generator* of the U(1) group, the Dirac equation for a free particle,

$$i\gamma^\mu\partial_\mu\psi = m\psi, \tag{2.10}$$

becomes:

$$i\gamma^\mu\partial_\mu(e^{iq\chi(x)}\psi) = i\gamma^\mu(\partial_\mu + iq\partial_\mu\chi)\psi = m\psi. \tag{2.11}$$

This differs from the free particle Dirac equation by the term $-q\gamma^\mu\partial_\mu\chi\psi$. It can be seen that the free particle Dirac equation cannot be local gauge invariant due to this additional term. The solution here is to introduce a field, A_μ which transforms as

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu\chi, \tag{2.12}$$

such that the original form of the Dirac equation becomes

$$i\gamma^\mu(\partial_\mu\psi + iqA_\mu) = m\psi. \tag{2.13}$$

This idea can be applied to the QCD, which obeys the SU(3) group. Suppose a SU(3) transformation is applied on the wave function, i.e.

$$\psi(x) \rightarrow \psi(x)' = \exp\left[ig_S\alpha(x)\cdot\hat{\mathbf{T}}\right]\psi(x), \tag{2.14}$$

where g_S is some coupling constant, the $\hat{\mathbf{T}} = \{T^a\}$ are the eight generators of the SU(3), which are related to the *Gell-Mann matrices* by $T^a = \frac{1}{2}\lambda^a$ [8], and $\alpha(x)$ are eight functions of the space-time coordinate x, corresponding to each of the eight SU(3) generators. Due to the fact that the SU(3) group is represented by 3 by 3 matrices, the additional degrees of freedom is accounted by a vector of three components, namely red, green and blue.

Hence, the idea of color charge comes naturally from requiring the local gauge invariance. Finally, the concept of the gluon also comes out when requiring the local gauge invariance, which is the quanta of the eight introded fields. The Dirac equation with interactions with the eight type of gluons becomes:

$$i\gamma_\mu \left[\partial_\mu + ig_S G^a_\mu T^a\right] \psi - m\psi = 0, \tag{2.15}$$

which is invariant under local SU(3) transformation if the new fields transform as:

$$G^k_\mu \to G'^k_\mu = G^k_\mu - \partial_\mu \alpha_k - g_S f_{ijk} \alpha_i G^j_\mu, \tag{2.16}$$

where the f_{ijk} is the structure constant to account for the fact that the SU(3) generators do not commute (and therefore, QCD is a *non-Abelian* theory).

An important result of the extra $g_S f_{ijk} \alpha_i G^j_\mu$ term is the gluon can interact with itself, which is the origin of colour confinement. So far, there is no free quark observed in the nature, and the reason might well possibly be colour confinement. The qualitative explaination of this hypothesis is as follows: consider two quarks are in a bound state, in order to create a free quark one would need to pull the two quarks far away from each other until they become 'free'. However, as gluons can interact with themselves (as attraction), and the interaction between the two quarks can be thought of as exchanging gluons, the exchanged gluons actually attract themselves. The effect is that the gluon field is 'squeezed' into the shape of a tube, which has an energy density approximately constant over the distance. Therefore, the energy stored in the field is proportional to the separation of the quarks, giving a term in the potential of the form: $V(\mathbf{r}) \sim \kappa r$, where experimentally $\kappa \sim 1$ GeV/fm. This corresponds to the a force of the oder of 10^5 N, and consequently, the gluon field can store enough energy to create new pairs of quarks when the two quarks are far apart. The newly created quarks pair can become new bound states with the quarks being pulled apart. This process goesn on if the initial quarks are pulled further apart, as shown in Figure 2.2:

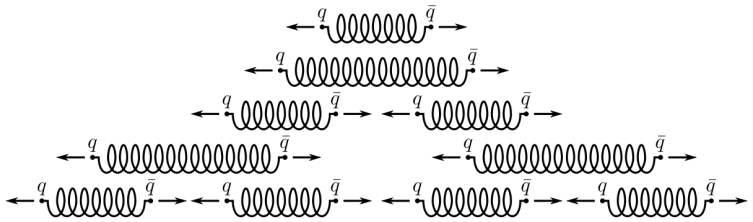


Figure 2.2: An illustration of a pair of quarks being pulled away: new pair of quarks are created and become new bound states with the two quarks being pulled.

Another consequence of the colour confinement is that the coloured gluons are confined

to the colourless objects, which is the reason why gluons are massless but the strong force range is not macroscopic like photons and the electromagnetic force. In addition, at short distances (or equivalently, high energy scales), the coupling strength of the strong force $\alpha_S = \frac{g_s^2}{4\pi}$ is small, and in bound states quarks behave like free particles. This is reffered to as *asymptotic freedom*. For example, for momentum transfer at the scale of the mass of the Z boson, α_S has a value of around 0.12. In modern particle detectors, the α_S value is sufficiently small for the pertubation theory to be used. Finally, the Lagrangian of QCD describing the quarks interactions via gluons and the gluon self-interaction in a compact form is:

$$\mathcal{L}_{QCD} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a, \tag{2.17}$$

with the covariant derivatives D_μ given by:

$$D_\mu = \partial_\mu + ig_s T^a G^a_\mu. \tag{2.18}$$

2.1.4 Electroweak theory

Following a similar approach to that described in the previous section, consider the non-abelian SU(2) transformation, i.e.

$$\psi(x) \rightarrow \psi(x)' = \exp\left[ig_W \alpha(x) \cdot \mathbf{T}\right]\psi(x), \tag{2.19}$$

with g_W being the weak coupling constant, $\mathbf{T}$ being the three generators of the SU(2) group, which are related to the Pauli spin matrices by $\mathbf{T} = \frac{1}{2}\sigma$, and $\alpha(x)$ are three functions which specify the local phase at each point in space-time. To satisfy local gauge invariance, three gauge fields must be introduced: W^k_μ with $k = 1, 2, 3$, corresponding to three gauge bosons $W^{(1)}, W^{(2)}, W^{(3)}$. Fermions are comprised of components with negative and positive chirality, referred to as left- and right-handed particles, respectively.

Since the weak force only interacts with left-handed (LH) chiral particles or right-handed (RH) chiral antiparticles, and the the generators are 2×2 spin matrices, the LH particles and RH antiparticles states can be expressed as a weak isospin doublet, i.e.

$$\psi_L^{\ell=e,\mu,\tau} = \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L, \psi_L^{q=1,2,3} = \begin{pmatrix} u_q \\ d'_q \end{pmatrix}_L, \tag{2.20}$$

where the d'_q are the flavour states representing the three generations of the up-type quarks and the d'_q are the down-type quarks. Notice the flavour eigenstates d'_q differ from the mass eigenstates d_q, where the former is a mixture of the latter using the Cabibbo-Kobayashi-Maskawa (CKM) matrix [9]. The RH chiral particles and LH antiparticles are represented

by a weak isospin singlets, with

$$\psi_R^{\ell=e,\mu,\tau} = \ell_R \text{ and } \psi_R^{q=u,c,t,d,s,b} = q_R. \tag{2.21}$$

Analogous to the QCD formulation, an extra interaction term arises, which is

$$ig_W T_k \gamma^\mu W_\mu^k \psi_L = ig_W \frac{1}{2} \sigma_k \gamma^\mu W_\mu^k \psi_L, \tag{2.22}$$

where ψ_L is the LH weak isospin doublet. The physical W bosons are in fact the linear combinations of the two gauge fiels $W^{(1)}$ and $W^{(2)}$:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^{(1)} \mp iW_\mu^{(2)}). \tag{2.23}$$

It is natural to think the physical Z boson corresponds to the third W_μ^k, as it implies a neutral current which can be related to the chargeless Z. However, experimentalally the Z boson does not only couple to LH particles, but also to RH particles, although not equally. To solve this conflict, electromagnetism is introduced into the story, which was so far not considered.

In the electroweak theory by Glashow, Salam and Weinberg [10–12], the U(1) gauge symmetry of electromagnetism is regarded by a new $\mathrm{U}(1)_Y$ local gauge symmetry, and it transforms as:

$$\psi(x) \to \psi(x)' = \hat{U}\psi(x) = \exp\left[ig' \frac{Y}{2} \zeta(x)\right] \psi(x), \tag{2.24}$$

with g' being a new coupling constant (its relation will become clear in the following), $\zeta(x)$ being some fuction in x. Requiring local gauge invariance necessitates the interaction term:

$$g' \frac{Y}{2} \gamma^\mu B_\mu \psi \tag{2.25}$$

(notice the same form as the simple example in section 2.1.3). Using the interaction term, one can now write the photon and Z boson in terms of linear combinations of the new B_μ field and the third W_μ^k:

$$A_\mu = +B_\mu \cos\theta_W + W_\mu^{(3)} \sin\theta_W, \text{ and} \tag{2.26}$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^{(3)} \cos\theta_W, \tag{2.27}$$

where θ_W is called the *weak mixing angle*.

One can deduce the $Y = 2(Q - I_W^3)$ relation using the following logic: the electroweak theory is invariant under $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ transformation, and the corresponding hypercharge Y is conserved. One can assume that the relation between the charge, weak isospin

and the hypercharge is linear, i.e.

$$Y = \alpha Q + \beta I_W. \tag{2.28}$$

Since the Y must be the same for a LH electron and a LH neutrino, i.e. $Y_{e_L} = Y_{\nu_L}$ (otherwise the U(1) transformation will break the symmetry of the isospin doublet), using their charge and weak isospin values respectively one can conclude that:

$$Y = 2(Q - I_W^3). \tag{2.29}$$

The full formulation might not be as important, but one can deduce the electromagnetic current j^μ_{em} has terms equal to:

$$\begin{aligned} \bar{e}_L\gamma^\mu e_L : \qquad & Q_e e = \frac{1}{2}g' Y_{eL}\cos\theta_W - \frac{1}{2}g_W \sin\theta_W, \\ \bar{\nu}_L\gamma^\mu \nu_L : \qquad & 0 = \frac{1}{2}g' Y_{\nu L}\cos\theta_W - \frac{1}{2}g_W \sin\theta_W. \end{aligned}$$

Since $Y_{e_L} = Y_{\nu_L} = -1$ and $Y = 2(Q - I_W^3)$, the coupling constant g' follows the relation:

$$e = g'\cos\theta_W = g_W \sin\theta_W. \tag{2.30}$$

The expected ratio of the weak to electromagnetic coupling constants is

$$\frac{\alpha}{\alpha_W} = \frac{e^2}{g_W^2} = sin^2\theta_W \sim 0.23. \tag{2.31}$$

Finally, the Lagrangian of the electroweak theory is:

$$\mathcal{L}_{\text{electroweak}} = \bar{\psi}_L\gamma^\mu D^L_\mu \psi_L + \bar{\psi}_R\gamma^\mu D^R_\mu - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}\vec{W}_{\mu\nu}\vec{W}^{\mu\nu}, \tag{2.32}$$

with

$$D^L_\mu = i\partial_\mu - \frac{g}{2}\vec{\sigma}\cdot\vec{W}_\mu - \frac{g'}{2}YB_\mu, \text{ and } D^R_\mu = i\partial_\mu - \frac{g'}{2}YB_\mu, \tag{2.33}$$

where $\vec{\sigma}$ are the three Pauli matrices.

2.1.5 The Higgs mechanism

The local gauge invariance is preserved in $SU(2)_L$ group only if the bosons are massless. Consider if the photon were massive, the QED Lagrangian becomes:

$$\mathcal{L} \rightarrow \bar{\psi}(i\gamma^\mu\partial_\mu - m_e)\psi + e\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m_\gamma^2 A_\mu A^\mu \tag{2.34}$$

where the new term $\frac{1}{2}m_\gamma^2 A_\mu A^\mu$ arises assuming a massive photon. It is clear that this new term is not gauge invariant under the U(1) group transformation. This simple example can be applied to the $SU(2)_L$, and to solve the conflict that experimental observations show the weak bosons are massive while the local gauge invariance requires the weak bosons to be massless, the Higgs mechanism is proposed.

Now consider a complex scalar field (ϕ):

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2), \tag{2.35}$$

with a Lagrangian of the form:

$$\mathcal{L} = (\partial_\mu \phi)^*(\partial^\mu \phi) - V(\phi) \text{ with } V(\phi) = \mu^2(\phi^*\phi) + \lambda(\phi^*\phi)^2. \tag{2.36}$$

For the potential $V(\phi)$ to be physical, it should have a finite minimum, and therefore $\lambda > 0$. However, the coefficient μ^2 can be either positive or negative. When $\mu^2 < 0$, the potential has inifinite minima defined by

$$\phi_1^2 + \phi_2^2 = \frac{-\mu^2}{\lambda} = v^2, \tag{2.37}$$

which is a circle on the $\phi_1 - \phi_2$ plane, as shown in Figure 2.3. The physical vacuum

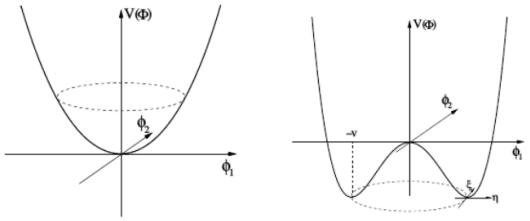


Figure 2.3: The potential $V(\phi)$ for a complex scalar field for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right). Image taken from [13].

state corresponds to a particular point on the circle, where the $SU(2) \times U(1)$ symmetry is *spontaneously broken*.

Without loss of generality, one can pick the vacuum state to be

$$(\phi_1, \phi_2) = (v, 0), \tag{2.38}$$

and the complex scalar field can be expanded about the vacuum state as

$$(\phi_1, \phi_2) = (v + \eta(x), \zeta(x)), \tag{2.39}$$

where $\eta(x)$ and $\zeta(x)$ are pertubations of real fields in the ϕ_1 and ϕ_2 direction. The Lagrangian can now be written in the form of:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) + \frac{1}{2}(\partial_\mu \zeta)(\partial^\mu \zeta) - V(\eta, \zeta), \tag{2.40}$$

with $V(\eta, \zeta)$ given by:

$$V(\eta, \zeta) = \mu^2\phi^2 + \lambda\phi^4, \text{ and } \phi^2 = \phi\phi^* = \frac{1}{2}[(\mu + \eta)^2 + \zeta^2], \tag{2.41}$$

which after expanding $V(\eta, \zeta)$ gives:

$$V(\eta, \zeta) = -\frac{1}{4}\lambda v^4 + \lambda v^2\eta^2 + \lambda v\eta^3 + \frac{1}{4}\lambda\eta^4 + \frac{1}{4}\lambda\zeta^4 + \lambda v\eta\zeta^2 + \frac{1}{2}\lambda\eta^2\zeta^2. \tag{2.42}$$

The second term $\lambda v^2\eta^2$ can be seen as the mass term of field η, i.e. $\frac{1}{2}m_\eta^2\eta^2 = \lambda v^2\eta^2$, while the rest can be seen as interaction terms. Notice that the field ζ along the ϕ_2 direction (the direction that the potential does not change) does not have a mass term, and therefore it is massless. The massless particle corresponding to this field is called a *Goldstone boson*.

The full formulation of the Higgs mechanism is rather long and not appropriate in the context of this thesis. The general idea is that, by requiring symmetry in a particular group, one can use the vacuum state function with pertubations in the gauge invariant Lagrangian and derive the kinematic terms of the massive field η and massless ζ, and the massive gauge field B (which was massless originally). In this process, the massless field B has acquired mass, and by choosing the gauge carefully (known as the *Unitary gauge*), the massless ζ field can be absorbed into the now massive gauge field B.

In the SM, the Higgs mechanism is embedded in the U(1) and $\mathrm{SU}(2)_L$ group, and to account for the three degrees of freedom of the $W^\pm$, Z bosons, three Goldstone bosons are required. Therefore, the simplest method would be to have two complex scalar fields, and since one of the electroweak bosons is neutral, one of the field needs to be neutral as well, which would be denoted by ϕ^0. The second must be charged to account for the $W^\pm$ and one can denote the charged scalar field as ϕ^+ such that $(\phi^+)^* = \phi^-$. The scalar field can now be written as:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{2.43}$$

For the Higgs potential with the form:

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2, \tag{2.44}$$

the vacuum state satisfies:

$$\phi^\dagger\phi = \frac{1}{2}\sum_{i=1,2,3,4}\phi_i^2 = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}. \tag{2.45}$$

Because the photon is required to remain chargeless, the minimum of the potential must correspond to a non-zero vacuum expectation value only of the neutral scalar field ϕ^0. Writting the doublet in unitary gauge, one gets:

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.46}$$

The resulting Lagrangian is known as the Salam-Weinberg model [10–12].

To preserve the $\mathrm{SU}(2)_L \times \mathrm{U}(1)$ symmetry, the derivatives needs to be replaced by appropriate covariant derivatives:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ig_W\mathbf{T}\cdot\mathbf{W}_\mu + ig'\frac{Y}{2}B_\mu. \tag{2.47}$$

By subsituting $\phi(x)$ in the kinematic term $(D_\mu\phi)^\dagger(D^\mu\phi)$, the Lagrangian of the Higgs field becomes:

$$\begin{aligned}\mathcal{L} \;=& \frac{1}{2}\partial_\mu h\partial^\mu h + \frac{g_W^2}{4}(v+h)^2 W_\mu^+ W^{-\mu} + \frac{1}{8}(g_W^2 + g'^2)(v+h)^2 Z_\mu Z^\mu \\ & - \lambda v^2 h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4.\end{aligned} \tag{2.48}$$

As a result, the mass of the W boson is determined by second term on the first row of the Lagrangian:

$$m_W = \frac{1}{2}g_W v, \tag{2.49}$$

and the mass of the Z boson is given by the third term on the first row:

$$m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2} = \frac{m_W}{\cos\theta_W}. \tag{2.50}$$

The mass of the Higgs boson is given by the first term on the second row:

$$m_H = \sqrt{2\lambda}v. \tag{2.51}$$

In addition to the mass terms, the Lagrangian includes the interaction terms of $VVhh$, VVh (V for $W^\pm$ or Z) and the Higgs self-interaction h^3, h^4 terms (trilinear and quadlinear). The Lagrangian does not depend in A_μ, and therefore the $\mathrm{U}(1)$ symmetry is unbroken, and the

photon remains massless. The vacuum expectation value of the Higgs field determined experimentally is given by $v \approx 246$ GeV [7].

2.1.6 Yukawa coupling

The last missing piece of the mass mystery is the orgin of the mass of the fermions. Naively the mass term in the Lagrangian would look like $-m\bar{\psi}\psi$, however, this term is obviously not invariant under SU(2) transformations. Instead, one can construct a term: $\bar{L}\phi$ which is invariant under SU(2), because ϕ transforms as: $\phi \rightarrow \phi' = (I + ig_W\epsilon(x) \cdot \mathbf{T})\phi$ and $\bar{L} \equiv L^\dagger\gamma^0$ transforms as: $\bar{L} \rightarrow \bar{L}' = L(I - ig_W\epsilon(x) \cdot \mathbf{T})$. When combined with the RH doublet, the $\bar{L}\phi R$ term is invariant under $\mathrm{SU}(2)_L$ transformation, and so is its Hermitian conjugate: $\bar{R}\phi^\dagger L$. For the Higgs field of the form:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h \end{pmatrix}, \tag{2.52}$$

and taking the example of an electron, the Lagrangian is:

$$\mathcal{L} = -g(\bar{L}\phi R + \bar{R}\phi^\dagger L) = -\frac{g_e}{\sqrt{2}}v(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{g_e}{\sqrt{2}}h(\bar{e}_L e_R + \bar{e}_R e_L), \tag{2.53}$$

where the first term has the form required for the fermion masses. The g_e, which is the *Yukawa coupling constant*, takes the form of:

$$g_e = \sqrt{2}\frac{m_e}{v}. \tag{2.54}$$

Rewriting the Lagrangian:

$$\mathcal{L} = -m_e\bar{e}e - \frac{m_e}{v}\bar{e}eh, \tag{2.55}$$

one can see the first term is again the mass term, which originates from the interaction of the massless electron with the vacuum expectation of the Higgs field, and the second term of the electron and the Higgs boson.

One may notice that, the mass term is only acquired through the interaction of the lower part of the weak doublets and of the complex scalar field, which means in this process, only leptons and the down-type quark obtain masses. What about the up-type quark and the neutrinos? Ignoring the neutrinos for now, the up-type quark can acquired by writing the scalar field in its conjugate form of:

$$\phi_c = -i\sigma_2\phi^* = \frac{1}{\sqrt{2}}\begin{pmatrix} -\phi^{0*} \\ \phi^- \end{pmatrix}. \tag{2.56}$$

And with the same Lagrangian, just by replacing ϕ by ϕ_c, the up-type quark can also acquire mass. In conclusion, the Yukawa couplings of the fermions to the Higgs field are given by:

$$g_f = \sqrt{2}\frac{m_f}{v}. \tag{2.57}$$

$g_f = \sqrt{2}\frac{m_f}{v}$. Interestingly, for the top quark with mass $\sim 173.5GeV$, the coupling strength of the top quark to the Higgs field very close to unity. While the neutrinos have such a small mass that they are oftenly considered as massless, the Yukawa coupling will be unnaturally small, suggesting that they might be acquiring their masses in a different way. A possbiliy is the *seesaw mechanism* [14], but it is outside the scope of this chapter.

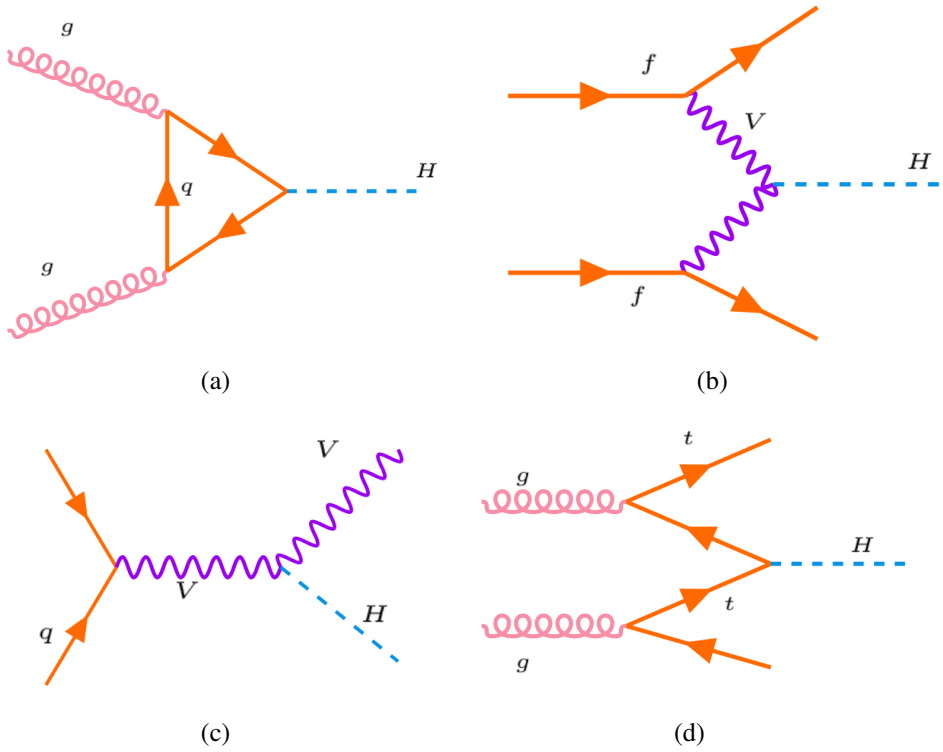


Figure 2.4: Feynman diagrams of the four production mechanism: (a) ggF, (b) VBF, (c) VBFH and (d) ttH.

2.1.7 Higgs boson production at the LHC

In the proton-proton collisions, Higgs bosons are produced via four main mechanism: gluon-gluon fusion (ggF), vector boson fusion (VBF), production associated with a vector boson (VBFH) and production associated with a top or bottom quark-antiquark pair (ttH), as shown in the four Feynman diagrams in Figure 2.4. All these production modes have been observed with cross-sections compatible with the SM prediction, as shown in Figure 2.5.

The dominant production mode is the ggF, which is an order of magnitude greater

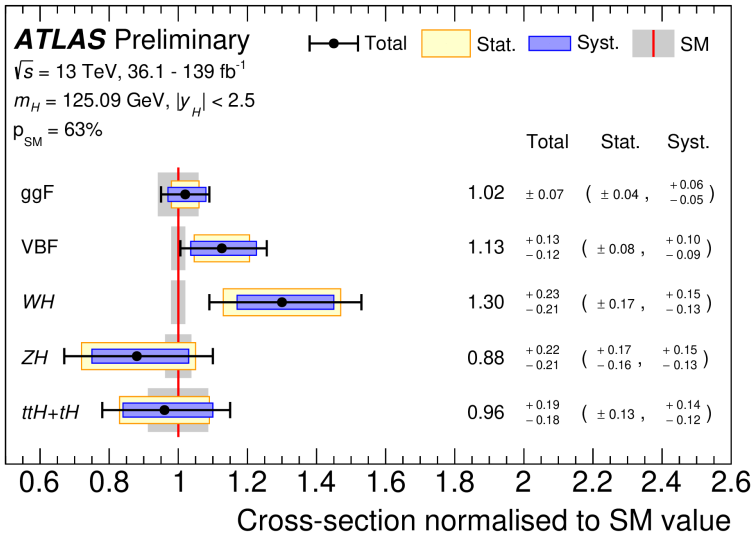


Figure 2.5: Cross sections for ggF, VBF, WH, ZH and ttH+tH production modes. The cross sections are normalised to their SM predictions, measured assuming SM values for the decay branching fractions. The black error bars, blue boxes and yellow boxes show the total, systematic, and statistical uncertainties in the measurements, respectively. The gray bands indicate the theory uncertainties on the SM cross-section predictions. The level of compatibility between the measurement and the SM prediction corresponds to a p-value (p-value is discussed in more detail in section 2.4) of 63%. Image taken from Ref. [15].

than the next largest production mode VBF. The production cross-sections are shown in Figure 2.6 as a function of the center-of-mass energy of proton-proton collisions $\sqrt{s}$.

Once Higgs bosons are produced, it is possible to detect them from their decay products. As discussed in the above sections, the couplings of the Higgs boson to the fermions and bosons are proportional to the mass of the particles, i.e. for fermions: $\alpha \propto m_f \frac{g_W}{2m_W}$ and for W and Z bosons: $\alpha \propto m_W g_W$ and $\alpha \propto m_Z \frac{g_W}{\cos\theta}$ respectively. Given the *branching ratio* is the fraction of all decays that result in a particular final state, $\mathrm{BR}(h \to x) = \frac{\Gamma(h \to x)}{\Gamma}$ with Γ being the decay width, the largest decay branching ratio predicted by the SM of the Higgs boson is to bottom quark (for the Higgs boson mass of 125 GeV), of 58.2%; the next largest decay branching ratio is 21.4% of the decay to a pair of W bosons, where one of them is off-shell [16]. The branching ratios of the other final states are listed in Table 2.1.

The first observations of the Higgs boson were based on approximately 20 fb^{-1} of data (ATLAS and CMS combined) collected from 2011 to 2012, corresponding to a total of approximately 400000 Higgs bosons produced. While this number may seem large, only a very small fraction is picked up by the detector, and even worse, most of the decays involve QCD production of multi-jets final states. Hence it is difficult to distinguish the decays of the Higgs boson from the large QCD background in proton-proton collisions.

Decay mode	Branching ratio
$H \to b\bar{b}$	58.24%
$H \to WW^*$	21.37%
$H \to gg$	8.19%
$H \to \tau^+\tau^-$	6.27%
$H \to c\bar{c}$	2.89%
$H \to ZZ^*$	2.62%
$H \to \gamma\gamma$	0.23%
$H \to Z\gamma$	0.15%
$H \to \mu^+\mu^-$	0.02%

Table 2.1: The SM predicted branching ratios in descending order of the Higgs boson for m_H = 125 GeV. Values taken from Ref. [16].

For this reason, physicists focused on the more distinitive decay channels, such as $H \to \gamma\gamma$ and $H \to WW^*/ZZ^*$, despite the small branching ratio. The results show statistically compelling evidence for the discovery of a new particle with the expected properties of the Higgs boson, which has a significance of 5.9(5.0) σ of the ATLAS (CMS) observations. [1, 2].

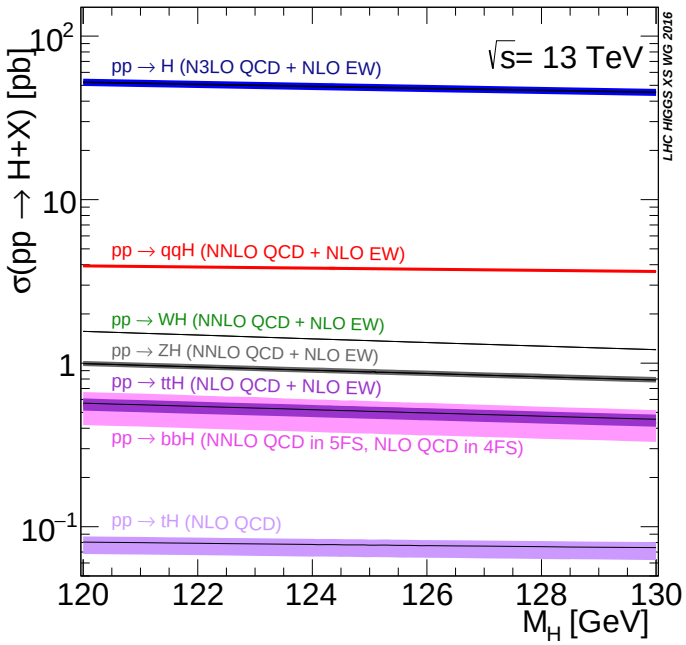


Figure 2.6: The production cross-sections of the SM Higgs boson at $\sqrt{s}$ = 13 TeV. The $pp \to H$ corresponds to the ggF production and the $pp \to qqH$ corresponds to the VBF production. The $pp \to WH, ZH, ttH$ corresponds to the VBFH, ttH. Image taken from Ref. [17].

The interactions between the Higgs boson and fermions were first established by the observations of the Higgs decaying to a pair of τ leptons with a combined significance

of 5.5 σ [18, 19]. While the interaction with bottom quarks were observed a bit latter, the Higgs bosons were observed to decay to two bottom quarks in 2018 by ATLAS and CMS [20, 21]. Even though this decay channel should account for nearly 60% of all Higgs decays at the LHC, it is extremely difficult to spot it amongst the vast number of background QCD particles produced by proton-proton collisions.

To show the interaction strength of the Higgs boson to other particles, it is convenient to quote the reduced coupling-strength modifiers, defined as $\gamma_F = \kappa_F \frac{g_F}{\sqrt{2}} = \kappa_F \frac{m_F}{v}$ for fermions ($F = t, b, \tau, \mu$) and $\gamma_v = \sqrt{\kappa_V \frac{g_V}{2v}} = \sqrt{\kappa_V} \frac{m_V}{v}$ for weak gauge bosons ($V = W, Z$), with κ being the coupling scale factors. The modifiers are shown as a function of their masses m_F and m_V, respectively in Figure 2.7 with the vacuum expectation value of the Higgs field $v = 246$ GeV, where a nice linear connection is drawn across the different particles.

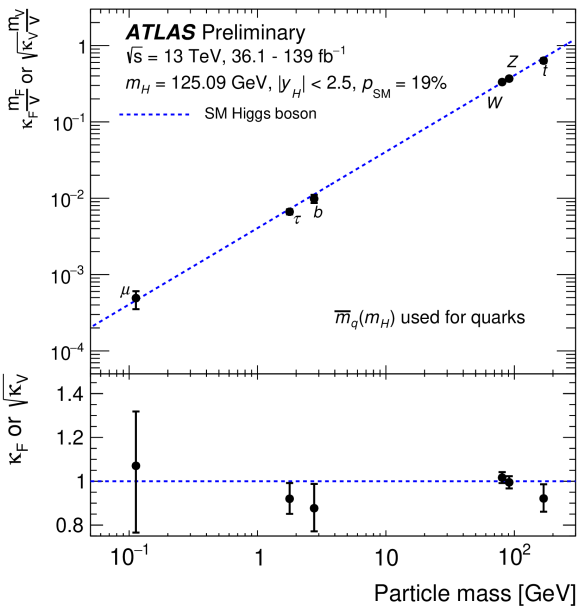


Figure 2.7: The reduced coupling-strength modifiers $\kappa_F \frac{m_F}{v}$ and $\sqrt{\kappa_V} \frac{m_V}{v}$ as a function of their masses m_F and m_V, for vacuum expectation value $v = 246$ GeV. The SM prediction for both cases is also shown (dashed line). The black error bars represent 68% CL intervals for the measured parameters. The lower panel shows the ratios of the values to their SM predictions. Image taken from Ref.[15].

2.1.8 Higgs boson pair production at the LHC

The Higgs boson self-coupling provides direct access to the shape of the Higgs potential, and its measurement is a primary physics goal of the LHC and its forthcoming upgrade, the High Luminosity LHC. It also has crucial implication in electroweak symmetry breaking

mechanism, as discussed earlier in section 2.1.5. To measure it, a direct probe would to measure the Higgs boson pair production, which is the main goal of this thesis.

The di-Higgs production discussed in the section is referred to as the *non-resonant* production, in contrast to the *resonant* production via resonance of an anomalous particle that is not predicted by the SM. The resonant production will be discussed in more detail in section 2.2.

As defined in section 2.1.5, the Higgs potential is given by:

$$V(h) = \frac{1}{2}m_h^2 h^2 + \lambda v h^3 + \frac{\lambda}{4}h^4, \tag{2.58}$$

and to be explicit, the second term is the trilinear self-interaction of the Higgs boson with self-coupling constant $\lambda \equiv \lambda_{HHH}$, responsible for the di-Higgs production and the third term is the quadlinear term, responsible for the triple-Higgs production.

At the LHC, the dominant di-Higgs production mode is via ggF, which contributes approximately 90% of the total cross-section. The second most significant production is via VBF which is also considered in this thesis.

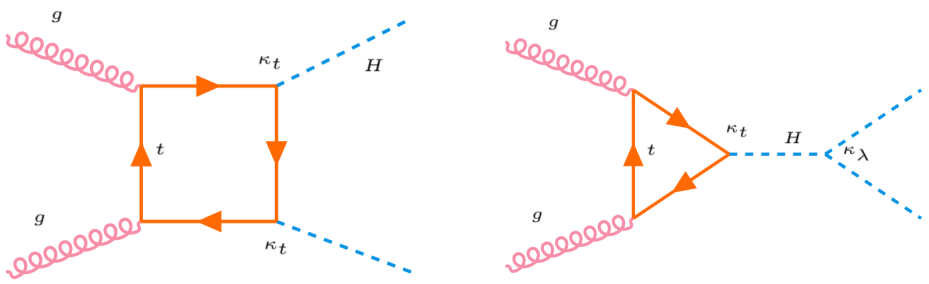
The leading order Feynman diagrams via ggF production are shown in Figure 2.8, where the left is referred to as the *box diagram* and the right is called the *triangle diagram*. In the box diagram, the two Higgs bosons are produced via two ttH vertices, and hence the interaction amplitude is proportional to the square of the top yukawa coupling, g_t^2. As a result, the box diagram is not sensitive to the self-interaction constant λ_{HHH}. In contrast, the triangle diagram has direct acess to λ_{HHH}, where a Higgs boson decays to two Higgs bosons. The interaction amplitude is proportional to the multiple of the top yukawa and the Higgs self-interaction constant, $g_t\lambda_{HHH}$.

It would be convenient to define the coupling modifers κ_t and κ_λ which will be used throughout the text. The modifiers are defined as $\kappa_t = g_t/g_t^{\mathrm{SM}}$ and $\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{\mathrm{SM}}$, where g_t and λ_{HHH} are the measured values and g_t^{SM} and $\lambda_{HHH}^{\mathrm{SM}}$ are the values predicted by the SM.

While the pair production of the Higgs boson occurs at a very small rate due to the small phase space of decaying to two off-shell Higgs, these two diagrams interferes destructively, and therefore make the pair production cross-section even smaller. The dominant production mode is via ggF, and the production cross-section calculated at next-to-next-to-leading order (NNLO), taking into account the finite top-quark mass assumption (FTApprox) [22] is given by:

$$31.05^{+2.2\%}_{-5.0\%}(\mathrm{scale}) \pm 2.1\%(\alpha_S) \pm 2.1\%(\mathrm{PDF}) \pm 2.6\%(\mathrm{m_{top}})\ \mathrm{fb}, \tag{2.59}$$

at $\sqrt{s} = 13$ TeV and $\mathrm{m}_H = 125$ GeV [17]. The scale uncertainty is due to the finite order

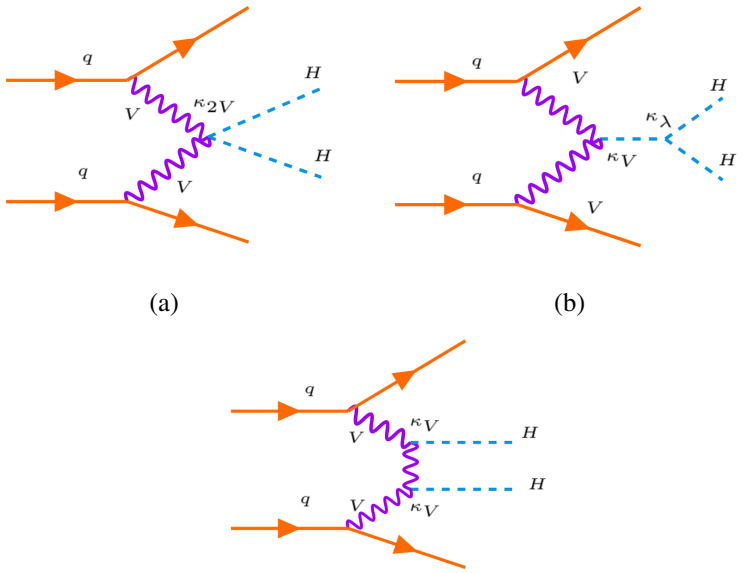


(a) (b)

Figure 2.8: Leading order Feynman diagrams of (a) Box diagram and (b) triangle diagram of the di-Higgs production.

of QCD calculations, the α_s and PDF terms account for the uncertainties on the strong coupling constant and parton distribution functions respectively, and the m_{top} uncertainty is related to the top-quark mass scheme.

The VBF di-Higgs production is also considered in this thesis. The Leading order Feynman diagrams are shown in Figure 2.9. The vertices denoted by κ_{2V}, κ_V and κ_λ represent the $VVHH$, VVH and HHH couplings modifiers, respectively. The cross-



(a) (b)

(c)

Figure 2.9: Leading order Feynman diagrams of the VBF HH production.

section is calculated at next-to-next-to-next-to-leading order (N3LO) in QCD in the limit

in which there is no partonic exchange between the two protons [23], which is given by:

$$1.726^{+0.03\%}_{-0.04\%}(\text{scale}) \pm 2.1\%(\text{PDF} + \alpha_S)\ \text{fb}, \tag{2.60}$$

at $\sqrt{s} = 13$ TeV and $m_H = 125$ GeV [17].

2.1.8.1 The $b\bar{b}\tau^+\tau^-$ decay channel

The final states of the di-Higgs production can be one of many possible combinations of single-Higgs decays. The dominant decay mode is to $b\bar{b}b\bar{b}$, with a branching ratio of 33%, as shown in Figure 2.10. The main focus of this thesis is the $b\bar{b}\tau^+\tau^-$ decay channel which accounts for 7.3% of the total decay channels. In chapter **??**, a search for the Higgs pair production in the $b\bar{b}\tau^+\tau^-$ channel is presented. The results of the search is also combined with the $b\bar{b}b\bar{b}$ channel and the $b\bar{b}\gamma\gamma$ to maximise the sensitivity. The $b\bar{b}b\bar{b}$ channel has the largest branching ratio, however, it had been shown to be very difficult to extract the signal from the vast QCD background, which can be seen from the late discovery of the $H \to b\bar{b}$ decay [20, 21]. The opposite is the $b\bar{b}\gamma\gamma$ channel, which has very clean backgrounds but with a much smaller branching ratio, which is only 0.26%. The $b\bar{b}\tau^+\tau^-$ channel has the advantage of both two channels, that it has a relatively high branching ratio and relatively clean background. Therefore the $b\bar{b}\tau^+\tau^-$ channel has great sensitivity.

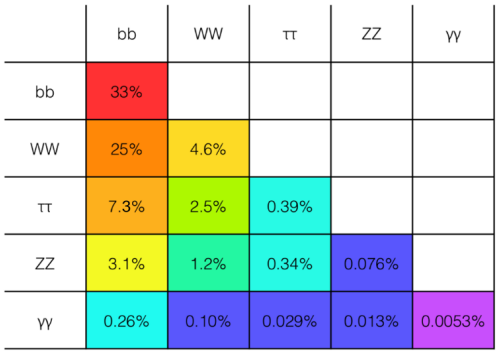


Figure 2.10: The most common di-Higgs decay channels and the corresponding branching ratios. Image made by Alessandra Betti.

The two τ-leptons in the final state subsequently decay either leptonically or hadronically, as described in section **??**. Processes with final state with a leptonically decaying τ and a hadronically decaying τ, which accounts for 42.0% of the di-τ decays are categorised as the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. Processes with both τ decay hadronically are categorised as the $\tau_{\text{had}}\tau_{\text{had}}$ channel, which accounts for 45.6%.

The search for di-Higgs production in the $b\bar{b}\tau^+\tau^-$ channel using the early Run 2 data recorded by the ATLAS detector during 2015 and 2016 [24] set the world-best observed

upper limit at that time on the production cross-section. The cross-section times branching ratio for non-resonant di-Higgs production is constrained to be less than 30.9 fb, 12.7 times the Standard Model expectation, at 95% confidence level. As combined with the results in the $b\bar{b}b\bar{b}$, $b\bar{b}W^+W^-$, $W^+W^-W^+W^-$, $b\bar{b}\gamma\gamma$ and $W^+W^-\gamma\gamma$ channels, the limit is tightened to 6.9 times the SM expectation [25]. With the full Run 2 data and various improvements, the author will present readers the exciting results in chapter **??**.

In the search for di-Higgs production presented in chapter **??**, the normalisation of ggF non-resonant HH production is set to the production cross-section times the $b\bar{b}\tau^+\tau^-$ branching ratio,

$$\sigma_{ggF} \times BR_{b\bar{b}\tau^+\tau^-} = 31.05\,\text{fb} \times 0.073 = 2.268\,\text{fb}; \tag{2.61}$$

while similarly for the VBF production the normalisation is set to:

$$\sigma_{VBF} \times BR_{b\bar{b}\tau^+\tau^-} = 1.726\,\text{fb} \times 0.073 = 0.1261\,\text{fb}. \tag{2.62}$$

Other non-resonant HH production modes are not considered as their contributions to the analysis sensitivity are expected to be negligible.

2.2 Beyond the Standard Model

Despite the huge success of the SM, there are many unanswered questions in the SM which were outlined at the beginning of this chapter. Many theories beyond the SM have been developed, which try to include the phenomena not explained in the SM and to solve some of the theoretical issues. In this section the limitations of the SM theory are summarised and a few BSM models, considered in this thesis work as they could be observed in the $b\bar{b}\tau^+\tau^-$ final state, are introduced.

2.2.1 Big questions in the Standard Model

Some of the important questions unanswered by the SM are outlined in the following:

- How to integrate gravity into the SM? Only three fundamental interactions are considered in the SM. Attempts to include gravity in the quantum field theory result in a theory which is not *renormalisable* (which predicts infinite values for observables such as particle masses). Hence gravity is ignored in the SM due to its small interaction strength, but that means that the SM will break down at large gravitational scales.

- What is the nature of dark matter and dark energy? Measurement of the cosmic

microwave background radiation show that the directly observable SM matter is only 5% of the energy of the universe [26]. while the dark matter and dark energy makes up about 27% of the universe and 68% of the universe, respectively, where the ratio of dark matter is estimated using measurements of galaxy rotation curves [27] and gravitational lensing [28], and the ratio of dark energy is estimated based on the rate of the universe expansion [29].

- Where does the matter-antimatter assymmetry come from? The physical world is made of matter instead of antimatter. At the early stage of the universe, the creation and annihilation of the matter-antimatter was in equilibrium state, but when the universe started to cool down matter and antimatter could only annihilate. For matter to survive the annihilation, one of the three Sakharov [30] condition requires the *CP-symmetry* to be violated, which is the combination of the charge symmetry and the parity symmetry. This phenomenon is observed during certain types of weak decay, however the violation in the SM is too small to account for the matter-antimatter assymmetry.

- What is the origin of neutrino masses? In the SM, there are no RH neutrinos, and there are only Higgs doublets of $\mathrm{SU}(2)_L$, and therefore neutrinos are required to be massless. However, observations in neutrino oscillations [31, 32] imply that the neutrinos have non-zero mass. If neutrinos are normal Dirac particles, this will imply an unnaturally small Yukawa coupling to the Higgs field. An other possibility to explain its smallness of mass is that neutrinos are *Majorana* particles, meaning that neutrinos are their own particles. If this is true, processes violating the lepton-number conservation can be allowed, such as the neutrino-less double β-decay.

- Why there are so many free parameters? As mentioned in section 2, in the SM, there are 26 free parameters that have to be input by hand, including the masses of the 12 fermions, the three coupling constants describing the strengths of the gauge interactions, the two parameters describing the Higgs potential, the eight mixing angles of the CKM and the *PMNS* (the Pontecorvo-Maki-Nakagawa-Sakata matrix, which accounts for neutrino oscillation) matrices, and the CP violation phase in the strong interaction.

- How to solve the *hierarchy problem*? The Higgs mass is much smaller than the Planck mass $\mathcal{O}(10^{19}$ GeV). The Higgs mass is corrected by quantum loops which are proportional to the square of the energy scale, Λ, and at Planck scale $\Lambda_P \sim 10^{19}$ GeV these correction becomes very large (if the SM is still valid). These very large corrections need to be precisely canceled out to leave the observed Higgs mass

of $m_H = 125$ GeV. This cancellation requires a high level of fine-tuning and it is therefore considered unnatural, which is known as the hierarchy problem.

- Can the forces be unified? In the mid 1970s, it was suggested by Georgi and Glashow that the observed gauge symmetries of the SM can be accommodated within a larger SU(5) symmetry group. In this Grand Unified Theory (GUT), the coupling constants of the SM are found to converge, although not exactly, at the energy scale of about 10^{15} GeV.

2.2.2 Two-Higgs-doublet Model

In the SM, the Higgs mechanism assumes a doublet of complex scalar fields. While this is the simplest choice, it is not unique. The most relavant BSM model to this thesis is the two-Higgs-doublet model (2HDM), which is one of the simplest possible extensions of the SM, first proposed by Tsung-Dao Lee in 1973 [33]. He assumed two doublets of complex scalar fields to create a spontaneously CP-violating theory. Introducing an additional scalar field might induce flavour-changing neutral currents, but there are several ways to arrange the yukawa couplings so that there is natural flavour conservation [34].

Nowadays, there are many motivations for 2HDMs, the best known of which might be supersymmetry [35]. The basic idea of supersymmetric theories is that fermions and bosons are related to their super-partners which differ in spin by 1/2. The introduction of these new particles can solve some of the limitations in the SM described above, such as the hierarchy problem, force unification and the origin of dark matter. In supersymmetric theories, the scalars belong to chiral multiplets and their complex conjugates belong to multiplets of the opposite chirality; since multiplets of different chiralities cannot couple together in the Lagrangian, a single Higgs doublet is unable to give mass simultaneously to the charge 2/3 and charge -1/3 quarks. Therefore, an additional doublet is required.

With some simplifying assumptions, the potential of the two Higgs doublet Φ_1 and Φ_2 with hyper charge +1 can be written as [36]:

$$\begin{aligned} V =& m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - m_{12}^2 (\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1) + \frac{\lambda_1}{2} (\Phi_1^\dagger \Phi_1)^2 + \frac{\lambda_2}{2} (\Phi_2^\dagger \Phi_2)^2 \\ &+ \lambda_3 \Phi_1^\dagger \Phi_1 \Phi_2^\dagger \Phi_2 + \lambda_4 \Phi_1^\dagger \Phi_2 \Phi_2^\dagger \Phi_1 + \frac{\lambda_5}{2} \left[(\Phi_1^\dagger \Phi_2)^2 + (\Phi_2^\dagger \Phi_1)^2 \right], \end{aligned}$$

where all parameters are real. The vacuum state is then given by:

$$\Phi_1 = \begin{pmatrix} 0 \\ \frac{v_1}{\sqrt{2}} \end{pmatrix}, \Phi_2 = \begin{pmatrix} 0 \\ \frac{v_2}{\sqrt{2}} \end{pmatrix}, \tag{2.63}$$

which give eight fields:

$$\Phi_a = \begin{pmatrix} \phi_a^+ \\ (v_a + \rho_a + i\eta_a)/\sqrt{2} \end{pmatrix}, a = 1, 2, \tag{2.64}$$

where the ϕ_a^+, v_a, ρ_a η_a are the four fields of pertubations around the vacuum state Φ_1 and Φ_2, which makes a total of eight. Three of the eight scalar fields are Goldstone bosons that give mass to the W and the Z bosons, the remaining five fields correspond to five physical Higgs bosons: two CP-even neutral scalars h and H_0, two charged scalar particles $H^\pm$, and a CP-odd neutral pseudoscalar A_0.

Like in the single complex field case, the mass terms arise from the square of the field and the mass terms of the neutral scalars h and H_0 can be represented in matrix form as:

$$\mathcal{L}_{mass}^{\psi=\rho} = (\psi_1, \psi_2)\ M(v_1, v_2, \lambda_{1,2,3,4}, m_{12})_\psi \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{2.65}$$

where $M(v_1, v_2, \lambda_{1,2,3,4})_\psi$ is the matrix encapsulating the coefficients of the scalar mass terms, which can be diagonalised by a rotation angle α [36]. Similarly for the pseudoscalar A_0 and the two charged scalars $H^\pm$, the mass term is given by:

$$\mathcal{L}_{mass}^{\psi=\eta,\phi^\pm} = (\psi_1^{(+)}, \psi_2^{(+)})\ M(v_1, v_2, \lambda_{4,5}, m_{A,12})_\psi \begin{pmatrix} \psi_1^{(-)} \\ \psi_2^{(-)} \end{pmatrix}, \tag{2.66}$$

and the diagonalisation angle is defined as $\tan\beta \equiv \frac{v_2}{v_1}$ [36]. These two parameters α and β determine the interactions of the different Higgs field with the vector bosons and the fermions (once the Yukawa coupling strength is provided). In the limit where $\cos(\beta-\alpha) \to 0$, referred to as 'decoupling limit' (or 'alignment limit'), the lighter neutral scalar h presents almost the same properties as the SM Higgs boson [37]. The heavier neutral scalar H, on the other hand, could be generated at the LHC which subsequently decays to two Higgs boson, as shown in the Feynman diagram at leading order in Figure 2.11. Therefore, the resonant di-Higgs production is of interest of this thesis.

In chapter **??**, searches for the di-Higgs production from resonance of a generic heavy spin-0 neutral scalar is presented. The resonance particle is assumed to have a narrow decay width, and the normalisation of resonant production in the analysis is set to

$$\sigma \times BR_{b\bar{b}\tau^+\tau^-} = 1\,\text{pb} \times 0.073 = 0.073\,\text{pb}, \tag{2.67}$$

where dummy cross-section of 1 pb is chosen for the resonant production.

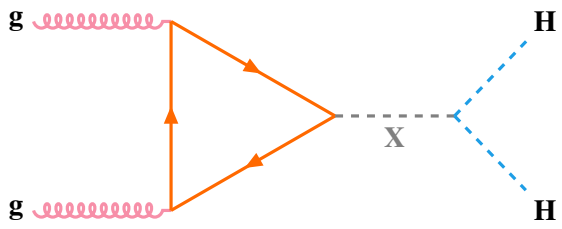


Figure 2.11: SM leading-order Feynman diagrams for Higgs boson pair production through resonance of a generic spin-0 scalar particle.

2.2.3 Effective field theory interpretation

Instead of a *theory of everything*, physicists' focus these days is on less ambitious but more practical *theories of something*, which describe particular physical systems in particular conditions. Such theories are seen as effective fields theories (EFTs) because they are not meant to be valid at all energy scales, and often the degrees of freedom they describe are emergent rather than fundamental.

The basic idea behind EFTs is that things simplify when viewed from a distance. In particle physics, 'viewing from far away' is equivalently saying that the energy scale being considered is low. The SM is commonly accpeted as an effective theory applicable up to energies not exceeding a certain scale Λ. For a field theory valid above that scale, it needs to satisfy a few requirements:

- it should have a gauge group that contains the SM $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ group,
- it should incorporate all degrees of freedom of the SM,
- it should reduce to the SM at low energies.

In most cases, the reduction to the SM is achieved by decoupling the heavy particles with masses of order Λ or larger. If one writes down the Lagrangian of a BSM theory, the BSM part will be suppressed by powers of Λ:

$$\mathcal{L}_{BSM} = \mathcal{L}_{SM}^{(4)} + \frac{1}{\Lambda}\sum_k c_k^{(5)} O_k^{(5)} + \frac{1}{\Lambda^2}\sum_k c_k^{(6)} O_k^{(6)} + \mathcal{O}\left(\frac{1}{\Lambda^3}\right), \tag{2.68}$$

where $\mathcal{L}_{SM}^{(4)}$ is the SM Lagrangian which contains dimension-two and -four operators only, and the BSM physics is encapsulated by operators of dimension-five, dimension-six $O_k^{(5)}$, $O_k^{(6)}$ and of higher dimensions (suppressed by higher order of Λ) [38]. The $c_k^{(n)}$ are the dimensionless coupling constants (*Wilson coefficients*).

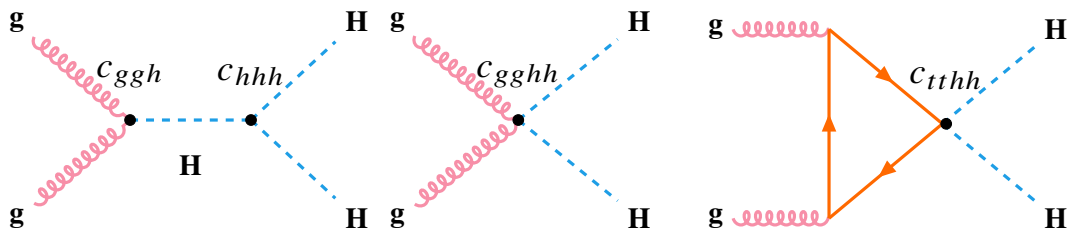
It has been proven in Ref. [38] that the dimension-five operator (insterstingly, there is only one of such operator) violates the lepton number conservation, while at the LHC this

effect is almost unobservable, therefore it is not considered. On the other hand, it is possible to write down a large number of dimension-6 operators, for example, just attaching an extra term $\phi^\dagger\phi$ to any of the terms in the SM Lagrangian will be dimension-six. It was shown in Ref. [39] there are 59 of them. Some BSM operators have the effect of rescaling the Higgs boson coupling to other particles, and some create new anomalous couplings which were not allowed in the SM. In both scenarios the di-Higgs production receives corrections on the production cross-section, in some cases the cross-section is greatly enhanced, making the observation of di-Higgs possible with current LHC data.

One major benefit of the EFT approach is that it is 'model independent'. Same set of low energy operators (the ones of the SM) can be reused for different BSM theories, in contrast to the 'model dependent' approach, where one searches for the signatures of a specific new particle(s).

There are two approaches to treating the SM as an EFT: Standard Model EFT (SM EFT) [39] and Higgs EFT (HEFT) [40, 41]. The SM EFT in general is a simpler framework due to more restrictive symmetries: the operators follow the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry. On the other hand, the only manifest gauge symmetry of HEFT is $\mathrm{SU}(3)_C \times \mathrm{U}(1)_{\mathrm{em}}$, while the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry is non-linearly realised. In this formalism, anomalous Higgs boson couplings are expected to be the dominant effects on new physics in the electroweak sector. Deviations from SM predictions can potentially be observed via di-Higgs production using the HEFT framework.

In the HEFT Lagrangian, ggF HH production is described at LO with 5 operators and their corresponding Wilson coefficients: $c_{hhh}, c_{tth}, c_{tthh}, c_{ggh}$ and c_{gghh}. The first two Wilson coefficients are the couplings modifiers of the HHH, ttH vertices, i.e. $c_{hhh} \equiv \kappa_\lambda = \lambda_{HHH}/\lambda^{\mathrm{SM}}_{HHH}$ and $c_{tth} = \lambda_{ttH}/\lambda^{\mathrm{SM}}_{ttH}$, with λ_{HHH} being the Higgs boson self-coupling constant and λ_{ttH} being the top-quark Yukawa coupling. While these two modifiers are responsible for SM vertices, the rest accounts for the Non-SM interactions, affect respectively the $ttHH$, gHH and $ggHH$ vertices. The Feynman diagrams for such interactions and the corresponding Wilson coefficients are shown in Figure 2.12.



(a) ggh and hhh (b) gghh (c) tthh

Figure 2.12: BSM HEFT leading order Feynman diagrams for Higgs boson pair production through gluon-gluon fusion.

The HEFT Lagrangian reduces to the SM Lagrangian for $c_{hhh} = c_{tth} = 1$ and $c_{tthh} = c_{ggh} = c_{gghh} = 0$, when the Higgs boson self-coupling and top-quark Yukawa coupling have SM values and none of the BSM production modes are present.

The di-Higgs production gives a unique access to the coefficients c_{hhh}, c_{tthh} and c_{gghh} (the c_{ggh} and c_{tth} coefficients can be probed by single-Higgs boson production, with even higher sensitivity). The HEFT formalism allows one to interpret general searches in different BSM scenarios by simultaneously varying multiple Wilson coefficients. In section **??**, results one dimensional scans on coefficients c_{hhh}, c_{tthh} and c_{gghh} is presented, assuming that all other couplings take their SM values. The resultsm also represent the first dedicated scan on the c_{tthh} and c_{gghh} coefficients in ATLAS. In addition, *cluster analysis* [42] can be used to group the shapes of the m_{HH} distribution that are predicted by HEFT. At next-to-leading order (NLO), 7 benchmarks (BM) [43], defined in Table 2.2, describe representative shape features of the HEFT BSM m_{HH} distribution (for example, a high peak in the low m_{HH} region, double peaks, enhanced tail), and can be used to explore multiple BSM scenarios. In section **??**, upper limits are set on the production cross-section assuming the 7 benchmarks.

Benchmark	c_{hhh}	c_{tth}	c_{tthh}	c_{ggh}	c_{gghh}
SM	1	1	0	0	0
BM 1	3.94	0.94	$-1/3$	0.5	$1/3$
BM 2	6.84	0.61	$1/3$	0.0	$-1/3$
BM 3	2.21	1.05	$-1/3$	0.5	0.5
BM 4	2.79	0.61	$1/3$	-0.5	$1/6$
BM 5	3.95	1.17	$-1/3$	$1/6$	-0.5
BM 6	5.68	0.83	$1/3$	-0.5	$1/3$
BM 7	-0.10	0.94	1	$1/6$	$-1/6$

Table 2.2: HEFT Wilson coefficient values in the SM and in seven BSM benchmark hypotheses defined in Ref [43].

2.3 Machine learning theory

A machine learning (ML) algorithm, which is widely considered to be general in nature and task-independent, 'learns' from more and more training data and improve its performance on some given tasks. High energy physics research and analysis have been using ML algorithms for some time; for example, ML algorithms play an important role in boosting the physics performance of reconstruction; they help reducing the execution time of computationally-expensive event simulation and calibration; they provide strong analysis power for searches for BSM physics or probing the SM with increasing precision.

ML algorithms are commonly used for two types of problems, *classification* and *regression* [44]. In a classification problem, variables relevant to the physics problem are selected and a ML model is trained using signal and background events (or instances). Then the model learns how to assign a class label to the data, for example to classify whether an event is signal or background. In a regression problem, a continuous function is learned, an example is to obtain the best estimate of a particle's energy based on the measurements from multiple detectors.

Early ML applications in HEP often used decision trees: a tree like model for decisions, starting at the root, climbing up the branches and reaching the leaves, where each leaf represents a decision [45]. For classification problems, each leaf represents one's decision assigning a data item to a class (binary or multiclass problems). In high energy physics, the most widely used trees are boosted decision trees (BDT), which combine many individual trees with weights assigned to each tree, where the weights are 'boosted' if the event is classified successfully.

Another class of commonly used ML algorithms is the artificial neural networks (ANN or just NN), which is the ML algorithm used in this thesis. As compared to the traditional cut-based approach which sets requirements on one or more variables, the NN exploits information on multiple variables and the correlations between them. The cut-based approach is analogous to cutting a hyper-cube in the phase space formed by the variables of an event, while the NN is similar to averaging many hyper-cubes in the phase space, and therefore it can describe the truth shape better.

Neural networks aim to mimic the biological brains in a simplified manner, where the neurons and synapses are represented with connected layers of nodes and their connections. The connections between nodes are quantified by weights. A positive weight reflects an excitatory connection, while negative values mean inhibitory connections. All inputs are modified by a weight and summed as a linear combination. Finally, an *activation function* controls the amplitude of the output. For example, an acceptable range of output is usually between 0 and 1, or it could be -1 and 1. A simple example of a neural network is shown in Figure 2.13.

More specifically, layers in a NN include an input, an output, and one or multiple hidden layers. These connections are quantified by weights w_i, where i represents the index of the input node in the previous layer. The input to a node in $n+1^{\text{th}}$ layer is given by the weighted sum of the outputs of the nodes in the previous layer (the n^{th} layer):

$$y^{n+1} = \sum_{i}^{N^n} w_i^n x_i^n + b^n, \tag{2.69}$$

where N_n is the number of nodes, y^{n+1} is the intput to the node, x_i^n is the output of the i^{th}

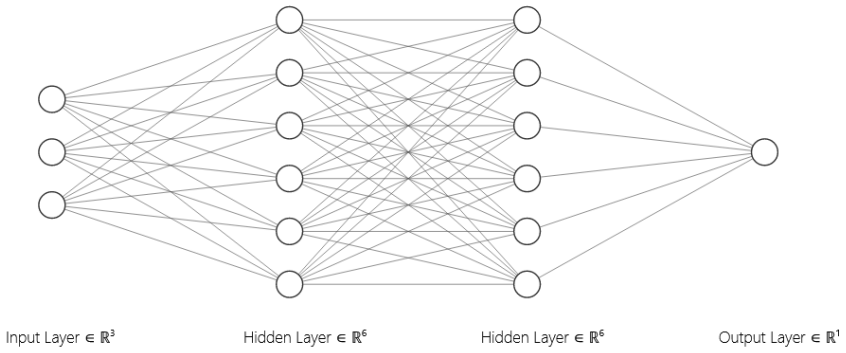


Figure 2.13: A simple neural network with 3 input nodes, two hidden layers each with 6 nodes, and a output node.

node, b^n is the *bias* which shifts the $w_i^n x_i^n$ by a certain amount. A node then takes this input and performs a non-linear transformation using an activation function to form its output. There are a few frequently used activation functions, such as the sigmoid function: $x^{n+1} = \frac{e^{y^{n+1}}}{e^{y^{n+1}}+1}$, or simply $\tanh(y^{n+1})$, where the output is limited to range $[-1, 1]$; or the rectified linear unit (ReLU), given by $x^{n+1} = max(0, y^{n+1})$, that has output in range $[0, 1]$.

Once the architecture of the NN is chosen, the next question is how to determine the weights and bias of the network. The performance of the network for a certain task can be quantified by the *loss function*. While many different types of functions serve the purpose, one can consider a simple example that $L(y, \hat{y}) = \frac{1}{m}\sum_{i=1} m(y_i - \hat{y}_i)^2$, where $L(y, \hat{y})$ is the loss function, y is the expected output and $\hat{y}$ is the actual output that the network gives, for m test data. To improve the performance of the NN, or equivalently minimising the loss function, one can 'train' the NN provided with data. The training data for the NN used in this thesis consists of pairs of lablled input variables and expected output, meaning the training is *supervised*. The data set used for training can be categorised into training, validation and test data sets, where the first two are frequently combined together for cross-validation.

During training, the gradient of the loss function with respect to the weight and bias is computed. The weight and bias are varied in the negative gradient direction by a small amount. This process is performed for multiple iterations (*epochs*) until the loss function reaches its minimum. The size of each step is control by the *learning rate*. This iterative process is known as *gradient descent*.

For a large training set, the computational time for the gradient of the loss function might be very long, since the loss function is function of all data set. In this thesis, the *stochastic gradient descent* [46] method is used. Instead of computing the gradient using all

data, the data set is divided into small *batches* which are used to calculate an approximate gradient.

The similar problem happens when the number of connections become large, which can happen quite easily even with a small number of nodes in each layer. Commonly the *backpropagation* algorithm is used, which adjust the weights and bias of one layer according to the expected output of the next layer. For example, if the final output of the network is 0.5 while one expects 1, the output can be increased by boosting the weights of the connections to the previous layer with positive weights and to nodes with large values, and vice versa. The weights are then updated from one layer to the layer before it (hence the term 'backpropagation') using the chain rule.

In addition, a *momentum* is used in this thesis to the gradient descent algorithm that allows the search to build inertia in a direction in the search space and overcome the oscillations of noisy gradients and cross over flat spots of the search space.

The architecture of the NN (number of layers and nodes), learning rate, number of batches, choice of the activation functions and momentum are called collectively as *hyper-parameters*. Unlike the weights which are learned during training, the hyperparameters are set before training is performed. Depending on the data patterns to be learned or abstracted, different values of the hyperparameters will be needed for the same ML tool. The optimisation of the hyperparameters of the NN used in this thesis is outlined in section **??**.

2.4 Statistical Interpretation

Once the experiment has been conducted and the data has been recorded, physicists might be interested in asking questions such as, did one or did one not establish a discovery? Or, how well does an alternate model describe this discovery? The first question has to do with the goodness of the fit of the observed data to the good and old Standard Model, while the second question has to do with hypotheses testing and the derivation of confidence intervals and upper limits [47]. In order to answer these questions, one needs to first define the null hypothesis H_0 and the alternative hypothesis H_1. The definition of H_1 and H_0 depends of the specific physics problem, using the tone of the particle physics, for search for a new signal process, the null hypothesis H_0 is assuming only the background is observed, the alternative hypothesis H_1 is assuming one observes both signal and background; for problems such as setting upper limits, this definition is reversed.

It is common to refer to the *p-value* to quantity the deviation from the null hypothesis, which is defined as:

$$p = \int_{x_{obs}}^{\infty} g(x|H_0)dx, \tag{2.70}$$

where $g(x|H_0)$ is the probability density function of observing a quantity x given the null

hypothesis, and x_{obs} is the observed value. The p-value derivation is shown illutratively in Figure 2.14, together with α, which is defined as:

$$\alpha = \int_{x_{5\sigma}}^{\infty} g(x|H_0)dx, \tag{2.71}$$

where $x_{5\sigma}$ is the observed value that corresponds 5σ deviation from the null hypothesis. One can therefore claim a discovery if the observed p-value $p < \alpha$ (equivalently, the background-only null hypothesis is rejected with a probability $1 - p$). Another commonly used statistic jargon is the *confidence level* (CL), which is defined as:

$$CL = \int_{-\infty}^{x_{5\sigma}} g(x|H_0)dx \equiv 1 - p. \tag{2.72}$$

Lastly, the p-value is oftenly converted to a significance Z, defined as the number of standard deviations above the mean of a Gaussian distribution. The relation between the p-value and Z is given by:

$$Z = \Phi^{-1}(1 - p), \tag{2.73}$$

where Φ^{-1} is the quantile function of the standard Gaussian distribution. The discovery of signal, as mentioned before, needs to be rejecting the background-only hypothesis with at least $Z \geq 5$, which corresponds to a p-value of 2.87×10^{-7}. In the case of setting upper limits, for excluding the signal hypothesis, p-value is by convention required be less than 0.05 (95% confidence level), which is equivalent to $Z = 1.64$. This value is also used for setting upper limits in chapter **??**. N.B. In chapter **??**, the CL used for setting upper limits is calculated by the CL_s method [48] where CL_s is (unfortunately named CL_s) given by:

$$CL_s = \frac{p_{s+b}}{1 - p_b}, \tag{2.74}$$

where p_{s+b} and p_b are the confidence leveles for the signal-plus-background hypothesis and the background-only hypothesis, respectively. The reason for using the CL_s instead of the p_{s+b} is that, in case of an experiment with low sensitivity, $1 - p_b$ becomes small, and hence the signal will not be excluded even with a small p-value.

2.4.1 Test statistics

A *test statistic* is a quantity calculated from data, which can be used to estimate how probable is the result that we observe with respect to some null hypothesis. In terms of hypothesis testing, the observable x is usually replaced by a test statistic t, so that the

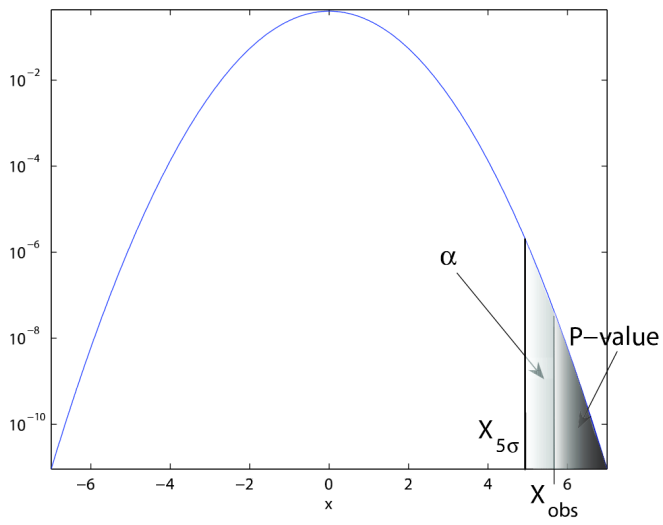


Figure 2.14: An illustration showing the control area α and the p-value of a Gaussian distribution. Note, in this example $X_{obs} > X_{5\sigma}$. Image taken from Ref.[47]

.

p-value becomes

$$p = \int_{t_{obs}}^{\infty} g(t|H_0)dt. \tag{2.75}$$

Due to the Neyman-Pearson lemma [49], the most powerful test statistic one can construct is given by:

$$t = \frac{L(H_1)}{L(H_0)}, \tag{2.76}$$

where $L(H_1)$ and $L(H_0)$ are the likelihood fuction of the hypothesis H_1 and H_0, respectively. The analsysis discussed in this thesis in most of the time involves signal and background, distributed in some histograms $\mathbf{n}$ of N bins: $\mathbf{n} = (n_1, n_2, ..., n_N)$. In such case, the expectation of each bin is consists of the background b and the signal s with some signal strength μ:

$$E[n_i] = \mu s_i + b_i. \tag{2.77}$$

In addition to the signal strength, it is common that the signal and background will depend on some additional parameters, as referred to *nuisance parameters*, that is not of direct insterst of the analysis but needs to be fitted from the data. Hence, the expected signal and background is determined by the total signal and background and the respective probability function f and nuisance parameters θ [50]:

$$s_i = s_{tot} \int_i f_s(x; \theta_s)dx, \tag{2.78}$$

$$b_i = b_{tot} \int_i f_b(x; \theta_b) dx. \tag{2.79}$$

In addition, these nuisance parameters are oftenly constrained by additional subsidiary measurements. Suppose the measurements are distributed in a new histogram $\mathbf{m} = (m_1, ..., m_M)$, and the expectation values of bin m_i is given by $E[m_i] = u_i(\theta)$.

Since in a counting experiment, the data follows a Poisson distribution, the likelihood function can be written as:

$$L(\mu, \theta) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{N} \frac{u_k^{m_k}}{m_k!} e^{-u_k}. \tag{2.80}$$

One can define a profile likelihood ratio, such that:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})}, \tag{2.81}$$

where $\hat{\theta}$ and $\hat{\mu}$ are the maximum likelihood estimators, that

$$(\hat{\mu}, \hat{\theta}) = argmax\ L(\mu, \theta). \tag{2.82}$$

and $\hat{\hat{\theta}}$ is the θ that maximise L for the specific μ. As a result, the profile likelihood ratio has value $0 \leq \lambda \leq 1$, and for λ value close to 1, the hypothesis is well compatibile with the observed data.

Using the profile likelihood ratio, one can define the test statistic as:

$$t_\mu = -2ln\lambda(\mu), \tag{2.83}$$

and therefore, t_μ approaches to 0 from the positive direction, and corresponds to a large p-value. For establishing a discovery of a positive signal, one can assume the signal strength $\mu \geq 0$. If data fluctuates below the expected background, i.e. $\hat{\mu} < 0$, it may constitute evidence against the background-only model, but this does not show that the data contain signal events and the most physical explaination is that this is due to some systematic error. Therefore, for **signal discovery**, the test statistic of the background-only null hypothesis given by:

$$t_0 = q_0 = \begin{cases} -2ln\lambda(0) & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0. \end{cases} \tag{2.84}$$

Consider the opposite case, for **setting upper limits** on a signal, one needs the test statistic for the signal-plus-background null hypothesis with signal strength μ, and if data fluctuates above signal plus background, i.e. $\hat{\mu} > \mu$, the signal is not likely to be excluded

and most likely is due to systematic errors. Therefore:

$$t_\mu = q_\mu = \begin{cases} -2ln\lambda(\mu) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} \tag{2.85}$$

2.4.2 Look elsewhere effect

Suppose one specify an hypothesis with a specific Higgs mass and observed some signal at some mass range, one should also take into account that this signal could be a fluctuation which could be observed anywhere in the sensitivity range [51]. This effect is referred to as *look-elsewhere effect*. A common belief is that the such effect is the reason for the habit of defining a discovery as a 5 σ and not for example 4 σ, because even if one quote 5 σ the effective significance might be lower. In this thesis, the look-elsewhere effect is accounted by calculating the global significance following Ref.[52], using the up-crossing method:

$$p_{\text{global}} = p_{\text{local}} + N_{\text{up}} e^{-1/2(Z^2_{\text{local}} - Z^2_{\text{ref}})}, \tag{2.86}$$

where the p_{local} and p_{global} are the local and global p-values, and Z_{ref} is chosen for p=0.5 (corresponding to 0 σ significance level), and N_{up} is the number of times the local p-value curve crossing the reference line (in this case, p=0.5) in the upwawrd direction. In this method, the p-value is degraded by adding an extra term N_{up} that accounts for the range of the search, and such term is penalised for large local significance.

Bibliography

[1] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[2] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett. B* 716 (2012), p. 30. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[3] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[4] E. D. Bloom et al. "High-Energy Inelastic $e - p$ Scattering at 6° and 10°". In: *Phys. Rev. Lett.* 23 (16 Oct. 1969), pp. 930–934. DOI: 10.1103/PhysRevLett.23.930. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.930.

[5] Martin Breidenbach et al. "OBSERVED BEHAVIOR OF HIGHLY INELASTIC ELECTRON-PROTON SCATTERING". In: *Physical Review Letters* 23 (1969), pp. 935–939.

[6] P. A. M. Dirac. "The Quantum Theory of the Electron". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117.778 (1928), pp. 610–624. ISSN: 09501207. URL: http://www.jstor.org/stable/94981.

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

[8] Murray Gell-Mann. "Symmetries of Baryons and Mesons". In: *Phys. Rev.* 125 (3 Feb. 1962), pp. 1067–1084. DOI: 10.1103/PhysRev.125.1067. URL: https://link.aps.org/doi/10.1103/PhysRev.125.1067.

[9] Makoto Kobayashi and Toshihide Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. URL: https://doi.org/10.1143/PTP.49.652.

[10] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(61)90469-2. URL: https://www.sciencedirect.com/science/article/pii/0029558261904692.

[11] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc. C* 680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034.

[12] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[13] Özer Özdal. "THE HIGGS BOSON AND RIGHT-HANDED NEUTRINOS IN SUPER-SYMMETRIC MODELS". PhD thesis. July 2016. DOI: 10.13140/RG.2.2.18314.52165.

[14] R.N. MOHAPATRA. "SEESAW MECHANISM AND ITS IMPLICATIONS". In: *SEESAW 25* (Apr. 2005). DOI: 10.1142/9789812702210_0003. URL: http://dx.doi.org/10.1142/9789812702210_0003.

[15] *Combined measurements of Higgs boson production and decay using up to* 139 fb^{-1} *of proton-proton collision data at* $\sqrt{s}$ = 13 *TeV collected with the ATLAS experiment*. Tech. rep. Geneva: CERN, Nov. 2021. URL: https://cds.cern.ch/record/2789544.

[16] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector". In: (2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[17] ATLAS. "LHC Higgs Cross Section Working Group DiHiggs subgroup webpage". 2021. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWGHH?redirectedfrom=LHCPhysics.LHCHXSWGHH#Latest_recommendations_for_gluon.

[18] ATLAS Collaboration. "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector". In: *JHEP* 04 (2015), p. 117. DOI: 10.1007/JHEP04(2015)117. arXiv: 1501.04943 [hep-ex].

[19] ATLAS and CMS Collaborations. "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV". In: *JHEP* 08 (2016), p. 045. DOI: 10.1007/JHEP08(2016)045. arXiv: 1606.02266 [hep-ex].

[20] ATLAS Collaboration. "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector". In: *Phys. Lett. B* 786 (2018), p. 59. DOI: 10.1016/j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex].

[21] CMS Collaboration. "Observation of Higgs Boson Decay to Bottom Quarks". In: *Phys. Rev. Lett.* 121 (2018), p. 121801. DOI: 10.1103/PhysRevLett.121.121801. arXiv: 1808.08242 [hep-ex].

[22] Massimiliano Grazzini et al. "Higgs boson pair production at NNLO with top quark mass effects". In: *JHEP* 05 (2018), p. 059. DOI: 10.1007/JHEP05(2018)059. arXiv: 1803.02463 [hep-ph].

[23] Frédéric A. Dreyer and Alexander Karlberg. "Vector-boson fusion Higgs pair production at N^{3}LO". In: *Phys. Rev. D* 98.11 (2018), p. 114016. DOI: 10.1103/PhysRevD.98.114016. arXiv: 1811.07906 [hep-ph].

[24] ATLAS Collaboration. "A search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: *Phys. Rev. Lett.* 121 (2018), p. 191801. DOI: 10.1103/PhysRevLett.121.191801. arXiv: 1808.00336 [hep-ex]. Erratum: in: *Phys. Rev. Lett.* 122 (2019), p. 089901. DOI: 10.1103/PhysRevLett.122.089901.

[25] ATLAS Collaboration. "Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: *Phys. Lett. B* 800 (2020), p. 135103. DOI: 10.1016/j.physletb.2019.135103. arXiv: 1906.02025 [hep-ex].

[26] P. A. R. Ade, N. Aghanim, et al. "Planck2013 results. I. Overview of products and scientific results". In: *Astronomy and Astrophysics* 571 (Oct. 2014). ISSN: 1432-0746. DOI: 10.1051/0004-6361/201321529. URL: http://dx.doi.org/10.1051/0004-6361/201321529.

[27] V. C. Rubin, Jr. Ford W. K., and N. Thonnard. "Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc)." In: 238 (June 1980), pp. 471–487. DOI: 10.1086/158003.

[28] Richard Massey, Thomas Kitching, and Johan Richard. "The dark matter of gravitational lensing". In: *Reports on Progress in Physics* 73.8 (July 2010), p. 086901. ISSN: 1361-6633. DOI: 10.1088/0034-4885/73/8/086901. URL: http://dx.doi.org/10.1088/0034-4885/73/8/086901.

[29] Supernova Search Team Collaboration. "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". In: *The Astronomical Journal* 116.3 (Sept. 1998). ISSN: 0004-6256. DOI: 10.1086/300499. URL: http://dx.doi.org/10.1086/300499.

[30] A. D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35. DOI: 10.1070/PU1991v034n05ABEH002497.

[31] K. et al. Eguchi. "First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance". In: *Physical Review Letters* 90.2 (Jan. 2003). ISSN: 1079-7114. DOI: 10.1103/physrevlett.90.021802. URL: http://dx.doi.org/10.1103/PhysRevLett.90.021802.

[32] Q. R. et al. Ahmad. "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory". In: *Physical Review Letters* 89.1 (June 2002). ISSN: 1079-7114. DOI: 10.1103/physrevlett.89.011301. URL: http://dx.doi.org/10.1103/PhysRevLett.89.011301.

[33] T. D. Lee. "A Theory of Spontaneous T Violation". In: *Phys. Rev. D* 8 (4 Aug. 1973), pp. 1226–1239. DOI: 10.1103/PhysRevD.8.1226. URL: https://link.aps.org/doi/10.1103/PhysRevD.8.1226.

[34] Sheldon L. Glashow and Steven Weinberg. "Natural conservation laws for neutral currents". In: *Phys. Rev. D* 15 (7 Apr. 1977), pp. 1958–1965. DOI: 10.1103/PhysRevD.15.1958. URL: https://link.aps.org/doi/10.1103/PhysRevD.15.1958.

[35] H.E. Haber and G.L. Kane. "The search for supersymmetry: Probing physics beyond the standard model". In: *Physics Reports* 117.2 (1985), pp. 75–263. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(85)90051-1. URL: https://www.sciencedirect.com/science/article/pii/0370157385900511.

[36] G.C. Branco et al. "Theory and phenomenology of two-Higgs-doublet models". In: *Physics Reports* 516.1-2 (July 2012), pp. 1–102. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2012.02.002. URL: http://dx.doi.org/10.1016/j.physrep.2012.02.002.

[37] Howard E. Haber and Yosef Nir. "Multi-scalar models with a high-energy scale". In: *Nuclear Physics B* 335.2 (1990), pp. 363–394. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(90)90499-4. URL: https://www.sciencedirect.com/science/article/pii/0550321390904994.

[38] H Georgi. "Effective Field Theory". In: *Annual Review of Nuclear and Particle Science* 43.1 (1993), pp. 209–252. DOI: 10.1146/annurev.ns.43.120193.001233. eprint: https://doi.org/10.1146/annurev.ns.43.120193.001233. URL: https://doi.org/10.1146/annurev.ns.43.120193.001233.

[39] B. Grzadkowski et al. "Dimension-six terms in the Standard Model Lagrangian". In: *Journal of High Energy Physics* 2010.10 (Oct. 2010). ISSN: 1029-8479. DOI: 10.1007/jhep10(2010)085. URL: http://dx.doi.org/10.1007/JHEP10(2010)085.

[40] R. Alonso et al. "The effective chiral Lagrangian for a light dynamical "Higgs particle"". In: *Physics Letters B* 722.4-5 (May 2013), pp. 330–335. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2013.04.037. URL: http://dx.doi.org/10.1016/j.physletb.2013.04.037.

[41] Gerhard Buchalla, Oscar Catà, and Claudius Krause. "Complete electroweak chiral Lagrangian with a light Higgs at NLO". In: *Nuclear Physics B* 880 (Mar. 2014), pp. 552–573. ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2014.01.018. URL: http://dx.doi.org/10.1016/j.nuclphysb.2014.01.018.

[42] Alexandra Carvalho et al. "Higgs pair production: choosing benchmarks with cluster analysis". In: *Journal of High Energy Physics* 2016.4 (Apr. 2016), pp. 1–28. ISSN: 1029-8479. DOI: 10.1007/jhep04(2016)126. URL: http://dx.doi.org/10.1007/JHEP04(2016)126.

[43] M. Capozi and G. Heinrich. "Exploring anomalous couplings in Higgs boson pair production through shape analysis". In: *Journal of High Energy Physics* 2020.3 (Mar. 2020). ISSN: 1029-8479. DOI: 10.1007/jhep03(2020)091. URL: http://dx.doi.org/10.1007/JHEP03(2020)091.

[44] Piero Altoe Kim Albertsson et al. *Machine Learning in High Energy Physics Community White Paper*. 2019. arXiv: 1807.02876 [physics.comp-ph].

[45] Dimitri Bourilkov. "Machine and deep learning applications in particle physics". In: *International Journal of Modern Physics A* 34.35 (Dec. 2019), p. 1930019. ISSN: 1793-656X. DOI: 10.1142/s0217751x19300199. URL: http://dx.doi.org/10.1142/S0217751X19300199.

[46] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016.

[47] Eilam Gross. "LHC Statistics for Pedestrians". In: (2008). DOI: 10.5170/CERN-2008-001.205. URL: https://cds.cern.ch/record/1099994.

[48] A L Read. "Presentation of search results: the CL_s technique". In: *J. Phys. G* 28 (2002), pp. 2693–704. DOI: 10.1088/0954-3899/28/10/313. URL: https://cds.cern.ch/record/722145.

[49] J. Neyman and E. S. Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses". In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 231 (1933), pp. 289–337. URL: http://www.jstor.org/stable/91247.

[50] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[51] Eilam Gross and Ofer Vitells. "Trial factors for the look elsewhere effect in high energy physics". In: *The European Physical Journal C* 70.1-2 (Oct. 2010), pp. 525–530. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1470-8. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1470-8.

[52] Eilam Gross and Ofer Vitells. "Trial factors for the look elsewhere effect in high energy physics". In: *The European Physical Journal C* 70.1-2 (Oct. 2010), pp. 525–530. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1470-8. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1470-8.