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

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Zhiyuan Li

May 30, 2022

Contents

Abstract

A search for Higgs boson pair (di-Higgs) production in events with two b-jets and two τ-leptons ($b\bar{b}\tau^+\tau^-$) is presented, using a proton–proton collision data set with an integrated luminosity of 139 fb^{-1} collected at $\sqrt{s} = 13$ TeV by the ATLAS experiment at the LHC. Di-Higgs produced non-resonantly, or in the decay of a generic narrow-width scalar resonance in the mass range 251 to 1600 GeV is targeted. Events in which one τ-lepton decays leptonically and the other decays hadronically (the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ decay channel) are considered. Multivariate discriminants are used to extract the signals. No significant excess of events above the expected background is observed in the non-resonant search, therefore upper limits are set on the non-resonant di-Higgs production cross-section. Upper limits are also set on the resonant di-Higgs production cross-section. The search results are then reinterpreted in the framework of Higgs Effective Fields Theory (HEFT), where the non-resonant di-Higgs production is modified by assuming seven HEFT benchmark models. One-dimensional scans are performed on the Higgs boson self-coupling modifier κ_λ and two HEFT coupling parameters associated with the couplings of a Higgs boson pair with two gluons, c_{gghh}, and with a top quark pair, c_{tthh}.

The $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ results are statistically combined with the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ results, where both τ-leptons decay hadronically, and with the results of the $b\bar{b}b\bar{b}$ and $b\bar{b}\gamma\gamma$ di-Higgs decay channels. Observed (expected) upper limits are set at the 95% confidence-level on the non-resonant di-Higgs production cross-section of 3.1 (3.1) times the Standard Model prediction, and on the resonance production cross-section between 1.1 and 595 fb (1.2 and 392 fb), depending on the heavy resonance mass. The largest excess in the resonant search is observed at a resonance mass of 1 TeV, with a local (global) significance of 3.1 σ ($2.1^{+0.4}_{-0.2}\,\sigma$). The κ_λ, c_{gghh} and c_{tthh} are constrained to $-1.0 < \kappa_\lambda < 6.6$, $-0.3 < c_{gghh} < 0.4$ and $-0.2 < c_{tthh} < 0.6$, respectively.

Declaration

I hereby confirm this work is my own, except where other works are referenced. This work has not previously been submitted to any institute, including this one. This thesis does not exceed the relevant word count.

Zhiyuan Li

Acknowledgements

I would like to express my biggest gratitude to my supervisors, Carl Gwilliam, Nikolaos Romopotis and Andrew Mehta. It is my honour to have three great supervisors at the same time. I am extremely grateful to Carl and would like to thank him for being endlessly patient with me. Regardless of the number and level of my questions, you always takes them the most seriously and find the best comprehensive way to answer them. Your optimistic attitude in research has inspired me to sail upright to where I am now. I would like to thank Nikos for his great help along the way; his rigour shows me the proper altitude for conducting research, and I will try to chase up your standard in my future academic career. Thank you to Andy, you have inspired me with countless ideas in the calibration project.

I would also express gratitude to the rest of the Liverpool ALTAS team, which has been extremely supportive, even during the pandemic. Also, I want to say a big thank you to the ATLAS $b\bar{b}\tau^+\tau^-$ group, especially to Allison McCarn, that you have been helping me throughout my entire research project, not to mention writing a great reference for me.

Finally, I would like to thank my parents for their support, and my beloved partner Zhiming. You have supported me since I first came to the UK with your unreserved love, and given me all the courage that I need to pursue my dream.

Chapter 1

Introduction

It has been a decade since the discovery of the Higgs boson with a mass of about 125 GeV at the LHC, observed by the ATLAS and CMS experiments [1, 2]. It was the last fundamental particle predicted by the the Standard Model of Particle Physics (SM), and so far, all measured properties of this boson are consistent with those predicted by the SM. The SM is therefore hailed by many as the most successful quantum field theory ever constructed.

Despite its success, the SM still cannot answer many important questions. For example, the SM does not explain the existence of Dark Matter (DM) or Dark Energy, and it leads to some theoretical issues such as the hierarchy problem. Therefore the SM is widely considered as an incomplete model. Physicists inspired by this knowledge have put in a massive effort to develop new theories Beyond the Standard Model (BSM).

One of the most pressing tasks of the LHC since the discovery of the Higgs boson is to measure its couplings precisely, in particular the Higgs boson self-coupling. The precise measurement is vital to test the Higgs mechanism [3, 4], the electroweak theory [5] and the SM theory itself. Particularly, di-Higgs production provides direct access to the self-coupling constant. The cross-section of this process (via ggF production) as predicted by the SM is only 31.05 fb [6], three orders of magnitude smaller than the cross-section for single Higgs-boson production. Therefore, the observation of the di-Higgs process is extremely challenging, and it is one of the main goals for the High Luminosity LHC (HL-LHC).

Nevertheless, some BSM theories modify the Higgs boson self-coupling strength, or provide anomalous couplings to the Higgs boson, which can enhance the non-resonant di-Higgs production cross-section. Moreover, some BSM theories predict new heavy resonance particles that can decay to a pair of Higgs bosons. The presence of one of these resonances would also alter the di-Higgs production rate, making the observation of the di-Higgs process possible with the current dataset.

In this thesis, a search for di-Higgs production using the $b\bar{b}\tau^+\tau^-$ final state is presented. The search utilises a proton–proton collision data set with an integrated luminosity of 139 fb^{-1} collected at $\sqrt{s}$ = 13 TeV by the ATLAS experiment at the LHC. The search results are presented as upper limits on the di-Higgs production cross-section. The search results are then reinterpreted using the HEFT framework, where upper limits are set on the cross-section for non-resonant ggF di-Higgs production assuming 7 benchmark models, and one-dimensional scans are performed on κ_λ, c_{gghh}, and c_{tthh}.

This thesis is organised as follows: Chapter 2 outlines the theoretical background and motivation for searches presented in Chapter 7. An introduction of this chapter is given in Section 2.1. Chapter 3 describes the LHC and the ATLAS detector. The methods used for reconstructing the physics objects are detailed in Chapter 4. Chapter 5 details the signal and background simulation samples used for the calibration presented in Chapter 6 and the di-Higgs searches presented in Chapter 7. In Chapter 6, a calibration is performed by the author on the measurement of the c-jet mis-tagging efficiency of the b-tagging algorithm, using $t\bar{t}$ events. The calibration result has become the official recommendation for all ATLAS users who apply b-tagging in their analyses. Chapter 7 details the search for di-Higgs production, where a brief introduction is given in Section 7.1. Finally, this thesis is summarised in Chapter 8.

Chapter 2

Theory and motivation

2.1 Introduction

Particle physics is at 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 and represents one of the triumphs of modern physics.

The last fundamental particle predicted by the SM, the Higgs boson, has been observed by the ATLAS and CMS experiments at the Large Hadron Collider [1, 2]. The SM is therefore hailed by many as the most successful quantum field theory ever constructed. 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) [7] that have to be input by hand; what is the particle content of dark matter; what is the origin of the matter-antimatter asymmetry in the universe.

This chapter is structured as follows. Section 2.2 introduces the basic concepts of the SM, including the gauge theory and fundamental forces and the Higgs mechanism. Section 2.3 gives a hint of beyond the SM (BSM) theories that address some of the important but unanswered questions. Section 2.4 and Section 2.5 outline the machine learning algorithms and the statistical methods used in this thesis, respectively.

2.2 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 [5]. 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 finalised in 1969 upon experimental confirmation of the existence of quarks [8, 9].

In the SM, most of the everyday phenomena we see in the physical world are 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 a positively charged nucleus, bound by electromagnetic attraction. Such a 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 often 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 [7]), gravity must be integrated into the theory. To date, no clear approach is available for combining the two behemoths. A huge effort has been made, and there is large progress, but there are still unanswered questions both theoretically and experimentally.

2.2.1 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-energy 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 $m_\tau \approx 3500\ m_e$ [10]. The three generations of particles are collectively called *fermions*. Fermions have a half-integer spin 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 [11],

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

where the γ^μ are the Dirac 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 and spin, but opposite charge. 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 spin, and do not obey the Pauli exclusion principle. In particular, the fundamental SM 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 seen as the excitations of the quantum field of each force. For example, the familiar photon is the gauge boson of the QED, and for the strong interaction, the force-carrying particle 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 (electrically) 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 intrinsically 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 [7], due to *colour confinement* (which will be discussed in more detail in Section 2.2.3). 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 quarks and gluons themselves.

The final element of the SM is its scalar sector. 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 important role in the SM by providing mass for all known particles (if neutrinos also acquire their masses from the Higgs boson): without it, all fundamental 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 2.2.5.

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.

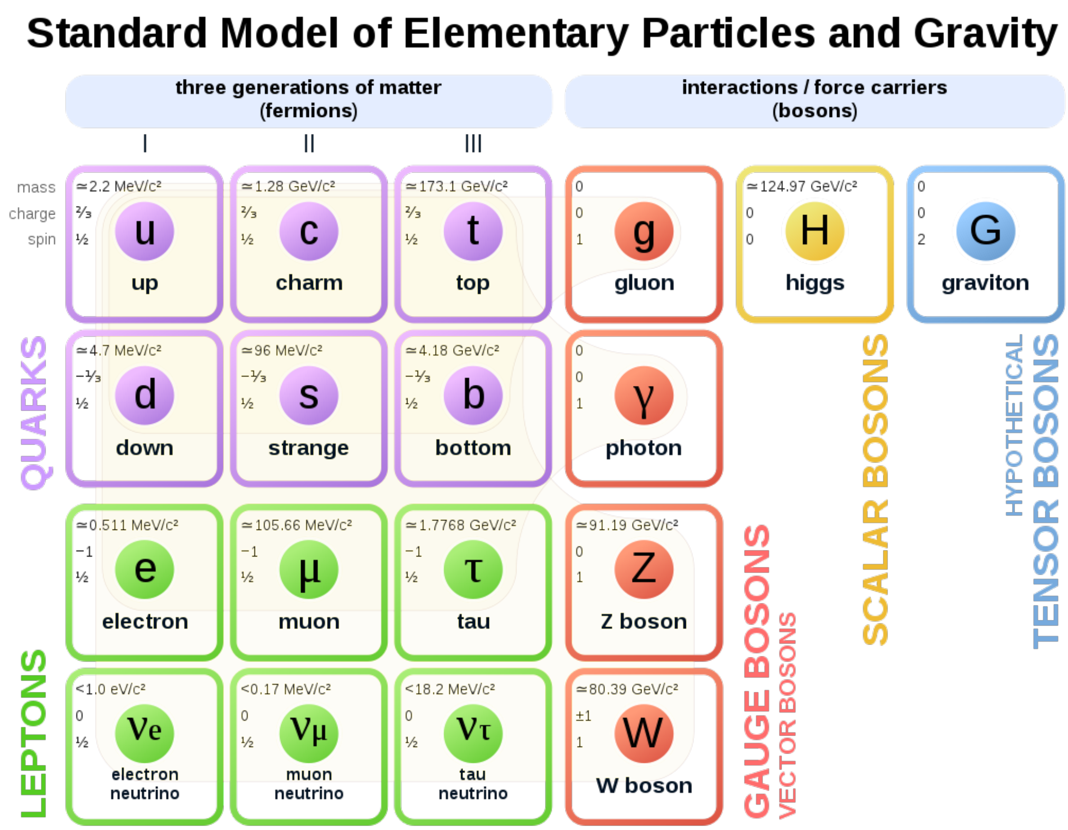


Figure 2.1: SM of elementary particles: the 12 fundamental fermions and 5 fundamental bosons (along with the hypothetical Graviton). The mass, charge and spin of each particle are given inside the particle boxes [10]. Image credit to MissMJ, Cush.

2.2.2 Symmetries in the Standard Model

Symmetry plays a crucial role in modern physics, particularly in the SM. The 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 colour 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 chirality particles, while the letter Y stands for the weak hypercharge, which is the charge associated with the $U(1)_Y$ group like the electric charge, which is the charge associated with the $U(1)_{EM}$ group. The weak hypercharge is related to the electric charge (Q) and the weak isospin (I_W) as Q

$= I_W^3 + Y/2$, with I_W^3 being the third component of the weak isospin and is conserved due to the $\mathrm{SU}(2)_L$ symmetry.

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

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

where $\mathcal{L}_{\mathrm{QCD}}$ describes the dynamics of the strong force, $\mathcal{L}_{\mathrm{Electroweak}}$ describes the dynamics of the electroweak force and $\mathcal{L}_{\mathrm{Higgs}}$ and $\mathcal{L}_{\mathrm{Yukawa}}$ are the terms which introduce mass to gauge bosons and fermions, 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 $\delta\psi$:

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

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.

2.2.3 Quantum Chromodynamics

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

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

where ψ is the electric potential, **A** is the vector potential and χ is any twice continuously differentiable function that depends on position and time. 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 $\mathrm{U}(1)$ gauge 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 $\mathrm{U}(1)$ 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 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 QCD, which obeys the $\mathrm{SU}(3)$ group. Suppose an $\mathrm{SU}(3)$ gauge transformation is applied on the wave function, i.e.

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

where g_S is some coupling constant, the $\hat{\mathbf{T}} = \{T^a\}$ are the eight generators of the $\mathrm{SU}(3)$, which are related to the *Gell-Mann matrices* by $T^a = \frac{1}{2}\lambda^a$ [12], and $\boldsymbol{\alpha}(x)$ are eight functions of the space-time coordinate x, corresponding to each of the eight $\mathrm{SU}(3)$ generators. Representing the $\mathrm{SU}(3)$ group by 3×3 matrices, the additional degrees of freedom are accounted by a vector of three components, namely red, green and blue. Hence, the idea of colour charge comes naturally from requiring the gauge invariance. Finally, the concept of the gluon also comes out when requiring the gauge invariance, which is the quanta of the eight 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:

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

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 nature, and the reason might well possibly be colour confinement. The qualitative explanation 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 a force of the order 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 quark pairs can become new bound states with the quarks being pulled apart. This process goes on if the initial quarks are pulled further apart, as shown in Figure 2.2.

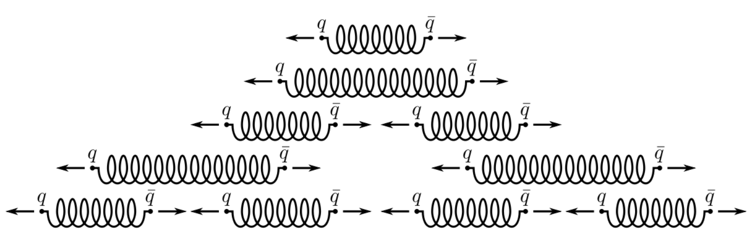


Figure 2.2: An illustration of a pair of quarks being pulled away: new pairs of quarks are created and become new bound states with the two quarks being pulled.

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 referred 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 [7]. In modern particle detectors, the α_S value is sufficiently small for pertubation theory to be used. Finally, the Lagrangian of QCD describing the interactions of the quarks via gluons and the gluon self-interactions 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:

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

2.2.4 Electroweak theory

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

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

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 $\mathbf{T} = \frac{1}{2}\boldsymbol{\sigma}$, and $\boldsymbol{\alpha}(x)$ are three functions which specify the local phase at each point in space-time. To satisfy 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 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 are a mixture of the latter using the Cabibbo-Kobayashi-Maskawa (CKM) matrix [13]. The RH chiral particles and LH antiparticles are represented by 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 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 fields $W^{(1)}$ 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, experimentally 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 [14–16], the U(1) gauge symmetry of electromagnetism is represented by a new $\mathrm{U}(1)_Y$ 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 to g will become clear in the following), $\zeta(x)$ being a function in x. Requiring 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 2.2.3). 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 $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge 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.

$$Y = \alpha Q + \beta I_W^3. \tag{2.28}$$

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) gauge 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. 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 [10]

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

Finally, the Lagrangian of the electroweak theory is:

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

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.

2.2.5 The Higgs mechanism

The gauge invariance is preserved in $\mathrm{SU}(2)_L$ 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 gauge transformation. This simple example can be applied to the $\mathrm{SU}(2)_L$, and to solve the conflict that experimental observations show the weak bosons are massive, while the gauge invariance requires the weak bosons to be massless, the Higgs mechanism is proposed.

Now consider a complex scalar field, ϕ:

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

where ϕ_1 and ϕ_2 are real scalar fields, 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 a set of 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

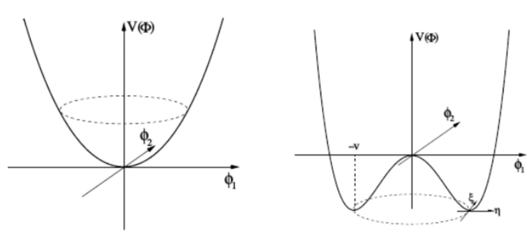


Figure 2.3: The potential $V(\phi)$ for a complex scalar field for $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right). Image taken from [17].

state corresponds to a particular point on the circle, where the $\mathrm{SU}(2) \times \mathrm{U}(1)$ 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 the 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 in electroweak symmetry 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 perturbations in the gauge invariant Lagrangian and derive the kinematic terms of the massive field η and massless ζ, and the massive gauge field (which was massless originally). In this process, the massless field 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.

In the SM, the Higgs mechanism is embedded in the U(1) and 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 fields 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. Writing 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 [14–16].

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

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

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 the 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 quadrilinear). The Lagrangian does not depend on A_μ, and therefore, the 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 [10].

2.2.6 Yukawa coupling

The last missing piece of the mass mystery is the origin 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: $\bar{L}\phi$, where L are left-handed chiral fermions placed in SU(2) doublets (with R being right-handed chiral fermions placed in SU(2) singlets), which is invariant under SU(2). This term is invariant because ϕ transforms as: $\phi \to \phi' = (I + ig_W\boldsymbol{\epsilon}(x)\cdot\mathbf{T})\phi$ and $\bar{L} \equiv L^\dagger\gamma^0$ transforms as: $\bar{L} \to \bar{L}' = L(I - ig_W\boldsymbol{\epsilon}(x)\cdot\mathbf{T})$. When combined with the RH singlet, 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 after symmetry breaking:

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

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 Higgs field, and the second term corresponding to the interaction 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 the complex scalar field, which means in this process, only the charged leptons and the down-type quarks obtain masses. What about the up-type quarks and the neutrinos? Ignoring the neutrinos for now, the up-type quarks can acquire mass 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 quarks can also acquire mass. In conclusion, the Yukawa couplings of the fermions to the Higgs field are given by:

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

Interestingly, for the top quark with mass $\sim$ 173.5 GeV [10], the coupling strength of the top quark to the Higgs field is very close to unity. While the neutrinos have such a small mass that they are often considered as massless, the Yukawa coupling will be unnaturally small, suggesting that they might be acquiring their masses in a different way. A possibility is the *seesaw mechanism* [18], but it is outside the scope of this chapter.

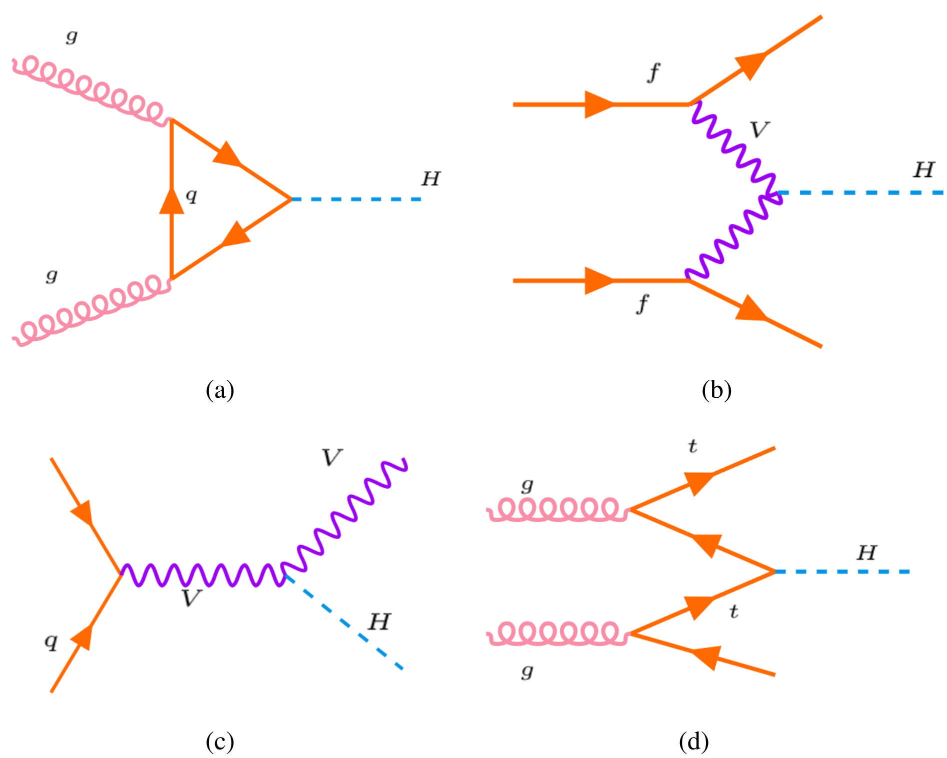


Figure 2.4: The lowest order Feynman diagrams of the four production mechanisms: (a) ggF, (b) VBF, (c) VH and (d) ttH.

2.2.7 Higgs boson production at the LHC

In the proton-proton collisions, Higgs bosons are produced via four main mechanisms: gluon-gluon fusion (ggF), vector boson fusion (VBF), production associated with a vector boson (VH) and production associated with a top-anti-top-quark 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.

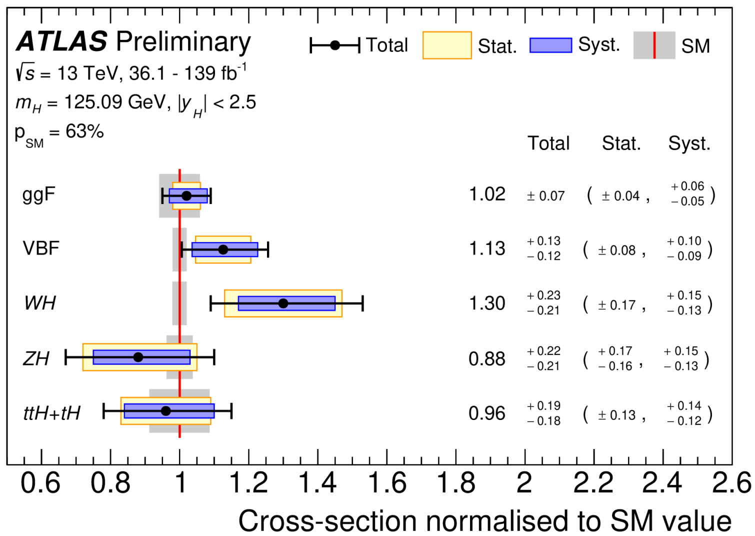


Figure 2.5: Cross sections for ggF, VBF, WH, ZH and ttH+tH production modes. The cross sections are normalised to their SM predictions, measured assuming SM values for the decay branching fractions. The black error bars, blue and yellow boxes show the total, systematic, and statistical uncertainties in the measurements, respectively. The grey 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* (more details in Section 2.5) of 63%. Image taken from Ref. [19].

The dominant production mode is the ggF, an order of magnitude greater than the next largest production mode VBF. The production cross-sections are shown in Figure 2.6 as a function of the centre-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 quarks (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 [20]. The main branching ratios of the Higgs boson are listed in Table 2.1.

Decay mode	Branching ratio
$H \to b\bar{b}$	58.2%
$H \to WW^*$	21.4%
$H \to gg$	8.19%
$H \to \tau^+\tau^-$	6.27%
$H \to c\bar{c}$	2.89%
$H \to ZZ^*$	2.62%
$H \to \gamma\gamma$	0.227%
$H \to Z\gamma$	0.153%
$H \to \mu^+\mu^-$	0.0218%

Table 2.1: The SM predicted branching ratios in descending order of the Higgs boson for m_H = 125 GeV. Values taken from Ref. [20].

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 [7]. 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-jet final states. Hence it is difficult to distinguish the decays of the Higgs boson from the large background from multi-jet production in proton-proton collisions. For this reason, physicists focused on the more distinctive 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].

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 [22, 23]. While the interaction with bottom quarks was observed later, when the Higgs bosons were observed to decay to two bottom quarks in 2018 by ATLAS and CMS [24, 25]. 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 multi-jet background 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_F and κ_V 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

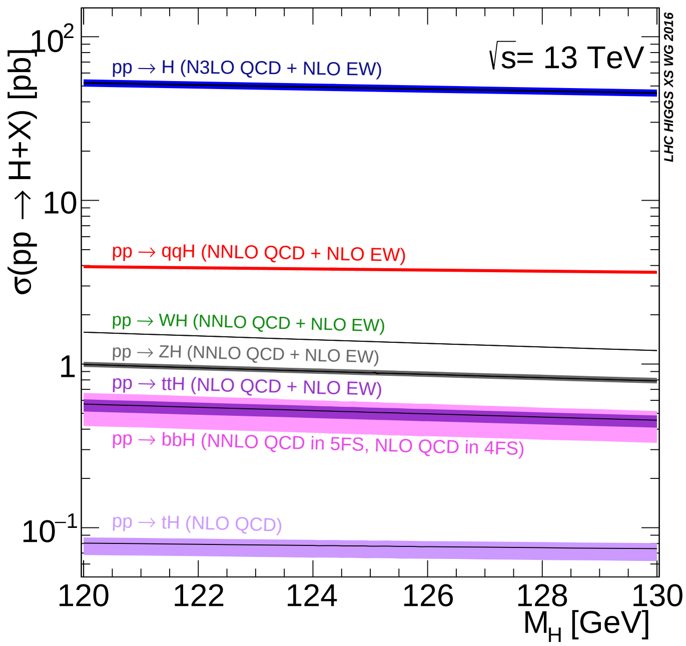


Figure 2.6: The production cross-sections of the SM Higgs boson at $\sqrt{s}$ = 13 TeV. The $pp \to H$ corresponds to the ggF production and the $pp \to qqH$ corresponds to the VBF production. The $pp \to WH, ZH, ttH$ corresponds to the VH, ttH. Image taken from Ref. [21].

Higgs field v = 246 GeV, where a straight line is drawn across the different particles.

2.2.8 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. It also has crucial implications in electroweak symmetry breaking mechanism, as discussed earlier in Section 2.2.5. To measure it, a direct probe would be to measure Higgs boson pair production, which is the main goal of this thesis.

The di-Higgs production process discussed in this 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 2.3.

As defined in Section 2.2.5, the Higgs potential is given by:

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

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

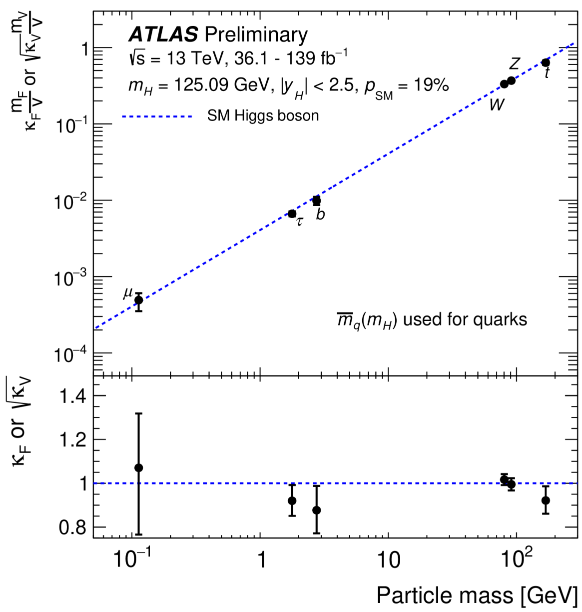
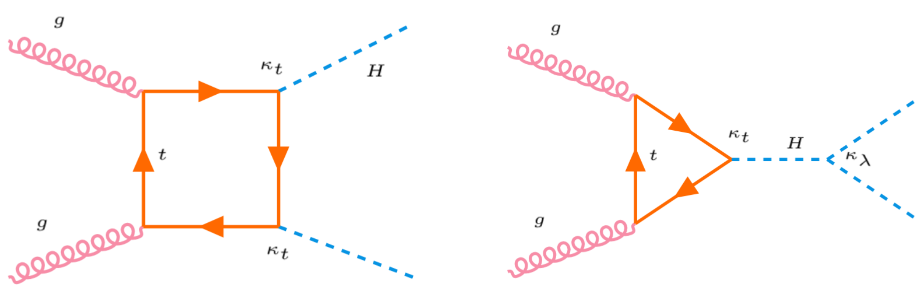


Figure 2.7: The reduced coupling-strength modifiers $\kappa_F \frac{m_F}{v}$ and $\sqrt{\kappa_V} \frac{m_V}{v}$ as a function of their masses m_F and m_V, for a vacuum expectation value $v = 246$ GeV. The SM prediction for both cases is also shown (dashed line). The black error bars represent 68% CL intervals for the measured parameters. The lower panel shows the ratios of the values to their SM predictions. Image taken from Ref. [19].

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 access 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 $\kappa_t = g_t/g_t^{\text{SM}}$ and $\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{\text{SM}}$, where g_t and λ_{HHH} are the measured values and g_t^{SM} and $\lambda_{HHH}^{\text{SM}}$ are the values predicted by the SM.



(a) (b)

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

The pair production of the Higgs boson occurs at a minimal rate due to the small phase space of decaying to two on-shell Higgs. Moreover, these two diagrams interfere destructively, making the 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) [6] is given by:

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

at $\sqrt{s} = 13$ TeV and $\text{m}_H = 125$ GeV [21]. The scale uncertainty is due to the finite order 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 process 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-

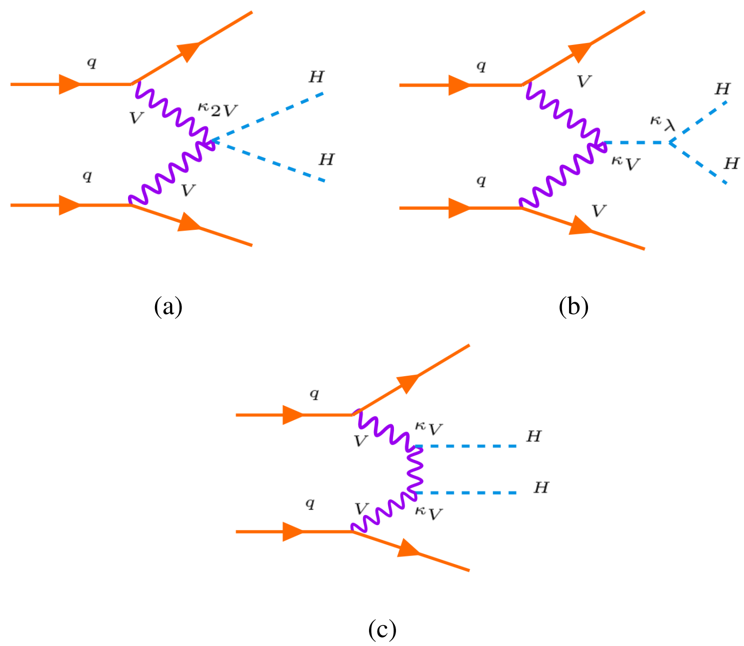


Figure 2.9: Leading order Feynman diagrams for VBF HH production.

section is calculated at next-to-next-to-next-to-leading order (N3LO) in QCD, which is given by [26]:

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

at $\sqrt{s} = 13$ TeV and $m_H = 125$ GeV [21].

2.2.8.1 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 7, a search for di-Higgs production in the $b\bar{b}\tau^+\tau^-$ channel is presented. The results of the search are also combined with the $b\bar{b}b\bar{b}$ channel and the $b\bar{b}\gamma\gamma$ channel to maximise the sensitivity. Even though the $b\bar{b}b\bar{b}$ channel has the largest branching ratio, it has been shown to be very difficult to extract the signal from the vast multi-jet background, which can be seen from the late observation of the $H \to b\bar{b}$ decay [24, 25]. 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 channels. It has a relatively high branching ratio and relatively clean background, making the $b\bar{b}\tau^+\tau^-$ channel one of the most sensitive channels.

	bb	WW	ττ	ZZ	γγ
bb	33%				
WW	25%	4.6%			
ττ	7.3%	2.5%	0.39%		
ZZ	3.1%	1.2%	0.34%	0.076%	
γγ	0.26%	0.10%	0.029%	0.013%	0.0053%

Figure 2.10: The most common di-Higgs decay channels and the corresponding branching ratios. Image credit: K. Leney.

The two τ-leptons in the final state subsequently decay either leptonically or hadronically, as described in Section 4.5. Processes with final state with a leptonically decaying τ and a hadronically decaying τ, which account for 42.0% of the di-τ decays, are categorised as the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. Processes with both τ decaying hadronically are categorised as the $\tau_{\text{had}}\tau_{\text{had}}$ channel, which account for 45.6% [20].

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 [27] 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 was constrained to be less than 30.9 fb, 13 times the Standard Model expectation, at 95% confidence level. When 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 was tightened to 6.9 times the SM expectation [28]. With the full Run 2 data and various improvements, the author will present readers the exciting results in chapter 7.

In the search for di-Higgs production presented in chapter 7, 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.

2.3 Beyond the Standard Model

Despite the huge success of the SM, there are many unanswered questions in the SM. In order to include the phenomena not explained in the SM and to solve some of the theoretical issues, many theories beyond the SM have been developed. In this section, the limitations of the SM theory are summarised and a few BSM models, considered in this thesis work on the $b\bar{b}\tau^+\tau^-$ final state, are introduced.

2.3.1 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? Measurements of the cosmic microwave background radiation show that the directly observable SM matter is only 5% of the energy of the universe [29]. Measurements of galaxy rotation curves [30] and gravitational lensing [31] show that dark matter makes up about 27% of the universe, while measurements of the universe expansion rate [32] show that dark energy makes up about 68%.

- Where does the matter-antimatter asymmetry 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 an 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 [33] conditions requires the *CP-symmetry* (the combination of the charge symmetry and the parity symmetry) to be violated. 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 asymmetry.

- What is the origin of neutrino masses? In the SM, there are no RH neutrinos, and there are only Higgs doublets of $\mathrm{SU}(2)_L$. Therefore neutrinos are required to be massless. However, observations in neutrino oscillations [34, 35] imply that the neutrinos have a non-zero mass. If neutrinos are normal Dirac particles, this will imply an unnaturally small Yukawa coupling to the Higgs field. Another 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 2.1, 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 the Planck scale ($\Lambda_P \sim 10^{19}$ GeV) this correction becomes very large (if the SM is still valid). This significant correction needs 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.

2.3.2 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 relevant 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 [36]. 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 [37].

Nowadays, there are many motivations for 2HDMs, the best known of which might be supersymmetry [38]. 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 hypercharge +1 can be written as [39]:

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

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, two charged scalar particles $H^\pm$, and a CP-odd neutral pseudoscalar A.

Like in the single complex doublet case, the mass terms arise from the square of the field and the mass terms of the neutral scalars h and H 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 $M(v_1, v_2, \lambda_{1,2,3,4})_\psi, m_{12}$ is the matrix encapsulating the coefficients of the scalar mass terms, which can be diagonalised by a rotation angle α [39]. Similarly for the pseudoscalar A 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 β is defined as $\tan\beta \equiv \frac{v_2}{v_1}$ [39]. These two parameters α and β determine the interactions of the different Higgs fields 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 'weak decoupling limit' (or 'alignment limit') [40], the lighter neutral scalar h presents almost the same properties as the SM Higgs boson [41]. The heavier neutral scalar H, on the other hand, could be generated at the LHC, which subsequently decays to two Higgs bosons, as shown in the Feynman diagram at leading order in Figure 2.11. Therefore, the resonant di-Higgs production is of interest to this thesis.

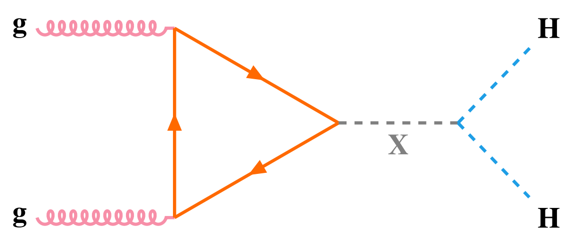


Figure 2.11: SM leading-order Feynman diagrams for Higgs boson pair production through resonance of a generic spin-0 scalar particle.

In chapter 7, 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 a nominal cross-section of 1 pb is chosen for the resonant production.

2.3.3 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 equivalent to saying that the energy scale being considered is low. The SM is commonly accepted 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:

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

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 $O_k^{(5)}$, dimension-six $O_k^{(6)}$ and of higher dimensions (suppressed by higher orders of Λ) [42]. The $c_k^{(n)}$ are the dimensionless coupling constants (*Wilson coefficients*).

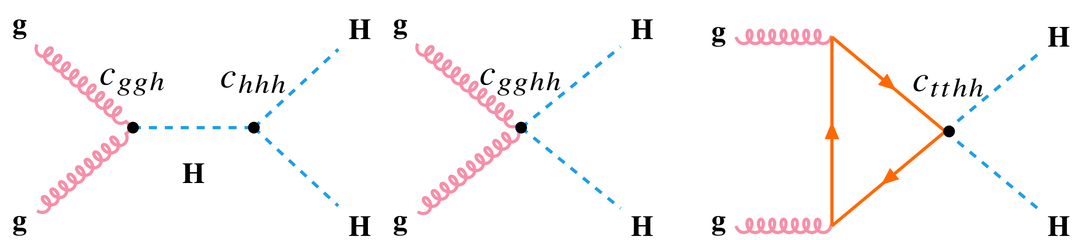
It has been proven in Ref. [42] that the dimension-five operator (interestingly, there is only one such operator) violates the lepton number conservation. At the LHC, this effect is almost unobservable, therefore it is not considered. On the other hand, it is possible to write down a large number of dimension-6 operators, for example, just attaching an extra term $\phi^\dagger \phi$ to any of the terms in the SM, the Lagrangian will be dimension-six. It was shown in Ref. [43] 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 to 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', assuming the theories considered are irrelevant at higher energy. The 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) [43] and Higgs EFT (HEFT) [44, 45]. 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

$SU(3)_C \times U(1)_{em}$, while the $SU(2)_L \times U(1)_Y$ 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 as shown in Figure 2.8, i.e. $c_{hhh} \equiv \kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{\text{SM}}$ and $c_{tth} \equiv \kappa_t = \lambda_{ttH}/\lambda_{ttH}^{\text{SM}}$, 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 account for the non-SM interactions, affecting the $ttHH$, gHH and $ggHH$ vertices respectively. 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 process gives unique access to the coefficients c_{hhh}, c_{tthh} and c_{gghh} (the c_{ggh} and c_{tth} coefficients can also be probed by single-Higgs boson production, with higher sensitivity). The HEFT formalism allows one to interpret general searches in different BSM scenarios by simultaneously varying multiple Wilson coefficients. In Section 7.6, results of one dimensional scans on coefficients c_{hhh}, c_{tthh} and c_{gghh} (assuming that all other couplings take their SM values) are presented. The results also represent the first dedicated scan on the c_{tthh} and c_{gghh} coefficients in ATLAS.

In addition, *cluster analysis* [46] 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) [47], 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. Upper limits are set on the di-Higgs production cross-section for each of these 7 benchmarks. The m_{HH} distributions of these 7 BM are shown together with the m_{HH} distribution of the SM non-resonant ggF HH production in Figure 2.13.

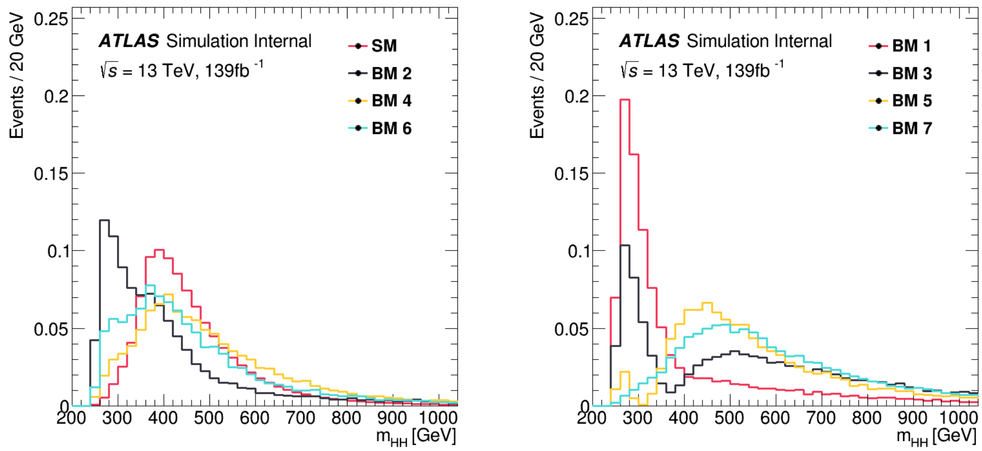


Figure 2.13: The m_{HH} distributions for the SM HH production and HEFT shape benchmarks. All distributions are normalised to unity at truth level (derived directly from events without detector response simulation). Image taken from Ref. [48].

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. [47].

2.4 Machine learning theory

Machine learning (ML) algorithms are a generic and task-independent method which, rather than being explicitly programmed, learn from the dataset relevant to the task. 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 [49], and they help reduce the execution time of computationally-expensive event simulation and calibration [50].

ML algorithms are commonly used for two types of problems: *classification* and *regression* [51]. In a classification problem, variables relevant to the physics problem are selected and an ML model is trained by separating the data into classes, e.g. signal and background events. 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 [52]. For classification problems, each leaf represents the model's decision to assign a data item to a class. 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 that of artificial neural networks (ANN/NN 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 true shape of the event better.

Neural networks were inspired by the biological brains, 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 to a given node in the neural network 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 2.14.

More specifically, layers in an 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 the $n+1^{\text{th}}$ 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 input to the node, x_i^n is the output of the i^{th}

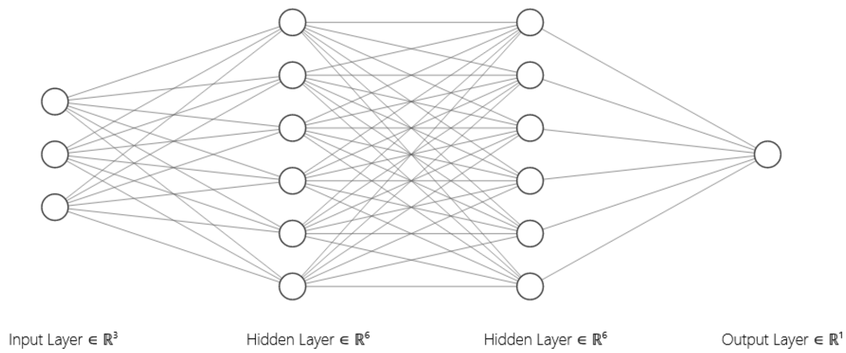


Figure 2.14: A simple neural network with 3 input nodes, two hidden layers each with 6 nodes, and an output node.

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 the range $[-1, 1]$; or the rectified linear unit (ReLU), given by $x^{n+1} = max(0, y^{n+1})$, that has output in the range $[0, 1]$. The sigmoid function and the ReLU are shown in Figure 2.15.

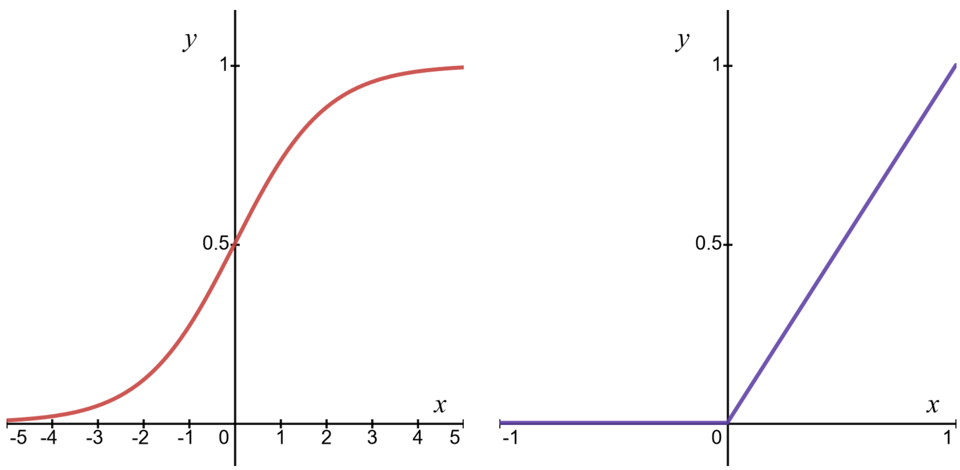


Figure 2.15: The sigmoid (left) function and the rectified linear unit (right).

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, such as the binary cross entropy which is defined as $L(y, \hat{y}) = \frac{1}{m} \sum_{i=1} (y_i - \hat{y}_i)^2$, 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 for the task, an iterative process through which the loss is minimised, known as 'training', is employed. The training data for the NN used in this thesis consists of pairs of labelled 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 for multiple iterations (*epochs*) until the loss function reaches its minimum. The size of each step is controlled 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 will be very long, since the loss function is a function of all data set. In this thesis, the *stochastic gradient descent* [53] 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.

A 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 adjusts the weights and bias of one layer according to the expected output of the next layer. For example, if the network's final output 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 *hyperparameters*. 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 choice of hyperparameters is non-trivial, and is generally chosen as the model which provides the best performance. The optimisation of the hyperparameters of the NN used in this thesis is outlined in Section 7.4.2.

An NN is used for the search for the non-resonant HH production. The resonant signal, however, is not a single signal hypothesis, but rather a set of continuous signal hypotheses parameterised by the mass of the heavy resonance, which decays to Higgs boson pairs. As a result, a single classifier would not be enough; while using a set of single classifiers trained on each of the simulated resonant signal points would provide good discrimination for each of them, it can not be used to interpolate well between these points. For this reason, and to reduce the number of algorithms that require training, Parametric Neural Networks (PNNs) [54] are used to extract the resonant signal. PNNs are neural networks connected to one or more physics parameters which enable the optimal signal-background classification for a continuous spectrum of the signal.

2.5 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 Standard Model, while the second question has to do with hypothesis testing and the derivation of confidence intervals and upper limits [55].

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 on the specific physics problem. For a search for a new signal process, the null hypothesis H_0 assumes only the background is observed, while the alternative hypothesis H_1 assumes one observes both signal and background; for problems such as setting upper limits, this definition is reversed.

Once the null hypothesis and the alternative hypothesis are defined and data is measured, one can test the hypothesis with a *test statistic*. It is a quantity calculated from data, which can be used to estimate how probable is the result that one observes with respect to the null hypothesis.

The analysis discussed in this thesis consists of signal and background, distributed in a histograms $\mathbf{n}$ of N bins: $\mathbf{n} = (n_1, n_2, ..., n_N)$. In such a case, the expectation of each bin consists of the background b and the signal s with some signal strength μ:

$$E[n_i] = \mu s_i + b_i. \tag{2.70}$$

Since in a counting experiment, the data follows a Poisson distribution, the *likelihood function* of the signal strength is defined as [56]:

$$L(\mu) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)}. \tag{2.71}$$

In addition to the signal strength, it is common that the signal and background will depend on some additional parameters, referred to as *nuisance parameters*, which are not of direct interest of the analysis but need to be fitted from the data. They are used to parametrise the effect of each systematic uncertainty on the expected number of signal and background events in each bin, and are usually modelled by Gaussian functions with width representing the size of the uncertainty. Nuisance parameters θ introduce additional flexibility to the model, which reflects the loss of information due to the systematic uncertainties.

The nuisance parameters are usually constrained by subsidiary measurements, distributed in a new histogram $\mathbf{m} = (m_1, ..., m_M)$, where the expectation values of bin m_j is given by $E[m_j] = u_j(\theta)$.

Together with the nuisance parameters, equation 2.71 becomes:

$$L(\mu, \theta) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \prod_{j=1}^{N} \frac{u_j^{m_j}}{m_j!} e^{-u_j}. \tag{2.72}$$

One can define a profile likelihood ratio, such that:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})}, \tag{2.73}$$

where $\hat{\theta}$ and $\hat{\mu}$ are defined as:

$$(\hat{\mu}, \hat{\theta}) = argmax\ L(\mu, \theta). \tag{2.74}$$

and $\hat{\hat{\theta}}$ are the values of the nuisance parameters 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 observed data is well compatible with the hypothesis when the signal strength has the specific value μ. Using the profile likelihood ratio, one can define the test statistic as:

$$t_\mu = -2 \ln \lambda(\mu), \tag{2.75}$$

in order that t_μ is positive, and a small value of t_μ suggests the data is not compatible with the hypothesis.

To establish a positive signal discovery, 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 likely explanation is that this is due to some systematic error. Therefore, for **signal discovery**, the test statistic of the background-only null hypothesis is given by:

$$t_0 = q_0 = \begin{cases} -2 \ln \lambda(0) & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0. \end{cases} \tag{2.76}$$

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:

$$t_\mu = q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} \tag{2.77}$$

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

$$p = \int_{t_{obs}}^{\infty} f(t|H_0)dt, \tag{2.78}$$

where t_{obs} is the observed value of the test statistic, $f(t|H_0)$ is the probability density of measuring test statistic of value t given the null hypothesis. Equivalently speaking, the null hypothesis is rejected with a probability of $1-p$, which is referred to as *confidence level*. This thesis sets upper limits on the di-Higgs production cross-section. Taking this as an example, the null hypothesis (signal-plus-background hypothesis) is rejected if the p-value is smaller than a conventional threshold of 0.05, corresponding to rejecting the null hypothesis at 95% confidence level.

The p-value is often 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:

$$Z = \Phi^{-1}(1-p), \tag{2.79}$$

where Φ^{-1} is the quantile function of the standard Gaussian distribution. For setting upper limits, the conventional threshold of 0.05 corresponds to $Z = 1.64$. While for signal discovery, the requirement is much more stringent, that the Z values are required to be $Z \geq 5$, which corresponds to a p-value of 2.87×10^{-7}. The sensitivity of an experiment can be measured by the *expected* significance for a given signal hypothesis. For example, the sensitivity to discovering a given signal process, H_1, can be quantified by the expectation value of Z obtained by testing the background-only model, H_0, under the assumption of H_1, and vice versa. In chapter 7, the upper limits are set using the CL_s method at 95% confidence level, with CL_s defined as [57]:

$$CL_s = \frac{p_{s+b}}{1-p_b}, \tag{2.80}$$

where p_{s+b} and p_b are the p-values for the signal-plus-background hypothesis and the background-only hypothesis, respectively. The reason for using CL_s instead of just p_{s+b} is to avoid excluding a signal falsely in the case where the analysis has little or no sensitivity to the signal. In such a case, the probability density functions of $f(t|s+b)$ and $f(t|b)$ almost entirely overlap, the denominator $1 - p_b$ becomes small and p_{s+b} is penalised, preventing the signal model being excluded.

2.5.1 Look elsewhere effect

When searching for a BSM resonance, one usually scans a mass range. However, suppose one specified a hypothesis with a specific resonance mass and observed an excess at this mass point, one should also take into account that this signal could be a fluctuation which could be observed anywhere in the sensitivity range [58]. This effect is referred to as the *look-elsewhere effect*. A common belief is that this 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 quotes 5 σ, the effective significance might be lower. In this thesis, the look-elsewhere effect is accounted by calculating the *global* significance following Ref. [59], using the up-crossing method:

$$p_{\text{global}} = p_{\text{local}} + N_{\text{up}} e^{-1/2(Z^2_{\text{local}} - Z^2_{\text{ref}})}, \tag{2.81}$$

where the p_{local} are the *local* p-values measured at specific mass points, 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 crosses the reference line (in this case, p=0.5) in the upward 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.

Chapter 3

The ATLAS experiment at the Large Hadron Collider

3.1 The Large Hadron Collider

The Large Hadron Collider [60] is the world's largest and most powerful particle accelerator. It started operation in 2008 and retains its crucial role in the many accelerators in the world. The collider is located in a ring tunnel of circumference 26.7 km, which lies beneath the French-Swiss border near Geneva, with superconducting magnets along the tunnel to keep the particle beam in orbit and accelerating structures to boost the beam to the desired energy. Inside the tunnel, *bunches* of up to 1.15×10^{11} protons travelling at close to the speed of light in opposite directions are collided 40 million times per second at four crossing points, around which are positioned four main detetors: ATLAS (A Toroidal LHC ApparatuS) [61], CMS [62] (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) [63] and LHCb (b stands for beauty) [64].

3.1.1 The LHC Accelerator Complex

The proton beams colliding in the LHC have an energy of the order of a few TeV. To reach such high energy, a series of acceleration steps are required for the beams before entering the LHC ring. The protons are supplied by the injector chain Linac 2 — Proton Synchrotron Booster (PSB) — Proton Synchrotron (PS) — Super Proton Synchrotron (SPS), as shown in Figure 3.1.

The protons are produced by stripping off the electrons from hydrogen gas in an electric field. Linac 2, the first accelerator in the chain, is a linear accelerator which accelerates the protons to an energy of 50 MeV. They then enter the PSB, which accelerates the protons to 1.4 GeV, followed by the PS, where they reach 25 GeV. The series of radio frequency cavities in the PS splits the beam into discrete bunches of protons of 25 ns spacing. These

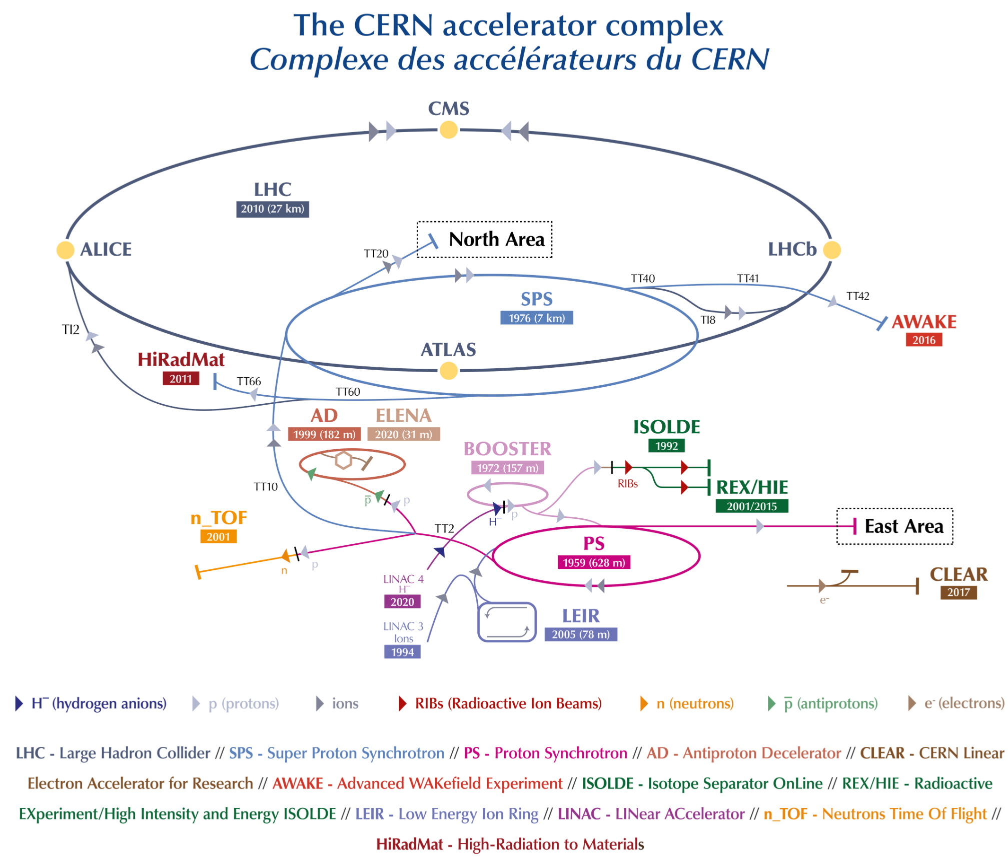


Figure 3.1: The LHC is the last ring (dark blue line) in a complex chain of particle accelerators. The smaller machines are used in a chain to help boost the particles to their final energies and provide beams to a whole set of smaller experiments.

bunches are then accelerated to 450 GeV in the SPS, from which they are finally injected into the beam pipes of the LHC. Each proton beam contains 2808 bunches, arranged in "trains" with 72 bunches in each "carriage", with a gap of around 320 ns between carriages. The beams are required to have well defined transverse and longitudinal emittance.

The beam pipes are kept at ultra-high vacuum, a vacuum thinner than interstellar void, maintained for 24 km of low-temperature section and 3 km of room-temperature section. For the low-temperature section, the vacuum is achieved by pumping in 9000 m^3 of cryogenic gas, which later will be condensed and adhered to the surface of the beampipe. For the room temperature section, the vacuum is achieved by use of a non-evaporable getter (NEG) that absorbs residue gas particles when heated. More residue is absorbed by an ion pumper.

The beam pipes are installed in the existing tunnel that was constructed between 1984 and 1989 for the Large Electron-Positron Collider (LEP) [65], which lies between 45 m and 170 m below the surface on a plane inclined at 1.4% sloping towards the Léman lake.

There are advantages and disadvantages for a proton-proton collider (LHC) compared to the particle-antiparticle colliders (electron-positron collider like LEP, or proton-anti-proton collider). One disadvantage is two rings are needed to accommodate the two counter-rotating beams for the LHC, while the particle-antiparticle colliders can have both beams sharing the same phase space in a single ring. On the other hand, LHC is able to achieve very high collision energy, which is not possible using electron-positron colliders, neither linear nor circular ones. Moreover, LHC is able to achieve very high luminosity, which is not possible using proton-anti-proton colliders.

The electron-positron colliders cannot achieve high energy because, in a circular electron collider, synchrotron radiation lost is proportional to the Lorentz factor $\gamma = E/m$ to the power of four, where E and m are the energy and mass of the particle, respectively. Since electrons are about 2000 times lighter than protons, synchrotron radiation lost is at the order of 10^{13} faster for electrons than for protons. For a linear electron-positron collider, an extremely long acceleration section is required with current technologies, which makes it an impractical option.

As for the proton-anti-proton collider, it would not be possible to achieve such high luminosity using anti-proton beams, since it is much more difficult to produce anti-protons than to produce protons. In addition, at high energies the proton anti-proton collider starts losing one of its advantages of having higher cross-section, which is due to the quark sea and anti-quarks in protons becoming more "visible" at high energies.

As explained above, two separate rings are required to accommodate the two beam pipes, while the internal diameter of the tunnel is only about 3.8 m. It's technically challenging to install them in such small space. LHC therefore adopted the twin-bore magnet design [66], as shown in Figure 3.2. It was first proposed by John Blewett at the Brookhaven laboratory in 1971 due to cost considerations [67], but in the case of the LHC the overriding reason for adopting this solution is the lack of space in the tunnel.

3.1.2 Luminosity and pileup

Luminosity is an important measure of a collider's performance. It relates closely to the number of events generated per second, given by:

$$N = \mathcal{L}\sigma, \tag{3.1}$$

where σ is the scattering cross-section for the process under study and $\mathcal{L}$ is the machine instantaneous luminosity. For the cross-section, it is more common to use barns as the unit, where $1\text{b} = 10^{-28}\ \text{m}^2 = 10^{-24}\ \text{cm}^2$, since particle interactions usually have very small cross-sections. The machine luminosity depends on the beam parameters and can

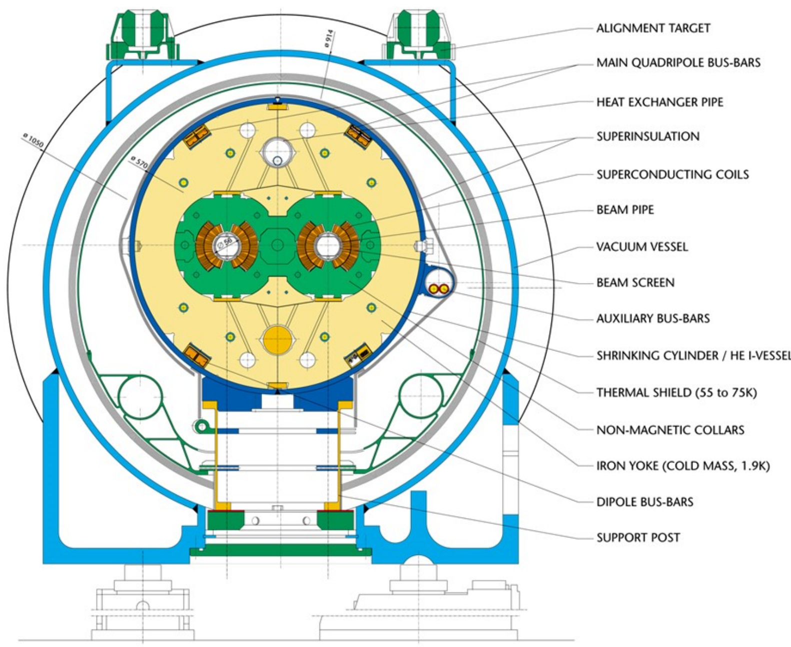


Figure 3.2: Double-bore magnet configuration of the LHC superconducting magnets [66].

be written for a Gaussian beam distribution as:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta *} F, \tag{3.2}$$

where N_b refers to the number of protons per bunch, n_b is number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ϵ_r is the normalised transverse beam emittance, $\beta *$ is the beta function at the collision point which describes the size of the beam, and F refers to the geometric luminosity reduction factor due to the crossing angle at the interaction point. While the instantaneous luminosity measures the rate of collisions (per unit cross-section), the total number of collisions (per unit cross-section) is measured by the integrated luminosity L, given by:

$$L = \int \mathcal{L} dt, \tag{3.3}$$

which is the integral of the instantaneous luminosity over time.

The two general-purpose experiments, ATLAS and CMS are both aiming at a peak luminosity of $\mathcal{L} = 10^{34}$ cm^{2}s^{-1} for proton-proton collisions, which corresponds to about one billion pp collisions per second. The instantaneous luminosity was much improved in real-time operations, reaching about twice the nominal value (from 2015 to 2018) thanks to the effort of the LHC experts.

Another important parameter for the LHC experiments is the pileup, which is a measure of the number of inelastic pp interactions that occur per bunch crossing. Higher pileup gives more luminosity (for a fixed number of bunches) but makes physics analysis more difficult due to the signals in the detector from the additional interactions. The distribution of the recorded luminosity in terms of the pileup is shown in Figure 3.3 for operations from 2015 to 2018 (Run 2), where the $<\mu>$ stands for the mean number of interactions per bunch crossing. There are two main sources of pileup: in-time pileup and out-of-time pileup. The former refers to the additional proton-proton collisions occurring in the same bunch-crossing as the collision of interest. The latter refers to the additional proton-proton collisions occurring in bunch-crossings just before and after the collision of interest.

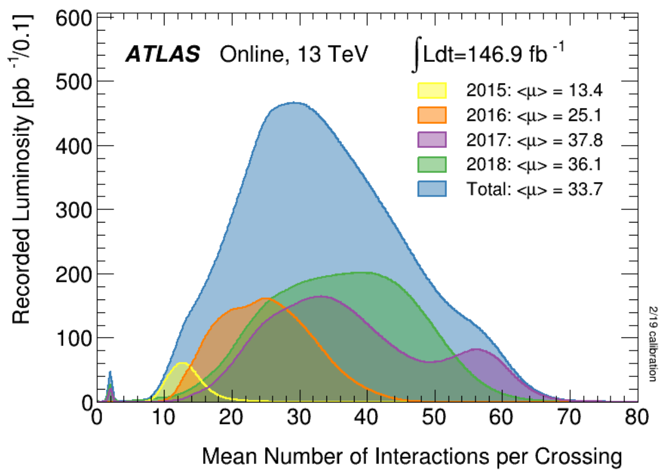


Figure 3.3: Shown is the luminosity-weighted distribution of the mean number of interactions per crossing for the data collected by the ATLAS from 2015 to 2018. All data recorded by ATLAS during stable beams are shown, and the integrated luminosity and the mean μ value is given in the figure.

3.1.3 Operation schedule

The LHC operation and shutdown schedule is shown in Figure 3.4. Following the downtime after an incident in one of the main dipole circuits during the first commissioning in 2008 [68], the operation restarted at lower beam energy to minimise the risk. Therefore, the first proton run (2010-2013) [69] was carried out at 3.5-4 TeV per beam (centre-of-mass energy 7-8 TeV). Furthermore, a bunch spacing of 50 ns was used instead of the nominal 25 ns, with peak luminosity of 0.8×10^{34} cm^{-2}s^{-1}, and pileup larger than nominal.

In Run 1, the LHC delivered about 30 fb^{-1} of proton data and important physics results, most notably the discovery of the Higgs boson [1, 70]. Run 1 was followed by a long shutdown (LS1, 2013-2014) with a large number of consolidation and upgrade activities [71]. The bus-bar splices between the superconducting magnets were improved,

Figure 3.4: LHC operation schedule and luminosity targets.

in order to make sure that the LHC could operate at higher energy without risk of repeating the 2008 incident.

Run 2 (2016-2018) was carried out at 6.5 TeV per beam (centre of mass energy 13 TeV) [72]. As shown in Figure 3.5, out of the 156 fb^{-1} of data LHC has delivered at 13 TeV centre-of-mass energy, the ATLAS detector has recorded 147 fb^{-1} and 139 fb^{-1} of data is certified to be good quality data. The 156 fb^{-1} data accounts for the luminosity delivered from the start of stable beams until the LHC requests ATLAS to put the detector in a safe standby mode to allow a beam dump or beam studies. The recorded luminosity is slightly smaller than the delivered luminosity due to the inefficiency of the Data Acquisition and the so-called "warm start": when the stable beam flag is raised, the tracking detectors undergo a ramp of the high-voltage and, for the pixel system, turning on the pre-amplifiers. More details of the ATLAS detector can be found in the following sections. The recorded data is checked carefully to exclude possible hardware or software issues. This is achieved by monitoring detector-level quantities and reconstructed collision event characteristics at key stages of the data processing chain. This procedure led to the high efficiency of good quality data: 95.6% [73].

This thesis uses the 139 fb^{-1} data recorded by the ATLAS detector of Run 2. The nominal bunch spacing of 25 ns was used, with slightly fewer bunches (2500) per beam. The LHC experts have continually improved the running scenario to increase the luminosity, and during Run 2, the luminosity surpassed the designed luminosity by a factor of 2. As well as improving the instantaneous luminosity, the availability of the machine was dramatically improved during Run 2, which is an essential factor enabling the high effi-

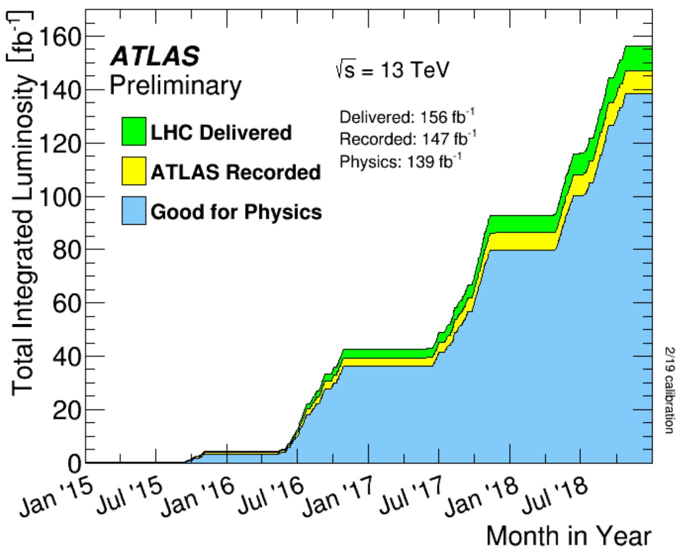


Figure 3.5: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at 13 TeV centre-of-mass energy in Run 2.

ciency of good quality data as mentioned above. During Run 2, the machine was providing physics collisions during 50% of the allocated physics time, which is very impressive for a superconducting collider.

The operation of CERN's accelerators is subject to scheduled shutdowns to allow necessary repair and upgrade work to take place. The present shutdown, LS2, is devoted to preparations for Run 3 of the LHC, which will have an integrated luminosity equal to the two previous runs combined, and for the High-Luminosity LHC (HL-LHC), the successor to the LHC, which will begin operation at the end of 2027, and eventually, generate 10 times the integrated luminosity of all Run 1, 2 and 3 combined!

The LS2 schedule has had to be modified due to the COVID-19 pandemic, which the new schedule anticipates that the first test beams will circulate in the LHC at the end of September 2021, four months later than the date planned before the COVID-19 crisis, to give the LHC's main experiments time to prepare their own upgrade. Run 3 of the LHC will begin at the start of March 2022. The third long shutdown (LS3) will begin at the start of 2025 and end in mid-2027. This is when the equipment for the HL-LHC and its experiments will be installed.

3.2 The ATLAS Detector

ATLAS is one of the two general-purpose detectors built for probing proton-proton collisions. This detector represents the work of a large collaboration of several thousand physicists, engineers, technicians, and students over a period of fifteen years of dedicated design, development, fabrication, and installation. The overall layout of the detector is shown in Figure 3.6 [61]. It has the shape of a cylinder, 46 m long, 25 m in diameter, and sits in a cavern 100 m below ground. The ATLAS detector weighs 7000 tonnes, similar to the weight of the Eiffel Tower. The detector itself is a many-layered instrument designed to detect some of most energetic particles ever created on earth. It consists of six different detecting subsystems wrapped concentrically in layers around the collision point of nearly 4π solid angle coverage to record the trajectory, momentum, and energy of particles, allowing them to be individually identified and measured. These six subsystems are the pixel detector [74], the semiconductor tracker (SCT) [75], the transition radiation tracker (TRT) [75], the electromagnetic (EM) calorimeter [76], the hadronic calorimeter [77] and the muon spectrometer (MS) [78]. The first three sub-detectors are collectively known as the inner detector (ID), described in Section 3.2.3, and it is used for tracking charged particles. The electromagnetic and the hadronic calorimeter, described in Section 3.2.4, are responsible for measuring the energies of the electromagnetic and hadronic particles, respectively. The MS, described in Section 3.2.5, is a unique sub-detector used for measuring the momentum of muons leaving the calorimeters.

A huge magnet system bends the paths of the charged particles so that their momenta can be measured as precisely as possible.

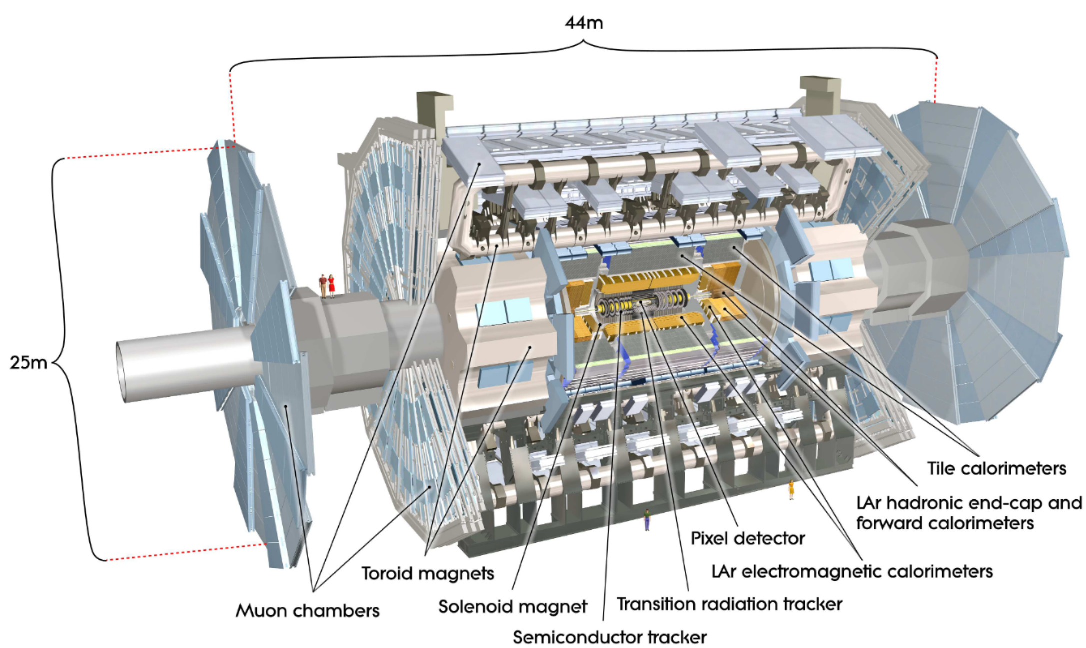


Figure 3.6: Cut-away view of the ATLAS detector. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tonnes. The figure is taken from Ref. [61].

The high interaction rates, radiation doses, particle multiplicities and energies, as well as the requirements for precision measurements have set stringent standards on the design of the ATLAS detector. Therefore the ATLAS detector is designed to fulfil the following requirements:

- Fast, radiation-resistant electronics and sensor elements and high detector granularity. This is due to the high frequency of collisions, high particle fluxes and high radiation environment of the detector.
- Large acceptance in polar angle with almost full azimuthal angle coverage, due to the geometry of the detector (more details in Section 3.2.1).
- Good energy resolution calorimetry, as required to enable accurate physical object reconstruction. The high resolution of energy can be obtained with very good electromagnetic calorimetry for electron and photon identification and measurements, complemented by full-coverage hadronic calorimetry for accurate jet and missing transverse energy measurements.
- Tracking of precision in the ID, as required to provide high momentum resolution and to allow the reconstruction of secondary vertices to identify b-hadrons and τ-leptons.
- Good muon identification and momentum resolution over a wide range of momenta and the ability to determine the charge of high-p_{T}muons unambiguously in the muon spectrometer.
- Trigger system with high efficiency for low p_{T} objects with sufficient background rejection, which is a prerequisite to achieving an acceptable trigger rate for most physics processes of interest.

The main performance goals of the detector are listed in Table 3.1.

3.2.1 Coordinate system

The ATLAS coordinate system is a right-handed Cartesian system with the nominal interaction point defined as the origin of the coordinate system, while the beam direction defines the z-axis and the x-y plane is transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the centre of the LHC ring and the positive y-axis is defined as pointing upwards, as shown in Figure 3.7.

The azimuthal angle ϕ is measured as usual around the beam axis, and the polar angle θ is the angle from the beam axis. In high energy physics, it's more common to use the

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\%\ p_T \bigoplus 1\%$	± 2.5	None
EM calorimetry	$\sigma_E/E = 10\%\ /E \bigoplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry			
barrel and end-cap	$\sigma_E/E = 50\%/E \bigoplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\%\ /E \bigoplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T =$ 1 TeV	± 2.7	± 2.4

Table 3.1: General performance goals of the ATLAS detector. The units for energy of the particle, E and transverse momentum, p_{T} (detailed definition in Section 3.2.1) are in GeV [61]. Note that, for high-p_{T} muons, the muon-spectrometer performance is independent of the inner-detector system.

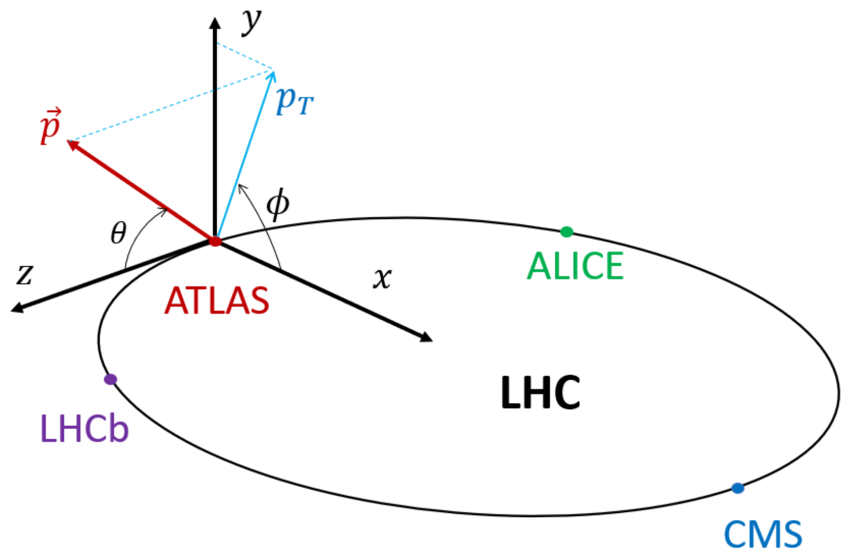


Figure 3.7: Illustration of the coordinate system used at the ATLAS experiment in the geographical context of the LHC.

pseudorapidity instead of the polar angle θ, defined as:

$$\eta = -\ln\ \tan(\theta/2). \tag{3.4}$$

When the particle is travelling close to the beam pipe, $\eta \to \pm\infty$ (positive for along the beam pipe, negative for the opposite direction) and $\eta \to 0$ when p_z is small. In the case of highly relativistic particles (which is the common case in high energy physics), the pseudorapidity approaches the rapidity,

$$y = 1/2 \ln[(E + p_z)/(E - p_z)], \tag{3.5}$$

where E is the energy of the particle, m is its mass and p_z is the momentum along the

z-axis.

It can be derived that the difference in rapidity of two particles is invariant under a Lorentz transformation along the beam axis, which is not the case for θ. This is the main reason for not using θ. The reason for using pseudorapidity but not rapidity is that due to the limited angle coverage of the detector, it's usually hard to determine the total energy and the momentum along the z-axis, especially when the direction of the particles are close to the beam pipe. While the pseudorapidity is determined only by the polar angle, which is much easier and faster to compute.

Another commonly used variable, transverse momentum p_{T}, is defined as the momentum of a particle transverse to the beam direction (z-direction):

$$\vec{p_T} = (p_x, p_y). \tag{3.6}$$

The reason for using the transverse momentum is that, because the partons that make up a proton share the momentum, the initial longitudinal momentum is unknown; we do know, however, that the initial transverse momentum was zero. And hence we can look for the missing transverse momentum, defined as

$$\vec{E_T}^{miss} = -\sum_i \vec{p_{T_i}} \tag{3.7}$$

for visible particles i, where $E_{\mathrm{T}}^{\mathrm{miss}}$ is the magnitude of $\vec{E_T}^{miss}$ (Confusingly E_T^{miss} is commonly called missing transverse energy or MET. Missing transverse energy is equivalent to missing transverse momentum only if the missing particle(s) were massless.).

Finally, the distance ΔR in the pseudorapidity-azimuthal angle space between two particles' directions is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}, \tag{3.8}$$

where $\Delta\eta$ and $\Delta\phi$ are the difference in the η and ϕ coordinates, respectively.

3.2.2 Magnets

ATLAS has a unique hybrid system of four large superconducting magnets [79]. This magnetic system is 22 m in diameter and 26 m in length, with a stored energy of 1.6 GJ. Figure 3.8 shows the real scale of the magnets system compared to a person. Figure 3.6 shows the general layout, the four main layers of detectors and the four superconducting magnets which provide the magnetic field over a volume of approximately 12000 m^3.

The spatial arrangement of the coil windings is shown in Figure 3.9. The ATLAS

Figure 3.8: A picture showing the real size of the toroidal magnets compared to a person.

magnet system consists of two parts:

- a solenoid, which is aligned on the beam axis and provides a 2 T axial magnetic field in the z-direction for the ID. Because the magnet is located in front of the EM calorimeter, it is imperative to minimise possible interactions between the magnet and the particles being studied. This is achieved by embedding over 9 km of niobium-titanium superconductor wires into strengthened, pure aluminium strips, which is capable of providing such a powerful magnetic field in just 4.5 cm thickness.
- A barrel toroid and two end-cap toroids, which produce a toroidal magnetic field of approximately 0.5 T and 1 T for the muon detectors in the central and end-cap regions, respectively. The barrel toroid generates the magnetic field in the central zone of the muon spectrometer, along the tangential direction of the circumferences centred on the z-axis (ϕ direction). The end-cap toroids are two smaller toroids designed to provide the magnetic field in the forward areas of the muon spectrometer. This magnet configuration provides a field that is mostly orthogonal to the muon trajectories.

3.2.3 Inner detector

The inner detector is the closest sub-detector to the beamline, designed to track the early trajectories of charged particles for momentum calculations and locate their primary and secondary vertices with extremely high precision. The ID is required to deal with a large

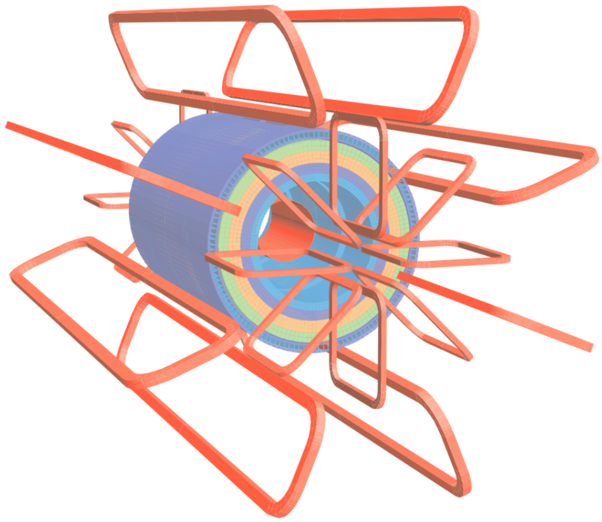


Figure 3.9: Geometry of magnet windings and tile calorimeter steel. Image taken from [61].

number of tracks promptly, of the order of 1000 particles per collision arising every 25 ns. The ID has full coverage in azimuthal angle ϕ and $|\eta| < 2.5$ acceptance in pseudorapidity. As mentioned briefly above, it consists of three parts: the pixel detector and the insertable B-Layer (IBL) [80] (as one part), the semiconductor tracker and the transition radiation tracker. The layout of the ID is shown in Figure 3.10, with a charged track (in red) traversing the sensors and structural elements.

3.2.3.1 Pixel detector and IBL

The silicon pixel detector is the closest ATLAS component to the collision. It is composed of layers of silicon pixels and designed to have a very high granularity for reconstructing primary and secondary interaction vertices. The detector layers are formed of silicon sensor modules and in total, there are approximately 92 million pixels (consequently, 92 million readout channels) in the system. It consists of three cylindrical layers in the barrel region positioned at the radial distances of 50.5, 88.5 and 122.5 mm, and of disks perpendicular to the beams in the end-caps at the longitudinal distances of 49.5, 58.0 and 65.0 mm. In 2014, during the first LHC long shutdown, a fourth pixel layer was installed inside the existing detector, the insertable B-Layer (IBL) at a radius of 33 mm from the beam axis. The new pixel layer provides an additional space point very close to the interaction point, which significantly improves the identification of jets coming from b quark hadronisation (b-jets). Particles with $|\eta| < 2.5$ traverses the four layers usually produce four space-points. The pixel detector provides a resolution of $\sigma_\phi = 10\ \mu$m in the bending direction ($R - \phi$), and a resolution of $\sigma = 115\ \mu$m in the $z(R)$ direction in the barrel (end-cap) region.

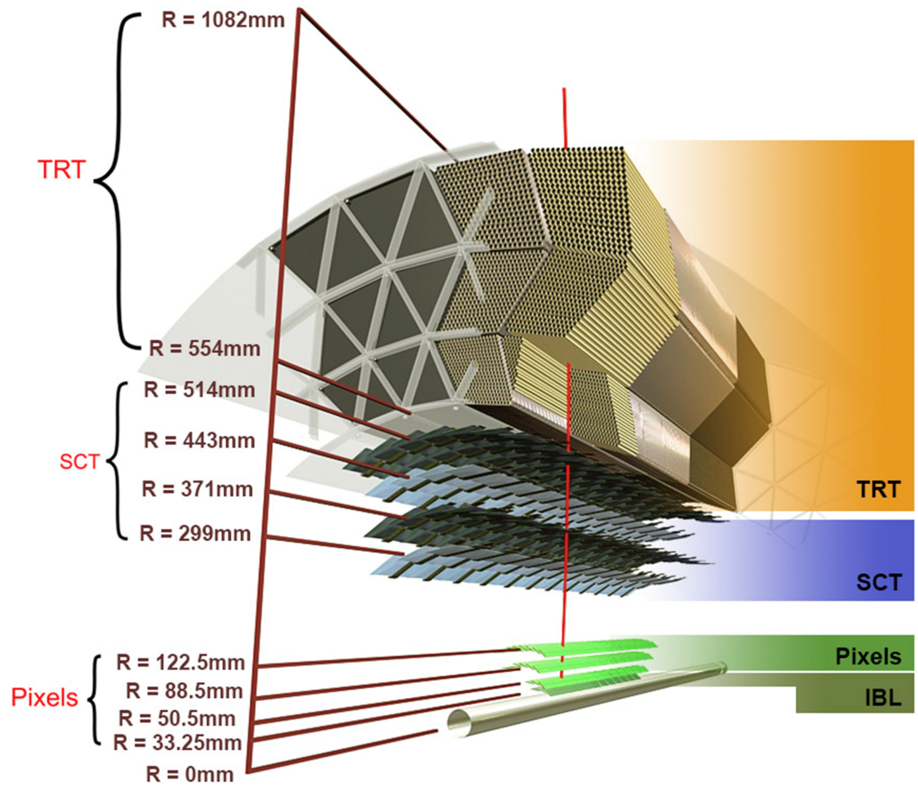


Figure 3.10: Cut-away view of the inner detector. Image taken from [81].

3.2.3.2 Semiconductor tracker

The next constituent of the inner detector is the SCT. It is a silicon microstrip detector with over six million readout channels, which surrounds the pixel detector and covers the region of radius between 299 mm and 560 mm. It consists of four layers of strips located axially on the beam direction in the barrel region and placed along the z-direction in the end-cap region. This configuration allows the particles along the beam pipe to be constructed. Each layer of strips is glued back to back with an angle of 40 mrad to form a two-sided module and make possible the measurement of the second coordinate. The sensors are 285 μm thick and are constructed of high-resistivity n-type bulk silicon with p-type implants. Readout strips are positioned every 80 μm, providing a spatial resolution of $\sigma_\phi = 17\ \mu m$ in the bending direction $(R - \phi)$ and $\sigma_\phi = 580\ \mu m$ in the z (barrel) and R (end-cap) direction.

3.2.3.3 Transition radiation tracker

The outermost part of the inner detector is the TRT, which covers the radial region between 563 mm and 1066 mm. It is a straw drift tube tracker, which consists of modules of 4 mm diameter polyimide straws, filled with a mixture of gas of 70% Xe, 27% CO_2 and 3% O_2

and a gold-plated tungsten wire in the centre. The straws are interleaved with propylene fibres (foils) in the barrel (end-cap) region. With a spatial resolution of $\sigma_\phi = 130\ \mu$m, the TRT measures the track position only in the bending direction ($R - \phi$). This is because when a charged particle passes through a straw tube, electrons from the gas are liberated through ionisation; under high voltage, these electrons then drift toward the wire in the centre, where a current flow is created and registered as a hit. Since a hit can happen on any location along the wire, the information of the z position of the particle is lost. In addition, the TRT provides the capability of distinguishing electrons from other charged particles. When a highly-relativistic charged particle traverses the polymer straws interface, the particle emits transition radiation, which is then absorbed by the Xeon gas. The intensity of the radiation depends on the gamma factor of the particle (strongest for lighter particles). Hence this information can be exploited for electron identification.

3.2.4 Calorimeter system

Calorimeters are used to measure the energy of both charged and neutral particles. The ATLAS calorimeters [82], as shown in Figure 3.11, consist of three major components, the Electromagnetic calorimeter, the Hadronic calorimeter and the Forward calorimeter (FCal). The fine granularity of the EM calorimeter is ideal for precision measurements of electrons and photons; the coarser granularity of the hadronic calorimeter is sufficient for hadronic jet reconstruction; the FCal provides coverage of large pseudorapidity region: these calorimeters cover the range $|\eta| < 4.9$. All three calorimeters are sampling calorimeters. Sampling calorimeters use different materials for the absorber and the active part: the absorber (or passive material) is responsible for producing particle showers where the active part then measures their energy. Note that the fraction of total particles energy deposited in the passive material is not measured; the overall energy must be deduced from the definite measurements taken in the active detector layers.

The electromagnetic and hadronic showers must be contained in the calorimeter to ensure precise measurement of the total energy of the particle and to avoid punch-through into the muon system. The calorimeter depth is hence an important design consideration. The thickness of the calorimeter is measured in radiation length X_0, which is the mean length of a material over which an electron will lose all but $1/e$ of its initial energy through radiative processes; and nuclear interaction length λ, which is the mean distance travelled by a hadronic particle before undergoing an inelastic nuclear interaction. The total thickness of the EM calorimeter is greater than 22 radiation lengths (X_0) in the barrel and greater than 24 X_0 in the end-caps. The total thickness of the calorimeters, including 1.3 λ from the outer support, is 11 λ at $\eta = 0$ and has been shown both by measurements and simulations to be sufficient to reduce punch-through well below the irreducible level

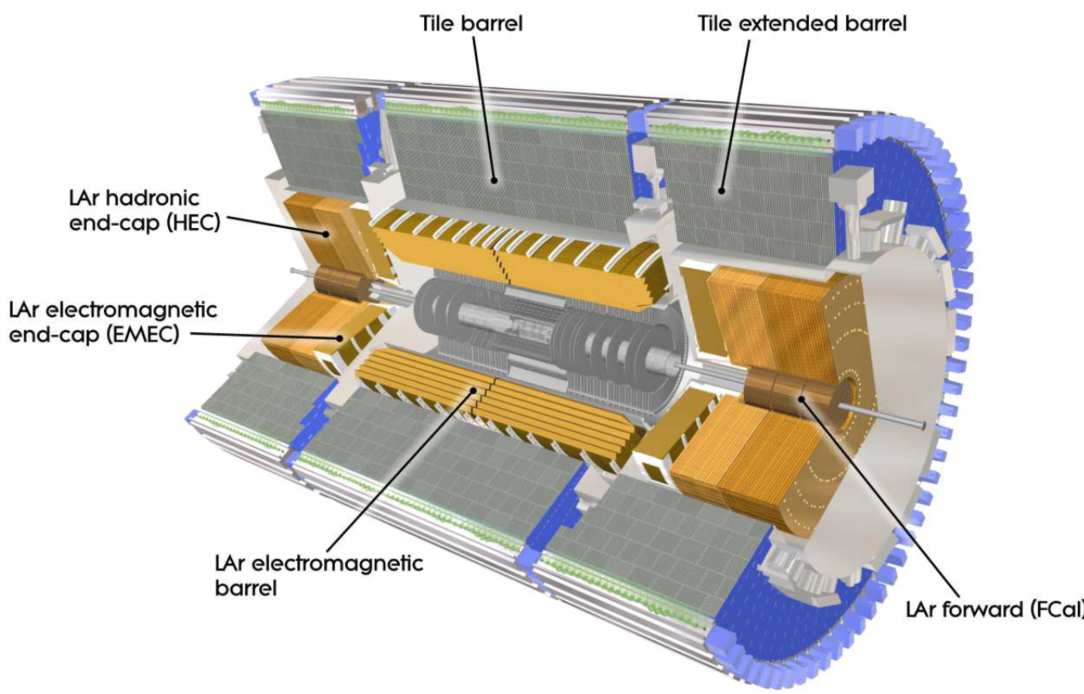


Figure 3.11: Cut-away view of the ATLAS calorimeter system. Image taken from [82].

of prompt or decay muons.

3.2.4.1 Electromagnetic calorimeter

The EM calorimeter is a lead-Liquid argon (lead-LAr) detector [76] with accordion-shaped (as shown in Figure 3.12) kapton electrodes and lead absorber plates over its full coverage, while using liquid argon (LAr) as the active material. The lead thickness in the absorber plates has been optimised as a function of η in terms of EM calorimeter performance in energy resolution. The EM calorimeter is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$), each housed in their own cryostat. Additional material needed to instrument and cool the detector creates a "crack" region at $1.375 < |\eta| < 1.52$, where the energy resolution is significantly degraded.

The barrel calorimeter consists of two identical half-barrels, separated by a small gap (4 mm) at $z = 0$. Each end-cap calorimeter is mechanically divided into two coaxial wheels: an outer wheel covering the region $1.375 < |\eta| < 2.5$, and an inner wheel covering the region $2.5 < |\eta| < 3.2$.

The calorimeter has three layers along the transverse direction: a pre-sampler with very high granularity in η, in order to reconstruct the neutral pions decaying to two photons and particles which already starts showering in the inner detector. The pre-sampler is followed by longer towers of relatively high granularity, which is the major part of detecting EM showers, and reponsible for measuring the η and ϕ coordinates of the particles. The last

Figure 3.12: A figure of the accordion-shaped electrodes.

layer detects showers generated from particles other than electrons or photons that start showering inside the EM calorimeter before leaving it.

3.2.4.2 Hadronic calorimeter

The Hadronic calorimeter is comprised of the Tile Hadronic calorimeter (HCAL) and the LAr hadronic end-cap calorimeter (HEC). The HCAL is placed directly outside the EM calorimeter envelope. Its barrel covers the region $|\eta| < 1.0$, and its two extended barrels the range $0.8 < |\eta| < 1.7$. It uses steel as the absorber and scintillating tiles as the active material. It is segmented in-depth in three layers, approximately 1.5, 4.1 and 1.8 λ thick for the barrel and 1.5, 2.6, and 3.3 λ for the extended barrel. The total detector thickness at the outer edge of the tile-instrumented region is 9.7 λ at $\eta = 0$.

The HEC is similar to the construction of the ECAL, using LAr as the active material, but instead of using lead, it uses copper as the absorber. It consists of two independent wheels per end-cap, located directly behind the end-cap electromagnetic calorimeter and sharing the same LAr cryostats. The HEC covers the range of $1.5 < |\eta| < 3.2$, slightly overlapping with the forward calorimeter which will be described in the following paragraph (around $|\eta|= 3.1$) and the tile calorimeter ($|\eta| < 1.7$). This overlap is to reduce the drop in material density at the transition between the different calorimeters.

3.2.4.3 Forward calorimeter

The Forward Calorimeter (FCal) covers $3.1 < |\eta| < 4.9$ and is approximately 10 interaction lengths deep. It consists of three modules in each end-cap: the first, made of copper, is optimised for electromagnetic measurements, while the other two, made of tungsten, measure predominantly the energy of hadronic interactions. All three modules use LAr as active material. Due to high particle fluxes and energies in the forward region, the calorimeter must contain relatively long showers in the small volume allowed by design constraints, and thus must be very dense.

3.2.5 Muon Spectrometer

The muon spectrometer is the outermost and largest sub-detector of ATLAS. A cut-away view of the MS is shown in Figure 3.13. It fully covers the calorimeter system and occupies a large part of the ATLAS cavern. It is based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers. Over the range $|\eta| < 1.4$, magnetic bending is provided by the large barrel toroid. For $1.6 < |\eta| < 2.7$, muon tracks are bent by two smaller end-cap magnets inserted into both ends of the barrel toroid. Over $1.4 < |\eta| < 1.6$, usually referred to as the transition region, magnetic deflection is provided by a combination of barrel and end-cap fields. The configuration of magnets provides a field mostly orthogonal to the muon trajectories, hence minimising the degradation of resolution due to multiple scattering.

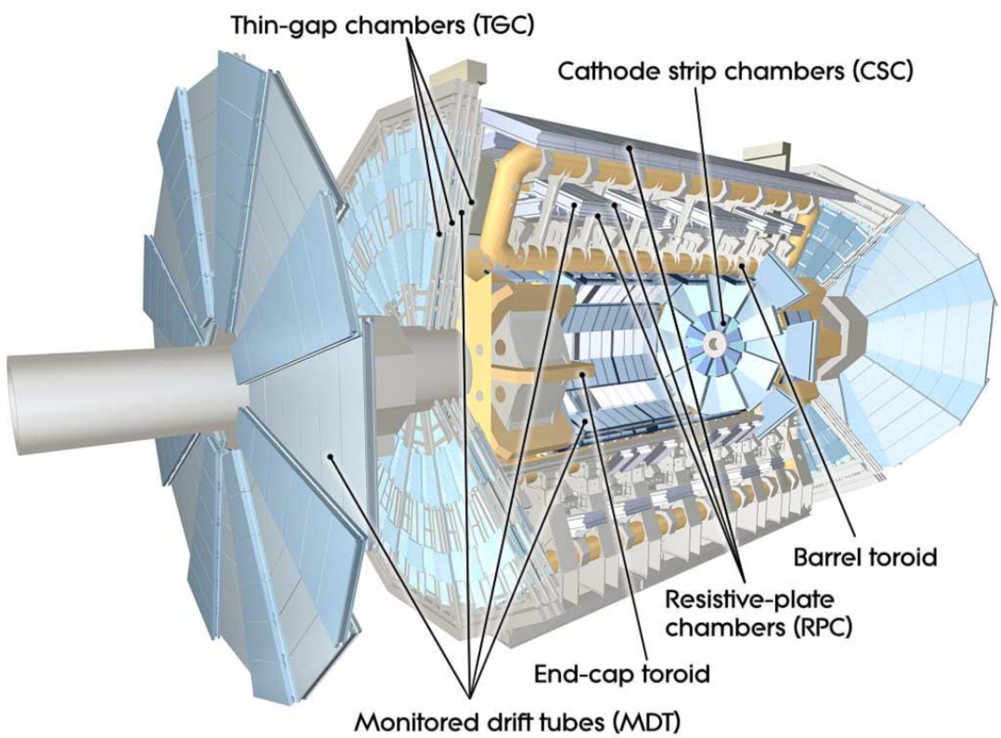


Figure 3.13: Cut-away view of the ATLAS muon spectrometer system. Image taken from [61].

The MS consists of four subsystems which rely on four different gas detector technologies. Two of them, the resistive plate chambers (RPC) in the barrel region and the thin gap chambers (TGC) in the end-cap region, provide trigger signals, while the other two, the monitored drift tubes (MDT) in the barrel and the cathode strip chambers (CSC) in the end-cap region provide the momentum measurement. The MDT chambers provide high precision measurements in the bending direction over most of the detector acceptance while the CSC is used in the forward region where the particle flux is too high for the MDT chambers. The muon chambers are arranged in the barrel ($|\eta| < 1.05$) in three cylindrical layers around the beam axis, while in the end-cap regions ($1.05 < |\eta| < 2.7$), they are placed in three wheels. The resolution of muons tracks momentum measurement varies from typically 2-3% over most of the kinematic range, to about 10% at $p_{\mathrm{T}} = 1$ TeV.

3.2.6 Trigger system

As mentioned in section 3.1.3, the spacing between bunches is 25 ns, which translates to a 40 MHz bunch-crossing frequency, with up to 80 collisions per bunch crossing. This is far beyond the data collection bandwidth and storage capacity of ATLAS. Therefore, it's necessary to adopt a trigger system that makes fast decisions as to whether an event is high quality, rare or "interesting" and to save the event or not. The ATLAS trigger system consists of two consecutive parts: the Level 1 trigger [83] which is hardware-based, followed by the software-based High-Level Trigger (HLT) [84].

The L1 trigger searches for signatures from high-p_{T} muons, electrons/photons, jets, and τ-leptons decaying into hadrons. It also selects events with large MET and large total transverse energy. The L1 trigger uses reduced-granularity information from a subset of detectors: the RPC and TGC for high-p_{T} muons, and all the calorimeter subsystems for electromagnetic clusters, jets, τ-leptons, E_{miss}^{T} , and large total transverse energy. As a result, the L1 trigger reduces the event rate from 40 MHz to a maximum of 100 kHz. The decision is made by Central Trigger Processor (CTP), which operates on signals from dedicated hardware in the calorimeter and muon detector systems. The decision time, at under 2.5 μs, is faster than the ID can process events so ID information is omitted. For each data-taking period, the L1 trigger is loaded with a trigger menu, a list of up to 256 criteria used to determine whether an event is accepted. The trigger menus are designed to accommodate a broad physics programme, with high acceptance for both BSM searches and SM precision measurements. The L1 trigger also uses detector information with reduced granularity to identify Regions of Interest (RoI) [85] in ϕ and η. The RoI information with full granularity and precision and all the available detector data (including the ID information) within the RoI's are provided to the HLT. This trigger level reduces the rate of events by two orders of magnitude, reaching an average of 1 kHz with a latency of 300 ms [86]. These events are passed on to a data storage system for offline analysis.

Chapter 4

Physics Object Reconstruction

Particles produced in the pp collisions inside ATLAS can interact with the detector sub-systems with each type of particle leaving a unique signature. Reconstructing and identifying these particles precisely and efficiently using the information recorded in each sub-detector is a building block of physics analysis. Therefore, the following chapter outlines the reconstruction procedure of the particles important for the analysis present in this thesis.

4.1 Track and vertex reconstruction

Track and vertex reconstruction is the starting point of physics objects reconstruction, which makes it crucial to understand how they are implemented in ATLAS. Track reconstruction [87, 88] is performed mainly with the so-called "inside-outside" procedure, complemented by the "outside-in" tracking and and the reconstruction of TRT-standalone tracks. The inside-out stage starts by assembling the raw measurements into *clusters*: an algorithm called connected component analysis [89] groups pixels and strips in a given sensor, where the deposited energy yields a charge above the threshold, with a common edge or corner into clusters.

From clusters, three-dimensional measurements, referred to as *space points*, are created (the yellow points in Figure 4.1). They represent the point where the charged particle traversed the active material of the ID. Each space point equates to one cluster in the pixel detector, while in the SCT, clusters from both sides of a strip layer must be combined to obtain a three-dimensional measurement. Three space points are combined to form track seeds (circled in blue in Figure 4.1). A combinatorial Kalman filter [90] is then used to build track candidates from the chosen seeds by incorporating additional space points from the remaining layers of the pixel and SCT detectors which are compatible with the preliminary trajectory (circled in a blue dashed line in Figure 4.1). A track score computed by the

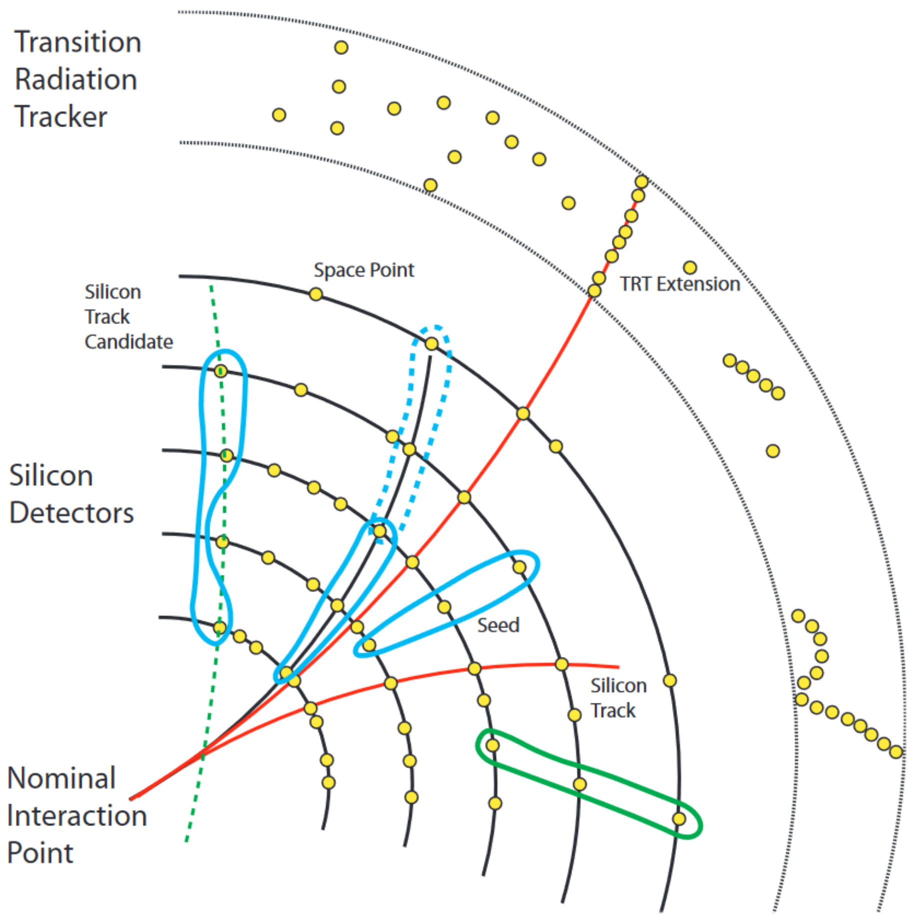


Figure 4.1: An example of track reconstruction. Image reproduced from Ref. [91].

quantities of the fitted track is assigned to each track. The track candidates are processed in descending order of track score and those that fail a set of minimum requirements on p_{T}, the number of holes, and the number of clusters are rejected; candidates that have too many bad quality clusters are stripped down and re-scored, then returned to the list of remaining candidates. This process is referred to as "ambiguity solving" [88]. Finally, the tracks are extended into the TRT and, by using the full information of all three sub-detectors, the tracks are fitted again to determine the final track parameters. The track can be fully represented by five parameters measured at the perigee, which are the impact parameter (IP, transverse distance from the interaction point) d_0, the distance from the interaction point along the z axis z_0, the azimuthal angle ϕ, the polar angle θ and the charge-momentum ratio q/p_T.

The complementary "outside-in" algorithm starts with searching for tracks with segments reconstructed in the TRT, and extend the tracks inwards by adding silicon hits. The tracks are built with the combinatorial Kalman filter and passed to the ambiguity solving procedure. Finally, tracks with a TRT segment but no extension into the silicon detectors are referred to as TRT-standalone tracks.

After the tracks are reconstructed, primary vertices, which are the points where the hard-scattering processes occurred, are reconstructed in two steps [92]: a) the primary vertex finding algorithm, dedicated to associate reconstructed tracks to the vertex candidates, and b) the vertex fitting algorithm, dedicated to reconstructing the vertex position and its corresponding error matrix. It also refits the associated tracks constraining them to originate from the reconstructed interaction point. The vertex finding algorithm works as follows: first, vertex seeds are obtained from the z-position at the beamline of the reconstructed tracks; and then, an iterative χ^2 fit is then performed using the vertex seed and nearby tracks. Vertices are required to contain at least two tracks, and tracks displaced by more than 7σ from the vertex are used to seed a new vertex. The procedure is repeated until no additional vertices can be found, and no unassociated tracks are left in the event. The primary vertex for each event is selected as the vertex with the highest $\sum_{tracks}(p_T^{track})^2$.

4.2 Electrons

When a high energy electron (or positron) enters the detector, it interacts with the detector material primarily via bremsstrahlung. This results in radiation of photons, which subsequently convert into electron-positron pairs that continue to interact with the detector material, leading to a cascade of particles of decreasing energy. These cascade particles are usually referred to as an electromagnetic *shower*. They are very collimated and frequently create neighbouring signals in the calorimeter component. These interactions can occur inside the ID volume or even in the beam pipe, generating multiple tracks in the ID, or can instead occur downstream of the ID, only impacting the shower in the calorimeter.

4.2.1 Reconstruction

The reconstruction of an electron is based on three fundamental components: a) localised clusters of energy deposits found within the EM calorimeter, b) charged tracks identified in the ID (as described in detail in chapter 4.1), and c) close matching in $\eta \times \phi$ space of the tracks to the clusters [93]. Figure 4.2 provides a schematic illustration of the elements that enter into the reconstruction and identification of an electron. The reconstruction starts from EM cluster seeding from localised energy deposits using a sliding-window algorithm [94]. The $\eta \times \phi$ space of the EM is divided into *towers* of 200×256 elements of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, consistent with the granularity of the second layer of the EM calorimeter. Then algorithm then "slides" a rectangular window of size 3×5 towers whose summed transverse energy exceeds 2.5 GeV and is a local maximum to form a seed-cluster. The centre of the seed moves in steps of 0.025 in either the η or ϕ direction

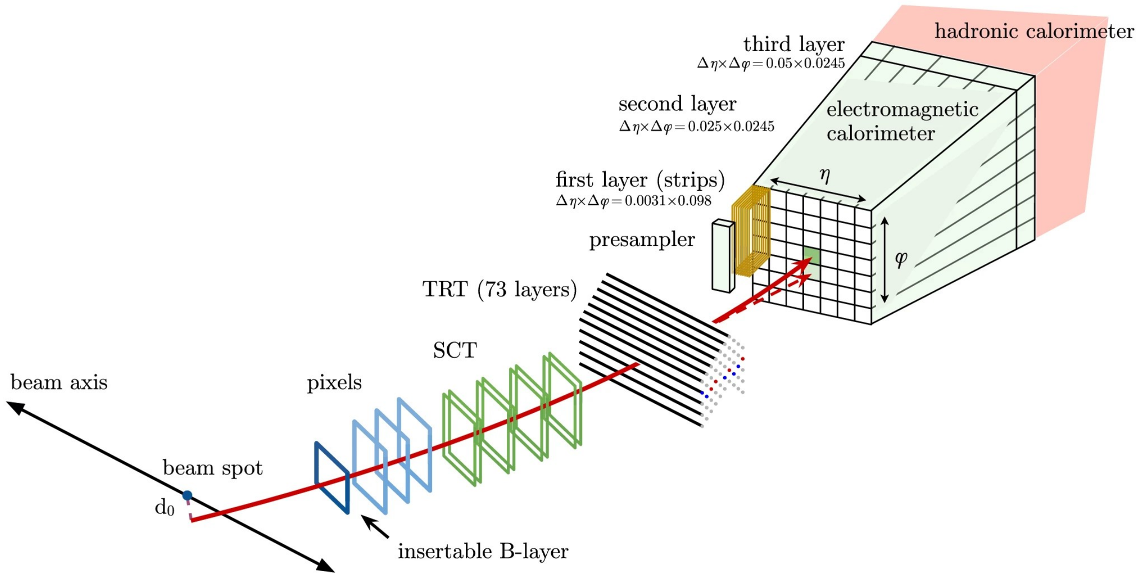


Figure 4.2: A schematic illustration of the path of an electron through the detector. The red trajectory shows the hypothetical path of an electron, which first traverses the tracking system (pixel detectors, then SCT and finally the TRT) and then enters the electromagnetic calorimeter. The dashed red trajectory indicates the path of a photon produced by the interaction of the electron with the material in the tracking system. Image reproduced from Ref. [93].

to search for localised energy deposits; this process is repeated until every element of the calorimeter has been covered. To better account for the energy loss of charged particles in the detector material, a subsequent fitting procedure using an optimised Gaussian-sum filter [95] is performed on tracks that are "loosely" matched to the EM clusters. This matching requires the tracks and clusters to satisfy:

- $|\eta_{cluster} - \eta_{track}| < 0.05$,

and one of the two requirements:

- $-0.20 < \Delta\phi < 0.05$, or
- $-0.10 < \Delta\phi_{res} < 0.05$,

where $\Delta\phi \equiv -q \times (\phi_{cluster} - \phi_{track})$ with q being the charge of the particle, and $\phi_{cluster}$, ϕ_{track} and $\eta_{cluster}$, η_{track} are the ϕ, η coordinates of the cluster barycentre and the position of the track extrapolated from the perigee to the second layer of the calorimeter, respectively; $\Delta\phi_{res}$ is similar to $\Delta\phi$ but with the momentum of the track rescaled to the energy of the cluster. The asymmetry in the condition is to account for the energy loss due to bremsstrahlung where tracks with negative (positive) electric charge bend due to the magnetic field in the positive (negative) ϕ direction.

The matching of the fitted tracks to the candidate calorimeter seed-cluster is the final step of electron reconstruction. The matching requires $-0.10 < \Delta\phi < 0.05$, with the other alternative requirement remaining the same. If several tracks fulfil the matching criteria, the track considered to be the primary electron track is selected using an algorithm that takes into account the distance in η and ϕ between the extrapolated tracks and the cluster barycentres (again, measured in the second layer of the calorimeter), the number of hits in the silicon detectors and in the ID layer; a candidate with an associated track with at least four hits in the silicon layers and no association with a vertex from a photon conversion (photon converting into electron-positron pair) is considered as an electron candidate. However, if the primary candidate track can be matched to a secondary vertex and has no pixel hits, then this object is classified as a photon candidate (likely a conversion).

A further classification is performed using the candidate electron's E/p and p_{T}, the presence of a pixel hit, and the secondary-vertex information to determine unambiguously whether the object is only to be considered as an electron candidate or if it should be ambiguously classified as potentially either a photon candidate or an electron candidate.

4.2.2 Identification

The reconstruction algorithm is very efficient in reconstructing electrons, however, this is not necessarily what is needed for many ATLAS analyses, where they are interested in prompt electrons. Prompt electrons are electrons coming from the primary interaction of the event, while non-prompt electrons may come from the semileptonic decays of heavy quarks or from photon conversions. Other objects such as hadrons can be mis-reconstructed as electrons as well. This necessitates the identification of prompt electrons, making use of the differences between prompt electrons and non-prompt electrons/mis-reconstructed electrons. A multivariate likelihood technique, taking advantage of the correlations among the variables describing the differences, is employed to select prompt electrons [93]. The input variables to the likelihood describe the following characteristics: a) shower shape, b) properties of the track, c) matching of the track and clusters.

To quantify the performance of the identification, the identification efficiency is measured, which represents the probability of a true prompt electron is passing the identification requirements. Different cuts are applied to the final discriminant to define *working points* (WP), which can specify the identification efficiency, as a function of E_T or η of the electron. Three WPs, *Loose*, *Medium* and *Tight* are defined. Each WP is utilising a separate multivariate discriminant formed from a different selection of discriminating variables, and applying a different requirement on the resulting discriminant output. An example of the electron identification efficiency is shown in Figure 4.3.

In addition, many analyses further require the electron to pass some *isolation* requirements. Isolation is built exploiting a characteristic signature of the prompt electrons that there is relatively little activity surrounding the prompt electrons, as compared to non-

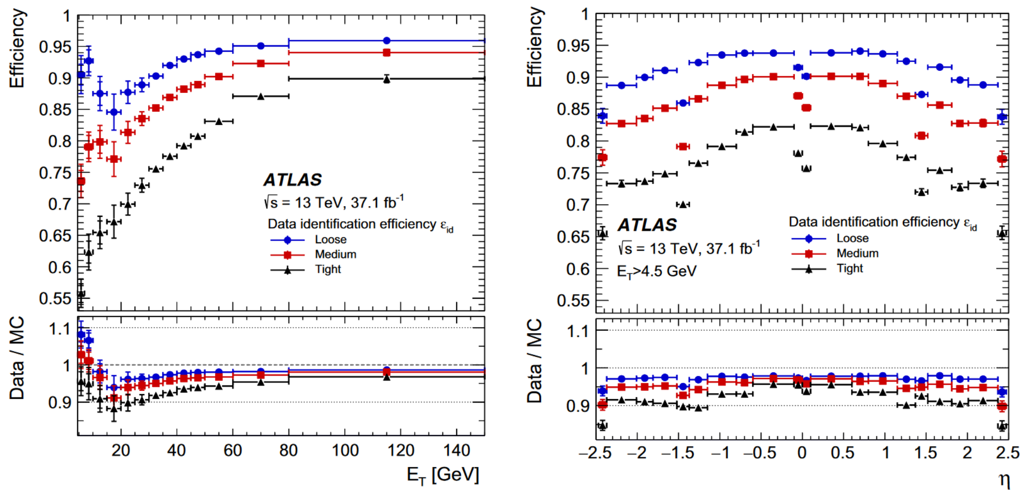


Figure 4.3: Measured electron identification efficiencies in $Z \to ee$ events for the 'Loose' (blue circles), 'Medium' (red squares), and 'Tight' (black triangles) WPs as a function of E_T (left) and η (right). The vertical uncertainty bars (barely visible because they are small) represent the statistical (ID) and total (outer bars) uncertainties. For both plots, the bottom panel shows the data-to-simulation ratios. Reproduced from Ref. [93].

prompt electrons [96]. For the electron isolation, four WPs: *Gradient*, *HighPtCaloOnly*, *Loose* and *Tight* are defined, each targeting a fixed value of isolation efficiency or imposing fixed requirements on the isolation variables.

In the analysis presented in Chapter 7 the 'Loose' identification WP is used, which, in combination with the additional track hit requirements provides an electron efficiency of 95%. The electron candidates are also required to have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.47$. Further cuts are used to define signal electrons, as defined in Section 7.2.2. In addition, the electron candidates are required to pass the 'Loose' isolation WP which has an efficiency of 99%. The isolation requirement is also inverted to provide control regions for estimating backgrounds, as described in Section 7.3.1. In the FTAG calibration effort presented in Chapter 6 the 'Medium' identification WP is used, where the electron efficiency is roughly 85%, with additional requirements on the electron $p_{\mathrm{T}} > 27$ GeV and $|\eta| < 2.47$ (while unlike requirements in Chapter 7, no further cuts are applied). Electrons are required to pass the 'Gradient' isolation WP, designed to give an efficiency of 90% at $p_{\mathrm{T}} = 25$ GeV and 99% at $p_{\mathrm{T}} = 60$ GeV.

4.3 Muons

The characteristic behaviour of a muon in the detector is a particle ionising minimally. The muon reconstruction is done taking advantage of this feature, with information from the ID and MS tracking detectors being combined, while information from the calorimeters is also used to account for cases of large energy loss in the calorimeters, and for MS-independent tagging of ID tracks as muon candidates in regions of limited MS coverage [97]. On the orther hand, the muon identification is performed by applying quality requirements that suppress background, mainly from pion and kaon decays, that muon candidates often leave a distinctive "kink" topology in the reconstructed track, and hence degrade the quality of the track.

4.3.1 Reconstruction

The muon reconstruction consists of two steps. The first step is to reconstruct the stand-alone track in the MS, followed by reconstruction with complete detector information. The first step of reconstruction starts with the identification of short straight-line local track segments reconstructed from hits in an individual MS *stations* (layers). At least two of these segments are then required to build preliminary track candidates, with information from precision measurements in the bending plane and measurements of the second coordinate from the MS trigger detectors to create three-dimensional track candidates. A global χ^2 fit of the muon trajectory through the magnetic field is performed, outlier hits are removed and hits along the trajectory that were not assigned to the original track candidate are added. Finally, the tracks are fitted again with the updated hits information, and ambiguities are resolved by removing tracks that share a large fraction of hits with higher-quality tracks.

The second step of reconstruction is to combine the track candidates with the complete information of all sub-detectors. The reconstruction proceeds according to five main reconstruction strategies, leading to the corresponding muon types:

- *Combined muons* are identified by matching MS tracks to ID tracks and performs a combined track fit based on the ID and MS hits, taking into account the energy loss in the calorimeters, in the region $|\eta| < 2.5$.

- *Inside out muons* are reconstructed in the region $|\eta| < 2.5$ using a complementary inside-out algorithm, which extrapolates ID tracks to the MS and searches for at least three loosely-aligned MS hits. The ID track, the energy loss in the calorimeters and the MS hits are then used in a combined track fit.

- *Muon-spectrometer extrapolated muons* are muons reconstructed when an MS track cannot be matched to an ID track. Its parameters are extrapolated to the beamline and used to define a Muon-spectrometer extrapolated muon. Such muons are used to extend the acceptance outside that of the ID ($|\eta| < 2.5$), thus exploiting the full MS coverage up to $|\eta| = 2.7$.

- *Segment-tagged muons* are identified by requiring that an ID track extrapolated to the MS satisfies the tight angular matching requirements to at least one reconstructed MS segment, in the $|\eta| < 2.5$ region.

- *Calorimeter-tagged muons* are identified by extrapolating ID tracks through the calorimeters to search for energy deposits consistent with a minimum-ionising particle, in the region $|\eta| < 0.1$.

4.3.2 Identification

As a consequence, it is expected that the fit quality of the resulting combined track will be poor and that the momentum measured in the ID and MS may not be compatible [98]. Therefore, a set of requirements are applied on the number of hits in the different ID subdetectors and different MS stations, on the track fit properties, and on variables that test the compatibility of the individual measurements in the two detector systems [97]. A given set of requirements for each of the muon types defined above is referred to as a working point (WP). The main metrics considered for designing the WPs are the selection efficiency and purity in simulation, where the prompt muon efficiency of a selection WP represents the probability that a real prompt muon traversing the detector is reconstructed as a muon and satisfies the WP; the purity of a selection WP is one minus the hadron misidentification rate (the fraction of light hadrons reconstructed as muons and satisfying the WP). Three standard selection WPs: *Loose*, *Medium*, and *Tight* are designed to cover the majority of physics analysis, while two additional WPs, *High-*p_T and *Low-*p_T, are designed to accommodate the analyses targeting extreme phase space. An example of the reconstruction and isolation efficiency of muon is shown in Figure 4.4.

In Chapter 7, muons are selected with $p_\mathrm{T} > 7$ GeV and $|\eta| < 2.7$, and passing the 'Loose' identification criteria with an efficiency of about 95%, as well as the 'pflowLoose_VarRadIso' isolation criteria [99] with an efficiency of roughly 97%. (further cuts are applied to define signal muons, as defined in Section 7.2.2); while in Chapter 6 muons are selected with $p_\mathrm{T} > 27$ GeV and $|\eta| < 2.5$, and passing the 'medium' identification criteria as well as a track-based isolation criteria, 'TightTrackOnly' with an efficiency of around 94% [97] (with no further cuts applied).

4.4 Jets

A jet can be defined as a collimated spray of particles arising from the fragmentation and hadronisation of a parton (quark or gluon) after a collision. Jets provide a link between the observed colourless stable particles and the underlying physics at the partonic level. A basic illustration of a collision of two protons, the subsequent particle shower and a reconstructed jet is shown in Figure 4.5.

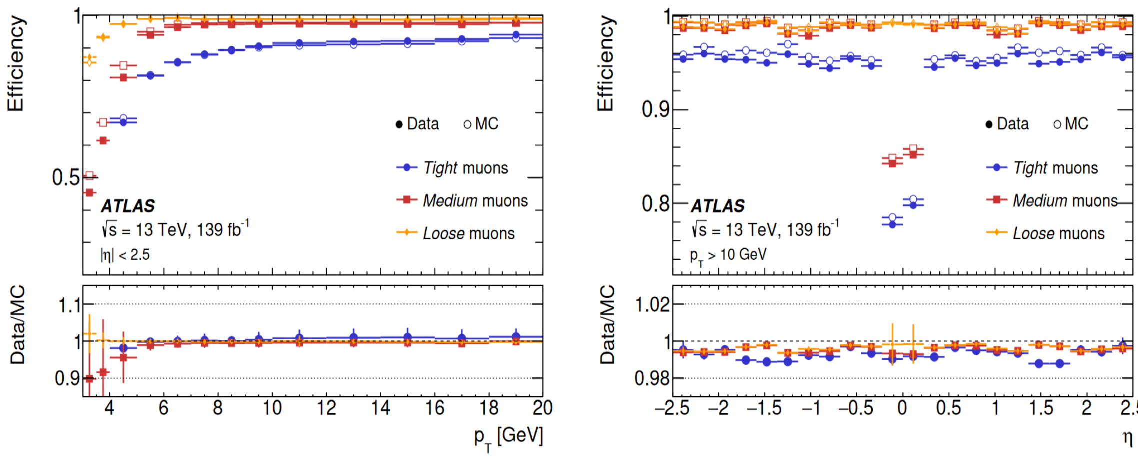


Figure 4.4: Muon reconstruction and identification efficiencies for the Loose, Medium, and Tight criteria. The left plot shows the efficiencies measured in $J/\psi \rightarrow \mu\mu$ events as a function of p_{T}. The right plot displays the efficiencies measured in $Z \rightarrow \mu\mu$ events as a function of η, for muons with $p_{\mathrm{T}} > 10$ GeV. The predicted efficiencies are depicted as open markers, while filled markers illustrate the result of the measurement in collision data. The statistical uncertainty in the efficiency measurement is smaller than the size of the markers, and thus not displayed. The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties. Reproduced from Ref. [97].

4.4.1 Reconstruction

The jet reconstruction starts by forming clusters of energy deposit in the calorimeters by performing a three-dimensional topological clustering of individual calorimeter cell signals [101]. The clustering begins with a seed cell and builds a cluster by iteratively adding neighbouring cells, providing these cells have significant energy relative to the expected noise. This algorithm clusters the energy deposits into so-called "topo-clusters" and combines their four-momenta. In Run 1 of the LHC, the ATLAS experiment used either solely the calorimeter or solely the tracker to reconstruct hadronic jets and soft particle activity, and the vast majority of analysis utilised jets that were built from topo-clusters, referred to *EMTopo jets*.

During Run-2, an alternative approach, called "*Particle flow*" (PFlow), became the default in ATLAS, as described in detail in Ref. [102]. It is also the default approach used in this thesis. Measurements from both the tracker and the calorimeter are combined to form the signals. The particle flow algorithm provides a list of tracks and a list of topo-clusters. Then, well-measured tracks are selected following a set of stringent quality criteria. The algorithm then attempts to match each track to a single topo-cluster in the calorimeter. The expected energy in the calorimeter, deposited by the particle that also created the track, is computed based on the topo-cluster position and the track momentum. As a result, a new set of clusters called *PFlow clusters* are produced, matching the topo-clusters to the

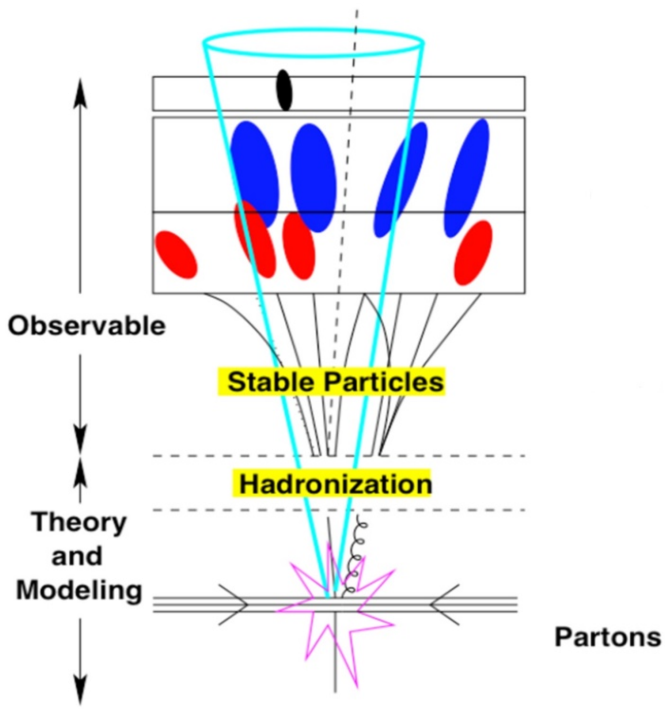


Figure 4.5: A simple example of an event showing the point of collision, the fragmentation and hadronization of the quarks and gluons and the resulting jet found through the detection of the stable particles. Image reproduced from Ref. [100].

particles that created the high-quality tracks.

Using the PFlow clusters as inputs, PFlow jets are then reconstructed using the anti-k_t algorithm [103]. The anti-k_t algorithm sequentially combines PFlow clusters into larger clusters based on the momentum-weighted distance between clusters. For two clusters i and j the algorithm defines:

$$d_{i,j} = min(p_{T,i}^{2p}, p_{T,j}^{2p})\frac{\Delta R_{i,j}^2}{R^2} \tag{4.1}$$

and

$$d_{i,beam} = p_{T,i}^{2p} \tag{4.2}$$

where p is a exponent parameter and the value of -1 is used for the anti-k_t algorithm (as the name 'anti' suggests), $\Delta R_{i,j}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, and $p_{T,i}$, η_i and ϕ_i are the transverse momentum, pseudorapidity and azimuthal coordinate of cluster i, respectively. The parameter R controls the size of the jet, and for standard jets in ATLAS this is chosen to be $R = 0.4$. The algorithm combines objects i, j with minimum value of $d_{i,j}$ into one object iteratively, until no more $d_{i,j}$ can be found smaller than $d_{i,beam}$. The combined object is considered the final jet candidate.

4.4.2 Calibration

The calibration of jets starts prior to the reconstruction. Since the energy of the calorimeter cells is measured at the electromagnetic scale, the local cluster weighting (LCW) calibration is applied to topo-clusters (before they are used to form jets) to account for the differences in the detector response to hadronic and electromagnetic showers [104].

After the jets are reconstructed, the four-momenta of jets are calibrated with the jet energy scale (JES) calibration, which consists of several consecutive stages derived from a combination of MC-based methods and in-situ techniques [105]. MC-based calibrations correct the reconstructed jet four-momentum to that found from the simulated stable particles within the jet, and in-situ techniques are used to measure the difference in jet response between data and simulation, with residual corrections applied to jets in data only. The calibrations account for features of the detector, the jet reconstruction algorithm, jet fragmentation, the busy data-taking environment resulting from multiple pp interactions, and the difference in jet response between data and simulation.

In order to calibrate the jet energy resolution (JER), the jet momentum must be measured precisely. As described in detail in Ref. [106], a *dijet balance* approach is used for this purpose, based on a well-defined dijet system, where the two jets are expected to have p_{T} that sum up to zero precisely.

Furthermore, jets arising from pileup are suppressed by using the *jet vertex tagger*, which is a multivariate combination of track-based variables developed to separate hard-scatter jets from pileup jets [107].

A jet cleaning selection is applied in order to veto any ‘fake’ jets, which arise from non-collision background events, such as cosmic rays, or from detector effects. Finally, all jets in the analysis are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$.

4.4.3 Identification of heavy-quark flavoured jets

Known as *flavour tagging*, the identification of jets containing b-hadrons (b-jets) against the large background of jets containing c-hadrons (c-jets) or jets coming from the hadronization of light (u,d,s) quarks or gluons (light jets) is of major importance in many areas of the physics programme of the ATLAS experiment at the LHC. It is crucial in a large number of SM precision measurements, studies of the Higgs boson properties, and searches for new phenomena [108–110]; it also plays an important role in the $HH \to bb\tau\tau$ searches presented in Chapter 7.

The ATLAS Collaboration uses various algorithms to identify b-jets [111], referred to as b-tagging algorithms, when analysing data recorded during Run 2 of the LHC. These algorithms exploit the long lifetime, high mass and high decay multiplicity of b-hadrons, as well as the properties of the b quark fragmentation. Given a lifetime of the order of 1.5 ps, b-hadrons have a significant mean flight length ($\langle c\tau \rangle \approx 450\ \mu m$), in the detector before decaying, generally leading to at least one vertex displaced from the hard-scatter collision point, as illustrated in Figure 4.6.

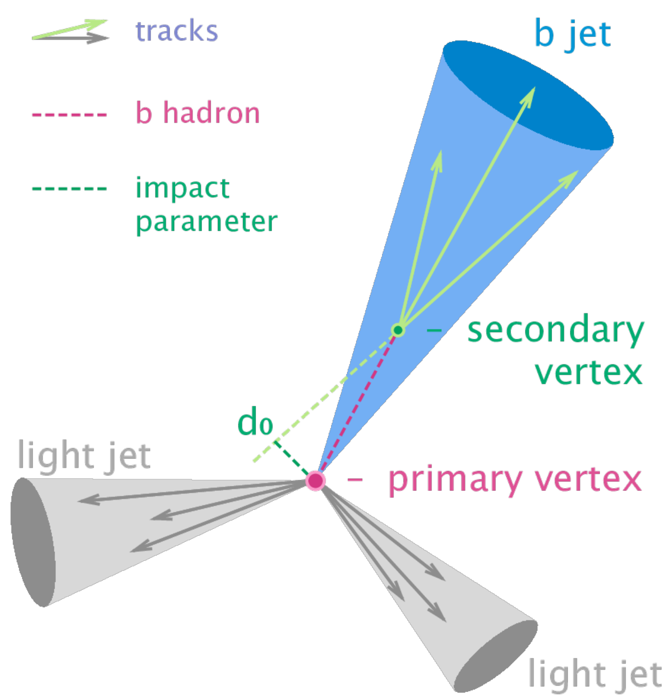


Figure 4.6: A diagram showning the b hadron decay initiated jets.

The strategy developed by the ATLAS Collaboration is based on a two-stage approach. Firstly, low-level algorithms reconstruct the characteristic features of the b-jets via two complementary approaches. A first approach, implemented in the IP2D and IP3D algorithms, is inclusive and based on exploiting the large impact parameters of the tracks originating from the b-hadron decay [112]. The second approach explicitly reconstructs displaced vertices. The SV1 algorithm [113], attempts to reconstruct an inclusive secondary vertex, while the JetFitter algorithm [114], aims to reconstruct the full b- to c-hadron decay chain. These algorithms, first introduced during Run 1 [111], have been improved and retuned for Run 2 [115]. Secondly, in order to maximise the b-tagging performance, the results of the low-level b-tagging algorithms are combined into high-level algorithms via multivariate classifiers.

The most performant algorithms presently in use in physics analyses at ATLAS are based on multivariate combinations of the available information (MV2) or additionally using a deep feed-forward neural network (DL1) [112, 116]. This is illustrated in Figure 4.7, where the performance is characterised by the probability of tagging a b-jet (b-jet tagging efficiency, ϵ_b) and the probability of mistakenly identifying a c-jet (light-flavour jet) as a

b-jet labelled $\epsilon_c(\epsilon_l)$.

The distributions of the output discriminant of the MV2 and DL1 tagger for b-jets, c-jets, and light-flavour jets in simulated $t\bar{t}$ events are shown in Figure 4.8. The evaluation of the performance of the algorithms is carried out using b-jet tagging single-cut operating points (OPs, or working points, WP). These are based on a fixed selection requirement on the b-tagging algorithm output discriminant distribution ensuring a specific b-jet tagging efficiency, for the b-jets present in simulated $t\bar{t}$ sample. The discriminant distributions are also divided into five 'pseudo-continuous' bins, delimited by the selections used to define the b-jet tagging single-cut WPs for 85%, 77%, 70% and 60% efficiency, and bounded by the trivial 100% and 0% selections.

Depending on the low-level algorithm, the DL1 tagger can be further separated into two taggers: DL1 and DL1r, where the DL1 tagger uses traditional track-based impact parameter taggers IP2D and IP3D [117] and the DL1r tagger uses a Recurrent Neural Network Impact Parameter tagger (RNNIP) [112].

The calibration of DL1 and DL1r algorithms has been an original contribution of the author of this thesis, and is described in more detail in Chapter 6. The DL1r tagger is now the default b-tagging algorithm used for flavour tagging in ATLAS and is utilised in the analysis presented in this thesis.

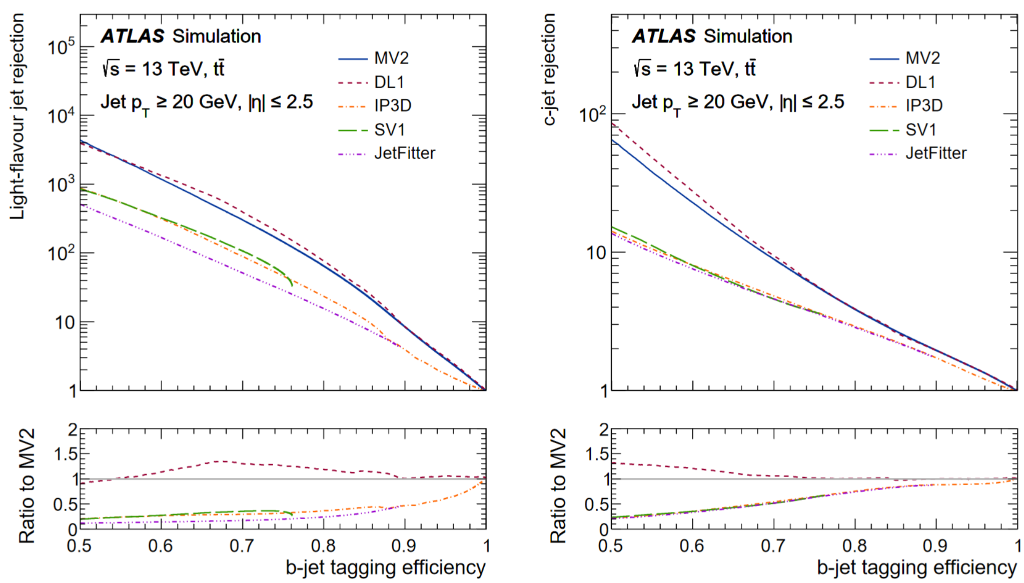


Figure 4.7: The light-flavour jet (left) and c-jet (right) rejections versus the b-jet tagging efficiency for the IP3D, SV1, JetFitter, MV2 and DL1 b-tagging algorithms evaluated on $t\bar{t}$ events [115].

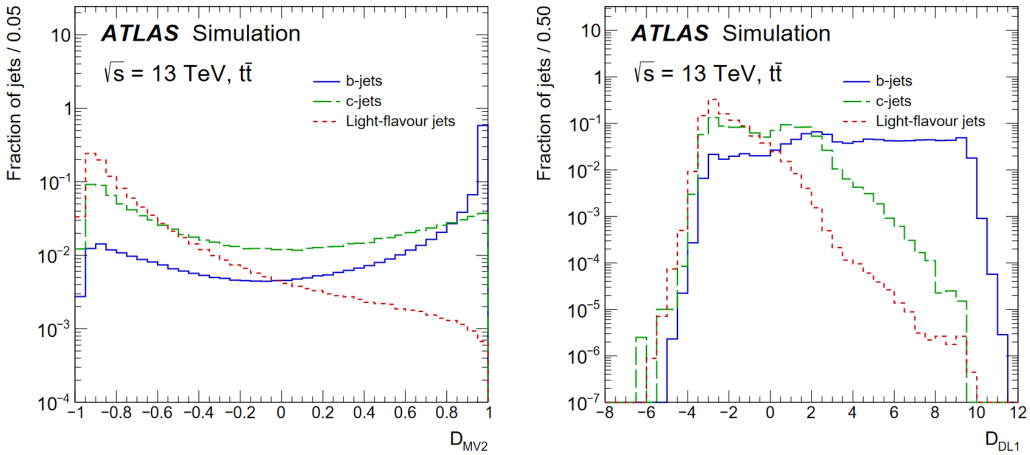


Figure 4.8: The fraction of light-flavour jets and c-jets versus the b-jets in the MV2 (left) and DL1 (right) b-tagging algorithms output distribution evaluated on $t\bar{t}$ events [115].

4.5 Hadronically decaying τ lepton

With a mass of 1.777 GeV and a proper decay length of 87 μm [10], tau leptons decay either leptonically ($\tau_{\text{lep}} \rightarrow \ell\nu_{\ell}\nu_{\tau}$, $\ell = e, \mu$) or hadronically ($\tau_{\text{had}} \rightarrow$ hadrons ν_{τ}) and do so typically before reaching active regions of the ATLAS detector. The leptonically decaying τ_{lep} is simply reconstructed as either an electron or muon, with the neutrinos contributing to the real component of the $E_{\text{T}}^{\text{miss}}$. On the other hand, the hadronically decaying τ_{had}can be identified via their decay products. The hadronic tau lepton decays represent 65% of all possible decay modes [10]. In these decay modes, the hadronic decay products are one or three charged pions in 72% and 22% of all cases, respectively. Charged kaons are present in the majority of the remaining hadronic decays. In 78% of all hadronic decays, up to one associated neutral pion is also produced. This results in an experimental signature of a collimated calorimeter shower with either one or three associated tracks (*prongs*). The neutral and charged hadrons stemming from the tau lepton decay make up the visible decay products of the tau lepton, and are in the following referred to as τ_{had}.

4.5.1 Reconstruction

The τ_{had} candidates are seeded by jets formed using the anti-k_t algorithm, with a jet-size parameter $R = 0.4$. For events with multiple interactions, the chosen primary vertex may not be the one where the tau lepton originated. Here the *tau vertex association* algorithm is used with all tau candidate tracks within $\Delta R < 0.2$ around the jet seed direction as input. The p_{T} of these tracks is summed and the primary vertex candidate to which the largest fraction of the p_{T} sum is matched to is chosen as the tau vertex [107].

Tracks are associated with the τ_{had} if they are in the *core* region $\Delta R < 0.2$ around the

Working point	Signal efficiency 1-prong	3-prong	Background rejection BDT 1-prong	3-prong	Background rejection RNN 1-prong	3-prong
Tight	60%	45%	40	400	70	700
Medium	75%	60%	20	150	35	240
Loose	85%	75%	12	61	21	90
Very loose	95%	95%	5.3	11.2	9.9	16

Table 4.1: List of defined working points with fixed true τ_{had} selection efficiencies and the corresponding background rejection factors for misidentified τ_{had} in dijet events for the BDT and RNN classifiers. The BDT was used for a previous analysis and is only shown here for comparison purposes. Table reproduced from Ref. [49].

τ_{had} direction and satisfy the following criteria: $p_{\text{T}} > 1$ GeV, at least two associated hits in the pixel layers of the inner detector, and at least seven hits in total in the pixel and the SCT layers. Furthermore, requirements are imposed on the distance of closest approach of the track to the vertex in the transverse plane, $|d_0| < 1.0$ mm, and longitudinally, $|z_0 sin\theta|$ < 1.5 mm. Tracks in the *isolation* region $0.2 < \Delta R < 0.4$ are used for the calculation of identification variables and are required to satisfy the same selection criteria. The number of tracks (prongs) is susceptible to underestimation due to tracking inefficiency, or overestimation due to tracks from photon conversions passing the track selection criteria.

4.5.2 Identification

The τ_{had} reconstruction algorithm alone provides no discrimination against other particles that result in jet-like signatures in the detector. Therefore, dedicated algorithms are used to identify hadronic tau lepton decays. Here, a recurrent neural network (RNN) classifier is used as described in Ref. [49]. The RNN uses a combination of low-level input variables for individual tracks and clusters that are associated to the τ_{had} candidate as well as several high-level observables calculated from track and calorimeter quantities. Due to the distinct signatures of 1- and 3-prong τ_{had}decays, the τ_{had} identification is split into dedicated algorithms for 1- and 3-prong τ_{had}. Four working points with increasing background rejection (*Very loose*, *Loose*, *Medium* and *Tight*) are defined to be used by physics analyses. The corresponding signal selection efficiencies and rejection powers are given in Table 4.1.

In the previous analysis using the 36.1 fb^{-1} data [27], a different τ_{had} identification (tau-ID) algorithm was used, which used a similar set of input variables to the RNN tau-ID, such as the invariant mass of the track system, the fraction of energy deposited in the central region. This algorithm is based on a boosted decision tree (BDT tau-ID) [118]. The RNN tau-ID shows better performance and allows moving to a looser WP with increased efficiency (about 24% and 11% in case of two $\tau_{\text{had}}$and one τ_{had}in the final state, respectively) and without losing jet rejection rate, as shown in Figure 4.9.

Selected τ_{had} candidates in the analysis are required to have $p_{\text{T}} > 20$ GeV, $|\eta| < 2.5$, with candidates in the barrel-endcap transition region of the calorimeter ($1.37 < |\eta| < 1.52$) vetoed due to poor detector instrumentation in this region, one or three tracks, unit charge, and to pass the Loose τ_{had}–ID working point. The Loose WP corresponds to 85% efficiency for 1-prong and 75% efficiency for 3-prong (the efficiency is flat in p_{T} by definition).

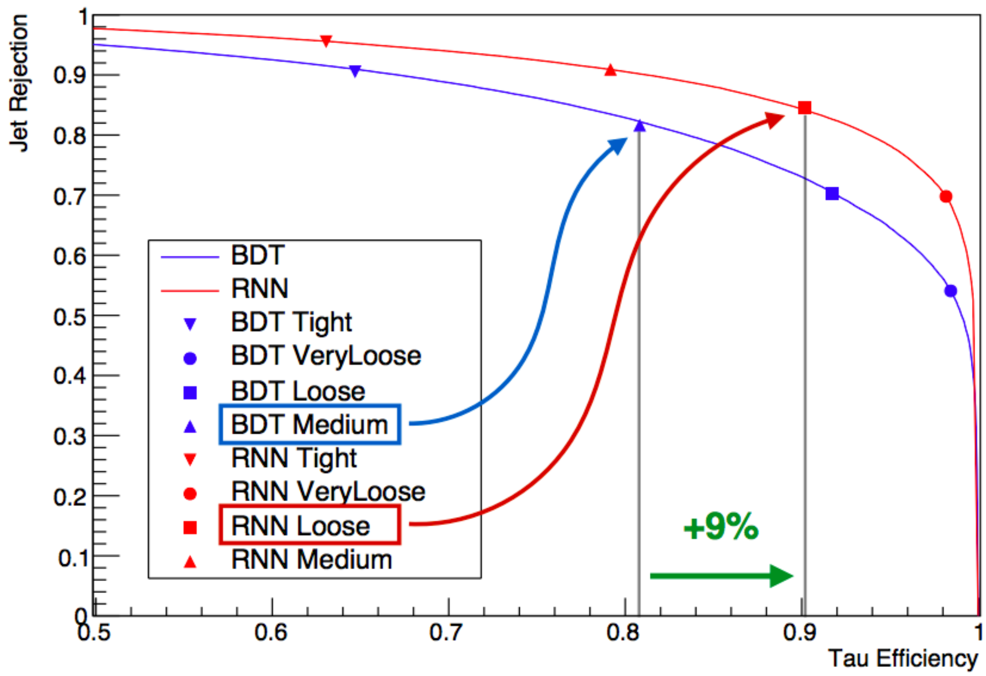


Figure 4.9: Jet rejection and tau efficiency of tau candidate, measured in $\gamma^* \to \tau\tau$ sample for RNN-ID [49] and $Z/\gamma^* \to \tau\tau$ sample for BDT-ID [119]. Jet rejection represents the probability of a jet not originating from a tau lepton being rejected by the identification algorithm. The green arrow indicates the increase in tau efficiency. Figure reproduced from the analysis internal notes.

Additional rejection of τ_{had} candidates originating from electrons is provided by a BDT employing track and shower shape information. The Loose working point is used for the analysis presented in Chapter 7, corresponding to a selection efficiency of about 95% efficiency for true τ_{had} [120].

4.6 Missing transverse energy

As defined in Section 3.7, the $\vec{E}_{\text{T}}^{\text{miss}}$ is given as $\vec{E}_{\text{T}}^{\text{miss}} = -\sum_i \vec{p_{T_i}}$, which is the negative vector sum of transverse momentum collected from the detector, from which one or more "invisible" particle(s) can be inferred. The reconstruction of $\vec{E}_{\text{T}}^{\text{miss}}$ is comprised of two contributions [121], within the region of $|\eta| < 4.9$. The first one is from *hard-event* signals, combining information from fully reconstructed and calibrated physics objects, i.e. electrons, muons, photons, jets, hadronically decaying τ-leptons and jets. The second one is from the *soft-event* signals, consisting of reconstructed charged-particle tracks associated with the hard-scatter vertex but with no physics objects. Hence, $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as:

$$E_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{\substack{\text{selected}\\ \text{electrons}}} p_T^e - \sum_{\substack{\text{accepted}\\ \text{photons}}} p_T^\gamma - \sum_{\substack{\text{accepted}\\ \tau\text{-leptons}}} p_T^{\tau_{\mathrm{had}}} - \sum_{\substack{\text{selected}\\ \text{muons}}} p_T^\mu - \sum_{\substack{\text{accepted}\\ \text{jets}}} p_T^{\mathrm{jet}} - \sum_{\substack{\text{unused}\\ \text{tracks}}} p_T^{\mathrm{track}} \tag{4.3}$$

.

A non-zero $E_{\mathrm{T}}^{\mathrm{miss}}$ suggests the existence of a non-interacting particle. In the SM, the only particle that does not deposit energy in the detector are neutrinos, as they are weakly interacting leptons which do not undergo strong or electromagnetic forces. Nevertheless, some BSM theories predict the existence of additional weakly-interacting particles.

4.7 Overlap removal

After the event is reconstructed, an overlap-removal procedure is applied to resolve ambiguities when a physical object is reconstructed as multiple particles in the ATLAS detector. The angular distance ΔR is used to measure the overlap of two reconstructed objects. Overlaps between most of the detector objects used in the analysis are resolved by using the standard overlap removal tools AssociationUtils [122], with analysis-specific procedures for the reconstructed τ_{had}, anti-τ_{had} objects and jets. The step-by-step procedure that is used to resolve ambiguities in the reconstructed objects is summarised in the following:

- e_1–e_2: For two electrons e_1 and e_2 in an event, reject e_1 if both electrons share the track and $p_{T1} < p_{T2}$
- τ_{had}-e: Reject τ_{had} if $\Delta R < 0.2$
- τ_{had}-μ: Reject τ_{had} if $\Delta R < 0.2$:
  Case 1 (τ_{had} $p_T > 50$GeV): $p_T, \mu > 2$GeV and combined muon
  Case 2 (τ_{had} $p_T \leq 50$GeV): $p_T, \mu > 2$GeV
- μ–e: Reject μ if calo-muon and shared ID track
- e–μ: Reject e if shared ID track
- jet–e: Reject jet if $\Delta R < 0.2$
- e–jet: Reject e if $\Delta R < 0.4$
- jet–μ: Reject jet if $N_{track} < 3$ ($p_T^{track} > 500$MeV), and $\Delta R < 0.2$

- μ–jet: Reject μ if $\Delta R < 0.4$

Additionally, an analysis-specific overlap-removal procedure for τ_{had} anti-τ_{had} and jets is implemented:

- jet–τ_{had}: Reject jet if $\Delta R < 0.2$
- anti–τ_{had}–jet: Reject anti–τ_{had}if jet is b-tagged and $\Delta R < 0.2$
- jet–anti–τ_{had}: Reject jet if $\Delta R < 0.2$

This establishes the following priority: τ_{had} > b-tagged jet > anti-τ_{had} > un-tagged jet.

Chapter 5

Data and Monte Carlo Simulation

5.1 Simulation of physics processes

A precise and reliable theoretical prediction is the key to interpreting and analysing the data recorded by the ATLAS experiment. It enables the quantification of agreement between the data and the SM, and test of possible new physics beyond the SM. In this thesis, the signal and background processes are simulated by Monte Carlo (MC) event generators. The full simulation process is comprised of a few consecutive steps: starting with the simulation of the hard-scattering process, followed by the simulation of the parton-showering and finally the simulation of the interaction of the particles with the detector and the response of the detector.

The hard-scattering process happens when two partons which carry a fraction of the protons' momentum collide inelastically. The partons can be valence quarks, which are the quarks or anti-quarks that determine the quantum numbers of the proton or gluons which mediate the strong force, or sea quarks which are virtual quark-anti-quark pairs that are created and annihilated promptly. The hard-scatter process is typically characterised by large momentum transfer, and the probability of a parton carrying a given fraction of the total proton momentum is described by the *parton distribution function* (PDF). In the MC simulation, the hard-scattering process is modelled by the *matrix element*, using leading order (LO) or next-to-leading order (NLO) Feynman diagrams. It can also be simulated at next-to-next-to-leading order (NNLO) for better theoretical approximation.

After the hard-scattering process, each parton will undergo the showering process where it hadronises or radiates further partons. This process is simulated by dedicated algorithms. The hadronisation process happens mostly in low-energy regimes, which is non-perturbative and requires phenomenological modelling exploiting specific hadronisation models. On the other hand, the radiation process happens at higher energy, and it stops when the parton loses enough energy to reach the confinement energy-scale. The hard-scattering and showering processes of partons are described by the combination of the matrix element generator and the parton shower algorithms.

In addition to the hard-scattering and showering processes, interaction can happen prior or after them, as referred to initial state radiation (ISR) or final state radiation (FSR). The energy scale of these additional processes is typically a few GeV, much smaller compared to the hard-scattering energy scale. The radiated gluons or photons, together with other particles originating from soft scattering interactions, are described as the *underlying event*.

In addition, pileup events, as described in Section 3.1.2, need to be simulated. They are separate collisions and are simulated separately from the hard-scattering event. However, in the reconstruction of the event, they cannot be separated from the hard-scattering event, therefore they are later overlaid on the hard-scattering events.

A variety of MC generator programs were developed. Some of them are called multi-purpose generators and can simulate a full event on their own, while some others are dedicated only to hard-scattering or parton shower and need to be used combined with other generators. After simulation, simulated events are passed to Geant 4 [123, 124], a software package for simulating the interactions of particles with matter.

Finally, event generators require *tuning* to match data. The tuning parameters are based on phenomenological models and are applied on hadronisation simulation and underlying event.

5.2 Data samples

The data used in this search were collected at a centre-of-mass energy of 13 TeV between 2015 and 2018. For the c-jet mis-tagging efficiency calibration (FTAG calibration) presented in Chapter 6, the data sample was collected using a set of single-electron [125] and single-muon triggers [126]. Requirements on p_{T} over a range of 24–30 GeV are applied to the single-electron triggers, with additional quality and isolation requirements depending on the p_{T} threshold and the data-taking period. While for the single-muon triggers, requirements on p_{T} over a range 20–26 GeV are applied on the isolated muons and a tighter cut of 50 GeV is applied for muons without any isolation requirement.

For the HH searches presented in Chapter 7, single-lepton triggers and lepton-plus-τ_{had} triggers are used, as discussed in detail in Section 7.2.

Events are selected for analysis only if they are of good quality and if all the relevant detector components are known to be in operating conditions [127]. The total integrated luminosity of the data, after meeting the good quality criteria, is $139.0 \pm 2.4\ \mathrm{fb}^{-1}$ [128, 129]. The recorded events contain an average of 34 simultaneous inelastic pp collisions per bunch-crossing.

5.3 Simulated event samples

As mentioned in Section 5.1, the signal and background processes are modelled by Monte Carlo simulation. No signal sample is defined for the FTAG calibration. The signal targeted in the HH searches includes the SM-like non-resonant HH production via ggF and VBF, and the BSM resonant HH production. To simulate these processes, the full ATLAS detector simulation [124] (FS) is applied, including the GEANT 4 package. The only exception is the resonant signal samples, where a fast simulation (AF2) [124] is used, that instead of fully simulating the response of the calorimeters, uses pre-simulated showers to save computation time. For unstable hadrons i.e. b- and c-hadrons, the decay process is simulated by the EVTGEN v1.6.0 package[130], with the exception of the VBF non-resonant samples and samples generated by SHERPA [131]. The resulting events were then processed through the same reconstruction programs as the data. To simulate the pileup effects, minimum bias events are overlaid on the simulation, exploiting the PYTHIA 8.186 generator [132] for soft QCD processes using the A3 tune [133]. The NNPDF2.3LO [134] PDFs are used. In addition, the Higgs boson mass was fixed to 125 GeV for all simulated samples that contain this particle. The mass of the Higgs boson is assumed to be 125 GeV for all samples containing an SM Higgs boson. This value is also used for calculating the single- and pair-production cross-sections of the Higgs boson, as well as the decay branching ratio of the Higgs boson.

A summary of the event samples used for the simulation of the signal and background processes is shown in Table 5.1.

Process	ME generator	ME QCD	PDF	PS	Tuned parameters	Cross-section
Signal						
non-resonant $gg \to HH$ (ggF)	Powheg-Box v2	NLO	PDF4LHC15 NLO	Pythia 8.244	A14	NNLO FTApprox
non-resonant $qq \to qqHH$ (VBF)	MadGraph5_aMC@NLO v2.7.3	LO	NNPDF3.0NLO	Pythia 8.244	A14	N3LO(QCD)
resonant $gg \to X \to HH$	MadGraph5_aMC@NLO v2.6.1	LO	NNPDF2.3LO	Herwig v7.1.3	H7.1-Default	–
Top quark						
$t\bar{t}$	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.230	A14	NNLO+NNLL
single top(t-, s-, Wt-channels)	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.230	A14	NLO
$t\bar{t}Z$	Sherpa 2.2.1	NLO	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NLO
$t\bar{t}W$	Sherpa 2.2.8	NLO	NNPDF3.0NNLO	Sherpa 2.2.8	Default	NLO
Single Higgs boson						
ggF	Powheg-Box v2	NNLO	NNPDF3.0NLO	Pythia 8.212	AZNLO	N3LO(QCD)+NLO(EW)
VBF	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.212	AZNLO	NNLO(QCD)+NLO(EW)
$qq \to WH, ZH$	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.212	AZNLO	NNLO(QCD)+NLO(EW)
$gg \to ZH$	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.212	AZNLO	NLO+NLL
$t\bar{t}H$	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.230	A14	NLO
Vector boson + jets						
W/Z+jets	Sherpa 2.2.1	NLO ($\leq$ 2 jets), LO (3,4 jets)	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NNLO
Diboson						
WW, WZ, ZZ	Sherpa 2.2.1	NLO ($\leq$ 1 jet), LO (2,3 jets)	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NLO

Table 5.1: The generators used for the simulation of the signal and background processes. The order of the cross-section calculation refers to the expansion in the strong coupling constant (α_S). The acronyms ME, PS and UE are used for matrix element, parton shower and underlying event, respectively. The terms ggF, VBF refer to gluon-gluon fusion and vector-boson fusion respectively. The cross-section for resonant production is not shown as a nominal cross-section of 1 pb was chosen. Reproduced from Ref. [135].

5.3.1 Simulated signal samples

In the HH searches, contributions from both the ggF and VBF processes to the SM non-resonant HH signal production are included, each simulated with different generators and PDFs. The expansion order of Feynman diagrams is also different. For the resonant HH signal, only the ggF contribution is considered. It was simulated for 20 values of the resonance mass, m_X, between 251 GeV and 1600 GeV (251, 260, 280, 300, 325, 350, 375, 400, 450, 500, 550, 600, 700, 800, 900, 1000, 1100, 1200, 1400 and 1600).

The ggF events were generated with the Powheg-Box v2 generator [136] at NLO with finite top quark mass, using the PDF4LHC15 NLO PDF set [137]. Parton showers and hadronisation were interfaced to Pythia 8.244 [132] with the A14 set of tuned parameters [138, 139] and the NNPDF2.3LO PDF set.

On the other hand, the VBF events generated using the MadGraph5_aMC@NLO v2.7.3 [140] generator at LO with the NNPDF3.0NLO [141] PDF set. Parton showering and hadronisation were simulated using Pythia 8.244 with the A14 tune and the NNPDF2.3LO PDF set. To simulate the decays of the b- and c-hadrons, the EvtGen v1.7.0 program was used.

Finally, the resonant signal of a heavy spin-0 narrow width resonance via ggF production was simulated with the MadGraph5_aMC@NLO v2.6.1 generator using the NNPDF2.3LO PDF set at LO. The parton shower and hadronisation were simulated to Herwig 7.1.3 [142, 143], using the H7.1-Default tune [144] and the NNPDF2.3LO PDF set.

5.3.2 Non-resonant signal reweighting and combination

In the BSM scenarios, the non-resonant di-Higgs production is sensitive to the self-coupling constant and other possible anomalous couplings, as described in Section 2.2.8. In this thesis, a reweighting method is used to evaluate the non-resonant di-Higgs production with a range of possible values of self-coupling modifier, $\kappa_\lambda \equiv c_{hhh}$, two Wilson coefficients c_{gghh}, c_{tthh} and 7 benchmark models of a set of five HEFT couplings defined in Table 2.2.

κ_λ Reweighting For the ggF non-resonant HH production, MC samples are generated at $\kappa_\lambda = 1$ and 10, while for VBF production, MC samples are generated at $\kappa_\lambda = 0$, 1, 2 and 10. A sample combination technique is used to model the signal hypothesis at different κ_λ values. For the ggF di-Higgs production, a reweighting method described in Ref. [145] is used to obtain predictions at different κ_λ values in the range $\kappa_\lambda \in [-30, 30]$ in increments of 0.2 based on a linear combination of generator samples at $\kappa_\lambda = 0, 1$ and 20. The remaining $\kappa_\lambda = 10$ sample is used to validate the method. For each κ_λ value, a set of weights $w(m_{HH}, \kappa_\lambda)$ is evaluated by dividing the binned m_{HH} distribution of the κ_λ

target sample by the SM distribution. They can be used to reweight the SM non-resonant sample to any κ_λ value, which is performed at the analysis level–after reconstruction and the selection steps defined in Section 7.2.2. Given the assumption that the kinematics of the ggF events and their acceptance depends only on the m_{HH} variable, using the weights $w(m_{HH}, \kappa_\lambda)$, the reweighted sample describes correctly any kinematics distribution of a given target κ_λ value. Good closure is found between the distributions obtained from κ_λ generated and reweighted, as shown in Section 7.5.3.3.

For the VBF di-Higgs production, the above reweighting procedure is not valid, because the kinematic of the events can not be defined just using a single variable such as m_{HH}. Instead, three fully-reconstructed MC samples with $\kappa_\lambda = 1, 2, 10$ are used. The event distributions and the multivariate algorithm output (more details in Section 7.4) for any κ_λ value are obtained from the linear combination of the corresponding distributions of the three samples at the analysis level. As defined in Section 2.2.8, the full cross-section for the VBF HH production involves three diagrams, and expanding the absolute squared of the amplitude yields six terms:

$$\sigma = \kappa_V^2\kappa_\lambda^2 a_1 + \kappa_V^4 a_2 + \kappa_{2V}^2 a_3 + \kappa_V^3\kappa_\lambda a_4 + \kappa_V\kappa_\lambda\kappa_{2V} a_5 + \kappa_V^2\kappa_{2V} a_6. \tag{5.1}$$

In the case of a κ_λ scan, this formula is reduced to

$$\sigma = \kappa_\lambda^2 a_1 + \kappa_\lambda a_2 + a_3. \tag{5.2}$$

Using the basis of $\kappa_\lambda = 1, 2, 10$, the linear coefficients for combining the three samples defined are then given by:

$$\sigma(\kappa_\lambda) = \left(\frac{\kappa_\lambda^2}{9} - \frac{4\kappa_\lambda}{3} + \frac{20}{9}\right) \times \sigma(1) + \left(-\frac{\kappa_\lambda^2}{8} + \frac{11\kappa_\lambda}{8} - \frac{5}{4}\right) \times \sigma(2)$$

$$+ \left(\frac{\kappa_\lambda^2}{72} - \frac{\kappa_\lambda}{24} + \frac{1}{36}\right) \times \sigma(10) \tag{5.3}$$

HEFT interpretation signals reweighting It is extremely computationally expensive to simulate all of the possible variations of the Wilson coefficients described by the HEFT. Therefore, a similar reweighting production is applied to derive the signal samples assuming the 7 BM, and varying the two Wilson coefficients c_{gghh}, c_{tthh} (while the other HEFT couplings take their SM values). For simplicity, only the ggF non-resonant production is considered.

The reweighting is performed with weights in the m_{HH} distribution. As derived in

Ref. [146], the NLO differential cross-section for the HH production via gluon-gluon fusion $\mathrm{d}\sigma_{HH}/\mathrm{d}m_{HH}$ is parametrised in terms of coupling combinations as Eq. (5.4):

$$\begin{aligned}\frac{\mathrm{d}\sigma_{HH}}{\mathrm{d}m_{HH}} &= \mathrm{Poly}(\mathbf{A}, c_{hhh}, c_{tth}, c_{tthh}, c_{ggh}, c_{gghh}|m_{HH}) \\ &= A_1 c_{tth}^4 + A_2 c_{tthh}^2 + (A_3 c_{tth}^2 + A_4 c_{ggh}^2) \cdot c_{hhh}^2 + A_5 c_{gghh}^2 + (A_6 c_{tthh} + A_7 c_{tth} c_{hhh}) \cdot c_{hhh}^2 \\ &+ (A_8 c_{tth} c_{hhh} + A_9 c_{ggh} c_{hhh}) \cdot c_{tthh} + A_{10} c_{tthh} c_{gghh} + (A_{11} c_{ggh} c_{hhh} + A_{12} c_{gghh}) \cdot c_{tth}^2 \\ &+ (A_{13} c_{hhh} c_{ggh} + A_{14} c_{gghh}) \cdot c_{tth} c_{hhh} + A_{15} c_{ggh} c_{gghh} c_{hhh} + A_{16} c_{tth}^3 c_{ggh} \\ &+ A_{17} c_{tth} c_{tthh} c_{ggh} + A_{18} c_{tth} c_{ggh}^2 c_{hhh} + A_{19} c_{tth} c_{ggh} c_{gghh} + A_{20} c_{tth}^2 c_{ggh}^2 \\ &+ A_{21} c_{tthh} c_{ggh}^2 + A_{22} c_{ggh}^3 c_{hhh} + A_{23} c_{ggh}^2 c_{gghh}. \end{aligned} \tag{5.4}$$

where the coefficients A_i are evaluated in the $240\,\mathrm{GeV} < m_{HH} < 1040\,\mathrm{GeV}$ range in bins of 20 GeV. The binned m_{HH} distribution from simulated SM events can be reweighted to any point in the HEFT space using the event weights defined in Eq. (5.5).

$$w(m_{HH}) = \frac{\mathrm{d}\sigma_{HH}/\mathrm{d}m_{HH}}{\mathrm{d}\sigma_{HH}^{\mathrm{SM}}/\mathrm{d}m_{HH}} = \frac{\mathrm{Poly}(\mathbf{A}|m_{HH})}{\mathrm{Poly}^{\mathrm{SM}}(\mathbf{A}|m_{HH})} \tag{5.5}$$

The weight $w(m_{HH})$ encapsulates both shape and overall normalisation with respect to the simulated SM HH sample. In Section 7.6, upper limits are set on the production cross-section assuming the 7 BM. Therefore it is preferred to normalise these BSM signals to the same cross-sections, so that the signal strength is with respect to the same standard. To remove the normalisation effect, normalised weights $w^{\mathrm{norm}}(m_{HH})$ are defined in Eq. (5.6), where the cross-section is normalised to the SM cross-section value.

$$w^{\mathrm{norm}}(m_{HH}) = w(m_{HH}) \cdot \frac{\sigma_{HH}^{\mathrm{SM}}}{\sigma_{HH}} = \frac{\mathrm{Poly}(\mathbf{A}|m_{HH})}{\mathrm{Poly}^{\mathrm{SM}}(\mathbf{A}|m_{HH})} \cdot \frac{\mathrm{Poly}^{\mathrm{SM}}(\mathbf{A})}{\mathrm{Poly}(\mathbf{A})} \tag{5.6}$$

In this thesis, the BSM processes are reweighted from simulated SM events using the event-per-event weight $w^{\mathrm{norm}}(m_{HH})$. Events with $m_{HH} > 1040$ GeV are reweighted using the weight corresponding to the highest available m_{HH} bin with a central value of 1030 GeV. Systematic uncertainties stemming from reweighting are derived at truth level (directly from events without detector response simulation) to cover possible discrepancies between the reweighted distributions and simulated distributions, as described in Section 7.5.3.4.

5.3.3 Simulated background samples

The major background processes for the HH searches are the $t\bar{t}$, single-top, a boson produced in association with jets and jets faking a τ_{had} (more details in Section 7.3.1). Minor background processes include the Drell-Yan processes, processes with diboson final states and processes associated with one Higgs boson (referred to as single Higgs boson processes). These background processes, except for the jets faking a τ_{had} background, are modelled with the full ATLAS detector simulation. For the FTAG calibration, the events used for calibration mostly originate from the $t\bar{t}$ processes. Other minor backgrounds include the single-top, diboson, production of $t\bar{t}$ in association with a boson and a boson produced in association with jets. Likewise, all processes are passed through the full simulation. The generators, PDFs, expansion order and the tune used in the HH searches and the FTAG calibration for the samples in common are in general identical, unless specified.

The $t\bar{t}$ production and the single top-quark events in the Wt-, s- and t-channels were simulated by the Powheg-Box v2 generator [136, 147–149] together with the NNPDF3.0NLO PDF set [141]. The showering, hadronisation and underlying event are modeled by Pythia 8.230, with parameters set according to the A14 tune [150] and using the NNPDF2.3LO PDF set. The top quark mass was set to 172.5 GeV, with top-quark spin correlations preserved. To achieve better accuracy, the $t\bar{t}$ production cross-section is calculated at next-to-next-to-leading order and next-to-next-to-leading-logarithm (NNLO+NNLL) [151]. While for the single-top process, the cross-sections are calculated at NLO [152–154]. The $t\bar{t}$-Wt interference is removed using the diagram removal scheme [155].

The production of bosons in association with jets (W/Z+jets) are simulated by Sherpa 2.2.1 generator [131] using the NNPDF3.0NNLO [141] PDF set. The tuning used for the parton shower is developed by the Sherpa authors. The matrix elements are simulated for up to two partons at NLO and up to four partons at LO, calculated with the Comix [156] and OpenLoops [157] libraries. The samples are normalised to NNLO prediction [158].

Similar settings apply to the diboson (WW, WZ and ZZ) events. Sherpa 2.2.1 generator is used to simulate these processes for the HH searches, while Sherpa 2.2.1 and Sherpa 2.2.2 are used for the FTAG calibration depending on the process. For both the HH searches and FTAG calibration, diboson samples are simulated using matrix elements at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. The cross-section is calculated at NLO accuracy, while the rest of the settings remains the same as the ones for the boson+jets background.

In addition, the events where a vector boson is produced in association with $t\bar{t}$ ($t\bar{t}Z$ and $t\bar{t}W$) are generated differently between the HH searches and the FTAG calibration. In the former, Sherpa 2.2.1 (2.2.8) is used to simulate the $t\bar{t}Z$ ($t\bar{t}W$) production at NLO. The cross-sections are calculated at NLO accuracy and the rest remains the same as the settings for the boson+jets. While for the FTAG calibration, the $t\bar{t}Z$ and $t\bar{t}W$ events are modelled using the MadGraph5_aMC@NLO v2.3.3 [140] generator at NLO with the

NNPDF3.0NLO [141] PDF. The events are interfaced to Pythia 8.210 [159] using the A14 tune [138] and the NNPDF2.3LO [141] PDF set. The decays of the bottom and charm hadrons are simulated using the EvtGen v1.2.0 program [130].

In the HH searches presented in this thesis, the SM single Higgs boson production is considered as part of the background. It has played a non-negligible role as a background to the analysis, especially in the non-resonant search. This is due to the similar kinematics between the single- and double- Higgs production. The simulated events of the single Higgs boson production, in various processes, were generated using Powheg-Box v2 generator and NNPDF3.0NNLO PDF set.

The ggF single Higgs production is interfaced to the Powheg NNLOPS program [160, 161] at NNLO accuracy. The cross-section is calculated at Next-to-Next-to-Next-to-Leading order (N3LO) for the QCD processes and NLO for the electroweak expansion [162–166]. On the other hand, the VBF single Higgs production events are interfaced to Pythia 8.212 using the CTEQ6L1 PDF set [167], together with the AZNLO tune [168]. The cross-section for VBF process is calculated at NNLO for QCD processes and NLO for electroweak expansion [162, 169–171].

For processes where a single Higgs is produced in association with a vector boson, i.e. the $qq \rightarrow ZH$, $gg \rightarrow ZH$ and $qq \rightarrow WH$, processes, Pythia 8.212 is used to simulate the parton shower and hadronisation using the AZNLO tune, together with the CTEQ6L1 PDF set. For the $qq \rightarrow WH$ and the $qq \rightarrow ZH$ samples, the cross-section is calculated at NNLO for QCD processes and NLO for electroweak expansion [172–178]; while for $gg \rightarrow ZH$ the cross-section is calculated at NLO for QCD and next-to-leading logarithmic (NLL) for electroweak expansion [179–183].

Finally, the simulated events with a single Higgs produced in association with a pair of top quarks are interfaced to Pythia 8.230. The parton shower and hadronisation are set to the A14 tune, using the NNPDF2.3LO PDF set. The cross-section is calculated at NLO accuracy [162].

In the HH searches and FTAG calibration presented in this thesis, additional samples were produced with alternative generators or settings, in order to estimate systematic uncertainties in the event modelling. These alternative samples will be defined once they are introduced.

Chapter 6

Charm jet mis-tagging calibration

6.1 Calibration methods for b-jets and light-jets

MC simulations are not able to model exactly the performance of the b-tagging algorithms in data. For this reason calibration is required, i.e. correcting MC to recover the data in terms of b-tagging efficiency, c-jet mis-tagging and light-jet mis-tagging rates [115]. The calibration is performed for all supported jet collections described in Section 4.4 and working points, which are cuts in the b-tagging algorithm output identifying the different tagging efficiencies and corresponding light-jet and c-jet rejection rates.

This chapter presents the author's work on the calibration of the c-jet mis-tagging efficiency. The calibration results are published in Ref. [184], and have become the official recommendation for all ATLAS users applying b-tagging.

In general, the efficiency is calculated with data and simulations, and scale factors are then calculated to match the efficiency extracted from simulations to the data.

The production of $t\bar{t}$ pairs at the LHC provides an abundant source of b-jets by virtue of the high cross-section and the $t \to Wb$ branching ratio being close to 100%. A very pure sample of $t\bar{t}$ events can be selected by requiring that both W bosons decay leptonically, referred to as di-leptonic $t\bar{t}$ decays in the following.

For the b-jet calibration, the performance of the b-tagging algorithms is evaluated in the simulation and the efficiency with which these algorithms identify jets containing b-hadrons is measured in collision data. The measurement uses a likelihood-based method in the di-leptonic $t\bar{t}$ sample, where events with exactly 2 jets and 2 opposite-sign leptons are selected. The data b-jet efficiency is then extracted from a combined likelihood fit and subsequently compared with that predicted by the simulation. Scale factors are then calculated to emulate the performance of the algorithms to the data [115].

For the light-jet mis-tagging calibration, two methods are used to measure the mis-tagging rate from the data [185]. The first is the negative tag method, which uses a high statistics data sample enriched in light-jets with the application of a modified algorithm which reverses some of the criteria used in the nominal identification algorithm. The second is the adjusted Monte Carlo (adjusted-MC) method, which adjusts the characteristic track observables in the simulation to the data and then compares the adjusted simulation to the 'standard' simulation. The scale factors are then calculated using these two methods. The scale factors of the two different methods are in good agreement within the systematic uncertainties.

6.2 Calibration method for c-jet

It is worth mentioning that the author's qualification task to become an ATLAS author was to calibrate the rate of a c-jet being mis-identified as a b-jet, which is a part of the calibration of the b-tagging algorithm. During the task, the calibration range has been extended down to 20 GeV (previously 25 GeV) in jet p_{T} and a new selection category has been developed to increase the data statistics of the scale factors in the high-p_T ($p_T^{jet} >$ 70 GeV) region. The calibration is performed on the PFlow jets (as defined in Section 4.4) and *VR-Track jets* reconstructed using the variable radius jet algorithm [186].

As determined by the CKM matrix [187, 188], the W boson decays dominantly to a pair of light quarks (u quark and d quark) or to a s quark and a c quark. The W boson decays very rarely to pairs containing a b quark. More specifically, the branching ratio of W boson decays to a u quark and d quark pair or a s quark and c quark pair is 33.1%, and to pairs containing a b quark is only 0.057% [10]. Therefore, b-tagged jets from the W decay are most likely to be mis-tagged c-jets or light-jets.

Furthermore, given the ratio between the DL1 light-jet rejection and the corresponding c-jet rejection ranges from 10 to 40 (Figure 4.7), the c-jet is much more likely to be mis-tagged than the light-jet. This allows for a source of mis-tagged c-jets to be obtained in the $t\bar{t}$ events, requiring that one W boson decays leptonically and the other decay hadronically (referred to as semi-leptonic $t\bar{t}$ decay in the following), where the b-tagged jets from the W decay are candidates of mis-tagged c-jets. Requiring a W boson decaying leptonically reduces the number of combinations of jets of different flavour, and allows triggering with the lepton.

The event kinematics are shown by the diagram in Figure 6.1, where the $t\bar{t}$ pair decays to a b and a $\bar{b}$ quark, circled in red. One of the W bosons, circled in blue, decays hadronically to quarks, and the other W boson decays leptonically to either an electron or a muon and the corresponding neutrinos, circled in green and purple, respectively. The lepton in the final state is used for triggering. The following notation will be used: the jets that are the decay products of the W boson are referred to as W-jets and the remaining two jets are referred to as top-jets.

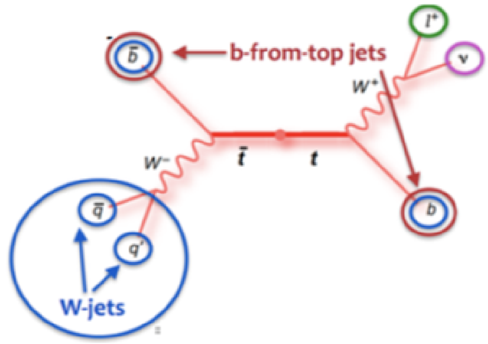


Figure 6.1: Feynman diagram of the semi-leptonic $t\bar{t}$ events.

A kinematic likelihood technique referred to as KLFitter [189], is used to assign jets to the proper $t\bar{t}$ decay product (more details in Section 6.4).

The c-jet efficiency is defined as the ratio of events with either of the W-jets tagged. The efficiency is evaluated in four p_{T} intervals, with boundaries of 20, 40, 65, 140 and 250 GeV for the PFlow jets and 15, 20, 40, 140 GeV for the VR-Track jets; and for four b-tagging working points with b-tagging efficiencies of 85%, 77%, 70% and 60%.

The choice of the bin boundaries ensures enough statistics for each bin and hence relatively flat statistical uncertainty, given the underlying c-jet p_{T} spectrum as shown in Figure 6.8. The boundaries for the VR-Track jets are lower than for PFlow jets, since the track jets miss the neutral particles the reconstructed energy is significantly below the true jet energy.

The main method described in this chapter is for the 'fixed-cut' calibration, where the efficiency is defined as the fraction of b-jets passing the tagger. Jets are said to be tagged (untagged) at a particular working point if they have DL1r scores greater (less) than the DL1r score of that working point. The events with both W-jets b-tagged are discarded to simplify the fit described in the following.

To extract the scale factors of the c-jet mis-tagging, a fit is performed by minimising the χ^2 defined as:

$$\chi^2 = \sum_{t=1}^{4}\sum_{i=1}^{4}\sum_{j=1}^{4}\Big\{N^t_{\mathrm{data}}(i,j) - p(i,j)\big[c^t(i)N^t_C(i,j)$$

$$+N^t_J(i,j) + \sum_k c^{t=4}(k)N^t_X(i,j,k)\big]\Big\}^2 / N^t_{\mathrm{data}}(i,j)$$

$$+\sum_{i=1}^{4}\sum_{j=i}^{4}\left[N_{\text{data}}^{\text{untag}}(i,j)-p(i,j)N_{\text{MC}}^{\text{untag}}(i,j)\right]^2/N_{\text{data}}^{\text{untag}}(i,j). \tag{6.1}$$

The basic idea behind this formula is to compare the data to the total expectation, which is dominated by $t\bar{t}$ events. The $c^t(i)$ is the main floating parameter in the fit, which is the c-jet mis-tagging scale factor at working point t, with a b-tagged W-jetin p_{T} bin labelled i. Another floating parameter is $p(i,j)$, which is the overall normalisation factor scaling the MC to data, with a b-tagged W-jet in the p_{T} bin i and the other (untagged) W-jet in the p_{T} bin labelled j.

In the first row of the formula, $N_{\text{data}}^t(i,j)$ is the number of data events with a b-tagged W-jet in p_{T} bin i and the untagged W-jet in p_{T} bin j. Similarly, $N_C^t(i,j)$ is the number of MC events with a b-tagged W-jet, and this b-tagged W-jet is indeed a c-flavoured jet, which is checked by the truth information of the MC simulation. This type of events can be seen as 'signal'.

In contrast, $N_J^t(i,j)$ in the second row is the number of events when neither the b-tagged W-jet nor the jets assigned to the top quark (referred to as top jets) are c-flavoured jet (checked by MC truth information); and $N_X^t(i,j,k)$ in the second row is the number of events where one of the top jets is a c-flavoured jet. These two type of events can be seen as 'background'. The latter term applies to events where the KLFitter wrongly assigns a c-flavoured jet from the W to a top quark. This background requires a specific treatment since this c-flavoured jet is in a p_{T} bin, labelled k, different from those of the W-jets and so must be binned in this variable in addition to i and j. In addition, this background depends only on $c^{t=4}(k)$ becasue the two top jets are required to be b-tagged at the 60% working point. Finally, $N_{\text{data}}^{\text{untag}}(i,j)$ and $N_{\text{MC}}^{\text{untag}}(i,j)$ are the numbers of data and MC events with no b-tagged W-jets. The calibration is then given as the scale factors of the four working points in bins of p_{T} defined in the above text.

6.3 Alternative Monte Carlo samples

The events that are used in this study originate mostly due to $t\bar{t}$ production. This process is produced with settings defined in Section 5.3.3. While all samples were produced using the ATLAS simulation infrastructure and Geant4 software, this process is also produced with fast simulation, using the same generators and settings. This alternative sample is used to estimate the uncertainty due to different choice of simulation.

For the nominal setting, the $t\bar{t}$ process is modelled using the h_{damp} parameter[1] set

[1]The h_{damp} parameter is a resummation damping factor and one of the parameters that controls the matching of Powheg matrix elements to the parton shower and thus effectively regulates the high-p_{T} radiation against which the $t\bar{t}$ system recoils.

to 1.5 m_{top} [190]. An alternative sample is produced by varying the h_{damp} and the renormalisation and factorisation simultaneously, and choosing the Var3c up/down variants of the A14 tune. This alternative sample is produced in order to estimate the uncertainty due to ISR. While the impact of the FSR is estimated by varying the renormalisation scale for emissions from the parton shower by a factor of two up or down.

Another alternative sample with replacing the PYTHIA 8.230 interface to the HERWIG 7.04, while keeping the same generator (POWHEG-BOX v2) and the same parameters for the generation. The HERWIG 7.04 is used together with the H7UE tune and the MMHT2014LO. This alternative sample is produced to estimate the uncertainty due diffenrent choice of parton shower and hadronisation model.

All the alternative samples are generated using the fast simulation.

6.4 Kinematic Likelihood Fitter

The simulated $t\bar{t}$ events are split according to the flavour of the W-jets. The notation '$t\bar{t}$, ll' denotes that both W-jets are light-flavour jets. Similarly, '$t\bar{t}$, cl' ('$t\bar{t}$, bl') indicates that one of the W-jets is a c-jet (b-jet) whereas the other is a light-flavour jet. W-jets with flavour other than what is discussed above fall into the category denoted by '$t\bar{t}$, other'. This category includes events in which at least one of the W-jets comes from a hadronically decaying τ-lepton.

The four-vectors of the four highest p_{T} jets, the lepton and the event $E_{\text{T}}^{\text{miss}}$ are used as inputs to a likelihood-based $t\bar{t}$ event reconstruction algorithm, which is described in more detail in Ref. [189]. This algorithm uses a likelihood function to assign the four jets to the $t\bar{t}$ decay topology. In particular, the algorithm assigns one jet to be the b-jet from the leptonically decaying top quark ($t \to Wb \to \ell\nu b$), another to the b-jet from the hadronically decaying top quark ($t \to Wb \to qq'b$, where qq' are the quarks in which the W boson decays) and the remaining two jets to the jets that come from the hadronic W boson decay. The jet assignment does not use any b-tagging information to avoid bias.

6.5 Maximising likelihood

Taking only four jets in the event limits the total number of possible jet assignments (permutations) in the event. For example, in the semi-leptonic channel, four jets can be permuted a total number of times equal to $4! = 24$. However, the two W-jets are kinematically indistinguishable. This reduces the possible number of permutations to 12. For every combination of jet assignments, the likelihood is maximised over its free parameters, the energy of the four jets, the lepton energy and the three components of the momentum of the neutrino, and provides a value based on how closely the kinematic information from the reconstructed objects for a specific jet assignment resembles the expected kinematic behaviour of the decay of a Standard Model semi-leptonic $t\bar{t}$ event. The likelihood therefore distinguishes the possible permutations on an event-by-event basis. The best permutation, given by the largest negative log-likelihood value, is adopted as the jet assignments for the event. An additional requirement of log-likelihood > -48 is placed

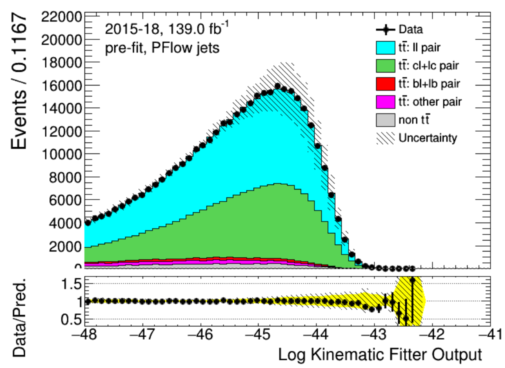


Figure 6.2: Distribution of the negative logarithm of the likelihood that is used to reconstruct the $t\bar{t}$ decay. The distribution shown is for the combination, in each event, which has the highest value of the negative log-likelihood.

on the output of the likelihood value for the chosen event permutation. An example of the distribution of log-likelihood of the best permutations is shown in Figure 6.2. In this figure, the data events are compared against the simulation. The majority of the events come from $t\bar{t}$ production. There is only a very small fraction of events, which is denoted as 'non $t\bar{t}$' on the figure, that come from other processes like W or Z production in association with jets or single-top production.

6.6 Event selection

6.6.1 Standard selection

Events are required to contain exactly one trigger-matched lepton with $p_T > 27$ GeV and exactly four jets with $p_T > 25$ GeV. Leptons are required to have p_T above 27 GeV in order to avoid the turn-on curve for the single lepton triggers. Events which contain an additional lepton with $p_T > 27$ GeV are rejected. The events are also required to have $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV, which is assumed to be the result of the neutrino from the leptonically decaying W boson. The transverse mass m_T between the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$, is constrained as follows:

$$m_T = \sqrt{2p_T^{\ell}E_{\mathrm{T}}^{\mathrm{miss}}(1-\cos\Delta\phi)} > 40 \text{ GeV}, \tag{6.2}$$

where $\Delta\phi = \phi(E_{\mathrm{T}}^{\mathrm{miss}}) - \phi(\ell)$ is the azimuthal difference between the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$. The

	PFlow jets	Track jets
Data	227118	218351
$t\bar{t}$	235670 ± 200	223770 ± 180
Non $t\bar{t}$	7610 ± 120	7280 ± 100
Data/MC	0.934 ± 0.002	0.945 ± 0.002

Table 6.1: Standard selection: prefit comparison of the number of events in data and in simulation considering the PFlow jets and the VR-Track jets for events with exactly 4 jets.

yields of the data and the MC are given in Table 6.1. An example of the p_{T} distributions before any tagging or fitting and after the standard selection is shown in Figure 6.3. More plots can be found in Appendix A.1.1. The yellow band in the lower pad shows the overall systematic uncertainties, combining the experimental uncertainties and the $t\bar{t}$ modelling uncertainties, as described in Section 6.7. The data/MC ratio shows good agreement within the systematic uncertainties.

6.6.2 Low-p_{T} selection

The author has developed an orthogonal selection to extend the calibration in the low-p_T region so that the calibration can be applied to PFlow jets with $20 < pt < 25$ GeV. The p_{T} threshold of the VR-Track jets is 10 GeV therefore the low-p_{T} selection is not needed. Instead of requiring events to have exactly 4 jets $p_{\mathrm{T}} > 25$ GeV, events are required to have exactly 3 jets with $p_T > 25$ GeV and exactly 1 jet with 20 GeV $< p_T <$ 25 GeV. Other than that, all requirements for the selection are the same. This additional region provides candidates for the PFlow W-jets that are used for calibration in the $20 - 25$ GeV region. The inclusive yields of the low-p_{T} selection of the data and the MC are given in Table 6.2, and the p_{T} distributions of the W-jets are shown in Figure 6.4. More plots of the kinematic distributions are shown in Appendix A.1.2. Good agreement between MC and data is shown in these distributions, and the p_T range of the sub-leading has gone down to 20 GeV.

	PFlow jets
Data	59987
$t\bar{t}$	56530 ± 90
Non $t\bar{t}$	3340 ± 60
Data/MC	1.002 ± 0.004

Table 6.2: Low-p_{T} selection: prefit comparison of the number of events in data and MC for the PFlow W-jets. Events are required to have exactly 3 jets with $p_{\mathrm{T}} > 25$ GeV and one jet with $20 < p_{\mathrm{T}} < 25$ GeV.

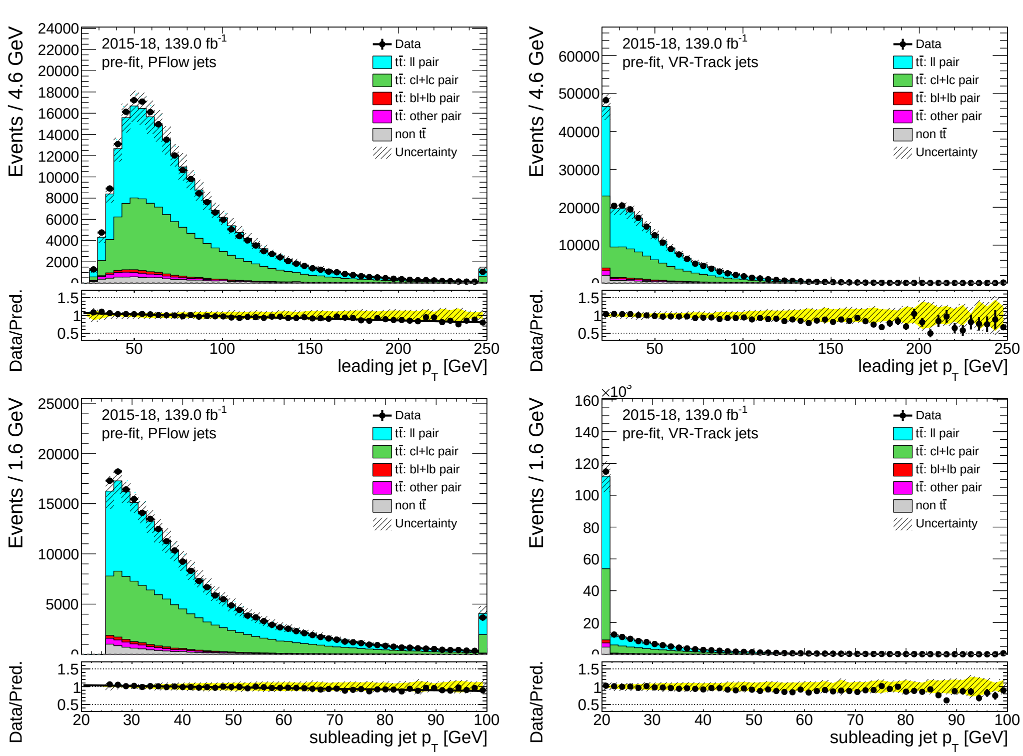


Figure 6.3: Standard selection: data versus simulation of the leading and sub-leading W jet p_{T} for the PFlow jets in the left column and for VR-Track jets in the right column. The leading jet and sub-leading jet refer to the highest p_{T} W jet and the second highest p_{T} jet, respectively. The 'non $t\bar{t}$' background indicates background comes from non-$t\bar{t}$ processes like W or Z production in association with jets or single-top production. The error in the table (and the following yields tables for different selection) is stats-only.

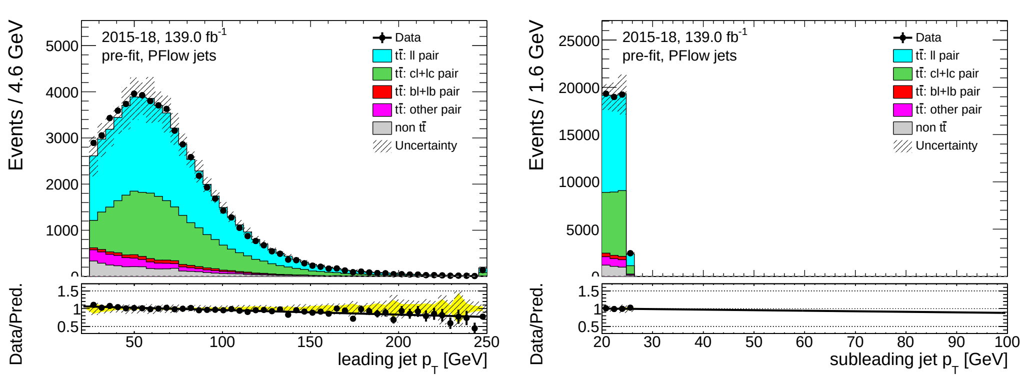


Figure 6.4: Low-p_{T} selection: data versus simulation of the PFlow W-jets p_{T}.

6.6.3 High-p_{T} selection

It has been observed in the previous calibrations that the statistics are relatively low for the high-p_{T} region (e.g. jet $p_{\mathrm{T}} > 100$ GeV). Therefore, the author has worked on an orthogonal selection to improve this situation. Instead of requiring events to have exactly 4 jets, events are required to have at least 5 jets with $p_T > 25$ GeV, in which at least 1 jet has $p_\mathrm{T} > 70$ GeV. Other than that, all requirements for the selection remain the same.

As the 'signal' of the c-jet calibration is the event with a true c-flavoured jet, it would be useful to introduce the c-jet purity, defined as

$$c\text{−jet purity} = \frac{N_{\text{true } c\text{−jet}}}{N_{\text{all}}} \tag{6.3}$$

where $N_{\text{true } c\text{−jet}}$ stands for the number of events with a true c-flavoured jet from the W decay, and N_{all} stands for the number of all events.

The affect of loosening the 4-jet-requirement and requiring at least one jet with $p_\mathrm{T} >$ 70 GeV on the c-jet purity is checked; the value of the 70 GeV is chosen such that the c-jet purity is relatively unaffected compared to the standard selection, while achieving a high gain in statistics. The c-jet purity and the gain in statistics are shown in Figure 6.6.

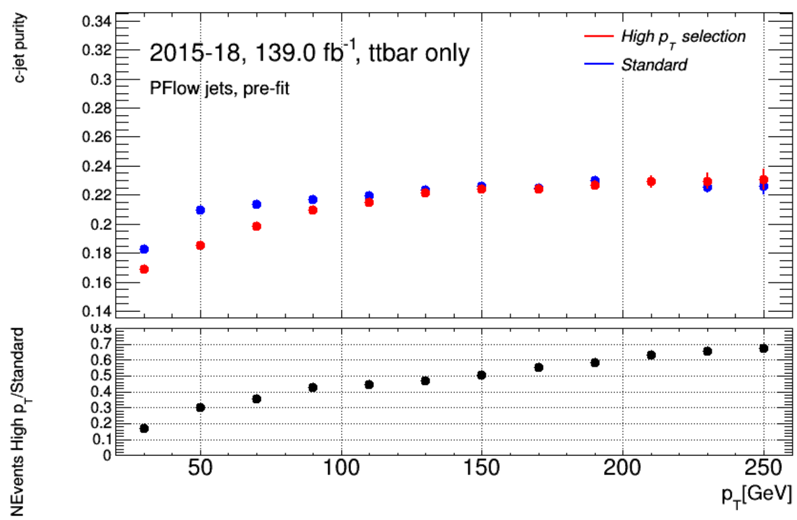


Figure 6.5: Cut value: 70 GeV

Figure 6.6: Comparison of c-jet purity of the high-p_T selection to the standard selection. The bottom pad shows the ratio of number of events of the high-p_T selection to the standard selection.

The yields of the data and the MC are given in Table 6.3. An example of the p_T distributions before any tagging or fitting, applying the high-p_T selection is shown in Figure 6.7. In general the event statistics improve about 80% in the region with $p_\mathrm{T} >$ 70 GeV as desired. More plots can be found in Appendix A.2.

	PFlow jets	Track jets
Data	98273	83957
$t\bar{t}$	99430 ± 120	87476 ± 110
Non $t\bar{t}$	1842 ± 21	1570 ± 20
Data/MC	0.97 ± 0.003	0.94 ± 0.003

Table 6.3: High-p_{T} selection: prefit comparison of the number of events in data and in simulation considering the PFlow W-jets and the VR-Track jets.

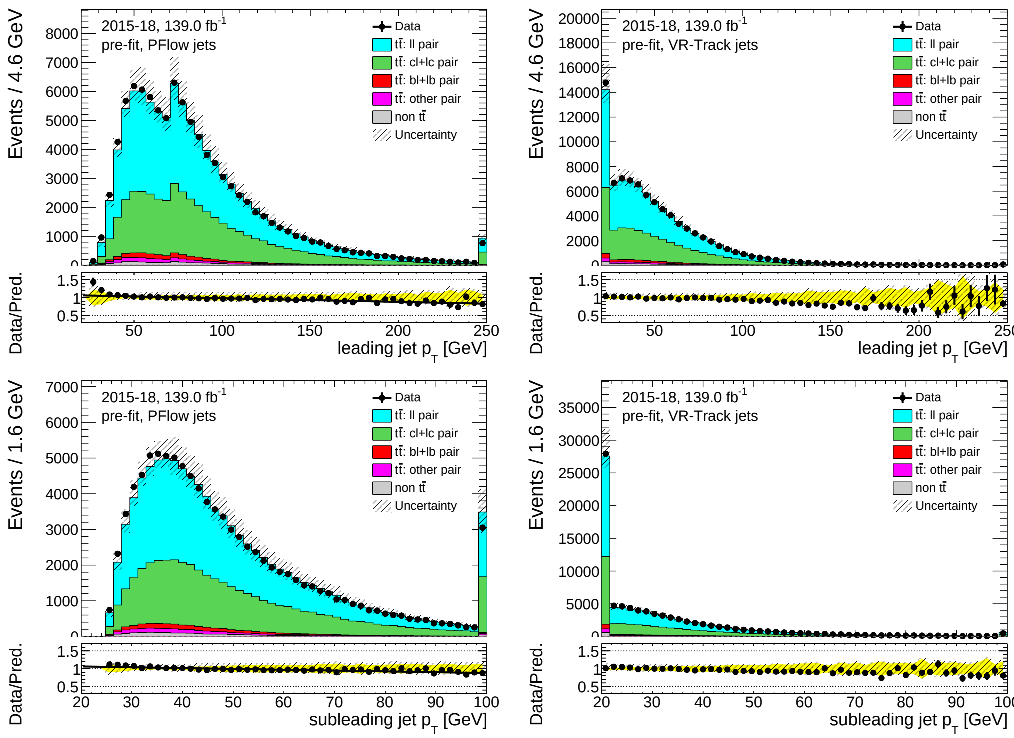


Figure 6.7: High-p_{T} selection: data versus simulation of W-jets p_{T} for PFlow jets in the left column and for VR-Track jets in the right column.

6.6.4 Combined selection

As the standard selections, low-p_{T} selection and high-p_{T} selection are orthogonal to each other, all the selections are combined to provide the maximum range and statistics for the calibration. The yields of the data and the MC are given in Table 6.4. An example of the p_{T} distributions before any tagging or fitting and after the combined selection is shown in Figure 6.8. More plots can be found in Appendix A.3.

	PFlow jets	Track jets
Data	385378	302308
$t\bar{t}$	383520 ± 230	302690 ± 200
Non $t\bar{t}$	12420 ± 120	8570 ± 100
Data/MC	0.973 ± 0.002	0.971 ± 0.002

Table 6.4: Combined selection: prefit comparison of the number of events in data and in simulation considering the PFlow jets and the VR-Track jets for an inclusive selection.

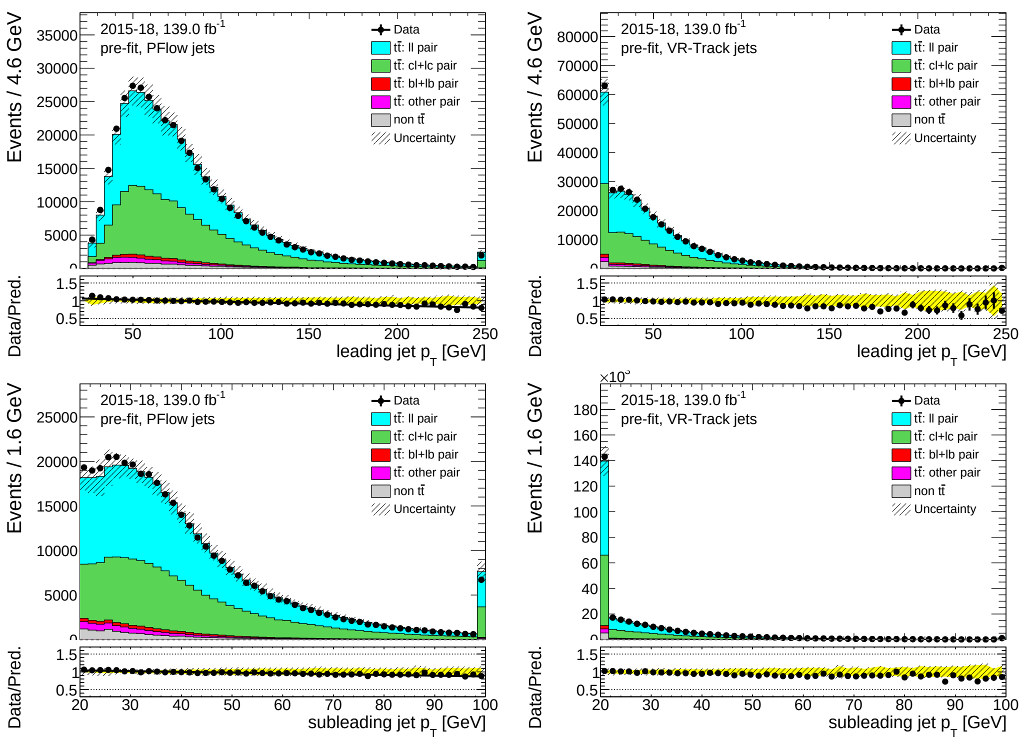


Figure 6.8: Combined selection: data versus simulation of W-jets p_{T} for PFlow jets in the left column and for VR-Track jets in the right column.

6.7 Systematic uncertainties

The systematic uncertainties considered and propagated in this calibration can be broadly categorised into experimental and modelling systematic uncertainties.

6.7.1 Experimental uncertainties

Experimental uncertainties are related to the detector and estimated using data-driven methods or simulations. The lepton energy scale and resolution are corrected to provide a better agreement between MC predictions and data, uncertainties due to these corrections are considered. Uncertainties are taken into account on the electron and muon trigger, identification and reconstruction efficiencies, and for uncertainties associated with the isolation requirements.

The JES uncertainty depends on p_T and η and takes into account uncertainties due to pile-up effects. Uncertainties on JER are taken into account. Uncertainties on the energy scale and resolution of the electrons, muons, jets and taus are propagated to the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$, which also has additional dedicated uncertainties on the scale, resolution, and reconstruction efficiency of tracks not associated to any of the reconstructed objects, along with the modelling of the underlying event. Uncertainties on the b-tagging (mis-tagging) probabilities for b- (light-) jets are considered both for the tagging jets assigned to the b quark from the top decay and for the jets associated to the hadronically decaying W boson. Supporting material for this section can be found in the Appendix, Table A.1.

6.7.2 Modelling uncertainties

The uncertainty due to different choices of the parton shower model is estimated by comparing the MC samples generated with the nominal parton shower model and with the alternative parton shower model. More specifically, it is derived by comparing the prediction from Powheg interfaced either to Pythia or Herwig++. The uncertainty due to additional radiation in the initial state and the final state is estimated by comparing the nominal MC samples with the MC samples with alternative scales of renormalisation and factorisation. The uncertainty on modelling of initial state radiation (ISR) is assessed with two alternative Powheg+Pythia 8 samples. The samples include one with an increase in radiation which has the renormalisation and factorisation scales decreased by a factor of two, and the *hdamp* parameter doubled (which controls the p_{T} of the first additional emission), while the sample with a decrease in radiation has the scales increased by a factor of two. In all cases, MC-to-MC SFs are taken into account. In addition, the uncertainty due to the variations samples being produced by fast simulation while the nominal samples being produced full simulation is also considered. The comparisons of the nominal $t\bar{t}$ sample and the samples with each systematic uncertainty are shown in Table 6.5.

6.8 Under-estimation of $t\bar{t}$ + Heavy flavour background

Despite the fact that the true nature of most of the reconstructed W-jets are either c-jets or light-jets, there is still a very small amount which are true b-jets.

There are two main sources of these true b-jets. The first is a W boson decays to a b and a c quark. The second is when the $t\bar{t}$ plus a gluon process (referred to as $t\bar{t}$ + heavy flavour process) is selected, and the gluon splits into a pair a b quarks, and one of them is assigned as a W jet. The first source can be excluded by requiring no c-jets in the W-jets, meaning the true b-jet in the W-jets can only come from the $t\bar{t}$ + heavy flavour process.

	PFlow jets		Track jets	
	Yields	Ratio of difference to nominal sample	Yields	Ratio of difference to nominal sample
$t\bar{t}$ Nominal Data/MC	385378 ± 230 0.973 ± 0.002		302690 ± 200 0.971 ± 0.002	
$t\bar{t}$ AF2 DATA/MC(AF2)	386260 ± 250 0.967 ± 0.002	0.716%	304860 ± 230 0.965 ± 0.002	0.716%
$t\bar{t}$ ISR DATA/MC(ISR)	377130 ± 220 0.989 ± 0.002	-1.665%	297960 ± 200 0.986 ± 0.002	-1.562%
$t\bar{t}$ Herwig DATA/MC(Herwig)	331960 ± 220 1.119 ± 0.002	-13.443%	259940 ± 190 1.126 ± 0.002	-14.123%

Table 6.5: Comparison of the number of events in data and in simulation considering the PFlow jets and the VR-Track jets for an inclusive selection. The uncertainty due to the variations samples being produced by fast simulation is included in the table as $t\bar{t}$ AF2.

This process is underestimated by the MC by about 30% for both the PFlow and VR-Track jets collections, as shown in Table 6.6 and Figure 6.9, where an extra cut requiring at least one W jet with DL1r > 8 is added to the combined selection to reject most of the true c-jets and true light-jets. A more thorough study is done in Ref. [191], where the mis-modelling factor is measured to be 1.25 ± 0.25, which is also consistent with the 30% mis-modelling observed in the previous study. Therefore, events in the simulation in which the top jets and at least one of the W-jets are b-jets (referred to as 3 true b-jets events), are scaled by 1.25 ± 0.25. All results shown in this chapter have this scale factor implemented, and the full difference between the simulation before applying this scale factor and after is taken as a systematic uncertainty. This uncertainty has been added in quadrature to the systematic uncertainties described in Section 6.7 in all the plots in this chapter.

	PFlow jets	VR-Track jets
Data	1589	1336
$t\bar{t}$	1100 ± 13	940 ± 12
Non $t\bar{t}$	83 ± 6	69 ± 5
Data/MC	1.34 ± 0.04	1.32 ± 0.04

Table 6.6: Yields of the 2018 data and MC of the combined selection, requiring at least 1 PFlow or track W jet with DL1r > 8 to reject most of the light- and c-jets.

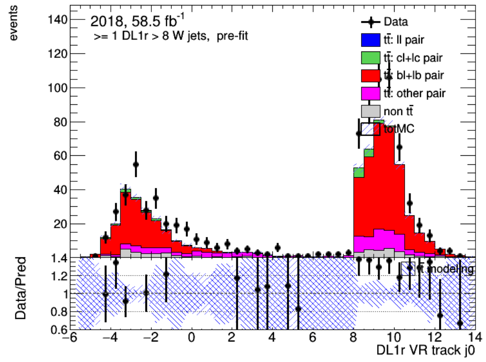


Figure 6.9: The DL1r score distribution of the leading VR-Track jet, requiring at least 1 VR-Track jets have DL1r > 8 to reject most of the light and the c jets, with $t\bar{t}$ modelling and statistical uncertainties.

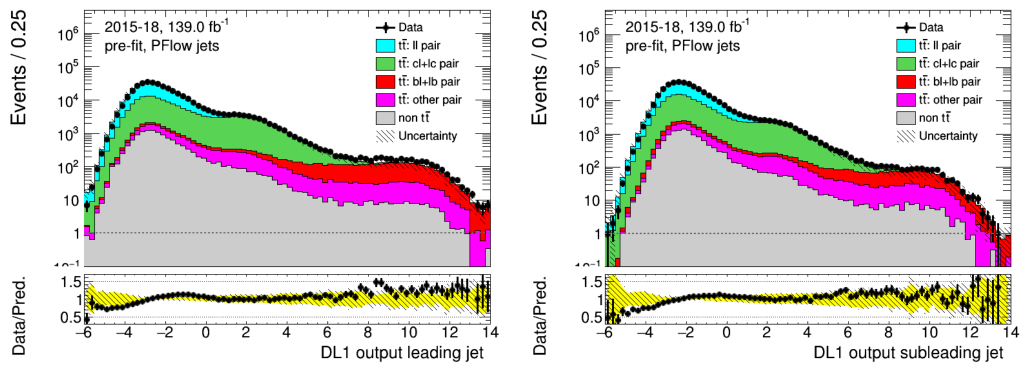
6.9 Results

6.9.1 Overview

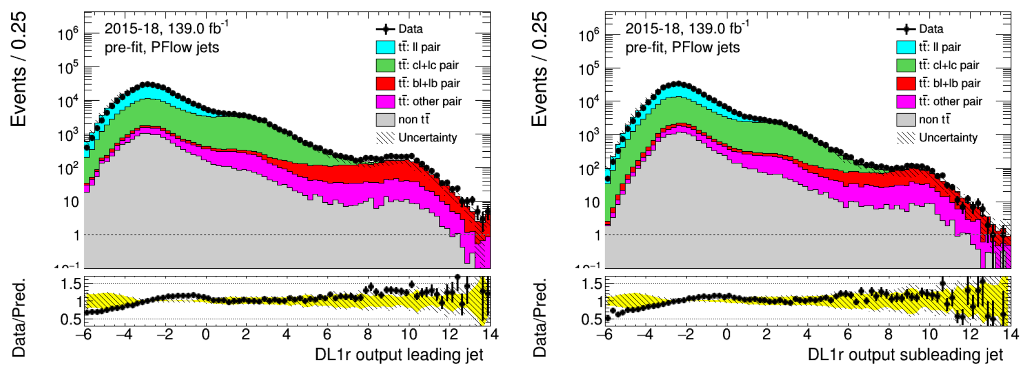
Four rounds of calibrations have been carried out, containing different jet collections, Monte Carlo samples, analysis framework and b-jet identification algorithm. In the latest round, the calibration includes the PFlow jet and the VR-Track jet collection, and MV2c10, DL1 and DL1r taggers. The low-p_T selection and the standard selection are carried out for all four calibrations, while the high-p_T selection is only implemented in the latest calibration.

6.9.2 b-tagging algorithms output distribution

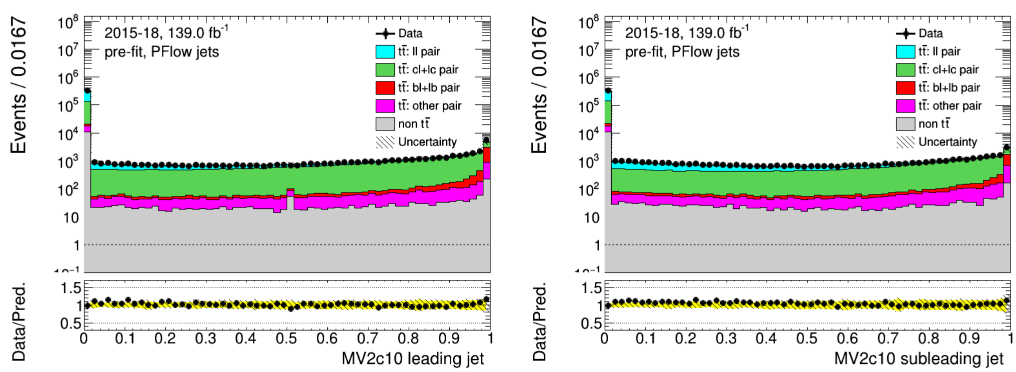
The distributions of the b-tagging algorithms' output of the MC and the data of the latest calibration (December 2020) are shown in Figure 6.10 for the PFlow jets and Figure 6.11 for the VR-Track jets, combining the standard selection, low p_{T} and the high-p_T selection. In these figures, the data events are compared against the simulation. The majority of the events come from $t\bar{t}$ production. There is only a very small fraction of non $t\bar{t}$ events. The W-jet pairs are mostly light-jets pairs and c-jet light-jet pairs, and a very small fraction of the pairs are b-jet light-jet pairs or pairs containing one or more τ hadron(s). The yellow band in the lower pad indicates the overall systematic uncertainties and the black band represents the $t\bar{t}$ modelling systematic uncertainty, which dominates at low b-tagging discriminant (DL1 or DL1r < 4). The experimental systematic uncertainty is in general very small. At high b-tagging discriminant (DL1 or DL1r > 4), the uncertainty due to the 1.25 ± 0.25 scale factor becomes more important.



(a) DL1 tagger output

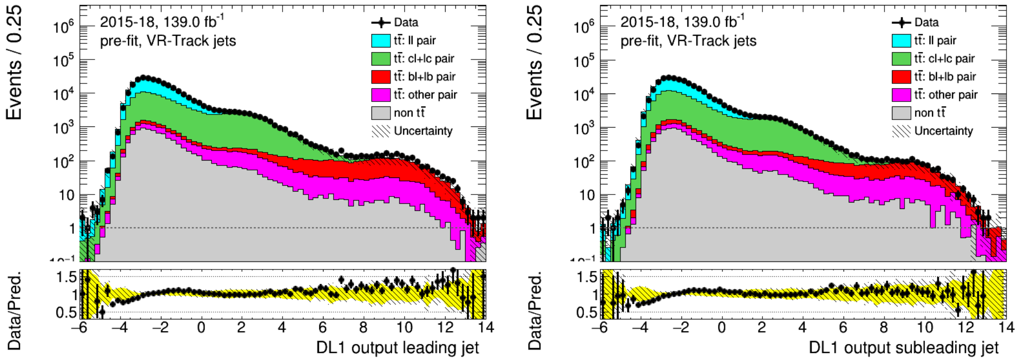


(b) DL1r tagger output

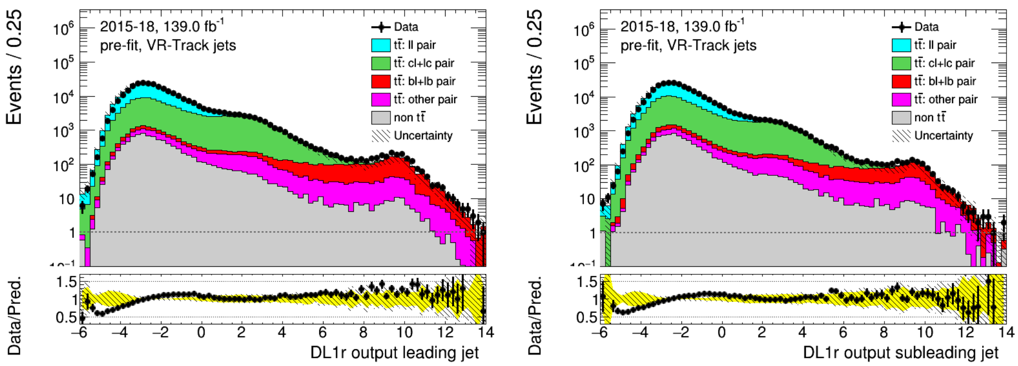


(c) MV2c10 tagger output

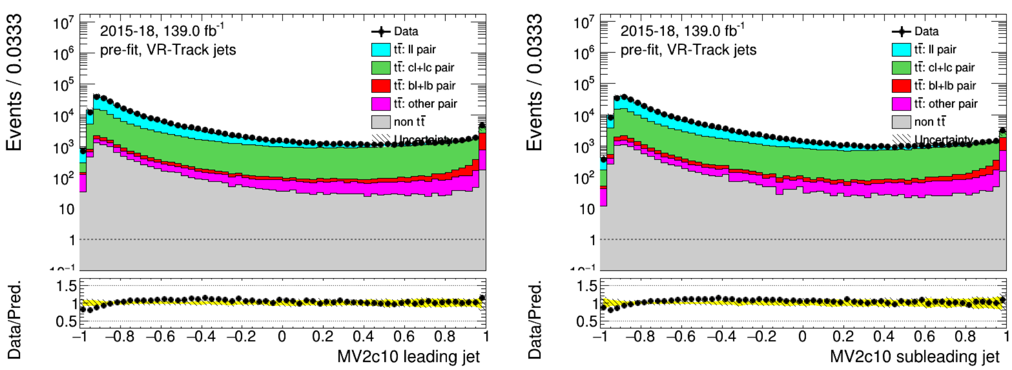
Figure 6.10: PFlow jets: distributions of the DL1, DL1r and MV2c10 tagger outputs of the combined selection, leading jet in the left column and sub-leading jet in the right column, before fitting or tagging with full uncertainties.



(a) DL1 tagger output



(b) DL1r tagger output



(c) MV2c10 tagger output

Figure 6.11: VR-Track jets: distributions of the DL1, DL1r and MV2c10 tagger outputs of the combined selection, leading jet in the left column and sub-leading jet in the right column, before fitting or tagging with full uncertainties.

6.9.3 Efficiencies and Scale Factors

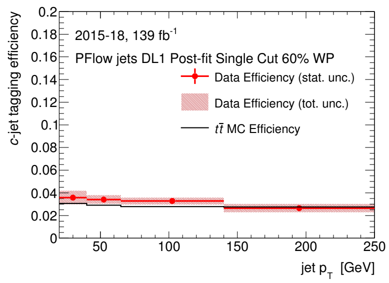
The DL1 and DL1r c-jet efficiencies and scale factors with systematic uncertainties are calculated with four fixed cut working points for the PFlow and VR VR-Track jets collection in the latest derivation in December 2020.

The c-jet mis-tagging efficiencies are shown in Figures 6.12-6.15 for the PFlow jet collections and the VR-Track jets with the DL1 and the DL1r tagger. For PFlow jets, these results combine the standard selection, low-p_{T} selection and the high-p_T selection and for the VR-Track jets, they combine the standard selection and the high-p_T selection.

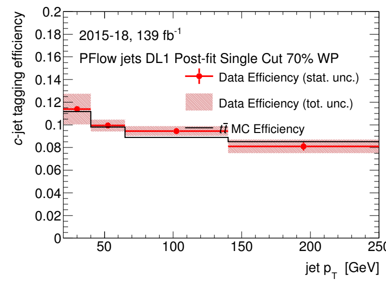
The 1.25 ± 0.25 scale factor is applied on events with 3 true b-jets. The overall uncertainties are shown in the red band. The scale factors are shown in Figure 6.16-6.19 for the PFlow jets and the VR-Track jets with the DL1 and DL1r tagger. The tighter working points (60%, 70%) show larger uncertainties and bigger deviation from 1, while the looser working points (77%, 85%) have much smaller uncertainty, and the simulation is able to recover the data well due to more abundant event statistics. For the PFlow jets, in most of the working points, the systematic uncertainties dominate in the low-p_{T} bins (p_{T} < 150 GeV) and the statistical error, represented by the error bars on the markers become more important in the last bin. For the VR-Track jets, the statistical uncertainty is relatively constant for all bins, while the systematic uncertainty increases as the p_{T} increases. To demonstrate the effect on statistics with the high-p_{T} selection, the fractional statistical uncertainties of 60% working point scale factor are shown in Table 6.7 for the standard and the combined selection. In some bins, the statistical uncertainty can decrease by up to 30%, suggesting that the high-p_{T} selection is successful at increasing events statistics.

	PFlow jets			VR-Track jets		
	Standard selection	High-p_{T} selection	Fractional decrease	Standard selection	High-p_{T} selection	Fractional decrease
Bin No.1	3.3%	3.3%	0.0%	5.6%	5.3%	5.7%
Bin No.2	3.1%	2.8%	10.7%	4.2%	3.7%	13.5%
Bin No.3	3.4%	2.6%	30.8%	5.8%	4.9%	18.4%
Bin No.4	12.1%	9.3%	30.1%	7.2%	5.6%	28.6%

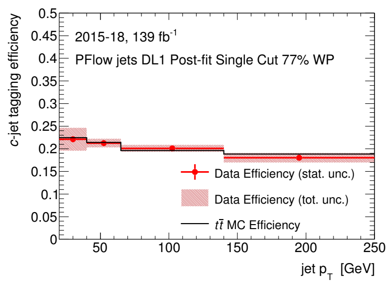
Table 6.7: Comparison of the fractional statistical uncertainty in the DL1r 60% working point scale factor. The p_{T} range of each bin can be found in section 6.2.



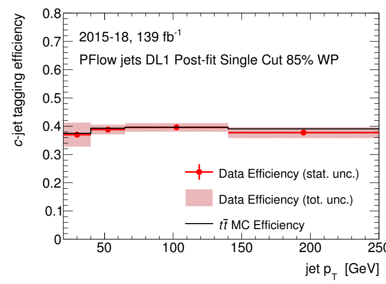
(a) 60% working point



(b) 70% working point

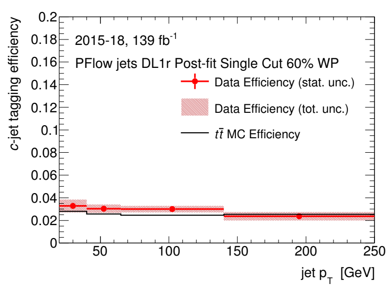


(c) 77% working point

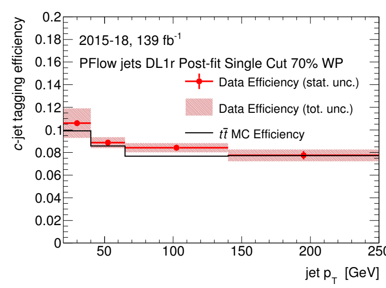


(d) 85% working point

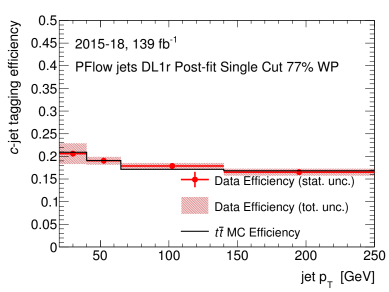
Figure 6.12: Efficiencies for c-jets to be mis-tagged as b-jets for the PFlow jets collection with the DL1 tagger.



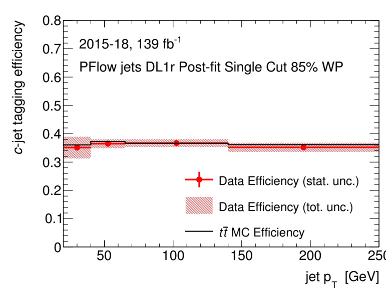
(a) 60% working point



(b) 70% working point

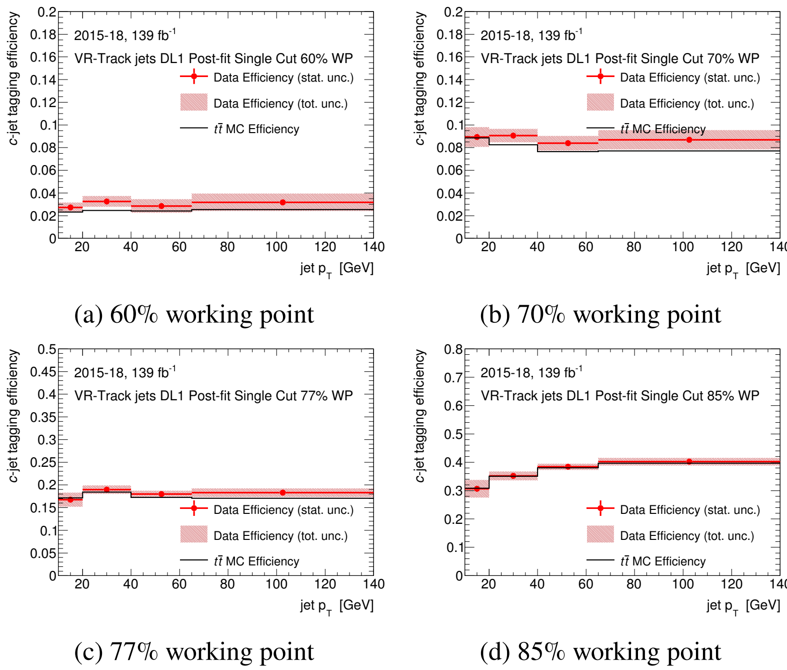


(c) 77% working point



(d) 85% working point

Figure 6.13: Efficiencies for c-jets to be mis-tagged as b-jets for the PFlow jets collection with the DL1r tagger.



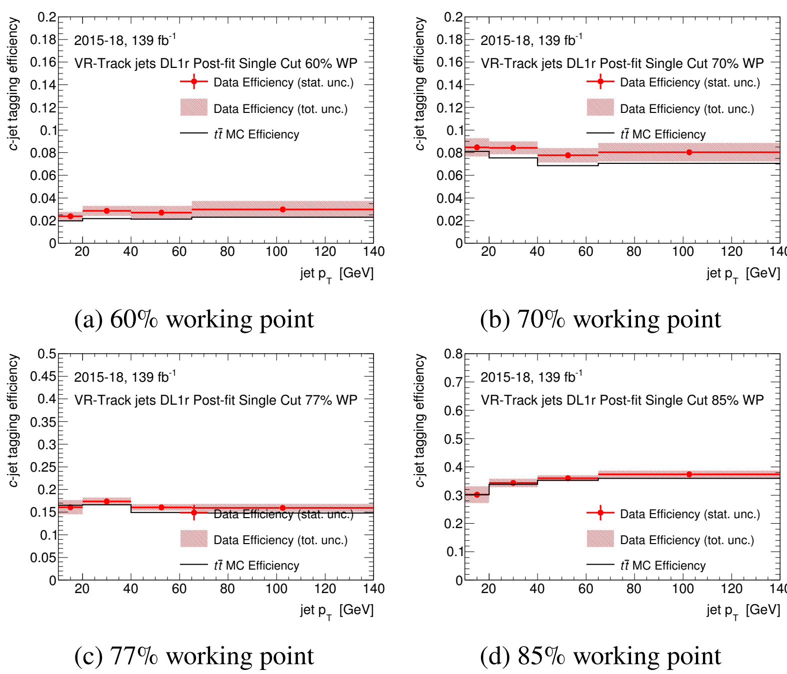
(a) 60% working point

(b) 70% working point

(c) 77% working point

(d) 85% working point

Figure 6.14: Efficiencies for c-jets to be mis-tagged as b-jets for the VR-Track jets collection with the DL1 tagger.



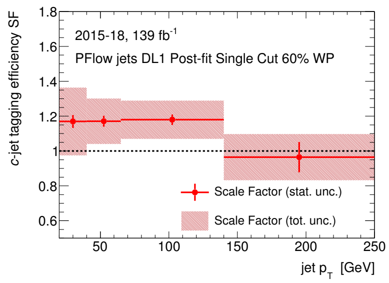
(a) 60% working point

(b) 70% working point

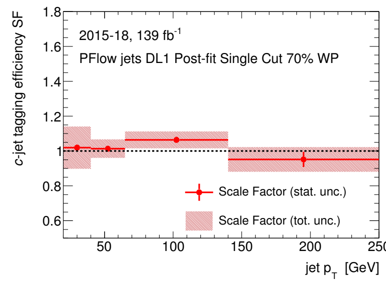
(c) 77% working point

(d) 85% working point

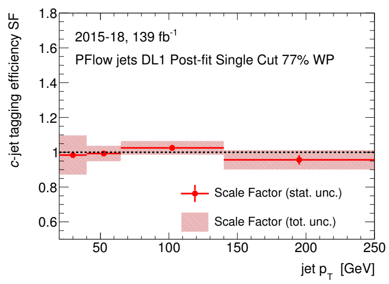
Figure 6.15: Efficiencies for c-jets to be mis-tagged as b-jets for the VR-Track jets collection with the DL1r tagger.



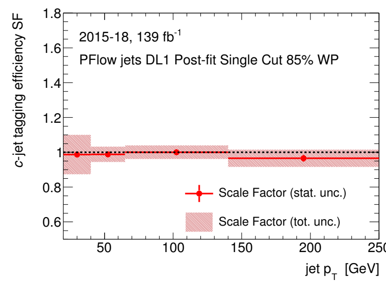
(a) 60% working point



(b) 70% working point

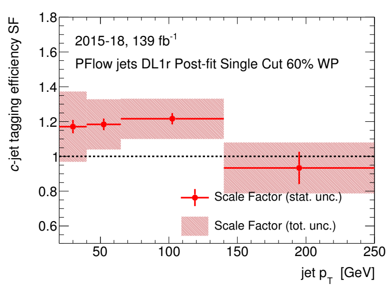


(c) 77% working point

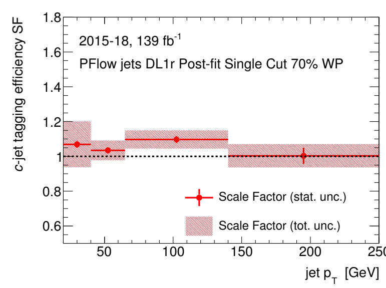


(d) 85% working point

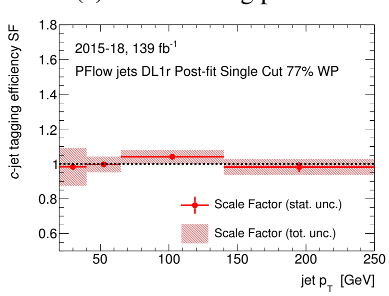
Figure 6.16: c-jet mis-tagging scale factors for the PFlow jets collection with the DL1 tagger.



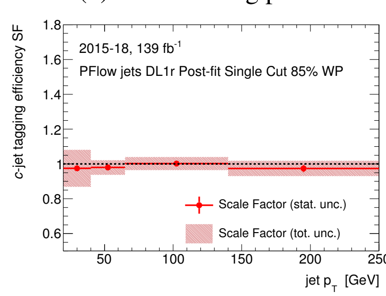
(a) 60% working point



(b) 70% working point

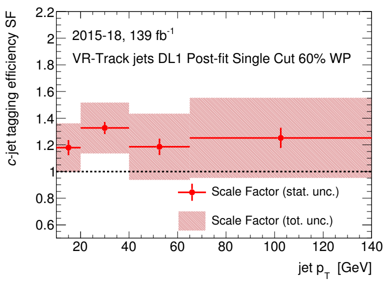


(c) 77% working point

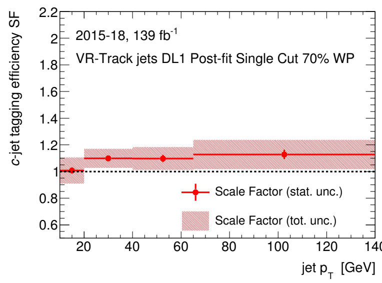


(d) 85% working point

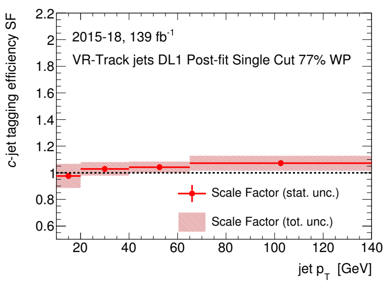
Figure 6.17: c-jet mis-tagging scale factors for the PFlow jets collection with the DL1r tagger.



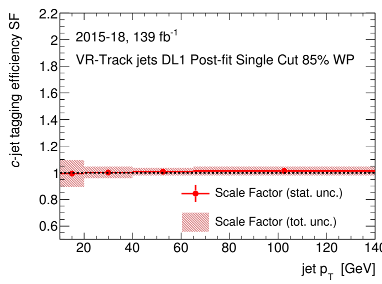
(a) 60% working point



(b) 70% working point

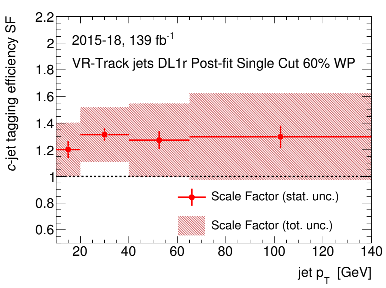


(c) 77% working point

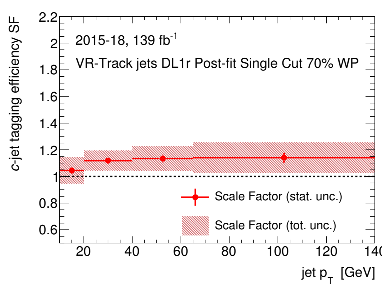


(d) 85% working point

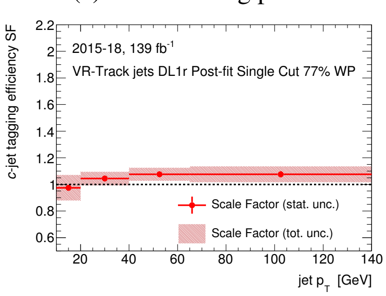
Figure 6.18: c-jet mis-tagging scale factors for the VR-Track jets collection with the DL1 tagger.



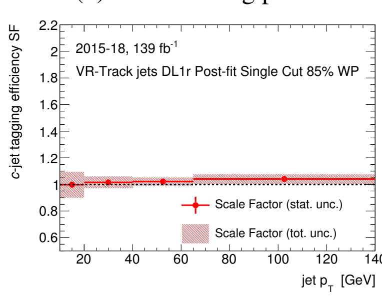
(a) 60% working point



(b) 70% working point



(c) 77% working point



(d) 85% working point

Figure 6.19: c-jet mis-tagging scale factors for the VR-Track jets collection with the DL1r tagger.

Chapter 7

Search for Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ channel

7.1 Introduction

This chapter describes the search for Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ channel, where one Higgs boson decays to a b quark pair and the other to a τ-lepton pair. As the two τ-leptons decay either leptonically or hadronically, the analysis is divided further into two sub-channels depending on their decay mode, the $b\bar{b}\tau^{\pm}_{\mathrm{lep}}\tau^{\mp}_{\mathrm{had}}$ channel (for simplicity, referred to as $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel) where one of the τ-leptons decays leptonically and the other decays hadronically, and the $b\bar{b}\tau^{+}_{\mathrm{had}}\tau^{-}_{\mathrm{had}}$ channel (or referred to as $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel) where both of the τ-leptons decay hadronically. The decay mode where both τ-leptons decay leptonically is not considered in this analysis due to its insignificant contribution. In this thesis, the author will present his work in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel, and will also show the combination results with the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel.

The results of this search are interpreted in terms of resonant and non-resonant production of the di-Higgs. For the non-resonant production, upper limits are set on the SM di-Higgs production cross-section, and exclusion limits are set on the Higgs self-coupling λ_{HHH}. The non-resonant search is also interpreted assuming 7 HEFT benchmark models, and one-dimensional scans are performed on two Wilson coefficients c_{gghh}, c_{tthh}. For the resonant search, upper limits are set on the resonance production cross-section as a function of the resonance mass, targeting a generic spin-0 neutral scalar resonance.

The $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ combination results are presented in Ref. [135]. The non-resonant and resonant search and the one-dimensional scan on λ_{HHH} results are combined with other final states of the di-Higgs production, including $b\bar{b}b\bar{b}$ and $b\bar{b}\gamma\gamma$. The combined result is presented in Ref. [192]. The upper limits on the 7 HEFT benchmarks and the exclusion limits on c_{gghh}, c_{tthh} are combined with the $b\bar{b}\gamma\gamma$ and are presented in Ref. [48].

The analysis strategy is as follows: first, a set of selections are applied on the reconstructed physics objects and kinematics variables, which define the signal regions and the control region, as described in Section 7.2; neural networks trained on the signal regions events are then used to extract the various signals, and the output of the algorithm is used as the final discriminant, as described in Section 7.4. The estimations of the various backgrounds considered in this analysis, in particular, the fake-τ_{had} background are detailed in Section 7.3. The systematic uncertainties considered in this analysis are discussed in Section 7.5; a profile likelihood fit is then performed simultaneously on all $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ signal regions and the control region used to constrain the various background contributions, with all systematic uncertainties served as nuisance parameters. The statistical combination result of $b\bar{b}\tau^+\tau^-$, $b\bar{b}\gamma\gamma$ and $b\bar{b}b\bar{b}$ is also studied. The setup and the results are shown in Section 7.6.

The author's contributions to the analyses presented in this chapter are as follows. As one of the main analysers in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the author produced all data and MC (including all signals and background) histograms with full experimental and theoretical systematics, which are the inputs to the likelihood fit. The author has derived the systematic uncertainties for the following major backgrounds: $t\bar{t}$, Z + HF and single-top. The author has also contributed to the estimation of the fake-τ_{had} background and its uncertainties. Finally, the author implemented and validated a reweighting method in order to scan the non-resonant HH production varying the self-coupling modifier κ_λ and two Wilson coefficients, c_{gghh} and c_{ttHH}, and set upper limits on the production cross-sections assuming 7 benchmark models. The corresponding uncertainties are also derived by the author.

7.2 Trigger and event Selection

7.2.1 Trigger selection

The ATLAS trigger system, as described in Section 3.2.6, consists of the Hardware-based level (L1) triggers and the software-based level triggers (HLT) to select characteristic events. Events in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel are recorded using a combination of single-lepton triggers (SLTs) and lepton-plus-τ_{had} triggers (LTTs). The offline objects are required to pass additional p_{T} cuts. For the leptons (e or μ), a threshold that is 1 GeV higher than the HLT p_{T} threshold is applied to reach its plateau region, where the cut efficiency is approximately constant. For the τ_{had}, the threshold is set at 5 GeV above the HLT p_{T} requirement for the same reason. The offline e, μ and τ_{had} must be matched to the corresponding trigger objects.

The priority is given to SLT events if a lepton fulfils the offline lepton p_{T} requirements.

Single Lepton Triggers (SLT)		
Period	**Single Electron Triggers (SET)**	**Single Muon Triggers (SMT)**
2015	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 HLT_mu50
2016 & 2017 & 2018	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0	HLT_mu26_ivarmedium HLT_mu50

Table 7.1: SLT triggers used in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel, along with the trigger-dependent offline p_{T} thresholds, for each year/period are shown. List taken from official ATLAS recommendation for triggers [193].

A combination of various un-prescaled[1] single-electron triggers (SET) and single-muon triggers (SMT) are used to maximise the efficiency.

The SLT triggers used are listed in Table 7.1. In this table, the following naming conventions are used: the trigger names begin with HLT, followed by the lepton type ('e' or 'mu' for an electron or muon respectively) and the HLT p_{T} threshold. What comes next is the identification WP ('loose', 'medium' or 'tight'), prefixed with 'lh', which stands for the likelihood-based trigger. The suffix of 'nod0' indicates that no impact parameter cuts are applied. For SMT, no identification is applied and therefore it is not specified. The following is the isolation requirement, specified by 'i' and the working point ('varloose', 'varmedium'). If not specified, no isolation requirement is applied. Finally, the suffix which starts 'L1EM' ('L1MU') indicates the object is seeded with L1 electromagnetic (muonic) trigger items, with p_{T} threshold specified by the number after; the 'VH' indicates that the p_{T} threshold varies with eta to account for energy loss and hadronic core isolation is applied. If not specified, the L1 seed is using the default setting.

As the p_{T} threshold is raised by 1 GeV for the HLT p_{T} requirements, the single electron triggers (SET) require an event with an electron with $p_{\mathrm{T}} > 25$ or > 27 GeV, whereas the single muon triggers (SMT) require an event with a muon with $p_{\mathrm{T}} > 21$ or > 27 GeV, depending on the data-taking period.

The LTT triggers used are listed in Table 7.2. For the LTT triggers, an electron (muon) is required with $p_{\mathrm{T}} > 18$ (15) GeV, together with a τ_{had} with $p_{\mathrm{T}} > 30$ GeV or 40 GeV depending on the period. The triggers also have L1 requirements on the jet p_{T} to reduce trigger rate. The turn-on curve for L1 is slow, therefore offline cuts are applied on the jet at 80 (45) GeV for L1 requirements of 25 (12) GeV. The same naming conventions are used for the LTT triggers. In addition, the trigger names are expanded by the 'tau' which stands for τ_{had} followed by its p_{T}, identification requirements and pre-selection settings. Again, unless specified, the default L1 seeds are used. After the lepton, the number of τ_{had}

[1]The trigger rate can be reduced by *prescaling* which randomly vetos events that pass the trigger.

is specified by a number before 'TAU', followed by a number indicating the p_T threshold. For example, '2TAU12I' stands for two τ_{had} with $p_T > 12$, with electromagnetic isolation applied ('I'). Finally, the number of jets and the p_T requirements are specified in the same fashion. If unspecified, at least one jet with $p_T > 25$ GeV is required (the offline cut is raised to 80 GeV).

Lepton Tau Triggers (LTT)
Electron Tau Triggers (ETT)
HLT_e17_lhmedium_nod0_tau25_medium1_tracktwo
HLT_e17_lhmedium_nod0_ivarloose_tau25_medium1_tracktwo
HLT_e17_lhmedium_nod0_ivarloose_tau25_medium1_tracktwo_L1EM15VHI_2TAU12IM_4J12
HLT_e17_lhmedium_nod0_ivarloose_tau25_medium1_tracktwoEF
HLT_e17_lhmedium_nod0_ivarloose_tau25_medium1_tracktwoEF_L1EM15VHI_2TAU12IM_4J12
HLT_e17_lhmedium_nod0_ivarloose_tau25_mediumRNN_tracktwoMVA
HLT_e17_lhmedium_nod0_ivarloose_tau25_mediumRNN_tracktwoMVA_L1EM15VHI_2TAU12IM_4J12
Muon Tau Triggers (MTT)
HLT_mu14_tau25_medium1_tracktwo
HLT_mu14_ivarloose_tau25_medium1_tracktwo
HLT_mu14_ivarloose_tau25_medium1_tracktwo_L1MU10_TAU12IM_3J12
HLT_mu14_ivarloose_tau25_medium1_tracktwoEF_L1MU10_TAU12IM_3J12
HLT_mu14_ivarloose_tau25_mediumRNN_tracktwoMVA_L1MU10_TAU12IM_3J12
HLT_mu14_ivarloose_tau35_medium1_tracktwo
HLT_mu14_ivarloose_tau35_medium1_tracktwoEF
HLT_mu14_ivarloose_tau35_mediumRNN_tracktwoMVA

Table 7.2: LTT triggers used in the $\tau_{lep}\tau_{had}$ channel. For different periods, a mixture of triggers from this list are used. Table reproduced from analysis internal note. List taken from official ATLAS recommendation for triggers [193].

7.2.2 Event pre-selection

Before passing the events to the multivariate algorithm for signal extraction, they are required to pass a loose *pre-selection* criteria to select events consistent with the signal topology and to remove obvious backgrounds. Events passing the pre-selection criteria are referred to as signal region (SR) events.

At least one primary vertex is required. Events are also required to contain exactly one electron or muon, an oppositely charged τ_{had} and exactly two b-tagged jets. The selected electron (muon) must pass a tight (medium) identification requirement with an efficiency of around 80% (97%) [93, 98].

To reject background events from low-mass Drell-Yan events, the invariant mass of the τ-lepton pair ($m_{\tau\tau}^{\text{MMC}}$), is required to be above 60 GeV, which is estimated from the four-momenta of the electron or muon, the τ_{had} and the E_T^{miss} using the Missing Mass Calculator (MMC) [194].

To reject $t\bar{t}$ events, the b-tagged jet pair invariant mass (m_{bb}) is required to be less than 150 GeV. In addition, by reversing this cut, a $t\bar{t}$-enriched region can be defined which is used in the estimation of fake-τ_{had} backgrounds as described in Section 7.3.1. A τ_{had} with $p_{\text{T}} > 20$ GeV and $|\eta| < 2.3$ is required in the SLT category. A τ_{had} with $p_{\text{T}} > 30$ GeV, or higher if required by the LTT triggers (as defined in Section 7.2.1), and $|\eta| < 2.3$ is required in the LTT events. In the SLT channel, the (sub-)leading b-tagged jet must have $p_{\text{T}} > 45$ (20) GeV, while in the LTT channel, the jet p_{T} requirements are trigger-dependent, as defined in the previous section. The full event pre-selection is summarised in Table 7.3.

$\tau_{\text{lep}}\tau_{\text{had}}$ categories	
SLT	LTT
e/μ selection	
Exactly one tight e or medium μ	
$p_{\text{T}}^{e} > 25, 27$ GeV	18 GeV $< p_{\text{T}}^{e} <$ SLT cut
$p_{\text{T}}^{\mu} > 21, 27$ GeV	15 GeV $< p_{\text{T}}^{\mu} <$ SLT cut
$|\eta^{e}| < 2.47$, not $1.37 < |\eta^{e}| < 1.52$	
$|\eta^{\mu}| < 2.7$	
τ_{had} selection	
One loose τ_{had}	
$|\eta| < 2.3$	
p_{T}> 20 GeV	p_{T}> 30 GeV
Jet selection	
≥ 2 jets with $|\eta| < 2.5$	
p_{T}> 45 (20) GeV	Trigger dependent
Event-level selection	
Trigger requirements passed	
Collision vertex reconstructed	
$m_{\tau\tau}^{\text{MMC}} > 60$ GeV	
Opposite-sign electric charges of $e/\mu/\tau_{\text{had}}$ and τ_{had}	
Exactly two b-tagged jets	
$m_{bb} < 150$ GeV	

Table 7.3: Summary of the event pre-selections, shown separately for the SLT and LTT. Thresholds on the (sub-)leading p_{T} object are given outside (within) parentheses. The possible values of the requirements in the SLT are separated by commas which depends on the year of the data-taking. For the jet selection in the LTT channel multiple selection criteria are used. The trigger p_{T} thresholds shown correspond to the offline requirements. Table reproduced from Ref. [135].

The fraction of events accepted by the detector is quantified by the *acceptance* A, and the fraction of events selected by the analysis selection is quantified by the *selection efficiency* ϵ. These two rates are usually multiplied together ($A \times \epsilon$) to quantify the rate of a simulated event to be accepted by the detector and pass the analysis selection. The cumulative $A \times \epsilon$ in each step of the pre-selection is summarised in Table 7.4 and Table 7.5 for the SLT and LTT channels, respectively, for the non-resonant signal and three example resonance mass points.

The $A \times \epsilon$ for all resonance mass points, m_X are shown in Figure 7.1. The decrease in $A \times \epsilon$ for m_X greater than about 1000 GeV is due to the boost of the Higgs bosons causing their decay products to become highly collimated more often, and therefore harder

Selection	Non-resonant signal ggF HH	VBF HH	Resonant signal (300 GeV)	(500 GeV)	(1000 GeV)
Basic selection	19%	16%	13%	22%	30%
Trigger	12%	9.2%	6.0%	14%	22%
Object selections	9.7%	7.2%	5.0%	11%	20%
Trigger specific offline p_{T} cuts	9.5%	6.8%	4.4%	11%	20%
Opposite-charged τ and lepton	9.4%	6.6%	4.4%	11%	20%
Two b-tagged jets	4.4%	2.7%	1.8%	4.9%	10%
$m_{\tau\tau}^{\mathrm{MMC}} > 60$ GeV	4.3%	2.7%	1.8%	4.8%	9.7%
$m_{bb} < 150$ GeV	4.1%	2.6%	1.7%	4.6%	9.2%

Table 7.4: Cumulative $A \times \epsilon$ for simulated signal events to pass each stage of the event pre-selection in the SLT channel. The efficiencies are calculated with respect to $HH \to b\bar{b}\tau^+\tau^-$ decays in which one τ-lepton decays hadronically, and one decays leptonically. The term 'Basic selection' refers to selections with at least one τ_{had} candidate and one lepton passing loose kinematic requirements. The 'Object selections' requires exactly one τ_{had} candidate, at least two jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. The 'Trigger specific offline p_{T} cuts' are cuts placed on the p_{T} of the reconstructed jet or τ_{had} that is geometrically matched to the HLT objects to ensure the efficiencies of the HLT objects reach the plateau region.

Selection	Non-resonant signal ggF HH	VBF HH	Resonant signal (300 GeV)	(500 GeV)	(1000 GeV)
Basic selection	19%	16%	13%	22%	30%
Trigger	3.4%	3.2%	2.8%	4.0%	3.4%
Object selections	2.9%	2.3%	2.2%	2.8%	2.4%
Trigger specific offline p_{T} cuts	2.3%	1.8%	1.4%	2.6%	2.3%
Opposite-charged τ and lepton	2.3%	1.8%	1.4%	2.6%	2.3%
Two b-tagged jets	1.1%	0.76%	0.61%	1.2%	1.1%
$m_{\tau\tau}^{\mathrm{MMC}} > 60$ GeV	1.0%	0.75%	0.61%	1.2%	1.1%
$m_{bb} < 150$ GeV	0.98%	0.71%	0.57%	1.1%	1.0%

Table 7.5: Cumulative $A \times \epsilon$ for simulated signal events to pass each stage of the event pre-selection in the LTT channel. The efficiencies are calculated with respect to $HH \to b\bar{b}\tau^+\tau^-$ decays in which one τ-lepton decays hadronically and one decays leptonically. The term 'Basic selection' refers to selections with at least one τ_{had} candidate and one lepton passing loose kinematic requirements. The 'Object selections' requires exactly one τ_{had} candidate, at least two jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. The 'Trigger specific offline p_{T} cuts' are cuts placed on the p_{T} of the reconstructed jet or τ_{had} that is geometrically matched to the HLT objects to ensure the efficiencies of the HLT objects reach the plateau region.

to identify.

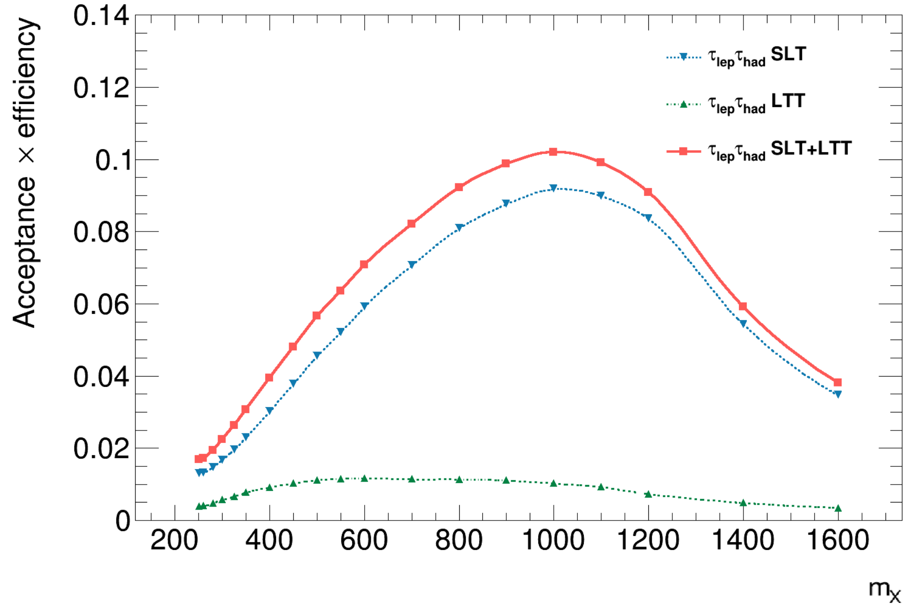


Figure 7.1: $A \times \epsilon$ for the resonant Di-Higgs production as a function of the resonance mass m_X in SLT, LTT and SLT LTT combined.

After the pre-selection, the event yields of the data, signal and background and the corresponding statistical error are shown in Table. 7.6 for the SLT channel, and Table. 7.7 for the LTT channel. The 'Fake' background in the table represents the background due to a jet faking a τ_{had}, which is estimated by a data-driven method described in detail in Section 7.3.1.

7.2.3 Anti–τ_{had} selection

In order to provide fake–τ_{had}–enriched regions used for background estimation, an "anti–τ_{had}" selection is defined: those τ_{had} objects that fail the RNN Loose τ_{had}–ID and have an RNN score greater than 0.01 are labeled as anti–τ_{had} candidates. The RNN cut is used and recommended by the Fake-Tau-Task-Force [195]. The minimum RNN score cut ensures that the jet has features somewhat similar to a true τ_{had} and hence that the composition of the jet (either quark- or gluon-initiated) would be more similar to that in the SR.

Exactly one anti–τ_{had} object is selected when there are no τ_{had} passing the offline τ_{had}–ID requirement. This is to make sure only one τ_{had} object (either true τ_{had} or anti–τ_{had}) is selected. For the LTT channel where τ_{had}–ID is applied at trigger level (more details in Section 7.2.1), only the anti–τ_{had} object that is matched to the trigger τ_{had} is considered, and thus there are no multiple selection possibilities. However, for the SLT channel where a τ_{had} trigger is not used, an anti–τ_{had} candidate is chosen randomly when there are more

SampleName	N_{MC}	Yield
Signal Samples		
ggF Non-resonant	151522	5.872 ± 0.018
VBF Non-resonant	36457	0.2002 ± 0.0013
Resonant at m_X =300 GeV	18398	77.2 ± 0.6
Resonant at m_X =500 GeV	21332	210.7 ± 1.5
Resonant at m_X =1000 GeV	35119	425.2 ± 2.3
Resonant at m_X =1600 GeV	14421	160.6 ± 1.4
Background Samples		
Fake	2089776	33920 ± 110
$t\bar{t}$	490058	61620 ± 90
single top	33158	3689 ± 32
$Z \to \tau\tau + (cc, bc, bb)$	23827	1316 ± 59
other	31839	1150 ± 120
SM Higgs	234087	154 ± 2
total bkg	3054267	102440
data	98456	98456

Table 7.6: Numbers of MC simulation events and pre-fit event yields in the di-Higgs $bb\tau_{\text{lep}}\tau_{\text{had}}$ SLT signal region for the data, background and ggF, VBF non-resonant HH signal along with a few resonance signals as examples. The 'Fake' background represents the background due to a jet faking a τ_{had} and the 'other' background includes the W/Z+jets and the diboson background contributions. The largest and second-largest contributions to the 'SM Higgs' background are a Higgs boson produced via ttH and ggF production, respectively.

SampleName	N_{MC}	Yield
Signal Samples		
ggF Non-resonant	35045	1.416 ± 0.009
VBF Non-resonant	10188	0.0548 ± 0.0006
Resonant at m_X =300	6319	26.48 ± 0.34
Resonant at m_X =500	5177	51.7 ± 0.7
Resonant at m_X =1000	3820	47.4 ± 0.8
Resonant at m_X =1600	1412	16.1 ± 0.4
Background Samples		
Fake	41151	1752 ± 33
$t\bar{t}$	32873	4207 ± 24
single top	1715	197 ± 8
$Z \to \tau\tau + (cc, bc, bb)$	8079	439 ± 29
other	1997	97 ± 29
SM Higgs	30508	24 ± 1
total bkg	117806	6876.79
data	6351	6351

Table 7.7: Numbers of MC simulation events and pre-fit event yields in the di-Higgs $bb\tau_{\rm lep}\tau_{\rm had}$ LTT signal region for signal, data, and background. The naming conventions are the same as Table 7.6.

reconstructed $\tau_{\rm had}$ satisfying the anti–$\tau_{\rm had}$ definition. Any anti–$\tau_{\rm had}$ objects that are not selected in this process are also not considered when performing the overlap removal of detector objects, as discussed in Section 4.7. Derived variables used in the analysis, such as the $E_{\rm T}^{\rm miss}$, $m_{\tau\tau}^{\rm MMC}$ and $E_{\rm T}^{\rm miss}\phi$ centrality (more details in Section 7.4) are calculated in the same way as for signal events, but with the anti-$\tau_{\rm had}$ taking the place of the loose $\tau_{\rm had}$ candidate.

7.2.4 Z + HF control region event selection

The Z boson production in association with heavy flavour jets, i.e. b-, c-jets (Z + HF background) is known to be not well modelled by the Sherpa generator. Therefore, a dedicated control region event selection is defined to select $Z \to \mu\mu/ee$ + heavy flavour jets events. Since the production of jets is independent of the decay mode of the Z boson, this selection provides an orthogonal region with high purity to the SR, which requires two τ leptons in the final state. The normalisation is extracted from this region, and it also provides constraints on the normalisation of the $t\bar{t}$ background.

The event selection is defined as follows:

- Events are selected using the single-lepton and di-lepton triggers, as specified in Ref. [196].

- Exactly two leptons (e/μ) with opposite-sign charges and $p_{\mathrm{T}} > 9$ GeV are required, these leptons are also required to be compatible with the primary vertex. In addition, they are required to pass the medium and loose isolation requirements.

- Exactly two b-tagged jets (tagged with DL1r 77% working point).

- The reconstructed mass of the two leptons is required to be between 75 GeV and 110 GeV, to be consistent with the Z boson mass.

- $m_{bb} < 40$ GeV or $m_{bb} > 210$ GeV to veto Higgs mass peak.

The data and background $m_{\ell\ell}$ distributions are shown in Figure 7.2. In this figure, the normalisation and shape of the backgrounds and the uncertainty on the total background are shown as determined from the likelihood fit to data in the non-resonant HH search, as described in Section 7.6.3. The uncertainty band includes statistical and systematic uncertainties on the total background. The 'Top-quark' background includes the $t\bar{t}$ and the single-top background; the 'other' background includes the W/Z bosons + jets and the diboson background. The dashed histogram shows the total pre-fit background, which shows a large deviation from the data due to the mis-modelling issue in Z+HF background. After performing the fit and applying the extracted normalisation and shape correction, simulation and data agree well within the uncertainties.

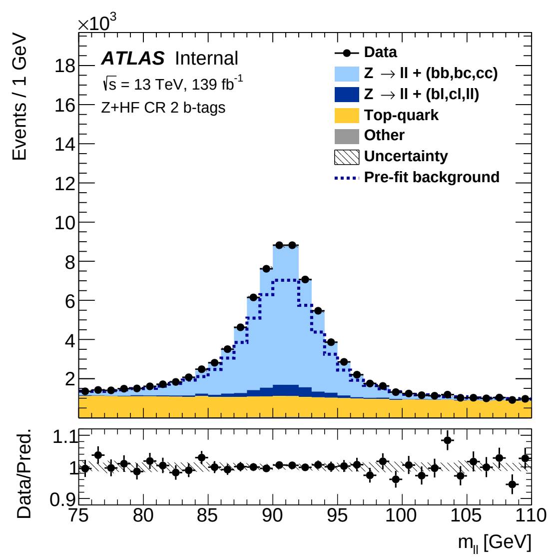


Figure 7.2: Post-fit data and background $m_{\ell\ell}$ distributions in the Z + HF control region. Figure taken from Ref. [135].

7.3 Background estimation

This section describes the background estimation methods used in the di-Higgs analysis. The dominant background is the $t\bar{t}$ background, followed by the single top and the $Z+$HF backgrounds, as shown in Table 7.6 and 7.7. The simulated event samples summarised in Section 5.3 are used to model all background processes, except for the fake-τ_{had} background where a jet fakes a τ_{had}. It is estimated using data-driven techniques as discussed below. In particular, the lepton faking τ_{had} background is modelled by simulation, as its contribution is found to be very small compared to the jet faking τ_{had} background. The $t\bar{t}$ with true-τ_{had} and $Z+$HF templates are taken from the MC prediction but their normalisations are derived from data as included as freely floating parameters in the final fit, as described in Section 7.6.

7.3.1 Background with a jet misidentified as a τ_{had}

The fake-τ_{had} background can have different origins. In Figure 7.3, two Feynman diagrams are shown for the two dominant processes contributing to the fake-τ_{had} background, which are the $t\bar{t}$ and multi-jet (referred to as QCD) processes.

In the $t\bar{t}$ events, the fake-τ_{had} background typically originates from quark-initiated jets from top quark decay; in multi-jet events, jets initiated from both quarks and gluons can be misidentified as τ_{had}. In the following text, the fake background initiated by the $t\bar{t}$ (multi-jet) events is referred to as $t\bar{t}$ (multi-jet) fakes.

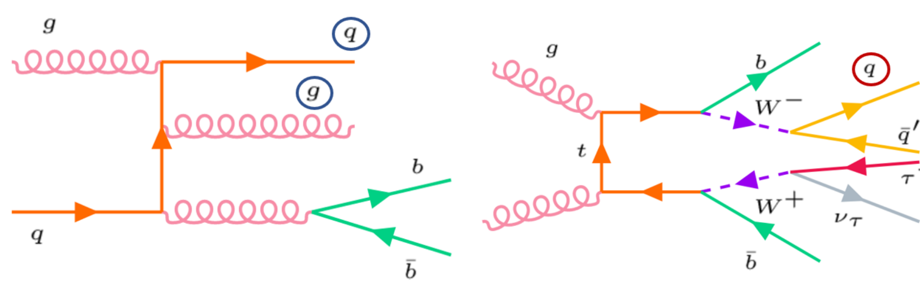


Figure 7.3: Feynman diagrams on the left (right) for the multi-jet ($t\bar{t}$) originated fake-τ_{had} background. The gluon and quark circled in blue in the left diagram fake a τ_{had} and a lepton, respectively, and the quark circled in red in the right diagram fakes a τ_{had}.

7.3.1.1 Fake factor method

The fake-τ_{had} background events are estimated using a data-driven method, the 'fake factor' (FF) method, due to the imperfect simulation of these processes. In short, the fake factor is the ratio of the number of events with fake-τ_{had} in one region to another region. In this analysis, the fake factor for a given background source is defined as:

$$\mathrm{FF} = \frac{N(\text{ID selection})}{N(\text{anti-ID selection}),} \tag{7.1}$$

where the numerator is the number of fake-τ_{had} background events passing the nominal signal region τ_{had}-ID selection (referred to as ID selection in the following), and the denominator is the number of fake-τ_{had} background events passing the anti-τ_{had} selection (referred to as anti-ID selection, as defined in Section 7.2.3). To obtain a correct estimation of the N((anti-)ID /selection), events with a true τ_{had} are subtracted from the data events, i.e. $N = N(\text{data}) - N(\text{true } \tau_{\mathrm{had}}, \text{MC})$. The fake factor for a given source is calculated in its dedicated background-enriched control regions, and then fake factors calculated in different regions are combined and used to normalise the fake-τ_{had} background events distributions from the anti-ID selection to the ID selection.

The fake factors are derived separately for the SLT and LTT channels. Due to the different origins of the fake-τ_{had}, the FFs are calculated separately for the $t\bar{t}$ and multi-jet, and also separately for 1- and 3-prong τ_{had} candidates. The fake factor is parameterised in bins of p_{T} of the τ_{had}, while the dependence on η is also checked, but no obvious trend is observed.

The dedicated control region for each source is referred to as FF-CR. The FF-CR for each process is defined as follows:

- $t\bar{t}$ FF-CR: same selection as the ID/anti-ID selection but with m_{bb} cut reversed: m_{bb} > 150 GeV.
- Multi-jet FF-CR: same selection as the ID/anti-ID selection but with lepton isolation requirements reversed: ‘tight’ electrons and ‘medium’ muons are required to fail their respective ‘loose’ isolation working points.

The combined fake factor is determined using the fake factors calculated in the individual FF-CRs, defined as:

$$\mathrm{FF(comb)} = \mathrm{FF(multi-jet)} \times \mathrm{r_{QCD}} + \mathrm{FF}(t\bar{t}) \times (1 - \mathrm{r_{QCD}}), \tag{7.2}$$

where the FF(multi – jet) (FF($t\bar{t}$)) is the fake factor calculated in the multi-jet ($t\bar{t}$) FF-CR. The value of $\mathrm{r_{QCD}}$ is defined as the fraction of multi-jet fakes in the total number of fake-τ_{had} background. It is measured as a function of the τ_{had} p_{T}, split into 1-prong and 3-prong, and into the type of light lepton (e or μ) since an electron is more easily mis-indentified as a jet than a muon. $\mathrm{r_{QCD}}$ is measured for events passing the anti-ID selection as:

$$\mathrm{r_{QCD}} = \frac{N(\text{multi-jet, data})}{N(\text{data}) - N(\text{true } \tau_{\mathrm{had}}, \text{MC})} \tag{7.3}$$

where the N(multi-jet, data) is calculated by subtracting all background contributions apart from multi-jet, regardless of whether they contain fake or true-τ_{had} candidates, from the data in the anti-τ_{had} selection:

$$N(\text{multi-jet, data}) = N(\text{data}) - N(\text{true } \tau_{\text{had}}, \text{MC}) - N(\text{fake } \tau_{\text{had}}, \text{MC}) \quad (7.4)$$

The subtracted backgrounds are taken from the MC predictions, and the MC simulated fake background has solely the contribution from the $t\bar{t}$ fakes.

In graphical form, the various FF-CRs where the fake factors are measured and applied can be seen in Figure 7.4.

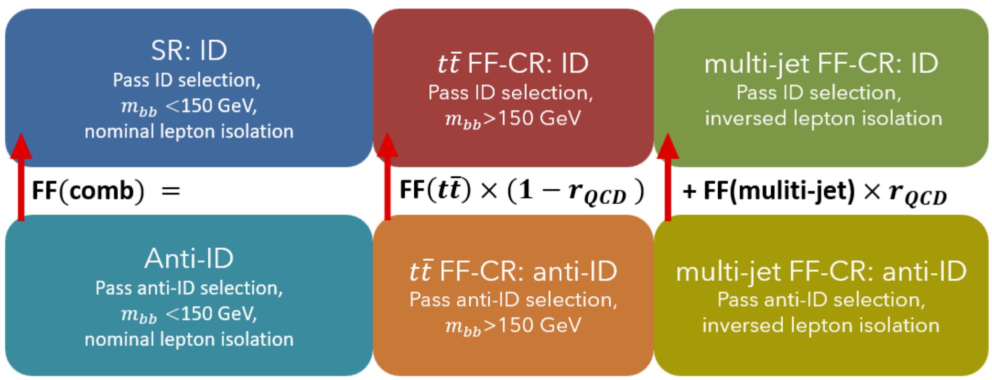


Figure 7.4: Graphical representation of the fake factor method. The fake factors are calculated independently for the $t\bar{t}$ FF-CR and the multi-jet FF-CR, and the combined fake factor is calculated using these fake factors and the value of r_{QCD}. The direction of the arrow indicates the direction of extrapolation, that the fake factor is applied on the bottom regions to extrapolate to the top regions.

7.3.1.2 $t\bar{t}$ background reweighting

The determination of the combined fake factor is sensitive to the modelling of simulated $t\bar{t}$ events with true-τ_{had} given that this is the dominant background that is subtracted from data in the derivation of the fake factors and r_{QCD}. The $t\bar{t}$ modelling also affects the fake background estimation in the SR, since the fake factor is applied on the anti-ID events where the true-τ_{had} $t\bar{t}$ background is subtracted from. It was observed that mis-modelling in the true $t\bar{t}$ background especially in the high jet multiplicity and high top-quark p_{T} regions can cause issues in the calculation of the fake factors, giving non-physical negative values in the high τ_{had} p_{T} region. To mitigate this issue, simulated events from $t\bar{t}$ production are differentially reweighted depending on the jet multiplicity and the scalar sum of the transverse momentum of all visible final state objects (H_T) in the event.

These reweighting factors are determined bin-by-bin in distribution of jet multiplicity and H_T, from another $t\bar{t}$ FF-CR, $t\bar{t}$ FF-CR2, which is defined using a selection identical to the SR selection, but with the $t\bar{t}$ FF-CR m_{bb} requirements ($m_{bb} > 150$ GeV) and an additional $m_{\mathrm{T}}^{W} > 40$ GeV requirement, with m_{T}^{W} defined in Section 7.4.1.

$m_{\mathrm{T}}^{W} > 40$ GeV requirement. Furthermore, events in this region are required to have a reconstructed τ_{had} candidate, but this candidate is not required to pass any RNN τ_{had} requirement. The m_{T}^{W} requirement is introduced to remove any potential contamination from multi-jet events. The reweighting factors are shown in Appendix B.1, Figure B.1

The reweighting method is validated in two additional validation regions. The reweighting only applies to the fake background estimation process while it's not applied to the true $t\bar{t}$ background in the SR, as there is no significant mis-modeling seen there, and the uncertainties due to the $t\bar{t}$ modelling have been taken care of using a different approach, as described in Section 7.5.2.1. The difference between the reweighted and original fake-τ_{had} background estimation is taken as a systematic uncertainty due to $t\bar{t}$ background modelling, with more details available in Section 7.5.4.

7.3.1.3 Fake factor calculation

The data and MC events in $t\bar{t}$ FF-CR passing the anti-ID selection before and after the $t\bar{t}$ reweighting are shown in Figures 7.5 and 7.6, respectively. The fake background is simulated by MC, and it only includes the contributions from the $t\bar{t}$-initiated fakes. A clear mis-modelling can be seen in the plots before the reweighting, and it is much mitigated after the reweighting process. At this stage, data and MC are not expected to agree well due to the $t\bar{t}$ fake is known to be poorly modelled by the MC, which is also the main reason for using the data-driven fake factor method. Nevertheless, a large number of $t\bar{t}$ fakes is observed, suggesting the high purity of $t\bar{t}$ fakes in this region. With true τ_{had} contributions subtracted from the data (and with true τ_{had} $t\bar{t}$ reweighted), events in this region are used in $t\bar{t}$ FF-CR fake factor calculation (as the denominator).

Similarly, the data and MC events in multi-jet FF-CR passing the anti-ID selection before and after the $t\bar{t}$ reweighting are shown in Figures 7.7 and 7.8, respectively. As the multi-jet fakes are not modelled by the MC, a large discrepancy is shown between data and the MC distributions, which is comprised of the multi-jet fakes. For consistency, the $t\bar{t}$ reweighting is applied, however, this region is dominated by the multi-jet fakes, and the $t\bar{t}$ contribution is negligible. The contribution from $t\bar{t}$ fakes is also very small, suggesting the purity of multi-jet fakes is high in this region. With true τ_{had} contributions subtracted from data, events in this region are used for the multi-jet FF-CR fake factor calculation (as the denominator).

The calculated fake factors are shown on Figure 7.9 and Figure 7.10, for the SLT and LTT channels respectively. The fake factors obtained with reweighted and un-weighted $t\bar{t}$ contributions are shown on the same graph.

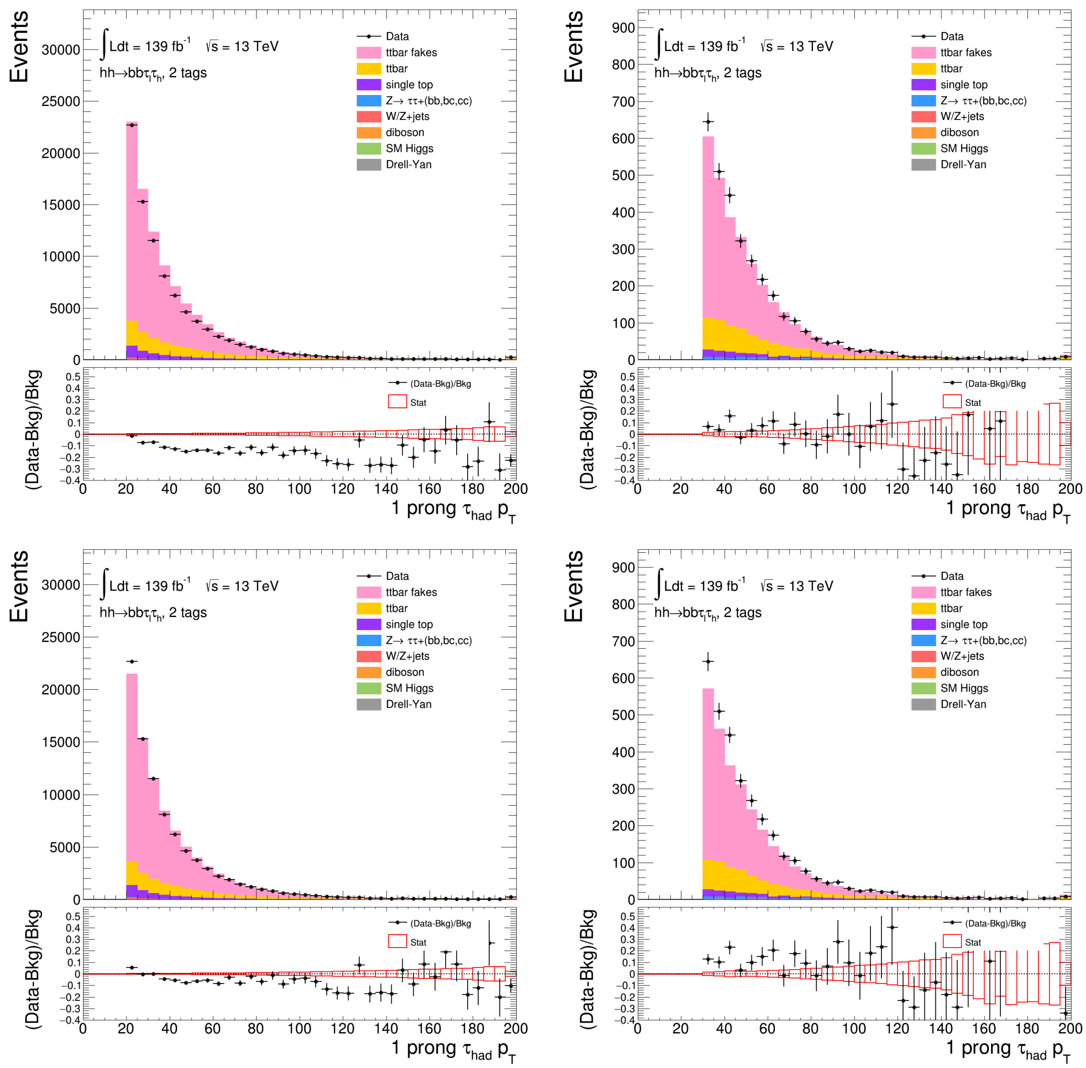


Figure 7.5: In the $t\bar{t}$ FF-CR, the τ_{had} p_T distributions are plotted for the SLT (left) and LTT channel (right) with $t\bar{t}$ un-weighted (top) and with $t\bar{t}$ re-weighted (bottom). The τ_{had} in these events has 1 prong. The MC $t\bar{t}$ fakes is labelled as 'ttbar fakes' in pink. Only statistical uncertainties are included in the error bands.

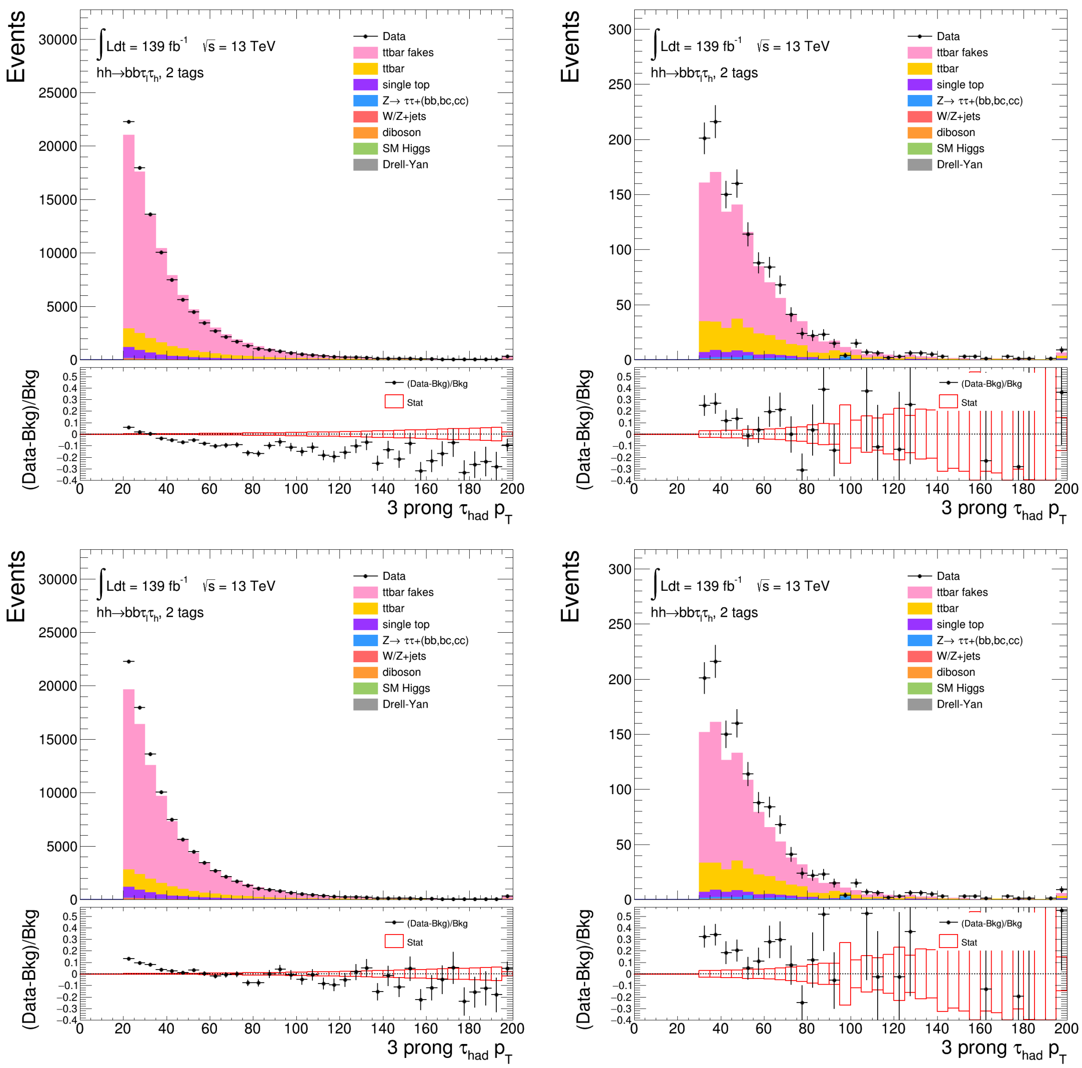


Figure 7.6: In the $t\bar{t}$ FF-CR, the τ_{had} p_T distributions are plotted for the SLT (left) and LTT channel (right) with $t\bar{t}$ un-weighted (top) and with $t\bar{t}$ re-weighted (bottom). The τ_{had} in these events has 3 prongs. The MC $t\bar{t}$ fake-τ_{had} background is labelled as 'ttbar fakes' in pink. Only statistical uncertainties are included in the error bands.

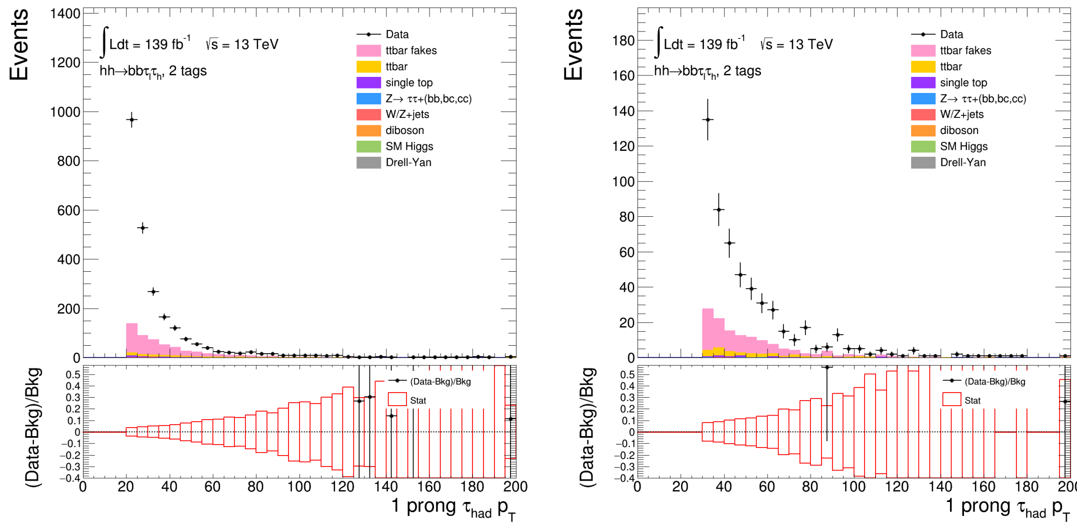


Figure 7.7: In the multi-jet FF-CR, the τ_{had} p_T distributions are plotted for the SLT (left) and LTT channel (right) with $t\bar{t}$ re-weighted. The τ_{had} in these events has 1 prong. The MC $t\bar{t}$ fake-τ_{had} background is labelled as 'ttbar fakes' in pink. Only statistical uncertainties are included in the error bands.

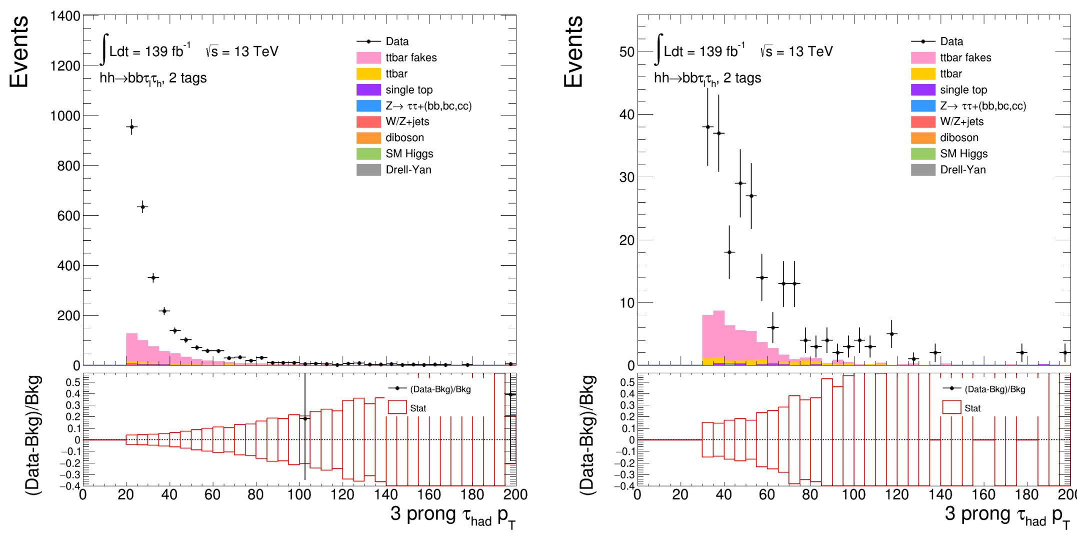


Figure 7.8: In the multi-jet FF-CR, the τ_{had} p_T distributions are plotted for the SLT (left) and LTT channel (right) with $t\bar{t}$ re-weighted. The τ_{had} in these events has 3 prongs. Only statistical uncertainties are included in the error bands.

The reweighting procedure has a minor impact on the multi-jet fake factor, as expected. It is found in the un-reweighted fake factor that the last bin of the $t\bar{t}$ fake factor is negative. To avoid an unphysical fake factor, these negative values are set to 0, which leads to poor estimation of the fake-τ_{had} background in high $\tau_{\rm had}$ $p_{\rm T}$ region. This problem is much mitigated by the reweighting procedure, where the fake factor is being pulled up visibly.

The binning is optimised for the $\mathrm{FF}_{t\bar{t}}$ and the same binning is used for the $\mathrm{FF}_{\rm QCD}$. A smooth and obvious trend is observed in the $\mathrm{FF}_{t\bar{t}}$, but some artefacts in the shapes at mid- and high-$p_{\rm T}$ range are shown in the $\mathrm{FF}_{\rm QCD}$. This issue has no visible impact on the fakes estimation because the $t\bar{t}$ FF dominates over the multi-jet FF in the combined FF, which can be seen in the $\mathrm{r}_{\rm QCD}$ distribution, as shown in Figures 7.11 and 7.12 for the SLT and LTT channels, respectively. These figures show the $\mathrm{r}_{\rm QCD}$ for 1-prong and 3-prong $\tau_{\rm had}$ candidates in $e\tau_{\rm had}$ and $\mu\tau_{\rm had}$ channels. The $\mathrm{r}_{\rm QCD}$ in general has small values in the SLT channel (in most bins the $\mathrm{r}_{\rm QCD}$ is 0), suggesting that the contribution from the multi-jet fake-τ_{had} background is small. While in the LTT, the $\mathrm{r}_{\rm QCD}$, in general, has higher values, due to the fact that the LTT channel focuses more on the low $p_{\rm T}$ region, where the multi-jet production is significant.

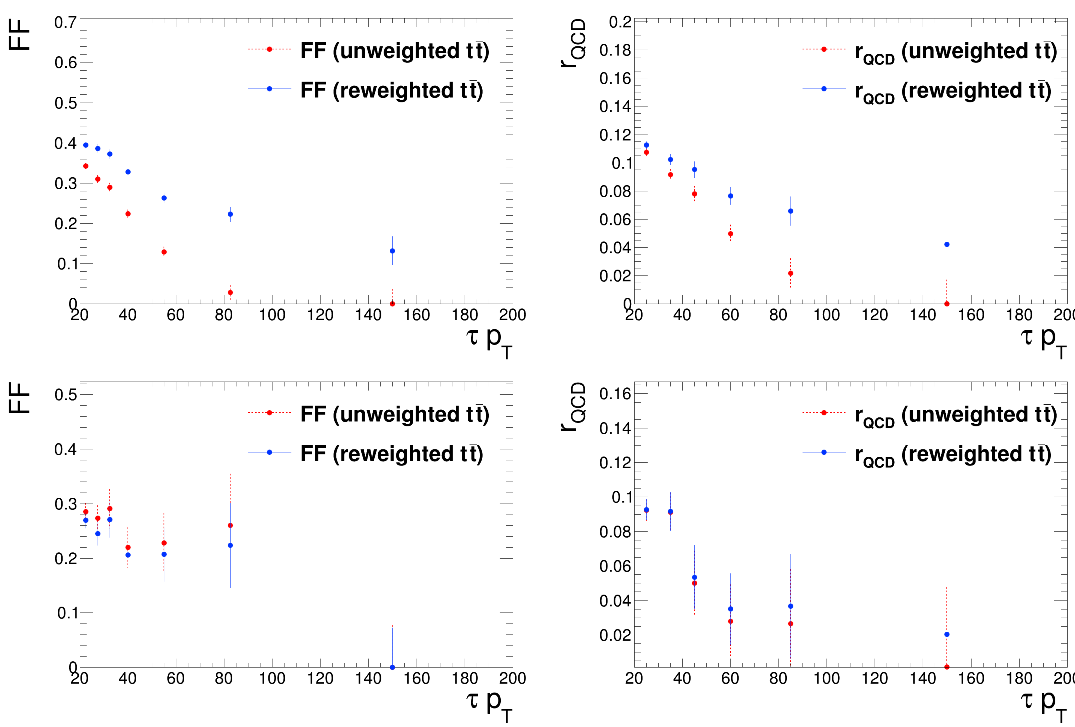


Figure 7.9: Fake factors calculated in $t\bar{t}$ FF-CR (top) and multi-jet FF-CR (bottom) for 1-prong (left) and 3-prong (right) $\tau_{\rm had}$ candidates for the SLT category.

Statistical uncertainties in $\mathrm{FF}_{t\bar{t}}$, $\mathrm{FF}_{\rm QCD}$ and $\mathrm{r}_{\rm QCD}$ are evaluated and propagated to the final result, and a conservative 30% modelling uncertainty is assigned to simulated non-$t\bar{t}$

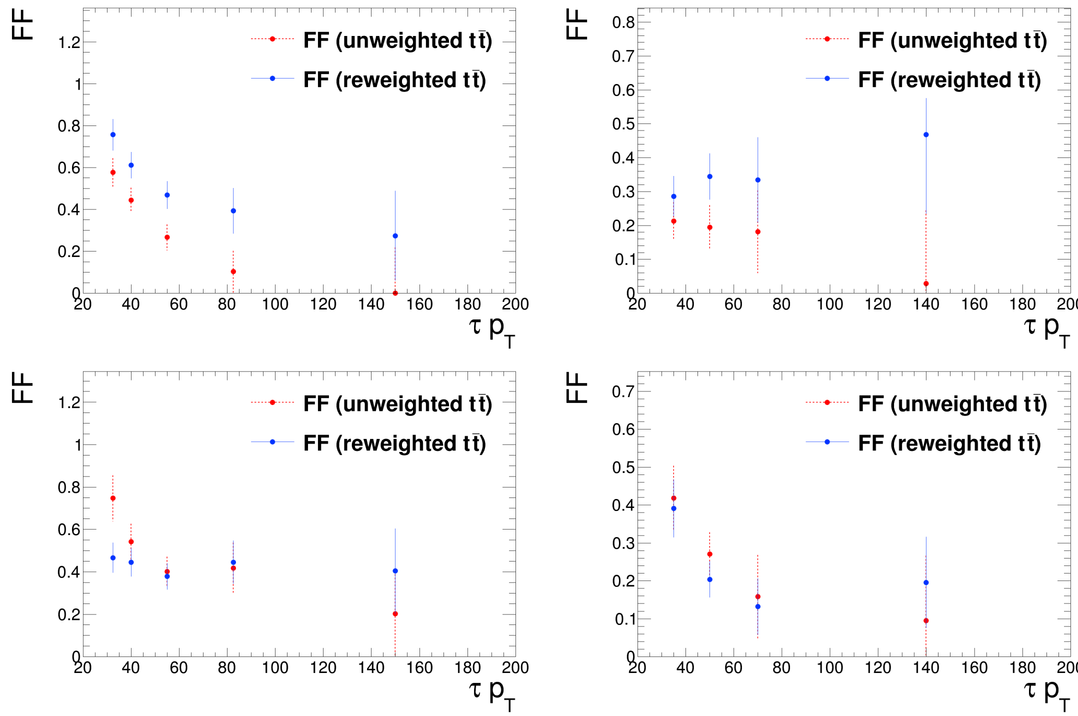


Figure 7.10: Fake factors calculated in $t\bar{t}$ FF-CR (top) and multi-jet FF-CR (bottom) for 1-prong (left) and 3-prong (right) τ_{had} candidates for the LTT category.

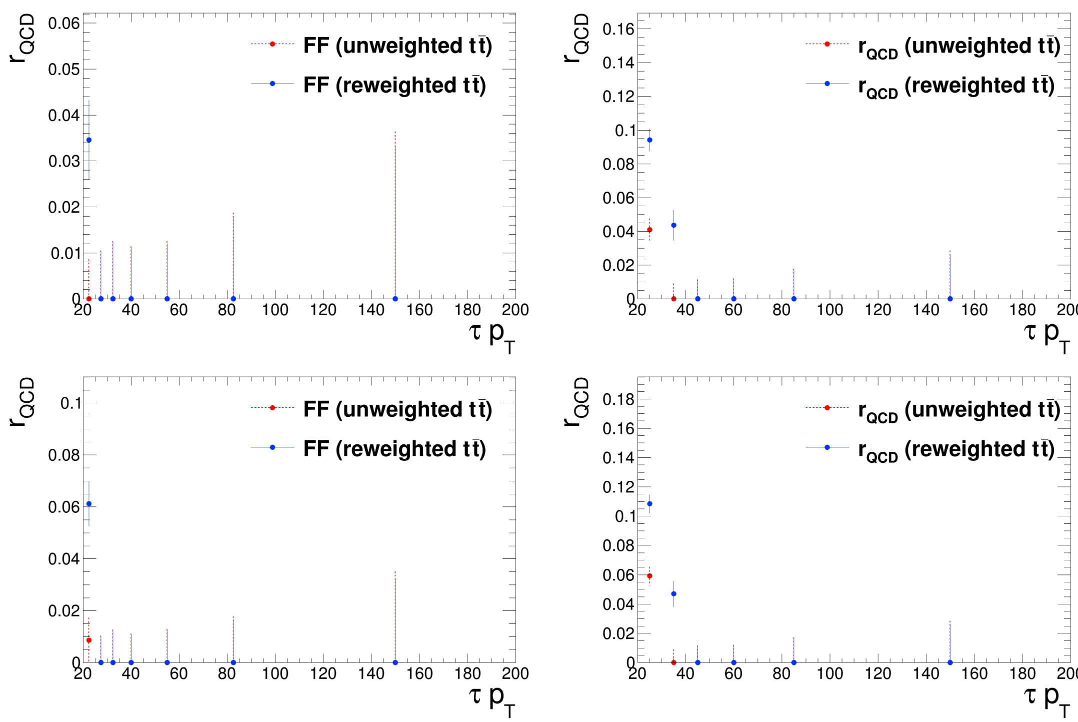


Figure 7.11: r_{QCD} for 1-prong (left) and 3-prong (right) τ_{had} candidates with $e\tau_{\text{had}}$ (top) and $\mu\tau_{\text{had}}$ (bottom) final states for the SLT channel.

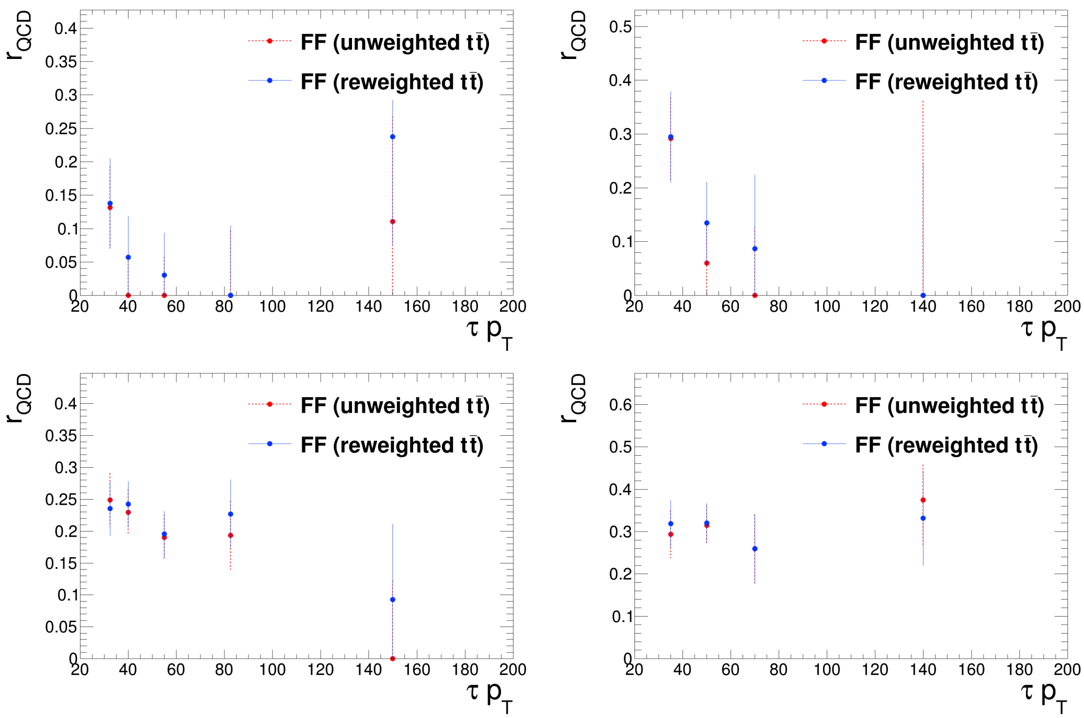


Figure 7.12: r_{QCD} for 1-prong (left) and 3-prong (right) τ_{had} candidates with $e\tau_{\text{had}}$ (top) and $\mu\tau_{\text{had}}$ (bottom) final states for the LTT channel.

backgrounds which are subtracted from data. The uncertainties due to the $t\bar{t}$ modelling issue and its subtraction are discussed in more detail in Section 7.5.4.

7.3.1.4 Fake factor method validation

The combined FF method is validated in the 0-b-tagged and 1-b-tagged regions, where the same event selection as SR is applied but with different numbers of b-tagged jets required. The fakes in each validation region are estimated with FF calculated in each validation region with the same method used in the 2-b-tagged region. These two validation regions are chosen as the signal contamination in the 0-b-tagged and 1-b-tagged regions is negligible, and the 0-b-tagged region can benefit from its rich statistics and the 1-b-tagged region can benefit from being closer to the SR. The estimated background distributions agree well with the observed distributions in all validation regions. The data and MC comparison with fakes estimated with the FF method are shown in Figure 7.13 for the SLT channel and Figure 7.14 for the LTT channel. For simplicity, only statistical uncertainties are included. Small fluctuations in these plots are not fully covered in the uncertainty band, which is acceptable as long as no shape dependence is observed.

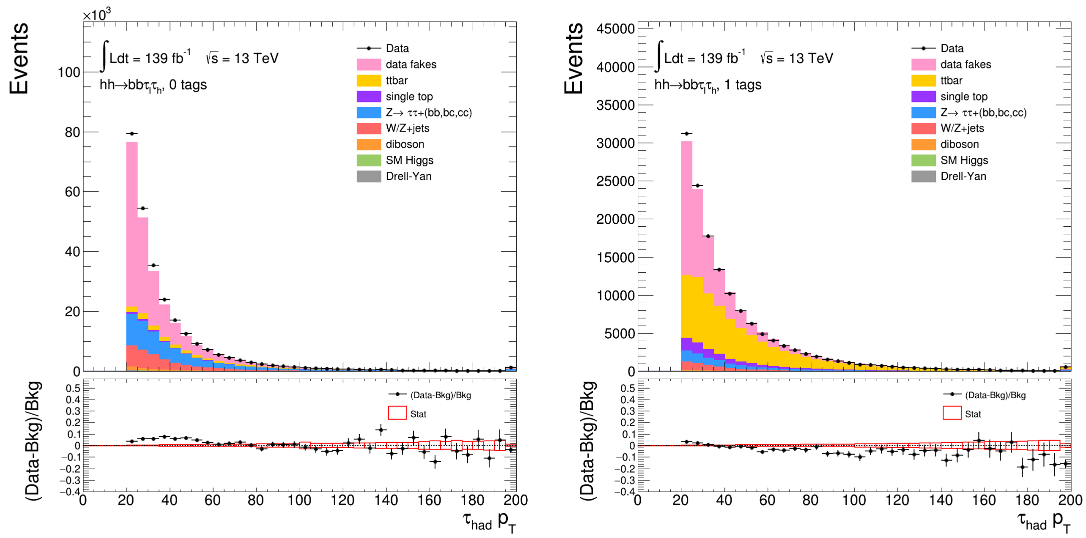


Figure 7.13: A comparison of data and background with fakes estimated with combined FF evaluated at the 0-b-tagged (left) and 1-b-tagged region (right) for the SLT channel. Only statistical uncertainties are included in the error bands.

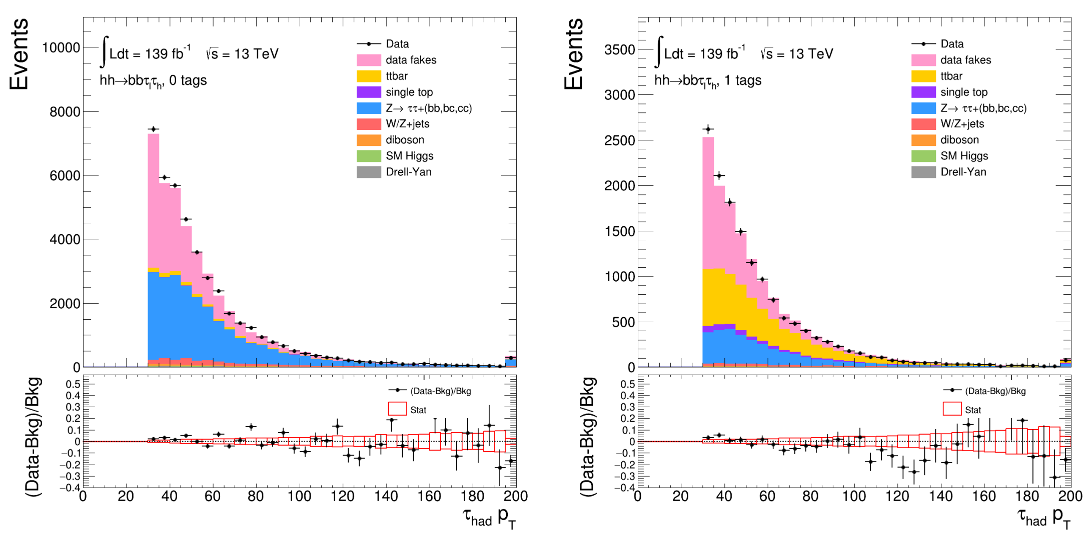


Figure 7.14: A comparison of data and background with fakes estimated with combined FF evaluated at the 0-b-tagged (left) and 1-b-tagged region (right) for the LTT channel. Only statistical uncertainties are included in the error bands.

7.4 Signal extraction

The signal extraction in the HH analysis is based on Multivariate techniques (MVA), and in particular, to extract the non-resonant HH signal, a neural network is used. The NN is trained with the signal and background defined in Section 5. During the training, the background due to a jet faking a τ_{had} is modelled by simulation, instead of the data-driven method described in Section 7.3.

7.4.1 MVA input variables

The same set of MVA input variables is used for the resonant and non-resonant production modes, though different input variables are used for the SLT and LTT channels. The choice of input variables is based on Ref. [27]. Three new high-level variables, $\Delta\phi(\ell\tau, E_{\mathrm{T}}^{\mathrm{miss}})$, $\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ and S_{T} are introduced, which have shown discriminating power in the LTT channel.

These variables are listed in Table 7.8 and are defined as follows:

- m_{HH} is the invariant mass of the HH system as reconstructed from the τ-lepton pair (calculated using the MMC) and the b-tagged jet pair;
- $m_{\tau\tau}^{\mathrm{MMC}}$ is the invariant mass of the di-τ system, calculated using the MMC.
- m_{bb} is the invariant mass of the di-b-jetsystem.
- $\Delta R(\tau, \tau)$ is evaluated between the electron or muon and the τ_{had};
- $\Delta R(b, b)$ is evaluated between the b-tagged jets;
- $\Delta p_{\mathrm{T}}(\ell, \tau)$ is the difference between the transverse momenta of the lepton and the τ_{had}
- $m_{\mathrm{T}}^{W} = \sqrt{2p_{\mathrm{T}}^{\ell} E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi_{\ell,E_{\mathrm{T}}^{\mathrm{miss}}})}$ is the transverse mass of the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$, with $\Delta\phi_{\ell,E_{\mathrm{T}}^{\mathrm{miss}}}$ defined in the following;
- the $E_{\mathrm{T}}^{\mathrm{miss}}$ ϕ centrality specifies the angular position of the $E_{\mathrm{T}}^{\mathrm{miss}}$ relative to the τ_{had} in the transverse plane [22] and is defined as $(A + B)/\sqrt{A^2 + B^2}$, where $A = \sin(\phi_{E_{\mathrm{T}}^{\mathrm{miss}}} - \phi_{\tau_2})/\sin(\phi_{\tau_1} - \phi_{\tau_2})$, $B = \sin(\phi_{\tau_1} - \phi_{\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}})/\sin(\phi_{\tau_1} - \phi_{\tau_2})$, and τ_1 and τ_2 represent the electron or muon and τ_{had};
- $\Delta\phi(\ell\tau, bb)$ is the azimuthal angle between the $\ell + \tau_{\text{had-vis}}$ system and the b-tagged jet pair;
- $\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ is the azimuthal angle between the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$;

- $\Delta\phi(\ell\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ is the azimuthal angle between the electron or muon and τ_{had} system and the $E_{\mathrm{T}}^{\mathrm{miss}}$;
- S_{T} is the total transverse energy in the event, summed over all jets, τ_{had} and leptons in the event and $E_{\mathrm{T}}^{\mathrm{miss}}$.

Variable	SLT	LTT
m_{HH}	✓	✓
$m_{\tau\tau}^{\mathrm{MMC}}$	✓	✓
m_{bb}	✓	✓
$\Delta R(\tau, \tau)$	✓	✓
$\Delta R(b, b)$	✓	
$\Delta p_{\mathrm{T}}(\ell, \tau)$	✓	✓
Sub-leading b-tagged jet p_{T} (p_T^{b2})	✓	
m_{T}^{W}	✓	
$E_{\mathrm{T}}^{\mathrm{miss}}$	✓	
$E_{\mathrm{T}}^{\mathrm{miss}}$ ϕ centrality	✓	
$\Delta\phi(\ell\tau, bb)$	✓	
$\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$		✓
$\Delta\phi(\ell\tau, E_{\mathrm{T}}^{\mathrm{miss}})$		✓
S_{T}		✓

Table 7.8: Variables used as inputs to the MVAs in the SLT and LTT channels. The same choice of input variables is used for the resonant and non-resonant production modes.

The data versus MC background comparison of these variables is shown in Figure 7.15 for the SLT channel and Figure 7.16 for the LTT channel, respectively. These figures are plotted after performing the fit and applying the extracted normalisation and shape correction. The uncertainty band includes the statistical uncertainty and all the systematic uncertainties, as described in Section 7.5. Good agreement between the data and the simulation is present, suggesting the background distributions are modelled well.

In both the NN used for non-resonant signal and the PNNs used for resonant signal, the three most significant and discriminating variables are m_{HH}, $m_{\tau\tau}^{\mathrm{MMC}}$ and m_{bb}.

7.4.2 MVA hyper-parameter optimisation and trainning

The input variables each have a different range of values and units of measure, such as GeV, radian, etc. Some input variables have very large values (such as m_{HH}) compared to the others (such as angle related variables), the large values can dominate the machine learning algorithms. As a result, the algorithms may pay most of their attention to the large values and ignore the variables with smaller values. Therefore, the input variables are standardised by subtracting the median and dividing by the interquartile range, to improve the performance of the ML algorithm.

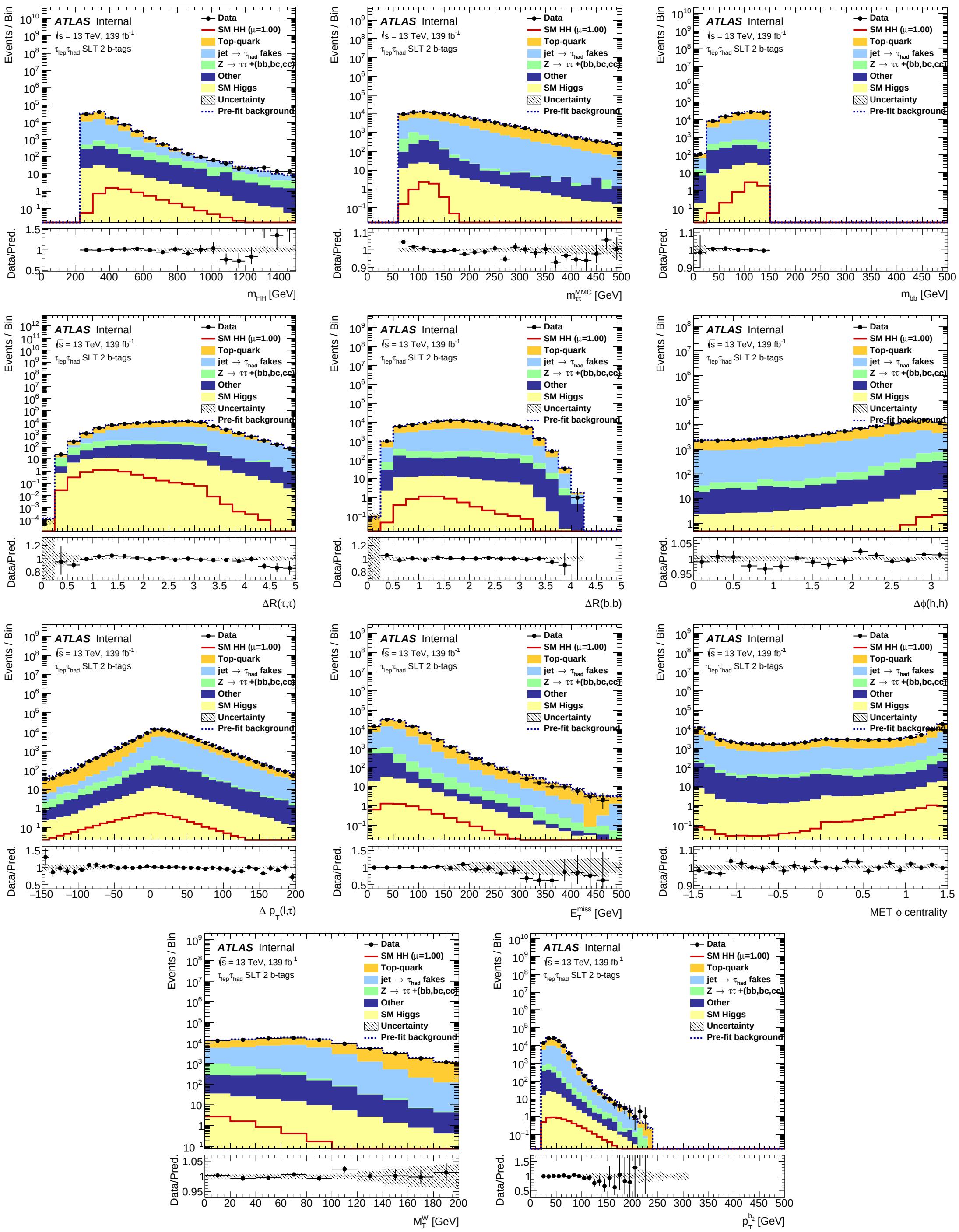


Figure 7.15: Post-fit PNN/NN input variable distributions in the SLT signal region.

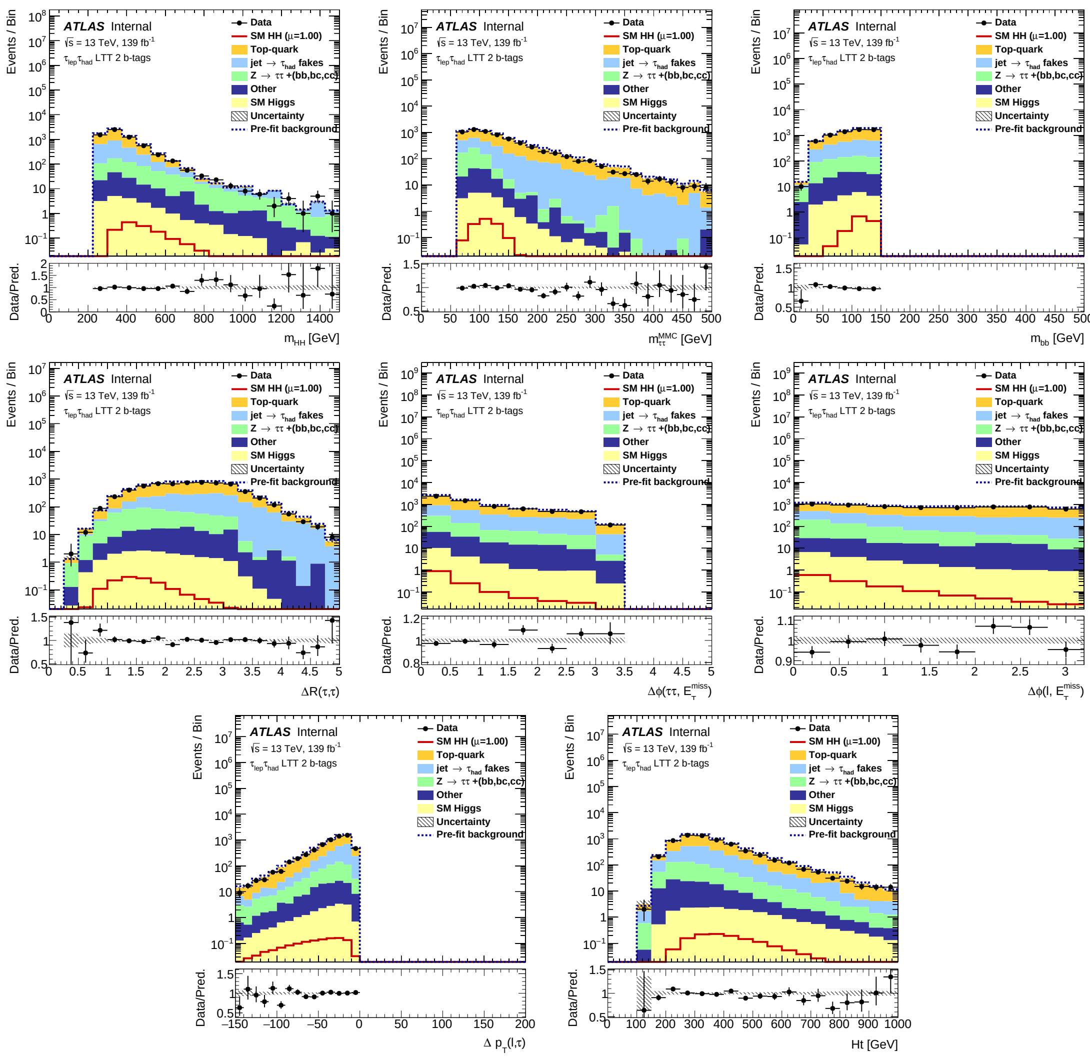


Figure 7.16: Post-fit PNN/NN input variable distributions in the LTT signal region.

Parameter	Value for SLT	Value for LTT
Layer sizes	512 512	512 512 512
Epochs	241	238
Batch-size	34	35
Learning rate	0.500	0.500
Learning rate decay	4.15×10^{-3}	4.20×10^{-3}
Nesterov momentum	0.944	0.791
L2 regularisation weight	1.00×10^{-5}	1.65×10^{-5}

Table 7.9: Training parameters used for the di-Higgs $\tau_{\text{lep}}\tau_{\text{had}}$ non-resonant NNs. Table reproduced from Ref. [135].

Parameter	Value for SLT	Value for LTT
Layer sizes	512 512	256 256 256
Epochs	245	296
Batch-size	33	24
Learning rate	0.500	0.377
Learning rate decay	4.55×10^{-3}	3.38×10^{-3}
Nesterov momentum	0.975	0.975
L2 regularisation weight	5.45×10^{-7}	8.64×10^{-6}

Table 7.10: Training parameters used for the di-Higgs $\tau_{\text{lep}}\tau_{\text{had}}$ resonant PNNs. Table reproduced from Ref. [135].

The neural networks are trained using Keras [197] interfaced to Tensorflow [198]. During training, the sum of all backgrounds normalised to their respective cross-sections are used.

The hyperparameters used for training the NN (PNN) are listed in Table 7.9 (7.10). The ReLU and sigmoidal activation functions are used for the hidden layers and the output layer, respectively. The loss function is the binary cross-entropy, using L2 regularisation ('Ridge regression', which adds squared amplitude for the penalty term), and the optimiser used is stochastic gradient descent [53]. The numbers of epochs, batch size, learning rate and learning rate decay are again listed in Tables 7.9, 7.10. The Nesterov momentum is used in the gradient descent algorithm.

These hyperparameters were optimised separately for each MVA by a Bayesian optimisation procedure with a Gaussian process [199]. A metric for the NN hyperparameters optimisation is defined as the quadrature sum of approximated significances of bins with at least 5 expected background events, while for the PNN, a metric is defined as the approximated significance after an optimised cut (a cut that maximises the significance), and the bins after the cut are required to have at least 5 expected background events. The metric is evaluated on the validation and the testing sets, and the hyper-parameters giving the best value are chosen. The NN and PNN are retrained when the optimisation is done.

During training for the NN (PNNs), a two-fold validation is used. The data set is partitioned into two sets depending on the event numbers being even or odd, where one set is used for training with the other set being used for validation, vice versa. Checks for overtraining of the NNs can be found in Appendix B.2.

In addition, the PNNs are trained on all the signals simultaneously, and are parameterised by the HH sample resonance mass. The PNNs, therefore, provide near-optimal sensitivity over the range of signal masses considered. Since the background does not have a well-defined value of the resonance mass, the background events are assigned a random value generated from the signal mass parameter during training.

The PNNs interpolate well between the mass points used for the training. This is checked by plotting the PNN response in terms of Asimov significance as a function of the PNN mass parameter, which is shown in Figure 7.17. The Asimov significance is calculated following the method described in Ref. [56]. The PNNs show very high significance over the whole scanning mass spectrum. The minimum significance appears in the vicinity of low signal masses reaching a value of 0.7.

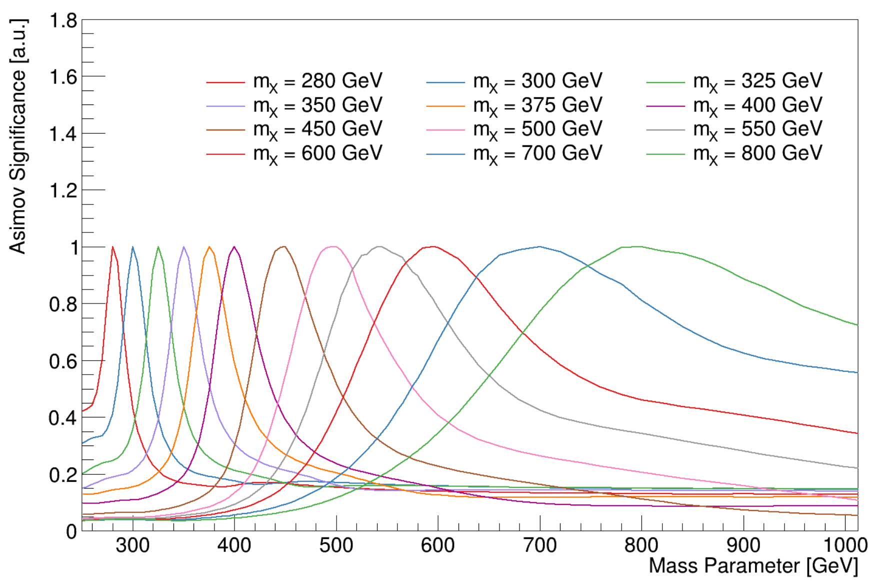


Figure 7.17: PNN response in terms of Asimov significance as a function of the PNN mass parameter obtained for signal samples with different masses. Due to a large difference in the acceptance times efficiency of the $\tau_{\text{lep}}\tau_{\text{had}}$ channel signal region selection depending on the mass of the resonance, all significances are scaled so that the maximum significance is 1 for better visibility.

7.4.3 MVA output distributions

The output distributions of the NN and the PNNs are shown in Figures 7.18 and 7.19 for the SLT and LTT channels, respectively. Similar to the input variables plots, they are plotted after performing the fit and with correction applied, with all uncertainties considered.

The non-resonant HH signal assuming various κ_λ values are passed to the same NN classification used for the SM HH signal. The NN score of non-resonant HH with κ_λ = -5, -3, -1.6, 1, 2, 7 and 10 samples are shown in Figure 7.20 (7.21) for the SLT (LTT) channel. Due to time constraints, the NN has only been trained with SM signal. Therefore it is expected that the separation power will degrade for samples with κ_λ deviating far from 1.

The same NN classification is also applied on non-resonant HH signal assuming 7 benchmark in BSM scenarios. The NN scores of non-resonant HH assuming BM1, BM2, BM3, BM5 together with the SM non-resonant HH signal are shown in Figure 7.22 (7.23) for the SLT (LTT) channel. Again, as the NN has only been trained with SM signal, it is expected that the separation power will degrade for the HEFT signals.

To validate the MC modelling, the post-fit PNN score distributions in the 1-b-tag control region are shown in Figures 7.24 and 7.25 for the SLT and LTT channels, respectively. The NN/PNN distributions of the background agree well with the data. In addition, the data/MC agreement for the resonance signals not shown is also satisfying.

The binning of the MVA output discriminant is defined in Section 7.6.1. The event yields in the last three bins of the MVA discriminant are shown in Table B.1 (Table B.2) in Appendix B.2 for the SLT (LTT) channel, which are the most significant bins for signal extraction.

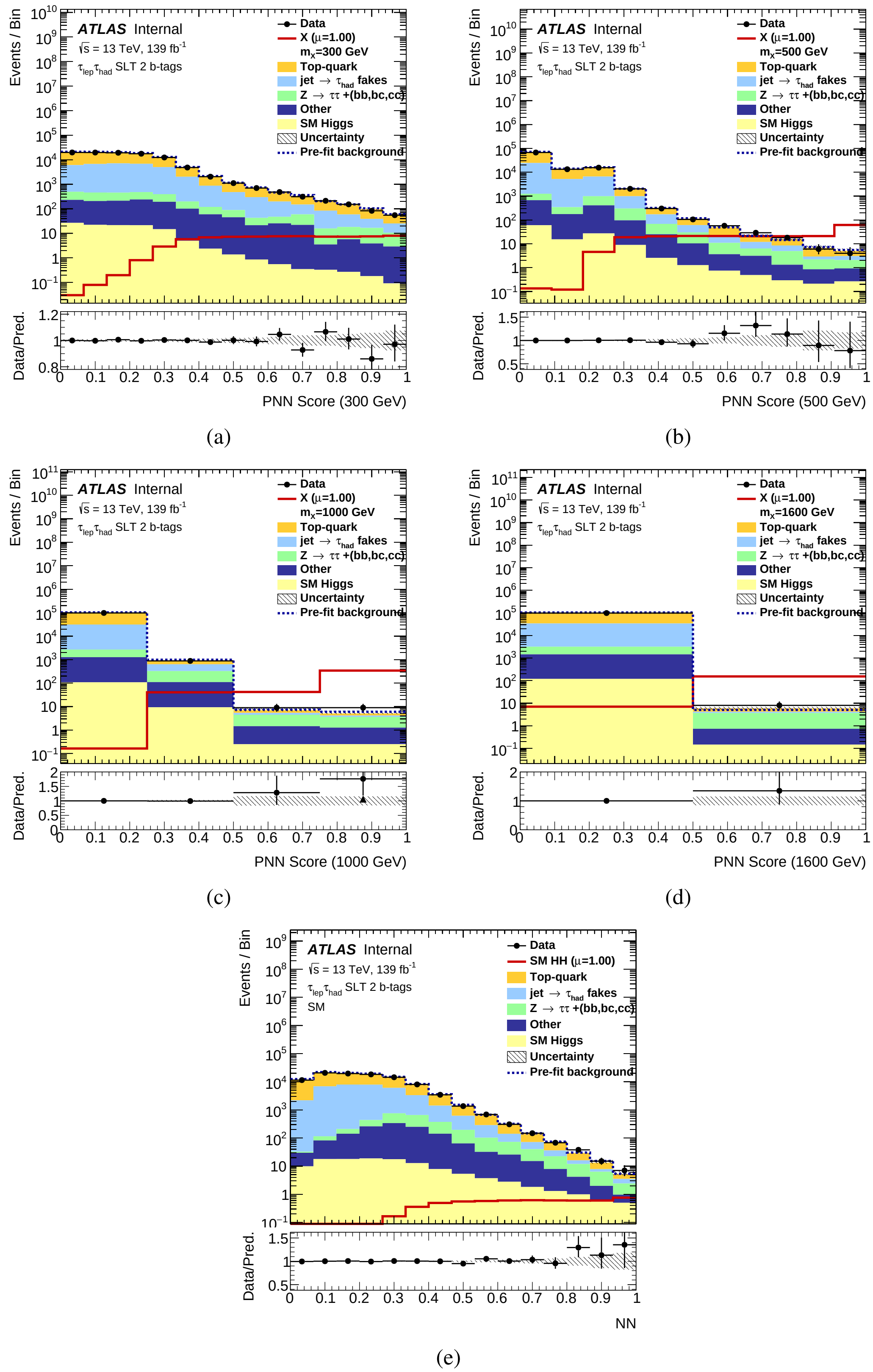
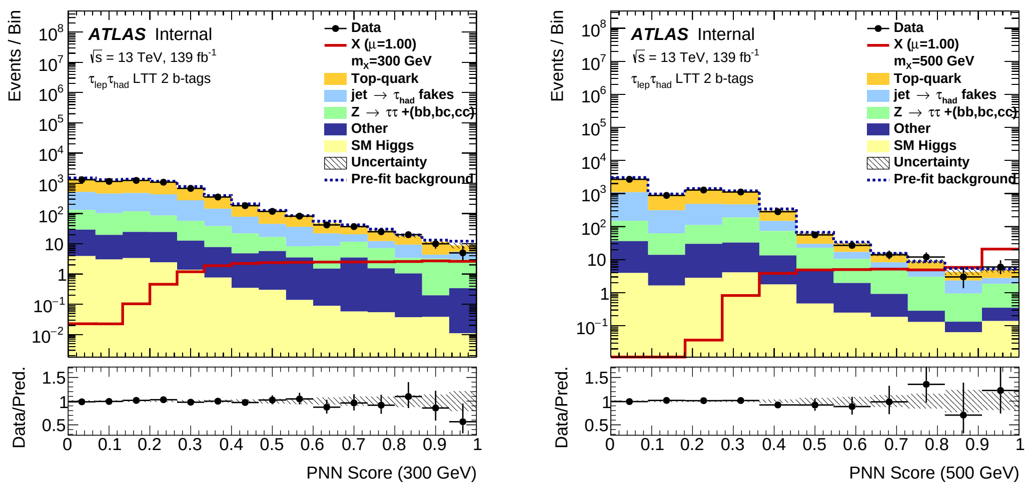
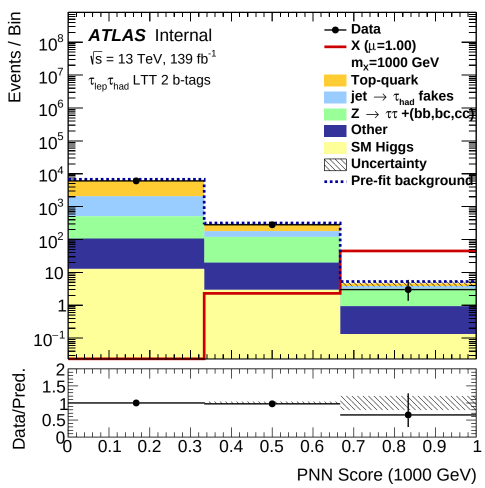


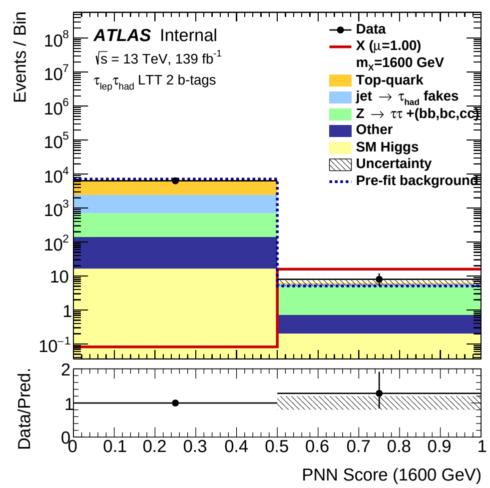
Figure 7.18: Post-fit PNN score distributions for the 300, 500, 1000, 1600 GeV mass points and non-resonant NN in the SLT channel.



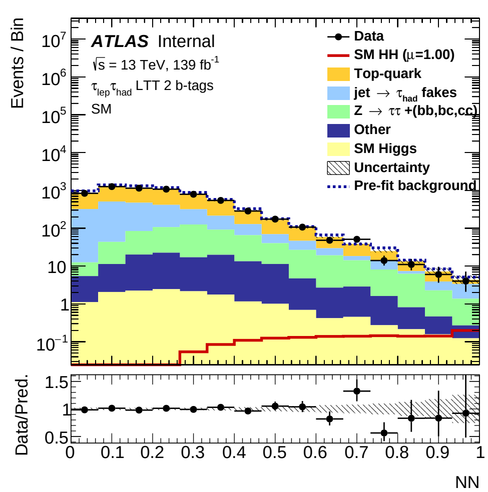
(a)



(b)



(c)



(d)

Figure 7.19: Post-fit PNN score distributions for the 300, 500, 1000, 1600 GeV mass points and non-resonant NN in the LTT channel.

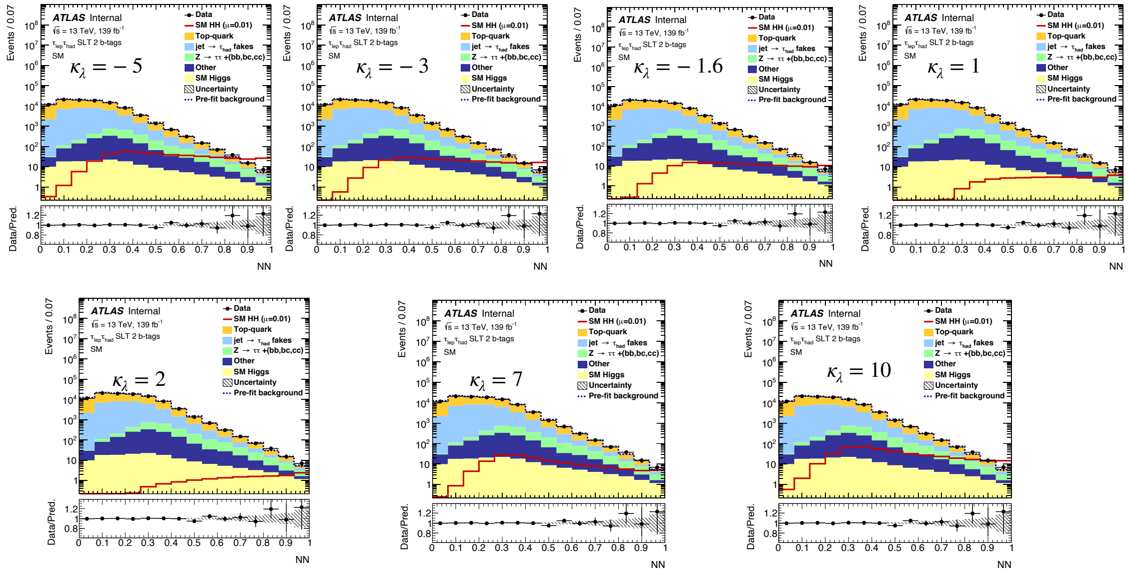


Figure 7.20: Post-fit plots of the neural network output in the $\tau_{\rm lep}\tau_{\rm had}$ SLT signal region, showing the agreement of the background-only hypothesis with the data as well as the signal prediction for different κ_λ values.

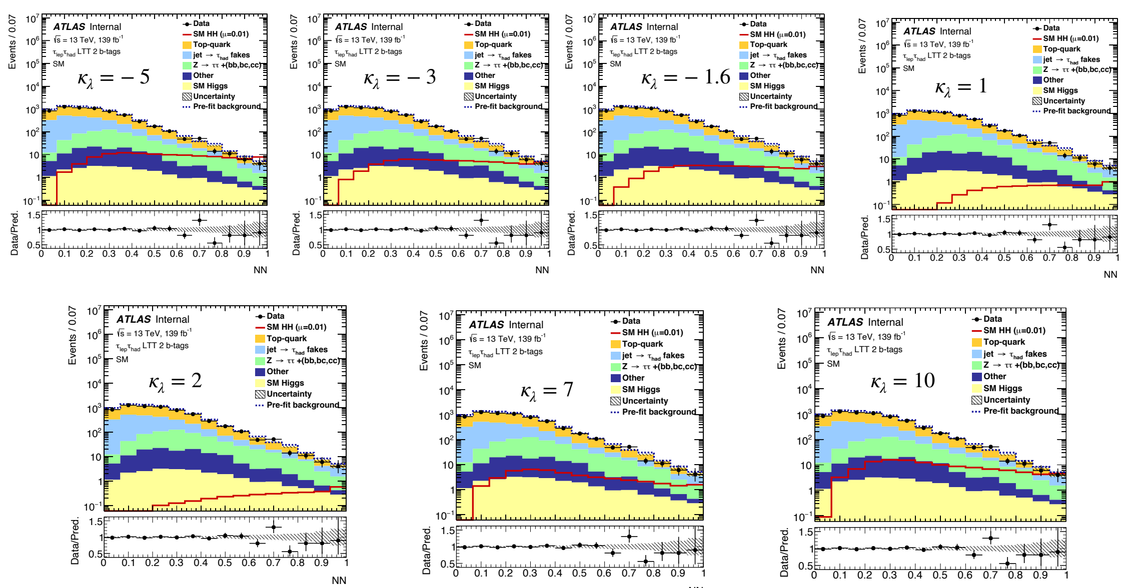


Figure 7.21: Post-fit plots of the neural network output in the $\tau_{\rm lep}\tau_{\rm had}$ LTT signal region, showing the agreement of the background-only hypothesis with the data as well as the signal prediction for different κ_λ values.

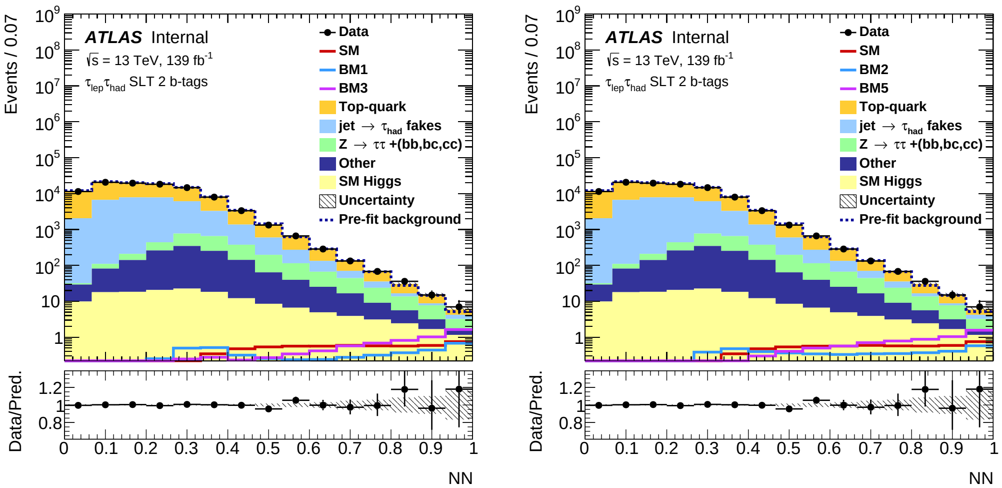


Figure 7.22: Post-fit plots of the neural network output in the $\tau_{\text{lep}}\tau_{\text{had}}$ SLT signal region, showing the agreement of the background-only hypothesis with the data as well as the signal prediction assuming BM1, BM2, BM3, BM5. Plots credit: Matthew Sullivan.

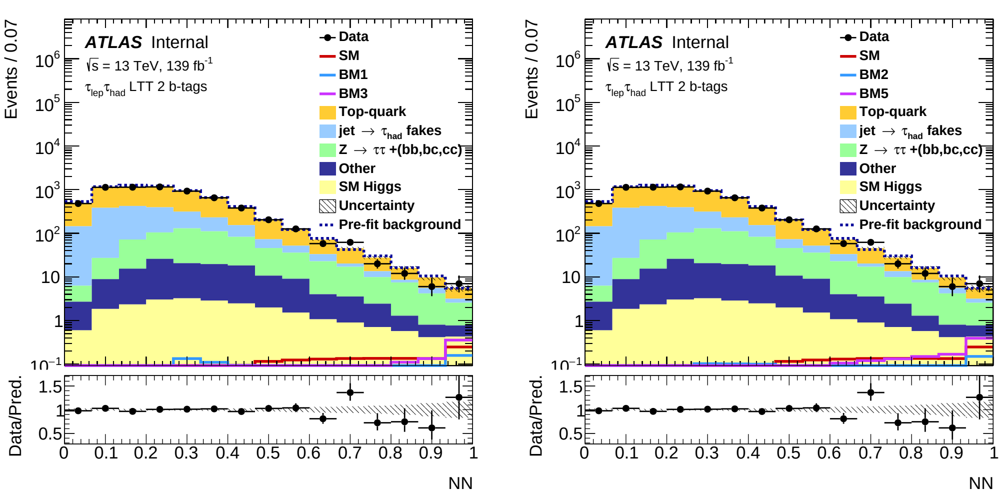


Figure 7.23: Post-fit plots of the neural network output in the $\tau_{\text{lep}}\tau_{\text{had}}$ LTT signal region, showing the agreement of the background-only hypothesis with the data as well as the signal prediction assuming BM1, BM2, BM3, BM5. Plots credit: Matthew Sullivan.

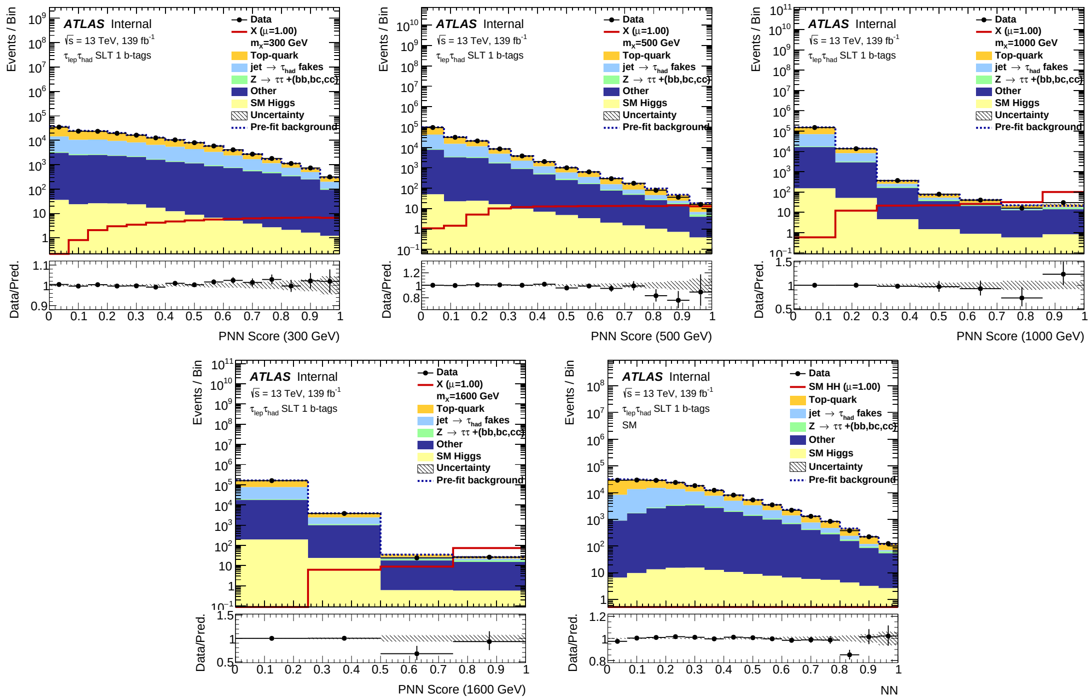


Figure 7.24: Post-fit 1-*b*-tag region PNN score distributions for the 300, 500, 1000, 1600 GeV mass points and non-resonant NN in the SLT channel.

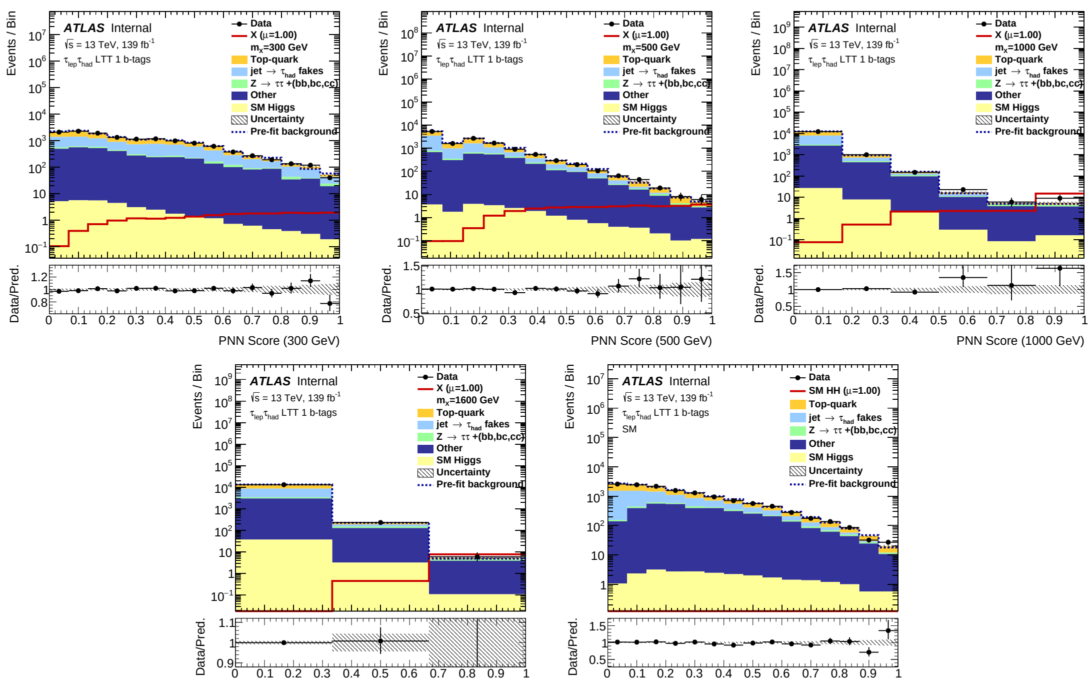


Figure 7.25: Post-fit 1-*b*-tag region PNN score distributions for the 300, 500, 1000, 1600 GeV mass points and non-resonant NN in the LTT channel.

7.5 Systematic uncertainties

The systematic uncertainties considered in the HH analysis are outlined in this section. Due to the same reason mentioned in Section 6.7, the systematic uncertainties can be broadly categorised into experimental and modelling systematic uncertainties. The former arises from experimental sources related to the detector response and the physics object reconstruction and identification, while the modelling systematic uncertainties account for theoretical uncertainties related to the limited knowledge of the MC simulated processes and uncertainties on data-driven backgrounds.

The experimental and modelling uncertainties are derived for all backgrounds and signals, manifesting in variations of the overall yields or the shape of the final discriminant, or both together. All uncertainties considered are propagated through the analysis, and are included in the statistical analysis discussed in Section 7.6.

The experimental systematic uncertainties are outlined in Section 7.5.1, and the background and signal modelling systematic uncertainties are outlined in Section 7.5.2 and 7.5.3, respectively. The systematic uncertainties impact on the likelihood fit is summarised in Section 7.5.5.

7.5.1 Experimental uncertainties

The experimental systematic uncertainties have been briefly discussed in Section 6.7, while many of the systematics considered there will be included in this analysis as well. In Table 7.11 and Table 7.12, the full list of the experimental systematics are listed.

An uncertainty on measuring the luminosity arises from calibrating it with the van der Meer method [200]. The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [128], obtained using the LUCID-2 detector [201] for the primary luminosity measurements.

A reweighting is applied to the simulations to match the pile-up condition in each data-taking period during Run 2 [202]. The uncertainty associated with the reweighting is estimated by varying the reweighting factors used in the simulation.

The uncertainties in single lepton triggers, uncertainties due to lepton reconstruction, identification and isolation are estimated by varying the selection used in the *tag-and-probe* method. The idea of the tag-and-probe method can be explained using a simple example: consider $Z \to e^+e^-$ events; one can apply stringent selection criteria on one electron ('tag'), and loosely identifying another electron which is compatible with the first electron (when these two electrons have a reconstructed mass close to the mass of Z). The second electron has a very high probability of being a real electron from Z decay and can then be used for efficiency measurement ('probe'). For electrons, $Z \to e^+e^-$ events are used [96] while for muons, both $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events are used [97]. The same events are used to estimate the uncertainties on electron or muon energy resolution by smearing the electron or muon energies in simulations.

Similarly, for the τ_{had}, the tag-and-probe method is used to measure the reconstruction and identification efficiency scale factors, using $Z \to \tau^+\tau^-$ events [120]. Uncertainties are estimated in the efficiencies of the electron veto for τ_{had}, using $Z \to e^+e^-$ events.

The systematic uncertainties associated with the jet energy scale calibration [106], as described in Section 4.4, are grouped into a reduced-size set of nuisance parameters. The reduced set is referred to as 'CategoryReduction', and is the ATLAS recommended set of uncertainties for analyses aiming to perform combinations with other searches. Similarly, the jet energy resolution uncertainties [106] are grouped into the 'FullJER' scheme for this reason. The jet energy resolution uncertainties therefore are comprised of 30 nuisance parameters and 13 nuisance parameters for the jer energy resolution uncertainties.

The uncertainty associated with the JVT algorithm (see Section 4.4) is estimated by varying the requirements on the algorithm output discriminant.

As described in Section 6, scale factors are measured to correct the b-tagging efficiency in simulations to match the one in data. Uncertainties in the b-tagging efficiency are derived as a function of p_{T} for b-jets and c-jets using $t\bar{t}$ events [115, 184], and as a function of p_{T} and η for light-flavour jets using di-jet events [185]. The 'medium reduction' scheme is used, which is the preferred reduction scheme used for analyses aiming for combination, which contains 13 nuisance parameters.

The energy scale and resolution uncertainties mentioned above, in addition to uncertainty in the tracks matched to the primary vertex but not associated to the other reconstructed objects (the soft event, as defined in Section 4.6), are propagated to the $E_{\text{T}}^{\text{miss}}$calculation [203].

7.5.2 Background modelling uncertainties for MC-based processes

The modelling systematic uncertainties include the uncertainties due to theoretical cross-section calculations and acceptance. The former only affects the normalisation of the samples, and is applied on all MC backgrounds; the later affects both the expected event yields and the shape of kinematic distributions, and therefore the MVA discriminant output shapes. In order to disentangle these effects, the impact from shape variation is considered separately from the normalisation effect, as it does not change the number of expected events. The shape uncertainties are derived comparing the distributions of several kinematic variables or the MVA discriminant output of the alternative samples to the ones of the nominal sample. The comparison is performed by taking the normalised distributions (to exclude the normalisation effects that are considered separately) and calculating the

NP Name	Description
Event	
ATLAS_LUMI_Run2	Uncertainty on the total integrated luminosity (1.7%)
ATLAS_PU_PRW_DATASF	Pile-up reweighting
electrons	
EL_EFF_TRIG_TOTAL_1NPCOR_PLUS_UNCOR	electron-trigger
EL_EFF_RECO_TOTAL_1NPCOR_PLUS_UNCOR	electron-reconstruction
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	electron-identification
EL_EFF_ISO_TOTAL_1NPCOR_PLUS_UNCOR	electron-isolation
muons	
MUON_EFF_TrigStatUncertainty	mun-trigger
MUON_EFF_TrigSystUncertainty	
MUON_EFF_RECO_STAT	muon-reconstruction
MUON_EFF_RECO_SYS	
MUON_EFF_RECO_STAT_LOWPT	
MUON_EFF_RECO_SYS_LOWPT	
MUON_EFF_ISO_STAT	muon-isolation
MUON_EFF_ISO_SYS	
MUON_EFF_TTVA_STAT	muon-TTVA
MUON_EFF_TTVA_SYS	
MUON_ID	muon-momentum calibration
MUON_MS	
MUON_SCALE	
MUON_SAGITTA_RHO	
MUON_SAGITTA_RESBIAS	
τ-leptons	
TAUS_TRUEHADTAU_EFF_TRIGGER_STATDATA161718	τ-trigger
TAUS_TRUEHADTAU_EFF_TRIGGER_STATDATA1718	
TAUS_TRUEHADTAU_EFF_TRIGGER_STATDATA2016	
TAUS_TRUEHADTAU_EFF_TRIGGER_STATDATA2018	
TAUS_TRUEHADTAU_EFF_TRIGGER_STATDATA2018AFTTS1	
TAUS_TRUEHADTAU_EFF_TRIGGER_STATMC161718	
TAUS_TRUEHADTAU_EFF_TRIGGER_STATMC1718	
TAUS_TRUEHADTAU_EFF_TRIGGER_STATMC2016	
TAUS_TRUEHADTAU_EFF_TRIGGER_STATMC2018	
TAUS_TRUEHADTAU_EFF_TRIGGER_STATMC2018AFTTS1	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYST161718	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYST1718	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYST2016	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYST2018	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYST2018AFTTS1	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYSTMU161718	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYSTMU1718	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYSTMU2016	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYSTMU2018	
TAUS_TRUEHADTAU_EFF_TRIGGER_SYSTMU2018AFTTS1	
TAUS_TRUEHADTAU_EFF_RECO_HIGHPT	τ-reconstruction
TAUS_TRUEHADTAU_EFF_RECO_TOTAL	
TAUS_TRUEHADTAU_EFF_RNNID_1PRONGSTATSYSTPT2025	τ-identification
TAUS_TRUEHADTAU_EFF_RNNID_1PRONGSTATSYSTPT2530	
TAUS_TRUEHADTAU_EFF_RNNID_1PRONGSTATSYSTPT3040	
TAUS_TRUEHADTAU_EFF_RNNID_1PRONGSTATSYSTPTGE40	
TAUS_TRUEHADTAU_EFF_RNNID_3PRONGSTATSYSTPT2025	
TAUS_TRUEHADTAU_EFF_RNNID_3PRONGSTATSYSTPT2530	
TAUS_TRUEHADTAU_EFF_RNNID_3PRONGSTATSYSTPT3040	
TAUS_TRUEHADTAU_EFF_RNNID_3PRONGSTATSYSTPTGE40	
TAUS_TRUEHADTAU_EFF_RNNID_HIGHPT	
TAUS_TRUEHADTAU_EFF_RNNID_SYST	
TAUS_TRUEHADTAU_SME_TES_INSITUEXP	τ-energy scale
TAUS_TRUEHADTAU_SME_TES_INSITUFIT	
TAUS_TRUEHADTAU_SME_TES_MODEL_CLOSURE	
TAUS_TRUEHADTAU_SME_TES_PHYSICSLIST	
TAUS_TRUEELECTRON_EFF_ELEBDT_STAT	τ-electron veto
TAUS_TRUEELECTRON_EFF_ELEBDT_SYST	
TAUS_TRUEHADTAU_EFF_ELEOLR_TOTAL	
egamma	
EG_RESOLUTION_ALL	egamma resolution
EG_SCALE_ALL	egamma scale
EG_SCALE_AF2	
$E_{\mathrm{T}}^{\mathrm{miss}}$	
MET_SoftTrk_ResoPara	MET soft term resolution
MET_SoftTrk_ResoPerp	
MET_SoftTrk_ScaleDown	MET soft term scale
MET_SoftTrk_ScaleUp	

Table 7.11: List of experimental systematic uncertainties on leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$. Table reproduced from analysis internal notes.

NP Name	Description
Jets	
JET_EtaIntercalibration_Modelling	Jet energy scale (CategoryReduction)
JET_EtaIntercalibration_TotalStat	
JET_EtaIntercalibration_NonClosure_highE	
JET_EtaIntercalibration_NonClosure_negEta	
JET_EtaIntercalibration_NonClosure_posEta	
JET_Pileup_OffsetMu	
JET_Pileup_OffsetNPV	
JET_Pileup_PtTerm	
JET_Pileup_RhoTopology	
JET_Flavor_Composition	
JET_Flavor_Response	
JET_PunchThrough_MC16	
JET_PunchThrough_AFII	
JET_EffectiveNP_Detector1	
JET_EffectiveNP_Detector2	
JET_EffectiveNP_Mixed1	
JET_EffectiveNP_Mixed2	
JET_EffectiveNP_Mixed3	
JET_EffectiveNP_Modelling1	
JET_EffectiveNP_Modelling2	
JET_EffectiveNP_Modelling3	
JET_EffectiveNP_Modelling4	
JET_EffectiveNP_Statistical1	
JET_EffectiveNP_Statistical2	
JET_EffectiveNP_Statistical3	
JET_EffectiveNP_Statistical4	
JET_EffectiveNP_Statistical5	
JET_EffectiveNP_Statistical6	
JET_SingleParticle_HighPt	
JET_RelativeNonClosure_AFII	
JET_BJES_Response	
JET_EtaIntercalibration_NonClosure_2018data	
JET_JER_DataVsMC_MC16	Jet energy resolution (FullJER)
JET_JER_DataVsMC_AFI	
I JET_JER_EffectiveNP_1	
JET_JER_EffectiveNP_2	
JET_JER_EffectiveNP_3	
JET_JER_EffectiveNP_4	
JET_JER_EffectiveNP_5	
JET_JER_EffectiveNP_6	
JET_JER_EffectiveNP_7	
JET_JER_EffectiveNP_8	
JET_JER_EffectiveNP_9	
JET_JER_EffectiveNP_10	
JET_JER_EffectiveNP_11	
JET_JER_EffectiveNP_12restTerm	
JET_JVT_EFF	Jet vertex tagging
JET_FJVT_EFF	Forward jet vertex tagging
b-tagging	
FT_EFF_Eigen_B_0	b-tagging scale factors
FT_EFF_Eigen_B_1	
FT_EFF_Eigen_B_2	
FT_EFF_Eigen_C_0	c-jet mis-tagging scale factors
FT_EFF_Eigen_C_1	
FT_EFF_Eigen_C_2	
FT_EFF_Eigen_C_3	
FT_EFF_Eigen_Light_0	light-jet mis-tagging scale factors
FT_EFF_Eigen_Light_1	
FT_EFF_Eigen_Light_2	
FT_EFF_Eigen_Light_3	
FT_EFF_extrapolation	b-tagging scale factors extrapolation to high p_{T}
FT_EFF_extrapolation_from_charm	c-jet mis-tagging scale factors extrapolation to high p_{T}

Table 7.12: List of experimental systematic uncertainties on jets and b-tagging. Table reproduced from analysis internal notes.

ratio of the alternative and the nominal number of events bin per bin.

Acceptance uncertainties for $t\bar{t}$, Z+heavy flavour jets, single-top (Wt-channel), ttH and ZH are estimated in the SR and CR, using the nominal and alternative MC simulated samples. The acceptance uncertainties on minor backgrounds (Z+light flavour jets, W+jets, and di-boson) are taken from Ref. [204], as described in Section 7.5.2.5.

The normalisation acceptance uncertainties are derived comparing the acceptance obtained for each alternative sample to the nominal acceptance:

$$\sigma_A = \frac{A_{\mathrm{var}} - A_{\mathrm{nom}}}{A_{\mathrm{nom}}} = \frac{N_{\mathrm{var}} - N_{\mathrm{nom}}}{N_{\mathrm{nom}}}, \tag{7.5}$$

where A_{var} (A_{nom}) is the acceptance of the variation (nominal) sample and N_{var} (N_{nom}) is the expected event yields of the variation (nominal) sample. Acceptance uncertainties on the normalisation of the backgrounds are applied in all regions and treated as correlated across regions except for the $t\bar{t}$ and Z + HF backgrounds whose normalisations are left unconstrained (floated) in the global likelihood fit. For these two processes, a relative acceptance uncertainty is needed to take into account the correlation and the differences in the normalisations between regions. The relative acceptance uncertainty is estimated from the comparison of the relative amount of events predicted by the alternative model in one region 'i' with respect to another region 'j' and the same fraction in the nominal model:

$$\sigma_{A^i/A^j} = \frac{\frac{A^i_{var}}{A^j_{var}} - \frac{A^i_{nom}}{A^j_{nom}}}{\frac{A^i_{nom}}{A^j_{nom}}} = \frac{\frac{N^i_{var}}{N^j_{var}} - \frac{N^i_{nom}}{N^j_{nom}}}{\frac{N^i_{nom}}{N^j_{nom}}}. \tag{7.6}$$

The list of all uncertainties applied to all MC backgrounds is reported in Table B.3 and Table B.4 in Appendix B.3.

7.5.2.1 Uncertainties on the $t\bar{t}$ background

The normalisation of the $t\bar{t}$ background is estimated from data as included as a freely floating parameter in the final fit. This normalisation is correlated across all regions included in the fit, and it is determined mainly from the tails of the $m_{\ell\ell}$ distribution in the Z + HF CR, which have a high purity of $t\bar{t}$ events.

Relative acceptance uncertainties on the normalisation are applied on $t\bar{t}$ in the SR to take into account potential differences in the normalisation between the SRs and the CR. The shape variations are checked and applied, correlated with the relative acceptance uncertainties on the normalisations, where they are found to be relevant as described in the following. All these uncertainties are derived by MC-to-MC comparison. The alternative samples are all generated using the AF2 simulation as described in Section 5.3, and for

Source	SLT	LTT
ME	± 0.0026	± 0.009
PS	± 0.072	± 0.088
ISR	+0.0005, -0.0081	-0.0052, +0.013
FSR	+0.014, -0.0097	+0.0096, -0.032
PDF+α_s	± 0.006	± 0.0073
Total	± 0.074	-0.094, +0.090

Table 7.13: Relative size of relative acceptance normalisation uncertainties on $t\bar{t}$.

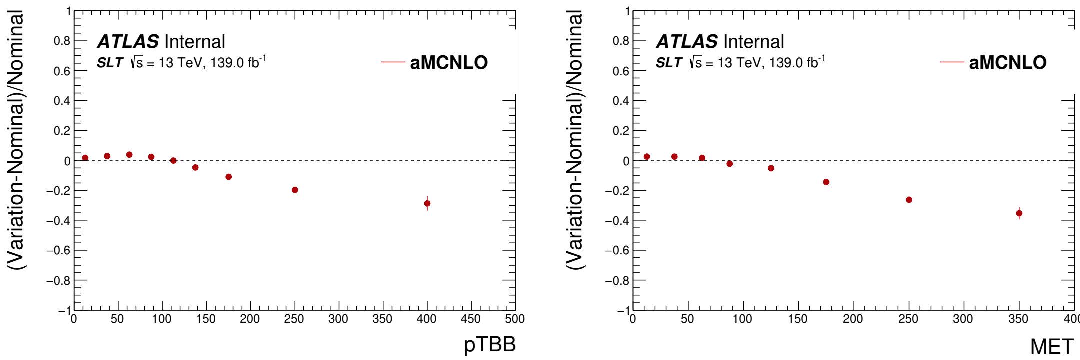
consistency, the fast simulation nominal $t\bar{t}$ sample is also generated and is used in the comparison. The derived uncertainties are then applied on the $t\bar{t}$ sample generated with full simulation.

Uncertainties due to PDF, α_s, FSR and scales are evaluated using internal alternative weights present in all the $t\bar{t}$ nominal samples.

Uncertainties from the Parton Shower (PS), Matrix Element (ME) and Initial State Radiation (ISR) are estimated using the differences between the nominal samples and the corresponding variations samples: parton shower uncertainties are evaluated comparing the nominal samples showered with Pythia to alternative samples showered with Herwig; uncertainty due to the matrix element is evaluated comparing the nominal Powheg +Pythia samples to alternative aMC@NLO +Pythia samples. The ISR up variation is evaluated by varying up the h_{damp} parameter (the model parameter that controls ME/PS matching in Powheg and effectively regulates the high-p_{T} radiation) of the fast simulation Powheg +Pythia, at the same time dividing by two the renormalisation and factorisation scales and varying the showering; while the ISR down variation is evaluated by comparing the nominal Powheg +Pythia with the samples obtained doubling the renormalisation and factorisation scales and varying the showering.

Relative acceptance uncertainties on the normalisation between the CR and the SR are calculated using Equation 7.6 and are reported in Table 7.13.

The shapes of $t\bar{t}$ parton shower and matrix element modelling uncertainties are checked in the MVA discriminant input variables and output distributions for the SLT and LTT channels. The shape in matrix element/parton shower uncertainties is found to be insignificant in the NN/PNN scores in the LTT channel and therefore not considered in this analysis. To avoid running these variation samples for every analysis iteration and to estimate the shape of the systematic uncertainties, the parton shower/matrix element uncertainties are parametrised in the MVA output distributions for the SLT channel. Two parametrisation methods were studied, one using parametrisation in a combination of MVA discriminant input variables and one using parametrisation directly in the MVA output distributions. The latter was adopted for the final result. Both methods will be discussed in the following.



(a) a

(b) b

Figure 7.26: The shape-only parametrisation in bins of p_T^{bb} (a) and in $E_{\mathrm{T}}^{\mathrm{miss}}$ for the matrix element uncertainty, derived in the SLT channel.

The first method is based on a parametrisation calculated as the ratio of the expected event yields of the variation sample to the nominal sample in bins of the MVA discriminant input variables. For the matrix element uncertainty, obvious trends are seen in p_T of the 2 b-jet system (p_T^{bb}) and $E_{\mathrm{T}}^{\mathrm{miss}}$. In order to mimic the MVA output distribution of the variation sample, the nominal sample is weighted by the parametrisation in bins of p_T^{bb} and $E_{\mathrm{T}}^{\mathrm{miss}}$, and the weighted distributions are passed to the MVA classification to check the shape of the output. However, the parametrisation using neither the p_T^{bb} nor $E_{\mathrm{T}}^{\mathrm{miss}}$ alone recovers the shape of the MVA output of the variation sample very well. In addition, all other MVA discriminant input variables are checked for parametrisation, but none of them gives good closure in MVA output either. Therefore, a sequential parametrisation in p_T^{bb} and $E_{\mathrm{T}}^{\mathrm{miss}}$ is checked. The weights are applied on the nominal sample first using the p_T^{bb} parametrisation. The $E_{\mathrm{T}}^{\mathrm{miss}}$ parametrisation is then extracted by taking the ratio of the $E_{\mathrm{T}}^{\mathrm{miss}}$ of the variation sample to the $E_{\mathrm{T}}^{\mathrm{miss}}$ of the weighted nominal sample. The additional step is to take into account the possible change in $E_{\mathrm{T}}^{\mathrm{miss}}$ after the first weighting step. This sequential parametrisation is shown in Figure 7.26 for the SLT channel.

The p_T^{bb} and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions are checked again for the weighted samples as compared to the variation samples. Good closure is found in both distributions and in other MVA discriminant input variables. To check how well the weighted nominal sample can recover the MVA output distributions of the variation sample, the NN score distribution of the variation and the weighted nominal is shown in Figure 7.27, where the event yields of the variation sample are normalised to those of the nominal. The closure in the NN score is also very good.

The parton shower uncertainty is first checked using a parametrisation in a single MVA variable, but none of the MVA discriminant inputs alone gives good closure in the MVA output distributions. Therefore, a similar sequential parametrisation approach is used. The

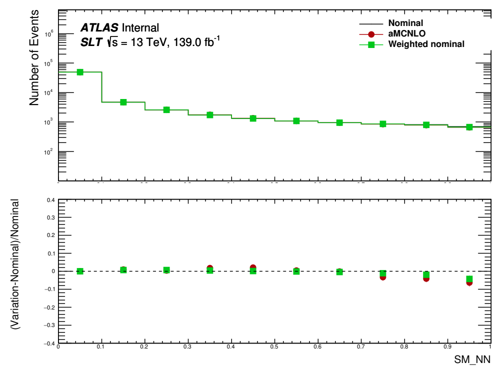
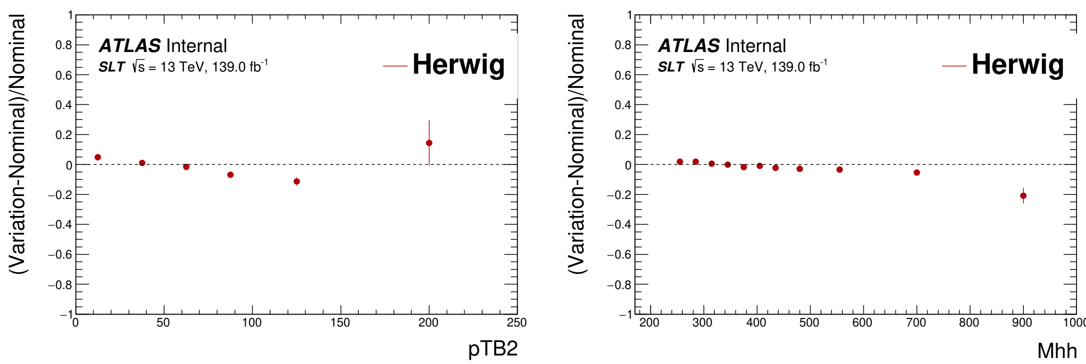


Figure 7.27: SLT channel: shape-only NN score of the matrix element variation sample and the weighted nominal, using sequential parametrisation in p_T^{bb} and $E_{\mathrm{T}}^{\mathrm{miss}}$. The binning shown here is in equal width bins of NN score.



(a) a

(b) b

Figure 7.28: The shape-only parametrisation in bins of p_T^{b2} (a) in m_{HH} (b), for the parton shower uncertainty, derived in the SLT channel.

parametrisation is performed first in bins of the p_{T} of the subleading jet (p_T^{b2}) followed by parametrising the residue in bins of mass of the di-Higgs system (m_{HH}), as shown in Figure 7.28. The weighted sample is compared with the variation sample, good closure in the MVA output is obtained as shown in Figure 7.29.

The sequential parametrisation method is compared with the second method, which is based on parametrising directly in the ratio of MVA discriminant output of the variation samples to the ones of the nominal. By definition, this method recovers exactly the variation shape in the NN/PNN output distribution; therefore, no closure check is required. The binning used for parametrisation is the same as the binning used in the final fit. Due to the very fine binning in the high MVA output region, large statistical fluctuations are observed, therefore, a smoothing process is applied to these uncertainties during limit setting to reduce the effect of the fluctuations. The SM NN parametrisation for the matrix element and parton shower variation in the SLT channel is shown in Figure 7.30. Plots for more

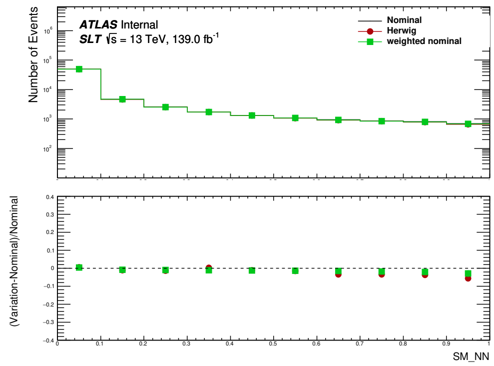


Figure 7.29: The shape-only NN score of the parton shower variation sample and the weighted nominal. The binning shown here is in equal width bins of NN score.

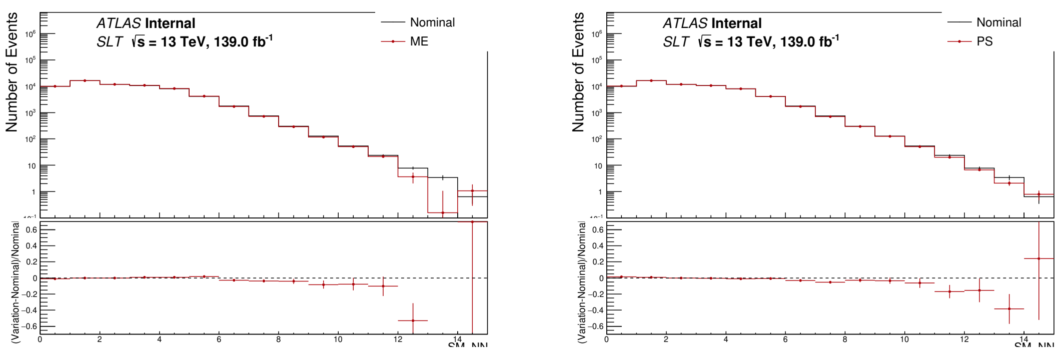


Figure 7.30: Parametrisation in SM NN score distribution for the ME (left) and PS (right) systematic uncertainties. The binning is the same as the final fit binning.

resonant mass points can be found in Appendix B.3.

The main reason for using the second method is to ease the combination with the other channels, and to reduce complexity, even though the second method has some drawbacks, such as the larger fluctuations. Also, in the high MVA score region, where the binning becomes much finer, the first method starts to fail to cover the large difference between the variation and the nominal samples.

The ISR up and down variations have not shown obvious shape in the NN/PNN distribution; therefore, only a normalisation uncertainty is considered.

Uncertainties on $t\bar{t}$ in the $Z+$HF control region

Shape uncertainties on the $t\bar{t}$ background are neglected in the $Z+$HF control region as they are found to be negligible in the $m_{\ell\ell}$ distribution included in the fit for this region, as shown in Figure B.8 in Appedix B.3.

Source	SLT	LTT
ME	± 0.021	± 0.10
Scales	-0.029, +0.053	-0.054, +0.085
CKKW	± 0.07	± 0.071
QSF	± 0.016	± 0.016
PDF+α_s	± 0.0026	± 0.0033
PDF choice	± 0.0097	± 0.011
Total	-0.081, +0.092	-0.14, +0.15

Table 7.14: Relative size of relative acceptance normalisation uncertainties on Z + HF.

7.5.2.2 Uncertainties on the Z + HF background

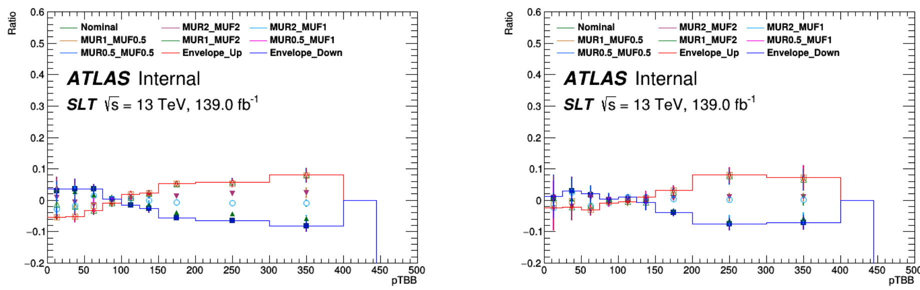
The normalisation of the Z + HF background is estimated from data as a freely floating parameter in the final fit. This normalisation is mainly determined from the dedicated Z + HF dilepton control region. Only the Z + HF backgrounds are considered, including $Z \to \tau\tau$+bb, bc, cc. Contributions from Z + light flavour jets, including $Z \to \tau\tau$+bl, cl, ll are excluded.

Uncertainties due to the choice of matrix element generators are evaluated by comparing the Sherpa 2.2.1 samples to the MadGraph + Pythia samples. Uncertainties on the Z+jets background modelling related to the choice of renormalisation and factorisation scales are evaluated using event weights included in the Sherpa 2.2.1 samples, varying the scales either together or independently up and down by a factor of two, leading to 7-point scale variations. The uncertainty due to the choice of PDF set is evaluated by comparing the `NNPDF3.0` PDF set to two other PDF sets, to two `NNPDF3.0` PDF varied α_s values. The difference between these PDF sets are covered in an envelope and the envelope is taken to be the intra-PDF uncertainty. The inter-PDF set uncertainty is estimated by calculating the standard deviation of 101 replicas of `NNPDF3.0` PDF set.

From the nominal Sherpa configuration, there are two other parameters that can be varied to introduce uncertainties on the modelling of the W/Z+jets process: matrix element matching (*CKKW*), which controls the scale taken for the calculation for the overlap between jets from the matrix element and the parton shower. The nominal value for this parameter is 20 GeV. The up variation increases this to 30 GeV, while the down variation decreases the nominal value to 15 GeV. The second parameter is the resummation scale (*QSF*), which controls the scale used for the resummation of soft gluon emission μ_{QSF} and is varied from 2 and with respect to the nominal.

Relative acceptance normalisation uncertainties between the CR and the SR are calculated from the single channel acceptances using Equation 7.6 and reported in Table 7.14.

For the 7-point scale variation, the ratios of the normalised MVA output distributions of the variations to the one of the nominal are checked. An obvious shape is observed, and a parametrisation approach is used, similar to the one used for the $t\bar{t}$ systematic uncertainties. All MVA discriminant input variables were considered for the parametrisation. Instead of parametrising all seven points of variation, the parametrisation is considered for the envelope encapsulating the maximum shift (in the upward or downward direction) of the variation. The variable to parametrise is chosen such that the weighted nominal samples can best recover the envelope covering the maximum shift in the PNN scores. As a result, the envelope is parametrised by p_T^{bb}, as shown in Figure 7.31. In the figure, the maximum shift of the 7-point variations in the MVA output distributions are covered by the envelope labelled as 'Envelop_Up(Down)'.



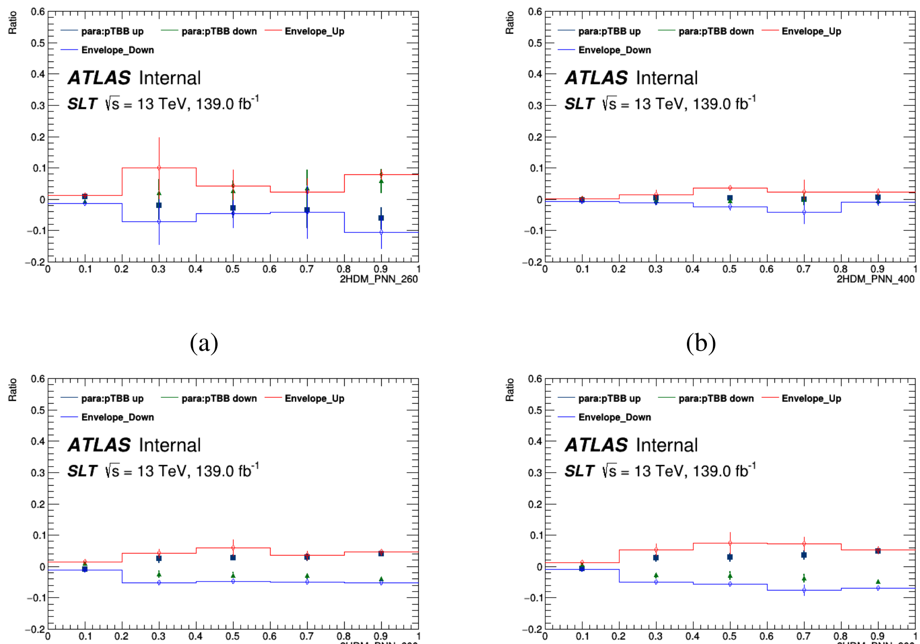
(a) (b)

Figure 7.31: SLT (left) and LTT channel (right): Taking an envelope of the Z + HF scale variation with p_T of the 2 b-jet system.

The PNN score distribution is shown in Fig.7.32 for the SLT channel and Fig.7.33 for the LTT channel. The maximum shift of the variation samples in the PNN score is encapsulated by envelopes labelled as 'Envelop_Up(Down)', and the PNN scores of the weighted nominal sample are labelled as 'para:pTBB Up(Down)'. The parametrisation can recover the shape of the envelope in the PNN scores in general. In some bins of the weighted distribution, the PNN score does not match the envelope. However, these bins have larger uncertainties, and therefore large fluctuations are expected. As a result, one can still conclude a good closure from these plots.

For the uncertainty due to generator variation, the ratios of the variance to the nominal are checked in the MVA output. For most of the bins and mass points, no significant deviation from unity is observed within the statistical uncertainty. Similarly, the shape dependence on the uncertainty due to choice of intra/inter-PDF set, α_s variations, CKKW and QSF are checked. No obvious trend is observed.

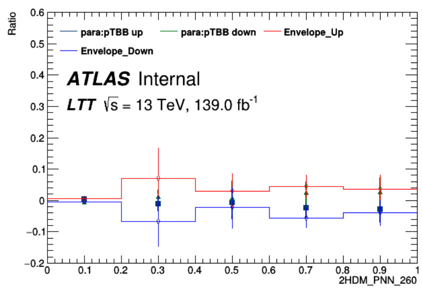
The normalised ratio of the PNN distributions of the variation mentioned above to the nominal are shown in Appendix B.3, in Figure B.9 (B.10) for the generator uncertainty, Figure B.11 (B.12) for the intra-PDF set uncertainties, Figure B.13 (B.14) for the inter-PDF set and α_s uncertainty, Figure B.15 (B.16) for the CKKW and QSF uncertainties in the



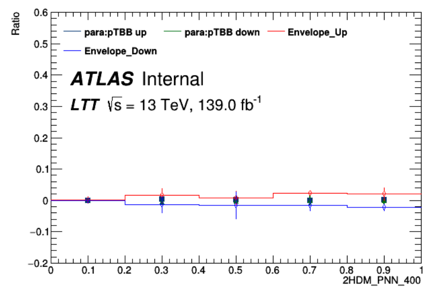
(a) (b)

(c) (d)

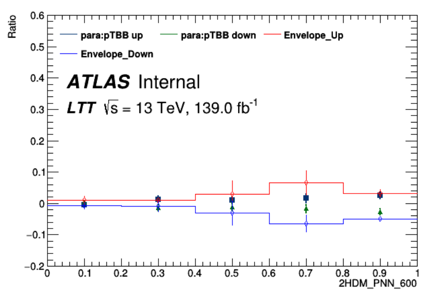
Figure 7.32: SLT channel: PNN output distributions of the nominal sample weighted by parametrisation in p_T^{bb}, labelled as ‘para:pTBB Up(Down)’ versus envelope encapsulating the maximum shift of the variation for various resonant mass values: (a) 260 GeV, (b) 400 GeV, (c) 600 GeV, (d) 800 GeV.



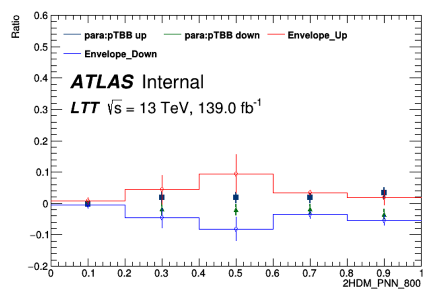
(a)



(b)



(c)



(d)

Figure 7.33: LTT channel: PNN output distributions of the nominal sample weighted by parametrisation in p_T^{bb}, labelled as 'para:pTBB Up(Down)' versus envelope encapsulating the maximum shift of the variation for various resonant mass values: (a) 260 GeV, (b) 400 GeV, (c) 600 GeV, (d) 800 GeV.

SLT (LTT) channel.

The yields of the up and down variations for the CKKW and QSF uncertainties are both smaller than the nominal sample yields. The solution for this issue is to normalise the central value of the up and down variations to the nominal sample, and the uncertainty is taken as the difference from the normalised up and down variations.

Uncertainties on Z + HF background in the Z + HF control region

Shape uncertainties on the Z + HF background are neglected in the Z + HF control region as they are found to be negligible in the $m_{\ell\ell}$ distribution included in the fit for this region, as shown in Appendix B.3.

7.5.2.3 Uncertainties on the single top background

All single-top uncertainties are derived by MC-to-MC comparison, and split into normalisation acceptance uncertainties and shape uncertainties.

As the Wt-channel contribution dominates over the t-channel and s-channel contributions, only the contribution from the Wt-channel is considered. Uncertainties due to PDF, ISR and FSR are evaluated using internal alternative weights present in all single-top nominal samples.

For PDF and α_s variations, the `PDF4LHC_NLO_30` PDF set is used for evaluation. The PDF set is comprised of 30 eigenvectors parameterising the uncertainties for all the PDFs and 1 parameter for the α_s variations. The variations of the PDF and α_s are combined by taking the square root of the quadrature sum of the 30 eigenvectors variations and the α_s variation in bins of PNN score distributions. The method used here follows the `PDF4LHC` recommendations [137].

Using the `PDF4LHC15_30` PDF set, the uncertainties are estimated by combining the difference between the error sets and the nominal set following the recommendations in Ref. [137].

Uncertainties from the parton shower, matrix element and single top interference are estimated using the same method described in Section 7.5.2.1 for estimating the uncertainties for the $t\bar{t}$ background.

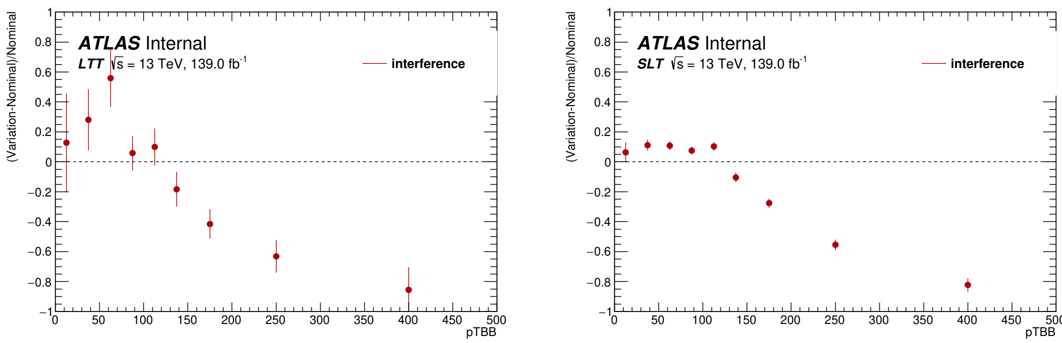
Uncertainty due to single-top interference is evaluated comparing the full simulated nominal Powheg +Pythia samples with diagram removal (DR) scheme to the alternative full simulation Powheg +Pythia sample with diagram subtraction (DS) scheme.

The normalisation acceptance uncertainties are shown in Table 7.15.

The shape uncertainty is derived separately for the SLT channel and the LTT channel. Only normalisation acceptance uncertainty is considered for the parton shower and matrix

Source	SLT	LTT
ME	± 0.022	± 0.15
PS	± 0.077	± 0.093
Single top interference	± 0.078	± 0.11
ISR	-0.047, +0.064	-0.045, +0.062
FSR	-0.054, +0.043	-0.069, +0.035,
PDF	± 0.032	± 0.032
Total	±13.7%	±21.1%

Table 7.15: Size of normalisation acceptance uncertainties for single-top background for the HH analysis.



(a) a

(b) b

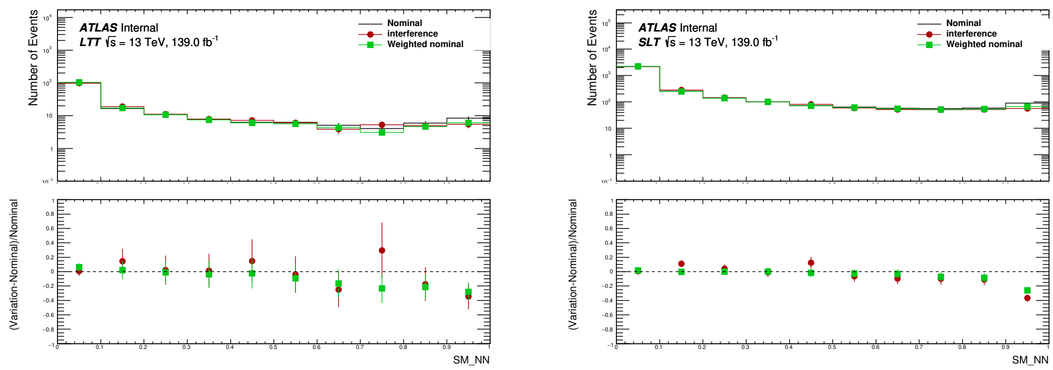
Figure 7.34: LTT (a) and SLT channels (b): shape only Parametrisation in bins of p_T^{bb} distribution for Wt DS/DR interference uncertainty.

element uncertainties as no obvious shape is observed in the NN or the PNN scores, as shown in Appendix B.3 in Figure B.18, B.20 in the SLT channel and in Figure B.19, B.21 in the LTT channel for the parton shower, matrix element uncertainties, respectively.

For the single-top interference uncertainty, the parametrisation is extracted from the ratio of the variation to the nominal full simulation samples in bins of p_T^{bb} where it shows the most obvious trend and largest shift from unity. Other MVA discriminant inputs are also tested, but most of them either do not have an obvious shape or the variation size is much smaller compared to the p_T^{bb} variation, and therefore cannot provide sufficient closure in the MVA output distributions. The parametrisation is shown in Figure 7.34 (a), (b) for the SLT and the LTT channels respectively.

The parametrisation is applied on the nominal sample as weights and the weighted sample is passed through the MVA classification. The NN score distributions of the weighted sample and of the variation sample are shown in Fig 7.35.

The closure in the PNN scores for various resonance mass are shown in Appendix B.3 in Figure B.22 and in Figure B.23 for the SLT and LTT channels, respectively. Good closure is seen across various mass points of NN classification and in most of the bins,



(a) (b)

Figure 7.35: LTT (left) and SLT (right) channels : shape only SM NN score of the Wt DS/DR interference uncertainty .

therefore, the parametrisation in p_T^{bb} is adopted.

7.5.2.4 Uncertainties on single Higgs boson processes

Uncertainties on cross-section and BR are applied to all single Higgs boson processes included in the analysis as backgrounds (following recommendations from the LHC Higgs Working Group [21]).

For all single Higgs boson processes, uncertainties arising from PDF, α_s and scales are derived using internal alternative weights present in all the single Higgs boson nominal samples.

The uncertainties due to the choice of PDF and α_s variations are estimated using the same method as described in Section 7.5.2.3 for evaluating the single top background uncertainties.

The ISR and FSR uncertainties are estimated using the same method as described in Section 7.5.2.1.

Variations of the renormalisation and factorisation scales are used to estimate the uncertainty due to missing higher order corrections. For these scale variations, the 7-point scale variations of the renormalisation and the factorisation scale are used and they are combined by taking an envelope of all of the variations, similar to the method used in Section 7.5.2.2.

The uncertainties due to parton shower and matrix element are evaluated using the same method described in the previous sections.

Uncertainties for processes that have variations less than 1% are not considered.

For the ttH process, the matrix element NLO matching uncertainty is evaluated comparing the nominal Powheg + Pythia samples to alternative aMC@NLO + Pythia samples.

For the ZH process where the Higgs boson decays to a pair of b-jets ($ZHbb$), the ISR

Process	SLT	LTT	Description
ttH	± 0.006	± 0.008	PDF+α_s
ttH	± 0.003	± 0.006	Scales
ttH	± 0.013	± 0.067	PS
ttH	± 0.001	± 0.01	ISR
ttH	-0.051, +0.032	-0.15, +0.055	FSR
ttH	± 0.0025	± 0.019	ME NLO matching
$ZHbb$	± 0.006	± 0.005	PDF+α_s
$ZHbb$	± 0.030	± 0.025	Scales
$ZHbb$	± 0.11	± 0.037	PS
$ZH\tau\tau$	± 0.0077	± 0.012	PDF+α_s
$ZH\tau\tau$	± 0.022	± 0.028	Scales
$ZH\tau\tau$	± 0.055	± 0.15	PS
VBF$H\tau\tau$	-0.003, +0.004	-0.003, +0.004	Scales
VBF$H\tau\tau$	± 0.021	± 0.021	PDF+α_s
ggF$H\tau\tau$	± 0.039	± 0.039	Scale
ggF$H\tau\tau$	± 0.032	± 0.032	PDF+α_s

Table 7.16: Relative size of single Higgs boson acceptance uncertainties. $ZHbb$, $ZH\tau\tau$, VBF$H\tau\tau$ and ggF$H\tau\tau$ refer to the $Z(H \to bb)$, $Z(H \to \tau\tau)$, VBF Higgs production with $H \to \tau\tau$ and ggF Higgs production with $H \to \tau\tau$ processes, respectively. Table reproduced from analysis internal notes.

and FSR uncertainties are found to be negligible in the analysis presented in Ref. [205], thus they are neglected in this analysis.

The size of the acceptance uncertainties on the normalisations are reported in Table 7.16. Normalisation uncertainties smaller than 1% are neglected in the analysis. After taking out the normalisation effects, shape effects were checked on the PNN score distributions and found to be negligible and therefore are not considered in the analysis.

7.5.2.5 Uncertainties on other minor backgrounds

For minor backgrounds, i.e. single-top (s- and t-channels), Z+light jets, W+jets and Diboson, acceptance uncertainties are only applied on the normalisation, and their size is taken from the the analysis presented in Ref. [204, 205]. On single-top production, an acceptance uncertainty of 20% is applied in the s- and t-channels. An acceptance uncertainty of 23% is applied on Z+light jets. On W+jets an acceptance uncertainty of 37% is applied in order to cover in addition for the fake-τ contribution. This uncertainty was estimated comparing the MC and the data-driven prediction for W+jets fakes in the 0-b-tag region of the $\tau_{\text{lep}}\tau_{\text{had}}$ channel in the previous iteration of this analysis. Acceptance uncertainties of 25%, 26% and 20% are applied on WW, WZ and ZZ, respectively.

7.5.3 Signal uncertainties

Acceptance uncertainties on the resonant and non-resonant signals are evaluated by comparing the nominal signal samples to alternative MC samples for parton shower, PDF and scale variations.

7.5.3.1 Resonant di-Higgs production

The parton shower uncertainties for the resonant signals are estimated by comparing the nominal samples showered with HERWIG to alternative samples showered with PYTHIA which are available in full-simulation for the $m_X = 500$ GeV and the $m_X = 1000$ GeV mass points. The parton shower normalisation acceptance uncertainty is found to be 6% (3%) for $m_X = 500$ ($m_X = 1000$) GeV in the SLT and 4% (7%) for $m_X = 500$ ($m_X = 1000$) GeV in the LTT. Thus, a conservative 6% (7%) uncertainty is adopted on the normalisation for all mass points in the SLT (LTT) channel.

After taking out the normalisation effect, the acceptance uncertainty on the shape of the PNN distribution is evaluated. Figures B.24 and B.25 show the comparisons of the PNN distributions obtained from the nominal (black) and alternative (blue) signal samples for the $m_X = 500$ GeV and the $m_X = 1000$ GeV mass points for the SLT and LTT SRs respectively. A linear fit to the ratio of the two distributions is performed to parameterise the shape variation as a function of the PNN score. The linear function in the PNN score obtained from the fit was performed at $m_X = 500$ GeV, which has the largest slope, and this function is used to obtain the templates for the variations for all mass points. Relevant plots can be found in Appendix B.3, Figures B.24 and B.25.

Scale, PDF and α_s uncertainties on the resonant signal are found to be negligible (less than 1%) and therefore not included in the analysis.

The resonant samples are generated with AF2 for masses up to 1 TeV (from 1.1 TeV full simulation is used). Therefore the difference between the AF2 samples and the full simulation samples is checked, using the 400 GeV mass point sample corresponding to 2017 and 2018 data-taking periods. The difference is less than 1% in the SLT channel and 2.9% in the LTT channel, and there is no obvious shape difference between the two types of simulation, as shown in Figure 7.36. The larger difference in acceptance times efficiency in LTT originates from the τ_{had} used in LTT triggers, for which no dedicated recommendations for AF2 exist. Therefore a 3% acceptance uncertainty is assigned to all signal samples in the LTT channel simulated with AF2 (up to 1 TeV mass point).

All acceptance uncertainties on the signal normalisation obtained in the different signal regions are shown in Table 7.17.

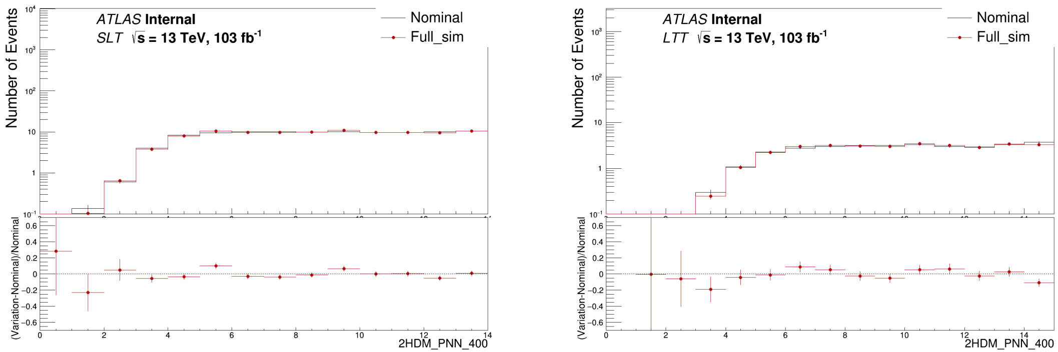


Figure 7.36: Comparison of the di-Higgs $\tau_{\text{lep}}\tau_{\text{had}}$ signal PNN distributions obtained from the nominal (AF2) and alternative (full simulation) signal samples for the m_X = 400 GeV for the SLT (left) and the LTT (right) channel. The binning is the same as the final fit binning.

Process X, GeV	SLT	LTT	Description
m_X = 400	<0.01	0.029	AF2
m_X = 500	0.06	0.05	PS
m_X = 1000	0.03	0.07	PS

Table 7.17: List and relative size of the ggF HH resonant signal acceptance uncertainties in SLT and LTT. Table reproduced from analysis internal notes.

7.5.3.2 Non-resonant di-Higgs production

ggF non-resonant di-Higgs production The PS uncertainties for the ggF non-resonant signal are estimated by comparing the nominal samples showered with Pythia to alternative samples showered with Herwig. The overall PS acceptance uncertainty on the normalisation is found to be 7.6% in SLT and 7.5% in LTT.

PDF and α_s uncertainties are evaluated using the same method as described in the previous sections. The variations have less than 1% shift from the nominal, and therefore these uncertainties are neglected.

Scale uncertainties are evaluated using the 7-point scale variations. An envelope is used to encapsulate all of the variations. The normalisation uncertainty of the envelope is 1.2% (1%) in SLT (LTT).

The acceptance normalisation uncertainties are summarised in Table 7.18. After taking out the normalisation effects, shape effects were checked on the MVA score distributions and found to be negligible and therefore are not considered in the analysis.

VBF non-resonant di-Higgs production The PS uncertainty for the VBF signal sample is evaluated by comparing the Pythia nominal sample to an alternative sample showered using Herwig. The normalisation uncertainty is found to be 6.3% (2.1%) for the SLT (LTT) channel. No significant shape impact was observed on the final discriminant output.

Process	SLT	LTT	Description
SM	0.076	0.075	PS
SM	0.012	0.010	Scales

Table 7.18: List and relative size of ggF HH non-resonant signal acceptance uncertainties in SLT and LTT.

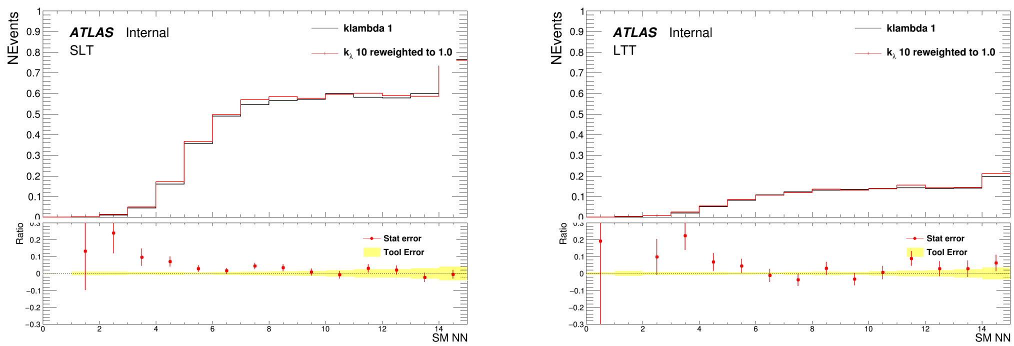


Figure 7.37: Comparison of the κ_λ = 1 and 10 ggF samples in the NN output for the SLT (left) and the LTT (right) channel The "Tool Error" band shows the statistical uncertainty on the applied weights, originating from the limited size of the samples used to derive them.

Scale uncertainties are evaluated using the internal weights of 7-point scale variations, the normalisation uncertainty is found to be negligible.

PDF and α_s uncertainties are evaluated using the same approach as described in previous sections. The normalisation uncertainty is less than 1% and therefore is neglected.

7.5.3.3 Uncertainties on κ_λ scan

As described in Section 5.3.2, the non-resonant HH signal assuming various κ_λ values are produced by a reweighting method. The uncertainties on each sample due to this reweighting needs to be accounted for. This section outlines how the uncertainties on the various κ_λ samples via ggF and VBF production are calculated.

Uncertainties on ggF As the weights are derived using truth-level samples, additional shapes might be introduced when these truth-level samples are passed to the analysis-level pre-selection; therefore the reweighting method needs to be validated. This is done by comparing the SM non-resonant HH sample with the κ_λ =10 sample reweighted to κ_λ =1. The comparison is shown in Figure 7.37. Good closure is observed between the two distributions, suggesting that the pre-selection does not introduce visible effects to the distribution after the weights are applied.

Nevertheless, it was found that the choice of sample used as input for the reweighting procedure affects the obtained signal acceptance at different values of κ_λ. This dependence is shown in Fig. 7.38, using either the generated κ_λ =1 or κ_λ =10 sample as input. The maximal discrepancy of 3% (SLT) or 8% (LTT) between the measured values is applied as the uncertainties on the normalisation of the ggF sample after the reweighting. Finally, the acceptance uncertainties originating from the choice in PS, PDF and α_s value or QCD scales have been derived for both the $\kappa_\lambda = 1$ and the $\kappa_\lambda = 10$ sample, using the same method described in previous sections. No shape dependence has been found in these uncertainties. The acceptance uncertainties are summarised in Tab. 7.19. Since these could not be evaluated for all κ_λ values investigated in the scan, the greater uncertainty is applied to all signals with κ_λ not equal to 1 for each of the three considered sources of uncertainty.

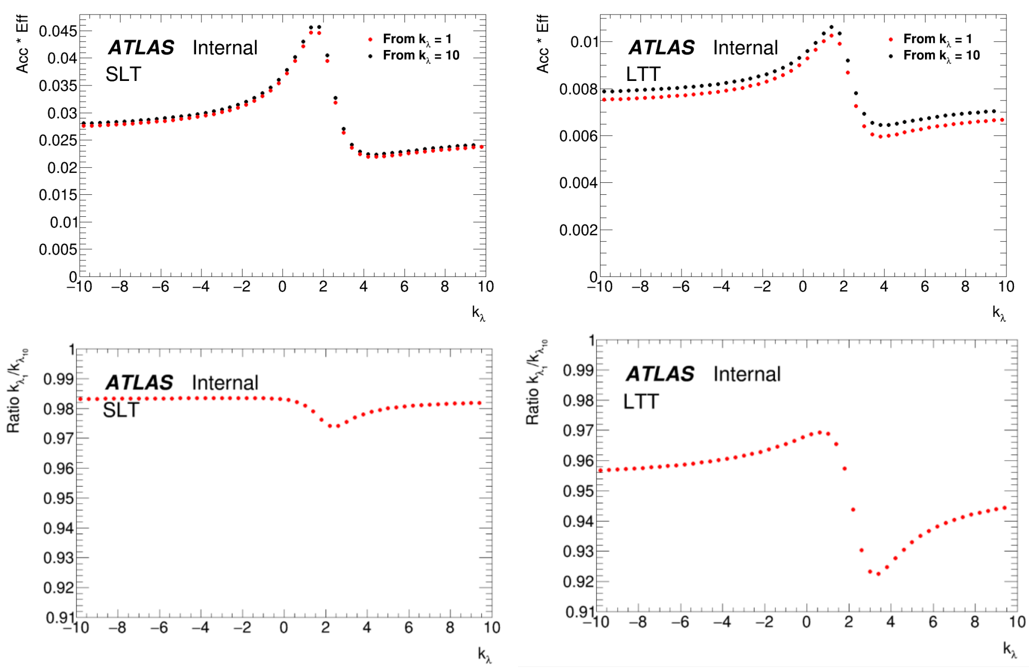


Figure 7.38: Top row shows the acceptance × efficiency at different κ_λ values depending on whether the reweighting procedure is applied on the ggF sample with $\kappa_\lambda = 1$ (nominal) or $\kappa_\lambda = 10$ (alternative), bottom row shows the ratio of these acceptance × efficiencies.

Uncertainties on VBF The QCD scale and PDF+α_s uncertainties for the SM and $\kappa_\lambda \in \{0, 2, 10\}$ samples are listed in Tab. 7.20. Additionally, the PS uncertainty is evaluated for the SM sample and at $\kappa_\lambda = 10$. For each of these sources of uncertainty, the largest one is used to define the respective variation for the prediction in all non-SM cases.

An uncertainty on the closure of the linear combination of VBF samples is evaluated in both SLT and LTT and applied to the normalisation of the VBF sample after the linear

Process	SLT	LTT	Description
SM	0.076	0.075	PS
SM	0.0061	0.0072	PDF+α_s
SM	0.012	0.010	Scales
$\kappa_\lambda = 10$	0.10	0.085	PS
$\kappa_\lambda = 10$	0.0074	0.0072	PDF+α_s
$\kappa_\lambda = 10$	0.0087	0.0074	Scales

Table 7.19: List and relative size of ggF HH non-resonant signal acceptance uncertainties in SLT and LTT.

combination. Four different bases of κ_λ values are tested. Closure plots are shown in Figure B.26 in Appendix B.3. The maximal difference is found to be of 2.2% (SLT) and 5.8% (LTT) between the multivariate discriminator distributions for the MC and reweighted κ_λ samples, and it comes from the linear combination of VBF samples ($\kappa_\lambda = 0, 1, 10$) to $\kappa_\lambda = 2$ in all cases.

Process	SLT	LTT	Description
SM	6.3 %	2.1 %	PS
SM	0.15 %	0.23 %	PDF+α_s from NNPDF30NLO
SM	0.86 %	0.63 %	Scales
$\kappa_\lambda = 0$	0.28 %	0.21 %	PDF+α_s from NNPDF30NLO
$\kappa_\lambda = 0$	1.16 %	0.91 %	Scales
$\kappa_\lambda = 2$	0.27 %	0.23 %	PDF+α_s from NNPDF30NLO
$\kappa_\lambda = 2$	1.08 %	0.89 %	Scales
$\kappa_\lambda = 10$	7.79 %	5.73 %	PS
$\kappa_\lambda = 10$	0.26 %	0.37 %	PDF+α_s from NNPDF30NLO
$\kappa_\lambda = 10$	1.25 %	0.99 %	Scales

Table 7.20: List and relative size of VBF HH non-resonant signal acceptance uncertainties in SLT and LTT. Table reproduced from analysis internal notes.

7.5.3.4 Uncertainties on HEFT intepretation signals

To evaluate the uncertainties due to the reweighting procedure described in Section 5.3.2, truth-level samples of 100,000 events are generated with the Powheg-Box [v2] generator [206] interfaced with Pythia 8.244 for parton showering and hadronisation. The samples are generated for the ggF non-resonant signal, and for signals assuming the 7 BM and varying the values of the two Wilson coefficients c_{gghh}, c_{tthh} to ± 0.5 and 1.0 (with the other HEFT couplings taking their SM values). To avoid expensive simulation of these events, they were instead compared using a truth-level analysis.

In order to estimate the uncertainties due to reweighting, a set of selections are applied on the truth level physics objects of these samples to mimic the event pre-selection applied in the SLT and LTT channels:

- Common to SLT and LTT:
  - $m_{bb} < 150$ GeV
  - $m_{\tau\tau} > 60$ GeV
  - τ_{had} $|\eta| < 2.3$
  - 2 truth b quarks, with leading b quark with $p_{\mathrm{T}} > 45$ GeV, subleading b quark with $p_{\mathrm{T}} > 20$ GeV (and require parents of both to be Higgs bosons)
  - One electron/muon, electron with $|\eta| < 2.47$ and not $1.37 < |\eta| < 1.52$, muon with $|\eta| < 2.7$
  - Minimum lepton cut: e/μ with $p_{\mathrm{T}} > 7$ GeV in order to be considered in the selection
- Specific to SLT: e/μ with $p_{\mathrm{T}} > 27$ GeV, $\tau p_{\mathrm{T}} > 20$, veto events with additional leptons with $p_{\mathrm{T}} < 27$
- Specific to LTT: e/μ with $p_{\mathrm{T}} > 18/15$ GeV, and with $p_{\mathrm{T}} <$ SLT cut, τ with $p_{\mathrm{T}} > 30$ GeV, veto events with additional leptons with $p_{\mathrm{T}} < 18/15$ GeV

No NN/PNN is applied, because it is hard to mimic all the inputs and their correlations at truth level.

After the selection defined above, the m_{HH} distributions of the generated BM signals are compared to the ones of the BM signals reweighted from the SM signal, as shown in Figures B.35 and B.36 in Appendix B.3. Similarly, the m_{HH} distributions are checked for the generated $c_{gghh}, c_{tthh} = \pm 0.5, 1.0$ signals and reweighted from the SM signal, as shown in Figures B.37 and B.38 for the SLT and LTT channel. In both checks, no obvious shape dependence is found between the generated and reweighted distributions. Therefore, only the normalisation acceptance uncertainties are considered. For the c_{gghh}, c_{tthh} scans, the largest uncertainty is taken from each scan and is applied on the whole range of each scan. The normalisation uncertainties are summarised in Table 7.21 for the 7 BM and Table 7.22 for the c_{gghh}, c_{tthh} scans.

7.5.4 Uncertainties on fake background

The following uncertainties are considered for the fake background:

- The statistical uncertainty on the $\mathrm{FF}_{t\bar{t}}$, $\mathrm{FF}_{\mathrm{QCD}}$ and $\mathrm{r}_{\mathrm{QCD}}$ values are considered. They are propagated to the final estimates of the fake background.
- A conservative 30% uncertainty is assigned to all non-$t\bar{t}$ background being subtracted from the data. This uncertainty is derived by varying the FF up and down by 30% when applying on the non-$t\bar{t}$ background passing the anti-ID selection.

Region	$\tau_{\text{lep}}\tau_{\text{had}}$ (SLT)	$\tau_{\text{lep}}\tau_{\text{had}}$ (LTT)
BM 1	15%	8%
BM 2	13%	6%
BM 3	10%	14%
BM 4	2%	8%
BM 5	8%	17%
BM 6	3%	14%
BM 7	1%	12%

Table 7.21: Uncertainties on the yield for the 7 benchmarks from the HEFT reweighting. They are evaluated from the difference in yields between the generated and reweighted samples for each HEFT benchmark.

Region	$\tau_{\text{lep}}\tau_{\text{had}}$ (SLT)	$\tau_{\text{lep}}\tau_{\text{had}}$ (LTT)
c_{tthh}	4%	10%
c_{gghh}	10%	11%

Table 7.22: Uncertainties on the yield of the c_{tthh} and c_{gghh} scans from the HEFT reweighting. The envelope of the yield differences between the generated and reweighted samples with the coefficient values of ±0.5 and 1.0 is used for the full scan.

- The uncertainty due to $t\bar{t}$ modelling was estimated by the difference between the fake background derived with and without the $t\bar{t}$ reweighting (more details in Section 7.3.1). An alternative approach was studied by the author: using the $t\bar{t}$ variation samples to evaluate the PS, matrix element, ISR and FSR uncertainties to replace the nominal $t\bar{t}$ samples when deriving the FF. The fake background derived using the variation samples is correlated to the corresponding ones in $t\bar{t}$ modelling uncertainties estimations. The PNN score distributions of the fake background derived using the variation samples are shown in Appedix B.3, Figures B.27- B.34.

  The author compared these two methods, where the first one was adopted in the end since the variation is observed to be higher and hence covers all fake backgrounds derived using variation samples.

- The r_{QCD} value is highly sensitive to the $t\bar{t}$ background normalisation and shape. Given the fact that the FF_{QCD} and $\text{FF}_{t\bar{t}}$ are very similar, the r_{QCD} in practice does not have big impact on the combined fake factor. Therefore the uncertainty on r_{QCD} is estimated by varying the value from 0 to 0.5. The impact of assigning such a conservative uncertainty was checked and was found to be very small.

7.5.5 Summary of Systematics Uncertainties

In Figure 7.39, the rankings of the nuisance parameters considered in the combined fit of SLT and LTT (more details in Section 7.6.1) are shown. The nuisance parameters are arranged in descending order of their fractional impact on the non-resonant HH production signal strength μ, $\Delta\mu/\Delta\mu_{tot}$. The $\Delta\mu/\Delta\mu_{tot}$ is represented by the horizontal bars in the plot with values read on the top blue x-axis. The dashed (un-dashed) blue area represents the fractional impact on μ when the nuisance parameter is varied up (down) by one standard deviation (with this parameter fixed to the variation and fit again allowing all other nuisance parameters to vary). The black round dots, read on the bottom x-axis, represent the shift of the fitted value θ of the nuisance parameters from their nominal value θ_0, with respect to their standard deviation $\Delta\theta$. The error bars corresponding to the dots are the fitted uncertainties, relative to the nominal uncertainties. The red open circle, again read on the bottom x-axis represents the fitted value of the $t\bar{t}$ and Z + HF normalisation which are freely floating in the fit, with the error bars representing the uncertainty. The statistical uncertainty is also included as nuisance parameters with Poissonian priors. The name 'LepHad_SLT(LTT)_SR_MVA_bin_14' represents the statistical error on the 14^{th} bin in the SM NN score distribution. More ranking plots can be found in Appendix B.4.

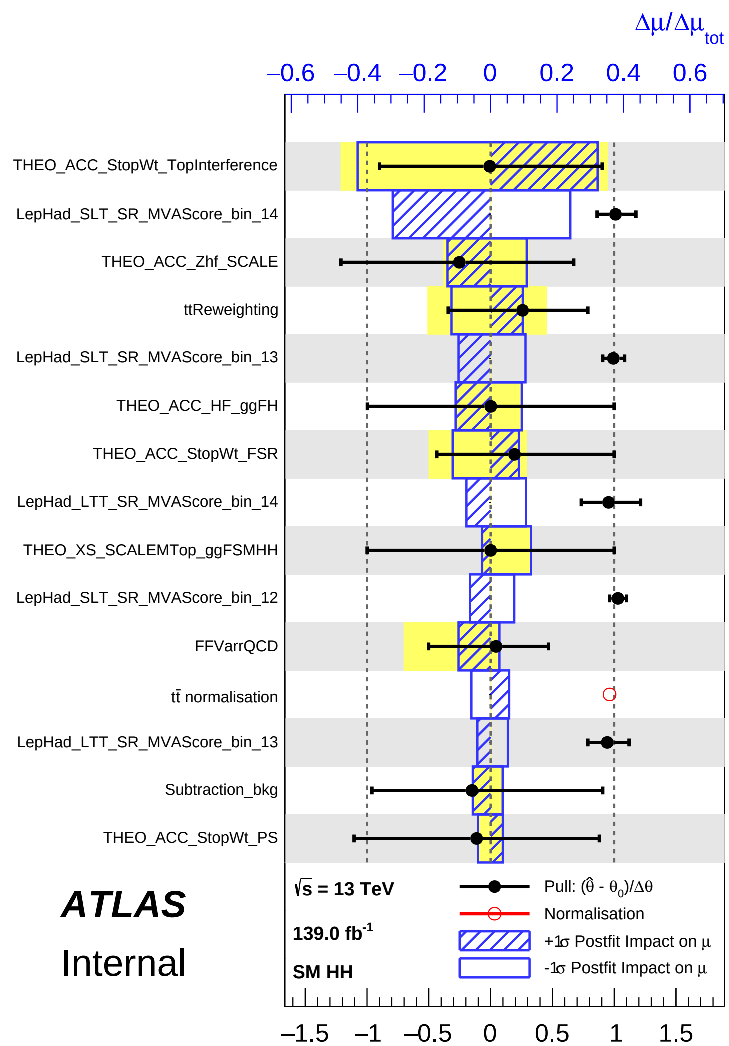


Figure 7.39: Nuisance parameters rankings for the di-Higgs $\tau_{\text{lep}}\tau_{\text{had}}$ fit to data for the SM fit.

7.6 Results

The set-up used for the profile likelihood fit is described in Section 7.6.1. In Section 7.6.2, results obtained from $\tau_{\text{lep}}\tau_{\text{had}}$-only fit (using only the $\tau_{\text{lep}}\tau_{\text{had}}$ signal regions and the $Z+\text{HF}$ control region) are presented. The result from the likelihood fit shows that the data is compatible with the null hypothesis for all signals. Therefore, the data is used to set upper limits at 95% CL on the signal cross-sections. The likelihood fit is performed again on the combined $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ signal regions and the $Z+\text{HF}$ control region, and the results are presented in Section 7.6.3. The fit results in $b\bar{b}\tau^+\tau^-$ channel are then combined with the $b\bar{b}\gamma\gamma$ channel for upper limits on the non-resonant SM HH production cross-section and limits on the self-coupling modifier κ_λ, and with the $b\bar{b}\gamma\gamma$ and $b\bar{b}b\bar{b}$ channels for upper limits on the narrow-width scalar X production cross-section as a function of the resonance mass. These combined results are presented in Section 7.6.4.

7.6.1 Fit setup

The profile likelihood fit is first performed on the MVA output in the $\tau_{\text{lep}}\tau_{\text{had}}$ SLT, LTT regions, together with the $m_{\ell\ell}$ distribution in the $Z+\text{HF}$ control region referred to as the $\tau_{\text{lep}}\tau_{\text{had}}$-only fit. After checking the $\tau_{\text{lep}}\tau_{\text{had}}$-only fit results, a likelihood fit is then performed simultaneously on the MVA output in all three signal regions, the $\tau_{\text{lep}}\tau_{\text{had}}$ SLT, LTT and the $\tau_{\text{had}}\tau_{\text{had}}$ regions, together with the $m_{\ell\ell}$ distribution in the Z+HF control region. As described in Section 7.4, for the resonant HH production, PNNs are used for signal extraction, and for the non-resonant production, a neural network is used in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel (a boosted decision tree is used for the non-resonant production in the $\tau_{\text{had}}\tau_{\text{had}}$ signal region). Therefore, fits are performed separately on the non-resonant and resonant production, and for the resonant, fits are performed separately on each mass hypothesis.

The normalisation of the $t\bar{t}$ background and the normalisation of the $Z+\text{HF}$ background are determined in the fit as freely floating parameters. Relative acceptance uncertainties between CR and SRs are applied on these normalisations in the signal regions included in the fit as described in Section 7.5.

The parameter of interest (POI) is the signal strength μ. As described in Section 2.2.8.1, for the non-resonant SM case, the μ is relative to the ggF+VBF input signal cross-section times the $b\bar{b}\tau^+\tau^-$ channel decay branching ratio, 2.394 fb. For the resonant case, the μ is relative to the input cross-section of 1 pb times the $b\bar{b}\tau^+\tau^-$ branching ratio, which is 73 fb.

In the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the MVA outputs are binned in 1090 bins, where 990 bins with equal width are in the range from 0 to 0.99 and 100 bins with equal width are in the range from 0.99 to 1. A rebinning procedure is performed on the MVA output. First, a rebinning metric is defined as:

$$Z(I[k,l]) = 10\frac{n_s(I[k,l])}{N_s} + 5\frac{n_b(I[k,l])}{N_b}, \tag{7.7}$$

where

- $I[k,l]$ is an interval of the histograms, containing the bins between bin k and bin l;
- N_s is the total number of signal events in the histogram;
- N_b is the total number of background events in the histogram;
- $n_s(I[k,l])$ is the total number of signal events in the interval $I[k,l]$;
- $n_b(I[k,l])$ is the total number of background events in the interval $I[k,l]$.

The re-binning is then conducted using the following procedure:

1. Starting from the first bin k_0 on the right end of the MVA output histogram (MVA output = 1), increase the range of the interval $I(k_0,l)$ by adding one bin to the left;
2. Calculate the value of Z at each step;
3. Once $Z(I[k_0,l]) > 1$, rebin all the bins in the interval $I(k_0,l)$ into a single bin;
4. Repeat steps 1-3, starting this time from the last bin on the right (bin $l+1$, the next bin to the left of bin l) until the interval reaches the left end (0) of the histogram.

In addition, the binning algorithm requires at least 5 background events in each bin.

All sources of systematic uncertainties described in Section 7.5 are considered as nuisance parameters in the profile likelihood. Correlations of the nuisance parameters (NP) across the four regions are taken into account in the fit. All experimental uncertainties, modelling uncertainties from the same source (except for the fake background, which was estimated using a different method in the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channel), $t\bar{t}$ and $Z+\text{HF}$ background normalisations and relative acceptance uncertainties on $t\bar{t}$ and $Z+\text{HF}$ backgrounds are correlated across the three signal regions and one control region.

Each NP is split into shape and normalisation effects. The shape uncertainties are included as alternative histograms, while the normalisation effects are implemented as either flat (for the floating backgrounds) or Gaussian priors. The NPs are then processed to be symmetrised, smoothed and pruned: for one-sided experimental systematic uncertainties and uncertainties with up and down variation going to the same side, symmetrisation is applied to avoid under-constraint problems and improve the stability of fitting; then, smoothing is applied on systematic uncertainties with large statistical fluctuations to avoid instabilities; finally, systematic uncertainties which have less impact are pruned away to speed up the fitting process.

7.6.2 $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel results

The profile likelihood fit is first performed on the two signal regions in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and the Z + HF control region. This section describes the fit results obtained with this setting.

7.6.2.1 Post-fit background event yields

The expected signal and background event yields passing the signal region event pre-selection after the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$-only fit are shown in Table 7.23 (7.24) for the SLT (LTT) channel, extracted from the non-resonant HH fit result (using the NN discriminant). The overall background yield shows a high level of agreement with the data yield.

SampleName	Yield
Signal Samples	
ggF+VBF Non-resonant	6 ± 1
Background Samples	
Fake	33100 ± 1600
$t\bar{t}$	58500 ± 1500
single top	3730 ± 480
$Z \to \tau\tau + (cc, bc, bb)$	1640 ± 140
other	1301 ± 198
SM Higgs	153 ± 19
total Bkg	98450 ± 680
data	98456

Table 7.23: Post-fit event yields in the SLT signal region for the data, background and signal. Background names conventions are the same as used in Table 7.6.

In the SM non-resonant HH fit, the floating normalisations of the $t\bar{t}$ and Z + HF backgrounds are determined to be 0.96 ± 0.03 and 1.38 ± 0.11, respectively. The normalisation factors depend on the value of m_X, but they are found to be consistent with the numbers in non-resonant HH fit within uncertainties, as shown in Table 7.25.

7.6.2.2 Limit results: $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$-only fit

Table 7.26 shows the limit results obtained using SLT-only fit, LTT-only fit, and SLT, LTT combined fit ($\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$-only fit).

Upper limits at 95% CL are set on the non-resonant HH production cross-section as a function of κ_λ, in the range of -10 to 10. The κ_λ is excluded at 95% CL outside the range of [-3.8, 12.2] ([-3.8, 11.9]) for the observed (expected) limit. The results are shown in Figure 7.40.

SampleName	Yield
Signal Samples	
ggF+VBF Non-resonant	1.44 ±0.24
Background Samples	
Fake	1540 ± 180
$t\bar{t}$	3940 ± 190
single top	203 ± 40
$Z \to \tau\tau + (cc, bc, bb)$	537 ± 65
other	121.4 ±27
SM Higgs	24.0 ±4.8
total Bkg	6359 ± 80
data	6351

Table 7.24: Post-fit event yields in the LTT signal region for the data, background and signal. Background names conventions are the same as used in Table 7.7.

Mass point	Z + HF norm. factor	$t\bar{t}$ norm. factor
300 GeV	1.38 ± 0.11	0.97 ± 0.04
500 GeV	1.37 ± 0.11	0.97 ± 0.04
1000 GeV	1.39 ± 0.11	0.97 ± 0.04
1600 GeV	1.37 ± 0.11	0.97 ± 0.04

Table 7.25: Post-fit normalisation factors of $t\bar{t}$ and Z + HF background obtained from the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$-only fit for m_X=300, 500, 1000 and 1600 GeV fit.

	Obs.	$-2.0\,\sigma$	$-1.0\,\sigma$	Exp.	$1.0\,\sigma$	$2.0\,\sigma$
SLT-only fit						
$\sigma_{\mathrm{ggF+VBF}}$ [fb]	316.9	132.2	177.5	246.4	342.9	459.7
$\mu_{\mathrm{ggF+VBF}}$	10.9	4.5	6.0	8.3	11.6	15.6
LTT-only fit						
$\sigma_{\mathrm{ggF+VBF}}$ [fb]	498.7	386.6	519.0	720.3	1002.5	1343.9
$\mu_{\mathrm{ggF+VBF}}$	16.5	13.1	17.6	24.4	33.9	45.5
SLT + LTT combined fit						
$\sigma_{\mathrm{ggF+VBF}}$ [fb]	267.2	124.5	167.1	231.9	322.8	432.7
$\mu_{\mathrm{ggF+VBF}}$	9.2	4.2	5.7	7.9	10.9	14.7

Table 7.26: 95% CL expected and observed upper limits on the non-resonant HH production cross-section assuming SM kinematics using SLT only, LTT only and $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel fit. The signal strength is with respect to the SM cross-section, i.e. $\mu_{\mathrm{ggF+VBF}} = \sigma_{\mathrm{ggF+VBF}}/\sigma^{\mathrm{SM}}_{\mathrm{ggF+VBF}}$.

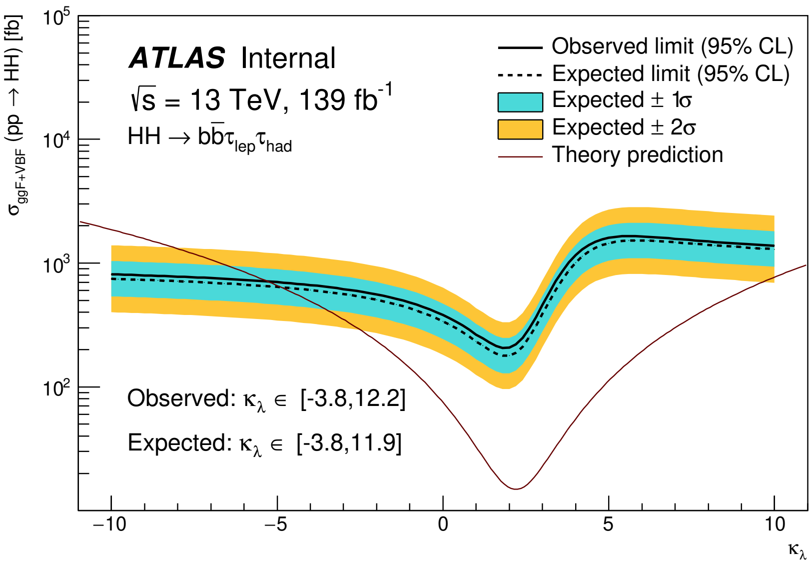


Figure 7.40: Observed and expected 95% CL exclusion limits on non-resonant HH production as a function of κ_λ in the using $\tau_{\text{lep}}\tau_{\text{had}}$-only fit. The observed (expected) limit on κ_λ is [-3.8, 12.2] ([-3.8, 11.9]).

Figures 7.41 show the 95% CL expected limits on the resonant HH production cross-section as a function of the resonance mass from the $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, respectively. No significant excess is observed in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel (largest local significance at 1.1 TeV corresponding to 1.6 σ).

7.6.3 Combined results of $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$

A combined fit is then performed on all three signal regions and the $Z+$HF control region. The combined results are shown in this section.

Table 7.27 shows 95% CL expected limit on the non-resonant HH production cross-section assuming SM kinematics from the combined fit.

	Obs.	$-2.0\,\sigma$	$-1.0\,\sigma$	Exp.	$1.0\,\sigma$	$2.0\,\sigma$
$\sigma_{\text{ggF+VBF}}$ [fb]	135.3	61.4	82.4	114.3	159.1	213.3
$\mu_{\text{ggF+VBF}}$	4.7	2.1	2.8	3.9	5.4	7.2

Table 7.27: 95% CL expected and observed upper limits on the non-resonant HH production cross-section assuming SM kinematics using $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ combined fit. The signal strength is with respect to the SM cross-section, i.e. $\mu_{\text{ggF+VBF}} = \sigma_{\text{ggF+VBF}}/\sigma^{\text{SM}}_{\text{ggF+VBF}}$.

The 95% CL expected and the observed limits on the resonant HH production cross-section, as a function of the resonance mass from the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ combined fit

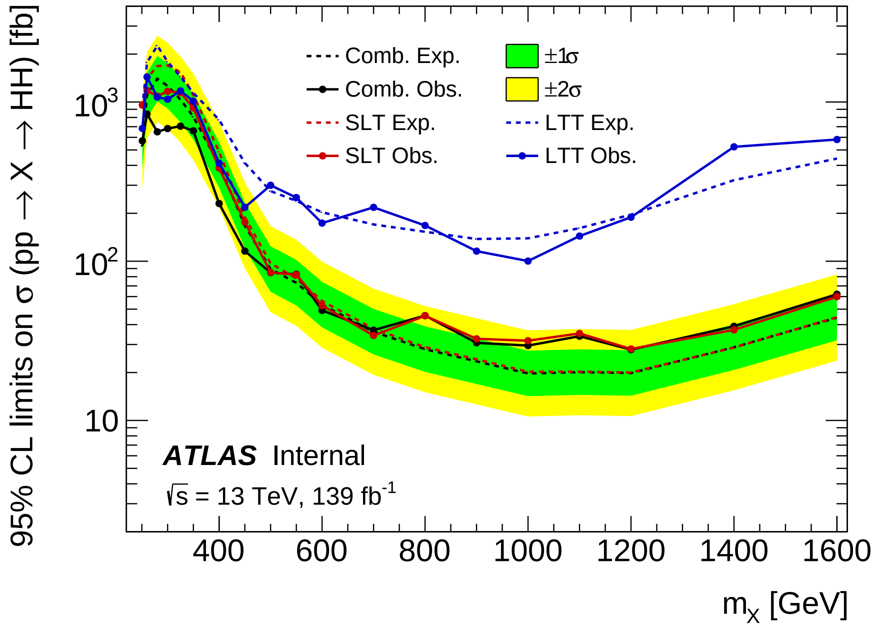


Figure 7.41: 95% CL expected and observed upper limitss on the resonant HH production cross-section as a function of the resonance mass in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. The black lines show the $\tau_{\text{lep}}\tau_{\text{had}}$ channel combined limit, whereas the red and blue lines show the limit for SLT and LTT channels separately. The solid and dashed lines represent the observed and expected limits, respectively.

are shown in Figure 7.42. Observed (expected) upper limits range from 26 – 950 fb (12 – 850 fb), depending on the mass of the resonance. In the combined fit, an excess (observed limit deviates greater than 2 σ from the expected limit) is present in the mass range between 800 and 1100 GeV. The maximum excess is found at 1000 GeV, which has a local deviation of 3.1 σ. The local p-values as a function of resonance mass are shown in Figure 7.43. The look-elsewhere effect is accounted by calculating the global significance using the up-crossing method described in Section 2.5.1. The estimated global significance is $2.1^{+0.4}_{-0.2}\sigma$.

The 95% CL expected and observed upper limits on the cross-section of non-resonant HH production as a function of κ_λ are shown in Figure 7.44. The 95% CL observed (expected) exclusion limit of κ_λ is $-2.4 < \kappa_\lambda < 9.2$ ($-2.0 < \kappa_\lambda < 9.0$).

Observed and expected upper limits are set at 95% CL on the ggF HH cross-section for the different HEFT shape benchmarks combining the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ result. The limits are shown in Figure 7.45 and Table 7.28. The least stringent limit is set for benchmark 2, at 199.8 fb (161.9 fb) for the observed (expected) limit, and the most stringent upper limits are set for benchmark 7 at 65.3 fb (50.6 fb). No SM HH production is considered. In addition, Figure 7.46 shows the 95% CL upper limits on the ggF HH cross-section as a function of the c_{gghh} and c_{tthh} HEFT Wilson coefficients. The observed (expected)

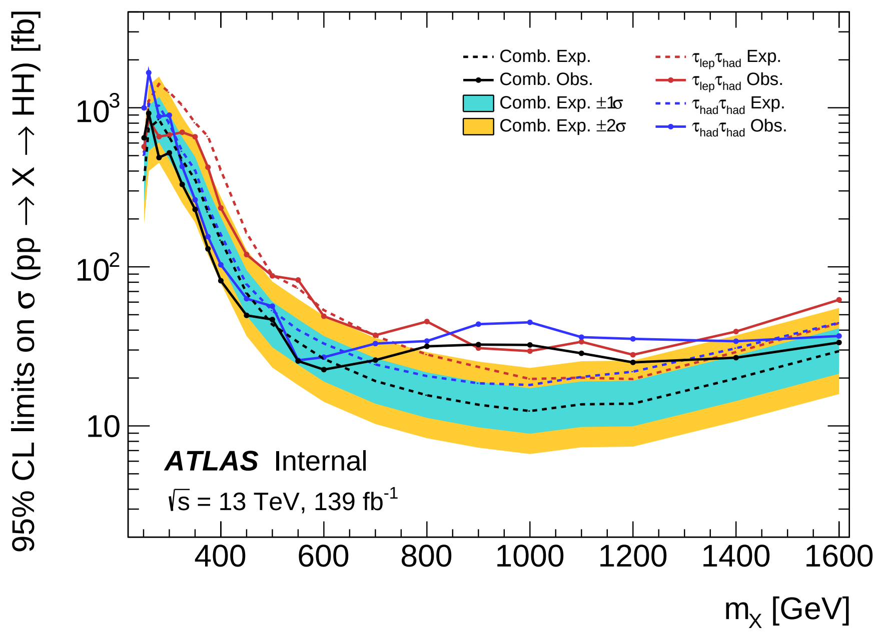


Figure 7.42: 95% CL expected and observed upper limitss on the resonant HH production cross-section as a function of the resonance mass. The dashed lines show the expected limits, while the solid lines show the observed limits. Black, red and blue represents the upper limit from the combination, $\tau_{\text{lep}}\tau_{\text{had}}$ channel and $\tau_{\text{had}}\tau_{\text{had}}$ channel, respectively.

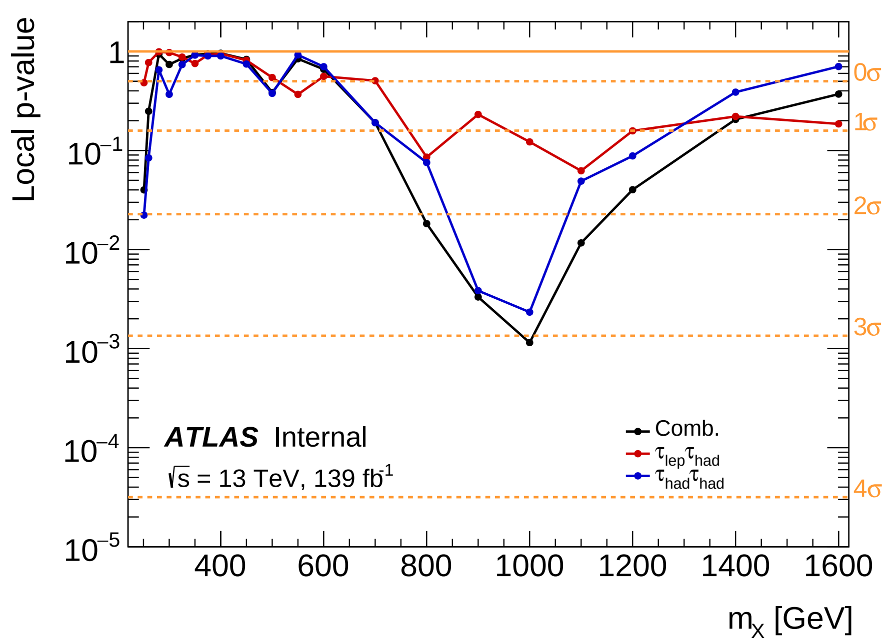


Figure 7.43: Local p-value of the background-only hypothesis test as a function of the mass of the resonance. The black line represents the p-value from the combination, while the red and blue lines are those for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channel, respectively. The horizontal dashed orange lines show the corresponding (2-sided) significance level.

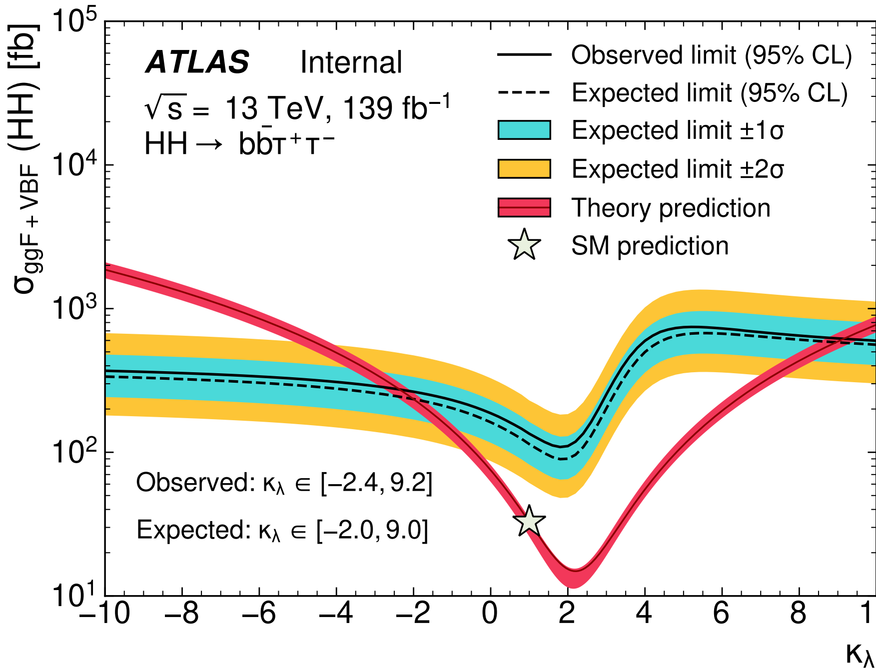


Figure 7.44: Observed and expected 95% CL upper limits on non-resonant $\sigma(pp \rightarrow HH)$ cross-section as a function of κ_λ using $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ combined fit. The observed (expected) exclusion limit at 95% CL on κ_λ in the $b\bar{b}\tau^+\tau^-$ channel is: [-2.4, 9.2] ([-2.0, 9.0]).

95% CL intervals on c_{gghh} and c_{tthh} are $-0.4 < c_{gghh} < 0.4$ ($-0.4 < c_{gghh} < 0.4$) and $-0.3 < c_{tthh} < 0.7$ ($-0.2 < c_{tthh} < 0.6$), respectively.

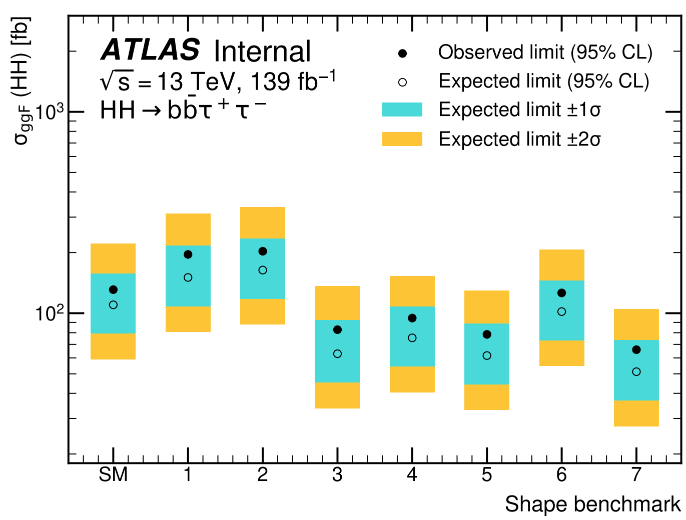
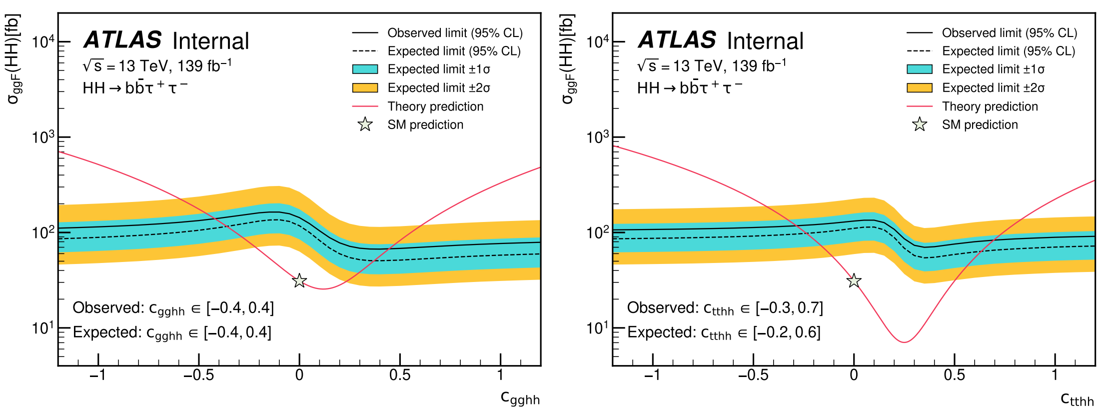


Figure 7.45: Observed and expected 95% CL upper limit on $\sigma_{\text{ggF}}(HH)$ for the seven HEFT shape benchmarks and SM, combining the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels. The expected limits assume no HH production.

	95% CL upper limit on ggF HH cross-section [fb]					
Benchmark	Obs.	-2σ	-1σ	Exp.	$+1\sigma$	$+2\sigma$
BM 1	195.8	80.8	108.4	150.5	216.9	312.0
BM 2	203.1	87.9	118.0	163.7	235.0	335.5
BM 3	82.8	33.8	45.3	62.9	92.1	136.3
BM 4	94.6	40.5	54.4	75.4	107.9	153.0
BM 5	78.5	33.1	44.4	61.6	89.1	129.1
BM 6	126.1	54.7	73.4	101.8	145.6	206.4
BM 7	65.9	27.5	36.9	51.2	73.5	104.8

Table 7.28: Observed and expected 95% CL upper limit on $\sigma_{\text{ggF}}(HH)$ for the seven HEFT shape benchmarks and SM, combining the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels. The expected cross-section limits assume no HH production.



(a) (b)

Figure 7.46: 95% CL upper limits on the ggF HH cross-section as a function of the c_{gghh} (a) and c_{tthh} (b) Wilson coefficients obtained from combining the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ results. The expected limits assume no HH production. The red curve represents the theory prediction where all Wilson coefficients are set to their SM values except for the one being scanned.

7.6.4 Combination of results with other HH decay channels

To maximise the sensitivity in the HH searches, the $b\bar{b}\tau^+\tau^-$ results are combined with other HH decay channels results. The $b\bar{b}\tau^+\tau^-$ non-resonant HH result is combined to the $b\bar{b}\gamma\gamma$ channel, which also has high sensitivity due to the high mass resolution in di-photon systems. For the resonant searches, the $b\bar{b}\tau^+\tau^-$ results are combined with $b\bar{b}\gamma\gamma$ and $b\bar{b}b\bar{b}$ channels. The $b\bar{b}\gamma\gamma$ channel is expected to have high sensitivity in the low mass regime and the $b\bar{b}b\bar{b}$ channel dominates over the high mass regime, which benefits from high selection efficiency in the more boosted event topology.

7.6.4.1 Combined results of $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$

The $b\bar{b}\tau^+\tau^-$ results are combined with the results in $b\bar{b}\gamma\gamma$ channel in terms of upper limits on the non-resonant HH production signal strength μ, and on cross-section as a function of κ_λ [192].

Systematic uncertainties relating to the data-taking conditions, such as those associated with the integrated luminosity and the pileup mis-modeling, are considered fully correlated among the input searches. Uncertainties related to physics objects used by multiple searches, such as jets and flavour-tagging, are treated as correlated. Theoretical uncertainties on simulated signal and background processes, such as the HH and single Higgs boson parton shower, scale and PDFs are treated as correlated where possible. HEFT reweighting uncertainties are also correlated between the two analyses.

The upper limits on signal strength are shown in Table 7.29, and the limits on the cross-section are shown in Figure 7.47. This combined result sets the current world-best observed upper limit on the SM HH cross-section [207–210], 3.1 times SM cross-section, at 95% CL. The κ_λ is excluded at 95% CL outside the range of [-1.0, 6.6] ([-1.2 , 7.2]) for the observed (expected) limit.

	Obs.	$-2.0\,\sigma$	$-1.0\,\sigma$	Exp.	$1.0\,\sigma$	$2.0\,\sigma$
$b\bar{b}\gamma\gamma$	4.3	3.1	4.1	5.7	8.8	14.3
$b\bar{b}\tau^+\tau^-$	4.6	2.1	2.8	3.9	5.9	9.4
Combined	3.1	1.7	2.2	3.1	4.7	7.3

Table 7.29: Observed and expected 95 % CL upper limits on the signal strength for SM HH production derived from the $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$ searches, and their statistical combination. Table reproduced from Ref. [192].

The 95% CL upper limits on the ggF HH cross-section for the combination of the $HH \to b\bar{b}\gamma\gamma$ and $HH \to b\bar{b}\tau^+\tau^-$ searches for the HEFT shape benchmarks are shown on Figure 7.48 and Table 7.30. The least stringent limit is set for benchmark 2, at 133.6 fb (133.5 fb) for the observed (expected) limit, and the most stringent upper limits are set for benchmark 7 at 50.2 fb (45.5 fb). No SM HH production is considered. The 95% CL combined upper limits on the ggF HH cross-section as a function of c_{gghh} and c_{tthh} are shown in Figure 7.49 and Table 7.31. The observed (expected) 95% CL intervals on Wilson coefficients c_{gghh} and c_{tthh} are $-0.3 < c_{gghh} < 0.4$ ($-0.3 < c_{gghh} < 0.3$) and $-0.2 < c_{tthh} < 0.6$ ($-0.2 < c_{tthh} < 0.6$).

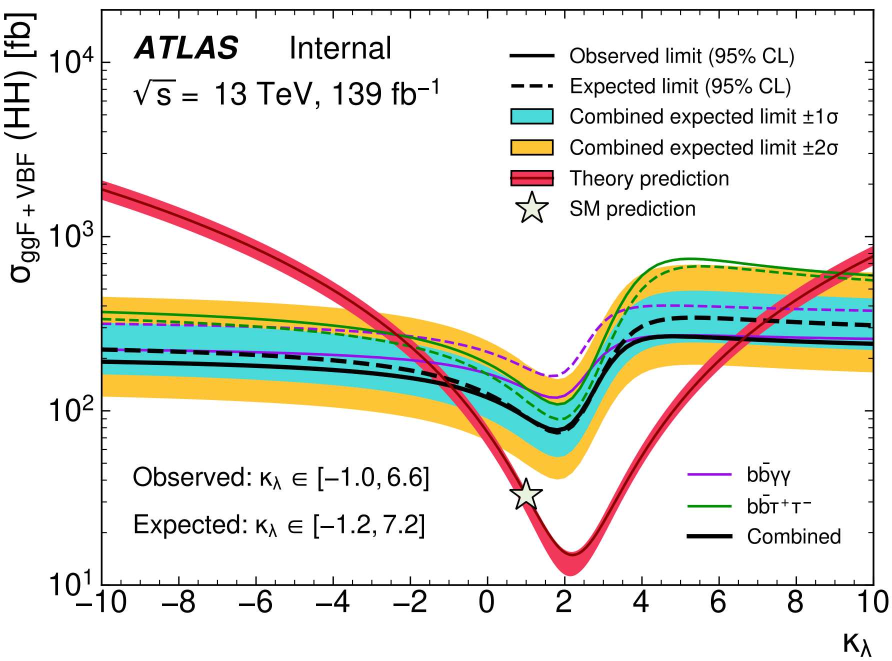


Figure 7.47: Observed and expected 95% CL exclusion limits on non-resonant $\sigma(pp \to HH)$ as a function of κ_λ for the individual channels and the combination of $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$. The observed (expected) combined limit on κ_λ is: [-1.0, 6.6] ([-1.2 , 7.2]). Image taken from Ref. [192].

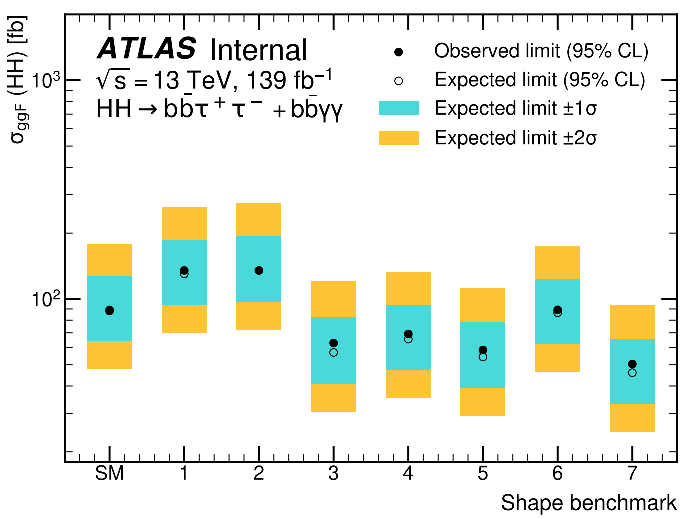


Figure 7.48: Observed and expected 95% CL upper limit on $\sigma_{\text{ggF}}(HH)$ for the seven HEFT shape benchmarks and SM obtained from $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$ analyses, and their combination. The expected limits assume no HH production. Image taken from Ref. [48].

7.6.4.2 Combined results of $b\bar{b}\tau^+\tau^-$, $b\bar{b}\gamma\gamma$ and $b\bar{b}b\bar{b}$

The $b\bar{b}\tau^+\tau^-$ results are combined with the results in $b\bar{b}\gamma\gamma$ and $b\bar{b}b\bar{b}$ channels in terms of upper limits on the resonant HH production cross-section. The correlation between

95% CL Upper limit on ggF HH cross-section [fb]										
	$b\bar{b}\gamma\gamma$		$b\bar{b}\tau^+\tau^-$		Combination					
Benchmark	Obs.	Exp.	Obs.	Exp.	Obs.	-2σ	-1σ	Exp.	$+1\sigma$	$+2\sigma$
BM 1	189.3	293.0	195.8	150.5	135.0	69.9	93.8	130.2	186.3	264.6
BM 2	183.0	267.3	203.1	163.7	134.9	72.5	97.3	135.0	192.9	273.1
BM 3	109.8	163.1	82.8	62.9	62.9	30.6	41.0	56.9	82.7	120.8
BM 4	111.2	155.7	94.6	75.4	69.2	35.2	47.2	65.6	93.5	132.0
BM 5	97.9	137.8	78.5	61.6	58.4	29.2	39.2	54.3	78.2	112.1
BM 6	134.5	189.0	126.1	101.8	89.2	46.4	62.3	86.5	123.4	174.1
BM 7	88.9	126.2	65.9	51.2	50.4	24.7	33.1	46.0	65.8	93.2

Table 7.30: Observed and expected 95% CL upper limits on $\sigma_{\text{ggF}}(HH)$ for the seven HEFT shape benchmarks and SM, obtained from $HH \to b\bar{b}\gamma\gamma$, $HH \to b\bar{b}\tau^+\tau^-$, and their combination. The expected cross-section limits assume no HH production. Table reproduced from Ref. [48].

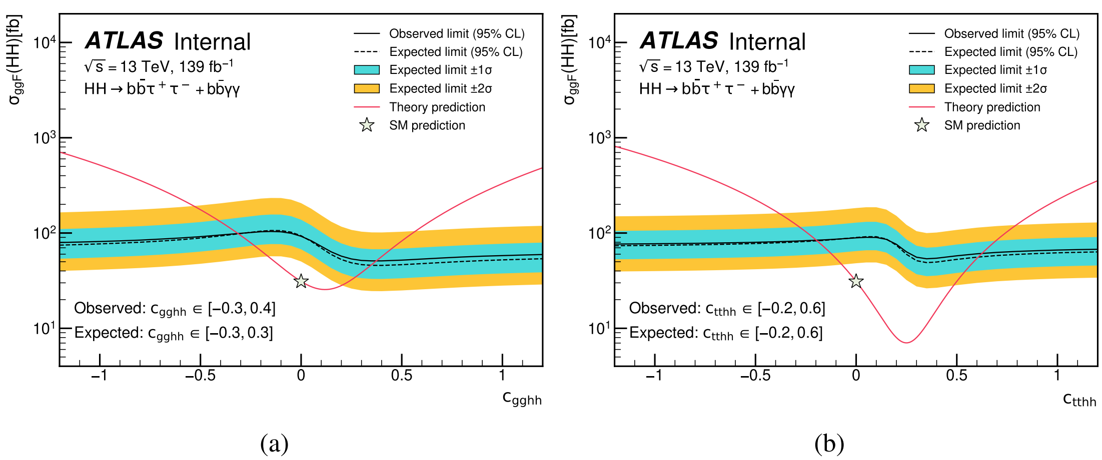


Figure 7.49: 95% CL upper limits on the ggF HH cross-section as a function of the c_{gghh} (a) and c_{tthh} (b) HEFT Wilson coefficients obtained for the $HH \to b\bar{b}\gamma\gamma$ and $HH \to b\bar{b}\tau^+\tau^-$ combination. The expected limits assume no HH production. The red curve represents the theory prediction where all Wilson coefficients are set to their SM values except for the one being scanned. Image taken from Ref. [48].

	$b\bar{b}\gamma\gamma$		$b\bar{b}\tau^+\tau^-$		Combination	
Wilson coefficient	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
c_{gghh}	$[-0.4, 0.5]$	$[-0.5, 0.7]$	$[-0.4, 0.4]$	$[-0.4, 0.4]$	$[-0.3, 0.4]$	$[-0.3, 0.3]$
c_{tthh}	$[-0.3, 0.8]$	$[-0.4, 0.9]$	$[-0.3, 0.7]$	$[-0.2, 0.6]$	$[-0.2, 0.6]$	$[-0.2, 0.6]$

Table 7.31: Allowed ranges for c_{gghh} and c_{tthh} for $HH \to b\bar{b}\gamma\gamma$ and $HH \to b\bar{b}\tau^+\tau^-$ analyses, and their combination at 95% CL. Table reproduced from Ref. [48].

the nuisance parameters are taken care of using the same procedure as described in the previous section.

Observed and expected limits at 95% CL on the resonant HH production cross-section for the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$, and $b\bar{b}b\bar{b}$ searches, and their statistical combination [192], are shown in Figure 7.50. The observed (expected) combined limits on cross-section range from 1.1 to 595 fb (1.2 to 392 fb) depending on the resonance mass. As mentioned before, the $b\bar{b}\gamma\gamma$ results show the largest sensitive low m_X, and the $b\bar{b}b\bar{b}$ limits dominate in high m_X. The $b\bar{b}\tau^+\tau^-$ search is most sensitive in the intermediate range, around 400–800 GeV. These three channels are therefore complementary to each other.

The largest excess is observed at 1.1 TeV, corresponding to a local significance of 3.2 σ. While a similar excess is observed in the $b\bar{b}\tau^+\tau^-$ channel, the compatibility is checked with the $b\bar{b}b\bar{b}$ excess. At 1.1 TeV, the level of compatibility of the fitted signal strength in the two channels corresponds to a p-value of 0.34 (>> 0.05), and hence the excess is compatible.

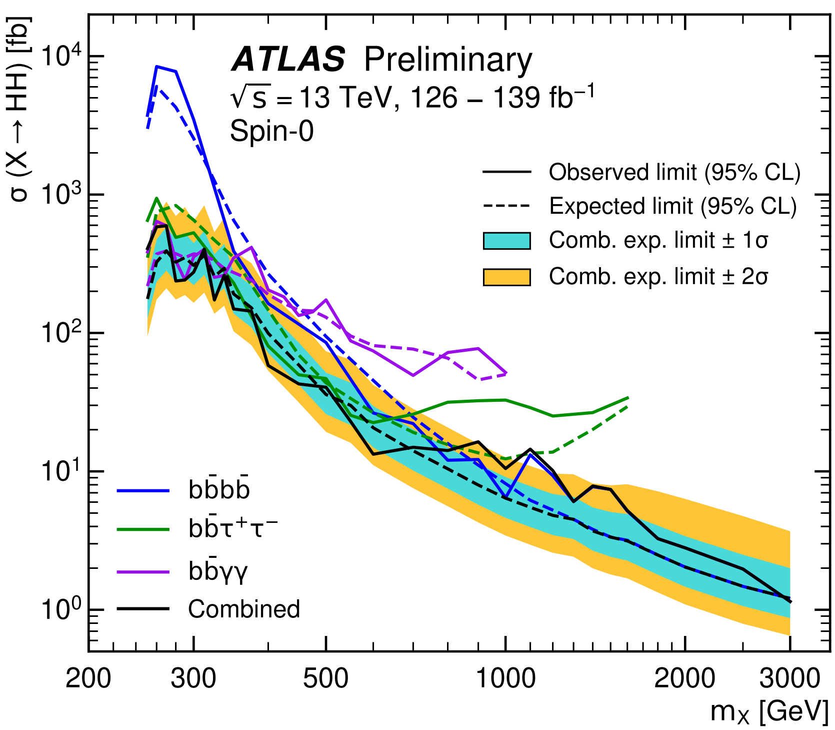


Figure 7.50: Expected and observed 95% CLupper limits on $\sigma(X \rightarrow HH$ for a spin-0 resonance as a function of its mass m_X in the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}b\bar{b}$ searches, and their statistical combination. Image taken from Ref. [192].

Chapter 8

Summary

This thesis presented the calibration on the c-jet mis-tagging efficiency in the DL1r b-tagging algorithm, followed by a search for Higgs boson pair (di-Higgs) production in the $b\bar{b}\tau^+\tau^-$ channel.

In the calibration work, the calibration strategy has been reviewed, the systematic uncertainties have been carefully taken into account, and a new orthogonal selection has been developed to reduce the statistical uncertainties in the high-p_{T} region of the scale factors. The calibration results on the DL1 and DL1r b-tagging algorithms with the PFlow jets and VR-Track jets have become the official recommendation for all ATLAS users who apply b-tagging in their analyses.

In the search for di-Higgs production, both non-resonant and resonant di-Higgs production are targeted. The background simulations are checked carefully, and a data-driven method is adopted to estimate the background due to a jet faking the τ_{had} lepton. An event pre-selection is defined to select the signal region events. Systematic uncertainties stemming from each source of the experiment, theory and simulation are thoroughly checked and applied in the analysis. A neural network and a set of parametric neural networks are used for the extraction of the signal process. Using the output of the NN and PNN discriminants, profile-likelihood fits are performed on $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel alone and the combination with the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channels. No significant excess of events above the expected background is observed in the non-resonant di-Higgs production. Observed (expected) upper limits are set at 95% CL on the non-resonant production cross-section of 4.7 (3.9) times the SM prediction. The resonant di-Higgs production cross-section is constrained to 95% to 26 – 950 fb (12 – 850 fb) for the observed (expected) limit, depending on the mass of the resonance. The maximum excess is found at 1000 GeV with a local deviation of 3.1 σ, corresponding to a global significance of $2.1^{+0.4}_{-0.2}$ σ. In addition, observed (expected) exclusion limits are set at 95% CL on κ_λ of $-2.4 < \kappa_\lambda < 9.2$ ($-2.0 < \kappa_\lambda < 9.0$); the observed (expected) 95% CL intervals on c_{gghh} and c_{tthh} are $-0.4 < c_{gghh} < 0.4$

$(-0.4 < c_{gghh} < 0.4)$ and $-0.3 < c_{tthh} < 0.7$ $(-0.2 < c_{tthh} < 0.6)$, respectively.

The $b\bar{b}\tau^+\tau^-$ results are then combined statistically with the other HH decay channels, $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^+\tau^-$. Again, no significant excess of events above the expected background is observed in the non-resonant search. When combined with the $b\bar{b}\gamma\gamma$ channel, the observed (expected) upper limits on the non-resonant production cross-section at 95% CL are tightened to 3.1 (3.1) times the SM prediction, achieving the current world-best result. The κ_λ at 95% CL is further excluded outside the range of $-1.0 < \kappa_\lambda < 6.6$ $(-1.2 < \kappa_\lambda < 7.2)$ for the observed (expected) limit. The observed (expected) 95% CL intervals on Wilson coefficients c_{gghh} and c_{tthh} are tightened to $-0.3 < c_{gghh} < 0.4$ ($-0.3 < c_{gghh} < 0.3$) and $-0.2 < c_{tthh} < 0.6$ ($-0.2 < c_{tthh} < 0.6$), respectively. The combination of the $b\bar{b}\tau^+\tau^-$, $b\bar{b}\gamma\gamma$ and $b\bar{b}b\bar{b}$ results sets observed (expected) limits on the resonant di-Higgs production cross-section ranging from 1.1 to 595 fb (1.2 to 392 fb) depending on the resonance mass. The largest excess is observed at 1.1 TeV, corresponding to a local significance of 3.2 σ. The excess is checked with the $b\bar{b}\tau^+\tau^-$ excess and is found to be compatible.

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] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[4] P.W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Physics Letters* 12.2 (1964), pp. 132–133. ISSN: 0031-9163. DOI: https://doi.org/10.1016/0031-9163(64)91136-9. URL: http://www.sciencedirect.com/science/article/pii/0031916364911369.

[5] 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.

[6] 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].

[7] Mark Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. ISBN: 978-1-107-03426-6.

[8] 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.

[9] Martin Breidenbach et al. "OBSERVED BEHAVIOR OF HIGHLY INELASTIC ELECTRON-PROTON SCATTERING". In: *Phys. Rev. Lett.* 23 (1969), pp. 935–939.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] Sheldon L. Glashow. “Partial-symmetries of weak interactions”. In: *Nucl. Phys.* 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.

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

[16] 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.

[17] Özer Özdal. “THE HIGGS BOSON AND RIGHT-HANDED NEUTRINOS IN SUPERSYMMETRIC MODELS”. PhD thesis. July 2016. DOI: 10.13140/RG.2.2.18314.52165.

[18] 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.

[19] *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.

[20] 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].

[21] 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.

[22] 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].

[23] 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].

[24] 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].

[25] 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].

[26] 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].

[27] 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.

[28] 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].

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

[30] 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.

[31] Richard Massey, Thomas Kitching, and Johan Richard. “The dark matter of gravitational lensing”. In: *Rep. Prog. Phys* 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.

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

[33] 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.

[34] K. et al. Eguchi. “First Results from KamLAND: Evidence for Reactor Antineutrino Disappearance”. In: *Phys. Rev. Lett.* 90.2 (Jan. 2003). ISSN: 1079-7114. DOI: 10.1103/physrevlett.90.021802. URL: http://dx.doi.org/10.1103/PhysRevLett.90.021802.

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

[36] 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.

[37] 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.

[38] H.E. Haber and G.L. Kane. “The search for supersymmetry: Probing physics beyond the standard model”. In: *Phys. Rep.* 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.

[39] G.C. Branco et al. “Theory and phenomenology of two-Higgs-doublet models”. In: *Phys. Rep.* 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.

[40] Howard E. Haber. *The Higgs data and the Decoupling Limit*. 2014. DOI: 10.48550/ARXIV.1401.0152. URL: https://arxiv.org/abs/1401.0152.

[41] Howard E. Haber and Yosef Nir. “Multi-scalar models with a high-energy scale”. In: *Nucl. Phys. 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.

[42] 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.

[43] B. Grzadkowski et al. “Dimension-six terms in the Standard Model Lagrangian”. In: *JHEP* 2010.10 (Oct. 2010). ISSN: 1029-8479. DOI: 10.1007/jhep10(2010)085. URL: http://dx.doi.org/10.1007/JHEP10(2010)085.

[44] 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.

[45] Gerhard Buchalla, Oscar Catà, and Claudius Krause. “Complete electroweak chiral Lagrangian with a light Higgs at NLO”. In: *Nucl. Phys. 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.

[46] Alexandra Carvalho et al. “Higgs pair production: choosing benchmarks with cluster analysis”. In: *JHEP* 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.

[47] M. Capozi and G. Heinrich. “Exploring anomalous couplings in Higgs boson pair production through shape analysis”. In: *JHEP* 2020.3 (Mar. 2020). ISSN: 1029-8479. DOI: 10.1007/jhep03(2020)091. URL: http://dx.doi.org/10.1007/JHEP03(2020)091.

[48] Alessandra Betti et al. *HEFT interpretations of Higgs boson pair searches in $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau\tau$ final states and of their combination in ATLAS*. Tech. rep. Geneva: CERN, Mar. 2022. URL: https://cds.cern.ch/record/2802963.

[49] *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*. Tech. rep. Geneva: CERN, Aug. 2019. URL: https://cds.cern.ch/record/2688062.

[50] Michele Faucci Giannelli. “FastCaloGAN: a fast simulation for the ATLAS calorimeter system using GANs”. In: (Oct. 2020). URL: https://cds.cern.ch/record/2742369.

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

[52] 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.

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

[54] Pierre Baldi et al. "Parameterized neural networks for high-energy physics". In: *Eur. Phys. J. C* 76.5 (2016), p. 235. DOI: 10.1140/epjc/s10052-016-4099-4. arXiv: 1601.07913 [hep-ex].

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

[56] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. 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.

[57] 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.

[58] Eilam Gross and Ofer Vitells. "Trial factors for the look elsewhere effect in high energy physics". In: *Eur. Phys. J. 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.

[59] Eilam Gross and Ofer Vitells. "Trial factors for the look elsewhere effect in high energy physics". In: *Eur. Phys. J. 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.

[60] Lyndon Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[61] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[62] S. Chatrchyan et al. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[63] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[64] A. Augusto Alves Jr. et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[65] Stephen Myers. “The LEP collider, from design to approval and commissioning”. In: (Oct. 1991). DOI: 10.5170/CERN-1991-008.

[66] L Rossi. “The LHC superconducting magnets”. In: 1 (2003), pp. 141–145.

[67] J. P. Blewett. “200-GeV Intersecting Storage Accelerators”. In: *eConf* C710920 (1971), p. 501.

[68] Ph Lebrun et al. “Report of the Task Force on the Incident of 19 September 2008 at the LHC”. In: *CERN LHC Project Rep* 1168 (2009).

[69] R Alemany-Fernandez et al. “Operation and Configuration of the LHC in Run 1”. In: (2013).

[70] ATLAS Collaboration. “A Particle Consistent with the Higgs Boson Observed with the ATLAS Detector at the Large Hadron Collider”. In: *Science* 338 (2012), p. 1576. DOI: 10.1126/science.1232005.

[71] F Bordry et al. “The first long shutdown (LS1) for the LHC”. In: (2013).

[72] Jorg Wenninger. “Operation and Configuration of the LHC in Run 2”. In: (Mar. 2019). URL: https://cds.cern.ch/record/2668326.

[73] Georges Aad et al. “ATLAS data quality operations and performance for 2015–2018 data-taking”. In: *JINST* 15.04 (2020), P04003.

[74] ATLAS Collaboration. *ATLAS Pixel Detector: Technical Design Report*. ATLAS-TDR-11; CERN-LHCC-98-013. 1998. URL: https://cds.cern.ch/record/381263.

[75] ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report, Volume 1*. ATLAS-TDR-4; CERN-LHCC-97-016. 1997. URL: https://cds.cern.ch/record/331063.

[76] ATLAS Collaboration. *ATLAS Liquid Argon Calorimeter: Technical Design Report*. ATLAS-TDR-2; CERN-LHCC-96-041. 1996. URL: https://cds.cern.ch/record/331061.

[77] ATLAS Collaboration. *ATLAS Tile Calorimeter: Technical Design Report*. ATLAS-TDR-3; CERN-LHCC-96-042. 1996. URL: https://cds.cern.ch/record/331062.

[78] ATLAS Collaboration. *ATLAS Muon Spectrometer: Technical Design Report*. ATLAS-TDR-10; CERN-LHCC-97-022. CERN, 1997. URL: https://cds.cern.ch/record/331068.

[79] ATLAS Collaboration. *ATLAS Magnet System: Magnet Project Technical Design Report, Volume 1*. ATLAS-TDR-6; CERN-LHCC-97-018. 1997. URL: https://cds.cern.ch/record/338080.

[80] ATLAS Collaboration. *ATLAS Insertable B-Layer: Technical Design Report*. ATLAS-TDR-19; CERN-LHCC-2010-013. 2010. URL: https://cds.cern.ch/record/1291633.

[81] J. Pequenao. *Computer generated image of the ATLAS inner detector*. 2008. URL: https://cds.cern.ch/record/1095926.

[82] ATLAS Collaboration. *ATLAS Calorimeter Performance: Technical Design Report*. ATLAS-TDR-1; CERN-LHCC-96-040. 1996. URL: https://cds.cern.ch/record/331059.

[83] ATLAS Collaboration. *ATLAS Level-1 Trigger: Technical Design Report*. ATLAS-TDR-12; CERN-LHCC-98-014. 1998. URL: https://cds.cern.ch/record/381429.

[84] ATLAS Collaboration. *ATLAS High-Level Trigger, Data Acquisition and Controls: Technical Design Report*. ATLAS-TDR-16; CERN-LHCC-2003-022. 2003. URL: https://cds.cern.ch/record/616089.

[85] Robert Blair et al. "The ATLAS High Level Trigger Region of Interest Builder". In: *JINST* 3 (2008), P04001. DOI: 10.1088/1748-0221/3/04/P04001. arXiv: 0711.3217 [physics.ins-det].

[86] Catrin Bernius. *The ATLAS Trigger Algorithms Upgrade and Performance in Run-2*. 2017. DOI: 10.48550/ARXIV.1709.09427. URL: https://arxiv.org/abs/1709.09427.

[87] ATLAS Collaboration. "Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in High Pile-Up LHC Environment". In: (2012). URL: https://cds.cern.ch/record/1435196.

[88] ATLAS Collaboration. "Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2". In: *Eur. Phys. J. C* 77 (2017), p. 673. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex].

[89] Azriel Rosenfeld and John L. Pfaltz. "Sequential Operations in Digital Picture Processing". In: *J. ACM* 13.4 (Oct. 1966), pp. 471–494. ISSN: 0004-5411. DOI: 10.1145/321356.321357. URL: https://doi.org/10.1145/321356.321357.

[90] R. Frühwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(87)90887-4. URL: https://www.sciencedirect.com/science/article/pii/0168900287908874.

[91] ATLAS Collaboration. "Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV". In: (2010). URL: https://cds.cern.ch/record/1281363.

[92] ATLAS Collaboration. "Performance of primary vertex reconstruction in proton–proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment". In: (2010). URL: https://cds.cern.ch/record/1281344.

[93] ATLAS Collaboration. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [hep-ex].

[94] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. Geneva: CERN, Apr. 2008. URL: https://cds.cern.ch/record/1099735.

[95] ATLAS Collaboration. "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung". In: (2012). URL: https://cds.cern.ch/record/1449796.

[96] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data". In: *JINST* 14 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[97] ATLAS Collaboration. "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 81 (2020), p. 578. DOI: 10.1140/epjc/s10052-021-09233-2. arXiv: 2012.00578 [hep-ex].

[98] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 76 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[99] ATLAS Collaboration. "Recommended isolation working points". In: (). URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/RecommendedIsolationWPs#Muon_working_points.

[100] Ryan Atkin. "Review of jet reconstruction algorithms". In: *J. Phys. Conf. Ser*. Vol. 645. 1. IOP Publishing. 2015, p. 012008.

[101] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77.7 (July 2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5004-5. URL: http://dx.doi.org/10.1140/epjc/s10052-017-5004-5.

[102] ATLAS Collaboration. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *Eur. Phys. J. C* 77 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[103] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 2008.04 (Apr. 2008), pp. 063–063. DOI: 10.1088/1126-6708/2008/04/063. URL: https://doi.org/10.1088/1126-6708/2008/04/063.

[104] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[105] ATLAS Collaboration. "Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 96 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[106] ATLAS Collaboration. "Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (2020). arXiv: 2007.02645 [hep-ex].

[107] ATLAS Collaboration. "Tagging and suppression of pileup jets with the ATLAS detector". In: (2014). URL: https://cds.cern.ch/record/1700870.

[108] ATLAS Collaboration. "Summary of the ATLAS experiment's sensitivity to supersymmetry after LHC Run 1 — interpreted in the phenomenological MSSM". In: *JHEP* 10 (2015), p. 134. DOI: 10.1007/JHEP10(2015)134. arXiv: 1508.06608 [hep-ex].

[109] ATLAS Collaboration. "Combination of searches for Higgs boson pairs in pp collisions at 13 TeV with the ATLAS experiment". In: (2018). URL: https://cds.cern.ch/record/2638212.

[110] Dean Carmi et al. "Interpreting LHC Higgs results from natural new physics perspective". In: *JHEP* 2012.7 (July 2012). ISSN: 1029-8479. DOI: 10.1007/jhep07(2012)136. URL: http://dx.doi.org/10.1007/JHEP07(2012)136.

[111] "Performance of b-jet identification in the ATLAS experiment". In: *JINST* 11.04 (Apr. 2016), P04008–P04008. ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/04/p04008. URL: http://dx.doi.org/10.1088/1748-0221/11/04/P04008.

[112] ATLAS Collaboration. "Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run". In: (2017). URL: https://cds.cern.ch/record/2273281.

[113] ATLAS Collaboration. "Secondary vertex finding for jet flavour identification with the ATLAS detector". In: (2017). URL: https://cds.cern.ch/record/2270366.

[114] ATLAS Collaboration. "Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC". In: (2018). URL: https://cds.cern.ch/record/2645405.

[115] ATLAS Collaboration. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].

[116] Luca Scodellaro. "b tagging in ATLAS and CMS". In: (2017). arXiv: 1709.01290 [hep-ex]. URL: https://arxiv.org/abs/1709.01290.

[117] "Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run". In: ATL-PHYS-PUB-2016-012 (June 2016). URL: https://cds.cern.ch/record/2160731.

[118] ATLAS Collaboration. "Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s} = 8$ TeV". In: *Eur. Phys. J. C* 75 (2015), p. 303. DOI: 10.1140/epjc/s10052-015-3500-z. arXiv: 1412.7086 [hep-ex].

[119] ATLAS Collaboration. "Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC". In: (2015). URL: https://cds.cern.ch/record/2064383.

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

[121] M. *et. al.* Aaboud. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 78.11 (Nov. 2018). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-018-6288-9. URL: http://dx.doi.org/10.1140/epjc/s10052-018-6288-9.

[122] ATLAS Collaboration. "ATLAS Overlap Removal Tool: AssociationUtils". In: (). URL: https://gitlab.cern.ch/atlas/athena/tree/21.2/PhysicsAnalysis/AnalysisCommon/AssociationUtils/.

[123] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[124] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C* 70 (2010), p. 823. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[125] ATLAS Collaboration. "Performance of electron and photon triggers in ATLAS during LHC Run 2". In: *Eur. Phys. J. C* 80 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex].

[126] ATLAS Collaboration. "Performance of the ATLAS muon triggers in Run 2". In: *JINST* 15.09 (2020), P09015. DOI: 10.1088/1748-0221/15/09/p09015. arXiv: 2004.13447 [hep-ex].

[127] ATLAS Collaboration. "ATLAS data quality operations and performance for 2015–2018 data-taking". In: *JINST* 15 (2020), P04003. DOI: 10.1088/1748-0221/15/04/P04003. arXiv: 1911.04632 [physics.ins-det].

[128] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC". In: (2019). URL: https://cds.cern.ch/record/2677054.

[129] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017.

[130] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A* 462 (2001), p. 152. DOI: 10.1016/S0168-9002(01)00089-4.

[131] Enrico Bothmann et al. "Event generation with Sherpa 2.2". In: *SciPost Phys.* 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph].

[132] T. Sjöstrand, S. Mrenna, and P. Skands. "A brief introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[133] ATLAS Collaboration. "The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model". In: (2016). URL: https://cds.cern.ch/record/2206965.

[134] Richard D. Ball et al. "Parton distributions with LHC data". In: *Nucl. Phys. B* 867 (2013), p. 244. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].

[135] *Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-030. Geneva: CERN, July 2021. URL: https://cds.cern.ch/record/2777236.

[136] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[137] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[138] ATLAS Collaboration. "ATLAS Pythia 8 tunes to 7 TeV data". In: (2014). URL: https://cds.cern.ch/record/1966419.

[139] ATLAS Collaboration. *Summary of ATLAS Pythia 8 tunes*. ATL-PHYS-PUB-2012-003. 2012. URL: https://cds.cern.ch/record/1474107.

[140] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

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

[142] M. Bähr et al. "Herwig++ physics and manual". In: *Eur. Phys. J. C* 58 (2008), p. 639. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[143] Johannes Bellm et al. “Herwig 7.0/Herwig++ 3.0 release note”. In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[144] Stefan Gieseke, Christian Röhr, and Andrzej Siodmok. “Colour reconnections in Herwig++”. In: *Eur. Phys. J. C* 72 (2012), p. 2225. DOI: 10.1140/epjc/s10052-012-2225-5. arXiv: 1206.0041 [hep-ph].

[145] “Validation of signal Monte Carlo event generation in searches for Higgs boson pairs with the ATLAS detector”. In: (2019).

[146] G. Buchalla et al. “Higgs boson pair production in non-linear Effective Field Theory with full mt-dependence at NLO QCD”. In: *JHEP* 2018.9 (Sept. 2018). ISSN: 1029-8479. DOI: 10.1007/jhep09(2018)057. URL: http://dx.doi.org/10.1007/JHEP09(2018)057.

[147] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. “A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”. In: *JHEP* 09 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph].

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

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

[150] ATLAS Collaboration. “Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO”. In: (2017). URL: https://cds.cern.ch/record/2261938.

[151] Michal Czakon and Alexander Mitov. “Top++: A program for the calculation of the top-pair cross-section at hadron colliders”. In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[152] N. Kidonakis. “Next-to-next-to-leading logarithm resummation for s-channel single top quark production”. In: *Phys. Rev. D* 81 (2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028. arXiv: 1001.5034 [hep-ph].

[153] N. Kidonakis. “Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production”. In: *Phys. Rev. D* 83 (2011), p. 091503. DOI: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792 [hep-ph].

[154] Nikolaos Kidonakis. “Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-”. In: *Phys. Rev. D* 82 (2010), p. 054018. DOI: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451 [hep-ph].

[155] Stefano Frixione et al. "Single-top hadroproduction in association with a W boson". In: *JHEP* 07 (2008), p. 029. DOI: 10.1088/1126-6708/2008/07/029. arXiv: 0805.3067 [hep-ph].

[156] Tanju Gleisberg and Stefan Höche. "Comix, a new matrix element generator". In: *JHEP* 12 (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].

[157] Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

[158] Charalampos Anastasiou et al. "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order". In: *Phys. Rev. D* 69 (2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266.

[159] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[160] Keith Hamilton et al. "NNLOPS simulation of Higgs boson production". In: *JHEP* 10 (2013), p. 222. DOI: 10.1007/JHEP10(2013)222. arXiv: 1309.0017 [hep-ph].

[161] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. "Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator". In: *JHEP* 05 (2015), p. 140. DOI: 10.1007/JHEP05(2015)140. arXiv: 1501.04637 [hep-ph].

[162] 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].

[163] Charalampos Anastasiou et al. "High precision determination of the gluon fusion Higgs boson cross-section at the LHC". In: *JHEP* 05 (2016), p. 058. DOI: 10.1007/JHEP05(2016)058. arXiv: 1602.00695 [hep-ph].

[164] Charalampos Anastasiou et al. "Higgs Boson Gluon-Fusion Production in QCD at Three Loops". In: *Phys. Rev. Lett.* 114 (2015), p. 212001. DOI: 10.1103/PhysRevLett.114.212001. arXiv: 1503.06056 [hep-ph].

[165] Falko Dulat, Achilleas Lazopoulos, and Bernhard Mistlberger. "iHixs 2 – Inclusive Higgs cross sections". In: *Comput. Phys. Commun.* 233 (2018), pp. 243–260. DOI: 10.1016/j.cpc.2018.06.025. arXiv: 1802.00827 [hep-ph].

[166] Stefano Actis et al. "NLO electroweak corrections to Higgs boson production at hadron colliders". In: *Phys. Lett. B* 670 (2008), pp. 12–17. DOI: 10.1016/j.physletb.2008.10.018. arXiv: 0809.1301 [hep-ph].

[167] J. Pumplin et al. "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis". In: *JHEP* 07 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195 [hep-ph].

[168] ATLAS Collaboration. "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *JHEP* 09 (2014), p. 145. DOI: 10.1007/JHEP09(2014)145. arXiv: 1406.3660 [hep-ex].

[169] M. Ciccolini, Ansgar Denner, and S. Dittmaier. "Strong and Electroweak Corrections to the Production of Higgs + 2 Jets via Weak Interactions at the Large Hadron Collider". In: *Phys. Rev. Lett.* 99 (2007), p. 161803. DOI: 10.1103/PhysRevLett.99.161803. arXiv: 0707.0381 [hep-ph].

[170] Mariano Ciccolini, Ansgar Denner, and Stefan Dittmaier. "Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC". In: *Phys. Rev. D* 77 (2008), p. 013002. DOI: 10.1103/PhysRevD.77.013002. arXiv: 0710.4749 [hep-ph].

[171] Paolo Bolzoni et al. "Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD". In: *Phys. Rev. Lett.* 105 (2010), p. 011801. DOI: 10.1103/PhysRevLett.105.011801. arXiv: 1003.4451 [hep-ph].

[172] M.L. Ciccolini, S. Dittmaier, and M. Krämer. "Electroweak radiative corrections to associated WH and ZH production at hadron colliders". In: *Phys. Rev. D* 68 (2003), p. 073003. DOI: 10.1103/PhysRevD.68.073003. arXiv: hep-ph/0306234 [hep-ph].

[173] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. "NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders". In: *Phys. Lett. B* 579 (2004), pp. 149–156. DOI: 10.1016/j.physletb.2003.10.112. arXiv: hep-ph/0307206 [hep-ph].

[174] Giancarlo Ferrera, Massimiliano Grazzini, and Francesco Tramontano. "Associated Higgs-W-Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at NNLO". In: *Phys. Rev. Lett.* 107 (2011), p. 152003. DOI: 10.1103/PhysRevLett.107.152003. arXiv: 1107.1164 [hep-ph].

[175] Oliver Brein et al. "Top-quark mediated effects in hadronic Higgs-Strahlung". In: *Eur. Phys. J. C* 72 (2012), p. 1868. DOI: 10.1140/epjc/s10052-012-1868-6. arXiv: 1111.0761 [hep-ph].

[176] Giancarlo Ferrera, Massimiliano Grazzini, and Francesco Tramontano. "Higher-order QCD effects for associated WH production and decay at the LHC". In: *JHEP* 04 (2014), p. 039. DOI: 10.1007/JHEP04(2014)039. arXiv: 1312.1669 [hep-ph].

[177] Giancarlo Ferrera, Massimiliano Grazzini, and Francesco Tramontano. "Associated ZH production at hadron colliders: The fully differential NNLO QCD calculation". In: *Phys. Lett. B* 740 (2015), pp. 51–55. DOI: 10.1016/j.physletb.2014.11.040. arXiv: 1407.4747 [hep-ph].

[178] John M. Campbell, R. Keith Ellis, and Ciaran Williams. "Associated production of a Higgs boson at NNLO". In: *JHEP* 06 (2016), p. 179. DOI: 10.1007/JHEP06(2016)179. arXiv: 1601.00658 [hep-ph].

[179] Lukas Altenkamp et al. "Gluon-induced Higgs-strahlung at next-to-leading order QCD". In: *JHEP* 02 (2013), p. 078. DOI: 10.1007/JHEP02(2013)078. arXiv: 1211.5015 [hep-ph].

[180] B. Hespel, F. Maltoni, and E. Vryonidou. "Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM". In: *JHEP* 06 (2015), p. 065. DOI: 10.1007/JHEP06(2015)065. arXiv: 1503.01656 [hep-ph].

[181] Robert V. Harlander et al. "Soft gluon resummation for gluon-induced Higgs Strahlung". In: *JHEP* 11 (2014), p. 082. DOI: 10.1007/JHEP11(2014)082. arXiv: 1410.0217 [hep-ph].

[182] Robert V. Harlander, Stefan Liebler, and Tom Zirke. "Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-doublet model". In: *JHEP* 02 (2014), p. 023. DOI: 10.1007/JHEP02(2014)023. arXiv: 1307.8122 [hep-ph].

[183] Oliver Brein, Robert V. Harlander, and Tom J. E. Zirke. "vh@nnlo - Higgs Strahlung at hadron colliders". In: *Comput. Phys. Commun.* 184 (2013), pp. 998–1003. DOI: 10.1016/j.cpc.2012.11.002. arXiv: 1210.5347 [hep-ph].

[184] ATLAS Collaboration. "Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s}$ = 13 TeV collected with the ATLAS detector". In: (Sept. 2021). arXiv: 2109.10627 [hep-ex].

[185] ATLAS Collaboration. "Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s}$ = 13 TeV". In: (2018). URL: https://cds.cern.ch/record/2314418.

[186] David Krohn, Jesse Thaler, and Lian-Tao Wang. "Jets with variable R". In: 2009.06 (June 2009), pp. 059–059. DOI: 10.1088/1126-6708/2009/06/059. URL: https://doi.org/10.1088/1126-6708/2009/06/059.

[187] G. Abbiendi et al. "A measurement of the rate of charm production in W decays". In: *Physics Letters B* 490.1-2 (Sept. 2000), pp. 71–86. ISSN: 0370-2693. DOI: 10.1016/s0370-2693(00)00971-0. URL: http://dx.doi.org/10.1016/S0370-2693(00)00971-0.

[188] K.A. Olive. "Review of Particle Physics". In: *Chinese Physics C* 38.9 (Aug. 2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001. URL: https://doi.org/10.1088/1674-1137/38/9/090001.

[189] Johannes Erdmann et al. "A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 748 (2014), pp. 18–25. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2014.02.029.

[190] ATLAS Collaboration. "Studies on top-quark Monte Carlo modelling for Top2016". In: (2016). URL: https://cds.cern.ch/record/2216168.

[191] ATLAS Collaboration. "Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: *JHEP* 04 (2019), p. 046. DOI: 10.1007/JHEP04(2019)046. arXiv: 1811.12113 [hep-ex].

[192] *Combination of searches for non-resonant and resonant Higgs boson pair production in the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}b\bar{b}$ decay channels using pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-052. Geneva: CERN, Oct. 2021. URL: https://cds.cern.ch/record/2786865.

[193] ATLAS. "LHC Trigger Recommendations For Analysis Groups Full Run2 ". 2020. URL: https://twiki.cern.ch/twiki/bin/view/Atlas/TriggerRecommendationsForAnalysisGroup

[194] A. Elagin et al. "A new mass reconstruction technique for resonances decaying to $\tau\tau$". In: *Nucl. Instrum. Meth. A* 654 (2011), pp. 481–489. DOI: 10.1016/j.nima.2011.07.009. arXiv: 1012.4686 [hep-ex].

[195] ATLAS. "LHC Fake Taus Task Force webpage". 2020. URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/FakeTausTaskForce.

[196] ATLAS Collaboration. "Search for non-resonant Higgs boson pair production in the $bb\ell\nu\ell\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s}$ = 13 TeV". In: *Phys. Lett. B* 801 (2020), p. 135145. DOI: 10.1016/j.physletb.2019.135145. arXiv: 1908.06765 [hep-ex].

[197] François Chollet et al. *Keras*. https://keras.io. 2015.

[198] *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/.

[199] Jonas Mockus. *Bayesian approach to global optimization: theory and applications*. Vol. 37. Springer Science & Business Media, 2012.

[200] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 76 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[201] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017.

[202] ATLAS Collaboration. "Measurement of the Inelastic Proton–Proton Cross Section at $\sqrt{s}$ = 13 TeV with the ATLAS Detector at the LHC". In: *Phys. Rev. Lett.* 117 (2016), p. 182002. DOI: 10.1103/PhysRevLett.117.182002. arXiv: 1606.02625 [hep-ex].

[203] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 78 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[204] ATLAS Collaboration. "Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector". In: (2020). URL: https://cds.cern.ch/record/2714885.

[205] Konie Al Khoury et al. *Data/Monte Carlo Modelling Studies for the Standard Model $VH, H \to b\bar{b}$ Analysis*. Tech. rep. ATL-COM-PHYS-2019-1205. Geneva: CERN, Sept. 2019. URL: https://cds.cern.ch/record/2690042.

[206] Gudrun Heinrich et al. "A non-linear EFT description of $gg \to HH$ at NLO interfaced to POWHEG". In: *JHEP* 10 (2020), p. 021. DOI: 10.1007/JHEP10(2020)021. arXiv: 2006.16877 [hep-ph].

[207] CMS Collaboration. "Combination of Searches for Higgs Boson Pair Production in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. Lett.* 122 (12 Mar. 2019), p. 121803. DOI: 10.1103/PhysRevLett.122.121803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.121803.

[208] *Search for nonresonant Higgs boson pair production in the 4 leptons plus 2 b jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2725691.

[209] *Search for Higgs boson pair production in the four b quark final state*. Tech. rep. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2771912.

[210] A. M. Sirunyan et al. "Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 2021.3 (Mar. 2021). ISSN: 1029-8479. DOI: 10.1007/jhep03(2021)257. URL: http://dx.doi.org/10.1007/JHEP03(2021)257.

Appendix A

Supplementary material for c-jet calibration

A.1 Additional plots for kinematic variables

A.1.1 Standard selection

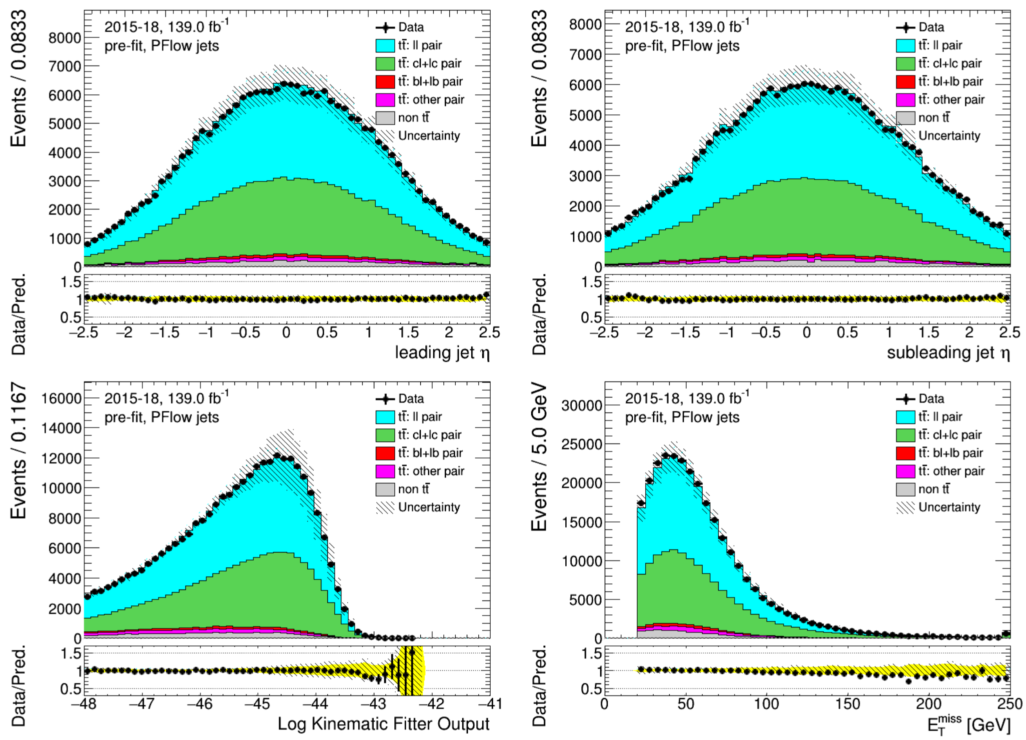


Figure A.1: PFlow jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the standard selection, before fitting or tagging with full uncertainties.

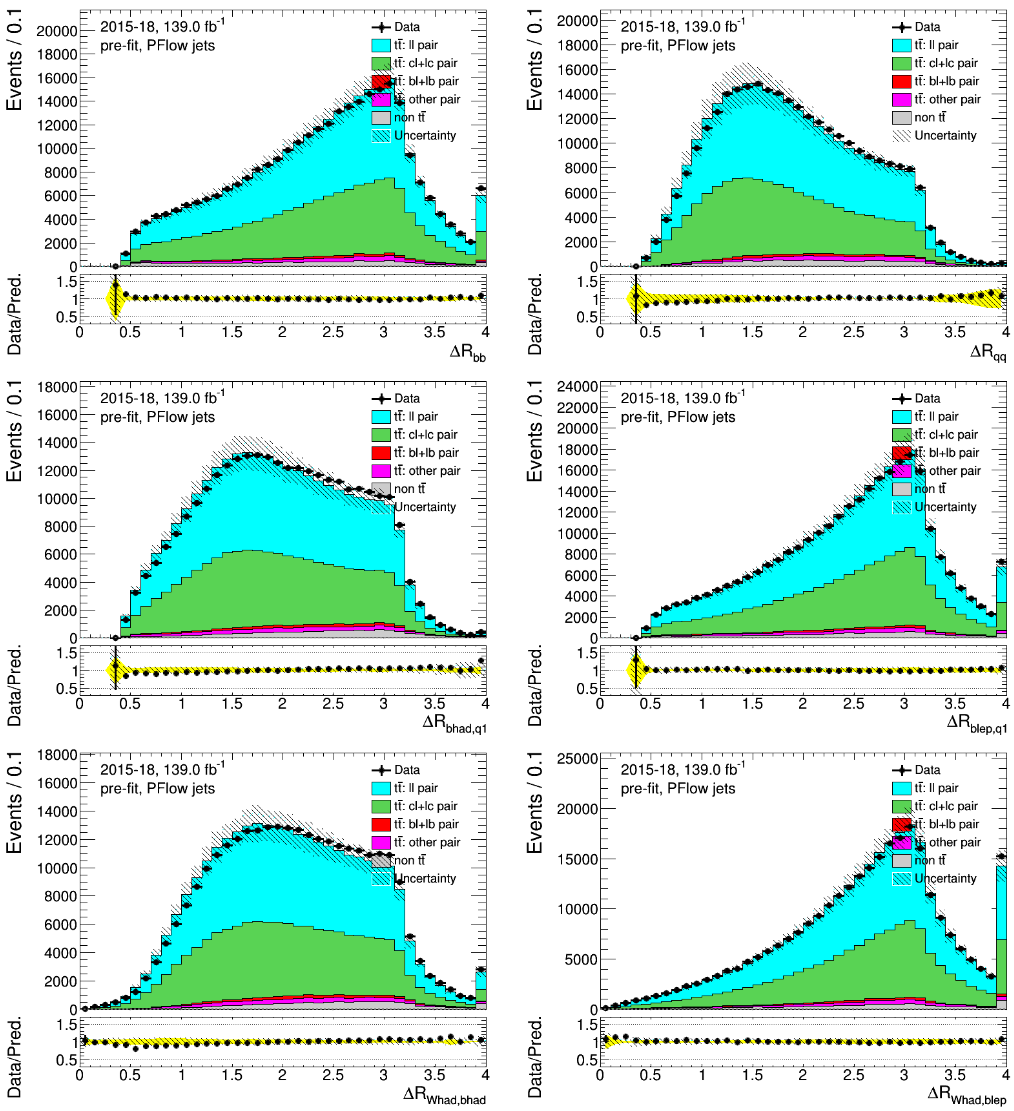


Figure A.2: PFlow jets: distributions of angle related variables of the combination of the standard selection, before fitting or tagging with full uncertainties.

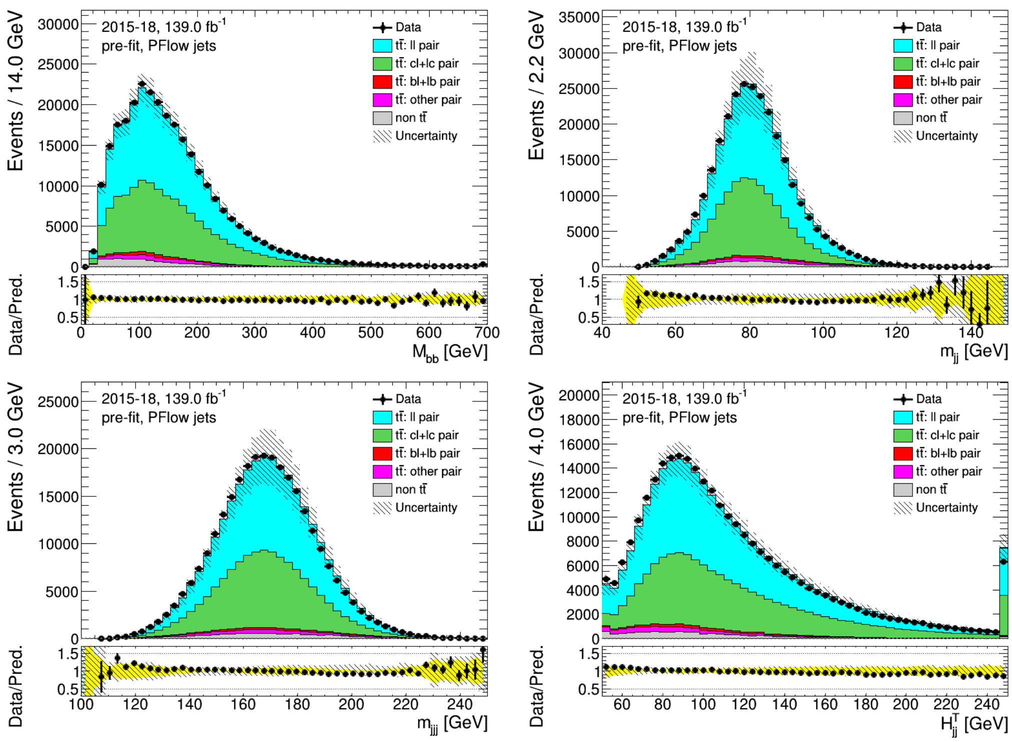


Figure A.3: PFlow jets: distributions of mass related variables of the standard selection, before fitting or tagging with stat-only uncertainties.

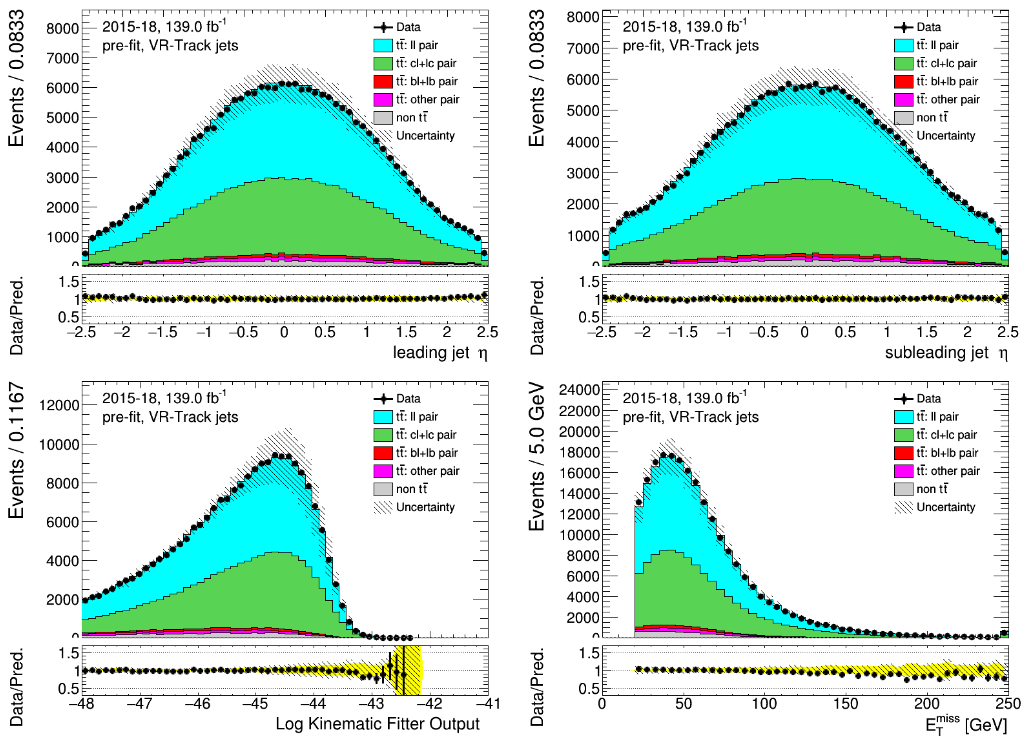


Figure A.4: VR-Track jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the standard selection, before fitting or tagging with full uncertainties.

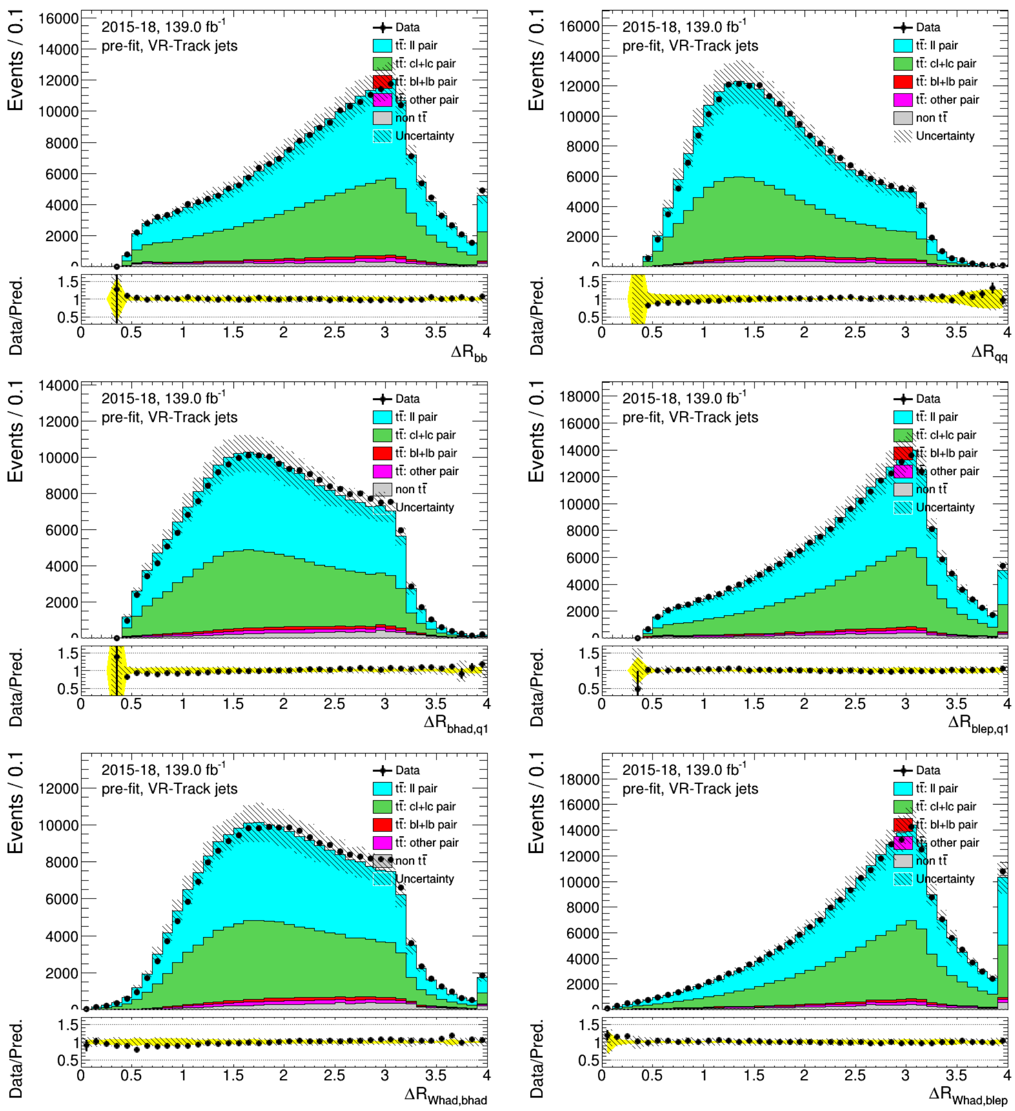


Figure A.5: VR-Track jets: distributions of angle related variables of the combination of the standard selection, before fitting or tagging with full uncertainties.

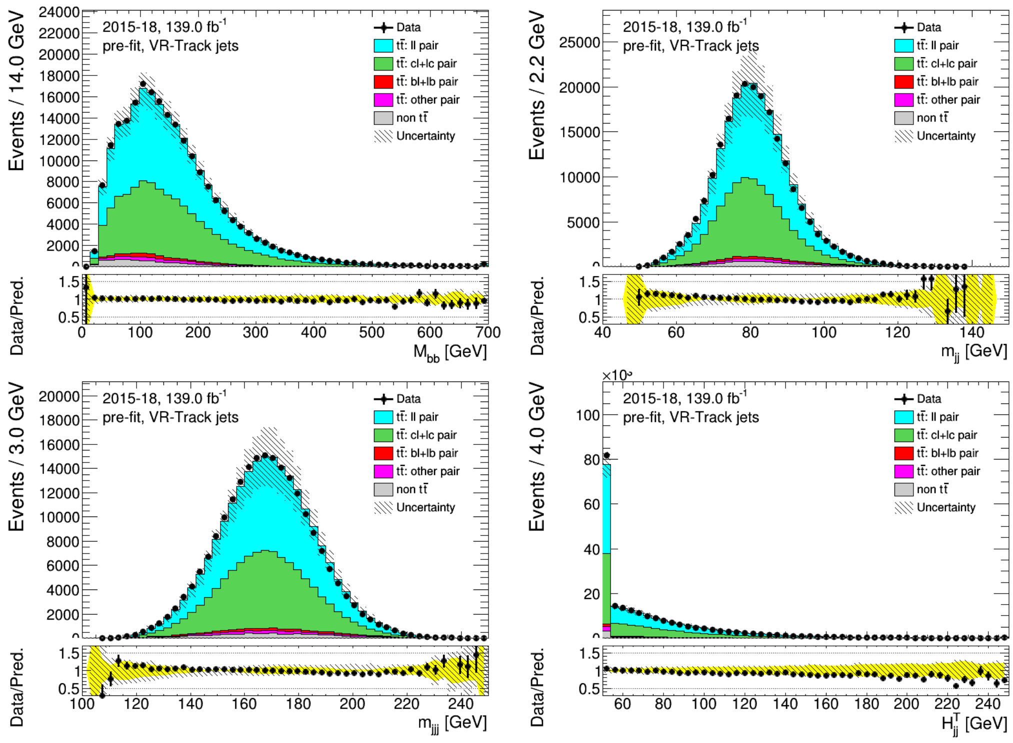


Figure A.6: VR-Track jets: distributions of mass related variables of the standard selection, before fitting or tagging with stat-only uncertainties.

A.1.2 Low-p_T selection

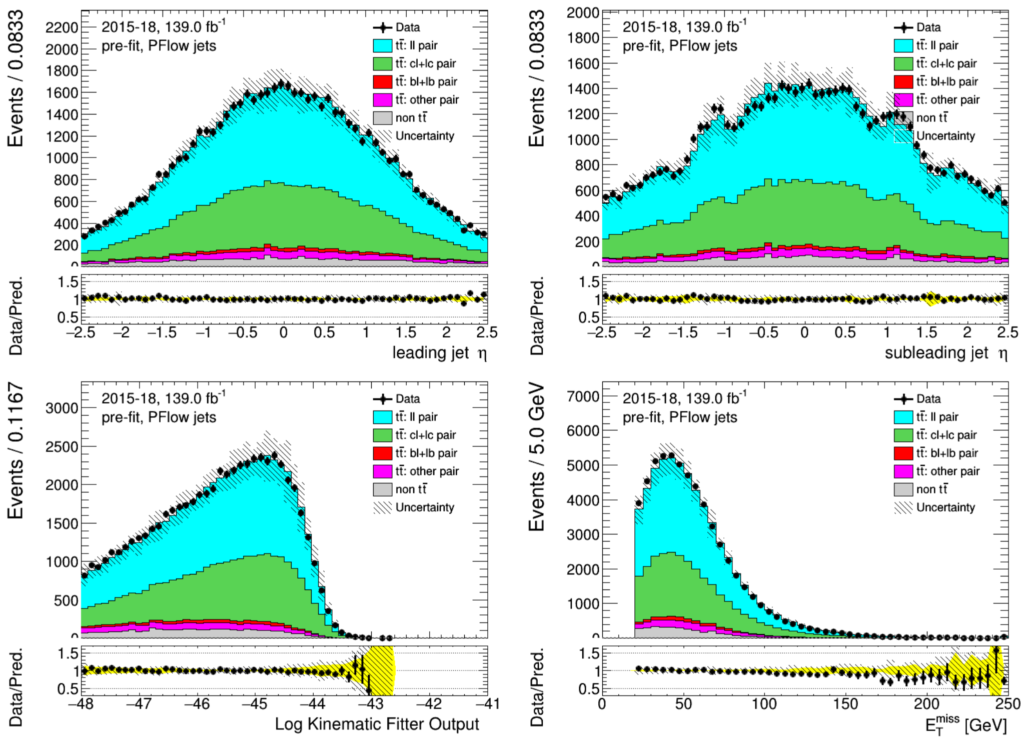


Figure A.7: PFlow jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the low-p_{T} selection, before fitting or tagging with full uncertainties.

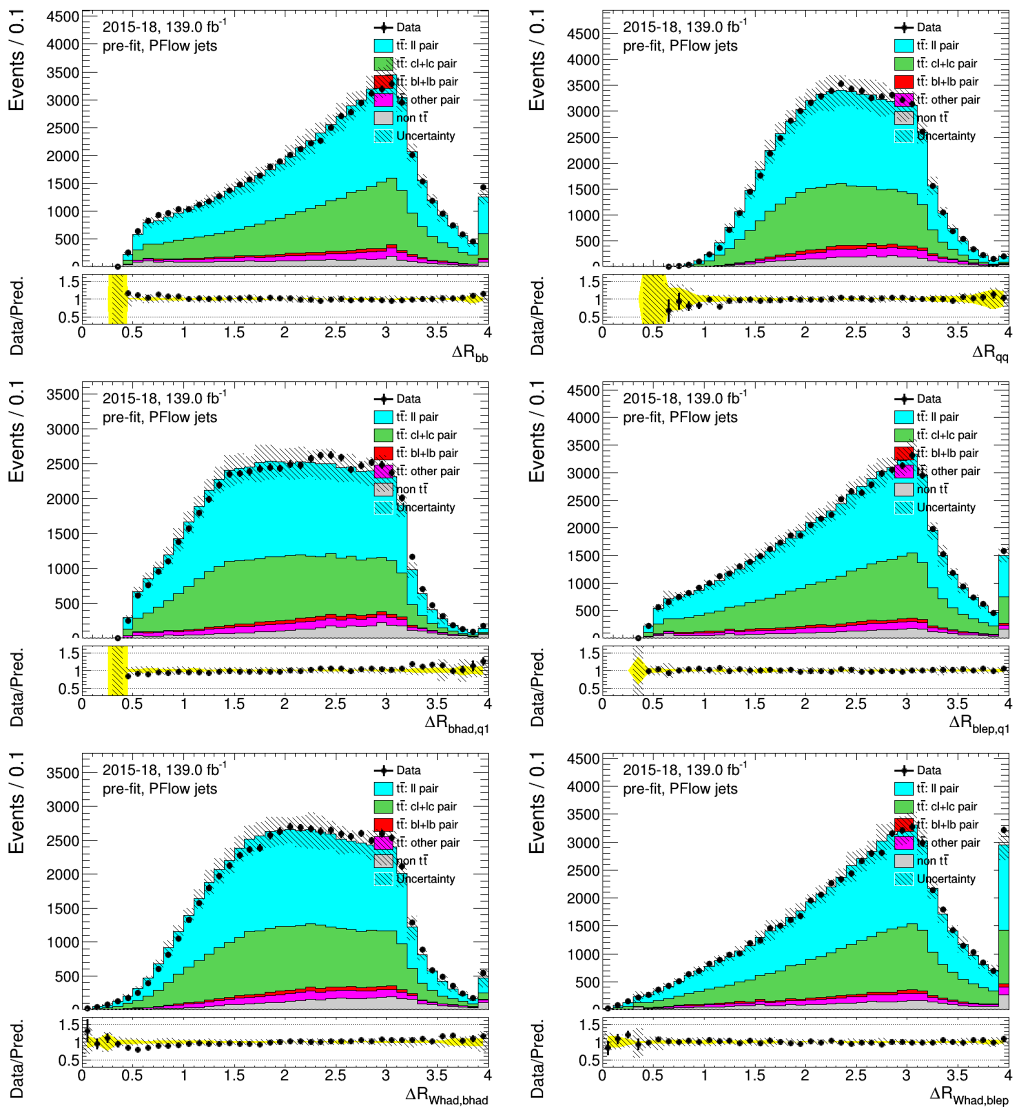


Figure A.8: PFlow jets: distributions of angle related variables of the combination of the low-p_{T} selection, before fitting or tagging with full uncertainties.

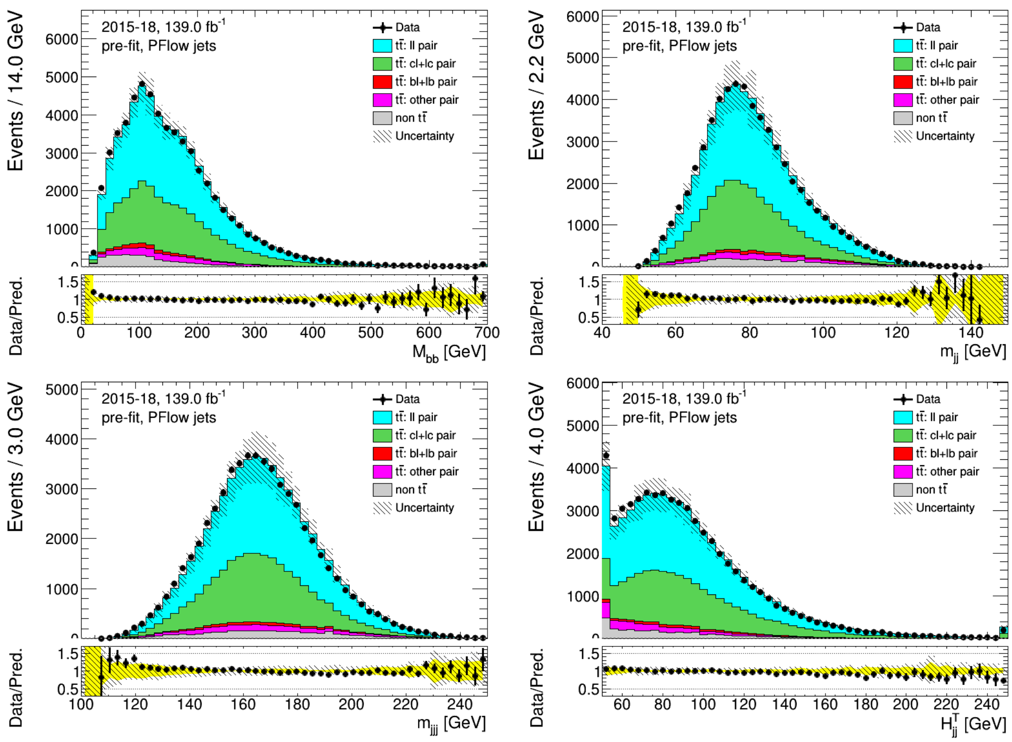


Figure A.9: PFlow jets: distributions of mass related variables of the low-p_T selection, before fitting or tagging with stat-only uncertainties.

A.2 High-p_T selection

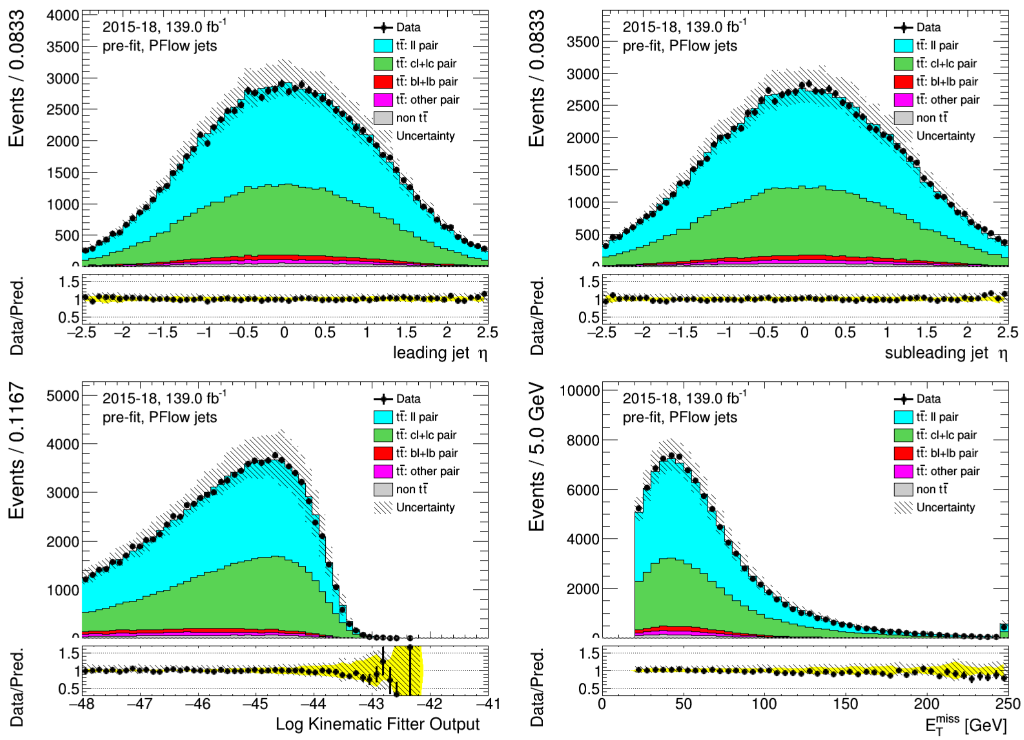


Figure A.10: PFlow jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the high-p_T selection, before fitting or tagging with full uncertainties.

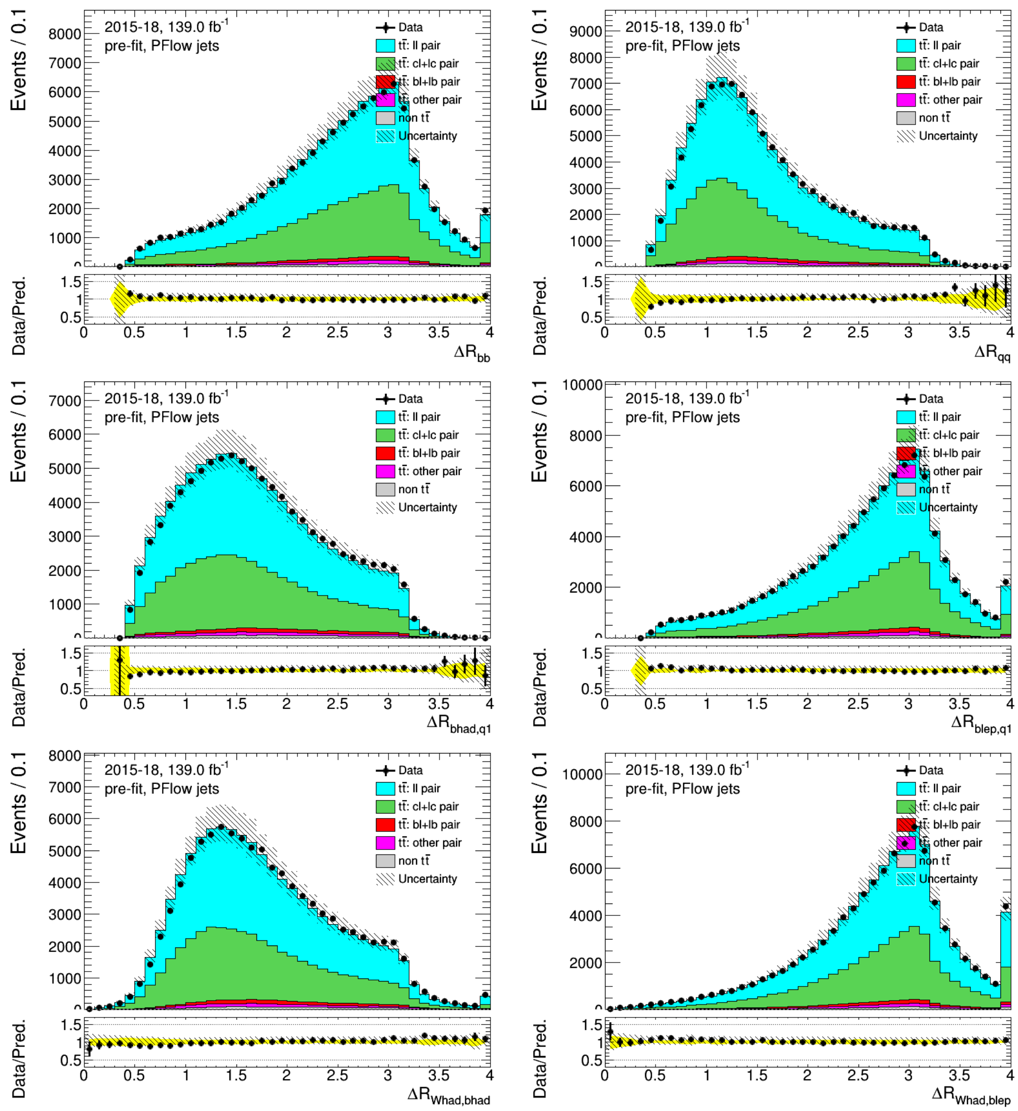


Figure A.11: PFlow jets: distributions of angle related variables of the combination of the high-p_T selection, before fitting or tagging with full uncertainties.

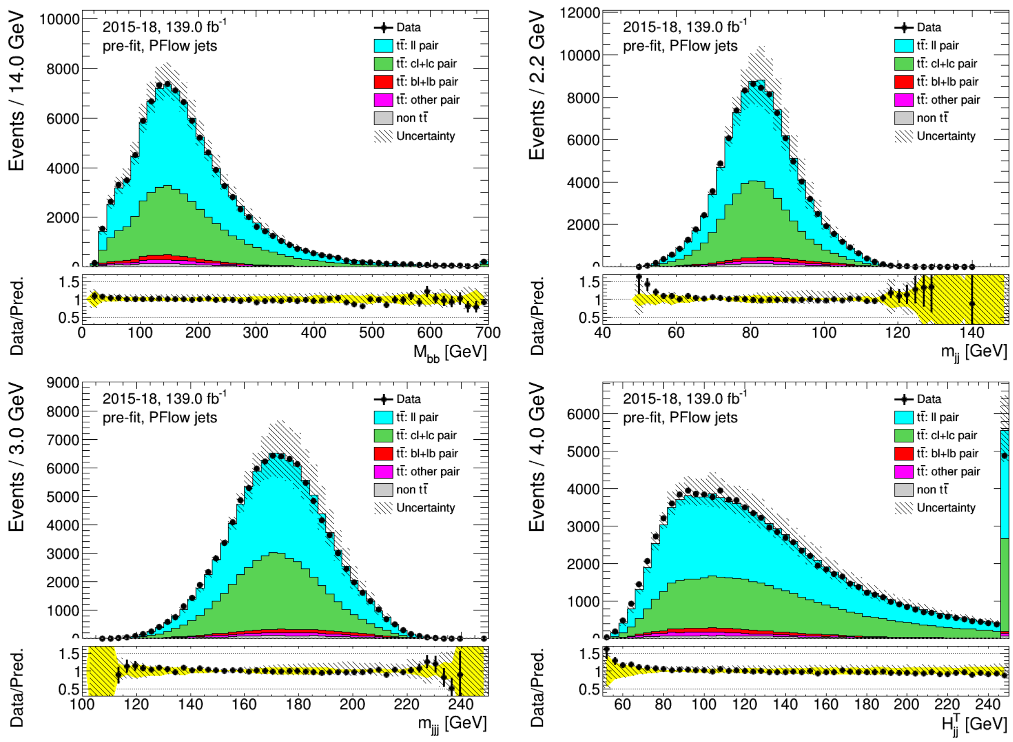


Figure A.12: PFlow jets: distributions of mass related variables of the high-p_{T} selection, before fitting or tagging with stat-only uncertainties.

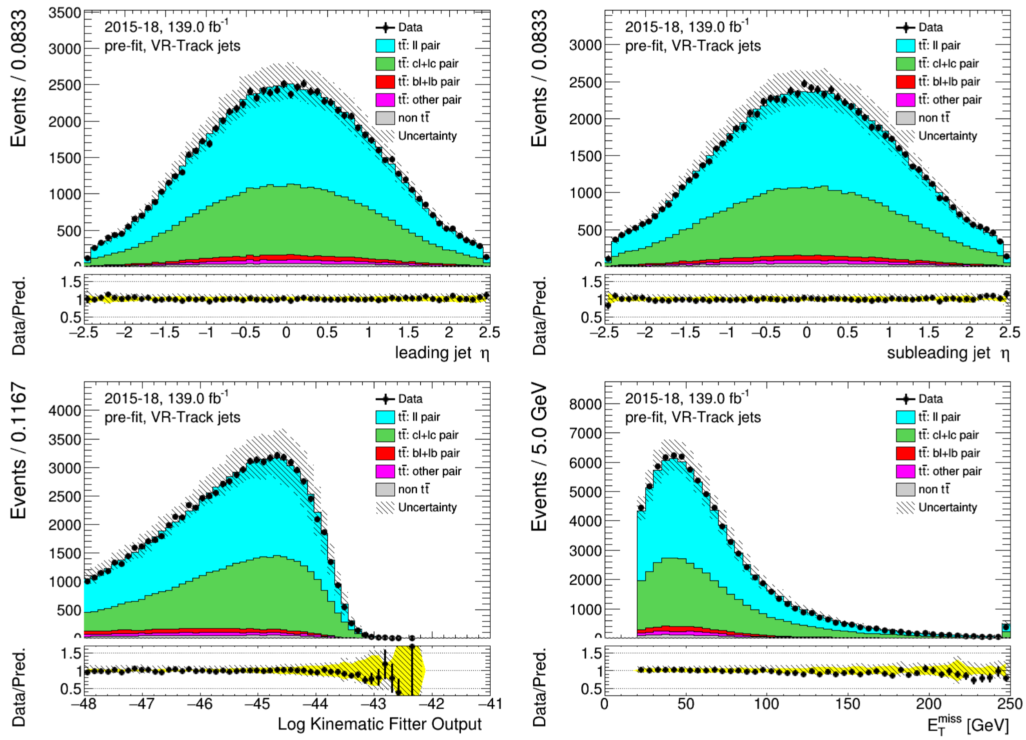


Figure A.13: VR-Track jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the high-p_T selection, before fitting or tagging with full uncertainties.

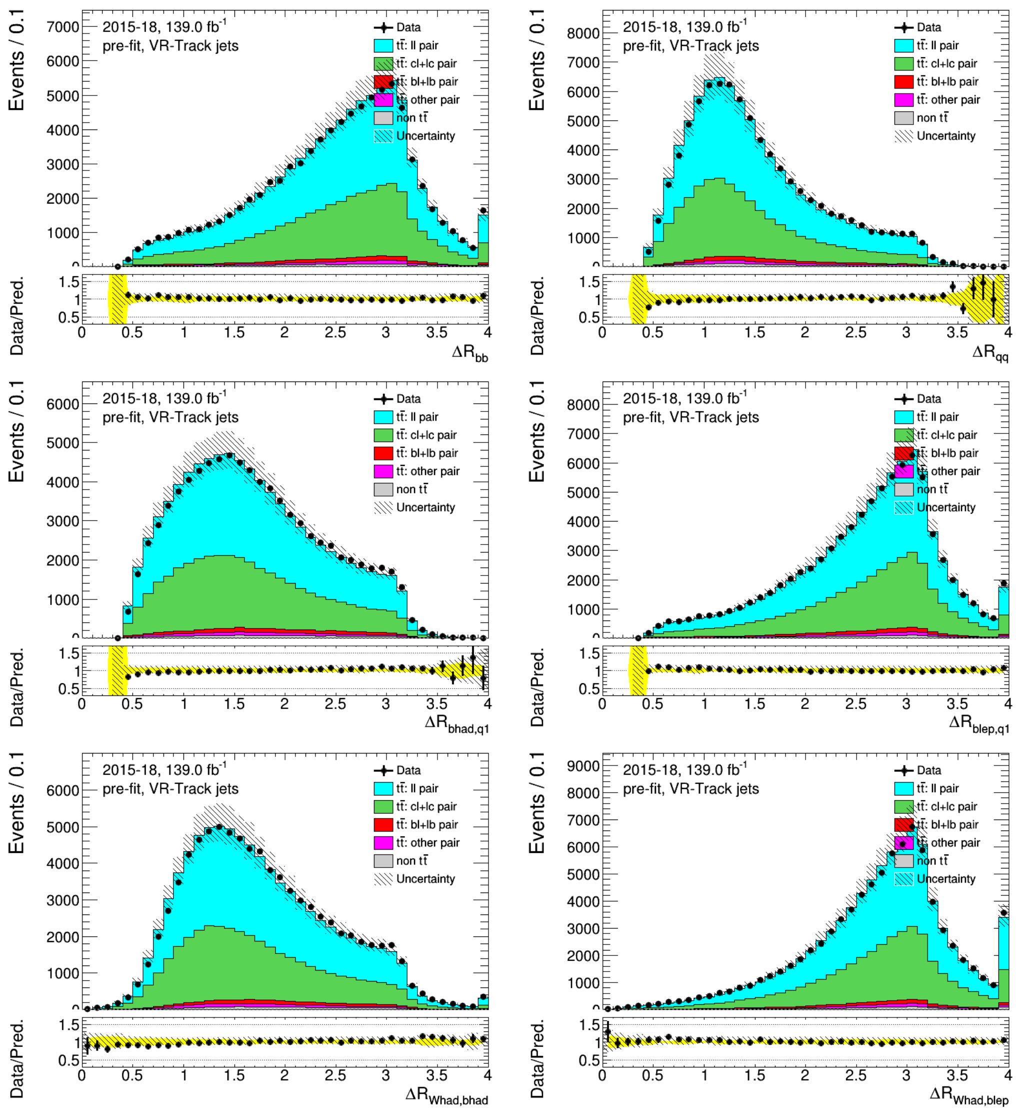


Figure A.14: VR-Track jets: distributions of angle related variables of the combination of the high-p_{T} selection, before fitting or tagging with full uncertainties.

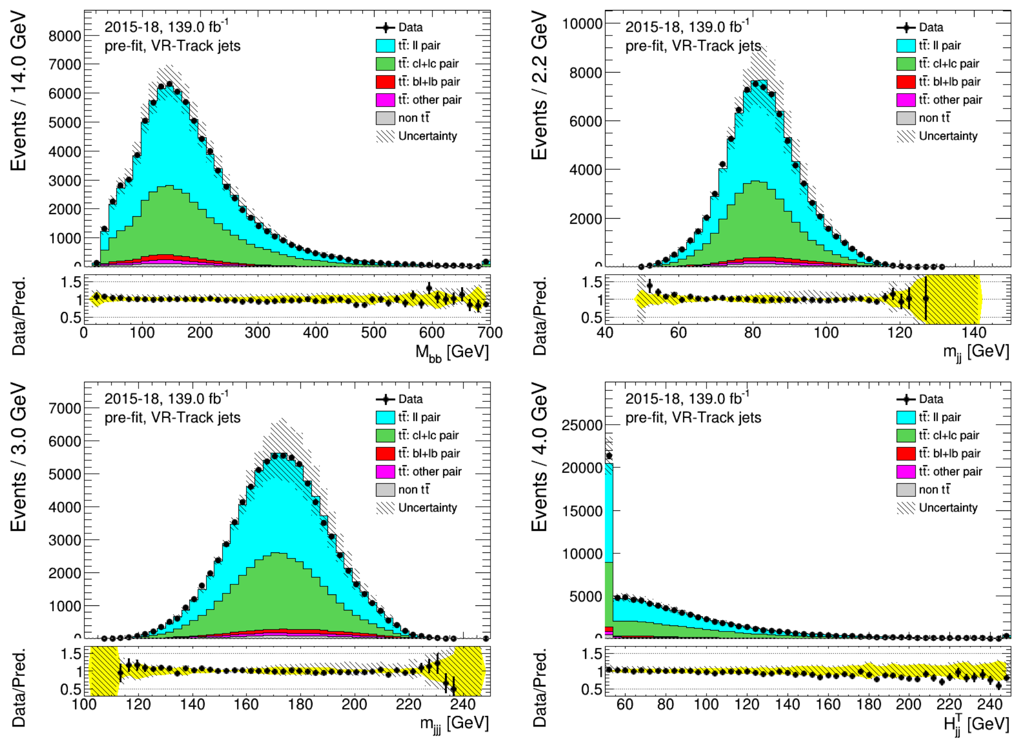


Figure A.15: VR-Track jets: distributions of mass related variables of the high-p_{T} selection, before fitting or tagging with stat-only uncertainties.

A.3 Combined selection

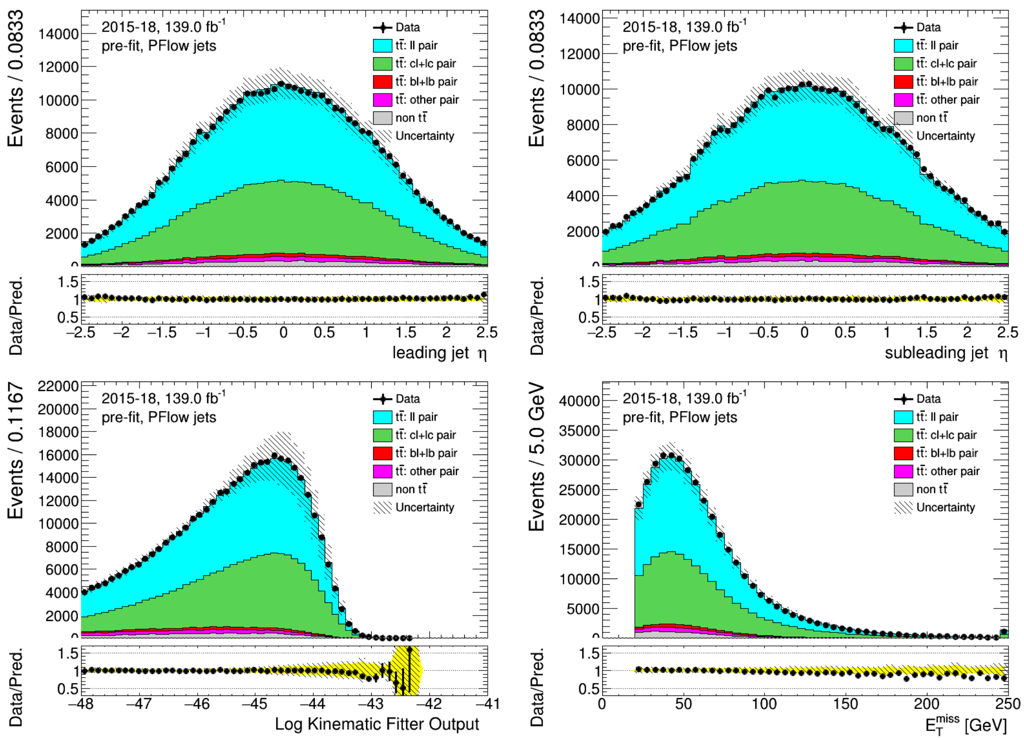


Figure A.16: PFlow jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the combined selection, before fitting or tagging with full uncertainties.

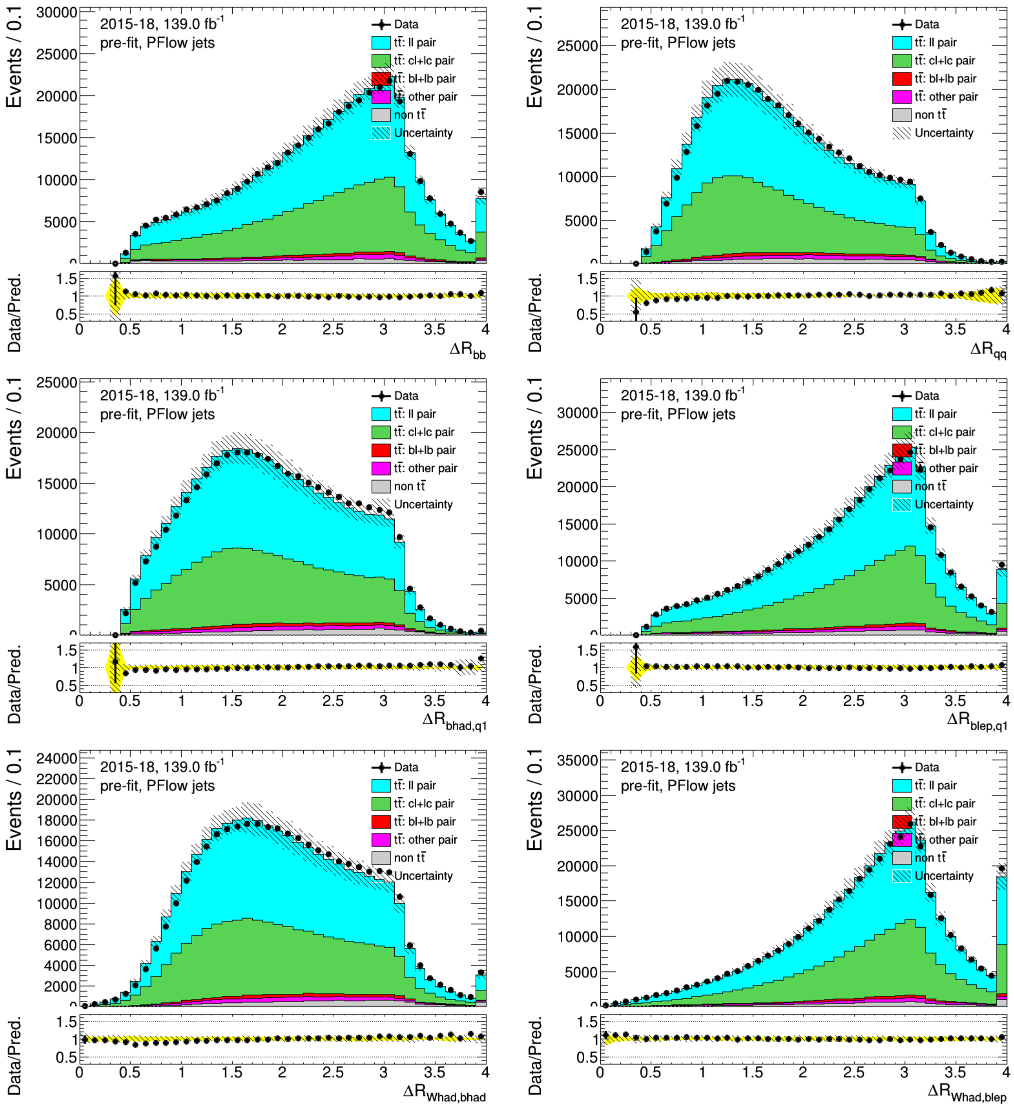


Figure A.17: PFlow jets: distributions of angle related variables of the combination of the combined selection, before fitting or tagging with full uncertainties.

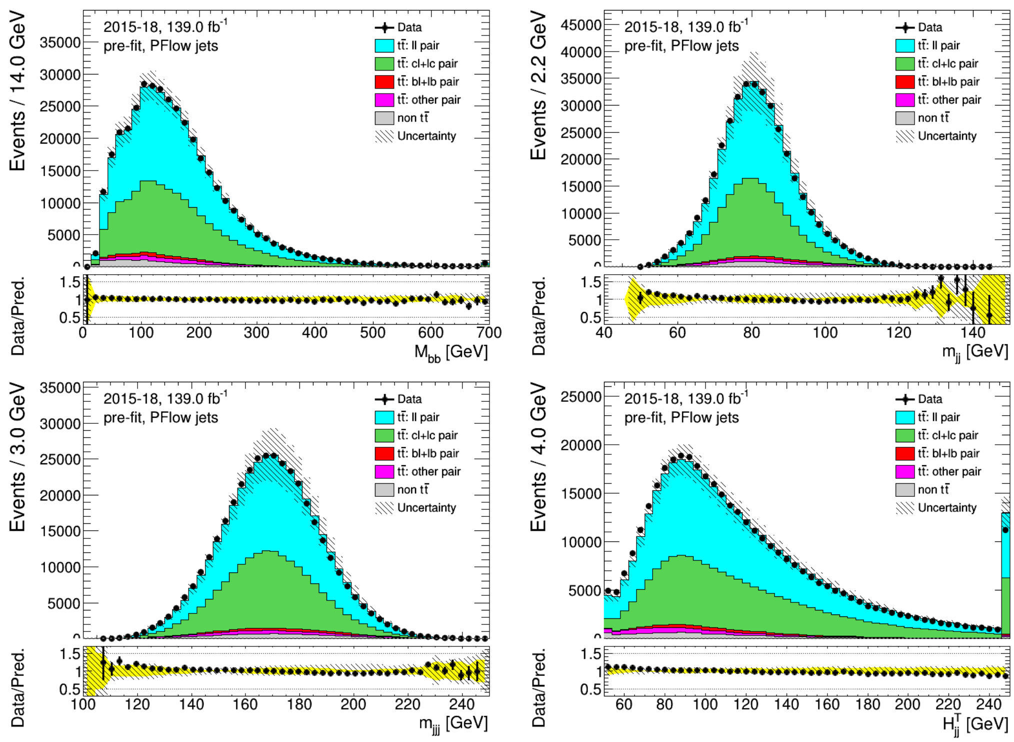


Figure A.18: PFlow jets: distributions of mass related variables of the combined selection, before fitting or tagging with stat-only uncertainties.

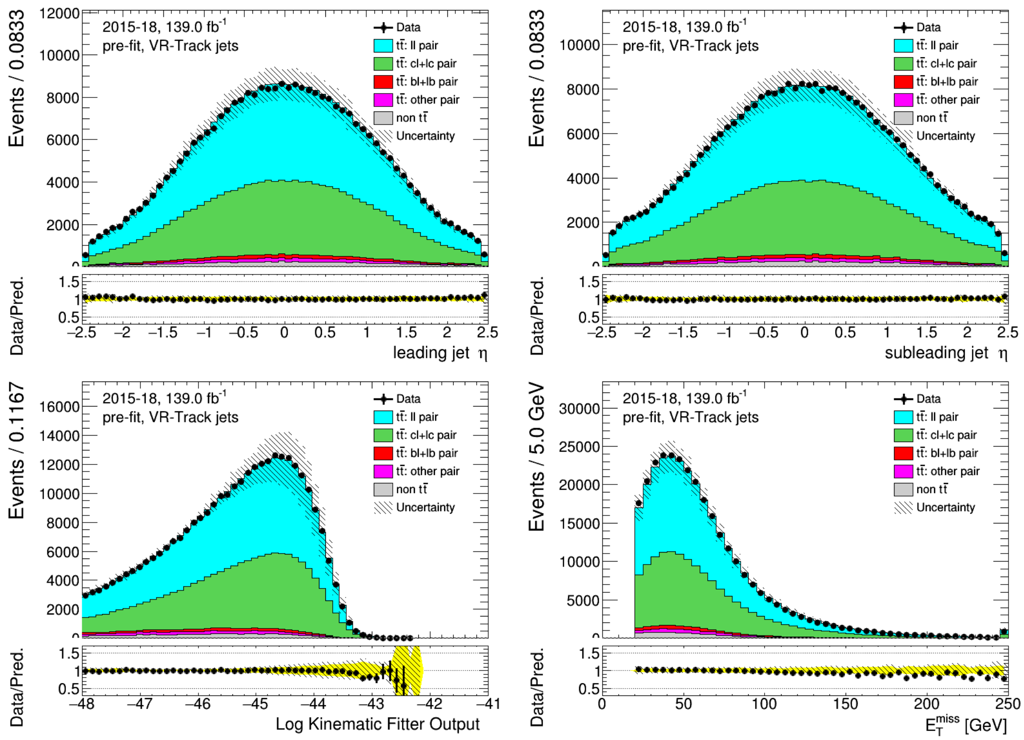


Figure A.19: VR-Track jets: distributions of the leading and sub-leading jets from W decay, KLFitter output and the transverse missing transverse energy of the combined selection, before fitting or tagging with full uncertainties.

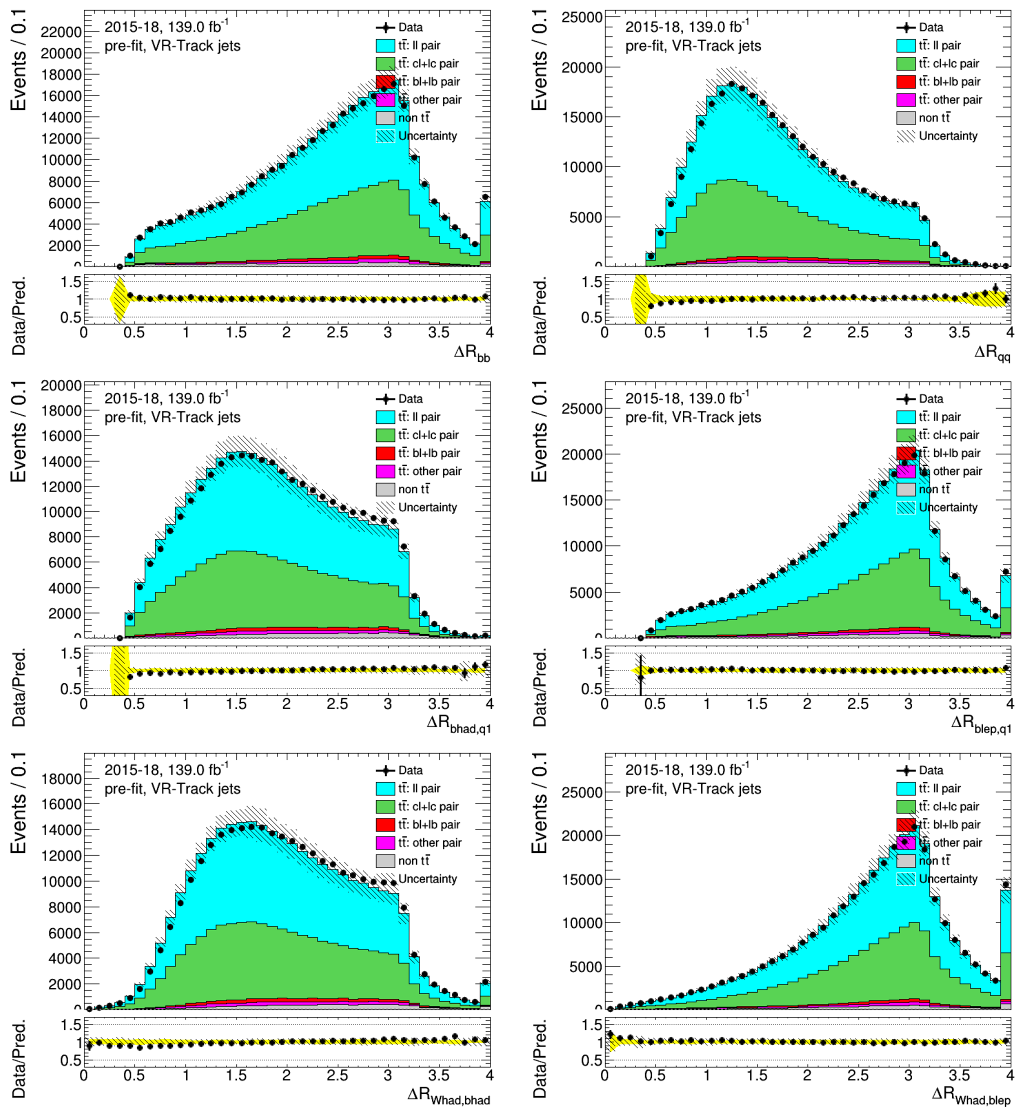


Figure A.20: VR-Track jets: distributions of angle related variables of the combination of the combined selection, before fitting or tagging with full uncertainties.

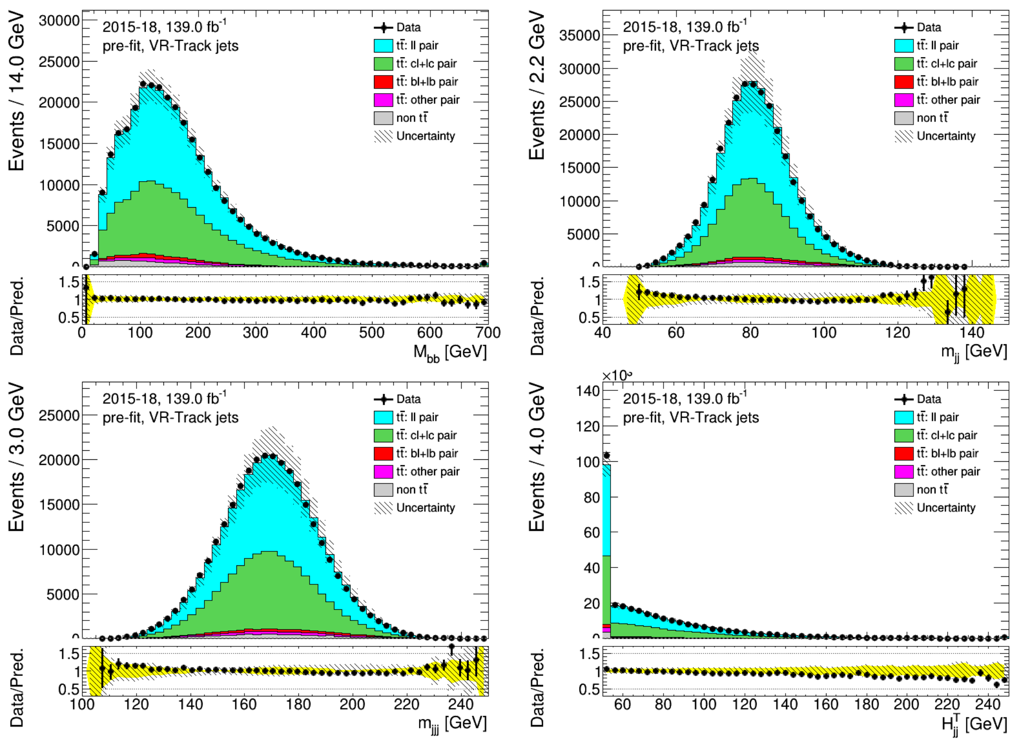


Figure A.21: VR-Track jets: distributions of mass related variables of the combined selection, before fitting or tagging with stat-only uncertainties.

A.4 Experimental uncertainties

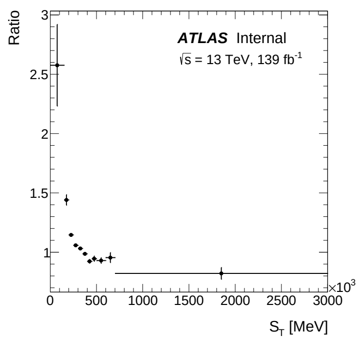
Systematic uncertainty
EG_RESOLUTION_ALL
MUON_ID
MUON_MS
MET_SoftTrk_ResoPara
MET_SoftTrk_ResoPerp
MET_SoftTrk_ScaleDown
MET_SoftTrk_ScaleUp
JET_Pileup_OffsetNPV
JET_Pileup_RhoTopology
JET_EffectiveNP_Modelling1
JET_EffectiveNP_Modelling2
JET_EffectiveNP_Modelling3
JET_EffectiveNP_Modelling4
JET_EffectiveNP_Statistical4
JET_EffectiveNP_Detector1
JET_JER_EffectiveNP_1
JET_JER_EffectiveNP_2
JET_JER_EffectiveNP_3
JET_JER_EffectiveNP_4
JET_BJES_Response
JET_Flavor_Composition
JET_Flavor_Response

Table A.1: List of experimental systematics.

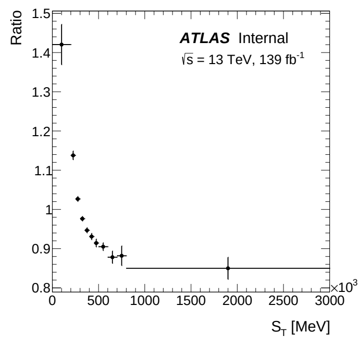
Appendix B

Supplementary material for HH searches

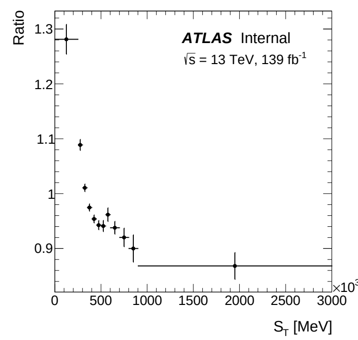
B.1 Additional plots for fake-background estimation



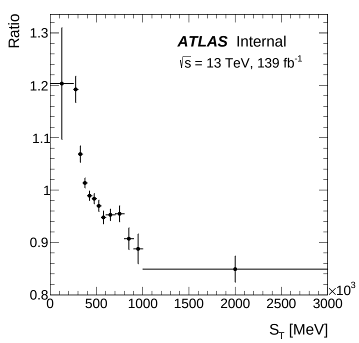
(a) $N_{\mathrm{jets}} = 2$



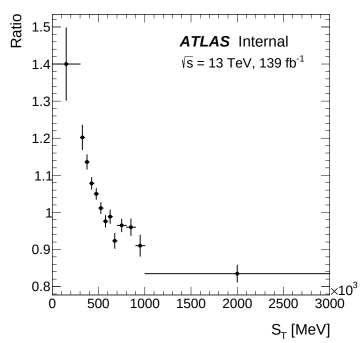
(b) $N_{\mathrm{jets}} = 3$



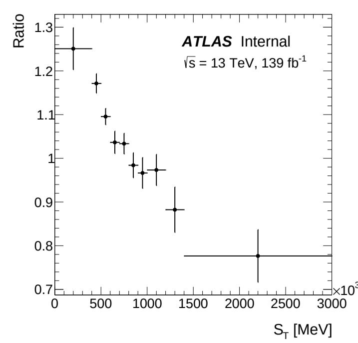
(c) $N_{\mathrm{jets}} = 4$



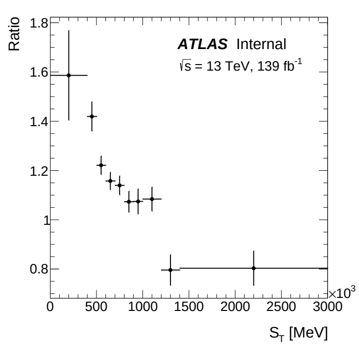
(d) $N_{\mathrm{jets}} = 5$



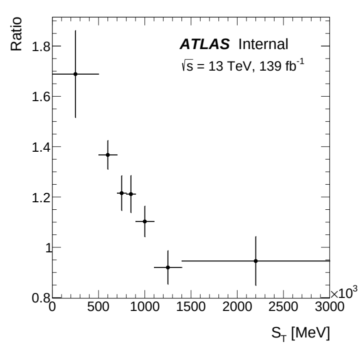
(e) $N_{\mathrm{jets}} = 6$



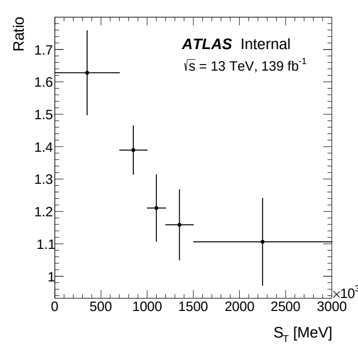
(f) $N_{\mathrm{jets}} = 7$



(g) $N_{\mathrm{jets}} = 8$

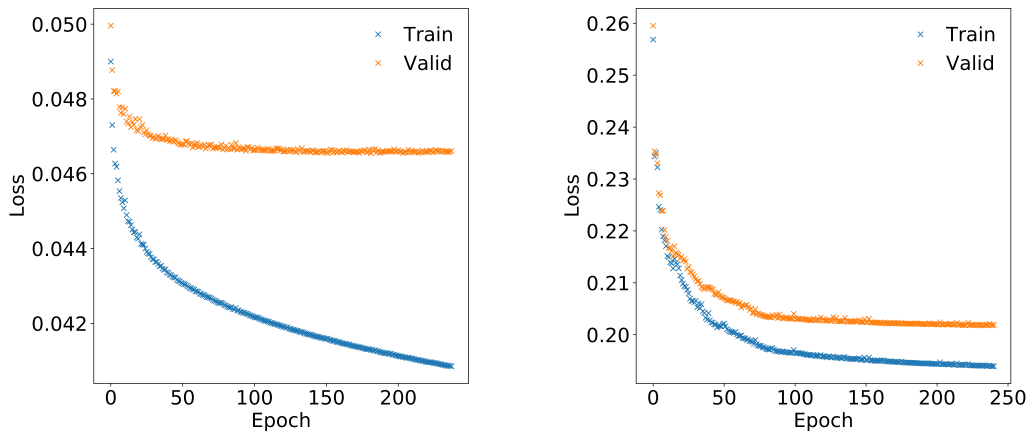


(h) $N_{\mathrm{jets}} = 9$



(i) $N_{\mathrm{jets}} \leq 10$

Figure B.1: The $t\bar{t}$shape correction scale factor as functions of H_{T} in different N_{jets}. The error bars are calculated from the statistical uncertainties of data and simulated samples. Figures reproduced from analysis internal notes.



(a) (b)

Figure B.2: Non-resonant NN loss functions as a function of training epochs evaluated using the two-fold cross validation, for the (a) LTT and (b) SLT categories. Figures reproduced from analysis internal notes, study performed by Elliot Reynolds.

B.2 Additional material for MVA signal extraction

To check if the neural networks are overtrained, the loss functions as a function of the training epoch for the NN (PNNs) are shown in Figure B.2 (Figure B.3) for the SLT and LTT channels. A stable generalisation gap is seen between the training and validation, indicating that the overtraining is small. Further checks are done using a much simpler architecture consisting of three layers of 32 nodes for the LTT channel. A large gap is still seen, indicating the generalisation gap is due to limited data rather than effective capacity of the neural networks. Additional plots of the NN (PNNs) output distribution of the signal and summed background are shown in Figure B.4 and Figure B.5 for the SLT and LTT channel, respectively.

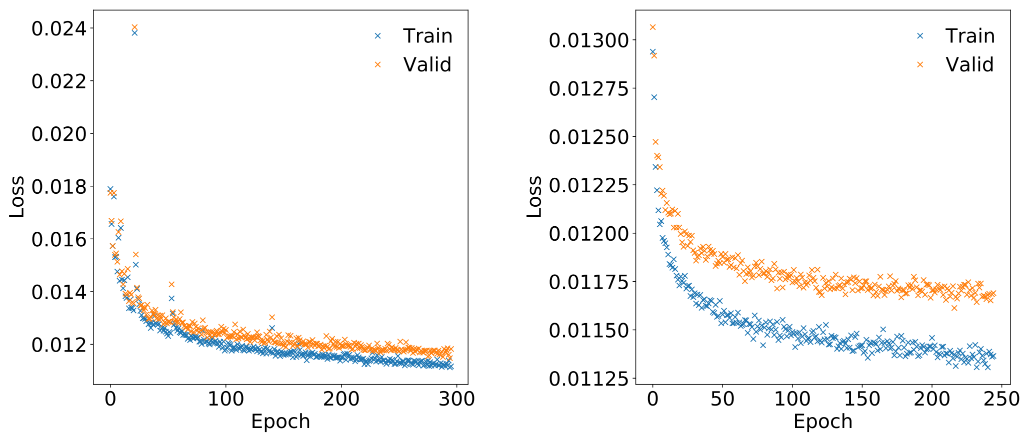
The folloing figures are additional material for the MVA section.

Last bin				
Process	Last bin SM	Last bin 300 GeV	Last bin 500 GeV	Last bin 1000 GeV
ttbar	0.46 ± 0.22	30.9 ± 2	1.3 ± 0.4	0.68 ± 0.28
Fake	1.2 ± 0.7	16.7 ± 2.4	1.1 ± 0.6	0.8 ± 0.6
Stop	0.93 ± 0.35	3 ± 0.6	0.96 ± 0.35	0.9 ± 0.34
ZHF	1.32 ± 0.26	8.6 ± 3.2	1.36 ± 0.28	1.94 ± 0.25
ZLF	0.04 ± 0.025	−0.25 ± 0.25	0.029 ± 0.029	0.27 ± 0.18
WHbb	0.0005 ± 0.00035	0	0.0021 ± 0.0012	0.0033 ± 0.0013
WHtautau	0	0	0.006 ± 0.006	0
qqZHbb	0.278 ± 0.006	0.048 ± 0.008	0.078 ± 0.004	0.393 ± 0.007
ggZHbb	0.054 ± 0.004	0.0008 ± 0.0005	0.09 ± 0.05	0.0342 ± 0.0034
qqZHtautau	0.28 ± 0.04	0.059 ± 0.016	0.084 ± 0.023	0.158 ± 0.028
ggZHtautau	0.052 ± 0.014	0.00016 ± 0.00016	0.061 ± 0.015	0.032 ± 0.011
ggFHtautau	0.25 ± 0.05	0	0.111 ± 0.034	0.15 ± 0.04
VBFHtautau	0.0047 ± 0.0028	0.0025 ± 0.0019	0.0018 ± 0.0018	0.011 ± 0.004
ttH	0.221 ± 0.019	0.09 ± 0.012	0.166 ± 0.016	0.082 ± 0.012
Wjets	0	0	0	0
Diboson	0.2 ± 0.09	0.24 ± 0.09	0.15 ± 0.09	0.53 ± 0.13
DY	0	0	0	0
signal ggF	0.764 ± 0.007	8.03 ± 0.19	61.2 ± 0.8	342.5 ± 2.1
signal VBF	0.00743 ± 0.00024	NA	NA	NA
Second-to-last bin				
Process	Last-1 bin SM	Last-1 bin 300 GeV	Last-1 bin 500 GeV	Last-1 bin 1000 GeV
ttbar	4.5 ± 0.8	61.4 ± 2.9	3.1 ± 0.6	0.82 ± 0.34
Fake	1.9 ± 0.8	25.4 ± 3.2	0.8 ± 0.6	1.4 ± 0.8
Stop	2.4 ± 0.6	3 ± 0.6	1.3 ± 0.4	1.05 ± 0.34
ZHF	4 ± 0.6	13 ± 4	1.44 ± 0.33	2.64 ± 0.32
ZLF	0.3 ± 0.12	0.29 ± 0.27	0.11 ± 0.08	−0.03 ± 0.08
WHbb	0.0044 ± 0.0015	0	0.0026 ± 0.0015	0.0029 ± 0.001
WHtautau	0.013 ± 0.009	0	0	0
qqZHbb	0.448 ± 0.008	0.086 ± 0.01	0.075 ± 0.004	0.232 ± 0.005
ggZHbb	0.16 ± 0.05	0.0022 ± 0.0008	0.0331 ± 0.0034	0.06 ± 0.04
qqZHtautau	0.44 ± 0.05	0.07 ± 0.018	0.099 ± 0.021	0.14 ± 0.026
ggZHtautau	0.126 ± 0.022	0	0.059 ± 0.015	0.017 ± 0.008
ggFHtautau	0.25 ± 0.05	0.069 ± 0.029	0.093 ± 0.034	0.17 ± 0.04
VBFHtautau	0.012 ± 0.005	0.009 ± 0.004	0.0031 ± 0.0022	0.01 ± 0.004
ttH	0.27 ± 0.021	0.11 ± 0.012	0.128 ± 0.014	0.07 ± 0.01
Wjets	0	0	0	0
Diboson	0.56 ± 0.17	0.55 ± 0.2	0.12 ± 0.07	0.82 ± 0.18
DY	0	0	0	0
signal ggF	0.599 ± 0.006	7.53 ± 0.18	21 ± 0.5	42 ± 0.7
signal VBF	0.00869 ± 0.00027	NA	NA	NA
Third-to-last bin				
Process	Last-2 bin SM	Last-2 bin 300 GeV	Last-2 bin 500 GeV	Last-2 bin 1000 GeV
ttbar	9.1 ± 1.1	88 ± 3.4	5.2 ± 0.9	209 ± 5
Fake	4.7 ± 1.5	44 ± 4	3.6 ± 1.2	395 ± 13
Stop	5.3 ± 0.8	4.3 ± 0.7	1.2 ± 0.4	104 ± 4
ZHF	7.1 ± 0.7	11 ± 4	3.3 ± 0.7	212 ± 6
ZLF	0.54 ± 0.28	0.9 ± 0.7	0.04 ± 0.11	21.9 ± 2.9
WHbb	0.021 ± 0.004	0.013 ± 0.008	0.0053 ± 0.002	0.511 ± 0.018
WHtautau	0	0	0.005 ± 0.005	0.101 ± 0.029
qqZHbb	0.677 ± 0.011	0.123 ± 0.012	0.121 ± 0.006	3.355 ± 0.022
ggZHbb	0.24 ± 0.06	0.0024 ± 0.0009	0.043 ± 0.004	1.31 ± 0.11
qqZHtautau	0.45 ± 0.05	0.063 ± 0.018	0.159 ± 0.028	1.71 ± 0.09
ggZHtautau	0.161 ± 0.024	0	0.059 ± 0.015	0.57 ± 0.05
ggFHtautau	0.41 ± 0.07	0.089 ± 0.033	0.092 ± 0.031	4.23 ± 0.22
VBFHtautau	0.033 ± 0.008	0.009 ± 0.004	0.01 ± 0.004	0.402 ± 0.026
ttH	0.482 ± 0.028	0.159 ± 0.015	0.188 ± 0.017	4.45 ± 0.08
Wjets	0.08 ± 0.08	0.07 ± 0.07	0	6 ± 0.9
Diboson	1.03 ± 0.19	0.95 ± 0.22	0.26 ± 0.08	22.2 ± 1.5
DY	0	0	0	0.7 ± 0.4
signal ggF	0.596 ± 0.006	7.36 ± 0.18	20.6 ± 0.5	40.5 ± 0.7
signal VBF	0.00997 ± 0.00029	NA	NA	NA

Table B.1: Pre-fit event yields in the last, second-to-last and third-to-last MVA bin of the di-Higgs

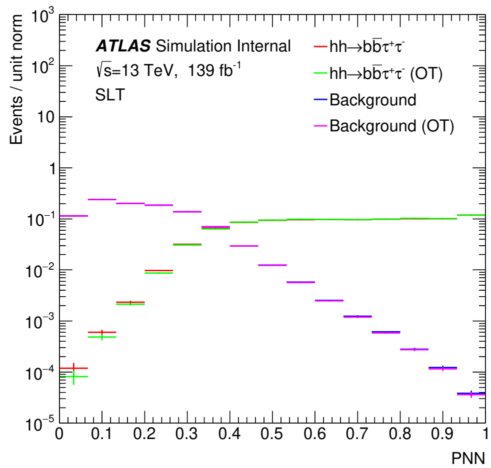
Last bin				
Process	Last bin SM	Last bin 300 GeV	Last bin 500 GeV	Last bin 1000 GeV
ttbar	0.76 ± 0.31	5 ± 0.8	2.3 ± 0.6	0.57 ± 0.29
Fake	2.6 ± 1.4	3.7 ± 1.7	1 ± 0.8	0.9 ± 0.9
Stop	0.32 ± 0.19	0.14 ± 0.14	0.17 ± 0.12	0.47 ± 0.25
ZHF	1.03 ± 0.23	3.3 ± 2.2	1.3 ± 0.4	2.1 ± 0.5
ZLF	0.07 ± 0.07	0	0.05 ± 0.09	0.52 ± 0.24
WHbb	0.0013 ± 0.0009	0	0	0.004 ± 0.004
WHtautau	0	0	0	0
qqZHbb	0.0656 ± 0.0032	0.0073 ± 0.0028	0.0416 ± 0.0027	0.0785 ± 0.0032
ggZHbb	0.0191 ± 0.0026	0.0004 ± 0.0004	0.0182 ± 0.0024	0.0063 ± 0.0014
qqZHtautau	0.065 ± 0.018	0.019 ± 0.01	0.051 ± 0.016	0.052 ± 0.018
ggZHtautau	0.02 ± 0.009	0	0.022 ± 0.009	0
ggFHtautau	0.076 ± 0.034	0	0.076 ± 0.03	0.11 ± 0.04
VBFHtautau	0.005 ± 0.0029	0	0.0018 ± 0.0018	0.0035 ± 0.0025
ttH	0.044 ± 0.008	0.014 ± 0.005	0.065 ± 0.01	0.031 ± 0.007
Wjets	0	0	0	0
Diboson	−0.01 ± 0.05	−0.05 ± 0.04	0.03 ± 0.04	0.44 ± 0.14
DY	0	0	0	0
signal ggF	0.1977 ± 0.0035	2.61 ± 0.11	20.8 ± 0.5	45.1 ± 0.8
signal VBF	0.00344 ± 0.00017	NA	NA	NA
Second-to-last bin				
Process	Last-1 bin SM	Last-1 bin 300 GeV	Last-1 bin 500 GeV	Last-1 bin 1000 GeV
ttbar	3.8 ± 0.7	8.5 ± 1.1	2.5 ± 0.6	108 ± 4
Fake	2 ± 1.4	1.4 ± 1.1	1.7 ± 1.2	84 ± 9
Stop	0.37 ± 0.21	0.45 ± 0.24	0.08 ± 0.08	18 ± 1.7
ZHF	1.6 ± 0.4	2.5 ± 1.1	0.82 ± 0.19	86.4 ± 3.4
ZLF	0.15 ± 0.1	0.07 ± 0.07	0	11 ± 4
WHbb	0	0.0027 ± 0.0027	0.0016 ± 0.0011	0.0074 ± 0.002
WHtautau	0	0	0	0.047 ± 0.021
qqZHbb	0.083 ± 0.004	0.022 ± 0.005	0.0384 ± 0.0033	0.826 ± 0.013
ggZHbb	0.0262 ± 0.0029	0.00022 ± 0.00016	0.0154 ± 0.0023	0.44 ± 0.06
qqZHtautau	0.088 ± 0.021	0.011 ± 0.008	0.06 ± 0.018	0.57 ± 0.05
ggZHtautau	0.027 ± 0.01	0	0.028 ± 0.01	0.212 ± 0.028
ggFHtautau	0.094 ± 0.032	0.015 ± 0.011	0.022 ± 0.015	1.48 ± 0.14
VBFHtautau	0.005 ± 0.0029	0.0025 ± 0.0024	0.0043 ± 0.0027	0.12 ± 0.014
ttH	0.062 ± 0.01	0.02 ± 0.005	0.046 ± 0.009	1.45 ± 0.05
Wjets	0	0	0	0.83 ± 0.31
Diboson	0.14 ± 0.08	0.1 ± 0.09	0 ± 0.04	3.7 ± 0.4
DY	0	0	0.15 ± 0.15	0.15 ± 0.15
signal ggF	0.1402 ± 0.0028	2.72 ± 0.11	5.87 ± 0.25	2.28 ± 0.17
signal VBF	0.00294 ± 0.00016	NA	NA	NA
Third-to-last bin				
Process	Last-2 bin SM	Last-2 bin 300 GeV	Last-2 bin 500 GeV	Last-2 bin 1000 GeV
ttbar	6.3 ± 0.9	9.7 ± 1.2	3.4 ± 0.7	4104 ± 24
Fake	1.4 ± 1.3	6.6 ± 2.4	2.4 ± 1.3	2060 ± 40
Stop	0.85 ± 0.34	0.3 ± 0.21	0.47 ± 0.25	179 ± 5
ZHF	4.8 ± 0.9	2.6 ± 1.1	2.1 ± 0.4	368 ± 14
ZLF	0.18 ± 0.09	0.25 ± 0.2	0.05 ± 0.04	37 ± 9
WHbb	0.0012 ± 0.0008	0	0.001 ± 0.0007	0.162 ± 0.026
WHtautau	0	0.006 ± 0.006	0	0.131 ± 0.033
qqZHbb	0.124 ± 0.006	0.041 ± 0.007	0.0457 ± 0.0032	2.95 ± 0.05
ggZHbb	0.0323 ± 0.0032	0.0013 ± 0.0006	0.0221 ± 0.0026	1.01 ± 0.13
qqZHtautau	0.129 ± 0.026	0.015 ± 0.009	0.052 ± 0.016	1.56 ± 0.09
ggZHtautau	0.039 ± 0.012	0.004 ± 0.004	0.028 ± 0.01	0.53 ± 0.04
ggFHtautau	0.098 ± 0.029	0.005 ± 0.005	0.085 ± 0.035	2.98 ± 0.2
VBFHtautau	0.014 ± 0.005	0.0019 ± 0.0019	0.0033 ± 0.0023	0.248 ± 0.02
ttH	0.111 ± 0.013	0.024 ± 0.006	0.043 ± 0.008	9.52 ± 0.11
Wjets	0	0	0	3.2 ± 1
Diboson	0.18 ± 0.08	0 ± 0.08	0.09 ± 0.07	16.3 ± 1
DY	0	0	0	0.68 ± 0.29
signal ggF	0.1382 ± 0.0028	2.61 ± 0.11	4.93 ± 0.23	0.009 ± 0.009
signal VBF	0.00331 ± 0.00017	NA	NA	NA

Table B.2: Pre-fit event yields in the last, second-to-last and third-to-last MVA bin of the di-Higgs

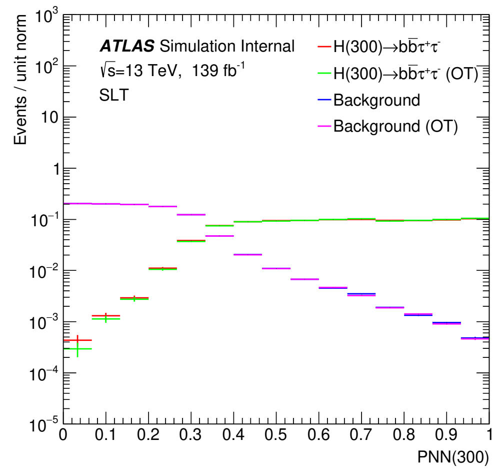


(a) (b)

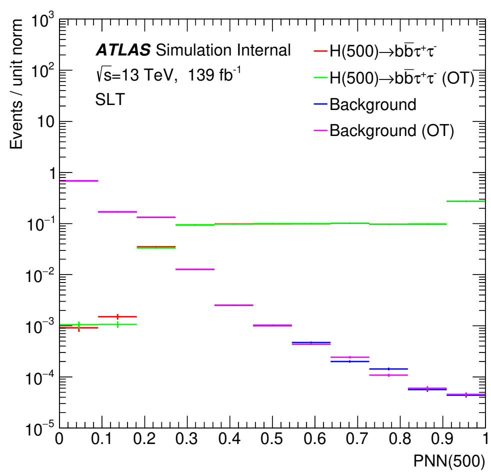
Figure B.3: PNN loss functions as a function of training epochs evaluated using the two-fold cross validation, for the (a) LTT and (b) SLT categories. Figures reproduced from analysis internal notes, study performed by Elliot Reynolds.



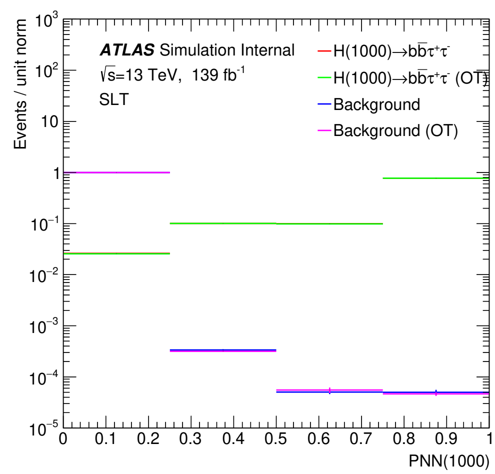
(a)



(b)

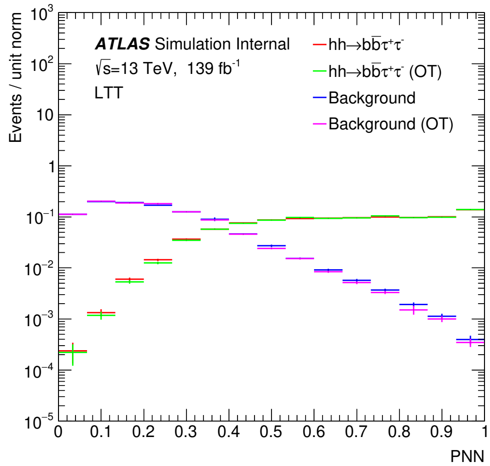


(c)

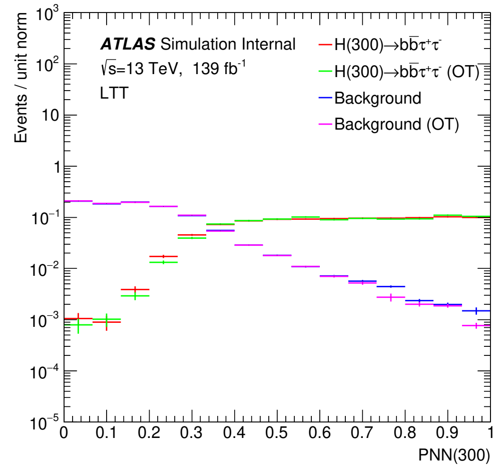


(d)

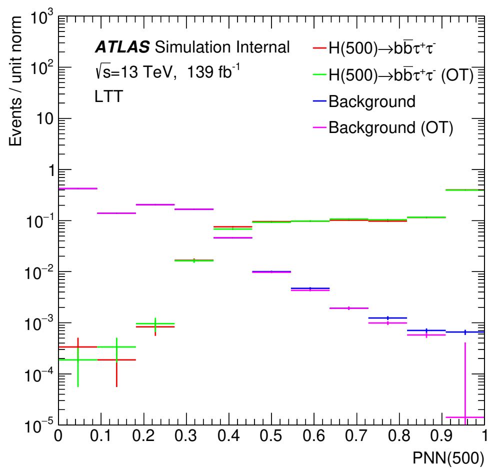
Figure B.4: (P)NN output distributions for (a) non-resonant signal, and target signal masses of (b) 300 GeV, (c) 500 GeV, and (d) 1000 GeV obtained evaluating the MVA with even-odd events crossing or evaluating the MVA on the training datasets for the SLT category. The "OT" histograms refer to the "OverTraining" check in which the (P)NN is applied to the same data on which it was trained, the other histograms are obtained evaluating the MVA with even-odd event crossing. Images reproced from analysis internal notes.



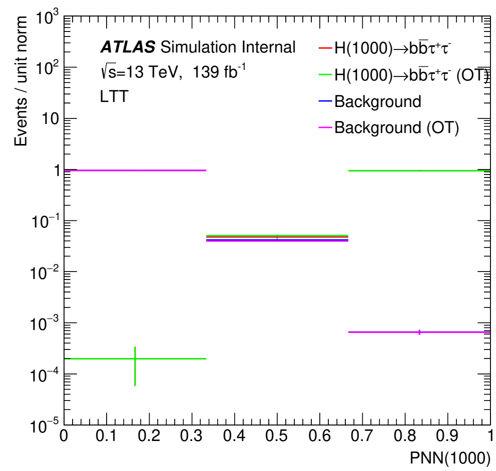
(a)



(b)



(c)



(d)

Figure B.5: (P)NN output distributions for (a) non-resonant signal, and target signal masses of (b) 300 GeV, (c) 500 GeV, and (d) 1000 GeV obtained evaluating the MVA with even-odd events crossing or evaluating the MVA on the training datasets for the LTT category. The “OT” histograms refer to the “OverTraining” check in which the (P)NN is applied to the same data on which it was trained, the other histograms are obtained evaluating the MVA with even-odd event crossing. Images reproced from analysis internal notes.

Process	Name	LepHad SLT	LepHad LTT	Comment
ttbar	THEO_ACC _TTBAR_ME	N: +0.0026, -0.0026, S	N: -0.009, +0.009	Matrix element acceptance
ttbar	THEO_ACC _TTBAR_PS	N: -0.072, +0.072, S	N: -0.088, +0.088	Parton shower acceptance
ttbar	THEO_ACC _TTBAR_ISR	N: +0.0005, -0.0081	N:-0.0052, +0.013	ISR acceptance
ttbar	THEO_ACC _TTBAR_FSR	N: +0.014, -0.0097	N: +0.0096, -0.032	FSR acceptance
ttbar	THEO_ACC _TTBAR_PDFalphas	N: -0.006, +0.006	N: -0.0011, +0.0011	PDF+α_S acceptance
Z+hf	THEO_ACC_Zhf_GENERATOR	N: +0.021, -0.021	N: +0.10, -0.10	Matrix element acceptance
Z+hf	THEO_ACC_Zhf_SCALE	N: -0.029, +0.053, S	N: -0.054, +0.085, S	Scale acceptance
Z+hf	THEO_ACC_Zhf_CKKW	N: +0.07, -0.07	N: +0.071, -0.071	CKKW acceptance
Z+hf	THEO_ACC_Zhf_QSF	N: -0.016, +0.016	N: -0.016 , +0.016	QSF acceptance
Z+hf	THEO_ACC_Zhf_PDFalphas	N: -0.0026, +0.0026	N: -0.0033 , +0.0033	PDF+α_S acceptance
Z+hf	THEO_ACC_Zhf_PDFChoice	N: -0.0097, +0.0097	N: -0.011, +0.011	PDF choice acceptance
stopWt	THEO_XS_Stop	N: -0.054, +0.054	N: -0.054, +0.054	cross section
stopWt	THEO_ACC_StopWt_ME	N: - 0.022, +0.022	N: -0.15, 0.15	Matrix element acceptance
stopWt	THEO_ACC_StopWt_PS	N: +0.077, -0.077	N: -0.093, +0.093	Parton shower acceptance
stopWt	THEO_ACC_StopWt_ISR	N: -0.047, +0.064	N: -0.045, +0.062	ISR acceptance
stopWt	THEO_ACC_StopWt_FSR	N: -0.054, +0.043	N: -0.069, +0.035	FSR acceptance
stopWt	THEO_ACC_StopWt_PDF	N: -0.032, +0.032	N: -0.032, +0.032	PDF acceptance
stopWt	THEO_ACC_StopWt_TopInterference	N: +0.078, -0.078, S	N: +0.11, +0.11, S	top interference acceptance
ttH	THEO_XS_SCALE_ttH	N: -0.092, +0.058	N: -0.092, +0.058	Scale cross section
ttH	THEO_XS_PDFalphas_ttH	N: -0.036, +0.036	N: -0.036, +0.036	PDF+α_S cross section
ttH	THEO_ACC_GEN_ttH	-	N: -0.019, +0.019	Matrix element acceptance
ttH	THEO_ACC_PS_ttH	N: -0.013, +0.013	N: -0.067, +0.067	Parton shower acceptance
ttH	THEO_ACC_SCALE_ttH	-	-	Scale acceptance
ttH	THEO_ACC_ISR_ttH	-	N: -0.01, +0.01	ISR acceptance
ttH	THEO_ACC_FSR_ttH	N: -0.051, +0.032,	N: -0.15 , +0.055	FSR acceptance
ggFHtautau	THEO_XS_SCALE_ggFH	N: -0.039, +0.039	N: -0.039, +0.039	Scale cross section
ggFHtautau	THEO_XS_PDFalphas_ggFH	N: -0.032, +0.032	N: -0.032, +0.032	PDF+α_S section
ggF, ZH, WH, VBFH with Htautau	THEO_BR_Htautau	N: -0.02, +0.02	N: -0.02, +0.02	Htautau BR
ggFHtautau	THEO_ACC_HF_ggFH	N: -1.0, +1.0	N: -1.0, +1.0	Higgs + HF mod unc
qqZHbb, qqZHtautau	THEO_XS_SCALE_qqZH	N: -0.006,+0.005,	-0.006, +0.005	Scale cross section
qqZHbb, qqZHtautau	THEO_XS_PDFalphas_qqZH	N: -0.019, +0.019	N: -0.019, +0.019	PDF+α_S cross section
ggZHbb, ggZHtautau	THEO_XS_SCALE_ggZH	N: -0.19 +0.25,	N: -0.19 , +0.25	Scale cross section
ggZHbb, ggZHtautau	THEO_XS_PDFalphas_qqZH	N: -0.024, +0.024	N: -0.024, +0.024	PDF+α_S cross section
ZHbb, WHbb	THEO_BR_Hbb	N: -0.013, +0.013	N: -0.013, +0.013	Hbb BR
ZHbb	THEO_ACC_PS_ZHbb	N: -0.11, +0.11	N: -0.037, +0.037	Parton shower acceptance
ZHbb	THEO_ACC_SCALE_ZHbb	N: -0.030, +0.030	N: -0.025, +0.025	Scale acceptance
ZHtautau	THEO_ACC_PS_ZHtautau	N: -0.055, +0.055	N: -0.15, +0.15	Parton shower acceptance
ZHtautau	THEO_ACC_PDFalphas_ZHtautau	-	N: -0.012, +0.012	PDF+α_S acceptance
ZHtautau	THEO_ACC_SCALE_ZHtautau	N: -0.022, +0.022	N: -0.028, +0.028	Scale acceptance
WHbb, WHtautau	THEO_XS_SCALE_WH	N: -0.007, +0.005	N: -0.007 , +0.005	Scale cross section
WHbb, WHtautau	THEO_XS_PDFalphas_WH	N: -0.019, +0.019	N: -0.019, +0.019	PDF+α_S cross section
WHtautau	THEO_ACC_HF_WH	N: -1.0, +1.0	N: -1.0, +1.0	Higgs + HF mod unc
VBFHtautau	THEO_XS_SCALE_VBFH	N: -0.003, +0.004	-0.003, +0.004	Scale cross section
VBFHtautau	THEO_XS_PDFalphas_VBFH	N: -0.021, +0.021	N: -0.021, +0.021	PDF+α_S cross section
VBFHtautau	THEO_ACC_HF_VBFH	N: -1.0, +1.0	N: -1.0, +1.0	Higgs + HF mod unc

Table B.3: List of MC background uncertainties for major backgrounds. The table shows the process, the name of the uncertainty, the relative size of the normalisation uncertainty (N) and whether the uncertainty includes also a shape variation (S) in the different SRs.

B.3 Addional material for systematic uncertainties

In the following plots, the parametrisation of different MVA scores are shown. In Fig. B.6 (B.7), the shape only ratio of the PS (ME) variation sample to the nominal sample (both AF2) in the final fit binning is shown. This ratio is applied on the full sim nominal sample to mimic the effect from the systematic varations.

The resonant signal PS uncertainties are parametrised by linear functions. The function used in the SLT SR is:

- Down variation = 1.82 - 0.95 * PNN Score.
- Up variation = 0.18 + 0.95 * PNN Score;

The function used in the LTT SR is:

- Down variation = 1.84 - 1.01 * PNN Score;
- Up variation = 0.16 + 1.01 * PNN Score.

Process	Name	Size	Comment
V+jets	THEO_XS_V	-0.05, +0.05	cross section
Z+lf	THEO_ACC_Zlf	-0.23, +0.23	Acceptance from VHbb
W+jets	THEO_ACC_W	-0.37, +0.37	Acceptance from VHbb in LepHad SRs
W+jets	THEO_ACC_W	-0.50, +0.50	Acceptance inflated from VHbb analysis for tau fakes in HadHad SR
WW, WZ, ZZ	THEO_XS_Diboson	-0.06, +0.06	cross section
WW	THEO_ACC_Diboson	-0.25, +0.25	Acceptance from VHbb
WZ	THEO_ACC_Diboson	-0.26, +0.26	Acceptance from VHbb
ZZ	THEO_ACC_Diboson	-0.20, +0.20	Acceptance from VHbb

Table B.4: List of MC background uncertainties for minor backgrounds. The table shows the process, the name of the uncertainty, the relative size of the normalisation uncertainty (N). Table reproduced from analysis internal notes.

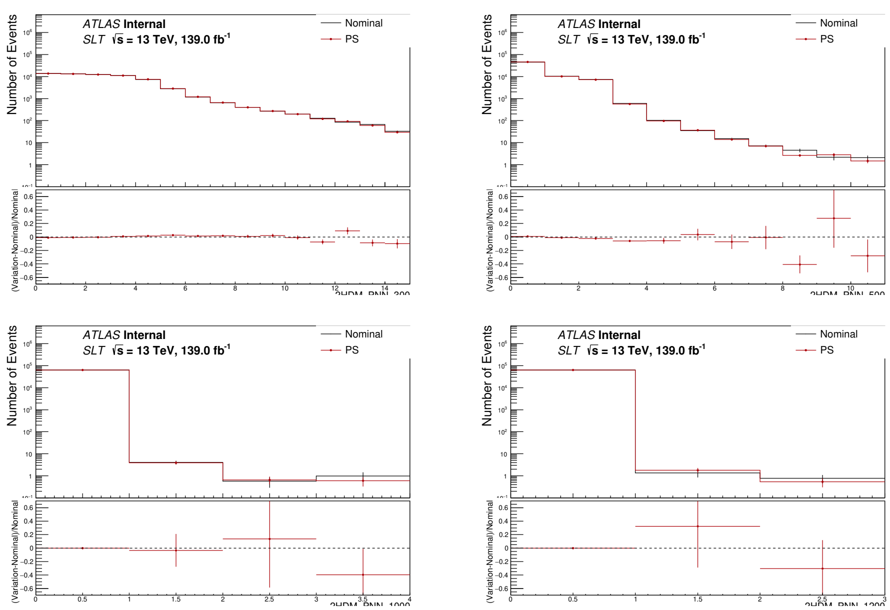


Figure B.6: Shape-only ratio of the PS variation to the nominal in PNN score of various resonant mass.

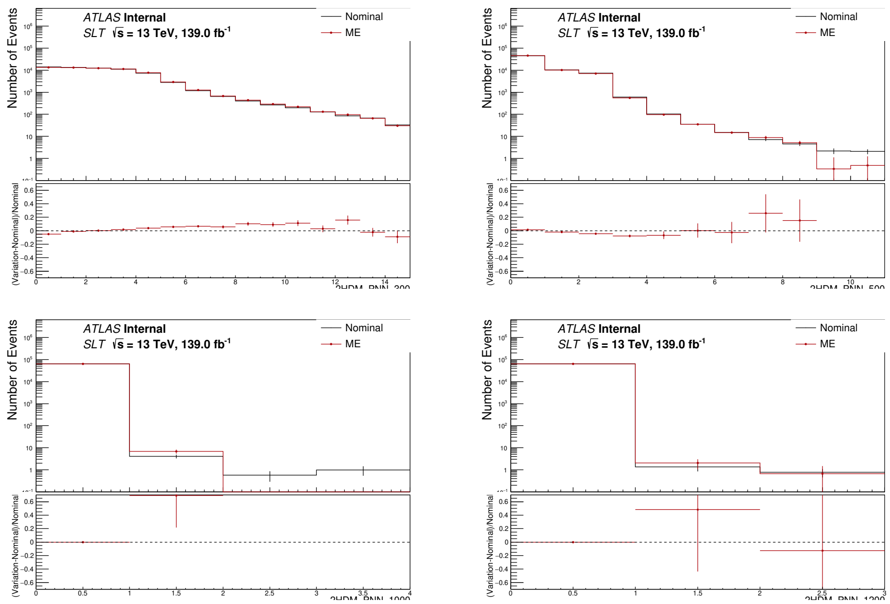


Figure B.7: Shape-only ratio of the ME variation to the nominal in PNN score of various resonant mass.

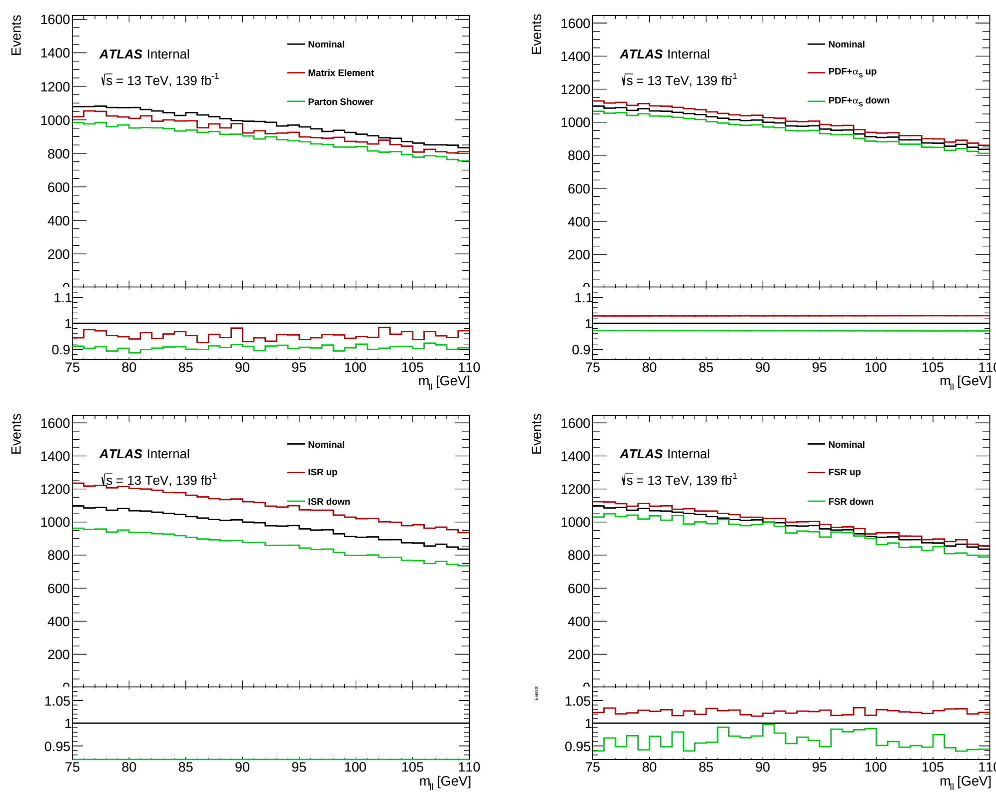
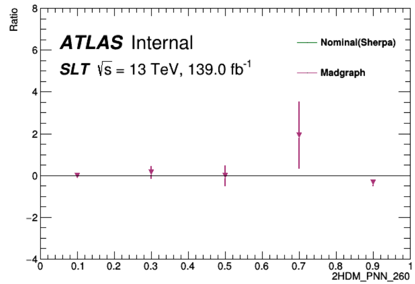
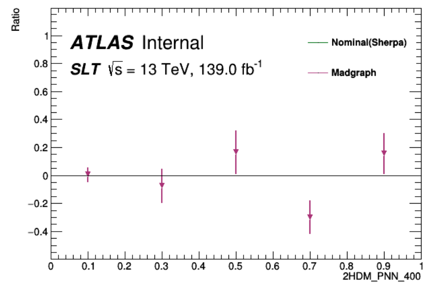


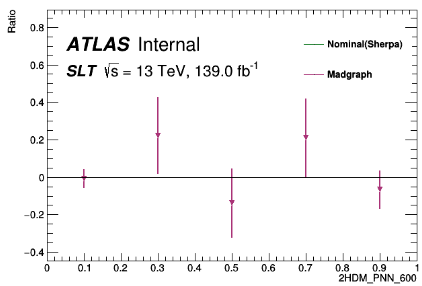
Figure B.8: m_{ll} spectra comparisons for nominal and systematics from matrix element and parton shower (top left), PDF+α_S (top right), ISR (bottom left) and FSR (bottom right). Images reproced from analysis internal notes.



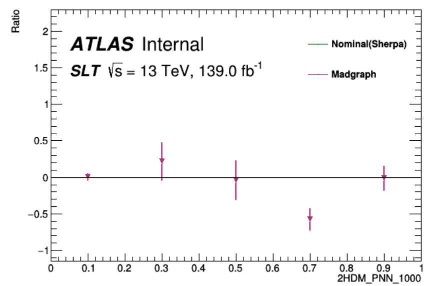
(a)



(b)

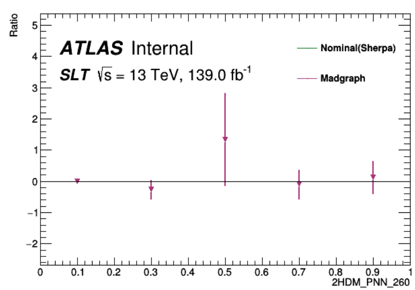


(c)

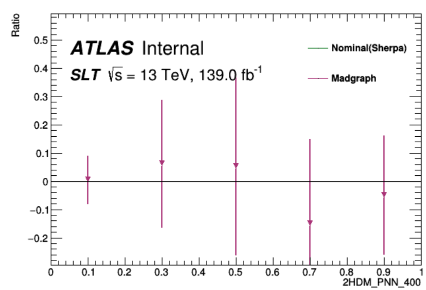


(d)

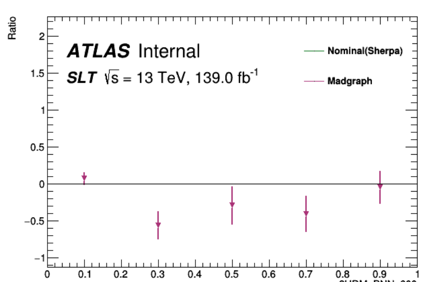
Figure B.9: SLT channel: comparision of the nominal Sherpa sample verse the variation Madgraph in the PNN score distribtuion.



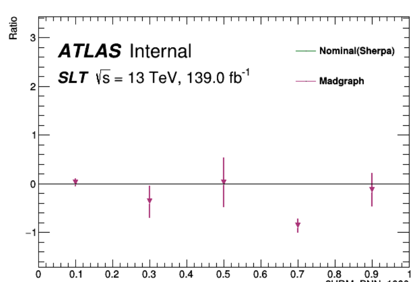
(a)



(b)

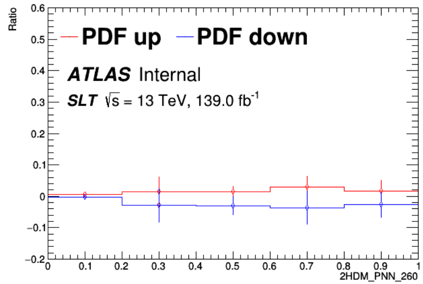


(c)

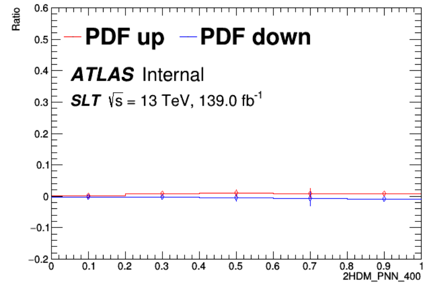


(d)

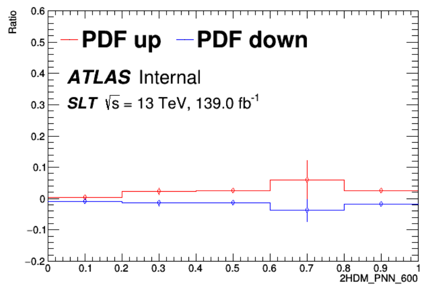
Figure B.10: LTT channel: comparision of the nominal Sherpa sample verse the variation Madgraph in the PNN score distribtuion.



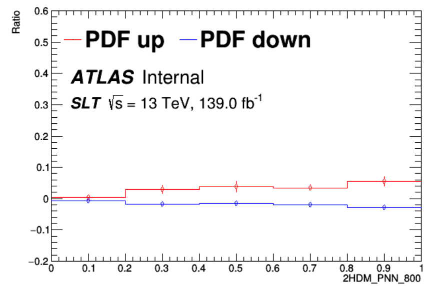
(a)



(b)

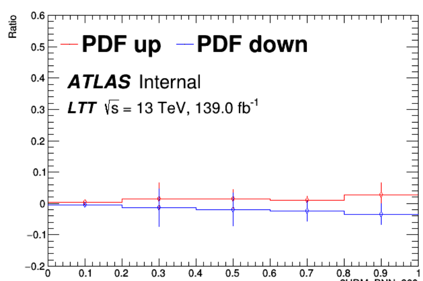


(c)

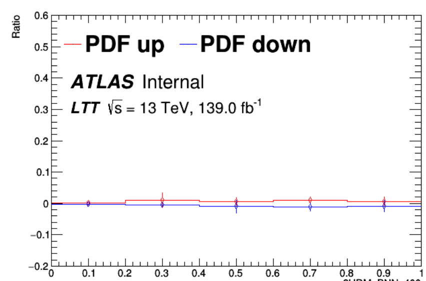


(d)

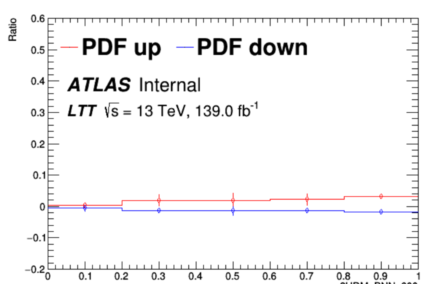
Figure B.11: SLT channel: Z+HF intra-PDF variation in PNN score of various resonant mass.



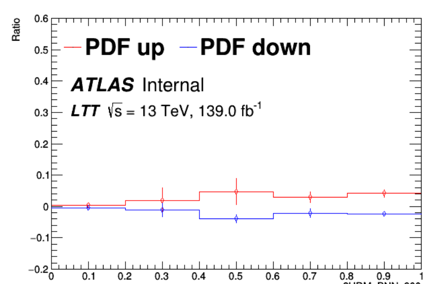
(a)



(b)

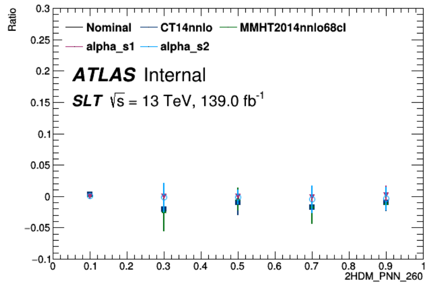


(c)

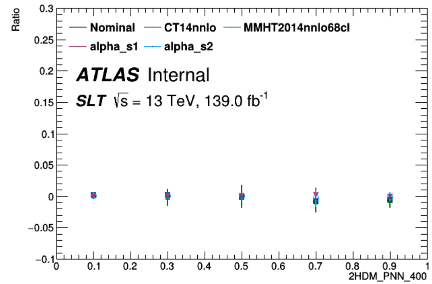


(d)

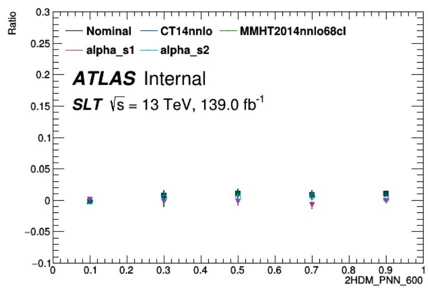
Figure B.12: LTT channel: Z+HF intra-PDF variation in the PNN score of various resonant masses.



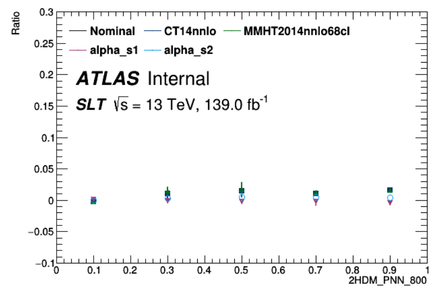
(a)



(b)

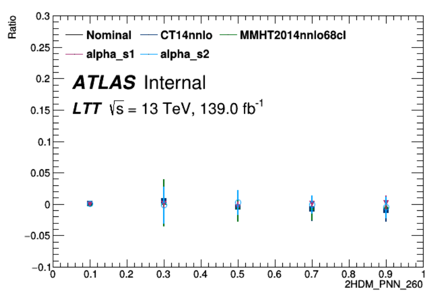


(c)

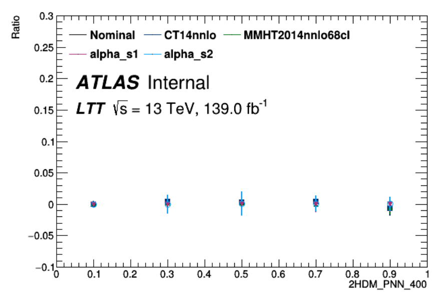


(d)

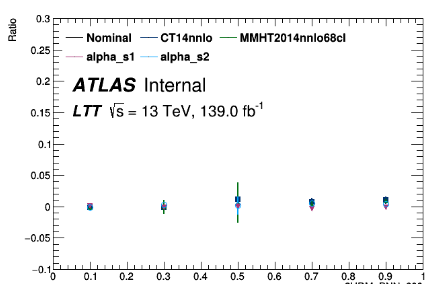
Figure B.13: SLT channel: Z+HF inter-PDF variations and α_s variations in PNN score of various resonant mass.



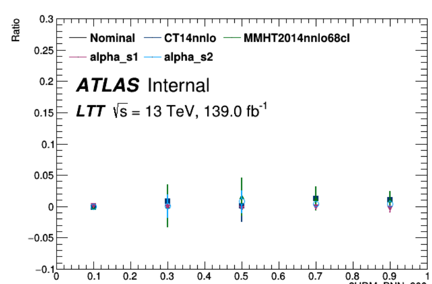
(a)



(b)

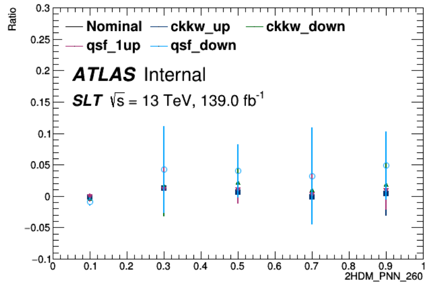


(c)

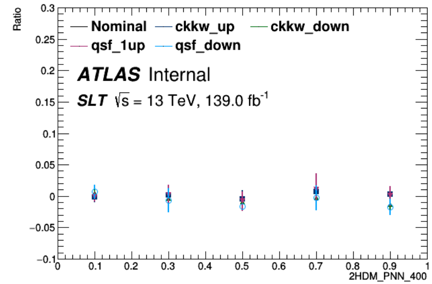


(d)

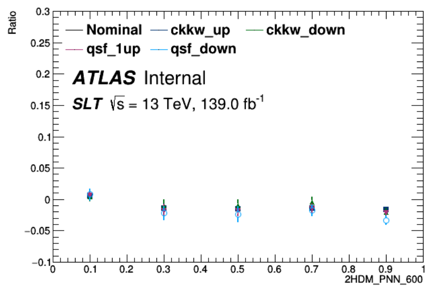
Figure B.14: LTT channel: Z+HF inter-PDF variations and α_s variations in the PNN score of various resonant masses.



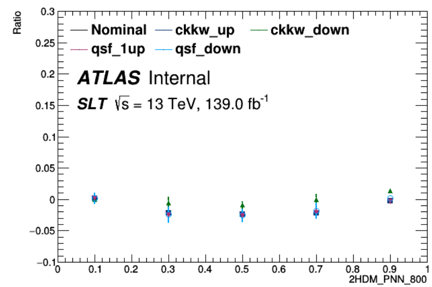
(a)



(b)

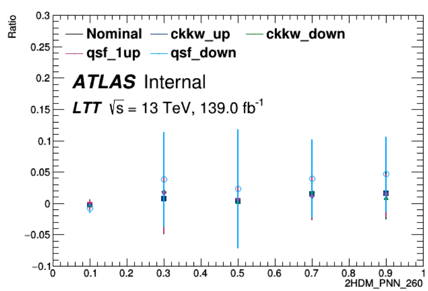


(c)

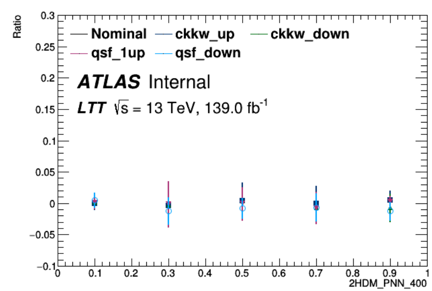


(d)

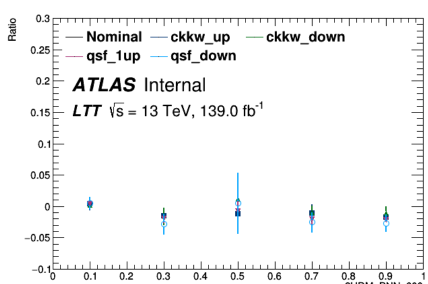
Figure B.15: SLT channel: Z+HF ckkw and qsf variations in PNN score of various resonant mass.



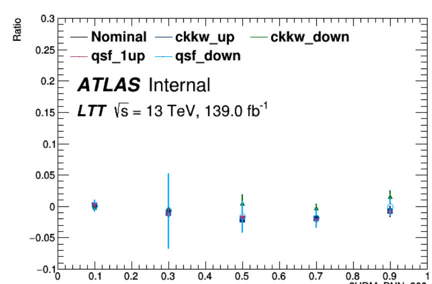
(a)



(b)



(c)



(d)

Figure B.16: LTT channel:Z+HF ckkw and qsf variations in the PNN score of various resonant masses.

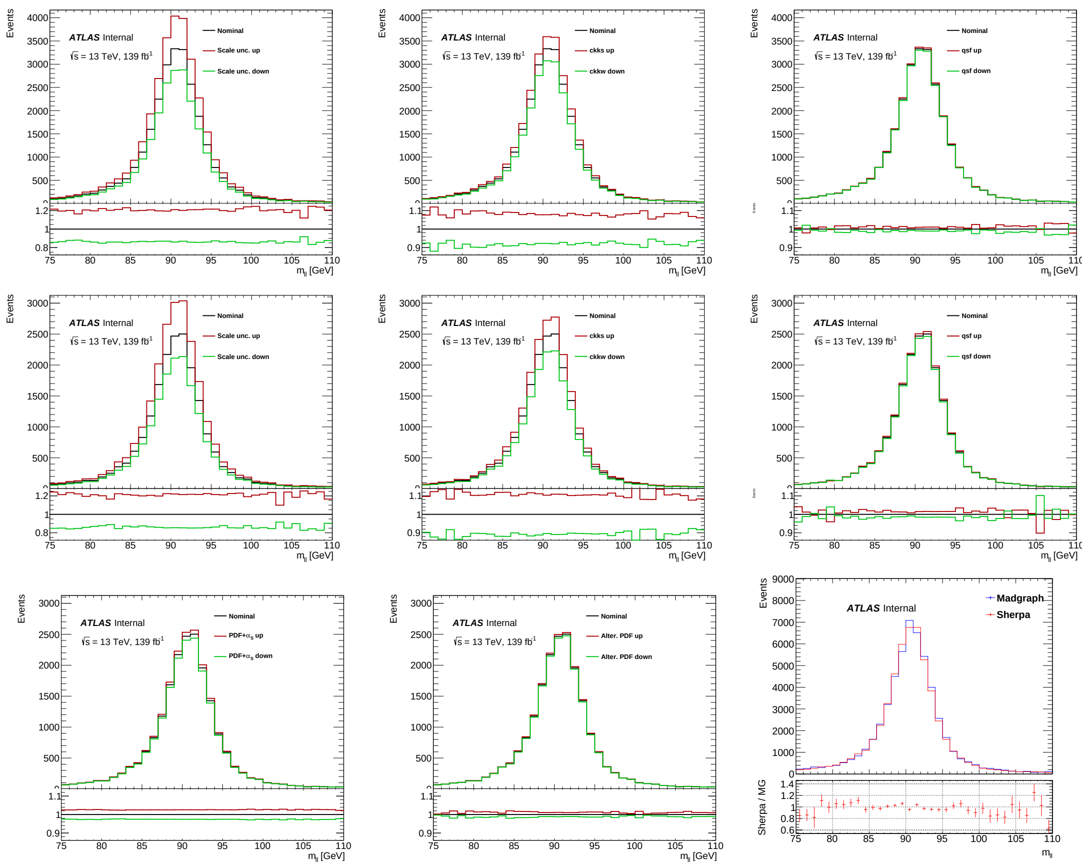


Figure B.17: Top row: the m_{ll} spectra comparisons for nominal and systematics from scale variations (left), ckkw (middle) and qsf (right) on Z($\rightarrow$ $\mu\mu$) + HF (first row) and $Z(\rightarrow \ell\ell)$+HF (second row) processes. Bottom row: the m_{ll} spectra comparisons for nominal and systematics from PDF+α_S (left) and alternative PDF (middle) variations for $Z(\rightarrow ee)$+HF process and on the right for the nominal Sherpa generator and alternative MadGraph generator for $Z(\rightarrow \ell\ell)$+HF process. Images reproduced from analysis internal notes.

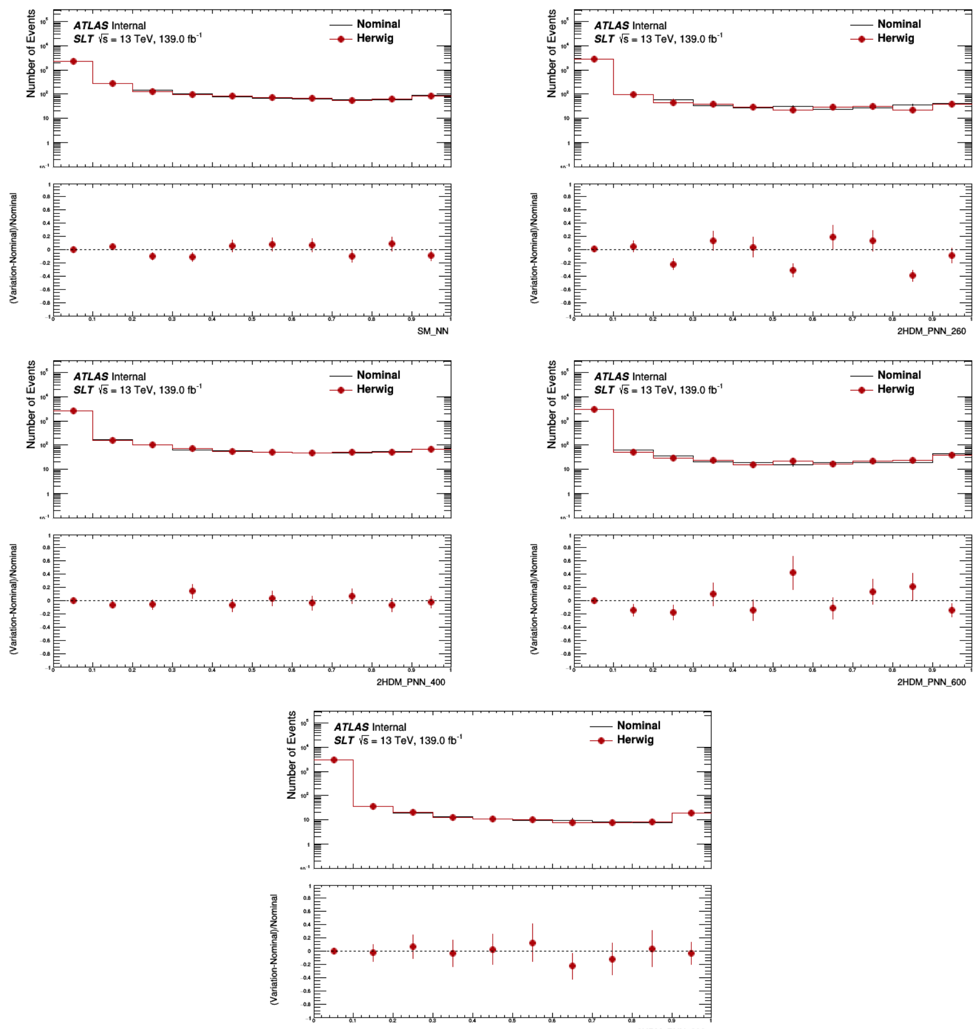


Figure B.18: SLT channel: shape-only PS uncertainty in the NN and PNN score of various resonant masses.

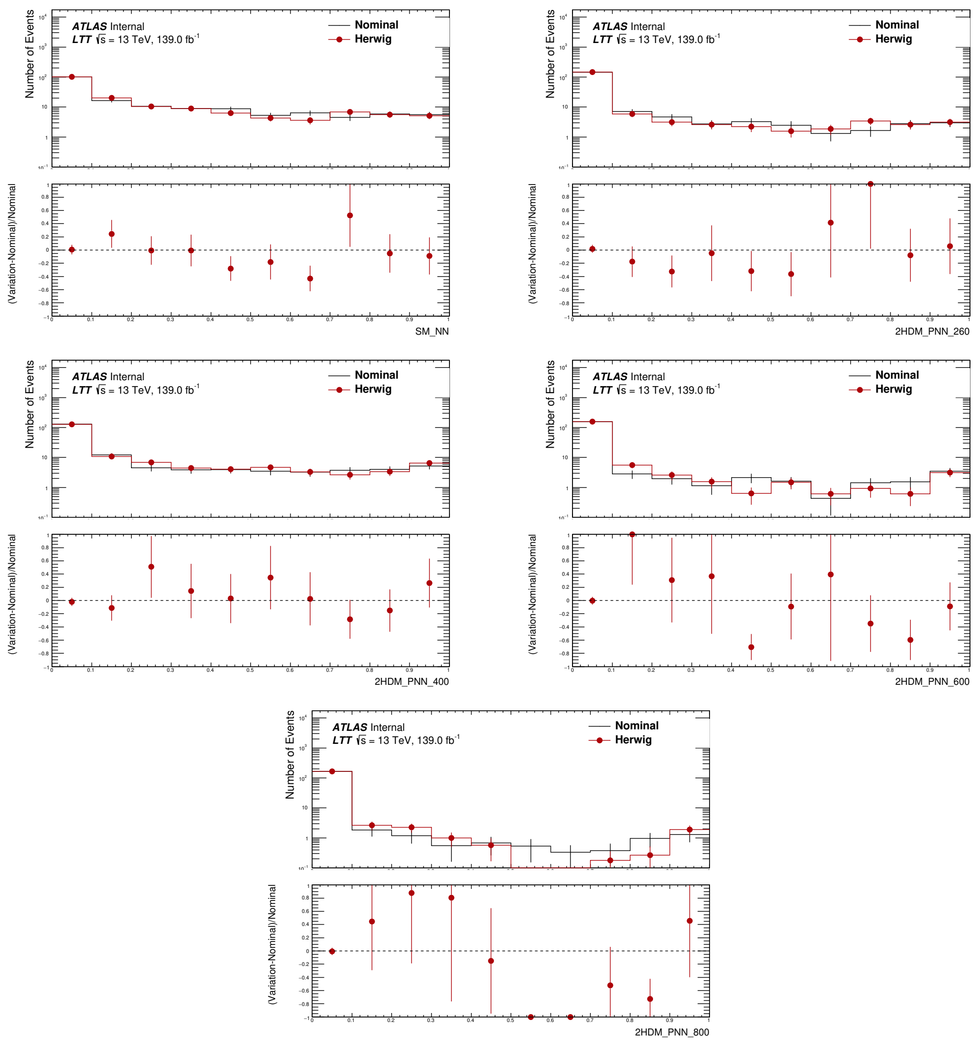


Figure B.19: LTT channel: shape-only PS uncertainty in the NN and PNN score of various resonant masses.

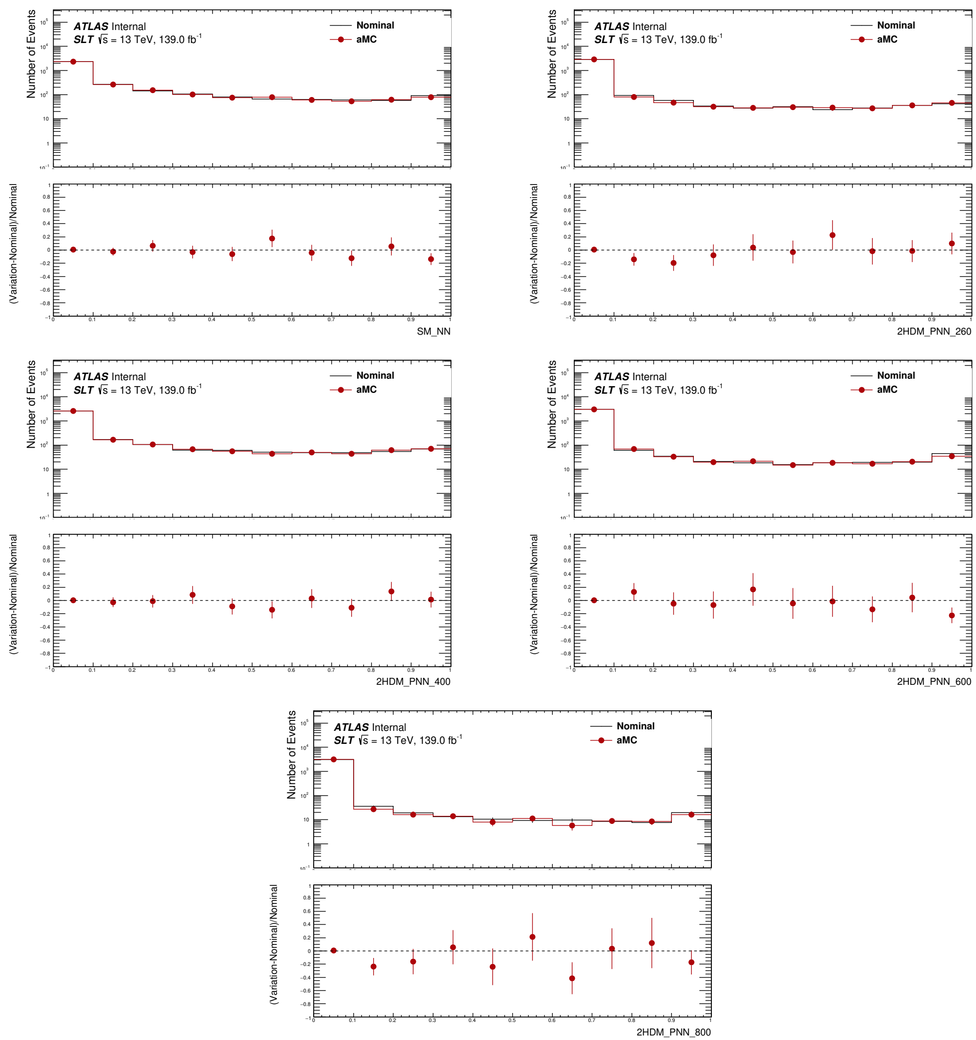


Figure B.20: SLT channel: shape-only ME uncertainty in the NN and PNN score of various resonant masses.

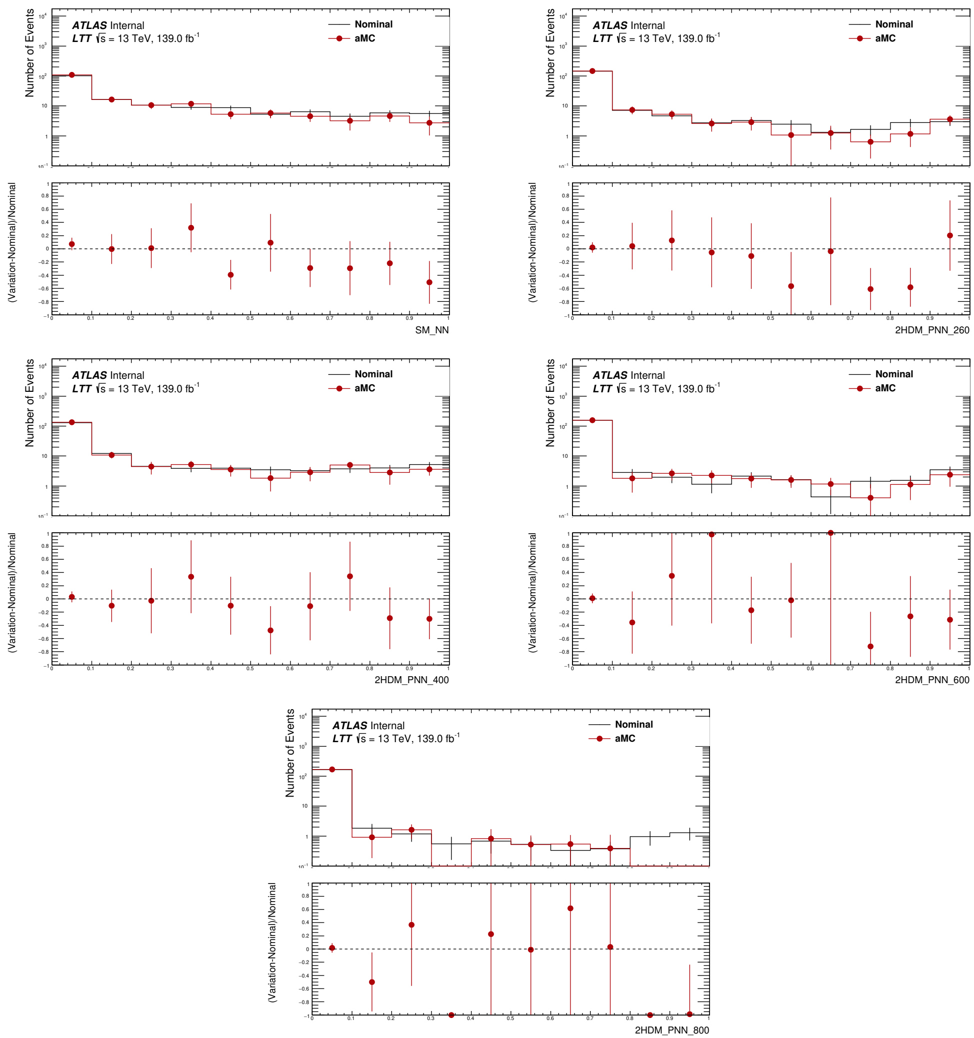


Figure B.21: LTT channel: shape-only ME uncertainty in the NN and PNN score of various resonant masses.

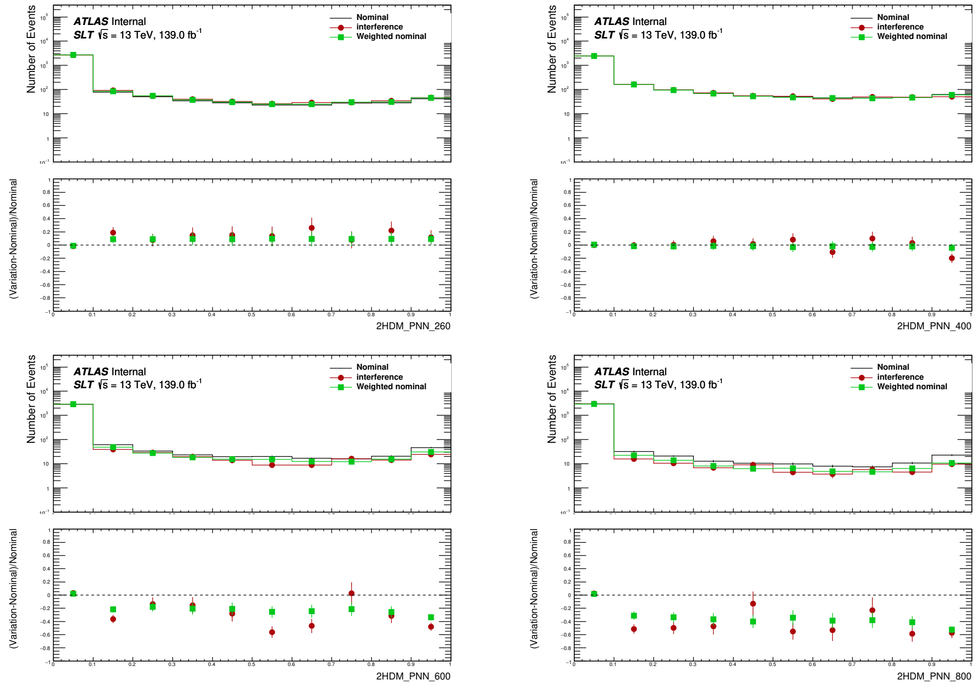


Figure B.22: SLT channel: shape-only interference uncertainty in the PNN score of various resonant masses.

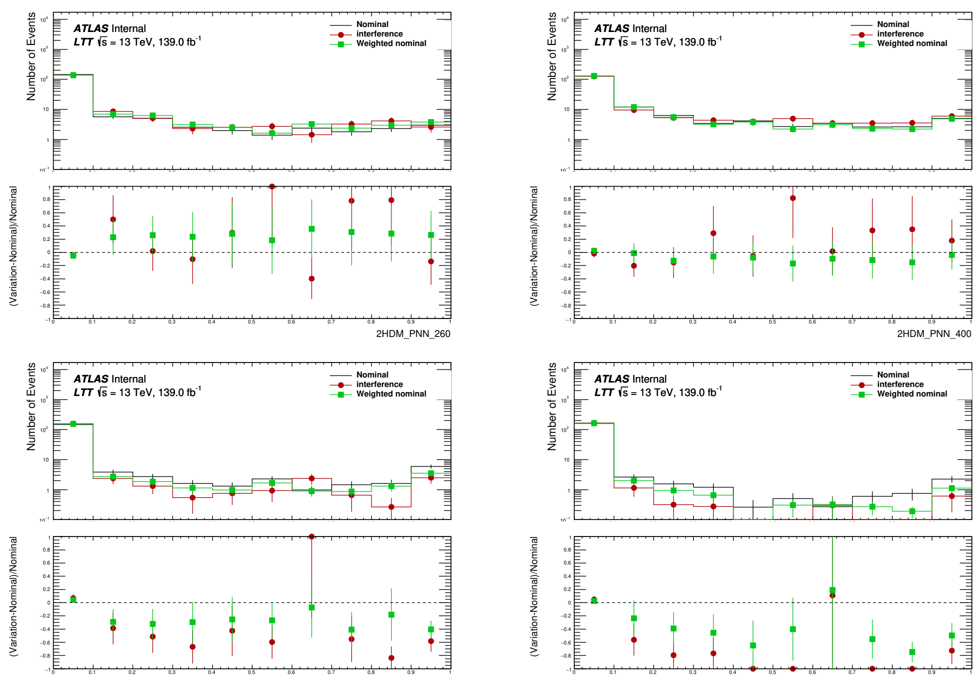


Figure B.23: LTT channel: shape-only interference uncertainty in the PNN score of various resonant masses.

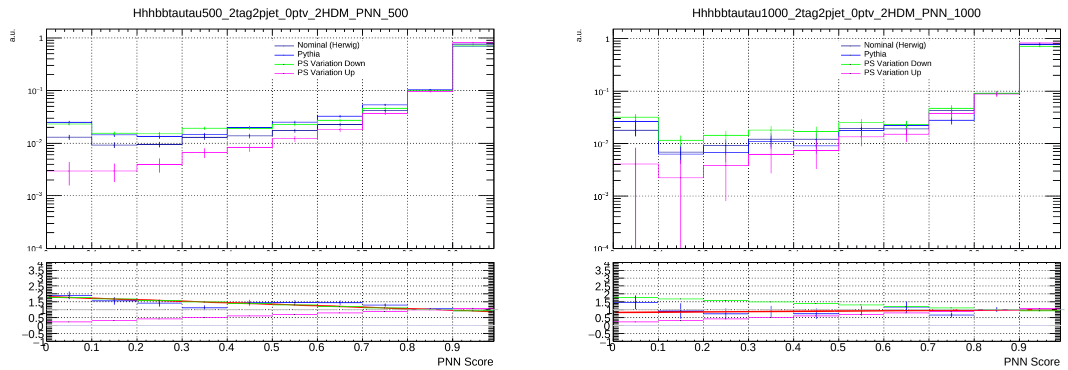


Figure B.24: Comparison of the SLT signal PNN distributions obtained from the nominal (black) and alternative (blue) signal samples for the PS variations for the m_X = 500 GeV (left) and the m_X = 1000 GeV (right) mass points. A linear fit to the ratio of the two distributions is performed and shown in the lower panel of the figures (red line). The variations obtained from the linear function obtained from the m_X = 500 GeV fit are also shown in the figures (green and magenta). Image reproduced from analysis internal notes.

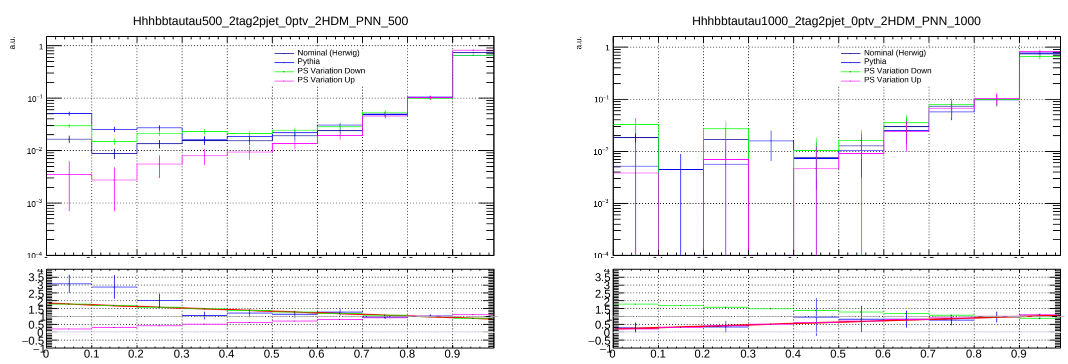
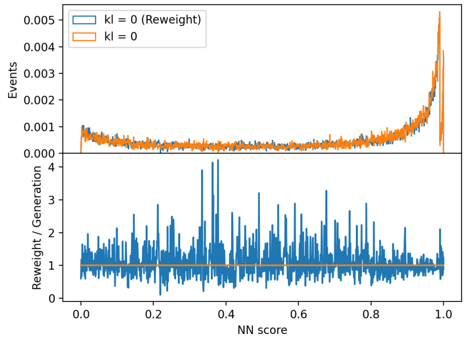
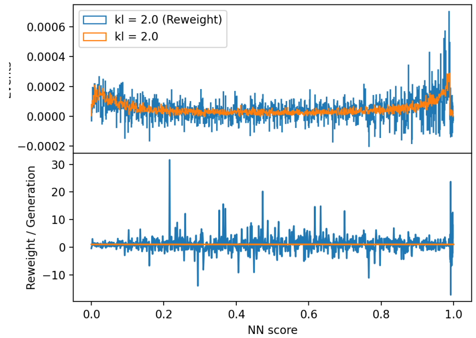


Figure B.25: Comparison of the LTT signal PNN distributions obtained from the nominal (black) and alternative (blue) signal samples for the PS variations for the m_X = 500 GeV (left) and the m_X = 1000 GeV (right) mass points. A linear fit to the ratio of the two distributions is performed and shown in the lower panel of the figures (red line). The variations obtained from the linear function obtained from the m_X = 500 GeV fit are also shown in the figures (green and magenta). Image reproduced from analysis internal notes.



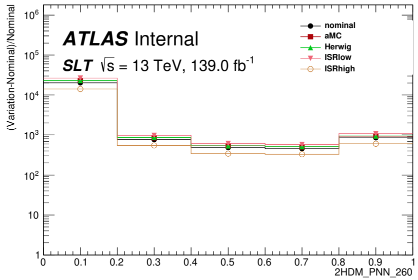
(a) $\kappa_\lambda = 0$



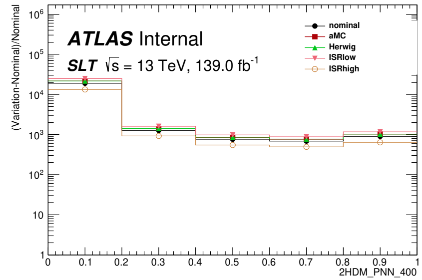
(b) $\kappa_\lambda = 2$

Figure B.26: Closure plots for the linear combination of VBF samples for two different bases in the $\tau_{\rm lep}\tau_{\rm had}$ channel. NN score distributions for the MC $\kappa_\lambda = 0, 2$ samples (orange) and the reweighted ones (blue) are shown together with their ratio. The largest difference between the two distributions is of 2.2% (SLT) and stems from the linear combination of VBF samples ($\kappa_\lambda = 1, 10, 0$) to $\kappa_\lambda = 2$. Figure reproduced from analysis internal notes.

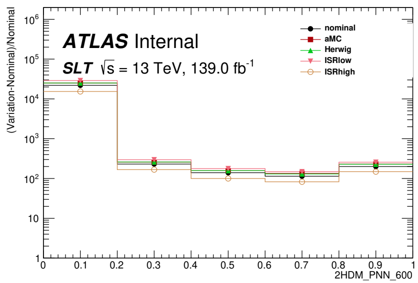
The data driven fakes estimated with $t\bar{t}$ variation samples are shown in Fig.B.27 for the SLT channel and Fig.B.28 for the LTT channel. In Fig.B.29 and Fig.B.30, the y-axis shows the rate of differences between the fakes estimated with $t\bar{t}$ variation samples and the nominal $t\bar{t}$ sample, to the nominal $t\bar{t}$ sample. The nominal $t\bar{t}$ sample is fast simulation sample and it's compared to the fast simulation variation $t\bar{t}$ samples. These variation samples includes the uncertainties from the matrix elements, parton shower selection and the initial state radiation. The final state radiation radiation on the other hand is estimated with the full simulation $t\bar{t}$ sample internal weights, as shown in Fig.B.31 for the SLT channel and Fig.B.32 for the LTT channel. Similarly the ratios are shwon in Fig.B.33 for the SLT channel and Fig.B.34 for the LTT channel.



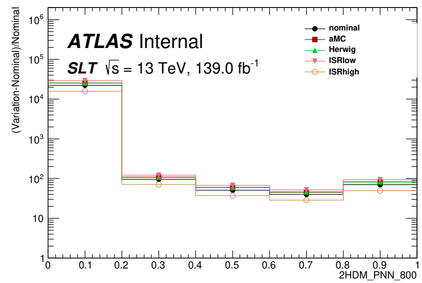
(a)



(b)

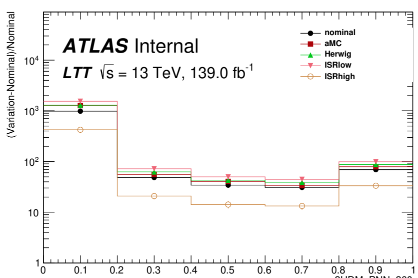


(c)

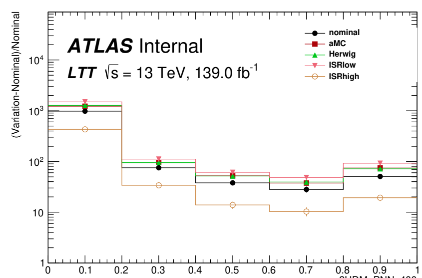


(d)

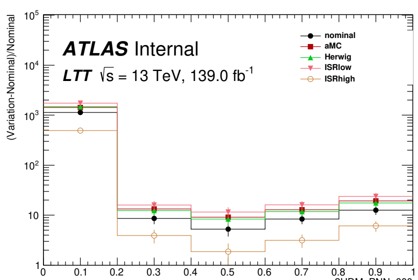
Figure B.27: SLT channel: data driven fakes estimated with the nominal $t\bar{t}$ sample and the variation $t\bar{t}$ sample.



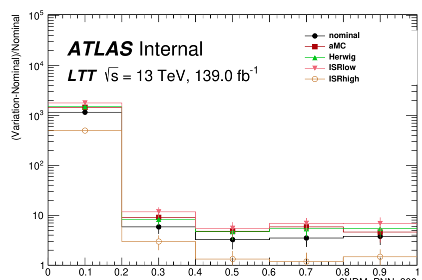
(a)



(b)

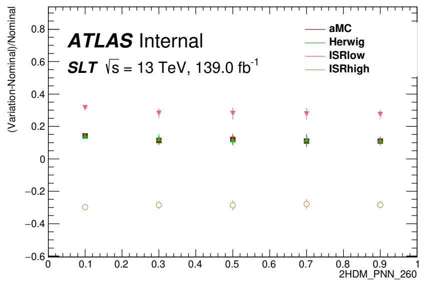


(c)

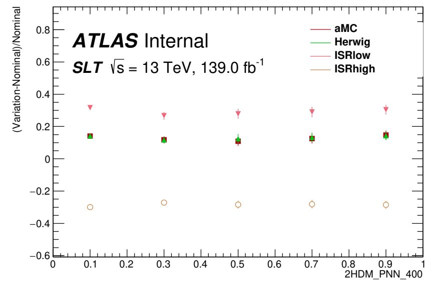


(d)

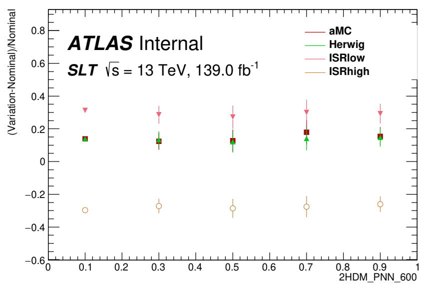
Figure B.28: LTT channel: data driven fakes estimated with the nominal $t\bar{t}$ sample and the variation $t\bar{t}$ sample.



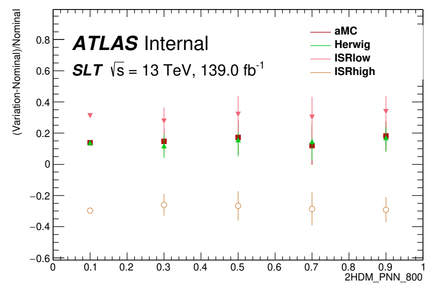
(a)



(b)

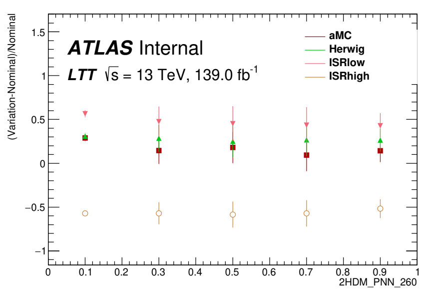


(c)

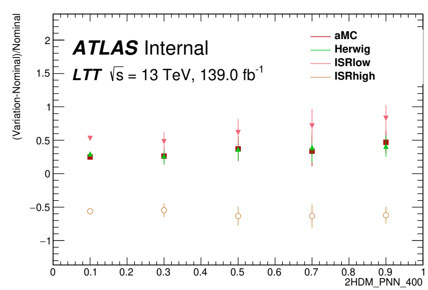


(d)

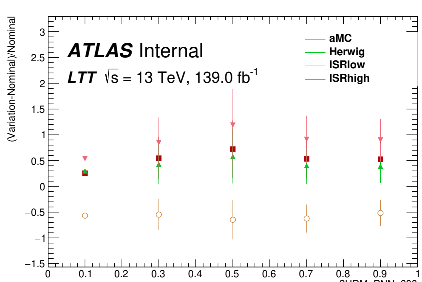
Figure B.29: SLT channel: data driven fakes estimated with the nominal $t\bar{t}$ sample and the variation $t\bar{t}$ sample.



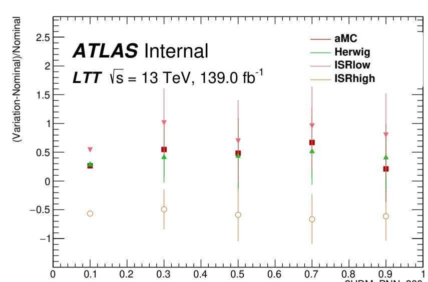
(a)



(b)

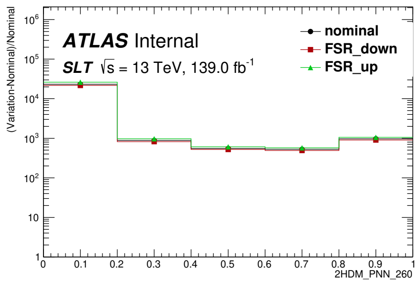


(c)

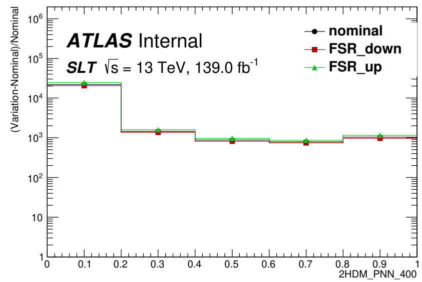


(d)

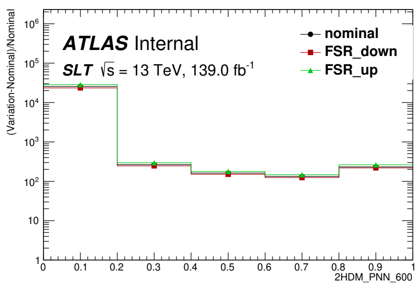
Figure B.30: LTT channel: data driven fakes estimated with the nominal $t\bar{t}$ sample and the variation $t\bar{t}$ sample.



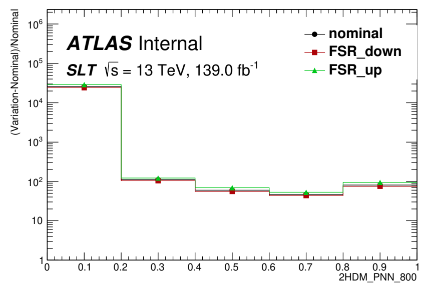
(a)



(b)

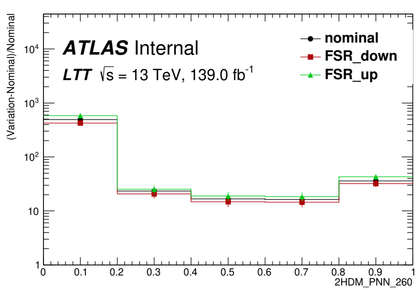


(c)

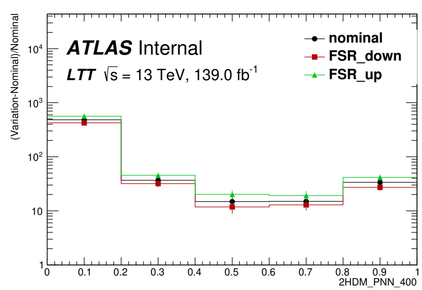


(d)

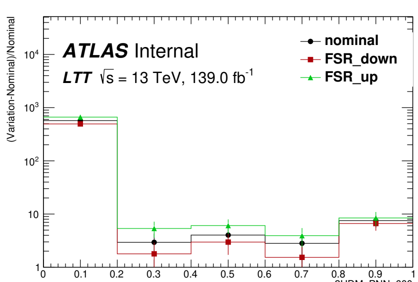
Figure B.31: SLT channel: data driven fakes estimated with the nominal $t\bar{t}$ sample and the variation $t\bar{t}$ sample.



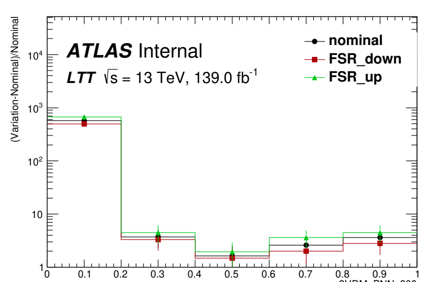
(a)



(b)

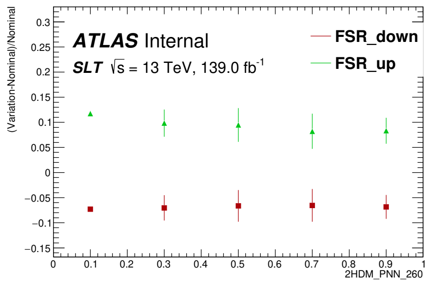


(c)

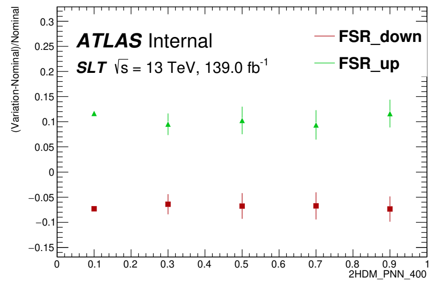


(d)

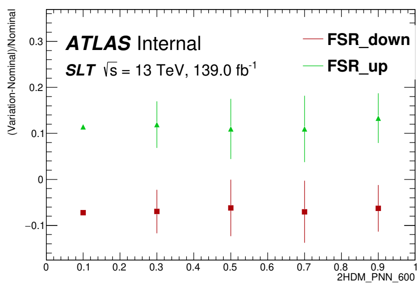
Figure B.32: LTT channel: data driven fakes estimated with the nominal $t\bar{t}$ sample and the variation $t\bar{t}$ sample.



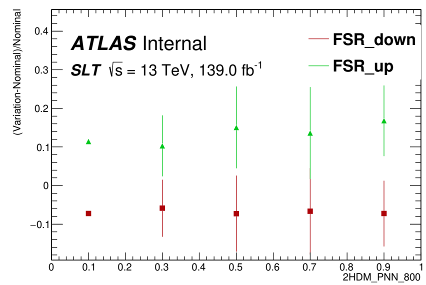
(a)



(b)

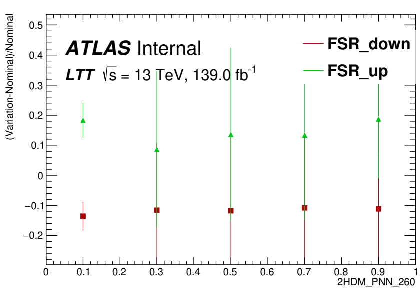


(c)

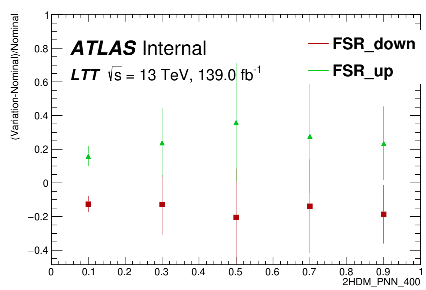


(d)

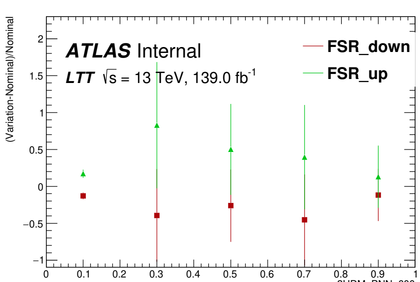
Figure B.33: SLT channel: data driven fakes estimated with the nominal $t\bar{t}$ sample and the variation $t\bar{t}$ sample.



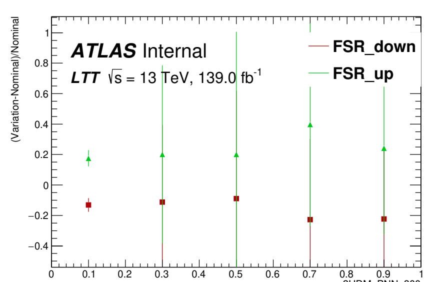
(a)



(b)



(c)



(d)

Figure B.34: LTT channel: data driven fakes estimated with the nominal $t\bar{t}$ sample and the variation $t\bar{t}$ sample.

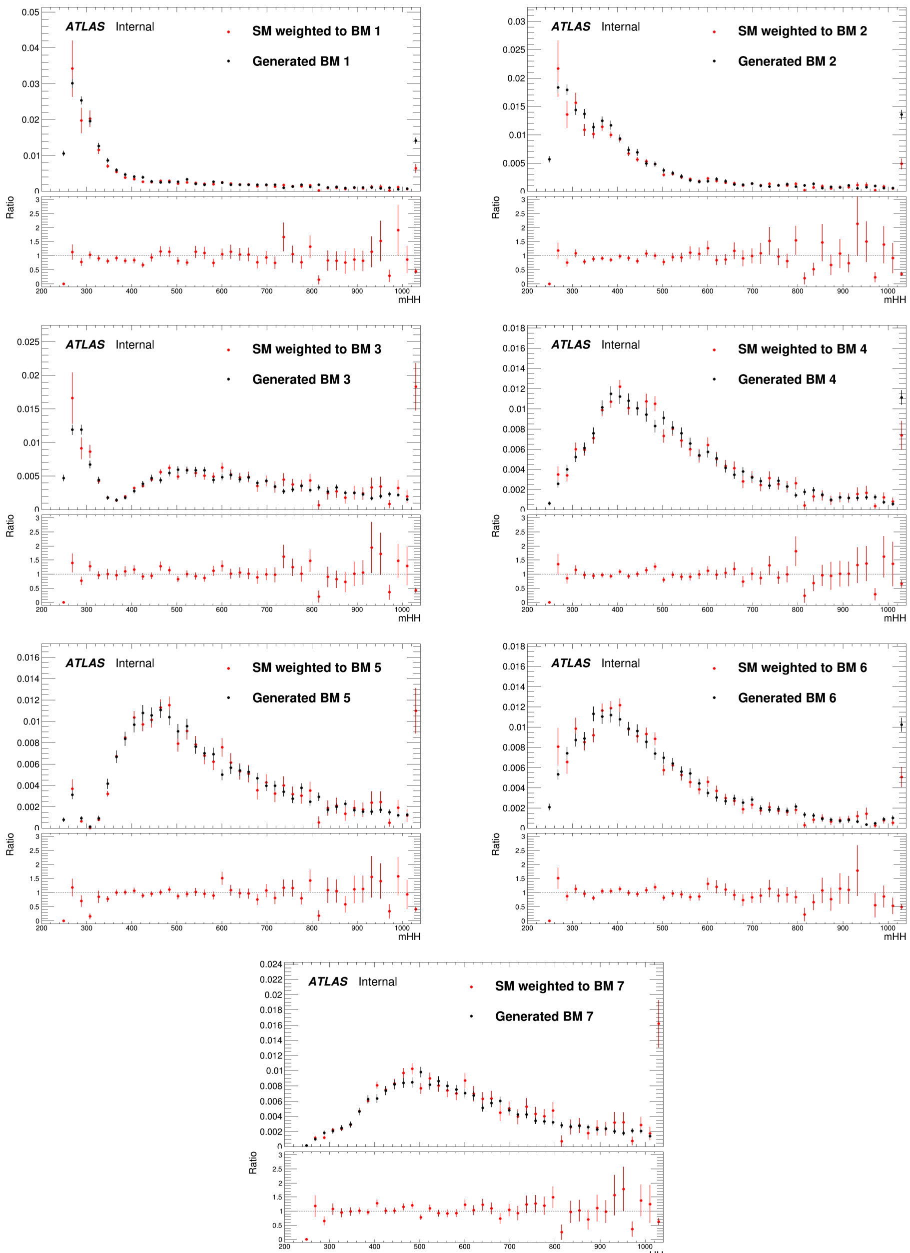


Figure B.35: SLT channel: m_{HH} distributions of the 7 HEFT BM signals, generated from simulation and reweighted from the SM non-resonant ggF siganl.

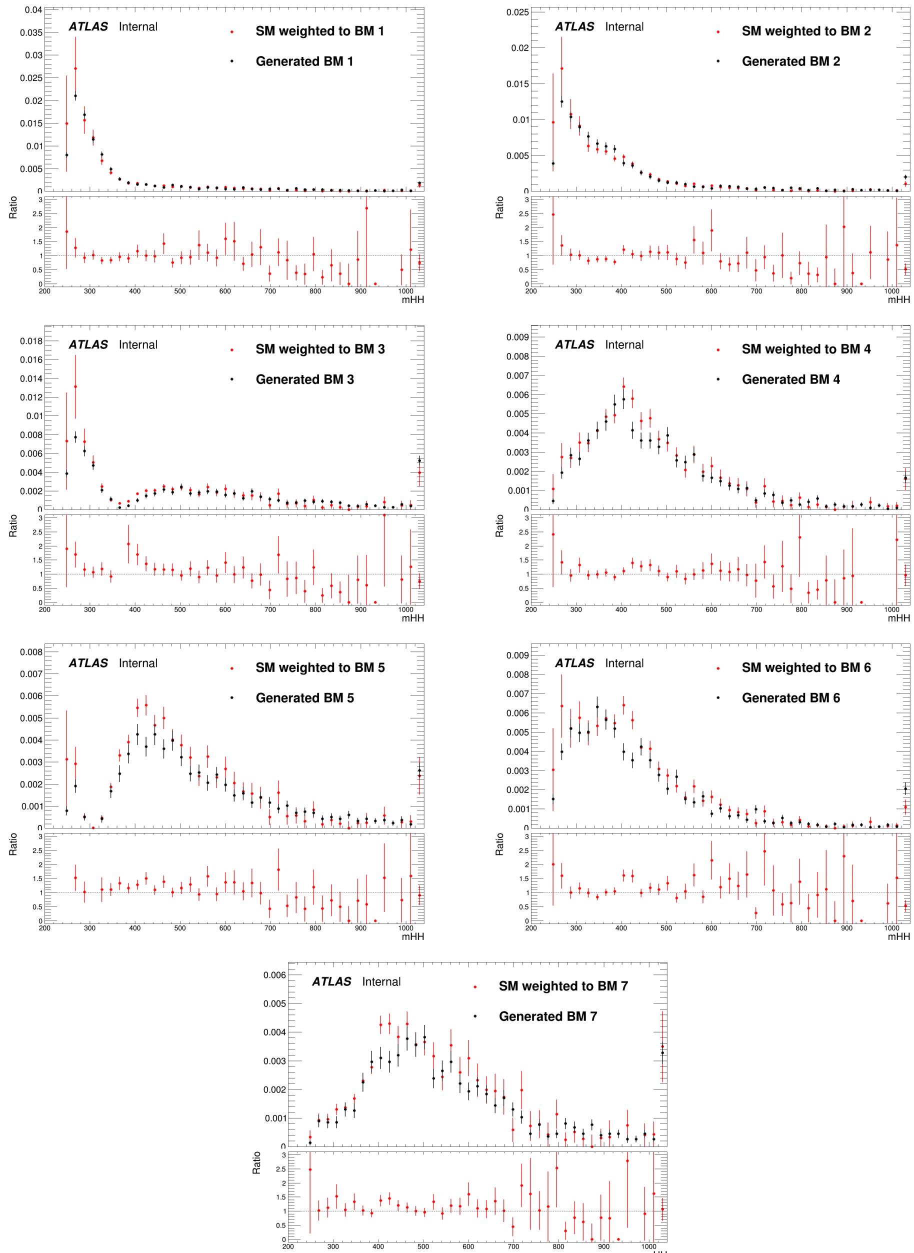


Figure B.36: LTT channel: m_{HH} distributions of the 7 HEFT BM signals, generated from simulation and reweighted from the SM non-resonant ggF siganl.

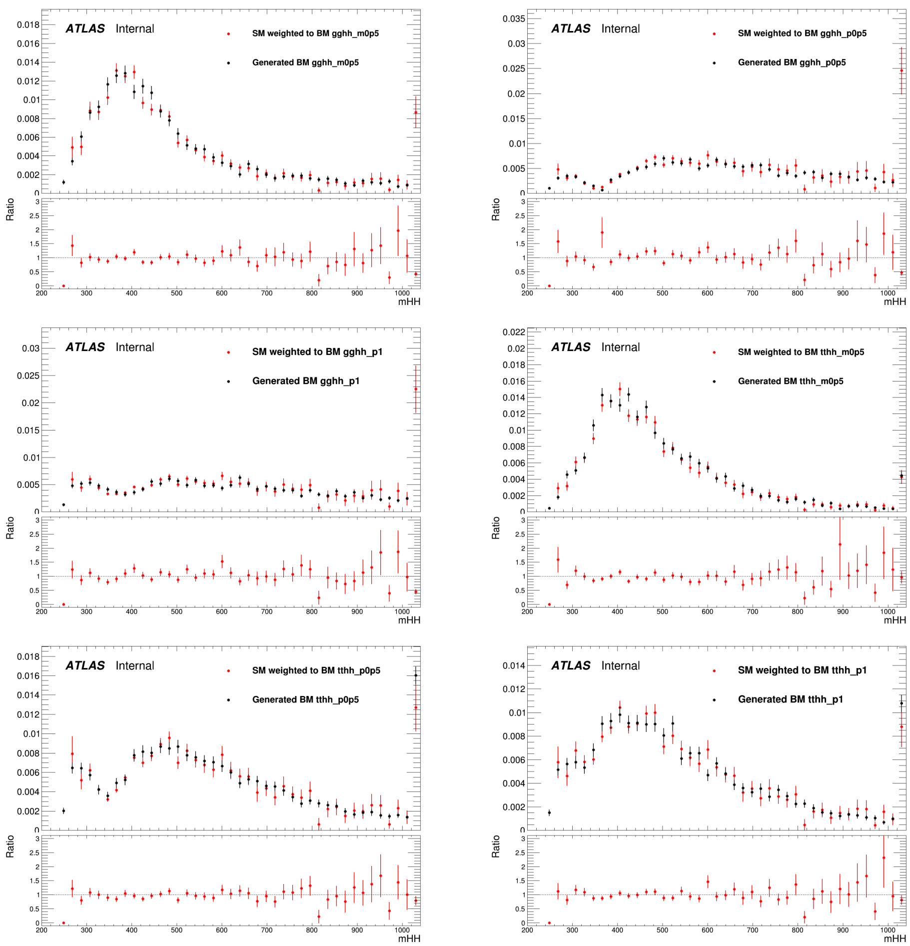


Figure B.37: SLT channel: m_{HH} distributions of the $c_{gghh} = \pm 0.5, 1.0$ ($c_{tthh} = \pm 0.5, 1.0$) signals in the top (bottom) row, generated from simulation and reweighted from the SM non-resonant ggF siganl.

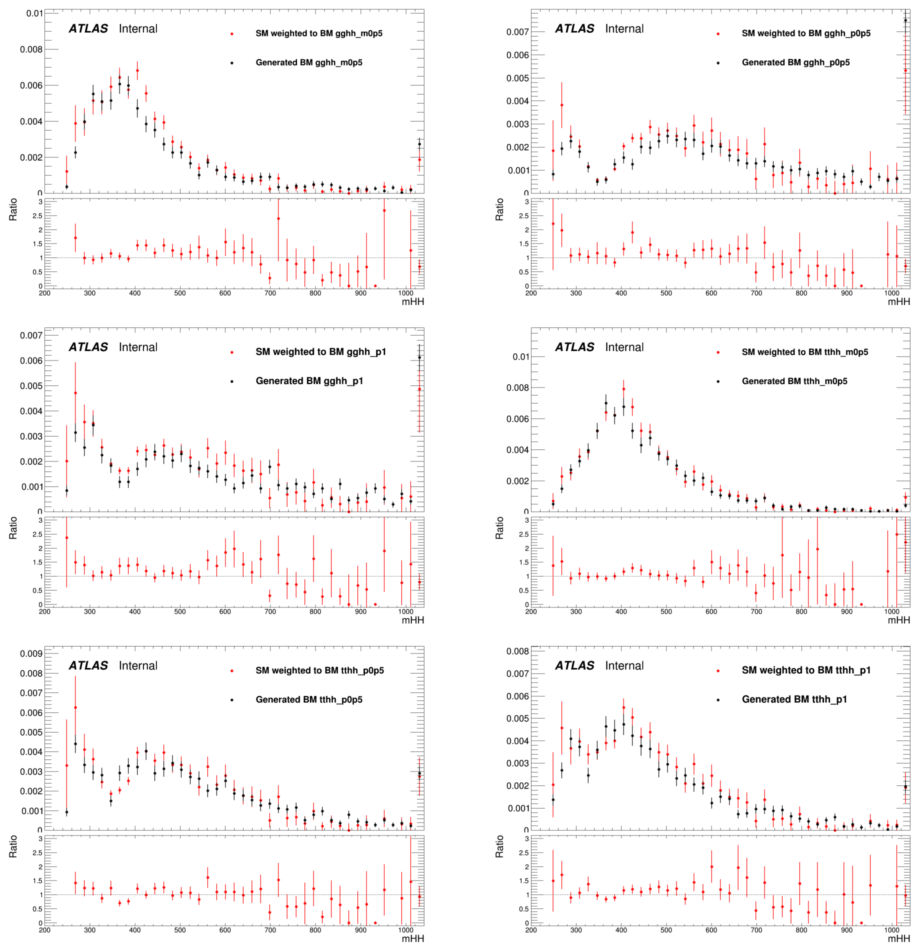


Figure B.38: LTT channel: m_{HH} distributions of the $c_{gghh} = \pm 0.5, 1.0$ ($c_{tthh} = \pm 0.5, 1.0$) signals in the top (bottom) row, generated from simulation and reweighted from the SM non-resonant ggF siganl.

B.4 Additional material for results

Figures B.39- B.43 show the NP rankings for the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$-fit (SLT + LTT) to data.

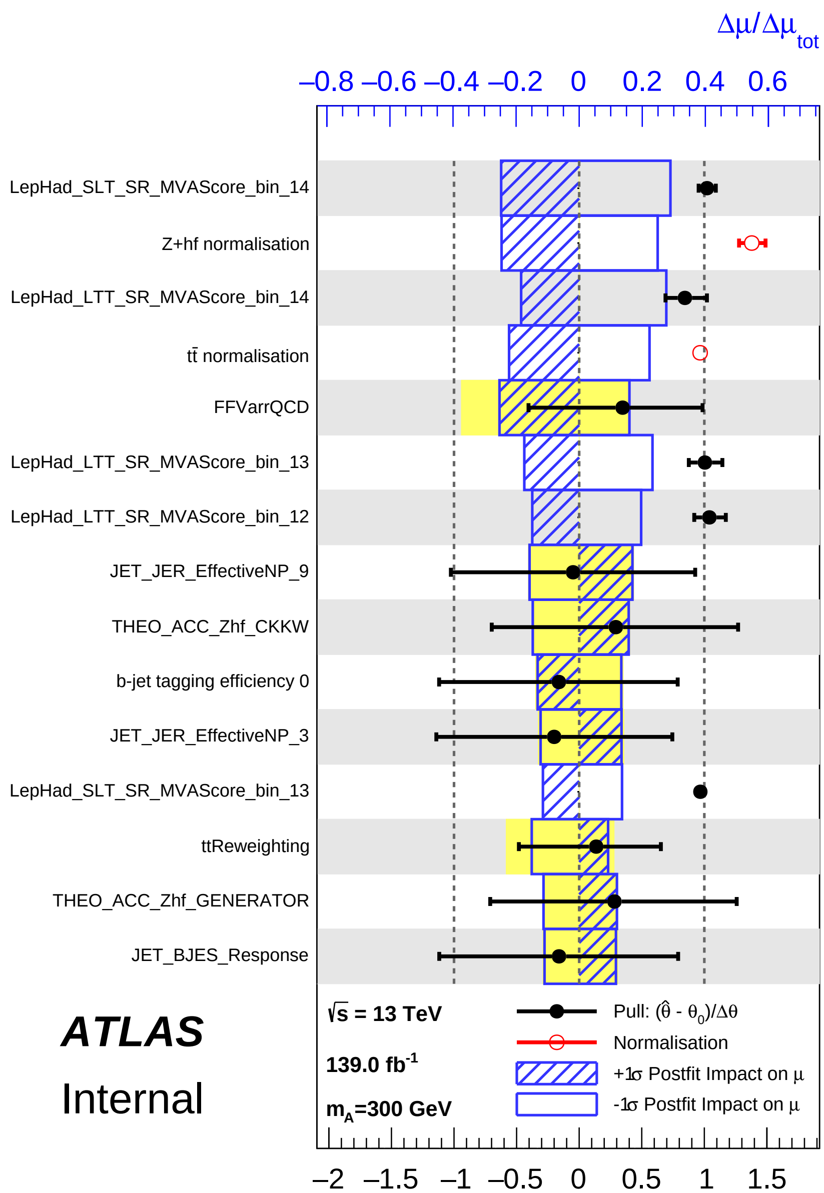


Figure B.39: NP rankings for the di-Higgs $\tau_{\text{lep}}\tau_{\text{had}}$-fit to data for the 300 GeV resonant fit.

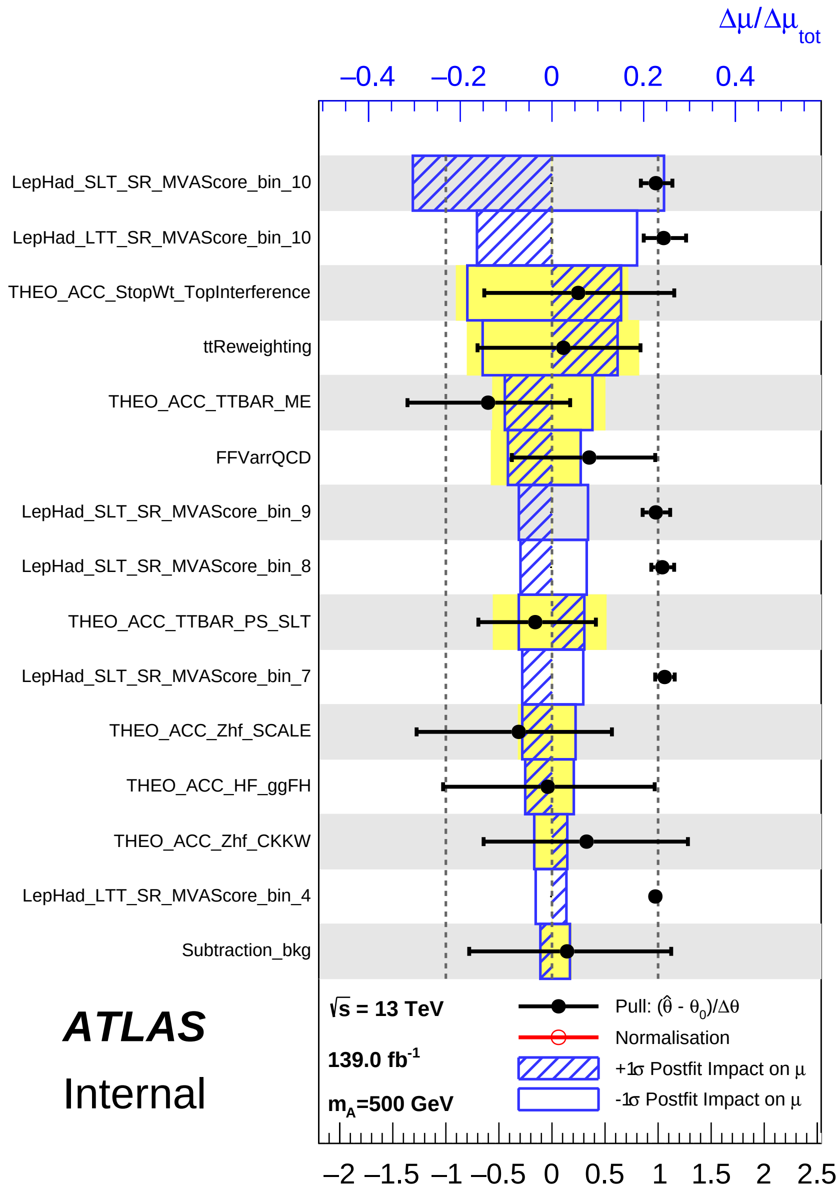


Figure B.40: NP rankings for the di-Higgs $\tau_{\text{lep}}\tau_{\text{had}}$-fit to data for the 500 GeV resonant fit.

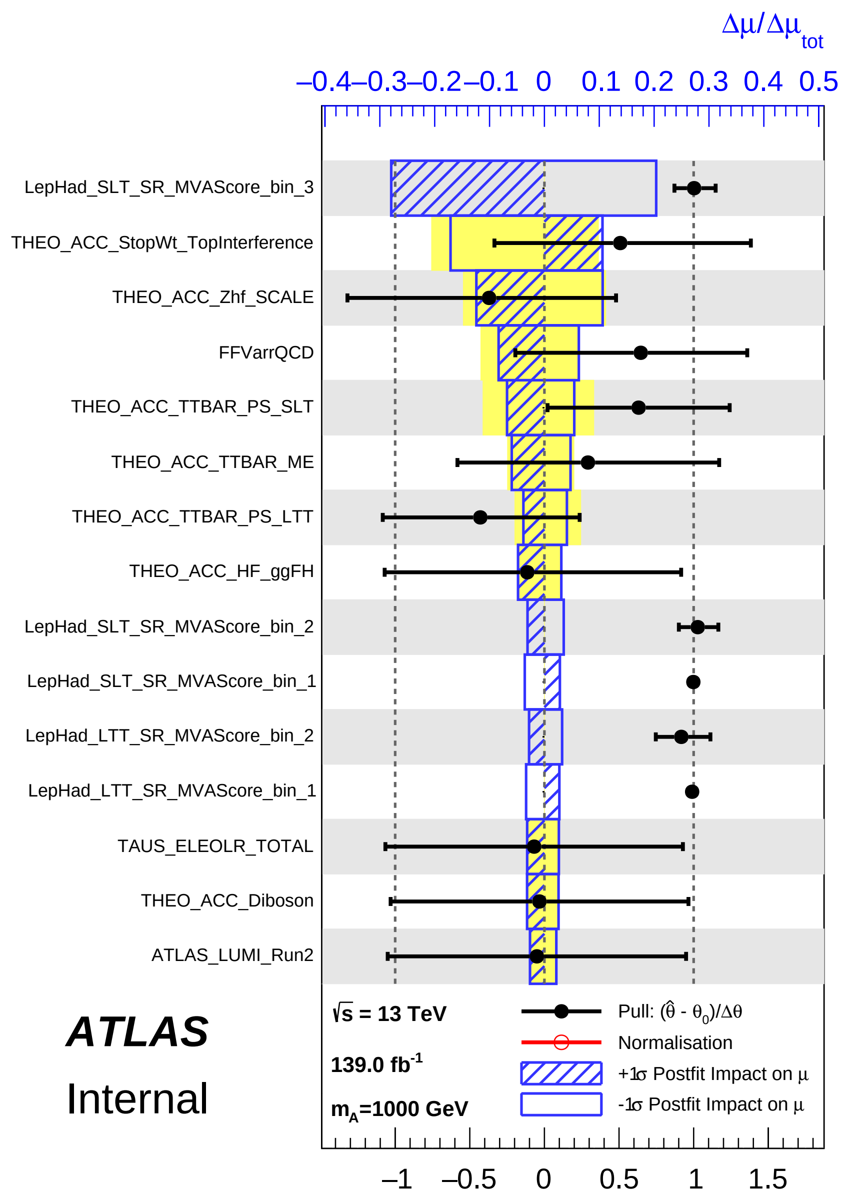


Figure B.41: NP rankings for the di-Higgs $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$-fit to data for the 1000 GeV resonant fit.

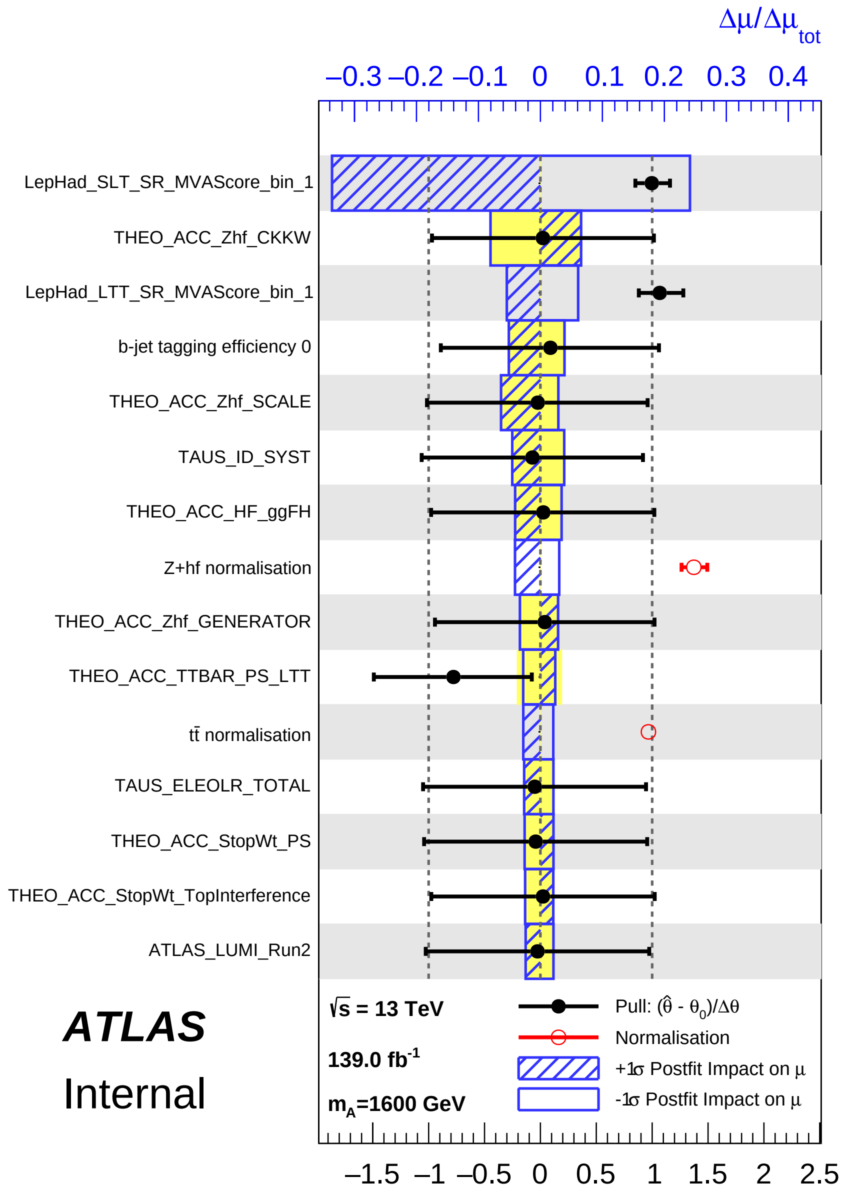


Figure B.42: NP rankings for the di-Higgs $\tau_{\text{lep}}\tau_{\text{had}}$-fit to data for the 1600 GeV resonant fit.

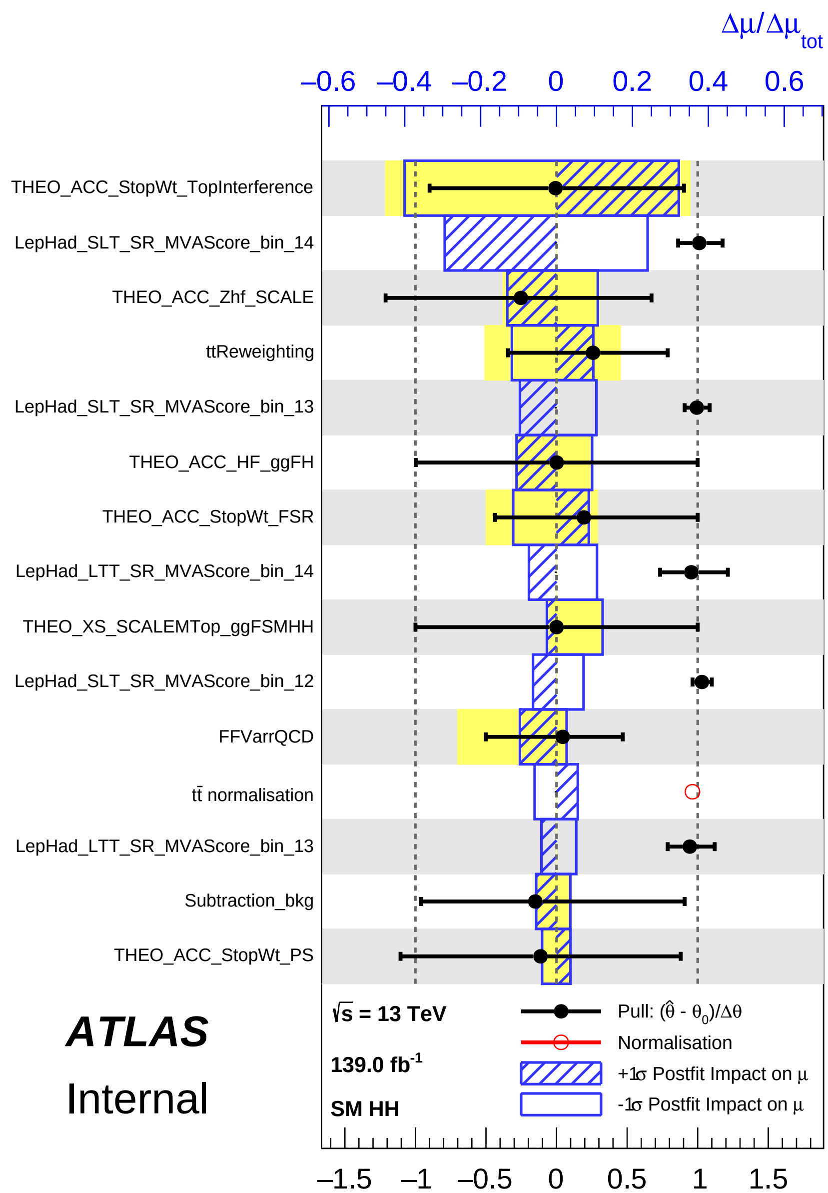


Figure B.43: NP rankings for the di-Higgs $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$-fit to data for the SM fit.

Acronyms

2HDM Two-Higgs-doublet Model 29

AF2 Fast Simulation 146

ALICE A Large Ion Collider Experiment 42

ANN/NN Artificial Neural Networks 35

ATLAS A Toroidal LHC ApparatuS 89

BDT Boosted Decision Tree 35

BM Benchmark 33

BSM Beyond the Standard Model 4

CKM Cabibbo-Kobayashi-Maskawa 13

CMS Compact Muon Solenoid 4

CSC Cathode Strip Chambers 60

CTP Central Trigger Processor 60

DM Dark Matter 4

DR Diagram Removal 155

DS Diagram Subtraction 155

EFT Effective Fields Theory 31

EM Electromagnetic 49

FCal Forward Calorimeter 56

FF Fake Factor 119

FS Full ATLAS detector simulation 81

FSR Final State Radiation 80

ggF Gluon-Gluon Fusion 4, 21

GUT Grand Unified Theory 29

HCAL Tile Hadronic calorimeter 58

HEC Hadronic End-cap Calorimeter 58

HEFT Higgs Effective Fields Theory 1

HL-LHC High Luminosity LHC 4

HLT High-Level Trigger 60

IBL Insertable B-Layer 54

ID Inner Detector 49

ISR Initial State Radiation 80

JER Jet Energy Resolution 71

JES Jet Energy Scale 71

L1 Hardware-based level 110

LAr Liquid Argon 57

LCW Local Cluster Weighting 71

LEP Large Electron-Positron Collider 43

LH Left-Handed 13

LHCb Large Hadron Collider beauty 42

LO Leading Order 79

LS Long Shutdown 46

LTT Lepton-plus-Tau Trigger 110

MC Monte Carlo 79

MDT Monitored Drift Tubes 60

ME Matrix Element 147

ML Machine Learning 34

MS Muon Spectrometer 49

MVA Multivariate Techniques 130

NEG Non-Evaporable Getter 43

NLL Next-to-Leading-Logarithm 87

NLO Next-to-Leading Order 33

NNLL Next-to-Next-to-Leading-Logarithm 86

NNLO Next-to-Next-to-Leading Order 25

NP Nuisance Parameter 169

PDF Parton Distribution Function 79

PMNS Pontecorvo-Maki-Nakagawa-Sakata 29

PNN Parametric Neural Network 134

POI Parameter Of Interest 168

PS Parton Shower 147

PSB Proton Synchrotron Booster 42

QCD Quantum Chromodynamics 7

QED Quantum Electrodynamics 7

QFT Quantum Field Theory 8

ReLU Rectified Linear Unit 36

RH Right-Handed 13

RNN Recurrent Neural Network 75

RNNIP Recurrent Neural Network Impact Parameter tagger 73

RPC Resistive Plate Chambers 59

SCT Semiconductor Tracker 49

SET Single-Electron Trigger 111

SLT Single-Lepton Trigger 110

SM the Standard Model of Particle Physics 4

SMT single-Muon Trigger 111

SPS Super Proton Synchrotron 42

SR Signal Region 112

tau-ID τ_{had} identification 75

TGC Thin Gap Chambers 59

TRT Transition Radiation Tracker 49

ttH Higgs boson production associated with a top-anti-top-quark pair 21

VBF Vector Boson Fusion 21

VH Higgs boson production associated with a vector boson 21

WP Working Point 65