GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Master's Thesis

Searches for resonant Higgs boson pair production in the $b\bar{b}WW^{(*)}$ decay channel for the boosted single-lepton final state at ATLAS

Suche nach resonanter Higgs-Boson-Paarproduktion im $b\bar{b}WW^{(*)}$-Zerfallskanal für den „boosted" Einzellepton-Endzustand bei ATLAS

prepared by

Naman Kumar Bhalla

from New Delhi, India

at the II. Institute of Physics

Thesis number:	II.Physik-UniGö-MSc-2021/03
Thesis period:	12th October 2020 until 11th October 2021
First referee:	Prof. Dr. Stan Lai
Second referee:	Prof. Dr. Ariane Frey

Abstract

This thesis presents searches for a heavy scalar resonance (X) decaying into a pair of scalar bosons in the $b\bar{b}WW^{(*)}$ decay channel for the single-lepton final state and boosted topologies using the full Run 2 ATLAS proton-proton collision dataset. The scalar boson pair can either be two Standard Model Higgs bosons (HH) or an additional heavy scalar in conjunction with a Standard Model Higgs boson (SH). Due to the interplay between the masses of the scalars, various topologies, extending from fully resolved to fully boosted, are possible. Each of these topologies have their possibilities and challenges. Thus, having orthogonal cuts to define regions, where one or the other topology dominates is crucial to optimise each topology independently. Moreover, the definitions of signal, validation and control regions play a key role in any analysis.

A statistical framework, based on profile likelihood fits, is used to set expected 95% C.L. exclusion limits on the HH production cross section, $\sigma(pp \to X \to HH)$, and on the cross section of SH production times its branching ratio to the $b\bar{b}WW$ decay channel, $\sigma(pp \to X \to SH) \times \mathcal{BR}(SH \to b\bar{b}WW)$, for the HH and SH scenarios, respectively. It is also used to perform an iterative procedure of simultaneous single-bin fits to constrain the leading backgrounds in the control regions. The expected 95% C.L. exclusion limits are set by fitting the shape of the sum of the mass of the visible decay products of X and the missing transverse energy ($m_{\text{vis+met}}$) in the blinded signal regions, with only the statistical uncertainties included. For the HH samples, the best limit is expected for $m_X = 5.0\,\text{TeV}$ at 2.8 fb, which is in the same order of magnitude as the limits obtained by the boosted 0-lepton final state of the $b\bar{b}WW^{(*)}$ analysis and by the boosted $b\bar{b}b\bar{b}$ analysis. For the SH samples, the best limit is expected for $m_X = 3.0\,\text{TeV}$ and $m_S = 240\,\text{GeV}$ at 0.87 fb.

Keywords: Heavy scalar resonances, Higgs boson pair production, $b\bar{b}WW^{(*)}$ decay channel, single-lepton final state, boosted topologies

Contents

1. Introduction

"Science is not everything, but science is very beautiful."
- J. Robert Oppenheimer

The question of what matter is made up of is one that has engaged human civilisations for millennia. The ancient Greeks tried to answer this with earth, water, air and fire as the four elements of matter, while the Hindus also had an additional element of space in Ayurveda, calling them *Panchamahabhuta* (Sanskrit for "five great elements"). With time, the concepts of an indivisible atom with the early periodic tables, splitting the atom into electrons and a nucleus, the nucleus further into protons and neutrons, and protons and neutrons into quarks have been introduced. Currently, quarks and leptons are believed to be the elemental components of matter with the Standard Model (SM) of particle physics being the most successful model at explaining matter and its interactions at the most fundamental level.

The SM predicts a total of 17 fundamental particles (12 fermions and 5 bosons), each having a corresponding antiparticle. It explains three out of the four fundamental forces of nature. The electromagnetic and the weak interactions are unified by the electroweak theory [1–3], while the strong interaction is described in the theory of quantum chromodynamics (QCD) [4, 5]. The SM compliments these two theories with the Higgs mechanism [6–9]. The experimental discovery of the Higgs boson (H) in 2012 by the Atlas and Cms experiments [10, 11] at the Large Hadron Collider (Lhc) at Cern acted as the final piece of the puzzle in the search for SM particles.

Even with all its merits, it is known that the SM is not the final word on the subject due to multiple reasons. The biggest shortcoming of the SM is that it is not a complete theory. The SM is unable to explain the fourth fundamental force of nature, i.e. the gravitational force, which is known to play an important role at higher energy regimes. In addition to this, there are also multiple open questions that the SM is unable to answer. These include, but are not restricted to, questions about dark matter [12–15], dark energy [16–18], neutrino oscillations [19–21] and baryon asymmetry [22, 23].

To answer these questions a plethora of theories beyond the Standard Model (BSM)

have been proposed, which try to preserve the experimentally verified aspects of the SM, while suggesting certain extensions to it in order to resolve some of its deficiencies. Many of these BSM theories focus on extending the Higgs sector by inserting additional terms to the SM Lagrangian. Some examples of these are the addition of a real or complex singlet field in the RxSM and CxSM, respectively [24], the Two-Real-Scalar-Singlet extension of the SM (TRSM) [25] and the addition of a complex doublet field in the Two-Higgs-Doublet Model (2HDM) [26]. As a consequence of these additional fields, these theories predict the existence of additional scalar bosons.

This thesis presents the searches conducted for heavy scalar resonances (X) with a possible decay into either a pair of Higgs bosons or an additional heavy scalar (S) in conjunction with a Higgs bosons using the full Run 2 ATLAS data. The case where one H decays into a $b\bar{b}$ pair, while the other scalar boson (H or S) decays into a pair of W bosons is considered in this thesis and focusses primarily on the scenario, where one W boson decays leptonically and the other decays hadronically. The scalar X is assumed to have sufficient mass for the decay particles to be boosted, such that the final state particles cannot be fully resolved in the detector.

The thesis is structured as follows: Chapter 2 presents the theoretical aspects of the thesis, which comprises of a detailed description of the SM and the Higgs mechanism, followed by some of the relevant BSM scenarios, the main production modes of the Higgs boson along with its possible decays, and the theory of profile likelihood fits. Chapter 3 describes the experimental setup consisting of the LHC and the ATLAS detector. Monte-Carlo (MC) event generators and a description of the treatment of jets in boosted topologies is also included. Chapter 4 portraits the analysis strategy, starting with the description of the $b\bar{b}WW^{(*)}$ event topology, followed by the object definitions and the event selection. Chapter 5 illustrates the attempts made in order to optimise the event selection for the SH signal samples. This includes the challenges posed by the various SH topologies, the definitions and efficiencies of various classifier techniques that were studied and also the reasons behind requiring separate boosted and split-boosted regions. Next, the description of the statistical framework and the statistical interpretation is presented in Chapter 6 including the background normalisation fits and exclusion fits. This chapter also includes a discussion about the expected exclusion limits with only the statistical uncertainties along with the effect of including preliminary systematic uncertainties. The thesis is summarised with a few words on the outlook in Chapter 7.

2. Theory and Phenomenology

> "I would rather have questions that can't be answered than answers that can't be questioned."
>
> - Richard P. Feynman

The ultimate goal of particle physics is to understand the elemental constituents of matter along with their interactions, which, when put together, could shed light on the functioning of the universe at a larger scale. Even though the SM of particle physics is the best theoretical framework available currently, it still leaves many questions unanswered.

2.1. The Standard Model and Higgs Mechanism

This section provides a brief introduction to the SM and the Higgs mechanism. A more detailed explanation can be found in textbooks such as Refs. [27, 28] and in reviews such as Ref. [29].

The Standard Model

The SM is a quantum field theory and consists of 17 fundamental particles, which are all described as excitations in quantum fields. These include twelve spin-½ fermionic particles, which are represented by spinor fields, ψ, and make up the matter around us. The four spin-1 gauge bosons result from the gauge fields, V_μ. These are the force carriers and mediate the interactions between matter. The spin-0 Higgs boson results from the only scalar field in the SM, ϕ, and is responsible for giving mass to these fundamental particles. All the particles are graphically represented in Figure 2.1.

Starting with the force carriers, the photon mediates the electromagnetic force and interacts with particles that carry an electric charge. Next, the gluon, which is the mediator of the strong force and couples to particles possessing colour charge (analogous to electric charge for the electromagnetic force). Unlike the photon, which is electrically neutral, the gluon itself also possesses colour charge leading to gluon self-interactions.

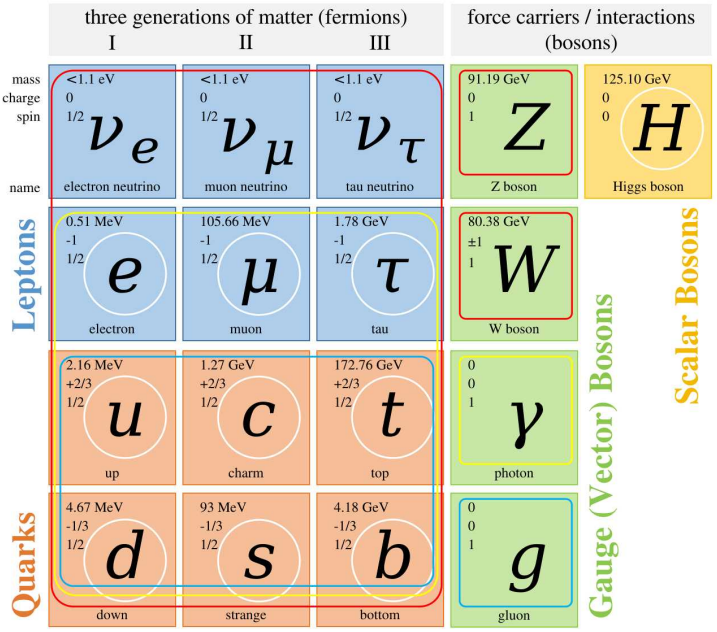


Figure 2.1.: Particle content of the SM. The red, yellow and blue outlines represent the possible fermionic couplings of the gauge bosons, while the white circles represent the possible fermionic interactions of the Higgs boson.

Since the gluon exists as a colour octet, there are a total of eight types of gluons. The W boson is responsible for charged currents of the weak interaction and interacts only with particles possessing left-handed chirality (or antiparticles with right-handed chirality) and is, thus, the origin of parity violation in the SM. While the photon and gluon are massless, the W boson is massive and provides the only vertex transformation of a fermion to a different flavour fermion. Finally, the Z boson, like the W boson is also massive and mediates the neutral currents of the weak interactions. It couples with both left- and right-handed particles.

On the other hand, the twelve fermions are further divided into two groups, quarks and leptons, based on their participation in the strong interaction. While the quarks exist as colour triplets and interact with the gluon, the leptons are colour singlets and hence do not participate in strong interactions. Both, quarks and leptons exist in three generations, with fermion masses increasing from one generation to the next. Each generation has an up-type quark with a charge of $^{+2}/_{3}$, a down-type quark possessing a charge of $^{-1}/_{3}$, a charged lepton (down-type lepton) with a charge of -1 and a neutrino (up-type lepton) which is electrically neutral. Since the weak force couples to the chirality of a particle, there is a further distinction between left- and right-handed particles. While left-handed particles are placed in weak isospin doublets, right-handed particles exist in weak isospin singlets.

The foundation of the SM was laid by Glashow, Weinberg and Salam in the 1960s with the unification of the electromagnetic and weak forces into an electroweak theory [1–3] using the $SU(2)_L \times U(1)_Y$ symmetry gauge group for this description. For this, hypercharge was introduced which is defined as

$$Y = 2Q - 2I_3, \tag{2.1}$$

where Q is the electrical charge and I_3 denotes the third component of the weak isospin. A few years later, the theory of QCD [4, 5] was also added, described by the $SU(3)_C$ gauge group, where C denotes the colour charge.

Mathematically, the interactions in QCD are represented by an octet of fields, $G_\mu^{1,\ldots,8}$. The generators of the $SU(3)_C$ gauge group here are proportional to the Gell-Mann matrices. This results in eight gluon fields with different colour states. In the electroweak sector, the B_μ field corresponds to Y, where Y is the generator of the $U(1)_Y$ gauge group. The $SU(2)_L$ gauge group relates to the three fields $W_\mu^{1,2,3}$. Here the generators are proportional to the Pauli matrices, τ^a. The mixing between the W_μ^1 and W_μ^2 fields results in the W^+ and W^- bosons, while the mixing between the W_μ^3 and B_μ fields leads to the photon, γ, and the Z boson. The field strengths of these fields are given by

$$\begin{aligned} G_{\mu\nu}^a &= \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_3 f^{abc} G_\mu^b G_\nu^c \\ W_{\mu\nu}^a &= \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g_2 \varepsilon^{abc} W_\mu^b W_\nu^c \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu. \end{aligned} \tag{2.2}$$

Due to the non-abelian nature of the $SU(3)$ and $SU(2)$ gauge groups, self-interactions between the gauge fields are required, leading to triple and quartic gauge couplings.

The interaction between gauge fields and spinor fields in then described by the covariant derivative. This is defined as

$$\mathcal{D}_\mu \psi = \left(\partial_\mu + \underbrace{i g_3 \lambda_a G_\mu^a}_{SU(3)_C} + \underbrace{i g_2 \tau_b W_\mu^b}_{SU(2)_L} + \underbrace{i g_1 Y B_\mu}_{U(1)_Y} \right) \psi, \tag{2.3}$$

where Y, τ_b and λ_a are the generators of their respective gauge groups. The second term would be zero for lepton spinor fields, since leptons are colour singlets and do not interact strongly. Exploiting the Lagrange formalism, the expression for the SM Lagrangian is

obtained as

$$\mathcal{L} = \underbrace{-\frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}}_{\text{kinematics of gauge fields}} + \underbrace{\bar{\psi}^f_L i\mathcal{D}_\mu\gamma^\mu\psi^f_L + \bar{\psi}^f_R i\mathcal{D}_\mu\gamma^\mu\psi^f_R}_{\text{kinematics and interactions of spinor fields}} \quad , \tag{2.4}$$

where γ^μ represents the Dirac matrices and ψ^f_L denotes all left-handed fermion doublets, while ψ^f_R stands for all right-handed fermion singlets.

The above Lagrangian is invariant under $SU(3)_C \times SU(2)_L \times U(1)_L$ local gauge transformations but does not include any mass terms. However, it is known that most fermions along with the W and Z bosons are massive [29]. But in the electroweak theory, apart from the issue of gauge bosons being massive, the addition of fermionic masses leads to terms like

$$-m\left(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R\right). \tag{2.5}$$

These are not invariant under $SU(2)_L \times U(1)_Y$ local gauge transformations, since ψ_R is a weak isospin singlet while ψ_L is a weak isospin doublet. Therefore, a mechanism to generate masses without violating the $SU(2)_L \times U(1)_Y$ local gauge invariance is needed. This must also break the symmetry between the weak and electromagnetic forces, since only the W and Z bosons are observed to be massive, while the photon is seen to be massless. This is known as electroweak symmetry breaking.

The Higgs Mechanism

In the 1960's, a simultaneous solution was proposed to the problem of gauge boson masses by Brout, Englert, Higgs, Guralnik, Hagen and Kibble [6–9] which is popularly known as the Brout-Englert-Higgs (BEH) mechanism or simply the Higgs mechanism. This introduces a field, Φ, as a complex scalar doublet

$$\Phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2\\ \phi_3 + i\phi_4\end{pmatrix}. \tag{2.6}$$

The Lagrangian resulting from this field in the most general form can be written as

$$\mathcal{L}_H = \left(\partial_\mu\Phi\right)^\dagger\partial^\mu\Phi - V\left(\Phi\right) \quad \text{with} \quad V\left(\Phi\right) = \mu^2\Phi^\dagger\Phi + \lambda\left(\Phi^\dagger\Phi\right)^2. \tag{2.7}$$

$V\left(\Phi\right)$ is known as the Higgs potential and has a total of two free parameters, μ^2 and λ. By considering positive values for λ to bound the potential from below and a negative

value of μ^2, the potential has an infinite set of non-zero degenerate minima at

$$\Phi^\dagger\Phi = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}, \tag{2.8}$$

where v is known as the vacuum expectation value (VEV) of Φ. In order to attain a more stable state, any one of these degenerate points is spontaneously chosen by the vacuum state, breaking the symmetry of the Lagrangian. This is shown in Figure 2.2 for a complex scalar field $\phi = \phi_1 + i\phi_2$ with $\lambda > 0$ and $\mu^2 < 0$. The minimum must be electrically neutral in order to preserve the symmetry of quantum electrodynamics and is conventionally assumed to be real, which results in

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.9}$$

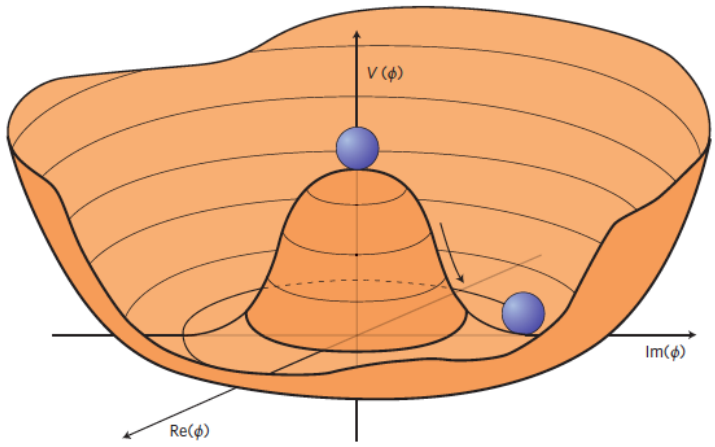


Figure 2.2.: The potential $V(\phi) = \mu^2(\phi^*\phi) + \lambda(\phi^*\phi)^2$ for a complex scalar field $\phi = \phi_1 + \phi_2$ with $\lambda > 0$ and $\mu^2 < 0$.

In order to interpret the Lagrangian more clearly, the fields can be expanded about this minimum as

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ v + H(x) + i\phi_4(x) \end{pmatrix}. \tag{2.10}$$

Setting $\phi_1(x) = \phi_2(x) = \phi_4(x) = 0$ such that the three massless Goldstone bosons [30, 31] result in the longitudinal degrees of freedom of the $W^\pm$ and Z bosons, Equation 2.10 can directly be written in the unitary gauge as

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.11}$$

To obtain the masses of the gauge bosons while respecting the $SU(2)_L \times U(1)_Y$ local gauge invariance of the electroweak theory, the derivatives in the Lagrangian (Equation 2.7) are replaced by the covariant derivatives (Equation 2.3), but without the term for the $SU(3)_C$ gauge group. By defining the observed electroweak bosons as mass eigenstates of W^a_μ and B_μ as

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp i W^2_\mu\right), \quad Z_\mu = \frac{g_2 W^3_\mu - g_1 B_\mu}{\sqrt{g_2^2 + g_1^2}}, \quad \text{and} \quad A_\mu = \frac{g_1 W^3_\mu + g_2 B_\mu}{\sqrt{g_2^2 + g_1^2}}, \tag{2.12}$$

the Lagrangian can be expanded as

$$\begin{aligned}\mathcal{L}_H = &\underbrace{\frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2}_{\text{massive scalar}} \underbrace{-\lambda v H^3 - \frac{1}{4}\lambda H^4}_{\text{self-interactions}} \underbrace{+\frac{g_2^2 v^2}{4} W^-_\mu W^{+\mu} + \frac{\left(g_2^2 + g_1^2\right) v^2}{8} Z_\mu Z^\mu}_{\text{mass terms of gauge bosons}} \\ &\underbrace{+\frac{g_2^2 v}{2} W^-_\mu W^{+\mu} H + \frac{\left(g_2^2 + g_1^2\right) v}{4} Z_\mu Z^\mu H + \frac{g_2^2}{4} W^-_\mu W^{+\mu} H^2 + \frac{g_2^2 + g_1^2}{8} Z_\mu Z^\mu H^2}_{\text{interactions between the scalar and gauge fields}}.\end{aligned} \tag{2.13}$$

From the mass terms in the above Lagrangian, the masses of the Higgs boson and the gauge bosons are given as

$$m_H = \sqrt{2\lambda v^2}, \quad m_W = \frac{g_2 v}{2}, \quad m_Z = \frac{v}{2}\sqrt{g_2^2 + g_1^2} \quad \text{and} \quad m_A = 0. \tag{2.14}$$

In addition to this, triple and quartic couplings between the gauge bosons and the Higgs boson along with the self-interactions of the Higgs boson appear in the Lagrangian, which are also invariant under local gauge transformations. Figure 2.3 depicts these self-couplings of the Higgs boson with the respective coupling strengths.

(a) trilinear ***(b)*** quartic

Figure 2.3.: Feynman diagrams depicting the self-couplings of the SM Higgs boson. The coupling strengths of the vertices are denoted in red.

Until now, only the masses of gauge bosons and their interactions with the Higgs boson

have been discussed. But the Higgs mechanism also provides mass to the fermions. These interactions are described by another term in the Lagrangian

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_f \left(\bar{\psi}_L^f \Phi \psi_R^f + \bar{\psi}_R^f \Phi^\dagger \psi_L^f \right) \quad \text{with} \quad m_f = \frac{\lambda_f v}{\sqrt{2}}, \tag{2.15}$$

where λ_f represents the Yukawa coupling of the corresponding fermion.

In the SM, neutrinos are considered to be massless. However, it is known that at least two out of the three neutrino mass eigenstates must be non-zero in order to explain the observed neutrino oscillations [19–21]. The SM must thus be extended to explain neutrino masses, but this is out of the scope of this thesis.

In 2012, a scalar boson was discovered by the ATLAS and CMS experiments at the LHC [10, 11] which is consistent with the SM Higgs boson. Currently, the world average of its mass measurements yields a value of $m_H = 125.10 \pm 0.14\,\text{GeV}$ [29].

2.2. Theories Beyond the Standard Model

As discussed in Section 2.1, the Higgs sector in the SM comprises of just one $SU(2)$ doublet, which is the simplest of all possible scalar structures to solve the problem of particle masses in the SM. However, other sectors of the SM, such as the fermion structure with 3 generations, are not so simple. This provides a reason to look for BSM theories in the Higgs sector to also solve some more phenomena not explained by the SM. This section aims at providing a description of the most relevant BSM scenarios for this thesis. These include the 2HDM [26], with a few words on the Minimal Supersymmetric extension of the Standard Model (MSSM) [32], and the TRSM [25].

Two-Higgs-Doublet Models

The 2HDM, as the name suggests, includes two Higgs doublet fields, Φ_1 and Φ_2, instead of just one, as in the SM. To simplify the model, some commonly made assumptions are:

- CP invariance of the Higgs sector,
- CP not being spontaneously broken, and
- all quartic terms, which are odd in either of the doublets, being eliminated by discrete symmetries.

With these assumptions, the most general scalar potential for Φ_1 and Φ_2 with $Y = +1$ is

$$\begin{aligned} V = {} & m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - m_{12}^2 \left(\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1 \right) + \frac{\lambda_1}{2} \left(\Phi_1^\dagger \Phi_1 \right)^2 + \frac{\lambda_2}{2} \left(\Phi_2^\dagger \Phi_2 \right)^2 \\ & + \lambda_3 \Phi_1^\dagger \Phi_1 \Phi_2^\dagger \Phi_2 + \lambda_4 \Phi_1^\dagger \Phi_2 \Phi_2^\dagger \Phi_1 + \frac{\lambda_5}{2} \left[\left(\Phi_1^\dagger \Phi_2 \right)^2 + \left(\Phi_2^\dagger \Phi_1 \right)^2 \right], \end{aligned} \tag{2.16}$$

with all parameters being real. The minimisation of the above potential yields

$$\langle \Phi_1 \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_1 \end{pmatrix} \quad \text{and} \quad \langle \Phi_2 \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_2 \end{pmatrix}. \tag{2.17}$$

With two complex scalar doublets, there are now a total of eight fields, of which three are again required for the longitudinal degrees of freedom of the $W^\pm$ and Z bosons, resulting in five physically observable "Higgs" bosons: two electrically neutral scalars (h^0 and H^0), two electrically charged scalars ($H^\pm$) and an electrically neutral pseudoscalar (A^0), with masses of the charged scalar ($m_\pm$) and of the pseudoscalar (m_A) given as

$$m_\pm^2 = \left(\frac{m_{12}^2}{v_1 v_2} - \lambda_4 - \lambda_5 \right) \left(v_1^2 + v_2^2 \right) \quad \text{and} \quad m_A^2 = \left(\frac{m_{12}^2}{v_1 v_2} - 2\lambda_5 \right) \left(v_1^2 + v_2^2 \right), \tag{2.18}$$

respectively. The masses of the two neutral scalars can be obtained by diagonalising the mass-squared matrix

$$M_H^2 = \begin{pmatrix} m_{12}^2 \dfrac{v_2}{v_1} + \lambda_1 v_1^2 & -m_{12}^2 + \lambda_{345} v_1 v_2 \\ -m_{12}^2 + \lambda_{345} v_1 v_2 & m_{12}^2 \dfrac{v_1}{v_2} + \lambda_2 v_2^2 \end{pmatrix}, \tag{2.19}$$

where $\lambda_{345} = \lambda_3 + \lambda_4 + \lambda_5$. The angle α is defined as the angle of rotation to perform the diagonalisation. The theory of 2HDM with the above-mentioned assumptions results in six free parameters including the four masses of the physical bosons, the rotation angle, and the ratio of the VEVs of the two doublets

$$\tan \beta \equiv \frac{v_2}{v_1}. \tag{2.20}$$

The 2HDM can be divided into various classes based on the couplings of the two doublets, Φ_1 and Φ_2, to the different types of fermions. In order to ensure the absence of flavour-changing neutral currents (FCNC), the Glashow-Weinberg condition [33] is enforced. This requires each group of fermions (up-type quarks, down-type quarks and charged or down-type leptons) to couple to exactly one doublet and also acts as a necessary and sufficient condition for the same. Without any FCNC, a total of four different

classes of 2HDM can exist. By convention Φ_2 is considered to be the doublet coupling to the up-type quarks in all four classes. These are then divided as:

- Type I, where all three fermion groups couple to Φ_2,
- Type II, where the down-type fermions couple to Φ_1,
- Lepton-specific, where Φ_2 couples to quarks and Φ_1 couples to charged leptons, and
- Flipped, where Φ_1 couples to down-type quarks and Φ_2 couples to charged leptons.

These four classes are also summarised in Table 2.1.

Table 2.1.: The four possible classes of the 2HDM with the absence of FCNC along with the couplings of each fermion group to the two doublet fields. By convention up-type quarks always couple to Φ_2.

Class	ψ_R^u	ψ_R^d	ψ_R^l
Type I	Φ_2	Φ_2	Φ_2
Type II	Φ_2	Φ_1	Φ_1
Lepton-specific	Φ_2	Φ_2	Φ_1
Flipped	Φ_2	Φ_1	Φ_2

Apart from some other motivations such as for axion models [34] and for baryon asymmetry [35–41], one of the biggest motivations of 2HDM is Supersymmetry (SUSY) [42, 43], which introduces a fermionic counterpart to each SM boson and *vice versa*. Since in SUSY theories the scalar and its complex conjugates belong to multiplets of opposites chiralities, a single Higgs doublet cannot simultaneously generate masses for up- and down-type quarks. Therefore, a SUSY theory must have at least two Higgs doublets. The MSSM [32] is the simplest case of SUSY with exactly two Higgs doublets and is hence an example of 2HDM, particularly of the Type-II class.

In addition to the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry of the SM, the MSSM introduces conservation of an additional discrete symmetry known as the R-parity defined as

$$R_p = (-1)^{2s+3B+L}, \tag{2.21}$$

where s is the spin of the particle and B and L are the baryon and lepton numbers, respectively. By definition, the R-parity of ordinary SM particles is given as $R_p = +1$, while for their supersymmetric counterparts it is given as $R_p = -1$. The conservation of this discrete symmetry ensures the stability of the lightest supersymmetric particle (LSP) and provides a candidate for dark matter. Moreover, the MSSM is also able to solve

the hierarchy problem, which refers to the large radiative corrections to the Higgs mass. By the introduction of additional scalar bosons, the MSSM does this with a very precise cancellation (in the order of 10^{-30}) to the Planck scale [44].

In this search the $X \to HH$ scenario attempts to set limits on such 2HDM models, by identifying X as the neutral heavy scalar, H^0, and H as the SM-like light neutral scalar h^0.

Two-Real-Scalar-Singlet Extension of Standard Model

The TRSM attempts to extend the scalar sector of the SM with the addition of two real scalar singlet degrees of freedom written as two real singlet fields S and X. Two discrete $\mathbb{Z}_2$ symmetries are introduced to reduce the number of free parameters. These are

$$\mathbb{Z}_2^S : S \to -S,\ X \to X,\ \text{SM} \to \text{SM} \quad \text{and} \quad \mathbb{Z}_2^X : X \to -X,\ S \to S,\ \text{SM} \to \text{SM}. \tag{2.22}$$

Then the most general renormalisable scalar potential invariant under the $\mathbb{Z}_2^S \otimes \mathbb{Z}_2^X$ symmetry is given as

$$\begin{aligned} V = {} & \mu_\Phi^2 \Phi^\dagger \Phi + \lambda_\Phi \left(\Phi^\dagger \Phi\right)^2 + \mu_S^2 S^2 + \lambda_S S^4 + \mu_X^2 X^2 + \lambda_X X^4 \\ & + \lambda_{\Phi S} \Phi^\dagger \Phi S^2 + \lambda_{\Phi X} \Phi^\dagger \Phi X^2 + \lambda_{SX} S^2 X^2, \end{aligned} \tag{2.23}$$

where all coefficients are real. Again, the fields can be decomposed in the unitary gauge as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \phi_h + v \end{pmatrix}, \quad S = \frac{\phi_S + v_S}{\sqrt{2}}, \quad X = \frac{\phi_X + v_X}{\sqrt{2}}, \tag{2.24}$$

where $v = 246\,\text{GeV}$ is the SM VEV. For cases where v_S and v_X are non-zero, the $\mathbb{Z}_2$ symmetries are spontaneously broken resulting in three physical scalars from the mixing between the three fields ϕ_h, ϕ_S and ϕ_X, also known as the *broken phase*. In the broken phase, the mass eigenstates $h_{1,2,3}$, ordered by their masses as

$$m_1 \leq m_2 \leq m_3, \tag{2.25}$$

are related to the fields $\phi_{h,S,X}$ via the 3×3 orthogonal mixing matrix R

$$\begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = R \begin{pmatrix} \phi_h \\ \phi_S \\ \phi_X \end{pmatrix}. \tag{2.26}$$

The mixing matrix R is parametrised by three mixing angles $\theta_{hS},\ \theta_{hX},\ \theta_{SX}$ and is given as

$$R = \begin{pmatrix} c_1c_2 & -s_1c_2 & -s_2 \\ s_1c_3 - c_1s_2s_3 & c_1c_3 + s_1s_2s_3 & -c_2s_3 \\ c_1s_2c_3 + s_1s_3 & c_1s_3 - s_1s_2c_3 & c_2c_3 \end{pmatrix}, \tag{2.27}$$

where the short-hand notations

$$\begin{aligned} &s_1 \equiv \sin\theta_{hS}, \quad s_2 \equiv \sin\theta_{hX}, \quad s_3 \equiv \sin\theta_{SX}, \\ &c_1 \equiv \cos\theta_{hS}, \quad c_2 \equiv \cos\theta_{hX}, \quad \text{and} \quad c_3 \equiv \cos\theta_{SX}, \end{aligned} \tag{2.28}$$

is used. From Equation 2.23, it can be seen that the scalar potential of the TRSM has a total of nine free parameters. These include the masses of the three physical scalars ($m_{1,2,3}$), the VEVs of the fields $\phi_{h,S,X}$ (v, v_S and v_X), and the three mixing angles of the mixing matrix R (θ_{hS}, θ_{hX} and θ_{SX}).

It can be shown that parametrisation of a single complex scalar singlet field has no physical difference compared to that of two real scalar singlet fields, since it is not possible to write gauge invariant and renormalisable interactions between a scalar singlet and any of the SM fermions. This also removes any CP-violating interactions that one would naively expect due to the imaginary parts of the complex scalar fields. Therefore, any pure singlet extension of the SM is a theory of only CP-even scalars. However, the more stringent symmetry of the CxSM (addition of a complex scalar singlet field) [24] leads to one of the physical scalar masses and one of the singlet VEVs being dependent on the remaining free parameters. Moreover, since the scalar singlet fields do not have any direct gauge or fermion couplings, the singlets can only interact with the SM Higgs boson via the couplings of the scalar potential and will inherit some of its gauge and Yukawa couplings due to their mixing with the SM Higgs boson in the broken phase. This also means that if one or both of the singlet VEVs are exactly zero, the singlet fields cannot interact with the SM particles and are stabilised by the $\mathbb{Z}_2$ symmetries. This results in possible dark matter candidates. Furthermore, the TRSM also allows symmetrical and asymmetrical cascade decays of the scalars which include

$$h_3 \to h_2h_2, \quad h_3 \to h_1h_1, \quad h_3 \to h_2h_1 \quad \text{and} \quad h_2 \to h_1h_1, \tag{2.29}$$

so long as they are not kinematically forbidden.

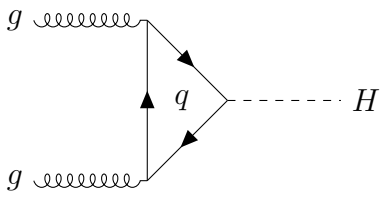
The $X \to SH$ scenario of this search attempts to set limits on the TRSM, by identifying X as the heaviest scalar, h_3, S as the intermediary scalar h_2, and H as the SM-like light scalar, h_1.

2.3. Higgs Production Modes and Decays Channels

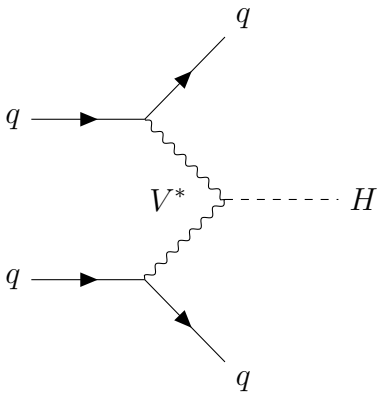
This section provides details about the possible production modes for a single Higgs boson and Higgs boson pair production at the LHC with the latest cross-section predictions. It also discusses the primary decay modes with the current branching ratio ($\mathcal{BR}$) calculations.

Single Higgs Boson Production

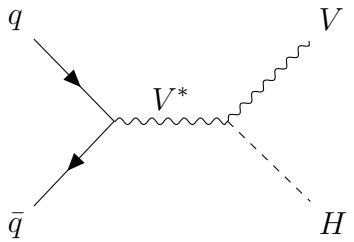
At the LHC, the SM Higgs boson can be produced in various processes. Figure 2.4 shows the Feynman diagrams of the four most important of these processes and also lists their respective production cross sections at centre-of-mass energy, $\sqrt{s} = 13\,\text{TeV}$ for the SM Higgs boson with $m_H = 125.10\,\text{GeV}$.



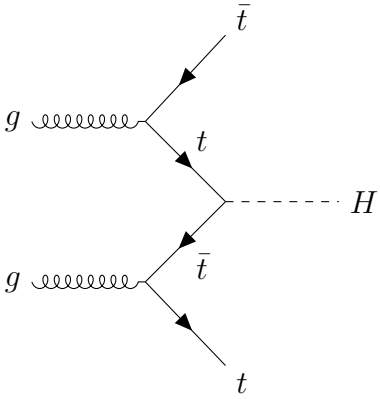
(a) ggF ($\sigma = 48.61\,\text{pb}$)



(b) VBF ($\sigma = 3.77\,\text{pb}$)



(c) VH ($\sigma = 1.36\,\text{pb}$ ($W^{\pm}$) or $0.88\,\text{pb}$ (Z))



(d) $t\bar{t}H$ ($\sigma = 0.61\,\text{pb}$)

Figure 2.4.: Feynman diagrams of the main production modes of the SM Higgs boson at the LHC.

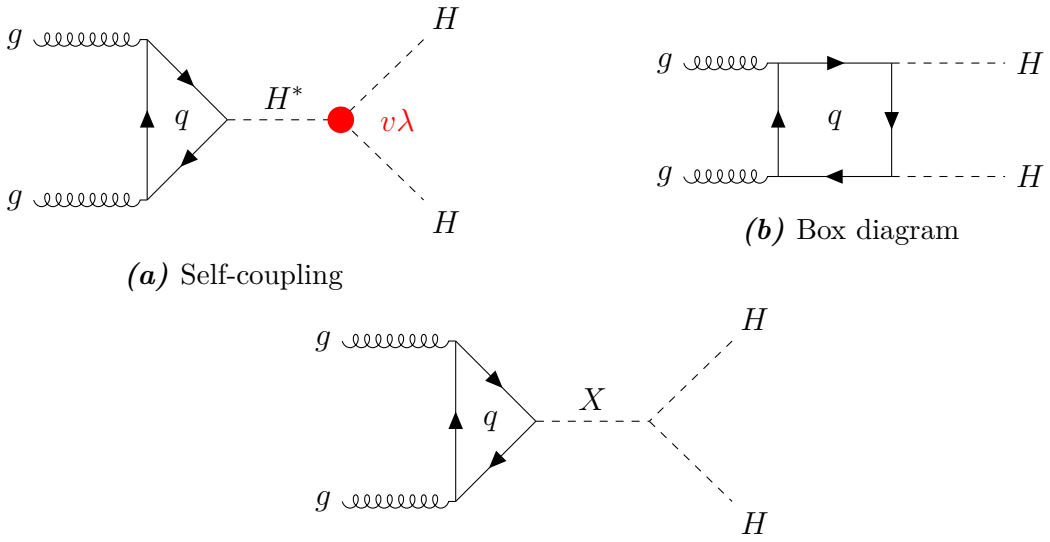
In proton-proton collisions with $\sqrt{s} = 13\,\text{TeV}$, which are the conditions for the LHC

Run 2, the highest cross section for the production of a Higgs boson is provided by the gluon-gluon fusion process (ggF, Figure 2.4a) with $\sigma = 48.61\,\text{pb}$ [45]. The second most dominant is the vector boson fusion process (VBF, Figure 2.4b) with $\sigma = 3.77\,\text{pb}$ [45], which is an order of magnitude smaller. The VBF process has a very distinct signature in the detector due to the two quarks having a very small scattering angle. Third in line is the associated production of the Higgs boson with a vector boson (VH, Figure 2.4c), where V can either be a $W^{\pm}$ boson ($\sigma = 1.36\,\text{pb}$ [45]) or a Z boson ($\sigma = 0.88\,\text{pb}$ [45]). The fourth process in the associated production of Higgs boson with a top quark pair ($t\bar{t}H$, Figure 2.4d) with $\sigma = 0.61\,\text{pb}$ [45].

Higgs Boson Pair Production

In the context of the LHC, the Higgs boson pair production can dominantly take place in two ways according to the SM. These are shown in Figures 2.5a and 2.5b and are together known as the non-resonant modes of Higgs boson pair production. Both of these processes start with two gluons interacting via a quark loop, but in the self-coupling diagram (Figure 2.5a), the quark loop radiates one off-shell Higgs boson which further decays into a pair of Higgs bosons. This point of self-interaction between three Higgs bosons, which was also discussed in Section 2.1, provides a handle to directly probe the SM Higgs potential by measuring λ from the coupling strength of the triple-point vertex. The box diagram (Figure 2.5b), however, interferes destructively with the self-coupling diagram and brings the overall cross section for Higgs boson pair production at $\sqrt{s} = 13\,\text{TeV}$ to $\sigma^{\text{SM}}_{gg\to HH} = 31.05\,\text{fb}$ [46, 47], which might be measurable in the future runs of the LHC, such as the high-luminosity phase (HL-LHC), but is too small to be measured with the current Run 2 data.

However, as discussed in the previous section, there are many BSM theories that predict a heavy scalar X (heavy neutral Higgs boson from the 2HDM or the scalar X from the TRSM) with a possible decay into a pair of Higgs bosons. This is called resonant production and is shown in Figure 2.5c. Due to the presence of many theories predicting such processes, the cross section of resonant Higgs boson pair production is unknown. This could then be large enough to be measurable with the current Run 2 data at the LHC. Therefore, such searches for resonant Higgs boson pair production could lead to discoveries of new physics beyond the SM [48, 49].



(a) Self-coupling

(b) Box diagram

(c) Resonant

Figure 2.5.: Feynman diagrams of the dominant Higgs boson pair production modes. Here, (a) and (b) correspond to non-resonant production, while (c) represents the case of resonant production of a pair of Higgs bosons.

Decay Channels

Since the Higgs boson is an unstable particle, it further decays into lighter, more stable particles. With its mass of $m_H = 125.10\,\text{GeV}$, the Higgs boson is massive enough to decay into any SM particle, except the top quark. The decay into top quarks is kinematically forbidden due to its large mass. The $\mathcal{BR}$s for different decay modes are shown in Table 2.2 for a single Higgs boson and in Figure 2.6 for a pair of Higgs bosons.

Table 2.2.: $\mathcal{BR}$s of several decay channels of a single Higgs boson assuming $m_H = 125.10\,\text{GeV}$ [50].

Decay Channel	$\mathcal{BR}$ [%]
$b\bar{b}$	58.07
WW	21.54
gg	8.18
$\tau\tau$	6.26
$c\bar{c}$	2.88
ZZ	2.64
$\gamma\gamma$	0.23
$Z\gamma$	0.15
$\mu\mu$	0.02

From Table 2.2, it is visible that the highest $\mathcal{BR}$ of the Higgs boson is for the decay

	bb	WW	ττ	ZZ	γγ
bb	33%				
WW	25%	4.6%			
ττ	7.4%	2.5%	0.39%		
ZZ	3.1%	1.2%	0.34%	0.076%	
γγ	0.26%	0.10%	0.029%	0.013%	0.0005%

Figure 2.6.: Branching ratios of a pair of Higgs bosons for several decay channels.

into a $b\bar{b}$ pair with 58.07% probability [50]. This is expected since the coupling of the Higgs boson to fermions is directly proportional to the fermion mass and the bottom quark is the heaviest of all the fermions the Higgs boson can decay into. The second most dominant decay mode is the WW decay channel with a probability of 21.54% [50]. Even though the $H \to WW$ coupling is much stronger than the $H \to b\bar{b}$ coupling, the kinematic suppression changes the $\mathcal{BR}s$. In this case, since the mass of the Higgs boson is smaller than twice the mass of a W boson, at least one of the two W bosons from the decaying Higgs boson will be off-shell and unstable. This causes the $H \to WW$ decay channel to have a lower preference. Other decay channels for the Higgs boson include the decays into a pair of gluons via a quark loop, a pair of τ leptons, a $c\bar{c}$ pair, a pair of Z bosons, and a pair of photons via a fermionic or bosonic loop.

In the case of a pair of Higgs bosons (Figure 2.6), both can decay independently of each other. Hence, the dominant decay channel here is the $b\bar{b}b\bar{b}$ channel with 33% probability, followed by the $b\bar{b}WW^*$ channel with 25% probability, and the $b\bar{b}\tau\tau$ channel with 7.4% probability. This thesis focusses on the $b\bar{b}WW^*$ decay channel, particularly for the case where one W boson decays hadronically and the other decays leptonically. The total $\mathcal{BR}$ for this process can be calculated as

$$\mathcal{BR}\left(HH \to b\bar{b}\ell\nu q\bar{q}\right) = \mathcal{BR}\left(HH \to b\bar{b}WW^*\right) \times \mathcal{BR}\left(WW \to \ell\nu q\bar{q}\right) = 10.86\%. \quad (2.30)$$

The complete event topology will be discussed in more detail in Section 4.1.

In this thesis, scenarios with two additional scalar bosons (X and S, which will refer to the mass eigenstates instead of the fields from this point on) apart from the SM Higgs

boson, H were also considered with $m_X > m_S > m_H$. Following the benchmark scenario, **BP3** of Ref. [25], searches were conducted for the $X \to SH$ signatures in the $b\bar{b}WW^{(*)}$ decay channel. Since S is expected to have Higgs-like couplings to SM particles, for $m_S > 135\,\text{GeV}$, the decay of S to WW dominates over the $b\bar{b}$ decay channel. This makes $b\bar{b}WW^{(*)}$ the decay channel with the largest $\mathcal{BR}$ which can also be seen in Figure 2.7.

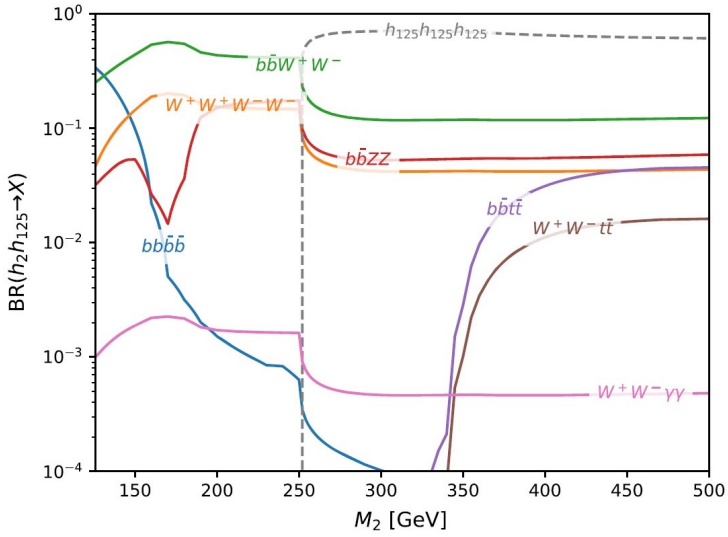


Figure 2.7.: $\mathcal{BR}$s of the SH state (h_2h_{125}) into various decay channels as a function of m_S (M_2) [25]. In the TRSM, the cascade decay $S \to HH$, giving three SM-like Higgs bosons in the final state can be tuned away, which leaves the $b\bar{b}WW^{(*)}$ as the dominant decay mode.

2.4. Profile Likelihood Fits

This section briefly describes the method of maximum likelihood estimators and how they are used for setting upper limits in particle physics experiments. A brief description of the CL_s method is also included in the end of the section. More detailed description can be found in Refs. [51, 52].

Likelihood Estimators

The idea of profile likelihood ratios builds upon the probability theory, where instead of calculating the probability of a certain event with some given parameters, the likelihood of the parameters is calculated based on the given dataset. The value of the parameter that maximises this likelihood function, or minimises the negative-log-likelihood (NLL) is called the maximum likelihood estimator.

In particle physics, often the cross-section for signals is measured in the presence of backgrounds, and then compared with theory predictions. Hence, the number of expected events is then $\mu s + b$, where s is the predicted signal yield, b is the background prediction and μ is the signal strength, which is also the parameter of interest. The signal strength is usually allowed to take any value in order to remove bias from the theory.

The observed and expected events are then binned into a multinomial distribution, followed by the addition of systematic uncertainties via nuisance parameters $\vec{\theta}$ in the likelihood function, which are typically assumed to be Gaussian in nature. Put together, this results into the likelihood function. The NLL function then yields

$$\begin{aligned} -\ln\left[L\left(\mu|\vec{r},\vec{s},\vec{b},\vec{\theta}\right)\right] = & \left(\mu s_{\text{tot}}\left(\vec{\theta}\right) + b_{\text{tot}}\left(\vec{\theta}\right)\right) \\ & - \sum_{i=1}^{n} r_i \ln\left[\mu s_i\left(\vec{\theta}\right) + b_i\left(\vec{\theta}\right)\right] \\ & + \sum_{k=1}^{m} \frac{\left(\theta_k - \theta_{k_0}\right)^2}{2\sigma_{\theta_k}^2}, \end{aligned} \tag{2.31}$$

where $\vec{r}$ represents the observed events, n is the number of bins and the last term runs over all m systematic uncertainties. The minimum of NLL then gives the maximum likelihood estimator for the signal strength, $\hat{\mu}$.

Upper Limits

In cases where analyses are not sensitive enough to claim discovery or evidence, upper limits are set on the production cross-section to exclude regions of phase-space with a quantified certainty. In order to do this, a test statistic needs to be defined to compare two hypotheses. According to the Neyman-Pearson lemma [53], the optimal test statistic for this purpose is the likelihood ratio given by

$$\lambda\left(\mu\right) = \frac{L\left(\mu, \hat{\hat{\vec{\theta}}}\right)}{L\left(\hat{\mu}, \hat{\hat{\vec{\theta}}}\right)}, \tag{2.32}$$

where $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the set of values that maximise L, and $\hat{\hat{\vec{\theta}}}$ are the values of nuisance parameters that maximise L for a specific value of μ.

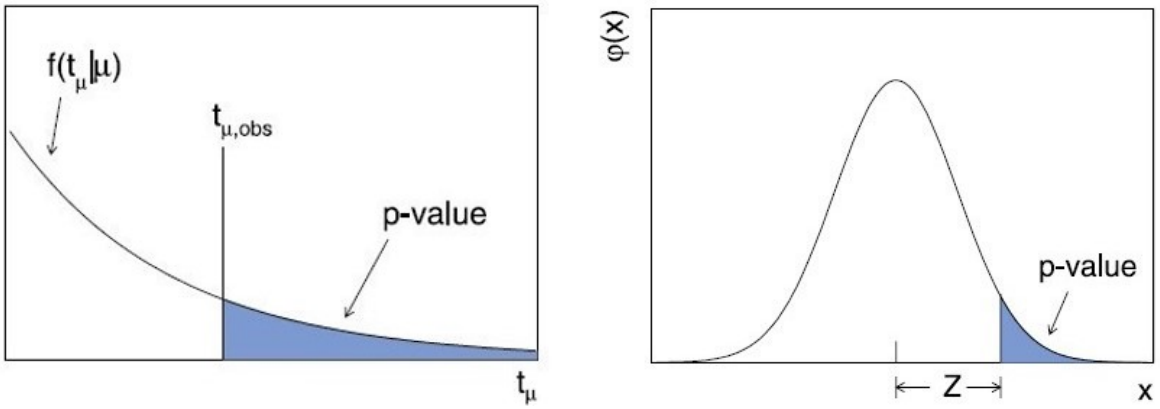
In this case a test against the most probable scenario ($\mu = \hat{\mu}$) is performed, while

looking for a deficit in events using the test statistic q_μ defined as

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \text{if } \hat{\mu} \leq \mu \\ 0 & \text{if } \hat{\mu} \geq \mu \end{cases} \quad \text{with} \quad p_\mu = \int_{q_\mu^{\text{obs}}}^{\infty} \mathcal{P}(q_\mu|\mu)\, dq_\mu, \tag{2.33}$$

where $\mathcal{P}$ is the probability distribution function (PDF) of q_μ and the p-value represents the probability for a given value of μ to yield the observed or even less compatible data. Generally, the 95% confidence level (C.L.) upper limit on μ is quoted, which is the highest value of μ, where $p_\mu \geq 0.05$. The p-value is graphically represented in Figure 2.8a.

The p-value is then used to calculate the significance, Z, which corresponds to the number of standard deviations, σ, after which the fraction of area under the remaining Gaussian tail is equal to the p-value. This is shown in Figure 2.8b. A 95% C.L. exclusion corresponds to at least 2σ excess.



(a) Relation between the p-value and q_μ^{obs} $(t_{\mu,\text{obs}})$

(b) Relation between significance Z and p-value

Figure 2.8.: Graphical representation of PDF $\mathcal{P}(q_\mu|\mu)$ $(f(t_\mu|\mu))$, observed test statistic q_μ^{obs} $(t_{\mu,\text{obs}})$, p-value and significance Z [51].

CL_s Procedure

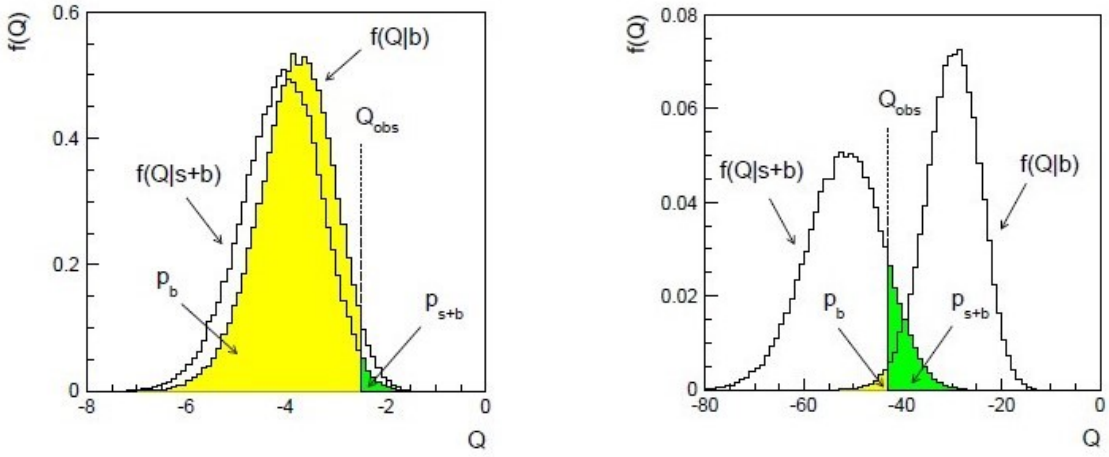
Sometimes analyses have very low sensitivity to the signal. In such cases $p_\mu < 0.05$ would imply exclusion of μ, but this might also be inconsistent with $\mu = 0$. The CL_s procedure [54] avoids such spurious exclusions by testing both, $\mu = 0$ (background-only) and $\mu = 1$ (signal+background) hypotheses simultaneously. This is represented in Figures 2.9a and 2.9b for analyses with low and high sensitivity, respectively. The test statistic used

here is defined as

$$Q = -2\ln\left(\frac{L_{\mathrm{s+b}}}{L_{\mathrm{b}}}\right) \quad \text{with} \quad \mathrm{CL_s} = \frac{\mathrm{CL_{s+b}}}{\mathrm{CL_b}} = \frac{p_{\mathrm{s+b}}}{1-p_{\mathrm{b}}}. \tag{2.34}$$

Then the signal+background hypothesis can safely be rejected if $\mathrm{CL_s} \leq 0.05$, which is quoted as 95% C.L. exclusion. The $\mathrm{CL_s}$ procedure however loses its coverage characteristic, i.e. 95% C.L. limits no longer mean that 5% of the experiments will exclude the true value. Instead, the $\mathrm{CL_s}$ procedure is more conservative as more than 95% of the experiments will cover the true value.

In this thesis, the $\mathrm{CL_s}$ method was used to set 95% C.L. upper limits on the cross section of resonant Higgs boson pair production, $\sigma(pp \to X \to HH)$, and on the cross section of SH production times its branching ratio to the $b\bar{b}WW$ decay channel, $\sigma(pp \to X \to SH) \times \mathcal{BR}(SH \to b\bar{b}WW)$.



(a) Analysis with low sensitivity

(b) Analysis with high sensitivity

Figure 2.9.: Graphical representation of the ingredients for the $\mathrm{CL_s}$ procedure [52]. Here Q_{obs} is the observed test statistic, and $f\,(Q|\mathrm{s+b})$ and $f\,(Q|\mathrm{b})$ are the PDFs for the signal+background ($\mu = 1$) and background-only ($\mu = 0$) hypotheses, respectively.

3. Experimental Setup

> "Give me six hours to chop down a tree and I will spend the first four sharpening the axe."
>
> - Abraham Lincoln

This chapter presents the experimental setup for this thesis. This includes a description of the Large Hadron Collider (LHC) at CERN and the ATLAS detector. A description of Monte Carlo (MC) event generators and the treatment of jets in boosted topologies is also provided.

3.1. The Large Hadron Collider

The LHC [55–58] is a circular proton-proton and heavy-ion synchrotron operated by the European Organisation for Nuclear Research (CERN). It has two rings with symmetrical proton beams accelerated in opposite directions, which intersect at four collision points corresponding to the four major experiments on the LHC. The LHC was installed in the 26.7 km tunnel originally used for the Large Electron Proton (LEP) collider, and is the latest addition to the CERN accelerator complex shown in Figure 3.1 and by far the largest collider to exist around the globe.

The first operation of the LHC started in late 2009 and continued in 2010 and 2011 with proton-proton (pp) collision at a center-of-mass energy, $\sqrt{s} = 7\,\text{TeV}$. In 2012 the beam energy was increased to 4 TeV per beam resulting in $\sqrt{s} = 8\,\text{TeV}$. These run periods together make up the Run 1 of the LHC. This was then followed by the Long Shutdown 1 for upgrades and maintenance from early 2013 to early 2015. In 2015, the LHC began the Run 2 data taking period with $\sqrt{s} = 13\,\text{TeV}$ which continued until 2018. At the time of writing of this thesis, the LHC is in the Long Shutdown 2 phase and preparing for Run 3, tentatively with $\sqrt{s} = 13.6\,\text{TeV}$.

The LHC, like other recent accelerators, relies on superconducting dipole magnets made of superconducting NbTi cables for bending the high-energy proton beams around the circular track. To attain the dipole magnetic field of above 8 T, superfluid helium is

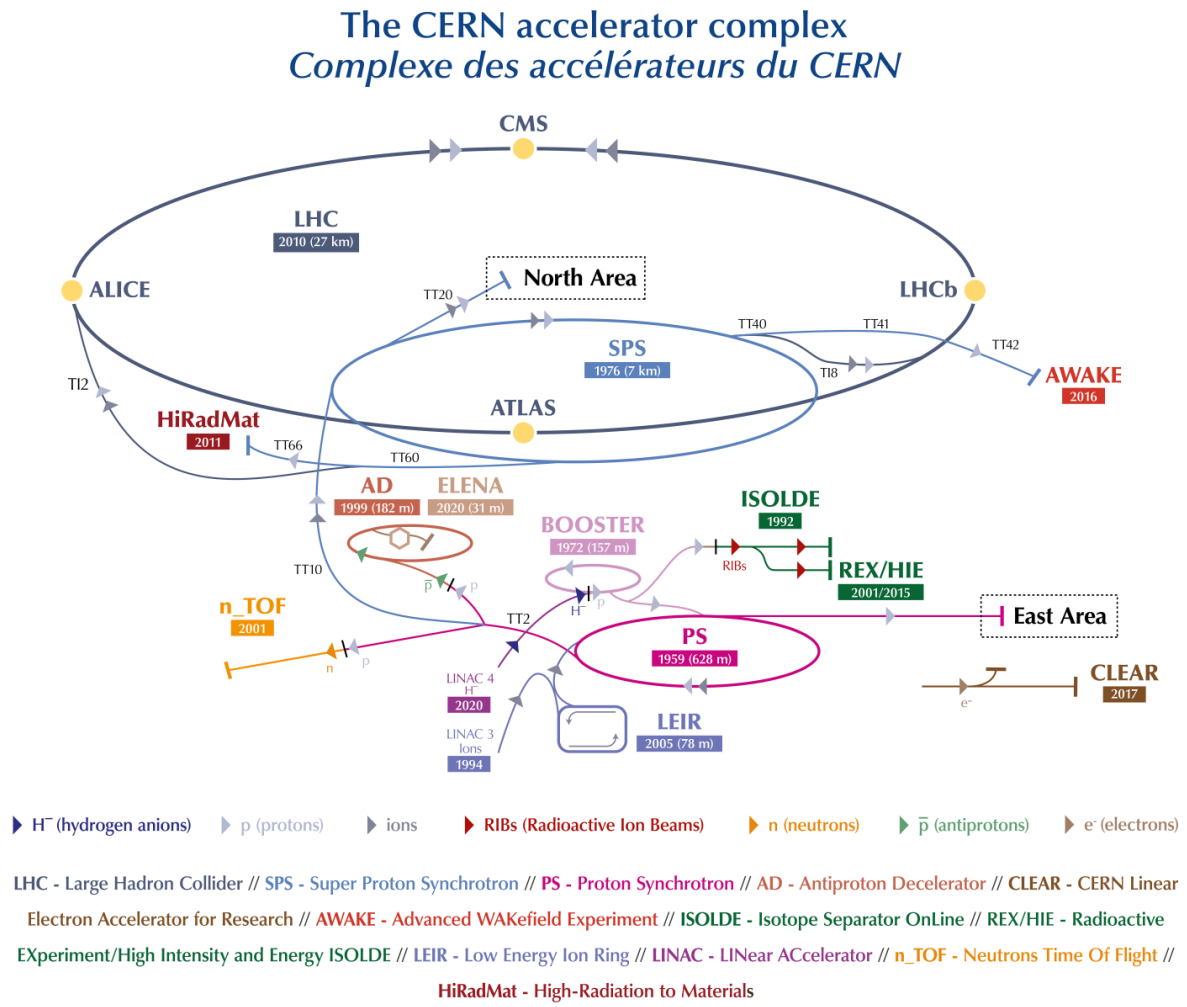


Figure 3.1.: CERN accelerator complex showing all the steps leading upto the proton beams in the LHC. © CERN

used in order to cool the magnets to 1.9 K. In addition to this, quadrupole magnets are inserted to focus these beams in order to minimise the beam loss. Many other types of magnets are also inserted for small corrections and for the stability of the beams. Due to space restrictions of recycling the LEP collider tunnel for the LHC, a "twin-bore" design was adopted for the LHC magnets, which allows it to accelerate two same-sign beams in opposite directions within a single ring.

An important measure to judge the performance of a collider is its luminosity ($\mathcal{L}$), which is defined as the ratio of the interaction rate to the interaction cross section. The performance goals of the LHC include reaching a peak luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for proton-proton collisions. To achieve this, the LHC fits about 10^{11} protons per bunch

with a total of 2808 bunches per beam. Due to its small bunch spacing time, the LHC witnesses a bunch crossing every 25 ns.

As depicted in Figure 3.1, the LHC injection chain, which in the end results in the high-energy proton beams, is a relatively complex setup. The procedure starts with the linear accelerator, LINAC 4, which accelerates negative hydrogen atoms (hydrogen atom with an additional electron) to 160 MeV and prepares them for the Proton Synchrotron Booster (PSB). The two electrons are stripped-off the during injection into the PSB, leaving only protons. The PSB accelerates the protons to 2 GeV and acts as a pre-accelerator to the Proton Synchrotron (PS), which raises the beam energy to 26 GeV. The proton beams are then transferred to the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV and are then, finally injected into the LHC. All of these accelerators have been upgraded over time to meet the stringent needs of the LHC.

The LHC has four major experiments located at the intersection points of the two proton beams. These include the two bigger general-purpose detectors, ATLAS (A Toroidal LHC ApparatuS) [59] and CMS (Compact Muon Solenoid) [60], which simultaneously claimed the discovery of the Higgs boson in 2012 [10, 11]. Additionally, there are two specialised detectors, the LHCb (Large Hadron Collider beauty) experiment [61], which focusses on B-physics and ALICE (A Large Ion Collider Experiment) [62], which focusses on heavy-ion physics.

3.2. The ATLAS Detector

Since this thesis analyses the full Run 2 data recorded by the ATLAS experiment, which corresponds to a total of 139 fb^{-1} as shown in blue in Figure 3.2 as "Good for physics", this section is dedicated to the detector itself, providing a brief description of the coordinate system used, the different components and the trigger system. A more detailed description is provided in Ref. [59].

The Coordinate System

The ATLAS detector defines its origin for the coordinate system at the interaction point of particles and is forward-backward symmetric with respect to this point. The positive x-axis is then defined to point towards the centre of the LHC ring and the positive y-axis points upwards. The detector uses a right-handed system to define the positive z-axis along the direction of the beam. From this, a particle's transverse momentum p_T can be calculated as

$$p_T = \sqrt{p_x^2 + p_y^2}, \tag{3.1}$$

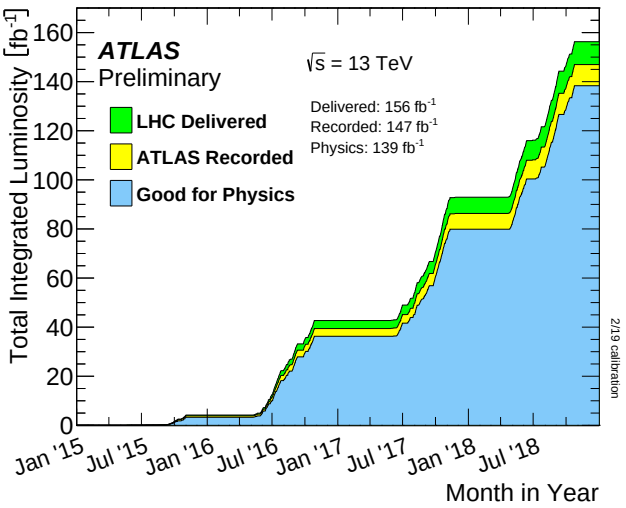


Figure 3.2.: The total integrated luminosity delivered by LHC (green), recorded by ATLAS (yellow) and considered good for physics analysis (blue) during the entire Run 2 as a function of time [63].

where p_X and p_y are the components of the particle's momentum along the x and y directions, respectively.

Due to cylindrical structure of the ATLAS detector, the azimuthal angle ϕ in the x-y plane around the beam axis and the polar angle θ between the positive z direction and the outgoing particle are also introduced. Using this, the pseudorapidity is defined as

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) = \frac{1}{2}\ln\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right), \tag{3.2}$$

where $\vec{p}$ is the momentum of the particle and p_z is the momentum in the z direction. In the relativistic limit ($m \ll |\vec{p}|$) the pseudorapidity approaches the rapidity defined as

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{3.3}$$

with E as the energy of the particle. A special property of y is that the differences in this quantity are Lorentz invariant in the z-direction, which is particularly important at hadron colliders, where the fundamental interactions involve partons carrying a fraction of the proton's total momentum. This leads to a boost along the z-axis with respect to the centre-of-mass frame of reference. However, for practical purposes η is more commonly used, since it is a purely geometric measure and can be easily expressed in terms of θ. Additionally, differences in η are approximately Lorentz invariant along the z-axis in the relativistic limit.

Another important variable is the angular distance between two points in the η-ϕ plane, defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{3.4}$$

Since $\Delta\eta$ is approximately invariant under Lorentz boosts in the z-direction and $\Delta\phi$ has no z dependence, ΔR is also approximately invariant under Lorentz boosts along the z-direction in the relativistic limit.

Components of the Detector

With a length of 46 m and a diameter of 25 m, the ATLAS detector is the largest of the four main experiments on the LHC. It weighs approximately 7000 tonnes and is composed of several layers. At its core lies the inner detector (ID) immersed in a solenoid magnet to measure the momentum and charge of charged particles. This is surrounded by the electromagnetic (ECAL) and hadronic (HCAL) calorimeters for energy measurements. The outermost layer is the muon spectrometer (MS) for identification and momentum measurements of muons. The overall layout of the ATLAS detector is split into a central barrel region and two end-cap region in the forward and backward directions. This proves to be useful in providing a nearly 4π solid angle coverage and maximises the detection efficiency. A diagram of the entire detector is shown in Figure 3.3. The dependencies of the relative resolution of each component, along with their η coverage are also summarised in Table 3.1.

Table 3.1.: Expressions for required resolution, and η coverage of each component of the ATLAS detector [59]. Here the resolutions for ID and MS are relevant for tracks, while the ones for ECAL and HCAL are relevant for electromagnetic and hadronic showers respectively.

Detector Component	η coverage	
ID	$\sigma_{p_T}/p_T = 0.05\% \cdot p_T \oplus 1\%$	$\lvert\eta\rvert < 2.5$
ECAL	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$\lvert\eta\rvert < 3.2$
HCAL (barrel, end-cap)	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\lvert\eta\rvert < 3.2$
HCAL (forward)	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$
MS	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1\,\text{TeV}$	$\lvert\eta\rvert < 2.7$

The ID is positioned closest to the interaction point for momentum measurements of charged particles. From the Lorentz force, it is known that trajectories of charged particles bend perpendicular to the magnetic field. The p_T of a particle can hence be calculated from the particle's charge and the curvature of its path. Therefore, the ID is immersed

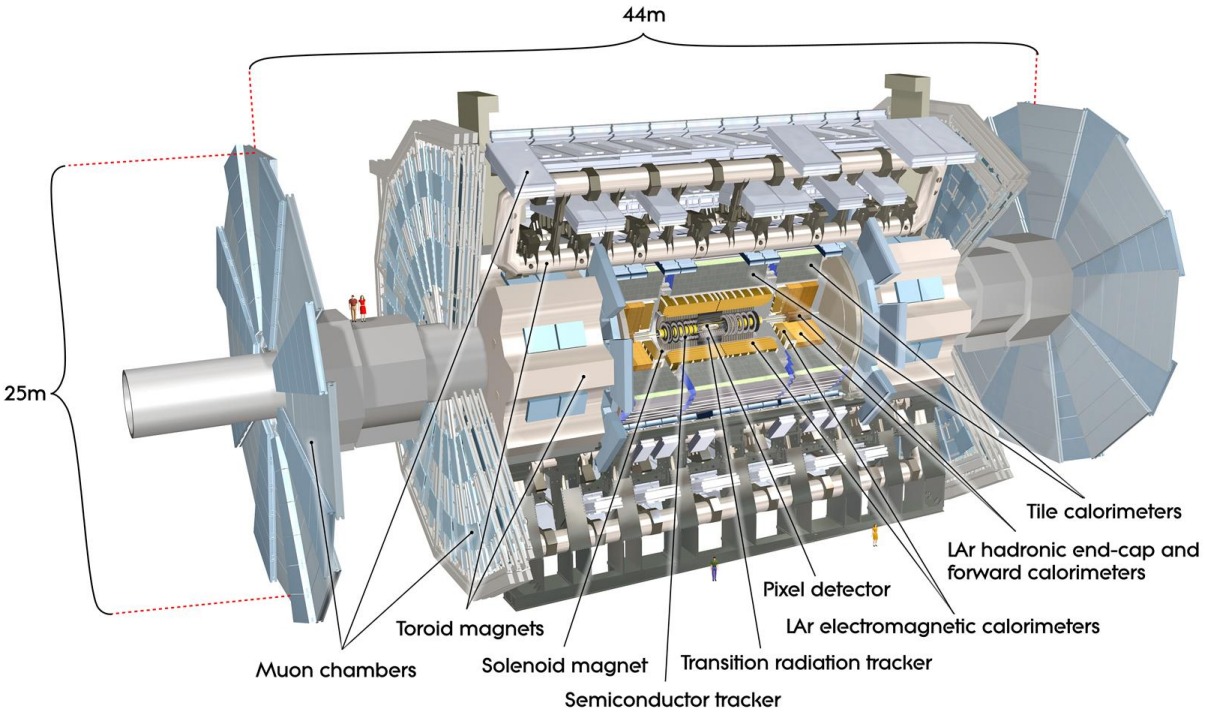


Figure 3.3.: Cut-away diagram of the ATLAS detector depicting its dimensions and the different components [59].

in a 2 T solenoid field. To achieve a high spatial resolution, the ID is made of three subcomponents. Going from the innermost to the outermost, the first subcomponent is the Silicon Pixel Detector. The innermost layer of the Pixel detector is called the Insertable B-Layer (IBL) [64] and was inserted during the Long Shutdown 1 phase of the LHC in order to improve the resolution of the primary and secondary vertex reconstruction. The pixels of the IBL are $50 \times 250\,\mu\text{m}^2$ in size. Every pixel in the layers beyond the IBL is $50 \times 400\,\mu\text{m}^2$ in size and provides a resolution of $14 \times 115\,\mu\text{m}^2$. In total, the Pixel detector consists a total of 92 million pixels. Secondly, the Semiconductor Tracker (SCT), which is a silicon microstrip tracker with a readout strips in every $80\,\mu\text{m}^2$ of area. Due to the increase in spatial separation between tracks with an increasing distance from the interaction point, the degradation of resolution from the pixel detector to the SCT is not problematic. Finally, the Transition Radiation Tracker (TRT), which is the outermost part of the ID, is made up of many smaller straw tubes. It also assists in distinguishing between electrons and pions by its enhanced electron identification capabilities provided by the detection of transition-radiation photons in the xenon-based gas mixture of the straw tubes. While the Silicon detectors pick up charge deposits from charged particles, the gas drift tubes are ionised by charged particles passing through them. These induce

charges in the detection elements, which could be in the form of wires or plates.

Outside the ID is the ECAL, which is responsible for measuring the energies of mainly electrons, positrons and photons. This is achieved by allowing the particles to interact with the calorimeter material, causing a cascade of Bremsstrahlung and pair production processes. These result in particle showers until the shower particles have energies low enough for the material to absorb them. The ECAL is a sampling calorimeter with alternate sections of lead as the absorber material, and liquid argon (LAr) as the active material. The ATLAS ECAL has a relatively fine granularity, and is designed with an accordion-like structure to have uniform response in each direction. The total thickness of the ECAL is $> 22\,(24)$ radiation lengths, X_0, in the barrel (end-cap) region.

The ECAL is further surrounded by the HCAL, which is also a sampling calorimeter responsible for stopping hadrons by absorbing their energies. It is composed of three sections, each with different materials. The scintillator tile calorimeters in barrel region employ iron as the absorber with plastic scintillators as the active material. The end-cap and forward regions have copper and tungsten as absorbers, respectively, with both using LAr as the active material. The spatial resolution of the HCAL is worse than the ECAL, but it has a larger shower depth with larger η coverage. The HCAL has a thickness of approximately $9.7\,(10)$ interaction lengths, λ_{int}, of active calorimeter in the barrel (end-cap) region. Including the outer support, the total thickness is $11\lambda_{\text{int}}$ at $\eta = 0$.

It is important to mention that even with the distinction between ECAL and HCAL, hadrons can shower in the ECAL and the electrons and photons can leak into the HCAL. But generally, the electromagnetic showers are shorter and narrower, whereas the hadronic showers start later and are longer and wider, which enables accurate identification.

Finally, the outermost part of the ATLAS detector is the MS dedicated to momentum measurements of muons, since they often traverse through the calorimeters. The muons travel through another toroidal magnetic field, which again bends them in curved trajectories to measure the transverse momentum with higher precision. The MS is composed of different types of gaseous drift chambers in the central region with three large wheels at the end caps. Due to their long size, these chambers contribute towards improving the momentum resolution.

Trigger System

In Section 3.1, it was mentioned that the LHC has bunch crossings every 25 ns, i.e. a frequency of 40 MHz, which corresponds to a data rate of approximately 600 TB/s, which is too large and unmanageable for the current data processing and storage techniques. However, the overwhelming majority of these collisions are scattering events that are

physically uninteresting. To reduce this rate to a manageable amount, a two-level trigger system is deployed such that only a smaller subset of the total data is stored for further physics analysis [65]. The purpose of these triggers is to identify potentially interesting events with high-energy outgoing particles.

The level-1 trigger (L1) is a hardware based trigger, which uses the information from the calorimeters and the muon spectrometer to look for events having large energy deposits or high momentum tracks. The information for these events is then collected and compared to predefined trigger items. The L1 trigger reduces the event rate to approximately 100 kHz.

For the second level ATLAS uses a software based high-level trigger (HLT). The HLT uses algorithms for reconstruction and signature analyses in order to identify particles, allowing the application of particle specific p_T or E_T thresholds. This brings down the event rate to about 2 kHz. The information passing the HLT is implemented using offline analysis procedures and stored permanently for further physics analysis.

3.3. Monte Carlo Event Generators

In order to perform physics analysis with the ATLAS detector it is necessary to have detailed modelling of physics processes along with the precise knowledge of the detector's response to these processes. This is achieved by simulating samples using Monte Carlo (MC) event generators. A multi-step procedure needs to be followed to reach a stage where the simulated events can be directly compared to the real data recorded by the detector [66]. The steps are also represented in Figure 3.4. The final step of detector simulation is, however, not included in the figure.

The first step to event generation is to calculate the matrix element and hence the cross section corresponding to the relevant Feynman diagrams. This also includes the prompt decays of unstable particles such as the top quark, tau lepton, and W/Z and Higgs bosons, which have extremely short lifetimes ($c\tau < 10\,\text{mm}$) and are neither affected by the magnetic field nor by the geometry of the detector. These calculations generally range from leading order (LO) to next-to-next-to leading order (NNLO). Since the interacting partons only carry a fraction of the proton's total momentum, parton distribution functions (PDFs) are included in the calculations of the matrix element, which predict the probability for a certain parton's momentum. Common examples of matrix element generators include SHERPA [68], MADGRAPH5_AMCNLO [69] and PYTHIA [70].

In the next step, QCD and electroweak corrections are applied which include higher order initial and final state radiation due to the self-coupling of the gluon and the in-

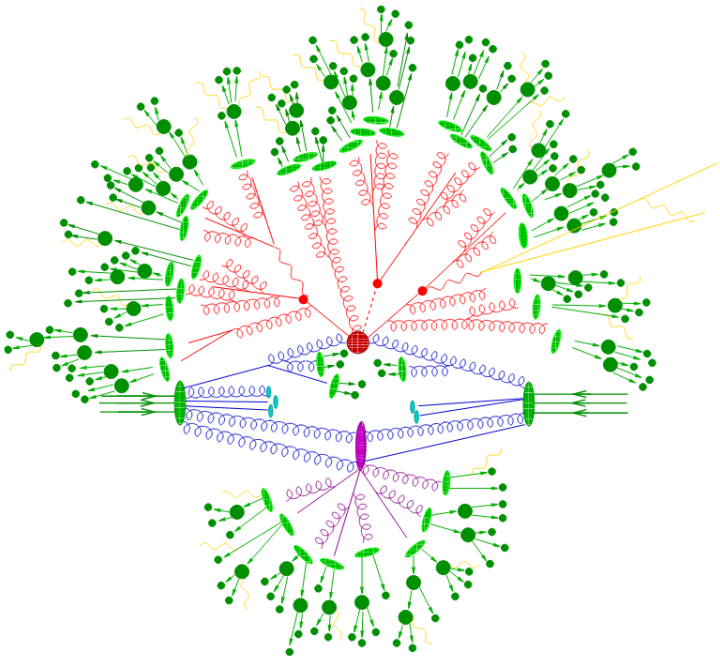


Figure 3.4.: Sketch of a hadron-hadron collision as simulated by a MC event generator [67]. The incoming partons are shown as green lines with arrows with the blue lines representing the initial state radiations. The central red blob represents the hard collision surrounded by components from simulated parton showers. The light green blobs represent the parton-to-hadron transition and the dark green blobs depict hadron decays. Yellow lines represent the soft photon radiations. The purple blob shows an underlying event.

teractions between the electroweak bosons. The electroweak corrections generally have a much smaller effect compared to the QCD corrections.

Since proton beams are composed of bunches with many protons, it is very likely that multiple interactions take place in a single bunch crossing. Therefore, multiple proton-proton interactions are also included in the simulation process. To simulate the pile-up in the detector, actual data from minimum bias events is overlaid in the MC after the last step of detector simulation

As a result of colour confinement, the field energy between the isolated quarks and gluons increases until it is energetically more favourable for the production of another $q\bar{q}$ pair. This iterative procedure, also known as *parton shower*, results in a spray of particles which propagate at high energies through the detector. As soon as the initial energy is below the requirement for additional parton creation, they combine to form colour singlet hadrons. This process is known as *hadronisation* and is prohibitively non-

perturbative for matrix element calculations except for the hardest partons. This calls for the need of parton shower models, which approximate the process by calculating the average number of radiation lengths for which no additional particles are radiated by the parton. These parton shower models can be implemented with programs such as HERWIG [71], SHERPA [68] and PYTHIA [70].

The final step to make the simulated events resemble real data events is a simulation of the detector. This requires detailed information of the involved magnetic fields and a precise model of the detector's geometry along with the materials used in its construction. This tracks the simulated particles' trajectory through the various detector components and creates readout signals that a real event would have generated. For the ATLAS detector, GEANT4 [72] is the simulation package used for the full simulation and ATLFAST-II (AFII) [73] is used for fast simulation, since the full simulation is generally computer intensive and highly time consuming. However, since the AFII setup only has a complete simulation of the inner detector and the muon spectrometer, it needs to be validated with the full simulation for a subset of samples. A list of all the MC samples used in this thesis can be found in Appendix C.

3.4. Jets in Boosted Topologies

Since quarks and gluons hadronise very soon after their production, they are observed in the detector via the shower of these hadrons as energy deposits in the calorimeters and, if charged, as tracks in the inner detector. Ideally, all the hadrons originating from the same parton should be reconstructed as a single object called a *jet*.

For the reconstruction of jets, sequential jet clustering algorithms are generally used [74–76] which are commonly based on the metric

$$d_{ij} = \min\left[p_{T,i}^{2p}, p_{T,j}^{2p}\right] \frac{\Delta R_{i,j}^2}{R^2} \quad \text{with} \quad p = \begin{cases} +1, & k_t \\ 0, & \text{Cambridge/Aachen} \\ -1, & \text{anti-}k_t \end{cases}, \tag{3.5}$$

where p determines the type of the algorithm, R is known as the size parameter of the jet, and $\Delta R_{i,j}$ is defined in the y-ϕ plane instead of the η-ϕ plane, contrary to Equation 3.4. The algorithm combines constituents i and j to a proto-jet after calculating a cut-off value $d_{iB} = p_{T,i}^{2p}$, along with d_{ij} for every pair of constituents. This proto-jet is then added to a modified list of constituents if $d_{ij} < d_{iB}$. Otherwise i is labelled as a jet and removed from the list of constituents. This procedure is then iterated for the modified list of

constituents. In ATLAS, the standard jet collection is created using the anti-k_t algorithm with $R = 0.4$. This is typically sufficient to reconstruct a single parton.

Since this thesis focusses on heavy resonances, which can decay into massive intermediate decay products with large transverse momenta, the particles in the final state from the decay of the intermediate decay products are close together in the detector. The effect of the p_T of such intermediate products on the mean angular separation between the final state particles can be approximated as [77]

$$\Delta R \approx \frac{2m}{p_T}, \tag{3.6}$$

where m is the mass of the resonance. For cases with a low-p_T intermediary particle, the decay products are well separated in the detector and can be individually reconstructed. This is known as the *resolved* topology. On the other hand, decay products from high-p_T resonances are *boosted*, i.e. they tend to appear so close to each other in the detector that individual reconstruction is no longer possible. Therefore, for boosted topologies, large-R jets are used, which have a larger size parameter and can contain all decay products.

The large-R jet collection also uses the anti-k_t algorithm but with $R = 1.0$. However, due to the larger radius, a larger number of unwanted particles from pile-up processes are also included in the jet. To control this, a procedure called *jet trimming* is performed [78]. This procedure first reclusters the large-R jet constituents with the k_t algorithm ($R = 0.2$) in smaller subjets and then removes every subjet having less than 5% of the total p_T of the large-R jet.

The large-R jets also contain information about the incident partons in the form of substructure, which can be used to gain information about the decay of the heavy particle forming a large-R jet. The first substructure variable is the N-subjettiness, τ_N [79], defined as

$$\tau_N(\beta) = \frac{\sum_{i \in J} p_T, i \cdot \min\left(\Delta R_{k,i}^{\beta} | k = 1, 2, \ldots, N\right)}{\sum_{i \in J} p_{T,i} \cdot R^{\beta}}, \tag{3.7}$$

where the jet constituents i are clustered in N subjets, and the k in $\Delta R_{k,i}$ denotes the jet axis of the k^{th} subjet. Here, β is a positive, tunable parameter. For cases with smaller distances between all constituents and their nearest subjet axis, τ_N has a smaller value. However, if an original subjet is missing, the distances are significantly larger. Thus, for practical purposes, the ratio $\tau_N/\tau_{N-1} \ll 1$ indicates that a large-R jet is better described by N subjets rather that $N - 1$ subjets. This indicates the presence of N initial partons.

Energy correlation functions (ECFs) [80] are another set of substructure variables de-

fined as

$$\begin{aligned} \mathrm{ECF}_0(\beta) &= 1 \\ \mathrm{ECF}_n(\beta) &= \sum_{j_1<\cdots<j_n\in J} \left(\prod_{i=1}^{n} p_{T,j_i}\right) \left(\prod_{i=1}^{n-1}\prod_{k=i+1}^{n} \Delta R_{j,j_k}\right)^{\beta}. \end{aligned} \tag{3.8}$$

In contrast to N-subjettiness, ECFs are based only on energy and angular information of particles within the jet and do not rely on an explicit ordering of constituents that subjets require. Again, the ratios between these functions prove to be useful. These are defined as

$$r_n(\beta) = \frac{\mathrm{ECF}_{n+1}(\beta)}{\mathrm{ECF}_n(\beta)}, \tag{3.9}$$

as well as

$$C_n(\beta) = \frac{r_n(\beta)}{r_{n-1}(\beta)} = \frac{\mathrm{ECF}_{n+1} \times \mathrm{ECF}_{n-1}}{\mathrm{ECF}_n^2} \tag{3.10}$$

and

$$D_n(\beta) = \frac{\mathrm{ECF}_{n+1} \times \mathrm{ECF}_{n-1} \times \mathrm{ECF}_1^n}{\mathrm{ECF}_n^3}. \tag{3.11}$$

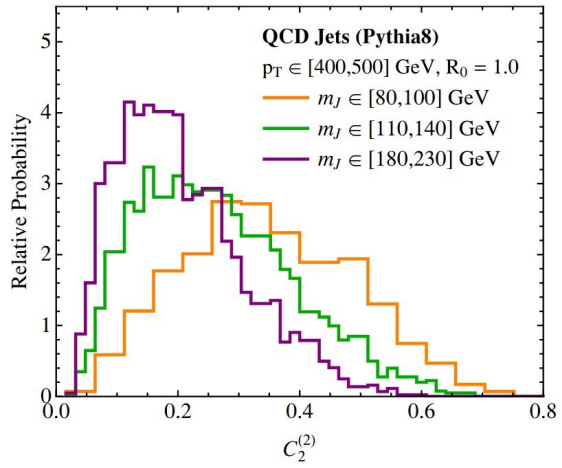
In previous years, it has been shown that the variables $C_2(\beta)$ and $D_2(\beta)$ are useful in identifying 2-prong jets ($n = 2$) [80–82]. Examples of distributions of the $C_2(2)$ variable with different jet size parameters are also shown in Figure 3.5, which compares the distributions of jets originating from QCD and those from a Z boson or Higgs boson. Distributions peaking at lower $C_2(2)$ values are more likely to be 2-prong jets.

Track-Assisted Reclustered Jets

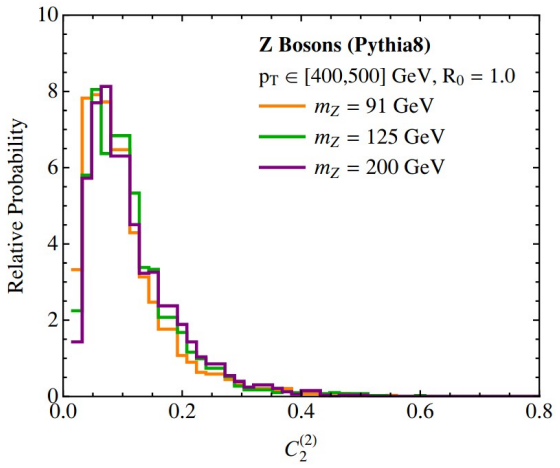
Due to the coarse angular resolution of the ATLAS calorimeter, the reconstruction of jet substructure variables from large-R jets becomes increasingly difficult for boosted topologies with jets appearing very close in the detector, especially when using AFII, which poorly models the substructure compared to data. To overcome this limitation, the jet reconstruction can be assisted using track information from the ID, which provides a much better angular resolution. ATLAS employs the track-assisted (TA) jet mass for standard large-R jets as a standard procedure. Conventionally, the mass of a jet J reconstructed in the calorimeter is defined as

$$m^{\mathrm{calo}} = \sqrt{\left(\sum_{i\in J} E_i\right)^2 - \left(\sum_{i\in J} \vec{p}_i\right)^2}, \tag{3.12}$$

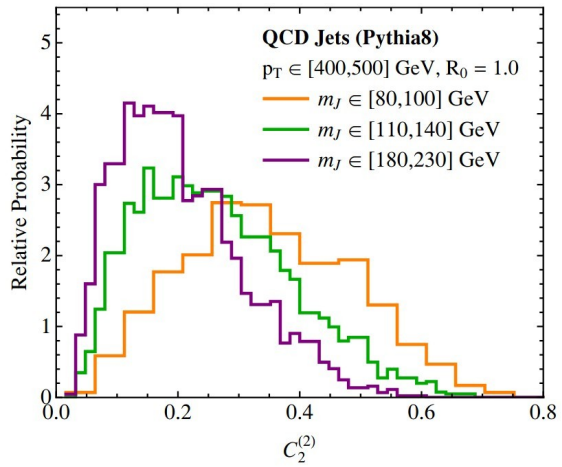
where E_i and $\vec{p}_i$ are the energy and momentum of the i^{th} jet component. In order to include track information, the particle tracks from the ID are associated to the calorimeter



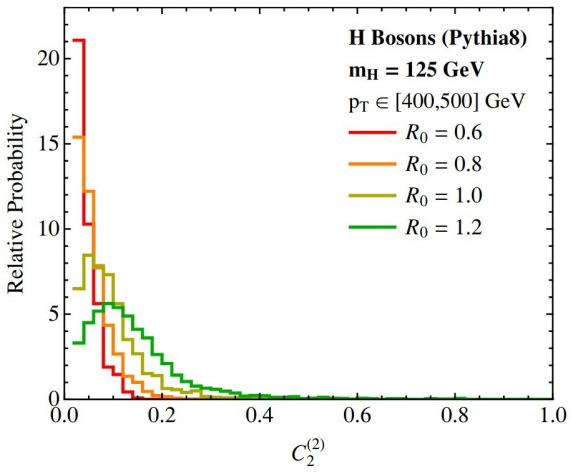
(a) QCD Jets



(b) Z bosons decaying to jets



(c) QCD Jets decaying to $b\bar{b}$



(d) H bosons decaying to $b\bar{b}$

Figure 3.5.: Distributions of $C_2(2)$ variable with different jet size parameters comparing QCD jets with jets originating from a Z boson or Higgs boson [80]. Distributions peaking at lower $C_2(2)$ values are more likely to be 2-prong jets.

clusters through a procedure known as *ghost-association* [83]. For this, the scaled tracks are added to the input list of the corresponding cluster algorithm after setting the p_T of the tracks to an infinitesimal value. The scaling ensures the addition of only angular information without distorting the jet reconstruction. Any ghosts that are included in the jet are then linked to the original tracks that are then considered associated to the jet.

The TA jet mass is then defined as [84]

$$m^{\mathrm{TA}} = \frac{p_T^{\mathrm{calo}}}{p_T^{\mathrm{track}}} \cdot m^{\mathrm{track}}, \tag{3.13}$$

where m^{track} and p_T^{track} are the mass and transverse momentum of the associated tracks. The reweighting $p_T^{\mathrm{calo}}/p_T^{\mathrm{track}}$ ensures the inclusion of the energy of neutral hadrons to which the ID is not sensitive. Furthermore, m^{TA} is also useful when using the AFII detector simulation, since the parametrised detector response of the calorimeter results in a poorly modelled calorimeter-based mass.

However, it was observed that, while resolutions for both m^{calo} and m^{TA} degrade at high p_T, for W and Z boson jets, m^{TA} has a superior resolution to m^{calo} for $p_T > 1\,\mathrm{TeV}$. The charged-to-neutral fluctuations dominate the resolution of m^{TA}, which is worse than that of m^{calo} for $p_T < 1\,\mathrm{TeV}$ [84]. In order to further improve the resolution of the large-R jet mass, the combined mass [84], m^{comb}, is used. This combines the good resolution of m^{calo} for low p_T jets with that of m^{TA} for high-p_T jets by an inverse-resolution weighted mean using the expected jet mass resolutions determined from MC simulations. This is mathematically defined as

$$m^{\mathrm{comb}} = \frac{\sigma_{\mathrm{calo}}^{-2}}{\sigma_{\mathrm{calo}}^{-2} + \sigma_{\mathrm{TA}}^{-2}} \cdot m^{\mathrm{calo}} + \frac{\sigma_{\mathrm{TA}}^{-2}}{\sigma_{\mathrm{calo}}^{-2} + \sigma_{\mathrm{TA}}^{-2}} \cdot m^{\mathrm{TA}}, \tag{3.14}$$

where σ_{calo} and σ_{TA} are the calorimeter-based jet mass resolution function and the track-assisted mass resolution function respectively.

This procedure can also be extended to further jet substructure variables, however, it does not account for the local differences between energy deposits from neutral and charged hadrons within the jet. This is resolved by the use of another jet collection algorithm known as track-assisted reclustered (TAR) jets [85], by applying local corrections to the p_T of close-by tracks.

The TAR jet algorithm uses tracks and calibrated anti-k_t $R = 0.2$ jets to reconstruct large-R anti-k_t jets. The large-R jets are formed by reclustering the $R = 0.2$ jets, which are then trimmed. The tracks are then matched to the $R = 0.2$ jets that survive the trimming procedure. This is followed by the scaling of the p_T of each constituent track using the p_T of the constituent jet to which it is matched. This procedure is also depicted graphically in Figure 3.6. The mass, m^{TAR}, and other substructures are then calculated using these constituent tracks. This approach offers an excellent resolution from tracks, the flexibility of using calibrated inputs and a straightforward method to calibrate and assign uncertainties to jet observables by the bottom-up propagation of uncertainties. The

resolution comparison of the jet mass and $D_2(1)$ substructure variable between different jet definitions is also shown in Figure 3.7 for W jets and $H \to b\bar{b}$ jets. The resolutions are defined using the interquantile range (IQnR) method. For $D_2(1)$ it is defined as one half of the 50% IQnR, whereas for mass it is defined as one half of the 68% IQnR divided by the median. It can be seen that TAR jets (orange) perform the best overall, particularly for the $D_2(1)$ substructure variable.

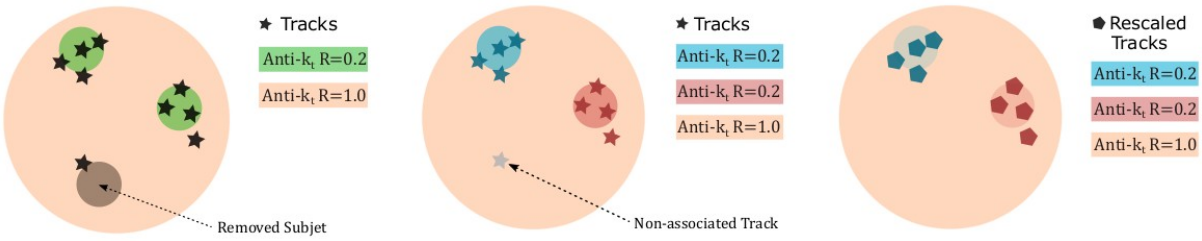
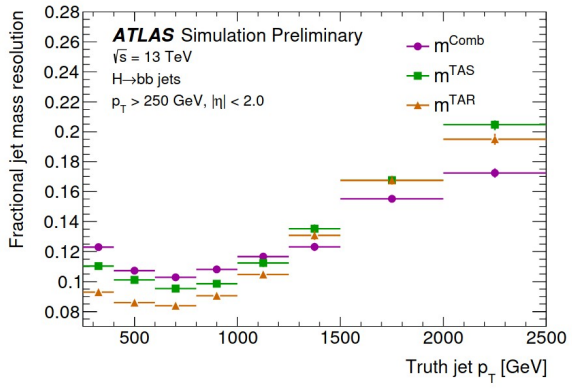
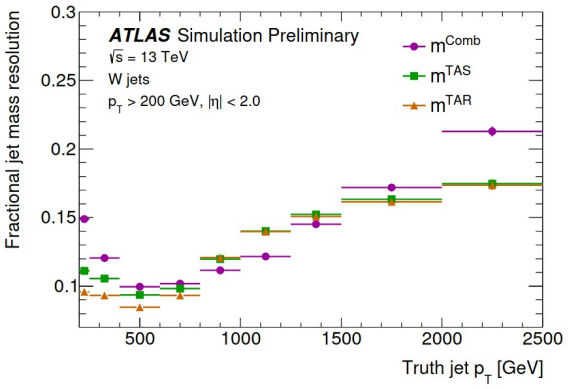


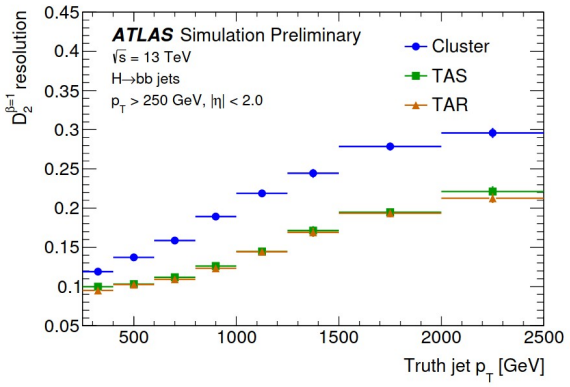
Figure 3.6.: Schematic diagram depicting the track-assisted reclustering algorithm.



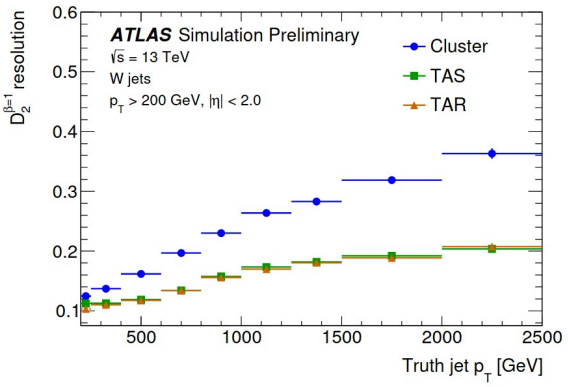
(a) Jet mass for $H \to b\bar{b}$ jets



(b) Jet mass for W jets



(c) $D_2(1)$ variable for $H \to b\bar{b}$ jets



(d) $D_2(1)$ variable for W jets

Figure 3.7.: p_T-dependent resolution comparison of jet mass and $D_2(1)$ variable between different jet definitions for W jets and $H \to b\bar{b}$ jets [85]. For $D_2(1)$ the resolution is defined as one half of the 50% IQnR, whereas for mass it is defined as one half of the 68% IQnR divided by the median.

4. $X \to HH(SH) \to b\bar{b}WW^{(*)}$ Searches in ATLAS Data

> "Measure what can be measured, and make measurable what cannot be measured."
>
> - Galileo Galilei

This chapter describes the analysis strategies followed for this thesis. This includes a detailed description of the event topology of the $b\bar{b}WW^{(*)}$ decay channel along with the definitions of the final state objects and the selection cuts applied in order to improve the sensitivity of the analysis.

4.1. The $b\bar{b}WW^{(*)}$ Event Topology

As already mentioned at the end of Section 2.3, this thesis focusses on the scenario where two Higgs bosons are produced in pp collisions and one decays into a pair of bottom quarks, while the other decays into a pair of W bosons. This decay channel has the second highest $\mathcal{BR}$ for the HH scenario and the highest $\mathcal{BR}$ for the SH scenario for $m_S > 135\,\text{GeV}$ (see Section 2.3). Since the W boson is also an unstable particle, it can further decay, either hadronically into a pair of quarks, or leptonically into a charged lepton and a neutrino. Depending on how the two W bosons decay, there are three generalised final state signatures, which are:

- *0-lepton*: Both the W bosons decay hadronically.
- *1-lepton*: One W boson decays hadronically, while the other decays leptonically.
- *2-lepton*: Both the W bosons decay leptonically.

This thesis puts its focus on the 1-lepton final state.

In the case of the HH signal samples, this thesis considers the mass of the heavy scalar resonance, X, to be $0.8\,\text{TeV} \leq m_X \leq 5.0\,\text{TeV}$. Due to such high masses, the two Higgs

bosons are produced with a very high transverse momenta. This causes the final state objects to overlap with each other and, therefore, the signature falls into the boosted regime. Thus, the two b-jets are collected in one large-R jet, while the two quark jets from the hadronically decaying W boson (W_{had}) are reconstructed as another large-R jet. In most cases, the lepton (ℓ) from the leptonically decaying W boson (W_{lep}) also overlaps with the large-R jet from the W_{had}. The entire decay chain is also graphically depicted in Figure 4.1.

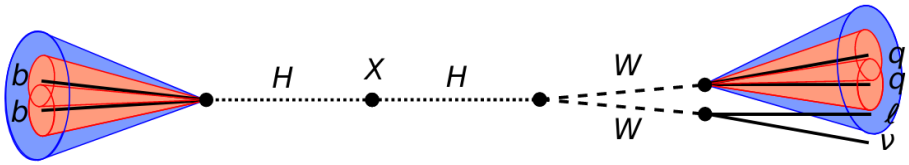


Figure 4.1.: Graphical representation of the boosted $X \to HH \to b\bar{b}WW^* \to b\bar{b}qq\ell\nu$ topology as seen in the detector. The heavy resonance X shown in the center decays into a pair of Higgs bosons, which are seen back-to-back. They further decay into a pair of b-jets and two W bosons. The W bosons then finally decay into two quark jets, a lepton and a corresponding neutrino.

For the case of SH signal samples, the mass of X, m_X, as well as of the Higgs-like scalar boson S, m_S, are free parameters. This thesis considers $0.75\,\text{TeV} \leq m_X \leq 3.0\,\text{TeV}$ and $2m_W < m_S < m_X - m_H$ as the mass ranges for these scalars. The two variable masses have a crucial impact on the event topology, which will be discussed in more detail in Section 5.1. This chapter will mostly focus on the fully boosted topology based on HH samples. The decay chain for the SH signal samples is shown in Figure 4.2.

Since only two large-R jets, one of them being b-tagged, a lepton and missing transverse energy, E_T^{miss}, are observed as the final state objects from the signal process, there are certain background processes that imitate this signature in the detector. The main background process considered in this analysis include the $t\bar{t}$ process with one W boson decaying leptonically and the other hadronically, and the W+jets process with the W boson decaying leptonically. The corresponding detector signatures for these two processes are also shown in Figure 4.3. Other backgrounds include the non-prompt lepton backgrounds (largest contribution from multijet production, hence also referred to as QCD background), the Z+jets process, the SM diboson production and the single top process.

4.2. Object Definition

This section provides the object reconstruction definitions that are used for the various objects relevant for this analysis.

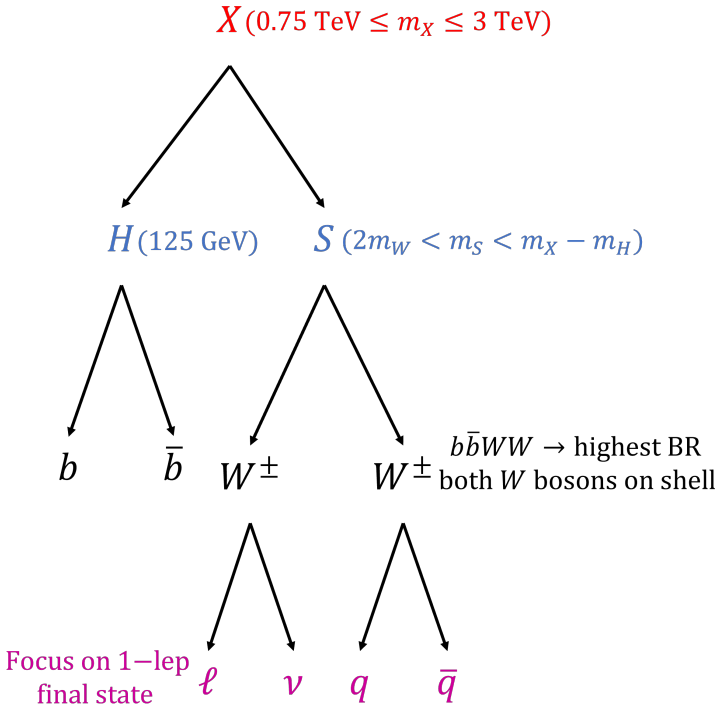
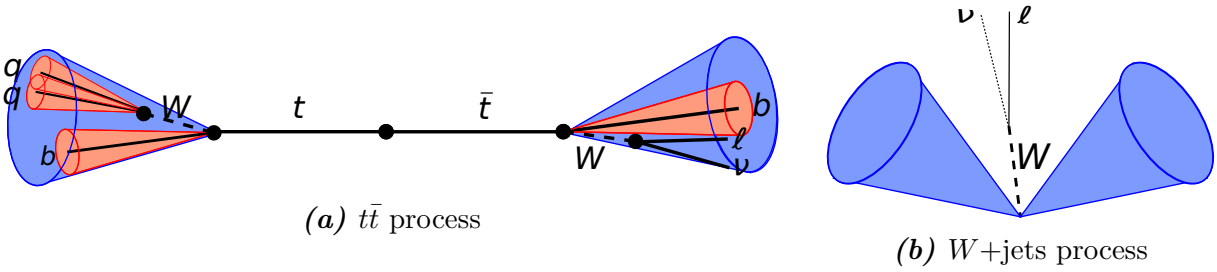


Figure 4.2.: Graphical representation of the $X \to SH \to b\bar{b}WW \to b\bar{b}qq\ell\nu$ decay chain.



(a) $t\bar{t}$ process

(b) W+jets process

Figure 4.3.: Graphical representation of the main background processes as seen in the detector.

Jets

This analysis uses various types of jet definitions for different purposes. These are defined as below:

- *Standard jets* recluster the jets using the anti-k_t algorithm [76] on particle flow objects with $R = 0.4$ [86]. These jets must fulfil the requirements of $p_T > 20\,\text{GeV}$, $|\eta| < 4.5$ and pass the medium working point of the jet vertex tagger [87]. Standard jets are only used for reconstructing the missing transverse energy as well as in the overlap removal, to avoid double counting of the leptonic and hadronic energies, since $R = 0.4$ is too small to collect all hadronic decay products in a single jet

but also too large to resolve the individual hadronic decay products in a boosted topology.

- *Small-R jets* build the jets from locally calibrated topological clusters from the calorimeter using the anti-k_t algorithm but with $R = 0.2$ [88]. As inputs for the TAR jets, only small-R jets satisfying $p_T > 15\,\text{GeV}$ and $|\eta| < 2.5$ will be considered.

- *Large-R jets*, as in the case of small-R jets, are also built using the anti-k_t algorithm on locally calibrated topological clusters but with $R = 1.0$. In order to make the jets more robust against pile-up and underlying event contributions, jet trimming [78] is also applied as explained in Section 3.4. As per the standard ATLAS recommendation, these jets are required to satisfy $p_T > 200\,\text{GeV}$ and $|\eta| < 2.0$. In this analysis, the large-R jets are used mainly in the trigger selection and to preserve the orthogonality between different decay channels for the Higgs boson pair production searches.

- *VR Track jets* refer to the variable-R (VR) track jets [89] and are necessary for flavour tagging in boosted topologies. These cluster tracks by using the anti-k_t algorithm with R dependent on the jet p_T as $R(p_T) = \rho/p_T$, with $\rho = 30$ determining the scale with which the effective jet size decreases. The jet size is bounded between $R = 0.02$ and $R = 0.4$. These values have been found to be optimal for boosted $H \to b\bar{b}$ tagging [90]. The jets are also required to satisfy $p_T > 10\,\text{GeV}$, $|\eta| < 2.5$ and to have at least two constituents. Furthermore, events are vetoed for cases with $\Delta R(\text{jet}_i, \text{jet}_j) < \min\left[R_{\text{jet}_i}, R_{\text{jet}_j}\right]$, where i runs over all jets passing the above-mentioned requirements and being considered for b-tagging, while j runs over all jets also passing the same requirements but with a loosened $p_T > 5\,\text{GeV}$ with $i \neq j$.

- *Track-Assisted Reclustered (TAR) jets* [85] (see Section 3.4) are chosen as the main jet collection for the event reconstruction due to their optimised performance in dense environments. Contrary to the other jet collections, TAR jets are reclustered from calibrated small-R jets after performing overlap removal against electrons. Thus, the TAR jets themselves do not need to be calibrated, allowing great flexibility in clustering algorithm, size parameter and grooming. For this analysis the anti-k_t algorithm with a fixed $R = 0.75$ (see Figure 4.5), and trimming with $f_{\text{Cut}} = 0.05$ applied were found to be optimal. Additionally, TAR jets must satisfy $p_T > 100\,\text{GeV}$ and $|\eta| < 2.0$. For calculating quantities such as the m_{TAR}, it is also necessary that every track has at least one constituent jet associated to it. In order to allow flavour tagging, VR track jets are ghost-associated [83] to the TAR jets providing a robust

matching procedure making use of the catchment area of the untrimmed TAR jet.

Electrons

Electrons are reconstructed by matching energy clusters in the ECAL to tracks in the Inner Detector [91]. Electron candidates are required to satisfy the following criteria:

- `Medium` likelihood ID
- `TightTrackOnly` isolation, which uses the tracks from the inner detector to ensure that the electron is well isolated while keeping the cone radius dependent on p_T of the electron.
- $p_T > 10\,\text{GeV}$ and $|\eta| < 2.47$ excluding the crack region between the barrel and end-cap calorimeters ($1.37 < |\eta| < 1.52$)
- $|d_0^{\text{sig}}| < 5$ and $|z_0 \sin\theta| < 0.5\,\text{mm}$ to stay consistent with the primary vertex as per standard ATLAS recommendations.

Muons

Muons are reconstructed by matching tracks from the Inner Detector to the tracks in the Muon Spectrometer [92]. Muon candidates must fulfil the following requirements:

- `Medium` ID
- `FCTightTrackOnly` (Fixed-Cut Tight Track Only) isolation, which again works in a similar fashion as for the case of electrons.
- $p_T > 10\,\text{GeV}$ and $|\eta| < 2.5$
- $|d_0^{\text{sig}}| < 3$ and $|z_0 \sin\theta| < 0.5\,\text{mm}$, again, to stay consistent with the primary vertex as per standard ATLAS recommendations.

Missing Transverse Energy

The missing transverse energy (E_T^{miss}) is defined as the negative vector sum of the transverse momenta of all other reconstructed objects, i.e. standard jets, electrons and muons [93]. Furthermore, soft terms not included in above-mentioned objects are still taken into account by the track soft term (TST) algorithm by adding the momentum of soft tracks originating from the hard-scatter vertex. Thus, these terms only account for low p_T charged particles but not for neutral ones.

Overlap Removal

Since the object reconstruction algorithms of ATLAS run independently of each other on all tracks and energy clusters from the calorimeter, energies can be included in multiple reconstructed objects when they overlap geometrically. As an attempt to avoid this, a ΔR[1] based overlap removal as per standard ATLAS recommendations is applied in the following sequential steps:

1. Reject a muon, i.e. remove from consideration for the subsequent steps, if it deposits significant amount of energy in the calorimeters, also known as calorimeter-tagged (CT) muons, and shares the inner detector track with an electron.

2. Reject an electron if it shares the inner detector track with a muon.

3. Due to issues in the particle flow algorithm a special overlap removal for particle flow jets against muons is needed. Reject a particle flow jet if ΔR(jet,muon) < 0.4 and

   - the muon is isolated and the sum of transverse energies of topological clusters in a cone of $\Delta R = 0.4$ around the muon, $E_{T,0.4}^{\text{topo}} < y_0$,
   - the muon p_T exceeds a certain fraction of the summed track p_T of the jet, $\frac{p_T^{\mu}}{\sum_{\text{tracks}} p_T^{\text{tracks}}} > x_2$, or
   - a mixture of both requirements, $E_{T,0.4}^{\text{topo}} < y_0 + \frac{y_2 - y_1}{x_2 - x_1} \cdot \left(\frac{p_T^{\mu}}{\sum_{\text{tracks}} p_T^{\text{tracks}}} - x_1 \right)$.

   The values that were found to be optimal for $n_{\text{tracks in jet}} < (\geq)\ 4$ are as follows: $x_1 = 0.7\,(0.6)$, $x_2 = 0.85\,(0.9)$, $y_0 = y_1 = 15\,(5)$ GeV and $y_2 = 30\,(30)$ GeV.

4. A jet is rejected if ΔR(jet,electron) < 0.2.

5. An electron is rejected if ΔR(jet,electron) $< \min\left[0.4,\ 0.04 + \frac{10\,\text{GeV}}{p_T^e}\right]$.

6. A jet is rejected if $n_{\text{tracks in jet}} < 3$ and either the muon is ghost-associated to the jet or ΔR(jet,muon) < 0.2.

7. A muon would be rejected if ΔR(jet,muon) $< \min\left[0.4,\ 0.04 + \frac{10\,\text{GeV}}{p_T^{\mu}}\right]$.

8. Finally, a small-R jet is rejected against an electron if $\Delta R < 0.2$.

[1] In the overlap removal, ΔR is defined in y-ϕ plane instead of the η-ϕ plane.

b-Tagging

Since the decay of the bottom quark is suppressed by the CKM Matrix [94, 95], it has a long enough lifetime to be seen with the spatial resolution of the detectors before it decays into a charm quark. This leads to a secondary vertex which is displaced from the primary vertex. Because of this distinct signature it is possible to extract the flavour of the quark that initiates the jet by the so-called flavour tagging. There are two main algorithms available: `MV2c10` [96] and `DL1r` [97]. In this analysis, b-tagging is performed on VR track jets, for which the `DL1r` working point (WP) with a 77% efficiency in $t\bar{t}$ MC was found to perform best.

Event Reconstruction

For the reconstruction of a signal event, the lepton candidate is defined as a muon passing the above-mentioned quality recommendations. Only the muon channel is used due to lower sensitivity and complications in the electron channel (see Section 4.3). The transverse component of the neutrino candidate is reconstructed as the total E_T^{miss} in the event. Further, the three p_T-leading jets in the event are considered for the jet classification. Of these, the TAR jet closest in ΔR to the lepton candidate is classified as the W_{had} candidate, whereas the p_T-leading TAR jet of the remaining jets is classified as the $H \to b\bar{b}$ candidate. The entire event reconstruction is summarised in Table 4.1.

Table 4.1.: Reconstruction definitions of all final state objects in a signal event.

Object	Definition/Reconstruction
ℓ	Muon passing quality recommendations
ν	E_T^{miss} for transverse component
W_{had}	TAR jet closest to lepton
$H \to b\bar{b}$	p_T-leading TAR jet which is not W_{had}

4.3. Event Selection

This section aims to define and motivate the selection cuts applied to suppress the background events and, at the same time, gain sensitivity towards the signal events. Since a basic skeleton of the event selection already existed at the time this thesis began, this section starts with stating this initial event selection, and then motivates the changes and

optimisations made over the course of the investigations. In order to avoid any ambiguity, the finalised event selection is also explicitly stated at the end of this section. This includes definitions for the preselection, control regions (CRs), validation regions (VRs) and signal regions (SRs).

Initial Event Selection

Given the event topology under investigation with two TAR jets in the signal events corresponding to the $H \to b\bar{b}$ and W_{had} candidates, large-R jets are great candidates for selecting signal events. The analysis uses high-level triggers (HLT) on single large-R jets in the preselection for all the region definitions. The jet p_T thresholds of these HLTs vary depending on the year and span from 360 GeV to 480 GeV [98]. To ensure that only events in the trigger plateau are selected, an additional cut is placed, requiring the p_T-leading TAR jet to satisfy $p_T > 500\,\text{GeV}$. In addition to this, exactly one signal lepton is required which could either be an electron or a muon passing quality recommendations. The preselection also includes the requirement for the event to have at least two TAR jets with $R = 1.0$, which is necessary for the jet classification in the event reconstruction to be well defined.

Three CRs are also defined, one for each dominant background:

- The $t\bar{t}$ CR requires exactly two b-tagged TAR jets in the event since in a $t\bar{t}$ event both the TAR jets would mostly have a b-jet from the decay of a top quark. An additional requirement of $m_{\text{TAR}}^{W_{\text{had}}} < 20\,\text{GeV}$ is also applied. However, this does not mean that the W_{had} candidate in a $t\bar{t}$ event cannot be on-shell. Instead, since the W_{had} candidate is defined as the TAR jet closest to the lepton, this will mostly be a b-jet in a $t\bar{t}$ event.

- The W+jets CR is defined by the requirement of exactly zero b-tagged TAR jets in the event. This requirement is useful since the W boson must decay leptonically to imitate the signal signature and the other jets, particularly from pile-up events, are rarely b-jets. Additionally, the events must satisfy $m_{\text{TAR}}^{H\to b\bar{b}} < 90\,\text{GeV}$ in order to reduce signal contamination, and the transverse mass, $m_T^{W_{\text{lep}}}$ must satisfy $60\,\text{GeV} < m_T^{W_{\text{lep}}} < 120\,\text{GeV}$ to ensure the presence of a leptonically decaying W boson.

- Finally, the QCD CR is defined exactly like the W+jets CR except, in this case, the $m_T^{W_{\text{lep}}}$ must satisfy either $m_T^{W_{\text{lep}}} < 60\,\text{GeV}$ or $m_T^{W_{\text{lep}}} > 120\,\text{GeV}$ in order to maintain orthogonality between both CRs.

For the signal regions, exactly one b-tagged TAR jet is required in the event, which also ensures orthogonality between the SRs and the CRs. In addition to this, the events must satisfy $90\,\text{GeV} < m_{\text{TAR}}^{H\to b\bar{b}} < 140\,\text{GeV}$ in order to contain the mass peak of the $H \to b\bar{b}$ candidate. Finally, the b-tagged jet is required to also be characterised as the $H \to b\bar{b}$ candidate. There is a further split into four signal regions based on the lepton flavour (electron or muon) and the b-tag multiplicity of the $H \to b\bar{b}$ candidate (exactly one b-tag or at least two b-tags). The event distributions of in the CRs and SRs can also be found in Ref. [98].

Background Estimation

While the other backgrounds can be simulated well with MC generators, the non-prompt lepton background requires a data-driven approach for a reasonable estimate. For this the number of loose and tight leptons, given as

$$N_L = N_{\text{prompt}} + N_{\text{QCD}} \quad \text{and} \quad N_T = \epsilon \times N_{\text{prompt}} + f \times N_{QCD}, \tag{4.1}$$

are used. Here, N_{prompt} is the number of prompt leptons and N_{QCD} is the number of non-prompt leptons. The prompt and non-prompt leptons have a real and fake efficiency of ϵ and f, respectively, to pass the tight lepton selection, where tight leptons are the signal leptons such that the QCD estimate in the SRs can be obtained by weighting each data event with

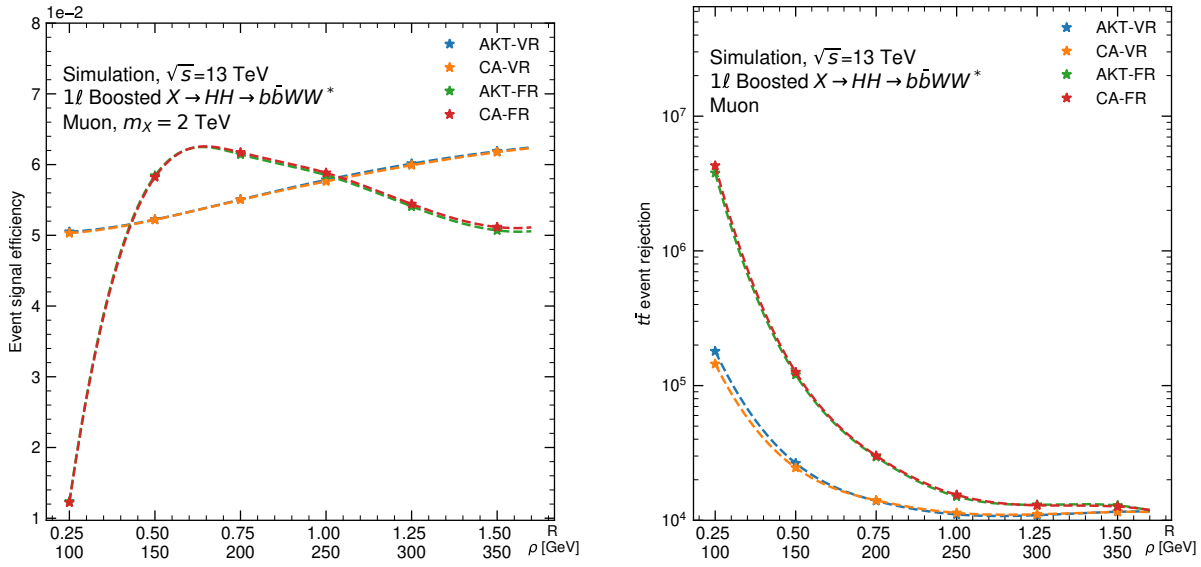
$$w(N_T, NLnT) = f \times N_{\text{QCD}} = \frac{(\epsilon - 1)f}{\epsilon - f} N_T + \frac{\epsilon f}{\epsilon - f} N_{LnT}. \tag{4.2}$$

Considering events with exactly one loose lepton, they can either pass the tight (N_T) or the loose-not-tight (N_{LnT}) requirement. The real efficiency ϵ is obtained from prompt lepton MC simulation in the SRs, where the reconstructed lepton must be truth matched to the real prompt lepton. The fake efficiency f, on the other hand, is measured in the QCD CR using data, from which the prompt lepton MC backgrounds are subtracted. Further details about the background estimation can be found in Ref. [98].

Optimisation

An additional cut at on the ΔR between the lepton candidate and the closest of the three p_T-leading TAR jets, $\min\left[\Delta R\left(\ell, J\right)\right] \leq 1.0$, is also introduced in order to ensure that the selected events exhibit boosted topologies. This will be discussed in much more detail in Section 5.3.

Next, the size parameter of TAR jets was optimised to enhance the sensitivity. Rigorous studies were conducted to understand the effect of changing R in the TAR jets definition [98]. The conclusions can be summarised by Figure 4.4, which shows the event signal efficiency for $m_X = 2.0\,\text{TeV}$ and the background rejection efficiency for the $t\bar{t}$ background for the boosted 1-lepton final state of the $X \to HH \to b\bar{b}WW^*$ decay channel, particularly for the case where the lepton is a muon. The figure compares the performance of fixed radius (FR) and variable radius (VR) for the Cambridge-Aachen (CA) and anti-k_t (AKT) algorithms and conveys that the FR $R = 0.75$ TAR jets have a higher event selection and background rejection efficiencies compared to the $R = 1.0$ TAR jets. A smaller R was not considered in order to ensure the full reconstruction of a large-R jet. To reinforce this observation, a comparison between expected exclusion limits[2] with $R = 1.0$ and $R = 0.75$ is drawn. This can be seen in Figure 4.5, which shows the expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ as a function of m_X. The figure also makes it clear that TAR jets with $R = 0.75$ (black dashed line) perform slightly better than with $R = 1.0$ (red dashed line) and, therefore, supports the previous conclusion.



(a) Event signal efficiencies for $m_X = 2.0\,\text{TeV}$ ***(b)*** Rejection efficiencies for $t\bar{t}$ background

Figure 4.4.: Performance comparison of fixed-radius (FR) and variable-radius (VR) Cambridge-Aachen (CA) and anti-k_t (AKT) algorithms for the boosted 1-lepton final state of the $X \to HH \to b\bar{b}WW^*$ decay channel, particularly for the case where the lepton is a muon [98].

Validation regions are necessary to ensure good agreement between data and the background estimation. It is crucial that the phase space covered by the VRs is as close as

[2]The detailed description of the entire limit setting procedure will be given later in Chapter 6.

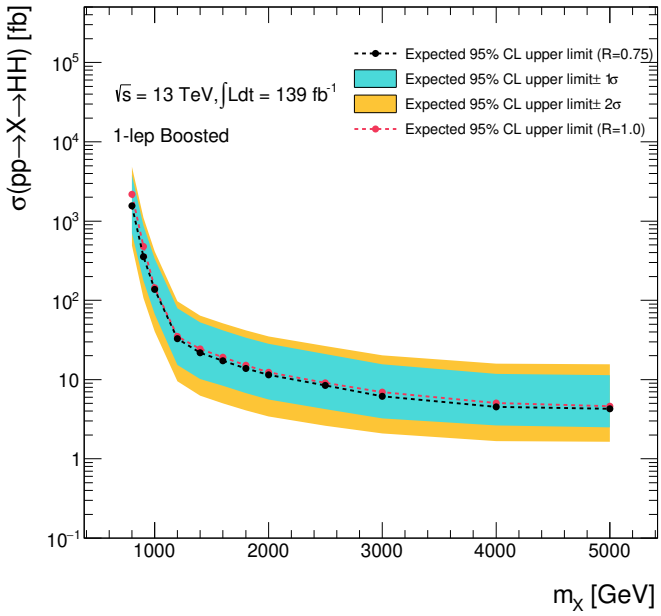
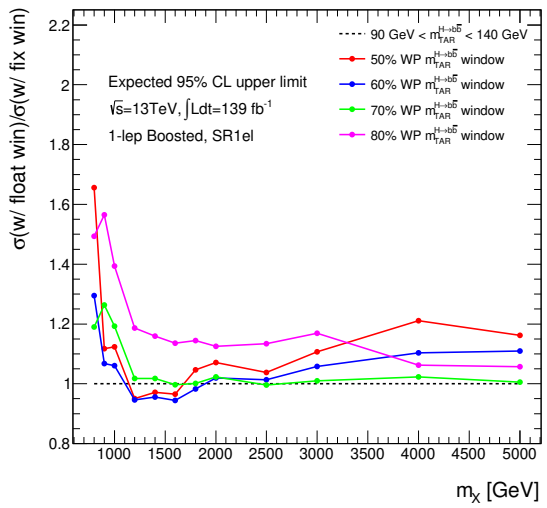
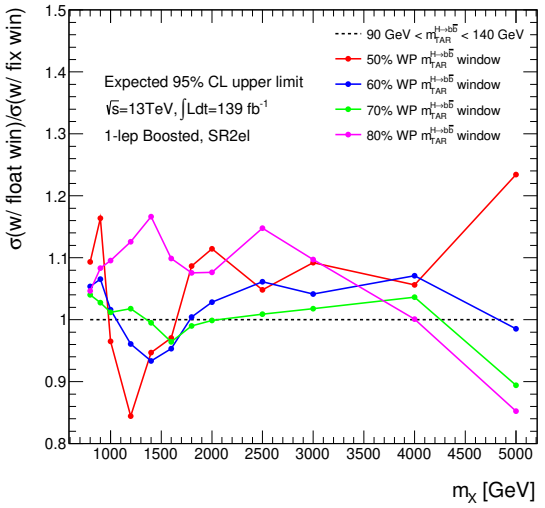


Figure 4.5.: Comparison of expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ between TAR jets with $R = 1.0$ (dark pink dashed line) and $R = 0.75$ (black dashed line). Only statistical uncertainties are included in shown exclusion limits.

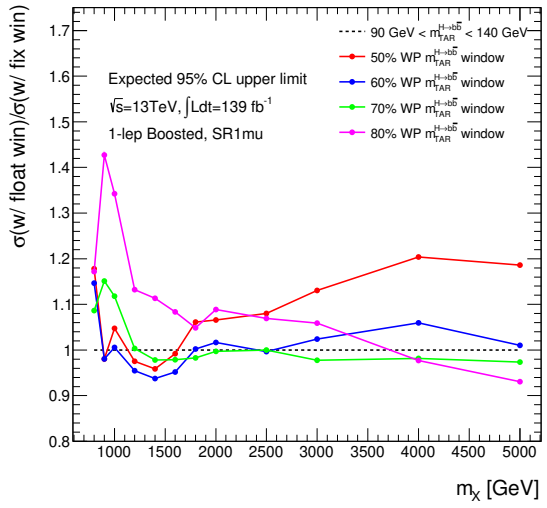
possible to the SRs, while at the same time maintaining minimal signal contamination to avoid any bias in the analysis. Hence, the VRs are also defined with exactly one b-tagged TAR jet in the event but with the cut on $m_{\text{TAR}}^{H\to b\bar{b}}$ in the SRs inverted. This would then result in one VR for every corresponding SR. However, to ensure low signal contamination in the VRs, varying $m_{\text{TAR}}^{H\to b\bar{b}}$ window cuts were proposed dependent on $p_T^{H\to b\bar{b}}$. To do this, all signal events for all HH samples passing the preselection were added together. The $p_T^{H\to b\bar{b}}$-dependent windows were then obtained about the median containing 50, 60, 70 and 80% of all signal events [98]. These windows will be referred to as the 50, 60, 70 and 80% WPs and are shown as a function of $p_T^{H\to b\bar{b}}$ in Appendix B.1. To compare the performance of these four WPs, again the expected exclusion limits are obtained in all four SRs individually. The ratio of these limits to the ones obtained by the fixed window cut on $m_{\text{TAR}}^{H\to b\bar{b}}$ are shown in Figure 4.6. From this figure, it can be seen that for most mass points, the 70% WP (green solid line) results in the best expected limits of the four WPs. Further it can also be seen, that this WP performs almost as well as the fixed $m_{\text{TAR}}^{H\to b\bar{b}}$ window (black dashed line). Hence the fixed cuts on $m_{\text{TAR}}^{H\to b\bar{b}}$ in the SR definition as well as in the definitions of the W+jets and QCD CRs are replaced by the varying 70% WP $m_{\text{TAR}}^{H\to b\bar{b}}$ window.



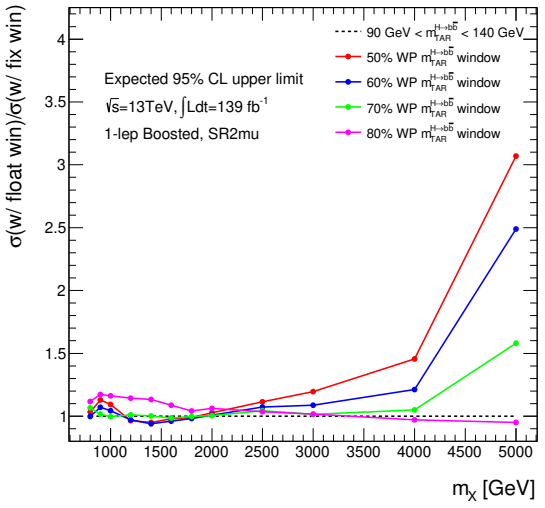
(a) SR1el



(b) SR2el



(c) SR1mu



(d) SR2mu

Figure 4.6.: Ratio of expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ with floating $m_{\text{TAR}}^{H\to b\bar{b}}$ window to those with fixed $90\,\text{GeV} < m_{\text{TAR}}^{H\to b\bar{b}} < 140\,\text{GeV}$ in all four SRs. Only statistical uncertainties are included in shown expected exclusion limits.

In order to verify the relative sensitivity of these eight regions, as defined in Table 4.2 by the $m_{\text{TAR}}^{H\to b\bar{b}}$ window (pass or fail), lepton flavour (electron or muon) and the b-tag multiplicity (exactly one or at least two), the exclusion limits in all these regions are compared. This can be seen from the expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ in Figure 4.7a, which shows that the regions with at least two b-tags (blue and magenta lines) perform better than the one with exactly one b-tag (red and green lines)

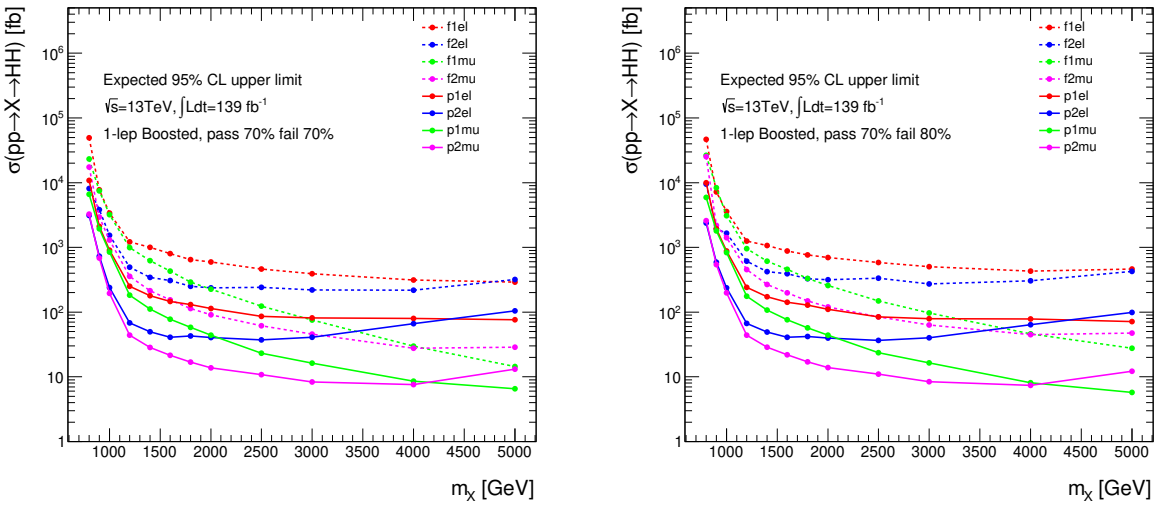
for most mass points except for $m_X = 5\,\text{TeV}$. This could be a result of the highly boosted topology, where the two b-jets cannot be resolved even as VR track jets. Moreover, to further reduce the signal contamination in the VRs, and hence minimise the bias, the definition of regions requiring events to fail the $m_{\text{TAR}}^{H\to b\bar{b}}$ window was changed from the 70% WP to the 80% WP. The expected exclusion limits for all individual regions are shown in Figure 4.7b. However, a striking observation is that the region requiring events to fail the $m_{\text{TAR}}^{H\to b\bar{b}}$ window with at least two b-tags for the muon channel (magenta dashed line) performs better than the region requiring events to pass the $m_{\text{TAR}}^{H\to b\bar{b}}$ window with exactly one b-tag for the electron channel (red solid line) for $m_X > 2.5\,\text{TeV}$. Since this is the region of high sensitivity for the analysis, the definition of SRs and VRs cannot be entirely based on passing or failing the $m_{\text{TAR}}^{H\to b\bar{b}}$ window and needs to be reconsidered.

Table 4.2.: Summary of all eight regions defined by $m_{\text{TAR}}^{H\to b\bar{b}}$ window, lepton flavour and the b-tag multiplicity.

Region	$m_{\text{TAR}}^{H\to b\bar{b}}$ window	$N_{b\text{-tags}}$	ℓ flavour
f1el	fail	$= 1$	e
f1mu	fail	$= 1$	μ
f2el	fail	≥ 2	e
f2mu	fail	≥ 2	μ
p1el	pass	$= 1$	e
p1mu	pass	$= 1$	μ
p2el	pass	≥ 2	e
p2mu	pass	≥ 2	μ

There are, however, various possible ways to combine subsets of the eight regions as SRs, whereas those not a part of this combination will be defined as VRs. An optimal trade-off is required between gaining sensitivity towards the signal events and having good VRs to verify the background modelling, since VRs are not a part of the final likelihood fits. Several possibilities are considered and their performance is compared relative to the nominal case where all regions requiring events to pass the $m_{\text{TAR}}^{H\to b\bar{b}}$ window are considered as SRs, while those requiring events to fail the $m_{\text{TAR}}^{H\to b\bar{b}}$ window are considered as VRs. This comparison is shown in Figure 4.8 and the possibilities that were considered are defined as follows:

- *4P+4F* (black line): All eight regions combined as SRs with no VRs available. This is considered only as a baseline scenario to compare the loss in sensitivity.
- *4P+2F* (red line): All four regions requiring events to pass the $m_{\text{TAR}}^{H\to b\bar{b}}$ window and both the muon channel regions requiring events to fail the $m_{\text{TAR}}^{H\to b\bar{b}}$ window combined



(a) pass 70% or fail 70% WP of $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window ***(b)*** pass 70% or fail 80% WP of $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window

Figure 4.7.: Comparison of expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ for all eight regions defined by the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window (pass or fail), lepton flavour (electron or muon) and the b-tag multiplicity (exactly one or at least two). Only statistical uncertainties are included in shown exclusion limits.

as SRs. Only the electron channel regions requiring events to fail the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window would be available as VRs.

- *4P+1F* (blue line): All four regions requiring events to pass the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window and the muon channel region requiring events to fail the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window with at least two b-tags combined as SRs. VRs include both electron channel and at least two b-tag muon regions requiring events to fail the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window.

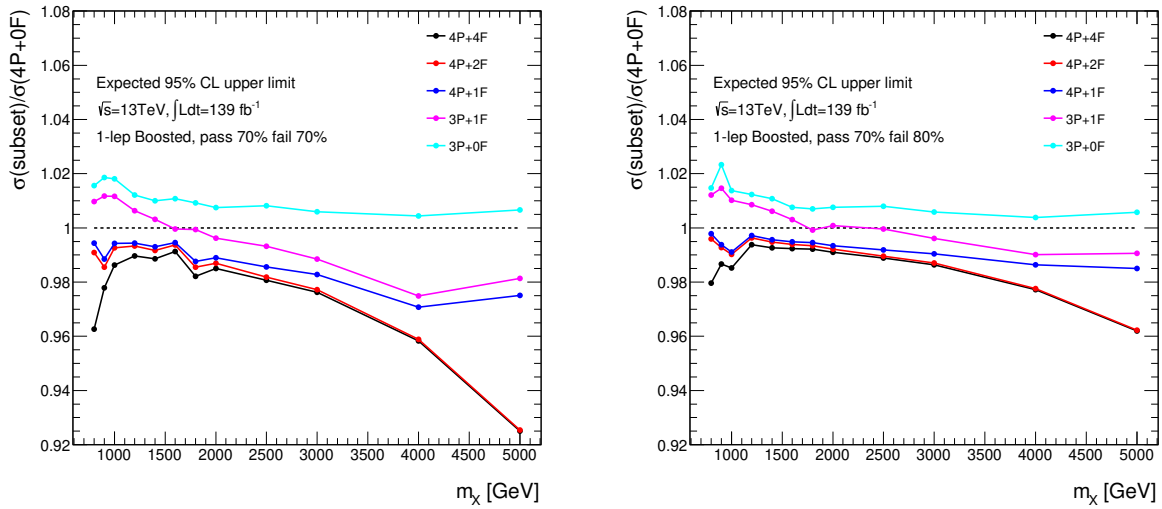
- *4P+0F* (dashed line): All four regions requiring events to pass the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window combined as SRs, and all four regions requiring events to fail the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window as VRs. This is the nominal case to which was used so far. This will also be used as the reference point to compare the performance of other possible combinations.

- *3P+1F* (magenta line): All regions requiring events to pass the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window except the exactly one b-tag electron region, and the muon channel region requiring events to fail the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window with at least two b-tags combined as SRs. The VRs include the exactly one b-tag electron channel region requiring events to pass the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window, both electron channel regions requiring events to fail the $m_{\mathrm{TAR}}^{H\to b\bar{b}}$ window and the exactly one b-tag muon channel region requiring events to fail the

$m_{\text{TAR}}^{H\to b\bar{b}}$ window.

- *3P+0F* (cyan line): All regions requiring events to pass the $m_{\text{TAR}}^{H\to b\bar{b}}$ window except the exactly one b-tag electron region combined as SRs. All regions requiring events to fail the $m_{\text{TAR}}^{H\to b\bar{b}}$ window and the exactly one b-tag electron channel region used as VRs.

This is also summarised in Table 4.3.



(a) pass 70% or fail 70% WP of $m_{\text{TAR}}^{H\to b\bar{b}}$ window ***(b)*** pass 70% or fail 80% WP of $m_{\text{TAR}}^{H\to b\bar{b}}$ window

Figure 4.8.: Comparison of expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ with different subset of regions assumed as SRs. Only statistical uncertainties are included in shown exclusion limits.

Table 4.3.: Summary of various possibilities of combinations considered.

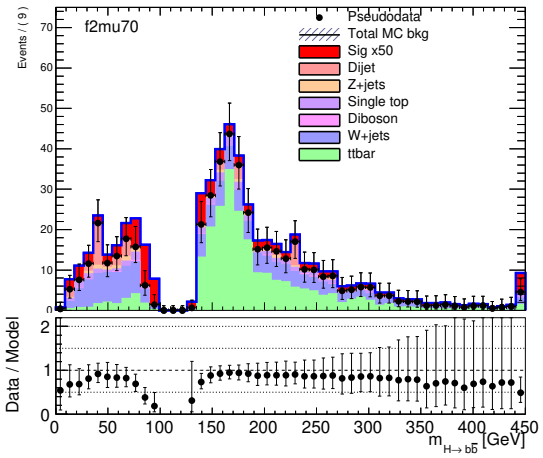
Combination	f1el	f1mu	f2el	f2mu	p1el	p1mu	p2el	p2mu
4P+4F	SR	SR	SR	SR	SR	SR	SR	SR
4P+2F	VR	SR	VR	SR	SR	SR	SR	SR
4P+1F	VR	VR	VR	SR	SR	SR	SR	SR
4P+0F	VR	VR	VR	VR	SR	SR	SR	SR
3P+1F	VR	VR	VR	SR	VR	SR	SR	SR
3P+0F	VR	VR	VR	VR	VR	SR	SR	SR

From Figure 4.8b it can be seen that the limits for 4P+2F case are very close to the 4P+4F case, which goes on to show that the electron channel regions requiring events to fail the $m_{\text{TAR}}^{H\to b\bar{b}}$ window have a very low sensitivity to the signal. Then, by removing

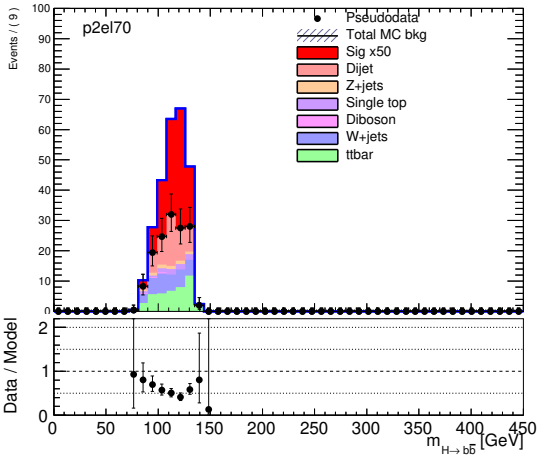
the exactly one b-tag muon channel region requiring events to fail the $m_{\text{TAR}}^{H\to b\bar{b}}$ window, the 4P+1F case shows a degradation in the performance of about 2% for $m_X > 3.0\,\text{TeV}$. Further, it can be seen that the 3P+1F case is another 1 to 2% weaker than the 4P+1F case. This connects to the sensitivity of the exactly one b-tag electron channel region requiring events to pass the $m_{\text{TAR}}^{H\to b\bar{b}}$ window. However, this case still performs better than the nominal case (4P+0F) for $m_X > 2.5\,\text{TeV}$, and is only about 1% worse than the nominal case for lower masses. On further removing the region with at least two b-tag muon channel requiring events to fail the $m_{\text{TAR}}^{H\to b\bar{b}}$ window, the 3P+0F case results in limits about 2% weaker than the 3P+1F case, particularly for higher m_X. From these observations, it was decided to use the 3P+1F case, since it provides a better overall performance than the nominal 4P+0F case as well as allows the usage of useful VRs. Additionally, since the muon channel region requiring events to fail the $m_{\text{TAR}}^{H\to b\bar{b}}$ window and having at least two b-tags is now converted to a SR, the 70% WP is used also for this particular region to make the best of the SRs.

Figures 4.9 and 4.10 also show the event distributions for the $m_{\text{TAR}}^{H\to b\bar{b}}$ variable in the four SRs and four VRs, respectively, for the 3P+1F case with dijet MC sample as a proxy for the QCD estimate. The data points in these plots correspond to pseudodata. In all of these plots, the signal represents the sum of all mass points and is then scaled up by a factor of 50. Apart from the $m_{\text{TAR}}^{H\to b\bar{b}}$ window in all eight regions, the plots also depict the amount of signal contribution in each of these SRs and VRs.

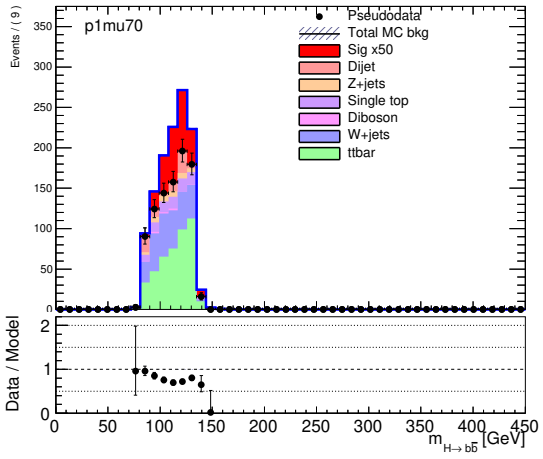
The initial search strategy incorporated both electron and muon channels. However, the electron channel was eventually left out due to several reasons. Firstly, the isolation scale factors on identification efficiencies for electrons in TAR jets have not been made available yet. Secondly, the data-driven estimate of the non-prompt lepton background in case of electrons was much more involved as compared to the muon channel [98]. Thirdly, in order to include the electron channel, additional systematic uncertainties would be required for the lepton-in-jet topology. Furthermore, the time required to overcome these challenges posed by the electron channel did not fit into the analysis timescales. Moreover, as shown in Figure 4.11, which shows the comparison of expected exclusion limits on $\sigma(pp \to X \to HH)$ with (dark pink dashed line) and without (black dashed line) the electron channel included, the sensitivity gained by the inclusion of the electron channel is negligible, especially in for the region where the analysis is particularly sensitive ($m_X > 2.0\,\text{TeV}$). Hence, the electron channel was removed from consideration in the analysis. Thus, it was decided to continue with the muon channel exclusively. Therefore, for the rest of this thesis, lepton refers to a muon, unless specified otherwise. A summary of the remaining SRs and VRs is also provided in Table 4.4. However, the definition of SRs



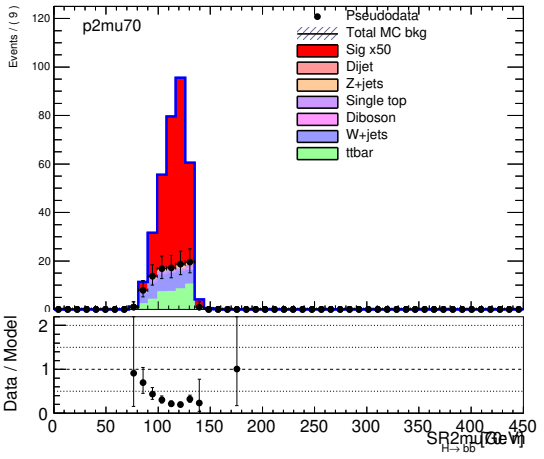
(a) f2mu



(b) p2el



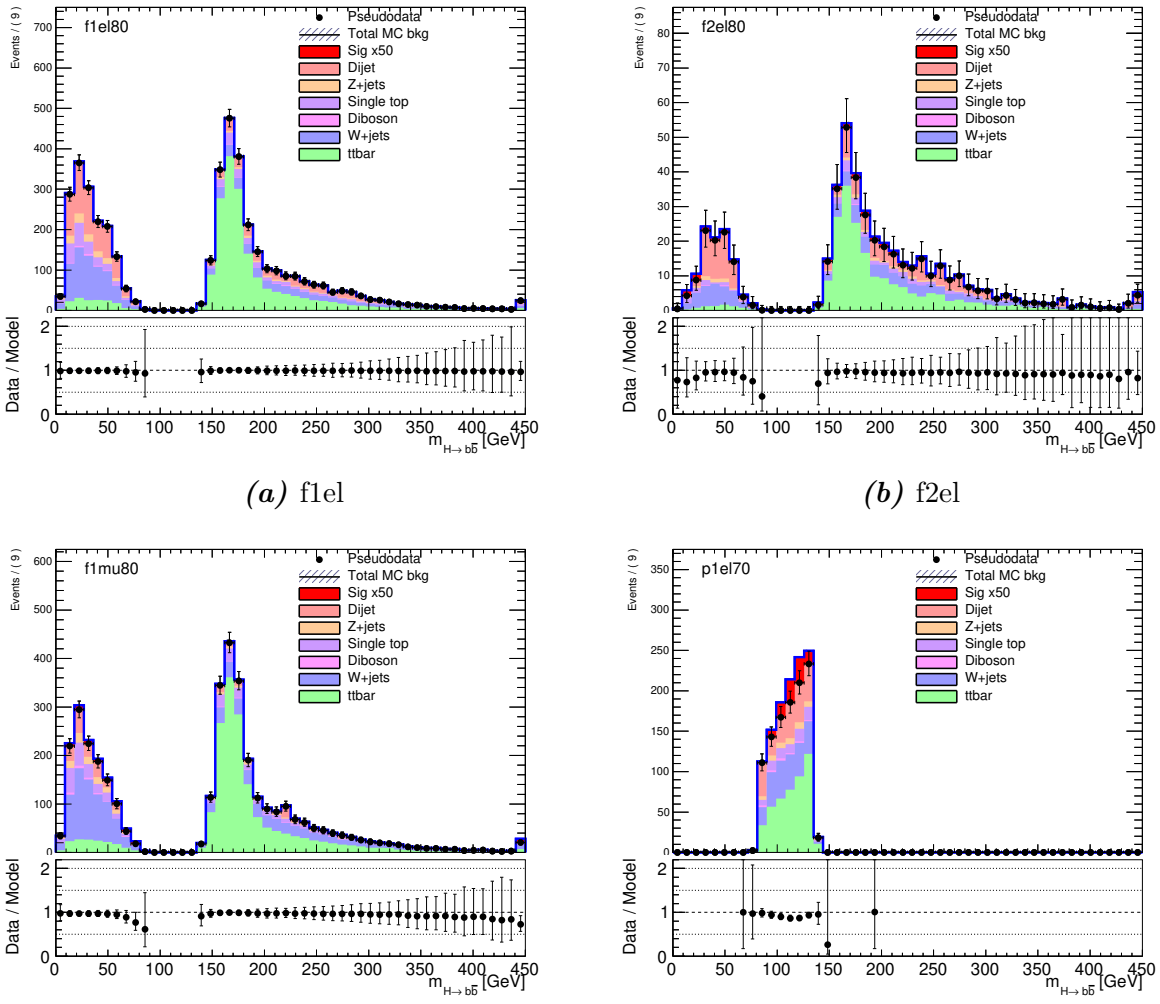
(c) p1mu



(d) p2mu

Figure 4.9.: Distributions of the $m_{\text{TAR}}^{H\to b\bar{b}}$ variable in the four SRs of the 3P+1F case with dijet MC sample as a proxy for the QCD background. The data points correspond to the pseudodata which is set to the background only estimate. The signal here is the sum of all HH signal samples and scaled up by a factor of 50. The deviations in the ratio plots show the signal contribution.

and VRs is not changed, even with the presence of just one VR, since all three SRs have a considerable amount of signal contribution and the also the background composition in all three SRs is fairly similar.



(a) f1el *(b)* f2el

(c) f1mu *(d)* p1el

Figure 4.10.: Distributions of the $m_{\text{TAR}}^{H\to b\bar{b}}$ variable in the four VRs of the 3P+1F case with dijet MC sample as a proxy for the QCD background. The data points correspond to the pseudodata which is set to the background only estimate. The signal here is the sum of all HH signal samples and scaled up by a factor of 50. The deviations in the ratio plots show the signal contribution.

Additional Modelling Checks

Additional attempts were made to improve the modelling of the non-prompt lepton background with the help of selection cuts rejecting regions that were either poorly modelled or highly populated by this background. For this, different kinematic variables were investigated [98]. Figure 4.12 shows distributions of two of these variables, the p_T of the $H \to b\bar{b}$ candidate and $\Delta R\,(\ell, W_{\text{had}})$, in the VR. From Figure 4.12a, it can be seen that for $p_T^{H\to b\bar{b}} > 500\,\text{GeV}$, data is fairly well modelled except for a shift in normalisation.

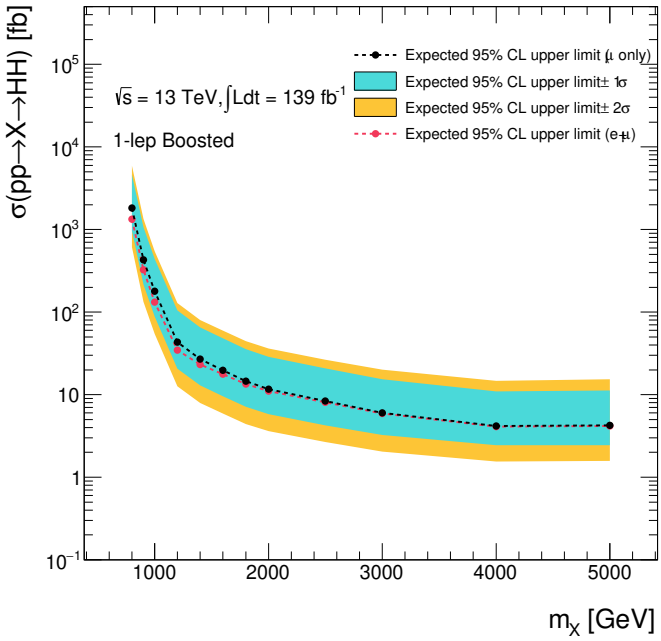


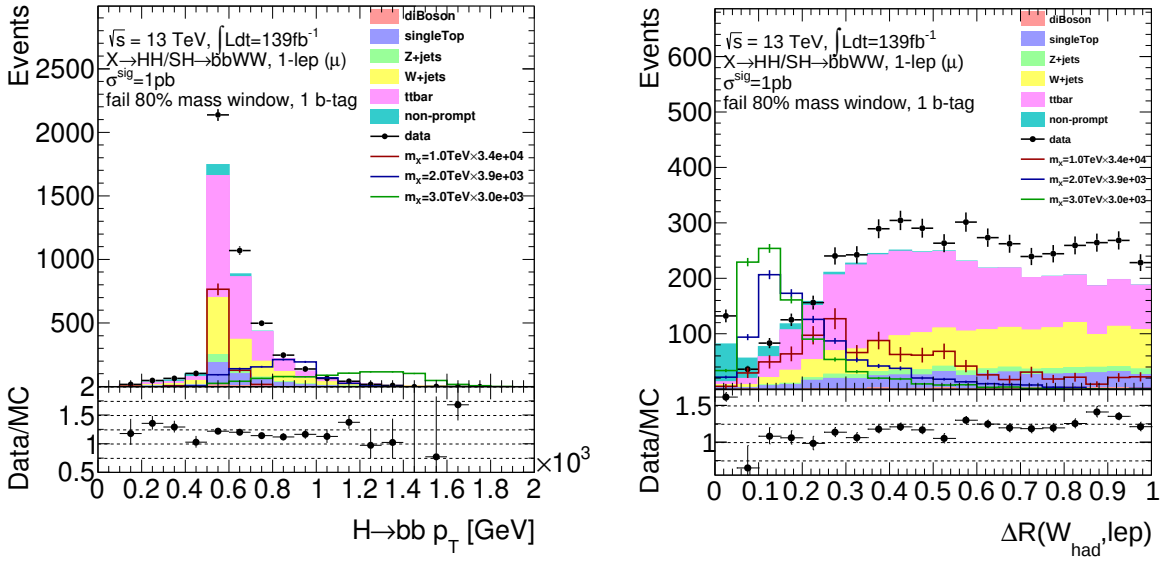
Figure 4.11.: Comparison of expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ with and without the electron channel included. Only statistical uncertainties are included in shown exclusion limits.

Table 4.4.: Summary of all SRs and VRs as per the 3P+1F combination with just the muon channel.

Region	$m_{\text{TAR}}^{H\to b\bar{b}}$ window	$N_{b\text{-tags}}$
VRf1	fail 80%	$= 1$
SRf2	fail 70%	≥ 2
SRp1	pass 70%	$= 1$
SRp2	pass 70%	≥ 2

However, this is not the case for $p_T^{H\to b\bar{b}} < 500\,\text{GeV}$. From Figure 4.12b, one can also see that the first bin corresponding to $\Delta R\,(\ell, W_{\text{had}}) < 0.05$ has a very high contribution from the non-prompt background with a very low signal contamination. Therefore, these two cuts are investigated.

The expected exclusion limits are again compared to test the performance of these cuts. Figure 4.13 shows the expected exclusion limits of just the $p_T^{H\to b\bar{b}} > 500\,\text{GeV}$ selection cut applied (red solid line) and the $\Delta R\,(\ell, W_{\text{had}}) > 0.05$ selection cut applied in addition to the $p_T^{H\to b\bar{b}}$ cut (blue solid line), both normalised to the limits obtained without these two cuts. From the figure, it can be seen that the addition of the requirement for $p_T^{H\to b\bar{b}} > 500\,\text{GeV}$ (red solid line) improves the limits across the entire mass spectrum, but particularly for



(a) $p_T^{H\to b\bar{b}}$ *(b)* $\Delta R\,(\ell, W_{\text{had}})$

Figure 4.12.: Distributions of $p_T^{H\to b\bar{b}}$ and $\Delta R\,(\ell, W_{\text{had}})$ in the validation region [98]. The distributions for signal samples are scaled up to 25% of the total background events.

lower m_X. With the requirement of $\Delta R\,(\ell, W_{\text{had}}) > 0.05$ in addition (blue solid line), the limits further improve for $m_X < 1.4\,\text{TeV}$. However, for $m_X > 1.6\,\text{TeV}$, the limits begin to worsen with the selection cut on $\Delta R\,(\ell, W_{\text{had}})$, since the cut also rejects highly boosted signal events with the lepton candidate very close to the W_{had} TAR jet, which correspond to high m_X. Hence only the $p_T^{H\to b\bar{b}}$ cut is retained for the final event selection, whereas the $\Delta R\,(\ell, W_{\text{had}})$ cut is dropped.

Finally, to optimise the SR selection, additional kinematic and substructure variables are investigated [98]. The distributions for the most promising of these variables, the $\Delta R\,(\ell, W_{\text{had}})$ and $C_2^{H\to b\bar{b}}$, in the three SRs are also shown in Figure 4.14. From the figure it can be seen that the signal samples generally peak at lower values of $\Delta R\,(\ell, W_{\text{had}})$ and $C_2^{H\to b\bar{b}}$, while the backgrounds are much more broadly spread. These are again implemented as p_T-dependent floating window cuts with 50, 60, 70 and 80% WPs (see Appendix B.1), as the $m_{\text{TAR}}^{H\to b\bar{b}}$ window cut discussed earlier in this subsection. However, for the $\Delta R\,(\ell, W_{\text{had}})$ window, the $p_T^{H_{\text{vis}}}$ (dashed lines in Figure 4.15a) was used, where $H_{\text{vis}} = W_{\text{had}} + \ell$. As before, the ratio of expected exclusion limits is obtained with and without additional cuts for the HH and SH samples, which is shown in Figure 4.15. From Figure 4.15a it is evident that the $\Delta R\,(\ell, W_{\text{had}})$ selection cuts (dashed lines) performs better than the $C_2^{H\to b\bar{b}}$ cuts (solid lines), while the 80% WP (magenta lines) is seen to perform best for both cases. Hence, the 80% WP $\Delta R\,(\ell, W_{\text{had}})$ window cut is added to

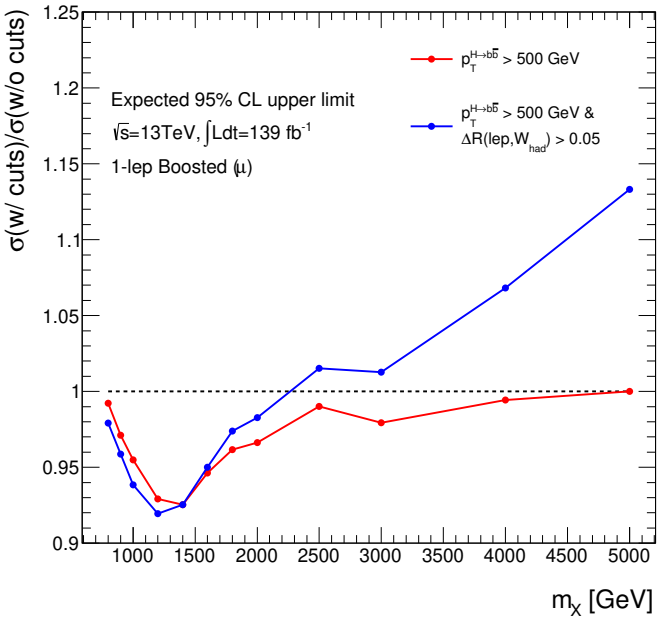


Figure 4.13.: Ratio of expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ with and without additional cuts. Only statistical uncertainties are included in shown exclusion limits. The red line represents the limits with the additional requirement of $p_T^{H\to b\bar{b}} > 500\,\text{GeV}$ and the blue line represents the limits with both the additional selection cuts, $p_T^{H\to b\bar{b}} > 500\,\text{GeV}$ and $\Delta R\,(\ell, W_{\text{had}}) > 0.05$, applied.

the HH signal region definition.

For the SH samples, due to variable m_S, a variety of topologies can be observed which will be discussed in detail in Section 5.1. However, due to these variations in topologies, the $\Delta R\,(\ell, W_{\text{had}})$ is not expected to perform well for the SH samples. Thus, for the SH samples, only the $C_2^{H\to b\bar{b}}$ variable was tested. The $C_2^{H\to b\bar{b}}$, however, can be defined by either adding all the HH signal samples or by adding all the SH samples. Hence, both of these are tried. The comparison of expected exclusion limits with and without these cuts is shown in Figures 4.15b, 4.15c and 4.15d for $m_X = 1.0$, 2.0 and 3.0 TeV, respectively. From these plots, it can be concluded that the 80% WP $C_2^{H\to b\bar{b}}$ window determined on the SH samples (magenta dashed line) performs overall the best of all the possibilities and provides an improvement of approximately 10% for most mass points. This same cut is then also tested on the HH samples in addition to the $\Delta R\,(\ell, W_{\text{had}})$ cut. This is represented by the black solid line in Figure 4.15a, which further improves the limit. The total improvement for the HH samples with both the selection cuts applied together goes up to about 30% for $m_X \geq 1.4\,\text{TeV}$.

Final Event Selection

The finalised preselection, as it stands, includes the requirement for the single large-R jet trigger to have fired and the presence of exactly one signal muon in the event. It also requires at least two TAR jets in the event. Furthermore, the events must satisfy the selection cuts $\min[\Delta R(\ell, J)] \leq 1.0$ and $p_T^{H\to b\bar{b}} > 500\,\text{GeV}$.

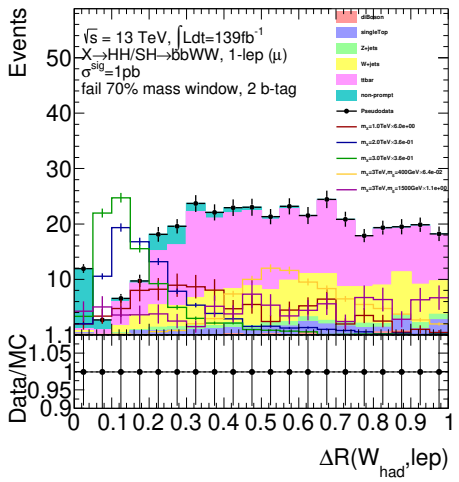
The SRs require the event to also have exactly one b-tagged TAR jet in the event. The events must pass the 80% WP $C_2^{H\to b\bar{b}}$ window. Additionally, for the HH samples, the events must also pass the 80% WP $\Delta R(\ell, W_{\text{had}})$ window. These are then split into three SRs as follows:

- *SRp1*: Pass 70% WP $m_{\text{TAR}}^{H\to b\bar{b}}$ window with the $H \to b\bar{b}$ candidate having exactly one b-tag.
- *SRp2*: Pass 70% WP $m_{\text{TAR}}^{H\to b\bar{b}}$ window with the $H \to b\bar{b}$ candidate having at least two b-tags.
- *SRf2*: Fail 70% WP $m_{\text{TAR}}^{H\to b\bar{b}}$ window with the $H \to b\bar{b}$ candidate having at least two b-tags.

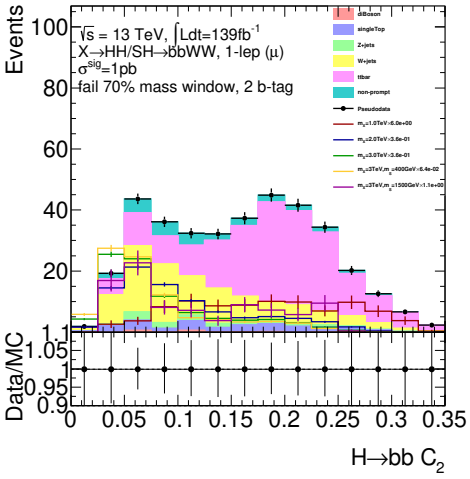
The three CRs are then defined as follows:

- The $t\bar{t}$ CR requires exactly two b-tagged TAR jets in the event and further requires $m_{\text{TAR}}^{W_{\text{had}}} < 20\,\text{GeV}$.
- The W+jets CR is defined by the requirement of exactly zero b-tagged TAR jets in the event, and the events must fail the 70% WP $m_{\text{TAR}}^{H\to b\bar{b}}$ window. Finally, it requires the events to also satisfy $60\,\text{GeV} < m_T^{W_{\text{lep}}} < 120\,\text{GeV}$.
- Finally, the QCD CR is defined exactly like the W+jets CR except, in this case, the cut on $m_T^{W_{\text{lep}}}$ is vetoed.

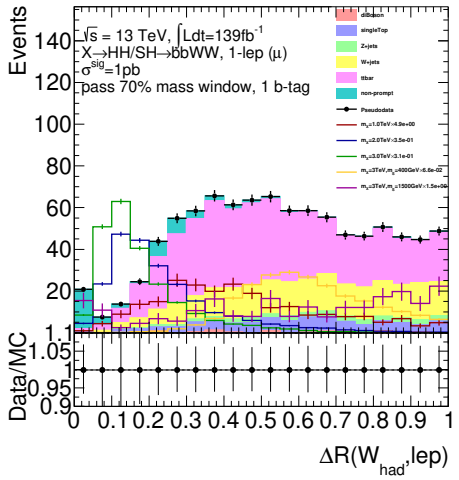
The only VR is defined by the requirement of exactly one b-tagged TAR jet in the event, where the $H \to b\bar{b}$ candidate must also have exactly one b-tag. Finally, the events must fail the 80% WP $m_{\text{TAR}}^{H\to b\bar{b}}$ window.



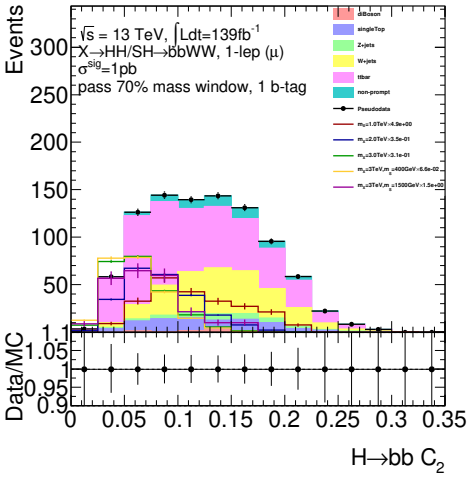
(a) SRf2, $\Delta R\,(\ell, W_{\mathrm{had}})$



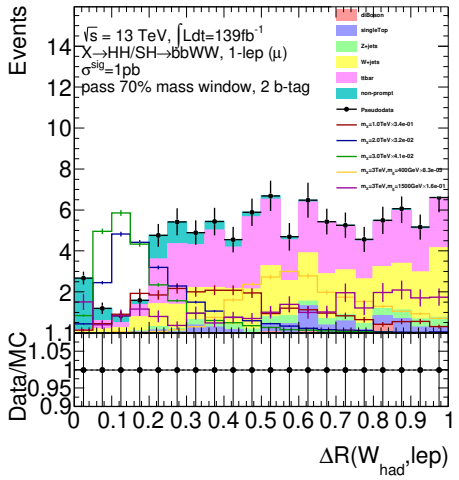
(b) SRf2, $C_2^{H\to b\bar{b}}$



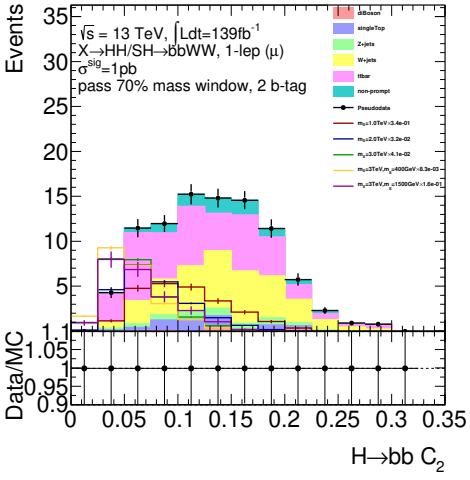
(c) SRp1, $\Delta R\,(\ell, W_{\mathrm{had}})$



(d) SRf2, $C_2^{H\to b\bar{b}}$

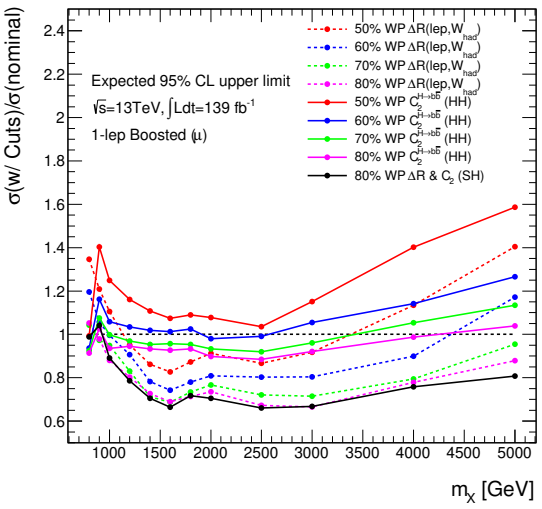


(e) SRp2, $\Delta R\,(\ell, W_{\mathrm{had}})$

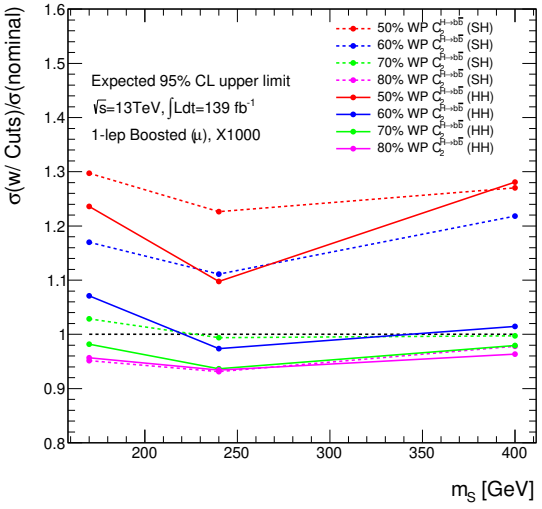


(f) SRf2, $C_2^{H\to b\bar{b}}$

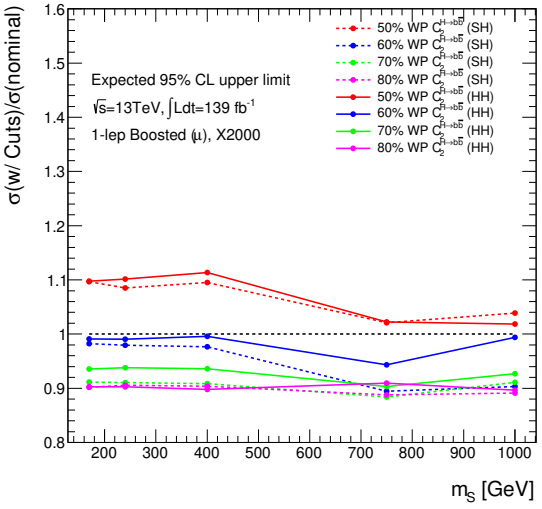
Figure 4.14.: Distributions of $\Delta R\,(\ell, W_{\mathrm{had}})$ and $C_2^{H\to b\bar{b}}$ in the three SRs [98]. The distributions for signal samples are scaled up to 25% of the total background events.



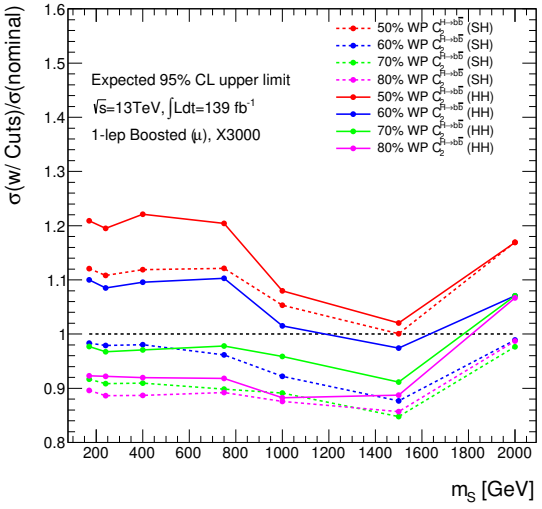
(a) HH samples



(b) SH samples with $m_X = 1.0\,\text{TeV}$



(c) SH samples with $m_X = 2.0\,\text{TeV}$



(d) SH samples with $m_X = 3.0\,\text{TeV}$

Figure 4.15.: Ratio of expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ with and without additional SR optimisation cuts. Only statistical uncertainties are included in shown exclusion limits.

5. Event Selection Optimisation in $X \to SH \to b\bar{b}WW$

"There's a way to do it better - find it."
- Thomas A. Edison

This chapter elaborates on the studies conducted towards optimising the jet classification for the SH signal samples. For this analysis, the variables $0.75\,\text{TeV} \leq m_X \leq 3.0\,\text{TeV}$, and $2m_W < m_S < (m_X - m_H)$ define the phase space of the $X \to SH \to b\bar{b}WW$ search. Throughout this chapter lepton, ℓ, refers to either an electron or a muon.

5.1. Challenges with SH Topologies

Due to the interplay of m_X and m_S, a variety of topologies can be observed, depending on their values and their difference ($\Delta M_{X,S}$). These are broadly classified into the following five types:

1. *Resolved* (Figure 5.1a): This topology is observed for low m_X and corresponding low m_S, where all the decay products are well separated in the detector.

2. *Split-resolved* (Figure 5.1b): This case is observed when $\Delta M_{X,S}$ is very small and m_X is sufficiently large. It is characterised by the overlap of the two quark jets from the decay of the W_{had}, which can then be reconstructed as a single large-R jet. The two b-jets and the lepton are, however, still well separated in the detector.

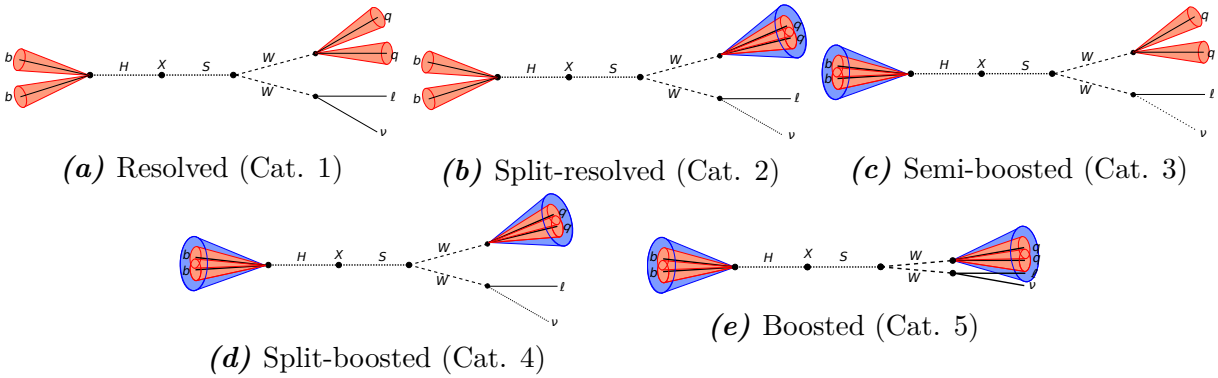
3. *Semi-boosted* (Figure 5.1c): This topology acts as the transition region between high and low m_X, where the two b-jets from the $H \to b\bar{b}$ decay can be reconstructed as a single large-R jet. The decay products of the W_{had} and the lepton are, however, resolved in this case.

4. *Split-boosted* (Figure 5.1d): The split-boosted topology is observed for high m_X and high m_S, while the difference is not too small ($0.3m_X \lessapprox m_S \lessapprox 0.8m_X$). The

two quark jets from the W_{had} are reconstructed as one single large-R jet, and the two b-jets from the $H \to b\bar{b}$ decay are reconstructed as another single large-R jet. Additionally, the lepton from the decay of the W_{lep} is well isolated from the decay products of the W_{had}.

5. *Boosted* (Figure 5.1e): Finally, the boosted topology is observed for high m_X and large $\Delta M_{X,S}$ ($m_S \lessapprox 0.3 m_X$). Its characteristics are similar to split-boosted case, except that the lepton also overlaps with the large-R jet from the decay of the W_{had}.

Figure 5.2 depicts these five cases in the m_X vs. m_S plane. From this plot, it is also evident that since only $0.75\,\text{TeV} \leq m_X \leq 3.0\,\text{TeV}$ is considered in this thesis, the most relevant topologies for this analysis are the split-boosted and the boosted scenarios. Further, it is also worth noting that the transition between these two topologies occurs when $m_S \approx 0.3\, m_X$.



(a) Resolved (Cat. 1) ***(b)*** Split-resolved (Cat. 2) ***(c)*** Semi-boosted (Cat. 3)

(d) Split-boosted (Cat. 4) ***(e)*** Boosted (Cat. 5)

Figure 5.1.: Graphical representation of the various possible SH topologies as they appear in the detector.

These changing topologies, however, pose an interesting challenge on the event selection, which can be seen in Figure 5.3. The figure shows the successive efficiencies after applying the various selection cuts as described in Section 4.3. It can be seen that with increasing m_S for a constant m_X, the lepton identification improves. This can be easily explained by the fact that high m_S corresponds to split-boosted topologies, where the lepton is well isolated. But it is even more striking that for split-boosted cases, the efficiency of the requirement of the $H \to b\bar{b}$ candidate to have at least 1 b-tag drops significantly. This is likely to be a consequence of the jet classification applied in the object definitions, where the TAR jet closest to the ℓ candidate is classified as the W_{had} candidate. In a split-boosted scenario, however, the $H \to b\bar{b}$ candidate could end up geometrically closer to the ℓ candidate than the W_{had} candidate. Thus, the jet classification used so far would not be able to classify the two TAR jets correctly in such cases.

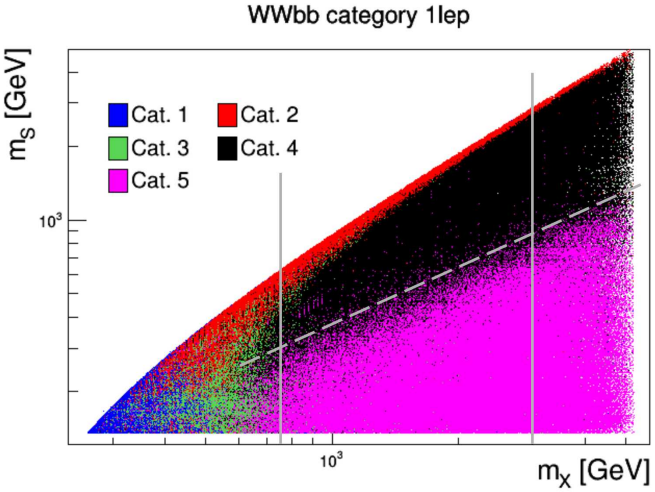


Figure 5.2.: Distribution of the various possible SH topologies in the m_X vs. m_S plane [98]. Here "Cat. 1", "Cat. 2", "Cat. 3", "Cat. 4" and "Cat. 5" refer to resolved, split-resolved, semi-boosted, split-boosted and boosted topologies, respectively. The two solid vertical lines correspond to $m_X = 0.75\,\text{TeV}$ and $3.0\,\text{TeV}$, while the dashed diagonal line represents the transition region between boosted and split-boosted topologies at $m_S \approx 0.3 m_X$.

This can also be seen in Figure 5.4, which shows the distributions of m_{TAR} for the jet candidate closest to the ℓ candidate normalised to unity for all signal samples. From the figure it is evident that for samples with high m_S for each m_X, another peak near the mass of the Higgs boson starts appearing in the mass distributions. Since the three p_T-leading jet candidates in the event are used for the classification, the peak at low masses ($m_{\text{TAR}} < 30\,\text{GeV}$) most likely corresponds to the scenarios where either the W_{had} is not entirely contained in a large-R jet (especially for low m_X and low m_S) or the third jet originating from QCD pile up is closest to the ℓ candidate.

This, however, shows that the classification method used for the boosted HH signal is not suitable for split-boosted SH topologies and demonstrates the need to employ better classification methods.

5.2. Classifier Definitions and Efficiencies

To tackle the challenge discussed in the previous section (Section 5.1), a few classifier definitions were tried in an attempt to correctly label the W_{had} and the $H \to b\bar{b}$ candidates across all signal samples. It is worth mentioning that for all of these classifiers only the three p_T-leading jets are considered. The classifier methods studied are listed below.

- p_T method: This is the default classifier used for the boosted HH analysis. The

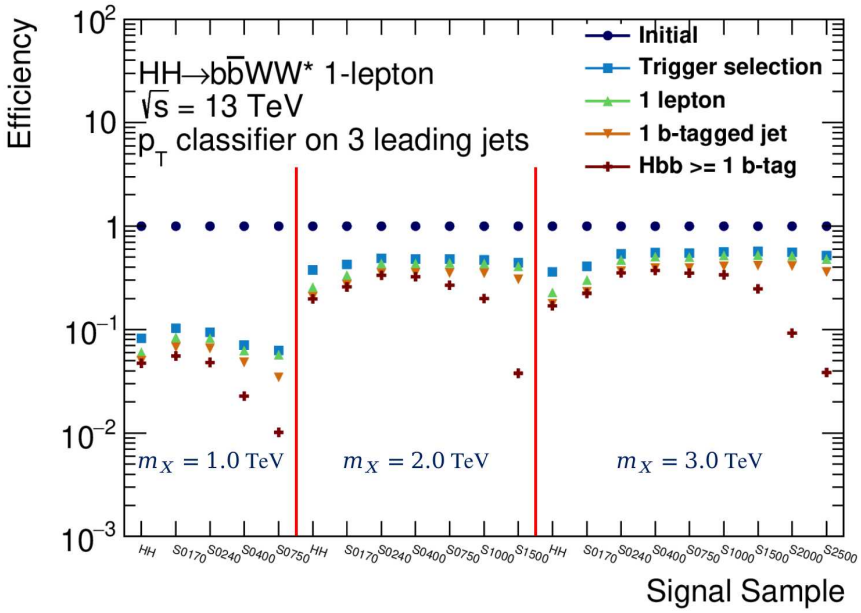


Figure 5.3.: Successive efficiencies after various event selection cuts for different signal mass points [98]. Here the "Trigger selection" refers to the requirement of the event to pass a large-R jet trigger, "1 lepton" refers to the requirement of a signal lepton (electron or muon), "1 b-tagged jet" is the requirement of at least 1 TAR jet in the event to be b-tagged and "$Hbb \geq 1$ b-tag" is the requirement for the TAR jet classified as $H \to b\bar{b}$ to have at least 1 b-tag.

W_{had} candidate is classified as the jet closest to the ℓ candidate and the $H \to b\bar{b}$ candidate is classified as the p_T-leading jet of the remaining jets. This is motivated by the close proximity of the lepton to the W_{had} in boosted topologies.

- m_H method: The jet with m_{TAR} closest to m_H is classified as the $H \to b\bar{b}$ candidate and the p_T-leading jet of the remaining jets is classified as the W_{had} candidate.

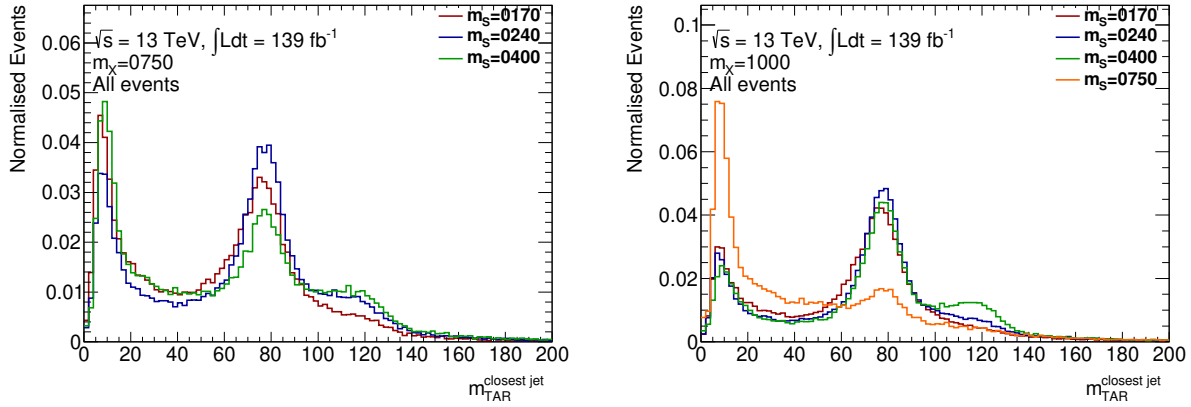
- m_W method: The jet with m_{TAR} closest to m_W is classified as the W_{had} candidate and the p_T-leading jet of the remaining jets is classified as the $H \to b\bar{b}$ candidate.

- $\sum p_T$ method: The $H \to b\bar{b}$ and the W_{had} candidates are classified as the combination of jets that minimise

$$\left| E_T^{\text{miss}} + p_T^{\ell} + p_T^{W_{\text{had}}} - p_T^{H \to b\bar{b}} \right|, \tag{5.1}$$

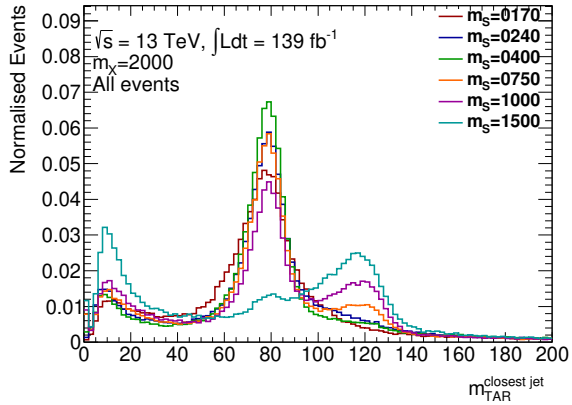
  which is essentially the expression for the conservation of transverse momentum.

- λ^2 method: The $H \to b\bar{b}$ and the W_{had} candidates are classified as the combination

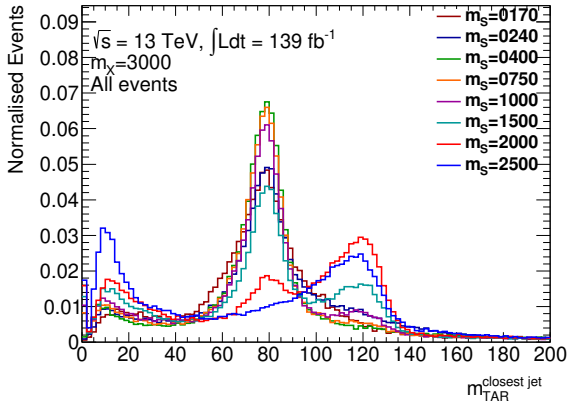


(a) $m_X = 0.75\,\text{TeV}$

(b) $m_X = 1.0\,\text{TeV}$



(c) $m_X = 2.0\,\text{TeV}$



(d) $m_X = 3.0\,\text{TeV}$

Figure 5.4.: Distributions of m_{TAR} of the jet closest to the ℓ candidate normalised to unity for all signal samples. All m_S here are given in GeV.

of jets that minimise λ^2 defined as

$$\lambda^2 = \left(\frac{m_{\text{TAR}}^{H\to b\bar{b}} - m_H}{\sigma_H}\right)^2 + \left(\frac{m_{\text{TAR}}^{W_{\text{had}}} - m_W}{\sigma_W}\right)^2, \tag{5.2}$$

which tries to combine the m_H and m_W methods, such that minimising λ^2 means better matches for the $H \to b\bar{b}$ and W_{had} candidates. The two terms are normalised by $\sigma_{H/W}$, which is the root mean square (RMS) of the ratio of the m_{TAR} and m_{parton} distributions of the respective truth-matched jet. It should be noted that since $\sigma_{H/W}$ are the RMS of the ratio of m_{TAR} and m_{parton} distributions, they are dimensionless quantities. However, this leads to λ having dimensions of mass. Tables 5.1 and 5.2 show the values of σ_H and σ_W, respectively, for each signal sample. From the tables, it can also be seen that the resolution of the ratio of the m_{TAR} and m_{parton} distributions of the respective truth-matched $H \to b\bar{b}$ candidate is generally smaller

for more boosted samples, while for the W_{had} candidate the resolution of this ratio is generally smaller for more split-boosted samples.

It should also be noted that these classification methods do not apply any selection on the events. To test and compare the performance of these classifiers, the truth-matching efficiencies of all the classifiers for individual signal samples are obtained. These are shown in Figure 5.5.

Table 5.1.: Values of σ_H used in the λ^2 method for all signal samples. Since σ_H is defined as the RMS of the ratio of the m_{TAR} and m_{parton} distributions of the truth-matched $H \to b\bar{b}$ jet, it has no units.

m_S \ m_X	750 GeV	1000 GeV	2000 GeV	3000 GeV
170 GeV	0.1617 ± 0.0020	0.1383 ± 0.0009	0.1458 ± 0.0004	0.1922 ± 0.0008
240 GeV	0.1664 ± 0.0020	0.1465 ± 0.0010	0.1438 ± 0.0003	0.1929 ± 0.0006
400 GeV	0.1903 ± 0.0024	0.1544 ± 0.0012	0.1421 ± 0.0003	0.1953 ± 0.0006
750 GeV	-	0.2112 ± 0.0018	0.1384 ± 0.0003	0.1869 ± 0.0006
1000 GeV	-	-	0.1369 ± 0.0003	0.1748 ± 0.0005
1500 GeV	-	-	0.1848 ± 0.0005	0.1521 ± 0.0005
2000 GeV	-	-	-	0.1416 ± 0.0004
2500 GeV	-	-	-	0.1794 ± 0.0006

Table 5.2.: Values of σ_W used in the λ^2 method for all signal samples. Since σ_W is defined as the RMS of the ratio of the m_{TAR} and m_{parton} distributions of the truth-matched W_{had} jet, it has no units.

m_S \ m_X	750 GeV	1000 GeV	2000 GeV	3000 GeV
170 GeV	0.2220 ± 0.0032	0.2228 ± 0.0016	0.2164 ± 0.0006	0.2405 ± 0.0010
240 GeV	0.1924 ± 0.0025	0.1900 ± 0.0014	0.2150 ± 0.0005	0.2635 ± 0.0009
400 GeV	0.1420 ± 0.0018	0.1668 ± 0.0014	0.1654 ± 0.0004	0.1938 ± 0.0006
750 GeV	-	0.1115 ± 0.0008	0.1538 ± 0.0004	0.1810 ± 0.0006
1000 GeV	-	-	0.1516 ± 0.0004	0.1779 ± 0.0005
1500 GeV	-	-	0.1513 ± 0.0004	0.1779 ± 0.0005
2000 GeV	-	-	-	0.1898 ± 0.0006
2500 GeV	-	-	-	0.2039 ± 0.0006

From the plot, it is clear that the p_T method is the most efficient for $m_S \leq 0.3\, m_X$, which was observed to be the transition point between boosted and split-boosted topologies in Section 5.1. However, as m_S increases, this efficiency for the p_T method decreases steadily.

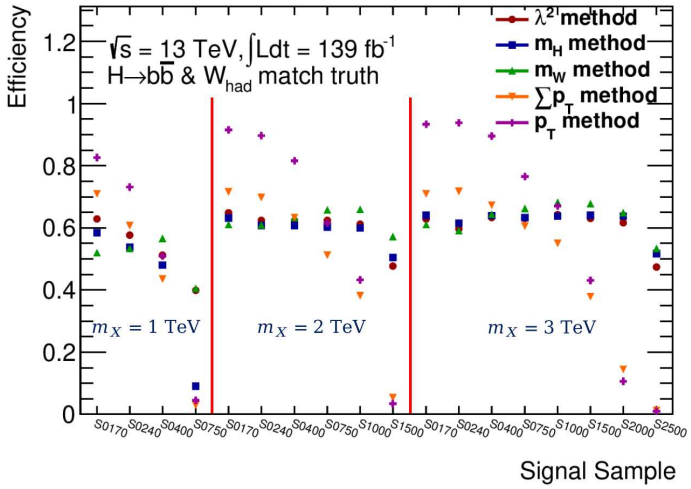


Figure 5.5.: Truth-matching efficiencies for different classifier definitions.

In addition to this, the m_W method shows the most consistent performance of all the classifiers, but its efficiency is too sub-optimal to be used in the analysis.

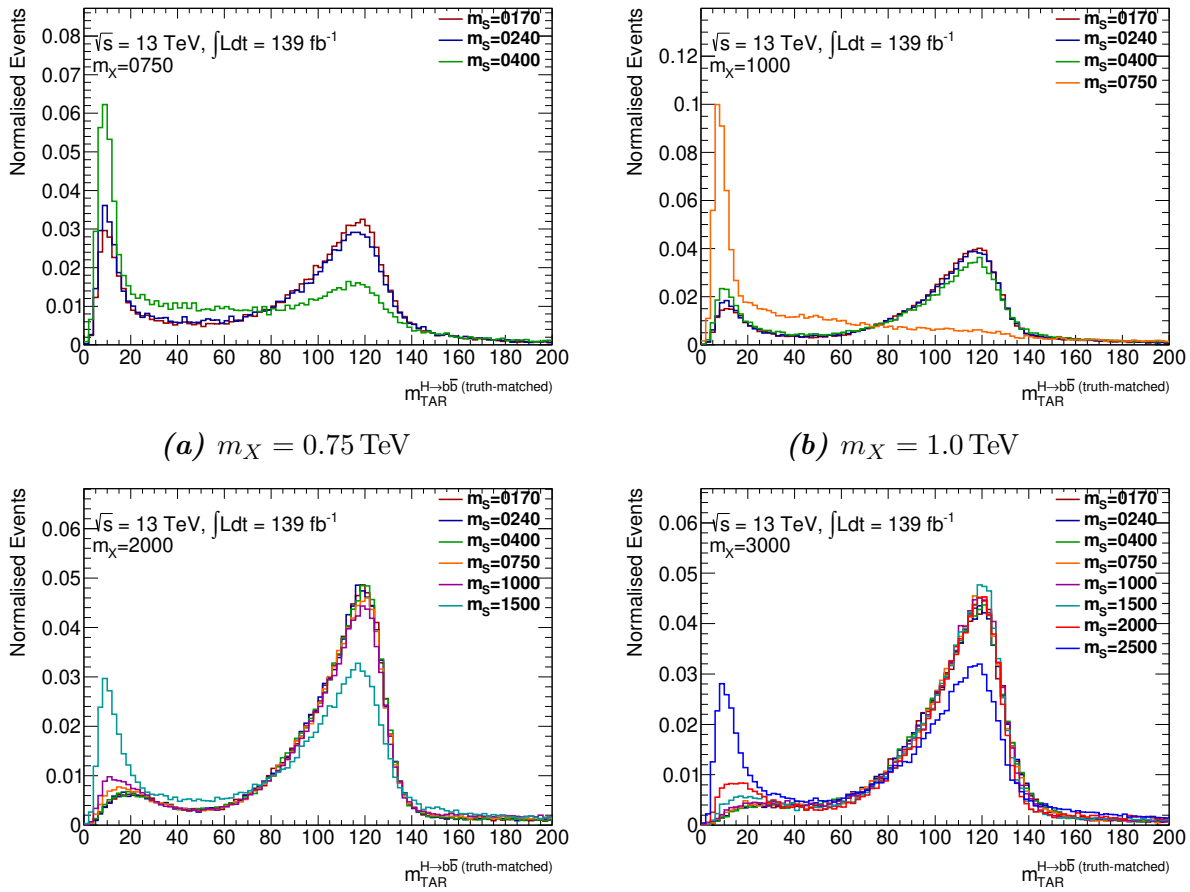
Since most of the classifier definitions are dependent on the mass of each individual jet, the distributions of m_{TAR} normalised to unity of the truth-matched $H \to b\bar{b}$ and W_{had} candidates are also shown for all the signal samples in Figures 5.6 and 5.7, respectively. Looking at Figure 5.6, one can see a low-mass peak in the mass spectrum of the truth-matched $H \to b\bar{b}$ candidate, which is more prominent for cases with high m_S. In these cases, $\Delta M_{X,S}$ becomes small enough to fall into the split-resolved topology, where the full $H \to b\bar{b}$ decay cannot be reconstructed as a single large-R jet.

From Figure 5.7, again a low-mass peak can be seen in the mass spectrum of the truth-matched W_{had} candidate in a few cases, with low m_X and low m_S. Referring back to Figures 5.1 and 5.2, it is visible that these tend towards the semi-boosted topology, where the two quark jets from the W_{had} are usually not reconstructed in a single large-R jet.

It is, however, evident from the plot in Figure 5.5 that no single classification method is efficient across all signal samples and hence, enforces the need to distinguish the boosted and split-boosted events, to be able to optimise both topologies independently.

5.3. Boosted and Split-Boosted Regions

A promising candidate variable to distinguish the boosted and split-boosted events is the ΔR between the ℓ candidate and the closest of the three p_T-leading jets, $\min[\Delta R(\ell, J)]$. Figure 5.8 shows the distributions for the $\min[\Delta R(\ell, J)]$ normalised to unity for all the signal samples. Looking at the plots, this variable shows a promising distinguishing



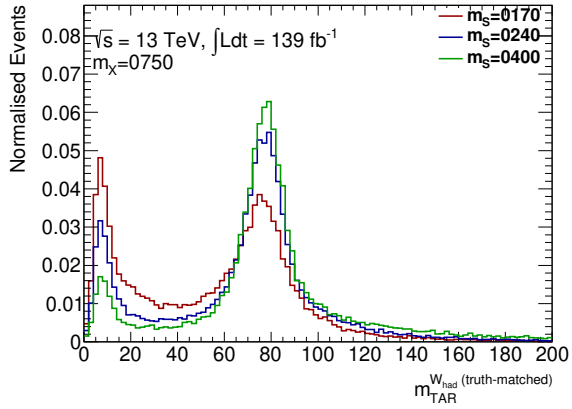
(a) $m_X = 0.75\,\text{TeV}$ ***(b)*** $m_X = 1.0\,\text{TeV}$

(c) $m_X = 2.0\,\text{TeV}$ ***(d)*** $m_X = 3.0\,\text{TeV}$

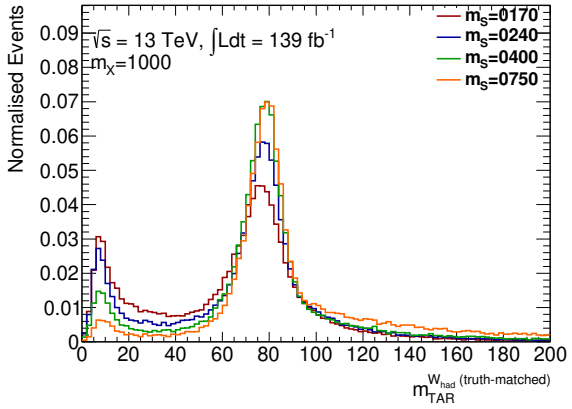
Figure 5.6.: Distributions of m_{TAR} for the truth-matched $H \to b\bar{b}$ jet candidate normalised to unity for all signal samples. All m_S here are given in GeV.

power between the different signal samples and, hence, between boosted and split-boosted topologies. Looking also at Figure 5.9, which shows the truth-matching efficiencies of the different classifiers dependent on the $\min[\Delta R(\ell, J)]$, it is clear that the p_T method has the highest efficiency for $\min[\Delta R(\ell, J)] < 1.0$. Therefore, this value was chosen as the point to separate between boosted and split-boosted topologies. From Figure 5.8 it can now also be seen which signal samples fall into the boosted region ($\min[\Delta R(\ell, J)] \leq 1.0$) and can therefore be classified with a reasonable efficiency by the p_T method. Since both topologies are sensitive to different regions of phase-space, it was considered best to optimise each of them independently to maximise the optimised area of the m_X vs. m_S plane (Figure 5.2).

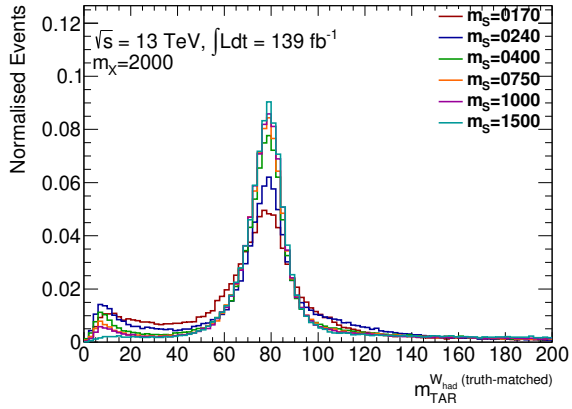
Furthermore, from Figure 5.10, which shows signal significance (ratio of signal yields to the square root of the estimated background yields) of the boosted HH signal samples with and without a cut on the $\min[\Delta R(\ell, J)]$, a clear improvement can be observed



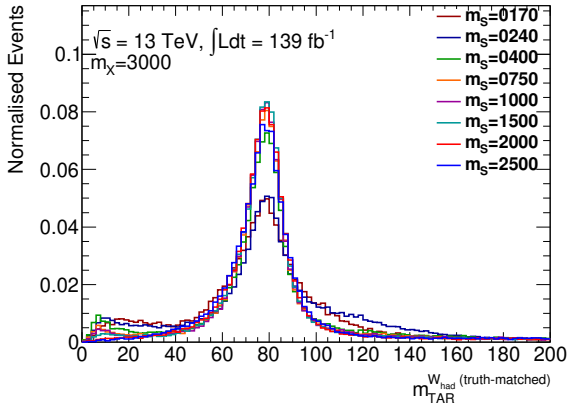
(a) $m_X = 0.75\,\text{TeV}$



(b) $m_X = 1.0\,\text{TeV}$



(c) $m_X = 2.0\,\text{TeV}$



(d) $m_X = 3.0\,\text{TeV}$

Figure 5.7.: Distributions of m_{TAR} for the truth-matched W_{had} jet candidate normalised to unity for all signal samples. All m_S here are given in GeV.

with the $\min[\Delta R(\ell, J)]$ cuts. Hence, the HH analysis also benefits from this cut at $\min[\Delta R(\ell, J)] < 1.0$. Therefore, the cut was implemented for in the boosted 1-lepton event selection. However, optimising the classification for the split-boosted scenario is beyond the scope of this thesis.

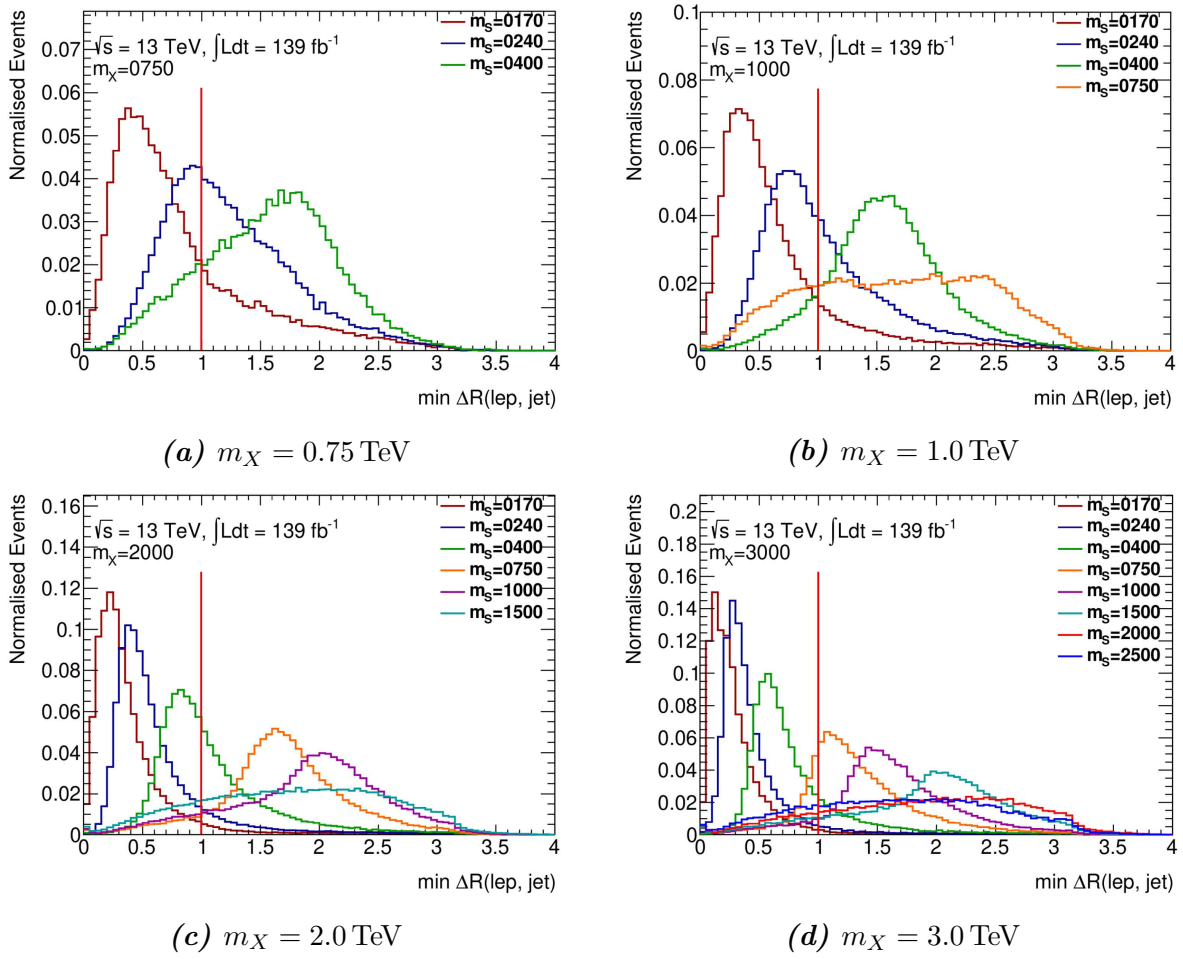


Figure 5.8.: Distributions for min $[\Delta R\,(\ell, J)]$ normalised to unity for all signal samples. All m_S here are given in GeV.

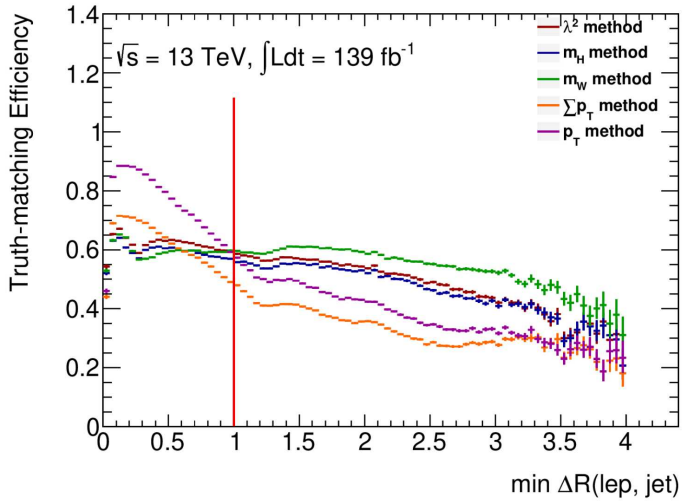


Figure 5.9.: Truth-matching efficiencies for different classifier methods dependent on the min $[\Delta R\,(\ell, J)]$.

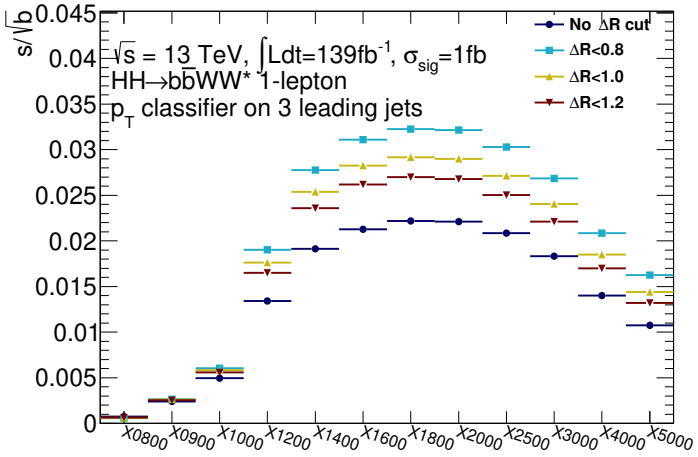


Figure 5.10.: Signal significance of the boosted HH signal samples with promising cut values for $\min\left[\Delta R\left(\ell, J\right)\right]$ [98].

6. Statistical Framework and Interpretation

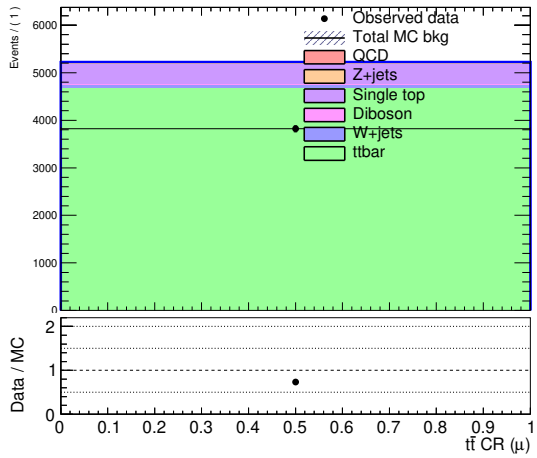
"Once you eliminate the impossible, whatever remains,
no matter how improbable, must be the truth."
- Sir Arthur Conan Doyle *(Sherlock Holmes)*

This chapter aims to provide details about the statistical framework and the fitting procedure, based on maximum likelihood fits, which are used for normalising the backgrounds and obtaining the signal strengths in this analysis. The chapter ends with a discussion about the expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ and $\sigma(pp \to X \to SH) \times \mathcal{BR}(SH \to b\bar{b}WW)$ for the HH and SH samples, respectively, in the blinded SRs with only statistical uncertainties included.

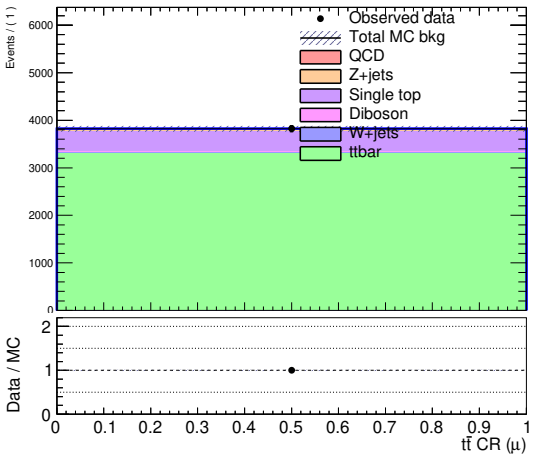
For the statistical analysis, the HistFitter [99] framework was used, which makes use of the HistFactory [100] and RooStat [101] packages. These are further based on the RooFit [102] and Root [103, 104] C++ libraries. For all the fits, the minimisation of the goodness-of-fit quantity was performed with Minuit [105, 106].

6.1. Background Normalisation Fits

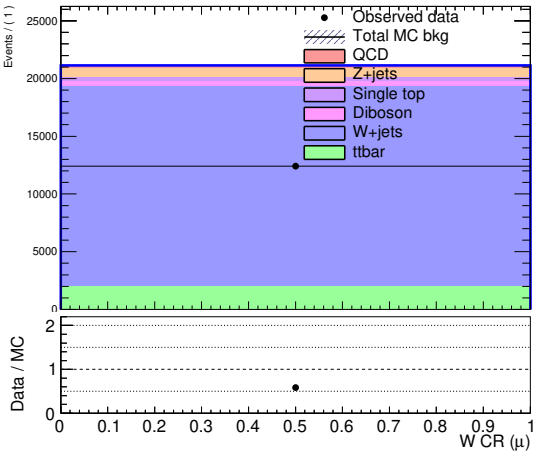
The motivation behind performing background normalisation fits can be realised by looking at the pre-fit composition of the $t\bar{t}$, W+jets and QCD CRs, which is shown in Figures 6.1a, 6.1c and 6.1e. From the plots, it is evident that the MC background overestimates the data in all three CRs. Thus, if a data-driven approach is used to estimate the non-prompt background by subtracting the MC backgrounds from the observed data, it will result in negative non-prompt events, which is unphysical. Therefore, the main backgrounds, i.e. $t\bar{t}$ ($\mu_{t\bar{t}}$), W+jets ($\mu_{W+\text{jets}}$) and multijet (μ_{QCD}) backgrounds need to be constrained by their respective normalisation factors (NFs). This is done by performing simultaneous single-bin fits in all three CRs.



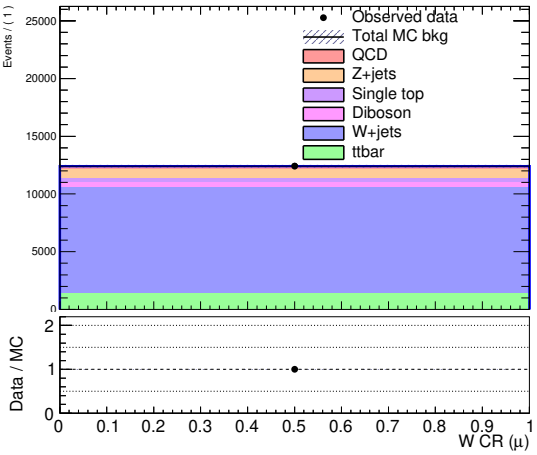
(a) Pre-fit $t\bar{t}$ CR



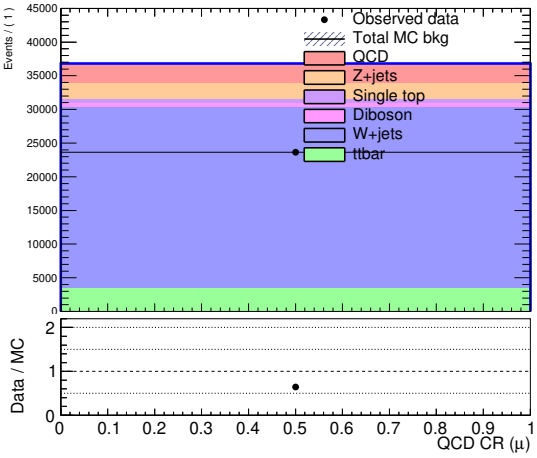
(b) Post-fit $t\bar{t}$ CR



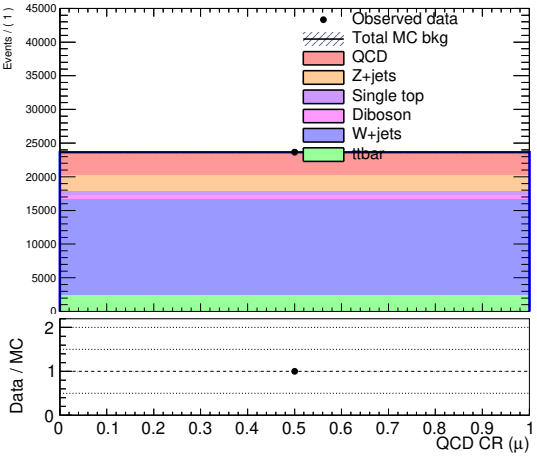
(c) Pre-fit W+jets CR



(d) Post-fit W+jets CR



(e) Pre-fit QCD CR



(f) Post-fit QCD CR

Figure 6.1.: Background composition in the $t\bar{t}$, W+jets and QCD CRs before and after performing the background normalisation fits. These plots correspond to the iteration with the data-driven non-prompt lepton backgrounds.

To obtain the correct NFs, an accurate estimate of the non-prompt background is also required. Therefore, an iterative procedure is employed in order to tackle this issue, where the first iteration is performed using a free-floating dijet MC sample as proxy for the non-prompt background in the fit. The prompt lepton backgrounds are then scaled by their respective NFs received from this first iteration, and these normalised backgrounds are used to estimate the multijet background. Finally, the fit is repeated with the non-prompt background estimate to obtain the final NFs. Figures 6.1b, 6.1d and 6.1f show the post-fit background composition in the $t\bar{t}$, W+jets and QCD CRs, respectively, which now also agrees with the observed data.

Additionally, the yields of all the background processes before and after the background normalisation fit are shown in Table 6.1 for all three CRs. Since systematic uncertainties, or uncertainties on MC statistics were not included in the background normalisation fits to avoid double-counting them in the analysis fits, only the samples that are allowed to vary within a defined range have an uncertainty, while others have zero uncertainties. The table also confirms the agreement of the post-fit yields with the observed data.

Table 6.1.: Yields for all background processes before and after the background normalisation fit in all three CRs.

Region	$t\bar{t}$ CR	W+jets CR	QCD CR
Observed events	3820	12404	23633
$\sum$ Post-fit events	3820 ± 60	12400 ± 110	23630 ± 150
Post-fit $t\bar{t}$ events	3300 ± 60	1414 ± 26	2425 ± 46
Post-fit W+jets events	39.6 ± 0.6	9200 ± 130	14270 ± 200
Post-fit diboson events	2.99 ± 0.00	396.11 ± 0.00	629.69 ± 0.00
Post-fit single top events	434.11 ± 0.00	352.33 ± 0.00	587.06 ± 0.00
Post-fit Z+jets events	11.26 ± 0.00	775.96 ± 0.00	2364.07 ± 0.00
Post-fit QCD events	32.7 ± 2.6	265 ± 21	3360 ± 260
$\sum$ Pre-fit events	5224.6 ± 0.5	21134.7 ± 1.8	36817.7 ± 2.7
Pre-fit $t\bar{t}$ events	4673.8 ± 0.5	2003.01 ± 0.20	3434.83 ± 0.34
Pre-fit W+jets events	74.911 ± 0.007	17383.9 ± 1.7	26965.7 ± 2.7
Pre-fit diboson events	2.99 ± 0.00	396.11 ± 0.00	629.69 ± 0.00
Pre-fit single top events	434.11 ± 0.00	352.34 ± 0.00	587.06 ± 0.00
Pre-fit Z+jets events	11.26 ± 0.00	775.96 ± 0.00	2364.07 ± 0.00
Pre-fit QCD events	27.5825 ± 0.0028	223.403 ± 0.022	2836.34 ± 0.28

The NFs obtained after each iteration of the fit are listed in Table 6.2. From the table, it can be seen that the NF for the $t\bar{t}$ process remains fairly stable at around 70% for both iterations, while the W+jets NF increases slightly from 48% to 53%. The QCD

background was increased by 18% of the pre-fit value by the fit.

Table 6.2.: Normalisation factors for the $t\bar{t}$, W+jets and dijet/QCD backgrounds obtained after each iteration of the background normalisation fit.

Iteration	$\mu_{t\bar{t}}$	$\mu_{W+\text{jets}}$	$\mu_{\text{dijet/QCD}}$
with dijet MC	0.72 ± 0.01	0.48 ± 0.01	0.67 ± 0.04[1]
with QCD estimate	0.71 ± 0.01	0.53 ± 0.01	1.18 ± 0.09

Finally, Figure 6.2 shows the correlation between these NFs after the fit. As expected, all the three NFs are anti-correlated to each other, since to get an agreement with the observed data, if one NF increases, the others must decrease. However, the anti-correlation is strongest between the W+jets NF and the QCD NF ($\approx$ 80%), since the QCD CR includes a considerable number of W+jets events (see Figures 6.1e and 6.1f). However, the $t\bar{t}$ NF has only 18% and 5% anti-correlation with the W+jets and QCD NFs, respectively. This also explains why the $t\bar{t}$ NF stays fairly constant when replacing the dijet MC background with the QCD estimate in the second iteration of the fit, while the W+jets NF shows a larger change.

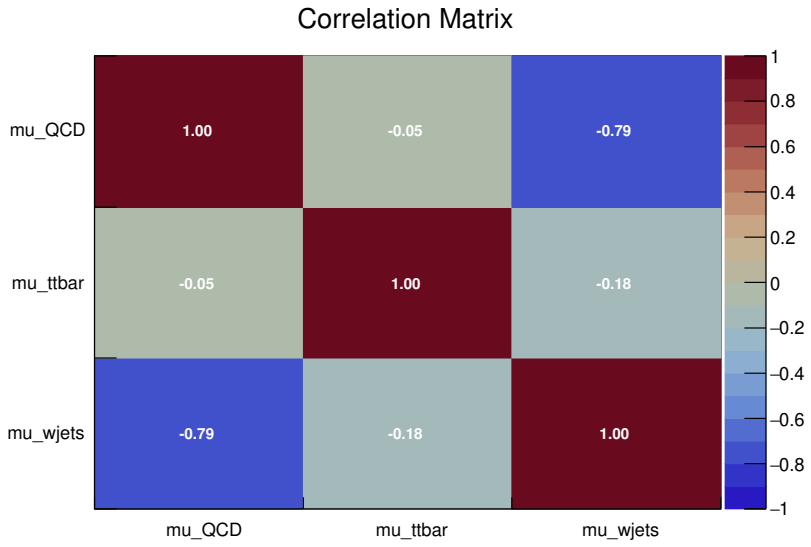


Figure 6.2.: Correlation matrix between all three NFs for the background normalisation fit with the QCD estimate.

[1]The NFs for dijet MC after the first iteration is only listed here for completion and was not used for the estimate of the non-prompt background.

6.2. Exclusion Fits

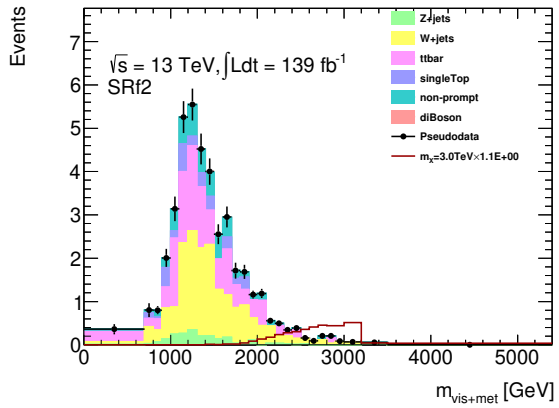
The actual likelihood fits for the analysis are performed using MC samples with $\int \mathcal{L}\,\mathrm{d}t = 139\,\mathrm{fb}^{-1}$ in the three SRs defined at the end of Section 4.3. The nature of the fits is defined by profile likelihood fits (see Section 2.4) with the asymptotic approximation method [51]. For this, pseudodata is used, which is set to the background-only expectation and defined as the sum of all weighted and normalised background events.

In the exclusion fits, the $t\bar{t}$, W+jets and QCD backgrounds are fixed to the NFs obtained from the background normalisation fit. The uncertainties on the NFs will be included as a separate systematic uncertainty. However, this is outside the scope of this thesis, since the implementation of systematics did not fit into the timescale, but is a key task for the publication of the analysis results. Hence, the only uncertainties included in these analysis fits are the uncertainties on the the MC statistics.

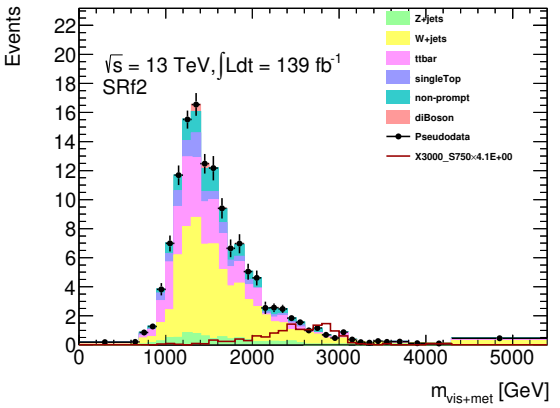
Further, the shape of the $m_{\text{vis+met}}$ variable is used for fitting in the SRs, which was chosen as the final discriminant for the analysis. This is simply defined as the sum of the invariant mass of the selected visible objects and the missing transverse energy. The pre-fit distributions for this variable in the three SRs for $m_X = 3.0\,\text{TeV}$ HH sample can be seen in Figure 6.3. Using these fits, the 95% C.L. expected exclusion limits were derived on $\sigma(pp \to X \to HH)$ and $\sigma(pp \to X \to SH) \times \mathcal{BR}(SH \to b\bar{b}WW)$ for the HH and SH samples, respectively, in the blinded SRs with only statistical uncertainties included.

However, it was also noticed that by just changing the binning in the SRs, the fits behaved differently, yielding a difference in the final limits. This can be seen in Figure 6.4, which shows the expected exclusion limits obtained from fits using 18 (red solid line) and 54 (blue solid line) bins between 0 GeV and 5400 GeV in the $m_{\text{vis+met}}$ distribution for all SRs. The fluctuations in these expected exclusion limits, such as the ones at 2.0 TeV in the limits corresponding to 54 bins, indicate an instability in the fit. This was found to originate from statistical fluctuations in bins with low background statistics. In order to resolve this issue, an attempt to fit the $\log(m_{\text{vis+met}})$ was also made, such that the bins in the high-$m_{\text{vis+met}}$ tails are wider. The black solid line in Figure 6.4 shows the exclusion limits obtained with this implementation, and it is evident that the exclusion limits still exhibit a jump at $m_X = 1.8$ and 3.0 TeV.

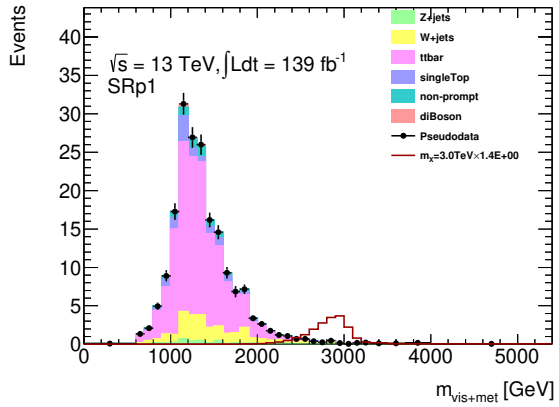
Hence, it is necessary to have well-motivated definition for the binning. Thus, variable-width bins are used, with all bins having at least ten unweighted background events. The post-fit distributions in all three SRs with variable bin widths for the $m_X = 3.0\,\text{TeV}$ HH sample, and the $m_X = 3.0\,\text{TeV}$ and $m_S = 750\,\text{GeV}$ SH sample are shown in Figure 6.5. Since the backgrounds remain the same for all mass points, only a single mass point



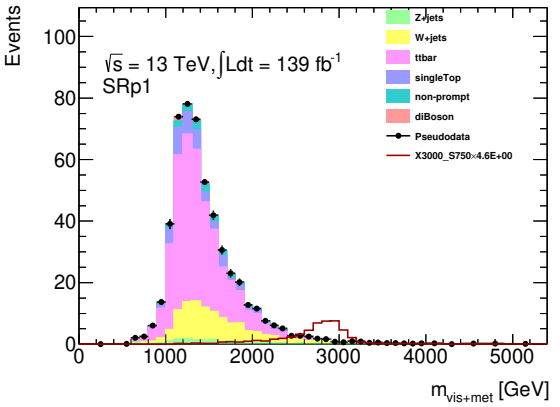
(a) SRf2, HH sample



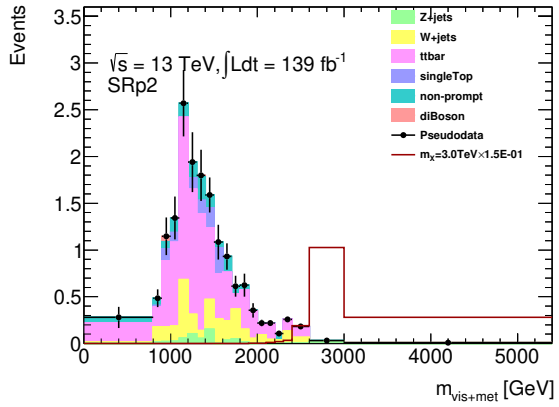
(b) SRf2, SH sample



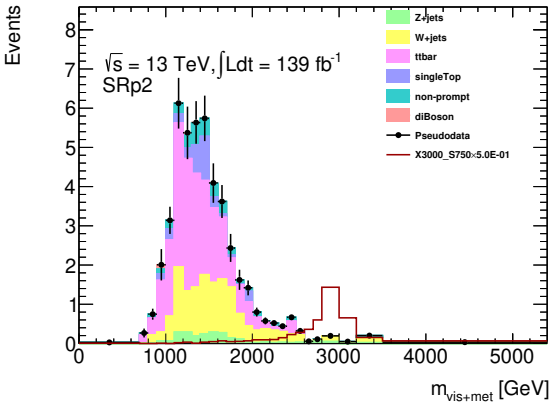
(c) SRp1, HH sample



(d) SRp1, SH sample



(e) SRp2, HH sample



(f) SRp2, SH sample

Figure 6.3.: Pre-fit distributions of $m_{\text{vis+met}}$ in all three SRs with variable bin widths for $m_X = 3.0\,\text{TeV}$ HH sample, and $m_X = 3.0\,\text{TeV}$ and $m_S = 750\,\text{GeV}$ SH sample. The signal has been scaled up to 10% of the total background.

is shown here for the HH and SH samples, each. Distributions for a few more mass points can be found in Appendix B.2. However, apart from the variable bin widths, these

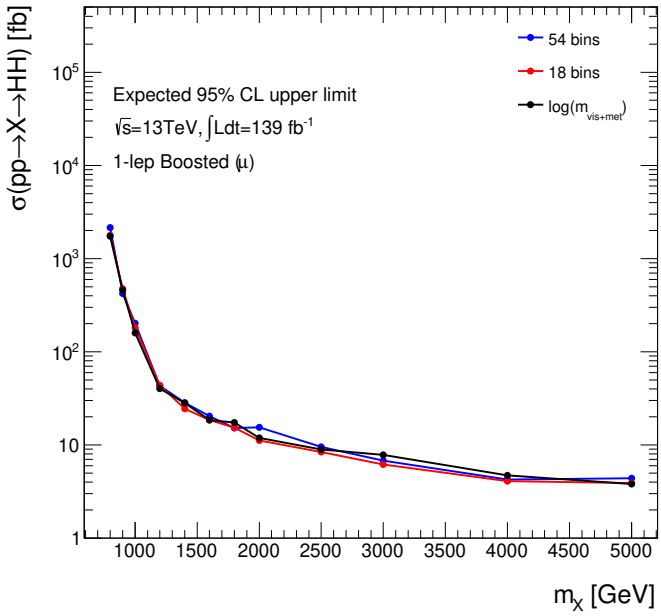


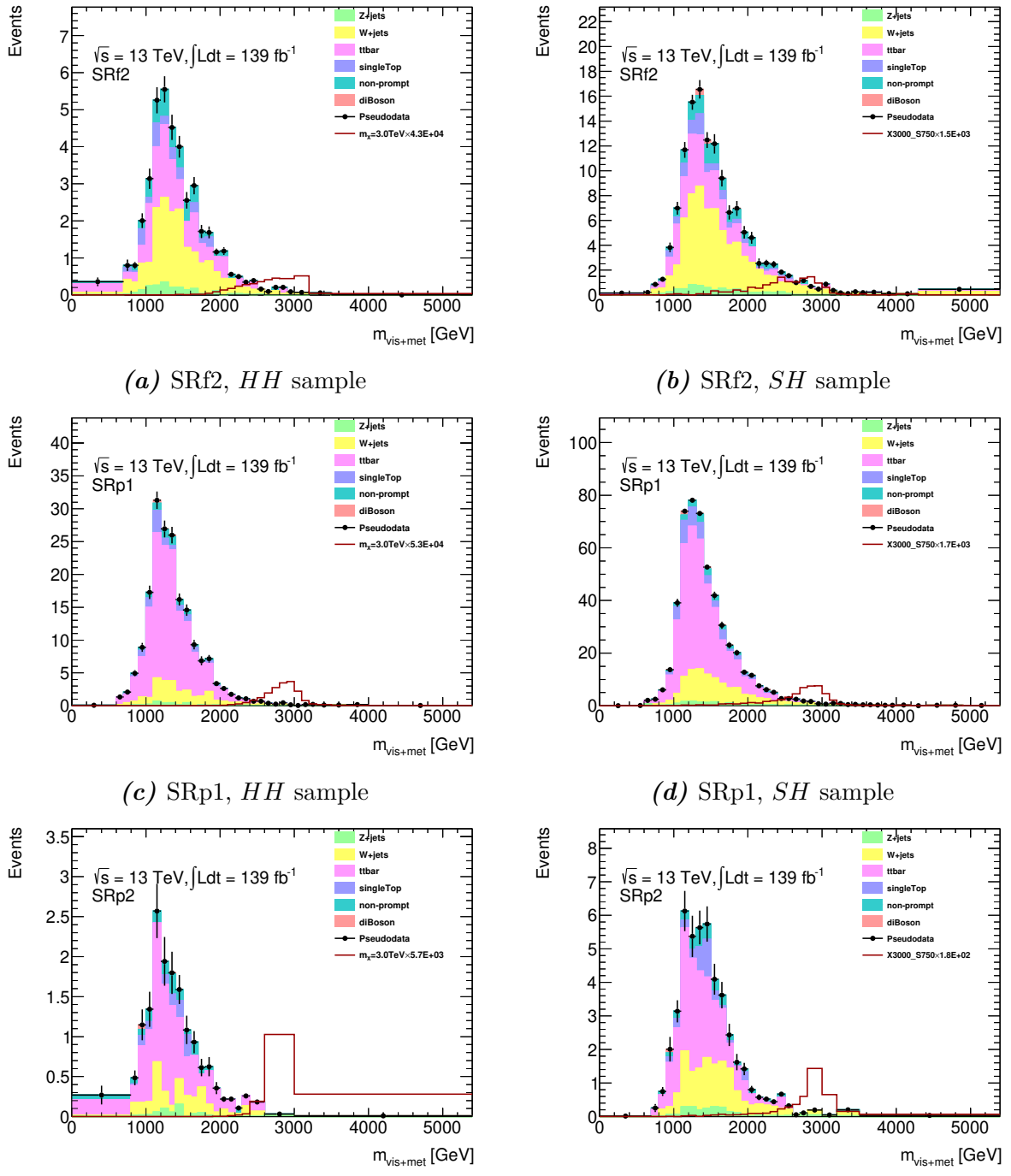
Figure 6.4.: Expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ using different binning definitions for the $m_{\text{vis+met}}$ distribution in the SRs while performing the analysis fits. Only statistical uncertainties are included in the shown limits. These limits do not include the selection cuts requiring events to pass the 80% $\Delta R\,(\ell, W_{\text{had}})$ and $C_2^{H\to b\bar{b}}$ windows defined in Section 4.3

distributions also show that for most cases, the distribution of the signal is spread across at least ten bins for most cases.

Additionally, the yields for these particular cases are also shown in Tables 6.3 and 6.4. Yields for a few more mass points can be found in Appendix A.1. The background yields before the fit are exactly the same for other mass points of the respective case. However, the post-fit background yields can vary within the MC statistical uncertainties. The expected exclusion limits with variable bin widths are shown in the following section.

6.3. Expected Exclusion Limits with Statistical Uncertainties Only

This section provides the final expected exclusion limits with statistical uncertainties only using the full Run 2 ATLAS data for the HH and SH scenarios. Only plots of these limits are shown in this section. The tables with the exact numbers can be found in Appendix A.2.



(a) SRf2, HH sample

(b) SRf2, SH sample

(c) SRp1, HH sample

(d) SRp1, SH sample

(e) SRp2, HH sample

(f) SRp2, SH sample

Figure 6.5.: Post-fit distributions of $m_{\text{vis+met}}$ in all three SRs with variable bin widths for $m_X = 3.0\,\text{TeV}$ HH sample, and $m_X = 3.0\,\text{TeV}$ and $m_S = 750\,\text{GeV}$ SH sample. The signal has been scaled up to 10% of the total background.

$X \to HH \to b\bar{b}WW^*$

Figure 6.6 shows the expected exclusion limit on $\sigma(pp \to X \to HH)$ with the finalised event selection for the boosted 1-lepton final state (black dashed line) along with the

Table 6.3.: Signal and background yields in all three SRs before and after the analysis fit for $m_X = 3.0\,\text{TeV}$ HH sample. The pre-fit background yields for other mass points are exactly the same, while the post-fit backgrounds can vary within the MC statistical uncertainties.

Region	SRf2	SRp1	SRp2
Expected events	40	186	15
$\sum$ Post-fit events	40.3 ± 1.6	186 ± 4	15.8 ± 1.3
Post-fit $t\bar{t}$ events	12.8 ± 0.5	134.8 ± 3.1	9.5 ± 0.7
Post-fit W+jets events	16.7 ± 0.7	25.8 ± 0.6	3.08 ± 0.21
Post-fit diboson events	0.0280 ± 0.0027	0.887 ± 0.034	0.073 ± 0.015
Post-fit single top events	3.00 ± 0.16	11.48 ± 0.30	1.00 ± 0.11
Post-fit Z+jets events	2.13 ± 0.09	3.55 ± 0.08	0.56 ± 0.05
Post-fit QCD events	5.61 ± 0.22	9.88 ± 0.22	1.56 ± 0.11
Post-fit Sig (3.0 TeV) events	0.00 ± 0.29	0.0 ± 1.1	0.0 ± 0.8
$\sum$ Pre-fit events	44.0 ± 1.7	200 ± 4	26.4 ± 1.1
Pre-fit $t\bar{t}$ events	12.8 ± 0.5	134.8 ± 3.2	9.5 ± 0.7
Pre-fit W+jets events	16.8 ± 0.7	25.8 ± 0.6	3.08 ± 0.22
Pre-fit diboson events	0.0280 ± 0.0028	0.89 ± 0.04	0.073 ± 0.016
Pre-fit single top events	3.00 ± 0.17	11.48 ± 0.31	1.00 ± 0.11
Pre-fit Z+jets events	2.13 ± 0.09	3.55 ± 0.09	0.56 ± 0.05
Pre-fit QCD events	5.61 ± 0.23	9.88 ± 0.23	1.57 ± 0.11
Pre-fit Sig (3.0 TeV) events	3.63 ± 0.09	13.69 ± 0.07	10.60 ± 0.04

boosted 0-lepton final state (dark pink dashed line) for comparison[2]. It can be seen that the 1-lepton final state is about three times more sensitive for $m_X \leq 2.5\,\text{TeV}$, while for $m_X \geq 3.0\,\text{TeV}$, the 0-lepton final state starts to dominate and is also up to three times better than the 1-lepton final state for high values of m_X. This trend can be associated to the unique lepton-in-jet topology of the 1-lepton final state which helps in rejecting the backgrounds. However, as m_X increases, the final states become more boosted and lepton identification efficiency drops. In addition to this, the background event yield also reduces for the 0-lepton final state at high m_X, making the 0-lepton final state more dominant in this region.

The best limit for the 1-lepton final state is still expected for $m_X = 5.0\,\text{TeV}$ at 2.8 fb. This value is also compared to the limits obtained by the boosted $HH \to b\bar{b}b\bar{b}$ analysis which also has the highest branching ratio ($\mathcal{BR}(HH \to b\bar{b}b\bar{b}) = 33\%$). The limits for the boosted $b\bar{b}b\bar{b}$ decay channel are also shown in Figure 6.6 (purple dashed line) [107, 108].

[2]The 0-lepton expected exclusion limits also include a 10% background estimation systematic uncertainty in addition to the statistical uncertainties.

Table 6.4.: Signal and background yields in all three SRs before and after the analysis fit for $m_X = 3.0\,\text{TeV}$ and $m_S = 750\,\text{GeV}$ SH sample. The pre-fit background yields for other mass points are exactly the same, while the post-fit backgrounds can vary within the MC statistical uncertainties.

Region	SRf2	SRp1	SRp2
Expected events	131	517	46
$\sum$ Post-fit events	131 ± 4	518 ± 6	46 ± 4
Post-fit $t\bar{t}$ events	34.1 ± 1.0	324.9 ± 3.2	22.9 ± 1.2
Post-fit W+jets events	65.2 ± 1.8	103.8 ± 1.4	13.5 ± 0.7
Post-fit diboson events	1.11 ± 0.06	0.887 ± 0.034	4.83 ± 0.10
Post-fit single top events	9.06 ± 0.27	45.0 ± 0.5	3.57 ± 0.25
Post-fit Z+jets events	7.03 ± 0.20	14.52 ± 0.29	2.34 ± 0.12
Post-fit QCD events	14.9 ± 0.4	24.75 ± 0.32	3.69 ± 0.19
Pre-fit Sig ($m_X = 3.0\,\text{TeV}$, $m_S = 750\,\text{GeV}$) events	0.0 ± 1.1	0.0 ± 3.4	0.0 ± 2.9
$\sum$ Pre-fit events	135 ± 4	529 ± 6	55.3 ± 2.6
Pre-fit $t\bar{t}$ events	34.1 ± 1.0	324.9 ± 3.5	22.9 ± 1.3
Pre-fit W+jets events	65.3 ± 2.0	103.9 ± 1.5	13.5 ± 0.7
Pre-fit diboson events	1.108 ± 0.06	4.826 ± 0.11	0.205 ± 0.026
Pre-fit single top events	9.07 ± 0.29	45.0 ± 0.5	3.57 ± 0.27
Pre-fit Z+jets events	7.04 ± 0.21	14.54 ± 0.33	2.34 ± 0.13
Pre-fit QCD events	15.0 ± 0.5	24.75 ± 0.35	3.70 ± 0.20
Pre-fit Sig ($m_X = 3.0\,\text{TeV}$, $m_S = 750\,\text{GeV}$) events	3.59 ± 0.17	10.90 ± 0.32	9.06 ± 0.31

At $m_X = 5.0\,\text{TeV}$, the limit from the $b\bar{b}b\bar{b}$ analysis is 1.6 fb, which is 1.8 times better than the limits from the 1-lepton $b\bar{b}WW^*$ limits.

$X \to SH \to b\bar{b}WW$

Figure 6.7 shows the expected exclusion limit on $\sigma(pp \to X \to SH) \times \mathcal{BR}(SH \to b\bar{b}WW)$ as a function of m_S also with the finalised event selection for the boosted 1-lepton final state. These are shown as independent plots for each m_X. From the plots, a general trend of limits being better for smaller m_S values can be seen. Furthermore, the limits for a particular m_S also improves with increasing m_X. Both of these scenarios (high m_X and low m_S) correspond to boosted topologies, where the analysis was also expected to be more sensitive due to the $\min[\Delta R(\ell, J)] \leq 1.0$ cut applied in the preselection, as discussed in Section 5.3. An interesting observation is that for $m_X \geq 2.0\,\text{TeV}$, the

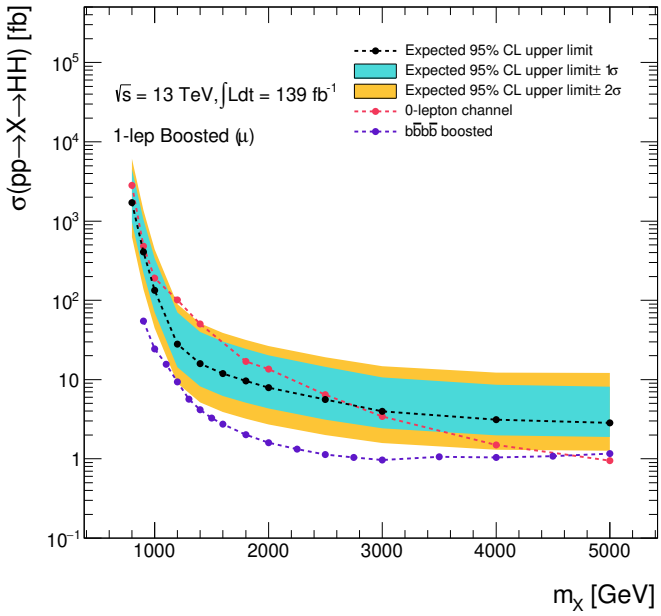


Figure 6.6.: Expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ with the finalised event selection and full Run 2 data. The exclusion limits from the 0-lepton final state and the boosted $b\bar{b}b\bar{b}$ channel are also shown in dark pink and purple, respectively, for comparison. The expected limits from the 0-lepton final state include a 10% systematic uncertainty on the background estimate, while the other limits only include statistical uncertainties.

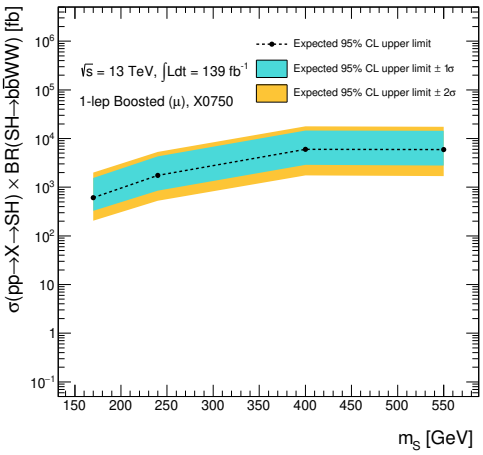
limits for $m_S = 240\,\text{GeV}$ is better than for $m_S = 170\,\text{GeV}$ after which it again starts to worsen. This can again be explained by the argument that for highly boosted scenarios, the lepton identification efficiency drops. The best limit for the SH scenario are expected for $m_X = 3.0\,\text{TeV}$ and $m_S = 240\,\text{GeV}$ at $0.87\,\text{fb}$.

The limits for the SH scenario can also be compared with the limits for the HH samples by assuming the SM Higgs boson to be the scalar particle, S, with $m_S = 125.10\,\text{GeV}$. However, since the final selections for the HH and SH are not identical, the comparison was made between expected exclusion limits obtained for both cases without the selection cuts requiring events to pass the 80% $\Delta R\,(\ell, W_{\text{had}})$ (only for HH samples) and $C_2^{H\to b\bar{b}}$ windows defined in Section 4.3. Additionally, since the HH limits take into account the $\mathcal{BR}(HH \to b\bar{b}WW^*) = 25\%$, this factor must be corrected before comparing both the scenarios. After applying this correction, the HH limits ($m_S = 125.1\,\text{GeV}$) are about 20-30% worse than the SH limits at $m_S = 170\,\text{GeV}$. This can again be due to the lepton identification efficiency for highly boosted cases, but for lower m_X, this is mostly due to one of the two W bosons being off-shell in case of the HH scenario, while for the SH

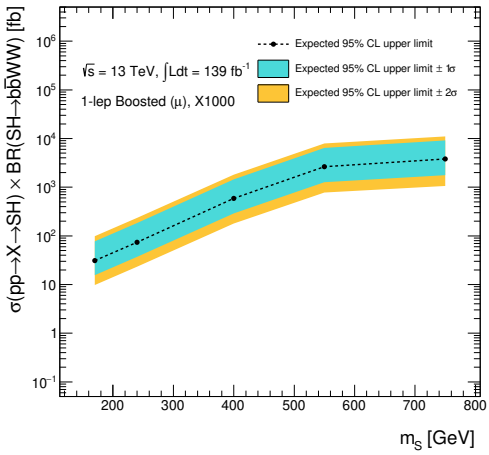
case, they can both be simultaneously on-shell.

6.4. Addition of Preliminary Systematic Uncertainties

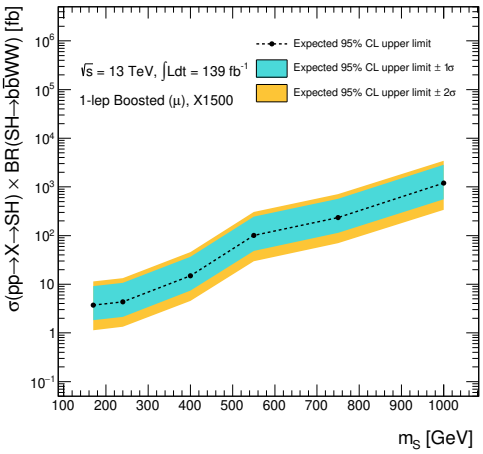
Until now, the discussed limits only included uncertainties originating from data and MC statistics. The inclusion of systematic uncertainties is one key task that still remains to be implemented. However, the analysis is expected to be dominated by statistical uncertainties rather than systematic uncertainties, which can also be seen in Figure 6.8. The figure shows the effect of adding dummy systematics as ratios of limits with and without the systematics. These dummy systematics were added as fully uncorrelated systematics across all background samples. From the figure it is evident that the addition of a flat 10% dummy systematic (red lines) has only a 2% effect on the expected exclusion limits, while with a flat 20% dummy systematic (blue lines), the limits worsen by about 6-10%. Further, it can be seen that the addition of uncertainties on MC statistics (black solid line), which were already included in the expected exclusion limits shown in the previous section, has about 5% effect. This goes on to show that the analysis is dominated by statistical uncertainties rather than systematic uncertainties, which are expected to have a small effect on the final expected exclusion limits. It should, however, be noted that this study was conducted before the selection was finalised and even before it was decided to switch the exclusion fits to have variable bin widths. Hence the limits used for this plot were obtained using uniform bin widths in the SRs without the optimisation cuts on the 80% $\Delta R\,(\ell, W_{\mathrm{had}})$ and $C_2^{H\to b\bar{b}}$ windows. However, these caveats are only expected to have a minor effect on the overall picture.



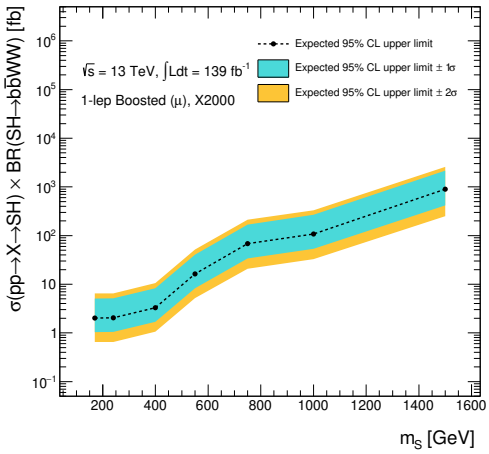
(a) $m_X = 0.75\,\text{TeV}$



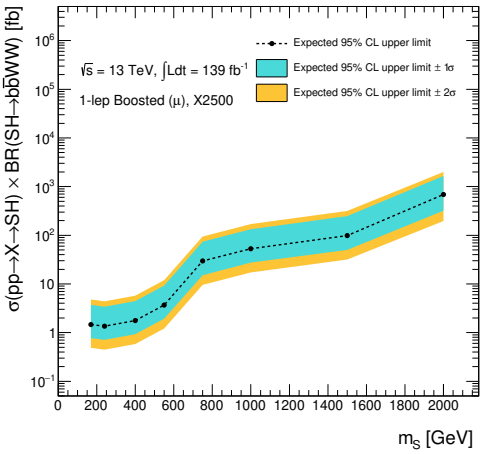
(b) $m_X = 1.0\,\text{TeV}$



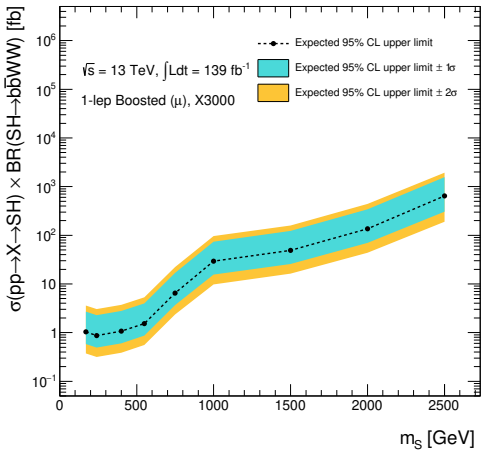
(c) $m_X = 1.5\,\text{TeV}$



(d) $m_X = 2.0\,\text{TeV}$



(e) $m_X = 2.5\,\text{TeV}$



(f) $m_X = 3.0\,\text{TeV}$

Figure 6.7.: Expected 95% C.L. exclusion limits using the finalised event selection on $\sigma(pp \to X \to SH) \times \mathcal{BR}(SH \to b\bar{b}WW)$ with full Run 2 data. The only uncertainties included in shown exclusion limits are the ones originating from data and MC statistics.

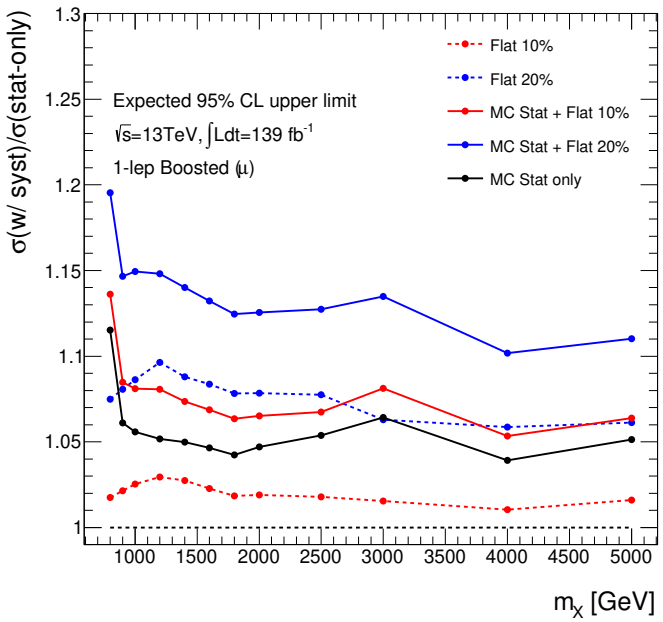


Figure 6.8.: Ratio of expected 95% C.L. exclusion limits with and without fully uncorrelated dummy systematics. The solid black line depicts the effect on MC statistical uncertainties on the exclusion limits.

7. Conclusion and Outlook

> "One never notices what has been done; one can only see what remains to be done."
>
> - Marie Curie

This thesis presented the searches conducted for heavy scalar resonances (X) decaying into two SM-like Higgs bosons, $X \to HH$, or an intermediate mass scalar in conjunction with a SM Higgs boson, $X \to SH$, in the $b\bar{b}WW^{(*)}$ decay channel using the full Run 2 ATLAS pp collision data. The 1-lepton final state in boosted topologies was discussed. This included a discussion about the event topology of the channel under investigation followed by the definition of all final state objects. The finalised event selection cuts were also motivated along with detailed study about the need to optimise boosted and split-boosted topologies independently. For this, an additional cut was placed at $\min\left[\Delta R\left(\ell, J\right)\right] \leq 1.0$ to be consistent with the the boosted selection.

Further, the statistical framework was described including the background normalisation fits as an iterative procedure of simultaneous single-bin fits in the CRs, and analysis fit as simultaneous shape fits in the final discriminant, $m_{\text{vis+met}}$, with variable bin widths in the SRs. Finally, the HH exclusion limits were discussed and compared with the $b\bar{b}WW^{(*)}$ 0-lepton final state and the $b\bar{b}b\bar{b}$ decay channels, and found to be in the same order of magnitude. The SH exclusion limits were seen to be generally better for more boosted topologies (large $\Delta M_{X,S}$). The best limits are expected for $m_X = 5.0\,\text{TeV}$ at 2.8 fb for HH case, and for $m_X = 3.0\,\text{TeV}$ and $m_S = 240\,\text{GeV}$ at 0.87 fb for the SH case.

These exclusion limits, however, only include uncertainties originating from data (via the QCD estimate) and MC statistics. The effect of dummy systematics was also studied, which showed that the analysis is dominated by statistical uncertainties and addition of systematic uncertainties only has a minor effect on the overall picture. However, the inclusion of full systematics still remains a key task that needs to be implemented.

Further, the final step in the analysis would be to combine the limits obtained by the boosted 1-lepton and 0-lepton final states. As discussed in Chapter 6, the 1-lepton channel is more sensitive for $m_X \leq 2.5\,\text{TeV}$, while the 0-lepton channel dominates for

$m_X \geq 3.0\,\text{TeV}$. Hence, both the channels would have a significant contribution to the combined limits for the boosted $X \to HH$ limits in the $b\bar{b}WW^*$ decay channel. Future plans to combine the exclusion limits obtained by the the different HH decay channels are also in place.

Overall, the search for heavy scalars decaying into a pair of Higgs bosons is a crucial task in order to test the predictions of various BSM theories. For this, exploiting every feasible decay channel, such as the $b\bar{b}WW^{(*)}$ decay channel considered here, is of great importance to obtain a strong combination of exclusion limits. These further are important in understanding how rare such interactions could be, and how significant their effect would be in shaping the understanding of the universe now and in the future.

A. Additional Tables

A.1. Signal and Background Yields

Table A.1.: Signal and background yields for additional SH mass points in the three SRs

m_X	SR	Expected events	Signal+Background	Signal
1.0 TeV	SRf2	40	40.2 ± 1.9	0.0 ± 0.8
	SRp1	186	186 ± 7	0 ± 5
	SRp2	15	16 ± 9	0 ± 9
5.0 TeV	SRf2	40	40.3 ± 1.8	0.00 ± 0.35
	SRp1	186	186 ± 5	0.0 ± 3.1
	SRp2	15	15.7 ± 1.6	0.0 ± 1.1

Table A.2.: Signal and background yields for additional SH mass points in the three SRs

m_X	m_S	SR	Expected events	Signal+Background	Signal
3.0 TeV	240 GeV	SRf2	131	131.5 ± 3.5	0.0 ± 0.6
		SRp1	517	518 ± 5	0.0 ± 1.9
		SRp2	46	46.3 ± 2.6	0.0 ± 1.6
3.0 TeV	1500 GeV	SRf2	131	132 ± 4	0.0 ± 1.8
		SRp1	517	518 ± 8	0 ± 5
		SRp2	46	46 ± 6	0 ± 5

A.2. Expected Exclusion Limits

Table A.3.: Expected 95% C.L. exclusion limits on $\sigma(pp \to X \to HH)$ with only statistical uncertainties included.

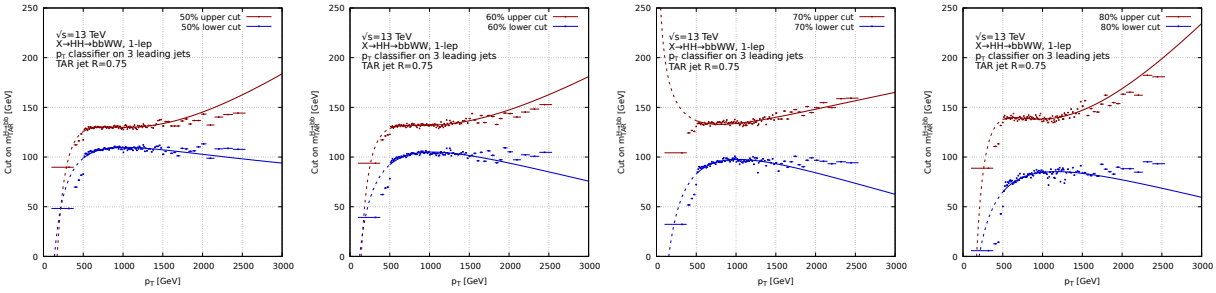
m_X [TeV]	median$^{+1\sigma+2\sigma}_{-1\sigma-2\sigma}$ [fb]
0.8	$1700^{+1100+2700}_{-600-1000}$
0.9	$410^{+230+550}_{-140-220}$
1.0	$130^{+70+170}_{-40-70}$
1.2	28^{+14+34}_{-9-14}
1.4	16^{+8+19}_{-5-8}
1.6	12^{+6+15}_{-4-6}
1.8	$9.6^{+5.2+12.5}_{-3.2-5.1}$
2.0	$7.9^{+4.4+10.7}_{-2.7-4.3}$
2.5	$5.6^{+3.3+8.0}_{-2.0-3.1}$
3.0	$4.0^{+2.7+6.8}_{-1.6-2.4}$
4.0	$3.1^{+2.4+6.1}_{-1.3-2.0}$
5.0	$2.8^{+2.4+6.4}_{-1.3-1.9}$

Table A.4.: Expected 95% C.L. exclusion limits on $\sigma(pp \to X \to SH) \times \mathcal{BR}(SH \to b\bar{b}WW)$ with only statistical uncertainties included.

m_S \ m_X	0.75 TeV	1.0 TeV	1.5 TeV	2.0 TeV	2.5 TeV	3.0 TeV
170 GeV	$610^{+330+780}_{-200-330}$	31^{+15+}_{-10-16}	$3.7^{+1.7+4.0}_{-1.1-1.8}$	$2.0^{+1.0+2.4}_{-0.7-1.0}$	$1.5^{+0.8+1.9}_{-0.5-0.8}$	$1.0^{+0.6+1.5}_{-0.4-0.6}$
240 GeV	$1800^{+800+1800}_{-500-800}$	74^{+36+}_{-23-37}	$4.3^{+2.0+4.7}_{-1.3-2.1}$	$2.0^{+1.0+2.4}_{-0.7-1.0}$	$1.3^{+0.7+1.7}_{-0.4-0.7}$	$0.87^{+0.53+1.29}_{-0.32-0.49}$
400 GeV	$6000^{+2600+5800}_{-1800-2900}$	$590^{+270+}_{-180-290}$	15^{+7+16}_{-5-7}	$3.3^{+1.7+4.0}_{-1.1-1.7}$	$1.8^{+0.9+2.2}_{-0.6-0.9}$	$1.1^{+0.6+1.6}_{-0.4-0.6}$
550 GeV	$6000^{+2400+5500}_{-1700-2800}$	$2600^{+1200+}_{-800-1300}$	$100^{+44+104}_{-30-48}$	16^{+8+19}_{-5-8}	$3.7^{+1.9+4.6}_{-1.2-1.9}$	$1.5^{+0.9+2.2}_{-0.6-0.9}$
750 GeV	-	$3800^{+1500+}_{-1100-1800}$	$230^{+100+240}_{-70-110}$	68^{+32+74}_{-21-34}	30^{+15+35}_{-10-15}	$6.5^{+3.8+9.3}_{-2.4-3.7}$
1000 GeV	-	-	$1190^{+470+1060}_{-340-560}$	$108^{+50+116}_{-33-53}$	53^{+27+64}_{-17-27}	29^{+16+38}_{-10-16}
1500 GeV	-	-	-	$900^{+350+790}_{-250-420}$	$98^{+50+119}_{-32-50}$	49^{+26+62}_{-16-26}
2000 GeV	-	-	-	-	$690^{+280+620}_{-200-320}$	$140^{+70+167}_{-40-70}$
2500 GeV	-	-	-	-	-	$650^{+280+630}_{-190-310}$

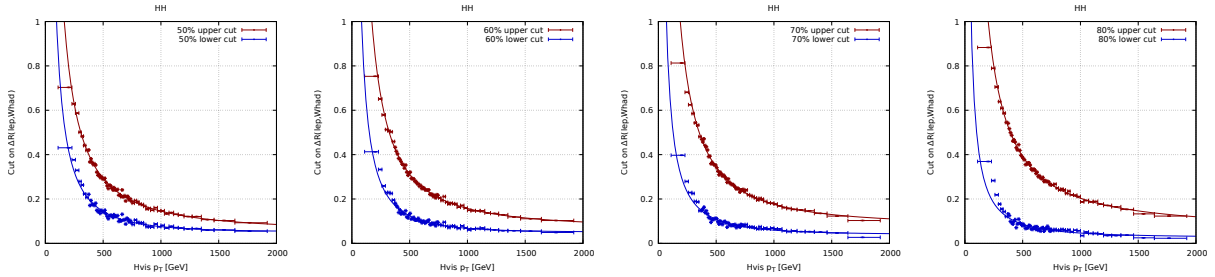
B. Additional Figures

B.1. Floating Windows Dependent on p_T



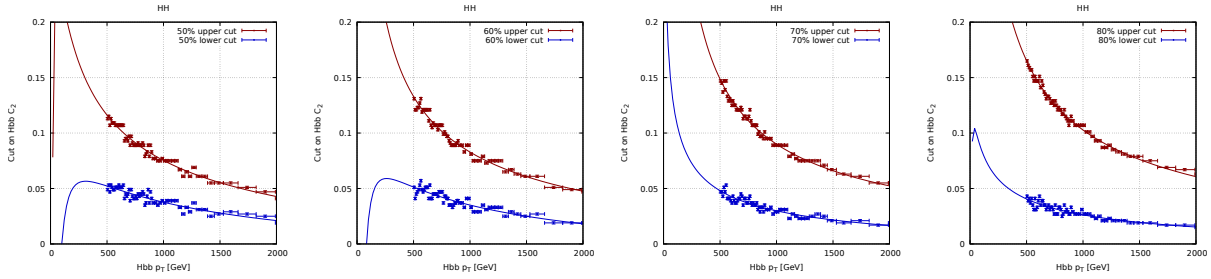
(a) 50% WP ***(b)*** 60% WP ***(c)*** 70% WP ***(d)*** 80% WP

Figure B.1.: Different WPs for $m_{\text{TAR}}^{H\to b\bar{b}}$ windows dependent on $p_T^{H\to b\bar{b}}$ [98].



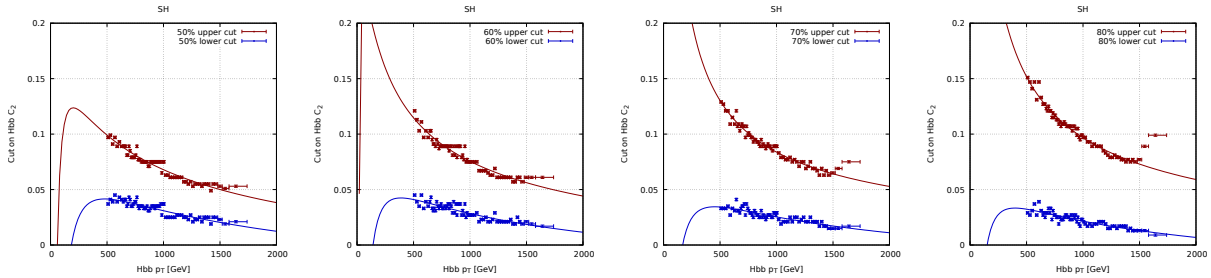
(a) 50% WP ***(b)*** 60% WP ***(c)*** 70% WP ***(d)*** 80% WP

Figure B.2.: Different WPs for $\Delta R\,(\ell, W_{\text{had}})$ windows dependent on $p_T^{H_{\text{vis}}}$ [98].



(a) 50% WP ***(b)*** 60% WP ***(c)*** 70% WP ***(d)*** 80% WP

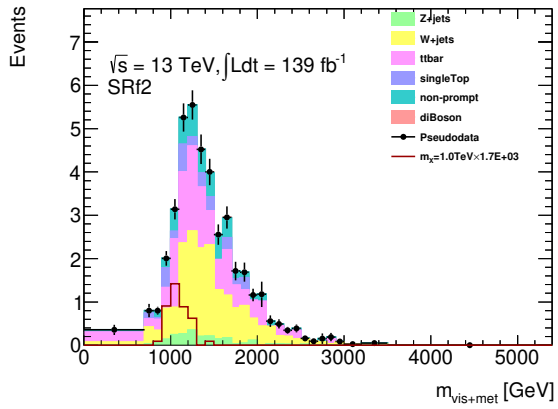
Figure B.3.: Different WPs for $C_2^{H\to b\bar{b}}$ windows dependent on $p_T^{H\to b\bar{b}}$ using HH signal samples [98].



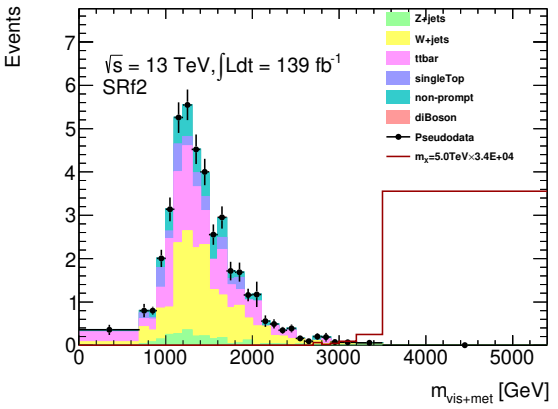
(a) 50% WP ***(b)*** 60% WP ***(c)*** 70% WP ***(d)*** 80% WP

Figure B.4.: Different WPs for $C_2^{H\to b\bar{b}}$ windows dependent on $p_T^{H\to b\bar{b}}$ using SH signal samples [98].

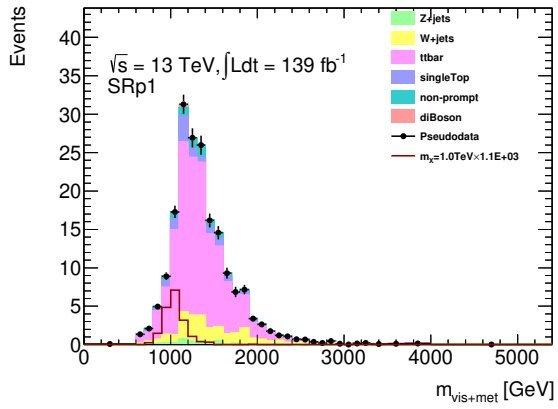
B.2. Post-fit Distributions of $m_{\text{vis+met}}$ in SRs



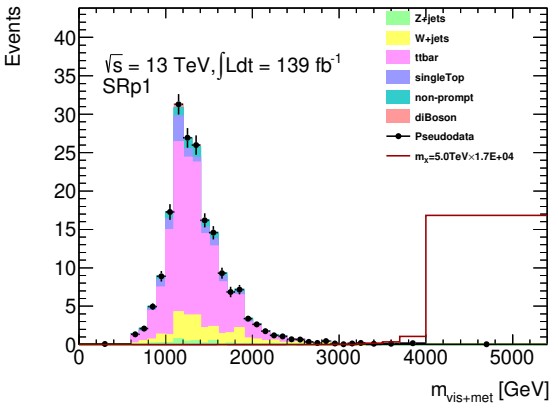
(a) SRf2, $m_X = 1.0\,\text{TeV}$



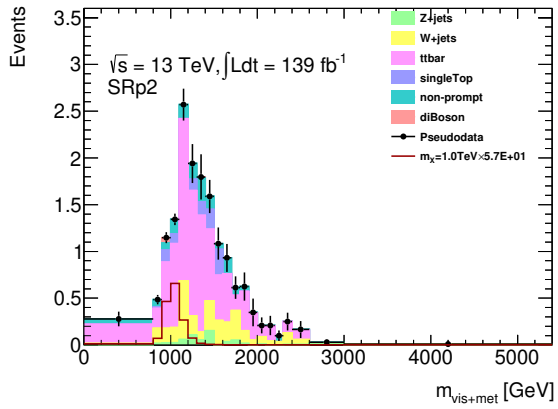
(b) SRf2, $m_X = 5.0\,\text{TeV}$



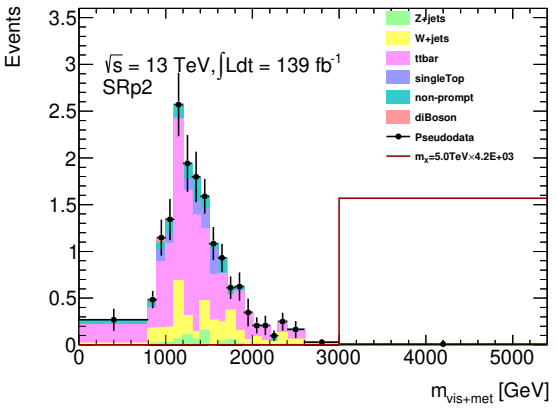
(c) SRp1, $m_X = 1.0\,\text{TeV}$



(d) SRp1, $m_X = 5.0\,\text{TeV}$

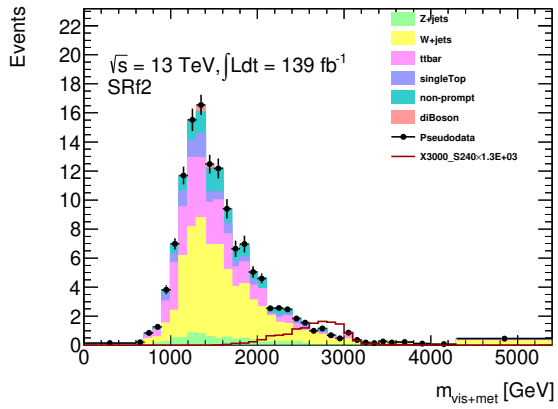


(e) SRp2, $m_X = 1.0\,\text{TeV}$

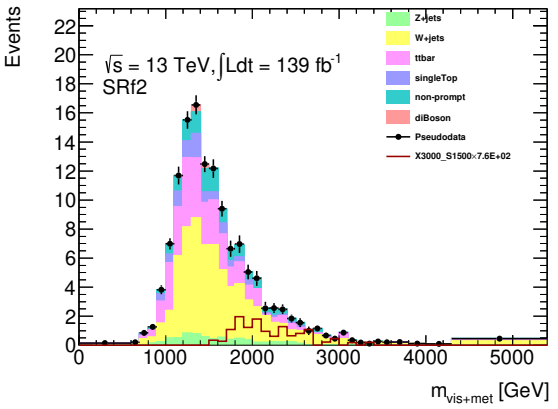


(f) SRp2, $m_X = 5.0\,\text{TeV}$

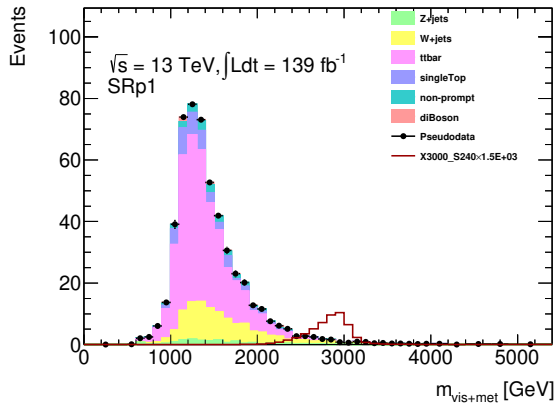
Figure B.5.: Post-fit distributions of $m_{\text{vis+met}}$ in all three SRs with variable bin widths for $m_X = 1.0$ and $5.0\,\text{TeV}$ HH samples. The signal has been scaled up to 10% of the total background.



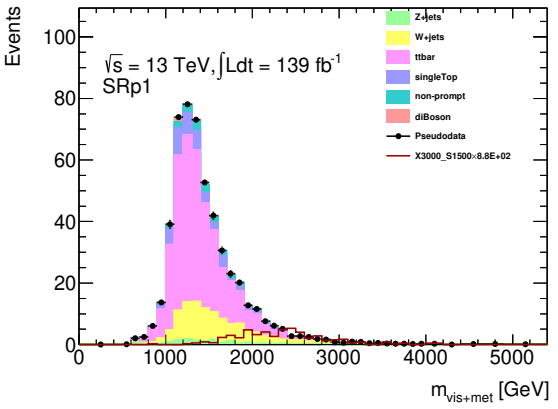
(a) SRf2, $m_X = 3.0\,\text{TeV}$ and $m_S = 240\,\text{GeV}$



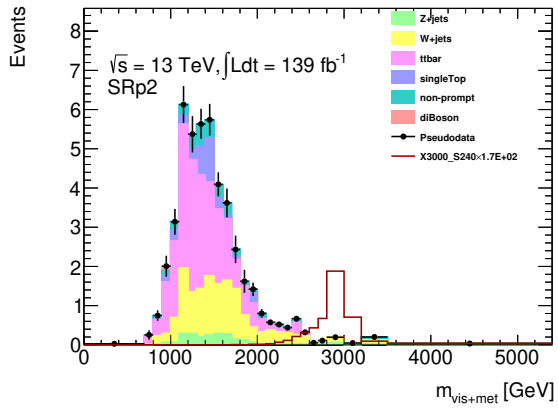
(b) SRf2, $m_X = 3.0\,\text{TeV}$ and $m_S = 1500\,\text{GeV}$



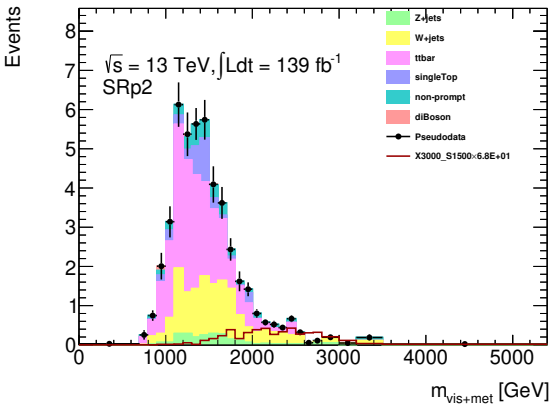
(c) SRp1, $m_X = 3.0\,\text{TeV}$ and $m_S = 240\,\text{GeV}$



(d) SRp1, $m_X = 3.0\,\text{TeV}$ and $m_S = 1500\,\text{GeV}$



(e) SRp2, $m_X = 3.0\,\text{TeV}$ and $m_S = 240\,\text{GeV}$



(f) SRp2, $m_X = 3.0\,\text{TeV}$ and $m_S = 1500\,\text{GeV}$

Figure B.6.: Post-fit distributions of $m_{\text{vis+met}}$ in all three SRs with variable bin widths for $m_S = 240$ and $1500\,\text{GeV}$ SH samples, both having $m_X = 3.0\,\text{TeV}$. The signal has been scaled up to 10% of the total background.

C. List of MC Samples

This section lists all of the Monte Carlo samples in this analysis. All of the samples listed below are available in the mc16a, mc16d, and mc16e campaigns. These are indicated by the r-tag (rXXXX), where mc16a corresponds to r9364, mc16d corresponds to r10201, and mc16e corresponds to r10724.

C.1. Signal samples

$X \to HH \to bbWW^*$ 1-lepton:

mc16_13TeV.450220.MadGraphHerwig7EvtGen_PDF23LO_X800tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7592_a875_rXXXX_p4128
mc16_13TeV.450221.MadGraphHerwig7EvtGen_PDF23LO_X900tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7592_a875_rXXXX_p4128
mc16_13TeV.450219.MadGraphHerwig7EvtGen_PDF23LO_X1000tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7329_a875_rXXXX_p4128
mc16_13TeV.450222.MadGraphHerwig7EvtGen_PDF23LO_X1200tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7592_a875_rXXXX_p4128
mc16_13TeV.450223.MadGraphHerwig7EvtGen_PDF23LO_X1400tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7592_a875_rXXXX_p4128
mc16_13TeV.450224.MadGraphHerwig7EvtGen_PDF23LO_X1600tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7592_a875_rXXXX_p4128
mc16_13TeV.450225.MadGraphHerwig7EvtGen_PDF23LO_X1800tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7592_a875_rXXXX_p4128
mc16_13TeV.450229.MadGraphHerwig7EvtGen_PDF23LO_X2000tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7329_a875_rXXXX_p4128
mc16_13TeV.450226.MadGraphHerwig7EvtGen_PDF23LO_X2500tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7592_a875_rXXXX_p4128
mc16_13TeV.450239.MadGraphHerwig7EvtGen_PDF23LO_X3000tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7329_a875_rXXXX_p4128
mc16_13TeV.450227.MadGraphHerwig7EvtGen_PDF23LO_X4000tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7592_a875_rXXXX_p4128
mc16_13TeV.450228.MadGraphHerwig7EvtGen_PDF23LO_X5000tohh_WWbb_1lep.deriv.DAOD_EXOT8.e7592_a875_rXXXX_p4128

$X \to SH \to bbWW$ 1-lepton:

mc16_13TeV.800751.Py8EG_A14NNPDF23LO_XHS_X350_S170_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800752.Py8EG_A14NNPDF23LO_XHS_X500_S170_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800753.Py8EG_A14NNPDF23LO_XHS_X500_S240_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800754.Py8EG_A14NNPDF23LO_XHS_X750_S170_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800755.Py8EG_A14NNPDF23LO_XHS_X750_S240_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800756.Py8EG_A14NNPDF23LO_XHS_X750_S400_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800757.Py8EG_A14NNPDF23LO_XHS_X750_S550_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800758.Py8EG_A14NNPDF23LO_XHS_X1000_S170_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800759.Py8EG_A14NNPDF23LO_XHS_X1000_S240_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800760.Py8EG_A14NNPDF23LO_XHS_X1000_S400_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800761.Py8EG_A14NNPDF23LO_XHS_X1000_S550_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800762.Py8EG_A14NNPDF23LO_XHS_X1000_S750_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800763.Py8EG_A14NNPDF23LO_XHS_X1500_S170_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800764.Py8EG_A14NNPDF23LO_XHS_X1500_S240_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800765.Py8EG_A14NNPDF23LO_XHS_X1500_S400_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800766.Py8EG_A14NNPDF23LO_XHS_X1500_S550_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800767.Py8EG_A14NNPDF23LO_XHS_X1500_S750_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800768.Py8EG_A14NNPDF23LO_XHS_X1500_S1000_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800769.Py8EG_A14NNPDF23LO_XHS_X2000_S170_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800770.Py8EG_A14NNPDF23LO_XHS_X2000_S240_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800771.Py8EG_A14NNPDF23LO_XHS_X2000_S400_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800772.Py8EG_A14NNPDF23LO_XHS_X2000_S550_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800773.Py8EG_A14NNPDF23LO_XHS_X2000_S750_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800774.Py8EG_A14NNPDF23LO_XHS_X2000_S1000_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800775.Py8EG_A14NNPDF23LO_XHS_X2000_S1500_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800776.Py8EG_A14NNPDF23LO_XHS_X2500_S170_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800777.Py8EG_A14NNPDF23LO_XHS_X2500_S240_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800778.Py8EG_A14NNPDF23LO_XHS_X2500_S400_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800779.Py8EG_A14NNPDF23LO_XHS_X2500_S550_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800780.Py8EG_A14NNPDF23LO_XHS_X2500_S750_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800781.Py8EG_A14NNPDF23LO_XHS_X2500_S1000_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800782.Py8EG_A14NNPDF23LO_XHS_X2500_S1500_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800783.Py8EG_A14NNPDF23LO_XHS_X2500_S2000_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800784.Py8EG_A14NNPDF23LO_XHS_X3000_S170_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128

mc16_13TeV.800785.Py8EG_A14NNPDF23LO_XHS_X3000_S240_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800786.Py8EG_A14NNPDF23LO_XHS_X3000_S400_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800787.Py8EG_A14NNPDF23LO_XHS_X3000_S550_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800788.Py8EG_A14NNPDF23LO_XHS_X3000_S750_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800789.Py8EG_A14NNPDF23LO_XHS_X3000_S1000_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800790.Py8EG_A14NNPDF23LO_XHS_X3000_S1500_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800791.Py8EG_A14NNPDF23LO_XHS_X3000_S2000_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128
mc16_13TeV.800792.Py8EG_A14NNPDF23LO_XHS_X3000_S2500_bbWW_1lep.deriv.DAOD_EXOT8.e8312_a875_rXXXX_p4128

C.2. Background samples

Bulk $t\bar{t}$:
mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_EXOT8.e6337_s3126_rXXXX_p4004
mc16_13TeV.410471.PhPy8EG_A14_ttbar_hdamp258p75_allhad.deriv.DAOD_EXOT8.e6337_s3126_rXXXX_p4004
All-hadronic $m_{t\bar{t}}$-sliced $t\bar{t}$:
mc16_13TeV.410284.PhPy8EG_A14_ttbar_hdamp258p75_allhad_mtt_1100_1300.deriv.DAOD_EXOT8.e6603_s3126_rXXXX_p4004
mc16_13TeV.410285.PhPy8EG_A14_ttbar_hdamp258p75_allhad_mtt_1300_1500.deriv.DAOD_EXOT8.e6686_s3126_rXXXX_p4004
mc16_13TeV.410286.PhPy8EG_A14_ttbar_hdamp258p75_allhad_mtt_1500_1700.deriv.DAOD_EXOT8.e6686_s3126_rXXXX_p4004
mc16_13TeV.410287.PhPy8EG_A14_ttbar_hdamp258p75_allhad_mtt_1700_2000.deriv.DAOD_EXOT8.e6686_s3126_rXXXX_p4004
mc16_13TeV.410288.PhPy8EG_A14_ttbar_hdamp258p75_allhad_mtt_2000_14000.deriv.DAOD_EXOT8.e6686_s3126_rXXXX_p4004
Non-allhadronic $m_{t\bar{t}}$-sliced $t\bar{t}$:
mc16_13TeV.410633.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad_1100_1300.deriv.DAOD_EXOT8.e6602_s3126_rXXXX_p4004
mc16_13TeV.410634.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad_1300_1500.deriv.DAOD_EXOT8.e6602_s3126_rXXXX_p4004
mc16_13TeV.410635.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad_1500_1700.deriv.DAOD_EXOT8.e6685_s3126_rXXXX_p4004
mc16_13TeV.410636.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad_1700_2000.deriv.DAOD_EXOT8.e6685_s3126_rXXXX_p4004
mc16_13TeV.410637.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad_2000_14000.deriv.DAOD_EXOT8.e6685_s3126_rXXXX_p4004
Single top:
mc16_13TeV.410644.PowhegPythia8EvtGen_A14_singletop_schan_lept_top.deriv.DAOD_EXOT8.e6527_s3126_rXXXX_p4004
mc16_13TeV.410645.PowhegPythia8EvtGen_A14_singletop_schan_lept_antitop.deriv.DAOD_EXOT8.e6527_s3126_rXXXX_p4004
mc16_13TeV.410646.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_top.deriv.DAOD_EXOT8.e6552_s3126_rXXXX_p4004
mc16_13TeV.410647.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_antitop.deriv.DAOD_EXOT8.e6552_s3126_rXXXX_p4004
mc16_13TeV.410654.PowhegPythia8EvtGen_A14_Wt_DS_inclusive_top.deriv.DAOD_EXOT8.e6552_s3126_rXXXX_p4004
mc16_13TeV.410655.PowhegPythia8EvtGen_A14_Wt_DS_inclusive_antitop.deriv.DAOD_EXOT8.e6552_s3126_rXXXX_p4004
mc16_13TeV.410658.PhPy8EG_A14_tchan_BW50_lept_top.deriv.DAOD_EXOT8.e6671_s3126_rXXXX_p4004
mc16_13TeV.410659.PhPy8EG_A14_tchan_BW50_lept_antitop.deriv.DAOD_EXOT8.e6671_s3126_rXXXX_p4004
$W \to e\nu$+jets:
mc16_13TeV.364170.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364171.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364172.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364173.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364174.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364175.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364176.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364177.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364178.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364179.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364180.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364181.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364182.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV500_1000.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364183.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV1000_E_CMS.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
$W \to \mu\nu$+jets:
mc16_13TeV.364156.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364157.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364158.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364159.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364160.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364161.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364162.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364163.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364164.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364165.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364166.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364167.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364168.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV500_1000.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364169.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV1000_E_CMS.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
$W \to \tau\nu$+jets:
mc16_13TeV.364184.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364185.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364186.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364187.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364188.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364189.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364190.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004

mc16_13TeV.364191.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364192.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364193.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364194.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364195.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_BFilter.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364196.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV500_1000.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004
mc16_13TeV.364197.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV1000_E_CMS.deriv.DAOD_EXOT8.e5340_s3126_rXXXX_p4004

$W \to \ell\nu$+jets t-channel:
mc16_13TeV.308096.Sherpa_221_NNPDF30NNLO_Wenu2jets_Min_N_TChannel.deriv.DAOD_EXOT8.e5789_s3126_rXXXX_p4004
mc16_13TeV.308097.Sherpa_221_NNPDF30NNLO_Wmunu2jets_Min_N_TChannel.deriv.DAOD_EXOT8.e5767_s3126_rXXXX_p4004
mc16_13TeV.308098.Sherpa_221_NNPDF30NNLO_Wtaunu2jets_Min_N_TChannel.deriv.DAOD_EXOT8.e5767_s3126_rXXXX_p4004

$Z \to ee$+jets:
mc16_13TeV.364114.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364115.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364116.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_BFilter.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364117.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364118.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364119.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_BFilter.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364120.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364121.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364122.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_BFilter.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364123.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364124.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364125.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_BFilter.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364126.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004
mc16_13TeV.364127.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS.deriv.DAOD_EXOT8.e5299_s3126_rXXXX_p4004

$Z \to \mu\mu$+jets:
mc16_13TeV.364100.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364101.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364102.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_BFilter.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364103.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364104.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364105.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_BFilter.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364106.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364107.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364108.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_BFilter.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364109.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364110.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364111.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_BFilter.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364112.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV500_1000.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004
mc16_13TeV.364113.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV1000_E_CMS.deriv.DAOD_EXOT8.e5271_s3126_rXXXX_p4004

$Z \to \tau\tau$+jets:
mc16_13TeV.364128.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364129.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364130.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_BFilter.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364131.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364132.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364133.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_BFilter.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364134.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364135.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364136.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_BFilter.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364137.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364138.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_EXOT8.e5313_s3126_rXXXX_p4004
mc16_13TeV.364139.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_BFilter.deriv.DAOD_EXOT8.e5313_s3126_rXXXX_p4004
mc16_13TeV.364140.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004
mc16_13TeV.364141.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS.deriv.DAOD_EXOT8.e5307_s3126_rXXXX_p4004

$Z \to \ell\ell$+jets t-channel:
mc16_13TeV.308092.Sherpa_221_NNPDF30NNLO_Zee2jets_Min_N_TChannel.deriv.DAOD_EXOT8.e5767_s3126_rXXXX_p4004
mc16_13TeV.308093.Sherpa_221_NNPDF30NNLO_Zmm2jets_Min_N_TChannel.deriv.DAOD_EXOT8.e5767_s3126_rXXXX_p4004
mc16_13TeV.308094.Sherpa_221_NNPDF30NNLO_Ztautau2jets_Min_N_TChannel.deriv.DAOD_EXOT8.e5767_s3126_rXXXX_p4004

$W + \gamma$:
mc16_13TeV.364413.MadGraphPythia8EvtGen_WqqgammaNp0123_DP140_280.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004
mc16_13TeV.364414.MadGraphPythia8EvtGen_WqqgammaNp0123_DP280_500.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004
mc16_13TeV.364415.MadGraphPythia8EvtGen_WqqgammaNp0123_DP500_1000.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004
mc16_13TeV.364416.MadGraphPythia8EvtGen_WqqgammaNp0123_DP1000_2000.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004
mc16_13TeV.364417.MadGraphPythia8EvtGen_WqqgammaNp0123_DP2000_inf.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004

$Z + \gamma$:
mc16_13TeV.364418.MadGraphPythia8EvtGen_ZqqgammaNp0123_DP140_280.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004
mc16_13TeV.364419.MadGraphPythia8EvtGen_ZqqgammaNp0123_DP280_500.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004
mc16_13TeV.364420.MadGraphPythia8EvtGen_ZqqgammaNp0123_DP500_1000.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004
mc16_13TeV.364421.MadGraphPythia8EvtGen_ZqqgammaNp0123_DP1000_2000.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004
mc16_13TeV.364422.MadGraphPythia8EvtGen_ZqqgammaNp0123_DP2000_inf.deriv.DAOD_EXOT8.e5969_s3126_rXXXX_p4004

Dijet:

mc16_13TeV.364700.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ0WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364701.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ1WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364702.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ2WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364703.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ3WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364704.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ4WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364705.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ5WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364706.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ6WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364707.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ7WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364708.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ8WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364709.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ9WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364710.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ10WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364711.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ11WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

mc16_13TeV.364712.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ12WithSW.deriv.DAOD_EXOT8.e7142_s3126_rXXXX_p4004

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22**, 579 (1961)

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967)

[3] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519**, 367 (1968)

[4] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8**, 214 (1964)

[5] H. Fritzsch, M. Gell-Mann, H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. B **47**, 365 (1973)

[6] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12**, 132 (1964)

[7] F. Englert, R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13**, 321 (1964)

[8] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13**, 508 (1964)

[9] G. S. Guralnik, C. R. Hagen, T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13**, 585 (1964)

[10] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716**, 1 (2012)

[11] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716**, 30 (2012)

[12] J. F. Navarro, C. S. Frenk, S. D. M. White, *The Structure of cold dark matter halos*, Astrophys. J. **462**, 563 (1996)

[13] G. Bertone, D. Hooper, J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys. Rept. **405**, 279 (2005)

[14] LUX Collaboration, *Results from a search for dark matter in the complete LUX exposure*, Phys. Rev. Lett. **118(2)**, 021303 (2017)

[15] XENON Collaboration, *Dark Matter Search Results from a One Ton-Year Exposure of XENON1T*, Phys. Rev. Lett. **121(11)**, 111302 (2018)

[16] P. J. E. Peebles, B. Ratra, *The Cosmological Constant and Dark Energy*, Rev. Mod. Phys. **75**, 559 (2003)

[17] Supernova Search Team, *Type Ia supernova discoveries at $z > 1$ from the Hubble Space Telescope: Evidence for past deceleration and constraints on dark energy evolution*, Astrophys. J. **607**, 665 (2004)

[18] E. J. Copeland, M. Sami, S. Tsujikawa, *Dynamics of dark energy*, Int. J. Mod. Phys. D **15**, 1753 (2006)

[19] Super-Kamiokande Collaboration, *A Measurement of atmospheric neutrino oscillation parameters by SUPER-KAMIOKANDE I*, Phys. Rev. D **71**, 112005 (2005)

[20] K2K Collaboration (K2K), *Measurement of Neutrino Oscillation by the K2K Experiment*, Phys. Rev. D **74**, 072003 (2006)

[21] KamLAND Collaboration, *Precision Measurement of Neutrino Oscillation Parameters with KamLAND*, Phys. Rev. Lett. **100**, 221803 (2008)

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

[23] M. Dine, A. Kusenko, *The Origin of the matter - antimatter asymmetry*, Rev. Mod. Phys. **76**, 1 (2003)

[24] R. Costa, M. Mühlleitner, M. O. P. Sampaio, R. Santos, *Singlet Extensions of the Standard Model at LHC Run 2: Benchmarks and Comparison with the NMSSM*, JHEP **06**, 034 (2016)

[25] T. Robens, T. Stefaniak, J. Wittbrodt, *Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios*, Eur. Phys. J. C **80(2)**, 151 (2020)

[26] G. C. Branco, et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516**, 1 (2012)

[27] M. Thomson, *Modern particle physics*, Cambridge University Press, New York (2013)

[28] M. E. Peskin, D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA (1995)

[29] Particle Data Group, *Review of Particle Physics*, PTEP **2020(8)**, 083C01 (2020)

[30] J. Goldstone, *Field Theories with Superconductor Solutions*, Nuovo Cim. **19**, 154 (1961)

[31] J. Goldstone, A. Salam, S. Weinberg, *Broken Symmetries*, Phys. Rev. **127**, 965 (1962)

[32] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model*, Phys. Rept. **459**, 1 (2008)

[33] S. L. Glashow, S. Weinberg, *Natural Conservation Laws for Neutral Currents*, Phys. Rev. D **15**, 1958 (1977)

[34] R. D. Peccei, H. R. Quinn, *CP Conservation in the Presence of Instantons*, Phys. Rev. Lett. **38**, 1440 (1977)

[35] N. Turok, J. Zadrozny, *Electroweak baryogenesis in the two doublet model*, Nucl. Phys. B **358**, 471 (1991)

[36] M. Joyce, T. Prokopec, N. Turok, *Nonlocal electroweak baryogenesis. Part 2: The Classical regime*, Phys. Rev. D **53**, 2958 (1996)

[37] A. T. Davies, C. D. froggatt, G. Jenkins, R. G. Moorhouse, *Baryogenesis constraints on two Higgs doublet models*, Phys. Lett. B **336**, 464 (1994)

[38] J. M. Cline, K. Kainulainen, A. P. Vischer, *Dynamics of two Higgs doublet CP violation and baryogenesis at the electroweak phase transition*, Phys. Rev. D **54**, 2451 (1996)

[39] J. M. Cline, P.-A. Lemieux, *Electroweak phase transition in two Higgs doublet models*, Phys. Rev. D **55**, 3873 (1997)

[40] M. Laine, K. Rummukainen, *Two Higgs doublet dynamics at the electroweak phase transition: A Nonperturbative study*, Nucl. Phys. B **597**, 23 (2001)

[41] L. Fromme, S. J. Huber, M. Seniuch, *Baryogenesis in the two-Higgs doublet model*, JHEP **11**, 038 (2006)

[42] H. E. Haber, G. L. Kane, *The Search for Supersymmetry: Probing Physics Beyond the Standard Model*, Phys. Rept. **117**, 75 (1985)

[43] S. P. Martin, *A Supersymmetry primer*, Adv. Ser. Direct. High Energy Phys. **18**, 1 (1998)

[44] R. K. Kaul, P. Majumdar, *Cancellation of Quadratically Divergent Mass Corrections in Globally Supersymmetric Spontaneously Broken Gauge Theories*, Nucl. Phys. B **199**, 36 (1982)

[45] M. Cepeda, et al., *Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC*, CERN Yellow Rep. Monogr. **7**, 221 (2019)

[46] M. Grazzini, et al., *Higgs boson pair production at NNLO with top quark mass effects*, JHEP **05**, 059 (2018)

[47] J. Baglio, et al., $gg \to HH$ *: Combined uncertainties*, Phys. Rev. D **103(5)**, 056002 (2021)

[48] CMS Collaboration, *Combination of searches for Higgs boson pair production in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. Lett. **122(12)**, 121803 (2019)

[49] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Lett. B **800**, 135103 (2020)

[50] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector* **2/2017** (2016), arXiv:1610.07922 [hep-ph]

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

[52] G. Cowan, *Statistics for Searches at the LHC*, in *69th Scottish Universities Summer School in Physics: LHC Physics*, pages 321–355 (2013)

[53] J. Neyman, E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, Phil. Trans. Roy. Soc. Lond. A **231(694-706)**, 289 (1933)

[54] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. G **28**, 2693 (2002)

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

[56] O. S. Brüning, et al., *LHC Design Report Vol.1: The LHC Main Ring* (2004)

[57] O. S. Brüning, et al., *LHC Design Report. 2. The LHC infrastructure and general services* (2004)

[58] M. Benedikt, P. Collier, V. Mertens, J. Poole, K. Schindl, *LHC Design Report. 3. The LHC injector chain* (2004)

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

[60] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3**, S08004 (2008)

[61] LHCb Collaboration (LHCb), *The LHCb Detector at the LHC*, JINST **3**, S08005 (2008)

[62] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3**, S08002 (2008)

[63] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15(04)**, P04003 (2020)

[64] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-19 (2010)

[65] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, CERN-LHCC-2013-018, ATLAS-TDR-023 (2013)

[66] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70**, 823 (2010)

[67] S. Höche, *Introduction to parton-shower event generators*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pages 235–295 (2015)

[68] Sherpa Collaboration, *Event Generation with Sherpa 2.2*, SciPost Phys. **7(3)**, 034 (2019)

[69] J. Alwall, et al., *Computing decay rates for new physics theories with FeynRules and MadGraph 5_aMC@NLO*, Comput. Phys. Commun. **197**, 312 (2015)

[70] T. Sjöstrand, et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**, 159 (2015)

[71] J. Bellm, et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76(4)**, 196 (2016)

[72] GEANT4 Collaboration, *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506**, 250 (2003)

[73] ATLAS Collaboration, *Performance of the Fast ATLAS Tracking Simulation (FATRAS) and the ATLAS Fast Calorimeter Simulation (FastCaloSim) with single particles*, Technical report (2014)

[74] S. Catani, Y. L. Dokshitzer, M. H. Seymour, B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. B **406**, 187 (1993)

[75] Y. L. Dokshitzer, G. D. Leder, S. Moretti, B. R. Webber, *Better jet clustering algorithms*, JHEP **08**, 001 (1997)

[76] M. Cacciari, G. P. Salam, G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04**, 063 (2008)

[77] J. M. Butterworth, A. R. Davison, M. Rubin, G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100**, 242001 (2008)

[78] D. Krohn, J. Thaler, L.-T. Wang, *Jet Trimming*, JHEP **02**, 084 (2010)

[79] J. Thaler, K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP **03**, 015 (2011)

[80] A. J. Larkoski, G. P. Salam, J. Thaler, *Energy Correlation Functions for Jet Substructure*, JHEP **06**, 108 (2013)

[81] A. J. Larkoski, I. Moult, D. Neill, *Power Counting to Better Jet Observables*, JHEP **12**, 009 (2014)

[82] A. J. Larkoski, I. Moult, D. Neill, *Analytic Boosted Boson Discrimination*, JHEP **05**, 117 (2016)

[83] M. Cacciari, G. P. Salam, G. Soyez, *The Catchment Area of Jets*, JHEP **04**, 005 (2008)

[84] ATLAS Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, ATLAS-CONF-2016-035 (2016)

[85] ATLAS Collaboration, *Track assisted techniques for jet substructure*, ATL-PHYS-PUB-2018-012 (2018)

[86] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77(7)**, 466 (2017)

[87] ATLAS Collaboration, *Tagging and suppression of pileup jets*, ATLAS-CONF-2014-018 (2014)

[88] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77**, 490 (2017)

[89] D. Krohn, J. Thaler, L.-T. Wang, *Jets with Variable R*, JHEP **06**, 059 (2009)

[90] ATLAS Collaboration, *Variable Radius, Exclusive-k_T, and Center-of-Mass Subjet Reconstruction for Higgs($\rightarrow b\bar{b}$) Tagging in ATLAS*, ATL-PHYS-PUB-2017-010 (2017)

[91] 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*, Eur. Phys. J. C **79(8)**, 639 (2019)

[92] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. C **76(5)**, 292 (2016)

[93] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **78(11)**, 903 (2018)

[94] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10**, 531 (1963)

[95] M. Kobayashi, T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49**, 652 (1973)

[96] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012 (2016)

[97] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013 (2017)

[98] ATLAS Collaboraion, *Search for resonant boosted HH and SH production in the $bbVV$ decay channel with 0 or 1 leptons in the final state using the full Run 2 ATLAS data*, ATL-COM-PHYS-2021-761 (2021)

[99] M. Baak, et al., *HistFitter software framework for statistical data analysis*, Eur. Phys. J. C **75**, 153 (2015)

[100] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, W. Verkerke (ROOT), *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, CERN-OPEN-2012-016 (2012)

[101] L. Moneta, et al., *The RooStats Project*, PoS **ACAT2010**, 057 (2010)

[102] W. Verkerke, D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241**, MOLT007 (2003)

[103] R. Brun, F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389**, 81 (1997)

[104] I. Antcheva, et al., *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, Comput. Phys. Commun. **182**, 1384 (2011)

[105] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, CERN-D-506 (1994)

[106] F. James, M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10**, 343 (1975)

[107] ATLAS Collaboration, *Search for resonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, ATLAS-CONF-2021-035 (2021)

[108] ATLAS Collaboration, *Supporting Document: The Search for Resonant HH Production Decaying to the 4b Final State Using the Full Run-2 Data and the Boosted Analysis Channel*, ATL-COM-PHYS-2020-083 (2020)

Acknowledgements

Research is a collective effort and this thesis is no exception. Hence, I would like to extend my gratitude to the most respected Prof. Dr. Stan Lai for providing me the opportunity to work on this topic and for his wise guidance at every step of this thesis. He also deserves the credits for broadening my horizons, not just of physics, but also of wine and culinary skills. I am also grateful to Kira (even though I can never thank her enough) for her never-ending support, and for always being so patient with my constant questions and my unacceptable working hours. It goes unsaid, that none of this work would have been "fruitful" without her support. Next, I would also like to thank Jason for taking out the time from his busy schedule to provide his insightful comments and feedback on the analysis, and also on this document. I am also thankful to Joshua for agreeing to take this work forward and would also like to apologise to him for having to deal with my immature code and scripts.

In addition to this, I am also thankful to all the other group and institute members for making this past year so enjoyable. I am glad that I could meet many of them on multiple occasions, despite the on-going pandemic. A very special thanks also goes to my landlady/housemate for pampering me with such delicious food, especially while I was engrossed in writing this thesis. I would also like to thank my girlfriend, Monisha, for helping me with all the graphics and design related issues during my thesis. And of course, I cannot end this acknowledgements without thanking all my friends and my family for their constant support at all times. To my parents, I remain forever indebted for all that I have, and everything that I will ever achieve.

Erklärung nach §17(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen:
Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.
Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 12. Januar 2022

(Naman Kumar Bhalla)