GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Master's Thesis

Konzeptionalisierung einer geboosteten $HH \to b\bar{b}WW^*$ Analyse im 1-Lepton Endzustand Run 3 des ATLAS Experiments

Conceptualization of a Boosted $HH \to b\bar{b}WW^*$ Analysis in the 1-lepton final state for ATLAS Run 3

prepared by

Lina Buschmann

from Winsen (Luhe)

at the II. Physikalischen Institut

Thesis number:	II.Physik-UniGö-MSc-2024/11
Thesis period:	11th June 2024 until 11th December 2024
First referee:	Prof. Dr. Stan Lai
Second referee:	PD Dr. Benjamin Schwenker

Zusammenfassung

Das Standardmodell der Teilchenphysik ist das erfolgreichste physikalische Modell. Dennoch ist bekannt, dass es unvollständig ist. Mit der Entdeckung des Higgs-Bosons im Jahr 2012 wurde das Standardmodell vervollständigt. Dies führte zu zahlreichen Studien, die seine Eigenschaften mit den Vorhersagen des Standardmodells vergleichen. Da das Higgs-Boson an Masse koppelt, bietet es Möglichkeiten, nach Physik jenseits des Standardmodells zu suchen.
Diese Arbeit konzentriert sich auf die resonante Produktion von Higgs-Boson-Paaren im $b\bar{b}WW^*$-Zerfallskanal mit einem geladenen Lepton im Endzustand unter Verwendung simulierter ATLAS Run 3-Daten. Die hohe Masse der skalaren Resonanz führt zu einer „boosted“ Topologie, bei der die Zerfallsprodukte des Higgs-Bosons, das in ein $b\bar{b}$-Quarkpaar zerfällt, sowie die Zerfallsprodukte des hadronisch zerfallenden W-Bosons nicht einzeln aufgelöst werden können. Außerdem überlappt das geladene Lepton mit einem der large-R Jets, die aus diesen hadronischen Objekten rekonstruiert wurden. Dies macht die Topologie einzigartig, aber auch schwierig in der Rekonstruktion.
Das Ziel dieser Arbeit ist es, eine neue Analyse für diesen Zerfallskanal zu konzipieren, wobei zunächst eine Vorselektion etabliert werden soll, um anschließend die Ergebnisse der auf Schnitten basierenden Selektion mit einer Klassifikation auf Basis eines Feedforward Neuronalen Netzes zu vergleichen.

Stichwörter: Teilchenphysik, Masterarbeit, Higgs Boson Paarproduktion, Neuronale Netze

Abstract

The Standard Model of Particle Physics is the most successful physics model in describing fundamental physics phenomena. Still, it is known to be incomplete. With the discovery of the Higgs boson in 2012, all Standard Model particles have been observed. This led to many studies testing its properties against the Standard Model prediction. Since the Higgs boson couples to mass, it can provide opportunities to search for physics beyond the Standard Model.
This thesis focuses on resonant Higgs boson pair production in the $b\bar{b}WW^*$ decay channel with one charged lepton in the final state using simulated ATLAS Run 3 data. The large mass of the scalar resonance leads to the boosted topology, in which the decay products of the Higgs boson decaying to a $b\bar{b}$ quark pair as well as the decay products of the hadronically decaying W boson cannot be resolved. The charged lepton also overlaps with one of the large-R jets reconstructed from these hadronic objects. This makes the topology unique but challenging to reconstruct.
The aim of this thesis is to conceptualise a new analysis for this decay channel, focusing on first establishing a preselection to then compare the performance of the cut-based selection with a feed-forward neural network based classification.

Keywords: Particle Physics, Master thesis, Higgs boson pair production, Neural Networks

Contents

1. Introduction

The Standard Model of particle physics (SM) [1–4] is a highly precise and successful model. It is widely accepted and intensively tested. Since the discovery of the Higgs boson in 2012 [5, 6], many studies have been performed to test the Higgs boson's properties against Standard Model predictions. In particular, a direct measurement of the Higgs self-coupling is important to characterise the Higgs potential. This coupling strength could be measured in Higgs boson pair production, motivating the search for such processes, although, the SM predicted cross section $\sigma_{HH}^{SM} = 31.05\,\text{fb}$ [7–9] is very small, due to destructive interference.

Furthermore, it is known that there must be physics beyond the SM (BSM) as the SM fails to account for the existence of dark matter and other observed phenomena [10]. Many BSM theories predict additional, highly massive narrow-width particles, X, which can decay to two Higgs bosons, H, enhancing the Higgs boson pair production cross section. Since the mass of X is not predicted by theory in this approach, a wide range of masses, m_X, is considered, resulting in various decay topologies.

This Master's thesis focuses on the boosted topology of the BSM decay channel $X \to HH \to b\bar{b}WW^*$ in the 1-lepton final state. In this context, "boosted" means that the two Higgs bosons have large kinematic energies of several hundred GeV, resulting in their respective decay products being highly collimated. The aim of this thesis is to develop an analysis for this decay for Run 3 data recorded by the ATLAS detector from 2022 until 2025 with an expected integrated luminosity of $\int \mathcal{L}dt = 150\,\text{fb}^{-1}$. A few years ago, Peer Drescher showed in his Bachelor's thesis [11] for the Run 2 analysis that using a model based on a neural network for the separation of signal and background outperforms a cut-based approach. Given these results, a similar model should be developed for Run 3.

In chapter 2 of this thesis, the theoretical foundations are first discussed, where an overview of the Standard model of particle physics, as well as a brief explanation of the Higgs mechanism with a BSM extension, is given. Afterwards, in chapter 3, an introduction to Machine Learning and neural networks is given. Then, in chapter 4, the experimental setup of the LHC and the ATLAS detector with all its subdetectors is de-

scribed. In chapter 5, the studied decay channel, the analysis objects and the data samples used for this analysis are explained in detail. In chapter 6, the baseline signal selection for the analysis as well as the neural network model are presented. Lastly, in chapter 7, the results of this Master's thesis are summarized and an outlook is provided.

2. Theoretical Background

In this chapter, the theoretical background of this Master's thesis is covered. First, the Standard Model of Particle Physics (SM) [1–4] is described in detail, including three of the four known fundamental forces. The Higgs boson was the last discovered particle from the SM [5, 6]. The Higgs mechanism [12–17], which gives mass to particles, is explained in detail as well. After that, Higgs boson pair production is discussed.

2.1. The Standard Model of Particle Physics

There exist four fundamental forces in the universe: the gravitational, the electromagnetic, the weak and the strong force. Einstein's General Theory of Relativity [18] explains the gravitational force, while the other three interactions have been unified into quantum field theory known as the SM. It also describes all known elementary particles as excitations of quantum fields.

2.1.1. Particles and Forces

The elementary particles in the SM can be split into two categories: fermions with spin-$\frac{1}{2}$ and bosons with integer spin. Fermions are the elementary particles that make up matter and bosons are the mediators of the fundamental forces and the Higgs mechanism.

There are twelve fermions and five bosons in the SM. In Figure 2.1, all elementary particles of the SM and some of their properties are listed. The twelve fermions can be further split into quarks and leptons, each consisting of six particles in three generations. Each generation includes two quarks and two leptons in weak isospin doublets with differing masses. There are electrically charged leptons with an electrical charge of -1 and also neutral leptons called neutrinos. The quarks can be further divided into up-type quarks with an electrical charge of $+\frac{2}{3}$ and down-type quarks with an electrical charge of $-\frac{1}{3}$. Quarks also carry the charge of the strong interaction called colour. The three distinct values of the colour charge are red, blue and green. While the charges stay the same over the three generations of fermions, their masses and flavours change. To every fermion there exists an anti-particle with exactly opposite charges.

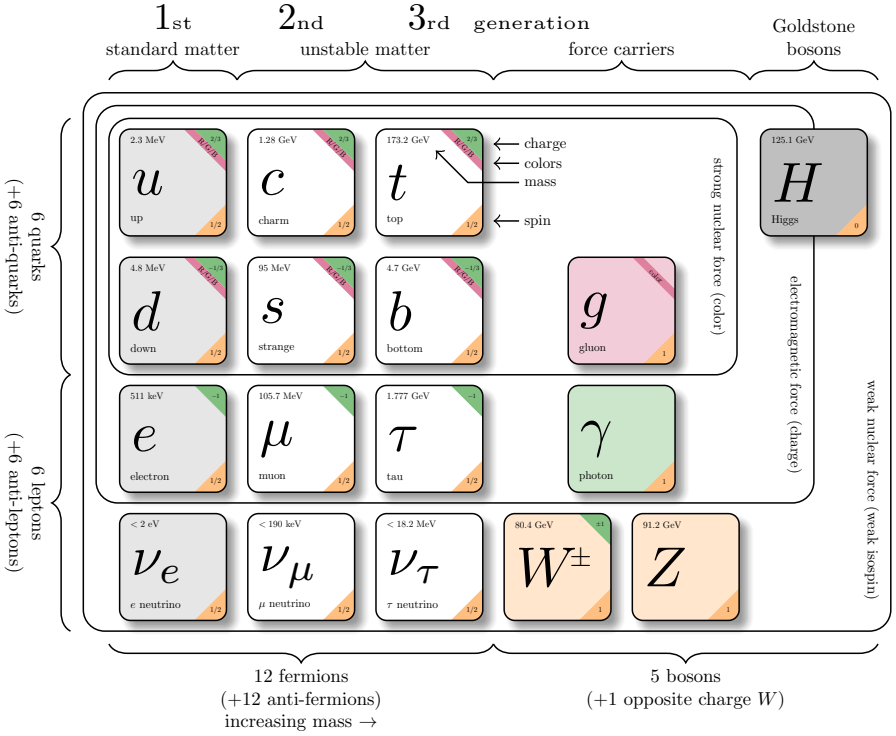


Figure 2.1.: The elementary particles of the SM. In the upper half the quarks and in the lower half the leptons are shown. Both are grouped into the three generations with particles from higher generations increasing in mass and decreasing in lifetime. The gauge bosons that mediate the three fundamental forces are to the right of the fermions. The Higgs boson is also depicted in grey. Their colours correspond to the mediated force, indicated by coloured corners for each particle. Also the electric charge, the spin and the particle masses are given. Taken from Ref. [1].

The fermions in the first generation are stable, while particles from the second and third generation have increasingly higher masses and decay into fermions of lower masses. The three lepton flavours are called electron (e), muon (μ) and tau (τ), each having a corresponding neutrino of their respective flavour. Quarks have six different flavours. The up-type quarks are called up (u), charm (c) and top (t) while the down-type quarks are called down (d), strange (s) and bottom (b).

All bosons but the Higgs boson are spin 1 gauge bosons. These are called gluon (g), photon (γ), W^- boson, W^+ boson and Z boson. The gluon mediates the strong force, the photon the electromagnetic force, and the two W bosons as well as the Z boson mediate the weak force. The last boson, the Higgs boson, is a spin 0 Goldstone boson. Goldstone bosons stem from spontaneous symmetry breaking of massless particles (see Section 2.1.2). The interaction with the Higgs field generates masses to the SM particles.

Quantum Electrodynamics

The quantum field theory describing the interaction between electrically charged particles via photon exchange is Quantum Electrodynamics (QED) [19]. The electromagnetic Lagrangian for a massive fermion ψ and a massless spin-1 photon field A^μ is

$$\mathscr{L}_{QED} = \overline{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi + g_e\overline{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{2.1}$$

where $\overline{\psi} = \psi^\dagger\gamma^0$ is the adjoint spinor, γ^μ are the gamma matrices, m is the particles mass and

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A\mu, \tag{2.2}$$

denotes the electromagnetic field strength tensor. The second term of the electromagnetic Lagrangian introduces interactions between the photon and the fermion with a coupling of g_e.

Furthermore, the electromagnetic Lagrangian is invariant under local $U(1)$ transformations with a phase α:

$$\psi \to \psi^{'} = e^{i\alpha}\psi; \quad A_\mu \to A_\mu^{'} = A\mu + \frac{1}{g_e}\partial_\mu\alpha. \tag{2.3}$$

Quantum Chromodynamics and Electroweak Unification

The quantum field theory describing the interaction between particles carrying colour charge via gluon exchange is Quantum Chromodynamics (QCD)[20, 21]. In analogy to QED, the Lagrangian can be written as

$$\mathscr{L}_{QCD} = \overline{q}_i\left(i\gamma^\mu\left(D_\mu\right)_{ij}\right)q_j - \frac{1}{4}G^{\mu\nu}G_{\mu\nu}, \tag{2.4}$$

with $G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc}G^b_\mu G_\nu$ and

$$\left(D_\mu\right)_{ij} = \partial_\mu\delta_{ij} - ig_s\left(T_a\right)_{ij}G^a_\mu, \tag{2.5}$$

representing the kinematic terms of the eight gluon fields G^a_μ. Here g_s denotes the coupling strength of the strong interaction and the indices i and j correspond to the colour state of the quarks q. QCD has an underlying $SU(3)$ symmetry with respect to the colour charge, meaning it is invariant under local $SU(3)_C$ gauge transformations.

QCD has some special properties. As the gluon itself also carries a colour charge, it can interact with itself resulting in the coupling constant becoming larger for smaller energies [22, 23]. This means that increasing the distance between quarks leads to the creation of quark-antiquark pairs because the energy in the coupling increases. This process is called

hadronisation. Another interesting phenomenon of QCD is the fact that a free quark has never been observed, this phenomenon is called colour confinement [24]. This means that hadronised quarks only exist in colour neutral states, either by three differently coloured quarks forming a bound state called baryon or by a quark anti-quark pair with the same colour forming a bound state called meson.

Finally, there is the weak interaction. It is transmitted by the either positively or negatively charged W boson or the neutral Z boson. It is the only force that neutrinos interact with, as they neither have an electric charge nor a colour charge.
The Glashow-Salam-Weinberg (GSW) Theory [2–4] allows the simultaneous description of electromagnetic and weak interaction through introduction of the hypercharge

$$Y = 2\left(Q - I_w^{(3)}\right), \tag{2.6}$$

where Q denotes the electric charge and $I_w^{(3)}$ the third component of weak isospin. This leads to the description of the photon field A_μ and the Z boson field Z_μ via a new boson field B_μ:

$$A_\mu = B_\mu \cos\theta_W + W_\mu^{(3)} \sin\theta_w \tag{2.7}$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^{(3)} \cos\theta_w, \tag{2.8}$$

with θ_w being the weak mixing angle. It is apparent that the photon field A_μ and the Z boson field Z_μ are superpositions of $W_\mu^{(3)}$ and B_μ. In a similar way, the $W^\pm$ bosons are superpositions of the $W_\mu^{(1)}$ and $W_\mu^{(2)}$ fields:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^{(1)} \mp i W_\mu^{(2)}\right). \tag{2.9}$$

2.1.2. The Brout-Englert-Higgs Mechanism

The combined SM Lagrangian

$$\mathscr{L}_{\mathrm{SM}} = \underbrace{-\frac{1}{4}G^a_{\mu\nu}G_a^{\mu\nu} - \frac{1}{4}W^a_{\mu\nu}W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}}_{\text{kinematics of gauge fields}} + \underbrace{\overline{\psi}^f_L i\gamma^\mu D_\mu \psi^f_L + \overline{\psi}^f_R i\gamma^\mu D_\mu \psi^f_R}_{\text{kinematics and interactions of fermions}} \tag{2.10}$$

with

$$D_\mu\psi = (\partial_\mu - \underbrace{ig_s T_a G^a_\mu}_{SU(3)_C} - \underbrace{ig_2 T_a W^a_\mu}_{SU(2)_L} - \underbrace{ig_1 T B_\mu}_{U(1)_Y})\psi, \tag{2.11}$$

contains all previously discussed quantum field theories and is invariant under $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry. For both gauge bosons and fermions, the SM Lagrangian does not contain mass terms to preserve this invariance. Thus neither gauge bosons nor fermions should have masses if the Lagrangian was complete. But particle masses are clearly observed for various elemental particles.
The mass terms can be achieved through the Higgs mechanism, developed by Brout, Englert, Guralnik, Hagen, Higgs and Kibble [12–17], which introduces a doublet of complex scalar fields Φ. The Langrangian of Φ is given by

$$\mathscr{L}_{\text{Higgs}} = \left(D^{\mu}\Phi\right)^{\dagger}\left(D_{\mu}\Phi\right) - V\left(\Phi\right), \tag{2.12}$$

and the Higgs potential is

$$V\left(\Phi\right) = \mu^2\Phi^{\dagger}\Phi + \lambda\left(\Phi^{\dagger}\Phi\right)^2. \tag{2.13}$$

In Figure 2.2 the Higgs potential is sketched for $\lambda > 0$ and $\mu^2 < 0$. Under such conditions, the Higgs potential has an infinite number of minima around the origin at $\Phi^{\dagger}\Phi = v^2 = -\mu^2/\lambda$, where v is the vacuum expectation value. Due to the rotational symmetry of the Higgs potential, when expanding the field around the minimum, the ground state can arbitrarily be chosen to be real:

$$\langle\Phi\rangle = \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.14}$$

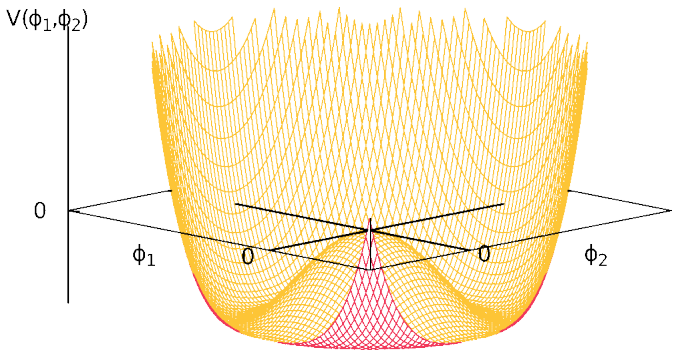


Figure 2.2.: Illustration of the Higgs potential for $\lambda > 0$ and $\mu^2 < 0$.

The field expansion

$$\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} \theta_1(x) + i\theta_2(x) \\ v + H(x) + i\theta_3(x) \end{pmatrix} \tag{2.15}$$

results in three massless Goldstone bosons $\theta_i(x)$ and one massive Higgs boson $H(x)$.

Applying unitary gauge, the Goldstone bosons are absorbed by the longitudinal degrees of freedom of the W and Z bosons and vanish. The field takes the form

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.16}$$

and the Lagrangian can be written as

$$\mathscr{L}_{\text{Higgs}} = \frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4}\lambda H \tag{2.17}$$

$$+ \frac{1}{8}\left(v + H^2\right)\begin{pmatrix} W_\mu^1 & W_\mu^2 & W_\mu^3 & B_\mu \end{pmatrix} \begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & gg' \\ 0 & 0 & gg' & g^2 \end{pmatrix} \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \\ W_\mu^3 \\ B_\mu \end{pmatrix}. \tag{2.18}$$

The terms in the first line of the Lagrangian describes a free massive scalar with cubic and quartic self-interactions, where the interaction strength is proportional to λ. The second line describes the interaction of the W_μ and B_μ fields with the scalar Higgs field. The second line of the Lagrangian also gives rise to a massless photon and massive $W^\pm$, Z and Higgs boson H with masses

$$m_H = \sqrt{2\lambda v^2},\ m_W = \frac{gv}{2},\ m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} \text{ and } m_A = 0. \tag{2.19}$$

From Equation (2.18), it can be seen that the couplings of the gauge bosons to the Higgs boson are proportional to the squared mass of the gauge bosons. Thus, the more massive the gauge boson, the greater the coupling to the Higgs boson.
The Higgs field also generates masses for fermions via the Yukawa coupling. Both the Higgs field and the fermions form a SU(2) doublet. In unitary gauge, the Higgs doublet projects onto the $T_3 = -1/2$ part of the weak isospin doublet of the fermion. Thus, the fermion masses are generated separately for down-type quarks together with charged leptons and for up-type quarks together with neutrinos, according to their weak isospin properties. For down-type quarks as well as charged leptons, the Lagrangian density includes terms such as

$$\mathscr{L}_d = -\frac{g_f v}{\sqrt{2}}\left(\overline{d}_L d_R + \overline{d}_R d_L\right) - \frac{g_f}{\sqrt{2}} H \left(\overline{d}_L d_R + \overline{d}_R d_L\right). \tag{2.20}$$

The fermion mass arises in the first term as

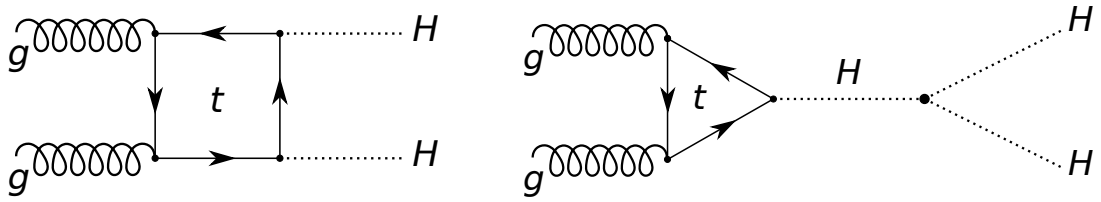
$$m_f = g_f \frac{v}{\sqrt{2}}, \tag{2.21}$$

where g_f is the Yukawa coupling constant [4, 25, 26]. The interaction of the fermion with the Higgs field is described by the second term. For up-type quarks and neutrinos the Lagrangian density includes

$$\mathscr{L}_u = -\frac{g_f v}{\sqrt{2}} \left(\overline{u}_L u_R + \overline{u}_R u_L\right) - \frac{g_f}{\sqrt{2}} H \left(\overline{u}_L u_R + \overline{u}_R u_L\right), \tag{2.22}$$

giving the same result as for down-type quarks and charged leptons.

2.2. Higgs Boson Pair Production



(a) box diagram **(b)** cubic self-coupling

Figure 2.3.: Dominant Feynman diagrams for SM Higgs pair production.

In pp collisions at a centre-of-mass energy of $\sqrt{s} = 13.6\,\text{TeV}$, SM Higgs boson pair production without any additional particles occurs predominantly via two production mechanisms at hadron colliders [27], both initiated by gluon-fusion processes. The Feynman diagrams for both processes are shown in Figure 2.3. For both processes, the Higgs boson pair production is non-resonant and the mediating Higgs boson can have an off-shell mass. The two processes interfere destructively, reducing the overall cross section. The cross section is [7–9]

$$\sigma_{gg\to HH} = 31.05^{+2.2\%}_{-5.0\%}\,\text{fb}. \tag{2.23}$$

This cross section is three orders of magnitude smaller than the single Higgs gluon-gluon fusion cross section, making a discovery at LHC challenging.
While SM Higgs boson pair production is solely possible in a non-resonant mode, beyond the Standard Model (BSM) theories allow the resonant production of Higgs boson pairs by introducing a heavy particle X that can decay into Higgs bosons. A possible BSM theory to describe this extra scalar particle is briefly discussed.

The two-Higgs-doublet model [28] introduces a second complex scalar doublet Φ_2. The

potential for two complex scalar doublets is

$$\begin{aligned} V = & m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - m_{12}^2 \left(\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1 \right) \\ & + \frac{\lambda_1}{2} \left(\Phi_1^\dagger \Phi_1 \right)^2 + \frac{\lambda_2}{2} \left(\Phi_2^\dagger \Phi_2 \right)^2 + \lambda_3 \Phi_1^\dagger \Phi_1 \Phi_2^\dagger \Phi_2 \\ & + \lambda_4 \Phi_1^\dagger \Phi_2 \Phi_2^\dagger \Phi_1 + \frac{\lambda_5}{2} [\left(\Phi_1^\dagger \Phi_2 \right)^2 + \left(\Phi_2^\dagger \Phi_1 \right)], \end{aligned} \tag{2.24}$$

which becomes minimal for

$$\langle \Phi_1 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_1 \end{pmatrix}, \langle \Phi_2 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_2 \end{pmatrix}. \tag{2.25}$$

Due to doubling the number of doublets with respect to the SM, the extension around the vacuum expectation values yields double the amount of fields, totalling in eight fields. In analogy to the SM case, three of those fields are absorbed when applying unitary gauge resulting in the masses of the $W^\pm$ and Z bosons. The remaining five fields are all physical scalar fields: two neutral scalar particles (light and heavy Higgs H, X), two charged scalar particles ($H^\pm$) and one neutral pseudoscalar particle (A). All these scalar particles couple to each other. For this thesis, the heavy scalar Higgs boson X and the light scalar Higgs boson H are of interest. The light scalar Higgs boson should have very similar properties to the SM Higgs boson to reflect the good agreement of the Higgs signal rate to the SM prediction.

	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 2.4.: Higgs pair branching ratios for $m_H = 125\,\text{GeV}$ [29].

The Higgs boson has a decay width of $\Gamma_H = 4.097\,\text{MeV}$ [29] and thus a lifetime of $\tau_H = 1.6 \times 10^{-22}\,\text{s}$. It decays too quickly to reach the detector, such that only its decay products are visible in the detector. Figure 2.4 lists the branching ratios of Higgs pairs. As the Higgs coupling strength is proportional to the mass of particles, it is more

likely for the Higgs boson to decay into heavy particles.
The decay to W^+W^- and ZZ boson pairs are kinematically suppressed though, as the invariant mass of such systems surpasses the Higgs boson mass. For these decay modes to occur, one gauge boson must be off-shell. This makes the decay of a Higgs boson pair into the $b\bar{b}b\bar{b}$ final state most likely with 33 %. But this decay channel is hard to separate from background making it difficult to identify. The decay channel with the second highest branching ratio is $b\bar{b}WW^*$, which is the channel of interest for this Master's thesis, with a branching ratio of 25 %. The W boson then has a 32.75 % chance of decaying leptonically. If it does not, it decays into a pair of quarks. This analysis focuses on the 1-lepton final state, meaning one W boson decays leptonically and the other hadronically. All other decay channels have significantly lower branching ratios.

3. Fundamentals of Neural Networks

In recent years, machine learning has become increasingly popular in the ATLAS, CMS, and high-energy physics communities. Researchers have used machine learning techniques such as neural networks to better separate signal events from background noise [30]. The performance of neural networks has improved due to new training methods and more efficient use of computing resources.

This chapter discusses the fundamentals of neural networks. Neural networks excel in recognising and learning patterns, making them applicable for a variety of different tasks, such as image or speech recognition and even complex scientific analyses. The pattern recognition ability makes neural networks invaluable to particle physics research as they have no problem with processing large amounts of data. In this thesis, neural networks are used to classify signal and background events.

First, a short introduction into the large field of machine learning is given. Afterwards, neural networks will be discussed in detail focusing on classification problems.

3.1. Introduction to Machine Learning

The term machine learning was first used by an IBM programmer, who was a pioneer in the field of artificial intelligence, in 1959 [31]. In machine learning, an algorithm is created and adapted for learning some kind of statistical model and make decisions based on what it learned. This is different from the ”classical” approach, where an algorithm is given explicit instructions to solve a specific problem. The machine learning algorithm solves a problem without explicit instructions.

This requires training of the algorithm. There are three different training approaches used for training depending on the task at hand. The first approach is supervised learning [32], the algorithm is presented with vast amounts of data which were already divided into different classes by a human. The algorithm is supposed to learn properties of the input

to differentiate between the classes.
The second approach is unsupervised learning [32], were the algorithm is presented with data without any labels applied. The algorithm is supposed to find its own structure in the presented input. This learning technique can uncover hidden patterns in data.
The third and final learning approach is the reinforcement learning [33]. It is a completely different approach in which the idea is to reward or punish the algorithm for the action it takes. Simple examples for reinforcement learning are self-driving cars or non-playable characters in video games.
Independently of the training approach, the more data is presented to the algorithm during training, the better the training results are. After the training, the algorithm is confronted with new data and is now supposed to apply what it learned.

In this thesis, a neural network is trained using the supervised training approach. The training process is explained in detail in Section 3.3.

3.2. Neural Networks

The origin of Neural networks dates back to the late 1950s. They are a class of machine learning algorithms that mimics the neural structure of human brain cells. A neural network made up of a single layer is called a perceptron [32, 34]. Figure 3.1 shows a schematic of a perceptron.

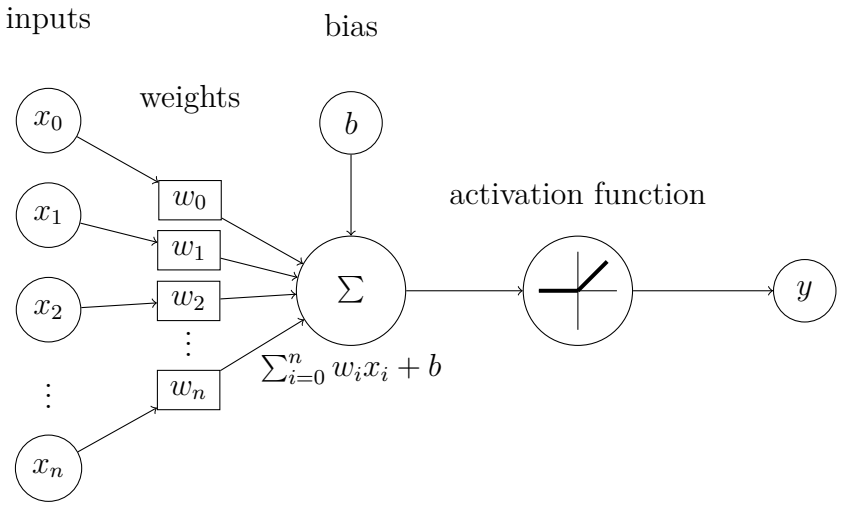


Figure 3.1.: Illustration of a perceptron. Information travels from the input layer towards the binary output y.

Perceptrons are the basis of neural networks. They are able to solve linearly separable problems. This works by summing over n weighted input variables $w \cdot x$ and adding a bias b. Applying the activation function f to the biased sum, which provides the possibility to describe non-linear relations between different variables, leads to the output

$$y = f\left(\sum_{i=1}^{n} w_i \cdot x_i + b\right). \tag{3.1}$$

Many perceptrons running in parallel are called a layer. A layer takes n inputs and runs them through m perceptrons to produce m outputs. Thus, a layer can be expressed by the weights matrix W, the bias vector $\vec{b}$ and an activation function:

$$\vec{y} = f\left(W\vec{x} + \vec{b}\right) \tag{3.2}$$

Applying multiple perceptrons after one another creates a multi-layer perceptron (MLP), also known as neural network. MLP are able to solve non-linear separable problems if the right structure for the given problem is chosen. The MLP consists of several layers and several nodes within one layer. First is the input layer, where the data enters the neural network. After that comes one or more hidden layers, which take the outputs of the input layer or the preceding hidden layer as input. Finally there is the output layer, which gives one binary output for a simple binary classification problem. For a multi-class problem, multiple outputs are needed. The number of nodes and layer can be chosen freely. In general, it is better to have more hidden layers than needed rather then too few, as the neural network could not be flexible enough to be able to deal with the non-linearities in the data. Extra weights that form excess hidden units vanish when applying fitting regularisation.

A MLP is a fully connected neural network. Each node in one layer is connected with every node in the preceding and following layer. In Figure 3.2, a schematic of a MLP with one hidden layer is shown.

The choice of activation function is crucial to the performance of a neural network. It has direct influence on the network's ability to learn and classify datasets [35]. Non-linear activation functions enable the network to learn complex relations within the data and handle high-dimensional datasets. Additionally, the activation function should be differentiable (almost everywhere) due to the gradient based optimisation of neural networks. In Figure 3.3, some common activation functions are depicted.

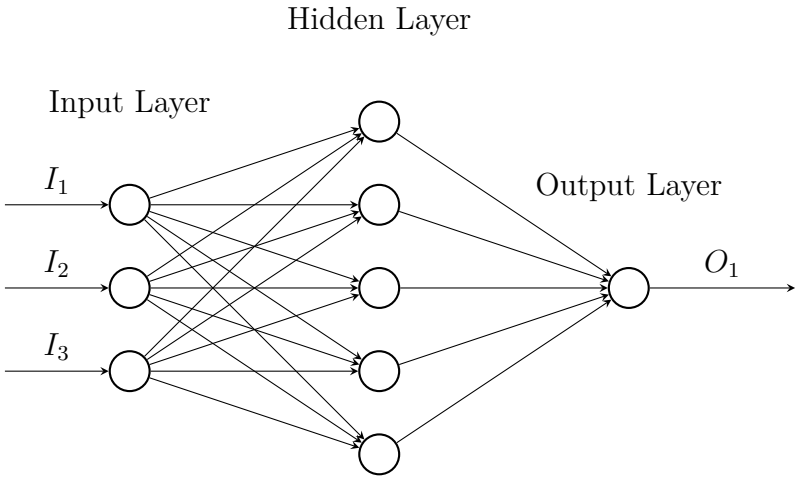


Figure 3.2.: Illustration of a Multi-Layer Perceptron consisting of an input layer, one hidden layer and the output layer with a single node. Each node in the hidden layer represents a perceptron.

The sigmoid activation function [32]

$$\sigma = \frac{1}{1 + e^{-x}} \tag{3.3}$$

is a widely used activation function for binary classification problems with binary variables. It transforms the input values into the range $[0, 1]$.
The preferred activation for most use cases is the Rectified Linear Unit (ReLU)[36]:

$$ReLU(x) = \max(0, x). \tag{3.4}$$

It hinders negative outputs to propagate through the network further. By that it also is more efficient than other activation functions as it causes only a subset of nodes to activate at the same time, reducing computational complexity. However, when the gradient becomes zero, it can be problematic and lead to training instabilities. To counter this problem, similar, but fully differentiable functions have been developed [37].
The softmax activation function

$$Softmax(x_i) = \frac{e^{x_i}}{\sum_n e^{x_n}}, \tag{3.5}$$

where x_i denotes the output for class i, is the activation function of choice for multi-class classification problems. It is an extension of the sigmoid function converting the network output into a probability distribution over all possible classes. The softmax activation function is usually applied to the output layer of a multi-class classification neural network to obtain the final class probabilities.

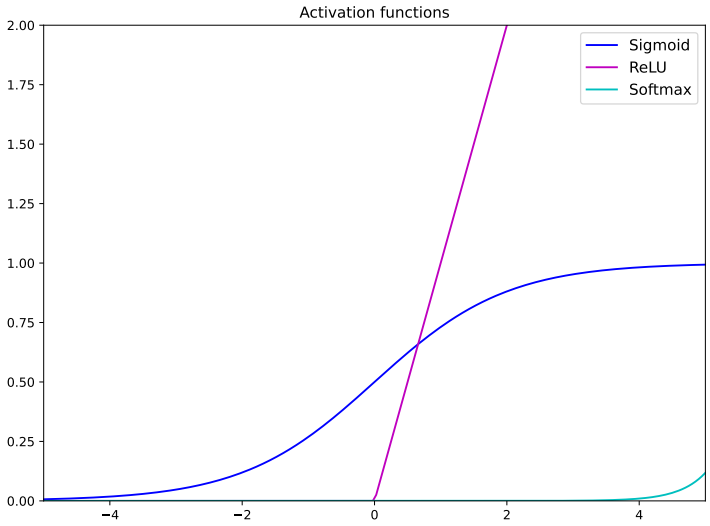


Figure 3.3.: Common activation functions.

Depending on the way nodes are connected, different types of neural networks exist. The most commonly used type of neural network is the Feed-forward neural network (FNN). In FNNs, information is only passed forward through the network layers, from the input to the output layer without looping back.
Deep neural networks (DNNs) are a type of neural network with multiple hidden layers, making them more complex compared to conventional neural networks. There is a variety of DNN architectures, such as Recurrent Neural Networks (RNNs) [38], Convolutional Networks (CNNs) [38] and Graph Neural Networks (GNNs) [39].

3.3. Neural Network Training and Optimisation

To train and evaluate a neural network, the data is split into three groups: training, validation and testing. The training dataset contains the data used to train the neural network. During training, the validation dataset is used to cross check the performance of the neural network. Finally, the test dataset is used to evaluate the performance of the network after training. Thus, the test data cannot be visible to the model during training.
Neural network training is done in epochs. An epoch refers to one complete iteration of the entire training dataset through the learning algorithm. This means that when all the data passed through the neural network, one epoch is completed.
During training, overfitting can occur if the training dataset is too small or the model is not well optimised. In such cases, the neural network performs well on the training data

but not on the test data. It has thus learned the training data too well. In contrast, an underfit model is unable to properly learn the training data and performs poorly in general.
In this section, the loss function, the main metric to evaluate neural network training, is explained in detail. Some optimisation techniques are also presented.

3.3.1. Loss Function

Neural networks are trained by minimising a loss function [32]. The choice of loss function has a huge impact on the model training. Loss functions describe how much the model output and the actual target value deviate from each other. They also offer the means to evaluate model performance during training and provide a metric for the model optimisation. For this, the loss function should be differentiable almost everywhere and furthermore it should have a global minimum. Many different functions obeying these conditions are used as loss functions in machine learning. For classification problems with n classes the Cross-Entropy-Loss

$$\text{Cross-Entropy}(p, t) = -\sum_{i=1}^{n} t_i \cdot \log p_i \tag{3.6}$$

is commonly used. Here p_i denotes the model prediction for class i and t_i the truth label for class i. Combining the Cross-Entropy-Loss with a softmax activation function in the output layer, ensures that the output is $p \in [0, 1]^n$. In this special case the Cross-Entropy-Loss becomes the categorical Cross-Entropy-Loss. Due to its logarithmic nature, the Cross-Entropy-Loss mitigates large deviations from the desired value. Also, minimising the Cross-Entropy-loss is equivalent to maximising the likelihood, which is a well known and often used statistical optimisation technique [32].
In this analysis, the categorical Cross-Entropy loss is used together with the softmax activation function.

3.3.2. Optimisation techniques

For developing efficient and high accuracy neural networks, optimisation plays an important role. Optimisation depends highly on the network's task and the data provided. In this context, optimisation describes the fine-tuning of various network parameters with the goal to find the best performing model. Especially in complex applications, optimisation techniques are essential to obtain high-quality results. Optimisation of neural networks is a challenging and time-consuming task, highly benefiting from experience of

working with neural networks.

Hyperparameters
Parameters that are fixed before the network training are called hyperparameters. The following hyperparameters define the training stability and performance of an FNN:

- **Number of layers**: defines the depth of the neural network.
- **Number of nodes per layer**: defines the width of the neural network.
- **Activation functions used per layer**: influences layer output and network usage.
- **Loss function**: measure of the training performance. The aim is to minimise the loss during training.
- **Optimiser**: enhances the speed and accuracy of training by adjusting the models weights based on gradients. The optimiser's learning rate determines the variation of weights during optimisation. The ADAM [40] optimiser is favoured for its adaptability and efficiency. It is an first-order gradient-based optimiser that is named after Adaptive Moment estimation.

Learning Rate based Optimisation
The learning rate of the optimiser can be set to a constant value or varied during training through scheduling methods. Learning rate scheduling techniques reduce the learning rate over time and the `ReduceLearningRateOnPlateau` (see Appendix A.1.3) method lowers the learning rate when the training loss does not change more than a defined value over a set number of epochs, which helps to improve the training accuracy.

Early Stopping
Early stopping is a method in neural network training to manage computation resources more efficiently. It also prevents overfitting by monitoring the validation loss. If it diverges from the training loss for a set number of epochs the training is stopped and reverted to the point of divergence.

4. Experimental Setup

In this chapter, the machines that allow us to study the $b\bar{b}WW^*$ channel are discussed in detail. This includes an overview of all the accelerators leading up to the Large Hadron Collider (LHC) in Section 4.1. The LHC provides accelerated protons for collision in the ATLAS detector, which is described in detail in Section 4.2. There, the properties of the particles generated in the collision, such as momenta, energies and charges, are recorded through multiple detection layers.

4.1. The Large Hadron Collider

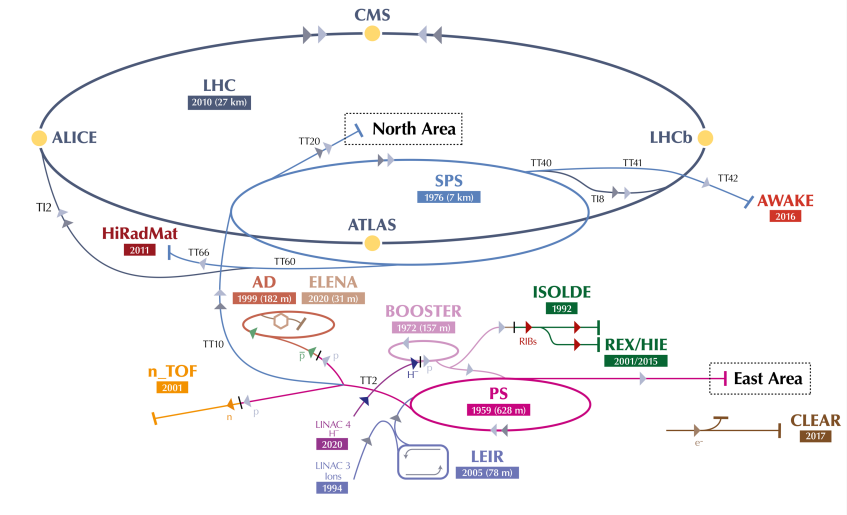


Figure 4.1.: Schematic of the full acceleration complex of the LHC [41].

The Large Hadron Collider is a synchrotron accelerator with a circumference of 27 km based at CERN in Geneva [42]. It is designed as a proton-proton (pp) as well as heavy ion collider with a current pp-collision centre-of-mass energy of 13.6 TeV. To achieve such high energies, the protons need to be accelerated through a series of acceleration steps seen in Figure 4.1. The start is the Linear Accelerator 4 (LINAC4) where negative hydrogen ions are accelerated to 160 MeV and passed to the Proton Synchrotron Booster (PSB). The PSB strips away the electrons, leaving only the protons, which are after that

accelerated further to 2 GeV and passed to the Proton Synchrotron (PS). The protons then reach energies of 26 GeV before being injected into the Super Proton Synchrotron (SPS), which accelerates the protons to energies of 250 GeV. Following that, the protons are passed to the LHC. In the LHC, it takes roughly 20 minutes for the protons to reach their final energy. At that point, the protons travel at almost the speed of light.

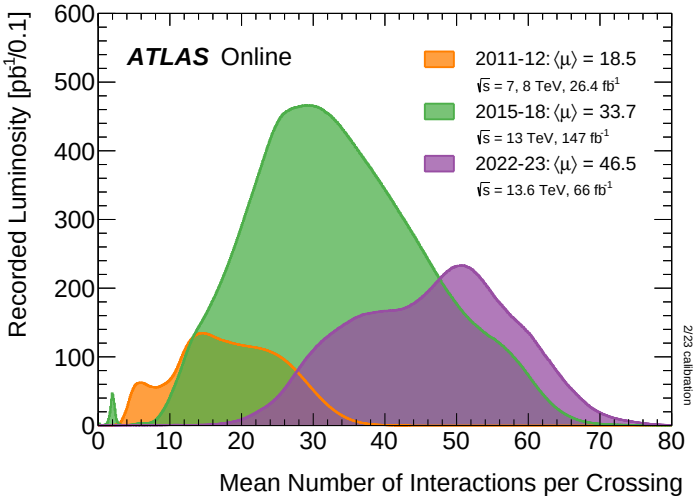


Figure 4.2.: The average number of interactions per bunch crossing through Run 1 (2011-12) to Run 3 (2022-23) [43].

Two proton beams, going in opposite directions, are accelerated in this way. The dipole magnets used for forcing the protons onto a circular orbit need field strengths of $B = 8\,\mathrm{T}$. Hence, the magnets are superconducting and need to be cooled down to 2 K. The beams are focused via quadrupole and higher order multipole magnets. The protons travel through the accelerator in bunches of 1.15×10^{11} protons per bunch. Bunches follow one another with a time interval of 25 ns. Collisions take place with a rate of 40 MHz. Every time the two beams cross, many interactions occur. For Run 3, there are on average 46.5 interactions per bunch crossing, as can be seen in Figure 4.2. The figure shows the recorded luminosity with respect to the mean number of interactions per crossing for Run 1, Run 2 and Run 3.

From the frequency of the bunch crossings f, the particle number per bunch in both beams N_1 N_2, the number of bunches n_b, the transverse beam dimension $\sigma_{x/y,i}$ and a reduction factor S, the luminosity delivered by the LHC can be determined by [44]

$$\mathscr{L} = \frac{N_1 N_2 n_b f S}{2\pi\sqrt{\sigma_{x,1}^2 + \sigma_{x,2}^2}\sqrt{\sigma_{y,1}^2 + \sigma_{y,2}^2}}. \tag{4.1}$$

During operation, the proton beams lose protons due to collisions and beam losses, such

that after approximately 15 h the beams are dumped and new beams are created.
An experiment is situated at each of the four beam crossings of LHC. The two large experiments are ATLAS [45] and CMS [46], both general purpose experiments focusing on precision measurements and search for BSM physics. The LHCb [47] experiment focuses on b-physics and the ALICE [48] experiment studies lead-lead collisions, providing conditions that mime the state of the universe shortly after the Big Bang.
The LHC in not running constantly. In between run periods there are long shutdown phases where the machine is maintained and upgrades are installed. The LHC started operation in 2010 with lower luminosity of $36\,\text{fb}^{-1}$ and a centre-of-mass energy of 7 to 8 TeV. In Run 2, from 2015-2018, both luminosity and centre-of-mass energy increased further to an integrated luminosity of $140\,\text{fb}^{-1}$ and $\sqrt{s} = 13\,\text{TeV}$. In the current Run 3 period, which is scheduled from 2022 to 2025, the target integrated luminosity is $250\,\text{fb}^{-1}$, nearly double the amount of Run 2, and the centre-of-mass energy is at 13.6 TeV. As of now, roughly $80\,\text{fb}^{-1}$ of integrated luminosity have been accumulated for Run 3. The long-term aim is to repair and upgrade the LHC and ATLAS in the next long shutdown to reach an integrated luminosity of $3000\,\text{fb}^{-1}$ in Run 4.

4.2. Atlas Detector

Since this Master's thesis analyses simulated data of the ATLAS experiment, an overview of the detector is provided based of Ref. [45].
The ATLAS (A Toroidal LHC ApparatuS) [45, 49] detector is the largest particle detector at the LHC with a length of 46 m and a diameter of 25 m. It weighs 7000 t. It has a cylindrical shape and the beam axis corresponds to its symmetry axis. The detector covers nearly the full solid angle. As it is a general-purpose detector, it comprises of several detector layers in an onion like structure with different tasks each. It can be divided in three main detector parts: the Inner Detector (ID), focusing on tracking information, the Calorimeter System, focusing on energy measurements, and the Muon Spectrometer for muons which do not deposit enough energy in the previous detector parts to be stopped and thus, are measured again. Figure 4.3 shows a schematic illustration of the ATLAS detector.
The detector components are described in more details in the following including the detector coordinate system and the most common variable definitions used in this thesis.

Coordinate System

The ATLAS detector is a cylindrical detector with an orthogonal right-handed coordinate

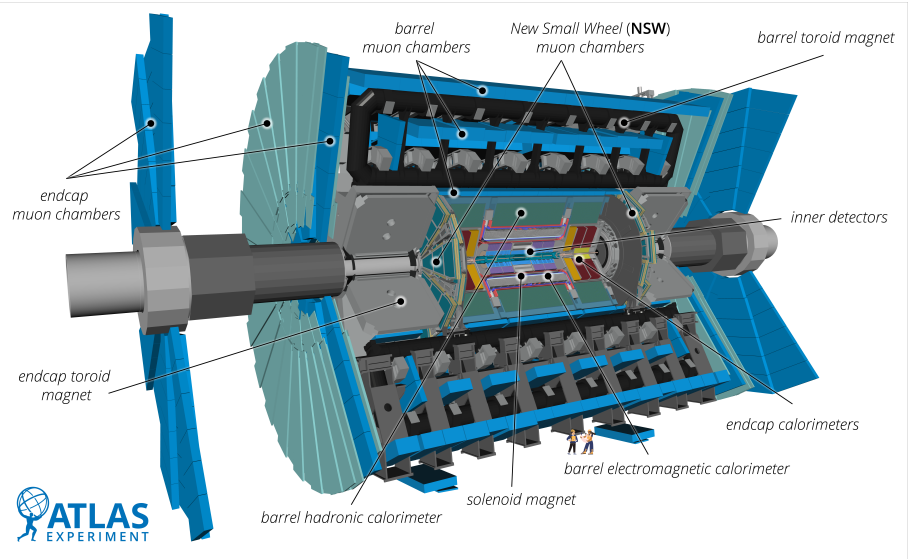


Figure 4.3.: Computer-generated overview of the ATLAS detector. People for scale. Taken from Ref. [50].

system, where the origin is placed at the center of the detector. In Figure 4.4, the ATLAS coordinate system is shown. In most cases a spherical coordinate system is used with the azimuthal angle ϕ around the z-axis in the x-y-plane and the polar angle θ between the z-axis and the x-y-plane. The azimuthal angle ϕ ranges from $-\pi$ to π and the polar angle can take values from 0 to π.

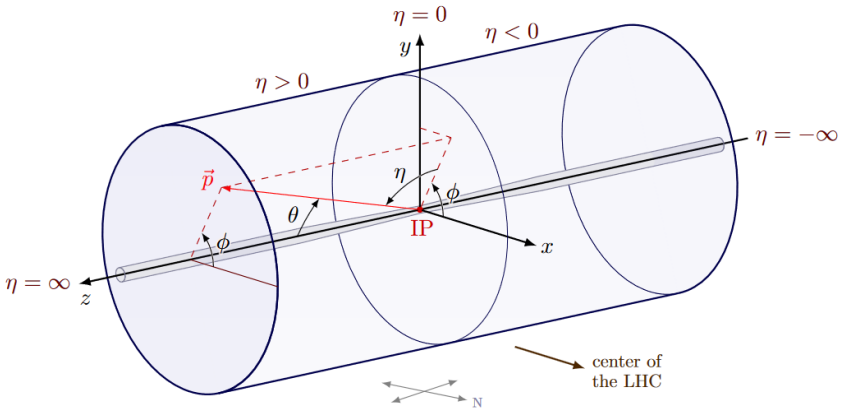


Figure 4.4.: Sketch of the ATLAS coordinate system. Cartesian and spherical coordinates are shown, as well as the pseudorapidity η. The red arrow illustrates a particle with momentum $\vec{p}$ travelling through the detector.

As the partons inside the protons collide with one another in the ATLAS detector the exact momentum of the decay products is unknown. Thus, transverse variables play an

important role for analyses, as they are invariant under Lorentz-boosts in the z-direction. One important variable is the transverse momentum of a particle

$$p_T = \sqrt{p_x^2 + p_y^2}, \tag{4.2}$$

where p_X and p_y are the respective momenta in x- and y-direction.
Another key variable is the pseudorapidity

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) = \frac{1}{2}\ln\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right). \tag{4.3}$$

Here $\vec{p}$ is the total momentum and p_z the momentum in z-direction of the particle.
The variable ΔR describes distances between two particles a and b. It is defined as

$$\Delta R(a,b) = \sqrt{\Delta\eta^2 + \Delta\phi^2} = \sqrt{(\eta_a - \eta_b)^2 + (\phi_a - \phi_b)^2} \tag{4.4}$$

and is approximately invariant under boosts in the z-direction.

Inner Detector

Going outward from the interaction point, the first detector layer is the Inner Detector. The Inner Detector is 6.2 m long and has a diameter of 2.1 m. Its purpose is the detection of charged particles trajectories by measuring ionisation signals in the detector layers which are combined into tracks. Tracks are combined signals of multiple hits in the detectors of the Inner Detector to determine the path a particle took through the detector and to determine some of its properties such as its transverse momentum. The Inner Detector starts at a radial distance of 3.3 cm from the interaction point and consists of multiple layers of pixel and strip detectors as well as a transition radiation tracker (TRT).
The pixel part is the most inner part of the Inner Detector providing a high resolution for reconstruction of the interaction point and tracks. The inner most pixel layer is the Insertable B-Layer [51], which was installed shortly before Run 2. It has a pixel size of $50 \times 250\,\mu\text{m}^2$ to improve the vertex resolution even further. The pixels in the subsequent layers all have a pixel size of $50 \times 400\,\mu\text{m}^2$. The strip detector part follows the pixel part. It consists of four layers in the barrel region and nine end-cap disks on each side. These strips are placed every $80\,\mu$m and slightly rotated with respect to the preceding one. The pixel and the strip detector parts cover a range of $|\eta| < 2.5$. The last part of the Inner Detector is the TRT, covering the $|\eta| < 2.0$ range. The TRT is made out of long straw tubes filled with a gas mixture following the principle of drift tubes. When a particle passes through the TRT it excites the gas mixture which is the transferred into an electrical signal.

The Inner Detector lies within the 2 T magnetic field of a solenoid magnet parallel to the beam axis. Due to this magnetic field, the Lorentz force acts on charged particles, forcing their tracks to bend according to their charge and momentum. The spatial resolution of the Inner Detector lessens with the distance from the interaction point. The overall momentum resolution of the Inner Detector is $\sigma_{p_T}/p_T = 0.05\%\, p_T\,[\text{GeV}] \oplus 1\%$

Calorimeters

After going through the Inner Detector, particles enter the calorimeter, which measures the energy of the particle independently of its charge by stopping it via a mechanism called showering. Particles passing through dense materials, interact with the material and emit radiation and undergo pair-production, by which the number of particles increases. The newly produced particles again interacting and create a particle shower.

The calorimeter is divided into the electromagnetic calorimeter (ECAL), dominantly measuring the energy of electrons, positrons and photons, and the hadronic calorimeter (HCAL), measuring the energy of hadrons. The ECAL covers the η range $|\eta| < 3.2$ in barrel and endcaps region. The HCAL has an additional η coverage in the forward region of $3.1 < |\eta| < 4.5$ for precision physics analysis. The exact η coverage for the two subsystems can be seen in Table 4.1.

The ECAL has a total thickness of >22 radiation lengths (X_0) in the barrel and >24 X_0 in the end-caps and a high granularity, providing good resolution for precision measurements electrons and photons. The ECAL is surrounded by the HCAL. The HCAL has a coarser granularity, sufficient for jet construction. As electrons, positrons and photons only shower via electromagnetic processes, whereas hadrons mostly interact via the strong force, different materials, both active and passive, are used for ECAL and HCAL. The ECAL consists of alternating layers of liquid argon active material and lead passive material. Analogous to the ECAL, the HCAL is split up into the barrel, end-cap and forward calorimeter parts. In the end-caps and in the forward region of the HCAL, also liquid argon is used as the active material with copper passive material, and additional tungsten passive material in the forward region. The barrel region of the HCAL uses scintillation tiles as active material and steel as the passive material. The energy resolution of the ECAL is $\sigma_E/E = 10\%/\sqrt{E}\,[\text{GeV}] \oplus 0.7\%$. For the HCAL the energy resolution in the barrel and end-cap parts are $\sigma_E/E = 50\%/\sqrt{E}\,[\text{GeV}] \oplus 3\%$ and in the forward part $\sigma_E/E = 100\%/\sqrt{E}\,[\text{GeV}] \oplus 10\%$.

Muon Spectrometer

Muons are minimal ionising particles. Muons however, are tracked in the Inner Detector

and only deposit minimal energy in the calorimeter system. Since the muon's momentum is typically very high, it is not measured accurately in the Inner Detector. To improve the muon momentum measurement an additional tracking detector is used, the muon spectrometer. It is the most outer part of the ATLAS detector and surrounded by three toroid magnets, which bent the muon's track, facilitating the momentum measurement. The muon spectrometer is comprised of 4000 individual muon chambers applying various detector technologies. Monitored Drift Tubes measure the track curvature and Resistive Plate Chambers are used for triggering and tracking in the Muon Spectrometer barrel region, while Cathode Strip Chambers and Thin-Gap Chambers provide precise spatial resolution in the end-caps. The Muon Spectrometer has a momentum resolution of $\sigma_{p_T}/p_T = 10\%$ at $p_T = 1\,\text{TeV}$.

The schematic in Figure 4.5 illustrates the behaviour of different particles in the individual detector layers of the ATLAS detector, and in Table 4.1, the resolutions and η coverage of all the detector parts are summarised.

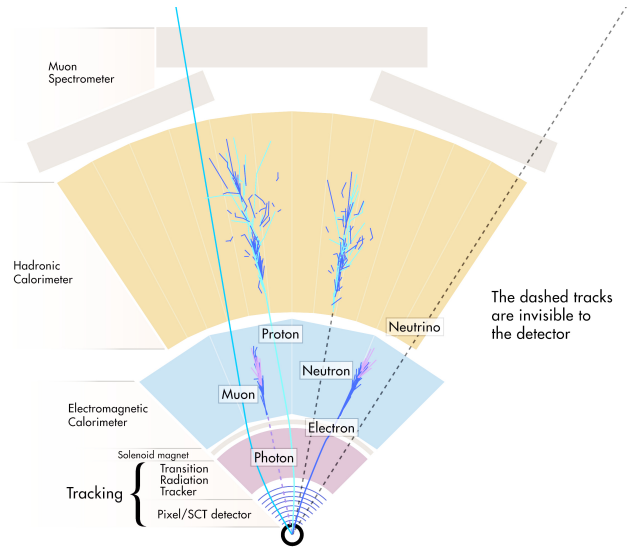


Figure 4.5.: Sketch of the ATLAS detector layers with all particle interactions. ©CERN (colours inverted)

Trigger System

The amount of data produced by the ATLAS detector is far too much to store. Per second, 1.348 billion interactions occur. Though most of the interactions occur at low energy scales and thus do not feature interesting physics for ATLAS analyses. Therefore, it is highly important to filter the events in real time and only store interesting ones.

Detector component	**Required resolution [GeV]**	**η coverage**	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	± 2.5	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1\,\text{TeV}$	± 2.7	± 2.4

Table 4.1.: Summary table of resolutions and η coverage of the individual parts of the ATLAS detector [45].

For this purpose, there is a two-stage trigger system in place [52, 53]. The first trigger stage is the Level-1 (L1) trigger which uses hardware-level information from the calorimeters and the Muon Spectrometer to select events with large energy deposits in the calorimeter or a high energy muon. After the L1 trigger, the event rate is reduced to 100 kHz.
The second trigger stage is the High-Level Trigger (HLT), which is a software based trigger and uses all available information and simplified reconstruction algorithms. The HLT runs through multiple steps. Each step can cause the current event to be removed. The HLT reduces the event rate further to 3 kHz, resulting in approximately 3 GB data stored per second.

5. The $X \to HH \to b\bar{b}WW^*$ Channel

This Master's thesis investigates the resonant production of Higgs boson pairs via a scalar resonance X in the $b\bar{b}WW^*$ decay channel with one charged lepton in the final state. The purpose of this chapter is to first introduce the possible topological signatures of this decay channel. Afterwards the simulated data used for this analysis is described, and in the end an overview of the relevant reconstructed objects is given.

5.1. Decay Topologies

Depending on the resonant mass m_X, the topology of this decay chain differs. For very high resonant masses $m_X \gtrsim 2\,\text{TeV}$, the final state will be boosted. A schematic for the boosted topology can be seen in Figure 5.1. The hadronic decay products of the Higgs boson decaying into an b quark pair are collimated and cannot be reconstructed as separate objects due to detector resolution. The same is true for the hadronically decaying W boson, which is one of the two W bosons into which the other Higgs boson decays. Furthermore, the lepton from the leptonically decaying W boson is very close to the decay products of the hadronically decaying W boson and might even overlap with them. The boosted topology is very unique and is easier to distinguish from backgrounds but comes with challenges in the object reconstruction.
For lower resonant masses, the topology becomes more and more resolved: All objects are well separated from one another. This topology is also relevant for non-resonant Higgs boson pair production as predicted by the SM.
This Master's thesis focuses on the boosted topology.

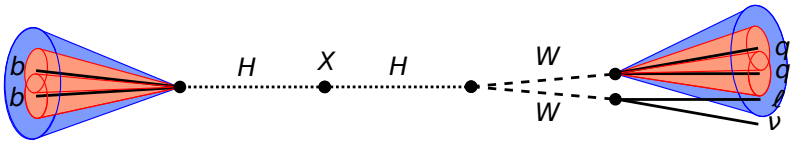


Figure 5.1.: The boosted decay topology for the decay $X \to HH \to b\bar{b}WW$ in the 1-lepton final state.

5.2. Monte Carlo Simulation of Signal and Background

The raw data from the ATLAS detector consists of spatial coordinates, the time and signal amplitudes from multiple detector parts. Monte Carlo (MC) samples imitate such raw data of the detector with the advantage of knowing which particle caused which signal called truth information. This study only uses MC simulated data.

$X \to HH$ Samples

For the signal decay channel three different mass-points $m_X = 2,\ 3$ and $4\,\text{TeV}$ of the heavy scalar X are analysed. The signal samples are generated using MADGRAPH 3.5.1 [54] with the `NNPDF23LO` tune [55]. The hadronisation and parton shower are simulated with PYTHIA 8.308 [56] with the heavy quark flavours being generated with EVTGEN 2.1.1 [57]. The $H \to b\bar{b}$ and $H \to WW^*$ branching ratios are set to be 50 % each. A final state filter ensures that one Higgs boson decays into $b\bar{b}$ and the other to WW^* and that there is one charged lepton in the final state. The MC samples are then converted to another file format for use in the analysis framework using the. Each of the three files contains 50.000 events. The exact filenames for these samples are listed in Appendix A.2.

Background Samples

For the scope of this thesis, three different kinds of background events are analysed: top quark pair production ($t\bar{t}$), W boson production in association with jets (W + jets) and multi-jet production (dijet).

For the production of $t\bar{t}$ samples, the POWHEG v2 generator is used with the `NNPDF23LO` tune. PYTHIA 8.307 is used for parton shower simulation with EVTGEN 2.1.1 for heavy quark flavours. In the $t\bar{t}$ background sample a filter is applied to ensure an one lepton final state. It contains 50 million events.

The W + jets background sample is generated using SHERPA 2.2.14 [58]. It is comprised of multiple sub-samples containing several million events each, and each applying different jet flavor filters. These were added together to result in the final background sample.

The dijet background sample is generated using PYTHIA 8.308 with EVTGEN 2.1.1 for patron shower simulation. The dijet sample is also divided into multiple sub-samples, filtered by the transverse momentum of the leading truth jet. The size of the sub-samples varies from 1 million to 30 million events per file.

All the background samples are converted to another file format again and their names are also listed in Appendix A.2.

5.3. Object Reconstruction

To interpret raw data of the ATLAS detector or MC data, multiple algorithms are used to identify the particles responsible for the real or simulated detector responses. In Figure 4.5, the detector responses for different particles are shown. To end up with accurate values for transverse momenta and particle masses, energy calibrations are applied. Lastly, to prevent double counting energy, an overlap removal is usually applied. In this thesis, however, there is no overlap removal applied yet, because the overlap removal was not yet correctly implemented for large radius jet objects used in this analysis.

5.3.1. Leptons

In pp-collisions, there are two ways to produce leptons. They can come from the decay of heavy particles such as W bosons and top quarks produced in the collision or they stem from the decay of hadrons. For the first possibility, the decay happens very quickly and close to the primary vertex of the collision such that is considered as identical. These leptons are called prompt leptons. For leptons produced in hadron decays on the other hand, the decay happens after the hadron travelled a considerable distance from the primary vertex, and the two vertices can be distinguished. These leptons are called non-prompt leptons.
For prompt leptons, only electron and muons are considered in this analysis, as tau leptons decay very quickly and require more advanced reconstruction methods. Electrons and muons from tau decays are also considered prompt for this analysis. Neutrinos do not interact with the detector.
To be able to distinguish prompt and non-prompt leptons, various selection criteria are applied on the leptons. These criteria can be seen in Table 5.1 and are explained in the following.

		Electrons	Muons
p_T^{min}		$>10\,\text{GeV}$	$>10\,\text{GeV}$
$\lvert\eta\rvert^{\text{max}}$		<2.47 w/o $1.37 < \eta < 1.52$	<2.5
Identification	all	`LooseBLayer`	`Loose`
	prompt	`Medium`	`Medium`
Isolation	all	`NonIso`	`NonIso`
	prompt	`TightTrackOnly_VarRad`	`PflowLoose_VarRad`

Table 5.1.: Current lepton selection criteria for all leptons and prompt leptons. The prompt lepton selection is applied on top of the selection for all leptons. For identification and isolation, the working point names are given. The definitions of the isolation working points are found below and in Ref. [59].

Electrons

Electrons are reconstructed from Inner Detector tracks and energy deposits in the electromagnetic calorimeter. To ensure that the reconstructed object is indeed an electron, all electron candidates need to pass a likelihood identification (ID). Many discriminating variables are used for this identification procedure which can be found in Table 1 of Ref. [60]. Based on these variables, different working points are defined for the ID. Here, the available working points are `Loose`, `Medium` and `Tight`, with the `Loose` working point putting the least and the `Tight` working point the most constraints on the electron candidates.

Additionally to the ID, the isolation of the electron is important to determine if it is prompt or non-prompt. It gives a measure on how many other particles are in the vicinity of the electron by putting a cone based on transverse momentum or energy around the electron's track or calorimeter cluster and calculate how much transverse momentum or energy of other particles is inside that cone. The less transverse momentum or energy that is inside such a cone, the more isolated the electron is. The isolation can be expressed in terms of the transverse energy as [60]

$$E_{\mathrm{T}}^{\mathrm{coneXX}} = E_{\mathrm{T,raw}}^{\mathrm{XX}} - E_{\mathrm{T,core}} - E_{\mathrm{T,leak}}^{\mathrm{XX}}(E_{\mathrm{T}}, \eta) - E_{\mathrm{T}}^{\mathrm{XX}}(\eta) \tag{5.1}$$

where XX refers to the cone size $\Delta R = \mathrm{XX}/100$. The track isolation variable $\left(p_{\mathrm{T}}^{\mathrm{coneXX}}\right)$ is calculated analogously in terms of the transverse momentum. For the isolation for PFOs [61], selected tracks are combined with neutral PFOs $\left(E_{\mathrm{T,neflow}}^{\mathrm{coneXX}}\right)$ into cones around the electron's track.

It is also possible to define isolation variables based a cone with variable radius. This variable size cone (varcone) is defined as

$$\Delta R = \min\left(\frac{10\,\mathrm{GeV}}{p_{\mathrm{T}}}, \frac{\mathrm{XX}}{100}\right). \tag{5.2}$$

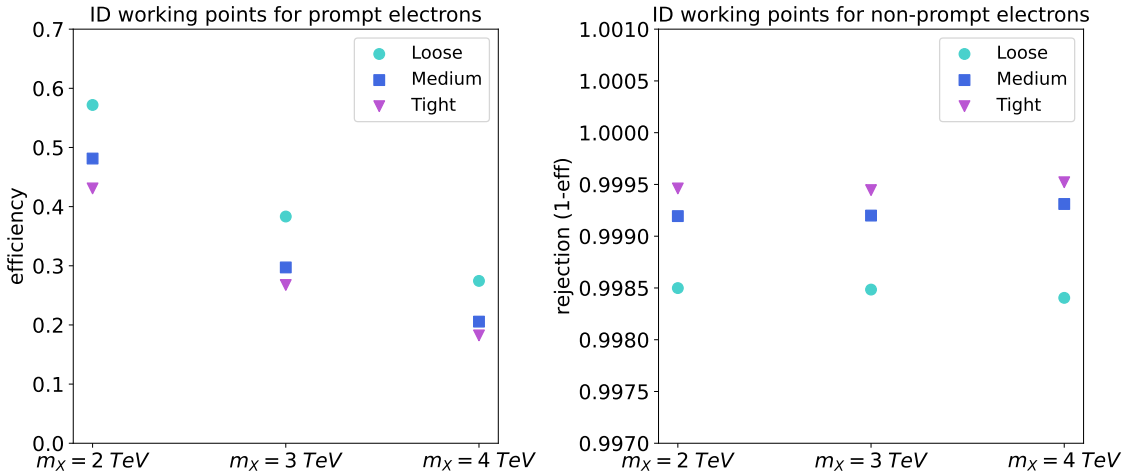
The definitions of the different isolation working points based of these variables can be seen in Table 5.2.

WP name	Definition
`HighPtCaloOnly`	$E_{\mathrm{T}}^{\mathrm{cone20}} < \max\left(0.015 \cdot \mathrm{p_T}, 3.5\,\mathrm{GeV}\right)$
`Loose`	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}} < 0.15$ && $E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.3$
`Tight`	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}} < 0.06$ && $E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.06$
`TightTrackOnly_VarRad`	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}} < 0.06$
`TightTrackOnly_FixedRad`	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}} < 0.06$ ($p_{\mathrm{T}} < 50\,\mathrm{GeV}$), $p_{\mathrm{T}}^{\mathrm{cone30}}/p_{\mathrm{T}} < 0.06$ ($p_{\mathrm{T}} > 50\,\mathrm{GeV}$)

Table 5.2.: Definition of the available isolation working points for electrons [59].

To be used in this analysis, certain selection criteria are placed on the electron. These selection criteria are seen in Table 5.1. Previous studies [62] showed that these criteria work well with the given topology. In the electron selection, the crack region between the barrel and the end-caps is excluded in the $1.37 < \eta < 1.52$ window because in that detector region, instrumentation is lacking, due to the electronics and readout cables there.

A brief study of the available electron ID and isolation working points was performed to decide which ones to choose. For this study, electrons candidates are divided into prompt and non-prompt electrons by truth level information provided by the framework.
The electron ID efficiencies for the signal samples are shown in Figure 5.2. As is expected, the efficiencies decrease for tighter electron ID working points across all available m_X. For prompt electrons, the efficiency also decreases with increasing m_X, because with increasing m_X the overlap between the lepton and the hadronically decaying W boson also increases, making it harder to identify the electron. This behaviour is not observed for non-prompt electrons. The rejection of the non-prompt electrons is above 99.7 % for all three ID working points, with the `Tight` working point having the highest rejection, as would be expected.
Despite having a lower efficiency than the `Loose` working point, the `Medium` ID working point was chosen for electrons, as previous studies [62] showed how the analysis benefits from it.

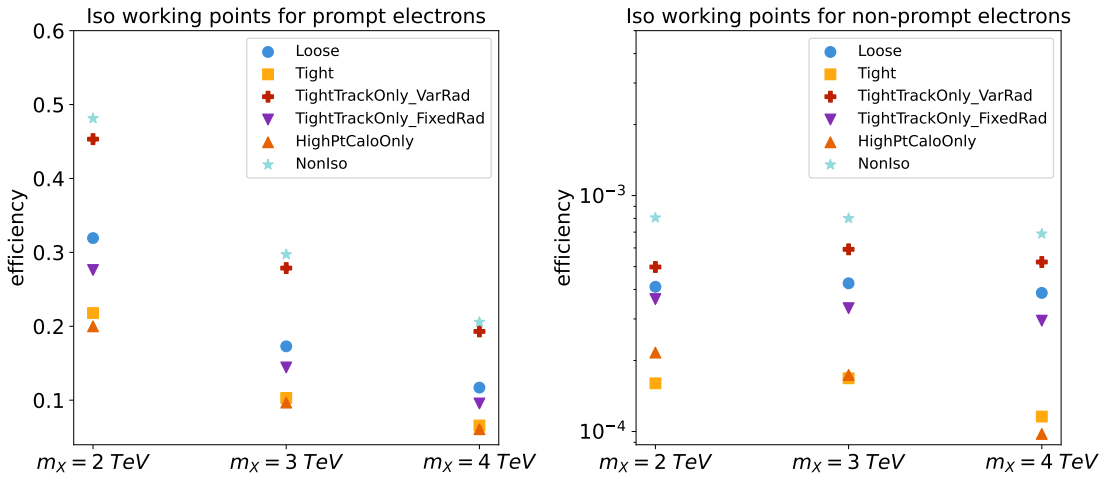


(a) prompt electrons **(b)** non-prompt electrons

Figure 5.2.: Prompt electron efficiencies and non-prompt electron rejections of ID working points.

Next, a similar study was performed for many electron isolation working points. The results are shown in Figure 5.3 for electrons passing the `Medium` ID working point. The

isolation working point with the best performance is the `TightTrackOnly_VarRad` isolation working point. It has the highest prompt electron efficiency, except for when no isolation is applied, while the non-prompt electron efficiency is similar to the `Loose` isolation working point which makes the `TightTrackOnly_VarRad` the best working point to use for this analysis.



(a) prompt electrons **(b)** non-prompt electrons

Figure 5.3.: Prompt and non-prompt electron efficiencies of isolation working points. `Medium` ID is applied to both prompt and non-prompt electrons.

Muons

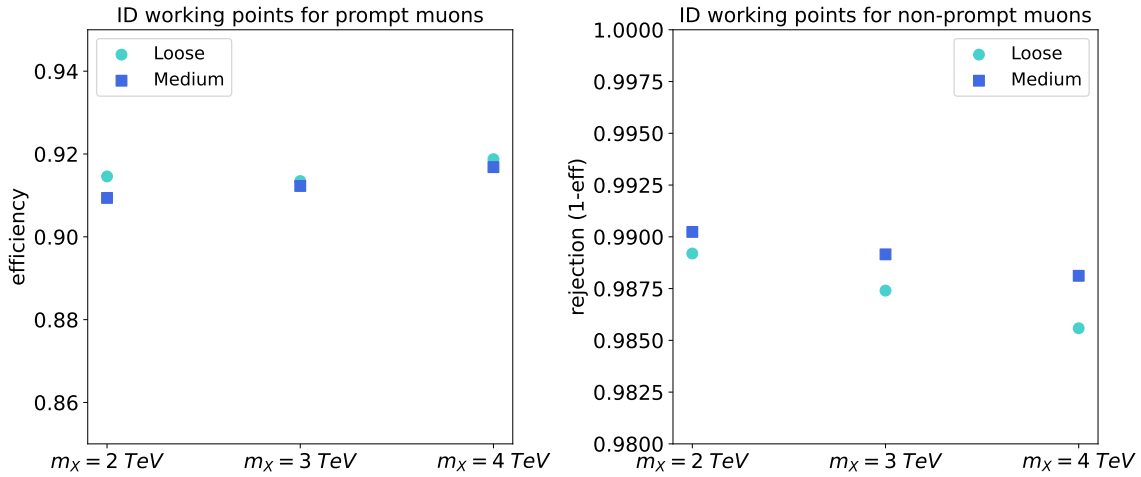
Muons are reconstructed using information from all parts of the ATLAS detector, including tracks from the Inner Detector, energy deposits in the calorimeters and also tracks in the Muon Spectrometer. For muons, there are also criteria placed on all muon candidates to be able to distinguish between prompt and non-prompt muons. These selection criteria can also be seen in Table 5.1. In contrast to electrons, the crack region is not excluded as the muon spectrometer covers that η region.

Likelihood identification and isolation working points are also applied on all muon candidates. For muons, only the `Loose` and `Medium` ID working points are available and the definitions of the muon isolation variables can be found in Table 5.3. Again, a study of the available ID and isolation working points is performed for muons. The split into prompt and non-prompt muon candidates is also done by looking at truth level information.

WP name	Definition
`PflowLoose_VarRad`	$(p_{\mathrm{T}}^{\mathrm{varcone30}} + 0.4\ E_{\mathrm{T,neflow}}^{\mathrm{cone20}})/\mathrm{p_T} < 0.16$
`PflowTight_VarRad`	$(p_{\mathrm{T}}^{\mathrm{varcone30}} + 0.4\ E_{\mathrm{T,neflow}}^{\mathrm{cone20}})/\mathrm{p_T} < 0.045$
`Loose`	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}} < 0.15$ && $E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.3$
`Tight`	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}} < 0.04$ && $E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.15$
`TightTrackOnly_VarRad`	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}} < 0.06$

Table 5.3.: Definition of the available isolation working points for muons [59].

Figure 5.4 shows the results of the study of muon ID working point efficiencies. For both working points, the efficiencies are above 90 % for prompt muons across all studied values of m_X. In contrast to the prompt electron efficiencies, the dependence on m_X in negligible for prompt muons. For non-prompt muons, the rejection for both working points is very high with roughly 99 %. For the `Medium` ID working point, the rejection is slightly higher than for the `Loose` ID working point. The `Medium` ID working point is chosen for further studies due to its high efficiency in prompt muon identification in combination with a high non-prompt muon rejection. Also some later investigations of the dijet background, which can be found in Section 6.3.2, strongly suggest the use of the `Medium` ID working point.

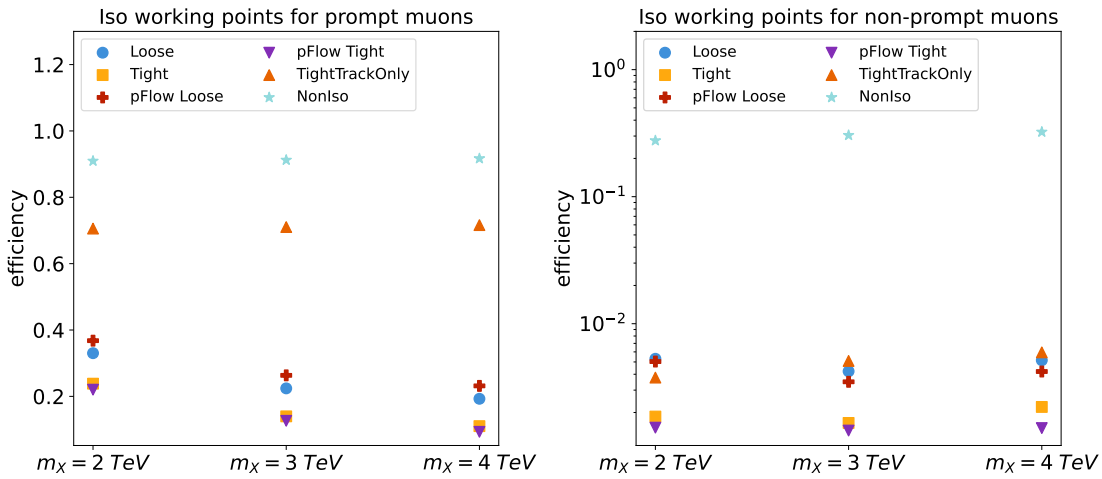


(a) prompt muons **(b)** non-prompt muons

Figure 5.4.: Prompt muon efficiencies and non-prompt muon rejections of ID working points.

Next, the available isolation working points are checked. The isolation efficiencies for these working points for all three signal samples can be seen in Figure 5.5. The `TightTrackOnly_VarRad` working point performs best. Isolation efficiencies decrease significantly for all the other working points. There is also a decrease in efficiency with an

increase in m_X for all but the `TightTrackOnly_VarRad` working point. The reason for this is the increasing overlap of the hadronically decaying W boson and the prompt muon. This is illustrated in Figure 5.6 which shows the efficiencies of the isolation working points



(a) prompt muons **(b)** non-prompt muons

Figure 5.5.: Efficiencies of isolation working points for muons. `Medium` ID is applied to both prompt and non-prompt muons.

for the $m_X = 3\,\text{TeV}$ sample with respect to the muon p_T for muons passing the `Medium` ID working point. Up until $p_T^\mu \lesssim 50\,\text{GeV}$, all isolation working point efficiencies increase but only the `TightTrackOnly_VarRad` isolation efficiency continues to increase beyond that and reaches nearly full efficiency around $p_T^\mu \gtrsim 400\,\text{GeV}$. Again, the muon isolation working point with the highest efficiency is the `TightTrackOnly_VarRad` working point with about 70 % efficiency.

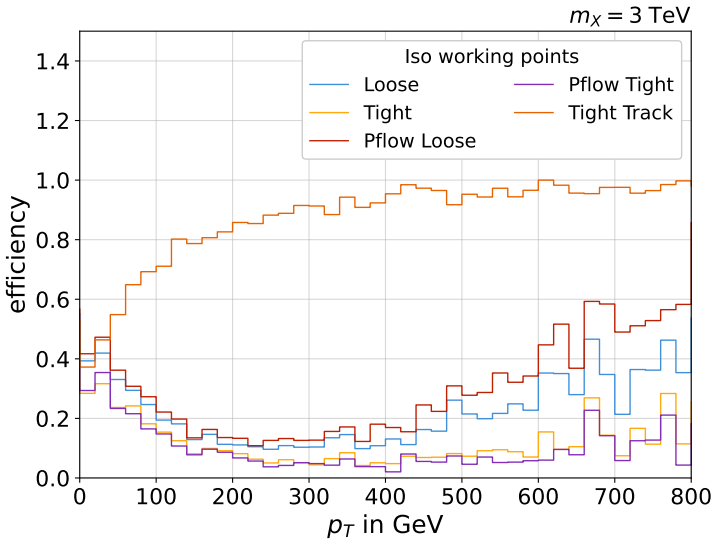


Figure 5.6.: Fraction of muons passing the isolation working point with respect to their p_T for the $m_X = 3\,\text{TeV}$ signal events.

Unfortunately, the `TightTrackOnly_VarRad` isolation working point is not yet recommended for muons in Run 3, so it cannot be used currently. To obtain the recommendations for this working point, the corresponding muon scale factors need to be determined which was beyond the scope of this thesis.
For now, we use the recommended `PflowLoose` isolation working point, which has the highest efficiency for prompt muons of the recommended working points.

5.3.2. Jets

A jet develops as the result of the hadronisation of quarks and gluons. Quarks and gluons produced in the collision (partons) hadronise and cause collimated showers of hadrons. These hadrons leave energy deposits in the calorimeters and tracks in the Inner Detector if they are charged. Usually each parton from the collision produces its own jet.
Jets are reconstructed from multiple objects. For the reconstruction of the energy deposits usually Topological Clusters [63] are used. Jets can also be reconstructed from tracks in the Inner Detector. Another possibility is to use combinations of tracks and clusters for reconstruction.

For reconstruction, multiple algorithms have been developed [64, 65] and the most com-

mon algorithms use the formula

$$d_{ij} = \min\left(p_{\mathrm{T},i}^k, p_{\mathrm{T},j}^k\right) \frac{\Delta R_{ij}^2}{R^2} \text{ with } k = \begin{cases} -2 & \text{anti-}\ k_t \\ 0 & \text{Cambridge-Aachen} \\ 2 & k_t \end{cases}, \tag{5.3}$$

where k defines the type of algorithm used and R denoting the size parameter.

In a boosted topology, it is possible that multiple partons create a single jet. To find out how many partons initiated a given jet, information on the energy distribution inside the jet is needed, also called substructure. For that, two types of substructure variables can be used. The first type uses the Energy Correlator Functions (ECFs) which are defined as

$$\mathrm{ECF}(N,\beta) = \sum_{i_1<i_2<\ldots<i_N\in J} \left(\prod_{a=1}^{N} p_{T,i_a}\right) \left(\prod_{b=1}^{N-1}\prod_{c=b+1}^{N} \Delta R_{i_b i_c}\right)^{\beta}. \tag{5.4}$$

The ECFs indicate the energy distribution inside the jet. The ratios

$$C_2(\beta) = \frac{\mathrm{ECF}_3 \times \mathrm{ECF}_1}{\mathrm{ECF}_2^2} \text{and} \tag{5.5}$$

$$D_2(\beta) = \frac{\mathrm{ECF}_3 \times \mathrm{ECF}_1 \times \mathrm{ECF}_1^2}{\mathrm{ECF}_2^3} \tag{5.6}$$

can be used to discriminate between two-prong and one-prong jets. Two-prong jets could correspond to the boosted decay of a hadronically decaying Higgs or W boson and one-prong jets would be more consistent with background events.
The second type of variable is a quantity called N-subjettiness τ_N which is defined as

$$\tau_N(\alpha) = \frac{1}{d_0(\alpha)} \sum_{i\in J} p_{T,i} \cdot \min\left(\Delta R_{1i}^{\alpha}, \Delta R_{2i}^{\alpha}, ..., \Delta R_{Ni}^{\alpha}\right) \tag{5.7}$$

$$\text{with } d_0(\alpha) = \sum_{i\in J} p_{T,i} \cdot \Delta R^{\alpha}. \tag{5.8}$$

N-subjettiness quantifies how likely a jet is made up of at least N subjets. The ratio $\tau_{MN} = \tau_M/\tau_N$ with $M > N$ becomes small if a jet contains M or more subjets.

In Table 5.4, a summary of the different kinds of jets used in this thesis can be found. The following paragraphs describe all used jet types in more detail.

Collection	Inputs	Size R	p_T[GeV]	$\|\eta\|$	Use case
PFlow Jet	Particle Flow Objects	0.4	> 20	< 4.5	used in $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation and overlap removal
UFO Jets	Unified Flow Objects	1.0	> 200	< 2.0	reconstruction of hadronically decaying analysis objects

Table 5.4.: Summary of the jet collections used in this analysis.

Particle Flow Jets

Particle flow (PFlow) jets are reconstructed from Particle Flow Objects (PFOs) calibrated at the electromagnetic scale [66]. PFOs are reconstructed from both tracking and cluster information, which makes their resolution superior to jets reconstructed only from clusters. PFlow jets are reclustered with the anti-k_T algorithm using a radius of $R = 0.4$ and to be used in this analysis they have to fulfil $p_T > 50\,\mathrm{GeV}$ and $|\eta| < 4.5$. These jets are used in the overlap removal and the determination of the missing transverse energy.

Unified Flow Object Jets

The Unified Flow Object (UFO) [67] was introduced by the ATLAS Collaboration as a jet object that performs well over the whole kinematic range by combining tracking information, PFOs, Topo Clusters and Track-CaloClusters [68] as input. An illustration of the algorithm is shown in Figure 5.7.

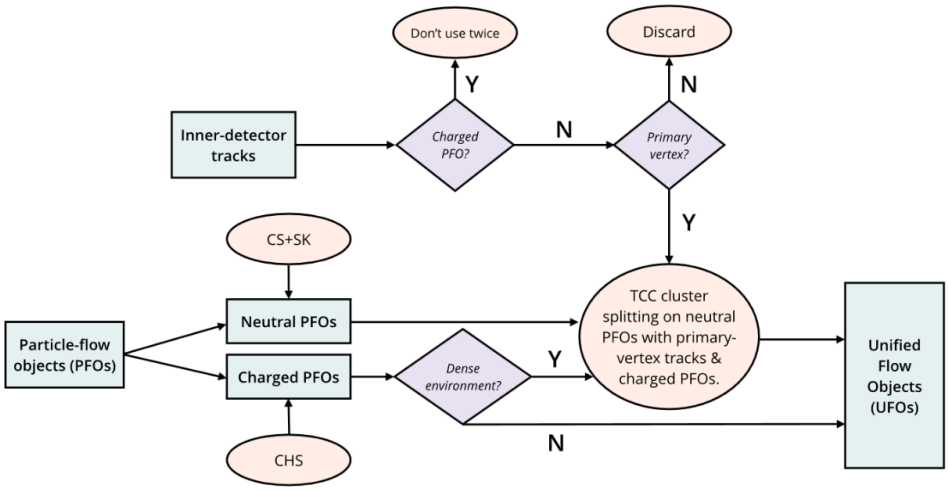


Figure 5.7.: Reconstruction algorithm of unified flow object jets. From Ref. [67].

UFO jets are large radius jets which collect all hadronic decay products of the boosted decay in a single jet and are reclustered with the anti-k_T algorithm using a radius of $R = 1.0$. To be used in this analysis, they need to pass the selection criteria of $p_T > 200\,\mathrm{GeV}$ and $|\eta| < 2.0$. UFO jets are used for the reconstruction of hadronically decaying analysis

objects such as the $H \to b\bar{b}$ and W_{had} candidate.

6. The Search for $X \to HH \to b\bar{b}WW^*$ in the 1-Lepton Final State with the boosted Topology

This Master's thesis focuses on the boosted 1-lepton togology in the search for resonant HH production in the $b\bar{b}WW^*$ decay channel. A schematic of the studied topology can be seen in Figure 5.1. As the final state particles are highly boosted, four objects are reconstructed in the event in general. These objects are the $b\bar{b}$-pair from the Higgs boson decay ($H \to b\bar{b}$), the hadronically decaying W boson (W_{had}) and the charged lepton (ℓ) with the corresponding neutrino (ν) from the leptonically decaying W boson. The event is characterised by the $H \to b\bar{b}$ and the W_{had} being reconstructed as single jet objects. In this thesis UFO jets are used, and the prompt lepton, which is selected as the signal lepton, is very close to the W_{had} or even overlapping with it. Since both the $H \to b\bar{b}$ and the W_{had} are reconstructed as the same type of jet object, a classification algorithm is developed to distinguish them by their properties.
This section first discusses and compares these classification methods. It also introduces a novel b-tagging tool used in this analysis. After that, the cut based signal region selection for this search is studied in detail, and in the end two different neural network based classifications are established and their performance is compared.

6.1. Analysis Objects

In this Master's thesis, two different jet classification methods are studied. The first one was already used in the corresponding Run 2 analysis and performed very well [62]. The second method uses the GN2x tagger [69], which is a novel b-tagging tool and under active development. The first classification method classifies the UFO jets by their distance to the signal lepton. The jet with the smallest distance ΔR to the signal lepton is labelled W_{had}. The leading p_T jet of the remaining jets is then labelled $H \to b\bar{b}$.
The second classification method makes use of b-tagging. b-tagging is a method of jet

flavour tagging to identify jets originating from a b-quark by looking for specific jet properties such as high p_T and secondary vertices. In Run 2, b-tagging was performed on individual small radius jets rather than large radius jets, such that the $H \to b\bar{b}$ candidate could be b-tagged multiple times, as the constituent jets coul be b-tagged independently. The GN2x tagger [69] performs tagging on the whole UFO jet rather than its constituents, meaning every UFO jet can only be tagged once. It is based on a Graph Neural Network (GNN) which assigns probability values p_{Hbb}, p_{Hcc}, p_{top} and p_{qcd} to each UFO jet in the input file. From these probabilities, the D_{Hbb} discriminant can be calculated, which describes how likely a given UFO jet originates from the decay of a Higgs boson to a $b\bar{b}$-pair. The D_{Hbb} discriminant is defined as

$$D_{\text{Hbb}} = \ln\left(\frac{p_{\text{Hbb}}}{f_{\text{Hcc}} \cdot p_{\text{Hcc}} + f_{\text{top}} \cdot p_{\text{top}} + (1 - f_{\text{top}} - f_{\text{Hcc}}) \cdot p_{\text{qcd}}}\right), \tag{6.1}$$

where $f_{\text{Hcc}} = 0.02$ and $f_{\text{top}} = 0.25$. It can be seen from the formula that high values indicate a high probability of a given jet originating from $H \to b\bar{b}$ and low/negative values indicate a low probability. Standard b-tagging working points can then be matched to the values of the D_{Hbb} discriminant. Table 6.1 shows the preliminary discriminant values with the corresponding b-tagging working point.

Working Point	50 %	55 %	60 %	65 %	70 %	75 %	80 %	85 %
Cut Value on DHbb	4.335	4.087	3.818	3.518	3.166	2.735	2.211	1.560

Table 6.1.: Standard b-tagging working points with corresponding preliminary D_{Hbb} score [69].

The D_{Hbb} discriminant values for all UFO jets in all signal events is displayed in Figure 6.1. It can be seen that most jets in the events are unlikely to be $H \to b\bar{b}$ jets. It also shows that the lowest mass signal sample is shifted slightly towards lower values with respect to the other two samples. That might be because this sample contains the least boosted events which might not be boosted enough for the GNN to perform as well as for the other two samples.
The UFO jet classification based on the D_{Hbb} score first searches for the UFO jet with the highest D_{Hbb} score and labels that one $H \to b\bar{b}$. Afterwards the jet with the smallest distance ΔR to the signal lepton is labelled W_{had}.

A quick comparison of the two jet classification methods was performed with respect to the masses of the $H \to b\bar{b}$ and W_{had} candidate. Figure 6.2 displays the W_{had} jet mass and the $H \to b\bar{b}$ jet mass for both classification methods on the $m_X = 3\,\text{TeV}$ mass point. Both classification methods perform well. For the D_{Hbb} score based classification, the

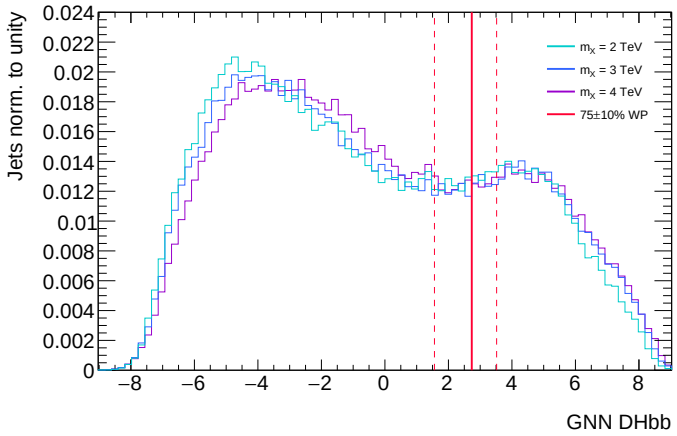
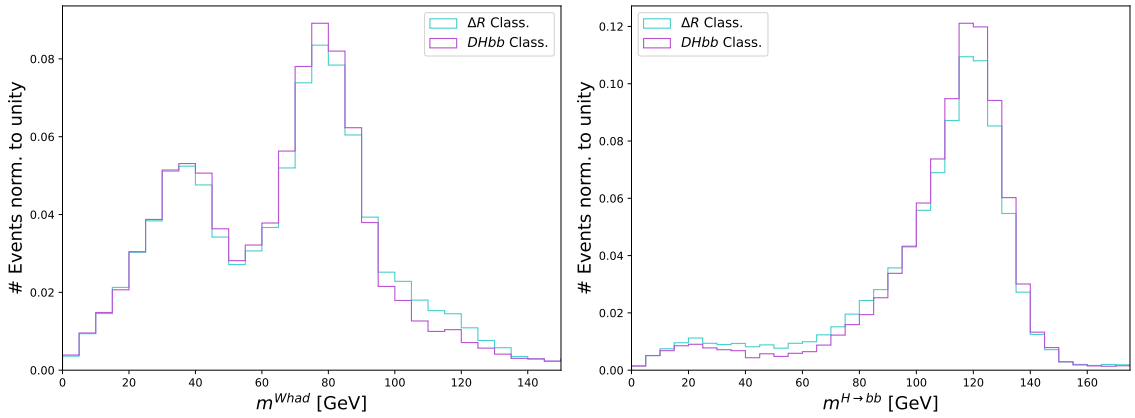


Figure 6.1.: D_{Hbb} discriminant score for all UFO jet in all signal events.

mass distributions for both jet objects are more focused on the W and Higgs boson mass respectively. For the distance based classification, both distributions have higher tails and a less pronounced peak. Nonetheless, the distanced based classification method is used for further studies, as it was also used in the corresponding Run 2 analysis [62] and because the GN2x tagger is under active development, meaning that recommendations regarding the b-tagging working points can change.



(a) distance class. **(b)** GNN class.

Figure 6.2.: Mass distributions of the two jet objects for both classification methods.

The efficiency of the GN2x b-tagging working point values was also studied briefly. For this study, the distance based classification is applied. The results of this study are shown in Figure 6.3. In general, the efficiencies behave as expected with the loosest b-tagging working point having the highest tagging efficiency and the tightest working point having the lowest efficiency. It can also be seen that the efficiencies slightly depends on m_X as was also noticed for the D_{Hbb} discriminant distribution.

However, all efficiencies are a little lower as would be expected. Usually, the percentage of the working point corresponds to its efficiency. The reason for this could be that the GNN is trained on pure $H \to b\bar{b}$ events and not the $X \to HH \to b\bar{b}WW^*$ decay studied in this thesis. This can also lead to some of the $H \to b\bar{b}$ jets not being classified as $H \to b\bar{b}$ candidates and conversely, other jets being classified as the $H \to b\bar{b}$ candidate.

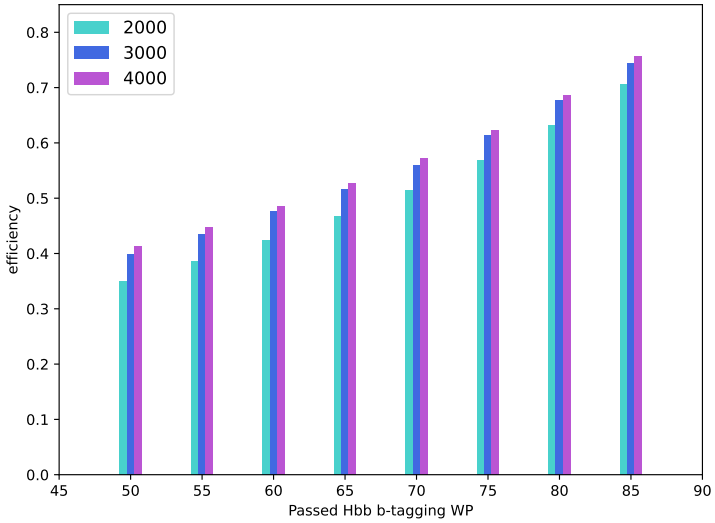


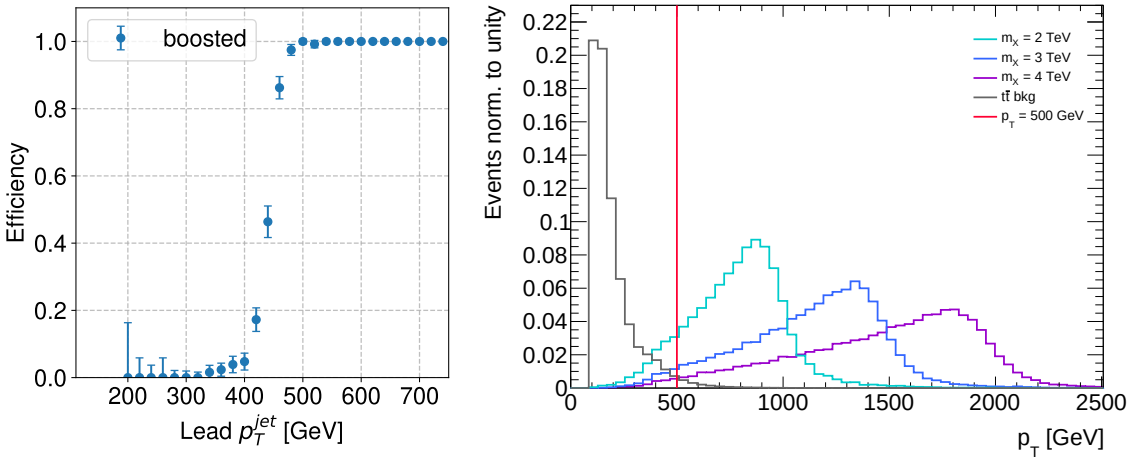
Figure 6.3.: Passed b-tagging working points for $H \to b\bar{b}$ jet.

6.2. Preselection

Based on the event topology and the analysis objects used, a preselection of the signal events is defined. The impact of these cuts is examined in this section with respect to the three signal samples and the $t\bar{t}$ background sample, which is expected to be the leading background contribution. The cuts are based on the corresponding Run 2 analysis [62] but modified to fit the structure of the Run 3 analysis.

The first cut applied requires $p_T > 500\,\text{GeV}$ for the leading UFO jet to ensure it is well within the trigger plateau of the UFO jet trigger. An illustration of the dependence of the trigger dependence on the transverse momentum can be seen in Figure 6.4a. The trigger efficiency with respect to the p_T of the leading jet is shown. The trigger reaches full efficiency around $p_T \gtrsim 480\,\text{GeV}$. The effect of this cut on the signal samples as well as on a $t\bar{t}$ background sample can be seen in Figure 6.4b. The figure shows the distribution of the leading jet p_T for all three signal samples and a $t\bar{t}$ background sample. The red vertical line indicates the cut value. This first cut already removes most of the

$t\bar{t}$ background events while keeping most of the signal events.



(a) UFO jet trigger efficiency **(b)** Leading jet p_T

Figure 6.4.: Leading jet p_T distributions for all signal mass points and $t\bar{t}$ background. The vertical red line indicates the cut value. The UFO jet trigger efficiency motivates the cut value.

The next two cuts are based on the event topology. To pass these cuts, there need to be at least two UFO jets in the event to be able to classify them as $H \to b\bar{b}$ and W_{had} and exactly one signal lepton.

To force the boosted topology, the next cut is placed on the distance between signal

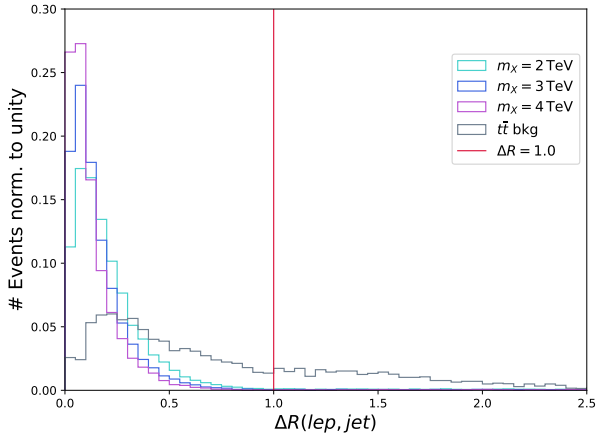


Figure 6.5.: ΔR distributions between the signal lepton and the closest jet for all signal mass points and $t\bar{t}$ background. The vertical red line indicates the cut value.

lepton and the closest UFO jet being $\Delta R(\ell, \text{closest jet}) < 1.0$. In Figure 6.5, the effect of this cut on the ΔR distribution between the signal lepton and the closest jet after passing all previous cuts is shown. As the $t\bar{t}$ background was already reduced drastically by the

cut on the leading jet p_T, the reduction is not as significant for this cut. Although, it cuts away almost no signal events while remove some more background events. The next

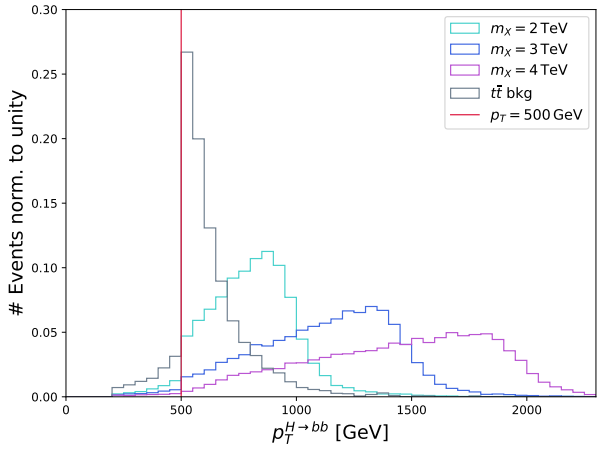


Figure 6.6.: $H \to b\bar{b}$ jet p_T distributions for all signal mass points and $t\bar{t}$ background. The vertical red line indicates the cut value.

cut requires that the $H \to b\bar{b}$ jet fulfils $p_T > 500\,\text{GeV}$. The resulting p_T distribution is displayed in Figure 6.6 where the red line indicates the cut value again. The step in the distribution results from the first cut, as the $H \to b\bar{b}$ jet is often the leading p_T jet.

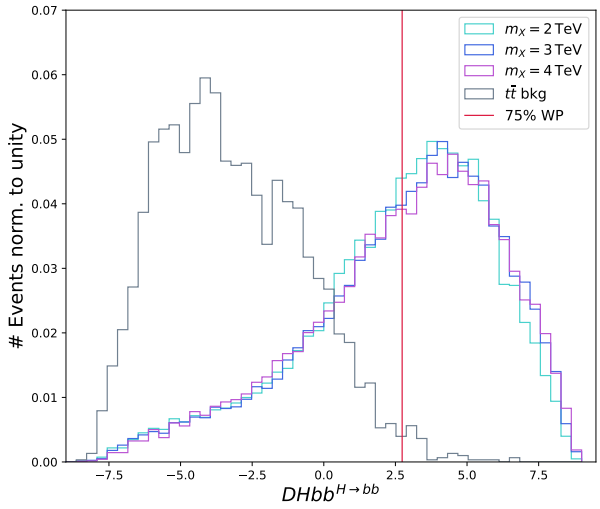


Figure 6.7.: D_{Hbb} score distributions for all signal mass points and $t\bar{t}$ background. The vertical red line indicates the cut value.

For passing the last cut, the $H \to b\bar{b}$ jet needs to pass a certain working point of the GN2x tagger. Currently, the 75 % working point is chosen as it has a similar signal efficiency as the b-tagging working point used in the Run 2 analysis [62]. The effects of this cut are

shown in Figure 6.7 which displays the distribution of the D_{Hbb} score. The distributions for signal and $t\bar{t}$ background look drastically different. The signal distributions are mostly positive while the $t\bar{t}$ background distribution is shifted to the negative values. This cut removes almost all of the remaining $t\bar{t}$ background events while keeping a good amount of signal events.

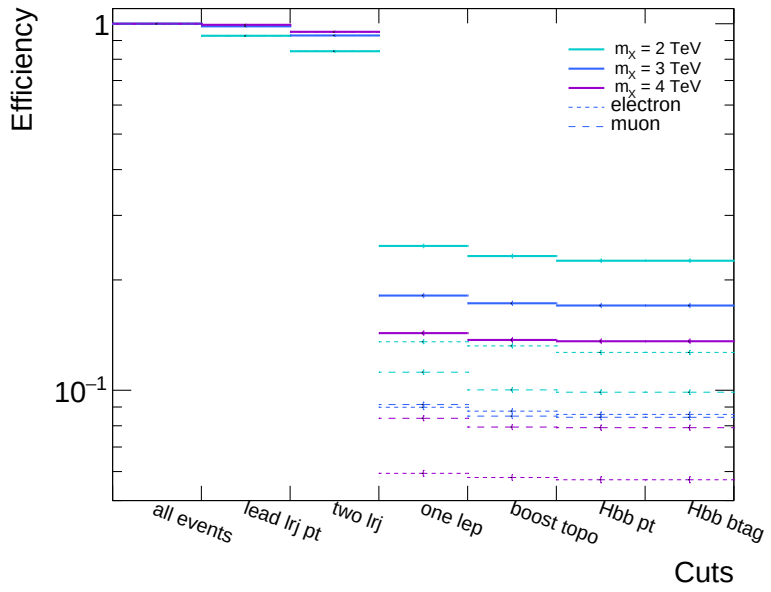


Figure 6.8.: Signal event efficiencies after successively applying all the cuts. The solid lines show the efficiencies for electron and muon events combined, while the different dashed lines represent electron and muon events individually.

In Figure 6.8, the event efficiencies after successively applying all the cuts is shown. The solid lines show the event efficiencies for electrons and muons combined, while for the dashed lines, electron and muons are considered individually. Currently, the biggest loss in signal efficiency is in the requirement of exactly one lepton. The reason for this are the identification and isolation efficiencies for electrons and muons. For electrons, most efficiency is lost when applying the medium ID cut (see Figure 5.2). For muons, the issue lies in the choice of isolation working point. As discussed in Section 5.3.1, the `TightTrackOnly_VarRad` isolation working point outperforms all others significantly, but that working point is not yet recommended for muon candidate selection. So instead the recommended `PflowLoose` isolation working point is used, which has a significantly lower efficiency for prompt muons (see Figure 5.5a). Obtaining the `TightTrackOnly_VarRad` isolation working point as a valid isolation criteria will improve the cut efficiency for muons drastically.

Assuming an integrated luminosity for Run 3 of $\int \mathcal{L} dt = 150\,\mathrm{fb}^{-1}$ and a signal cross

section of $\sigma = 10\,\text{fb}$, this rather simple baseline selection would yield 316 signal events for $m_X = 2\,\text{TeV}$, 245 signal events for $m_X = 3\,\text{TeV}$ and 199 signal events for $m_X = 4\,\text{TeV}$ for electron and muon events combined.

6.3. Signal & Background Classification using Neural Networks

After studying the cut-based selection, a machine learning based selection is investigated as well. To achieve this, appropriate input variables for the neural network training need to be found. Afterwards the neural network architecture, its optimisations and the training itself are discussed. Lastly the network performance is evaluated by testing it on individual signal mass points.

6.3.1. Neural Network Input Variables

For the neural network to be trained efficiently, proper input variables need to be defined. For that, various different variables were studied for all three signal and the $t\bar{t}$, W+jets and dijet background samples, to look for distributions that differ substantially for signal and background. As a guideline, the distributions of the variables, which were found to have a high impact on the training in a previous study [11] for Run 2 of the ATLAS experiment, are studied here as well. Apart from distributions of kinematic variables, distributions of jet substructure variables, as well as angular and radial distributions between different analysis objects and also b-tagging information are analysed.
For all possible input variables, the preselection from Section 6.2 is applied to the variables without the cut on the p_T and the b-tagging working point of the $H \to b\bar{b}$ candidate, as they reduce the background too drastically, not leaving enough statistics for the network to learn efficiently. The variable distributions are also weighted by taking into account different cross sections, lepton scale factors, generator weights and filter efficiencies of the different samples and dividing them by the number of weighted events. Weighting is important to compare distributions of different samples.

The variable with the highest impact on the training for Run 2 is the number of b-tagged jets $N^{\text{jets}}_{\text{b-tagged}}$. For Run 3, this information is not available, so the b-tagging of the $H \to b\bar{b}$ candidate is used instead. In Figure 6.9, the efficiency of the b-tagging working points for the $H \to b\bar{b}$ candidate are shown for all three signal mass points and $t\bar{t}$, W +jets and dijet background. As expected, The efficiency decreases for tighter working points. For the background samples, the b-tagging efficiency is significantly lower than for the signal, as was expected from the preselection study (see Figure 6.7), making this a suited variable for neural network training.

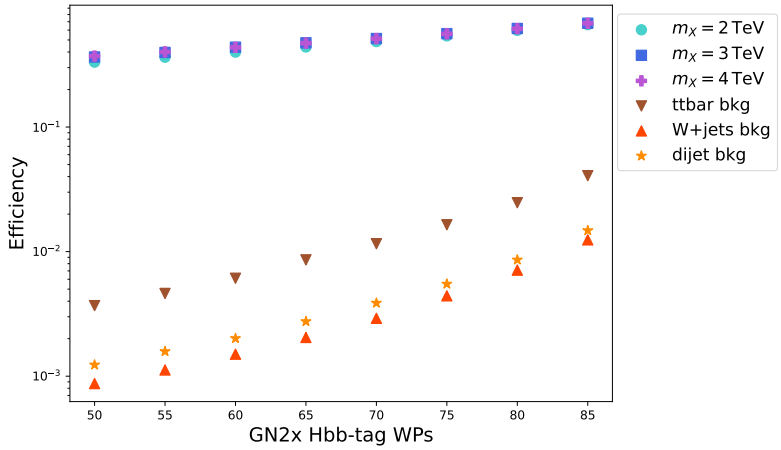
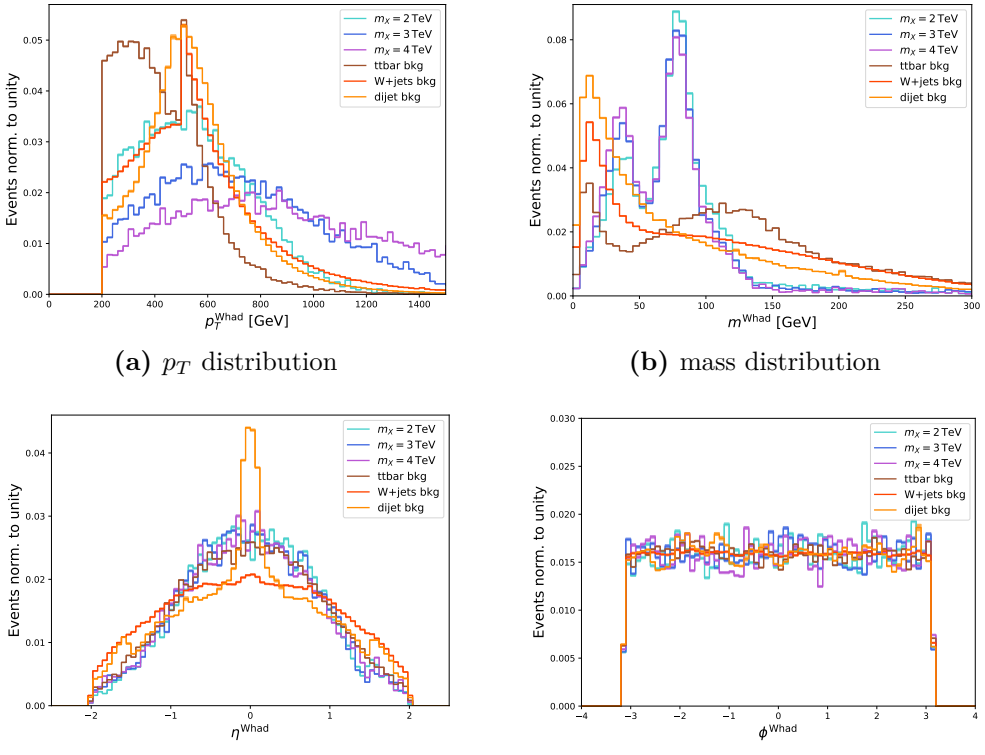


Figure 6.9.: $H \to b\bar{b}$ b-tagging working point efficiencies for signal and background.



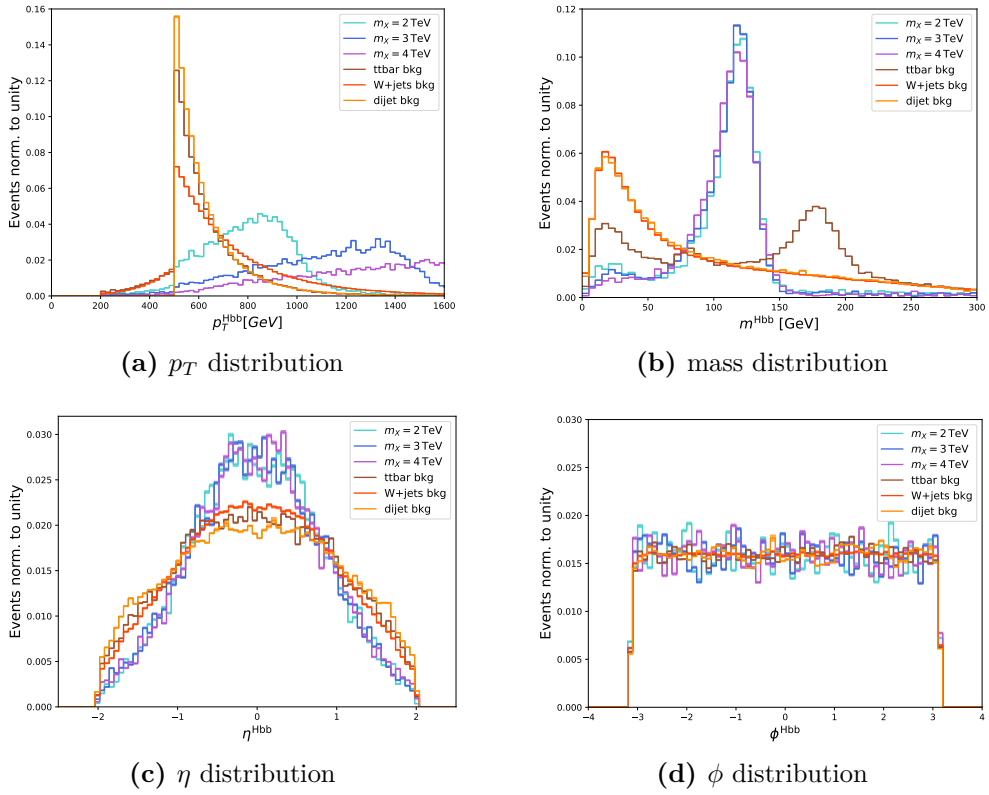
(a) p_T distribution **(b)** mass distribution

(c) η distribution **(d)** ϕ distribution

Figure 6.10.: Distributions of kinematic variables for the W_{had} jet candidates for signal and background.

Next, most of the distributions for the kinematic variables of the W_{had} candidate, the $H \to b\bar{b}$ candidate and the signal lepton performed well in Run 2. In Figure 6.10 and in Figure 6.11, the weighted distributions for p_T, m, η and ϕ for the W_{had} and for the $H \to b\bar{b}$ can be seen. The non-physical peak in the η spectrum for the W_{had} jet is discussed in

Section 6.3.2. For these distributions, the best separation between signal and background can be seen in the $H \to b\bar{b}$ mass distribution. There, a clear peak around the Higgs boson mass is observed for the signal samples. The $t\bar{t}$ background sample features a peak around the top quark mass and a second peak at low masses. For both the $W+$jets and the dijet background the mass distribution peaks at low values.



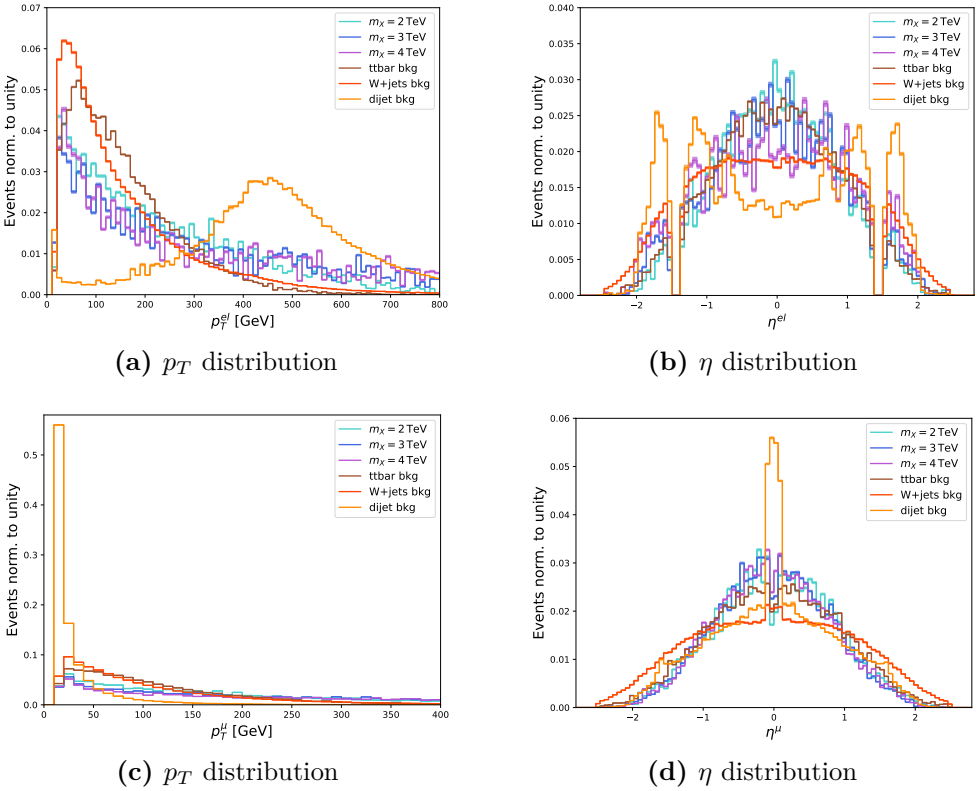
(a) p_T distribution **(b)** mass distribution

(c) η distribution **(d)** ϕ distribution

Figure 6.11.: Distributions of kinematic variables for the $H \to b\bar{b}$ jet candidates for signal and background.

Other useful kinematic variables are $p_T^{W_{\text{had}}}$, $p_T^{H \to b\bar{b}}$ and $m_{W_{\text{had}}}$, where the background and signal distributions also differ in shape.

The distributions of the kinematic variables for the signal lepton candidate can be found in Figure 6.12, for both lepton flavours separately. Looking at the p_T distributions for both lepton types, the dijet background contributions are distinctly different for both lepton types. The extreme peak of the dijet background towards zero stems predominantly from the non-prompt muons. For electrons this effect is not as drastic, but also visible. There, the p_T distribution is dominated by fake electron candidates. The distribution is shifted to much higher values due to electron candidates actually being jets.



(a) p_T distribution **(b)** η distribution

(c) p_T distribution **(d)** η distribution

Figure 6.12.: Distributions of kinematic variables for the signal lepton for signal and background.

In the η distributions, the dijet background contributions for both lepton flavours also look very different. An explanation for the non-physical nature of the muon η distribution in the dijet sample is given in Section 6.3.2.

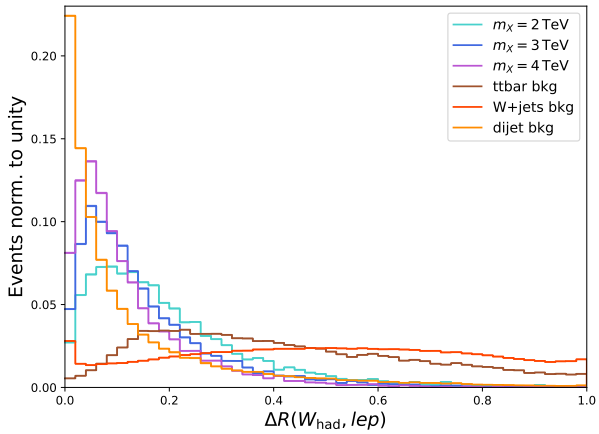
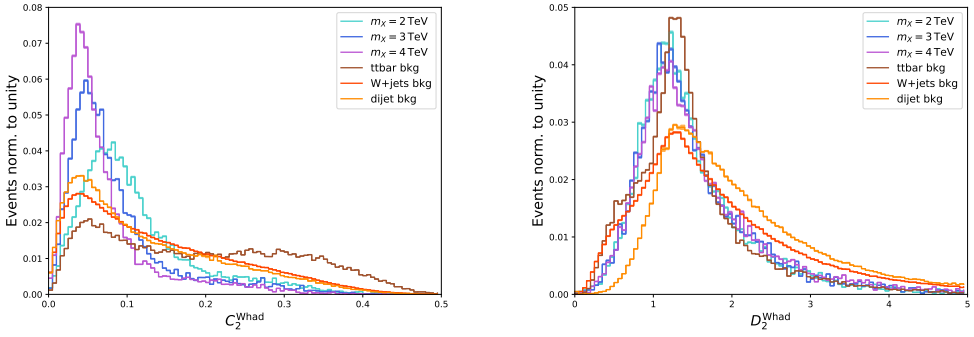


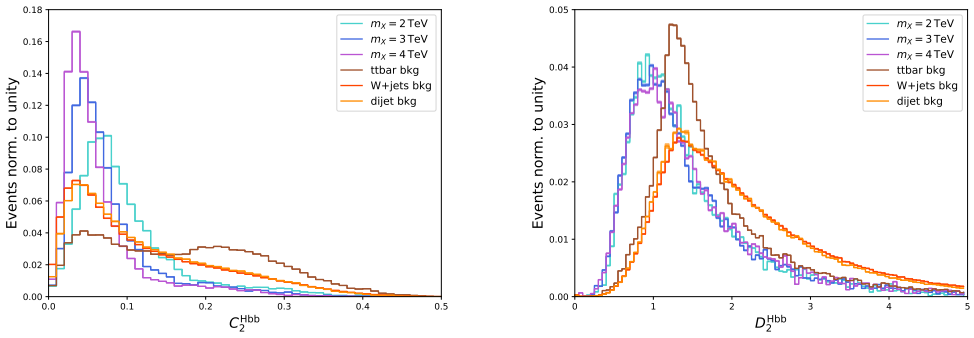
Figure 6.13.: Distribution of the distance between the signal lepton and the closest jet for signal and background.

For Run 2, the distance ΔR between the signal lepton and the closest jet was also an important variable for training. In Figure 6.13, the distributions for signal and background can be seen. The signal distributions peak towards lower values, showcasing the boosted topology. The highest signal mass point is the most boosted and thus, the peak is most pronounced. The dijet background distribution shows similar behaviour to the signal. This occurs due to misidentification of either a jet or non-prompt lepton as a signal lepton because they were very close to another jet. For the $t\bar{t}$ and the $W+$jets background, the distributions are very flat.



(a) C_2 distribution **(b)** D_2 distribution

Substructure variables for W_{had}



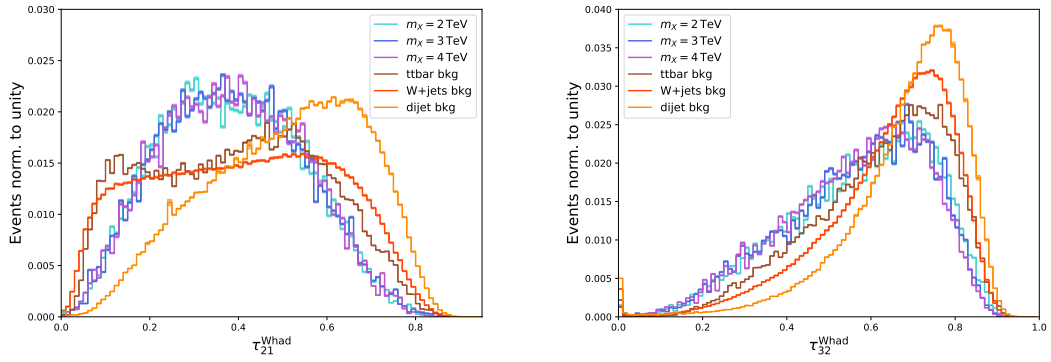
(c) C_2 distribution **(d)** D_2 distribution

Substructure variables for $H \to b\bar{b}$

Figure 6.14.: Distributions of C_2/D_2 jet substructure variables for signal and background.

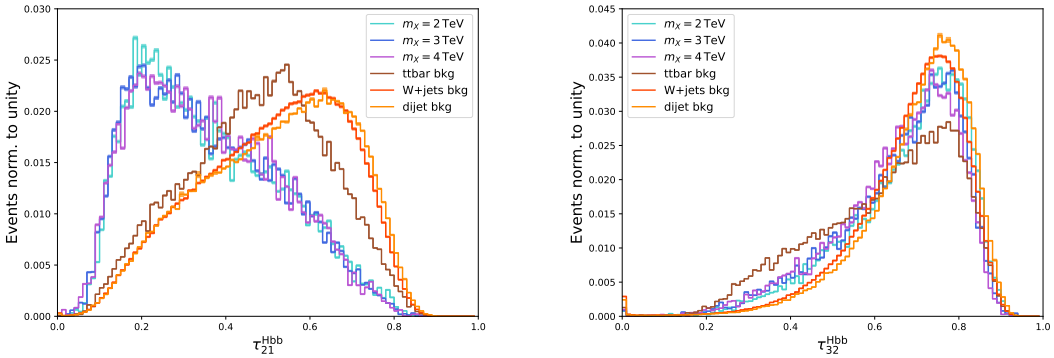
Lastly, in addition to the previously studied variables, jet substructure variables were analysed. Figure 6.14 shows the C_2 and D_2 jet substructure variables for the W_{had} and the $H \to b\bar{b}$ candidate. The distribution for C_2 of both jet candidates look quite similar for signal and background. The background distributions have a higher tail. The $t\bar{t}$ background distribution contains two separate peaks. The three signal distributions

also differ. For the D_2 distributions, the signal sample distributions agree well with one another, but the W_{had} distributions all peak around the same value. The D_2 distributions of the Hbb candidate for signal and background have a small separation between signal and background.



(a) τ_{21} distribution **(b)** τ_{32} distribution

Substructure variables for W_{had}



(c) τ_{21} distribution **(d)** τ_{32} distribution

Substructure variables for $H \to b\bar{b}$

Figure 6.15.: Distributions of n-subjettiness jet substructure variables for signal and background.

For the τ_{21} and τ_{32} substructure variables, seen in Figure 6.15, better separation is achieved. The τ_{21} distributions for both jet candidates look substantially different for signal and background, especially the $\tau_{21}^{H\to b\bar{b}}$ variable has a good separation between signal and background.

After performing a study of the important training variables for Run 2, as well as for the jet substructure variables, the following variables will be used for the training of the neural network:

- $H \to b\bar{b}$ passing 80 % b-tagging WP
- ΔR(lep,closest jet)
- τ_{21} for $H \to b\bar{b}$
- mass of W_{had} and $H \to b\bar{b}$
- η of W_{had}, $H \to b\bar{b}$ and lepton

6.3.2. Dijet Background η Study

During the study of the neural network input variables, unexpected behaviour showed for the η distributions of the analysis objects. Figure 6.16 shows the normalised η distributions for the signal lepton and $H \to b\bar{b}$ and W_{had} candidates for the signal as well as selected backgrounds. The η distributions for the $H \to b\bar{b}$ candidate look like they should for all signal and background samples. For the W_{had} and signal lepton however, the distribution of the dijet background peaks strongly around $\eta = 0$. For the signal lepton distribution, the peak is more pronounced than for the W_{had} distribution. This behaviour is non-physical and thus needed some investigation.

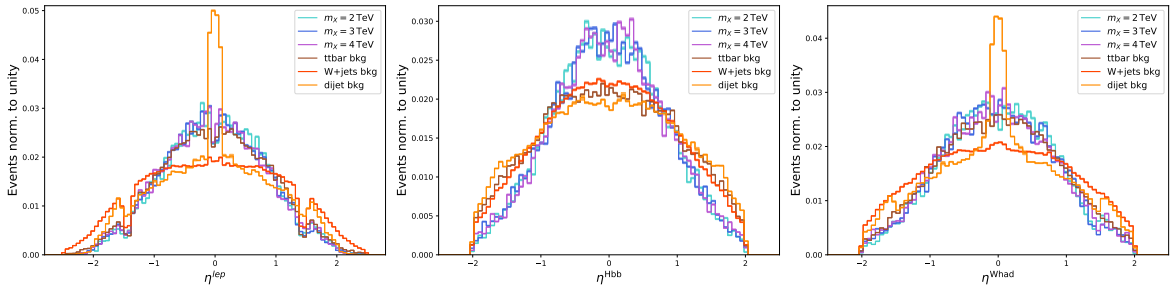
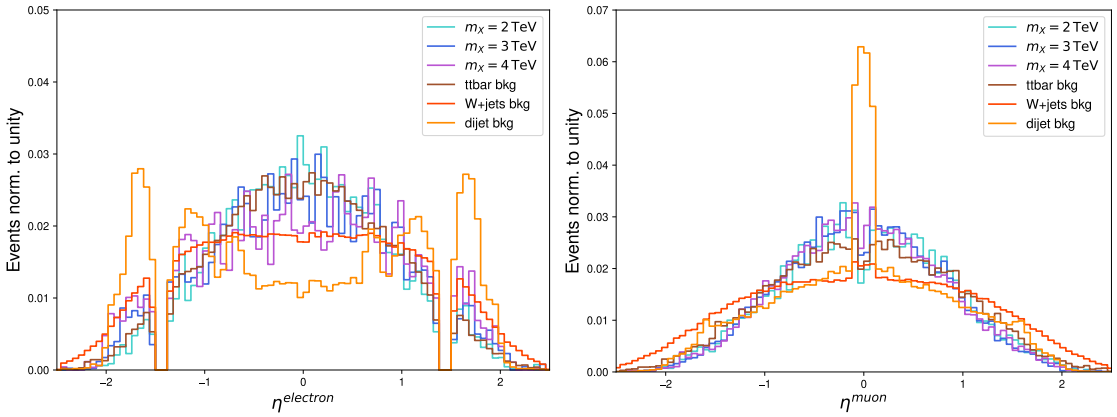


Figure 6.16.: Normalised η distributions for analysis objects for signal and main backgrounds.

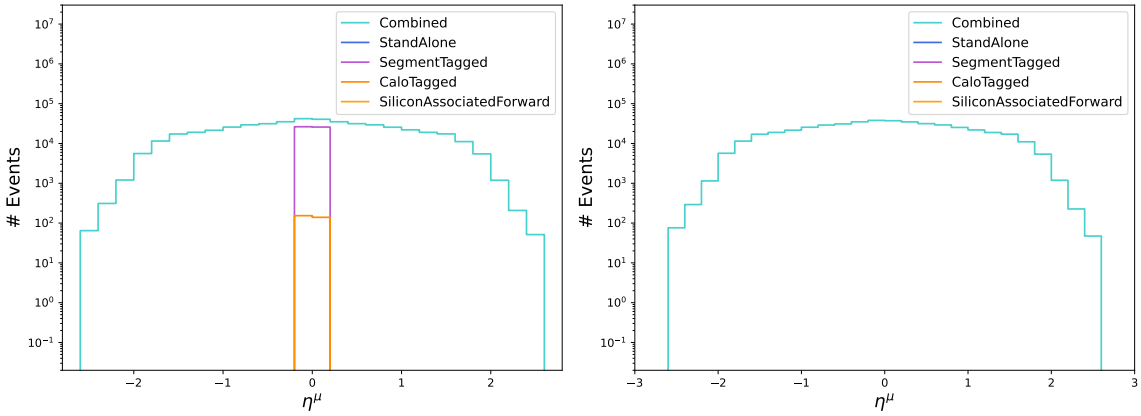
Looking at the signal lepton η distribution for electrons and muons separately, in Figure 6.17, it becomes clear that the problem is exclusive to muons. After some investigating, the issue originates from the reconstructed muon types in combination with the applied muon ID working point. In Figure 6.18, the η distributions for muons passing the `Loose` and `Medium` ID working points can be seen. When applying `Loose` ID to the muons, the segmenttagged muons have a significant influence on the η distribution. For `Medium` ID working point, only the combined muons remain, which are reconstructed from Inner Detector plus corresponding Muon Spectrometer track with extrapolation to the interaction point. Segmenttagged muons are reconstructed from a track in the Inner Detector if the



(a) electrons **(b)** muons

Figure 6.17.: Normalised η distributions for electrons and muons separately.

track matches a signal from a Muon Spectrometer segment. For the dijet events this can occur, when a high energy non-prompt muon is very close to a jet and the non-prompt muon signal is matched to the jet's track. The application of the `Loose` ID instead of



(a) Loose ID **(b)** Medium ID

Figure 6.18.: η distributions of reconstructed muon types for `Loose` and `Medium` ID muons on partial dijet background.

the `Medium` ID working point to the signal muons was responsible for this unexpected behaviour. Due to larger changes to the analysis framework to harmonise it across all analysed channels, this change in working point went unnoticed until the investigation of this issue. Due to lack of time it was not possible to repeat the neural network input variable study for this Master's thesis. Instead, the neural network training is performed with the available MC data. A veto is applied to all samples that have $|\eta_{\text{lep}}| < 0.1$ to prevent the learning of this non-physical behaviour.

6.3.3. Neural Network Setup & Training

For the machine learning based approach for the separation of signal and background, the results of a previous study [11] for Run 2 were reused and adjusted for Run 3.
A simple feed-forward neural network was trained on the data discussed previously. The network comprises of the input layer with 265 nodes, two hidden layers with 128 and 64 nodes respectively and the output layer with 2 nodes, which is considerably smaller than in previous studies, but for this study, the network is just supposed to classify two classes instead of four. An overview of the network architecture can be seen in Table 6.2.

Layer	Dimension
Input Layer	$\mathbb{R}^9 \rightarrow \mathbb{R}^{256}$
Hidden Layer 1	$\mathbb{R}^{256} \rightarrow \mathbb{R}^{128}$
Hidden Layer 2	$\mathbb{R}^{128} \rightarrow \mathbb{R}^{64}$
Output Layer	$\mathbb{R}^{64} \rightarrow \mathbb{R}^2$

Table 6.2.: Neural network architecture with dimensionality of each layer.

For training, the data is resampled using the `sklearn.utils.resample` function (see Appendix A.1.3) to have equal amounts of signal and background data for the network to train on. During training, 70 % of the data is visible to the network. The remaining 30 % of the data are reserved for testing and validation, with 20 % and 10 % respectively.

For training the model, the ADAM [40] optimiser is used with a learning rate of 0.001 and as loss function categorical Cross Entropy is used, because the model's task is to categorise the data into two classes. Due to a lack of time and expertise, only very rudimentary optimisation options are implemented. These are the learning rate schedulers `ReduceLROnPlateau` and `EarlyStopping` to prevent overfitting. There are many more optimisation options available to improve the network performance even further.

To later compare the model's performance, two different models were trained. One was trained on the 3 TeV signal sample and all three backgrounds and the other was trained on all three signal samples (2, 3 and 4 TeV) and all three background samples. Both models were trained over 50 epochs. The training history in terms of loss and accuracy for the model trained on only the 3 TeV signal sample can be seen in Figure 6.19. It can be seen that the loss rapidly decreases over the first few epochs, while the accuracy increases. After this steep decrease, the loss stagnates until the learning rate scheduler is triggered

and reduces the learning rate of the optimiser, resulting in the step in the loss function. The same behaviour is observed in the accuracy. After this step, the loss continues to decrease until the training is finished.

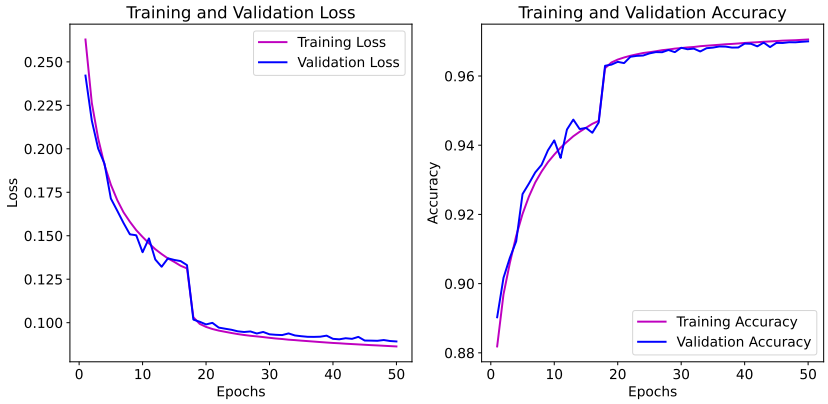


Figure 6.19.: Training and validation loss and accuracy curves for the single mass point trained model.

The same plots can be seen for the model trained on all signal samples in Figure 6.20. The overall behaviour of loss and accuracy is the same here, but the validation loss and accuracy are more erratic, pointing towards the model not being well optimised. Both loss and accuracy do not reach values as good as for the training on one signal sample. The total training metrics for both models can be seen in Table 6.3.

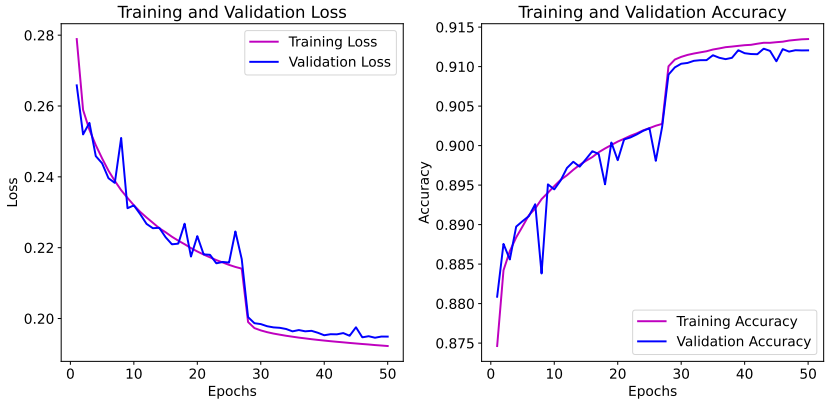


Figure 6.20.: Training and validation loss and accuracy curves for the multi mass point trained model.

6.3.4. Neural Network Performance Evaluation

After training the single mass point and the multi mass point neural networks, both are evaluated on all three signal mass points individually. The neural networks output class

	Single mass point model	Multi mass point model
Total training loss	0.0858	0.1922
Total training accuracy	0.9708	0.9133

Table 6.3.: Total training loss and accuracy for both trained models.

probabilities for the signal and background class for each tested event. The events are classified by the maximum predicted class probability. For each model, the distributions of the prediction probabilities for signal and background class, the ROC curve and the confusion matrix are analysed for each mass point.

Single Mass Point neural network

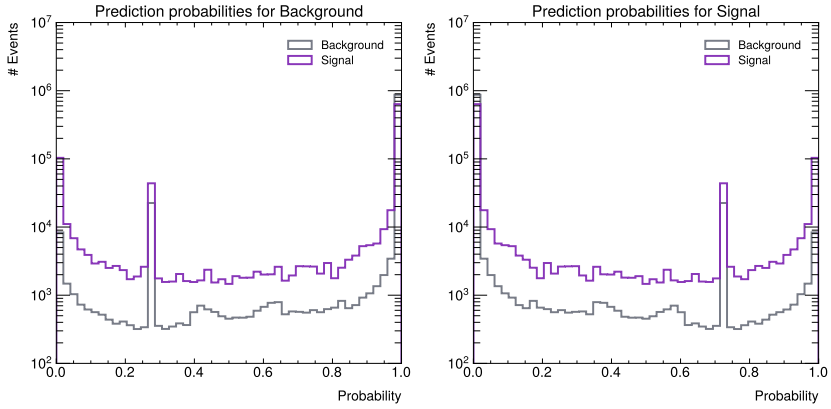
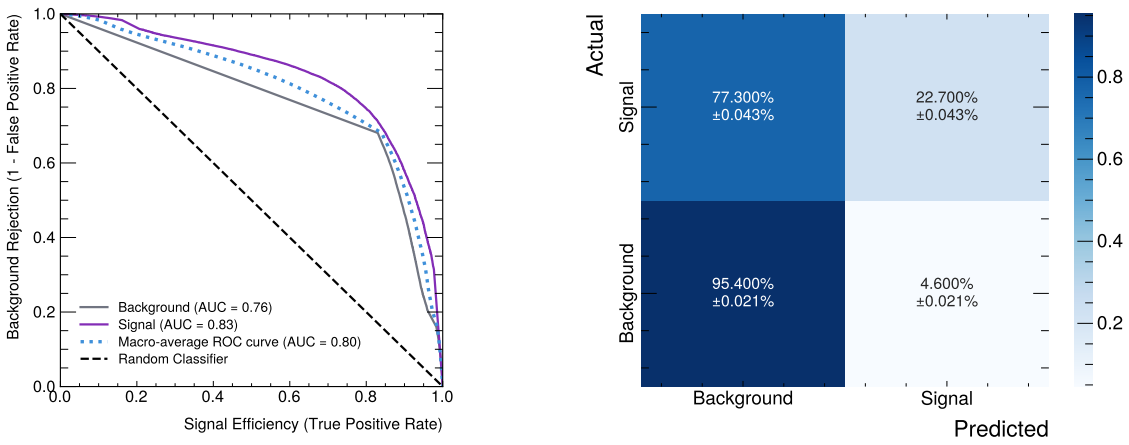


Figure 6.21.: Prediction probabilities for signal and background classes for the one mass point evaluated at 2 TeV. The colours indicate the true classes.

First, the single mass point network, trained on the 3 TeV mass point, is evaluated for the 2 TeV signal mass point. The prediction probability distributions for the two actual classes can be seen in Figure 6.21. The colours indicate the true class of the respective event. For this mass point, the model fails to correctly classify into signal and background. For this test, the model often classifies signal events as background. It also misidentifies a considerable amount of background events as signal events. The peaks around 0.3 and 0.7 are remnants of the fact that the network is not well optimised. With proper optimisation they should vanish. These peaks occur for both models and all tested samples.
Looking at the ROC curve and the confusion matrix for the 2 TeV test in Figure 6.22, the same behaviour is observed. Due to the low background rejection for this tested sample, the Area-under-curve (AUC) for the background class is only at 0.76, while for the signal class it is 0.83. The macro-average AUC, which takes the average of the two classes, is 0.80. For the confusion matrix, it shows that the model is still good at classifying



(a) ROC curve **(b)** confusion matrix

Figure 6.22.: Evaluation of the one mass point model with the 2 TeV mass point.

background as background. The model also tends to misidentify the 2 TeV signal as background, which could be expected, as the input variable distributions for this mass point are more background-like than for the other signal mass points.

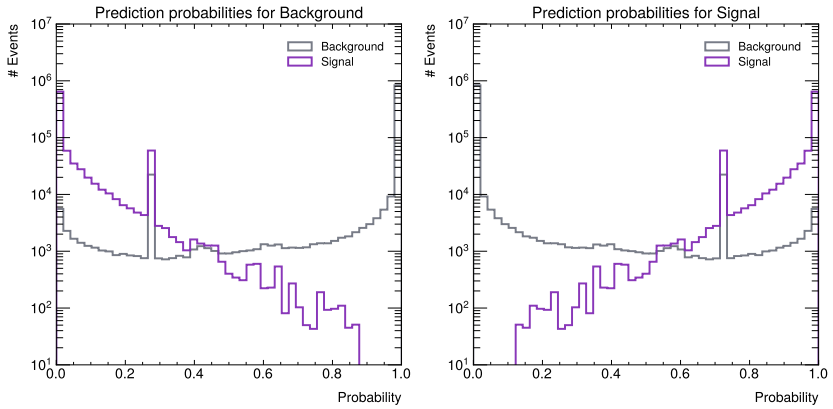
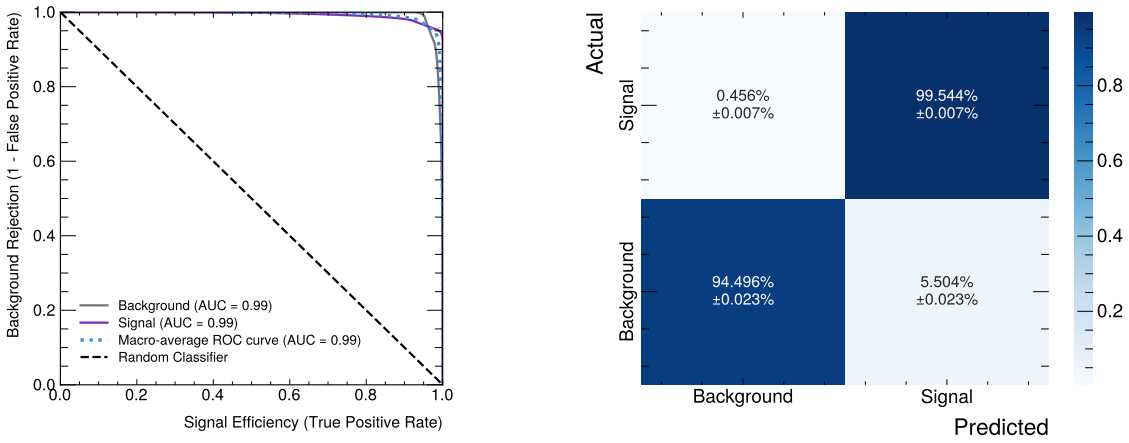


Figure 6.23.: Prediction probabilities for signal and background classes for the one mass point evaluated at 3 TeV. The colours indicate the true classes.

Next, the 3 TeV mass point was evaluated on the one mass point model. The distributions for the signal probabilities can again be seen in Figure 6.23. This is the mass point the network was trained on and thus it performs significantly better. Looking at the plot for the background prediction probability, the network does not output prediction probabilities above 88 % for true signal events. It also classifies the signal class with signal prediction probability values of at least 12 %. The behaviour for the true background class

is also good, but here the network still misidentifies some background events as signal. The overall distribution of prediction probabilities for true background events looks similar to the 2 TeV test.



(a) ROC curve **(b)** confusion matrix

Figure 6.24.: Evaluation of the one mass point model with the 3 TeV mass point.

The ROC curve and the confusion matrix in Figure 6.24 agree with the prediction probability distributions. The AUC for both classes is 0.99 and so is the average, which is a very good result. The confusion matrix shows that only about 0.5 % of signal events are falsely identified as background events. For the background class, approximately 5 % of actual background events are falsely identified as signal events.

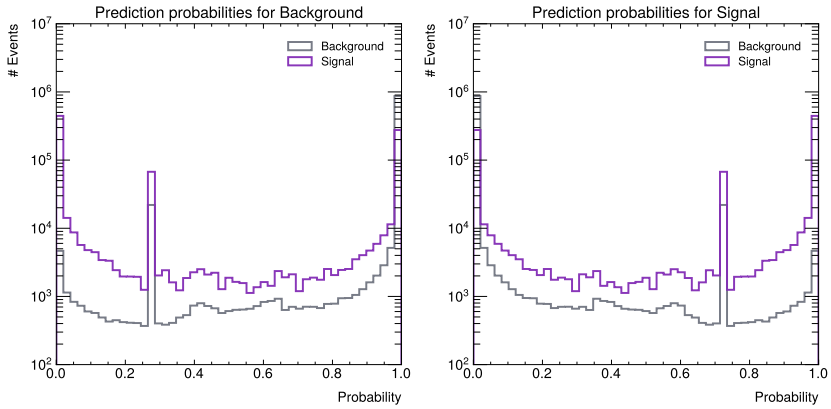
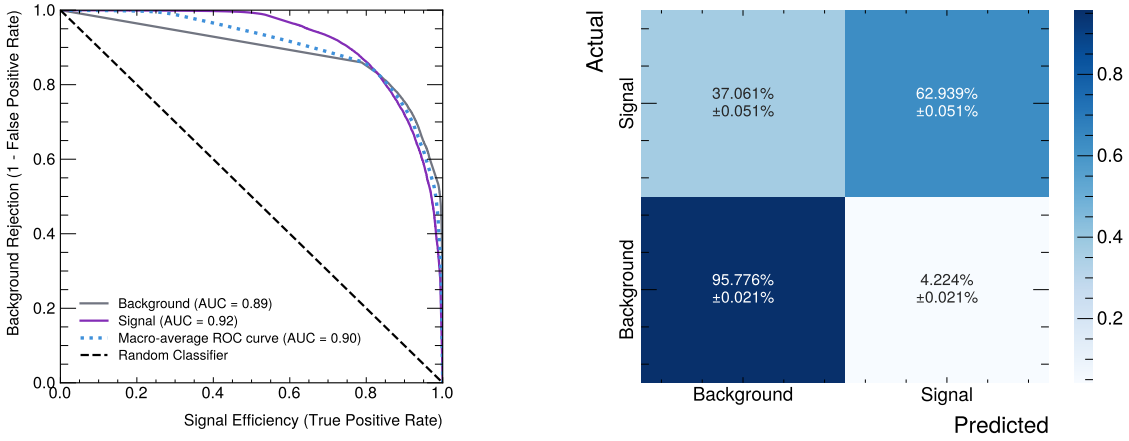


Figure 6.25.: Prediction probabilities for signal and background classes for the one mass point evaluated at 4 TeV. The colours indicate the true classes.

Lastly for the one mass point model, the model was evaluated with the 4 TeV mass point.

Figure 6.25 shows the prediction probabilities. For this mass point the performance worsened compared to the 3 TeV mass point, but is not as bad as for the 2 TeV mass point. Again, the general shape of the distribution for true background events looks similar as before. The model does a better job at identifying the background class than the signal class. For the background predictions, a considerable amount of true signal events have high prediction probabilities. The separation of true signal and background increases towards higher values of the signal prediction probability.



(a) ROC curve **(b)** confusion matrix

Figure 6.26.: Evaluation of the one mass point model with the 4 TeV mass point.

Figure 6.26 shows the ROC curve and the confusion matrix for the 4 TeV mass point evaluation. Here, the AUC for the background class is with 0.89 lower than for the signal class with 0.92, because the background rejection is low. Looking at the confusion matrix it can be seen that the background class is well identified. The signal classification efficiency is $\approx 63\,\%$.

As expected, the one mass point neural network performs well for the mass point it was trained on and considerably worse on the other two mass points.

Multi Mass Point neural network

The multi mass point network was trained on all three signal mass points simultaneously, and is evaluated on each signal mass point individually.

The multi mass point model is first evaluated with the 2 TeV mass point. The prediction probabilities for this evaluation can be seen in Figure 6.27. Similarly to the one mass

point model, this model performs worst for this mass point. Also, very similar peaks around 0.3 and 0.7 show up in the prediction probabilities for this model as well, again stemming from the lack of model optimisation. The trend of favouring the background class continues here as well, as both actual signal and actual background events have high background class prediction probabilities. For the prediction probability of the signal class, the separation of true signal and background increases marginally towards higher values.

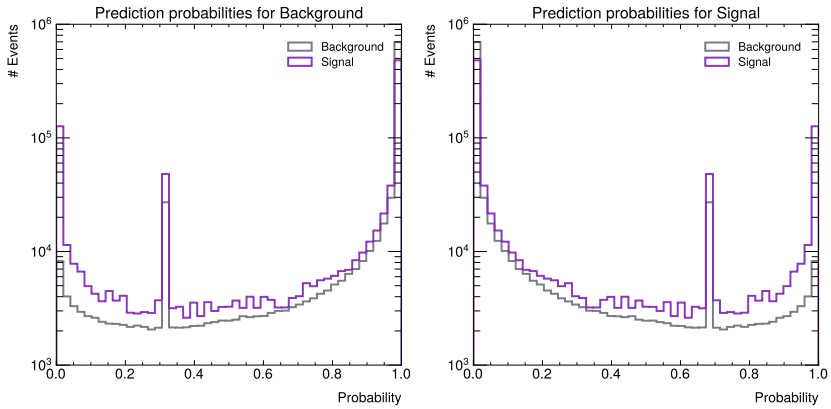
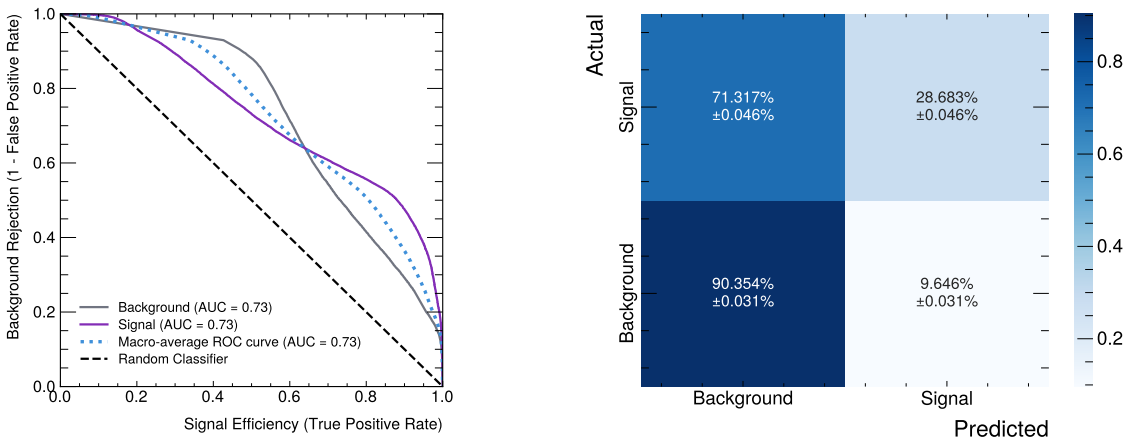


Figure 6.27.: Prediction probabilities for signal and background classes for the multi mass point evaluated at 2 TeV. The colours indicate the true classes.

The ROC curve and confusion matrix in Figure 6.28 confirm the poor performance. The AUC for the average of both classes is only at 0.73, which is the worst performance out of all tests. In the confusion matrix it can be seen that more than 70 % of the signal events are misidentified as background events. The reason for this even worse performance of the multi mass point model is two fold. First of all, the 2 TeV sample is the least boosted out of all three signal samples and thus the input variable distributions look most similar to the background out of the three. Second, the network architecture might not be deep and wide enough for the model to fully learn all three signal samples.
Next, the 3 TeV signal sample is tested on the multi mass point model. The corresponding prediction probabilities can be seen in Figure 6.29. The predictions are significantly better for this mass point. The predictions for the true signal events are better that for the true background. Nonetheless, these results are not as good as for the one mass point model. There still remain ambiguities for this model.
This result is reflected in the ROC curve and the confusion matrix for the 3 TeV mass point evaluation as shown in Figure 6.30. The macro-average AUC is 0.97, which is only slightly worse than for the one mass point model for the 3 TeV mass point. The confusion matrix also looks comparably good. The signal classification efficiency is at roughly 93 %,



(a) ROC curve **(b)** confusion matrix

Figure 6.28.: Evaluation of the multi mass point model with the 2 TeV mass point.

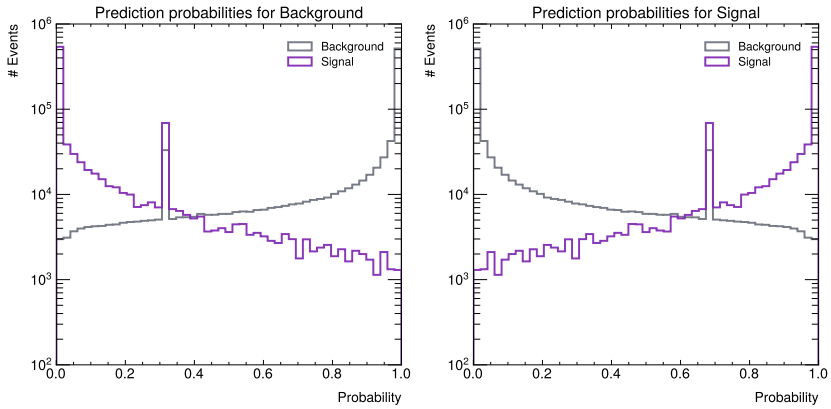
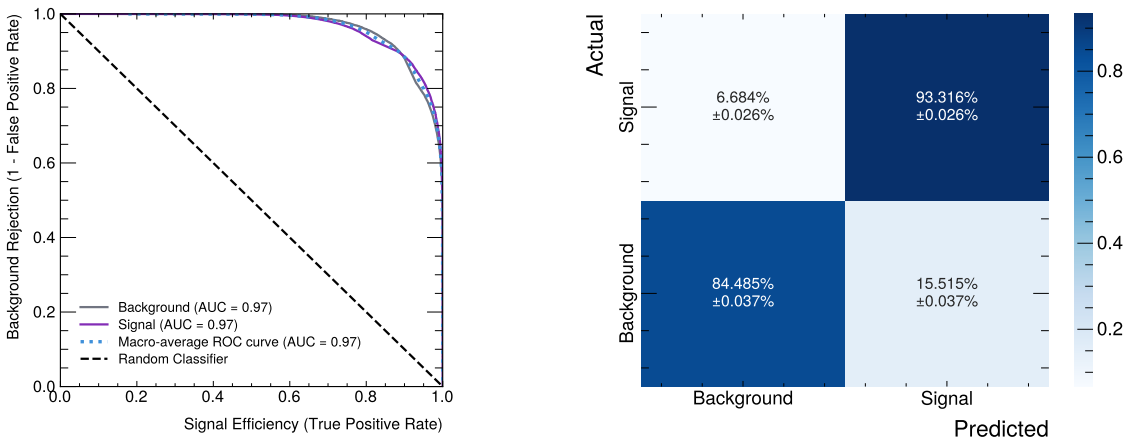


Figure 6.29.: Prediction probabilities for signal and background classes for the multi mass point evaluated at 3 TeV. The colours indicate the true classes.

while the background classification has an efficiency of almost 85 %.

Finally, the multi mass point model is tested with the 4 TeV signal mass point. For this mass point test, better results than for the one mass point model are expected, as the multi mass point model is familiar with $m_X = 4\,\text{TeV}$ events from training. The expectation is met, as can be seen in the prediction probabilities in Figure 6.31. Even though the model again performs better at identifying the background class than at identifying the signal class, the separation between true background and signal increases significantly towards higher signal prediction probability values.

This result becomes more clear when looking at the ROC curves and the confusion matrix in Figure 6.32, where the AUC for both classes is at 0.92, a value comparable to the 3 TeV testing. For the confusing matrix however things are a little worse. Before, the model



(a) ROC curve **(b)** confusion matrix

Figure 6.30.: Evaluation of the multi mass point model with the 3 TeV mass point.

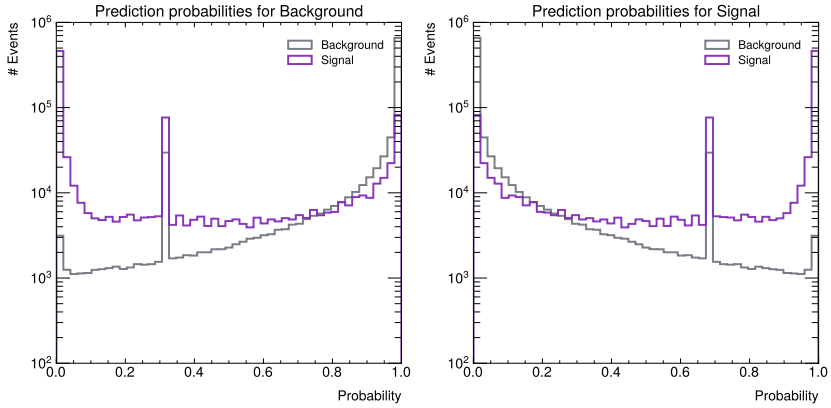
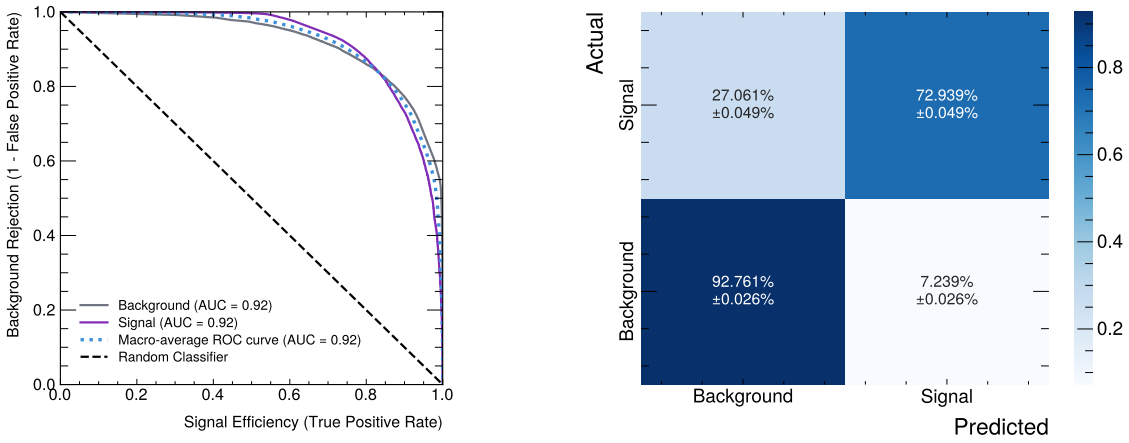


Figure 6.31.: Prediction probabilities for signal and background classes for the multi mass point evaluated at 4 TeV. The colours indicate the true classes.

was better at classifying signal correctly, now it is better at classifying background. The Background classification efficiency is at 93 % efficiency and signal classification at only 73 % efficiency. One possible reason for this could be, that the model confuses the 4 TeV signal with the dijet background. Another reason could again be that the chosen network architecture is not appropriate for the multi mass point model.

In Table 6.4, the test results are summarised for both models and each tested mass point. The overall best performance was achieved by the one mass point model when testing on the trained mass point. Similarly good results are achieved for the multi mass point model for the same mass point. The multi mass point model also performed well on the 4 TeV mass point. Both models had trouble identifying the signal events in the 2 TeV signal sample, as it contains the least boosted events.



(a) ROC curve **(b)** confusion matrix

Figure 6.32.: Evaluation of the multi mass point model with the 4 TeV mass point.

For the multi mass point model better results could potentially be achieved for all mass point when the neural network has more hidden layers and nodes per layer. For both models proper optimisation as well as increased training time could also improve the results. For the multi mass point model, larger improvements across all signal samples are expected than for the single mass point model, hence the focus of this analysis should be on the improvement and advancement of the multi mass point model.

	One mass point model			Multi mass point model		
Test mass point	2 TeV	3 TeV	4 TeV	2 TeV	3 TeV	4 TeV
Average AUC	0.80	0.99	0.90	0.73	0.97	0.92
Signal classification efficiency	0.23	0.99	0.63	0.29	0.93	0.73
Bkg classification efficiency	0.95	0.94	0.96	0.90	0.84	0.93

Table 6.4.: Results of the tests on the different mass points of the two models.

7. Conclusion and Outlook

This Master's thesis presents the ongoing search for resonant Higgs boson pair production in the using ATLAS Run 3 data. As the Higgs boson coupling strength is proportional to the mass of the particle, studying the Higgs sector with BSM extensions is a good possibility to search for new particles that do not couple to other SM particles.
The analysis focuses on the $b\bar{b}WW^*$ decay channel with one charged lepton in the final state. Due to the large mass of the scalar boson X, the two Higgs bosons from the decay are highly boosted. The boosted topology is characterised by one charged lepton and two large radius jets. One of them is the $H \to b\bar{b}$ candidate and the other the W_{had} candidate. The overlap between the lepton and the W_{had} candidate leads to a distinct signature that can easily be distinguished from most backgrounds.
The result of the cut-based selection presented in this thesis seems promising. Considering that the best performing isolation working point for muons is not yet recommended and thus cannot be used, improved signal event yields can be expected when the working point becomes recommended for muon candidate selection.
In the scope of this thesis, two neural networks were trained. The single mass point model was trained on only the 3 TeV signal mass point. while the multi mass point model was trained on all three signal mass points. Both neural networks perform well on the $m_X = 3\,\text{TeV}$ signal mass point and poorly on the $m_X = 2\,\text{TeV}$ mass point. The multi mass point network model outperforms the single mass point model for the $m_X = 4\,\text{TeV}$ mass point. Given proper optimisation, which was not possible in the scope of this thesis, and longer training time, it can be expected that both models achieve even better results. For the multi mass point model, improved results can be expected across all three mass points, whereas for the single mass point model the expected improvement would be exclusive to the 3 TeV mass point.
A final evaluation of the cut-based model against the neural network models can be made after the recommendation for the best performing muon isolation working point is obtained. The lack of that working point is responsible for the highest loss in efficiency in the cut-based model. Considering that a cut-based preselection is applied to the neural network input variables, this drastic loss in efficiency carries over to the neural network

training, impacting the efficiency.

The next steps for this analysis should be to first obtain the recommendation for the needed muon isolation working point by determining the corresponding muon scale factors. After that, the efficiencies for the cut-based model and the neural network models should be compared for the new ATLAS background sample recommendations, and with the correct ID and isolation working points, while applying overlap removal to all reconstructed leptons and jet objects. Then effort should be placed into optimising the neural networks, focusing on optimisation of the multi mass point model, as its expected improvement exceeds the improvement of the single mass point model. It could also be considered to set up two different neural networks, one for the high mass points and one for low mass, once more mass points are available for Run 3. If the neural network the outperforms the cut-based selection, as it was shown in Ref. [11] for Run 2, the neural network should be implemented in the analysis framework for Run 3 for signal and background event classification. This feature is currently in development, which is why this could not be done in the scope of this thesis.
It could also be interesting to train the neural networks independently for the two signal lepton flavours or to setup neural networks with more output classes as it was done for Run 2 [11] and see how this changes the network's performance.
While doing this analysis, new analysis tools were developed by the ATLAS collaboration, some of which should also be used for this analysis. It is now recommended to use the mass dependent GN2x tagger for b-tagging, as the mass independent version cannot handle dijet background very well, and thus cannot be calibrated. There are also some new algorithms for electron identification. A neural network based approach is currently being tested and could replace the likelihood method in the near future. It should be considered to adapt these changes for this analysis as well.
The results of this thesis provide a good basis for the ATLAS Run 3 boosted $HH \rightarrow b\bar{b}WW^*$ analysis with one charged lepton in the final state. A promising preselection was established and also the necessity of the `TightTrackOnly_VarRad` isolation working point for muons was discovered. The neural network based selection offers a new approach for this analysis, replacing the cut-based approach, and thus provides new opportunities in the search for resonant HH production.

A. Appendix

A.1. Technical Setup

The technical setup of the analysis presented in this thesis is described in this section. The samples used in this analysis are derived xAOD files (DAOD_PHYS), in which events that are not interesting and objects not passing certain criteria are removed from the file. This derivation also adds information and object collections, like specific jet collections, which are not available at xAOD level.
The DAOD_PHYS samples are then processed by the Easyjet framework which is used in other Higgs boson pair production analyses as well and is the foundation of a future common di-Higgs framework. In Appendix A.1.1 more details are given on the framework. The Easyjet framework outputs data trees, which are then passed to the HHARD framework for post-processing. A description of the HHARD framework is given in Appendix A.1.2. For the neural network study performed in this thesis, the software TensorFlow is used. It is introduced in Appendix A.1.3.

A.1.1. Easyjet Framework

The Easyjet [70] framework is a highly modular framework, that runs within Athena [71]. It is used by many di-Higgs analyses and it is in active development. It consists of a core framework embedded in the EasyjetHub and analysis specific developments contained within separate packages and varying for each analysis. For this analysis, the Easyjet framework is first used to apply the specific selections to the analysis objects and events, while also keeping track of the efficiencies of each selection cut. Afterwards, the output variables, e.g. weights, truth information and analysis specific variables, are set, filled and stored in ROOT n-tuples

A.1.2. HHARD Framework

For post-processing of the Easyjet framework, the HHARD [72] framework developed by the $HH \to b\bar{b}\tau\tau$ analysis is used. It efficiently converts the ROOT n-tuples to histograms

and also hdf5 files and can handle large amounts of data. It is also applies weights to the generated events and it is possible to define different regions for the output, such as individual regions for the two signal lepton flavours in this analysis.

A.1.3. TensorFlow

TensorFlow [73] is a well established and popular machine learning open-source software library focusing on the training of neural networks. In this thesis, TensorFlow was used for neural network setup and training as it includes all required features, such as loss functions, optimisers and learning rate schedulers like `ReduceLearningRateOnPlateau`. It is also very user friendly and simple to run on an available GPU. Even though TensorFlow has many helpful features for training, some other features were missing for this analysis. In these cases, the python library scikit-learn [74] was used in addition.

A.2. Monte Carlo Samples

A.2.1. Signal Samples

mc23_13p6TeV.525948.MGPy8EG_PDF30NLO_X2000tohh_WWbb_1lep.deriv.DAOD_PHYS.e8551_s4162_r14622_p5855
mc23_13p6TeV.525949.MGPy8EG_PDF30NLO_X3000tohh_WWbb_1lep.deriv.DAOD_PHYS.e8551_s4162_r14622_p5855
mc23_13p6TeV.525950.MGPy8EG_PDF30NLO_X4000tohh_WWbb_1lep.deriv.DAOD_PHYS.e8551_s4162_r14622_p5855

A.2.2. Background Samples

$t\bar{t}$ background:

mc21_13p6TeV.601229.PhPy8EG_A14_ttbar_hdamp258p75_SingleLep.deriv.DAOD_PHYS.e8453_s3873_r13829_p5855

W+jets background:

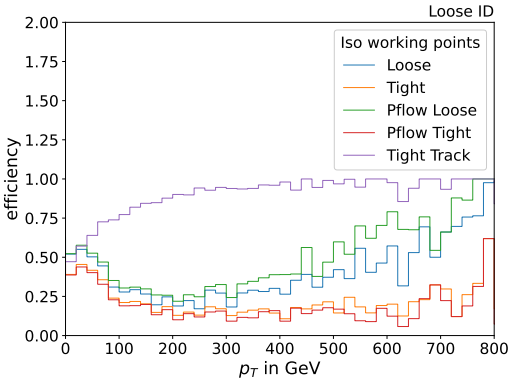
mc23_13p6TeV.700777.Sh_2214_Wenu_maxHTpTV2_BFilter.deriv.DAOD_PHYS.e8514_s4162_r14622_p5855
mc23_13p6TeV.700778.Sh_2214_Wenu_maxHTpTV2_CFilterBVeto.deriv.DAOD_PHYS.e8514_s4162_r14622_p5855
mc23_13p6TeV.700779.Sh_2214_Wenu_maxHTpTV2_CVetoBVeto.deriv.DAOD_PHYS.e8514_s4162_r14622_p5855
mc23_13p6TeV.700780.Sh_2214_Wmunu_maxHTpTV2_BFilter.deriv.DAOD_PHYS.e8514_s4162_r14622_p5855
mc23_13p6TeV.700781.Sh_2214_Wmunu_maxHTpTV2_CFilterBVeto.deriv.DAOD_PHYS.e8514_s4162_r14622_p5855
mc23_13p6TeV.700782.Sh_2214_Wmunu_maxHTpTV2_CVetoBVeto.deriv.DAOD_PHYS.e8514_s4162_r14622_p5855
mc23_13p6TeV.700783.Sh_2214_Wtaunu_maxHTpTV2_BFilter.deriv.DAOD_PHYS.e8514_s4162_r14622_p5855
mc23_13p6TeV.700784.Sh_2214_Wtaunu_maxHTpTV2_CFilterBVeto.deriv.DAOD_PHYS.e8514_s4162_r14622_p5855
mc23_13p6TeV.700785.Sh_2214_Wtaunu_maxHTpTV2_CVetoBVeto.deriv.DAOD_PHYS.e8514_s4162_r14622_p5855

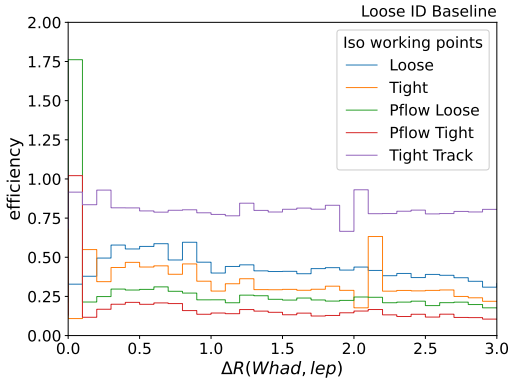
Dijet background:

mc23_13p6TeV.801165.Py8EG_A14NNPDF23LO_jj_JZ0.deriv.DAOD_PHYS.e8514_s4162_r14622_p5980
mc23_13p6TeV.801166.Py8EG_A14NNPDF23LO_jj_JZ1.deriv.DAOD_PHYS.e8514_s4162_r14622_p5980
mc23_13p6TeV.801167.Py8EG_A14NNPDF23LO_jj_JZ2.deriv.DAOD_PHYS.e8514_s4162_r14622_p5980
mc23_13p6TeV.801168.Py8EG_A14NNPDF23LO_jj_JZ3.deriv.DAOD_PHYS.e8514_s4162_r14622_p5980
mc23_13p6TeV.801169.Py8EG_A14NNPDF23LO_jj_JZ4.deriv.DAOD_PHYS.e8514_s4162_r14622_p5980
mc23_13p6TeV.801170.Py8EG_A14NNPDF23LO_jj_JZ5.deriv.DAOD_PHYS.e8514_s4162_r14622_p5980
mc23_13p6TeV.801171.Py8EG_A14NNPDF23LO_jj_JZ6.deriv.DAOD_PHYS.e8514_s4162_r14622_p5980
mc23_13p6TeV.801172.Py8EG_A14NNPDF23LO_jj_JZ7.deriv.DAOD_PHYS.e8514_s4162_r14622_p5980

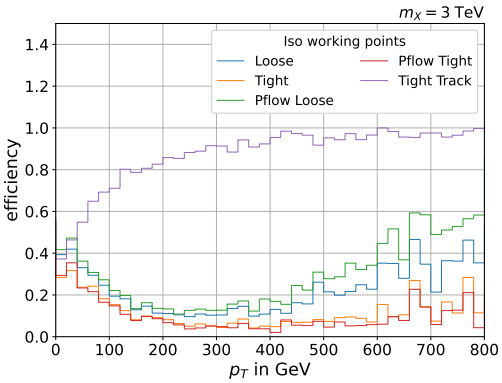
A.3. Additional Plots

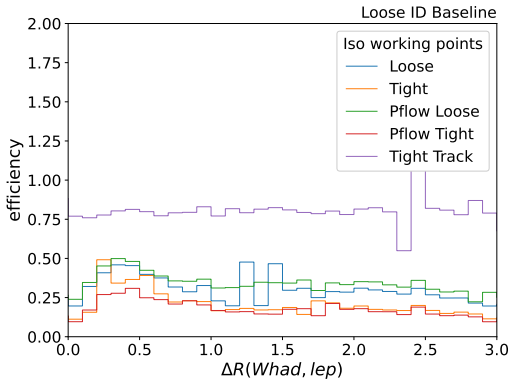
A.3.1. Lepton WP Study



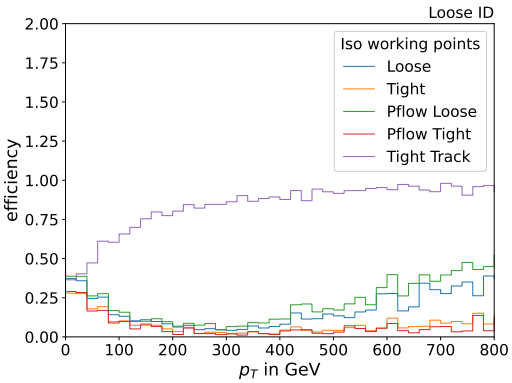


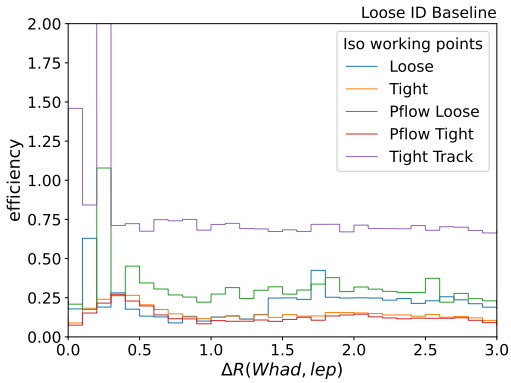
for $m_X = 2\,\mathrm{TeV}$





for $m_X = 3\,\mathrm{TeV}$

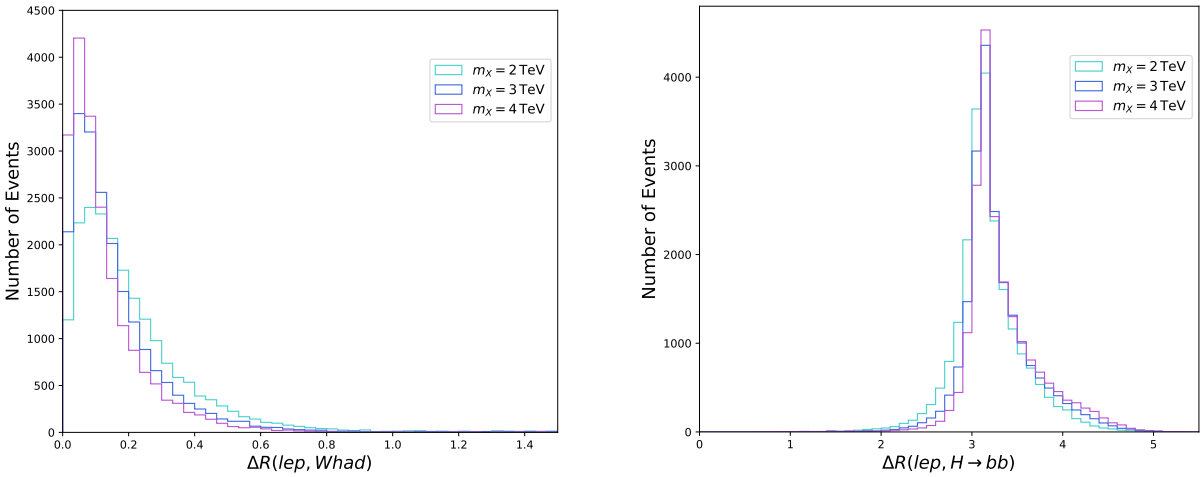




for $m_X = 4\,\mathrm{TeV}$

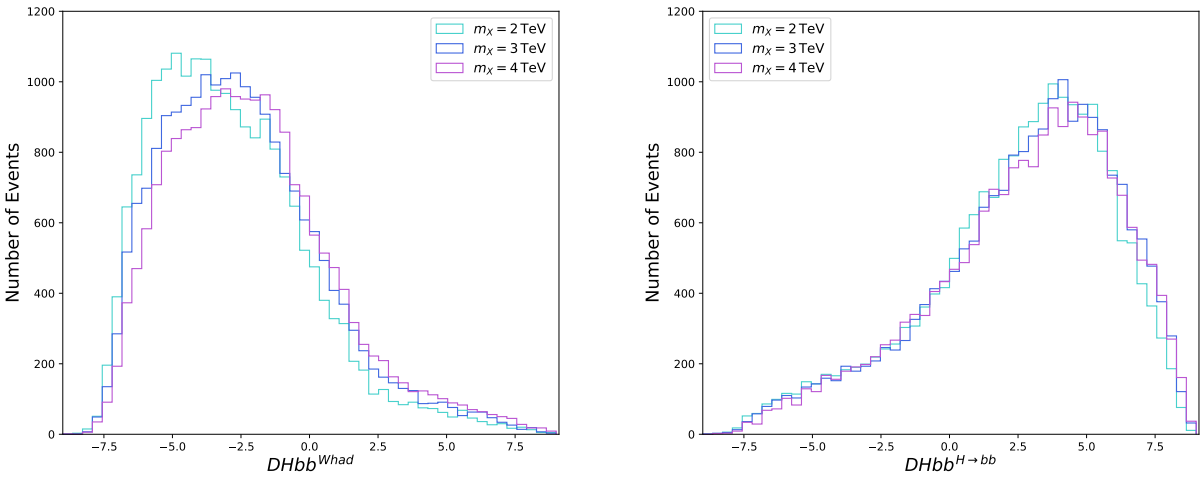
Figure A.1.: Fraction of muons passing the isolation working point with respect to their p_T and ΔR(Whad,lep).

A.3.2. Jet Object Classification



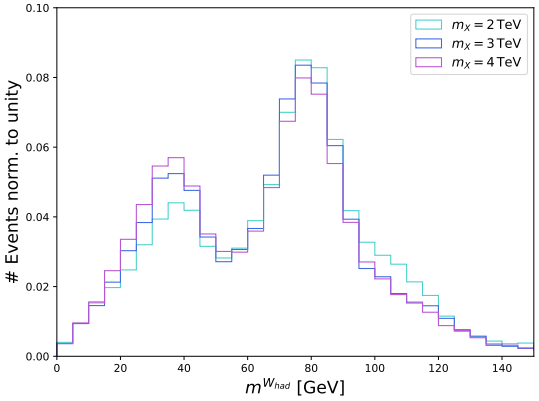
(a) W_{had} Jet **(b)** $H \to b\bar{b}$ Jet

Figure A.2.: ΔR(lep,jet) distributions for W_{had} and $H \to b\bar{b}$, distance classification applied.

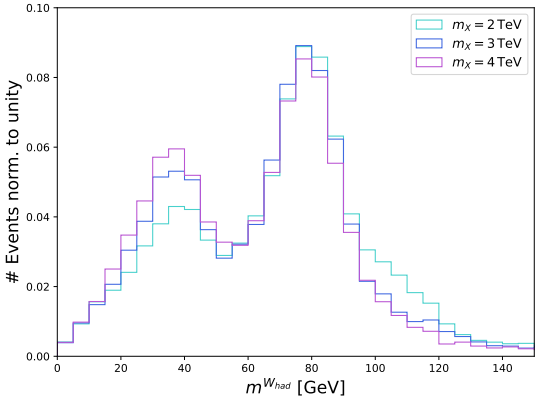


(a) W_{had} Jet **(b)** $H \to b\bar{b}$ Jet

Figure A.3.: D_{Hbb} distributions for W_{had} and $H \to b\bar{b}$, GNN classification applied.

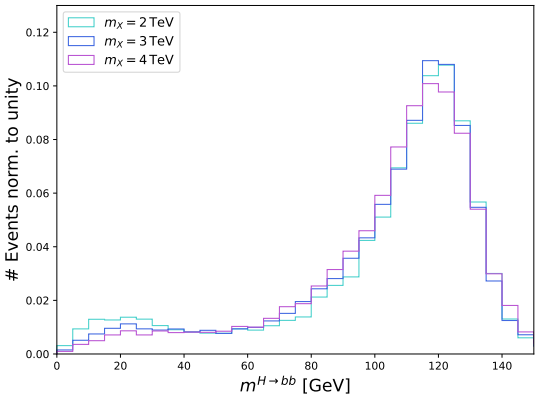


(a) distance class.

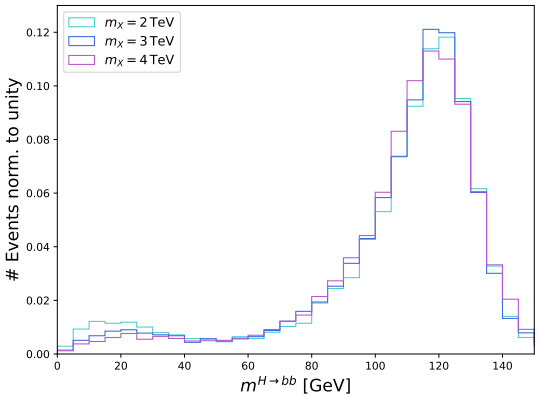


(b) GNN class.

W_{had} mass distributions



(c) distance class.



(d) GNN class.

$H \to b\bar{b}$ mass distributions

Figure A.4.: Mass distributions of the two jet objects for both classification methods.

A.3.3. Neural Network Input Variable Study

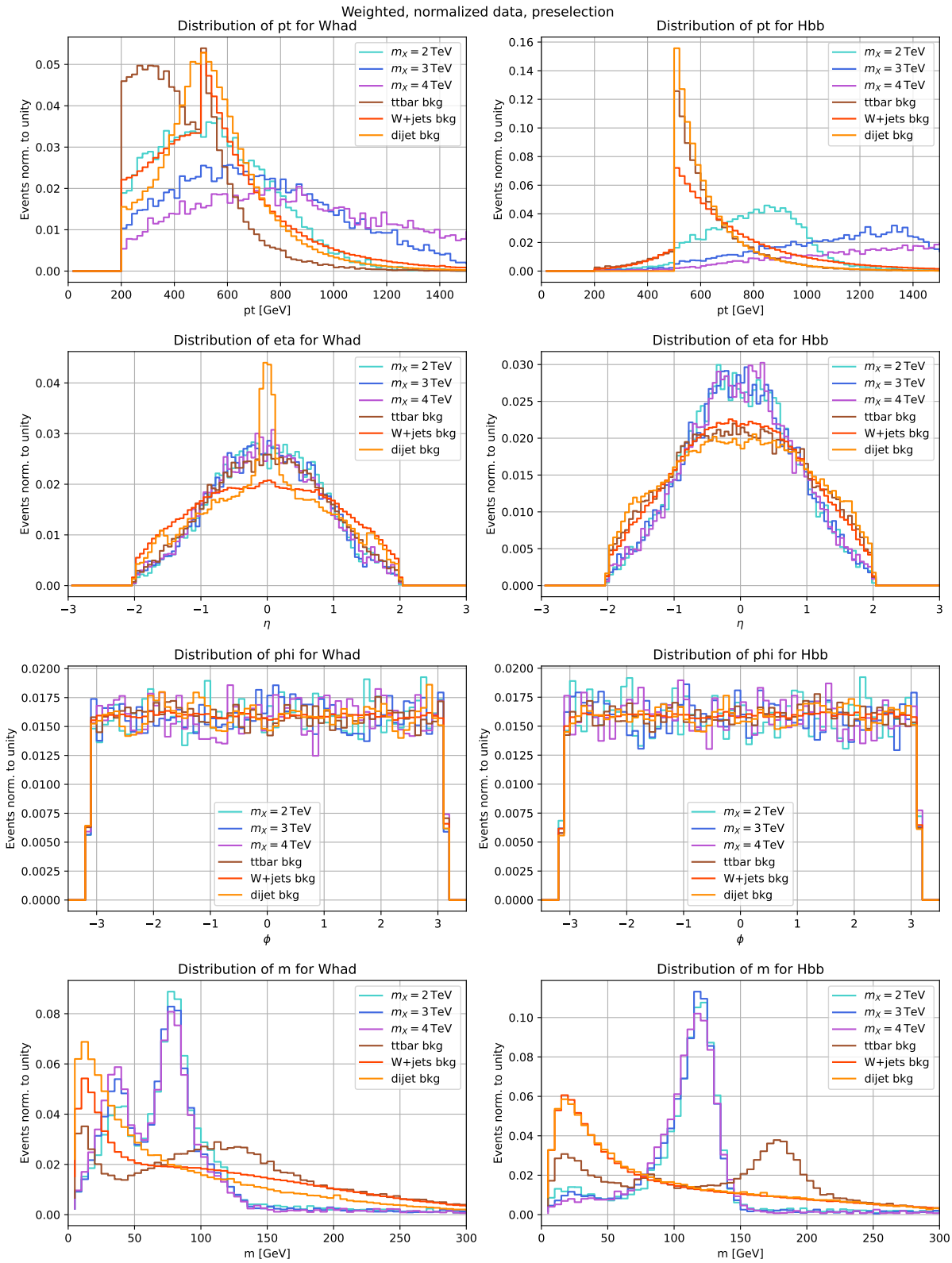


Figure A.5.: Distribution of kinematic variables of W_{had} and $H \to b\bar{b}$ candidate, weights and preselection applied.

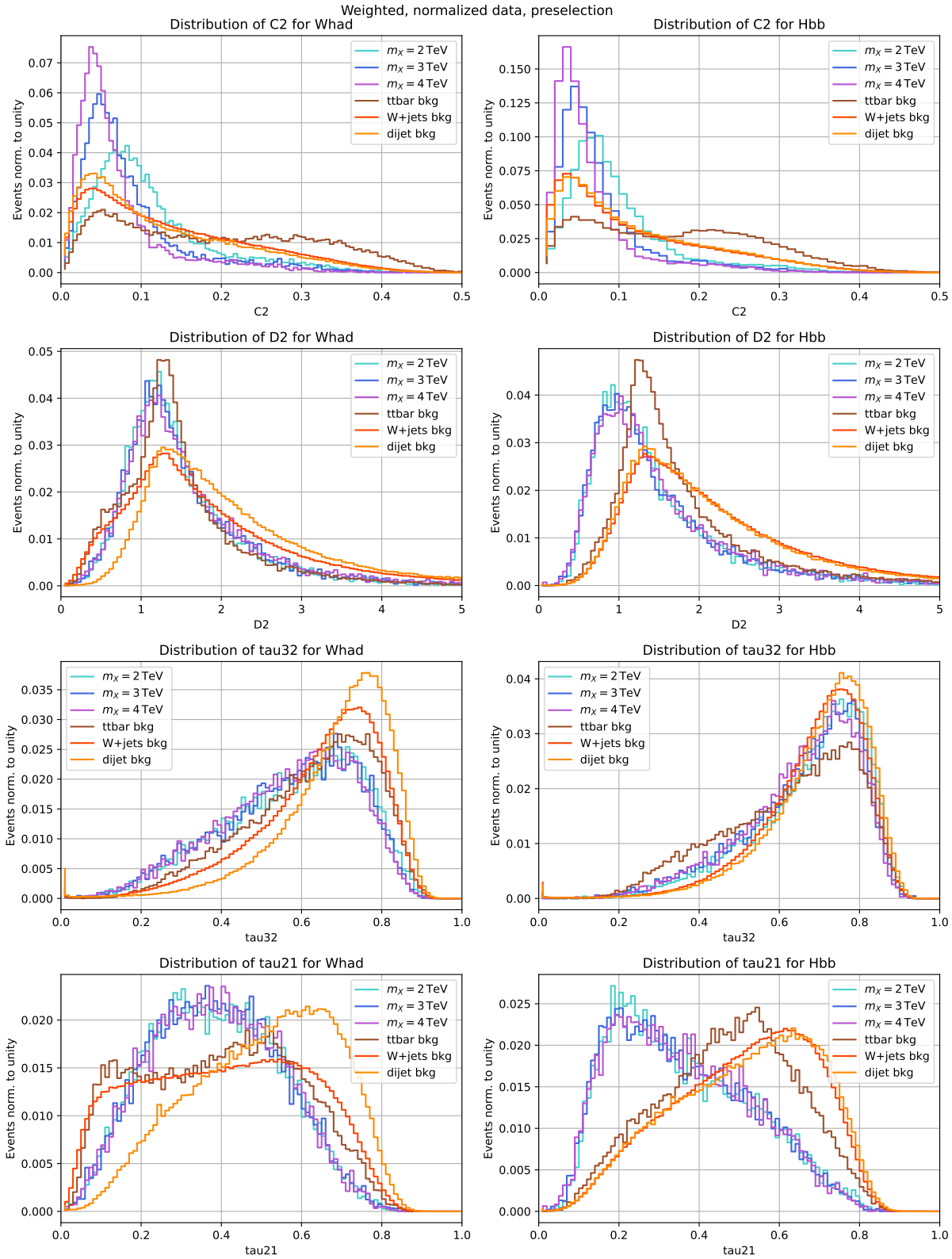


Figure A.6.: Distribution of substructure variables of W_{had} and $H \to b\bar{b}$ candidate, weights and preselection applied.

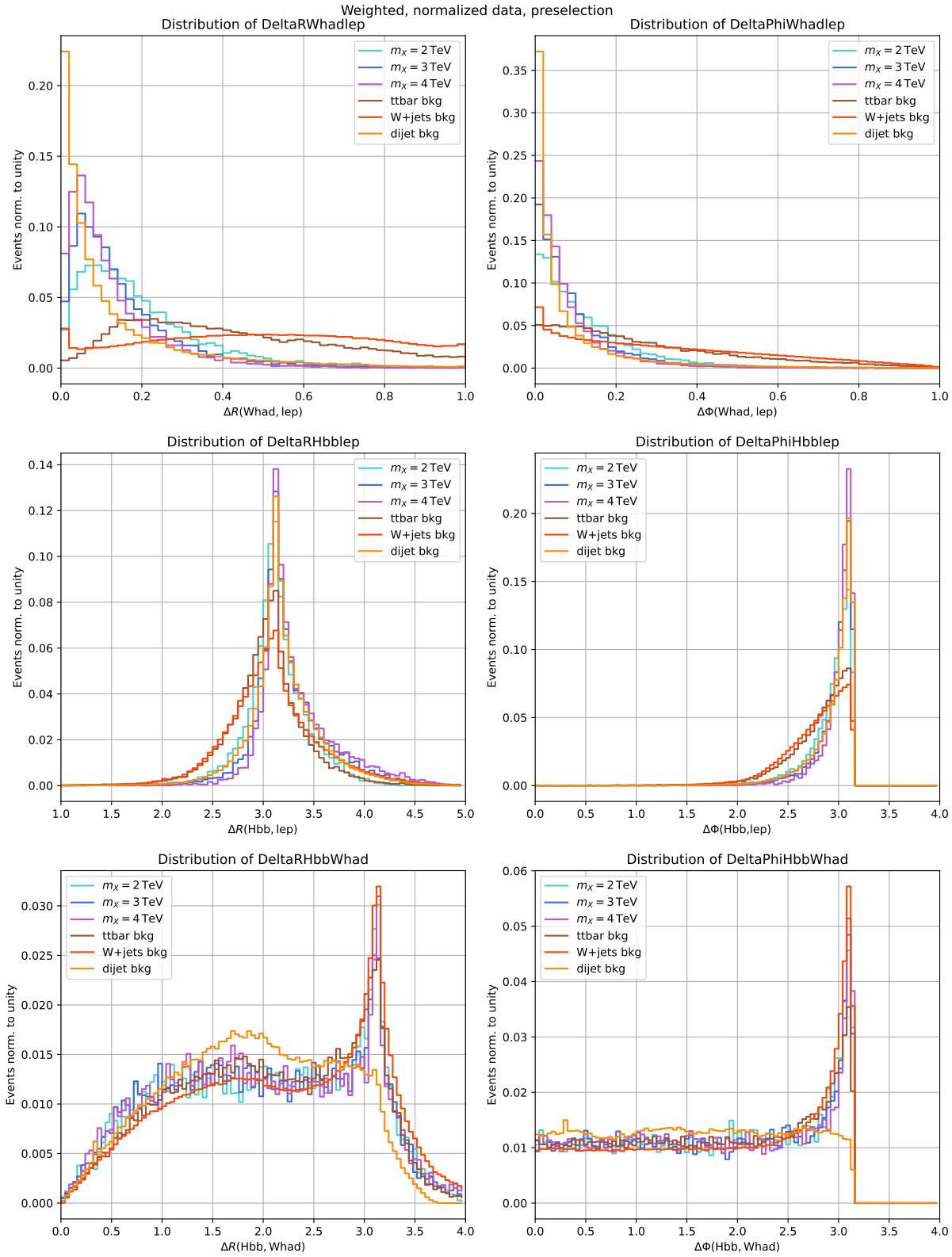


Figure A.7.: Distribution of geometrical and angular distance between signal lepton, W_{had} and $H \to b\bar{b}$ candidate, weights and preselection applied.

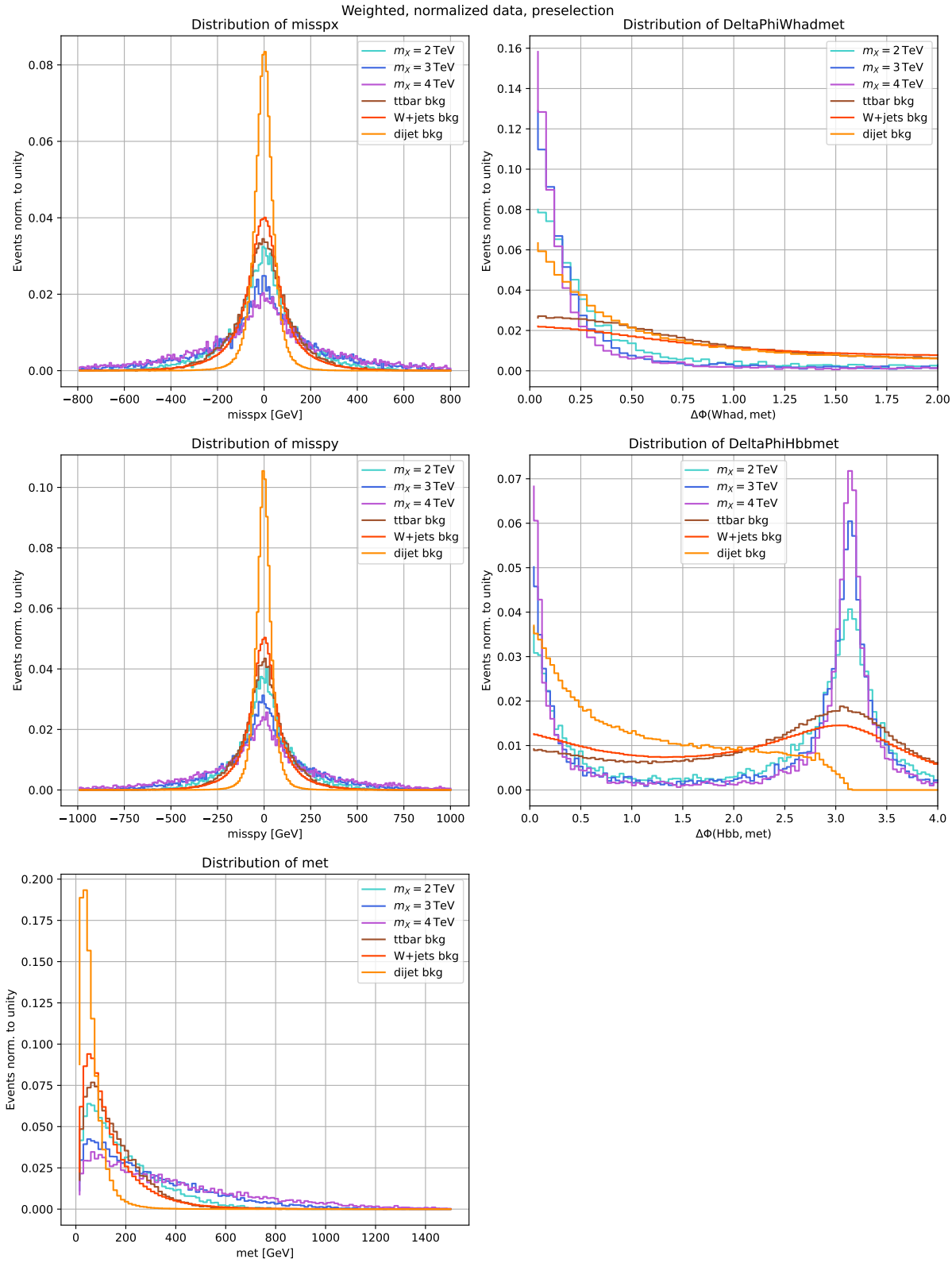


Figure A.8.: Distribution of missing E_T/momentum and geometrical and angular distance between missing E_T, W_{had} and $H \to b\bar{b}$ candidate, weights and preselection applied.

Bibliography

[1] P. A. Zyla, et al. (Particle Data Group), Prog. Theor. Exp. Phys. **2020** (2020)

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

[3] A. Salam, *Gauge unification of fundamental forces*, Rev. Mod. Phys. **52**, 525 (1980)

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

[5] 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)**, 1 (2012)

[6] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716(1)**, 30 (2012)

[7] S. Dawson, S. Dittmaier, M. Spira, *Neutral Higgs-boson pair production at hadron colliders: QCD corrections*, Phys. Rev. D **58**, 115012 (1998)

[8] D. de Florian, J. Mazzitelli, *Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD*, Phys. Rev. Lett. **111**, 201801 (2013)

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

[10] G. Bertone, D. Hooper, J. Silk, *Particle dark matter: evidence, candidates and constraints*, Phys. Rep. **405(5)**, 279 (2005)

[11] P. Drescher, *Signal/Background separation in resonant boosted HH and SH production in the bbVV* decay channel with 1-lepton in the final state* (2023), Thesis number: II.Physik-UniGö-BSc-2023/02

[12] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12(2)**, 132 (1964)

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

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

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

[16] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145**, 1156 (1966)

[17] T. W. B. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*, Phys. Rev. **155**, 1554 (1967)

[18] A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*, Ann. Phys. **354(7)**, 769 (1916)

[19] R. P. Feynman, *The Theory of Positrons*, Phys. Rev. **76**, 749 (1949)

[20] M. Gell-Mann, *The Eightfold Way: A Theory of strong interaction symmetry* (1961)

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

[22] D. J. Gross, F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30**, 1343 (1973)

[23] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30**, 1346 (1973)

[24] H. Fritzsch, M. Gell-Mann, H. Leutwyler, *Advantages of the color octet gluon picture*, Phys. Lett. B **47(4)**, 365 (1973)

[25] Y. Nambu, G. Jona-Lasinio, *Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. I*, Phys. Rev. **122**, 345 (1961)

[26] Y. Nambu, G. Jona-Lasinio, *Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. II*, Phys. Rev. **124**, 246 (1961)

[27] T. Plehn, M. Spira, P. Zerwas, *Pair production of neutral Higgs particles in gluon-gluon collisions*, Nucl. Phys. B **479(1)**, 46 (1996)

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

[29] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector* (2014)

[30] P. Baldi, et al., *Jet substructure classification in high-energy physics with deep neural networks*, Phys. Rev. D **93(9)** (2016)

[31] A. L. Samuel, *Some Studies in Machine Learning Using the Game of Checkers*, IBM Journal of Research and Development **3(3)**, 210 (1959)

[32] J. F. Trevor Hastie, Robert Tibshirani, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer New York (2009)

[33] L. P. Kaelbling, M. L. Littman, A. W. Moore, *Reinforcement Learning: A Survey*, Journal of Artificial Intelligence Research **Vol. 4**, 237 (1996)

[34] F. Rosenblatt, *The perceptron: A probabilistic model for information storage and organization in the brain.*, Psychological Review (1958)

[35] X. Glorot, A. Bordes, Y. Bengio, *Deep Sparse Rectifier Neural Networks*, in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* (2011)

[36] V. Nair, G. E. Hinton, *Rectified linear units improve restricted boltzmann machines*, International Conference on Machine Learning (2010)

[37] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, Mathematics of Control, Signals and Systems **2(4)**, 303 (1989)

[38] S. Leijnen, F. Veen, *The Neural Network Zoo*, Proceedings **47**, 9 (2020), Proceedings of the Conf. on Theoretical Information Studies 2019

[39] L. Wu, et al., *Graph Neural Networks: Foundations, Frontiers, and Applications*, Springer Singapore (2023)

[40] D. P. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization* (2017), `1412.6980`

[41] E. Mobs, *The CERN accelerator complex in 2019. Complexe des accélérateurs du CERN en 2019* (2019), general Photo, URL `https://cds.cern.ch/record/2684277`

[42] L. Evans, P. Bryant, *LHC Machine*, JINST **3(08)**, S08001 (2008)

[43] ATLAS Collaboration, *Public ATLAS Luminosity Results for Run-3 of the LHC*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3, accessed: 2024-05-03

[44] CERN, *CAS - CERN Accelerator School: Intermediate Accelerator Physics: Zeuthen, Germany 15 - 26 Sep 2003. CAS - CERN Accelerator School: Intermediate Course on Accelerator Physics*, CERN, Geneva (2006), URL https://cds.cern.ch/record/603056

[45] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3(08)**, S08003 (2008)

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

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

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

[49] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, Technical report, CERN, Geneva (2017)

[50] ATLAS Collaboration, *ATLAS experiment schematic illustration* (2022), general Photo, URL https://cds.cern.ch/record/2837191

[51] ATLAS Collaboration, *Production and integration of the ATLAS Insertable B-Layer*, JINST **13(05)**, T05008 (2018), URL https://dx.doi.org/10.1088/1748-0221/13/05/T05008

[52] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020)

[53] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77(5)**, 317 (2017)

[54] J. Alwall, et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **2014(7)**, 79 (2014)

[55] R. D. Ball, et al., *Parton distributions with LHC data*, Nucl. Phys. B **867(2)**, 244 (2013)

[56] C. Bierlich, et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3* (2022), URL `https://arxiv.org/abs/2203.11601`

[57] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462(1)**, 152 (2001), bEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines

[58] E. Bothmann, et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7**, 034 (2019)

[59] *Recommended isolation working points (Rel. 22)*, `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/RecommendedIsolationWPsRel22`, accessed: 2024-05-03

[60] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data*, JINST **14(12)**, P12006 (2019)

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

[62] K. Abeling, *Search for resonant Higgs boson pair production in the $b\bar{b}WW^*$ decay channel in the boosted 1-lepton final state using the full Run 2 ATLAS dataset*, Ph.D. thesis, Georg-August-Universität Göttingen (2022), Thesis number: II.Physik-UniGö-Diss-2022/01

[63] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77(7)**, 490 (2017)

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

[65] M. Cacciari, G. P. Salam, G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72(3)**, 1896 (2012)

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

[67] ATLAS Collaboration, *Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton–proton collisions*, Eur. Phys. J. C **81(4)**, 334 (2021)

[68] ATLAS Collaboration, *Improving jet substructure performance in ATLAS using Track-CaloClusters*, Technical report, CERN, Geneva (2017), URL `https://cds.cern.ch/record/2275636`

[69] *ATLAS Flavour Tagging Xbb Documentation*, `https://xbb-docs.docs.cern.ch/Xbb/GN2_track/`, accessed: 2024-05-03

[70] ATLAS Collaboration, *Easyjet*, URL `https://gitlab.cern.ch/easyjet/easyjet`

[71] ATLAS Collaboration, *Athena*, URL `https://gitlab.cern.ch/atlas/athena`

[72] ATLAS Collaboration, *HHARD*, URL `https://gitlab.cern.ch/atlas-physics/HDBS/DiHiggs/bbtautau/HHARD`

[73] M. Abadi, et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems* (2015), URL `https://www.tensorflow.org/`

[74] F. Pedregosa, et al., *Scikit-learn: Machine learning in Python*, Journal of machine learning research **12(Oct)**, 2825 (2011)

Danksagung

> Look up at the stars and not down at your feet. Try to make sense of what you see, and wonder about what makes the universe exist. Be curious.
>
> *Stephen Hawking*

First of all, I would like to thank **Prof. Dr. Stan Lai** for all the support he provided me with throughout my entire academic career. He guided me through several lectures, an internship, and the past year, during which I worked on this thesis. He was a great mentor to me, and I learned a lot from him. He probably does not remember, but he was also part of the reason I came to Göttingen in the first place. We spoke at Ideen-Expo 2017, and he gave me his business card, which I still have lying around somewhere.
I would also like to thank **Dr. Benjamin Schwenker**, not only because he was my supervisor for my bachelor thesis in AG Frey, who taught me a lot during the Corona pandemic, but also for agreeing to be my second referee again.
Next, I would like to thank the entire **AG Lai** and the whole Second Institute. Above all, I thank **Dr. Kira Abeling**, who has supported me for one and a half years now. Working with you was incredibly productive and always fun. I will miss our office. I would also like to thank **André Wilhahn**, my number one hater (jk), and **Athul Dev Sudhakar Ponnu** for technical support.
I also thank **Peer Christian Drescher** and **Josephine Stanitzok** for helping me understand neural networks and setting up the training framework.
Furthermore, my thanks go to the **bbVV analysis group** for their suggestions and discussions during meetings. I would like to explicitly thank **Valerio D'Amico** and **Jared Little** for helping me with technical and/or physical issues.
Special thanks go to **Greg Myers and Jordy Degens** from the $bb\tau\tau$ analysis group, who helped me set up the HHARD framework, and even after that, helped with fixing many bugs.
Additionally, I would like to thank **Lena Schulz** for being the best thesis buddy and DPG roommate, for listening to me complain about my code crashing, and for just being there and experiencing the same things as me. I wish you all the best.

Abschließend möchte ich mich noch bei einigen Leuten außerhalb der Physik bedanken. Zunächst bedanke ich mich beim gesamten **Team der Spieleburg**, die mich nun auch schon seit einigen Jahren neben meinem Studium begleiten. Insbesondere bedanke ich mich bei **Arne und Gabi Soltendieck**, die mir in diesen letzten, entscheidenden Wochen meines Studiums den Rücken so gut es geht freigehalten haben, sodass ich mich auf diese Arbeit konzentrieren konnte.

Außerdem danke ich meiner Familie. **Danke Mama, danke Jens, danke Papa, danke Opa, danke Oma, danke Steffi, danke Marco**, dass ihr mein physikalisches Interesse nie hinterfragt und mich mit allen Mitteln unterstützt habt. Ohne euch wäre ich jetzt nicht hier.

Ich danke meinen Plüschtieren für ihren moralischen Support.

Liebe Sophie,
danke, dass du schon mein ganzes Leben an meiner Seite bist und immer ein offenes Ohr für mich hast. Hab dich lieb, Cousinchen.

Marc,
danke für alles. Du bist immer für mich da und hast mich während der gesamten Masterarbeitszeit unterstützt. Danke, dass du mir selbst in den schwierigsten Situationen beigestanden hast. Und das Wichtigste: Danke für die oben erwähnten Plüschtiere. Ich liebe dich. <3

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 5. März 2025

(Lina Buschmann)