



School of Physics and Astronomy בית הספר לפיזיקה ולאסטרונומיה
The Raymond and Beverly Sackler הפקולטה למדעים מדויקים
Faculty of Exact Sciences ע"ש ריימונד ובברלי סאקלר
Tel Aviv University אוניברסיטת תל אביב

Research Thesis for a Degree "Doctor of Philosophy"

SEARCH FOR HIGGS BOSONS IN THE STANDARD MODEL AND BEYOND

Submitted to the Senate of Tel Aviv University by

HADAR COHEN

The work was carried out under the supervision of

PROF. EREZ ETZION

Hadar Cohen: *Search for Higgs Bosons in the Standard Model and Beyond,*

©

Dedicated to the loving memory of my grandfather,
for his bedtime stories planted the curiosity seed.

Boris Rahav (Rudizki)

1928 – 1998

ABSTRACT

The detailed study of the Higgs sector is a crucial milestone for the comprehensive understanding of the latest discovered part in the Standard Model. Using the data collected with the ATLAS experiment at CERN's Large Hadron Collider, two searches are performed, probing the Standard Model Higgs Boson and beyond it.

A search for a heavy charged Higgs boson ($m_{H^\pm} > m_{top}$), which decays predominantly into a top and bottom quarks, is performed based on ATLAS Run 2 $\sqrt{s} = 13$ TeV pp collision dataset, corresponding to an integrated luminosity of 36 fb^{-1} , collected during 2015 to 2016. Different final states with one or two leptons (electrons or muons) are studied, and exclusion limits are set.

With increasing and more sophisticated data-taking, ATLAS is becoming more sensitive to new decay channels. A Standard Model Higgs boson with a mass of 125 GeV is predicted to decay to a pair of charm quarks around 2.9% of the time. This thesis describes a search for the decay of the Higgs boson to charm quarks, based on the full ATLAS Run 2 $\sqrt{s} = 13$ TeV pp collision dataset, corresponding to an integrated luminosity of 139 fb^{-1} , collected during 2015 to 2018. The $H \rightarrow c\bar{c}$ decay candidates are reconstructed from a pair of jets identified by a c -tagging algorithm. The expected upper limit is $26 \times$ SM prediction.

In 2024 the LHC will be upgraded into a High Luminosity LHC (HL-LHC). The HL-LHC will deliver about 300 fb^{-1} of data per year. This puts challenging requirements on the trigger algorithms. Novel techniques using Deep Neural Networks are presented in this thesis to increase the detector's ability to identify a τ -lepton correctly. These techniques exhibit a promising ability to explore new (low p_T) regions never been analysed before.

*Do not imagine that these most difficult problems can be
thoroughly understood by any one of us. This is not the case.
At times the truth shines so brilliantly, that we perceive it as clear as day.
Our nature and habit then draw a veil over our perception,
and we return to a darkness almost as dense as before.
We are like those who, though beholding frequent flashes of lightning,
still find themselves in the thickest darkness of the night.*

— Maimonides, *Guide for the Perplexed*

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor *Prof. Erez Etzion*. Erez is truly one of a kind, not your typical run-of-the-mill professor, instead of a suit and tie he sticks to speedo and goggles. Erez has a true passion for science and hard work. It was a joy hunting together (sometimes even fishing) for particles in and outside all kinds of caverns, pursuing deeper knowledge, especially on the beach with a beer, watching the sunset and counting muons. Thank you for these wonderful years of science, I will cherish them forever.

To *Dr. Liron Barak*, for being this energizer bunny who never runs out of juice, even in the middle of a pandemic with four kids at home. Thank you for always keeping your door open and allowed me to embark the tau lepton journey with you.

I gratefully thank *Prof. Marek Karliner*. I vividly remember Marek constantly saying in class: "*do not memorise anything, understand what you are doing, this is physics*". You truly left a mark as an astonishing educator, thank you.

To *Meny Ben-Moshe-Raviv*, my "mother" in the university. For putting up with me claiming every sane piece of the lab. Thank you for it all, whether it is taking care of me as an under-graduate or working hard together on a project, it was genuinely a pleasure.

To my loving parents, *Daniel* and *Ilana Cohen*, for laying the foundations and supporting me throughout this long journey.

Lastly, to the most important person in my life, my wife, *Yamit*. Whether it is sailing a fierce thunderstorm in the Adriatic Sea with a broken helm or making it through this academic journey, you were always by my side, pushing me forward, lifting me when needed and dreaming with me towards future goals. No language is cable of phrasing my love to you. Without you, none of this would come to life.

CONTENTS

1	INTRODUCTION	1
2	THEORETICAL BACKGROUND	7
2.1	The Standard Model of Particle Physics	7
2.1.1	The BEH Mechanism	9
2.1.2	Higgs-Boson Production	12
2.1.3	Higgs-Boson Decay Modes	15
2.2	The 2-Higgs Doublet Model	17
2.3	Production and Decay of Charged Higgs Bosons	19
3	THE LHC AND THE ATLAS EXPERIMENT	21
3.1	The Large Hadron Collider	21
3.2	LHC experiments	22
3.3	The ATLAS Detector	24
3.3.1	Overview	24
3.3.2	Conventions	25
3.3.3	Inner Detector	26
3.3.4	Calorimeters	27
3.3.5	Muon Spectrometer	29
3.3.6	Trigger System	32
4	STATISTICAL FORMALISM	35
4.1	Binned Profile Likelihood	35
4.2	Test Statistic	37
4.3	Asimov Dataset	38
4.4	CLs Limits	39
4.5	Blinding and Unblinding Analysis	40
5	SEARCH FOR CHARGED HIGGS BOSON DECAYING TO TOP AND BOTTOM QUARKS	43
5.1	Introduction	43
5.2	Previous and Future Results	43

5.3	Analysis strategy	44
5.4	Signal and Background Production	46
5.4.1	Data Sample	47
5.4.2	Signal Samples	47
5.4.3	Background Samples	49
5.5	Object Reconstruction	53
5.5.1	Electrons	54
5.5.2	Muons	54
5.5.3	Jets	55
5.5.4	Missing Transverse Momentum, E_T^{miss}	56
5.6	Single Lepton Final State	57
5.6.1	Selection	57
5.7	Background suppression using BDT	59
5.7.1	INC vs HF BDT	62
5.7.2	BDT Performance in the Signal Regions	63
5.7.3	BDT Variable Ranking and Importance	67
5.7.4	BDT Background and Data Correlation	68
5.7.5	Summary of Choices for the Analysis	70
5.8	Systematic Uncertainties	71
5.8.1	Luminosity and Pile-up	72
5.8.2	Reconstructed Objects	72
5.8.3	Signal Modelling	76
5.8.4	Background Modelling	78
5.9	Results	84
5.9.1	Upper Cross-Section Limits as a Function of H^+ Mass	84
5.9.2	Interpretation of the Limits	87
6	SEARCH FOR THE DECAY OF THE HIGGS BOSON TO CHARM QUARKS	89
6.1	Introduction	89
6.2	Data and Monte Carlo Production	92
6.2.1	Data Sample	92
6.2.2	Modelling of the VH Process	92

6.2.3	Background Samples	93
6.3	Event Selection	96
6.3.1	Event Selection Common to All Channels	97
6.3.2	1-lepton Channel Specific Selection	99
6.3.3	Analysis Strategy and Event Categorisation	99
6.3.4	Additional Selection Criteria	101
6.4	Flavour Tagging	104
6.4.1	Working Point Optimisation	106
6.5	Systematic Uncertainties	109
6.5.1	Experimental Systematic Uncertainties	109
6.5.2	Signal and Background Modelling Systematic Uncertainties	113
6.6	VH(cc) Fit Results	127
6.7	Limits on the c -Quark Yukawa Coupling	133
6.8	Conclusion	135
7	τ TRIGGER USING DEEP LEARNING	137
7.1	Introduction	137
7.2	L1 Calorimeter Trigger	138
7.3	Machine Learning and Neural Networks	140
7.3.1	Data Propagation	141
7.3.2	Deep Learning	143
7.4	Architectures for Triggering	144
7.4.1	Monte Carlo Samples	144
7.4.2	Benchmark Algorithms	145
7.4.3	Convolutional Neural Networks	146
7.4.4	Deep-Sets	151
7.4.5	Training	154
7.5	Results	155
7.6	Conclusion	157
8	SUMMARY	159

INTRODUCTION

"Science, my boy, is made up of mistakes, but they are mistakes which it is useful to make because they lead little by little to the truth."

— Jules Verne, *Journey to the Center of the Earth*

The discovery of the Higgs boson [1–3] in 2012 by the ATLAS [4] and CMS [5] experiments is a result of knowledge accumulated during centuries of persistent scientific work done by many people around the globe. This is only one discovery, but such a discovery cannot come without the trial and error of many before. Every attempt to search for new phenomena brings us, the human race, into a clearer understanding of the surrounding universe, regardless of its outcome. This process allowed us to better understand matter's 'elementary' constituents, from atoms to electrons and nuclei from protons to quarks.

To date, our clearest and simplest knowledge of the fundamental rules of particle physics is laid out in the Standard Model (SM) [6–8], a theoretical framework that allows us to describe the properties, but more importantly, the interactions of elementary particles. While the Standard Model has been found to be in agreement with experimental results to an impressive level of accuracy, additional theories, referred to as beyond the Standard Model (BSM), have been developed to extend this framework and account for unexplained phenomena, e.g., neutrino masses, dark matter and gravitation.

This thesis presents three aspects of the author's research activities performed as part of the (A) Toroidal LHC ApparatuS (ATLAS) experiment at the Large Hadron Collider (LHC) at CERN. As these projects are done in collaboration, the author's personal contributions are given extra attention in the chapters and summarised here. In addition, a summary of the author's ATLAS qualification task is presented below.

1. **BSM, $H^\pm \rightarrow tb$** Even without including gravity, there are good experimental and theoretical reasons to believe that the SM is not the whole story. A deeper conceptual structure may be lurking behind it. Any sign of a charged scalar would be a clear indication of new physics. Numerous attempts have been made throughout the years, each result narrowing down the phase-space and excluding more and more regions, thus, focusing the effort. With each search, new and more sophisticated tools have been developed to find any sign of deviation from the SM. In this thesis, a search for the charged Higgs boson is performed in the decay channel $H^\pm \rightarrow tb$, whose branching ratio can range anywhere up to 100% depending on the type of Yukawa couplings. In a result [9], based on this thesis work, no significant excess above the expected SM background is found, and the appropriate upper limits were set as a function of $H^\pm$ mass.

Personal Contribution: The author was part of ATLAS's previous analysis run and continued with the new analysis team from the very start. During the course of the analysis, the author contributed to (lepton+jets) Data/MC comparisons, discriminating variables studies, signal MC studies (yields, negative weights, S/B). BDT (HF+INC) design and implementation. Studies of Deep Neural Networks and other methods as alternative multivariate discriminators. Finally, he was highly involved in developing the statistical infrastructure framework as part of the larger $t\bar{t}Hbb$ group.

2. $\text{SM}, H \rightarrow c\bar{c}$ After the discovery of the Higgs boson, attention has turned to more detailed measurements of its properties and couplings as means of testing the predictions of the SM. As the Higgs boson couples differently to each particle, the careful study of each interaction is crucial for the profound understanding of its nature. According to the SM, the Higgs boson has only a 2.9% chance of decaying directly into a pair of charm quarks. This low branching ratio makes this decay mode highly susceptible to contribution from new physics; thus, deviation from expected SM prediction could lead to signs of new physics. The first-ever result [10] of this direct measurement yielded an upper limit of $110\times$ the SM prediction. Since the search is an effort of the collaboration, the analysis team is not allowed to unblind the data before the official paper of ATLAS is published. Therefore, the analysis will be described in great detail without stating the observed results. The expected upper limit in the integrated luminosity of 139 fb^{-1} is $26\times$ the SM prediction. Within the next year, ATLAS is scheduled to publish an independent result along with a combined effort of $VH(c\bar{c})$ and $VH(b\bar{b})$ which is expected to improve any existing result significantly.

Personal Contribution: As a member of the analysis team from the first moment, the author has contributed to the majority of studies conducted in the analysis. Among these are background estimation with data and Monte-Carlo simulations, control and signal regions division, charm tagging optimisation, analysis and statistical framework development, fit studies including channels combination and lastly, shape and normalisation systematics estimation.

3. τ TRIGGER The tau lepton has a mass of 1.776 GeV and decays either leptonically ($\tau \rightarrow \nu_\tau \nu_l, l = e, \mu$) or hadronically ($\tau \rightarrow \nu_\tau$ hadrons) after a decay length of $87\mu m$. Such a short decay length means that tau leptons typically decay before they reach the innermost layer of the detector, and thus, only its decay products can be observed. The τ s, being the heaviest, third-generation leptons, provide a unique experimental opportunity to search for low-lying states that would otherwise go undetected. Proper identification is a crucial step to efficiently explore new phenomena, allowing to record events from unexplored regions. In this thesis, novel techniques for identifying tau lepton at the trigger level are being developed. Two methods are being presented, and both are based on Deep-Neural-Nets (DNN), which use only the raw calorimeter energy deposits as input, without any reconstructed object. Compared to current algorithms, both provide a significant improvement, thus, leading the way for future detector upgrades.

Personal Contribution: Besides professional guidance, the entire idea and execution are an original contribution of the author, done solely by him. Furthermore, the author also served as an advisor on developing a simpler approach (yet still improving over the existing methods) by using Boosted Decision Trees (BDTs) over Field-Programmable Gate Array (FPGA) chip so it could be done in the near future.

4. ATLAS QUALIFICATION Each member of the ATLAS collaboration is required to contribute to a subject different from their main work. Tel-Aviv University, together with Weizmann Institute of Science and the Technion, are responsible in Israel for the development and construction of ATLAS's muon spectrometers named small-strip Thin Gap Chambers (sTGC). The author's contribution to the project is two-fold, first, involved in the testing of prototype detectors, and second, in the assembly of the final New Small Wheel (NSW) [11] comprised of these detectors. For the first part, the author developed

As a second-generation to the former and larger Thin Gap Chamber (TGC), the new detector is given a prefix of lower-case 's' to emphasise its smaller size.

a real-time measuring framework for the use of Tel-Aviv University's cosmic-rays hodoscope. This framework enabled the detailed examination of prototype detectors as they are hit with cosmic rays. Thus, providing a crucial understanding of the working status of each part of the detector. Later on, this framework was extended to provide an online trajectory analysis of multiple detector layers and was put to use in a several weeks long beam experiment at Fermilab [12], which provided meaningful inputs to further research and development. The second part was NSW sTGC wedge integration. The NSWs aim to replace the present Small Wheels with new detectors and electronics, which can cope with the High-Luminosity LHC rates and will provide improved triggering. The author developed software to analyse and monitor chamber data during the noise and pulser tests at CERN. The integration of multiple institutes from different countries, each with its own detector conventions and tests had to be properly managed to construct a fully working wedge.

The work in this thesis is presented as follows.

Chapter 2 lays out the theoretical foundations being used throughout this thesis. Chapter 3 describes the experimental landscape at the LHC, with a detailed focus on the ATLAS experiment. Chapter 4 summarises the statistical formalism used by both search analyses. Chapter 5 lays out a search for $H^\pm \rightarrow tb$, focusing on the development of classification techniques. Chapter 6 describes the $H^0 \rightarrow c\bar{c}$ direct search with associated production of a vector boson. Finally, chapter 7 outlines novel techniques for triggering on τ particles using DNNs.

THEORETICAL BACKGROUND

2.1 THE STANDARD MODEL OF PARTICLE PHYSICS

The Standard Model of elementary particles, developed in the 1960s and 1970s, has stood for years as "the" theory of particle physics, passing numerous stringent tests. It describes the fundamental constituents, the building blocks of matter and their interactions through electromagnetic, strong, and weak forces. The matter particles are elementary, meaning that they do not have an inner structure. They are called fermions, have half integer spin, and are further categorized into leptons and quarks. The force carriers are bosons, particles with an integer spin. A schematic illustration of the particle content of the SM is given in Figure 2.1.

The fermions are organized in three families or generations. The CERN Large-Electron-Positron (LEP) experiments showed no more than three families with light neutrinos in the SM [13]. Each family consists of two quarks and two leptons. The quarks can be categorized into one up-type quark with an electric charge of $+\frac{2}{3}e$ and one down-type quark with an electric charge of $-\frac{1}{3}e$. The leptons consist of an electron-like lepton and a neutrino. The massive leptons (electron, muon and tau) have an electric charge of $-e$. Neutrinos are electrically neutral and much lighter than the charged leptons. All leptons with an electric charge interact electromagnetically, while the neutrino can only interact via the weak force.

Quarks, unlike leptons, also carry a colour charge, the quanta of the strong interaction. Quarks come in three colors, "red", "green" and "blue". Additionally, they only exist in a confined state for which the sum of the colour charges is "white". A Baryon is a three-quarks bound

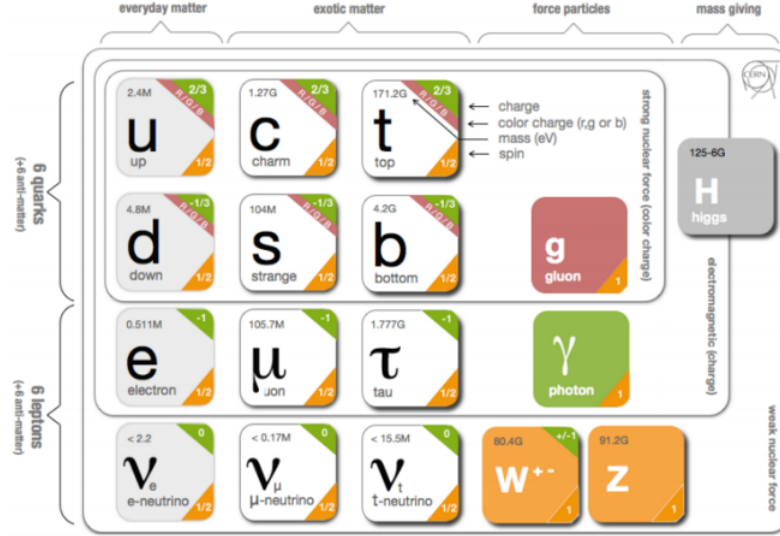


Figure 2.1: Standard Model particle content. Each square stands for an elementary particle and states its characteristics - charge, colour charge, mass and spin. The twelve fermions are divided into six quarks and six leptons sorted in three generations with increasing mass from left to right. The force mediator bosons share a rectangle with the fermions affected by the force they carry[14].

state. A proton, for example, is a baryon that consists of three quarks, two up and one down. A Meson is a two-quarks bound state which must have a quark and anti-quark pair. In recent years, bound states of four and five quarks (tetraquark- and pentaquark-like structure) have been predicted (e.g. [15]) and have been observed at the LHCb and Belle-II experiments (e.g. [16]).

In the SM, the forces are being carried out by the gauge bosons. The photon, γ , is the carrier of the electromagnetic force, the W and Z bosons are the carriers of the weak force and the gluons, g , for the strong force.

Up until 2012, almost all of the SM particles have been observed and studied. All but one missing piece, the Higgs boson. It was discovered in 2012 by the ATLAS and CMS experiments at CERN, Geneva [4, 5]. Unlike the other gauge bosons, the Higgs have a different role in the Standard Model, as explained further in the following chapters.

To claim a discovery, it is agreed that an experiment needs to demonstrate a deviation with a 5σ significance level. This will establish that a new particle was discovered with 99.9999% probability and $\leq 0.0001\%$

probability for being just a background fluctuation. For the collaborations to get enough sensitivity to claim a discovery, they combined several channels, $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ$, $H \rightarrow WW$, $H \rightarrow \tau\tau$, $H \rightarrow b\bar{b}$. The discovery was announced with the first three channels only, while the last two were added later.

Figure 2.2 presents the distribution of the four-lepton invariant mass m_{4l} measured by ATLAS and the invariant mass of di-photon candidates measured by CMS, as a selection of channels in which the Higgs boson with a mass of around 125 GeV was discovered. In July of 2012, both ATLAS and CMS presented results confirming the existence of the Higgs boson. Thereafter, in October of 2013, Peter Higgs and Francois Englert have been awarded the Nobel prize for physics.

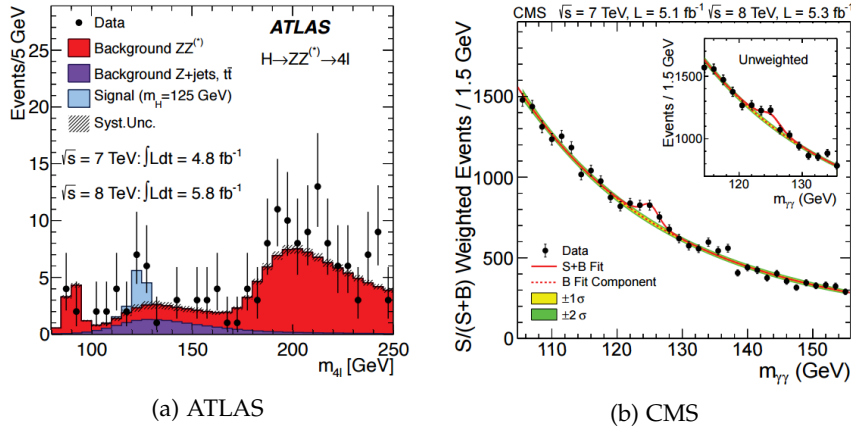


Figure 2.2: Four-lepton invariant mass, m_{4l} , for the selected candidates, measured in ATLAS (left) and the invariant mass distribution of di-photon candidates in CMS (right) as presented in the Higgs discovery papers [4, 5].

2.1.1 The BEH Mechanism

Local gauge invariance is integral in the construction of the SM Lagrangians. It can be shown that by adding a Dirac mass term ($m\bar{\psi}\psi$) to either the Quantum Chromo Dynamics (QCD) or Electroweak (EW) Lagrangians that the principle of local gauge invariance is broken. For Quantum Electro Dynamics (QED) and QCD, this is not a problem, as the respective vector bosons for these forces are assumed to be

massless. However, for the massive electroweak bosons, this poses a problem. As the concept of local gauge invariance is well-founded, with the experimentally verified predictions of QED and QCD, it was desirable to propose an alternative mechanism to generate the $W^\pm$ and Z^0 masses conserving the principle of local gauge invariance. The Brout-Englert-Higgs (BEH) mechanism [1–3] was originally formulated in 1964 and is explained in the following.

*BEH mechanism will
be referred to as
Higgs mechanism for
now on for shortness.*

Spontaneous Symmetry Breaking

Let us consider a Klein-Gordon Lagrangian for a scalar field with a modified mass term, μ , and a field self-interaction term λ : $\mathcal{L}_H = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda^2\phi^4$. When compared to the unmodified Klein-Gordon Lagrangian, it is seen that this modified Lagrangian has an unphysical (complex) mass term. The potential term can be written as: $V_H = -\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda^2\phi^4$ which can be graphically represented as shown in Figure 2.3.

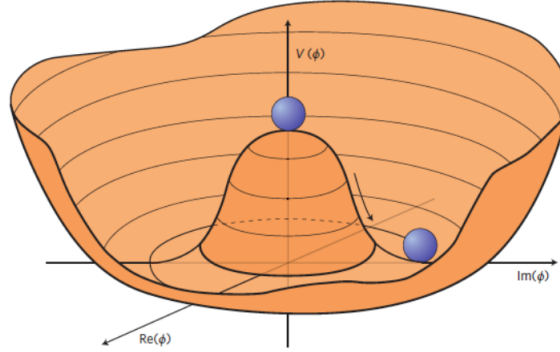


Figure 2.3: Graphical representation of the Higgs potential. Taken from [17].

The ground state of this potential is then found by expanding around the minima which gives the Lagrangian the form of: $\mathcal{L}_H = \frac{1}{2}\partial_\mu\eta\partial^\mu\eta + \frac{1}{2}\mu^2\eta^2 + \mu\lambda\eta^3 - \frac{1}{4}\left(\lambda^2\eta^4 + \frac{\mu^4}{\lambda^2}\right)$. The Lagrangian now contains a mass term ($\frac{1}{2}\mu^2\eta^2$), corresponding to a field with a boson of mass $m = \sqrt{2\mu}$. The higher-order terms in η are the field self-coupling terms corresponding to triple and quartic couplings.

Upon comparing the equations above, the rewritten Lagrangian is no longer invariant under the transformation $\eta \rightarrow -\eta$. The symmetry of the Lagrangian has been broken by the act of choosing to work with one of the ground (or vacuum) states. This is known as spontaneous symmetry breaking.

Spontaneous Symmetry Breaking of the SU(2) Group

Extending the scalar singlet in the previous example to a complex scalar SU(2) doublet will lead to the spontaneous symmetry breaking of the SU(2) group, and eventually a method for the generation of the $W^\pm$ and Z^0 boson masses

$$\Phi = \begin{pmatrix} \phi_a \\ \phi_b \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix},$$

leading to a Lagrangian for this field of the form: $\mathcal{L}_H = \frac{1}{2}(\partial_\mu \phi)^\dagger (\partial^\mu \phi) + \frac{1}{2}\mu^2(\phi^\dagger \phi) - \frac{1}{4}\lambda^2(\phi^\dagger \phi)^2$. Requiring the above Lagrangian to be invariant under local SU(2) transformations implies the introduction of the proper QED covariant derivative. The Lagrangian now becomes:

$$\mathcal{L}_H = \frac{1}{2} \left[(\partial_\mu + \frac{ig_W}{2} \vec{\tau} \vec{W}_\mu + ig_Y \frac{Y}{2} B_\mu) \phi \right]^\dagger \left[(\partial_\mu + \frac{ig_W}{2} \vec{\tau} \vec{W}_\mu + ig_Y \frac{Y}{2} B_\mu) \phi \right] + \frac{1}{2}\mu^2(\phi^\dagger \phi) - \frac{1}{4}\lambda^2(\phi^\dagger \phi)^2 - \frac{1}{4}W_{\mu\nu}W^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}.$$

Using the same approach as above, it is found that the potential is minimized when

$\phi^\dagger \phi = -\mu^2/\lambda^2$. A gauge is chosen, such that $\phi_1 = \phi_2 = \phi_4 = 0$, with $\phi_3 = -\mu^2/\lambda^2 \equiv v^2$, where v is known as the vacuum expectation value (VEV) and it is found that the equation contains a mass term for the ϕ -field, corresponding to a scalar boson with $m_h = \sqrt{2\lambda v^2}$.

With the Lagrangian above, it is now possible to "give" mass to the $W^\pm$ and Z^0 bosons by taking the covariant derivative and expanding

the above equation by performing the summation over the SU(2) generators and introducing the ground-state Higgs field. This leads to the definition of the physical $W^\pm$ as a linear superposition of the W_1 and W_2 fields of the forms: $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \pm iW_\mu^2)$. Expanding the above expression and including the Hermitian conjugate terms yields:

$$\frac{1}{4}\nu^2 g_W^2 (W_\mu^+ W^{\mu-}) + \frac{1}{8}\nu^2 (g_Y B_\mu - g_W W_\mu^3)(g_Y B^\mu + g_W W^{\mu3}).$$

The first term in this expression is identified as the mass term for the $W^\pm$ fields, corresponding to $m_W = \frac{1}{2}g_W\nu$. The final term in the expression concerns the Z^0 boson and the photon. It can be rewritten in matrix form and be diagonalized to give the massless photon state (A_μ) and the physical Z-boson state, with mass given by: $m_Z = \frac{1}{2}\nu\sqrt{g_Y^2 + g_W^2}$.

Fermion Masses

An additional attribute of the Higgs field is that it also gives rise to the fermion masses in a simple manner by introducing Yukawa mass terms into the Lagrangian. The terms can be written of the kind: $g_f \bar{\psi}_L \phi \psi_R$. Here g_f is the Yukawa coupling between the fermion and the Higgs field. After spontaneous symmetry breaking, in the ground state of the Higgs potential, this becomes $g_f \frac{\nu}{\sqrt{2}} \bar{\psi}_L \phi \psi_R$, resulting in a mass term for the fermion: $m_f = \frac{g_f \nu}{\sqrt{2}}$. The Yukawa couplings, hence the fermion masses, are not predicted in the SM and must be experimentally measured.

2.1.2 Higgs-Boson Production

The main production mechanisms at the LHC for the Higgs boson are:

- Gluon-Gluon fusion (ggF) - the process with the largest cross-section, $gg \rightarrow H^0 + X$ is mediated mostly by the exchange of

a virtual, heavy top quark. Contributions from lighter quarks propagating in the loop are suppressed proportional to m_q^2/m_t^2 , thus allowing to consider only the contributions of top and bottom quarks.

- Vector boson fusion production (VBF) - is the second-largest cross-section process at the LHC. Higgs production via VBF, $qq \rightarrow qqH^0$, proceeds by the scattering of two (anti-)quarks, mediated by t - or u -channel exchange of a W or Z boson, with the Higgs boson, radiated off the weak-boson propagator. Despite the smaller cross-section, VBF is an important mechanism of Higgs-boson production due to two forward jets which can be exploited to suppress relevant background processes.
- VH^0 associated production - Higgs Strahlung, $pp \rightarrow VH^0 + X$, $V = W^\pm, Z^0$. If an elementary fermion collides with an anti-fermion – e.g., a quark with an anti-quark, the two can form a virtual W or Z boson, which, if it carries sufficient energy, can then emit a Higgs boson. At the LHC, this process is only the third-largest.
- Higgs production in association with $t\bar{t}$ - Higgs radiation off top quarks, $pp \rightarrow t\bar{t}H$, provides a direct probe of the top-Higgs Yukawa coupling and was only recently measured by ATLAS [18] and CMS [19].
- Associated production with a single top quark - The Higgs production in association with a single top quark is one of only a few processes that can directly provide valuable information, in particular regarding the sign of the top Yukawa coupling. This sensitivity comes from the interference of the leading order Feynman diagrams where the Higgs boson couples either to the top quark or a W boson.

Figure 2.4 show representative diagrams for these dominant Higgs production processes.

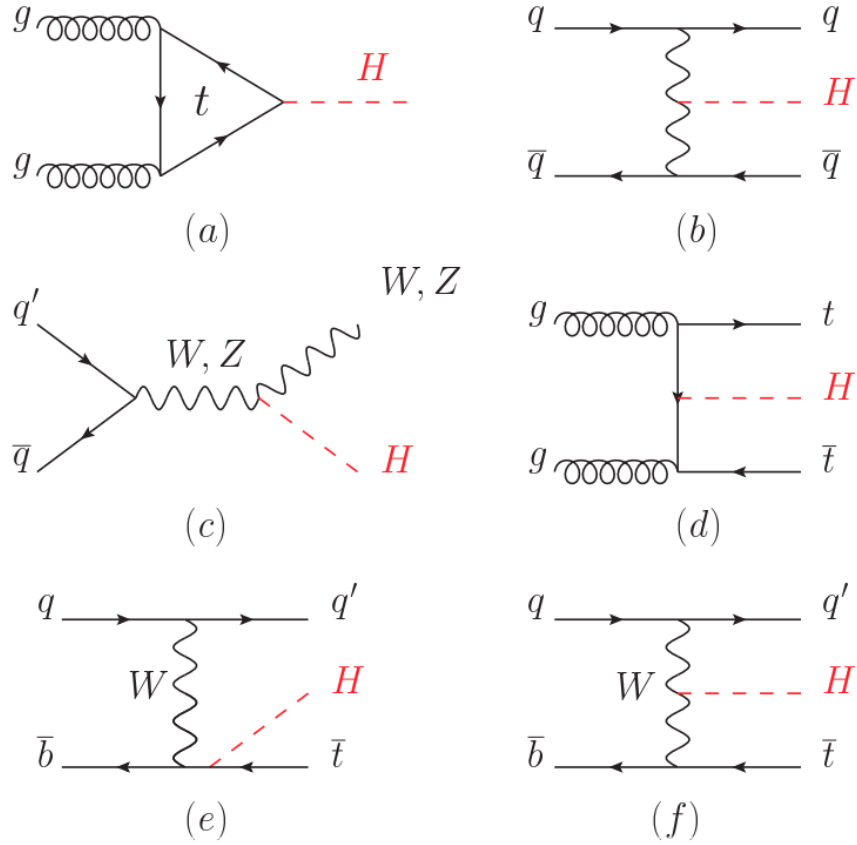


Figure 2.4: Leading order Feynman diagrams for the production mechanism of the SM Higgs boson in (a) gluon fusion, (b) Vector-boson fusion, (c) Higgs-strahlung (or associated production with a gauge boson), (d) associated production with a pair of top (or bottom) quarks, (e-f) production in association with a single top quark. with top quarks. Figure from [14].

2.1.3 Higgs-Boson Decay Modes

As laid out in the previous sections, the coupling strength of the Higgs boson is proportional to the mass of the interacting particles; heavier particles are preferred for the Higgs boson decay if the decay is kinematically allowed. Therefore, in the case of a Higgs boson decay into a massive vector boson pair, at least one of the bosons will be off-shell. Table 2.1 and Figure 2.5 summarise the various branching ratios for the Higgs boson with a mass of $m_H = 125\text{GeV}$.

In particular, the dominant decay modes are $H \rightarrow b\bar{b}$ and $H \rightarrow WW^*$, followed by $H \rightarrow gg$, $H \rightarrow \tau\bar{\tau}$, $H \rightarrow c\bar{c}$ and $H \rightarrow ZZ^*$. With much smaller rates, follow the Higgs boson decays into $H \rightarrow \gamma\gamma$, $H \rightarrow Z\gamma$ and $H \rightarrow \mu^-\mu^+$. The decay to massless particles like the photons or gluons arises from loop corrections, giving indirect information on the Higgs boson couplings to WW , ZZ and $t\bar{t}$ in different combinations.

A high branching ratio does not necessarily correspond to experimental sensitivity. The decay into bottom quark pairs is the dominant but also due to large QCD multijet background, difficulties in identification (tagging, see Section 6.4) at the LHC, the most challenging mode. The two channels that provided discovery for the Higgs boson were $\gamma\gamma$ and $ZZ(\rightarrow 4l)$. This can be understood by looking at the topology of the events having a clean signature within the detector. These two channels are considered high mass-resolution ones ($\sim 2\text{GeV}$) due to a possible complete mass reconstruction.

Channel	Branching Ratio
$b\bar{b}$	$58.2 \pm 1.3\%$
W^+W^-	$21.4 \pm 0.9\%$
gg	8.57 ± 0.87
$\tau\bar{\tau}$	$6.27 \pm 0.36\%$
$c\bar{c}$	$2.89 \pm 0.36\%$
ZZ	$2.62 \pm 0.11\%$
$\gamma\gamma$	$0.227 \pm 0.014\%$

Table 2.1: Branching ratios and their uncertainties for Higgs boson with a mass of $m_H = 125\text{GeV}$ [20]

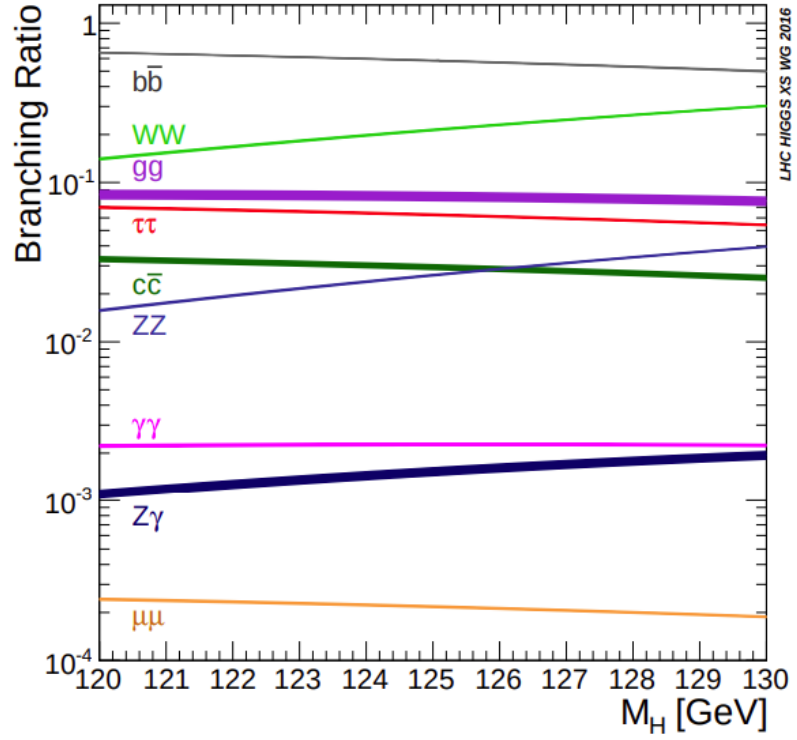


Figure 2.5: Higgs boson branching ratios and their uncertainties (expressed by the width of the curves) for the mass range around 125GeV [20].

2.2 THE 2-HIGGS DOUBLET MODEL

While the BEH mechanism is clearly the simplest one, there is no reason to believe that nature has to behave in such a manner. There are strong theoretical and experimental reasons to believe that the SM is just part of the truth, a yet to be completed theory.

A simple and natural extension of the SM is the inclusion of two Higgs doublets instead of one:

$$\phi_1 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix}, \quad \phi_2 = \begin{pmatrix} \phi_2^- \\ \phi_2^0 \end{pmatrix},$$

where ϕ_1 and ϕ_2 have positive hypercharge like the SM Higgs doublet and the superscripts $\pm$ and 0 denote the electric charge of the scalar fields ϕ . The set of models, which include two Higgs doubles, are therefore called 2-Higgs-Doublet Models (2HDMs) [21]. The most general 2HDM Higgs potential can be written [22]:

$$\begin{aligned} V_{2HDM} = & m_{11}^2(\phi_1^\dagger \phi_1) + m_{22}^2(\phi_2^\dagger \phi_2) - m_{12}^2(\phi_1^\dagger \phi_2 + \phi_2^\dagger \phi_1) + \frac{1}{2}\lambda_1(\phi_1^\dagger \phi_1)^2 \\ & + \frac{1}{2}\lambda_2(\phi_2^\dagger \phi_2)^2 + \lambda_3(\phi_1^\dagger \phi_1)(\phi_2^\dagger \phi_2) + \lambda_4(\phi_1^\dagger \phi_2)(\phi_2^\dagger \phi_1) + \frac{1}{2}\lambda_5[(\phi_1^\dagger \phi_2)^2 + (\phi_2^\dagger \phi_1)^2], \end{aligned}$$

where all the parameters are real. For a region of parameter space, the minimisation of this potential allows for the definition of v_1, v_2 as the vacuum expectation values and $\tan \beta \equiv \frac{|v_1|}{|v_2|}$ as the ratio between them. With two complex scalar SU(2) doublets, there are eight fields. From these eight degrees of freedom, three are getting "eaten" to give mass for the $W^\pm$ and Z^0 gauge bosons, and the remaining five are the physical scalar Higgs fields. This results in two charged scalars ($H^\pm$), two CP-even neutral scalars H^0 and h^0 , and one CP-odd pseudoscalar A .

Similar to the SM, the particle masses are free parameters of the theory and a particular choice of the masses m_h, m_H, m_A and $m_{H^\pm}$ of the physical particles, the ratio of the vevs, $\tan \beta$, the mixing angle α

and the soft breaking parameter m_{12} fix all parameters which are left in the 2HDM potential and therefore define a particular model.

The most general Yukawa Lagrangian for the interactions of the two Higgs doublets with the SM fermions can be written as:

$$\mathcal{L}_Y = - \sum_{i=1,2} \left(Q\tilde{\phi}_i y_i^u \bar{u} + Q\phi_i y_i^d \bar{d} + L\phi_i y_i^e \bar{e} + h.c. \right),$$

	Type I	Type II	Type III	Type IV
u	ϕ_2	ϕ_2	ϕ_2	ϕ_2
d	ϕ_2	ϕ_1	ϕ_2	ϕ_1
L_R^i	ϕ_2	ϕ_1	ϕ_1	ϕ_2

Table 2.2: Most common 2HDM types which lead to natural flavour conservation. The up and down type quarks fields are denoted by u and d , respectively. The right-handed lepton fields are denoted by L_R^i

A feature of all 2HDMs is the possibility of tree-level flavour-changing neutral currents (FCNCs), which are avoided in the SM. It is possible to remove FCNCs from the theory by assuming the discrete symmetry $\phi_1 \rightarrow \phi_1, \phi_2 \rightarrow -\phi_2$ for all right-handed quarks, which does not permit FCNCs at tree level since it forces any given type of fermions to couple to not more than one doublet (also known as Glashow-Weinberg condition) [23]. Regarding the quark sector of a 2HDM, there are only two possibilities to realise such a behaviour: either both, right-handed u - and d -type quarks, couple to the same doublet or they each couple to a different one, e.g. d -type to Φ_1 and u -type to Φ_2 .

Extending this symmetry to right-handed leptons, which are also allowed to couple to exactly one doublet to prevent FCNCs, gives two more possibilities for coupling fermions to Higgs doublets. Table 2.2 shows the possible combinations, usually denoted as Type I to Type IV.

2.3 PRODUCTION AND DECAY OF CHARGED HIGGS BOSONS

After the absorption of three degrees of freedom to give mass to W^+ , W^- and Z gauge bosons, two charged Higgs bosons ($H^\pm$) and three neutral Higgs bosons (h^0 , H^0 and A) arise in the spectra of the 2HDM. The production and decay of the charged Higgs boson depend on the parameters of the models in question and the charged Higgs boson mass, $m_{H^\pm}$. For a light H^+ ($m_{H^+} \lesssim m_{\text{top}}$), the dominant production mechanism is through $t\bar{t}$ production and the decay of a top quark into a charged Higgs boson and a b quark, $t \rightarrow bH^+$. If the H^+ is heavy, $m_{H^+} > m_{\text{top}}$, the relevant mass range for this analysis, the dominant H^+ production mode for most scenarios is in association with a top quark, as shown in Figure 2.6(a) and (b).

Two distinct modes exist, depending on whether the proton is assumed to carry b quarks as sea quarks or not. In the four-flavour scheme (4FS), the b -quarks are created from gluon splitting (four flavours in the proton), while in the five-flavour scheme (5FS), b -quarks are assumed to stem directly from the proton sea quarks (five flavours in the proton). In this thesis, the simulation is based on the 4FS only, and no further reference will be made to the 5FS.

Charged Higgs bosons with a mass lighter than the top quark mass can be dominantly produced at the LHC through gluon-gluon fusion in which top quark pairs are created with at least one top quark decay into a light charged Higgs boson instead of a W

Experimental data (through Deep Inelastic Scattering) revealed a rich structure of quarks, antiquarks and gluons that are far more abundant than the three valence quarks. These quarks and antiquarks are usually referred to as sea quarks.

$$gg \rightarrow t\bar{t} \rightarrow H^+ b W^- \bar{b}, H^+ b H^- \bar{b} \text{ and } W^+ b H^- \bar{b}.$$

Figure 2.7 show the branching ratios for the main decay channels of the charged Higgs. Charged Higgs bosons lighter than the top quark mass decay dominantly to $\tau\nu$. A charged Higgs boson that is heavier than the top quark will decay dominantly into tb .

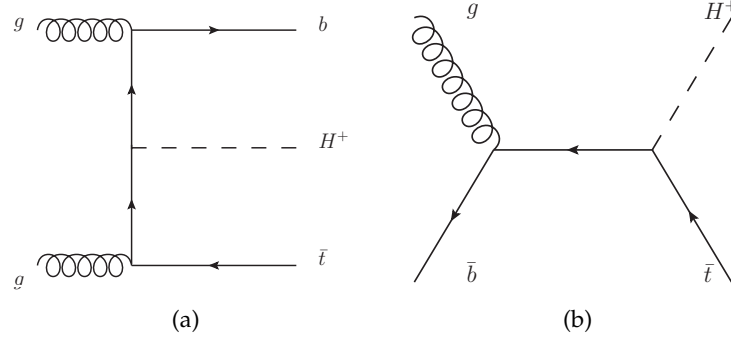


Figure 2.6: Feynman diagrams for the top-associated production of a heavy charged Higgs boson ($m_{H^+} > m_{\text{top}}$), in (a) the four-flavour scheme (4FS), and in (b) the five-flavour scheme (5FS).

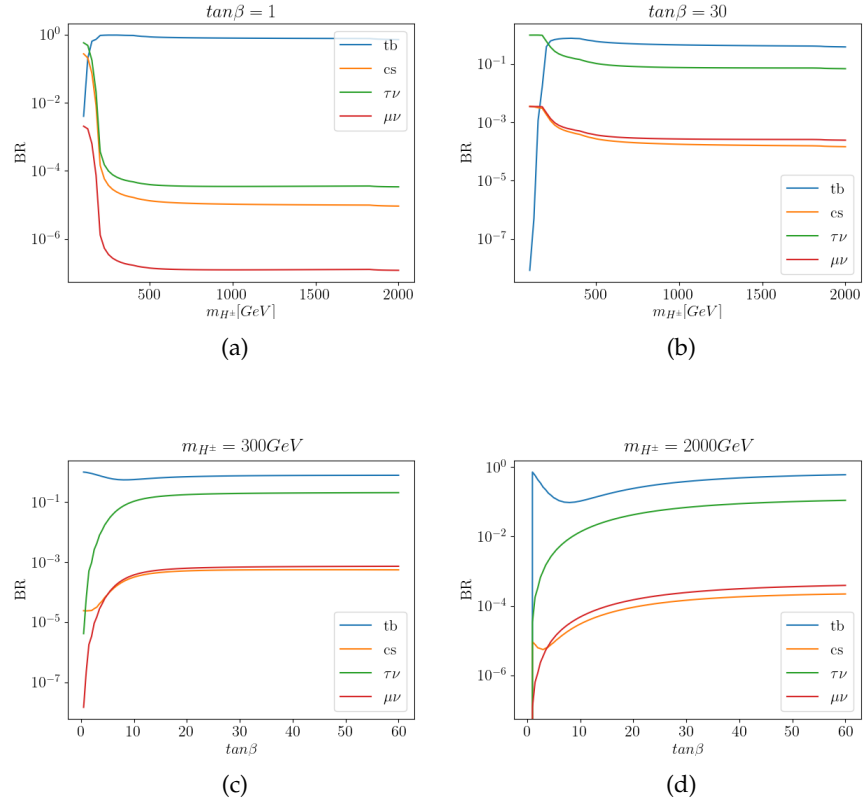


Figure 2.7: Branching ratios for a selected number of decay channels of the charged Higgs boson, shown as a function of mass, at (a) $\tan \beta = 1$, (b) $\tan \beta = 30$, and as a function of the $\tan \beta$ parameter, at (c) $m_{H^+} = 300$ GeV and (d) $m_{H^+} = 2000$ GeV. These figures have been produced with the *hplusmssm* package [24]

THE LHC AND THE ATLAS EXPERIMENT

3.1 THE LARGE HADRON COLLIDER

The proton-proton (pp) Large Hadron Collider (LHC) is the final stage of the accelerator complex at the European Centre for Nuclear Research (CERN), located near Geneva, on the Franco-Swiss border. It is a ring collider, 27 km long in perimeter, inside a tunnel, dug originally for the Large Electron Positron collider (LEP) in a depth that varies from 50 to 175 meters underneath the ground. The LHC and its associated accelerator chain culminate over 100 years of research into the production, acceleration, and focusing of charged particles. The LHC is the world's largest and highest energy accelerator, currently being upgraded after a successful run at 13 TeV. The LHC is capable of running in proton-proton, proton-ion, ion-proton, or ion-ion configurations.

To maximise the potential for new measurements and discoveries at CERN's LHC, the machine delivers as high as possible collision rates. As a direct consequence, multiple proton-proton collisions occur whenever two bunches of protons cross. Therefore, the "hard" collisions, meaning those with interesting physics processes, are contaminated by several "soft" ones. This is known as *pile-up* and introduces difficulties in reconstructing the final states as well as experimental uncertainties.

The dataset at $\sqrt{s} = 13$ TeV in the Run2 phase of the LHC corresponds to an integrated luminosity of about 156 fb^{-1} , see Table 3.1. The datasets effectively useful for analysis need to consider the data-taking efficiency with fully operational detectors and the data quality efficiency. The typical total inefficiency for both ATLAS and CMS

Most operations is dedicated to proton-proton collisions, which is the only type relevant to the work presented in this thesis.

Year	$\sqrt{s}$ TeV	$\int L dt$ fb ⁻¹	Period
2010	7	0.04	Run 1
2011	7	6.1	Run 1
2012	8	23.3	Run 1
2015	13	4.2	Run 2
2016	13	40.8	Run 2
2017	13	50.2	Run 2
2018	13	60.6	Run 2

Table 3.1: The LHC pp collision centre-of-mass energies and delivered luminosity.

experiments is approximately 10%, where approximately half is due to the data-taking inefficiency and a half from data quality.

The physics programme of the LHC is dominated by the four largest experiments: ALICE (A Large Ion Collider Experiment), LHCb (Large Hadron Collider beauty), CMS (Compact Muon Solenoid), and ATLAS.

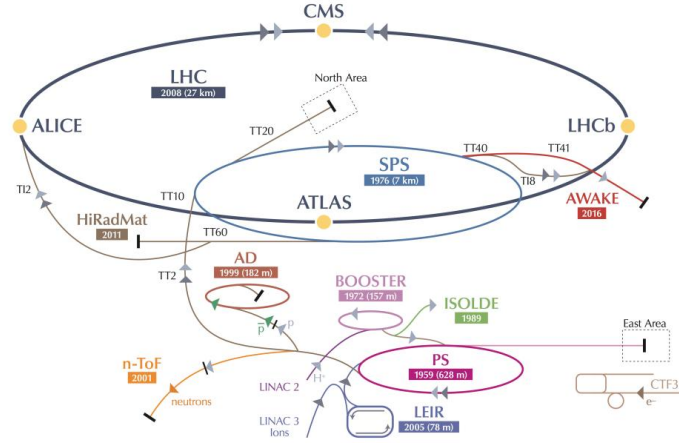


Figure 3.1: The CERN accelerator complex.

3.2 LHC EXPERIMENTS

ALICE [25] is a heavy ion experiment designed to study the conditions of the universe directly after the Big Bang. Heavy ion runs are used to

produce temperatures 100 times higher than inside the sun to produce quark-gluon plasma.

CMS [26] is in size the second biggest experiment at the LHC. It is a general purpose detector designed for searching the Higgs boson, for supersymmetry, dark matter and for precision measurements of the Standard Model. The detector is built inside a big 4 T solenoid magnet. CMS was the only detector at the LHC built over ground and installed in 15 slices in the CMS cavern. The detector system has a length of 21 m, a width of 15 m, a height of 15 m, and it weighs 12500 tons.

LHCb (Large Hadron Collider beauty) [27] is a detector system designed to study the B-mesons in detail. The focus is the study of the matter-antimatter symmetry and rare decays of B-mesons, which are sensitive to new physics beyond the Standard Model. It is the smallest detector at the LHC built as a single arm spectrometer detecting particles propagating in the forward direction. The experiment does not use the full luminosity of the LHC to avoid pile-up.

ATLAS (A Toroidal LHC ApparatuS) [28] is the largest detector at the LHC. It is like CMS, a general purpose detector with the same physics goals. More details about the ATLAS detector system, which was used to collect the data used in this thesis, are described in the next section.

3.3 THE ATLAS DETECTOR

3.3.1 Overview

The ATLAS detector is centred on one of the LHC collision points. As shown in Figure 3.2, ATLAS is over 24 meters high and almost 45 meters long. It weighs approximately 7000 tons. ATLAS is built around the LHC beam pipe, 100 meters underground. The beam pipe is centred on the cylindrical axis of symmetry of the detector. ATLAS offers a rich set of opportunities for new physics and a possibility to study strongly interacting matter at high energy densities.

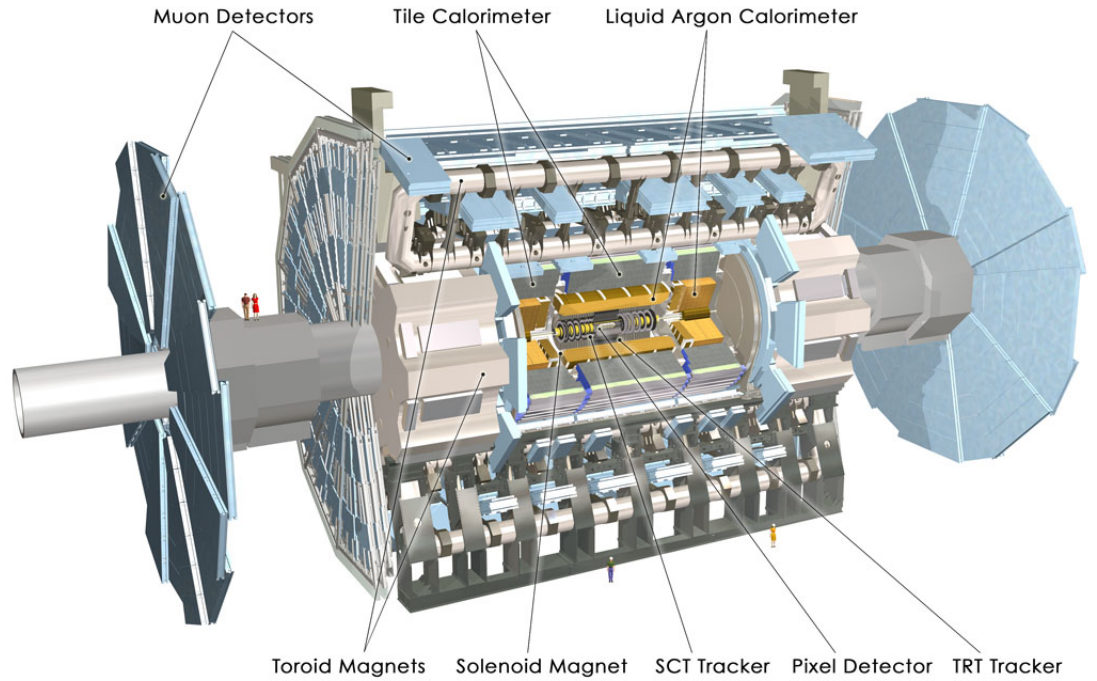


Figure 3.2: Cut-away view of the ATLAS detector.

Particles produced in the collisions emerge from the centre of the detector in all directions. ATLAS has been designed to record the paths and energies of the particles emerging from the collisions. ATLAS is composed of a series of concentric sub-systems, each sensitive to different types of particles produced in the collisions. The Inner Detector (ID) [29, 30] is closest to the interaction point and measures

the trajectories of charged particles. The ID is composed of the Pixel Detector [31, 32], the Semiconductor Tracker (SCT) [33–35], and the Transition Radiation Tracker (TRT) [36–38]. The ID operates in a 2 Tesla magnetic field provided by the solenoid magnet [39]. Surrounding the ID is the calorimeter system [40]. The calorimeter system is composed of liquid argon electromagnetic calorimeters [41], the tile calorimeters [42], the liquid argon hadronic end-cap calorimeters, and the forward calorimeters. These are indicated in Figure 3.2. The calorimeters are designed to measure the energy of electrons, photons, and hadrons.

The Muon Spectrometer (MS) [43] surrounds the calorimeters. The calorimeter system stops all particles except muons and neutrinos. The MS is designed to measure the trajectories of muons leaving the calorimeter. The MS is composed of muon chambers operating in a magnetic field, provided by the toroid magnetics [44, 45].

3.3.2 Conventions

A common coordinate system is used throughout ATLAS. The interaction point is defined as the origin of the coordinate system. The z -axis runs along the beamline. The $x - y$ plane is perpendicular to the beamline and is referred to as the transverse plane. Particle momenta measured in the transverse plane are referred to as the transverse momenta, p_T . The positive x -axis points from the interaction point to the centre of the LHC ring; the positive y -axis points upward to the earth's surface. The detector half at positive z -values is referred to as the "A-side", the other half the "C-side". The transverse plane is often described in terms of $r - \phi$ coordinates. The azimuthal angle ϕ is measured from the x -axis around the beam. The radial dimension, r , measures the distance from the beamline. The polar angle θ is defined as the angle from the positive z -axis. The polar angle is often reported in terms of pseudorapidity, defined as $\eta = -\ln \tan(\theta/2)$. The distance ΔR is defined in $\eta - \phi$ space as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. Transverse

quantities like transverse momentum p_T , transverse energy E_T and missing transverse energy E_T^{miss} are defined in the $x - y$ plane. The 4-momentum vector $\mathbf{P} = (E, p_x, p_y, p_z)$ can be expressed in terms of E, η, ϕ and p_T by $\mathbf{P} = [E, p_T \cos(\phi), p_T \sin(\phi), p_T \sinh(\eta)]$ for massless particles

3.3.3 Inner Detector

The inner detector (ID) is designed to precisely measure tracks of charged particles within $|\eta| < 2.5$. As it is immersed in a 2 T solenoidal magnetic field, the trajectories are curved, allowing to measure the momentum of charged particles. By extrapolating the tracks, the inner detector reconstructs the primary vertex and secondary vertices, which originate from the decay of long-lived particles, such as B-mesons. The inner detector provides electron identification for electrons with $0.5 \text{ GeV} < p_T < 150 \text{ GeV}$ within $|\eta| < 2.0$. The ID consists of three subsystems: the silicon pixel detector, the silicon microstrip tracker and the transition radiation tracker. Figure 3.3 shows a cut-away of the ID barrel, and Figure 3.4 shows a cut-away of one of the ID end-caps. The total length of the inner detector is about 6.2 m, and its diameter is about 2.1 m.

The silicon pixel detector is the innermost subsystem and is structured in three cylindrical layers in the barrel and three discs as end-caps on either side. It comprises about 80.4 million pixels with a minimum pixel size of $50 \times 400 \mu\text{m}^2$, which have an intrinsic accuracy of $10 \mu\text{m}$ in the $R - \phi$ transverse plane and of $115 \mu\text{m}$ in the z -direction (R -direction) for the barrel (for the discs). About 80.4 million channels are necessary for readout.

The silicon microstrip detector (or semiconductor tracker, SCT) is the intermediate subsystem. It is structured in four stereo layers of silicon microstrips in the barrel and nine discs as end-caps on either side. Each layer consists of two layers of silicon microstrips with a mean

pitch width of about $80\ \mu\text{m}$. The two layers are arranged back-to-back with an angle of $40\ \text{mrad}$ between them, allowing the determination of the coordinates ϕ and z (ϕ and R) in the barrel (end-caps). The SCT provides an intrinsic accuracy of $17\ \mu\text{m}$ in $R - \phi$ and of $580\ \mu\text{m}$ in the z -direction (R -direction) for the barrel (for the discs). About 6.3 million channels are necessary for readout.

The transition radiation tracker (TRT) is the outermost subsystem and provides measurements within $|\eta| < 2.0$. It consists of drift tubes with a diameter of $4\ \text{mm}$ and a length of $144\ \text{cm}$ in the barrel and $37\ \text{cm}$ in the end-caps. In the barrel, the drift tubes are arranged parallel to the beam axis, while in the end-caps, they point radially to it. The tubes are filled with a gas mixture of 70% xenon, 27% carbon dioxide and 3% oxygen. The produced charges are collected on a gold-plated tungsten wire acting as the anode. The space between drift tubes is filled with polypropylene material, which causes particles to emit low energy transition photons, detected then by the drift tubes. Since the transition radiation is proportional to $1/m$, where m is the particle's mass, light particles such as electrons produce more transition radiation than charged hadrons, allowing electrons to be separately identified.

3.3.4 Calorimeters

The calorimeter system measures the energy of hadrons, electrons and photons. It provides coverage up to $|\eta| = 4.9$, using several different technologies. An overview of the calorimeter system is shown in Figure 3.5. The calorimeter system provides containment for both electromagnetic and hadronic showers, stopping particles before reaching the muon system.

The ATLAS calorimeters are a type known as "sampling" calorimeters. Incident particles produce showers of energy in the calorimeter. Only a fraction of the energy produced by the particle is measured by

Muons leave a small fraction of energy in the calorimeters as well. But, their energy cannot be said to be measured by the calorimeters. It is the combined measurement alongside the spectrometer's data that provides the full muon energy.

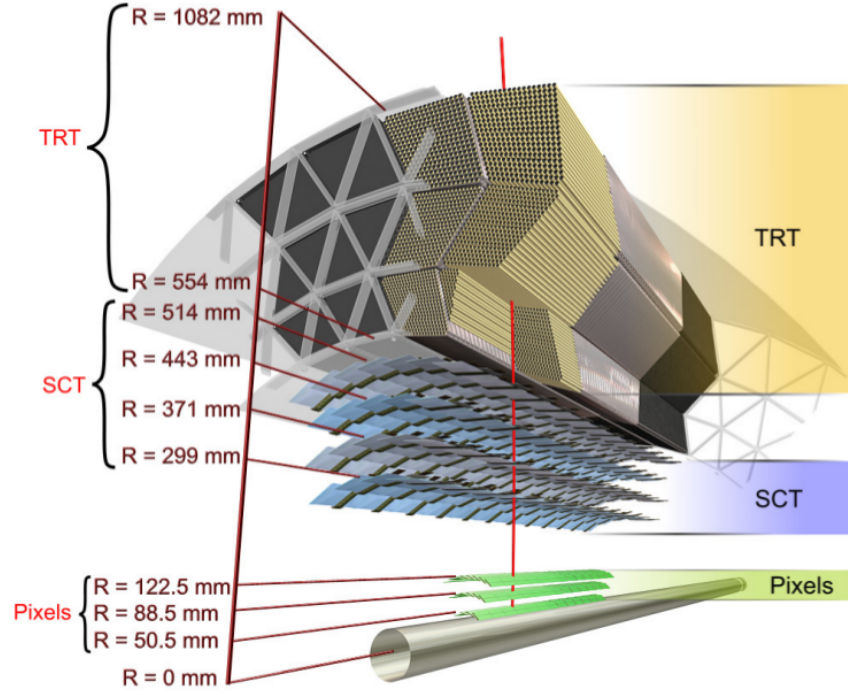


Figure 3.3: Drawing showing the detector elements crossed by a charged particle (red line) with 10 GeV p_T in the barrel of the Inner Detector. The particle emerges from the interaction point and traverses the beam-pipe, three pixel layers, four double layers of SCT sensors, and around 35 TRT straws

active detector sensors. The energy of the full shower can be inferred from the observed energy.

The energies of electrons and photons are measured by the liquid-argon (LAr) electromagnetic (EM) barrel and end-cap calorimeters. The EM calorimeter is a lead-LAr detector with a specialized geometry that provides complete and uniform ϕ coverage and fast readout. These detectors provide high granularity measurements, critical for particle identification in the range $|\eta| < 2.5$. The EM calorimeter is segmented into three radial sections with different $\eta - \phi$ granularities. Figure 3.6 shows a cutaway of the different layers in the EM barrel calorimeter. The first layer, referred to as the "strips", provides very fine segmentation in η . The strips can separate between showers initiated by electrons or photons and showers initiated by neutral pions. The second sampling provides the most energy measurement and has fine segmentation in both η and ϕ . The third sampling is coarser and

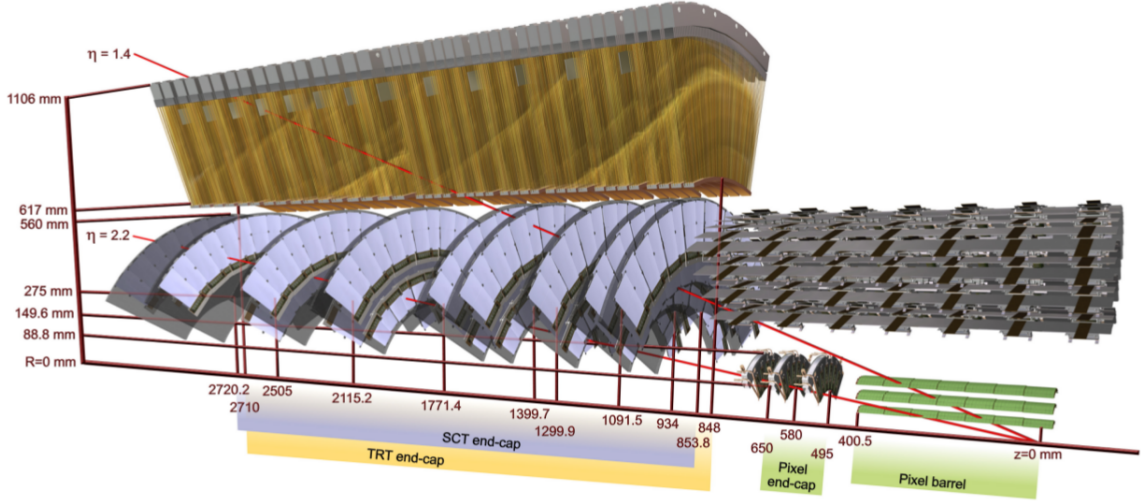


Figure 3.4: Drawing showing the detector elements crossed by two charged particles (red lines) of 10 GeV p_T in the end-cap of the Inner Detector. A particle at $|\eta| = 1.4$ traverses the beam-pipe, three pixel layers, four SCT disks with double layers of sensors, and approximately 40 straws in the TRT end-cap. A particle at $|\eta| = 2.2$ traverses the beam-pipe, only the first layer of the pixel detector, two end-cap pixel disks and the last four disks of the SCT end-cap. The coverage of the end-cap TRT does not extend beyond $|\eta| = 2$.

adds additional depth to the calorimeter. The EM calorimeters cover the pseudorapidity range $|\eta| < 3.2$.

The Tile calorimeters and the LAr hadronic end-cap calorimeter are designed to measure the energy of hadrons. The range $|\eta| < 1.7$ is covered by the Tile calorimeter. The scintillator-tile calorimeter is separated into a barrel and two extended barrel cylinders. In the end-caps, $1.5 < |\eta| < 3.2$, LAr technology is used for the hadronic calorimeters. The LAr forward calorimeters provide both electromagnetic and hadronic energy measurements and extend the pseudorapidity coverage to $|\eta| = 4.9$.

3.3.5 Muon Spectrometer

In the ATLAS muon system, the momentum of muons is determined in a large superconducting air-core toroid magnet system. It contains detector technologies for fast triggers and high precision tracking with $60 \mu m$ intrinsic resolution. Emphasis is put on reliable, high resolution,

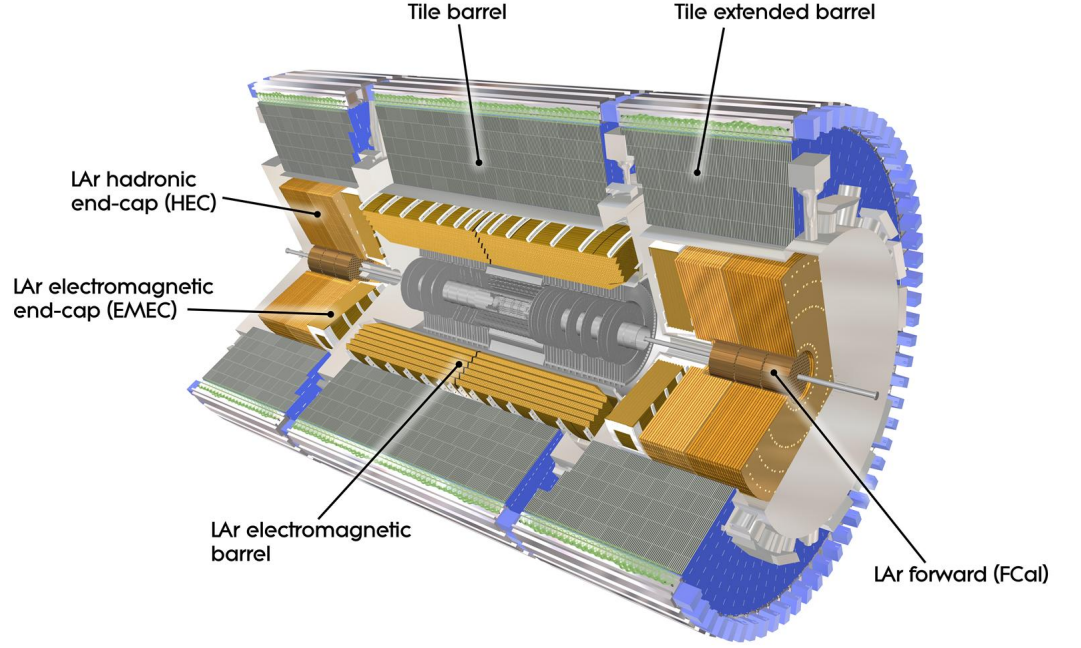


Figure 3.5: Cut-away view of the ATLAS calorimeter system.

stand-alone performance over a p_T range from 5 GeV to 1000 GeV. Good momentum resolution is needed, above large backgrounds, for the detection of decays involving muons. The resolution can be further improved by combining the stand-alone results of the muon spectrometer with these of the inner detector. An example is a light charged Higgs boson searches, where the decay product has one or two leptons. A schematic view of the muon system is shown in Figure 3.7.

A large central toroid magnet bends the muons in the region $|\eta| < 1.4$. In the high- η regions ($1.6 < |\eta| < 2.7$), two smaller toroidal superconducting magnets are inserted. Each of the three magnets is made from eight coils which are placed symmetrically around the beam axis. Magnetic bending power of 1.4 to 5.5 Tm is provided by the barrel toroid in the range $|\eta| < 1.4$. In the endcap region, a bending power of 1 to 7.5 Tm is provided by the end-cap magnets. The region $1.4 < |\eta| < 1.6$ is called the transition region, where the magnetic fields overlap. Muon chambers arranged in three cylindrical layers

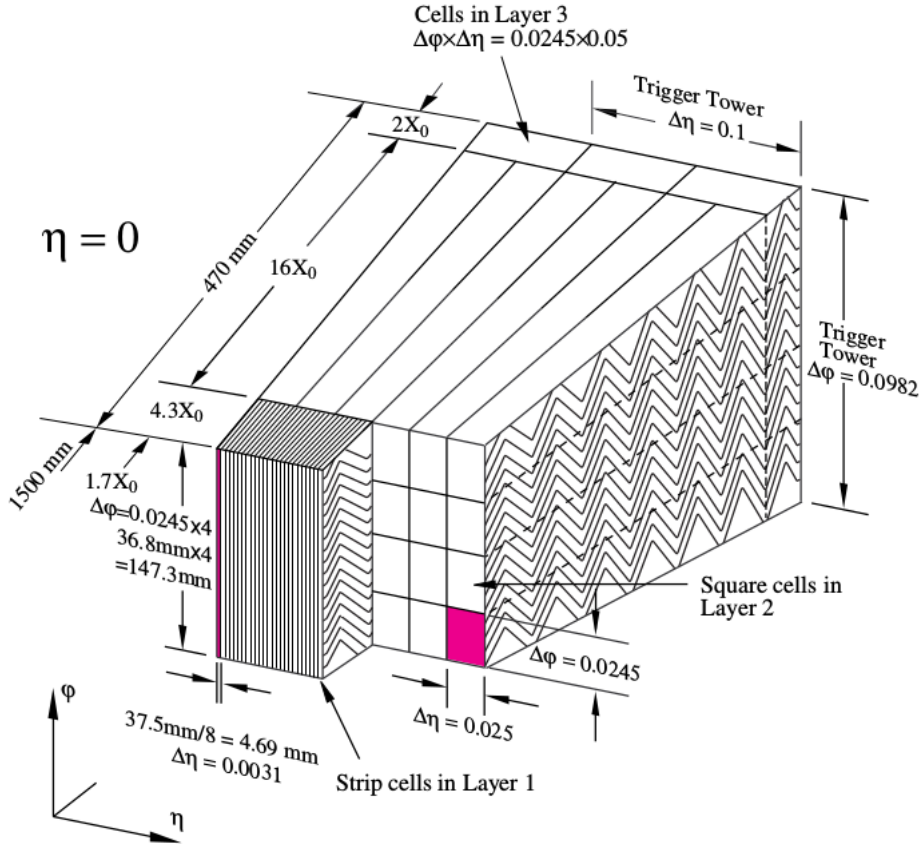


Figure 3.6: Sketch of a section of the LAr EM barrel where the different layers are clearly visible. The granularity in η and ϕ of the cells of each of the three layers is shown.

around the beam pipe are used in the barrel region to measure muon tracks. The transition and end-cap regions are covered by three layers of muon chambers arranged along ϕ , meaning the centre point of the tubes are oriented tangentially to the circle around the beam axis.

The region with $|\eta| < 2.7$ is covered by **Monitored Drift Tubes** (MDT). In the high- η region, the rates and background conditions are expected to be up to 30 kHz per tube at full LHC luminosity. The MDTs, filled with Ar/CO_2 gas, have a diameter of 29.97 mm, providing an average resolution of 60-80 μm . The maximal count rate of the MDT is 150 Hz/cm^2 .

Cathode Strip Chambers (CSC) are used in the region $2 < |\eta| < 2.7$ where the particle flux and the density of tracks are highest. The CSC can handle counting rates up to 1000 Hz/cm^2 . The CSC system is made of two disks of eight chambers, which enable the measurement

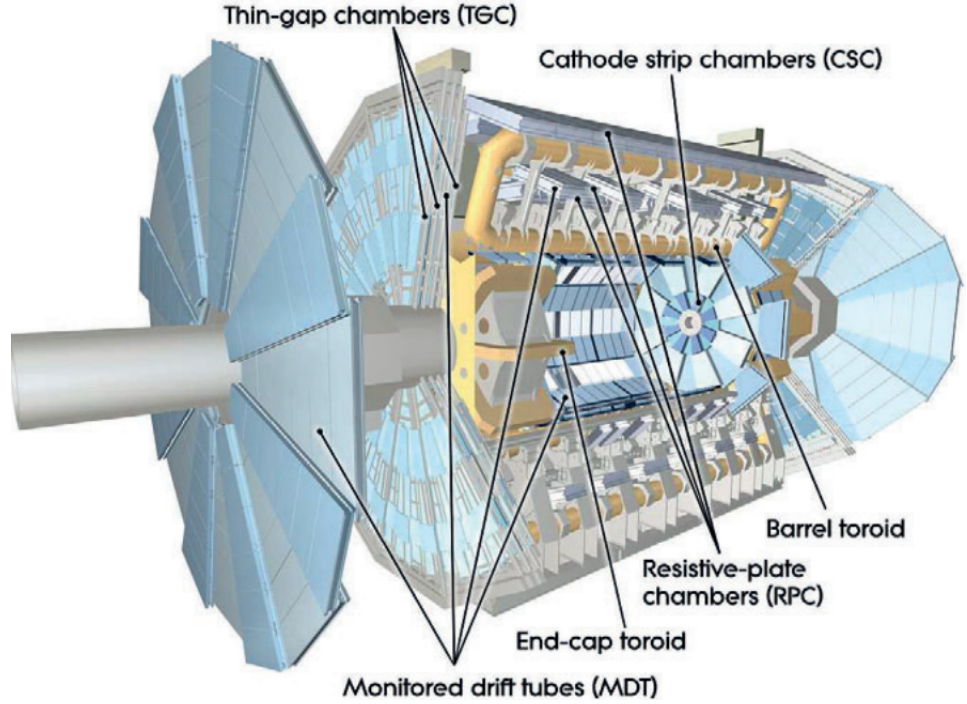


Figure 3.7: Overview of the ATLAS muon system

of η and ϕ coordinates. The CSC is filled with an Ar/CO_2 mixture. The resolution reached is $60 \mu m$.

Trigger chambers provide fast information to the Level 1 trigger logic about muon transverse momentum and bunch-crossing identification. The trigger chambers measure the ϕ coordinate to give a full space trajectory in combination with the η coordinate of the MDTs. The region $|\eta| < 2.4$ is covered using two different types of technologies. In the barrel region $|\eta| < 1.05$, **Resistive Plate Chambers (RPC)** are used and, in the region $1.05 < |\eta| < 2.4$, **Thin Gap Chambers (TGC)** are used.

3.3.6 Trigger System

The ATLAS detector in Run 2 has a three-level trigger system: Level 1 (L1), Level 2 (L2) and event filter (EF) level. The levels are staged, meaning that the next level refines the event selection from the previous level by applying more precise or additional requirements. The

Level 1 trigger uses high transverse momentum muons, electrons, photons, jets, τ -leptons, large missing and total transverse energies for triggering. Only reduced detector information is used to decide with a latency shorter than $2.5 \mu\text{s}$. The L1 trigger defines Regions Of Interest (ROI) that seed the L2 trigger. At L2, the full precision of all detector data in the ROI is used. The L2 trigger is configured to reduce the data rate to approximately 3.5 kHz with a latency of less than 40 ms per event. At the EF level, the data rate is further reduced to about 200 Hz for storage and offline analysis. The EF uses advanced near offline quality algorithms for processing and requires around 4 seconds per event. The stored data size is approximately 1.3 MB per event. A schematic layout of the trigger system is shown in Figure 3.8.

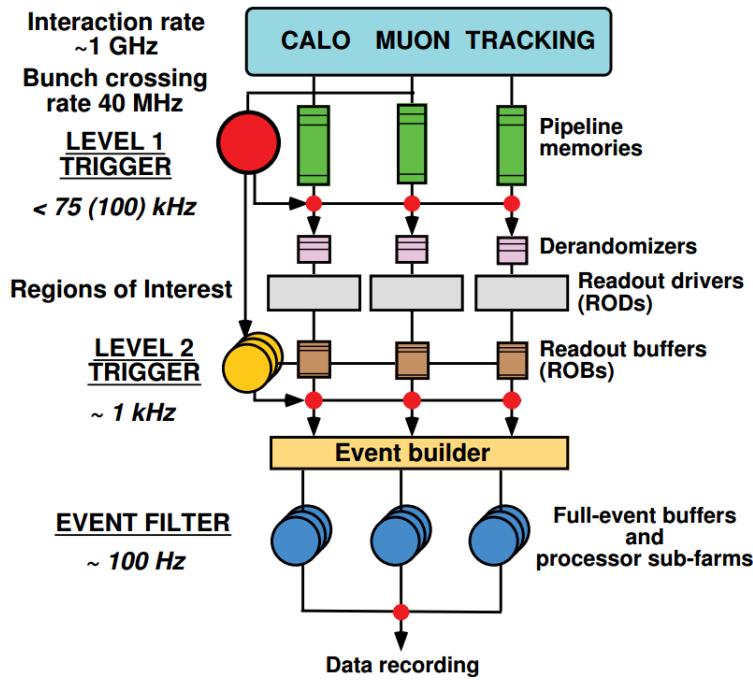


Figure 3.8: Schematics layout of the trigger system

STATISTICAL FORMALISM

A fundamental quest of High Energy Physics is the search for new physics processes that have been predicted but not yet experimentally seen. One constructs a physics model describing known processes, referred to as background, and potential non-SM processes, referred to as signal, in order to statistically test it against data. Under a given hypothesis, the statistical analysis of the model against data allows us to make a statement either of discovery or exclusion of new physics at a certain confidence level. In the following, the statistical formalism is used for the $H^\pm \rightarrow tb$ and $VH \rightarrow Vcc$ analysis to search for SM and new physics phenomena following References [46, 47].

4.1 BINNED PROFILE LIKELIHOOD

Consider an experiment that measures an observable (e.g. the H^0 candidate mass) for each selected event. The data can be represented as a binned histogram in designed regions, e.g. in the signal region denoted as SR. The expected event yield in the i -th bin can be expressed as

$$n_i^{exp} = \mu s_i + b_i, \quad (4.1)$$

where b_i denotes the expected background yield, s_i the expected signal yield in the i -th bin and μ denotes the signal strength defined as the ratio of the signal cross-section to the reference signal cross-section. The signal rate or signal cross-section is proportional to the signal strength. The value $\mu = 0$ corresponds to the background-only

One should expect $\mu = 1$ for a SM process.

hypothesis (or null hypothesis), describing only known processes, and $\mu = 1$ to the nominal signal hypothesis, describing background and the sought after signal.

The definition of the likelihood function is given in the following step-by-step. The foundation is a counting experiment in each bin using a set of Poissonian probability terms.

$$\mathcal{L}(\mu) = \prod_{i \in \text{bins}} \text{Pois}(N_i | \mu s_i + b_i) = \prod_{i \in \text{bins}} \frac{(\mu s_i + b_i)^{N_i}}{N_i!} e^{-(\mu s_i + b_i)}, \quad (4.2)$$

where N_i is the number of measured data events in bin i and s_i and b_i are expected signal and background counts, estimated from Monte Carlo simulation. The parameter μ is a common scale factor to the signal yield in all bins. It can be understood as the ratio of the measured $\sigma \times \text{BR}$ for the signal process divided by its (SM) expectation. One of the goals of the fit is the determination of this parameter, also known as the *Parameter of Interest* (PoI), and its uncertainty.

Systematic uncertainties are parameterised as a set of nuisance parameters (NPs) $\vec{\theta}$ that modify the expected signal and background yields, i.e. $\{s_i, b_i\} \rightarrow \{s_i(\vec{\theta}), b_i(\vec{\theta})\}$. Their effect is determined beforehand in auxiliary measurements and implemented in the likelihood via a set of Gaussian penalty terms $\mathcal{L}_{\text{Aux}}$ with mean zero and a variance of unity.

$$\mathcal{L}_{\text{Aux}} = \prod_{\theta \in \vec{\theta}} \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2}. \quad (4.3)$$

These terms constrain each NP's effect to the certainty with which its nominal value is known. The NPs with no corresponding Gaussian constraint (usually normalisation factors) are called *free-floating* with their value being solely determined from data in the fit. The problem of limited simulation statistics for the background processes is addressed by introducing γ -parameters γ_i in each bin that are allowed

to modify the expected background yield as $b_i(\vec{\theta}) \rightarrow \gamma_i b_i(\vec{\theta})$. They are constrained by multiplying the likelihood with

$$\mathcal{L}_{\text{BkgStat}}(\vec{\gamma}) = \prod_{i \in \text{bins}} \text{Gauss}(\beta_i | \gamma_i \beta_i, \sqrt{\gamma_i \beta_i}), \quad (4.4)$$

with

$$\beta_i = \frac{1}{\sigma_{\text{rel}}^2}, \quad (4.5)$$

where σ_{rel} is the relative statistical uncertainty on the expected total background yield.

The full likelihood can be (schematically) written as

$$\mathcal{L}(\mu, \vec{\theta}, \vec{\gamma}) = \prod_{i \in \text{bins}} \text{Pois}(N_i | \mu s_i(\vec{\theta}) + \gamma_i b_i(\vec{\theta})) \times \prod_{\theta \in \vec{\theta}} \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2} \times \prod_{i \in \text{bins}} \text{Gauss}(\beta_i | \gamma_i \beta_i, \sqrt{\gamma_i \beta_i}). \quad (4.6)$$

A maximum likelihood procedure is used to determine the point of interest (POI) and its uncertainty.

4.2 TEST STATISTIC

A Test-Statistic can be considered as a numerical summary of a given data-set. Such a reduction of the data to a singular value can be used in order to perform hypothesis testing. A proper test-statistic would allow us to quantify a characteristic that would distinguish the null (background only) from the alternative (signal and background) hypothesis. The used test-statistic is the log-likelihood ratio, constructed as

$$q_\mu = 2 \ln \left(\frac{\mathcal{L}(\mu, \hat{\hat{\theta}}, \hat{\hat{\gamma}})}{\mathcal{L}(\hat{\mu}, \hat{\hat{\theta}}, \hat{\hat{\gamma}})} \right), \quad (4.7)$$

where $\{\hat{\mu}, \hat{\hat{\theta}}, \hat{\hat{\gamma}}\}$ are the values that maximise the likelihood overall and $\{\hat{\hat{\theta}}, \hat{\hat{\gamma}}\}$ the ones that maximise $\mathcal{L}$ for a fixed value of μ . The profile

This formalism allows for the treatment of systematic uncertainties through the use of the profile likelihood ratio.

*A null hypothesis
can be one with
 $\mu = 0$, thus, a
background only
hypothesis.*

*PDF in HEP could
also mean parton
distribution function,
such distinction shall
be made when
necessary.*

likelihood ratio takes values of $0 \leq \lambda \leq 1$, with λ near 1 implying good agreement between the data and hypothesised value of the signal strength μ .

The *p-value* is used to quantify the compatibility of the background-only hypothesis ($\mu = 0$) with the fitted data. A smaller *p-value* indicates an increasing incompatibility between the data and the hypothesis. Under the signal and background hypothesis, the *p-value*, p_{s+b} , is defined as the probability of obtaining a value of the test-statistic that equals or rarer than the observed test-statistic q^{obs} . The asymptotic approximation of a non-central χ^2 distribution for the probability distribution function (PDF) of the test-statistic [46] is used leading to the identification $p_\mu = \sqrt{q_\mu}$.

According to the convention, the results are quoted in terms of (Gaussian) standard deviations Z , calculated from the *p-value* as

$$Z = \Phi^{-1}(1 - p), \quad (4.8)$$

where Φ^{-1} is the inverse of the cumulative distribution of a Gaussian distribution.

4.3 ASIMOV DATASET

It is often helpful to quantify the sensitivity of an experiment by giving the expected significance obtained from a given measurement under the assumption of different hypotheses. For example, the sensitivity to detecting a given signal process H_1 could be characterized by the expected value, assuming H_1 , of the value of Z obtained by a test of H_0 .

In the SM, there is only one Higgs Boson with well-defined couplings. To find the detection sensitivity, one must generate an ensemble of experiments containing the Higgs Boson at the tested mass. However, if one goes beyond the SM, e.g., for SUSY models, one faces a multidimensional parameter space in which the couplings of the Higgs

Boson, thus, their production cross-section and decay properties, vary as a function of the parameters. For each point in the parameter space, one has to estimate the detection sensitivity of the experiment. One is faced with the necessity of generating an enormous number of ensembles of experiments and evaluating the average sensitivity for each ensemble. In [46], it was shown that one could replace each ensemble of alternative-hypothesis experiments with a dataset representing the typical experiment. This "Asimov" dataset provides the desired average sensitivity. Thus, one is exempted from the need to perform an ensemble of experiments for each parameter set. The Asimov data set is constructed so that one will obtain the true parameter values if one uses them to evaluate the estimators for all parameters.

4.4 CLS LIMITS

With the definition of p_{s+b} , which quantifies the disagreement between the data and the signal plus background hypothesis, one may exclude a signal model at confidence level of $1 - \alpha = 95\%$ given $p_{s+b} < \alpha = 0.05$. However, such exclusion criteria may lead to the exclusion of signal model that the experiment has little or no sensitivity. If a search has little sensitivity to a specific signal model, the number of expected signal yields is minimal with respect to the expected background yield. In such case, the distributions of the test-statistic for both hypotheses (null and alternative) are closer together, having a large overlap. If the observed number of events has a sufficiently large downward fluctuation below the expected background, p_{s+b} decreases so that the analysis may exclude a signal model to which it has no or little sensitivity. In order to solve this, the CL_S probability is defined as:

$$CL_S \equiv \frac{p_{s+b}}{1 - p_b}. \quad (4.9)$$

In the scenario of a downward fluctuation, if an analysis is sensitive to a signal model, the two test-statistic distributions are well separated,

Using a single representative individual to stand in for the entire population can help evaluate the sensitivity of a statistical process. Franchise, a science fiction short story by Isaac Asimov, has been cited as the inspiration of the term "Asimov dataset" in which a single representative one can replace an ensemble of simulated experiments.

CL_S is a conservative limit setting.

$1 - p_b$ is small, but p_{s+b} is much smaller, yielding $CL_S \ll 1$. On the other hand, if a search has little sensitivity to a given signal model, the two distributions are then close together, so CL_S increases, reaching values of order 1, which in turn avoids exclusion of the signal model.

The upper limit on the signal strength μ , which can be translated into an upper limit on a cross-section, is the largest μ for which $CL_S(\mu) \geq \alpha$.

4.5 BLINDING AND UNBLINDING ANALYSIS

Blinding is a crucial tool of the scientific method. Its purpose is to reduce as much as possible the scientist confirmation bias (sometimes subconsciously), or discovery bias in this case.

This is widely known as a Double-blinded experiment where both the researcher and the participant are unaware of the treatment, usually placebo vs test drug.

ATLAS (and the rest of CERN's experiments) performs searches for new particles by looking for the slightest signal among a background of known physical processes. In case the data begin to indicate something more interesting than the mere background, then it is crucial to make sure that such observation is statistically significant.

There is a strong human tendency to perfect one's analysis based on the data already observed. A researcher working on discovering a new particle might be biased towards a specific region in the mass range due to previously known results and will work extensively to maximise (or minimise, depending on one's personal tendency) any sign of new physics. To avoid such bias while analysing new data, physicists "blind" themselves at certain regions, meaning one does not calculate any data-dependent observable at these regions .

Regions can be either mass ranges or selection cuts that maximise signal purity vs background.

Once the analysis is blinded, researchers optimise both the MC simulations and the methods of characterizing the irreducible background processes using the data from regions where no signal should be present.

The unblinding is done once the analysis is certain in the full chain of production (e.g. background estimation, selection cuts and systematics

treatment), only then, and when no further changes are allowed, the "curtain" is lifted, and the final result is calculated.

SEARCH FOR CHARGED HIGGS BOSON DECAYING TO TOP AND BOTTOM QUARKS

5.1 INTRODUCTION

After the discovery of a Higgs boson, h , with a mass of around 125 GeV consistent with the SM at the LHC in 2012 [4, 5, 48], the question of whether this is the only Higgs boson, or the first observed physical state of an extended Higgs sector, is a key issue. No charged scalar bosons exist within the SM. Still, many beyond the Standard Model (BSM) scenarios contain an extended Higgs sector with at least one set of charged Higgs bosons, $H^\pm$, in particular 2HDMs (Section 2.2) and models containing Higgs triplets [49–53].

The existence of a charged scalar boson would thus be a clear sign of BSM physics. Charged Higgs bosons are predicted by several BSM theories. Of importance for this analysis is the class of BSM scenarios in 2HDM [54–57], as well as models containing Higgs triplets [49–53].

In the following sections, the previous and future searches are described. The analysis strategy will be laid out with a detailed description of the samples and objects being used. Last, the single-lepton channel will be presented, followed by in-depth background suppression and final results.

Charged Higgs bosons are always denoted by H^+ , with the charge-conjugate H^- implied. The same goes for leptons, quarks and antiquarks q and $\bar{q}$ charges.

5.2 PREVIOUS AND FUTURE RESULTS

Previous searches for charged Higgs bosons have been done at the LEP experiments, where upper limits on the H^+ production in the mass range 40–100 GeV [58] were set. The Tevatron experiments CDF and

D0 set upper limits on $\text{BR}(t \rightarrow bH^+)$ for $80 < m_{H^+} < 150$ GeV [59, 60]. The D0 experiment made a dedicated search for the tb final state, for masses $180 < m_{H^+} < 300$ GeV [61]. The CMS experiment has investigated the tb final state in 8 TeV pp collisions and the mass range $180 < m_{H^+} < 600$ GeV [62]. The ATLAS experiment performed a search in 8 TeV analysis [63], with an observed (expected) upper limit on $\sigma \times \text{BR}$ for $H^+ \rightarrow tb$ between 6 and 0.2 pb (4 and 0.2 pb) in the mass range 200–600 GeV.

At the time this analysis is performed, the most recent published ATLAS result is the preliminary 13 TeV study for ICHEP 2016 [64]. This is the benchmark study and will be referred to as the “ICHEP” result throughout the chapter. These chapter results were published as ATLAS publication [9].

*The author was also
part of the ICHEP
result*

The scientific effort on the $H^\pm \rightarrow tb$ channel still continues, using tools developed in this work. Preliminary results of the follow-up study, using Run 2 dataset of 139 fb^{-1} are published at [65].

5.3 ANALYSIS STRATEGY

The analysis searches for the production of a charged Higgs boson through $gg \rightarrow tbH^+$, and its subsequent decay $H^+ \rightarrow tb$ for H^+ masses in the range 200 to 2000 GeV. The analysis considers two final states: when both top quarks decay semi-leptonically (the dilepton final state) and when one decays semi-leptonically and one hadronically (lepton-plus-jets). The final states contain four b -quarks and either two charged leptons and E_T^{miss} , or one charged lepton, E_T^{miss} and two light jets. Feynman diagrams are given in Figure 5.1. The b -jet from the gluon splitting is generally very soft and seldom reconstructed.

*The decay modes
 $t \rightarrow ev_e b$ and
 $t \rightarrow \mu\nu_\mu b$ are
considered.*

The analysis uses an object definition similar to the one used in the $t\bar{t}H(bb)$ analysis [66, 67], which is based on recommendations from the ATLAS Top Reconstruction Group [68]. The most challenging background is the SM $t\bar{t}$ production in association with $b\bar{b}$, since this

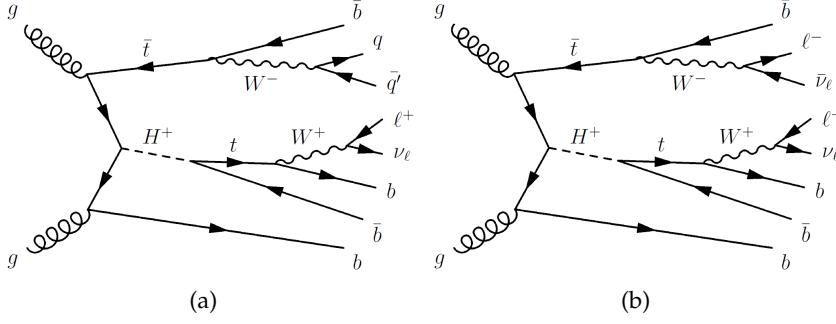


Figure 5.1: Representative Feynman diagrams of the production and decay of a heavy charged Higgs boson $gg \rightarrow H^+ \bar{t}b$, showing the (a) lepton-plus-jets and (b) dilepton final states.

process contains exactly the same final state particles as $t\bar{t}H^+ \rightarrow t\bar{t}b\bar{b}$. In order to gain maximal control over the backgrounds, the signal region is split into sub-categories based on the number of reconstructed b -tags. The baseline approach in the lepton-plus-jets final state follows the strategy from ICHEP, using the signal regions (SR) $5j3b$, $5j\geq 4b$, $\geq 6j3b$, $\geq 6j\geq 4b$ and the control regions (CR) are $5j2b$ and $\geq 6j2b$. The CRs with exactly four jets which were used in the ICHEP analysis, $4j2b$ and $4j\geq 3b$, are omitted because of region overlap with the $t\bar{t}$ -based c -tagging calibration scheme. For the dilepton final state, the baseline signal regions are $3j3b$, $\geq 4j3b$, $\geq 4j\geq 4b$, with control regions $3j2b$ and $\geq 4j2b$. Alternative SR divisions using pseudo-continuous b -tagging (PCBT) have been discussed but have been discarded since the gain in terms of signal strength is very marginal. This is true for both the lepton-plus-jets and dilepton final states. Additionally, calibration of the PCBT variables only exists for low- p_T (< 300 GeV) jets, which excludes the high H^+ mass points.

To increase the signal sensitivity, BDTs are trained for each and every mass point and final state (SR), as discussed in Section 5.7 (lepton-plus-jets final state). Limits to the $gg \rightarrow t\bar{t}H^+$ production cross-section are set using a simultaneous fit to signal and control regions in all final states. In the SR, the BDT output is fitted. In the CR, the fitted distribution is a good sensitive kinematic variable: H_T^{jets} , defined as the scalar sum of the p_T of all selected jets, in the lepton-plus-jets final

$\geq Xj \geq Yb$ means that at least X jets are found in the event, and at least Y of them are b -tagged.

PCBT relies on the information contained in the m_{V2C10} b -tagging score, instead of just selecting “ b -tagged” or not.

BDTs are among the most common used machine-learning algorithms given their robustness, simplicity and explainability.

state, and H_T^{all} , the scalar sum of all p_T selected jets, both charged leptons and the E_T^{miss} , in the dilepton final state.

5.4 SIGNAL AND BACKGROUND PRODUCTION

ATLAS pp collision data and simulations of the H^+ signal at various masses, as well as SM background simulations, are used in this analysis, as described below. Normalisation calculations are denoted as next-to-leading-order (NLO), next-to-next-to-leading-order (NNLO) and next-to-next-to-leading logarithmic order (NNLL). Table 5.1 provides an overview of the simulated samples used, with more details given in Section 5.4.2 and Section 5.4.3.

Sample	Generator	Showering	Normalisation
H^+	MG5_AMC [69]	PYTHIA8.2 [70]	NLO
$t\bar{t}$ + jets	POWHEG-Box v2 [71–73]	PYTHIA8.2	NNLO+NNLL [74–79]
$t\bar{t}$ +HF RW*	SHERPA+OPENLOOPS [80, 81]	SHERPA+OPENLOOPS [82]	NNLO+NNLL
W + jets	SHERPA [80, 83, 84]	SHERPA2.2.1	NNLO
Z + jets	SHERPA	SHERPA2.2.1	NNLO
Single top (s-ch., Wt)	POWHEG-Box v2	PYTHIA6.428 [85]	NNLO [86–88]
Single top (t -ch.)	POWHEG-Box v1	PYTHIA6.428	NNLO
$t\bar{t} + V$	MG5_AMC	PYTHIA8.2	NLO
$t\bar{t}H$	MG5_AMC	PYTHIA8.2	(N)NLO
Diboson	SHERPA	SHERPA2.1.1	NLO

Table 5.1: Overview of the basic generator parameters used to simulate the nominal signal and main background processes. The $t\bar{t}$ +HF component of the inclusive $t\bar{t}$ sample is reweighted to SHERPA+OPENLOOPS, see Section 5.4.3.1 for details. Detailed descriptions of all samples are given in Section 5.4.2 and Section 5.4.3.

Normalisation types

are defined as

follows:

NLO - next-to-

leading-order

NNLO next-to-next-

to-leading-order

All MC samples are generated with the assumption that the top quark mass is $m_{\text{top}} = 172.5$ GeV and the Higgs boson mass is $m_H = 125$ GeV. The detector simulation [89] is done with GEANT4 [90] (full simulation, FULLSIM) or through a fast simulation of the calorimeter response, ATLFastII [91]. Data and simulation are treated equally in terms of reconstruction software, and the simulation is corrected such that the object identification efficiencies, energy scales and energy resolutions match data.

Angular distance in the detector is measured in units of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ (where ϕ is the azimuthal angle and η the pseudo-rapidity). Alternatively, the distance $\Delta R_y \equiv \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$ is used, where $y = \frac{1}{2} \ln \left(\frac{E+p_z}{E-p_z} \right)$ is the rapidity of a particle of energy E and momentum component along the beam axis p_z . The variable ΔR_y is used in the object overlap removal only (Section 6.3).

5.4.1 Data Sample

For this analysis, pp collision data at $\sqrt{s} = 13$ TeV from the LHC and collected with the ATLAS detector have been used. The data were collected in 2015 and 2016 and correspond to an integrated luminosity of 36.1 fb^{-1} .

5.4.2 Signal Samples

The H^+ signal samples are generated with MADGRAPH5_AMCATNLO (MG5_AMC) + PYTHIA8.2 [69, 70], which is a 4FS NLO generator. The NNPDF2.3 parton distribution function (PDF) set [92] is used. The width of the H^+ has been set to zero. Static QCD scales $\mu_R, \mu_F = m_{H^+}/3$ are used. Differences with respect to the latest LHC Higgs Cross-Section Working Group [93, 94] were found to have a negligible impact. 18 different H^+ mass points between 200 and 2000 GeV are generated, with the ATLFastII (AFII) simulation. Additionally, the mass points 300 GeV and 1400 GeV are also available as FULLSIM (FS) for comparison.

Figure 5.2 shows the H^+ production cross-section as a function of the H^+ mass for three different values of $\tan \beta$ [20, 93–99]. Naturally, due to phase-space, the higher the mass of the charged boson, the harder it is to produce. This affects the analysis sensitivity, which will decrease at higher masses due to expected low statistics. Thus, in the case of limit-setting, it will result in weaker limits.

This simulated data assists in analysing the actual data produced by particle accelerators, colliders, and detectors. The abundance of data from these sources requires prior knowledge to help guide analysis, knowledge that comes from theoretical simulations. These simulations are produced using software such as MADGRAPH and SHERPA.

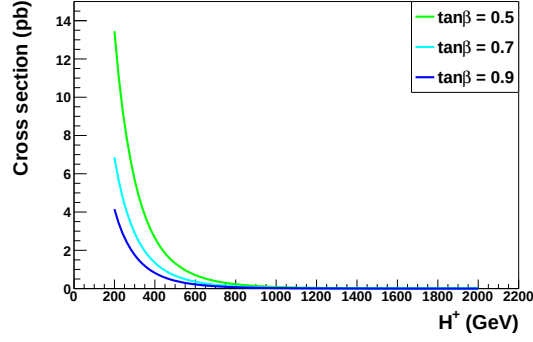


Figure 5.2: H^+ production cross-section as a function of the H^+ mass for three different values of $\tan\beta$. (Figure from [100], numbers from [95], based on [20, 93, 94, 96–99].)

The analysis relies on an accurate description of the signal by the AFII parametrisation.

Since MC are such a narrow resource in HEP analysis, the analysis teams wish to use a “lighter” version (AFII) of fast simulation and verify it holds to the same standards as full simulation.

In Figure 5.3 the number of jets after a lepton-plus-jets pre-selection (at least four jets, at least two b -jets tagged) is compared for the AFII and FS treatments of the two mass points. As can be seen in the plots below, an application of *smoothing* is applied. Smoothing is a technique used to smooth out the distribution in question by re-binning it, thus enabling better treatment of systematic uncertainties and observable collisions.

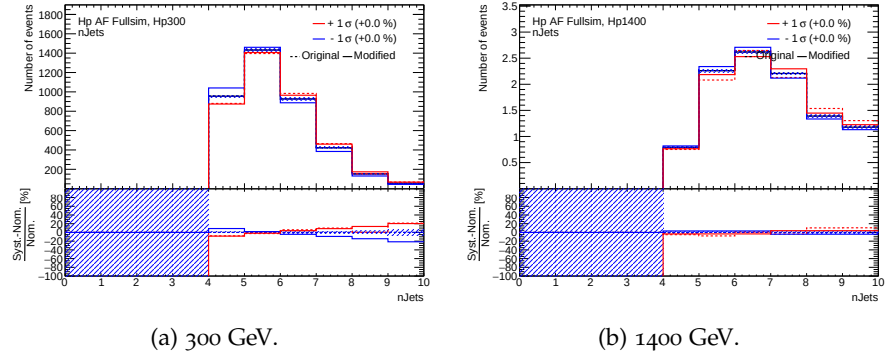


Figure 5.3: Comparison of the number of jets after a lepton-plus-jets preselection, comparing ATLFastII (black line) and FULLSIM (“systematic” symmetrised up and down in blue and red) for H^+ mass (a) 300 GeV and (b) 1400 GeV. The “original” is before smoothing, “modified” is after smoothing.

Similar comparisons for the BDT outputs, can be found in Figure 5.4.

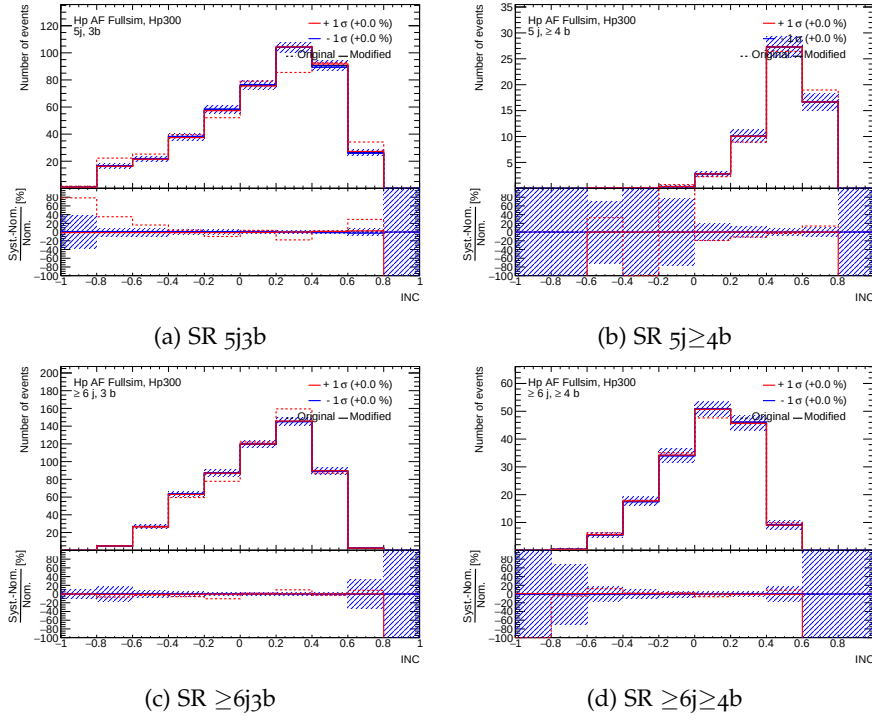


Figure 5.4: Comparison of the BDT output after a lepton-plus-jets preselection, comparing `AtlFASTII` (black line) and `FULLSIM` (“systematic” symmetrised up and down in blue and red) for H^+ mass 300 GeV. Here the SR are shown. The “original” is before smoothing, “modified” is after smoothing.

As can be seen, the agreement between the two is very good, and no overall trends can be seen.

5.4.3 Background Samples

The simulated background samples largely follow the treatment in the $t\bar{t}H(bb)$ analysis [66, 67], with some exceptions in the treatment of reweighting and systematic uncertainties.

The main SM background to the $gg \rightarrow H^+tb \rightarrow tbtb$ process is $t\bar{t}$ +jets production, where the $t\bar{t}$ +HF background (where HF stands for heavy flavour jets, i.e. jets from mainly b quarks) is irreducible. Additional background contributions come from single top quark production, the production of a W or Z boson in association with

jets (W/Z +jets), diboson production (WW , WZ , ZZ), as well as the associated production of a vector boson and a $t\bar{t}$ pair, $t\bar{t}V$ (where $V = W, Z$) and the associated production of a Higgs boson and top quarks, $t\bar{t}H$. Multijet events also contribute to the selected sample via the misidentification of a jet or a photon as an electron or with the presence of a non-prompt lepton.

5.4.3.1 $t\bar{t}$ +jets Background

hdamp controls the p_T of the first additional emission beyond the Born configuration. The main effect of the hdamp setting is to regulate the high- p_T emission against which the $t\bar{t}$ system recoils.

This is the dominant SM background, and thus the most important one. The nominal $t\bar{t}$ +jets sample is generated with the POWHEG-Box v2 NLO generator [71–73, 101] with the NNPDF3.0 parton distribution function (PDF) set [102]. The hdamp parameter is set to $\text{hdamp} = 1.5 \times m_t$. Parton shower and hadronisation are modelled by PYTHIA8.2 with the A14 [103] tune. The sample contains only “non-all hadronic” W boson decays. Dilepton decays can be found in a special sample, which is a subset of the “non-all hadronic” sample. The sample is normalised to the TOP++2.0 [74] theoretical cross-section of 832^{+46}_{-51} pb, calculated at next-to-next-to-leading order (NNLO) in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [75–79].

The generation of the $t\bar{t}$ +jets sample is done inclusively, with all possible flavours in the extra jets present. The decay of bottom and charm hadrons is simulated with the EVTGEN v1.2.0 [104] program. Additional samples with b -jet filters are produced, assuring reasonable statistics after multiple cuts on b -jets. The $t\bar{t}$ +HF category is subdivided into $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$, depending on whether the additional jets originate from hadrons containing b or c quarks, in the following way: if at least one jet ($p_T > 15$ GeV, $|\eta| < 2.5$) in the event is matched ($\Delta R < 0.3$) to a b -hadron with $p_T > 5$ GeV (not originating from a t decay), the event is categorised as $t\bar{t}+ \geq 1b$. If, instead, at least one jet is matched to a charm quark (not originating from a W decay) but no b quark, the event is categorised as $t\bar{t}+ \geq 1c$.

$t\bar{t}$ + jets events which belong to neither the $t\bar{t} + \geq 1b$ nor the $t\bar{t} + \geq 1c$ categories are called $t\bar{t}$ + light events. The $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$ categories are further subdivided into $t\bar{t} + b$ or $t\bar{t} + c$ (when exactly one jet is matched to exactly one b - or c - quark), or $t\bar{t} + b\bar{b}$ or $t\bar{t} + c\bar{c}$ (when two jets are each matched to two b - or two c - quarks), or $t\bar{t} + B$ or $t\bar{t} + C$ (when one single jet is matched to a $b\bar{b}$ or $c\bar{c}$ pair). This categorisation nomenclature has been used in the previous iterations of the analysis [64, 100] and facilitates the comparison with other generators and the treatment of systematic uncertainties.

Given the importance of the $t\bar{t} + \geq 1b$ background, this subset is further subjected to reweighting. The $t\bar{t} + \geq 1b$ events simulated with the nominal POWHEG+PYTHIA8 are reweighted in two ways, to reproduce the relative normalisation from SHERPA+OPENLOOPS [80, 81], and to reproduce its shape prediction. The reweighting is done for the various $t\bar{t} + \geq 1b$ categories in such a way that the relative normalisation changes, but the overall $t\bar{t} + \geq 1b$ normalisation is unchanged. The normalisation reweighting is done by changing the nominal prediction by POWHEG+PYTHIA8 for each subcategory to an NLO prediction including parton showering [82], based on SHERPA+OPENLOOPS using the CT10 [105] PDF set. The shape reweighting is done in several steps: first, the p_T of the $t\bar{t}$ system is reweighted, and secondly, the p_T of the top quarks. A third reweighting is done based on the type of $t\bar{t} + \geq 1b$ event. If there is only one additional heavy flavour jet, the p_T of that jet is used in the final reweighting. If there is more than one additional heavy flavour jet, first, the ΔR between the heavy flavour jets is reweighted, and then the p_T of the HF dijet system. In this analysis, both the shape and the normalisation reweighting is used.

Alternative $t\bar{t}$ samples are generated for the evaluation of systematic uncertainties:

- a SHERPA sample, generated with 5FS PDFs and a b -filter;
- two different POWHEG+PYTHIA8 samples, halving (doubling) the renormalisation and factorisation scales, doubling (halving) the h_{damp} parameter and increasing (decreasing) the A_{14} eigentune parameter

Var3c [103, 106], respectively a POWHEG sample which has been showered either by PYTHIA8 (nominal) or HERWIG7 [107, 108];

- a AMCATNLO sample interfaced to HERWIG++ [107], using either an inclusive $t\bar{t}$ production or having $t\bar{t} + c\bar{c}$ in the matrix element [109].

These samples are used for the evaluation of the systematic uncertainties related to the generator choice, the initial and final state radiation, the parton shower and hadronisation, and the $t\bar{t} + \geq 1c$ prediction, respectively.

5.4.3.2 W/Z +jets and Diboson+Jets

Samples of W/Z +jets events, and diboson production in association with jets, are generated using SHERPA. In the W/Z +jets samples, matrix elements are calculated using the COMIX [110] and OPENLOOPS matrix element generators and merged with the SHERPA parton shower [83] using the ME+PS@NLO prescription [84]. The CT10 PDF set is used together with dedicated parton shower tuning developed by the SHERPA authors. The W/Z +jets events are normalised to the NNLO cross-sections [111].

The diboson+jets samples are generated following the same approach but with up to one (ZZ) or zero (WW, WZ) additional partons at NLO and up to three additional partons at LO. They are normalised to their respective NLO cross-sections calculated by the generator.

5.4.3.3 Single Top Production

Overlaps between the $t\bar{t}$ and Wt final states are removed [112]. The t -channel single top-quark events are generated using the POWHEG-Box v1 generator which uses the four-flavour scheme for the NLO matrix elements calculations together with the fixed four-flavour PDF set CT10f4. The top quarks are decayed with MADSPIN [113], which preserves the spin correlations. The samples are interfaced to PYTHIA6.428 [85] with the PERUGIA 2012 underlying-event tune [114].

The single top quark t - and s -channel samples are normalised to the approximate NNLO theoretical cross-sections [86–88].

5.5 OBJECT RECONSTRUCTION

The physics objects used in the analysis are electrons, muons, jets and b -jets, and missing transverse momentum (E_T^{miss}). The basic object reconstruction follows the recommendations of the ATLAS Top Reconstruction Group [68], with some exceptions since there are four b -jets in the final state. Many of the object selection criteria have been developed in close collaboration with the $t\bar{t}H(bb)$ group [66, 67, 115, 116]. The deviation from the central recommendations are the following: in this analysis, there is a pre-selection requirement that each event in all final states contains at least two b -tagged jets; the p_T cut for the subleading charged leptons in the dilepton channel is lower than the central recommendations, and there is no E_T^{miss} cut in the lepton-plus-jets selection (although E_T^{miss} is used in the event reconstruction in certain cases). The deviations from the Top Reconstruction Group recommendations are driven by the fact that our final state is $tb\bar{t}b$, not $t\bar{t}$. The pre-selection of two b -jets is motivated by the existence of two extra b -jets in the final state. The lowered p_T cut of the second lepton in the dilepton final state increases both the H^+ signal (and $t\bar{t}H$) and the $t\bar{t}$ background, so signal (S) to background (B) ratio, S/B , remains constant. However, $S/\sqrt{B}$ increases. The E_T^{miss} cut is dropped since this mostly reduces the multi-jet background in the $t\bar{t}$ final state. For high- b environments, the multi-jet contribution is low anyway (see the 8 TeV $t\bar{t}H(bb)$ internal documentation [116] especially section 4 for more information).

5.5.1 Electrons

Electrons [117–119] are selected from energy deposits (clusters) in the electromagnetic calorimeter, which are associated with reconstructed tracks in the inner detector.

tight, medium and loose are ATLAS definitions for different levels of certainty in object reconstruction, with tight being the stringest.

Electron candidates must pass the “tight” [119] likelihood identification criterion (TightLH) and the following constraints on the transverse and longitudinal impact parameters: $|d_0^{\text{BL}}/\sigma(d_0^{\text{BL}})| < 5$ and $|\Delta z_0^{\text{BL}} \sin \theta| < 0.5$ mm. The transverse and longitudinal impact parameters are defined with respect to the beamline (BL) [120].

Additionally, the electrons must fulfil the “Gradient” isolation criterion [121]. The working points have been derived from $Z \rightarrow \ell\ell$ MC samples and give the efficiency (ϵ) for calorimeter and track isolation of $\epsilon = 0.1143p_{\text{T}}[\text{GeV}] + 92.14\%$. The total efficiency is 90% at $p_{\text{T}} = 25$ GeV, and 99% efficiency at $p_{\text{T}} = 60$ GeV.

5.5.2 Muons

Muons are reconstructed from tracks in the muon spectrometer, which are matched with tracks in the inner detector, and finally refitted using all the track information from both systems. All muons must fulfil $|\eta| < 2.5$. Furthermore, the muons must be of the “medium” quality [68] and fulfil the “Gradient” isolation criterion, with the same definition as in the electron case. Muons closer than $\Delta R_y = 0.4$ to a jet are removed, except if the jet has fewer than three associated tracks, in which case the jet is removed and the muon kept (this is relevant for high-energy muons which might lose significant energy in the calorimeter). The muon transverse and longitudinal impact parameters must fulfil $|d_0^{\text{BL}}/\sigma(d_0^{\text{BL}})| < 3$ and $|\Delta z_0^{\text{BL}} \sin \theta| < 0.5$ mm.

5.5.3 Jets

Jet candidates are formed from energy depositions in the calorimeter system, topo-clusters [122], using the anti- k_t jet algorithm [123] with a distance parameter of 0.4. Before jet formation, the topo-clusters are calibrated to the electromagnetic energy scale. After reconstruction, the jets are calibrated to the particle level using a jet energy scale (JES) which has been derived from simulations and data-based *in situ* corrections. All jets fulfil $|\eta| < 2.5$ and $p_T > 15$ GeV in the pre-selection (used e.g. for heavy flavour categorisation). For the analysis, the jet p_T cut is > 25 GeV.

To assure that the jets come from the primary pp interaction, a requirement on the Jet Vertex Tagger (JVT > 0.59) is imposed on central low- p_T jets ($p_T < 60$ GeV, $|\eta| < 2.4$). The quality criteria, commonly known as jet cleaning, are applied, and any event which contains at least one LooseBad jet is removed [124].

Since electrons and jets are reconstructed from energy deposits in the calorimeters, the same physical object can end up in both the electron and the jet collections. To avoid double-counting of energy, the closest reconstructed jet, which lies within ΔR_y of 0.2 from a selected electron, is removed. Thereafter, electrons that are closer than $\Delta R_y = 0.4$ to a jet are removed.

5.5.3.1 b -jets

A jet that originates from the hadronisation of a b -quark is identified through the process known as b -tagging [125]. It relies on an algorithm using multivariate techniques to combine information from the impact parameters of displaced tracks with respect to the primary interaction point, as well as topological properties of secondary and tertiary decay vertices within the jet. The algorithm used is the mv2c10 at various working points [126]. Working points are defined by a single cut value on the discriminant output distribution and are chosen to provide

a specific b -jet efficiency on an inclusive $t\bar{t}$ sample. The 70% (85%) working point has a rejection factor of 12 (3.1) and of 381 (33) on charm and light jets, respectively, as stated in Ref. [126], Table 2.

For the pre-selection, the 85% working point is used, but this is just a method for reducing the ntuple size, and its only consequence is that this is the loosest b -tagging requirement available when using these ntuples.

When there is a reference to “a b -jet” without further specification, the 70% working point is understood. This WP was also used in the ICHEP analysis. For the charged Higgs searches, it does not make a huge difference if the 70% or 77% WP is used [127]. The $t\bar{t}H$ analysis used 70% since a tighter requirement on the b -jets is better for their analysis, and the analysis chooses the same WP for simplicity.

PCBT variables are indirectly used as input variables to the BDT in the lepton-plus-jets final state via the discriminant for low masses. This means that for leading- p_T jets, the mv2c10 b -tagging score, binned in 60%, 70%, 77%, 85% and 100% efficiency points, is considered. It should be noted that the PCBT calibration only is valid for jets with p_T below 300 GeV. For conservative reasons, the analysis has chosen to use PCBT variables only for H^+ masses up to and including 300 GeV (where 95% of the jets have a p_T below 250 GeV). The conclusion is that the discriminant and its PCBT input variables are only used for H^+ masses at or below 300 GeV, in the lepton-plus-jets final state.

5.5.4 Missing Transverse Momentum, E_T^{miss}

The missing transverse momentum, E_T^{miss} , in the event is computed as the negative magnitude of the vector sum of the p_T of all physics objects in the event, with correction for “soft” energy in the event, which is not associated with any of the hard objects.

The missing transverse momentum is not used for event selection but as input to the event reconstruction.

5.6 SINGLE LEPTON FINAL STATE

The final state of the one-lepton decay of the H^+ with an associated top quark, the lepton-plus-jets final state, is shown in Figure 5.5 and contains one neutrino (E_T^{miss}), one charged lepton and six jets of which four originate from b -quarks. The b -quark from the initial gluon splitting is generally very soft and hard to reconstruct.

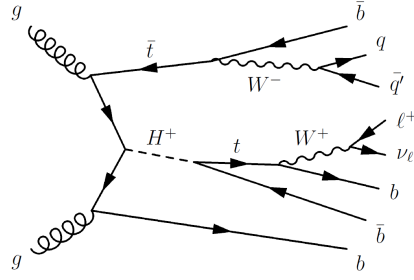


Figure 5.5: Representative Feynman diagram of the production and decay of a heavy charged Higgs boson $gg \rightarrow H^+ t \bar{t}$ in the lepton-plus-jets final states. It should be noted that either of the two top quarks can decay semi-leptonically in the lepton-plus-jets final state.

5.6.1 Selection

In the lepton-plus-jets final state, exactly one charged lepton, electron or muon, is selected.

The simulated events are randomly assigned a run number corresponding to the 2015 or 2016 data set, weighted such that the relative amount of 2015 and 2016 data is preserved in the simulation. The simulated event is then required to pass the trigger criteria of its assigned year.

The selected trigger-matched charged lepton must have a p_T above 27 GeV. Additionally, each event must contain at least four jets with p_T above 25 GeV, of which at least two are b -tagged. Events passing the dilepton selection are vetoed.

In order to gain a handle on the irreducible $t\bar{t} + b\bar{b}$ background, in the baseline analysis, the selected sample is divided into several SR

Meaning that the selected lepton has a distance $\Delta R < 0.15$ to the lepton reconstructed by the trigger.

and control regions (CR), based on the number of jets and b -jets tagged in the event. The SR for the lepton-plus-jets final state are $5j3b$, $5j \geq 4b$, $\geq 6j3b$ and $\geq 6j \geq 4b$, and the CR are $4j2b$, $4j \geq 3b$, $5j2b$ and $\geq 6j2b$. The CRs with exactly four jets which were used in the ICHEP analysis, $4j2b$ and $4j \geq 3b$, are omitted because of region overlap with the $t\bar{t}$ -based c -tagging calibration scheme. As an alternative, signal region division based on pseudo-continuous b -tagging (PCBT) has been explored but discarded.

For $t\bar{t}$ production, the number of events with high leading jet p_T is overestimated in the simulation, and reweighting factors for the leading jet p_T distribution are determined by comparing simulation and data in a lepton-plus-jets CR $\geq 4j2b$, and applied in all lepton-plus-jets and dilepton SRs and CRs. The reweighting factors are, for example, 0.99 for jet p_T up to 200 GeV and 0.89 for jet p_T of 500 GeV.

The most challenging background is $t\bar{t} + b\bar{b}$ since it contains exactly the same final state as the signal. In order to enhance the signal over background, multivariate techniques and discriminators are utilised, in the form of Boosted Decision Trees (BDT), which have been trained and used in the various SR.

5.7 BACKGROUND SUPPRESSION USING BDT

Upon performing hypothesis testing (see Chapter 4), one must choose the variable that shows the best separation between signal and background. However, a single kinematic variable is often not suited for this purpose. In fact, the features of the signal can be embedded in correlations between multiple variables. It is partly possible to overcome this issue by combining multiple features into engineered variables. For example, one could sum the transverse momenta of multiple particles to include kinematic information for all the objects in a given event. Unfortunately, it is difficult to create the optimal variable. Therefore this analysis makes use of a more refined technique. A set of variables, which encode some of the most important kinematic features of the H^+ decay, is chosen, and then a BDT is trained on the entire set, using MC events.

BDT is one of the most popular methods of combining multiple variables into one single optimised discriminator. In this analysis, the TMVA toolkit [128], ver 4.2.1, is used for this purpose. A decision tree is a binary tree-structured classifier trained on a Monte Carlo simulation sample of known composition and tested on another independent known sample. All provided variables are scanned at each step to find the single cut that offers the best signal-background separation, according to a chosen criterion. The sample is divided by the selected cut, and the algorithm is re-applied on each resulting sub-sample. This sample splitting continues until a stopping condition is reached, for example, the minimum number of events allowed in a sub-sample (also known as minimum leaf size). This means that a single decision tree makes a series of cuts, although no events are discarded because even an event that fails a particular cut is considered for other cuts by the algorithm. The result is that the phase space is divided into signal- and background-like regions and each event receives a score describing how signal- or background-like it is.

The “boosting” of the decision tree is done to improve the performance and the stability of their behaviour. After the training of the first tree, those events misclassified by the first tree have their weights increased. A second tree, which is trained on this re-weighted sample, then naturally focuses on the events which were problematic in the first tree. This procedure constitutes a single “boost”, and the number of such boosts is a parameter to be optimised in the BDT training. At the end of the boosting cycle, a score is assigned to each event using a weighted average of the score calculated in each tree. The weight used is derived from the misclassification rate of each tree.

Two different approaches were tried, either where the background is the $tt + bb$ background (heavy-flavour background, “HF”) or all SM backgrounds are used, i.e. tt , tth , ttW , ttZ , di-boson, single top, Z +jets and W +jets (inclusive “INC” BDT). This analysis uses the INC BDT for all mass points, and for the masses at and below 300 GeV, additionally also the discriminant is included as an extra variable.

All backgrounds used in the BDT training are appropriately weighted according to their best known cross-sections, with the normalisations.

- `weight_normalise` - relative portion of the sample w.r.t. the rest of the backgrounds.
- `weight_pileup` - pileup reweighing.
- `weight_mc` - generator weight.
- `weight_leptonSF` - lepton scale factors.
- `weight_bTagSF_70` - b-tagging scale factor for 70% WP.

The resulting BDTs have been trained on one sample and then tested on another, statistically independent sample. The analyses apply cross-training, i.e. first the BDT is trained on the even events and tested on the odd ones, and then the training is done on the odd events and the testing on the even ones. The basic settings of the BDT are listed in Table 5.2.

BDT variable	Value
BoostType	Grad
UseBaggedBoost	False
nTrees	165 (SR $5j \geq 3b$, $\geq 6j3b$), 120 (SR $\geq 6j \geq 4b$).
Shrinkage	0.13
MaxDepth	5
MinNodeSize	1%

Table 5.2: Settings for the BDT trained in the dilepton final state. TMVA 4.2.1 and ROOT version 6.04 were used. The nTree setting for $\geq 4j \geq 4b$ was lowered to 200 because of a lack of statistics.

Table 5.3 lists the variables used as input to the BDT.

Symbol	meaning
jet_pt[0]	Leading jet p_T , ordered by b -tagged jets and then non- b -tagged jets
Mbb_MindR	Invariant mass of closest b -jet pairs
pT_jet5	p_T of fifth leading jet, ordered by b -tagged jets and non- b -tagged jets
H1_all	Second Fox-Wolfram moment calculated using all jets and leptons
dRbb_avg	Average ΔR between all b -jet pairs in event
dRlepbb_MindR	ΔR between the lepton and the pair of b -jets with smallest ΔR
Muu_MindR	Invariant mass of the untagged jet-pair with minimum ΔR
HT_jets (H_T^{jets})	Scalar sum of E_T of all jets
Mbb_MaxPt	Invariant mass of the b -jet pair with maximum p_T
Mbb_MaxM	Maximal invariant mass of b -jets
Mjjj_MaxPt	Invariant mass of the jet triplet with maximum p_T
Centrality_all	Centrality calculated using all jets and leptons
HpDiscriminant	Kinematic discriminant ($m_{H^+} \leq 300$ GeV)

Table 5.3: List of variables for the lepton-plus-jets BDT. The discriminant, which includes PCBT variables, is used for masses up to and including 300 GeV only due to PCBT calibration is only valid for jet p_T up to 300 GeV.

5.7.1 INC vs HF BDT

Comparing the lepton-plus-jets BDT choices: INC (trained on all background samples) vs HF (trained on $t\bar{t}$ +HF only). Here the comparison is made only for mass points between 200 and 275 GeV. In the ICHEP analysis, there was a potential gain for the HF BDTs at low masses only, and preliminary studies in this incarnation of the analysis (not explicitly shown here) confirmed that there definitely was no gain in using the HF BDTs at masses above 275 GeV.

Table 5.4 lists the uncertainty on the signal strength μ_{H^+} for various settings: the BDTs have been trained either against the full set of backgrounds (“INC”), or against $t\bar{t}$ +HF only (“HF”). The prescaled triggers are *HLT_e24_lhmeduim_nod0_L1EM18VH* and *HLT_mu24*, for the electron and muon channels, respectively.

Mass	INC PS	INC UPS	HF PS	HF UPS
200 GeV	1.92	1.82	1.64	1.74
225 GeV	1.10	1.12	1.66	1.32
250 GeV	1.00	0.89	1.30	1.14
275 GeV	0.89	0.85	1.08	1.06

Table 5.4: Uncertainty on the signal strength μ_{H^+} , after fits of BDTs trained at masses 200 – 275 GeV. The table compares BDTs trained against all backgrounds (“INC”) with $t\bar{t}$ +HF as background (“HF”). The fake estimation is made with prescaled triggers (“PS”) or un-prescaled triggers (“UPS”).

The nominal value of
 μ_{H^+} is 0.

A fit is considered “better” if it gives a stricter constraint (lower uncertainty) on the signal strenght μ_{H^+} . Comparing “INC UPS” with “HF UPS” in Table 5.4, one can observe that the constraint on μ_{H^+} only is stricter for the mass point 200 GeV, and the same is true when comparing “INC PS” with “HF PS”.

When comparing prescaled triggers for the fake determination (PS) with un-prescaled triggers, most of the time, there is little difference. However, especially for the low masses and the HF BDT, sometimes the μ_{H^+} constraint changes a lot when moving to the UPS, i.e. at HF

225 GeV (an improvement from 1.66 to 1.32 when moving to UPS) or at HF 250 GeV (an improvement from 1.30 to 1.14). This kind of instability when changing the fakes estimate is a cause for concern.

It is desired to keep the correlations in the fits low, especially between the $t\bar{t}$ +HF background k -factor $k_{t\bar{t}+\geq 1b}$ and the signal strength μ_{H^+} , since the irreducible $t\bar{t}+\geq 1b$ background should not be correlated with the kinematically similar signal. Table 5.5 shows the anti-correlation between μ_{H^+} and $k_{t\bar{t}+\geq 1b}$, for masses 200 – 275 GeV.

Mass	INC PS	INC UPS	HF PS	HF UPS
200 GeV	–59%	–54%	–56%	–59%
225 GeV	–53%	–66%	–69%	–68%
250 GeV	–54%	–61%	–60%	–60%
275 GeV	–60%	–63%	–65%	–65%

Table 5.5: Values of the (anti-)correlation between μ_{H^+} and $k_{t\bar{t}+\geq 1b}$, for masses 200 – 275 GeV. The table compares BDTs trained against all backgrounds (“INC”) with $t\bar{t}$ +HF as background (“HF”). The fake estimation is made with prescaled triggers (“PS”) or un-prescaled triggers (“UPS”).

Here, the option INC with prescaled triggers (“INC PS”) is the only option that keeps the absolute values of the anti-correlations at or below 60%. The use of un-prescaled triggers generally increases the anti-correlation.

Conclusions: The INC BDT will be used for all mass points. The HF BDT, in fact, shows much higher instabilities when changing the fake modelling. This is caused by the fact that the HF BDT is trained only against the HF part of the $t\bar{t}$ background, and it is therefore not able to separate fakes and signal and the inclusive BDT.

The analysis also uses the prescaled triggers for the fake estimation since this option gives the lowest correlations between μ_{H^+} and $k_{t\bar{t}+\geq 1b}$.

5.7.2 BDT Performance in the Signal Regions

Figure 5.6 – Figure 5.8 show the BDT output in the various signal regions. The masses shown are $m_{H^+} = 225, 400, 1200, 2000$ GeV. Two

overtraining plots are shown, “even” which means that the training was done on even events and then tested on odd events, and “odd” in which the training was done on the odd events and tested on the even ones. It should be noted that at high H^+ masses, the separation is extremely good because of the large differences in kinematics between the SM background and high-mass H^+ signal.

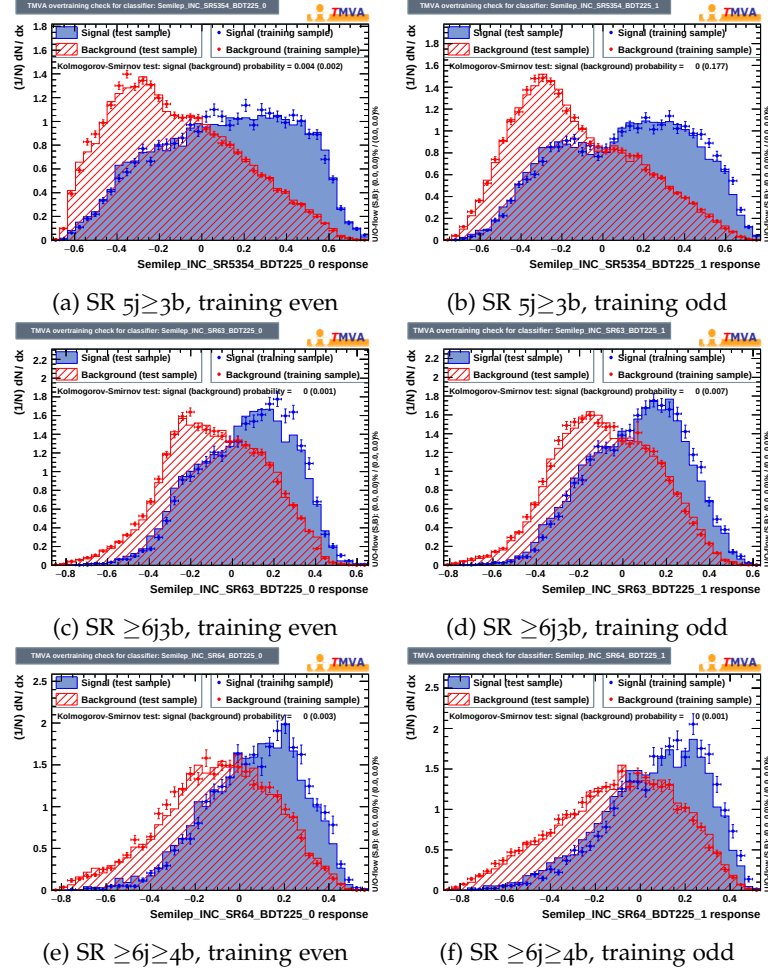


Figure 5.6: BDT output for the $m_{H^+} = 225$ GeV mass point. The signal regions considered are $5j3b$ and $5j \geq 4b$ combined ($5j \geq 3b$), $\geq 6j3b$ and $\geq 6j \geq 4b$.

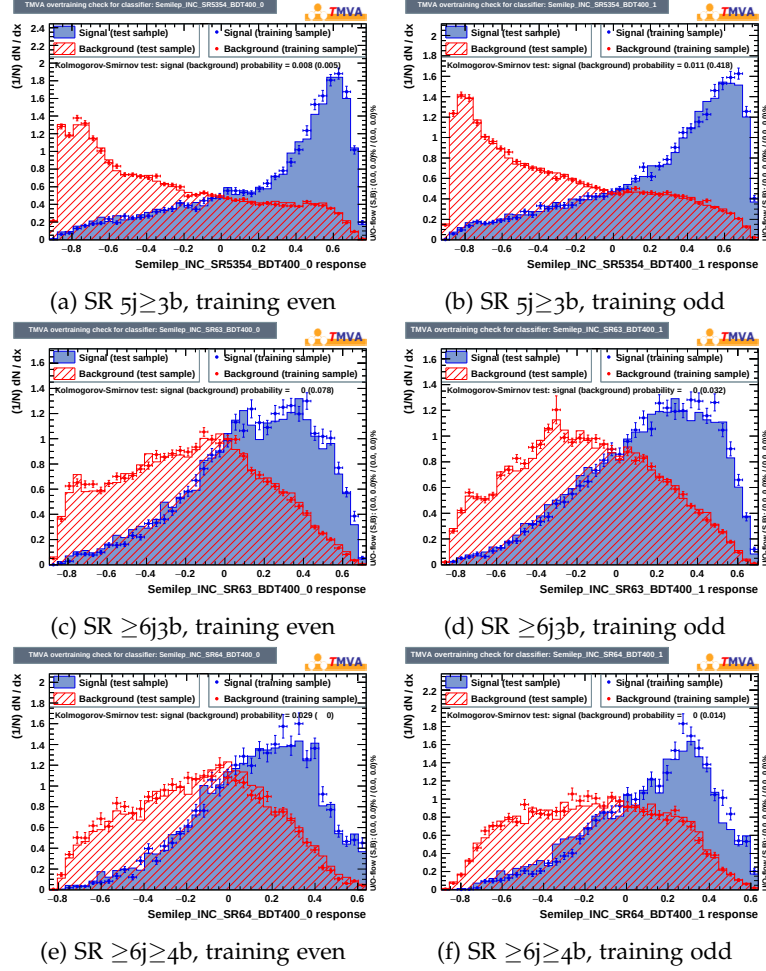


Figure 5.7: BDT output for the $m_{H^+} = 400$ GeV mass point. The signal regions considered are $5j3b$ and $5j \geq 4b$ combined ($5j \geq 3b$), $\geq 6j3b$ and $\geq 6j \geq 4b$. The training has been done on even or odd events, as indicated in the plots, and then tested on the remaining sample.

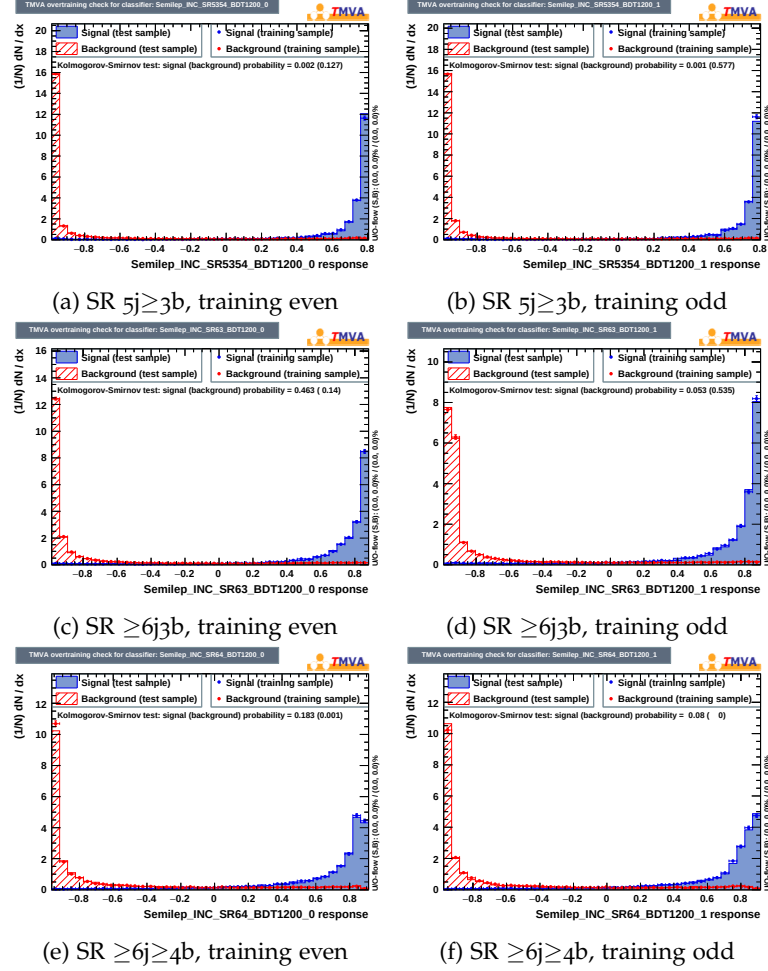


Figure 5.8: BDT output for the $m_{H^+} = 1200$ GeV mass point. The signal regions considered are $5j3b$ and $5j \geq 4b$ combined ($5j \geq 3b$), $\geq 6j3b$ and $\geq 6j \geq 4b$. The training has been done on even or odd events, as indicated in the plots, and then tested on the remaining sample.

5.7.3 BDT Variable Ranking and Importance

The importance of each BDT variable varies among the variables and as a function of H^+ mass. Figure 5.9 shows the relative ranking and the importance of the BDT input variables.

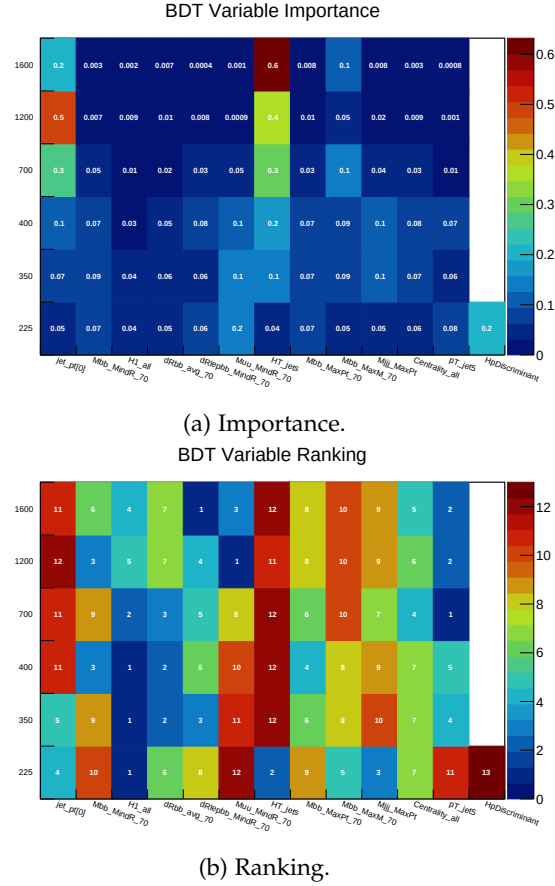


Figure 5.9: BDT input variable (a) importance and (b) ranking, for a selection of masses (y-axis) and all variables (x-axis). Ranking 12 or 13 means higher ranked than 1 (i.e. more important). The discriminant is only used for masses below 300 GeV and is thus only shown for mass point 225 GeV in this selection. The results have been averaged over all SR.

The most important variable at low mass is the discriminant, which is only used for masses up to 300 GeV. Other important variables at low masses are M_{bb_MindR} and M_{bb_MindR} (invariant mass of closest untagged and b -tagged jet pairs, respectively). These variables explore the kinematic differences between the $t\bar{t} + b\bar{b}$ background and the H^+ signal. Additionally, p_{T_jet5} (p_T of fifth leading jet) is important at

the lowest masses, as well as M_{jjj_MaxPt} (invariant mass of the jet triplet with maximum p_T) for intermediate masses.

In contrast, the important variables at high masses are the ones exploring energy: the E_T sum of all jets (H_T_jets) and the leading-jet p_T ($jet_pt[0]$), which are not very important at the very lowest masses. This reflects that the total energy of the decay products of a high-mass H^+ event is higher than in a typical $t\bar{t} + b\bar{b}$ background event.

Few variables have absolute importance higher than 15% for any mass. For masses above 400 GeV, only H_T_jets , $jet_pt[0]$ and M_{bb_maxM} (the maximal invariant mass of any b -jet pair) have importance greater than about 10%. The variable pT_jet5 looks like it is imperative at low masses but lose importance at high masses. While this is true for the relative ranking (Figure 5.9b), looking at the importance (Figure 5.9a), one can see that the absolute importance is 6–8% for all low mass points. Still, the ranking varies dramatically due to small changes in the importance of other variables.

5.7.4 BDT Background and Data Correlation

It is important to know that the BDTs behave reasonably in data. One way of checking this before unblinding is to look at the variable correlations in the control regions. In Figure 5.10 and Figure 5.11, the variable correlations when using either data or background simulation as background can be found. The figures show the variable correlations in the CRs $5j2b$ and $\geq 6j2b$, respectively.

The variable correlations are very similar when comparing background simulation and data in all control regions.

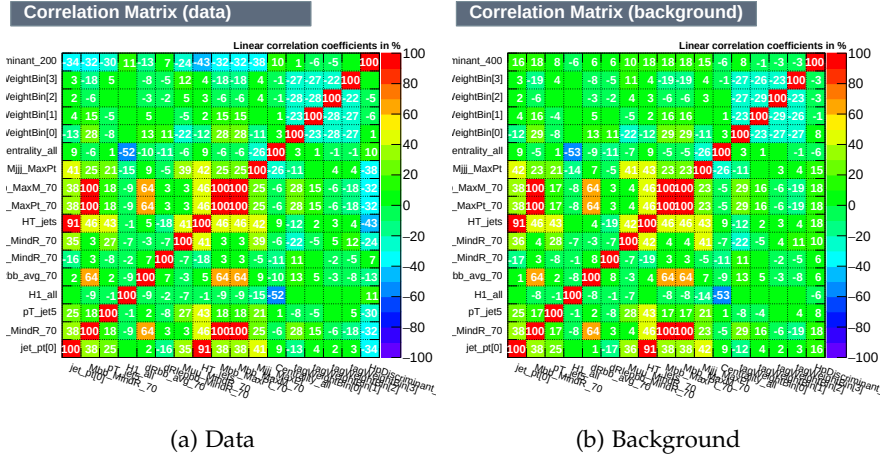


Figure 5.10: Comparison of the variable correlations in CR 5j2b using (a) data or (b) background simulation as background.

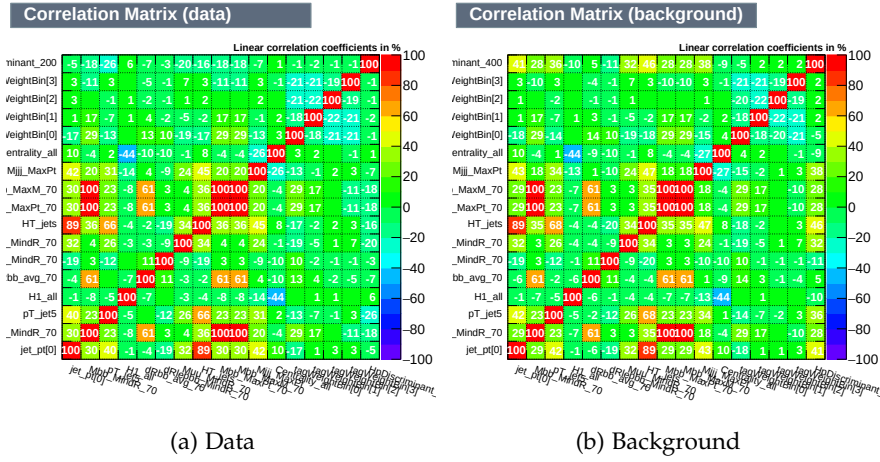


Figure 5.11: Comparison of the variable correlations in CR $\geq 6j2b$ using (a) data or (b) background simulation as background.

5.7.5 *Summary of Choices for the Analysis*

A multitude of different settings, variable lists and discriminating methods have been tried out. The choice for the analysis is the following:

- The BDT variable list contains twelve optimal variables, largely the same as used in the ICHEP analysis (see Table 5.3).
- The analysis added a H^+ mass reconstruction discriminating variable as an input variable to the BDT. However, the discriminant can only be used for low H^+ masses (at or below 300 GeV) since it contains the PCBT variables (the mv2c10 b -tagging score, binned in 60%, 70%, 77%, 85% and 100% efficiency points) with calibrations only valid for low jet p_T .
- The BDT is trained against all backgrounds, not only the $t\bar{t}$ +HF, for reasons explained in Section 5.7.1.
- To gain statistics in the BDT training, the 5j signal regions, 5j3b and 5j $\geq$ 4b are combined into a single SR, 5j $\geq$ 3b. This reduces the overtraining seen in previous iterations.

5.8 SYSTEMATIC UNCERTAINTIES

Systematic uncertainties are uncertainties in the knowledge of background and signal processes that manifest themselves as uncertainties in the yields and probability density (referred to as *shape*) of the observable signal and background. Systematic uncertainties arise from sources related to the reconstruction, identification and calibration of objects, referred to as *experimental systematic uncertainties*. Lastly, uncertainties related to theoretical predictions and modelling of background and signal are referred to as *theoretical systematic uncertainties*.

The evaluation of the systematic uncertainties follows the top working group recommendations [129], and are largely the same as in the $t\bar{t}H(bb)$ analysis, except for the H^+ signal systematic uncertainties, the $t\bar{t}H$ modelling (which is the background to us) and a few details in the break-down of the systematic uncertainties. The main deviation from the ICHEP analysis is in the systematic uncertainties associated with the simulated $t\bar{t}$ background, which is now POWHEG+PYTHIA8. The following systematic uncertainties apply to both the lepton-plus-jets and the dilepton final states unless otherwise stated.

Multiple sources of uncertainty are being considered, and they can affect both the normalisation of the simulated samples and the shape of their corresponding final discriminant distributions. Individual sources of systematic uncertainty are considered uncorrelated. Correlations of a given systematic uncertainty are maintained across processes and channels. Table 5.6 presents a list of all systematic uncertainties considered in the analysis and indicates whether they affect only normalisation or both shape and normalisation. The dominant sources of uncertainty for this analysis are related to the imperfect modelling of $t\bar{t}$ +jets events, in particular $t\bar{t}$ produced in association with extra heavy flavour jets, as well as b -tagging scale factor uncertainties and jet energy scale uncertainties. Lepton, E_T^{miss} and trigger

uncertainties are typically subdominant, as well as the uncertainties on smaller backgrounds.

5.8.1 *Luminosity and Pile-up*

The combined uncertainty on the integrated luminosity for the data collected in 2015 and 2016 is 2.1%, and it is applied as a flat scale normalisation on all processes estimated from simulation. It is derived, following a methodology similar to that detailed in [130], from a preliminary calibration of the luminosity scale using $x - y$ beam-separation scans performed in August 2015 and May 2016.

A variation in the pileup reweighting of the MC simulation is included to cover the uncertainty on the ratio between the predicted and measured inelastic cross-section in the fiducial volume defined by $M_X > 13 \text{ GeV}$ where M_X is the mass of the hadronic system [131]. The reweighting is done with the `CP::PileupReweightingTool` tool in order to correct for differences in the pileup distribution between the simulation and collision data and to correct the average number of interactions per bunch crossing ($\langle \mu \rangle$) to account for the latest luminosity calculation recommendations [132, 133].

5.8.2 *Reconstructed Objects*

5.8.2.1 *Charged Leptons (e and μ)*

Uncertainties associated with charged leptons arise from the trigger selection, the object reconstruction, the identification, and the isolation criteria, as well as the lepton momentum scale and resolution. The reconstruction, identification and isolation efficiency of electrons and muons, as well as the efficiency of the trigger used to record the events, differ slightly between data and simulation, which is compensated for by dedicated scale factors (SFs).

<i>Systematic uncertainty</i>	Type
Luminosity	N
Pile-up reweighting	SN
Reconstructed objects	
Electron trigger+reco+ID+isolation	SN
Electron energy scale+resolution	SN
Muon trigger+reco+ID+isolation+TTVA	SN
Muon momentum scale+resolution	SN
Jet vertex tagger (JVT)	SN
Jet energy scale (JES)	SN
Jet energy resolution (JER)	SN
E_T^{miss} resolution and scale	SN
b -tagging efficiency, fixed WP	SN
c mistag rate, fixed WP	SN
Light mistag rate, fixed WP	SN
Mistag extrapolation $c \rightarrow \tau$, fixed WP	SN
High- p_T b -tagging extrapolation, fixed WP	SN
b -tagging efficiency, PCBT	SN
c mistag rate, PCBT	SN
Light mistag rate, PCBT	SN
Mistag extrapolation $c \rightarrow \tau$, PCBT	SN
Signal model	
Renormalisation and factorisation scales	N
Parton distribution function (PDF)	SN
Background models	
$t\bar{t}$ cross-section	N
$t\bar{t} + \geq 3b$: normalisation	N
$t\bar{t} + b\bar{b}$: NLO shape	SN
$t\bar{t}$ modelling: residual NLO generator	SN
$t\bar{t}$ modelling: residual parton shower+hadronisation	SN
$t\bar{t}$ leading jet p_T reweighting	SN
Single top Wt channel model and cross-section	SN
$t\bar{t}V$ QCD and PDF	N
$t\bar{t}V$ generator	SN
$t\bar{t}H$ QCD and PDF	N
$t\bar{t}H$ model	SN
tZ QCD and PDF	N
$t\bar{t}WW$ QCD and PDF	N
tWZ cross-section	N
W +jets normalisation (ℓ +jets final state)	N
Z +jets normalisation (ℓ +jets final state)	N
Z +jets normalisation (dilepton final state)	N
Diboson normalisation	N

Table 5.6: Systematic uncertainties considered. “N” means that the uncertainty is normalisation-only for all processes and channels affected, while “SN” means that the uncertainty affects both shape and normalisation.

The efficiency of the SFs is derived by comparing $Z \rightarrow \ell^+ \ell^-$ ($\ell = e, \mu$) in data and simulation with the tag-and-probe method and used to correct for differences in the simulation with respect to data. The total uncertainty on the efficiency of the SFs is smaller than 0.5% for muons in the p_T range considered in this analysis [134]. For the muons, the reconstruction uncertainty is combined with the ID uncertainty, and track-to-vertex association (TTVA) is considered, as well as special uncertainties to consider the reconstruction of low- p_T muons.

The uncertainties on the combined electron reconstruction and identification SF is smaller than 1% for electrons with $p_T > 30$ GeV, while it reaches 2–3% for the lowest p_T values considered in this analysis [119].

5.8.2.2 *Jets and Missing Transverse Momentum*

Uncertainties associated with jets in principle arise from the efficiency of jet reconstruction and identification based on the JVT variable, as well as the jet energy scale (JES) and resolution (JER). The jet energy scale and its uncertainty were derived by combining information from test-beam data, LHC collision data (in situ techniques) and simulation [135]. The JES uncertainty is smaller than 1% in the central calorimeter region ($|\eta| < 1.2$) for jets with $100 < p_T < 500$ GeV. For low- p_T jets, $p_T = 20$ GeV, in the central region, the uncertainty is dominated by uncertainties in the corrections for multiple proton-proton interactions (pile-up), and the uncertainty is about 4.5% [136].

For the JES uncertainties, the analysis has 23 nuisance parameters. 21 of them are the ones recommended by the ATLAS Jet/ETmiss group [137] and they contain eight nuisance parameters from in situ analyses (Z+jet balance, γ +jet balance, and multi-jet balance), three nuisance parameters from η inter-calibration, one nuisance parameter for high- p_T jets extrapolated from single hadrons, four nuisance parameters which relate to pile-up, one nuisance parameter for the non-closure in the comparison between AtlFastII and FullSim simula-

tion calibration, one nuisance parameter from the flavour composition, one nuisance parameter from the flavour response uncertainty, one nuisance parameter from b -jets, which is applied to b -jets only, and one nuisance parameter for punch-through jets. In addition, there is one more nuisance parameter by de-correlating the flavour composition of $t\bar{t}$ +light and all other backgrounds, and also two different nuisance parameters for the punch-through effects in the FullSim and AFII samples, i.e. one in addition to the standard recommendation.

The E_T^{miss} reconstruction is affected by uncertainties associated with leptons and jet energy scales and resolutions, which are propagated to E_T^{miss} and thus are included under the corresponding per-object uncertainty category. Uncertainties in the resolution of the soft terms, as well as the soft scale, are considered. There are three components: one for each momentum direction x and y and one for the overall energy scale. Since E_T^{miss} is not used in selection but only in event reconstruction, its uncertainties typically have a small effect on the analysis.

5.8.2.3 Heavy- and Light-Flavour Tagging

Correction factors are applied to the simulated event samples to compensate for differences between data and simulation in the b -tagging efficiency for b , c and light-jets. The correction for b -jets is derived from $t\bar{t}$ events with final states containing two leptons. The corrections are consistent with uncertainties at the level of a few percent over most of the jet p_T range [125, 126, 138].

For the fixed (cumulative) working points, a total of 6, 3 and 16 independent eigen-variations (EVs) are considered for b -, c - and light-jets, respectively. For the PCBT variations, these numbers are multiplied by the 5 weight bins used, resulting in 30, 15 and 80 EVs for b -, c - and light-flavour, respectively. These systematic uncertainties are taken as uncorrelated between b -jets, c -jets, and light-jets. The efficiency of tagging the hadronic decays of τ -leptons in simulations is treated in

the same way as the efficiency of tagging c -jets, with an additional uncertainty accounting for the fact that the correction factors derived for c -jets are also used for hadronic decays of τ -leptons.

When using fixed WP tagging, additional uncertainty is included due to the extrapolation of SFs for jets with p_T beyond the kinematic reach of the data calibration samples used. This uncertainty is taken to be correlated among the three jet flavours. This source of uncertainties is, however, not integrated into the PCBT uncertainties, which are valid for jet p_T up to 300 GeV only. The analysis uses PCBT variables and the corresponding uncertainties for H^+ masses up to and including 300 GeV, in the lepton-plus-jets final state. To enable a combination with the dilepton final state, the PCBT calibration and uncertainties are also used in the dilepton final state for these masses, even though no PCBT variables are used in the analysis.

For H^+ masses above 300 GeV, the default fixed-WP calibration is used.

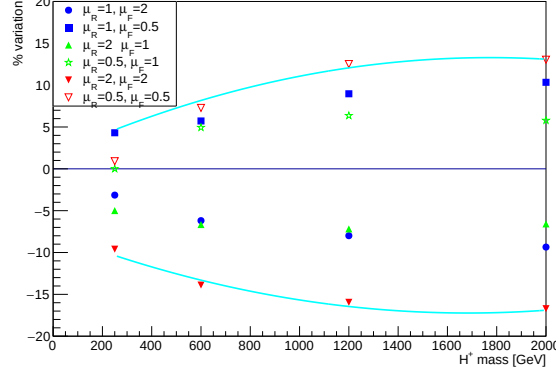
5.8.3 Signal Modelling

The signal uncertainty is modelled in two ways: by using the PDF uncertainties and through the variation of the renormalisation and factorisation scales μ_R and μ_F .

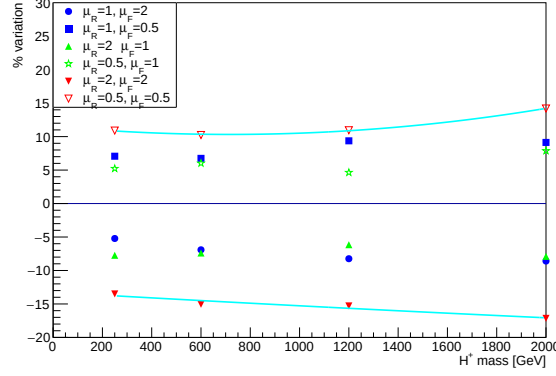
The uncertainties from the modelling of the *parton density function* (PDF) in the signal generation (which is done with the NNPDF2.3 PDF set [92]) is made using a symmetrised Hessian set of 30 components, PDF4LHC15_nlo_30, following the PDF4LHC recommendations for LHC Run II [139].

The signal *scale* uncertainty is modelled by varying the renormalisation and factorisation scales μ_R and μ_F up and down (by a factor of 2 or 0.5, respectively). Specially dedicated samples were created containing this information for four representative mass points, 250, 600, 1200 and 2000 GeV. A representative subset of plots to illustrate

the impact of the scale variations is shown here. The lepton-plus-jets and dilepton final states are treated separately but with the same method. The final result can be found in Figure 5.12.



(a) Lepton-plus-jets final state



(b) Dilepton final state

Figure 5.12: Scale variation, as a function of H^+ mass, for (a) lepton plus jets merged SR $\geq 5j3b$ and $\geq 5j\geq 4b$ and (b) the dilepton final state. The variations have been grouped in “up” and “down” variations, with $\mu_{R(F)}$ fixed at 1 and $\mu_{F(R)}$ varied to 2 or 0.5, or $\mu_{R(F)} = 2$, $\mu_{F(R)} = 0.5$. The envelope of the uncertainty is indicated in cyan. The change in acceptance can be seen as the variation is changing, thus affecting the fraction of events.

All possible combinations of μ_R and μ_F , each taking the values 1 (nominal), 2 (shift up) and 0.5 (shift down) are tried. This gives eight possible combinations in total, and $\mu_R = 1$, $\mu_F = 1$ is the nominal setting. In order to gain statistics, lepton-plus-jets SR are merged into $\geq 5j3b$ ($\geq 6j3b + 5j3b$) and $\geq 5j\geq 4b$ ($\geq 6j\geq 4b + 5j\geq 4b$). The dilepton SR are merged into $\geq 3j3b$ ($3j3b + \geq 4j3b$) to be used along with the

nominal $\text{SR} \geq 4j \geq 4b$ for masses 250 GeV and 600 GeV. For 1200 and 2000 GeV, where $3j3b$ is a CR, the SRs $\geq 4j3b$ and $\geq 4j \geq 4b$ are used. The later nomenclature is used for all plots, but for the low-mass ones, SR $3j3b$ is included in the $4j3b$ region.

In Figure 5.13 the variation of the scales is shown as a function of the BDT output, compared with the nominal scale ($\mu_R = 1$, $\mu_F = 1$) in the lepton-plus-jets final state, for both the lepton-plus-jets and dilepton final states. Figures 5.14 and 5.15 are derived from Figure 5.13, showing the relative variation for each final state. Figures 5.13 to 5.15 indicate that the scale variation for each set of variations is reasonably flat in the BDT distribution.

In Figure 5.12, the acceptance change as a consequence of the scale variations is shown, for the lepton-plus-jets final state (Figure 5.12a) and the dilepton final state (Figure 5.12b). In this plot the variations have been grouped in “up” and “down” variations, with $\mu_{R(F)}$ fixed at 1 and $\mu_{F(R)}$ varied to 2 or 0.5, or $\mu_{R(F)} = 2$, $\mu_{F(R)} = 0.5$. The maximal acceptance change, the “envelope” is also indicated.

The fitted envelopes of Figure 5.12 for each final state are used to estimate the uncertainty. These uncertainties have been symmetrised for each mass point and vary from 7% at low masses to 15% at masses above 1300 GeV for the lepton-plus-jets final state, ranging from 12% to 16.5% in the dilepton final state. This uncertainty is applied as a flat scale to the signal for each mass point. The impact of the signal systematic uncertainties (scale and PDF) on the final expected limit is small.

5.8.4 Background Modelling

$t\bar{t} + \text{Jets}$

The $t\bar{t}$ background, which is the largest SM background, is modelled using POWHEG+PYTHIA8, as described in Section 5.4.3.1. A large number of systematic uncertainties are considered. The cross-section

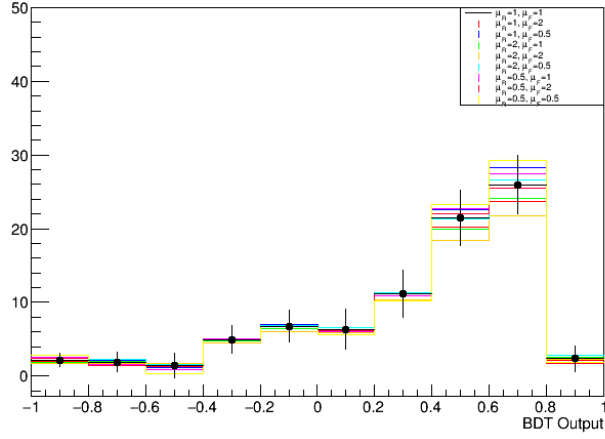
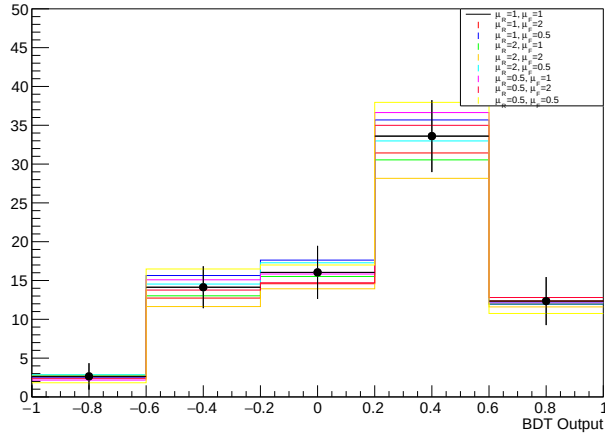
(a) Lepton-plus-jets, $SR \geq 5j \geq 4b$ (b) Dilepton, $SR \geq 4j \geq 4b$

Figure 5.13: Scale variation, BDT output. $m_{H^+} = 600$ GeV, in the (a) lepton-plus-jets merged $SR \geq 5j \geq 4b$ and (b) dilepton $SR \geq 3j \geq 3b$. The black lines indicate the nominal scale ($\mu_R = 1, \mu_F = 1$). The errors have been computed for the nominal scale only, as the sum of squared weights using `TH1::SetDefaultSumw2()`.

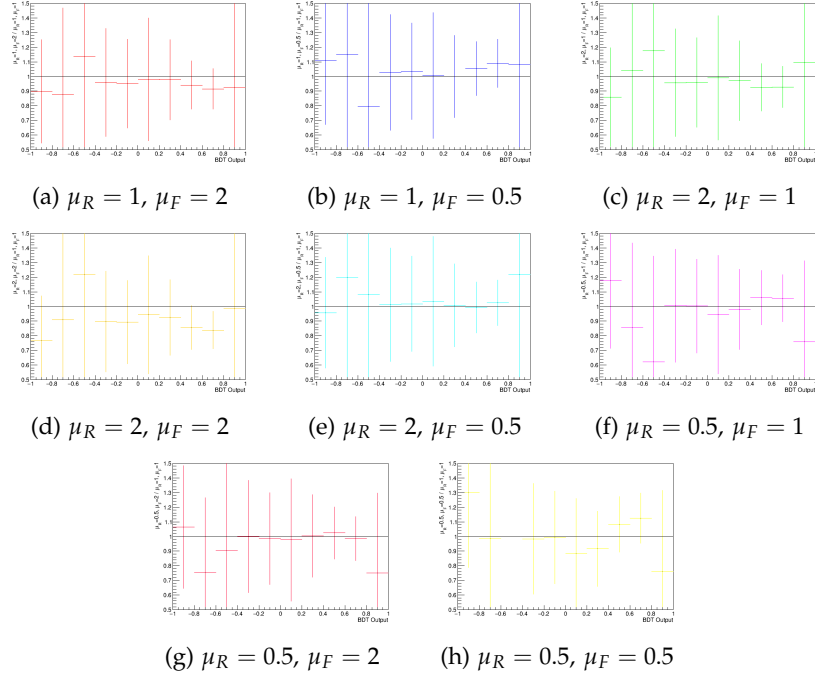


Figure 5.14: Relative change due to scale variation on the BDT output of the lepton-plus-jets final state for $m_{H^+} = 600$ GeV in the $\text{SR} \geq 5j \geq 4b$. The errors are the same relative uncertainties as in Figure 5.13.

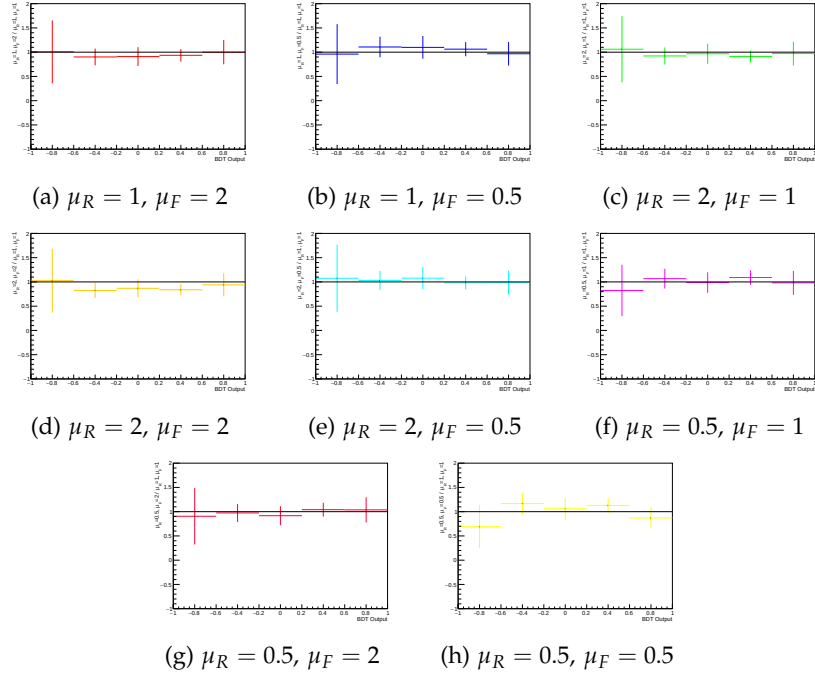


Figure 5.15: Relative change due to scale variation on the BDT output of the dilepton final state for $m_{H^+} = 600$ GeV in the $\text{SR} \geq 3j \geq 3b$. The errors are the same relative uncertainties as in Figure 5.13.

uncertainty is taken to be $\pm 6\%$, applied as a flat normalisation on the $t\bar{t}$ sample, for the inclusive $t\bar{t}$ NNLO+NNLL production [74]. This includes effects from varying the factorisation and renormalisation scales, the PDF, α_s , and the top quark mass. Furthermore, the $t\bar{t}$ +HF background cross-sections, $t\bar{t}+ \geq 1c$ and $t\bar{t}+ \geq 1b$ are allowed to float freely in the fit. The uncertainty of the $t\bar{t}$ background with 3 or more b quarks is taken to be 50% of the cross-section, to model the relative uncertainty with respect to the overall $t\bar{t}+ \geq 1b$ normalisation.

The uncertainties in the choice of a generator, in the radiation modelling and the parton shower and hadronisation process are modelled with the help of the alternative $t\bar{t}$ samples described in Section 5.4.3.1. The uncertainty inherent in the choice of NLO generator is modelled by comparing the nominal POWHEG+PYTHIA8 prediction with the SHERPA sample, generated with 5FS PDFs and a b -filter. The uncertainty of the initial and final state radiation modelling (ISR and FSR) is evaluated with the help of the two different POWHEG+PYTHIA8 samples, where the radiation has been increased (decreased) by halving (doubling) the renormalisation and factorisation scales, doubling (halving) the $hdamp$ parameter and increasing (decreasing) the A_{14} eigentune parameter $Var3c$ [103, 106]. The parton shower and hadronisation modelling uncertainties are estimated by comparing the two POWHEG samples, showered either by PYTHIA8 (nominal) or HERWIG7 [107, 108].

For the modelling of the $t\bar{t}+ \geq 1b$ backgrounds, the alternative samples are reweighted to the NLO prediction of SHERPA+OPENLOOPS, in the same way as the nominal sample (both normalisation and shape are reweighted) before the uncertainty is evaluated. Differences concerning the unweighted alternative samples are called *residual* uncertainties.

An additional uncertainty covering the 4FS vs 5FS production modes for $t\bar{t}+ \geq 1b$ is modelled by comparing the nominal POWHEG+PYTHIA8 sample with a 4F SHERPA sample with the FSR component taken from POWHEG+PYTHIA8.

To evaluate the uncertainty of the $t\bar{t} + \geq 1c$ modelling, an alternative NLO prediction with $t\bar{t} + c\bar{c}$ in the matrix element has been produced with AMCATNLO interfaced to HERWIG++, according to the method outlined in Ref. [109], as described in Section 5.4.3.1. The uncertainty is estimated as the difference between this sample and an inclusive $t\bar{t}$ sample produced with the same generator.

To ameliorate the mismodelling at high H^+ masses, the $t\bar{t}$ prediction is reweighted as mentioned in Section 5.6.1. This comes with an additional set of six systematic uncertainties, two for each of the categories $t\bar{t} + \text{light}$, $t\bar{t} + \geq 1c$ and $t\bar{t} + \geq 1b$:

- the shape uncertainty is modelled by taking the symmetrised difference between the reweighted and the original distributions (i.e. turning the reweighting “off”),
- the scale uncertainty at high p_T is estimated from the difference in the overall normalisations between the reweighted and original, for jet $p_T > 400$ GeV. This uncertainty is flat 15%.

Other SM Top Backgrounds

The total single-top cross-section uncertainty is 5% [86–88]. This is applied as a flat uncertainty to the Wt , t and s channels uncorrelated. Additionally, ISR and FSR uncertainties for the single top Wt and t channel backgrounds are estimated similarly to the $t\bar{t}$ modelling. The uncertainty on the interference between Wt and $t\bar{t}$ production at NLO [112] is assessed by comparing the default “diagram removal” scheme to an alternative “diagram subtraction” scheme. The parton showering and hadronisation modelling uncertainties in the single top t channel are estimated by comparing with samples where the parton shower generator is HERWIG instead of PYTHIA. It should be noted that the single top contributions in the t and s channel are minimal and combined with other tiny backgrounds, see the end of this subsection.

The uncertainty on the $t\bar{t}V$ NLO cross-section prediction is 15% [140], separated in QCD and PDF components, Z and W uncorrelated. The

generator uncertainty in the $t\bar{t}V$ production is estimated through comparisons with samples generated with SHERPA.

The $t\bar{t}H$ modelling uncertainty is assessed through an uncertainty on the cross-section (QCD and PDF components), and the modelling of the parton shower and hadronisation by varying the showering from PYTHIA8 to HERWIG++.

The minor $tH + X$ backgrounds, $tHjb$ and WtH are treated as one background. Its cross-section uncertainty is evaluated by considering the QCD and PDF contributions separately, each as a flat uncertainty of 20%.

The remaining minor top-related backgrounds, tWZ , $4t$, tZ , and $t\bar{t}WW$, as well as single top in the s - and t -channels, are combined to avoid spurious uncertainties from low statistics. However, uncertainties on these backgrounds are considered individually since there are different theoretical and modelling effects to consider. The treatment of the single top backgrounds is described at the beginning of the subsection. The cross-section uncertainties on the $4t$ background and tWZ are taken to be 50%, decorrelated for the backgrounds. The tZ cross-section uncertainty is decorrelated into a PDF and a QCD part, which translates to flat uncertainties of $\pm 0.9\%$ and $\pm 7.8\%$, respectively. The $t\bar{t}WW$ cross-section is also decorrelated into flat PDF and QCD parts, with normalisation $^{10.9}_{11.8}\%$ and $\pm 2.1\%$, respectively.

W/Z+Jets and Dibosons

A flat uncertainty of 40% is assumed for the $W/Z + \text{jets}$ cross-section, which is decorrelated among the jet bins. This number derives from variations of the scales and matching parameters in SHERPA simulations. In the dilepton final state, only the $Z + \text{jets}$ component is implemented, and the $W + \text{jets}$ uncertainty is included in the multi-jet background uncertainty.

An additional 30% is applied for the lepton-plus-jets final state for $W + \text{HF jets}$ components, decorrelated between 2, 3 and $\geq 4b$ -jets per

event, using the nHFJets tool. A similar uncertainty on Z+HF jets of 35% is applied in the dilepton final state for jet multiplicities larger than 2.

A 50% normalisation uncertainty on the diboson background is used, including uncertainties on the inclusive cross-section and additional jet production [141].

5.9 RESULTS

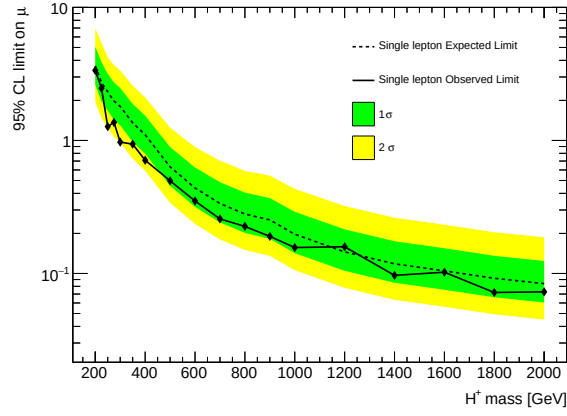
5.9.1 Upper Cross-Section Limits as a Function of H^+ Mass

The analysis set upper CLs [47] limits on the production of H^+ decaying into the $t\bar{b}$ final state. In Figure 5.16 the expected upper cross-section limits for the lepton-plus-jets and dilepton final states are shown individually, as a function of the H^+ mass. The limits are higher at low masses, flattening out at high masses. As expected, the dilepton limits are less strong for all masses but do bring a bit of improvement to the limits everywhere.

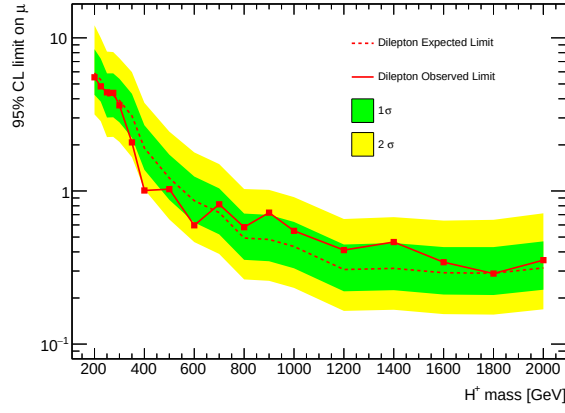
Figure 5.17a shows the evolution of the norm factors ($k_{t\bar{t}+\geq 1b}$ and $k_{t\bar{t}+\geq 1c}$) and the signal strength μ as a function of the H^+ mass, after a fit of the BDT output to collision data. The norm factors are allowed to float freely in the fit. The mean values and uncertainties are very stable as a function of H^+ mass. The k values fluctuate a bit as a function of mass, but they are fairly consistent over the entire range.

The signal strength after the fit to data, Figure 5.17b, is consistently lower than the null hypothesis (i.e. no H^+ signal), meaning that data is consistent with a negative H^+ signal.

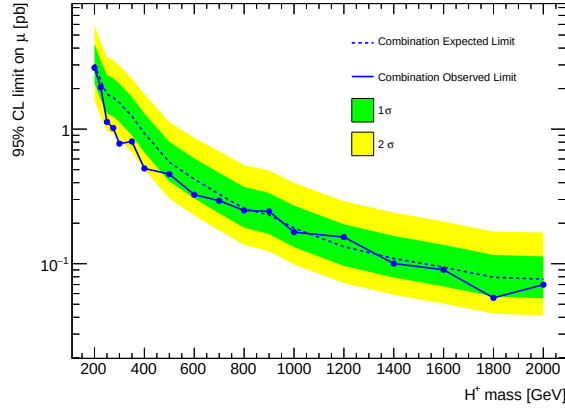
The limit plot with all limits in one single plot can be found in Figure 5.18.



(a) Lepton-plus-jets



(b) Dilepton



(c) Combined

Figure 5.16: Expected and observed upper cross-section limits on the production of H^+ as a function of the H^+ mass, in the (a) lepton-plus-jets, for (b) dilepton final states, and for (c) a combined fit to both final states.

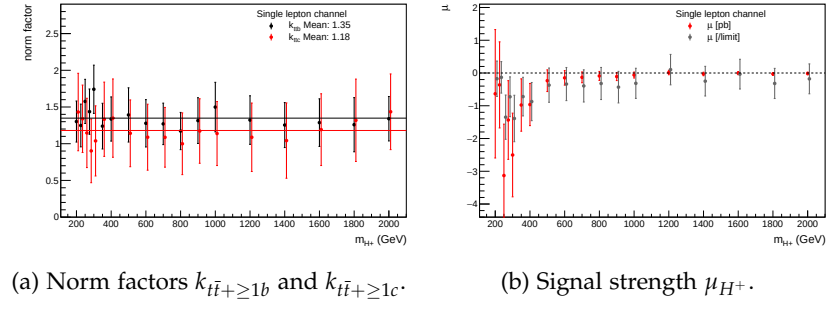


Figure 5.17: Evolution of (a) norm factors and (b) the signal strength, after fitting the output of BDTs trained at various H^+ masses to collision data, lepton-plus-jets final state. The x-axis shows the H^+ mass. In (a), the y-axis shows the norm factors $k_{t\bar{t}+\ge 1b}$ (black) and $k_{t\bar{t}+\ge 1c}$ (red), which are freely floating in the fit. The horizontal lines indicate the average overall mass points. In (b), the y-axis shows the signal strength μ_{H^+} , which indicates how much signal was found in the fit. The plot also shows the signal strength divided by the expected upper cross-section limit at that mass point.

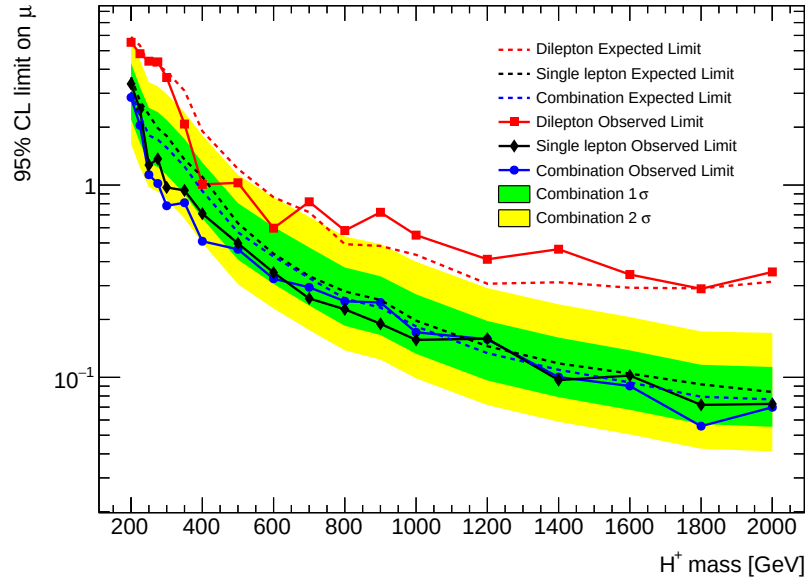


Figure 5.18: Expected and observed upper cross-section limits as a function of H^+ mass, comparing the lepton-plus-jets and dilepton final states and the combination of them.

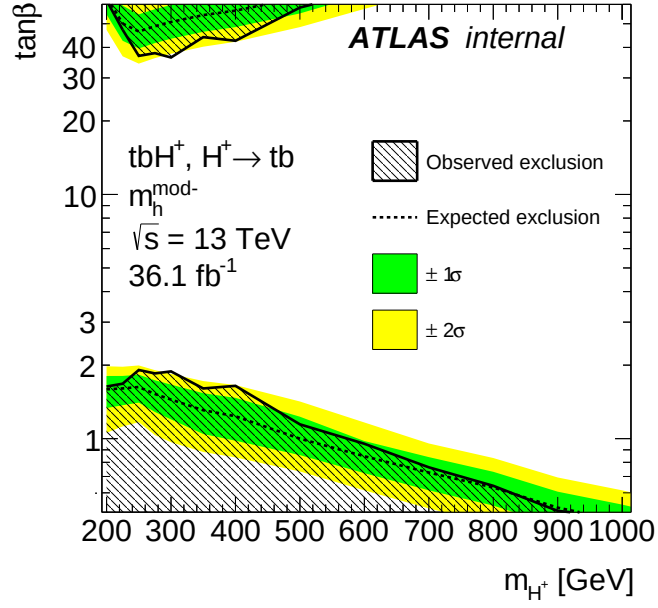
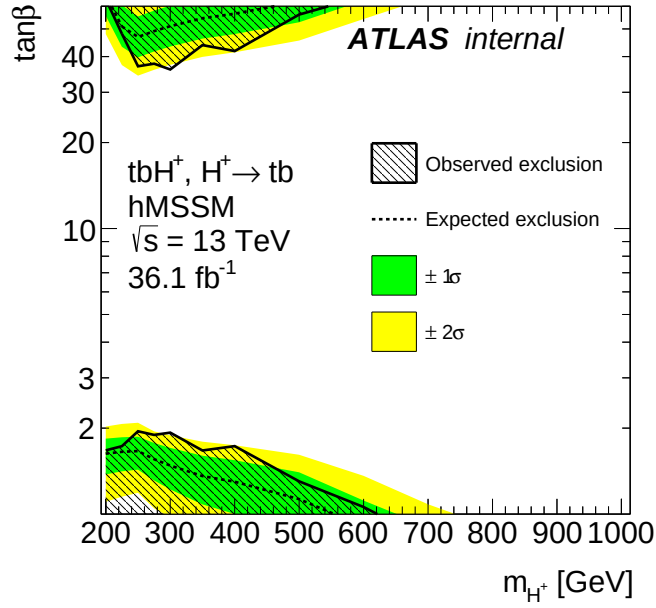
5.9.2 Interpretation of the Limits

Though this work focused solely on a simpler extension of the SM such as 2HDM, the analysis also interprets the limits for more advanced scenarios such as MSSM and hMSSM.

In Figure 5.19, limits are set on $\tan\beta$ for the m_h^{mod-} scenario of the MSSM [20, 96, 142] and the hMSSM model [143–145]. Beyond tree level, the Higgs sector is affected by choice of parameters in addition to Higgs boson masses and $\tan\beta$. For the m_h^{mod-} benchmark scenario, the top-squark mixing parameter is chosen such that the mass of the lightest CP-even Higgs boson, m_h , is close to the measured mass of the Higgs boson that was discovered at the LHC.

For H^+ masses of 200-970 GeV (200-990 GeV), some values of $\tan\beta$ in the range 0.5-1.7 (0.5-1.8) are excluded in the m_h^{mod-} model. For H^+ masses around 250 GeV the most stringent limits on $\tan\beta$ are set, and high values of $\tan\beta$ are excluded in the H^+ mass range 200-455 (200-470) GeV.

The Minimal Supersymmetric Standard Model (MSSM) is an extension to the Standard Model that realises supersymmetry. MSSM consists of taking the fields of the 2HDM extension and adding the corresponding supersymmetric partners.

(a) m_h^{mod-} model

(b) hMSSM model

Figure 5.19: Expected and observed limits on $\tan\beta$ as a function of m_{H^+} as well as bands for the 68% (in green) and 95% (in yellow) confidence intervals in the m_h^{mod-} scenario of the MSSM (a) and the hMSSM model (b). The limits are based on the combination of the lepton+jets and dilepton final states. The production cross-section of ttH and tH , as well as the branching ratios of the SM Higgs boson, are fixed to their SM values at any point in the plane. Uncertainties on the predicted H^+ cross-sections or BR are not considered.

SEARCH FOR THE DECAY OF THE HIGGS BOSON TO CHARM QUARKS

6.1 INTRODUCTION

The discovery of a Higgs boson by both the ATLAS and CMS experiments in 2012 confirmed the electroweak symmetry breaking mechanism, the process through which elementary particles gain mass via interaction with the Brout-Englert-Higgs field. The Higgs boson decays in $\gamma\gamma, ZZ, WW, bb$, and $\tau\tau$ channels have been observed. In addition, the couplings of Higgs to top (tH) [146, 147] were measured, hence the coupling to all third generation particles, top bottom and tau, were observed as well as to all the gauge bosons, but, the couplings to first and second generations are yet to be measured. Only recently ATLAS and CMS published first hints for $H \rightarrow \mu^+\mu^-$ [148, 149]. All measured properties are within the uncertainties consistent with the predictions of the Standard Model. Nevertheless, there remains much to be learned about the properties of this new particle. One of the highest priorities of the LHC physics program is the measurement of the couplings of the H boson to SM particles.

A Higgs boson with a mass of 125 GeV is predicted to decay to a pair of charm quarks around 2.9% of the time. This is the third-largest branching ratio of the Higgs boson to fermions after the decays $H \rightarrow b\bar{b}$ and $H \rightarrow \tau^+\tau^-$. Measuring the coupling of the Higgs boson to the charm quarks is of special interest as it can directly probe the Higgs mechanism for the 2nd generation of fermions. The Yukawa coupling

between the Higgs (denoted as ϕ) field and fermions (denoted as ψ) can be written as:

$$\mathcal{L}_f = -y_f \cdot [\bar{\psi}_L \phi \psi_R + \bar{\psi}_R \bar{\phi} \psi_L].$$

Expanding around the non-zero VEV leads to:

$$\mathcal{L}_f = \frac{-y_f v}{\sqrt{2}} \cdot \bar{\psi} \psi - \frac{y_f}{\sqrt{2}} \cdot h \bar{\psi} \psi,$$

where the first term can be identified as the mass term, while the second is the Yukawa coupling term (which is proportional to the fermion's mass). Measurements of the Yukawa coupling of the Higgs boson to quarks, y_c , in generations other than the third are difficult at hadron colliders due to small branching fractions, large backgrounds, and challenges in jet flavour identification.

Such a direct observation is critically important to confirm the pattern of non-universal Yukawa couplings and search for deviations from the SM as predicted in theories with an extended Higgs sector and explained in Section 2.2. BSM theories might provide a more profound explanation to the origin of the fermion mass hierarchy (i.e. why is $m_t/m_c \approx O(135)$).

Instead of looking for direct $H \rightarrow c\bar{c}$ via gluon-gluon fusion (ggF) production, one can use the associated production of the Higgs with a vector boson in order to eliminate the multi-jet background more easily. The ATLAS collaboration has previously performed a direct search for the decay $H \rightarrow c\bar{c}$ in the $Z(\ell^+\ell^-)H$ channel [10]. The original analysis, based on a $\sqrt{s} = 13$ TeV pp collision dataset corresponding to an integrated luminosity of 36.1 fb^{-1} , obtained a 95% CL limit of $\sigma(pp \rightarrow ZH) \times \mathcal{B}(H \rightarrow c\bar{c}) < 2.7 \text{ pb}$, corresponding to around $110\times$ the SM prediction.

This newer, more inclusive analysis is performed in three channels denoted by the number of leptons from the gauge boson decay, 0, 1

and 2 lepton, shown in fig. 6.1 while the most dominant background sources are $t\bar{t}$ and V +jets, due to their high cross-section at the LHC.

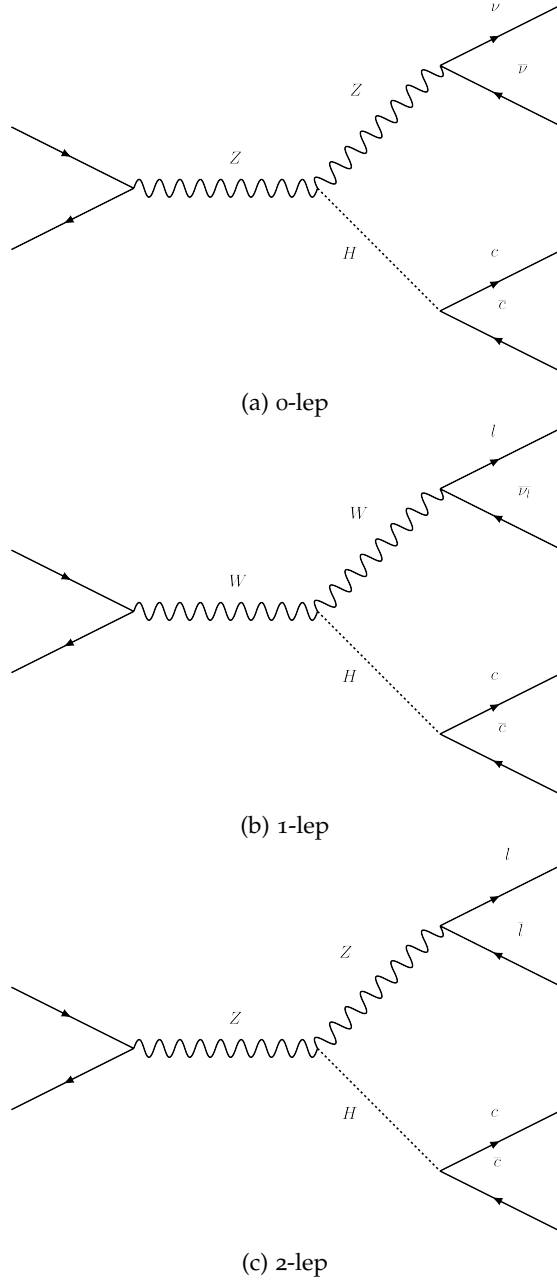


Figure 6.1: Tree level Feynman diagrams for $VHc\bar{c}$ three channels.

This thesis will focus on the 1-lepton channel analysis and includes final combined results for all three channels.

6.2 DATA AND MONTE CARLO PRODUCTION

6.2.1 Data Sample

The dataset used in this analysis includes only $\sqrt{s} = 13 \text{ TeV}$ pp collision data recorded in stable beam conditions and with all relevant sub-detectors operating normally. The corresponding integrated luminosity is 139 fb^{-1} for the combined 2015, 2016, 2017 and 2018 datasets. The dataset was collected with a suite of E_T^{miss} , single-electron and single-muon triggers.

6.2.2 Modelling of the VH Process

The VH signals include three main Higgs production modes: $ZH \rightarrow \nu\nu b\bar{b}$, $ZH \rightarrow ll b\bar{b}$ and $WH \rightarrow l\nu b\bar{b}$, where $l = e, \mu, \tau$. The ZH production modes are furthermore split into two contributions depending on the partonic initial state: $q\bar{q} \rightarrow ZH$ and $gg \rightarrow ZH$.

All $q\bar{q}$ -initiated production processes are simulated using the POWHEG generator with the MiNLO (Multiscale Improved NLO) procedure applied [150], interfaced to the PYTHIA 8 [70] MC generator applying the AZNLO tune [151] with the NNPDF3.0 PDF [102] set, while for the gg -initiated ZH contribution the analysis relies on the LO QCD POWHEG prediction interfaced with PYTHIA 8 for the modelling of parton shower, underlying event and multiple parton interactions, applying the AZNLO tune [151] with the NNPDF3.0 PDF [92] for both the matrix-element (ME) and the parton shower.

WH signal samples are normalised to the production cross-section at NNLO (QCD) and NLO (EW). The inclusive cross-section of ZH production is calculated at NNLO (QCD), and NLO (EW), the cross-section of gluon induced ZH production is then calculated at NLO (QCD), and quark induced production is taken as the difference of the two in order to avoid double counting [20]. Differential NLO EW cor-

rections are applied to qq-initiated VH productions as a function of the transverse vector boson momentum. They are derived from the HAWK Monte-Carlo generator [152, 153]. All cross-sections and branching ratios are obtained from the LHC Higgs Cross-Section Working Group prescriptions for a Higgs boson mass of $m_H = 125$ GeV.

6.2.3 Background Samples

Modelling of $t\bar{t}$

The modelling of top-quark pair production within the phase space of the analysis differs between the 0-/1-lepton and 2-lepton channel due to the differing definitions of the V -boson system. The V -boson system in the 0-/1-lepton channel typically represents a single W -boson from a single top-quark decay. Meanwhile, in the 2-lepton channel, the V -boson candidate is constructed from the ll -system, which corresponds to two leptons from two different top-quarks. This difference is reflected in the different flavour compositions of the $t\bar{t}$ process for the 0+1-lepton and 2-lepton channels, as the possibility of a hadronically decaying W -bosons in the 0-/1-lepton channel yields a larger non- bb flavour sub-component. Alternative $t\bar{t}$ MC generator configurations are used to assess the theoretical uncertainties on the top-quark pair production process for the 0-/1-lepton channel by serving as the point of comparison to the default MC configuration.

Modelling of V +jets

The production of W and Z bosons in association with jets represent a dominant background in the analysis. The ATLAS recommendations are followed for the modelling of V +jets processes [154]. The V +jet processes are simulated with SHERPA 2.2.1 [155] interfaced with the NNPDFs [92] for the ME calculation and the parton shower tuning. In order to provide good modelling of large jet multiplicities (events with many jets produced in association with the W or the Z boson,

which largely contributes to the background in the phase space considered by these analyses) SHERPA 2.2.1 provides a combination of different matrix elements with different parton multiplicities: up to two extra partons are included in the next-to-leading order (NLO) ME, and three or four extra partons are included at leading order (LO) in QCD. The merging of different parton multiplicities is achieved through a matching scheme based on the CKKW-L [156] [157] merging technique using a merging scale of $Q_{cut} = 20$ GeV. The modelling of even higher jet multiplicities relies on the parton shower algorithm. The parton shower and underlying event models used are the ones provided internally by SHERPA.

The SHERPA 2.2.1 generator adopts a full 5-flavour scheme, with massless b-quarks and c-quarks in the matrix elements, while massive quarks can be produced in the parton shower (heavy flavours can be produced directly in the scattering processes of the underlying event as well).

Alternative samples for the MC prediction and the modelling of V +jets processes are considered. The alternative samples provide a cross-check for the nominal prediction used in the analyses (understanding how much the nominal MC prediction is robust for certain variables and phase space regions can provide useful insights to check possible data/MC discrepancies or MC mismodelling) and are used to estimate modelling uncertainties on the nominal prediction. The alternative samples considered for V +jets processes are generated using MADGRAPH 5 [69] interfaced to PYTHIA 8 for the modelling of the parton shower and the underlying event.

Modelling of Di-boson Processes

The diboson background consists of final states arising from WW , WZ , and ZZ events. Several diboson processes can give a significant contribution in the analysis phase space: $Z \rightarrow c\bar{c}$ and $Z \rightarrow \nu\bar{\nu}$ for the o-lepton channel; $Z \rightarrow c\bar{c}$ and $Z \rightarrow ll$ for the 2-lepton channel; and

$Z \rightarrow c\bar{c}$ and $W \rightarrow l\nu$ for the 1-lepton channel. Additional processes can give smaller contributions arising from the mistagging of a jet from the W decay or the failed reconstruction of one of the leptons, for example, in $Z \rightarrow \nu\bar{\nu}$ and $W \rightarrow q\bar{q}$, $Z \rightarrow ll$ and $W \rightarrow q\bar{q}$, and $W \rightarrow l\nu$ and $W \rightarrow q\bar{q}$ production.

The qqVZ samples are generated using SHERPA 2.2.1 and the ggZZ samples using SHERPA 2.2.2. The default SHERPA 2.2 generator provides an NLO calculation for the diboson cross-sections for 0,1 extra partons and a LO calculation for 2+ extra partons in the hard scatter. Using these cross-sections, each diboson is normalised according to its MC generator predicted cross-section, including the k -factor corrections to account for higher order diagrams.

Modelling of Single-Top Processes

The default single-top (s-channel, t-channel, and Wt -channel) configuration used within the analysis utilises the POWHEG MC generator for the hard scattering prediction with the NNPDF30 NLO PDF set, interfaced with the PYTHIA 8 algorithm for the simulation of parton shower, underlying event and multiple parton interactions. The A14 parton shower tune is applied to the parton shower generator with the corresponding LO form of the PDF set, NNPDF23 LO. Single top samples in the t- and s-channels are generated applying a lepton filter to require the leptonic decay of the W boson, while Wt -channel samples are generated both applying a dilepton filter (both W bosons decay leptonically) and inclusively; in these analyses, the inclusive samples are used.

6.3 EVENT SELECTION

First of all, events are required to have good data quality and passing event cleaning criteria, i.e. the GRL requirement (data only), primary vertex, bad-jets rejection, and corrupted events rejection of LAr, Tile and SCT bursts, as recommended by the Data Preparation Group. Then, the events are categorised according to the number of VH-loose leptons, defined in Table 6.1 and Table 6.2: the o-lepton channel is composed of events with zero VH-loose leptons, the 1-lepton channel with exactly one WH-signal lepton, and the 2-lepton channel with exactly two ZH-signal leptons.

Electron Selection	p_T	η	ID	Isolation
VH-loose	>7 GeV	$ \eta < 2.47$	LH Loose	FCLoose
ZH-signal	>27 GeV	$ \eta < 2.47$	LH Loose	FCLoose
WH-signal	>27 GeV	$ \eta < 2.47$	LH Tight	FixedCutHighPtCaloOnly

Table 6.1: Electron Selection

Muon Selection	p_T	η	ID	Isolation
VH-Loose	>7 GeV	$ \eta < 2.7$	Loose quality	FixedCutLoose
ZH-Signal	>27 GeV	$ \eta < 2.7$	Loose quality	FixedCutLoose
WH-Signal	>25 GeV	$ \eta < 2.5$	Medium quality	FixedCutHighPtTrackOnly

Table 6.2: Muon Selection

In the 1-lepton channel, single electron events are required to be triggered by at least one of the lowest un-prescaled single-electron triggers. Events in the single muon channel and with $p_T^V > 150$ GeV are triggered using the same E_T^{miss} trigger as used in the o-lepton channel to obtain high trigger efficiency for single high p_T muons.

In all three channels, only events in which the Higgs candidate mass, $m_{c\bar{c}}$, is higher than 50 GeV are considered.

6.3.1 Event Selection Common to All Channels

In all three lepton channels, events are required to contain at least two signal jets, defined in Table 6.3. Exclusive categories of events, depending on the number of selected jets — signal+forward jets — they contain, are defined in order to maximise the signal significance: events containing two jets comprise the 2-jet category, three jets the 3-jet category, three or more jets 3+-jets category. In the 0- and 1-lepton channels, the 2- and 3-jet categories are used, and events with four or more jets are rejected due to the high $t\bar{t}$ background contamination. In the 2-lepton channel, where the high jet multiplicity regions result in a 5-6% gain in expected signal significance, the 2-jet and 3+-jet categories are used.

Jet Category		Selection Requirements
Forward Jets		$p_T > 30 \text{ GeV}$
		$2.5 \leq \eta < 4.5$
Signal Jets	JVT > 0.59 for $ \eta < 2.4$ and > 0.11 for $2.4 < \eta < 2.5$ if $p_T < 120 \text{ GeV}$	$p_T \geq 20 \text{ GeV}$ and $ \eta < 2.5$

Table 6.3: AntiKt4EMTopoJets selection requirements.

The kinematic fit based on Run-1 VH(bb) analysis has been developed for Run-2. Transfer functions were derived separately for jets that contain a muon and jets that do not. The kinematic fit yields an additional 20-30% improvement in $m_{c\bar{c}}$ mass resolution with respect to the muon-in-jet corrected mass resolution. The kinematic fit correction is applied to 2-jet and 3-jet events. It is not used for events with 4 or more jets because the improvement is smeared out by the additional jets in the events.

The mass of the Higgs candidate, $m_{c\bar{c}}$, is reconstructed using the two leading p_T jets, and the event is categorised according to the number of c-tags among the two leading p_T jets: 0 c-tag, 1 c-tag or 2 c-tag.

Though this thesis is focusing on the 1-lepton channel, for completeness, Table 6.4 summarises the selections applied in each of the three channels.

The momentum resolution of c-tagged jets is improved by applying the *muon-in-jet* described in the next paragraph. These corrections are only applied to the jet-derived variables after the events have passed the analysis selection and are not propagated to the E_T^{miss} variables.

MUON-IN-JET CORRECTION In practice, muons and electrons are typically removed by the overlap removal procedure. Nevertheless, they influence jet energy reconstruction. The electron deposits energy in the electromagnetic calorimeter, and its energy is accounted for in the jet. Muons are minimum ionising particles, depositing deficient energy in the calorimeter. To correct for it, the muon 4-vector is added to the jet 4-vector, while the energy deposited by the muon is removed.

If more than one muon is found within the jet cone, the muons are ordered according to the distance to the jet axis and only the muon closest to the jet axis is considered for the jet energy correction. The muon in jet selection is summarised as follows:

- $p_T > 4\text{GeV}, |\eta| < 2.7$
- pass Medium quality cut
- $\Delta R(\text{jet}, \text{muon}) < \min(0.4, 0.4 + 10/p_T^\mu)$

The $p_T > 4\text{GeV}$ and the quality cut was studied in the context of the VH(bb) analysis and was not optimised further for VH(cc). The variable radius cone ΔR is introduced to avoid prompt muons. The selection is applied without interference with preselection and overlap removal. The MET and MET significance, as well as the $\Delta R(\text{leading jet}, \text{sub-leading jet})$ in the event selection, is applied using uncorrected jets.

Reconstructed jets with the muon-in-jet correction are labelled *OneMu*.

Systematic uncertainties related to the muon selection for the muon-in-jet correction have been studied in Run-1 and found to be negligible.

The muon-in-jet correction is used as a default in the 0-, 1-, and 2-lepton channels and the improvement on the VH(cc) signal resolution is up to 6%.

6.3.2 1-lepton Channel Specific Selection

In addition to the selections common to the 0-, 1- and 2-lepton channels, additional event selections are implemented to confine events containing a W decaying to a lepton and a neutrino in addition to the Higgs boson. In the 1-lepton channel, events are required to contain exactly 1 WH-signal lepton; events containing additional VH-loose leptons are vetoed. Due to the increased multi-jet contribution at low transverse momenta, only events with $p_T^W > 150$ GeV are considered. Furthermore, a QCD background reduction selection is employed to events in the electron channel, requiring the event to have $E_T^{\text{miss}} > 30$ GeV. Events in the 1-lepton channel are also required to satisfy $m_T^W < 120$ GeV.

6.3.3 Analysis Strategy and Event Categorisation

One of the main challenges in this analysis is to achieve orthogonality with respect to the VH(bb) analysis. Therefore, the analysis has moved from a traditionally binary usage of jet flavour tagging (*tagged* or *not tagged*) to a ternary one (*b-tagged*, *c-tagged* or *not b-tagged and not c-tagged*). Such a flavour tagging paradigm did not exist in the ATLAS collaboration. In this analysis, the flavour tagging is applied sequentially, first the b-tagging requirement, then the c-tagging requirement if the b-tagging requirement is not satisfied.

Since the VH(cc) analysis is very similar to the VH(bb) one, designing a proper flavour tagging strategy is necessary in order to remove

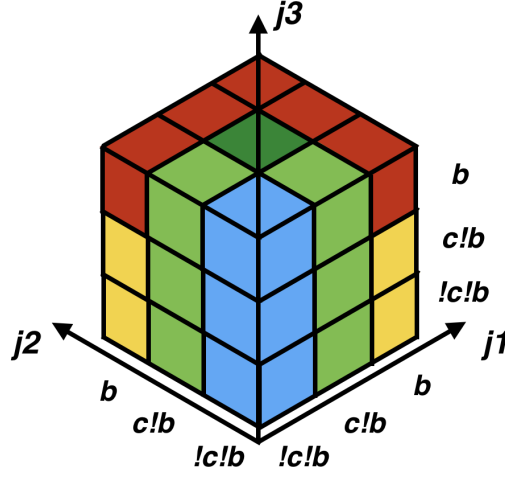


Figure 6.2: VH(cc) analysis categorisation in events with three jets. Red cubes represent events with two b-tags defining the VH(bb) signal region. Dark green cubes represent events with two c-tags. Light green cubes events with one c-tag are both considered for the VH(cc) signal region. Blue cubes are events failing both sequential b-tagging and c-tagging requirements. Yellow cubes correspond to events failing the VH(cc) signal region tag categorisation as they include a b-tagged jet among the two leading jets in the event.

the overlap between the two analyses, achieve orthogonality and reduce the VH(bb) contamination in the VH(cc) analysis. The VH(bb) event selection and categorisation is based on events with exactly two b-tagged jets ($n_{b-tags} = 2$). Therefore, the easiest way to remove the overlap between the two analyses is to make sure that the VH(cc) signal region always contains less than two b-tagged jets ($n_{b-tags} < 2$).

In all three channels of the analysis, c-tagging is applied to the two leading p_T signal jets using a combination of DL1 and MV2c10 algorithms to define a c-tagging requirement simultaneously with a b-tagging veto. A more detailed description of the flavour tagging paradigm used in this analysis can be found in Section 6.4.

Alternative strategies to reconstruct the Higgs candidate with all the jets in the events tested for the c-tagging requirements were shown to be less sensitive. This can be explained by the low efficiency of the c-tagger and the kinematic topology of signal events where the two leading p_T jets are most likely produced in the decay of the

Higgs boson. Events with three jets require special treatment due to the possibility of finding two b-tags on the second and third jet in the event. This effect is illustrated on Figure 6.2. An extra b-tagging veto requirement defined with the MV2c10 algorithm at 70% efficiency is applied on all jets $n_{\text{Jets}} > 2$ in the event. This extra requirement guarantees $n_{\text{b-tags}} < 2$ for every event in the analysis therefore effectively achieving orthogonality with the $VH(b\bar{b})$ analysis. The effect of this extra selection was studied and shown to have a marginal impact on the 2-lepton analysis while mostly removing $t\bar{t}$ background in the 0- and 1-lepton channels.

6.3.4 Additional Selection Criteria

In this cut-based analysis, the invariant mass distribution m_{cc} is used as a discriminant to calculate the final results. In order to improve the sensitivity of the measurements, events are separated in different p_T^V categories. For the 0- and 1-lepton channels, two p_T^V categories are considered: from 150 GeV to 250 GeV, greater than 250 GeV. For the 2-lepton channel, an additional p_T^V category, from 75 GeV to 150 GeV, is also included in this analysis.

The $\Delta R(\text{jet1}, \text{jet2})$ variable is used for further reducing the background contamination. Jets originating from a Higgs, vector boson resonance or produced via gluon splitting are expected to have a lower ΔR value in the p_T^V regimes considered in this analysis. Therefore, optimisation on the upper selection of the $\Delta R(\text{jet1}, \text{jet2})$ is performed and the significance of the $VH(cc)$ signal is used as a figure of merit. The primary goal of this additional selection is to remove $t\bar{t}$ background and the non gluon splitting contribution of V +jets. Figure 6.3 illustrates the signal efficiency and significance trend with $\Delta R(\text{jet1}, \text{jet2})$ performed on the 140 fb^{-1} dataset where the signal region selection at the exception of the ΔR upper selection. The optimal values are found to be consistent across the 0-, 1- and 2-lepton channels.

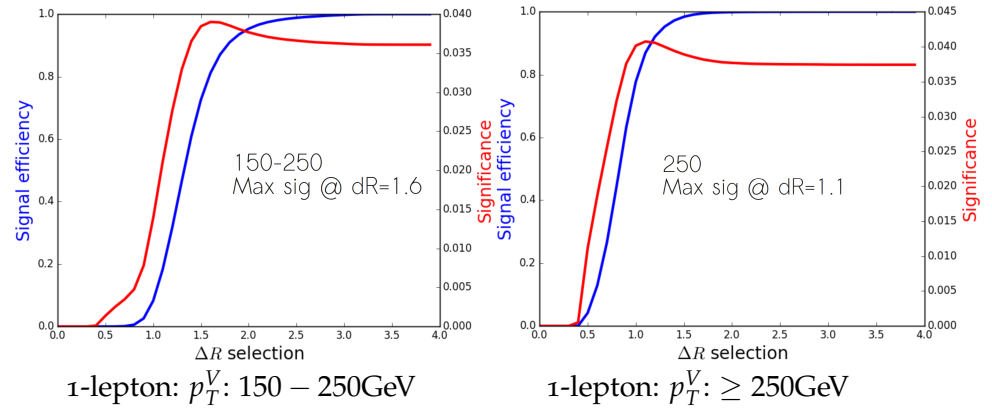


Figure 6.3: Working point optimisation for $\Delta R(\text{jet1}, \text{jet2})$ performed on 140 fb^{-1} with the signal region selection applied at the exception of the ΔR upper selection. The blue line gives the signal efficiency and the red line - the significance.

Common Selections	
Jets	≥ 2 signal jets
c -jets	1 or 2 c -tagged signal jets
b -tagging veto	≥ 3 signal or forward jets
Leading jet p_T	> 45 GeV
$m_{c\bar{c}}$	> 50 GeV
$\Delta R(\text{jet1}, \text{jet2})$	[75,150] GeV, $\Delta R \leq 2.3$
	[150,250] GeV, $\Delta R \leq 1.6$
	[250,∞] GeV, $\Delta R \leq 1.2$
0 Lepton	
Trigger	lowest un-prescaled E_T^{miss} triggers
Leptons	0 VH-loose lepton
E_T^{miss}	> 150 GeV
$E_{T, \text{trk}}^{\text{miss}}$	> 30 GeV
S_T	> 120 (2 jets), > 150 GeV (3 jets)
$ \min \Delta \phi(E_T^{\text{miss}}, \text{jet}) $	$> 20^\circ$ (2 jets), $> 30^\circ$ (3 jets)
$ \Delta \phi(E_T^{\text{miss}}, h) $	$> 120^\circ$
$ \Delta \phi(\text{jet1}, \text{jet2}) $	$< 140^\circ$
$ \Delta \phi(E_T^{\text{miss}}, E_{T, \text{trk}}^{\text{miss}}) $	$< 90^\circ$
p_T^V regions	[150, 250] GeV, [250, ∞] GeV
1 Lepton	
Trigger	e channel: un-prescaled single electron μ channel: above see 0-lepton triggers
Leptons	1 WH-signal lepton > 1 VH-loose lepton veto
E_T^{miss}	> 30 GeV (e channel)
m_T^W	< 120 GeV
p_T^V regions	[150, 250] GeV, [250, ∞] GeV
2 Lepton	
Trigger	un-prescaled single lepton
Leptons	2 VH-loose leptons (≥ 1 ZH-signal lepton)
	Same flavour, opposite-charge for $\mu\mu$
$m_{\ell\ell}$	$81 < m_{\ell\ell} < 101$ GeV
p_T^V regions	[75,150], [150, 250], [250, ∞] GeV

Table 6.4: Summary of the signal event selection in the 0-, 1- and 2-lepton analyses.

6.4 FLAVOUR TAGGING

The jets from the Higgs boson in the $H \rightarrow c\bar{c}$ signal process originate from c -quarks. Therefore, distinguishing c -jets from light-, τ - and b -jets is necessary in order to maximise the signal sensitivity. This procedure is referred to as c -tagging.

The algorithm for c -jet tagging used in this analysis is a combination of the DL1 and MV2 discriminants. Both are described below.

*IP2D and IP3D
exploit the large
impact parameters of
tracks originating
from b and c
hadrons.*

MV2 The MV2 is a high-level algorithm [158] consists of a BDT algorithm that combines the outputs of the low-level tagging algorithms IP2D and IP3D [158]. The kinematic properties of the jets, p_T and $|\eta|$, are included in the training in order to take advantage of the correlations with the other input variables. The output is one final discriminant expressing the likelihood for a jet to be a b -jet as opposed to both c - and light-jets, and the output of the MV2c10 algorithm is a score between -1 and 1.

$$p_b + p_c + p_u = 1$$

DL1 The second high-level tagging algorithm, DL1 [158], is a more advanced deep neural network (DNN). The input variables of DL1 are identical to MV2 with the addition of JETFITTER c -tagging variables such as the properties of the secondary vertex, invariant mass and number of tracks. The DL1 is a multi-label classifier, thus, providing three outputs corresponding to the probabilities of a jet being a b -jet p_b , a c -jet p_c and a light-jet p_u , each score between 0 and 1 .

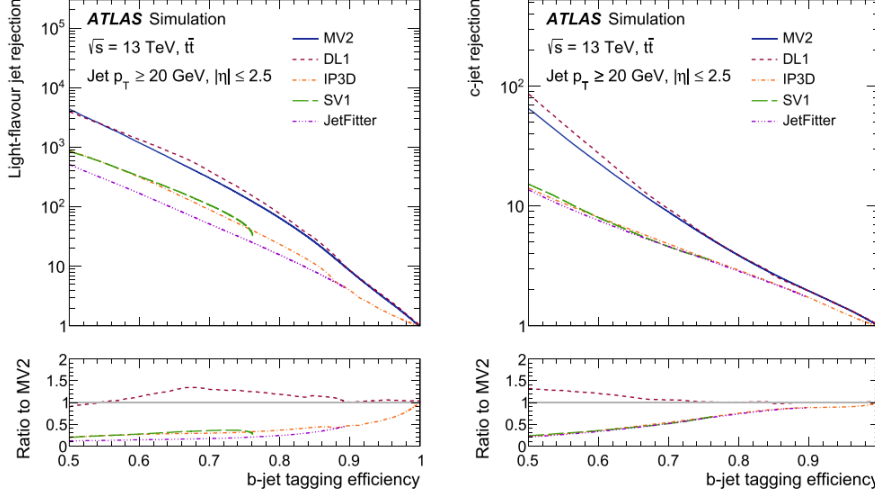


Figure 6.4: The light-flavour jet and c -jet rejections versus the b -jet tagging efficiency for the IP3D, SV1, JETFITTER, MV2 and DL1 b -tagging algorithms evaluated on the baseline $t\bar{t}$ events. Figure from [159].

Both algorithms are trained on *hybrid samples* composed of $t\bar{t}$ events with at least one lepton from the leptonic decay of the W boson and hadronically decaying Z' events to compensate the steep falling of the $t\bar{t}$ spectrum above 250 GeV. In total, 8 million events have been used for the training, 5 million for $t\bar{t}$ and 3 million for the Z' .

While primarily designed for b -jet identification, the DL1 outputs are used following the Neyman-Pearson lemma [160] to define a one-dimensional c -tagging discriminant:

$$DL_1 = \ln \frac{p_c}{f p_b + (1 - f) p_u}, \quad (6.1)$$

where f corresponds to a weight attributed to the probabilities p_b and p_u . A large value of f will increase the discrimination of c -jets and b -jets, while a low value will predominantly increase the discrimination of c -jets and light-jets.

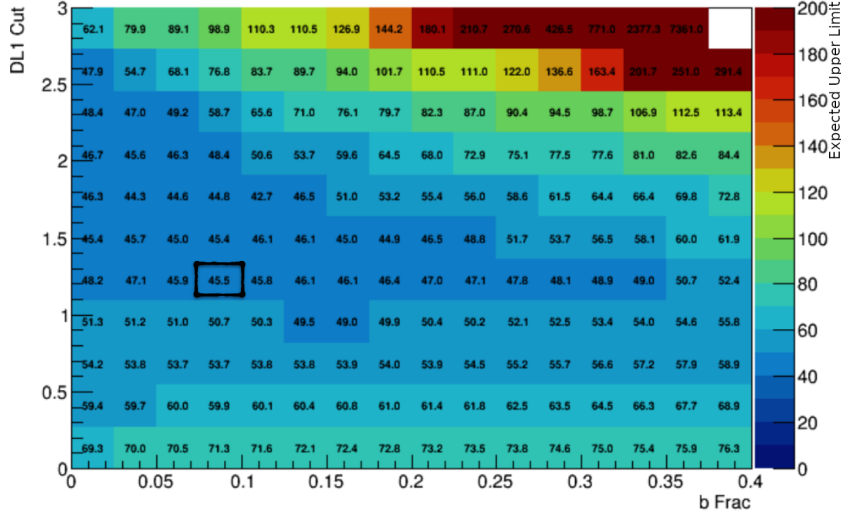
In the analysis, the MV2 discriminant is used as a b -tagger in order to define a b -tag veto. The working point considered has a 70% efficiency on b -jets and corresponds to the same b -tagger used in the $VH(b\bar{b})$ analysis.

The DL1 and MV2 algorithms are combined in a two-dimensional distribution in order to achieve c -tagging. In the events considered in the analysis, the b -tagging veto is required by default on all jets in the events in order to achieve orthogonality with the $H \rightarrow b\bar{b}$ analysis. For the two leading p_T jets, the b -tagging veto is required together with the DL1 selection. Dedicated flavour tagging calibrations are required to account for the (c -tagging and b -tag veto) to account for the correlation between the c -tagging requirement and the b -tagging veto directly in the calibrations. For all the other jets in the event, the b -tag veto is applied in a stand-alone way and the standard flavour tagging calibrations derived by the ATLAS flavour tagging CP group are used.

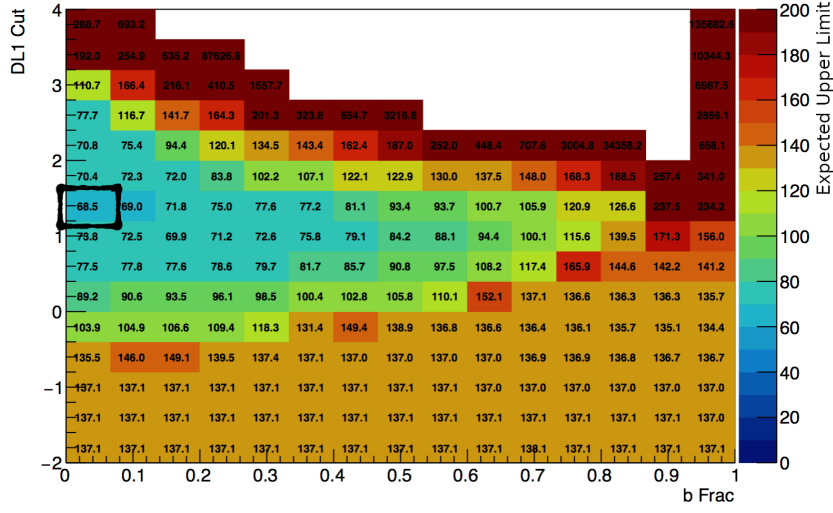
6.4.1 Working Point Optimisation

An optimisation study is performed to define the optimal working point for the DL1 discriminant assuming a default b -tag veto using the MV2c10 70% working point. From the Equation (6.1), two parameters are optimised to achieve the best c -tagging performance: the selection working point on the DL1 discriminant and the weight f . A grid optimisation is performed considering the events passing the signal region selections of the 0-, 1- and 2-lepton channels. The optimisation is performed to maximise the sensitivity of the signal, and two different figures of merit are used considering only statistical uncertainties: the expected limit at 95% confidence level (CLs) for the 0- and 1-lepton channels and the significance $z = \frac{s}{\sqrt{s+b}}$ for the 2-lepton channel.

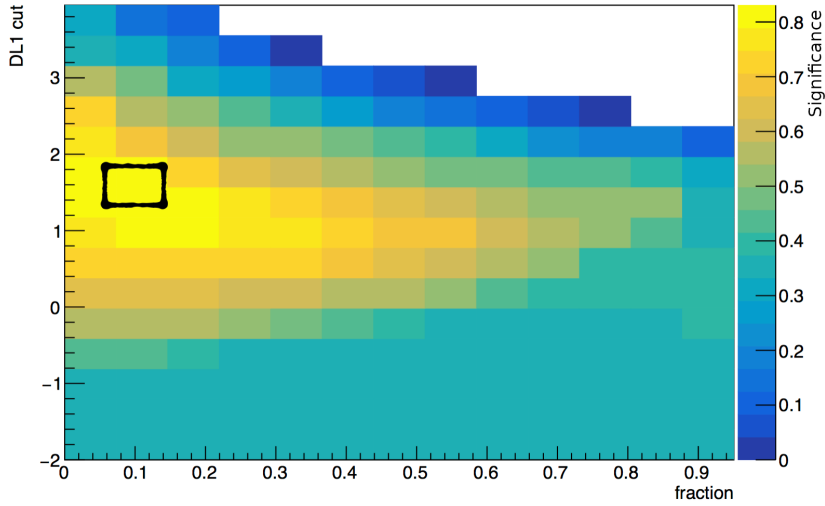
Figure 6.5 shows the result of the grid optimisation on the working point selection and the weight f of the DL1 definition. The three optimisations across all the channels converged towards the same working point, and the optimum lies in a stable region. Therefore, a unique working point is defined for all three channels and the optimal value for the DL1 selection is found to be $DL1 > 1.3$ with a weight



o-lepton



1-lepton



2-lepton

Figure 6.5: Working point optimisation for the simultaneous c -tagging and b -tag veto for the three channels, o-lepton, 1-lepton and 2-lepton. The black rectangle corresponds to the optimal value.

value of $f = 0.08$. This low optimal value for the weight f indicates that the signal sensitivity is more sensitive to the rejection of light-jets than b-jets. This can be explained by the default b-tag veto requirement removing up to 70% of b-jets before the optimisation. The performance of the optimal working point of the c-tagging, including the b-tag veto working point, is summarised in Table 6.5. Due to the similarities between the tau-jets and c-jets signature in the detector, their c-tagging performance is similar.

The impact of the $VH(\tau\tau)$ background on the analysis sensitivity is estimated assuming a branching ratio $BR(H \rightarrow \tau\bar{\tau}) \approx 2 \times BR(H \rightarrow c\bar{c})$ and a tau hadronic decay $BR(\tau \rightarrow \text{hadronic}) \approx 65\%$. The τ -identification used for the overlap removal procedure is the 60-75% working point (for 3 prongs or 1 prong), effectively resulting in a conservative 40% of hadronic τ -jets reconstructed and labelled as jets in the events passing the signal region selection. Therefore, the fraction of $VH(\tau\bar{\tau})$ with respect to $VH(c\bar{c})$ is:

$$\frac{VH(\tau\bar{\tau})}{VH(c\bar{c})} \approx 2 \times (0.4)^2 \times (0.6)^2 \approx 10\%. \quad (6.2)$$

For an expected sensitivity of $20 - 30 \times \text{SM}$ at 95% CL, the $VH(\tau\bar{\tau})$ background represents an impact of $\mathcal{O}(0.1\%)$ on the sensitivity. A similar estimate is performed for $VZ(\tau\bar{\tau})$ where the branching ratio is $BR(Z \rightarrow \tau\bar{\tau}) \approx 0.25 \times BR(Z \rightarrow c\bar{c})$ and the impact is $\mathcal{O}(1\%)$ for an expected significance $z \approx 3\sigma$. The $VH(\tau\tau)$ and $VZ(\tau\tau)$ are considered negligible in this analysis. A validation was performed in the 0-lepton channel by extracting the $Z(\nu\nu)Z(\tau\tau)$ yield from a diboson fully leptonic sample and comparing it to the $Z(\nu\nu)Z(cc)$ yield in the signal region of the analysis. The ratio $Z(\nu\nu)Z(\tau\tau)/Z(\nu\nu)Z(cc) = 1\%$ is found to be in agreement with the estimate.

Requirement	Efficiency			
	c-jets	b-jets	light-jets	tau-jets
$DL1 \geq 1.3$ and $f = 0.08$ and $MV2c10 \leq 0.83$	27%	8%	1.6%	25%

Table 6.5: Efficiencies of the c-tagging including the b-tag veto working point chosen by the optimisation study.

6.5 SYSTEMATIC UNCERTAINTIES

The origin of such uncertainties was already discussed in Section 5.8. As the uncertainties' final values are only derived during the fit, their values shall be presented in Section 6.6.

6.5.1 Experimental Systematic Uncertainties

Experimental systematic uncertainties associated with the simulated response of the detector, in terms of trigger efficiency, object reconstruction efficiency, object energy scale and resolution uncertainties, as well as flavour tagging, are all assessed following the relevant ATLAS combined performance (CP) groups.

Luminosity and Pile-up Simulation

The uncertainty on the measurement of the combined 2015-2018 integrated luminosity is 1.7%. It is derived, following a methodology similar to that detailed in [161], from a preliminary calibration of the luminosity scale using x-y beam-separation scans performed during dedicated running periods in each year.

The uncertainty on the distribution of pile-up vertices in the simulation is computed by changing the nominal value of the $\langle \mu \rangle$ data rescaling factor (1.0/1.09) to 1.0/1.0 and 1.0/1.18 to obtain a change identified with a $\pm 1\sigma$ variation [162].

Electron Reconstruction Efficiency

Electron reconstruction and identification efficiency and their associated systematic uncertainties are determined in bins of transverse momentum and rapidity from data and MC using $Z \rightarrow e^+e^-$, $W \rightarrow ev$, $J/\psi \rightarrow e^+e^-$ events [163]. The lower p_T threshold recommended for electrons is 4.5 GeV. In this range, the electron reconstruction efficiency is 97-99%. The identification efficiency scale factor is available from $p_T > 4.5$ GeV. The scale factor of the combined cuts of impact parameter significance and on $|z_0 \sin \theta|$ is included in the ID efficiency scale factor. The isolation efficiency scale factor is also provided. In the current recommendation scale, factors for $p_T > 150$ GeV are unity due to the low $Z \rightarrow e^+e^-$ yields in that region. Therefore, a systematic uncertainty of an additional $\pm 2\%$ is assigned above 150 GeV.

Electron Energy Scale and Resolution

Over 60 systematic variations are considered in the electron energy calibration. However, the electron energy scale and resolution systematic uncertainties have sub-leading contributions to the overall sensitivity of the analysis and a simplified systematic uncertainty model using just two (combined) nuisance parameters are used.

Muon Reconstruction Efficiency

Reconstruction, isolation and track-to-vertex association (d_0 significance and $z_0 \sin \theta$) scale factors are calculated with samples of $Z \rightarrow \mu^+\mu^-$ and $J/\psi \rightarrow \mu^+\mu^-$ events in the 2015-2018 dataset [164]. The scale factors are valid in the full p_T range relevant for this analysis. The J/ψ measurement is important below $p_T < 15$ GeV while the Z measurement is more accurate for $p_T > 15$ GeV and separate uncertainties are provided for each region. The isolation scale factor and its systematic uncertainties are supported in the range of $3 < p_T < 500$ GeV. For muons with $p_T > 500$ GeV, the SF measured in the 300-500 GeV region is provided, with a total systematic uncertainty of 2%. The

scale factor for the combined requirements on the impact parameter significance and the $|z_0 \sin \theta|$ is also provided.

Muon Momentum Scale and Resolution

Muons have been calibrated with a sample of $Z \rightarrow \mu^+ \mu^-$ events for the high p_T region ($p_T > 20$ GeV) and a sample of $J/\psi \rightarrow \mu^+ \mu^-$ events for the low p_T region ($p_T < 20$ GeV).

Jet Energy Scale and Resolution

Jet related systematic uncertainties are implemented through the ATLAS common used tool `JetUncertaintiesTool`. The number of uncertainties is divided into different reduction categories. An analysis may choose between the full set of nuisance parameters (~50 baseline parameters), a globally reduced set of 21 parameters or a strongly reduced set of four parameters. The baseline set uses *in-situ* analyses, η intercalibration, high- p_T jets, pile-up, flavour composition, flavour response, b -jets, and punch-through jets. Each of these areas of concern may be accessed separately using the tool.

The reduced sets comprise nuisance parameters made from performing a principal component analysis and combining the resulting variations in such a way as to preserve correlations in certain jet-kinematic regions. If an analysis will later combine their result with another group, these reduced sets may not be used. The 21 parameter uncertainty set is used in this analysis.

The most prominent sources of jet-related uncertainty are the uncertainties from the jet energy scale (JES) and jet energy resolution (JER). The determination of the JES is documented in [165] and the JER is determined from data-MC comparisons. While the uncertainty scheme includes 21 parameters, one is only relevant for samples produced with a fast simulation (AFII).

The VHcc analysis aims from the beginning to a combined result with the higher sensitivity analysis VHbb. Thus, every choice is made with the combined result in mind.

Flavour Tagging

The uncertainties related to jet c -tagging are implemented as event-weight systematic uncertainties (scale factors) and are summarised in Table 6.6.

The individual systematics for c -tagging contain up to 20 elements per jet flavour (b , c , τ , and light flavour). In an effort to alleviate the burden of accounting for ~ 160 systematics variations, the ATLAS flavour tagging CP group has performed a principal component analysis to reduce the number of systematic variations.

Systematic Uncertainty Name	Description
FT_EFF_Eigen_Light_{0-3}	Light flavour jet c -tagging efficiency uncertainty
FT_EFF_Eigen_C_{0-3}	c -jet c -tagging efficiency uncertainty
FT_EFF_Eigen_B_{0-7}	b -jet c -tagging efficiency uncertainty
FT_EFF_extrapolation	c -tagging efficiency uncertainty for high- p_T b - and c -jets
FT_EFF_extrapolation_from_charm	c -tagging efficiency uncertainty on τ -lepton jets

Table 6.6: Summary of the jet c -tagging efficiency uncertainties.

6.5.2 *Signal and Background Modelling Systematic Uncertainties*

The analysis makes use of simulated samples generated according to the recommendations from the ATLAS Physics Modelling Group. These generators are the so-called ‘nominal’ ones. Differences arise from the use of alternative generators, due to different orders of perturbation theory corrections, for instance, and have to be taken into account. Moreover, variations of the factorisation and renormalisation scales in the nominal generators also need to be considered.

These modelling uncertainties are taken into account through the study of discrepancies in the processes normalisations for the nominal and alternative generators, as well as differences in the acceptance ratios and $m_{c\bar{c}}$ shape distributions between the generators (including scale variations).

Normalisation uncertainties are the result of calculations taking into consideration the cross-sections and branching ratios of the respective processes. These are constrained in the global likelihood fit for well understood processes and free to float otherwise.

The $m_{c\bar{c}}$ shape uncertainties come from the ratios between the normalised distributions for nominal and alternative generators in each region and in the fit range. The final systematic uncertainty is then applied for each process separately, and its value is parametrised as a fit to the ratio of the alternative over nominal normalised distributions.

A study was performed by splitting all the possible tag, p_T^V and jet multiplicity categories and looking at the various shapes of the background considered. It was decided to combine various categories if the different shapes were in agreement within their statistical uncertainty in order to derive the fitted curve used as an uncertainty. A selection $m_{c\bar{c}} > 50$ GeV is applied on all the histograms.

For V +jets, acceptance ratios and shape uncertainties are derived for Vcc , Vcl and Vl only and applied to the group where they are floated together. For example, acceptance ratios and shape uncertain-

ties derived on W_{cc} are applied to both W_{bb} and W_{cc} . Generally, an agreement is observed between the various backgrounds ratios. The main goal of this procedure is to ensure that the dominant backgrounds V_{cc} and V_{cl} are described accurately in the most sensitive regions.

This section will only include the 1-lepton channel most relevant backgrounds, $t\bar{t}$ and $W + jets$. Modelling uncertainties considered in this analysis are summarised in Table 6.11.

6.5.2.1 $W+jets$

The systematics uncertainties for the modelling of the $W + jets$ backgrounds are estimated by comparing the nominal SHERPA 2.2.1 $W + jets$ samples to the alternative MADGRAPH 5 ones (including scale variations of the former). The W boson backgrounds are categorised according to the jets flavour, namely in $W + bb$, $W + bc$, $W + bl$, $W + cc$, $W + cl$ and $W + l$.

Tables 6.7 and 6.8 show the fraction of each process with respect to the total $W + jets$ background in the 1 c -tag and 2 c -tag regions respectively. The tables show that in the 1 c -tag region, the dominant processes are events in which the Higgs candidate is reconstructed from either a pair of *light* jets or a *light* jet and a c -jet. In the 2-tag region, the largest contribution to $m_{c\bar{c}}$ is from a pair of c -jets or a *light* jet and a c -jet.

The strategy for the final fit and the resulting choices made to derive the background modelling is to group similar physics processes together and then describe the acceptance ratios and shape uncertainties by using the most dominant background. The main assumption is that various flavour compositions originating from the same physics will be subject to the same modelling issues and should therefore be accounted for together.

The heavy flavour ('hf') jets produced in association with the W boson, namely in $W + bb$ and $W + cc$, are generated through gluon

1 c-tag	mid p_T^V		high p_T^V	
Process	2-jet	3-jet	2-jet	3-jet
$W + cc$	$(6.2 \pm 0.2)\%$	$(6.3 \pm 0.2)\%$	$(7.8 \pm 0.4)\%$	$(7.7 \pm 0.3)\%$
$W + cl$	$(52.1 \pm 0.2)\%$	$(52.6 \pm 0.2)\%$	$(45.4 \pm 0.4)\%$	$(46.9 \pm 0.4)\%$
$W + l$	$(35.6 \pm 0.2)\%$	$(35.1 \pm 0.2)\%$	$(39.6 \pm 0.4)\%$	$(38.3 \pm 0.4)\%$
$W + bb$	$(1.2 \pm 0.2)\%$	$(1.2 \pm 0.2)\%$	$(1.9 \pm 0.4)\%$	$(1.7 \pm 0.3)\%$
$W + bc$	$(1.0 \pm 0.2)\%$	$(1.0 \pm 0.2)\%$	$(0.9 \pm 0.4)\%$	$(1.0 \pm 0.3)\%$
$W + bl$	$(3.8 \pm 0.2)\%$	$(3.8 \pm 0.2)\%$	$(4.4 \pm 0.4)\%$	$(4.3 \pm 0.3)\%$
$W(\tau_h \nu) + l$	$(0.0 \pm 0.2)\%$	$(0.0 \pm 0.2)\%$	$(0.0 \pm 0.4)\%$	$(0.0 \pm 0.3)\%$
$W(\tau_h \nu) + c$	$(0.0 \pm 0.2)\%$	$(0.0 \pm 0.2)\%$	$(0.0 \pm 0.4)\%$	$(0.0 \pm 0.3)\%$
$W(\tau_h \nu) + b$	$(0.0 \pm 0.2)\%$	$(0.0 \pm 0.2)\%$	$(0.0 \pm 0.4)\%$	$(0.0 \pm 0.3)\%$

Table 6.7: Fractions of the individual W +jets components in the 1 c -tag region and their statistical uncertainty for mid p_T^V and high p_T^V . The dominant components are $W + cl$ and $W + l$.

2 c-tag	mid p_T^V		high p_T^V	
Process	2-jet	3-jet	2-jet	3-jet
$W + cc$	$(41.6 \pm 0.5)\%$	$(42.4 \pm 0.5)\%$	$(43.8 \pm 0.9)\%$	$(43.9 \pm 0.8)\%$
$W + cl$	$(36.5 \pm 0.5)\%$	$(36.5 \pm 0.5)\%$	$(32.1 \pm 0.8)\%$	$(33.1 \pm 0.8)\%$
$W + l$	$(13.1 \pm 0.4)\%$	$(12.8 \pm 0.5)\%$	$(14.2 \pm 0.8)\%$	$(13.7 \pm 0.8)\%$
$W + bb$	$(2.8 \pm 0.4)\%$	$(2.6 \pm 0.5)\%$	$(4.3 \pm 0.8)\%$	$(3.7 \pm 0.8)\%$
$W + bc$	$(3.4 \pm 0.4)\%$	$(3.3 \pm 0.5)\%$	$(2.8 \pm 0.8)\%$	$(2.9 \pm 0.8)\%$
$W + bl$	$(2.6 \pm 0.4)\%$	$(2.4 \pm 0.5)\%$	$(2.9 \pm 0.8)\%$	$(2.8 \pm 0.8)\%$
$W(\tau_h \nu) + l$	$(0.0 \pm 0.4)\%$	$(0.0 \pm 0.5)\%$	$(0.0 \pm 0.8)\%$	$(0.0 \pm 0.8)\%$
$W(\tau_h \nu) + c$	$(0.0 \pm 0.4)\%$	$(0.0 \pm 0.5)\%$	$(0.0 \pm 0.8)\%$	$(0.0 \pm 0.8)\%$
$W(\tau_h \nu) + b$	$(0.0 \pm 0.4)\%$	$(0.0 \pm 0.5)\%$	$(0.0 \pm 0.8)\%$	$(0.0 \pm 0.8)\%$

Table 6.8: Fractions of the individual W +jets components in the 2 c -tag region and their statistical uncertainty for mid p_T^V and high p_T^V . The dominant components are $W + cc$ and $W + cl$.

splitting, whose modelling is known to be problematic, and its normalisations are therefore floated together - $W + hf\{bb, cc\}$. In addition, the normalisation of the backgrounds with mixtures of light and heavy flavour jets are floated jointly as $W + mf\{bc, bl, cl\}$, again these jets being produced in the same way. The $W + l$ background has its normalisation constrained with a prior due to its well understood modelling.

The normalised $m_{c\bar{c}}$ shapes of the W +jets background categories in the 1-lepton channel are shown in Figure 6.6 for the combined (1+2) c -tag, (2+3) jet, $p_T^V > 150$ GeV region. The three categories considered in the 1-lepton channel exhibit a different $m_{c\bar{c}}$ shape, though in this

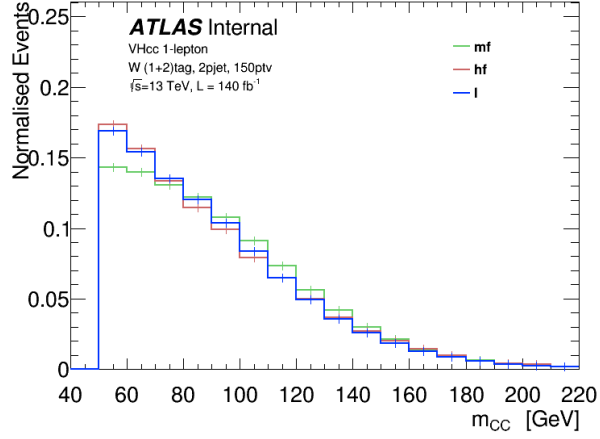


Figure 6.6: Normalised $m_{c\bar{c}}$ distributions of the various W +jets background categories for 1+2 tag, 2+3 $p_T^V > 150$ GeV. $W + hf = \{W + cc, W + bb\}$ is shown in red, $W + mf = \{W + cl, W + bl, W + bc\}$ in green, $W + l$ in blue.

case the distinction between the $W + hf$ and $W + l$ components is less pronounced.

In the final fit, the 1 c-tag and 2 c-tag categories allow to effectively constrain two main processes (that will be floated). The $W+l$ background is known to be well-modelled. Therefore it will be left constrained in the fit, and the Whf and Wmf will be floated. As Whf is mostly composed of Wcc , it will be best constrained by the 2 c-tag categories of the fit. As Wmf is mostly composed of Wcl , it will be constrained by the 1 c-tag. A test will be performed to include 0 c-tag as a control region in the analysis, in which case the Wl component will be floated too.

Normalisation Uncertainties

Normalisation uncertainties are derived as a single ratio between alternative and nominal samples. The normalisation of the $W + l$ background is constrained with a prior due to its well understood modelling. The normalisation uncertainty estimate is derived taking into account the yields for the nominal and alternative samples in the combined 1+2 tag, 2+3 jets, $p_T^V > 150$ GeV region.

MadGraphPythia8 vs Sherpa comparisons

The shape comparisons between two generators are performed by taking the ratio of the normalised shapes of the alternative and nominal generators in the range of the fit. The alternative generator investigated here is MADGRAPHPYTHIA8. The final systematic uncertainty is applied for each process separately before the merging. Its value is parametrised as a fit to the ratio of the alternative over the nominal sample.

A study was performed by splitting all the possible tag, p_T^V and jet multiplicity categories and looking at the various shapes of the background considered. It was decided to combine various categories if the different shapes were in agreement within their statistical uncertainty in order to derive the fitted curve used as an uncertainty. Figure 6.7 show the MADGRAPHPYTHIA8 to SHERPA $W + cc$ and $W + cl$ shape ratios for the 1 c -tag and 2 c -tag regions. In Figure 6.8 the ratios are split by p_T^V regions for $W + l$. Based on these studies, the following combinations were found to agree:

- Common to all process: combine 1 c -tag and 2 c -tag
- $W + cc$: Combine p_T^V and jet multiplicity regions (single shape for 1+2 c -tag, 2+3 jets, $p_T^V > 150$ GeV)
- $W + cl$: Combine p_T^V and jet multiplicity regions (single shape for 1+2 c -tag, 2+3 jets, $p_T^V > 150$ GeV)
- $W + l$: Combine jet multiplicity regions, and split p_T^V regions (two shapes for 1+2 c -tag, 2+3 jets, $150 < p_T^V < 250$ GeV and 1+2 c -tag, 2+3 jets, $p_T^V > 250$ GeV)

6.5.2.2 $t\bar{t}$

The modelling systematic uncertainties for the $t\bar{t}$ background are derived from comparisons between the nominal POWHEG + PYTHIA 8 and the alternative MC samples (POWHEG +HERWIG 7, MADGRAPH

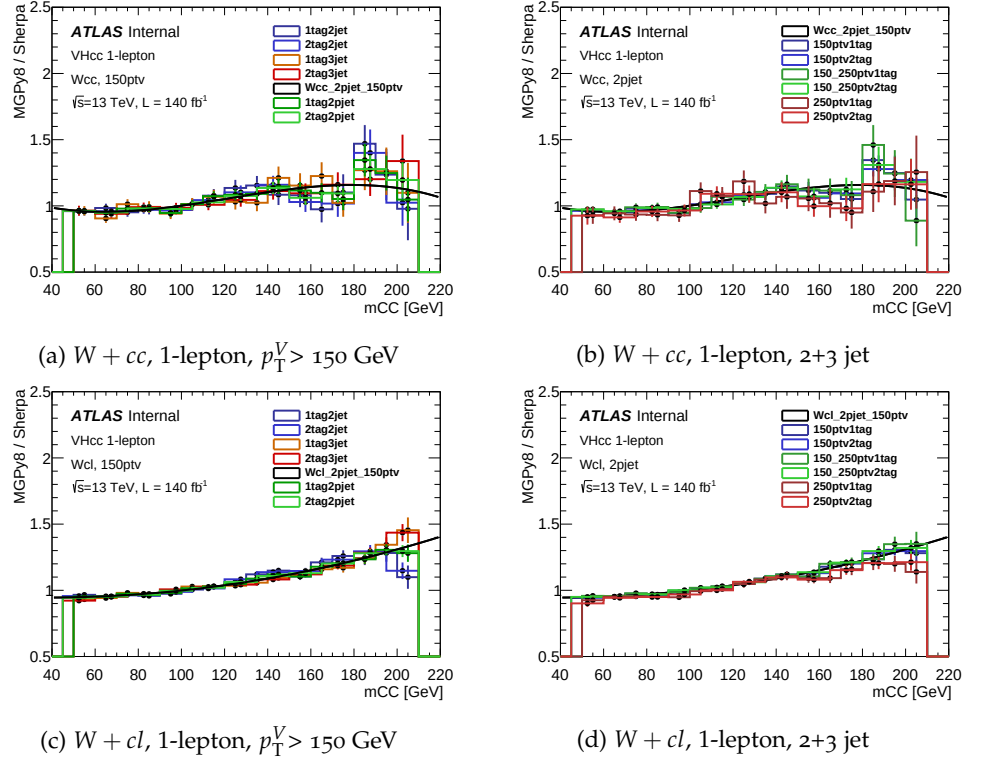


Figure 6.7: Ratio of normalised $m_{c\bar{c}}$ distributions of the alternative MadGraphPythia8 to nominal Sherpa MC generator for the $W + cc$ and $W + cl$ category with (a) and (c) for $p_T^V > 150$ GeV while (b) and (d) for 2+3jet regions. The black curve in each plot represents a fit to the combined region of a 3rd order polynomial.

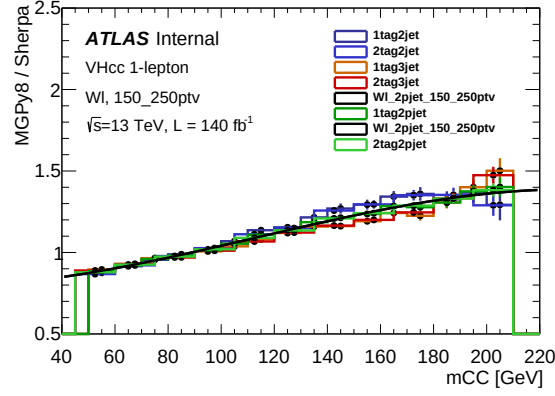
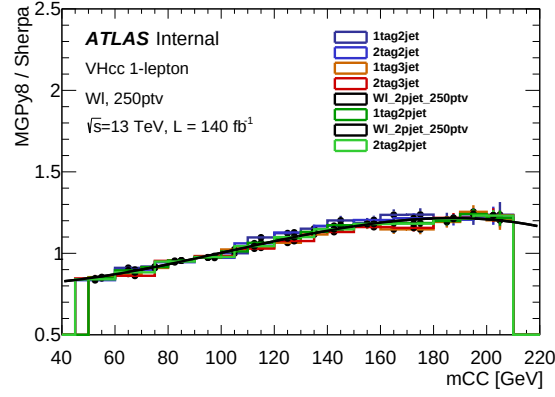
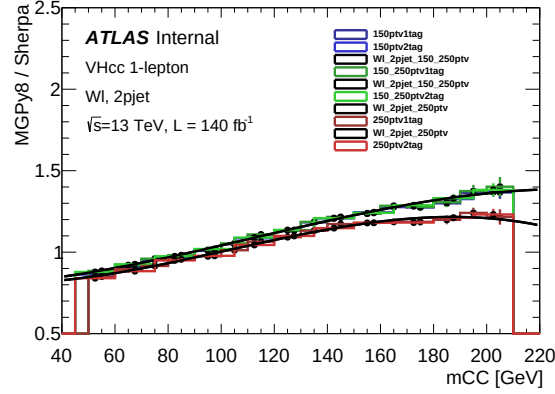
(a) $W + l$, 1-lepton, medium p_T^V (b) $W + l$, 1-lepton, high p_T^V (c) $W + l$, 1-lepton, 2+3 jet

Figure 6.8: Ratio of normalised $m_{c\bar{c}}$ distributions of the alternative Mad-GraphPythia8 to nominal Sherpa MC generator for the $W + l$ category. The black curve in each plot represents a fit to the combined region of a 3rd order polynomial.

5_aMC@NLO+PYTHIA 8.2 and POWHEG+PYTHIA 8 with Initial/Final State Radiation), and by considering the scale variation on the POWHEG + PYTHIA 8 sample.

Normalisation

In the 1-lepton channel, the normalisation of the $t\bar{t}$ is allowed to float freely and is decorrelated between the 2-jet and 3-jet categories. This prior corresponds to the acceptance uncertainty for the analysis phase space, as this is expected to be a dominant source of uncertainty compared to the theory uncertainties.

Flavour composition

The details of the di-jet flavour composition of the $t\bar{t}$ background were studied. Figure 6.9 shows the contributions of the different flavour components to the overall $t\bar{t}$ background for different analysis regions in the 1-lepton channel. The exact composition depends on the region, but the main contributions are the bc/bl components, as well as the cl contribution, which corresponds to the decay of one of the W-bosons in the $t\bar{t}$ decay.

The different flavour components are grouped into three categories based on their shapes and physics processes. Figure 6.10 shows the normalised $m_{c\bar{c}}$ distributions for the different components. Based on this, the following categories are defined:

- **cl+l**: Both jets are produced in the decay of the W-boson. Accordingly, $m_{c\bar{c}}$ is peaked around 80 GeV.
- **bc+bl**: Here, one of the b-jets of the $t\bar{t}$ decay is reconstructed together with a c/l-jet from the W-decay. This component peaks at around 130 GeV.
- **bb+b τ +cc**: The remaining, relatively small components are grouped together.

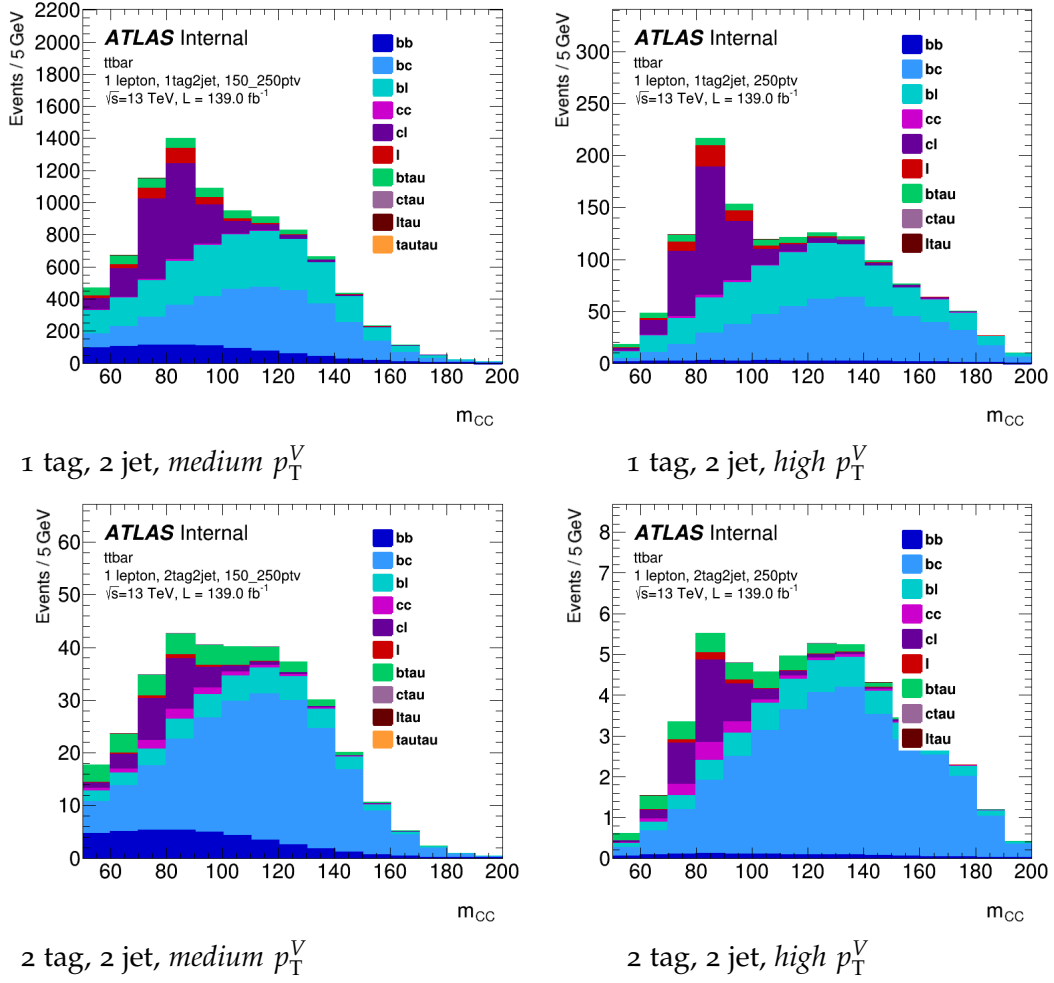


Figure 6.9: Flavour composition of the $t\bar{t}$ $m_{c\bar{c}}$ distributions in the different analysis categories in the 1-lepton channel.

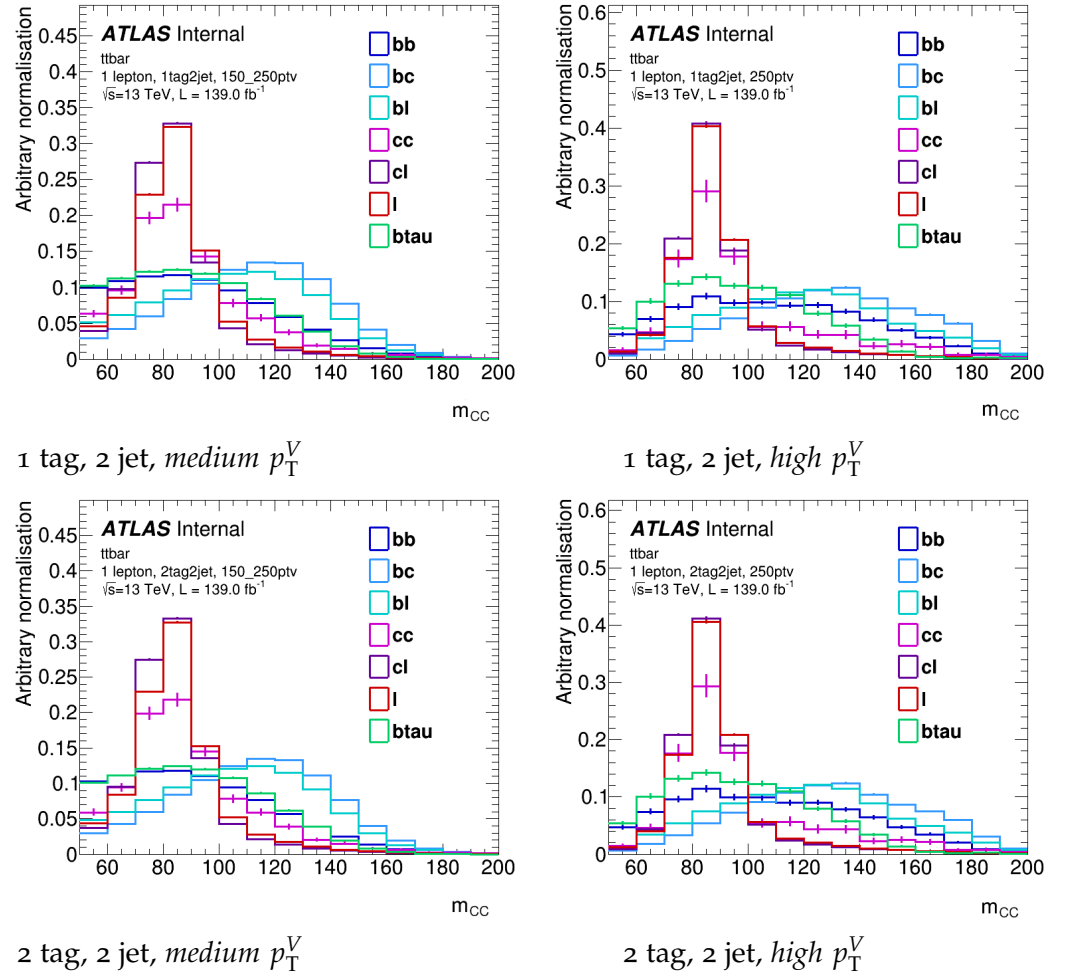


Figure 6.10: Normalised $m_{c\bar{c}}$ distributions of the different flavour components of $t\bar{t}$ in the different analysis categories in the 1-lepton channel.

Uncertainty	2 jets,	3 jets,	2 jets,	3 jets,	Total
$\frac{cl+l}{bl+bc}$	$(4 \pm 2)\%$	$(5 \pm 1)\%$	$(6 \pm 5)\%$	$(4 \pm 3)\%$	$(4 \pm 1)\%$
$\frac{bb+cc+b\tau}{bl+bc}$	$(9 \pm 2)\%$	$(14 \pm 2)\%$	$(21 \pm 12)\%$	$(11 \pm 7)\%$	$(13 \pm 2)\%$
$\frac{bb+cc+b\tau}{cl+l}$	$(12 \pm 2)\%$	$(16 \pm 2)\%$	$(17 \pm 11)\%$	$(8 \pm 7)\%$	$(14 \pm 1)\%$

Table 6.9: Flavour composition uncertainties of $t\bar{t}$.

Based on this categorisation, flavour composition uncertainties are calculated between the different categories. These are calculated on 1+2 c-tags and the uncertainties for the three alternative generators. The resultant uncertainties are shown in Table 6.9, for the different analysis categories and in total. The uncertainty between the two main components, cl/l and bc/bl is fairly small, and the values agree well between the different analysis categories. As a consequence, the inclusive numbers, derived in the 1-lepton channel, are used in the fit, which is calculated by considering all of the analysis categories together. The contributions of the different alternative generators and scale variations to the flavour composition uncertainties are detailed in Table 6.10. As only two acceptance uncertainties are needed between three categories, the ones for the cl/l and $bb/b\tau/cc$ components with respect to the bc/bl component are applied in the fit.

Uncertainty	AMCP8	PP8 RadHi	PH7	PP8 RadLo	μ_R down	μ_R up	Total
$\frac{cl+l}{bl+bc}$	$(-2.4 \pm 0.6)\%$	$(0.8 \pm 0.3)\%$	$(3.0 \pm 0.4)\%$	$(0.2 \pm 0.2)\%$	$(-1.0 \pm 0.3)\%$	$(1.1 \pm 0.2)\%$	$(4 \pm 1)\%$
$\frac{bb+cc+b\tau}{bl+bc}$	$(-7.8 \pm 0.8)\%$	$(-7.2 \pm 0.4)\%$	$(-6.8 \pm 0.5)\%$	$(0.3 \pm 0.3)\%$	$(0.2 \pm 0.3)\%$	$(1.1 \pm 0.2)\%$	$(13 \pm 2)\%$
$\frac{bb+cc+b\tau}{cl+l}$	$(-5.3 \pm 0.9)\%$	$(-8.0 \pm 0.5)\%$	$(-10.0 \pm 0.6)\%$	$(0.1 \pm 0.3)\%$	$(1.2 \pm 0.3)\%$	$(-1.7 \pm 0.3)\%$	$(14 \pm 1)\%$

Table 6.10: Flavour composition uncertainties of $t\bar{t}$ for the different generators.

Shape uncertainties

Uncertainties on the $m_{c\bar{c}}$ shape are derived by comparing the shapes of the nominal POWHEG + PYTHIA 8 sample to the three alternative generators. Two categories are considered. The cl/l flavour component is treated separately, as its $m_{c\bar{c}}$ distribution is peaked around 80 GeV. Thus the corresponding shape uncertainty is likely to impact Diboson

measurements. The remaining components are grouped together, i.e. all of the non-resonant components (bc, bl, bb, b τ , cc), and the second set of shape uncertainties is derived. Given that the bc/bl component is peaked at around 130 GeV, this shape uncertainty is expected to have a larger impact on the VH(cc) measurement.

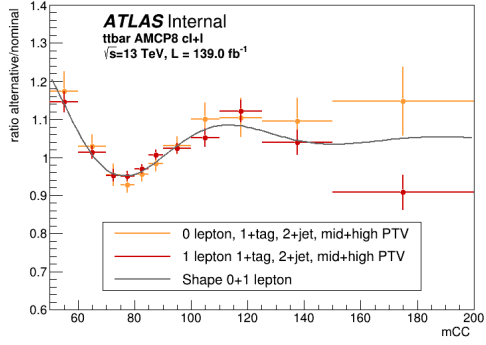
Shape uncertainties are derived from the 1+2 c-tag categories. While the shape of the flavour inclusive $t\bar{t}$ differs between the 1 and 2 c-tag categories, the shapes of the different flavour components are very consistent, as can be seen in Figure 6.10. The shape uncertainties are derived inclusively in the number of jets, and where the ratios of the $m_{c\bar{c}}$ distributions agreed, the *medium* p_T^V and *high* p_T^V categories are combined.

The shape uncertainties for the cl+l flavour component are presented in Figure 6.11. For MADGRAPH 5_aMC@NLO+PYTHIA 8.2 and POWHEG +HERWIG 7, one inclusive shape was used to describe the ratios of the alternative and nominal generator. The shape was parameterised as $f(x) = a \sin(bx + c)e^{-dx} + e$. For the POWHEG +PYTHIA 8 RadHi sample, the shapes were split by p_T^V and parametrised by a second order polynomial.

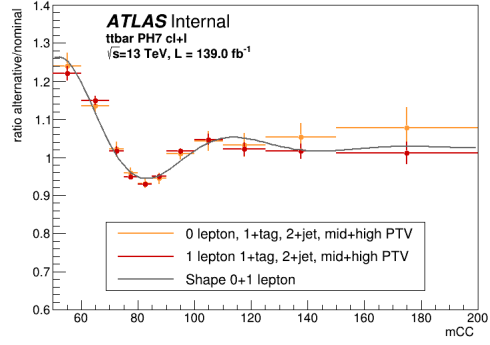
For the non-resonant flavour components (bc, bl, bb, b τ , cc) the derived shapes are presented in Figure 6.12. Here, separate shape uncertainties were derived for the *medium* p_T^V and *high* p_T^V regions. For MADGRAPH 5_aMC@NLO+PYTHIA 8.2 and POWHEG +HERWIG 7 the *medium* p_T^V and *high* p_T^V shapes were parameterised by a fourth and third order polynomial, respectively. For POWHEG +PYTHIA 8 RadHi a second and third order polynomial was sufficient to describe the *medium* p_T^V and *high* p_T^V $m_{c\bar{c}}$ ratios.

The scale variations of the nominal sample were also studied, but the differences in $m_{c\bar{c}}$ shape were minimal. Nevertheless, the $\mu_{R,FSR}$ variations are directly considered in the fit, i.e. as histograms as opposed to parameterised shapes. Due to the otherwise small weights, which decrease the statistical power of the sample, the second largest

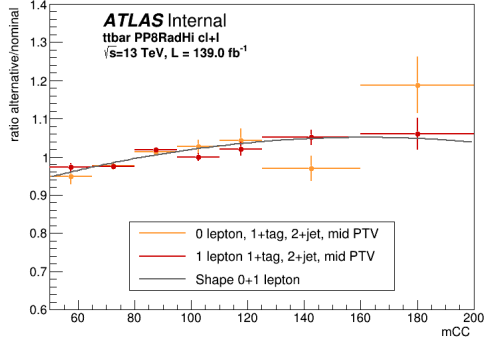
variation is used (i.e. MURp75, which corresponds to $\mu_{R,FSR} = 0.625$ and $\mu_{R,FSR} = 1.75$).



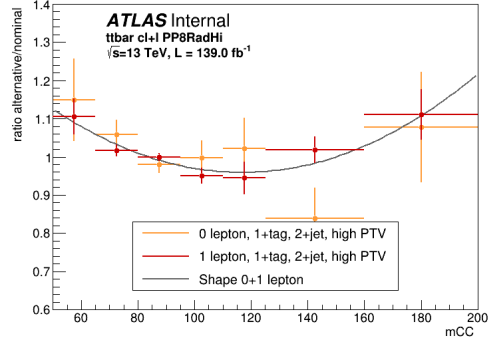
(a) aMC@NLO+PYTHIA 8.2, inclusive in p_T^V



(b) POWHEG +HERWIG 7, inclusive in p_T^V



(c) POWHEG +PYTHIA 8 RadHi, medium p_T^V



(d) POWHEG +PYTHIA 8 RadHi, high p_T^V

Figure 6.11: Ratio of $m_{c\bar{c}}$ distributions for $t\bar{t}$ (cl/l) of alternative to nominal generators with the derived shapes.

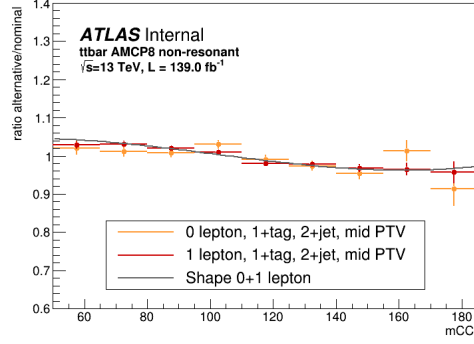
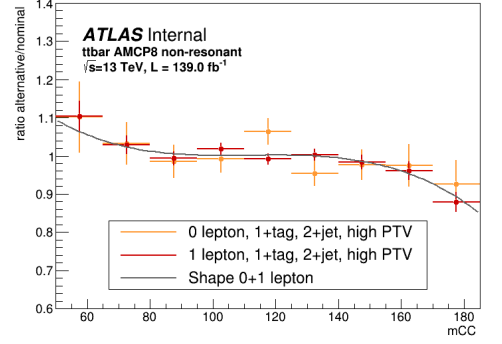
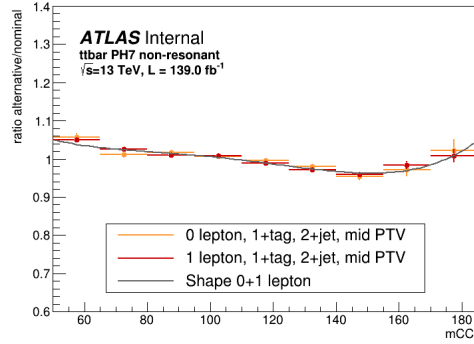
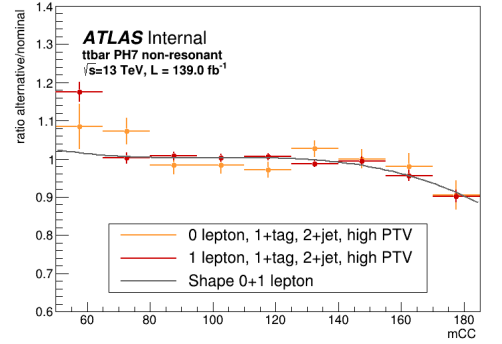
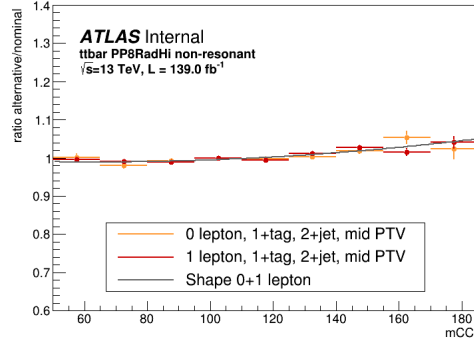
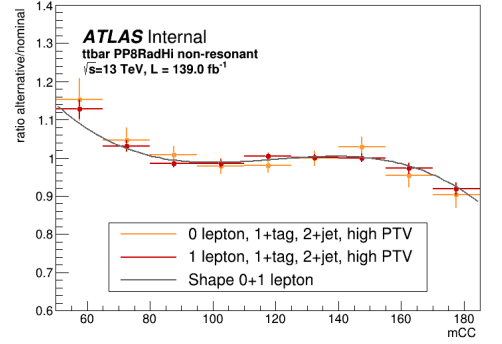
(a) aMC@NLO+PYTHIA 8.2, medium p_T^V (b) aMC@NLO+PYTHIA 8.2, high p_T^V (c) POWHEG +HERWIG 7, medium p_T^V (d) POWHEG +HERWIG 7, high p_T^V (e) POWHEG +PYTHIA 8 RadHi, medium p_T^V (f) POWHEG +PYTHIA 8 RadHi, high p_T^V

Figure 6.12: Ratio of $m_{c\bar{c}}$ distributions for $t\bar{t}$ (bc/bl/bb/b τ /cc) of alternative to nominal generators with the derived shapes.

6.6 VH(CC) FIT RESULTS

In order to measure the signal yield, a binned likelihood fit is performed in the invariant mass distribution of the Higgs candidate $m_{c\bar{c}}$. This fit is done simultaneously to all analysis regions and all three lepton sub-channels. The technical implementation is done within the WorkspaceMaker framework building upon HistFactory [166] together with RooFit [167] and RooStats [168].

The di-jet invariant mass $m_{c\bar{c}}$ is used as the final discriminant to separate the signal from the background. The distribution is fitted in the range 50 GeV to 210 GeV. The lower limit of the fit range is motivated by the larger background modelling uncertainties observed at $m_{c\bar{c}} < 50$ GeV. This low mass region also corresponds to a lower $\Delta R(cc)$ in the case of V+jets and is subject to a higher disagreement between the truth and direct tagging.

Figure 6.13 show the post-fit $m_{c\bar{c}}$ distributions after the combined conditional $\mu = 1$ fit to data. The signal mass window between 70 GeV and 140 GeV in the signal regions is still blinded. Due to an analysis being a combined effort of the collaboration, we cannot unblind without the collaboration's approval as described in Chapter 1. Overall, a good agreement between measured data and simulation can be observed in the signal region sidebands.

A fit is performed with the SM Higgs in mind, thus $\mu = 1$.

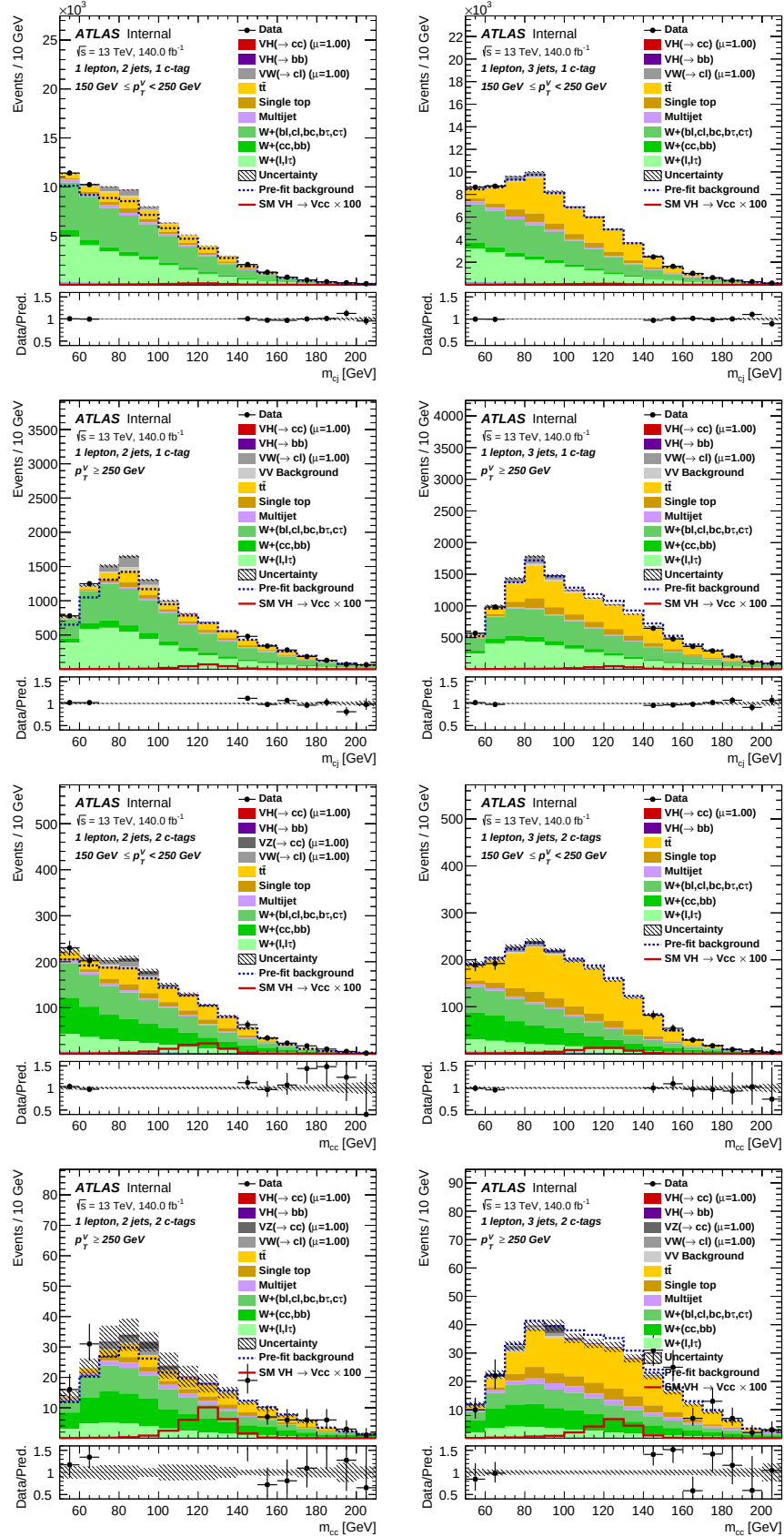


Figure 6.13: Post-fit distributions for the 1-lep channel in the 1- and 2-tag region, obtained in the combined conditional $\mu = 1$ fit to data.

$VH(\rightarrow b\bar{b})$	
$WH(\rightarrow b\bar{b})$ normalisation	27%
$ZH(\rightarrow b\bar{b})$ normalisation	25%
Diboson	
$WW/ZZ/WZ$ acceptance	10/5/12%
p_T^V acceptance	4%
N_{jet} acceptance	7 – 11%
Z+jets	
$Z+hf$ normalisation	Floating
$Z+mf$ normalisation	Floating
$Z+lf$ normalisation	Floating
$Z + bb$ to $Z + cc$ ratio	20%
$Z + bl$ to $Z + cl$ ratio	18%
$Z + bc$ to $Z + cl$ ratio	6%
p_T^V acceptance	1 – 8%
N_{jet} acceptance	10 – 37%
High ΔR CR to SR	12 – 37%
o- to 2-lepton ratio	4 – 5%
W+jets	
$W+hf$ normalisation	Floating
$W+mf$ normalisation	Floating
$W+l$ normalisation	13%
$W + bb$ to $W + cc$ ratio	4 – 10 %
$W + bl$ to $W + cl$ ratio	31 – 32 %
$W + bc$ to $W + cl$ ratio	31 – 33 %
$W \rightarrow \tau\nu(+c)$ to $W + cl$ ratio	11%
$W \rightarrow \tau\nu(+b)$ to $W + cl$ ratio	27%
$W \rightarrow \tau\nu(+l)$ to $W + l$ ratio	8%
N_{jet} acceptance	8 – 14%
High ΔR CR to SR	15 – 29%
$W \rightarrow \tau\nu$ SR to high ΔR CR ratio	5 – 18%
o- to 1-lepton ratio	1 – 6 %
Top quark (o- and 1-lepton)	
top(b) normalisation	Floating
top(other) normalisation	Floating
N_{jet} acceptance	7 – 9%
o- to 1-lepton ratio	4%
SR/top CR acceptance ($t\bar{t}$)	9%
SR/top CR acceptance (Wt)	16%
Wt / $t\bar{t}$ ratio	10%
Top quark (2-lepton)	
Normalisation	Floating
Multi-jet (1-lepton)	
Normalisation	20 – 100%

Table 6.11: Summary of the background modelling systematic uncertainties considered. The values given refer to the size of the uncertainty affecting the acceptance of each background.

Based on the fit results, the expected and observed 95% CL upper limits on the VH(cc) signal strength are given in Table 6.12, for the individual channels and the combined 0+1+2 lepton 3 POI fit. The pre-fit and post-fit expected limits show some differences in the individual channel fits, in particular in the 2-lepton channel. This is expected, as the $Z + jets$ background in the 2-lepton channel is underestimated pre-fit. The pre-fit and post-fit expected limits agree well in the combined fit.

	95% CL CL_s upper limit on the signal strength	
	Expected Limit (Pre-fit)	Expected Limit (Post-fit)
0-lepton	$38.2^{+15.0}_{-10.7}$	$41.7^{+16.3}_{-11.6}$
1-lepton	$46.7^{+18.3}_{-13.0}$	$46.8^{+18.3}_{-13.0}$
2-lepton	$51.0^{+19.9}_{-14.2}$	$58.8^{+23.0}_{-16.4}$
0+1+2-lepton	$25.1^{+9.8}_{-7.0}$	$26.0^{+10.1}_{-7.3}$

Table 6.12: The expected 95% CL CL_s upper limit on the cross-section and signal strength for the 3 analysis channels and in the combined fit.

The effect of systematic uncertainties and free-floating background normalisation parameters obtained from the likelihood fit to data is summarised in Tables 6.13 and 6.14. Being the main background (as can be seen in the post-fit Figure 6.13), $V + jets$ is the dominant source of uncertainty throughout the channels with 20-30%. The statistical uncertainty is obtained from a fit in which all nuisance parameters are fixed to their post-fit values. The total systematic uncertainty is found by subtracting the statistical uncertainty from the total uncertainty in quadrature. The impact of a set of systematic uncertainties is assessed by performing the fit with only their nuisance parameters fixed to their post-fit values.

Background	p_T^V	Jets	Value
top(b)			0.91 ± 0.06
top(other)			0.94 ± 0.08
$t\bar{t}$ (2 lep)	$p_T^V > 150$ GeV	2	0.76 ± 0.22
		3	0.96 ± 0.13
	$75 < p_T^V < 150$ GeV	2	1.08 ± 0.08
		3	1.06 ± 0.07
$W+hf$			1.16 ± 0.35
$W+mf$			1.28 ± 0.14
$Z+hf$	$p_T^V > 150$ GeV		1.19 ± 0.22
	$75 < p_T^V < 150$ GeV		1.25 ± 0.25
$Z+mf$	$p_T^V > 150$ GeV		1.10 ± 0.15
	$75 < p_T^V < 150$ GeV		1.11 ± 0.15
$Z+lf$	$p_T^V > 150$ GeV	2	1.07 ± 0.03
		3	1.08 ± 0.05
	$75 < p_T^V < 150$ GeV	2	1.12 ± 0.04
		3	1.07 ± 0.06

Table 6.13: Values of the free-floating background normalisation parameters obtained from the likelihood fit to data. The uncertainties represent the combined statistical and systematic uncertainties. Unless otherwise stated, normalisation parameters are correlated across all p_T^V and number-of-jets analysis regions.

Source of uncertainty		$\mu_{VH(c\bar{c})}$
Total		15.3
Statistical		10.0
Systematics		11.5
Statistical uncertainties		
Data statistics only		7.8
Floating normalisations		5.1
Theoretical and modelling uncertainties		
$VH(\rightarrow c\bar{c})$		2.1
$Z+\text{jets}$		7.0
Top-quark		3.9
$W+\text{jets}$		3.0
Diboson		1.0
$VH(\rightarrow b\bar{b})$		0.8
Multi-Jet		1.0
Simulation statistics		4.2
Experimental uncertainties		
Jets		2.8
Leptons		0.5
$E_{\text{T}}^{\text{miss}}$		0.2
Pile-up and luminosity		0.3
Flavour tagging	$c\text{-jets}$	1.6
	$b\text{-jets}$	1.1
	light-jets	0.4
	$\tau\text{-jets}$	0.3
Truth-flavour tagging	ΔR correction	3.3
	Residual non-closure	1.7

Table 6.14: Breakdown of contributions to the uncertainty in the fitted values of $\mu_{VH(c\bar{c})}$, $\mu_{VW(cq)}$ and $\mu_{VZ(c\bar{c})}$. The sum in quadrature of different sources of uncertainty may differ from the total due to correlations. In the case that the up and down systematic variations have different values, the mean of the absolute values is shown.

6.7 LIMITS ON THE c -QUARK YUKAWA COUPLING

The expected sensitivity of the analysis is considered within the context of the κ framework [169, 170] for a generic modified Yukawa couplings scenario. The coupling modifiers κ_i (where i represents any particles coupling to the Higgs boson directly) are introduced. In the context of the VH(cc) analysis, the κ_V and κ_c coupling modifiers appear explicitly as represented on Figure 6.14.

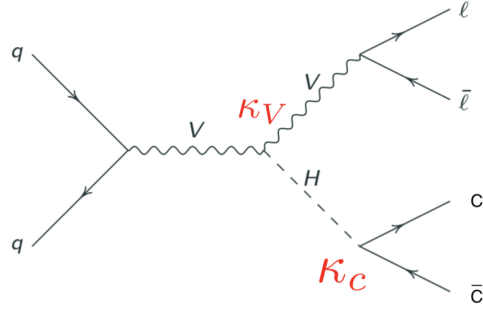


Figure 6.14: Feynman diagram of the VH(cc) process for the 2-lepton channel, red represent the two vertices where the κ_V and κ_c coupling modifiers appear explicitly.

The signal strength for the VH(cc) process can then be parametrised as

$$\mu(\kappa_i) = \frac{B(\kappa_i)\sigma_{pp \rightarrow VH}(\kappa_i)}{B^{SM}\sigma_{pp \rightarrow VH}^{SM}} = \frac{\kappa_V^2 \kappa_c^2}{\kappa_H^2}, \quad (6.3)$$

where $B = B(H \rightarrow c\bar{c})$ is the branching fraction of the Higgs boson decay to a pair of charm quarks, $\sigma_{pp \rightarrow VH}$ is the cross-section of the VH production mode in pp collisions and $\kappa_H^2 = \frac{\sum_j \Gamma_j^{SM} \times \kappa_j^2}{\Gamma_H^{SM}}$ with Γ^{SM} the SM width for various processes of the Higgs decay.

For the case that all $\kappa_i = 1, \forall i \neq c$, the signal strength parametrisation in Equation (6.3) becomes

$$\mu(\kappa_i) = \frac{\kappa_c^2}{B(H \rightarrow c\bar{c})\kappa_c^2 + (1 - B(H \rightarrow c\bar{c}))}. \quad (6.4)$$

Equation (6.4) shows that at a small value of the signal strength, the constraint on κ_c scales as $\mu(\kappa_c) \approx \kappa_c^2$ while at very high values of the signal strength the constraint on κ_c reaches an asymptotic limit defined as $\mu(\kappa_c) = \mu = \frac{1}{B(H \rightarrow c\bar{c})}$. Therefore, any expected signal strength above this threshold will have no constraining power over κ_c . This effect is illustrated on Figure 6.15.

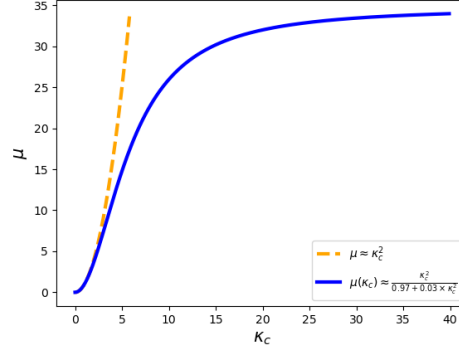


Figure 6.15: Parameterisation of the signal strength μ as a function of the charm quark Yukawa coupling modifier κ_c in the context of a scenario where all other coupling modifiers are considered unity, see Equation (6.4).

In order to extract a limit on the value of κ_c , the $VH(cc)$ signal strength is reparameterised directly in the fit according to Equation (6.4), and a profile likelihood scan is performed. This can be done in a fit configuration with 1 POI (κ_c), or with 3 POI (κ_c , $\mu_{VZ(cc)}$ and $\mu_{VW(cl)}$). The results for both cases are shown in Figure 6.16 for the individual channels and the combined fit to the Asimov dataset Section 4.3. The individual channel fits are not able to set 95% CL limits on κ_c . However, the combined fit is. In the case of the 3 POI fit, floating the Diboson signal strengths loosens the limit by around 12%, leading to a 95% CL limit of 9.1 and 10.2 on $|\kappa_c|$ for the 1 POI and 3 POI configuration, respectively.

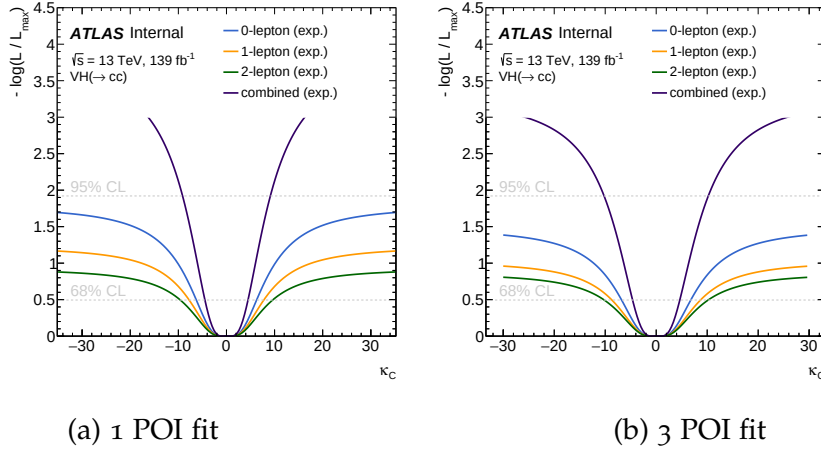


Figure 6.16: Profile likelihood scan for κ_c for the 0-lepton, 1-lepton, 2-lepton and combined fit to the Asimov dataset for the 1 POI and 3 POI fit setup.

6.8 CONCLUSION

Constraining the Higgs to charm coupling via direct measurement of the $VH(cc)$ decay is a challenging measurement due to a low predicted signal yield in the SM and the large irreducible backgrounds. Such a direct search for the decay of the Standard Model Higgs boson to a charm quark pair was performed using a pp collision dataset, collected at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 139 fb^{-1} recorded by the ATLAS experiment at the LHC. The production of a Higgs boson in association with a W or Z boson was considered. The decay products of the Higgs boson were identified using charm jet tagging algorithms.

At the time of writing, the signal region is still blinded. Nevertheless, the analysis still provides a glimpse at the upcoming result based solely on the expected signal. The current result is an expected upper limit of 26 times the SM prediction for the $VH(cc)$ production rate. The analysis was also validated with a measurement of VZ and VW production where the sample of hadronic W/Z decays is enriched in $Z \rightarrow c\bar{c}$ and $W \rightarrow cs, cd$. The expected signal significance is 2.6 for the $VZ(cc)$ process and 5.4 for $VWcl$.

τ TRIGGER USING DEEP LEARNING

7.1 INTRODUCTION

Tau leptons play an essential role in the physics program at the LHC. They are used in the studies for the Higgs boson (strongest signal for fermions) as well as in electroweak measurements and are one of the key ingredients to search for processes predicted by BSM theories. Identifying tau leptons with good performance is an essential part of these analyses.

With a mass of 1.7 GeV the tau is the heaviest lepton, and due to its short lifetime of $2.9 \times 10^{-13}\text{ s}$ and a decay length of $c\tau = 87\mu\text{m}$, it decays inside the beam pipe and never reaches the outermost parts of the detector. The tau is the only lepton that decays into leptons and hadrons, decaying leptonically in 35% of the cases and hadronically in the remaining 65%. The hadronic decay products are typically one ("1-prong" decay) or three ("3-prong" decay) charged pions and zero or one neutral pion. The neutral and charged hadrons stemming from the tau lepton decay make up the visible decay products of the tau lepton in the detector. The kinematics of quark or gluon jets (QCD) is similar to that of hadronically decaying tau leptons, leading to a potentially high probability of misidentifying them as taus. Therefore, a well performing tau lepton identification is crucial for the physics program.

This chapter describes novel algorithms to identify the presence of a tau lepton in a collision event at the ATLAS experiment. The algorithms are based on Deep Neural Networks (DNN) acting only on raw calorimeter data. The expected performance is evaluated in simu-

lated proton–proton collisions (MC) at $\sqrt{s} = 13$ TeV and compared to current algorithms.

This chapter is structured as follows. Section 7.2 describes the current status of tau triggering in ATLAS, both hardware and algorithms. Section 7.3 lays out the basics of neural networks. Section 7.4 portrays the proposed new algorithms for triggering. Section 7.5 and 7.6 conclude with results.

7.2 L1 CALORIMETER TRIGGER

The ATLAS Trigger and Data Acquisition (TDAQ) system is detailed briefly in Section 3.3.6 and in [171]. In preparation for Run 3, when the LHC will run at higher energy and instantaneous luminosity, the L1 calorimeter trigger (L1Calo) (which is in the focus of this work) [172] is currently implementing a significant programme of planned upgrades. The existing hardware will be replaced by a new system of feature extractor modules, which will process finer-granularity information from the calorimeters and execute more sophisticated algorithms to identify physics objects. These upgrades will permit better performance in a challenging high-luminosity and high-pileup environment.

ATLAS is planning an upgrade to the Trigger and Data Acquisition (TDAQ) system for the Phase-I period (Run 3, from 2021 to 2024) to address the new challenging accelerator machine conditions, namely an increase in luminosity to $2.5 \times 10^{34} / \text{cm}^2\text{s}$ and to maintain the sensitivity to electroweak physics without being affected by the increased number of pileup events.

The L1Calo trigger system for Run-2 is too coarse, and the reconstruction of Trigger Objects (TOB), which are electrons, photons, tau leptons, jets, and E_T^{miss} , is challenging at high instantaneous luminosities. High pile-up levels degrade the calorimeter resolution and isolation requirements for particles, resulting in worsening trigger

performance. This degradation can be alleviated by transferring off-line algorithms to the trigger level. The L1Calo trigger consists of electronics hardware systems with embedded logic based on Application Specific Integrated Circuit (ASICs) and Field-Programmable Gate Array (FPGAs) that discriminates and counts candidate particles above the threshold and pass them to the higher trigger levels. For an algorithm to run and trigger efficiently, it must be implemented over FPGA and not exceed strict computation time.

To separate the desired rare processes from the predominant QCD jet production and other backgrounds, L1 searches mainly for exclusive signatures that could identify isolated high transverse-energy electrons, photons, muons, and tau leptons. The 25 ns interval between successive bunch-crossings is far too short for processing and selecting events. In fact, given the size of the detectors and with much of the L1 electronics off the detector in a separate cavern, even the transmission delays for the signals are much longer than that. Therefore, the scheme adopted is for the detector data to be held in buffers while L1 decides. If the bunch-crossing passes the L1 criteria, a signal is sent, and the data is kept; if not, deleted. The time allowed for the L1 stage depends on the size of the on-detector buffers, and that requires a compromise between a long enough processing time to allow effective trigger algorithms. In ATLAS, the allowed decision time, or latency, for L1 was chosen to be 2.5 μs . Since it is unacceptable to exceed this, L1 was designed to have a nominal latency of about 2.0 μs to ensure an adequate safety margin. It is important to note that a large fraction of this time is consumed by the signal transmission delays from and back to the detector front-end electronics.

L1Calo uses calorimeter energy deposits to identify various types of high- E_T objects as well as energy sums of interest. Its input data comes from trigger towers of reduced granularity, mostly 0.1×0.1 in $\Delta\eta \times \Delta\phi$ from all the ATLAS electromagnetic and hadronic calorimeters. It comprises five layers denoted as EM0, EM1, EM2, EM3 and HAD,

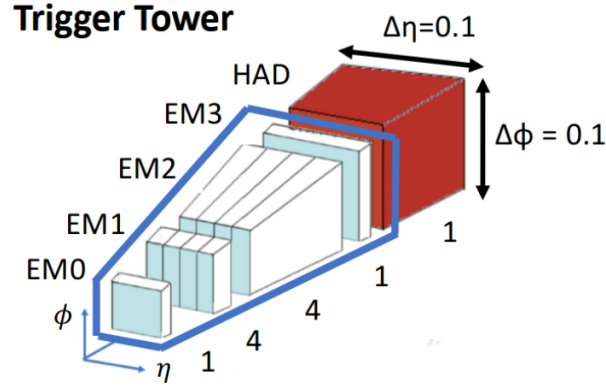


Figure 7.1: L1Calo trigger tower is composed of electromagnetic and hadronic calorimeters.

where the first four are electromagnetic, and the last is hadronic, as the name suggests. The tower layers have different granularity:

- coarse layers (3×3): EM0, EM3 and HAD
- fine layers (12×3): EM1 and EM2

This leads to the definition of the TOB, an object containing 99 energy deposits values throughout the calorimeter trigger tower layers. An L1Calo trigger algorithm is then defined as a boolean function acting on such TOB.

7.3 MACHINE LEARNING AND NEURAL NETWORKS

Neural Networks are a key piece of some of the most successful machine learning algorithms. A neural network is a manufactured attempt to reproduce and imitate the functionalities of the human brain, where neurons are interconnected among themselves. Conceived initially as the *Perceptron* [173], it has only gained fame in recent years due to advances in computer hardware, enabling more and more complex calculations. A machine learning algorithm in its core is no more than a mapping, $f : \vec{X} \rightarrow \vec{Y}$ where $\vec{X}, \vec{Y} \in \mathbb{R}^n$. The *learning* part is finding such mapping that can accurately make a future prediction. A neural network has input and output neurons, which are connected by weighted synapses. Although a network can have an arbitrary

*Originated from
perceptual
recognition*

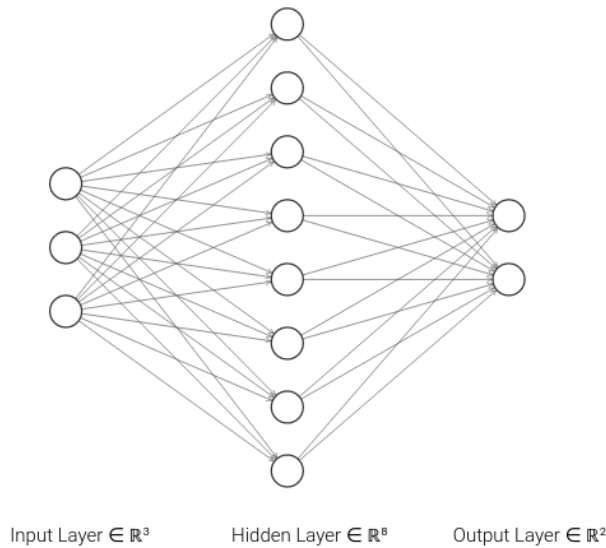


Figure 7.2: A simple, fully connected neural network with input nodes, single hidden layer and output layer.

architecture, for most applications, a simpler structure like the one shown in Figure Figure 7.2 will suffice.

The trick is to design a network architecture that can achieve high accuracy using relatively little computational power, with minimal data. One hidden layer is enough to represent an approximation of any function to an arbitrary degree of accuracy. This is known as the Universal Approximation Theorem [174]. The question then arises, why do people use deep neural networks if one layer is enough? The answer is that this network would need to have a neural architecture that is very wide since shallow networks require (exponentially) more width than a deep network. Furthermore, shallow networks have a higher affinity for overfitting, i.e. not generalizing enough for unseen data.

7.3.1 Data Propagation

As mentioned above, a neural network is merely a function, and a function requires input in order to provide the desired output. The input layer is made up of neurons corresponding to the algorithm's

specific data. It is a tensor, with a rank depending on the problem and architecture of choice. A rank-1 tensor could be a simple vector containing a particle's characteristics such as p_T, η, ϕ while a rank-2 tensor could be a grey-scale image and a rank-3 tensor might be a colour image with dimensions of $Width \times Height \times Colors$. This layer takes in data and propagates it to the next layer.

The following layers are called hidden layers, as they are hidden within the network. The rule here is simple, the input tensors from the previous layer are multiplied by a weight tensor and summed. These weight tensors are the "learnable" parameters of the model and are constantly changing with new data. Next, every multiplication product is passed through an activation function which decides whether the result is significant enough to 'fire' the node, i.e. passing it to the next layer.

The process comes to an end at the output layer. The algorithm's output should be understandable and correspond to decision making. When dealing with classification problems, an output layer is usually a rank-1 tensor, providing the relative probabilities of the input type.

This step is named *Forward-Propagation* as the data propagates forward through the network, making the network provide a prediction to a given input.

The "learning" part, or training, is changing the network weight tensors to respond to the level of deviation from the ground truth. In constructing a network, one is aspiring to an accurate as possible model, generally meaning a lowest as possible error. Updating the weights is made through a process of gradient descent [175], where the magnitude of change is dependent on the gradient at each step. This makes sure the update is in the right "direction" towards a minimal error. This is called *Back-Propagation*. The reason is that the gradient is being calculated through the chain rule, where every layer depends on the one after it. The data propagates backwards, enabling the correct update of the weights.

In classification problems, the network has to classify a given input to a specific label from a set of known categories.

One has to take into account the possibilities of local minima or an elaborate saddle point in the loss space.

7.3.2 *Deep Learning*

For a machine learning model to work properly (i.e. make the correct prediction), it must be supplied with proper data. A BDT acting on classifying a charged Higgs event (see Section 5.7) will need the event's information in order to make a prediction. This data is made of different variables named features. Feature engineering refers to a process of human-made features from existing ones, i.e. invariant mass of a particle from its decay products. By knowing the physics, researchers can come up with different higher level variables from basic ones.

Deep Learning is changing that. With deep learning, one can start with raw data, as features will be automatically created by the neural network in the learning process. The feature engineering approach was the dominant one until recently, when deep learning techniques demonstrated recognition performance better than the carefully crafted feature detectors. Deep learning shifts the burden of feature design also to the underlying learning system along with classification learning typical of earlier multiple layer neural network learning. From this perspective, a deep learning system is a fully trainable system beginning from raw input to the final output of recognised objects.

One can think about a Deep Learning model as made out of two adjacent networks. The first, a feature extractor, its sole purpose is to understand the underlying relations between the raw data, producing meaningful features with a clear distinction between classes (i.e. signal vs background). This information is then being "fed" into the second network to make a prediction similar to a more classical neural network.

Where "normal" neural networks might have up to three layers of neurons, a DNN can have 40 or 500 layers of neurons. The added depth was made possible only thanks to advances in Graphical Computer

Units (GPUs) which made tensor operations more efficient, enabling a significant reduction in training time.

7.4 ARCHITECTURES FOR TRIGGERING

As ATLAS approaches Run-3, different groups are devising new algorithms for τ triggering in L1Calo. Every proposition must meet strict demands for rate and latency, making sure not to exceed the Run 2 algorithm in any way. In the following, one of these algorithms, an adaptive clustering titled "*bigCluster*" (BC), is selected to act as a benchmark due to its promising results in the ATLAS trigger group.

7.4.1 Monte Carlo Samples

MC simulations are used to train the networks and evaluate their performance. The signal sample consists of $Z \rightarrow \tau\bar{\tau}$ generated with POWHEG+PYTHIA8 while the background sample consists of QCD dijet events produced with SHERPA. The pp collisions were simulated at a centre-of-mass energy of $\sqrt{s} = 13\text{TeV}$, with a 25 ns bunch spacing. Generated events were processed through the ATLAS detector simulation [176] based on Geant4 [90].

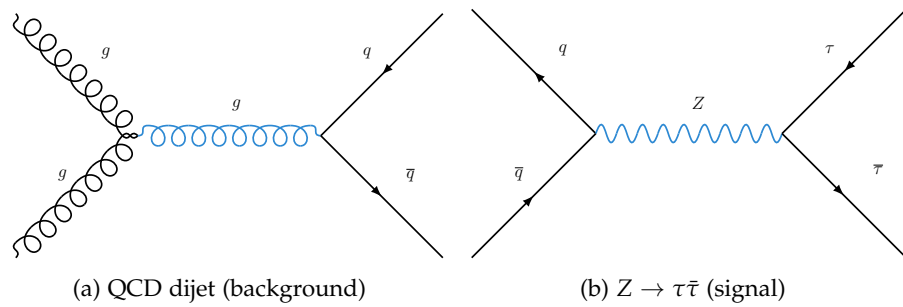


Figure 7.3: Feynman diagrams for the signal and the major background

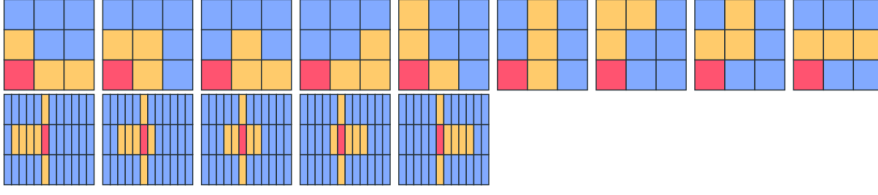


Figure 7.4: Adaptive cluster shape scanning over the different tower layers.

7.4.2 Benchmark Algorithms

As mentioned above, all decision making is based on a TOB, an object containing 99 energy deposits within a certain trigger tower. All current algorithms use this information (raw data) in order to decide whether to trigger or not. The basic principle is the same for all: try to sum up the cells in a specific manner and make a decision based on the p_T of this summation.

By studying MC samples of signal and background processes, teams develop elaborate shapes that mostly match tau lepton decay products and not a background event. Thus, by using a predefined recipe, one can, for example, search for the hottest cell in EM1+EM2, then cluster and adding from adjacent layers around it to provide a single value for threshold triggering. The motivation comes from understanding the event topology. As the tau decays to pions, the neutral π^0 will immediately decay into photons, $\pi^0 \rightarrow \gamma\gamma$, leaving a strong signature in the electromagnetic calorimeter. In contrast, the charged pions, $\pi^\pm$ will leave their mark in the hadronic calorimeter. Background events, such as a di-jet, should have a stronger hadronic signature.

As previously described, any new way of triggering cannot exceed the rate of the current algorithm, titled simply Run 2. Figure 7.5 shows the fake rate response for both Run 2 and BC. This is essentially the false positives the two algorithms output. Even though BC triggers fewer background events (fakes) in total, the decision is very p_T depended as can be seen. It is a huge goal for physicists to push the p_T threshold lower in order to enable low- p_T τ s to pass the trigger.

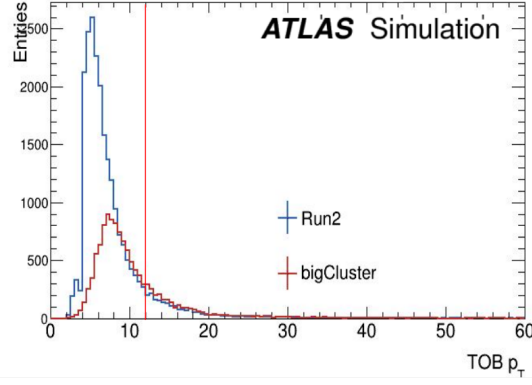


Figure 7.5: Fake rate estimation (QCD di-jet background sample) for the current Run 2 algorithm with the new proposed BC. Setting the threshold at $p_T = 12$ GeV yields an equal rate result.

By understanding the constraints of the problem at hand, the proper metric can be chosen. It is common to use and state an algorithm's accuracy, but this metric becomes misleading when one has to pay attention and limit the fake rate (i.e. false positives). Recall, on the other hand, defined as $\frac{TP}{TP+FP}$ will serve as a better metric, going to 1 as fake rate (FP) drops and reaching 0 when the fake rate is higher.

7.4.3 Convolutional Neural Networks

Convolutional Neural Networks (CNN) played a major role in making Deep Learning so popular by bridging the gap between the capabilities of humans and machines. CNN are known for their abilities to learn and perform over visual data such as images and video, making predictions once seemed impossible for a computer. Examples and illustrations will follow visual samples as they can be easily understood.

There are several drawbacks of traditional neural networks, especially when it comes to image processing. NNs use one neuron for each input (e.g. pixel in an image, multiplied by three in RGB case). The amount of weights rapidly becomes unmanageable for large images. For a 224×224 pixel image with all three colour channels, there are

around 150,000 weights that must be trained. As a result, difficulties arise whilst training, and overfitting can occur.

Another common problem is that NN reacts differently to an input (image) and its shifted version, meaning they are not translation invariant. If a picture of a dog appears in the top left of the image in one picture and the bottom right of another picture, it will try to correct itself and assume that a dog will always appear in this section of the image.

Thus, there is a need to leverage the spatial correlation of the image features (pixels) in such a way that we can see the object in the picture no matter where it may appear.

CNN, being a Deep Learning model (see Section 7.3.2), is constructed out of a feature extractor and a classifier. The ingenuity lies in the way this network is able to extract meaningful information.

When analysing images, the influence of nearby pixels is taken into account by using a filter. A filter is simply a small matrix of a size specified by the user (a rule of thumb is 3×3 or 5×5). This matrix is then convolved across the image from top left to bottom right. For each point on the image, a value is calculated based on the filter using a convolution operation producing what is called - a feature map.

The main objective of a convolution is to extract features such as edges, colours and corners from the input. As one goes deeper inside the network, the network starts identifying more complex features such as shapes, digits and face parts as well. If the image is RGB, then the filter will have a smaller height and width compared to the image, but it will have the same depth as the image - rank 3 tensor ($Height \times Width \times 3$).

The number of filters is a hyper-parameter to the network, affecting its ability to extract complex shapes and its training time. These filters are the weights of the model, being updated during the training process.

After passing several layers of convolution and activation functions, the output is then flattened into a rank-1 tensor, thus, summarizing the

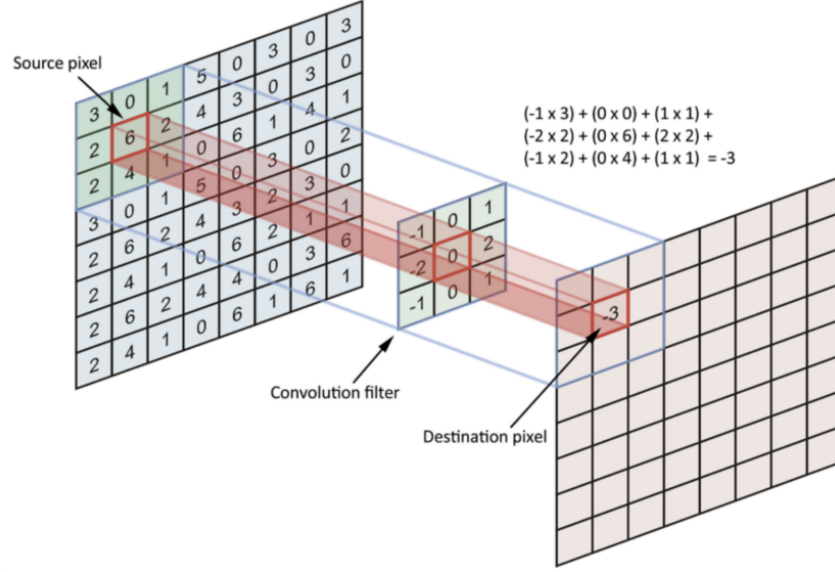


Figure 7.6: The convolution operation. Source pixels originate from the input image.

network's extracted features. Afterwards, it passes to a fully connected neural network in order to output a final probability.

7.4.3.1 CNN For $L1Calo$

The $L1Calo$ data is not an image, but its different cells do contain information related to one another. More so, the decay products trajectories can vary from one event to another and can appear anywhere within the TOB. It is their pattern that makes an event look more like a tau lepton event than a di-jet. In order for this kind of network to be able to classify tau lepton events properly, the data and the network usual architecture need to be modified.

The TOB object contains 99 cells of energy deposits. This data needs to be constructed in a way that represents an image-like formation. The trigger tower consists of five layers with different granularities. The initial step is to transform the coarser layers into a finer layer while still trying to preserve spatial information between adjacent layers. The method chosen here is to transform the data in a way that every deposit in a coarser layer would be copied into four cells as illustrated in Figure 7.7. Another method would be to zero-pad

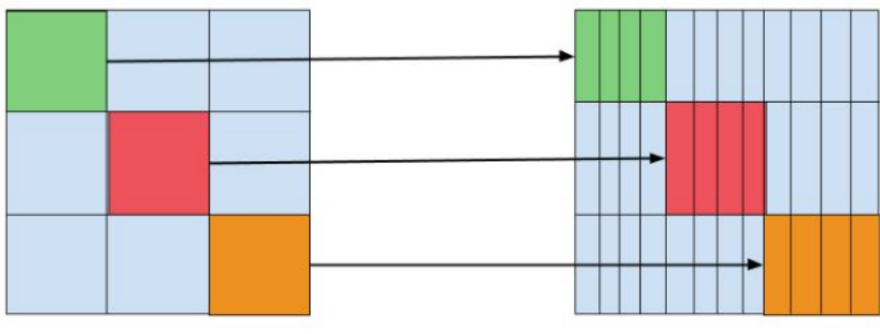


Figure 7.7: Preprocessing step for transforming the calorimeter coarser layers (EMo, EM₃, HAD) into a similar dimension as a finer one.

around existing cells, but this would lead to a loss of relations present in the calorimeter cells' physical domain.

The data at this stage has the dimensions of $Layers \times \Delta\eta \times \Delta\phi$ which corresponds to $5 \times 3 \times 12 = 120$ cells. This data cannot be simply sent into a standard CNN model. One has to think about the dimensions the convolution operation acts on. The filter has a fixed size, e.g. 5×5 , but, when dealing with an RGB picture, it has an additional dimension of depth, corresponding to the number of colour channels, making it a $5 \times 5 \times 3$ tensor. The convolution operation, more specifically a 2D convolution, is done simultaneously to all colour channels and then summed into a single value. A single neural network layer of convolution, consisting of n_f filters, transforms the input data dimensionality from $Height \times Width \times 3 \rightarrow Height \times Width \times n_f$. The motivation for that is the understanding that all colour channels are related, contributing to the final object and having a strong spatial relation. The analogue to a colour channel is a calorimeter layer. Together, all five of them in that specific order have a meaning derived from the tau lepton daughter particles. Thus, a *calorimeter pixel* is made up of five values as illustrated in Figure 7.8, while the entire event can be flattened into a very sparse rank-2 tensor as shown in Figure 7.9.

There are many off-the-shelf pre-trained models which can be used immediately with standard input.

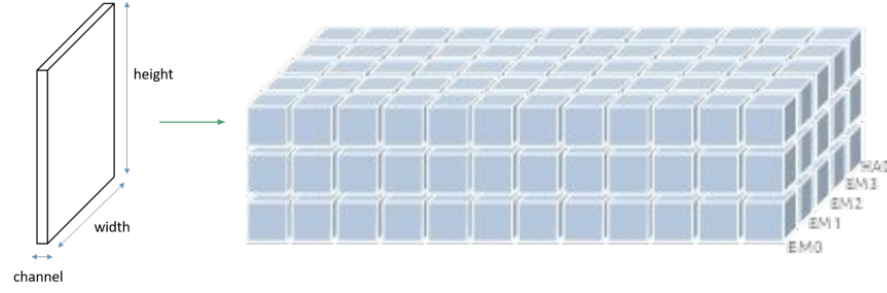


Figure 7.8: An image-like representation of the calorimeter layers.

7.4.3.2 Final Architecture

Once the data formation is set, the model can then be designed. The architecture is illustrated at Figure 7.10 and is made out of eight residual blocks [177]. Residual blocks are made of a set of stacked layers, where the inputs are added back to their outputs with the aim of creating identity mappings. In practice, such identity mappings are accomplished by means of the "skip" or residual connections. Their purpose is to enable the training of deeper models. Each block is made up of three consecutive stride-1 convolutions with a filter size of 3×3 . The filter size is kept constant throughout the model due to the input's small dimension.

The network has 39 layers that are trainable (the rest are fixed operations), corresponding to a 43,314 parameters. All layers, except the output layer, use the ReLU activation function.

The training and evaluation part of the network will be discussed in detail at Section 7.4.5.

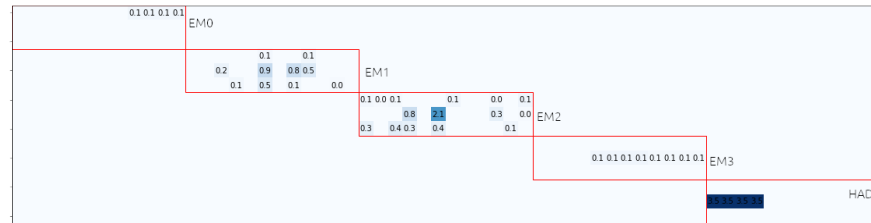


Figure 7.9: Illustration of an event's TOB different layers (after preprocessing) showing the sparse cells being activated along the particle trajectory. Using this raw data, the network should make a prediction.

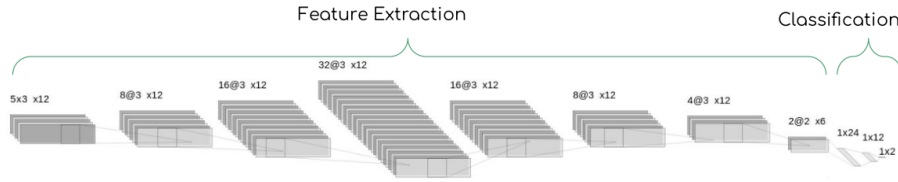


Figure 7.10: Final CNN model architecture for tau triggering in L1Calo. The first part is made of feature extraction, which then passes to a second and final classification layer.

7.4.4 Deep-Sets

As shown above (Figure 7.9) CNN representation is a very sparse one, with only 10% of the cells being active with the calorimeter in every TOB. A different approach to the same problem can lead to a better and deeper understanding. One can try and think of the data in the form of nodes in space, nodes with a special relationship with one another. This leads to defining an event as a *graph*. But, how should one define the graph's edges and the adjacency matrix? Are all layer nodes (cells) fully connected? Are the edges only close by nodes? Casting the data into a known pattern must come with a definite and reasonable characterisation. Within a layer, there is no temporal order. No single cell comes before nor after a different one. Thus, an event type (signal vs background) should not depend on the cells order, but merely on their properties and the fact they exist (energy deposited). Taking this into account, the data can be viewed as a *set*. The event is then formed in the following manner:

- each cell with an energy deposit is included in the set.
- every member of the set has the basic raw features:
 - energy in MeV
 - coordinates - a simple, discrete, relative location in the trigger tower grid with the x and y axes corresponding to η and ϕ and z -axis as the layer number.

Figure 7.11 shows an example of several data events.

e[MeV]	x	y	z
0.2	1	1	0
0.9	2	0	0
1.2	3	2	1
0.8	9	0	1
2.5	7	1	4

e[MeV]	x	y	z
0.2	1	1	0
0.9	2	0	0
1.2	3	2	1
2.8	9	0	1
0.2	5	1	2
1.6	6	1	2
0.1	2	2	3
3.1	0	1	4

e[MeV]	x	y	z
0.2	1	1	0
2.2	10	2	1

e[MeV]	x	y	z
2.2	1	1	0
0.1	2	0	0
2.1	3	2	1
3.8	9	0	1
0.5	11	2	2
0.9	6	2	3

Figure 7.11: Examples of TOB data in the form of a set. The length of the set is determined by the number of active cells, making each event of a different length.

Though this representation of the event is very efficient, it poses a problem - a set has no fixed size. Thus, an event could contain 15 or 33 cells. Neural networks act over a fixed size input.

The primary goal here is to find a function that maps the calorimeter input to a prediction. Thus, in this case, a *set function*.

Set function is a function whose input is a set and output is usually a number.

DeepSets [178] is a deep neural network generic framework meant to solve such problems by operating on sets. The network handles the case when the inputs to the network are permutation invariant sets rather than fixed dimensional vectors. As shown in [178], a function $S(X)$ is a valid set function if and only if it can be decomposed in the form $\rho(\sum_{x \in X} \psi(x))$ for a suitable transformations of ψ and ρ . Replacing ψ and ρ by universal approximators (see Section 7.3) leaves matters unchanged. Then, it remains to learn these approximators, meaning, training neural nets to approximate these functions.

Similar to CNN, DeepSet can be looked at as a two-stage operation, feature extraction (ψ) and classification. As Figure 7.12 shows, the model structure is defined as:

- each instance of the set's elements are transformed into some representation by $\psi(x)$.
- the summation $\sum_m \psi(x_m)$ of these representations are processed using the ρ network in the same manner as any deep neural network.

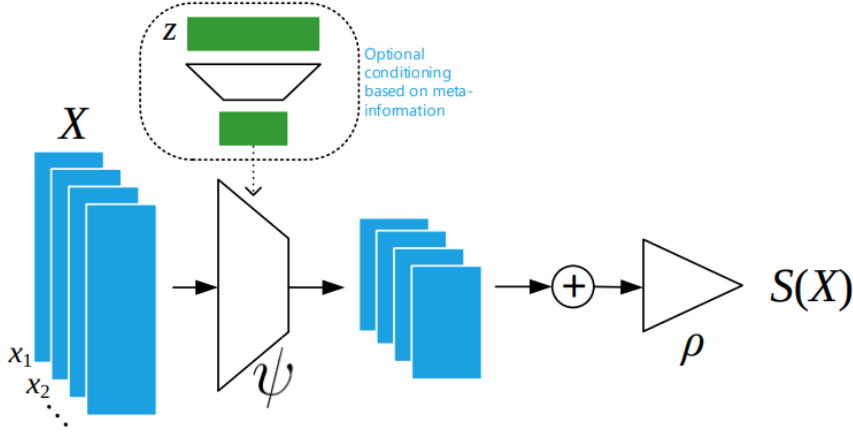


Figure 7.12: Generic architecture for DeepSets [178]

The second part of the summation is the key step. This allows the network to be both permutations invariant and to handle variable length input. By taking all of the set elements representation and adding them together into a single fixed length vector, the network can then proceed into classifying it in the standard way. All is left now is to find a suitable representation of the data, that is, extract meaningful features at ψ network, and then classify it using ρ .

7.4.4.1 Final Architecture

A common choice for feature extraction of data in such form (similar to time series analysis) is, once again, convolution, this time in one dimension. The final architecture is made up of three blocks of *DeepSetLayer* which consist of: $Conv1D \rightarrow ReLU \rightarrow Conv1D \rightarrow Tanh \rightarrow Conv1D \rightarrow ReLU \rightarrow Conv1D \rightarrow Tanh$, thus, resulting in 12 layers of convolution with $n_f = 256$, this is the ψ part of the network. Next, the feature representation is summed into a rank-1 tensor and passes to a fully connected, five-layer neural network, forming the ρ part. Figure 7.13 shows an illustration of the various parts of the network.

The entire network has 646,734 trainable parameters, which is an order of magnitude more than the CNN solution. The inference time,

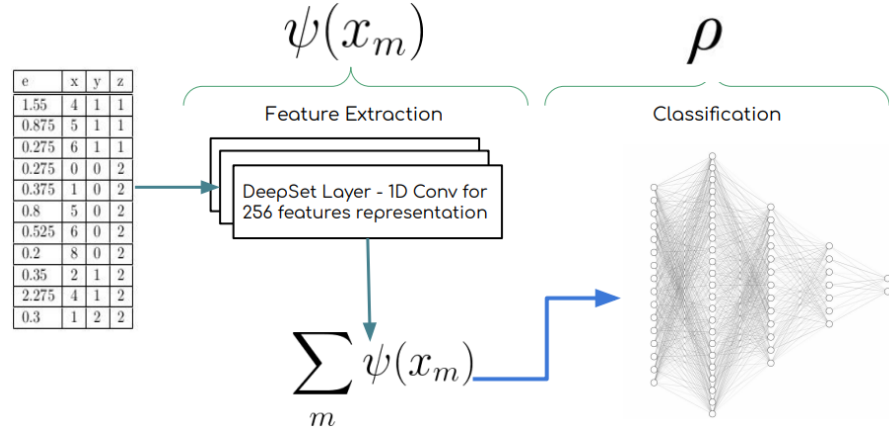


Figure 7.13: Illustration of DeepSet architecture and data flow for L1Calo tau identification.

or more speciously, computational operations is not necessarily higher. This will be discussed in Section 7.5.

7.4.5 Training

Both networks were constructed using PyTorch [179] and trained over Nvidia's Tesla K80 GPU. The normal procedures of data normalisation are applied in order for the networks to converge easily. The training is performed by minimising the weighted binary cross-entropy loss function, measuring the classification performance, with respect to the randomly initialised weights of the network. For the minimisation, stochastic gradient descent (SGD) is used. The convergence of the optimisation is improved by the Adam [180] optimiser algorithm that uses momentum and a stepwise reduction of the optimiser's learning rate when the loss on the monitoring sample stopped decreasing.

Both algorithms were trained for over 150 epochs. At each epoch, the algorithm is evaluated over the validation set and being saved. This process allows picking the best trained algorithm according to its recall and validation loss. Figure 7.14 shows the algorithms recall and losses as the training progress.

The loss function is weighted due to the MC samples ratio for signal and background.

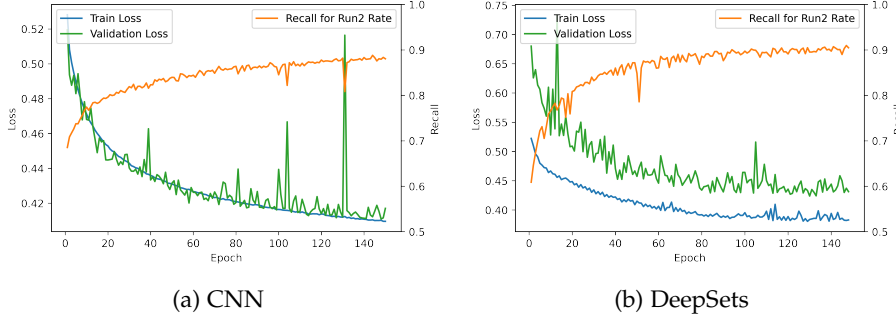


Figure 7.14: Training and validation loss vs recall for both algorithms. Some spikes occur in the validation, but one should consider the overall trend, which does not exhibit overtraining.

7.5 RESULTS

The performance of the tau identification algorithms is evaluated on statistically independent test samples of $Z \rightarrow \tau\bar{\tau}$ (signal) and di-jet (background) events. The networks performance is compared to that of Run 2 and BC. Before the final evaluation, the background rejection as compared to the signal efficiency (Figure 7.15) and a working point has been chosen with $\epsilon_{CNN} = 0.91$, $\epsilon_{DeepSet} = 0.93$.

Both DNN differ from the benchmark in the way they make the final decision. The benchmark algorithms estimate the τ 's p_T and set a threshold over it (to keep the fake rate low). The DNNs, on the other hand, are classifying the event, providing a probability for it to be a τ and the threshold is set over this probability in order to provide an equal Run 2 fake rate.

The final evaluation is summarised in a *Turn-On* curve, defined as the trigger efficiency as a function of the true tau p_T . As previously mentioned, the higher the tau lepton p_T , the easier it is to identify it and suppress the background processes. Thus, an improvement over existing algorithms would be increased efficiency in low- p_T regions while still keeping the fake rate the same as Run 2. Figure 7.16 shows the higher performance of the new algorithms compared to the benchmarks.

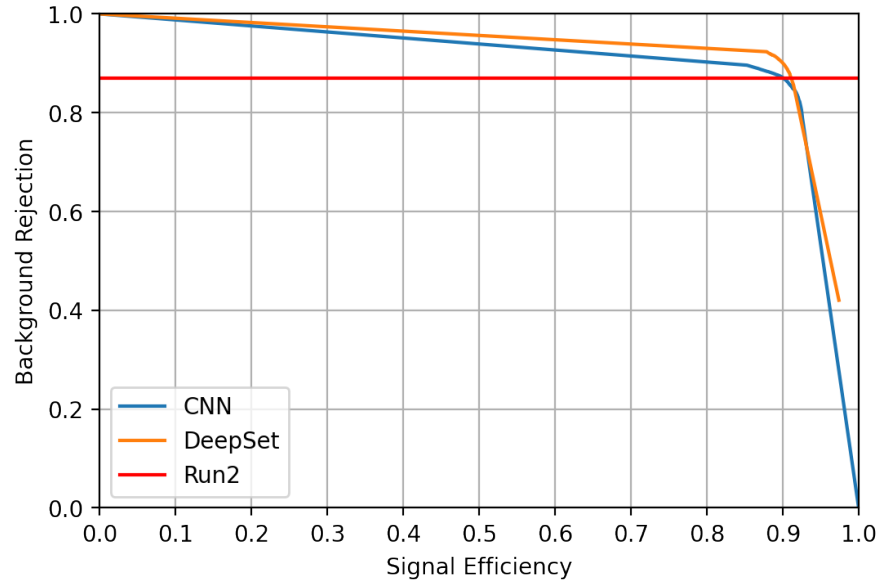


Figure 7.15: Signal efficiency vs background rejection for CNN and DeepSet architectures. Run 2 threshold for fake rate is in red. The point of intersection is chosen for final evaluation.

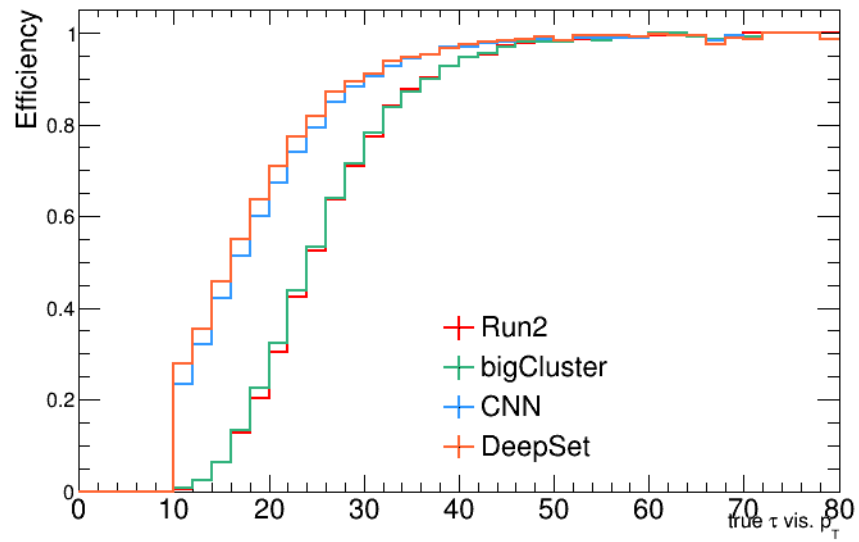


Figure 7.16: Turnon curve comparing DNN algorithm with current benchmarks. A significant gain in trigger efficiency is present in the low- p_T part.

Algorithm	FLOPS
Run-2	$0.1 \cdot 10^2$
bigCluster	$0.5 \cdot 10^2$
CNN	$1.5 \cdot 10^6$
DeepSet	$1.3 \cdot 10^7$

Table 7.1: Floating Point Operations per Second for each algorithm are measured on a standard CPU. The estimate included pre-processing steps. DeepSet value corresponds to the average set size of 12 cells active.

7.6 CONCLUSION

Trigger algorithms are not reconstruction algorithms. There are strict demands for trigger algorithm in terms of the number of computer cycles allowed; any slightest deviation will cause the trigger to become a bottleneck, and the entire chain of operation will cease to work.

Table 7.1 compares the number of Floating Point Operations per Second (FLOPS), a common way of measuring an algorithm's computational cost. First, DeepSet slightly better performance comes with a price of a full order of magnitude more computational operations. Second, the FLOPS of CNN and DeepSet are significantly higher than the benchmarks. Each implementation over an FPGA requires a dedicated conversion to logical gates, making every processing step a high cost and a possible bottleneck.

This leads to the conclusion that for the Run 3 hardware currently installed, the prospect of upgrading to a DNN based decision over FPGA is not realistic. While using this trigger architecture in the next ATLAS run is not feasible, the purpose of such researches is their ability to steer future decisions, specifically, ones concerning hardware upgrades. As the industry evolves, newer and more powerful GPUs are coming out, making them a prime choice for any future upgrade of the experiment. No doubt, these now-complicated algorithms shall become the next benchmark and will continue to push the boundaries further.

Reconstruction algorithms run offline and are not generally limited by hardware constraints.

SUMMARY

To date, the Standard Model is the theory that best encompasses our knowledge of particle physics; it elegantly describes matter in terms of elementary constituents and their interactions and provides a coherent description of three of the four fundamental forces of nature. There are many unanswered questions in particle physics, which can not be answered by the Standard Model and require extensions to the Standard Model. Questions about gravity, unification of fundamental forces and the origin of dark matter and energy are some of these. The high-quality data collected by ATLAS during Run 1 and Run 2 have been used in a plethora of physics analyses, including precision measurements of Standard Model processes, the discovery of the Higgs boson, and many Beyond the Standard Model searches. These analyses have provided results consistent with the predictions made by the Standard Model. Analyses performed searching for the direct production of particles expected by various Beyond the Standard Model theories, such as 2HDM, have not yet yielded discoveries. This thesis contains one such Beyond the Standard Model search ($H^\pm \rightarrow t\bar{b}$) and a precision measurement ($H \rightarrow c\bar{c}$), both probing the Higgs sector.

Introducing a 2HDM extension to the Standard Model would lead to five Higgs bosons, of which three are neutral, and two are charged. With the discovery of a neutral Higgs boson in 2012, the last particle of the Standard Model has been found. The remaining question is, of course, if it is really the neutral Higgs boson of the Standard Model or maybe a Higgs boson in a model beyond the Standard Model. The answer may be found through future precision measurements of this particle (e.g. the coupling strength of the Higgs to charm quarks). On the other hand, discovering a charged Higgs boson would be a clear

indication for physics beyond the Standard Model. Searches for such a new charged scalar particle were part of the physics program of ATLAS from the beginning.

Chapter 5 describes the search for charged Higgs bosons $H^\pm$ produced in association with a top and a bottom quark and decaying into a top and a bottom quark, $pp \rightarrow tbH^+ \rightarrow tbtb$, in the lepton-plus-jets and dilepton final states. The mass range $200 \text{ GeV} \leq m_{H^\pm} \leq 2000 \text{ GeV}$ was investigated. No significant excess above the background-only hypothesis was observed, and observed (expected) upper cross-section limits between 2.9 pb (3.2 pb) and 0.04 pb (0.08 pb) were set.

Chapter 6 contains a direct search for the decay of a Higgs boson to a charm quark-antiquark pair. The search was performed using 139 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13 \text{ TeV}$ by the ATLAS detector. It uses three channels, $ZH \rightarrow \nu\nu c\bar{c}$, $WH \rightarrow \ell\nu c\bar{c}$ and $ZH \rightarrow \ell\ell c\bar{c}$. The analysis yields an expected limit of $26.0^{+10.1}_{-7.3}$ times the predicted cross-section times branching fraction for a Higgs boson, with a mass of 125 GeV, decaying to a charm quark-anti-quark pair. Observing this decay would give new insight into the Higgs boson's relationship with the second generation of matter particles. Further, measuring the coupling of the Higgs boson's interaction to the charm quark could give insight into new physics processes. But spotting such decay is quite a challenge. While its branching fraction of 3% seems small, the main difficulty lies in the fact the search is heavily dominated by background processes.

Both of this thesis's searches had to focus on a thorough study of the background and their corresponding systematics to understand their impact on the fit. Only with such careful handling one can derive as accurate as possible results. In $H^\pm$ the focus was on $t\bar{t}b\bar{b}$ and resulted in designing a unique discriminator which targeted this background, thus enabling a better constrain of its modelling uncertainty in the final fit. In $H \rightarrow c\bar{c}$ it was V +jets with a similar signature to the signal being the governing background in all channels. Different Monte-Carlo generators and flavour splitting using sophisticated tagging

algorithms were used to estimate its uncertainty both in terms of normalisation and shape.

Chapter 7 is complementary to the analyses presented here. Interpreting the detector's raw data and Identifying the τ -lepton using Deep Neural Networks provides a different angle to work at particle accelerators from the analysis side. While working on a search, physicists analyse well-processed data. The very simple statement that an event contained an electron or a photon with a precise measurement of its momentum encapsulates an enormous amount of work being done behind the scenes. Reconstruction and trigger algorithms have been perfected over the years, all in order to go from the discharge of a tiny capacitor in the sub-detectors complex electronics to a track reconstruction in a very precise manner. Contributing to the better identification of any physics object would benefit the entire collaboration.

The use of Deep Learning in High Energy Physics today is embedded in various aspects. In the early days, most of the effort was put into the mathematical coding of the algorithms. Nowadays, the status of the Deep Learning frameworks and the availability of high-performance Graphical Processing Units enable physicists to focus mainly on understanding the physical event, the data that construct it and the physics-related feature that can be derived.

Properly identifying a τ -lepton within a Trigger Object has proven to be a challenging task. Two novel methods were presented. Both make use of Deep Learning in order to obtain a classification. The fact that both methods ran over the same data and yet yielded different results only emphasises the importance of properly understanding the data used. An improvement of 2% can be meaningful to an analysis, especially when seeking an elusive event.

The DeepSet architecture yielded a 93% signal efficiency while still keeping the same background rate as the Run 2 current algorithm. This efficiency improved significantly the Turn-On curve, which in turn is enabling the exploration of low- p_T events that will contribute to many searches such as light charged Higgs decaying into tau and neutrino ($H^\pm \rightarrow \tau\nu$), double production of light charged Higgs, low mass search to $\tau\bar{\tau}$ and more.

BIBLIOGRAPHY

- [1] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (16 1964) 508, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [2] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132.
- [3] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (9 1964) 321, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [4] 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** (2012) 1, arXiv: 1207.7214 [hep-ex].
- [5] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30, arXiv: 1207.7235 [hep-ex].
- [6] S. L. Glashow, *Partial-symmetries of weak interactions*, *Nuclear Physics* **22** (1961) 579 , ISSN: 0029-5582, URL: <http://www.sciencedirect.com/science/article/pii/0029558261904692>.
- [7] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (21 1967) 1264, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [8] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, *Nuovo Cim.* **11** (1959) 568.
- [9] ATLAS Collaboration, *Search for charged Higgs bosons decaying into top and bottom quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Journal of High Energy Physics* **2018** (2018), ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP11\(2018\)085](http://dx.doi.org/10.1007/JHEP11(2018)085).

- [10] ATLAS Collaboration, *Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment*, *Phys. Rev. Lett.* **120** (21 2018) 211802, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.120.211802>.
- [11] B. Stelzer, *The New Small Wheel Upgrade Project of the ATLAS Experiment*, (2014), URL: <https://cds.cern.ch/record/1958265>.
- [12] A. Abusleme et al., *Performance of a full-size small-strip thin gap chamber prototype for the ATLAS new small wheel muon upgrade*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **817** (2016) 85, ISSN: 0168-9002, URL: <https://www.sciencedirect.com/science/article/pii/S0168900216001285>.
- [13] ALEPH Collaboration, *Determination of the number of light neutrino species*, *Physics Letters B* **231** (1989) 519, ISSN: 0370-2693, URL: <http://www.sciencedirect.com/science/article/pii/0370269389907041>.
- [14] Particle Data Group, *Review of Particle Physics*, *Phys. Rev. D* **98** (3 2018) 030001, URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- [15] M. Karliner and J. L. Rosner, *Discovery of the Doubly Charmed Ξ_{cc} Baryon Implies a Stable $bb\bar{u}\bar{d}$ Tetraquark*, *Phys. Rev. Lett.* **119** (20 2017) 202001, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.119.202001>.
- [16] G. Cowan and T. Gershon, *Tetraquarks and Pentaquarks*, 2399-2891, IOP Publishing, 2018, ISBN: 978-0-7503-1593-7, URL: <http://dx.doi.org/10.1088/978-0-7503-1593-7>.
- [17] J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, *A Historical Profile of the Higgs Boson*, (2012), arXiv: 1201.6045 [hep-ph].
- [18] ATLAS Collaboration, *Measurement of Higgs boson production in association with a $t\bar{t}$ pair in the diphoton decay channel using*

- 139 fb^{-1} of LHC data collected at $\sqrt{s} = 13\text{ TeV}$ by the ATLAS experiment, (2019), URL: <http://cds.cern.ch/record/2668103>.
- [19] A. M. Sirunyan et al., *Measurements of $t\bar{t}H$ production and the CP structure of the Yukawa interaction between the Higgs boson and top quark in the diphoton decay channel*, (2020), arXiv: [2003.10866](https://arxiv.org/abs/2003.10866) [hep-ex].
- [20] J. R. Andersen et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, (2013), ed. by S Heinemeyer et al., arXiv: [1307.1347](https://arxiv.org/abs/1307.1347) [hep-ph].
- [21] T. D. Lee, *A Theory of Spontaneous T Violation*, *Phys. Rev. D* **8** (4 1973) 1226, URL: <https://link.aps.org/doi/10.1103/PhysRevD.8.1226>.
- [22] M. Maniatis, A. von Manteuffel, and O. Nachtmann, *On the phenomenology of a two-Higgs-doublet model with maximal CP symmetry at the LHC - synopsis and addendum*, (2010), arXiv: [1009.1869](https://arxiv.org/abs/1009.1869) [hep-ph].
- [23] S. L. Glashow and S. Weinberg, *Natural conservation laws for neutral currents*, *Phys. Rev. D* **15** (7 1977) 1958, URL: <https://link.aps.org/doi/10.1103/PhysRevD.15.1958>.
- [24] J. Schaarschmidt, *HPLUSMSSM*, GitHub repository ().
- [25] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008) S08002, URL: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08002>.
- [26] CMS Collaboration, *The CMS experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008) S08004, URL: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08004>.
- [27] LHCb Collaboration, *The LHCb Detector at the LHC*, *Journal of Instrumentation* **3** (2008) S08005, URL: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08005>.

- [28] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** (2008) So8003, URL: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003>.
- [29] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 1*, Technical design report. ATLAS, CERN, 1997, URL: <https://cds.cern.ch/record/331063>.
- [30] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 2*, Technical design report. ATLAS, CERN, 1997, URL: <https://cds.cern.ch/record/331064>.
- [31] ATLAS Collaboration, *ATLAS pixel detector: Technical Design Report. Technical Design Report ATLAS*, 1998.
- [32] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, *JINST* **3** (2008) Po7007, URL: <https://cds.cern.ch/record/1119279>.
- [33] ATLAS Collaboration, *The Barrel Modules of the ATLAS Semiconductor Tracker*, (2006), URL: <https://cds.cern.ch/record/974073>.
- [34] ATLAS Collaboration, *The ATLAS semiconductor tracker end-cap module*, *Nucl. Instrum. Methods Phys. Res., A* **575** (2007) 353, URL: <https://cds.cern.ch/record/1063618>.
- [35] ATLAS Collaboration, *The Silicon microstrip sensors of the ATLAS semiconductor tracker*, (2007), URL: <https://cds.cern.ch/record/1019885>.
- [36] ATLAS Collaboration, *The ATLAS TRT barrel detector*, *JINST* **3** (2008) Po2014, URL: <https://cds.cern.ch/record/1094548>.
- [37] ATLAS Collaboration, *The ATLAS TRT end-cap detectors*, *JINST* **3** (2008) P10003.
- [38] ATLAS Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, *JINST* **3** (2008) Po2013.

- [39] ATLAS Collaboration, *ATLAS central solenoid: Technical Design Report*, Technical design report. ATLAS, Electronic version not available, CERN, 1997, URL: <https://cds.cern.ch/record/331067>.
- [40] ATLAS Collaboration, *ATLAS calorimeter performance: Technical Design Report*, Technical design report. ATLAS, CERN, 1996, URL: <https://cds.cern.ch/record/331059>.
- [41] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, Technical design report. ATLAS, CERN, 1996, URL: <https://cds.cern.ch/record/331061>.
- [42] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*, Technical design report. ATLAS, CERN, 1996, URL: <https://cds.cern.ch/record/331062>.
- [43] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*, Technical design report. ATLAS, CERN, 1997, URL: <https://cds.cern.ch/record/331068>.
- [44] ATLAS Collaboration, *ATLAS barrel toroid: Technical Design Report*, Technical design report. ATLAS, Electronic version not available, CERN, 1997, URL: <https://cds.cern.ch/record/331065>.
- [45] ATLAS Collaboration, *ATLAS end-cap toroids: Technical Design Report*, Technical design report. ATLAS, Electronic version not available, CERN, 1997, URL: <https://cds.cern.ch/record/331066>.
- [46] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554, [Erratum: *Eur. Phys. J. C* 73,2501(2013)], arXiv: [1007.1727](https://arxiv.org/abs/1007.1727) [[physics.data-an](https://arxiv.org/archive/physics)].
- [47] A. L. Read, *Presentation of search results: the CL_s technique*, *J. Phys. G* **28** (2002) 2693.

- [48] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, *Phys. Rev. Lett.* **114** (2015) 191803, arXiv: 1503.07589 [hep-ex].
- [49] T. P. Cheng and L.-F. Li, *Neutrino Masses, Mixings and Oscillations in $SU(2) \times U(1)$ Models of Electroweak Interactions*, *Phys. Rev. D* **22** (1980) 2860.
- [50] J. Schechter and J. W. F. Valle, *Neutrino Masses in $SU(2) \times U(1)$ Theories*, *Phys. Rev. D* **22** (1980) 2227.
- [51] G. Lazarides, Q. Shafi, and C. Wetterich, *Proton Lifetime and Fermion Masses in an $SO(10)$ Model*, *Nucl. Phys. B* **181** (1981) 287.
- [52] R. N. Mohapatra and G. Senjanovic, *Neutrino Masses and Mixings in Gauge Models with Spontaneous Parity Violation*, *Phys. Rev. D* **23** (1981) 165.
- [53] M. Magg and C. Wetterich, *Neutrino Mass Problem and Gauge Hierarchy*, *Phys. Lett. B* **94** (1980) 61.
- [54] T. D. Lee, *A Theory of Spontaneous T Violation*, *Phys. Rev. D* **8** (1973) 1226.
- [55] G. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, *Phys.Rept.* **516** (2012) 1, arXiv: 1106.0034 [hep-ph].
- [56] K. Inoue et al., *Aspects of Grand Unified Models with Softly Broken Supersymmetry*, *Prog.Theor.Phys.* **68** (1982) 927.
- [57] T. Cheng and L.-F. Li, *Neutrino Masses, Mixings and Oscillations in $SU(2) \times U(1)$ Models of Electroweak Interactions*, *Phys. Rev. D* **22** (1980) 2860.
- [58] LEP working group for Higgs boson searches: ALEPH, DELPHI, L3 and OPAL Collaborations, *Search for Charged Higgs bosons: Combined Results Using LEP Data*, *Eur. Phys. J. C* **73** (2013) 2463, arXiv: 1301.6065 [hep-ex].

- [59] CDF Collaboration, T. Aaltonen and others, *Search for charged Higgs bosons in decays of top quarks in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev. Lett.* **103** (2009) 101803, arXiv: 0907.1269 [hep-ex].
- [60] D0 Collaboration, V. M. Abazov and others, *Search for charged Higgs bosons in top quark decays*, *Phys. Lett. B* **682** (2009) 278, arXiv: 0908.1811 [hep-ex].
- [61] D0 Collaboration, V. M. Abazov and others, *Search for charged Higgs bosons decaying to top and bottom quarks in $p\bar{p}$ collisions*, *Phys. Rev. Lett.* **102** (2009) 191802, arXiv: 0807.0859 [hep-ex].
- [62] CMS Collaboration, *Search for a charged Higgs boson in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **11** (2015) 018, arXiv: 1508.07774 [hep-ex].
- [63] ATLAS Collaboration, *Search for charged Higgs bosons in the $H^\pm \rightarrow tb$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *JHEP* **03** (2016) 127, arXiv: 1512.03704 [hep-ex].
- [64] ATLAS Collaboration, *Search for charged Higgs bosons in the $H^\pm \rightarrow tb$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, ATLAS-CONF-2016-089, 2016, URL: <https://cds.cern.ch/record/2206809>.
- [65] ATLAS Collaboration, *Search for charged Higgs bosons decaying into a top-quark and a bottom-quark at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2020), URL: <https://cds.cern.ch/record/2727860>.
- [66] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2017), ATLAS-CONF-2017-079, internal documentation for CERN-EP-2017-291, URL: <https://cds.cern.ch/record/2244360>.
- [67] M. Aaboud et al., *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair*

- in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, (2017), arXiv: 1712.08895 [hep-ex].
- [68] ATLAS Collaboration, “Top common objects 13 TeV,” twiki page, 2016, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopCommonObjects13TeV>.
- [69] 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* **1407** (2014) 079, arXiv: 1405.0301 [hep-ph].
- [70] T. Sjostrand, S. Mrenna, and P. Skands, *A brief introduction to {PYTHIA} 8.1*, *Computer Physics Communications* **178** (2008) 852, ISSN: 0010-4655, URL: <http://www.sciencedirect.com/science/article/pii/S0010465508000441>.
- [71] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **0411** (2004) 040, arXiv: hep-ph/0409146.
- [72] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **0711** (2007) 070, arXiv: 0709.2092 [hep-ph].
- [73] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **1006** (2010) 043, arXiv: 1002.2581 [hep-ph].
- [74] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930, arXiv: 1112.5675 [hep-ph].
- [75] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, *B710* (2012) 612, arXiv: 1111.5869 [hep-ph].
- [76] P. Bärnreuther, M. Czakon, and A. Mitov, *Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$* , *Phys. Rev. Lett.* **109** (13 2012) 132001,

- URL: <https://link.aps.org/doi/10.1103/PhysRevLett.109.132001>.
- [77] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, *JHEP* **1212** (2012) 054, arXiv: 1207.0236 [hep-ph].
 - [78] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the quark-gluon reaction*, *JHEP* **1301** (2013) 080, arXiv: 1210.6832 [hep-ph].
 - [79] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_s^4)$* , *Phys. Rev. Lett.* **110** (25 2013) 252004, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.110.252004>.
 - [80] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **0902** (2009) 007, arXiv: 0811.4622 [hep-ph].
 - [81] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601, arXiv: 1111.5206 [hep-ph].
 - [82] F. Cascioli et al., *NLO matching for $t\bar{t}b\bar{b}$ production with massive b -quarks*, *Phys. Lett. B* **734** (2014) 210, arXiv: 1309.5912 [hep-ph].
 - [83] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **0803** (2008) 038, arXiv: 0709.1027 [hep-ph].
 - [84] S. Höche et al., *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027, arXiv: 1207.5030 [hep-ph].
 - [85] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **0605** (2006) 026, arXiv: hep-ph/0603175.
 - [86] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with aW-orH-*, *Physical Review D* **82** (2010), ISSN: 1550-2368, URL: <http://dx.doi.org/10.1103/PhysRevD.82.054018>.

- [87] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, *Physical Review D* **81** (2010), ISSN: 1550-2368, URL: <http://dx.doi.org/10.1103/PhysRevD.81.054028>.
- [88] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for s-channel single top quark production*, *Physical Review D* **83** (2011), ISSN: 1550-2368, URL: <http://dx.doi.org/10.1103/PhysRevD.83.091503>.
- [89] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det].
- [90] GEANT4 Collaboration, S. Agostinelli and others., *GEANT4: A Simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [91] W Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*, (2012), ATL-SOFT-PROC-2012-065, URL: <https://cds.cern.ch/record/1458503>.
- [92] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244, arXiv: 1207.1303 [hep-ph].
- [93] C. Degrande et al., *Heavy charged Higgs boson production at the LHC*, *JHEP* **10** (2015) 145, arXiv: 1507.02549 [hep-ph].
- [94] C. Degrande et al., *Accurate predictions for charged Higgs production: Closing the $m_{H^\pm} \sim m_t$ window*, *Phys. Lett. B* **772** (2017) 87, arXiv: 1607.05291 [hep-ph].
- [95] M. Flechl et al., “Santander-matched cross sections for the process $gg \rightarrow tH + X$,” web page, 2016, URL: http://www.hephy.at/user/mflechl/hp_xsec/xsec_13TeV_tHp_2016_2_5.txt.
- [96] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph].
- [97] M. Flechl et al., *Improved cross-section predictions for heavy charged Higgs boson production at the LHC*, *Phys. Rev. D* **91** (2015) 075015, arXiv: 1409.5615 [hep-ph].

- [98] S. Dittmaier et al., *Charged-Higgs-boson production at the LHC: NLO supersymmetric QCD corrections*, *Phys. Rev. D* **83** (2011) 055005, arXiv: 0906.2648 [hep-ph].
- [99] E. L. Berger et al., *Associated production of a top quark and a charged Higgs boson*, *Phys. Rev. D* **71** (2005) 115012, arXiv: hep-ph/0312286 [hep-ph].
- [100] E. Bergeaas Kuutmann, L.-M. Mir, J. Schaarschmidt, et al., *Search for $H^\pm \rightarrow tb$ with the ATLAS detector in 13 TeV pp collisions: Internal documentation of the analysis*, (2016), ATL-COM-PHYS-2016-097, internal documentation for ATLAS-CONF-2016-089, URL: <https://cds.cern.ch/record/2128622>.
- [101] J. M. Campbell et al., *Top-pair production and decay at NLO matched with parton showers*, *JHEP* **04** (2015) 114, arXiv: 1412.1828 [hep-ph].
- [102] R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **1504** (2015) 040, arXiv: 1410.8849 [hep-ph].
- [103] ATLAS Collaboration, *ATLAS Run 1 Pythia8 tunes*, (2011), ATL-PHYS-PUB-2014-021, URL: <http://cdsweb.cern.ch/record/1966419>.
- [104] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152.
- [105] H.-L. Lai et al., *New parton distributions for collider physics*, *Phys. Rev. D* **82** (2010) 074024, arXiv: 1007.2241 [hep-ph].
- [106] ATLAS Collaboration, *Simulation of top-quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2016-004, 2016, URL: <https://cds.cern.ch/record/2120417>.
- [107] M. Bahr et al., *Herwig++ Physics and Manual*, *Eur. Phys. J. C* **58** (2008) 639, arXiv: 0803.0883 [hep-ph].
- [108] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196, arXiv: 1512.01178 [hep-ph].

- [109] ATLAS Collaboration, *Studies of $t\bar{t}c\bar{c}$ production with MADGRAPH5_AMC@NLO and HERWIG++ for the ATLAS experiment*, ATL-PHYS-PUB-2016-011, 2016, URL: <https://cds.cern.ch/record/2153876>.
- [110] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **0812** (2008) 039, arXiv: 0808.3674 [hep-ph].
- [111] J Butterworth et al., *Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s}=7$ TeV*, ATL-COM-PHYS-2010-695 (2010), <https://cds.cern.ch/record/1287902>.
- [112] S. Frixione et al., *Single-top hadroproduction in association with a W boson*, *JHEP* **0807** (2008) 029, arXiv: 0805.3067 [hep-ph].
- [113] P. Artoisenet et al., *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, *JHEP* **03** (2013) 015, arXiv: 1212.3460 [hep-ph].
- [114] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, *Phys. Rev. D* **82** (2010) 074018, arXiv: 1005.3457 [hep-ph].
- [115] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **75** (2015) 349, arXiv: 1503.05066 [hep-ex].
- [116] S. Guindon, E. Shabalina, J. Adelman, et al., *Search for the Standard Model Higgs boson produced in association with top quarks in pp collisions at 8 TeV with the ATLAS detector at the LHC*, (2014), ATL-COM-PHYS-2014-1471, internal documentation for EPJ C75 (2015) 349, URL: <https://cds.cern.ch/record/1969527>.
- [117] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton–proton collision data*, *Eur. Phys. J. C* **72** (2012) 1909, arXiv: 1110.3174 [hep-ex].
- [118] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton–proton collision data*, *Eur. Phys. J. C* **74** (2014) 2941, arXiv: 1404.2240 [hep-ex].

- [119] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016, URL: <https://cds.cern.ch/record/2157687>.
- [120] ATLAS Collaboration, *Tracking-CP analysis guidelines for DC14*, (2015), URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/InDetTrackingDC14#Impact_parameters_z0_d0_definiti.
- [121] ATLAS Collaboration, “Official Isolation Working Points,” twiki page, 2017, URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/IsolationSelectionTool>.
- [122] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, 2008, URL: <https://cds.cern.ch/record/1099735>.
- [123] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_t jet clustering algorithm*, **JHEP** **04** (2008) 063, arXiv: [0802.1189](https://arxiv.org/abs/0802.1189) [hep-ph].
- [124] ATLAS Collaboration, *Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015, URL: <https://cds.cern.ch/record/2037702>.
- [125] G. Aad et al., *Performance of b -Jet Identification in the ATLAS Experiment*, **JINST** **11** (2016) P04008, arXiv: [1512.01094](https://arxiv.org/abs/1512.01094) [hep-ex].
- [126] ATLAS Collaboration, *Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, 2016, URL: <https://cds.cern.ch/record/2160731>.
- [127] L-M Mir o.b.o the $H^+(\text{tb})$ analysis team, “Status Report on $H^+ \rightarrow tb$,” HTop plenary meeting, 20 April 2016, 2016, URL: https://indico.cern.ch/event/521393/contributions/2135304/attachments/1261015/1863852/Hplus_status_Htop_20April2016.pdf.

- [128] A. Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS ACAT (2007) 040, arXiv: [physics/0703039](#), URL: http://pos.sissa.it/archive/conferences/050/040/ACAT_040.pdf.
- [129] ATLAS Collaboration, *Top Systematic Uncertainties for early 13 TeV Analyses*, (2017), URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopSystematics13TeV>.
- [130] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, *Eur. Phys. J. C* **76** (2016) 653. 71 p, arXiv: [1608.03953](#), URL: <https://cds.cern.ch/record/2208146>.
- [131] M. Aaboud et al., *Measurement of the Inelastic Proton-Proton Cross section at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC*, *Phys. Rev. Lett.* **117** (2016) 182002, arXiv: [1606.02625](#) [hep-ex].
- [132] W. Buttinger, *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*, ATL-COM-SOFT-2015-119 (2015), URL: <https://cds.cern.ch/record/2014726>.
- [133] ATLAS Collaboration, *Extended Pileup Reweighting*, (2017), URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ExtendedPileupReweighting>.
- [134] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **76** (2016) 292, arXiv: [1603.05598](#) [hep-ex].
- [135] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, (2015), URL: <https://cds.cern.ch/record/2037613>.
- [136] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2017), arXiv: [1703.09665](#) [hep-ex].

- [137] ATLAS Collaboration, *Uncertainty release for analyses using 2015 and 2016 data in 20.7 (MC15c) releases*, (2017), URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertainties20152016Data20p7>.
- [138] ATLAS Collaboration, *Flavour tagging recommendations in release 20.7*, (2017), URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTagCalib2015>.
- [139] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys.* **G43** (2016) 023001, arXiv: 1510.03865 [hep-ph].
- [140] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{+-}$ production and decay at NLO*, *JHEP* **1207** (2012) 052, arXiv: 1204.5678 [hep-ph].
- [141] ATLAS Collaboration, *Multi-boson simulation for 13 TeV ATLAS analyses*, ATL-PHYS-PUB-2016-002, 2016, URL: <https://cds.cern.ch/record/2119986>.
- [142] M. Carena et al., *MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios after the Discovery of a Higgs-like Particle*, *Eur. Phys. J.* **C73** (2013) 2552, arXiv: 1302.7033 [hep-ph].
- [143] A. Djouadi et al., *The post-Higgs MSSM scenario: Habemus MSSM?* *Eur. Phys. J.* **C73** (2013) 2650, arXiv: 1307.5205 [hep-ph].
- [144] A. Djouadi et al., *Fully covering the MSSM Higgs sector at the LHC*, *JHEP* **06** (2015) 168, arXiv: 1502.05653 [hep-ph].
- [145] E. Bagnaschi et al., *Benchmark scenarios for low $\tan \beta$ in the MSSM*, (2015), LHCHSWG-2015-002, URL: <https://cds.cern.ch/record/2039911>.
- [146] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in pp collisions at $s=13$ TeV with the ATLAS detector*, *Physics Letters B* **800** (2020) 135103, ISSN: 0370-2693, URL: <http://dx.doi.org/10.1016/j.physletb.2019.135103>.
- [147] C. Collaboration, *Combination of Searches for Higgs Boson Pair Production in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. Lett.* **122** (12 2019) 121803, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.122.121803>.

- [148] ATLAS Collaboration, *A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector*, (2020), arXiv: [2007.07830 \[hep-ex\]](#).
- [149] C. Collaboration, *Evidence for Higgs boson decay to a pair of muons*, (2020), arXiv: [2009.04363 \[hep-ex\]](#).
- [150] G. Luisoni et al., *$HW \pm/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*, *Journal of High Energy Physics* **2013** (2013) 83, ISSN: 1029-8479, URL: [http://dx.doi.org/10.1007/JHEP10\(2013\)083](http://dx.doi.org/10.1007/JHEP10(2013)083).
- [151] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **09** (2014) 145, arXiv: [1406.3660 \[hep-ex\]](#).
- [152] A. Denner et al., *HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders*, *Comput. Phys. Commun.* **195** (2015) 161, arXiv: [1412.5390 \[hep-ph\]](#).
- [153] A. Denner et al., *Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with Hawk*, *JHEP* **03** (2012) 075, arXiv: [1112.5142 \[hep-ph\]](#).
- [154] ATLAS Collaboration, *Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2*, ATL-PHYS-PUB-2016-003, 2016, URL: <https://cds.cern.ch/record/2120133>.
- [155] E. Bothmann et al., *Event generation with Sherpa 2.2*, *SciPost Physics* **7** (2019), ISSN: 2542-4653, URL: <http://dx.doi.org/10.21468/SciPostPhys.7.3.034>.
- [156] L. Lonnblad, *Correcting the color dipole cascade model with fixed order matrix elements*, *JHEP* **05** (2002) 046, arXiv: [hep-ph/0112284 \[hep-ph\]](#).

- [157] N. Lavesson and L. Lonnblad, *W+jets matrix elements and the dipole cascade*, *JHEP* **07** (2005) 054, arXiv: [hep - ph / 0503293](#) [[hep-ph](#)].
- [158] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, (2017), URL: <https://cds.cern.ch/record/2273281>.
- [159] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *The European Physical Journal C* **79** (2019), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-019-7450-8>.
- [160] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **231** (1933) 289, ISSN: 02643952, URL: <http://www.jstor.org/stable/91247>.
- [161] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, (2019), URL: <https://cds.cern.ch/record/2677054>.
- [162] ATLAS Collaboration, *Extended Pileup Reweighting*, 2015, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ExtendedPileupReweighting> (visited on 11/30/2015).
- [163] ATLAS Collaboration, *Electron Efficiencies for Run2 (scale factors and uncertainties)*, 2015, URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronEfficiencyRun2> (visited on 12/09/2016).
- [164] ATLAS Collaboration, *MCP Analysis Guidelines MC16*, 2019, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisWinterMC16> (visited on 11/13/2019).
- [165] “Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV,” tech. rep.

- ATL-PHYS-PUB-2015-015, CERN, 2015, URL: <https://cds.cern.ch/record/2037613>.
- [166] K. Cranmer et al., *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, (2012), URL: <https://cds.cern.ch/record/1456844>.
- [167] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: [physics/0306116](https://arxiv.org/abs/physics/0306116) [[physics.data-an](https://arxiv.org/archive/physics)].
- [168] L. Moneta et al., *The RooStats Project*, **PoS ACAT2010 (2010) 057**, arXiv: [1009.1003](https://arxiv.org/abs/1009.1003) [[physics.data-an](https://arxiv.org/archive/physics)].
- [169] LHC Higgs Cross Section Working Group, *LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle*, (2012), arXiv: [1209.0040](https://arxiv.org/abs/1209.0040) [[hep-ph](https://arxiv.org/archive/hep)].
- [170] S Heinemeyer, C Mariotti, et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*, CERN Yellow Reports: Monographs, 2013, URL: <http://cds.cern.ch/record/1559921>.
- [171] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, **Eur. Phys. J. C 72 (2012) 1849**, arXiv: [1110.1530](https://arxiv.org/abs/1110.1530) [[hep-ex](https://arxiv.org/archive/hep)].
- [172] R Achenbach et al., *The ATLAS Level-1 Calorimeter Trigger*, **Journal of Instrumentation 3 (2008) P03001**, URL: <https://doi.org/10.1088%2F1748-0221%2F03%2F03%2Fp03001>.
- [173] F. Rosenblatt, *The Perceptron: A Probabilistic Model for Information Storage and Organization in The Brain*, **Psychological Review (1958) 65**.
- [174] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, **Mathematics of Control, Signals, and Systems (MCSS) 2 (1989) 303**, ISSN: 0932-4194, URL: <http://dx.doi.org/10.1007/BF02551274>.
- [175] S. Ruder, *An overview of gradient descent optimization algorithms*, (2017), arXiv: [1609.04747](https://arxiv.org/abs/1609.04747) [[cs.LG](https://arxiv.org/archive/cs)].

- [176] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, arXiv: 1005.4568 [hep-ex].
- [177] K. He et al., *Deep Residual Learning for Image Recognition*, (2015), arXiv: 1512.03385 [cs.CV].
- [178] S. H. Rezatofighi et al., *DeepSetNet: Predicting Sets with Deep Neural Networks*, (2017), arXiv: 1611.08998 [cs.CV].
- [179] A. Paszke et al., *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, (2019) 8024, ed. by H. Wallach et al., URL: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- [180] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, (2017), arXiv: 1412.6980 [cs.LG].

תקציר

חיבור זה מציג חיפוש אחר חלקיקי בוזון היגס במודל הסטנדרטי ומעבר לו בהתנגשויות פרוטונים במאיץ ה-LHC תוך שימוש במאורעות מגלאי ATLAS. התזה מתארת מחקר בשלוש חזיתות: הגדלת רגישות במדידות של פיזיקה במסגרת המודל הסטנדרטי של חלקיקים, חיפוש אחר סוג חדש של חלקיקים שחורג מהמודל הסטנדרטי והגדרת טכניקות לזיהוי חלקיקים לשימוש בריצות הבאות של המאיץ. החיפוש במסגרת המודל הסטנדרטי התמקד בדעיכת חלקיק בוזון ההיגס לזוג חלקיקי קווארק מסוג קסם. המדידה הראשונה אי פעם של דעיכה זו הניבה גבול עליון של פי 110 מהמודל הסטנדרטי ובמחקר זה הגבול העליון החזוי הינו פי 26 המהווה חסם איכותי יותר לאפקטים של פיזיקה חדשה בצורת דעיכה זו. החיפוש אחר פיזיקה חדשה התמקד בחלקיק בוזון ההיגס הטעון אשר דועך לזוג חלקיקי קווארק מסוג למעלה ולמטה. חיפוש זה השתמש בשיטות של למידת מכונה לצורך סיווג הנתונים והשמת חסם עליון. כל תוצאות המחקר היו עקביות עם המודל הסטנדרטי ללא סטיה משמעותית. החיפושים התבצעו עם מאורעות בהם אנרגיית מרכז המסה הייתה 13TeV והארה של 139fb^{-1} . בנוסף, חיבור זה מסכם מאמץ של ביסוס טכניקות חדשניות עבור זיהוי חלקיק הטאו בקלורימטר בגלאי ATLAS. במהלך מחקר זה נבנו מודלים של למידת מכונה עמוקה המבוססים על התמרת הפקדות האנרגיה של תוצרי דעיכת החלקיק לכדי קבוצה מתמטית. שיטות חדשות אלו הראו שיפור משמעותי ביחס לשיטות הקיימות.



School of Physics and Astronomy בית הספר לפיזיקה ולאסטרונומיה
The Raymond and Beverly Sackler הפקולטה למדעים מדויקים
Faculty of Exact Sciences ע"ש ריימונד ובברלי סאקלר
Tel Aviv University אוניברסיטת תל אביב

תזת מחקר עבור התואר "דוקטור לפילוסופיה"

חיפוש אחר חלקיקי בוזון היגס במודל הסטנדרטי ובפיזיקה חדשה

הוגש לסנאט אוניברסיטת תל אביב על ידי

הדר כהן

עבודה זו בוצעה בהנחיתו של פרופ. ארז עציון