Constraining the Top Quark Yukawa Coupling in Single-Top Higgs Production

Thomas Michael Carter

Doctor of Philosophy
The University of Edinburgh
May 2023

Abstract

The work presented in this thesis consists of the first search for the associated production of a top quark and a Higgs boson (tH) with the ATLAS detector at the LHC. It also establishes an important contribution to the fast simulation of the ATLAS detector.

A novel analysis of the tH process has been developed using the di-photon ($H \rightarrow \gamma\gamma$) decay channel. The analysis is designed to simultaneously optimise sensitivity to the Standard Model (SM) tH production, and to the scenario in which the top quark Yukawa coupling to the Higgs boson has a negative sign. The ratio of the measured and predicted cross-sections is reported as:

$$\sigma_{tH}/\sigma_{tH}^{\mathrm{SM}} = 2.6^{+4.2}_{-3.2} = 2.6^{+3.7}_{-2.9}(\mathrm{stat.})^{+1.6}_{-1.1}(\mathrm{syst.}). \tag{1}$$

An upper limit on the tH cross-section is set at 10 times the SM expectation, and the negative values of the top quark Yukawa coupling are excluded at 2.2σ. At the time of submission, these results apply the most stringent constraints on the tH cross-section and the top quark Yukawa coupling.

The muon punch-through component of the ATLAS fast simulation tool, AtlFastIII (AF3), has been developed. AF3 improves the speed of the ATLAS detector simulation by a factor of 5, facilitating the production of large sets of simulated data that aid in future ATLAS physics results.

Muon punch-through occurs when particles created in electromagnetic showers in the ATLAS calorimeters are not fully contained. This can result in signatures that may be mis-reconstructed as muons. Accurate modelling of this effect is essential for understanding muon fakes and for hadronic jet calibration.

Lay Summary

The Standard Model of particle physics is a theory of the fundamental particles of our Universe and their interactions. It describes nearly all known physical observations at the smallest energy scale, and has stood up to over 50 years of experimental scrutiny. Prior to 2012 the final missing component yet to be discovered was the Higgs boson.

The Large Hadron Collider (LHC) was constructed with the primary motivation of discovering the Higgs boson. The LHC is a superconducting particle accelerator that accelerates protons to nearly the speed of light, and collides them head-on. The proton collisions provide the conditions necessary to produce rare sub-atomic particles. Observing these particles using detectors gives us a glimpse into the sub-atomic world, and importantly helps us to understand the nature of the Universe.

In 2012, the ATLAS and CMS experiments at the LHC announced the discovery of the Higgs boson, completing the Standard Model. The discovery led to the Nobel Prize being awarded to Peter W. Higgs and François Englert, for "the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles". The profound role that the Higgs boson plays in the Standard Model led to it receiving the nickname the "god particle".

With the Higgs boson discovered, efforts have shifted focus towards understanding it further and measuring its properties. This thesis presents a novel analysis of the properties of the Higgs boson that contributes to these efforts. The result is the most sensitive measurement on the production of a Higgs boson in association with a single top quark, and the most sensitive constraint on the interaction between the Higgs boson and the top quark. These results contribute to the further understanding of the SM and the Higgs boson.

In particle physics, the fundamental processes and their interactions with the detector are simulated to obtain a full understanding. The expectation of the simulations are compared with observations and conclusions are made about nature. However, these simulations are computationally expensive. Therefore, a fast and computationally cheap alternative is in development. A significant contribution is made towards this alternative; a component that models the muon punch-through effect in the ATLAS detector. The fast simulation toolkit will enable the future simulation program of the ATLAS experiment.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Parts of this work have been published in [1], [2].

A significant contribution was made to [1]. Therefore, the publication has been replicated for the readers understanding.

(*Thomas Michael Carter, May 2023*)

Acknowledgements

First, and foremost I would like to thank Liza Mijovic for her unwavering support throughout my Ph.D. She was the pioneer of the idea behind my analysis work. She always found the time to support me, answering any questions no matter how naive, and occasionally working into the late hours with me to achieve results. Without her, I would not be the physicist I am today.

Thank you to Victoria Martin for her invaluable advice, continuous support and patience during my PhD. Her extensive knowledge and experience have guided me through my academic study, and her encouragement to attend and talk at conferences and events has helped me develop as a young researcher.

Thank you to Hasib Ahmed for his introduction to the world of fast simulation and generative modelling. I have enjoyed the field so much that I decided to pursue a professional career in it following my PhD.

I also would like to thank my colleagues at CERN and around the world, for their patience and trust in my ability to deliver, and for the opportunities they gave me.

Thank you to Vin Dhanak and Sergey Burdin from The University of Liverpool. They both inspired and encouraged me to pursue a Ph.D.

I would like to thank, Fytos and Sara. They have been the best possible friends and colleagues to have on this journey, to share ideas, troubles, and laughs with.

To my parents, for everything you do for me, for always encouraging me to push myself, and for inspiring me to pursue science from a young age. You have always been there for me, supportive and reliable, and a great sounding board for making big decisions in life. I couldn't have wished for more.

Last, but by no means least, a very special thank you goes to my partner Alice. She has been by my side throughout the whole experience and endured all the ups and downs. Without you, this whole endeavour would not have been possible. Thank you for all the amazing adventures you have shared with me in Scotland. I can't wait to start our new exciting chapter in life together.

Contents

I Foundations 3

Acronyms

p-p proton-proton.

AF2 AtlFastII.

AF3 AtlFastIII.

AUC Area Under Curve.

BCE Binary Cross-Entropy.

BDT Boosted Decision Tree.

BSM Beyond the Standard Model.

CDF cumulative density function.

CERN The European Organisation for Nuclear Research ("Conseil Européen pour la Recherche Nucléaire").

DSCB *double-sided* Crystal Ball.

EM electromagnetic.

EW Electroweak.

ID inner detector.

IP interaction point.

LAr Liquid-Argon.

LEP Large Electron Positron.

LHC Large Hadron Collider.

LO leading order.

MC Monte-Carlo.

MET Missing Transverse Momentum.

ML Machine Learning.

MPT muon punch through.

MS Muon Spectrometer.

NLO next-to-leading order.

NN Neural Network.

NSW New Small Wheel.

NTI Non-Tight-Isolated.

PCA Principal Component Analysis.

PCBtag pseudo-continuous b-tagging.

PDF parton distribution functions.

pdf probability density function.

PDGId Particle Data Group particle numbering scheme.

QCD Quantum Chromodynamics.

QED Quantum Electrodynamics.

QFT Quantum Field Theory.

SM Standard Model.

STXS Simplified Template Cross Sections.

TI Tight-Isolated.

Glossary

barrel The central region of the ATLAS detector.

barycentre The point at the centre of a system; an average point, weighted according to mass or other attribute.

bending plane The plane in which a charged particle trajectory is bent in a magnetic field.

end-cap The forward regions of the ATLAS detector.

hard-process The non-elastic scattering of particles in a proton-proton (p-p) collision.

hard-scatter The non-elastic scattering of particles in a p-p collision..

jet A reconstructed object that acts as a proxy for the underlying hadron.

muon punch-through The process by which particles created in showers exit out of the back of the ATLAS calorimeters.

muon segment A reconstructed quantity from energy deposits in the Muon Spectrometer (MS).

non-elastic A collision that results in the annihilation of particles.

shower A cascade of secondary particles produced as the result of a high-energy particle interacting with dense matter.

List of Figures

List of Tables

Chapter 1

Introduction

The final missing piece of Standard Model (SM) of particle physics was completed when a Higgs like particle was discovered in 2012 by the ATLAS and CMS experiments. Since discovery, experimental efforts have been geared towards understanding the particle in detail, and discerning whether it really is the SM Higgs boson. Part of the work completed for this thesis contributes to these efforts using the ATLAS experiment, through a novel analysis of the rare single-top Higgs (tH) process in the Higgs diphoton decay channel. At the time of writing this thesis, the analysis set the most stringent constraints on the tH cross-section and the top quark Yukawa coupling strength. The work advances the experimental knowledge of the Higgs boson, and lays the foundations for future analyses of the tH process using the ATLAS experiment.

Throughout the Run 3 operation of the Large Hadron Collider (LHC) and beyond, the amount of data collected by the ATLAS experiment will significantly increase. In this period, Higgs physics will move into the precision regime. However, in order to analyse these large datasets to conduct precise measurements, enormous sets of simulated data will be required. This presents an issue, as simulating data is very computationally expensive and time consuming. To address the issue, a fast and cheap alternative simulation solution has been developed for the ATLAS experiment, AtlFastIII (AF3). AF3 uses a parameterised approach to simulate the response of the ATLAS calorimeters, which are traditionally highly computationally expensive. Part of the work completed for this thesis contributes to AF3, developing a fast and accurate parameterised simulation of themuon punch through (MPT) effect. This work will play an important role in the future simulation program of the ATLAS experiment.

Part I presents the foundations on which the work in this thesis builds upon. It also gives the reader the required knowledge to understand the results and techniques that are presented throughout. Chapter 2 presents the particles and forces of the SM, and how they can be described using Quantum Field Theory (QFT). The largrangian formalism of QFT is then used to introduce the Higgs mechanism and how it gives the other SM particles mass. Properties of the Higgs boson are then discussed, including the dominant production and decay modes of the Higgs boson at the LHC. Chapter 3 gives an overview of the LHC and the ATLAS detector. Chapter 4 details the physical processes that happen inside electromagnetic (EM) and hadronic showers, this information is essential for understanding the MPT effect detailed in Part II. Finally, Chapter 5 gives a brief overview of neural networks which are used in the work completed for this thesis. The chapter covers the basics of how neural networks work, and some key aspects one should consider when using them.

Part II details the MPT simulation that was developed for the ATLAS fast simulation toolkit, AF3. Chapter 6 gives an overview of the simulation scheme used by the ATLAS experiment, and discusses how AF3 is able to achieve fast and accurate simulation of the ATLAS calorimeters. Chapter 7 then gives a detailed account of the development of the MPT simulation for AF3, covering the techniques that were used in its development, how it integrates within AF3, and results showing its modelling performance.

Part III presents the novel analysis of the tH production process in the diphoton decay channel. Chapter 8 gives an overview of the encompassing analysis in which measurements of the Higgs boson properties in the $\mathrm{H} \to \gamma\gamma$ decay channel are reported. Chapter 9 then presents the novel analysis of the tH process which uses three neural network models to categorise events. Using these neural network models, the analysis is able to simultaneously optimise for sensitivity to the tH SM cross-section measurement, and the sign of the top quark Yukawa coupling strength.

Part I

Foundations

Chapter 2

The Standard Model

The Standard Model (SM) is arguably the most well-tested effective theory to ever exist. It describes the physical interactions observed in nature at the smallest accessible energy scale. The SM has been constructed over decades of theoretical extension and experimental verification. Remarkably, the SM can describe almost all experimental results in particle physics, making it one of the great triumphs of modern science.

2.1 The Fundamental Particles and Forces

The SM describes the wide variety of physical phenomena that are observed at the sub-atomic scale, with a set of fundamental particles and their interactions. These fundamental particles are divided into two distinct groups: the fermions which make up the matter content of our Universe, and the gauge bosons that act as force carriers for interactions between fermions (Figure 2.1). The final component is the Higgs boson, which plays a special role in the SM as it provides the mechanism by which all other particles acquire mass (Section 2.3).

The fermions consist of quarks and leptons, and are categorised into three generations of increasing mass. The electron ($e^{\pm}$), muon ($\mu^{\pm}$), and tau ($\tau^{\pm}$) leptons each have a charge of $\pm e$, and a spin of $\frac{1}{2}$. For each of these leptons there exists an associated neutrino: the electron neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ). The masses of the neutrinos, and the hierarchy of their masses are still experimentally undetermined. The final component to the matter fermions are the quarks. For each generation, there is an up-type and down-type quark with non-integer charge $+\frac{2}{3}$ and $-\frac{1}{3}$ respectively. For each fermion in the SM, there exists an antiparticle with opposite charge.

There are four spin-1 gauge bosons within the standard model, these are: the photon (γ), the gluon (g), the W boson ($W^{\pm}$), and the neutral Z boson (Z). Each gauge boson is s mediator of one of the fundamental forces of nature. The electromagnetic force is mediated by the photon, the strong force by the gluon, and the weak force is mediated by the W and Z bosons. The particles of the SM are shown in Figure 2.1.

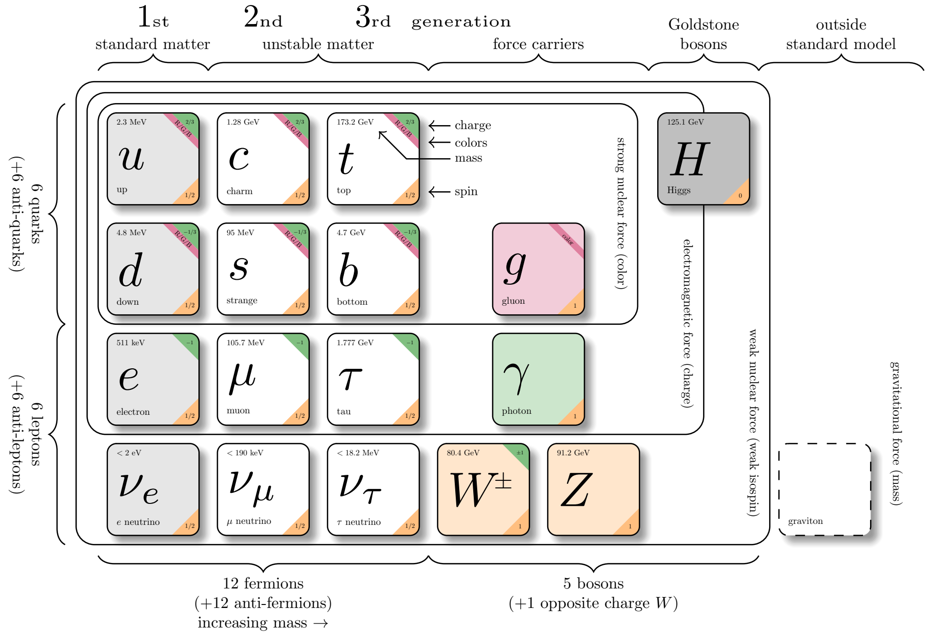


Figure 2.1 *The particles of the standard model. Showing the three generations of fermions, the four force-carrying gauge bosons, and the Higgs boson. The quarks are shown in their mass eigenstates, and the leptons are shown in their flavour eigenstates. Taken from [3].*

2.2 Quantum Field Theories

The SM particles and their interactions are described mathematically using Quantum Field Theory (QFT). QFT is a field theory that incorporates quantum mechanics and the principles of the theory of relativity. The SM QFT combines the strong interaction and the Electroweak (EW) interaction. The strong interaction is described by Quantum Chromodynamics (QCD), and the EW interaction is described by the unification of Quantum Electrodynamics (QED) and the weak force.

In QFT, particles are described as excitations of a quantum field $\phi(x^\mu)$ that is a continuous function of space-time x^μ, where $\mu = 0, 1, 2, 3$ labels the four dimensions of space-time. Throughout the rest of this section $\phi(x^\mu)$ may be referred to as ϕ, where the dependency on space-time is implied.

In classical mechanics the dynamics of a physical system can be described by the Lagrangian. Similarly, the dynamics of a quantum field can be described by a Lagrangian density. The Lagrangian density $\mathcal{L}$ is related to the Lagrangian L by

$$L = \int \mathcal{L} \, \mathrm{d}^3\mathbf{x}. \tag{2.1}$$

The Lagrangian density $\mathcal{L}$ will be referred to as simply the Lagrangian in the following text for simplicity.

Using the principal of least action, the Euler-Lagrange equation applied to fields can be shown to be

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0, \tag{2.2}$$

where ∂_μ is the partial derivative with respect to the space-time coordinates μ. Solving the Euler-Lagrange equations for a given Lagrangian will result in the equations of motion of the system.

As an initial example, we can look at the simple case for spin-0 scalar particles, which are described as excitations of the scalar field ϕ that satisfies the Klein-Gordon equation. The Lagrangian for the free non-interacting scalar field is given by,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2 \tag{2.3}$$

Solving the Euler-Lagrange equations (Equation 2.1) for this Lagrangian we then arrive at the Klein-Gordon equation that describes the evolution of the system,

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0. \tag{2.4}$$

Noether's theorem states that every differentiable symmetry of the action of a physical system with conservation forces has an associated conservation law [4]. Interpreted in terms of the Lagrangian describing a physical system, Noether's theorem states that if the Lagrangian is invariant under this symmetry then there must exist a conservation law. In the SM, symmetries of nature are incorporated into the Lagrangian. The resulting conservation laws correspond to the conserved currents of the force-carrying particles.

The SM Lagrangian can be written as a sum of the QCD Lagrangian and the EW Lagrangian,

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW}. \tag{2.5}$$

The conserved currents in the SM are described by the local gauge transformations belonging to the symmetry group

$$SU(3) \otimes SU(2)_L \otimes U(1)_Y. \tag{2.6}$$

Where the SU(3) symmetry group corresponds to the conservation of colour charge in the strong interaction, the SU(2) group corresponds to the conservation of weak hypercharge in the EW interaction, and U(1) corresponds to the conservation of weak hypercharge in the electromagnetic interaction. In the following sections, we cover these symmetries in more detail, and how they result in the interaction terms between the matter fermions and the force-carrying gauge bosons.

2.2.1 Quantum Electrodynamics

QED is the QFT describing the electromagnetic interaction between the spin-$\frac{1}{2}$ fermions and the photon. The equations of motion for a free fermion are described by the Dirac equation,

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \tag{2.7}$$

where ψ is a Dirac spinor state with four components, γ^μ are the Dirac matrices,

∂_μ is the 4-dimensional differential operator and m is the mass of the fermion.

The corresponding Lagrangian of the free fermion is,

$$\mathcal{L} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi, \tag{2.8}$$

where $\bar{\psi}$ is the adjoint spinor, defined as $\bar{\psi} = \psi^\dagger\gamma^0$.

Suppose we introduce a local U(1) phase transformation $\psi(x) \rightarrow \psi'(x) = \hat{U}\psi(x) = e^{iq\theta(x)}\psi(x)$, where q is electric charge. Under this transformation, we expect the equations of motion to stay the same, reflecting the fundamental symmetry of the universe that we observe. If we substitute this local phase transformation into the free-particle Lagrangian equation we get the following:

$$\begin{aligned} \mathcal{L} \rightarrow \mathcal{L}' = ie^{-iq\theta(x)}\bar{\psi}\gamma^\mu[e^{i\theta(x)}\partial_\mu\psi + iq(\partial_\mu\theta(x))e^{iq\theta(x)}\psi] - me^{-iq\theta(x)}\bar{\psi}e^{iq\theta(x)}\psi \\ = \mathcal{L} - q\bar{\psi}\gamma^\mu(\partial_\mu\theta(x))\psi \end{aligned} \tag{2.9}$$

We can see that the new Lagrangian of the system differs by the term $q\bar{\psi}\gamma^\mu(\partial_\mu\theta(x))\psi$. Therefore, the Lagrangian for a free-particle Dirac field is not invariant under a U(1) local phase transformation. To establish the required invariance we replace the derivative ∂_μ with the covariant derivative,

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + iqA_\mu \tag{2.10}$$

where A_μ is a new field. The required invariance is then satisfied if the new field transforms as $A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu\theta(x)$.

Hence the modified gauge invariant Lagrangian for a spin-half fermion takes the form,

$$\mathcal{L} = i\bar{\psi}\gamma_\mu\partial_\mu\psi - m\bar{\psi}\psi - q\bar{\psi}\gamma^\mu A_\mu\psi \tag{2.11}$$

where the extra term $-q\bar{\psi}\gamma^\mu A_\mu\psi$ represents the interaction between the free fermion field and the new field A_μ which can be identified as the photon. This is a profound statement, by requiring that physics is invariant under local U(1) phase transformations we can arrive at the interaction between the fermionic matter and the photon.

Introducing the kinetic term for the photon $-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$, the Lagrangian for QED becomes,

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi - q\bar{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{2.12}$$

where $F_{\mu\nu}$ is the electromagnetic field strength tensor defined as $F_{\mu\nu} = \partial^{\mu}A^{\nu} - A^{\nu}\partial^{\mu}$.

2.2.2 Quantum Chromodynamics

QCD is the QFT of the strong force, where the associated SU(3) symmetry leads to the conservation of colour charge. The Lagrangian for a free quark can be written as

$$\mathcal{L} = i\bar{\psi}_j\gamma^{\mu}\partial_{\mu}\psi_j - m\bar{\psi}_j\psi_j, \tag{2.13}$$

where the index j represents a sum over all quark flavours.

Following the same methodology as for QED, but this time invoking the principle of gauge invariance for the SU(3) symmetry, we arrive at the QCD Lagrangian

$$\mathcal{L}_{QCD} = i\bar{\psi}_j\gamma^{\mu}\partial_{\mu}\psi_j - m\bar{\psi}_j\psi_j - g_S T^{a}\gamma^{\mu}G^{a}_{\mu}\psi_j - \frac{1}{4}G^{a}_{\mu\nu}G^{\mu\nu}_{a} \tag{2.14}$$

Where T^{α} are the generators of the SU(3) symmetry group and are related to the Gell-Mann matrices by

$$T^{\alpha} = \frac{1}{2}\lambda^{\alpha}, \tag{2.15}$$

and g_S is the strong coupling constant. Because the generators of the SU(3) symmetry group are a set of 3x3 matrices, the fermion wavefunction must now include three degrees of freedom. These degrees of freedom are labelled 'colour', where 'red', 'blue' and 'green' label each state. The local phase transformation in the SU(3) group corresponds to 'rotating' states in this colour space. The local gauge invariance can be achieved by introducing eight new fields G^a_μ, where the index a runs from 1 to 8, and each field corresponds to one of the eight generators of the SU(3) symmetry. The term $-g_S T^a \gamma^\mu G^a_\mu \psi_j$ in the QCD Lagrangian describes the interaction between the quark fields and the new gluon fields. The term $-\frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a$ contains the kinetic term for new the free gluon fields. The gauge invariance holds provided that the new fields transform as,

$$G^k_\mu \to G^{k\prime}_\mu = G^k_\mu - \partial_\mu \alpha_k - g_S f_{ijk} \alpha_i G^j_\mu. \tag{2.16}$$

The final term arises because the generators of the SU(3) group do not commute, and the f_{ijk} are the structure constants of the group. The presence of this extra term gives rise to gluon self-interactions in the SM.

2.2.3 Electroweak

In the 1960s, Glashow, Salam and Weinberg developed a unified model of the electromagnetic (EM) and weak forces. The model remarkably predicts the existence of the neutral Z boson. In the unified EW model the U(1) gauge symmetry of QED is replaced with a new $U(1)_Y$ local gauge symmetry, and the weak interaction is characterised by the symmetries under $SU(2)_L$. Two new quantum numbers are associated with each fermion: weak isospin T (the conserved current of the $SU(2)_L$ group), and weak hypercharge Y (the conserved current of the $U(1)_Y$ group). The electric charge Q is related to the weak isospin and weak hypercharge by,

$$Q = T^3 + \frac{Y}{2}, \tag{2.17}$$

where T^3 is the third component of weak isospin.

Only left-handed chiral fermion states and right-handed chiral anti-fermion states participate in the weak interaction. These are defined as,

$$\psi_L = \frac{1-\gamma^5}{2}\psi, \quad \text{and} \quad \psi_R = \frac{1+\gamma^5}{2}\psi, \tag{2.18}$$

where $\gamma^5 = i\gamma^1\gamma^0\gamma^2\gamma^3$. Right (left) handed states are defined as having spin in the same (opposite) direction to their momentum.

Following similar methodology used for QED and QCD under the unified $SU(1)_Y \otimes SU(2)_L$ group we introduce four new vector fields to maintain gauge invariance, B_μ and W^a_μ where $a = 1, 2, 3$. The resulting Lagrangian is,

$$\begin{aligned} \mathcal{L}_{EW} = i\bar{\psi}_j\gamma^\mu\partial_\mu\psi_j - m\bar{\psi}_j\psi_j - i\frac{g'}{2}Y\gamma^\mu B_\mu\psi_j - i\frac{g}{2}T^a\gamma^\mu W^a_\mu\psi_j \\ -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a \end{aligned} \tag{2.19}$$

where Y is the weak hypercharge, T^a is the weak isospin, $-i\frac{g'}{2}Y\gamma^\mu B_\mu\psi_j$ is the interaction term between the new field B_μ and left and right-handed fermions with a coupling strength g', and $-i\frac{g}{2}T^a\gamma^\mu W^a_\mu\psi_j$ is the interaction term between the new fields W^a_μ and left (right) handed (anti-)fermions.

The kinetic terms for each of the new fields are defined by,

$$\begin{gathered} B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \\ \text{and} \\ W_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\epsilon^{abc}W^b_\mu W^c_\nu. \end{gathered} \tag{2.20}$$

The physical W bosons can be identified as the linear combinations of the first two components of W^a_μ,

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp iW^2_\mu\right). \tag{2.21}$$

The photon and Z boson are then written as linear combinations of the B_μ and W^3_μ fields of the EW interaction,

$$\begin{aligned} A_\mu &= +B_\mu\cos\theta_W + W^3_\mu\sin\theta_W \\ Z_\mu &= -B_\mu\sin\theta_W + W^3_\mu\cos\theta_W, \end{aligned} \tag{2.22}$$

where θ_W is the weak mixing angle. The mixing of the neutral fields of the $U(1)_Y$ and $SU(2)_L$ gauge symmetries is a natural result of the Higgs mechanism.

Finally, the EW Lagrangian can be simplified by introducing the relevant covariant derivative $D_\mu = \partial_\mu - i\frac{g'}{2}YB_\mu \; - i\frac{g}{2}T^a W^a_\mu$ to be,

$$\mathcal{L}_{EW} = i\bar{\psi}_j \gamma^\mu D_\mu \psi_j - m\bar{\psi}_j \psi_j - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a. \tag{2.23}$$

2.3 The Higgs Mechanism

So far we have introduced the Lagrangian formalism of the SM, including QCD and the unification of the electromagnetic and weak forces. However, in this description, there is no mechanism to give the $W^\pm$ and Z bosons mass. Simply adding mass terms to the Lagrangian violates the required gauge invariance of the theory. Instead, we need to give these bosons mass using a more subtle approach.

We introduce a weak isospin doublet of two complex scalar fields,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.24}$$

where, the upper and lower component of the doublet differ by one unit of charge. The Lagrangian for this doublet of complex scalar fields is,

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi), \tag{2.25}$$

with the Higgs potential,

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{2.26}$$

For the potential to have a minimum, λ must be greater than 0, however, there is no such restriction for μ. If $\mu^2 > 0$ then the potential is has a single minima at $|\phi|^2 = 0$. However, if $\mu^2 < 0$ then the potential has a set of infinite minima that satisfy,

$$\phi^\dagger \phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}. \tag{2.27}$$

The Higgs potential can be visualised in the complex plane in Figure 2.2. We can see that the vacuum state (minima of the potential) exists at an infinite number of points along a circle in the real and imaginary plane.

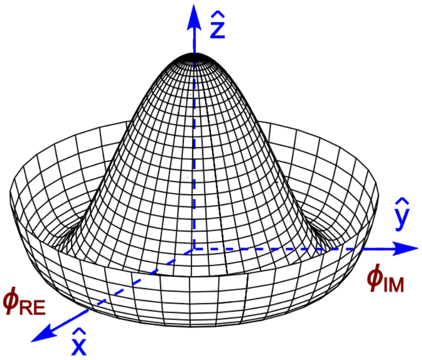


Figure 2.2 *The* $V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$ *Higgs potential for the complex scalar doublet where* $\mu^2 < 0$*. Taken from [5].*

The physical vacuum state will correspond to a particular point on this surface, breaking the $U(1)_Y \otimes SU(2)_L$ symmetry. Typically the minima is chosen so that $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = v$, then the vacuum state becomes,

$$\langle 0|\phi|0|0|\phi|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.28}$$

The fields can then be expanded about this minimum by writing,

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ v + \eta(x) + i\phi_4(x) \end{pmatrix}. \tag{2.29}$$

As a result of this spontaneous symmetry breaking, terms will be added to the Lagrangian describing a massive scalar particle and three massless *Goldstone bosons*. The three Goldstone bosons give the longitudinal degrees of freedom of the $W^\pm$ and Z bosons. The derivation will not be covered here, but these Goldstone bosons can be eliminated from the Lagrangian by making an appropriate gauge transformation. In this unitary gauge the Higgs doublet can be written as,

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.30}$$

The resulting Lagrangian describes a real massive scalar boson and three massive gauge bosons. The real scalar boson is the Higgs boson represented by the physical Higgs scalar field $H(x)$.

The terms in the associated Lagrangian that give masses to the gauge bosons are,

$$\begin{gathered}\frac{1}{8}v^2g_W^2(W_\mu^{(1)}W^{(1)\mu}+W_\mu^{(2)}W^{(2)\mu})\\ \text{and}\\ \frac{1}{8}v^2(g_WW_\mu^{(3)}+g'B_\mu)(g_WW^{(3)\mu}-g'B_\mu).\end{gathered} \tag{2.31}$$

In the Lagrangian, the mass terms for the $W^{(1)}$ and $W^{(2)}$ spin-1 fields appear as

$$\frac{1}{2}m_W^2W_\mu^{(1)}W^{(1)\mu} \quad \text{and} \quad \frac{1}{2}m_W^2W_\mu^{(2)}W^{(2)\mu}. \tag{2.32}$$

We can therefore determine the mass of the W boson to be,

$$m_W=\frac{1}{2}g_Wv, \tag{2.33}$$

where g_W is the $SU(2)_L$ gauge interaction strength, and v is the vacuum expectation value of the Higgs field (Equation 2.28).

The mass term associated with the neutral $W^{(3)}$ and B fields can be written as,

$$\begin{aligned}\frac{v^2}{8}(g_WW_\mu^{(3)}+g'B_\mu)(g_WW^{(3)\mu}-g'B_\mu) &= \frac{v^2}{8}\begin{pmatrix}W_\mu^{(3)} & B_\mu\end{pmatrix}\begin{pmatrix}g_W^2 & -g_Wg' \\ -g_Wg' & g'^2\end{pmatrix}\begin{pmatrix}W_\mu^{(3)} \\ B_\mu\end{pmatrix}\\ &= \frac{v^2}{8}\begin{pmatrix}W_\mu^{(3)} & B_\mu\end{pmatrix}\mathbf{M}\begin{pmatrix}W_\mu^{(3)} \\ B_\mu\end{pmatrix}\end{aligned} \tag{2.34}$$

where $\mathbf{M}$ is the non-diagonal mass matrix. The physical gauge bosons fields correspond to the eigenvectors of $\mathbf{M}$ and their masses are then given by the eigenvalues of $\mathbf{M}$, where we find

$$\begin{gathered}A_\mu=\frac{g'W_\mu^{(3)}+g_WB_\mu}{\sqrt{g_W^2+g'^2}} \quad \text{with} \quad m_A=0\\ \text{and}\\ Z_\mu=\frac{gW_\mu^{(3)}-g'_WB_\mu}{\sqrt{g_W^2+g'^2}} \quad \text{with} \quad m_Z=\frac{1}{2}v\sqrt{g_W^2+g'^2}.\end{gathered} \tag{2.35}$$

We see that the physical photon and Z boson fields are mixtures of the massless bosons associated with the $U(1)_Y$ and $SU(2)_L$ local gauge symmetries. We see that the Z boson has acquired mass through the Higgs mechanism, and the photon has remained massless. If we write the ratio of the couplings of the $U(1)_Y$ and $SU(2)_L$ gauge symmetries as

$$\frac{g'}{g_W} = \tan\theta_W \tag{2.36}$$

then we arrive at the relationship in Equation 2.22 between the physical and underlying fields.

The GSW model predicts that,

$$\frac{m_W}{m_Z} = \cos\theta_W. \tag{2.37}$$

meaning precise measurements of the weak mixing angle provide a stringent test of the SM, EW unification and the Higgs mechanism. In the complete model there are four parameters, the $U(1)_Y \otimes SU(2)_L$ couplings g' and g_W, and the free parameters of the Higgs potential μ and λ. By using the relation $m_W = \frac{1}{2} g_W v$, and the measured values of g_W and m_W, the vacuum expectation value of the Higgs potential is found to be

$$v = 246\ \text{GeV} \tag{2.38}$$

The parameter λ can be found by measuring the mass of the Higgs boson.

The coupling of the Higgs boson to the gauge bosons is proportional to the gauge boson mass. For the W boson the coupling strength is

$$g_{HWW} = \frac{1}{2} g_W^2 v = g_W m_W, \tag{2.39}$$

and likewise for the Z boson coupling.

2.4 Fermion Masses

The Higgs mechanism for the spontaneous symmetry breaking of the $U(1)_Y \otimes SU(2)_L$ gauge group of the SM is able to generate masses for the W and Z bosons. Remarkably, it can also be used to generate the masses of the fermions in the SM. In order to respect the $U(1)_Y \otimes SU(2)_L$ symmetry, the fermion mass term in the Lagrangian has to take the following form:

$$\mathcal{L} = -y_f(\bar{\psi_L}\phi\psi_R + \bar{\psi_R}\phi^\dagger\psi_L) \tag{2.40}$$

where y_f is the Yukawa coupling of the fermion f, ψ_L is the left-handed chiral doublet, ψ_R is the right-handed chiral singlet, and ϕ is the complex Higgs doublet.

More specifically for the electron field, we can write,

$$\mathcal{L}_e = -y_e \left[\begin{pmatrix} \bar{\nu_e} & \bar{e} \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e_R} \begin{pmatrix} \phi^{+*} & \phi^{0*} \end{pmatrix} \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right], \tag{2.41}$$

where y_e is the Yukawa coupling between the electron and the Higgs field. After spontaneous symmetry breaking the Higgs doublet in the unitary gauge is,

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.42}$$

and so the Lagrangian term becomes,

$$\mathcal{L}_e = -\frac{y_e}{\sqrt{2}} v(\bar{e_L}e_R + \bar{e_R}e_L) - \frac{y_e}{\sqrt{2}} H(\bar{e_L}e_R + \bar{e_R}e_L). \tag{2.43}$$

The first term can be identified as the mass-giving term for the electron, and the second term as the interaction between the electron field and the scalar Higgs field. Therefore the electron Yukawa coupling with the electron mass is related by,

$$y_e = \sqrt{2}\frac{m_e}{v}. \tag{2.44}$$

So far we have seen that the Higgs doublet can generate masses for the fermion in the lower component of the $SU(2)_L$ doublet. This presents an issue when we consider doublets of up and down type quarks. This can be remedied by constructing the conjugate of the Higgs doublet ϕ_c, which will not be shown here.

We find that for all Dirac fermions, gauge invariant mass terms can be constructed from either

$$\mathcal{L} = -y_f \left[\bar{L}\phi R + (\bar{L}\phi R)^\dagger\right] \quad \text{or} \quad y_f \left[\bar{L}\phi_c R + (\bar{L}\phi_c R)^\dagger\right], \tag{2.45}$$

where L and R represent the left and right-handed fermion doublet or singlet. These terms give rise to both the mass of the fermions and the interactions of the fermions with the Higgs boson. The Yukawa couplings of the fermions to the Higgs field are given by,

$$y_f = \sqrt{2}\frac{m_f}{v}. \tag{2.46}$$

2.5 Connecting Theory to Predictions

Ultimately the theories introduced in the previous sections of this chapter are used to make predictions of physical processes in the SM. Using the ATLAS experiment the predictions are tested experimentally. This section gives a brief overview of how predictions are made about collisions at the LHC.

The SM QFT is used to calculate cross-sections which describe the probability that certain particles are produced in a collision. The cross-section calculation of proton collisions at the Large Hadron Collider (LHC) is split into two components: the scattering cross-section of the interacting partons (calculated using perturbative QCD), and the cross-section of obtaining the incoming partons from the original colliding protons (modelled using parton distribution functions (PDF)s). The calculation can be split in this way because of the QCD factorisation theorem [6]. As an example, the calculation of the cross-section of the gluon-gluon fusion Higgs boson production mode is given by the following equation,

$$\sigma_{pp \to HX} = \int_0^1 \int_0^1 g(x_1)g(x_2)\sigma(gg \to H)dx_1dx_2. \tag{2.47}$$

Where the calculation is comprised of the cross-section for the $gg \to H$ process ($\sigma(gg \to H)$), and the gluon PDF ($g(x_1)$ and $g(x_2)$, where x_1 and x_2 are the fraction of the colliding protons momentum carried by each gluon).

To calculate the cross-section of two interacting partons, one starts from the free field description of each parton and includes a potential term that describes the interaction between them. The cross-section for the scattering is calculated as the integration over all possible paths that result in the correct outgoing particles, including all possible intermediate particle states. This integral is calculated as a perturbation expansion that is typically convergent. The number of times the interaction potential acts is the order of the perturbation expansion. Because the expansion is convergent, as the order of the perturbative expansion is increased, the accuracy of the prediction is increased. The lowest order calculation is referred to as leading order (LO), and subsequent increase in order is referred to as next-to-leading order (NLO) then NNLO and so on. Calculations are typically computed to at least NLO accuracy.

The Feynman rules were devised by Richard Feynman in the 1940s to make the construction of the scattering cross-section calculations easier. Incoming particles, outgoing particles, intermediary particles and their interactions all have associated Feynman rules. The accompanying Feynman diagrams provide a diagrammatic description of the interaction and the calculation. In a Feynman diagram, particles are represented by lines, with different types of particles (such as fermions or photons) represented by different types of lines. Interactions between particles are represented by vertices where two or more lines meet. The x direction represents the time dimension, and the y direction represents the space dimension. Antiparticles are interpreted as travelling backwards in time. The integral over all possible paths in the cross-section calculation can be interpreted as a sum over all possible Feynman diagrams for the interaction. Figure 2.3 shows two example Feynman diagrams for the $gg \to H$ process at LO and NLO.

PDFs describe the probability of finding a parton with a certain momentum fraction inside a hadron. PDFs cannot be calculated directly from QCD. Instead, they are determined through a combination of experimental measurements and theoretical models, such as global fits to data from deep inelastic scattering

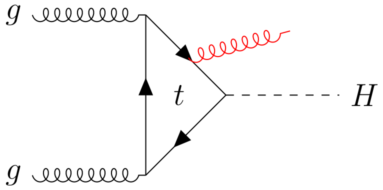


Figure 2.3 *Example Feynman diagrams for the $gg \rightarrow H$ process at LO (black lines) and NLO (black and red lines).*

experiments. Figure 2.4 shows an example graphical representation of a PDF, where the gluon PDF used in the gluon-gluon fusion cross-section calculation is is shown as g/10.

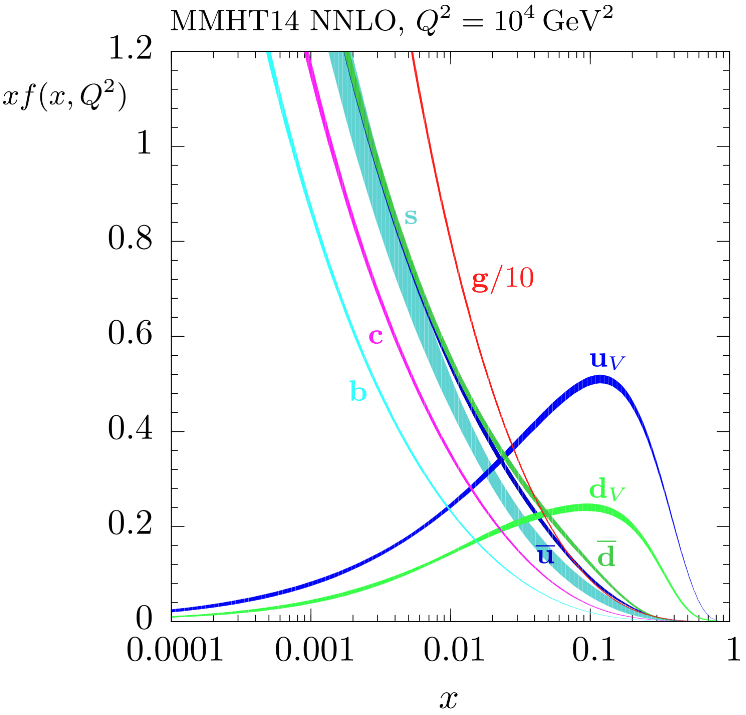


Figure 2.4 *Example PDF at $Q^2 = 10^4$ GeV^2. The g/10 line is the gluon PDF scaled down by a factor of 10. The width of the line represents the uncertainty on the PDF. Taken from [7].*

2.6 Properties of the Higgs Boson

The Higgs boson arises as a result of the Higgs mechanism. The Higgs boson is a neutral scalar particle. Its mass is a free parameter of the SM given by $m_H = 2v\lambda^2$, and so must be found experimentally.

The Higgs boson couples to all fermions with a coupling strength proportional to the fermion mass. It can therefore decay via $H \rightarrow ff$, for all kinematically allowed processes where m_h must be greater than $2m_f$. We also saw that the Higgs boson couples to the $W^{\pm}$ and Z gauge bosons, therefore if it is sufficiently massive, it can also decay via $H \rightarrow W^+W^-$ or $H \rightarrow ZZ$.

This proportionality of the coupling to mass determines the dominant processes through which the Higgs is produced, and through which it decays. The Higgs boson, therefore, couples preferentially to the most massive particles that are kinematically accessible.

2.7 The Higgs Boson at the LHC

Before the turn-on of the LHC, the unexplored mass window for a SM Higgs was fairly narrow. The lack of signal from the Large Electron Positron (LEP) experiment implied that $m_H > 114$ GeV and limits on quantum loop corrections from precision EW measurements at LEP and the Tevatron suggested that $m_H \lesssim 150$ GeV [8]. The discovery of the Higgs boson was therefore one of the main aims of the LHC, and the main motivation for its construction.

2.7.1 Production Modes

The Higgs boson can be produced at the LHC in many different ways. Here we outline the main Higgs boson production modes at the LHC and those that are relevant to the analyses discussed in Part III. Feynman diagrams for each of the production modes are shown, where the incoming quarks and or gluons are constituents of the protons in the collisions at the LHC.

ggF

Gluon-gluon fusion (ggF) is the dominant Higgs production mode at the LHC, where the Higgs is produced through a loop of virtual top quarks. The process can be represented by the equation $gg \rightarrow H$, the Feynman diagram is shown in Figure 2.5. The ggF process contributes ~87% of the total Higgs production cross-section at the LHC.

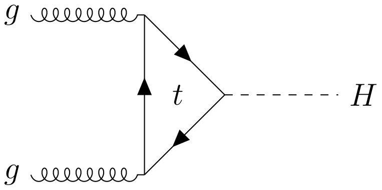


Figure 2.5 *Example Feynman diagram for the ggF Higgs boson production mode.*

VBF

Vector Boson Fusion (VBF) production is important at the LHC. This is because of the easily identifiable final state: two forward jets as a result of the break up of the colliding protons, and the decay products of the Higgs. The process can be represented by the equation $qq' \rightarrow Hqq'$, the Feynman diagram is shown in Figure 2.6. The VBF process contributes ∼7% of the total Higgs production cross-section at the LHC.

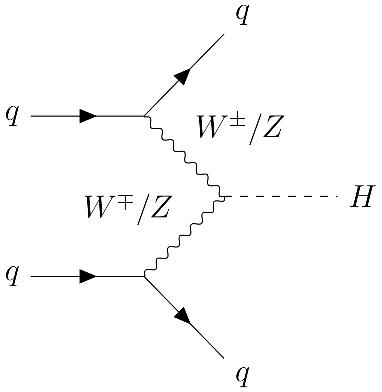


Figure 2.6 *Example Feynman diagram for the VBF Higgs boson production mode.*

VH

Higgs-Strahlung or VH production occurs when a vector boson ($W^{\pm}$ or Z), created via the interaction of two quarks, radiates a Higgs boson. The process can be represented by the equation $q\bar{q}' \rightarrow VH$, the Feynman diagram is shown in Figure 2.7. The VH process contributes ∼4% of the total Higgs production cross-section at the LHC.

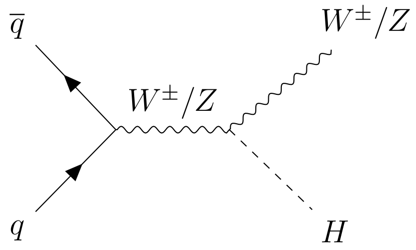


Figure 2.7 *Example Feynman diagram for the VH Higgs boson production mode.*

$t\bar{t}$H

Top quark pair associated production or $t\bar{t}H$ can be initiated by gluons or quarks from colliding protons, the Higgs is radiated from a top quark in the interaction. The decays of the short-lived top quarks are usually searched for and reconstructed. The process results in the production of many jets alongside the decay products of the Higgs making it difficult to distinguish from backgrounds. The Feynman diagram is shown in Figure 2.8. The $t\bar{t}H$ process contributes ~4% of the total Higgs production cross-section at the LHC.

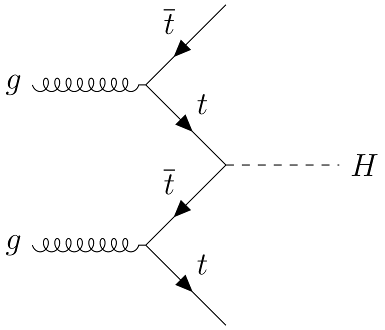


Figure 2.8 *Example Feynman diagram for the $t\bar{t}H$ Higgs boson production mode.*

$b\bar{b}$H

The $b\bar{b}H$ process is similar to the $t\bar{t}H$ process, but the Higgs boson is produced in association with two b-quarks. The process is very hard to separate from QCD backgrounds at the LHC. It is not explicitly targeted in the analysis discussed in Chapter 8 due to its similarity to the ggF process, however, it is considered when conducting the analysis measurement. The $b\bar{b}H$ process contributes ~1% of the total Higgs production cross-section at the LHC.

tH

Single top Higgs production or tH describes the production of a Higgs boson in association with a single top quark. It is typically split into two sub-processes: when the top and the Higgs are produced in association with a W boson (tHW), and when they are produced in association with a light quark and a b-quark ($tHqb$) (Figure 2.9). The tH process contributes $\sim$ 0.1% of the total Higgs production cross-section at the LHC, making it very rare.

These processes are of particular interest due to their sensitivity to the sign of the top quark Yukawa coupling strength. This sensitivity is used to constrain the allowed values of the top quark Yukawa coupling strength in the analysis detailed in Chapter 9.

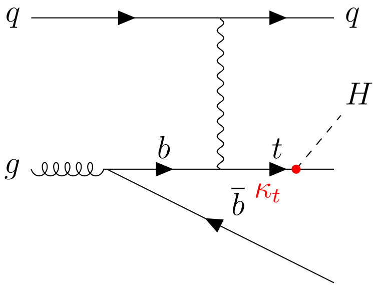


Figure 2.9 *Example Feynman diagram for the tHqb Higgs boson production mode.*

Figure 2.10 shows the cross-section of each of the Higgs boson production modes at the center of mass energies relevant to the LHC for a Higgs boson mass of $m_H = 125$ GeV.

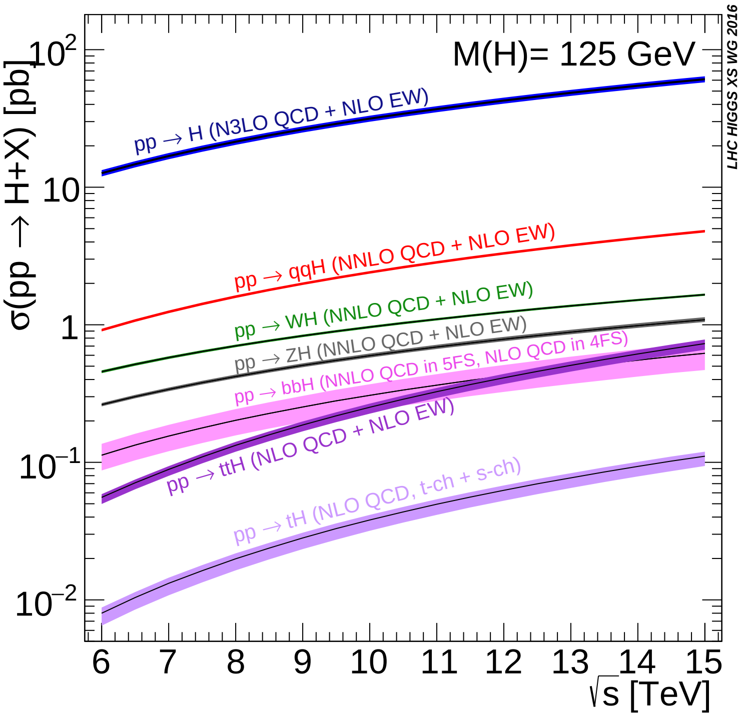


Figure 2.10 *Cross-section of the primary Higgs boson production modes as a function of centre-of-mass energy for a Higgs boson mass of* $m_H =$ 125 *GeV. Taken from [9].*

2.7.2 Decays

The lifetime of the Higgs boson is extremely small. This means that its existence has to be inferred through the observation of its decay products. The branching fractions of the dominant Higgs boson decay modes are shown in table 2.1 for a Higgs boson mass of $m_H = 125$ GeV.

The decay modes with the largest branching fractions are those where the Higgs couples to the most massive particles in the SM. Decays into a pair of WW^* and ZZ^* are kinematically allowed if one of the gauge bosons is produced *off-shell*. Even though the gluon and photon are massless, the Higgs can decay into pairs of each via a top quark loop. The Feynman diagram for a Higgs boson decaying into a pair of photons is shown in Figure 2.11.

Despite the low branching fraction of the $H \to \gamma\gamma$ decay channel, it provides a high sensitivity to the Higgs boson. This is due to the clean signature of two high-energy isolated photons that are easily identifiable in the typical QCD background at the LHC. Part III presents the state-of-the-art ATLAS analysis in this decay channel.

Decay mode	Branching Fraction
$H \to b\bar{b}$	5.824×10^{-01}
$H \to WW^*$	2.137×10^{-01}
$H \to gg$	8.187×10^{-02}
$H \to \tau^+\tau^-$	6.272×10^{-02}
$H \to c\bar{c}$	2.891×10^{-02}
$H \to ZZ^*$	2.619×10^{-02}
$H \to \gamma\gamma$	2.270×10^{-03}
$H \to Z\gamma$	1.533×10^{-03}
$H \to \mu^+\mu^-$	2.176×10^{-04}
Total Width [GeV]	4.088×10^{-03}

Table 2.1 *Branching fractions of the dominant Higgs boson decay modes for a Higgs mass of* $m_H = 125$ *GeV [10]*

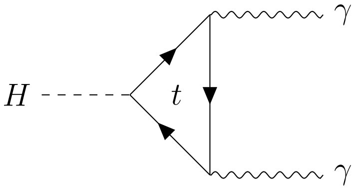


Figure 2.11 *The Feynman diagram for the Higgs boson decaying into a pair of photons via a top quark loop.*

2.7.3 Discovery

On the 4th of July 2012, the ATLAS and CMS experiments announced the discovery of a boson consistent with the SM Higgs boson [11, 12] using proton-proton (p-p) collision data at a centre-of-mass energy of $\sqrt{s} = 7$ and 8 TeV. Searches were conducted in the five decay modes: $\gamma\gamma$, ZZ^*, WW^*, $\tau^+\tau^-$ and $b\bar{b}$. The highest sensitivity was achieved in the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels, due to their distinguishable signatures. In both of these channels, the mass of Higgs boson candidate is reconstructed with high sensitivity. The combined mass of the Higgs boson candidate was later measured to be,

$$m_H = 125.09 \pm 0.21 \text{ (stat)} \pm 0.11 \text{ (syst) GeV [13].} \tag{2.48}$$

The discovery of the Higgs boson completed the SM picture that we know today (Figure 2.1). However, it marked only the start of a rich physics program at the LHC.

Chapter 3

The ATLAS Experiment

The European Organisation for Nuclear Research ("Conseil Européen pour la Recherche Nucléaire") (CERN) was set up in 1954 as a European research organisation of particle and nuclear physics. It now houses the Large Hadron Collider (LHC); the largest particle accelerator ever created. Four main experiments are situated along the LHC ring. This chapter will give an overview of the ATLAS experiment, including the ATLAS detector, physics object reconstruction, and upgrades.

3.1 The Large Hadron Collider

The LHC is a superconducting hadron accelerator located at CERN in Geneva. It is installed in a 26.7 km tunnel that was originally constructed for the Large Electron Positron (LEP) experiment. The tunnel lies underground at depths varying between 45m and 170m below the surface.

The LHC ring is connected to the CERN accelerator complex by two transfer tunnels, The tunnels inject proton beams travelling in opposite directions around the ring, contained by two separate beam pipes. These proton beams are accelerated to the required energy by a series of smaller accelerators: initial proton bunches are accelerated by LINAC2, and the booster, PS, and SPS accelerator rings before injection into the LHC tunnel (figure 3.1). A final "ramp up" phase is completed inside the LHC tunnel to accelerate the proton bunches to their final energy.

The LHC is designed to collide proton beams up to a center of mass energy of $\sqrt{s} = 14$ TeV and at a luminosity of 10^{34} cm^{2}s^{-1}. It can also collide heavy ions with an energy of $\sqrt{s} = 2.8$ TeV and a peak luminosity of 10^{27} cm^{2}s^{-1} [14]. In Run 1, between 2010 and 2012 the LHC operated at a center-of-mass energy of $\sqrt{s} = 7$ TeV and 8 TeV. Throughout Run 2, in 2015-2028 the LHC operated at a center of mass energy of $\sqrt{s} = 13$ TeV.

The proton beams collide at four interaction points on the LHC ring; a particle detector is situated at each point. The ATLAS and CMS particle detectors are general-purpose detectors, the LHCb detector is a forward arm detector designed for accurate b-physics and the ALICE experiment is designed around the high multiplicity conditions arising when colliding heavy ions.

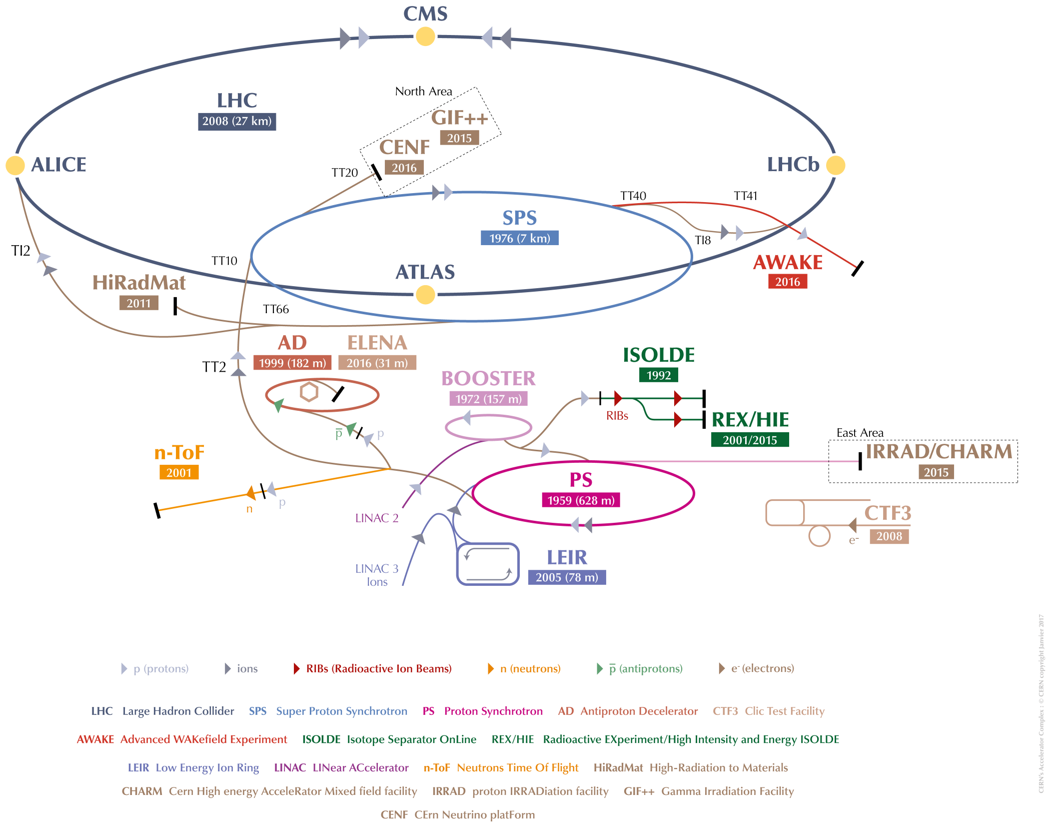


Figure 3.1 *Diagram illustrating the accelerator complex at CERN. Taken from [15]*

3.2 Runs, Luminosity & Pile-Up

The event rate dN/dt of a physical process created in collisions is given by:

$$dN/dt = L\sigma_{event}. \tag{3.1}$$

Where σ_{event} is the cross-section of the process in question, and L is the instantaneous luminosity.

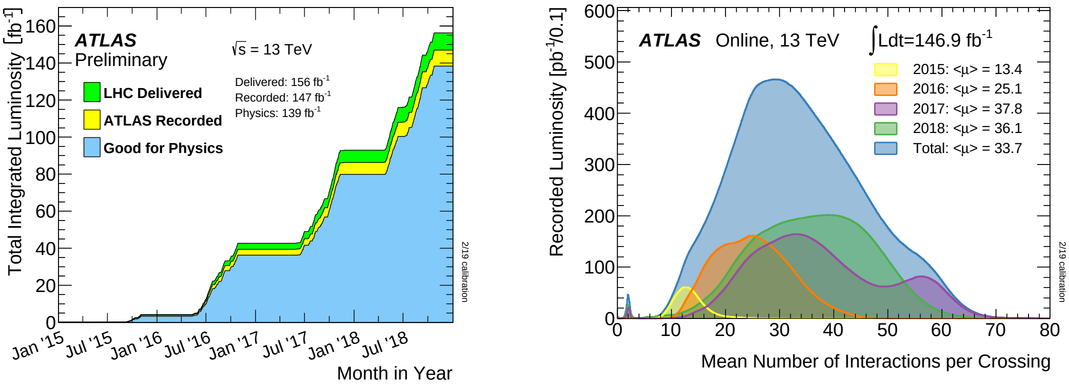
The luminosity of the LHC depends on many parameters related to the proton beam. For a Gaussian beam distribution, the instantaneous luminosity is given as,

$$L = \frac{N_b^2 n_b f_{rcv}}{4\pi\sigma_x^*\sigma_y^*}F = \frac{N_b^2 n_b f_{rcv}\gamma_r}{4\pi\epsilon_n\beta^*}F \tag{3.2}$$

where N_b is the number of particles per bunch in the proton beam, n_b is the number of bunches per beam, $f_{rcv} = 11.25$ kHz is the revolution frequency, and F is the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP). For circular beams at the IP the beam sizes in the horizontal and vertical plane σ_x^* and σ_y^* are identical, they can therefore be expressed as $\sigma_x^*\sigma_y^* = \beta^*\epsilon_n/\gamma_r$, where γ_r is the relativistic gamma factor, ϵ_n is the normalised transverse beam emittance, and β^* is the beta function at the IP [16].

The integrated luminosity measures of how many collisions occurred over a defined time period. It is defined as the integral of the instantaneous luminosity over this time period. The LHC Delivered, ATLAS Recorded, and Good for Physics integrated luminosity throughout Run 2 is shown in Figure 3.2 (a). The Good for Physics integrated luminosity of 139 fb^{-1} defines the amount of data available for Run 2 physics analysis.

Within each bunch crossing the number of proton-proton (p-p) collisions that occur is referred to as the pile-up. Typically, per bunch crossing, only one p-p collision will result in a hard scatter interaction that is of physics interest, other soft collisions result in Quantum Chromodynamics (QCD) backgrounds and present a challenge for object reconstruction (Section 3.10). The mean number of interactions per crossing recorded by the ATLAS detector throughout Run 2 is shown in Figure 3.2 (b).



(a) **(b)**

Figure 3.2 *The LHC Delivered, ATLAS Recorded, and Good for Physics integrated luminosity (a). The mean number of interactions per bunch crossing measured by the ATLAS detector for the Run 2 period (b) [17].*

3.3 The ATLAS Detector Overview

The ATLAS detector is a general-purpose particle detector designed to operate in the high luminosity conditions of the LHC. The high luminosity and energy of the LHC allow for more precise measurements of QCD, electroweak and flavour interactions. These expectations, along with the concrete goal of observing the Higgs boson were used as benchmarks in its design [16].

These considerations led to a set of general requirements for the ATLAS detector:

- fast, radiation-hard electronics and sensor elements
- high granularity
- large acceptance in pseudorapidity
- good charged particle momentum resolution and vertex detection
- very good electromagnetic (EM) calorimetry for photon and electron identification
- full coverage hadronic calorimetry for jet and $E_{\mathrm{T}}^{\mathrm{miss}}$measurements
- good muon identification
- highly efficient triggering

The following sections cover the ATLAS coordinate system the magnet system and the design of each detector component. A computer model of the ATLAS detector can be seen in Figure 3.3. The bulk of the detector is cylindrical, where the region forming the lateral surface of the cylinder is called the *barrel*, and the two faces of the cylinder are called the *end-caps*.

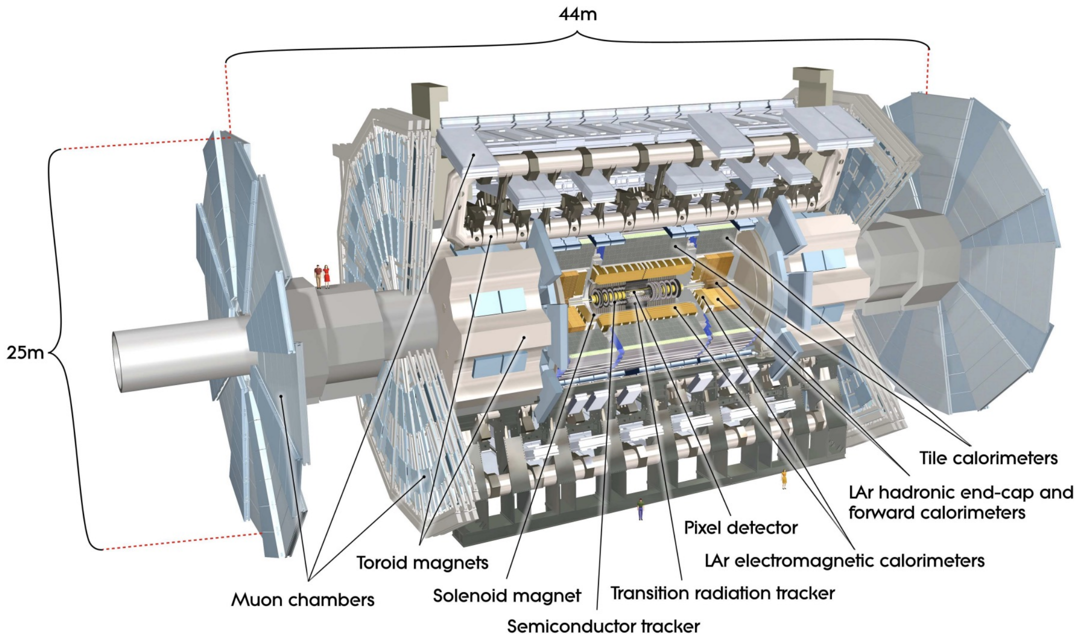


Figure 3.3 *Cut away computer model of the ATLAS detector, showing each of the detector components. Taken from [16].*

3.4 The Coordinate System

A unified coordinate system is used to describe the ATLAS detector and the particles created in p-p collisions. The origin is positioned at the nominal interaction point, the z-axis is aligned with the beam direction, and the x-y plane is perpendicular to the z-axis. The positive x-axis points towards the center of the LHC ring and the positive y-axis points upwards. The positive z-axis is in the direction of the A-Side of the ATLAS detector. The azimuthal angle $\phi \in [-\pi, \pi]$ is measured in the x-y plane around the beam axis, where the positive x-axis direction corresponds to $\phi = 0$, and the positive y-axis corresponds to $\phi = \pi/2$. The polar angle $\theta \in [0, \pi]$ is defined, where the positive z-axis corresponds to $\theta = 0$. Finally, pseudorapidity η is defined, given by Equation 3.3 [16].

$$\eta = -\ln\tan(\theta/2) \tag{3.3}$$

The coordinate system is summarised in Figure 3.4.

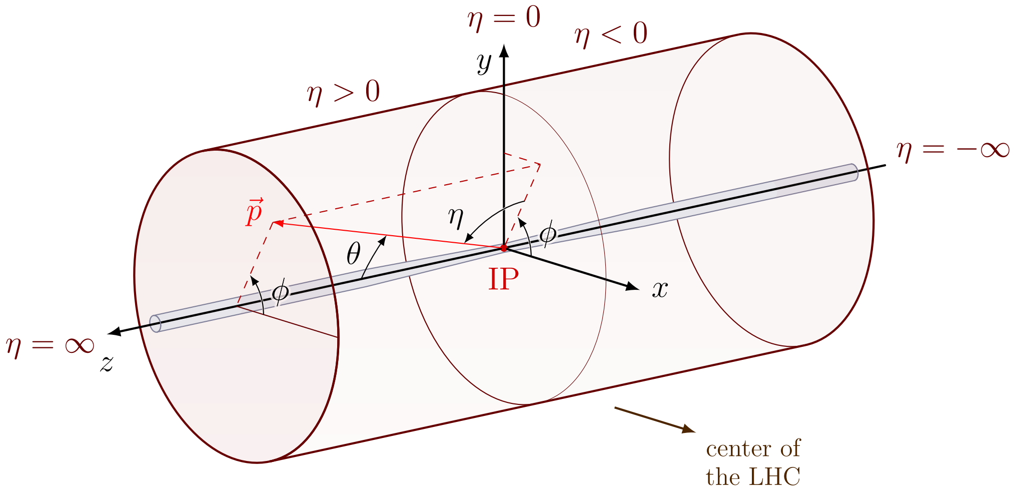


Figure 3.4 *The coordinate system used to define the ATLAS detector and physics objects. Taken from [18].*

Variables that further describe physics objects are defined in this coordinate system. The transverse momentum (p_T), defined as the magnitude of a particle's momentum perpendicular to the beam axis, is given by,

$$p_T = \sqrt{p_x^2 + p_y^2}. \tag{3.4}$$

ΔR is used extensively in the reconstruction of objects, overlap removal, and the definition of variables used in machine learning models. It is defined by the following equation,

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{3.5}$$

3.5 Magnet System

The ATLAS magnet system consists of one solenoid in the barrel region of the detector, and three toroids, one in the barrel and two in the endcap regions of the detector. These four large superconducting magnets allow for particle charge identification and momentum measurement in the ID and MS by bending the flight path of charged particles.

The single solenoid magnet surrounds the ID, providing a 2T axial magnetic field. The three toroidal magnets provide magnetic fields across each of the MS components. The barrel toroid surrounds the calorimeters and both endcap toroids generating the necessary magnetic field for the MS. The barrel toroid provides 1.5 to 5.5 Tm of bending power in the pseudorapidity range $0 \leq |\eta| \leq 1.4$, and the endcap toroids, approximately 1 to 7.5 Tm in the region $1.6 \leq |\eta| \leq 2.7$ [16].

3.6 Inner Detector

The inner detector (ID) is responsible for the momentum and trajectory measurements, and vertex tagging of charged particles as soon as they emerge from the IP. With approximately 1000 particles created every 25 ns within $\eta \leq 2.5$, the ID needs a very high resolution. This is achieved using thee complimentary sub-detectors and technologies: the Pixel detector, the silicon microstrip semi-conductor tracker (SCT), and the Transition Radiation Tracker (TRT) [16]. Each ID component can be seen in Figure 3.5.

The high granularity of the pixel detector and the SCT at the inner radii provide high-resolution tracking capabilities. At larger radii, the TRT allows for continuous tracking to enhance the Pixel and SCT, whilst also improving momentum resolution and particle identification [16].

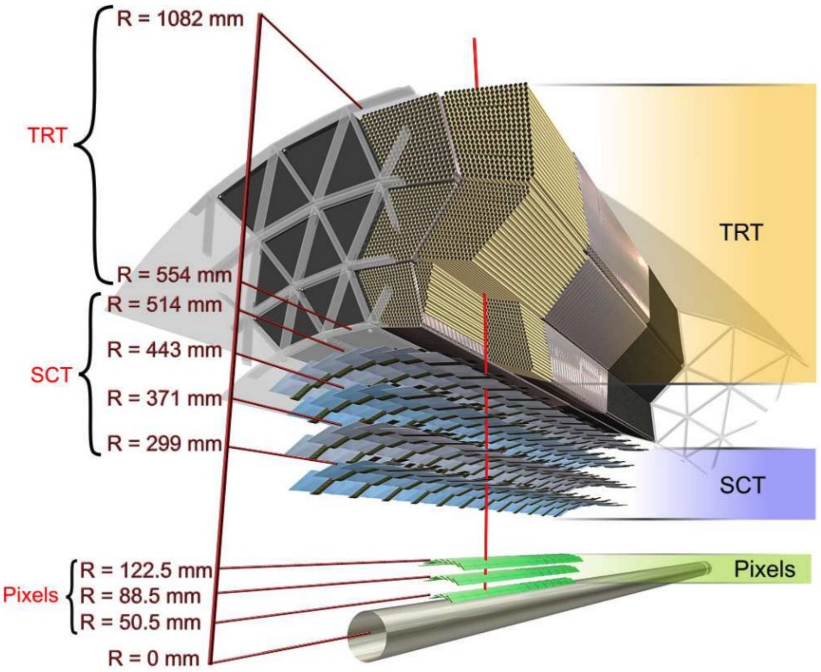


Figure 3.5 *Cut away computer model of the ID in the barrel region, showing each of the detector components. Taken from [16].*

3.6.1 Pixel Detector

The Pixel detector is the first detector a particle encounters on its flight path. It uses silicon pixel technology that can provide the necessary position resolution at the scale of tens of microns. The pixels each have a size of 50 μm in the $R-\phi$ direction, and 400 μm in the z direction. Charged particles traversing the pixels excite electron-hole pairs which are swept to opposite electrodes. The electrical signal is read out as a hit [16]. The Pixel detector was initially constructed from 3 layers in the barrel and the end-cap, however, in 2014 the insertable B-layer (IBL) was installed in the barrel region at 33 mm from the beam axis. The IBL improved the resolution of the Pixel detector close to the beam axis [19].

3.6.2 SCT Detector

The SCT uses the same semiconductor technology as the Pixel detector. It is composed of four cylindrical layers of silicon strip detectors in the barrel region, and 9 circular layers in the end-cap regions. The strip sensors have an 80 μm pitch, a length of ~6 cm, and a thickness of ~285 μm. The strips are grouped in pairs, in a back-to-back fashion, rotated at a stereo angle of 40 mrad. This arrangement significantly improves measurement resolution in the z direction. Due to their remarkable resolution hits in the SCT are referred to as "space-points" [16].

3.6.3 TRT Detector

The TRT uses a straw tube detector technology. The straws have a diameter of 4 mm. In the barrel straws are aligned with the beam axis, whereas in the end-cap they are positioned in wheels perpendicular to the beam radii. The straws are filled with Xenon gas, and a gold-plated tungsten anode wire runs down their middle. Charged particles traversing the tube will ionize the gas, creating electron-ion pairs. The negatively charged electrons drift towards the anode wire, accelerated by the electrical field. The electrical pulse is read out as a signal. In addition to providing spatial measurements of charged particles, the TRT can also be used for particle identification. The straws in the TRT are interleaved with polypropylene fibers in the barrel, and foils in the endcaps. When a relativistic charged particle traverses the boundary between the materials, transition radiation photons are emitted. The photons are absorbed by the gas in the straws and an additional signal is read out. Particles with a higher Lorentz factor ($\gamma = E/m$) will induce significantly larger signals. Paired with an energy measurement, the signal can be used to distinguish between pions and electrons [16].

3.7 Calorimeters

A cutaway view of the calorimeter system in the ATLAS detector is shown in Figure 3.6. The calorimeters cover the large pseudorapidity range of $|\eta| < 4.9$, each using different technologies suited to the wide range of physical processes of interest. The fine granularity of the EM calorimeter compliments that of the ID, allowing for precise identification and measurements of electrons and photons. The surrounding hadronic calorimeters are more coarse, but provide sufficient precision for jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements. The physical processes that occur in EM and hadronic showers are detailed in Chapter 4.

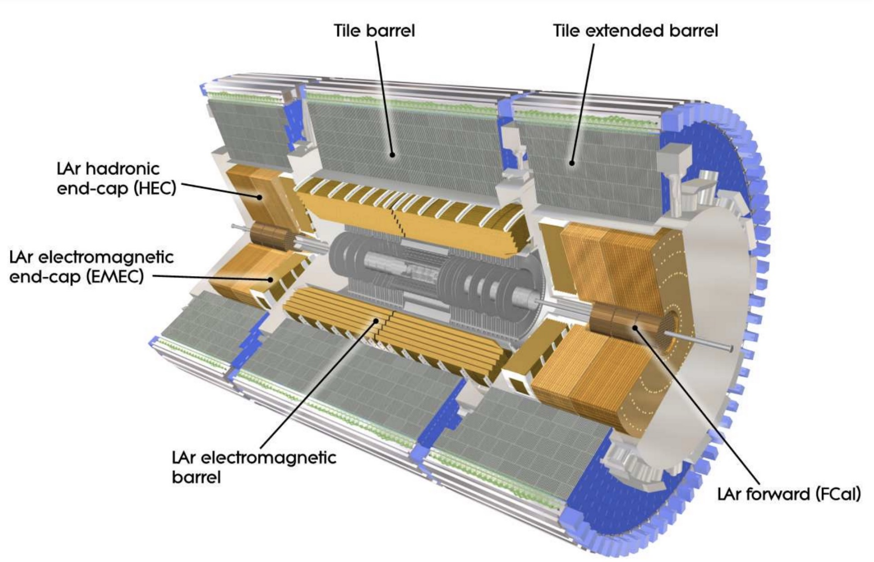


Figure 3.6 *Cut away model of the different calorimeters in the ATLAS detector. Taken from [16].*

3.7.1 LAr Electromagnetic Calorimeters

Two components comprise the EM calorimeters, these are: the barrel calorimeter ($|\eta| < 1.475$), and the two end-cap calorimeters ($1.375 < |\eta| < 2.5$). The EM calorimeter uses a lead-Liquid-Argon (LAr) technology, arranged in an accordion-like structure which allows for complete ϕ symmetry without any azimuthal cracks, and fast signal extraction (Figure 3.7). The lead strips act as an absorber causing incident particles to shower. Secondary particles created in the shower then ionize the LAr active material creating a large number of electrons that are collected at the electrodes. The electrons are interpreted as signals, and a proxy for the energy deposited.

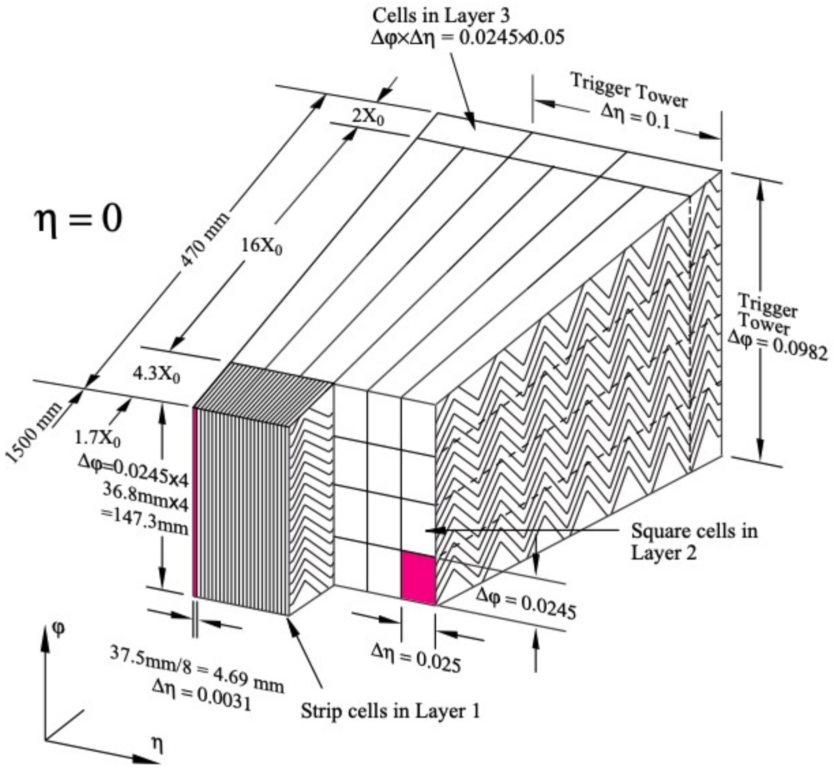


Figure 3.7 *Sketch of an EM calorimeter barrel module where the layers and accordion structure are clearly visible. Taken from [16].*

3.7.2 Hadronic Calorimeters

The Tile calorimeter sits directly outside the EM calorimeters and covers the barrel region of the detector. The tile calorimeter is comprised of two sections, these are: the tile barrel calorimeter ($|\eta| \leq 1.0$), and the extended barrel calorimeter ($0.8 \leq |\eta| \leq 1.7$). Each is a sampling calorimeter, made up of steel as the absorber and scintillating tiles as the active material. The sides of the scintillating tiles are read out by wavelength-shifting fibers feeding into photomultiplier tubes.

The LAr hadronic end-cap calorimeter (HEC) consists of two independent wheels per end-cap. They sit directly outside of the EM endcap calorimeters and cover the pseudorapidity region ($1.5 \leq |\eta| \leq 3.2$). The calorimeters use copper as the absorber material and LAr as the ionising material.

The LAr forward calorimeter (FCal) covers the most forward regions of the detector ($3.1 \leq |\eta| \leq 4.9$). It provides benefits to calorimetric coverage and reduces radiation background levels in the muon spectrometer. The FCal consists of three modules in each endcap: the innermost is made of copper and is optimised for EM measurements, while the other two are made of tungsten and designed for hadronic measurements. Each is made up of a matrix of absorber rods parallel to the beam axis and LAr as the active medium. This geometry allows for fast measurements in the high flux forward region.

3.8 Muon Spectrometer

The MS consists of precision-tracking monitor drift tube (MDT) chambers in the barrel region ($\eta \leq 1.05$), positioned in and around eight coils of superconducting toroid magnets. The endcap ($1.05 \leq \eta \leq 2.7$) sections of the MS lie in front and behind the endcap toroid magnets. The innermost detector in the endcap region is instrumented with cathode strip chambers (CSCs) instead of MDTs to withstand higher rate and background conditions. A trigger system, covering the range $|\eta| \leq 2.4$ consists of Resistive Plate Chambers (RPCs) in the barrel region and Thin Gap Chambers (TGCs) in the endcap regions. These trigger chambers serve multiple purposes: providing bunch-crossing identification, well-defined p_{T} thresholds, and allow for the improved measurement of muons, orthogonal to that of the MDTs and CSCs.

A cross-section of the MS in the $y - z$ plane is shown Figure 3.8, where each of the detectors described above can be seen.

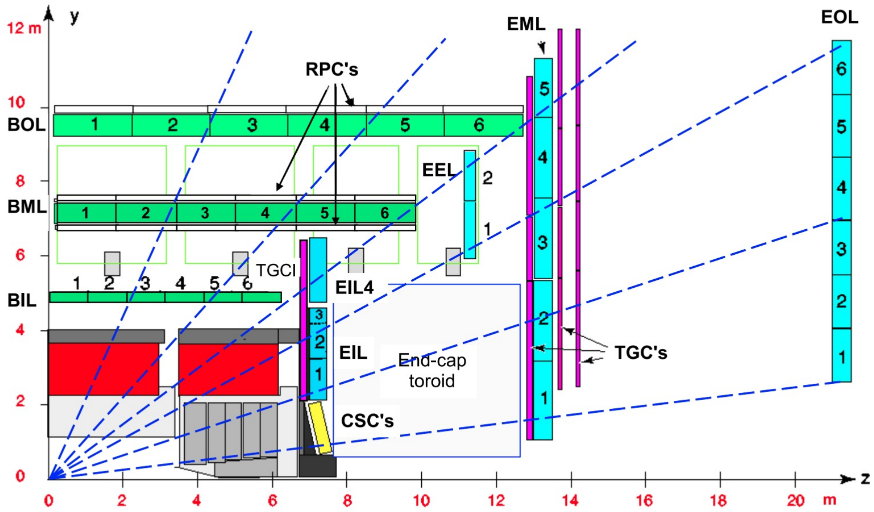


Figure 3.8 *Cross-section of the MS in the y-z plane, where dashed lines show pseudorapidity contours. BIL, BML, and BOL stand for Barrel Inner Layer, Barrel Middle Layer, and Barrel Outer Layer respectively. A similar description is used for the endcap layers. Taken from [16]*

.

3.9 Trigger System

The LHC operates at the design bunch crossing rate of 40 MHz. However, due to hardware and storage limitations, the ATLAS experiment cannot record data for all collisions. The rate is reduced to $\sim$1 kHz using a two-layered trigger system that ensures only collisions that are of physics interest are saved. The trigger event selection is based on physics signatures, such as the presence of energetic leptons, photons, jets or $E_{\mathrm{T}}^{\mathrm{miss}}$. The trigger system also exploits topological information and multivariate methods [20].

The trigger system used throughout Run 2 operates using two components:

- **The Level-1 Trigger (L1)** consists of the L1Calo and L1Muon subsystem triggers that process inputs from the calorimeters and the Muon Spectrometer (MS). In addition, the L1Topo trigger performs topology-based selections using inputs from the calorimeter and the MS. The L1 trigger cuts down the event rate from 40MHz to a maximum of 100kHz.
- **The High-Level Trigger (HLT)** is a software-based trigger that is applied to events that pass the L1 trigger selections. These events are temporarily buffered, and a limited online object reconstruction is applied.

The HLT then decides if the events should be permanently saved for analysis, reducing the rate of recorded events on average to ∼1 kHz.

3.10 Reconstruction and Physics Objects

Particles produced in the *p*-*p* collisions interact with the components of the ATLAS detector leaving signatures. Using these signatures, the following physics objects are reconstructed.

3.10.1 Tracks

Tracks play a key role in of the reconstruction of interaction vertices, the removal of pile up, the flavour tagging of jets, and the reconstruction of charged particles. They are formed from the signatures left in the ID by charged particles originating from the *p*-*p* collisions at the IP [21, 22]. The track reconstruction algorithm performs the following steps:

- **Formation of space-points:** Clusters of energy deposits in the Pixel and SCT, and drift circles in the TRT are transformed into 3D *space-points*.
- **Track finding seeded by space-points:** Sets of three space-points reconstructed in the Pixel and SCT are used to form seeds, these can either be *Pixel-only*, *SCT-only*, or *Mixed* seeds. Seeds are required to pass p_{T} and impact parameter resolution cuts, and to match a fourth trajectory compatible space-point. A Kalman filter [23] is applied to complete the track candidates.
- **Ambiguity solving:** Tracks are ranked using a *track-score* that is based on aspects such as cluster quality, number of holes in the track, and χ^2 penalisation for tracks with a poor fit. The logarithm of the track momentum is also combined into the track-score to favour higher energy tracks. The track-score is then used to remove any ambiguity with tracks that share clusters, favouring tracks with a higher score.
- **TRT extension:** Finally, track candidates are extended to include space-points in the TRT if there is a match. A track refit is then performed to improve momentum resolution.

The components of the track reconstruction algorithm are illustrated in Figure 3.9.

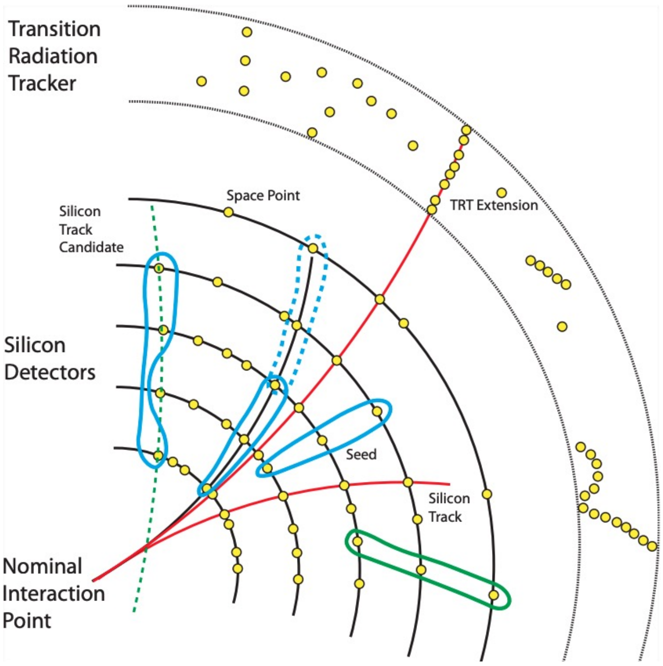


Figure 3.9 *Illustration of the components that are used to reconstruct tracks. Taken from [24].*

3.10.2 Primary Vertex

Efficient and precise reconstruction of the primary vertex, defined as the point at which a p-p interaction occurred, is essential in the analysis of data collected at the LHC. The primary vertex is used when deciding on which reconstructed particles and objects originated from the underlying hard process.

The primary vertex reconstruction algorithm uses tracks that have passed a set of quality requirements as detailed in [25]. Using these tracks, a seed position for the first vertex is selected. The tracks and vertex seed are then used to fit the best vertex position. This is applied iteratively, down-weighting less compatible tracks, and updating the vertex position until the final vertex position is determined. On completion, incompatible tracks are removed and are available to use in the selection of an alternative vertex. A vertex is required to have at least two compatible tracks. This process is repeated until all tracks are associated to a vertex, or no additional vertex can be found for the remaining tracks. The output of the vertex selection algorithm is a set of vertices, their coordinates, and their associated tracks.

3.10.3 Topo Clusters

Energy deposits inside the calorimeters are topologically-grouped into *topo-clusters*. These topo-clusters form the basis of the reconstruction of physics objects such as electrons, photons, and jets. The topo-clusters are formed by first selecting a calorimeter cell seed that has an energy deposit of more than 4σ of energy, where σ is the average amount of noise expected in the cell. All adjacent cells that contain at least 2σ of energy are then grouped into the cluster. This process is repeated until there are no more adjacent 2σ cells. The topo-cluster is completed by adding all adjacent calorimeter cells irrespective of their energy. The process is often referred to as 4-2-0 reconstruction, reflecting the energy requirements of adjacent cells [26].

3.10.4 Electrons and Photons

Topo-clusters and tracks form the input of the reconstruction of electrons and photons in ATLAS. The algorithm starts by considering topo-clusters in the entire calorimeter system. A cut is applied on the fraction of energy detected in the EM calorimeters to isolate topo-clusters that are primarily the result of EM showers. These clusters are used to create Regions of Interest (ROIs) within the ID where the standard tracking (Section 3.10.1) reconstruction is performed [27].

Track-matching for electron candidates is then performed on the loosely-matched tracks in a bremsstralung-aware refit to the EM clusters. A photon conversion vertex reconstruction is also performed on the loosely-matched tracks, and they are matched to topo-clusters. With this completed, a *supercluster* algorithm is run separately in parallel for electrons and photons. The supercluster algorithm accumulates *satellite* clusters together with a main *seed* cluster to form larger clusters. This allows for the recovery of low-energy photons radiated due to bremsstralung, that can cause satellite energy deposits in the calorimeters. A diagram illustrating the supercluster algorithm is shown in Figure 3.10.

Superclusters are built by first selecting the topo-clusters to be used as seeds. Topo-clusters are sorted by p_{T} such that higher p_{T} seeds only collect lower p_{T} satellites. For a topo-cluster to become an electron supercluster seed it must have a minimum energy of 1 GeV and must be matched to a track with ≥ 4 hits in the

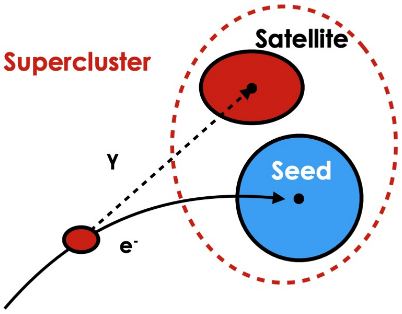


Figure 3.10 *Diagram illustrating the supercluster algorithm. Taken from [27]*

silicon tracking detector. To become a photon supercluster seed a topo-cluster must have an energy of at least 1.5 GeV [27].

With the supercluster seeds identified, the satellite matching procedure begins, where all remaining topo-clusters are examined. There is no p_{T} requirement placed on the satellite clusters, other than they must have a smaller p_{T} than the seed cluster. For both electrons and photons, a cluster is considered a satellite if it falls in within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster.

For electrons, an additional satellite cluster search is performed using a window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.3$ around the seed cluster *barycentre*. Within this window, a cluster is considered a satellite if it has ≥ 1 matched tracks, and the best-matched track is shared with the seed cluster. This approach is summarised in Figure 3.11.

For converted photons, a cluster is added as a satellite if: it has the same conversion vertex as the seed cluster, or a track match that is part of the conversion vertex of the seed cluster. Only the best-matched tracks and conversion vertices are used for the matching.

When all satellite clusters have been found, the process is repeated for the next highest p_{T} seed cluster. However, if it has already been used as a satellite, it is skipped. The process continues until all clusters have been examined [27].

As a result, the electron and photon reconstruction algorithm returns two collections of particles: the electron collection, and the photon collection. There is initially an overlap in each collection, where the analyser can choose to give preference to either electron or photon reconstruction efficiency.

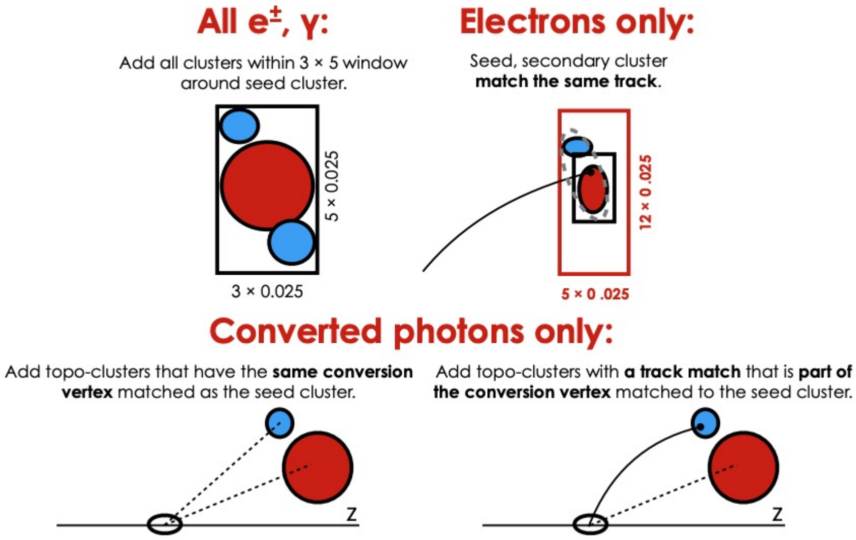


Figure 3.11 *Diagram illustrating the procedure used to add satellite clusters to a supercluster. Taken from [27]*

3.10.5 Jets

The vast majority of inelastic p-p collisions at the LHC result in the production of quarks and gluons [26]. The quarks and gluons undergo hadronisation, creating collimated streams of particles. Inside the ATLAS detector, these particles deposit energy inside the calorimeters. The collections of calorimeter energy deposits are reconstructed as *jets*, which give a useful representation of the original underlying process.

Topo-clusters are used as inputs to the anti-k_t algorithm [28] to reconstruct jets. A particle flow algorithm associates ID tracks with the topo clusters. If an association is made, the momentum measurement of the topo cluster is replaced with the more accurate momentum measurement from the ID track. Using this approach improves jet resolution and pile-up stability.

Two different R distance parameters are used when reconstructing the jets. Jets representing quarks and gluons typically called small-R jets use $R = 0.4$, and jets representing hadronically decaying massive particles are labeled large-R jets and use $R = 1.0$. Large-R jets are necessary to fully contain the energy deposited by massive hadronically decaying particles. However, these large-R jets are much more sensitive to pile-up effects, due to the jet radius covering much more of the calorimeter volume. Large-R jets are therefore *groomed* using the trimming procedure defined in [29].

Jet-Vertex-Tagger (JVT) The high pile-up conditions of collisions at the LHC result in challenges for identifying jets that arise from the hard process. Additional collisions, unrelated to the hard-scatter event, result in a background of soft energy deposits that have adverse and complex effects on jet reconstruction [30]. A jet vertex tagger is applied to jets to aid in distinguishing those that originate from the hard process from those from QCD backgrounds.

Flavour-Tagging The identification of jets containing b-hadrons and c-hadrons, referred to as b-jets and c-jets, is vital [31]. Jets are flavour-tagged using the DL1r algorithm [32], which uses a neural network that makes use of low-level information. The low-level information includes: track impact parameters, secondary vertex reconstruction, and displaced vertex information. These variables take advantage of the displaced decays of b and c hadrons due to their long lifetime.

3.10.6 Muons

Muons are reconstructed primarily from tracks in the ID and the MS supplemented by information from the calorimeters. The reconstruction is initially performed independently using information from the ID and the MS, this information is then combined to form the muon tracks that are used in physics analyses. In the ID, muons are reconstructed like any other charged particles [33].

Muon reconstruction in the MS Hit patterns inside the MS detector chambers are used to reconstruct *muon segments* inside the different detector layers. A Hough transform [34] is used to search for hits aligned on a trajectory in the bending plane of the ATLAS detector. The segments inside the MDT are reconstructed by performing a straight-line fit to the hits found in each layer. The RPC and TGC detectors measure the coordinate orthogonal to the bending plane. In the CSC detectors, segments are built using a separate combinatorial search in the η and ϕ planes. A loose requirement is made in the search algorithm on the compatibility of the track with the IP [33].

A segment-seeded combinatorial search is then used to reconstruct muon track candidates by fitting together hits from segments in the different MS detector layers. The algorithm starts first with segments in the middle layers of the MS detector where more trigger hits are available. The search is then expanded to use the segments from the outer and inner layers. When building a track segments are selected based on their hit multiplicity and fit quality. At least two segments are required to build a track, except in the barrel-endcap transition region of the detector, where a single high-quality segment can be used. Several track candidates can be built using the same segment. Later an overlap removal procedure is applied which selects the best track associated with the seed segment, or allows the segment to be shared between two tracks.

Segment hits are fitted using a global χ^2 fit to construct the tracks. Muons reconstructed using only information from the MS are referred to as *stand-alone muons*.

Combined Reconstruction Four different types of muons are reconstructed using various different algorithms applied to the information inside the ID, MS and calorimeters:

- **Combined (CB) muon:** track candidates are reconstructed independently in the ID and MS. A global refit is then performed using the segments from both of the track candidates. Segments can be added or removed in this step if they improve the fit quality. Typically an outside in procedure is used, where segments in the MS are fit, and the track is extrapolated to the ID.
- **Segment-tagged (ST) muons:** a track candidate in the ID is identified as a muon if when extrapolated it matches at least one segment in the MS MDT of CSC chambers. Segment tagged muons are reconstructed when a muon segment exists in only one of the MS layers, either because of low p_{T} or because they fall into a detector region with reduced MS acceptance.
- **Calorimeter-tagged (CT) muons:** a muon is reconstructed if a track candidate in the ID can be matched to a calorimeter energy deposit that is consistent with a *minimum-ionizing particle.* CT muons have the lowest purity, but this reconstruction allows for acceptance recovery in the regions where the MS is only partially instrumented.
- **Extrapolated (ME) muons:** a muon is reconstructed only from MS information if the trajectory of the track passes a loose requirement on

compatibility with the IP. The trajectory calculations takes into account estimated energy losses of the muon in the calorimeters. Segments in at least two of the MS layers are required, except in the forward region which requires three. These ME muons are mainly used to recover acceptance in the region $2.5 \leq |\eta| \leq 2.7$ which is not covered by the ID.

An overlap removal procedure is applied across each of the different muon types before producing a collection of muon objects that are used in physics analyses. When two types of muons share the same ID track candidate preference is given in the following order: CB muons, ST muons, CT muons. ME muons overlap is resolved by analysing the fit quality of tracks and selecting the track with a better fit quality and larger number of hits.

3.10.7 Missing Transverse Momentum

Missing Transverse Momentum (MET) ($E_{\mathrm{T}}^{\mathrm{miss}}$) serves as an important proxy for the transverse momentum carried by undetected particles originating from the collision event. In Standard Model (SM) processes MET typically serves as a proxy for any neutrinos in the event.

MET is comprised of two components: the component associated with the hard objects in the event, and the component associated with soft objects in the event. The hard component is constructed using fully reconstructed and calibrated particles and objects. The soft component is constructed using charged-particle tracks associated with the hard-scatter vertex [35]. MET is then defined as the negative sum of the components $p_{x(y)}$ of the transverse momentum vectors ($\mathbf{p_T}$) of the hard and soft objects in the event. The MET components in each of the transverse directions $E_{x(y)}^{\mathrm{miss}}$ are defined by,

$$E_{x(y)}^{\mathrm{miss}} = - \sum_{i \,\in\, hard\ objects} p_{x(y),i} - \sum_{j \,\in\, soft\ objects} p_{x(y),j}. \tag{3.6}$$

From this definition the following set of observables that are defined,

$$\mathbf{E_T^{\mathrm{miss}}} = (E_x^{\mathrm{miss}}, E_y^{\mathrm{miss}}), \tag{3.7}$$

$$E_T^{\text{miss}} = |\mathbf{E}_{\mathbf{T}}^{\text{miss}}| = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}, \tag{3.8}$$

$$\phi^{\text{miss}} = \arctan\left(E_y^{\text{miss}}/E_x^{\text{miss}}\right). \tag{3.9}$$

3.11 Upgrades

Since the ATLAS detector first started collecting data in 2008, it has undergone a number of hardware upgrades. In the Long Shutdown 1 (LS1) period, between Run 1 and Run 2, the most significant upgrades were: the insertion of the IBL (Insertable B-Layer), and the reduction of the diameter of the beam pipe at the ATLAS interaction point [19].

The LHC has recently completed Long Shutdown 2 (LS2), where a number of upgrades have taken place in preparation for Run 3. The most significant upgrade was the installation of the New Small Wheel (NSW) which replaced the small wheel, located as the innermost layer of the forward MS that was used throughout Run 1 and Run 2. The NSW is able to trigger on moderate momentum leptons in the high background environment throughout Run 3 and the HL-LHC [36].

After Run 3, upgrades to the ATLAS detector will be performed in anticipation of the high-luminosity LHC (HL-LHC). Significant upgrades include the Inner Tracker (ITk) [37], the optional high-granularity timing (HGTD) detector [38], and upgrades to the ATLAS trigger system [39].

Chapter 4

Particle Showers

Calorimetry is a detection principle used extensively throughout particle physics to measure the energy of electrons, photons, and hadrons. Calorimeters are instrumented materials, designed to absorb incoming particles. The incoming particle interacts with the dense material and produces a *shower* of *secondary* particles. These secondary particles interact with the active material and deposit energy which is detected as either charge or light. The charge or light serves as a measurement for the energy of the incoming particle [40].

In this section, the evolution of electromagnetic (EM) and hadronic interactions of particles in materials is described. This information is key to understanding the work detailed in Part II.

4.1 Electromagnetic Showers

EM showers describe the evolution of energy loss of electrons and photons when traversing dense material. The dominant source of energy loss of electrons with an energy greater than ~ 10 MeV is *bremsstrahlung.* For photons in the same energy range the dominant source of energy loss is *pair-production.* At energies above 1 GeV both of these processes become roughly energy-independent. As a result, an incident electron or photon with sufficient energy will produce secondary photons by bremsstrahlung, or electrons and positrons by pair-production. These secondary particles in turn produce more particles by the same mechanisms. The process is repeated, creating a *shower* of particles. The number of particles in the shower increases until the energy of the electron component falls below a critical threshold. Energy is then lost mainly through ionisation, and does not result in the creation of more particles [40]. An example of the typical evolution of an EM shower is shown in Figure 4.1

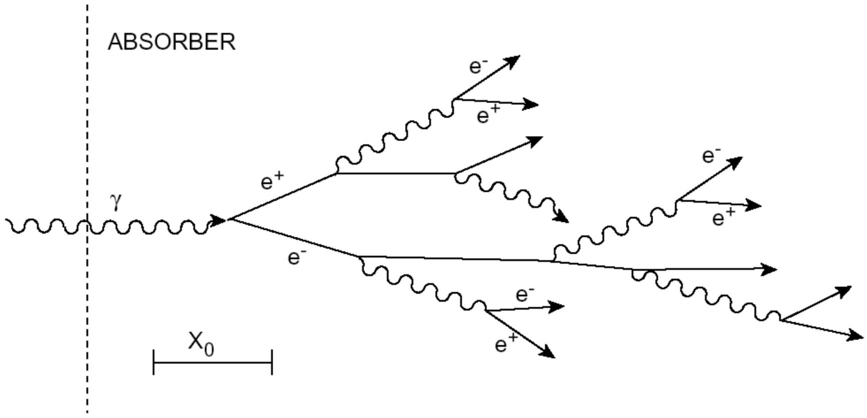


Figure 4.1 *The typical evolution of a EM shower. Taken from [41].*

The characteristic longitudinal and lateral profile of an EM shower in a material can be described by a single parameter, the radiation length X_0.

$$X_0(\mathrm{g/cm^2}) \simeq \frac{716\mathrm{gcm^{-2}}A}{Z(Z+1)\ln{(287/\sqrt{Z})}} \tag{4.1}$$

Where Z is the atomic number of the material, and A is the weight of the material. The radiation length is defined by the rate at which electrons lose energy by Bremsstrahlung. It represents the average distance that an electron needs to travel in a material to reduce its energy to $1/e$ of its original energy E_0.

$$\langle E(x)\rangle = E_0 e^{-x/X_0} \tag{4.2}$$

Similarly, a photon beam is reduced to $1/e$ of its initial intensity after travelling $\frac{9}{7}X_0$ of distance through the material.

$$\langle I(x)\rangle = I_0 e^{-(7/9)(x/X_0)} \tag{4.3}$$

The longitudinal and lateral profile of EM showers of different energies as a function of the radiation length X_0 are shown in Figure 4.2. Where the Molière radius is related to the radiation length by $R_M = 0.0265X_0(Z+1.2)$. The Molière radius is a characteristic constant of a material and gives a useful measure of lateral shower development.

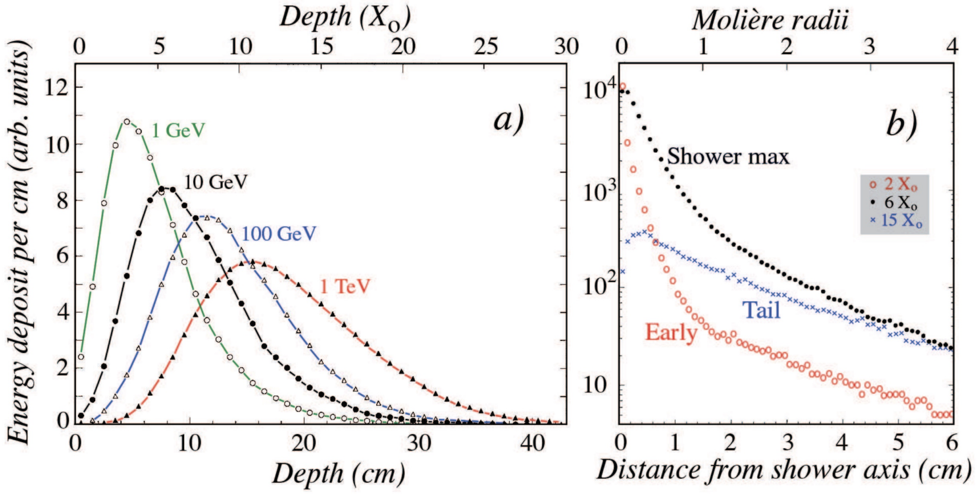


Figure 4.2 *The energy deposited as a function of depth for 1, 10, 100 and 100 GeV electron showers developing in a block of copper. In order to compare the shower profiles, the integrals of these curves have been normalised to the same value (a). The radial distributions of the energy deposited by 10 GeV electron showers in copper, at various depths (b). Results of EGS4 calculations. Taken from [42].*

4.2 Hadronic Showers

Hadronic showers are much more complex than EM showers. This is due to hadronic and nuclear interactions that produce a multitude of effects. Typically an incident hadron will traverse the material, and after an average distance it will interact strongly with the atoms in the material. This average distance is called the *interaction length* and is related to the mass of the material A by the following.

$$\lambda_{int} \approx \frac{35\text{cm}}{\rho} A^{1/3} \tag{4.4}$$

The interaction produces secondary hadrons that typically carry a large fraction of the energy of the incident hadron. Another significant fraction of the incident energy is consumed by nuclear processes. The result of the interaction is a cascade of particles referred to as the *hadronic shower.*

Typically a hadronic shower will have an EM component and a hadronic component. The EM component is initiated by effects such as the creation of electrons and photons from the excitation of atoms, radiation, and the decay of hadrons. The EM component can also be initiated via the decay of neutral hadrons such as the neutral pion decay $\pi^0 \rightarrow \gamma\gamma$. The hadronic component is initiated though effects such as the decay of charged hadrons, the break-up of nuclei, and the production of neutrons [40]. The typical evolution of a hadronic shower is illustrated in Figure 4.3.

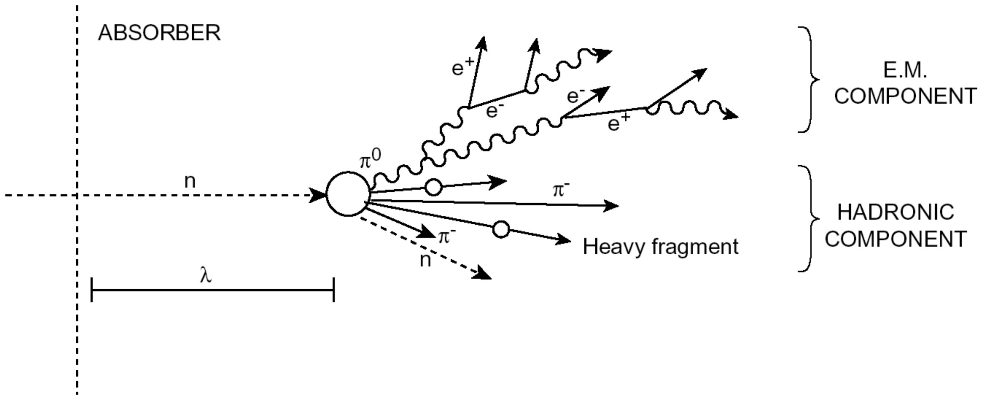


Figure 4.3 *The typical evolution of a hadronic shower. Taken from [41].*

Due to the complex nature of hadronic showers, large fluctuations in the longitudinal and lateral profiles are observed. This can result in a significant fraction of particles created in hadronic showers not being fully contained within a calorimeter, an effect referred to as *punch-through* (Part II).

Chapter 5

Neural Networks

Machine Learning (ML) describes the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyse and draw inferences from patterns in data. It has had profound success in the field of particle physics, and presents exciting prospects for its future [43]. The highly complex problems and data in particle physics make the field uniquely suited to ML algorithms. They are extensively used across nearly all aspects of the ATLAS analysis chain: from the initial trigger, in detector simulation, and to state-of-the-art analyses. This section will cover the foundations of neural networks, a popular type of machine learning algorithm that is used in the work completed for this thesis.

5.1 Neural Networks

Neural networks are a powerful tool in learning non-linear functions and relationships, often achieving better performance than more classical approaches [44]. The name originates from their analogy to connected biological neurons, where "signals" propagate from one neuron to another.

A neural network is constructed from a collection of neurons that are connected by weights w, and are modulated by biases b. Each neuron has an activation value that is determined by these weights and biases. Neurons are arranged in layers, where all neural networks must contain an *input layer* and an *output layer*, and generally contain a number of *hidden layers*. An illustration of two fully connected neural network layers is shown in Figure 5.1. The Figure also details how the activation values of the neurons in one layer depend on the weights, biases, and activations in the previous layer.

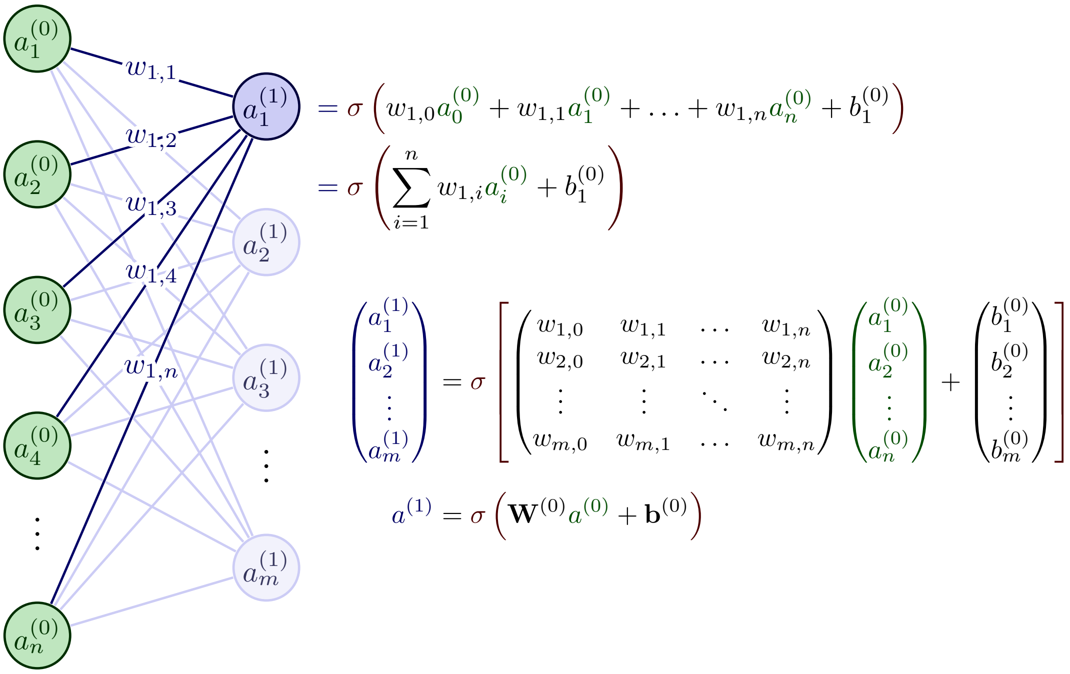


Figure 5.1 *Illustration of two fully connected neural network layers, showing how the activation values of each neuron in layer 1 depend on the weights, biases, and activations in layer 0. Taken from [18].*

The complete network maps an input feature space $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ to an output space $\boldsymbol{y} = (y_1, y_2, ..., y_m)$ by propagating values through the network of weights and biases. To allow the network to apply non-linear mappings between $\boldsymbol{x}$ and $\boldsymbol{y}$, a non-linear activation function $\sigma(a)$ is applied to the activation of the nodes in the hidden layers of the network.

5.1.1 Learning Algorithm

Data is fed through a neural network, using the data a learning algorithm trains its behaviour. There are two steps in the learning algorithm, these are called, back-propagation and optimisation.

Back-propagation refers to the process of updating the weights and biases in the network. In supervised learning, the weights and biases are updated so that the prediction of the neural network more closely matches its target value. Closeness to the target value is measured using a loss function. As an example, the Binary

Cross-Entropy loss function, used in binary classification tasks, is detailed. The BCE loss and is given by the following equation,

$$BCE = \frac{1}{n}\sum_{i=1}^{n} y_i \log(p(y_i)) + (1 - y_i)\log(1 - p(y_i)) \tag{5.1}$$

where n represents all of the data samples considered when calculating the loss, y_i is the true class of each sample (either 1 or 0), and $p(y_i)$ is the prediction of the neural network.

During training, the loss is computed over each batch of data. The gradient of the weights and biases are calculated with respect to the loss. The *optimiser* algorithm guides the neural network model on how to update the weights and biases using this gradient information. This process is repeated on consecutive batches of data, each time the weights and biases are updated to improve the models prediction. After the neural network has seen all batches of data the process is repeated, each repetition of the full dataset is called an epoch. The training cycle continues until the model is terminated, or reaches some user-defined performance threshold.

5.1.2 Setup

There are many components that enter the setup of a neural network. Specific choices are influenced by the nature of the data, objective, and often trial and error. The following detail some of the commonly considered components and best practices of setting up a neural network model.

Data is divided into training, validation and test sets. The model is exclusively trained on the training dataset and validated after each epoch using the validation dataset. Finally, when training is complete, the model performance is tested using the test dataset. This approach allows for the checking of under/overfitting of the model, and any biases.

Feature Scaling is applied to the input features of the neural network model to ensure that they span the same order of magnitude. This helps the neural network learn. In the neural network model developed in Chapter 7 a *min-max* feature scaling is used that linearly scaled features to lie in a defined range.

Class Weights are used when there is an imbalance in the number of samples of each class in the data. The weights enter the loss calculation, ensuring the contribution of each class is equal.

Batch Normalisation is often used as a pre-processing step before each layer in the neural network. It ensures that the inputs to each node are normalised to the same order of magnitude. This helps the neural network learn and avoid the exploding gradient problem.

Hidden Layers are used in neural network models. The term refers to any layers of the model between the input and output layers. Neural network models with hidden layers are considered "deep". The hidden layers, along with the hidden activation allow the model to learn complex multi-dimensional relationships.

Hidden Activation functions are applied to the hidden nodes of the neural network. They apply a non-linear transformation to the activation value of the node. The hidden activations allow the neural network to learn non-linear relationships.

Final Layer Activation functions are applied to the output nodes of the neural network. The choice of activation function depends on the objective of the neural network. For binary classication models the *sigmoid* activation is typically used, for regression models the *linear* activation is typically used.

The Optimiser is responsible for updating the weights and biases in the model using the gradient information. The ADAM (Adaptive Moment Estimation) [45] optimiser is the current most popular choice because of how well it performs. It uses momentum information information from past training iterations to inform updates to the weights and biases.

Regularisation techniques are used to try and avoid overfitting. An example is the *Dropout* regularisation, used in the neural networks developed in this thesis.

During training a random set of nodes, and associated weights and biases, are excluded from the back-propagation step. This forces the neural network to use a different set of nodes, and can help it to generalise.

Callbacks are used to perform actions after each training step (typically an epoch). Examples are: the *ReduceLearningRateOnPlateau* callback which reduces the learning rate of the optimiser if the loss has not decreased after a defined number of training steps, and the *EarlyStopping* callback that stops the training of the neural network if if the loss has not decreased after a defined number of training steps. Both are used in the neural networks developed in this thesis.

Hyperparameter Tuning is often conducted when fine-tuning the performance of a neural network model. Hyperparameters of the neural network, such as the number of nodes, the number of layers, and the hidden activation function are optimised. A Bayesian hyperparameter tuning was conducted for the neural network models detailed in Chapter 9.

Part II

A Fast Simulation Solution for the Muon Punch Through Effect

Chapter 6

Simulation in ATLAS

Analyses of data collected by the ATLAS experiment rely on the accurate simulation of physics processes. This includes simulation of the initial hard scatter event of interest, particle fragmentation and decay, detector interaction, and detector response. ATLAS employs a chain of steps that simulate each of these components.

This section will give an overview of the ATLAS simulation scheme, from event generation, through to digitisation. It will then focus on the ATLAS fast simulation tool, AtlFastIII (AF3) [2].

6.1 ATLAS Simulation Scheme

The simulation of physical processes that occur in the ATLAS experiment is handled using a chain of steps. This allows each component to be factored out and configured separately. An overview of the simulation chain is given in the following points.

Event Generation is in the following at the core of the simulation chain, it is responsible for modelling the fundamental interactions that take place in p-p collisions at the LHC. Theoretical calculations of the underlying process are used to construct a matrix element. These are typically calculated at LO or NLO accuracy. Using the matrix element a set of particles, and their corresponding momenta are generated. Particles are then decayed or hadronised. The final outgoing set of particles and their momenta are then passed on to the detector simulation step.

Detector Simulation is responsible for modelling the interaction of particles as they pass through the detector components. In the case of *full*-simulation, particles take steps through the detector volumes. On each step interactions between the particle and the detector components are modelled, contributing to energy loss, momentum deflection, and the creation of secondary particles. This approach to simulation is referred to as Monte-Carlo (MC). ATLAS uses the Geant4 simulation toolkit [46] for the *full*-simulation of the ATLAS detector. *Full*-simulation is the most accurate but also the slowest detector simulation option. *Fast*-simulation however, models the response of the detector using a parametric solution. Modelling is less accurate, however, computing time is significantly reduced. The output of the detector simulation is the raw energy deposited in each detector component.

Digitisation is the process where the response of the detector electronics is simulated. The simulated data is converted into a format corresponding to that of real data read out by the ATLAS DAQ system. The digitised output is then passed through the same reconstruction steps as real data, as detailed in Section 3.10.

6.2 AtlFastIII (AF3)

AtlFast is a fast detector simulation solution developed for the ATLAS experiment. It offers a fast and computationally cheap alternative to Geant4. Reducing computational cost is important as for Run3 and the High-Luminosity LHC, where computing resources become a limiting factor.

Throughout the Run2 phase of the LHC, the simulation of physics processes required approximately 40% of the computing resources available to the ATLAS experiment [47].

The vast majority of the computing time is spent simulating interactions in the calorimeters. This is due to the intricate accordion structure of the calorimeters,

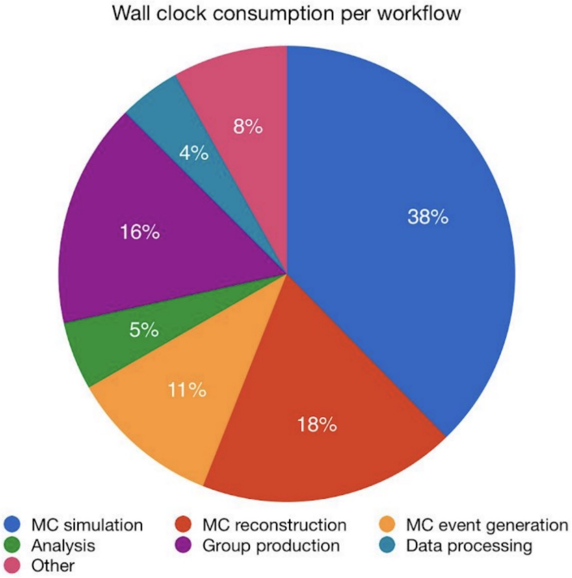


Figure 6.1 *Percentage of CPU time spent on different components of the ATLAS analysis chain in 2018. Taken from [47].*

and the complex development of electromagnetic and hadronic showers. In the simulation of the top anti-top pair production, 80% of the simulation time is spent modelling shower development in the calorimeters [48].

The latest version of AtlFast, AF3, is composed of different components. Each each deal with the simulation of particles in different regions of the detector. Figure 6.2 illustrates how each of these components come together to form the full AF3 detector simulation package.

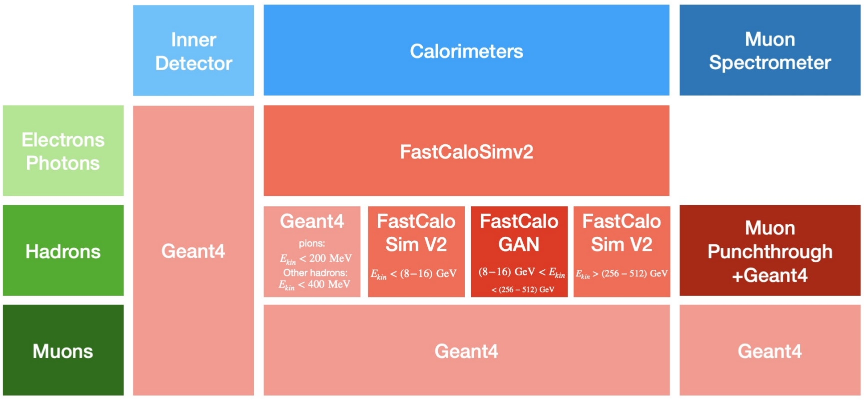


Figure 6.2 *The configuration of the different simulation components that comprise AF3. Taken from [2]*

The simulation of different particle types are handled separately in AF3. A different simulation scheme is used for electrons and photons, hadrons, and muons.

Geant4 is used to simulate the interactions of all particle types in the inner detector - simulation in the inner detector is far less computationally expensive than the calorimeters. Geant4 is also used to simulate muons throughout all detector components, because they have a minimal interaction cross-section.

For the simulation of electrons, photons, and hadrons in the calorimeters, two models are used, these are FastCaloSimV2 and FastCaloGAN. The simulation of electrons and photons is handled entirely by FastCaloSimV2, and a combination of FastCaloSimV2, FastCaloGAN, and Geant4 is used for the simulation of Hadrons. FastCaloSimV2 uses a parameterised approach to model the calorimeter response, whereas FastCaloGAN uses a deep learning approach using Generative Adversarial Networks. The final component of AF3 is the muon punch through (MPT) simulation. The development of the MPT simulation is covered in detail in Chapter 7.

AF3 is able to achieve excellent simulation accuracy of electrons, photons and hadrons. It makes important improvements in the modelling accuracy of physics variables when compared to AtlFastII (AF2). Specifically the modelling of high energy jets is vastly improved, this is achieved with the addition of FastCaloGAN and improvements in the techniques used in FastCaloSimV2. AF3 also now models the muon punch-through effect using the simulation methods detailed in Chapter 7. The modelling of punch-through was not present in AF2. Details of these modelling improvements can be found in [2].

AF3 also significantly reduces simulation time when compared to full simulation. Figure 6.3 shows the average CPU time spent simulating single photons at three different energies using Geant4, AF2, and AF3, in the ATLAS calorimeters. It can be seen that for an 8 GeV photon AF3 is approximately 20 times faster than Geant4, and for a 256 GeV photon, approximately 600 times faster. When simulating all of the ATLAS detector components, the AF3 simulation results in an approximate 5 times speed up when compared to Geant4 [2].

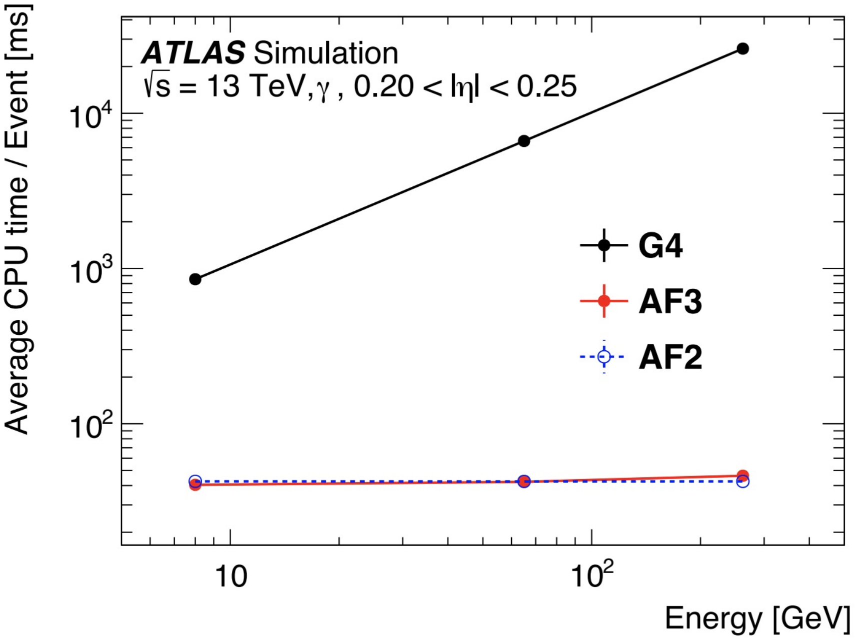


Figure 6.3 *Comparison of the CPU performance of AF3 with Geant4 and AF2. The average CPU time to simulate an event is estimated using 10 000 single photons at* $0.20 < |\eta| < 0.25$ *for three different energies: 8 GeV, 65 GeV, and 256 GeV. These photons are generated on the calorimeter surface and provide a comparison for calorimeter-only simulation time. Taken from [2].*

Chapter 7

Fast Simulation of the Muon Punch-Through Effect

This chapter presents the muon punch through (MPT) simulation component of AtlFastIII (AF3), which was developed for this thesis. The chapter covers aspects such as dataset choice, event selection, and the statistical techniques that are used to model the MPT effect. Excellent modelling accuracy is observed, and is summarised in the results (Section 7.5).

The work here is based on that detailed in [49], but improved and redesigned for the requirements of AF3 and Run 3 of the Large Hadron Collider (LHC).

7.1 The Muon punch-through Effect

The term MPT describes the effect where secondary particles created in particle showers in the ATLAS calorimeters, aren't fully contained. The escaping particles are named *punch-though* particles. The *punch-though* particles travel into the Muon Spectrometer (MS) leaving track signatures that can be missidentified as muons (Figure 7.1). Punch-through also impacts the energy calibration of high-energy jets. The punch-through effect is therefore important to model in the AF3 simulation.

The punch-through effect is almost completely dependent on hadronic showers in the calorimeters. This is because hadronic showers are much more elongated than electromagnetic (EM) showers, and fluctuations in shower depth are larger [42] (Chapter 4).

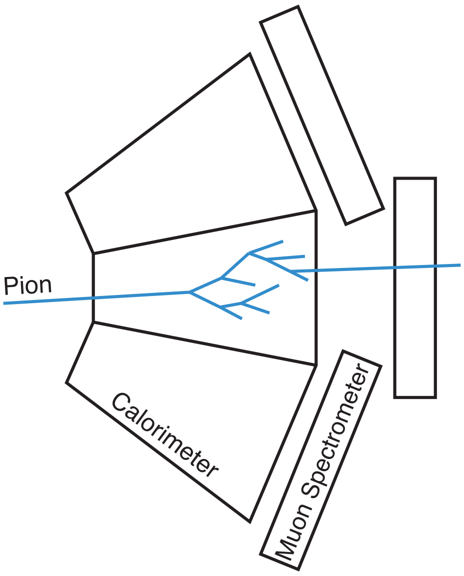


Figure 7.1 *Schematic illustrating the MPT effect.*

The ATLAS calorimeters were designed taking into account the trade-off between hadronic shower containment, and detector size and cost. Figure 7.2 shows the results of studies that were completed for the ATLAS experiment. The energy deposited by pions as a function length of the hadronic Tile Calorimeter was investigated. It can be seen that the longitudinal shower depth is dependent on the energy of the incoming pion. A significant fraction of the energy of the incoming pion is deposited beyond 10 interaction lengths.

Figure 7.3 shows the material budget of the ATLAS calorimeters. It can be seen that there is a significant dependency on η, and typically there is ~ 10 interaction lengths of calorimeter material before the MS ($|\eta| < 3.0$). It is therefore expected that a significant punch-through will arise from high-energy hadrons.

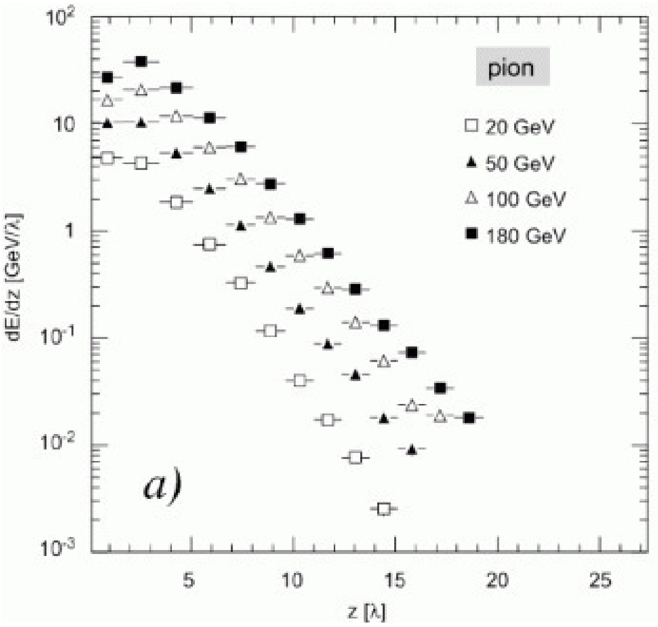


Figure 7.2 *Average longitudinal shower profiles of pions of different energies in the iron/plastic-scintillator calorimeter used in ATLAS. Taken from [42].*

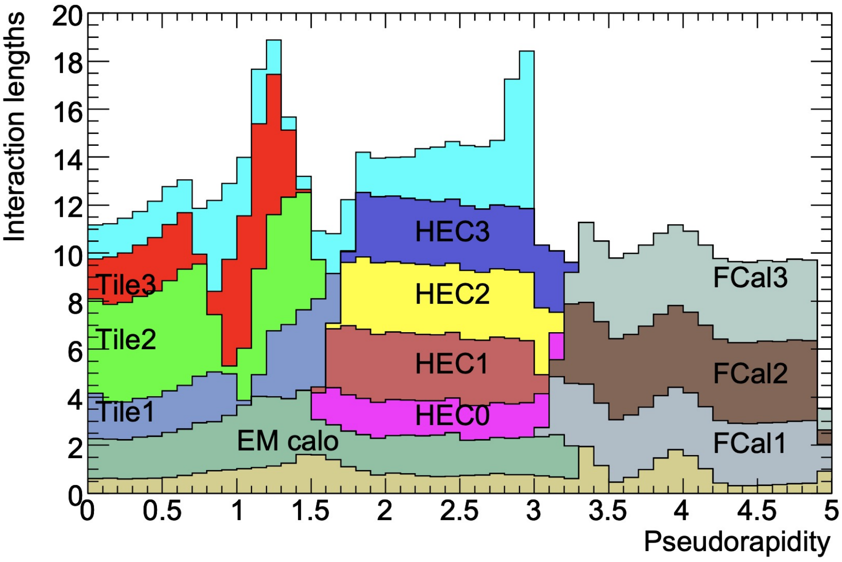


Figure 7.3 *Cumulative amount of material, in units of interaction length, as a function of $|\eta|$, in front of the electromagnetic calorimeters (off-yellow), in the electromagnetic calorimeters (labelled), and in the hadronic calorimeters (labelled). Also shown for completeness is the total amount of material in front of the first active layer of the muon spectrometer (light-blue) (up to $|\eta| < 3.0$). Taken from [16].*

7.2 Punch-Through Parameterisation

The development of the AtlFast fast simulation tool is derived from fully-simulated Geant4 samples. These samples are used to determine the energy response of different particles inside the ATLAS calorimeters, and to derive the parameterisation of the MPT effect. This section describes the analysis of these Geant4 samples, where different aspects of the MPT effect are investigated. The analysis is then used to derive a parameterisation of the effect for use in AF3 (Section 7.3).

7.2.1 Input Samples and Event Selection

A set of Geant4-simulated reference samples are used to create the AF3 calorimeter and MPT parameterisations. Single particle events are generated at the entrance to the calorimeters, equating to the surface of a cylinder with $r = 1148$ mm and $|z| \leq 3550$ mm. This approach allows for the examination of the energy response inside the calorimeters, without considering any effects of the inner detector (ID). All detector components affecting shower development in the calorimeters are considered, including the solenoid magnet and cryostat. A detailed list of settings for the simulation of these reference samples can be found in [2].

The initial momentum direction of the single particles is chosen to be consistent with a particle that originated from the interaction point (IP). The impact of the spread in the z direction of beam bunches is negligible. The single particles are simulated with a uniform distribution in ϕ and η. Single-photon (γ) and electron ($e^{\pm}$) samples are simulated to model electromagnetic showers, and single-pions ($\pi^{\pm}$) are simulated to model hadronic showers [2].

The single-particle samples are divided based on η and initial particle momentum. At each initial particle momentum, samples are simulated in 100 slices of $|\eta|$, where each slice covers a pseudorapidity range of 0.05. The 100 samples in η are simulated at a number of discrete initial particle momenta in the range of 16 MeV to 4.2 TeV. Intermediate momenta are derived using equal spacing on a logarithmic scale. The segmentation allows for different modelling in different momentum regimes and $|\eta|$ detector regions. The result is a total of 5100 samples. Ten thousand events are generated for each of these samples up to a momentum of 256 GeV, and 1000 events are generated above 256 GeV. This is due to the significant simulation time needed for the higher-momentum samples.

Because only hadronic showers result in a considerable punch-through, the MPT effect is parameterised using only the single-pion samples. Samples in the input momentum range 8 GeV - 4.2 TeV are used, as there is minimal punch-through below 8 GeV. Samples in the pseudorapidity range $|\eta| \leq 3.2$ are considered, allowing for full coverage of the MS ($|\eta| \leq 2.7$). The result is 640 samples used for the MPT parameterisation.

7.2.2 Particle Selection

The *MuonEntryLayer* is introduced in the Geant4 simulation of the ATLAS detector as a cylindrical surface between the hadronic calorimeter and the MS. The MuonEntryLayer records the position, momentum, energy, and Particle Data Group particle numbering scheme (PDGId) of any particle that passes through it. Particles recorded by the MuonEntryLayer are labeled *Punch-Through Particles*. They are required to have their momentum direction towards the MS, this ensures that back-scattered particles aren't considered.

7.2.3 Punch-Through Frequencies

An important component to understand in the MPT effect is the frequency at which certain events occur. The following are considered: the rate at which punch-through leads to a significant signal in the MS, and the frequency and type of Punch-Through Particles that exist.

Events that result in at least one reconstructed muon segment in the MS are deemed significant. Figure 7.4 shows the probability that a single-pion event leads to at least one reconstructed muon segment as a function of the incoming pion momentum and pseudorapidity. As can be seen, there is a strong dependency on the incoming pion momentum and $|\eta|$.

Figure 7.5 shows the frequency of punch-through particle types. The most frequent punch-through particles are neutrons, protons, and pions, however there is also a significant contribution from electrons, muons, photons, and kaons ($\mathrm{K}^+, \mathrm{K}^0, \mathrm{K}^0_\mathrm{S}, \mathrm{K}^0_\mathrm{L}$).

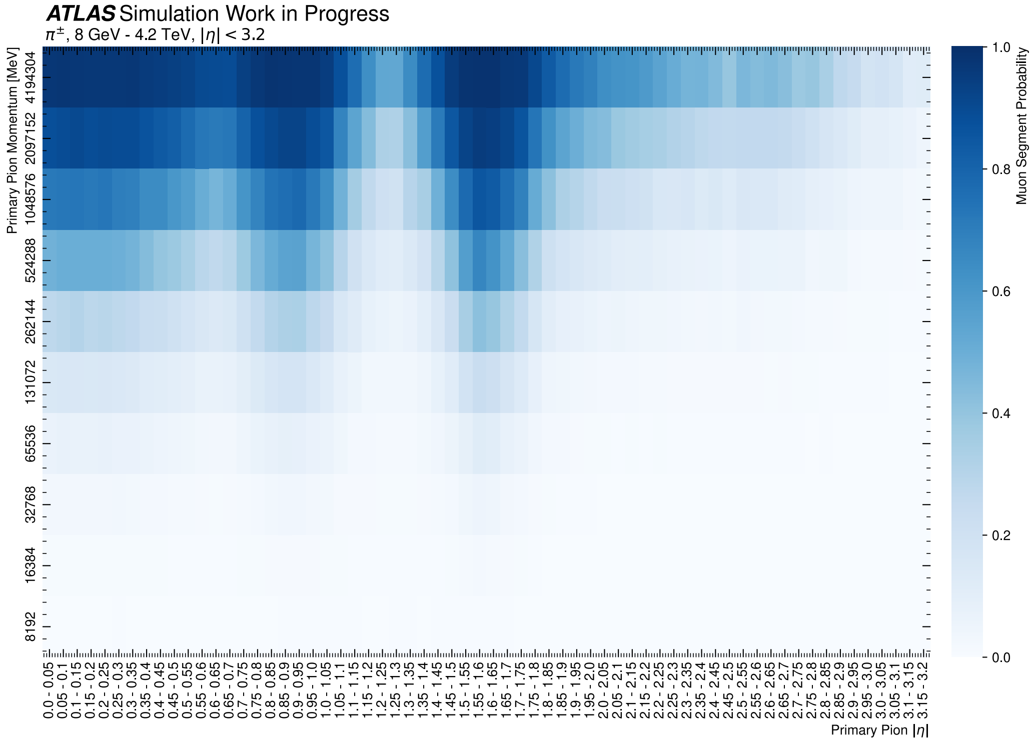


Figure 7.4 *The probability that an incoming single-pion creates at least one reconstructed muon segment in the MS as a function of the initial momentum and* $|\eta|$ *of the incoming pion.*

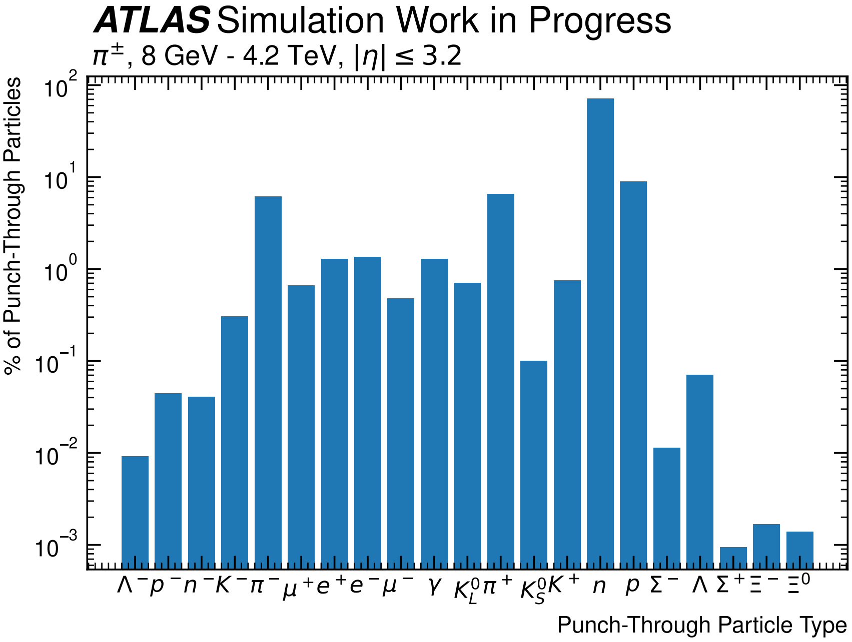


Figure 7.5 *The frequency of each punch-through particle type originating from single-pion events with an momentum of 64 GeV - 4.2 TeV and in the pseudorapidity range* $|\eta| \leq 3.2$.

Figure 7.6 shows the multiplicity of punch-through neutrons originating from an initial pion with $0.00 \leq |\eta| \leq 0.40$ and an momentum of 262.144 GeV. It can be seen that a significant number of events result in a high multiplicity of punch-through neutrons. Similar distributions are observed for the other punch-through particle types.

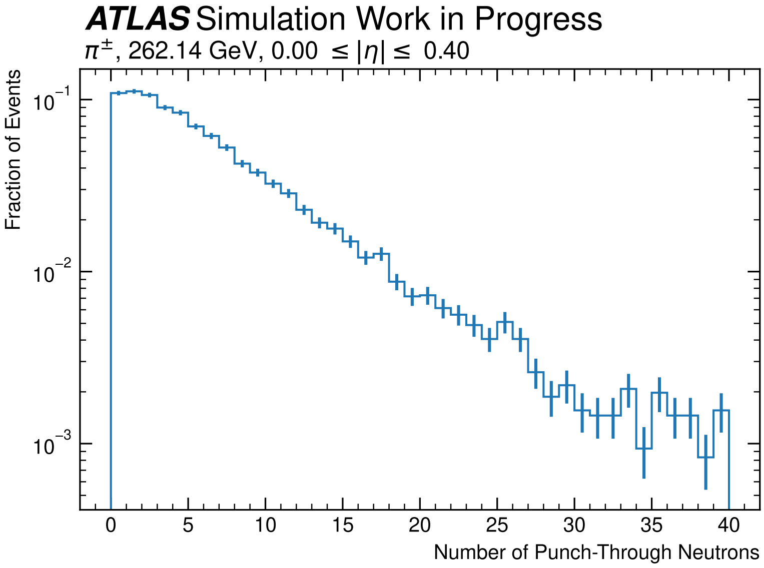


Figure 7.6 *The multiplicity of punch-through neutrons in single-pion events with an incoming momentum of 262.14 GeV and in the eta range* $0.00 \leq |\eta| \leq 0.40$.

7.2.4 Kinematics of Punch-Through Particles

Five variables are chosen to describe the punch-through particles, these are: energy, positional deflection in the θ and ϕ directions, and momentum deflection in the θ and ϕ directions.

Particle Energy

Figure 7.7 shows the energy distribution of punch-through neutrons originating from single-pion events with an incoming momentum of 262.14 GeV in the pseudorapidity range $0.00 \leq |\eta| \leq 0.40$.

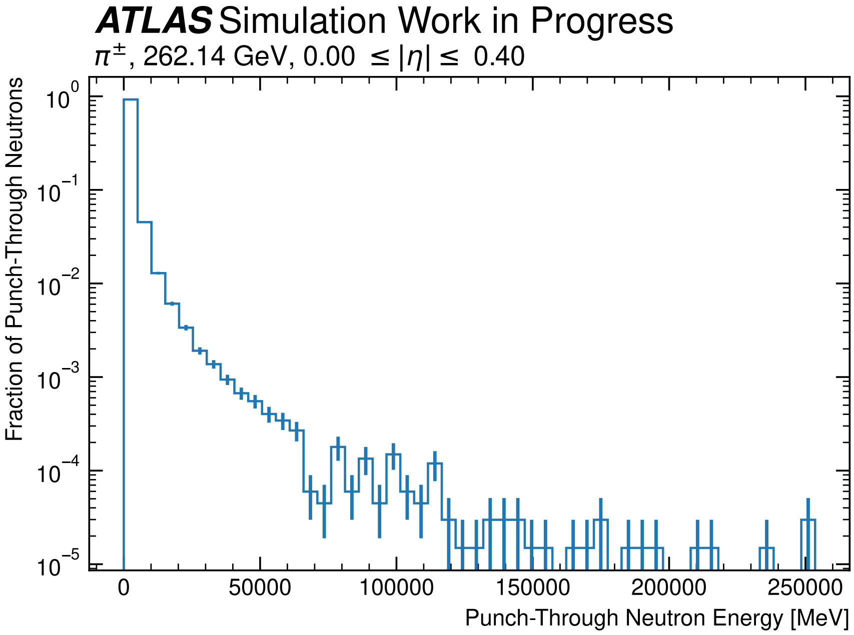
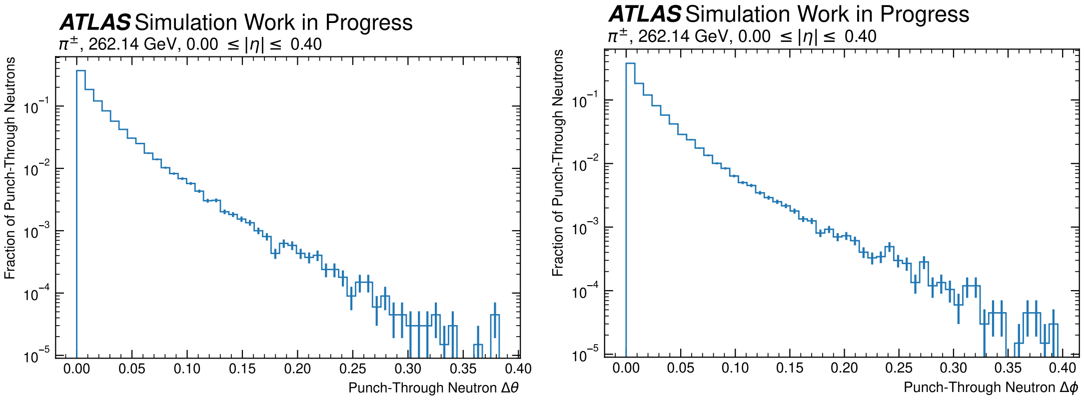


Figure 7.7 *The energy of punch-through neutrons in single-pion events with an incoming momentum of 262.14 GeV and in the eta range* $0.00 \leq |\eta| \leq 0.40$.

Position Deflection Angles

Two deflection angles are defined, $\Delta\phi$ and $\Delta\theta$, for each angular direction in the detector coordinate system. Two angles are used because of the complex structure of the ATLAS calorimeters, where there may be a different response in each direction. Figure 7.8 shows the $\Delta\theta$ and $\Delta\phi$ deflection angles of punch-through neutrons originating from an initial pion with $0.00 \leq |\eta| \leq 0.40$ and momentum of 262.14 GeV.

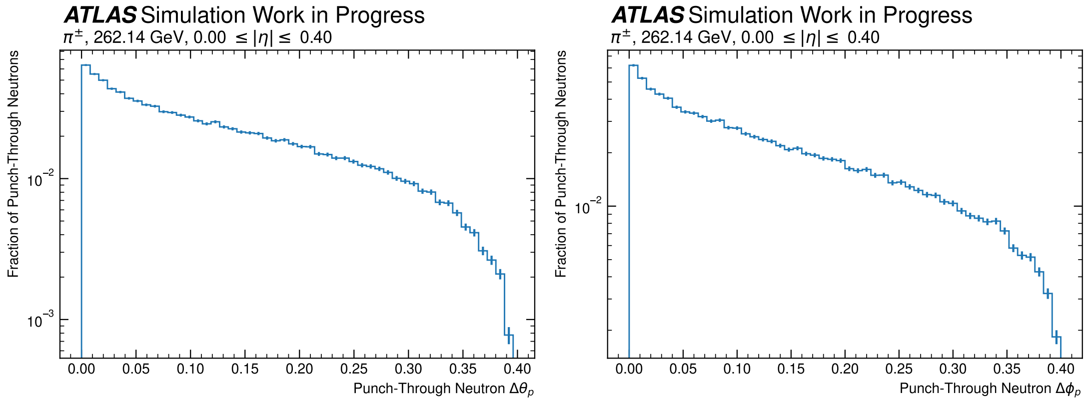


(a) **(b)**

Figure 7.8 *The deflection angles of punch-through neutrons in single-pion events with an incoming momentum of 262.14 GeV and in the pseudorapidity range* $0.00 \leq |\eta| \leq 0.40$*. Deflection angle in the* θ *direction (a), deflection angle in the* ϕ *direction (b).*

Momentum Deflection Angles

Similarly, two momentum deflection angles are defined, $\Delta\theta_p$ and $\Delta\phi_p$. These are the difference in θ and ϕ between the punch-through particle position on the MuonEntryLayer surface, and the direction of the punch-through particle momentum. Figure 7.9 shows the $\Delta\theta_p$ and $\Delta\phi_p$ momentum deflection angles of punch-through neutrons originating from an incoming pion with $0.00 \leq |\eta| \leq 0.40$ and a momentum of 262.14 GeV.



(a) **(b)**

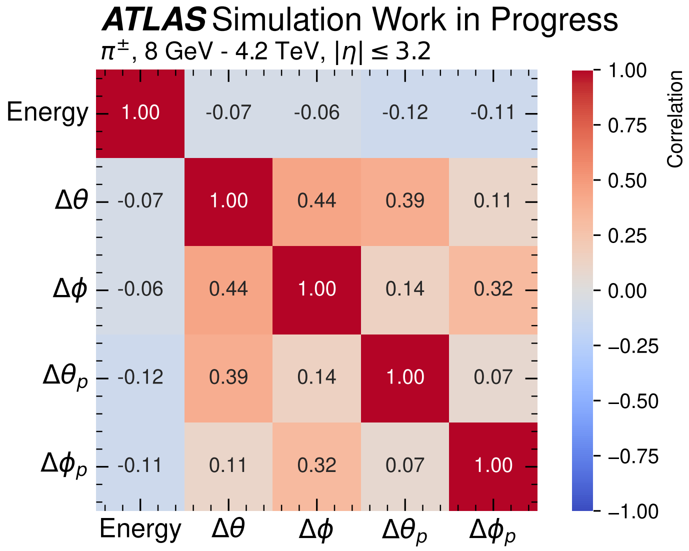
Figure 7.9 *The deflection angles of punch-through neutrons in single-pion events with an incoming momentum of 262.14 GeV and in the pseudorapidity range* $0.00 \leq \eta \leq 0.40$*. Momentum deflection angle in the* θ *direction (a), momentum deflection angle in the* ϕ *direction (b).*

7.2.5 Correlations

Two types of correlations are considered, these are: correlations between the kinematics used to describe each punch-through particle, and the correlation of the punch-through effect to the energy deposited in the calorimeter by the underlying hadronic shower.

Kinematics Correlations

Figure 7.10 shows the correlation between the five kinematic variables that were chosen to describe each punch-through particle. The correlations are calculated considering all punch-through particles produced by single-pions with input momentum of 8 GeV - 4.2 TeV and in the pseudorapidity range $|\eta| \leq 3.2$. Significant correlations can be seen: large and positive between deflection angles, and small and negative between the punch-through particle energy and each deflection angle.



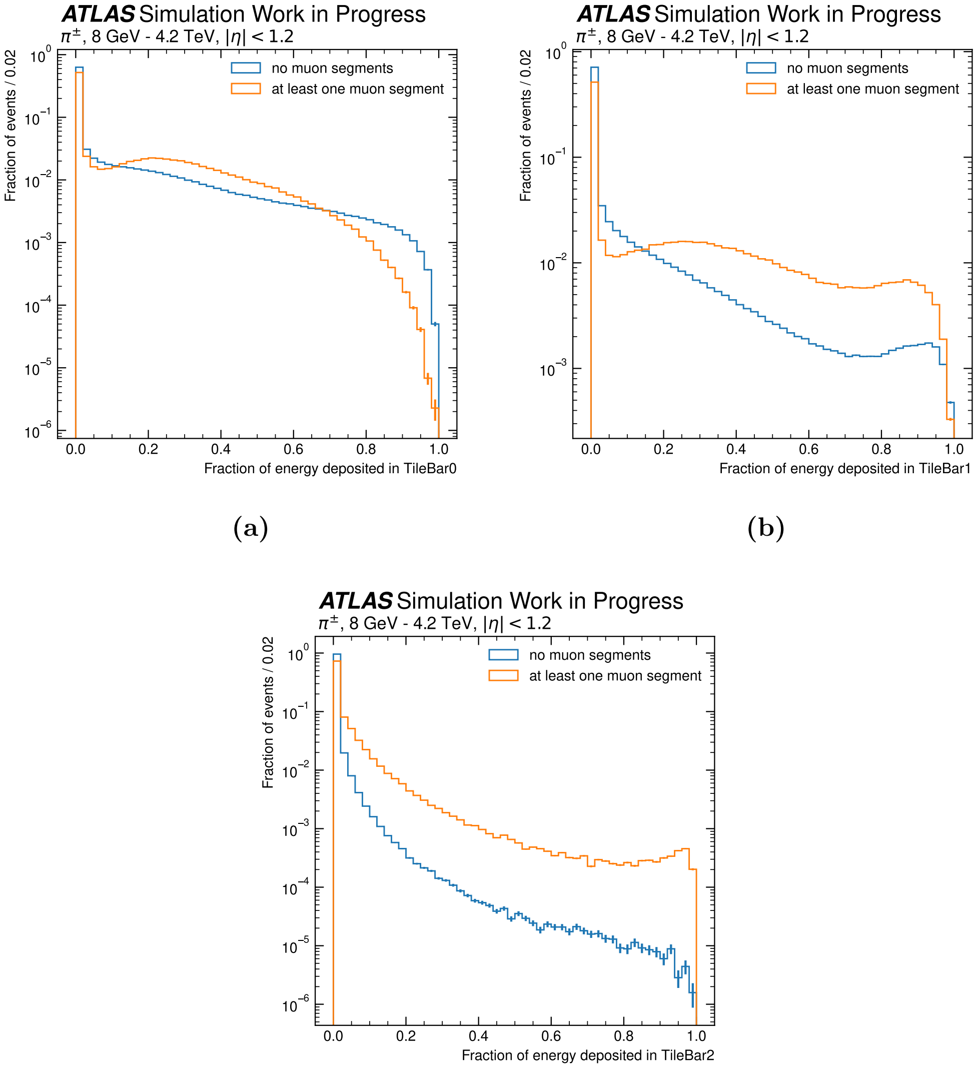
	Energy	$\Delta\theta$	$\Delta\phi$	$\Delta\theta_p$	$\Delta\phi_p$
Energy	1.00	-0.07	-0.06	-0.12	-0.11
$\Delta\theta$	-0.07	1.00	0.44	0.39	0.11
$\Delta\phi$	-0.06	0.44	1.00	0.14	0.32
$\Delta\theta_p$	-0.12	0.39	0.14	1.00	0.07
$\Delta\phi_p$	-0.11	0.11	0.32	0.07	1.00

Figure 7.10 *The correlation between each of the kinematics of punch-through particles originating from single-pions in the momentum range 8 GeV - 4.2 TeV and in the pseudorapidity range* $|\eta| \leq 3.2$.

Correlations to The Underlying Shower

The punch-through effect is correlated to the original hadronic shower inside the calorimeters. For example, a hadronic shower that penetrates deeper into the calorimeter is more likely to result in a significant punch-through. Because of the large fluctuations in the longitudinal depth of hadronic showers, the magnitude of the punch-through effect can be significantly different for identical hadrons. The correlation of the punch-through to the hadronic shower is therefore important to model.

Each of the ATLAS calorimeters are arranged in layers. The energy deposited in each layer gives an indication of the longitudinal depth of a hadronic shower. Figure 7.11 shows the fraction of energy deposited in the three layers of the hadronic barrel calorimeter for single-pion events. The fraction of energy deposited is shown for events that result in no reconstructed muon segments and at least one reconstructed muon segment. It can be seen that events that result in at least one reconstructed muon segment are more likely to deposit more energy in the outermost barrel calorimeter layers, TileBar1 and TileBar2.



(a) **(b)**

(c)

Figure 7.11 *The fraction of energy deposited in the (a) Tile Barrel 0, (b) Tile Barrel 1, (c) Tile Barrel 2 calorimeter layers in single-pion events with an initial momentum of 8 GeV - 4.2 TeV in the barrel region* ($|\eta| \leq 1.2$).

7.3 Parameterisation

The punch-through parameterisation is made up of three main components, these are: the punch-through classifier, the kinematics PCA, and the punch-through paramterisation file. The punch-through classifier is a calibrated neural network model that captures the correlation between the energy deposited in the calorimeter layers, and the punch-through effect. The kinematics PCA rotates the kinematic variables of the punch-through particles into a new basis, allowing linear correlations to be captured. The paramterisation file acts as a *lookup-table*, storing the frequency and kinematic PCA components of the Punch-Through Particles. This section details the implementation of these three components.

Punch-through neutrons, pions, protons, electrons, muons, photons and kaons are considered in the MPT parameterisation. These particles are selected due to their frequency, and potential to leave a signature in the MS.

7.3.1 Punch-Through Classifier

To capture the correlation between the underlying shower and MPT effect, a calibrated neural network model is trained. The model tries to predict whether a single-pion event will result in at least one reconstructed muon segment. It uses information about the incoming pion, and the amount of energy deposited in each calorimeter layer. The calibrated neural network prediction is used as a proxy for a probability inside the simulation.

Setup

The Keras [50] package is used to develop the neural network model. Single-pion events are given binary labels indicating whether they result in at least one reconstructed muon segment or not. The full set of single-pion data in the range 8 GeV - 4.2 TeV and $|\eta| \leq 3.2$ is used to train, validate, test, and calibrate the model. The data is first shuffled and then split into these exclusive sets. 28% of the data is used to train the model, 12% for validation, 30% for testing and the final 30% to calibrate the model.

Twenty-seven input variables are used, these are: the momentum, the absolute η, and the ϕ of the incoming pion, the total energy deposited in all of the calorimeter layers, and the energy fraction deposited in each of the 23 calorimeter layers. All input variables are scaled so that they lie in the range 0.2 to 0.8, using a *min-max* scaling procedure.

The model consists of four hidden layers, the first three each have 54 nodes, and the final hidden layer has 10. The hidden nodes of the network use the LeakyReLu activation function with an α of 0.3. The final layer uses the sigmoid activation function.

Training

The Binary Cross-Entropy (BCE) loss function, the ADAM optimiser with a learning rate of 0.01, and a batch size of 256 are used when training the network. The *ReduceLearningRateOnPlateau* and *EarlyStopping* callbacks are used to ensure that the same minimum is reached each time the network is trained. The *Dropout* regularisation technique is used with a probability of 0.05 to avoid overfitting the model.

Figure 7.12 shows the training and validation loss of the network as a function of the training epoch. It can be seen that the model reaches a minimum after around 150 epochs where the *EarlyStopping* callback terminates training. The validation loss gets less and less noisy as the training progresses, this is the result of using the *ReduceLearningRateOnPlateau* callback. It can also be seen that the validation loss is consistently less than the training loss, indicating that the model is not overfitting.

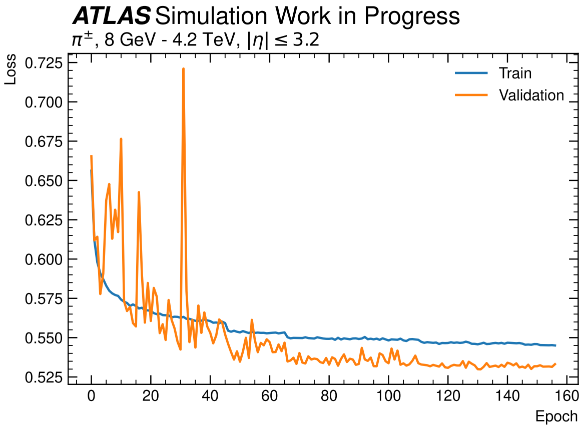
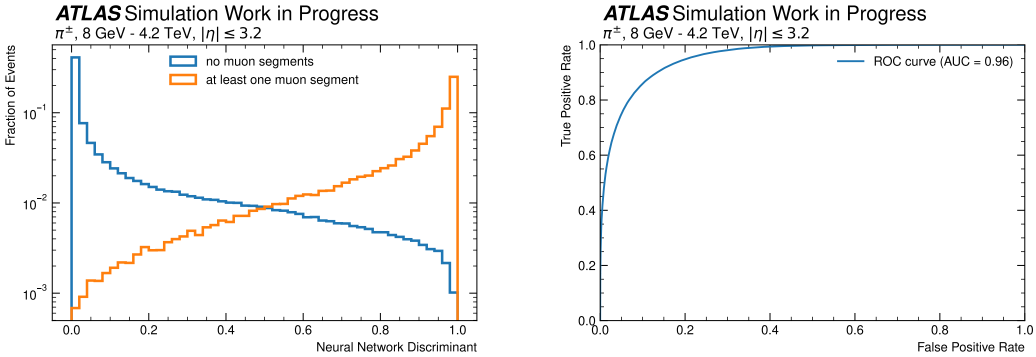


Figure 7.12 *Training and validation loss as a function of training epoch for the punch-through classifier.*

Performance

The performance of the trained model is evaluated using the ROC curve, and the distributions of each class in the neural network discriminant (Figure 7.13). It can be seen that the neural network is able to separate the two classes very well. This is reflected in the AUC value of 0.96.



(a) **(b)**

Figure 7.13 *Metrics to monitor the performance of the punch-through classifier neural network model. Class separations in the neural network discriminant (a). ROC curve (b).*

Calibration

Currently, the neural network prediction is not guaranteed to be a good proxy for probability. This can be tested using a *calibration curve*. Figure 7.14 shows the calibration curve for the neural network model. It can be seen that the model tends to overpredict the probability that an event should result in at least one reconstructed muon segment. For a perfectly calibrated model, the curve will exactly follow the diagonal.

An isotonic regression model is used to calibrate the output score of the neural network. The isotonic regressor maps the output of the uncalibrated network so that it closely matches the true probability of each class. Figure 7.15 shows the calibration curve for the calibrated neural network model. It can be seen that the model now very closely follows the diagonal indicating that it is well-calibrated.

Figure 7.16 shows the ratio between the average calibrated probability and the average true probability for each single-pion momentum and η. In the detector regions most likely to have punch-through, there is good agreement,

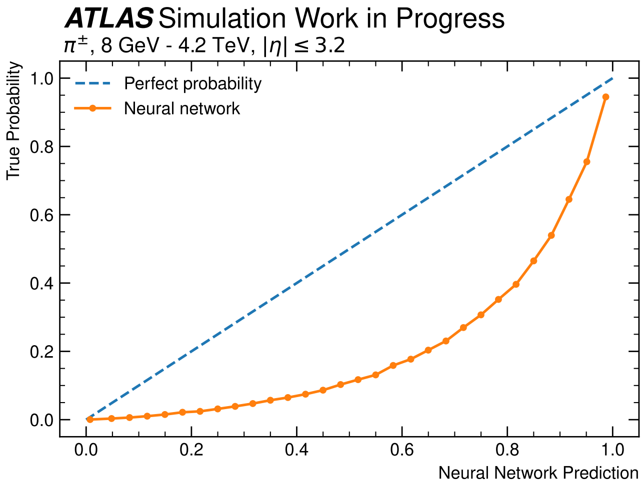


Figure 7.14 *Calibration curve of the punch-thorugh classifier neural network output before calibration.*

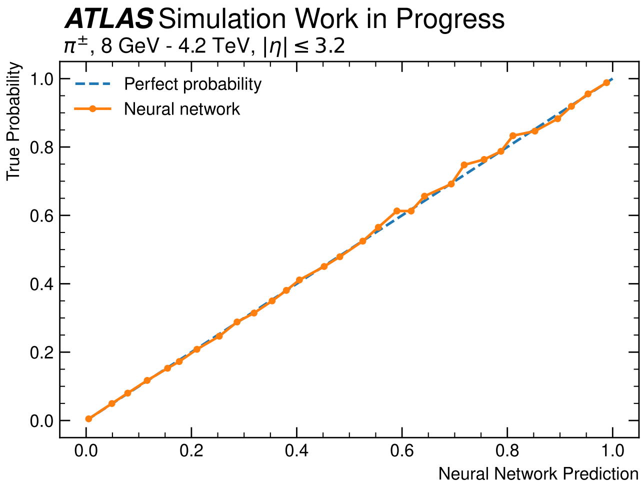


Figure 7.15 *Calibration curve of the punch-through classifier neural network output after calibration.*

with a maximum deviation of 25%. Modelling is worse in the regions where the probability of punch-through is lower.

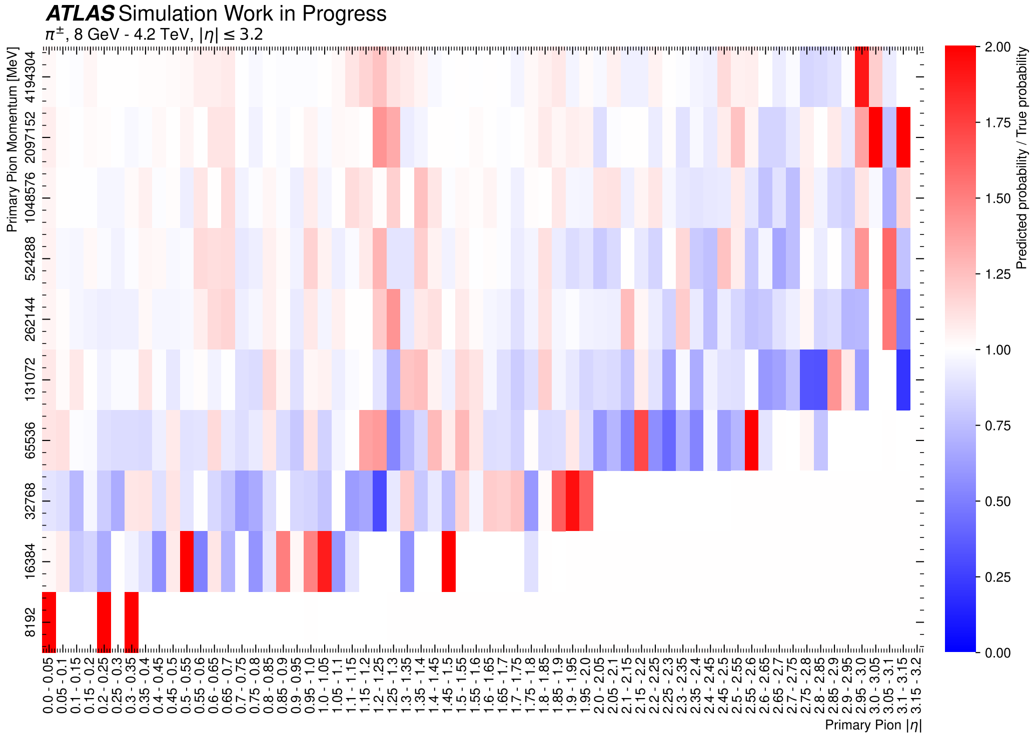


Figure 7.16 *Ratio of the neural network average predicted probability and true probability as a function of input pion momentum and η.*

7.3.2 Kinematics PCA

The correlations between the punch-through particle kinematics are captured using a Principal Component Analysis (PCA). The PCA rotates the kinematics into a new basis where any linear correlations are minimised.

The five kinematics are first transformed using a cumulative density function (CDF) to approximate a Gaussian. Figure 7.17 shows each of the transformed kinematics inclusively for all 9 punch-through particle types and their correlations.

A PCA rotation is then applied to the transformed kinematic distributions, where five principal components are used. Figure 7.18 shows the kinematics after applying the PCA rotation. The explained variances captured by each principal component are 43.28%, 21.22%, 18.47%, 10.43%, and 6.59% respectively. The PCA rotation shows consistent behavior for each punch-through particle type, input pion momentum, and input pion pseudorapidity.

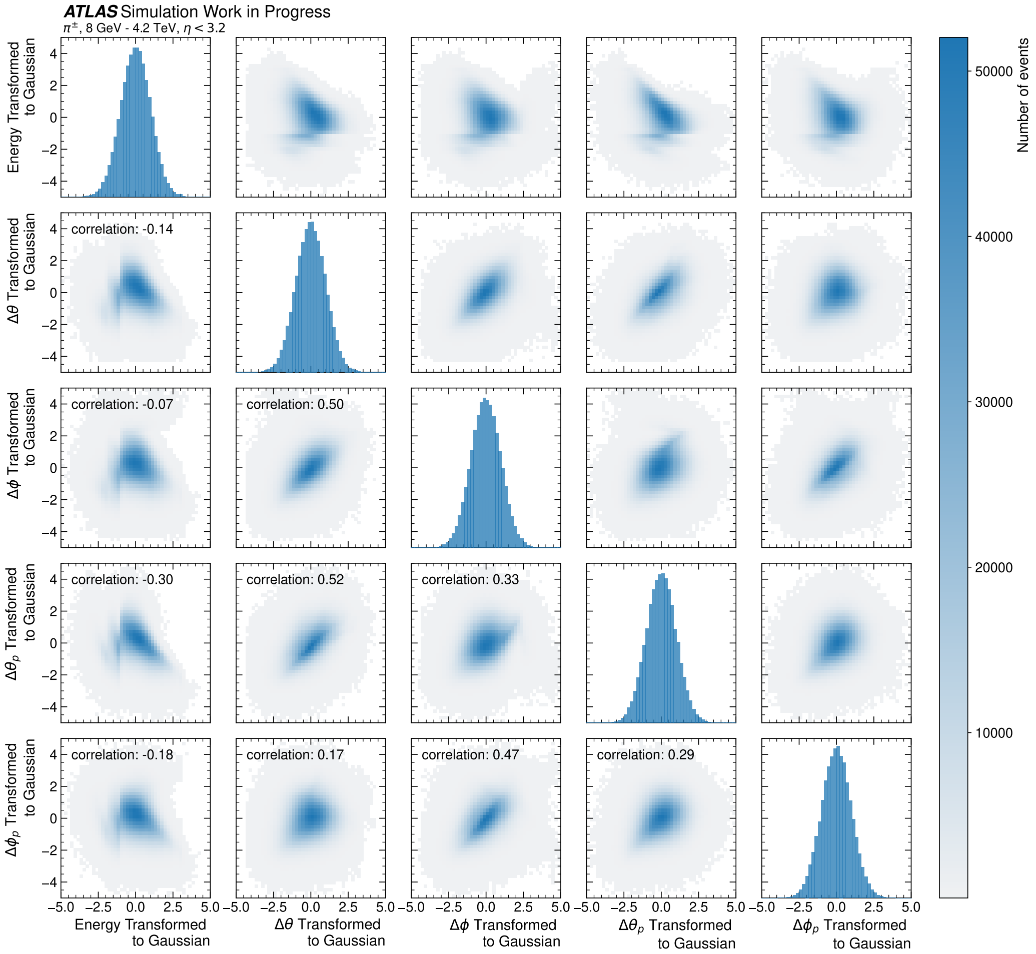


Figure 7.17 *The five punch-through kinematics transformed to Gaussian distributions. Plots are derived from all 9 punch-through particle types originating from single-pion events in the range 8 GeV - 4.2 TeV and* $|\eta| \leq 3.2$.

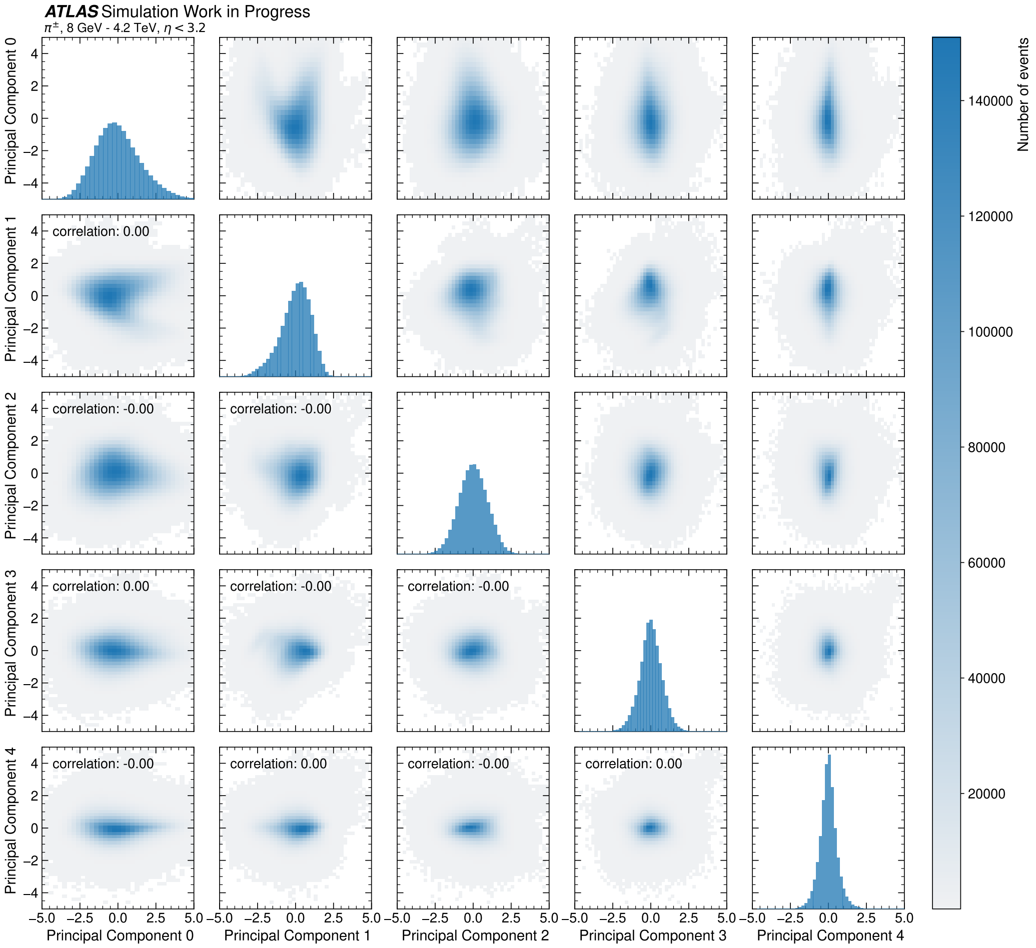


Figure 7.18 *The five principal components derived by a PCA rotation of the punch-through kinematics. Plots derived from all 9 punch-through particle types originating from single-pion events in the range 8 GeV - 4.2 TeV and* $|\eta| \leq 3.2$.

7.3.3 Parameterisation File

The parameterisation file contains a set of directories that segment the stored information and allow it to be used as a *lookup-table* (Figure 7.19). A directory is created for the frequency, and the 5 pricipal components of the each punch-through particle type. Within each of these directories a set of CDF distributions are stored, where each one models the indicated quantity for a specific momentum and psudorapidity range. For example, the directory `PCA0_PDG211`, contains the CDF distributions that model the first kinematic principal component of punch-through pions. Within this directory, there is the CDF distribution `E65536_etaMin000_etaMax040` which models momentum of 65536 MeV and the pseudorapidity range $0.00 \leq |\eta| \leq 0.40$.

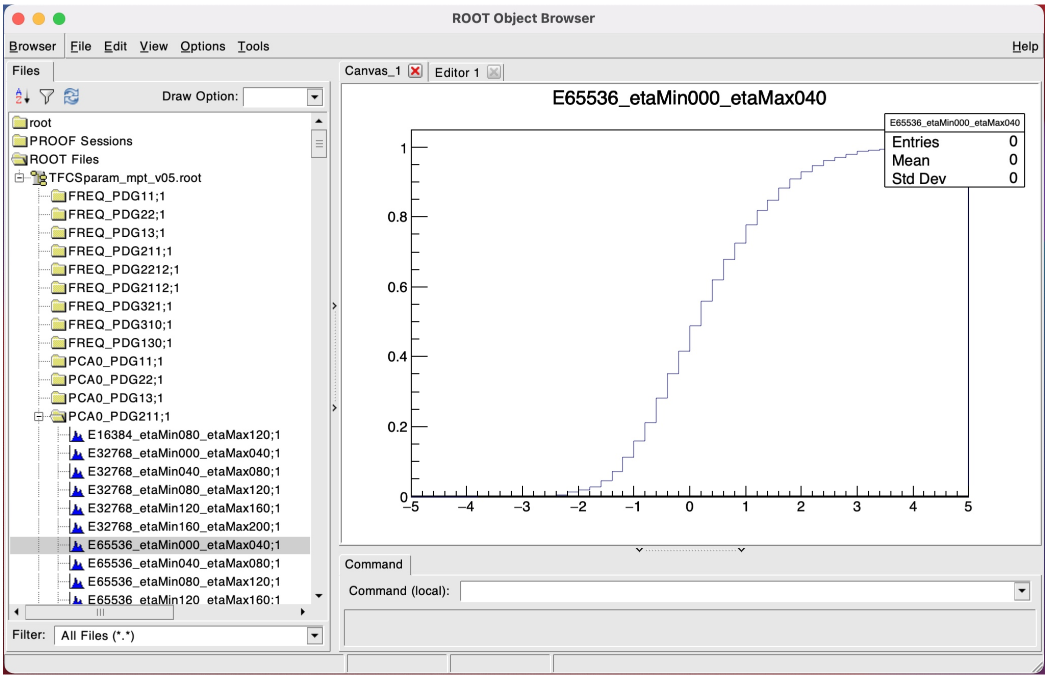


Figure 7.19 *The structure of the MPT parameterisation file that stores the frequency and kinematic principal components of different punch-through particles. CDFs in each directory correspond to different input momentum and pseudorapidity ranges.*

7.4 Simulation of the punch-through Effect

The punch-through simulation is fully integrated into the ATLAS simulation software framework as a standalone package and is used as default in AF3. This section describes how the parameterisations developed in Section 7.3 are used to simulate the punch-through effect.

7.4.1 Simulation Scheme

The punch-through simulation is initialised when an incoming particle enters the calorimeters. The punch-through simulation considers one incoming particle at a time. If the particle meets certain requirements, then it is considered further. Information about the particle and the calorimeter shower is then passed through the punch-through classifier neural network. If classified to create a punch-through, the simulation generates a set of punch-through particles using the parameterisation file. These punch-through particles are then positioned on the MuonEntryLayer and returned back to Geant4 for further simulation. The following points describe each of these steps in more detail.

Check Incoming Particle

A particle that enters the calorimeters is checked to see if it meets certain requirements. These requirements can be user-defined within the simulation code. However, the default requires the particle to be either a pion, a proton, a neutron, or a kaon with a momentum of ≥ 8 GeV and $|\eta| \leq 3.2$. The simulation also checks that the particle momentum direction is towards the MS. If the particle passes all of these checks it is considered a *punch-through initiator*.

Predict Using the Punch-Through Classifier

The kinematic properties of each *punch-through initiator*, and the calorimeter shower energy deposits are input into the punch-through classifier neural network model. The punch-through classifier returns a calibrated probability score. A random number is then drawn with a flat distribution between 0 and 1. If this random number is less than the punch-through probability score then punch-through particles are generated using the parameterisation file.

Match Energy and Pseudorapidity

Punch-through initiators are matched to the CDFs stored in the parameterisation file. First, the closest values are selected, then a linear interpolation is applied to the $|\eta|$ selection, and a linear interpolation in $\ln(momentum)$ is applied to the momentum selection. These interpolated values are used to select the CDFs for use in the following steps.

Number and Type of Punch-Through Particles

Random values are drawn from each selected frequency CDF to determine the frequency of each punch-through particle type to generate. If relevant, a random choice is made between creating the particle or antiparticle. The kinematics of each of these punch-through particles is determined in the following steps.

PCA Components

For each punch-through particle, a set of five numbers is drawn the CDFs corresponding to the principal components of the particle's kinematics. The inverse PCA transform is then applied to retrieve the energy and set of deflection angles for each particle.

Particle Position and Momentum

The position of the punch-through particle on the MuonEntryLayer surface is calculated using the deflection angles. The four-momentum of the particle is calculated using the momentum deflection angles and the energy. A random choice is made as to whether the deflection angles and momentum deflection angles are positive or negative in θ and ϕ.

Pass to Geant4

The punch-through particles that are created are passed over to Geant4 for further simulation in the MS.

7.5 Results

The MPT simulation has been extensively validated throughout the development process. In these validations, physics distributions are compared in events simulated with AF3 and compared to Geant4 as a reference. The following section displays some of these results.

7.5.1 Single-Pion Results

The MPT simulation is validated by simulating single-pion events. Single-pions are simulated using the same initial particle setup as the input samples (Section 7.2.1). Figure 7.20 shows the number of muon segments associated with reconstructed jets in these events. There is excellent agreement between AF3 with the MPT simulation and Geant4, it shows a vast improvement over AF3 without the MPT simulation. The configuration of AF3 without the MPT simulation can be considered as a proxy for AtlFastII (AF2) which does not include a MPT simulation component.

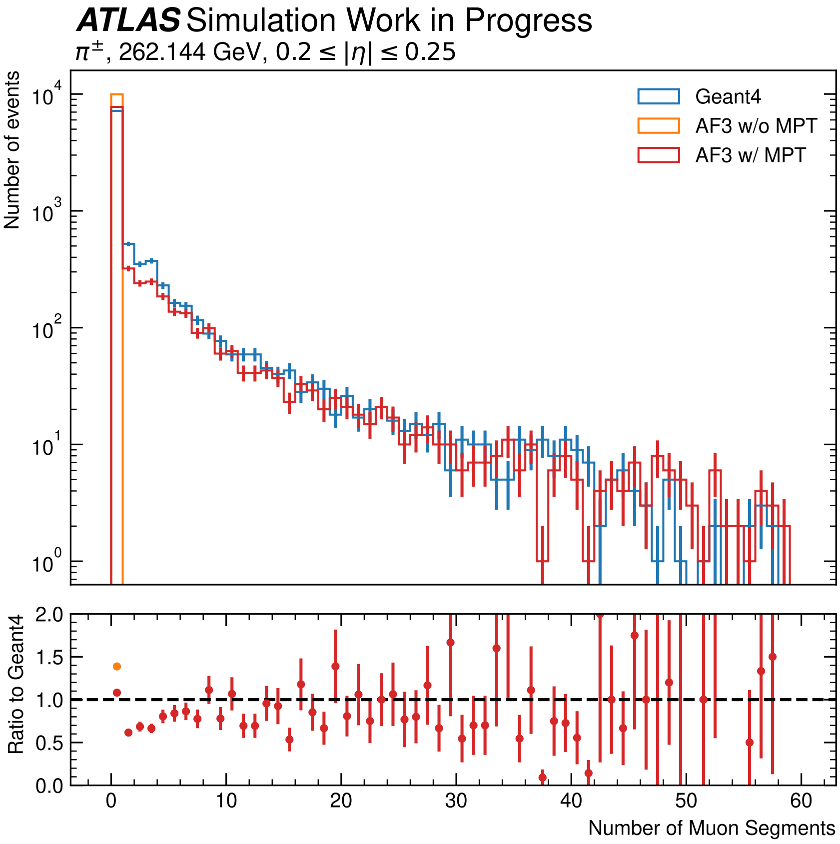


Figure 7.20 *The number of ghost muon segments associated with jets originating from single-pion events with an energy of 262.144 GeV and* $0.20 \leq |\eta| \leq 0.25$*. Geant4 is shown in blue, AF3 without the MPT simulation is shown in orange, and AF3 with the MPT simulation is shown in red.*

7.5.2 Z' Results

The MPT simulation is also validated by simulating Z' events. The Z' is a hypothetical particle similar to the Z boson but much heavier. It has similar decays to the Z boson, decaying to high energy hadrons $\sim 70\%$ of the time. This makes it a good process to test the punch-through simulation. Figure 7.21 shows the number of muon segments associated with reconstructed jets in these events. The agreement between AF3 with the MPT simulation and Geant4 is sufficiently good when considering how rare the puch-through effect is, and is vastly improved over AF3 without the MPT simulation. The configuration of AF3 without the MPT simulation can be considered as a proxy for AF2 which does not include a MPT simulation component.

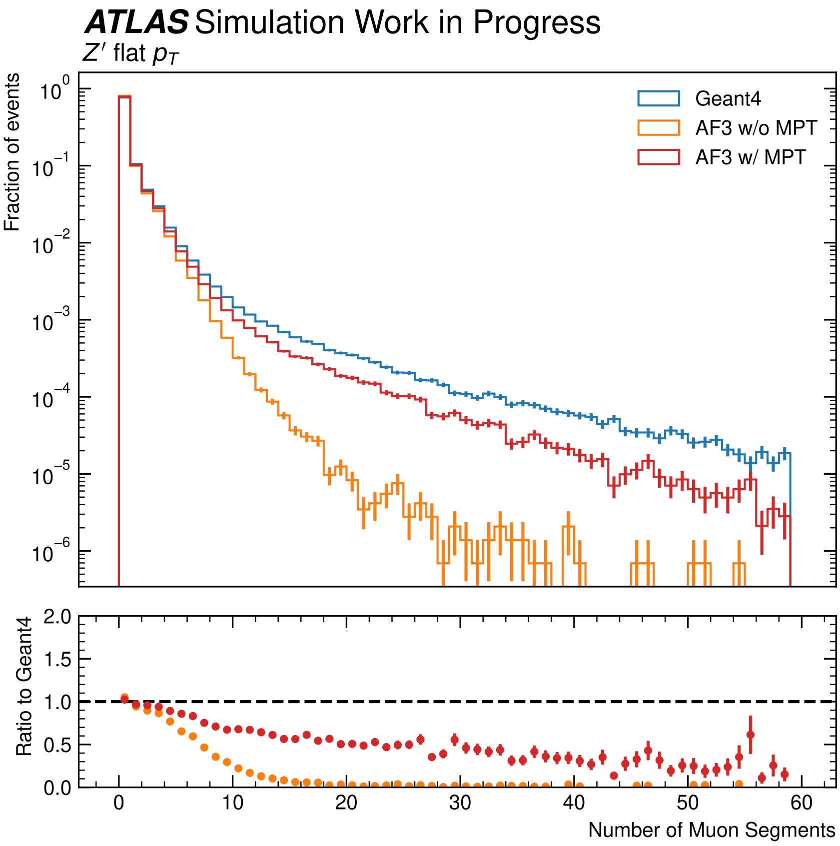


Figure 7.21 *The number of ghost muon segments associated with jets originating from Z' events. Geant4 is shown in blue, AF3 without the MPT simulation is shown in orange, and AF3 with the MPT simulation is shown in red.*

7.6 Summary

For this thesis, a fast simulation solution for the punch-through effect was developed for AF3. The solution is able to capture the intricacies of the punch-though effect, including: correlation to the underlying hadronic shower, correlations in the kinematics of punch-through particles, and the complex dependency of the punch-through effect on momentum and pseudorapidity. The solution is able to accurately model the number of muon segments associated with reconstructed jets in both single-pion events and Z' events. The muon-punch-through simulation will be used extensively in the future fast simulation program of the ATLAS experiment and will play a key role in meeting ATLAS computing requirements.

Part III

Constraining the Top Quark Yukawa Coupling in tH Production

Chapter 8

Measurement of the Higgs boson properties in the H $\rightarrow \gamma\gamma$ decay channel

This chapter presents an overview of an analysis of 139 fb^{-1} of proton-proton (p-p) collision data at $\sqrt{s} = 13$ TeV collected by the ATLAS experiment at the LHC, published in [1]. The analysis reports measurements of the cross-sections of Higgs boson production modes in the diphoton decay channel. The analysis is based on 101 signal regions, each targeting a different production mode and decay phase space. This includes the novel treatment of the $tHqb$ process, in which events are categorised to maximise the analysis sensitivity to the sign of the top quark Yukawa coupling as well as the Standard Model (SM) tH cross-section. The development of this novel categorisation is part of the work completed for this thesis, and is discussed in detail in Chapter 9. Using this approach the analysis is able to exclude the $\kappa_t = -1$ hypothesis at 2.2σ whilst simultaneously setting an upper limit on the tH cross-section at 10× the SM prediction.

The inclusive Higgs boson signal strength (defined as the ratio of the observed cross-section to the SM expected cross section) in the diphoton decay channel is measured to be $1.04^{+0.10}_{-0.09}$. Cross-sections for the gluon-gluon fusion, vector-boson fusion, associated production with a W and Z boson, and top associated production processes are reported (Section 8.6). The Higgs boson production cross-sections are also measured according to the Simplified Template Cross Sections (STXS) [51], comprising a total of 28 regions. Results are consistent with the SM predictions, with a p-value of 93%. The measurements are also used to set constraints on the Higgs boson coupling strengths in the kappa framework. The results presented in the kappa framework (Section 8.7) constitute part of the work completed for this thesis.

The purpose of this chapter is to provide an overview of the analysis strategy and results, giving enough information to understand the work in detailed chapter 9 and where it fits into the larger analysis.

8.0.1 Theoretical Frameworks

Simplified Template Cross-Sections

The STXS framework [51] is designed to be common across analyses and experiments, with the aim of reducing theoretical uncertainties in measurements, and ensuring results can be easily combined. The analysis targets the Stage 1.2 STXS binning scheme. Bins are defined at the particle level, in the fiducial Higgs boson rapidity region $|y_H| < 2.5$. Bins are mutually exclusive for each of the Higgs boson production processes:

- the $gg \to H$ process, which includes ggF and $gg \to ZH$ followed by a hadronic decay of the Z boson.
- the $qq' \to Hqq'$ process, which includes both VBF production and $q\bar{q}' \to VH$ where the vector boson decays hadronically.
- the $V(\to \ell\ell)H$ process, corresponding to the $pp \to VH$ process, where the vector boson decays leptonically (in the case of ZH, both decays of the Z boson to charged leptons and neutrinos are included)
- the top-associated production modes $t\bar{t}H$ and tH

Each production mode is split further into different regions based on the kinematics of the Higgs boson, the associated jets, and the W and Z bosons. Splits are also made on the number of jets, the number of leptons and the number of top quarks in the event. The STXS 1.2 Scheme is summarised in Figure 8.1.

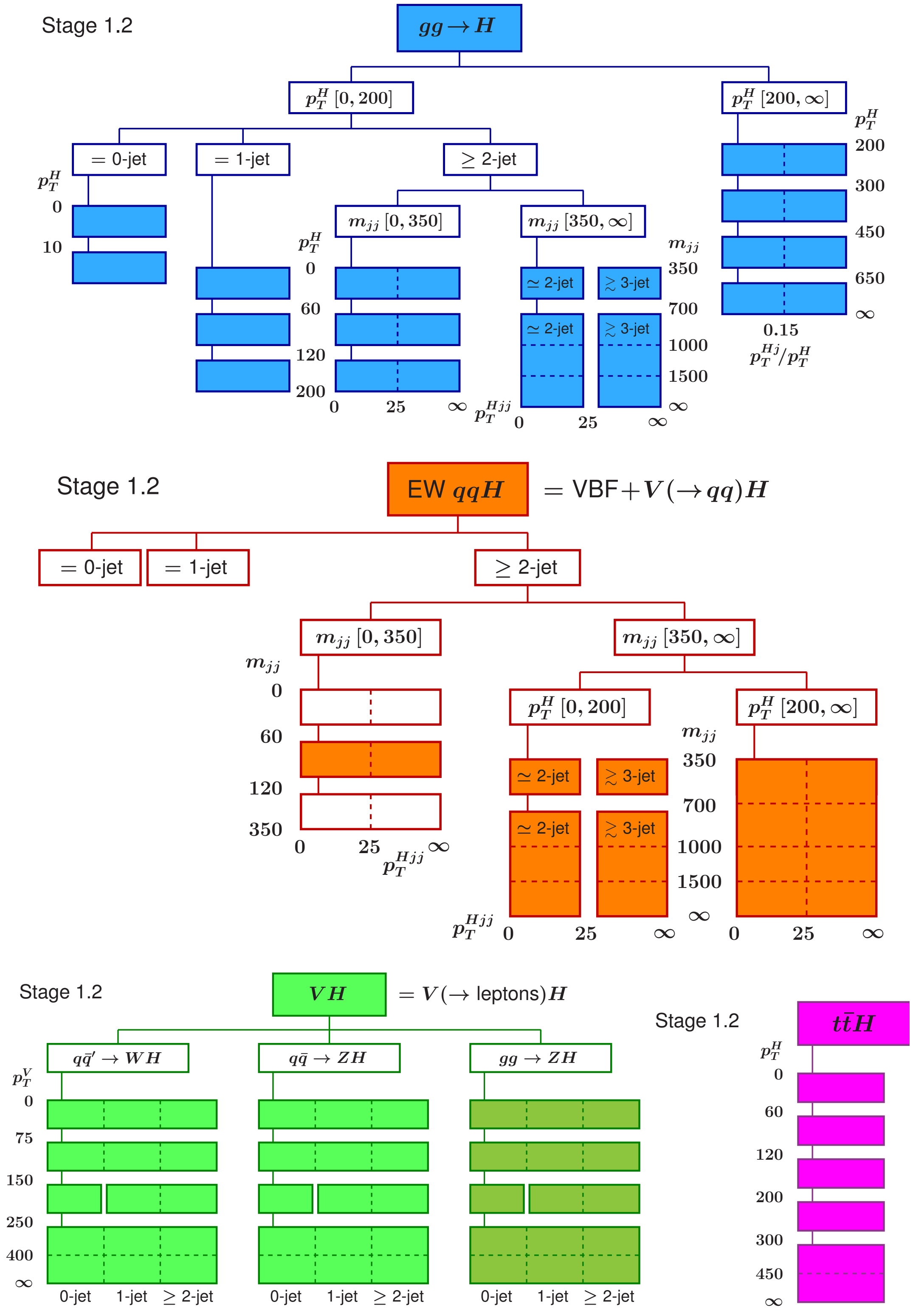


Figure 8.1 *Schematic showing the STXS 1.2 Scheme for the main Higgs boson production modes. The analysis measurement also targets the* tH *process without further splitting, and doesn't split the* $b\bar{b}H$ *process. Taken from [1].*

Kappa Framework

The couplings of the Higgs boson cannot be directly measured, however they can be extracted from the measured cross-sections of Higgs processes. This is performed by parameterising each of the SM Higgs production cross-sections and partial decay widths by multiplicative factors called kappas. In the parameterisation, it is assumed that there is only one underlying Higgs state that is SM like [52]. The zero width approximation is used for the Higgs state, and hence the signal cross-section of a given Higgs process is:

$$(\sigma \cdot BR)(ii \to H \to \gamma\gamma) = \frac{\sigma_{ii}(\kappa) \cdot \Gamma_{\gamma\gamma}(\kappa)}{\Gamma_H(\kappa)} \tag{8.1}$$

where $\sigma_{ii}(\kappa)$ is the production cross-section through the initial state ii, $\Gamma_{\gamma\gamma}(\kappa)$ is the partial decay width of the Higgs boson decaying to two photons, and $\Gamma_H(\kappa)$ is the total width of the Higgs boson as a function of the kappa coupling modifiers.

Modifiers for the Higgs boson couplings to the W and Z bozon are denoted κ_W and κ_Z, The couplings to the charm, bottom, and top quarks, and muon, and τ leptons are denoted κ_c, κ_b, κ_t, κ_μ, κ_τ respectively. The couplings of the Higgs to the other SM particles are assumed equal to their SM predictions.

The parameterisations of $\sigma_{ii}(\kappa)$, $\Gamma_{\gamma\gamma}(\kappa)$ and $\Gamma_H(\kappa)$ in terms of the individual kappa modifiers is shown in Table 8.1. The ggF and $H \to \gamma\gamma$ processes can be parameterised in terms of the resolved tree level Higgs boson couplings or by the effective κ_g and κ_γ coupling modifiers. Both cases are considered in the interpretation of results in this framework. The parameterisations of the partial decay widths Γ_{gg} and $\Gamma_{Z\gamma}$ are also considered as these have a significant impact on the total Higgs width Γ_H. The total Higgs width is expressed as a function of the κ modifiers, assuming no contributions from Higgs boson decays other than those present in the SM.

The $tHqb$ and tHW production cross-sections are parameterised independently in each analysis category by the expected signal yield. The parameterisations take into account event shape differences that result in a category efficiency dependency. The parameterisations constitute part of the work completed for this thesis, and are discussed in more detail in Section 9.8.

Table 8.1 *The parameterization of the Higgs boson production cross-sections σ_{ii}, the partial decay widths $\Gamma^{\gamma\gamma}$, Γ^{gg} and $\Gamma^{Z\gamma}$, and the total width Γ_H, normalized to their SM values, as functions of the coupling-strength modifiers κ. The parameterisation of the $\sigma(tHW)$ and $\sigma(tHqb)$ cross-sections are derived independently for each analysis category taking into account acceptance effects. Other coefficients are obtained from Refs. [10, 53]. Taken from [1].*

Production cross-section	Main interference	Effective modifier	Resolved modifier
$\sigma(\text{ggF})$	t–b	κ_g^2	$1.040\,\kappa_t^2 + 0.002\kappa_b^2 - 0.038\,\kappa_t\kappa_b - 0.005\,\kappa_t\kappa_c$
$\sigma(\text{VBF})$	-	-	$0.733\,\kappa_W^2 + 0.267\,\kappa_Z^2$
$\sigma(q\bar{q} \to ZH)$	-	-	κ_Z^2
$\sigma(gg \to ZH)$	t–Z	-	$2.456\,\kappa_Z^2 + 0.456\,\kappa_t^2 - 1.903\,\kappa_Z\kappa_t$ $- 0.011\,\kappa_Z\kappa_b + 0.003\,\kappa_t\kappa_b$
$\sigma(WH)$	-	-	κ_W^2
$\sigma(t\bar{t}H)$	-	-	κ_t^2
$\sigma(tHW)$	t–W	-	$A\,\kappa_t^2 + B\,\kappa_W^2 + C\,\kappa_t\kappa_W$, category-dependent
$\sigma(tHqb)$	t–W	-	$A\,\kappa_t^2 + B\,\kappa_W^2 + C\,\kappa_t\kappa_W$, category-dependent
$\sigma(b\bar{b}H)$	-	-	κ_b^2
Partial and total decay widths			
$\Gamma^{\gamma\gamma}$	t–W	κ_γ^2	$1.589\,\kappa_W^2 + 0.072\,\kappa_t^2 - 0.674\,\kappa_W\kappa_t + 0.009\,\kappa_W\kappa_\tau$ $+ 0.008\,\kappa_W\kappa_b - 0.002\,\kappa_t\kappa_b - 0.002\,\kappa_t\kappa_\tau$
Γ^{gg}	t–b	κ_g^2	$1.111\,\kappa_t^2 + 0.012\,\kappa_b^2 - 0.123\,\kappa_t\kappa_b$
$\Gamma^{Z\gamma}$	t–W		$1.118\,\kappa_W^2 + 0.004\,\kappa_t^2 - 0.125\,\kappa_W\kappa_t + 0.003\,\kappa_W\kappa_b$
Γ_H	-	κ_H^2	$0.581\,\kappa_b^2 + 0.215\,\kappa_W^2 + 0.063\,\kappa_\tau^2$ $+ 0.026\,\kappa_Z^2 + 0.029\,\kappa_c^2 + 0.0023\,\kappa_\gamma^2$ $+ 0.0004\,\kappa_s^2 + 0.00022\,\kappa_\mu^2$ $+ 0.082\,(\Gamma^{gg}/\Gamma^{gg}_{\text{SM}})$ $+ 0.0015\,(\Gamma^{Z\gamma}/\Gamma^{Z\gamma}_{\text{SM}})$

8.1 Datasets

8.1.1 Experimental Data

The analysis is performed on the full Run 2 set of p-p collision data collected by the ATLAS experiment at the LHC. The protons are collided at a centre of mass energy of $\sqrt{s} = 13$ TeV over the Run 2 period from 2015 to 2018, amounting to 139 ± 2.4 fb^{-1}of integrated luminosity.

The events are required to pass a diphoton trigger, which requires transverse energy above 35 GeV and 25 GeV for the two highest-p_{T} photon candidates in the event. The trigger uses photon identification selections based on calorimeter shower shape variables: in 2015–2016, a loose photon identification requirement was used, and in 2017–2018, a tighter requirement was used to cope with higher instantaneous luminosity.

8.1.2 Simulated Data

Simulated data is used in the analysis design and results, including: signal and background modelling, event selection, event categorisation, and the Asimov dataset [54]. The *nominal* datasets for each process are simulated with different Monte-Carlo (MC) generators that are chosen based on their modelling acccuracy. *Alternative* signal samples are also simulated with different generators and are used to estimate systematic uncertainties. All of the simulated signal samples and most of the background samples are simulated using the full Geant4 simulation of the ATLAS detector. The QCD diphoton background samples are simulated using the AFII (AtlFast2) fast simulation due to the high amount of statistics required.

The samples for the ggF, VBF, VH, and $t\bar{t}H$ proceses are generated using Powheg Box v2 [55–58].

The $tHqb$ and tHW samples are produced with MadGraph5_aMC@NLO [59]. The $tHqb$ (tHW) samples are produced in the four-flavour (five-flavour) scheme with the NNPDF3.0nnlo PDF. The same flavour scheme was used in the matrix element calculation and the PDF. The top quark and W boson decays are handled by MadSpin [60] to account for spin correlations in the decay products. The overlap between the tHW process and the $t\bar{t}H$ process at NLO was handled using the diagram removal technique detailed in [61, 62]. The $pp \to tHb$ process was not considered in the modelling of tH production due to its small cross-section.

The generated events are interfaced to Pythia [63] to model the decay of the Higgs boson, parton showering, and hadronization.

In addition, alternative samples generated with MadGraph5_aMC@NLO or showered with Herwig [64] are produced to evaluate the signal generator modelling uncertainties.

The decay of the Higgs boson into two photons is simulated by either Pythia [63] or Herwig [64]. The mass of the Higgs boson is set to $m_H = 125$ GeV, with a width of $\Gamma_H = 4.07$ MeV.

Each sample is normalised to the integrated luminosity using the latest theoretical calculations for each Higgs production mode cross-section [53]. The branching ratio of the Higgs decay to two photons of 2.270×10^{-3} is also used to normalise the samples. Table 8.2 shows the cross-section for each nominal sample.

Prod. Mode	σ [pb]
ggF	48.5
VBF	3.78
WH	1.37
$qq \to ZH$	0.76
$gg \to ZH$	0.12
$t\bar{t}H$	0.51
$tHqb$	0.074
tHW	0.015

Table 8.2 *Cross-sections of each production mode. Taken from [1].*

Three irreducible backgrounds are simulated: QCD $\gamma\gamma$ production, $V\gamma\gamma$ and $t\bar{t}\gamma\gamma$. The production of $t\bar{t}\gamma\gamma$ events is modelled using the MadGraph5_aMC@NLO 2.3.3 generator at Leading Order (LO) with the NNPDF2.3lo [7] PDF. The parton-showering and underlying-event simulation were performed using Pythia 8.2. The background samples are used to model the continuous background shape in $m_{\gamma\gamma}$ in each of the analysis categories (Section 8.4.2).

8.2 Object Reconstruction and Event Selection

Firstly, reconstructed photon candidates are required to pass a set of *preselection*-criteria. The two highest p_{T} photon objects are then used to define the diphoton system in the event. A neural network algorithm reconstructs the primary vertex from the diphoton system [65]. Finally, the photons are required to satisfy selection and additional identification criteria detailed in 8.2.2.

8.2.1 Photon Reconstruction & Identification

Photons are reconstructed from topo-clusters as detailed in Section 3.10.4. Reconstructed photons must fall inside the region of the electromagnetic (EM) calorimeter ($|\eta| < 2.37$), and outside of the range $1.37 < |\eta| < 1.52$ corresponding to the calorimeter transition region. Photon candidates are distinguished from jet backgrounds using shower shape variable information, using loose identification criteria in event preselection and tight identification criteria for the final event selection. The efficiency of the tight identification for unconverted (converted) photons ranges from 84% (85%) at $p_{\mathrm{T}} = 25$ GeV to 94% (98%) for $p_{\mathrm{T}} > 100$ GeV. In addition, isolation requirements on the surrounding activity within $\Delta R = 0.2$ of the photon are made. The total energy of calorimeter clusters and the sum of track transverse momenta are requested to be below 6.5% and 5% of the photon's transverse energy respectively.

8.2.2 Event Selection

Following the pre-selection requirements, the selected photon candidates are required to satisfy the *tight* identification and isolation criteria described in Section 8.2.1. Finally, the primary and secondary photon candidates are required to have a $p_T/m_{\gamma\gamma} > 0.35$ and > 0.25 respectively. Events that pass these selections and requirements are referred to as Tight-Isolated (TI) data. Events that fail the tight identification or isolation requirements are used throughout the analysis as a data-driven background estimate, referred to as Non-Tight-Isolated (NTI) data.

The result of these selections is a dataset of about 1.2 million events in the diphoton invariant mass in the range 105 to 160 GeV. The events in this mass range are used for further analysis of the properties of the Higgs boson. The predicted efficiency for the selection for a SM Higgs boson with $|y_H| < 2.5$ is 39%.

8.2.3 Jet, b-jet, Lepton, Top Quark and $E_{\mathrm{T}}^{\mathrm{miss}}$ Reconstruction and Selection

Jets (including b-jets), muons, electrons, and $E_{\mathrm{T}}^{\mathrm{miss}}$ objects are then reconstructed and used in the analysis to categorise events and measure the Higgs boson properties.

Jets are reconstructed using the Particle Flow algorithm [66] from energy deposits in the calorimeters, using the topological clustering of the anti-k_t algorithm [28] as detailed in Section 3.10. Jets are required to have $p_T > 25$ GeV and $|\eta| < 4.4$. The jet-vertex-tagger (JVT) [30] is used to suppress the jet background from pileup when reconstructing jets with $p_T < 60$ GeV and $|\eta| < 2.4$. For forward jets ($|\eta| > 2.5$) the forward-JVT based on the PFlow algorithm is used [67]. Jets containing b-hadrons are identified using the DL1r b-tagging algorithm, where a pseudo-continuous b-tagging score is used based on selection efficiency working points [68].

Electrons are reconstructed using topo-clusters as detailed in Section 3.10.4. Electron candidates are required to have $p_T > 10$ GeV and $|\eta| < 2.47$. They are also required to be outside of the EM calorimeter transition region. Electron candidates must satisfy the medium selection criteria laid out in [27], and isolation criteria are applied based on track and calorimeter information. The track associated with the electron candidate must be consistent with the selected diphoton vertex.

Muons are reconstructed using tracks in the inner detector (ID) and Muon Spectrometer (MS) as detailed in Section 3.10.6. Muons are selected using the *medium* selection criteria and are also required to meet calorimeter and track-based isolation criteria [33]. Muons are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.7$.

Overlap Removal is performed on the reconstructed objects. The procedure prioritises photons, and then removes any overlapping objects in ΔR.

Missing Transverse Energy or $E_{\mathrm{T}}^{\mathrm{miss}}$ is the final object to be reconstructed as the negative vector sum of the four momenta of the other reconstructed objects, as detailed in 3.10.

Top Quarks are reconstructed using a BDT algorithm trained with the $t\bar{t}H$ sample using the XGBoost package [69]. The BDT extracts the three-jet triplet that best matches the hadronic products of a top quark. Input variables to the BDT consist of: the p_{T} η, ϕ, and energy of the W boson and b-jet in consideration (where the W boson is formed from a pair of jets). Angular variables between the constituents are also considered, as well as b-tagging information and the invariant mass of the triplet. The BDT score is used to reconstruct top quark candidates in hadronic and leptonic events in the following way.

In hadronic events, where no leptons are reconstructed, the triplet with the highest BDT score is selected as the primary reconstructed top quark (t_1). In leptonic events, where exactly one lepton is reconstructed, the W boson candidate is first reconstructed from the lepton and the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$. The t_1 is then reconstructed from the W boson candidate and the jet that gives the higest BDT score. No top quark candidate is reconstructed in events with more than one reconstructed lepton. After t_1 is selected, if there are at least three additional jets remaining, a second top quark candidate (t_2) is reconstructed from the triplet with the highest BDT score. If there are two jets remaining after the selection of t_1, then a t_2 is reconstructed as the sum of these two jets, otherwise no t_2 is reconstructed [70].

The reconstructed top quark candidates t_1 and t_2 are also referred to as “recotop1” and “hybridtop2” respectively.

8.3 Event Categorisation

8.3.1 Overview

The events that pass the selection described in Section 8.2.2 are classified into categories each targeting an STXS region. The procedure uses a novel strategy to maximise the sensitivity of the measurement. The steps are as follows:

- First, a multiclass BDT is trained to separate signal events into categories that are sensitive to each targeted STXS region. Simulated signal samples are used to model each of the Higgs boson production modes. Backgrounds are not considered in this step.
- Next, a novel procedure titled "D-optimality" optimises the categorisation of events to maximise sensitivity across the whole STXS scheme.
- For category defined in the previous step a binary BDT is trained to separate the signal class from the background. This includes the non-resonant background and the other signal classes.
- A significance scan is then conducted across each of the binary BDT discriminants to define up to three analysis categories sensitive to the targeted STXS region.

The categorisation strategy is summarised in Figure 8.2. A more detailed description of each step and any additional or alternative treatment is discussed in the following sections. The variables used to train the models in each step are summarised in Table 8.3.

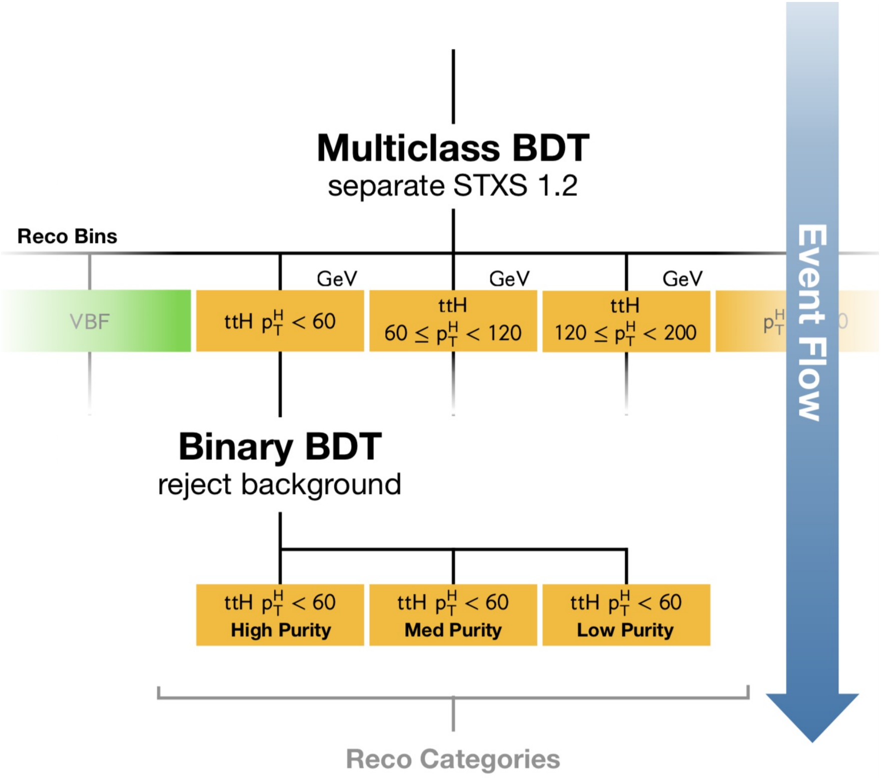


Figure 8.2 *Schematic illustrating the analysis categorisation procedure. "Reco Bins" represents the STXS regions that are targeted by the categorisation. "Reco Categories" represents the analysis categories defined using the BDT algorithms that are sensitive to the targeted STXS regions. The multiclass BDT considers each STXS signal class simultaneously, defining a category for each. A binary BDT is trained in each multiclass BDT category to reject the non-resonant background. Up to three categories are defined using each binary BDT with varying signal purity.*

Table 8.3 *The training variables used in the multiclass BDT and the set of binary classifiers. The first column lists the variables used in the multiclass BDT. The second column lists the variables used by the different binary classifier models, where different input variables are used for models targeting different STXS classes. The dagger symbol † denotes variables that have two versions with different jet p_{T} requirements. One version of such a variable is defined using jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, and the other version is defined using jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$. Both versions are used in the training of the multiclass BDT. The asterisk symbol * denotes $tHqb$ training variables that are only used for the classifiers suppressing the continuum background. Other $tHqb$ training variables are used in all three $tHqb$ classifiers. Taken from [1].*

Multiclass BDT	Binary classifier	
	STXS classes	Variables
η of the 2 highest-p_{T} photons transverse momentum ($p_{\mathrm{T}}^{\gamma\gamma}$) and rapidity ($y_{\gamma\gamma}$) of the diphoton system, di-jet p_{T}†, invariant mass, Δy, $\Delta\phi$, $\Delta\eta$ between the 2 jets, p_{T}, invariant mass of the $\gamma\gamma + j$ system, where j denotes the highest-p_{T} jet, p_{T}†, invariant mass of $\gamma\gamma + jj$ system, where jj denotes the two highest-p_{T} jets, Δy, $\Delta\phi$ between $\gamma\gamma$ and jj, minimum ΔR between jets and photons, invariant mass of the system comprising all jets in the event, di-lepton p_{T}, di-e or di-μ invariant mass (leptons are required to be oppositely charged), $E_{\mathrm{T}}^{\mathrm{miss}}$, p_{T} and transverse mass of the lepton + $E_{\mathrm{T}}^{\mathrm{miss}}$ system, p_{T}, η, ϕ of top candidates, invariant mass of the two highest-p_{T} top candidates; Number of jets †, of central jets ($\lvert\eta\rvert < 2.5$)†, of b-jets† and of leptons, p_{T} of the highest-p_{T} jet, scalar sum of the p_{T} of all jets, scalar sum of the transverse energies of all particles ($\sum E^{\mathrm{T}}$), $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, $\left\lvert E_{\mathrm{T}}^{\mathrm{miss}} - E_{\mathrm{T}}^{\mathrm{miss}}(\text{primary vertex with the highest } \sum p_{\mathrm{T,track}}^{2})\right\rvert > 30\,\mathrm{GeV}$ Score of the top reconstruction BDT, ΔR between the W and b candidates within the second top candidate ($\Delta R(W, b)$), η of the most forward jet, invariant mass of the diphoton and most forward jet system, Average number of interactions per bunch crossing.	individual STXS classes from $gg \to H$ $qq' \to Hqq'$ $qq \to H\ell\nu$ $pp \to H\ell\ell$ $pp \to H\nu\bar{\nu}$	All multiclass BDT variables, $\Delta\eta$ between the 2 photons ($\Delta\eta_{\gamma\gamma}$), Number of electrons and muons, $\gamma\gamma$ $\boldsymbol{p}_{\mathrm{T}}$ projected to its thrust axis ($p_{\mathrm{Tt}}^{\gamma\gamma}$), $\eta^{Zepp} = \frac{\eta_{\gamma\gamma}-\eta_{jj}}{2}$, where $\eta_{\gamma\gamma}$ and η_{jj} are the η of respectively the diphoton and jj systems, $\phi^{*}_{\gamma\gamma} = \tan(\frac{\pi-\lvert\Delta\phi_{\gamma\gamma}\rvert}{2})\sqrt{1-\tanh^{2}(\frac{\Delta\eta_{\gamma\gamma}}{2})}$, $\cos\theta^{*}_{\gamma\gamma} = \left\lvert\frac{(E^{\gamma_1}+p_z^{\gamma_1})\cdot(E^{\gamma_2}-p_z^{\gamma_2})-(E^{\gamma_1}-p_z^{\gamma_1})\cdot(E^{\gamma_2}+p_z^{\gamma_2})}{m_{\gamma\gamma}+\sqrt{(m_{\gamma\gamma}^{2}+(p_{\mathrm{T}}^{\gamma\gamma})^{2})}}\right\rvert$ where E^{γ_1} and $p_z^{\gamma_1}$ (E^{γ_2} and $p_z^{\gamma_2}$) are respectively the energy and longitudinal momentum of the highest-p_{T} (second-highest-p_{T}) photon and $\Delta\phi_{\gamma\gamma}$ is the ϕ difference between the photons.
	all $t\bar{t}H$ and tHW STXS classes combined	p_{T}, η, ϕ of 2 leading photons, p_{T}, η, ϕ and b-tagging scores of 6 leading jets, $E_{\mathrm{T}}^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, $E_{\mathrm{T}}^{\mathrm{miss}}$ azimuthal angle, Score of the top reconstruction BDT, p_{T}, η, ϕ of 2 leading leptons.
	$tHqb$	$p_{\mathrm{T}}^{\gamma\gamma}/m_{\gamma\gamma}$, $\eta_{\gamma\gamma}$ p_{T}, invariant mass, BDT score and $\Delta R(W, b)$ of the leading top candidate (t_1); p_{T}, η of the sub-leading top candidate (t_2); $p_{T}^{j_F}$, η_{j_F}, the transverse momentum and pseudo-rapidity of the most forward jet (j_F) Angular variables: $\Delta\eta_{\gamma\gamma,t_1}$, $\Delta\theta_{\gamma\gamma,t_2}$, $\Delta\theta_{j_F,t_1}$, $\Delta\theta_{j_F,t_2}$, $\Delta\theta_{j_F,\gamma\gamma}$ Invariant mass variables: $m_{\gamma\gamma,j_F}$, m_{t_1,j_F}, m_{t_2,j_F}, $m_{\gamma\gamma,t_1}$ Number of jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, Number of b-jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}^{*}$; Number of leptons*, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance*

8.3.2 Multiclass BDT

The multiclass BDT model is trained using LightGBM [71]. The model is trained on the simulated signal samples to separate each of the STXS truth bins targeted by the analysis. The targeted classes differ slightly from the STXS 1.2 scheme by the following points, either to improve classification performance, or because of limited sensitivity. This includes:

- The $b\bar{b}H$ and $gg \to H$ processes are hard to separate, and have a very similar analysis acceptance and selection efficiency. They are therefore measured as a single process in the analysis.
- The tH process is split into regions corresponding to the $pp \to tHqb$ and $pp \to tHW$ processes because of their differing kinematics. The $pp \to tHb$ process is ignored due to its very small cross-section.

The targeted classes are summarised in Figure 8.3. Each target class is weighted in the training to normalise their importance in the model loss function. The training is performed by minimising the categorical cross-entropy loss across each class.

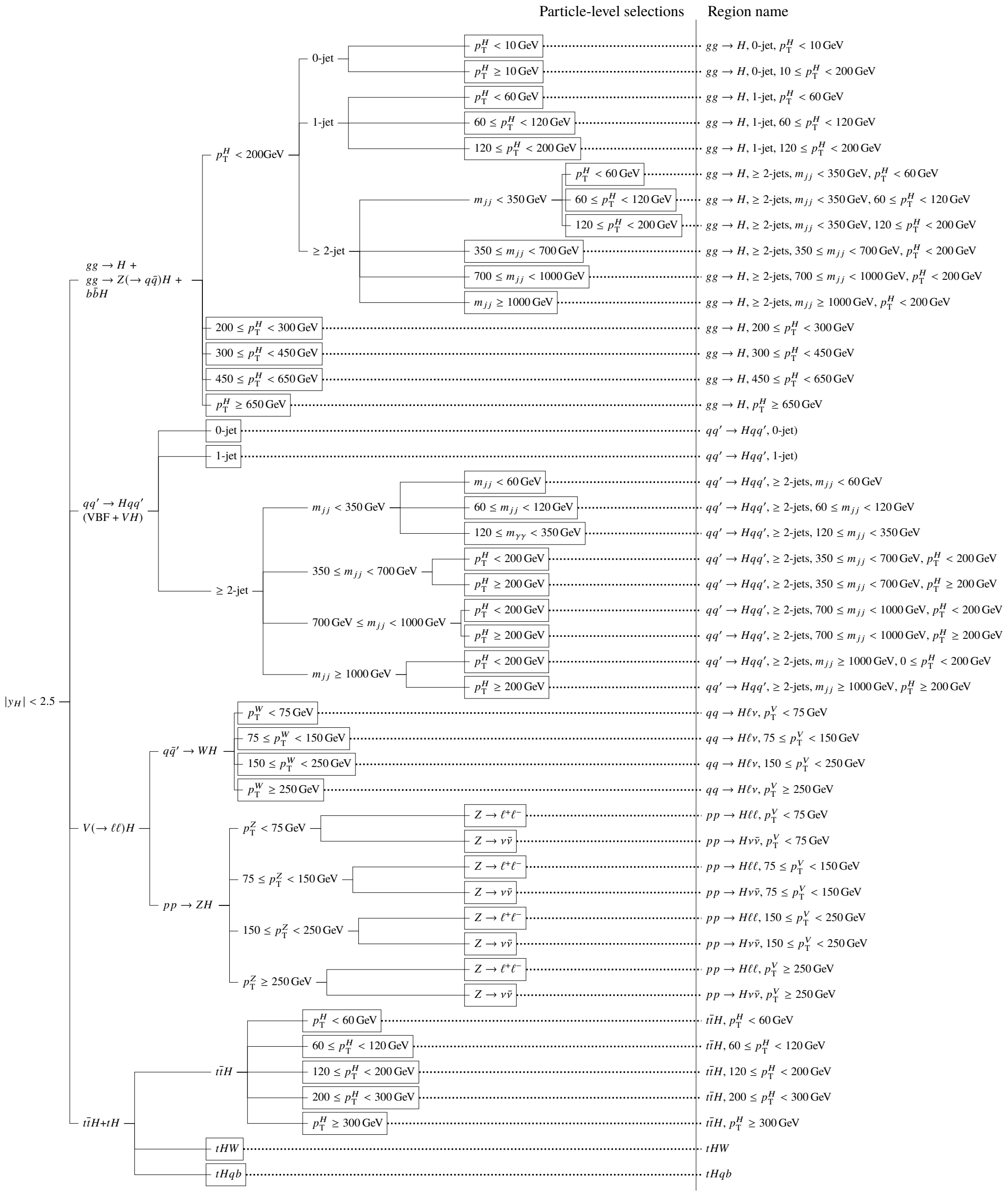
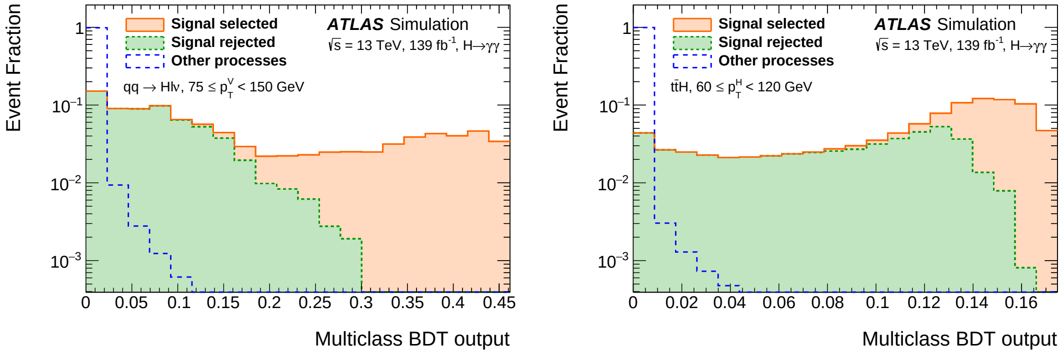


Figure 8.3 *The targeted STXS regions in the analysis. The tree indicates particle level selections that are made to define each region. Final regions are indicated by a box. The name of each region used in the analysis is shown on the right of the tree. Taken from [1].*

8.3.3 D-Optimality

In multi-classification, the class with the highest probability is selected as the model prediction. In the D-optimality procedure, weights are applied to the model prediction for each class to adjust how events are classified. The weights are adjusted to maximise the inverse determinant of the covariance matrix of the expected measurement. This results in a categorisation that achieves the maximum possible sensitivity across all targeted STXS regions. Figure 8.4 shows the weighted output score for two categories targeting different STXS regions.

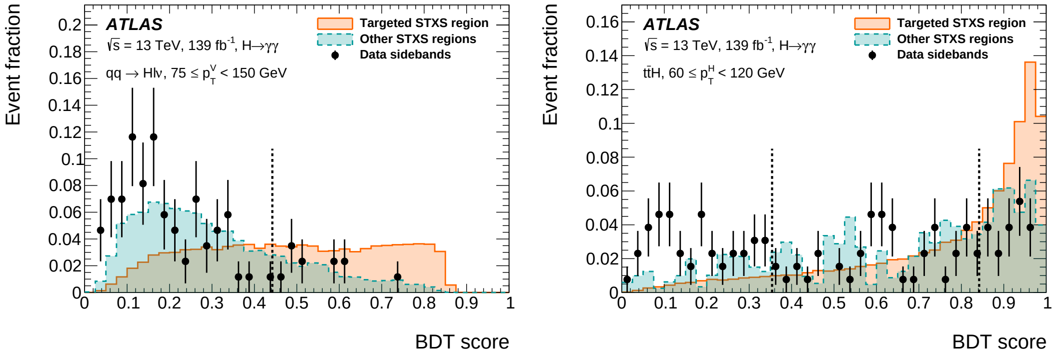


(a) $qq \to H\ell\nu$, $75 \leq p_{\mathrm{T}}^{V} < 150\,\mathrm{GeV}$ **(b)** $t\bar{t}H$, $60 \leq p_{\mathrm{T}}^{H} < 120\,\mathrm{GeV}$

Figure 8.4 *The weighted output distributions of the multiclass BDT for two STXS categories. In each plot, events corresponding to the targeted process are shown in solid colour, whereas the signal of other STXS classes is shown in long-dashed with no fill. The signal is further broken down into those events that are selected in the class (solid-orange), and those events that are classified into another category (dashed-green). The solid-orange component is stacked on top of the dashed-green component. Events are assigned to the class in which they have the largest weighted score. Taken from [1].*

8.3.4 Global Binary BDTs

Binary background rejection models are trained in each region. The models are used to reject the non-resonant background and increase sensitivity to the targeted STXS class. For the $gg \to H$, $qq' \to Hqq'$ and $V(\to \ell\ell)H$ processes, a binary BDT is trained to separate the targeted signal class from the continuum background and the other signal classes. Figure 8.5 shows the BDT score of two representative binary BDT models.



(a) $qq \to H\ell\nu$, $75 \le p_{\mathrm{T}}^{V} < 150\,\mathrm{GeV}$ **(b)** $t\bar{t}H$, $60 \le p_{\mathrm{T}}^{H} < 120\,\mathrm{GeV}$

Figure 8.5 *The binary BDT discriminant in two STXS categories. Events corresponding to the targeted STXS region are shown in solid-orange, and events from other STXS regions are shown in dashed-blue. Events in the data sidebands are shown as black dots. The vertical dashed lines represent cuts applied on the BDT score to define analysis categories. Taken from [1].*

8.3.5 $t\bar{t}H$ and tHW Binary BDTs

For the $t\bar{t}H$ and tHW processes, a different approach is taken. Here a binary BDT is trained inclusively on all events that enter any of the multiclass *top-like* categories. The BDT is trained to separate the $t\bar{t}H$ process from the continuum background, and other top processes ($tHqb$ and tHW). A similar procedure is used for the tHW process.

8.3.6 $tHqb$ Categorisation Treatment

A novel approach is used to categorise the $tHqb$ process, with the objective of maximising sensitivity to the sign of the top Yukawa coupling strength κ_t. The approach takes advantage of the kinematic differences between $tHqb$ events with $\kappa_t = +1$ (SM) and $\kappa_t = -1$. A neural network is trained to separate these two cases, and two analysis categories are introduced, one targeting the $tHqb$ $\kappa_t = +1$ (SM) events labeled "$tHqb$ (SM-like)", and the other targeting $\kappa_t = -1$ events labeled "$tHqb$ (BSM-like)". Two further neural networks are trained to separate each signal process ($tHqb$ $\kappa_t = +1$ (SM) and $\kappa_t = -1$) from the non-resonant background, and are used to maximise sensitivity to the two objectives. This novel categorisation is covered in detail in Chapter 9.

8.3.7 Categorisation Summary

The result is a total of 101 categories. The expected signal and background yields, purity, and significance in each category are shown in Figure 8.4. The expected event yields each category contributes to the targeted STXS class is shown in Figure 8.6.

Table 8.4 *The expected signal (S) and background (B) event yields in each analysis category in the* $m_{\gamma\gamma}$ *window containing 90% of the fitted signal events. The half width of the fitted signal events is given by* σ*. The signal purity* $f = S/(S+B)$ *and the expected significance* $Z = \sqrt{2((S+B)\ln(1+S/B)-S)}$ *are also shown. Only the signal yield originating from the targeted STXS region is considered in each category.*

Category	S	B	σ [GeV]	f [%]	Z
$gg \to H$					
0-jet, $p_T^H < 10$ GeV	695	26 000	3.43	2.6	4.3
0-jet, $p_T^H \geq 10$ GeV	1440	47 000	3.41	3.0	6.6
1-jet, $p_T^H < 60$ GeV, High-purity	168	4250	3.20	3.8	2.6
1-jet, $p_T^H < 60$ GeV, Med-purity	197	11 500	3.38	1.7	1.8
1-jet, $60 \leq p_T^H < 120$ GeV, High-purity	186	3310	3.10	5.3	3.2
1-jet, $60 \leq p_T^H < 120$ GeV, Med-purity	180	7780	3.37	2.3	2.0
1-jet, $120 \leq p_T^H < 200$ GeV, High-purity	23.0	182	2.61	11	1.7
1-jet, $120 \leq p_T^H < 200$ GeV, Med-purity	40.7	717	3.00	5.4	1.5
$\geq$ 2-jets, $m_{jj} < 350$ GeV, $p_T^H < 60$ GeV, High-purity	23.5	1050	3.08	2.2	0.72
$\geq$ 2-jets, $m_{jj} < 350$ GeV, $p_T^H < 60$ GeV, Med-purity	43.1	4360	3.39	0.98	0.65
$\geq$ 2-jets, $m_{jj} < 350$ GeV, $p_T^H < 60$ GeV, Low-purity	47.5	16 800	3.51	0.28	0.37
$\geq$ 2-jets, $m_{jj} < 350$ GeV, $60 \leq p_T^H < 120$ GeV, High-purity	49.1	901	3.03	5.2	1.6
$\geq$ 2-jets, $m_{jj} < 350$ GeV, $60 \leq p_T^H < 120$ GeV, Med-purity	93.9	6440	3.30	1.4	1.2
$\geq$ 2-jets, $m_{jj} < 350$ GeV, $120 \leq p_T^H < 200$ GeV, High-purity	15.5	74.8	2.64	17	1.7
$\geq$ 2-jets, $m_{jj} < 350$ GeV, $120 \leq p_T^H < 200$ GeV, Med-purity	22.7	343	2.97	6.2	1.2
$\geq$ 2-jets, $350 \leq m_{jj} < 700$ GeV, $p_T^H < 200$ GeV, High-purity	4.31	47.5	2.72	8.3	0.62
$\geq$ 2-jets, $350 \leq m_{jj} < 700$ GeV, $p_T^H < 200$ GeV, Med-purity	15.4	380	3.02	3.9	0.78
$\geq$ 2-jets, $350 \leq m_{jj} < 700$ GeV, $p_T^H < 200$ GeV, Low-purity	10.5	1080	3.31	0.97	0.32
$\geq$ 2-jets, $700 \leq m_{jj} < 1000$ GeV, $p_T^H < 200$ GeV, High-purity	2.34	33.3	2.84	6.6	0.40
$\geq$ 2-jets, $700 \leq m_{jj} < 1000$ GeV, $p_T^H < 200$ GeV, Med-purity	4.23	136	3.07	3.0	0.36
$\geq$ 2-jets, $700 \leq m_{jj} < 1000$ GeV, $p_T^H < 200$ GeV, Low-purity	3.34	429	3.26	0.77	0.16
$\geq$ 2-jets, $m_{jj} \geq 1000$ GeV, $p_T^H < 200$ GeV, High-purity	1.14	14.5	2.97	7.3	0.30
$\geq$ 2-jets, $m_{jj} \geq 1000$ GeV, $p_T^H < 200$ GeV, Med-purity	2.52	47.5	3.10	5.0	0.36
$\geq$ 2-jets, $m_{jj} \geq 1000$ GeV, $p_T^H < 200$ GeV, Low-purity	2.49	142	3.37	1.7	0.21
$200 \leq p_T^H < 300$ GeV, High-purity	15.3	38.0	2.28	29	2.3
$200 \leq p_T^H < 300$ GeV, Med-purity	29.4	236	2.64	11	1.9
$300 \leq p_T^H < 450$ GeV, High-purity	1.52	2.13	2.02	42	0.95
$300 \leq p_T^H < 450$ GeV, Med-purity	6.75	17.7	2.16	28	1.5
$300 \leq p_T^H < 450$ GeV, Low-purity	4.66	43.1	2.46	9.8	0.70
$450 \leq p_T^H < 650$ GeV, High-purity	1.00	1.25	1.85	45	0.81
$450 \leq p_T^H < 650$ GeV, Med-purity	0.800	2.00	1.98	29	0.53
$450 \leq p_T^H < 650$ GeV, Low-purity	0.830	10.7	2.19	7.2	0.25
$p_T^H \geq 650$ GeV	0.220	1.08	1.73	17	0.20
$qq' \to Hqq'$					
0-jet, High-purity	0.330	25.0	3.33	1.3	0.07
0-jet, Med-purity	1.27	471	3.35	0.27	0.06
0-jet, Low-purity	10.7	18 800	3.48	0.06	0.08
1-jet, High-purity	1.08	2.78	2.99	28	0.61
1-jet, Med-purity	3.50	26.1	3.11	12	0.67
1-jet, Low-purity	2.88	145	3.24	2.0	0.24
$\geq$ 2-jets, $m_{jj} < 60$ GeV, High-purity	0.350	2.10	2.71	14	0.24
$\geq$ 2-jets, $m_{jj} < 60$ GeV, Med-purity	0.670	19.0	2.79	3.4	0.15
$\geq$ 2-jets, $m_{jj} < 60$ GeV, Low-purity	1.92	243	2.93	0.78	0.12
$\geq$ 2-jets, $60 \leq m_{jj} < 120$ GeV, High-purity	3.45	6.34	2.65	35	1.3
$\geq$ 2-jets, $60 \leq m_{jj} < 120$ GeV, Med-purity	4.99	43.0	2.85	10	0.75
$\geq$ 2-jets, $60 \leq m_{jj} < 120$ GeV, Low-purity	2.99	87.3	3.01	3.3	0.32
$\geq$ 2-jets, $120 \leq m_{jj} < 350$ GeV, High-purity	2.98	24.4	2.93	11	0.59
$\geq$ 2-jets, $120 \leq m_{jj} < 350$ GeV, Med-purity	6.73	204	2.94	3.2	0.47
$\geq$ 2-jets, $120 \leq m_{jj} < 350$ GeV, Low-purity	8.78	1360	2.99	0.64	0.24
$\geq$ 2-jets, $350 \leq m_{jj} < 700$ GeV, $p_T^H < 200$ GeV, High-purity	2.52	2.75	2.96	48	1.4
$\geq$ 2-jets, $350 \leq m_{jj} < 700$ GeV, $p_T^H < 200$ GeV, Med-purity	9.15	34.7	3.06	21	1.5
$\geq$ 2-jets, $350 \leq m_{jj} < 700$ GeV, $p_T^H < 200$ GeV, Low-purity	5.97	106	3.27	5.3	0.57
$\geq$ 2-jets, $700 \leq m_{jj} < 1000$ GeV, $p_T^H < 200$ GeV, High-purity	2.91	3.00	2.90	49	1.5
$\geq$ 2-jets, $700 \leq m_{jj} < 1000$ GeV, $p_T^H < 200$ GeV, Med-purity	5.60	22.7	3.11	20	1.1
$\geq$ 2-jets, $m_{jj} \geq 1000$ GeV, $p_T^H < 200$ GeV, High-purity	10.8	3.89	3.01	74	4.2
$\geq$ 2-jets, $m_{jj} \geq 1000$ GeV, $p_T^H < 200$ GeV, Med-purity	10.7	19.0	3.23	36	2.3
$\geq$ 2-jets, $350 \leq m_{jj} < 700$ GeV, $p_T^H \geq 200$ GeV, High-purity	1.31	2.19	2.48	37	0.81
$\geq$ 2-jets, $350 \leq m_{jj} < 700$ GeV, $p_T^H \geq 200$ GeV, Med-purity	1.40	9.22	2.49	13	0.45
$\geq$ 2-jets, $350 \leq m_{jj} < 700$ GeV, $p_T^H \geq 200$ GeV, Low-purity	1.16	65.5	2.54	1.7	0.14
$\geq$ 2-jets, $700 \leq m_{jj} < 1000$ GeV, $p_T^H \geq 200$ GeV, High-purity	2.51	3.02	2.43	45	1.3
$\geq$ 2-jets, $700 \leq m_{jj} < 1000$ GeV, $p_T^H \geq 200$ GeV, Med-purity	1.49	47.4	2.54	3.0	0.22
$\geq$ 2-jets, $m_{jj} \geq 1000$ GeV, $p_T^H \geq 200$ GeV, High-purity	5.65	1.57	2.39	78	3.3
$\geq$ 2-jets, $m_{jj} \geq 1000$ GeV, $p_T^H \geq 200$ GeV, Med-purity	2.96	6.31	2.55	32	1.1
$qq \to H\ell\nu$					
$p_T^V < 75$ GeV, High-purity	1.91	4.91	3.17	28	0.81
$p_T^V < 75$ GeV, Med-purity	2.59	20.2	3.28	11	0.57
$75 \leq p_T^V < 150$ GeV, High-purity	2.62	2.05	3.02	56	1.6
$75 \leq p_T^V < 150$ GeV, Med-purity	2.08	12.4	3.23	14	0.58
$150 \leq p_T^V < 250$ GeV, High-purity	1.74	2.06	2.78	46	1.1
$150 \leq p_T^V < 250$ GeV, Med-purity	0.16	2.90	3.17	5.2	0.09
$p_T^V \geq 250$ GeV, High-purity	1.36	1.79	2.41	43	0.91
$p_T^V \geq 250$ GeV, Med-purity	0.02	3.12	3.15	0.78	0.01
$pp \to H\ell\ell$					
$p_T^V < 75$ GeV, High-purity	1.14	1.82	3.25	39	0.78
$p_T^V < 75$ GeV, Med-purity	1.06	215	3.29	0.49	0.07
$75 \leq p_T^V < 150$ GeV, High-purity	1.07	1.58	3.08	40	0.77
$75 \leq p_T^V < 150$ GeV, Med-purity	0.02	1.81	3.06	1.2	0.02
$150 \leq p_T^V < 250$ GeV, High-purity	0.71	1.79	2.78	28	0.50
$150 \leq p_T^V < 250$ GeV, Med-purity	0.10	16.5	2.88	0.62	0.03
$p_T^V \geq 250$ GeV	0.27	2.06	2.48	12	0.18
$pp \to H\nu\bar{\nu}$					
$p_T^V < 75$ GeV, High-purity	0.60	170	3.50	0.35	0.05
$p_T^V < 75$ GeV, Med-purity	1.15	1020	3.57	0.11	0.04
$p_T^V < 75$ GeV, Low-purity	0.87	2630	3.67	0.03	0.02
$75 \leq p_T^V < 150$ GeV, High-purity	0.58	2.30	2.97	20	0.37
$75 \leq p_T^V < 150$ GeV, Med-purity	1.83	17.8	3.26	9.3	0.43
$75 \leq p_T^V < 150$ GeV, Low-purity	2.18	288	3.44	0.75	0.13
$150 \leq p_T^V < 250$ GeV, High-purity	0.92	2.00	2.75	32	0.61
$150 \leq p_T^V < 250$ GeV, Med-purity	0.75	2.54	2.94	23	0.45
$150 \leq p_T^V < 250$ GeV, Low-purity	0.26	11.7	3.28	2.2	0.08
$p_T^V \geq 250$ GeV, High-purity	0.67	1.55	2.46	30	0.50
$p_T^V \geq 250$ GeV, Med-purity	0.05	1.97	3.05	2.6	0.04
$t\bar{t}H$					
$p_T^H < 60$ GeV, High-purity	3.04	4.01	3.18	43	1.4
$p_T^H < 60$ GeV, Med-purity	2.78	13.3	3.37	17	0.74
$60 \leq p_T^H < 120$ GeV, High-purity	4.30	4.09	3.06	51	1.9
$60 \leq p_T^H < 120$ GeV, Med-purity	2.99	8.61	3.31	26	0.97
$120 \leq p_T^H < 200$ GeV, High-purity	4.65	3.52	2.73	57	2.1
$120 \leq p_T^H < 200$ GeV, Med-purity	1.66	4.16	2.93	29	0.77
$200 \leq p_T^H < 300$ GeV	3.39	2.26	2.46	60	1.9
$p_T^H \geq 300$ GeV	2.73	1.66	2.12	62	1.8
tH					
$tHqb$, High-purity	0.55	2.16	3.04	20	0.36
$tHqb$, Med-purity	0.14	2.78	3.45	4.9	0.09
$tHqb$, BSM ($\kappa_t = -1$)	0.12	1.86	3.25	6.0	0.09
tHW	0.16	6.91	2.74	2.3	0.06
Low-purity top	5.18	65.8	3.32	7.3	0.63

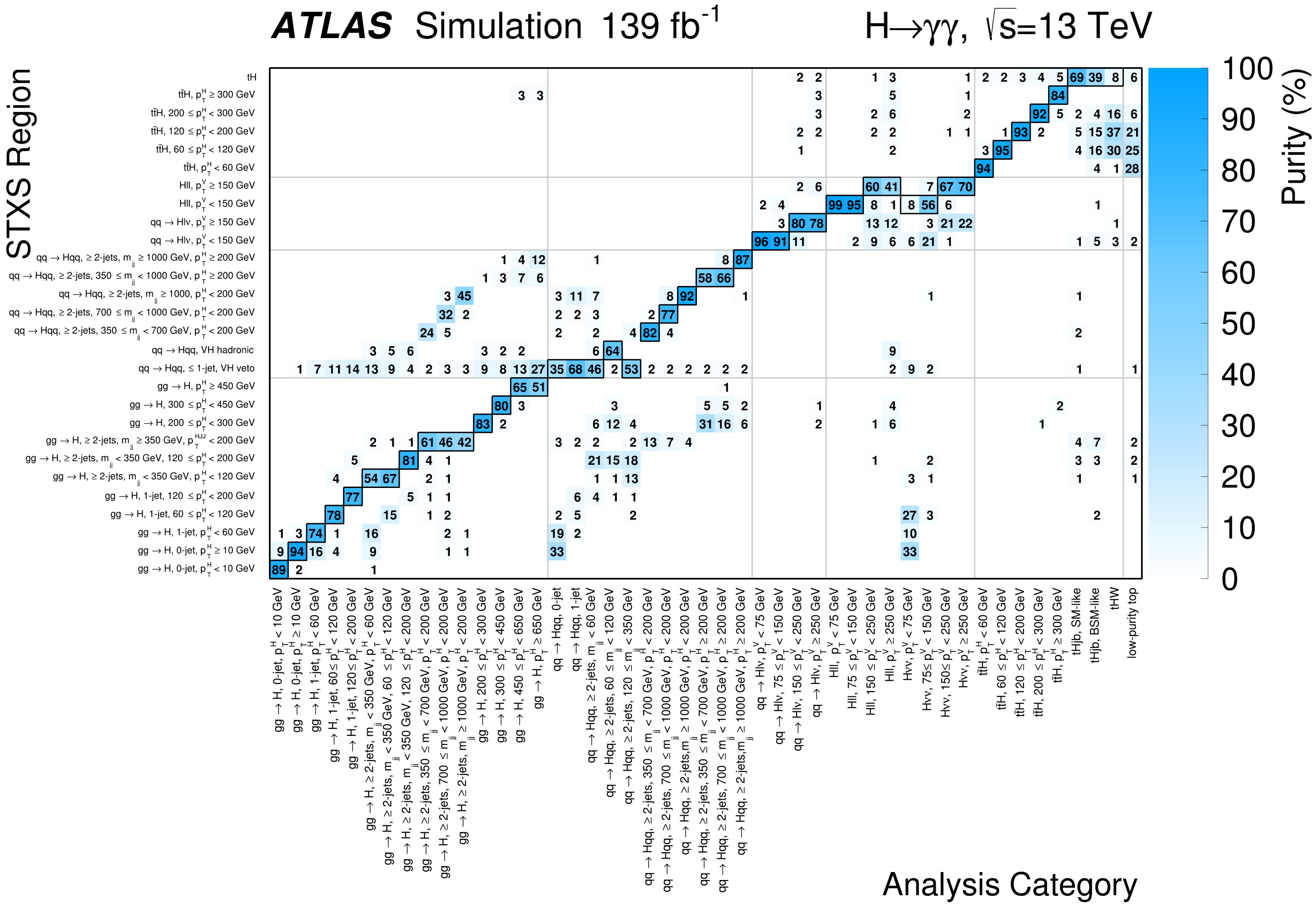


Figure 8.6 *Contribution of the grouped analysis categories to the event yield in each of the targeted STXS regions. The vertical axis lists the 28 targeted STXS regions, while the horizontal axis lists the grouped analysis categories. Contributions from individual categories are weighted by their purity shown in Table 8.4. Vertical entries sum to 100(%). Entries with a purity of less than 1% are not shown.*

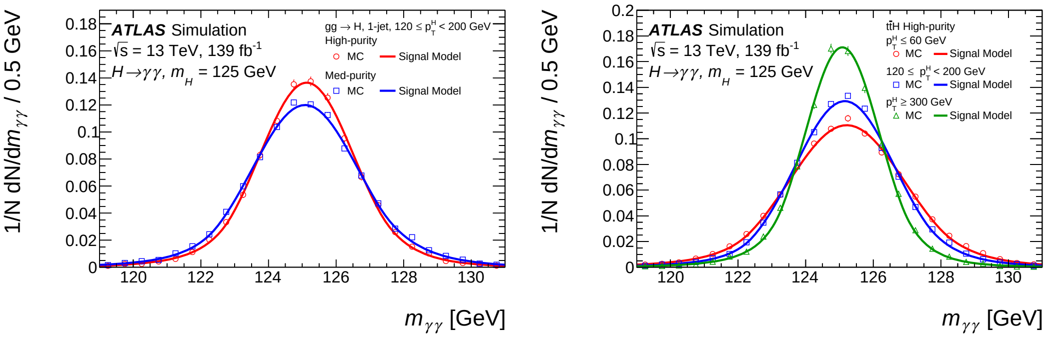
8.4 Signal & Background Modelling

In each of the analysis categories, the $m_{\gamma\gamma}$ distribution is described using an extended probability density function (pdf) in which the signal and background are modelled using analytical functions. The following section describes the function used to model the signal shape, and the procedure used to select the best function to model the non-resonant background in each analysis category.

8.4.1 Signal Modelling

The signal shape in each analysis category is modelled using the *double-sided* Crystal Ball (DSCB) function. The DSCB function is made up of three components; a Gaussian core that models the signal peak, and two power-law components that model the tails of the signal shape.

The signal model is fit to the signal sample events in the $m_{\gamma\gamma}$ spectrum in the mass range 110 to 140 GeV. The contribution of each signal sample is normalised to the cross-section of the given process. All relevant production processes are considered: ggF, VBF, VH, $t\bar{t}H$ and tH. Examples of the signal model in two groups of categories can be seen in Figure 8.7.



(a) *gg → H, 1-jet,* 120 ≤ p_T^H < 200 GeV   **(b)** $t\bar{t}H$

Figure 8.7 *Examples of the signal $m_{\gamma\gamma}$ distribution using signal MC in two groups of categories. 8.7a shows the signal $m_{\gamma\gamma}$ shapes for two categories targeting the same STXS region with different purities. 8.7b compares the signal $m_{\gamma\gamma}$ shapes for three 'high-purity' categories targeting different p_T^H regions of the $t\bar{t}H$ process. The markers represent distributions in MC samples with $m_H = 125\,\text{GeV}$, while the solid lines represent the corresponding fitted DSCB functions. Taken from [1]*

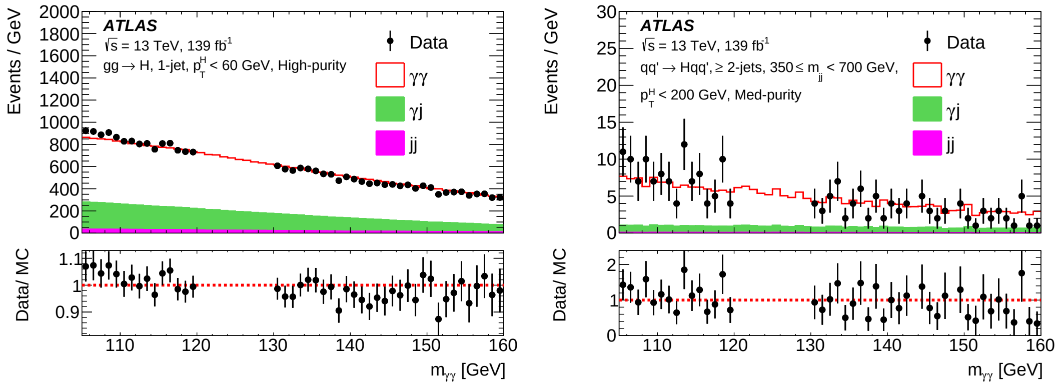
8.4.2 Background Modelling

The analysis uses a functional approach to model the background in each of the analysis categories. This constitutes two main steps: first a background template is constructed using a combination of simulated samples and data control regions, secondly, a function is selected to model the background using the *spurious signal* test.

Background Templates

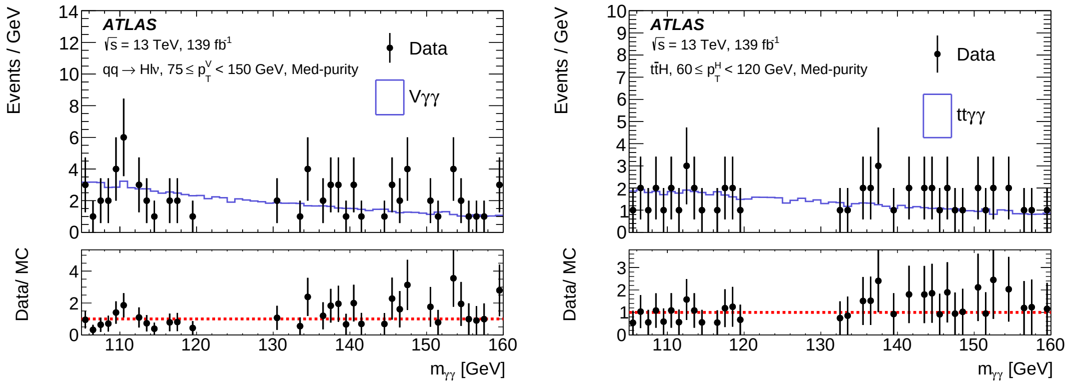
The samples used to model the continuum background depend on which process a category is targeting. In categories targeting the $gg \to H$ and $qq' \to Hqq'$ processes, the background template is defined using the $\gamma\gamma$, γj and jj samples. For categories targeting the $V(\to \ell\ell)H$ process the background template is constructed using the $V\gamma\gamma$ sample. For categories targeting the $t\bar{t}H$ and tH processes the background template is constructed using the $t\bar{t}\gamma\gamma$ sample. The background templates constructed for four analysis categories targeting the $gg \to H$ $qq' \to Hqq'$ VH and $t\bar{t}H$ processes are shown in Figure 8.8.

The background templates are defined in the range $105 < m_{\gamma\gamma} < 160\,\text{GeV}$ with 220 uniform-width bins. A smoothing technique using a Gaussian kernel [72] is applied to background templates where the average bin occupancy has at least 20 entries. This smoothing has the effect of reducing statistical fluctuations in the templates.



(a) $gg \to H$, *1-jet*, $p_{\mathrm{T}}^{H} < 60\,\mathrm{GeV}$, *High-purity*

(b) $qq' \to Hqq'$, ≥ 2*-jets*, $350 \leq m_{jj} < 700\,\mathrm{GeV}$, $p_{\mathrm{T}}^{H} < 200\,\mathrm{GeV}$, *Med-purity*



(c) $qq \to H\ell\nu$, $75 \leq p_{\mathrm{T}}^{V} < 150\,\mathrm{GeV}$, *Med-purity*

(d) $t\bar{t}H$, $60 \leq p_{\mathrm{T}}^{H} < 120\,\mathrm{GeV}$, *Med-purity*

Figure 8.8 *The diphoton invariant mass* $m_{\gamma\gamma}$ *distribution in data (black points) and continuum background templates (histograms) in four representative STXS categories. The data are shown excluding the region* $120 < m_{\gamma\gamma} < 130\,\mathrm{GeV}$ *containing the signal. In panels 8.8a and 8.8b, stacked histograms corresponding to the* $\gamma\gamma$ *(white),* γj *(green) and* jj *(magenta) background contributions are shown. In panel 8.8c, the white histogram represents contributions from* $V\gamma\gamma$ *and other sources of prompt* $\gamma\gamma$ *production. In 8.8d, the white histogram corresponds to simulated* $t\bar{t}\gamma\gamma$ *events. The templates do not represent the background shapes used in the analysis, but are used to identify flexible functions used to model the background in each category as described in the text. Taken from [1].*

Spurious Signal

With the background templates defined, the *spurious signal* test is used to select a function to model the background. In this choice there is a trade off, a function needs to be flexible enough to describe the background template, but not so flexible that incorrectly it captures statistical fluctuations in the background template. In the spurious signal test a number of functions are considered, these are summarised in Table 8.5.

Table 8.5 *Summary of the functions used for the modelling of the continuum background component.* N_{pars} *denotes the number of free parameters used to define the function shape. For the definition of the Bernstein polynomials,* $x = (m_{\gamma\gamma} - m_{min})/(m_{max} - m_{min})$*, where* $m_{min} = 105\,\text{GeV}$ *and* $m_{max} = 160\,\text{GeV}$ *are respectively the lower and upper bounds of the fitted* $m_{\gamma\gamma}$ *range. Taken from [1].*

Type	Function	N_{pars}	Acronym
Power law	$m_{\gamma\gamma}^{a}$	1	`PowerLaw`
Bernstein polynomial	$(1-x)^n + a_1 n x(1-x)^{n-1} + \cdots + a_n x^n$	$n = 1$–5	`Bern1`–`Bern5`
Exponential	$\exp(a m_{\gamma\gamma})$	1	`Exp`
Exponential of second-order polynomial	$\exp\big(a_1 m_{\gamma\gamma} + a_2 m_{\gamma\gamma}^2\big)$	2	`ExpPoly2`
Exponential of third-order polynomial	$\exp\big(a_1 m_{\gamma\gamma} + a_2 m_{\gamma\gamma}^2 + a_3 m_{\gamma\gamma}^3\big)$	3	`ExpPoly3`

To perform the *spurious signal* test for a given function the full analytical signal plus background model is fitted to the background only template in the range $105 \leq m_{\gamma\gamma} \leq 160\,\text{GeV}$. The maximum absolute value of fitted signal yield in the range $123 \leq m_{\gamma\gamma} \leq 127$ GeV, using a step of 0.5 GeV, is taken as a spurious signal. Negative fluctuations in the background model are considered in the fit. The number of *spurious signal* events N_{sp} is then taken as the maximum absolute number of signal events fitted by the signal model. The spurious signal test is passed if at least one of the following conditions is met:

- $N_{sp} < 10\% N_{s,exp}$, where $N_{s,exp}$ is the expected number of signal events in the category
- $N_{sp} < 20\% \sigma_{bkg}$. where σ_{bkg} is the statistical uncertainty on the fitted number of signal events when fitting the background plus signal model to the background template

A further requirement is made on the analytical function in question; the fit is required to yield a χ^2 probability of at least 1%.

If multiple functions pass the *spurious signal* requirement then the function with the smallest number of degrees of freedom is chosen to model the background.

In categories where the average number of events per bin in $m_{\gamma\gamma}$ is less than 20, the candidate functions are limited to the exponential, and exponential of second and third order polynomial. This restricted set of functions allows us to avoid potential nonphysical fits due to statistical fluctuations. The function is selected using the Wald test [73].

8.5 Systematic Uncertainties

Two categories of systematic uncertainties are considered: those that affect the modelling of the $m_{\gamma\gamma}$ distribution of both the signal and background, and those that affect the expected signal yield in each of the analysis categories. Uncertainties can come from both experimental and theoretical sources. Each of the systematic uncertainties is incorporated into the likelihood model as a nuisance parameter.

Table 8.6 shows a subset of the expected experimental and theoretical systematic uncertainties of the cross-section measurements in the SM hypothesis.

Table 8.6 *Expected contributions from the main sources of systematic uncertainty to the total uncertainty in the measurement of the cross-section times $H \to \gamma\gamma$ branching ratio for each of of the main Higgs boson production processes. The uncertainty from each source ($\Delta\sigma$) is shown as a fraction of the total expected cross-section (σ). Only uncertainty sources that contribute $\geq$ 1% to each expected cross-section are shown. Adapted from [1].*

	ggF + $b\bar{b}$H	VBF	WH	ZH	$t\bar{t}H$	tH
Uncertainty source	$\Delta\sigma$[%]	$\Delta\sigma$[%]	$\Delta\sigma$[%]	$\Delta\sigma$[%]	$\Delta\sigma$[%]	$\Delta\sigma$[%]
Theory uncertainties						
Higher-order QCD terms	±1.4	±4.1	±4.1	±12	±2.8	±16
Underlying event and parton shower	±2.5	±16	±2.5	±4.0	±3.6	±48
Experimental uncertainties						
Photon energy resolution	±3.0	±3.0	±3.8	±4.8	±3.0	±12
Photon efficiency	±2.7	±2.7	±3.3	±3.6	±2.9	±9.3
Luminosity	±1.8	±2.0	±2.4	±2.7	±2.2	±6.6
Pile-up	±1.4	±2.2	±2.0	±2.3	±1.4	±7.3
Background modelling	±2.0	±4.6	±3.6	±7.2	±2.5	±63

8.6 Results

The main results of this analysis are presented in terms of the following measurements of Higgs boson production.

- The overall signal strength of Higgs boson production in the diphoton decay channel (Section 8.6.1).
- The measurement of the cross-sections for the main Higgs boson production modes (Section 8.6.2).
- Cross-section measurements in a set of merged STXS regions defined for each production process (Section 8.6.3).

The following sections detail each of these results.

8.6.1 Overall Higgs boson signal strength

The overall Higgs boson signal strength, μ, including the inclusive production cross-section, and $H\rightarrow\gamma\gamma$ branching ratio in the fiducial region $|y_H| < 2.5$ is measured with respect to the SM expectation. The measurement is conducted by simultaneously fitting the $m_{\gamma\gamma}$ distributions in each of the 101 analysis categories. The signal strength is measured to be,

$$\mu = 1.04^{+0.10}_{-0.09} = 1.04 \pm 0.06\ (\text{stat.})^{+0.06}_{-0.05}\ (\text{theory syst.})^{+0.05}_{-0.04}\ (\text{exp. syst.})$$

The inclusive $m_{\gamma\gamma}$ distribution is shown in Figure 8.9, where events are weighted by the $\ln(1 + S/B)$ to reflect the sensitivity they bring to the measurement.

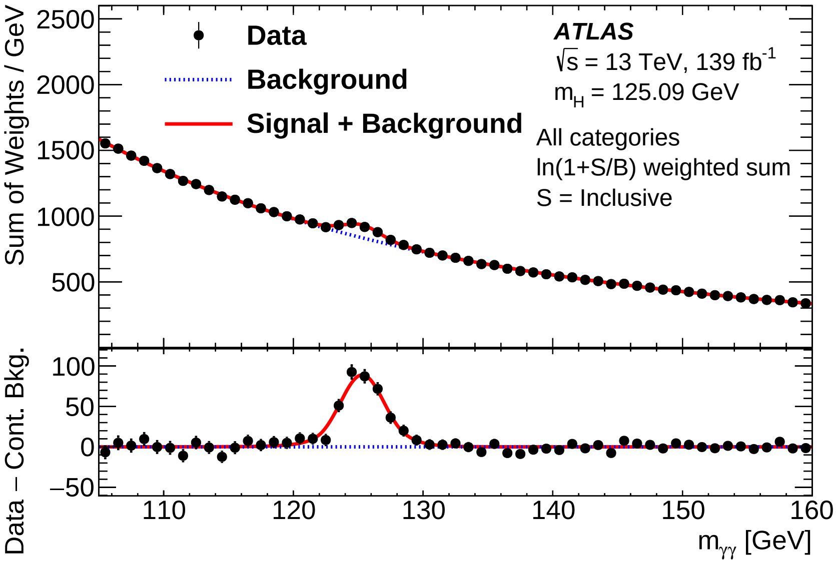
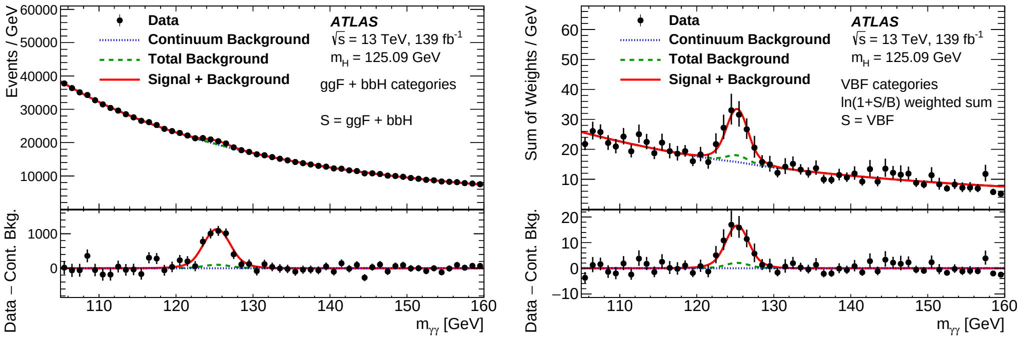


Figure 8.9 *The inclusive diphoton invariant mass distribution of events from all analysis categories. The data events (dots) in each category are weighted by* $\ln(1 + S/B)$*, where* S *and* B *are the expected signal and background yields in this category within the smallest* $m_{\gamma\gamma}$ *window containing 90% of the signal events. The expected signal is considered inclusively over all STXS regions. The fitted signal-plus-background pdfs from all categories are also weighted and summed, shown as the solid line. The blue dotted line represents the weighted sum of the fitted background functions from all categories. Taken from [1].*

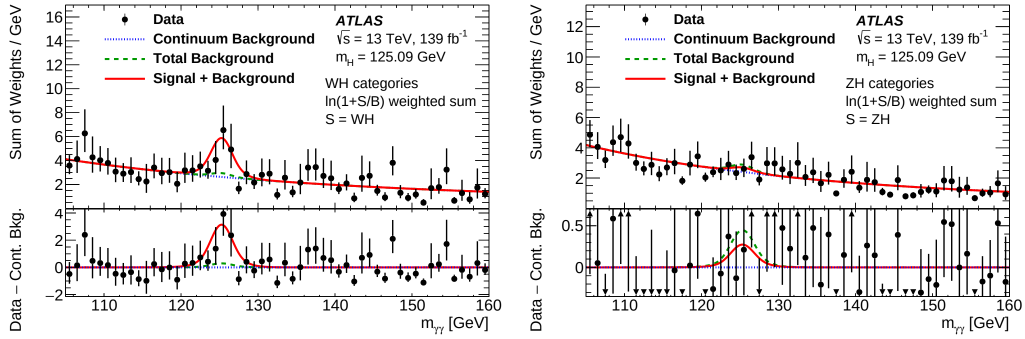
8.6.2 Production cross-sections

The cross-section for the following Higgs boson production modes is also measured separately: ggF, VBF, WH, ZH, $t\bar{t}H$, tH. Figure 8.10 shows the $m_{\gamma\gamma}$ spectrum associated with each measurement.

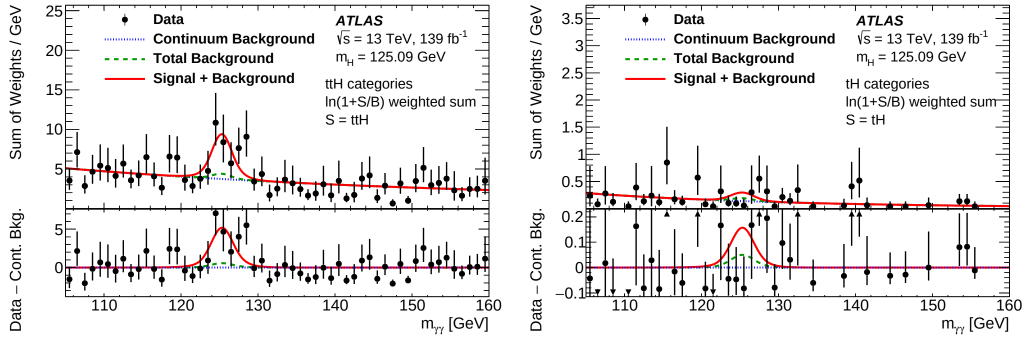
The cross-section measurements for each of the six Higgs boson production modes can be found in Figure 8.11. The total measurement is compatible with the SM, with a p-value of 55%. The correlations between each of the measurements are shown in Figure 8.12.



(a) $ggF + b\bar{b}H$ **(b)** VBF



(c) WH **(d)** ZH



(e) $t\bar{t}H$ **(f)** tH

Figure 8.10 *Combined diphoton invariant mass distributions for categories targeting the same production processes. The data (black dots) are weighted by* $\ln(1 + S/B)$ *where* S *and* B *are respectively the expected signal and background yields in the smallest* $m_{\gamma\gamma}$ *window containing 90% of the signal events. In this calculation, only Higgs boson events from the targeted production processes are considered as signal events. Higgs boson events from other processes as well as the continuum background events are considered as background. The fitted signal-plus-background pdfs from the relevant categories are summed, and represented by a solid line. The blue dotted line represents the weighted sum of the fitted continuum background pdfs, while the dashed line combines the contributions of continuum background and other Higgs boson events. Taken from [1].*

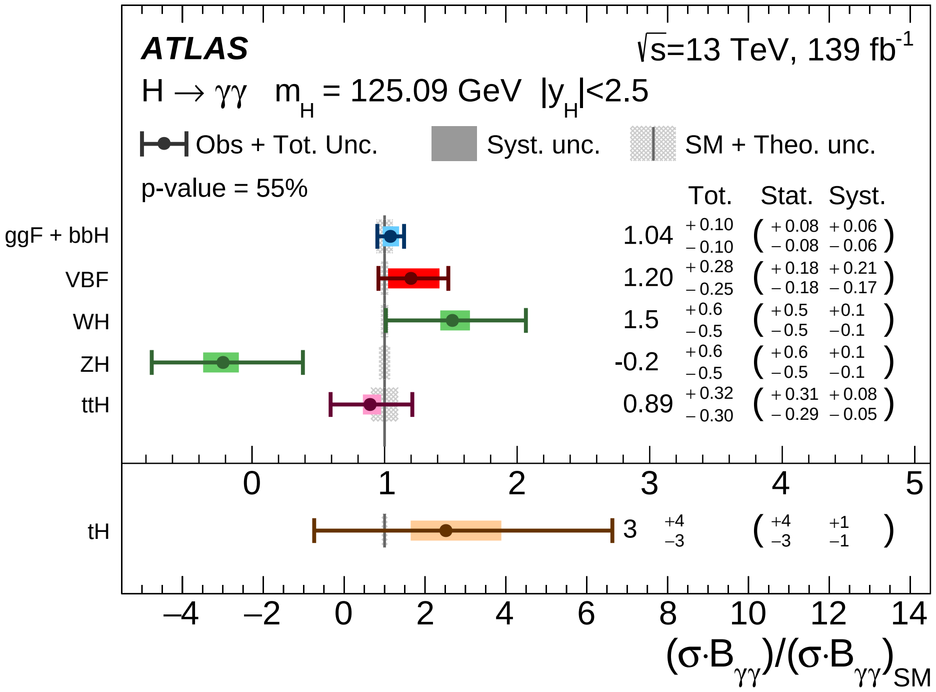


Figure 8.11 *Cross-sections times $H \to \gamma\gamma$ branching ratio for ggF+ $b\bar{b}H$, VBF, VH, $t\bar{t}H$, and tH production, normalized to their SM predictions. The values are obtained from a simultaneous fit to all categories. The error bars, hatched areas and shaded areas show respectively the total, systematic, and statistical uncertainties in the measurements. The grey bands show the theory uncertainties in the predictions, including uncertainties due to missing higher-order terms in the perturbative QCD calculations and choices of parton distribution functions and value of , as well as the $H \to \gamma\gamma$ branching ratio uncertainty. Taken from [1].*

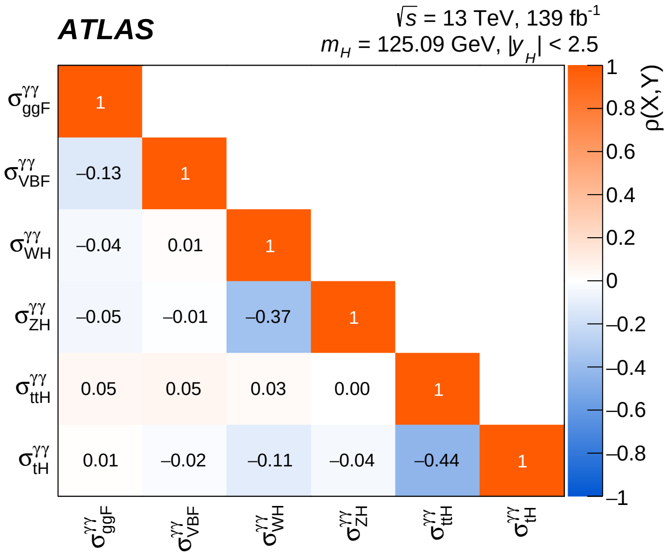


Figure 8.12 *Correlation matrix for the measurement of production cross-sections of the Higgs boson times the* $H \to \gamma\gamma$ *branching ratio. Taken from [1].*

8.6.3 Cross-sections in STXS regions

A measurement of the cross-sections in the 28 STXS regions is performed. The scheme in which cross-section measurements are performed differs slightly from the scheme targeted by the categorisation; regions are merged to avoid large statistical uncertainties, and large correlations between measurements. The merging scheme is chosen based on the expected sensitivity under the SM hypothesis.

Results are shown in Figure 8.13, and the correlation matrix of the measurement in Figure 8.14. Generally small correlations are observed between the cross-section measurements, with exceptions coming from processes with event topologies that are difficult to separate. Uncertainties on measurements range from 20% to greater than 100%. The statistical uncertainty component is generally much larger than the systematic component across the measurements. No significant deviations from the SM prediction in seen across the measurements, corresponding to a p-value of 93%.

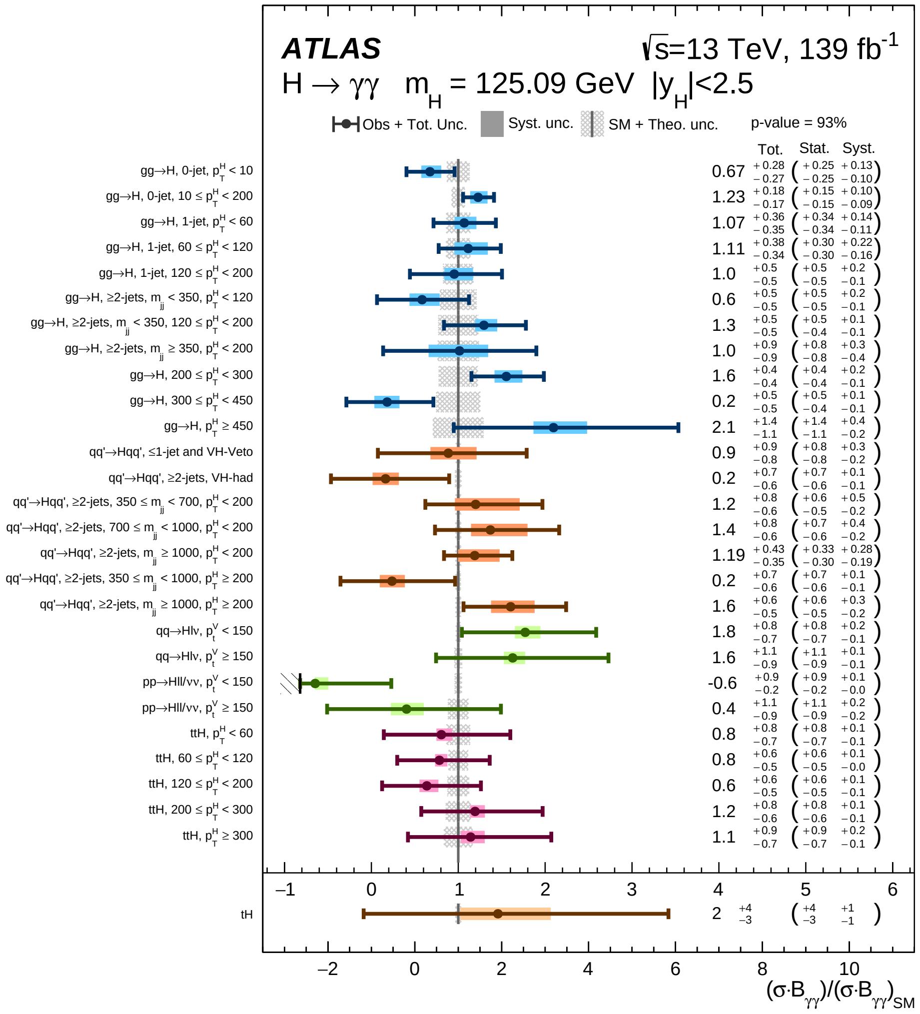


Figure 8.13 *Best-fit values and uncertainties for STXS parameters in each of the 28 regions considered, normalised to their SM predictions. The values for the $gg \to H$ process also include the contributions from $b\bar{b}H$ production. The black error bars, hatched boxes and shaded boxes show the total, systematic, and statistical uncertainties in the measurements, respectively. The uncertainties for the $pp \to H\ell\ell/\nu\bar{\nu}, p_{\mathrm{T}}^{V} < 150\,\mathrm{GeV}$ region are truncated at the value for which the model pdf becomes negative. The grey bands around the vertical line at $\sigma^{\gamma\gamma}/\sigma^{\gamma\gamma}_{SM}$ =1 show the theory uncertainties in the predictions, including uncertainties due to missing higher-order terms in the perturbative QCD calculations and choices of parton distribution functions and value of , as well as the $H \to \gamma\gamma$ branching ratio uncertainty. The p_{T} and m_{jj} values in the region definitions are indicated in* GeV, *and the 0J, 1J and 2J shorthands refer respectively to the 0-jet, 1-jet and ≥ 2-jets selections. Taken from [1].*

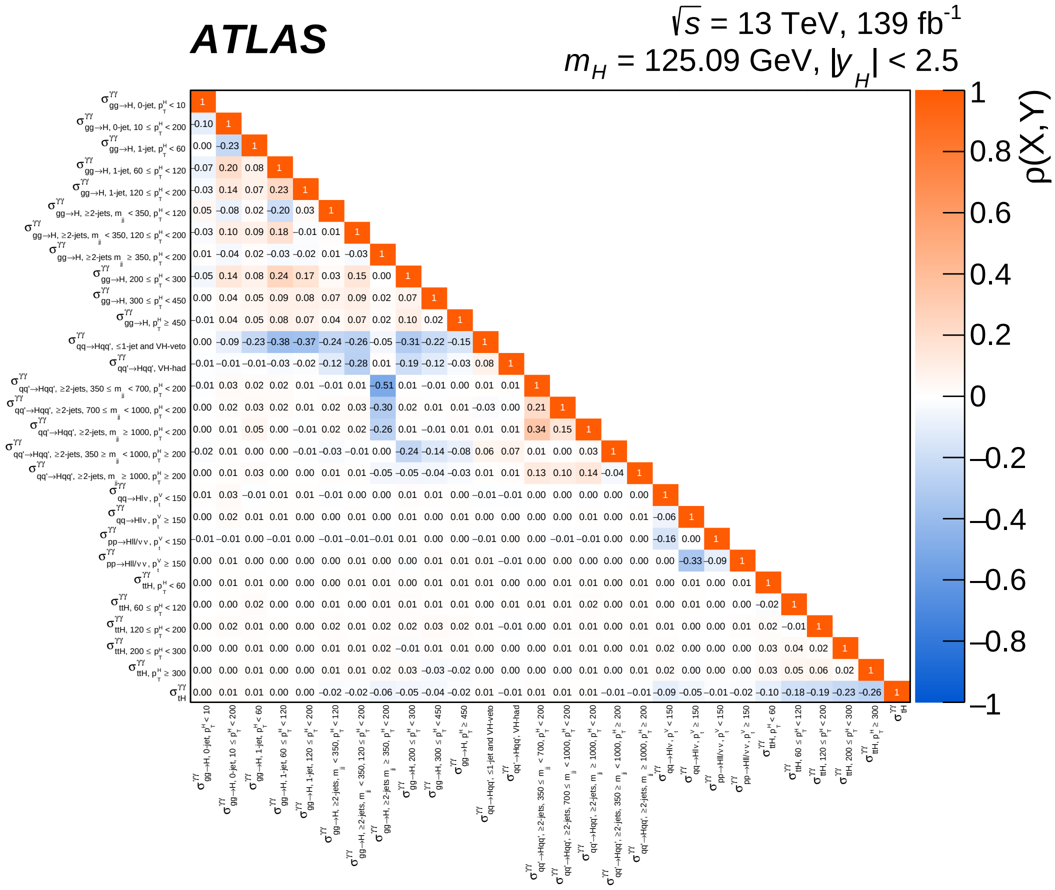


Figure 8.14 *Correlation matrix for the measurement of STXS parameters in each of the 28 regions considered. Taken from [1].*

8.7 Interpretation of the Results in the κ-framework

In this section, the interpretation of the analysis results in the κ framework (Section 8.0.1) is presented. The *resolved* parameterisation of the κ_g and κ_γ coupling modifiers are used in all models unless explicitly stated otherwise. Four specific models are considered in these results.

The first model focuses on the κ_t coupling modifier, where two configurations for the $gg \rightarrow H$ and $H \rightarrow \gamma\gamma$ loop processes are used. The first considers the case where their parameterisation are *resolved* in terms of the individual kappa modifiers. The second considers the case where the *effective* coupling modifiers κ_g and κ_γ are used. In each configuration, all other coupling modifiers are fixed to their SM expectation. Both configurations allow the sign of κ_t to be probed, where in each case contributions from different processes are considered. These contributions come from the tH and $gg \rightarrow ZH$ processes, as well as $H \rightarrow \gamma\gamma$ in the resolved configuration.

The negative log-likelihood scans for both configurations are shown in Figure 8.15. A good agreement can be seen with the SM value of $\kappa_t = +1$.

In the first configuration, where the $gg \rightarrow H$ and $H \rightarrow \gamma\gamma$ loops are resolved, negative values of κ_t are excluded with a significance of $> 6.7\sigma$. Values of κ_t outside of the range $0.87 < \kappa_t < 1.20$ are excluded at 95% CL ($0.85 < \kappa_t < 1.19$ expected).

In the second effective configuration, negative values of κ_t are excluded with $> 2.2\sigma$, where the majority of the sensitivity comes from the tH process and the novel categorisation of the $tHqb$ sub-process (Chapter 9). A small contribution also comes from the $gg \rightarrow ZH$ process. Values of κ_t outside of the range $0.65 < \kappa_t < 1.25$ are excluded at 95% CL ($0.71 < \kappa_t < 1.29$ expected).

The second model looks at the effective parameterisation of the $gg \rightarrow H$ and $H \rightarrow \gamma\gamma$ loops using the coupling modifiers κ_g and κ_γ. Both are assumed to be positive as the analysis measurement provides no sensitivity to their relative sign. All other coupling modifiers are fixed to their SM expectation. The negative log-likelihood contours and best fit values in the plane of $(\kappa_g, \kappa_\gamma)$ are shown in Figure 8.16. The best fit values of the two parameters are,

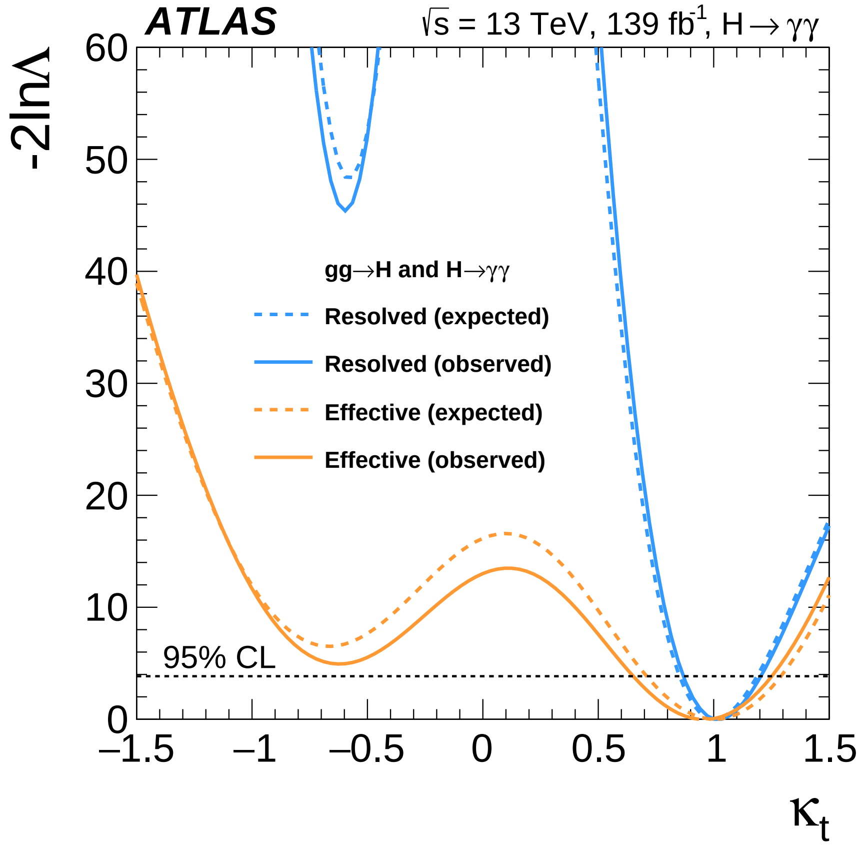


Figure 8.15 *Negative log-likelihood scans as a function of κ_t in a model where other coupling modifiers are fixed to their SM values. The $H \to \gamma\gamma$ and $gg \to H$ loops are either parameterised as a function of κ_t (blue) or fixed to their SM expectation (orange). In the latter case, sensitivity to the sign of κ_t is provided by the tH process, and to a lesser degree by the $gg \to ZH$ process. The solid curves correspond to observed data, and the dotted curves to an Asimov data set generated under the SM hypothesis. Taken from [1].*

$$\kappa_g = 1.01^{+\,0.11}_{-\,0.09},$$
$$\kappa_\gamma = 1.02^{+\,0.08}_{-\,0.07}.$$

Where a linear correlation between the two parameters of -79% is observed.

The third model considers two universal coupling modifiers: $\kappa_V = \kappa_W = \kappa_Z$ representing the coupling between Higgs boson and the gauge bosons, and $\kappa_F = \kappa_t = \kappa_b = \kappa_c = \kappa_\tau = \kappa_\mu$ representing the Higgs boson coupling to fermions. The $gg \to H$, $H \to \gamma\gamma$ and $gg \to ZH$ are described using their resolved parameterisations as a function of κ_V and κ_F. The negative log-likelihood

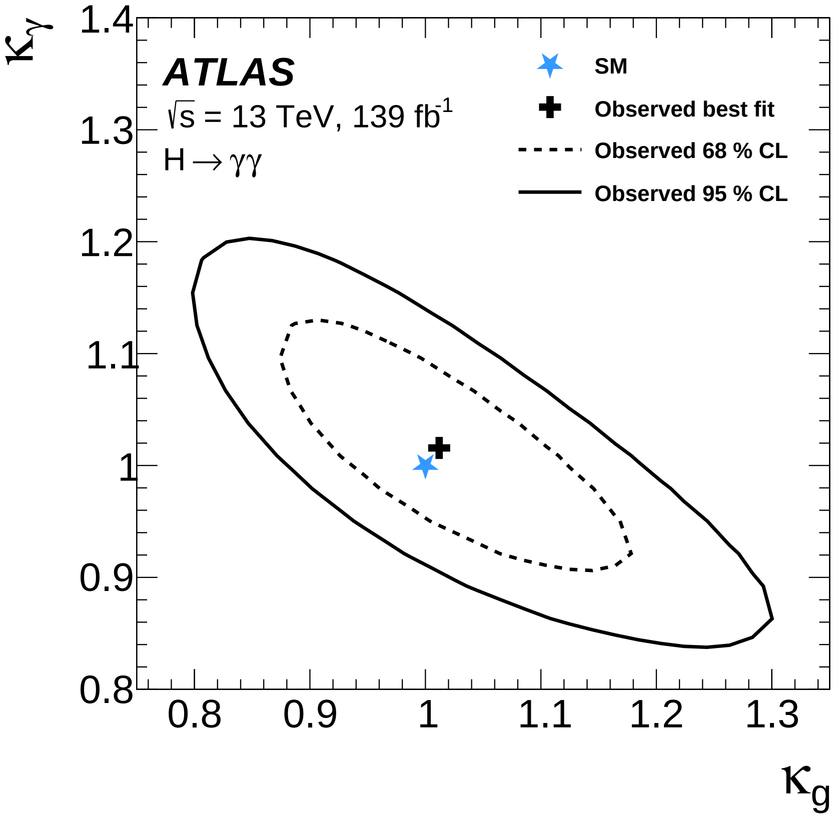


Figure 8.16 *Negative log-likelihood contours at 68% (dashed line) and 95% CL (solid line) in the (κ_g, κ_γ) plane, assuming that all other coupling-strength modifiers take their SM values. All other κ modifiers are fixed to their SM values. The best-fit point is indicated by a cross while the SM prediction is indicated by a star. Taken from [1].*

contours and best fit values in the plane (κ_V, κ_F) are shown in Figure 8.17. The best fit values of the two parameters are

$$\kappa_V = 1.02^{+\,0.06}_{-\,0.05}$$
$$\kappa_F = 1.00^{+\,0.16}_{-\,0.13}.$$

A final model is considered where ratios of coupling modifiers are constructed. This allows the κ_t coupling modifier to be probed without the need to make any assumptions about the total Higgs width. The κ_τ coupling modifier is fixed to 1 and $\kappa_b = \kappa_t$ is assumed. The total width of the Higgs boson is parameterised using the effective modifier κ_H, defined as $\Gamma_H = \kappa_H^2 \Gamma_H^{\text{SM}}$, where Γ_H^{SM} is the SM value of the Higgs width. The following measurement parameters are defined

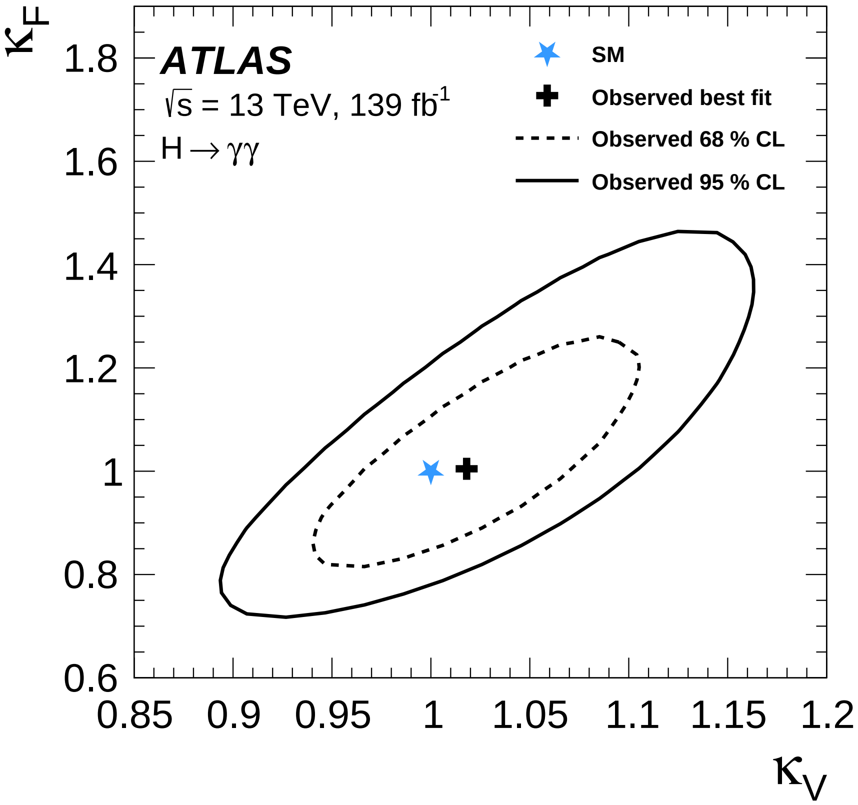


Figure 8.17 *Negative log-likelihood contours at 68% CL (dashed line) and 95% CL (solid line) in the (κ_V, κ_F) plane of modifiers applied to Higgs boson couplings to gauge bosons (κ_V) and fermions (κ_F). Loop processes and the Higgs boson total width Γ_H are parameterised as a function of κ_V and κ_F. The best-fit point is indicated by a cross, and the SM prediction by a star. Taken from [1].*

$$\begin{aligned}
\kappa_{g\gamma} &= \kappa_g \kappa_\gamma / \kappa_H \\
\lambda_{Vg} &= \kappa_V / \kappa_g \\
\lambda_{tg} &= \kappa_t / \kappa_g,
\end{aligned}$$

where the $gg \to H \to \gamma\gamma$ process is used as a reference to define the coupling modifier ratios. The λ_{tg} parameter is allowed to take positive or negative values, the other two are fixed to be positive. The negative log-likelihood scan of the λ_{tg} parameter is shown in Figure 8.18. Best fit values are shown in Table 8.7, and the linear correlation between the parameters is shown in Figure 8.19. Negative values of λ_{tg} are excluded with a significance of $> 2.1\sigma$, where sensitivity is provided by the tH and $gg \to ZH$ processes.

Table 8.7 *Best-fit values and uncertainties in the coupling-modifier ratio model. The second column expresses the measured parameters in terms of the coupling modifiers. The SM corresponds to* $\kappa_{g\gamma} = \lambda_{tg} = \lambda_{Vg} = 1$.

Parameter	Definition in terms of κ modifiers	Result
$\kappa_{g\gamma}$	$\kappa_g \kappa_\gamma / \kappa_H$	1.02 ± 0.06
λ_{Vg}	κ_V / κ_g	1.01 ± 0.11
λ_{tg}	κ_t / κ_g	$0.95\,^{+\,0.15}_{-\,0.16}$

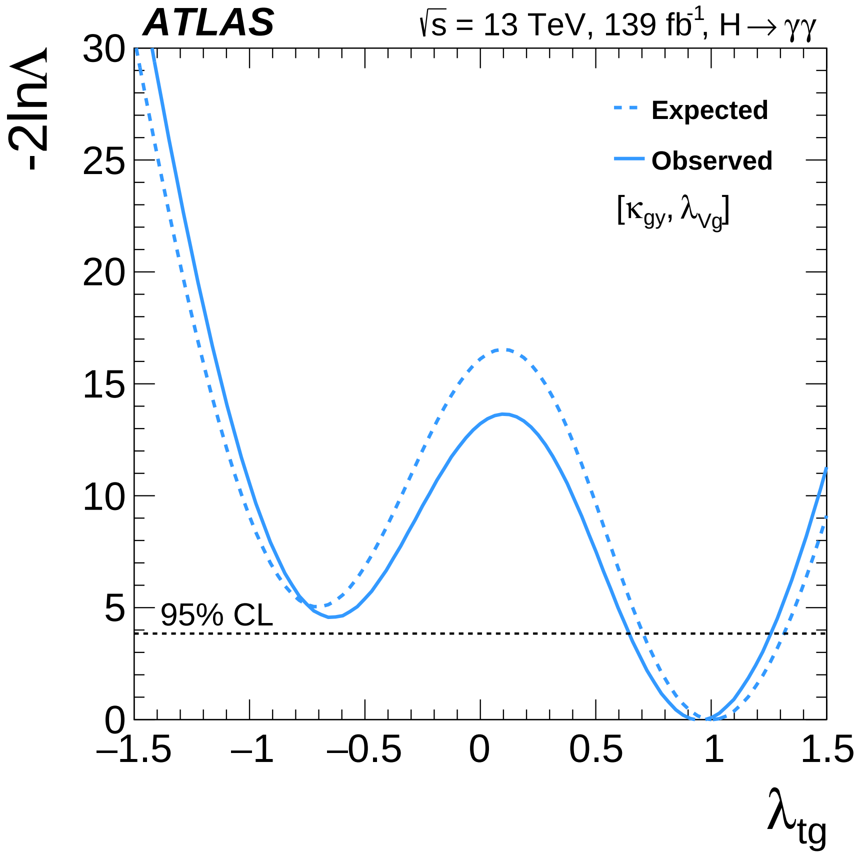


Figure 8.18 *Negative log-likelihood scan as a function of* $\lambda_{tg} = \kappa_t/\kappa_g$. *The solid curve corresponds to observed data, and the dotted curve to an Asimov data set generated under the SM hypothesis. Taken from [1].*

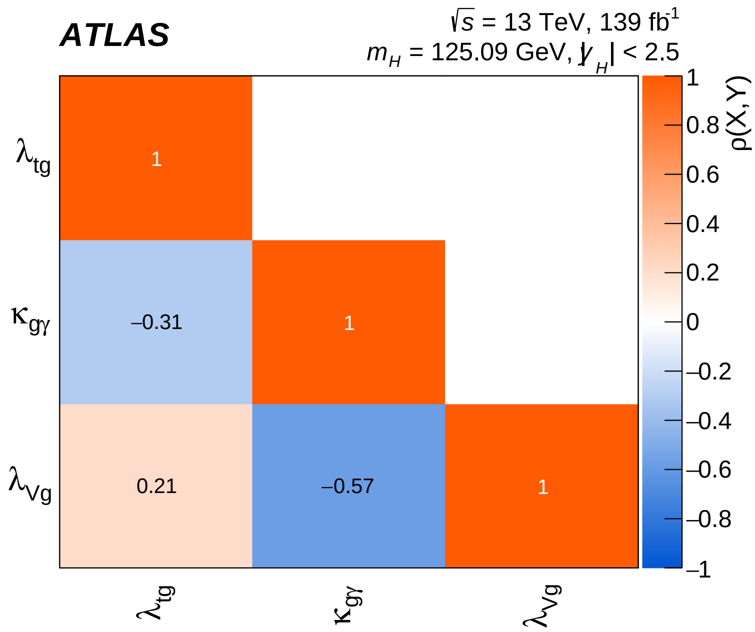


Figure 8.19 *The observed linear correlation between the three paramters,* $\kappa_{g\gamma}$*,* λ_{Vg} *, and* λ_{tg}*.*

Chapter 9

Constraining the Top Quark Yukawa Coupling in tH Production

This chapter presents an analysis of the tH process designed to optimise sensitivity to the sign of the top quark Yukawa coupling strength κ_t and the Standard Model (SM) tH cross-section. The development of a novel categorisation of the $tHqb$ sub-process using three neural network models is described. The categorisation is used to define three signal regions in the analysis detailed in Chapter 8 that target the $tHqb$ process. The categorisation enhances the expected sensitivity of the analysis detailed in Chapter 8 to the sign of κ_t, whilst simultaneously enhancing expected sensitivity to the SM tH cross-section measurement. Negative values of κ_t are directly excluded at $> 2.2\sigma$, and an upper limit on the tH cross-section is set at 10 times the SM expectation (Section 8.6).

9.1 The Top Quark Yukawa Coupling

The strength of the Yukawa coupling between the Higgs boson and the top quark in the SM is:

$$y_t^{SM} = \sqrt{2}\frac{m_t}{v}. \tag{9.1}$$

Measurements of the strength of y_t are conducted in the kappa framework (Section 8.0.1). Where κ_t is the ratio between the measured coupling strength and the SM expectation ($\kappa_t = y_t/y_t^{\mathrm{SM}}$). Throughout the rest of this chapter, the top quark Yukawa coupling strength will be referred to as κ_t.

9.2 Previous Results

The most stringent constraints on κ_t prior to those published in [1], are detailed in [74]. The paper combines the ATLAS analysis results of the Higgs boson in the $H \rightarrow \gamma\gamma$, ZZ^*, WW^*, $\tau\tau$, $b\bar{b}$, and $\mu\mu$ decay channels. The results of a generic fit in the kappa framework are shown in Figure 9.1, where the $gg \rightarrow H$ and $H \rightarrow \gamma\gamma$ production modes are parameterised by their effective couplings κ_g and κ_γ. In each model considered, negative values of κ_t are not excluded. This is because sensitivity to κ_t is predominantly achieved through $\sigma_{t\bar{t}H}$, which only depends on κ_t^2. This chapter targets negative values of κ_t using the tH process.

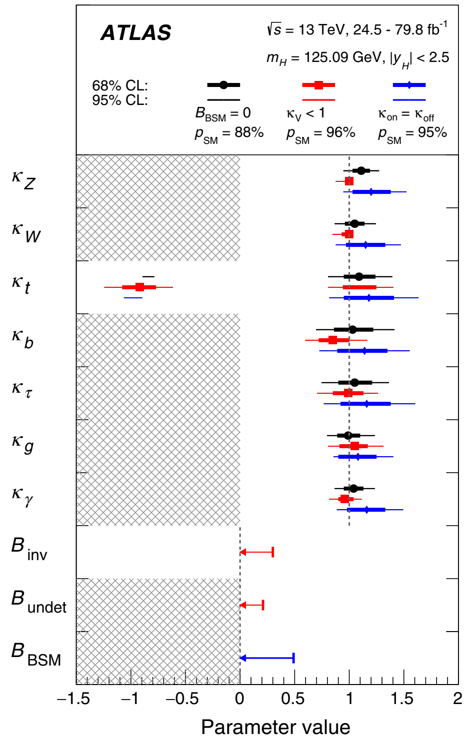


Figure 9.1 *Best-fit values and uncertainties for Higgs boson coupling modifiers per particle type using effective photon and gluon couplings κ_g and κ_γ. Three scenarios are considered: where the Higgs branching ratio to invisible (B_{inv}) and undetected (B_{undet}) decays are set to zero (black); where B_{inv} and B_{undet} are included as free parameters and κ_V is fixed to be less than or equal to 1 (red); where where B_{inv} and B_{undet} are included as free parameters and assumptions are applied to the off-shell coupling-strength scale factors as detailed in [74]. The SM corresponds to $B_{inv} = B_{undet} = 0$ and all κ parameters set to unity. All parameters except κ_t are assumed to be positive. Taken from [74].*

9.3 The tH Process

The tH process, introduced in Section 2.7.1, is extremely rare, contributing only ∼0.1% of Higgs boson production at the Large Hadron Collider (LHC). Only now, with the Full Run 2 dataset, is there enough sensitivity to design analyses targeting it.

The tH process can be split into two sub-processes, $tHqb$ process and tHW. The splitting is applied because the final states of the tHW process are hard to separate from the final states of the $t\bar{t}H$ process. The $tHqb$ sub-process is dominant, with a cross-section ∼ 3 times larger than the tHW sub-process (Table

8.2).

The dominant (t-channel) Feynman diagrams for the $tHqb$ process are shown in Figure 9.2. In the interaction, the Higgs boson can couple to either the outgoing top quark, or the exchanged W boson. It is this degeneracy that makes the $tHqb$ process sensitive to the sign of κ_t.

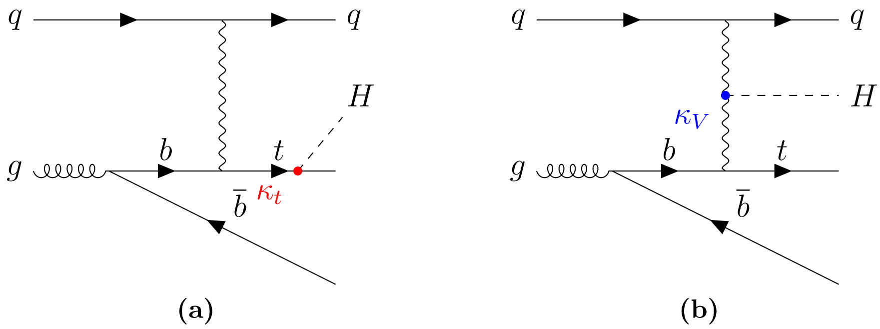


Figure 9.2 *The dominant leading order Feynman diagrams for the $tHqb$ t-channel sub-process.*

Parameterising the $tHqb$ cross-section as a function of κ_t and κ_V (as in Section 8.0.1) gives the following [9]:

$$\sigma_{tHqb}/\sigma_{tHqb}^{SM} = 2.633{\kappa_t}^2 + 3.578{\kappa_V}^2 - 5.211\kappa_t\kappa_V \tag{9.2}$$

Where σ_{tHqb} is the observed $tHqb$ cross-section, σ_{tHqb}^{SM} is the SM expected $tHqb$ cross-section, and κ_V is the vector boson (W/Z) coupling modifier.

When changing the value of κ_t, there are changes in the event topology. Differences in event topology are used to enhance sensitivity to κ_t further in the analysis, detailed in Section 9.5.

The dominant (t-channel) Feynman diagrams for the tHW process are shown in Figure 9.3. Similar to $tHqb$ the Higgs boson can couple to either the outgoing top quark, or the outgoing W boson. This makes the tHW process also sensitive to the sign of κ_t.

Similarly, parameterising the tHW cross-section in terms of κ_t and κ_V gives the following [9]:

$$\sigma_{tHW}/\sigma_{tHW}^{SM} = 2.909{\kappa_t}^2 + 2.310{\kappa_V}^2 - 4.220\kappa_t\kappa_V \tag{9.3}$$

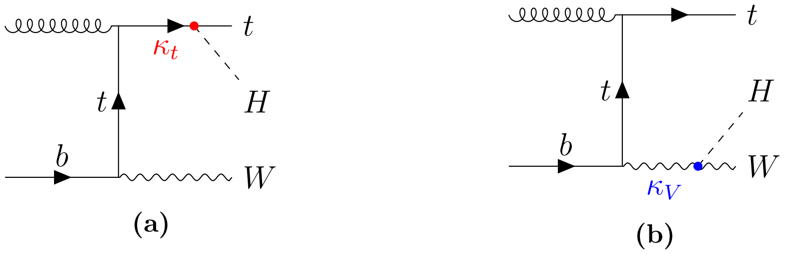


Figure 9.3 *The dominant leading order Feynman diagrams for the tHW t-channel sub-process.*

The cross-section of the $tHqb$ process, the tHW process and the $t\bar{t}H$ process as a function of κ_t is shown in Figure 9.4.

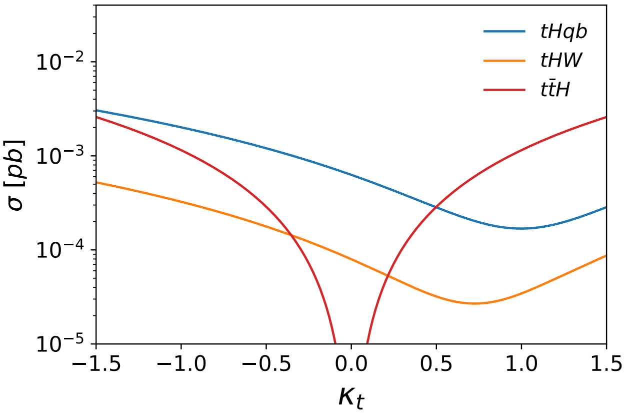


Figure 9.4 *The cross-section of the $tHqb$ (blue), tHW (orange) and $t\bar{t}H$ (red) Higgs boson production modes as a function of κ_t at the LHC.*

Because of the very low cross-section of the tHW, and the limited number of expected events that pass the analysis selection, a dedicated analysis is only performed for the $tHqb$ sub-process.

9.4 Object Reconstruction

Extra physics objects, in addition to those detailed in 8.2, are reconstructed. They are used to define variables that are sensitive to the $tHqb$ process (Section 9.5). Specifically, objects are defined that represent the W boson and b-quark decays of any top quarks in the event, and the most forward jet in the event j_F. The reconstruction procedure for each object is detailed in the following:

b quark and W boson objects are reconstructed for each reconstructed top in the event:

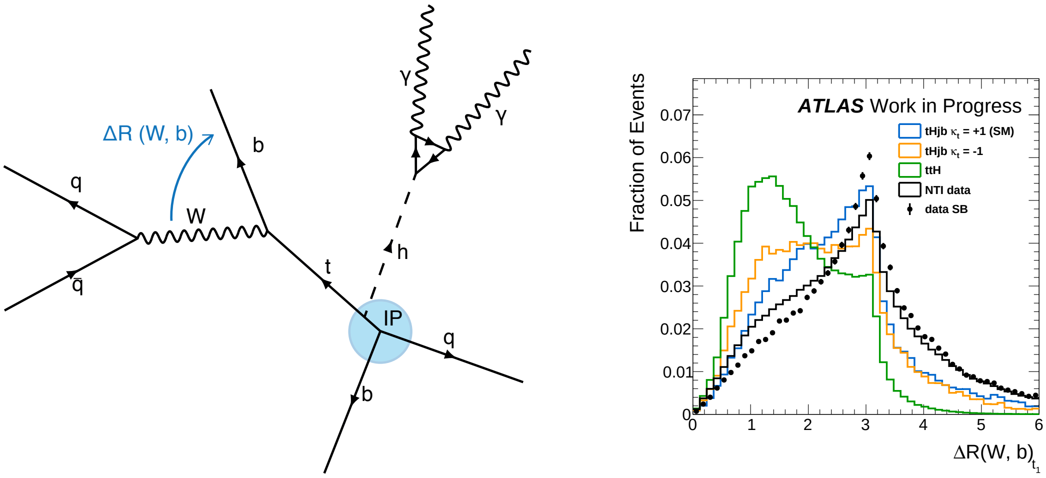
- If the first top (t_1) is reconstructed hadronically the constituent jet with the highest pseudo-continuous b-tagging (PCBtag) score is selected as the associated b quark object. If two jets have the same PCBtag score, the jet with the highest p_{T} is selected. The remaining two jets are reconstructed as the associated W boson object.
- If t_1 is reconstructed leptonically, the single constituent jet is selected as the associated b quark object. The associated W boson object is selected as the remaining four momentum of the top after subtracting the b quark object four momentum.
- If a second top quark (t_2) is reconstructed from 3 jets, the jet with the highest PCBtag score and p_{T} is selected as the associated b quark object. The remaining two jets are selected as the associated W boson object.
- If t_2 is reconstructed from 2 jets, the associated W boson object is selected as the jet with the highest p_{T}. The associated b quark object is selected as the remaining jet.

The most forward jet (j_F) object is selected as the jet with the largest absolute pseudorapidity $|\eta|$. Reconstructing this object targets the outgoing quark in $tHqb$ (Figure 9.2). Where because the dominant $tHqb$ process is a t-channel scattering process, the outgoing quark is expected to emerge in the forward direction.

9.5 $tHqb$ κ_t Sensitive Event Variables

Forty custom variables are defined to target differences in $tHqb$ events with $\kappa_t = +1$ and $\kappa_t = -1$. They are designed by considering the topology of $tHqb$ events for different κ_t values. Figure 9.5 (a) shows a schematic of the typical topology of a $tHqb$ event in the ϕ plane of the detector.

The thought process behind the $\Delta R(W, b)_{recotop1}$ variable is given as an example. Under the different κ_t hypotheses, the Higgs boson coupling to the top quark will result in a different amount of momentum transfer. This will result in a change in the opening angle between the decay products of the top quark. This can be seen in Figure 9.5 (b).



(a) **(b)**

Figure 9.5 *(a) Schematic showing the event topology of a typical* $tHqb$ *event in the* ϕ *plane. The opening angle* $\Delta R(W, b)$ *between the* W *and* b *decays of the top quark is indicated. IP signifies the interaction point. (b) Histogram of the* $\Delta R(W, b)_{t_1}$ *variable for the* $tHqb$ $\kappa_t = +1$ *vs* $\kappa_t = -1$ *model. Showing* $tHqb$ $\kappa_t = +1$ *(SM)(blue-line),* $tHqb$ $\kappa_t = -1$*(orange-line),* $t\bar{t}H$ *(green-line), Non-Tight-Isolated (NTI) data (black-line) and the side-bands of the Tight-Isolated (TI) data (black-dots).*

The top six most important variables in the $tHqb$ $\kappa_t = +1$ vs $\kappa_t = -1$ model are shown in Figure 9.6. Some variables depend on the reconstruction of physics objects, if these objects are not available in the event the variable is set to -99, and is not plotted.

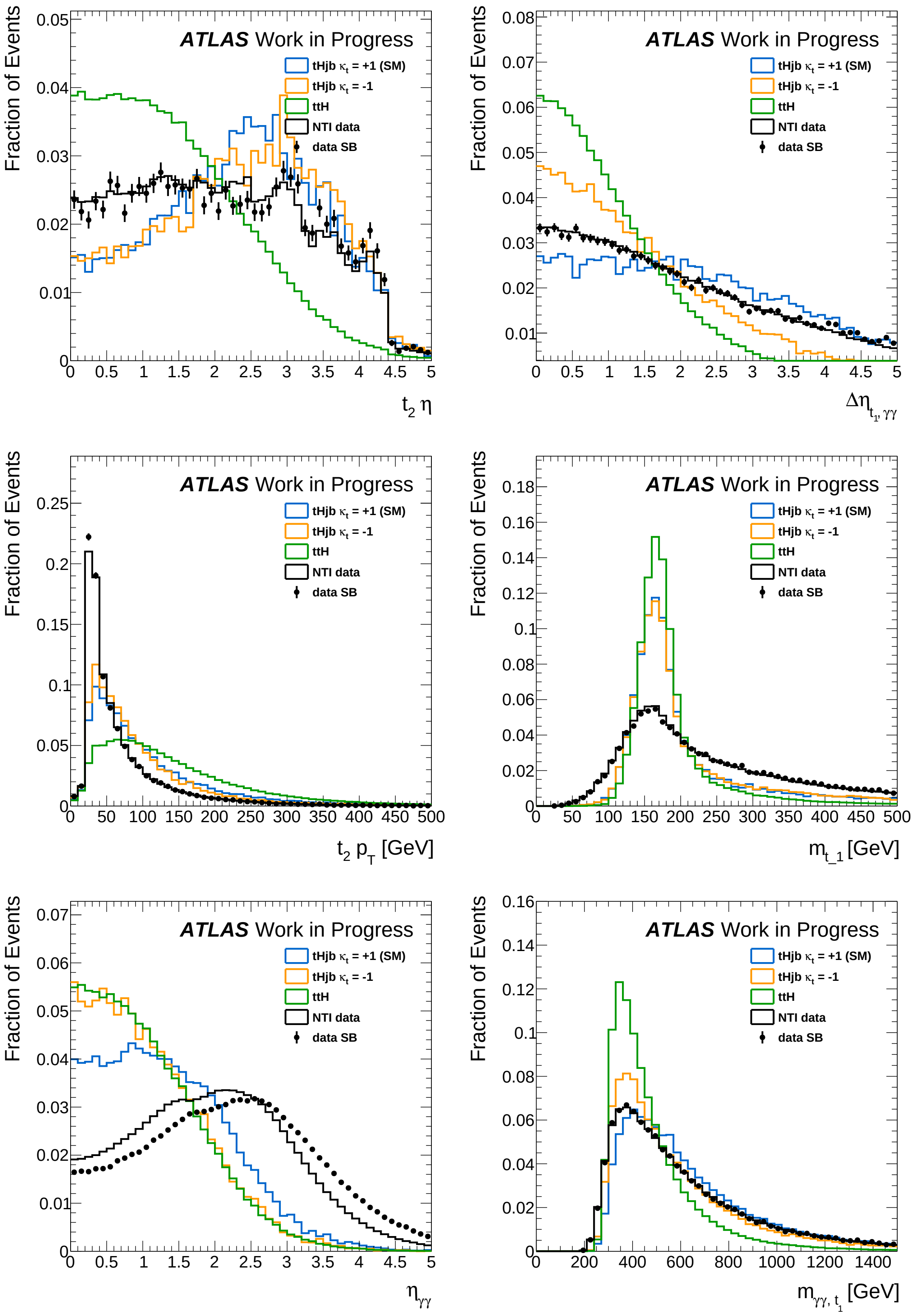


Figure 9.6 *The six most important variables for the tHqb* $\kappa_t = +1$ *vs* $\kappa_t = -1$ *model. Showing tHqb* $\kappa_t = +1$ *(SM)(blue-line), tHqb* $\kappa_t = -1$*(orange-line),* $t\bar{t}H$*(green-line), NTI data (black-line) and the side-bands of the TI data (black-dots with statistical error bars).*

9.6 Neural Network Classifiers

Three binary classifier neural network models are trained, these are:

- the $tHqb$ $\kappa_t = +1$ vs $\kappa_t = -1$ model, that classifies the two $tHqb$ signals simulated with different κ_t values,
- the $tHqb$ $\kappa_t = +1$ vs NTI data model, that separates the SM $tHqb$ process from the non-resonant background,
- and the $tHqb$ $\kappa_t = -1$ vs NTI data model, that separates the $tHqb$ $\kappa_t = -1$ process from the non-resonant background.

9.6.1 Neural Network Setup

The simulated $tHqb$ $\kappa_t = +1$ and $\kappa_t = -1$ samples are used to model each signal. The NTI data that is categorised into the *top-like* categories by the multiclass BDT (Section 8.3.2) is used to model the non-resonant background. Samples are split into exclusive *train* (50%), *test* (25%), and *validation* (25%) sets. The neural network models are developed using the Keras [50] package. In the loss function, each class is weighted to normalise its importance. The *ReduceLearningRateOnPlateau* and *EarlyStopping* callbacks are used to avoid overfitting, and to ensure the optimal minimum is reached. The output node uses the sigmoid activation function. The Binary Cross-Entropy (BCE) loss is used, with the Adam optimiser and a learning rate of $lr = 0.01$

The hyperparameters in Table 9.1 were selected by running a Bayesian optimisation procedure for each classifier with the objective of maximising the Area Under Curve (AUC) metric. The scikit-optimize package [75] was used to conduct the procedure.

Hyperparameter	Selected Value
number of hidden layers	4
hidden layer nodes	(200, 200, 200, 24)
hidden activation	LeakyReLU(*alpha*=0.3)
dropout rate	0.05

Table 9.1 *The selected hyperparameters of each neural network classifier.*

9.6.2 Variable Selection Procedure

For each model, all 40 input variables are ranked using a Boosted Decision Tree (BDT) developed with the XGBoost [69] package. The BDT is trained on the same classification objective as the neural network model in consideration. The neural network model is then trained recursively using the list of input variables, each time increasing the number of variables in order of their ranking. For each iteration, the neural network is trained 5 times so that statistical fluctuations in the training procedure can be accounted for. The performance of each neural network training is monitored using the AUC metric. The minimum number of training variables is selected after which no significant improvement in the AUC is observed.

The full set of selected variables used in each model can be found in Table 8.3 in Section 8.

9.6.3 Variable Selection Results

The results of the variable selection procedure are presented for each neural network model. This includes the variable ranking obtained from the associated BDT model, and the recursive feature selection results. For each model, the top 20 variables are selected.

The results of the variable selection procedure for the $tHqb$ $\kappa_t = +1$ vs $\kappa_t = -1$ model are shown in Figure 9.7.

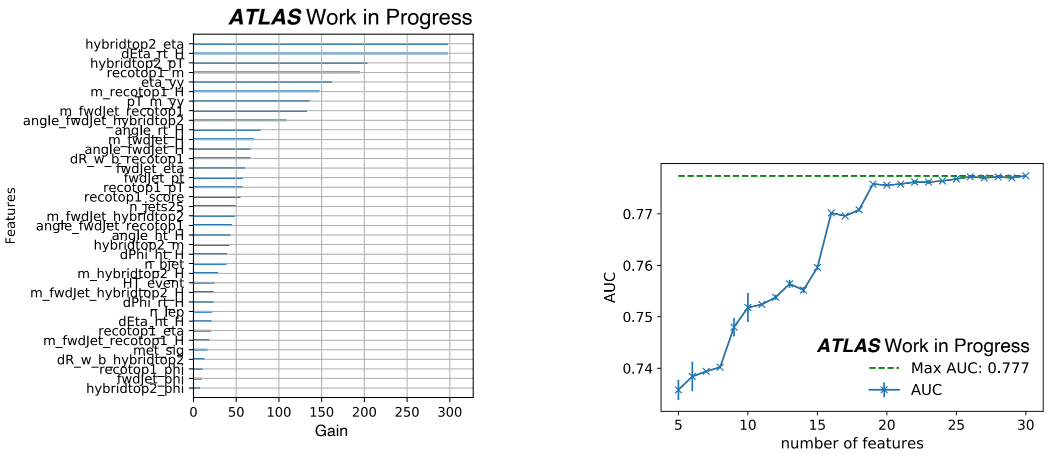


Figure 9.7 *Variable ranking and recursive feature selection results in the $tHqb$ $\kappa_t = +1$ vs $\kappa_t = -1$ model.*

The results of the variable selection procedure for the $tHqb$ $\kappa_t = +1$ vs NTI data model are shown in Figure 9.8. Only variables that have a correlation with $m_{\gamma\gamma}$ of $< 5\%$ are considered, avoiding any mass sculpting effects.

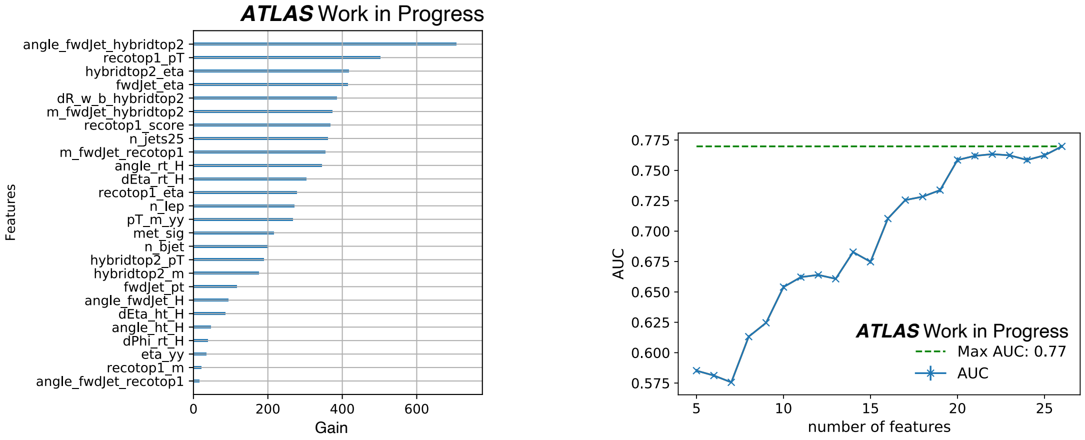


Figure 9.8 *Variable ranking and recursive feature selection results for the* $tHqb\kappa_t = +1$*vs NTI data model.*

The results of the variable selection procedure for the $tHqb$ $\kappa_t = -1$ vs NTI data model are shown in Figure 9.9. Only variables that have a correlation with $m_{\gamma\gamma}$ of $< 5\%$ are considered, avoiding any mass sculpting effects.

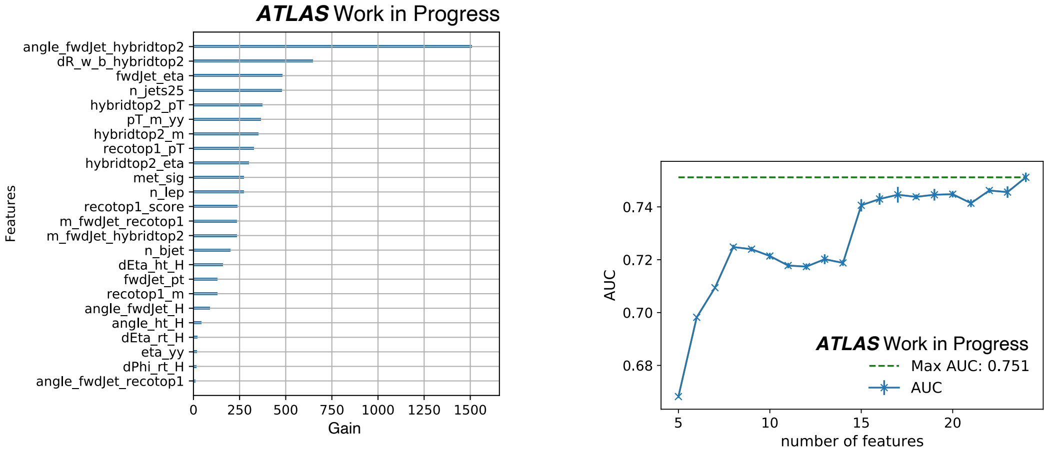


Figure 9.9 *Variable ranking and recursive feature selection results for the* $tHqb\kappa_t = -1$*vs NTI data model.*

9.6.4 Model Performance

The performance of each model is evaluated on the subset of events that enter the $tHqb$ multiclass category. The performance of the $tHqb$ $\kappa_t = +1$ vs $\kappa_t = -1$

model, the $tHqb\ \kappa_t = +1$ vs NTI data model, $tHqb\ \kappa_t = -1$ vs NTI data model and the are shown in Figures 9.10, 9.11, and 9.12 respectively.

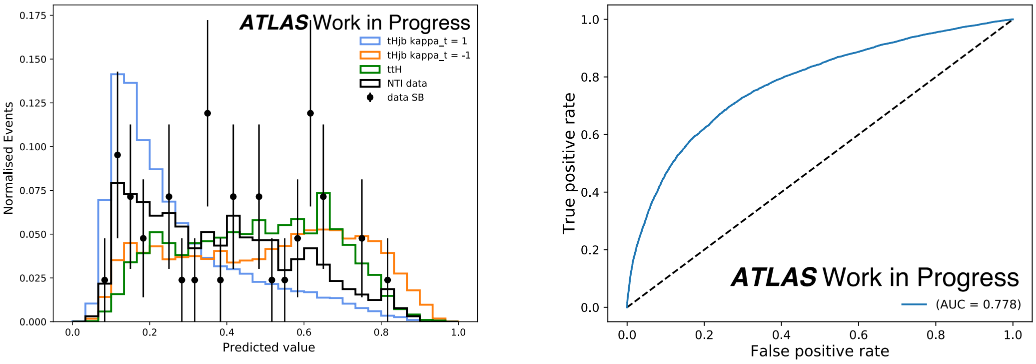


Figure 9.10 *Performance of the* $tHqb\ \kappa_t = -1$ *vs* $\kappa_t = -1$ *model in terms of class separation and ROC.*

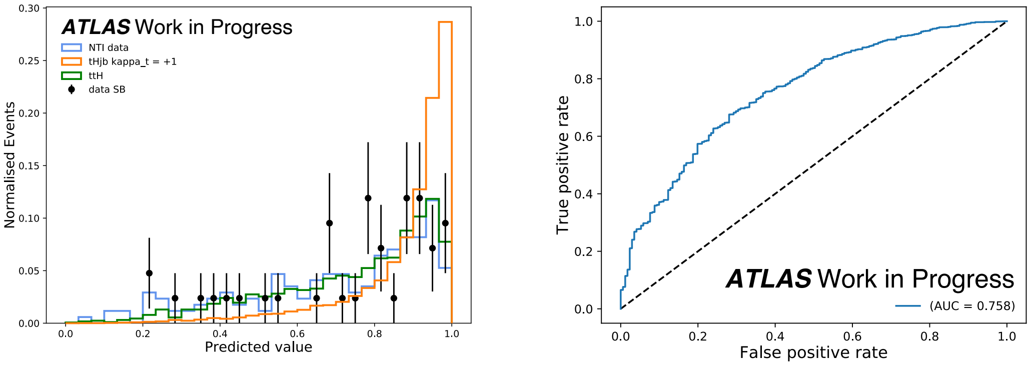


Figure 9.11 *Performance of the* $tHqb\ \kappa_t = +1$*vs NTI data model in terms of class separation and ROC.*

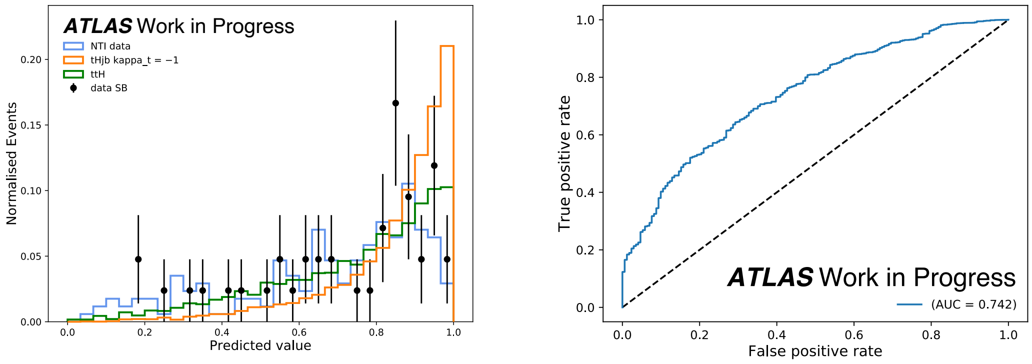


Figure 9.12 *Performance of the* $tHqb$ $\kappa_t = -1$*vs NTI data model in terms of class separation and ROC.*

9.6.5 Mass Sculpting

Each model was tested for mass sculpting of the non-resonant background. Any sculpting can affect the background fit in the statistical model of the analysis. The $t\bar{t}\gamma\gamma$ simulated sample was used to test for sculpting. Three different cut values on each neural network discriminant were considered. No significant mass sculpting was observed.

9.7 Event Categorisation Optimisation

To select the optimal event categorisation of the $tHqb$ process, a sensitivity investigation that accounts for statistical uncertainty is conducted. The sensitivity of each categorisation option is evaluated under the two objectives: maximising sensitivity to the SM $tHqb$ process, and maximising sensitivity to the $\kappa_t = -1$ hypothesis. The most sensitive event categorisation is then selected. The statistical stability of each categorisation is also considered. The following section describes this procedure in more detail, including how the statistics-only counting experiment is conducted, how different categorisations are constructed, and the results of the procedure. The exclusive *validation* dataset is used in this process to avoid potential biases.

9.7.1 Stat-Only Counting Experiment

A binned statistics-only fit is conducted using the following steps. The results of the fit are used to assess the sensitivity of the categorisation.

1. Signal Monte-Carlo (MC) events are categorised, and then normalised using the normalisation factors in Table 8.2.
2. A *double-sided* Crystal Ball (DSCB) is fitted to the inclusive signal MC in the $m_{\gamma\gamma}$ spectrum in each category.
3. A window is defined in the $m_{\gamma\gamma}$ spectrum that contains 90% of the signal yield; the yield in this window is labeled S90.
4. A background sample, either NTI data or MC, is categorised.
5. In each category a probability density function (pdf) is fit to the background sample. An exponential of a second-order polynomial is used if the yield is greater than 400 events, or an exponential otherwise.
6. The fitted pdf is scaled to the TI data side-band yield and by a factor of 55/45 to account for the relative size of the Higgs signal region. The factor is obtained by taking the ratio of the total number of bins and the number of bins in the signal region in the $m_{\gamma\gamma}$ spectrum.
7. The background yield B90 is then defined as the intergral of the background pdf in the S90 $m_{\gamma\gamma}$ window.

An Asimov fit is conducted, using the calculated S90 and B90 yields. When conducting a measurement of a single process, other signal yields and the B90 yield are considered as background.

In the fit, measurements of the cross-section of the $t\bar{t}H$ process and the tH processes are conducted. The sensitivity to the $\kappa_t = -1$ hypothesis is also evaluated. This is done by parameterising the yields of the $tHqb$ and tHW processes as a function of κ_t in each of the analysis categories, following the same procedure as detailed in Section 9.8. The expected sensitivity to the $\kappa_t = -1$ hypothesis is evaluated as the likelihood ratio between a fit conducted under the $\kappa_t = -1$ hypothesis, to a fit conducted under the SM hypothesis.

9.7.2 Categorisation Construction

The following procedure is used to build event categorisation options.

All events that are categorised to be $tHqb$-like by the multiclass BDT are considered. First, two categories are built by scanning through the $tHqb$ κ_t neural network discriminant, creating a $tHqb$ $\kappa_t = +1$ (SM) like category and a $tHqb$ $\kappa_t = -1$ like category for each cut choice. Cuts that result in less than 9 TI data side-band events in either category are rejected.

Next, to maximise sensitivity to the two objectives, the non-resonant background in each of these categories is rejected using the two background rejection neural network models developed in Section 9.6.

In this background rejection procedure, the non-resonant background is modelled using the $t\bar{t}\gamma\gamma$ MC sample, as it was determined to best model the TI data side-band in each NN discriminant. Prior to rejecting background, the $t\bar{t}\gamma\gamma$ yield is scaled to the yield of the TI data side-band, and by a factor of 55/45 to account for the relative size of the Higgs signal region. The factor is obtained by taking the ratio of the total number of bins and the number of bins in the signal region in the $m_{\gamma\gamma}$ spectrum.

The non-resonant background is then rejected in each of the categories by scanning through the NN discriminant and selecting the cut that maximises the S90/B90 significance S. The following significance calculation is used:

$$S = \sqrt{2 \cdot \left((s+b) \cdot \log\left(1 + \frac{s}{b}\right) - s\right)} \tag{9.4}$$

Where s is the number of signal events in the category, and b is the number of background events in the category.

Categories are required to have at least 9 events in the side-band integral of the background fit pdf. In each category the S90 window is assumed to be constant, saving computation time. If there are enough MC side-band events, the procedure is repeated, creating two sensitive categories to the corresponding objective. The sensitivity of each categorisation option is then evaluated using the stat-only counting experiment procedure detailed in Section 9.7.1.

Figure 9.13 illustrates how an example categorisation is constructed using the three neural network discriminants.

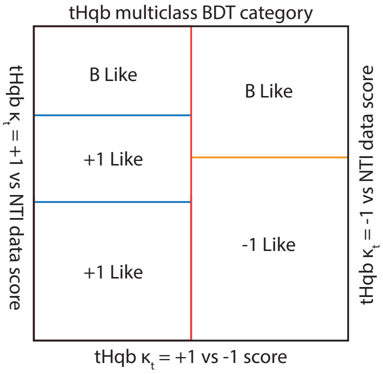


Figure 9.13 *Schematic showing how categorisations are constucted using each neural network discriminant. "+1 Like" indicates a category that is sensitive to the tHqb* $\kappa_t = +1$ *(SM) process, "-1 Like" indicates a category that is sensitive to the tHqb* $\kappa_t = -1$ *process, and "b Like" indicates a region that is determined to be background like.*

9.7.3 Results

The results of the categorisation optimisation procedure are in Figure 9.14. The selected categorisation is chosen on a compromise between the sensitivity of the two objectives, and statistical stability.

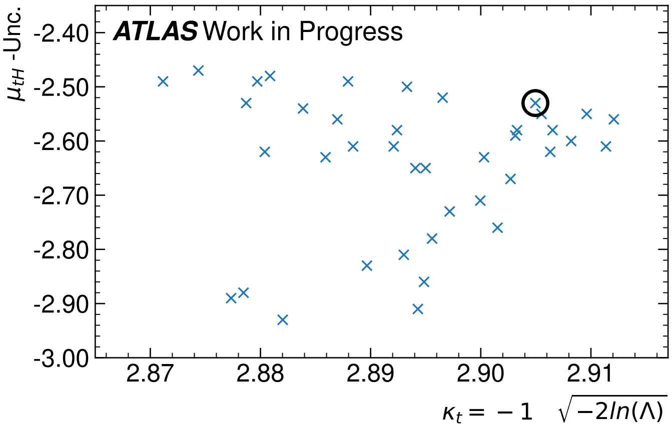


Figure 9.14 *Results of the tHqb categorisation investigation in terms of the negative uncertainty on the tH cross-section and the constraining power on the* $\kappa_t = -1$ *hypothesis. Categorisations to the upper right show a better performance. Categorisation options are indicated with a blue cross. The circled cross shows the selected categorisation.*

The chosen categorisation is defined by two categories that are sensitive to the $tHqb$ $\kappa_t = +1$ (SM) process, and one category that is sensitive to the $\kappa_t = -1$ hypothesis. Events that do not pass either of these three selections are deemed to be background-like, and are merged with other similar categories produced by the $t\bar{t}H$ and tHW categorisations as detailed in Section 8.3.

9.8 Signal Yield Parameterisation

As demonstrated in Section 9.5, the event shapes of the $tHqb$ and tHW processes have a strong dependency on κ_t and κ_V. This affects the expected tH yields (y_i) in each analysis category in addition to the cross-sectional dependencies detailed in Section 8.3:

$$y_i(\kappa_t, \kappa_V) = L \cdot \sigma_{tH}(\kappa_t, \kappa_V) \cdot \mathcal{A}(\kappa_t, \kappa_V) \cdot \epsilon_i(\kappa_t, \kappa_V), \tag{9.5}$$

where i denotes the target category, $\mathcal{A}$ is the acceptance of the ATLAS, detector and ϵ_i the efficiency for the tH events to pass the requirements of category i, and L is the luminosity of the LHC.

To capture the efficiency effects, the expected event yield $y_i(\kappa_t, \kappa_V)$ is parameterised in each of the analysis categories. Importantly, the parameterisation captures the dependencies $\mathcal{A}(\kappa_t, \kappa_V)$ and $\epsilon_i(\kappa_t, \kappa_V)$, which is an improvement over previous methods, which only account for the dependence in $\sigma_{tH}(\kappa_t, \kappa_V)$.

Each parameterisation is calculated via the analytical solution to the following equation,

$$y_i(\kappa_t, \kappa_V) = {\kappa_t}^2 A_i + \kappa_V^2 B_i + \kappa_t \kappa_V C_i \tag{9.6}$$

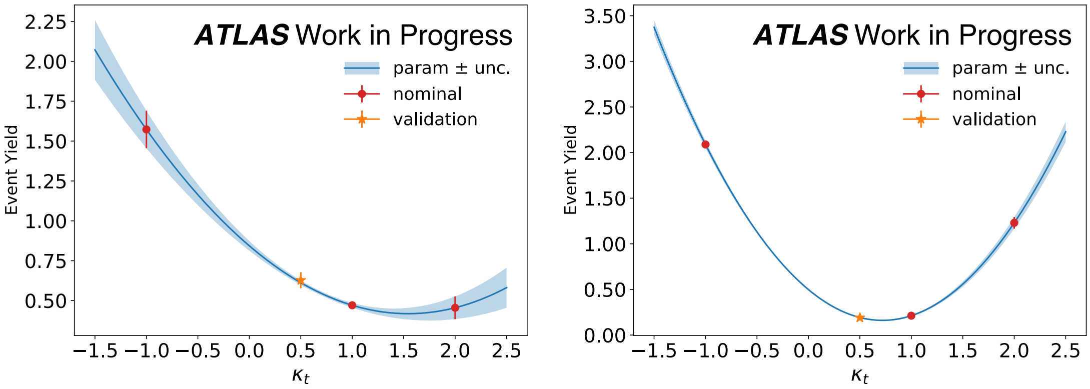
where y_i represents the predicted yield in category i, and A_i, B_i, and C_i are three unknown factors. Three $tHqb$ MC samples simulated with different κ_t values ($\kappa_t = -1$, $\kappa_t = 1$ and $\kappa_t = 2$) are used. The expected yields of each sample are used to solve for A_i, B_i, and C_i in Equation 9.6. A bootstrapping technique is used to estimate the statistical uncertainty in each yield, with 100 draws. The uncertainty in each parameterisation is then inferred from the yield uncertainties.

A closure test is conducted by comparing the yield predicted by the parameterisation to the yield of a $tHqb$ MC validation sample simulated with $\kappa_t = 0.5$. Any deviations are explained by statistical fluctuation.

The yield parameterisations of the tH and tHW processes in the tH categories can be found in Table 9.2. The full set of parameterisations in all $H \rightarrow \gamma\gamma$ analysis categories are provided in Appendix A. Two example parameterisations in graphical form can be found in Figure 9.15.

Process	Category	A	B	C
$tHqb$	$tHqb$, BSM ($\kappa_t = -1$)	8.220	8.760	-15.980
	$tHqb$, High-purity	0.380	1.790	-1.170
	$tHqb$, Med-purity	1.100	2.600	-2.700
	tHW	6.100	6.470	-11.560
tHW	$tHqb$, BSM ($\kappa_t = -1$)	2.510	2.170	-3.690
	$tHqb$, High-purity	3.820	2.460	-5.280
	$tHqb$, Med-purity	2.770	3.020	-4.790
	tHW	3.080	2.350	-4.430

Table 9.2 *Yield parametrisations as a ratio to the SM expected yield for the $tHqb$ and tHW processes in the tH analysis categories.*



(a) *Yield parameterisation for the $tHqb$ process in the $tHqb$ High-purity analysis category.*

(b) *Yield parameterisation for the tHW process in the tHW analysis category.*

Figure 9.15 *Yield parameterisations in the $tHqb$ and tHW high-purity analysis categories. Red points indicate the yield and κ_t values that were used to analytically solve the parameterisation. The blue solid line indicates the nominal parameterisation, and the light blue band indicates the statistical uncertainty of the parameterisation. The yellow point indicates the validation.*

Some yield parameterisations have large statistical uncertainties. The effect of these uncertainties on the κ_t exclusion limits was tested. Effects were found to be negligible when compared to the statistical uncertainty on the SM tH cross-section [9].

9.9 Summary

For this thesis, a novel search for tH production has been delivered. The search is performed in the framework of an analysis targeting all production modes which can be measured with high sensitivity in $H \to \gamma\gamma$ decays, detailed in Chapter 8. The main new developments presented in this chapter are: event categorisation of the $tHqb$ process was using three neural network models, and a parameterisation to capture the efficiency dependency of the $tHqb$ and tHW processes on κ_t and κ_V.

The measured tH production signal strength is:

$$\mu_{tH} = 2.6^{+4.2}_{-3.2} = 2.6^{+3.7}_{-2.9}(stat.)^{+1.6}_{-1.1}(syst.). \tag{9.7}$$

This pioneering work achieves the most stringent constraint on the tH SM cross-section, and the $\kappa_t = -1$ hypothesis to date. It also establishes the ability of ATLAS to probe tH production; Figure 9.16 shows the unique contribution to the ATLAS combination of all Higgs decay modes [76]. It highlights the $H \to \gamma\gamma$ channel as the only contributor to the tH sensitivity. This work has thus contributed to the experimental understanding of the Higgs boson.

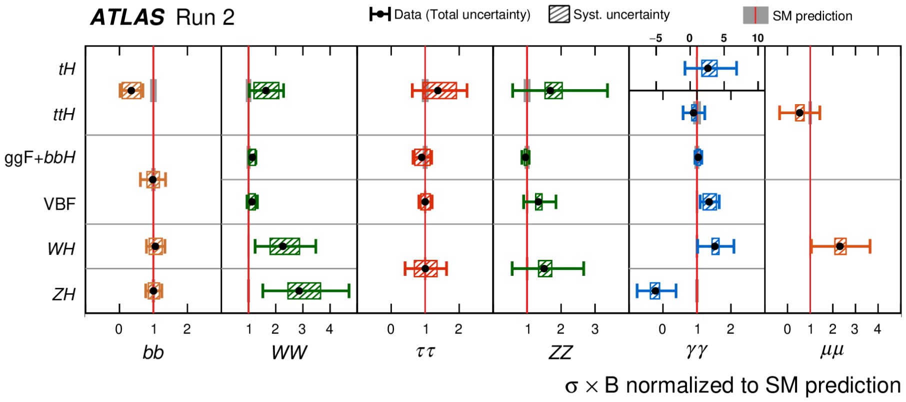


Figure 9.16 *Ratio of observed rate to predicted SM event rate for different combinations of Higgs boson production and decay processes. The horizontal bar on each point denotes the 68% confidence interval. The narrow grey bands indicate the theory uncertainties in the SM cross-section times the branching fraction predictions. The p-value for compatibility of the measurement and the SM prediction is 72%. Taken from [76].*

Part IV

Conclusion & Outlook

The discovery of the Higgs boson by the ATLAS and CMS collaborations in 2012 completed the Standard Model of particle physics, but it marked only the start of a rich physics program at the Large Hadron Collider (LHC). Since the discovery, the focus has moved toward precision measurements of the Higgs boson and its interactions.

This thesis presented the latest precision analysis of the Higgs boson properties using the di-photon decay channel ($H \to \gamma\gamma$). With the LHC Run 2 data of 139 fb^{-1}, the channel presented a newfound sensitivity to the rare single-top Higgs (tH) production process. The tH process also is sensitive to the sign of the top quark Yukawa coupling strength κ_t, and presented a unique opportunity to constrain it. Previous constraints on κ_t from the the ATLAS analysis of the combined measurements of Higgs boson production and decay using up to 80 fb^{-1}, left the negative values un-excluded in the region $-1.2 \leq \kappa_t \leq -0.6$ at the 95% CL level.

A novel analysis in the $H \to \gamma\gamma$ decay channel was developed for the $tHqb$ process to capitalise on this opportunity. The analysis used three neural network models. A set of original variables were designed and used to train the models, bringing enhanced sensitivity to the tH process and the κ_t hypotheses. Using the models, the analysis event categorisation was simultaneously optimised for sensitivity to the Standard Model (SM) tH cross-section measurement, and the sign of κ_t. A parameterisation to capture the efficiency dependency of the $tHqb$ and tHW processes on κ_t and κ_V was also developed, improving the sensitivity to κ_t. The ratio of the measured and predicted tH cross-sections was reported as:

$$\sigma_{tH}/\sigma_{tH}^{\mathrm{SM}} = 2.6^{+4.2}_{-3.2} = 2.6^{+3.7}_{-2.9}(\mathrm{stat.})^{+1.6}_{-1.1}(\mathrm{syst.}). \tag{9.8}$$

An upper limit on the tH cross-section was set at 10 times the SM expectation at the 95% CL, and negative values of the top quark Yukawa coupling were excluded at 2.2σ.

The results show significantly higher sensitivity to the tH process and κ_t than those reported by the CMS experiment. The CMS analysis of the Higgs boson properties in final states with electrons, muons, and hadronically decaying tau leptons using 139 fb^{-1} of data, reported the ratio of the measured and predicted tH cross-sections as:

$$\sigma_{tH}/\sigma_{tH}^{\mathrm{SM}} = 5.7 \pm 4.0 = 5.7 \pm 2.7(\mathrm{stat.}) \pm 3.0(\mathrm{syst.}). \tag{9.9}$$

Negative values of the top quark Yukawa coupling were not excluded at the 95% CL. The CMS analysis of the Higgs boson properties using $H \to \gamma\gamma$ decay channel using 139 fb^{-1} of data, reported the ratio of the measured and predicted tH cross-sections as:

$$\sigma_{tH}/\sigma_{tH}^{\mathrm{SM}} = 6.3^{+3.7}_{-3.4} = 6.3^{+3.8}_{-3.4}(\mathrm{stat.})^{+0.5}_{-0.7}(\mathrm{syst.}). \tag{9.10}$$

An upper limit on the tH cross-section was set at 14 times the SM expectation at the 95% CL.

The cutting-edge work in this thesis, therefore, has set the most stringent constraints on the tH cross-section and the $\kappa_t = -1$ hypothesis to date. It has also established the sensitivity of the ATLAS experiment to the tH production process. The work furthers the understanding of the SM and the Higgs boson, and builds the foundations for future analyses of the tH process.

In order to conduct even more precise measurements of the Higgs boson and its processes in the future, very large sets of data will be required. Even larger amounts of simulated data will be needed to design the associated analyses. This presents an issue for the ATLAS experiment's future, in which it is required to stay within a strict computing budget. Currently, a significant fraction of computing time is spent simulating processes in the ATLAS detector. The AtlFastIII (AF3) simulation toolkit solves the issue, as a fast and accurate alternative to traditional slower Monte-Carlo simulation techniques.

This thesis presented a significant contribution to AF3; the development of the muon punch-through component. The component uses a parameterised approach to accurately model the muon punch-through effect in the ATLAS detector. Accurate modelling of this effect is essential for understanding muon fakes and for hadronic jet calibration. The component uses a collection of techniques to achieve the required accuracy, including a novel calibrated neural network model. These contributions will play a significant role in the future simulation program of the ATLAS experiment, enabling more precise measurements of the Higgs boson.

Part V

Appendices

Appendix A

Yield Parameterisations

Category	*A*	*B*	*C*
GG2H_GE2J_MJJ_0_350_PTH_0_60__0	2.240	3.850	-5.090
GG2H_0J_PTH_0_10__0	3.520	1.990	-4.520
GG2H_0J_PTH_GT10__0	1.560	4.120	-4.680
GG2H_1J_PTH_0_60__0	2.540	6.490	-8.030
GG2H_1J_PTH_0_60__1	2.170	3.610	-4.780
GG2H_1J_PTH_120_200__1	2.710	4.510	-6.220
GG2H_1J_PTH_60_120__0	5.120	7.710	-11.830
GG2H_1J_PTH_60_120__1	1.810	2.610	-3.410
GG2H_GE2J_MJJ_0_350_PTH_0_60__1	3.170	3.520	-5.690
GG2H_GE2J_MJJ_0_350_PTH_0_60__2	2.680	3.370	-5.050
GG2H_GE2J_MJJ_0_350_PTH_120_200__0	5.890	7.460	-12.350
GG2H_GE2J_MJJ_0_350_PTH_120_200__1	5.110	6.590	-10.700
GG2H_GE2J_MJJ_0_350_PTH_60_120__0	2.850	4.000	-5.850
GG2H_GE2J_MJJ_0_350_PTH_60_120__1	3.170	4.070	-6.250
GG2H_GE2J_MJJ_350_700_PTH_0_200__0	6.600	2.970	-8.570
GG2H_GE2J_MJJ_350_700_PTH_0_200__1	2.670	3.690	-5.360
GG2H_GE2J_MJJ_350_700_PTH_0_200__2	3.570	3.660	-6.230
GG2H_GE2J_MJJ_700_1000_PTH_0_200__0	1.150	2.990	-3.140
GG2H_GE2J_MJJ_700_1000_PTH_0_200__1	1.550	2.700	-3.250
GG2H_GE2J_MJJ_700_1000_PTH_0_200__2	1.810	2.530	-3.340
GG2H_GE2J_MJJ_GT1000_PTH_0_200__0	1.640	3.060	-3.710
GG2H_GE2J_MJJ_GT1000_PTH_0_200__1	0.620	2.130	-1.750
GG2H_GE2J_MJJ_GT1000_PTH_0_200__2	1.010	2.540	-2.560
GG2H_PTH_200_300__0	5.240	5.470	-9.710
GG2H_PTH_200_300__1	5.080	6.010	-10.090
GG2H_PTH_300_450__1	7.840	7.020	-13.860
GG2H_PTH_300_450__2	6.190	5.700	-10.890
GG2H_PTH_450_650__0	5.450	10.020	-14.470
GG2H_PTH_450_650__1	5.900	9.060	-13.970
GG2H_PTH_450_650__2	6.890	6.210	-12.100
GG2H_PTH_GT650__0	16.240	14.570	-29.810
QQ2HLL_PTV_0_75__1	2.390	3.910	-5.300
QQ2HLL_PTV_150_250__0	6.550	7.330	-12.880
QQ2HLL_PTV_150_250__1	4.660	5.660	-9.310
QQ2HLL_PTV_75_150__0	3.790	5.400	-8.190
QQ2HLL_PTV_GT250__0	8.510	7.250	-14.760
QQ2HLNU_PTV_0_75__0	5.050	4.210	-8.260
QQ2HLNU_PTV_0_75__1	3.210	2.790	-5.000
QQ2HLNU_PTV_150_250__0	9.350	13.960	-22.310
QQ2HLNU_PTV_150_250__1	4.490	1.640	-5.130
QQ2HLNU_PTV_75_150__0	5.330	6.220	-10.560
QQ2HLNU_PTV_75_150__1	3.380	4.010	-6.390
QQ2HLNU_PTV_GT250__0	14.030	13.880	-26.920
QQ2HNUNU_PTV_0_75__0	2.060	0.930	-1.990
QQ2HNUNU_PTV_0_75__1	2.130	4.230	-5.360
QQ2HNUNU_PTV_0_75__2	2.160	2.210	-3.360
QQ2HNUNU_PTV_150_250__0	10.320	10.180	-19.500
QQ2HNUNU_PTV_150_250__1	22.940	20.700	-42.650
QQ2HNUNU_PTV_150_250__2	4.960	4.450	-8.410
QQ2HNUNU_PTV_75_150__1	4.320	7.650	-10.970
QQ2HNUNU_PTV_75_150__2	4.060	5.200	-8.260
QQ2HNUNU_PTV_GT250__0	11.250	3.050	-13.300
QQ2HNUNU_PTV_GT250__1	9.480	3.250	-11.730
QQ2HQQ_0J__1	2.650	6.080	-7.730
QQ2HQQ_0J__2	2.870	3.130	-5.000
QQ2HQQ_1J__2	1.770	2.670	-3.440
QQ2HQQ_GE2J_MJJ_0_60__0	1.950	5.900	-6.840
QQ2HQQ_GE2J_MJJ_0_60__2	7.630	9.450	-16.080
QQ2HQQ_GE2J_MJJ_120_350__0	1.360	4.330	-4.690
QQ2HQQ_GE2J_MJJ_120_350__1	3.440	3.780	-6.220
QQ2HQQ_GE2J_MJJ_120_350__2	4.460	4.980	-8.440
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200__0	18.620	15.290	-32.910
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200__2	0.750	3.230	-2.980
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200__0	4.080	7.910	-10.990
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200__1	2.550	3.190	-4.740
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200__2	3.900	3.240	-6.140
QQ2HQQ_GE2J_MJJ_60_120__0	16.420	12.730	-28.150
QQ2HQQ_GE2J_MJJ_60_120__1	36.870	32.920	-68.790
QQ2HQQ_GE2J_MJJ_60_120__2	8.450	7.970	-15.410
QQ2HQQ_GE2J_MJJ_700_1000_PTH_0_200__1	0.350	0.230	0.420
QQ2HQQ_GE2J_MJJ_700_1000_PTH_GT200__0	1.140	2.630	-2.770
QQ2HQQ_GE2J_MJJ_700_1000_PTH_GT200__1	2.610	3.800	-5.400
QQ2HQQ_GE2J_MJJ_GT1000_PTH_GT200__0	0.630	2.100	-1.730
QQ2HQQ_GE2J_MJJ_GT1000_PTH_GT200__1	0.190	2.100	-1.290
THJB_MINUS1__0	8.220	8.760	-15.980
THJB__0	0.380	1.790	-1.170
THJB__1	1.100	2.600	-2.700
THW__0	6.100	6.470	-11.560
TTH_PTH_0_60__0	5.290	4.280	-8.580
TTH_PTH_0_60__1	6.150	5.560	-10.710
TTH_PTH_120_200__0	4.100	2.560	-5.660
TTH_PTH_120_200__1	4.620	3.140	-6.760
TTH_PTH_200_300__0	5.410	5.850	-10.260
TTH_PTH_60_120__0	4.390	4.230	-7.620
TTH_PTH_60_120__1	4.130	3.710	-6.840
TTH_PTH_GT300__0	6.520	5.840	-11.360
BKG_LIKE_TOP	4.260	4.130	-7.390

Table A.1 *Yield parametrisations as a ratio to the SM expected yield for the* $tHqb$ *process in all of the analysis categories. Parameterisations are not reported where the expected SM yield is less than or equal to zero.*

Category	*A*	*B*	*C*
GG2H_GE2J_MJJ_0_350_PTH_0_60__0	1.370	1.670	-2.040
GG2H_1J_PTH_120_200__0	3.830	0.450	-3.280
GG2H_1J_PTH_120_200__1	1.330	3.190	-3.520
GG2H_1J_PTH_60_120__1	2.150	1.030	-2.180
GG2H_GE2J_MJJ_0_350_PTH_0_60__1	1.750	0.810	-1.550
GG2H_GE2J_MJJ_0_350_PTH_0_60__2	1.770	0.810	-1.580
GG2H_GE2J_MJJ_0_350_PTH_120_200__0	3.630	4.490	-7.120
GG2H_GE2J_MJJ_0_350_PTH_120_200__1	2.290	1.940	-3.230
GG2H_GE2J_MJJ_0_350_PTH_60_120__0	2.770	1.510	-3.280
GG2H_GE2J_MJJ_0_350_PTH_60_120__1	2.240	1.140	-2.380
GG2H_GE2J_MJJ_350_700_PTH_0_200__0	4.940	3.270	-7.220
GG2H_GE2J_MJJ_350_700_PTH_0_200__1	4.400	2.550	-5.950
GG2H_GE2J_MJJ_350_700_PTH_0_200__2	3.060	1.400	-3.460
GG2H_GE2J_MJJ_700_1000_PTH_0_200__0	3.600	2.430	-5.030
GG2H_GE2J_MJJ_700_1000_PTH_0_200__1	4.570	2.430	-5.990
GG2H_GE2J_MJJ_700_1000_PTH_0_200__2	5.440	1.520	-5.960
GG2H_GE2J_MJJ_GT1000_PTH_0_200__1	1.040	3.650	-3.690
GG2H_GE2J_MJJ_GT1000_PTH_0_200__2	4.040	1.960	-5.000
GG2H_PTH_200_300__0	2.390	3.980	-5.370
GG2H_PTH_200_300__1	3.800	3.910	-6.710
GG2H_PTH_300_450__0	5.000	6.620	-10.620
GG2H_PTH_300_450__1	5.520	6.530	-11.050
GG2H_PTH_300_450__2	7.370	5.390	-11.760
GG2H_PTH_450_650__0	7.140	11.530	-17.670
GG2H_PTH_450_650__1	10.850	11.650	-21.490
GG2H_PTH_450_650__2	9.740	10.090	-18.830
GG2H_PTH_GT650__0	12.540	13.280	-24.820
QQ2HLL_PTV_0_75__0	1.130	0.080	-0.220
QQ2HLL_PTV_0_75__1	2.300	1.120	-2.430
QQ2HLL_PTV_150_250__0	4.010	4.150	-7.160
QQ2HLL_PTV_150_250__1	3.110	2.350	-4.470
QQ2HLL_PTV_75_150__0	2.110	1.810	-2.920
QQ2HLL_PTV_75_150__1	4.500	1.280	-4.770
QQ2HLL_PTV_GT250__0	5.230	5.390	-9.620
QQ2HLNU_PTV_0_75__1	1.360	2.360	-2.720
QQ2HLNU_PTV_150_250__0	3.520	3.270	-5.790
QQ2HLNU_PTV_150_250__1	3.420	1.580	-4.000
QQ2HLNU_PTV_75_150__1	1.730	2.150	-2.890
QQ2HLNU_PTV_GT250__0	10.210	8.250	-17.470
QQ2HLNU_PTV_GT250__1	4.900	0.610	-4.500
QQ2HNUNU_PTV_0_75__0	1.080	0.810	-0.900
QQ2HNUNU_PTV_0_75__1	2.150	0.090	-1.240
QQ2HNUNU_PTV_150_250__0	3.040	2.220	-4.260
QQ2HNUNU_PTV_150_250__1	5.780	5.720	-10.500
QQ2HNUNU_PTV_150_250__2	3.650	2.270	-4.910
QQ2HNUNU_PTV_75_150__1	1.830	2.070	-2.900
QQ2HNUNU_PTV_75_150__2	2.770	2.140	-3.910
QQ2HNUNU_PTV_GT250__0	6.700	5.360	-11.060
QQ2HNUNU_PTV_GT250__1	7.760	3.910	-10.660
QQ2HQQ_0J__2	0.740	1.140	-0.880
QQ2HQQ_GE2J_MJJ_0_60__1	0.870	1.160	-1.030
QQ2HQQ_GE2J_MJJ_0_60__2	1.620	1.680	-2.300
QQ2HQQ_GE2J_MJJ_120_350__1	2.800	2.430	-4.240
QQ2HQQ_GE2J_MJJ_120_350__2	2.560	1.870	-3.430
QQ2HQQ_GE2J_MJJ_350_700_PTH_0_200__2	5.260	1.150	-5.410
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200__1	6.610	6.000	-11.600
QQ2HQQ_GE2J_MJJ_350_700_PTH_GT200__2	5.090	4.120	-8.210
QQ2HQQ_GE2J_MJJ_60_120__0	16.120	5.720	-20.830
QQ2HQQ_GE2J_MJJ_60_120__1	4.430	2.850	-6.280
QQ2HQQ_GE2J_MJJ_60_120__2	3.190	2.290	-4.480
QQ2HQQ_GE2J_MJJ_700_1000_PTH_GT200__0	9.520	2.780	-11.300
QQ2HQQ_GE2J_MJJ_700_1000_PTH_GT200__1	5.920	5.030	-9.950
THJB_MINUS1__0	2.510	2.170	-3.690
THJB__0	3.820	2.460	-5.280
THJB__1	2.770	3.020	-4.790
THW__0	3.080	2.350	-4.430
TTH_PTH_0_60__0	1.190	1.450	-1.640
TTH_PTH_0_60__1	1.720	0.760	-1.480
TTH_PTH_120_200__0	2.910	2.250	-4.160
TTH_PTH_120_200__1	2.660	2.080	-3.740
TTH_PTH_200_300__0	3.980	3.860	-6.840
TTH_PTH_60_120__0	2.090	1.250	-2.340
TTH_PTH_60_120__1	2.000	0.940	-1.950
TTH_PTH_GT300__0	6.410	6.670	-12.080
BKG_LIKE_TOP	2.580	1.560	-3.130

Table A.2 *Yield parametrisations as a ratio to the SM expected yield for the tHW process in all of the analysis categories. Parameterisations are not reported where the expected SM yield is less than or equal to zero.*

Bibliography

[1] ATLAS Collaboration. Measurement of the properties of Higgs boson production at $\sqrt{s} = 13$ TeV in the $H \to \gamma\gamma$ channel using 139 fb^{-1} of pp collision data with the ATLAS experiment. Technical report, CERN, Geneva, 2022. URL `https://cds.cern.ch/record/2814435`. 78 pages in total, 25 figures, 20 tables, submitted to JHEP. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGG-2020-16.

[2] ATLAS Collaboration. AtlFast3: the next generation of fast simulation in ATLAS. 2021. doi: 10.48550/ARXIV.2109.02551. URL `https://arxiv.org/abs/2109.02551`.

[3] Carsten Burgard. Example: Standard model of physics. URL `https://texample.net/tikz/examples/model-physics/`.

[4] A. K. Halder, Andronikos Paliathanasis, and P. G. L. Leach. Noether's Theorem and Symmetry, 2018.

[5] Mexican hat potential polar with details - higgs field. `https://www.pngitem.com/middle/mhxixb_mexican-hat-potential-polar-with-details-higgs-field`, Last accessed on 2023-04-20.

[6] John C. Collins, Davison E. Soper, and George Sterman. Factorization of hard processes in qcd, 2004.

[7] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions with LHC data. *Nuclear Physics B*, 867(2):244–289, feb 2013. doi: 10.1016/j.nuclphysb.2012.10.003. URL `https://doi.org/10.1016%2Fj.nuclphysb.2012.10.003`.

[8] Mark Thomson. *Modern particle physics*. Cambridge University Press, New York, 2013. ISBN 978-1-107-03426-6. doi: 10.1017/CBO9781139525367.

[9] CERN. CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector. 2017. doi: 10.23731/CYRM-2017-002. URL `https://e-publishing.cern.ch/index.php/CYRM/issue/view/32`.

[10] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group. 2013. doi: 10.5170/CERN-2013-004. URL `http://cds.cern.ch/record/1559921`.

[11] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, sep 2012. doi: 10.1016/j.physletb.2012.08.020. URL `https://doi.org/10.1016%2Fj.physletb.2012.08.020`.

[12] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, sep 2012. doi: 10.1016/j.physletb.2012.08.021. URL `https://doi.org/10.1016%2Fj.physletb.2012.08.021`.

[13] 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. *Physical Review Letters*, 114(19), may 2015. doi: 10.1103/physrevlett.114.191803. URL `https://doi.org/10.1103%2Fphysrevlett.114.191803`.

[14] Lyndon Evans and Philip Bryant. LHC machine. *Journal of Instrumentation*, 3(08):S08001–S08001, aug 2008. doi: 10.1088/1748-0221/3/08/s08001. URL `https://doi.org/10.1088/1748-0221/3/08/s08001`.

[15] Esma Mobs. The CERN accelerator complex. Complexe des accélérateurs du CERN. Jul 2016. URL `https://cds.cern.ch/record/2197559`. General Photo.

[16] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi: 10.1088/1748-0221/3/08/S08003.

[17] Public atlas luminosity results for run-2 of the LHC. URL `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`.

[18] Izaak Neutelings. Graphics with TikZ in LaTeX, Apr 2022. URL `https://tikz.net/`.

[19] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical report, 2010. URL `https://cds.cern.ch/record/1291633`.

[20] Javier Montejo Berlingen. Triggering in the ATLAS experiment. Technical report, CERN, Geneva, Oct 2020. URL `https://cds.cern.ch/record/2742661`.

[21] KyungEon Choi. Tracking and Vertexing with the ATLAS Inner Detector in the LHC Run-2. In Zhen-An Liu, editor, *Proceedings of International Conference on Technology and Instrumentation in Particle Physics 2017*, pages 400–403, Singapore, 2018. Springer Singapore. ISBN 978-981-13-1316-5.

[22] ATLAS Collaboration. Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2. *The European Physical Journal C*, 77(10):673, 2017. doi: 10.1140/epjc/s10052-017-5225-7. URL `https://doi.org/10.1140/epjc/s10052-017-5225-7`.

[23] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2): 444–450, 1987. ISSN 0168-9002. doi: https://doi.org/10.1016/0168-9002(87)90887-4. URL `https://www.sciencedirect.com/science/article/pii/0168900287908874`.

[24] Noemi Calace. Track and Vertex reconstruction in ATLAS for LHC Run-3 and High-Luminosity phases. 2021. URL `https://cds.cern.ch/record/2777660`.

[25] F. Meloni. Primary vertex reconstruction with the ATLAS detector. *Journal of Instrumentation*, 11(12):C12060–C12060, dec 2016. doi: 10.1088/1748-0221/11/12/c12060. URL `https://doi.org/10.1088/1748-0221/11/12/c12060`.

[26] Steven Schramm. ATLAS Jet Reconstruction, Calibration, and Tagging of Lorentz-boosted Objects. Technical report, CERN, Geneva, Nov 2017. URL `http://cds.cern.ch/record/2291608`.

[27] Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach. Technical report, CERN, Geneva, Dec 2017. URL `https://cds.cern.ch/record/2298955`. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-022.

[28] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, apr 2008. doi: 10.1088/1126-6708/2008/04/063. URL `https://doi.org/10.1088%2F1126-6708%2F2008%2F04%2F063`.

[29] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jet trimming. *Journal of High Energy Physics*, 2010(2), feb 2010. doi: 10.1007/jhep02(2010)084. URL `https://doi.org/10.1007%2Fjhep02%282010%29084`.

[30] ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in *pp* collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *The European*

Physical Journal C, 76(11), oct 2016. doi: 10.1140/epjc/s10052-016-4395-z. URL `https://doi.org/10.1140%2Fepjc%2Fs10052-016-4395-z`.

[31] Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2018(8):89, 2018. doi: 10.1007/JHEP08(2018)089. URL `https://doi.org/10.1007/JHEP08(2018)089`.

[32] 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(11), nov 2019. doi: 10.1140/epjc/s10052-019-7450-8. URL `https://doi.org/10.1140%2Fepjc%2Fs10052-019-7450-8`.

[33] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 76(5), may 2016. doi: 10.1140/epjc/s10052-016-4120-y. URL `https://doi.org/10.1140%2Fepjc%2Fs10052-016-4120-y`.

[34] J. Illingworth and J. Kittler. A survey of the Hough transform. *Computer Vision, Graphics, and Image Processing*, 44(1):87–116, 1988. ISSN 0734-189X. doi: https://doi.org/10.1016/S0734-189X(88)80033-1. URL `https://www.sciencedirect.com/science/article/pii/S0734189X88800331`.

[35] ATLAS Collaboration. Performance of missing transverse momentum reconstruction with the atlas detector using proton–proton collisions at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 78(11):903, 2018. doi: 10.1140/epjc/s10052-018-6288-9. URL `https://doi.org/10.1140/epjc/s10052-018-6288-9`.

[36] Bernd Stelzer. The New Small Wheel Upgrade Project of the ATLAS Experiment. Technical report, CERN, Geneva, Oct 2014. URL `https://cds.cern.ch/record/1958265`.

[37] Giovanni Calderini. The ATLAS ITk detector for High Luminosity LHC Upgrade. Technical report, CERN, Geneva, 2022. URL `https://cds.cern.ch/record/2798838`.

[38] Corentin Allaire. A High-Granularity Timing Detector (HGTD) in ATLAS : Performance at the HL-LHC. Technical report, CERN, Geneva, 2018. URL `https://cds.cern.ch/record/2302827`.

[39] Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System. Technical report, CERN, Geneva, 2017. URL `https://cds.cern.ch/record/2285584`.

[40] Christian Wolfgang Fabjan and F Gianotti. Calorimetry for Particle Physics. *Rev. Mod. Phys.*, 75:1243–1286, 2003. doi: 10.1103/RevModPhys.75.1243. URL `https://cds.cern.ch/record/692252`.

[41] last lecture 1) silicon tracking detectors 2 ... - birmingham. URL `http://epweb2.ph.bham.ac.uk/user/newman/appt08/lecture10.pdf`.

[42] R. Wigmans. *Calorimetry : energy measurement in particle physics / Richard Wigmans.* International series of monographs on physics ; 168. Oxford University Press, Oxford, second edition. edition, 2018. ISBN 9780191828652.

[43] Helge Voss. Successes, Challenges and Future Outlook of Multivariate Analysis In HEP. *Journal of Physics: Conference Series*, 608:012058, 2015. doi: 10.1088/1742-6596/608/1/012058.

[44] Johann Brehmer, Kyle Cranmer, Gilles Louppe, and Juan Pavez. A Guide to Constraining Effective Field Theories with Machine Learning. 2018. doi: 10.1103/PhysRevD.98.052004.

[45] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. 2014. doi: 10.48550/ARXIV.1412.6980. URL `https://arxiv.org/abs/1412.6980`.

[46] S. Agostinelli et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. ISSN 0168-9002. doi: https://doi.org/10.1016/S0168-9002(03)01368-8. URL `https://www.sciencedirect.com/science/article/pii/S0168900203013688`.

[47] P Calafiura, J Catmore, D Costanzo, and A Di Girolamo. ATLAS HL-LHC Computing Conceptual Design Report. Technical report, CERN, Geneva, Sep 2020. URL `http://cds.cern.ch/record/2729668`.

[48] ATLAS Collaboration. The ATLAS Simulation Infrastructure. *The European Physical Journal C*, 70(3):823–874, sep 2010. doi: 10.1140/epjc/s10052-010-1429-9. URL `https://doi.org/10.1140%2Fepjc%2Fs10052-010-1429-9`.

[49] Elmar Ritsch. Fast Calorimeter Punch-Through Simulation for the ATLAS Experiment, 2011. URL `https://cds.cern.ch/record/1388275`. Presented 28 Sep 2011.

[50] Francois Chollet et al. Keras, 2015. URL `https://github.com/fchollet/keras`.

[51] Nicolas Berger et al. Simplified template cross sections - stage 1.1. 2019. doi: 10.48550/ARXIV.1906.02754. URL `https://arxiv.org/abs/1906.02754`.

[52] LHC Higgs Cross Section Working Group. LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle. 2012. doi: 10.48550/ARXIV.1209.0040. URL `https://arxiv.org/abs/1209.0040`.

[53] CERN. CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector. 2017. doi: 10.23731/CYRM-2017-002. URL `https://e-publishing.cern.ch/index.php/CYRM/issue/view/32`.

[54] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), feb 2011. doi: 10.1140/epjc/s10052-011-1554-0. URL `https://doi.org/10.1140%2Fepjc%2Fs10052-011-1554-0`.

[55] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower monte carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010(6), jun 2010. doi: 10.1007/jhep06(2010)043. URL `https://doi.org/10.1007%2Fjhep06%282010%29043`.

[56] Paolo Nason. A new method for combining NLO QCD with shower monte carlo algorithms. *Journal of High Energy Physics*, 2004(11):040–040, nov 2004. doi: 10.1088/1126-6708/2004/11/040. URL `https://doi.org/10.1088%2F1126-6708%2F2004%2F11%2F040`.

[57] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *Journal of High Energy Physics*, 2007(11):070–070, nov 2007. doi: 10.1088/1126-6708/2007/11/070. URL `https://doi.org/10.1088%2F1126-6708%2F2007%2F11%2F070`.

[58] H. B. Hartanto, B. Jäger, L. Reina, and D. Wackeroth. Higgs boson production in association with top quarks in the POWHEG BOX. *Physical Review D*, 91(9), may 2015. doi: 10.1103/physrevd.91.094003. URL `https://doi.org/10.1103%2Fphysrevd.91.094003`.

[59] 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. *Journal of High Energy Physics*, 2014(7), jul 2014. doi: 10.1007/jhep07(2014)079. URL `https://doi.org/10.1007%2Fjhep07%282014%29079`.

[60] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *Journal of High Energy Physics*, 2013(3), mar 2013. doi: 10.1007/jhep03(2013)015. URL `https://doi.org/10.1007%2Fjhep03%282013%29015`.

[61] Stefano Frixione, Eric Laenen, Patrick Motylinski, Chris White, and Bryan R Webber. Single-top hadroproduction in association with a W boson. *Journal of High Energy Physics*, 2008(07):029–029, jul 2008. doi: 10.1088/1126-6708/2008/07/029. URL `https://doi.org/10.1088%2F1126-6708%2F2008%2F07%2F029`.

[62] Federico Demartin, Benedikt Maier, Fabio Maltoni, Kentarou Mawatari, and Marco Zaro. tWH associated production at the LHC. *The European Physical Journal C*, 77(1), jan 2017. doi: 10.1140/epjc/s10052-017-4601-7. URL `https://doi.org/10.1140%2Fepjc%2Fs10052-017-4601-7`.

[63] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, jun 2008. doi: 10.1016/j.cpc.2008.01.036. URL `https://doi.org/10.1016%2Fj.cpc.2008.01.036`.

[64] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J.*, C76(4):196, 2016. doi: 10.1140/epjc/s10052-016-4018-8.

[65] ATLAS Collaboration. Measurement of higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Physical Review D*, 90(11), dec 2014. doi: 10.1103/physrevd.90.112015. URL `https://doi.org/10.1103%2Fphysrevd.90.112015`.

[66] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS detector. *The European Physical Journal C*, 77(7), jul 2017. doi: 10.1140/epjc/s10052-017-5031-2. URL `https://doi.org/10.1140%2Fepjc%2Fs10052-017-5031-2`.

[67] Forward jet vertex tagging using the particle flow algorithm. Technical report, CERN, Geneva, Jul 2019. URL `https://cds.cern.ch/record/2683100`. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2019-026.

[68] ATLAS Collaboration. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{13}$ TeV. *The European Physical Journal C*, 79(11), nov 2019. doi: 10.1140/epjc/s10052-019-7450-8. URL `https://doi.org/10.1140%2Fepjc%2Fs10052-019-7450-8`.

[69] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4232-2. doi: 10.1145/2939672.2939785. URL `http://doi.acm.org/10.1145/2939672.2939785`.

[70] ATLAS Collaboration. CP properties of higgs boson interactions with top quarks in the $t\bar{t}H$ and tH processes using $H \to \gamma\gamma$ with the ATLAS detector. *Physical Review Letters*, 125(6), aug 2020. doi: 10.1103/physrevlett.125.061802. URL `https://doi.org/10.1103%2Fphysrevlett.125.061802`.

[71] Guolin Ke et al. Lightgbm: A highly efficient gradient boosting decision tree. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach,

R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL `https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf`.

[72] Meghan Frate, Kyle Cranmer, Saarik Kalia, Alexander Vandenberg-Rodes, and Daniel Whiteson. Modeling Smooth Backgrounds and Generic Localized Signals with Gaussian Processes. 2017. doi: 10.48550/ARXIV.1709.05681. URL `https://arxiv.org/abs/1709.05681`.

[73] A. Wald. Sequential Tests of Statistical Hypotheses. *The Annals of Mathematical Statistics*, 16(2):117 – 186, 1945. doi: 10.1214/aoms/1177731118. URL `https://doi.org/10.1214/aoms/1177731118`.

[74] ATLAS Collaboration. Combined measurements of higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment. *Physical Review D*, 101(1), jan 2020. doi: 10.1103/physrevd.101.012002. URL `https://doi.org/10.1103%2Fphysrevd.101.012002`.

[75] scikit-optimize. URL `https://scikit-optimize.github.io/stable/index.html`.

[76] T. Carter. ATLAS Collaboration. A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery. *Nature*, 607(7917): 52–59, jul 2022. doi: 10.1038/s41586-022-04893-w. URL `https://doi.org/10.1038%2Fs41586-022-04893-w`.