



COPYRIGHT AND USE OF THIS THESIS

This thesis must be used in accordance with the provisions of the Copyright Act 1968.

Reproduction of material protected by copyright may be an infringement of copyright and copyright owners may be entitled to take legal action against persons who infringe their copyright.

Section 51 (2) of the Copyright Act permits an authorized officer of a university library or archives to provide a copy (by communication or otherwise) of an unpublished thesis kept in the library or archives, to a person who satisfies the authorized officer that he or she requires the reproduction for the purposes of research or study.

The Copyright Act grants the creator of a work a number of moral rights, specifically the right of attribution, the right against false attribution and the right of integrity.

You may infringe the author's moral rights if you:

- fail to acknowledge the author of this thesis if you quote sections from the work
- attribute this thesis to another author
- subject this thesis to derogatory treatment which may prejudice the author's reputation

For further information contact the University's Copyright Service.

sydney.edu.au/copyright

Search for the Higgs Boson decaying to tau leptons at ATLAS using multi-variate analysis techniques

Mark Scarcella

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy
Faculty of Science
University of Sydney
2016

Abstract

This thesis presents three differing approaches to the search for the Standard Model Higgs boson decaying to tau leptons using $\sqrt{s} = 7$ TeV proton-proton collision data from the ATLAS experiment at the LHC. Multi-variate analysis techniques involving boosted decision trees are used to extend an existing cut-based analysis procedure. The expected 95% confidence level upper limit on the observed cross-section is compared between the analyses. The upper limit at a Higgs mass of $m_H = 125$ GeV is improved from $2.9_{-2.1}^{+4.3}$ to $2.3_{-1.7}^{+3.3}$ times the Standard Model prediction, after implementing multi-variate techniques. No significant excess is seen in data for any analysis strategy. The most sensitive measurement of the signal strength normalised to the Standard Model prediction was observed to be $\hat{\mu} = 1.6 \pm 1.1$, corresponding to 1.4σ upward fluctuation of the background-only model to match the data.

Acknowledgements

First and foremost, I must thank my supervisor Professor Kevin Varvell, not only for his guidance during my thesis, but also for inspiring me to take on my 3rd year special project on the tau reconstruction algorithm back in my undergrad days. Although it has taken me a little longer than planned, his expert knowledge of particle physics and subtle hand has kept me on track to submit the work presented in this thesis. I must also thank the other members of the Sydney Particle Physics group, with special mention to my associate supervisors Dr. Aldo Saavedra and Dr. Geng-Yuan Jeng for their experienced knowledge of the ATLAS experiment and analyses, and Dr. Bruce Yabsley for all manner of advice on statistics and particle physics theory.

To my physics office mates and friends, Curtis, Cameron, Nik, Ian, and The Chuz – I would not have lasted a week without being able to bounce around ideas and concepts with you, and having you keep me sane with coffee and beer.

My family and friends have given me unwavering support throughout this endeavour. Elysha, Tom, Sammy, Kirsten, Pete, Kieran, Yasmin, Justin, and Olivia – thanks for always being available for a meal, a beer, and a laugh. I must of course give special thanks to my parents Roslyn and Carl for always being there for anything and everything, providing me the encouragement to continue when it all seemed a bit much.

Finally, I must thank the other members of the ATLAS collaboration, in particular the HSG4 working group, for this thesis would not exist without them.

Declaration

The work presented in this thesis is the culmination of studies completed by the author between March 2010 and December 2015. Due to the size and scope of the experiment, it is not possible for any one researcher to undertake an analysis alone. This work is only made possible by the efforts of countless members of the ATLAS and LHC collaborations.

The analysis in Chapter 4 is based on the work presented in [1]. While the author had no direct input on the paper, the analysis was recreated, with modifications to the details of the background estimation and limit extraction methods, to be used as a benchmark for further research. The hybrid analysis method outlined in Chapter 5 was first proposed and presented by the author, and grew into a collaborative effort alongside analysts within the ATLAS $H \rightarrow \tau_\ell \tau_h$ working group (who have subsequently published their results in [2]) as the start of the fully multi-variate analysis in Chapter 6. The official ATLAS analysis and that of the author were conducted in parallel, after the basics of a multi-variate procedure had been established. The optimisation of the event categorisation, background estimation methods, boosted decision trees, and final signal region selections for the fully multi-variate analysis were all carried out by the author. All ATLAS monte-carlo samples were generated by the relevant working groups. Unless otherwise stated, all other work presented in this thesis was completed by the author.

Contents

1. Introduction	1
2. The Standard Model	3
2.1. Particle Content	4
2.2. Interactions of the Standard Model	4
2.2.1. Electromagnetic Interaction	6
2.2.2. Strong Interaction	7
2.2.3. Weak Interaction and Electroweak Unification	9
2.2.4. The Higgs Mechanism	11
2.3. The Higgs Boson	16
3. The ATLAS Experiment	23
3.1. The Large Hadron Collider	23
3.1.1. LHC Design	25
3.1.2. Injection chain	27
3.2. The ATLAS Detector	28
3.2.1. Magnet System	30
3.2.2. Inner Detector	30
3.2.3. Calorimetry	33
3.2.4. Muon Spectrometer	36
3.2.5. Trigger System	39
3.3. ATLAS Monte-Carlo and Data	43
3.3.1. Generation	44
3.3.2. Simulation	44
3.3.3. Reconstruction and Identification	45
4. $H \rightarrow \tau_\ell \tau_h$ Search with the ATLAS Detector	52
4.1. Signal and Background Processes	52
4.1.1. $H \rightarrow \tau_\ell \tau_h$ Signal	52

4.1.2.	Background Contributions	54
4.2.	Monte-Carlo Samples	57
4.3.	ATLAS Data	58
4.4.	Object Selection and m_{MMC} Calculator	59
4.4.1.	Muons	59
4.4.2.	Electrons	61
4.4.3.	Jets	61
4.4.4.	Hadronic Taus	62
4.4.5.	Missing Transverse Energy	62
4.4.6.	Missing Calorimeter Region and Overlap Removal	62
4.4.7.	Missing Mass Calculator (MMC)	63
4.5.	Event Pre-Selection	64
4.6.	Analysis Strategy	68
4.6.1.	VBF Category	68
4.6.2.	Boosted Category	69
4.6.3.	n-Jet Categories	70
4.6.4.	Analysis Level Selection	70
4.7.	Background Estimation Methods	76
4.7.1.	OS-rSS and the k -Factor Method	76
4.7.2.	$Z \rightarrow \tau\tau$ Background Estimation	79
4.7.3.	$Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$ Background Estimation	82
4.7.4.	$t\bar{t}$ and single- t Background Estimation	83
4.7.5.	$W(\rightarrow \ell\nu) + \text{jets}$ Background Estimation	83
4.7.6.	Fake Factor Method	84
4.8.	Systematic Uncertainties	89
4.8.1.	$H \rightarrow \tau\tau$ Branching Fraction Uncertainties	91
4.8.2.	QCD Scale and PDF Uncertainties	91
4.8.3.	Electron Uncertainties	92
4.8.4.	Muon Uncertainties	93
4.8.5.	Tau Uncertainties	93
4.8.6.	Jet and MET Uncertainties	94
4.8.7.	Trigger and Luminosity Uncertainties	94
4.8.8.	Background Estimate Uncertainty	94
4.9.	Limit Setting Procedure	95
4.9.1.	Exclusion Limits and the Signal Strength	95
4.9.2.	Statistical Significance	99

4.10. Results	100
4.10.1. Fit Validation	101
4.10.2. Exclusion Limits and Signal Strength	103
5. Multi-Variate Analysis Techniques and an $H \rightarrow \tau_\ell \tau_h$ Hybrid Analysis	110
5.1. Multi-Variate Analyses	110
5.1.1. Decision Trees	110
5.1.2. Boosting	113
5.2. Hybrid Multi-Variate/Cut-Based Analysis	118
5.2.1. Categorisation and Background Modelling	120
5.2.2. Growing BDT _{HBD}	121
5.2.3. Results	128
6. $H \rightarrow \tau_\ell \tau_h$ Multi-Variate Analysis	139
6.1. MVA Categorisation	139
6.2. Background Estimation Methods	140
6.2.1. Background k -Factors	142
6.2.2. Fake-Factor for Boosted Category	143
6.2.3. Background Estimation for TMVA	145
6.3. Growing BDTs	146
6.3.1. Training and Testing Samples	146
6.3.2. Training Parameters	151
6.3.3. Optimising BDT Variable List	153
6.3.4. Cutting on the BDTs	157
6.3.5. BDT Interpretation	166
6.4. Results	168
6.4.1. Exclusion Limits and Signal Strength	176
6.5. Comparison Between Analysis Strategies	178
7. Conclusions	183
A. Non-Jet Combination Results	185
B. Category Specific Pull Distributions	187
Bibliography	199
List of Figures	208

List of Tables 220

List of Algorithms 224

Chapter 1.

Introduction

The Standard Model of particle physics is the theory describing the elementary particles of nature, and how they interact with each other. While it does not offer an explanation for the existence of dark matter or gravity, the Standard Model has been extensively validated by experiment. The Large Hadron Collider (LHC) at CERN, operating since 2010, is the world's most powerful particle accelerator, colliding proton beams inside of four detecting experiments, including the ATLAS detector. A primary objective of ATLAS was to discover one of the last missing pieces of the Standard Model – the Higgs boson.

First proposed in 1964, the Higgs mechanism provides a way to give masses to the particles of the Standard Model. The detection of the Higgs boson at a particle accelerator eluded physicists for nearly 50 years, until 2012 when the ATLAS and CMS experiments on the LHC both announced the observation of a Higgs-like particle near 126 GeV. The confirmation of its existence earned both François Englert and Peter Higgs the Nobel Prize in 2013, for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles. Although it has been discovered, the properties of the Higgs boson must now be studied, including its coupling strengths to the various particles of the Standard Model.

The main objective of this thesis is to investigate the use of *multi-variate analysis* (MVA) techniques as an extension to an existing search for the Standard Model Higgs boson in the di-tau decay mode, using data from the ATLAS detector. MVA methods involve the study of multiple variables in a system *simultaneously*, exploiting potentially unseen correlations in a higher dimensional parameter space. These methods are often applied via machine learning algorithms, such as a *boosted decision tree* (BDT), to produce classifiers that can discriminate analysis objects into different classes. Due to the highly correlated nature of particles in high energy physics events, and the many variables constructed from the information read out by the detectors, MVAs can be an extremely useful tool in particle physics analyses.

This thesis presents three approaches to the search for the Standard Model Higgs boson decaying to two tau leptons. The analyses are performed on proton-proton collision data from the ATLAS detector during 2011, at a centre-of-mass energy of 7 TeV.

In Chapter 2, the physics of the Standard Model is reviewed. The fundamental particles of the theory and their interactions are described using the Lagrangian formalism. The theory of the Higgs boson is also discussed, along with a summary of its properties, and the searches leading up to its discovery in 2012.

An overview of the LHC and the ATLAS experiment is presented in Chapter 3, detailing the construction and operation of the various sub-detectors, and other detector components. The ATLAS particle reconstruction and identification methods are then described, with a particular focus on the algorithms used in this thesis.

Chapter 4 gives a detailed review of an ATLAS search for the Higgs Boson decaying to two tau leptons, using the so-called *cut-based* method. After outlining the signal and background contributions for the analysis, the methods by which events are selected and categorised is described. The approaches taken to estimate the various Standard Model backgrounds are then illustrated, followed by a discussion of the treatment of systematic uncertainties in these estimates. Finally, the procedure for extracting exclusion limits and signal strengths is explained, and the final results are presented, along with the statistical significance.

An extension to the cut-based analysis is presented in Chapter 5. Starting with an overview of MVA techniques – namely the implementation of BDTs – their usefulness in selecting Higgs bosons produced via *vector boson fusion* is then investigated. The exclusion limits obtained from this *hybrid* analysis are then compared back to the cut-based analysis.

The MVA approach is expanded in Chapter 6 by using BDTs to select events from four topologically different analysis categories. The background estimation methods from the cut-based analysis are updated here, to provide a better agreement between ATLAS data and simulation. A more detailed look at how the BDT is constructed is taken, studying the how the input variables and parameters can affect the sensitivity of the analysis. The results of this *full-MVA* analysis are presented, and comparisons made between it and the cut-based and hybrid analyses.

Finally, Chapter 7 reviews the conclusions drawn from the analyses presented in this thesis, and some final remarks are made.

Chapter 2.

The Standard Model

The Standard Model (SM) is the current best attempt at explaining the elementary particles of the universe, and the forces that govern their interactions. In this chapter, Section 2.1 summarises the particle content of the SM, followed by a description of the interactions of the theory in Section 2.2. Section 2.2.4 describes the Higgs mechanism, followed by a discussion of the searches for the SM Higgs boson, and its eventual discovery in Section 2.3.

Throughout this thesis, *natural units* are used, whereby the speed of light, the reduced Planck constant, and the Boltzmann constant, are all set to unity ($c = \hbar = k_B = 1$). All charges are expressed as multiples of the charge on the electron (e), and mass, energy and momentum are all given in terms of electron volts (eV).

The mathematical structure of the SM is described using the Lagrangian formalism. A Lorentz invariant Lagrangian density \mathcal{L} describes the masses and interactions of all the particles in the theory. From Noether's theorem [3], imposing *local gauge invariance* on the system (that is, ensuring that physics is the same across all space-time) results in a conserved current. This conserved current is associated with the introduction of the force carriers of the SM that mediate the fundamental interactions of the theory. The Lagrangian density of the SM is given by

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD} + \mathcal{L}_{Higgs} \quad (2.1)$$

where \mathcal{L}_{EW} (Sections 2.2.1 and 2.2.3) describes the electroweak part of the theory, \mathcal{L}_{QCD} (Section 2.2.2) describes the strong interaction, and \mathcal{L}_{Higgs} (Section 2.2.4) describes the Higgs interaction, as well as the masses for the particles¹.

2.1. Particle Content

The SM consists of two types of particle; the spin- $1/2$ *fermions* that make up matter, and the integer spin *bosons* that mediate their interactions. The particles of the SM and some basic properties are listed in Table 2.1.

The fermions of the SM are split into the *leptons* and the *quarks*, each of which are split into three *generations*. Each generation is grouped into *weak isospin doublets*, where weak isospin is a conserved quantity for the Weak interaction (described below). The leptons include the charged variety; the electron (e), muon (μ), and tau (τ), and their corresponding neutral partners, the electron-neutrino (ν_e), muon-neutrino (ν_μ), and tau-neutrino (ν_τ). The neutrinos have been inferred to be massive, albeit with much smaller masses than the other particles [4]. The quarks come in 6 flavours; the *up-type* quarks with a charge of $2/3$, up (u), charm (c), and top (t), and the *down-type* quarks with a charge of $-1/3$, down (d), strange (s), and bottom (b). All fermions have a corresponding *anti-particle*, in which the quantum numbers associated with the internal symmetries of the particle (the *generalised charges*) have been reversed. The quarks are never seen as free particles, rather, they must be grouped into the *hadrons*, which themselves come in two types; the *baryons* that contain three quarks or anti-quarks (qqq , $\bar{q}\bar{q}\bar{q}$), and the *mesons* that contain a quark anti-quark pair ($q\bar{q}$).

2.2. Interactions of the Standard Model

The interactions of the fermions are mediated by the spin-1 bosons of the theory. These particles come about through a requirement that the Lagrangian be invariant under local transformations. The underlying structure of the SM is a symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The $SU(3)_C$ part of the theory describes the *strong* interaction, which is mediated by the 8 gluons (g), and only exists between *coloured* particles (the quarks). The $SU(2)_L \times U(1)_Y$ is a unification of the *weak* and *electromagnetic* forces into the *electroweak* force. This symmetry is *spontaneously*

¹Except neutrino masses, which are yet to be added into the SM.

		Generation	Flavour	Charge [e]	Mass [GeV]
Fermions	Leptons	I	ν_e	0	$\lesssim 2 \times 10^{-9}$
			e	-1	5.110×10^{-4}
		II	ν_μ	0	$\lesssim 1.9 \times 10^{-4}$
			μ	-1	1.057×10^{-1}
		III	ν_τ	0	$\lesssim 18.2 \times 10^{-3}$
			τ	-1	1.777
	Quarks	I	u	2/3	$(2.3_{-0.5}^{+0.7}) \times 10^{-3}$
			d	-1/3	$(4.8_{-0.3}^{+0.5}) \times 10^{-3}$
		II	c	2/3	1.275 ± 0.025
			s	-1/3	$(95 \pm 5) \times 10^{-3}$
III		t	2/3	$173.07 \pm 0.52 \pm 0.72$	
		b	-1/3	4.18 ± 0.03	
Bosons	Force	Mediator	Charge [e]	Mass [GeV]	
	Strong	g	0	0	
	Weak	Z	0	91.1876 ± 0.0021	
	Weak	W^\pm	± 1	80.385 ± 0.015	
	Electromagnetic	γ	0	0	
		H	0	125.09 ± 0.24	

Table 2.1.: Particle Content of the Standard Model [5]. Errors are statistical and systematic when given. Masses of the neutrinos are inferred to be very small, but non-zero.

broken by a process known as the Higgs mechanism into the $U(1)_{EM}$, with the associated electromagnetic force carrier the photon (γ), and $SU(2)_L$, mediated by the weak bosons (W^\pm and Z^0). This breaking process also gives rise to the spin-0 Higgs boson H , which leads to the SM particles acquiring mass terms in the theory.

2.2.1. Electromagnetic Interaction

The electromagnetic force, formalised in quantum electrodynamics (QED), describes the interaction of charged fermions with photons. QED, perfected in the 1940's, has been experimentally verified to very high precision.

In the Lagrangian formalism, the free propagation of a fermion of mass m can be described by the Dirac Lagrangian

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad (2.2)$$

where ψ is a 4-component spinor field representing the fermion, and γ^μ are the Dirac matrices [6]. This Lagrangian density is invariant under global transformations of the form

$$\psi \rightarrow e^{i\theta} \psi \quad (2.3)$$

where θ is any real number, and the set of which make up the unitary transformations $U(1)$. The SM is required to also be locally gauge invariant, that is, invariant under phase transformations of the type

$$\psi \rightarrow e^{i\theta(x)} \psi \quad (2.4)$$

where $\theta(x)$ is now a function of space-time coordinates. The form of the Lagrangian in equation 2.2 is not invariant under such a transformation. To preserve locality, writing $\theta(x) = q\alpha(x)$, where q is the charge of the fermion, the derivative ∂_μ can be replaced by the *covariant derivative*

$$D_\mu = \partial_\mu + iqA_\mu \quad (2.5)$$

where A_μ is a new vector field that transforms like

$$A_\mu \rightarrow A_\mu - \partial_\mu \alpha(x) \quad (2.6)$$

To complete the Lagrangian, a free term for the vector field A_μ must be added, which from the Proca Lagrangian² is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m_A^2 A^\mu A_\mu \quad (2.7)$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, and m_A is the mass of the vector field. In order to maintain local gauge invariance, m_A must equal zero, and thus the massless vector field can be identified as the photon. The final Lagrangian density describing the electromagnetic interaction is therefore

$$\begin{aligned} \mathcal{L}_{QED} &= \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \\ &= \underbrace{\bar{\psi}i\gamma^\mu \partial_\mu \psi}_{\text{fermion kinetic}} - \underbrace{m\bar{\psi}\psi}_{\text{fermion mass}} - \underbrace{\frac{1}{4}F^{\mu\nu}F_{\mu\nu}}_{\text{photon kinetic}} - \underbrace{q\bar{\psi}\gamma^\mu \psi A_\mu}_{\text{interaction}} \end{aligned} \quad (2.8)$$

which contains kinetic terms for the fermion and photon, a mass term for the fermion, and an interaction term, with the strength of the interaction given by the electric charge q .

2.2.2. Strong Interaction

While QED describes the interaction of charged particles, Quantum Chromo-Dynamics (QCD) describes the interaction of *coloured* particles. Quarks come in three colours – red (r), green (g) and blue (b) – as well as their anti-colours ($\bar{r}, \bar{g}, \bar{b}$), while gluons, the force carrier of the strong interaction, have a unit of colour and anti-colour (for example, $r\bar{g}$). All other particles in the SM are colourless, and so do not take part in the strong interaction.

²The Proca Lagrangian describes a particle of spin 1 and mass m .

Similar to QED being based on the $U(1)$ symmetry group, QCD is based on the $SU(3)_c$ symmetry group, where c refers to the colour charge. Now, writing the quarks as spinors of the form

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}, \quad \psi_r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \psi_g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \psi_b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (2.9)$$

the Lagrangian density for quarks is the same as in Equation 2.2, however it must now be invariant under local $SU(3)_c$ transformations

$$\psi \rightarrow e^{ig_s \alpha^a(x) T^a} \psi \quad (2.10)$$

where g_s is the strong coupling constant, $\alpha^a(x)$ are again functions of space-time, and T^a are the (8) generators of $SU(3)$. The derivative ∂_μ in the Lagrangian density is then replaced with the covariant derivative

$$D_\mu = \partial_\mu + ig_s T^a G_\mu^a \quad (2.11)$$

where each G_μ^a is one of the 8 colour/anti-colour combinations, required to transform like

$$G_\mu^a \rightarrow G_\mu^a - \partial_\mu \alpha^a(x) - g_s f^{abc} \alpha^b(x) G_\mu^c \quad (2.12)$$

where f^{abc} are the structure constants. The final Lagrangian for QCD is then given by

$$\mathcal{L}_{QCD} = \underbrace{\bar{\psi} i \gamma^\mu \partial_\mu \psi}_{\text{quark kinetic}} - \underbrace{m \bar{\psi} \psi}_{\text{quark mass}} - \underbrace{\frac{1}{4} G^{a\mu\nu} G_{\mu\nu}^a}_{\text{gluon kinetic}} - \underbrace{\frac{1}{2} g_s \bar{\psi} \gamma^\mu T^a G_\mu^a \psi}_{\text{interaction}} \quad (2.13)$$

While the form looks similar to the QED case, it must be noted that here the gluon kinetic term contains trilinear and quartic terms, allowing gluons to self-interact. This self-interaction

is also responsible for the phenomenon of *colour confinement*, which will not be discussed here, however its effect is to only allow colourless particles (equal amounts of a colour and its anti-colour for mesons, or one of each colour for baryons) to exist freely in nature.

2.2.3. Weak Interaction and Electroweak Unification

The weak force is responsible for interactions between the left-handed fermions, and can be described by the $SU(2)_L$ symmetry group, with L referring to left-handedness. The fermions can be written in terms of their left-handed weak isospin doublets. For example, the first generation of leptons is given by

$$\psi_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad (2.14)$$

Once again, by analogy to Equation 2.2, the Lagrangian density for a pair of weakly interacting massless left-handed fermions is given by

$$L = \bar{\psi}_L (i\gamma^\mu \partial_\mu) \psi_L \quad (2.15)$$

which must transform like

$$\psi_L = e^{i\alpha^a(x) \cdot \sigma^a} \psi_L \quad (2.16)$$

where σ^a are the generators of $SU(2)$ (the Pauli matrices), and $\alpha^a(x)$ are functions of space-time. Again, imposing local gauge invariance of the Lagrangian under $SU(2)_L$, the derivative ∂_μ is replaced by a new covariant derivative

$$D_\mu = \partial_\mu + i\frac{g}{2} \sigma^a W_\mu^a \quad (2.17)$$

where g here is the strength of the weak interaction, and W_μ^a are three new massless vector fields that transform as

$$W_\mu^a \rightarrow \partial_\mu \alpha^a(x) - g \varepsilon^{abc} \alpha^b(x) W_\mu^c \quad (2.18)$$

where ε^{abc} is the Levi-Civita tensor.

While the $SU(2)_L$ theory explains much of the weak interaction, it is incomplete – for example, the weak gauge bosons are experimentally verified to have non-zero mass. A new symmetry can therefore be imposed, $SU(2)_L \times U(1)_Y$ (where Y is the *hypercharge*), which unifies the electromagnetic and weak forces into the *electroweak* interaction via the Weinberg-Salam-Glashow Model [7]. Electroweak unification requires the SM Lagrangian density to be invariant under *both* $U(1)_Y$ and $SU(2)_L$ simultaneously.

Unlike $SU(2)_L$, the $U(1)_Y$ theory couples to both the left- and right-handed isospins, with a local transformation of the form

$$\psi \rightarrow e^{i\frac{g'}{2}\alpha(x)} \psi \quad (2.19)$$

where g' is the strength of the interaction. Once again, to ensure local gauge invariance, the covariant derivative is replaced by

$$D_\mu = \partial_\mu + i\frac{g'}{2}B_\mu \quad (2.20)$$

resulting in a new massless gauge field B_μ that couples to both left and right handed fermions.

The combined $SU(2)_L \times U(1)_Y$ symmetry of the electroweak theory therefore gives the covariant derivative

$$D_\mu = \partial_\mu + i\frac{g}{2}\sigma^a W_\mu^a + i\frac{g'}{2}B_\mu \quad (2.21)$$

The complete Lagrangian density for the theory is then given by

$$\begin{aligned}
\mathcal{L}_{EW} = & \underbrace{\frac{1}{4}W^{a\mu\nu}W_{\mu\nu}^a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu}}_{\text{gauge field kinetic and self-interaction}} - \underbrace{\bar{\psi}_L\gamma^\mu(i\partial_\mu - \frac{g'}{2}B_\mu - i\frac{g_2}{2}\sigma^a W_\mu^a)}_{\text{left-handed fermion kinetic and interaction}}\psi_L \\
& + \underbrace{\bar{\psi}_R\gamma^\mu(i\partial_\mu - \frac{g'}{2}B_\mu\sigma^0)}_{\text{right-handed fermion kinetic and interaction}}\psi_R \quad (2.22)
\end{aligned}$$

where σ^0 is the 2×2 unit matrix. In its current state, the theory still has massless weak bosons and massless fermions. To introduce mass into the theory, the $SU(2)_L \times U(1)_Y$ symmetry must be *spontaneously broken* by the Higgs mechanism.

Note that the matrices W_μ^a can be written as

$$\mathbf{W}_\mu = \begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & -W_\mu^3 \end{pmatrix} \quad (2.23)$$

and it is useful to define (for reasons that will be clear soon)

$$\sigma^\pm = \frac{1}{\sqrt{2}}(\sigma_\mu^1 \pm i\sigma_\mu^2) \quad \text{and} \quad W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (2.24)$$

where W^\pm are identified as the charged gauge boson fields. As such the covariant derivative in equation 2.21 can be written as

$$D_\mu = \partial_\mu + i\frac{g}{2}(W_\mu^+ \sigma^+ + W_\mu^- \sigma^-) + \left(i\frac{g'}{2}B_\mu + i\frac{g}{2}W_\mu^3 \sigma^3 \right) \quad (2.25)$$

2.2.4. The Higgs Mechanism

The Higgs mechanism provides a way to give the gauge bosons of the weak force mass, as observed in nature [8–13]. As in the previous section, a Lagrangian density is constructed that is invariant under $SU(2) \times U(1)$. A doublet of complex scalar fields, Φ , is defined as

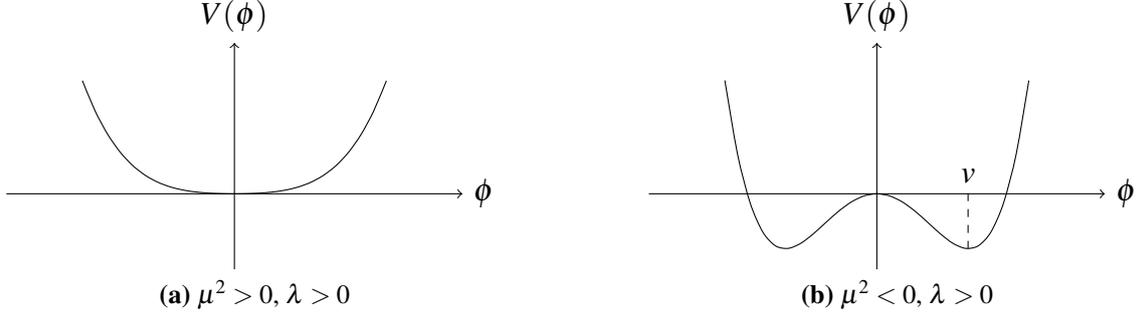


Figure 2.1.: Higgs potential $V(\phi) = \mu^2\phi^2 + \lambda^2\phi^4$ of a scalar field ϕ , and v is the vacuum expectation value.

$$\Phi = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (2.26)$$

A Lagrangian density satisfying local $SU(2) \times U(1)$ symmetry is

$$\mathcal{L}_\Phi = (D_\mu \Phi)^\dagger D^\mu \Phi - V(\Phi^\dagger \Phi) \quad (2.27)$$

where D_μ is as defined in equation 2.25 acting on Φ , and the potential $V(\Phi^\dagger \Phi)$ is chosen to be

$$V(\phi) = \mu^2 \Phi^\dagger \Phi + \lambda^2 (\Phi^\dagger \Phi)^2 \quad (2.28)$$

where μ and λ are real numbers. Figure 2.1 shows the form of $V(\phi)$ with different choices of μ and λ . By choosing $\mu^2 < 0$, the potential of this form has degenerate minima at

$$\Phi^\dagger \Phi = -\frac{\mu^2}{2\lambda} \quad (2.29)$$

Choosing any one of these minima will spontaneously break the $U(1)$ rotational symmetry of the theory. A particular ground state can therefore be adopted by carefully choosing the values of α^a in equation 2.16 such that $\phi_1^2 = \phi_3^2 = \phi_4^2 = 0$ and $\phi_2^2 = -\frac{\mu^2}{2\lambda}$ so that

$$\Phi_{ground} = \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.30)$$

where it is useful to define $v^2 = -\frac{\mu^2}{2\lambda}$ to be the *Vacuum Expectation Value* (VEV) of the potential. Expanding around the ground state in the direction of ϕ_1, ϕ_3 or ϕ_4 leaves the Lagrangian density at a minimum, while expanding in the direction of ϕ_2 results in a massive scalar field [14]. Calling this field the Higgs field H , the Lagrangian density can be expanded about the minimum

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+H \end{pmatrix} \quad (2.31)$$

Substituting equation 2.31 into the potential in equation 2.28 gives

$$V(\Phi) = V(H) = \frac{1}{2}\mu^2(v+H)^2 + \frac{1}{4}\lambda^2(v+H)^4 \quad (2.32)$$

Once expanded, the coefficient of the H^2 term describes the mass of the Higgs boson $m_H = \sqrt{-2\mu^2}$, and the remaining terms describe self-interactions of a new spin-0 Higgs boson. Substituting this into equation 2.27 gives

$$\begin{aligned} \mathcal{L}_\Phi &= (D_\mu \Phi)^\dagger D^\mu \Phi - V(\Phi) \\ &= \frac{1}{2} \partial_\mu H \partial^\mu H + \frac{g^2}{8} W_\mu^- W^{+\mu} \left(v + \frac{1}{\sqrt{2}} H \right)^2 \\ &\quad + \left[\frac{g^2}{4} W_\mu^3 W^{3\mu} - \frac{gg'}{2} W_\mu^3 B^\mu + \frac{g'^2}{4} B_\mu B^\mu \right] \left(v + \frac{1}{\sqrt{2}} H \right)^2 \\ &\quad - V(H) \end{aligned} \quad (2.33)$$

The W_μ^\pm fields are present, but the observed Z_μ and A_μ are not. Notice that a change of basis can be performed of the type

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \quad (2.34)$$

to produce the required fields, where θ_W is a parameter known as the Weinberg angle. Now, equation 2.27 can be modified to give

$$\begin{aligned} \mathcal{L}_\Phi = & \underbrace{\frac{1}{2} \partial_\mu H \partial^\mu H}_{\text{Higgs kinetic}} \\ & + \underbrace{\frac{1}{4} g^2 W_\mu^- W^{+\mu} \left(v + \frac{1}{\sqrt{2}} H \right)^2 + \frac{1}{4} (g^2 + g'^2) Z_\mu Z^\mu \left(v + \frac{1}{\sqrt{2}} H \right)^2}_{\text{interaction and gauge boson masses}} \\ & - \underbrace{V(H)}_{\text{Higgs potential}} \end{aligned} \quad (2.35)$$

where the first term represents the kinetic energy of the Higgs field, those involving W/Z and H are the interaction terms, and finally the terms not involving H can be interpreted as the mass terms of the gauge bosons, that is

$$m_W = \frac{1}{\sqrt{2}} v g \quad \text{and} \quad m_Z = \frac{1}{\sqrt{2}} v (g^2 + g'^2)^{1/2} \quad (2.36)$$

The value of μ in the Higgs mass term $m_H = \sqrt{-2\mu^2}$ is not known and so the Higgs mass cannot be predicted from theory. The VEV however can be predicted from the mass of the Z boson and its coupling strength as

$$v = \frac{2m_Z}{g^2 + g'^2} = 246 \text{ GeV} \quad (2.37)$$

It can be seen then, that the introduction of the complex scalar field doublet and the Higgs potential broke the $SU(2)_L \times U(1)_Y$ symmetry, resulting in a new massive spin-0 boson – the Higgs boson – while simultaneously giving mass to the weak bosons (W^\pm and Z).

Fermion Masses

The fermions are still massless at this point in the theory. Mass terms can be added to the SM while still maintaining local gauge invariance, if they are of the form

$$L_{\text{Yukawa}} = -\lambda_f \bar{\phi}_L \Psi \phi_R - \lambda_f \bar{\phi}_R \Psi \phi_L \quad (2.38)$$

where λ_f are the unknown Yukawa coupling constants for each fermion f . These Yukawa terms are added for each fermion into the SM, except for the neutrino. Expanding Φ around the ground state as in Section 2.2.4 gives

$$\mathcal{L}_{\text{Yukawa}} = - \underbrace{\frac{v\lambda_f}{\sqrt{2}} (\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L)}_{\text{fermion mass}} - \underbrace{\frac{\lambda_f}{\sqrt{2}} H (\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L)}_{\text{interaction}} \quad (2.39)$$

It can be seen here that the fermion masses depend on the VEV of the Higgs field, thus, their masses are generated through coupling to the Higgs as

$$m_f = \frac{v\lambda_f}{\sqrt{2}} \quad (2.40)$$

Similarly, the coupling strength of a fermion with the Higgs can be read off the second term, as

$$g_{Hff} = \frac{m_f}{v} \quad (2.41)$$

The coupling strength increases therefore with fermion mass. As will be clear in the next section, this fact has driven the approach to many of the experimental searches for the Higgs boson.

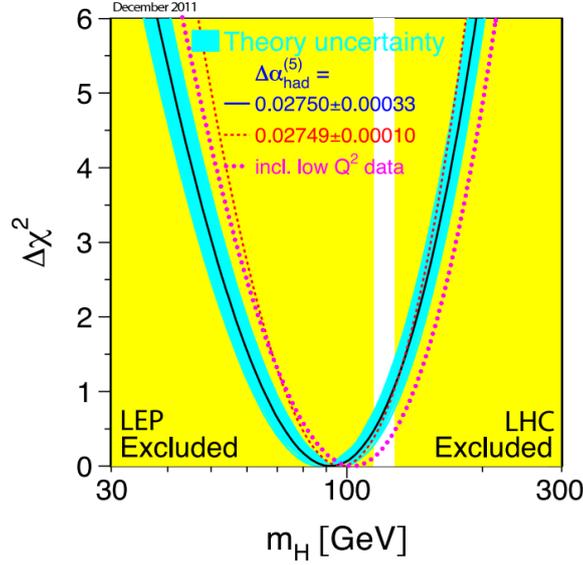


Figure 2.2.: $\Delta\chi^2 = \chi^2 - \chi_{min}^2$ vs. m_H curve. The line is the result of the fit to all electroweak precision data, with the preferred value at $m_H = 94_{-24}^{+29}$ GeV; the band represents an estimate of the theoretical error due to missing higher order corrections. The vertical band shows the 95% CL exclusion limit on m_H from the direct search. [16].

2.3. The Higgs Boson

As described in the previous section, the mass of the Higgs boson cannot be calculated from theory. For this reason, although being first proposed in 1964, it has evaded physicists for many years, until the discovery of a Higgs-like particle at the Large Hadron Collider (LHC – described in detail in Chapter 3) at CERN in July of 2012. This section describes Higgs phenomenology, and the various searches and studies undertaken to date.

While there are theoretical arguments that can be used to place bounds on m_H [15], experimental searches are needed to prove the existence of a Higgs boson, and determine its precise properties. Indirect constraints on m_H can be placed from global fits of precision electroweak measurements [16]. Figure 2.2 shows the preferred value for m_H as a function of $\Delta\chi^2$ of the fit, as well as excluded bands from direct searches at the Large Electron Positron Collider (LEP) and LHC (as of late 2011). The best fit value is $m_H = 94_{-24}^{+29}$ GeV. To claim discovery however, the Higgs must be produced and measured in a direct search. The searches at LEP, the Tevatron, and the ultimate discovery at the LHC are described below.

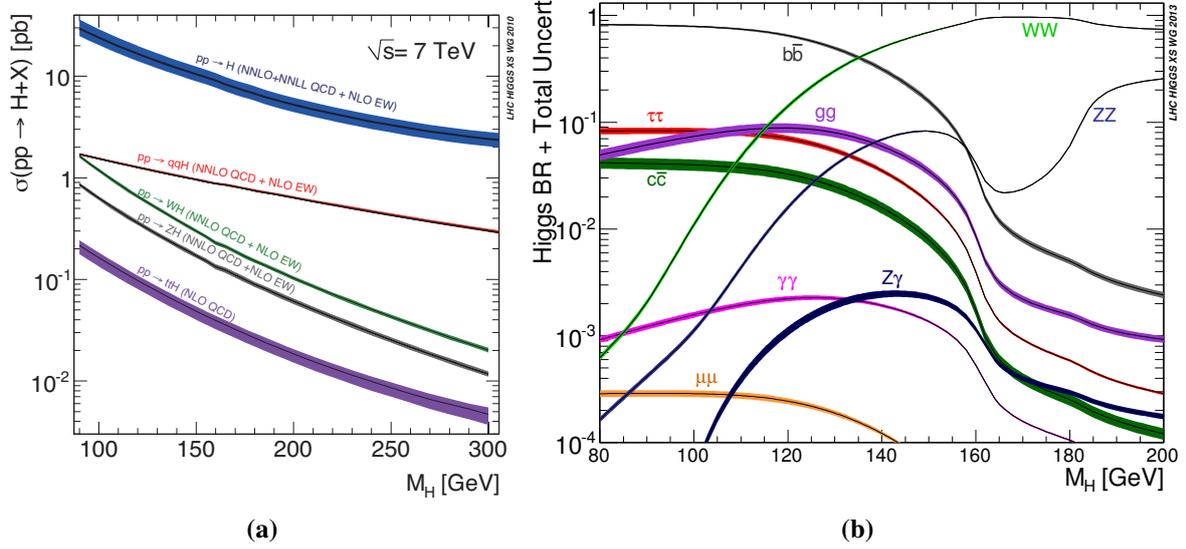


Figure 2.3.: (a) Standard Model Higgs boson production total cross sections at $\sqrt{s} = 7$ TeV [22] and (b) branching ratios and their uncertainties [23].

LEP Searches

The LEP collider was an e^+e^- collider operating at centre-of-mass energies between 91 and 209 GeV at CERN, between 1989 and 2000. Higgs searches were conducted by all four LEP detectors; ALEPH [17], OPAL [18], DELPHI [19] and L3 [20], providing a combined final lower bound on m_H at 114.4 GeV at a 95% confidence level³ [21].

At the energies involved at LEP, the dominant production mode for the Higgs was associated production (VH) with a Z boson ($e^+e^- \rightarrow HZ$). Figure 2.3b shows that a 115 GeV Higgs would decay predominantly into $b\bar{b}$ pairs (74%). A search was conducted by reconstructing the Higgs mass from the decay products and looking for an excess over the background. Constructing a likelihood ratio $Q = \mathcal{L}_{s+b}/\mathcal{L}_b$, where \mathcal{L}_{s+b} and \mathcal{L}_b are the signal plus background, and background hypotheses respectively. No excess above the background was observed (see Figure 2.4a, plots of this nature will be described in more detail in Section 4.9), and a lower limit of 114.4 GeV was set for m_H .

³The term *confidence level* will be described in Section 4.9.

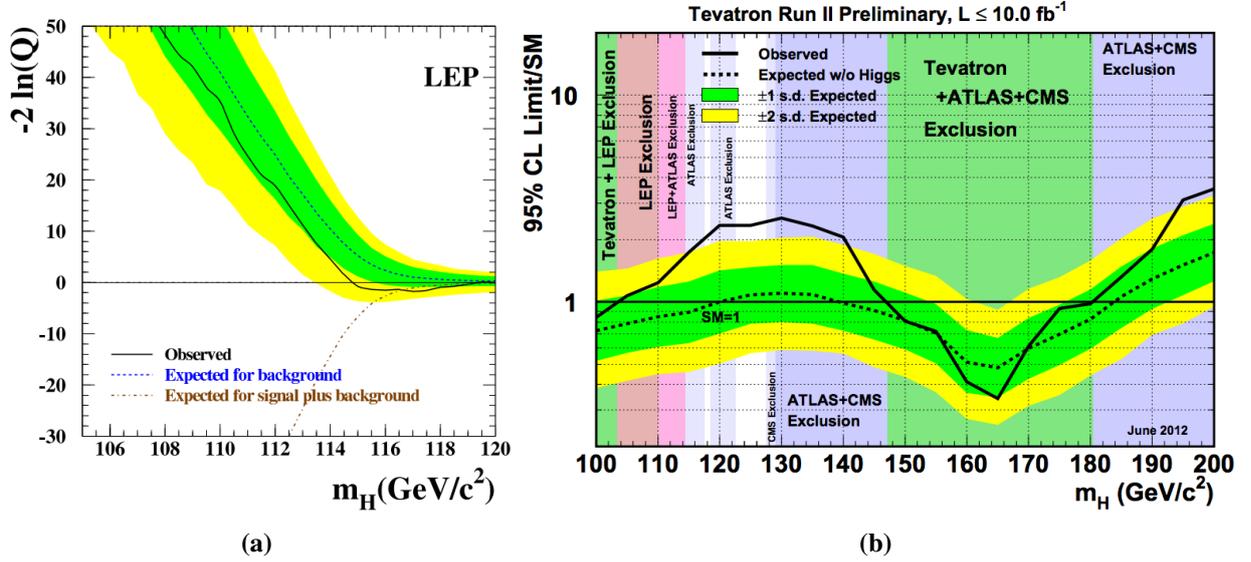


Figure 2.4.: (a) Observed and expected behaviour of the test statistic $2 \ln Q$ as a function of the test mass m_H , for the LEP experiments. The full curve represents the observation; the dashed curve shows the median background expectation; the dark and light shaded bands represent the 68% and 95% probability bands. A lower limit is placed at $m_H = 114.4$ GeV [27]. (b) Observed and expected 95% C.L. upper limits on the ratios to the SM cross section, as functions of m_H for the combined Tevatron analyses. The limits are expressed as a multiple of the SM prediction, with the bands indicating the 68% and 95% probability regions. [26].

Tevatron Searches

The Tevatron was a $p\bar{p}$ collider operating at the Fermi National Accelerator Laboratory (Fermilab) in Illinois, USA. Running from 1983 – 2011, the Tevatron collided beams at a maximum centre-of-mass energy of 1.96 TeV at two detectors – CDF [24] and DØ [25]. Together, the experiments were able to set a combined exclusion on m_H in the ranges $100 < m_H < 103$ GeV and $147 < m_H < 180$ GeV [26] at the 95% confidence level.

The main production mode at the Tevatron was via VH, gluon-gluon fusion (ggF – $gg \rightarrow H$) and vector boson fusion (VBF– $q\bar{q} \rightarrow q'\bar{q}'H$). Decays into $b\bar{b}$, W^+W^- , ZZ , $\tau^+\tau^-$ and $\gamma\gamma$ were all studied. Figure 2.4b shows the combined limit on m_H for the full Tevatron data set, with the excluded regions from LEP and LHC added.

LHC Discovery

Discussed in detail in the next chapter, the LHC is currently the world’s most powerful hadron collider. A pp machine, the LHC collided beams at a centre-of-mass energy of 7 TeV during

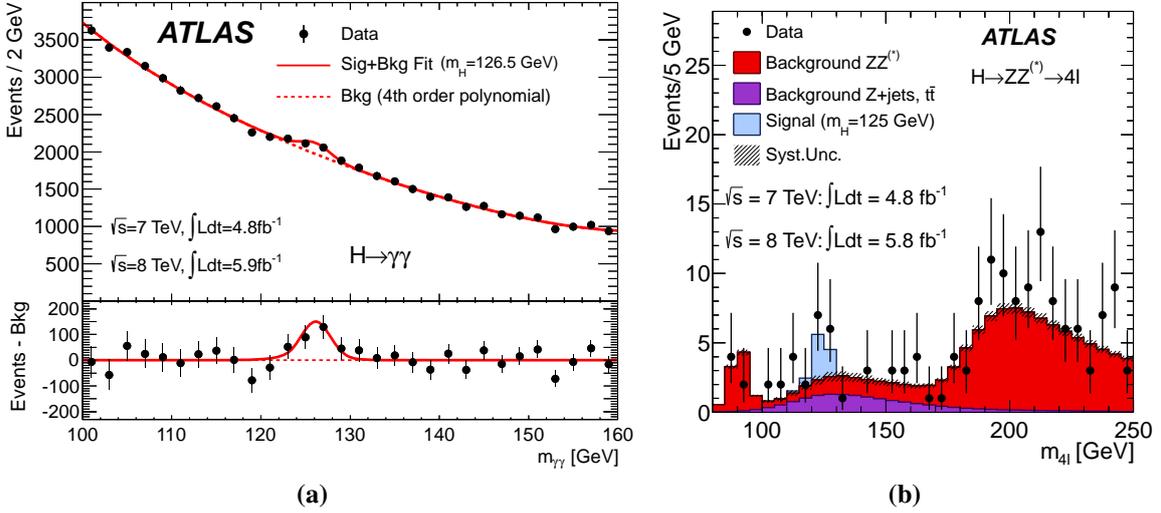


Figure 2.5.: Invariant mass distributions for the combined $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples for (a) the di-photon candidates, and (b) the four-lepton candidates. The result of a fit to the data with the signal component fixed to $m_H = 126.5$ GeV and a background component described by a fourth-order Bernstein polynomial is superimposed in the di-photon analysis. The signal expectation for a SM Higgs with $m_H = 125$ GeV is shown in the four-lepton analysis. [22]

2011, and increased this energy to 8 TeV in 2012. In July of 2012, the two general purpose detectors, ATLAS and CMS, together announced the observation of a new Higgs-like particle at the LHC at a mass of $126.0 \pm 0.4(\text{stat.}) \pm 0.4(\text{syst.})$ GeV with a significance of 5.9 standard deviations (meaning the probability of the background fluctuating upward to produce the signal is 1.7×10^{-9}) [22].

During 8 TeV running, the main production modes for the SM Higgs were gluon-gluon fusion (87.2%), vector boson fusion (7.1%), associated production with a vector boson (5.1%), and associated production with $t\bar{t}$ (0.6%), as shown in Figure 2.3a (these production modes will be described in Chapter 4). The most sensitive decay modes studied with the ATLAS detector were the $H \rightarrow \gamma\gamma$ and the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ channels. Although these channels do not have the highest branching ratios, their experimental signatures provided the best opportunity to discriminate a Higgs signal from the SM backgrounds. Figure 2.5 shows the invariant mass distributions for the $\gamma\gamma$ and 4ℓ searches with the 7 and 8 TeV ATLAS data samples, where significant excesses were seen at around 125 GeV.

Figure 2.6 shows the p_0 value as a function of m_H , as well as the combined observed signal strength μ for the ATLAS results. The p_0 value gives the probability that the background can give a random fluctuation as big as the observed data, while μ is the ratio of the observed

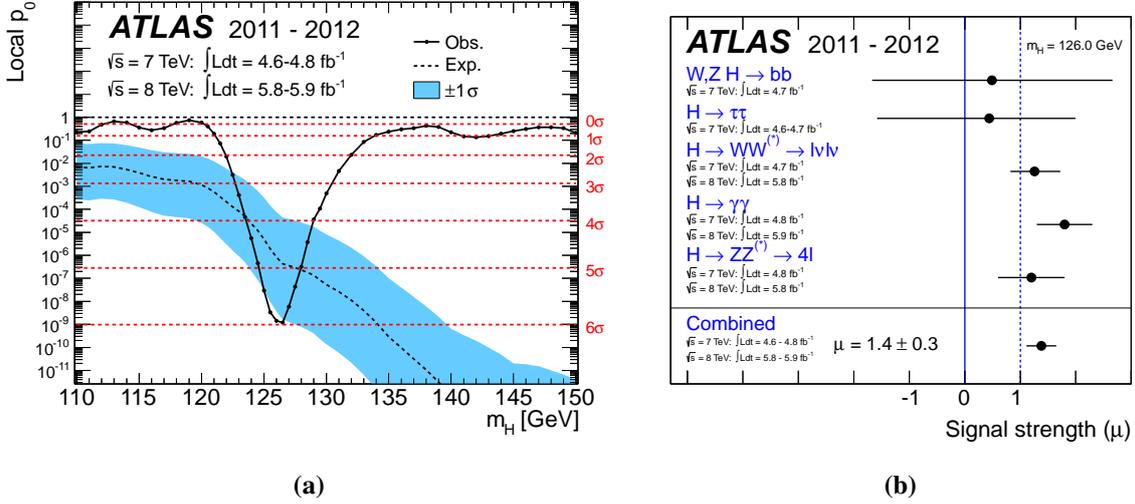


Figure 2.6.: (a) The observed (solid) p_0 as a function of m_H in the low mass range. The dashed curve shows the expected p_0 under the hypothesis of a SM Higgs boson signal at that mass with its plus/minus one sigma band. The horizontal dashed lines indicate the p-values corresponding to significances of 1 to 6 sigma. (b) Measurements of the signal strength parameter μ for $m_H = 126 \text{ GeV}$ for the individual channels and their combination. [22]

production rate to that predicted by the SM; a $\mu = 1$ is therefore consistent with the SM, while $\mu = 0$ is consistent with a background-only hypothesis. The observed excess is an almost 6σ deviation from the background, with a signal strength consistent with the SM at $\mu = 1.18^{+0.15}_{-0.14}$.

Of particular importance to the work in this thesis are the fermionic decay modes of the Higgs boson. To fully confirm that the observed new particle is *the* SM Higgs boson, and not *a* Higgs boson, all the couplings must be measured. This thesis focuses on the $H \rightarrow \tau_{lep}\tau_{had}$ (subscripts explained below) decay mode, with the work completed in parallel to a since published ATLAS search [2]. Here, the results of this dedicated $H \rightarrow \tau\tau$ search using the full 7 and 8 TeV ATLAS dataset is presented.

The analysis splits the $\tau\tau$ decays into three channels depending on the tau decay modes; $\tau_{lep}\tau_{lep}$, $\tau_{lep}\tau_{had}$ and $\tau_{had}\tau_{had}$, where τ_{lep} represents $\tau \rightarrow \ell\nu\bar{\nu}$ with $\ell = e, \mu$ and τ_{had} represents $\tau \rightarrow \text{hadrons } \nu$. The major production modes studied were VBF and VH, with their properties exploited by splitting the channels into different analysis categories; the VBF category, focusing on a high η separation of the two jets produced in association with the Higgs, and the boosted category, focusing on events with a highly boosted Higgs candidate. A multi-variate technique (an in-depth discussion of such techniques can be found in Section 5.1) was used to separate signal from background events.

The results from all channels and analysis categories were combined, to give an excess with an observed (expected) significance of 4.54σ (3.43σ) at 125.36 GeV. Figure 2.7 shows the signal strengths μ for the separate, and combined categories at this mass, with a combined signal strength of $\mu = 1.43_{-0.26}^{+0.27}(\text{stat.})_{-0.25}^{+0.32}(\text{syst.}) \pm 0.09(\text{theory})$, which is consistent with a SM Higgs boson. Detailed searches for the Higgs boson in 7 TeV ATLAS data decaying to two taus using a cut-based technique, a so-called *hybrid* technique, and multi-variate-analysis techniques are described in Chapters 4, 5, and 6 respectively.

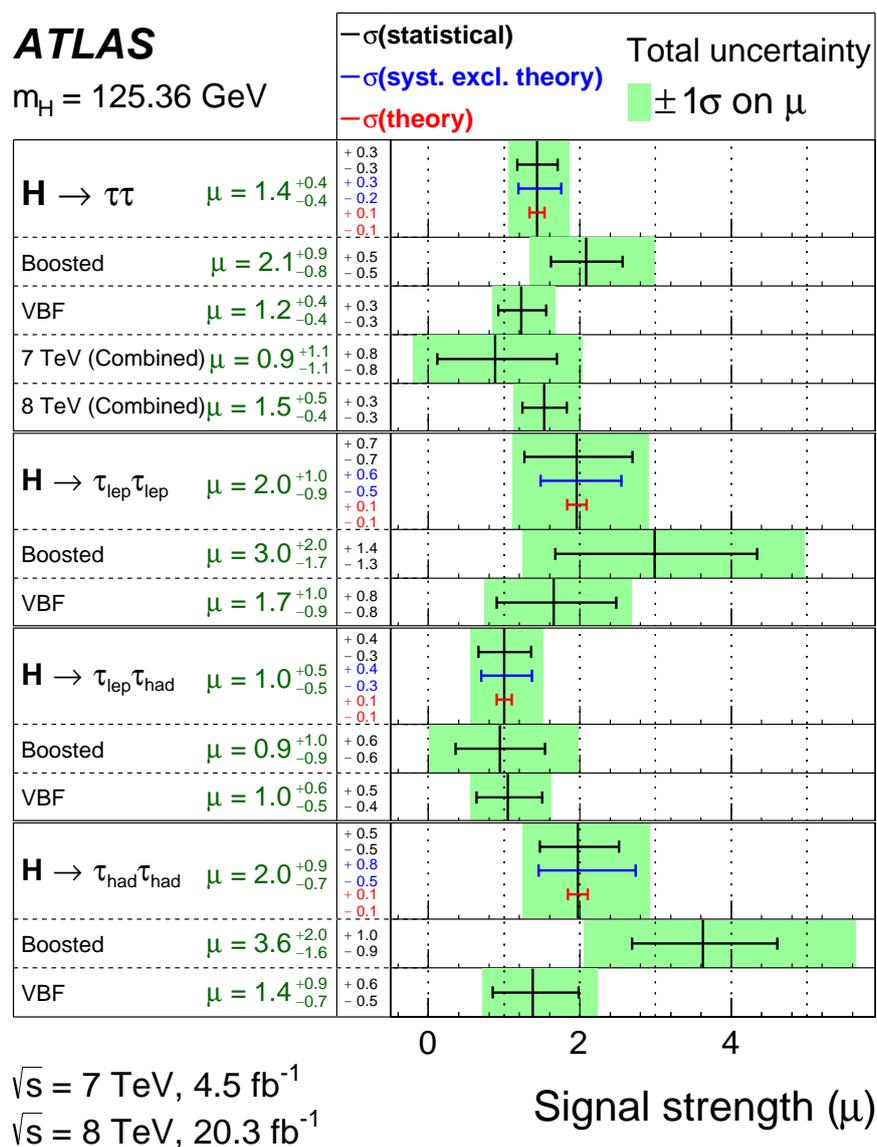


Figure 2.7.: The best-fit value for the signal strength μ in the individual channels and their combination for the full ATLAS data-sets at $\sqrt{s} = 7 \text{ TeV}$ and $\sqrt{s} = 8 \text{ TeV}$. The total $\pm 1\sigma$ uncertainty is indicated by the shaded green band, with the individual contributions from the statistical uncertainty (top, black), the experimental systematic uncertainty (middle, blue), and the theory uncertainty (bottom, red) on the signal cross section (from QCD scale, PDF, and branching ratios) shown by the error bars and printed in the central column. [2]

Chapter 3.

The ATLAS Experiment

In order to study the physics described in the previous chapter and probe the distance scales involved, extremely high energies are required. Particle accelerators are a tool to allow such experimentation, by accelerating charged particles, and colliding them within detecting materials. Based at the European Organisation for Nuclear Research (CERN) outside of Geneva, Switzerland, the Large Hadron Collider (LHC) is currently the world's most powerful proton-proton collider. Designed to collide proton beams at four points around its 27 km ring at a design energy of 14 TeV, the LHC was envisioned as a tool to help search for the (at the time) undiscovered Higgs boson, as well as to probe physics beyond the Standard Model (BSM). The ATLAS detector was designed as one of the general purpose detector on the LHC ring, tasked with making these discoveries by collecting the collision data.

3.1. The Large Hadron Collider

The design of the LHC was largely driven by the number of proton-proton interactions per-second required to produce extremely rare processes. Figure 3.1 shows the cross-sections and rates of various processes at a luminosity (defined below) of $1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ plotted against the centre of mass energy of the system (cms, \sqrt{s}). As it can be seen, as the rarity of a process increases (the cross-section decreases), higher \sqrt{s} and luminosities are required. The number of events per second seen for a given process cross-section (σ_{event}) is given by:

$$N_{event} = \mathcal{L} \sigma_{event} \tag{3.1}$$

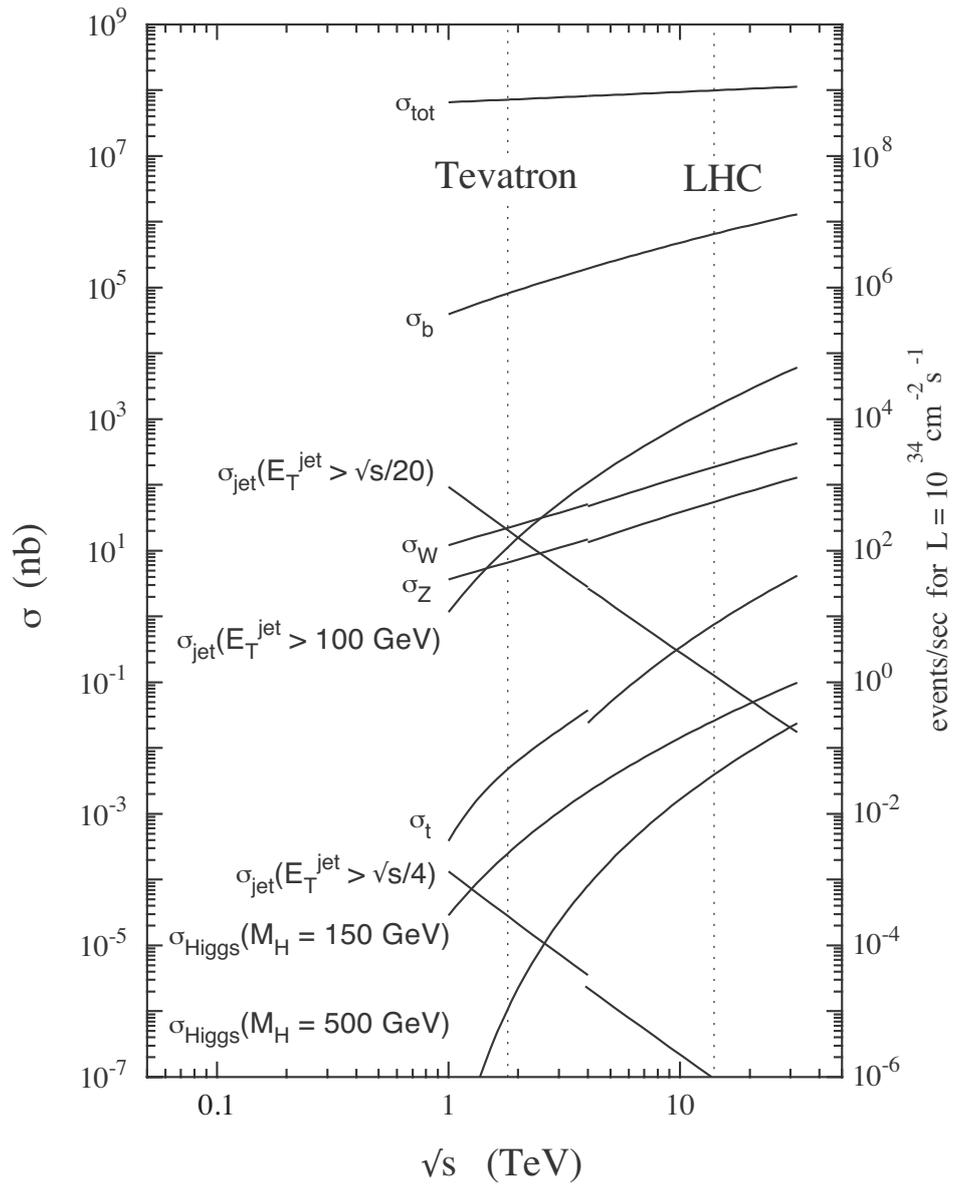


Figure 3.1.: Cross-section and event rates (for a luminosity of $1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ for various processes in proton(anti)proton collisions, as a function of the centre-of-mass energy \sqrt{s} . The dashed lines represent the design energies of the Tevatron (1.8 TeV) and the LHC (14 TeV) [28]. This thesis uses data from the 2011 7 TeV run of the LHC.

where L is the machine luminosity, defined as:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (3.2)$$

N_b and n_b are the number of particles per bunch, and bunches per beam respectively, f_{rev} is the revolution frequency of the beam, and γ_r is the relativistic gamma-factor. The parameters ϵ_n and β^* define the shape of the beam, with ϵ_n being the normalised emittance (average spread of the particles in position-momentum space, normalised to beam energy) and β^* being the optical beta-function at the interaction point (the distance from the focus point of the beam to the point at which the beam width is twice as large). The crossing angle of the beams is included in F , the geometric luminosity reduction factor, defined as:

$$F = \left[1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*} \right)^2 \right]^{-\frac{1}{2}} \quad (3.3)$$

where θ_c is the beam crossing angle, and σ_z and σ^* are the RMS bunch length and transverse RMS beam size at the interaction point. Thus, in order to achieve such a high luminosity, the LHC must have a high number of bunches, and protons per bunch, and a small ϵ_n and β^* - that is, highly squeezed bunches, with a small crossing angle at the interaction point. Table 3.1 shows the design values for these parameters, and the values as of late 2011.

3.1.1. LHC Design

The LHC was designed to occupy the previous CERN Large Electron Position (LEP) collider geometry; a 26.7 km tunnel, 70 – 100 m under the Swiss-French border. Unlike LEP however, the LHC was designed as a hadron machine, and to be the most energetic particle accelerator to date - colliding two 7 TeV proton beams at a luminosity of $1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ (or 2.8 TeV/nucleon lead ions at a luminosity of $1 \times 10^{27} \text{ cm}^{-2}\text{s}^{-1}$) at four interaction points around the ring:

- **ATLAS** (A Toroidal Lhc ApparatuS), a general purpose detector, discussed fully in section 3.2 [28].

Parameter	Nominal	2011
Proton energy [TeV]	7	3.5
Relativistic Gamma (γ_r)	7461	
Number of protons per bunch (N_b) [$\times 10^{11}$]	1.2	1.5
Number of bunches (n_b)	2808	1380
Bunch spacing [ns]	25	75 \rightarrow 50
Normalised emittance (ϵ_n) [$\mu\text{m rad}$]	3.75	1.9–2.4
Beta function (β^*) [m]	0.55	1.5 \rightarrow 1
Beam crossing angle (θ_c) [μrad]	285	240
RMS bunch length (σ_z) [cm]	7.55	
Transverse RMS beam size (σ^*) [μm]	16.7	
Instantaneous luminosity (L) [$\times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$]	1	0.4

Table 3.1.: LHC machine parameters for both design, and running as of 2011.

<https://acc-stats.web.cern.ch/acc-stats/>

- **CMS** (Compact Muon Solenoid), a general purpose detector, characterised by a large 4 T superconducting solenoid, with the inner detector inside the bore of the magnet, and the muon system inside the return yoke [29].
- **LHCb**, an experiment dedicated to heavy flavour (b) physics, designed to look for indirect evidence of new physics in CP violation and rare decays of bottom and charm hadrons [30].
- **ALICE** (A Large Ion Collider Experiment), a general purpose heavy-ion detector, designed to study lead-lead collisions in order to address the physics of strongly interacting matter and quark-gluon plasma [31].

The LHC ring consists of straight sections and 8 arc sections. The straight sections are approximately 528 m in length, containing the experiments and utility insertions, and are labelled Point x (Px), with x being the octant number. P1 and P5 house the high luminosity experiments ATLAS and CMS. LHCb and ALICE are housed in P2 and P8, where the beams 1 and 2 are also injected into the LHC (see Figure 3.2 and Section 3.1.2). The remaining straight sections contain collimation systems (P3 and P7), radio-frequency (RF) systems for accelerating the beams in P4, and a beam dump insertion where the beams can be extracted from the machine at P6.

The main RF accelerating system (ACS) is made of two 400 MHz superconducting systems (one for each beam), each with 8 cavities. The RF cavities generate a longitudinally oscillating voltage across a vacuum gap. As a proton arrives at the cavity it sees an accelerating voltage, which then oscillates to a decelerating voltage as it moves away. The frequency of the RF cavities (f_{RF}) is related to the revolution frequency of the proton beams by $f_{RF} = hf_{rev}$, where h is the harmonic number, which here gives the number of *buckets* that can be filled with bunches in the LHC ($h = 35640$ for the LHC). Protons with too high or low energy will arrive late or early at the RF cavity, and as such will oscillate around the nominal energy proton (a synchronous particle), in synchrotron oscillations. The length of these oscillations is related to the beam emittance.

The magnets of the LHC are in the arc sections, each made from 23 *cells*, which themselves are two 106.9 m long half cells, containing short straight sections (SSS) between three 14.3 m dipole magnets. The magnets are NbTi superconductors, cooled to 1.9 K using superfluid helium, capable of producing a magnetic field of 8.33 T. The dipole magnets contain two coils wound in opposite directions around rings 1 and 2, to bend the proton beams as they are accelerated around the ring. In addition to the dipoles, quadrupole magnets are used to focus and de-focus the proton beams, while sextupole and octupole magnets are used for beam corrections. For example; chromaticity sextupole magnets are used for tuning ellipticity of the beam cross-section, and Landau damping octupoles for correcting coherent oscillations due to collective beam effects.

3.1.2. Injection chain

Figure 3.2 shows a schematic of the CERN site layout, and the accelerators used to feed the LHC. The protons collided in the LHC start off as hydrogen molecules in a bottle, that are then stripped of their electrons and fired towards the first accelerator, LINAC2. LINAC2 is a linear accelerator that uses electric fields to accelerate the protons, pulsed at one bunch per second, to exit at 50 MeV into the Proton Synchrotron Booster (PSB). The PSB is a circular accelerator containing four rings. Protons enter with one bunch per ring, and are accelerated to 1.4 GeV. The protons then exit into the Proton Synchrotron (PS), which takes 6 bunches from the PSB, splitting them into smaller bunches depending on the LHC filling scheme. Bunches are grouped into *trains*, with 36 bunches per train for a 50 ns LHC filling scheme. The protons are kicked from the PS with 26 GeV of energy and enter the Super Proton Synchrotron (SPS). Bunch trains are combined together in the SPS, which accelerates the protons to an LHC injection energy of 450 GeV. Finally, the bunch trains are injected into the LHC, directed along either

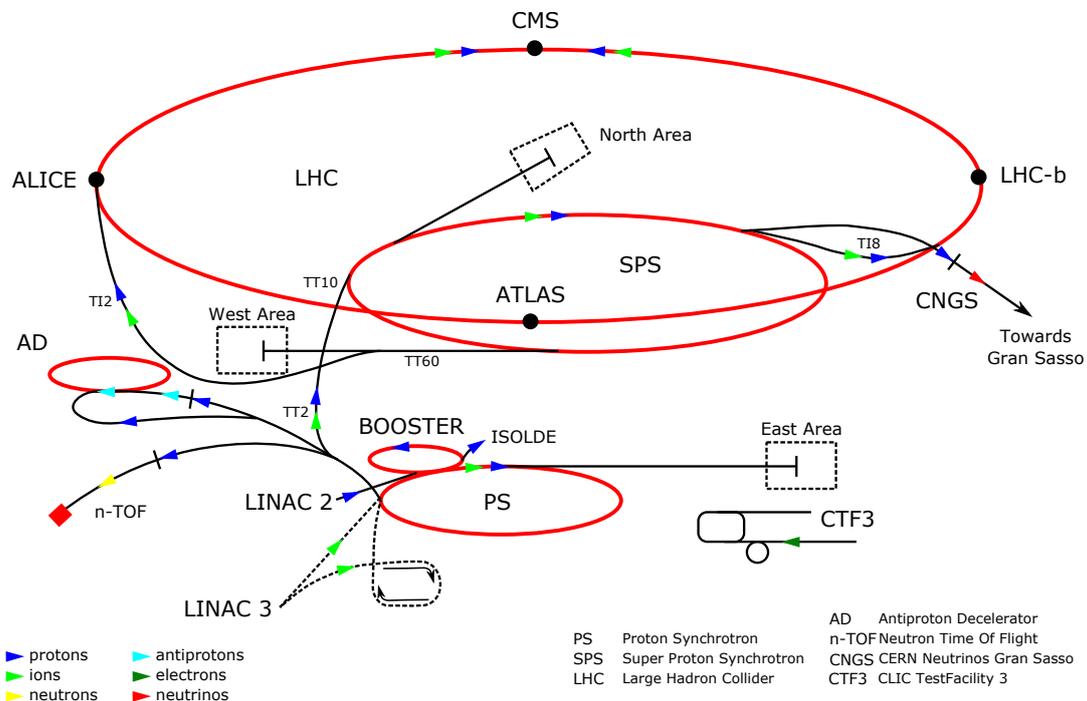


Figure 3.2.: Schematic of the CERN collider complex, including the LHC and the position of the major experiments. [32]

ring. The largest trains are injected first, and then smaller ones are injected to fill the gaps. The beams are accelerated around the ring to a collision energy of 7 TeV each (6.5 TeV as of 2013 running) until *stable beams* are announced, when collision and data taking can begin.

3.2. The ATLAS Detector

The design of the ATLAS detector was largely influenced by the physics goals of the experiment. The main requirements of the detector included;

- **Radiation hard electronics, with high granularity** were required due to the extremely high luminosity provided by the LHC. Detector components along the beam-pipe experience huge particle fluxes, and so need to be able to not only be robust enough to withstand extended running, but also accurate enough to discriminate between very close particles and identify multiple interaction vertices within a single bunch crossing.
- **High charged-particle momentum resolution** around the beam-pipe was necessary for the detection and accurate reconstruction of high energy leptons. Such leptons are not only produced from many important SM processes, such as t -quark decay and associated

production of Higgs bosons with t -quarks or W and Z bosons, but also many BSM particles, such as W' and Z' . Vertices displaced from the primary interaction vertex also need to be identified to be able to tag b -quark jets, produced from t -quark decays, or in decays from BSM particles such as a charged Higgs boson in the Minimal Supersymmetric Standard Model (MSSM).

- **High resolution electromagnetic calorimetry** was essential for SM Higgs searches. A light SM Higgs boson decaying into two photons would be the easiest to discern from the large multi-jet backgrounds, and so accurate photon energy measurements were needed.
- **Fully hermetic hadronic calorimetry** was also required, as many important physics processes, such as a Higgs boson produced by vector boson fusion, result in very forward jets as decay products. Also, by measuring the total energy deposited in the detector, and assuming that the incoming protons had zero net momentum transverse to the beam-pipe, the *missing transverse energy*¹ (\cancel{E}_T) of an event can be calculated. Events with high \cancel{E}_T are characteristic of much BSM physics, such as in Supersymmetric (SUSY) models with stable Lightest Supersymmetric Particles (LSP), and Extra-Dimension models where gravitons can be produced and escape into the extra dimensions.
- **High momentum resolution muon detection** for physics processes involving high momentum leptons. As muons will generally not deposit much energy within the calorimetry, a system to ensure high quality muon reconstruction was a must.
- **Low threshold trigger**² **systems** were needed to be able to select interesting events with low energy decay products from within the large QCD background, with a high efficiency.

With the above requirements in mind, the ATLAS detector was designed to consist of a solenoid magnet surrounding an inner detector consisting of pixel detectors, silicon strips, and transition radiation trackers. Around the inner detector are electromagnetic and hadronic calorimeters, and a muon spectrometer (MS) with superconducting toroid magnets. The detector was built with a forward-backward symmetry, with the beam-pipe being labelled the z -axis, and the $x - y$ -plane being transverse to the beam-pipe. A (η, ϕ) coordinate system is used, with ϕ being the azimuthal angle measured around the beam-axis, and the pseudorapidity η is defined as $\eta = -\ln \tan(\theta/2)$, where θ is the polar angle from the beam-axis. Differences in pseudorapidity are Lorentz invariant for high momentum particles. Distances between particles are given in terms of ΔR , defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

¹The term missing transverse energy will be described in Section 3.3.3.

²The ATLAS trigger systems are described in Section 3.2.5.

As well as influencing the detector components, the physics goals of ATLAS also defined how the data acquisition and *trigger* systems needed to perform. Due to the high luminosity of the LHC, it was not possible to store all the detector readouts for every bunch crossing. Rather, a trigger system was developed to record only those events that may contain interesting physics, such as high missing energy or high momentum leptons. The trigger system was defined to have three levels – Level 1 (L1), Level 2 (L2) and the Event Filter (EF) – with each level performing a more complicated event selection, and lowering the bandwidth until a manageable data rate is obtained. The ATLAS trigger systems are explained in detail in Section 3.2.5.

3.2.1. Magnet System

The ATLAS magnet system consists of four large superconducting magnets, designed to impart charged particles with a Lorentz force as they traverse the detector, in order to measure their momentum. The system includes a 2 T solenoidal magnet providing an axial magnetic field for the inner detector, and a 0.5 T toroidal barrel magnet and two 1 T toroidal end cap magnets providing toroidal magnetic fields for the muon detectors. The central solenoid lies in between the inner detector and the calorimetry, and as such was designed with the aim of keeping material thickness to a minimum (approximately 0.66 radiation lengths). The magnet is made from a coil of NbTi conductor, extending 5.8 m, with the flux returning through the hadronic calorimetry.

The barrel toroid makes up much of the volume of the MS, consisting of 8 coils of Nb/Ti/Cu conductor positioned radially to the beam direction. The system extends 25.3 m in length, with an outer diameter of 20.1 m. Finally, the end-cap toroids are made using the same technology as the barrel toroid, however the bending power is optimised for the high particle flux region of the muon end-caps. For a further description of the toroidal magnet system, see section 3.2.4.

3.2.2. Inner Detector

The ATLAS Inner Detector (ID) was designed to deliver precise pattern recognition for track reconstruction with high momentum resolution, and accurate vertex measurements for particles with $p_T > 100 - 500$ GeV, within $|\eta| < 2.5$. The ID consists of three sub-detectors as shown in Figure 3.3; silicon pixel detectors closest to the beam-pipe, surrounded by semiconducting silicon microstrip (SCT) detectors, and finally a transition radiation tracker (TRT) at high radii.

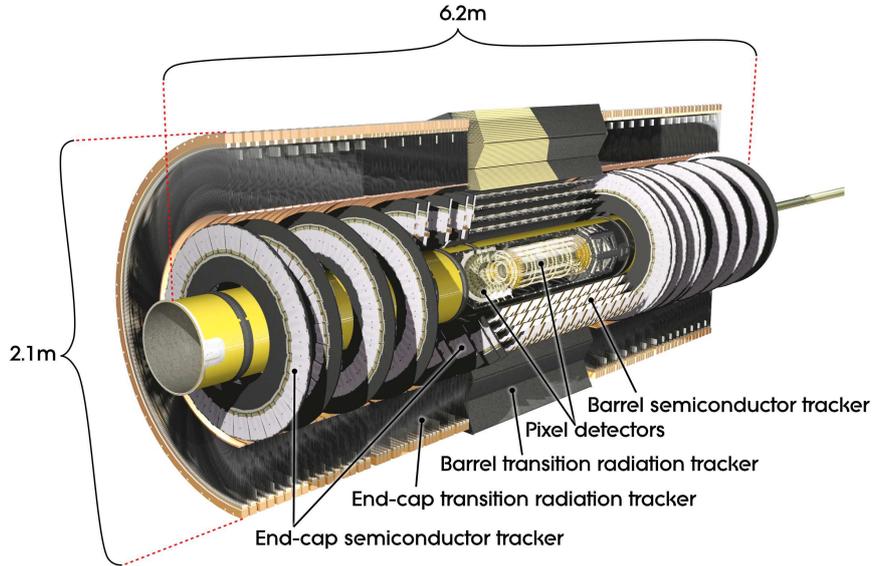


Figure 3.3.: Cut-away of the ATLAS Inner Detector. [32]

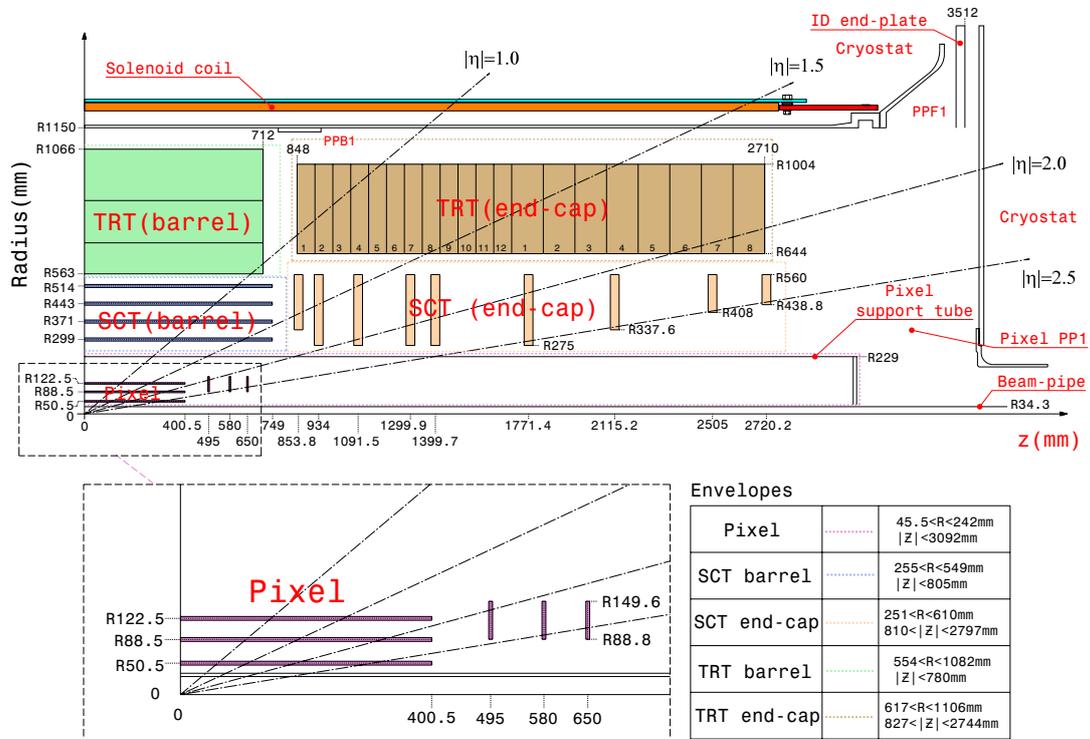


Figure 3.4.: Plan view of a quarter-section of the ATLAS inner detector showing each of the major detector elements with its active dimensions and envelopes. [28]

The ATLAS pixel system extends from a radius of 45.5 mm out to 242 mm from the nominal interaction point, arranged in three cylindrical layers within the barrel, and two sets of 3 discs at the end-caps, as shown in Figure 3.4. The modules themselves are 250 μm thick silicon detectors, designed to be radiation hard (due to their close proximity to the interaction point), and work well in the high occupancy (occupancy here referring to many particles striking the sensitive electronics) environment. There are 1744 sensors, each with 46080 readout channels (resulting in over 80 million channels over the entire pixel detector), operating in a temperature range of -5°C to -10°C . The pixel detectors typically provide three measurement points with a spatial resolution of 12 μm for particles originating from the interaction region, reading out a hit if the signal exceeds a predefined threshold.

The SCT system extends from a radius of 255 mm to 610 mm in the barrel, and consists of 4088 sensor modules arranged in 4 coaxial cylindrical layers in the barrel, and two end-caps of 9 wheels each (as shown in Figure 3.4), creating an almost hermetic coverage in the ID. The barrel modules are made from four sensors, each consisting of 788 $285 \pm 15 \mu\text{m}$ thick p-in-n silicon strips with a pitch (distance between neighbouring strips) of 80 μm , arranged in two layers (two each on the top and bottom), with a relative angle between the layers of ± 20 mrad (to provide accurate space-point resolution in $R - \theta$). The end-cap modules are similar to the barrel design, each with two sets of back-to-back sensors again with a relative rotation of ± 20 mrad. The spatial resolution at normal incidence for the SCT modules is $\approx 16 \mu\text{m}$, and at least a 4 space-point measurement is made for a particle emerging from the interaction region.

The TRT barrel is also shown in Figure 3.4, and consists of three rings of 32 modules of polyimide drift (straw) tubes interlaced with polypropylene fibres used as transition radiation material. The straw tubes are 72 cm long, and filled with a gas mixture of 70% Xe, 27% CO₂, and 3% O₂, with a tungsten wire anode through the centre. Electrons passing through the TRT will interact with the polypropylene fibres creating transition radiation photons, which are absorbed by the Xe-based gas with a much larger signal yield than other traversing particles. The signal is then read out at both ends of the wire anode, with hits from transition radiation being separated from others as high-threshold hits by the front-end electronics. Particles with $p_T > 0.5 \text{ GeV}$ and within $|\eta| < 2.0$ will traverse at least 36 straws, with electrons with $p_T > 2 \text{ GeV}$ producing approximately 7 high-threshold hits from the transition radiation. The TRT end-cap contains two independent wheels, with 12 near the interaction point, and a further 8 at higher radii. The transition radiation material here is polypropylene foils between the straw layers of 768 radially orientated straws. The resolution of the TRT is limited by the inter-straw spacing, with a $R - \phi$ resolution of 130 μm .

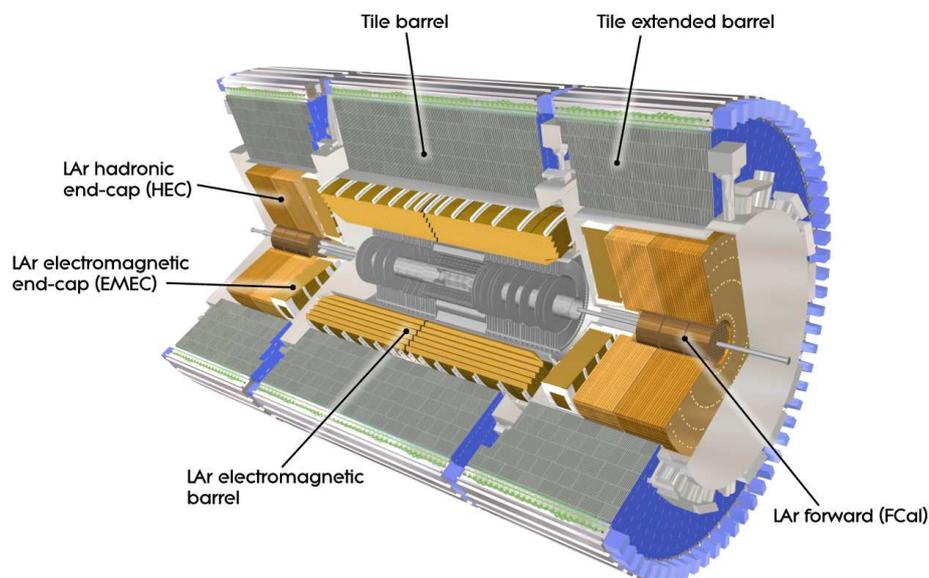
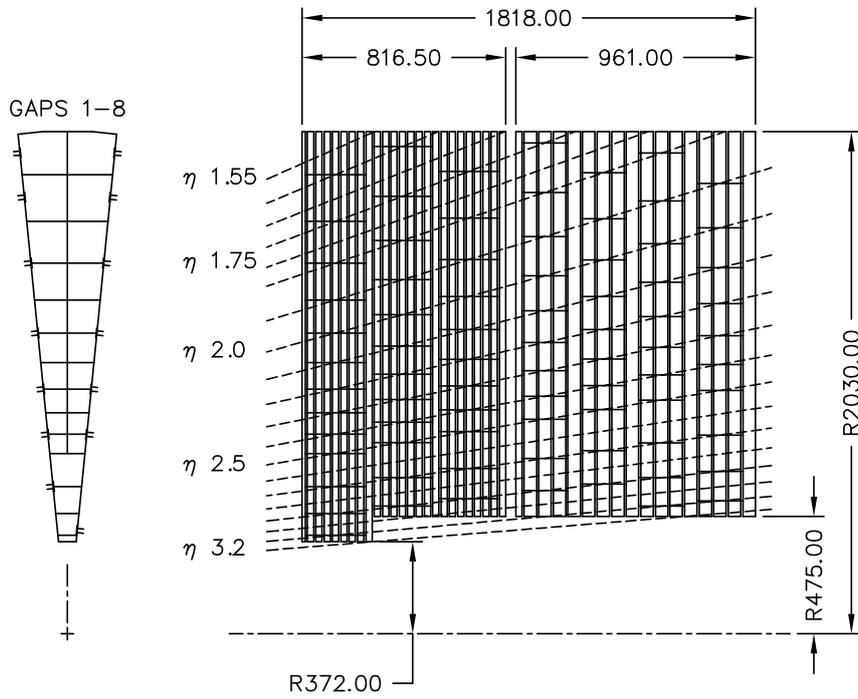


Figure 3.5.: Cut-away view of the ATLAS calorimeter system. The overall dimensions of the system are approximately 8 m in diameter, and 12 m in length. [32]

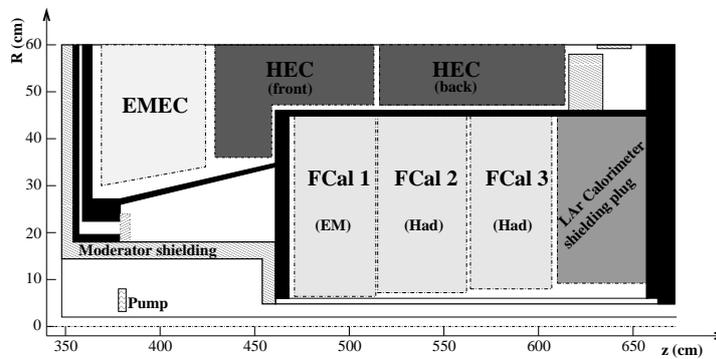
3.2.3. Calorimetry

The ATLAS calorimetry extends to $|\eta| < 4.9$, consisting of a fine granularity electromagnetic (EM) calorimeter for identifying electrons and photons, surrounded by a coarser granularity hadronic (HAD) calorimeter for jet and missing transverse energy detection, and is shown in Figure 3.5. The calorimeters are not only for containing electromagnetic and hadronic jet showers, but also limiting the punch-through to the muon system (described in section 3.2.4). The EM calorimeter therefore has a depth of > 22 radiation lengths (X_0).

The EM calorimetry is made from a liquid argon (LAr) detector, with lead absorption material and read-out electrodes. The LAr barrel extends from $|\eta| < 1.475$ (with a small gap at $z = 0$), and the LAr end-caps (EMEC) from $1.375 < |\eta| < 3.2$. The EM calorimetry is an accordion design, running axially in ϕ , to provide complete ϕ symmetry without any azimuthal cracks, with the size of the folds varied with η to maintain a constant LAr gap size (the thickness of the LAr gaps are optimised for EM performance). Within the central region of the detector ($|\eta| < 2.5$ - the extent of the ID), the LAr calorimeter is 3 layers in depth, with a fourth presampler layer in front of it within $|\eta| < 1.8$ to correct for energy lost within the ID, made from active LAr. Figure 3.6a shows a sketch of a LAr barrel module. The granularity of the EM barrel is highest in the first LAr layer ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$) to provide position information for EM showers, and decreases in the outer layers and with high



(a)



(b)

Figure 3.7.: (a) Schematic $R - \phi$ (left) and $R - z$ (right) views of the hadronic end-cap calorimeter. The semi-pointing layout of the readout electrodes is indicated by the dashed lines. Dimensions are in mm. (b) Schematic diagram showing the three FCal modules located in the end-cap cryostat. The material in front of the FCal and the shielding plug behind it are also shown. The black regions are structural parts of the cryostat. The diagram has an expanded vertical scale for clarity. [28]

$|\eta|$. The EMECs are similar in design to the barrel calorimeters, and consist of two wheels on either side of the EM barrel. An end-cap presampler also exists here within the precision region of $1.5 < |\eta| < 1.8$. Together, the EM barrel and EMECs provide over 237,000 readout channels.

The HAD barrel calorimetry is a sampling tile calorimeter, housed directly outside the EM barrel calorimeter envelope. It is split into barrel modules ($|\eta| < 1$) and extended barrel modules ($0.8 < |\eta| < 1.7$). Steel absorbers are placed between scintillating tiles as the active material, with light collected at their edges into waveguide fibres, and sent to photomultiplier tubes to read out the signal. There are three layers of calorimetry, with differing interaction lengths in the barrel and extended barrel. The fibres are grouped according to $\eta \times \phi = 0.1 \times 0.1$ cells for the first two layers, and 0.2×0.1 in the third.

To provide hadronic identification in the forward regions, the HAD barrel has a LAr hadronic end-cap (HEC) on either side, consisting of two wheels (HEC1 and HEC2) per end-cap, directly behind the EMEC. To ensure full coverage without gaps, the HEC overlaps with the HAD barrel and the EMEC, with a range of $1.5 < |\eta| < 3.2$. The wheels consist of 32 wedges, with those closest to the interaction point (HEC1) containing 25 mm copper plates, interlaced with a LAr active sampling medium, and 50 mm copper plates in HEC2 (as shown in Figure 3.7a). The readout cell size for the ID region of $|\eta| < 2.5$ is $\eta \times \phi = 0.1 \times 0.1$, with 0.2×0.2 for larger η . Besides containing hadronic showers, the HEC is also used for measuring the radiative energy loss of muons before they reach the MS.

The final piece of hadronic calorimetry is the LAr forward calorimeter (FCal). The FCal is designed to detect extremely forward particles with very high flux, with an extent of $3.1 < |\eta| < 4.9$. The detector is therefore 10 interaction lengths deep, and limits particles other than muons reaching the MS. The FCal is integrated into the end-cap cryostats, and each end-cap is divided into three modules, as shown in Figure 3.7b. The first module layer (FCal1) is made with copper electrodes, and is optimised for electromagnetic showers. FCal2 and FCal3 use tungsten electrodes, and are designed for absorption of hadrons. The design of the detector uses concentric electrode rods and tubes parallel to the beam axis, with active LAr in between. This design limits the size of the LAr gaps, helping to deal with ion buildup and provide a high density for the absorbing material.

3.2.4. Muon Spectrometer

The ATLAS MS was designed to detect muons (and other charged minimum ionising particles) that pass through the calorimetry, within the range $|\eta| < 2.7$, and to trigger on such particles within $|\eta| < 2.4$. The structure consists of large, superconducting air-cooled toroid magnets to provide a magnetic field orthogonal to the muon trajectory, with separate high-precision tracking chambers and trigger chambers, as shown in Figure 3.8. The high-precision tracking

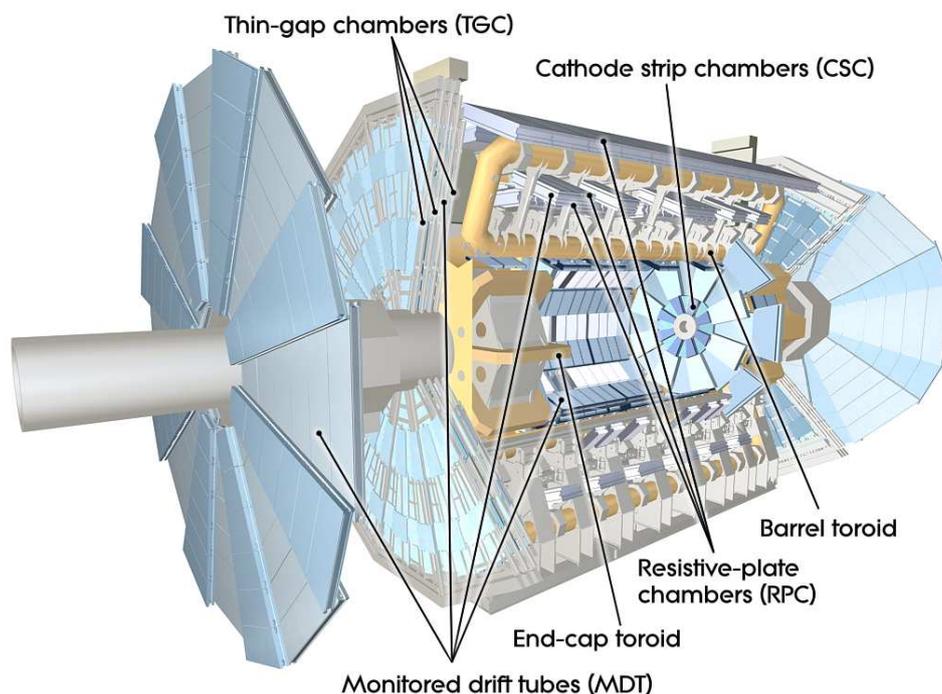


Figure 3.8.: Cut-away view of the ATLAS muon system. Total dimensions of the system are approximately 25 m in diameter, and 44 m in length. [32]

chambers consist of Monitored Drift Tubes (MDTs) and Cathode Strip Chambers (CSCs), while the triggering is performed by Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs).

The muon magnet system consists of three toroid barrel magnets, each containing 8 coils, positioned radially and symmetrically around the beam axis. Two end-cap toroids are inserted within the barrel toroid at each end. The bending power of the magnets (measured by $\int Bdl$, where B is the field in the component normal to the muon direction), is higher in the forward regions where the particle flux is largest, with a power of 1.0–7.5 Tm in the region $1.6 < |\eta| < 2.7$, and 1.5–5.5 Tm in the region $0 < |\eta| < 1.4$.

Precision tracking is performed by MDTs in the region $|\eta| < 2.7$, except for the innermost layer of the end-cap, where it is performed by the CSCs ($2 < |\eta| < 2.7$). Three concentric shells of chambers are positioned within the barrel toroid around the beam axis, and four wheels in the end-cap regions, as outlined in Figures 3.9a and 3.9b. The MDTs consist of 3-8 layers of pressurised drift tubes, filled with Ar/CO₂ gas, and a tungsten-rhenium wire for collecting electrons ionised by the traversing muon, and reading out the signal. This setup gives an average resolution of 35 μm per chamber, driven by the main physics goal of the detector to provide a p_T resolution of 10% for 1 TeV tracks (see Table 3.2).

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	± 2.5	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry (jets)			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \eta < 4.9$	$3.1 < \eta < 4.9$
Muon Spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 3.2.: General performance goals of the ATLAS detector. Note that, for high- p_T muons, the muon-spectrometer performance is independent of the inner-detector system. [33]

MDTs cannot be used in the high flux region of the end-caps, due to their safe running limit of approximately 150 Hz/cm^2 . CSCs however have a safe running limit of around 1000 Hz/cm^2 , and so make up the innermost layer of the end-cap chambers in $|\eta| > 2$. As with the end-cap MDTs, the CSCs are segmented in ϕ , with 8 chambers per disk. Each CSC has multiple wires extending radially, which induce charge in cathode strips, running both perpendicular to the wires to provide the precision coordinate, and parallel for the transverse coordinate. The CSCs can therefore simultaneously measure η and ϕ of the track. Although the CSCs provide fewer measurements along the track compared to the MDTs, this is made up for by the higher rate capacity and time resolution.

The second function of the MS is to provide fast information on muon tracks traversing the detector, and to then send the information on directions, multiplicities, and energies to the L1 trigger, as well as providing bunch-crossing identification (BCID) and being robust to background noise. The muon trigger system extends to $|\eta| < 2.4$, and due to higher particle fluxes in the forward regions, an η -dependent granularity is required to maintain p_T resolution across the detector. The design is therefore RPCs in the barrel ($|\eta| < 1.05$) and TGCs in the end-caps ($1.05 < |\eta| < 2.40$), with a requirement for a coincidence of hits in both types of chambers for the trigger to fire. RPCs consist of parallel electrode plates, filled with a gas of $\text{C}_2\text{H}_2\text{F}_4/\text{Iso-C}_4\text{H}_{10}/\text{SF}_6$, where avalanches form along the ionising tracks towards the anode. They are arranged in three concentric layers around the beam, referred to as trigger stations. The TGCs are multi-wire chambers filled with CO_2 and $n\text{-C}_5\text{H}_{12}$, and are characterised by their wire-to-cathode distance (1.4 mm) being smaller than their wire-to-wire distance (1.8 mm). As well as providing high granularity for muon p_T triggering, the TGCs also provide ϕ -coordinate measurements for the MDTs. TGCs are mounted in two concentric rings for the outer end-cap

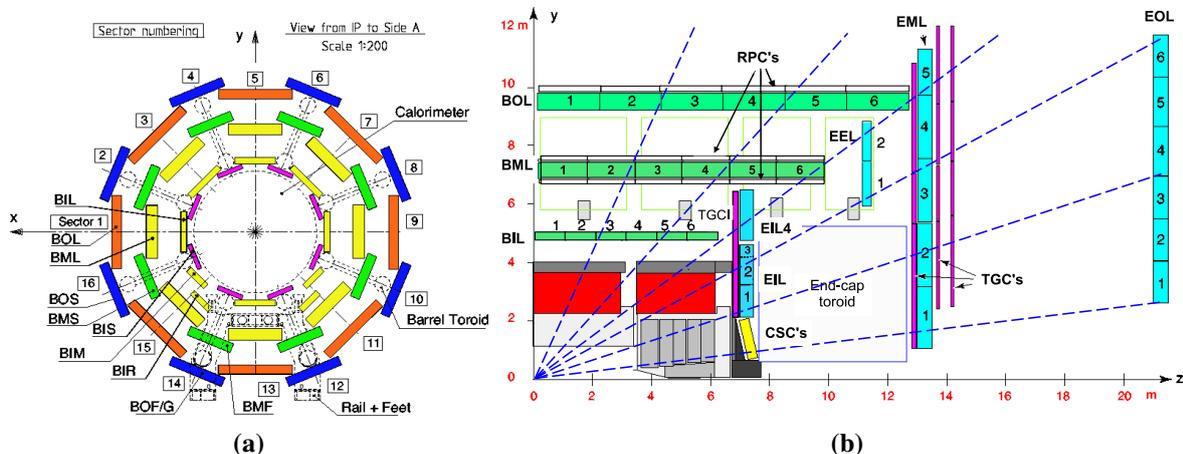


Figure 3.9.: (a) Cross-section of the barrel muon system perpendicular to the beam axis (non-bending plane), showing three concentric cylindrical layers of eight large and eight small chambers. The outer diameter is about 20m. (b) Cross-section of the muon system in a plane containing the beam axis (bending plane). Infinite-momentum muons would propagate along straight trajectories which are illustrated by the dashed lines and typically traverse three muon stations. [28]

($1.05 < |\eta| < 1.92$) and the inner or forward ring ($1.92 < |\eta| < 2.40$). Section 3.2.5 describes in more detail the layout and function of the Level 1 muon trigger.

3.2.5. Trigger System

At the design filling scheme of 2808 bunches and a 25 ns spacing, the bunch-crossing rate in the LHC is approximately 40 MHz. ATLAS however, is only capable of recording data at approximately 200 Hz. The majority of the collision events are *minimum-bias* - usually soft partonic scatters of the proton-proton collisions, referred to as the *underlying event* - which are important to study for physics reasons such as improving the understanding of QCD effects, total cross-sections, and so on. These events are also used for understanding the experiment in terms of pile-up, occupancy, and beam backgrounds. Interesting rare and new physics processes however will be completely swamped by minimum-bias events, and so the ATLAS triggers were designed to make decisions on an event-by-event basis on whether the event was worth writing to disk. The trigger system is made up of the hardware based L1 trigger, and the software based High Level Trigger (HLT), consisting of the L2 and EF, and the Data Acquisition System (DAQ).

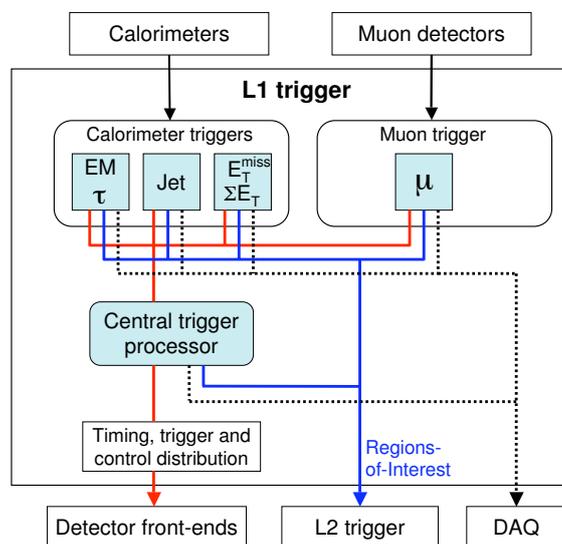


Figure 3.10.: Block diagram of the L1 trigger. The overall L1 accept decision is made by the central trigger processor, taking input from calorimeter and muon trigger results. The paths to the detector front-ends, L2 trigger, and data acquisition system are shown from left to right in red, blue and black, respectively. [28]

Level 1

The L1 trigger has three main components; the L1 Calorimeter Trigger (L1Calo), the L1 muon (tracking) trigger, and the Central Trigger Processor (CTP). The CTP combines information gathered from the L1Calo and the L1 muon trigger, and makes an overall L1 accept decision, L1A (the logical OR of all trigger items). The flow of the L1 trigger is shown in Figure 3.10. The maximum L1A rate that can be handled by the readout systems is 75 kHz, with the decision reaching the front-end electronics within $2.5 \mu\text{s}$ after the bunch-crossing.

L1Calo is a system of approximately 7000 analogue trigger towers (of 0.1×0.1 in $\Delta\eta \times \Delta\phi$) in the electromagnetic (EM) and hadronic (HAD) calorimeters. The purpose of L1Calo is to identify high transverse energy (E_T) electrons/photons, jets, and taus, with potential isolation requirements (a minimum object separation in the detector), and events with high scalar transverse energy sums (ΣE_T) or missing transverse energy. The architecture of L1Calo is shown in Figure 3.11; a pre-processor digitises the analogue input signal, and associates it to a bunch crossing (this association is very important, as the pulse width of the calorimeter signals can be several times larger than the bunch crossing interval). The pre-processor uses a look-up table to perform a pedestal subtraction, set a noise threshold, calibrates the E_T of the pulse and sends this information to the Cluster Processor (CP), and sums the trigger towers into 0.2×0.2 jet elements for the Jet/Energy Sum Processor (JEP).

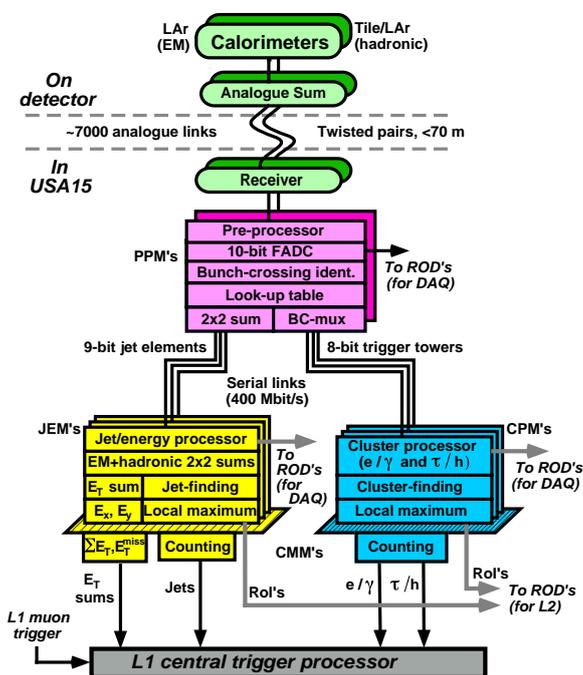


Figure 3.11.: Architecture of the L1 calorimeter trigger. [28]

The CP searches for electrons/photons and taus by looking for maxima in the EM and HAD trigger towers' E_T using overlapping, sliding windows [34], and sends this information to the CTP every bunch crossing. Similarly to the CP, the JEP takes jet trigger elements, and calculates E_T sums within overlapping windows. The energy summation algorithm sums E_x , E_y and E_z of the event and reports the total E_T and \cancel{E}_T to the CTP. In both the CP and the JEP, the 2×2 local maximum of the trigger elements defines the coordinates of the Region of Interest (RoI) to be sent to L2.

The L1 muon trigger consists of finely segmented RPC and TGC detectors (see Section 3.2.4), with fast and accurate timing to select the bunch-crossing producing the muon candidate. The basic algorithm governing the trigger is to look for a coincidence of hits in different trigger stations along the track, or the road of the muon. RPC trigger chambers are used for the muon barrel trigger ($|\eta| < 1.05$), while the TGC detectors are used for the muon end-cap trigger ($1.05 < |\eta| < 2.4$).

The barrel muon trigger consists of three stations of (doublets of) RPC detectors (schema is shown in Figure 3.12a). A muon is selected by finding a hit in the second station (RPC2), and then looking for a hit in the first (RPC1), within the road of the candidate (the path between the hit in the RPC2 and the interaction point). For low- p_T candidates, a 3-out-of-4 coincidence in RPC1 and RPC2, with an additional requirement of a 1-out-of-2 coincidence in the third station

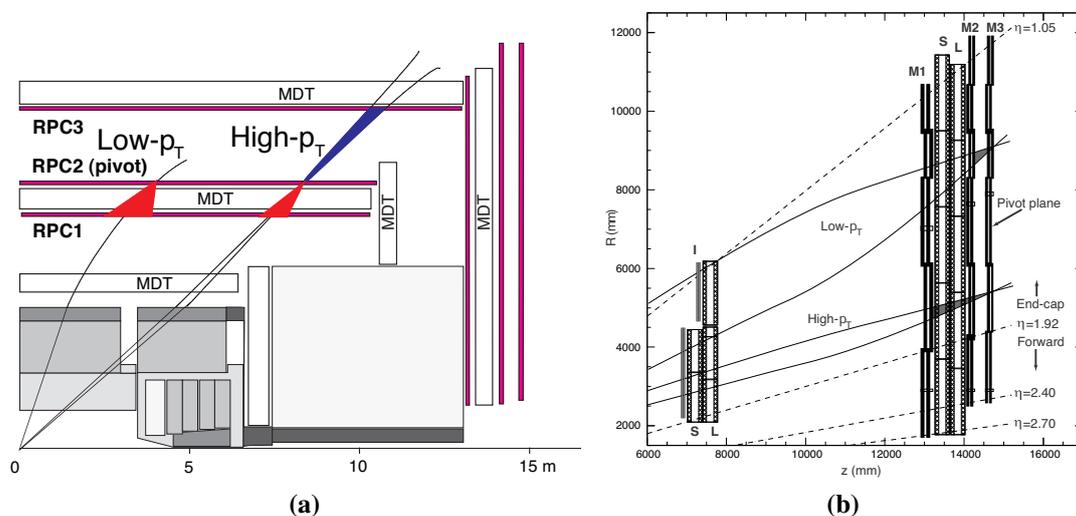


Figure 3.12.: (a) Schema of the L1 muon barrel trigger. The RPCs are arranged in three stations of doublets: RPC1, RPC2, and RPC3, and (b) Schema of the L1 muon end-cap trigger. The TGCs are arranged in three stations of doublets: I, M2, and M3, and one station of triplets: M1. Also shown are the low- and high-transverse momenta roads for both systems. [28]

(RPC3) for high- p_T candidates. The end-cap muon trigger is made up of four stations of TGC detectors (schema is shown in Figure 3.12b). Much like the barrel trigger, the end-cap trigger looks for a hit in the fourth station (M3), and extrapolates back to the interaction point to find further hits, again depending on the p_T of the muon candidate. A more complex requirement on the coincidences is applied separately for the radial distance from the interaction point (R), and ϕ , to make the final trigger decision. The information from the muon barrel and end-cap triggers is sent to the Muon to Central Trigger Processor Interface (MUCTPI), which combines all the results of track multiplicities and candidate p_T , and sends them to the CTP for each bunch-crossing.

High Level Trigger (HLT) and the Data Acquisition System (DAQ)

The HLT/DAQ system is responsible for reading out the detector information from L1, applying the L2 trigger, building the event, and further filtering the event selection with the EF. The DAQ also handles the overall configuration, control, and monitoring of the experiment.

The flow through the HLT/DAQ begins with the L1 trigger. During the L1 latency ($2.5 \mu\text{s}$), event data is buffered within the detector specific electronics, and only sent to HLT/DAQ upon a L1A. For every L1A, the L1Calo, muon triggers, and CTP provide RoI information to the RoI builder, which creates a single data structure to be sent to a L2 Processing Unit (L2PU) to

perform an event selection. Based in software, the L2 trigger reconstructs the RoIs from L1 using the full detector granularity. Additional information is added allowing for more detailed selections, for example shower shape in the calorimeters, to matching from tracks to clusters.

If the event is accepted, it is built by the event-builder before being sent to the EF, at a rate of approximately 5 kHz. The EF is a vast processing farm, which processes events based on the standard ATLAS event reconstruction and analysis frameworks, which are described in Section 3.3.3. The EF then classifies the events according to a pre-defined set of event streams, and if selected, EF output nodes (SFOs) take the data and interface with the HLT/DAQ to be recorded to disk, at a rate of up to 400 Hz.

The selection applied by all trigger levels is defined by the trigger *menu* – a stable trigger configuration defined for the physics run. An example of a trigger used in the analysis described in Chapter 4 is the EF_tau16_loose_e15_medium trigger. The menu defined by this trigger is as follows. Firstly, at L1 the trigger L1_2TAU6_EM10 is applied, which requires two 6 GeV threshold clusters in the HAD calorimeter, and a 10 GeV threshold cluster in the EM calorimeter. Next in the stream is the L2_tau16_loose_e15_medium trigger, which itself takes the RoIs from L1 and fully reconstructs them, ensuring they satisfy the requirements of a *medium* defined electron with a transverse momentum (p_T) threshold of 15 GeV, and a *loose* defined tau with p_T threshold of 16 GeV. The labels loose, medium, and tight, refer to the quality demanded by the selection. Finally, these RoIs are passed to the EF which conducts a full event reconstruction, writing out the event if all other requirements of the trigger are satisfied.

3.3. ATLAS Monte-Carlo and Data

Almost all ATLAS analyses utilise Monte-Carlo (MC) simulation to study the expected detector response to proton-proton interactions and the resulting particles. In fact, due to the immense luminosities of the LHC, there is generally more data available than MC. By having a full model of the detector, one is able to study the expected signal and background data distributions to develop the analysis strategies. The full event simulation [35] process is three steps. Firstly, events are *generated* (Section 3.3.1), whereby the hard interaction of the proton-proton collision is calculated using monte-carlo methods. Secondly, the propagation of the generated particles are *simulated* (Section 3.3.2) as they traverse the ATLAS detector. Finally, the simulated particles undergo *reconstruction* (Section 3.3.3) into physics *objects*. An object is not necessarily a particle, rather it refers to the detector response to the physical object produced in a hard

process, for example, a *jet*. The event reconstruction is shared between MC and real collision data. All of these steps are run through the ATLAS software framework Athena [36].

3.3.1. Generation

At the generation stage, hard processes are calculated by sampling the appropriate matrix element of the process. Higher order corrections such as gluon radiation are also generated in a shower approximation step, depending on the piece of software being used. Immediate decays, such as those of W or Z bosons are generated, and all *stable* particles are saved. As this occurs very close to the interaction point, no knowledge of the ATLAS geometry is needed.

As a proton-proton interaction is not purely the hard process, generators also add in underlying event partons involved in soft processes. Also, due to the high luminosity of the LHC, many collisions may occur during the single bunch crossing (the mean number of interactions per bunch crossing for 7 TeV collisions was approximately 6 [37]). These pileup events are also generated, so that the final MC matches true collision data as closely as possible.

Most of the ATLAS generation is done by software packages maintained externally to ATLAS, including PYTHIA [38, 39], Herwig++ [40], ALPGEN [41], and MC@NLO [42]. Tau decays and photon emissions are handled by TAUOLA [43] and PHOTOS [44].

3.3.2. Simulation

Generated events are read into the full ATLAS simulation software to be propagated through the detector. This propagation is done by GEANT4 [45], which includes a full description of the ATLAS sub-detectors, support material, as well as the solenoidal and toroidal magnetic fields. The configuration of the detector, including any misalignments and distortions can also be specified. The interactions of the particles with the detecting volumes as they traverse the detector are simulated, and the detector outputs are recorded, resulting in an output file of the same form as those produced from real data. A record of the *truth* data is also kept – the true positions and energies of all the simulated particles as they propagate through the detector – for use in performing checks on the success and efficiencies of the reconstruction algorithms.

The ATLAS digitisation software converts the simulated deposits of energy by particles in the sensitive areas of the detector (referred to as *hits*) into *digits*. These digits can be time dependent voltage signals, or simply record when a voltage threshold has been exceeded. The

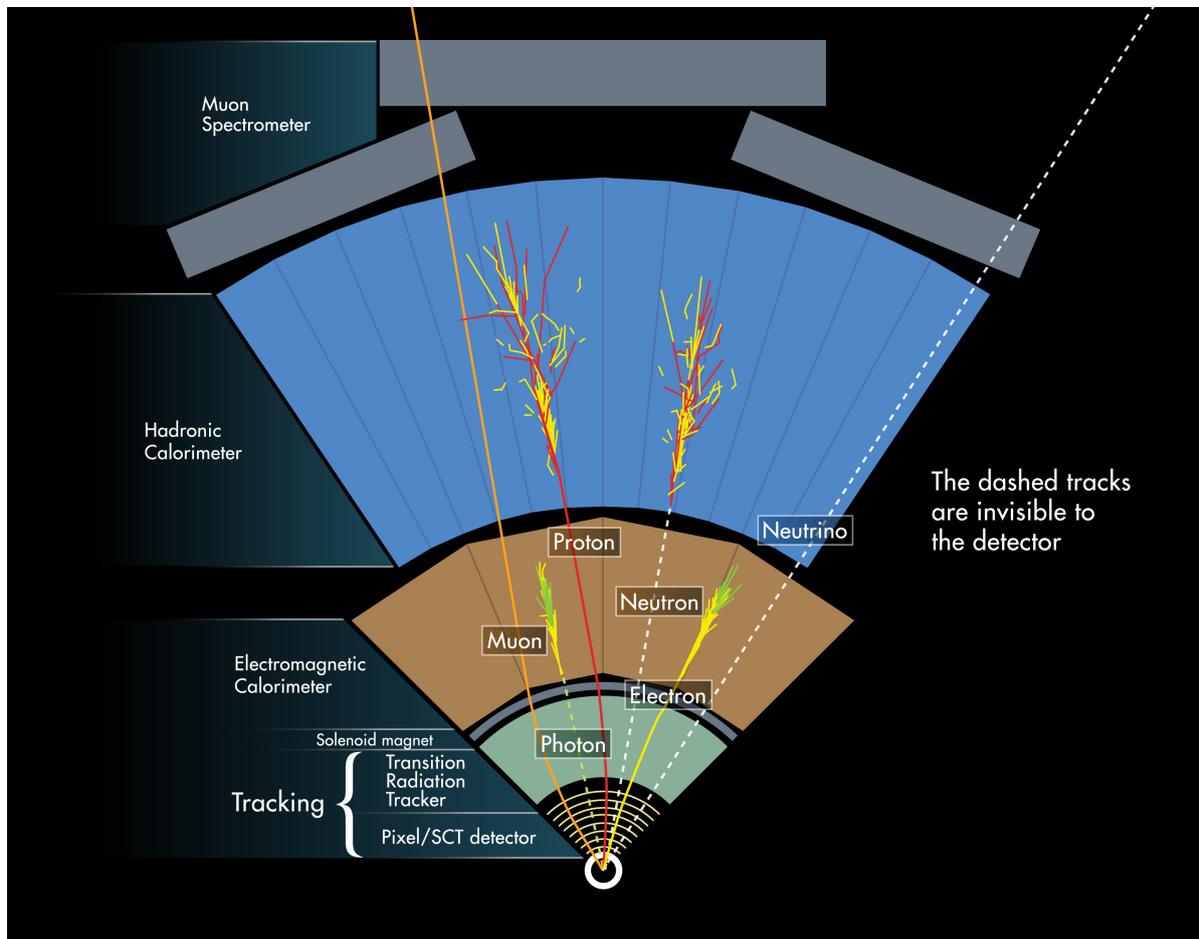


Figure 3.13.: Cross-section diagram of the ATLAS detector, showing how different particles interact with the different sub-detectors [46].

digits are recorded as Raw Data Objects (RDOs). Once created, the RDOs are submitted to the full reconstruction software.

3.3.3. Reconstruction and Identification

The ATLAS reconstruction algorithms take the raw detector outputs in the form of energy deposits and track hits (RDOs), from either simulated or real data, and combine them to form physics objects useful for analyses. Figure 3.13 shows how different particles interact with different parts of the ATLAS detector. Once reconstructed, the candidate objects undergo an *identification* process, whereby they are assigned a particle type (or a useful designation such as jet or missing energy) based on their properties. Here the basic reconstruction and identification algorithms used in this thesis will be explained.

Electrons

Correct identification of isolated electrons has been an essential ingredient in the success of many ATLAS analyses, including many Standard Model measurements, and searches for beyond Standard Model physics. Due to similar detector responses from sources such as misidentified hadrons, or electrons from photon conversions, isolated electrons can have large backgrounds in the ATLAS detector. The electron reconstruction algorithms therefore must have the ability to efficiently reconstruct and identify electrons, while simultaneously rejecting as much of the backgrounds as possible [47].

The electron reconstruction algorithms differ for the barrel ($|\eta| < 2.5$) and end-cap regions of the detector. In the barrel, EM clusters are selected with energy deposits greater than 2.5 GeV, in a window of 0.025×0.025 in (η, ϕ) . These clusters are then associated with tracks in the ID; tracks with $p_T > 0.5$ GeV are extrapolated from their last measured point to the middle layer of the EM calorimeter, at which point it is matched with a cluster if the distance from the track impact point and the cluster centre is $|\Delta\eta| < 0.05$. An electron candidate is classified as reconstructed if at least one track is matched to the cluster, with priority given to tracks with hits in the pixel and SCT detectors for multi-track matches. Once reconstructed, the total electron energy is calculated from the energy left in the EM calorimeter, corrected for material losses, and the spatial coordinates taken from the track parameters.

Forward electrons ($2.5 < |\eta| < 4.9$) do not pass through any tracking volume, and so any deposits in the EMECs are not distinguishable as electrons or photons. Also, due to the lack of detecting material, candidates are required to have $p_T > 20$ GeV. For these reasons, forward electrons are not used in the analyses in this thesis.

Once reconstructed, candidates undergo identification, using further calorimeter and track properties, to reject backgrounds from non-isolated electrons, and misidentified hadrons. Three levels of identification are used – `loose++`, `medium++`, and `tight++` – with each level increasing the background rejection, however with loss of signal efficiency. Loose identification relies on the shower-shape in the EM calorimeter to reject backgrounds, while the medium selection also requires a hit in the innermost layer of the pixel detector, as well as detection by the TRT when available. Finally, the tight selection uses as many identification tools as possible, including requirements on track quality, TRT hits, and the ratio of EM cluster energy to the track momentum.

Muons

Like electrons, correct identification of muons is essential for many ATLAS analyses. The MS, in combination with the ID, is used for detecting the charged tracks of muons as they traverse the detector. There are several different sets of criteria used for muon identification; *standalone* muons are reconstructed only in the MS, *combined* muons are reconstructed independently in the MS and ID, and then combined, and finally *tagged* muons are only reconstructed if the candidate track is associated with either a track segment within the MDT or CSC chambers of the MS, or with an energy deposit within a calorimeter cell [48]. Two algorithms are then used to perform the reconstruction, Staco [49] or Muid [50]. The two algorithms differ in approaches such as the direction in which tracks are extrapolated, and the way ID and MS tracks are combined. The default ATLAS muon reconstruction uses Staco muons.

Like the electron identification, three levels of muons are available, *loose*, *medium*, and *tight*. These levels differ in how many tagging segments are required in the MS, with more tags required for a higher identification level. Also, a series of track quality cuts are placed to further ensure good muons; at least 1 pixel hit, at least 5 SCT hits, at most 2 pixel or SCT sensors traversed by the track but without hits, and, in the TRT acceptance region ($0.1 < |\eta| < 1.9$), at least 9 TRT hits.

Jets

Jets refer to collimated bunches of particles produced from the hadronisation of quarks or gluons emerging from the $p - p$ collisions. Being comprised of many different particles, reconstructed jets are clusterings of these particles into objects in their own right. Jets leave complex signatures in the EM and HAD calorimeters, as well as potential tracks in the ID.

There are many jet clustering algorithms available, all of which combine single tracks and energy clusters into a single object. The ATLAS default jet algorithm is the *anti- k_T* algorithm [51]. Given particles with momenta $p_{T,1}, p_{T,2}, \dots$, the distance between particles i and j can be calculated as

$$d_{ij} = \min(p_{T,i}^{-2}, p_{T,j}^{-2}) \frac{\Delta R_{ij}}{R^2} \quad (3.4)$$

clustering particles with the smallest d_{ij} . It can be seen that if particle j is soft (low momenta), it will cluster with harder particles before it clusters with other soft particles. This behaviour

produces conical jets around the hardest particles, that are infrared safe (presence of a soft particle between two hard particles does not affect their recombination), collinear safe (splitting a particle's momentum between two collinear particles does not affect the jet) and order independent (the same jet is reconstructed at generator to detector level). In ATLAS, jets are reconstructed over the whole range of the calorimeters ($|\eta| < 4.9$), with ID tracks being associated with the jet if their distance from the jet barycentre is $\Delta R < 0.4$.

Due to the relatively long lifetime of b -hadrons, it is possible to *tag* jets originating from these particles due to their displaced vertices compared to the primary vertex. This is particularly useful for identifying $t\bar{t}$ events, as t -quarks decay almost 100% of the time into a W boson and a b -quark. Tagging of b -jets is done by the JetFitterCOMBNN algorithm, which exploits topologies of weak b - and c -hadron decays inside a jet (JetFitter), along with impact parameter based algorithms (IP3D and secondary vertex (SQ1)), combined into a single discriminant using a neural network method [52]. For the analyses outlined in this thesis, a jet is tagged as a b -jet with a JetFitterCOMBNN > 0.35 , corresponding to a b -tagging efficiency of 70%.

As there are on order 5 interactions per bunch crossing in the 7 TeV data, many pile-up jets are expected in each event. A quantity called the Jet Vertex Fraction (JVF), defined as

$$\text{JVF} = \frac{\sum_{\text{PV tracks}} p_{\text{T}}^{\text{trk}}}{\sum_{\text{All jet tracks}} p_{\text{T}}^{\text{trk}}} \quad (3.5)$$

is used to reject jets that are not associated with the hard interaction of interest. The JVF checks the fraction of jet track momentum coming from the primary vertex, compared to the total track momentum of a jet. The higher the JVF, the higher the probability that the jet is *not* a pile-up jet.

Hadronic Taus

As described in Section 2.1, the tau is the heaviest of the charged leptons. Unlike electrons and muons, taus do not travel through the detector, rather, they decay mostly within the ATLAS beam pipe with a mean lifetime of 2.9×10^{-13} s (or a mean decay length of $87 \mu\text{m}$), assuming the tau travels at the speed of light [53]. The dominant decay modes for taus are shown in Table 3.3. Taus decay leptonically (to e/μ) approximately 35% of the time, and hadronically

Decay Mode	Branching Fraction [%]
$\tau \rightarrow e \nu_e \nu_\tau$	17.8
$\tau \rightarrow \mu \nu_\mu \nu_\tau$	17.4
$\tau \rightarrow \pi^\pm \nu_\tau$	10.8
$\tau \rightarrow \pi^\pm (\geq 1 \pi^0) \nu_\tau$	36.6
$\tau \rightarrow 3 \pi^\pm \nu_\tau$	9.3
$\tau \rightarrow 3 \pi^\pm (\geq 1 \pi^0) \nu_\tau$	5.1

Table 3.3.: Dominant decay modes and branching fractions for tau leptons [5].

approximately 64% of the time. Throughout this thesis, the decay products of a hadronically decaying tau may be written as τ_h . Leptonic tau decay products are almost indistinguishable from electrons and muons produced from other prompt decays. For this reason, ATLAS has dedicated hadronic tau reconstruction algorithms.

Hadronic tau decay products generally consist of 1- or 3-*prong* decays, referring to the number of charged decay products, and some number of neutral decay products. The vast majority of the time, the products are charged and neutral pions (π^\pm, π^0), with rarer kaon (K^\pm, K^0) decays. The complication with reconstructing hadronic taus at the LHC is the high amount of QCD jet activity (many orders of magnitude higher), which can leave similar signatures to taus in the detector. Similarly, electrons can *fake* 1-prong tau candidates. The hadronic tau reconstruction algorithms therefore must be able to discriminate between these cases.

As a tau decay is topologically similar to a jet, the tau reconstruction begins from an anti- k_T seed jet (as described above), for all seed jets within $|\eta| < 2.5$ with $p_T > 10$ GeV. The tau candidate 4-momentum is then calculated, using (η, ϕ) of the calorimeter cluster. Tracks are then associated with the tau candidate if they are within $\Delta R < 0.2$ of the cluster centre, and pass a set of quality criteria; $p_T > 1$ GeV, at least 2 pixel hits, at least 7 pixel + SCT hits, the distance of closest approach to the primary vertex in the transverse plane $|d_0| < 1.0$ mm, and the longitudinal distance of approach z_0 satisfying $|z_0 \sin \theta| < 1.5$ mm. Isolation is also imposed on the tau candidate, with all tracks within $0.2 < \Delta R < 0.4$ also required to pass the quality criteria.

The identification step provides most of the discrimination between real taus and QCD jets, exploiting the generally more collimated and lower track multiplicity signature of tau decays compared to jets. Many track and calorimeter variables are derived to study the differences

between the two (and are explained in detail in [53]), including variables such as the ratio of energy deposited in the EM calorimeter to the total tau energy, the p_T weighted distance of EM cells from the tau centre, the p_T weighted distances of tracks from the tau centre, the leading track momentum fraction, and the energy within the core ($\Delta R < 0.1$) compared to the full tau energy. These variables are all combined together into one discriminant using a Boosted Decision Tree (BDT – see Chapter 5 for an in-depth discussion of BDTs). Like the electron and muon identification algorithms, three working points are defined – BDTLoose, BDTMedium and BDTTight – corresponding to signal efficiencies of 60%, 45% and 30% respectively [53].

Discrimination against electrons is also achieved through the use of a BDT, again with three working points corresponding to signal efficiencies of 95%, 85% and 75%, using variables derived from the TRT hits around the tau candidate [53]. Finally, although rare, muons may also fake hadronic tau decays. Muons will leave a higher track-to-calorimeter energy ratio than a true tau candidate, and so a veto has been defined, giving a 50% muon rejection rate, while retaining approximately 96% of tau candidates.

Missing Energy

In hadron colliders, while the total momentum of a collision longitudinal to the beamline cannot be known, the total momentum in the plane transverse to the beamline must be conserved. The missing energy, \cancel{E}_T , is therefore the momentum imbalance in the transverse plane, left over after summing all other momenta in the event. In the SM, missing energy is due to neutrinos escaping the detector, but it can also give evidence of new, weakly interacting particles from beyond SM theories [54].

The calculation of the \cancel{E}_T takes all fully calibrated energies of the object described above, using calorimeter deposits, tracks in the MS, and recovering low p_T tracks that do not reach the calorimeters using the ID. In principle, the total \cancel{E}_T can be defined as

$$\begin{aligned} \cancel{E}_T = & \cancel{E}_T^e + \cancel{E}_T^\gamma + \cancel{E}_T^\tau + \cancel{E}_T^{\text{jets}} + \cancel{E}_T^\mu \\ & + \cancel{E}_T^{\text{softjets}} + \cancel{E}_T^{\text{CellOut}} + \cancel{E}_T^{\text{calo } \mu} \end{aligned} \quad (3.6)$$

however in reality, the contributions to the \cancel{E}_T cannot be measured directly. Cells not associated with any reconstructed objects are included in the CellOut term, while energy losses from muons in the calorimeters are included in the calo μ term. Finally, a softjets term is included for other topological clusters not associated with any reconstructed objects. The \cancel{E}_T resolution,

calculated using $Z \rightarrow \ell\ell$ events in ATLAS data, can be given by $\sigma = k \cdot \sqrt{(\Sigma E_T)}$, where ΣE_T is the total transverse energy of the event, and $k \approx 0.5 \text{ GeV}^{1/2}$ [55].

Chapter 4.

$H \rightarrow \tau_\ell \tau_h$ Search with the ATLAS Detector

This chapter outlines an implementation of the 7 TeV $H \rightarrow \tau\tau \rightarrow \tau_h \nu_\tau \ell \nu_\ell$ cut-based analysis as originally completed by the ATLAS Collaboration [1]. The decay of interest has one tau decaying hadronically and the other leptonically, and will be referred to by the shorthand $H \rightarrow \tau_\ell \tau_h$. While the author had no direct input into this paper, the analysis has been modified in this thesis, and the results serve as a benchmark for our multi-variate approach to be compared against.

The cut-based analysis takes $\sqrt{s} = 7$ TeV ATLAS data, and after applying appropriate pre-selections (Section 4.5), splits it into different analysis categories based on the event topologies of the Higgs production mechanisms (Section 4.6). Once categorised, upper limits on the production cross-section of the Higgs boson with respect to the SM prediction are calculated. Section 4.7 outlines the background estimation methods used for this and the following analyses, and Sections 4.8 and 4.9 show the treatment of systematic uncertainties and the limit setting procedure used. Finally, the upper limits obtained from the analysis are presented in Section 4.10.

4.1. Signal and Background Processes

4.1.1. $H \rightarrow \tau_\ell \tau_h$ Signal

At the energies involved at the LHC, the SM Higgs boson is produced mainly via gluon-gluon fusion (ggF) through an intermediate top-loop, Vector Boson Fusion (VBF), or with an associated vector boson (VH , $V = W, Z$). Other production modes, including those with associated t -quarks, are not considered here due to their relatively low cross-section. Figure 2.3

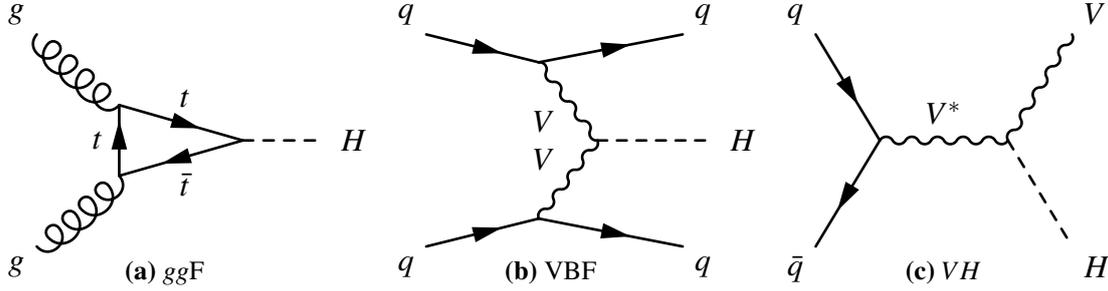


Figure 4.1.: Leading order Higgs production modes at the LHC; (a) gluon-gluon fusion (ggF), (b) vector boson fusion (VBF), and (c) associated production (VH , $V=W, Z$)

shows the production cross-section of the Higgs boson at the LHC at 7 TeV centre-of-mass energy, as well as the branching ratios, as functions of the Higgs mass. Leading order Feynman diagrams of the major modes shown in Figure 4.1.

At leading order, the ggF process produces a low- p_T Higgs. In the tau decay mode, the Higgs decay products will be produced almost back to back in ϕ . Due to the large boost of the taus, the neutrinos from their decays emerge similarly almost back to back, which when summed, result in a low missing energy¹ (\cancel{E}_T) measurement for the system. At NLO however, the ggF process occurs with associated jets, which can give the Higgs a boost transverse to the beamline. In this situation, as the tau decay products are no longer back to back, the missing energy from the neutrinos sums constructively, giving a higher \cancel{E}_T to the system. Reconstructing the Higgs candidate in the boosted scenario therefore improves the mass resolution, as more information about the escaping neutrinos can be inferred.

Although ggF occurs with a significantly larger cross-section than VBF , the topology of VBF production allows for the most significant suppression of the background processes. VBF events have the distinctive feature of two high- p_T jets on opposite sides of the detector (high separation in η , defined as $\Delta\eta$), with the Higgs decaying centrally. Tagging events with very forward jets gives a handle with which to extract signal events. The desire to fully exploit the ggF and VBF topologies is the motivating factor for the categorisation step of the analysis strategy, described in Section 4.6.

When produced with an associated vector boson, Higgs decay products again become boosted, much like with NLO ggF events. The associated W or Z also adds a charged lepton and a neutrino or a pair of oppositely charged leptons to the event topology. The analysis outlined here does not take these extra leptons into account, however, there are dedicated analyses that focus on such events [56].

¹Although we say energy as per convention, we are really referring to the vector sum of the neutrino momenta.

Process		$\sigma \times \mathcal{B}$ [pb]
$W(\rightarrow \ell\nu) + \text{jets}$		1.046×10^4
$Z/\gamma^*(\rightarrow \ell\ell) + \text{jets} (m_{\ell\ell} > 40 \text{ GeV})$		1.07×10^3
$Z/\gamma^*(\rightarrow \ell\ell) + \text{jets} (10 < m_{\ell\ell} < 40 \text{ GeV})$		3.89×10^3
$t\bar{t}$		1.646×10^2
single- t	t -channel	58.7
	Wt -channel	15.6
	s -channel	3.90
$VV \rightarrow \ell\tau_h X$	WW	46.2
	WZ	18.0
	ZZ	5.6
$H \rightarrow \tau_\ell \tau_h$	ggF	5.4×10^{-1}
	VBF	4.1×10^{-2}
	WH	1.5×10^{-3}
	ZH	8.3×10^{-4}

Table 4.1.: Cross-sections (σ) times branching fraction (\mathcal{B}) of the various background and signal processes studied.

4.1.2. Background Contributions

In order to observe the presence of any signal, all background physics processes that produce a similar detector response must be understood. The main background processes studied in this analysis are outlined below. Their cross-sections times the branching fraction (\mathcal{B}), and those of the $H \rightarrow \tau_\ell \tau_h$ signal processes are shown in Table 4.1. The most significant background processes are $W(\rightarrow \ell\nu) + \text{jets}$ and $Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$. Also important are fake- τ events from QCD di-jet processes; the estimation of such events is taken from data, as described in Section 4.7.

$Z \rightarrow \tau\tau$ processes

Due to having the same final state as the signal processes, $Z \rightarrow \tau_h \tau_\ell$ events form a largely irreducible background for the analysis (with example Feynman diagrams shown in Figure 4.2). Events where both taus decay leptonically can enter the selection if one lepton is mis-reconstructed as a tau.

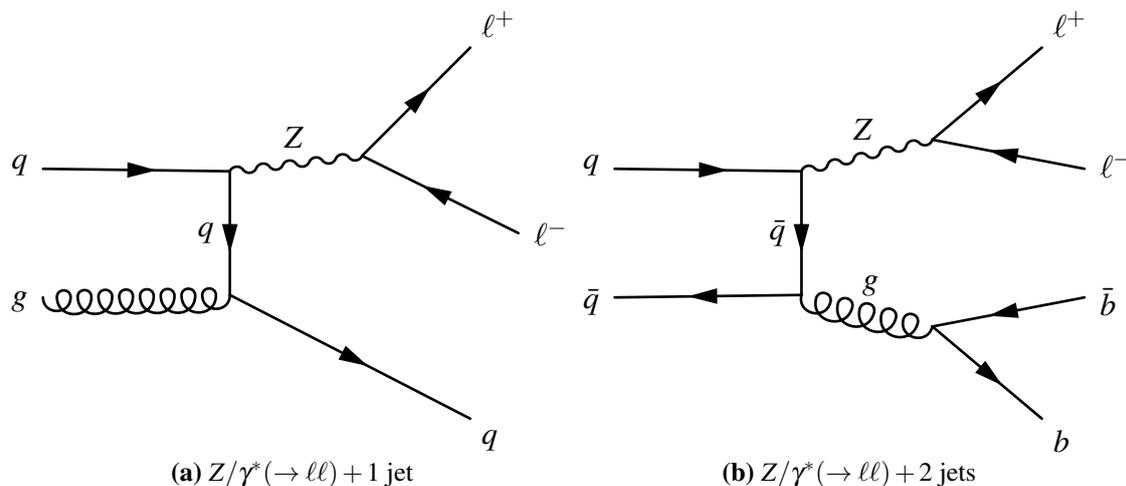


Figure 4.2.: Example Feynman diagrams for $Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$ processes

$Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$ processes ($\ell = e/\mu$)

$Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$ processes (also shown in Figure 4.2) contribute to the background when either one of the leptons, or one of the associated jets, is reconstructed as a hadronic tau. This situation is of particular importance, as the parent mass of such a system typically peaks closer to the expected signal mass than the true $Z \rightarrow \tau_h \tau_\ell$ events, due to the lack of the extra tau neutrino. How to deal with these events is described in section 4.5.

$W(\rightarrow \ell\nu) + \text{jets}$ processes ($\ell = e/\mu/\tau$)

Similar to $Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$ events, $W(\rightarrow \ell\nu) + \text{jets}$ (as shown in Figure 4.3) can enter the backgrounds when one of the accompanying jets is reconstructed as a hadronic tau. The final event topology here is therefore similar to the signal topology, as the extra neutrino from the W decay also results in a comparable \cancel{E}_T . Considering the relatively high cross-section of $W(\rightarrow \ell\nu) + \text{jets}$ events, they contribute significantly to the backgrounds.

$t\bar{t}$ processes

Many $t\bar{t}$ pairs are produced at the LHC. Top quark pairs will produce a high jet multiplicity final state, almost always decaying $t\bar{t} \rightarrow W^+ b W^- \bar{b}$ (as shown in Figure 4.4). If the W decays leptonically, the final state will include leptons, \cancel{E}_T , and possible hadronic taus. As with most of

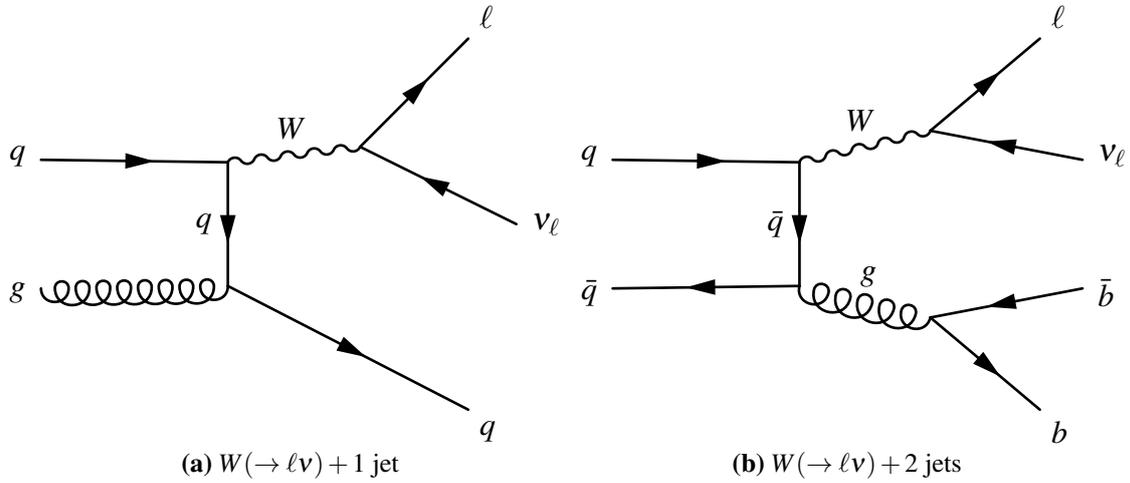


Figure 4.3.: Example Feynman diagrams for $W(\rightarrow \ell\nu) +$ jets processes.

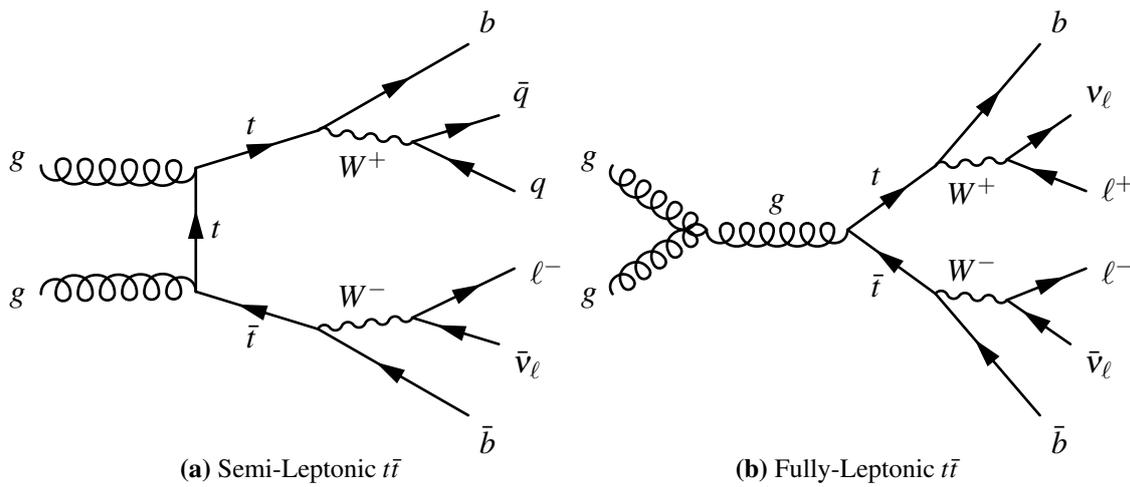


Figure 4.4.: Example Feynman diagrams for $t\bar{t}$ processes.

the major backgrounds, fake hadronic tau candidates can also arise from the mis-reconstruction of light leptons or jets.

Single-top processes

Single top contributions can occur as s - or t -channel processes, or in association with a W (as shown in Figure 4.5). In all cases, any jet or lepton from the associated or prompt W decay can fake a hadronic tau.

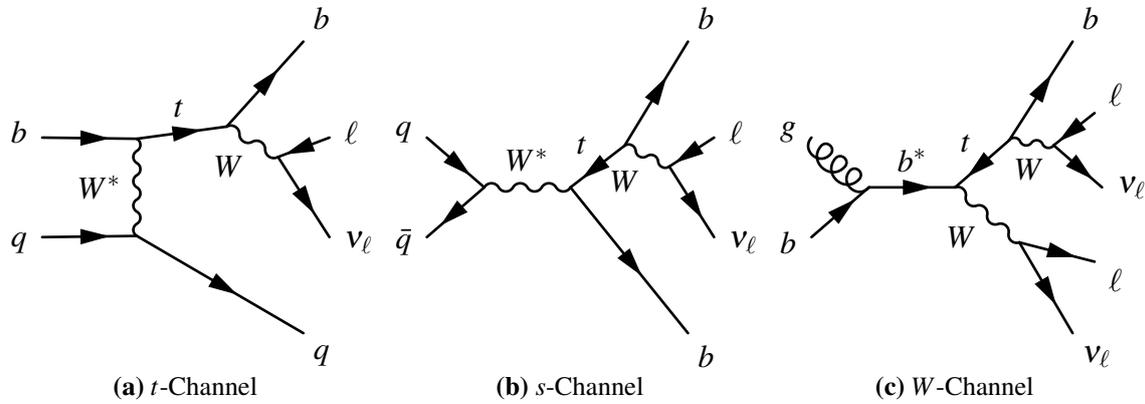


Figure 4.5.: Example Feynman diagrams for single- t processes.

Di-boson processes

Di-boson events produce a pair of vector bosons (WW , WZ , ZZ). Final states with one lepton and one hadronic tau can arise if one or both of the bosons decay leptonically. Although the cross-section of di-boson processes is low, in events containing W bosons, the extra neutrino brings the \cancel{E}_T closer to that of the signal. Events with only light leptons as decay products can also become backgrounds if one is reconstructed as a hadronic tau. Finally, hadronic decays of the bosons can also lead to the same final state as the signal if the hadronic jet is reconstructed as a tau.

QCD jets processes

QCD di-jet processes occur at high rates at the LHC. Such events can replicate the signal final state if one jet is reconstructed as a hadronic tau.

4.2. Monte-Carlo Samples

Monte-Carlo samples are used to study features of the signal processes, and also to investigate the background contributions.

The Higgs signal processes (ggF , VBF , and VH) are modelled using PYTHIA [38], which handles the hard interaction, as well as the final state parton showers and decays at leading-order (LO). Next-to-leading-order (NLO) corrections for ggF and VBF production modes are

provided by POWHEG [57, 58]. Samples are generated for a range of Higgs masses, from 100 – 150 GeV, in increments of 5 GeV.

The $W(\rightarrow \ell\nu) + \text{jets}$ and $Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$ processes are generated with ALPGEN [41], calculating LO matrix elements for up to 5 radiated jets. ALPGEN uses the MLM matching scheme² [59] to ensure there is no double counting of events between hard processes that have equivalent amounts of radiation. HERWIG [40] is used to simulate the parton showers.

The $t\bar{t}$, single-top, and di-boson events are generated at NLO by MC@NLO [42], also using HERWIG for parton shower and hadronisation. The underlying event for the above processes is simulated by JIMMY [60]. Finally, the QCD jet processes are generated by PYTHIA. The tunes (parameters of the generators) for HERWIG, JIMMY, and PYTHIA are described in [61]. As well as MC samples, $Z \rightarrow \tau\tau$ are also simulated using a semi-data-driven method known as the *embedding* technique, which is explained in Section 4.7.2.

To simulate tau decays, TAUOLA [43] is used, along with PHOTOS [44] to simulate any extra photon radiation from charged leptons to leading-log (LL) approximation. To account for pile-up conditions (multiple events from the same or neighbouring bunch crossings) events are re-weighted to match LHC run conditions according to the average number of interactions per bunch crossing $\langle\mu\rangle$, obtained from ATLAS data.

Once generated, all MC events are passed through a full ATLAS simulation using GEANT4 [35, 45], and are reconstructed with the same software as for data (as described in Chapter 3).

4.3. ATLAS Data

The data used was collected from the ATLAS detector during 2011 with the LHC running proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 7$ TeV. The data collected from the trigger streams of interest amounted to a total of $4.60 \pm 0.08 \text{ fb}^{-1}$.

As the desired final state includes a lepton and hadronic tau candidate, a combination of single lepton, and lepton+tau triggers was used. Table 4.2 lists all the triggers running in the data used in this analysis. A *period* represents a period of data taking when the LHC machine conditions were approximately the same and the trigger menu was stable. Triggers are named by their level, object selection, and the tightness of the selection. The details of the ATLAS trigger algorithms are described in Section 3.2.5. For example, EF_tau16_loose_e15_medium

²named after its author, Michelangelo L. Mangano.

is an $e + \tau$ trigger at the event-filter level, with a loose selection of a tau with $p_T^\tau > 16$ GeV and an electron with $p_T^e > 15$ GeV and a medium selection criteria. The table shows that the electron energy threshold for the lepton+tau trigger (LTT) is lower than that of the single lepton trigger (SLT). Thus, by collecting both streams one can recover events with high- and low- p_T electrons. Only a SLT for the muon stream was used. As the instantaneous luminosity increased during the runs, later periods had to use higher threshold triggers, or employ a tighter identification scheme to maintain a constant data rate. The early (later) runs required either a medium (medium1 – the 1 here is used to indicate that the algorithm changed between trigger menus) electron candidate of $p_T > 15(20)$ GeV, a loose (medium) hadronic tau candidate with $p_T > 16(20)$ GeV and an electron candidate with $p_T > 15$ GeV, or a (medium) muon candidate with $p_T > 18$ GeV.

4.4. Object Selection and m_{MMC} Calculator

An *object* refers to the detector response that corresponds to the physical particle produced in the hard process. In this analysis, the objects studied are the hadronic tau candidate, lepton, leading and sub-leading jets, and the \cancel{E}_T . The selection of these objects follows the recommendations of the ATLAS combined performance groups. The selection criteria will be outlined below, and also summarised in table 4.3. The reconstruction algorithms used are described in detail in section 3.3.3.

4.4.1. Muons

Muons are reconstructed using the Staco algorithm (as described in Section 3.3.3), which selects candidates with tracks in both the inner detector and the muon spectrometer, with a loose identification level. Fiducial selections of $p_T > 10$ GeV and $|\eta| < 2.5$ are applied. The muon tracks are also required to be isolated in both the electromagnetic and hadronic calorimeters, as well as the tracking detectors. Within a cone of $\Delta R = 0.2$ around the track, the ratio of energy deposited in the the calorimeters to the muon p_T ($\sum p_T^{clr}/p_T^\mu$) must be less than 2%. Additional tracks of $p_T > 1$ GeV within a cone of $\Delta R = 0.4$ of the muon track ($\sum p_T^{trk}/p_T^\mu$) must also not sum to more than 6% of the muon p_T . Differences in p_T resolution [62] and identification efficiency [63] between data and simulation are corrected for in Monte Carlo.

Period	# Runs	Run Range	e -trigger	$e + \tau$ -trigger	μ -trigger	\mathcal{L} [pb^{-1}]
B2	3	178044–178109	EF_e20_medium	EF_tau16_loose_e15_medium	EF_mu18_MG	11.7
D	22	179725–180481	EF_e20_medium	EF_tau16_loose_e15_medium	EF_mu18_MG	161.9
E	5	180614–180776	EF_e20_medium	EF_tau16_loose_e15_medium	EF_mu18_MG	48.8
F	15	182013–182519	EF_e20_medium	EF_tau16_loose_e15_medium	EF_mu18_MG	136.1
G	28	182726–183462	EF_e20_medium	EF_tau16_loose_e15_medium	EF_mu18_MG	537.6
H	13	183544–184169	EF_e20_medium	EF_tau16_loose_e15_medium	EF_mu18_MG	259.5
I	23	185353–186493	EF_e20_medium	EF_tau16_loose_e15_medium	EF_mu18_MG	324.8
J	9	186516–186755	EF_e20_medium	EF_tau16_loose_e15_medium	EF_mu18_MG_medium	226.4
K	16	186873–187815	EF_e20_medium	EF_tau20_medium_e15_medium	EF_mu18_MG_medium	561.7
L	44	188921–190343	EF_e20vh_medium1	EF_tau20_medium_e15vh_medium	EF_mu18_MG_medium	1387.7
M	30	190608–191933	EF_e20vh_medium1	EF_tau20_medium_e15vh_medium	EF_mu18_MG_medium	1005.1
Total	208	1780044–191933				4.6 fb^{-1}

Table 4.2.: Run, trigger information, and integrated luminosities for data used. The terminology used here is explained in the text.

	Muon	Electron	Jet	Hadronic Tau
Selection	$p_T > 10 \text{ GeV}$	$p_T > 15 \text{ GeV}$	$p_T > 25 \text{ GeV}$	$p_T > 20 \text{ GeV}$
	$ \eta < 2.5$	$ \eta < 1.32, 1.52 < \eta < 2.47$	$ \eta < 4.5$	$ \eta < 2.5$
	$\sum p_T^{clr}/p_T^\mu < 0.04$ ^a	$\sum p_T^{clr}/p_T^e < 0.08$ ^a	$ JVF > 0.75$ ^c	BDTMedium
	$\sum p_T^{trk}/p_T^\mu < 0.06$ ^b	$\sum p_T^{trk}/p_T^e < 0.06$ ^b		1 or 3 tracks charge ± 1 e/μ Veto

^a $\sum p_T^{clr}/p_T$ is sum over LAr calorimeter cells within $\Delta R < 0.2$ of the candidate track
^b $\sum p_T^{trk}/p_T$ is sum over tracks with $p_T > 1 \text{ GeV}$ within $\Delta R < 0.4$ of the candidate track
^c for jets within $|\eta| < 2.4$

Table 4.3.: Object selection criteria

4.4.2. Electrons

Electrons are reconstructed from clusters of cells in the electromagnetic calorimeter that are matched to a track in the inner detector, and further identified using information from the electromagnetic shower shape. The loose++ identification algorithm is used, as described in Section 3.3.3. Selected electrons are required to have $p_T > 15 \text{ GeV}$ and be within the detector region $|\eta| < 2.47$, excluding the gap region of $1.32 < |\eta| < 1.52$. Much like the muon selection, isolation requirements are imposed in both the calorimetry and the tracking detectors, with maximum momentum fractions of 8% and 6% respectively. Mis-modelling of reconstruction and identification efficiencies, energy resolutions and calorimeter isolation are corrected for in Monte Carlo [47].

4.4.3. Jets

Although not used in the event pre-selection, jet candidates are used to separate the events into their analysis categories. Jets are reconstructed from EM calorimeter clusters using the anti- k_T algorithm (as described in Section 3.3.3). Selected jets are required to be within $|\eta| < 4.5$ with a $p_T > 25 \text{ GeV}$. Additionally, jets within the tracking detectors ($|\eta| < 2.5$) are subject to a vertex cut. The jet-vertex-fraction (JVF) is used to suppress pile-up contributions to the jet selection, and a cut of $|JVF| > 0.75$ is required.

Calibration schemes were used to convert the EM calibration of the calorimeter clusters to a the energy scale of the hadronic calorimeters, (in both η and p_T) [64], with corrections applied to account for out-of-cluster energy and dead material losses [65].

4.4.4. Hadronic Taus

Tau candidates are required to be within $|\eta| < 2.5$ with a $p_T > 20$ GeV. Track multiplicity is required to be 1 or 3, with a charge of ± 1 . For identification, the BDTMedium algorithm was applied, which gives a signal efficiency of approximately 60%. An electron veto (BDTMedium) is also applied, to reject electrons faking tau candidates, maintaining a signal efficiency of approximately 85%. Similarly, the muon veto is applied, rejecting approximately 50% muons with almost no signal loss.

4.4.5. Missing Transverse Energy

Missing energy gives a measure of the amount of neutrino energy in the event, and is calculated from the momentum imbalance in the plane transverse to the beamline, by summing all of the observed momenta. This calculation is described in more detail in Section 3.3.3.

4.4.6. Missing Calorimeter Region and Overlap Removal

During approximately 20% of the data taking, part of the LAr calorimeter was unavailable (due to issues with the optical readout electronics). To reject any spurious missing energy, any events with a selected jet within $\Delta R = 0.2$ of the LAr “hole” were rejected. Similarly, any events with tau candidates within $\Delta R = 0.1$ were also rejected. It should be noted that in MC no events were rejected, rather the total luminosity was scaled to account for the lost events.

Following object identification, any objects overlapping in a region $\Delta R < 0.2$ are subject to overlap removal. After ignoring lepton isolation requirements and requiring a looser electron identification, the order of priority in deciding on the identity of the object is muons $>$ electrons $>$ tau candidates $>$ jets. If two objects overlap, the lower priority object is removed.

4.4.7. Missing Mass Calculator (MMC)

The final event topology of the $H \rightarrow \tau_\ell \tau_h$ system includes two tau neutrinos from the tau decays, and a further neutrino from the leptonically decaying tau. In order to reconstruct the di-tau mass ($m_{\tau\tau}$), one needs to determine the x -, y -, and z -momentum components of the tau neutrinos (\mathbf{p}_{mis_1} and \mathbf{p}_{mis_2}), and the invariant mass of the neutrinos from the leptonically decaying tau system (m_{mis_1} and m_{mis_2}). These 7 unknowns however only have 4 equations available:

$$\begin{aligned}
 \mathcal{E}_{T,x} &= p_{mis_1} \sin \phi_{mis_1} \cos \phi_{mis_1} + p_{mis_2} \sin \phi_{mis_2} \cos \phi_{mis_2} \\
 \mathcal{E}_{T,y} &= p_{mis_1} \sin \phi_{mis_1} \sin \phi_{mis_1} + p_{mis_2} \sin \phi_{mis_2} \sin \phi_{mis_2} \\
 M_{\tau_1}^2 &= m_{mis_1}^2 + m_{vis_1}^2 + 2\sqrt{p_{vis_1}^2 + m_{vis_1}^2} \sqrt{p_{mis_1}^2 + m_{mis_1}^2} - 2p_{vis_1} p_{mis_1} \cos \Delta\theta_{vm_1} \\
 M_{\tau_2}^2 &= m_{vis_2}^2 + 2\sqrt{p_{vis_2}^2 + m_{vis_2}^2} \sqrt{p_{mis_2}^2 + m_{mis_2}^2} - 2p_{vis_2} p_{mis_2} \cos \Delta\theta_{vm_2}
 \end{aligned} \tag{4.1}$$

where $\mathcal{E}_{T,x/y}$ are the x - and y - components of the \mathcal{E}_T , and $p_{vis_{1/2}}$, $m_{vis_{1/2}}$, $\theta_{vis_{1/2}}$, and $\phi_{vis_{1/2}}$ are the momenta, invariant mass, polar, and azimuthal angles of the visible tau decay products. The tau invariant mass is fixed to $M_{\tau_{1/2}} = 1.777$ GeV, and $\Delta\theta_{vm_{1/2}}$ is the polar angle between the missing and visible components ($\mathbf{p}_{mis_{1/2}}$ and $\mathbf{p}_{vis_{1/2}}$) of the momentum vectors of the tau decay products. Clearly this system of equations is under-constrained, however, a probabilistic approach can be taken to estimate the most likely solutions for the unknowns [66]. The extra information included is the 3-dimensional angle between the visible and missing tau decay products,

$$\Delta\Theta = \cos^{-1} \left(\frac{\mathbf{p}_{vis} \cdot \mathbf{p}_{mis}}{|\mathbf{p}_{vis}| |\mathbf{p}_{mis}|} \right) \tag{4.2}$$

This knowledge is encoded in a series of PDFs; $\mathcal{P}(\Delta\Theta, p_\tau)$, which are parameterised by fitting $\Delta\Theta$ distributions (obtained for a range of p_τ s in simulated $Z/\gamma^* \rightarrow \tau\tau$ events) with a linear combination of Gaussian and Landau functions of the form $a_0(e^{-a_1 p_\tau} + a_2/p_\tau)$, where a_i are parameters of the fit.

Once the appropriate PDFs have been made, the system of equations 4.1 is solved for any point in $(\phi_{mis_1}, \phi_{mis_2}, m_{mis_1})$ space³. The angle $\Delta\Theta$ is therefore fully defined, and a global event probability⁴ can be determined

$$\mathcal{P}_{event} = \mathcal{P}(\Delta\theta_1, p_{\tau_1}) \times \mathcal{P}(\Delta\theta_2, p_{\tau_2}) \quad (4.3)$$

For each event, a $m_{\tau\tau}$ distribution is produced, weighted by the \mathcal{P}_{event} for each scan point. The maximum of this distribution is used as the final value of $m_{\tau\tau}$, which will be referred to from here as m_{mmc} .

The efficiency of the MMC calculator can be estimated by the number of events for which a solution is found, divided by the total number of events. While the efficiency is very high for $Z \rightarrow \tau\tau$ events (approximately 99%), it is very low for backgrounds. This allows one to reject backgrounds based on the MMC output. Deviations from 100% efficiency arise from neutrino vectors being outside the scan window, and missing energy values fluctuating more than the allowed 3σ .

4.5. Event Pre-Selection

After applying the trigger requirements outlined in section 4.3, it is necessary to impose further quality criteria to the data, known as pre-selections. Various *event cleaning* cuts are placed to account for bad quality events. Missing energy originating from sources besides proton-proton collisions is suppressed by *jet cleaning*, where fake jets are reconstructed due to issues in hardware, LHC beam conditions, cosmic-ray showers, and so on [67]. These corrections reject approximately 19% of all reconstructed data, however this results in a much higher consistency and quality of data.

After the cleaning cuts, topological cuts are placed to focus only on events with a final state like that of the signal process. Using the object selection as described in Section 4.4, the following requirements are placed:

³ m_{mis_1} is only required in the scan for the leptonically decaying tau, as there is less information available due to the extra neutrino.

⁴As explained in [66], the MMC performance is in fact highly correlated with \mathcal{E}_T resolution. To account for this, the resolution of $\mathcal{E}_{T,x}$ and $\mathcal{E}_{T,y}$ can also be included in the parameter scan, and the resolution functions $\mathcal{P}(\Delta\mathcal{E}_{x/y}) = \exp[-(\Delta\mathcal{E}_{x/y})/2\sigma^2]$ can be included into the event probability, where σ is the \mathcal{E}_T resolution provided by ATLAS and $\Delta\mathcal{E}_{x,y}$ is the difference between the measured and scanned values of $\mathcal{E}_{T,x/y}$.

Process	$e\tau_h$ Events	$\mu\tau_h$ Events
ggF 125	57.75 ± 0.59	46.78 ± 0.53
VBF 125	5.586 ± 0.051	4.596 ± 0.046
WH 125	1.956 ± 0.071	1.742 ± 0.060
ZH 125	1.098 ± 0.039	0.974 ± 0.037
SS-Data	24210 ± 160	15040 ± 130
$Z \rightarrow \tau\tau$	18154 ± 79	13391 ± 53
$W(\rightarrow \ell\nu) + \text{jets}$	15030 ± 190	13750 ± 240
$Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$	1584 ± 31	920 ± 25
$t\bar{t}$	1342.7 ± 9.7	1067.1 ± 8.6
$VV \rightarrow \ell\tau_h X$	305.3 ± 6.0	278.6 ± 5.8
Signal	66.47 ± 0.60	54.09 ± 0.54
Background	60620 ± 260	44450 ± 280
Data	60810	45021

Table 4.4.: Number of events at pre-selection level for the $e\tau_h$ and $\mu\tau_h$ channels. The various backgrounds are estimated using the methods described in Section 4.7, with SS-Data being a data driven estimate of the multi-jet contribution. Uncertainties are statistical only.

- **1 light lepton** (e or μ) with $p_T > 25$ GeV is required for events passing the SLT, with $17 < p_T < 25$ GeV for (e) events passing the LTT (SLT and LTT are defined in Section 4.3). Events with > 1 lepton (after loosening the electron ID and muon isolation) are rejected to suppress $Z \rightarrow \ell\ell$ ($\ell = e/\mu$) and single-top processes.
- **1 hadronic tau** candidate with $p_T > 20(25)$ GeV for the SLT (LTT) trigger.
- **Charge correlation** between the tau and lepton is required (opposite charges).

The number of events in data, and in each of the simulated signal and background samples is shown in Table 4.4. Figures 4.6 and 4.7 show the transverse momenta of the selected objects (hadronic tau candidate, lepton, leading and sub-leading jets, and \cancel{E}_T), as well as the m_{MMC} distribution, for both the e and μ channels. The methods used for estimating the various background contributions are described in Section 4.7. A good agreement between data and MC is seen for all variables.

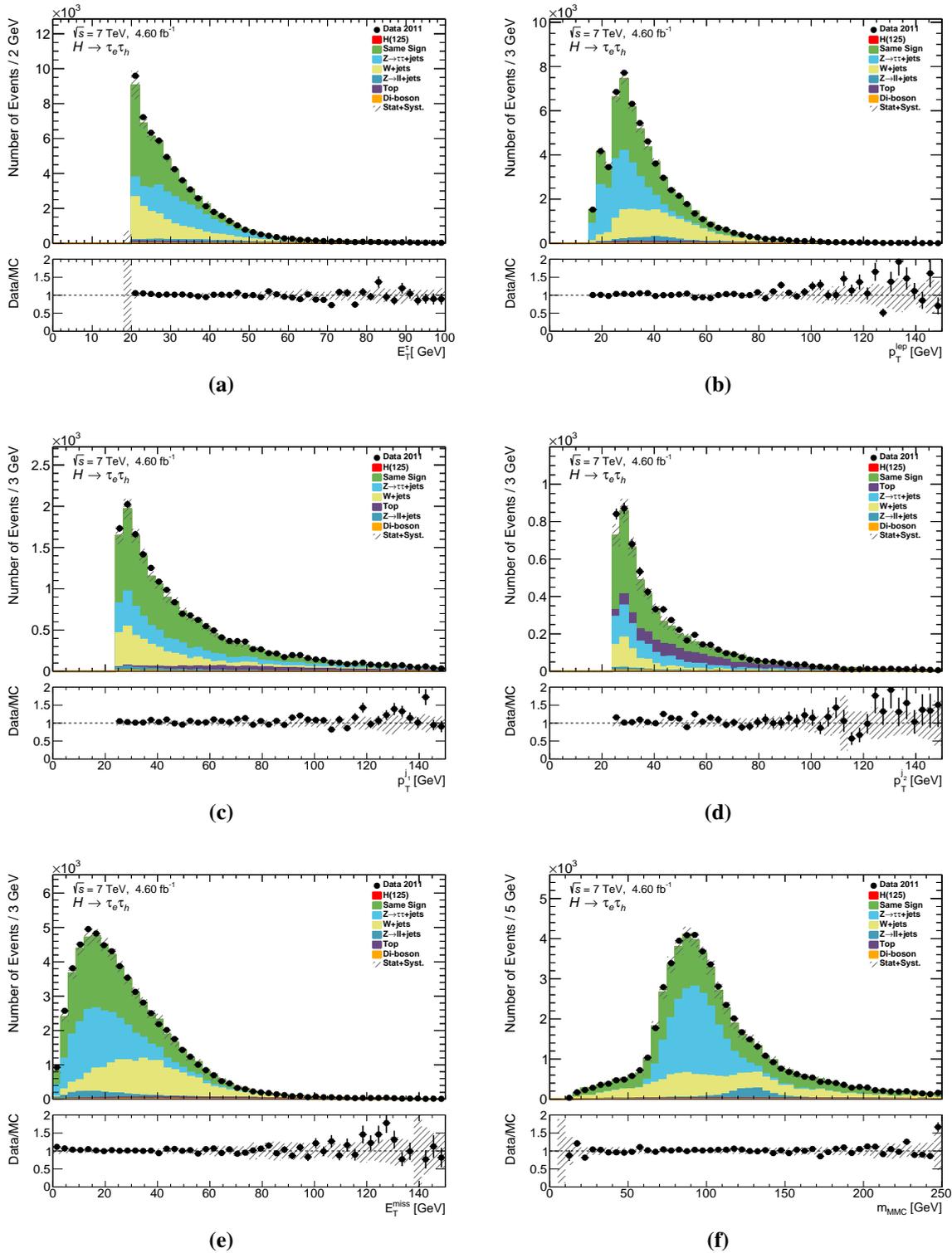


Figure 4.6.: Kinematic variables for the selected objects, and the m_{MMC} distribution, at pre-selection level for the e channel. Uncertainties are statistical and systematic.

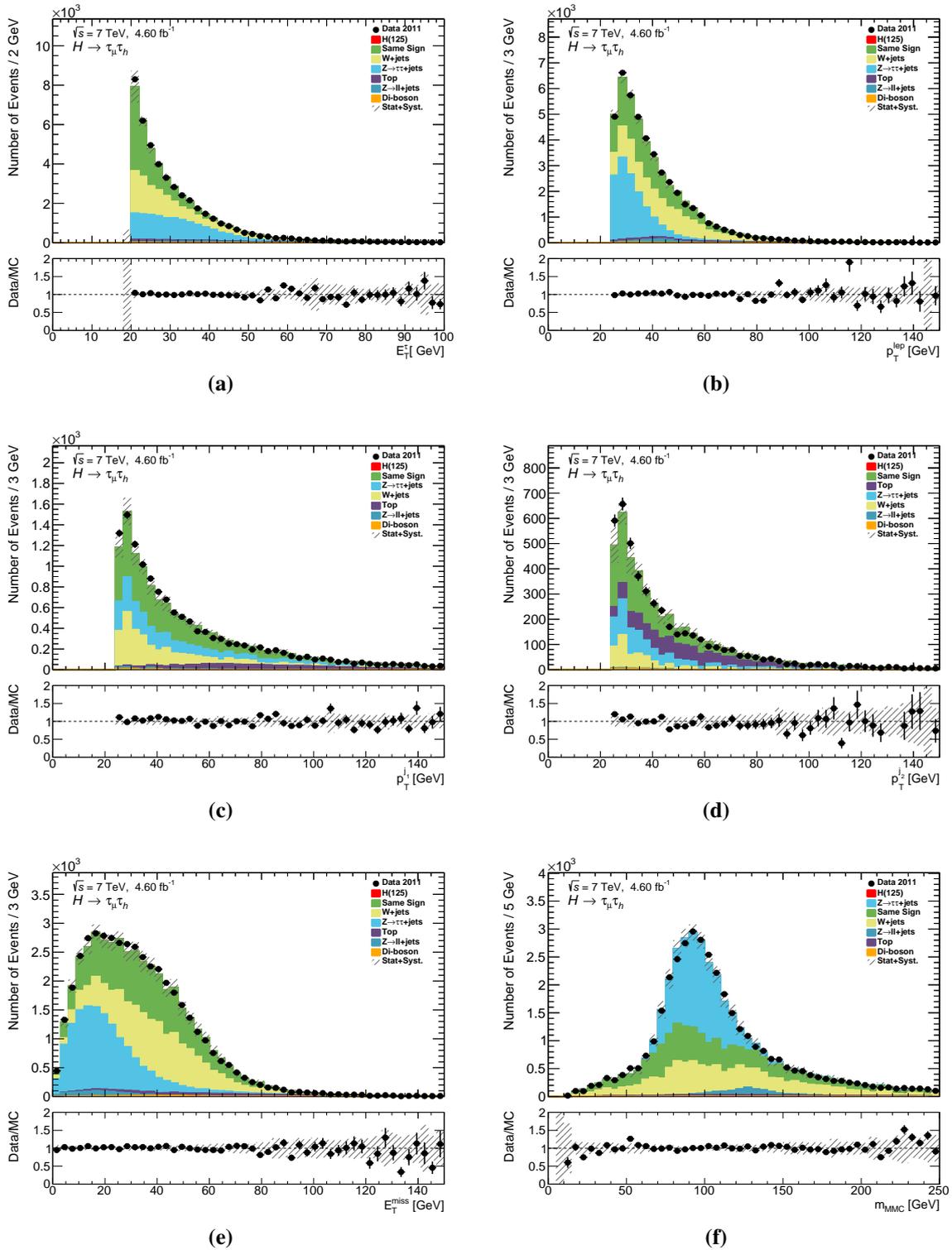


Figure 4.7.: Kinematic variables for the selected objects, and the m_{MMC} distribution, at pre-selection level for the μ channel. Uncertainties are statistical and systematic.

4.6. Analysis Strategy

After pre-selection, the events are sorted into *categories*, designed to exploit specific topological properties of the Higgs production mechanisms. The cut-based approach to the categorisation is described here, however the motivation for, and basic features of the categories are still valid for the hybrid and multi-variate analyses described in the following chapters.

Final states produced with Higgs boson decays can have jets in the final state. Production via VBF and VH (when V decays hadronically) processes can result in two jets in the final state at LO. The ggF production process can also give rise to jets through initial state radiation (ISR – where one of the initial partons can *radiate* a gluon). The major background ($Z \rightarrow \tau\tau$) is produced via $q\bar{q}$ at LO, and as the rate of ISR for gluons is higher than for quarks, the rate of signal events with jet signatures should be higher than that of the backgrounds. The backgrounds are split into four categories; the VBF category for events with two hard, forward jets, the Boosted category for events with a boosted Higgs candidate, the 1-Jet category for events with one associated jet in the final state, and the 0-Jet category for events with no jets in the final state. The optimisation process for each category is described below. A summary of the categorisation and analysis level cuts is shown in Table 4.5.

4.6.1. VBF Category

As discussed in section 4.1.1, the topology of the VBF Higgs production mode provides good discrimination against background processes. The selection of events with a final state involving two highly forward jets is optimised in a dedicated VBF event category. A multi-dimensional scan was made on jet and tau candidate kinematic variables to decide upon the best cut values to select VBF event candidates. The cut values chosen were those that maximised the significance of the signal (s/\sqrt{b}) for all 125 GeV Higgs signal and background processes in the range $100 < m_{MMC} < 150$ GeV. The background models used are those described in section 4.7.

The variables used in the scan were: leading and sub-leading jet momentum ($p_T^{j_1(j_2)}$), di-jet invariant mass (m_{j_1, j_2}), di-jet pseudo-rapidity separation ($|\Delta\eta_{j_1, j_2}|$), tau candidate transverse momentum (p_T^τ), missing energy (\cancel{E}_T), and the magnitude of the total transverse momentum of the system

$$p_T^{tot} = |\mathbf{p}_T^\ell + \mathbf{p}_T^{\tau_h} + \mathbf{p}_T^{j_1} + \mathbf{p}_T^{j_2} + \cancel{\mathbf{E}}_T| \quad (4.4)$$

Final cut values are shown in Table 4.5.

4.6.2. Boosted Category

The Boosted category is designed to exploit topologies with a boosted Higgs candidate. Such topologies provide good discrimination against non-resonant backgrounds in particular. The mass resolution for events in the Boosted category also is better than others due to the collinear nature of the tau decay products. This allows for greater separation between Higgs signal and the irreducible $Z \rightarrow \tau\tau$ background. Events that fail the VBF selection are subject to the Boosted categorisation.

To select Boosted events, the Higgs candidate transverse momentum, defined as

$$p_T^H := |\mathbf{p}_T^\ell + \mathbf{p}_T^{\tau_h} + \cancel{\mathbf{E}}_T| \quad (4.5)$$

was used. As this definition does not depend on the jet kinematics, it is not subject to corrections from the jet energy scale and resolutions (these corrections are a major source of uncertainties in the analysis, as discussed in Section 4.8). Selecting only events with $p_T^H > 100$ GeV (chosen again by studying s/\sqrt{b}) gave a high rejection rate of non-resonant backgrounds, including multi-jet and $t\bar{t}$ processes.

The boosted scenario also allows the use of the collinear approximation [66], as the assumption the tau decay products are all travelling in the same direction is more valid. It is possible to define the energy fractions of the visible decay products of the two tau decays as

$$x_{\ell(\tau_h)} = \frac{p_{vis}^{\ell(\tau_h)}}{p_{vis}^{\ell(\tau_h)} + p_{mis}^{\ell(\tau_h)}} \quad (4.6)$$

where p_{vis} is the momentum of the visible decay products of the tau decay, and p_{mis} is the invisible decay products (that is, the neutrinos). The x - and y -components of the missing energy can be written in terms of the energy fractions as

$$\cancel{E}_{T,x(y)} = \left(\frac{1}{x_\ell} - 1\right) p_{x(y)}^\ell + \left(\frac{1}{x_{\tau_h}} - 1\right) p_{x(y)}^{\tau_h} \quad (4.7)$$

Inverting these equations gives the energy fractions in terms of measurable quantities

$$x_\ell = \frac{p_x^{\tau_h} \cdot p_y^\ell - p_y^{\tau_h} \cdot p_x^\ell}{p_x^{\tau_h} \cdot \cancel{E}_{T,y} - p_y^{\tau_h} \cdot \cancel{E}_{T,x} + p_x^{\tau_h} \cdot p_y^\ell - p_y^{\tau_h} \cdot p_x^\ell} \quad (4.8)$$

and

$$x_{\tau_h} = \frac{p_x^{\tau_h} \cdot p_y^\ell - p_y^{\tau_h} \cdot p_x^\ell}{p_y^{\tau_h} \cdot \cancel{E}_{T,x} - p_x^{\tau_h} \cdot \cancel{E}_{T,y} + p_x^{\tau_h} \cdot p_y^\ell - p_y^{\tau_h} \cdot p_x^\ell} \quad (4.9)$$

These variables tend to 0.5 for signal type processes, while non-resonant backgrounds tend higher or lower, due to lack of or other sources of missing energy in the event. The optimised cut values for the Boosted category are found in Table 4.5.

4.6.3. n-Jet Categories

Events failing the VBF and Boosted categories fall into the 1- or 0-Jet categories. The events are separated into 1- and 0-Jet samples by cutting on the number of jets with $p_T > 25$ GeV. The 0-Jet category is very much dominated by $Z \rightarrow \tau\tau$ events, while the 1-Jet category has large fake- τ contributions, and these differences can be targeted with analysis level cuts. The n-jet categories are the least sensitive, however provide a large set of events with which to check the background models against data. Due to this large available data-set, the e and μ channels are treated separately in these categories. The optimised cut values for the n-jet categories are found in table 4.5.

4.6.4. Analysis Level Selection

After selecting categories based on specific event topologies, a series of *analysis level* cuts were placed to reject specific non-resonant background components, and also to define control regions for the various background estimation methods described in section 4.7.

Transverse Mass of Lepton and Missing Energy

The transverse mass of the lepton and missing energy is defined as

$$m_T = \sqrt{2p_T^\ell \cancel{E}_T [1 - \cos\Delta\phi(\ell, \cancel{E}_T)]} \quad (4.10)$$

where $\Delta\phi(\ell, \cancel{E}_T)$ is the angle between the lepton and the \cancel{E}_T vectors in the transverse plane of the detector. This variable is used to reject $W(\rightarrow \ell\nu) + \text{jets}$ and $t\bar{t}$ events which have high m_T , from signal events with low m_T . A cut of $m_T < 50$ GeV is placed for the VBF, Boosted, and 1-jet categories, with a lower cut of $m_T < 30$ GeV for the 0-jet category.

The m_T is also useful for defining control regions. The 0-jet category in the range $30 < m_T < 40$ GeV is used as a $Z \rightarrow ee$ control region, while the $m_T > 70$ GeV region in all categories is used to select $W(\rightarrow \ell\nu) + \text{jets}$ events for the so called *k-factor method* of background normalisation, described in Section 4.7.

Angular Separation of Tau Candidates

The angular separation of the visible tau decay products is given by

$$\Delta R = \sqrt{|\phi_{\tau_h} - \phi_\ell|^2 + |\eta_{\tau_h} - \eta_\ell|^2} \quad (4.11)$$

This separation is dependent on the boost of the parent Higgs candidate. The correlation between ΔR and p_T^H is different for signal, $Z \rightarrow \tau\tau$, and non-resonant background processes. Parameterising the dependence in signal events by a Landau function, an estimate of the predicted $\Delta R^{pred.}(p_T^H)$ is obtained. The deviation of the measured ΔR from the prediction is then given by

$$\Delta(\Delta R) = |\Delta R - \Delta R^{pred.}(p_T^H)| \quad (4.12)$$

Table 4.5 shows the $\Delta(\Delta R)$ cuts used for each analysis category.

Azimuthal Angles of Tau Decay Products

A measure of the angular separation between the visible tau decay products and the missing energy vector is given by

$$\sum \Delta\phi = |\phi_\ell - \cancel{\phi}| + |\phi_{\tau_h} - \cancel{\phi}| \quad (4.13)$$

where $\cancel{\phi}$ is the azimuthal angle of the missing energy in the transverse plane. Typically, the missing energy points between the visible decay products in ϕ for signal (and $Z \rightarrow \tau\tau$) events, giving $\sum \Delta\phi < \pi$. This variable provides extra discrimination power against $W(\rightarrow \ell\nu) + \text{jets}$ events which generally have much larger values of $\sum \Delta\phi$. Table 4.5 shows the $\sum \Delta\phi$ cuts used for each analysis category.

Fake Tau Candidates from e and μ

Suppression of $Z \rightarrow \ell\ell$ ($\ell = e/\mu$) events can be achieved through minimising the fake rate of $e/\mu \rightarrow \tau_h$ candidates. The lepton vetoes placed in the tau candidate object selection provide much of the power, however there are extra cuts that can be placed at the analysis level in the 1- and 0-jet categories, referred to as $Z \rightarrow \ell\ell$ rejection cuts in Table 4.5.

For muon fakes, which are often accompanied by an extra track due to Bremsstrahlung radiation, the electromagnetic fraction (f_{EM} ; the fraction of energy of the tau candidate deposited in the EM calorimeters) is usually lower than that of true tau candidates. A cut of $f_{EM} > 0.1$ is therefore applied to 2-prong tau candidates (before the tau object selection is placed) where the invariant mass of the track-muon pair is $80 < m_{\mu, \text{trk}} < 100$ GeV. The cut on the invariant mass was chosen to retain as many true tau candidates as possible.

Electrons fake taus at the highest rates in the very centre of the detector, $|\eta| < 0.05$, due to the lack of coverage of the TRT and EM calorimeters. To reject these fakes, any 1-prong tau candidates from this region are rejected.

Momentum Asymmetry of Tau Decay Products

The extra neutrino from the leptonic tau decay can cause a momentum asymmetry between the visible tau decay products. This is exploited in the 0-jet category, by requiring that $p_T^\ell - p_T^{\tau_h} < 0$.

	VBF	Boosted	1-Jet	0-Jet	Loose VBF
Category	$p_T^\tau > 30$ GeV	not VBF	not Boosted	not 1-Jet	$p_T^{j_1(j_2)} > 30$ GeV
	$\cancel{E}_T > 20$ GeV	$\cancel{E}_T > 20$ GeV	$\cancel{E}_T > 20$ GeV	$\cancel{E}_T > 20$ GeV	$\Delta\eta_{j_1, j_2} > 2$
	$p_T^{j_1(j_2)} > 40$ GeV	$p_T^H > 100$ GeV	$p_T^{j_1} > 25$ GeV	$p_T^{j_1} < 25$ GeV	$m_{j_1, j_2} > 200$ GeV
	$\eta_{j_1} \times \eta_{j_2} < 0$	$0 < x_\ell < 1$			
	$\Delta\eta_{j_1, j_2} > 3$	$0.2 < x_{\tau_h} < 1.2$			
	$m_{j_1, j_2} > 500$ GeV				
	jet centrality				
	$p_T^{tot} < 40$ GeV				
Analysis	$m_T < 50$ GeV	$m_T < 50$ GeV	$m_T < 50$ GeV	$m_T < 30$ GeV	
	$\Delta(\Delta R) < 1$	$\Delta(\Delta R) < 0.6$	$\Delta(\Delta R) < 0.6$	$\Delta(\Delta R) < 0.5$	
	$\Sigma\Delta\phi < 2.9$	$\Sigma\Delta\phi < 1.6$	$\Sigma\Delta\phi < 3.5$	$\Sigma\Delta\phi < 3.5$	
			$Z \rightarrow \ell\ell$ rejection	$Z \rightarrow \ell\ell$ rejection	
				$p_T^\ell - p_T^{\tau_h} < 0$	
$t\bar{t}$	$m_T > 50$ GeV b-tag	$m_T > 50$ GeV b-tag	$m_T > 50$ GeV b-tag	$m_T > 50$ GeV b-tag	$m_T > 50$ GeV b-tag
W		$m_T > 70$ GeV no b-tag no x_ℓ cut no x_{τ_h} cut $p_T^H > 50$ GeV	$m_T > 70$ GeV no b-tag	$m_T > 70$ GeV no b-tag	
$Z \rightarrow \tau\tau$	$m_T < 70$ GeV $40 < m_{\text{vis}} < 80$ GeV no b-tag				$m_T < 70$ GeV $40 < m_{\text{vis}} < 80$ GeV no b-tag
QCD					no iso. lepton $m_T < 50$ GeV no b-tag

Table 4.5.: Category and analysis cuts for the cut-based analysis. $e\tau_h$ and $\mu\tau_h$ channels are combined for VBF and Boosted categories, and treated separately for the 1– and 0–jet categories. The Loose VBF category is used in the fake-factor method (described in Section 4.7.6). $t\bar{t}$, W, $Z \rightarrow \tau\tau$ and QCD refer to control regions applied on top of the category selection for use in calculating k –factors (described in Section 4.7.1). Any cuts in the control regions supersede those in the categorisation.

Events with two real tau candidates (signal and $Z \rightarrow \tau\tau$) are assumed to have a negative momentum imbalance, while $W(\rightarrow \ell\nu) + \text{jets}$ and multi-jet events will peak in the positive.

Process	VBF	Boosted
ggF 125	$0.162 \pm 0.024 \pm \begin{smallmatrix} 0.038 \\ 0.023 \end{smallmatrix}$	$3.94 \pm 0.12 \pm \begin{smallmatrix} 0.22 \\ 0.23 \end{smallmatrix}$
VBF 125	$0.836 \pm 0.020 \pm \begin{smallmatrix} 0.109 \\ 0.101 \end{smallmatrix}$	$1.436 \pm 0.026 \pm 0.074$
WH 125	< 0.01	$0.528 \pm 0.036 \pm \begin{smallmatrix} 0.031 \\ 0.029 \end{smallmatrix}$
ZH 125	< 0.01	$0.297 \pm 0.021 \pm \begin{smallmatrix} 0.019 \\ 0.018 \end{smallmatrix}$
$Z \rightarrow \tau\tau$	$6.46 \pm 0.44 \pm \begin{smallmatrix} 1.47 \\ 1.30 \end{smallmatrix}$	$535 \pm 15 \pm \begin{smallmatrix} 26 \\ 25 \end{smallmatrix}$
Fake- τ	$1.23 \pm 0.22 \pm \begin{smallmatrix} 0.11 \\ 0.14 \end{smallmatrix}$	
SS-Data		$65.5 \pm 8.4 \pm 3.7$
$W(\rightarrow \ell\nu) + \text{jets}$		$34.3 \pm 5.5 \pm \begin{smallmatrix} 9.6 \\ 6.5 \end{smallmatrix}$
$t\bar{t}$	$0.82 \pm 0.24 \pm \begin{smallmatrix} 0.18 \\ 0.17 \end{smallmatrix}$	$41.3 \pm 1.4 \pm \begin{smallmatrix} 4.4 \\ 4.1 \end{smallmatrix}$
$Z \rightarrow \ell\ell$	$0.39 \pm 0.22 \pm \begin{smallmatrix} 0.20 \\ 0.23 \end{smallmatrix}$	$4.5 \pm 1.3 \pm \begin{smallmatrix} 1.5 \\ 1.7 \end{smallmatrix}$
Di-boson	$0.085 \pm 0.031 \pm \begin{smallmatrix} 0.022 \\ 0.031 \end{smallmatrix}$	$7.39 \pm 0.34 \pm \begin{smallmatrix} 0.68 \\ 0.73 \end{smallmatrix}$
Signal	$0.998 \pm 0.031 \pm 0.113$	$6.20 \pm 0.13 \pm \begin{smallmatrix} 0.23 \\ 0.24 \end{smallmatrix}$
Background	$9.31 \pm 0.60 \pm 1.59$	$688 \pm 18 \pm \begin{smallmatrix} 29 \\ 27 \end{smallmatrix}$
Data	10	692

Table 4.6.: Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to VBF filtered MC for the VBF category, and the embedding sample for all others. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method, and Fake- τ is the estimate of fakes from multi-jets and $W(\rightarrow \ell\nu) + \text{jets}$ contributions using the fake factor method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.

Process	1-Jet (e)	1-Jet (μ)	0-Jet (e)	0-Jet (μ)
ggF 125	$8.74 \pm 0.21 \pm \begin{smallmatrix} 0.75 \\ 0.86 \end{smallmatrix}$	$6.44 \pm 0.18 \pm \begin{smallmatrix} 0.57 \\ 0.63 \end{smallmatrix}$	$9.40 \pm 0.25 \pm \begin{smallmatrix} 0.61 \\ 0.71 \end{smallmatrix}$	$4.61 \pm 0.18 \pm \begin{smallmatrix} 0.41 \\ 0.53 \end{smallmatrix}$
VBF 125	$1.677 \pm 0.028 \pm \begin{smallmatrix} 0.088 \\ 0.093 \end{smallmatrix}$	$1.353 \pm 0.025 \pm \begin{smallmatrix} 0.065 \\ 0.067 \end{smallmatrix}$	$0.086 \pm 0.006 \pm \begin{smallmatrix} 0.009 \\ 0.007 \end{smallmatrix}$	$0.038 \pm 0.004 \pm \begin{smallmatrix} 0.003 \\ 0.002 \end{smallmatrix}$
WH 125	$0.449 \pm 0.034 \pm \begin{smallmatrix} 0.030 \\ 0.035 \end{smallmatrix}$	$0.402 \pm 0.032 \pm \begin{smallmatrix} 0.022 \\ 0.029 \end{smallmatrix}$	$0.027 \pm 0.008 \pm \begin{smallmatrix} 0.010 \\ 0.007 \end{smallmatrix}$	$0.029 \pm 0.008 \pm 0.004$
ZH 125	$0.285 \pm 0.020 \pm 0.020$	$0.264 \pm 0.019 \pm \begin{smallmatrix} 0.017 \\ 0.019 \end{smallmatrix}$	$0.017 \pm 0.005 \pm \begin{smallmatrix} 0.002 \\ 0.003 \end{smallmatrix}$	< 0.01
$Z \rightarrow \tau\tau$	$1478 \pm 21 \pm \begin{smallmatrix} 79 \\ 69 \end{smallmatrix}$	$1074 \pm 15 \pm 52$	$2357 \pm 25 \pm \begin{smallmatrix} 120 \\ 110 \end{smallmatrix}$	$784 \pm 13 \pm 35$
SS-Data	$811 \pm 28 \pm 41$	$308 \pm 19 \pm 18$	$396 \pm 20 \pm 20$	$57.7 \pm 8.2 \pm 3.4$
$W(\rightarrow \ell\nu)+\text{jets}$	$226 \pm 18 \pm \begin{smallmatrix} 34 \\ 33 \end{smallmatrix}$	$157 \pm 18 \pm \begin{smallmatrix} 31 \\ 24 \end{smallmatrix}$	$102 \pm 13 \pm \begin{smallmatrix} 11 \\ 10 \end{smallmatrix}$	$51.0 \pm 12.3 \pm \begin{smallmatrix} 9.3 \\ 3.8 \end{smallmatrix}$
$t\bar{t}$	$96.4 \pm 2.6 \pm \begin{smallmatrix} 4.3 \\ 4.8 \end{smallmatrix}$	$74.0 \pm 2.3 \pm 12.3$	$0.46 \pm 0.19 \pm \begin{smallmatrix} 0.17 \\ 0.18 \end{smallmatrix}$	$0.33 \pm 0.14 \pm \begin{smallmatrix} 0.16 \\ 0.12 \end{smallmatrix}$
$Z \rightarrow \ell\ell$	$34.1 \pm 4.8 \pm \begin{smallmatrix} 15.6 \\ 12.7 \end{smallmatrix}$	$15.6 \pm 2.9 \pm \begin{smallmatrix} 11.3 \\ 9.5 \end{smallmatrix}$	$48.4 \pm 5.5 \pm \begin{smallmatrix} 15.86 \\ 15.7 \end{smallmatrix}$	$12.3 \pm 2.8 \pm \begin{smallmatrix} 1.9 \\ 1.2 \end{smallmatrix}$
Di-boson	$12.3 \pm 1.1 \pm \begin{smallmatrix} 1.1 \\ 1.0 \end{smallmatrix}$	$8.9 \pm 1.0 \pm \begin{smallmatrix} 1.0 \\ 1.1 \end{smallmatrix}$	$2.13 \pm 0.58 \pm \begin{smallmatrix} 0.38 \\ 0.26 \end{smallmatrix}$	$2.23 \pm 0.30 \pm \begin{smallmatrix} 0.22 \\ 0.21 \end{smallmatrix}$
Signal	$11.15 \pm 0.22 \pm \begin{smallmatrix} 0.76 \\ 0.86 \end{smallmatrix}$	$8.45 \pm 0.19 \pm \begin{smallmatrix} 0.57 \\ 0.64 \end{smallmatrix}$	$9.53 \pm 0.25 \pm \begin{smallmatrix} 0.61 \\ 0.71 \end{smallmatrix}$	$4.68 \pm 0.18 \pm \begin{smallmatrix} 0.41 \\ 0.43 \end{smallmatrix}$
Background	$2658 \pm 40 \pm \begin{smallmatrix} 97 \\ 88 \end{smallmatrix}$	$1633 \pm 31 \pm \begin{smallmatrix} 66 \\ 62 \end{smallmatrix}$	$2896 \pm 35 \pm \begin{smallmatrix} 125 \\ 116 \end{smallmatrix}$	$907 \pm 19 \pm \begin{smallmatrix} 37 \\ 36 \end{smallmatrix}$
Data	2795	1688	3006	950

Table 4.7.: Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to the embedding sample. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.

4.7. Background Estimation Methods

The major background contributions for this analysis come from $Z \rightarrow \tau\tau$ and QCD multi-jet events. Wherever possible, a data-driven approach is used to model these events, allowing larger statistics samples with underlying event properties that more fully match those of true ATLAS data. The $Z \rightarrow \tau\tau$ events can be modelled using a method known as *embedding*. The QCD multi-jet estimates can be modelled by the OS-rSS method, which exploits the charge-correlation between the selected lepton and hadronic tau, or the fake-factor method, which uses a more direct measurement of the $jet \rightarrow \tau_h$ fake rate from data for use in the statistically limited VBF category. The aim of the QCD multi-jet estimations are not necessarily to increase statistics, but rather to better model the fake rates of $e \rightarrow \tau_h$, $\mu \rightarrow \tau_h$ and $jet \rightarrow \tau_h$ mis-reconstructions in data.

Where data-driven estimates cannot be made, MC event numbers are normalised to data in appropriate control regions (CRs) via the k -factor method. The CRs were chosen such that the purity of the background being studied was as high as possible, whilst maintaining a high efficiency. The cuts used to select the CRs are shown along with the category definitions in Table 4.5. All analysis variables of interest were checked in each CR to ensure an accurate modelling of the data with MC. Data-driven and MC-based methods are detailed below.

4.7.1. OS-rSS and the k -Factor Method

Except for the VBF category (in which the fake-factor method is used), the OS-rSS method provides the overall background estimation framework for the analysis. As the analysis requires the charges of the selected tau candidate and the lepton to have *opposite-sign* (OS), *same-sign* (SS) events can be used to model QCD multi-jet events. After accounting for events from other background processes that may be double-counted in the SS sample, the OS events are taken as estimates of all other background processes. The process of calculating any normalisation required for the MC samples is referred to as the k -factor method.

The charge correlation between the selected lepton and hadronic tau (i.e., $q_\ell q_{\tau_h}$) is different between the major backgrounds. Backgrounds with a real hadronic tau ($Z \rightarrow \tau\tau$, $t\bar{t}$, single- t and $VV \rightarrow \ell\tau_h X$) or those where a lepton fakes the hadronic tau ($Z/\gamma^*(\rightarrow \ell\ell) + jets$) have strong charge correlation, so the final signal region will mainly contain OS events where $q_\ell q_{\tau_h} < 0$. However, events where a jet fakes the hadronic tau (mainly QCD multi-jets, $W(\rightarrow \ell\nu) + jets$ and $t\bar{t}$ events) contain some charge asymmetry, resulting in SS events where $q_\ell q_{\tau_h} > 0$ in the

signal region. As the event selection *requires* $q_\ell q_{\tau_h} < 0$, the SS events from data are orthogonal to the final selection, and can be used to estimate the QCD multi-jets background.

It must be assumed that the shape of the m_{MMC} distribution for the QCD background in the signal region is the same for OS and SS events (this can be seen to be true by looking at the OS and SS events in data in Figure 4.8). The number of events in each selection however is not expected to be equal, as the charge correlation depends on the production mechanism of the event (QCD jets formed from either $q\bar{q}$, qq' , or $q\bar{q}'$ are expected to be charge correlated, while qg or gg processes are not). Thus, a correction factor $r_{QCD} = N_{OS}^{QCD} / N_{SS}^{QCD}$ is applied to account for any biases.

r_{QCD} Calculation

The calculation of r_{QCD} , described in [1], is obtained from a QCD-enriched control region. By selecting events with $\cancel{E}_T < 15$ GeV, $m_T < 30$ GeV, no requirement of lepton isolation, and a BDT_{loose} tau-ID (as described in Section 3.3.3), a set of events with $< 35\%$ contamination from electroweak and top processes is obtained (which are subsequently subtracted from MC estimates after appropriate normalisation corrections). The value r_{QCD} was calculated separately for the electron and muon channels in [1], as $r_{QCD}^{e\tau_h} = 1.000 \pm 0.051$ and $r_{QCD}^{\mu\tau_h} = 1.177 \pm 0.066$. Checks of the OS and SS event shapes of the key kinematic distributions in the QCD control region after subtracting electroweak and top contributions were conducted and are shown in Figure 4.8.

Total Background Estimate Using the OS-rSS Method

Using the method described above to estimate the QCD multi-jet contribution, and estimating the other backgrounds as described below, the number of background events is given by

$$N^{bkd} = r_{QCD} \cdot N_{SS}^{data} + N_{corr.}^{Z \rightarrow \tau\tau} + N_{corr.}^{W+jets} + N_{corr.}^{Z \rightarrow \ell\ell(\rightarrow\tau)} + N_{corr.}^{Z \rightarrow \ell\ell+jet(\rightarrow\tau)} + N_{corr.}^{t\bar{t}+single-t} + N_{corr.}^{VV} \quad (4.14)$$

where the terms $N_{corr.}^X$ refer to contributions from the other backgrounds (MC or the embedding sample) corrected for events already included in N_{SS}^{data} , and any MC-data discrepancies due to mis-modelling of the $e \rightarrow \tau_h$ and $jet \rightarrow \tau_h$ fake rates or other normalisation uncertainties. These contributions are generally estimated by

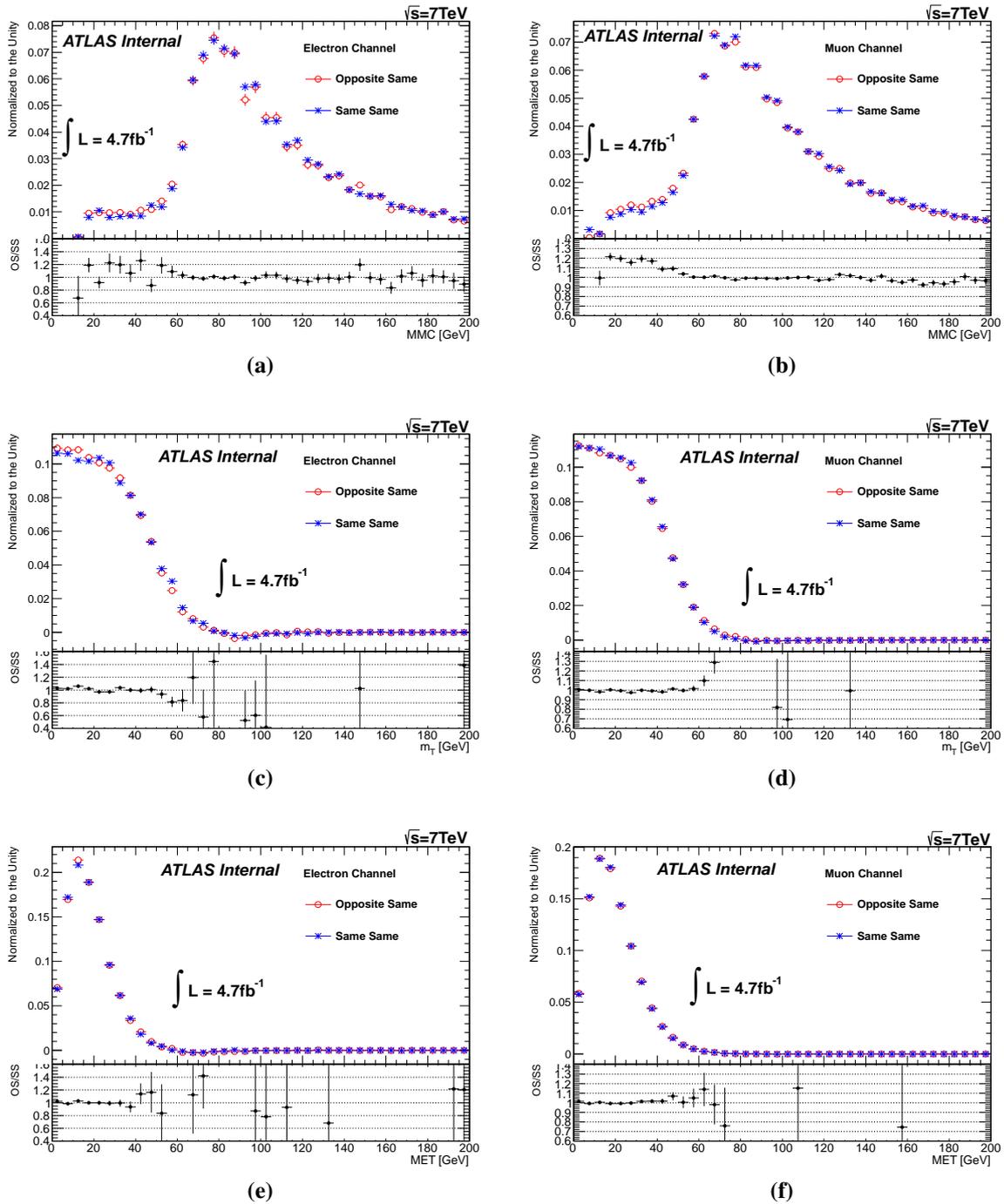


Figure 4.8.: Comparison of shapes for m_{MMC} , E_T and m_T distributions in OS and SS events from the QCD control region for the $e\tau_h$ (left) and $\mu\tau_h$ (right) channels. Contributions from electroweak and top backgrounds are subtracted from data and all distributions are normalised to the unit area. [1]

$$N_{corr.}^X = k_{OS}^X \cdot N_{OS}^X - r_{QCD} \cdot k_{SS}^X \cdot N_{SS}^X \quad (4.15)$$

for some background X , where a scale factor $k_{OS(SS)}$ provides the corrections described above, and is calculated using

$$k^X = \frac{N^{data} - N^{others}}{N^X} \quad (4.16)$$

As the OS and SS components of the backgrounds can differ in shape and number, this scale factor is calculated for both (that is, k_{OS} and k_{SS}).

The k -factors and their CR definitions are described below for all MC-based estimates. As the k -factors depend on other background estimations to be correctly calculated, the order in which they are estimated can make a difference. Therefore, the factors were calculated in order of $k^{Z\ell\ell} \rightarrow k^{top} \rightarrow k^W \rightarrow k^{Z\tau\tau}$, using the previous factor in each subsequent calculation. The following sections outline the specifics of the calculation for each background category. Table 4.8 shows the k -factors for each analysis category.

4.7.2. $Z \rightarrow \tau\tau$ Background Estimation

Two methods are used to estimate the $Z \rightarrow \tau\tau$ background contributions, both the k -factor method and the embedding method. As it is a data-driven approach, the embedding method would be ideal to use for all background categories. However, as the procedure is limited to the number of reconstructed $Z \rightarrow \mu\mu$ events in data and other selection efficiencies as described below, the final number of embedding events passing the VBF event selection is only ≈ 10 times the number of expected $Z \rightarrow \tau\tau$ events. Thus, a special VBF filtered $Z \rightarrow \tau\tau$ MC data-set was created. This filtering, as well as the embedding method, is explained below.

τ -Embedding in $Z \rightarrow \mu\mu$ ATLAS Data

The embedding method involves taking $Z \rightarrow \mu\mu$ events from ATLAS data (which, except for the mass difference between the muon and tau are kinematically identical to $Z \rightarrow \tau\tau$ events) and replacing the muons with simulated tau decays. The benefit of such a method is that the jet activity, underlying event properties, and pile-up conditions are taken directly from

Category	$k_{OS}^{t\bar{t}}$	$k_{SS}^{t\bar{t}}$	k_{OS}^W	k_{SS}^W	$k^{Z \rightarrow \tau\tau}$	k^{emb}	$k^{Z \rightarrow \ell\ell}$	$k^{Z/\gamma^*(\rightarrow\ell\ell)+jets}$
VBF	e	1.00 ± 0.06			0.88 ± 0.09		0.98 ± 0.24	1.00 ± 0.34
	μ						1.00 ± 0.00	
Boosted	e	0.80 ± 0.10	0.47 ± 0.23	0.58 ± 0.09	0.61 ± 0.13	0.104 ± 0.001	0.98 ± 0.24	0.56 ± 0.03
	μ					0.074 ± 0.001	1.00 ± 0.00	
1-Jet	e	0.86 ± 0.04	1.08 ± 0.17	0.64 ± 0.001	0.80 ± 0.03	0.104 ± 0.001	0.98 ± 0.24	0.56 ± 0.03
	μ	0.81 ± 0.04	0.99 ± 0.67	0.62 ± 0.01	0.73 ± 0.03	0.074 ± 0.001	1.00 ± 0.00	
0-Jet	e	0.86 ± 0.04	1.08 ± 0.17	0.57 ± 0.01	0.81 ± 0.03	0.104 ± 0.001	0.98 ± 0.24	0.56 ± 0.03
	μ	0.81 ± 0.04	0.99 ± 0.67	0.54 ± 0.01	0.70 ± 0.001	0.074 ± 0.001	1.00 ± 0.00	

Table 4.8.: k -factors for each analysis category and each lepton channel. Unless explicitly stated, factors apply to both OS and SS events. Uncertainties are statistical only.

data, and only the well understood tau decays need to be simulated. It must be noted that while this method models the shapes of the distributions correctly, the overall normalisation of these events is not, and so MC-based $Z \rightarrow \tau\tau$ events are still required for this purpose. The embedding procedure involves five main steps:

1. $Z \rightarrow \mu\mu$ events from data are selected by requiring two oppositely charged muons (satisfying the object selection described in 4.4), with $p_T > 15$ GeV, and no other tracks of up to 2% of the muon p_T within a cone of $\Delta R < 0.2$ of the muon track. The muon pair is also required to come from a common vertex, and have an invariant mass, $m_{\mu\mu} > 40$ GeV.
2. The kinematics of the muons are then extracted from the $Z \rightarrow \mu\mu$ event. At the truth level, the muons are replaced with truth taus, after modifying their four-momenta due to the muon-tau mass difference, such that

$$p_\tau = \sqrt{E_\mu^2 - m_\tau^2}$$

with the production vertex set to that of the muons.

3. The $Z \rightarrow \tau\tau$ kinematics are then simulated, using TAUOLA to decay the tau (including both leptonic and hadronic decays) and PHOTOS to add any photon radiation, producing a clean decay with no underlying event, calorimeter noise, or pile-up. A filter is applied to only use events with $p_T^{e/mu/\tau} > 15$ GeV before any simulation/digitisation/reconstruction to increase the statistics in the final analysis event selection. This $Z \rightarrow \tau\tau$ decay is known as a *mini event*.
4. Before implanting the $Z \rightarrow \tau\tau$ mini event, the muons must be removed from the $Z \rightarrow \mu\mu$ data event. A $Z \rightarrow \mu\mu$ mini event is simulated with the kinematics of the real muons, in order to subtract the associated energy deposits in the calorimeter. All muon tracks are also removed from the event. Now the $Z \rightarrow \tau\tau$ mini event is added to the data event, adding all tracks and calorimeter deposits due to the tau decays.
5. Finally, the event is re-reconstructed, in order to recalculate the missing energy and object properties of the event.

After producing the embedding sample, it needs to be appropriately normalised. Taking the number of events in the $Z \rightarrow \tau\tau$ ALPGEN MC as correct, the embedding sample is scaled separately in the e and μ channels by factors $k_e^{emb} = N_e^{Z \rightarrow \tau\tau MC} / N_e^{emb} = 0.104 \pm 0.001$ and $k_\mu^{emb} = N_\mu^{Z \rightarrow \tau\tau MC} / N_\mu^{emb} = 0.074 \pm 0.001$.

VBF Filtered $Z \rightarrow \tau\tau$ Monte-Carlo

To produce enough statistics for the VBF category, a VBF-filtered ALPGEN $Z \rightarrow \ell\ell$ ($\ell = e/\mu/\tau$) sample was created. The full VBF selection was not applied, rather, a relaxed selection was applied to the truth jets at the MC particle level. Jets were required to be within $|\eta| < 5.0$ with a $p_T > 15$ GeV, and be reconstructed by the anti- k_T algorithm within a cone of $\Delta R < 0.4$. After removing overlap between leptons and jets, the events were fully simulated if they had at least two selected jets, a pair of jets with an invariant mass $m_{j_1 j_2} > 200$ GeV and a separation of $|\Delta\eta_{j_1 j_2}| > 2.0$.

A scaling factor $k^{Z \rightarrow \tau\tau}$ was then determined using a $Z \rightarrow \tau\tau$ control region, defined with $m_T < 70$ GeV, a visible mass near the Z peak, $40 < m_{\text{vis}} < 70$ GeV and a b -tag veto. A value of $k^{Z \rightarrow \tau\tau} = 0.88 \pm 0.09$ was obtained.

4.7.3. $Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$ Background Estimation

To model the backgrounds originating from $Z \rightarrow \ell\ell$ events, different approaches are required depending on whether the fake τ candidate comes from an electron, muon, or jet. The different methods are detailed here.

$$Z \rightarrow ee(e \rightarrow \tau_h)$$

To model the $Z \rightarrow ee(e \rightarrow \tau_h)$ contribution, a CR is defined with the invariant mass of the electron and tau candidate being $80 < m_{\tau,e} < 100$ GeV, and the same selection as the 0-jet category without the \cancel{E}_T cut. Finally, only single-prong tau candidates are selected, and the transverse mass of the system is required to be within $30 < m_T < 40$ GeV to maintain orthogonality to the signal region and W CR. The data/MC ratio is then extracted as $k^{Z \rightarrow ee(e \rightarrow \tau_h)} = 0.98 \pm 0.24$.

$$Z \rightarrow \mu\mu(\mu \rightarrow \tau_h)$$

The $Z \rightarrow \mu\mu(\mu \rightarrow \tau_h)$ events provide a very small proportion of the total background ($< 1\%$). As such, no suitable CR could be found and so a normalisation factor of $k^{Z \rightarrow \mu\mu(\mu \rightarrow \tau_h)} = 1.0 \pm 0.0$ was used. The normalisation therefore is taken directly from the ALPGEN simulation, with any associated uncertainties.

$$Z \rightarrow \ell\ell(\text{jet} \rightarrow \tau_h)$$

The normalisation estimate of $Z \rightarrow \ell\ell(\text{jet} \rightarrow \tau_h)$ events is taken directly from the ALPGEN simulation, however, due to the known mis-modelling of the $\text{jet} \rightarrow \tau_h$ fake rates, a further correction is needed. A sample of two OS leptons (e/μ) is obtained from data, and then one of the associated jets is required to be reconstructed as a tau candidate, passing the same ID as the analysis selection. The fake rate here is independent of any charge or flavour correlations between the taus leptons. The rate is calculated in regions corresponding to each of the analysis categories; VBF ($p_T^{j_1} > 50$ GeV, $p_T^{\bar{\tau}} > 30$ GeV), Boosted ($p_T^H > 50$ GeV), and 1- and 0-jet categories ($p_T^{j_1} > 25$ GeV). After subtracting all other backgrounds, the $Z \rightarrow \ell\ell$ MC where there is no truth tau is scaled to match data. Due to low statistics in the VBF category, the normalisation factor was calculated to be $k_{VBF}^{Z \rightarrow \ell\ell(\text{jet} \rightarrow \tau_h)} = 1.00 \pm 0.34$, while all other categories agreed within statistical uncertainties to give $k_{others}^{Z \rightarrow \ell\ell(\text{jet} \rightarrow \tau_h)} = 0.56 \pm 0.03$.

4.7.4. $t\bar{t}$ and single- t Background Estimation

The background from $t\bar{t}$ processes includes events with high jet multiplicity, including those originating from b -quarks. In order to test the agreement between data and MC, a control region rich in $t\bar{t}$ events and orthogonal to the signal regions is selected. For each analysis category, the categorisation cuts described in Section 4.6 are firstly applied. In order to enhance the $t\bar{t}$ purity and ensure orthogonality to the signal regions however, the transverse mass cut is inverted ($m_T > 50$ GeV), and a b -tagged jet is required. The $t\bar{t}$ background is then scaled by a factor $k_{t\bar{t}}$ such that the overall background yield in the control region matches that of the data. This scaling is done separately for opposite-sign and same-sign events, with the obtained values shown in Table 4.8.

4.7.5. $W(\rightarrow \ell\nu) + \text{jets}$ Background Estimation

The $W(\rightarrow \ell\nu) + \text{jets}$ background primarily consists of a real lepton from the W decay, plus an associated jet that is misidentified as a hadronic tau candidate. The available MC sample overestimates the $W(\rightarrow \ell\nu) + \text{jets}$ contribution across all analysis categories. A data-driven approach to the $W(\rightarrow \ell\nu) + \text{jets}$ background estimate can be used with the fake-factor method in the VBF category, however a scaling of the MC to match the data yield in an appropriate CR is used in all other categories. After the base categorisation cuts are applied, a W CR is selected by placing a transverse mass cut of $m_T > 70$ GeV, to select for the high neutrino content of

$W(\rightarrow \ell \nu) + \text{jets}$ events. To increase the available statistics in the Boosted category, the x_ℓ and x_{τ_h} cuts are dropped, and the Higgs candidate momentum cut is relaxed to $p_T^H > 50$ GeV.

As explained in section 4.7.1, hadronic tau fake rates for quark jets (dominant in OS events) and gluon jets (dominant in SS events) differ, and as such $k^{W(\rightarrow \ell \nu) + \text{jets}}$ is calculated separately for OS and SS events. Calculations are also done separately for $e\tau_h$ and $\mu\tau_h$ channels, and are shown in Table 4.8.

4.7.6. Fake Factor Method

In the OS-rSS method, SS events were used to estimate the number events with a fake-hadronic tau originating largely from QCD multi-jet events. A significant number of fake- τ_h candidates however also come from $W(\rightarrow \ell \nu) + \text{jets}$ processes. Due to the $W(\rightarrow \ell \nu) + \text{jets}$ MC sample in the VBF category being statistically limited, along with the potential mis-modelling of the tau fake rates, a data-driven way of estimating these events is also desired. One such way is the fake-factor (FF) method.

The FF method provides a way of estimating the number of fake- τ_h events in the signal region (SR) by assuming that the hadronic tau fake rate can be extrapolated from events where the hadronic tau candidate fails the the ID requirements of the selection. An anti- τ (or $\bar{\tau}$) sample is therefore obtained by inverting the tau-ID requirements in the SR data. The number of these events $N_{\bar{\tau}}^{data}$, when multiplied by the calculated fake factor f , give the estimated number of fake- τ_h events in the SR;

$$N_{est}^{fake} = f \cdot N_{\bar{\tau}}^{data} \quad (4.17)$$

As there are two sources of fake- τ_h s, two fake factors are calculated; f_W and f_{QCD} for fakes from $W(\rightarrow \ell \nu) + \text{jets}$ and QCD multi-jet processes respectively. The SR contains both of these events in some ratio r_W , and so a mixed fake factor f_{MIX} is obtained for the final fake estimate;

$$f_{MIX} = r_W \cdot f_W + (1 - r_W) \cdot f_{QCD} \quad (4.18)$$

	Signal Region	W Control Region	QCD Control Region
Selection	isolated lepton	isolated lepton	no isolated lepton
	$m_T < 50$ GeV	$m_T > 70$ GeV	$m_T < 50$ GeV
	$p_T^{j_1(j_2)} > 30$ GeV	$p_T^{j_1(j_2)} > 30$ GeV	$p_T^{j_1(j_2)} > 30$ GeV
	no b-tag	no b-tag	no b-tag
	$\Delta\eta_{j_1, j_2} > 2$	$\Delta\eta_{j_1, j_2} > 2$	$\Delta\eta_{j_1, j_2} > 2$
	$m_{j_1, j_2} > 200$ GeV	$m_{j_1, j_2} > 200$ GeV	$m_{j_1, j_2} > 200$ GeV

Table 4.9.: Selection regions for the FF calculation.

The values f_W , f_{QCD} , and r_W are calculated using a loose VBF category selection to increase the available statistics, requiring only two leading jets with $p_T^{j_1(j_2)} > 30$ GeV, $\Delta\eta_{j_1, j_2} > 2$, and $m_{j_1, j_2} > 200$ GeV to account for $Z \rightarrow \tau\tau$ MC filtering as discussed in Section 4.7.2, and a b -tag veto. The requirements on lepton isolation and the m_T of the system depend on which control region is being selected, and are shown in Table 4.9.

Figure 4.9 shows the m_{vis} distributions for the hadronic tau and anti- τ events in each FF control region. The large discrepancy between data and MC in the QCD CR comes from the fact that there is not yet any QCD multi-jet estimation included in the model. It is clear from these plots that the number of events available in the anti- τ sample is far greater than the hadronic tau sample, however in the W and QCD CRs, the relative amounts of each background contribution remains similar.

r_W Calculation

The ratio of $W(\rightarrow \ell\nu) + \text{jets}$ to QCD multi-jet events in the SR can be estimated by

$$r_W = \frac{N_{est}^W}{N_{est}^W + N_{est}^{QCD}} \tag{4.19}$$

Again using the anti- τ sample and assuming a similar background composition to the hadronic tau events, N_{est}^W can itself be calculated by

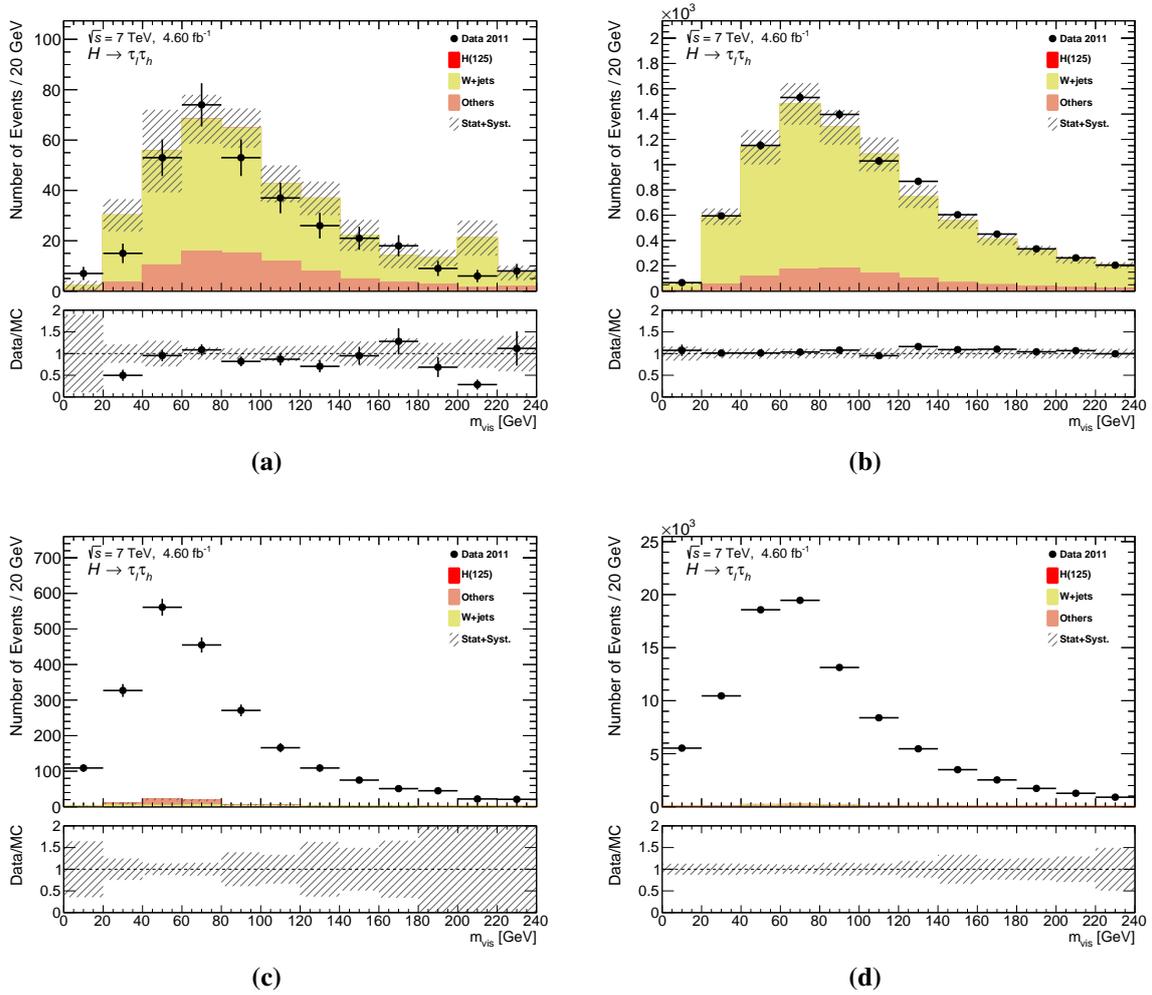


Figure 4.9.: Tau and anti- τ_h m_{vis} distributions for the loose VBF W CR in (a) and (b), and the QCD CR in (c) and (d). Difference between MC and data is due to missing QCD multi-jet MC. Uncertainties are statistical and systematic.

E_T^τ [GeV]	CR	$N^{\bar{\tau}data}$	$N^{\bar{\tau}WMC}$	$N^{\bar{\tau}XMC}$	N_{est}^W	N_{est}^{QCD}	r_W
(30, 40)	W	2479 ± 50	1667 ± 35				
	SR	2973 ± 55	1149 ± 29	431 ± 5	1709 ± 66	833 ± 86	0.67 ± 0.02
(40, 60)	W	2394 ± 49	1573 ± 34				
	SR	2574 ± 51	1017 ± 25	488 ± 5	1547 ± 60	539 ± 79	0.74 ± 0.03
(60, 80)	W	981 ± 31	673 ± 20				
	SR	1027 ± 32	415 ± 21	230 ± 3	604 ± 40	193 ± 52	0.76 ± 0.05
(80, 200)	W	939 ± 31	754 ± 23				
	SR	957 ± 31	512 ± 17	231 ± 3	637 ± 35	89 ± 47	0.88 ± 0.06

Table 4.10.: Number of events in each background and region required for the calculation of r_W , binned by E_T^τ . Uncertainties are statistical only.

$$N_{est}^W = N_{WCR}^{\bar{\tau}data} \times \frac{N_{SR}^{\bar{\tau}WMC}}{N_{WCR}^{\bar{\tau}WMC}} \quad (4.20)$$

where $N_{WCR}^{\bar{\tau}data}$ is the number of anti- τ events in the W CR data, and the multiplicative factor is the selection efficiency of anti- τ candidates in W MC between the signal region and the W CR.

The estimated number of QCD events is calculated taking anti- τ data in the SR, less the N_{est}^W and other background components $N_{SR}^{\bar{\tau}XMC}$,

$$N_{est}^{QCD} = N_{SR}^{\bar{\tau}data} - (N_{est}^W + N_{SR}^{\bar{\tau}XMC}) \quad (4.21)$$

As the fake rate is expected to change with the tau energy, r_W is calculated in bins of E_T^τ . The numbers of events in each background and control region are shown in Table 4.10, along with the value of r_W in each bin. The r_W curve is also shown in Figure 4.10.

f_{MIX} Calculation

The calculation of f_W and f_{QCD} are carried out in their respective control regions as per Equation 4.17, in bins of E_T^τ as with r_W . Table 4.11 shows the number of tau and anti- τ events

E_T^τ [GeV]	CR	N_τ	$N_{\bar{\tau}}$	f
(30,40)	W	58 ± 11	1940 ± 50	0.030 ± 0.006
	QCD	479 ± 22	25562 ± 161	0.019 ± 0.001
	MIX			$0.026 \pm 0.004^{+0.004}_{-0.007}$
(40,60)	W	38 ± 10	1774 ± 49	0.021 ± 0.006
	QCD	284 ± 17	20687 ± 145	0.014 ± 0.001
	MIX			$0.019 \pm 0.004^{+0.002}_{-0.006}$
(60,80)	W	18 ± 7	692 ± 32	0.026 ± 0.010
	QCD	106 ± 11	7195 ± 86	0.015 ± 0.001
	MIX			$0.023 \pm 0.008^{+0.003}_{-0.008}$
(80,200)	W	5 ± 5	673 ± 31	0.008 ± 0.008
	QCD	79 ± 9	6076 ± 79	0.013 ± 0.002
	MIX			$0.008 \pm 0.007^{+0.005}_{-0.001}$

Table 4.11.: Number of tau and anti- τ candidates in each FF control region, and the final $f_{W/QCD/MIX}$ numbers for each E_T^τ bin. Uncertainties are statistical only except for the f_{MIX} values which have a systematic uncertainty based on varying the r_W value up and down.

in each control region, in each bin of E_T^τ . The fake factor for each region is also given, with the final f_{MIX} as calculated per Equation 4.18. The fake factor curves are also shown in Figure 4.10.

The statistical uncertainty on the FF measurement is estimated by taking the extremes of f_{MIX} by assuming a maximum uncertainty on the r_W measurement, effectively saying the SR contains only either $W(\rightarrow \ell\nu) + \text{jets}$ or QCD multi-jet fakes exclusively. If more W MC was available, an independent validation for r_W could be conducted reducing its uncertainty, and in turn, the FF uncertainty.

Total Background Estimate Using the FF Method

Incorporating the fake- τ_h estimate into the total background model now gives

$$N^{bkd} = (N_{\bar{\tau}}^{data} - N_{\bar{\tau}}^{MC}) \times f_{MIX} + N_{\bar{\tau}}^{MC} \tag{4.22}$$

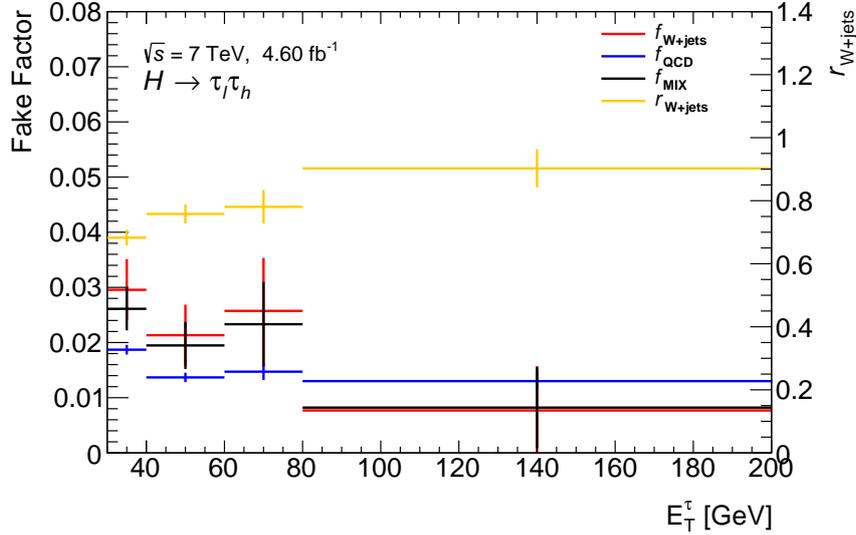


Figure 4.10.: Fake Factors (left axis) and r_W (right axis) values as a function of tau transverse energy.

where $N_{\bar{\tau}/\tau}^{MC}$ represents events from $Z \rightarrow \tau\tau$, $Z \rightarrow \ell\ell$, $t\bar{t}$, single- t , and $VV \rightarrow \ell\tau_h X$ backgrounds, in either the $\bar{\tau}$ or tau selected samples. The remaining backgrounds modelled by MC are corrected by a k_{OS}^X as calculated in Section 4.7.1; $k_{OS}^{t\bar{t}} = 1.00 \pm 0.06$, $k_{OS}^{Z \rightarrow \tau\tau} = 0.88 \pm 0.09$, and $k_{Z/\gamma^*(\rightarrow \ell\ell)+jets} = 1.00 \pm 0.34$.

Figure 4.11 shows the number of hadronic tau events in the W and QCD CRs, as well as the SR, after using f_W , f_{QCD} , and f_{MIX} for estimating the fake- τ_h contribution. A good agreement is seen across all regions, validating the assumption that the background composition in the anti- τ and hadronic tau samples are similar, and that the calculation of r_W is sound. Also shown is the FF SR region using the OS-rSS method for estimating the fake- τ_h contribution. The OS-rSS method overestimates the fake backgrounds in the region where a Higgs signal may be expected, as well as providing low event yields in the tails of the distribution.

4.8. Systematic Uncertainties

The main source of uncertainty in this analysis is statistical, due to the lack of available MC events. Other uncertainties arise however from both the theoretical models used to produce the background samples, and also experimental uncertainties from the ATLAS detector. This section outlines the various sources of systematic uncertainties. Section 4.9 explains how these uncertainties enter the limit extraction procedure as *nuisance parameters*.

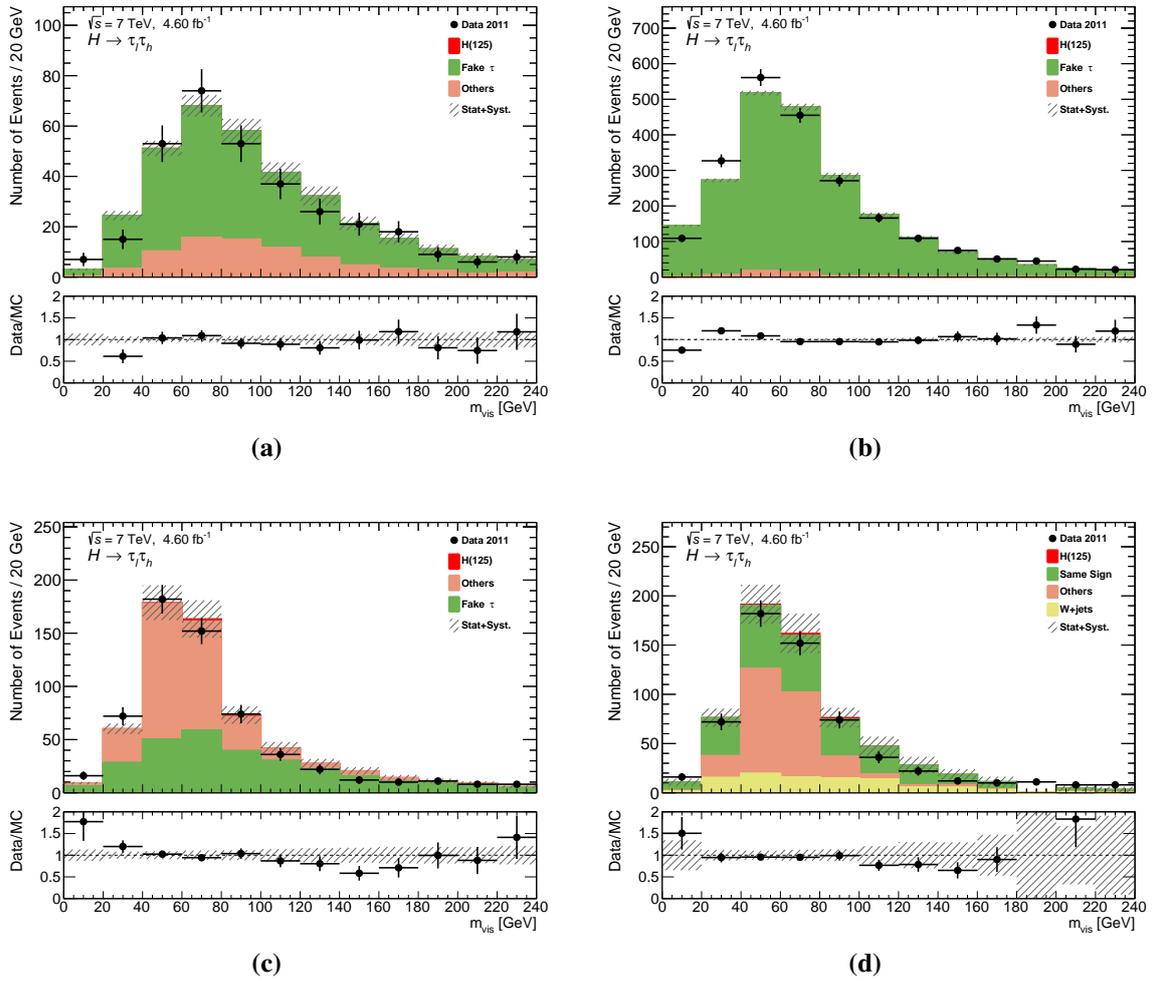


Figure 4.11.: m_{vis} distributions for (a) the W control region using f_{W+jets} , (b) the QCD control region using F_{QCD} , and (c) the loose VBF region using f_{MIX} . Uncertainties are statistical and systematic. Plot (d) shows the loose VBF region selection using the OS-rSS method. All other background contributions are included in the *others* category.

m_H [GeV]	\mathcal{B} Uncertainty [%]
110	+6.9 -6.7
115	+6.4 -6.3
120	+6.1 -6.0
125	+5.7 -5.7
130	+5.3 -5.2
135	+4.8 -4.8
140	+3.6 -3.6
145	+3.3 -3.3
150	+3.0 -3.1

Table 4.12.: Higgs branching ratio (\mathcal{B}) uncertainty variation (%) for each simulated Higgs mass point [68].

The size of the uncertainties for the VBF and Boosted analysis categories are shown as an example in Tables 4.14 and 4.15, giving the changes in the nominal event yields by varying the systematic up and down. The systematics are ordered by the impact of varying the parameter up and down on the final value of the signal strength, extracted by the fit described in Section 4.9. It can be seen that jet energy scale and resolution uncertainties, as well as the background estimation uncertainties generally are the most important systematics.

4.8.1. $H \rightarrow \tau\tau$ Branching Fraction Uncertainties

Provided by the Higgs cross-section working group [68], uncertainties are assigned to the branching fraction of $H \rightarrow \tau\tau$ (when referring to this uncertainty later, it will be identified as BR_atautau) for each simulated mass point, as shown in Table 4.12.

4.8.2. QCD Scale and PDF Uncertainties

To account for missing higher order QCD corrections to the Higgs production cross-sections, QCD scale uncertainties are provided by the LHC Higgs cross-section working group [69]. Factorisation and renormalisation scales are varied around the nominal value to estimate the uncertainty. Table 4.13 shows the percentage variation in the cross-section for each signal topology and each sample Higgs mass.

m_H [GeV]	QCDscale_qqH	QCDscale_ggH	QCDscale_ZH	QCDscale_WH
100	+0.4 -0.3	+7.8 -8.4	+1.1 -1.2	+0.6 -0.5
105	+0.3 -0.3	+7.7 -8.3	+1.3 -1.6	+0.3 -0.8
110	+0.5 -0.2	+7.5 -8.1	+1.2 -1.2	+0.3 -0.7
115	+0.2 -0.2	+7.4 -8.0	+1.3 -1.2	+0.4 -0.8
120	+0.3 -0.4	+7.2 -7.9	+1.5 -1.2	+0.4 -0.7
125	+0.3 -0.3	+7.1 -7.8	+1.4 -1.6	+0.2 -0.8
130	+0.3 -0.2	+7.0 -7.7	+1.5 -1.4	+0.3 -0.8
135	+0.5 -0.1	+6.9 -7.6	+1.7 -1.4	+0.7 -0.4
140	+0.2 -0.2	+6.8 -7.5	+1.5 -1.6	+0.5 -0.5
145	+0.4 -0.0	+6.7 -7.5	+1.8 -1.8	+0.2 -0.8
150	+0.2 -0.1	+6.6 -7.4	+1.8 -1.6	+0.4 -0.8

Table 4.13.: QCDscale uncertainty variations (%) for the VBF (QCDscale_qqH), ggF (QCDscale_ggH), and VH (QCDscale_ZH and QCDscale_WH) signal topologies, for each simulated Higgs mass point [69].

The PDF uncertainties are also provided by the Higgs cross-section working group. The variation of the signal production cross-sections when using different PDF sets are obtained from the MCFM generator [70]. The uncertainty on the ggF cross-section (pdf_gg) is calculated to be $\pm 8\%$, while the uncertainties on the VBF and VH production modes (pdf_qq) are $\pm 4\%$.

4.8.3. Electron Uncertainties

There are multiple sources of systematic uncertainties considered for electrons; energy scale and resolution, identification efficiency, and calorimeter isolation uncertainties [71]. For the electron energy scale (E_SCALE), a 1(3)% uncertainty is placed on the electromagnetic calorimeter energy clusters in the central (forward) region, with the calibration calculated in ATLAS data by using well studied $Z \rightarrow ee$ and $J/\Psi \rightarrow ee$ resonances [72]. To account for mis-modelling in the electron energy resolution (E_RES), a 1–4% energy smearing is applied to all energy clusters, again calculated by measuring the di-electron invariant mass from Z decays in ATLAS data.

Electron identification efficiency uncertainty has been measured to be approximately 3% using tag-and-probe⁵ methods in $Z \rightarrow ee$ data [33], while the calorimeter isolation uncertainty is propagated to the final event yields by varying the isolation energy correction by $\pm 2(\pm 4)\%$ for electrons with $p_T > (<) 20$ GeV. These uncertainties are treated as correlated, and enter the analysis together as an electron efficiency term, E_EFF.

4.8.4. Muon Uncertainties

Muon systematic uncertainties enter through varying the muon momentum resolution (M.RES) and identification efficiencies (M.EFF). A percentage uncertainty of $\pm 1\%$ is applied for M.EFF, calculated from a tag-and-probe analysis of $Z \rightarrow \mu\mu$ decays in ATLAS data [73]. Similarly, a $\pm 1\%$ uncertainty is assigned to M.RES, by measuring the width of $Z \rightarrow \mu\mu$ and $W \rightarrow \mu\nu_\mu$ decays in ATLAS data [62].

4.8.5. Tau Uncertainties

Hadronically decaying taus will decay inside the hadronic calorimeter, and as such a correct calibration of the tau jet energy scale is required. The systematic uncertainties of the tau energy scale TES are included in the analysis by smearing the p_T of all tau jets passing the object selection [74]. Corrections to the hadronic tau energy scale from the jet energy scale, are derived from simulated $Z \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$ events, providing corrections to be applied to the energy recorded by ATLAS. An uncertainty of on average 3% is assigned to the TES, which is assumed to be correlated with the jet energy scale (JES – see the next section), and so these uncertainties are added together in quadrature.

Tau identification efficiency uncertainties (T.EFF) were calculated for the various tau identification algorithms, again using a tag-and-probe approach [53]. For the BDTmedium algorithm used in this analysis, uncertainties are assigned to 4(8)% for taus with $p_T > 22(20 < p_T < 22)$ GeV.

⁵tag-and-probe refers to the method of taking a well known decay (such as $Z \rightarrow ee$), *tagging* one leg using tight identification algorithms, and *probing* the other leg using whichever identification method is to be tested.

4.8.6. Jet and MET Uncertainties

Jet property uncertainties arise from both global energy scale effects, and due to uncertainties in the flavour composition of each event. A conservative estimate of the jet energy scale (JES) uncertainty, derived from studying the average calorimeter energy response of calibrated jets, was applied by smearing the p_T of all jets passing the object selection by 7–10% [64]. As mentioned in the previous section, the hadronic tau energy scale uncertainty TES is added in quadrature to the jet energy scale uncertainty, and both are used as a correlated uncertainty JES_TES. The effect of the jet energy resolution (JER) uncertainty on the event yields, estimated using simulated di-jet data to be up to 10% [75], was also studied by smearing the jet p_T s before event selection.

Uncertainties in the electron, hadronic tau, and jet energy scales and resolutions are propagated to the missing energy uncertainty (by varying their contributions in Equation 3.6) as MET_SCALE and MET_RES.

4.8.7. Trigger and Luminosity Uncertainties

Systematic uncertainties on the electron trigger (E_TRIG) [76] and muon trigger (M_TRIG) [77] efficiencies are found to be of order 1%, obtained by varying the object selections in the efficiency measurements, and also studying the effects of using different MC generators in the simulated data. These uncertainties are added in quadrature to the E(M)_EFF terms in the final treatment.

A total uncertainty on the integrated luminosity of the 2011 run is taken as $\pm 1.8\%$, estimated using various luminosity detectors and algorithms [37].

4.8.8. Background Estimate Uncertainty

The data-driven background estimate in Section 4.7 gives rise to systematic uncertainties for each of the calculated k -factors (K_X , where X is the background component), as well as r_{QCD} (RQCD) and f_{MIX} (FF). The statistical uncertainty on the k -factor calculation is used as the systematic variation on the background estimate (as per the analysis outlined in [1]), with each shown in Table 4.8. The uncertainty on r_{QCD} is $\pm 5.1(\pm 5.6)\%$ for electron (muon) events. As explained in Section 4.7.6, the systematic uncertainty on the fake factor measurement is taken

as the maximum variation in the r_W estimate, leading to approximately $\pm 30\text{--}60\%$ uncertainty depending on the tau energy bin.

4.9. Limit Setting Procedure

This section outlines how an exclusion limit for the production cross-section of the SM Higgs boson, normalised to the theory cross-section, is determined. The *signal strength* parameter $\mu = \sigma / \sigma_H^{SM}$ gives a measure of deviations from the SM prediction. The question then asked is what is the largest signal strength that can exist, given the observed data? The approach taken is the standard method used by ATLAS to extract exclusion limits [78].

4.9.1. Exclusion Limits and the Signal Strength

To test the agreement of the observed data d_i to the model prediction p_i in some N -binned distribution, a *binned likelihood function* is defined as the product of the Poisson probabilities in each bin i ;

$$\mathcal{L} = \prod_i^N \frac{p_i^{d_i}}{d_i!} e^{-p_i} \quad (4.23)$$

This is a simplistic view of the likelihood, as we must also include the systematic uncertainties of the model into the test. These uncertainties enter as *nuisance parameters* θ (note that this is a vector of parameters), applied to the background b_i and signal s_i models. The model prediction in each bin therefore, depending on the signal strength μ (where $\mu = 1$ defines the nominal signal plus background hypothesis) can be given by

$$p_i(\mu, \theta) = b_i(\theta) + \mu s_i(\theta) \quad (4.24)$$

Now, the likelihood function can be written as

	VBF	ggF	$Z \rightarrow \tau\tau$	Fake- τ	Others
NOM	0.836	0.162	6.455	1.277	1.573
STAT	± 0.020	± 0.024	± 0.437	± 0.224	± 0.348
SYST	+0.109 -0.110	+0.028 -0.023	+1.467 -1.296	+0.503 -0.512	+0.297 -0.316
JER	± 0.003	± 0.002	± 0.590		± 0.130
JES_TES	+0.101 -0.102	+0.021 -0.022	+1.127 -0.865		+0.238 -0.086
FF				+0.486 -0.489	
E_RES	-0.001 -0.002		-0.036 -0.041		+0.000 -0.129
E_SCALE	+0.002 -0.000		+0.000 -0.043		-0.000 -0.015
MET_RES	+0.005 -0.002	+0.003 -0.000	-0.043 +0.212		-0.021 -0.049
K_ZTT			± 0.660		
MET_SCALE	-0.011 +0.014	-0.016 +0.003	-0.059 +0.049		-0.001 -0.159
K_TT					± 0.066
K_ZLL_JET					± 0.091
K_ZLL					+0.091 -0.028
T_EFF	± 0.033	± 0.006	± 0.256		± 0.024
M_EFF	± 0.006	± 0.001	± 0.043		± 0.009
BR_tautau	± 0.048	± 0.009			
M_SCALE	+0.000 -0.001	± 0.000	+0.000 -0.007		+0.000 -0.001
E_EFF	+0.002 -0.006	+0.000 -0.002	+0.094 -0.131		+0.012 -0.014
pdf_qq	± 0.033				
QCDScale_ggH		+0.012 -0.013			
pdf_gg		± 0.013			
QCDScale_qqH	± 0.003				

Table 4.14.: Statistical and systematic uncertainties as on the nominal event yields for the VBF category, ordered by the size of their influence on the fitted signal strength. NOM, STAT, and SYST refer to the nominal yield, statistical uncertainty, and total systematic uncertainty respectively.

	VBF	ggF	WH	ZH	Embedding	SS Data	Others
NOM	1.436	3.940	0.528	0.297	535.209	65.540	89.655
SYST	± 0.026	± 0.115	± 0.036	± 0.021	± 14.516	± 8.349	± 5.882
STAT	± 0.073	+0.215 -0.226	+0.031 -0.029	+0.019 -0.018	+26.262 -25.292	± 3.709	+15.367 -10.469
JER	± 0.009	± 0.019	± 0.006	± 0.003			± 2.971
JES_TES	+0.021 -0.020	+0.099 -0.114	+0.017 -0.010	+0.011 -0.009			+10.908 -7.576
K_EMB					± 3.030		
K_TT							± 3.151
K_W							± 4.280
K_ZLL_JET							± 0.010
K_ZLL							± 0.010
T_EFF	± 0.060	± 0.161	± 0.022	± 0.012	± 22.163		± 1.246
MET_SCALE	-0.002 +0.004	-0.014 -0.009	+0.004 +0.002	-0.005 +0.003			+3.826 -2.5163
RQCD					± 0.195	± 3.462	± 1.794
M_EFF	± 0.011	± 0.028	± 0.004	± 0.002	± 11.063	± 0.721	
E_RES	+0.002 +0.000	+0.017 +0.007	-0.001 +0.005	+0.003 +0.001	-0.955 -0.453		-0.149 +0.988
E_SCALE	+0.002 +0.002	-0.001 +0.004	+0.002 +0.005	+0.001 -0.000	+0.978 +0.181		+0.545 +1.562
BR_tautau	± 0.082	± 0.225	± 0.030	± 0.017			
M_SCALE	+0.000 -0.002	-0.003 -0.004	+0.000 -0.001	-0.000 -0.000	+0.539 -0.637		-0.005 -0.116
pdf_gg	± 0.115						
E_EFF	+0.004 -0.009	+0.011 -0.025	+0.001 -0.004	+0.001 -0.002	± 4.704	+0.337 -0.451	
QCDScale_ggH	+0.102 -0.112						
MET_RES	-0.003 -0.006	-0.027 -0.047	+0.001 +0.001	+0.003 -0.000			-1.619 -1.237
pdf_qq		± 0.158	± 0.021	± 0.012			
QCDScale_WH			+0.001 -0.004				
QCDScale_ZH				+0.004 -0.005			
QCDScale_qqH	± 0.004						

Table 4.15.: Statistical and systematic uncertainties as on the nominal event yields for the Boosted category, ordered by the size of their influence on the fitted signal strength. NOM, STAT, and SYST refer to the nominal yield, statistical uncertainty, and total systematic uncertainty respectively.

$$\mathcal{L}(\mu, \theta) = \prod_i^N \frac{(b_i(\theta) + \mu s_i(\theta))^{d_i}}{d_i!} e^{-(b_i(\theta) + \mu s_i(\theta))} \quad (4.25)$$

The best possible fit to the data is found by taking the values of μ and θ that maximise \mathcal{L} . This is known as the *unconditional maximum likelihood* $\mathcal{L}(\hat{\mu}, \hat{\theta})$, with $\hat{\mu}$ being the observed signal strength and $\hat{\theta}$ the best values for the nuisance parameters. For any other value of μ , the *conditional maximum likelihood* $\mathcal{L}(\mu, \hat{\theta})$ can be found by choosing $\hat{\theta}$ to again maximise \mathcal{L} . The ratio of these two likelihoods $\lambda(\mu)$ is the *profile likelihood ratio*

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \theta)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \quad (4.26)$$

with a value of $\lambda(\mu) \rightarrow 1$ meaning a good agreement between the data and the best possible model for the specified signal strength. The value of the likelihood ratio is then used to construct a test statistic q_μ as

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu) & \text{if } \hat{\mu} \leq \mu \\ 0 & \text{if } \hat{\mu} > \mu \end{cases} \quad (4.27)$$

The smaller the observed signal strength $\hat{\mu}$ is with respect to the hypothesised signal strength μ , the greater the incompatibility between the model and the data, and the larger q_μ becomes. In the region where $\hat{\mu} > \mu$, then the signal hypothesis is still valid, and $q_\mu = 0$. From here, a p -value can be calculated to give a measure of the probability of obtaining a result at least as large as the observed $\hat{\mu}$;

$$p_\mu = \int_{q_{\hat{\mu}}}^{\infty} f(q_\mu | \mu) dq_\mu \quad (4.28)$$

where $f(q_\mu | \mu)$ is the pdf of q_μ . Now it is possible to obtain an *upper limit* on the signal strength by finding the largest value of μ excluded by the data. The value of μ that gives $p_\mu = 0.05$ defines an upper limit at the 95% confidence level, which is interpreted as there being less than a 5% chance of missing as signal at least this large in the data.

In order to optimise the analysis, an expectation of the what the exclusion limit may be, without looking at the data, is required. To achieve this, *pseudo data* can be constructed from the modelled background and signal predictions as

$$d_i^A(\mu) = b_i + \mu s_i \quad (4.29)$$

where $d_i^A(\mu)$ is the *Asimov data* [78]. The most tightly constrained exclusion limit is then calculated by testing $d_i^A(0)$ in place of the observed data. By assuming the background-only model, the test is mainly sensitive to how well the nuisance parameters can be constrained in the fit, giving a handle on how to improve the model.

This Asimov data is used to create 95% confidence level exclusion limits⁶ on σ/σ_H^{SM} , that is, a limit on how sensitive the analysis is to the SM Higgs boson cross section (σ_H^{SM}). For instance, an exclusion limit of $3 \times \sigma_H^{SM}$ indicates that the analysis is only sensitive to a signal if it occurs at three times the expected SM rate. An analysis is deemed *sensitive* to the signal if the exclusion limit is below one.

4.9.2. Statistical Significance

Now, to test whether any excess of observed data came from background fluctuation and *not* from any signal, a new test is required. A test statistic q_0 is defined as

$$q_0 = \begin{cases} -2 \ln \lambda(\mu) & \text{if } \hat{\mu} \geq 0 \\ 0 & \text{if } \hat{\mu} < 0 \end{cases} \quad (4.30)$$

Now, any excess in the data ($\hat{\mu} \geq 0$) is not consistent with a background-only model. The probability that the background has fluctuated up in such a way to produce the observed data is

$$p_0 = \int_{q_{\hat{\mu}}}^{\infty} f(q_0|0) dq_0 \quad (4.31)$$

⁶Often dubbed ‘‘Brazilian Plots’’ due to the green and yellow bands for the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties, as seen in Section 4.10.2.

Bin Merging

1. Begin with histogram with 10 GeV bins from 0–400 GeV
2. For each background, look for worst bin, starting with score $s(b) = 0$:
 $n(b)$ = total bin content for bin b
 $\varepsilon(b)$ = total error on bin b
 - a) If $n(b) \leq 0$: $s(b) += 1 - n(b)$
 - b) If $\varepsilon(b)/n(b) > 1$: $s(b) += \varepsilon(b)/n(b)$
3. Take bin with the highest score b_{\max} , and merge:
 - a) If $n(b_{\max} + 1) > n(b_{\max} - 1)$: Merge bins b_{\max} and $b_{\max} + 1$
 - b) If $n(b_{\max} + 1) < n(b_{\max} - 1)$: Merge bins b_{\max} and $b_{\max} - 1$
4. Continue until all bins have $s(b) = 0$

Algorithm 4.1: Bin merging algorithm.

For ease, this value is often converted to the number of standard deviations (σ) that $\hat{\mu}$ is from $\mu = 0$. In particle physics, if the analysis is sensitive to potential signal, it is customary to claim an excess of 3σ as an *evidence*, with an excess of at least 5σ as *observation*.

4.10. Results

Exclusion limits are obtained by applying the method discussed in the previous section to the m_{MMC} distribution for each analysis category, and then combining the results. To ensure a converging fit, it is required that all bins of the distribution to be fit have non-zero values for each background component. For this reason, the $t\bar{t}$, $Z \rightarrow \ell\ell$, di-boson, and $W(\rightarrow \ell\nu) + \text{jets}$ (for the non-VBF categories) are combined into an *others* category, as there are insufficient statistics to populate the distribution otherwise. Also, a rebinning procedure is applied by finding the *worst* bin in the distribution, and then merging it with the bin above or below depending on which has a higher content. The bin ranking is done per background, with bins with negative⁷ or zero content being deemed worst, followed by those where the total uncertainty on the bin is greater than 100% of its content. Each bin b is given a score $s(b)$, based on its content, and the bin with the highest score is then merged with its neighbour. The procedure is outlined in Algorithm 4.1.

⁷a negative yield can arise if the MC event weight is less than zero, due to the specifics of the background estimation method, or the MC generator setting an event weight to -1 when calculating NLO decays [79]

Scale systematics are included as nuisance parameters in the fit by allowing the nominal shape to vary up and down by its $\pm 1\sigma$ uncertainty (scale systematics include K_X , R_{QCD} , FF , $E/M/T_EFF$, BR_tautau , pdf_qq/gg , and $QCDScale_qqH/ggH/WH/ZH$), while shape systematics are treated with separate $\pm 1\sigma$ histograms (shape systematics include JES_TES , $E/M/MET_SCALE$, JER , and E/MET_RES). The fitting procedure is carried out using the ROOT packages ROOSTATS [80] for implementing the likelihood function, and ROOFIT [81] for the minimisation. The m_{MMC} pre-fit distributions with their final binning are shown in Figure 4.12.

4.10.1. Fit Validation

After performing the fit, the reliability of the background model and the treatment of the systematics can be studied using *pull* distributions. A pull-distribution tests the post-fit values ($\hat{\theta}$) of the nuisance parameters, compared to their nominal pre-fit values (θ_0), and the expected uncertainty $\Delta\theta$;

$$\text{pull} = \frac{\hat{\theta} - \theta_0}{\delta\theta} \quad (4.32)$$

The systematics are shown to be well accounted for if the pull is consistent with 0. Any deviation suggests a possible mis-modelling of the background, or an over/underestimation of the size of the systematic. Furthermore, the *post-fit impact* ($\Delta\hat{\mu}$) of each nuisance parameter, calculated by setting each nuisance parameter value to its $\pm 1\sigma$ variation, and refitting, can be investigated. The post-fit impact gives a measure of the influence of each nuisance parameter on the fitted signal strength.

Figure 4.13 shows the pull distributions for the VBF and Boosted categories. The nuisance parameters can be seen to be well modelled, with the fitted values being consistent with the nominal values for all systematics to $\pm 1\sigma$ uncertainties. Both categories also show that the JER and JES_TES nuisance parameters have the largest effect on the fitted signal strength.

The pull distributions for the 1- and 0-Jet categories are shown in 4.14. A much larger discrepancy between the nominal and fitted nuisance parameters can be seen. The large upward shifts in the identification efficiency nuisance parameters T_EFF and E_EFF , are likely due to a potential mis-modelling of the large fake- τ content of the 1-Jet category, and the $Z \rightarrow \tau\tau$ contribution for the 0-Jet category. A mis-modelling of the jets in the 0-Jet category may

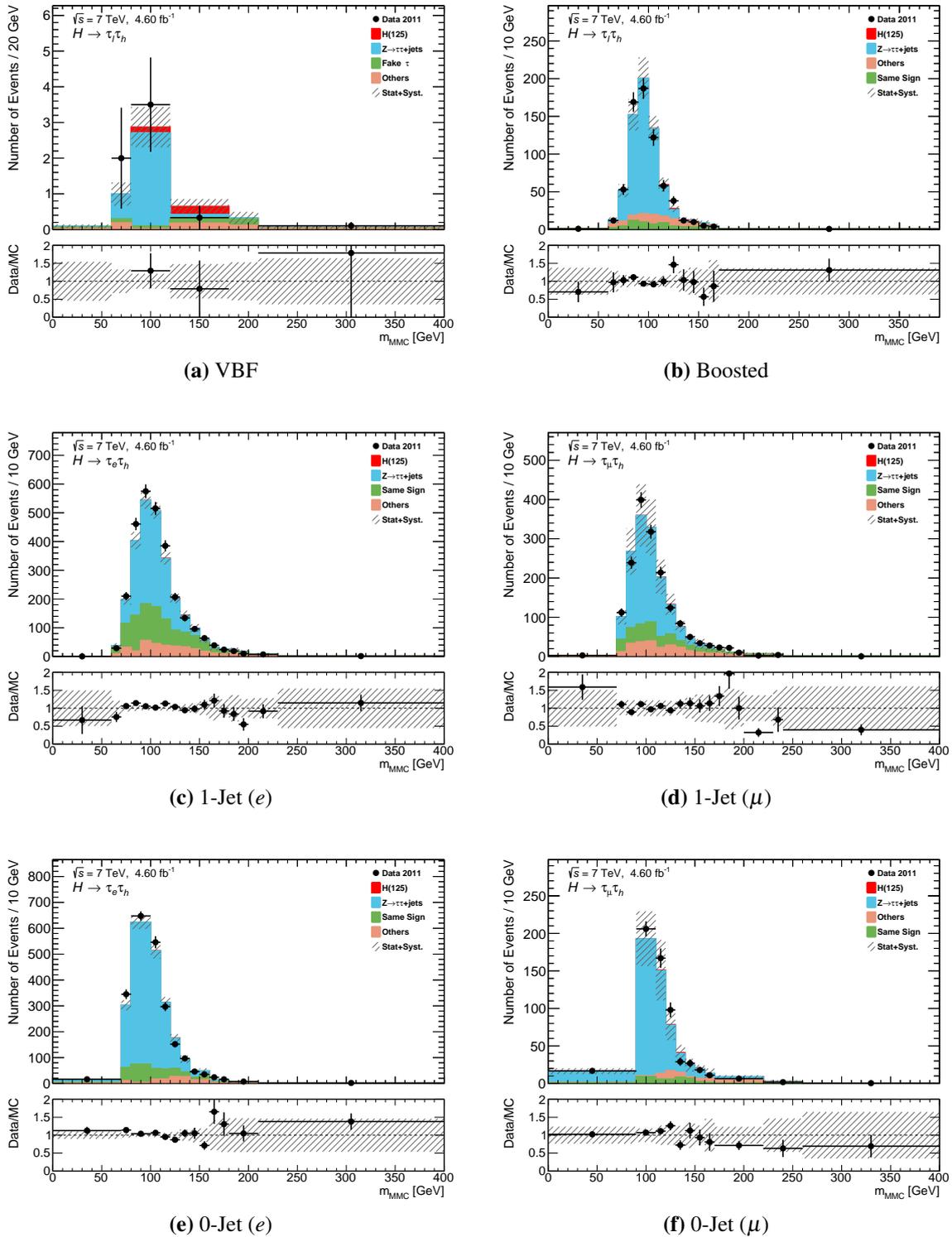


Figure 4.12.: m_{MMC} pre-fit distributions in the signal region for each analysis category. $Z \rightarrow \ell\ell, t\bar{t}$, di-boson, and $W(\rightarrow \ell\nu) + \text{jets}$ backgrounds are grouped into the *others* category. Errors include statistical errors and all systematic uncertainties. Bins are merged according to Algorithm 4.1, with bin contents scaled to number of events for every 10 GeV.

similarly be responsible for the large pulls on the jet and MET energy scale and resolution nuisance parameters JES_TES and MET_RES.

It should be emphasised here that, although there is room for improving the background model and systematic treatment for the 1- and 0-Jet categories, similar pull distributions were found for the analysis techniques discussed in the following chapters. Although the inclusion of the 1- and 0-Jet categories in the final fits may lessen the robustness of the observed exclusion limits and signal strengths, the primary focus of this thesis is a comparison of analysis techniques. Removing these categories lessens the information available to study the differences in these techniques, and as such, all analysis categories will be included in the final fits. The expected exclusion limits and signal strengths using only the VBF and Boosted categories for the various analysis techniques show the same trends as the full combination, and Figures are included in Appendix A.

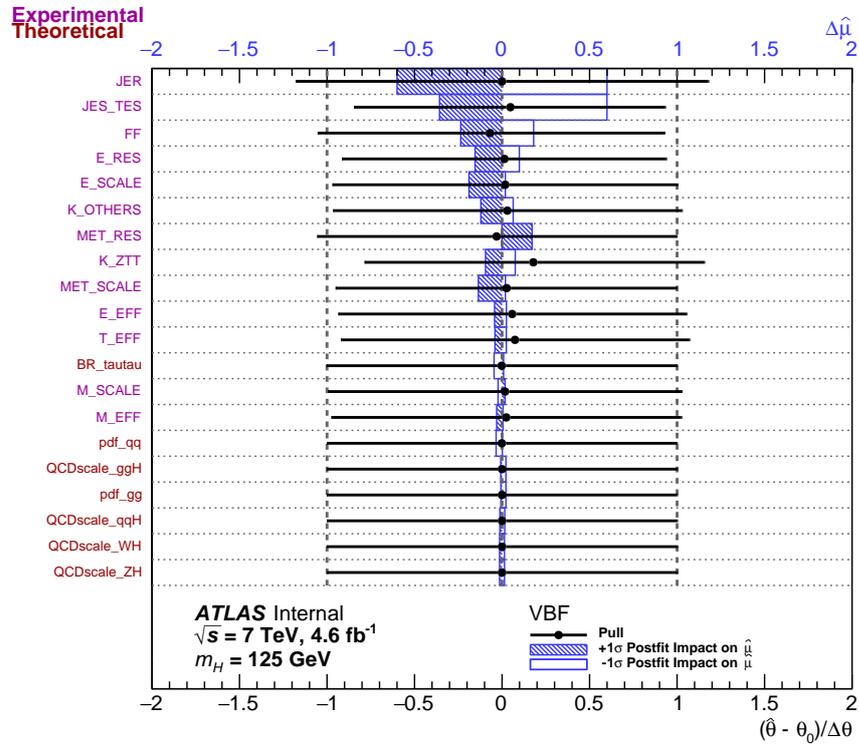
Finally, Figure 4.15 shows the pull distribution for the combined fit, as well as the purely VBF and Boosted combination, referred to as the *non-jet combined* category. The pulls for the combined fit indicate a large upward shift for T_EFF and a large downward shift for E_RES, driven primarily by the 1- and 0-Jet categories. This is supported by the non-jet combination fit, which shows a much more consistent agreement between the pre- and post-fit nuisance parameter values.

4.10.2. Exclusion Limits and Signal Strength

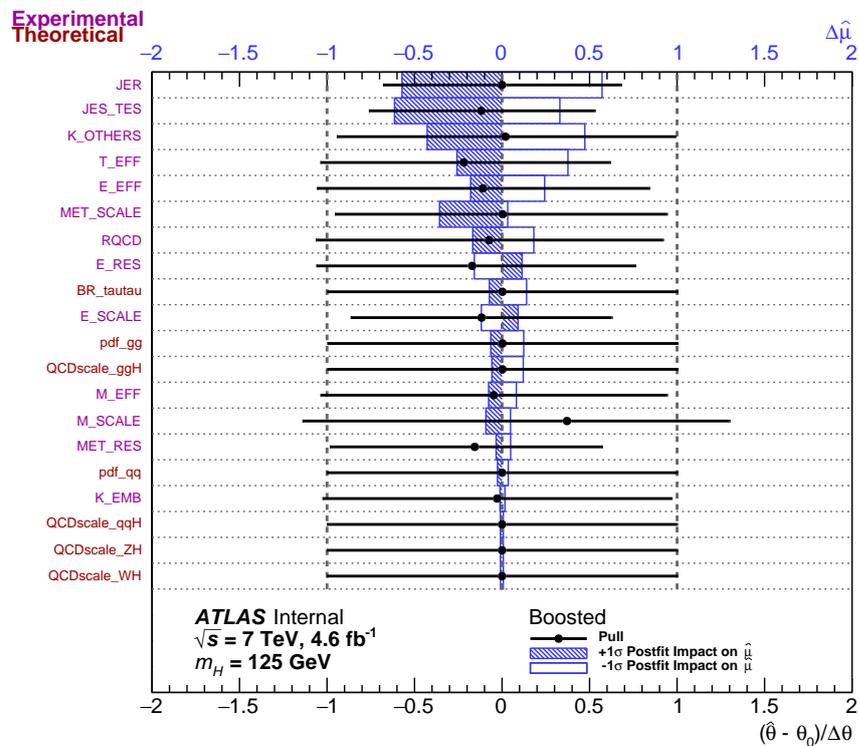
The exclusion limit for each category is calculated separately, with the 95% confidence level limits on σ/σ_H^{SM} as a function of assumed m_H shown in Figure 4.16. For each category, the highest expected limit is placed at $m_H = 125$ GeV, as this was the mass point at which the analysis was optimised. The Boosted category provides the most stringent expected limit at $5.3_{-3.9}^{+7.9} \times \sigma_H^{SM}$ at $m_H = 125$ GeV, followed by the VBF category with $5.6_{-4.1}^{+8.9} \times \sigma_H^{SM}$. It can be seen that even though the signal purity of the VBF category is highest, the high uncertainties placed on the background estimation compared to the Boosted category lower its significance.

The observed limits agree with the expected limits to within 2σ for all categories. The departure of the observed from expected limits in the 0- and 1-Jet categories can be attributed to the background model slightly over- and underestimating the data in each category respectively, as can be seen in Figure 4.12.

A combined exclusion limit as obtained as shown in Figure 4.17, with an expected upper limit on the Higgs production cross section placed at $2.9_{-0.8}^{+1.4} \times \sigma_H^{SM}$ for $m_H = 125$ GeV,

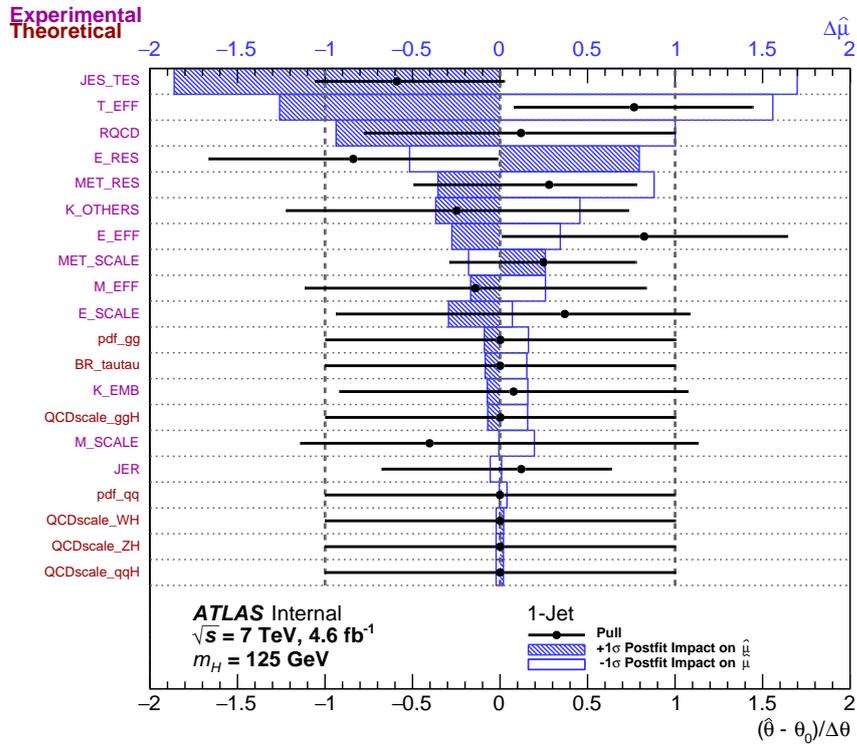


(a) VBF

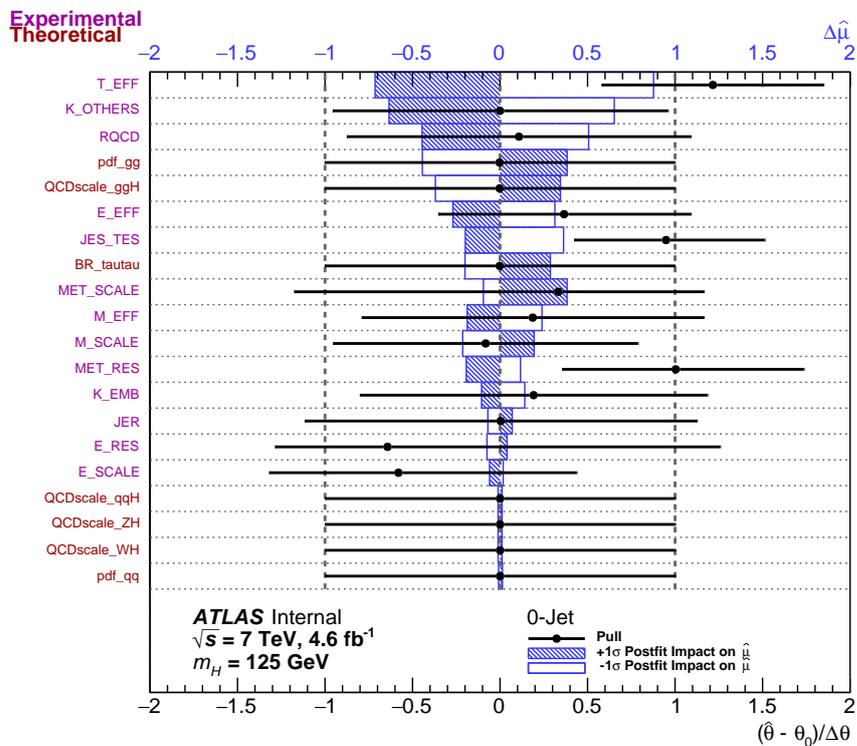


(b) Boosted

Figure 4.13.: Pull distributions for the (a) VBF and (b) Boosted categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

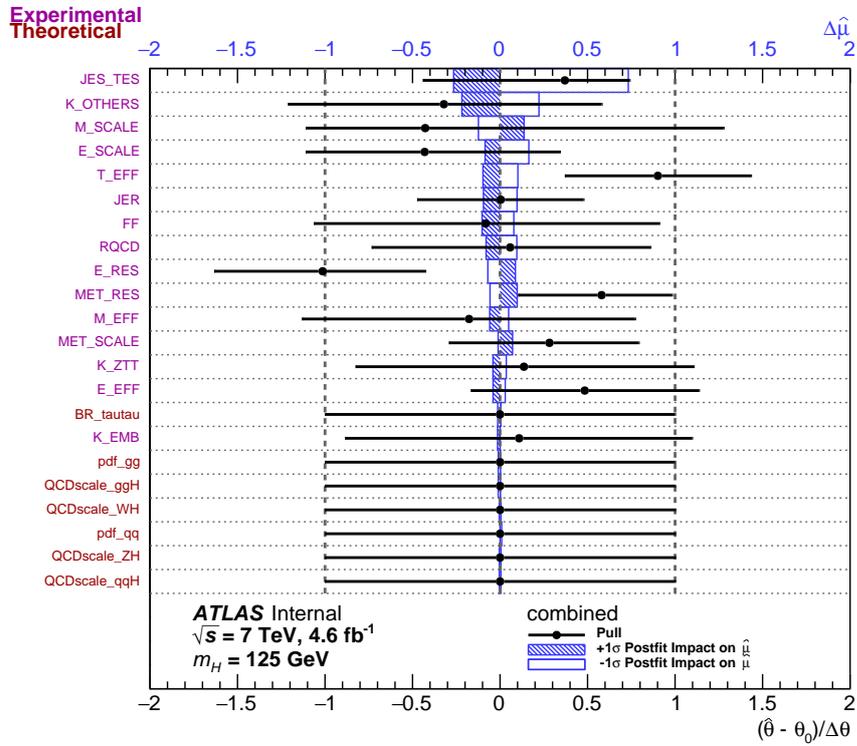


(a) 1-Jet

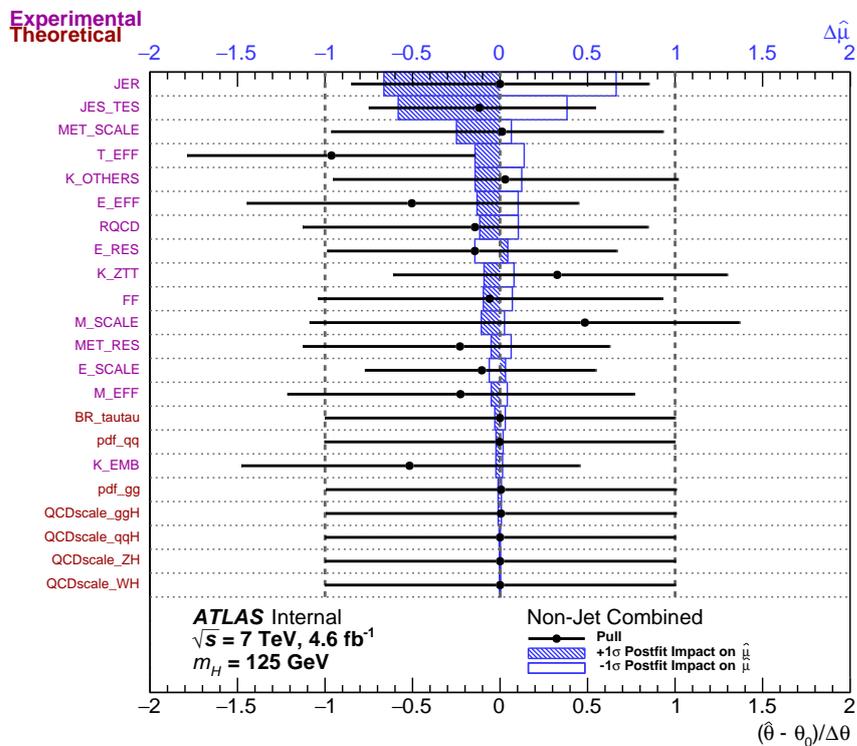


(b) 0-Jet

Figure 4.14.: Pull distributions for the (a) 1-Jet and (b) 0-Jet categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.



(a) Combined



(b) Non-Jet Combined

Figure 4.15.: Pull distributions for the (a) combined and (b) non-jet combined categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

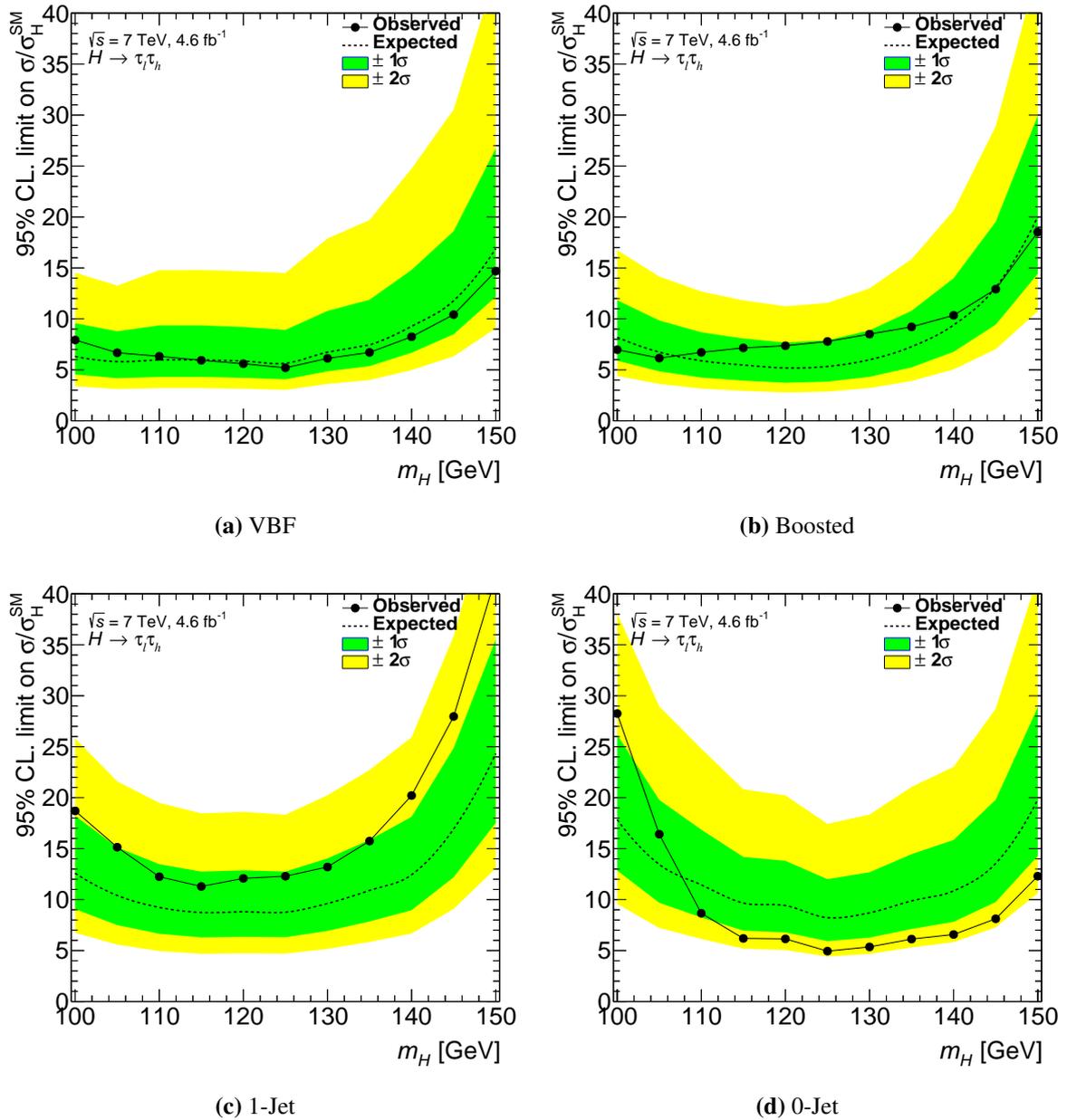


Figure 4.16.: Cut-based exclusion limits for VBF, Boosted, 1-Jet, and 0-Jet analysis categories. Uncertainties are statistical and systematic.

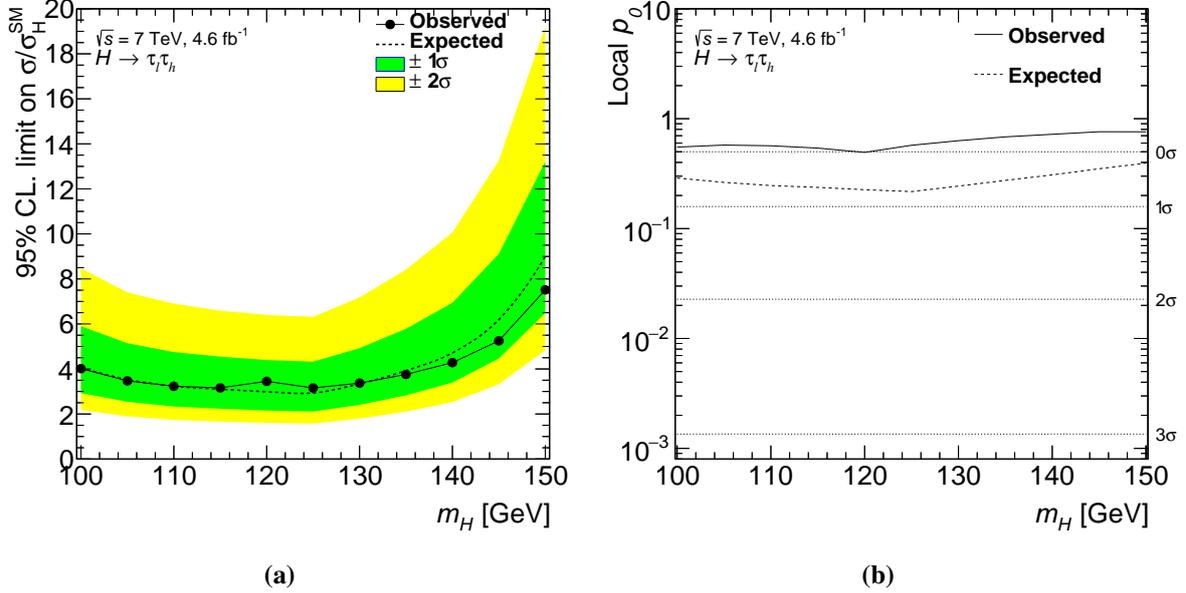
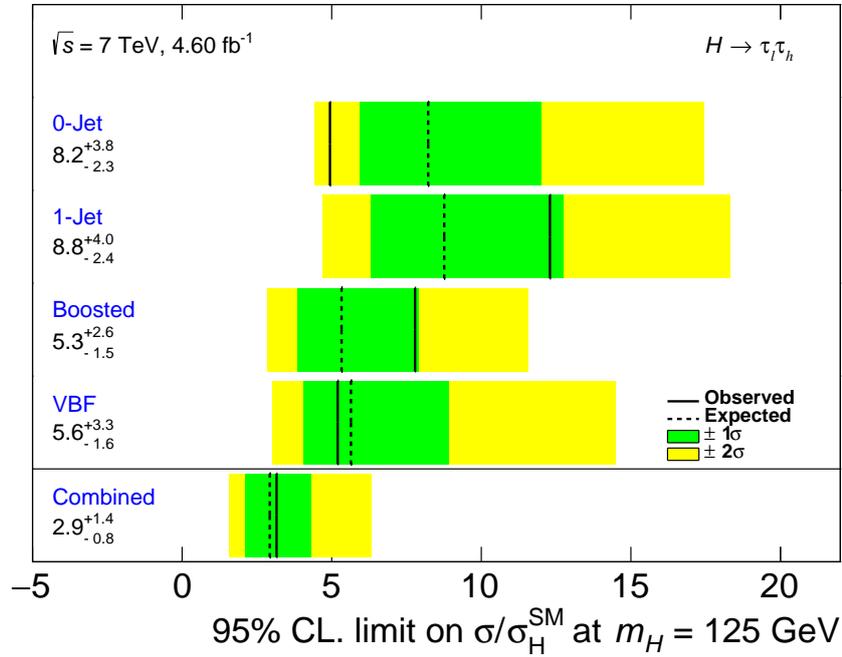


Figure 4.17.: Cut-based (a) exclusion limits and (b) signal significance for VBF, Boosted, 1-Jet, and 0-Jet analysis categories combined. Uncertainties are statistical and systematic.

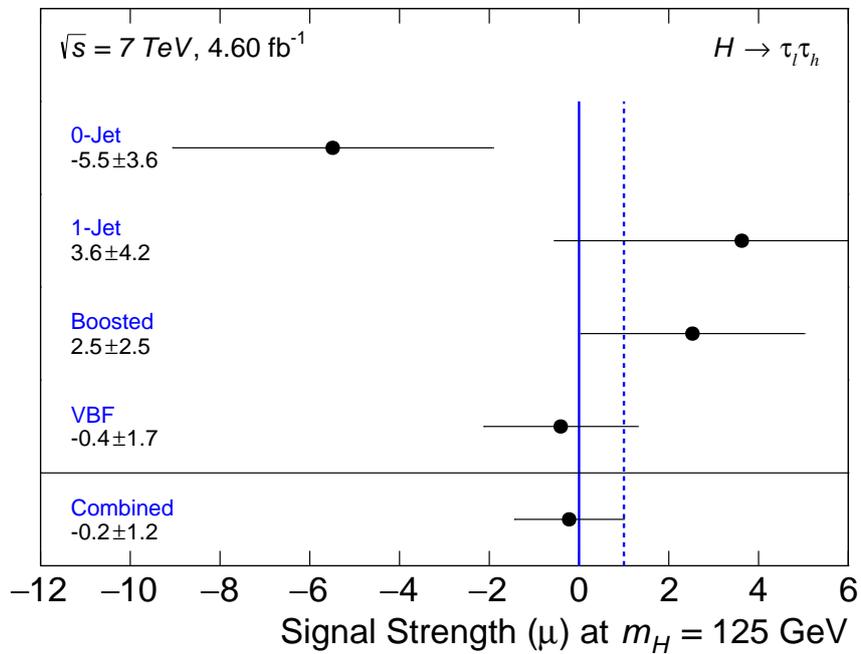
and an observed limit of $3.2 \times \sigma_H^{SM}$. The observed limit is within the $\pm 1\sigma$ uncertainties of the expected limit across the entire mass range. Also shown in Figure 4.17 is the signal significance as a function of m_H , assuming the background-only hypothesis. At $m_H = 125$ GeV, an expected significance of $p_0 = 0.22$ (0.78σ) is set. The observed significance however is below zero, which while consistent with the background-only hypothesis, it does indicate that the background model was overestimated. No significant excess of data is observed.

A comparison of the exclusion limits for each individual category at $m_H = 125$ GeV is shown in Figure 4.18, as well as the fitted value of the signal strength ($\hat{\mu}$) for each category, and the combined case. The combined result $\hat{\mu} = -0.2 \pm 1.2$ is consistent both with the existence of a 125 GeV mass Higgs boson ($\mu = 1$), and also the background-only hypothesis ($\mu = 0$) at the 1σ level. Individually, all categories are consistent with either hypothesis within 2σ .

In the next chapter, a multi-variate approach is taken to try and improve the cut-based analysis, and in particular the expected limits from the VBF category. This cut-based analysis, and specifically the expected exclusion limit at $m_H = 125$ GeV will serve as the benchmark upon which the so-called *hybrid* analysis will be compared.



(a)



(b)

Figure 4.18.: (a) Comparison of cut-based exclusion limits for each individual analysis category, and the combined limit, for $m_H = 125 \text{ GeV}$. Uncertainties are statistical and systematic. (b) Extracted signal strength μ at $m_H = 125 \text{ GeV}$ for each analysis category, and the combined signal strength. Solid line is $\mu = 0$ and dashed line is $\mu = 1$

Chapter 5.

Multi-Variate Analysis Techniques and an $H \rightarrow \tau_\ell \tau_h$ Hybrid Analysis

This chapter will give an overview of multi-variate analysis techniques, in particular the boosted decision tree. The application of these techniques to the $H \rightarrow \tau_\ell \tau_h$ analysis of the previous chapter will then be explored, with a comparison of the results obtained presented.

5.1. Multi-Variate Analyses

Multi-variate analyses (MVA) are a machine learning approach to data analysis. Rather than analysing data by hand to determine the best methods of signal extraction, an algorithm is trained to *learn* to make predictions about the data, given various inputs. Many MVA methods however can be subject to a problem known as *overtraining*, where specific features of the training sample are selected, rather than the general features of the data. Care must therefore be taken when constructing the MVA classifiers. The MVA technique discussed here is that of the Boosted Decision Tree (BDT), with the theory explained in more detail in [82].

5.1.1. Decision Trees

The standard classification problem in particle physics is the separation of signal and background events, given some detector response in the form of variables. A *cut-based* analysis will take these variables and apply certain requirements or *cuts* on them to preferentially select signal-type events and discard background-type events. The problem here is that the events

that fail the classification may very well be signal events. The simple cut-based approach is therefore not robust against event mis-classifications.

The decision tree approach is to take the otherwise discarded events and apply further selections, performing a series of sequential binary splits in the feature space to recover mis-classified events. A set of N events $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$ with variables $\mathbf{x}_i = \{x_1, x_2, \dots, x_n\}_i$ and known classifications

$$y_i = \begin{cases} +1, & \text{if event is signal} \\ -1, & \text{if event is background} \end{cases} \quad (5.1)$$

referred to as the *training* sample is used to *grow* the tree, while a separate set of events, the *testing* sample, is used to analyse the results.

A tree is grown by placing a cut on one of the variables, splitting the data into two branches. This point becomes a *node*. This splitting continues down each branch until some stopping condition is met, and the branch ends in a *leaf*. Figure 5.1 shows an example of a simple decision tree. Notice that signal events that fail the $x_1 > a$ cut at the first node may be recovered at the right node by the $x_3 < c$ cut.

The optimal splitting variable and value can be determined in many ways, with a common choice being the *gini index*. Given a region of feature space R_j , where j runs over the total number of regions of interest, the number of signal and background events are N_j^S and N_j^B . The gini index g_j is then given by

$$g_j = P_j(1 - P_j) \quad (5.2)$$

where P_j is the purity

$$P_j = \frac{N_j^S}{N_j^S + N_j^B} \quad (5.3)$$

To decide which variable to cut on and where, two new regions $R_j^{x < c}$ and $R_j^{x > c}$ are created by placing a cut c on variable x_i . The choice of (x_i, c) that maximises the separation of the two new regions, i.e. the quantity

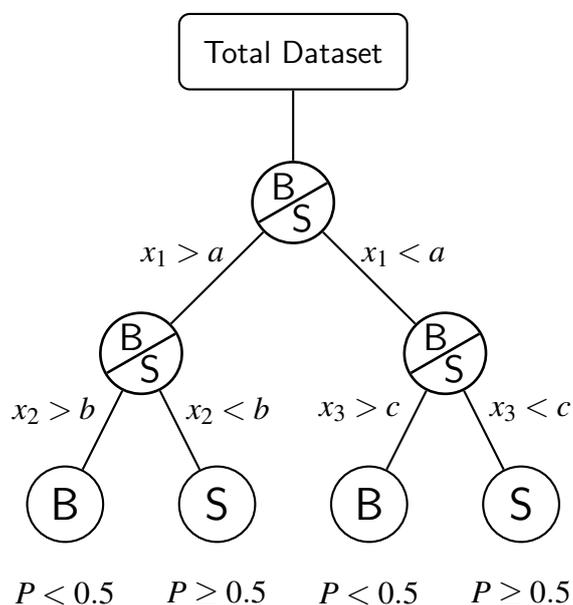


Figure 5.1.: Example of a simple decision tree, separating signal S from background B.

$$\Delta g_j = g_j - \frac{N_j^{x_i < c} g_j^{x_i < c} + N_j^{x_i \geq c} g_j^{x_i \geq c}}{N_j^{x_i < c} + N_j^{x_i \geq c}} \quad (5.4)$$

defines the node.

The stopping conditions provide the simplest way of controlling the power of a decision tree. One condition is the *minimum leaf size*, whereby a node becomes a leaf when it contains too few events, allowing one to protect against statistical fluctuations. Another stopping condition is the *maximum depth* - the maximum number of splits that can occur. Once a leaf is formed, it may be classified as signal according to its purity. A leaf is classified as signal if $P \geq 0.5$.

Decision trees have the potential to be powerful, but also unstable. A small change in training data can produce a large change in the final tree. Also, decision trees grown from small training samples, or those with highly correlated features, are likely to reach their stopping conditions early. Such trees may end up as simple rectangular cuts on the feature space, with poor discriminating power. By using many of these simple trees simultaneously however, a more detailed map of the space may be constructed. This remedy is known as *boosting*.

5.1.2. Boosting

The aim of boosting is to add many *weak* classifiers (trees) together to make a single *strong* classifier. The desired outcome is a function $F(\mathbf{x})$ that maps a set of variables \mathbf{x} to the known responses y , that is, we require a model that can predict whether an event is signal or background, given its properties. Consider now a weighted sum of M weak classifiers $h(\mathbf{x}; \mathbf{a}_m)$;

$$F(\mathbf{x}) = \sum_{m=1}^M \alpha_m h(\mathbf{x}; \mathbf{a}_m) \quad (5.5)$$

where each is built on some training sample of N events $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$, and has parameters $\mathbf{a}_m = \{a_1, \dots, a_k\}$. A straight sum is not ideal, as not every classifier will perform equally and a way to preferentially select the more accurate classifiers is needed. The contribution of each classifier in the sum is therefore weighted by some coefficient α_m , where

$$\sum_{m=1}^M \alpha_m = 1 \quad (5.6)$$

Calculating and optimising the classifiers and their coefficients may not be a trivial task. Rather than trying to optimise all the terms of the sum at once via some complicated fit, each term can be optimised individually, after taking into account all the preceding terms. The model at each boost step $F_m(\mathbf{x})$ is therefore given by

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \alpha_m h(\mathbf{x}; \mathbf{a}_m) \quad (5.7)$$

and each coefficient α_m and classifier parameters \mathbf{a}_m are optimised in such a way that the performance of the model improves. This process is known as *forward stagewise additive modelling*, and can be compared to iteratively adding corrections to a function expansion. The next step is to decide upon how the optimisation is done. At its simplest, optimisation involves the minimisation of some metric, which in our case we will call the *loss function*.

The loss function $L(y, F(\mathbf{x}))$ is defined to measure the deviation of the model response $F(\mathbf{x})$ from the true response y . Each boost step is therefore optimised by choosing the coefficient and classifier parameters such that the deviation from the true response is minimised, i.e.;

$$(\alpha_m, \mathbf{a}_m) = \arg \min_{\alpha, \mathbf{a}} \sum_{i=1}^N L(y_i, F_{m-1}(\mathbf{x}_i) + \alpha h(\mathbf{x}_i; \mathbf{a})) \quad (5.8)$$

A simple example of a loss function is that of exponential loss;

$$L(y, F(\mathbf{x})) = e^{-F(\mathbf{x})y} \quad (5.9)$$

where for each event, both the model and known responses are either signal ($S = 1$) or background ($B = -1$). A more forgiving model of loss is binomial log-likelihood loss,

$$L(y, F(\mathbf{x})) = \ln \left(1 + e^{-2F(\mathbf{x})y} \right) \quad (5.10)$$

The minimisation in equation 5.8 can be done explicitly in the case of exponential loss, while the binomial log-likelihood loss function requires an analytical approach, such as the *gradient descent method*, to minimise. The specifics of the minimisation depend on the boosting algorithm. Two algorithms studied in this thesis are the Adaptive Boost (AdaBoost) [83] and Gradient Boost [84] methods, as implemented by the ROOT Toolkit for Multivariate Data Analysis package, TMVA [85].

AdaBoost

The AdaBoost approach to boosting decision trees involves a reweighting of the data at each boost step before training a new tree. The weights are determined based on the classification error of the previous tree. Events that are classified correctly are weighted up, and those mis-classified are weighted down.

To build the forest of classifiers as in equation 5.5, we start with data as defined above and the exponential loss function. Each iteration m of the algorithm starts by setting the event weights according to the classification error of the forest so far;

$$w_i^{(m)} = e^{-y_i F_{m-1}(\mathbf{x}_i)} \quad (5.11)$$

To optimise the next decision tree $h(\mathbf{x}_i; \mathbf{a})$ and the coefficients α_m , we must now minimise the loss function

$$L = \sum_{i=1}^N e^{-y_i F_m(\mathbf{x}_i)} \quad (5.12)$$

$$= \sum_{i=1}^N w_i^{(m)} e^{-y_i \alpha_m h_m(\mathbf{x}_i; \mathbf{a})} \quad (5.13)$$

Now, splitting this sum into terms where the event is correctly classified ($y_i = h_m(\mathbf{x}_i; \mathbf{a})$) or incorrectly classified ($y_i \neq h_m(\mathbf{x}_i; \mathbf{a})$) gives

$$L = \sum_{y_i = h_m(\mathbf{x}_i; \mathbf{a})} w_i^{(m)} e^{-\alpha_m} + \sum_{y_i \neq h_m(\mathbf{x}_i; \mathbf{a})} w_i^{(m)} e^{\alpha_m} \quad (5.14)$$

It can then be shown the tree that minimises L must also minimise the classification error incorporated into each $w_i^{(m)}$. It can also be shown that, given the optimised tree, the coefficients α_m can then be found by minimising L by setting $\frac{\partial L}{\partial \alpha_m} = 0$ and solving for α_m , giving

$$\alpha_m = \frac{1}{2} \log \left(\frac{\sum_{y_i = h_m(\mathbf{x}_i; \mathbf{a})} w_i^{(m)}}{\sum_{y_i \neq h_m(\mathbf{x}_i; \mathbf{a})} w_i^{(m)}} \right) \quad (5.15)$$

Defining the weighted error rate to be

$$\epsilon_m = \frac{\sum_{y_i \neq h_m(\mathbf{x}_i; \mathbf{a})} w_i^{(m)}}{\sum_{i=1}^N w_i^{(m)}} \quad (5.16)$$

results in the boost weight for the AdaBoost algorithm

$$\alpha_m = \frac{1}{2} \log \left(\frac{1 - \epsilon_m}{\epsilon_m} \right) \quad (5.17)$$

The final result is a score $S(\mathbf{x}_i) \in [-1, 1]$ for each event, which is the weighted sum of classifier outcomes for each boost step. The AdaBoost algorithm is summarised in algorithm 5.1. Rather than relying on the binary classifier outcome $h(\mathbf{x}_i; \mathbf{a})$ to build the score, the *training purity* – the amount of signal or background events in each leaf compared to the total – can be used as the signal or background weights. When implemented here it results in the Real-AdaBoost algorithm, however it can be used for any boosting procedure.

The use of the exponential loss function results in poorly performing trees are weighted down quite harshly, meaning that AdaBoost can converge to a strong classifier quite quickly. This trait however also means that AdaBoost is also subject to overtraining, where specific properties of the training sample are emphasised in the final discriminator, rather than the global trends. The impact of statistical fluctuations can be lessened by changing the *learning rate* β , which can slow the boosting by multiplying α_m by some value $0 < \beta \leq 1$. This forces the algorithm to make smaller adjustments to the event weights at each boost step. Other boosting methods that are more robust against overtraining may also be used. One such example is gradient boosting.

Gradient Boost

Gradient boosting can accommodate a variety of loss functions, and as such, can be made less prone to overtraining. One caveat to this is that the number of boosts required to achieve a similar performance to AdaBoost can be greater. However, when dealing with small numbers of training events as is the case in rare physics searches, it will be seen in Section 6.3.2 that the loss in speed is well worth the increase in robustness.

Unlike AdaBoost which relies on the mis-classification rate, the Gradient boosting algorithm relies on calculating the gradient of the loss function at each boost step, and growing regression trees¹ to predict the calculated gradient at each leaf. The tree weights α_m are then those that minimise the loss after adding the new tree to the forest.

Once again, given data $\{\mathbf{x}_i, y_i\}_{i=1}^N$ and assuming binomial log-likelihood loss (Equation 5.10), the first step of the boosting procedure is to calculate the gradient of the loss function, referred to here as the *target*² t_{im}

¹Unlike a decision tree which acts as a simple classifier, the leaves of a regression tree can be continuous variables. The splitting procedure also differs compared to decision trees, with the splitting variables and values being those that minimise the average square error after the split.

²called the “pseudo” residual in the jargon.

AdaBoost (TMVA Implementation)

1. Begin with data $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$ where $y_i \in \{-1, 1\}$ and weights $w_i^{(0)} = 1$
2. Train classifier $h_0(\mathbf{x}_i; \mathbf{a}) \in \{-1, 1\}$
3. For each boost step $m \in \{1, 2, \dots, M\}$:
 - a) Train classifier $h_m(\mathbf{x}_i; \mathbf{a})$ with modified weights

$$w_i^{(m)} = e^{-y_i F_{m-1}(\mathbf{x}_i)}$$

- b) Calculate classification error rate ϵ_m and derive the boost rate

$$\alpha_m = \frac{1}{2} \log \left(\frac{1 - \epsilon_m}{\epsilon_m} \right)$$

- c) Add the new classifier to the forest

$$F_m(\mathbf{x}_i) = F_{m-1}(\mathbf{x}_i) + \beta \alpha_m h_m(\mathbf{x}_i, \mathbf{a})$$

where β is the learning rate

4. The final classifier score $S(\mathbf{x}_i)$ is then

$$S(\mathbf{x}_i) = \beta \sum_{m=1}^M \alpha_m h_m(\mathbf{x}_i; \mathbf{a}_m)$$

Algorithm 5.1: AdaBoost TMVA implementation [85]

$$t_{im} = - \left[\frac{\partial L(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)} \right]_{F_m(\mathbf{x}_i)=F_{m-1}(\mathbf{x}_i)} \quad (5.18)$$

A weak learner $h(\mathbf{x}; \mathbf{a}_m)$ is then fitted to the targets (can be thought of as a classifier being trained on the set $\{\mathbf{x}_i, t_{im}\}_{i=1}^N$, effectively replacing the true response y_i with the targets). The optimal value of α_m is then that which minimises the loss $L(y, F_{m-1}(\mathbf{x}_i) + \alpha_m h(\mathbf{x}_i; \mathbf{a}_m))$. This is a difficult optimisation problem when using binomial log-likelihood loss, however it can be simplified by the use of regression trees.

The result of a regression tree is a set of J regions $\{R_{jm}\}_{j=1}^J$ (like leaves of a decision tree), that return the average value of the targets within. By treating each region separately, the minimisation can be reduced to the problem of calculating the γ_{jm} that minimises the loss $L(y, F_{m-1}(\mathbf{x}_i) + \gamma_{jm})$, for $\mathbf{x}_i \in R_{jm}$. The model is then updated at each step, in each region separately, and the forest becomes

$$F_m(\mathbf{x}_i) = F_{m-1}(\mathbf{x}_i) + \beta \sum_{j=1}^J \gamma_{jm} \text{ for } \mathbf{x}_i \in R_{jm} \quad (5.19)$$

As the model response converges, the gradient of the loss function - the targets - also converge and become stable. Any new regression trees grown will not add any new discrimination power, although unlike AdaBoost, they will not likely lead to overtraining. The final outcome of the algorithm is a score $S(\mathbf{x}_i) \in [-1, 1]$, and its implementation as found in TMVA is described in Algorithm 5.2.

5.2. Hybrid Multi-Variate/Cut-Based Analysis

The multi-variate techniques outlined above may be used to increase the sensitivity of the search for the SM Higgs. The simplest extension of the analysis in the previous chapter is to simply take the variables used in the VBF categorisation, and add them to a BDT. This naive approach will be referred to as the *hybrid* method, and is used as a stepping stone to performing a full, dedicated multi-variate analysis (detailed in Chapter 6).

This section outlines the analysis structure of the hybrid method, and the process in which a BDT is constructed for selecting events in the hybrid VBF category, VBF_{HBD} . The choice of

 Gradient Boost (TMVA Implementation)

1. Begin with data $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$ where $y_i \in \{-1, 1\}$
2. Calculate initial target $t_{i0} = y_i/2$
3. Grow regression tree $h_0(\mathbf{x}; \mathbf{a}_m)$ to predict initial target
4. For each boost step $m \in \{1, 2, \dots, M\}$:
 - a) Calculate targets

$$t_{im} = \frac{2y_i}{e^{2y_i F_{m-1}(\mathbf{x}_i)} + 1}$$

- b) Fit $h_m(\mathbf{x}; \mathbf{a})$ to target making a set of regions $\{R_{jm}\}_{j=1}^J$
- c) Choose γ that minimises $L(y_i, F_{m-1}(\mathbf{x}_i) + \gamma)$ in every region R_{jm}

$$\gamma_{jm} = \frac{\sum_{i \in R_{jm}} w_i t_{im}}{\sum_{i \in R_{jm}} w_i^2 |t_{im}| (1 - |t_{im}|)}$$

- d) Update the model by adding γ_{jm} to the forest for every region R_{jm}

$$F_m(\mathbf{x}_i) = F_{m-1}(\mathbf{x}_i) + \beta \sum_{j=1}^J \gamma_{jm}$$

where β is the learning rate (shrinkage).

5. The final forest is then

$$F(\mathbf{x}_i) = \beta \sum_{m=1}^M \sum_{j=1}^J \gamma_{jm}$$

and the classifier score is

$$S(\mathbf{x}_i) = \frac{2}{1 + e^{-2F(\mathbf{x}_i)}} - 1$$

Algorithm 5.2: Gradient Boost TMVA implementation [85]

	VBF _{HBD}	Boosted	1-Jet	0-Jet
Category	$N_{jet} \geq 2$	not VBF	not Boosted	not 1-Jet
	$m_{j_1 j_2} > 200 \text{ GeV}$	$p_T^H > 100 \text{ GeV}$	$N_{jet} = 1$	$N_{jet} = 0$
	$\Delta\eta_{j_1 j_2} > 2$	$0 < x_l < 1$	$p_T^{j_1} > 25 \text{ GeV}$	$\cancel{E}_T > 20 \text{ GeV}$
	$\text{BDT}_{\text{HBD}} > x_{\text{HBD}}$	$0.2 < x_{\tau_h} < 1.2$	$\cancel{E}_T > 20 \text{ GeV}$	
Analysis	$m_T < 50 \text{ GeV}$	$m_T < 50 \text{ GeV}$	$m_T < 50 \text{ GeV}$	$m_T < 30 \text{ GeV}$
	$\Delta(\Delta R) < 1$	$\Delta(\Delta R) < 0.6$	$\Delta(\Delta R) < 0.6$	$\Delta(\Delta R) < 0.5$
	$\Sigma\Delta(\phi) < 2.9$	$\Sigma\Delta(\phi) < 1.6$	$\Sigma\Delta(\phi) < 3.5$	$\Sigma\Delta(\phi) < 3.5$
			$Z \rightarrow \ell\ell$ rejection	$Z \rightarrow \ell\ell$ rejection

Table 5.1.: Selection cuts for Hybrid categories.

boosting algorithm and training parameters used in the TMVA framework are also discussed, along with the process for choosing a cut on the discriminator BDT_{HBD} to select the VBF_{HBD} events. Finally, the exclusion limits obtained from this analysis are presented, and compared to those from the cut-based analysis.

5.2.1. Categorisation and Background Modelling

While in the cut-based analysis the categorisation cuts were optimised to provide the best signal efficiency for the VBF category, in the hybrid method, BDT_{HBD} is expected to perform the selection. For this reason the $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ category is defined loosely; requiring two jets of $p_T > 25 \text{ GeV}$, and with $m_{j_1 j_2} > 200 \text{ GeV}$ and $\Delta\eta(j_1, j_2) > 2$. The latter two cuts are placed in order to use the ALPGEN VBF filtered $Z \rightarrow \tau\tau$ sample as per the cut-based analysis, as these selections are present in the MC filtering. The Boosted, 1-jet and 0-jet categories are defined as they were in the cut-based analysis. All category and analysis level cut definitions are shown in Table 5.1.

The background estimation methods implemented here are the same as those used for the cut-based analysis, including the k -factors, fake-factor weights, r_{QCD} , and embedding correction factors (shown in Table 4.8).

5.2.2. Growing BDT_{HBD}

The task of building the decision tree and performing the boosting to produce BDT_{HBD} is handled by TMVA, as explained in Section 5.1. The user however must define the training and testing samples to be used, as well as the training variables, and training parameters.

Training and Testing Samples

To construct BDT_{HBD} , signal and background samples must first be selected. While the most obvious selection would be to use all the analysis backgrounds, it was decided to only train against the $Z \rightarrow \tau\tau$ contribution. Figure 5.2 shows the m_{MMC} distribution for $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ before, and after signal region analysis level cuts. The yields for each background are shown in Table 5.2, where it can be seen that the analysis level cuts reject $\gtrsim 70\%$ of all backgrounds except $Z \rightarrow \tau\tau$ (with only $\approx 30\%$ signal loss). As such, training against these other backgrounds would only serve to separate events in BDT space that the analysis cuts can already remove. This decision was also made for the first attempt as these background samples have negative weights in some bins of the distribution, and handling negatively weighted events for the fake-factor method in the TMVA framework is not trivial (this issue however was resolved for the full-MVA analysis, described in Section 6.2.3).

It is required that the samples be split into roughly equal-sized training and testing samples. In order to increase the statistics available to TMVA to limit overtraining, both the embedding and ALPGEN $Z \rightarrow \tau\tau$ VBF filtered samples were used; with ALPGEN for training, and embedding for testing. The signal sample consisted of $H \rightarrow \tau_\ell \tau_h$ VBF events only, for a Higgs mass of every 5 GeV in the range 100–150 GeV. The existence of a 126 GeV Higgs boson was not assumed, and so the analysis was designed to be sensitive over a large possible mass range.

Training Variables

The variables used to construct BDT_{HBD} are those used in the cut-based definition of the VBF category, as described in Section 4.6.1, and are listed in Table 6.4. In the cut-based analysis the η centrality is a Boolean value indicating whether or not the lepton or tau is between the two leading jets in pseudorapidity. As the BDT requires a continuous variable for training, the η centrality $C_{\eta_1 \eta_2}(\eta)$ is now defined as the distance between the lepton (or tau) and the middle of the two leading jets in pseudorapidity;

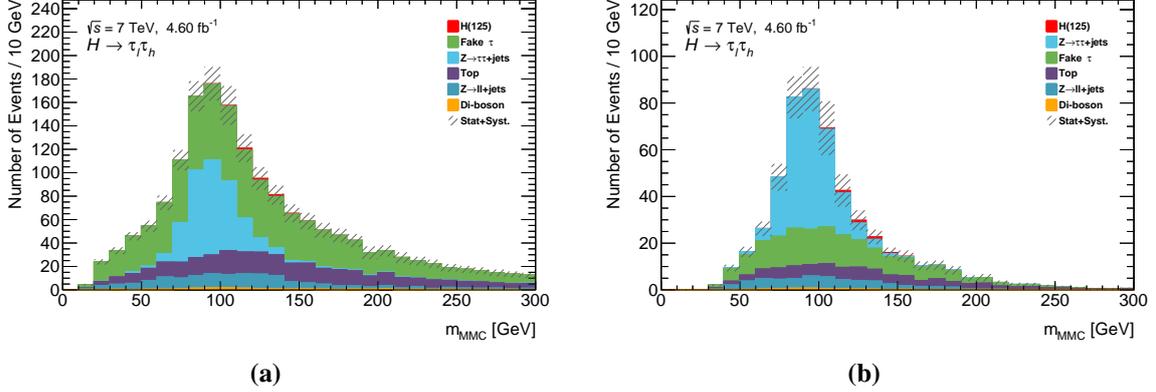


Figure 5.2.: $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ category (a) without and (b) with analysis level cuts. Uncertainties are statistical and systematic.

Process	VBF_{HBD}	$\text{VBF}_{\text{HBD}}+\text{SR}$	% Reduction
ggF 125	3.32 ± 0.12	2.16 ± 0.09	35
VBF 125	4.95 ± 0.05	3.57 ± 0.04	28
WH 125	0.12 ± 0.02	0.06 ± 0.01	50
ZH 125	0.06 ± 0.01	0.03 ± 0.01	50
$Z \rightarrow \tau\tau$	319.8 ± 3.0	223.7 ± 2.4	30
Fake- τ	893.8 ± 6.4	165.6 ± 2.9	81
$t\bar{t}$	318.4 ± 3.9	71.3 ± 1.9	78
$Z \rightarrow \ell\ell$	152.5 ± 4.4	47.4 ± 2.5	69
Di-boson	15.19 ± 0.93	4.24 ± 0.27	72
Signal	8.46 ± 0.13	5.83 ± 0.10	31
Background	1829 ± 10	552.4 ± 5.4	70

Table 5.2.: Yields for the $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ category before and after signal region (SR) analysis level cuts, as well as the percent reduction in yield. Uncertainties are statistical only.

$$C_{\eta_1 \eta_2}(\eta) = \exp \left[- \left(\frac{2}{\eta_1 - \eta_2} \right)^2 \left(\eta - \frac{\eta_1 + \eta_2}{2} \right)^2 \right] \quad (5.20)$$

where $\eta_{1(2)}$ are the pseudorapidities of the leading two jets, and η is the lepton or tau pseudorapidity. This definition leads to three cases:

$$C_{\eta_1 \eta_2}(\eta) \begin{cases} = 1 & \text{if } \ell/\tau \text{ is in the middle of the two leading jets} \\ = 1/e & \text{if } \ell/\tau \text{ is aligned with one of the two leading jets} \\ < 1/e & \text{if } \ell/\tau \text{ is outside of the two leading jets} \end{cases} \quad (5.21)$$

The BDT_{HBD} input variables are shown for the $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ selection in Figures 5.3 and 5.4, and show good agreement between Data/MC.

Training Parameters

The parameters of the BDT training algorithm were not optimised for the hybrid study; rather, the default values defined by TMVA for use with a BDT were used. It will be interesting however to compare these values with those used in the full MVA analysis of Chapter 6, which did undergo training parameter optimisation, and as such Table 6.3 shows the parameters of interest for the hybrid BDT (as well as the BDTs constructed in the full-MVA analysis).

These parameters are described in detail in Section 6.3.2, but simply, the `BoostType` refers to the particular boosting algorithm used, with the `Learning Rate` being the strength of each boost step. The `NTrees` parameter defines how many boost steps are applied in the training, with `MaxDepth` and `nCuts` being the maximum depth of decision tree, and the number of points in the available parameter space used to determine the optimal node splitting, for each trained classifier. Finally, the `MinNodeSize` defines the minimum number (as a percentage) of training events required in the final leaf.

BDT_{HBD} Validation

Figure 5.5 shows the BDT_{HBD} after training and testing. Also shown is the Receiver Operating Characteristic (ROC) curve, which gives the fraction of background rejected ($1 - \epsilon_B$) as a

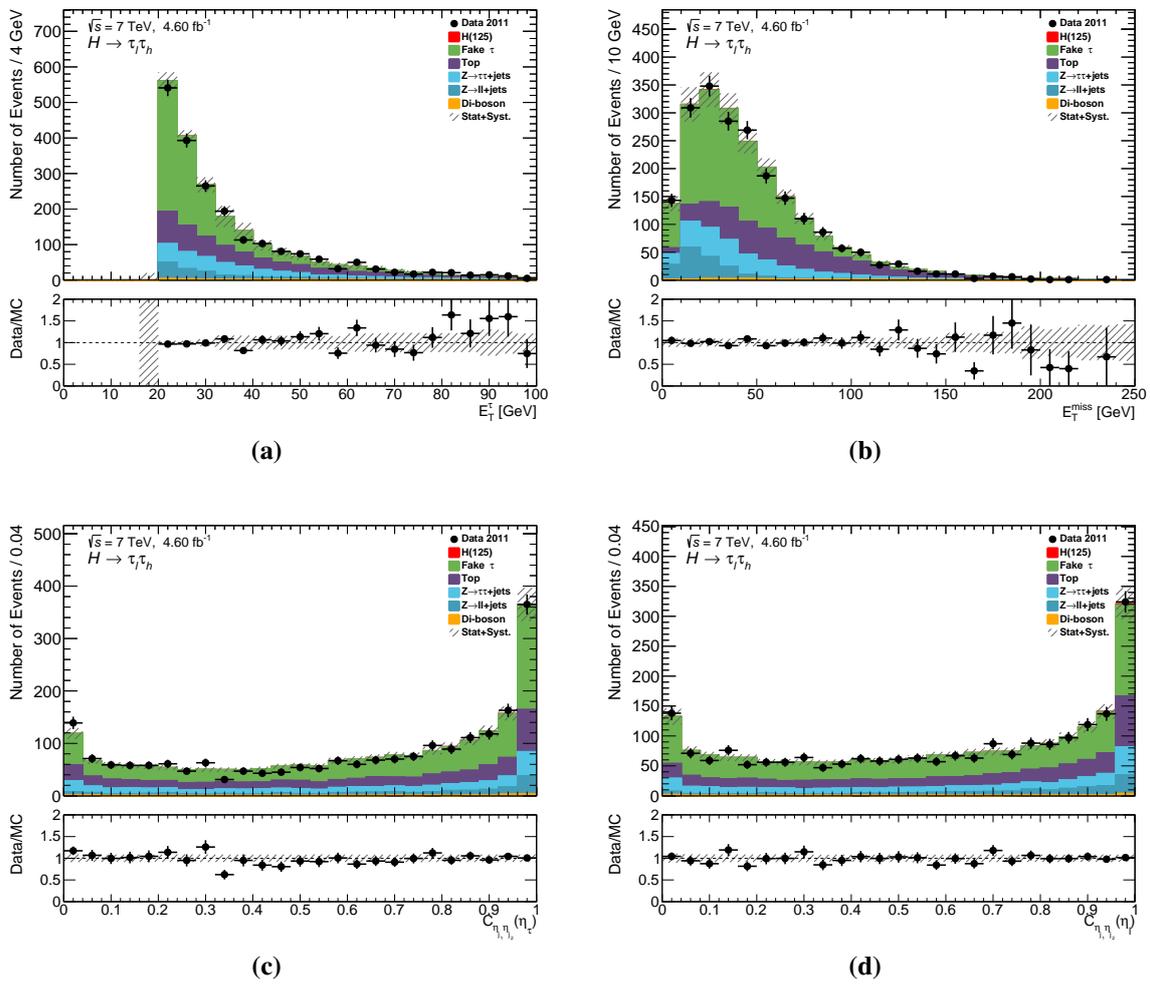


Figure 5.3.: Tau and lepton kinematic variables fed into BDT_{HBD} . Uncertainties are statistical and systematic.

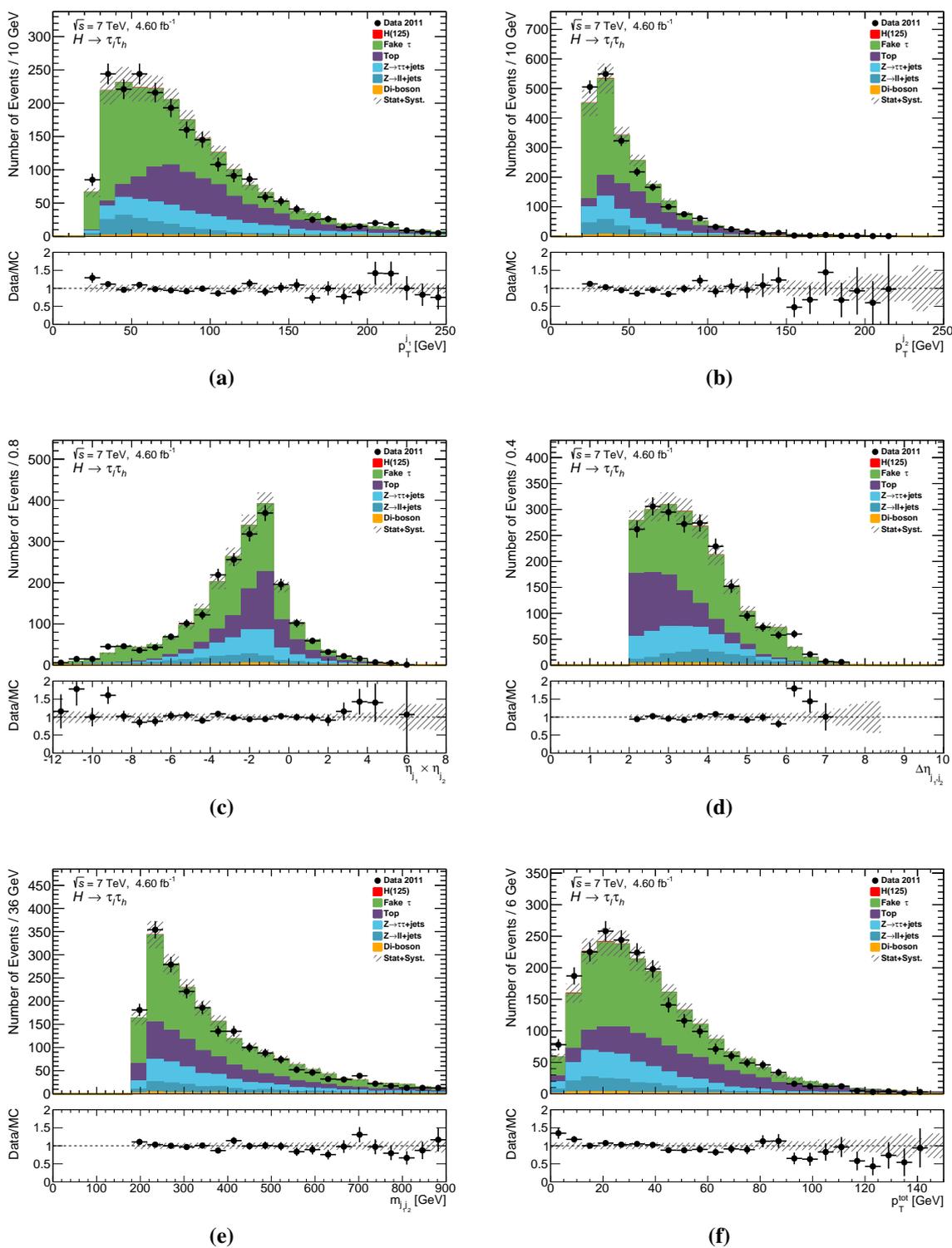


Figure 5.4.: Jet kinematic variables fed into BDT_{HBD} . Uncertainties are statistical and systematic.

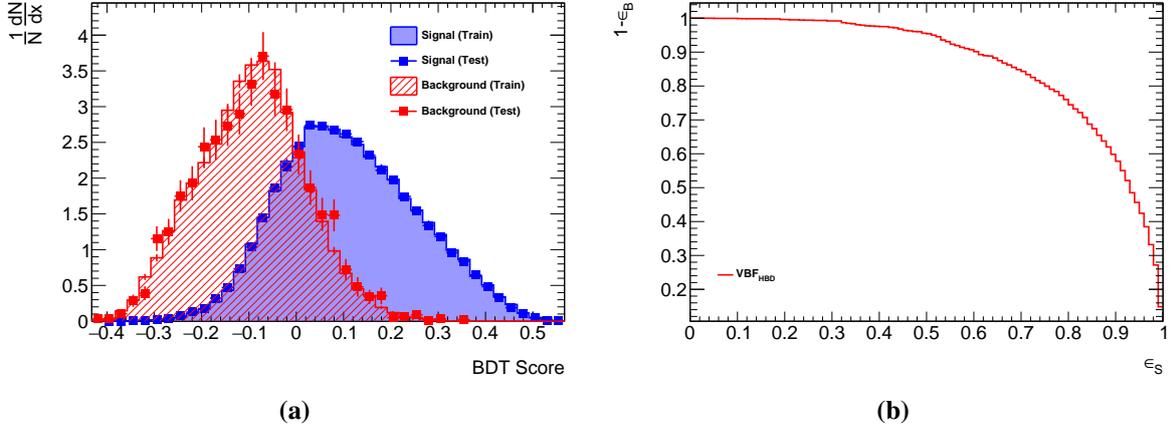


Figure 5.5.: (a) BDT score and (b) ROC curve outputs from TMVA.

function of signal efficiency (ϵ_S) for a series of cuts along the BDT score. As will be seen in Chapter 6, the ROC curve is a useful tool to analyse the performance of a BDT. The area under the ROC curve will tend towards 1 as the performance increases, approaching a step-like shape.

The BDT_{HBD} can be seen to separate the signal and background samples, with the signal events peaking at a BDT_{HBD} score of ≈ 0.1 , and the background events at ≈ -0.1 . It must be noted however that the signal and background samples here are both normalised to the same yield, and the background consists solely of $Z \rightarrow \tau\tau$ events. To see the true separation power of the discriminant, it must be applied back to all appropriately normalised background events.

Figure 5.6a shows the BDT_{HBD} score after being applied to the full data set, for the $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ category before any signal region selection. The apparent separation here is less significant than observed in Figure 5.5, however the location of the peak for the $m_H = 125$ GeV signal is still shifted from the peak from the major background components. In order to remain blind to any potential signal, data is not shown in this distribution. To validate the background model for BDT_{HBD} , a side-band region is defined as the $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ selection with $m_{\text{MMC}} < 100$ GeV or $m_{\text{MMC}} > 150$ GeV, with BDT_{HBD} shown in Figure 5.6b. While these events may make it into the final fit, they are away from any expected signal and so provide a suitable data-set to check the background model. Good agreement between data and MC is seen across the whole BDT_{HBD} range.

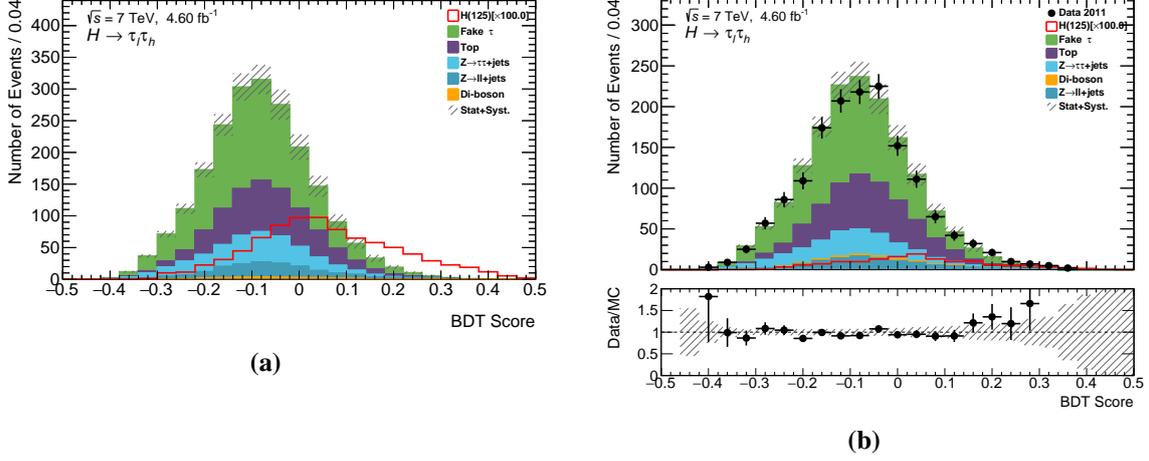


Figure 5.6.: BDT_{HBD} after (a) $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ categorisation and (b) after $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ and a side-band selection of $m_{\text{MMC}} < 100 \text{ GeV}$ or $m_{\text{MMC}} > 150 \text{ GeV}$. $H \rightarrow \tau\tau$ signal is $\times 100$. Uncertainties are statistical and systematic.

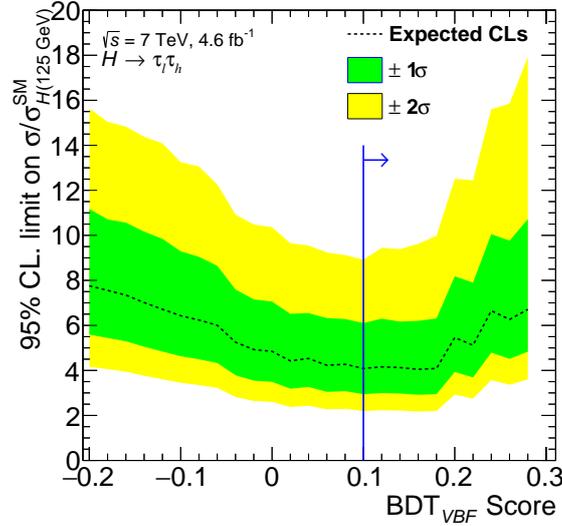


Figure 5.7.: Exclusion as a function of BDT_{HBD} with $m_H = 125 \text{ GeV}$. Blue line indicates final cut value chosen. Uncertainties are statistical and systematic.

BDT_{HBD} Cut

The final step of the VBF category selection is to cut on the BDT_{HBD} . In order to choose x_{HBD} , a series of cuts on the BDT_{HBD} score were placed in the range $(-0.2, 0.3)$ in steps of 0.02. An exclusion limit for the $\text{VBF}_{\text{HBD}}^x$ category was then calculated, where x is the cut value on BDT_{HBD} . The method used to extract the limit was identical to that in the cut-based

analysis; namely fitting the m_{MMC} distribution after following the rebinning procedure outlined in Algorithm 4.1. Figure 5.7 shows the 95% confidence level exclusion limit as a function of BDT_{HBD} with $m_H = 125$ GeV.

The value of x that achieved the most sensitive and stable limit, while retaining the highest level of statistics, was chosen as x_{HBD} ; thus, the final VBF categorisation cut is $\text{BDT}_{\text{HBD}} > 0.1$.

BDT_{HBD} Interpretation

As MVA methods can be somewhat of a black-box, it is important to try and develop an understanding of why the BDT is outperforming the cut-based approach. Figures 5.8 and 5.9 show the 2D correlations between the tau system variables, and the jet variables that are fed into BDT_{HBD} . Perhaps of most importance here is the row of plots on the far right, which shows the BDT_{HBD} score against each input variable. Take for instance the plot of BDT_{HBD} versus $C_{j_1 j_2}(\eta_\tau)$. In the cut-based analysis, a requirement of $C_{j_1 j_2}(\eta_\tau) > 1/e$ is placed³. From the figure it can be seen that, although it does reject a significant amount of the $Z \rightarrow \tau\tau$ background, a large amount still exists towards $C_{j_1 j_2}(\eta_\tau) = 1$. When viewed along the BDT_{HBD} axis however, the signal and background events much more clearly separated. The final categorisation cut of $\text{BDT}_{\text{HBD}} > 0.1$, is much more efficient at rejecting the $Z \rightarrow \tau\tau$ background and retaining signal. This same result can be seen for all the variables going into BDT_{HBD} .

5.2.3. Results

The event yields for the final hybrid categories are shown in Tables 5.3 and 5.4. Exclusion limits are extracted for the hybrid categories the same way as for the cut-based analysis, including the same systematic treatments. Figure 5.11 shows the binned m_{MMC} distributions for each analysis category that were passed into the fit. The fit model can be validated using the pull distribution for the combined fit in Figure 5.10. Although the impact of each nuisance parameter has changed slightly from the cut-based analysis, the pulls have remained similar, and can be compared to Figure 4.15. The pull distributions for each individual category, as well as the non-jet combination fit, can be found in Appendix B.

³In the cut-based analysis this variable is referred to as the tau η centrality, and the requirement for it to be true corresponds to the cut here.

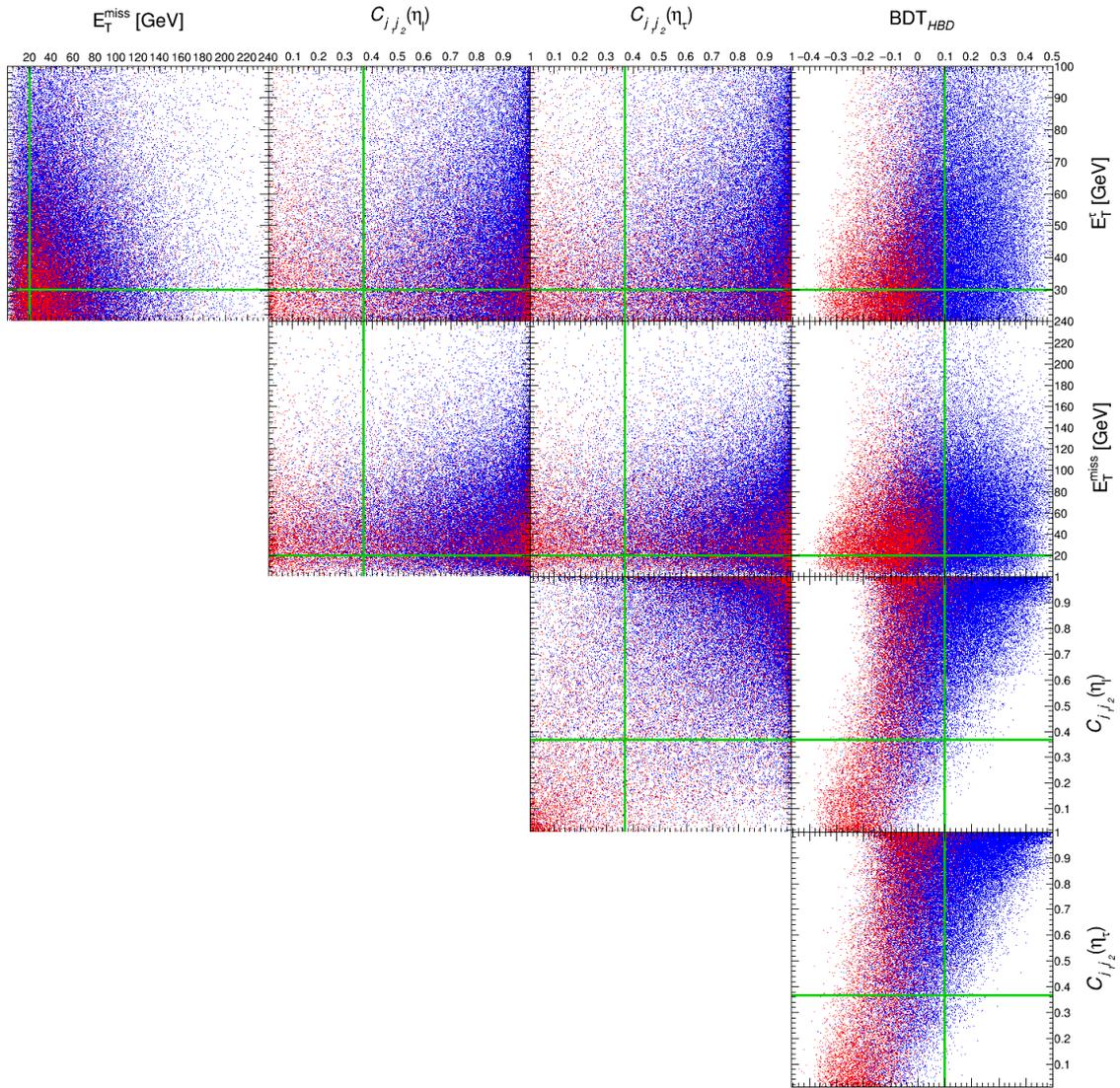


Figure 5.8.: Correlations between the tau system variables and BDT_{HBD} . Blue data points are VBF Higgs signal for every 5 GeV mass point within $m_H = [100, 150]$ GeV, red data points are ALPGEN VBF filtered $Z \rightarrow \tau\tau$ MC, and the green lines represent the positions of the cuts made in the cut-based analysis, and the final BDT_{HBD} cut.

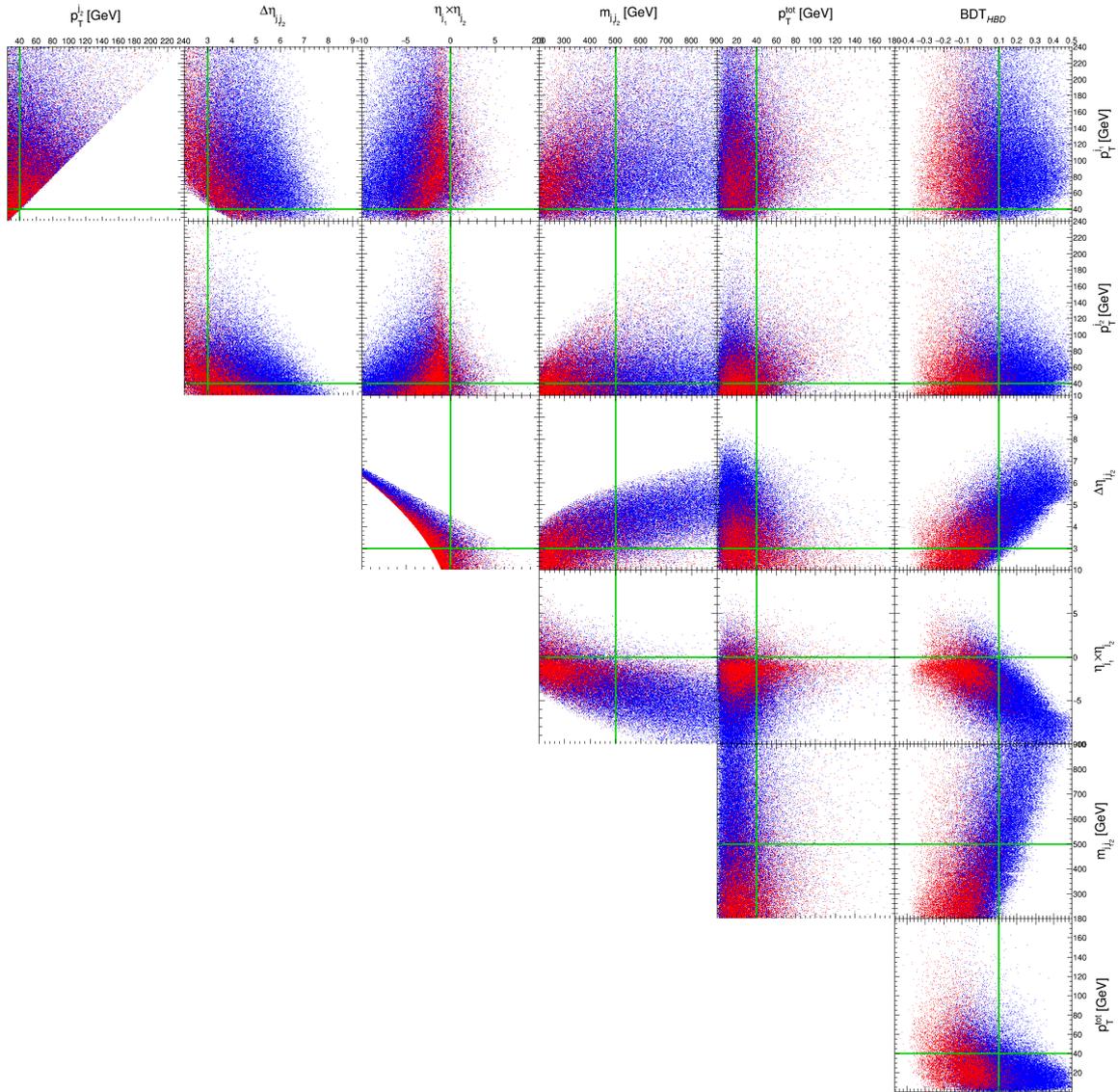


Figure 5.9.: Correlations between the jet variables and BDT_{HBD} . Blue data points are VBF Higgs signal for every 5 GeV mass point within $m_H = [100, 150]$ GeV, red data points are ALPGEN VBF filtered $Z \rightarrow \tau\tau$ MC, and the green lines represent the positions of the cuts made in the cut-based analysis, and the final BDT_{HBD} cut.

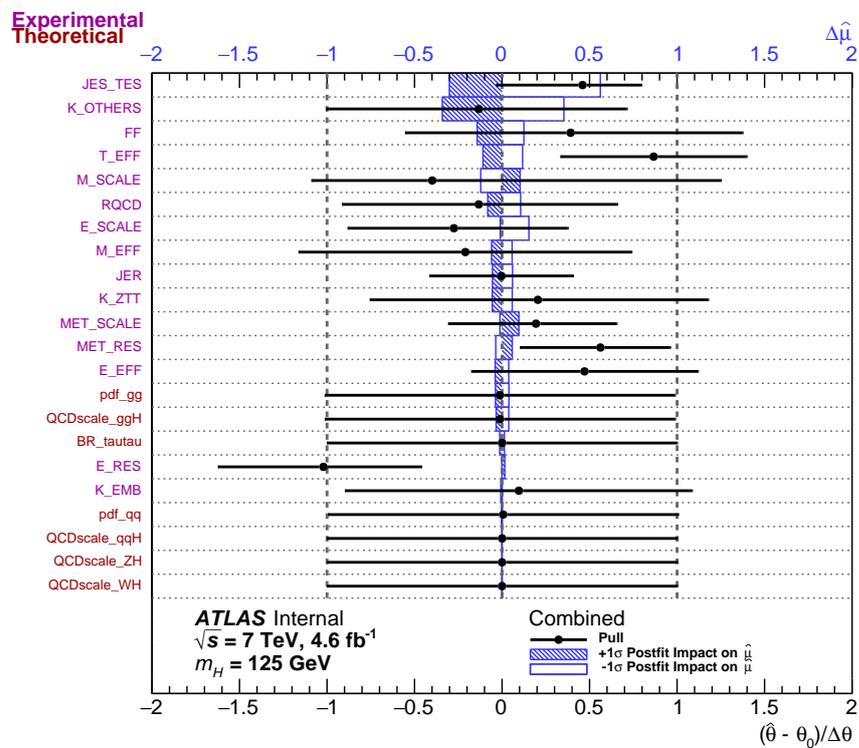


Figure 5.10.: Pull distribution for the combined categories in the hybrid analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

Process	VBF	Boosted
ggF 125	$0.389 \pm 0.039 \pm \begin{smallmatrix} 0.082 \\ 0.059 \end{smallmatrix}$	$3.88 \pm 0.11 \pm \begin{smallmatrix} 0.21 \\ 0.24 \end{smallmatrix}$
VBF 125	$1.853 \pm 0.029 \pm \begin{smallmatrix} 0.182 \\ 0.195 \end{smallmatrix}$	$1.165 \pm 0.023 \pm \begin{smallmatrix} 0.059 \\ 0.058 \end{smallmatrix}$
WH 125	< 0.01	$0.528 \pm 0.036 \pm \begin{smallmatrix} 0.031 \\ 0.029 \end{smallmatrix}$
ZH 125	< 0.01	$0.297 \pm 0.021 \pm \begin{smallmatrix} 0.019 \\ 0.018 \end{smallmatrix}$
$Z \rightarrow \tau\tau$	$9.75 \pm 0.55 \pm \begin{smallmatrix} 3.42 \\ 2.26 \end{smallmatrix}$	$532 \pm 14 \pm 26$
Fake- τ	$8.61 \pm 0.63 \pm \begin{smallmatrix} 1.90 \\ 1.78 \end{smallmatrix}$	
SS-Data		$65.5 \pm 8.3 \pm 3.8$
$W(\rightarrow \ell\nu) + \text{jets}$		$35.6 \pm 5.5 \pm \begin{smallmatrix} 9.6 \\ 6.6 \end{smallmatrix}$
$t\bar{t}$	$2.84 \pm 0.38 \pm \begin{smallmatrix} 0.59 \\ 0.30 \end{smallmatrix}$	$41.3 \pm 1.4 \pm \begin{smallmatrix} 4.3 \\ 4.1 \end{smallmatrix}$
$Z \rightarrow \ell\ell$	$0.25 \pm 0.18 \pm \begin{smallmatrix} 0.22 \\ 0.20 \end{smallmatrix}$	$4.7 \pm 1.3 \pm \begin{smallmatrix} 1.5 \\ 1.7 \end{smallmatrix}$
Di-boson	$0.369 \pm 0.073 \pm \begin{smallmatrix} 0.050 \\ 0.039 \end{smallmatrix}$	$7.42 \pm 0.34 \pm \begin{smallmatrix} 0.66 \\ 0.75 \end{smallmatrix}$
Signal	$2.244 \pm 0.049 \pm 0.200$	$5.87 \pm 0.12 \pm \begin{smallmatrix} 0.22 \\ 0.25 \end{smallmatrix}$
Background	$25.2 \pm 1.2 \pm \begin{smallmatrix} 4.2 \\ 3.5 \end{smallmatrix}$	$686 \pm 18 \pm \begin{smallmatrix} 28 \\ 27 \end{smallmatrix}$
Data	35	688

Table 5.3.: Number of events after hybrid categorisation and analysis level cuts for the VBF and Boosted categories. The various backgrounds are estimated using the methods described in Section 4.7. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.

Process	1-Jet (e)	1-Jet (μ)	0-Jet (e)	0-Jet (μ)
ggF 125	$8.91 \pm 0.21 \pm \begin{smallmatrix} 0.77 \\ 0.87 \end{smallmatrix}$	$6.54 \pm 0.18 \pm \begin{smallmatrix} 0.60 \\ 0.65 \end{smallmatrix}$	$9.40 \pm 0.25 \pm \begin{smallmatrix} 0.65 \\ 0.74 \end{smallmatrix}$	$4.61 \pm 0.18 \pm \begin{smallmatrix} 0.40 \\ 0.42 \end{smallmatrix}$
VBF 125	$2.414 \pm 0.033 \pm \begin{smallmatrix} 0.138 \\ 0.159 \end{smallmatrix}$	$1.904 \pm 0.030 \pm \begin{smallmatrix} 0.121 \\ 0.128 \end{smallmatrix}$	$0.086 \pm 0.006 \pm \begin{smallmatrix} 0.009 \\ 0.007 \end{smallmatrix}$	$0.038 \pm 0.004 \pm \begin{smallmatrix} 0.003 \\ 0.002 \end{smallmatrix}$
WH 125	$0.449 \pm 0.034 \pm \begin{smallmatrix} 0.028 \\ 0.034 \end{smallmatrix}$	$0.402 \pm 0.032 \pm \begin{smallmatrix} 0.023 \\ 0.030 \end{smallmatrix}$	$0.027 \pm 0.008 \pm \begin{smallmatrix} 0.010 \\ 0.007 \end{smallmatrix}$	$0.029 \pm 0.008 \pm \begin{smallmatrix} 0.004 \\ 0.004 \end{smallmatrix}$
ZH 125	$0.285 \pm 0.020 \pm 0.019$	$0.264 \pm 0.019 \pm \begin{smallmatrix} 0.017 \\ 0.020 \end{smallmatrix}$	$0.017 \pm 0.005 \pm \begin{smallmatrix} 0.002 \\ 0.003 \end{smallmatrix}$	< 0.01
$Z \rightarrow \tau\tau$	$1485 \pm 21 \pm \begin{smallmatrix} 81 \\ 71 \end{smallmatrix}$	$1077 \pm 15 \pm 53$	$2358 \pm 25 \pm 125$	$788 \pm 13 \pm 33$
SS-Data	$814 \pm 29 \pm 42$	$314 \pm 19 \pm 19$	$388 \pm 20 \pm 22$	$60.0 \pm 12.5 \pm 3.4$
$W(\rightarrow \ell\nu) + \text{jets}$	$234 \pm 18 \pm \begin{smallmatrix} 34 \\ 33 \end{smallmatrix}$	$166 \pm 18 \pm \begin{smallmatrix} 31 \\ 25 \end{smallmatrix}$	$108.3 \pm 13.4 \pm \begin{smallmatrix} 9.5 \\ 10.2 \end{smallmatrix}$	$55.9 \pm 12.3 \pm \begin{smallmatrix} 9.3 \\ 3.8 \end{smallmatrix}$
$t\bar{t}$	$103.2 \pm 2.7 \pm \begin{smallmatrix} 5.7 \\ 6.9 \end{smallmatrix}$	$78.3 \pm 2.3 \pm 12.9$	$0.62 \pm 0.21 \pm \begin{smallmatrix} 0.27 \\ 0.15 \end{smallmatrix}$	$0.43 \pm 0.15 \pm \begin{smallmatrix} 0.21 \\ 0.12 \end{smallmatrix}$
$Z \rightarrow \ell\ell$	$34.1 \pm 4.8 \pm \begin{smallmatrix} 15.6 \\ 12.7 \end{smallmatrix}$	$15.5 \pm 2.9 \pm \begin{smallmatrix} 11.3 \\ 9.5 \end{smallmatrix}$	$48.4 \pm 5.5 \pm \begin{smallmatrix} 15.9 \\ 15.8 \end{smallmatrix}$	$13.1 \pm 2.7 \pm \begin{smallmatrix} 2.0 \\ 1.3 \end{smallmatrix}$
Di-boson	$12.7 \pm 1.1 \pm 1.1$	$9.07 \pm 1.04 \pm \begin{smallmatrix} 0.98 \\ 1.01 \end{smallmatrix}$	$2.29 \pm 0.59 \pm \begin{smallmatrix} 0.47 \\ 0.28 \end{smallmatrix}$	$2.41 \pm 0.30 \pm \begin{smallmatrix} 0.33 \\ 0.22 \end{smallmatrix}$
Signal	$12.1 \pm 0.22 \pm \begin{smallmatrix} 0.79 \\ 0.89 \end{smallmatrix}$	$9.11 \pm 0.19 \pm \begin{smallmatrix} 0.61 \\ 0.66 \end{smallmatrix}$	$9.53 \pm 0.25 \pm \begin{smallmatrix} 0.65 \\ 0.74 \end{smallmatrix}$	$4.67 \pm 0.18 \pm \begin{smallmatrix} 0.40 \\ 0.42 \end{smallmatrix}$
Background	$2684 \pm 41 \pm \begin{smallmatrix} 98 \\ 90 \end{smallmatrix}$	$1657 \pm 31 \pm \begin{smallmatrix} 66 \\ 64 \end{smallmatrix}$	$2905 \pm 35 \pm 128$	$919 \pm 20 \pm \begin{smallmatrix} 35 \\ 33 \end{smallmatrix}$
Data	2825	1702	3025	957

Table 5.4.: Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to the embedding sample. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.

Hybrid Exclusion Limit

Figure 5.12 shows the 95% confidence level limits on σ/σ_H^{SM} as a function of assumed m_H for each analysis category. The observed limits are within 2σ of the expected limits for all analysis categories. Much like the cut-based analysis limits, the 1- and 0-Jet categories show the largest deviations between observed and expected limits. This is again expected – as can be seen from Table 5.4, these categories are almost identical between the two analysis in terms of event yields. The introduction of VBF_{HBD} only has an effect on the VBF and Boosted categories.

The VBF category has an improved expected exclusion limit of $4.2_{-1.2}^{+1.9} \times \sigma_H^{SM}$, with an observed limit of $5.5 \times \sigma_H^{SM}$. The limit from the Boosted category has actually increased compared to the cut-based analysis, with an expected (observed) limit of $5.7_{-1.6}^{+2.8} \times \sigma_H^{SM}$ ($8.0 \times \sigma_H^{SM}$). The source of these differences will be discussed in the next section.

The combined exclusion limit for the hybrid analysis, as well as the statistical significance of the result is given in Figure 5.13. An expected limit is set at $2.7_{-0.7}^{+1.2} \times \sigma_H^{SM}$ with an observed limit of $3.1 \times \sigma_H^{SM}$. No significant excess is seen in the data, with an observed significance of 0.2σ .

Finally, Figure 5.14 shows a comparison of the exclusion limits for each hybrid analysis category, and the combined limit, as well as the signal strengths. Much like the cut-based analysis, each category is consistent with the background-only hypothesis and the existence of a $m_H = 125$ GeV Higgs boson to within 2σ . The observed signal strength is $\hat{\mu} = 0.3 \pm 1.2$.

Comparison to Cut-Based Analysis

While the exclusion limits, signal strengths, and statistical significance of the results from both analyses agree within 1σ , the introduction of BDT_{HBD} to select VBF-type events has improved the central value of the combined expected exclusion limit and reduced the uncertainties from $2.9_{-0.8}^{+1.4}$ to $2.7_{-0.7}^{+1.2} \times \sigma_H^{SM}$. While not a significant decrease, it does show that a naive implementation of multi-variate techniques into the event categorisation process does improve the sensitivity of the analysis.

It is interesting to look at the events that end up in the new VBF_{HBD} category compared to the cut-based category. As seen in Table 5.3 and Figure 5.11, there is significantly more contribution fake- τ s in the VBF_{HBD} category – roughly equal to the $Z \rightarrow \tau\tau$ contribution – compared to $\approx 20\%$ of the $Z \rightarrow \tau\tau$ events in the cut-based categorisation. This however is not surprising, as BDT_{HBD} was only trained against $Z \rightarrow \tau\tau$ backgrounds, and so was less

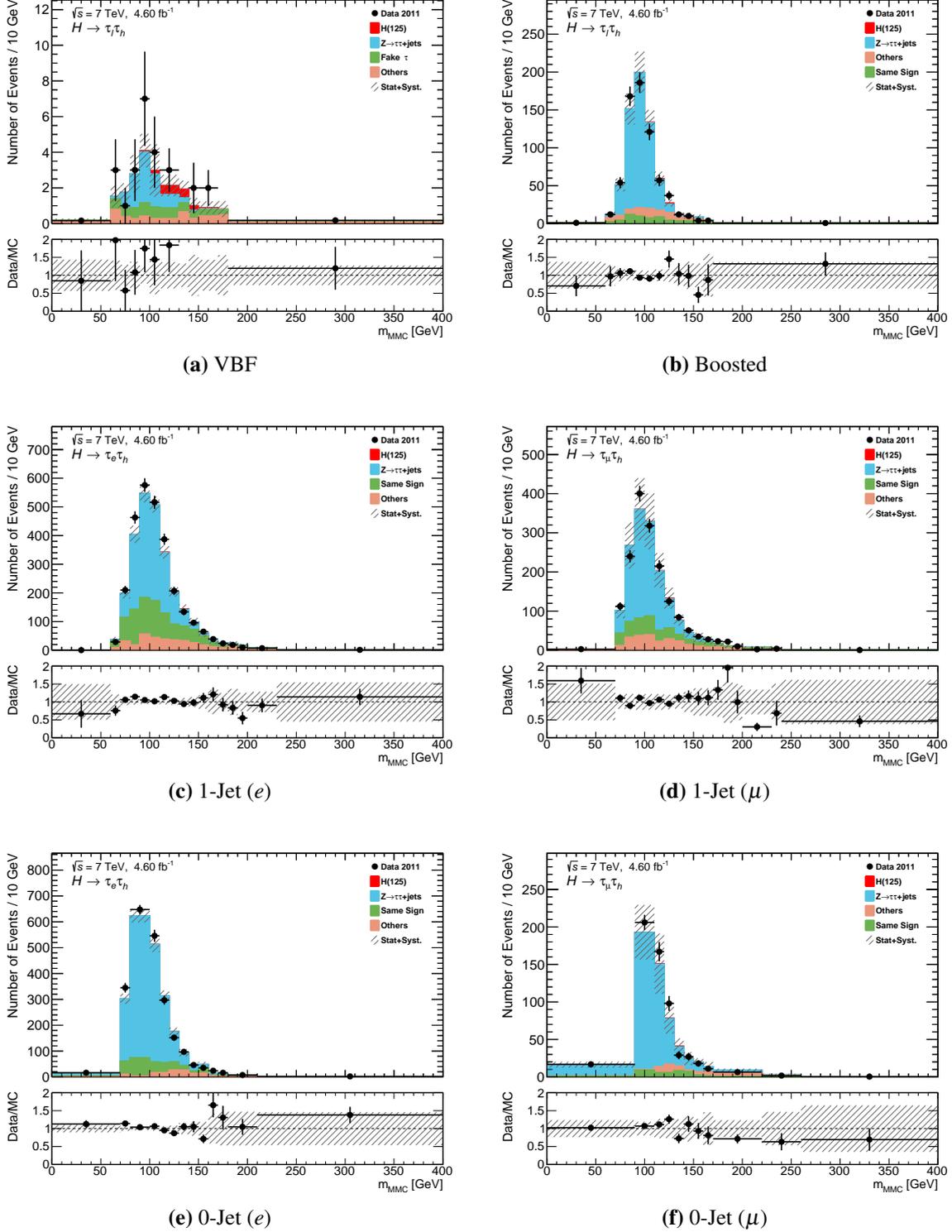


Figure 5.11.: m_{MMC} pre-fit distributions in the signal region for each analysis category. $Z \rightarrow \ell\ell$, $t\bar{t}$, di-boson, and $W(\rightarrow \ell\nu) + \text{jets}$ backgrounds are grouped into the *others* category. Uncertainties are statistical and systematic. Bins are merged according to Algorithm 4.1, with bin contents scaled to number of events for every 10 GeV.

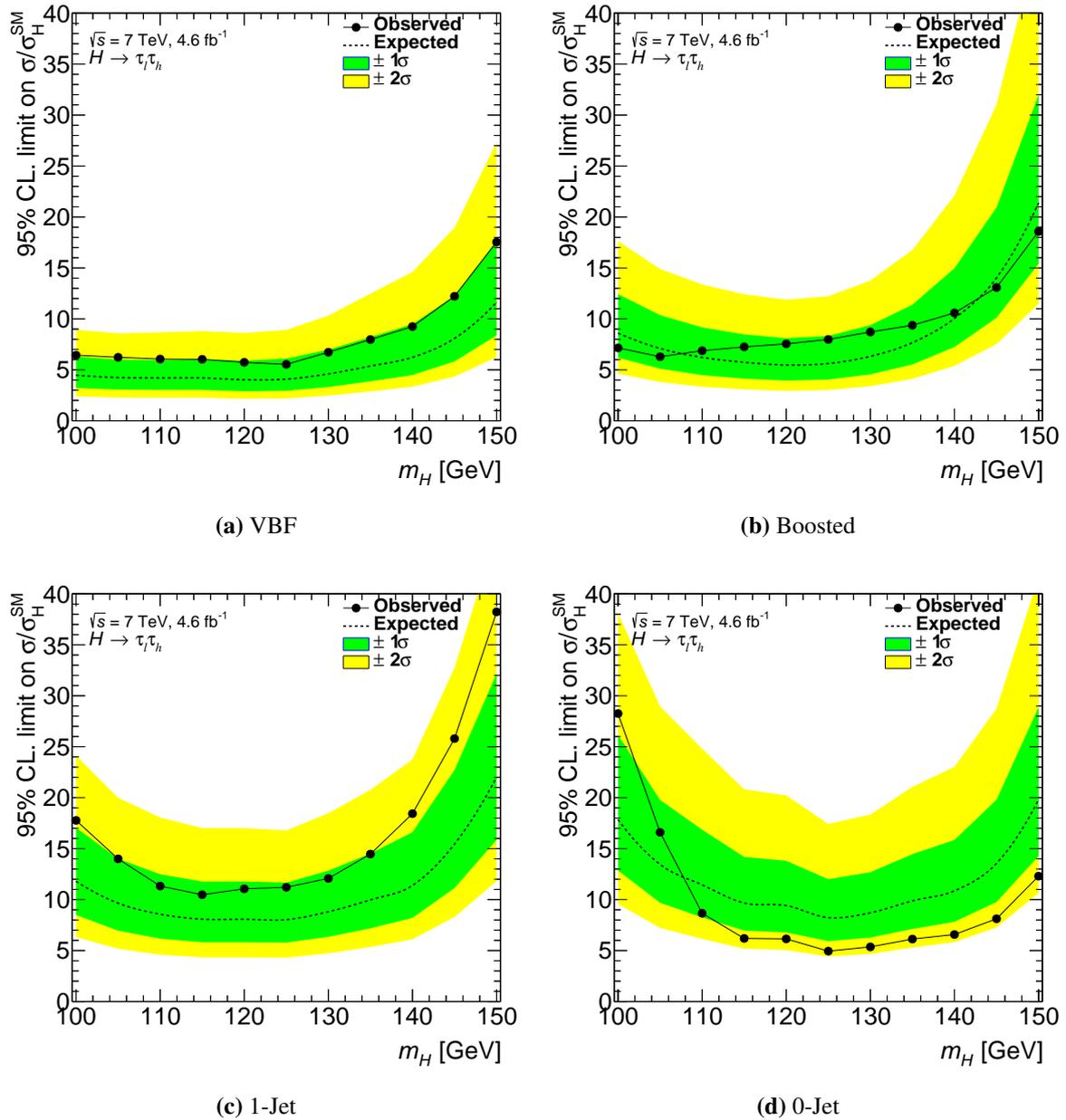


Figure 5.12.: Hybrid exclusion limits for VBF, Boosted, and 1-Jet analysis categories. Uncertainties are statistical and systematic.

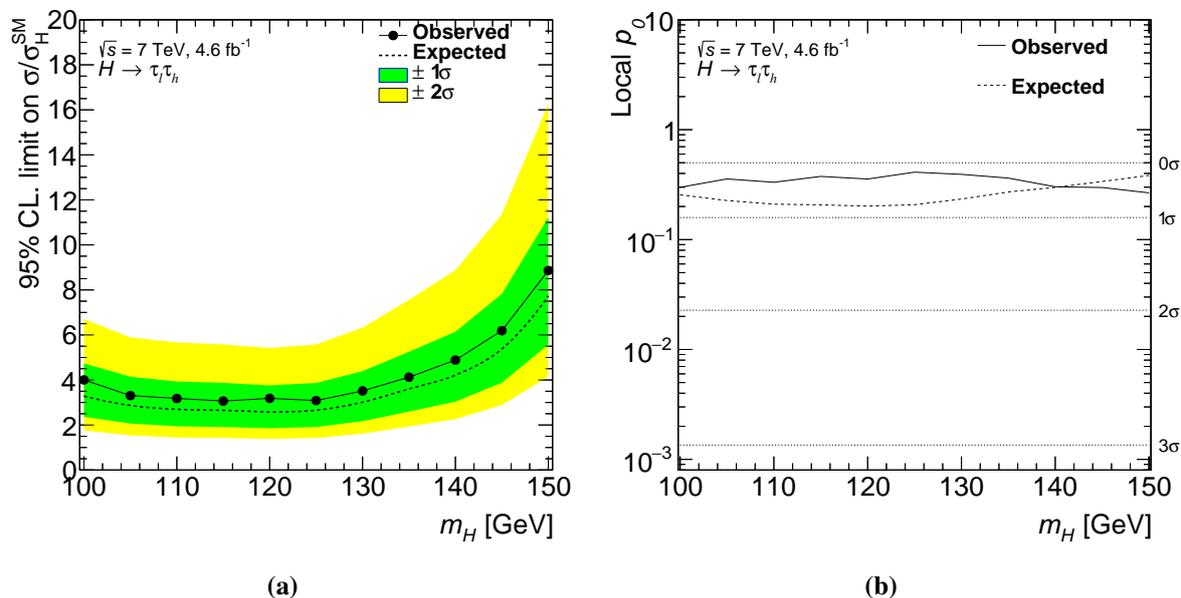
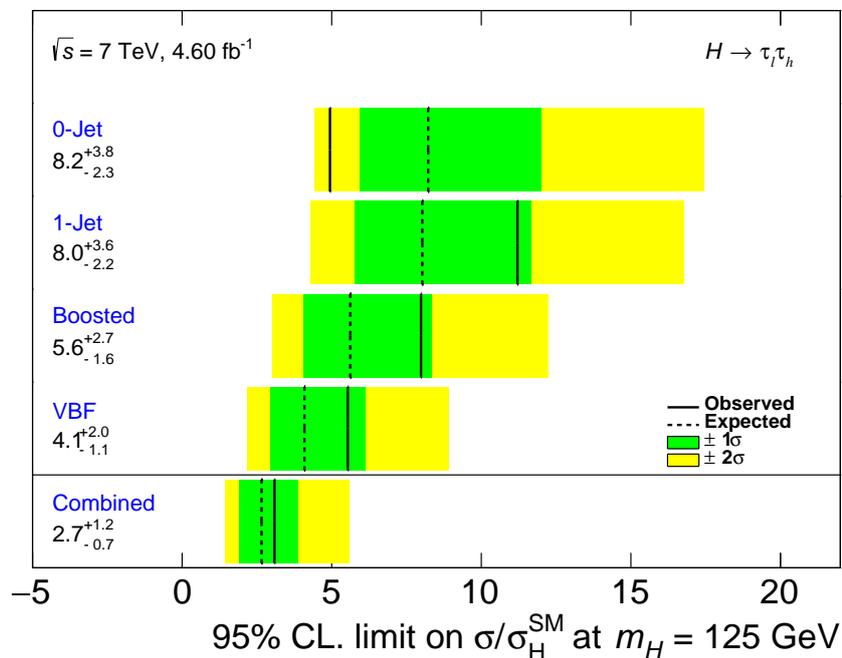


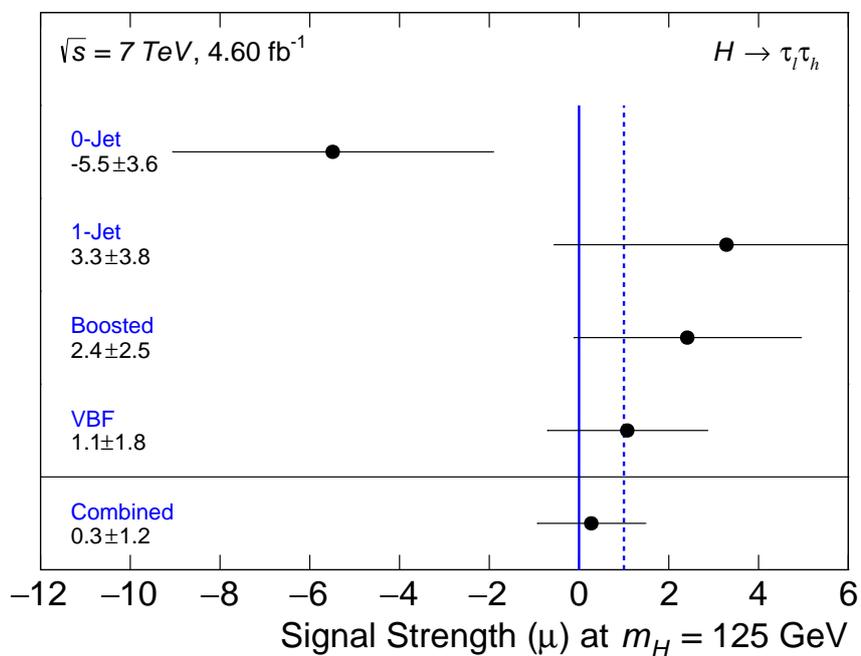
Figure 5.13.: Hybrid (a) exclusion limits and (b) signal significance for VBF, Boosted, 1-Jet, and 0-Jet analysis categories combined. Uncertainties are statistical and systematic.

likely to be able to discriminate signal events from fake- τ events. The increase in the Boosted limit can be attributed to ≈ 0.5 of its signal events moving to the VBF category in the new categorisation while the background yields stay the same, corresponding to roughly a 10% reduction in signal. The sensitivity of the boosted category can be recovered with further optimisation of the multi-variate procedure.

The next chapter takes the multi-variate approach one step further; BDTs for *each* analysis category will be constructed, and optimised for variable inputs and training parameters. It will be seen that the gain in sensitivity from the hybrid method can be improved even further by adopting a full-MVA approach.



(a)



(b)

Figure 5.14.: (a) Comparison of hybrid exclusion limits for each individual analysis category, and the combined limit, for $m_H = 125 \text{ GeV}$. Uncertainties are statistical and systematic. (b) Extracted signal strength μ at $m_H = 125 \text{ GeV}$ for each analysis category, and the combined signal strength. Solid line is $\mu = 0$ and dashed line is $\mu = 1$

Chapter 6.

$H \rightarrow \tau_\ell \tau_h$ Multi-Variate Analysis

This chapter describes a full multi-variate-analysis (MVA) approach to the $H \rightarrow \tau_\ell \tau_h$ search in ATLAS. Rather than simply using the cut-based analysis variables to construct a BDT as was done in the hybrid analysis, an optimised set of variables were chosen to build four BDTs – one for each analysis category. The analysis strategy presented here relies heavily on the methods and results of the previous two chapters, with its performance being compared to both the cut-based and hybrid methods.

6.1. MVA Categorisation

Like the previous analyses, the MVA method employs the analysis category technique to focus on the VBF production mode, the boosted Higgs scenario, and n -jet topologies. A base categorisation splits the events loosely into five groups – the VBF, Boosted, 1-Jet, and 0-Jet categories as before, and now a fifth *Rest* category, defined to recover events. The category definitions are shown in Table 6.1.

The VBF_{MVA} category is defined similarly to VBF_{HBD} , however the jet p_{T} thresholds are increased to $p_{\text{T}} > 40$ GeV, and a veto on events containing a b -tagged jet is applied. The $\text{Boosted}_{\text{MVA}}$ category is loosely defined to be those events that fail VBF_{MVA} , with $p_{\text{T}}^H > 100$ GeV, and again no b -tagged jets. The 1- and 0-Jet categories are defined by their jet multiplicity and are required to fail the proceeding category selections. Finally the *Rest* category contains all those events that do not fall into any other category.

The base categories are used for construction of the BDTs (Section 6.3), as well as calculation of k -factors (Section 6.2.1), and fake-factors (Section 6.2.2). For extracting limits however,

	VBF	Boosted	1-Jet	0-Jet	Rest
Base	$p_T^{j_1(j_2)} > 40 \text{ GeV}$ $\Delta\eta_{j_1, j_2} > 2$ $m_{j_1, j_2} > 200 \text{ GeV}$ no b-tag	not VBF $p_T^H > 100 \text{ GeV}$ no b-tag	not Boosted $p_T^{j_1} > 25 \text{ GeV}$	not 1-Jet $p_T^{j_1} < 25 \text{ GeV}$	not 0-Jet
SR	$\text{BDT}_{\text{VBF}} > x_{\text{VBF}}$	$\text{BDT}_{\text{BST}} > x_{\text{BST}}$	$\text{BDT}_{\text{1J}} > x_{\text{1J}}$	$\text{BDT}_{\text{0J}} > x_{\text{0J}}$	$m_T < 30 \text{ GeV}$ $\cancel{E}_T > 20 \text{ GeV}$
$t\bar{t}$	$m_T > 50 \text{ GeV}$ b-tag	$m_T > 50 \text{ GeV}$ b-tag	$m_T > 50 \text{ GeV}$	b-tag	$m_T > 50 \text{ GeV}$ b-tag
W	$p_T^{j_1(j_2)} > 25 \text{ GeV}$ $m_T > 70 \text{ GeV}$	$p_T^H > 50 \text{ GeV}$ $m_T > 70 \text{ GeV}$	$m_T > 70 \text{ GeV}$ no b-tag	$m_T > 70 \text{ GeV}$	$m_T > 70 \text{ GeV}$ no b-tag
$Z \rightarrow \tau\tau$	$m_T < 70 \text{ GeV}$ $40 < m_{\text{vis}} < 80 \text{ GeV}$ no b-tag				
QCD	$p_T^{j_1(j_2)} > 25 \text{ GeV}$ $m_T < 50 \text{ GeV}$ no isolated lepton	$p_T^H > 50 \text{ GeV}$ $m_T < 50 \text{ GeV}$ no isolated lepton			

Table 6.1.: Category and analysis cuts for the MVA analysis. SR cuts were not placed when training the category-specific BDTs. $t\bar{t}$, W, $Z \rightarrow \tau\tau$, QCD refer to control regions applied on top of the category selection for use in calculating k -factors and fake-factors (for the VBF and Boosted categories only). Any cuts in the CRs supersede those in the categorisation.

the Signal Region (SR) cuts are also applied. For the four main categories, this involves a cut on the category-specific BDT, with failing events flowing through to the next category. As a BDT is not constructed for the Rest category, SR cuts of $m_T < 30 \text{ GeV}$ and $\cancel{E}_T > 20 \text{ GeV}$ are applied to reduce Fake- τ and $Z \rightarrow \tau\tau$ background contributions.

6.2. Background Estimation Methods

As the analysis categories have changed significantly from the cut-based method, the background estimation methods described in Section 4.7 are repeated here, with the same control region definitions applied. Also recalculated are the embedding correction factors. For the fake-factor method, the same fake-factors calculated for the VBF category in Section 4.7.6 are used again here, however a second fake-factor estimation for the Boosted category is also calculated.

Category	$k_{OS}^{\tau\bar{\tau}}$	$k_{SS}^{\tau\bar{\tau}}$	k_{OS}^W	k_{SS}^W	$k^{Z \rightarrow \tau\tau}$	k^{emb}	$k^{Z \rightarrow \ell\ell}$	$k^{Z/\gamma^* (\rightarrow \ell\ell) + jets}$
VBF	e	0.76 ± 0.08			0.72 ± 0.10		0.98 ± 0.24	1.00 ± 0.34
	μ						1.00 ± 0.00	
Boosted	e	0.93 ± 0.04				0.119 ± 0.002	0.98 ± 0.24	0.56 ± 0.03
	μ					0.081 ± 0.002	1.00 ± 0.00	
1-Jet	e	0.85 ± 0.05	1.10 ± 0.20	0.63 ± 0.02	0.80 ± 0.03	0.119 ± 0.002	0.98 ± 0.24	0.56 ± 0.03
	μ	0.80 ± 0.05	1.02 ± 0.18	0.61 ± 0.01	0.74 ± 0.03	0.081 ± 0.002	1.00 ± 0.00	
0-Jet	e	0.85 ± 0.06	1.10 ± 0.17	0.57 ± 0.01	0.81 ± 0.03	0.119 ± 0.002	0.98 ± 0.24	0.56 ± 0.03
	μ	0.80 ± 0.05	1.02 ± 0.18	0.54 ± 0.01	0.70 ± 0.02	0.081 ± 0.002	1.00 ± 0.00	
Rest	e	0.87 ± 0.04	1.12 ± 0.16	0.64 ± 0.01	0.79 ± 0.03	0.119 ± 0.002	0.98 ± 0.24	0.56 ± 0.03
	μ	0.79 ± 0.04	0.99 ± 0.15	0.62 ± 0.01	0.73 ± 0.03	0.081 ± 0.002	1.00 ± 0.00	

Table 6.2.: k -factors for each analysis category and each lepton channel. Unless explicitly stated, factors apply to both OS and SS events.

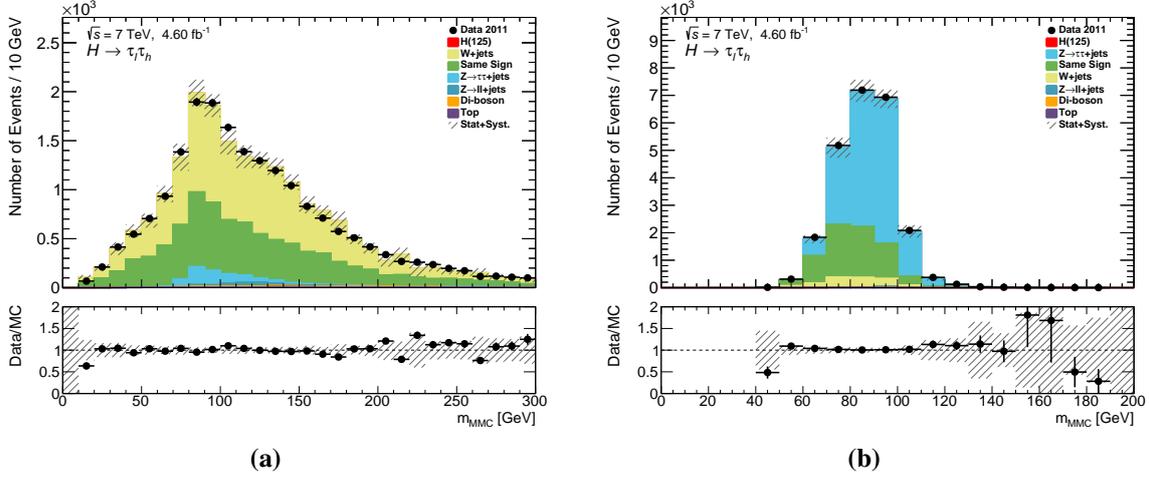


Figure 6.1.: m_{MMC} distributions for (a) the $W(\rightarrow \ell\nu) + \text{jets}$ control region at the pre-selection level after application of $k^{W,PS}$ using ALPGEN $Z \rightarrow \tau\tau$, and (b) the Embedding control region at the pre-selection level after application of k^{emb} . Uncertainties are statistical and systematic.

6.2.1. Background k -Factors

Embedding k -Factors

The embedding sample is normalised at the pre-selection level, however in this region there is a significant contribution from $W(\rightarrow \ell\nu) + \text{jets}$ events. For this reason, a pre-selection scaling of the $W(\rightarrow \ell\nu) + \text{jets}$ MC is applied in the region $m_T > 70$ GeV, and requiring no b -tagged jets, and using the ALPGEN $Z \rightarrow \tau\tau$ MC sample instead of the Embedding sample in the OS-rSS method. The $W(\rightarrow \ell\nu) + \text{jets}$ k -factors are estimated to be $k_{OS}^{W,PS} = 0.56 \pm 0.01$ and $k_{SS}^{W,PS} = 0.71 \pm 0.02$.

Using these scalings, an Embedding control region is defined with $m_T < 40$ GeV, and $40 < m_{vis} < 70$ GeV. Scaling the Embedding sample for electron and muon events separately in this region to match data gives $k_e^{emb} = 0.119 \pm 0.002$ and $k_\mu^{emb} = 0.081 \pm 0.002$. The $W(\rightarrow \ell\nu) + \text{jets}$ control region and the Embedding control region at the pre-selection level after application of $k^{W,PS}$ and k^{emb} respectively are shown in Figure 6.1.

MC k -Factors

The k -factors for the remaining MC samples are calculated as in the cut-based analysis. The factors are calculated for each category, before any application of BDTs, in order of purity

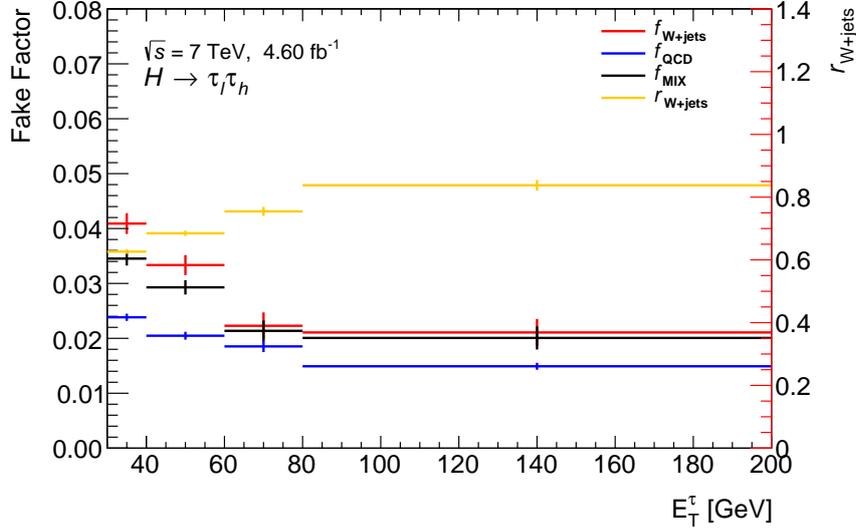


Figure 6.2.: Fake-factors and r_W values as a function of τ energy for the MVA Boosted category.

of control region; $k^{\bar{i}} \rightarrow k^W \rightarrow k^{Z\tau\tau}$. The control regions used are shown in Table 6.1. Scale factors for di-boson and $Z \rightarrow \ell\ell$ samples are taken from the cut-based analysis. The background estimation scale factors for the full-MVA analysis are shown in Table 6.2.

6.2.2. Fake-Factor for Boosted Category

In the cut-based and hybrid analyses, the Boosted category had a yield of approximately 700 background events, and 6 signal events. As will be seen later in this chapter, the purity of the MVA boosted category has greatly improved, with final yields of the order of 80 background events, and 5 signal events. For this reason, a more robust background estimation strategy was required, as the OS-rSS method becomes less reliable with low statistics, especially in the $W(\rightarrow \ell\nu) + \text{jets}$ MC sample. The fake-factor method described in Section 4.7.6 was therefore applied to the Boosted category. In the control regions defined below, f_{QCD} , f_W , and r_W were calculated to produce a Boosted-specific f_{MIX} .

Fake-factor control regions were defined as in Table 6.1, requiring $m_T > 70 \text{ GeV}$ and $p_T^H > 50 \text{ GeV}$ for the W CR, and $m_T < 50 \text{ GeV}$, $p_T^H > 50 \text{ GeV}$, and no isolated lepton requirement for the QCD CR, after failing the Loose VBF categorisation (as defined in Table 4.5). As for the VBF fake-factor, f_W and f_{QCD} were calculated from the number of anti- τ events in the two control regions, and these values combined via r_{W+jets} to form f_{MIX} . The calculation of r_{W+jets} was also performed as in the VBF case; using the number of anti- τ events in the SR in data, multiplied by the ratio of anti- τ events in the SR to the W CR in $W(\rightarrow \ell\nu) + \text{jets}$ MC

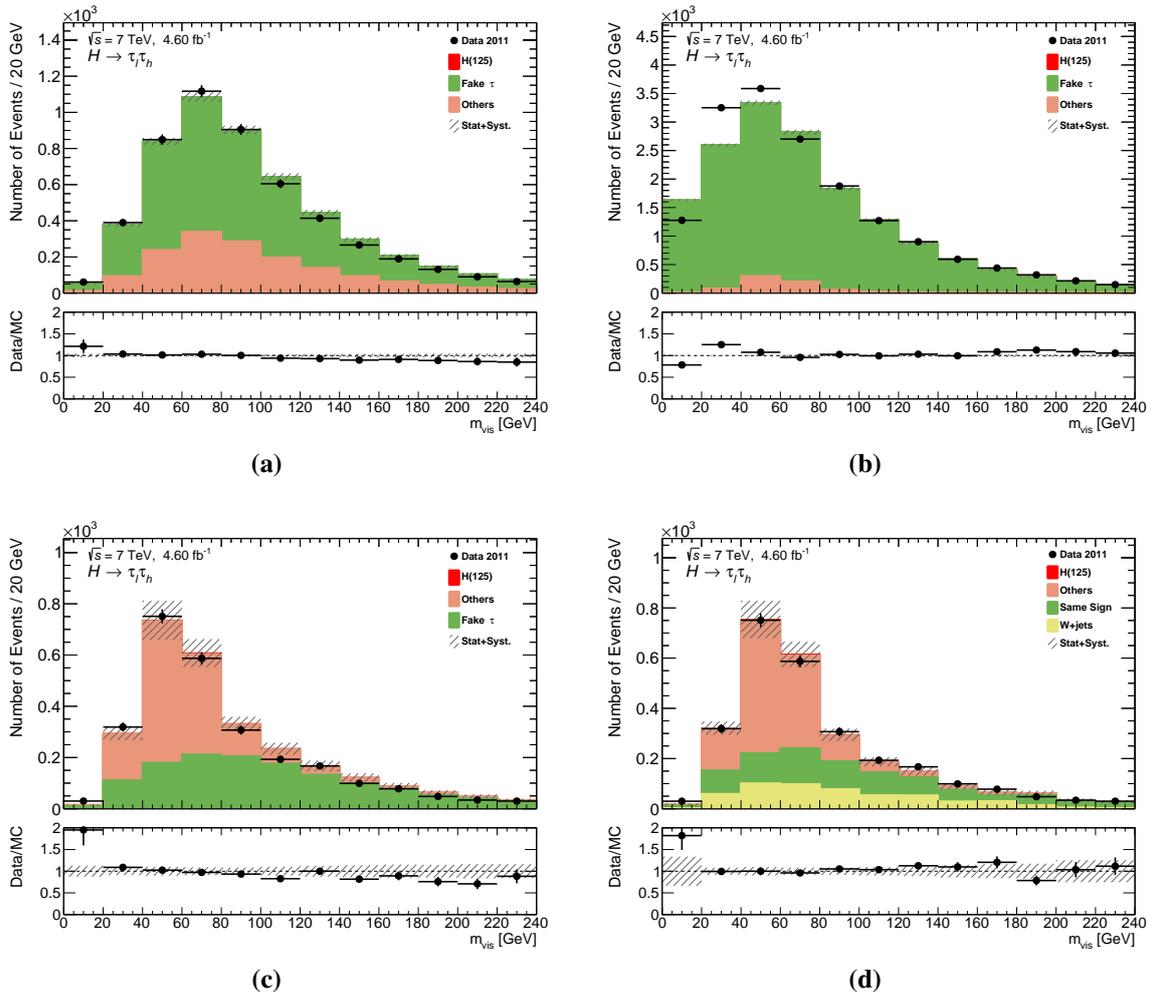


Figure 6.3.: m_{vis} distributions using the fake-factor method for the Boosted (a) W CR, (b) QCD CR, (c) SR, and (d) the SR using the OS-rSS method. The fake-factor for the W CR uses only f_W and the QCD CR uses only f_{QCD} , while the SR uses f_{MIX} . Uncertainties are statistical and systematic.

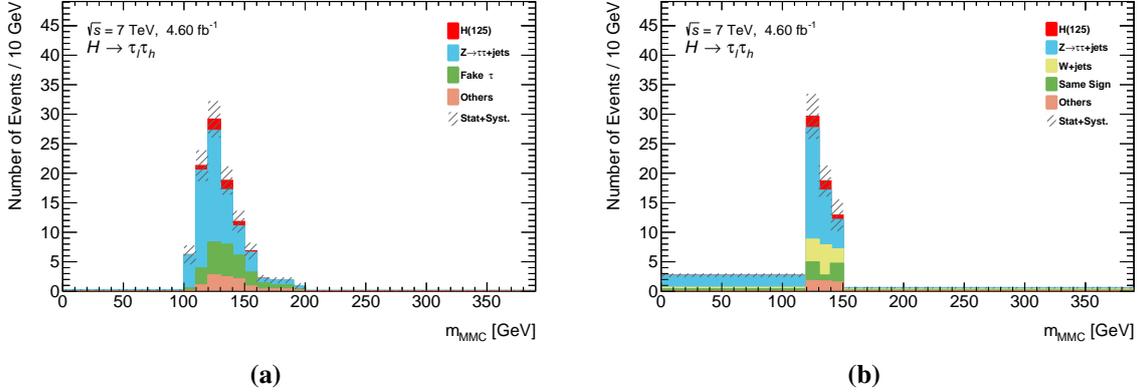


Figure 6.4.: Blinded m_{MMC} distributions after full SR cuts for the MVA Boosted category using (a) the fake-factor method and (b) the OS-rSS method, after running the rebinning procedure (Algorithm 4.1). Uncertainties are statistical and systematic.

(Equation 4.20). Figure 6.2 shows the fake-factors, and r_{W+jets} , binned by E_T^τ . Figure 6.3 shows m_{vis} distributions for the fake-factor control regions, as well as a SR comparison between the fake-factor method, and the OS-rSS method. The QCD CR shows a disagreement between the data and MC at low m_{vis} , as was also seen in the VBF case. While the source of this discrepancy is not clear, it is outside the expected signal region, and its effect on the data/MC agreement in the SR is minimal. A comparison between the two methods in the SR (after application of the BDT cuts and the rebinning procedure starting with 10 GeV bins) is shown in Figure 6.4. It can be seen that the fake-factor method results in a much better populated distribution, which allows a finer binning for the final limit extraction.

6.2.3. Background Estimation for TMVA

In the hybrid analysis, the BDT_{HBD} was constructed using only the $Z \rightarrow \tau\tau$ and Embedding samples as backgrounds. Here, it was decided to use the full background estimation. A complexity arises in the way that TMVA works in that each event must have a positive weight. Although TMVA can be instructed to ignore negative weights in the training, the current background estimation method will not work in this case. With the OS-rSS method, SS events in MC are given a negative weight ($-r_{QCD} \cdot k_{SS}^X$) so there is no double counting with the SS data sample. Similarly, in the fake-factor method, anti- τ MC events are added with negative weights so there is no double counting with the anti- τ data sample. In order to correctly estimate the backgrounds within the TMVA framework, a normalisation procedure is adopted.

For the fake-factor method, all anti- τ events in data are taken and normalised to the total yield from the fake-factor method, and the anti- τ events in MC are ignored. Similarly, for the OS-rSS method, OS events of each background are normalised to the final yields using the full OS-rSS estimation, and the SS events in MC are ignored. Both scenarios require that the background shapes of the BDT input variables for the normalised samples match those of the full background estimation samples. Figures 6.5 and 6.6 show a comparison between the background shapes and yields when using the fake-factor estimation for the VBF category, while Figures 6.7 and 6.8 show a comparison of the background shapes and yields for the OS-rSS method for the 1-Jet category, as representative examples of the estimation methods. The variables shown are described in Section 6.3.3. Except for small fluctuations, the agreement in shapes between the full estimation and the samples used in the normalisation procedure is consistent across all variables, and similar agreement is seen for the Boosted and 0-Jet categories.

6.3. Growing BDTs

This section describes the method by which BDTs were constructed for each analysis category. As the process is the same for each category, a detailed outline will be given for the VBF category only. The final BDT distributions will then be shown.

6.3.1. Training and Testing Samples

To grow the most discriminating BDT, whilst simultaneously limiting overtraining, high numbers of events are needed. The simplest way to define the training and testing samples would be to split the whole data set in half – one half for training, the other for testing. There are however a few instances where extra gains in statistics can be made.

Firstly, as in the hybrid analysis (Section 5.2.2), both the ALPGEN and Embedding samples can be used for $Z \rightarrow \tau\tau$ background estimation. For the VBF category, the Embedding sample was used to train BDT_{VBF} , while the ALPGEN $Z \rightarrow \tau\tau$ filtered MC sample was used for testing. As ALPGEN is used in the final VBF background model, it was used as the testing sample rather than the training sample to check that the tree was not overtrained. Conversely, as the Embedding sample is used in the final background model for the Boosted, 1-Jet and 0-Jet categories, it was used as the testing sample, with the unfiltered ALPGEN $Z \rightarrow \tau\tau$ MC as the training sample.

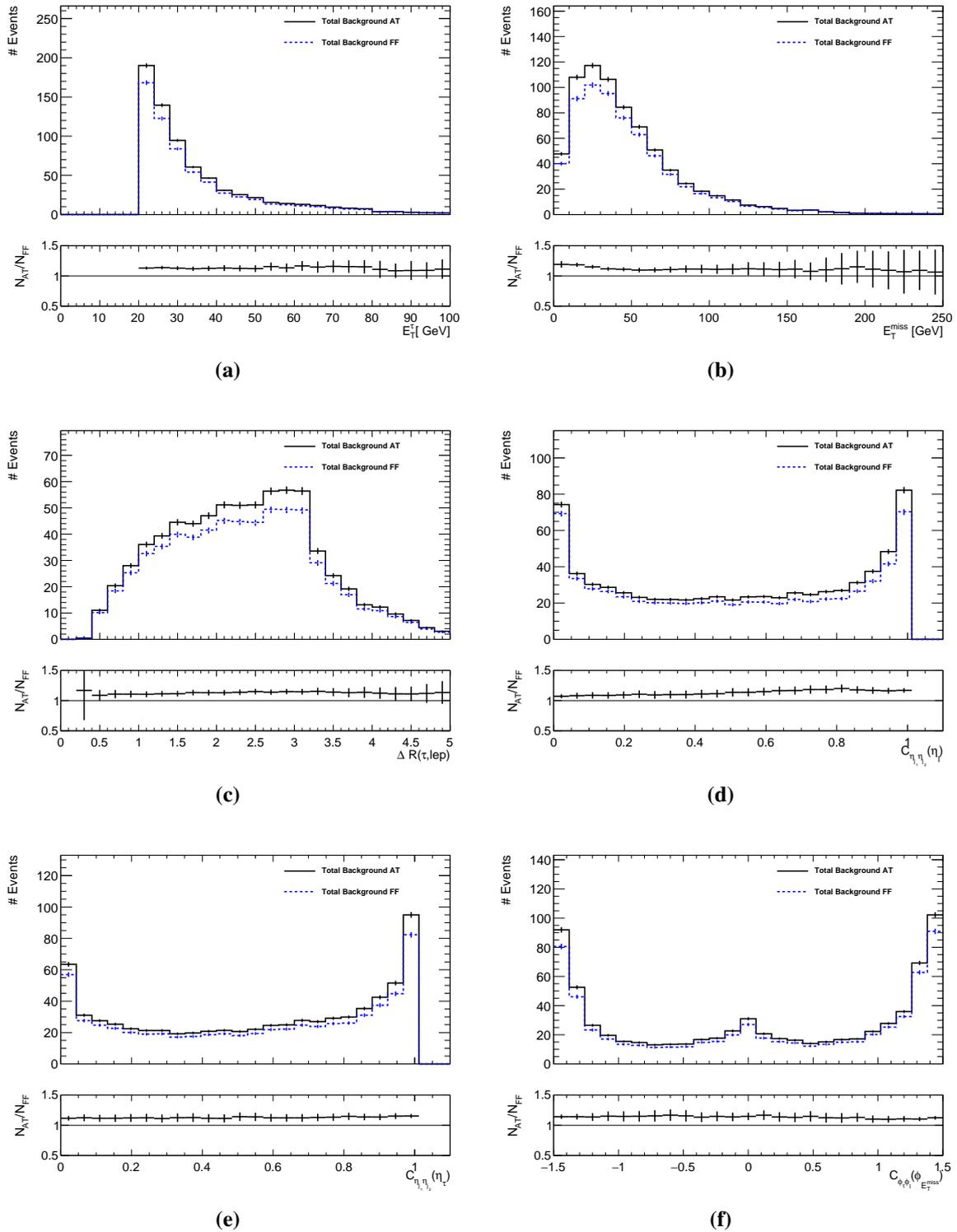


Figure 6.5.: Background shapes in the VBF category for full fake-factor method compared to using AT events only for the fake- τ background contribution, for all variables being fed into BDT_{VBF} . Uncertainties are statistical only.

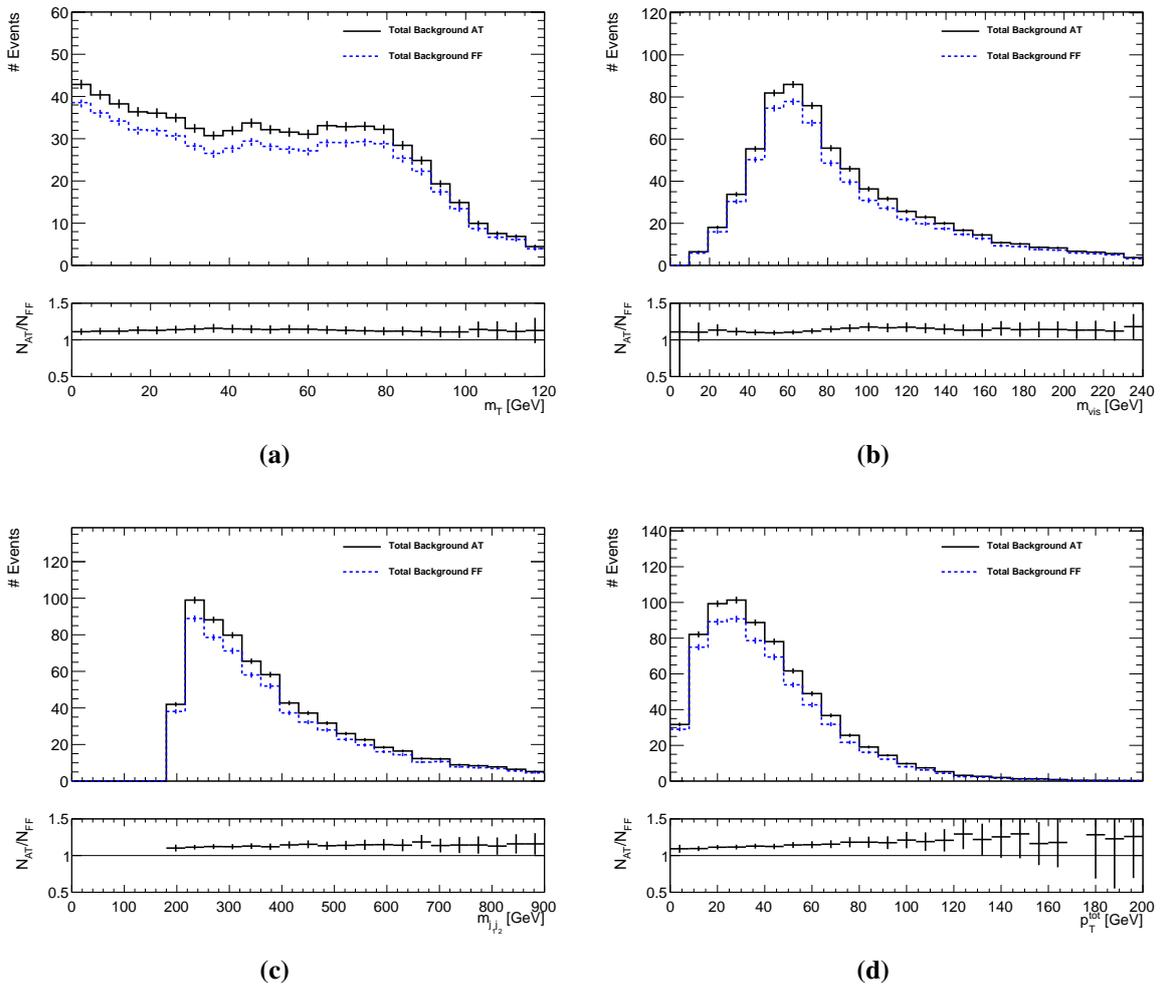


Figure 6.6.: Background shapes in the VBF category for full fake-factor method compared to using AT events only for the fake- τ background contribution, for all variables being fed into BDT_{VBF} . Uncertainties are statistical only.

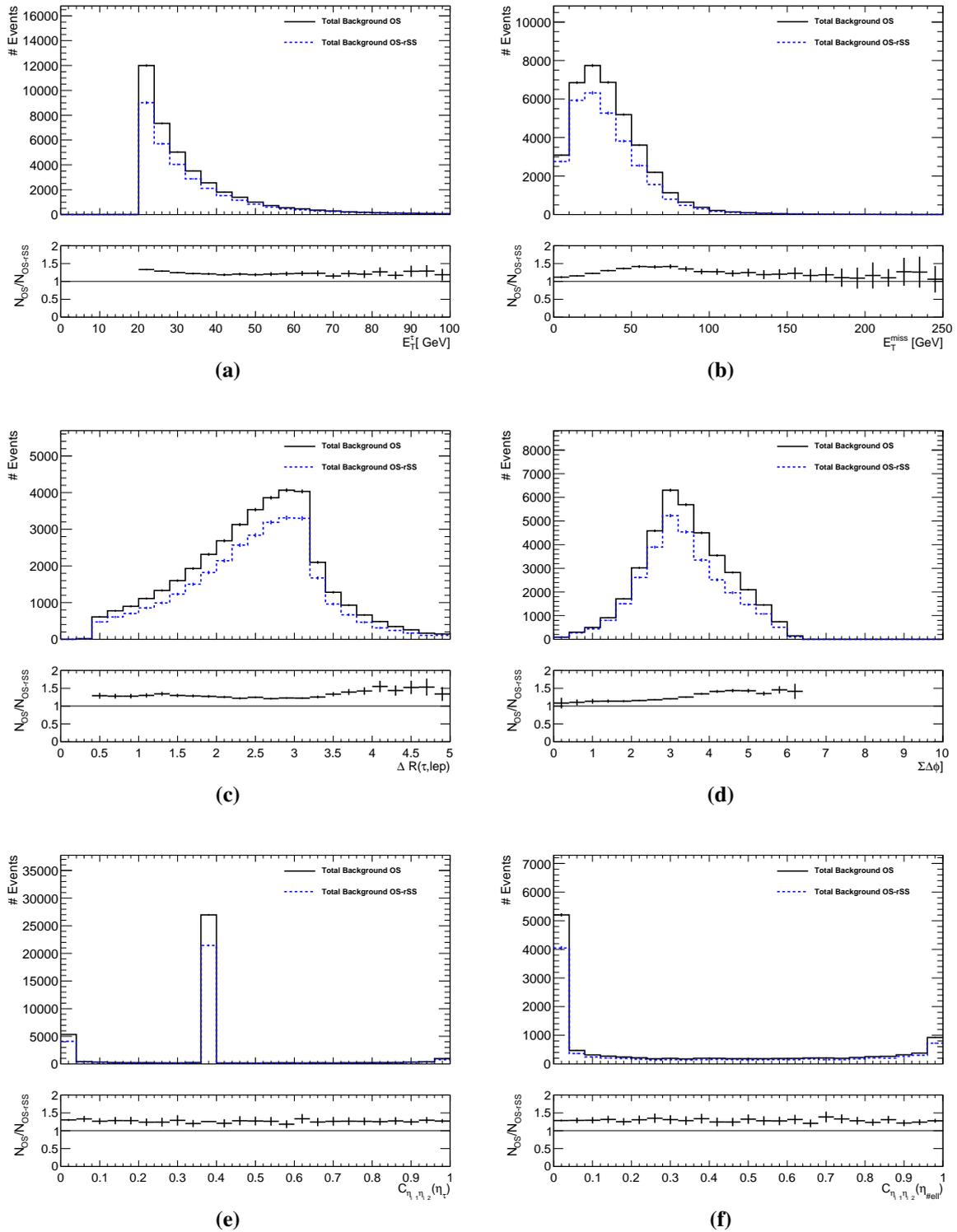


Figure 6.7.: Background shapes in the 1-Jet category for full OS-rSS method compared to using OS events only for the fake- τ background contribution, for variables being fed into BDT_{1J} . Uncertainties are statistical only.

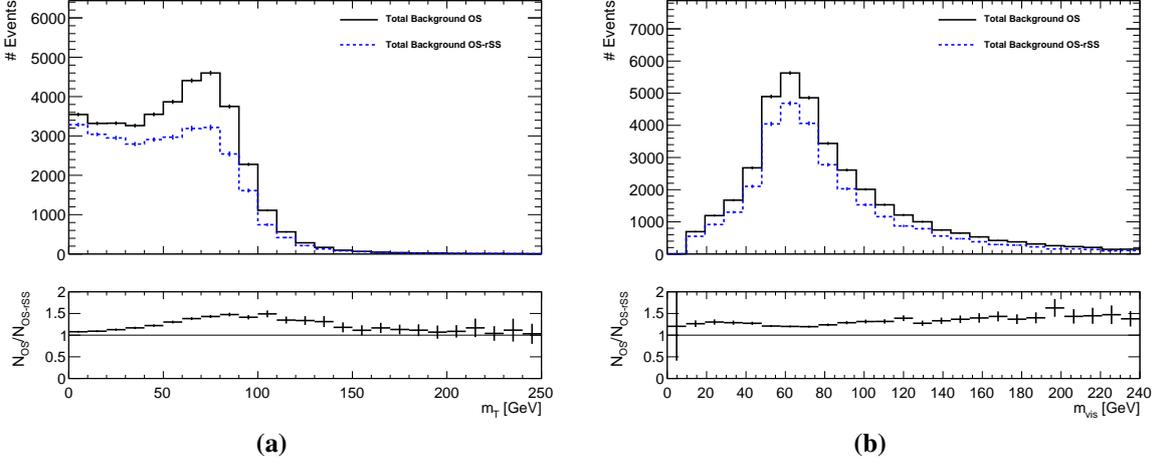


Figure 6.8.: Background shapes in the 1-Jet category for full OS-rSS method compared to using OS events only for the fake- τ background contribution, for variables being fed into BDT_{1J} . Uncertainties are statistical only.

The second largest background component is the fake- τ events. Since the VBF and Boosted categories use the fake-factor method to estimate these contributions, the estimate from OS-rSS acts as an independent sample. As these data sets are orthogonal, there is no double counting, and both estimation methods give matching yields within uncertainties. For these reasons, the OS-rSS method was used for fake- τ background estimation in the training samples, with the fake-factor method used for the testing samples.

All other background data sets are split approximately in half for use in testing and training. All event weights are modified such that the final event yields match for the training and testing samples.

For the signal events, a different approach was taken compared to the Hybrid analysis. As it is known that there exists a Higgs-type particle at around $m_H = 126$ GeV, it was decided to only train the BDTs against the $m_H = 125$ GeV MC sample¹. For the VBF category, only the events produced via VBF were used for training and testing. In the case of the boosted Higgs scenario, the responsible production mode may be VBF, ggF , or VH , and so all MC samples were used for the Boosted category. Finally, only the ggF sample was used for the 1- and 0-Jet categories as it is the only production mode expected to make a significant contribution. Categorisation cuts in Table 6.1 were placed before training.

¹Note that the MC was generated *before* the discovery in 2012, and so the $m_H = 125$ GeV mass point used is the one closest to the observed mass.

6.3.2. Training Parameters

As discussed in Section 5.1, there are many parameters that can be changed in order to optimise the discriminating power of a BDT. In this study, 5 parameters were optimised; the learning rate, the number of trees in the forest, the maximum depth of each tree, the granularity of the phase space in which cuts can be made, and the final allowed size of each leaf.

As we are yet to optimise a list of input variables, a full list of potential variables were used to choose the training parameters. Table 6.4 shows the full list of potential variables used, with those selected for the final BDT. These variables are explained in the next section.

The method used to optimise the training parameters was to take the integral of the ROC curve of the calculated BDT for the training sample. A higher ROC integral points towards a greater separation between the signal and background samples.

Boost Type

Starting with the default TMVA parameters, the first choice made was to move to a Gradient Boost as apposed to AdaBoost. Figures 6.9a and 6.9b show the same events and training parameters for the AdaBoost and Gradient Boost algorithms. Although both methods currently show significant overtraining (evident from the difference in shapes between the training and testing samples), this will disappear after further optimisation. The choice for the move was two fold; firstly, once optimised, the discriminating power of the gradient boost is much higher, and secondly, the shape of the gradient boost is more predictable. While the AdaBoost algorithm produces two Gaussian-type curves for the signal and background shapes, the Gradient Boost algorithm will always produce a multi-modal distribution with a high peak towards ± 1 for signal and background respectively. These distributions also tend to completely populate the score range $[-1, 1]$ for all major background groups, helping with potential *0-bin errors* in the fitting process.

Learning Rate

The learning rate (or shrinkage for the Gradient Boost algorithm) gives a measure of how powerful the boosts are. A more predictable performance can be achieved with a slower learning rate. Figure 6.9c shows the shrinkage as a function of the ROC integral. A low shrinkage value of 0.08 is chosen for the VBF category.

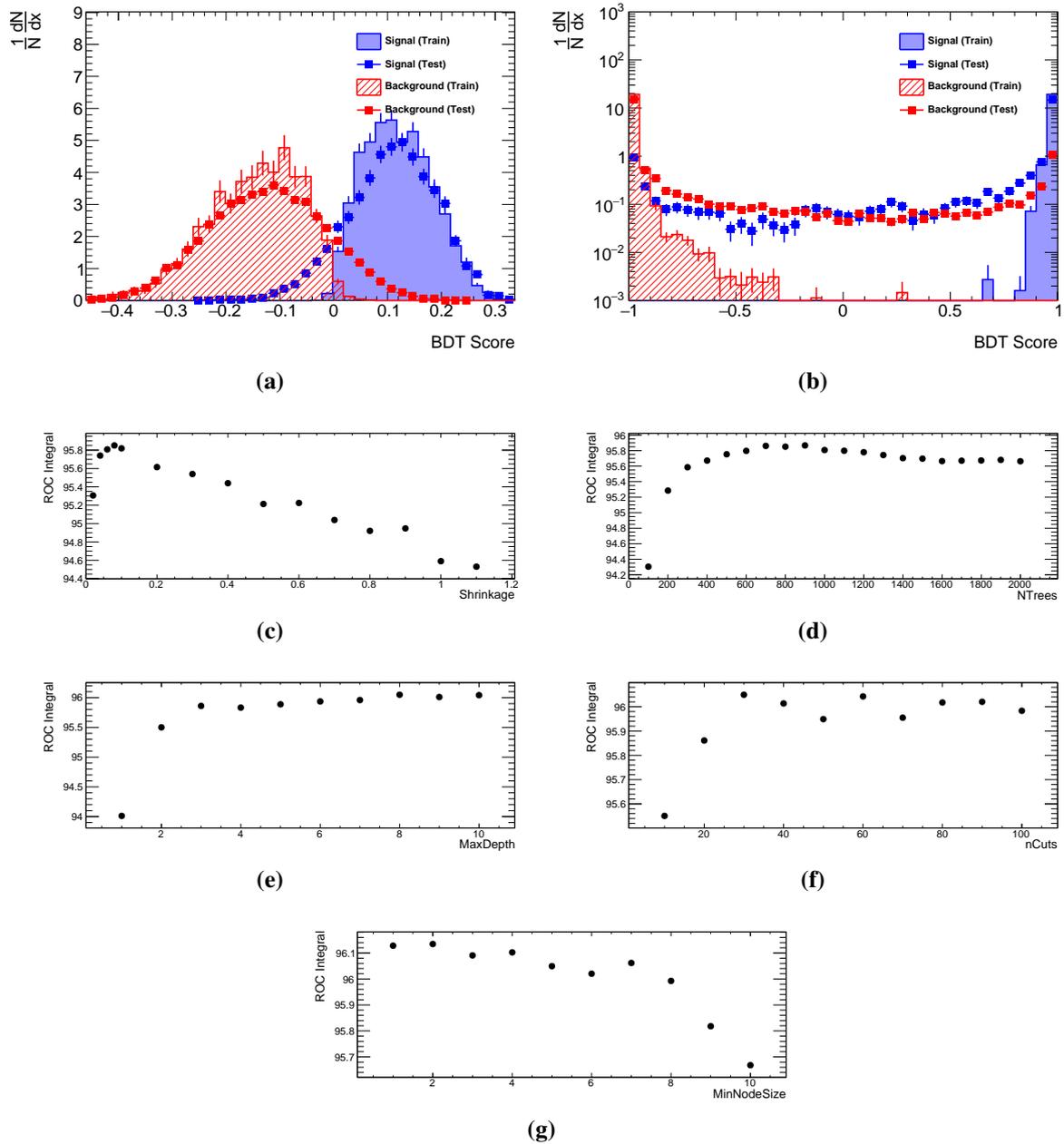


Figure 6.9.: (a) and (b) show the TMVA training and testing output for BDT_{VBF} for the AdaBoost and Gradient Boost algorithms, before training parameter optimisation. The remaining figures show the scans for the training parameters in Section 6.3.2.

Number of Trees

The number of trees (`NTrees`) refers to maximum number of boost steps allowed by the algorithm. With a small learning rate it is expected that a large number of trees will be needed to provide the best performance. Figure 6.9d shows the number of trees as a function of the ROC integral. As too many trees can result in overtraining, the first stable point was chosen with a value of 700.

Maximum Tree Depth

The maximum tree depth (`MaxDepth`) determines how many levels of node splittings can be applied in each tree. By letting the tree grow too deep, specific events can end up being selected, resulting in overtraining. Limiting tree depth therefore provides a more robust BDT. Given the trend in Figure 6.9e, a maximum tree depth of 3 was selected.

Number of Cuts

The number of cuts (`nCuts`) can be chosen to allow a finer granularity in the available phase space for the cut decision at each node. Figure 6.9f shows that at least 30 grid points are needed to find the optimal splitting point, and not much gain in discrimination is achieved for higher values.

Minimum Node Size

The minimum node size (`MinNodeSize`) is the percentage of training events required in a leaf node. Too few training events in each leaf can again result in overtraining. A value of 2% was chosen based off the result from Figure 6.9g.

6.3.3. Optimising BDT Variable List

Now that the BDT training parameters are selected, the variable list being fed into the BDT also needs to be optimised. It is important to not just throw all possible variables into the BDT and expect robust performance. By including too many variables, the available phase space becomes too large and it is again much easier for the BDT to become overtrained. Using too

Training Parameter	BDT _{HBD}	BDT _{VBF}	BDT _{BST}	BDT _{IJ}	BDT _{OJ}
BoostType	AdaBoost		Gradient Boost		
Learning Rate	0.5	0.08	0.4	0.8	0.9
NTrees	800	700	700	600	500
MaxDepth	3	3	4	3	3
nCuts	20	30	80	70	30
MinNodeSize	5	2	4	3	3

Table 6.3.: TMVA training parameters for BDTs grown for the hybrid VBF category, and the full MVA analysis categories. Parameters are described in Section 6.3.2.

many variables can have the same effect as growing the trees too deep – specific events can end up being selected for, which when applied to a new set of events (such as the ATLAS data) can give very different results.

Variable Selection

Table 6.4 shows the list of variables before optimisation. The table is split into five sections based on the type of variable; object kinematics, the di-tau system, mass, the di-jet system, and total momentum variables. These groups of variables allow the BDT to find areas of phase space that are specific to VBF-type events, boosted events, and to discriminate between Higgs signals and $Z \rightarrow \tau\tau$ backgrounds.

To aid discrimination between true tau decays and fake- τ events, the E_T^τ and \cancel{E}_T variables, and derived quantities, were included. The di-tau system variables include the p_T -ratio of the tau to the lepton (E_T^τ/p_T^ℓ), the (η, ϕ) separation of the tau and lepton ($\Delta R(\tau, \ell)$), the sum of angular separations of the visible tau decay products and the \cancel{E}_T ($\sum \Delta\phi$), the centrality variables (C), and the energy fractions ($x_{\tau/\ell}$). These derived variables were defined to exploit correlations that appear in the di-tau system, but not necessarily in fake- τ backgrounds. Three centrality variables are tested, the tau and lepton η centrality as defined in Section 5.2.2, as well as the \cancel{E}_T ϕ centrality, defined generally as

$$\begin{aligned}
 \Phi_1(\phi) &= \frac{\sin(\phi - \phi_1)}{\sin(\phi_2 - \phi_1)} \\
 \Phi_2(\phi) &= \frac{\sin(\phi - \phi_2)}{\sin(\phi_2 - \phi_1)} \\
 C_{\phi_1\phi_2}(\phi) &= \frac{\Phi_1 + \Phi_2}{\sqrt{\Phi_1^2 + \Phi_2^2}}
 \end{aligned} \tag{6.1}$$

Similar to the η centralities, this variable tests whether ϕ lies between ϕ_1 and ϕ_2 . In this case, the missing energy is expected to lie between the tau and lepton for signal events.

The mass variables are important for separating signal events from $Z \rightarrow \tau\tau$ events as the event properties are extremely similar, except for a shift in the mass peak of the mass of the Higgs compared to the Z decay products. Only the transverse mass and the mass of the visible decay products (m_{vis}) are tested however. The decision to omit m_{MMC} from the optimisation process was made as it will be the final distribution used to extract an exclusion limit. Including m_{MMC} in the BDT therefore may add biases into the method. The di-jet system variables are expected to play a role in providing discriminating power for the VBF category, selecting for events with two very forward jets. Finally, the vector sum of the total transverse momenta of the event (p_T^{tot}) is included, again to pick up events with forward jets.

Variable Optimisation

To optimise the variable list, a simple trimming procedure was employed. A BDT was grown for a list of variables. New BDTs were then grown after removing one variable from the list. The variable that, when removed had the smallest decrease in the ROC integral was permanently removed from the list. This iterative procedure continued until the ROC integral decreased by at most 1% compared to the full variable list.

The final variable list after running the optimisation procedure is shown in Table 6.4. It is important to ensure that no unnecessary correlations exist in the BDT variables. Figures 6.12a and 6.12b show the linear correlations between the final selected variables for signal and background processes, for the VBF category. Notice that, although some correlations exist (and are expected for the decay), the strongest correlation in the signal sample does not exist in the background, and vice versa. This allows the BDT to exploit these differences to improve the discrimination. Figures 6.10 and 6.11 show the input variables for BDT_{VBF} , as an example

	BDT _{HBD}	BDT _{VBF}	BDT _{BST}	BDT _{1J}	BDT _{0J}
E_T^τ	•	•		•	
\cancel{E}_T	•	•	•	•	•
$p_T^{j_1}$	•				
$p_T^{j_2}$	•				
E_T^τ/p_T^ℓ			•		•
$\Delta R(\tau, \ell)$		•	•	•	•
$\Sigma \Delta \phi$				•	•
$C_{\eta_{j_1} \eta_{j_2}}(\eta_\ell)$	•	•		•	
$C_{\eta_{j_1} \eta_{j_2}}(\eta_\tau)$	•	•		•	
$C_{\phi_\tau \phi_\ell}(\phi_{\cancel{E}_T})$		•	•		
m_T		•	•	•	
m_{vis}		•	•	•	•
$m(j_1, j_2)$	•	•			
$\Delta \eta(j_1, j_2)$	•				
$\eta_{j_1} \times \eta_{j_2}$	•				
p_T^{tot}	•	•			

Table 6.4.: Input variables for the Hybrid BDT, and the BDTs used for the full-MVA analysis.

of the validity of the background model. It can be seen, as it was for all analysis categories, that all variables are well modelled.

The final BDT output from TMVA for the VBF category, and the associated ROC curve, are shown in Figures 6.12c and 6.12d. The difference between background shapes for the training and testing samples can be attributed to the different background estimation methods used in each case.

BDT Validation

The BDTs for each analysis category are shown in Figure 6.13. As a BDT score $\rightarrow 1$ is more signal-like, the data has been blinded above the score where the signal efficiency reaches 30%. This *low-BDT* region is well modelled for all four analysis categories. It is interesting to note the shape of the various BDT scores. For the VBF category, the score peaks strongly at ± 1 , however this separation decreases as the categories progress, until the 0-Jet category where the distribution actually has a local peak near a score of 0. This points to the greater sensitivity of the VBF and Boosted categories, compared to the n -Jet categories.

To check that the BDT is well modelled in the *high-BDT* region (towards 1), while still staying blind to the signal, the variable from the BDT that TMVA deemed most *important* (a variables' importance is defined by how often it is used at a node) was cut on to reduce the signal efficiency to 30%, while retaining as much background as possible. The variables chosen were $C_{\eta_{j_1} \eta_{j_2}}(\eta_\ell)$ for VBF, $\Delta R(\tau, \ell)$ for Boosted, and $\Sigma \Delta \phi$ for 1- and 0-Jet. Figure 6.14 shows these high-BDT distributions. Again, a good data-MC agreement is seen for all categories.

The BDT scores in the control regions are also shown in Figure 6.15 ($t\bar{t}$ CR), Figure 6.16 ($W(\rightarrow \ell \nu) + \text{jets}$ CR), and Figure 6.17 ($Z \rightarrow \tau\tau$ CR). All regions show agreement between data and MC across the whole BDT score, within the background estimation uncertainties.

6.3.4. Cutting on the BDTs

As in the hybrid analysis, the BDT cut values are determined by finding the score that minimises the exclusion limit for the category. The limit setting procedure here is the same as in the hybrid analysis, with the m_{MMC} distribution being re-binned as per Algorithm 4.1, and then fitted via the methods outlined in Section 4.9, after placing a cut on the category BDT score.

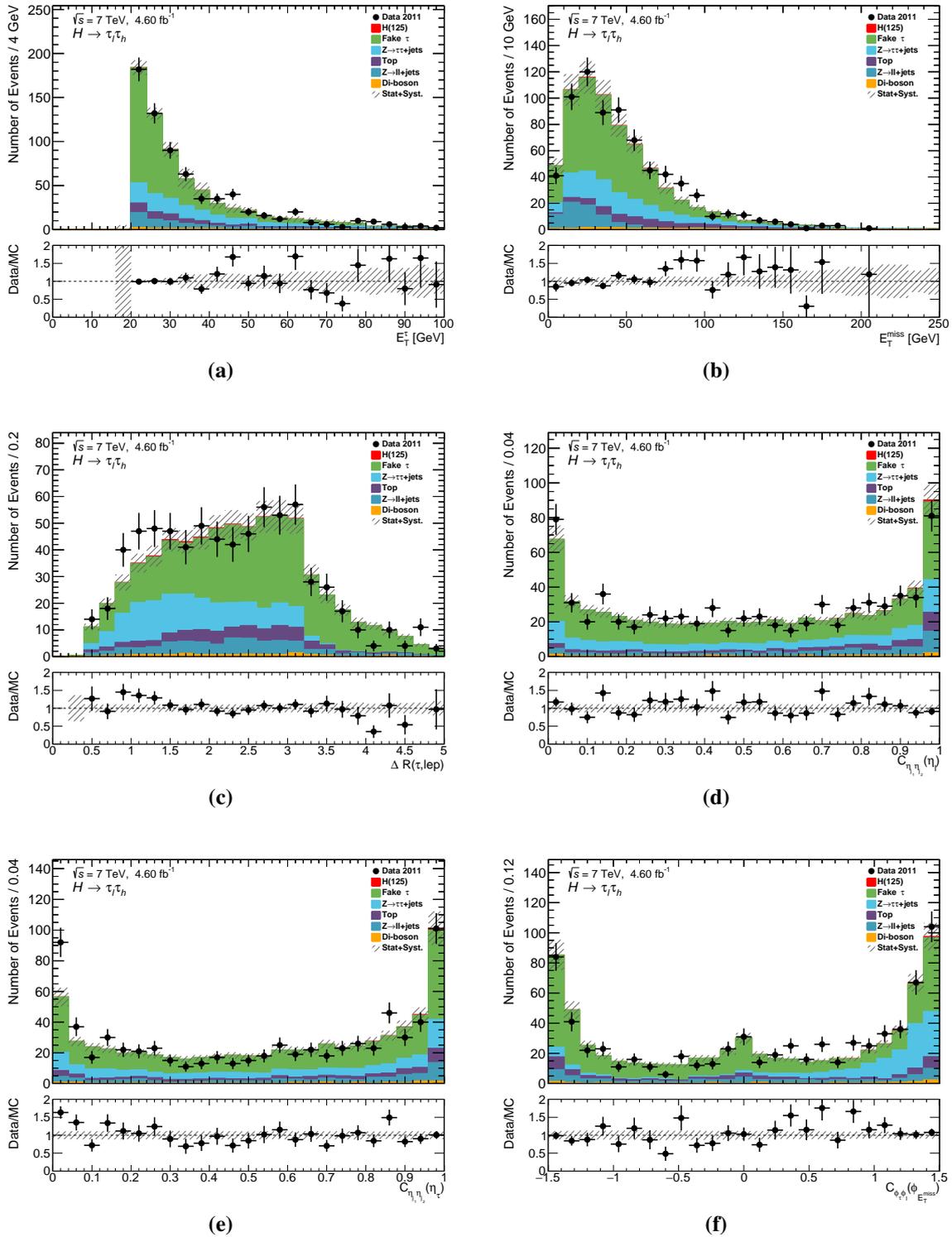


Figure 6.10.: Input variables for BDT_{VBF} after the base VBF_{MVA} selection. Uncertainties are statistical and systematic.

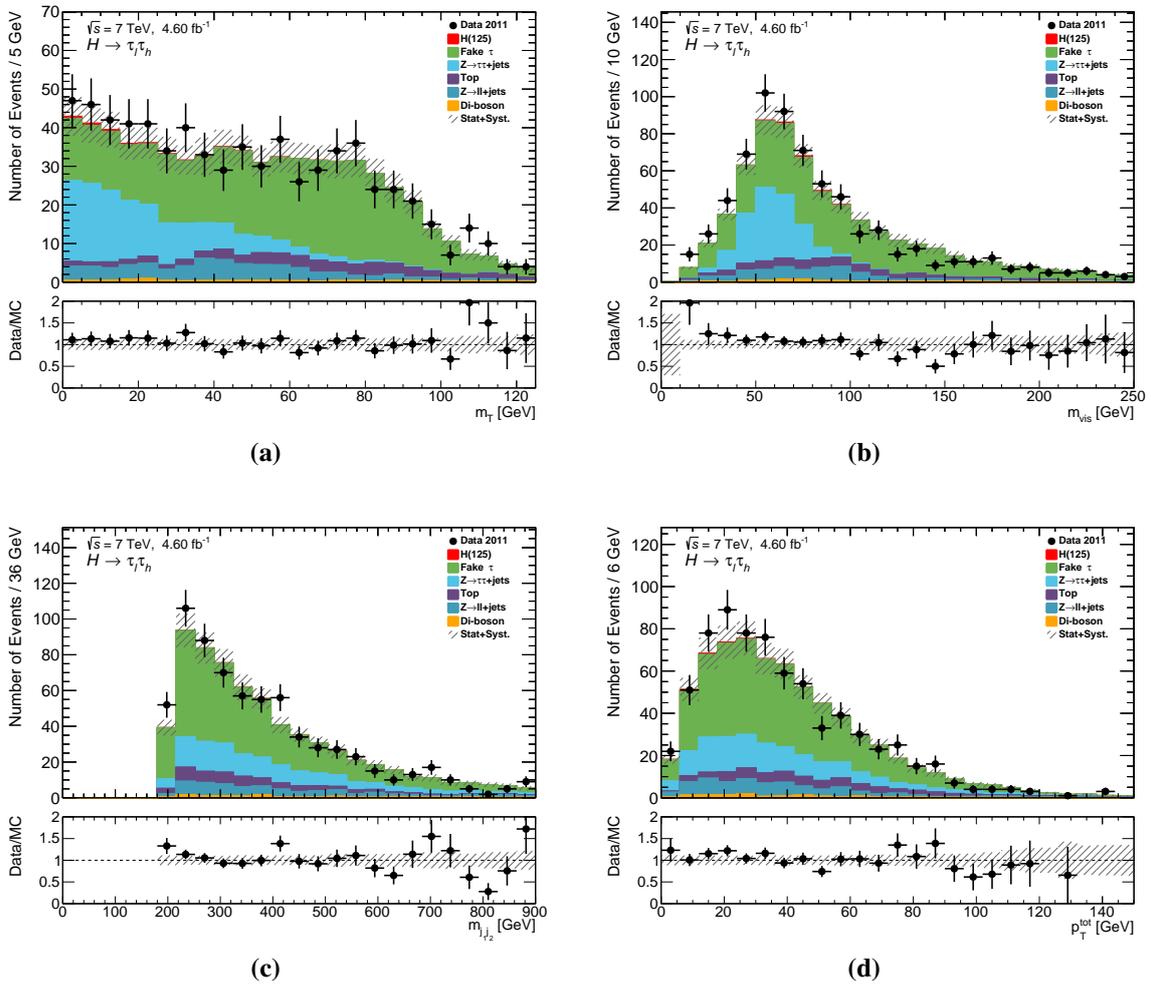


Figure 6.11.: Input variables for BDT_{VBF} after the base VBF_{MVA} selection. Uncertainties are statistical and systematic.

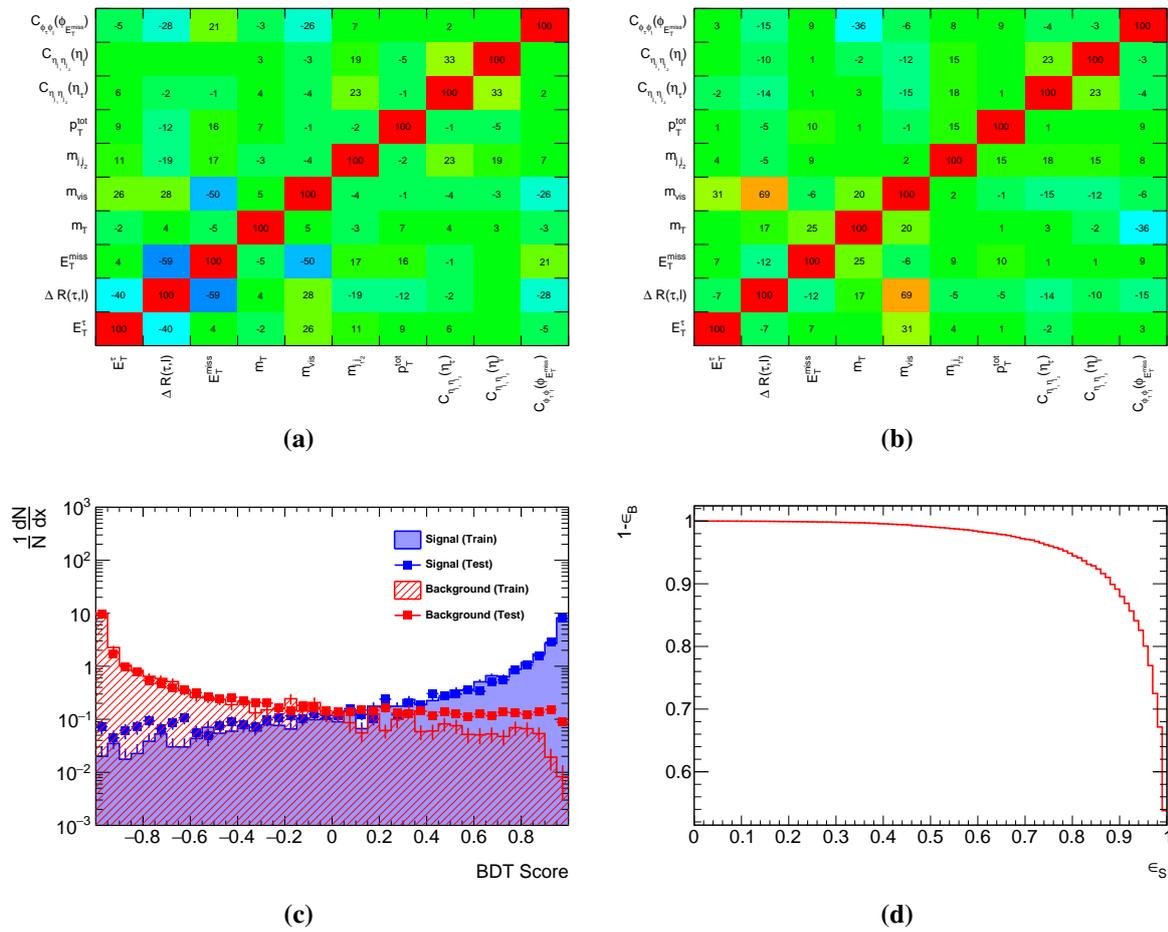


Figure 6.12.: Linear correlation matrices for (a) signal and (b) background events after training for BDT_{VBF}. Bins show level of correlation as a percentage. The variables compared are described in Sections 4.6 and 6.3.3. (c) and (d) show the BDT Score and ROC curves as outputted from TMVA for BDT_{VBF}.

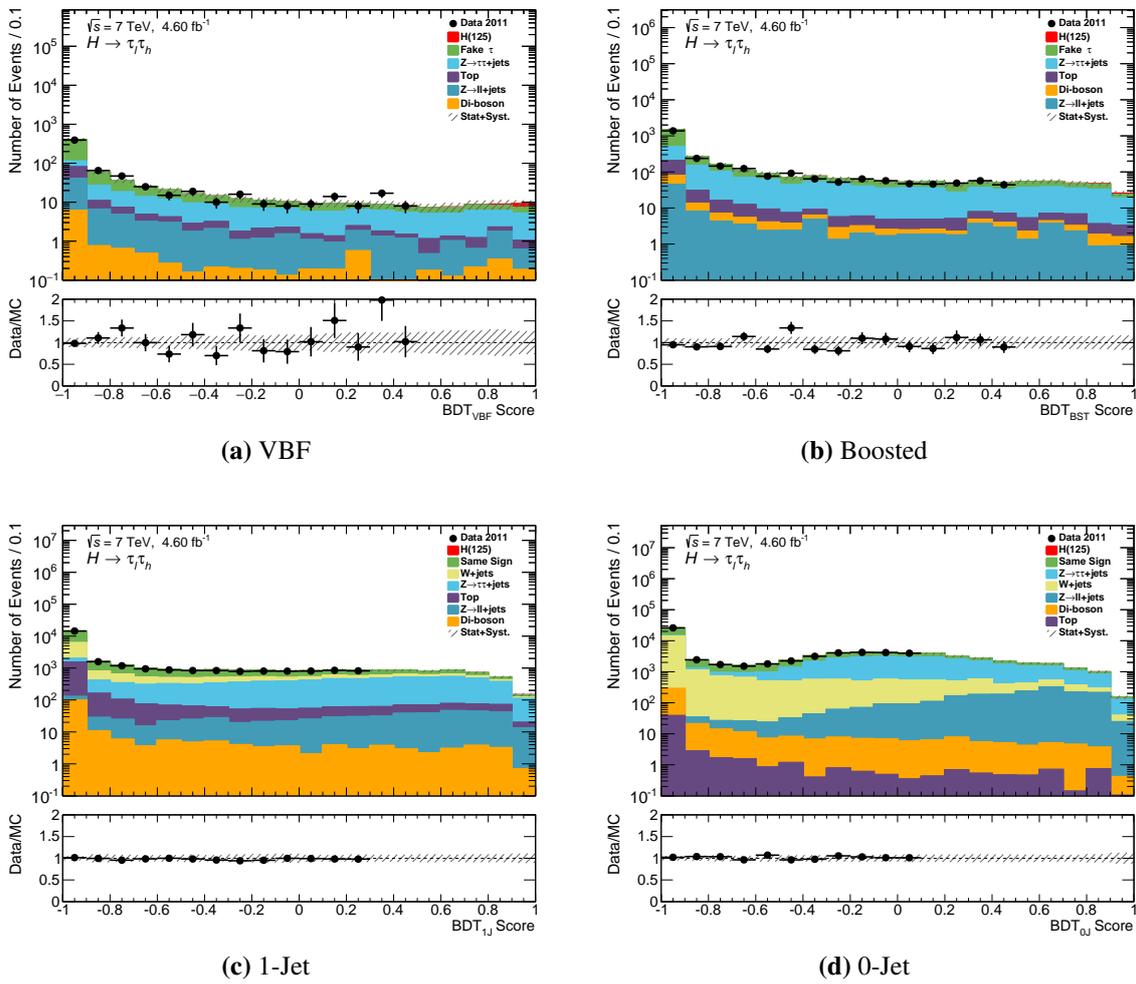


Figure 6.13.: BDT distributions for the MVA analysis in the low-BDT control region. Data is blinded for BDT scores above a signal efficiency of 30%. Uncertainties are statistical and systematic.

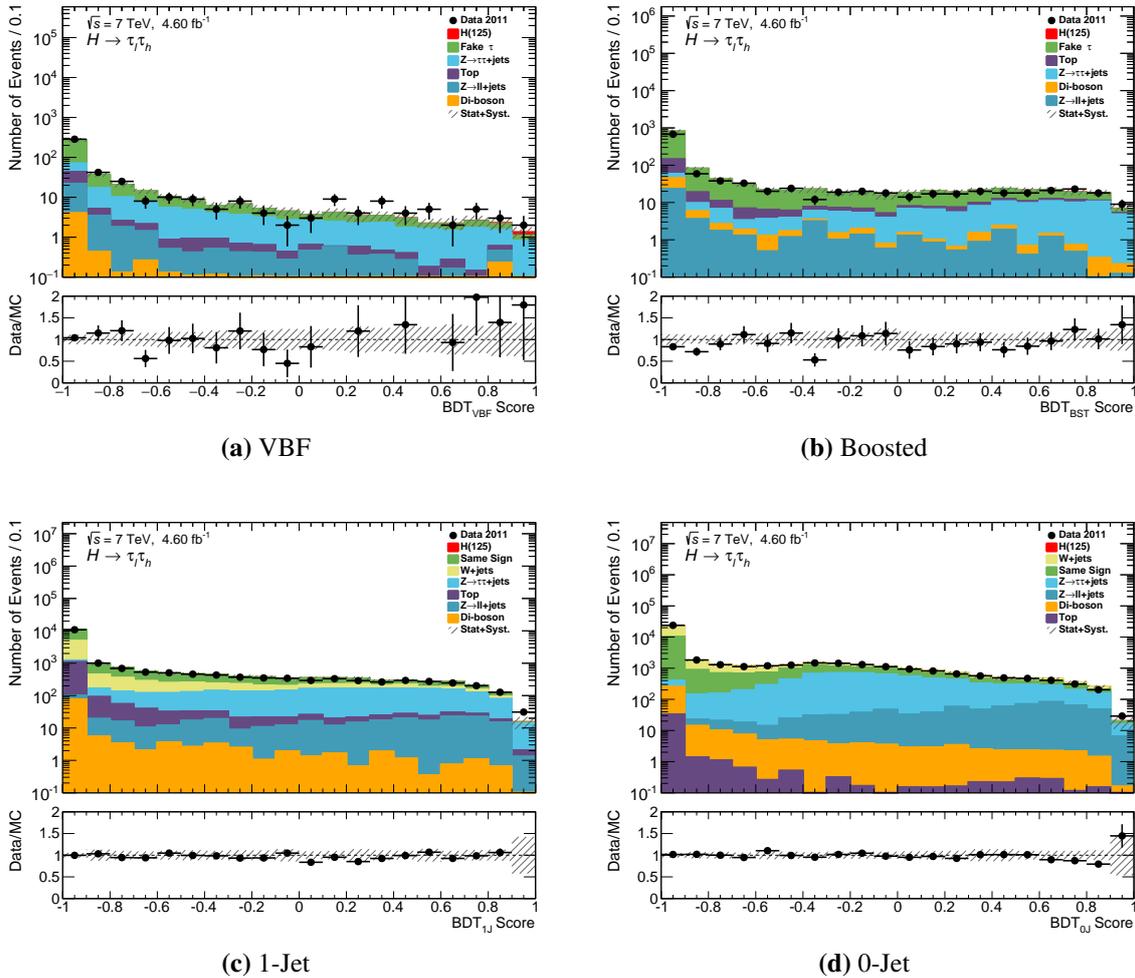


Figure 6.14.: BDT distributions for the MVA analysis in the high-BDT control region. Cuts are placed on the most important variable for each BDT from TMVA such that the signal efficiency is below 30%; $C_{\eta_{j_1} \eta_{j_2}}(\eta_\ell) < 0.7$ for VBF, $\Delta R(\tau, \ell) > 1.8$ for Boosted, $\Sigma \Delta\phi > 3.0$ for 1-Jet, and $\Sigma \Delta\phi > 3.2$ for 0-Jet. Uncertainties are statistical and systematic.

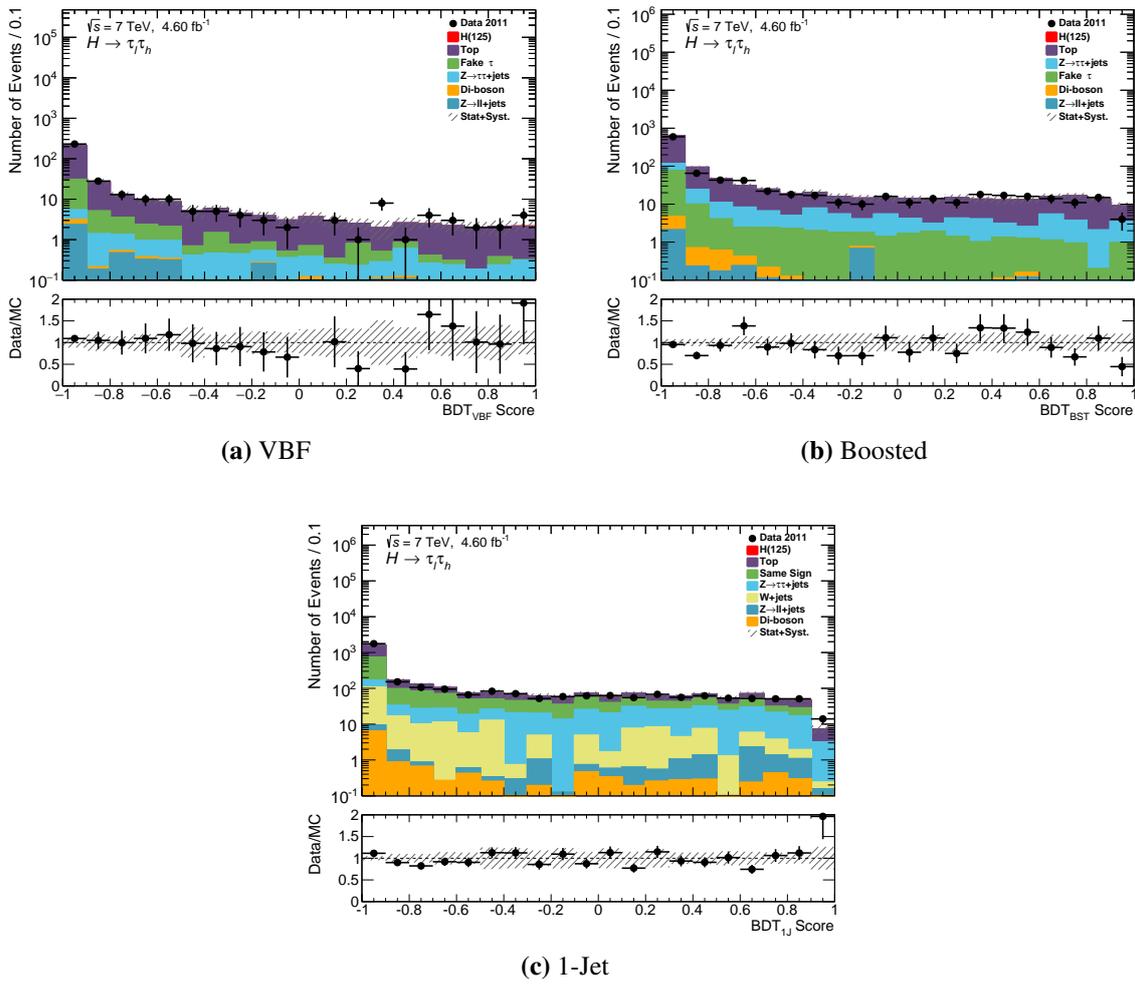


Figure 6.15.: BDT distributions for the MVA analysis in the $t\bar{t}$ control region. Uncertainties are statistical and systematic. Note that due to the requirement of no jets in the 0-Jet category, there is no $t\bar{t}$ control region available.

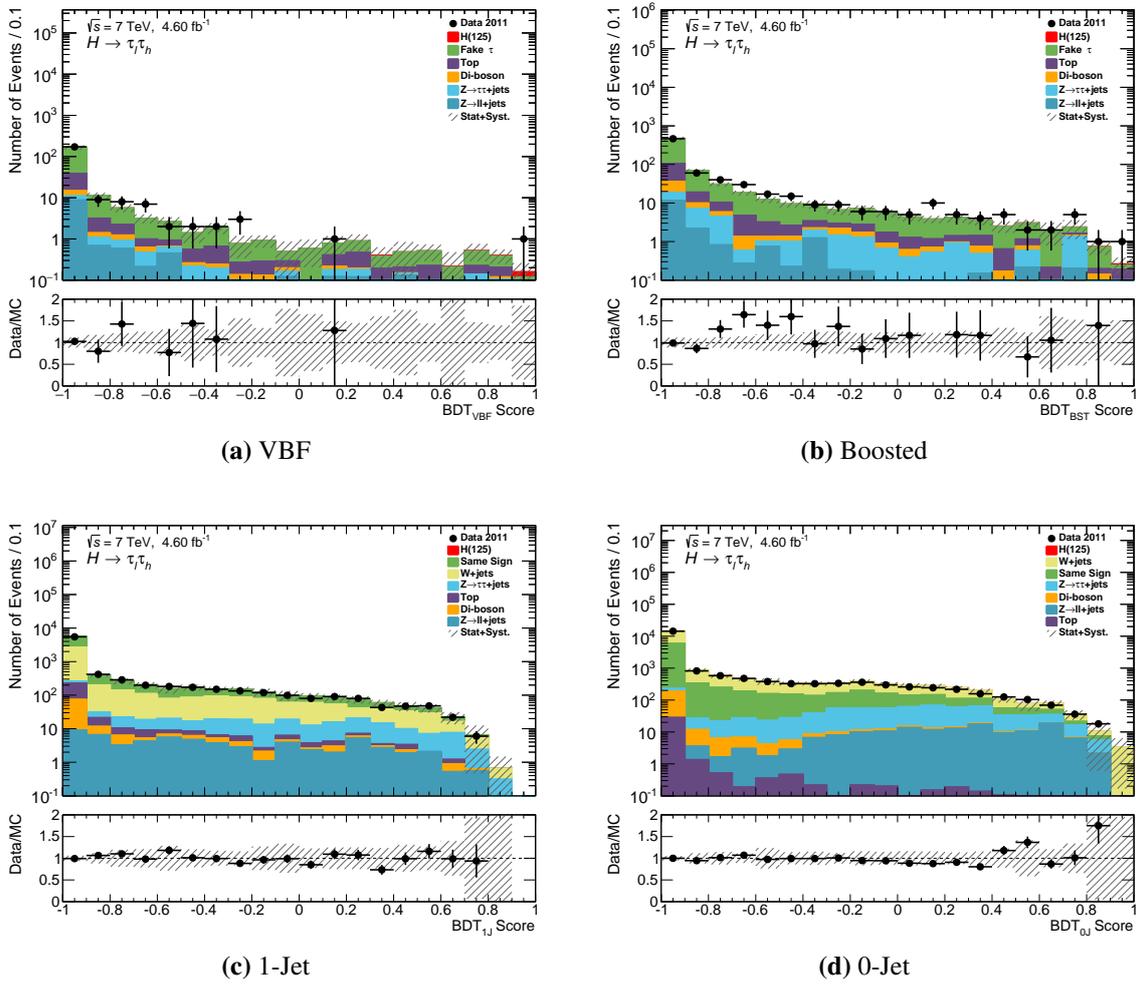


Figure 6.16.: BDT distributions for the MVA analysis in the $W(\rightarrow \ell\nu) + \text{jets}$ control region. Uncertainties are statistical and systematic.

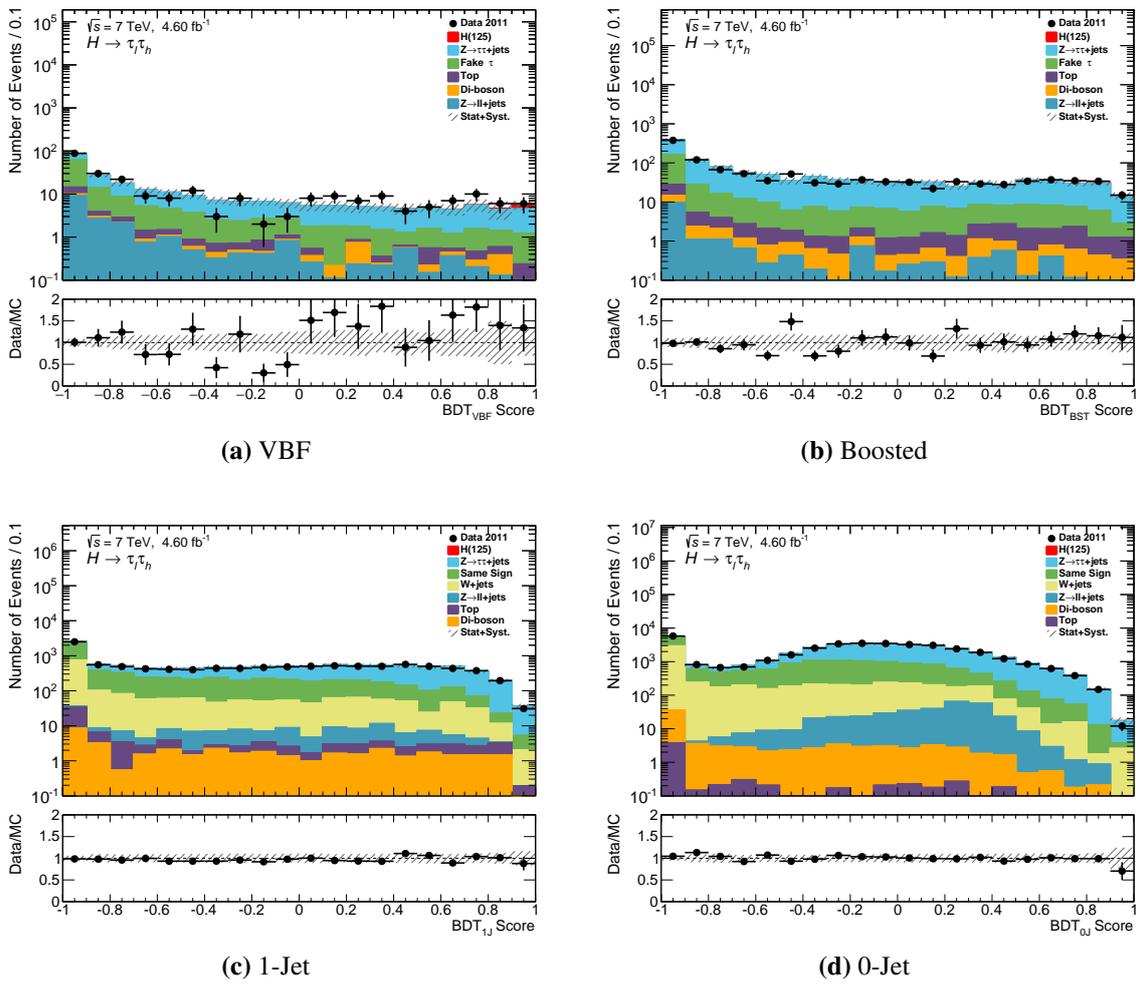


Figure 6.17.: BDT distributions for the MVA analysis in the $Z \rightarrow \tau\tau$ control region. Uncertainties are statistical and systematic.

Figure 6.18 shows the expected 95% confidence level exclusion limits on σ/σ_H^{SM} as a function of BDT score at $m_H = 125$ GeV, for the four analysis categories. In the final categorisation, events that fail the BDT cut will flow into the subsequent category. Due to this, each exclusion plot was only calculated once the BDT cut from the proceeding category was determined, and applied.

The sharp rises in the exclusion limits at the high score values are due to lack of available statistics left in the signal region. In this extreme end of the BDT score, the problem of backgrounds having zero or negative yields returns, which no amount of rebinning the m_{MMC} distribution can alleviate. The VBF category has a minimum exclusion limit at $BDT_{VBF} > 0.96$ however, again, with such a harsh cut the VBF signal region is too sparsely populated for the fit to converge. In an effort to retain as many events as possible in the signal region, cuts were placed at the lowest BDT score that still gave a comparable exclusion and the smallest error bands compared with the absolute minimum. The final cuts used to define the MVA categories are therefore $BDT_{VBF} > 0.70$, $BDT_{BST} > 0.70$, $BDT_{IJ} > 0.75$, and $BDT_{0J} > 0.75$.

6.3.5. BDT Interpretation

As in Chapter 5, we can again look at the correlations between the BDT input variables, and the classifier score. For the example of the VBF category, Figure 6.19 shows the tau system variables and Figure 6.20 shows the jet variables and tau system mass variables. Looking at the BDT_{VBF} column it is immediately obvious how discriminating BDT_{VBF} is. Projecting the signal and backgrounds onto the y-axis for these figures shows how most of the input variables provide far less separation between signal and background compared to BDT_{VBF} . The output from TMVA ranks the variables in terms of their importance and separation, defined as

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_s(y) - \hat{y}_B(y))^2}{\hat{y}_s(y) + \hat{y}_B(y)} dy \quad (6.2)$$

where \hat{y}_s and \hat{y}_B are the signal and background PDFs of y , $\langle S^2 \rangle = 0$ for fully overlapping sample shapes, and $\langle S^2 \rangle = 1$ for completely separated samples. The plots for the variables for which have the greatest separation – $C_{\phi_\tau \phi_\ell(\phi_{\cancel{E}_T})}$ and $m_{j_1 j_2}$ – also show significant separation along the BDT_{VBF} axis. More interestingly though is that the variable with the *worst* separation, \cancel{E}_T , still shows a strong separation of signal and background along the BDT_{VBF} axis. This indicates

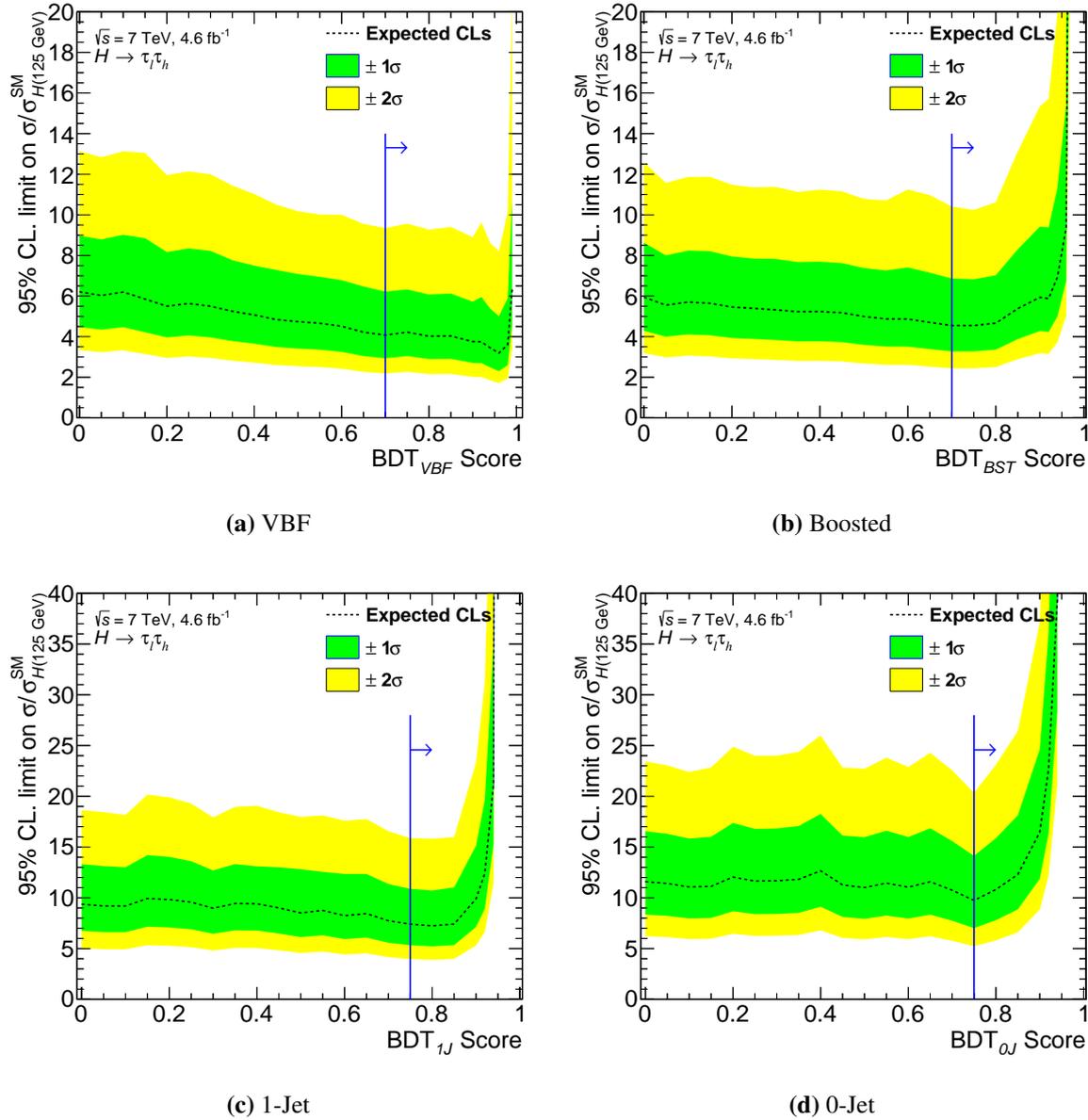


Figure 6.18.: Exclusion as a function of BDT score for the four analysis categories, with $m_H = 125$ GeV. Blue line indicates final cut values chosen. Errors include statistical and systematic uncertainties.

that while placing a one-dimensional cut on \cancel{E}_T may not provide any gains, the BDT can find higher dimensional relationships between the variables, and exploit them.

6.4. Results

The final event yields for the MVA analysis categories are shown in Tables 6.5 and 6.6. The VBF_{MVA} category has a total background yield comparable to the hybrid analysis, however with a $Z \rightarrow \tau\tau$ contribution greater than the fake- τ s. The signal yield has increased in the MVA analysis by $\approx 30\%$. The Boosted category has changed significantly from the cut-based and hybrid analyses after the introduction of BDT_{BST} . While the signal yield has only increased slightly, the background yields have decreased by more than 5 times, with the most significant reduction being in the $Z \rightarrow \tau\tau$ background. Similarly, the 1- and 0-Jet categories have seen dramatic background reduction, particularly in $Z \rightarrow \tau\tau$, with similar signal yields. Finally, the newly added Rest category has significantly more events than the other categories, with the $Z \rightarrow \tau\tau$ background dominating, and a high ggF signal yield. Roughly half of the events in the Rest category would have fallen into the 1-Jet category, and the other half into the 0-Jet category, for the cut-based and hybrid analyses. The much larger yields seen here is due to a looser signal region selection compared to these analyses.

The data/MC agreement is generally within 1σ of the background uncertainties, except for the VBF category. The rebinned pre-fit m_{MMC} distributions for each analysis category are shown in Figures 6.22 and 6.23. The excess in the VBF category is clearly seen at $m_{\text{MMC}} \approx 150$ GeV. In the expected signal region $100 < m_{\text{MMC}} < 150$ GeV however, the data is consistent with the background-only prediction to within 1σ of the statistical and systematic uncertainties. The impact of this excess will become apparent in the next section when studying the exclusion limit. All other analysis categories show smooth gaussian peaks in the expected signal region, in both data and MC.

As per the cut-based and hybrid analyses, the pull distribution for the combined categories can be used as a method of validating the fit model. The pulls and impacts of each nuisance parameter differ slightly between the MVA analysis and the cut-based and hybrid analyses, however similar trends can be seen. All pulls lie within a $\pm 2\sigma$ deviation of the nominal value. The pull distributions for each individual category, as well as the non-jet combination fit, can be found in Appendix B.

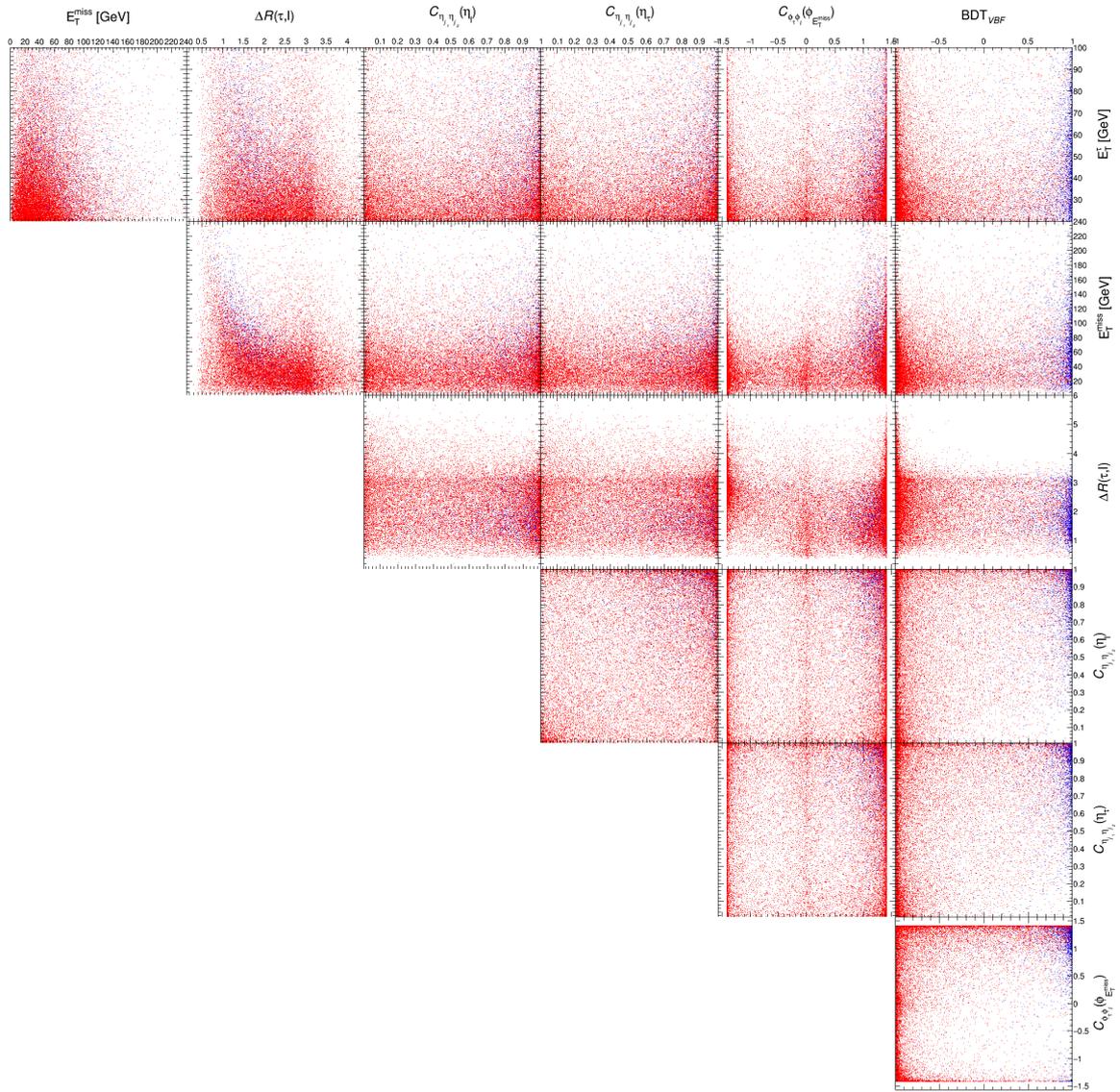


Figure 6.19.: Correlations between the tau system variables, and BDT_{VBF} . Blue data points are VBF Higgs signal at $m_H = 125$ GeV, and red data points are all backgrounds, estimated as per Section 6.2.3.

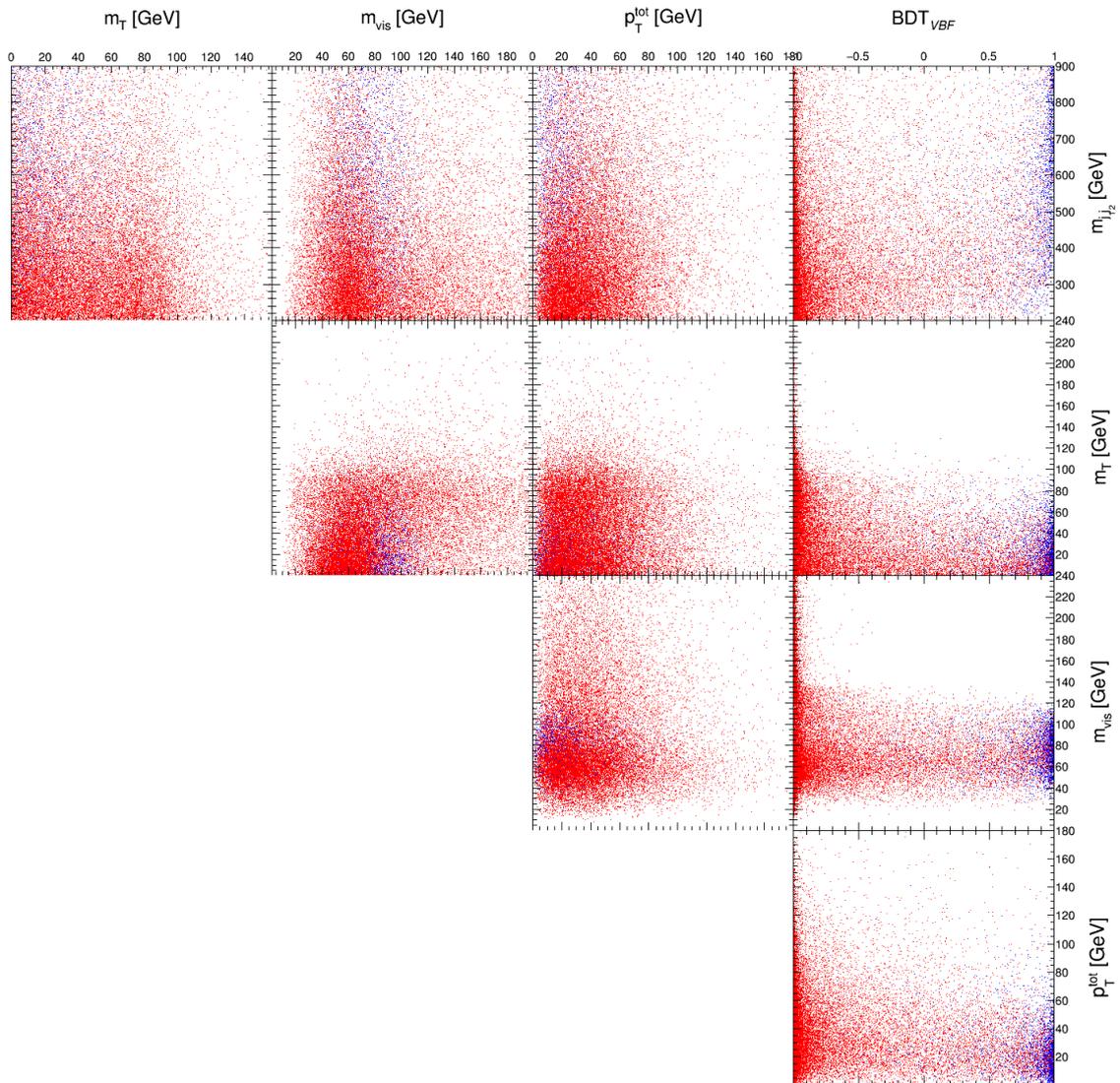


Figure 6.20.: Correlations between the jet variables, tau system mass variables, and BDT_{VBF} . Blue data points are VBF Higgs signal at $m_H = 125$ GeV, and red data points are all backgrounds, estimated as per Section 6.2.3.

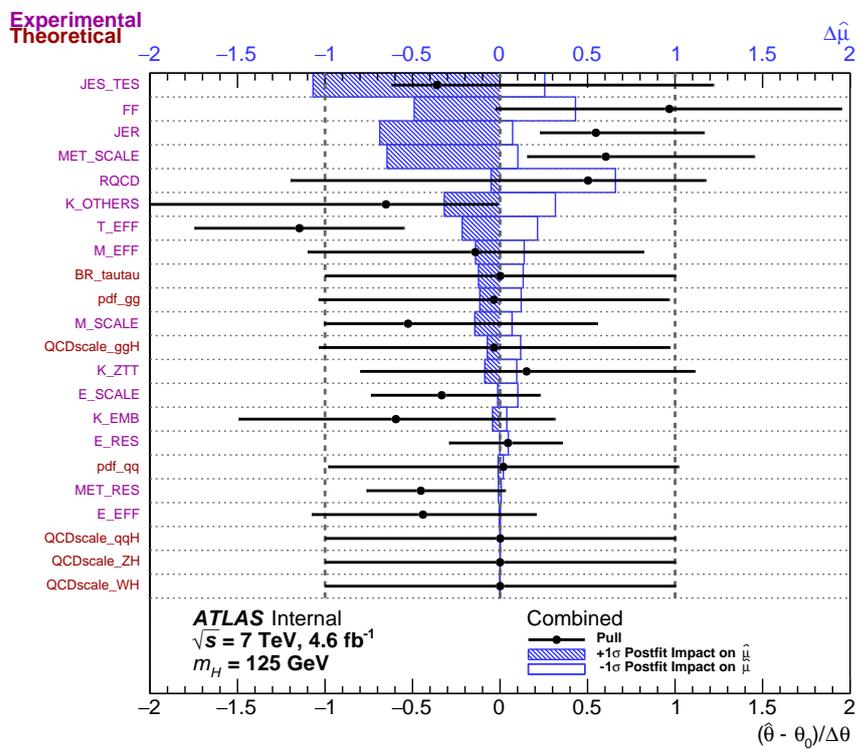


Figure 6.21.: Pull distribution for the combined categories in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

Process	VBF	Boosted	1-Jet (e)	1-Jet (μ)
ggF 125	$0.664 \pm 0.050 \pm^{0.090}_{0.097}$	$4.10 \pm 0.12 \pm^{0.47}_{0.72}$	$6.87 \pm 0.19 \pm^{0.85}_{1.07}$	$5.22 \pm 0.17 \pm^{0.66}_{0.81}$
VBF 125	$2.202 \pm 0.032 \pm^{0.248}_{0.265}$	$1.216 \pm 0.024 \pm^{0.109}_{0.153}$	$1.481 \pm 0.026 \pm^{0.122}_{0.132}$	$1.080 \pm 0.022 \pm^{0.099}_{0.121}$
WH 125	$0.023 \pm 0.008 \pm^{0.006}_{0.009}$	$0.582 \pm 0.038 \pm^{0.066}_{0.101}$	$0.269 \pm 0.026 \pm^{0.034}_{0.026}$	$0.239 \pm 0.025 \pm^{0.034}_{0.023}$
ZH 125	< 0.01	$0.266 \pm 0.019 \pm^{0.029}_{0.050}$	$0.161 \pm 0.015 \pm^{0.022}_{0.019}$	$0.169 \pm 0.015 \pm^{0.014}_{0.020}$
$Z \rightarrow \tau\tau$	$12.89 \pm 0.55 \pm^{4.36}_{4.02}$	$76.3 \pm 5.6 \pm^{4.5}_{4.0}$	$347 \pm 11 \pm^{18}_{19}$	$243.5 \pm 7.5 \pm^{11.7}_{11.7}$
Fake- τ	$6.25 \pm 0.53 \pm^{1.66}_{1.52}$	$30.9 \pm 1.2 \pm^{6.0}_{8.4}$		
SS-Data			$149.0 \pm 12.2 \pm^{8.3}$	$38.8 \pm 6.8 \pm^{2.2}$
$W(\rightarrow \ell\nu) + \text{jets}$			$70.6 \pm 7.8 \pm^{10.3}_{15.6}$	$6.1 \pm 7.1 \pm^{5.3}_{9.2}$
$t\bar{t}$	$1.44 \pm 0.23 \pm^{0.25}_{0.30}$	$7.54 \pm 0.58 \pm^{0.76}_{0.98}$	$29.2 \pm 1.3 \pm^{2.6}_{2.0}$	$22.2 \pm 1.1 \pm^{1.6}_{1.3}$
$Z \rightarrow \ell\ell$	$0.96 \pm 0.34 \pm^{0.57}_{0.62}$	$3.12 \pm 0.81 \pm^{2.40}_{1.98}$	$52.8 \pm 5.4 \pm^{17.1}_{15.3}$	$26.2 \pm 3.5 \pm^{6.1}_{9.2}$
Di-boson	$0.69 \pm 0.23 \pm^{0.13}_{0.14}$	$2.55 \pm 0.24 \pm^{0.35}_{0.48}$	$2.91 \pm 0.44 \pm^{0.75}_{0.70}$	$2.39 \pm 0.46 \pm^{0.40}_{0.37}$
Signal	$2.889 \pm 0.060 \pm^{0.264}_{0.282}$	$6.17 \pm 0.13 \pm^{0.49}_{0.75}$	$8.78 \pm 0.19 \pm^{0.85}_{1.08}$	$6.70 \pm 0.17 \pm^{0.67}_{0.82}$
Background	$23.5 \pm 1.0 \pm^{4.7}_{4.4}$	$121.8 \pm 5.8 \pm^{7.9}_{9.5}$	$652 \pm 19 \pm^{28}_{30}$	$378 \pm 13 \pm^{15}_{18}$
Data	35	142	658	373

Table 6.5.: Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to VBF filtered MC for the VBF category, and the embedding sample for all others. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method, and Fake- τ is the estimate of fakes from multi-jets and $W(\rightarrow \ell\nu) + \text{jets}$ contributions using the fake-factor method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.

Process	0-Jet (e)	0-Jet (μ)	Rest (e)	Rest (μ)
ggF 125	$9.32 \pm 0.25 \pm \begin{smallmatrix} 0.81 \\ 0.90 \end{smallmatrix}$	$7.89 \pm 0.23 \pm \begin{smallmatrix} 0.54 \\ 0.73 \end{smallmatrix}$	$16.87 \pm 0.30 \pm \begin{smallmatrix} 1.07 \\ 1.17 \end{smallmatrix}$	$10.08 \pm 0.23 \pm \begin{smallmatrix} 0.65 \\ 0.71 \end{smallmatrix}$
VBF 125	$0.083 \pm 0.006 \pm 0.007$	$0.064 \pm 0.006 \pm \begin{smallmatrix} 0.006 \\ 0.004 \end{smallmatrix}$	$2.964 \pm 0.037 \pm \begin{smallmatrix} 0.176 \\ 0.181 \end{smallmatrix}$	$2.212 \pm 0.032 \pm \begin{smallmatrix} 0.129 \\ 0.132 \end{smallmatrix}$
WH 125	$0.032 \pm 0.009 \pm \begin{smallmatrix} 0.007 \\ 0.006 \end{smallmatrix}$	$0.035 \pm 0.009 \pm \begin{smallmatrix} 0.004 \\ 0.006 \end{smallmatrix}$	$0.816 \pm 0.045 \pm \begin{smallmatrix} 0.054 \\ 0.055 \end{smallmatrix}$	$0.637 \pm 0.040 \pm \begin{smallmatrix} 0.046 \\ 0.042 \end{smallmatrix}$
ZH 125	$0.019 \pm 0.005 \pm 0.004$	$0.015 \pm 0.005 \pm \begin{smallmatrix} 0.001 \\ 0.003 \end{smallmatrix}$	$0.465 \pm 0.026 \pm \begin{smallmatrix} 0.030 \\ 0.028 \end{smallmatrix}$	$0.361 \pm 0.022 \pm \begin{smallmatrix} 0.019 \\ 0.020 \end{smallmatrix}$
$Z \rightarrow \tau\tau$	$523 \pm 14 \pm 28$	$387.8 \pm 9.7 \pm \begin{smallmatrix} 18.7 \\ 18.8 \end{smallmatrix}$	$4878 \pm 41 \pm 268$	$2311 \pm 23 \pm 116$
SS-Data	$223 \pm 15 \pm 12$	$88.3 \pm 10.2 \pm 5.1$	$2085 \pm 46 \pm 117$	$594 \pm 26 \pm 34$
$W(\rightarrow \ell\nu)+\text{jets}$	$102 \pm 14 \pm \begin{smallmatrix} 15 \\ 11 \end{smallmatrix}$	$69.5 \pm 15.2 \pm \begin{smallmatrix} 6.1 \\ 9.0 \end{smallmatrix}$	$565 \pm 30 \pm \begin{smallmatrix} 47 \\ 60 \end{smallmatrix}$	$367 \pm 29 \pm \begin{smallmatrix} 32 \\ 42 \end{smallmatrix}$
$t\bar{t}$	$0.42 \pm 0.18 \pm \begin{smallmatrix} 0.20 \\ 0.18 \end{smallmatrix}$	$0.40 \pm 0.16 \pm \begin{smallmatrix} 0.11 \\ 0.10 \end{smallmatrix}$	$150.5 \pm 3.3 \pm \begin{smallmatrix} 7.5 \\ 8.5 \end{smallmatrix}$	$107.0 \pm 2.7 \pm 5.8$
$Z \rightarrow \ell\ell$	$185 \pm 11 \pm \begin{smallmatrix} 66 \\ 55 \end{smallmatrix}$	$161.0 \pm 9.7 \pm \begin{smallmatrix} 22.3 \\ 15.5 \end{smallmatrix}$	$131.3 \pm 9.1 \pm \begin{smallmatrix} 35.1 \\ 34.9 \end{smallmatrix}$	$55.2 \pm 5.5 \pm \begin{smallmatrix} 14.0 \\ 10.2 \end{smallmatrix}$
Di-boson	$3.03 \pm 0.32 \pm \begin{smallmatrix} 0.29 \\ 0.35 \end{smallmatrix}$	$2.50 \pm 0.25 \pm \begin{smallmatrix} 0.26 \\ 0.43 \end{smallmatrix}$	$24.7 \pm 1.7 \pm \begin{smallmatrix} 1.9 \\ 2.1 \end{smallmatrix}$	$18.8 \pm 1.1 \pm \begin{smallmatrix} 1.3 \\ 2.0 \end{smallmatrix}$
Signal	$9.45 \pm 0.25 \pm \begin{smallmatrix} 0.81 \\ 0.90 \end{smallmatrix}$	$8.01 \pm 0.23 \pm \begin{smallmatrix} 0.54 \\ 0.73 \end{smallmatrix}$	$21.12 \pm 0.31 \pm \begin{smallmatrix} 1.09 \\ 1.18 \end{smallmatrix}$	$13.28 \pm 0.24 \pm \begin{smallmatrix} 0.66 \\ 0.72 \end{smallmatrix}$
Background	$1036 \pm 27 \pm \begin{smallmatrix} 75 \\ 64 \end{smallmatrix}$	$707.41 \pm 22.91 \pm \begin{smallmatrix} 30.15 \\ 26.41 \end{smallmatrix}$	$7835 \pm 69 \pm \begin{smallmatrix} 298 \\ 300 \end{smallmatrix}$	$3444 \pm 46 \pm \begin{smallmatrix} 126 \\ 129 \end{smallmatrix}$
Data	947	705	7580	3274

Table 6.6.: Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to the embedding sample. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.

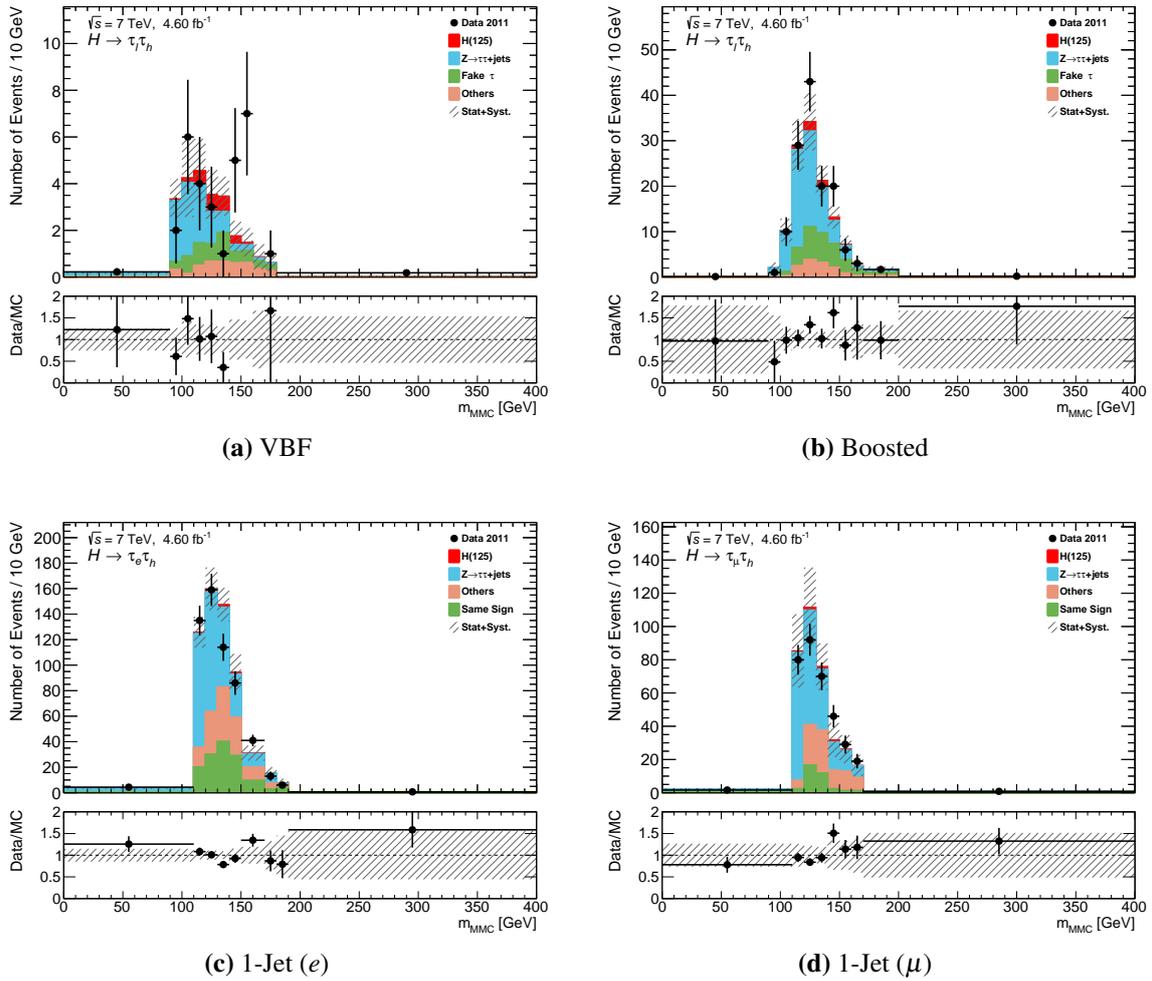


Figure 6.22.: m_{MMC} pre-fit distributions in the signal region for the VBF, Boosted, and 1-Jet analysis categories. $Z \rightarrow \ell\ell$, $t\bar{t}$, di-boson, and $W(\rightarrow \ell\nu) + \text{jets}$ backgrounds are grouped into the *others* category. Uncertainties are statistical and systematic. Bins are merged according to Algorithm 4.1, with bin contents scaled to number of events for every 10 GeV.

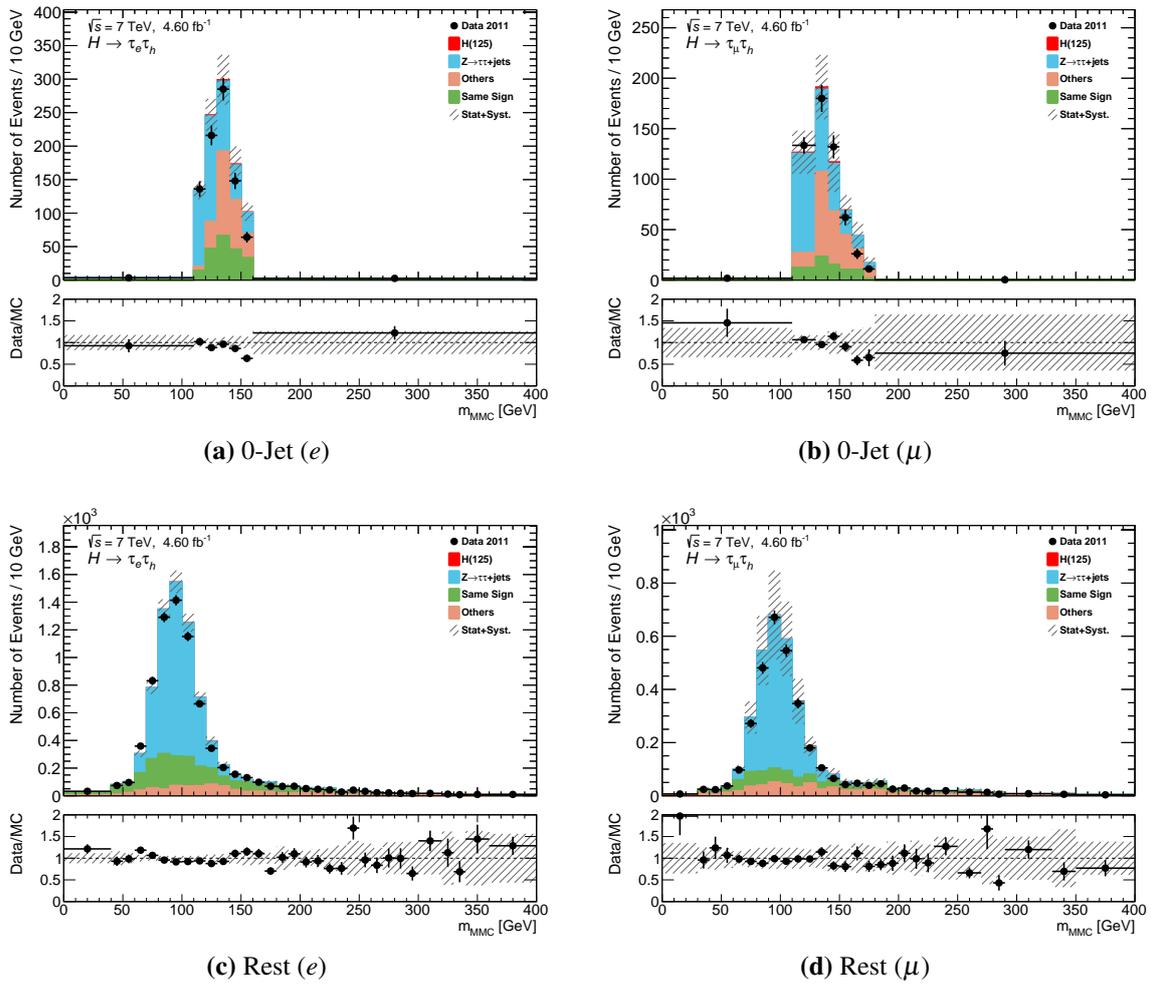


Figure 6.23.: m_{MMC} pre-fit distributions in the signal region for the 0-Jet and Rest analysis categories. $Z \rightarrow \ell\ell$, $t\bar{t}$, di-boson, and $W(\rightarrow \ell\nu) + \text{jets}$ backgrounds are grouped into the *others* category. Uncertainties are statistical and systematic. Bins are merged according to Algorithm 4.1, with bin contents scaled to number of events for every 10 GeV.

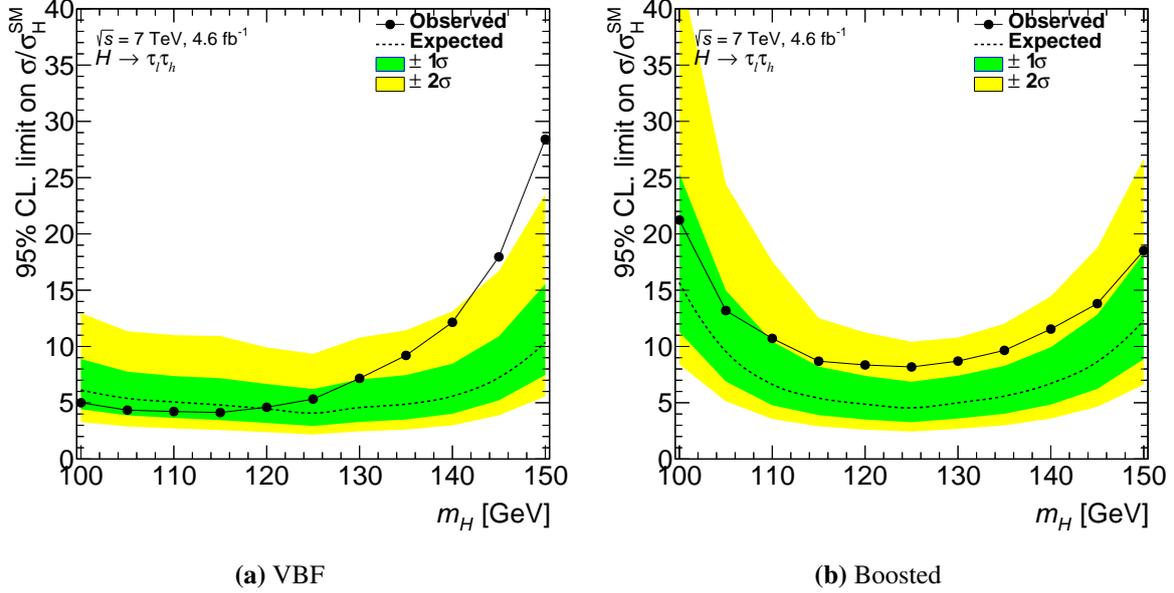


Figure 6.24.: Full-MVA observed and expected 95% CL. exclusion limits for the VBF and Boosted analysis categories, with 1σ and 2σ error bands. Uncertainties are statistical and systematic.

6.4.1. Exclusion Limits and Signal Strength

The category-specific exclusion limits are shown in Figures 6.24 and 6.25. The excess seen in the VBF m_{MMC} distribution in Figure 6.22 is evident here, with the observed exclusion limit ($28.4 \times \sigma_H^{SM}$) being outside the 2σ uncertainties of the expected value ($10.4^{+5.2}_{-2.9} \times \sigma_H^{SM}$) at 150 GeV. A similar excess towards 150 GeV is also seen in the 1-Jet category. The Boosted category has higher observed than expected limits across the entire m_H range, at about the $+1\sigma$ level. At the mass point of interest ($m_H = 125$ GeV), the observed exclusion limit for all categories is in agreement with the expected.

The combined exclusion limit for the full-MVA analysis is shown in Figure 6.26, along with the signal significance as a function of assumed m_H . At $m_H = 125$ GeV, the expected exclusion limit is $2.3^{+1.0}_{-0.6} \times \sigma_H^{SM}$, with an observed limit of $3.7 \times \sigma_H^{SM}$. The observed limit departs from the expected limit as m_H increases, due primarily to the excess seen in the VBF and 1-Jet categories. As the expected limits do not drop below $1 \times \sigma_H^{SM}$, the analysis is not sensitive to enough to observe the SM Higgs boson.

The excess towards $m_H = 150$ GeV is very evident in the signal significance plot, with the observed probability of the background giving a fluctuation as high as seen in the data of 2.9σ . It is worth noting however that as the analysis was optimised for $m_H = 125$ GeV, this fluctuation

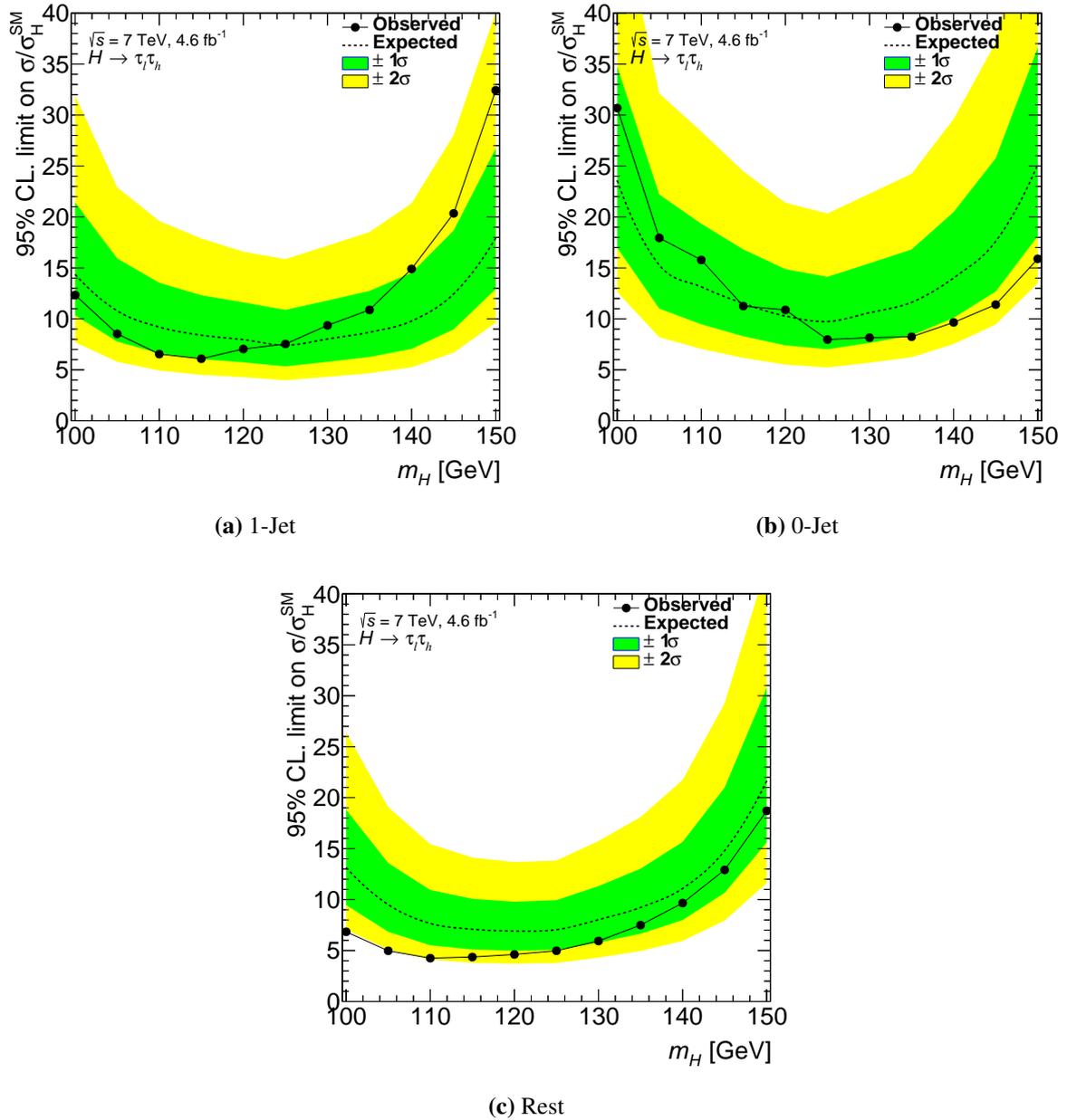


Figure 6.25.: Full-MVA observed and expected 95% CL. exclusion limits for the 1-Jet, 0-Jet, and Rest analysis categories, with 1σ and 2σ error bands. Uncertainties are statistical and systematic.

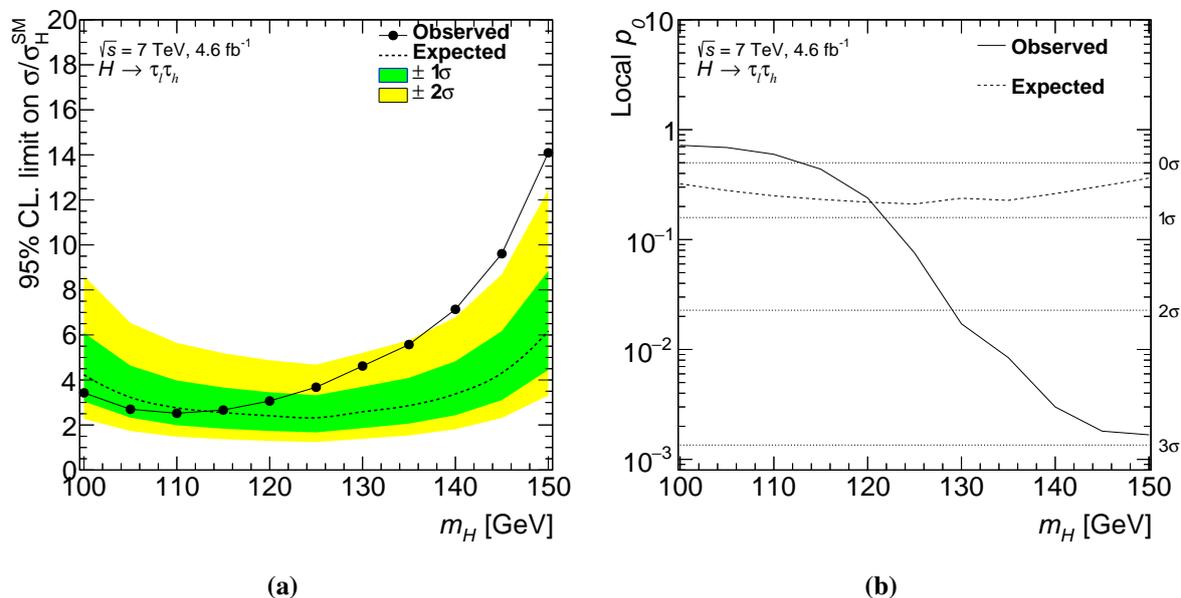


Figure 6.26.: Results from the combined analysis categories showing (a) the observed and expected 95% CL. exclusion limits with 1σ and 2σ error bands, and (b) the signal significance. Uncertainties are statistical and systematic.

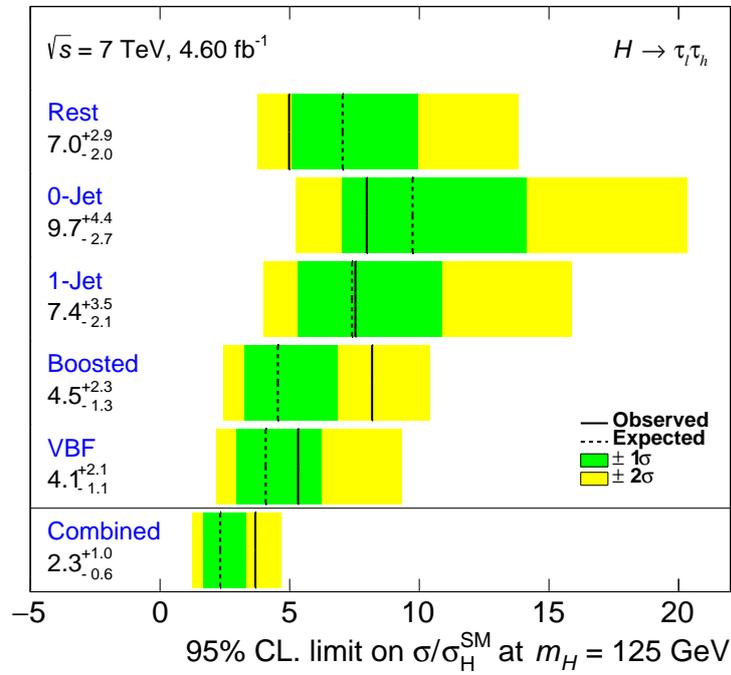
would need to be studied further to be deemed significant. At $m_H = 125$ GeV, the statistical significance of the data fluctuation is 1.4σ , with an expected value of 0.8σ . Therefore, no significant excess in the data is seen, across the whole m_H range.

Figure 6.27 shows a comparison of the exclusion limits at $m_H = 125$ GeV for all analysis categories, as well as the fitted signal strengths. It can be seen that the VBF and Boosted categories contribute the most to the combined exclusion limit by being the most sensitive, however interestingly the Rest category also proved to be quite sensitive.

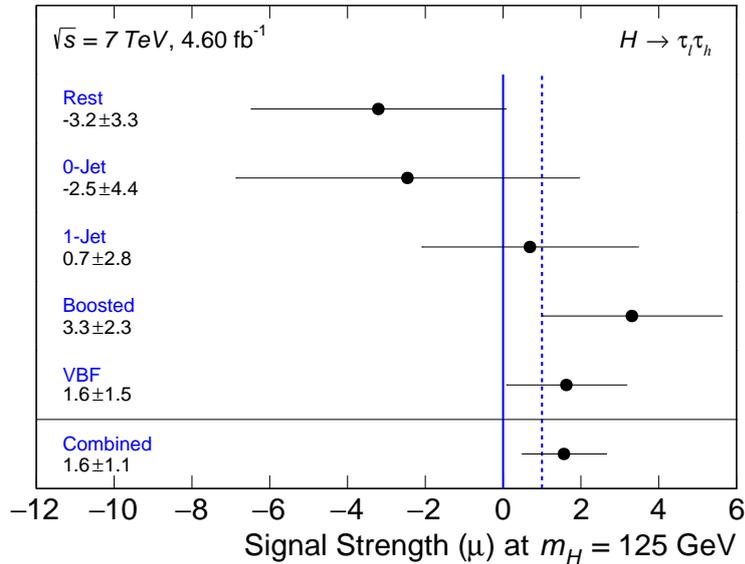
The fitted signal strength $\hat{\mu}$ is consistent with the background-only model, and the existence of the SM Higgs boson for all analysis categories, and the combined analysis, to within 2σ . The combined signal strength is observed to be $\hat{\mu} = 1.6 \pm 1.1$ at $m_H = 125$ GeV.

6.5. Comparison Between Analysis Strategies

The m_{MMC} distributions from the VBF category using the three analysis strategies are shown again in Figure 6.28, as an example of the difference in approaches. The main difference in the three cases is that, besides at least doubling the signal and background yields, the relative ratio



(a)



(b)

Figure 6.27.: (a) Comparison of full-MVA exclusion limits for each individual analysis category, and the combined limit, for $m_H = 125$ GeV. Uncertainties are statistical and systematic. (b) Extracted signal strength μ at $m_H = 125$ GeV for each analysis category, and the combined signal strength. Uncertainties are statistical and systematic.

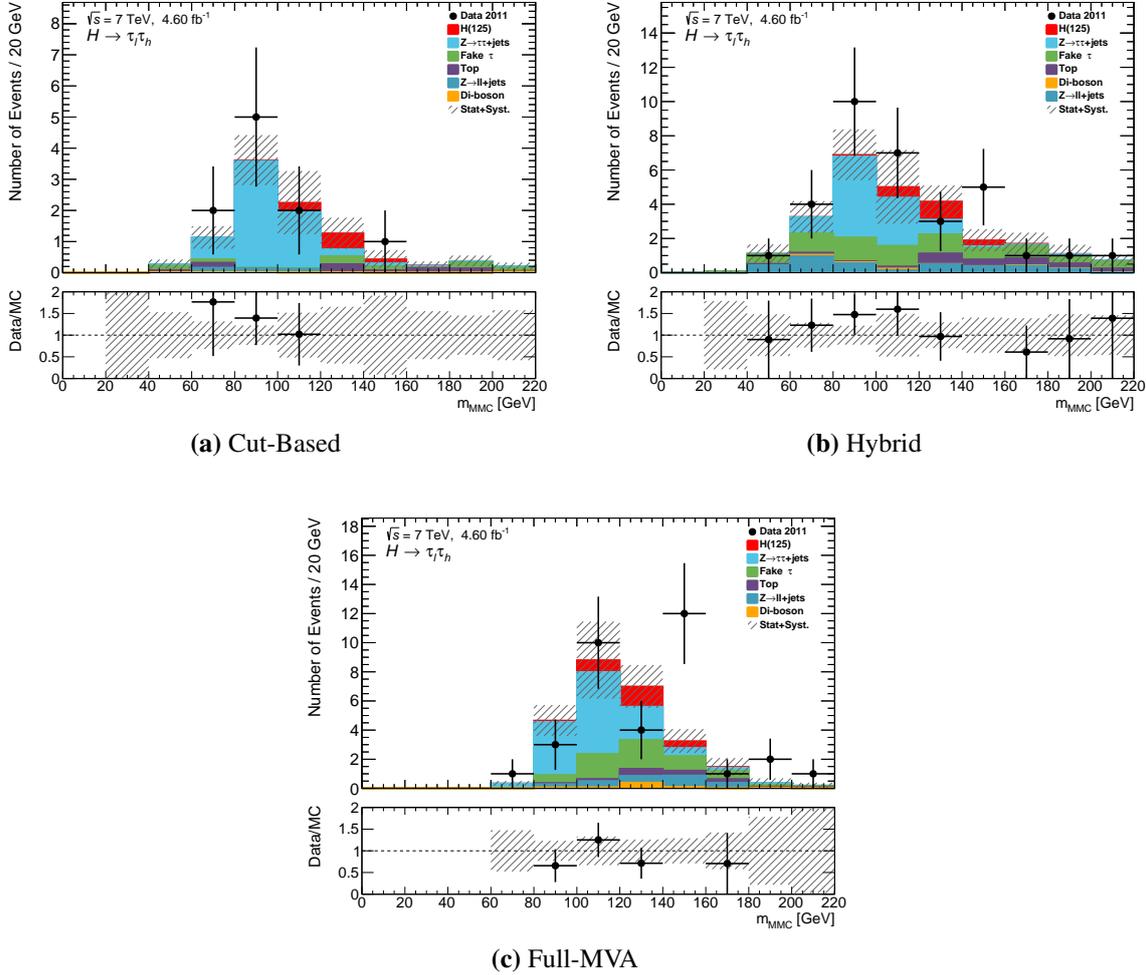
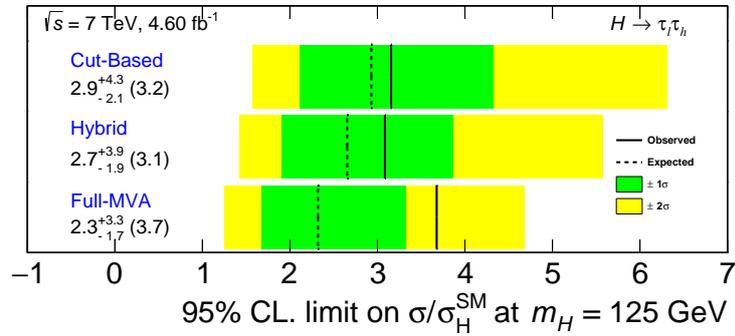


Figure 6.28.: m_{MMC} distributions in the VBF category with signal region selections for the three analysis strategies. Uncertainties are statistical and systematic.

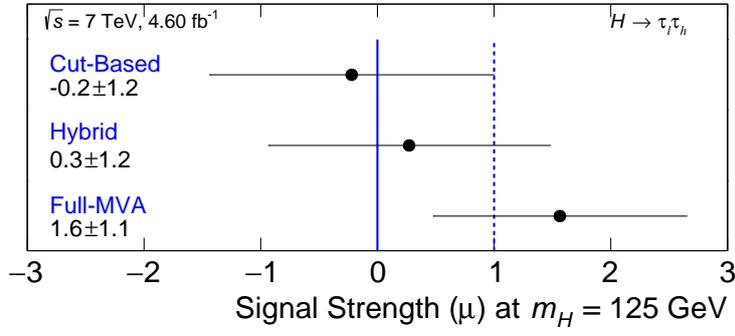
of the $Z \rightarrow \tau\tau$ background to all others decreases with the use of a BDT. The higher proportion of fake- τ backgrounds in the hybrid VBF category is due to BDT_{HBD} only being trained against $Z \rightarrow \tau\tau$ backgrounds.

Even after training and testing against the full background model however, BDT_{VBF} still did not recover the same background ratios as in the cut-based analysis. Rather, BDT_{VBF} selected events from both $Z \rightarrow \tau\tau$ and fake- τ backgrounds that were much closer to the signal Higgs mass, as is evident from the background shapes. The peak of the m_{MMC} distribution has moved from $m_H \approx 90$ GeV in the cut-based analysis to $m_H \approx 110$ GeV in the full-MVA. This can be attributed to BDT_{VBF} being trained exclusively against the $m_H = 125$ GeV Higgs sample, forcing the BDT to sculpt the backgrounds to the signal mass.

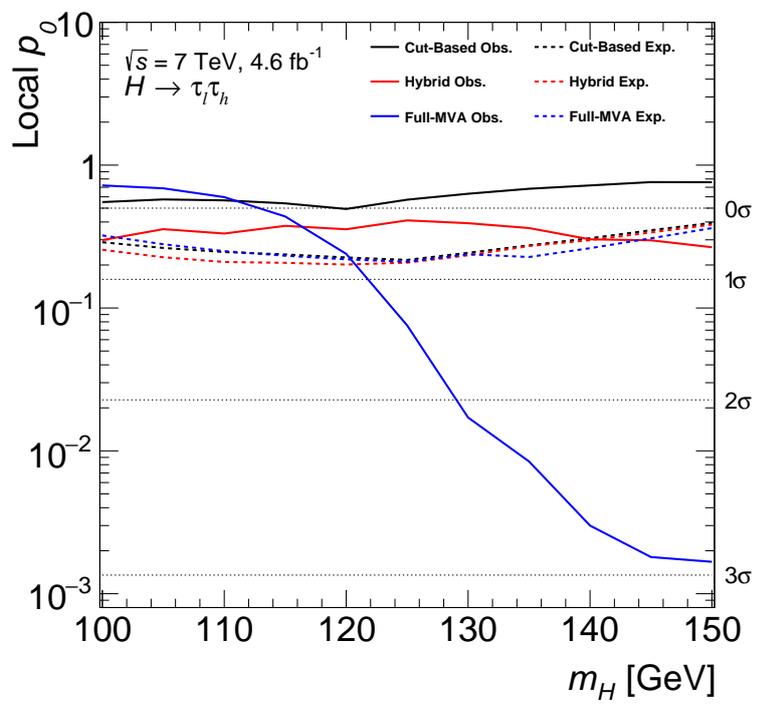
A summary of the main results of the three approaches taken to the $H \rightarrow \tau\tau$ analysis are shown in Figure 6.29. It is first important to note that the central values for the expected exclusion limits for each analysis are consistent to within 1σ uncertainties, although this is expected as the data used for each analysis is highly overlapping. As the MVA methods are applied to the analyses however, the uncertainties on the measurements decrease. The observed signal strengths for the three analyses also agree with each other to within 1σ , although the full-MVA method gives a $\hat{\mu}$ slightly more consistent with the existence of the SM Higgs boson. Finally, the statistical significances of the results are also shown. The expected significances of the three analyses are almost identical, however the observed significances vary greatly, especially with increasing m_H . The reason for this, as discussed previously, is the excess seen near $m_H = 150$ GeV in the VBF category of the full-MVA analysis, which is not seen in the others. At $m_H = 125$ GeV, the most statistically significant result is from the full-MVA analysis, with an signal strength of $\hat{\mu} = 1.6 \pm 1.1$, corresponding to an upward fluctuation of 1.4σ of the background-only expectation.



(a)



(b)



(c)

Figure 6.29.: Comparisons of the (a) 95% confidence level exclusion limits and (b) the extracted signal strength $\hat{\mu}$, at $m_H = 125$ GeV for the three analysis approaches. A comparison of the statistical significance of the results across the whole tested m_H range is shown in (c). Uncertainties are statistical and systematic.

Chapter 7.

Conclusions

The discovery of the Higgs boson in 2012 was one of the most important achievements in modern experimental physics. The LHC performed exceptionally, as did the detectors used to make the discovery. With its existence now confirmed, the properties of the Higgs boson need to be studied, most notably its coupling strength to the other particles of the SM. This thesis detailed three approaches to searching for the SM Higgs boson decaying to two tau leptons in $\sqrt{s} = 7$ TeV proton-proton collisions with the ATLAS detector.

The first analysis strategy used a cut-based method. Analysis categories were defined to separate the collision data into the distinct topologies of the Higgs production modes, namely the vector boson fusion, boosted, and gluon-gluon fusion scenarios. An expected (observed) 95% confidence level upper limit on σ/σ_H^{SM} was calculated to be $2.9_{-2.1}^{+4.3}$ (3.2) at $m_H = 125$ GeV. A fitted signal strength of $\hat{\mu} = -0.2 \pm 1.2$ was extracted, corresponding to a downward fluctuation of the data in the background-only model.

The hybrid method then improved on the cut-based approach by using MVA techniques to train a BDT to select the VBF category, based on the variables used in the cut-based selection. Working in a higher dimensional phase space, the BDT was able to improve the purity of the VBF selection, and subsequently the upper limit on σ/σ_H^{SM} , with an expected (observed) value of $2.7_{-1.9}^{+3.9}$ (3.1) at $m_H = 125$ GeV. The fitted signal strength was observed to be $\hat{\mu} = 0.3 \pm 1.2$, corresponding to a 0.2σ upward fluctuation of the data in the background-only model.

Finally, the analysis was extended further with the use of BDTs to define all analysis categories. The BDT training and testing was optimised for each category to maximise the discriminating power. Using the full-MVA method, an expected (observed) upper limit on σ/σ_H^{SM} was calculated to be $2.3_{-1.7}^{+3.3}$ (3.7) at $m_H = 125$ GeV. The fitted signal strength was

observed to be $\hat{\mu} = 1.6 \pm 1.1$, corresponding to a 1.4σ upward fluctuation of the data in the background-only model.

Although the implementation of MVA methods did not bring the expected upper limit on σ/σ_H^{SM} down to 1 as needed for the analysis to be sensitive, it did improve the expected limits and their uncertainties. This thesis shows the power of a BDT in exploiting complex correlations between variables to be able to separate events from otherwise irreducible backgrounds.

In the context of other ATLAS searches, MVA techniques have been used in recently published Higgs results. The analysis presented in [2] performed a search on the combined ATLAS $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data-set, looking for $H \rightarrow \tau\tau$ decays. The data was found to give a 4.5σ upward fluctuation of the background-only model, with an observed signal strength of $\mu = 1.43_{-0.37}^{+0.43}$, constituting evidence of a SM Higgs boson. A cut-based analysis was developed alongside this analysis as a cross-check, and it was *not* found to be sensitive to the Higgs signal. Again, this highlights the power of MVA techniques.

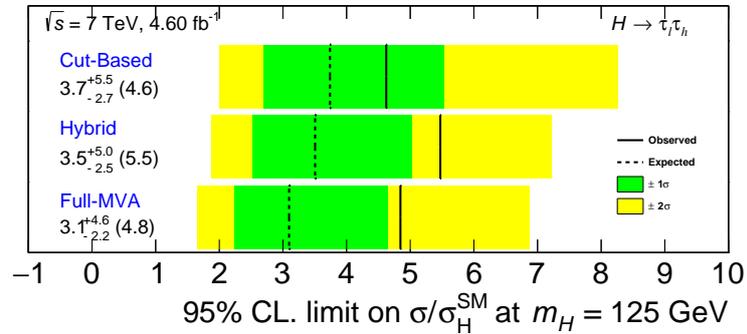
As the LHC continues its Run II at $\sqrt{s} = 13$ TeV, more precise measurements of the Higgs properties can be made. With over 300 fb^{-1} of data expected to be collected, analyses will begin to focus on measuring the Higgs mass, production cross-sections, branching fractions, and so on. MVA techniques, proven to give remarkable improvement on analysis sensitivity, will no doubt be utilised in the future.

Appendix A.

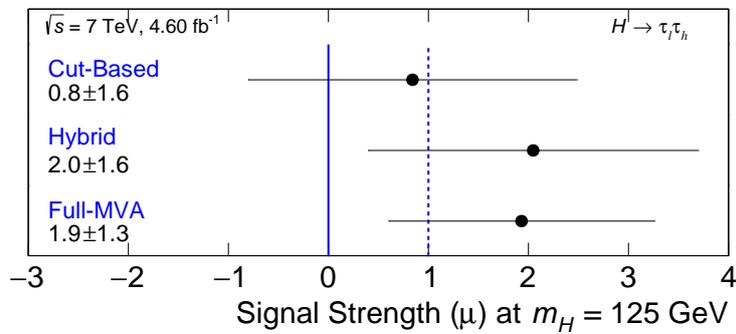
Non-Jet Combination Results

This chapter contains figures of the combined exclusion limits (Figure A.1a), signal strengths (Figure A.1b), and statistical significance (Figure A.1c) of the combined non-jet category fits. The procedure in which these figures are produced is discussed in Section 4.9. As outlined in Section 4.10.1, the non-jet combination fit was performed in order to study the effect of poor background modelling in the 1- and 0-Jet categories on the final exclusion limit. The non-jet combination fit for the full-MVA analysis returned an expected (observed) exclusion upper limit on σ/σ_H^{SM} of $3.1_{-2.2}^{+4.6}$ (4.8) at $m_H = 125$ GeV. The fitted signal strength was observed to be $\hat{\mu} = 1.9 \pm 1.3$, corresponding to a 1.5σ upward fluctuation of the data in the background-only model.

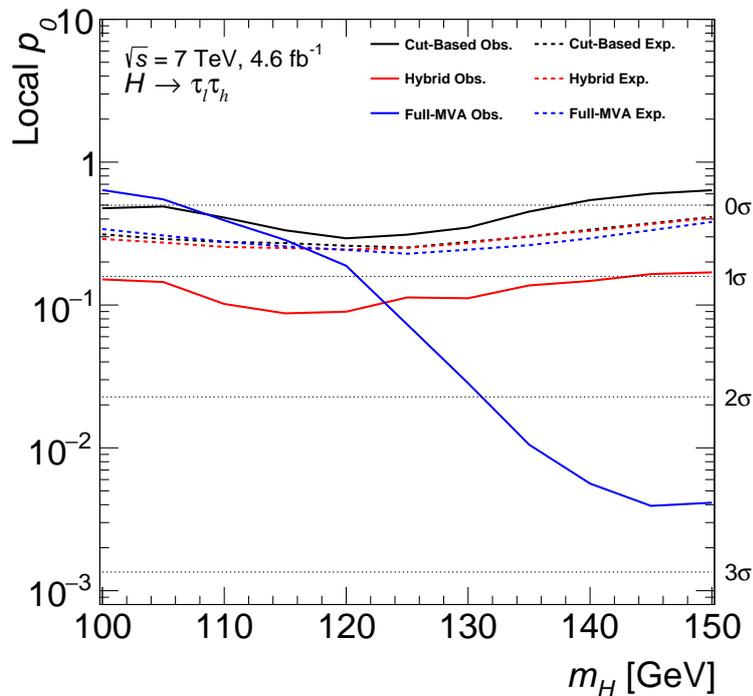
Compared to the full fit results presented in Section 6.4.1, the non-jet category fit is within statistical and systematic uncertainties. The trends in the exclusion limits and signal strengths between the three analysis approaches follow those of the full fit.



(a)



(b)



(c)

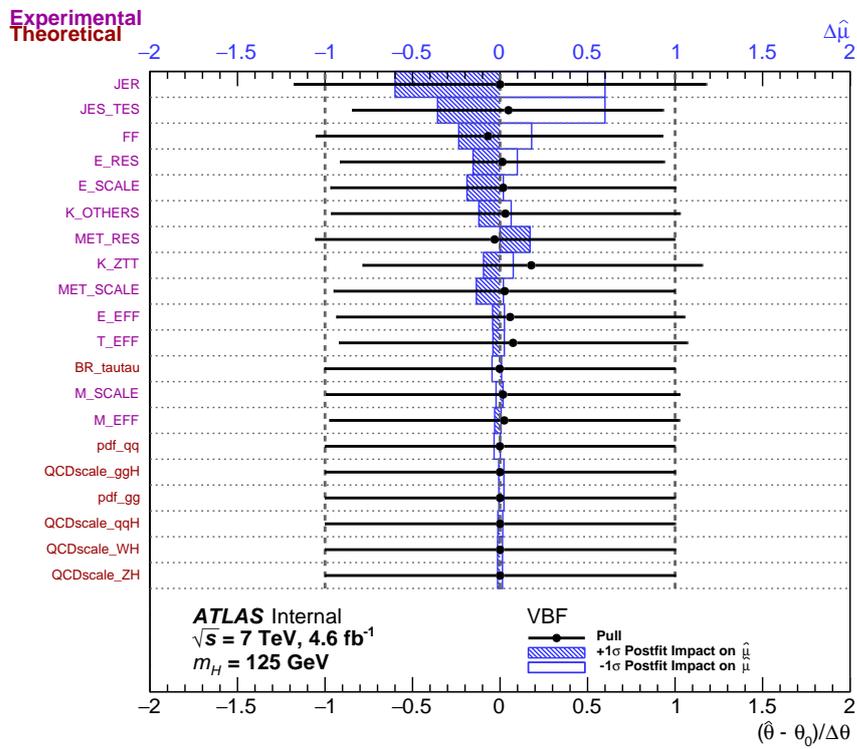
Figure A.1.: Comparisons of the (a) 95% confidence level exclusion limits and (b) the extracted signal strength $\hat{\mu}$, at $m_H = 125 \text{ GeV}$ for the three analysis approaches using only the VBF and Boosted categories in the combined fit. A comparison of the statistical significance of the results across the whole tested m_H range is shown in (c). Uncertainties are statistical and systematic.

Appendix B.

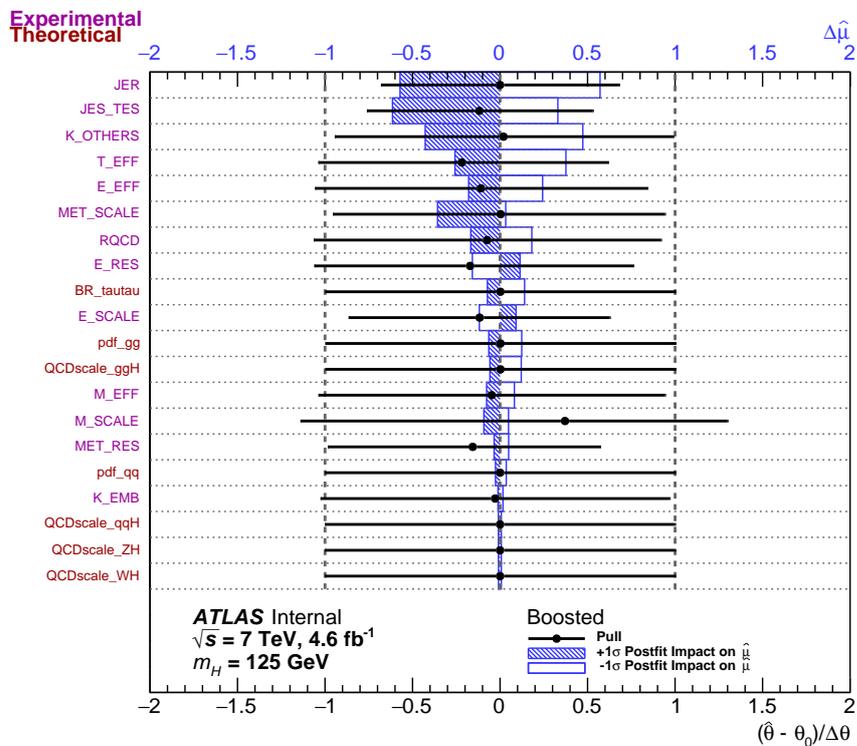
Category Specific Pull Distributions

This chapter contains the pull distributions for each analysis category, for the three analysis approaches (Figures [B.1–B.3](#) for the cut-based analysis, Figures [B.4–B.6](#) for the hybrid analysis, and Figures [B.7–B.10](#) for the full-MVA analysis). A discussion of pull distributions and their interpretation can be found in Section [4.10.1](#).

In general, the VBF and Boosted categories for all analysis approaches show the post-fit measurements of the nuisance parameters to be consistent with their nominal values. The 1-Jet, 0-Jet, and Rest categories show a greater deviation between the post-fit and nominal values, primarily in those nuisance parameters associated with the identification efficiency, energy scale, and energy resolution systematic uncertainties. All pulls lie within a 2σ deviation of the nominal value. Pull distributions for the non-jet combination fits discussed in Section [4.10.1](#) are also shown.

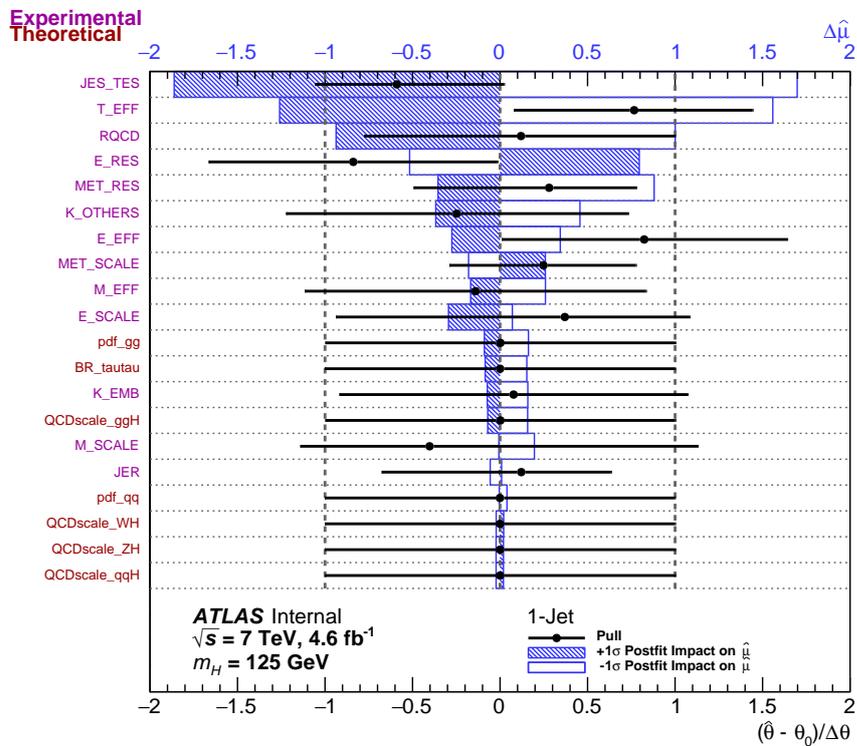


(a) VBF

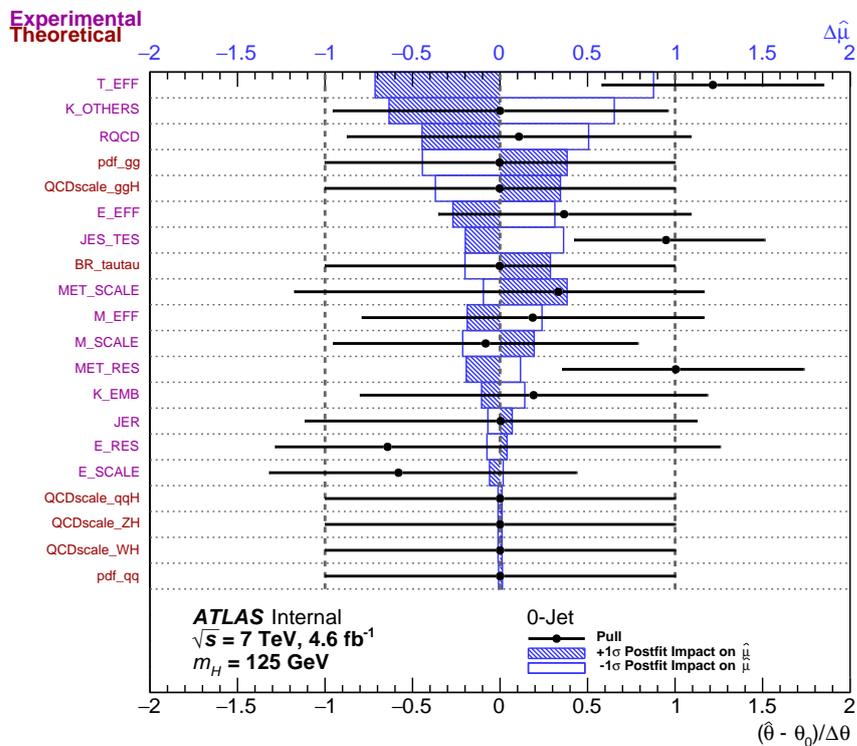


(b) Boosted

Figure B.1.: Pull distributions for the (a) VBF and (b) Boosted categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

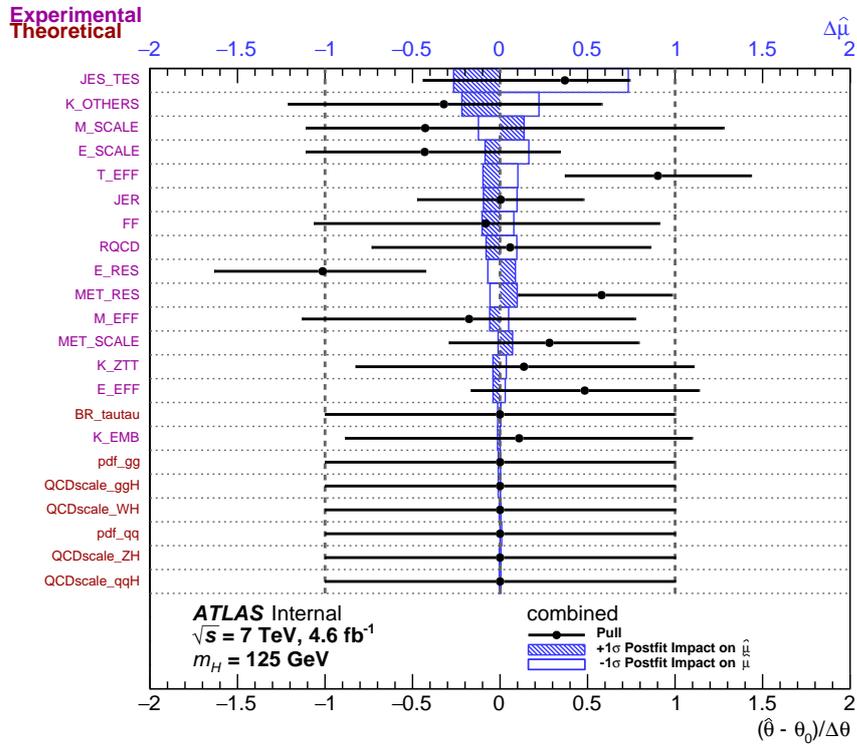


(a) 1-Jet

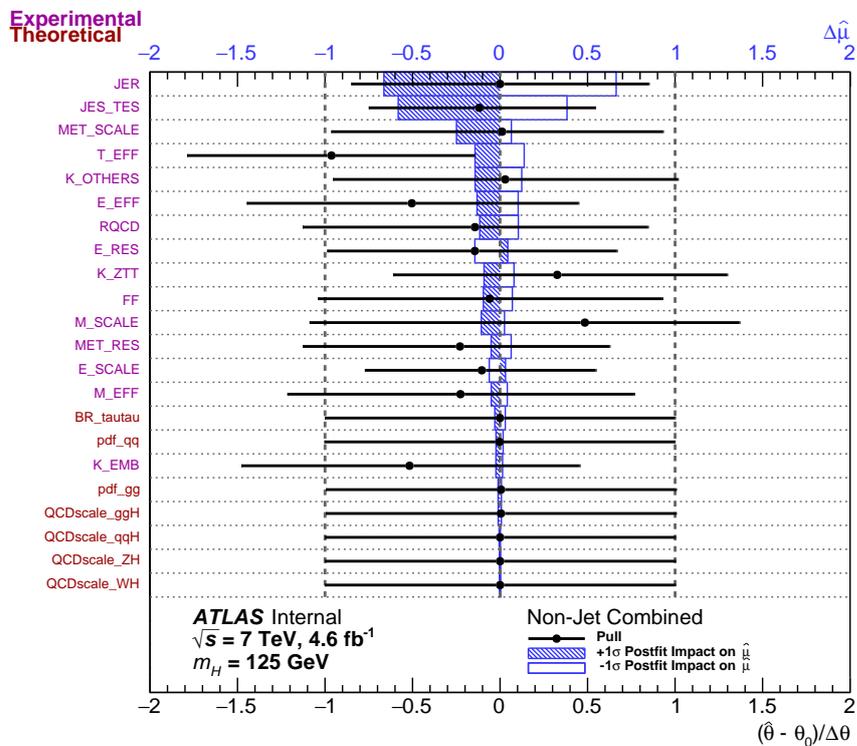


(b) 0-Jet

Figure B.2.: Pull distributions for the (a) 1-Jet and (b) 0-Jet categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

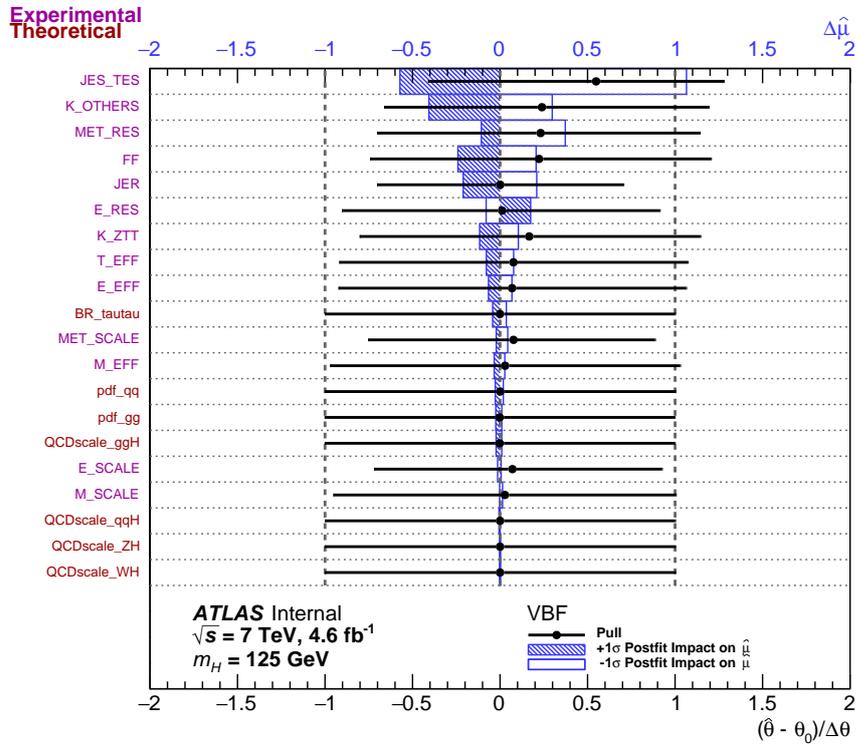


(a) Combined

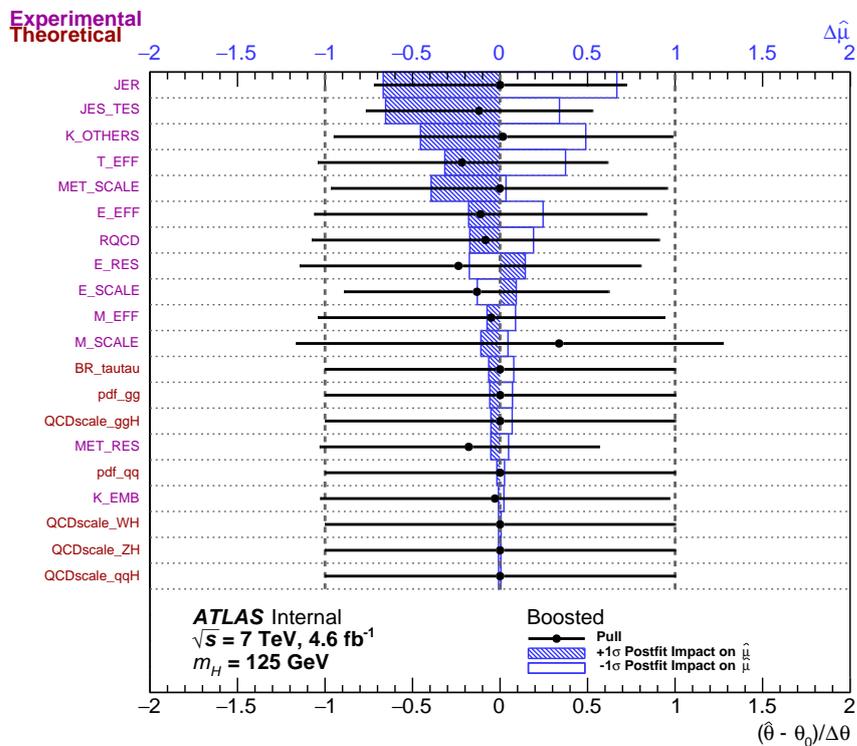


(b) Non-Jet Combined

Figure B.3.: Pull distributions for the (a) combined and (b) non-jet combined categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

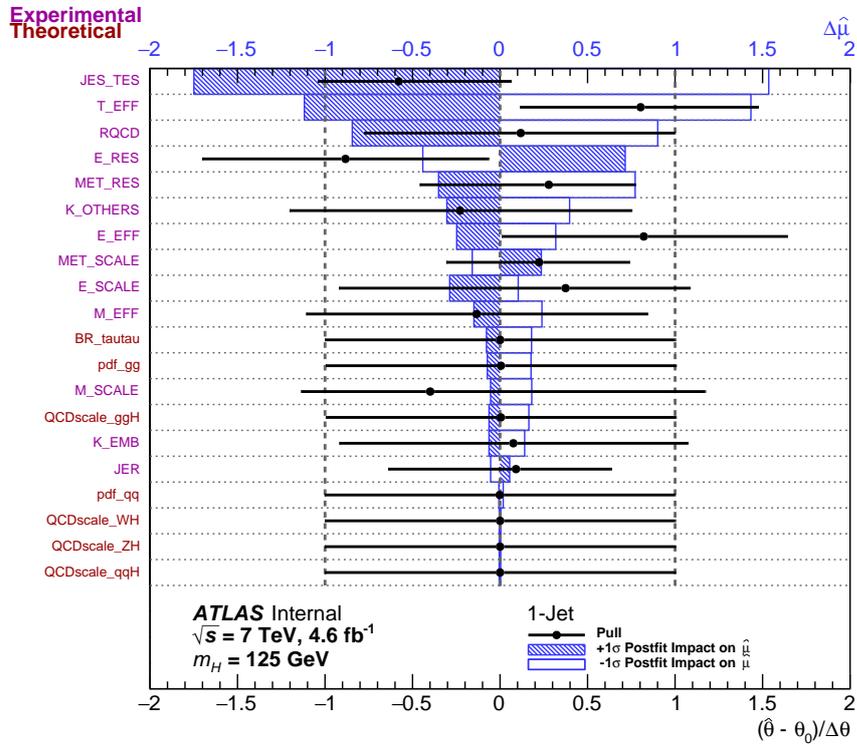


(a) VBF

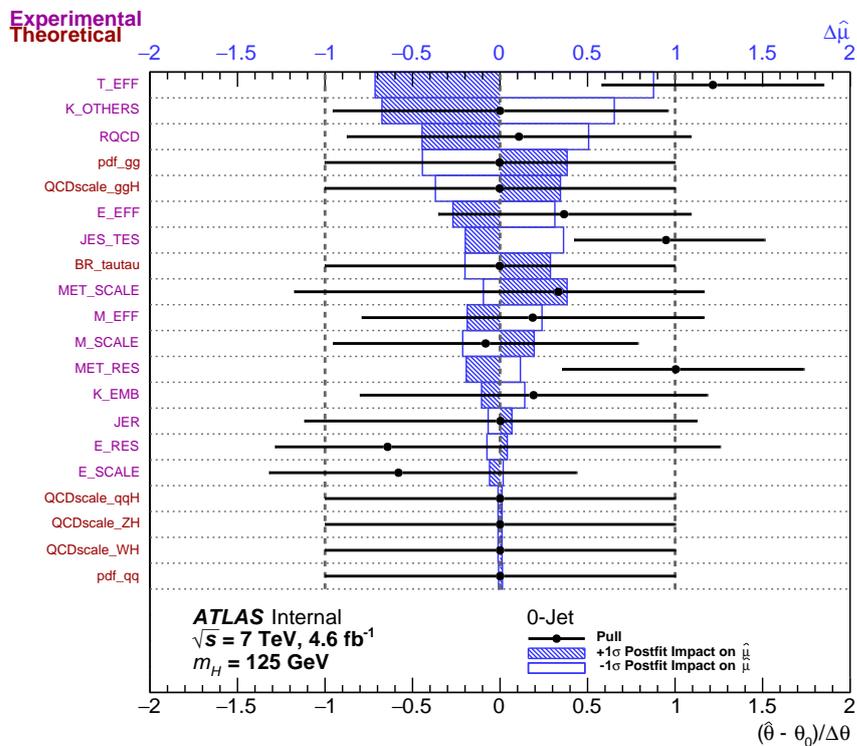


(b) Boosted

Figure B.4.: Pull distributions for the (a) VBF and (b) Boosted categories in the hybrid analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

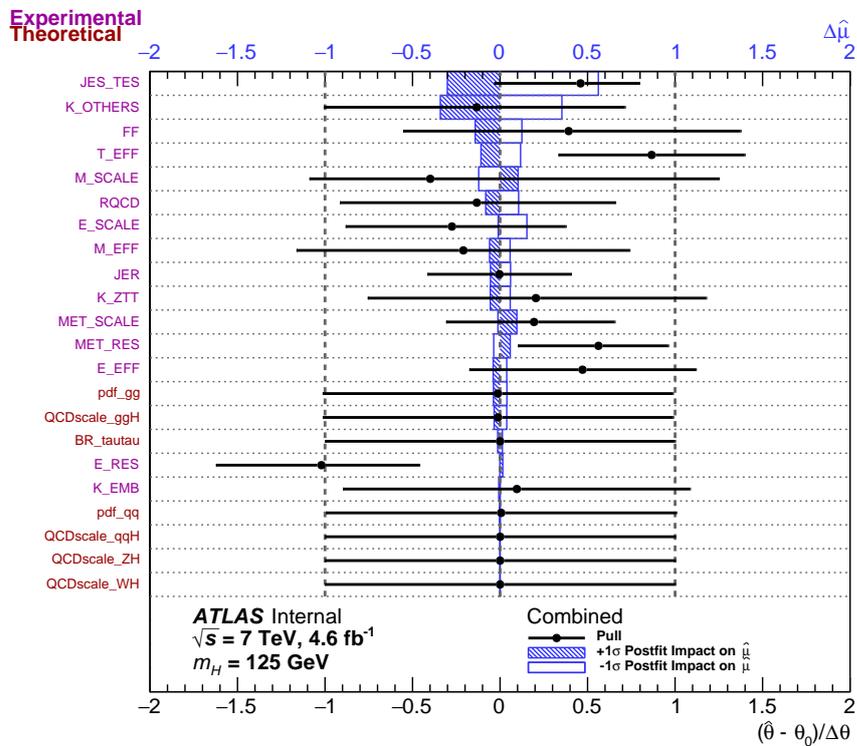


(a) 1-Jet

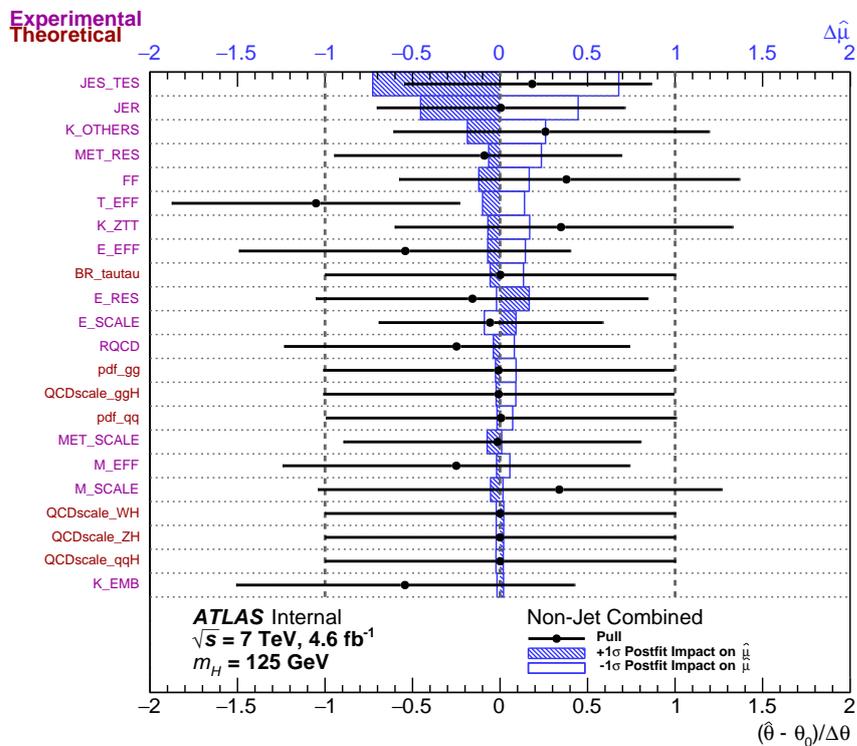


(b) 0-Jet

Figure B.5.: Pull distributions for the (a) 1-Jet and (b) 0-Jet categories in the hybrid analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

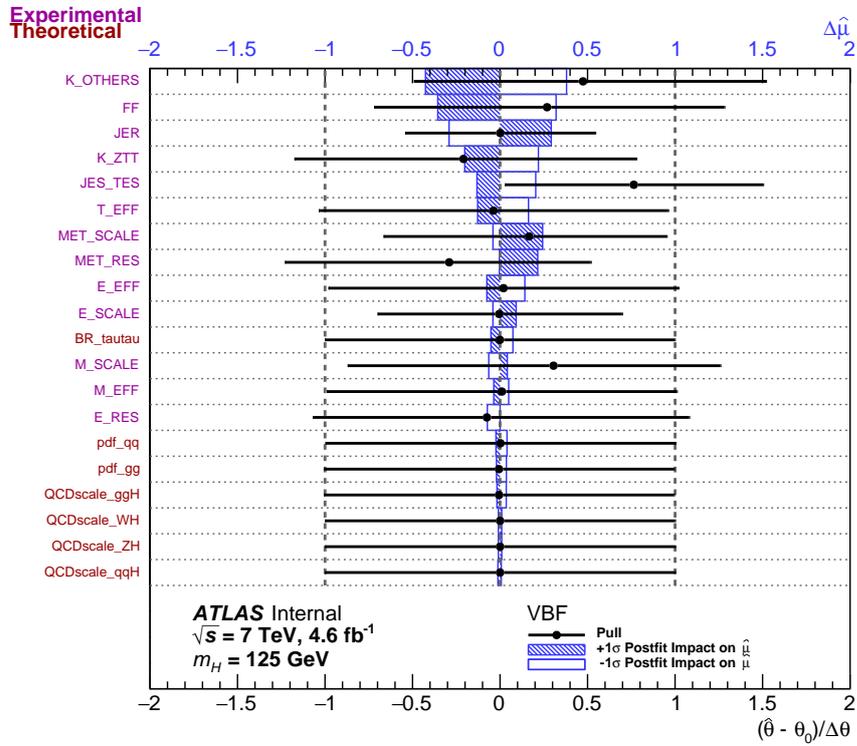


(a) Combined

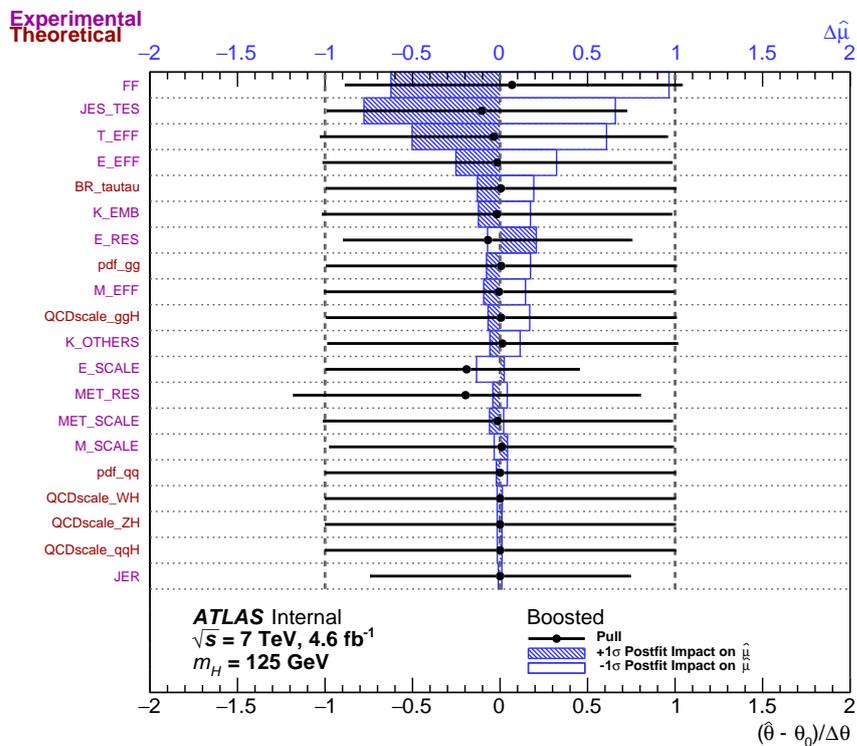


(b) Non-Jet Combined

Figure B.6.: Pull distributions for the (a) combined and (b) non-jet combined categories in the hybrid analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

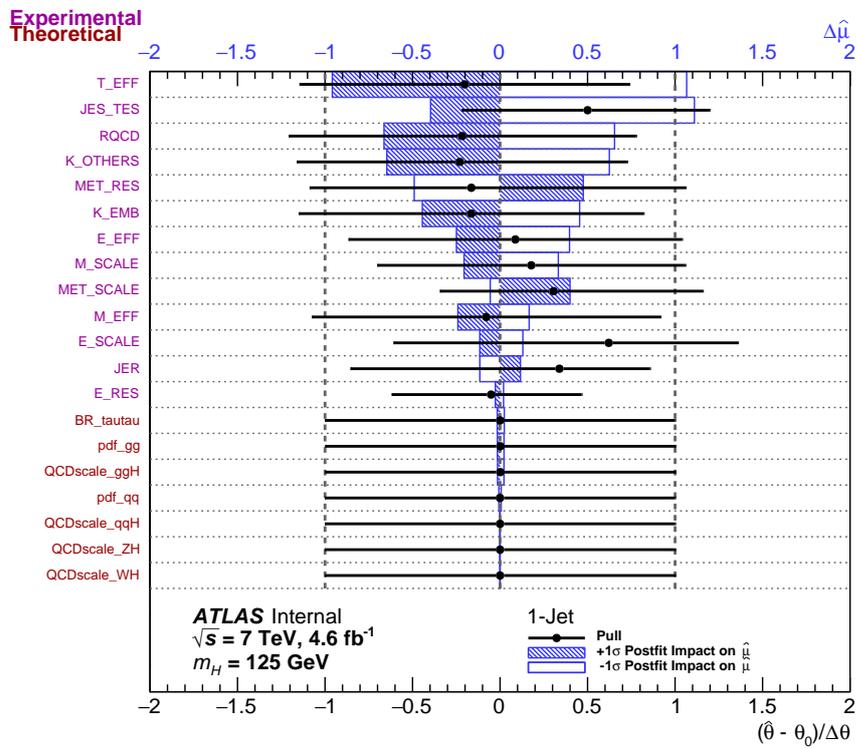


(a) VBF

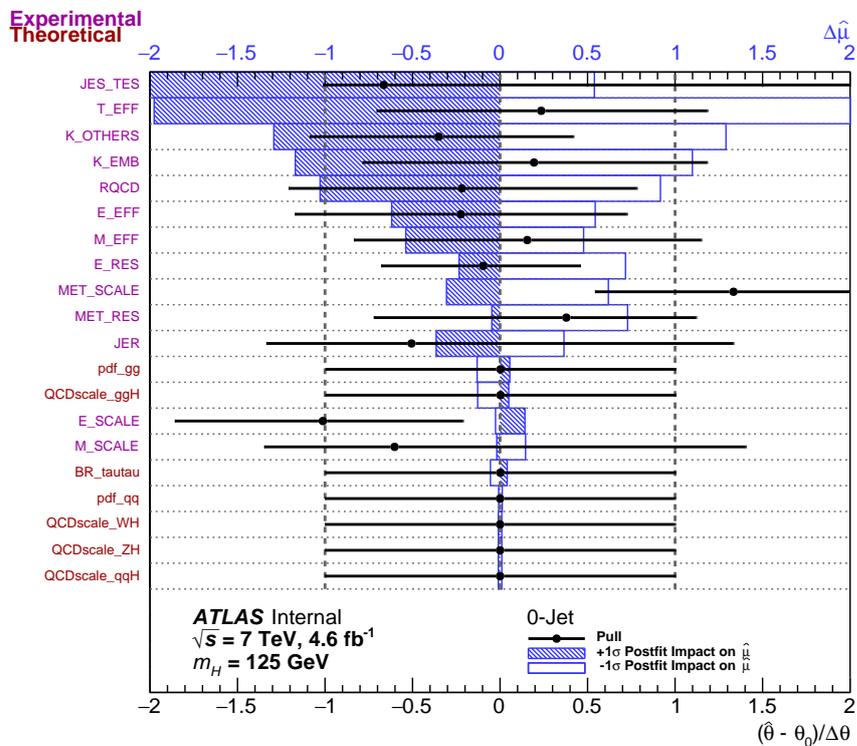


(b) Boosted

Figure B.7.: Pull distributions for the (a) VBF and (b) Boosted categories in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

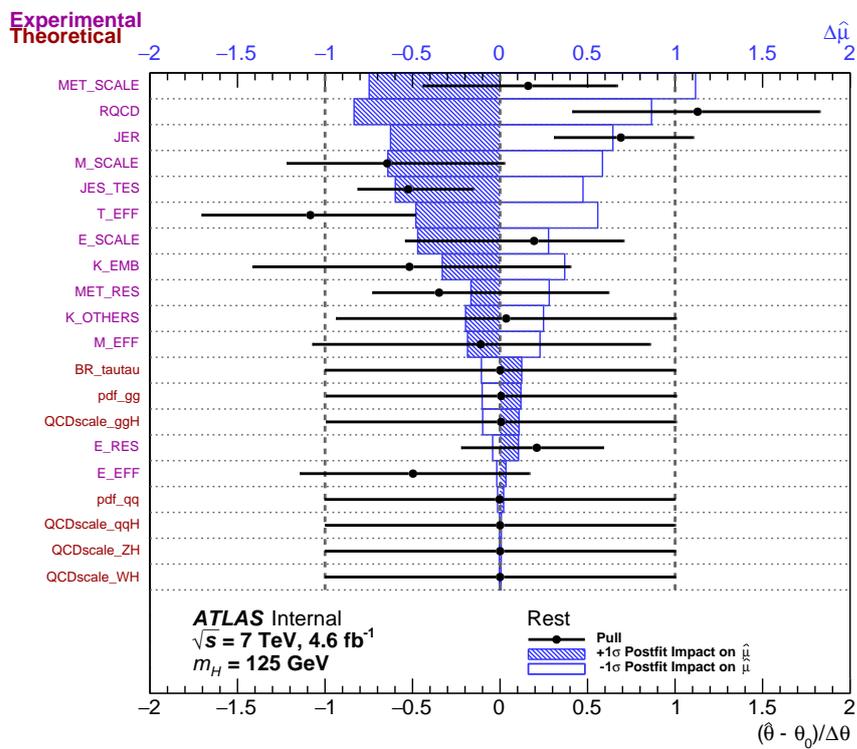


(a) 1-Jet



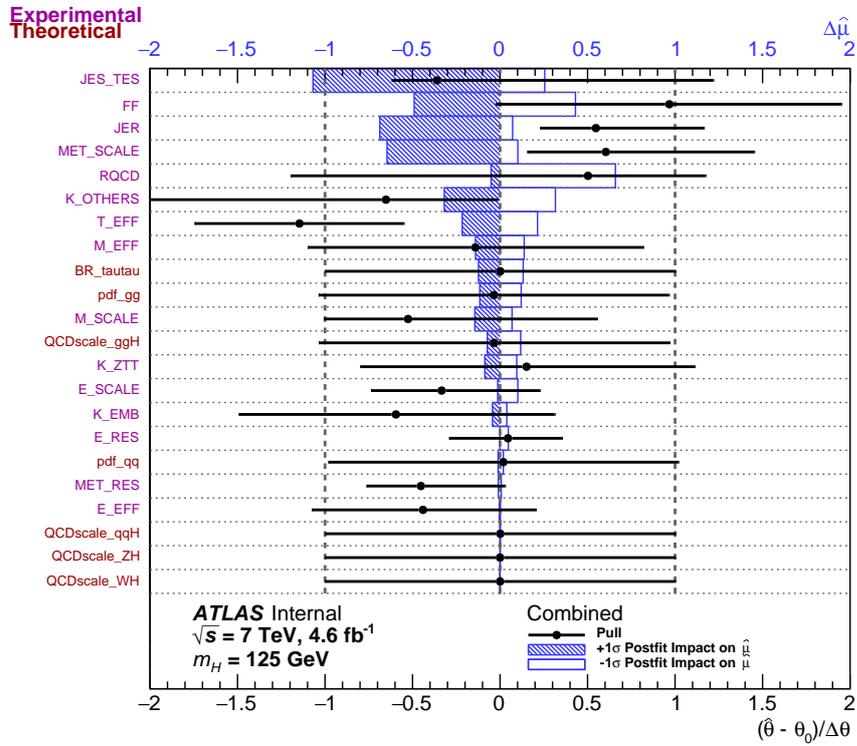
(b) 0-Jet

Figure B.8.: Pull distributions for the (a) 1-Jet and (b) 0-Jet categories in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

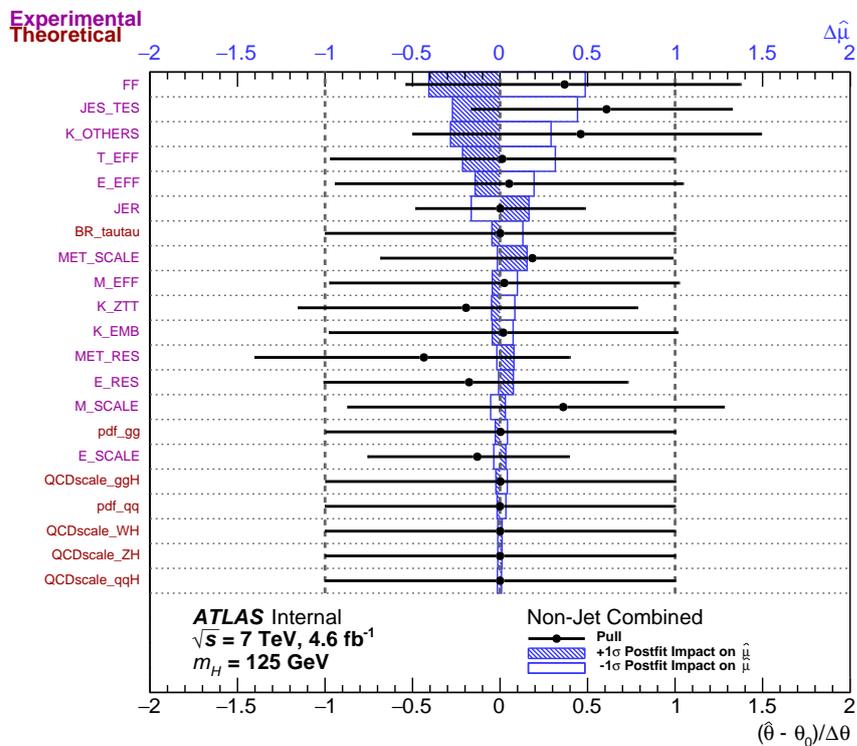


(a) Rest

Figure B.9.: Pull distribution for the Rest category in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.



(a) Combined



(b) Non-Jet Combined

Figure B.10.: Pull distributions for the (a) combined and (b) non-jet combined categories in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$.

Bibliography

- [1] S. Banerjee, E. Barberio, M. Beckingham, J. Biesiada, C. Boddy, D. Cavalli, S. Consonni, E. Coniavitis, E. Costaneda-Miranda, S. Farrington, A. Goussiou, K. Hanawa, H. Hass, Y. Heng, D. Jennens, G.-Y. Jeng, J. Keller, J. Kraus, J. Kroseberg, R. Madar, K. Nakamura, G. Nunes Hanninger, D. O’Neil, E. Pianori, C. Pizio, A. Pranko, D. Rousseau, N. Ruthmann, Y. Sakurai, T. Schwindt, R. Simoniello, K. Tan, A. Tanasijczuk, M. Trottier-McDonald, D. Varouchas, T. Vickey, M. Volpi, and S.-L. Wu, “Re-optimized Search for Standard Model $H \rightarrow \tau_{lep} + \tau_{had}$ with the ATLAS Detector in 7 TeV Proton-Proton Collisions,” Tech. Rep. ATL-COM-PHYS-2012-1087, CERN, Geneva, Jul, 2012. <https://cds.cern.ch/record/1463251>. Cited on pages [iv](#), [52](#), [77](#), [78](#), [94](#), and [211](#).
- [2] ATLAS Collaboration, G. Aad *et al.*, “Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector,” *JHEP* **04** (2015) 117, [arXiv:1501.04943 \[hep-ex\]](#). Cited on pages [iv](#), [20](#), [22](#), [184](#), and [209](#).
- [3] E. Noether, “Invariant Variation Problems,” *Gott. Nachr.* **1918** (1918) 235–257, [arXiv:physics/0503066 \[physics\]](#). [Transp. Theory Statist. Phys.1,186(1971)]. Cited on page [3](#).
- [4] Super-Kamiokande Collaboration, S. Fukuda *et al.*, “Solar B-8 and hep neutrino measurements from 1258 days of Super-Kamiokande data,” *Phys. Rev. Lett.* **86** (2001) 5651–5655, [arXiv:hep-ex/0103032 \[hep-ex\]](#). Cited on page [4](#).
- [5] Particle Data Group Collaboration, K. A. Olive *et al.*, “Review of Particle Physics,” *Chin. Phys.* **C38** (2014) 090001. Cited on pages [5](#), [49](#), and [220](#).
- [6] D. Griffiths, *Introduction to Elementary Particles*. Physics textbook. Wiley, 2008. Cited on page [6](#).
- [7] S. M. Bilenky and J. Hosek, “Glashow-Weinberg-Salam theory of electroweak interactions and the neutral currents,” *Phys. Rept.* **90** (1982) 73–157. Cited on page [10](#).

- [8] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons,” *Phys. Rev. Lett.* **13** (1964) 321–323. Cited on page 11.
- [9] P. W. Higgs, “Broken symmetries, massless particles and gauge fields,” *Phys. Lett.* **12** (1964) 132–133. Cited on page 11.
- [10] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons,” *Phys. Rev. Lett.* **13** (1964) 508–509. Cited on page 11.
- [11] P. W. Higgs, “Spontaneous Symmetry Breakdown without Massless Bosons,” *Phys. Rev.* **145** (1966) 1156–1163. Cited on page 11.
- [12] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles,” *Phys. Rev. Lett.* **13** (1964) 585–587. Cited on page 11.
- [13] T. W. B. Kibble, “Symmetry breaking in nonAbelian gauge theories,” *Phys. Rev.* **155** (1967) 1554–1561. Cited on page 11.
- [14] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries,” *Phys. Rev.* **127** (1962) 965–970. Cited on page 13.
- [15] T. Hambye and K. Riesselmann, “SM Higgs mass bounds from theory,” in *e^+e^- linear colliders: Physics and detector studies. Proceedings, Workshops, ECFA/DESY, Frascati, Italy, February 5-6, 1996, London, UK, July 4-6, 1996, Munich, Germany, September 16-18, 1996 and Hamburg, Germany, November 20-22, 1996. Pt.E.* 1997.
[arXiv:hep-ph/9708416](https://arxiv.org/abs/hep-ph/9708416) [hep-ph].
<http://alice.cern.ch/format/showfull?sysnb=0255880>. Cited on page 16.
- [16] SLD Heavy Flavor Group, DELPHI, ALEPH, OPAL, LEP Electroweak Working Group, L3 Collaboration, “A Combination of preliminary electroweak measurements and constraints on the standard model,” [arXiv:hep-ex/0212036](https://arxiv.org/abs/hep-ex/0212036) [hep-ex]. Cited on pages 16 and 208.
- [17] ALEPH Collaboration, A. Heister *et al.*, “Final results of the searches for neutral Higgs bosons in e^+e^- collisions at \sqrt{s} up to 209 GeV,” *Phys. Lett.* **B526** (2002) 191–205,
[arXiv:hep-ex/0201014](https://arxiv.org/abs/hep-ex/0201014) [hep-ex]. Cited on page 17.
- [18] OPAL Collaboration, G. Abbiendi *et al.*, “Search for the standard model Higgs boson with the OPAL detector at LEP,” *Eur. Phys. J.* **C26** (2003) 479–503,
[arXiv:hep-ex/0209078](https://arxiv.org/abs/hep-ex/0209078) [hep-ex]. Cited on page 17.
- [19] DELPHI Collaboration, J. Abdallah *et al.*, “Final results from DELPHI on the searches

- for SM and MSSM neutral Higgs bosons,” *Eur. Phys. J.* **C32** (2004) 145–183, [arXiv:hep-ex/0303013 \[hep-ex\]](#). Cited on page 17.
- [20] L3 Collaboration, P. Achard *et al.*, “Standard model Higgs boson with the L3 experiment at LEP,” *Phys. Lett.* **B517** (2001) 319–331, [arXiv:hep-ex/0107054 \[hep-ex\]](#). Cited on page 17.
- [21] W. D. Schlatter and P. M. Zerwas, “Searching for Higgs : From LEP towards LHC,” *Eur. Phys. J.* **H36** (2012) 579–600, [arXiv:1112.5127 \[physics.hist-ph\]](#). Cited on page 17.
- [22] ATLAS Collaboration, G. Aad *et al.*, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Phys. Lett.* **B716** (2012) 1–29, [arXiv:1207.7214 \[hep-ex\]](#). Cited on pages 17, 19, 20, 208, and 209.
- [23] LHC Higgs Cross Section Working Group Collaboration, J. R. Andersen *et al.*, “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties,” [arXiv:1307.1347 \[hep-ph\]](#). Cited on pages 17 and 208.
- [24] CDF Collaboration, F. Abe *et al.*, “The CDF Detector: An Overview,” *Nucl. Instrum. Meth.* **A271** (1988) 387–403. Cited on page 18.
- [25] D0 Collaboration, S. Abachi *et al.*, “The D0 Detector,” *Nucl. Instrum. Meth.* **A338** (1994) 185–253. Cited on page 18.
- [26] Tevatron New Physics Higgs Working Group, CDF, D0 Collaboration, “Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data,” 2012. [arXiv:1207.0449 \[hep-ex\]](#). http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-12-318. Cited on pages 18 and 208.
- [27] OPAL, DELPHI, LEP Working Group for Higgs boson searches, ALEPH, L3 Collaboration, R. Barate *et al.*, “Search for the standard model Higgs boson at LEP,” *Phys. Lett.* **B565** (2003) 61–75, [arXiv:hep-ex/0306033 \[hep-ex\]](#). Cited on pages 18 and 208.
- [28] ATLAS Collaboration, G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST* **3** (2008) S08003. Cited on pages 24, 25, 31, 34, 35, 39, 40, 41, 42, 209, and 210.
- [29] CMS Collaboration, S. Chatrchyan *et al.*, “The CMS experiment at the CERN LHC,”

- [JINST 3 \(2008\) S08004](#). Cited on page 26.
- [30] LHCb Collaboration, J. Alves, A. Augusto *et al.*, “The LHCb Detector at the LHC,” [JINST 3 \(2008\) S08005](#). Cited on page 26.
- [31] ALICE Collaboration, K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC,” [JINST 3 \(2008\) S08002](#). Cited on page 26.
- [32] ATLAS Collaboration, “ATLAS Collaboration Photo Gallery.” <http://www.atlas.ch/photos/>. Cited on pages 28, 31, 33, 37, 209, and 210.
- [33] ATLAS Collaboration, G. Aad *et al.*, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics,” [arXiv:0901.0512 \[hep-ex\]](#). Cited on pages 38, 93, and 220.
- [34] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, “Calorimeter Clustering Algorithms: Description and Performance,” Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008. <https://cds.cern.ch/record/1099735>. Cited on page 41.
- [35] ATLAS Collaboration, G. Aad *et al.*, “The ATLAS Simulation Infrastructure,” *Eur. Phys. J. C* **70** (2010) 823–874, [arXiv:1005.4568 \[physics.ins-det\]](#). Cited on pages 43 and 58.
- [36] ATLAS Collaboration Collaboration, *ATLAS Computing: technical design report*. Technical Design Report ATLAS. CERN, Geneva, 2005. <https://cds.cern.ch/record/837738>. Cited on page 44.
- [37] ATLAS Collaboration, G. Aad *et al.*, “Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC,” *Eur. Phys. J. C* **71** (2011) 1630, [arXiv:1101.2185 \[hep-ex\]](#). Cited on pages 44 and 94.
- [38] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175 \[hep-ph\]](#). Cited on pages 44 and 57.
- [39] T. Sjostrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, “An Introduction to PYTHIA 8.2,” *Comput. Phys. Commun.* **191** (2015) 159–177, [arXiv:1410.3012 \[hep-ph\]](#). Cited on page 44.
- [40] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, “HERWIG 6: An Event generator for hadron emission

- reactions with interfering gluons (including supersymmetric processes),” *JHEP* **01** (2001) 010, [arXiv:hep-ph/0011363 \[hep-ph\]](#). Cited on pages 44 and 58.
- [41] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, “ALPGEN, a generator for hard multiparton processes in hadronic collisions,” *JHEP* **07** (2003) 001, [arXiv:hep-ph/0206293 \[hep-ph\]](#). Cited on pages 44 and 58.
- [42] S. Frixione and B. R. Webber, “Matching NLO QCD computations and parton shower simulations,” *JHEP* **06** (2002) 029, [arXiv:hep-ph/0204244 \[hep-ph\]](#). Cited on pages 44 and 58.
- [43] S. Jadach, J. H. Kühn, and Z. Was, “TAUOLA—a library of Monte Carlo programs to simulate decays of polarized τ leptons,” *Computer Physics Communications* **64** no. 2, (1991) 275–299. Cited on pages 44 and 58.
- [44] E. Barberio and Z. Was, “PHOTOS—a universal Monte Carlo for QED radiative corrections: version 2.0,” *Computer Physics Communications* **79** no. 2, (1994) 291–308. Cited on pages 44 and 58.
- [45] GEANT4 Collaboration, S. Agostinelli *et al.*, “GEANT4: A Simulation toolkit,” *Nucl. Instrum. Meth.* **A506** (2003) 250–303. Cited on pages 44 and 58.
- [46] J. Pequeno and P. Schaffner, “An computer generated image representing how ATLAS detects particles.” Jan, 2013. Cited on pages 45 and 210.
- [47] ATLAS Collaboration, G. Aad *et al.*, “Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data,” *Eur. Phys. J.* **C74** no. 7, (2014) 2941, [arXiv:1404.2240 \[hep-ex\]](#). Cited on pages 46 and 61.
- [48] ATLAS Collaboration, G. Aad *et al.*, “Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC protonproton collision data,” *Eur. Phys. J.* **C74** no. 11, (2014) 3130, [arXiv:1407.3935 \[hep-ex\]](#). Cited on page 47.
- [49] S. Hassani, L. Chevalier, E. Lancon, J. F. Laporte, R. Nicolaidou, and A. Ouraou, “A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages,” *Nucl. Instrum. Meth.* **A572** (2007) 77–79. Cited on page 47.
- [50] T. Lagouri *et al.*, “A Muon Identification and Combined Reconstruction Procedure for the

- ATLAS Detector at the LHC at CERN,” *IEEE Trans. Nucl. Sci.* **51** (2004) 3030–3033.
Cited on page 47.
- [51] M. Cacciari, G. P. Salam, and G. Soyez, “The Anti-k(t) jet clustering algorithm,” *JHEP* **04** (2008) 063, [arXiv:0802.1189](https://arxiv.org/abs/0802.1189) [hep-ph]. Cited on page 47.
- [52] ATLAS Collaboration Collaboration, “Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data,”
<https://cds.cern.ch/record/1369219>. Cited on page 48.
- [53] ATLAS Collaboration Collaboration, “Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS,” Tech. Rep. ATLAS-CONF-2011-152, CERN, Geneva, Nov, 2011.
<https://cds.cern.ch/record/1398195>. Cited on pages 48, 50, and 93.
- [54] I. Vichou, “Performance of Missing Transverse Momentum Reconstruction in ATLAS with Proton-Proton Collisions at $\sqrt{s} = 7$ TeV,” Tech. Rep. ATL-PHYS-PROC-2012-212, CERN, Geneva, Oct, 2012. <https://cds.cern.ch/record/1482392>. Cited on page 50.
- [55] ATLAS Collaboration, G. Aad *et al.*, “Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS,” *Eur. Phys. J.* **C72** (2012) 1844, [arXiv:1108.5602](https://arxiv.org/abs/1108.5602) [hep-ex]. Cited on page 51.
- [56] ATLAS Collaboration Collaboration, “Search for the Standard Model Higgs boson produced in association with a vector boson and decaying into a tau pair in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector,” Tech. Rep. [arXiv:1511.08352](https://arxiv.org/abs/1511.08352).
CERN-PH-2015-226, Nov, 2015. <https://cds.cern.ch/record/2109123>.
Comments: 17 pages (+author list: 33 pages total), 9 figures, 7 tables. Submitted to Phys.Rev.D. All figures including auxiliary figures are available at <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/HIGGS-2014-01>. Cited on page 53.
- [57] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG,” *JHEP* **04** (2009) 002, [arXiv:0812.0578](https://arxiv.org/abs/0812.0578) [hep-ph]. Cited on page 58.
- [58] P. Nason and C. Oleari, “NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG,” *JHEP* **02** (2010) 037, [arXiv:0911.5299](https://arxiv.org/abs/0911.5299) [hep-ph]. Cited on page 58.

- [59] J. Alwall *et al.*, “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions,” *Eur. Phys. J.* **C53** (2008) 473–500, [arXiv:0706.2569 \[hep-ph\]](#). Cited on page 58.
- [60] J. Butterworth, J. R. Forshaw, and M. H. Seymour, “Multiparton interactions in photoproduction at HERA,” *Zeitschrift für Physik C: Particles and Fields* **72** no. 4, (1996) 637–646. Cited on page 58.
- [61] A. collaboration *et al.*, “ATLAS Monte Carlo tunes for MC09,” *ATLAS public note: ATL-PHYS-PUB-2010-002* (2010). Cited on page 58.
- [62] ATLAS Collaboration Collaboration, “Muon Momentum Resolution in First Pass Reconstruction of pp Collision Data Recorded by ATLAS in 2010,” <https://cds.cern.ch/record/1338575>. Cited on pages 59 and 93.
- [63] ATLAS Collaboration Collaboration, “Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector,” <https://cds.cern.ch/record/1345743>. Cited on page 59.
- [64] ATLAS Collaboration, G. Aad *et al.*, “Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV,” *Eur. Phys. J.* **C73** no. 3, (2013) 2304, [arXiv:1112.6426 \[hep-ex\]](#). Cited on pages 62 and 94.
- [65] T. Barillari, E. Bergeaas Kuutmann, T. Carli, J. Erdmann, P. Giovannini, K. J. Grahn, C. Issever, A. Jantsch, A. Kiryunin, K. Lohwasser, A. Maslennikov, S. Menke, H. Oberlack, G. Pospelov, E. Rauter, P. Schacht, F. Span, P. Speckmayer, P. Stavina, and P. Strzenec, “Local Hadronic Calibration,” <https://cds.cern.ch/record/1112035>. Due to a report-number conflict with another document, the report-number ATL-LARG-PUB-2009-001-2 has been assigned. Cited on page 62.
- [66] A. Elagin, P. Murat, A. Pranko, and A. Safonov, “A New Mass Reconstruction Technique for Resonances Decaying to di-tau,” *Nucl. Instrum. Meth.* **A654** (2011) 481–489, [arXiv:1012.4686 \[hep-ex\]](#). Cited on pages 63, 64, and 69.
- [67] ATLAS Collaboration Collaboration, “Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV,” <https://cds.cern.ch/record/1277678>. Cited on page 64.
- [68] S. Dittmaier *et al.*, “Handbook of LHC Higgs Cross Sections: 2. Differential Distributions,” [arXiv:1201.3084 \[hep-ph\]](#). Cited on pages 91 and 221.

- [69] LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier *et al.*, “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables,” [arXiv:1101.0593](https://arxiv.org/abs/1101.0593) [hep-ph]. Cited on pages 91, 92, and 222.
- [70] J. M. Campbell and R. K. Ellis, “An Update on vector boson pair production at hadron colliders,” *Phys. Rev.* **D60** (1999) 113006, [arXiv:hep-ph/9905386](https://arxiv.org/abs/hep-ph/9905386) [hep-ph]. Cited on page 92.
- [71] ATLAS Collaboration, G. Aad *et al.*, “Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data,” *Eur. Phys. J.* **C72** (2012) 1909, [arXiv:1110.3174](https://arxiv.org/abs/1110.3174) [hep-ex]. Cited on page 92.
- [72] M. Agustoni, M. Aharrouche, A. Ahmad, N. Besson, M. Boonekamp, L. Carminati, J. de Vivie, H. De La Torre, J. Del Peso, J. Farley, V. Gallo, S. Glazov, C. Handel, S. Haug, M. Karnevskiy, I. Koletsou, N. Lorenzo Martinez, N. Makovec, L. Mandelli, Y. Nakahama, R. Turra, G. Unal, and G. Pasztor, “Electromagnetic energy scale in-situ calibration and performance: Supporting document for the egamma performance paper,” Tech. Rep. ATL-COM-PHYS-2011-263, CERN, Geneva, Mar, 2011. <https://cds.cern.ch/record/1335395>. Cited on page 92.
- [73] ATLAS Collaboration Collaboration, “Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector,” Tech. Rep. ATLAS-CONF-2011-063, CERN, Geneva, Apr, 2011. <https://cds.cern.ch/record/1345743>. Cited on page 93.
- [74] ATLAS Collaboration Collaboration, “Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC in 2011,” Tech. Rep. ATLAS-CONF-2012-054, CERN, Geneva, Jun, 2012. <https://cds.cern.ch/record/1453781>. Cited on page 93.
- [75] ATLAS Collaboration Collaboration, “Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s} = 7$ TeV,” Tech. Rep. ATLAS-CONF-2010-054, CERN, Geneva, Jul, 2010. <https://cds.cern.ch/record/1281311>. Cited on page 94.
- [76] ATLAS Collaboration Collaboration, “Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\sqrt{s} = 7$ TeV in 2011,” Tech. Rep. ATLAS-CONF-2012-048, CERN, Geneva, May, 2012. <https://cds.cern.ch/record/1450089>. Cited on page 94.

- [77] ATLAS Collaboration Collaboration, “Performance of the ATLAS muon trigger in 2011,” Tech. Rep. ATLAS-CONF-2012-099, CERN, Geneva, Jul, 2012. <https://cds.cern.ch/record/1462601>. Cited on page 94.
- [78] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur. Phys. J.* **C71** (2011) 1554, [arXiv:1007.1727 \[physics.data-an\]](https://arxiv.org/abs/1007.1727). [Erratum: *Eur. Phys. J.*C73,2501(2013)]. Cited on pages 95 and 99.
- [79] G. Corcella and M. H. Seymour, “Initial state radiation in simulations of vector boson production at hadron colliders,” *Nucl. Phys.* **B565** (2000) 227–244, [arXiv:hep-ph/9908388 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9908388). Cited on page 100.
- [80] ROOT Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, “HistFactory: A tool for creating statistical models for use with RooFit and RooStats,” tech. rep., Jan, 2012. <https://cds.cern.ch/record/1456844>. Cited on page 101.
- [81] W. Verkerke and D. P. Kirkby, “The RooFit toolkit for data modeling,” *eConf* **C0303241** (2003) MOLT007, [arXiv:physics/0306116 \[physics\]](https://arxiv.org/abs/physics/0306116). [186(2003)]. Cited on page 101.
- [82] T. J. Hastie, R. J. Tibshirani, and J. H. Friedman, *The elements of statistical learning : data mining, inference, and prediction*. Springer series in statistics. Springer, New York, 2009. <http://opac.inria.fr/record=b1127878>. Autres impressions : 2011 (corr.), 2013 (7e corr.). Cited on page 110.
- [83] Y. Freund and R. E. Schapire, “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting,” *J. Comput. Syst. Sci.* **55** no. 1, (1997) 119–139. Cited on page 114.
- [84] J. H. Friedman, “Stochastic gradient boosting,” *Comput. Stat. Data Anal.* **38** (2002) 367–378. Cited on page 114.
- [85] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, “TMVA: Toolkit for Multivariate Data Analysis,” *PoS ACAT* (2007) 040, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039). Cited on pages 114, 117, 119, and 224.

List of Figures

2.1. Higgs potential $V(\phi) = \mu^2\phi^2 + \lambda^2\phi^4$ of a scalar field ϕ , and v is the vacuum expectation value.	12
2.2. $\Delta\chi^2 = \chi^2 - \chi_{min}^2$ vs. m_H curve. The line is the result of the fit to all electroweak precision data, with the preferred value at $m_H = 94_{-24}^{+29}$ GeV; the band represents an estimate of the theoretical error due to missing higher order corrections. The vertical band shows the 95% CL exclusion limit on m_H from the direct search. [16].	16
2.3. (a) Standard Model Higgs boson production total cross sections at $\sqrt{s} = 7$ TeV [22] and (b) branching ratios and their uncertainties [23].	17
2.4. (a) Observed and expected behaviour of the test statistic $2\ln Q$ as a function of the test mass m_H , for the LEP experiments. The full curve represents the observation; the dashed curve shows the median background expectation; the dark and light shaded bands represent the 68% and 95% probability bands. A lower limit is placed at $m_H = 114.4$ GeV [27]. (b) Observed and expected 95% C.L. upper limits on the ratios to the SM cross section, as functions of m_H for the combined Tevatron analyses. The limits are expressed as a multiple of the SM prediction, with the bands indicating the 68% and 95% probability regions. [26].	18
2.5. Invariant mass distributions for the combined $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples for (a) the di-photon candidates, and (b) the four-lepton candidates. The result of a fit to the data with the signal component fixed to $m_H = 126.5$ GeV and a background component described by a fourth-order Bernstein polynomial is superimposed in the di-photon analysis. The signal expectation for a SM Higgs with $m_H = 125$ GeV is shown in the four-lepton analysis. [22]	19

- 2.6. (a) The observed (solid) p_0 as a function of m_H in the low mass range. The dashed curve shows the expected p_0 under the hypothesis of a SM Higgs boson signal at that mass with its plus/minus one sigma band. The horizontal dashed lines indicate the p-values corresponding to significances of 1 to 6 sigma. (b) Measurements of the signal strength parameter μ for $m_H = 126$ GeV for the individual channels and their combination. [22] 20
- 2.7. The best-fit value for the signal strength μ in the individual channels and their combination for the full ATLAS data-sets at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The total $\pm 1\sigma$ uncertainty is indicated by the shaded green band, with the individual contributions from the statistical uncertainty (top, black), the experimental systematic uncertainty (middle, blue), and the theory uncertainty (bottom, red) on the signal cross section (from QCD scale, PDF, and branching ratios) shown by the error bars and printed in the central column. [2] 22
- 3.1. Cross-section and event rates (for a luminosity of $1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ for various processes in proton(anti)proton collisions, as a function of the centre-of-mass energy \sqrt{s} . The dashed lines represent the design energies of the Tevatron (1.8 TeV) and the LHC (14 TeV) [28]. This thesis uses data from the 2011 7 TeV run of the LHC. 24
- 3.2. Schematic of the CERN collider complex, including the LHC and the position of the major experiments. [32] 28
- 3.3. Cut-away of the ATLAS Inner Detector. [32] 31
- 3.4. Plan view of a quarter-section of the ATLAS inner detector showing each of the major detector elements with its active dimensions and envelopes. [28] 31
- 3.5. Cut-away view of the ATLAS calorimeter system. The overall dimensions of the system are approximately 8 m in diameter, and 12 m in length. [32] 33
- 3.6. (a) Sketch of a barrel module where the different layers are clearly visible with the ganging of electrodes in ϕ . The granularity in η and ϕ of the cells of each of the three layers and of the trigger towers (used in the L1 trigger as explained in Section 3.2.5). (b) Schematic showing how the mechanical assembly and the optical readout of the tile calorimeter are integrated together. The various components of the optical readout, namely the tiles, the fibres and the photomultipliers, are shown. [28] 34

3.7. (a) Schematic $R - \phi$ (left) and $R - z$ (right) views of the hadronic end-cap calorimeter. The semi-pointing layout of the readout electrodes is indicated by the dashed lines. Dimensions are in mm. (b) Schematic diagram showing the three FCal modules located in the end-cap cryostat. The material in front of the FCal and the shielding plug behind it are also shown. The black regions are structural parts of the cryostat. The diagram has an expanded vertical scale for clarity. [28]	35
3.8. Cut-away view of the ATLAS muon system. Total dimensions of the system are approximately 25 m in diameter, and 44 m in length. [32]	37
3.9. (a) Cross-section of the barrel muon system perpendicular to the beam axis (non-bending plane), showing three concentric cylindrical layers of eight large and eight small chambers. The outer diameter is about 20m. (b) Cross-section of the muon system in a plane containing the beam axis (bending plane). Infinite-momentum muons would propagate along straight trajectories which are illustrated by the dashed lines and typically traverse three muon stations. [28]	39
3.10. Block diagram of the L1 trigger. The overall L1 accept decision is made by the central trigger processor, taking input from calorimeter and muon trigger results. The paths to the detector front-ends, L2 trigger, and data acquisition system are shown from left to right in red, blue and black, respectively. [28] .	40
3.11. Architecture of the L1 calorimeter trigger. [28]	41
3.12. (a) Schema of the L1 muon barrel trigger. The RPCs are arranged in three stations of doublets: RPC1, RPC2, and RPC3, and (b) Schema of the L1 muon end-cap trigger. The TGCs are arranged in three stations of doublets: I, M2, and M3, and one station of triplets: M1. Also shown are the low- and high-transverse momenta roads for both systems. [28]	42
3.13. Cross-section diagram of the ATLAS detector, showing how different particles interact with the different sub-detectors [46].	45
4.1. Leading order Higgs production modes at the LHC; (a) gluon-gluon fusion (ggF), (b) vector boson fusion (VBF), and (c) associated production (VH , $V = W, Z$)	53
4.2. Example Feynman diagrams for $Z/\gamma^*(\rightarrow \ell\ell) + \text{jets}$ processes	55

4.3. Example Feynman diagrams for $W(\rightarrow \ell\nu) + \text{jets}$ processes.	56
4.4. Example Feynman diagrams for $t\bar{t}$ processes.	56
4.5. Example Feynman diagrams for single- t processes.	57
4.6. Kinematic variables for the selected objects, and the m_{MMC} distribution, at pre-selection level for the e channel. Uncertainties are statistical and systematic.	66
4.7. Kinematic variables for the selected objects, and the m_{MMC} distribution, at pre-selection level for the μ channel. Uncertainties are statistical and systematic.	67
4.8. Comparison of shapes for m_{MMC} , \cancel{E}_T and m_T distributions in OS and SS events from the QCD control region for the $e\tau_h$ (left) and $\mu\tau_h$ (right) channels. Contributions from electroweak and top backgrounds are subtracted from data and all distributions are normalised to the unit area. [1]	78
4.9. Tau and anti- τ_h m_{vis} distributions for the loose VBF W CR in (a) and (b), and the QCD CR in (c) and (d). Difference between MC and data is due to missing QCD multi-jet MC. Uncertainties are statistical and systematic.	86
4.10. Fake Factors (left axis) and r_W (right axis) values as a function of tau transverse energy.	89
4.11. m_{vis} distributions for (a) the W control region using $f_{W+\text{jets}}$, (b) the QCD control region using F_{QCD} , and (c) the loose VBF region using f_{MIX} . Uncertainties are statistical and systematic. Plot (d) shows the loose VBF region selection using the OS-rSS method. All other background contributions are included in the <i>others</i> category.	90
4.12. m_{MMC} pre-fit distributions in the signal region for each analysis category. $Z \rightarrow \ell\ell$, $t\bar{t}$, di-boson, and $W(\rightarrow \ell\nu) + \text{jets}$ backgrounds are grouped into the <i>others</i> category. Errors include statistical errors and all systematic uncertainties. Bins are merged according to Algorithm 4.1, with bin contents scaled to number of events for every 10 GeV.	102

4.13. Pull distributions for the (a) VBF and (b) Boosted categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$	104
4.14. Pull distributions for the (a) 1-Jet and (b) 0-Jet categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$	105
4.15. Pull distributions for the (a) combined and (b) non-jet combined categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$	106
4.16. Cut-based exclusion limits for VBF, Boosted, 1-Jet, and 0-Jet analysis categories. Uncertainties are statistical and systematic.	107
4.17. Cut-based (a) exclusion limits and (b) signal significance for VBF, Boosted, 1-Jet, and 0-Jet analysis categories combined. Uncertainties are statistical and systematic.	108
4.18. (a) Comparison of cut-based exclusion limits for each individual analysis category, and the combined limit, for $m_H = 125$ GeV. Uncertainties are statistical and systematic. (b) Extracted signal strength μ at $m_H = 125$ GeV for each analysis category, and the combined signal strength. Solid line is $\mu = 0$ and dashed line is $\mu = 1$	109
5.1. Example of a simple decision tree, separating signal S from background B.	112
5.2. $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ category (a) without and (b) with analysis level cuts. Uncertainties are statistical and systematic.	122

5.3. Tau and lepton kinematic variables fed into BDT_{HBD} . Uncertainties are statistical and systematic.	124
5.4. Jet kinematic variables fed into BDT_{HBD} . Uncertainties are statistical and systematic.	125
5.5. (a) BDT score and (b) ROC curve outputs from TMVA.	126
5.6. BDT_{HBD} after (a) $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ categorisation and (b) after $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ and a side-band selection of $m_{\text{MMC}} < 100$ GeV or $m_{\text{MMC}} > 150$ GeV. $H \rightarrow \tau\tau$ signal is $\times 100$. Uncertainties are statistical and systematic.	127
5.7. Exclusion as a function of BDT_{HBD} with $m_H = 125$ GeV. Blue line indicates final cut value chosen. Uncertainties are statistical and systematic.	127
5.8. Correlations between the tau system variables and BDT_{HBD} . Blue data points are VBF Higgs signal for every 5 GeV mass point within $m_H = [100, 150]$ GeV, red data points are ALPGEN VBF filtered $Z \rightarrow \tau\tau$ MC, and the green lines represent the positions of the cuts made in the cut-based analysis, and the final BDT_{HBD} cut.	129
5.9. Correlations between the jet variables and BDT_{HBD} . Blue data points are VBF Higgs signal for every 5 GeV mass point within $m_H = [100, 150]$ GeV, red data points are ALPGEN VBF filtered $Z \rightarrow \tau\tau$ MC, and the green lines represent the positions of the cuts made in the cut-based analysis, and the final BDT_{HBD} cut.	130
5.10. Pull distribution for the combined categories in the hybrid analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$	131
5.11. m_{MMC} pre-fit distributions in the signal region for each analysis category. $Z \rightarrow \ell\ell$, $t\bar{t}$, di-boson, and $W(\rightarrow \ell\nu) + \text{jets}$ backgrounds are grouped into the <i>others</i> category. Uncertainties are statistical and systematic. Bins are merged according to Algorithm 4.1, with bin contents scaled to number of events for every 10 GeV.	135
5.12. Hybrid exclusion limits for VBF, Boosted, and 1-Jet analysis categories. Uncertainties are statistical and systematic.	136

5.13. Hybrid (a) exclusion limits and (b) signal significance for VBF, Boosted, 1-Jet, and 0-Jet analysis categories combined. Uncertainties are statistical and systematic.	137
5.14. (a) Comparison of hybrid exclusion limits for each individual analysis category, and the combined limit, for $m_H = 125$ GeV. Uncertainties are statistical and systematic. (b) Extracted signal strength μ at $m_H = 125$ GeV for each analysis category, and the combined signal strength. Solid line is $\mu = 0$ and dashed line is $\mu = 1$	138
6.1. m_{MMC} distributions for (a) the $W(\rightarrow \ell\nu) + \text{jets}$ control region at the pre-selection level after application of $k^{W,PS}$ using ALPGEN $Z \rightarrow \tau\tau$, and (b) the Embedding control region at the pre-selection level after application of k^{emb} . Uncertainties are statistical and systematic.	142
6.2. Fake-factors and r_W values as a function of τ energy for the MVA Boosted category.	143
6.3. m_{vis} distributions using the fake-factor method for the Boosted (a) W CR, (b) QCD CR, (c) SR, and (d) the SR using the OS-rSS method. The fake-factor for the W CR uses only f_W and the QCD CR uses only f_{QCD} , while the SR uses f_{MIX} . Uncertainties are statistical and systematic.	144
6.4. Blinded m_{MMC} distributions after full SR cuts for the MVA Boosted category using (a) the fake-factor method and (b) the OS-rSS method, after running the rebinning procedure (Algorithm 4.1). Uncertainties are statistical and systematic.	145
6.5. Background shapes in the VBF category for full fake-factor method compared to using AT events only for the fake- τ background contribution, for all variables being fed into BDT_{VBF} . Uncertainties are statistical only.	147
6.6. Background shapes in the VBF category for full fake-factor method compared to using AT events only for the fake- τ background contribution, for all variables being fed into BDT_{VBF} . Uncertainties are statistical only.	148
6.7. Background shapes in the 1-Jet category for full OS-rSS method compared to using OS events only for the fake- τ background contribution, for variables being fed into BDT_{1J} . Uncertainties are statistical only.	149

6.8. Background shapes in the 1-Jet category for full OS-rSS method compared to using OS events only for the fake- τ background contribution, for variables being fed into BDT_{1J} . Uncertainties are statistical only.	150
6.9. (a) and (b) show the TMVA training and testing output for BDT_{VBF} for the Adaboost and Gradient Boost algorithms, before training parameter optimisation. The remaining figures show the scans for the training parameters in Section 6.3.2.	152
6.10. Input variables for BDT_{VBF} after the base VBF_{MVA} selection. Uncertainties are statistical and systematic.	158
6.11. Input variables for BDT_{VBF} after the base VBF_{MVA} selection. Uncertainties are statistical and systematic.	159
6.12. Linear correlation matrices for (a) signal and (b) background events after training for BDT_{VBF} . Bins show level of correlation as a percentage. The variables compared are described in Sections 4.6 and 6.3.3. (c) and (d) show the BDT Score and ROC curves as outputted from TMVA for BDT_{VBF}	160
6.13. BDT distributions for the MVA analysis in the low-BDT control region. Data is blinded for BDT scores above a signal efficiency of 30%. Uncertainties are statistical and systematic.	161
6.14. BDT distributions for the MVA analysis in the high-BDT control region. Cuts are placed on the most important variable for each BDT from TMVA such that the signal efficiency is below 30%; $C_{\eta_{j_1}\eta_{j_2}}(\eta_\ell) < 0.7$ for VBF, $\Delta R(\tau, \ell) > 1.8$ for Boosted, $\Sigma\Delta\phi > 3.0$ for 1-Jet, and $\Sigma\Delta\phi > 3.2$ for 0-Jet. Uncertainties are statistical and systematic.	162
6.15. BDT distributions for the MVA analysis in the $t\bar{t}$ control region. Uncertainties are statistical and systematic. Note that due to the requirement of no jets in the 0-Jet category, there is no $t\bar{t}$ control region available.	163
6.16. BDT distributions for the MVA analysis in the $W(\rightarrow \ell\nu) + \text{jets}$ control region. Uncertainties are statistical and systematic.	164
6.17. BDT distributions for the MVA analysis in the $Z \rightarrow \tau\tau$ control region. Uncertainties are statistical and systematic.	165

6.18. Exclusion as a function of BDT score for the four analysis categories, with $m_H = 125$ GeV. Blue line indicates final cut values chosen. Errors include statistical and systematic uncertainties.	167
6.19. Correlations between the tau system variables, and BDT_{VBF} . Blue data points are VBF Higgs signal at $m_H = 125$ GeV, and red data points are all backgrounds, estimated as per Section 6.2.3.	169
6.20. Correlations between the jet variables, tau system mass variables, and BDT_{VBF} . Blue data points are VBF Higgs signal at $m_H = 125$ GeV, and red data points are all backgrounds, estimated as per Section 6.2.3.	170
6.21. Pull distribution for the combined categories in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$	171
6.22. m_{MMC} pre-fit distributions in the signal region for the VBF, Boosted, and 1-Jet analysis categories. $Z \rightarrow \ell\ell$, $t\bar{t}$, di-boson, and $W(\rightarrow \ell\nu) + \text{jets}$ backgrounds are grouped into the <i>others</i> category. Uncertainties are statistical and systematic. Bins are merged according to Algorithm 4.1, with bin contents scaled to number of events for every 10 GeV.	174
6.23. m_{MMC} pre-fit distributions in the signal region for the 0-Jet and Rest analysis categories. $Z \rightarrow \ell\ell$, $t\bar{t}$, di-boson, and $W(\rightarrow \ell\nu) + \text{jets}$ backgrounds are grouped into the <i>others</i> category. Uncertainties are statistical and systematic. Bins are merged according to Algorithm 4.1, with bin contents scaled to number of events for every 10 GeV.	175
6.24. Full-MVA observed and expected 95% CL. exclusion limits for the VBF and Boosted analysis categories, with 1σ and 2σ error bands. Uncertainties are statistical and systematic.	176
6.25. Full-MVA observed and expected 95% CL. exclusion limits for the 1-Jet, 0-Jet, and Rest analysis categories, with 1σ and 2σ error bands. Uncertainties are statistical and systematic.	177

6.26. Results from the combined analysis categories showing (a) the observed and expected 95% CL. exclusion limits with 1σ and 2σ error bands, and (b) the signal significance. Uncertainties are statistical and systematic.	178
6.27. (a) Comparison of full-MVA exclusion limits for each individual analysis category, and the combined limit, for $m_H = 125$ GeV. Uncertainties are statistical and systematic. (b) Extracted signal strength μ at $m_H = 125$ GeV for each analysis category, and the combined signal strength. Uncertainties are statistical and systematic.	179
6.28. m_{MMC} distributions in the VBF category with signal region selections for the three analysis strategies. Uncertainties are statistical and systematic.	180
6.29. Comparisons of the (a) 95% confidence level exclusion limits and (b) the extracted signal strength $\hat{\mu}$, at $m_H = 125$ GeV for the three analysis approaches. A comparison of the statistical significance of the results across the whole tested m_H range is shown in (c). Uncertainties are statistical and systematic.	182
A.1. Comparisons of the (a) 95% confidence level exclusion limits and (b) the extracted signal strength $\hat{\mu}$, at $m_H = 125$ GeV for the three analysis approaches using only the VBF and Boosted categories in the combined fit. A comparison of the statistical significance of the results across the whole tested m_H range is shown in (c). Uncertainties are statistical and systematic.	186
B.1. Pull distributions for the (a) VBF and (b) Boosted categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$	188
B.2. Pull distributions for the (a) 1-Jet and (b) 0-Jet categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$	189

- B.3. Pull distributions for the (a) combined and (b) non-jet combined categories in the cut-based analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$ 190
- B.4. Pull distributions for the (a) VBF and (b) Boosted categories in the hybrid analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$ 191
- B.5. Pull distributions for the (a) 1-Jet and (b) 0-Jet categories in the hybrid analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$ 192
- B.6. Pull distributions for the (a) combined and (b) non-jet combined categories in the hybrid analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$ 193
- B.7. Pull distributions for the (a) VBF and (b) Boosted categories in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$ 194
- B.8. Pull distributions for the (a) 1-Jet and (b) 0-Jet categories in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$ 195

-
- B.9. Pull distribution for the Rest category in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$ 196
- B.10. Pull distributions for the (a) combined and (b) non-jet combined categories in the MVA analysis. Black points show the fitted value of the nuisance parameter $\hat{\theta}$ with $\pm 1\sigma$ error bars, with respect to the nominal value θ_0 . Blue bands show the post-fit impact of the nuisance parameter on the fitted signal strength $\hat{\mu}$ by varying up (hatched) or down (open) the post-fit nuisance parameter value by $\pm 1\sigma$ 197

List of Tables

2.1. Particle Content of the Standard Model [5]. Errors are statistical and systematic when given. Masses of the neutrinos are inferred to be very small, but non-zero.	5
3.1. LHC machine parameters for both design, and running as of 2011. https://acc-stats.web.cern.ch/acc-stats/	26
3.2. General performance goals of the ATLAS detector. Note that, for high- p_T muons, the muon-spectrometer performance is independent of the inner-detector system. [33]	38
3.3. Dominant decay modes and branching fractions for tau leptons [5].	49
4.1. Cross-sections (σ) times branching fraction (\mathcal{B}) of the various background and signal processes studied.	54
4.2. Run, trigger information, and integrated luminosities for data used. The terminology used here is explained in the text.	60
4.3. Object selection criteria	61
4.4. Number of events at pre-selection level for the $e\tau_h$ and $\mu\tau_h$ channels. The various backgrounds are estimated using the methods described in Section 4.7, with SS-Data being a data driven estimate of the multi-jet contribution. Uncertainties are statistical only.	65

4.5. Category and analysis cuts for the cut-based analysis. $e\tau_h$ and $\mu\tau_h$ channels are combined for VBF and Boosted categories, and treated separately for the 1- and 0-jet categories. The Loose VBF category is used in the fake-factor method (described in Section 4.7.6). $t\bar{t}$, W , $Z \rightarrow \tau\tau$ and QCD refer to control regions applied on top of the category selection for use in calculating k -factors (described in Section 4.7.1). Any cuts in the control regions supersede those in the categorisation.	73
4.6. Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to VBF filtered MC for the VBF category, and the embedding sample for all others. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method, and Fake- τ is the estimate of fakes from multi-jets and $W(\rightarrow \ell\nu) + \text{jets}$ contributions using the fake factor method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.	74
4.7. Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to the embedding sample. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.	75
4.8. k -factors for each analysis category and each lepton channel. Unless explicitly stated, factors apply to both OS and SS events. Uncertainties are statistical only.	80
4.9. Selection regions for the FF calculation.	85
4.10. Number of events in each background and region required for the calculation of r_W , binned by E_T^τ . Uncertainties are statistical only.	87
4.11. Number of tau and anti- τ candidates in each FF control region, and the final $f_{W/QCD/MIX}$ numbers for each E_T^τ bin. Uncertainties are statistical only except for the f_{MIX} values which have a systematic uncertainty based on varying the r_W value up and down.	88
4.12. Higgs branching ratio (\mathcal{B}) uncertainty variation (%) for each simulated Higgs mass point [68].	91

4.13. QCDscale uncertainty variations (%) for the VBF (QCDscale_qqH), ggF (QCDscale_ggH), and VH (QCDscale_ZH and QCDscale_WH) signal topologies, for each simulated Higgs mass point [69].	92
4.14. Statistical and systematic uncertainties as on the nominal event yields for the VBF category, ordered by the size of their influence on the fitted signal strength. NOM, STAT, and SYST refer to the nominal yield, statistical uncertainty, and total systematic uncertainty respectively.	96
4.15. Statistical and systematic uncertainties as on the nominal event yields for the Boosted category, ordered by the size of their influence on the fitted signal strength. NOM, STAT, and SYST refer to the nominal yield, statistical uncertainty, and total systematic uncertainty respectively.	97
5.1. Selection cuts for Hybrid categories.	120
5.2. Yields for the $\text{VBF}_{\text{HBD}}^{\text{pre-BDT}}$ category before and after signal region (SR) analysis level cuts, as well as the percent reduction in yield. Uncertainties are statistical only.	122
5.3. Number of events after hybrid categorisation and analysis level cuts for the VBF and Boosted categories. The various backgrounds are estimated using the methods described in Section 4.7. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.	132
5.4. Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to the embedding sample. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.	133
6.1. Category and analysis cuts for the MVA analysis. SR cuts were not placed when training the category-specific BDTs. $t\bar{t}$, W , $Z \rightarrow \tau\tau$, QCD refer to control regions applied on top of the category selection for use in calculating k -factors and fake-factors (for the VBF and Boosted categories only). Any cuts in the CRs supersede those in the categorisation.	140
6.2. k -factors for each analysis category and each lepton channel. Unless explicitly stated, factors apply to both OS and SS events.	141

6.3. TMVA training parameters for BDTs grown for the hybrid VBF category, and the full MVA analysis categories. Parameters are described in Section 6.3.2.	154
6.4. Input variables for the Hybrid BDT, and the BDTs used for the full-MVA analysis.	156
6.5. Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to VBF filtered MC for the VBF category, and the embedding sample for all others. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method, and Fake- τ is the estimate of fakes from multi-jets and $W(\rightarrow \ell\nu) + \text{jets}$ contributions using the fake-factor method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.	172
6.6. Number of events after categorisation and analysis level cuts. The various backgrounds are estimated using the methods described in Section 4.7. $Z \rightarrow \tau\tau$ refers to the embedding sample. SS-Data is a data driven estimate of the multi-jet contribution using the OS-rSS method. Quoted uncertainties are statistical and systematic respectively, as calculated in Section 4.8.	173

List of Algorithms

- 4.1. Bin merging algorithm. 100
- 5.1. AdaBoost TMVA implementation [85] 117
- 5.2. Gradient Boost TMVA implementation [85] 119