A search for the Standard Model Higgs boson via its decay to tau leptons and W bosons at the ATLAS detector

Christopher Boddy

Brasenose College, Oxford

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

University of Oxford, Trinity Term 2012

A search for the Standard Model Higgs boson via its decay to tau leptons and W bosons at the ATLAS detector

Christopher Boddy
Brasenose College, Oxford

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

May 2012

Abstract

Understanding the origin or Electroweak symmetry breaking within the Standard Model was a key motivation for the construction of the Large Hadron Collider (LHC) experiment at CERN. This thesis presents a search for evidence of Higgs boson production in the 4.7 fb^{-1} of collision data recorded at a centre-of-mass energy of 7 TeV at the ATLAS detector during 2011.

This search is focused on signal events in which a Higgs boson is produced in the mass range $100 < m_H < 180$ GeV/c^2 and subsequently decays to a pair of W bosons or a pair of tau leptons to final states with one hadronically decaying tau lepton and one light lepton. After an event selection criteria has been applied, the number of events in this data sample is consistent with the total background estimate and an upper limit is placed on the SM Higgs boson production rate at 95% confidence level. In addition, the prospects for measuring the SM Higgs coupling strength to tau leptons with the associated Higgs production channels and the full LHC dataset are also presented.

Dedicated to Sarah and my parents.

Acknowledgements

Firstly I would like to thank my supervisors Sinead Farrington and Chris Hays for their support and seemingly in-exhaustible patience, particularly when I asked lots of silly questions. The past three years were made very pleasant by the past and present members of the Oxford ATLAS group for their help and useful suggestions; especially to James Ferrando and Hugo Beauchemin for sharing their expertise in the beginning and latterly to Elias Coniavitis for all of the invaluable conversations about the analyses we've worked on together.

I also owe a debt to the HSG4 group for their useful comments and suggestions, in particular to Sasha Prankho, Jed Biesiada, Elisabetta Pianori and Dimitris Varouchas.

To Alan Barr I also give credit, in particular for an insignificant conversation over lunch one day that quickly lead to a grant proposal and a successful outreach project for the Science and Technology Facilities Council with guidance from the ATLAS outreach group[1].

I'd also like to thank the other students of the Oxford ATLAS group[2], in particular Alex Pinder, Dan Short, Sam Whitehead, Caterina Doglioni, Phil Jones, Gemma Wooden, Aimee Larner and James Buchanan for all the laughter, and soothing words[3] when things went wrong, particularly in the first year.

[1] Michael Barnett, Steven Goldfarb and Pete Watkins in particular.

[2] and students from other experiments whom we adopted.

[3] and often top notch profanity.

Contents

1 Introduction **1**

2 Theoretical overview **3**
2.1 Introduction . . . 3
2.2 The Standard Model . . . 3
2.2.1 Overview . . . 3
2.2.2 Particles of the Standard Model . . . 4
2.2.3 Electroweak theory . . . 5
2.2.4 QCD . . . 7
2.3 The Higgs mechanism . . . 8
2.3.1 Spontaneous symmetry breaking . . . 9
2.3.2 Fermion masses . . . 10
2.4 Higgs boson phenomenology . . . 11
2.4.1 Constraints on the Higgs mass . . . 11
2.4.2 Production and decay at the LHC . . . 12
2.4.3 Measuring the SM Higgs boson coupling strength . . . 16

3 The ATLAS detector at the LHC **18**
3.1 Introduction . . . 18
3.1.1 The Large Hadron Collider . . . 18
3.1.2 ATLAS coordinate system . . . 20
3.1.3 Detector requirements and specifications . . . 21
3.2 Magnets . . . 24
3.3 Inner detector . . . 24
3.3.1 Pixel detector . . . 24
3.3.2 Semiconductor tracker . . . 25
3.3.3 Transition radiation tracker . . . 26
3.4 Calorimetry . . . 27
3.4.1 Electromagnetic calorimeters . . . 27
3.4.2 Hadronic calorimeters . . . 29
3.5 Muon system . . . 30
3.6 Trigger . . . 31
3.6.1 The L1 trigger . . . 32
3.6.2 The HLT . . . 33
3.6.3 Data distribution . . . 33
3.7 Event reconstruction . . . 34

3.7.1 Inner detector track and vertex reconstruction . . . 35
3.7.2 Jet finding and heavy flavour tagging . . . 36
3.7.3 b-jet tagging . . . 38
3.7.4 Electrons and photons . . . 38
3.7.5 Muons . . . 40
3.7.6 Hadronically decaying tau leptons . . . 42
3.7.7 Missing transverse momentum . . . 42
3.8 Detector simulation . . . 43

4 LHSee 45
4.1 Introduction . . . 45
4.2 Visualising ATLAS events on a mobile platform . . . 45
4.3 3-D Techniques . . . 46
4.4 Visualising the detectors of the ATLAS experiment . . . 46
4.5 Event Visualisation . . . 52
4.5.1 Tracks . . . 52
4.5.2 Calorimeter histograms . . . 55
4.6 Conclusion . . . 55

5 Expected Standard Model Higgs boson coupling measurements using the $H \to \tau\tau$ channel 56
5.1 Introduction . . . 56
5.2 Event generation and simulation . . . 57
5.2.1 Parametric detector simulation . . . 58
5.2.2 Fake-rate application . . . 59
5.2.3 Triggering . . . 60
5.2.4 Pile-up . . . 60
5.3 Event selection and sensitivity . . . 60
5.3.1 WH analysis . . . 61
5.3.2 ZH analysis . . . 65
5.3.3 $t\bar{t}H$ analysis . . . 68
5.4 Sensitivity estimation . . . 71
5.4.1 Signal yield extraction . . . 71
5.5 Expected sensitivity to SM Higgs coupling ratios . . . 71
5.6 Conclusion . . . 73

6 Tau identification and mis-identification probability 74
6.1 Introduction . . . 74
6.2 Physics with tau leptons . . . 74
6.3 Hadronic tau reconstruction and identification in ATLAS . . . 76
6.3.1 Reconstruction . . . 76
6.4 Tau mis-identification probability . . . 79
6.4.1 Introduction . . . 79
6.4.2 Measuring the τ_h fake-rate from data . . . 80
6.4.3 Event selection . . . 81

6.4.4 Systematic uncertainties 86
6.4.5 Results 88
6.4.6 Summary 88

7 Search for the Standard Model Higgs boson in the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ channels **91**
7.1 Introduction 91
7.1.1 Signal and background processes 92
7.2 Datasets 94
7.2.1 Data 94
7.2.2 Simulation 95
7.3 Di-tau mass reconstruction in $H \to \tau\tau$ events 95
7.3.1 Visible mass 96
7.3.2 Collinear mass 96
7.3.3 Missing Mass Calculator 98
7.4 Common event selection criteria 101
7.4.1 Trigger 101
7.4.2 Pileup re-weighting 101
7.4.3 Muon selection 101
7.4.4 Electron selection 102
7.4.5 τ_h selection 102
7.4.6 Hadronic jet selection 103
7.4.7 Overlap removal 103
7.5 Optimisation of the $H \to \tau\tau$ event selection criteria 103
7.5.1 $H \to \tau\tau$ event selection criteria 106
7.5.2 VBF event selection 107
7.5.3 $H \to \tau\tau$ results 107
7.6 Optimisation of the $H \to WW^* \to \ell\nu\tau\nu$ event selection criteria 109
7.6.1 $H \to WW^* \to \ell\nu\tau\nu$ event selection cuts 109
7.7 Data-driven estimate of the $W(\to \ell\nu)$ +jets background 110
7.7.1 Introduction 110
7.7.2 $W(\to \ell\nu)$ +jets in the $H \to \tau\tau$ analysis 111
7.7.3 $W(\to \ell\nu)$ +jets in the $H \to WW^* \to \ell\nu\tau\nu$ analysis 112
7.7.4 Systematic uncertainty on the $W(\to \ell\nu)$ +jets data-driven background method 113
7.8 Data-driven estimate of the QCD multi-jet background 114
7.8.1 Introduction 114
7.8.2 QCD background in the $H \to \tau\tau$ analysis 115
7.8.3 QCD background in the $H \to WW^* \to \ell\nu\tau\nu$ analysis 115
7.8.4 Systematic uncertainty on the QCD multi-jet data-driven background method 116
7.9 Systematic uncertainties 116
7.9.1 Theory uncertainties 116
7.9.2 Trigger efficiencies 117
7.9.3 Electron candidates 118

7.9.4 Muon candidates . . . 118
7.9.5 τ_h candidates . . . 119
7.9.6 Jet candidates . . . 119
7.9.7 $p_{\mathrm{T}}^{\mathrm{miss}}$ systematic uncertainties . . . 119
7.9.8 Data driven background estimation systematic uncertainties . . . 120
7.9.9 Systematic uncertainties in the $H \to \tau\tau$ analysis . . . 120
7.9.10 Systematic uncertainties in the $H \to WW^* \to \ell\nu\tau\nu$ analysis . . . 120
7.10 Limit Setting . . . 120
7.10.1 Limit . . . 125

8 Conclusions and outlook **133**

A $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ analysis appendix **135**
A.1 Simulated Datasets . . . 135

References **140**

List of Figures

2.1 Illustration of the Higgs potential $V(\phi)$ of the complex scalar field $\phi = \phi_1 + i\phi_2$ for $\lambda > 0$ and $\mu^2 > 0$. 9
2.2 A χ^2 fit using indirect electroweak measurments for a range of Higgs boson masses. 12
2.3 Combined upper limits on the SM Higgs boson production cross section from measurements made at the Tevatron experiments. 13
2.4 Dominant Higgs boson production mechanisms at the LHC. 14
2.5 Mass dependence of (a) the SM Higgs production cross section, (b) the SM Higgs branching ratios, and (c) the total SM Higgs boson width. 15

3.1 The CERN accelerator complex. 19
3.2 Pseudo-rapidity η and polar angle θ in the ATLAS coordinate system. 20
3.3 Diagram showing the situation, scale and relative positions of the main ATLAS sub-detectors. 22
3.4 Cross section for various physical processes for a hadron collider operating at an instantaneous luminosity of $\mathcal{L} = 10^{33}$ cm^{-2}s^{-1}. 23
3.5 Cross-sectional diagram showing the ATLAS inner detector. 25
3.6 Calorimeters in the ATLAS experiment. 28
3.7 An ATLAS electromagnetic calorimeter barrel module. 29
3.8 The ATLAS Muon Spectrometer detectors in the $R - z$ plane. 30
3.9 Diagram showing the ATLAS trigger decision chain and data-flow model. . . 32
3.10 Problems that can arise during jet finding: (a) Infrared safety: soft particle (gluon) emission should not affect the result of the algorithm, (b) Collinear safety: one parton splitting into two should not affect the result. 37

4.1 Primitive graphs used to construct detector geometry: a Cuboid, an Annulus, a Toroid and a Cylinder. 46
4.2 LHSee projection of the inner detector in the $z - y$ plane. 47
4.3 LHSee projection of the inner detector in the $x - y$ plane. 48
4.4 LHSee 3-D visualisation of the inner detector Pixel barrel section. 48
4.5 LHSee longitudinal plane projection showing the ATLAS calorimeters. 49
4.6 LHSee transverse plane projection showing the ATLAS calorimeters. 49
4.7 LHSee 3-D visualisation of the ATLAS calorimeters. 50
4.8 LHSee transverse plane projection showing all detectors overlaid with a simulated $W \to \mu\nu$ event. 50
4.9 LHSee 3-D visualisation of the ATLAS muon detectors with a simulated $Z \to \mu^+\mu^-$ event. 51

4.10 3-D visualisation of the inner detector tracks. Only sub-detectors with $x < 0$ are shown. . . . 51
4.11 Projection of an inner detector track helix in the $x - y$ plane. . . . 53

5.1 The $l_W \tau_l \tau_h$ $p_{\mathrm{T}}^{\mathrm{miss}}$ and m_T distributions after all other cuts are applied. . . . 63
5.2 The $l_W \tau_l \tau_h$ $p_{\mathrm{T}}^{\mathrm{miss}}$ visible mass distribution of signal and background events passing the full event selection. . . . 64
5.3 The signal and background collinear mass distributions after the full ZH event selection assuming an intergrated luminosity of 100 fb^{-1}. . . . 67
5.4 The opposite-sign, same-flavour light-lepton invariant mass distribution of $t\bar{t}H$ $\tau_l \tau_h$ events pasing all other selection criteria. . . . 69
5.5 The visible mass distributions after the full $t\bar{t}H$ event selection for Case iii (a) and Case iv (b) events. . . . 70
5.6 The expected relative statistical uncertainties on $\sigma \times$ BR of VH (left, $V = W, Z$) and $t\bar{t}H$ (right) production for the nominal (top) and optimistic (bottom) tau identification performance scenarios. . . . 72
5.7 The expected relative statistical uncertainties on ratios of partial widths using measurements of associated Higgs boson production and decays to tau-lepton or bottom-quark pairs. Each partial width Γ_i corresponds to the trilinear interaction of a Higgs boson to another particle i. Shown are the nominal (left) and optimistic (right) detector performance scenarios. . . . 73

6.1 Tau lepton decays. . . . 75
6.2 $W(\to \ell\nu)$ +jets diagrams. . . . 79
6.3 γ-jet diagrams. . . . 81
6.4 Tau identification variables for reconstructed τ_h candidates after applying an event selection. . . . 83
6.5 Hadronic τ mis-identification rate in events with different vertex multiplicities. 84
6.6 Fraction of probe jets initiated by quarks as a function of the jet transverse momentum in Pythia γ+jet, Z+jet and QCD di-jet MC samples. . . . 85
6.7 Hadronic τ mis-identification probability, as a function of probe candidate p_{T}, for quark-initiated and gluon-initiated probe τ_h candidates. . . . 85
6.8 Hadronic τ mis-identification rates for events with an isolated photon candidate. 87
6.9 Affect on hadronic τ mis-identication rate of selection criteria variation. . . . 88
6.10 The hadronic τ mis-identification probability of hadronic jets using γ+jet topologies. . . . 89

7.1 SM Higgs decay via WW into final states with a τ lepton. . . . 92
7.2 The effect of spin correlations on the angular separation of the tau and light lepton decay products in $H \to WW^* \to \ell\nu\tau\nu$ events. . . . 93
7.3 Signal and background process visible mass distributions in the $H \to \tau\tau$ analysis. . . . 96
7.4 Signal and background process collinear-mass distributions in the $H \to \tau\tau$ analysis. . . . 98
7.5 Missing Mass Calculator probability distribution functions for original τ lepton momentum ($45 < p_\tau \leq 50$ GeV). . . . 99

7.6 Missing transverse momentum distributions of signal and background processes in the $H \to \tau\tau$ analysis. 104
7.7 Transverse mass distributions of signal and background processes in the $H \to \tau\tau$ analysis. 105
7.8 Exclusive n-jet distribution of events in the $H \to \tau\tau$ analysis. 106
7.9 Di-jet variable distributions used to further select VBF events passing all other selection cuts with two additional jets with $p_{\mathrm{T}} > 25$ GeV. The signal histograms from ggF and VBF SM Higgs production show the expected yield for 100 times the integrated luminosity of the analysis. 126
7.10 MMC mass distributions after the full $H \to \tau\tau$ events selection for events with no additional jets. 127
7.11 MMC mass distributions after the full $H \to \tau\tau$ events selection for events with at least one additional jet. 128
7.12 $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the $H \to WW^* \to \ell\nu\tau\nu$ signal and background distributions. 129
7.13 The $\Delta R(\ell, \tau_h)$ distribution of $H \to WW^* \to \ell\nu\tau\nu$ signal and background processes. 129
7.14 Visible mass distribution of the $H \to WW^* \to \ell\nu\tau\nu$ signal and background after the full selection criteria have been applied. 130
7.15 The $\Delta\phi(\ell, \tau_h)$ distribution of $H \to WW^* \to \ell\nu\tau\nu$ signal and background processes. 130
7.16 Variation of light lepton isolation on the QCD multijet scale factor. 131
7.17 The expected and observed 95% confidence-level upper limit on the SM Higgs production cross section in the $H \to \tau\tau$ analysis. 131
7.18 The expected and observed 95% confidence-level upper limit on the SM Higgs production cross section in the $H \to WW^* \to \ell\nu\tau\nu$ analysis. 132

List of Tables

2.1 Standard Model fermions and their associated quantum numbers. 5
2.2 Standard Model bosons with their electric charge, mass and the particles with which they interact. 5

3.1 ATLAS inner detector track parameter resolution. 36
3.2 Reconstructed electron cluster window sizes. 39

5.1 The calorimeter resolution parameters defined in Equation 5.3. 58
5.2 The acceptance criteria of reconstructed objects. 59
5.3 Object identification and mis-identification rates. 59
5.4 Triggering efficiencies. 60
5.5 WH analysis signal and background cross sections at $\sqrt{s} = 14$ TeV. 62
5.6 The numbers of WH signal and background events passing each set of requirements, for an integrated luminosity of 100 fb^{-1} and $m_H = 125$ GeV. Also shown is the signal over the square root of background, a measure of the statistical sensitivity to the signal. Additional sensitivity is gained from a fit to the visible mass distribution. 64
5.7 The contribution of each background to the $l_W \tau_l \tau_h E_{\mathrm{T}}^{\mathrm{miss}}$ final state for an integrated luminosity of 100 fb^{-1}. 64
5.8 The number of WH signal events for m_H in the range 115-135 GeV, and the statistical significance of the excess of signal events over background in 100 fb^{-1} of integrated luminosity. 64
5.9 $Z(\to \ell\ell)H(\to \tau\tau)$ final states. 65
5.10 ZH analysis signal and background cross sections at $\sqrt{s} = 14$ TeV. 66
5.11 The numbers of ZH signal and background events passing each set of requirements, for an integrated luminosity of 100 fb^{-1} and Higgs boson mass of 125 GeV. Also shown is the signal over the square root of background, a measure of the statistical sensitivity to the signal. Additional sensitivity is gained from a fit to the collinear mass distribution. 67
5.12 The number of ZH signal events for m_H in the range 115-135 GeV, and the statistical significance of the excess of signal events over background in 100 fb^{-1} of integrated luminosity. 67
5.13 The contribution of each background to the ZH final state for an integrated luminosity of 100 fb^{-1}. 67
5.14 $t\bar{t}(\to q\bar{q}'b\bar{b}l^{\pm}\nu)H(\to \tau^{\pm}\tau^{\mp})$ final states. 68
5.15 $t\bar{t}H$ analysis signal and background cross sections at $\sqrt{s} = 14$ TeV. 69

5.16 The number of ttH signal events in each channel for m_H in the range 115-135 GeV, and the statistical significance of the excess of signal events over background in 100 fb^{-1} of integrated luminosity. The $t\bar{t}$ pair is selected in the $l_W \nu q\bar{q}b\bar{b}$ final state. 70
5.17 The contribution of each background to the $t\bar{t}H$ final states for an integrated luminosity of 100 fb^{-1}. 70

6.1 Tau decay modes reconstructed at ATLAS and their associated branching fractions. 75
6.2 γ-jet cut-flow for events that fulfil the trigger and data-quality requirements for $\int \mathcal{L}\mathrm{dt} \simeq 34$ pb^{-1}. 82

7.1 The total number of signal and background events in each $H \rightarrow \tau\tau$ signal region. 108
7.2 The total number of signal and background events in each $H \rightarrow WW^* \rightarrow \ell\nu\tau\nu$ signal region. 110
7.3 $W(\rightarrow \ell\nu)$ +jets scale factors. 112
7.4 Exclusive $W(\rightarrow \ell\nu)$ +jets scale factors for events with an oppositely charged $\ell - \tau_h$ pair (OS) and a similarly charged $\ell - \tau_h$ pair (SS) in each signal region of the $H \rightarrow WW^* \rightarrow \ell\nu\tau\nu$ analysis. 113
7.5 Construction of QCD control regions. 114
7.6 QCD n_C/n_D ratios in each $H \rightarrow \tau\tau$ signal region. 115
7.7 QCD n_C/n_D ratios in each $H \rightarrow WW^* \rightarrow \ell\nu\tau\nu$ signal region. 115
7.8 Dominant theory systematic uncertainties. 117
7.9 Trigger scale factors and systematic uncertainties. 118
7.10 Systematic uncertainties in the $H \rightarrow \tau\tau$ analysis for events with no additional jets. 121
7.11 Systematic uncertainties in the $H \rightarrow \tau\tau$ analysis for events with at least one additional jet. 122
7.12 Systematic uncertainties in the $H \rightarrow WW^* \rightarrow \ell\nu\tau\nu$ analysis. 123

A.1 Simulated Z/γ^* background datasets. 136
A.2 Simulated W background datasets. 137
A.3 Simulated $t\bar{t}$, single top, single top in association with a W boson and Diboson datasets. 138
A.4 $H \rightarrow \tau\tau$ and $H \rightarrow WW^* \rightarrow \ell\nu\tau\nu$ signal process cross sections. 139
A.5 Theoretical uncertainties associated with the gluon-gluon fusion (ggF) and vector-boson fusion (VBF) signal processes. 139

Chapter 1

Introduction

This thesis was submitted in the Spring of 2012, a few months before the observation of a new boson consistent with the decay of a Standard Model Higgs boson at $m_H = 125$ GeV was jointly announced by the ATLAS and CMS collaborations.

The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (known as CERN) is the largest and highest energy particle collider in the world and has been regularly colliding proton and lead-ion beams since November 2009. Collision events are recorded by the four main LHC experiments: ALICE, ATLAS, CMS and LHCb. By the end of 2011, about 5 fb^{-1} of proton-proton collision data has been collected by both ATLAS and CMS at a centre-of-mass energy $\sqrt{s} = 7$ TeV, leading to the publication of many new search results and measurements. The data collected to date represent a small fraction of the full physics reach of the LHC which is expected to provide several hundred fb^{-1} of collision data at $\sqrt{s} = 14$ TeV. The analysis of these data is expected to expand the frontiers of today's knowledge of particle physics, which is condensed into the Standard Model (SM).

The SM of particle physics has been remarkably successful in explaining the results from decades of high energy physics experiments (including the LHC) in terms of elementary particles and their interactions. In spite of this, there are still as yet un-observed predictions of the SM: chief among them is the origin of particle masses. In the SM, particle masses arise through the so-called Higgs mechanism, evidence of which would be provided by the observation of the 'Higgs boson'. The discovery of Higgs boson production in proton-proton collisions is one of the main goals of the LHC.

The results from previous direct searches, indirect SM measurements and theoretical arguments require the mass of the SM Higgs boson to be narrowly constrained where the LHC can observe it. However, to verify that an observed neutral resonance is indeed the result of Higgs boson production, its spin and coupling strengths to other SM particles must be measured and compared with the SM predictions.

The work in this thesis is dedicated to the analysis of data collected in 2010, 2011 and a feasibility study of measuring the Higgs boson coupling strength to tau leptons. Therefore, much work has been done to understand the tau identification algorithms used at ATLAS and in particular to measure the probability of mis-identifying a hadronic jet as a hadronically decaying tau lepton (the tau 'fake rate').

The outline of this thesis is the following: Chapter 2 gives an overview of the SM, the Higgs mechanism and the physics of the Higgs boson at the LHC. A description of the LHC and the ATLAS experiment is given in Chapter 3. In Chapter 4, a new application for visualising collision events in ATLAS is described. A brief overview of tau lepton physics and an analysis measuring the tau fake rate in ATLAS data collected in 2010 are described in Chapter 6. Chapter 7 documents two searches for a light SM Higgs boson using 2011 ATLAS data using final states with a light charged lepton and a hadronically decaying tau lepton. In the absence of an observed excess, an upper limit on the signal cross section is placed at a 95% confidence level for Higgs boson masses in the range $100 < m_H < 180$ GeV/c^2. Finally, Chapter 5 presents the prospects for measuring the Higgs boson coupling strength to tau leptons using events in which the Higgs boson is produced in association with a weak gauge boson or a $t\bar{t}$ pair and subsequently decays to tau leptons, assuming 100 fb^{-1} of collision data at $\sqrt{s} = 14$ TeV.

Chapter 2

Theoretical overview

2.1 Introduction

This chapter gives a brief overview of the Standard Model, the Higgs mechanism and the phenomenology of Higgs boson production at a hadron collider.

2.2 The Standard Model

2.2.1 Overview

The Standard Model (SM) [1, 2, 3] of particle physics is used to describe the fundamental, point-like constituents of matter and three of the four forces through which they interact: the electromagnetic (EM) and weak nuclear forces that can be unified into the electroweak (EW) force and the strong nuclear force which is described by the theory of quantum chromodynamics (QCD).

The SM is a relativistic quantum theory in which quantum fields are used to describe particles with spin 1/2 known as fermions, and integer spin particles known as bosons. The SM is a gauge theory based on the symmetry group $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ in which interactions are introduced by requiring that the fermion fields remain invariant under a continuous group of local transformations. The dynamics of the fermion and boson fields are represented by a renormalisable Lagrangian

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD}, \tag{2.1}$$

which is a function of these fields. This Lagrangian cannot be complete, however, since the QCD and electroweak interaction terms cannot account for the experimentally observed masses of the fundamental particles.

The Higgs mechanism extends the SM Lagrangian to include a scalar field and its interactions

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}. \tag{2.2}$$

For a particular form of the scalar potential, this Lagrangian generates particle masses through the Higgs mechanism, also referred to as spontaneous symmetry breaking.

2.2.2 Particles of the Standard Model

Each particle field in the SM is defined by a unique set of quantum numbers. The twelve fermions, which are contained within three families of different flavour and increasing mass but otherwise identical quantum numbers, are shown in Table 2.1. For each fermion, there is a corresponding anti-particle with the same mass but opposite quantum numbers (such as the experimentally observed EM charge). Almost all visible matter in the universe is comprised of particles from Generation 1, since massive particles from higher generations are unstable. Fermions are further divided into leptons, which take part in electroweak interactions, and quarks that also participate in strong interactions due to their colour charge which can take three values: red, green and blue. Despite recent experimental results showing the neutrinos to have a tiny mass [4], in the SM they can be treated as massless. The boson particles that mediate the SM forces between the fermions are shown in Table 2.2.

Fermions	Generations			Charge	Weak Isospin		Hypercharge
	1	2	3	$Q[e]$	T	T_3	Y
Leptons (L)	$\left\{ \begin{matrix} e \\ \nu_e \end{matrix} \right\}$	$\left\{ \begin{matrix} \mu \\ \nu_\mu \end{matrix} \right\}$	$\left\{ \begin{matrix} \tau \\ \nu_\tau \end{matrix} \right\}$	0 -1	$1/2$ $1/2$	$1/2$ $-1/2$	-1 -1
Leptons (R)	e	μ	τ	-1	0	0	-2
Quarks (L)	$\left\{ \begin{matrix} u \\ d' \end{matrix} \right\}$	$\left\{ \begin{matrix} c \\ s' \end{matrix} \right\}$	$\left\{ \begin{matrix} t \\ b' \end{matrix} \right\}$	$2/3$ $-1/3$	$1/2$ $1/2$	$1/2$ $-1/2$	$1/3$ $1/3$
Quarks (R)	u d'	c s'	t b'	$2/3$ $-1/3$	0 0	0 0	$4/3$ $-2/3$

Table 2.1: Standard Model fermions and their associated quantum numbers. L and R refer to the left and right handed chirality states. The fermion-prime states are the physical eigenstates which are related to the weak eigenstates by the CKM matrix [5].

Bosons	Charge Q(e)	Mass (GeV)	Interactions
Photon: γ	$< 5 \times 10^{-30}$ [6]	$< 10^{-27}$ [6]	Electromagnetic: electrically charged particles.
Gluon: g	0	0	Strong: coloured particles (quarks and gluons)
W boson: W^+, W^-	± 1	80.385(15) [6]	Electroweak: fermions, W, Z,
Z boson: Z	0	91.1876(21) [6]	γ and H.

Table 2.2: Standard Model bosons with their electric charge, mass and the particles with which they interact.

2.2.3 Electroweak theory

Electromagnetic and weak interactions are unified in the SM using the symmetry group $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ into electroweak interactions that distinguish between particles with left- and right-handed chirality. For a given generation, left-handed fermions manifest themselves as a doublet of leptons or quarks under actions of the $\mathrm{SU(2)_L}$ group whose generator is the weak isospin operator, T. Weak Hypercharge, Y, is conserved quantity in electroweak interactions. The left- and right-handed fermions have different hypercharges under $\mathrm{U(1)_Y}$ rotations. The non-zero masses of the W and Z bosons break the $\mathrm{SU(2)_L}$ gauge symmetry, leaving a residual $\mathrm{U(1)_{EM}}$ symmetry with electromagnetic charge Q, defined as

$$Q = T_3 + \frac{Y}{2}. \tag{2.3}$$

.

The electroweak Lagrangian is given by

$$\mathcal{L}_{EW} = i\bar{L}\gamma^{\mu}D_{\mu}L + i\bar{R}\gamma^{\mu}D_{\mu}R - \frac{1}{4}W^{i}_{\mu\nu}W^{\mu\nu}_{i} - \frac{1}{4}B^{i}_{\mu\nu}B^{\mu\nu}_{i}, \tag{2.4}$$

where γ^{μ} are the Dirac matrices and L and R are the left- and right-handed projections of the fermion field. The four gauge fields, $W^{i}_{\mu}(i = 1, 2, 3)$ and B_{μ} are related to the $3 + 1$ degrees of freedom of the $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ group with the corresponding field strength tensors

$$W^{i}_{\mu\nu} = \partial_{\mu}W^{i}_{\nu} - \partial_{\nu}W^{i}_{\mu} - g\epsilon^{ijk}W^{j}_{\mu}W^{k}_{\nu} \quad \text{and} \quad B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu}. \tag{2.5}$$

The covariant derivatives that preserve local gauge invariance are

$$D_{\mu} = \partial_{\mu} + igT_{3}\frac{\sigma_{i}}{2}W^{i}_{\mu} + ig'YB_{\mu}, \tag{2.6}$$

where $\sigma_{i}(i = 1, .., 3)$ are the $\mathrm{SU(2)_L}$ group generators (i.e. the Pauli Matrices) and g and g' are coupling constants that determine the strength of the coupling to the $\mathrm{SU(2)_L}$ and $\mathrm{U(1)_Y}$ gauge fields, respectively. Since the Pauli matrices do not commute and the terms are non-Abelian, the Lagrangian contains self-interaction terms of the weak isospin gauge bosons.

The fields of the experimentally observable weak bosons $W^{\pm}_{\mu}, Z_{\mu}$ and the photon A_{μ} are given by a linear combination of the electroweak gauge fields

$$\begin{aligned} W^{\pm}_{\mu} &= \tfrac{1}{\sqrt{2}}(W^{1}_{\mu} \mp iW^{2}_{\mu}), \\ Z_{\mu} &= -B_{\mu}\sin\theta_{W} + W^{3}_{\mu}\cos\theta_{W}, \\ A_{\mu} &= B_{\mu}\cos\theta_{W} + W^{3}_{\mu}\sin\theta_{W}, \end{aligned} \tag{2.7}$$

where θ_{W} is the weak mixing angle. The electric charge and electroweak couplings are also related by

$$e = g \sin \theta_W = g' \cos \theta_W , \tag{2.8}$$

where the weak mixing angle has been measured experimentally using the Z pole observables; the Z-boson mass, m_Z[1], to be $\sin^2 \theta_W = 0.23146(12)$ [7] in the on-shell scheme and the strong coupling constant as measured at m_Z, $\alpha_s(m_Z)$.

These transformed fields show that the charged W bosons couple to all left-handed fermions and right-handed anti-fermions with the same coupling strength. In the neutral-current interactions, the Z boson couples differently to each fermion depending on its charge and weak isospin.

2.2.4 QCD

The strong nuclear interaction is based on the non-Abelian symmetry group $\mathrm{SU}(3)_\mathrm{C}$ and describes the interactions of particles that have colour charge, i.e. the quarks and gluons. Each quark forms a triplet in colour space that can have one of three colours with corresponding anti-colours for the anti-quarks (such that colour-anti-colour states are colour-neutral singlets). The group $\mathrm{SU}(3)_\mathrm{C}$ has eight generator matrices λ^a and hence there are eight different gluon gauge fields $G^a_\mu (a = 1, ..., 8)$ where each is a unique colour-anti-colour superposition. The QCD Lagrangian is given by

$$\mathcal{L}_{QCD} = \sum_f \bar{q}_f (i\gamma^\mu D_\mu - m_f) q_f - \frac{1}{4} G^{\mu\nu}_a G^a_{\mu\nu}, \tag{2.9}$$

where q_f and m_f denote the quark fields and masses. In the QCD Lagrangian, the gluon fields enter via the field strength tensors

$$G^{\mu\nu}_a = \partial^\mu G^\nu_a - \partial^\nu G^\mu_a - g_s f^{abc} G^\mu_b G^\nu_c , \tag{2.10}$$

[1]and the Fermi constant which is derived from the muon lifetime formula.

where g_s denotes the strong coupling constant and the structure constants f^{abc} determine the commutators of the $SU(3)_C$ generators. The covariant derivatives that can be chosen to preserve local gauge invariance are given by

$$D_\mu = \partial_\mu + ig_s \frac{\lambda_a}{2} G_\mu^a, \tag{2.11}$$

and the local gauge invariance requires that the gluons are massless. The non-Abelian nature of this theory results in non-zero commutators of the generator matrices, resulting in cubic and quartic gluon self-coupling terms. These self-interactions also account for two phenomena of the quarks and gluons: the so-called 'asymptotic freedom' (that at very small length scales they can be regarded as free particles) and 'confinement' (observable particles are bound colour-singlets states and free quarks and gluons cannot be observed).

2.3 The Higgs mechanism

The EW and strong interactions of the SM so far require the particles to be entirely massless[2] to ensure gauge invariance of the SM Lagrangian. This is completely at odds with a host of experimental results in which the fermion and weak gauge bosons are shown to indeed have mass.

In the SM, particle masses can be introduced by a phenomenon known as the Higgs mechanism [8, 9, 10, 11] in which the mass terms are generated by spontaneously breaking the electroweak symmetry with a scalar field potential.

[2] Even though mass terms are allowed for the fermion fields in $\mathcal{L}_{QCD}$, they are forbidden for the same fermion fields in $\mathcal{L}_{EW}$ in order to conserve gauge invariance.

2.3.1 Spontaneous symmetry breaking

A complex, two-component scalar field,

$$\Phi = \frac{1}{\sqrt{2}} \begin{Bmatrix} \phi_1^1 + i\phi_2^1 \\ \phi_1^2 + i\phi_2^2 \end{Bmatrix}, \tag{2.12}$$

is introduced, chosen to be an isospin doublet of $\mathrm{SU(2)_L}$ with weak hypercharge $Y = 1$. The self-dynamics of this new field are described by

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi), \tag{2.13}$$

with the same covariant derivatives as for the EW Lagrangian shown in Equation 2.6. The scalar field potential is defined as

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \tag{2.14}$$

If both $\lambda > 0$ and $\mu^2 > 0$ this potential takes the form shown in Figure 2.1 with a continuous non-zero minimum at $\Phi^\dagger \Phi = \mu^2/2\lambda$ and a vacuum expectation value (VEV) $v = \frac{\mu^2}{2\lambda}$.

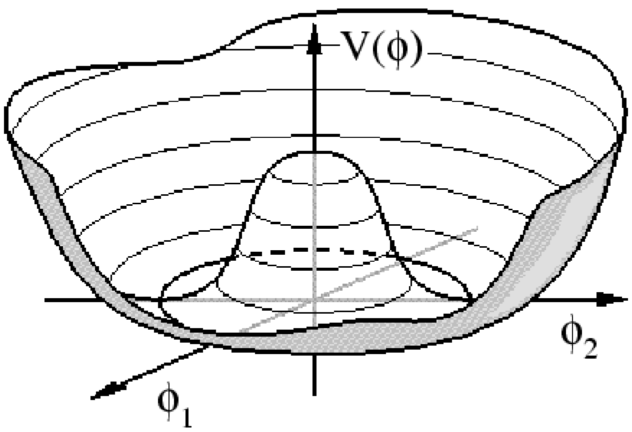


Figure 2.1: Illustration of the Higgs potential $V(\phi)$ of the complex scalar field $\phi = \phi_1 + i\phi_2$ for $\lambda > 0$ and $\mu^2 > 0$.

The ground state of this field can be arbitrarily chosen along a ring of minimum potential, the act of which is known as spontaneous symmetry breaking. One choice is

$$\Phi_0 = \frac{1}{\sqrt{2}} \left\{ \begin{array}{c} 0 \\ v \end{array} \right\} \qquad \text{with} \qquad v = \frac{\mu^2}{\lambda}, \tag{2.15}$$

with an expansion around this ground state being described by

$$\Phi = \Phi_0 + \delta\Phi = \frac{1}{\sqrt{2}} \left\{ \begin{array}{c} 0 \\ v + h(x) \end{array} \right\}, \tag{2.16}$$

leading to a physical Higgs field $h(x)$ (and three massless Goldstone bosons that are absorbed to give mass to the $W^{\pm}$ and Z gauge bosons). Substituting this into Equation 2.13 leads to terms where the Higgs field couples to the gauge fields $W_\mu^i W_i^\mu$ and $B_\mu B^\mu$. A non-zero VEV produces mass terms for these gauge fields with eigenstates given by Equation 2.7 and eigenvalues given by

$$m_W = m_Z \cos\theta_W = vg/2 \quad \text{and} \quad m_\gamma = 0. \tag{2.17}$$

The vacuum expectation value can also be calculated to be $v = 246$ GeV using the measured value of the Fermi constant [5].

2.3.2 Fermion masses

To account for fermion masses, the Lagrangian is extended by the so-called 'Yukawa' terms

$$\mathcal{L}_{Yukawa} = -g_f(\bar{f}_L \Phi f_R + \bar{f}_R \Phi^\dagger f_L). \tag{2.18}$$

Again substituting the expansion of the scalar field Φ around the chosen minimum leads to terms of the functional form

$$-\frac{g_f}{\sqrt{2}}(v + h(x))(\bar{f}_L f_R + \bar{f}_R f_L), \tag{2.19}$$

in the Lagrangian from which the fermion masses can be read

$$m_f = \frac{g_f v}{\sqrt{2}}. \tag{2.20}$$

2.4 Higgs boson phenomenology

In the Lagrangian of Equation 2.13, one physical scalar field of the original four remains, the quantum of which is known as Higgs boson. As a consequence of this, if the Higgs mechanism describes the realisation of electroweak symmetry breaking in nature, there is a CP-even, electrically neutral particle that has coupling to all massive SM fermions and bosons that can be experimentally observed. However, the mass of the Higgs boson is not predicted in the Higgs mechanism and must be added by hand, i.e. it must be directly measured in an experiment.

2.4.1 Constraints on the Higgs mass

Terms in the electroweak Lagrangian (Equation 2.4) lead to vertices where the weak bosons self-interact. Since the probability for any processes cannot exceed one, a constraint on the s-wave scattering amplitude can be made such that as upper limit on the Higgs mass can be estimated: $m_H \leq 850$ GeV.

Another indirect constraint comes from the precision electroweak measurements observed at other experiments, with the Higgs boson mass left as a free parameter to be fitted. One example of this comes from the very well measured weak boson masses since the Higgs boson contributes to the $W^{\pm}$ vacuum polarisations through loop effects. Combining this with other electroweak measurements from LEP, SLC and the Tevatron leads to a fit value of $m_H = 87^{+35}_{-26}$ GeV, or an upper limit of $m_H < 157$ GeV at 95% confidence level [12].

The indirectly constrained range of Higgs masses is also consistent with the results from direct searches. If the global fit value takes into account the lower limit placed on the SM Higgs mass by the LEP experiments of $m_H > 114.4$ GeV [12] then the indirect upper limit becomes $m_H < 186$ GeV at 95% confidence level. The global minimum is shown in Figure 2.2.

Other direct searches conducted at the Tevatron analysing up to 10.0 fb^{-1} of $p\bar{p}$ collision data at a centre of mass energy $\sqrt{s} = 1.96$ TeV exclude a 30 GeV wide mass region around $m_H = 160$ GeV in addition to the region already excluded by LEP, as shown in Figure 2.3.

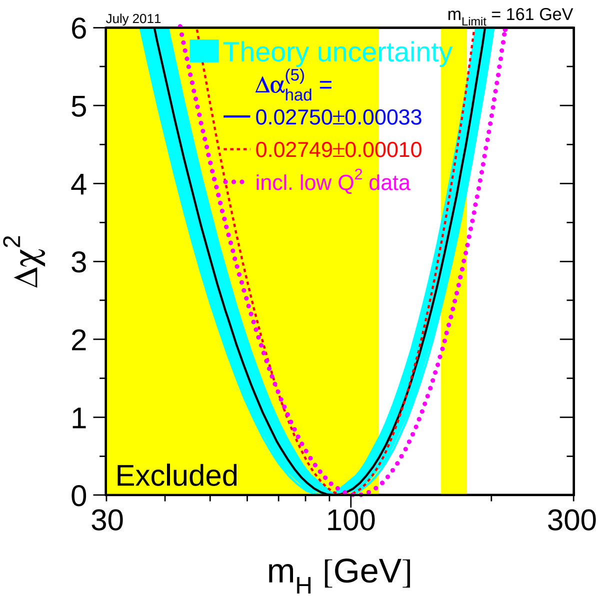


Figure 2.2: A global SM fit ($\Delta\chi^2 = \chi^2 - \chi^2_{min}$) for a range of Higgs boson masses to precision electroweak observables at LEP. The yellow area signifies masses excluded by direct searches (taken from reference [12]).

2.4.2 Production and decay at the LHC

At the LHC, there are several channels in which evidence of SM Higgs boson production can be sought. In particular, to identify any neutral resonance as a SM Higgs boson, its

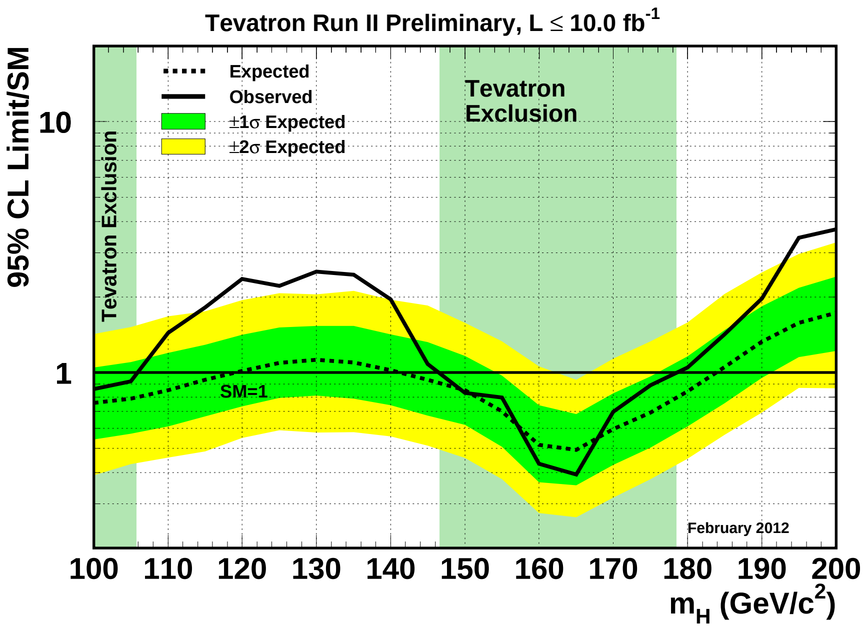
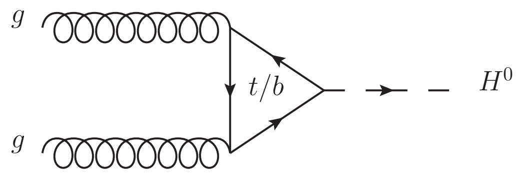


Figure 2.3: Combined upper limits on the SM Higgs boson production cross section divided by the SM expectation as a function of m_H (solid lines) obtained in a combination of results from the CDF and D0 experiments with an integrated luminosity of up to 10.0 fb^{-1}. The dashed line shows the median expected limit in the absence of a signal and the green and yellow bands indicate the corresponding 68% and 95% expected regions. Mass regions in which the observed limit is smaller than one, are excluded (taken from reference [13]).

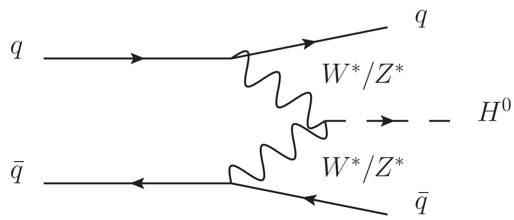
couplings to other SM particles and its spin must also be measured and compared with the SM predictions.

Higgs production

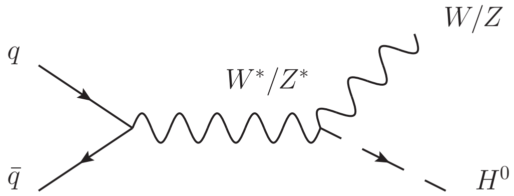
The m_H dependence of the cross section for the dominant Higgs boson production processes is shown in Figure 2.5(a). The most abundant source is expected to be the gluon-gluon fusion process, the diagram of which is shown in Figure 2.4(a). It is the dominant production mode even though it is loop-induced because of the strong coupling of the Higgs boson to the (virtual) top quark, the relatively large α_s for QCD processes and the huge flux of gluons at low Q^2 for LHC collisions. Despite the lower cross section for vector boson fusion diagrams (see Figure 2.4(b)), it is expected that they provide a more sensitive search channel due to the lack of colour exchange between the out-going partons. Similarly, despite the lower cross section for the associated production diagrams shown in Figures 2.4(c) and 2.4(d), the additional final state particles from, for example, a leptonically decaying W boson, tend to completely change the background yield and composition.



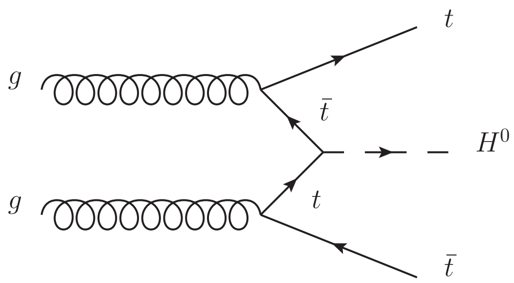
(a) Gluon-gluon fusion.



(b) Weak boson fusion.



(c) Associated production with a weak boson.



(d) Associated production with a $t\bar{t}$ pair.

Figure 2.4: Dominant Higgs boson production mechanisms at the LHC.

Higgs decay

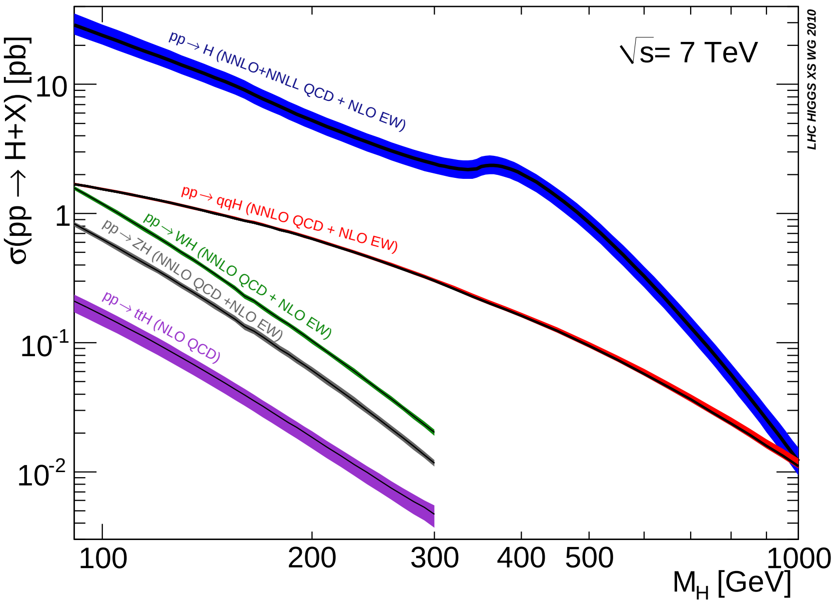
In the SM, the coupling strength g of a Higgs to fermion anti-fermion (weak gauge-boson) pair vertex is directly proportional to the mass of the fermion (weak gauge-boson). The partial width of a Higgs to fermion anti-fermion (gauge-boson pair) decay Γ_f (Γ_V) is expressed

$$\Gamma_f \propto m_f^2 \qquad \text{and} \qquad \Gamma_V \propto m_V^2. \tag{2.21}$$

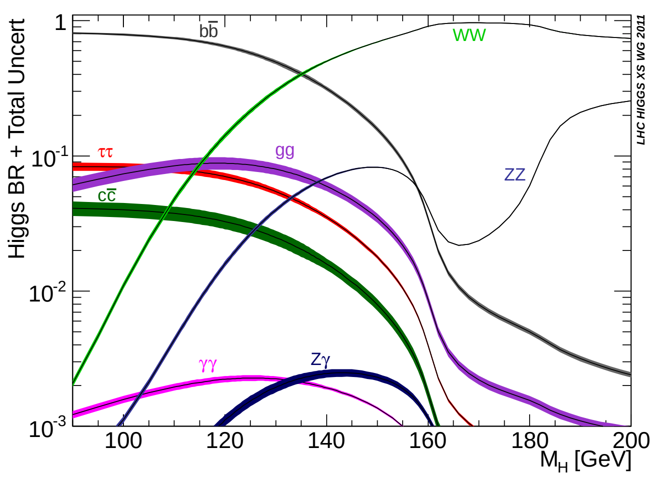
The expected relative decay rate of a SM Higgs boson to a given pair of SM particles can be expressed in terms of the branching ratio ($\mathcal{B}$). For a SM Higgs boson with a total decay width Γ_{tot} and particular decay $H \to X\bar{X}$ with partial width Γ_X the $\mathcal{B}$ is defined to be

$$\mathcal{B}(H \to X\bar{X}) = \frac{\Gamma_X}{\Gamma_{tot}}. \tag{2.22}$$

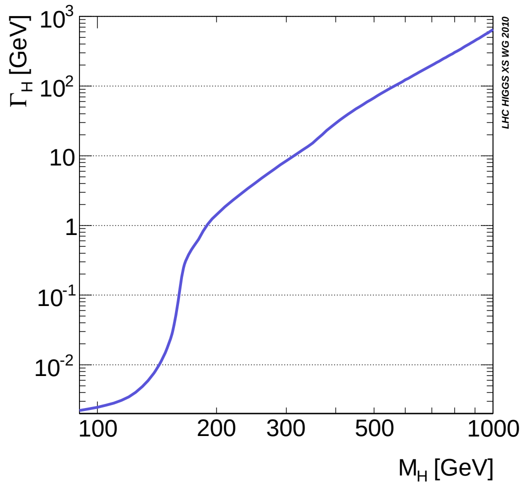
The total width increases rapidly as m_H increases, with a jump of about three orders of magnitude at the W^+W^- production threshold. For low masses preferred by fits to electroweak data, $\Gamma_{tot} << m_H$. The mass dependence of the $\mathcal{B}$ of the Higgs boson is shown in Figure 2.5(b).



(a) SM Higgs boson production cross section at $\sqrt{s} = 7$ TeV as a function of m_H.



(b) Expected branching ratio of the SM Higgs boson to SM particles as a function of the Higgs boson mass m_H.



(c) SM Higgs boson total width as a function of m_H.

Figure 2.5: Mass dependence of (a) the SM Higgs production cross section, (b) the SM Higgs branching ratios, and (c) the total SM Higgs boson width (taken from reference [14]).

For a low-mass SM Higgs boson with $115 \leq m_H \leq 135$ GeV, the expected discovery modes are the subleading decays $H \to \gamma\gamma$, $H \to ZZ$, and $H \to WW$ [15, 16, 17]. With more data, it is expected that the $H \to \tau\tau$ and $H \to b\bar{b}$ [18, 19, 20, 21, 22] decays can be observed in various production processes, allowing a wide variety of Higgs boson cross section measurements. Specific SM Higgs boson coupling ratio measurements could then be made [23, 24] and compared with the SM prediction.

For masses favoured by global fits of electroweak data ($m_H < 135$ GeV), the dominant decay mode is to b quarks, but due to the very high $b\bar{b}$ jet production cross section at the LHC, this is experimentally very challenging. At these low masses, one of the most attractive decay modes to study becomes that to the heaviest leptons in the SM, tau leptons. At higher masses, another interesting decay mode with the same final state as a Higgs boson decaying to tau leptons is a Higgs boson decaying to a W^+W^- pair that subsequently decay leptonically with at least one W boson decaying to a tau lepton and a neutrino.

2.4.3 Measuring the SM Higgs boson coupling strength

With large LHC datasets ($\int \mathcal{L}\ \mathrm{dt} > 30\ \mathrm{fb}^{-1}$), it should be possible to measure the relative coupling strength of Higgs to other SM particles by measuring the production cross section and branching ratio of different Higgs production and decay processes. Cross section measurements are proportional to the square of these couplings. For example,

$$\sigma(pp \to VH) \propto \Gamma_V \quad \text{and} \quad \sigma(pp \to t\bar{t}H) \propto \Gamma_t, \tag{2.23}$$

where t is the top quark, V is a vector boson and Γ_V and Γ_t are the partial widths of a Higgs to $V\bar{V}$ or $t\bar{t}$ decay.

The ratio of two Higgs cross section measurements can be used as a direct probe of the ratio of the Higgs boson coupling strengths to such fields, for example

$$\frac{\sigma(pp \to VH(\to \tau\tau))}{\sigma(pp \to t\bar{t}H(\to \tau\tau))} \propto \frac{\Gamma_V}{\Gamma_t} = \frac{g^2_{VVH}}{g^2_{t\bar{t}H}}. \tag{2.24}$$

or

$$\frac{\sigma(pp \to VH(\to b\bar{b}))}{\sigma(pp \to VH(\to \tau\tau))} \propto \frac{\Gamma_b}{\Gamma_t} = \frac{g^2_{b\bar{b}H}}{g^2_{\tau\tau H}}. \tag{2.25}$$

A ratio measurement has the advantage that some sources of systematic uncertainty common to both measurements will cancel when taken in ratio, for example the uncertainty on the integrated luminosity. Additionally, if both measurements have a common Higgs decay channel, systematic uncertainties on the measured Higgs decay products will also cancel.

Chapter 3

The ATLAS detector at the LHC

3.1 Introduction

In this chapter, a description of the LHC, the ATLAS[1] detector and the event reconstruction algorithms used in the analyses presented in this thesis are given.

3.1.1 The Large Hadron Collider

The LHC [25] is a subterranean double-ring superconducting hadron collider, installed in a circular tunnel of radius 4.25 km, located between 45 m and 180 m underground, near Geneva, Switzerland. Eventually, proton-proton collisions are planned to have $\sqrt{s} = 14$ TeV and an instantaneous luminosity $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1}. The first successful collisions occurred in Autumn 2009 at $\sqrt{s} = 450$ GeV. From March 2010 until the end of 2011, the first phase of the LHC physics research programme has been carried out with collisions at $\sqrt{s} = 7$ TeV, in which both the ATLAS and the Compact Muon Solenoid (CMS) experiments have collected an integrated luminosity of about 5 fb^{-1}. The lead ion collisions are not discussed here, since they are not relevant for the analyses in this thesis.

The LHC is designed to have up to 2808 circulating proton 'bunches', where each bunch

[1]Formerly the experiment's name was an acronym: A Toroidal LHC ApparatuS.

has a diameter of about 7 μm and is made up of about 10^{11} protons. There is a spatial distance of about 7.5 cm between bunches or equivalently 25 ns between bunch crossings. Proton bunches are directed by 8.3 T magnetic fields generated locally by more than 1200 superconductive dipole magnets, each of which is about 15 m long. The beams are brought to collision at four points around the ring, where the four experiments ALICE, ATLAS, CMS and LHCb are situated in underground caverns, as shown in Figure 3.1.

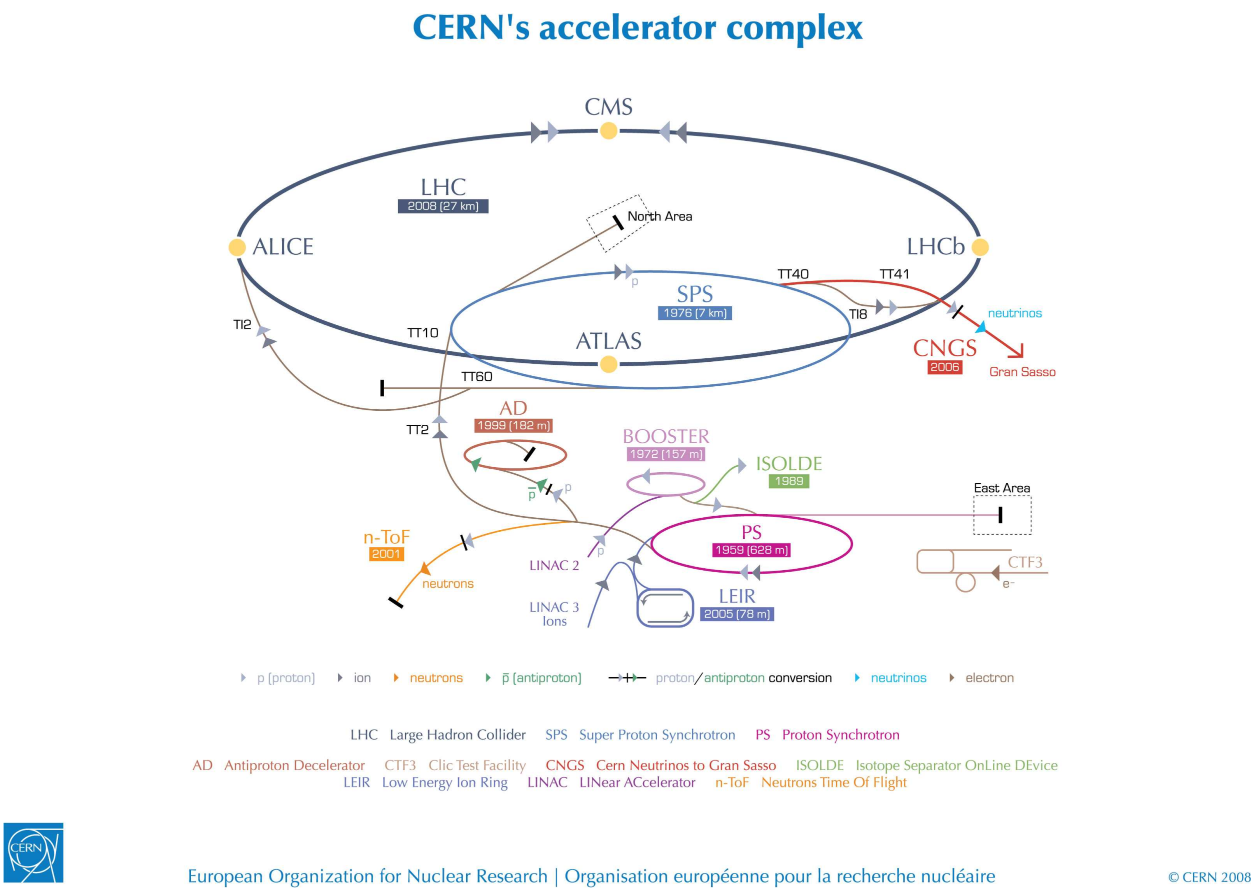


Figure 3.1: Schematic diagram of the CERN accelerator complex and the location of the four main LHC experiments at the beam-crossing points (taken from reference [26]).

On each side of the bunch crossing point, almost 400 quadrupole magnets 5-7 m long focus the proton bunches before collision. The LHC operates in 'fills' where protons are first accelerated from approximately at rest to 450 GeV, and then injected into the LHC. From there, it takes about 20 minutes for the beams to be accelerated to the target energy of the LHC in 2011: $E = 3.5$ TeV. Proton acceleration in the LHC occurs in eight superconducting radio frequency cavities around the ring, each with a field gradient of 5 MV/m. After acceleration to 3.5 TeV, the beams are suitable for physics data for up to 10 hours, after which they are dumped onto an absorber and the next fill is prepared.

3.1.2 ATLAS coordinate system

The coordinate system for events in the ATLAS detector is defined here and followed throughout the rest of this thesis. The counter-clockwise beam direction defines the positive z-axis while the $x-y$ plane is transverse to the beam direction. The positive x-axis points towards the centre of the LHC ring and the positive y-axis points directly upwards. Switching to a cylindrical polar coordinate system, the polar angle θ and azimuthal angle ϕ are measured with respect to the z-axis and x-axis respectively. Since the polar angle θ is not invariant under a Lorentz boost, it is also useful to define the pseudo-rapidity of a particle with four-momentum (p_X, p_Y, p_Z, E) as

$$\eta = \frac{1}{2} \ln \frac{|\vec{p}| + p_Z}{|\vec{p}| - p_Z} = -\ln \tan \frac{\theta}{2}, \tag{3.1}$$

which is a measure of the 'forwardness' of a particle's vector. A comparison of η and θ is shown in Figure 3.2.

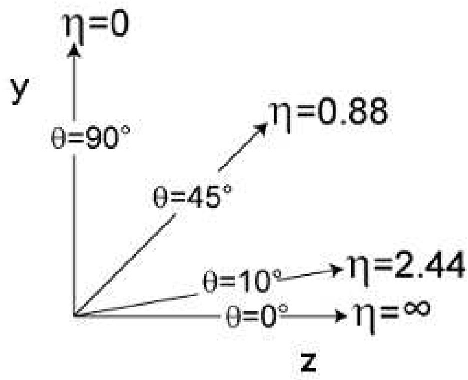


Figure 3.2: Pseudo-rapidity η and polar angle θ in the ATLAS coordinate system.

The transverse momentum (p_{T}) is defined to be the momentum in the plane transverse to the beam axis, i.e. the magnitude of the vector in the $x-y$ plane. The angular separation of any two objects is also defined to be

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}, \tag{3.2}$$

and is useful, for example, in removing overlapping reconstructed candidate objects or matching hits at the boundary between two sub-detectors.

3.1.3 Detector requirements and specifications

The ATLAS experiment [27, 28] is one of two general purpose experiments at the LHC and is designed in a layered configuration of nearly hermetic sub-detectors, as shown in Figure 3.3. ATLAS is intended to investigate a wide range of physical processes, some of which are shown in Figure 3.4.

ATLAS is therefore designed to handle more than 40×10^6 inelastic scattering interactions per second, the vast majority of which are 'minimum-bias' events i.e. QCD interactions at predominantly low momentum transfer. At $\mathcal{L} = 10^{33}$ cm^{-2}s^{-1} about 100 W and 10 Z gauge bosons are produced each second, with an expected Standard Model (SM) Higgs boson production rate several orders of magnitude below this. The high rate for weak boson production allows for an unprecedented increase in statistics from previous experiments allowing precision measurement of quantities predicted in the SM. At the same time, they also provide a significant source of background for SM Higgs searches. In order to successfully perform measurements of high-rate processes in a high-luminosity environment while also observing rare processes, the detector was designed with the following requirements:

- fast radiation-hard electronics to separate measurements of the nominal collision from those of adjacent bunch crossings (which produce 'out-of-time' pileup);
- highly granular tracking detectors and calorimeters to separate particles resulting from multiple parton-parton collisions in a given bunch crossing ('in-time' pileup);
- good charged-particle momentum and impact-parameter resolutions, with high recon-

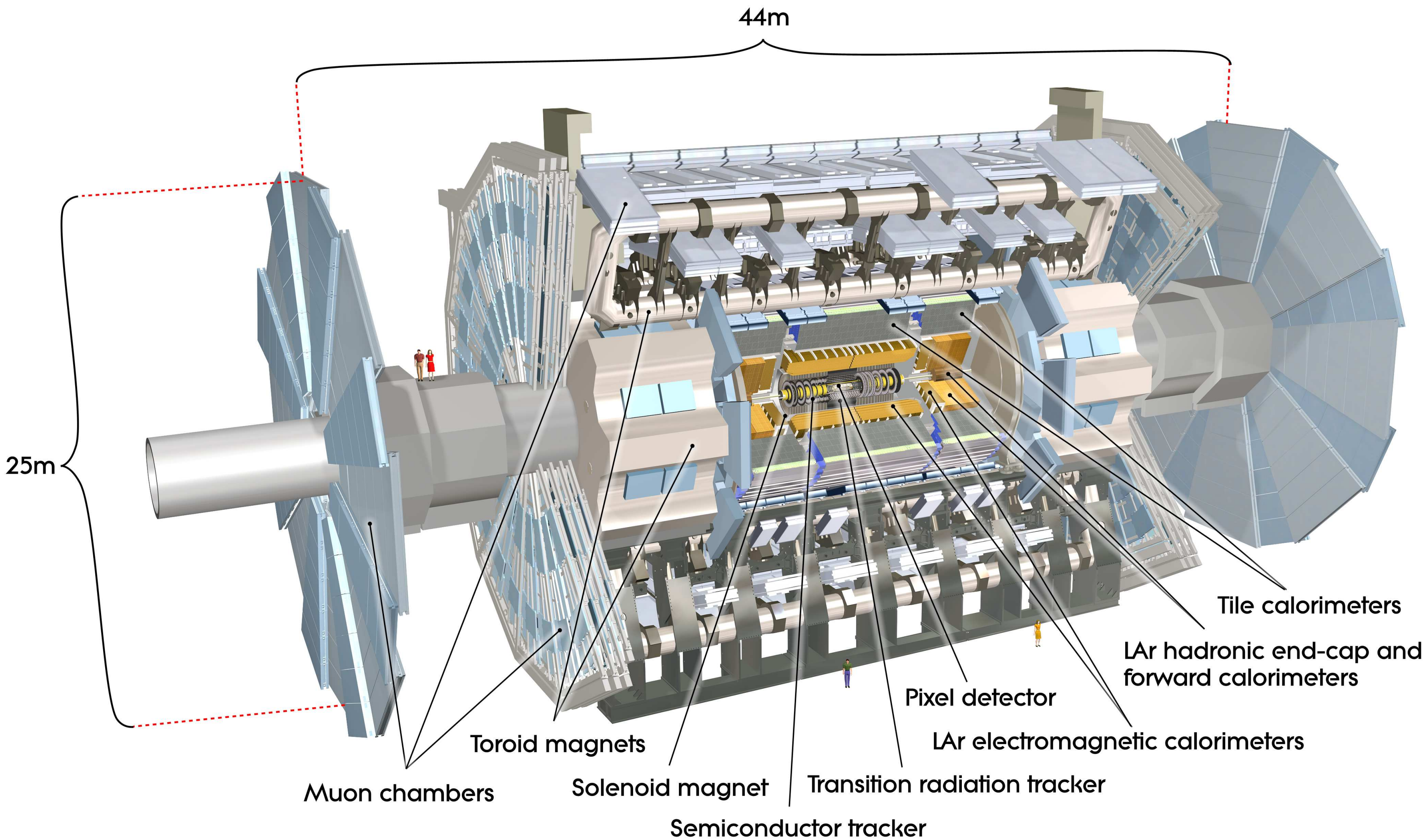


Figure 3.3: Diagram showing the situation, scale and relative positions of the main ATLAS sub-detectors (taken from reference [27]).

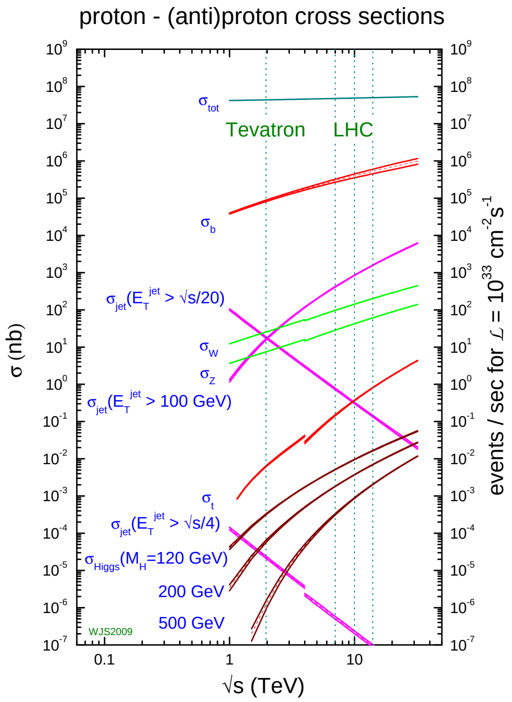


Figure 3.4: Cross section for various processes as a function of centre of mass collision energy and the event rate (taken from reference [29]).

struction efficiency;

- excellent electromagnetic calorimetry to measure the energy and shower shapes of electrons and photons, along with large angular coverage;

- hadronic calorimetry for accurate hadronic jet and missing transverse momentum measurements;

- muon detectors to provide good muon identification and momentum resolution over a wide range of momenta;

- high bandwidth for triggering, with efficient triggers selecting objects with p_{T} values typical of a W or Z boson decay and sufficient background rejection to achieve an acceptable trigger rate.

3.2 Magnets

In order to make precise charge and transverse momentum measurements, there are two magnetic fields in ATLAS: a central solenoid (CS) and an air-core toroidal system.

The superconducting CS (shown in Figure 3.3) surrounds the inner detector and provides a 2 T magnetic field, causing the path of charged particles to follow a helix within its volume. From the curvature of this helix, precise charge and transverse momentum measurements can be made. A separate magnetic field of strength 0.5-1 T is created in the outermost region of the detector by three superconducting air toroids (again, shown in Figure 3.3). Like most of the sub-detectors, the air toroids are designed to cover the central (low $|\eta|$) area with a cylindrical 'barrel' module flanked at either end in the z-axis by two 'endcap' modules extending to higher $|\eta|$.

3.3 Inner detector

The inner detector (ID) consists of the tracking detectors surrounding the interaction point (IP) at the heart of ATLAS. The three sub-detectors that make up the ID: the pixel, semi-conductor tracker (SCT) and transition radiation tracker (TRT) detectors are illustrated in Figure 3.5. As a charged particle propagates through these detector elements, space-point measurements are made while the path of the particle is bent by the CS magnetic field. From the curvature of the resulting helical path the particle describes, both the charge and the transverse momentum of the particle can be inferred.

3.3.1 Pixel detector

To ensure good vertex reconstruction, the tracking detector closest to the IP requires the highest possible resolution to accurately extrapolate the reconstructed tracks back to the IP. In ATLAS, a silicon pixel detector based system is employed with three barrel layers arranged as concentric cylinders around the z-axis. In the higher $|\eta|$ regions, two endcap

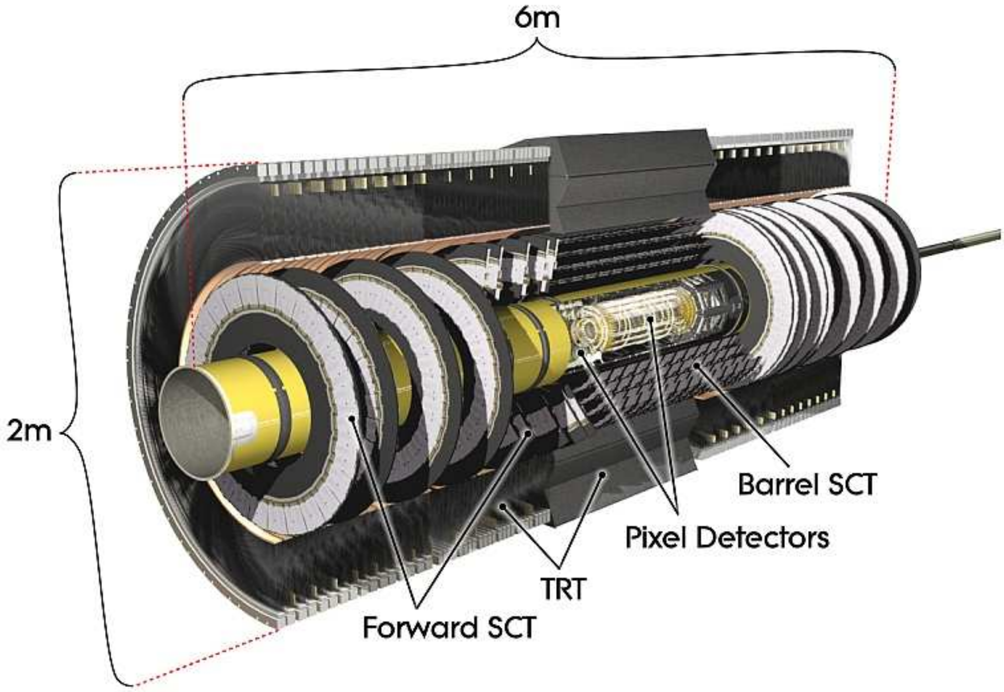


Figure 3.5: Cross-sectional diagram showing the ATLAS inner detector (taken from reference [27]).

modules consist of three disks placed perpendicular to the beam axis. The pixel detector provides an intrinsic position resolution of 10 μm in the $R - \phi$ plane and 115 μm in the z (R) plane in the barrel (endcap) region. In total, there are 1744 identical pixel elements with approximately 80.4 million readout channels.

In summary, a charged particle will produce up to three 'hits' in the pixel detector with very high spatial resolution that allow for primary and secondary vertex reconstruction, which can be used in b flavour jet and tau lepton tagging.

3.3.2 Semiconductor tracker

The ATLAS SCT barrel section comprises four cylindrical layers, each equipped with 8448 rectangular silicon-strip sensors. Both SCT endcap sections consist of nine disks such that any charged particle track with $|\eta| < 2.5$ will cross at least four of them (or the four barrel SCT layers). The silicon strips collect charge in a similar way to the pixel elements but with a coarser granularity, with an intrinsic resolution of 17 μm in the $R - \phi$ plane and 580 μm along the z (R) axis in the barrel (endcap) region. Each layer in the barrel and endcap is composed of two strip layers, crossed at a stereo angle of 40 mrad. The combination of measurements in the two strips provides a three dimensional position measurement of the traversed charged particle. In total there are about 6.4 million readout channels in the SCT.

3.3.3 Transition radiation tracker

The ATLAS TRT is the outermost radial ID sub-detector and performs measurements of tracks with $|\eta| \leq 2$, providing a different approach to the silicon-based Pixel and SCT detectors.

The TRT is composed of 4 mm polyimide, straw drift tubes filled with a gas mixture of 3% O_2, 27% CO_2 and 70% Xe. These straws are situated parallel to the beam direction in the barrel region and radially in the endcap regions and hence only provide $R - \phi$ information, with an intrinsic resolution of 130 μm per straw. This lower resolution is mitigated by a large number of measurements per track, since each track will pass through $\sim$ 36 straws. In total, there are approximately 351,000 TRT readout channels.

Transition radiation is produced when a relativistic particle traverses an inhomogeneous medium such as the boundary between materials of different electrical properties. In the TRT, polypropylene fibres (foils) are situated between the barrel (endcap) straws to provide this boundary. The radiation produced ionises the gas mixture of the straw tubes. The resulting drift electron current is amplified by about a factor of 10^4 as it is drawn to a central gold plated tungsten wire running along the tube. The total drift time is approximately 40 ns.

The intensity of the radiation produced is proportional to the particle's Lorentz factor $\gamma = E/m$. Hence, due to the mass difference between electrons and charged hadrons, the magnitude of transition radiation can also be used to identify the track.

3.4 Calorimetry

A more detailed view of the ATLAS calorimeters is shown in Figure 3.6. The main purpose of the calorimters is to determine the position and magnitude of energy deposited by particles. This is achieved using a highly granular liquid Argon (LAr) electromagnetic (EM) sampling calorimeter envelopes the solenoid with a central barrel section and two endcaps with coverage up to $|\eta| < 3.2$ and a hadronic calorimeter (HCAL) composed of different technologies for different regions of $|\eta|$, extending to a pseudo-rapidity of $|\eta| = 4.9$.

The energy resolution of the EM calorimeter is parametrised in terms of a so-called 'sampling' term[2] and a constant term[3] to be $\frac{\sigma_E}{E} = \frac{0.1(\mathrm{GeV}^{1/2})}{\sqrt{E}} \oplus 0.0017$ where $\oplus$ denotes addition in quadrature. The HCAL energy resolution is parametrised to be $\frac{\sigma_E}{E} = \frac{0.5(\mathrm{GeV}^{1/2})}{\sqrt{E}} \oplus 0.03$.

The high granularity of the EM calorimeter is essential to the identification of electrons and photons. The coarser granularity of the HCAL is designed for jet reconstruction and measurements of an imbalance in the transverse momentum vector-sum of all energy deposits, the missing transverse momentum[4]. The final function of the calorimeters is to contain the EM and hadronic shower shapes and hence limit any EM or hadronic particles from entering the muon detectors. This is accomplished with a minimum calorimeter depth of 22 interaction lengths (χ_0) in the barrel region and 24 interaction lengths in the endcaps.

3.4.1 Electromagnetic calorimeters

The EM calorimeter consists of a central barrel section covering $|\eta| < 1.475$ with two endcap sections ($1.375 < |\eta| < 3.2$). The barrel is composed of two identical halves symmetric about $z = 0$, each made of 16 modules covering $\pi/8$ of the ϕ plane. Each endcap has

[2]The sampling term has $\propto E^{1/2}$ dependence due to the statistical nature of the energy deposition in the calorimeters.

[3]A constant term independent of E is also present due to non-uniformities in the detector calibration.

[4]Since the the initial longitudinal momentum of the incoming partons is unknown, so is the longitudinal boost of the center of mass frame of the final state particles in each collision. Therefore, only the transverse projections of the event's momentum can be used to infer the production of particles that do not interact with the detector, such as neutrinos.

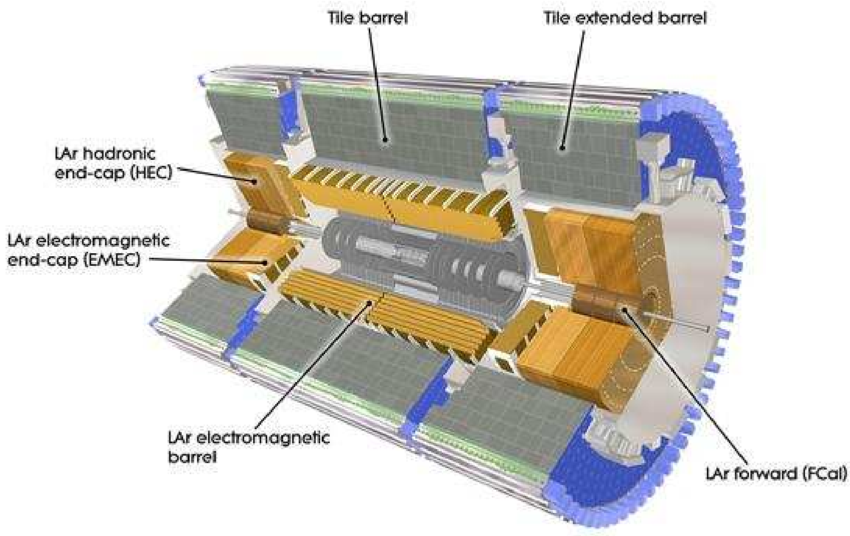


Figure 3.6: Calorimeters in the ATLAS detector (taken from reference [27]).

two coaxial wheels where the outer wheel makes precision measurements of particles with $1.375 < |\eta| < 2.5$, and the inner wheel makes lower resolution measurements in the region $2.5 < |\eta| < 3.2$. In the region designed to make precision measurements of photon and electron energy deposits, $|\eta| < 2.5$, the EM calorimeter is split longitudinally into three sections. The first is referred to as the '$|\eta|$ strip layer'. The middle layer has a depth of 16 χ_0, a coarser granularity than the $|\eta|$ strip layer and is designed to contain the main energy deposit of an EM shower. The back layer is twice as coarse in granularity again and stops EM energy leaking into the hadronic calorimeter. To correct for energy losses in the dead material in front of the calorimeters, an additional thin liquid argon layer, the 'presampler' layer, sits in front of the η strip layer in the region $|\eta| < 1.8$. An illustration of the EM calorimeter is shown in Figure 3.7.

The EM calorimeter is a LAr sampling calorimeter with lead absorber plates that provide complete ϕ symmetry without any cracks in ϕ. The lead absorber plates also initiate electromagnetic showers of incident electrons and photons in which a cascade of EM particles is produced, starting from a single e^+e^- pair from a photon or a bremstrahlung photon ejected from an electron. Sandwiched between the lead absorber layers are LAr sampling layers in which the incident electrons ionise the Argon. The resulting charged current is collected by copper electrodes with a drift time of about 250 ns for a 2000 V potential.

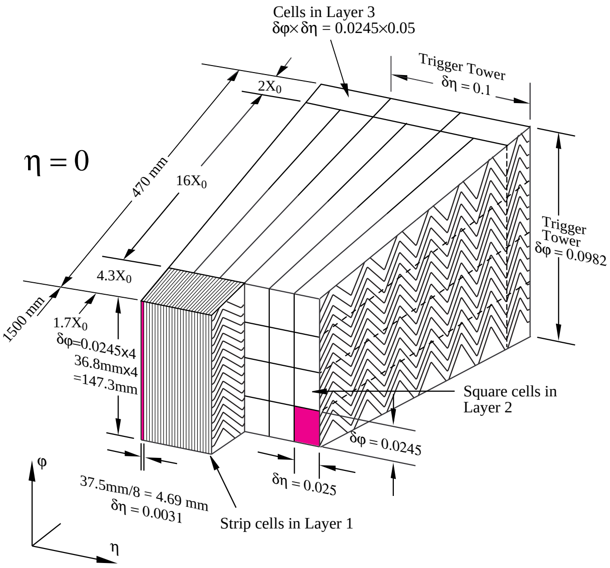


Figure 3.7: A sketch of an EM calorimeter barrel module. The granularity of the transverse and longitudinal layers is also shown (taken from reference [30].)

The EM calorimeter energy resolution is parametrised after noise subtraction as:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E(\mathrm{GeV})}} \oplus b, \tag{3.3}$$

where a and b are constants describing the stochastic nature of the energy deposition and local non uniformities in the response, respectively. These constants have been measured as a function of $|\eta|$ using electron test-beams with a known energy where $a = 0.1$ GeV$^{1/2}$ and $b = 0.0017$ for $|\eta| = 0.69$.

3.4.2 Hadronic calorimeters

The hadronic calorimeter system measures the energy and direction of hadronic particles that survive the EM calorimeter. Hadronic calorimetry in the central region ($|\eta| < 1.7$) is provided by a plastic scintillator-tile sampling calorimeter (TileCal). In the TileCal, stacks of steel plates sandwiched with polystyrene scintillator tiles form the active material that initiate hadron showers. Each side of the scintillating tiles is read out by a photomultiplier tube.

At larger pseudo-rapidities up to $|\eta| < 4.9$, hadronic calorimetry is provided by the LAr

Hadronic Endcap Calorimeter (HEC) and the Forward Calorimeter (FCal). The HEC is described above as the inner wheel of the EM endcap calorimeter. The FCal consists of three modules in each endcap. One is made of copper and is hence optimised for electromagnetic measurements whereas the other two are made of tungsten that is more suited to hadronic calorimetry. The total depth of the FCal is approximately ten interaction lengths.

3.5 Muon system

The muon spectrometer forms the outermost layer of detectors around the IP, designed to measure the transverse momentum of charged particles with $|\eta| \leq 2.7$ that have passed through the calorimeters by measuring the curvature of their path as they pass though a non-uniform toroidal magnetic field. Due to the statistical nature of the energy deposition in the calorimeters there are rare so-called 'punch through' hadrons. These charged particles are almost always muons since they lose far less energy in the calorimeters.

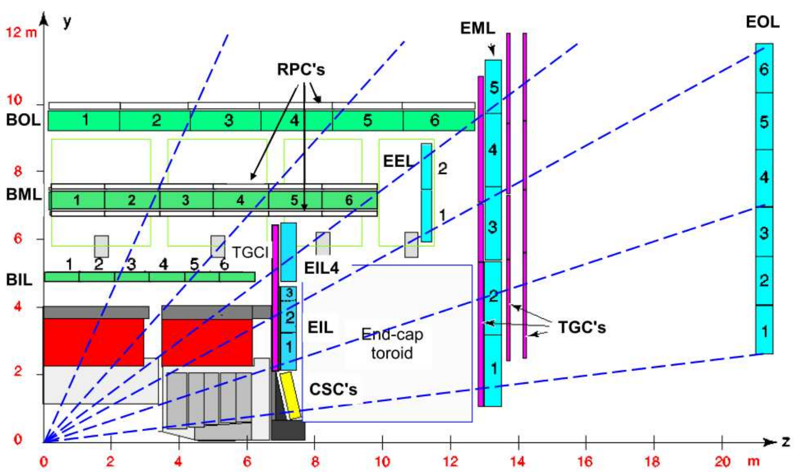


Figure 3.8: Cross section of the ATLAS Muon Spectrometer detectors in the $R - z$ plane. Infinite momentum muons would travel along straight trajectories such as the dashed lines. The letters B and E of the MDT naming scheme refer to the barrel and endcap chambers, respectively. The second and third letters refer to the layer (inner, middle and outer) and sector type (large and small), respectively (taken from reference [27]).

The muon spectrometer is designed to provide a transverse momentum resolution of 10%

for muon tracks with $p_T \simeq 1$ TeV and make track p_T measurements and identify the charge of tracks with $p_T \leq 3$ TeV. High p_T leptons are a canonical event signature for many new physics models, including several Higgs boson analyses such as $H \to \tau\tau \to \ell\nu\nu\ \tau_h\nu$.

The layout of the muon spectrometer sub-detectors is shown in Figure 3.8. Precision track p_T measurements are made by the monitored drift tube (MDT) chambers in the $R-z$ plane with $|\eta| \leq 2.7$, the basic element of which is a pressurised drift tube with diameter 29.97 mm filled with an Ar/CO_2 gas mixture at 3 bar. An MDT chamber is composed of three to eight tubes with an average spatial resolution of 80 μm per tube. The innermost MDT barrel wheels are replaced by Cathode Strip Chambers (CSC) with a higher spatial resolution than the MDTs to better cope with the high particle fluxes at high $|\eta|$.

As charged particles traverse each tube the electrons produced from the gas ionisation are collected by a central tungsten wire held at a potential of 3080 V. The CSCs are multi-wire proportional chambers with radially aligned anode wires and perpendicular cathode strips that record hits by interpolating the charge from gas ionisation on adjacent cathode strips.

The muon spectrometer also has detectors with a lower p_T resolution but with a much faster timing resolution of ≈ 4 ns used to make fast track p_T measurements up to $|\eta| < 2.4$ in the muon trigger. Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) are also present in the barrel and endcap regions, respectively.

3.6 Trigger

The ATLAS detector is designed to record far fewer events for offline analysis than there are bunch crossings at even modest instantaneous luminosities. An event is categorised as interesting and hence worthy of recording offline by a three level system known as the 'Trigger'. The Trigger is composed of a fast, online 'level 1' (L1) selection algorithm that passes event decisions to an offline 'high level trigger' (HLT) consisting of a 'level 2' (L2) and 'Event Filter' (EF) stage. When operating at the design luminosity of 10^{34} cm^{-2}s^{-1}, the event rate is expected to be approximately 1 GHz, of which about 200 Hz can be recorded offline. This requires the three stage trigger to reject approximately 5×10^6 events per second. Since bunch crossings at the IP can occur every 25 ns, the system must make the decision quickly. A diagram of the ATLAS trigger and data flow model is shown in Figure 3.9.

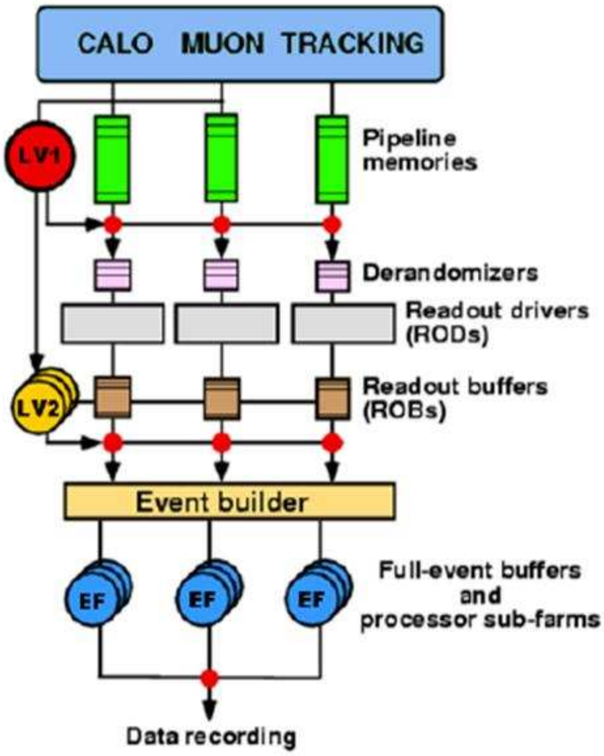


Figure 3.9: Diagram showing the ATLAS trigger decision chain and data-flow model (taken from reference [30]).

For many analyses (including those described in this thesis), the hardware-based L1 trigger system is used to find high p_{T} electrons, photons and muons. A reduced subset of the detectors including all the calorimeters and the muon RPCs and TPCs (but excluding the inner detector) is used to provide the L1 with this information with a lower resolution than is used in the offline selection. If a reconstructed object passing some p_{T} threshold is found, this event is passed on to the HLT, otherwise it is discarded and no further attempt is made to record the event information.

3.6.1 The L1 trigger

After each bunch crossing, the detector information is time stamped and buffered into the pipeline memories located on the readout electronics (see Figure 3.9). The L1 trigger decision must reach the electronics within 2.5 μs of the bunch crossing. The detector electronics can handle a maximum event rate of 75 kHz.

Events selected by an L1 trigger are read out from the detector electronics into readout drivers (RODs) and then into readout buffers (ROBs) of which there are about 1700 in total. Intermediate buffers ('derandomisers' in Figure 3.9) are used to ensure the data can be read out of the pipeline memories with the available bandwidth of the RODs.

3.6.2 The HLT

The HLT is entirely software based and processed on a dedicated farm of around 2000 computing elements using mostly commercially available hardware. Both the L2 and EF trigger algorithms use the full granularity of the calorimeter and muon sub-detectors combined with limited information from the inner detector.

All the detector data for a bunch crossing selected by the L1 trigger are held in the ROBs (see Figure 3.9) until they are either rejected by the L2 trigger or transferred to a storage area associated with the EF trigger. The L2 trigger receives so-called 'Regions of Interest' from the L1 algorithms that are geometric regions of the detector in which L1 physics objects have been identified. The L2 trigger runs more computationally intensive algorithms to further identify these objects with an average processing time of about 40 ms to reduce the event rate fed into the EF algorithms to about 3.5 kHz. The EF is a combination of offline algorithms that utilise the full detector information to decide which events are to be stored for further offline analysis at an expected rate of about 200 Hz[5], with an average processing time of about four seconds per event per computing node.

3.6.3 Data distribution

The ATLAS Data Acquisition System (DAQ) handles the intermediate buffering and data distribution. Events selected by EF trigger algorithms are written to a permanent storage

[5]In 2011 the EF trigger was often run at a highter rate, typically 300 Hz.

site (Tier 0) located in the CERN computing centre. A second copy of every event selected by the EF algorithms is also stored at one of ten ATLAS Tier-1 sites, external to CERN to ensure data security. These data are later processed using higher level event reconstruction algorithms and stored in formats optimised for efficient user analysis and distributed via the 'GRID' using the file format of the `ROOT` analysis framework [31]. The GRID is a very large distributed network of computing and storage elements available to users for rapid data analysis through parallelisation of the analysis jobs.

3.7 Event reconstruction

Physics analyses of ATLAS data are based on the reconstruction of physics objects from detector information in the form of signals from the detector read-out electronics.

From the inner detector, hits recorded from the sub-detectors are reconstructed into helical tracks, the radius of which is determined by the particle's p_{T} and the magnitude of the magnetic field produced by the solenoid. These tracks are then extrapolated to the IP to reconstruct vertices, i.e. the common origin of two or more inner detector tracks. Electrons, photons, jets and hadronically decaying tau leptons are reconstructed using the tracks from the inner detector combined with energy measurements made in the calorimeters. Using the reconstructed vertex information, jets are further classified by a hypothesis of their origin from either a b or c hadron or a hadronic jet initiated by a light parton (u ,d , s quark or gluon).

To reconstruct muons, inner detector tracks are extrapolated to tracks reconstructed in the muon spectrometer. Since the incoming protons have no transverse momentum, the vector p_{T} sum of all objects measured in the calorimeters and muon spectrometer is also used to measure the missing momentum or $p_{\mathrm{T}}^{\mathrm{miss}}$ vector. This is a measure of the negative vector p_{T} sum of all particles produced in each collision that do not interact with any of the ATLAS sub-detectors, for example neutrinos produced in weak boson decay or uncharged particles from new physics models.

3.7.1 Inner detector track and vertex reconstruction

Inner detector tracks are reconstructed in an algorithm with three steps. First, hits in the silicon detectors and the TRT are used to build space-points. Second, the default tracking exploits the high granularity of the Pixel and SCT detectors to find tracks that can be extrapolated to the interaction region. In this step, track seeds are formed using space-points from the three pixel layers and the first SCT layer. These seed tracks are then extended to hits in the other SCT layers to form candidate tracks. Following this, each candidate is fitted to a hypothesis helix, outlier space-points are removed and any space-point-to-track ambiguities are resolved and fake tracks discarded by applying quality requirements. The surviving track candidates are then extended into the TRT and refitted using the full information of all three detectors.

Each reconstructed track can be parametrised by five quantities defined at the point at which the track is closest to the centre of the coordinate system. The quantities z_0 and d_0 denote the displacements between the centre of the detector and the point of closest approach along the beam axis and in the transverse plane, respectively. The angles ϕ and $\cot\theta$ are the corresponding azimuthal angle and the cotangent of the polar angle. The other quantity is the 'curvature' of the helix, the inverse of the helix radius in the transverse plane, which is derived from the charge over the transverse momentum of the track, q/p_{T}. The measurement of these helix parameters is dependent on the number of multiple scattering interactions of the charged particle, and hence the amount of material traversed by the particle in the inner detector. The resolution of each parameter can be parametrised as a function of the track p_{T} by

$$\sigma(p_{\mathrm{T}}) = \sigma(\infty)\left(1 \oplus \frac{p'}{p_{\mathrm{T}}}\right), \tag{3.4}$$

where p' is the p_{T} at which the intrinsic and multiple scattering terms are equal for each parameter and $\sigma(\infty)$ is the resolution expected for a track of infinite momentum. Table 3.1 shows $\sigma(\infty)$ and p' for each helix parameter in two regions of the detector in which different amounts of material have been traversed.

Helix parameter	$0.25 < \|\eta\| < 0.75$		$1.5 < \|\eta\| < 1.75$	
	$\sigma(\infty)$	p' (GeV)	$\sigma(\infty)$	p' (GeV)
q/p_{T}	0.34 TeV^{-1}	44	0.41 TeV^{-1}	80
ϕ	70 μrad	39	92 μrad	49
$\cot\theta$	0.7×10^{-3}	5	1.2×10^{-3}	10
d_0	10 μm	14	12 μm	20
$z_0 \times \sin\theta$	91 μm	2.3	71 μm	3.7

Table 3.1: Track parameter resolutions (RMS) for two η regions with near minimal and maximal amounts of material traversed by charged particles through the inner detector. The momentum and angular resolutions are for muons and the impact-parameter resolutions are shown for pions.

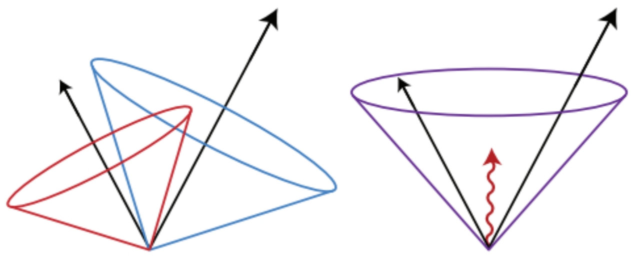
At the LHC, the nominal size of the bunches is $\sigma_{xy} = 15$ μm and $\sigma_z = 5.6$ cm, and so determining the coordinates of the primary interaction vertex (PV) along the z axis requires vertex reconstruction. In vertex reconstruction, reconstructed tracks are associated with a candidate vertex. In the vertex fitting step, the position of the PV is determined and parameters of the tracks of the associated tracks are recalculated, using constraints from the PV position. There are several vertex finding algorithms used in ATLAS but each is essentially finding the minimum of a χ^2 function using the vertex position and the parameters of the tracks at a chosen vertex position.

3.7.2 Jet finding and heavy flavour tagging

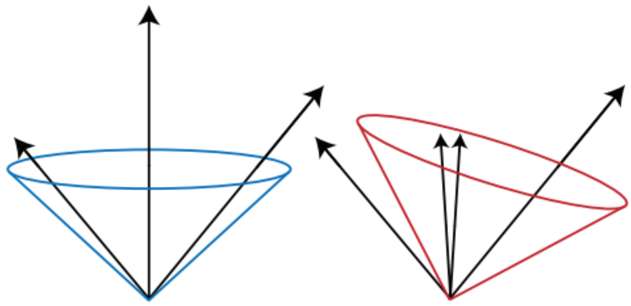
To reconstruct the final state partons produced in each collision, jet finding algorithms are applied to the energy deposited in the calorimeters by the products of the parton hadronisation. In this thesis, every analysis uses an implementation of a sequential recombination algorithm, the anti-k_t algorithm [32] that is both infrared and collinear safe, as shown in Figure 3.10.

The procedure for jet finding is:

- All calorimeter deposits are converted into a list of massless four-vectors that are combined into clusters using the η, ϕ and E of the cluster. Clusters are seeded using cells that are above a threshold signal-to-noise ratio, $\Gamma = E_{cell}/\sigma_{cell}^{noise}$ and combined



(a)



(b)

Figure 3.10: Problems that can arise during jet finding: (a) Infrared safety: soft particle (gluon) emission should not affect the result of the algorithm, (b) Collinear safety: one parton splitting into two should not affect the result.

with adjacent cells if a second, lower signal-to-noise ratio is exceeded for the combined object.

- A comparison is then made of the inverse square of the transverse momenta, $d_i = (1/p_{\mathrm{T}}^2)_i$, of all objects in the list with the inverse square of the transverse momentum of every other object in the list to form pairs defined by

$$d_{ij} = \min[(1/p_{\mathrm{T}}^2)_i, (1/p_{\mathrm{T}}^2)_j] \times \frac{\Delta R_{ij}^2}{R^2} \quad \text{with} \quad \Delta R_{ij}^2 = \Delta\phi_{ij}^2 + \Delta\eta_{ij}^2 \tag{3.5}$$

  where R is a free parameter that determines the size of the jets, chosen in these analyses to be 0.4.

- The minimum of every d_{ij} pair and d_i of all objects in the list is found, d_{ij}^{min}. If d_{ij}^{min} is a d_{ij} pair in the list, the two elements are removed from the list and combined into a new object that is added back into the list. Otherwise the object associated with d_{ij}^{min} is removed from the list and considered to be a jet. The step is repeated until the object list is empty.

Clusters energies (and hence the final jet objects) are initially defined at the electromagnetic energy scale[6] (EM scale)[33]. Corrections are made to account for the energy loss by particles traversing material in front of the calorimeter system and for deposits missed by the clustering algorithm.

In ATLAS, jet reconstruction is performed with a threshold of jet $p_{\mathrm{T}} > 7$ GeV. The free parameter R can be optimised for each analysis: small cone sizes acquire less contamination from objects close together in busy environments while larger cone sizes allow for a more precise energy resolution. Further quality criteria are required of the jet candidates in the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ analyses (see Chapter 7).

3.7.3 b-jet tagging

Hadronic jets originating from b-quarks can be experimentally distinguished from jets originating from light partons. b-hadrons have a relatively long lifetime of $\tau \approx 1.5$ ps, so b-hadrons with high p_{T} typically have a flight path in the transverse plane of several millimetres. The secondary vertex of the resulting jet can be identified by either reconstructing the decay vertex or combining the impact parameters of the charged hadron tracks into a discriminant [34].

3.7.4 Electrons and photons

In order to reconstruct electrons and photons, a 'sliding window' algorithm is used to look for EM calorimeter clusters produced by the electromagnetic particle showers. The algorithm is split into three steps: tower building, pre-clustering and cluster filling.

Initially, the calorimeter is segmented into a rectangular grid with $\Delta\eta\times\Delta\phi = 0.025\times0.025$ resolution. The window size of each reconstruction algorithm is shown in Table 3.2.

The energies deposited in every layer of a given geometrical unit of this grid are then

[6] this is defined using the measured calorimeter response to an electron test-beam. Corrections are applied to account for the differences in the energy cluster deposit shapes of hadronic jets and electrons.

Reconstruction algorithm	Barrel	Endcap
Electron	3×7	5×5
Conversion	3×7	5×5
Photon	3×5	5×5

Table 3.2: Cluster window sizes in $N_{\eta}^{tower} \times N_{\phi}^{tower} = 0.025 \times 0.025$ units for the middle EM calorimeter cell layer.

summed to form an energy 'tower'. For all towers with $E_T > 2$ GeV, the electron reconstruction algorithm then attempts to match an inner detector track within a $|\Delta\eta \times \Delta\phi|$ window of 0.05×0.1. The ratio of tower energy to track momentum, E/p, must be less than 10, and the tracks must not be consistent with $\gamma \rightarrow e^{+}e^{-}$ conversions. Energy corrections are made to account for energy loss due to bremsstrahlung in the inner detector. An electron candidate is created if an energy tower can be matched to a track, otherwise it is classified as a photon. In this way, approximately 93% of true, isolated electrons with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ are reconstructed as electron candidates. Since on average an electron candidate will have shed 20% - 50% [27] of its energy (dependent on its $|\eta|$) after passing through the SCT due to bremsstrahlung and multiple scattering, energy calibrations are then applied to account for these additional particles which are collinear with the electron candidate.

Following this, the calibrated electron candidates are then subjected to dedicated identification algorithms to provide separation of true electron and photon candidates from hadronic backgrounds such as charged pions [35]. For electron candidates, three levels of quality, referred to as 'loose', 'medium' and 'tight', correspond to decreasing levels of signal efficiency and simultaneously increasing levels of background rejection. The selection criteria for each level have been simultaneously optimised for up to seven η bins and up to six p_{T} bins. The selection criteria for each are as follows:

- The 'loose' identification criterion applies cuts on the total energy deposited in the candidate tower cells of the first layer of the hadronic calorimeter and on the lateral shower shape and shower width using information from the middle layer of the EM calorimeter.

- The 'medium' identification criterion requires the same selection as the 'loose' cuts but makes additional cuts on the inner detector track quality and uses information about the energy deposited in the first layer of the EM calorimeter strips. These cuts are optimised to separate single charged pions or electron clusters from electron/positron pairs from π^0 conversions. Conversions cause a specific energy-deposit pattern with two maxima that can be resolved in this strip layer if a $|\Delta\eta \times \Delta\phi| = 0.125 \times 0.2$ window around the tower cell with the highest E_{T} is considered. Therefore, such clusters can be rejected as electron candidates if they contain multiple maxima. The track quality is defined in terms of the number of pixel, SCT and TRT hits associated with the track and its impact parameters. The medium cuts increase the background rejection by about a factor of 4 with respect to the loose cuts while reducing the signal efficiency by about 10%.

- The 'tight' criterion requires candidates to pass the medium selection cuts and further requires a hit on the innermost Pixel detector layer to further reject conversion backgrounds. A cut is also made of the number of TRT hits (and the number of high threshold TRT hits) and a tighter cluster-to-track matching criterion is applied to help remove pion backgrounds.

Finally, a further cut is placed on the so called relative 'isolation energy' of the electron candidate which is defined as the energy sum of all additional EM calorimeter energy clusters within a cone of $\Delta R \leq 0.2$ of the candidate divided by the E_{T} of the candidate, since hadronic backgrounds will usually be part of a wider hadronic jet. A similar cut is made relative to the candidate p_{T} rather than an absolute cut on the isolation energy as it is more independent of the number of pileup interactions.

3.7.5 Muons

Two muon reconstruction and identification algorithms are used in ATLAS. So-called 'standalone' muon candidates are defined by a reconstructed track in the muon spectrometer with $|\eta| < 2.7$ extrapolated back to the beam line. 'Combined' muon candidates match reconstructed tracks in the muon spectrometer to reconstructed tracks in the inner detector and hence the pseudo-rapidity range is limited by the inner detector to $|\eta| < 2.5$. Despite the lower fiducial volume, the analyses described in Chapter 7 use combined muon candidates for the following reasons:

- The muon spectrometer has gaps in detector acceptance in several layers at $|\eta| \sim 0$ and $|\eta| \sim 1.2$ that reduce signal acceptance. Muons in this region can be better reconstructed using an inner detector track matched to a partially reconstructed muon detector track.

- Muon candidates with very low p_{T} (less than a few GeV) will generally not traverse the outermost muon spectrometer detector layers making them more difficult to reconstruct and identify without inner detector information.

- Muons produced from the weak decay of neutral mesons outside of the inner detector are often reconstructed by the standalone algorithms and constitute an additional source of physics background that is drastically reduced when using combined muon candidates.

By using measurements from both detectors, the p_{T} resolution is also improved. In the combined muon algorithm, a χ^2 function is used to define how well a muon spectrometer track is matched to an inner detector track based on the outer and inner track segments or by partially refitting the helix parameters starting from the inner detector track and adding measurements from the muon spectrometer track, taking into account energy losses in material between the inner detector and muon spectrometer. Typically, the combined reconstruction algorithm will have a signal efficiency of about 94% for simulated single muons in a sample of the leptonic W boson decay and $t\bar{t}$ production [27].

3.7.6 Hadronically decaying tau leptons

The first step in reconstructing hadronic tau decays (τ_h) is to 'seed' each reconstructed jet[7] that has $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$ as a candidate τ_h. Following this, the calorimeter clusters associated with the seed jet are refined and used to calculate kinematic quantities. Tracks reconstructed with the inner detector are associated with a τ_h candidate if they are reconstructed within a cone of radius $\Delta R = 0.2$ from the seed jet axis and have:

- $p_{\mathrm{T}} > 1$ GeV,
- ≥ 2 pixel hits and ≥ 7 silicon tracker hits and
- $|d_0| < 1.0$ mm and $|z_0 \sin\theta| < 1.5$ mm,

where d_0 and $|z_0 \sin\theta|$ are the track helix parameters described in Section 3.7.1. Tau candidates are then classified by the number of associated tracks, i.e. either single-prong or multi-prong for candidates with 1 or ≥ 1 track, respectively. A set of variables is then calculated from the tracking and calorimeter information. The variables are designed to provide good separation of hadronic tau decays from both hadronic jets produced in QCD interactions and electrons. The reconstructed variables are used to create a multivariate discriminant to reject backgrounds while accepting true τ_h.

A description of the τ_h identification algorithms and a data-driven analysis measuring the τ_h mis-identification rate at ATLAS are given in Chapter 6.

3.7.7 Missing transverse momentum

A large imbalance of transverse energy deposits in the calorimeters indicates the production of particles which have passed through the detectors without any interaction, such as neutrinos produced in the leptonic decay of a W boson. The missing transverse momentum is

[7]The jet reconstruction process is described in Section 3.7.2. For τ_h seed jets, a distance parameter of $R = 0.4$ is used.

thus defined:

$$\vec{p}_T^{miss} = -\sum \vec{p}_T \quad \text{and} \quad p_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{\vec{p}_T^{miss} \cdot \vec{p}_T^{miss}}. \tag{3.6}$$

In this definition, $\sum \vec{p}_T$ includes all energy deposits in the calorimeters and the $\vec{p}_T$ of reconstructed combined muon candidates. Corrections are applied for energy clusters associated with identified electrons, photons, muons, τ_h and jets [36]. Corrections are also made to account for gaps in detector acceptance and additional detector material. Noisy electronics in the calorimeters are also taken into account by only considering deposits above a cell's noise threshold and using clusters to which noise suppression cuts have also been applied, as described in reference [27].

3.8 Detector simulation

In order for the Monte Carlo (MC) samples to provide an accurate description of what is observed in data, generated MC events must take into account the response of the detectors. The ATLAS detector simulation is performed with the `GEANT4` [37] program using a detailed description of the material distributed in the ATLAS detector to simulate the full detector response, i.e. the signals provided by the detector electronics of the individual sub-detector modules. The particles produced in a MC event are propagated through the detectors and the simulation of the detector response is modelled using the `GEANT4` program. The simulated detector signals are then passed through the full event reconstruction process, in the same way as is done for the real data taken from the experiment. Therefore, detailed studies of e.g. the electron and photon shower shapes and reconstruction efficiency are possible using simulated samples. However, it is often possible to calibrate such quantities in a data-driven way that is independent of MC, in case of any generator or simulation mis-modeling.

To study some systematic uncertainties, additional samples were created with deliberate detector misalignment introduced and slight distortions in the solenoid / toroidal magnetic fields to model the effect of a symmetry axis not coincident with the z-axis. Similarly, MC samples were simulated with additional non-active material in front of the calorimeters to better estimate the jet energy scale.

Chapter 4

LHSee

4.1 Introduction

Often in analysis, visual investigation presents a unique way in which to understand an aspect of detector performance or an event topology. With this in mind, a new tool that allows analysers to visualise ATLAS events on mobile platforms and devices has been developed.

4.2 Visualising ATLAS events on a mobile platform

LHSee [38] is an event display package developed to provide an interactive, visual investigation of both the ATLAS detector and the high-energy physics events it records. It is designed to be used on any mobile device using the Android operating system [39], including hand-held phones and tablet devices, with an intuitive and user-friendly interface to produce both 2-dimensional (2-D) projected visualisations and fully 3-dimensional (3-D) event displays which the average physicist or educator can easily interpret. LHSee implements several of the techniques first developed in the ATLAS event display programme `ATLANTIS` [40].

4.3 3-D Techniques

LHSee uses an implementation of the 3-D graphical application programming interface `OpenGLES` [41] to render both the ATLAS detector and the event information as a 3-D visualisation. The ATLAS detector can be visualised by first defining a set of primitive geometrical shapes that can be thought of as embedded graphs[1]. These graphs, shown in Figure 4.1, are then used with specific detector geometry information (first implemented in `ATLANTIS` [40]) to build more complicated graphs that describe the various ATLAS sub-detectors. In total, $\mathcal{O}(1000)$ vertices and edges are used to render the ATLAS detector in each frame.

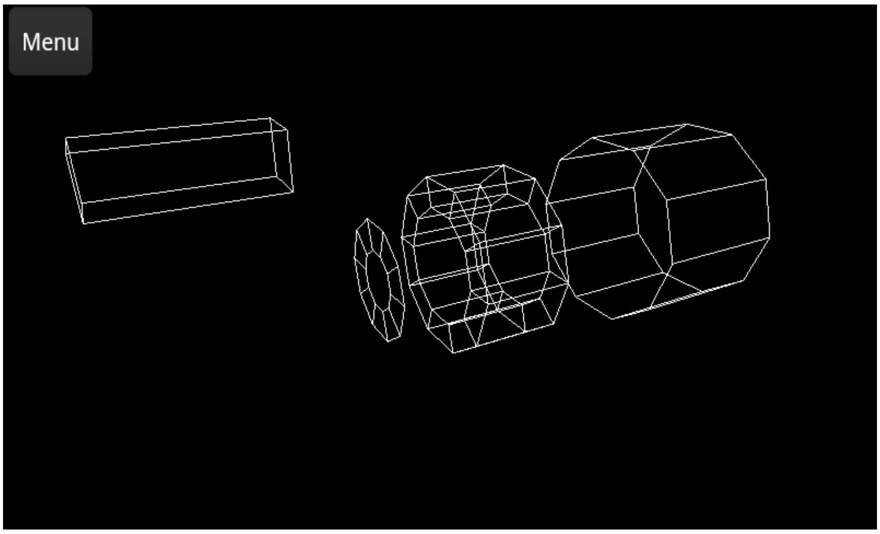


Figure 4.1: Primitive graphs used to construct detector geometry: a Cuboid, an Annulus, a Toroid and a Cylinder.

4.4 Visualising the detectors of the ATLAS experiment

The ATLAS detector and coordinate system are described in Chapter 3. LHSee can display projections of the various ATLAS sub-detectors that make up the inner detector in the

[1]A graph is composed of a set of vertices connected by a set of edges. For an embedded graph, each vertex also has a position relative to the centre of the coordinate system.

plane parallel to the incoming protons (the $z - y$ plane) and the plane perpendicular to the incoming protons (the $x - y$ plane) as shown in Figure 4.2 and Figure 4.3, respectively. Alternatively, these detectors can also be visualised in 3-D, as shown in Figure 4.4, using the method described in Section 4.3.

The Electromagnetic and Hadronic calorimeters are represented as radial barrel components along the z-axis enclosing the inner detector, flanked on either side by endcap rings oriented perpendicular to the z-axis. This is shown in Figure 4.5, Figure 4.6 and Figure 4.7 with the Electromagnetic calorimeter in green, and the Hadronic calorimeter in red.

The sub-detector systems that make up the ATLAS Muon Spectrometer are collectively displayed in blue in both the 2-D projections and the 3-D visualisation, as shown in Figure 4.8 and Figure 4.9.

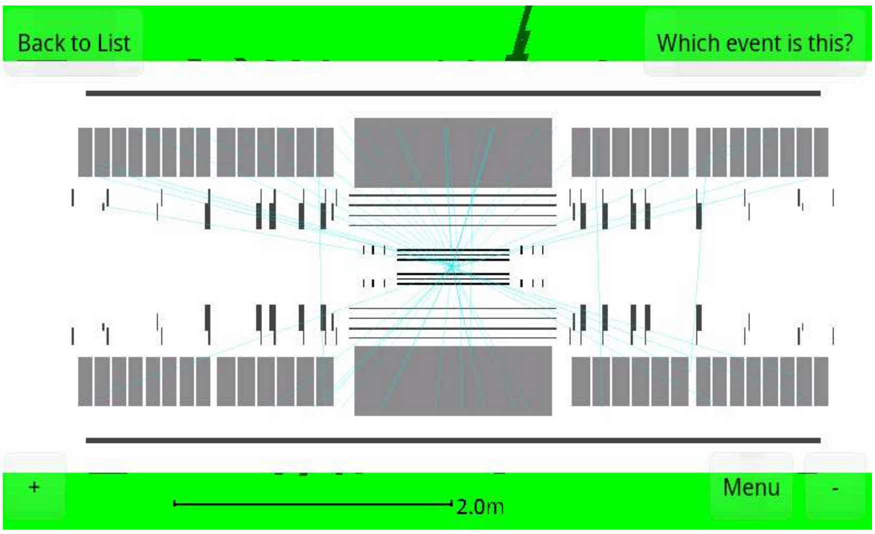


Figure 4.2: Projection of the inner detector in the $z - y$ plane. The silicon layers of the Pixel and SCT are shown in black and the TRT sections are shown in grey. The fitted inner detector tracks of a simulated QCD di-jet event are overlaid in cyan, showing their paths through the detectors.

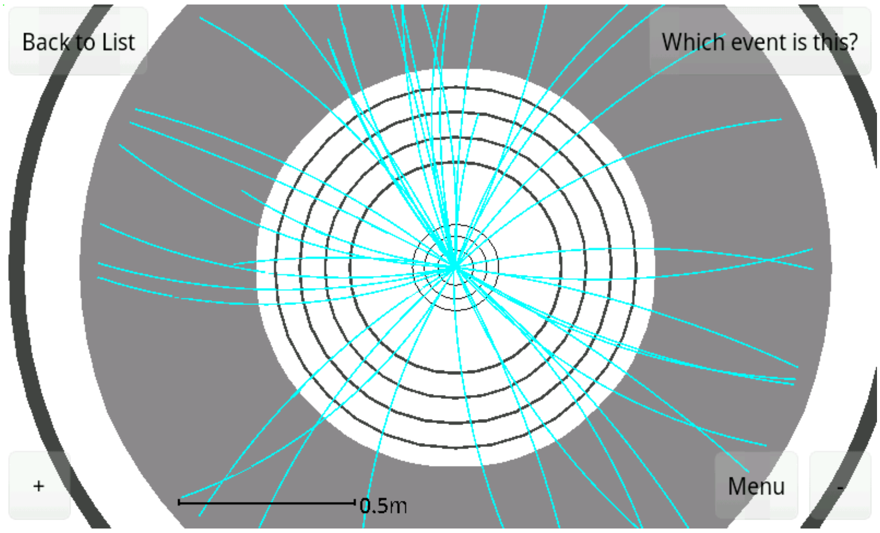


Figure 4.3: Projection of the inner detector in the $x - y$ plane. The silicon layers of the Pixel and SCT are shown in black and the TRT sections are shown in grey. The fitted inner detector tracks of a simulated QCD di-jet event are overlaid in cyan, showing their paths through the detectors.

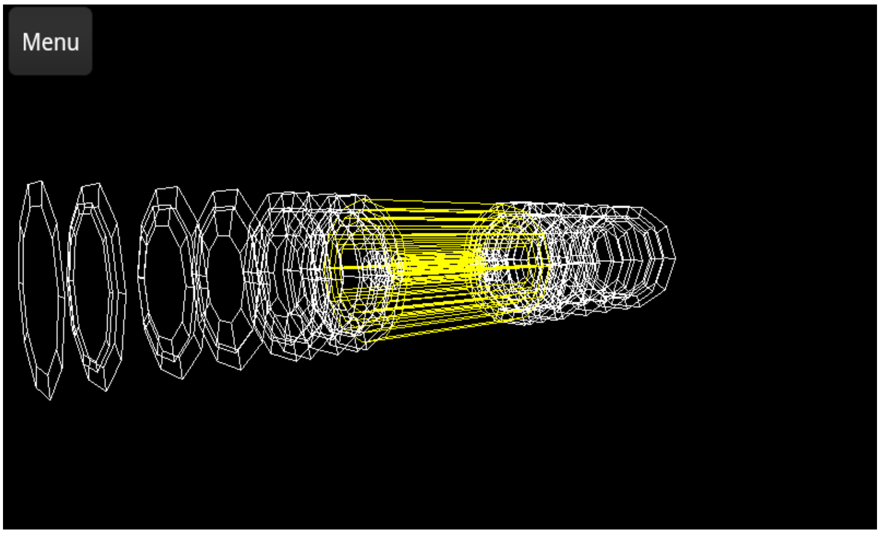


Figure 4.4: 3-D visualisation of the inner detector Pixel barrel section (yellow) and Endcap section (white).

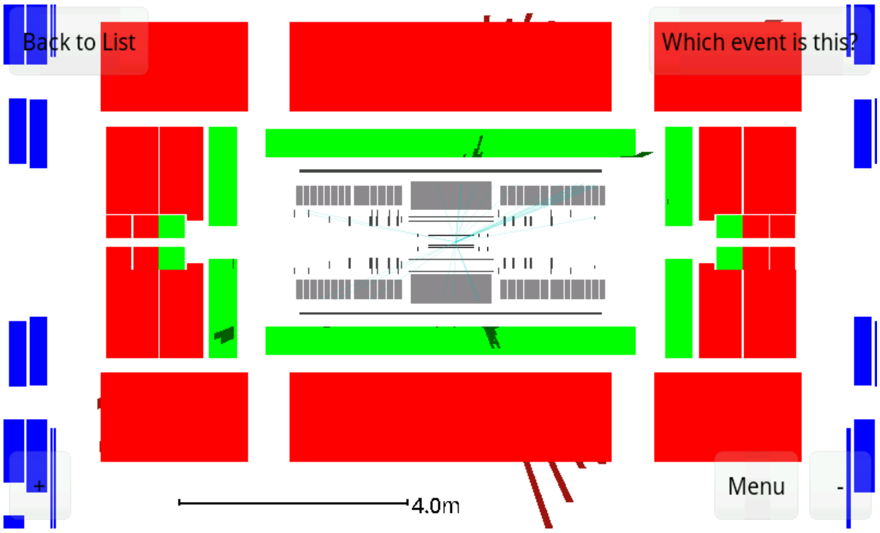


Figure 4.5: $z - y$ plane projection showing the Electromagnetic calorimeters (green) and Hadronic calorimeters (red) and transverse energy histograms for calorimeter cells with $E_T >$ 0.25 GeV from a simulated QCD di-jet event in dark green and dark red, respectively.

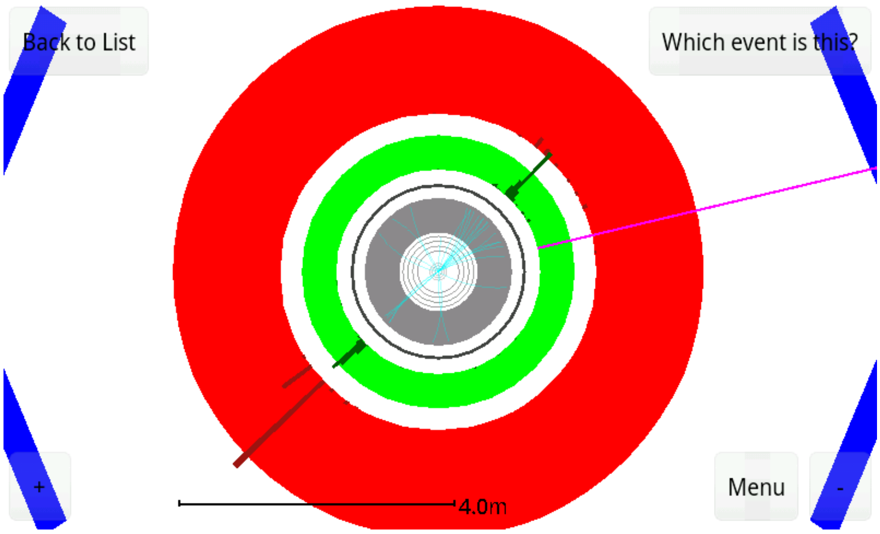


Figure 4.6: $x - y$ plane projection showing the Electromagnetic calorimeters (green) and Hadronic calorimeters (red) and transverse energy histograms for calorimeter cells with $E_T >$ 0.25 GeV from a simulated QCD di-jet event in dark green and dark red, respectively.

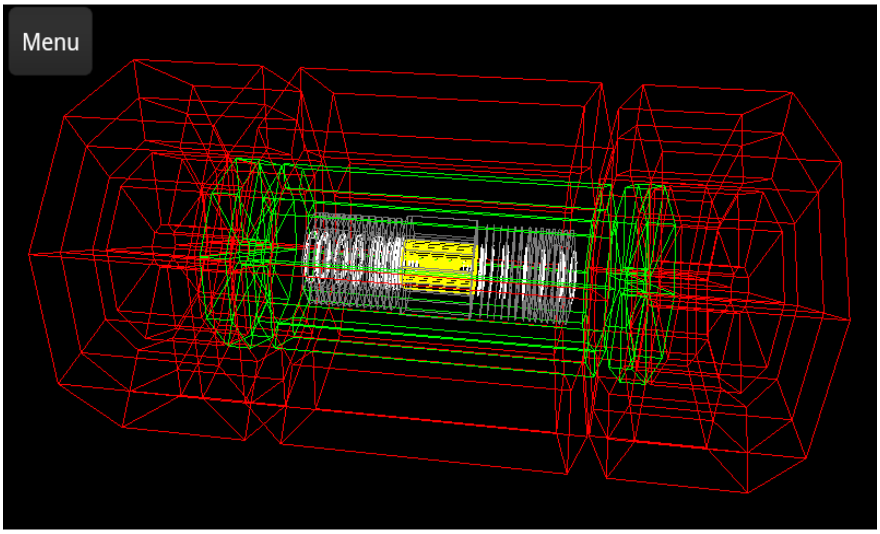


Figure 4.7: 3-D visualisation showing the calorimeters. The EM calorimeters are displayed in green, and the hadronic calorimeters are shown in red.

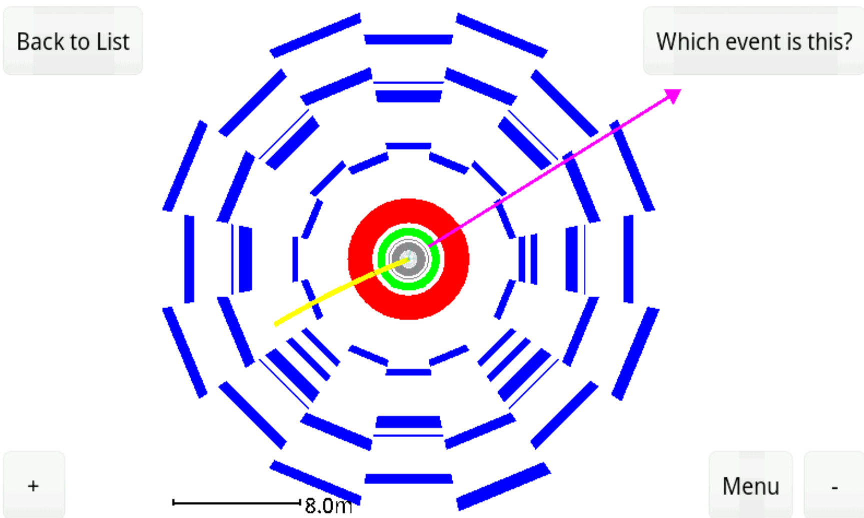


Figure 4.8: $x - y$ plane projection showing all detectors overlaid with a simulated $W \rightarrow \mu\nu$ event, with muon tracks in yellow and $\vec{p}_T^{miss}$ shown in magenta.

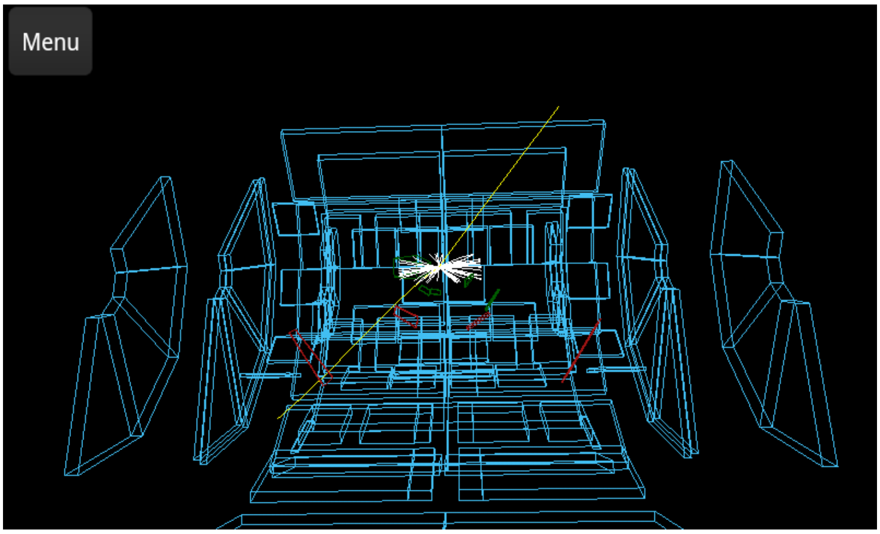


Figure 4.9: 3-D visualisation showing the muon detectors. Muon tracks are are shown in yellow from a simulated $Z \rightarrow \mu^+\mu^-$ event.

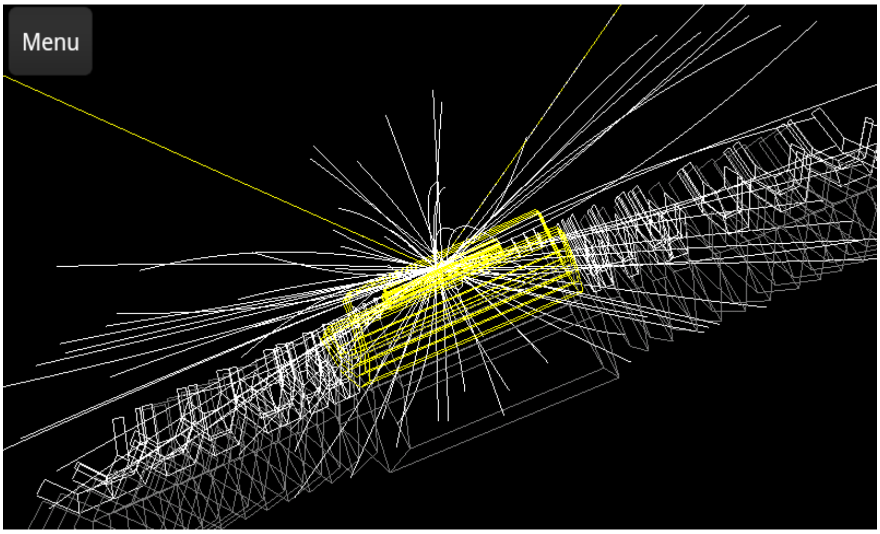


Figure 4.10: 3-D visualisation of the inner detector tracks. Only sub-detectors with $x < 0$ are shown.

4.5 Event Visualisation

4.5.1 Tracks

A charged particle passing through the inner detector travels along a curved path that can accurately be described by a helix, the radius of which is determined by the strength of the solenoidal magnetic field, the particle's charge and its transverse momentum. To visualise charged particle tracks, points lying on each track helix are generated using a set of helix equations that can be parametrised for each track. The track helices are drawn from the closest point on the helix to the centre of the coordinate system to the point at which it leaves the inner detector fiducial region. Figure 4.3 and Figure 4.10 show the tracks from a simulated event in the projected plane transverse to the incoming protons and in 3-D in cyan and white, respectively.

Figure 4.11 shows the projection of a track helix in the $x - y$ plane (i.e. a circle). The closest point on this circle to the centre of the coordinate system, P_0, is P_d with length D_0 from P_0. A negatively charged particle travels along this circle in the anti-clockwise direction, initially in the direction of ϕ_0. The direction of the vertex as seen from the centre of the coordinate system P_0 is $\phi_0 + \frac{\pi}{2}$.

First, let us assume that the length of $D_0 = 0$, and hence $P_d = P_0$. The negatively charged particle will travel anti-clockwise along the circle with $d\alpha > 0$ and $d\alpha$ increasing as it does so. This is defined as a positive turning circle with sign $s = +1$. A positively charged particle would have $s = -1$.

In a homogeneous solenoidal field, the radius of the circle, R, is inversely proportional to the strength of the magnetic field, B. According to the Lorentz force, R is given by

$$R = \frac{p_{\mathrm{T}}}{|e|B} \qquad \text{and} \qquad R_0 = sR, \tag{4.1}$$

where e is the charge of the particle. The centre of the circle P_c has coordinates (x_c, y_c) that

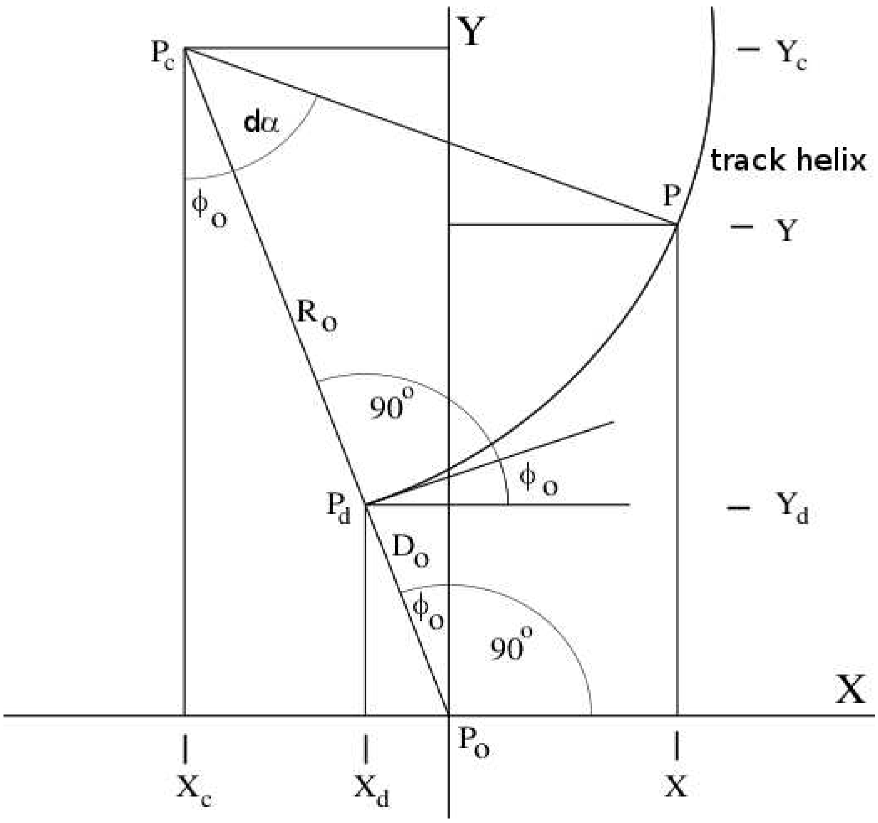


Figure 4.11: Projection of an inner detector track helix in the $x - y$ plane.

are given by

$$x_c = R\cos(\phi_0 + s\frac{\pi}{2}) \qquad \text{and} \qquad y_c = R\sin(\phi_0 + s\frac{\pi}{2}), \tag{4.2}$$

or

$$x_c = -R_0 \sin\phi_0 \qquad \text{and} \qquad y_c = R_0 \cos\phi_0. \tag{4.3}$$

If the circle does not pass through P_0 and instead has a distance of closest approach D_0 (see Figure 4.11), then these coordinates become

$$x_c = -(R_0 + D_0)\sin\phi_0 \qquad \text{and} \qquad y_c = (R_0 + D_0)\cos\phi_0, \tag{4.4}$$

as can be seen in Figure 4.11. The coordinates of the point of closest approach P_d are given by

$$x_d = -D_0 \sin\phi_0 \qquad \text{and} \qquad y_d = +D_0 \cos\phi_0, \tag{4.5}$$

as shown in Figure 4.11. The coordinates of a point P on the circle at $d\alpha$ as seen from P are given by

$$x = x_c + R_0 \sin(\phi_0 + d\alpha) \qquad \text{and} \qquad y = y_c - R_0 \cos(\phi_0 + d\alpha). \tag{4.6}$$

To turn this circle into a helix, a third coordinate is calculated in the z plane for every point on the circle, and is defined to be

$$z = z_0 + \cot\theta_0 \cdot R_0 \cdot d\alpha, \tag{4.7}$$

where z_0 is the z coordinate of P_d and the tangent of the helix and the $x - y$ plane is $\tan\theta_0 = p_Z/p_{\mathrm{T}}$.

An embedded graph is created for each track using the five helix parameters (p_{T}, e, $\cot\theta_0$, ϕ_0 and D_0) in the following way:

1. The radius of the track helix (and hence R_0) is calculated using p_{T} and e.
2. R_0 is used to calculate the coordinates of P_d using ϕ_0 and D_0.
3. The boundary value of $d\alpha$ where the helix leaves the fiducial volume of the inner detector, $d\alpha'$, is calculated iteratively using a recursive algorithm. This boundary is defined by a cylinder along the z-axis with the outermost radius and length of the TRT.
4. Using Equations 4.6 and 4.7 (and hence $\cot\theta_0$), a set of points is generated along the helix from the point of closest approach P_d to this boundary, with $\min(0, d\alpha') \leq d\alpha \leq \max(0, d\alpha')$. These points are equally spaced in $d\alpha$.
5. Consecutive points are connected by an edge and the track is stored as an embedded graph to be displayed.

This algorithm is used to define a graph for every reconstruced track in the event. Tracks with p_{T} greater than $p_{\mathrm{T}}^{\mathrm{min}}$ (a value specified by the user) can then displayed for any $p_{\mathrm{T}}^{\mathrm{min}} > 0.5$ GeV.

4.5.2 Calorimeter histograms

Projections into the transverse plane of the energy deposited in the cells of the Electromagnetic and Hadronic calorimeters with $E_T^{cell} > 0.25$ GeV are represented as histograms in the planes transverse and parallel to the incoming proton beams. These histograms are superimposed directly on top of the calorimeter visualisations, as shown in Figure 4.5 and Figure 4.6. The missing transverse momentum vector, $\vec{p}_T^{miss}$, is also represented in Figure 4.8 as a magenta arrow pointing along the direction of the missing transverse momentum vector where the width of the arrow is proportional to its magnitude.

4.6 Conclusion

A light-weight event display programme has been developed specifically to be used on mobile platforms that can be used to visually investigate the complex high-energy physics events that are recorded at the ATLAS detector at the LHC experiment and is freely available[2]. The intended primary use is as an educational tool and has been downloaded and installed by over 50,000 users worldwide.

[2]https://market.android.com/details?id=com.lhsee

Chapter 5

Expected Standard Model Higgs boson coupling measurements using the $H \to \tau\tau$ channel

5.1 Introduction

At the LHC, the expected discovery modes for a light SM Higgs boson favoured by global fits of Electroweak data are the subleading decays $H \to \gamma\gamma$, $H \to ZZ$, and $H \to WW$ [15, 16, 17]. However, to identify that a neutral resonance is consistent with a SM Higgs boson, its coupling strengths to other massive SM particles must be measured and compared with the SM predictions.

Extensive studies have been made of Higgs boson measurements at the LHC [42]. In this chapter, a new study in presented of the sensitivity of 100 fb^{-1} of $\sqrt{s} = 14$ TeV LHC data to the associated Higgs boson production processes; WH, ZH, and $t\bar{t}H$, followed by $H \to \tau\tau$ and at least one $W \to l\nu$ or $Z \to ll$ decay. The production diagrams of these processes are shown in Figures 2.4(c) and 2.4(d).

Measurements in the associated production channels can be used to improve the LHC sensitivity to the coupling ratios $g_{t\bar{t}H}/g_{ZZH}$ and $g_{t\bar{t}H}/g_{WWH}$, the relevant ratio is

$$\frac{\sigma(pp \to VH(\to \tau\tau))}{\sigma(pp \to t\bar{t}H(\to \tau\tau))} \propto \frac{\Gamma_V}{\Gamma_t}. \tag{5.1}$$

Furthermore, measurements in the $H \to \tau\tau$ decay channels can be combined with the expected sensitivity of measurements of associated Higgs production cross section in the $b\bar{b}$ decay channel [20, 21] to measure the Yukawa coupling ratio $g_{Hbb}/g_{H\tau\tau}$. This ratio is determined at leading order by the bottom-quark and tau-lepton masses and is sensitive to differences in the source of mass for quarks and leptons [43]. The relevant ratio measurements are

$$\frac{\sigma(pp \to VH(\to \tau\tau))}{\sigma(pp \to VH(\to b\bar{b}))} \propto \frac{\Gamma_\tau}{\Gamma_b} \quad \text{and} \quad \frac{\sigma(pp \to t\bar{t}H(\to \tau\tau))}{\sigma(pp \to t\bar{t}H(\to b\bar{b}))} \propto \frac{\Gamma_\tau}{\Gamma_b}. \tag{5.2}$$

This Chapter is structured as follows: Section 5.2 outlines the procedures for generating and simulating signal and background events; Section 5.3 describes the specific selection and expected signal and background yields for the WH, ZH, and $t\bar{t}H$ processes; Section 5.4 details the cross section determination of each channel; Section 5.5 presents the expected coupling-ratio sensitivities and Section 5.6 summarises the conclusions.

5.2 Event generation and simulation

Monte Carlo samples of signal and background events are used to calculate event selection efficiencies and event yields. To account for detector acceptances and resolutions, a parameteric detector simulation is performed using `Delphes`.

`Sherpa` [44] is used to generate all signal and background MC samples used in Section 5.3 where possible, with parton distribution functions taken from CTEQ6L [45]. In these samples, tau leptons are decayed within `Sherpa`. In the $t\bar{t}H$ analysis (Section 5.3.3), `Alpgen` [46] is used to generate the $t\bar{t}$+2 jets and W+6 jets background processes for the hard process and `Pythia` [47] is used for the hadronisation and showering due to the high multiplicity final states. For samples with jets in the final state, parton jets are included to leading order at the matrix-element level and additional jets modelled by parton showering within `Sherpa` and `Pythia`. The acceptances of Higgs-boson and diboson processes were cross-checked using the `Herwig++` [48] event generator.

5.2.1 Parametric detector simulation

The `Delphes` simulation package [49] is used to model the detector geometric acceptance. Detector parameters are based on the ATLAS detector: the tracker is assumed to reconstruct all charged tracks with $|\eta| < 2.5$ with 100 % efficiency. Calorimeter towers cover the range $|\eta| < 3.0$ with electromagnetic and hadronic tower granularity of $\delta\eta \times \delta\phi \simeq 0.1 \times 0.1$, and cover the region $|\eta| < 4.9$ with coarser granularity. The energy of each stable particle is summed in the calorimeter tower through which it propagates. This energy, E, is then smeared according to Gaussian resolution functions assigned to the central, endcap and forward region electromagnetic calorimeters (EC) and hadronic calorimeters (HC) according to

$$\frac{\sigma_E}{E} = C \oplus \frac{S}{\sqrt{E}} \oplus \frac{N}{E}, \tag{5.3}$$

where E is expressed in units of GeV. The values used for the constant, C, sampling, S, and noise, N term of each sub-detector are chosen to match the expected performance of the ATLAS detector [33] and are shown in Table 5.1 along with the η region covered by each sub-detector.

Detector	$\lvert\eta\rvert$	S (GeV$^{\frac{1}{2}}$)	N (GeV)	C
	$0 - 1.7$	0.101	0	0.0017
EC	$1.7 - 3.2$	0.1	0	0.0017
	$3.2 - 4.9$	0.285	0	0.035
	$0 - 1.7$	0.5205	1.59	0.0302
HC	$1.7 - 3.2$	0.70	0	0.05
	$3.2 - 4.9$	0.942	0	0.075

Table 5.1: The calorimeter resolution parameters defined in Equation 5.3.

The energy and p_{T} of reconstructed electrons, muons, hadronically decaying τ leptons and jets are smeared according to Gaussian resolution functions matched to the performance of the ATLAS detector [33]. The acceptance criteria for these objects are summarised in Table 5.2.

Object	$\lvert\eta^{\max}\rvert$	p_{T}^{min} (GeV)
e	2.5	10
μ	2.7	10
τ_h	2.5	10
Jet	3.5	15

Table 5.2: The acceptance criteria of reconstructed objects.

In addition to the criteria listed in Table 5.2, a jet is only tagged as a hadronically decaying τ lepton if more than 90% of its energy falls within a cone of radius $\delta R < 0.15$, with one or three charged particle track(s) with $p_T > 2$ GeV within a cone of radius $\delta R < 0.4$ from the jet axis.

The reconstruction of jets is performed using information from the calorimeter towers, with the anti-kt algorithm [32]. This algorithm is applied using the `FastJet` package [50], as implemented in `Delphes`. The b-tagging efficiency is assumed to be 60% for all jets with an associated generator level b-quark, with a b-tagging fake-rate of 10% for c-jets and 1% for light-quark or gluon jets. The jet acceptance is chosen to be conservative under high instantaneous luminosity conditions.

5.2.2 Fake-rate application

The dominant background to several signal channels consists of events where a light jet has been mis-identified as an e, μ or τ_h. In accordance with the description of the ATLAS detector given in [33], a separate identification efficiency and jet mis-identification rate is applied for each object. These values are summarised in Table 5.3.

Object	ID efficiency (%)	Jet mis-ID rate (%)
e	64.2	0.0108
μ	94.2	0.169
τ_h	40	1.0

Table 5.3: Object identification and mis-identification rates.

5.2.3 Triggering

To account for the effect of triggering, approximate trigger efficiencies given in [33] are applied for a single light lepton trigger to each analysis channel. The trigger efficiency for each, along with the minimum p_{T} assumed for the trigger are listed in Table 5.4.

Object	Triggering efficiency (%)	Object p_T^{min} (GeV)
e	94.3	25
μ	80	25

Table 5.4: Triggering efficiencies.

5.2.4 Pile-up

At high luminosity, the presence of pile-up events is expected to affect the detector response. At an instantaneous luminosity of $\mathcal{L} = 10^{34}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$, one can expect $\simeq$ 25 additional interactions with $\sum E_T = 30$ GeV per interaction [30], where $\sum E_T$ is the total E_T measured in the calorimeters. Since the $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution of ATLAS is predicted to be $\sim 0.5\sqrt{\sum \mathrm{E_T}}$ [33], we smear the projections p_x^{miss} and p_y^{miss} with a Gaussian term, with a width of 10 GeV or 15 GeV. The two choices correspond to an optimistic and nominal expectation of the $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution, respectively.

It is also expected that at this high instantaneous luminosity the tau-ID efficiency will be significantly degraded. We therefore also use two scenarios of tau-ID performance: 40% efficient and 28% efficient, with a corresponding tau fake-rate of 1% [51].

5.3 Event selection and sensitivity

Subdividing the three production channels WH, ZH and $t\bar{t}H$ by the decay of the tau leptons originating from the Higgs boson (i.e. hadronic or leptonic tau decay) leads to nine signal channels. For each decay channel, an event selection is applied to suppress the background in that channel. A one-dimensional binned-likelihood fit is then performed with a mass-based distribution of the surviving events to evaluate the channels' sensitivity. By using a binned-likelihood fit, normalisation uncertainties on the background are constrained.

In the following sections, all reconstructed objects are required to pass the acceptance criteria listed in Table 5.2 and are assumed to be reconstructed and identified with the efficiencies listed in Table 5.3. Events are also required to have one lepton passing the single lepton trigger selection criteria described in Section 5.2.1. Tables and figures of expected signal and background contributions assume the 'nominal' conditions, in which the tau identification efficiency is expected to be 40% and the effect of pile-up on the $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution is estimated by smearing the reconstructed $p_{x,y}^{miss}$ with an additional Gaussian term of width 15 GeV.

5.3.1 WH analysis

In the WH channels, only events in which the W boson decays leptonically are considered. This leads to final states containing one lepton, $p_{\mathrm{T}}^{\mathrm{miss}}$ from the neutrino and two tau leptons from the Higgs boson decay. Events in which both tau leptons decay hadronically ($H \to \tau\tau \to \tau_h\tau_h\nu\nu$) are not considered due to an overwhelming background contribution from $W(\to \ell\nu)$ +jets production in which two jets have been mis-identified as hadronic tau decays (using a $\simeq 1\%$ tau fake-rate). Events in which both tau leptons decay leptonically are also not included since the branching ratio for $\tau\tau \to \ell^+\ell^-\ 4\nu$ is much lower than that of $\tau\tau \to \ell\tau_h\ 3\nu$ and the leptons from tau decay are less likely to pass the acceptance cuts, both of which degrade the expected sensitivity.

The final states considered are then $\ell_W\tau_\ell\tau_h p_{\mathrm{T}}^{\mathrm{miss}}$ where ℓ_W is an e or μ assumed to come from a W-boson decay and τ_ℓ is an e or μ assumed to come from a tau-lepton decay. ℓ_W is defined to be the highest p_{T} lepton in the event since $p_T^{\ell_W} > p_T^{\tau_\ell}$.

Several background sources will contribute to these final states. W/Z + jets, $t\bar{t}$ and tW production can contribute when at least one hadronic jet is mis-identified as a lepton or a τ_h. To model these backgrounds, the jet mis-identification rates listed in Table 5.3 are applied separately to each jet in these events. Production of WZ background and WH signal are modelled using MC acceptances, with corrections for trigger and identification efficiencies listed in Table 5.3.

The WH signal and background process cross sections are listed in Table 5.5. The signal cross sections are calculated using V2HV [52] and include QCD corrections at NLO. The uncertainties on all signal cross sections are $\mathcal{O}(10\%)$, while the uncertainties on the branching ratios determined from HDECAY [53] are $\mathcal{O}(1\%)$. All background process cross sections are calculated using `MCFM` [54] at NLO. The $W/Z+$jets cross sections are calculated after requiring jets to have $p_T > 15$ GeV and $|\eta| < 3.5$, and, when there are two or more jets, $m_{jj} > 20$ GeV.

m_H (GeV)	$\sigma(pp \to WH)$	BR($H \to \tau\tau$)
115	1.98 pb	0.0739
120	1.74 pb	0.0689
125	1.53 pb	0.0620
130	1.35 pb	0.0537
135	1.19 pb	0.0444
Background process	$\sigma \times BR$	
$W(\to l\nu)Z/\gamma^*(\to ee/\mu\mu)$	52.4 pb $\times$ 2.18% = 1.04 pb	
$W(\to l\nu)Z/\gamma^*(\to \tau\tau)$	52.4 pb $\times$ 1.09% = 0.522 pb	
$W(\to l\nu) + 2$ jets	26772 pb $\times$ 32.4% = 8674 pb	
$Z(\to ll) + 1$ jet	24466 pb $\times$ 10.1% = 2470 pb	
$t\bar{t} \to \ell\nu\ell\nu b\bar{b}$	933 pb $\times$ 10.4% = 97.9 pb	

Table 5.5: WH analysis signal and background cross sections at $\sqrt{s} = 14$ TeV.

Since these events will produce neutrinos from multiple sources (i.e. the leptonic W decay and each τ decay), a conventional Higgs mass reconstruction is not possible. The visible mass (see Section 7.3.1) is defined as the invariant mass of the τ_ℓ and τ_h and background discrimination is obtained using the visible mass distribution in a binned-likelihood fit to extract the signal yield.

To further select events, a lower limit on the reconstructed $p_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV suppresses backgrounds from $Z \to \ell\ell/\tau\tau+$ jets production. To suppress the $t\bar{t}$ background, an upper limit is placed at $p_{\mathrm{T}}^{\mathrm{miss}} < 80$ GeV.

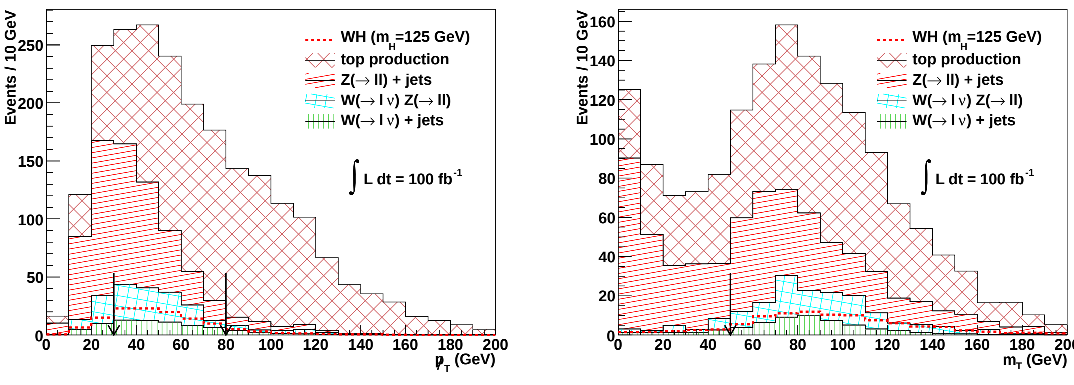
The transverse mass

$$m_T = \sqrt{2p_T^{l_W} p_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\delta\phi(l_W, p_{\mathrm{T}}^{\mathrm{miss}}))}, \tag{5.4}$$

can be used to partially reconstruct the mass of the W boson in $W(\to \ell\nu)$ +jets events and the $W(\to \ell\nu)$ +jets m_T distribution will peak at the end-point, $m_T \simeq m_W$. For $Z \to \tau^+\tau^-$ events, the m_T distribution has a double-peak structure depending on how close the high p_{T} lepton is from the $\vec{p}_T^{miss}$ in the transverse plane. Since it is not expected that the high p_{T} lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$ should be collinear in signal events, we require $m_T > 50$ GeV.

The dominant background after this selection comes from $Z \to \ell^+\ell^-$ and $Z \to \tau\tau \to \ell^+\ell^- 4\nu$ events. Since events with an opposite-sign, same-flavour light-lepton pair form a small fraction of the signal, these events are removed.

Table 5.6 shows the number of signal (N_s^{WH}) and background (N_b) events after each selection requirement, as well as $N_s^{WH}/\sqrt{N_b}$. In Table 5.7 the background yield from each process is shown separately. Table 5.8 shows the expected number of signal events passing the full selection as a function of m_H. Figure 5.1 shows the distributions of the individual cut variables after all prior cuts are applied. The selection gives reasonable statistical sensitivity to WH production, though the large number of events makes the search susceptible to systematic uncertainties. The visible mass distribution after the full $l_W\tau_l\tau_h$ $p_{\mathrm{T}}^{\mathrm{miss}}$ event selection is shown in Figure 5.2.



(a) $p_{\mathrm{T}}^{\mathrm{miss}}$ after all other cuts.

(b) m_T after all other cuts.

Figure 5.1: The $l_W\tau_l\tau_h$ $p_{\mathrm{T}}^{\mathrm{miss}}$ and m_T distributions after all other cuts are applied.

Selection	N_s^{WH}	N_b^{WH}	$N_s^{WH}/\sqrt{N_b^{WH}}$
$p_T^{l_W} > 25$ GeV, $p_T^{\tau_l,\tau_h} > 15$ GeV, $\sum q_l = \pm 1$ and no jet	233	171408	0.6
$30 < E_{\mathrm{T}}^{\mathrm{miss}} < 80$ GeV	137	19124	1.0
$m_T > 50$ GeV	103	1582	2.6
No opposite-sign same-flavour $l_W \tau_l$	92	1177	2.7

Table 5.6: The numbers of WH signal and background events passing each set of requirements, for an integrated luminosity of 100 fb^{-1} and $m_H = 125$ GeV. Also shown is the signal over the square root of background, a measure of the statistical sensitivity to the signal. Additional sensitivity is gained from a fit to the visible mass distribution.

Process	Number of events
$t\bar{t}(\rightarrow l\nu l\nu b\bar{b})$	573
$Z/\gamma^*(\rightarrow ll) + 1$ jet	330
$tW(\rightarrow l\nu bl\nu)$	112
$W(\rightarrow l\nu)Z/\gamma^*(\rightarrow \tau\tau)$	81
$W(\rightarrow l\nu) + 2$ jets	52
$W(\rightarrow l\nu)Z/\gamma^*(\rightarrow ee/\mu\mu)$	30
Total	1177

Table 5.7: The contribution of each background to the $l_W \tau_l \tau_h E_{\mathrm{T}}^{\mathrm{miss}}$ final state for an integrated luminosity of 100 fb^{-1}.

m_H (GeV)	N_s^{WH}	$N_s^{WH}/\sqrt{N_b^{WH}}$
115	122	3.6
120	109	3.2
125	92	2.7
130	70	2.0
135	52	1.5

Table 5.8: The number of WH signal events for m_H in the range 115-135 GeV, and the statistical significance of the excess of signal events over background in 100 fb^{-1} of integrated luminosity.

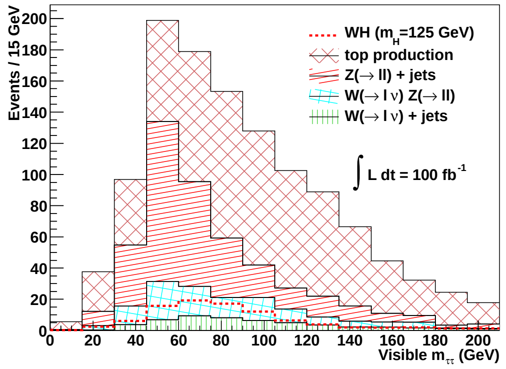


Figure 5.2: The $l_W \tau_l \tau_h$ $p_{\mathrm{T}}^{\mathrm{miss}}$ visible mass distribution of signal and background events passing the full event selection.

5.3.2 ZH analysis

Production of a Higgs boson in association with a Z boson has a lower production cross section than WH but has the advantage that it is less susceptible to backgrounds with fake τ_h production. The final state particle combinations considered are summarised in Table 5.9, where two of the light leptons must be of the same flavour and opposite charge. Note that events where both taus decay hadronically are considered.

Case	Process	Final state
i	$Z(\to l^{\pm}l^{\mp})\tau^{\pm}(\to \tau_h^{\pm}\nu)\tau^{\mp}(\to \tau_h^{\mp}\nu)$	two leptons, two τ_h and $p_{\mathrm{T}}^{\mathrm{miss}}$
ii	$Z(\to l^{\pm}l^{\mp})\tau^{\pm}(\to l^{\pm}2\nu)\tau^{\mp}(\to \tau_h^{\mp}\nu)$	three leptons, one τ_h and $p_{\mathrm{T}}^{\mathrm{miss}}$

Table 5.9: $Z(\to \ell\ell)H(\to \tau\tau)$ final states.

The ZH channels have relatively low signal statistics in 100 fb^{-1} of integrated luminosity, with an irreducible dominant background from ZZ production. Therefore, loose selection requirements are applied and the channels are combined. Additional sensitivity could be achieved by incorporating the $l_Z l_Z \tau_l \tau_l$ channel, but it is not considered here because of the small branching ratio and the increased ZZ background.

While the dominant background is expected to come from diboson production, where two Z bosons decay leptonically ($ZZ \to \ell\ell\tau\tau$), a small yield is expected from Z +jets and $t\bar{t}$ events in which two jets have been mis-identified as hadronically decaying tau leptons in Case i or a light-lepton and a τ_h in Case ii. These backgrounds are modelled by applying the product of jet mis-identification rates listed in Table 5.3 to jet pairs in these events. Production of ZZ background and ZH signal are modelled using the trigger- and identification-corrected MC acceptances listed in Table 5.3.

The ZH signal and background process cross sections are listed in Table 5.10. Like the WH process, the ZH signal cross sections is calculated using V2HV [52] and includes QCD corrections at NLO. Background process cross sections are calculated using `MCFM` [54] at NLO applying the same generator cuts as the WH backgrounds.

Due to the neutrinos produced in each tau decay, a conventional mass reconstruction

m_H (GeV)	$\sigma(pp \to ZH)$	BR($H \to \tau\tau$)
115	1.05 pb	0.0739
120	0.922 pb	0.0689
125	0.813 pb	0.0620
130	0.718 pb	0.0537
135	0.638 pb	0.0444
Background process	$\sigma \times BR$	
$Z/\gamma^*(\to \ell\ell)Z/\gamma^*(\to \tau\tau)$	17.7 pb $\times$ 0.340% = 60.2 fb	
$Z(\to \ell\ell/\tau\tau) + 2$ jets	9018 pb $\times$ 10.1% = 911 pb	
$t\bar{t} \to \ell\nu\ell\nu b\bar{b}$	933 pb $\times$ 10.4% = 97.9 pb	

Table 5.10: ZH analysis signal and background cross sections at $\sqrt{s} = 14$ TeV.

is not possible. However, since the physical sources of missing momentum originate from the Higgs boson decay, the collinear mass approximation (see Chapter 7) can be used to reconstruct the Higgs boson mass.

The highest (lowest) p_T lepton from the decay is required to have $p_T > 25$ (15) GeV. Events are selected by first requiring an opposite-charge same-flavour lepton pair; these are the Z decay products. In events with three light-leptons, the opposite-sign same-flavour lepton pair with an invariant mass closest to the Z boson mass are assumed to originate from the Z boson decay. The remaining object pair ($\tau_l\tau_h$ or $\tau_l\tau_l$) is assumed to originate from the Higgs decay and are also required to be of opposite charge. Table 5.11 shows the numbers of signal (N_s^{ZH}) and background (N_b^{ZH}) events, as well as $N_s^{ZH}/\sqrt{N_b^{ZH}}$, in each channel, after each selection requirement. The collinear mass requirement reduces the signal yield by nearly 20%; recovering these events with an alternative mass variable would improve the measurement.

Using these selection criteria, the expected number of signal and background events for the final states listed in Table 5.9 are shown in Tables 5.12 and 5.13. The signal and background collinear mass distributions of events passing the full event selections are shown in Figure 5.3 for Case i and Case ii events.

Selection	N_s^{ZH}	N_b^{ZH}	$N_s^{ZH}/\sqrt{N_b^{ZH}}$
Opposite-charge $\tau_h\tau_h$ and $l_Z l_Z$; highest (lowest) $p_T^{l_Z} > 25$ (15) GeV; $p_T^{\tau_h} > 25$ GeV	32	193	2.3
Collinear mass solution	26	144	2.1
Opposite-charge $\tau_h\tau_l$ and $l_Z l_Z$; highest (lowest) $p_T^{l_Z} > 25$ (15) GeV; $p_T^{\tau_h(\tau_l)} > 25$ (15) GeV	36	266	2.2
Collinear mass solution	30	188	2.2

Table 5.11: The numbers of ZH signal and background events passing each set of requirements, for an integrated luminosity of 100 fb^{-1} and Higgs boson mass of 125 GeV. Also shown is the signal over the square root of background, a measure of the statistical sensitivity to the signal. Additional sensitivity is gained from a fit to the collinear mass distribution.

m_H (GeV)	N_s^{ZH}	$N_s^{ZH}/\sqrt{N_b^{ZH}}$
115	77	4.2
120	71	3.9
125	56	3.1
130	45	2.4
135	33	1.8

Table 5.12: The number of ZH signal events for m_H in the range 115-135 GeV, and the statistical significance of the excess of signal events over background in 100 fb^{-1} of integrated luminosity.

Process	Number of events
$Z/\gamma^*(\to ll)Z/\gamma^*(\to \tau\tau)$	305
$Z/\gamma^*(\to ll) + 2$ jets	25
$t\bar{t}(\to l\nu l\nu b\bar{b})$	2
Total	332

Table 5.13: The contribution of each background to the ZH final state for an integrated luminosity of 100 fb^{-1}.

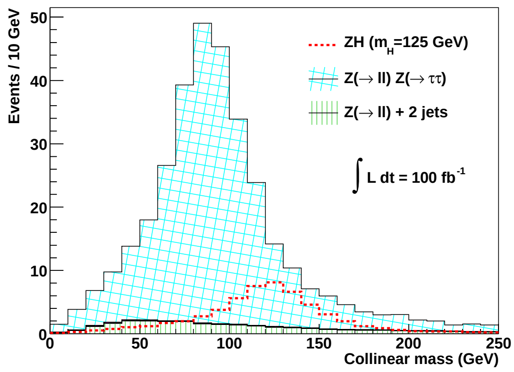


Figure 5.3: The signal and background collinear mass distributions after the full ZH event selection assuming an intergrated luminosity of 100 fb^{-1}.

5.3.3 $t\bar{t}H$ analysis

The last set of associated Higgs production channels considered are those with a $t\bar{t}$ pair. In particular, when the $t\bar{t}$ pair decays semi-leptonically, $t\bar{t} \to \ell\nu q\bar{q}' b\bar{b}$, the final states are the same as that of a WH analysis with the addition of 2 light jets and 2 b-jets. The final states considered are summarised in Table 5.14. Due to the high multiplicity of the final state and the presence of at least one lepton resulting from the decay of a top quark, the $H \to \tau\tau \to \tau_h\tau_h 2\nu$ decay is not overwhelmed by SM backgrounds and is also considered.

Case	Process	Final state
iii	$t\bar{t}(\to q\bar{q}b\bar{b}l^{\pm}\nu)\tau^{\pm}(\to \tau_h^{\pm}\nu)\tau^{\mp}(\to \tau_h^{\mp}\nu)$	one lepton, two τ_h, 4 jets and $p_{\mathrm{T}}^{\mathrm{miss}}$
iv	$t\bar{t}(\to q\bar{q}b\bar{b}l^{\pm}\nu)\tau^{\pm}(\to l^{\pm}2\nu)\tau^{\mp}(\to \tau_h^{\mp}\nu)$	two leptons, one τ_h, 4 jets and $p_{\mathrm{T}}^{\mathrm{miss}}$

Table 5.14: $t\bar{t}(\to q\bar{q}' b\bar{b}l^{\pm}\nu)H(\to \tau^{\pm}\tau^{\mp})$ final states.

The dominant background is expected to come from $t\bar{t}(\to l\nu q\bar{q}' b\bar{b})Z/\gamma^*$ where the Z/γ^* decays to a pair of leptons. A small contribution is also expected from $t\bar{t}$ +jets and $W(\to \ell\nu)$ +6 jets where two jets are mis-identified as a pair of τ_h in Case iii or a light-lepton and a τ_h in Case iv. These backgrounds are modelled by allowing jet pairs to be mis-identified as a pair of τ_h or a lepton and a τ_h, and weighting the event by the product of each jet mis-identification rate shown in Table 5.3. The $t\bar{t}Z/\gamma^*$ background and $t\bar{t}H$ signal are modelled using the trigger- and identification-corrected MC acceptances listed in Table 5.3.

The $t\bar{t}H$ signal and background cross sections are summarised in Table 5.15. The $t\bar{t}H$ signal cross section is calculated with NLO QCD corrections [55]. `Alpgen` is used to calculate the cross section of $W + 6$ jets and $t\bar{t} + 2$ jets.

As with the WH selection, ℓ_W from the semi-leptonic top quark decay is defined to be the highest p_{T} lepton ($p_T^{\ell_W} > p_T^{\tau_\ell}$) and in Case iv events, τ_l is therefore defined as the lower p_{T} lepton. The leptonic W decay and each τ decay will produce neutrinos. Therefore, the visible mass of the $\tau_l\tau_h$ ($\tau_h\tau_h$) pair is used in a binned-likelihood fit to extract the signal yield of each channel.

m_H (GeV)	$\sigma(pp \to t\bar{t}H)$	BR($H \to \tau\tau$)
115	0.785 pb	0.0739
120	0.694 pb	0.0689
125	0.623 pb	0.0620
130	0.559 pb	0.0537
135	0.501 pb	0.0444
$t\bar{t}(\to l\nu q\bar{q}b\bar{b})Z/\gamma^*(\to ee/\mu\mu)$	973 fb $\times$ 2.89% = 28.1 fb	
$t\bar{t}(\to l\nu q\bar{q}b\bar{b})Z/\gamma^*(\to \tau\tau)$	973 fb $\times$ 1.45% = 14.1 fb	
$t\bar{t}(\to l\nu q\bar{q}b\bar{b}) + 2$ jets	255 pb $\times$ 43.8% = 112 pb	
$W(\to \ell\nu) + 6$ jets	23.5 pb $\times$ 32.4% = 7.61 pb	

Table 5.15: $t\bar{t}H$ analysis signal and background cross sections at $\sqrt{s} = 14$ TeV.

To suppress the $t\bar{t}Z$ background, events with an opposite-charge same-flavour lepton pair in Case iv events are removed if they have $|m_Z - m_{\ell\ell}| < 10$ GeV (Figure 5.4). The two τ_h candidates and the $\tau_h - \tau_\ell$ in Cases iii and iv, respectively, are required to be oppositely charged to reduce the background from mis-identified jets. To suppress the contribution from $t\bar{t}$ +jets, the visible mass is required to be less than 150 GeV.

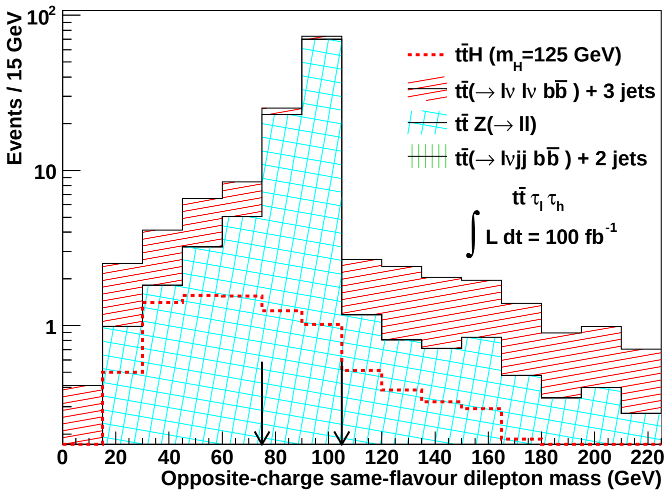


Figure 5.4: The opposite-sign, same-flavour light-lepton invariant mass distribution of $t\bar{t}H$ $\tau_l\tau_h$ events pasing all other selection criteria.

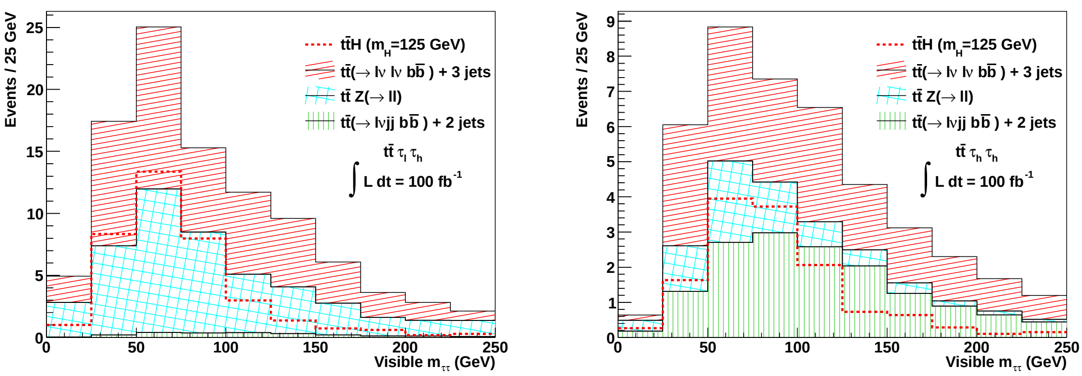
Using these selection criteria, the expected number of signal and background events for the final states listed in Table 5.14 are shown in Tables 5.16 and 5.17. The signal and background visible mass distributions of events passing the full event selection in Case iii and Case iv events are shown in and Figures 5.5(a) and 5.5(b), respectively.

m_H (GeV)	Channel	N_s^{ttH}	$N_s^{ttH}/\sqrt{N_b^{ttH}}$
115	$t\bar{t}+\tau_h\tau_l$	47	4.8
	$t\bar{t}+\tau_h\tau_h$	17	2.7
120	$t\bar{t}+\tau_h\tau_l$	47	4.8
	$t\bar{t}+\tau_h\tau_h$	16	2.5
125	$t\bar{t}+\tau_h\tau_l$	37	3.7
	$t\bar{t}+\tau_h\tau_h$	14	2.1
130	$t\bar{t}+\tau_h\tau_l$	30	3.0
	$t\bar{t}+\tau_h\tau_h$	11	1.7
135	$t\bar{t}+\tau_h\tau_l$	22	2.2
	$t\bar{t}+\tau_h\tau_h$	7	1.1

Table 5.16: The number of ttH signal events in each channel for m_H in the range 115-135 GeV, and the statistical significance of the excess of signal events over background in 100 fb^{-1} of integrated luminosity. The $t\bar{t}$ pair is selected in the $l_W\nu q\bar{q}b\bar{b}$ final state.

Process	$t\bar{t}+\tau_h\tau_l$ channel	$t\bar{t}+\tau_h\tau_h$ channel
$t\bar{t}(\rightarrow l\nu l\nu b\bar{b})+3$ jets	52	20
$t\bar{t}(\rightarrow l\nu q\bar{q}b\bar{b})+Z/\gamma^*(\rightarrow ee/\mu\mu)$	32	2
$t\bar{t}(\rightarrow l\nu q\bar{q}b\bar{b})+Z/\gamma^*(\rightarrow \tau\tau)$	13	5
$t\bar{t}(\rightarrow l\nu q\bar{q}b\bar{b})+2$ jets	2	15
Total	99	42

Table 5.17: The contribution of each background to the $t\bar{t}H$ final states for an integrated luminosity of 100 fb^{-1}.



(a) $t\bar{t}H$ $\tau_h\tau_h$ visible mass distribution.

(b) $t\bar{t}H$ $\tau_\ell\tau_h$ visible mass distribution.

Figure 5.5: The visible mass distributions after the full $t\bar{t}H$ event selection for Case iii (a) and Case iv (b) events.

5.4 Sensitivity estimation

After applying the full event selection, the sensitivity of each channel is evaluated by extracting the number of expected signal events from a binned-likelihood fit. The expected variation of measurements is determined using an ensemble of pseudo-experiments of the mass-based distribution of each channel.

5.4.1 Signal yield extraction

Expected sensitivities oo the cross section of a given process is determined using pseudoexperiments . In each pseudoexperiment, data are produced according to a Poisson distribution in each bin of the relevant mass-based fit distribution, where the mean of the Poisson is equal to the expected number of signal and background events in that bin. The number of signal events is determined by minimizing the negative log likelihood of the fit distribution. This procedure is performed for 10^4 pseudoexperiments for each process, and the statistical uncertainty is taken to be the root-mean square of the resulting signal-yield distribution. The relative statistical uncertainty on the $\sigma \times \mathrm{BR}$ of each signal process are shown in Figure 5.6. Sub-channels are combined according to $1/\sigma = 1/\sqrt{\sum_{channel} \sigma^2_{channel}}$. Since the ratio of the WWH and ZZH couplings are fixed by the SU(2)$_\mathrm{L}$ gauge symmetry of the SM, the WH and ZH channels are combined into the VH channel.

5.5 Expected sensitivity to SM Higgs coupling ratios

Measuring these channels allows us to directly probe the Γ_V/Γ_t ratio. The expected uncertainty on this ratio for a range of SM Higgs boson masses and detector conditions is shown in Figure 5.7, assuming the expected sensitivity of each of the WH, ZH and $t\bar{t}H$ sub-channels shown in Figure 5.6.

Recent predictions of the expected LHC sensitivity to the processes $\sigma(pp \to VH(\to b\bar{b}))$ [20] and $\sigma(pp \to t\bar{t}H(\to b\bar{b}))$ [55, 56, 57] determine a signal sensitivity $N_S/\sqrt{N_B + N_S}$ for

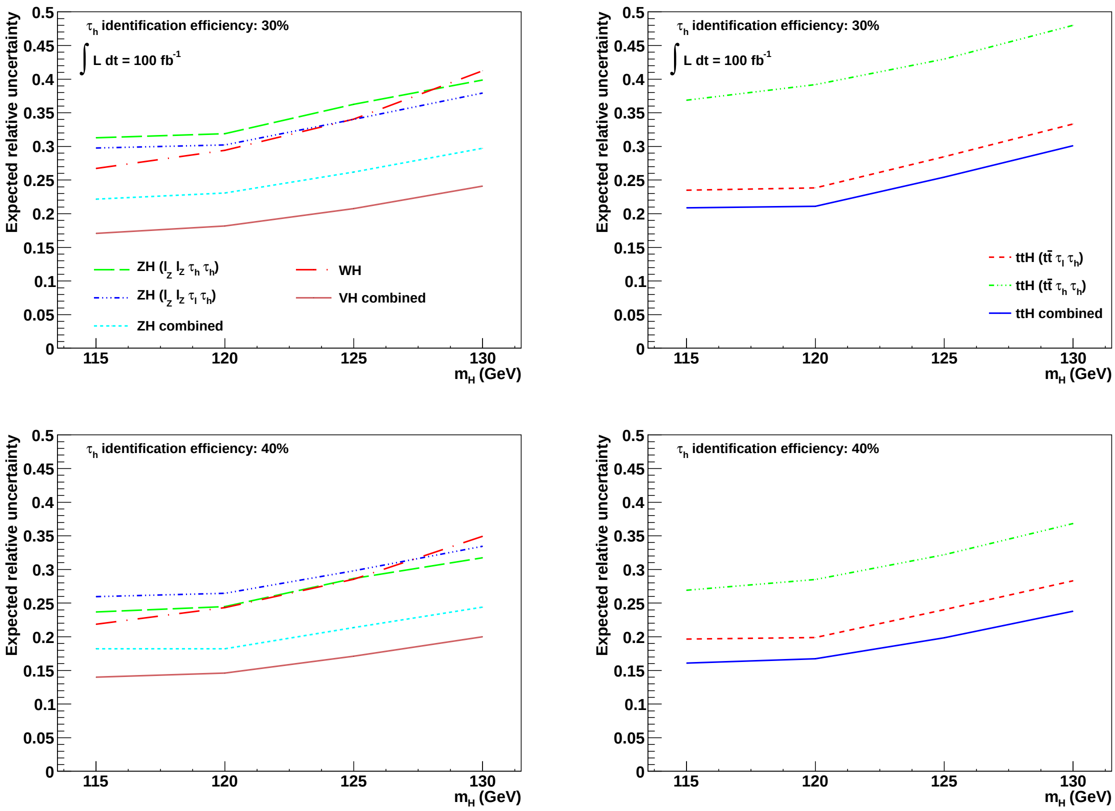


Figure 5.6: The expected relative statistical uncertainties on $\sigma \times$ BR of VH (left, $V = W, Z$) and $t\bar{t}H$ (right) production for the nominal (top) and optimistic (bottom) tau identification performance scenarios.

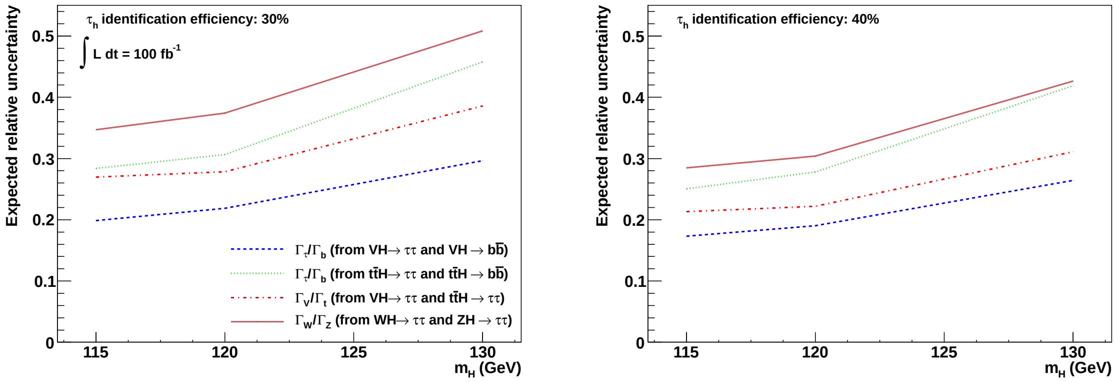


Figure 5.7: The expected relative statistical uncertainties on ratios of partial widths using measurements of associated Higgs boson production and decays to tau-lepton or bottom-quark pairs. Each partial width Γ_i corresponds to the trilinear interaction of a Higgs boson to another particle i. Shown are the nominal (left) and optimistic (right) detector performance scenarios.

a range of SM Higgs boson masses. By taking the inverse of this signal significance as the expected relative uncertainty of a measurement of the $\sigma \times BR$ of these channels (assuming no systematic uncertainties) an estimate of the expected uncertainty on a measurement of Γ_τ/Γ_b using both production mechanisms is made. This is shown as a function of SM Higgs boson mass, assuming the sensitivity of each sub-channel shown in Figure 5.7.

5.6 Conclusion

The associated SM Higgs production channels WH, ZH and $t\bar{t}H$ can be used to probe the Yukawa coupling strength $g_{\tau\tau H}$ and the coupling ratios Γ_τ/Γ_b and Γ_V/Γ_t. With an integrated luminosity of 100 fb^{-1}, experiments at the LHC are expected to be sensitive to these coupling ratios, with an expected relative uncertainty of 20% - 50%, depending on the SM Higgs boson mass and the detector performance.

Chapter 6

Tau identification and mis-identification probability

6.1 Introduction

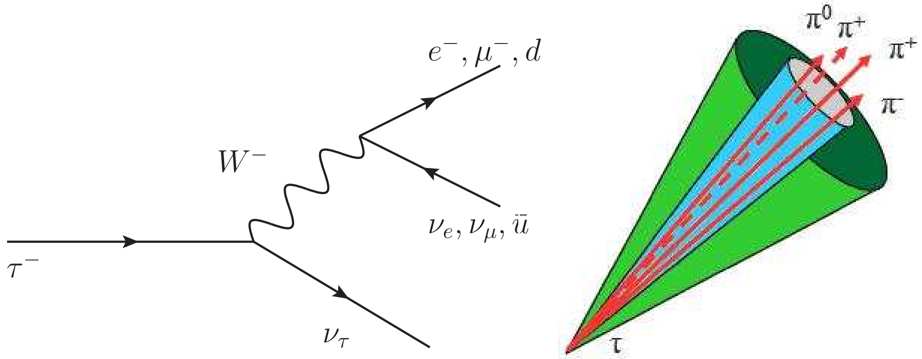
This chapter describes the physics of hadronic tau decay (τ_h) and τ_h reconstruction and identification in ATLAS. Finally, a measurement of the τ_h mis-identification probability for hadronic jets using data collected in 2010 is also presented.

6.2 Physics with tau leptons

The tau lepton was first discovered in experiments at the SPEAR e^+e^- collider experiment at SLAC in 1974 as an excess of events with an opposite sign electron-muon pair with large missing energy [6]. The tau rest mass has since been measured to a high degree of accuracy to be 1776 ± 0.17 MeV[1] with a mean lifetime, $\tau = (290.6 \pm 1.0) \times 10^{-15}$ s leading to an average decay length $c\tau = 87.11$ μm [6]. Tau leptons decay weakly and the experimental signature can be classified as either 'leptonic', where either an electron or muon is produced through the decay of a virtual W boson, or 'hadronic' where one or more charged and neutral

[1]In natural units where $c = 1$.

mesons can be detected, as shown in Figure 6.1(a). Both of these classes of tau decay cause an observed momentum imbalance due to the undetected neutrino(s) produced in the tau decay.



(a) Possible Tau decay diagrams.

(b) A hadronically decaying tau lepton illustration.

Figure 6.1: Tau lepton decays.

No distinction is made in ATLAS between a measured 'primary' lepton such as those produced in $Z \to e^+e^-/\mu^+\mu^-$ decay or a measured lepton produced in tau decay. The branching fractions of the leptonic and hadronic final states of tau decay have also been measured previously; the corresponding final states identified in ATLAS are shown in Table 6.1. The hadronic final states are classified into cases where either one or three reconstructed charged particle tracks are produced, since the tau lepton is electromagnetically charged with $|Q| = 1$. This covers about 98% of all possible hadronic tau decays.

τ^- Final State	BR (%)	Decay type
$e^-\bar{\nu_e}\nu_\tau$	17.85	leptonic
$\mu^-\bar{\nu_\mu}\nu_\tau$	17.36	
$\pi^-\nu_\tau$	10.91	hadronic, 1-track
$\pi^-\pi^0\nu_\tau$	25.51	
$\pi^-2\pi^0\nu_\tau$	9.29	
$\pi^-3\pi^0\nu_\tau$	1.04	
$K^-\nu_\tau$ + Neutrals	1.57	
$\pi^+2\pi^-\nu_\tau$	9.32	hadronic, 3-track
$\pi^+2\pi^-\nu_\tau\pi^0$	4.61	
$\pi^+K^-\pi^-\nu_\tau$ + Neutrals	0.48	

Table 6.1: Tau decay modes reconstructed at ATLAS and their associated branching fractions (taken from reference [6]).

6.3 Hadronic tau reconstruction and identification in ATLAS

Due to the Lorentz boost produced by the large mass difference between the tau lepton and its parent (typically a W or Z boson), hadronically decaying tau leptons are observed as narrow, pencil-like hadronic jets with one or three tracks, as shown in Figure 6.1(b).

6.3.1 Reconstruction

A description of τ_h reconstruction is given in Section 3.7.6.

Indentification variables

Every reconstructed hadronic jet will seed a corresponding τ_h candidate, offering no background rejection. The following variables are used to provide separation between true τ_h and hadronic jets where all energies are calibrated at the EM scale:

- Electromagnetic radius (R_{EM}): the transverse energy weighted shower width in the electromagnetic (EM) calorimeter:

$$R_{EM} = \frac{\sum_{i \in EM\ 0-2}^{\Delta R_i < 0.4} E_{T,i} \Delta R_i}{\sum_{i \in EM\ 0-2}^{\Delta R_i < 0.4} E_{T,i}}, \tag{6.1}$$

  where i is the index over EM calorimeter cells associated with the τ_h candidate (and hence its seed jet) in the pre-sampler and layers 1 and 2. The angular separation between the calorimeter cell and the τ_h axis is ΔR_i, and $E_{T,i}^{EM}$ is the cell transverse energy.

- Hadronic shower radius (R_{had}): the transverse energy weighted shower width in the hadronic calorimeter:

$$R_{had} = \frac{\sum_{i \in EM3+had}^{\Delta R_i < 0.4} E_{T,i} \Delta R_i}{\sum_{i \in EM3+had}^{\Delta R_i < 0.4} E_{T,i}}, \tag{6.2}$$

where the index i is over the hadronic calorimeter cells and the cells in layer 3 of the EM calorimeter that are associated with the seed jet.

- Track radius (R_{track}): the p_{T} weighted track width

$$R_{\text{track}} = \frac{\sum_{i}^{\Delta R^i < 0.4} p_{\text{T}}^i \Delta R^i}{\sum_{i}^{\Delta R^i < 0.4} p_{\text{T}}^i}, \tag{6.3}$$

where the index i runs over all inner detector tracks passing the selection criteria in Section 3.7.6 within a cone of $\Delta R = 0.4$ from the τ_h axis and p_{T}^i is the p_{T} of track i.

- Leading track momentum fraction (f_{track}):

$$f_{\text{track}} = \frac{p_{\text{T}}^{\text{track}}}{p_{\text{T}}^{\tau}}, \tag{6.4}$$

where $p_{\text{T}}^{\text{track}}$ and p_{T}^{τ} are, respectively, the highest p_{T} track of the τ_h candidate and the transverse momentum of the τ_h candidate.

- Core energy fraction f_{core}: the fraction of transverse energy in a cone of radius $\Delta R = 0.1$ from the τ_h candidate:

$$f_{\text{core}} = \frac{\sum_{i \in all}^{\Delta R_i < 0.1} E_{T,i}}{\sum_{j \in all}^{\Delta R_i < 0.4} E_{T,j}}, \tag{6.5}$$

where i and j are the indices over all calorimeter cells associated with the τ_h candidate (and hence its seed jet) within a cone of radius $\Delta R = 0.1$ and $\Delta R = 0.4$ from the jet-axis, respectively.

- Calorimeter radius (R_{cal}): the shower width in the EM and hadronic calorimeter weighted by transverse energy deposited in each calorimeter

$$R_{\text{cal}} = \frac{\sum_{i \in all}^{\Delta R_i < 0.4} E_{T,i} \Delta R_i}{\sum_{i \in all}^{\Delta R_i < 0.4} E_{T,i}}, \tag{6.6}$$

where i and j run over calorimeter cells associated with the τ_h candidate within a cone of radius $\Delta R = 0.4$ of the τ_h axis in the EM and hadronic calorimeter, respectively.

- Track mass (m_{tracks}): the invariant mass of the associated tracks.

- Transverse flight path significance ($S_{\mathrm{T}}^{\mathrm{flight}}$): the decay length significance in the transverse plane of the secondary vertex for τ_h candidates with more than one associated track

$$S_{\mathrm{T}}^{\mathrm{flight}} = \frac{L_{\mathrm{T}}^{\mathrm{flight}}}{\Delta L_{\mathrm{T}}^{\mathrm{flight}}}, \tag{6.7}$$

where $L_{\mathrm{T}}^{\mathrm{flight}}$ is the reconstructed, signed decay length and $\Delta L_{\mathrm{T}}^{\mathrm{flight}}$ is the estimated uncertainty.

- Maximum ΔR (ΔR_{max}): The maximal ΔR between any associated track within a cone of radius $\Delta R = 0.2$ and the τ_h axis.

- Leading track impact parameter significance ($S_{\mathrm{lead\ track}}$):

$$S_{\mathrm{lead\ track}} = \frac{d_0}{\Delta d_0}, \tag{6.8}$$

where d_0 is the distance of closest approach of the highest p_{T} track to the reconstructed primary vertex in the transverse plane, and its estimated uncertainty Δd_0.

These variables are then used to form three different discriminants: a cut-based selection, a projective likelihood identification and a boosted decision tree (BDT) identification [58, 59, 60]. The cut values, BDT and likelihood requirements are optimised using MC signal and QCD di-jet background from data [61] at three levels of signal efficiency known as ‘loose’, ‘medium’ and ‘tight’. The values of the BDT and likelihood scores required for these points are also parametrised as a function of the τ_h p_{T}, since several of the reconstruction variables’ shapes vary as a function of τ_h p_{T}.

6.4 Tau mis-identification probability

6.4.1 Introduction

In the search for processes such as $Z/H \to \tau\tau$, several SM processes such as $W \to \ell\nu$ in association with one or more jets will contribute as backgrounds if a jet is mis-identified as a hadronically decaying tau lepton (diagrams for $W(\to \ell\nu)$ +jets production are shown in Figure 6.2).

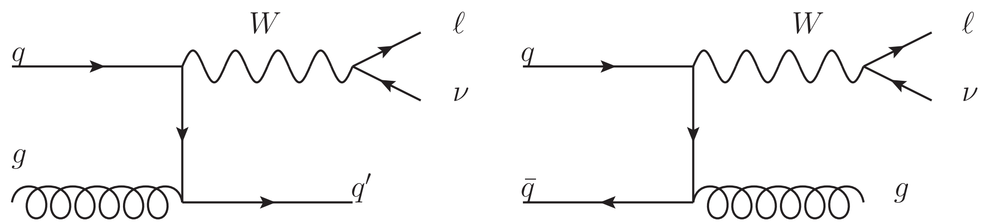


Figure 6.2: $W(\to \ell\nu)$ +jets diagrams. Such processes are an important background to $Z/H \to \tau\tau$ analyses when the jet formed by the outgoing parton hadronisation is mis-identified as a τ_h.

Parton hadronisation and fragmentation

Perturbative QCD is a theory for describing the interaction of quarks and gluons at very short distances. At long distances, the QCD coupling strength becomes much stronger and can no longer be successfully described by perturbation theory. In this region of confinement, the process of hadronisation transforms the coloured partons into colourless hadrons that are reconstructed as jets in the calorimeters. Hadronisation has yet to be understood from first principles and hence phenomenological models of a probabalistic and iterative nature are used, one example is the string fragmentation model used in the `Pythia` [47] MC generator.

However, these models merely aim to respresent existing data and cannot claim to be correct, particularly when extrapolating jet properties to higher energy experiments. It was observed that the hadronisation models used in many generators poorly reproduce the jet shapes in LHC data and this has a dramatic effect on the tau identification algorithms. For example, the `Pythia` MC prediction of the yield of W+jet events in which a hadronic jet has been mis-identified as a τ_h is almost double that observed in data (see Chapter 7). A precise data-driven determination of the number of 'fake' τ_h candidates is therefore crucial to these channels.

6.4.2 Measuring the τ_h fake-rate from data

The rate at which a hadronic jet is mis-identified as a hadronically decaying tau lepton is referred to as the tau mis-identification probability or tau fake-rate, f_{ID}. In this analysis, f_{ID} is defined to be:

$$f_{ID} = \frac{\text{number of jets reconstructed and identified as a } \tau_\mathrm{h}}{\text{number of jets reconstructed as a } \tau_\mathrm{h}}. \tag{6.9}$$

A hadronic jet originating from a quark is more likely to be mis-identified as a τ_h than a hadronic jet originating from a gluon since quark-initiated jets will, on average, hadronise into a narrower $\eta - \phi$ cone and have a lower track multiplicity than jets that originate from gluons.

To measure f_{ID}, events in which a photon is produced in association with a hadronic jet allow us to use the so-called tag and probe method with a large, clean sample of seed jets to be extracted from data. This is achieved by selecting a well identified photon as a 'tag' object that will have an associated, kinematically connected 'probe' hadronic jet which has not been directly subjected to any selection. The main advantage of this method is that it is largely independent of Monte Carlo or any bias introduced by using one of the tau triggers. Using the photon-jet channel also gives an estimate of the tau fake-rate appropriate for $W+$jet production. This is because neither the photon nor the W boson will interact via the strong force and hence the vast majority of the associated jets produced will therefore have been initiated by a quark, rather than a gluon. The leading order γ-jet production diagrams at the LHC are shown in Figure 6.3.

An arbitrarily low f_{ID} can be achieved by sacrificing the tau identification signal effi-

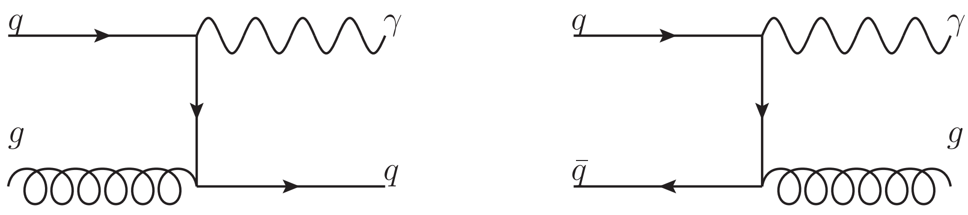


Figure 6.3: γ-jet diagrams.

ciency, which is defined as the efficiency of identifying true τ_h in Monte Carlo $W \to \tau\nu$ and $Z \to \tau\tau$ samples. Rather than attempting to minimise f_{ID}, the selection is defined for the following levels of signal efficiency:

- 70% signal efficiency ('loose' criteria);
- 50% signal efficiency ('medium' criteria);
- 40% signal efficiency ('tight' criteria).

Furthermore, 'looser' selection corresponds to 1-prong τ_h candidates passing the loose criteria and 3-prong τ_h candidates passing the medium criteria while 'tighter' selection corresponds to 1-prong τ_h candidates passing the medium criteria and 3-prong τ_h candidates passing the tight criteria.

6.4.3 Event selection

This analysis uses data collected during 2010, corresponding to an integrated luminosity of $\int \mathcal{L}\ \mathrm{dt} \simeq 34\ \mathrm{pb}^{-1}$ when the relevant sub-detectors were performing optimally. In order to select the γ+jet topology, events are required to pass Event-Filter level trigger in which a photon with transverse energy $E_\mathrm{T} > 15$ GeV has been identified.

Reconstructed photons in each event are required to the satisfy the following selection criteria:

- Cluster $E_T \geq 15$ GeV,
- $|\eta| \leq 1.37$ or $1.52 < |\eta| < 2.37$,

Cut	Data 2010
Exactly one photon candidate satisfying selection	481,790
Exactly one jet satisfying selection with tau candidate satisfying selection criteria	217,101
$\lvert\Delta\phi(\gamma, \mathrm{jet})\rvert \geq \pi - 0.3$	96,882
p_{T}-balance	90,393

Table 6.2: γ-jet cut-flow for events that fulfil the trigger and data-quality requirements for $\int \mathcal{L}\mathrm{dt} \simeq 34\ \mathrm{pb}^{-1}$.

- Well identified and isolated from other energy deposits in the calorimeter[2].

Jets are selected if they have pseudorapidity $|\eta| \leq 2.5$, transverse momentum $p_{\mathrm{T}} \geq 15$ GeV and satisfy the data quality criteria described in reference [63].

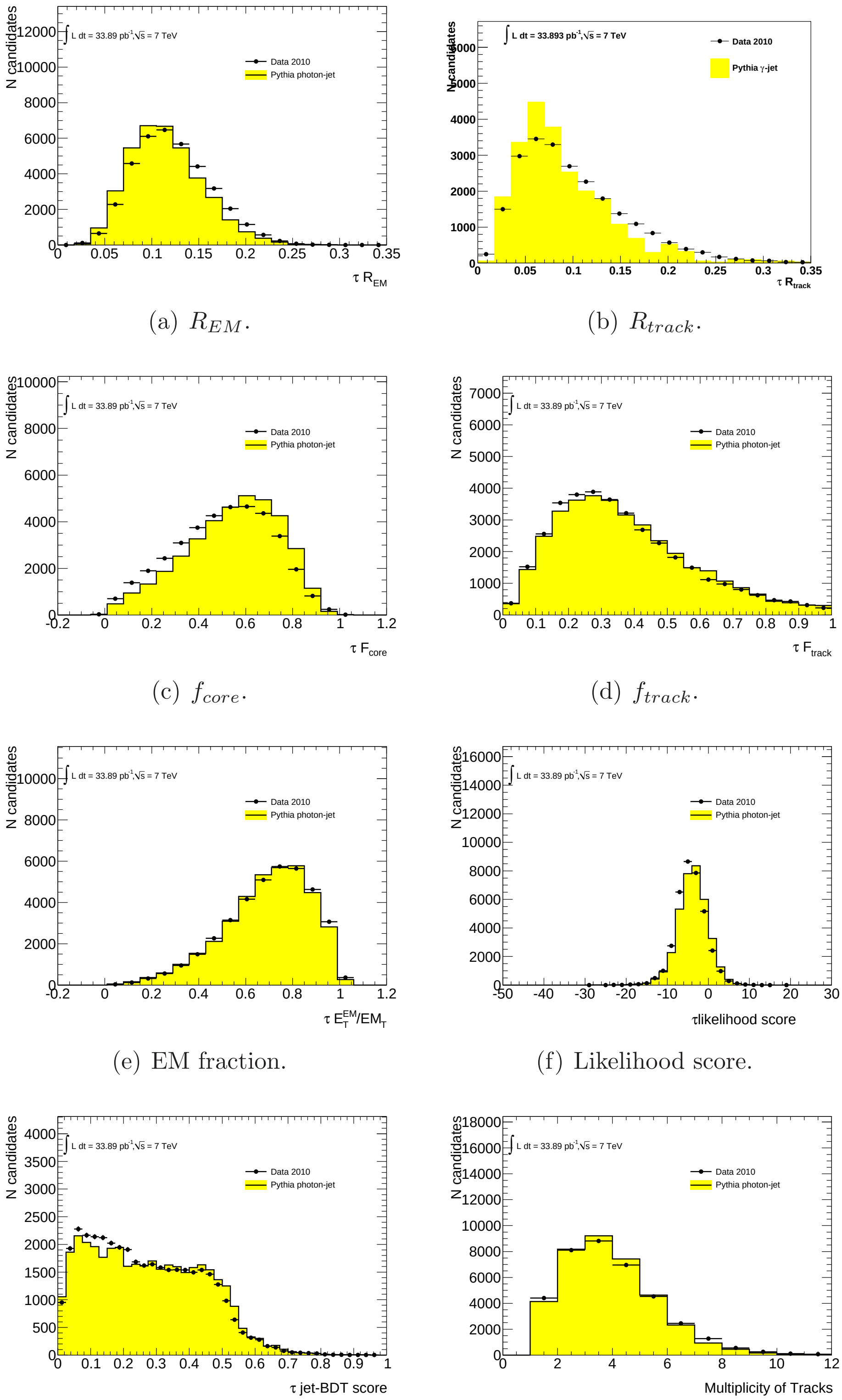
Each event is required to have exactly one photon and one jet passing these selection criteria, separated in the transverse plane such that the difference between their azimuthal angles $\Delta\phi > \pi - 0.3$ radians and balanced in p_{T} such that the difference between their transverse momenta is less then half of the transverse momentum of the photon. The cut flow is summarised in Table 6.2.

The identification variables for τ_h candidates with at least one associated inner detector track within $\Delta R \leq 0.2$ of a jet in events passing this selection are shown in Figure 6.4. The data / MC differences are attributed to poor hadronisation modelling in the MC, in which predictions of the hadron shower tend to be narrower than those observed in data.

Pile-up

The effects of pileup will become increasingly important as the instantaneous luminosity delivered by the LHC machine increases. To study these effects, the fake-rate was calculated for events with different amounts of pileup, by separating the events by the number of reconstructed vertices in the event. The results are shown in Figure 6.5. The fake-rate decreases with more p-p-interactions in the same bunch crossing. This result is expected, as

[2]This corresponds to the `isEM-TightIso` criteria defined in reference [62].



(a) R_{EM}.

(b) R_{track}.

(c) f_{core}.

(d) f_{track}.

(e) EM fraction.

(f) Likelihood score.

(g) BDT score.

(h) Number of associated tracks.

Figure 6.4: Hadronic tau identification variables for reconstructed hadronically-decaying tau-lepton candidates matched to the probe jets in data 2010 and γ-jet MC after applying the γ-jet selection criteria.

a more crowded environment leads to a lower fake-rate due to fewer tau candidates meeting the tau identification calorimeter criteria. Due to this dependence, it was decided to calculate the fake-rate for events with different number of vertices. The bins chosen are 1-2 vertices, and ≥3 vertices.

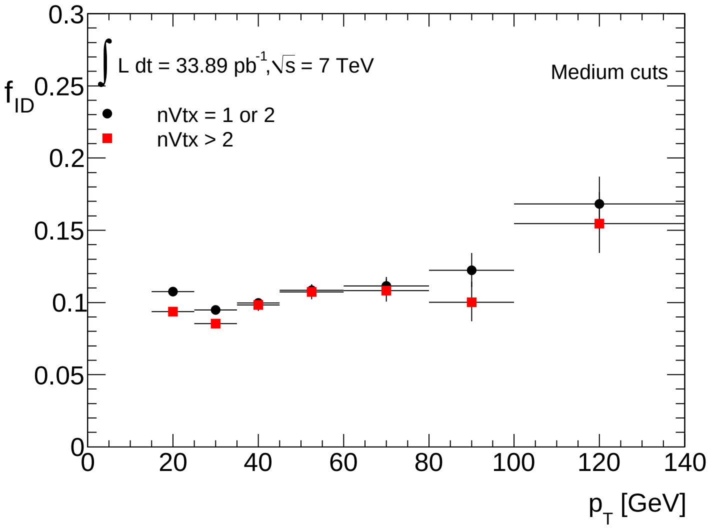


Figure 6.5: Comparison of f_{ID} in events with different vertex multiplicity for the cut-based τ_h identification with a medium selection, as a function of candidate τ_h p_{T}.

Probe jet origin

After the photon-jet event selection has been applied, the probe jet sample will consist of a mixture of quark- and gluon-initiated jets. The quark fraction is defined to be the probability that a probe jet is initiated by a quark. The quark fraction is estimated using the Monte Carlo information, looking for highest p_{T} gluon or quark within a cone of $\Delta R = 0.4$ from the reconstructed probe jet to determine if the jet originated from a quark or a gluon. The quark fraction is shown in Figure 6.6 for γ+jet, Z+jet and QCD di-jet MC samples [51]. This procedure reproduces the cross section ratio from each matrix element to within 3%.

Furthermore, good agreement is observed when calculating f_{ID} exclusively for quark- and gluon-initiated jets across the different event topologies using Monte Carlo samples [51]. This is shown in Figures 6.7(a) and 6.7(b) for quark- and gluon-initiated jets, respectively. The fake-rate for quark-initiated jets is much higher than that of gluon-initiated jets.

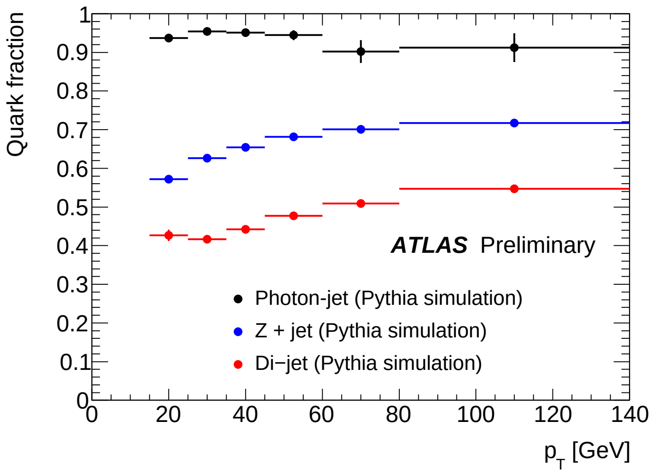
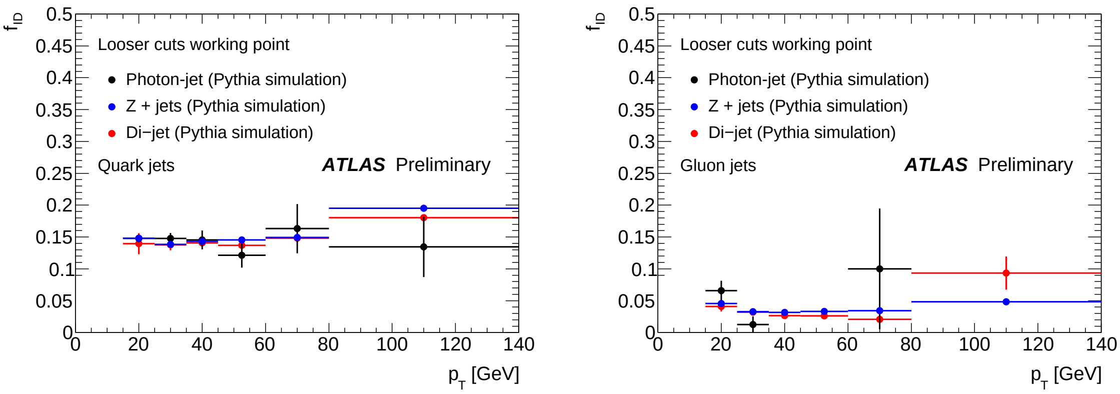


Figure 6.6: Fraction of probe jets initiated by quarks as a function of the jet transverse momentum in Pythia γ+jet, Z+jet and QCD di-jet MC samples.



(a) f_{ID} from quark-initiated τ_h candidates.

(b) f_{ID} from gluon-initated τ_h candidates.

Figure 6.7: Mis-identification probability, as a function of probe candidate p_{T}, for quark-initiated and gluon-initiated probe τ_h candidates for the loose cut-based tau identification criteria, as determined from Monte Carlo simulation for γ+jet, Z+jet and QCD di-jet event topologies. The errors shown are statistical only.

6.4.4 Systematic uncertainties

True tau leptons in the sample

The presence of events containing hadronic tau decays passing the photon-jet selection is estimated using W$\rightarrow \tau\nu$ and Z$\rightarrow \tau\tau$ MC samples, requiring events to pass the selection criteria described in Section 6.4.3. Normalising the expected yield of these samples to the integrated luminosity collected leads to a negligible expected contribution due to the requirement that there be a well identified, isolated photon candidate.

Di-jet contamination due to the photon selection

As described in section 6.4.3, the event selection requires a well identified and isolated photon candidate. By removing the photon candidate isolation requirement, the fraction of jets initiated by quarks is expected to decrease due to the contribution from QCD di-jet events where one jet has been falsely identified as a photon. The effect on the tau fake-rate in these events and the events selected using the default photon selection is shown in Figure 6.8. As expected, the fake-rate decreases since the fraction of probe jets initiated by a quark will have decreased.

A systematic uncertainty for this effect is assigned to each p_{T} bin by taking the f_{ID} ratio with and without the photon calorimeter isolation requirement.

Additional gluon-initiated probe jets

As discussed in Section 6.4.3, the event selection requires that the jet and the photon are back-to-back in the transverse plane ($\Delta\phi > \pi - 0.3$) and balanced in p_{T} ($|\Delta p_{\mathrm{T}}| < \frac{p_{\mathrm{T}}^{\gamma}}{2}$). The deviation of $\Delta\phi$ and Δp_{T} from their nominal values of π and 0, respectively, can have an effect on the fake-rate due to the increased presence of gluon probe jets i.e. events from higher order diagrams with an additional final state gluon that is selected as the probe jet. To study this effect, two sub-samples are created; one containing events which fulfil a more

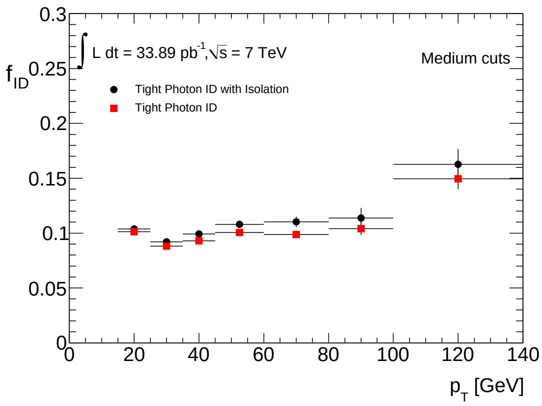


Figure 6.8: ID fake-rates $f_{\rm ID}$ for probe jets passing the medium cut based ID in selected events where the tag photons pass the default tight selection and the tight selection with isolation in the photon-jet method.

strict requirement on the variable in question, and one containing events which fulfil a looser requirement (up to the allowed tolerance). This is performed separately for the two variables, and the thresholds are chosen to produce sub-samples with similar numbers of events. The sub-samples are defined as follows:

- back-to-back cut:

  looser back-to-back: $\Delta\phi < \pi - 0.1$,

  tighter back-to-back: $\Delta\phi \geq \pi - 0.1$.

- $p_{\rm T}$ balance cut:

  loosely balanced: $|\Delta p_{\rm T}| < 0.44 * \frac{p_{\rm T}^{\gamma}}{2}$,

  tightly balanced: $|\Delta p_{\rm T}| \geq 0.44 * \frac{p_{\rm T}^{\gamma}}{2}$.

The effect on the medium cut-based identification criteria f_{ID} can be seen in Figure 6.9. In both cases the fake-rate for the sample which fulfils the tighter requirement is slightly higher than for the sample which fulfil the looser requirement. This suggests that events in which the γ-jet pair is not well balanced in $p_{\rm T}$ or are less back-to-back in the ϕ plane have a lower quark fraction. A likely explanation is that this is due to events from higher order diagrams in which the quark has radiated a gluon and one jet is outside the fiducial region of the detector (or fails the jet selection).

For each of these effects, a systematic uncertainty on f_{ID} is assigned separately in each p_{T} bin as the ratio of f_{ID} in which the photon and tau candidate p_{T} are loosely and tightly balanced.

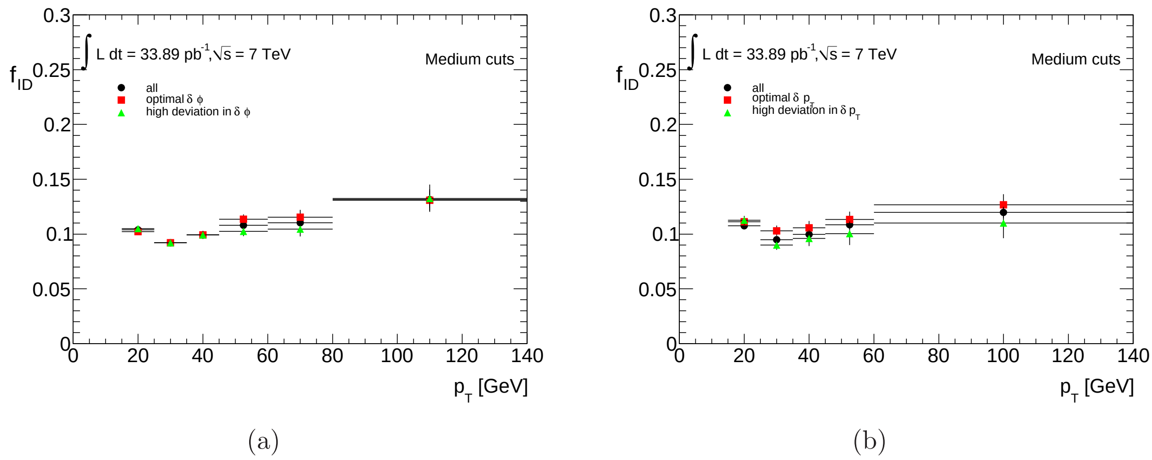


Figure 6.9: The effect on f_{ID} when the event selection criteria is varied for the cut-based τ_h identification criteria with medium selection, as a function of candidate τ_h p_{T}.

6.4.5 Results

The tau fake-rate is calculated using the probe jet for each level of the tau identification algorithms listed in Section 6.4.1. Figure 6.10 shows the fake-rate for each tau identification algorithm where the hatched band represents the sum in quadrature of the systematic uncertainty due to di-jet and multijet contamination.

6.4.6 Summary

A measurement of the rate at which the ATLAS τ_h identification algorithms mis-identify a hadronic jet as a τ_h has been made using data collected in 2010. The mis-identification probabilities range between 0.5% to 10% depending on the τ_h identification algorithm, the τ_h candidate p_{T}, number of reconstructed tracks associated with the τ_h candidate, and the number of pileup interactions in the event.

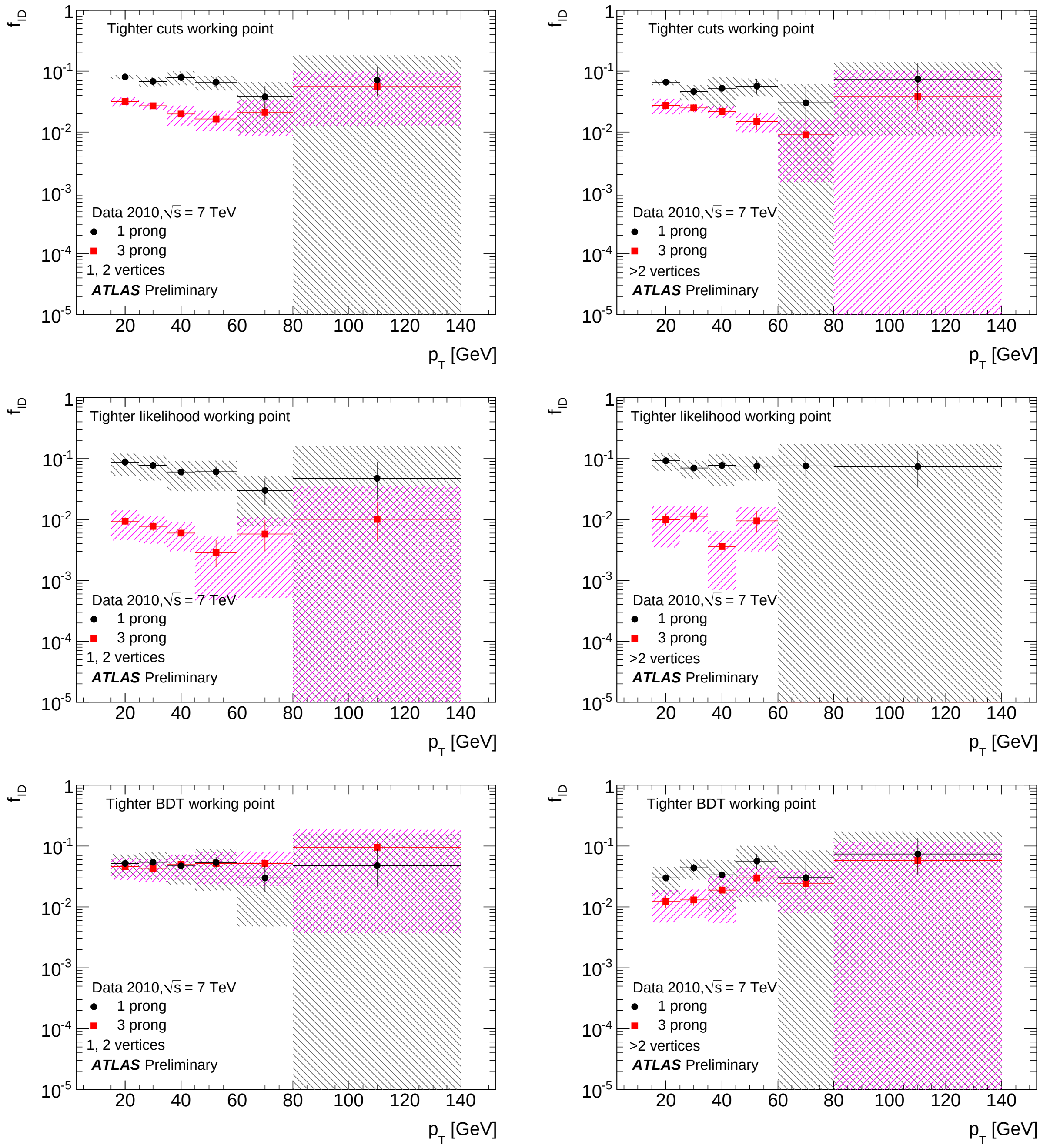


Figure 6.10: The tau mis-identification probability of hadronic jets from γ+jet topologies. These are shown as a function of τ-lepton p_{T} for 1-prong and 3-prong candidates in events with one or two primary vertices (left column) and more than two primary vertices (right column) and for the cut-based (top row), likelihood-based (middle row) and BDT-based (bottom row) identification algorithms. The level of the identification algorithms corresponds to the tighter working point. The statistical uncertainties are represented by the vertical bars; the shaded areas correspond to the total uncertainty.

The f_{ID} measured in γ-jet data was successfully used to cross check the data-driven estimate of the $W(\rightarrow \ell\nu)$ +jets background yield in the $Z \rightarrow \tau\tau \rightarrow \ell\tau_h$ cross section measurement [64].

Chapter 7

Search for the Standard Model Higgs boson in the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ channels

7.1 Introduction

This analysis is a direct search for SM Higgs boson production and subsequent decay to final states with at least one tau lepton at the ATLAS detector using data collected in 2011. In particular we search exclusively for the di-tau final state $H \to \tau^+\tau^- \to \ell\tau_h\ 3\nu$. The final state with a light lepton (e, μ) and a hadronic tau decay, is expected to give the greatest sensitivity of the di-tau final states since an isolated light lepton[1] can be easily triggered on and the branching ratio is much higher than that to two light leptons.

Another search is also made for events from the same production mechanisms and final state but through a different Higgs decay chain, in particular to a tau lepton and a light lepton via a pair of W bosons (Figure 7.1).

[1]The $e^- - \tau_h^+$, $\mu^- - \tau_h^+$ and their charge conjugate final states are considered throughout the rest of this Chapter.

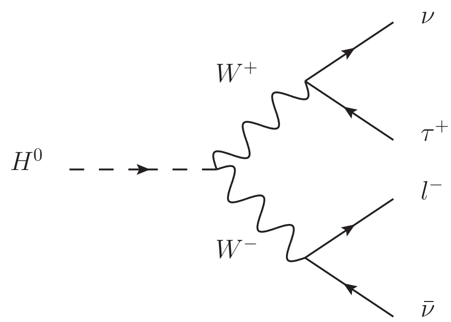


Figure 7.1: SM Higgs decay via WW into final states with a τ lepton.

7.1.1 Signal and background processes

Signal in $H \to \tau\tau$

At the LHC, the SM Higgs boson is predicted to be produced mostly in the gluon-gluon fusion ($gg \to H$) and vector-boson fusion (VBF) processes. The leading order Feynman diagrams for these processes are shown in Figures 2.4(a) and 2.4(b). In the leading-order (LO) diagrams, the two taus produced in the Higgs decay are back-to-back in the plane transverse to the beam axis. The tau decays lead to a relatively low missing transverse momentum from the undetected neutrinos. In next-to-leading-order (NLO) QCD, the Higgs boson can be produced in association with at least one hadronic jet. This has the effect of producing events with higher missing transverse momentum since the neutrinos are no longer back-to-back in the transverse plane.

Signal in $H \to WW^* \to \ell\nu\tau\nu$

In contrast, the light lepton and τ_h particles in signal events produced via a pair of W bosons will have a much smaller angular separation due to constraints on the final state particle helicities dictated by the spin of the Higgs boson parent, as illustrated in Figure 7.2. For a spin-0 parent particle such as a SM Higgs boson, the visble decay particles (in this case an $\ell - \tau_h$ pair) are emitted with a small angular separation. This offers the most powerful discriminant against the dominant background processes which tend to produce these particles almost back-to-back in the plane transverse to the incoming proton beams.

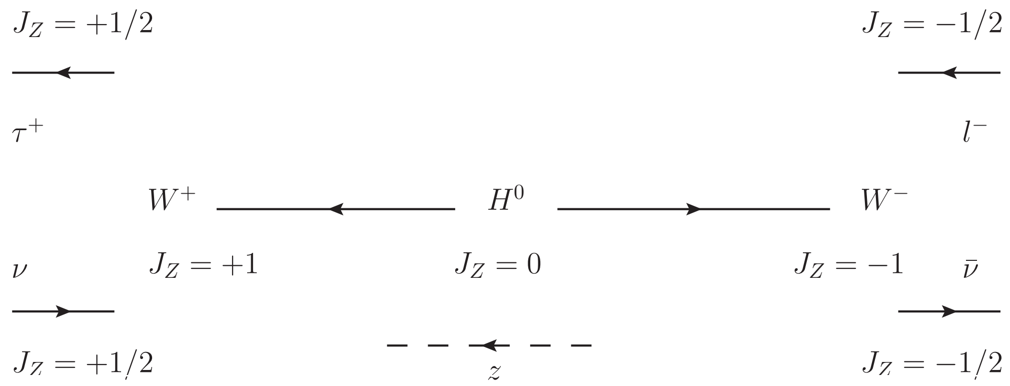


Figure 7.2: The effect of spin correlations on the angular separation of the tau and light lepton decay products in $H \to WW^* \to \ell\nu\tau\nu$ events. In this simplified example the solid lines represent the momentum vectors of the W bosons, τ and lepton in a one dimensional Higgs decay. In $\mathrm{SU(2)_L}$, the W can couple to left handed neutrinos and right-handed anti-neutrinos and so to satisfy angular momentum conservation in this example, the momentum vector of $\ell(\tau)$ must be parallel (anti-parallel) to its helicity vector such that both momentum vectors point in the positive z direction.

Background

The following processes are considered as backgrounds in these analyses:

- Inclusive $Z \to \ell\ell$: the largest source of background for the $H \to \tau\tau$ analysis is irreducible, with a di-tau final state produced in $Z \to \tau\tau$ (and to a lesser degree the Drell-Yan process $q\bar{q} \to \gamma^* \to \tau^+\tau^-$). The decay products in these events have similar kinematics to $H \to \tau\tau$ signal events due to the small mass difference between the Z-boson and a light SM Higgs boson. Similar processes where the Z boson or virtual photon decay to an electron-positron (or muon-anti-muon) pair also contribute if the light lepton or any additional jet produced in the event is mis-identified as a hadronically decaying tau lepton.

- $W(\to \ell\nu)$ +jets: the production of a W boson in association with jets forms a significant background due to the relatively large production cross section and branching ratio of a W boson decay to a charged lepton, significant missing transverse momentum from the neutrino produced in the leptonic W decay, and the fact that any jet in the event can be mis-identified as a hadronically decaying tau lepton. After making a requirement on the angular separation between the light lepton and the τ_h, events from $W(\to \ell\nu)$ +jets are the primary background in the $H \to WW^* \to \ell\nu\tau\nu$ analysis.

- $t\bar{t}$ production: the process $pp \to t\bar{t} \to W^+b\ W^-\bar{b}$ can lead to events with missing transverse momentum, a light lepton and a tau lepton of opposite charge if both W bosons decay leptonically. There is also a small probability that a jet produced by either b quark could be mis-identified as a τ_h.

- Single top production: s or t channel single top production or single top in association with a W boson can also imitate the signal process final state if a W boson decays to a light lepton and either a W boson decays to a τ_h or a jet arising from a b quark is mis-identified as a τ_h.

- Diboson production: production of a pair of electroweak bosons can result in a pair of charged leptons and missing transverse momentum from the boson decays.

- QCD multi-jet production: events with multiple hadronic jets arising from QCD processes form an important background since the cross section is very large and there is a non-negligible probability that in a given event one jet can be mis-identified as a light lepton and another as a τ_h.

7.2 Datasets

7.2.1 Data

The collision data were collected between March and October in 2011 during which time the LHC was operating at a centre-of-mass energy of $\sqrt{s} = 7$ TeV. In data periods used, all of the relevant sub-detectors were performing optimally. The dataset corresponds to an integrated luminosity of $\int \mathcal{L}\ \mathrm{dt} = 4.7 \pm 0.2\ \mathrm{fb}^{-1}$ selected with high p_{T} light-lepton triggers.

7.2.2 Simulation

To understand the signal and background processes and to estimate their contributions, Monte Carlo generators were employed. For the production of W and Z bosons, the `Alpgen` generator [46] was used to generate matrix elements with up to five additional partons. `MC@NLO` [65] was used to simulate diboson, $t\bar{t}$ and single top events. The parton shower and hadronisation for these samples was performed by the `Herwig` [66] generator. In addition, the underlying event from additional soft QCD interactions was simulated using `Jimmy` [67].

The `Powheg` [68] generator was used in combination with `Pythia` [47] to generate and correctly model the $gg \to H$ and VBF production signal events with subsequent $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ decays. Events with a SM Higgs produced in association with a weak boson were simulated using `Pythia` [47]. For all simulated samples, `TAUOLA` [69] and `PHOTOS` [70] were used to model the tau decay and photon radiation from the charged leptons, respectively. All Monte Carlo events are fully simulated using `GEANT4` [37] and reconstructed in the same way as the collision data. The cross sections for each signal and background process are summarised in Appendix A.1.

7.3 Di-tau mass reconstruction in $H \to \tau\tau$ events

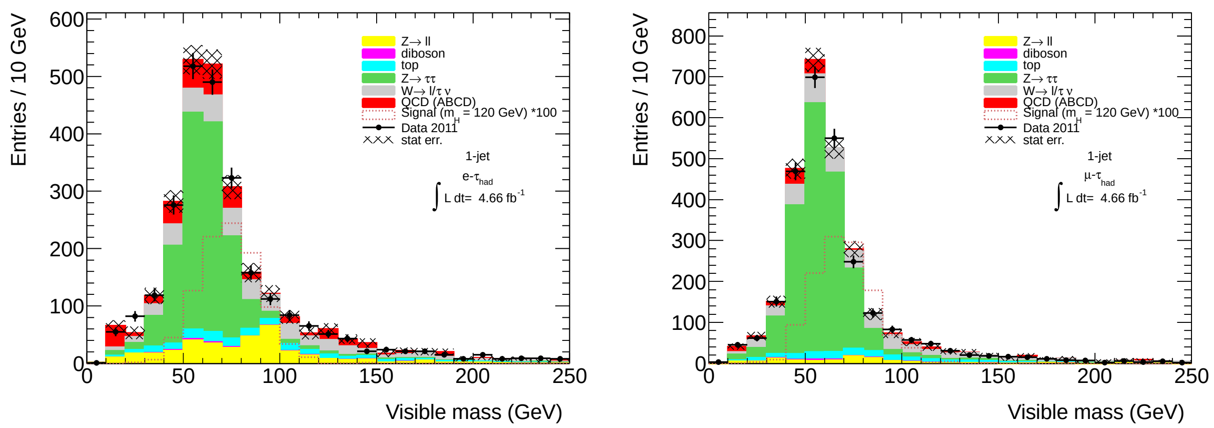
Each tau decay will produce at least one undetected neutrino, making a conventional reconstruction of the original Higgs four-momentum using the decay products alone not possible. Hence, the Higgs mass cannot be fully reconstructed. In this section, three methods to reconstruct the mass of the di-tau system are explored. For the $H \to \tau\tau$ analysis, the MMC method was used as it offers the best mass resolution of the three methods, and hence can better discriminate against the primary background process; $Z \to \tau\tau$.

7.3.1 Visible mass

Despite the mass difference between m_Z and a SM Higgs boson with $110 < m_H < 150$ GeV, the invariant mass distributions of the visible tau decay products,

$$\text{visible mass} = \sqrt{(p_{\tau_h} + p_\ell) \cdot (p_{\tau_h} + p_\ell)} = \sqrt{(E_{\tau_\text{h}} + E_\ell)^2 - (\vec{p}_{\tau_h} \cdot \vec{p}_\ell)^2}, \tag{7.1}$$

offer little separation between $Z \to \tau\tau$ and $H \to \tau\tau$ due to the undetected neutrino(s) produced in each tau decay. The visible mass of the tau decay products from signal and background processes is shown in Figure 7.3 for events passing the selection requirements described in Section 7.5.1, for events with at least one additional hadronic jet.



(a) $e - \tau_h$ (b) $\mu - \tau_h$

Figure 7.3: Data / MC comparison of the visible mass distribution of the $H \to \tau\tau$ signal regions with at least one additional hadronic jet after applying the selection criteria described in Section 7.5.1. Data-driven methods described in Sections 7.7 and 7.8 are used to estimate the $W(\to \ell\nu)$ +jets and QCD backgrounds. The dotted-line histograms show the expected signal yield of gluon-gluon fusion (ggF) and vector-boson fusion (VBF) Higgs production, for $m_H = 120$ GeV with 100 times the integrated luminosity of the analysis.

7.3.2 Collinear mass

Since the sources of missing momentum originate from the Higgs boson, the so-called collinear mass approximation can also be used to reconstruct the Higgs mass. In the collinear mass approximation, it is assumed that the decay products of each τ are collinear with the τ in the laboratory frame since $\frac{m_H}{2} >> m_\tau$, such that the taus are highly boosted in the laboratory frame. Neglecting the τ rest mass and imposing that the neutrinos in each tau decay are collinear with the visible tau-decay products, the collinear di-tau invariant mass of a system with $\tau \to \tau_h \nu$ and $\tau \to \ell\nu\nu$ can be written

$$m_{\tau\tau} = \sqrt{2(E_{\tau_h} + E_{\nu\tau_h})(E_\ell + E_{\nu\ell})(1 - \cos\theta_{\ell\tau_h})}, \tag{7.2}$$

which can also be written as

$$m_{\tau\tau} = \frac{m_{\ell\tau_h}}{\sqrt{\chi_\ell \chi_{\tau_h}}} \quad \text{for } \chi_{\ell,\tau_h} \geq 0, \tag{7.3}$$

where $E_{\nu\tau_h}$ and $E_{\nu\ell}$ is the energy (sum) of the neutrino(s) produced in the hadronic (leptonic) tau decay, χ_{l,τ_h} is the fraction of the tau's momentum taken by the visible decay products and $m_{\ell\tau_h}$ is the invariant mass of the $\ell - \tau_h$ pair. Since there are no other source of real $p_{\mathrm{T}}^{\mathrm{miss}}$ in the event, χ_{ℓ,τ_h} can be also be expressed in terms of the visible tau decay products momenta and the $p_{\mathrm{T}}^{\mathrm{miss}}$

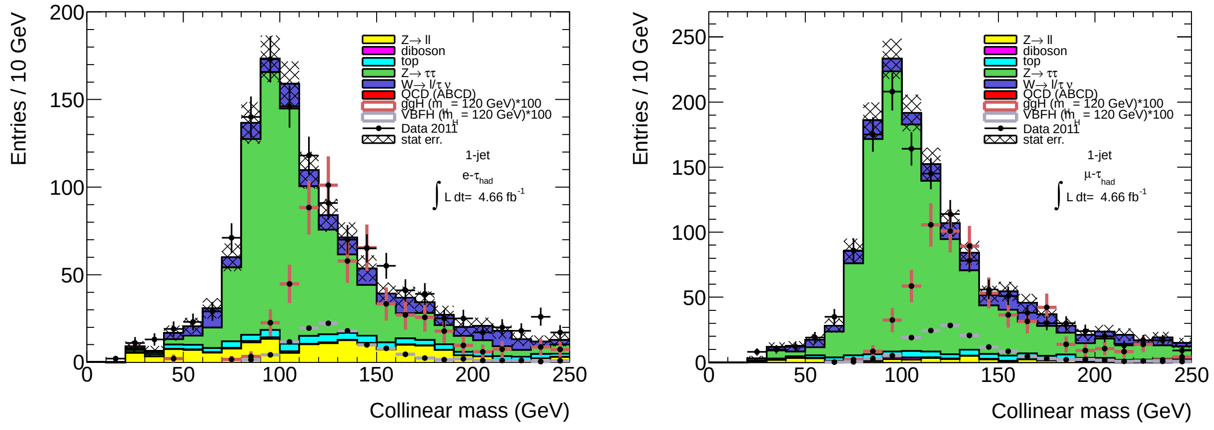
$$\chi_{\tau_l} = \frac{E_\ell}{E_\ell + E_{\nu\ell}} = \frac{p_x^\ell p_y^{\tau_h} - p_x^{\tau_h} p_y^\ell}{p_y^{\tau_h} p_x^\ell + p_x^{miss} p_y^{\tau_h} - p_x^{\tau_h} p_y^\ell - p_y^{miss} p_x^{\tau_h}}, \tag{7.4}$$

and

$$\chi_{\tau_h} = \frac{E_{\tau_h}}{E_{\tau_h} + E_{\nu\tau_h}} = \frac{p_x^\ell p_y^{\tau_h} - p_x^{\tau_h} p_y^\ell}{p_y^{\tau_h} p_x^\ell + p_y^{miss} p_x^\ell - p_x^{\tau_h} p_y^\ell - p_x^{miss} p_y^\ell}. \tag{7.5}$$

The collinear mass of the tau decay products from signal and background processes is shown in Figure 7.4 for events passing the selection requirements described in Section 7.5.1 for events with at least one additional hadronic jet.

The collinear mass was not used in the $H \to \tau\tau$ analysis for two reasons. First, it tends to over-estimate the mass of the system (seen in Figure 7.4 as an excessive tail in the $Z \to \tau\tau$ distribution) due to its sensitivity to the $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution. Second, the fraction of signal events which must be discarded due to finding a non-physical mass solution is about 30%. While this is an effective cut against backgrounds, it reduces the overall sensitivity of each channel.



(a) $e - \tau_h$. (b) $\mu - \tau_h$.

Figure 7.4: Data / MC comparison of the collinear mass distribution of the $H \to \tau\tau$ signal regions with at least one additional hadronic jet after applying the selection criteria described in Section 7.5.1. Data-driven methods described in Sections 7.7 and 7.8 are used to estimate the $W(\to \ell\nu)$ +jets and QCD backgrounds.

7.3.3 Missing Mass Calculator

Better discriminating power is obtained by reconstructing the di-tau invariant mass using the so-called Missing Mass Calculator (MMC) [71], which aims to fully reconstruct the event topology. The MMC algorithm can be thought of as an extension of the collinear mass approximation where a small opening angle is allowed between the neutrino(s) and the visible decay products of each tau decay. In this technique a set of simultaneous equations is constructed using the mass of each tau parent and orthogonal projections of the $p_{\mathrm{T}}^{\mathrm{miss}}$

$$
\begin{aligned}
p_x^{\mathrm{miss}} &= p_{\mathrm{miss1}} \sin\theta_{\mathrm{miss1}} \cos\phi_{\mathrm{miss1}} + p_{\mathrm{miss2}} \sin\theta_{\mathrm{miss2}} \cos\phi_{\mathrm{miss2}}, \\
p_y^{\mathrm{miss}} &= p_{\mathrm{miss1}} \sin\theta_{\mathrm{miss1}} \sin\phi_{\mathrm{miss1}} + p_{\mathrm{miss2}} \sin\theta_{\mathrm{miss2}} \sin\phi_{\mathrm{miss2}}, \\
m_{\tau 1}^2 &= m_{\mathrm{miss1}}^2 + m_{\mathrm{vis1}}^2 + 2\sqrt{p_{\mathrm{vis1}}^2 + m_{\mathrm{vis1}}^2}\sqrt{p_{\mathrm{miss1}}^2 + m_{\mathrm{miss1}}^2} - 2p_{\mathrm{vis1}}p_{\mathrm{miss1}} \cos\Delta\theta_{\nu m1}, \\
m_{\tau 2}^2 &= m_{\mathrm{miss2}}^2 + m_{\mathrm{vis2}}^2 + 2\sqrt{p_{\mathrm{vis2}}^2 + m_{\mathrm{vis2}}^2}\sqrt{p_{\mathrm{miss2}}^2 + m_{\mathrm{miss2}}^2} - 2p_{\mathrm{vis2}}p_{\mathrm{miss2}} \cos\Delta\theta_{\nu m2},
\end{aligned}
\tag{7.6}
$$

where p_x^{miss} and p_y^{miss} are projections of the $p_{\mathrm{T}}^{\mathrm{miss}}$ vector along the x and y axes, $p_{\mathrm{vis1,2}}$, $m_{\mathrm{vis1,2}}$ are the momenta and invariant mass of the visible tau decay products and $m_\tau = 1.777$ GeV. The neutrino 4-momenta for each tau decay are combined into $p_{\mathrm{miss1,miss2}}$ with invariant mass

m_{miss1} (in tau decays in which two neutrinos are produced i.e. leptonic tau decay), with the polar angular difference between $\vec{p}_{\text{vis1,2}}$ and $\vec{p}_{\text{miss1,2}}$ denoted by $\Delta\theta_{\nu m1,2}$ for each tau decay.

This system is under-constrained, with six degrees of freedom and four constraints. However, a likelihood can be obtained for a particular set of solutions using additional information the τ decay kinematics. For the MMC method, the expected three-dimensional angle between $\vec{p}_{\text{vis1,2}}$ and $\vec{p}_{\text{miss1,2}}$, $\delta\theta_{3D}$ is used. Figure 7.5 shows the expected distribution of $\delta\theta_{3D}$ for three different scenarios: hadronic tau decay with one or three associated charged hadron track(s) and leptonic tau decay.

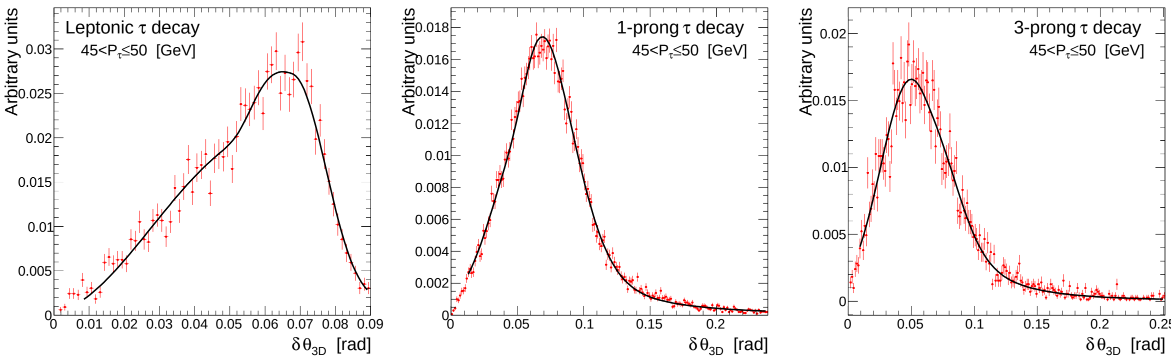


Figure 7.5: Example of the probability distribution functions $\mathcal{P}(\Delta\theta, p_\tau)$ for a particular value of the original τ lepton momentum ($45 < p_\tau \leq 50$ GeV). These functions are used in the calculation of the global event probability $\mathcal{P}_{event}$ for three cases: leptonic decays (left plot), 1-prong hadronic decays (middle plot), and 3-prong hadronic decays (right plot) of τ leptons. These distributions depend only on the decay type and initial transverse momentum of the τ lepton.

The system of equations can be solved simultaneously for any arbitrary choice of $(\phi_{\text{miss1}}, \phi_{\text{miss2}})$. With this information $\vec{p}_{\text{miss1,2}}$ and hence $\Delta\theta^{1,2}_{3D}$ are defined. The probability of each point in this parameter space is calculated using the $\delta\theta_{3D}$ distributions shown in Figure 7.5, defined for each tau decay type and initial tau momentum from $10 < p_{\text{T}}^{\tau} < 230$ GeV in 5 GeV bins of p_{T}^{τ}. Since this opening angle is proportional to Lorentz boost of the τ, these distributions are parametrised as a function of the τ_h p_{T}. For each region of this phase space, the $\delta\theta_{3D}$ distribution is fit by a linear combination of Gaussian and Landau functions with the mean and width of p_τ along with the relative Gaussian and Landau normalisations parametrised by

$$\mathcal{P}(p_\tau) = a_0(\exp^{-a_1 \cdot p_\tau} + a_2/p_\tau), \tag{7.7}$$

where a_i are the parametrisation coefficients. The full PDF for $\mathcal{P}(\Delta\theta_{3D}, p_\tau)$ can then be used to evaluate a probability for a given tau decay topology. A global event probability for a di-tau decay is then defined as

$$\mathcal{P}_{\text{event}} = \mathcal{P}(\Delta\theta^1_{3D}, p^1_\tau) \times \mathcal{P}(\Delta\theta^2_{3D}, p^2_\tau) \tag{7.8}$$

By scanning over a grid over the full range of possible $(\phi_{\text{miss1}}, \phi_{\text{miss2}})$, an $m_{\tau\tau}$ distribution is obtained with each value weighted by its corresponding $\mathcal{P}_{\text{event}}$. The position of the maximum of this distribution is then used to estimate the physical $m_{\tau\tau}$ of each event.

The performance of this algorithm is highly dependent on the $p_{\text{T}}^{\text{miss}}$ resolution. Mis-measurement of the $p_{\text{T}}^{\text{miss}}$ is incorporated by convolving P_{event} with Gaussian $p_{\text{T}}^{\text{miss}}$ resolution functions:

$$\mathcal{P}(p^{\text{miss}}_{x,y}) = \exp\frac{-(\Delta p^{\text{miss}}_{x,y})^2}{2\sigma^2}, \tag{7.9}$$

where σ is the $p_{\text{T}}^{\text{miss}}$ resolution and $\Delta p_{x,y}$ are the differences between $p_{\text{T}}^{\text{miss}}$ projections onto the x and y axis as determined from the choice of $(\phi_{\text{miss1}}, \phi_{\text{miss2}})$ and that of the event's measured $p_{\text{T}}^{\text{miss}}$ vector. The global event probability then becomes

$$\mathcal{P}_{\text{event}} = \mathcal{P}(\Delta\theta^1_{3D}, p^1_\tau) \times \mathcal{P}(\Delta\theta^2_{3D}, p^2_\tau) \times \mathcal{P}(p^{\text{miss}}_x) \times \mathcal{P}(p^{\text{miss}}_y). \tag{7.10}$$

For a given event, the so-called 'MMC mass' is the invariant mass of the hadronic tau, light lepton and hypothesised neutrino four-vectors that produce the highest value of $\mathcal{P}_{\text{event}}$, after a scan over $m_{\tau\tau}$ and the full range of $(\phi_{\text{miss1}}, \phi_{\text{miss2}})$.

For the $H \to \tau\tau$ analysis, the MMC mass is used since it provides the best discrimination of the signal events from the primary background, $Z \to \tau\tau$. In the $H \to WW^* \to \ell\nu\tau\nu$ analysis, the visible mass is used.

7.4 Common event selection criteria

7.4.1 Trigger

To preselect events from the collision data, triggers were used that select a reconstructed light lepton in a given event. For the $e - \tau_h$ channel, the Event Filter trigger required the event to have a cluster of transverse energy in the EM calorimeter $E_T > 20$ GeV or $E_T > 22$ GeV for the early or later data taking periods, respectively. This adjustment was necessary due to changing collision conditions throughout the year, with a greater number of pile-up interactions in the later periods. For the $\mu - \tau_h$ channel, the trigger required a muon candidate with $p_{\mathrm{T}} > 18$ GeV at the Event Filter level.

7.4.2 Pileup re-weighting

The total number of interactions per bunch crossing averaged over the luminosity for that block[2] is defined as the average μ for that block [72].

Since object reconstruction, identification and mis-identification efficiencies are sensitive to additional interactions, it is necessary for the Monte Carlo samples to simulate additional interactions as well as the primary simulated process. The events are weighted such that the average μ distribution for each period of data-taking is replicated in the simulated datasets.

7.4.3 Muon selection

A reconstructed muon candidate is required to have $p_{\mathrm{T}} > 10$ GeV with $|\eta| < 2.5$. It is further required that in a cone of radius $\Delta R = 0.2$ around the muon the sum of additional transverse energy deposited in the electromagnetic and hadronic calorimeters be less than 4% of the candidate's p_{T}. Additionally, the scalar p_{T} sum of additional inner detector tracks that have $p_{\mathrm{T}} > 1$ GeV in a cone of radius $\Delta R = 0.4$ around the muon candidate is required to be less

[2]Luminosity is measured in atomic units corresponding to approximately 2 minutes of ATLAS data-taking known as luminosity blocks, but the duration can vary due to run conditions and other operational issues.

than 6% of the muon's p_{T}. Simulated data are corrected to account for observed differences in the muon p_{T} resolution and identification efficiency between data and simulation [73]; the corrections are $\simeq 1\%$. If the event was triggered by a muon object, a muon candidate passing these selection criteria with $p_{\mathrm{T}} > 20$ GeV is required to be within a cone of radius $\Delta R = 0.2$ from the trigger object.

7.4.4 Electron selection

Reconstructed electrons are required to have $E_{\mathrm{T}} > 15$ GeV and $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$[3]. They are required to pass the tight selection criteria (described in Chapter 3) and are further required to be isolated in the calorimeter, such that the sum of additional transverse energy deposited in a cone of radius $\Delta R = 0.2$ from the electron is less than 8% of the electron's E_{T}. Candidate electrons are also required to be isolated in the inner detector; the scalar p_{T} sum of other inner detector tracks with $p_{\mathrm{T}} > 1$ GeV within a cone of radius $\Delta R = 0.4$ from the electron must be less than 6% of the electron's p_{T}.

In analogy with the muon-based corrections, simulated data are corrected to account for observed differences in the electron energy resolution and identification efficiency between data and simulation [74].

If the event was triggered by an electron object, an electron candidate passing this selection criteria with $p_{\mathrm{T}} > 25$ GeV is required to be within a cone of radius $\Delta R = 0.2$ from the trigger object.

7.4.5 τ_h selection

Reconstructed τ_h candidates are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$, with exactly 1 or 3 reconstructed inner detector tracks and track charge sum equal to ± 1. In the $H \to \tau\tau$ analysis, candidates are required to pass the medium BDT-based multivariate identification

[3]The transition region where the barrel and endcap electromagnetic calorimeters overlap ($1.37 < |\eta| < 1.52$) is not used since the identification criteria are less effective in this region.

requirement while τ_h candidates in the $H \to WW^* \to \ell\nu\tau\nu$ analysis are required to pass the tight BDT-based multivariate identification requirements. These are described in Section 6.3 and reference [27].

7.4.6 Hadronic jet selection

Hadronic jet candidates are required to pass several data quality cuts (described in reference [63]) in order to suppress detector backgrounds and backgrounds from out-of-time pileup. Jets are also required to have $p_\mathrm{T} > 25$ GeV and $|\eta| < 4.5$. Finally, a cut is placed on the 'jet vertex fraction' (JVF), defined to be the fraction of the jet's total track p_T originating from the primary interaction vertex: the requirement is JVF > 0.75 for jets with $|\eta| < 2.4$. This final cut allows for the suppression of jet backgrounds from pile-up without raising the jet p_T threshold.

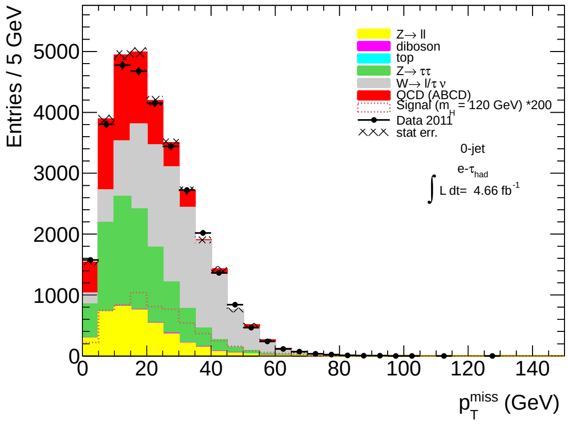
7.4.7 Overlap removal

Often the same final state object may be reconstructed as an object by more than one object reconstruction algorithm. In the case that different selected object candidates are reconstructed within a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.2$, one of them is discarded. These overlapping candidates are resolved by selecting muons, electrons, τ_h and finally jets, in that order of priority. For the overlap removal procedure only, the muon calorimeter isolation requirement is ignored and the electron identification requirement is relaxed to the medium selection criteria.

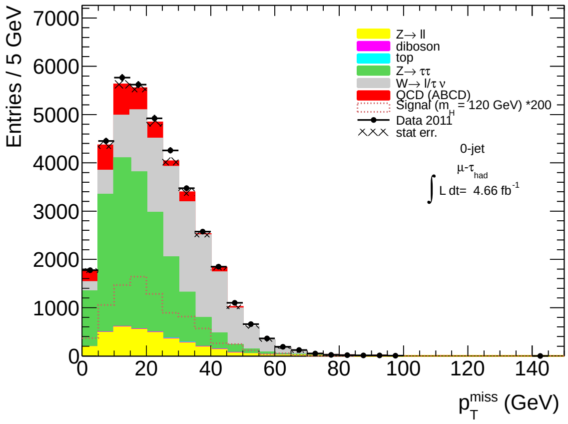
7.5 Optimisation of the $H \to \tau\tau$ event selection criteria

In order to maximise the expected sensitivity of the analysis to SM Higgs boson signal events, the signal significance is maximised under variation of the selection cuts. The signal significance is defined as $n_S/\sqrt{n_B}$ where n_S is the expected number of signal events and n_B is the expected number of background events from all sources. The W+jets and QCD multi-jet background yields for each set of cuts is estimated using the prescriptions in Section 7.7 and Section 7.8, respectively. The varied selection criteria are:

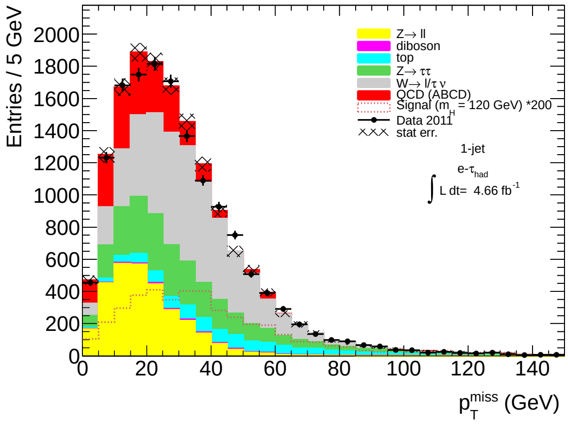
- The $p_{\mathrm{T}}^{\mathrm{miss}}$ requirement (shown in Figure 7.6 after the event pre-selection) was varied in steps of 10 GeV from 0 to 40 GeV.



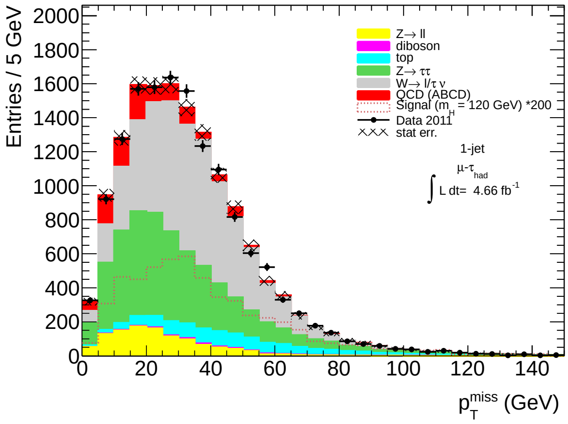
(a) $e - \tau$, 0 jet



(b) $\mu - \tau$, 0 jet



(c) $e - \tau$, ≥ 1 jet



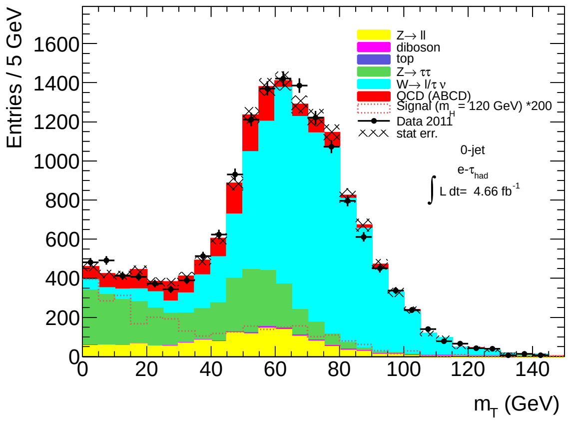
(d) $\mu - \tau$, ≥ 1 jet

Figure 7.6: Data / MC comparison of the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the 0-jet and 1-jet signal regions after applying the W+jets scale factor, as described in Section 7.7. The dotted-line histograms show the expected signal yield of gluon-gluon fusion and vector-boson fusion Higgs production, for $m_H = 120$ GeV with 200 times the integrated luminosity of the analysis.

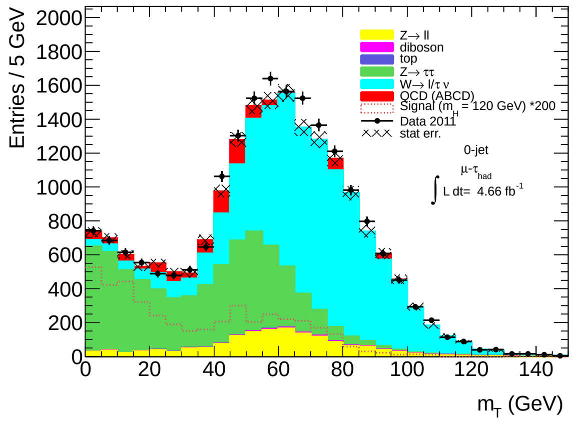
- The transverse mass of the light lepton and the $p_{\mathrm{T}}^{\mathrm{miss}}$, m_T, is defined to be

$$m_T = \sqrt{2 p_{\mathrm{T}}^{\ell} p_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos \Delta\phi[\ell, p_{\mathrm{T}}^{\mathrm{miss}}])}, \tag{7.11}$$

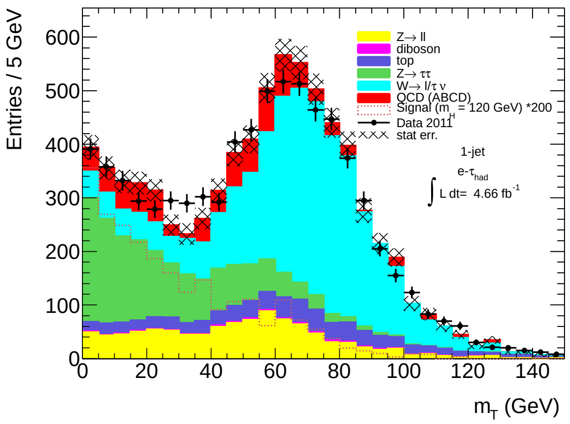
where p_{T}^{ℓ} is the light lepton p_{T}, and $\Delta\phi[\ell, p_{\mathrm{T}}^{\mathrm{miss}}]$ is the angular separation of the light lepton and the $p_{\mathrm{T}}^{\mathrm{miss}}$ in the transverse plane. The transverse mass is effective in separating W+jets and Z+jets events from signal. An upper limit of the transverse mass of the event was varied from 20 to 40 GeV in steps of 10 GeV. Signal and background m_T distributions of events with $p_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV are shown in Figure 7.7.



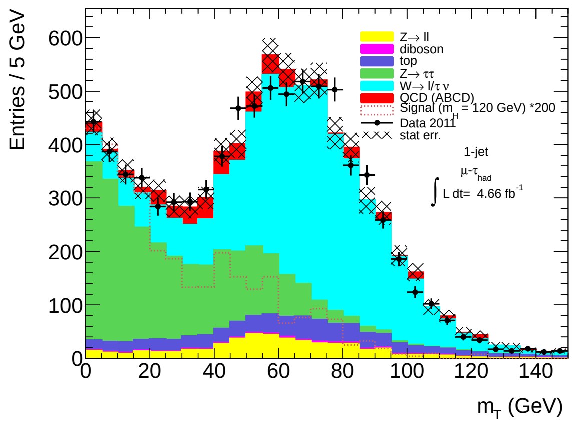
(a) $e - \tau$, high p_T^{miss}, 0 jet



(b) $\mu - \tau$, high p_T^{miss}, 0 jet



(c) $e - \tau$, high p_T^{miss}, ≥ 1 jet



(d) $\mu - \tau$, high p_T^{miss}, ≥ 1 jet

Figure 7.7: Data / MC comparison of the transverse mass distribution in the 0-jet and $\geq$ 1-jet signal regions after applying the f_W scale factor to the expected yield of $W(\to \ell\nu)$ +jets from simulation, as described in Section 7.7.

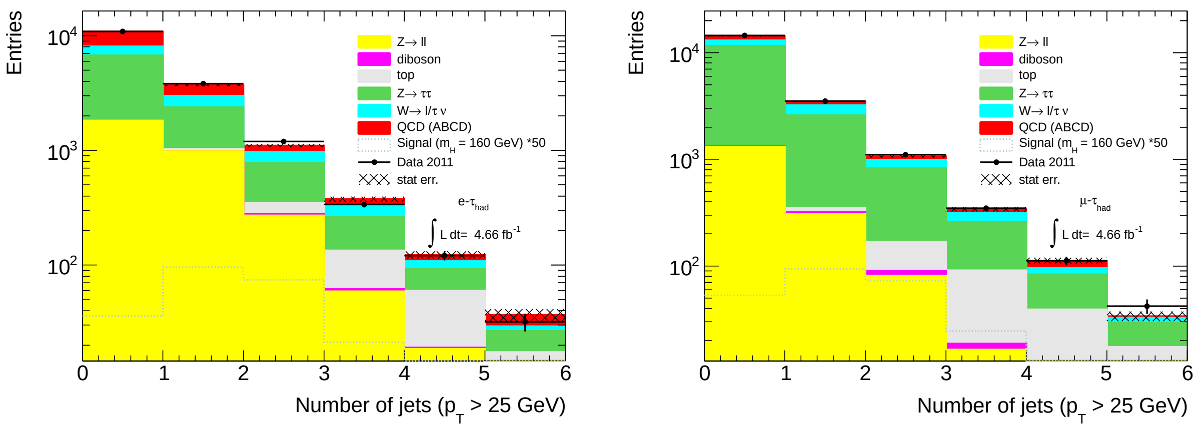
- The degree to which the reconstructed tau candidate has been identified as a τ_h is varied from 'medium' to 'tight' selection. These levels correspond to a τ_h identification efficiency of 35% and 45%, respectively.

7.5.1 $H \to \tau\tau$ event selection criteria

After optimising for $H \to \tau\tau$ signal, events are required to pass the following selection criteria:

- Exactly one identified τ_h and exactly one identified, isolated light lepton candidate. These candidates must have opposite charge, i.e. $q_{\tau_h} \times q_\ell < 0$.
- $m_T < 30$ GeV.

The exclusive number of additonal hadronic jets distribution of events passing these selection criteria are shown in Figure 7.8.



(a) $e - \tau$ events.

(b) $\mu - \tau$ events.

Figure 7.8: Data / MC comparison of the number of additional hadronic jets of events with $m_T < 30$ GeV. Data-driven methods described in Sections 7.7 and 7.8 are used to estimate the $W(\to \ell\nu)$ +jets and QCD backgrounds.

Additionally,

- Since the p_{T} thresholds for electrons and muons is different, $e - \tau_h$ and $\mu - \tau_h$ events are treated as separate signal regions, since this affects the background composition.
- Events with $p_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV and $p_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV are treated as separate signal regions.
- Also, events in the $e - \tau_h$ and $\mu - \tau_h$ regions that do not pass the VBF region selection criteria (see Section 7.5.2) are further split by whether the number of additional jets in the event passing the selection criteria described in Section 7.4.6 is either 0 or ≥ 1.

7.5.2 VBF event selection

Despite the lower cross section for the production of a Higgs boson via vector-boson-fusion, the distinct topology of these events due to the lack of colour exchange between the outgoing partons enhances discrimination of signal from background processes. In addition to the selection criteria above, the variables used to exploit these differences to create a VBF signal region using the highest and second-highest p_{T} jets (hereafter referred to as jet 1 and jet 2, respectively) are:

- The di-jet invariant mass ($m_{jet1,jet2}$), shown in Figure 7.9(a). Events in the VBF signal region are required to have $m_{\mathrm{jet1,jet2}} > 300$ GeV.

- The pseudorapidity difference of these jets, as shown in Figure 7.9(b), is required to be larger than 3, and $\eta_{\mathrm{jet1}} \times \eta_{\mathrm{jet2}} < 0$.

- A centrality requirement on the reconstructed lepton and τ_h pseudorapidity, such that

$$\min(\eta_{\mathrm{jet1}}, \eta_{\mathrm{jet2}}) < \eta^{\ell}, \eta^{\tau_h} < \max(\eta_{\mathrm{jet1}}, \eta_{\mathrm{jet2}}). \tag{7.12}$$

Events that pass this selection criteria are treated as a separate signal region. Events with two or more jets that fail the VBF criteria but pass the more general described selection cuts in Section 7.5.1 are added to the ≥ 1 jet signal regions.

7.5.3 $H \to \tau\tau$ results

The total number of expected signal and background events in each $H \to \tau\tau$ signal region is shown in Table 7.1. The signal yield in the ≥ 1 jet signal regions predominantly comes from the gluon-gluon fusion Higgs production mechanism. The MMC mass distribution for each of the seven $H \to \tau\tau$ signal regions are shown in Figures 7.10 and 7.11.

Process Group	0 jet			
	$e - \tau_h$		$\mu - \tau_h$	
	low p_T^{miss}	high p_T^{miss}	low p_T^{miss}	high p_T^{miss}
$Z/\gamma^* \to \tau^+\tau^-$	3546 ±199	1344 ±125	7467 ±411	2772 ±256
$W(\to \ell\nu)$ +jets	790 ±83	355 ±64	942 ±97	390 ±72
$Z/\gamma^* \to \ell\ell$	1382 ±344	331 ±132	895 ±224	184 ±74
(Single) top	1.2 ±0.6	1.7 ±0.6	1.3 ±0.6	2.6 ±0.9
Diboson	7.4 ±1.1	4.6 ±1.8	9.5 ±1.1	6.3 ±1.4
QCD multi-jet	2738 ±105	516 ±36	1450.0 ±21	269 ±4
Total	8465 ±409	2553±194	10765 ±477	3623 ±275
Data 2011	8248	2512	10886	3563
Signal ($m_H = 120$ GeV)	9.3 ±1.3	7.8 ±1.2	15.2 ±2.1	10.7 ±1.2

Process Group	≥ 1 jet		VBF
	$e - \tau_h$	$\mu - \tau_h$	$\ell - \tau_h$
$Z/\gamma^* \to \tau^+\tau^-$	1263 ±96	1843 ±133	45±4
$W(\to \ell\nu)$ +jets	411 ±58	465 ±71	24±9
$Z/\gamma^* \to \ell\ell$	355 ±92	90 ±32	12±5
(Single) top	168 ±12	172 ±12	12±0.9
Diboson	14.7 ±3.8	15.8 ±3.1	1.2±0.3
QCD multi-jet	354 ±20	120 ±5	20±2
Total	2567 ±145	2707 ±153	115±11
Data 2011	2574	2707	122
Signal ($m_H = 120$ GeV)	9.8 ±2.0	12.3 ±2.5	3.0±0.4

Table 7.1: The total number of signal and background events in each $H \to \tau\tau$ signal region where the uncertainties take into account the finite MC sample statistics and all sources of experimental error described in Section 7.9 . The last row shows the number of events found in data.

7.6 Optimisation of the $H \to WW^* \to \ell\nu\tau\nu$ event selection criteria

Due to the different topology of $H \to WW^* \to \ell\nu\tau\nu$ events relative to $H \to \tau\tau$ events, it is necessary to apply separate selection criteria. In this section, the optimisation method described in Section 7.5 is employed using variables designed to exploit the small angular separation expected between the τ_h and light lepton. As in Section 7.5.1, the W+jets and QCD background yields for each set of cuts is estimated using the data-driven methods described in Section 7.7 and Section 7.8, respectively. The variables used in this optimisation included:

- the $p_{\mathrm{T}}^{\mathrm{miss}}$ of the event (Figure 7.12), on which the cut was varied from $p_{\mathrm{T}}^{\mathrm{miss}} > 0$ GeV to $p_{\mathrm{T}}^{\mathrm{miss}} > 40$ GeV in 10 GeV steps;
- the transverse mass of the event (defined in Equation 7.11) was varied from $m_T > 30$ GeV to $m_T > 50$ GeV in 10 GeV steps;
- and the angular separation of the light lepton and τ_h, defined as

$$\Delta R(\ell, \tau_h) = \sqrt{\Delta\eta(\ell, \tau_h)^2 + \Delta\phi(\ell, \tau_h)^2}, \tag{7.13}$$

  and shown in Figure 7.13. The cut was varied from $\Delta R(\ell, \tau_h) < 2.05$ to $\Delta R(\ell, \tau_h) < 1$ in steps of 0.15.

7.6.1 $H \to WW^* \to \ell\nu\tau\nu$ event selection cuts

As a result of the $H \to WW^* \to \ell\nu\tau\nu$ signal optimisation with $m_H = 125$ GeV, events were required to pass the following selection criteria:

- Exactly one tightly identified τ_h and exactly one identified, isolated light lepton candidate. The candidates must be of opposite charge $q_{\tau_h} \times q_\ell < 0$.

- $m_T > 30$ GeV.
- $p_{\mathrm{T}}^{\mathrm{miss}} > 10$ GeV.
- $\Delta R(\ell, \tau_h) < 1.15$.

Due to the different lepton p_{T} cuts, events with a final state $e - \tau_h$ pair and $\mu - \tau_h$ pair are treated as separate signal regions.

The total numbers of expected signal and background events in each $H \to WW^* \to \ell\nu\tau\nu$ signal region are shown in Table 7.2 and the visible mass distribution for both of the $H \to WW^* \to \ell\nu\tau\nu$ signal regions are shown in Figure 7.14.

Process Group	$e - \tau$	$\mu - \tau$
$W(\to \ell\nu)$ +jets	598±71	757±77
(single) top	69 ±4	69± 3
$Z \to \ell\ell$	35±12	57 ±18
$Z \to \tau\tau$	32±10	33±8
Diboson	7.2±1.4	8.6±1.1
QCD multi-jet	19.1±4.7	12.2 ±0.9
Total	760±73	937±80
Signal ($m_H = 125$ GeV)	1.1±0.1	1.4±0.2
Signal ($m_H = 160$ GeV)	13.1±0.8	13.9±0.9
Data 2011	713	852

Table 7.2: The total number of signal and background events in each $H \to WW^* \to \ell\nu\tau\nu$ signal region where the uncertainties take into account the finite MC sample statistics and all sources of experimental error described in Section 7.9. The last row shows the number of events found in data.

7.7 Data-driven estimate of the $W(\to \ell\nu)$ +jets background

7.7.1 Introduction

To correctly model the yield of background events in which a hadronic jet has been misidentified as a τ_h, i.e. the $W(\to \ell\nu)$ +jets background, we cannot rely on Monte Carlo alone since the hadronisation of partons is not generally well modelled. Therefore, we construct regions of phase space that should be relatively free of signal and dominated by events arising from the $W(\to \ell\nu)$ +jets process. By comparing the expected yield of $W(\to \ell\nu)$ +jets from simulation in this region with the observed yield from data, a correction factor or scale factor (f_W) can be calculated and used to normalise the yield of these events in the signal regions.

For each signal region, (defined in Sections 7.5.1 and 7.6.1) the different kinematic selection requirements on, for example, the electron and muon transverse momenta, will cause differences in the expected ratios of jets originating from quarks or gluons in $W(\to \ell\nu)$ +jets events. Since jets originating from quarks are more likely to be mis-identified as a τ_h than jets originating from gluons, it is necessary to derive a separate f_W scale factor for each signal region. It is also necessary to derive separate f_W for events with the reconstructed charge product $q_\ell \times q_{\tau_h} < 0$ [opposite sign (OS) events] and $q_\ell \times q_{\tau_h} > 0$ [same sign (SS) events], since the fraction of mis-identified τ_h initiated by a quark or gluon is also different in these two regions.

7.7.2 $W(\to \ell\nu)$ +jets in the $H \to \tau\tau$ analysis

The transverse mass is defined in Equation 7.11 and shown for four signal regions in Figure 7.7. Since the $H \to \tau\tau$ signal peaks at $m_T \sim 0$ GeV, a $W(\to \ell\nu)$ +jets control region is constructed around the W mass end-point ($m_T \sim m_W$). For each signal region, f_W is defined as

$$f_W = \frac{N_{data}(70 < m_T < 110 \text{ GeV}) - N^{MC}_{non-W+\text{jets}}(70 < m_T < 110 \text{ GeV})}{N^{MC}_{W+\text{jets}}(70 < m_T < 110 \text{ GeV})}, \tag{7.14}$$

where $N_{data}(70 < m_T < 110 \text{ GeV})$, $N^{MC}_{non-W+\text{jets}}(70 < m_T < 110 \text{ GeV})$ and $N^{MC}_{W+\text{jets}}(70 < m_T < 110 \text{ GeV})$ are the yield of data, non W+jets events and W+jets events passing all the $H \to \tau\tau$ selection criteria in a given region passing except for the transverse mass which is required to lie between 70 GeV and 110 GeV and second and third quantities are derived

SS/OS region	Final state particles	Signal Region	f_W
OS	$e-\tau$	low p_T^{miss}, 0 jet	0.533 ± 0.070
		high p_T^{miss}, 0 jet	0.582 ± 0.006
		≥ 1 jet	0.610 ± 0.011
	$\mu-\tau$	low p_T^{miss}, 0 jet	0.431 ± 0.064
		high p_T^{miss}, 0 jet	0.541 ± 0.005
		≥ 1 jet	0.604 ± 0.010
	$\ell-\tau$	VBF cuts	1.000 ± 0.123
SS	$e-\tau$	low p_T^{miss}, 0 jet	1.562 ± 0.483
		high p_T^{miss}, 0 jet	0.850 ± 0.017
		≥ 1 jet	0.781 ± 0.024
	$\mu-\tau$	low p_T^{miss}, 0 jet	0.793 ± 0.256
		high p_T^{miss}, 0 jet	0.697 ± 0.012
		≥ 1 jet	0.726 ± 0.021
	$\ell-\tau$	VBF cuts	1.336 ± 0.263

Table 7.3: Exclusive $W(\to \ell\nu)$ +jets scale factors for events with an oppositely charged $\ell - \tau_h$ pair (OS) and a like charged $\ell - \tau_h$ pair (SS) in each signal region of the $H \to \tau\tau$ analysis.

from MC. These scale factors are shown in Table 7.3. The differences in f_W are due to the different fraction of jets originating from a quark or gluon in the sample due to

- the different light-lepton p_{T} cut;
- the different jet multiplicities;
- the OS and SS regions quark fraction difference.

Due to the large statistical uncertainties on the scale factors for the low $p_{\mathrm{T}}^{\mathrm{miss}}$, 0-jet region, the high $p_{\mathrm{T}}^{\mathrm{miss}}$, 0-jet region scale factors are used to normalise the yield of W+jets in all 0-jet regions.

7.7.3 $W(\to \ell\nu)$ +jets in the $H \to WW^* \to \ell\nu\tau\nu$ analysis

For the $H \to WW^* \to \ell\nu\tau\nu$ analysis, there is no such separation of signal and $W(\to \ell\nu)$ +jets events in the transverse mass distributions. This motivates a different definition of the $W(\to \ell\nu)$ +jets control region for the $H \to WW^* \to \ell\nu\tau\nu$ analysis.

The signal and background $\Delta\phi(\ell, \tau_h)$ distribution of events passing the object selection cuts and with $m_T > 30$ GeV are shown in Figure 7.15. In order to construct a control region that is suitably pure, events are first required to have $\Delta\phi(\ell, \tau_h) < 2$. This selects a region dominated by $W(\to \ell\nu)$ +jets production though this is also where the signal peaks. To isolate the $W(\to \ell\nu)$ +jets from the signal, a cut is placed on the event's $\Delta R(\ell, \tau_h)$ (Figure 7.13). The region $\Delta R(\ell, \tau_h) > 1.15$ and $\Delta\phi(\ell, \tau_h) < 2$ is used as a control region to obtain a $W(\to \ell\nu)$ +jets scale factor:

$$f_W = \frac{N_{data}(\Delta R(\ell, \tau_h) > 1.15) - N^{MC}_{non-W+\text{jets}}(\Delta R(\ell, \tau_h) > 1.15)}{N^{MC}_{W+\text{jets}}(\Delta R(\ell, \tau_h) > 1.15)}, \tag{7.15}$$

where $N_{data}(\Delta R(\ell, \tau_h) > 1.15)$, $N^{MC}_{non-W+\text{jets}}(\Delta R(\ell, \tau_h) > 1.15)$ and $N^{MC}_{W+\text{jets}}(\Delta R(\ell, \tau_h) > 1.15)$ are the yield of data, non $W(\to \ell\nu)$ +jets events and $W(\to \ell\nu)$ +jets events after all other $H \to WW^* \to \ell\nu\tau\nu$ selection criteria. The f_W scale factors for $e - \tau_h$ and $\mu - \tau_h$ events are shown in Table 7.4.

Region	Channel	f_W
OS	$e - \tau_h$	0.618 ± 0.014
	$\mu - \tau_h$	0.555 ± 0.008
SS	$e - \tau_h$	1.01 ± 0.10
	$\mu - \tau_h$	0.71 ± 0.08

Table 7.4: Exclusive $W(\to \ell\nu)$ +jets scale factors for events with an oppositely charged $\ell - \tau_h$ pair (OS) and a similarly charged $\ell - \tau_h$ pair (SS) in each signal region of the $H \to WW^* \to \ell\nu\tau\nu$ analysis.

7.7.4 Systematic uncertainty on the $W(\to \ell\nu)$ +jets data-driven background method

A systematic uncertainty on the yield of $W(\to \ell\nu)$ +jets in each signal region of the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ analysis due to this method is defined by varying the f_W for each region by its statistical uncertainty. These are shown in Tables 7.3 and 7.4 for the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ analyses, respectively.

7.8 Data-driven estimate of the QCD multi-jet background

7.8.1 Introduction

Simulated event samples cannot be used to model multi-jet backgrounds from QCD processes with several outgoing partons due to the very large multi-jet process cross sections. Instead, the data are used to estimate the contribution of QCD events using a so-called 'ABCD' method.

In this method, the charge product of the light lepton and the τ_h, $q_{\tau_h} \times q_\ell$, and the isolation requirements on the light lepton are used to split events into four regions as shown in Table 7.5. The signal region (A) has an isolated, light lepton and $q_{\tau_h} \times q_\ell < 0$. The control regions (B, C and D) are expected to contain a negligible number of signal events and hence are modelled as containing events originating only from QCD and other background processes.

	Light lepton isolation	
	Isolated	Not isolated
$q_\ell \times q_{\tau_h} < 0$ (OS)	A (Signal)	C
$q_\ell \times q_{\tau_h} > 0$ (SS)	B	D

Table 7.5: Construction of QCD control regions.

To estimate the yield of QCD events in each signal region, first the expected number of events from other background processes (Z+jets, W+jets, diboson and $t\bar{t}$ and single top) is calculated for regions B, C and D. For each region, this expected yield is subtracted from the observed number of data events to obtain the number of QCD events in each region, n_C and n_D respectively. A ratio is then taken to find the OS/SS ratio of non-isolated QCD events, n_C/n_D. To model the yield distribution shapes of QCD events in region A, region B is used, after normalising to the OS/SS ratio n_C/n_D.

7.8.2 QCD background in the $H \to \tau\tau$ analysis

The n_C/n_D ratio defined in Section 7.8 for each $H \to \tau\tau$ signal region is shown in Table 7.6. Figure 7.16 shows that these ratios have no significant dependence on the calorimeter isolation requirement for all events passing the tau selection requirements. As with the f_W determination in Section 7.7.2, the variations present are due to the different quark fraction of jets in each region.

Final State	Signal Region	Non-isolated OS/SS ratio
$e-\tau$	0 jet, low p_T^{miss}	1.070 ± 0.022
	0 jet, high p_T^{miss}	1.069 ± 0.039
	≥ 1 jet	1.025 ± 0.028
$\mu-\tau$	0 jet, low p_T^{miss}	1.154 ± 0.009
	0 jet, high p_T^{miss}	1.154 ± 0.018
	≥ 1 jet	1.141 ± 0.011
$\ell-\tau$	VBF cuts	1.209 ± 0.055

Table 7.6: QCD n_C/n_D ratios in each $H \to \tau\tau$ signal region.

7.8.3 QCD background in the $H \to WW^* \to \ell\nu\tau\nu$ analysis

The multi-jet background yield and shape are estimated in the same way for the $H \to WW^* \to \ell\nu\tau\nu$ signal regions as for the $H \to \tau\tau$ signal regions, with the n_C/n_D ratios shown in Table 7.7. After the $\Delta\phi(\ell, \tau_h)$ requirement, there are very few events expected from QCD processes, since in a di-jet event two highest p_T jets are expected to be back-to-back in the transverse plane.

Signal Region	Non-isolated OS/SS ratio
$e-\tau$	1.80 ± 0.44
$\mu-\tau$	1.41 ± 0.10

Table 7.7: QCD n_C/n_D ratios in each $H \to WW^* \to \ell\nu\tau\nu$ signal region.

7.8.4 Systematic uncertainty on the QCD multi-jet data-driven background method

A systematic uncertainty on the yield of QCD multi-jet events in each signal region of the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ analysis from this method is defined by varying the non-isolated SS/OS ratio (n_C/n_D) in each region by its statistical uncertainty. These uncertainties for each signal region of the $H \to \tau\tau$ analysis are summarised in Table 7.6, while those calculated for the $H \to WW^* \to \ell\nu\tau\nu$ analysis signal regions are shown in Table 7.7.

An additional uncertainty could have been defined by using MC samples with a different parton showering model to determine the yield of non-QCD events in the control regions n_B, n_C and n_D and taking the difference in the expected yield as a systematic uncertainty. Since the non-QCD yield of events in these regions is very small, the effect on the expected QCD multi-jet yield in the signal region is assumed to be negligible.

7.9 Systematic uncertainties

Due to the imperfect modelling of both the signal and background physics processes and the simulation of the ATLAS detector response, small data-based corrections are applied and systematic uncertainties are determined.

7.9.1 Theory uncertainties

In both analyses, the expected number of signal events are derived from simulated samples and scaled according to their next-to-next-to-leading-order (NNLO) cross sections listed in Appendix A.1 and the integrated luminosity of the collected data. All background samples (except for $W(\to \ell\nu)$ +jets and QCD multi-jet production) are scaled by the NLO cross sections listed in Appendix A.1 and the integrated luminosity.

In the NNLO Higgs production and NLO background calculations, the cross section for each simulated process (Appendix A.1) was re-calculated with the factorisation and renormalisation scales halved and doubled simultaneously, with the maximum variation of the resulting cross section in either direction taken as a systematic uncertainty. The parton density function systematic uncertainty was assigned by taking the difference in cross section when using a different PDF set [75]. Since the change in the visible mass and MMC mass distribution shapes from these variations is negligibly small, only the overall normalisation is treated as a source of systematic uncertainty. Since the $W(\rightarrow \ell\nu)$ +jets yield is normalised to data there is no need to account for an uncertainty on the overall cross-section. The dominant theory uncertainties on the main backgrounds are summarised in Table 7.8. The theory uncertainties on the signal processes as a function of m_H are shown in Appendix A.1; the gluon-gluon fusion process uncertainties sum in quadrature to be $\simeq$ 18% while the vector-boson-fusion process uncertainties sum in quadrature to be $\simeq$ 4% [14].

Process	QCD scale (%)	PDF variation (%)	q^2 scale (%)
$Z \rightarrow \ell\ell/\tau\tau$	1.0	2.0	12.5
$t\bar{t}$	1.0	8.0	3.0

Table 7.8: Dominant theory systematic uncertainties [76] on the primary backgrounds in both analyses.

The systematic uncertainty on the integrated luminosity of the data samples is estimated to be 3.9% by comparing the integrated luminosity calculated from various techniques, as described in reference [72].

7.9.2 Trigger efficiencies

The measured trigger efficiency scale factors is defined to be the observed difference between the MC expectation of the trigger efficency and data. These were calculated using $Z \rightarrow \ell\ell$ data with tag-and-probe methods. The electron and muon triggered events are shown in Table 7.9.

Trigger	Scale Factor (%)
Muon ($p_{\mathrm{T}} > 18$ GeV)	99.2 ± 0.5
Electron ($p_{\mathrm{T}} > 20(22)$ GeV)	99.5 ± 1.0

Table 7.9: Trigger scale factors and systematic uncertainties.

7.9.3 Electron candidates

For electrons, systematic uncertainties on the following are considered: energy scale and resolution, reconstruction and identification efficiency, and calorimeter isolation efficiency [74].

Energy clusters in the electromagnetic calorimeter identified as resulting from primary electrons have an energy scale uncertainty of about 1% (3%) in the barrel (endcap) regions, as determined by studies that calibrate data samples using resonances that decay to electrons, such as $Z \to e^+e^-$ and an estimate of energy deposited in the upstream material.

The electron identification efficiency and its associated systematic uncertainty are evaluated using tag-and-probe methods using electrons produced in the decay of J/ψ and Z resonances, and the uncertainty is found to be about 1.5%, varying with the candidate E_{T} and η. The uncertainty on the $e - \tau_h$ event selection is obtained by varying the identification efficiency by this amount.

The uncertainty on the electron calorimeter isolation scale factor (2%) is found using $Z \to e^+e^-$ events in which a comparison of the electron isolation efficiency for data and simulation is made. The effect of the incertainty is determined by varying the isolation scale factor by its uncertainty.

7.9.4 Muon candidates

Systematic uncertainties on the muon p_{T} resolution and the identification efficiency scale factor are applied by varying these quantities by their uncertainties [73]. The muon p_{T} resolution uncertainty is obtained by calibrating the simulated events to data using the measured Z boson peak width, while the identification efficiency is measured using a tag-and-probe analysis of muons produced in $Z \to \mu^+\mu^-$ events in data.

7.9.5 τ_h candidates

The τ_h energy scale systematic uncertainty [58] is based on studies of simulated $Z \to \tau\tau$ events. In these events the relative difference of the reconstructed τ_h p_{T} and that of the hadronic tau decay in the Monte Carlo truth record is studied using reconstructed, identified tau candidates within a cone of radius $\Delta R = 0.2$ from a hadronic tau decay in the MC truth record. The sources of uncertainty considered include variations in the hadronic shower model used, energy cluster noise thresholds, and the amount of additional dead material traversed before reaching the calorimeters in the simulated data.

A systematic uncertainty is also assigned to a scale factor for the mis-identification rate of an electron as a tau candidate [58], measured using $Z \to e^+e^-$ data events with a reconstructed $m_{e^+e^-}$ within a narrow window around m_Z. This scale factor is only used for simulated events in which the reconstructed τ_h candidate is matched to a simulated primary electron with $p_{\mathrm{T}} > 5$ GeV within a cone of radius $\Delta R = 0.2$.

7.9.6 Jet candidates

Systematic uncertainties arising from the estimation of the jet energy scale in [77, 78] are taken into account. The jet energy scale is measured at the electromagnetic (EM) scale and is calibrated using energy deposits in the calorimeters from electromagnetic showers. This energy is established using test-beam measurements for electrons in the barrel and endcap calorimeters. Corrections to this are applied to account for the hadronic calorimeter response as well as energy loss in dead material derived from Monte Carlo simulated events that restore the calorimeter response of the reconstructed jet to the simulated jet response.

7.9.7 $p_{\mathrm{T}}^{\mathrm{miss}}$ systematic uncertainties

Since both analyses have a cut on the $p_{\mathrm{T}}^{\mathrm{miss}}$ of the event, the energy scale variation of the jets, electrons and τ_h are propagated to the $p_{\mathrm{T}}^{\mathrm{miss}}$ vector and the effect on the acceptance of this cut is evaluated [79]. The energy scale variations of the clusters associated with a selected jet or τ_h candidate are treated as fully correlated. The effect of energy scale uncertainty on the acceptance is also evaluated for energy clusters not associated with a reconstructed object.

7.9.8 Data driven background estimation systematic uncertainties

The systematic uncertainties associated with these methods are described in Sections 7.7.4 and 7.8.4.

7.9.9 Systematic uncertainties in the $H \to \tau\tau$ analysis

The systematic uncertainties arising from the object selection are shown in Tables 7.10 and 7.11 for the $H \to \tau\tau$ signal regions for events with no additional selected jets and ≥ 1 additional selected jet in the final state, respectively. The $W(\to \ell\nu)$ +jets normalisation is taken from the control region, but uncertainties on the $W(\to \ell\nu)$ +jets yield account for any differences in the object selection efficiencies, energy scales or resolutions between the W control regions and the signal regions since the scale factors f_W are normalised to data.

7.9.10 Systematic uncertainties in the $H \to WW^* \to \ell\nu\tau\nu$ analysis

The systematic uncertainties arising from the object selection are shown in Table 7.12 for the $H \to WW^* \to \ell\nu\tau\nu$ signal regions.

7.10 Limit Setting

Since the number of observed data events in the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ analysis channels shown in Tables 7.1 and 7.2 are consistent with a background only hypothesis, no evidence of Higgs boson production is observed (though it is not expected that either analysis

0 jet ($p_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV)	Energy Scale		Electrons		Muons		τ_h		$p_{\mathrm{T}}^{\mathrm{miss}}$	
	Jet/τ_h	Electron	Res.	Id.	Res.	Id.	Id.	e-FR	Cluster	Pile-up
ggF	−4.3 +3.0	+0.0 +0.0	+0.2 +0.1	+1.1 −1.4	+0.1 −0.1	+1.2 −1.4	+4.1 −4.1	+0.0 −0.0	−9.6 +8.4	−6.0 +4.7
VBF	−24.4 +15.0	+0.1 −0.0	+1.0 −0.0	+1.2 −1.4	+0.0 −0.0	+1.2 −1.3	+4.2 −4.2	+0.0 −0.0	−19.5 +12.3	−9.8 +1.0
$Z \to \tau\tau$	+3.2 −2.9	+0.3 −0.4	−0.1 −0.1	+1.0 −1.2	+0.0 −0.1	+1.3 −1.5	+4.3 −4.3	+0.0 −0.0	−0.0 −0.0	−0.0 −0.0
$W(\to \ell\nu)$ +jets + jets	+1.3 −5.1	+0.0 −0.3	+0.0 −1.1	+0.0 −0.0	+0.1 −0.2	+0.1 −0.1	−0.7 +0.7	−0.5 +0.5	+3.4 −3.2	+1.4 −2.9
$Z \to \ell\ell$	+1.5 −2.4	−0.2 +0.4	+0.1 +0.3	+1.6 −2.0	+0.4 −0.5	+0.9 −1.0	+4.4 −4.4	+22.6 −22.6	−8.6 +8.5	−5.0 +5.2
(single) top	−20.3 +10.3	+0.0 −0.8	+0.0 −1.3	+1.3 −1.6	+4.7 −5.7	+1.1 −1.2	+4.6 −4.6	+0.0 −0.0	+12.7 −0.0	+2.7 −0.0
Diboson	−2.7 +0.6	+0.2 +0.2	+0.8 −0.0	+1.3 −1.5	+0.1 −0.2	+1.1 −1.3	+4.4 −4.4	+1.6 −1.6	+0.0 −0.9	+0.0 −0.4
0 jet (high $p_{\mathrm{T}}^{\mathrm{miss}}$)	**Energy Scale**		**Electrons**		**Muons**		**τ_h**		**$p_{\mathrm{T}}^{\mathrm{miss}}$**	
	Jet/τ_h	Electron	Res.	Id.	Res.	Id.	Id.	e-FR	Cluster	Pile-up
ggF	+1.6 −0.0	+0.3 −0.4	−0.3 −0.2	+1.2 −1.5	+0.1 −0.1	+1.2 −1.3	+4.0 −4.0	+0.0 −0.0	+6.2 −5.0	+4.5 −2.0
VBF	−12.4 +18.0	+0.0 −0.8	−0.9 +0.1	+1.1 −1.4	+0.0 −0.0	+1.2 −1.4	+4.1 −4.1	+0.0 −0.0	+6.6 −1.9	+5.1 −0.0
$Z \to \tau\tau$	+8.4 −8.2	+0.0 −0.2	−0.1 −0.1	+1.0 −1.2	+0.1 −0.1	+1.3 −1.5	+4.1 −4.1	+0.0 −0.0	−0.0 −0.0	−0.0 +0.0
$W(\to \ell\nu)$ +jets + jets	+0.0 −6.3	+0.0 −0.7	+0.0 −1.8	+0.1 −0.1	+1.2 −1.5	+0.1 −0.2	−1.0 +1.0	−0.5 +0.5	+6.2 −17.8	+4.4 −11.7
$Z \to \ell\ell$	+1.0 −5.8	−0.0 −1.0	+0.0 −0.3	+1.8 −2.2	+0.6 −0.9	+0.8 −0.9	+4.3 −4.3	+29.0 −29.0	+19.9 −16.5	+11.5 −8.7
(single) top	−21.7 +21.6	+0.0 −2.6	+0.0 −3.1	+1.2 −1.5	+3.2 −4.1	+1.2 −1.3	+4.1 −4.1	+0.0 −0.0	+2.8 −0.0	+5.1 −0.0
Diboson	+0.0 −2.9	−0.7 +0.8	+0.3 +0.4	+1.2 −1.5	+0.2 −0.2	+1.1 −1.2	+4.2 −4.2	+0.8 −0.8	+6.5 −6.9	+0.0 −5.3

Table 7.10: Uncertainties (in %) on the number of selected events for the different simulated background processes and for a signal sample with a simulated $m_H = 120$ GeV for $H \to \tau\tau$ signal regions with no additional reconstructed jets.

is sensitive to SM Higgs boson production with this amount of integrated luminosity). An upper limit is placed on the excluded cross section of SM like Higgs boson production, as a function of the Higgs mass using the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ signal regions separately. The limits are defined to have an exclusion at 95% confidence level and are calculated using the profile likelihood method [80]. The MMC mass and visible mass are used as the discriminating variables in the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ signal regions, respectively. Systematic uncertainties are included as nuisance parameters. Systematic uncertainties on the shape and normalisation of the MMC and visible mass distributions due to the variation of the jet and τ_h energy scales are included. Other uncertainties are described in Section 7.9 including the measurement of integrated luminosity, object energy scales, resolutions and the acceptance common to all samples and signal regions are treated

≥ 1 jet	Energy Scale		Electrons		Muons		τ_h		$p_{\mathrm{T}}^{\mathrm{miss}}$	
	Jet/τ_h	Electron	Res.	Id.	Res.	Id.	Id.	e-FR	Cluster	Pile-up
ggF	$^{+9.8}_{-10.0}$	$^{-0.3}_{+0.3}$	$^{+0.1}_{+0.1}$	$^{+1.2}_{-1.5}$	$^{+0.0}_{-0.0}$	$^{+1.1}_{-1.3}$	$^{+4.2}_{-4.2}$	$^{+0.0}_{-0.0}$	$^{-0.2}_{+1.3}$	$^{-0.1}_{+1.4}$
VBF	$^{+0.9}_{-2.3}$	$^{-0.1}_{+0.2}$	$^{+0.0}_{+0.0}$	$^{+1.3}_{-1.6}$	$^{+0.1}_{-0.1}$	$^{+1.1}_{-1.2}$	$^{+4.3}_{-4.3}$	$^{+0.0}_{-0.0}$	$^{-1.9}_{+0.5}$	$^{-0.9}_{+0.5}$
$Z \to \tau\tau$	$^{+5.0}_{-4.5}$	$^{+0.2}_{-0.1}$	$^{+0.1}_{+0.0}$	$^{+1.2}_{-1.5}$	$^{+0.0}_{-0.0}$	$^{+1.2}_{-1.3}$	$^{+4.4}_{-4.4}$	$^{+0.0}_{-0.0}$	$^{+0.0}_{-0.0}$	$^{-0.0}_{-0.0}$
$W \to \ell\nu$ + jets	$^{+0.0}_{-4.2}$	$^{-2.1}_{+1.1}$	$^{+0.1}_{-0.3}$	$^{+0.4}_{-0.5}$	$^{+1.3}_{-1.6}$	$^{+0.4}_{-0.5}$	$^{-1.7}_{+1.7}$	$^{-0.7}_{+0.7}$	$^{+0.0}_{-13.4}$	$^{+0.0}_{-13.0}$
$Z \to \ell\ell$	$^{+18.8}_{-14.1}$	$^{+0.0}_{-2.0}$	$^{-1.0}_{+0.1}$	$^{+2.3}_{-2.8}$	$^{+0.4}_{-0.5}$	$^{+0.4}_{-0.5}$	$^{+4.7}_{-4.7}$	$^{+14.1}_{-14.1}$	$^{+10.0}_{-7.7}$	$^{+5.8}_{-5.4}$
(single) top	$^{+4.2}_{-3.9}$	$^{+0.0}_{-0.0}$	$^{+0.0}_{-0.4}$	$^{+1.4}_{-1.8}$	$^{+0.0}_{-0.1}$	$^{+1.0}_{-1.1}$	$^{+4.4}_{-4.4}$	$^{+0.4}_{-0.4}$	$^{+0.6}_{-0.7}$	$^{+0.3}_{-0.6}$
Diboson	$^{+11.0}_{-4.8}$	$^{-0.1}_{+0.9}$	$^{+0.1}_{+0.3}$	$^{+1.4}_{-1.7}$	$^{+0.1}_{-0.1}$	$^{+1.0}_{-1.2}$	$^{+4.4}_{-4.4}$	$^{+0.7}_{-0.7}$	$^{+0.4}_{-1.6}$	$^{+0.0}_{-0.4}$
VBF signal region	**Energy Scale**		**Electrons**		**Muons**		τ_h		$p_{\mathrm{T}}^{\mathrm{miss}}$	
	Jet/τ_h	Electron	Res.	Id.	Res.	Id.	Id.	e-FR	Cluster	Pile-up
ggF	$^{+26.0}_{-19.7}$	$^{+0.0}_{-1.6}$	$^{+0.0}_{-0.0}$	$^{+1.9}_{-2.4}$	$^{+0.0}_{-0.0}$	$^{+0.6}_{-0.7}$	$^{+4.4}_{-4.4}$	$^{+0.0}_{-0.0}$	$^{+6.2}_{-0.0}$	$^{+6.2}_{-0.0}$
VBF	$^{+9.9}_{-9.2}$	$^{+0.0}_{-0.2}$	$^{-0.3}_{+0.1}$	$^{+1.3}_{-1.6}$	$^{+0.4}_{-0.5}$	$^{+1.1}_{-1.2}$	$^{+4.2}_{-4.2}$	$^{+0.0}_{-0.0}$	$^{-1.9}_{+0.0}$	$^{+0.0}_{-0.1}$
$Z \to \tau\tau$	$^{+5.2}_{-4.4}$	$^{+0.0}_{-0.0}$	$^{+0.0}_{-0.0}$	$^{+1.1}_{-1.3}$	$^{+1.3}_{-1.6}$	$^{+1.2}_{-1.4}$	$^{+4.3}_{-4.3}$	$^{+0.0}_{-0.0}$	$^{+0.0}_{-0.0}$	$^{+0.0}_{-0.0}$
$W \to \ell\nu$ + jets	$^{-39.4}_{+24.7}$	$^{+0.1}_{+0.1}$	$^{+0.0}_{+0.3}$	$^{+0.8}_{-0.9}$	$^{+4.8}_{-5.6}$	$^{+0.4}_{-0.5}$	$^{-2.6}_{+2.6}$	$^{-1.4}_{+1.4}$	$^{+0.0}_{-6.0}$	$^{+4.8}_{-3.2}$
$Z \to \ell\ell$	$^{+47.7}_{-7.6}$	$^{+0.0}_{-3.9}$	$^{+0.0}_{-3.9}$	$^{+2.6}_{-3.2}$	$^{+1.7}_{-2.1}$	$^{+0.2}_{-0.2}$	$^{+5.0}_{-5.0}$	$^{+2.0}_{-2.0}$	$^{+0.3}_{-0.9}$	$^{+0.0}_{-2.6}$
(single) top	$^{+2.2}_{-5.2}$	$^{+0.0}_{-1.8}$	$^{+0.0}_{-2.2}$	$^{+1.7}_{-2.0}$	$^{+1.8}_{-2.0}$	$^{+0.8}_{-0.9}$	$^{+4.5}_{-4.5}$	$^{+0.0}_{-0.0}$	$^{-2.0}_{+0.4}$	$^{+0.0}_{-0.9}$
Diboson	$^{+14.5}_{-14.2}$	$^{+0.2}_{-0.0}$	$^{+0.7}_{-0.5}$	$^{+1.1}_{-1.3}$	$^{+0.2}_{-0.2}$	$^{+1.2}_{-1.4}$	$^{+4.3}_{-4.3}$	$^{+0.6}_{-0.6}$	$^{+10.6}_{-0.0}$	$^{-3.7}_{+6.7}$

Table 7.11: Uncertainties (in %) on the number of selected events for the different simulated background processes and for a signal sample with a simulated $m_H = 120$ GeV for $H \to \tau\tau$ signal regions of events with ≥ 1 additional reconstructed jets failing the VBF selection and events passing the VBF selection.

as fully correlated and constrained using Gaussian functions.

The likelihood function for a particular bin of the MMC distribution in a given signal region is defined as:

$$\mathcal{L}(\mu, \vec{\beta}_{stat}, \vec{\theta}_s, \vec{\theta}_b, \vec{\theta}_{global}) = \mathrm{P}(n, \mu_T)\mathrm{P}(n_{stat}, \beta_{stat})\mathcal{L}(\vec{\theta}_s, \vec{\theta}_b, \vec{\theta}_{global}), \tag{7.16}$$

where:

Process	Energy Scale		Electrons		Muons		τ_h		$p_{\mathrm{T}}^{\mathrm{miss}}$	
	Jet/τ_{h}	Electron	Res.	Id.	Res.	Id.	Id.	e-FR	Cluster	Pile-up
ggF	$^{+0.5}_{-5.5}$	$^{+0.9}_{-0.6}$	$^{+0.4}_{0.0}$	$^{+0.8}_{-0.8}$	$^{+0.0}_{-0.1}$	$^{+0.2}_{-0.2}$	$^{+4.2}_{-4.2}$	$^{+1.7}_{-1.7}$	$^{+0.9}_{-0.4}$	$^{+0.8}_{-0.5}$
VBF	$^{+0.5}_{-5.5}$	$^{+0.9}_{-0.6}$	$^{+0.4}_{0.0}$	$^{+0.8}_{-0.8}$	$^{+0.1}_{-0.1}$	$^{+0.2}_{-0.2}$	$^{+4.2}_{-4.2}$	$^{+1.7}_{-1.7}$	$^{+0.9}_{-0.4}$	$^{+0.8}_{-0.5}$
$Z \to \tau\tau$	$^{+0.0}_{-0.6}$	$^{+1.9}_{-1.2}$	$^{+0.9}_{-0.2}$	$^{+0.8}_{-0.8}$	$^{+0.3}_{-0.4}$	$^{+0.3}_{-0.3}$	$^{+4.2}_{-4.2}$	$^{+0.1}_{-0.1}$	$^{+5.4}_{-4.1}$	$^{+3.7}_{-2.5}$
$W \to \ell\nu$ + jets	$^{+0.0}_{-6.2}$	$^{+0.5}_{-0.5}$	$^{+0.3}_{0.0}$	$^{+0.8}_{-0.8}$	$^{+0.0}_{-0.0}$	$^{+0.2}_{-0.2}$	$^{+4.2}_{-4.2}$	$^{+0.1}_{-0.1}$	$^{+0.5}_{-0.9}$	$^{+0.4}_{-0.5}$
$Z \to \ell\ell$ + jets	$^{+0.0}_{-4.3}$	$^{+1.3}_{-0.4}$	$^{+1.2}_{0.0}$	$^{+0.8}_{-0.8}$	$^{+13.2}_{-0.4}$	$^{+0.3}_{-0.3}$	$^{+4.2}_{-4.2}$	$^{+16.8}_{-16.8}$	$^{+3.0}_{-1.8}$	$^{+1.4}_{0.0}$
(Single) Top	$^{+0.0}_{-3.3}$	$^{+0.5}_{-0.2}$	$^{+0.2}_{0.0}$	$^{+0.9}_{-0.9}$	$^{+0.1}_{0.0}$	$^{+0.3}_{-0.3}$	$^{+4.2}_{-4.2}$	$^{+0.6}_{-0.6}$	$^{+0.2}_{-0.3}$	$^{+0.0}_{-0.2}$
Diboson	$^{+0.0}_{-1.6}$	$^{+0.3}_{-0.3}$	$^{+0.4}_{-0.0}$	$^{+0.8}_{-0.8}$	$^{+0.1}_{0.0}$	$^{+0.2}_{-0.2}$	$^{+4.2}_{-4.2}$	$^{+1.3}_{-1.3}$	$^{+0.6}_{-0.6}$	$^{+0.5}_{-0.0}$

Table 7.12: Uncertainties (in %) on the number of selected events for the different simulated background processes and for a signal sample with a simulated $m_H = 125$ GeV for $H \to WW^* \to \ell\nu\tau\nu$ signal regions passing the selection criteria described in Section 7.6.1.

- P(a, b) is a Poisson distribution

$$\mathrm{P(a,b)} = \frac{a^b \exp^{-b}}{a!}. \tag{7.17}$$

- n is the number of data events in a given bin.

- $\vec{\beta}_{stat}$ are the statistical uncertainties on the MC or data driven estimates.

- $\vec{\theta}_{s,b}$ are nuisance parameters (vectors of systematic uncertainties) for each signal and background process.

- $\vec{\theta}_{global}$ are the common systematic uncertainties that are fully correlated across all signal and background processes (e.g. the luminosity uncertainty).

- $\mathcal{L}(\vec{\theta}_s, \vec{\theta}_b, \vec{\theta}_{global})$ is a parameterisation of the nuisance parameters which are constrained using Gaussian functions

$$\mathcal{L}(\vec{\theta}_s, \vec{\theta}_b, \vec{\theta}_{global}) = \prod_{\theta = \vec{\theta}_s, \vec{\theta}_b, \vec{\theta}_{global}} \mathrm{Gaussian}(\theta | \mathrm{mean} = 0, \mathrm{sigma} = 1). \tag{7.18}$$

- μ_T is the number of expected events, as defined by

$$\mu_T = \sum_l \mu L \sigma_l(m_H) f_s(\vec{\theta}_s) + \sum_j L\beta_j f_b(\vec{\theta}_b), \tag{7.19}$$

  In the definition of μ_T,

  - L is the integrated luminosity,
  - μ is the scaling factor of the expected SM signal cross section (signal strength) and $\mu = 0$ (1) corresponds to the absence (presence) of a SM Higgs boson signal,
  - $\sigma_l(m_H)$ is the effective cross section for signal events in each SM Higgs production process, $l =$ ggF,VBF,WH and ZH,
  - β_j is the cross section for background process j,
  - and $f_{s,b}(\vec{\theta}_{s,b})$ is the dependence of the expected number of events on each nuisance parameter.

A test statistic is obtained from the profile-likelihood ratio using asymptotic formulae [80], as given by

$$\tilde{q}_\mu = \begin{cases} -2\ln\frac{\mathcal{L}(\mu,\hat{\hat{\vec{\theta}}}(\mu))}{\mathcal{L}(0,\hat{\hat{\vec{\theta}}}(0))} & \hat{\mu} < 0, \\ -2\ln\frac{\mathcal{L}(\mu,\hat{\hat{\vec{\theta}}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\vec{\theta}})} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu, \end{cases} \tag{7.20}$$

where:

- $\hat{\mu}$ is the maximum likelihood estimator of μ,
- $\hat{\vec{\theta}}$ represents the nuisance parameters evaluated at μ,
- $\hat{\hat{\vec{\theta}}}(\mu)$ represents the maximum likelihood estimators of $\vec{\theta}$ at a given μ.

Pseudo experiments are generated using the distributions of signal and background to obtain a PDF $f(\tilde{q}_\mu, \mu, \hat{\hat{\vec{\theta}}}(\mu))$ for a given signal strength, μ. Using this PDF, a p-value (the probability that a background-only experiment fluctuates more than the observation) for μ is obtained using:

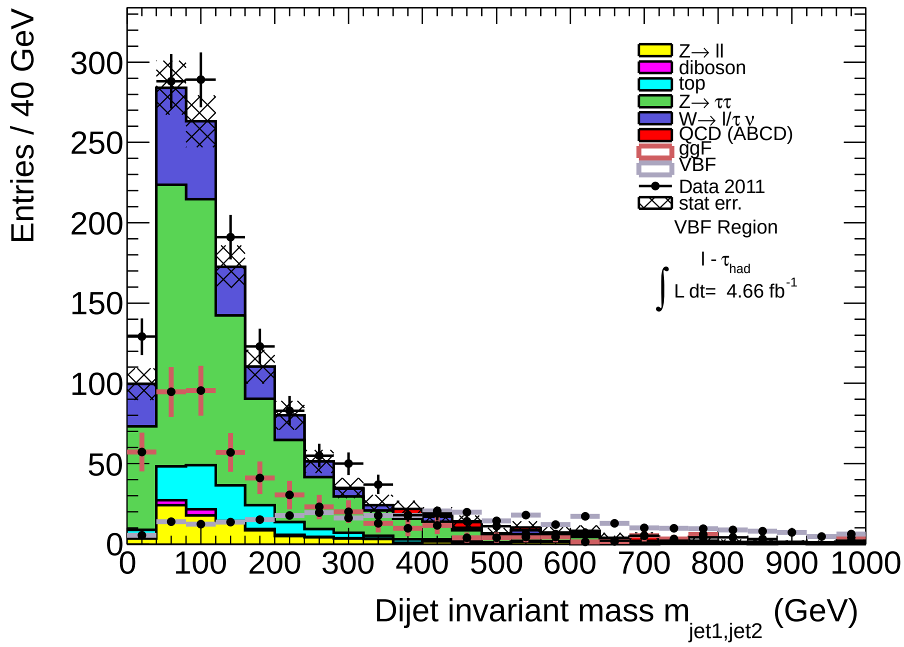
$$p_\mu = \int_{\tilde{q}_{\mu,obs}}^{\infty} f(\tilde{q}_\mu, \mu, \hat{\hat{\vec{\theta}}}(\mu)) d\tilde{q}_\mu. \tag{7.21}$$

An upper limit is set on the signal strength μ by iteratively evaluating this integral until $p_\mu = 0.05$. Similarly, pseudo experiments using background only distributions are used to evaluate an expected upper limit on μ, along with ± 1 and ± 2 standard deviation (SD) values at each assumed signal mass.

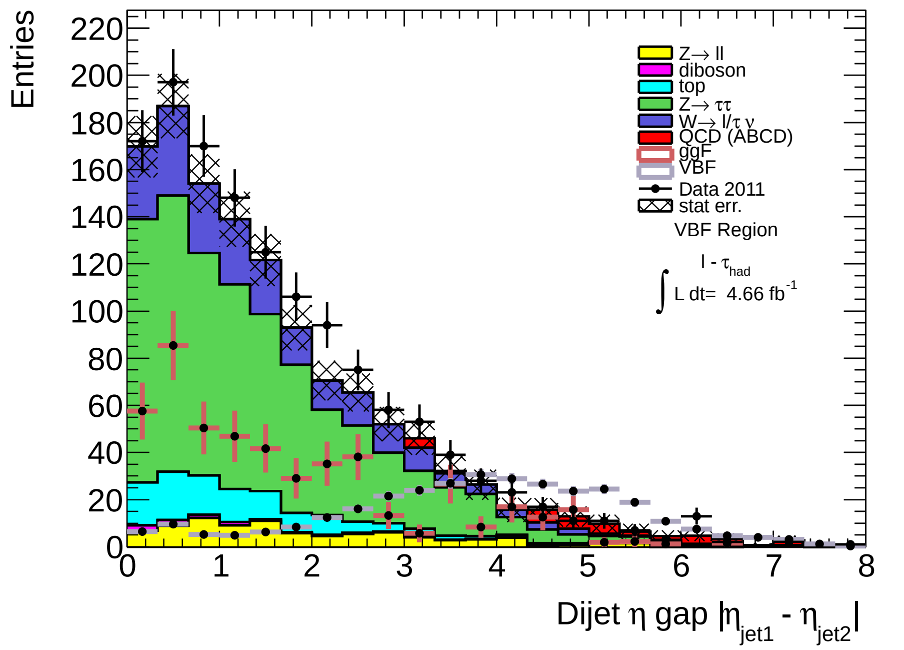
7.10.1 Limit

The expected and observed upper limits on μ as a function of assumed SM Higgs boson mass is shown in Figure 7.17 for the $H \to \tau\tau$ analysis. As m_H increases, the SM Higgs production cross section and $\mathcal{B}(H \to \tau\tau)$ both fall, leading to a higher expected upper limit at higher m_H. Conversely, the primary background process ($Z \to \tau\tau$) becomes less important as m_H increases due to the $\tau\tau$ mass resolution obtained using the MMC method. The shape of the expected upper limit as a function of Higgs mass is determined by these two effects.

The expected and observed upper limits on μ as a function of assumed SM Higgs boson mass is shown in Figure 7.18 for the $H \to WW^* \to \ell\nu\tau\nu$ analysis. As m_H increases, $\mathcal{B}(H \to WW^*)$ increases faster than the SM Higgs production cross section decreases, leading to a lower expected upper limit. The dominant uncertainty is due to the uncertainty on the primary background, W+jets production with a fake τ_h, due to the limited W+jets MC sample size and the uncertainty on the W+jets scale factor, f_W.

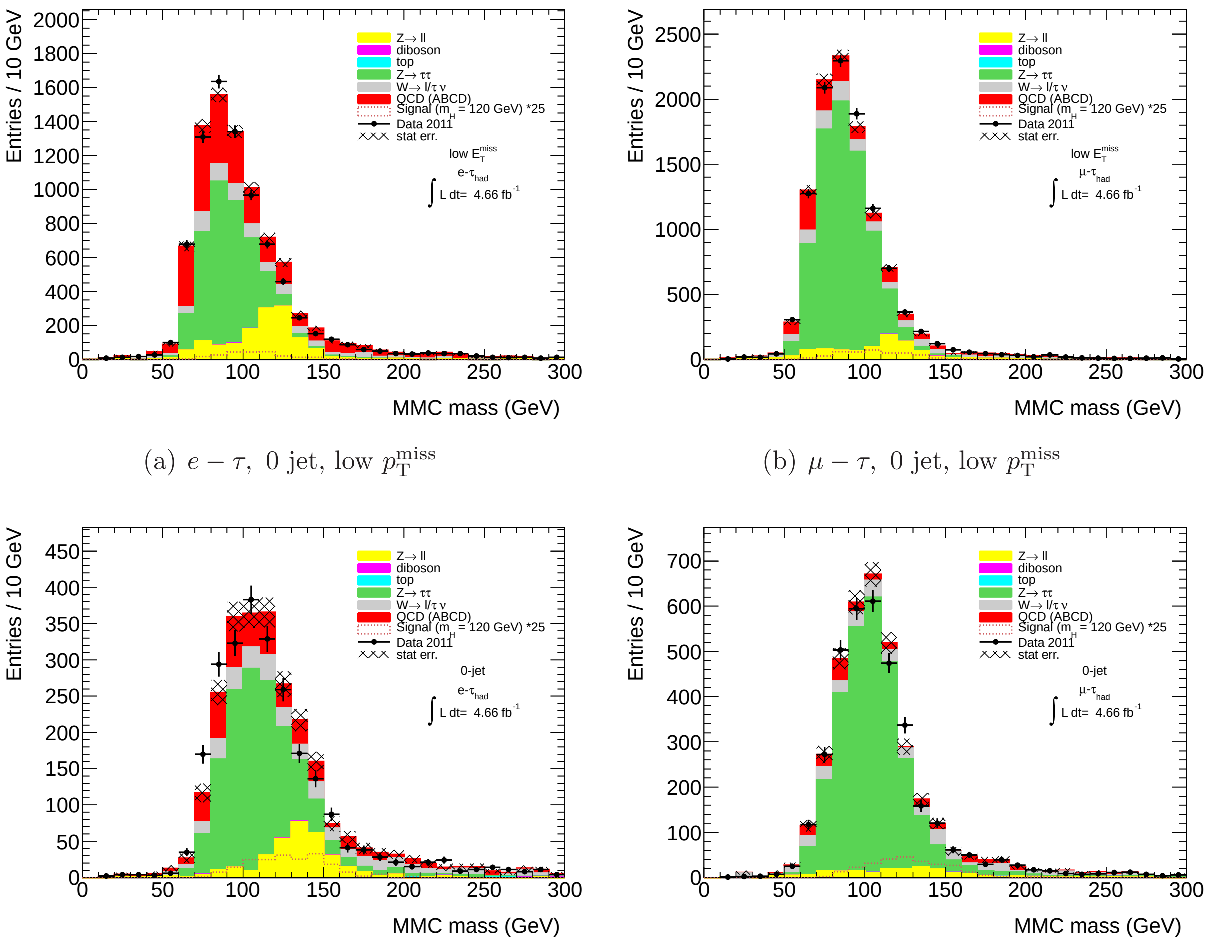


(a) $m_{\mathrm{jet1,jet2}}$



(b) $|\eta_{\mathrm{jet1}} - \eta_{\mathrm{jet2}}|$

Figure 7.9: Di-jet variable distributions used to further select VBF events passing all other selection cuts with two additional jets with $p_{\mathrm{T}} > 25$ GeV. The signal histograms from ggF and VBF SM Higgs production show the expected yield for 100 times the integrated luminosity of the analysis.



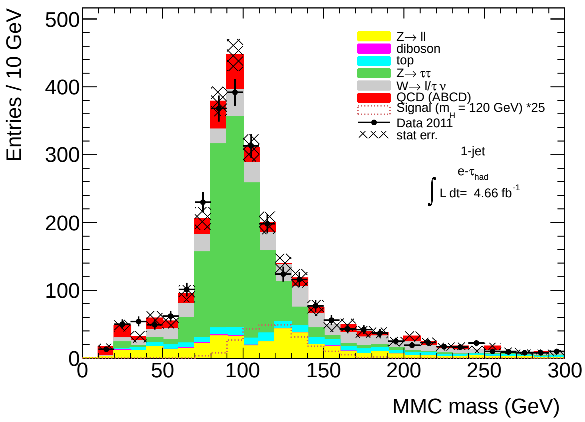
(a) $e-\tau$, 0 jet, low $p_{\mathrm{T}}^{\mathrm{miss}}$

(b) $\mu-\tau$, 0 jet, low $p_{\mathrm{T}}^{\mathrm{miss}}$

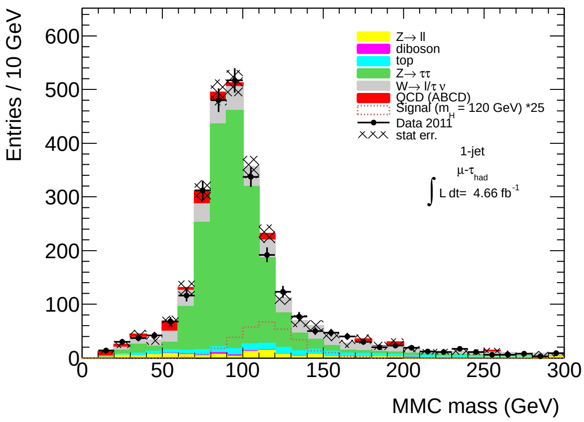
(c) $e-\tau$, 0 jet, high $p_{\mathrm{T}}^{\mathrm{miss}}$

(d) $\mu-\tau$, 0 jet, high $p_{\mathrm{T}}^{\mathrm{miss}}$

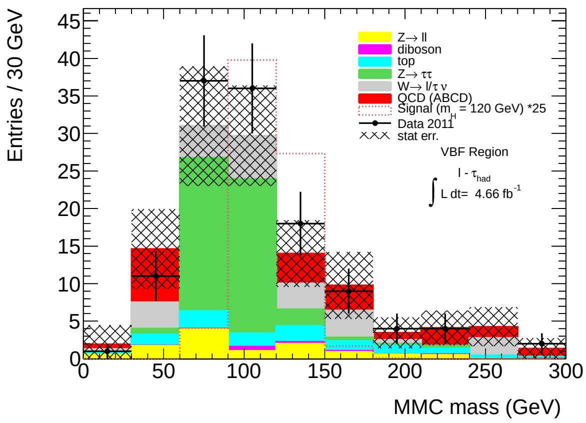
Figure 7.10: Comparison of the signal and background event distributions with the collision data for the $H \to \tau\tau$ signal regions with 0 additional jets. In each signal region, separate W+jets scale factors and QCD normalisations are used, as described in Sections 7.7 and 7.8. The black cross-hatched area indicates the statistical uncertainty on the sum of the backgrounds.



(a) $e-\tau$, ≥ 1 jet

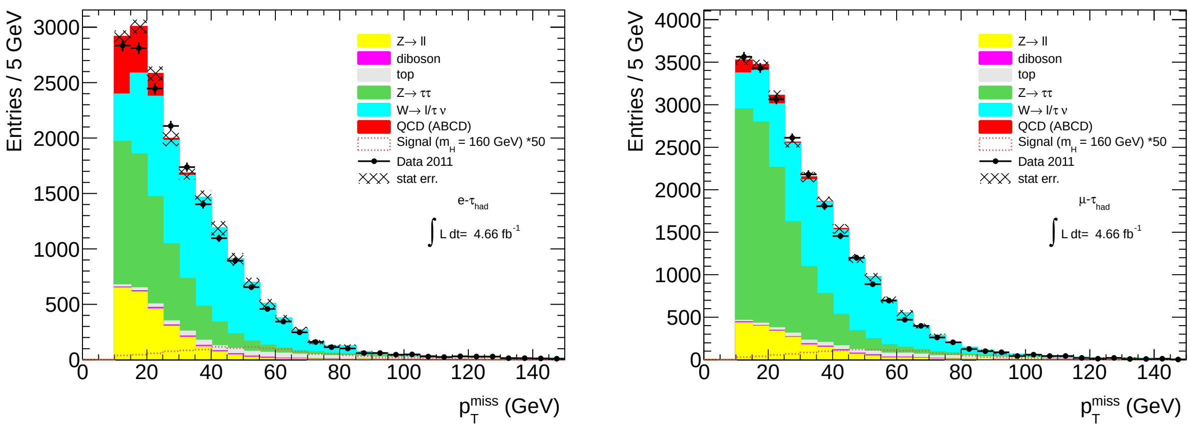


(b) $\mu-\tau$, ≥ 1 jet



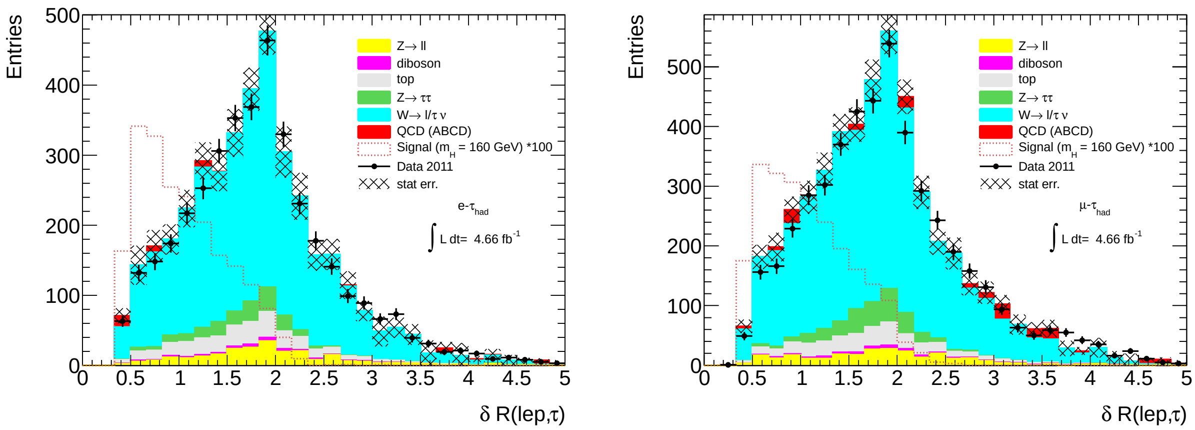
(c) $\ell-\tau$ VBF region

Figure 7.11: Comparison of the signal and background event distributions with the collision data for the $H \to \tau\tau$ signal regions with ≥ 1 additional jet. In each signal region, separate W+jets scale factors and QCD normalisation are used, as described in Sections 7.7 and 7.8. The black cross-hatched area indicates the statistical uncertainty on the sum of the backgrounds.



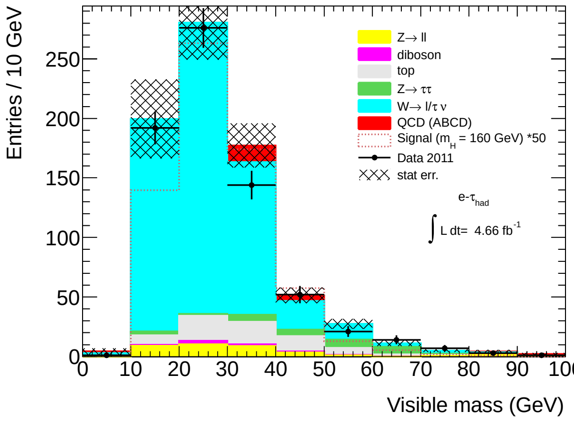
(a) $e-\tau$ (b) $\mu-\tau$

Figure 7.12: $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the $H \to WW^* \to \ell\nu\tau\nu$ signal and background distributions.

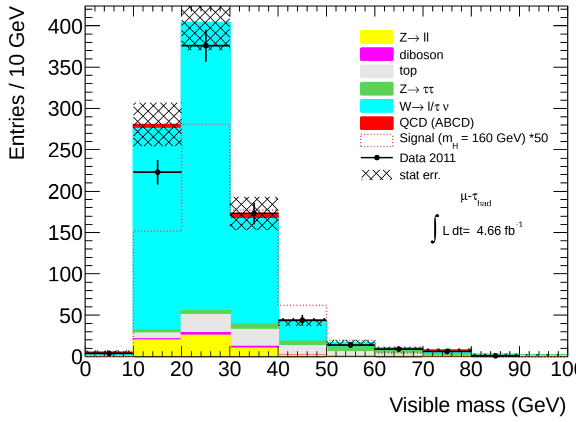


(a) $e-\tau$ events (b) $\mu-\tau$ events

Figure 7.13: The $\Delta R(\ell, \tau_h)$ distribution of $H \to WW^* \to \ell\nu\tau\nu$ signal, background and a comparison with data of events with $\Delta\phi(\ell, \tau_h) < 2.0$, $m_T > 30$ GeV and $p_{\mathrm{T}}^{\mathrm{miss}} > 10$ GeV after the f_W scale factor is applied. The QCD yield is calculated using the method described in Section 7.8.

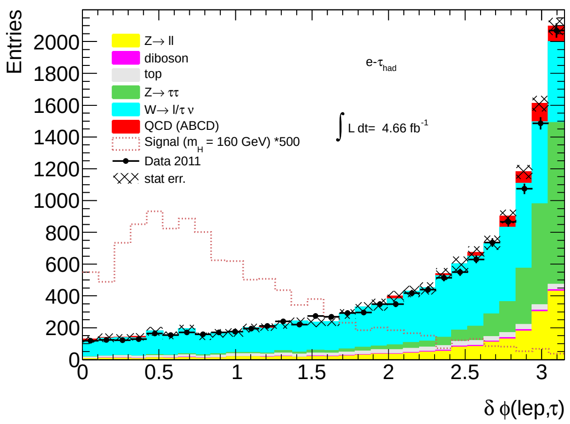


(a) $e - \tau$

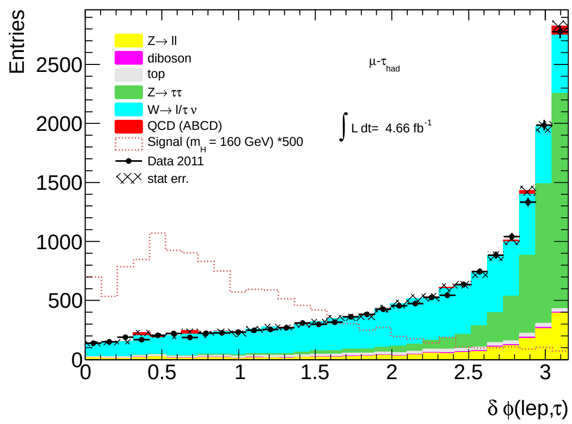


(b) $\mu - \tau$

Figure 7.14: Visible mass distribution of the $H \to WW^* \to \ell\nu\tau\nu$ signal and background after the full selection criteria have been applied.



(a) $e - \tau$ events



(b) $\mu - \tau$ events

Figure 7.15: The $\Delta\phi(\ell, \tau_h)$ distribution of $H \to WW^* \to \ell\nu\tau\nu$ signal, background and a comparison with data after the f_W scale factor is applied. The QCD yield is calculated using the method described in Section 7.8.

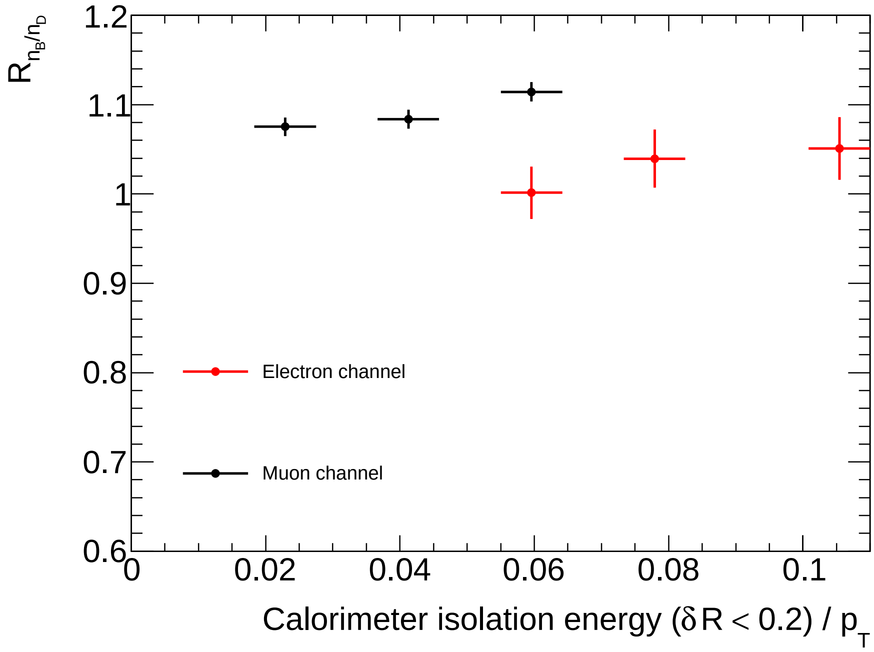


Figure 7.16: Effect on the non-isolated, opposite-sign to same-sign ratio of QCD events, n_C/n_D, under variation of the light lepton calorimeter isolation requirement.

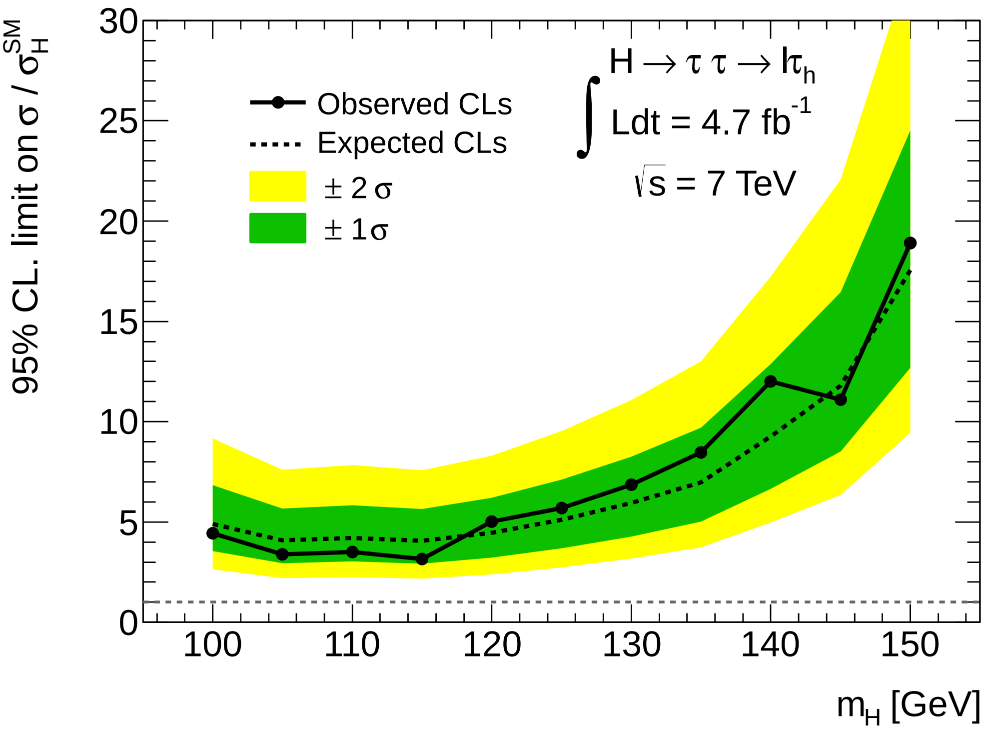


Figure 7.17: The expected and observed 95% confidence-level upper limit on the SM Higgs production cross section, normalised to the SM Higgs production cross section in the $H \to \tau\tau$ analysis. The yellow and green bands correspond to 1σ and 2σ variations due to statistical fluctuations at each mass point.

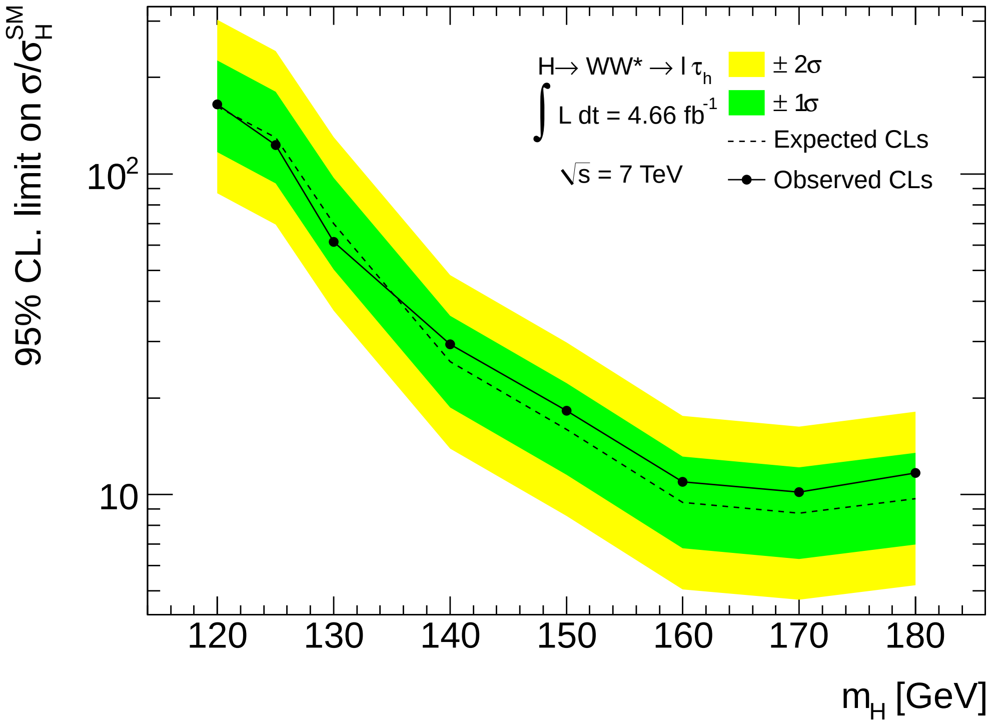


Figure 7.18: The expected and observed 95% confidence-level upper limit on the SM Higgs production cross section, normalised to the SM Higgs production cross section in the $H \to WW^* \to \ell\nu\tau\nu$ analysis. The yellow and green bands correspond to 1σ and 2σ variations due to statistical fluctuations at each mass point.

Chapter 8

Conclusions and outlook

In the Standard Model (SM) of particle physics, a mechanism is required to generate the observed masses of the SM fermions and electroweak gauge bosons. The discovery of Higgs boson production at the LHC would provide evidence for the so-called Higgs mechanism. A search for a light SM Higgs boson in the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ decay modes using 4.7 fb^{-1} of data collected at the ATLAS detector in 2011 has been presented. The number of events passing an event selection is consistent with total background estimation from other SM processes and an observed upper limit is placed on the SM Higgs boson production rate at 95% confidence level at around 5 times the SM cross section in the range $100 < m_H < 130$ GeV/c^2.

Global fits of indirect electroweak measurements combined with the results of previous direct searches favour a relatively light Higgs boson with a mass $114 < m_H < 157$ GeV/c^2. In this mass range, a search has been made using one of the most sensitive SM Higgs decay channels, the decay to tau leptons. This result was combined with other ATLAS SM Higgs searches [81] to exclude a SM Higgs boson with $110 < m_H < 117.5$ GeV/c^2, $118.5 < m_H < 122.5$ GeV/c^2 and $129 < m_H < 539$ GeV/c^2 at 95% confidence level while the range $120 < m_H < 555$ GeV/c^2 was expected to have been excluded.

With a larger data sample, there are several possible ways in which to improve the sensitivity of the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ analyses beyond the naive improvement due to increased statistics alone. By further specialisation of the signal regions, cuts can be designed to further improve the signal to background ratio. For example, it may be possible to design cuts to select the associated production modes $W/Z(\rightarrow \text{jet jet})H(\rightarrow \tau\tau)$ or events in which the Higgs boson is boosted in recoil against a jet. Another way would be to move to a multi-variate based discriminant such as a Boosted Decision Tree or a Neural Network in both analyses to exploit differences in the correlations of the signal and background event kinematics.

In 2012, it is expected that ~ 20 fb^{-1} will be collected at $\sqrt{s} = 8$ TeV by the ATLAS detector. With this increased sample size, it is expected that the combined results of the Higgs analyses proceeding at the ATLAS experiment should be able to either exclude at 95% confidence level or claim discovery of a light, SM Higgs boson.

Appendix A

$H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ analysis appendix

A.1 Simulated Datasets

The cross section at $\sqrt{s} = 7$ TeV used to normalise the expected yield of each process considered as background to the $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ analyses described in Chapter 7 are listed in Table A.1, Table A.2 and Table A.3 along with the generator used for that process.

The NNLO Higgs production cross section at $\sqrt{s} = 7$ TeV for the gluon gluon fusion (ggF) and weak boson fusion (VBF) simulated samples along with the branching ratios (BR) used for $H \to \tau\tau$ and $H \to WW^*$ are shown as a function of Higgs mass in Table A.4. The systematic uncertainties on the cross section under variation of α_s, PDF set and factorisation and renormalisation scales is shown in Table A.5.

Process	$\sigma \times \mathcal{BR} \times \mathrm{k-factor} \times \epsilon_{filter}$ (pb)	N_{events}	Generator
$Z \to \tau^+\tau^- \to \ell\tau_h$+Np0	835.5	10613180	`Alpgen`
$Z \to \tau^+\tau^- \to \ell\tau_h$+Np1	167.95	3334138	`Alpgen`
$Z \to \tau^+\tau^- \to \ell\tau_h$+Np2	50.45	804948	`Alpgen`
$Z \to \tau^+\tau^- \to \ell\tau_h$+Np3	14.06	509848	`Alpgen`
$Z \to \tau^+\tau^- \to \ell\tau_h$+Np4	3.49	145000	`Alpgen`
$Z \to \tau^+\tau^- \to \ell\tau_h$+Np5	0.96	45001	`Alpgen`
$Z/\gamma^* \to \tau^+\tau^- \to \ell\tau_h$+Np0 ($10 < m_{ll} < 40$)	3727.22	875000	`Alpgen`
$Z/\gamma^* \to \tau^+\tau^- \to \ell\tau_h$+Np1 ($10 < m_{ll} < 40$)	103.61	300000	`Alpgen`
$Z/\gamma^* \to \tau^+\tau^- \to \ell\tau_h$+Np2 ($10 < m_{ll} < 40$)	50.51	399000	`Alpgen`
$Z/\gamma^* \to \tau^+\tau^- \to \ell\tau_h$+Np3 ($10 < m_{ll} < 40$)	10.2	150001	`Alpgen`
$Z/\gamma^* \to \tau^+\tau^- \to \ell\tau_h$+Np4 ($10 < m_{ll} < 40$)	2.26	40000	`Alpgen`
$Z/\gamma^* \to \tau^+\tau^- \to \ell\tau_h$+Np5 ($10 < m_{ll} < 40$)	0.56	10001	`Alpgen`
$Z \to e^+e^-$+Np0	835.4	6218285	`Alpgen`
$Z \to e^+e^-$+Np1	167.95	1234998	`Alpgen`
$Z \to e^+e^-$+Np2	50.68	810000	`Alpgen`
$Z \to e^+e^-$+Np3	13.95	220001	`Alpgen`
$Z \to e^+e^-$+Np4	3.6	60001	`Alpgen`
$Z \to e^+e^-$+Np5	1.04	50001	`Alpgen`
$Z \to \mu^+\mu^-$+Np0	835.4	6615231	`Alpgen`
$Z \to \mu^+\mu^-$+Np1	167.68	1334297	`Alpgen`
$Z \to \mu^+\mu^-$+Np2	50.41	304948	`Alpgen`
$Z \to \mu^+\mu^-$+Np3	13.99	110001	`Alpgen`
$Z \to \mu^+\mu^-$+Np4	3.44	30001	`Alpgen`
$Z \to \mu^+\mu^-$+Np5	0.96	10001	`Alpgen`
$Z/\gamma^* \to e^+e^-$+Np0 ($10 < m_{ll} < 40$)	3727.1	994950	`Alpgen`
$Z/\gamma^* \to e^+e^-$+Np1 ($10 < m_{ll} < 40$)	103.6	299999	`Alpgen`
$Z/\gamma^* \to e^+e^-$+Np2 ($10 < m_{ll} < 40$)	50.51	799950	`Alpgen`
$Z/\gamma^* \to e^+e^-$+Np3 ($10 < m_{ll} < 40$)	10.22	149999	`Alpgen`
$Z/\gamma^* \to e^+e^-$+Np4 ($10 < m_{ll} < 40$)	2.26	40001	`Alpgen`
$Z/\gamma^* \to e^+e^-$+Np5 ($10 < m_{ll} < 40$)	0.56	10001	`Alpgen`
$Z/\gamma^* \to \mu^+\mu^-$+Np0 ($10 < m_{ll} < 40$)	3727.1	999850	`Alpgen`
$Z/\gamma^* \to \mu^+\mu^-$+Np1 ($10 < m_{ll} < 40$)	103.54	300001	`Alpgen`
$Z/\gamma^* \to \mu^+\mu^-$+Np2 ($10 < m_{ll} < 40$)	50.57	499998	`Alpgen`
$Z/\gamma^* \to \mu^+\mu^-$+Np3 ($10 < m_{ll} < 40$)	10.22	150001	`Alpgen`
$Z/\gamma^* \to \mu^+\mu^-$+Np4 ($10 < m_{ll} < 40$)	2.26	40000	`Alpgen`
$Z/\gamma^* \to \mu^+\mu^-$+Np5 ($10 < m_{ll} < 40$)	0.56	10001	`Alpgen`

Table A.1: Simulated Z/γ^* background datasets.

Process	$\sigma \times \mathcal{BR} \times \mathrm{k-factor} \times \epsilon_{filter}$ (pb)	N_{events}	Generator
$W \rightarrow e\nu$ +Np0	8245.2	3358886	`Alpgen`
$W \rightarrow e\nu$ +Np1	1551.6	2199646	`Alpgen`
$W \rightarrow e\nu$ +Np2	451.92	3768633	`Alpgen`
$W \rightarrow e\nu$ +Np3	121.56	908948	`Alpgen`
$W \rightarrow e\nu$ +Np4	30.3	250001	`Alpgen`
$W \rightarrow e\nu$ +Np5	8.55	70000	`Alpgen`
$W \rightarrow \mu\nu$ +Np0	8245.2	3462943	`Alpgen`
$W \rightarrow \mu\nu$ +Np1	1551.6	2348645	`Alpgen`
$W \rightarrow \mu\nu$ +Np2	451.92	3768738	`Alpgen`
$W \rightarrow \mu\nu$ +Np3	121.56	1008447	`Alpgen`
$W \rightarrow \mu\nu$ +Np4	30.3	254951	`Alpgen`
$W \rightarrow \mu\nu$ +Np5	8.55	70001	`Alpgen`
$W \rightarrow \tau\nu$ +Np0	8245.2	3358886	`Alpgen`
$W \rightarrow \tau\nu$ +Np1	1551.6	2249196	`Alpgen`
$W \rightarrow \tau\nu$ +Np2	451.92	3750987	`Alpgen`
$W \rightarrow \tau\nu$ +Np3	121.56	1009947	`Alpgen`
$W \rightarrow \tau\nu$ +Np4	30.3	249999	`Alpgen`
$W \rightarrow \tau\nu$ +Np5	8.55	65001	`Alpgen`

Table A.2: Simulated W background datasets.

Process	$\sigma \times \mathcal{BR} \times \mathrm{k-factor} \times \epsilon_{filter}$ (pb)	N_{events}	Generator
$W^-Z \to l\nu ll$	0.09	100001	MC@NLO
$W^-Z \to l\nu q\bar{q}$	0.92	25001	MC@NLO
$W^-Z \to l\nu\tau\tau$	0.04	25001	MC@NLO
$W^-Z \to q\bar{q}\prime ll$	0.27	100001	MC@NLO
$W^-Z \to q\bar{q}\prime\tau\tau$	0.26	25001	MC@NLO
$W^-Z \to \tau\nu ll$	0.04	25001	MC@NLO
$W^-Z \to \tau\nu\tau\tau$	0.02	25001	MC@NLO
$W^+Z \to l\nu ll$	0.16	25001	MC@NLO
$W^+Z \to l\nu q\bar{q}$	1.7	25001	MC@NLO
$W^+Z \to l\nu\tau\tau$	0.08	25001	MC@NLO
$W^+Z \to q\bar{q}\prime ll$	0.51	24951	MC@NLO
$W^+Z \to q\bar{q}\prime\tau\tau$	0.26	25001	MC@NLO
$W^+Z \to \tau\nu ll$	0.04	25001	MC@NLO
$W^+Z \to \tau\nu\tau\tau$	0.02	24951	MC@NLO
$ZZ \to 2l2\tau$	0.03	25001	MC@NLO
$ZZ \to 4\tau$	0.01	25001	MC@NLO
$ZZ \to llll$	0.03	50001	MC@NLO
$ZZ \to ll\nu\nu$	0.15	100000	MC@NLO
$ZZ \to llq\bar{q}$	0.53	25001	MC@NLO
$ZZ \to \tau\tau\nu\nu$	0.08	25001	MC@NLO
$ZZ \to \tau\tau q\bar{q}$	0.27	25001	MC@NLO
$q\bar{q}\prime \to W^+W^- \to e\nu e\nu$	0.52	199950	MC@NLO
$q\bar{q}\prime \to W^+W^- \to e\nu\mu\nu$	0.52	200001	MC@NLO
$q\bar{q}\prime \to W^+W^- \to e\nu\tau\nu$	0.52	200001	MC@NLO
$q\bar{q}\prime \to W^+W^- \to \mu\nu e\nu$	0.52	199950	MC@NLO
$q\bar{q}\prime \to W^+W^- \to \mu\nu\mu\nu$	0.52	199001	MC@NLO
$q\bar{q}\prime \to W^+W^- \to \mu\nu\tau\nu$	0.52	100001	MC@NLO
$q\bar{q}\prime \to W^+W^- \to \tau\nu e\nu$	0.52	199951	MC@NLO
$q\bar{q}\prime \to W^+W^- \to \tau\nu\mu\nu$	0.52	200001	MC@NLO
$q\bar{q}\prime \to W^+W^- \to \tau\nu\tau\nu$	0.52	199677	MC@NLO
$gg \to W^+W- \to e\nu e\nu$	0.02	10001	gg2WW
$gg \to W^+W- \to e\nu\mu\nu$	0.02	10001	gg2WW
$gg \to W^+W- \to e\nu\tau\nu$	0.01	10001	gg2WW
$gg \to W^+W- \to \mu\nu e\nu$	0.02	10001	gg2WW
$gg \to W^+W- \to \mu\nu\mu\nu$	0.02	10000	gg2WW
$gg \to W^+W- \to \mu\nu\tau\nu$	0.01	10001	gg2WW
$gg \to W^+W- \to \tau\nu e\nu$	0.01	10001	gg2WW
$gg \to W^+W- \to \tau\nu\mu\nu$	0.01	10001	gg2WW
$gg \to W^+W- \to \tau\nu\tau\nu$	0.01	10001	gg2WW
$t\bar{t}$ (No fully hadronic decays)	91.34	7146746	MC@NLO
$t\bar{t}$ (fully hadronic decays)	73.23	1199035	MC@NLO
single top : s − channel $W \to e\nu$	0.5	99901	AcerMC
single top : s − channel $W \to \mu\nu$	0.5	199851	AcerMC
single top : s − channel $W \to \tau\nu$	0.5	175001	AcerMC
single top : t − channel $W \to e\nu$	7.83	999949	AcerMC
single top : t − channel $W \to \mu\nu$	7.83	999949	AcerMC
single top : t − channel $W \to \tau\nu$	7.83	998996	AcerMC
single top : Wt − channel	15.6	769898	AcerMC

Table A.3: Simulated $t\bar{t}$, single top, single top in association with a W boson and Diboson datasets.

m_H (GeV)	ggF $\sigma(pp \to H)$ (pb)	VBF $\sigma(pp \to q\bar{q}'H)$ (pb)	$BR(H \to \tau\tau)$	$BR(H \to WW^*)$
100	24.02	1.546	0.0836	0.0111
105	21.78	1.472	0.0825	0.0243
110	19.84	1.398	0.0802	0.0482
115	18.13	1.332	0.0765	0.0867
120	16.63	1.269	0.0710	0.143
125	15.31	1.211	0.0637	0.216
130	14.12	1.6868	0.0548	0.305
135	13.08	1.10	0.0452	0.403
140	12.13	1.052	0.0354	0.504
145	11.27	1.004	0.0261	0.603
150	10.50	0.9617	0.0178	0.699
160	9.080	0.8787	0.00396	0.909
170	7.729	0.8173	0.000920	0.965
180	6.739	0.7480	0.000587	0.932

Table A.4: $H \to \tau\tau$ and $H \to WW^* \to \ell\nu\tau\nu$ signal process cross sections and branching ratios as a function of SM Higgs boson mass (taken from Reference [14]).

m_H (GeV)	q^2 variation (%)		PDF + α_s variation (%)	
	ggF	VBF	ggF	VBF
100	16.5	1.0	8.0	4.0
105	16.4	1.0	8.0	4.0
110	16.3	1.0	8.0	4.0
115	16.5	1.0	8.9	4.0
120	16.3	1.0	8.0	4.0
125	16.3	1.0	8.0	4.0
130	16.4	1.0	8.0	4.0
135	16.1	1.0	8.0	4.0
140	16.2	1.0	8.0	4.0
145	16.4	1.0	8.0	4.0
150	16.6	1.0	8.0	4.0
160	16.6	1.0	8.0	4.0
170	16.6	1.0	8.0	4.0
180	16.6	1.0	8.0	4.0

Table A.5: Theoretical uncertainties associated with the gluon-gluon fusion (ggF) and vector-boson fusion (VBF) signal processes as a function of SM Higgs boson mass, taken from reference [14].

References

[1] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[2] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[3] A. Salam and J. C. Ward. Electromagnetic and weak interactions. *Phys. Rev. Lett.*, pages 168–171, 1964.

[4] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998.

[5] F. Halzen and A. D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984.

[6] K. et al Nakamura. Review of Particle Physics. *Journal of Physics G*, 33:1+, 2006.

[7] C. Amsler et al. Review of particle physics. *Physics Letters B*, 667(15):1 – 6, 2008. Review of Particle Physics.

[8] P.W. Higgs. *Phys. Lett.*, 12:132, 1964.

[9] F. Englert and R. Brout. *Phys. Rev. Lett.*, 13:321, 1964.

[10] P.W. Higgs. *Phys. Rev. Lett.*, 13:508, 1964.

[11] P.W. Higgs. *Phys. Rev. Lett.*, 145:1156, 1966.

[12] LEP Electroweak Working Group. . 2011. `http://lepewwg.web.cern.ch/LEPEWWG/`.

[13] Tevatron New-Phenomena and Higgs Working Group. . *hep-ex/1203.3774*, 2011.

[14] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *CERN-2011-002*, CERN, Geneva, 2011.

[15] The CMS Collaboration. G. l. bayatian et al. *J. Phys. G: Nucl. Part. Phys.*, 34(995), 2007.

[16] S. Asai et al. *Eur. Phys. J. C*, 32S2(19), 2004.

[17] D. Rainwater F. Maltoni and S. Willenbrok. *Phys. Rev. D*, 66(034022), 2002.

[18] E. Richter-Was and Acta. M. Sapinski. *Phys. Polon. B*, 30(1001), 1999.

[19] T. Müller V. Drollinger and D. Denegri. *arXiv:hep-ph/0111312v1 (2001).*

[20] M. Rubin J. M. Butterworth, A. R. Davison and G. P. Salam. *Phys. Rev. Lett.*, 104(111801), 2010.

[21] K. Black et al. *arXiv:1010.3698v1*, 2010.

[22] E. Gabrielli et al. *Nucl. Phys. B*, 64(781), 2007.

[23] A. Nikitenko D. Zeppenfeld, R. Kinnunen and E. Richter-Was. *Phys. Rev. D*, 62(013309), 2000.

[24] M. Duhrssen et al. *Phys. Rev. D*, 70(113009), 2004.

[25] L. Evans and P. Bryant. LHC machine. *Journal of Instrumentation*, 3:S08001, 2008.

[26] C. Lefèvre (CERN). CERM-DI-0812015, 2008.

[27] G. Aad et al. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3:S08003, 2008.

[28] G. Aad et al. Expected performance of the ATLAS experiment: detector, trigger and physics. *arXiv*, 901:2008–20.

[29] R. S. Thorne A. D. Martin, W. J. Stirling and G. Watt. . `http://projects.hepforge.org/mstwpdf/`.

[30] ATLAS: Detector and physics performance technical design report. Volume 1. CERN-LHCC-99-14.

[31] R. Brun and F. Rademakers. ROOT–An object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1-2):81–86, 1997.

[32] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[33] The ATLAS Collaboration. *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*, volume CERN-OPEN-2008-020. CERN, Geneva, 2009. Pages 368-396.

[34] The ATLAS Collaboration. Commisioning of the atlas high-performance b-tagging algorithms in 7 tev data. 2011. ATLAS-CONF-2011-102.

[35] The ATLAS Collaboration. *Eur. Phys. J. C*, 72:1909, 2012.

[36] The ATLAS Collaboration. *Eur. Phys. J. C*, 72:1844, 2012.

[37] S. Agostinelli et al. The GEANT4 Collaboration. Geant4 - a simulation toolkit. *Nucl. Instrum. Meth*, A506:250, 2003.

[38] A. Barr and C.Boddy. *Phys. Educ.*, 47:270–273, 2012.

[39] Google Ic. 2011 Android. . `http://www.android.com`.

[40] Gary Taylor. Visualizing the atlas inner detector with atlantis. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 549.

[41] OpenGL-ES. 2011. . `http://www.khronos.org/opengles`.

[42] A. Djouadi. *Phys. Rep.*, 457(1), 2008.

[43] S. Peñaranda J. Guasch, W. Hollik. *Phys. Lett. B*, 367(505), 2001.

[44] T. Gleisberg et al. *J. High Energy Phys.*, 007(0902), 2009. We use Version 1.2.1.

[45] Pavel M. Nadolsky et al. Implications of CTEQ global analysis for collider observables. *Phys. Rev.*, D78:013004, 2008.

[46] F. Piccinini R. Pittau A. Polosa M.L. Mangano, M. Moretti. Alpgen, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[47] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[48] M. Bähr et al. *Eur. Phys. J. C*, 58(639), 2008.

[49] X. Rouby S. Ovyn and V. Lemaitre. *arXiv:hep-ph/0903.2225v3*, 2010.

[50] M. Cacciari and G. P. Salam. *Phys. Lett. B*, 57(641), 2006.

[51] The ATLAS Collaboration. Measurement of the mis-identification probability of tau leptons from hadronic jets and from electrons. 2011. ATLAS-CONF-2011-113.

[52] M. Spira. http://people.web.psi.ch/spira/v2hv/.

[53] J. Kalinowkski A. Djouadi and M. Spira. *Comput. Phys. Commun.*, 56(108), 1998.

[54] J. Campbell, R.K. Ellis, and D. Rainwater. Next-to-leading order QCD predictions for W+ 2 jet and Z+ 2 jet production at the CERN LHC. *Physical Review D*, 68(9):094021, 2003.

[55] S. Dawson et al. *Phys. Rev. D*, 68(034022), 2003.

[56] W. Beenakker et al. *Nucl. Phys. B*, 151(653), 2003.

[57] M. Spira. *Fortsch. Phys.*, 46(203), 1998.

[58] The ATLAS Collaboration. Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS. Technical Report ATLAS-COM-CONF-2011-152, CERN, Geneva, Nov 2011.

[59] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[60] Rene Brun et al. Root website, 1995-2011. `http://root.cern.ch/drupal/`.

[61] The ATLAS Collaboration. Identification of Hadronic Tau Decays for Summer 2011. Technical Report ATLAS-COM-CONF-2011-112, CERN, Geneva, Jul 2011.

[62] The ATLAS Collaboration. Expected photon performance in the ATLAS experiment. ATL-PHYS-PUB-2011-007.

[63] 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 sqrts=7 TeV. ATL-COM-PHYS-2010-247, ATLAS-CONF-2010-038.

[64] ATLAS Collaboration. *Phys. Rev. D*, 84:112006, Dec 2011.

[65] P. Nason S. Frixione and B.R. Webber. Matching NLO QCD and parton showers in heavy flavour production. *JHEP*, 0308:007, 2003.

[66] G. Corcella et al. HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes). *JHEP*, 01:010, 2001.

[67] J. R. Forshaw J. M. Butterworth and M. H. Seymour. Multiparton interactions in photoproduction at hera. *Z. Phys*, C72:637, 1996.

[68] A. S. et al. Higgs boson production in gluon fusion. *JHEP*, 02:029, 2009.

[69] J. H. Kuhn S. Jadach and Z. Was. Tauola - a library of monte carlo programs to simulate decays of polarized leptons. *Comput. Phys. Commun.*, 64:275, 1990.

[70] B. V. Eijk E. Barberio and Z. Was. Photos - a universal monte carlo for qed radiative corrections in decays. *Comput. Phys. Commun.*, 66:115, 1991.

[71] A. Elagin, P. Murat, A. Pranko, and A. Safonov. A New Mass Reconstruction Technique for Resonances Decaying to di-tau. *Nucl.Inst.Meth.*, A654:481, 2011.

[72] The ATLAS Collaboration. Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector in 2011. Technical Report ATLAS-COM-CONF-2011-130, CERN, Geneva, Jul 2011.

[73] The ATLAS Collaboration, 2012. `http://svn.cern.ch/reps/atlasoff/PhysicsAnalysis/MuonID/MuonIDAnalysis%/MuonEfficiencyCorrections/tags/MuonEfficiencyCorrections-01-00-10/`.

[74] The ATLAS Collaboration, 2012. `http://svn.cern.ch/reps/atlasoff/Reconstruction/egamma/egammaAnalysis/%tags/egammaAnalysisUtils-00-02-70/`.

[75] J Butterworth, E Dobson, U Klein, B Mellado Garcia, T Nunnemann, J Qian, D Rebuzzi, and R Tanaka. Single boson and diboson production cross sections in pp collisions at sqrts=7 tev. Technical Report ATL-COM-PHYS-2010-695, CERN, Geneva, Aug 2010.

[76] J Butterworth, E Dobson, U Klein, B Mellado Garcia, T Nunnemann, J Qian, D Rebuzzi, and R Tanaka. Single boson and diboson production cross sections in pp collisions at sqrts=7 tev. Technical Report ATL-COM-PHYS-2010-695, CERN, Geneva, Aug 2010.

[77] The ATLAS Collaboration. Close-by jet effects on jet energy scale calibration in pp collisions at sqrt(s)=7 tev with the atlas detector. Technical Report ATL-CONF-2011-062, CERN, 2011.

[78] The ATLAS Collaboration, 2012. `http://svn.cern.ch/reps/atlasoff/Reconstruction/Jet/JetUncertainties/t%ags/JetUncertainties-00-03-04/`.

[79] The ATLAS Collaboration, 2012. `http://svn.cern.ch/reps/atlasoff/Reconstruction/MissingETUtility/tags/%MissingETUtility-00-02-12/`.

[80] K. Cranmer E. G. G. Cowan and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Arxiv preprint hep-ex/10071727*, 2010.

[81] The ATLAS Collaboration. An update to the combined search for the standard model higgs boson with the atlas detector at the lhc using up to 4.9 fb1 of pp collision data at s = 7 tev. Technical Report ATLAS-CONF-2012-019, CERN, Geneva, Mar 2012.