

SEARCH FOR NEUTRAL MINIMALLY SUPERSYMMETRIC STANDARD  
MODEL HIGGS DECAYING TO TWO HADRONIC TAUS WITH THE  
ATLAS DETECTOR IN  $pp$  COLLISIONS AT 7 TEV CENTER OF MASS  
ENERGY

MSSM  $A/H/h \rightarrow \tau_h \tau_h$

by

Saminder Dhaliwal

A thesis submitted in conformity with the requirements  
for the degree of Doctor of Philosophy  
Graduate Department of Physics  
University of Toronto

Copyright © 2012 by Saminder Dhaliwal

# Abstract

Search for neutral Minimally Supersymmetric Standard Model Higgs decaying to two hadronic taus with the ATLAS detector in  $pp$  collisions at 7 TeV center of mass energy

MSSM  $A/H/h \rightarrow \tau_h \tau_h$

Saminder Dhaliwal

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2012

This thesis presents the first search for a neutral Minimal Supersymmetric Standard Model Higgs boson decaying to a pair of hadronic taus in proton-proton collisions with a center-of-mass energy of  $\sqrt{s} = 7$  TeV. Results are presented for an integrated luminosity of  $1.056 \text{ fb}^{-1}$  using data from the ATLAS detector at the Large Hadron Collider. After signal selection, 245 events are observed. The number of events is consistent the background estimate of  $256 \pm 26$ . An exclusion limit for the Higgs boson production is derived as a function of (cross-section) $\times$ (branching ratio). This limit is presented as a function of two parameters:  $m_A$  and  $\tan \beta$  in the  $m_h^{max}$  scenario. A point in the  $(m_A, \tan \beta)$  phase space is excluded if the signal hypothesis is rejected at the 95% confidence level.

*For my parents.*

## Acknowledgements

There are many people who deserve thanks for their support and encouragement during the course of my graduate studies. I'm indebted to my colleagues in the ATLAS collaboration, and in particular to the tau working group. I'm immensely grateful to the ATLAS group in Toronto for guidance every step of the way, especially to my supervisor Robert Orr for giving me space when needed, and challenging me to really interpret results when necessary.

I would like to thank all the friends and family who have seen me through the last five years. I owe a great deal to my mother who seems to have some sort of unwavering faith in my ability and always supports my endeavours. I'm lucky to have made some great friends both in Toronto and at CERN, without whom the experience wouldn't have been nearly as enjoyable. Finally, my thanks to Chris for brightening my days.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theory</b>	<b>5</b>
2.1	The Standard Model . . . . .	6
2.1.1	Quantum field theory and gauge invariance . . . . .	8
2.1.2	Spontaneous symmetry breaking . . . . .	10
2.1.3	Problems with the Standard Model . . . . .	13
2.2	Supersymmetry . . . . .	14
2.3	Minimal Supersymmetric Standard Model . . . . .	15
2.3.1	MSSM Higgs Sector . . . . .	19
<b>3</b>	<b>LHC</b>	<b>21</b>
<b>4</b>	<b>ATLAS Detector</b>	<b>25</b>
4.1	Inner Detector . . . . .	27
4.2	Calorimeters . . . . .	33
4.3	Muon Spectrometer . . . . .	39
4.4	Trigger . . . . .	43
4.5	Pile-up . . . . .	44
<b>5</b>	<b>Object Reconstruction</b>	<b>46</b>
5.1	Track Reconstruction with the Inner Detector . . . . .	46

5.2	Cluster formation in the Calorimeters . . . . .	48
5.3	Hadronic calibration . . . . .	48
5.4	Jets . . . . .	50
5.4.1	anti- $k_T$ algorithm . . . . .	50
5.5	Hadronically decaying $\tau$ leptons . . . . .	51
5.6	Electrons . . . . .	52
5.7	Muons . . . . .	52
5.8	Missing transverse energy . . . . .	53
5.9	Object Identification . . . . .	53
<b>6</b>	<b>Data and Monte Carlo</b>	<b>54</b>
6.1	Data sets used in this analysis . . . . .	54
6.2	Monte Carlo . . . . .	54
<b>7</b>	<b>MSSM Higgs signal</b>	<b>58</b>
7.1	Higgs production . . . . .	58
7.1.1	Production cross-section . . . . .	59
7.2	Higgs decay to taus . . . . .	63
7.3	Tau decay modes . . . . .	63
<b>8</b>	<b>Background Processes</b>	<b>65</b>
8.1	QCD multi-jet events . . . . .	66
8.2	$Z \rightarrow ee/\mu\mu$ and $W \rightarrow e\nu/\mu\nu + \text{jets}$ . . . . .	67
8.3	$W \rightarrow \tau\nu + \text{jets}$ . . . . .	68
8.4	$Z/\gamma^* \rightarrow \tau\tau (+ \text{jets})$ . . . . .	68
8.5	top quark production . . . . .	68
8.6	Diboson: $WW, WZ, ZZ$ . . . . .	70

<b>9</b>	<b>Tau Identification</b>	<b>73</b>
9.1	Tau Identification variables . . . . .	73
9.2	Log likelihood ratio discriminant . . . . .	81
9.3	$\tau_h$ Identification Efficiency in Data . . . . .	86
9.3.1	$W$ background scale factors . . . . .	88
9.3.2	QCD background . . . . .	89
<b>10</b>	<b>Event Selection</b>	<b>93</b>
10.1	Scale factors . . . . .	98
10.2	Summary of Event selection . . . . .	98
<b>11</b>	<b>Trigger scale factors</b>	<b>103</b>
11.1	Trigger Factorization . . . . .	103
11.2	Single trigger efficiency measurement . . . . .	105
11.3	Single trigger fake rate measurement . . . . .	110
<b>12</b>	<b>Electroweak Background</b>	<b>112</b>
12.1	$Z \rightarrow \tau_h \tau_h$ background estimation . . . . .	113
12.2	$W \rightarrow \tau \nu$ background estimation . . . . .	114
<b>13</b>	<b>QCD Background Estimation</b>	<b>121</b>
13.1	ABCD Method . . . . .	121
13.2	Verification of the underlying assumptions . . . . .	124
<b>14</b>	<b>Validation of the Combined Background Estimation</b>	<b>129</b>
<b>15</b>	<b>Systematic Uncertainties</b>	<b>139</b>
15.1	Monte Carlo cross-section predictions . . . . .	139
15.1.1	Signal cross-section . . . . .	139
15.1.2	$Z$ and $W$ Background cross-section . . . . .	143

15.2	$\tau_h$ identification efficiency . . . . .	144
15.3	Trigger efficiencies for $\tau_h$ decays . . . . .	144
15.4	Trigger fake rates of $\tau_h$ candidates . . . . .	145
15.5	Tau energy scale . . . . .	145
15.6	$E_T^{Miss}$ uncertainties . . . . .	146
15.6.1	Topological cluster energy scale . . . . .	146
15.6.2	$E_T^{miss}$ resolution . . . . .	147
15.7	Electron and Muon associated uncertainties . . . . .	148
15.7.1	Reconstruction and Identification efficiencies . . . . .	149
15.7.2	Energy scale . . . . .	150
15.8	Luminosity in data . . . . .	150
15.9	Summary of Systematic Uncertainties . . . . .	150
<b>16</b>	<b>Results</b>	<b>153</b>
<b>17</b>	<b>Conclusions</b>	<b>158</b>
	<b>Appendices</b>	<b>161</b>
<b>A</b>	<b>Run numbers and data periods</b>	<b>161</b>
<b>B</b>	<b>Monte Carlo Generators</b>	<b>163</b>
B.0.1	QCD interlude . . . . .	164
<b>C</b>	<b>Parton Distribution Functions</b>	<b>167</b>
<b>D</b>	<b>Signal cross-sections predicted from Monte Carlo</b>	<b>169</b>
<b>E</b>	<b>Electroweak and top Monte Carlo</b>	<b>172</b>
	<b>Bibliography</b>	<b>172</b>

# List of Tables

2.1	Fermion generations . . . . .	7
2.2	Gauge bosons . . . . .	7
2.3	Chiral supermultiplet fields in the MSSM . . . . .	17
2.4	Gauge supermultiplet fields in the MSSM . . . . .	17
2.5	MSSM Higgs couplings . . . . .	20
6.1	Data periods and runs in this analysis . . . . .	57
6.2	Generators used in this analysis . . . . .	57
7.1	Theoretical production cross-sections for MSSM neutral Higgs at the LHC	61
8.1	QCD dijet cross-sections . . . . .	66
8.2	Background cross-sections . . . . .	72
9.1	Identification variables that enter likelihood function . . . . .	82
9.2	Selection for tau identification efficiency study with data . . . . .	86
9.3	Division of data based on the ABCD background estimation method . . .	89
9.4	Tau identification efficiency in Monte Carlo and data . . . . .	92
10.1	Event selection cutflow from data and each Monte Carlo background . .	101
10.2	Event selection cutflow from data and EW and top Monte Carlo background	102
11.1	Selection for single trigger efficiency measurement in data . . . . .	107
11.2	Trigger efficiency scale factors applicable to Monte Carlo . . . . .	109

11.3	Selection for single trigger fake rate measurement in data . . . . .	110
11.4	Trigger fake rate scale factors applicable to Monte Carlo . . . . .	111
13.1	Signal and control regions defined by ABCD method for QCD background estimation . . . . .	122
13.2	Number of events from data and Monte Carlo in signal and QCD control regions . . . . .	123
15.1	Production cross-section uncertainty for signal . . . . .	140
15.2	Production cross-section uncertainty for $Z$ + jets and $W$ + jets . . . . .	143
15.3	Systematic uncertainties from trigger and $\tau_h$ identification . . . . .	145
15.4	Tau energy scale uncertainties . . . . .	146
15.5	Variation in acceptance from topological cluster energy scale . . . . .	148
15.6	Variation in acceptance from $E_T^{Miss}$ resolution uncertainty . . . . .	148
15.7	Variation in acceptance from electron and muon reconstruction and iden- tification uncertainties . . . . .	150
15.8	Summary of systematic uncertainties . . . . .	152
A.1	ATLAS data periods and runs used in this analysis. . . . .	162
D.1	$b$ -associated Higgs production cross-sections for all signal samples . . . .	170
D.2	gluon fusion Higgs production cross-sections for all signal samples . . . .	171
E.1	$W$ background samples and cross-sections . . . . .	173
E.2	$Z/\gamma^*$ background samples and cross-sections ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) . .	174
E.3	$Z/\gamma^*$ background samples and cross-sections ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) . .	175
E.4	$t\bar{t}$ and single top background samples and cross-sections . . . . .	176
E.5	$WW$ background samples and cross-sections . . . . .	177
E.6	$WZ$ and $ZZ$ background samples and cross-sections . . . . .	178

# List of Figures

1.1	Tevatron limits on MSSM Higgs in $m_A$ - $\tan \beta$ plane using $1.8 \text{ fb}^{-1}$ of data from CDF, and $2.2 \text{ fb}^{-1}$ from D0 . . . . .	2
1.2	Tevatron limits on MSSM Higgs in $m_A$ - $\tan \beta$ plane with $7.3 \text{ fb}^{-1}$ from D0 . . . . .	3
2.1	MSSM Higgs couplings to fermions and scalars . . . . .	15
3.1	Layout of the CERN accelerator complex and the LHC experiments . . . . .	23
4.1	Schematic of ATLAS detector . . . . .	26
4.2	Inner Detector . . . . .	29
4.3	Particle traversing Inner Detector . . . . .	32
4.4	Calorimeters . . . . .	35
4.5	Barrel electromagnetic calorimeter in construction . . . . .	36
4.6	Muon spectrometer . . . . .	40
7.1	MSSM Higgs production at the LHC . . . . .	58
7.2	Higgs masses as a function of $m_A$ . . . . .	60
7.3	Theoretical production cross-sections for MSSM neutral Higgs at the LHC . . . . .	62
7.4	Tau decay Feynman diagram . . . . .	64
8.1	$W$ and $Z$ decay Feynman diagrams . . . . .	67
8.2	$t\bar{t}$ decay to two hadronic taus and $b$ -jets . . . . .	69
8.3	$t$ -channel single top production . . . . .	69

8.4	$s$ -channel and $W$ associated single top production . . . . .	71
8.5	Diboson production . . . . .	71
9.1	Reconstructed track multiplicity of hadronic taus . . . . .	74
9.2	Tau identification variable: electromagnetic radius . . . . .	77
9.3	Tau identification variable: track radius . . . . .	78
9.4	Tau identification variable: cluster mass . . . . .	78
9.5	Tau identification variable: fraction of momentum carried by highest- $p_T$ tau track . . . . .	79
9.6	Tau identification variable: electromagnetic fraction . . . . .	79
9.7	Tau identification variable: track mass . . . . .	80
9.8	Tau discriminant log-likelihood-ratio score . . . . .	83
9.9	tau discriminant performance . . . . .	85
9.10	Visible mass from tau identification study with data . . . . .	91
10.1	$E_T^{Miss}$ distribution from signal, Monte Carlo based background and data	97
11.1	Trigger efficiency . . . . .	105
11.2	Factorized triggers compared to combined triggers . . . . .	106
11.3	Trigger efficiencies as a function of hadronic tau $p_T$ (data periods <b>B</b> <sub>2</sub> to <b>F</b> )	108
11.4	Trigger efficiencies as a function of hadronic tau $p_T$ (data periods <b>F</b> to <b>H</b> )	108
12.1	Embedded data kinematics compared to $Z \rightarrow \tau_h \tau_h$ Monte Carlo . . . . .	116
12.2	Visible mass from embedded data compared to $Z \rightarrow \tau_h \tau_h$ Monte Carlo . .	117
12.3	Embedded data kinematics compared to $W \rightarrow \tau_h \nu$ Monte Carlo . . . . .	118
12.4	Visible mass from embedded data compared to $W \rightarrow \tau_h \nu$ Monte Carlo . .	119
12.5	Hadronic tau multiplicity in embedded data and $W \rightarrow \tau_h \nu$ Monte Carlo .	119
13.1	Log-likelihood-ratio discriminant score for same and opposite sign taus .	125

13.2	Number of tracks associated to two highest- $p_T$ hadronic taus for same and opposite sign taus . . . . .	126
13.3	$p_T$ of two highest- $p_T$ hadronic taus for same and opposite sign taus . . .	127
13.4	Visible mass distribution for medium identified taus compared to loose identified taus . . . . .	128
14.1	Transverse momentum of highest and second highest- $p_T$ tau from data and estimated background . . . . .	131
14.2	Pseudorapidity of highest and second highest- $p_T$ tau from data and estimated background . . . . .	132
14.3	Azimuthal angle of highest and second highest- $p_T$ tau from data and estimated background . . . . .	133
14.4	Charge of the highest and second highest- $p_T$ tau from data and estimated background . . . . .	134
14.5	Number of tracks associated to highest and second highest- $p_T$ tau from data and estimated background . . . . .	135
14.6	Electromagnetic radius of highest and second highest- $p_T$ tau from data and estimated background . . . . .	136
14.7	Missing transverse energy. . . . .	137
14.8	Angular distance between highest and second highest- $p_T$ tau from data and estimated background . . . . .	137
15.1	Production cross-section uncertainty for $b\bar{b}A/H/h$ signal . . . . .	141
15.2	Production cross-section uncertainty for $gg \rightarrow A/H/h$ signal . . . . .	142
15.3	Functions used to scale topological cluster energy for uncertainty calculation	147
16.1	Visible mass of the two leading $\tau_h$ candidates for backgrounds, signal and data. Also shown is a $m_A = 200\text{GeV}$ Higgs Boson in the $m_h^{max}$ scenario with $\tan\beta = 20$ . . . . .	154

16.2	Derived limit for neutral MSSM Higgs in $m_A$ and $\tan\beta$ plane . . . . .	156
16.3	Derived limit for cross-section of a generic Higgs boson . . . . .	157
C.1	Parton Distribution functions from MSTW 2008 . . . . .	167

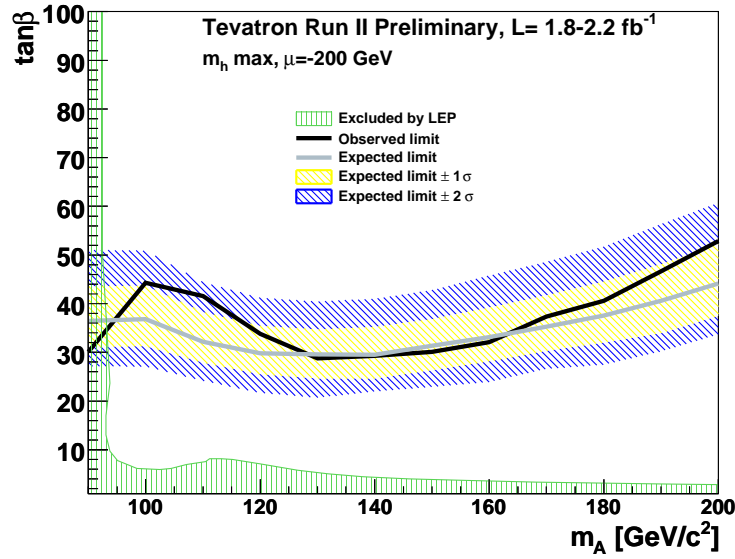
# Chapter 1

## Introduction

One of the greatest achievements of twentieth century physics was the development of the Standard Model of particle physics, a unified description of the elementary particles that make up our universe, their properties and interactions via four fundamental forces. Despite its great successes and the accuracy to which it has been verified experimentally, the theory has flaws. The mechanisms it describes are incomplete and it makes predictions only after the insertion of parameters measured by particle and astrophysics experiments. Most seriously, the theory must break down at higher energies as yet unexplored. The quest for a theory describing physics beyond the Standard Model has engaged physicists for the last few decades, with many new ideas postulated, attempting to address the big unanswered questions. The recently operational Large Hadron Collider will explore new energies and surely take us from conjecture to concrete discovery. One popular theory that becomes open to observation with the Large Hadron Collider is Supersymmetry, where each Standard Model particle is paired with a superpartner. As with the Standard Model, the Higgs mechanism is still employed to achieve electroweak symmetry breaking, however a more complex sector containing at least five Higgs bosons is required.

This dissertation describes the search for a neutral Higgs boson in the Minimal Supersymmetric extension to the Standard Model through the decay channel of two hadronic

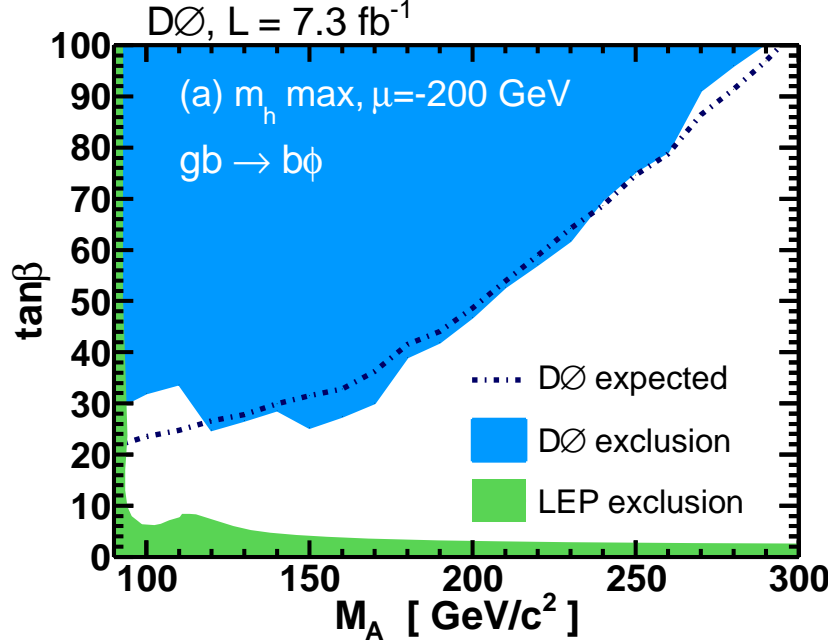
taus. It is the first search for any new physics involving two hadronic taus performed with the ATLAS detector at the Large Hadron Collider. The Minimally Supersymmetric Higgs has been sought at lower energy colliders, resulting in exclusion limits from both the Large Electron Positron Collider (LEP) and the Tevatron. These limits are expressed in terms of two free parameters from the Minimal Supersymmetric Model: the mass of one of the five observable Higgs,  $m_A$ , and the ratio of the two Higgs doublet's vacuum expectation values,  $\tan\beta$ . Limits are shown in Figure 1.1 for a combination from the Tevatron experiments using  $1.8 \text{ fb}^{-1}$  of data from CDF, and  $2.2 \text{ fb}^{-1}$  from D0 [1].



**Figure 1.1:** 95% Confidence limits on the MSSM Higgs in the  $m_A$ - $\tan\beta$  plane from the Tevatron. Limit set using  $1.8 \text{ fb}^{-1}$  of data from CDF, and  $2.2 \text{ fb}^{-1}$  from D0 recorded with a center-of-mass energy of 1.96 TeV. The black line denotes the observed limit, above which the phase space is excluded. The grey line shows the expected limit and the hatched yellow and blue regions denote the  $\pm 1$  and  $2\sigma$  bands around the expectation. The shaded light-green area shows the limits from LEP.

A recently updated limit from D0 alone is given in Figure 1.2 with  $7.3 \text{ fb}^{-1}$  of data. [2].

The analysis described in this thesis expands on the LEP and Tevatron searches



**Figure 1.2:** 95% Confidence limits on the MSSM Higgs in the  $m_A$ - $\tan\beta$  plane from an updated  $D0$  analysis using  $7.3 \text{ fb}^{-1}$  of data from the Tevatron. The solid coloured areas show the excluded phase space from  $D0$  in blue, and LEP in green.

using data recorded at a center-of-mass energy higher than ever available before. The dissertation begins by introducing the theoretical motivation for the search in Chapter 2. The ‘experimental apparatus,’ the Large Hadron Collider and the ATLAS detector are described in Chapters 3 and 4. Any expected signal in the Large Hadron Collider has been calculated theoretically, and will be discussed in Chapter 7. Other than the potential signal, there will be an abundance of other ‘background’ processes, detailed in Chapter 8, that must be suppressed for a successful analysis. This requires prudent event selection, documented in Chapter 10. Obviously, hadronic taus are of great importance in this analysis, and the method of identifying these particles in ATLAS is described in detail in Chapter 9.

This thesis uses natural units, in which the speed of light and the reduced Planck constant are set equal to one,  $c = \hbar = 1$ . As a result, all of energy, momentum and mass

are expressed in terms of giga-electron volts, GeV. Any discussion of the electroweak scale is characterized by an energy of  $\sim 100$  GeV, and the Planck scale at  $10^{19}$  GeV.

# Chapter 2

## Theory

The Standard Model of particle physics is the most complete theory of physics at the subatomic scale [3, 4, 5]. It describes the fundamental particles that form our universe, and their interactions with each other. These interactions are explained in terms of three forces: the weak, electromagnetic, and strong force. Gravity is not accounted for, the gravitational interaction between elementary particles is very small and safely neglected in the description of the other forces. One primary success of the Standard Model is the unified theory of the electroweak force, which combines electromagnetic and weak interactions.

The theory is not without its problems. Recently the phenomenon of neutrino flavour oscillation, which requires neutrinos to have non-zero mass, has been experimentally established [6, 7]. In the last few decades astrophysics experiments have collected evidence for the existence of dark matter [8]. Neither of these developments are described by the Standard Model. Removing the issues of gravity, dark matter, and neutrino oscillations, the Standard Model has inherent problems. One prominent concern is that of ‘fine tuning.’ If we assume the Standard Model reveals the structure of the universe up to the Planck scale, radiative corrections to the Higgs mass would inevitably make the mass extremely large, unless incredibly fine tuned cancellations are incorporated.

This Chapter describes the Standard Model and its particles. The limitations of the model are discussed and then a popular extension of the standard model, supersymmetry is introduced, with explanations as to why it fixes some of the problems associated with the Standard Model.

## 2.1 The Standard Model

The Standard Model (SM) is a gauge field theory that describes two types of point-like particles, fermions and bosons, and their interactions. Fermions are particles that carry an internal angular momentum, spin, of  $1/2$ . Elementary fermions fall into two categories, quarks and leptons. Bosons are particles with integer spin. Fermionic particles interact by exchanging vector bosons of spin 1.

In terms of field theories, the SM is a renormalizable quantum field theory that results from a spontaneously broken  $SU_C(3) \times SU_L(2) \times U_Y(1)$  local gauge invariance. Each of the three forces is associated with a local symmetry transformation. The strong force is characterized by  $SU_C(3)$  transformations, and the weak and electromagnetic forces, combined in the electroweak force, by an  $SU_L(2) \times U_Y(1)$  symmetry [11].

The strong force, often referred to as quantum chromodynamics (QCD), describes quarks, which form colour triplets of the QCD gauge group. Gluons are the bosons that mediate the strong force, and they themselves carry colour. The bosons of the electroweak force are the  $W^\pm$ ,  $Z$ , and the photon ( $\gamma$ ). In the Standard Model there are three copies (or families) of fermion particles, each copy of which differs only by mass, with each family successively heavier. All the elementary fermions are listed in Table 2.1 with their electromagnetic charges and observed masses. Each quark exists in three varieties, one for each of the colour charges. The mediating gauge bosons are given in Table 2.2. Every particle listed has been discovered and is thought to be fundamental. Not shown, but observed experimentally, are the associated antiparticles to each particle.

An antiparticle has the opposite internal quantum numbers, including electric charge, to its related particle.

	Leptons	Charge	Mass (GeV)	Quarks	Charge	Mass (GeV)
1st Generation	$e$	-1	0.000548	$u$	$+\frac{2}{3}$	0.0023
	$\nu_e$	0	$\sim 0$	$d$	$-\frac{1}{3}$	0.0048
2nd Generation	$\mu$	-1	0.113	$c$	$+\frac{2}{3}$	1.275
	$\nu_\mu$	0	$\sim 0$	$s$	$-\frac{1}{3}$	0.095
3rd Generation	$\tau$	-1	1.776	$t$	$+\frac{2}{3}$	173.5
	$\nu_\tau$	0	$\sim 0$	$b$	$-\frac{1}{3}$	4.65

**Table 2.1:** *The fundamental fermions divided into generations. The electromagnetic charge and mass, taken from Reference [12], is given for each particle.*

Force	Boson	Mass (GeV)
Electromagnetic	$\gamma$	0
Weak	$Z$	80.385
Weak	$W^\pm$	91.188
Strong	$g$	0

**Table 2.2:** *Mediating gauge bosons for each force, with masses in GeV [12]. The photon and gluons are electromagnetically neutral. There are two  $W$  bosons, with positive and negative charge.*

### 2.1.1 Quantum field theory and gauge invariance

The dynamics of particles can be described by quantum field theory (QFT), which combines quantum mechanics and special relativity in a Lagrangian formalism. The dynamics are contained in a single Lagrangian density,  $\mathcal{L}$ , where the Euler-Lagrange equations are solved to obtain the equations of motion. The fields in a Lagrangian are operators that create and annihilate particles [11].

The Lagrangian that correctly predicts the Dirac equation and describes the motion of a free fermion field,  $\psi(x)$ , with mass  $m$ , is given by Equation 2.1.

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi = \bar{\psi}i\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \quad (2.1)$$

We can introduce interactions of fermion fields with gauge bosons by forcing this Lagrangian to be invariant under local transformations of the field. Where ‘local’ indicates that the transformation depends on some space-time point  $x$ . For example, requiring invariance under a local U(1) phase transformation, as given in equation 2.2,

$$\psi \rightarrow \psi' = e^{ie\chi(x)}\psi, \quad (2.2)$$

necessitates the introduction of a field,  $A^\mu$ . The field enters the Lagrangian (Equation 2.1) with the following replacement (called the covariant derivative):

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ieA^\mu. \quad (2.3)$$

The theory is gauge invariant if the covariant derivative transforms as the matter fields themselves:

$$D_\mu\psi \rightarrow (D_\mu\psi)' = e^{ie\chi(x)}(D_\mu\psi). \quad (2.4)$$

This new field,  $A^\mu$ , known as a gauge field, must therefore transform according to Equa-

tion 2.5.

$$A^\mu = A'^\mu = A^\mu - \frac{1}{e} \partial^\mu \chi \quad (2.5)$$

The transformations above describe the quantum electrodynamic (QED) Lagrangian. Including the kinetic terms for the free electromagnetic field ( $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ ) we have:

$$\mathcal{L}_{QED} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}i\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi - eA_\mu \bar{\psi}i\gamma^\mu \psi. \quad (2.6)$$

Where we now have a description that includes the original free fermion field, and an interaction term with the newly introduced gauge field. We cannot add a mass term,  $m^2 A_\mu A^\mu$ , as such a term is not invariant under local U(1) transformations.

In a similar way to the Lagrangian of QED, the SM Lagrangian requires invariance under a  $SU_C(3) \times SU_L(2) \times U_Y(1)$  local gauge symmetry. The  $SU_C(3)$  term introduces eight spin 1 particles,  $G_\mu^\alpha$ , with  $\alpha$  from 1 to 8. These are the massless, coloured gluons, as indicated by the  $C$  subscript. The  $SU_L(2)$  invariance produces 3 massless gauge bosons,  $W_\mu^a$ , with  $a$  from 1 to 3. Finally,  $U_Y(1)$  invariance has one associated gauge boson,  $B_\mu$ . The subscript  $Y$  distinguishes the group associated with the quantum number of hypercharge from that of the electromagnetic charge. The four spin 1 gauge bosons from  $SU_L(2) \times U_Y(1)$  invariance are related to the physical bosons that mediate the weak interactions,  $W^\pm$ ,  $Z^0$ , and the photon, in a way that will be described further below.

From experimental observation we know that the charged weak vector bosons ( $W^\pm$ ) only interact with left-handed chiral states (which can be represented using the fifth gamma matrix as  $\Psi_L = \frac{1}{2}(1 - \gamma^5)\Psi$ ). The subscript  $L$  indicates left handed fermions, and the chirality of the SU(2) transformation.

In its unbroken form, the  $SU_C(3) \times SU_L(2) \times U_Y(1)$  gauge field theory gives a spectrum of massless particles. The fermion mass term  $m\bar{\psi}\psi$ , which was acceptable in the Lagrangian for QED, is no longer invariant because the right and left handed chiral components transform differently (as,  $m\bar{\psi}\psi = m\bar{\psi}_R\psi_L + m\bar{\psi}_L\psi_R$ ). As the particles in the

universe have mass, the symmetries cannot be left unbroken and a means for the particles to acquire mass is required. This method is described in the SM by the introduction of the Higgs sector. The Higgs mechanism spontaneously breaks the electroweak  $SU_L(2) \times U_Y(1)$  symmetry down to  $SU_L(2) \times U_{EM}(1)$ , and in the process introduces the Higgs boson.

### 2.1.2 Spontaneous symmetry breaking

Masses may be introduced to the gauge invariant  $SU_C(3) \times SU_L(2) \times U_Y(1)$  Lagrangian through a process called ‘Spontaneous Symmetry Breaking.’ The symmetry breaking materializes by the addition of scalar fields to the Lagrangian. The degrees of freedom from these new scalar particles transfer to the gauge boson fields to give them mass.

Spontaneous Symmetry breaking can be understood from a global  $U(1)$  symmetry by considering the Lagrangian given by Equation 2.7.

$$\mathcal{L} = (\partial^\mu \phi^\dagger)(\partial_\mu \phi) + \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2 \quad (2.7a)$$

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \quad (2.7b)$$

Where  $\phi(x)$  is a single complex scalar field, that transforms according to  $\phi(x) \rightarrow e^{i\alpha} \phi(x)$ . Equation 2.7b explicitly shows the potential term in the Lagrangian. If  $\mu^2 > 0$ , the potential has a single minimum value  $V_{\min} = 0$  at  $\phi = \phi_0 = 0$  and the potential preserves the symmetry of the Lagrangian. In the more interesting case, where  $\mu^2 < 0$ , the minimum of the potential lies on a circle on the complex plane with a radius  $\frac{\mu^4}{4\lambda}$ . This minimum occurs at  $\phi_0 = \sqrt{\frac{\mu^2}{2\lambda}} e^{i\theta}$

Only the magnitude of the potential is determined at this minimum,  $|\phi_0| = \sqrt{\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}$ , where  $v$  is the vacuum expectation value. The phase is arbitrary and there are an infinite number of minima, degenerate in all values of  $\theta$ . In a transition from a high energy density state to low energy density state (such as in the early universe), the field will

spontaneously fall into one of the degenerate minima. By realizing a particular vacuum the U(1) symmetry is spontaneously broken, since the vacuum expectation value is not gauge invariant; there results a symmetry of the Lagrangian that is not shared by the true, realized vacuum.

By introducing two real fields  $\sigma(x)$  and  $\eta(x)$ , both of which have zero vacuum expectation value, we can perturb  $\phi(x)$  around the chosen vacuum:

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \sigma(x) + i\eta(x)) \quad (2.8)$$

Expanding the Lagrangian in terms of these fields gives the Lagrangian in Equation 2.9

$$\mathcal{L} = \frac{1}{2}(\partial^\mu \sigma(x))(\partial_\mu \sigma(x)) - \frac{1}{2}(2\lambda v^2)\sigma(x)^2 + \frac{1}{2}(\partial^\mu \eta(x))(\partial_\mu \eta(x)) + \mathcal{L}_I + \text{constant} \quad (2.9)$$

we see that  $\sigma$  field has mass ( $m_\sigma = v\sqrt{2\lambda}$ ) but that the  $\eta$  field remains massless. (Interaction terms are denoted by  $\mathcal{L}_I$ ). One massive and one massless scalar boson have been obtained by introducing the complex scalar field. The massless boson is called a Goldstone boson.

If we shift to requiring local U(1) invariance,  $\phi(x) \rightarrow e^{i\alpha(x)}\phi(x)$ , where  $\alpha$  is now dependent on  $x$ , the potential term in Equation 2.7 is still invariant but the kinetic term is not. As previously for QED, we must introduce a gauge field to maintain invariance, and re-write the Lagrangian. By then considering the ground state for  $\mu^2 < 0$ , we are lead to spontaneous symmetry breaking, where now the goldstone boson can be absorbed to give the gauge field mass.

In the SM, the Higgs field,  $\phi$ , is required to be an SU(2) doublet of complex scalar fields, as represented by Equation 2.10. This means there are four real fields, or alternatively four degrees of freedom.

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad \bar{\Phi} = \begin{pmatrix} \bar{\phi}^0 \\ \phi^- \end{pmatrix} \quad (2.10)$$

The relevant terms in the initial SM Lagrangian are the  $W_\mu^a$  and  $B_\mu$  which each have 2 degrees of freedom, by including the Higgs field we begin with 12 degrees of freedom. After spontaneous symmetry breaking, we have a real scalar Higgs, with one degree of freedom, three massive weak bosons, the  $W^\pm$  and  $Z$ , with nine, and one massless photon with two degrees of freedom, summing again to 12. The scalar degrees of freedom are ‘eaten’ to give the  $W$  and  $Z$  bosons mass. Despite the added complexity in the SM, the basic structure of the Higgs mechanism remains unchanged. The vacuum expectation value,  $v$ , in the SM is given by:

$$v = \left(\frac{1}{\sqrt{2}G_F}\right)^{\frac{1}{2}} = 246 \text{ GeV} [12], \quad (2.11)$$

where  $G_F$  is the Fermi coupling constant. After acquiring a vacuum expectation value, the Higgs doublet can be used to give fermions mass through ‘Yukawa couplings.’ The masses take the form:

$$\sqrt{\frac{1}{2}}yv(\bar{\Psi}_L\psi_R + \bar{\psi}_R\Psi_L) \quad (2.12)$$

where  $y$  is a Yukawa coupling constant that is proportional to the fermion mass. Yukawa couplings are determined by the measured masses of the fermions, therefore the production and decay rates of the Higgs boson through fermion pairs are predicted in the SM.

As a side note we should observe that the quark mass eigenstates are not equal to the weak eigenstates, and this allows charged weak interactions to involve transitions between the three generations. This mixing is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix, and it allows charge parity (CP) violation (a symmetry in which particles are exchanged for their antiparticles, and spatial coordinates switched).

### 2.1.3 Problems with the Standard Model

It has been demonstrated that by imposing gauge invariance and spontaneous symmetry breaking, it is possible to predict the interactions of all the standard model particles, and their masses. The resulting masses are however, only defined in terms of the introduced model parameters, no absolute predictions for the interaction strengths are made.

By construction, rather than consequence, the SM describes three generations of quarks and leptons. The Higgs boson is a particle as yet unobserved<sup>1</sup> and gravity is entirely absent. Perhaps of greatest relevance for motivating supersymmetry, and therefore this thesis, the standard model contains a hierarchy problem.

Fermion and gauge boson masses have been discussed by considering the tree level couplings to the Higgs field. The tree level couplings generate a bare Higgs mass,  $m_H^{bare} = \frac{\sqrt{2}\mu h}{2\phi c}$ . To compute the physical mass of the Higgs, we must consider the radiative correction for every particle that couples to the Higgs field. Higher order terms involve particle loops in which the momenta of particles are not fixed, and an integration must be performed over all possible momenta. Such an integration leads to a quadratic divergence, and to fix this, the integral should be cut-off at an energy  $\Lambda_{UV}$ , the scale where new physics is expected. Doing so gives the result for the Higgs mass in Equation 2.13, where  $m_f$  is the mass of the fermion in the loop,  $\lambda_f$  the Higgs to fermion coupling and  $\Lambda_{UV}$  the cut-off energy. Suggesting the scale of new physics to be the Planck scale makes the Higgs mass very heavy.

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^3}(\Lambda_{UV}^2 + \dots) \quad (2.13)$$

Since there are contributions from other loops, including scalars and other particles coupling to the Higgs, these various terms can cancel out the divergent terms. To do so, the parameters must be precisely fine-tuned to get a suitable level of cancellation. Such

---

<sup>1</sup>Recent observations at the LHC have resulted in the discovery of a new boson at 125-126 GeV [9, 10], possibly the Higgs. Further study is needed to confirm the exact particle and couplings.

a fine-tuning is generally considered unphysical, since there is no known mechanism to motivate it. The problem arises since there are 16 orders of magnitude between the EW scale and the reduced Planck scale. This is known as the SM hierarchy problem, and a number of possible solutions have been suggested.

## 2.2 Supersymmetry

Supersymmetry is an extension to the standard model predicting a symmetry between fermions and bosons. This can be expressed by the transformation operator,  $Q$  as in Equation 2.14.

$$Q(Boson) = Fermion \quad (2.14a)$$

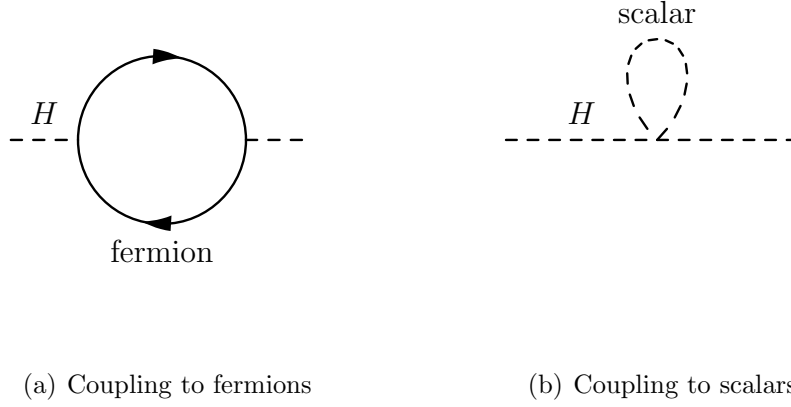
$$Q(Fermion) = Boson \quad (2.14b)$$

Every fermion has a scalar superpartner and every boson a fermionic superpartner. Each SM particle is put in a supermultiplet paired with its own superpartner. The superpartner must have the same gauge group representation, and the supermultiplet must have equal fermionic and bosonic degrees of freedom. Scalar superpartner contributions to the Higgs mass are given in Equation 2.15. Where  $m_s$  is the mass of the relevant scalar and  $\lambda_s$  is the Higgs to scalar coupling.

$$\Delta m_H^2 = \frac{\lambda_s}{8\pi^2}(\Lambda_{UV}^2 + \dots) \quad (2.15)$$

The form of scalar superpartner contributions suggests that if two scalars also exist for every existing fermion, the divergent fermionic terms (the form for which are given in 2.13) would cancel with the scalar terms. If  $|\lambda_f^2|$  is equal to  $\lambda_s$ , the cancellation would be exact, and even at the Planck scale, the Higgs mass would remain at the Electroweak scale.

Figure 2.1 shows the Higgs couplings for scalar and fermion particles, the leading terms for the two diagrams have a relative minus sign between them.



**Figure 2.1:** *MSSM Higgs couplings to fermions and scalar particles.*

Unbroken supersymmetry (SUSY) predicts superpartners with identical mass to their SM counterparts. Since no such particles has yet been discovered, SUSY must be a broken symmetry, but the method of symmetry breaking is unknown. However, to generate the correct SM Higgs vacuum expectation value, the breaking scale must be less than  $\sim 1$  TeV [13].

## 2.3 Minimal Supersymmetric Standard Model

Here, the minimal extension of the SM Higgs sector required for SUSY is discussed, which is the sector assumed in the Minimal Supersymmetric Standard Model (MSSM). The MSSM is specified by the choice of the superpotential given in equation 2.16 [13]. Where the fields that appear are the chiral superfields corresponding to Table 2.3, all colour indices are suppressed.

$$W = y_u^{ij} \bar{u}_i Q_j H_u - y_d^{ij} \bar{d}_i Q_j H_d - y_e^{ij} \bar{e}_i L_j H_d + \mu H_u H_d \quad (2.16)$$

In addition to the Yukawa terms, we require Higgs self couplings which necessitates

the introduction of a mass dimensional parameter  $\mu$ . In order to generate the correct electroweak breaking, we need  $\mu$  to be  $\mathcal{O}(m_W)$ . Additional gauge invariant terms could be added to this potential, but these are terms which violate either baryon or lepton number. By introducing a symmetry called ‘R-parity,’  $R = (-1)^{2s+3B+L}$ , where  $B$  and  $L$  are baryon and lepton numbers respectively, spin ( $s$ ), prevents these terms from being added to the superpotential.

The spin 0 partners of the standard model quarks and leptons are called squarks and sleptons. Both the left-handed and right-handed quarks and leptons have a superpartner. Although these are labelled by the handedness of their standard model counterpart, they are themselves spin-less. The squarks and sleptons are notated by drawing a ‘tilde’ over the top of the symbol for their standard model partners. The spin 1/2 partner of the spin 1 standard model gluon is called a gluino, notated  $\tilde{g}$ , the spin 1/2 partners the  $W^\pm$  and  $W^0$  gauge bosons, ‘winos’ ( $\tilde{W}^\pm$  and  $\tilde{W}^0$ ) and the spin 1/2 counterpart of the  $B$  boson, ‘bino’ ( $\tilde{B}$ ).

Tables 2.3 and 2.4 give the chiral and gauge supermultiplets of the MSSM. The Higgs particles (which have spin 0) must be put into chiral supermultiplets and the spin 1 vector bosons into gauge supermultiplets. Quarks and leptons must be placed in chiral supermultiplets so that the left and right-handed spin 1/2 particles transform differently under Lorentz transformation. Two Higgs doublets are needed to give mass to both the up and down type quarks in the standard model.

The states in the MSSM (Tables 2.3 and 2.4) mix to form the physical mass states. The neutral bino, wino and higgsino ( $\tilde{H}_u^0, \tilde{H}_d^0$ ) states mix to form four neutral particles called neutralinos  $\tilde{\chi}_{1,2,3,4}^0$ . The charged wino and higgsino ( $\tilde{H}_u^+, \tilde{H}_d^-$ ) states mix to form two charged charginos,  $\tilde{\chi}_{1,2}^\pm$ . Neutralinos and charginos are both gauginos. Of the quarks, only the third generation mixes significantly,  $\tilde{t}_L$  and  $\tilde{t}_R$  mixing to form the states  $\tilde{t}_1$  and  $\tilde{t}_2$ , and equivalently,  $\tilde{b}_L$  and  $\tilde{b}_R$  mixing to form the states  $\tilde{b}_1$  and  $\tilde{b}_2$ . Up-down family and charm-strange family mixing is often neglected. Mixing of the leptons in the first

Names	Symbol	$SU_C(3), SU_L(2), U_Y(1)$	spin 0	spin 1/2
squarks, quarks (3 families)	$Q$	$3, 2, \frac{1}{3}$	$(\tilde{q}_L, \tilde{d}_L)$	$(u_L, d_L)$
	$\bar{u}$	$\bar{3}, 1, -\frac{4}{3}$	$\tilde{u}_R$	$\bar{u}_R^\dagger$
	$\bar{d}$	$\bar{3}, 1, \frac{2}{3}$	$\tilde{d}_R$	$\bar{d}_R^\dagger$
sleptons, leptons (3 families)	$L$	$1, 2, -1$	$(\tilde{\nu}_L, \tilde{e}_L)$	$(\nu_L, e_L)$
	$\bar{e}$	$1, 1, 2$	$\tilde{e}_R$	$\bar{e}_R^\dagger$
Higgs, Higgsinos	$H_u$	$1, 2, 1$	$(H_u^+, H_u^0)$	$(\tilde{H}_u^+, \tilde{H}_u^0)$
	$H_d$	$1, 2, -1$	$(H_d^0, H_d^-)$	$(\tilde{H}_d^0, \tilde{H}_d^-)$

**Table 2.3:** *Chiral supermultiplet fields in the MSSM.*

Names	$SU_C(3), SU_L(2), U_Y(1)$	spin 1/2	spin 1
gluinos, gluons	$8, 1, 0$	$\tilde{g}$	$g$
winos, W bosons	$1, 3, 0$	$\tilde{W}^\pm, \tilde{W}^0$	$W^\pm, W^0$
bino, B boson	$1, 1, 0$	$\tilde{B}$	$B$

**Table 2.4:** *Gauge supermultiplet fields in the MSSM.*

two generations is also often neglected, but the left-handed and right-handed staus are commonly taken to mix significantly ( $\tilde{\tau}_L$  and  $\tilde{\tau}_R$  mixing to form  $\tilde{\tau}_1$  and  $\tilde{\tau}_2$ ). The large left-right mixing between the stops, sbottoms and staus means that they are usually denoted by subscripts 1 and 2 rather than L and R. This mixing is summarized below:

neutralinos:	$\tilde{H}_u^0, \tilde{H}_d^0, \tilde{B}^0, \tilde{W}^0 \rightarrow \tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$
charginos:	$\tilde{W}^\pm, \tilde{H}_u^\pm, \tilde{H}_d^\mp \rightarrow: \tilde{\chi}_1^\pm, \tilde{\chi}_2^\pm$
stau	$\tilde{\tau}_L, \tilde{\tau}_R \rightarrow \tilde{\tau}_1, \tilde{\tau}_2$
stop:	$\tilde{t}_L, \tilde{t}_R \rightarrow \tilde{t}_1, \tilde{t}_2$
sbottom:	$\tilde{b}_L, \tilde{b}_R \rightarrow \tilde{b}_1, \tilde{b}_2$

### 2.3.1 MSSM Higgs Sector

In the MSSM Higgs sector, we have seen there are two complex Higgs doublets. The  $H_u$  couples to up type fermions, and the  $H_d$  to down type fermions. After spontaneous symmetry breaking, the doublets acquire vacuum expectation values  $v_u$  and  $v_d$ , with a ratio defined as  $\tan \beta = \frac{v_u}{v_d}$ . Three degrees of freedom give masses to the  $W^\pm$  and  $Z$  bosons, and five physical Higgs states remain. There are two charged scalars,  $H^+$  and  $H^-$ , with equal mass. Two CP even scalars,  $h^0$ , and  $H^0$ , and one CP odd scalar,  $A^0$ .

There are six free parameters that describe the Higgs sector: four Higgs boson masses,  $\tan \beta$ , and the mixing angle  $\alpha$  between the two CP even states. In the MSSM, relationships between the parameters can be determined from the structure of the superpotential, leaving only two free parameters at tree-level, generally taken to be  $m_{A^0}$  and  $\tan \beta$ . All the coupling strengths to fermions and gauge bosons, as well as the masses, are completely defined at tree level with these two parameters. The relationships also result in limits of the Higgs scalars:

$$m_{A^0} \geq m_{h^0}$$

$$m_{h^0} \leq m_Z$$

$$m_{H^0} \geq m_Z$$

$$m_{H^\pm} \geq m_W$$

We have not observed the  $h^0$ , but this does not void the theory as loop corrections can be significant. Couplings from the top quark and stops can push the mass up to  $\sim 135$  GeV. This situation is known as the  $m_h^{\max}$  scenario [14].

The MSSM Higgs couplings differ from a SM Higgs. The ratio of the couplings for the  $W$ ,  $Z$ , up type and down type fermions with respect to the SM Higgs coupling are given in Table 2.5.

While decays into  $ZZ$  or  $WW$  are dominant in the SM for Higgs boson masses above  $m_H \sim 160$  GeV, in the MSSM these decay modes are either suppressed by  $\cos(\alpha - \beta)$  in

Higgs boson	up type fermions	down type fermions	$W/Z$
$A^0$	$1/\tan \beta$	$\tan \beta$	0
$H^0$	$\sin \alpha / \sin \beta$	$\cos \alpha / \cos \beta$	$\cos(\alpha - \beta)$
$h^0$	$\cos \alpha / \sin \beta$	$-\sin \alpha / \cos \beta$	$\sin(\alpha - \beta)$

**Table 2.5:** *Ratio of the Higgs couplings in the MSSM to the SM for up and down type fermions, and  $W$  and  $Z$  gauge bosons.*

the case of the  $H^0$ , or even absent as in the case of the  $A^0$ .

# Chapter 3

## LHC

To understand the mechanism behind electroweak symmetry breaking, we need to be able to conduct experiments beyond the electroweak scale. This energy scale,  $\mathcal{O}(1 \text{ TeV})$ , has recently become open to observation with the Large Hadron Collider.<sup>1</sup> High energy particle colliders are designed to produce inelastic collisions between relativistic particles. In circular colliders, such as the Large Hadron Collider (LHC), two beams of particles travelling in opposite directions collide head on. There is zero net momentum before and after the collision, and the center-of-mass energy is given by twice the beam energy. These conditions define the kinematic phase space available to the particles produced in the collision.

The LHC is the world's highest energy particle collider, capable of a center-of-mass energy of 14 TeV [15]. The machine collides two beams of protons,<sup>2</sup> and since protons are composite particles, a collision occurs between two packets of loosely bundled quarks and gluons. As the momentum of a proton is distributed amongst the constituent partons (described further in Appendix C), the effective energy available to produce new particles

---

<sup>1</sup>High energy events occur naturally and have been observed with cosmic ray particles [16]. However the rate of such events is too small to study electroweak symmetry breaking. (At 1 TeV, the rate is 1 particle per  $\text{m}^2/\text{s}$  [17]).

<sup>2</sup>The LHC can also collide heavy ions, using lead with an energy of 2.8 TeV per nucleon, up to an instantaneous luminosity of  $10^{27} \text{ cm}^{-2}\text{s}^{-1}$  [15].

is lower than 14 TeV.

In addition to the center-of-mass energy, the rate of collisions produced in an accelerator are of importance, particularly if we want to study rare processes. The luminosity of a machine, and the cross-section for an interaction determine the rate of collisions, as given in Equation 3.1.

$$\text{Number of interactions per second} = \mathcal{L} \times \sigma_{\text{interaction}} \quad (3.1)$$

Where the luminosity is denoted  $\mathcal{L}$  and has units  $\text{cm}^{-2}\text{s}^{-1}$ . In an accelerator, beams of particles are separated into bunches of  $N$  particles. The geometrical overlap of bunches as the beams cross is described by  $\sigma_x$  and  $\sigma_y$ , the transverse components of the collision area. For identical bunches colliding with a frequency given by  $f$ , the luminosity is defined as:

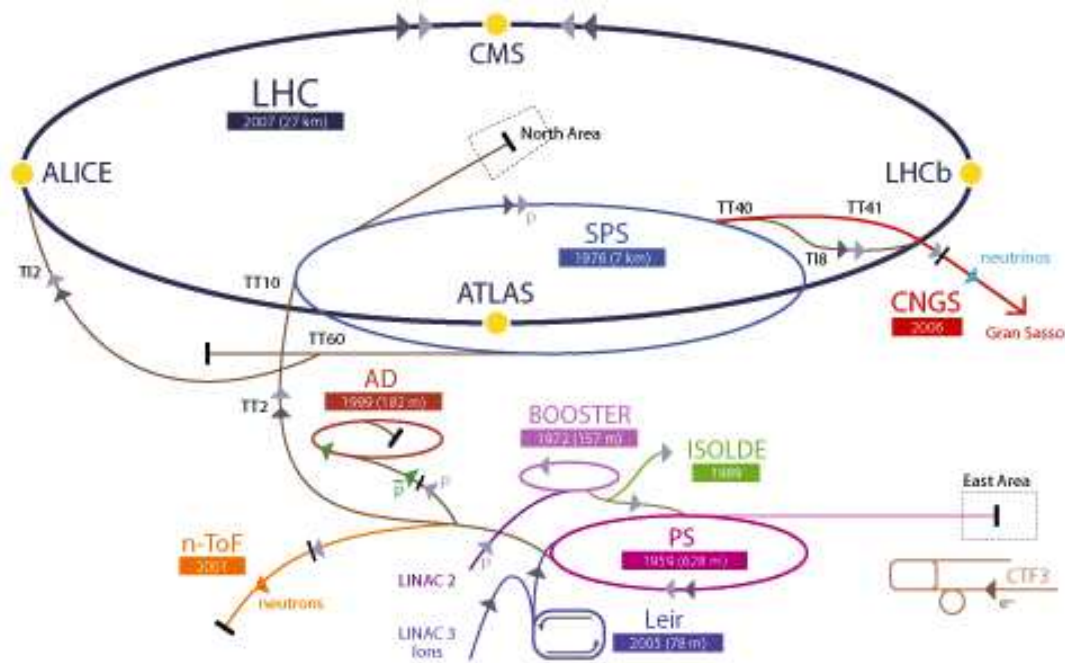
$$\mathcal{L} = f \frac{N^2}{4\pi\sigma_x\sigma_y} \quad (3.2)$$

The LHC is designed to reach a luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$  [15]. Bunches of protons will collide every 25 ns, a frequency of 40 MHz.

The main components of the accelerator are radio-frequency (RF) cavities and superconducting magnets. The RF cavities group protons together in up to 2808 bunches, and accelerate these bunches to the collision momentum. The cavities work by generating an oscillating voltage along the beamline. They use a standing wave whose frequency is set such that protons get bunched together and each bunch is given an accelerating push as they pass through the cavity. Two different types of magnets are used for specific operations on the beams. A total of 1232 dipole magnets curve the beams into their circular trajectories, and 392 quadrupole magnets keep the protons focused. The superconducting magnetic dipoles stretch the limits of current technology. In order to guide the beam they must produce a magnetic field of 8.33 T, and to do so, they operate in superfluid helium at 1.9 K [15].

The LHC is located at CERN, across the borders of France and Switzerland. The ac-

celerator is housed in the tunnel previously used by the Large Electron Positron Collider (LEP). The circular tunnel has a circumference of 27 km and is underground at a depth varying between 50 and 175 meters [15]. There are four points in the LHC, where the beams intersect for collisions. A detector, delivering data to an experiment, surrounds each collision point. ATLAS, an acronym for ‘**A** **T**oroidal **L**HC **A**pparatu**S**,’ and CMS, ‘**C**ompact **M**uon **S**olenoid;’ are multipurpose detectors, designed to collect high momentum events in the search for Higgs bosons and other phenomena beyond the Standard Model. LHCb, short for ‘**L**HC **b**eauty’ is (unsurprisingly) designed for  $b$ -physics, in order to examine discrepancies between the production of matter and antimatter. **ALICE**, shortened from ‘**A** **L**arge **I**on **C**ollider **E**xperiment,’ is built to study the formation of a quark-gluon plasma in heavy ion collisions. A schematic of the LHC experiments is given in Figure 3.1.



**Figure 3.1:** Layout of the CERN accelerator complex and the LHC experiments.

Protons are injected into the LHC with a momentum of 450 GeV and then accelerated to 3.5 TeV. The initial energy is achieved by a series of accelerators, also shown in

Figure 3.1. First the protons reach an energy of 50 MeV in the linear accelerator LINAC2. They are then fed into the Proton Synchrotron Booster and accelerated to 1.4 GeV. From here they enter the Proton Synchrotron (PS), and are accelerated to 26 GeV. Finally, the protons reach the Super Proton Synchrotron (SPS), achieving an energy of 450 GeV with which they enter the LHC. The entire injection chain and proton acceleration in the LHC takes  $\sim 20$  minutes.

At the time of writing this thesis, the maximum proton beam energy attained by the LHC was 3.5 TeV. Collisions were produced every 50 ns, with an intensity of  $1.51 \times 10^{11}$  protons per bunch, giving a peak luminosity of  $3.6 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ . Even at the 7 TeV centre-of-mass energy and lower luminosity, the LHC has surpassed all previous accelerators, unlocking energy regions with the potential for new physics.

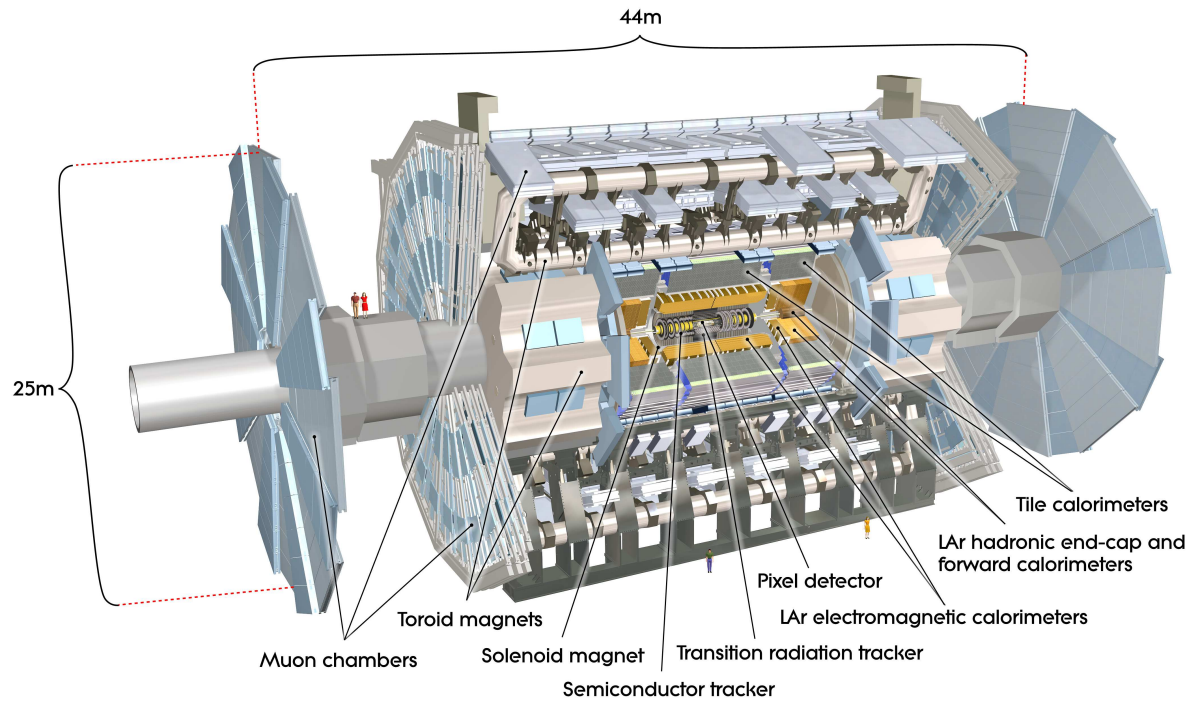
# Chapter 4

## ATLAS Detector

The ATLAS detector is designed to measure many different physics signatures while operating in the LHC's high luminosity environment. It is the largest volume detector ever built, with dimensions  $46 \times 25 \times 25$  m. The 7000 ton apparatus is housed in an underground cavern at a depth of 100 m, surrounding an LHC beam intersection point [18]. The detector, shown in Figure 4.1, measures the 4-vectors of particles produced in high energy collisions using a combination of sub-detectors.

Each sub-detector is designed to exploit the energy loss mechanisms of specific particles, to discern their nature and measure their kinematics. The Inner Detector is closest to the ATLAS collision point. It is used to measure the momenta and reconstruct the trajectory of charged particles. The calorimeters measure the energy of most traversing particles, by causing them to shower through electromagnetic and strong interactions. Muons exit the calorimeters, only losing a small amount of energy by ionization, their momenta and trajectories are reconstructed by the Muon Spectrometer.

The momentum of charged particles in both the Inner Detector and Muon Spectrometer are deduced from their trajectories in a magnetic field. A superconducting solenoidal magnet surrounds the Inner Detector, generating a field with the strength 2.6 T. The Muon Spectrometer incorporates a superconducting toroidal magnetic system, formed



**Figure 4.1:** *The ATLAS detector with magnets, sub-detectors and overall dimensions labelled.*

*The width of the detector, not shown, measures the same as the height: 25 m.*

by an arrangement of three sets of eight coils, the field strength is 4.1 T.

Analysis with ATLAS uses a reference frame centred on the nominal interaction point; this is defined as the origin of the coordinate system. The beam line defines the  $z$  direction, and the  $xy$  plane is transverse to the beams. The positive  $x$  axis points from the interaction point to the centre of the LHC, and the positive  $y$  direction points upwards. The detector is symmetric around the  $z$  axis, making a cylindrical coordinate system useful. The azimuthal angle,  $\phi$ , sweeps around the  $z$  axis. The polar angle,  $\theta$ , extends from the beam line; this is usually expressed as the Lorentz invariant variable of pseudorapidity,  $\eta = -\ln(\tan \frac{\theta}{2})$ .

Charged particles are bent in the  $r\phi$  plane when traversing the Inner Detector, and in the  $rz$  plane through the Muon Spectrometer. All transverse quantities are defined as projected values on the transverse ( $r\phi$  and  $xy$ ) plane. For example, the transverse momentum of a measured particle,  $p_T$ , is determined by the  $xy$  components of the particle's momentum vector,  $p$ .

## 4.1 Inner Detector

The Inner Detector (ID) is the innermost sub-detector of ATLAS, positioned closest to the LHC beam line. It is designed to measure charged track momenta, primary and secondary vertices; with minimal disturbance to particle energy. This is achieved with two types of technology: silicon semiconductor detectors, and proportional drift tubes.

The semiconductor detectors are thin wafers of silicon, doped to form diodes. Electrodes applied to the top and bottom of the wafers with a reverse bias configuration extend the depletion zone, which is free of charge carriers, to cover the full thickness of the silicon. A charged particle traversing the semiconductor ionizes atoms in the silicon bulk, forming electron-hole pairs. The electric field generated by the electrodes forces the charges away before they recombine. This motion of charge results in a current pulse

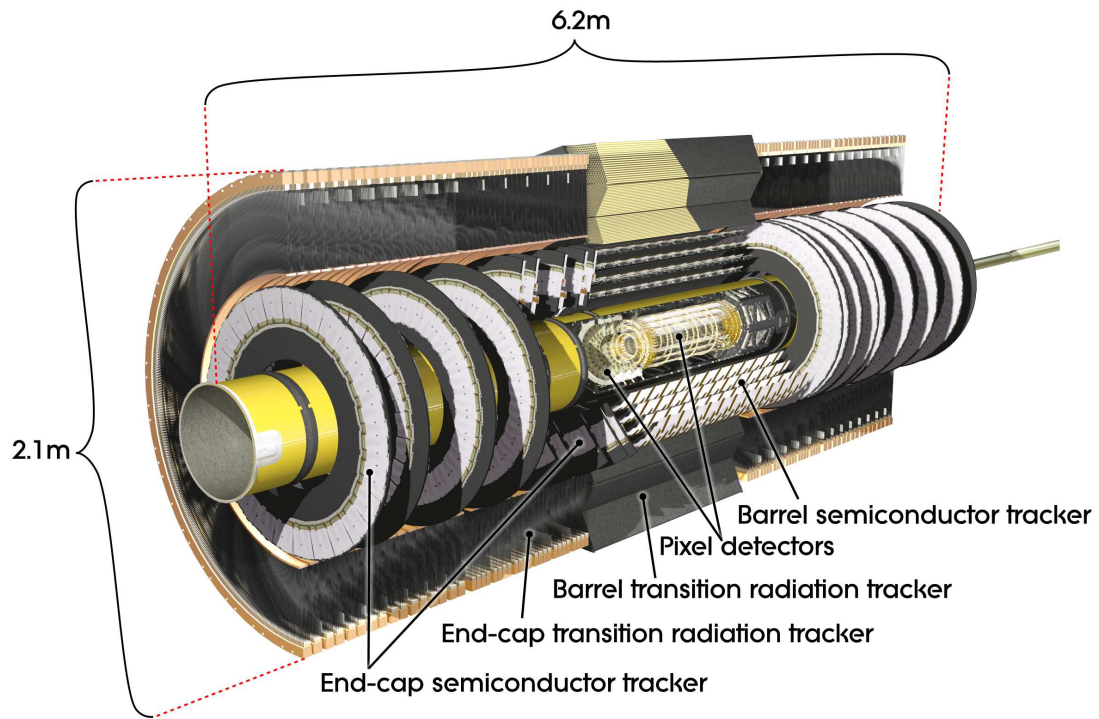
proportional to the energy lost by the incident particle. In ATLAS the silicon detectors register a binary ‘hit’ if a particle passes through the sensitive silicon region.

Proportional drift tubes are cylindrical proportional chambers. A voltage applied to an anode wire centred in a cathode tube creates an electric field that rises as  $1/r$  towards the wire. The tube contains a gas that is ionized by a crossing charged particle. Ions drift towards the tube wall, and free electrons accelerate towards the anode. The motion of charge produces the detected signal. With a sufficient voltage difference, the electric field close to the anode causes secondary ionization, enhancing the signal. This amplification is approximately constant and depends on the electric field and ionized gas.

The ID is divided into three parts: the pixel detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). The first two of these are silicon semiconductors, and the last is constructed from proportional drift tubes. All three sub-detectors are shown in Figure 4.2. The ID is immersed in an axial magnetic field of 2 T generated by a solenoidal magnet. The magnetic field curves the trajectory of a charged particle, allowing a measurement of the particle momentum, which is proportional to the particle bending radius.

An important aspect of silicon detectors is the occupancy, defined as the fraction of channels with a hit in the detector at any given moment. High occupancy can become a problem in a high particle flux environment, such as that close to the interaction point in ATLAS. One solution to this problem is to reduce the area of the sensitive region, which implies an increase in the total number of channels, as in the ATLAS pixel detector. Further away, where particles are less concentrated, the occupancy is reduced and larger detectors are an option.

The pixel detector is nearest to the collision point, and subject to intense radiation and particle flux. To support coverage along  $\eta$ , the detector is divided into a barrel component to measure particles produced centrally (with a large transverse component), and two end-caps on either side of the barrel for particles produced in the forward direction. The



**Figure 4.2:** Cross sectional view of the ATLAS inner detector. The pixel sub-detector is closest to the beam line, followed by the SCT. The TRT, the outermost layer, covers the largest volume.

pixel detector measures hits for tracks within  $|\eta| < 2.5$ , with full coverage in  $\phi$ . The barrel is constructed from three cylindrical layers and each end-cap is made of three disks. The main components of the detector are 1744 identical modules, arranged along the barrel cylinders and end-cap disks. The sensor is made of a silicon semiconductor junction  $256 \pm 3 \mu\text{m}$  thick. Each module contains  $47 \times 232$  pixels, with a typical pixel size of  $50 \times 400 \mu\text{m}^2$ , corresponding to 80.4 million channels for the entire detector. The longer sensing edge of the pixels is laid along the  $z$  direction in the barrel, and the radial direction in the end-caps. Combining measurements from three pixel hits, the sub-detector has an intrinsic spatial resolution of  $10 \mu\text{m}$  in the  $r\phi$  direction, and  $115 \mu\text{m}$  in the  $z$  ( $r$ ) direction for the barrel (end-caps).

The second ID sub-detector is the SCT. The technology is similar to that of the pixels, with differences in the size and shape of the sensitive regions. There are 4088 SCT modules spread across the barrel and end-caps. These are arranged on four cylinders in the barrel region, and on nine disks in each end-cap. Modules consist of four silicon strip sensors. Two sensors are chained together, and a second pair of identical sensors are glued back to back with the first, at a stereo angle of  $40 \text{ mrad}$  to provide space points. Each module is  $12 \text{ cm}$  in length, and has a total of 768 strips. Strips in the barrel are approximately parallel to the beam line, with a pitch of  $80 \mu\text{m}$ . In the end-caps the strip direction is radial and of variable pitch, averaging  $80 \mu\text{m}$ . In total the SCT readout consists of  $\sim 6.3$  million channels. The entire SCT has an intrinsic spatial resolution of  $17 \mu\text{m}$  in the  $r\phi$  direction, and  $580 \mu\text{m}$  in the  $z$  ( $r$ ) direction for the barrel (end-caps).

Currently both the pixels and SCT operate with an applied voltage of  $\sim 150 \text{ V}$ . This will be increased to maintain the depletion zones as more data is collected, and radiation damage is experienced. The signal readout from silicon detectors is very fast, approximately  $20 \text{ ns}$ .

The largest and outermost layer of the ID is the TRT. The TRT consists of 298 304 proportional drift tubes,  $4 \text{ mm}$  in diameter, readout by  $\sim 351\,000$  channels. The cathode

tubes are made from carbon polyimide, with layered films of aluminium and kapton. The anode wire at the centre of each TRT tube is made from gold-plated tungsten, and measures  $30\text{ }\mu\text{m}$  in diameter. The drift tubes operates with a gas mixture of 70% Xe, 27%  $\text{CO}_2$ , and 3%  $\text{O}_2$ . A voltage of 1530 V is applied between the central anode and the cathode tube walls to create an electric field between them. When a charged particle passes through a tube chamber it ionizes the gas molecules inside along its trajectory. Close to the anode, secondary ionization creates a gas gain of  $2.5 \times 10^4$ . The collection of these charged particles is registered as a hit.

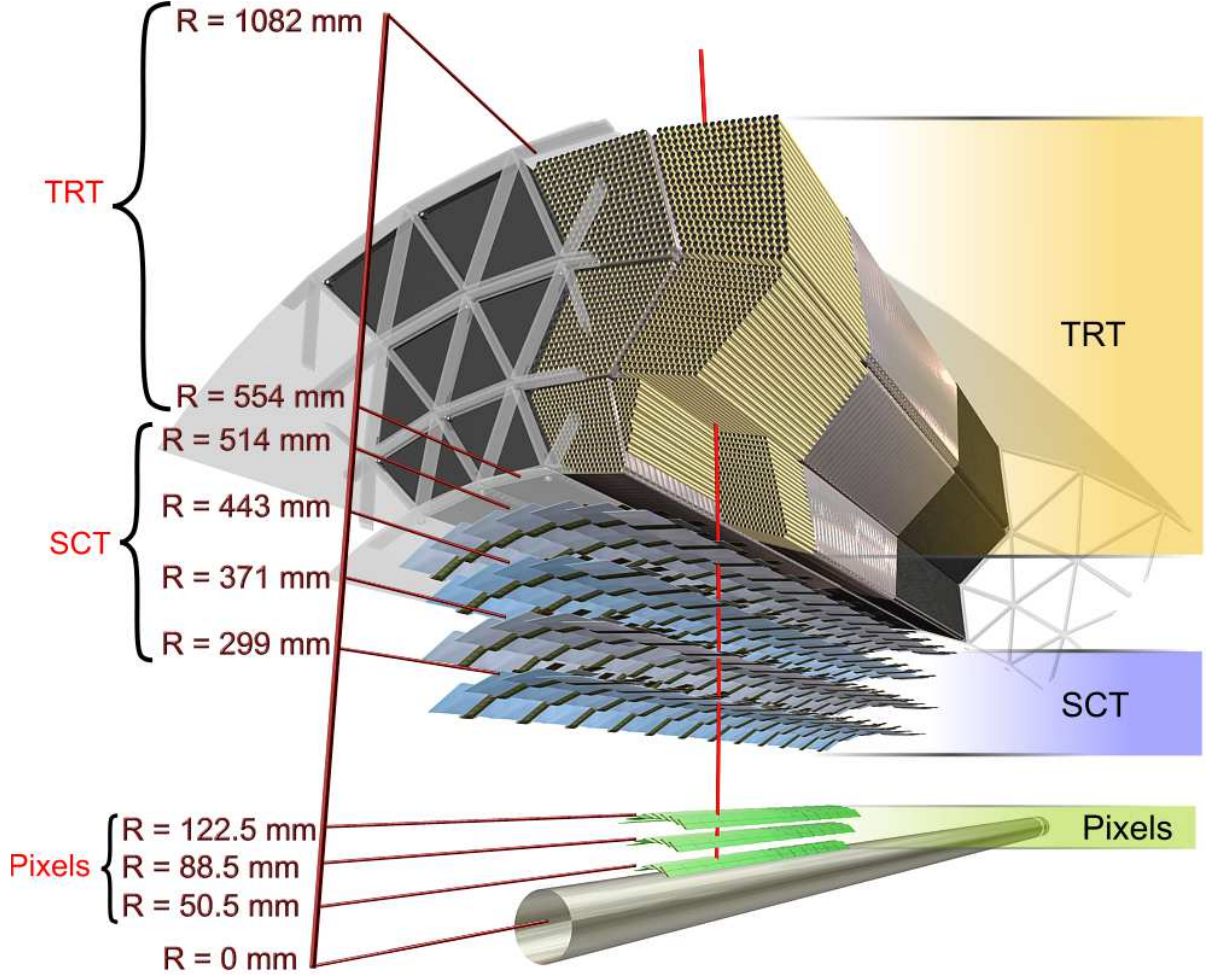
The straws in the barrel region are arranged in three cylindrical layers and orientated in the  $z$  direction. The straws in the end-caps are arranged in wheel structures, radially oriented. This sub-detector has an intrinsic accuracy of  $130\text{ }\mu\text{m}$  on the position of a track in the  $r$ - $\phi$  direction. The maximum drift time in the TRT is 48 ns, for charged particles furthest away from readout areas.

The TRT is also used for electron identification. All drift tubes are interleaved with layers of polypropylene fibres and foils. A charged particle that passes through the boundary region between materials with a different refraction index emits X-ray radiation with an intensity proportional to its relativistic  $\gamma$  factor. The TRT operates with hits based on two threshold levels. Lighter electrons, with a larger value of  $\gamma$  will emit a higher number of X-rays, producing more high threshold hits, as opposed to heavier charged pions, which usually only meet the lower threshold requirements.

Individually the silicon detectors give better spatial resolution than the TRT. However, silicon is denser (interfering with particle energy) and more costly than the proportional drift tubes. The TRT compensates for its lower resolution by providing a larger number of track hits over a greater distance. An average track with  $p_T > 0.5\text{ GeV}$  will have 36 ID hits, 30 of which come from the TRT. Figure 4.3 shows a schematic image of a charged particle traversing the entire ID. Including all sub-detectors, the ID gives a

momentum resolution of:

$$\frac{\sigma_{p_T}}{p_T} = 0.05\% p_T \oplus 1\%$$



**Figure 4.3:** Schematic diagram of a charged particle (with  $p_T > 0.5$  GeV) traversing the Inner Detector barrel. The particle will produce a maximum of seven silicon detector hits, and approximately 30 TRT hits.

## 4.2 Calorimeters

The purpose of the calorimeters is to measure the energy and position of particles. To do this, particles must shower within calorimeters, and this dictates the material and size requirements. The calorimeter system is divided into two parts: the electromagnetic (EM) calorimeter, and the hadronic (HAD) calorimeter. The first is dedicated to the measurement of photons and electrons, the latter to measurement of hadrons.

Electron and photon energy loss is best discussed in terms of radiation lengths,  $X_0$ . One  $X_0$  is the mean distance over which a high energy electron loses all but  $1/e$  of its energy by bremsstrahlung, and  $7/9$  of the mean free path for pair production by a high energy photon [12]. The unit of measurement is  $\text{g}\cdot\text{cm}^{-2}$ , and therefore in practice, the number of radiation lengths a particle encounters depends on the density of the medium, and the length of the material it passes through. Travelling through matter a high energy electron will radiate a photon, which in turn produces an electron-positron pair. This process is repeated, doubling the number of particles after each  $X_0$ , until the electrons and photons reach a critical energy below which no more particles are produced. Once the electron energy is below this threshold, of approximately 10 MeV, it loses energy by ionizing the medium. Low energy photons ( $\sim 1.02 \text{ MeV} \approx 2m_e$ ) lose energy through Compton scattering and the photoelectric effect.

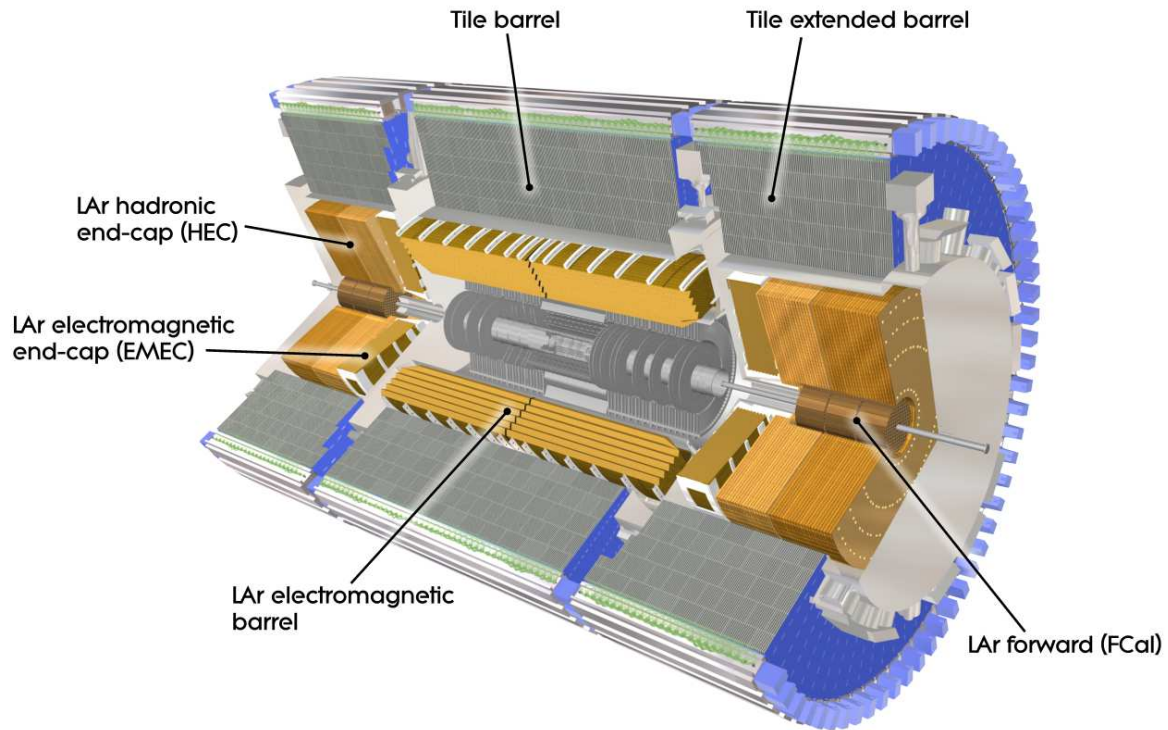
Hadronic particle energy loss is measured in terms of interaction lengths,  $\lambda$ , with units  $\text{g}\cdot\text{cm}^{-2}$ . An interaction length is the mean path length for a high energy baryon or meson to undergo a nuclear interaction [12]. It is longer than  $X_0$  because the cross-section of nuclear interactions is smaller than electromagnetic interactions. Energy loss through nuclear interactions are more complex and more difficult to measure accurately than electromagnetic processes. A large, and fluctuating, amount of energy lost by incident hadrons is converted into the excitation and break up of nuclei. However, only a small fraction of this energy is detectable, resulting in variations of the measured energy. Secondary particles produced in hadronic interactions are mostly pions and

kaons. Neutral pions decay to two photons, which shower electromagnetically. A fraction of the kaons and charged pions will decay to muons and neutrinos, and are not measured as part of hadronic interactions. Any emitted neutrons or protons interact strongly, with protons losing energy by ionization.

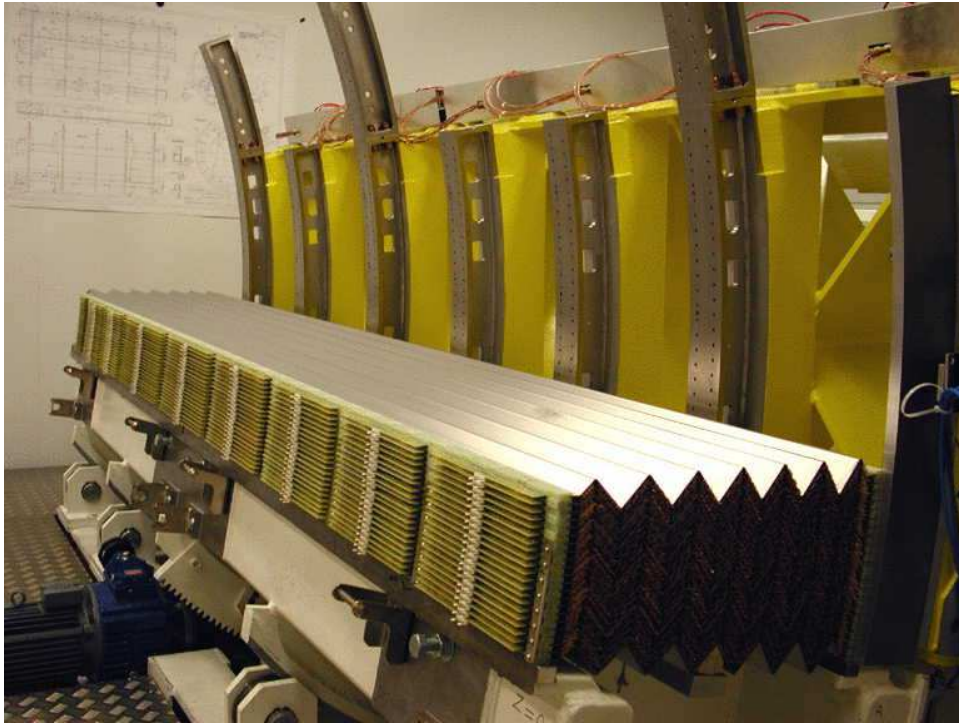
The EM calorimeter ranges from 22 to 33  $X_0$  in the barrel, and from 24 to 38  $X_0$  in the end-caps. The HAD calorimeter is  $\sim 11 \lambda$ , in the barrel and end-caps, including  $1.3 \lambda$  from the outer support material [18]. Both the EM and HAD calorimeters are sampling calorimeters, meaning the functions of particle absorption and active signal readout are separated, and the sensitive medium only samples shower energy. Each calorimeter system provides full coverage in azimuthal angle; this is essential for measuring missing energy from un-detected particles, such as neutrinos. The ATLAS calorimeters cover the range  $|\eta| < 4.9$ , using a variety of materials suited to the radiation environment. Figure 4.4 shows the layout of the ATLAS calorimeters.

The EM calorimeter is a lead and liquid Argon (LAr) detector. The dense absorber material, lead, contributes most of the radiation lengths, where one  $X_0$  corresponds to a 5.6mm thickness of lead [19]. The active material, LAr, produces an output signal from ionization that is proportional to the energy of the showering electron or photon. The detection mechanism is similar to gas ionization chambers, with the advantage of a higher density medium that has greater absorption. For full azimuthal coverage the EM calorimeter is designed with an accordion geometry, the structure is shown in Figure 4.5. Sheets of lead absorber are folded, and placed so that an incident particle from the interaction point traverses multiple absorber layers from a single sheet. The accordion waves vary with distance from the interaction point, to keep the LAr drift gap constant at 2.1 mm. This is necessary for uniform detector performance. Readout electrodes, made up of three copper layers, are positioned in the gaps between the absorber. The two outer copper layers form a voltage difference across the drift gap, and the middle layer serves to readout the signal. Segmentation in the  $\eta$  direction is etched onto the

readout surface, and  $\phi$  segmentation is obtained by grouping together the appropriate number of electrode layers.



**Figure 4.4:** *Cross sectional view of the ATLAS calorimeters, which surround the Inner Detector. All calorimeters, other than the hadronic Tile, use liquid argon based technology and must be contained in cryostats.*



**Figure 4.5:** *Barrel electromagnetic calorimeter in construction. The accordion waves are clearly visible.*

The EM calorimeter is divided into a barrel section, covering  $|\eta| < 1.47$ , and two end-caps in the range  $1.37 < |\eta| < 3.2$ . Each of these is housed in an individual cryostat to keep the LAr at a temperature of 88 K [18]. The barrel cryostat shares a vessel wall with the solenoidal magnet that surrounds the ID, this minimizes the amount of material a particle crosses before the calorimeter. The barrel calorimeter is made of two identical cylinders, with a gap between them of 4 mm at  $z = 0$ . Each end-cap calorimeter is divided into two coaxial wheels, an outer wheel that covers  $1.37 < |\eta| < 2.5$ , and an inner wheel from  $2.5 < |\eta| < 3.2$ . In the region devoted to precision physics,  $|\eta| < 2.5$ , the EM calorimeter is segmented into three longitudinal layers. The first layer is finely grained in the  $\eta$  direction, with values ranging from  $\Delta\eta = 0.0031$  for  $|\eta| < 1.4$  to  $\Delta\eta = 0.025$  for  $2.4 < |\eta| < 2.5$ . Azimuthal readout division is  $\Delta\phi = 0.1$  radians for all  $\eta$ , the unit radians will be omitted henceforth. The granular  $\eta$  readout is designed to resolve the two photons in  $\pi^0 \rightarrow \gamma\gamma$  decays, this helps reduce background for single photon signals. The second layer is the deepest,  $\sim 16 X_0$ , and captures the majority of an electromagnetic shower, the average cell size is  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ . Any residual showering is contained in the third and coarsest layer, with a cell size  $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$ . In the range  $|\eta| < 1.8$ , there is an additional presampler layer, 11mm of LAr with no absorber, used to correct for energy lost by electrons and photons before reaching the calorimeter. The transition between the barrel and end-cap EM calorimeters,  $1.37 < |\eta| < 1.52$ , gives poorer performance because of support structure and readout material in the area. The resulting energy resolution from the EM calorimeter is:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$$

The hadronic calorimeter is made of different materials in the central and forward regions. The Tile calorimeter covers the central region,  $|\eta| < 1.7$ , and is divided into a barrel component ( $|\eta| < 1.0$ ) and two extended barrels ( $0.8 < |\eta| < 1.7$ ). It sits directly outside the EM barrel, as seen in Figure 4.4. The Tile is a sampling calorimeter that uses a steel absorber, and active scintillating tiles. Such scintillation detectors work by

converting the excitation caused by incident particle energy loss into light. The light is transferred via a fibre light guide into a photomultiplier. In the Tile calorimeter, the sensitive elements are 3 mm thick polystyrene tiles placed in planes perpendicular to the beam line. Scintillation photons from these tiles causes the emission of wavelength shifted photons in fibres lining the sides of the tiles; these secondary photons are then detected by photomultiplier tubes located along the outer edge of the calorimeter. Shifting the wavelength of the primary photons into secondary photons with a wavelength of 400 to 450 nm traps light in the fibres and generates a more efficient readout from the photomultiplier tubes. The tiles are separated by 14 mm thick steel absorber plates; one  $\lambda$  in iron corresponds to a length of  $\sim 17$  cm [19]. Fibres are combined at the photomultipliers to subdivide the calorimeter into cells. The calorimeter is divided longitudinally into three layers  $1.5 \lambda$ ,  $4.1 \lambda$  and  $1.8 \lambda$  thick at  $\eta = 0$ . The cells have dimension of  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  in the first two layers, and  $0.2 \times 0.1$  in the final layer.

In the region  $1.5 < |\eta| < 3.1$ , hadronic energy measurement is performed with the Hadronic end-cap calorimeter (HEC). This consists of two independent wheels in two end-caps, located directly behind the EM end-cap calorimeters, sharing the same cryostats. The HEC uses similar technology to the EM calorimeter but substitutes copper, which has  $\lambda \sim 15$  cm [19], for the absorber material. Each wheel is divided into two longitudinal segments, for a total of four layers per end-cap. The wheels closest to the interaction point are built with 25 mm parallel copper plate absorbers, while those further away use 50 mm copper plates. The absorbers are interleaved with 8.5 mm LAr gaps, the active medium for this sampling calorimeter. The energy resolution of the HAD barrel and end-cap calorimeters is:

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$$

The Forward Calorimeter (FCal) covers  $3.1 < |\eta| < 4.9$  and is also a LAr based detector. It is integrated into the end-cap cryostats, as this provides clear benefits in terms of uniformity of the calorimetric coverage as well as reduced radiation background

levels in the muon spectrometer. The FCal is approximately  $10 \lambda$  deep, and consists of three modules in each end-cap. The first module uses a copper absorber, for EM shower containment. The other two modules use a tungsten absorber, where one  $\lambda \sim 10$  cm [19], to measure the energy of hadronic interactions. The location of the FCal is very close to the beam pipe, and the particle flux is high. Therefore the electrode structure is different from the accordion and plate geometry of the EM and HEC calorimeters. The structure consists of concentric rods and tubes parallel to the beam axis. The LAr in the gap between the rod and the tube is the sensitive medium. The energy resolution of the hadronic FCal is:

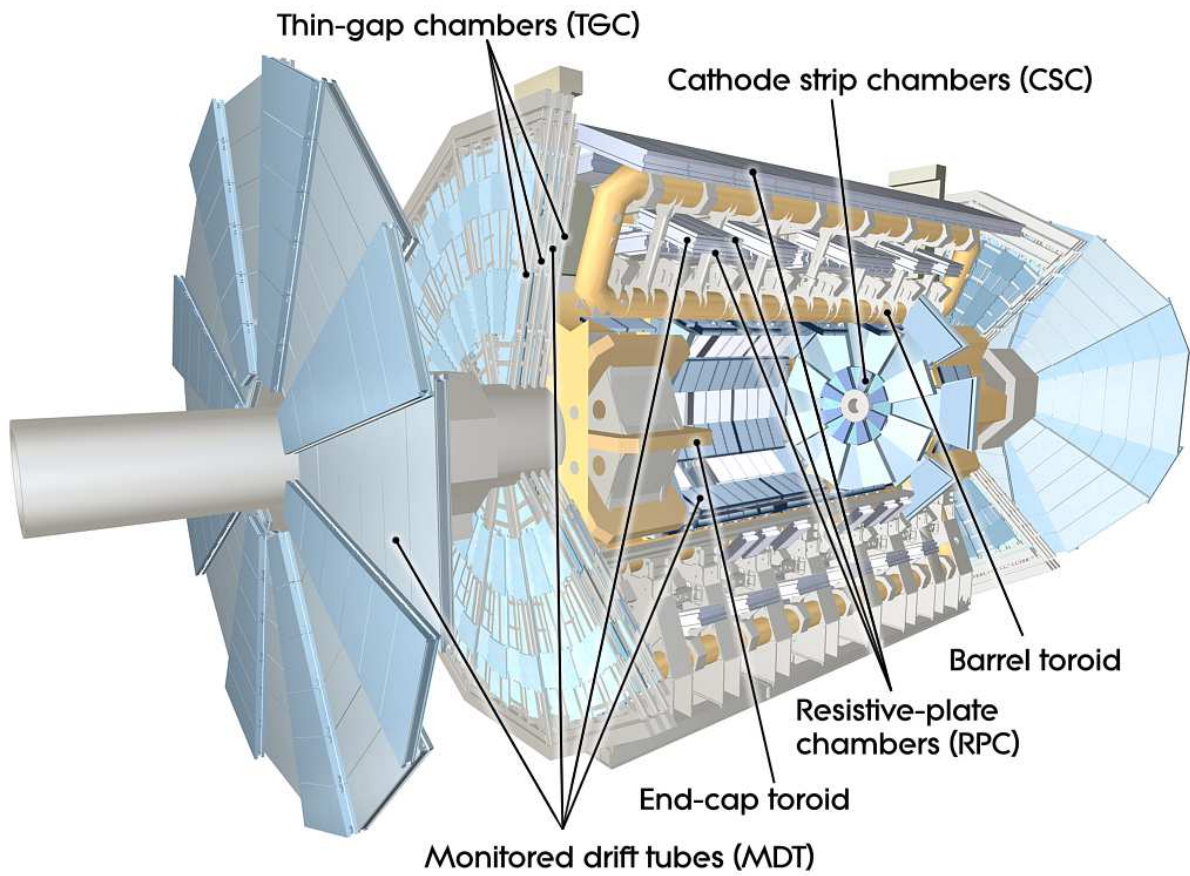
$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$$

All of the LAr calorimeters have a signal drift time and readout between 400 to 500 ns, but the signal is shaped for a faster response. The signal propagation for Tile calorimeter takes approximately 50 ns.

### 4.3 Muon Spectrometer

The layout of the muon spectrometer is shown in Figure 4.6. The spectrometer is used to trigger on muons and measure their momenta, muons being the only charged particles that can traverse the entire calorimeter system. Measurement is based on the magnetic deflection of muon trajectories. The muon sub-detectors are immersed in a magnetic field generated by three toroidal magnets. A barrel toroid serves the range  $|\eta| < 1.4$ , and two smaller end-cap toroids tilted by  $22.5^\circ$  with respect to the beam line are positioned in the regions  $1.6 < |\eta| < 2.7$ . A combination of the two form the magnetic field in the transition region  $1.4 < |\eta| < 1.6$ . This configuration gives a magnetic field that is mostly orthogonal to the muon trajectories.

Muon tracking is performed over most of the  $\eta$  range with Monitored Drift Tubes (MDTs). At large  $|\eta|$ , close to the beam line and interaction point, Cathode Strip



**Figure 4.6:** *ATLAS muon spectrometer, with all the sub-detectors visible and labelled. As these sub-detectors are positioned far from the interaction point, in order to envelope a collision they must cover a large volume.*

Chambers (CSCs) with higher granularity are used. The CSCs are designed to measure multiple tracks in a high particle flux region. Individual detector components are aligned relative to one another using an optical alignment system that is integrated into the supporting structure.

The MDTs are cylindrical ionization chambers operated at a pressure of 3 bar, limiting gas diffusion, covering  $|\eta| < 2.7$ . Tubes are bundled together into chambers, each containing 3 to 8 layers of drift tubes. The cathode tube is constructed from aluminum and has a diameter of 30 mm. The anode is a 50  $\mu\text{m}$  thick tungsten-rhenium wire. Each tube contains a gas mixture of Ar (93%) and  $\text{CO}_2$  (7%). The length of the tubes varies with location between 1 m to 6 m. The wire sag must be contained to within 10  $\mu\text{m}$ , and this provides a spatial resolution of 80  $\mu\text{m}$  per tube, or 35  $\mu\text{m}$  per chamber in the  $z$  direction. The layout is such that the chamber size, including the number of tubes contained, increases proportional to the distance from the interaction point. The operating voltage of 3080 V gives a gas gain of  $\sim 2 \times 10^4$  per tube. Since the chambers are located far from the interaction point they cover a large spatial area, this results in a signal that takes 700 ns to propagate and be read out.

The CSCs are positioned in the first layer of tracking in the forward region:  $2 < |\eta| < 2.7$ , replacing a layer of MDTs. The CSCs are better suited for the high rate of particle flux and muon track density; CSCs can function effectively up to a counting rate of 1000 Hz/cm<sup>2</sup>, while MDTs are limited to 150 Hz/cm<sup>2</sup>. The CSCs are multi-wire proportional chambers, operating in a gas mixture of Ar (80%) and  $\text{CO}_2$  (20%). Each anode wire is 30  $\mu\text{m}$  thick and positioned in the radial direction, with an applied voltage of 1900 V. The cathode planes are segmented in strips, orthogonal to one another on each plane; this allows two coordinates to be measured. The track position is obtained by interpolating between charges induced on neighbouring cathode strips. Those cathodes segmented perpendicular to the wire measure the  $z$  coordinate, and those parallel measure the transverse component. There is a gas gain of  $6 \times 10^4$ , caused by an avalanche induced

close to the anodes and spread out over 3 to 5 cathode readout strips. Each CSC chamber is constructed from four CSC planes, giving four independent measurements in  $\eta$  and  $\phi$  along each track path. Each plane has a readout time of  $\sim 7$  ns. Combined, the CSC chambers give an intrinsic accuracy of  $40\ \mu\text{m}$  in the radial direction and  $5\ \text{mm}$  in the  $\phi$  direction for a traversing muon.

The triggering components of the muon spectrometer cover  $|\eta| < 2.4$ . Resistive Plate Chambers (RPC) are used in the barrel and Thin Gap Chambers (TGC) in the end-caps. The trigger chambers for the muon spectrometer serve a two fold purpose: to provide well defined transverse momentum thresholds for triggering, and to measure the muon coordinate in the direction orthogonal to that determined by the MDT and CSC tracking chambers.

The RPCs are positioned on three cylinders concentric with the beam axis, at radii of about 5, 7.5, and 10 m, in the range  $|\eta| < 1.05$ . A cylinder holds two independent chambers, each measuring  $\eta$  and  $\phi$ . A track through all three stations delivers six measurements. Each chamber is constructed from a 3 mm plate of rigid polystyrene immersed in a gas mixture of  $\text{C}_2\text{H}_2\text{F}_4/\text{Iso-C}_4\text{H}_{10}/\text{SF}_6$  (94.7%/5%/0.3%). Measuring strips 25-35 mm wide and  $17\ \mu\text{m}$  thick made of copper are glued on top, separated by 2 mm, with a readout time of 7 ns. The two inner chambers are used for low  $p_T$  muon triggers, from 6 to 9 GeV. The outer chamber triggers on higher  $p_T$  muons, 9 to 35 GeV. The RPC gives an intrinsic spatial resolution of 10 mm in both the  $z$  and radial directions and 5 mm in the  $\phi$  direction.

The TGCs are multi-wire proportional chambers that give triggering capability, and also a measurement in the azimuthal coordinate to compliment the MDT detectors, within  $1.05 < |\eta| < 2.4$ . They are positioned in seven layers in the middle of the muon end-caps. The high quenching gas used is a mixture of  $\text{CO}_2$  and  $\text{n-C}_5\text{H}_{12}$  (55% and 45%), operated using a wire potential of 2900 V. The TGC has a readout time of 4 ns and gives a spatial resolution of 2-6 mm in the radial direction, and 3-7 mm in the  $\phi$  direction.

For muons with  $p_T$  of 1 TeV, the resolution of the muon spectrometer,  $\frac{\sigma_{p_T}}{p_T}$ , is 10%.

## 4.4 Trigger

At design luminosity LHC bunches will cross at a rate of 40 MHz. Each bunch crossing is referred to as an ‘event.’ Hardware limitations restrict the number of events ATLAS is capable of writing to disk to a rate of approximately 400 Hz [20]. To ensure collisions of interest are recorded within these constraints, events are ‘triggered’ in a three step system. These three stages of event selection are known as the Level 1, Level 2 and Event Filter (EF) trigger system. The latter two are collectively referred to as the High Level Trigger (HLT). Each trigger level refines the previous decision by applying further selection criteria. To trigger on an event, the detector is scanned in a search for specific physics signatures constructed from electrons, photons, quark/gluon jets, hadronically decaying taus, and missing energy signals. If the trigger requirements are met, an event is accepted, if not, it is rejected.

The Level 1 (L1) trigger decision is based on reduced granularity information from the ATLAS sub-detectors. The muon trigger chambers select muons, and basic information from the calorimeters crudely identify electrons, photons, jets and hadronic taus. A combination of L1 triggers reduce the event rate to 75 kHz. To meet this rate ‘prescaling’ is necessary for L1 triggers selecting events based on low momentum particles. Each prescale is a numerical value, designated to a particular trigger. Only one event from the prescaled number of events is kept.

During data taking there is latency period of  $2.5 \mu\text{s}$  to reduce dead time, when data is stored in buffers on front-end electronics attached to each sub-detector. If an event passes an L1 trigger, an accept signal is sent by the central trigger processor (CTP) that prompts the sub-detector read out drivers (RoDs) to send formatted information to the readout system. Here the data is stored temporarily on readout buffers (RoBs). The

stored information includes the specific L1 triggers that accepted the event each trigger's region of interest (RoI). The RoI is the locality of the detector that caused the trigger accept.

At Level 2 (L2), the trigger focuses on information in the RoIs, which contain  $< 2\%$  of the entire event information. This step allows more thorough signal reconstruction, including information from the ID. The rate of events is reduced to 3.5 kHz during a 40 ms processing time window. The EF level performs event reconstruction using the entire detector, with software similar to that used for analyses. This allows a complete and detailed picture of the event to which refined selection criteria is applied, ensuring the event acceptance rate is reduced down to the required 400 Hz. Once an event has passed the EF it is stored to disk.

## 4.5 Pile-up

Conditions at the LHC will produce  $\sim 23$  independent proton-proton collisions every bunch crossing at design luminosity [15]. This means there are multiple processes in a single event. These additional collisions, that usually do not cause a trigger accept, are called in-time pile-up. Furthermore, in the calorimeters and MDTs, the drift time for signals can be  $\geq 400$  ns. Therefore signal from more than 20 bunch crossings are superimposed. This is referred to as out-of-time pile-up. Pile-up adds spurious energy to an event, and thus degrades energy resolution.

In-time pile-up consists mainly of minimum bias collisions. There is a global structure to the events due to purely statistical effects. Such pile-up is modelled in **ATLAS** based on observations in data, as described in References [21] and [22]. To limit the effects of out-of-time pile-up signals are shaped for faster readout.

In analysis, pile-up is treated as a source of noise, although in reality it is not purely random [24]. However, the number of collisions in the LHC is very high, and pile-up can

be treated as a continuous background, and therefore as noise.

# Chapter 5

## Object Reconstruction

Events in ATLAS are recorded as energy measurements in calorimeter cells and hits in the ID and muon spectrometer. The stored electrical outputs are converted to the 4-vectors of physical particles, used in analyses, by various reconstruction algorithms. This Chapter describes the process of reconstructing quark and gluon jets, hadronically decaying taus, electrons, photons, muons, and missing energy. Before discussing particle reconstruction, there are brief explanations of track and ‘cluster’ formation. Clusters are a group of calorimeter cells containing energy deposits from a common particle. A synopsis of hadronic calibration methods is also given, the calibration of reconstructed hadronic particles is crucial for correct energy measurement.

### 5.1 Track Reconstruction with the Inner Detector

There are two complementary methods for track reconstruction in ATLAS. The first is an ‘inside-out’ approach for charged particles originating from the interaction region. It begins with hits from the pixel detector and progresses to the SCT and TRT. The second is an ‘outside-in’ approach, initiated in the TRT and continuing towards the pixels. The latter technique can reconstruct tracks from secondary vertex decays.

The inside-out algorithm begins by grouping hits in the silicon detectors that are

consistent with tracks emerging from the collision region. Grouping is constrained by the predicted curvature of a track, in order to restrict the number of combinations of individual silicon hits. In each group, the inside-out procedure takes the hit closest to the beam line, and then searches for successive hits on the track using a fitting procedure [25]. With every additional hit, information on the track curvature is updated to improve the fit. Not all initial groupings of hits result in a reconstructed track, the success rate is approximately 10% [26]. Additionally, if two track candidates share a large number of hits, both candidates are combined. Once a track is established from the silicon hits, the matching method is extended to the TRT.

The outside-in procedure addresses the reconstruction of tracks that originate beyond the interaction point. Track formation begins in the TRT, and is based on a histogramming method documented in Reference [27]. The method uses a transformation to convert the curved trajectory expected from a particle in the TRT straws to a straight line in a rotated coordinate system. The track segments built in the TRT are extended to the silicon detectors, with which space points can be formed.

Doubly counted tracks, built from both algorithms, are merged. For a track to be reconstructed by either of the two methods, a minimum transverse momentum of  $p_T > 0.5$  GeV is required throughout the ID.

The extrapolation of tracks back towards the beam line from the ID reconstructs the vertices at which physics interactions occurred. A primary vertex, as opposed to a pile-up vertex, is defined as one with at least four ‘quality’ tracks associated to it. Where a quality track is defined as one with:

- $p_T > 1$  GeV
- Number of pixel hits  $\geq 2$
- Number of pixel and SCT hits  $\geq 7$
- $|d_0| < 1.0$  mm

- $|z_0 \sin \theta| < 1.5 \text{ mm}$

## 5.2 Cluster formation in the Calorimeters

Particles that shower in the calorimeters deposit their energy among individual cells and these cells are combined in three dimensions across all calorimeters to produce clusters. There are multiple ways to construct a cluster, each giving a slightly different result. The method discussed here is called ‘topological clustering.’ Topological clusters are built from individual cells in the calorimeters by collecting neighbouring cells that contain significant energy compared to the expected noise. The cell by cell noise varies across the detector by orders of magnitude, depending on the electronics and pile up conditions [28]. Each cell that satisfies a signal to noise ratio greater than four ( $signal/noise > 4$ ) is referred to as the seed for a cluster. The neighbouring cells of each seed are included in the cluster provided they satisfy  $signal/noise > 2$ . This process is continued iteratively to include or reject all neighbours of the cells within a cluster. When all cells that satisfy the  $signal/noise$  condition have been included, a final layer of neighbouring cells are added to each cluster, regardless of energy. The inclusion of these bordering cells ensures that energy deposits from the tails of a particle shower are not discarded. The method of topological clustering suppresses electronics noise as well as pile up from minimum bias events.

## 5.3 Hadronic calibration

The ATLAS calorimeter is non-compensating; the response for electromagnetic showers and hadronic showers differs. An energy scaling procedure must thus be applied to each reconstructed jet and hadronic tau in order to obtain an estimate of the original parton or tau energy. The calibration can be performed on the jet/tau object, this is referred to as global hadronic calibration, or on the constituent clusters, this is known as local

hadronic calibration. Both methods correct the energy by applying weights that are derived from simulated data (the simulation of data is described later, in Chapter 6 and in Appendix B). Although not the case at the time of this thesis, the weights could also be derived using a data driven approach.

Local hadronic calibration works by designating clusters as either em-like or hadronic-like; the goal is to distinguish between clusters dominated by electromagnetic showers and those dominated by hadronic interactions. As a single final decay particle is usually contained within one to two clusters [29], this is essentially a separation of electrons, photons, and neutral pions from charged pions, and neutrons. The clusters are classified based on shape variables, which are a result of the way different particles interact with matter, as discussed briefly in Section 4.2. The most promising shape variables for separating hadronic and electromagnetic showers are the depth of the shower, the energy density, the width and the length of the shower. Hadronic showers are broader, deeper into the detector and less energy dense than electromagnetic showers. Each hadronic-like cluster has a weight assigned, dependent on its shape variables, its energy, and  $\eta$ .

The benefit of local calibration over global calibration is that it determines the energy with a better resolution [29]. This is due directly to the ability to separate the electromagnetic and hadronic components of an object and treat them accordingly.

Additionally, both the electromagnetic and hadronic clusters are corrected for energy lost in the dead material that comes before the calorimeter systems. The dead material includes the cryostats, the ID, and structural support material. Energy loss in dead material is also addressed with the global calibration approach, as the weighting method based on the comparison of reconstructed objects to generated objects in simulation must inherently do so.

## 5.4 Jets

Quarks and gluons produced from interactions in ATLAS fragment and hadronize to form collimated jets of hadronic particles. These jets manifest themselves as localized energy deposits in the calorimeters. Jets used in this analysis are reconstructed with the anti- $k_T$  algorithm [30, 31, 32]; this takes topological clusters [28] as inputs.

### 5.4.1 anti- $k_T$ algorithm

To reconstruct the jet, topological clusters are combined using the anti- $k_T$  algorithm. This combination process ensures soft particles cluster with harder transverse momentum particles, rather than amongst themselves [30]. The algorithm uses two sets of criteria. Firstly, the transverse momentum weighted distance between two clusters labelled  $i$  and  $j$ ,  $d_{i,j}$ , as given in Equation 5.1. Secondly, the inverse of the cluster  $p_T^2$ , labelled  $d_{i,B}$ , as in Equation 5.2.

$$d_{i,j} = \min\left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2}\right) \frac{\Delta R_{i,j}^2}{0.4^2} \quad (5.1)$$

$$d_{i,B} = \frac{1}{p_{T,i}^2} \quad (5.2)$$

Clusters are combined to form jets by calculating the smallest of  $d_{i,j}$  or  $d_{i,B}$ . If it is  $d_{i,j}$ , clusters  $i$  and  $j$  are combined and treated as a single entity. Note that, clusters physically close together (within  $\Delta R_{i,j} < 0.4$ ) are preferred due to the imposed factor of  $0.4^2$  in the denominator of  $d_{i,j}$ . If  $d_{i,B}$  is smaller, the cluster(s) represented by  $i$  is labelled a jet. This process is performed until no entities remain. Jets constructed by the anti- $k_T$  algorithm are not greatly altered by soft particles, and tend to be conical structures formed around isolated hard  $p_T$  objects.

The topological clusters used in the anti- $k_T$  algorithm are at the electromagnetic energy scale. Therefore the jets are corrected to the hadronic energy scale using a global

hadronic calibration method. As described in Reference [33], calibration weights to shift the jet reconstructed energy back to the generator level were extracted in fine bins of the jet  $p_T$  and  $\eta$  from simulated data, and applied to all reconstructed jets in data.

## 5.5 Hadronically decaying $\tau$ leptons

Hadronically decaying tau leptons are reconstructed using anti- $k_T$  jets. These jets are referred to as seeds for the tau reconstruction algorithm, which then associates tracks to each seed jet. Tau reconstruction uses all seed jets within  $|\eta| < 2.5$ , the range of the ID.

For taus, the clusters that compose an anti- $k_T$  jet are first calibrated with the local hadronic method. To correct for any residual calibration effects that are not sufficiently addressed by local hadronic calibration, a final global calibration correction is also applied [33]. As for jets, this is performed in separate regions of  $p_T$  and  $\eta$ .

The  $\eta$  and  $\phi$  directions of the  $\tau_h$  are determined by calculating the sum of the four-vectors of the constituent topological clusters, assuming zero mass for each cluster [33]. The tau mass is defined to be zero during reconstruction, meaning the  $\tau_h$  momentum and energy are identical. Quality tracks are associated to each tau candidate if they are within  $\Delta R < 0.2$  of the tau axis. The distance measurement,  $\Delta R$ , is defined in terms of  $\eta$  and  $\phi$  as  $\sqrt{\Delta\eta^2 + \Delta\phi^2}$ . The tau axis is the direction from the energy weighted centre of the constituent topological clusters to the interaction point. Reconstructed taus are classified as single or multi-prong depending on the number of associated tracks.

Taus that decay leptonically are reconstructed directly as an electron or muon, the contribution from neutrinos is measured by the missing energy term. In this thesis, the term tau refers only to hadronically decaying taus.

## 5.6 Electrons

Electrons from the interaction point leave a single track in the ID and shower in the EM calorimeters. Reconstruction begins with a calorimeter seed cluster of  $E_T > 2.5$  GeV. Seed clusters are formed using a sliding window algorithm [28], not with topological clustering. The sliding window algorithm begins by calculating the energy within a region of  $3 \times 5$  in units of  $\eta$  and  $\phi$  around high energy cells in the EM calorimeters. The region size corresponds to the shape of an electromagnetic shower. The algorithm proceeds by displacing the block of selected cells around the original position and recalculating the contained energy, until a stable energy cluster is established. An inner detector track with  $p_T > 0.5$  GeV is matched to the cluster within  $\Delta\eta \times \Delta\phi$  of  $0.05 \times 0.1$ . The larger allowance in the  $\phi$  direction is to compensate for the curvature of tracks bent by the field of the solenoidal magnet.

## 5.7 Muons

High energy muons produced in collisions will traverse the calorimeters, depositing very little energy. However, they create track hits in both the inner detector and the muon spectrometer. Initially muon track reconstruction algorithms use information solely from the muon spectrometer. Line track segments that point towards the interaction point are constructed in the magnetic bending plane ( $rz$ ) by linking hits in the MDTs and CSCs. These segments are combined to form a track. The tracks from the muon spectrometer are propagated back to the ID, by modelling the effect of detector material on the trajectory of the muon and on the measured or parametrized energy loss in the calorimeters. The resulting track parameters are compared to the reconstructed ID tracks using a  $\chi^2$  fitting procedure [34]. If a match is found with a  $\chi^2$  probability larger than 0.001, the ID and muon spectrometer tracks are combined in a global track refit, providing a muon candidate. The momentum of the combined track is corrected for the average  $\sim 3$  GeV

of energy that a muon will deposit in the calorimeters.

## 5.8 Missing transverse energy

Neutrinos do not interact with the ATLAS detector and their presence in an event leads to missing transverse energy ( $E_T^{Miss}$ ). Conceptually, the reconstruction of  $E_T^{Miss}$  in an event is based on the difference between the detected  $E_T$  and the initial  $E_T$  of the beams, which is zero.

To compute the total  $E_T$  in the detector, topological clusters are constructed in  $|\eta| < 4.5$ , and calibrated using the local hadronic method. The contribution from muons is given by their reconstructed  $p_T$ . The energy deposited by muons in the calorimeters is corrected for [35].

Energy from all the clusters and muons in an event is summed, where the cluster directions are taken from each topological cluster to the interaction point. Explicitly, the  $E_T^{Miss}$  term is defined as:

$$E_T^{Miss} = - \sum E_T^{cluster} + E_T^\mu - p_T^\mu \quad (5.3)$$

## 5.9 Object Identification

The reconstruction algorithms for particle objects are generally quite efficient since they are based on loose detector properties. For use in analysis particles are subject to additional ‘identification’ requirements. There are individual identification criteria for each particle object, pertaining to the detector cluster shapes and tracks expected for the physical particle. Identification criteria for taus are discussed in detail in Chapter 9. Identification of other particles used in this thesis are detailed in Chapter 10.

# Chapter 6

## Data and Monte Carlo

### 6.1 Data sets used in this analysis

Data recorded by the ATLAS detector corresponds to continuous time periods called ‘runs.’ This thesis uses data from 89 runs recorded in 2011, with an integrated luminosity of  $1.056 \text{ fb}^{-1}$ . Multiple runs are combined into chronological data taking periods for runs that share common parameters such as beam intensity and amount of pileup. The data periods are assigned letters, those used in this analysis are given in Table 6.1. A complete list of the run numbers is given in Appendix A.

### 6.2 Monte Carlo

Experimental observations in data are compared to theoretical predictions using simulated data, called Monte Carlo (MC). In simulated events, generators play the role of the LHC, and detector simulation programs the role of ATLAS. The observed and simulated data share the same event reconstruction framework and subsequent physics analysis. A general description of Monte Carlo methods is given in this Chapter.

Monte Carlo is produced in three sequential steps:

**Generation** A generator begins by producing an interaction between partons in the colliding protons. The process is specified in the input, and is usually chosen to be a hard scattering process. The production amplitudes are processed using matrix elements which can be calculated perturbatively using Feynman calculus. Additionally, any radiative emissions from the partons before and after the hard interaction are generated, as well as any interactions of the additional partons in the proton. Finally, all unstable particles produced that will decay before reaching a material layer in the ATLAS detector are forced to do so. For ‘free’ quarks and gluons this involves hadronization and fragmentation into jets of hadronic particles. The results of the generator are a list of final state particles, and their 4-vectors, as produced in the event. A filter can be applied to select only events with specific properties (e.g. events containing a leptonic decay, or more complicated requirements such as angular separations of decay products).

**Simulation** The particle record from the event generation stage is the input of the simulation. Each particle is propagated through a geometric model of the ATLAS detector by Geant4 [36, 37], which provides detailed models of physics interactions and the infrastructure to propagate particles through material layers. As the particles are propagated through the detector layers they may decay or radiate causing the generation of new particles, which are themselves propagated by the simulation. In particular, interactions in the calorimeters often lead to the development of large particle showers. Finally, energy deposited by particles in the sensitive regions of the detector are recorded as hits containing information on energy deposition, position, and time.

**Digitization** Hits from simulation are read into the digitization. Detector noise is added to the energy depositions and the decisions from the L1 trigger are evaluated and recorded, however no events are discarded. The energy depositions in the various

detector components are converted into voltages and currents. In the real detector these signals are converted into digital output via the read out drivers. In the digitization, the performance of these drivers is emulated to obtain a final data format identical to the output of the real detector, allowing it to be processed by the standard reconstruction algorithms.

The search for the MSSM Higgs would not be feasible without Monte Carlo. A comparison to the Standard Model expectations is needed in order to form a quantitative statement on the discovery or exclusion of the Higgs. Such a comparison requires an accurate description of theoretical predictions, in a form that is directly comparable with data. The Monte Carlo provides just that. Of course, it is naive to trust the Monte Carlo implicitly. Whenever possible, it is verified that the Monte Carlo provides an accurate description of the data it is attempting to model.

A list of the physics processes modelled with Monte Carlo for use in this analysis are given in Table 6.2, along with the generators used. A more detailed description of the generators is given in Appendix B.

Data Period	Included runs	Total number of runs
B <sub>2</sub>	178044 - 178109	3
D	179710 - 180481	23
E	180614 - 180776	5
F	182013 - 182519	17
G	182726 - 183462	28
H	183544 - 184169	13

**Table 6.1:** *Data periods used in this analysis, with the number of runs in each data period indicated.*

Physics process	Generator
QCD jets	PYTHIA
$Z + \text{jets}$	ALPGEN
$W + \text{jets}$	ALPGEN
$t\bar{t}$	MC@NLO
single top (s and t channel)	MC@NLO
single top (associated W)	MC@NLO
Diboson	HERWIG
$bbA$	SHERPA
$gg \rightarrow A$	MC@NLO

**Table 6.2:** *Complete list of the event generators in this analysis, along with the physics process they were used to emulate.*

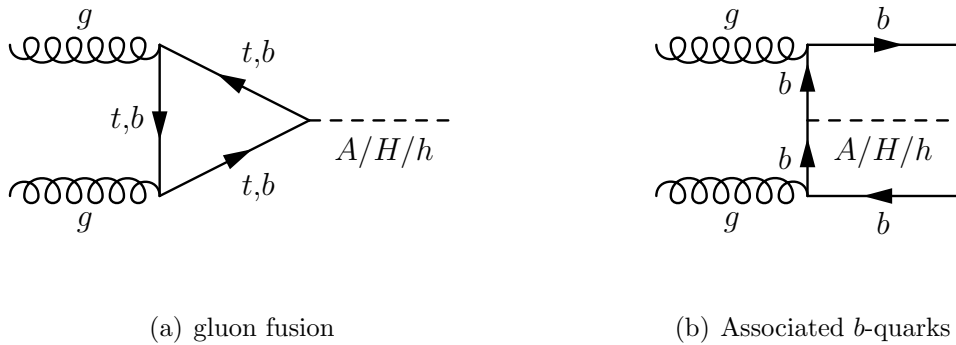
# Chapter 7

## MSSM Higgs signal

### 7.1 Higgs production

In the MSSM, neutral Higgs bosons are produced through gluon fusion or in association with  $b$ -quarks. Gluon fusion dominates for low Higgs boson masses and small values of  $\tan\beta$ . The leading order production processes are shown in Figure 7.1.

Exploring the entire MSSM phase space is difficult because of the large number of free parameters. Therefore, searches at the LHC, and previously at LEP [38] and the Tevatron [1], are performed in specific benchmark scenarios [14]. Parameters in this analysis are defined by the  $m_h^{\max}$  scenario; this gives maximum values for the CP even Higgs



**Figure 7.1:** Dominant MSSM Higgs production mechanisms at the LHC.

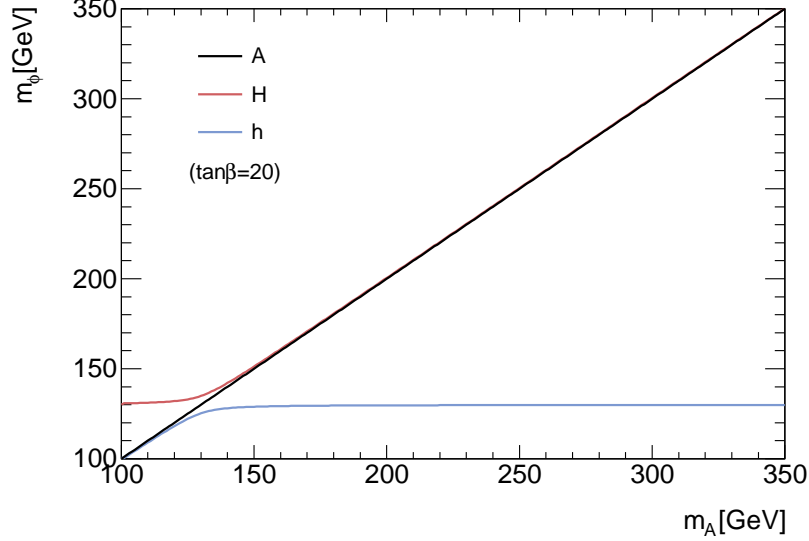
boson mass,  $m_{h^0}$ , as a function of  $\tan\beta$ . The fixed parameters in the  $m_h^{\max}$  benchmark are as follows:

- $m_{\text{top}} = 172.5 \text{ GeV}$
- $\bar{m}_b(\bar{m}_b) = 4.213 \text{ GeV}$
- $\alpha_S(M_Z) = 0.119$
- $M_{SUSY} = 1 \text{ TeV}$
- $X_t = 2 \text{ TeV}$  (the stop mixing parameter)
- $M_2 = 200 \text{ GeV}$  (the gaugino mass parameter)
- $\mu = 200 \text{ GeV}$

In the  $m_h^{\max}$  scenario the Higgs search depends on two unknown quantities: the mass of the CP odd Higgs,  $m_A$ , and the ratio of vacuum expectation values,  $\tan\beta$ . Monte Carlo samples are generated with  $\tan\beta = 20$  for fourteen different mass points,  $m_A$ , ranging from 110 to 350 GeV. Associated  $b$ -quark processes ( $b\bar{b}A/H/h$ ) are simulated with SHERPA [68], and gluon fusion production ( $gg \rightarrow A/H/h$ ) with MC@NLO [67]. The signal samples are listed in Appendix D. This Monte Carlo is also used to study the  $h/H$  bosons, assuming the same final state kinematics as for the  $A$ . Figure 7.2 shows the relationship between  $m_A$  and  $m_h/m_A$  for a representative value of  $\tan\beta = 20$ . Above  $\sim 140 \text{ GeV}$ , for all  $\tan\beta$ , the  $A$  and  $H$  bosons have similar masses. For all  $m_A$  and  $\tan\beta$ ,  $m_h$  remains close to 130 GeV.

### 7.1.1 Production cross-section

$A/H/h$  cross-sections, with  $m_h^{\max}$  parameters, are computed separately for each production mechanism. Details of the calculation methods are not given, but are available in the associated references. The gluon fusion production cross-section is calculated using



**Figure 7.2:** *Higgs  $A/H/h$  masses as a function of  $m_A$ , using  $\tan\beta = 20$ .*

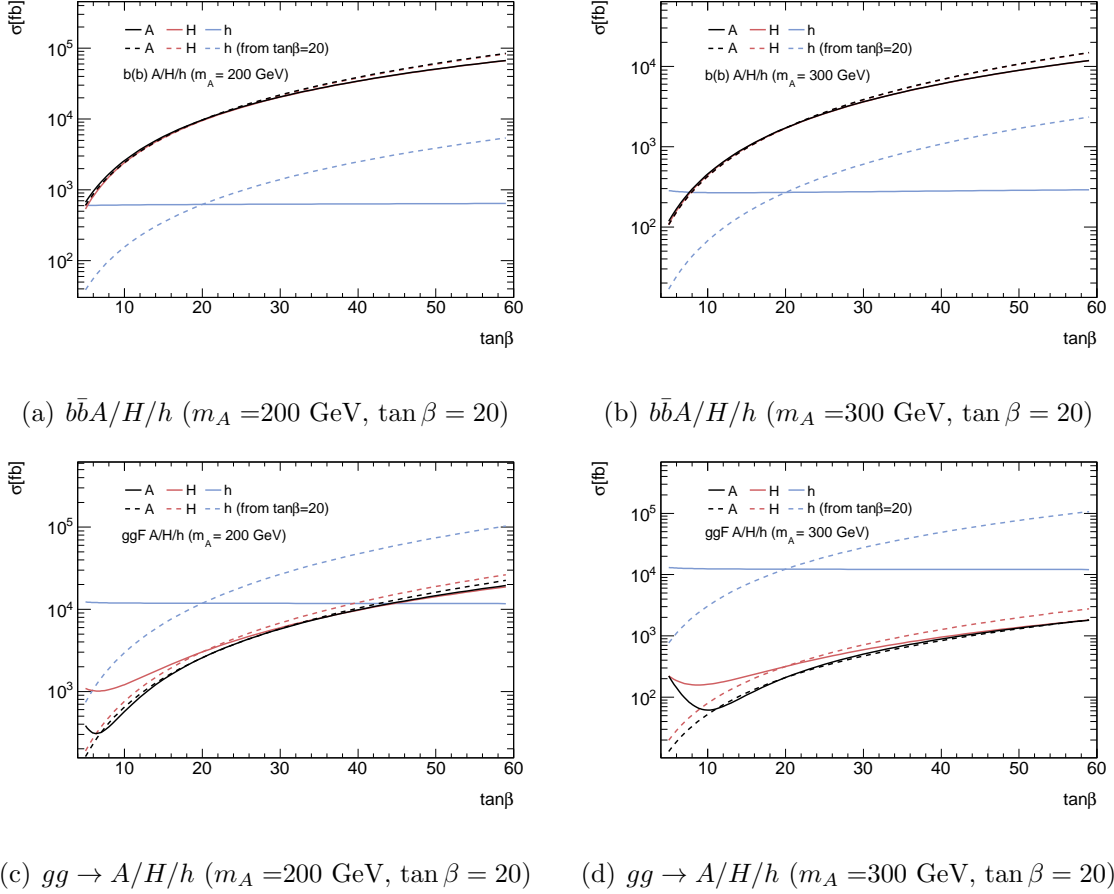
HIGLU [39] and ggh@nnlo [40]. The  $b$ -associated production cross-section is calculated with bbh@nnlo [41] for both the 4 flavour scheme [43], where  $b$ -quarks parton distribution functions are omitted in the calculation, and the 5 flavour scheme. A combined result from both  $b$ -associated production schemes is obtained using the Santander [45] matching method. Expected cross-sections are given in Table 7.1 for  $m_A = 200$  and 300 GeV, with  $\tan\beta = 20$  and 40. Cross-sections for all Monte Carlo generated mass points are given in Appendix D. Additional details for production cross-section calculations are available in Reference [42].

Theoretically, the natural width of the Higgs grows as  $\tan\beta$  increases. The expected width for  $m_A = 300$  GeV is  $\sim 1$  GeV for  $\tan\beta = 20$ , compared to  $\sim 10$  GeV for  $\tan\beta = 50$ . However, Monte Carlo samples are only generated with  $\tan\beta = 20$ , as the visible mass distribution from a Higgs decaying to two hadronic taus is broad compared to the natural width. To evaluate cross-sections for  $\tan\beta$  values other than 20 a multiplicative scaling of  $\tan^2\beta/20^2$  is applied to the  $\tan\beta = 20$  production cross-sections, for the  $A$  and  $H$  bosons. Figure 7.3 shows the theoretical production cross-

$m_A, \tan \beta$	$gg \rightarrow A/H/h$ cross-section (pb)	$b\bar{b}A/H/h$ cross-section (pb)
200 GeV, 20	2.58/3.03/11.90	9.72/9.52/0.62
300 GeV, 20	0.21/0.32/12.31	1.71/1.70/0.27
200 GeV, 40	9.84/9.76/11.52	34.41/34.20/0.63
200 GeV, 40	0.89/9.54/12.21	6.07/6.04/0.28

**Table 7.1:** *Theoretical production cross-sections for MSSM neutral Higgs at the LHC, with  $m_A = 200$  or  $300$  GeV and  $\tan \beta = 20$  or  $40$ .*

sections of  $b\bar{b}A/H/h$  and  $gg \rightarrow A/H/h$  as a function of  $\tan \beta$  for  $m_A = 200$  GeV and 300 GeV, as well as the scaled cross-section values derived from  $\tan \beta = 20$ . The scaling is not applicable for the  $h$  boson. Effectively, for a specific  $m_A$ , the production cross-section for the  $h$  remains constant. Uncertainties in the production cross-section of the MSSM neutral Higgs, including branching ratios, are obtained with *mssm\_xs\_tool* [46], which is based on HIGLU [39], ggh@nnlo [40], bbh@nnlo [41] and FeynHiggs [44]. These uncertainties are discussed further in Section 15.1.1.



**Figure 7.3:** Theoretical production cross-section of  $b\bar{b}A/H/h$  and  $gg \rightarrow A/H/h$  processes as a function of  $\tan\beta$  for  $m_A = 200$  GeV and 300 GeV. Dashed lines show estimated cross-sections from the values of  $\tan\beta = 20$  by multiplying with a factor of  $\tan\beta^2/20^2$ .

## 7.2 Higgs decay to taus

The Higgs couplings to taus in the MSSM are given by Equations 7.1 and 7.2. (The coupling has the same form for the  $b$  quark, with the obvious replacement of  $m_\tau$  with  $m_b$ ).

$$-\left(\frac{gm_\tau}{2m_W}\right)\bar{\Psi}\Psi\left(-\frac{\sin\alpha}{\cos\beta}h^0+\frac{\cos\alpha}{\cos\beta}H^0\right) \quad (7.1)$$

$$i\left(\frac{gm_\tau}{2m_W}\right)\tan\beta\bar{\Psi}\gamma_5\Psi A^0 \quad (7.2)$$

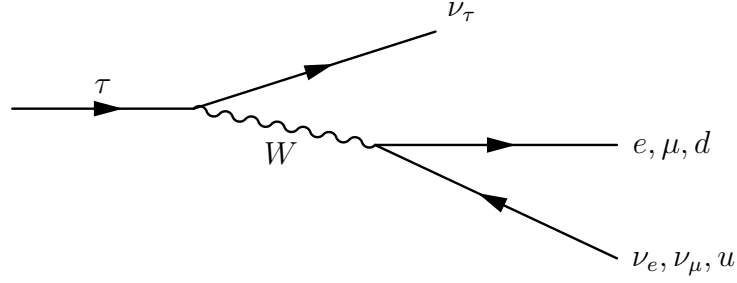
In contrast, the Standard Model couplings are given by Equation 7.3.

$$-\left(\frac{gm_\tau}{2m_W}\right)\bar{\Psi}\Psi H_{SM} \quad (7.3)$$

In particular, we see there can be a large enhancement by the value of  $\tan\beta$  in the MSSM coupling. Additionally, in Section 2.3.1, we found that the coupling between the  $A$  and  $W/Z$  bosons was zero. Numerically the branching fractions for  $A/H/h \rightarrow b\bar{b}$  and  $A/H/h \rightarrow \tau\tau$  are  $\sim 0.9$  and  $\sim 0.08$  respectively. Consequently, these are the relevant decay modes for neutral MSSM Higgs searches. Despite the large advantage in branching fraction of the  $b\bar{b}$  mode, the  $\tau\tau$  channel is very relevant. Because of the large QCD jet background at the LHC,  $b\bar{b}$  final states are extremely difficult signatures to analyze.

## 7.3 Tau decay modes

A tau decays to the lighter leptons with a branching fraction of 0.352. This is to electrons and the associated neutrinos with a fraction 0.178, and to muons (and associated neutrinos) with a fraction 0.174. The basic decay mode of the tau is shown in Figure 7.4.



**Figure 7.4:** *Tau decay modes. The leptonic decay process has two associated neutrinos.*

The remaining tau branching fraction ( $\sim 0.65$ ) is through hadronic decays. Taus are the only leptons with sufficient mass to decay hadronically. Of these hadronic decays, 77% result in one charged particle. This occurs either by the process shown in Figure 7.4, for about a quarter of the cases, and through resonance states of either the  $\rho(770)$  or  $a_1(1260)$  for the remaining.

Another 22% of the total hadronic decays produce three charged particles, and the small remainder go to five charged particles. The charged particles are mostly pions, but  $\sim 5\%$  of the decays involve kaons. The decays involving the strange quark are thought to proceed via kaon resonances. More than 60% of the hadronic decays also have neutral pions in the process [12].

# Chapter 8

## Background Processes

Collisions in the LHC result in a variety of physics processes, with relative rates dictated by the individual production cross-sections. We've seen from Section 7.1.1, that the  $A/H/h \rightarrow \tau_h \tau_h$  signal, if it exists, is expected with a cross-section of  $\mathcal{O}(0.5 \text{ pb})$ , including the relevant branching ratios. Other background processes are known to exist with significantly larger cross-sections. To detect a signal, events must be carefully selected from the background. This Chapter describes the background processes, and the event selection is detailed later, in Chapter 10.

Broadly speaking, any event with a reconstructed signature of two high momentum hadronic taus, and a significant amount of missing energy can contribute to background. This can occur because of mis-measured particles, a quark jet mis-identified as a  $\tau_h$  for example, or through irreducible backgrounds, where the signature is identical to signal.

The theoretical cross sections for Standard Model backgrounds are summarized in Table 8.2, where the signal process is also shown for comparison. Although the number of events from QCD jets are estimated with data (as will be described in Chapter 13), Table 8.1 lists the theoretical cross-sections for completeness. References [48, 49] describe the calculations of all cross sections. It is possible that SUSY processes themselves contribute to background, but the theoretical production cross-sections are in general

much lower than for the major Standard Model backgrounds.

## 8.1 QCD multi-jet events

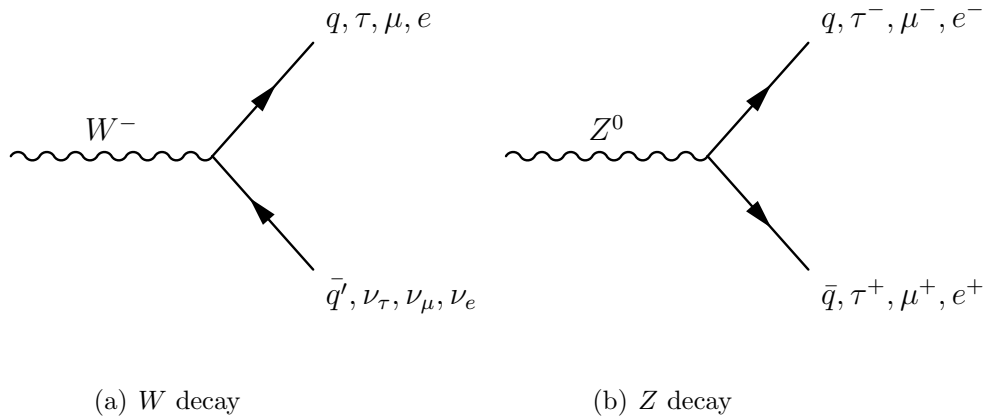
Jet production through strong interactions in the LHC occur at a rate many orders of magnitude larger than the signal. Table 8.1 gives the multi-jet cross-sections, as calculated with the generator PYTHIA [66], divided into categories of the highest- $p_T$  generated jet. Typically the signature from QCD events is different from the signal, but the volume of QCD means that any residual events that appear signal-like will contribute significantly to the background. In fact, mis-reconstructed dijet events with each jet incorrectly identified as a hadronically decaying tau constitute the dominant background source. Understanding this background process, and suppressing it effectively is crucial in this analysis. However it is not possible to produce Monte Carlo data with the equivalent number of dijet events expected in real events. Instead, an approach based on real data is used to obtain the number of background events and shapes of distributions coming from QCD. This method is detailed in Chapter 13.

Range of highest jet $p_T$	cross-section (nb)
< 17 GeV	9752970
17 - 35 GeV	673020
35 - 70 GeV	41194.7
70 - 140 GeV	2193.25
140 - 280 GeV	87.2487
280 - 560 GeV	2.32856
560 - 1120 GeV	0.03385

**Table 8.1:** *Cross-sections for inclusive QCD dijet events in the LHC using the PYTHIA generator. The cross-sections are divided into categories by the generated leading jet  $p_T$ .*

## 8.2 $Z \rightarrow ee/\mu\mu$ and $W \rightarrow e\nu/\mu\nu + \text{jets}$

$Z$  and  $W$  decays to electrons or muons are minor backgrounds in this analysis. Events with well reconstructed electrons or muons are rejected in this study, as will be described in Chapter 10. To assess the number of events that result in mis-identified taus forming a signal-like signature in  $W$  and  $Z$  boson decays, Monte Carlo data generated with ALPGEN [69] is analyzed. The ALPGEN generator is a collection of LO generators with up to six associated jets. It is used for  $W+$  and  $Z+$  jet processes in order to better model the influence of additional jets, which are abundant in the QCD dominated background at the LHC. Separate samples of  $W/Z + 0, 1, 2, 3, 4, 5$  partons are generated, and are summed together to get a complete  $W/Z + \text{jets}$  sample. The inclusive  $Z \rightarrow ee/\mu\mu$  cross-section is calculated as  $1.07 \pm 0.05$  nb [49], and inclusive  $W \rightarrow e\nu/\mu\nu$  cross-section as  $10.46 \pm 0.52$  nb [49]. Figure 8.1 displays all decay mechanisms of the  $Z$  and  $W$  bosons, no additional jets are shown.



**Figure 8.1:**  $W$  and  $Z$  decay processes.

### 8.3 $W \rightarrow \tau\nu + \text{jets}$

$W$  bosons that decay to taus are a source of background when the tau decays hadronically, and there is large  $E_T^{Miss}$  in the event, as expected from the associated neutrino. An additional QCD jet accompanying the  $W$  can be mis-identified as a  $\tau_h$ , leading to a signal-like signature. The inclusive  $W \rightarrow \tau\nu_\tau + \text{jets}$  background is analyzed with ALPGEN generated samples. The process has a calculated cross-section of  $10.46 \pm 0.52$  nb [49].

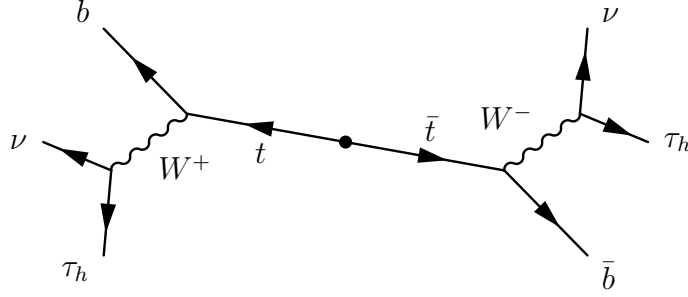
### 8.4 $Z/\gamma^* \rightarrow \tau\tau (+ \text{jets})$

$Z$  decays to hadronic taus are an irreducible background. The final state is identical to signal, with similar kinematics. In particular, the background is important for low mass Higgs bosons, when the signal falls in the tail region of the  $Z$  mass peak. The effect is enhanced in this analysis, since we do not consider the invariant mass from the  $\tau\tau$  process, but rather the visible mass ( $\tau_h\tau_h$ ), which omits neutrinos. As the contribution from neutrinos varies, the visible mass has a broader spread than the invariant mass. The calculated inclusive cross-section of  $Z \rightarrow \tau\tau + \text{jets}$  is  $1.07 \pm 0.05$  nb [49].  $Z/\gamma^*$  samples are used to assess this background, cross-sections are given for the full mass spectrum in Table 8.2. The number of background events is evaluated with Monte Carlo generated with ALPGEN.

### 8.5 top quark production

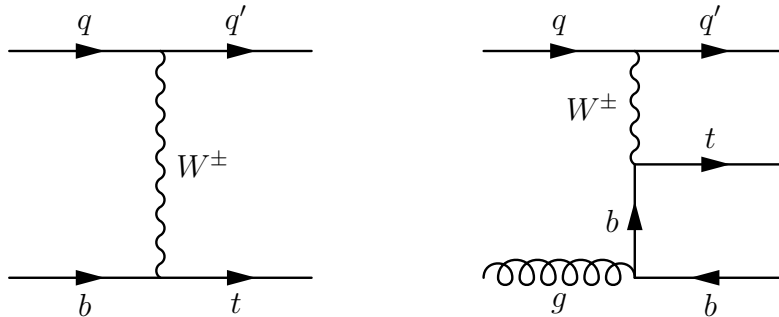
The branching ratio of  $t\bar{t}$  events decaying to  $b\bar{b}$  and two  $W$  bosons exceeds 98%, and tends to result in a high jet multiplicity. This leads to a large fraction of mis-identified taus, and a small fraction of events with real taus, as shown in Figure 8.2. However,  $t\bar{t}$  production has a small cross-section, and for that reason the number of expected background events is small. Calculated to NLO QCD, using a top mass of 172.5 GeV, the cross-section

is  $164.57 +11.45 \text{ pb} -15.78 \text{ pb}$  [48]. This background is evaluated with MC@NLO [67] generated samples.



**Figure 8.2:**  $t\bar{t}$  decay to two hadronic taus and  $b$ -jets.

Single tops are produced in  $t$ -channel processes, as shown in Figure 8.3;  $s$ -channel processes, shown on the left side of Figure 8.4; and also in association with a  $W$  boson, shown on the right side of Figure 8.4. Calculated cross-sections are 58.7 pb for  $t$ -channel, 3.94 pb for  $s$ -channel and 13.1 pb for associated  $W$  production, each with a 5% uncertainty [48]. Monte Carlo samples used for the analysis of background from these channels are generated with MC@NLO.

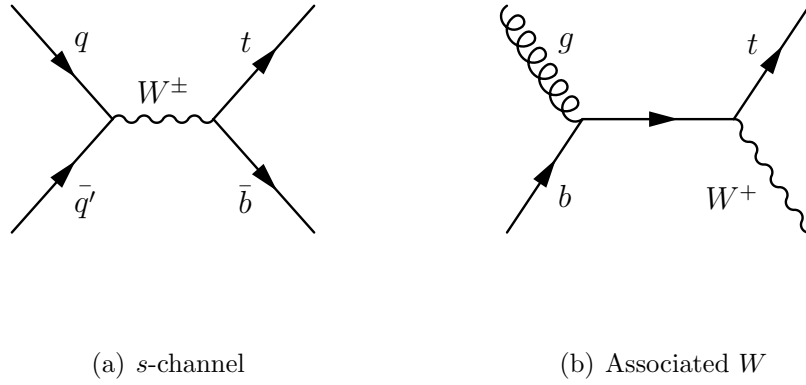


**Figure 8.3:**  $t$ -channel single top production.

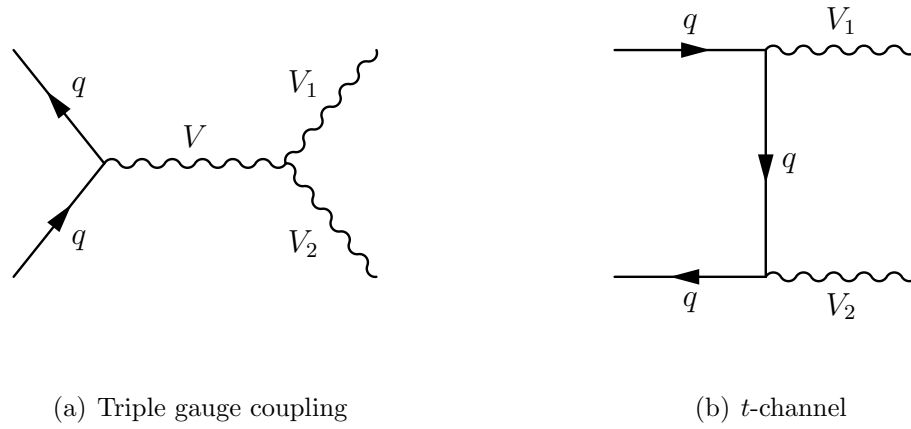
## 8.6 Diboson: $WW$ , $WZ$ , $ZZ$

The electroweak production of pairs of vector bosons ( $WW$ ,  $WZ$  and  $ZZ$ ) can lead to final states with hadronic taus, or to quark jets that can be misidentified as such. The inclusive  $WW$  cross-section is  $44.9 \pm 2.2$  pb,  $WZ$  is  $18.0 \pm 1.3$  pb, and  $ZZ$  is  $5.64 \pm 0.28$  pb [49]. Figure 8.5 shows the production modes for diboson events, which are studied with MC@NLO generated Monte Carlo samples. Overall the diboson background cross-section is small, as is the expected contribution of background events.

All background from  $W \rightarrow \tau\nu$ ,  $W \rightarrow e\nu$ ,  $W \rightarrow \mu\nu$ ,  $Z \rightarrow ee$ ,  $Z \rightarrow \mu\mu$ ,  $Z \rightarrow \tau\tau$  and diboson processes are collectively referred to as electroweak (EW) background. Top background consists of contributions from both  $t\bar{t}$  and single- $t$ . In this thesis, the number of events that pass selection for both the EW and top backgrounds are analyzed using Monte Carlo; there is no data driven approach. A summary of EW and top background cross-sections is given in Table. 8.2, signal cross-sections with  $m_A = 200$  GeV and  $\tan \beta = 20$  are also shown for comparison.



**Figure 8.4:**  $s$ -channel and  $W$  associated single top production.



**Figure 8.5:** Diboson production.

Process	cross-section (pb)
$W \rightarrow l+\text{jets} \ (l = e, \mu, \tau)$	$10.46 \times 10^3$
$Z/\gamma^* \rightarrow ll+\text{jets} \ (m_{ll} > 40 \text{ GeV})$	$1.07 \times 10^3$
$Z/\gamma^* \rightarrow ll+\text{jets} \ (10 < m_{ll} < 40 \text{ GeV})$	$3.89 \times 10^3$
$t\bar{t}$	164.6
Single top: $t, s$ , assoc. $W$	58.7, 3.9, 13.1
Diboson: $WW, WZ, ZZ$	46.2, 18.0, 5.6
$gg \rightarrow A/H/h \rightarrow \tau\tau$	0.31/0.37/1.09
$b\bar{b}A/H/h \rightarrow \tau\tau$	1.17/1.17/0.06

**Table 8.2:** Cross-sections (multiplied by relevant branching ratios). NNLO calculation for  $W/Z + \text{jets}$  and  $b\bar{b}A/H/h \rightarrow \tau\tau$ , NLO+NNLL for  $t\bar{t}$ , NLO for single top and dibosons, NNLO(top loop)+NLO(bottom loop and top/bottom loop interference) for  $gg \rightarrow A/H/h \rightarrow \tau\tau$ . Signal with  $m_A = 200 \text{ GeV}$  and  $\tan\beta = 20$ .

# Chapter 9

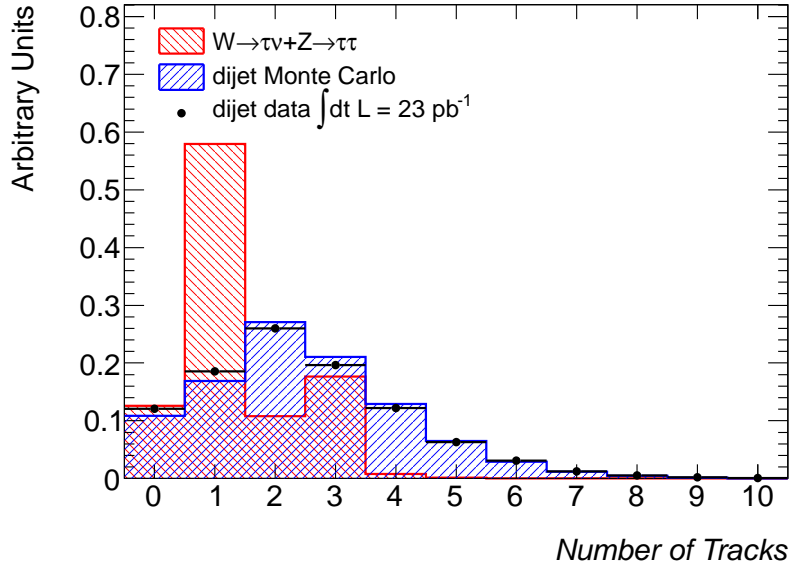
## Tau Identification

Candidates for reconstructed hadronic taus that are to be used in analyses are produced from all anti- $k_T$  jets within the fiducial region of the inner detector ( $|\eta| < 2.5$ ), as discussed in Section 5.5. The rejection of quark and gluon jets from the tau candidate pool is performed in a separate identification step using a set of discriminating variables based on distinct properties of hadronically decaying taus. These variables are input to a projective likelihood identification method, and the resulting discriminant separates hadronically decaying taus from QCD jets. The likelihood identification is trained using Monte Carlo input for signal  $\tau_h$  candidates, and dijet selected data events for background QCD jets. Efficiency values for the identification method obtained from Monte Carlo are verified with data.

### 9.1 Tau Identification variables

Taus decay hadronically in a relatively collimated region compared to QCD jets, which tend to be broader. Additionally, hadronic taus, on average have a larger fractional contribution of neutral pions than QCD jets. This results in large energy deposits in the electromagnetic calorimeters. Of the taus that decay hadronically, 77% do so with a single charged particle, and  $\sim 23\%$  to three charged particles. These are referred to

as single-prong and multi-prong taus respectively. The small fraction of taus that decay to five or more charged particles are ignored, as it becomes difficult to distinguish these from QCD background. Figure 9.1 shows the number of reconstructed tracks associated to a  $\tau_h$  candidate. Single-prong candidates dominate for generated taus, as expected. Reconstructed multi-prong taus usually have three tracks, as would a generated tau, however due to track reconstruction inefficiencies, these can also be reconstructed with two tracks. To avoid repeating terms in the description below, we'll define the *core cone* as the region within  $\Delta R < 0.2$  of the tau axis, and the *isolation annulus* as  $0.2 \geq \Delta R < 0.4$  from the tau axis.



**Figure 9.1:** Number of core tracks associated to  $\tau_h$  candidates coming from  $Z \rightarrow \tau\tau$ ,  $W \rightarrow \tau\nu$ , and dijet PYTHIA generated Monte Carlo. Data points shown with markers are from dijet selection applied to data.

Taus are divided into single-prong and multi-prong candidates to provide better separation from QCD background. Discriminating variables are produced individually for each category [50]. Variables are described below with corresponding figures of the distributions. In each figure the  $\tau_h$  is taken from Monte Carlo samples using the PYTHIA generated [66]  $Z \rightarrow \tau\tau$  and  $W \rightarrow \tau\nu$  events. The QCD background is shown from

PYTHIA dijet samples, as well as dijets selected events from  $23 \text{ pb}^{-1}$  of data. All  $\tau_h$  candidates are required to have  $15 \text{ GeV} < p_T < 60 \text{ GeV}$ . Low  $p_T$  taus are difficult to identify under the dense background of QCD jets. The QCD background cross-section is generally orders of magnitude larger than expected tau signals, and it falls off exponentially with  $p_T$ . All distributions are normalized to unit area. The discriminating variables between hadronic taus and QCD jets are as follows:

**Electromagnetic radius:** Transverse energy weighted shower width in the electromagnetic calorimeters

$$R_{\text{EM}} = \frac{\sum_i^{\Delta R_i < 0.4} E_{\text{T},i} \Delta R_i}{\sum_i^{\Delta R_i < 0.4} E_{\text{T},i}},$$

where  $i$  runs over all cells associated to the tau candidate in the EM calorimeter presampler, and the first and second layer.  $\Delta R_i$  is defined between a calorimeter cell and the tau jet seed axis. This variable is a measure of the lateral shower spread in the EM calorimeter. QCD jets are broader than hadronic taus as can be seen from Fig. 9.2.  $R_{\text{EM}}$  is used to suppress background for both single and multi-prong taus.

**Track radius:**  $p_T$  weighted track distribution

$$R_{\text{track}} = \frac{\sum_i^{\Delta R_i < 0.4} p_{\text{T},i} \Delta R_i}{\sum_i^{\Delta R_i < 0.4} p_{\text{T},i}},$$

where  $i$  runs over all core and isolation tracks of the tau candidate. The distance  $\Delta R_i$  is defined relative to the tau seed axis, and  $p_{\text{T},i}$  is the track transverse momentum. For candidates with only one track in the core cone and isolation annulus,  $R_{\text{track}}$  simplifies to the  $\Delta R$  between the track and the axis of the seed jet. The discriminating power of this variable is useful for single-prong taus only, and is shown in Figure 9.3.

**Cluster mass:** Invariant mass of the constituent clusters of the tau candidate calibrated at the EM scale.

Although the reconstructed  $\tau_h$  mass is set to zero during reconstruction, the mass calculated from the constituent clusters composing a single-prong tau aids in discriminating against QCD jets. Figure 9.4 shows the cluster mass for single-prong tau candidates.

**Inverse leading track momentum fraction:** Total transverse momentum of a tau candidate divided by the transverse momentum of the leading core track:

$$f_{\text{track}} = \frac{p_{\text{T}}^{\tau}}{p_{\text{T}}^{\text{track},1}}.$$

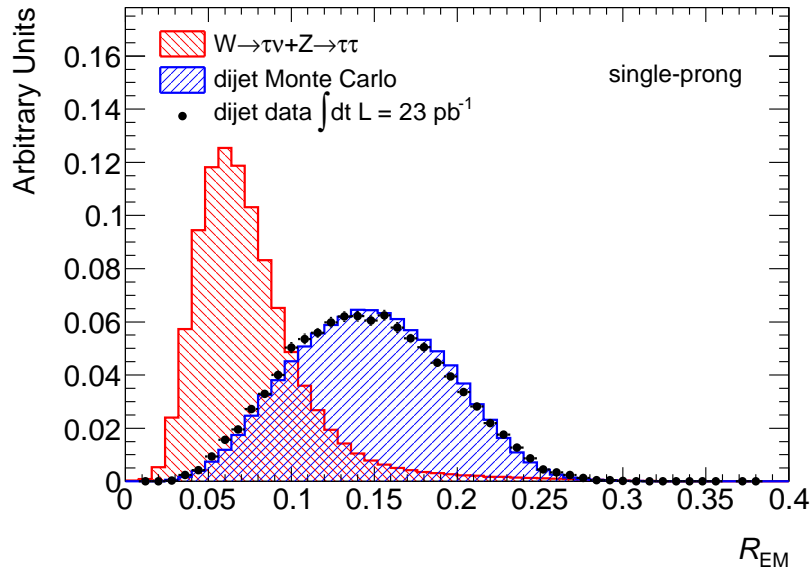
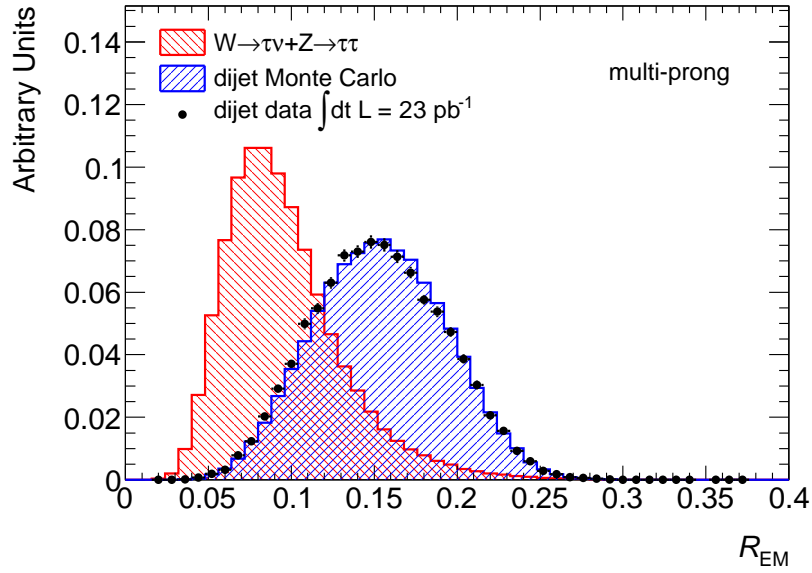
Here  $p_{\text{T}}^{\tau}$  is the total transverse momentum of the tau candidate as measured in the calorimeters using the assumption that the tau is massless. This term includes contributions from neutral pions.  $p_{\text{T}}^{\text{track},1}$  is that of the highest- $p_T$  single track. The variable distribution is shown in Figure 9.5 for multi-prong taus only, it is not used for single prong tau discrimination.

**Electromagnetic fraction:** Fraction of transverse energy of the tau candidate measured in the EM calorimeter

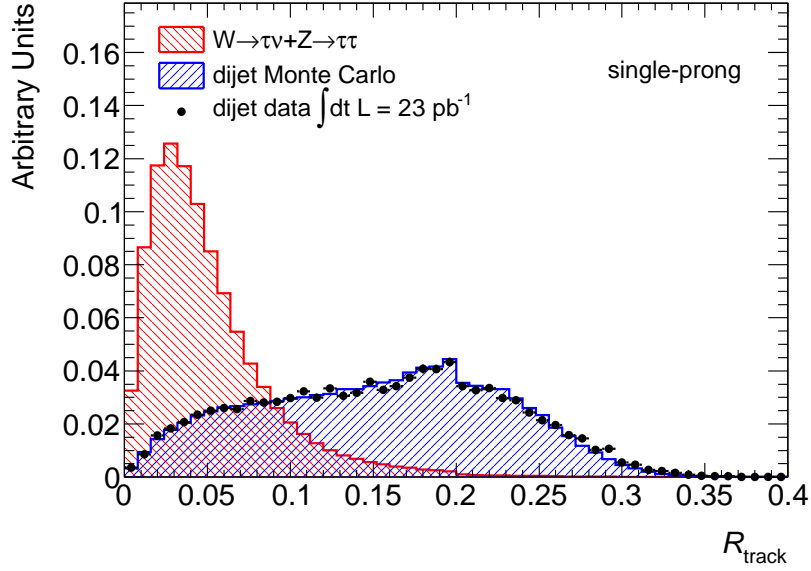
$$f_{\text{EM}} = \frac{\sum_i^{\Delta R_i < 0.4} E_{\text{T},i}}{\sum_j^{\Delta R_j < 0.4} E_{\text{T},j}}.$$

Where  $i$  runs over all cells associated to the tau candidate in the presampler and first two layers of the EM calorimeter, and  $j$  runs over the cells from all layers of both calorimeters. The cell energy is calibrated at the EM scale. The variable  $f_{\text{EM}}$  only gives significant discriminating power for multi-prong taus, and this distribution is shown in Figure 9.6.

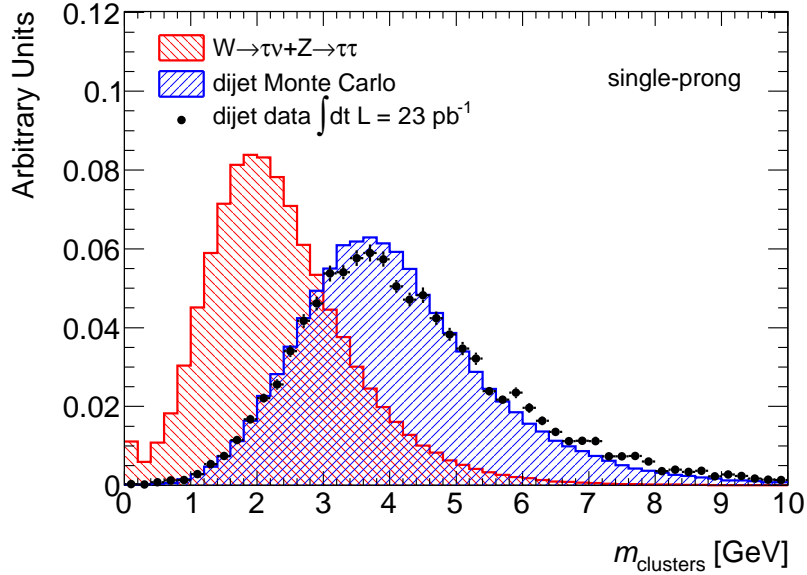
**Track mass:** Invariant mass of the tracks associated to a tau. Tracks are used from both the core and isolation regions, however this variable is only used to discriminate between multi-prong tau candidates and QCD jets. The corresponding distribution is shown in Figure 9.7.

(a)  $R_{EM}$  for single prong tau(b)  $R_{EM}$  for multi prong tau

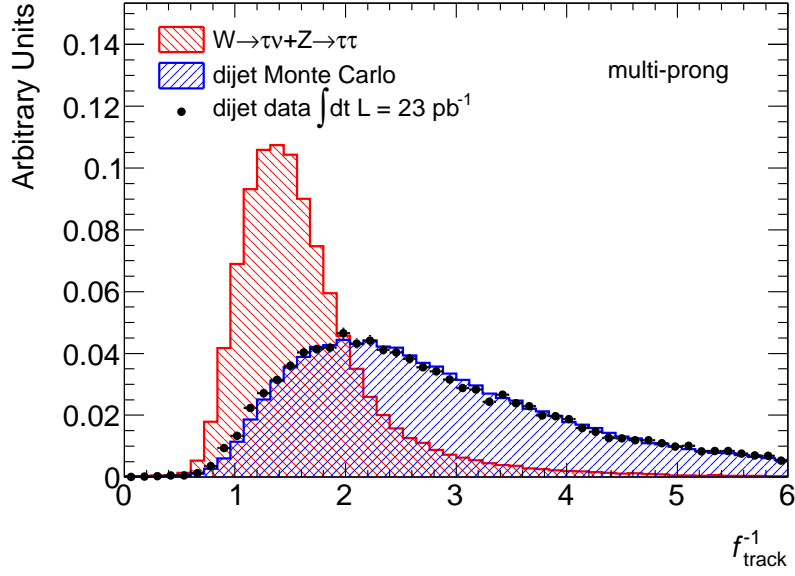
**Figure 9.2:** Electromagnetic radius,  $R_{EM}$ , distributions for tau candidates taken from  $Z \rightarrow \tau\tau$ ,  $W \rightarrow \tau\nu$ , and dijet PYTHIA generated Monte Carlo. Data points shown with markers are from dijet selection applied to data.



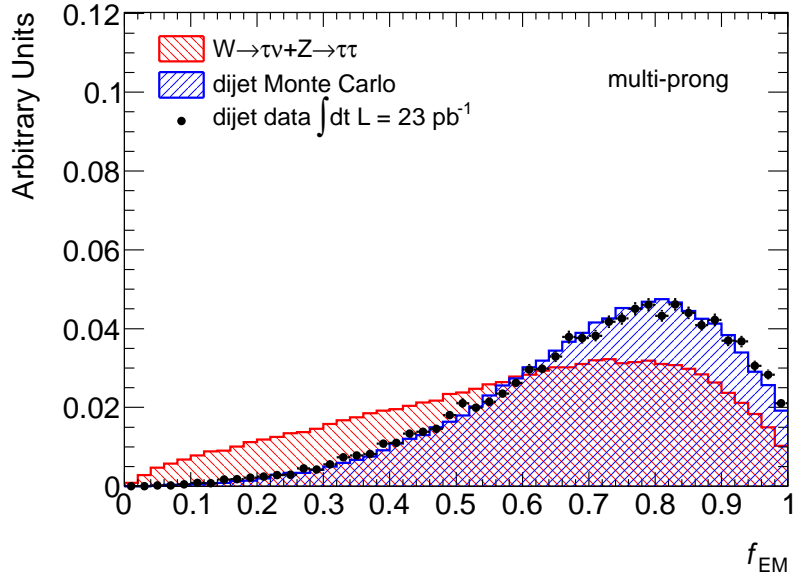
**Figure 9.3:** Track radius,  $R_{\text{track}}$ , distributions for single-prong tau candidates taken from  $Z \rightarrow \tau\tau$ ,  $W \rightarrow \tau\nu$ , and dijet PYTHIA generated Monte Carlo. Data points shown with markers are from dijet selection applied to data.



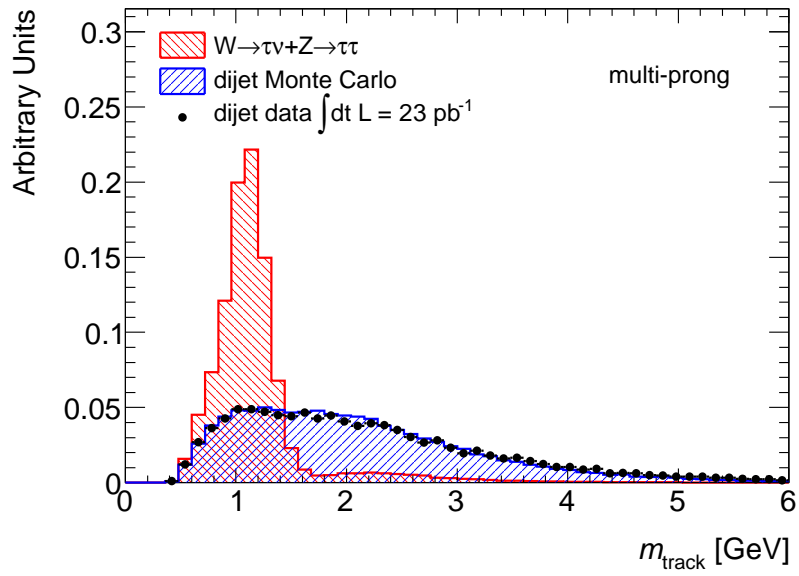
**Figure 9.4:** Mass constructed from topological clusters,  $m_{\text{clusters}}$ , that form hadronic tau. Distributions for single-prong tau candidates taken from  $Z \rightarrow \tau\tau$ ,  $W \rightarrow \tau\nu$ , and dijet PYTHIA generated Monte Carlo. Data points shown with markers are from dijet selection applied to data.



**Figure 9.5:** Total tau  $p_T$  from the calorimeters divided by the highest- $p_T$  track,  $f_{\text{track}}$ , distribution for multi-prong tau candidates. Distributions taken from  $Z \rightarrow \tau\tau$ ,  $W \rightarrow \tau\nu$ , and dijet PYTHIA generated Monte Carlo. Data points shown with markers are from dijet selection applied to data.



**Figure 9.6:** Electromagnetic fraction,  $f_{\text{EM}}$ , distribution for multi-prong tau candidates taken from  $Z \rightarrow \tau\tau$ ,  $W \rightarrow \tau\nu$ , and dijet PYTHIA generated Monte Carlo. Data points shown with markers are from dijet selection applied to data.



**Figure 9.7:** Mass constructed from tracks,  $m_{\text{tracks}}$ , distribution for multi-prong tau candidates taken from  $Z \rightarrow \tau\tau$ ,  $W \rightarrow \tau\nu$ , and dijet PYTHIA generated Monte Carlo. Data points shown with markers are from dijet selection applied to data.

## 9.2 Log likelihood ratio discriminant

The likelihood function for some variable  $x$  described by the probability density function  $p(x)$  is given by the joint probability density function in Equation 9.1.

$$L = \prod_{j=1} p(x_j), \quad (9.1)$$

The index  $j$  runs over all measurements of  $x$ . When considering the likelihood from multiple uncorrelated probability density functions, assigned the index  $i$ , this extends to:

$$L = \prod_{i=1} \prod_{j=1} p_i(x_j), \quad (9.2)$$

In tau identification the likelihood is constructed from the discriminating variables, under the assumption that these are uncorrelated. The discriminating variables are in histograms that are normalized to unity. For a particular discriminating variable, the measurement  $x_j$  has a value  $p(x_j) = n_j$ , where  $n_j$  is the number of entries in bin  $j$ . We define the likelihood function as given in Equation 9.3.

$$L_{S(B)} = \prod_{j=1}^N \prod_{i=1} p_i^{S(B)}(x_j). \quad (9.3)$$

The term  $p_i^{S(B)}(x_j)$  is the signal or background discriminating variable histogram for  $i$  of  $N$  variables, where  $p_i^{S(B)}(x_j)$  corresponds to some  $n_j$  for discriminating variable  $i$ . The function is constructed for signal ( $L_S$ ) and background ( $L_B$ ) individually. All identification variables used for  $\tau_h$  discrimination against QCD jets have been described in section 9.1, and are summarized in Table 9.1 for both single and multi-prong taus.

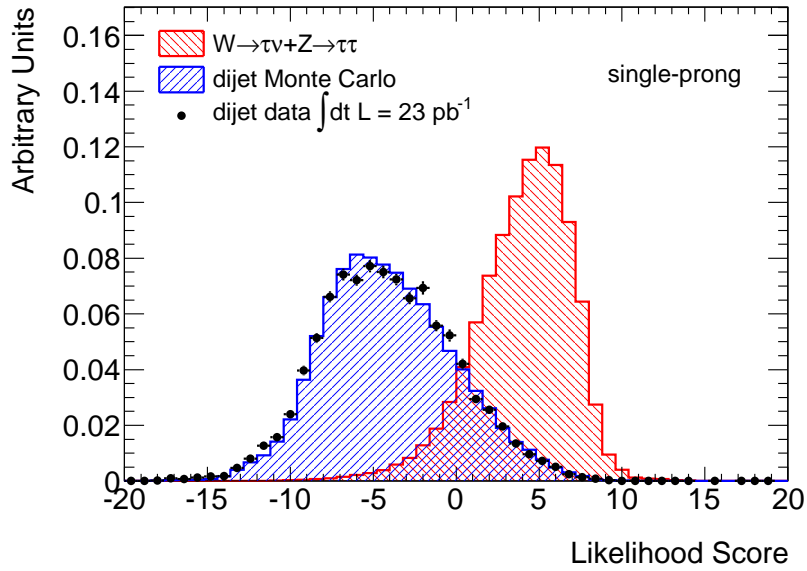
The test discriminant used is the log-likelihood-ratio between signal and background is given in Equation 9.4.

$$d = \ln \left( \frac{L_S}{L_B} \right) = \sum_{i=j} \sum_{i=1}^N \ln \left( \frac{p_i^S(x_j)}{p_i^B(x_j)} \right). \quad (9.4)$$

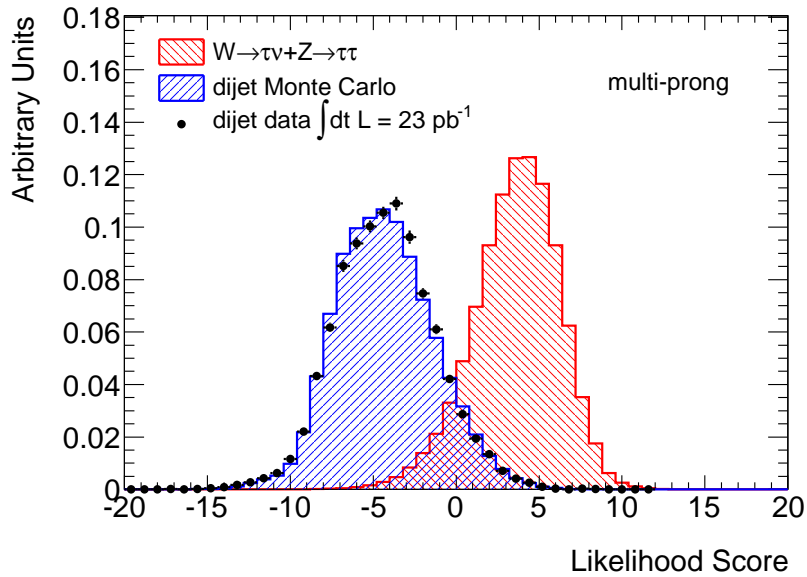
	$N$ (Number of variables)	$p_i$ (discriminating variables)
single-prong $\tau_h$	3	$R_{EM}, R_{\text{track}}, m_{\text{clusters}}$
multi-prong $\tau_h$	5	$R_{EM}, f_{\text{track}}, f_{\text{EM}}, m_{\text{tracks}}, S_{\text{T}}^{\text{flight}}$

**Table 9.1:** *Identification variables for single and multi-prong taus that are input for the likelihood functions.*

The log-likelihood-ratio is constructed in categories in order to maximize the method's discrimination power [50]. The categories are defined by the tau  $p_T$ ,  $\eta$ , and the number of vertices in an event. The  $\eta$  regions ( $0 < |\eta| \leq 1.3$ ,  $1.3 < |\eta| \leq 1.6$ ,  $1.6 < |\eta| \leq 2.5$ ) coarsely divide the barrel region of the detector from the endcap. The central bin corresponds to the transition region of the electromagnetic barrel and endcap calorimeters. There are three tau  $p_T$  bins (15-30, 30-60, 60+ GeV) giving equal statistics from  $Z \rightarrow \tau\tau$  and  $W \rightarrow \tau\nu$  Monte Carlo samples. Finally, in order to deal with the different pile-up conditions, the log-likelihood-ratio is produced for each category depending on the number of primary vertices in an event [50]. The resulting distributions from Equation 9.4, referred to as the Likelihood *score* are shown inclusively for low pile up tau candidates with  $15 \text{ GeV} < p_T < 60 \text{ GeV}$  in Fig. 9.8.



(a) single-prong taus



(b) multi-prong taus

**Figure 9.8:** Log-likelihood-ratio scores for tau candidates taken from  $Z \rightarrow \tau\tau$ ,  $W \rightarrow \tau\nu$ , and dijet PYTHIA generated Monte Carlo. Data points shown with markers are from dijet selection applied to data. All tau candidates displayed are required to have  $15 \text{ GeV} < p_T < 60 \text{ GeV}$ .

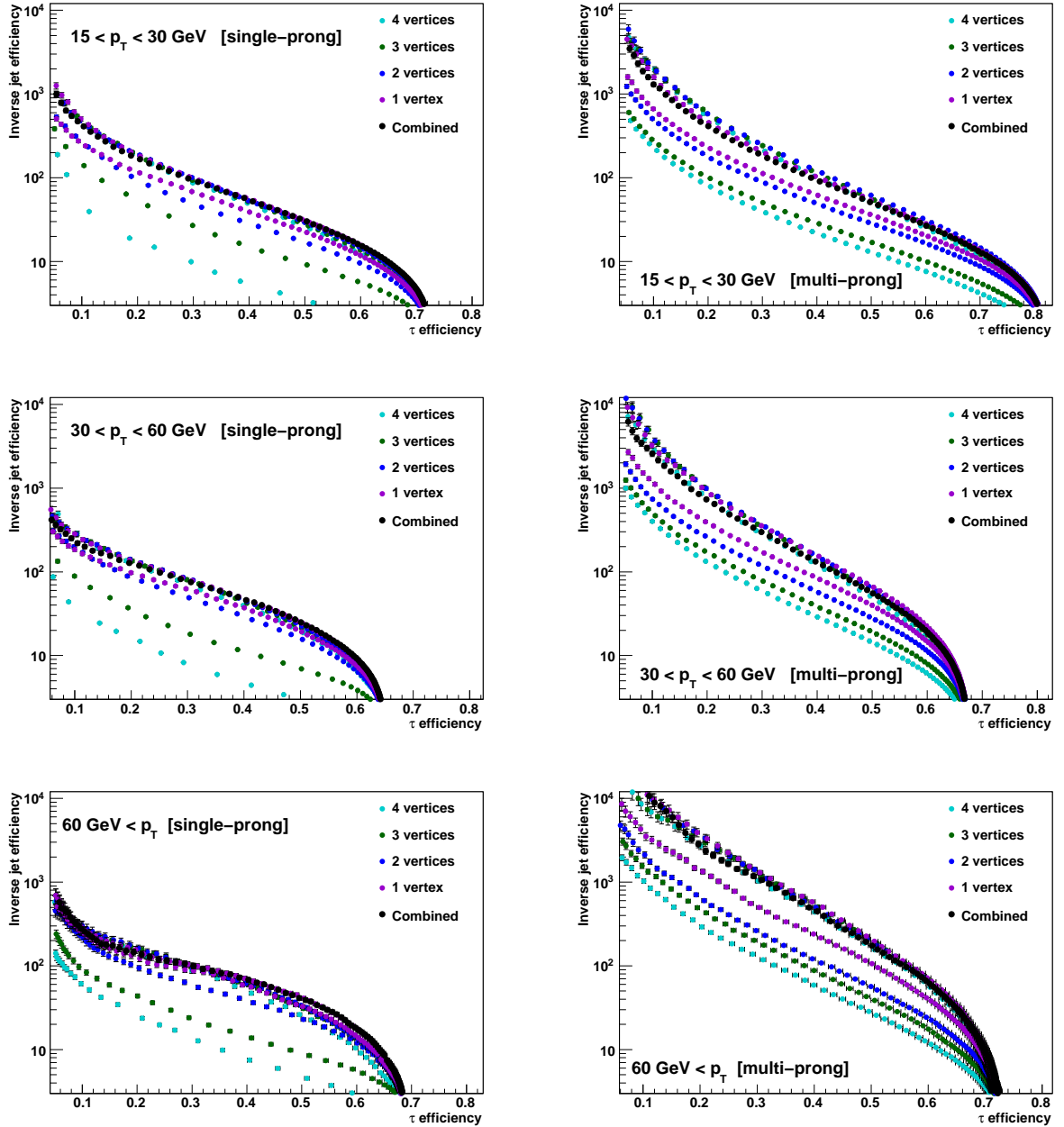
Three levels of identification, denoted **loose**, **medium**, and **tight**, are defined from the log-likelihood-score. These correspond to  $\tau_h$  efficiencies of 30%, 50%, and 60%. The efficiency is defined from Monte Carlo samples with real generated taus as:

$$\frac{\text{Number of reconstructed taus that are generator matched and pass identification}}{\text{Number of generated taus}}$$

Reconstructed taus are considered generator matched if there is a simulated hadronic tau shower within  $\Delta R < 0.2$ . The background efficiency, taken from dijet selected data events, is given as:

$$\frac{\text{Number of identified tau candidates}}{\text{Number of reconstructed tau candidates}}$$

Figure 9.9 displays the discriminant performance. Tau identification efficiency is shown against the inverse background efficiency. The reason the background efficiency is not simply referred to as the ‘rejection’ is because typically this term would be defined as the number of accepted tau candidates to rejected tau candidates. In most analyses it is more useful to consider the background efficiency since this involves the total number of reconstructed taus. Performance is shown for single-prong and multi-prong taus in three  $p_T$  bins, for inclusive  $\eta$ , separated for different numbers of primary vertices. The **medium** identification point (with 50% tau efficiency) is used in the bulk of this analysis, the **loose** point is used in QCD background estimation.



**Figure 9.9:** Log-likelihood-ratio efficiency against inverse of jet efficiency. Tau candidates taken from  $Z \rightarrow \tau\tau$ ,  $W \rightarrow \tau\nu$ , and jets from dijet PYTHIA generated Monte Carlo. Distributions are divided into single-prong (left column) and multi-prong (right column) taus. Tau candidates are shown in three  $p_T$  bins, from 15 - 30 GeV, 30 - 60 GeV and 60 GeV onwards. Each plot shows the effect of number of primary vertices on identification, the black markers show the inclusive result.

### 9.3 $\tau_h$ Identification Efficiency in Data

So far we have discussed the discrimination power given by the Monte Carlo. Data is used to verify these values and measure their uncertainty. The data driven approach employs a tag-and-probe method. When considering a two particle decay, where only one of the decay legs is of interest (the probe), we can employ stringent conditions on the other leg (the tag) to ensure we select appropriate events. The probe particle is largely independent of any requirements, providing an unbiased sample pool. The data based tag-and-probe analysis is outlined here, and full documentation is available in Reference [53].  $Z \rightarrow \tau\tau \rightarrow \mu\tau_h$  events were selected by an isolated muon trigger, giving an unbiased probe tau. The analysis was performed with  $730 \text{ pb}^{-1}$  of data, up to period **H**. Strict selection was applied to the reconstructed tag muon, and to event kinematics to ensure a pure sample of  $Z$  events. Minimal criteria were applied to the probe  $\tau_h$ . Object and event selection are listed in Table 9.2.

Event Selection	$\mu$ criteria	$\tau_h$ criteria
<ul style="list-style-type: none"> <li>• EF_mu18 trigger</li> <li>• Dilepton veto</li> <li>• <math>W</math> boson suppression</li> <li>• Opposite charged <math>\tau_h</math> and <math>\mu</math></li> <li>• <math>40 \text{ GeV} &lt; m_{\mu,\tau} &lt; 80 \text{ GeV}</math></li> </ul>	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 20 \text{ GeV}</math></li> <li>• <math> \eta  &lt; 2.5</math></li> <li>• No additional tracks within <math>\Delta R &lt; 0.2</math> of muon</li> <li>• Isolation requirement of <math>\frac{E_T(\Delta R &lt; 0.2)}{p_T} &lt; 0.04</math></li> </ul>	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 30 \text{ GeV}</math></li> <li>• <math> \eta  &lt; 2.5</math></li> <li>• Exactly 1 or 3 core tracks</li> <li>• Charge = <math>\pm 1</math></li> </ul>

**Table 9.2:** Selection for tau identification efficiency study with data.

Any events with selected probe taus within  $\Delta R < 0.2$  of an identified muon were eliminated to reduce contamination from  $Z \rightarrow \mu\mu$  processes with a mis-reconstructed  $\tau$  candidate. For events with multiple probe taus, the highest- $p_T$  tau was used. The EF\_mu18 trigger was the lowest  $p_T$  un-prescaled muon trigger available in the data taking

period. All events containing more than one muon or identified electron with  $p_T > 10$  GeV were rejected. Background from  $Z \rightarrow ee$  and  $Z \rightarrow \mu\mu$  is reduced by 60% with this selection, with negligible effect on signal.  $W \rightarrow \mu\nu$  suppression was based on the transverse mass,  $m_T$ , which is defined as  $\sqrt{2p_T^\mu E_T^{Miss} \cos(\Delta\phi)}$ . The angle  $\Delta\phi$  is the separation in the transverse plane between the tagged muon and event  $E_T^{Miss}$ . For a process such as  $W$  decay, where one decay leg is a neutrino,  $m_T$  tends to be larger than other two body decays, with a maximum attainable value of  $m_W$ . Specifically the  $W$  boson suppression requirement was  $m_T < 50$  GeV, which reduced  $W \rightarrow \mu\nu$  events to 12.5% of the original number, and retained 90% of signal events. The visible mass,  $m_{\mu,\tau}$ , is defined as the invariant mass of the probe tau and tagged muon. This mass term does not contain any contribution from the neutrinos associated to the tau decay. Events are kept if the visible mass has a value between 40 and 80 GeV. This window includes the bulk of the signal, but excludes background  $Z \rightarrow ee$  and  $Z \rightarrow \mu\mu$  events which are distributed tightly around the  $Z$  boson mass peak.

Despite the event selection requirements it is difficult to remove most background without any criteria on tau identification. The probe tau sample is contaminated with fake taus, and no more than 50% purity is attained. An accurate estimation of background contributions is necessary in order to subtract the fakes from the tau samples. All background estimations, other than QCD, come from a Monte Carlo based approach, as detailed in Reference [53]. Although  $W$  backgrounds are estimated from Monte Carlo, they are corrected to the expected rate in data, while QCD background is estimated directly from data. Once events and objects are selected the **medium** log-likelihood-ratio identification is applied to the probe taus, to obtain a sample of identified  $\tau_h$  objects. The identification efficiency is given by the ratio of identified taus to probe taus.

### 9.3.1 $W$ background scale factors

The probability for a QCD jet to be misidentified as a tau is known to be overestimated in the Monte Carlo simulation [51, 52]. This affects the Monte Carlo based prediction of  $W$ +jet background events. To correct the estimation, the number of  $W$  background events must be scaled to match the observation in data. A  $W$  control region is defined in data, containing events that pass the dilepton event selection but fail the  $W$  suppression cuts. To ensure adequate statistics none of the subsequent event level cuts are applied. The high transverse mass cut,  $m_T > 50$  GeV, makes the QCD jet background negligible in the control region. Other small backgrounds,  $Z \rightarrow ll$  and  $t\bar{t}$ , are removed by using predictions from Monte Carlo. The visible mass distribution from  $W \rightarrow \mu\nu/\tau\nu/e\nu$  Monte Carlo samples are then collectively scaled by a factor  $k_W$ , such that the predicted number of  $W$  events in the Monte Carlo is equal to the events observed in the visible mass from the data control region. The scale factors are derived for the probe tau sample, using the method just described, and also for samples with the additional requirement of **medium** identification on the selected probe taus. This allows us to correct for the  $W$  background expected in both the probe and identified  $Z$  samples individually.

### 9.3.2 QCD background

Multijet background in the selected  $Z$  events is estimated from data directly using what is called the ABCD method. The ABCD method divides the data into four sections using two uncorrelated parameters. One of these sections, referred to as region **A** in this analysis, contains the expected signal and is the phase space in which we want to estimate the shape and normalization of the background. Here the shape refers to any kinematic distribution or constructed mass distribution from the background, and the normalization is the overall number of events. The shapes can be taken from either of region **B** or **C**, where the choice usually depends on the available statistics. The normalization is taken from the ratio of background events in regions **B** to **D**, or **C** to **D**, depending on which region is used to extract background shapes. Table 9.3 illustrates the independent data regions constructed from two uncorrelated parameters (labelled 1 and 2), satisfying some requirements, generically referred to X and Y.

	Parameter 1 satisfying requirement X	Parameter 1 satisfying requirement NOT X
Parameter 2 satisfying requirement Y	<b>A</b> (signal)	<b>B</b>
Parameter 2 satisfying requirement NOT Y	<b>C</b>	<b>D</b>

**Table 9.3:** *Division of data based on the ABCD method. The method works provided that parameter 1 and 2 are uncorrelated.*

In this analysis, the parameters used to divide the data are the muon isolation, and the sign of the tag muon and probe tau. The key assumption in the ABCD method is that the  $\mu$  isolation is uncorrelated to the charge sign of the  $\mu$  and  $\tau$ . An isolated muon is defined by  $\frac{E_T(\Delta R < 0.2)}{p_T} < 0.04$ ; the calorimeter section traversed by the muon should

contain less than 4% of the muon transverse momentum. Specifically the regions are defined as:

**Region A:** Isolated muon and opposite signed charges for muon and  $\tau_h$  candidate

**Region B:** Non-isolated muon and opposite signed charges for muon and  $\tau_h$

**Region C:** Isolated muon and same signed charges for muon and  $\tau_h$

**Region D:** Non-isolated muon and same signed charges for muon and  $\tau_h$

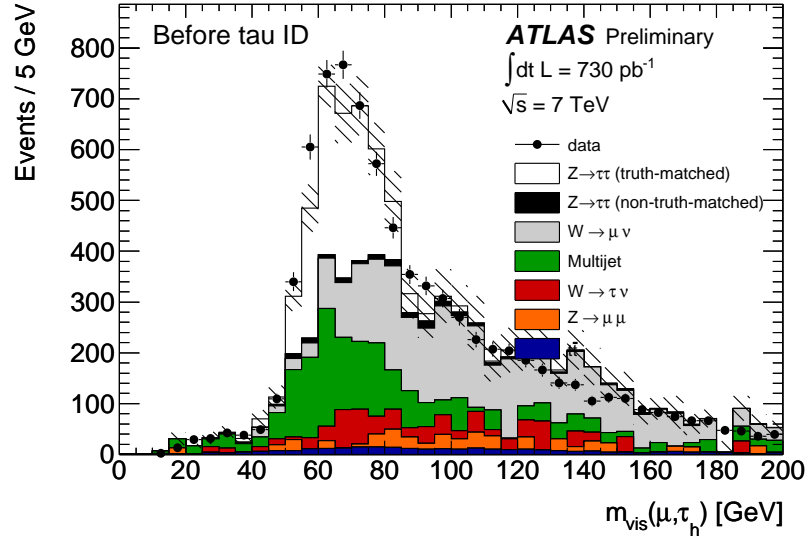
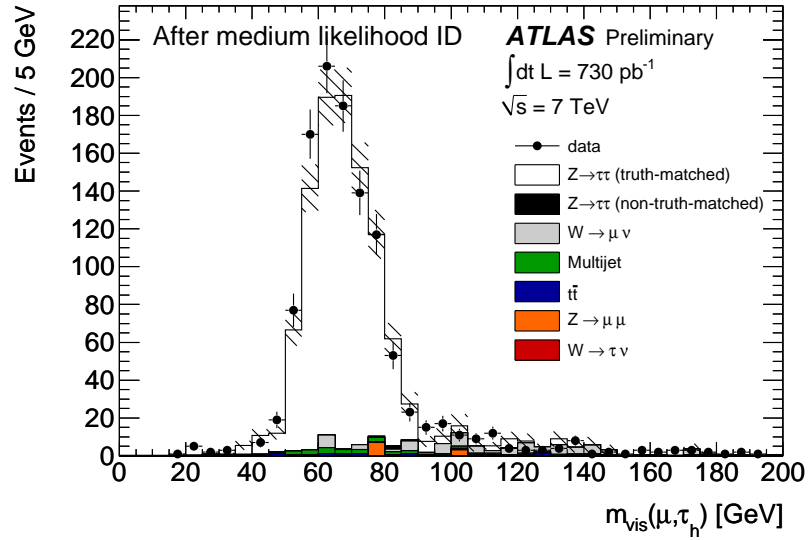
Region A contains the  $Z$  signal. The visible mass distribution from QCD is taken from region C. All remaining background distributions and normalizations in each region, including  $Z \rightarrow \mu\mu$ ,  $t\bar{t}$  and  $Z \rightarrow \tau\tau$ , come directly from Monte Carlo. The number of QCD background events in the signal region is predicted by taking the number in region C, and scaling by the ratio of events in region B to D.<sup>1</sup>:

$$N_A = N_C \left( \frac{N_B}{N_D} \right)$$

where  $N_i$ , for  $i = A, B, C$  and D, indicates the number of events in the corresponding region. As with the  $W$  scale factors, these values were calculated separately with and without  $\tau_h$  identification for background expected in both the probe and identified  $Z$  samples individually. The difference in statistics between the combined regions **BD** and **AC** is treated as a systematic uncertainty, denoted  $\Delta R_{bias}$ . Visible mass distributions after the event selection for probe and identified  $\tau_h$  candidates are shown in Figure 9.10.

---

<sup>1</sup>This is different to the procedure used in [53] for  $Z$  observation, where the initial number was taken from region B and scaled by the ratio in region C and D.

(a) Probe  $\tau_h$  candidates(b) Identified  $\tau_h$ 

**Figure 9.10:** Visible mass distributions,  $m_{\mu\tau_h}$ , after the  $\tau_h$  efficiency event selection for probe and identified  $\tau_h$  candidates.

The final  $\tau_h$  identification efficiency is calculated as:

$$\epsilon = \frac{N_D^P - N_W^P \Delta k_W^P - N_{QCD}^C \frac{N_{QCD}^{B,P}}{N_{QCD}^D} - N_{other}^P}{N_D^{P+F} - N_W^{P+F} \Delta k_W^{probe} - \frac{N_{QCD}^C}{N_{QCD}^D} \left( N_{QCD}^{B,P+F} \right) - N_{other}^{P+F}}, \quad (9.5)$$

$N_D$  gives the number of events found after event selection in data.  $N_W$  is the estimated number of  $W$  events, which only presents itself scaled by  $\Delta k_W$ . The terms  $N_{QCD}^B$ ,  $N_{QCD}^C$  and  $N_{QCD}^D$  are the number of QCD background events passing selection in regions B, C and D.  $N_{other}$  gives the number of non-truth matched events from  $Z \rightarrow \mu\mu$ ,  $t\bar{t}$  and  $Z \rightarrow \tau\tau$  Monte Carlo. The superscripts  $P$  and  $F$  refer to events with tau candidates passing or failing the **medium** log-likelihood-ratio discriminant.

There are two main sources of uncertainty in the study. The first comes from the statistical uncertainty on the  $W$  scale factors,  $\Delta k_W^{probe}$  and  $\Delta k_W^P$ . The second from the statistical uncertainty in the QCD background regions,  $\Delta R_{bias}^{probe}$  and  $\Delta R_{bias}^P$ . These are summarized in Table 9.4, which also give the statistical uncertainty for Monte Carlo and data.

$\epsilon_{MC}$	$\Delta\epsilon_{MC}^{stat} (\%)$	$\epsilon_{data}$	$\Delta\epsilon_{data}^{stat} (\%)$	$\Delta\epsilon_{data}^{k_W} (\%)$	$\Delta\epsilon_{data}^{R_{bias}} (\%)$	$\Delta\epsilon_{data}^{tot} (\%)$
0.536	$\pm 2.4$	0.497	$\pm 7.4$	$\pm 0.9$	$\pm 3.4$	$\pm 8.5$

**Table 9.4:** **Medium** tau identification efficiency in Monte Carlo,  $\epsilon_{MC}$ , and in data,  $\epsilon_{data}$ . Statistical uncertainty for Monte Carlo, and uncertainty for data are given.

In the analysis for  $A/H/h \rightarrow \tau_h \tau_h$ , the difference between tau identification efficiency predicted from Monte Carlo and measured in data is treated as a systematic uncertainty in this analysis. This is discussed with other sources of uncertainty in Chapter 15.

# Chapter 10

## Event Selection

This analysis searches for an MSSM Higgs through the process  $A/H/h \rightarrow \tau_h \tau_h$ . The search method relies on the mass constructed from the two hadronic taus, the visible mass. Reconstructing the invariant mass is difficult since the individual momenta carried by the neutrinos in the tau decays are unknown. Only the summed  $p_T$  of the two neutrinos is available from the  $E_T^{Miss}$  term in an event. Techniques to form the invariant mass exist, assuming the neutrinos from tau decay travel in the same direction as the tau. However, the methods place restrictions on the angular separation between the two hadronic taus, and also require confidence in the measured  $E_T^{Miss}$ . Criteria on the separation of the taus reduces the available statistics substantially. Additionally, to ensure the  $E_T^{Miss}$  term is correct, all other particles must be measured accurately. In particular the calibration of hadronic particles is very important, since any mis-calibration exhibits itself in the measurement of  $E_T^{Miss}$ . Due to these constraints, at the time of this analysis it was decided to consider the visible mass.

Before beginning with event selection, we should note that a given particle in the detector can be identified as different reconstructed objects. To select a single reconstructed object, ‘overlap’ removal is applied. This is achieved by removing reconstructed objects that coincide geometrically within  $\Delta R < 0.2$ . If a muon and jet are found in the

same location, the object is defined as a muon and no longer considered a jet. The overlap removal in this analysis prioritizes muons, followed by electrons, **medium** identified hadronic taus, and then jets.

Event selection is optimized for a neutral MSSM Higgs boson mass greater than 200 GeV. The search for a lower Higgs mass is difficult due to the irreducible  $Z/\gamma^* \rightarrow \tau\tau$  background. Below are given the criteria for selected events. Following this Table 10.1 presents the results of this selection on data and Monte Carlo, including expected signal for  $m_A = 200$  GeV and  $\tan\beta = 20$ . Table 10.2 gives results for data, expected signal, and combined Monte Carlo background. QCD background is unaccounted for in both Table 10.1 and 10.2. The data driven background estimates for QCD are presented in Section 13.

**Data quality.** We only use data within a run when the two proton beams are colliding bunches and all sub-detectors are recording data. Events that meet this criteria are stored in a good runs list (GRL)<sup>1</sup>. This is in contrast to data stored where the beams are still being accelerated to their full energy, or a sub-detector malfunctions.

**Event trigger.** All events in this analysis have satisfied the EF trigger

‘EF\_tau29\_medium1\_tau20\_medium1.’ The selection uses all three ATLAS trigger levels: Level-1, 2 and the event filter (EF) [20]. The associated L1 trigger (L1\_2TAU8\_TAU11) selects two narrow clusters in the calorimeters with a  $p_T$  threshold of 8 GeV and 11 GeV. At L2, tracks are reconstructed around the L1 clusters. L2 track selection differs between data periods **B<sub>2</sub>-E** and **F-H**. In the later periods, tracks were selected more rigorously to comply with trigger rate requirements for higher luminosity. During the EF stage, two reconstructed  $\tau_h$  candidates with  $p_T$  thresholds of 29 GeV and 20 GeV, both within  $|\eta| < 2.5$ , are required. The ‘medium1’ indicates comparable selections as the offline **medium** tau identifica-

---

<sup>1</sup>The GRL used in this analysis is: `data11_7TeV.periodAllYear_DetStatus-v18pro08-04_CoolRunQuery-00-03-98_Higgs_tautau.lh.xml`.

tion. This choice is the lowest  $p_T$  un-prescaled dihadronic tau trigger available in the data taking periods. Trigger efficiency and fake rates are analyzed with Monte Carlo and data, and will be described in Chapter 11.

**Event cleaning.** In addition to the good quality data requirements from the GRL, there are three other criteria the event must satisfy. The selection is designed to ensure a high rejection of events without proton-proton collisions.

Firstly, an event must contain at least one primary vertex from which originate at least four quality tracks.

Secondly, we reject events where the electronics in the LAr have malfunctioned causing significant noise bursts and other data integrity errors. Events containing such errors are flagged by data quality algorithms [58], and vetoed.

Finally, during data taking period **E** the electronics corresponding to part of the LAr calorimeter malfunctioned, and lost output signal [35]. This region, a strip along  $-0.05 < \eta < 1.45$  at  $\phi \sim -0.7$ , can contribute to fake  $E_T^{Miss}$  if a jet or  $\tau_h$  candidate points in its direction. Therefore an event is rejected if a jet of  $p_T > 20$  GeV is reconstructed inside the affected region,  $-0.1 < \eta < 1.5$  and  $-0.9 < \phi < -0.5$ , and also if at least one of the two leading  $\tau_h$  candidates is within  $-0.1 < \eta < 1.55$  and  $-0.9 < \phi < -0.5$ .

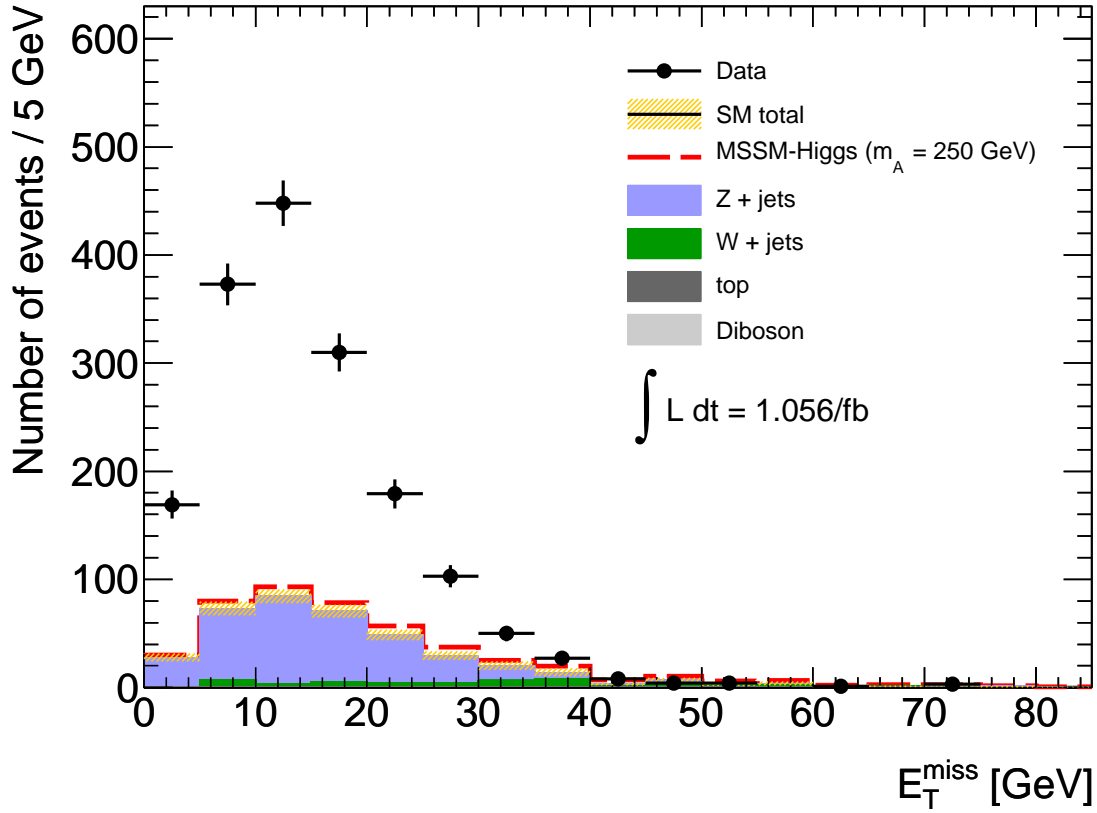
**Electron and muon veto.** In this analysis reconstructed electrons are required to pass what is termed the **medium** identification requirements [35]. This offers  $\sim 80\%$  efficiency for correctly identifying reconstructed electrons, and a jet rejection of  $10^3$ . Identification is based on electromagnetic particle shower shapes in the calorimeters, which are shorter and narrower in comparison to quark jets, denser in energy, and contained predominantly in the EM calorimeter [56]. Only electrons with  $p_T > 15$  GeV, within  $|\eta| < 2.5$  (and excluded from the electromagnetic calorimeter transition region,  $1.37 < |\eta| < 1.52$ ) are considered. Muons used in this analysis are required

to have  $p_T > 10$  GeV and  $|\eta| < 2.5$ . The over all efficiency for muon reconstruction is  $\sim 95\%$ , with a high rejection of all backgrounds. To suppress background events containing real electrons or muons, an event containing either a selected electron or muon is rejected.

**Selection of hadronic tau decays.** Two  $\tau_h$  candidates must be reconstructed in an event, both satisfying the **medium** log-likelihood-ratio requirements. The **medium** working point is chosen as a compromise between signal selection efficiency and background suppression. **Tight** identification results in a very small data sample and very few data events in the background control regions, limiting the precision of the QCD background estimation in Chapter 13. Only taus with one or three tracks, with a charge of  $\pm 1$ , are considered. An event is rejected if either leading  $\tau_h$  candidate is reconstructed in the pseudorapidity range  $1.37 < |\eta| < 1.52$ , corresponding to the transition region of the electromagnetic calorimeter. This avoids any mis-reconstruction resulting in fake  $E_T^{Miss}$ . Note that tau reconstruction necessarily is confined to  $|\eta| < 2.5$ , the range of the ID. Finally, the two leading taus must satisfy  $p_T > 45$  GeV and  $p_T > 30$  GeV.

If there are three or more  $\tau_h$  candidates satisfying all the requirements above, with the exception of the third tau meeting  $p_T > 20$  GeV, rather than 45 or 30 GeV, the event is also rejected. The two selected  $\tau_h$  candidates must have opposite charges, and each must match one of the two trigger RoIs within a cone of  $\Delta R < 0.2$ .

$E_T^{Miss} > \mathbf{25\ GeV}$ . Missing energy is expected in signal events from the neutrinos in  $\tau$  decays. From Monte Carlo simulation we know that  $E_T^{Miss} > 25$  GeV suppresses the QCD jet background, but retains adequate signal. Figure 10.1 shows the  $E_T^{Miss}$  distribution for data,  $Z/\gamma^* \rightarrow \tau\tau + \text{jets}$ ,  $W + \text{jets}$ , top and di-boson backgrounds. The difference between estimated Monte Carlo backgrounds and data is associated with the QCD background.



**Figure 10.1:**  $E_T^{\text{Miss}}$  distribution for data,  $Z/\gamma^* \rightarrow \tau\tau + \text{jets}$ ,  $W + \text{jets}$ ,  $\text{top}$  and  $\text{diboson}$  backgrounds after all event selection criteria other than  $E_T^{\text{Miss}}$  cut. Also included is the potential signal from an MSSM Higgs with  $m_A = 200$  GeV and  $\tan\beta = 20$ . The difference between the Monte Carlo backgrounds and data corresponds to the QCD background.

## 10.1 Scale factors

The Monte Carlo does not model the data perfectly, and for an effective analysis we apply various corrections to the simulated data. In most situations the Monte Carlo describes the shape of various distributions well, but not the rate of events. To correct for this, we apply a numerical scale factor to the Monte Carlo simulation. All scale factors in this analysis are described below.

**Pile-up** The average number of interactions per bunch crossing varies considerably during the course of data taking. It was not possible to generate adequate Monte Carlo statistics with pile-up conditions for all data periods. To scale the Monte Carlo to data, a re-weighting procedure based on the average number of bunch crossings per interaction is constructed, described in Reference [59]. This re-weighting procedure is applied in this analysis.

**Trigger efficiency** The Monte Carlo does not model the EF\_tau29\_medium1\_tau20\_medium1 performance in data accurately. To correct the Monte Carlo triggering of identified and generator matched taus, scale factors are constructed in Section 11.2.

**Trigger fake rates:** To correct the Monte Carlo triggering of mis-identified taus (fake taus), scale factors are constructed in Section 11.3. These are applied to taus reconstructed and identified in Monte Carlo, but not matched to a generator level tau.

## 10.2 Summary of Event selection

In summary, the key steps in the selection of candidate MSSM  $A/H/h \rightarrow \tau_h \tau_h$  events in data are:

1. Good quality data, dictated by beams colliding at full energy and all sub-detectors operational.
2. EF\_tau29\_medium1\_tau20\_medium1 trigger accept. This is a combination of two  $\tau_h$  triggers with 29 GeV and 20 GeV  $p_T$  thresholds.
3. Event cleaning that removes LAr detector malfunctions and requires a primary vertex.
4. No electrons reconstructed with  $p_T > 15$  GeV or muons reconstructed with  $p_T > 10$  GeV in the event.
5. First  $\tau_h$  requirements:
  - The highest- $p_T$   $\tau_h$  candidate in event must have  $p_T > 45$  GeV.
  - It must be identified as a **medium** candidate by the log-likelihood-ratio discriminant.
  - It must match either one of the  $\tau_h$  trigger RoIs within a cone of  $\Delta R < 0.2$ .
6. Second  $\tau_h$  requirements:
  - The second highest- $p_T$   $\tau_h$  candidate in event must have  $p_T > 30$  GeV.
  - It must be identified as a **medium** candidate by the log-likelihood-ratio discriminant.
  - It must match the remaining  $\tau_h$  trigger RoI within a cone of  $\Delta R < 0.2$ .
7. Reject the event if there are any additional identified **medium**  $\tau_h$  candidates with  $p_T > 20$  GeV
8. The two selected taus must have opposite charge.
9. Reconstructed  $E_T^{Miss} > 25$  GeV.

Tables 10.1 and 10.2 list the number of data events, simulated signal events, electroweak and top background events after applying each of the requirements described above. The uncertainties correspond only to the Monte Carlo statistical uncertainty. The selection results in 245 events in the data sample for an integrated luminosity of  $1.056 \text{ fb}^{-1}$ . From Monte Carlo simulation, the expected number of background events from electroweak processes and top is  $76 \pm 7$ . The largest contribution from electroweak background events after all selection criteria is from  $W \rightarrow \tau_h \nu$  + one or two jets and  $Z \rightarrow \tau_h \tau_h$  + zero jets. Electroweak background events with electrons or muons in the final signature are strongly suppressed by the lepton vetoes, with very few events satisfying the full event selection. The QCD background cannot be determined from simulated data because of a lack of Monte Carlo simulation and because of the large uncertainties on the QCD jet cross sections. This background contribution is estimated from data in Chapter 13. The full list of Monte Carlo samples used to estimate the electroweak and top backgrounds are in Appendix E.

	Data	$b\bar{b} \text{ A} \rightarrow \tau_h \tau_h$	$gg \rightarrow \text{A} \rightarrow \tau_h \tau_h$	W + jets	Z/ $\gamma^*$ + jets	Top	Di-boson
No cuts	$1.9 \times 10^8$	2622	1136	$3322 \times 10^4$	$15740 \times 10^3$	$2139 \times 10^2$	10725
Trigger/GRL	$8.7 \times 10^6$	$369.9 \pm 5.8$	$147.4 \pm 1.9$	$48195.4 \pm 272.9$	$(3428.4 \pm 3.9) \times 10^2$	$3958.7 \pm 36.0$	$891.8 \pm 8.1$
Event cleaning	$8.3 \times 10^6$	$344.3 \pm 5.6$	$136.2 \pm 1.8$	$44983.8 \pm 264.0$	$(3168.4 \pm 3.8) \times 10^2$	$3410.3 \pm 33.3$	$815.2 \pm 7.8$
Electron veto	$8.3 \times 10^6$	$341.6 \pm 5.6$	$135.8 \pm 1.8$	$44327.7 \pm 263.0$	$(3163.2 \pm 3.8) \times 10^2$	$2922.1 \pm 30.8$	$763.3 \pm 7.6$
Muon veto	$7.8 \times 10^6$	$331.3 \pm 5.5$	$131.3 \pm 1.8$	$11759.6 \pm 133.6$	$18493.5 \pm 91.4$	$850.9 \pm 16.6$	$68.8 \pm 2.5$
First $\tau_h$ candidate	$7.8 \times 10^6$	$331.3 \pm 5.5$	$131.3 \pm 1.8$	$11758.8 \pm 133.6$	$18493.2 \pm 91.4$	$850.9 \pm 16.6$	$68.8 \pm 2.5$
First $\tau_h$ $p_T > 45$ GeV	$3.9 \times 10^6$	$273.0 \pm 4.8$	$109.9 \pm 1.7$	$7622.2 \pm 99.9$	$7604.9 \pm 59.1$	$817.7 \pm 16.3$	$51.3 \pm 2.2$
First $\tau_h$ LLH medium	85311	$119.9 \pm 3.2$	$42.5 \pm 1.0$	$681.8 \pm 29.6$	$1352.5 \pm 24.8$	$30.0 \pm 2.9$	$8.4 \pm 0.7$
First $\tau_h$ trigger-matched	85197	$119.7 \pm 3.2$	$42.4 \pm 1.0$	$680.1 \pm 29.6$	$1347.4 \pm 24.7$	$29.5 \pm 2.9$	$8.4 \pm 0.7$
Second $\tau_h$ selection	85135	$119.7 \pm 3.2$	$42.4 \pm 1.0$	$678.0 \pm 29.5$	$1346.5 \pm 24.7$	$29.5 \pm 2.9$	$8.4 \pm 0.7$
Second $\tau_h$ $p_T > 30$ GeV	69258	$103.0 \pm 2.9$	$37.7 \pm 1.0$	$526.2 \pm 25.8$	$940.8 \pm 20.8$	$29.1 \pm 2.9$	$7.0 \pm 0.6$
Second $\tau_h$ LLH medium	3361	$53.8 \pm 2.1$	$16.6 \pm 0.7$	$47.5 \pm 6.7$	$328.1 \pm 12.2$	$2.9 \pm 1.0$	$2.1 \pm 0.4$
Second $\tau_h$ trigger-matched	3354	$53.8 \pm 2.1$	$16.6 \pm 0.7$	$47.5 \pm 6.7$	$326.0 \pm 12.2$	$2.9 \pm 1.0$	$2.1 \pm 0.4$
Third $\tau_h$ veto	3352	$53.8 \pm 2.1$	$16.6 \pm 0.7$	$47.5 \pm 6.7$	$326.0 \pm 12.2$	$2.9 \pm 1.0$	$2.1 \pm 0.4$
Opposite charge	1985	$53.4 \pm 2.1$	$16.5 \pm 0.7$	$36.3 \pm 5.8$	$319.6 \pm 12.1$	$2.8 \pm 0.9$	$2.0 \pm 0.3$
$E_T^{\text{miss}} > 25$ GeV	245	$17.9 \pm 1.2$	$5.2 \pm 0.3$	$24.9 \pm 4.8$	$47.6 \pm 4.6$	$2.0 \pm 0.9$	$1.4 \pm 0.3$

**Table 10.1:** Number of events passing the selection criteria for data and expected values for Monte Carlo signal, electroweak and top backgrounds normalized to the integrated luminosity of  $1.056 \text{ fb}^{-1}$ . The MSSM Higgs signal is shown in the  $m_h^{\text{max}}$  scenario for  $m_A = 200 \text{ GeV}$  and  $\tan \beta = 20$ .

	Data	Signal	Total MC bkg
No cuts	$1.9 \times 10^8$	3758	$4919 \times 10^4$
Trigger/GRL	$8.7 \times 10^6$	$517.3 \pm 6.1$	$(3958.8 \pm 4.8) \times 10^2$
Event cleaning	$8.3 \times 10^6$	$480.6 \pm 5.9$	$(3660.5 \pm 4.6) \times 10^2$
Electron veto	$8.3 \times 10^6$	$477.5 \pm 5.9$	$(3643.3 \pm 4.6) \times 10^2$
Muon veto	$7.8 \times 10^6$	$462.6 \pm 5.8$	$31172.8 \pm 162.7$
First $\tau_h$ candidate	$7.8 \times 10^6$	$462.6 \pm 5.8$	$31171.8 \pm 162.7$
First $\tau_h$ $p_T > 45$ GeV	$3.9 \times 10^6$	$382.9 \pm 5.1$	$16096.0 \pm 117.2$
First $\tau_h$ LLH medium	85311	$162.4 \pm 3.4$	$2072.7 \pm 38.7$
First $\tau_h$ trigger-matched	85197	$162.1 \pm 3.3$	$2065.4 \pm 38.7$
Second $\tau_h$ selection	85135	$162.1 \pm 3.3$	$2062.4 \pm 38.6$
Second $\tau_h$ $p_T > 30$ GeV	69258	$140.6 \pm 3.1$	$1503.1 \pm 33.3$
Second $\tau_h$ LLH medium	3361	$70.4 \pm 2.2$	$380.7 \pm 14.0$
Second $\tau_h$ trigger-matched	3354	$70.3 \pm 2.2$	$378.5 \pm 14.0$
Third $\tau_h$ veto	3352	$70.3 \pm 2.2$	$378.5 \pm 14.0$
Opposite charge	1985	$69.9 \pm 2.2$	$360.6 \pm 13.4$
$E_T^{\text{miss}} > 25$ GeV	245	$23.2 \pm 1.2$	$75.9 \pm 6.6$

**Table 10.2:** Number of events passing the selection criteria for data and expected values for Monte Carlo signal ( $m_A = 200$  GeV and  $\tan \beta = 20$ ), EW and top backgrounds normalized to the integrated luminosity of  $1.056 \text{ fb}^{-1}$ .

# Chapter 11

## Trigger scale factors

The scale factors applied to Monte Carlo for the EF\_tau29\_medium1\_tau20\_medium1 trigger rates are derived in this Chapter. There are two sets of scale factors constructed for two different situations. One for Monte Carlo that contains generated hadronic taus, and another for Monte Carlo with jets that are falsely identified as hadronic taus. The discrepancy between data and simulation differs for the two, and therefore each is calculated individually. The first scale factor, which we call the trigger efficiency, is applied to Monte Carlo in which there is a reconstructed hadronic tau that can be matched to a generator tau within  $\Delta R < 0.2$ . The second, called the trigger fake rate, is applied when the reconstructed tau is not matched to a generator tau. Both are applied to the taus as a function of  $p_T$ .

### 11.1 Trigger Factorization

Neither the trigger efficiency nor the fake rate of the EF\_tau29\_medium1\_tau20\_medium1 trigger, can be measured directly from data. Obviously, these rates are necessary when deriving scale factors to correct Monte Carlo to data. Since the EF\_tau29\_medium1\_tau20\_medium1 trigger is essentially a combination of two independent triggers, we solve the problem by estimating the values from the factorization of the corresponding single  $\tau_h$  triggers. These

are EF\_tau29\_medium1 and EF\_tau20\_medium1.

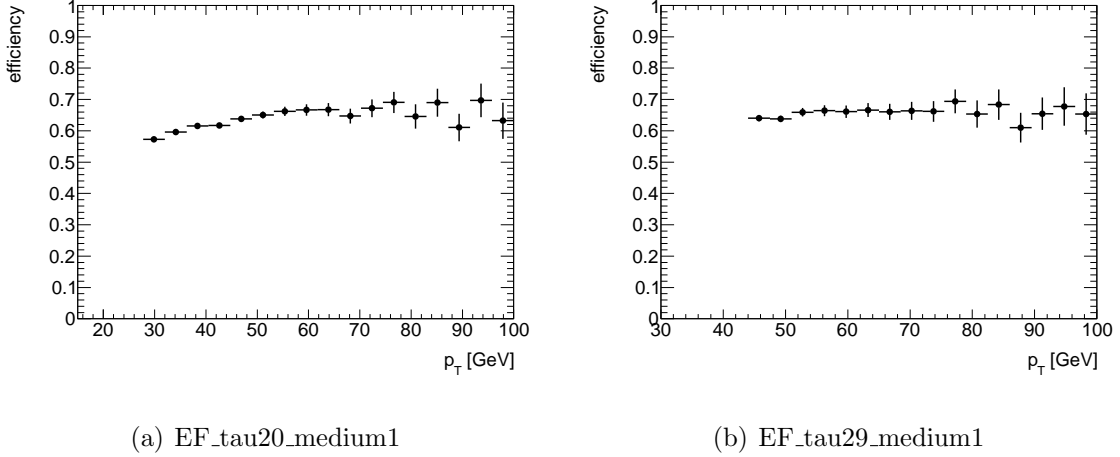
As the combined trigger contains two different  $p_T$  thresholds, the efficiency and fake rate must be determined as a function of both the leading and sub-leading tau candidate  $p_T$ . To prove that the combined trigger efficiency can be factorized by the single triggers, a Monte Carlo study using  $Z \rightarrow \tau\tau$  events was performed, and tested on  $A/H/h \rightarrow \tau\tau$  samples.

First the two efficiency distributions for EF\_tau29\_medium1 and EF\_tau20\_medium1 were determined as a function of the hadronic tau  $p_T$  in  $Z \rightarrow \tau\tau$  Monte Carlo events. For the high threshold trigger, EF\_tau29\_medium1, reconstructed  $\tau_h$  candidates were required to have  $p_T > 45$  GeV,  $|\eta| < 2.5$ , and pass **medium** tau identification. For the lower threshold trigger, EF\_tau20\_medium1, the  $p_T$  selection was dropped to  $> 30$  GeV. The efficiency distributions were constructed individually for each trigger by taking the number of taus that pass the criteria and are matched to the corresponding trigger RoI within  $\Delta R < 0.2$ , divided by the total number of selected taus. This ratio was calculated in bins of  $p_T$ , and both resulting distributions are shown in Figure 11.1.

The same tau selections were then made in  $A/H/h \rightarrow \tau\tau$  Monte Carlo. The factorized trigger efficiency was established by taking all selected taus and multiplying them by both single trigger efficiency values per  $p_T$  bin. Additionally, the combined trigger efficiency was determined directly, also in bins of  $p_T$ . This was calculated by taking all the selected taus satisfying the EF\_tau29\_medium1\_tau20\_medium1 trigger, divided by the number of selected taus in all events.

The factorized and direct efficiencies were compared bin by bin by to verify the validity of the approach. The results of the comparison are shown in Figure 11.2 for two different Higgs masses,  $m_A = 100, 200$ , and  $300$  GeV. The histograms display the difference in the calculated efficiency for each bin. As can be seen the discrepancy between the two approaches is compatible with zero.

To compute the EF\_tau29\_medium1\_tau20\_medium1 trigger scale factors from data,

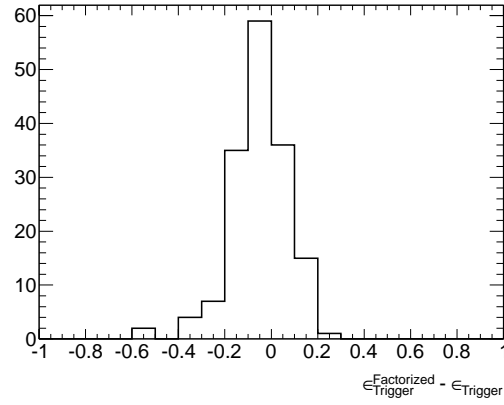
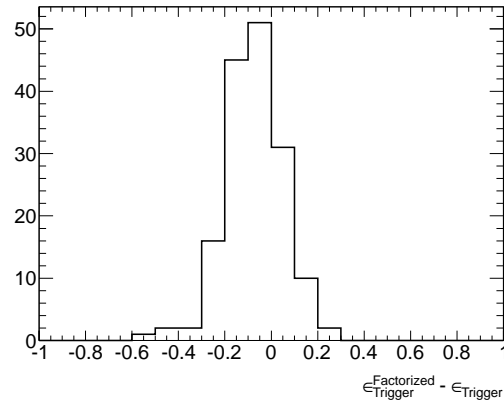


**Figure 11.1:** *Efficiency as a function of tau  $p_T$  for the EF\_tau29\_medium1 and EF\_tau20\_medium1 triggers.*

we calculate the trigger efficiency and fake rate for the two individual triggers, EF\_tau29\_medium1 and EF\_tau20\_medium1, and combine the results multiplicatively.

## 11.2 Single trigger efficiency measurement

In data the single tau trigger efficiency is measured with respect to an identified hadronic tau using a tag-and-probe analysis with  $Z \rightarrow \tau\tau \rightarrow \mu\tau_h$  events. The procedure follows directly from the method used in Section 9.3, to determine the tau identification efficiency. While previously the probe taus had very loose selection criteria, now we include a requirement of **medium** identification. This way the identification and trigger efficiencies are measured in two steps. By splitting the measurements, the calculated trigger efficiency has greater accuracy since the  $W$  and QCD backgrounds are reduced. The event, tag muon and probe tau selection criteria are listed in Table 11.1. Although the events are triggered using EF\_mu18, data is stored with the information of all other triggers that could cause the event to be accepted. We only consider events where there is the possibility of either the EF\_tau29\_medium1 or EF\_tau20\_medium1 trigger accept. Once the events are selected, and we have an unbiased sample of probe taus, the trigger efficiency is

(a)  $m_A = 200$  GeV(b)  $m_A = 300$  GeV

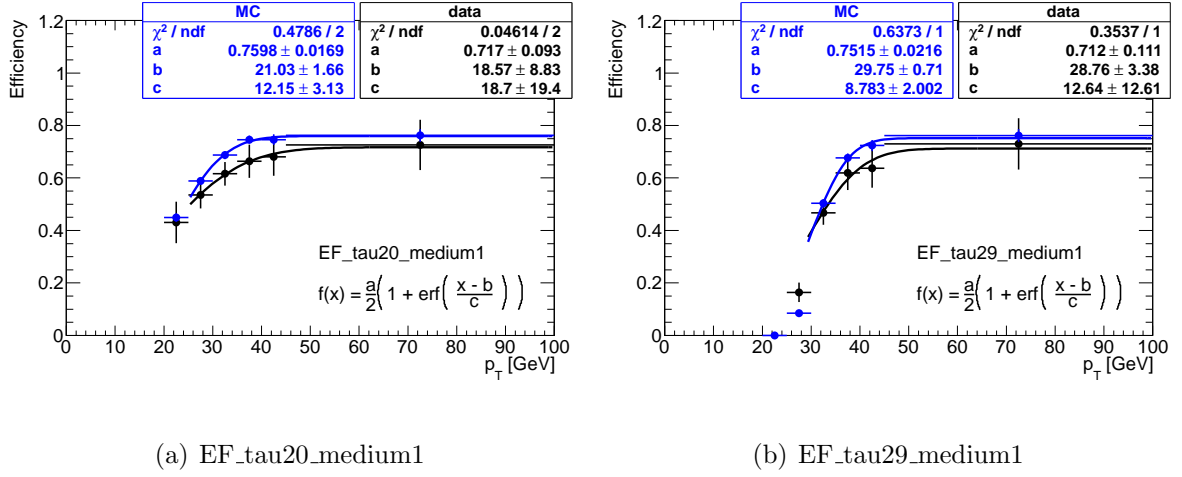
**Figure 11.2:** *Difference in the efficiency from factorized triggers, obtained from Z Monte Carlo samples, and the combined trigger when applied to two Higgs mass values. The difference is effectively zero, validating the factorization approach.*

determined. It is calculated separately for EF\_tau29\_medium1 and EF\_tau20\_medium1, in bins of hadronic tau  $p_T$ . The efficiency is the ratio of the number of hadronic taus that match the single tau trigger RoI within  $\Delta R < 0.2$  to the number of probe taus. The results from data are then compared to ALPGEN generated  $Z \rightarrow \tau\tau$  processes, where the reconstructed hadronic tau must match a generator tau within  $\Delta R < 0.2$ . Scale factors are constructed from the value in data divided by the value in Monte Carlo.

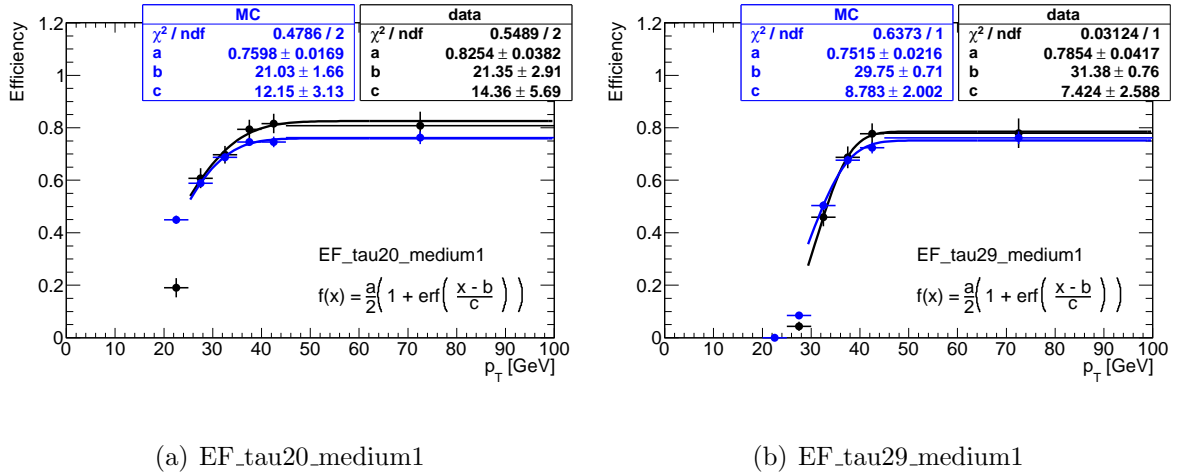
Event Selection	$\mu$ criteria	$\tau_h$ criteria
<ul style="list-style-type: none"> <li>• EF_mu18 trigger</li> <li>• Dilepton veto</li> <li>• <math>W</math> boson suppression</li> <li>• Opposite charged <math>\tau_h</math> and <math>\mu</math></li> <li>• <math>40 \text{ GeV} &lt; m_{\mu,\tau} &lt; 80 \text{ GeV}</math></li> <li>• Possible EF_tau29_medium1 or EF_tau20_medium1 accept</li> </ul>	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 20 \text{ GeV}</math></li> <li>• <math> \eta  &lt; 2.5</math></li> <li>• No additional tracks within <math>\Delta R &lt; 0.2</math> of muon</li> <li>• Isolation requirement of <math>\frac{E_T(\Delta R &lt; 0.2)}{p_T} &lt; 0.04</math></li> </ul>	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 30 \text{ GeV}</math></li> <li>• <math> \eta  &lt; 2.5</math></li> <li>• Exactly 1 or 3 core tracks</li> <li>• Charge = <math>\pm 1</math></li> <li>• <b>medium</b> identification</li> </ul>

**Table 11.1:** Selection of  $Z \rightarrow \tau\tau \rightarrow \tau_l\tau_h$  events for single tau trigger efficiency in data.

Efficiencies for both the EF\_tau20\_medium1 and EF\_tau29\_medium1 triggers are measured for run periods **B**<sub>2</sub> to **E** and **F** to **H** separately, because of a change made to the selection in the L2 tau trigger decisions from period **F** onwards. This change is not implemented in the Monte Carlo, so scale factors necessarily differ between periods. The efficiencies for each EF trigger are shown in Figures 11.3 and 11.4 as a function of the tau  $p_T$ . The data to Monte Carlo scale factors are given in Table 11.2 in bins of  $p_T$ .



**Figure 11.3:** Trigger efficiencies for hadronic taus as a function of  $p_T$ . Data periods B<sub>2</sub> to F.



**Figure 11.4:** Trigger efficiencies for hadronic taus as a function of  $p_T$ . Data periods F to H.

$p_T$ [GeV]	EF_tau20_medium1		EF_tau29_medium1	
	<b>B<sub>2</sub> to E</b>	<b>F to H</b>	<b>B<sub>2</sub> to E</b>	<b>F to H</b>
20-25	$0.96 \pm 0.17$	$0.55 \pm 0.04$	-	-
25-30	$0.91 \pm 0.09$	$1.03 \pm 0.07$	$1.30 \pm 0.16$	$0.68 \pm 0.15$
30-35	$0.90 \pm 0.07$	$1.02 \pm 0.05$	$0.93 \pm 0.10$	$0.91 \pm 0.07$
35-40	$0.89 \pm 0.09$	$1.06 \pm 0.05$	$0.92 \pm 0.10$	$1.02 \pm 0.07$
40-45	$0.91 \pm 0.10$	$1.09 \pm 0.06$	$0.88 \pm 0.11$	$1.07 \pm 0.06$
45-100	$0.95 \pm 0.13$	$1.06 \pm 0.08$	$0.96 \pm 0.13$	$1.02 \pm 0.08$

**Table 11.2:** *Trigger efficiency scale factors for **medium** identified hadronic taus in bins of  $p_T$ . The scale factors are constructed by taking the number of events in data divided by the events in Monte Carlo: Data/MC. They are applied to MC in the same bins to correct the number of events to data.*

### 11.3 Single trigger fake rate measurement

Falsely triggered events from QCD jets were studied using jets accompanying  $W \rightarrow \mu\nu$  production. The muon decay channel ensures a high purity sample, as muons are efficiently reconstructed and identified particles. The single tau triggers, EF\_tau29\_medium and EF\_tau20\_medium were studied separately, and the results combined as for the trigger efficiency. Event selection criteria are based on  $W$  studies from previous ATLAS analysis [54]. Data with an integrated luminosity of  $730 \text{ pb}^{-1}$ , with events selected by the un-prescaled EF\_mu18 trigger were used. The full event, muon and tau selection is given in Table. 11.3.

Event Selection	$\mu$ criteria	$\tau_h$ criteria
<ul style="list-style-type: none"> <li>• EF_mu18 trigger</li> <li>• Electron veto</li> <li>• <math>E_T^{Miss} &gt; 30 \text{ GeV}</math></li> <li>• <math>m_T(\mu, E_T^{Miss}) &lt; 60 \text{ GeV}</math></li> <li>• Possible EF_tau29_medium1 or EF_tau20_medium1 accept</li> </ul>	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 20 \text{ GeV}</math></li> <li>• <math> \eta  &lt; 2.5</math></li> <li>• No additional tracks within <math>\Delta R &lt; 0.2</math> of muon</li> <li>• Isolation requirement of <math>\frac{E_T(\Delta R &lt; 0.2)}{p_T} &lt; 0.04</math></li> </ul>	<ul style="list-style-type: none"> <li>• <math>p_T &gt; 15 \text{ GeV}</math></li> <li>• <math> \eta  &lt; 2.5</math></li> <li>• Exactly 1 or 3 core tracks</li> <li>• Charge = <math>\pm 1</math></li> <li>• <b>medium</b> identification</li> </ul>

**Table 11.3:** Selection for  $W \rightarrow \mu\nu + \text{jet}$  events for singel tau trigger fake rate in data.

Event selection criteria and efficient muon reconstruction gives inclusive  $W \rightarrow \mu\nu$  events with  $> 90\%$  purity and QCD contamination less than  $1\%$ . The fake rate is evaluated using the recoiling jet in an event. Results are obtained as a function of the falsely identified **medium** tau candidate  $p_T$ . The fake rate is determined by the ratio of the number of identified taus that match one of the single tau trigger RoIs within  $\Delta R < 0.2$  to the number of identified taus. Results from data are compared to inclusive ALPGEN generated  $W \rightarrow \mu\nu$  processes to construct scale factors. In the Monte Carlo, the identified tau must not be matched to any generator tau within  $\Delta R < 0.2$ .

Results are divided between two data taking periods, **B**<sub>2</sub> to **E**, and **F** onwards. The scale factors in bins of the fake tau  $p_T$  are given in Table 11.4.

$p_T$ [GeV]	EF_tau20_medium1		EF_tau29_medium1	
	<b>B</b> <sub>2</sub> to <b>E</b>	<b>F</b> to <b>H</b>	<b>B</b> <sub>2</sub> to <b>E</b>	<b>F</b> to <b>H</b>
20-30	$0.52 \pm 0.12$	$0.51 \pm 0.08$	-	-
30-45	$0.65 \pm 0.14$	$0.79 \pm 0.08$	$0.71 \pm 0.14$	$0.66 \pm 0.09$
45-100	$0.57 \pm 0.23$	$0.57 \pm 0.14$	$0.58 \pm 0.23$	$0.55 \pm 0.15$

**Table 11.4:** *Trigger fake rate scale factors for **medium** mis-identified hadronic taus in bins of  $p_T$ . The scale factors are constructed by taking the number of events in data divided by the events in Monte Carlo: Data/MC. They are applied to MC in the same bins to correct the number of events to data.*

Monte Carlo samples are corrected to model the trigger efficiency for data by applying the derived scale factors to reconstructed taus that are matched to a generator tau within  $\Delta R < 0.2$ . Alternatively, the trigger fake rate is applied to reconstructed taus that are not matched to a generator tau in this range. In each case the EF\_tau29\_medium1 scale factor is applied to the highest- $p_T$  tau, and the EF\_tau20\_medium1 scale factor to the sub-leading tau.

# Chapter 12

## Electroweak Background

All electroweak backgrounds are estimated from Monte Carlo in this analysis. Results for the expected event yield were given previously in Table 10.1. To ensure that the largest simulated EW backgrounds,  $Z \rightarrow \tau\tau$  and  $W \rightarrow \tau\nu$ , model the kinematic and mass shapes in data, an embedded data technique is used [60]. An embedded data sample uses either selected  $Z \rightarrow \mu\mu$  or  $W \rightarrow \mu\nu$  events, where the muons are replaced at the detector level, with full simulation  $\tau_h$  decays as follows:

1. All calorimeter cell energy in a cone of  $\Delta R < 0.1$  around the muon direction, defined by its track, is removed from the event.
2.  $Z \rightarrow \tau_h\tau_h$  ( $W \rightarrow \tau_h\nu$ ) is simulated in TAUOLA [77] using the same kinematics as the removed  $Z \rightarrow \mu\mu$  ( $W \rightarrow \mu\nu$ ).
3. Full detector simulation, without any calorimeter noise, underlying events, or pileup, is run for the extracted  $Z \rightarrow \tau_h\tau_h$  ( $W \rightarrow \tau_h\nu$ ) event.
4. The tracks and calorimeter cells from the simulated  $Z \rightarrow \tau_h\tau_h$  ( $W \rightarrow \tau_h\nu$ ) event are inserted into the initial data event, and a reconstruction of the hybrid event is performed.

Trigger information is not included in the embedding method, therefore such acceptances and efficiencies cannot be determined for the embedded samples themselves. Uncertainties are not derived from the embedded approach, since the samples are constructed with different event and object criteria than used in the  $A/H/h \rightarrow \tau_h \tau_h$  analysis. However, by a comparison of the embedded samples to Monte Carlo for the ALPGEN generated  $Z \rightarrow \tau\tau + \text{jets}$  and  $W \rightarrow \tau\nu + \text{jets}$  samples, we verify that the Monte Carlo models the data well.

## 12.1 $Z \rightarrow \tau_h \tau_h$ background estimation

Data samples are selected for  $Z \rightarrow \mu\mu$  by requiring two muons, both with  $p_T > 20$  GeV, originating from the same primary vertex. Combined, the muons must have an invariant mass of  $m_{\mu\mu} > 40$  GeV. The embedding technique is then applied to this data. ALPGEN generated  $Z \rightarrow \tau_h \tau_h$  Monte Carlo is compared to the embedded data with the following event selection:

- Event cleaning (as described in Chapter 10)
- Electron and muon lepton veto (as described in Chapter 10)
- Exactly two **medium** identified taus with opposite charge, with  $p_T > 25$  GeV and  $p_T > 20$  GeV

The  $E_T^{Miss}$  distribution for events passing selection, and the  $p_T$  distribution of the leading  $\tau_h$  candidates are shown in Fig. 12.1.

Figure 12.2 shows the visible mass distribution coming from the 4-momentum addition of the two  $\tau_h$  candidates in the event, ignoring any effect of the neutrinos. The plot shows both the ALPGEN generated Monte Carlo sample and the embedded data.

The agreement observed in Figures 12.1 and 12.2 between Monte Carlo and embedded data gives confidence that the ALPGEN generated samples can be used to estimate the

expected irreducible  $Z \rightarrow \tau_h \tau_h$  background in this analysis, including underlying events and pileup conditions.

## 12.2 $W \rightarrow \tau \nu$ background estimation

In this analysis,  $W \rightarrow \tau \nu$  + jet events are a source of background when a real  $\tau_h$ , and a mis-identified hadronic tau coming from a QCD jet, pass the event selection requirements. By studying this process with the embedded data technique, we compare the event kinematics in the simulation to data, including the QCD jets mis-identified as hadronic taus.

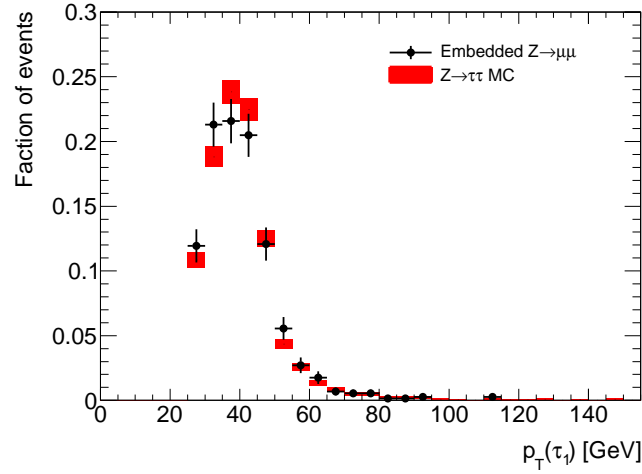
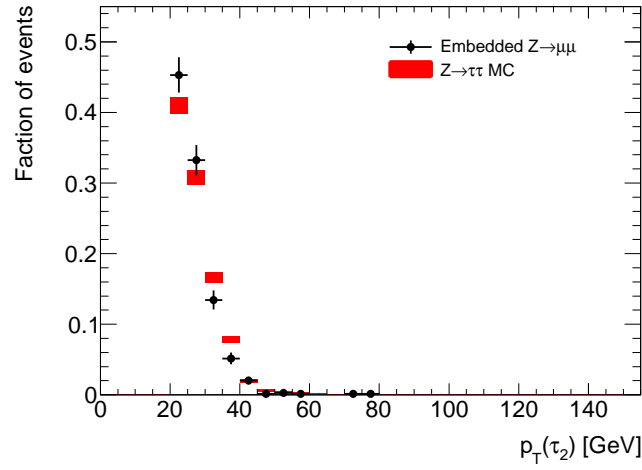
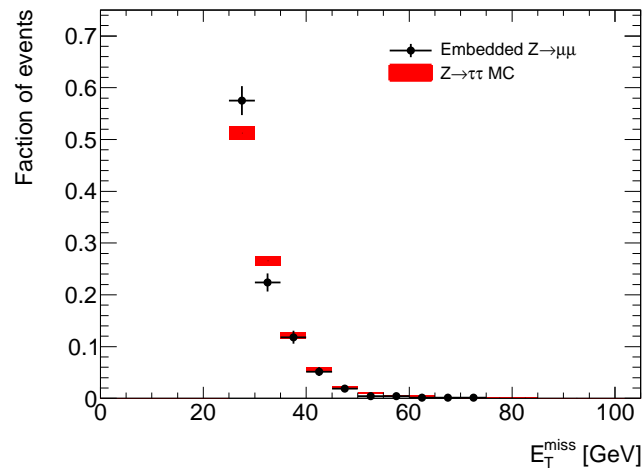
The  $W \rightarrow \mu \nu$  events are selected in data by requiring a muon with  $p_T > 20$  GeV, and  $E_T^{Miss} > 30$  GeV. The transverse mass generated by the muon and  $E_T^{Miss}$  is restricted to  $m_T > 40$  GeV. The transverse mass,  $m_T$ , is defined as  $\sqrt{2p_T^\mu E_T^{Miss} \cos(\Delta\phi)}$ , where  $\Delta\phi$  is the angular separation in the transverse plane between the  $\tau_h$  and  $E_T^{Miss}$ . ALPGEN generated Monte Carlo and the embedded  $W \rightarrow \tau_h \nu$  distributions were compared after the following event selection:

- Event cleaning (as described in Chapter 10)
- Electron and muon lepton veto (as described in Chapter 10)
- Exactly two **medium** identified taus with opposite charge, with  $p_T > 25$  GeV and  $p_T > 20$  GeV
- $E_T^{Miss} > 30$  GeV
- $m_T > 40$  GeV

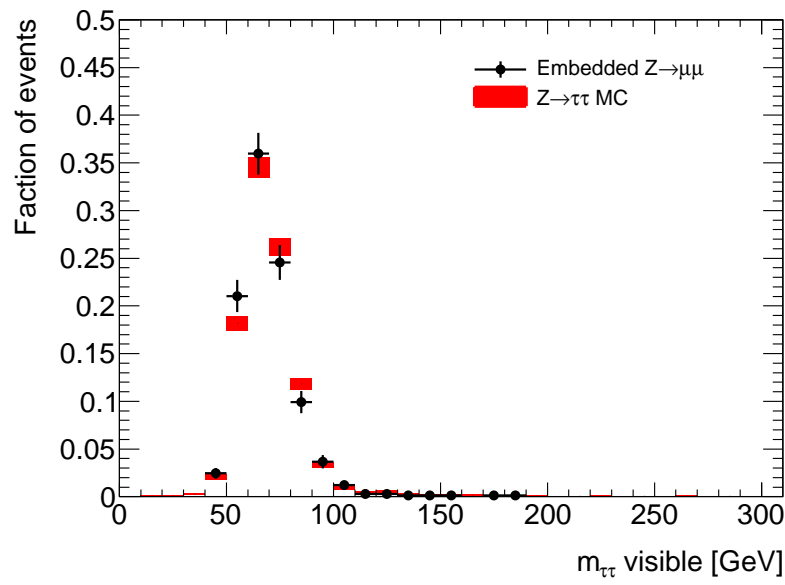
The transverse mass cut is necessary when comparing simulation to the embedded data, because of the original  $W \rightarrow \mu \nu$  data selection. Although there are no requirements on  $m_T$  in the Monte Carlo  $W$  background estimation for the  $A/H/h \rightarrow \tau_h \tau_h$  analysis,

it is needed here to select the same kinematic phase space in ALPGEN Monte Carlo as in the embedded events. Figure 12.3 compares the kinematic distributions of the Monte Carlo and the embedded data, the  $p_T$  distributions of the two leading  $\tau_h$  candidates and the event  $E_T^{Miss}$  are shown. Figure 12.4 shows the visible mass distribution of the two  $\tau_h$  candidates.

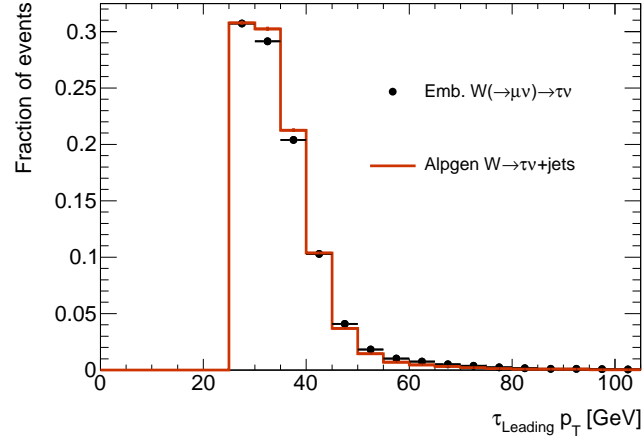
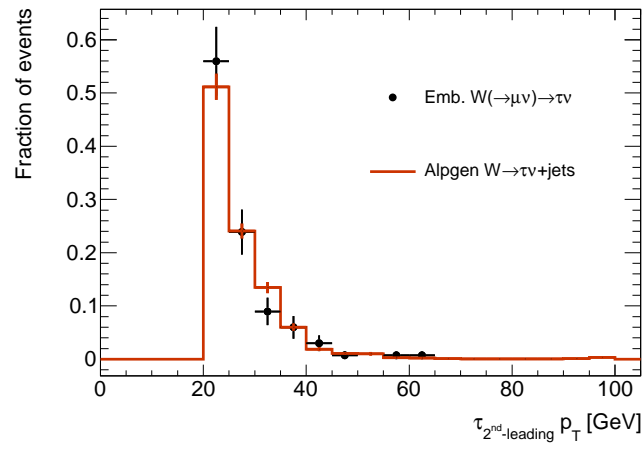
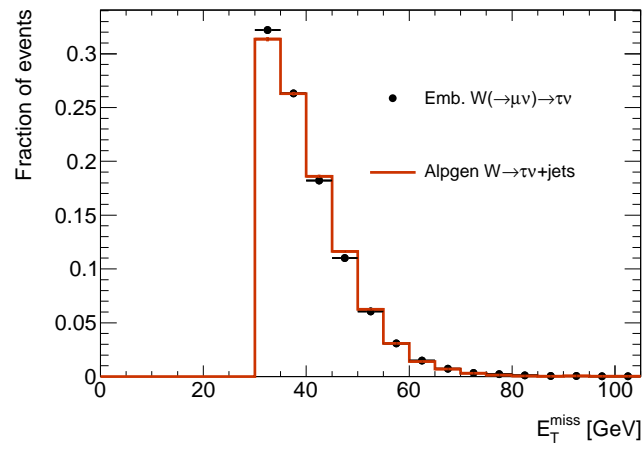
Figure 12.5 shows the jet multiplicity spectrum for  $W \rightarrow \tau_h \nu$  events selected requiring  $\tau_h$  candidates with log-likelihood-ratio derived **medium** identification.

(a)  $p_T$  of leading  $\tau_h$  candidate(b)  $p_T$  of sub-leading  $\tau_h$  candidate(c)  $E_T^{\text{Miss}}$  from selected events

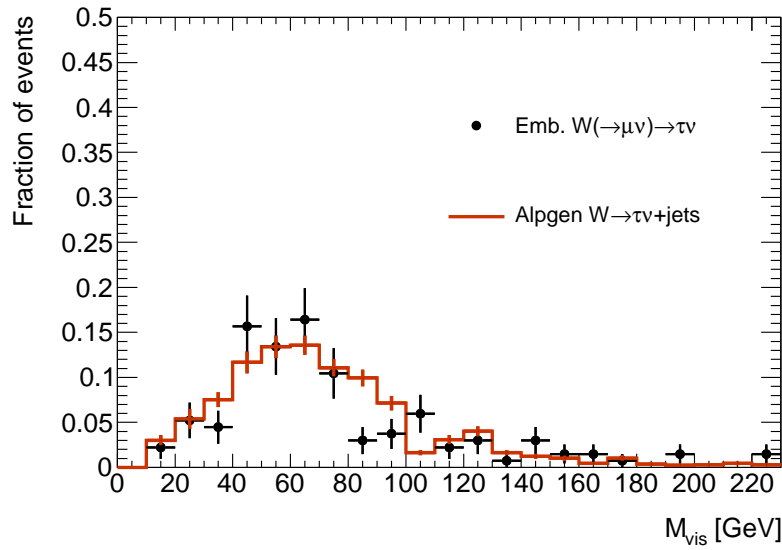
**Figure 12.1:** Comparison of the leading and sub-leading hadronic tau  $p_T$ , and  $E_T^{\text{Miss}}$  spectra, between ALPGEN generated  $Z \rightarrow \tau_h \tau_h$  Monte Carlo and embedded data sample.



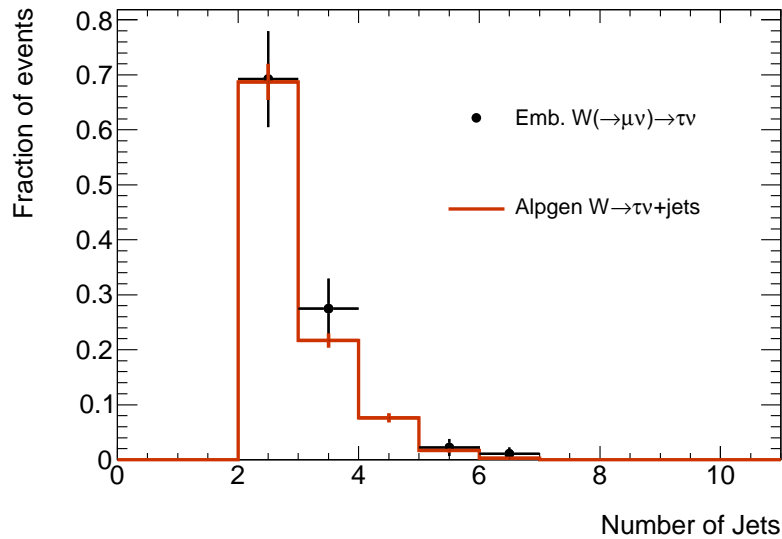
**Figure 12.2:** Comparison of the visible mass distribution from the ALPGEN generated  $Z \rightarrow \tau_h \tau_h$  Monte Carlo and the embedded data.

(a)  $p_T$  of leading  $\tau_h$  candidate(b)  $p_T$  of sub-leading  $\tau_h$  candidate(c)  $E_T^{Miss}$  from selected events

**Figure 12.3:** Comparison of the leading and sub-leading hadronic tau  $p_T$ , and  $E_T^{Miss}$  spectra, between ALPGEN generated  $W \rightarrow \tau_h \nu$  Monte Carlo and embedded data sample.



**Figure 12.4:** Comparison of the visible mass distribution from the ALPGEN generated  $W \rightarrow \tau_h \nu$  Monte Carlo and the embedded data. The visible mass is formed from one real hadronic tau, and one mis-identified tau from a QCD jet.



**Figure 12.5:** Comparison of the hadronic tau multiplicity from the  $W \rightarrow \tau_h \nu$  ALPGEN generated Monte Carlo and the embedded data sample for **medium** identified taus with  $p_T > 20$  GeV.

Again, good agreement is seen between data and the ALPGEN Monte Carlo distributions. Although the uncertainties from the figures in this section are not used in the final analysis, the values are within the final systematic uncertainty derived directly from altering ALPGEN generator conditions to construct Monte Carlo, as described in Section 15.1.2.

# Chapter 13

## QCD Background Estimation

It is possible to understand the QCD background directly from data, unlike other background processes. Consequently, this analysis does not rely on simulated event samples to predict the rate of QCD processes, but uses a data-driven approach.

### 13.1 ABCD Method

The data-driven approach is based on the ABCD method (as described in Section 9.3.2), by the selection of four independent data samples, three in QCD background-dominated regions (control regions) and one in the signal region. Regions are defined by same sign and opposite sign (OS and SS) taus, and the  $\tau_h$  log-likelihood-ratio identification level (**loose** and **medium**) of the two highest- $p_T$   $\tau_h$  candidates. The charge and identification variables are assumed to be uncorrelated. The following four regions are used:

**Region A:** events where the two highest- $p_T$   $\tau_h$  candidates satisfy the **medium**  $\tau_h$  identification requirements and have opposite charge.

**Region B:** events where the two highest- $p_T$   $\tau_h$  candidates satisfy the **medium**  $\tau_h$  identification requirements and have same charge.

**Region C:** events where the two highest- $p_T$   $\tau_h$  candidates satisfy the **loose**  $\tau_h$ , fail the **medium**  $\tau_h$  identification requirements and have opposite charge.

**Region D:** events where the two highest- $p_T$   $\tau_h$  candidates satisfy the **loose**  $\tau_h$ , fail the **medium**  $\tau_h$  identification requirements and have same charge.

In all regions the remaining analysis event selection, as described in Chapter 10, was also applied, with the exception of modifications in steps 5, 6 and 8 of Section 10.2 to comply with the definitions above. Region A contains the signal, and B, C and D, referred to as the ‘control regions’ primarily contain QCD jets. Table 13.1 illustrates the division of events by the ABCD method.

	OS	SS
<b>medium</b> $\tau_h$ -ID	A (Signal)	B
<b>loose</b> $\tau_h$ -ID	C	D

**Table 13.1:** *Definition of signal and control regions based on the charge correlation of the two highest- $p_T$  the  $\tau_h$  candidates (OS and SS) and the  $\tau_h$  identification, **medium**  $\tau_h$  identification and (**loose but not medium**)  $\tau_h$  identification.*

Table 13.2 shows the numbers of measured data events, the number of non-QCD background events, and the number of signal events predicted from simulation ( $m_A = 200$  GeV,  $\tan\beta = 20$ ). The term non-QCD background refers to the combined background from EW and top processes.

The shape of distributions from QCD background can be estimated from regions B and C. As there are fewer events in region B, and a higher associated statistical uncertainty, the shape is taken from region C. Variables related to the  $\tau_h$  identification (e.g.  $R_{EM}$ , log-likelihood-ratio output score) are correlated to the  $\tau_h$  identification level and the shape from region C is biased. For these variables the shape from region B is used even though this leads to a higher statistical uncertainty. Tau identification variables are only used to

	A	B	C	D
Data	245	131	624	480
$b\bar{b}A \rightarrow \tau_h \tau_h$	$14.4 \pm 1.0$	$0.1 \pm 0.1$	$1.6 \pm 0.3$	0
$gg \rightarrow A \rightarrow \tau_h \tau_h$	$3.3 \pm 0.3$	0	$0.3 \pm 0.1$	0
$Z/\gamma^* + \text{jets}$	$47.6 \pm 4.6$	$2.3 \pm 0.8$	$7.1 \pm 1.7$	$0.9 \pm 0.6$
$W + \text{jets}$	$24.9 \pm 4.8$	$3.6 \pm 1.5$	$15.1 \pm 4.6$	$1.3 \pm 0.6$
Top	$2.0 \pm 0.9$	$0.1 \pm 0.1$	$0.5 \pm 0.4$	0
Di-boson	$1.4 \pm 0.3$	$0.1 \pm 0.0$	$0.2 \pm 0.1$	0

**Table 13.2:** Observed number of events in data and expected number of events from simulation in the signal and the three control regions used for the QCD background estimation. The full event selection with exception of the criteria defining the control regions (charge product,  $\tau_h$  identification) is applied. For the simulated signal  $m_A = 200$  GeV and a value of  $\tan \beta = 20$  is used. Statistical uncertainties are quoted.

validate the background estimation approach, and not explicitly necessary for the MSSM Higgs search. The normalization is adjusted by the ratio of events in regions B and D. The estimate for the expected number of QCD background events in signal region A,  $N_A^{\text{QCD estimate}}$ , is then obtained by:

$$N_A^{\text{QCD estimate}} = \frac{N_B^{\text{data-non QCD}}}{N_D^{\text{data-non QCD}}} \times N_C^{\text{data-non QCD}}. \quad (13.1)$$

Where  $N_i^{\text{data-non QCD}} = N_i^{\text{data}} - N_i^{\text{non QCD}}$ ,  $N_i^{\text{data}}$  represents the number of observed events in region  $i$ , and  $N_i^{\text{non QCD}}$  gives the estimated number of events from non-QCD backgrounds taken from Monte Carlo, i.e.  $Z/\gamma^* + \text{jets}$ ,  $W + \text{jets}$ , top and di-boson backgrounds, in region  $i$  ( $i = B, C, \text{ or } D$ ).

The estimate of the total number of events from QCD background in the signal region obtained from 13.1 and the values from Table 13.2 is:

$$n_A^{\text{QCD estimate}} = 157 \pm 18. \quad (13.2)$$

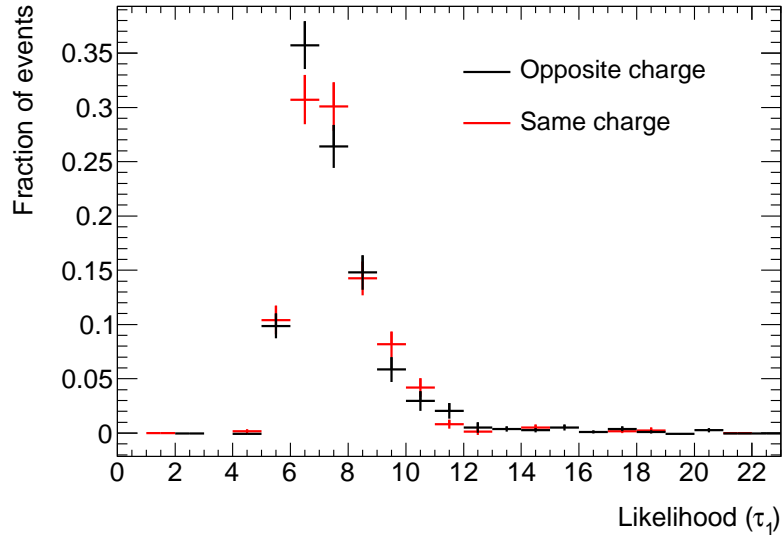
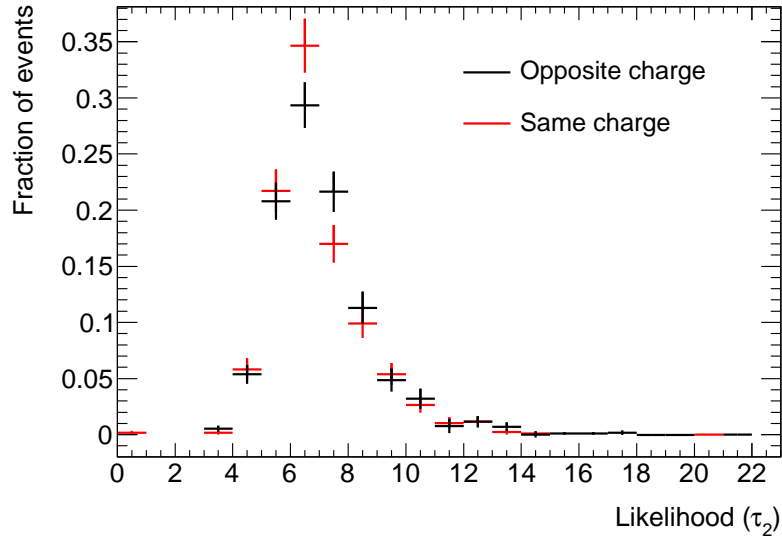
## 13.2 Verification of the underlying assumptions

The QCD background estimation in this method is based on three assumptions, namely:

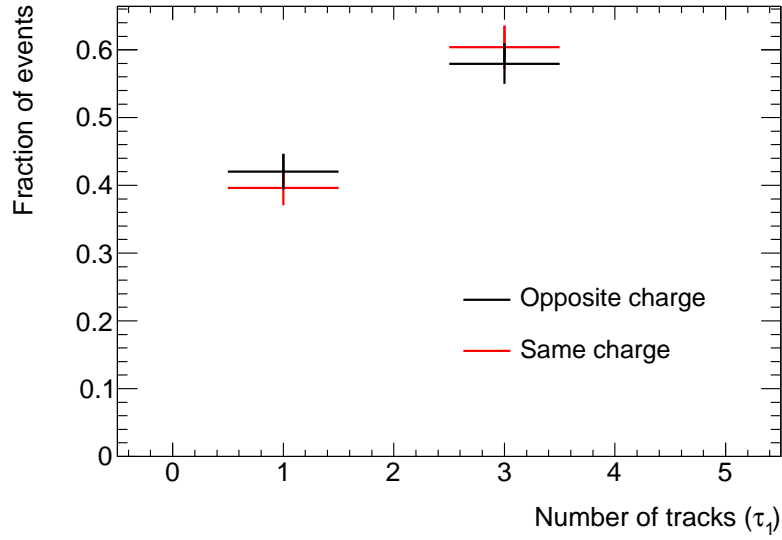
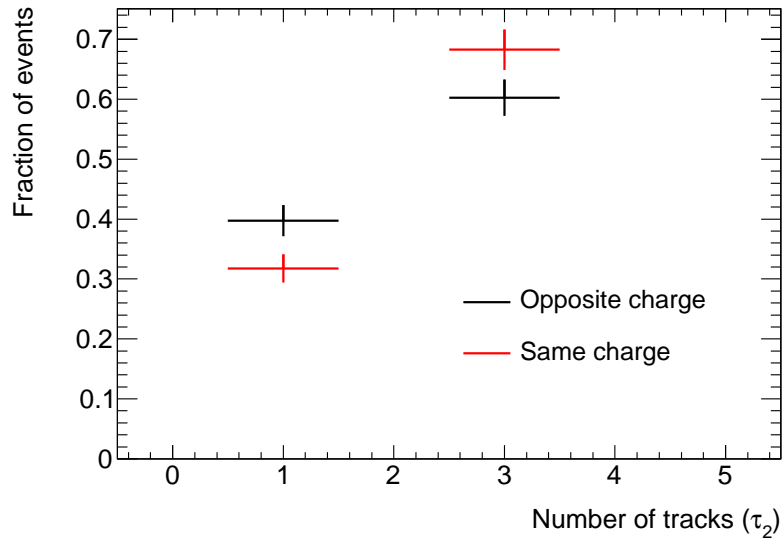
- The contribution of signal, electroweak and top backgrounds in the three control regions is small.
- The charge product between the two  $\tau_h$  candidates, and the  $\tau_h$  identification are uncorrelated.
- The shape of the  $m_{vis}$  distribution for QCD background (constructed from the two tau candidates) is the same in both regions **A** and **C**.

Table 13.2 shows that only a small contribution from non-QCD background events is present, at the level of 10%, in region B, and below 5% in all other control regions. As detailed in Section 13.1, these backgrounds are subtracted from the observed number of events in data.

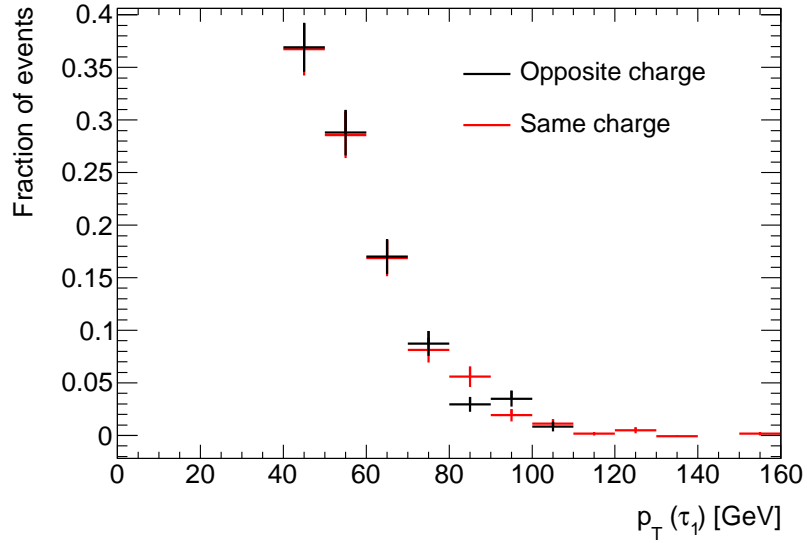
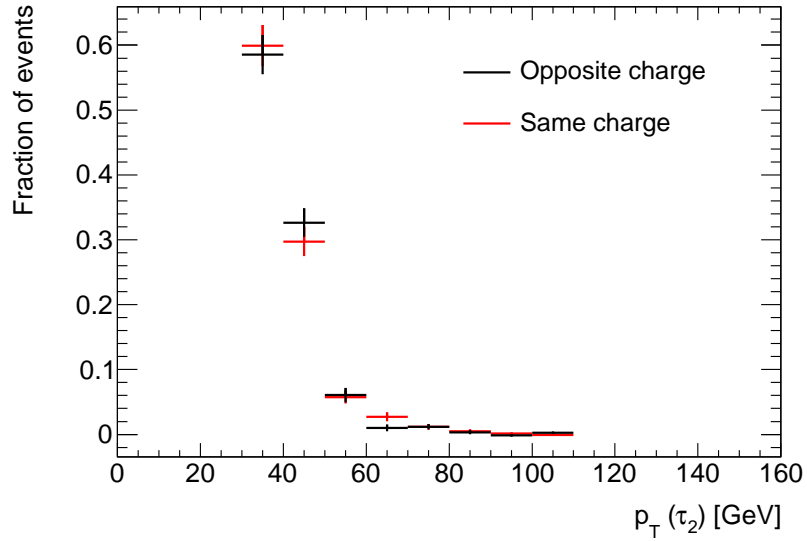
The  $\tau_h$  identification is dependent on the log-likelihood-ratio output score, calculated based on several input variables, discussed in Section 9.1. The assumption that the variables used to define the control regions for the QCD background estimation are uncorrelated is checked in Figures 13.1, 13.2 and 13.3 by comparing the distributions for events where the same sign highest- $p_T$   $\tau_h$  candidates to the opposite sign events.

(a) Highest- $p_T$   $\tau_h$  candidate(b) Second-highest- $p_T$   $\tau_h$  candidate

**Figure 13.1:** Comparison of the log-likelihood-ratio output score for  $\tau_h$  identification for events where the two highest- $p_T$  hadronic tau candidates have same and opposite charge. Shown is the observed number of events in data, non-QCD backgrounds are subtracted.

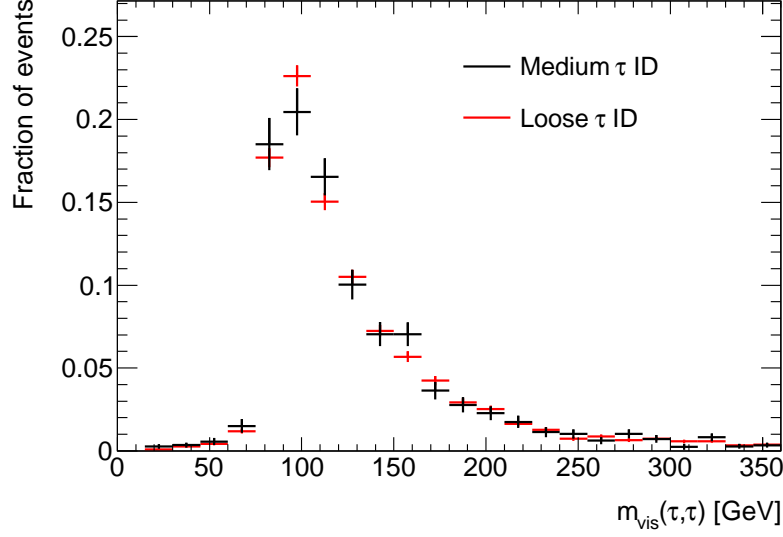
(a) Highest- $p_T$   $\tau_h$  candidate(b) Second-highest- $p_T$   $\tau_h$  candidate

**Figure 13.2:** Comparison of the number of tracks associated to the highest- $p_T$  and second-highest- $p_T$  hadronic tau candidates for events where the two highest- $p_T$  taus candidates have same and opposite charge. Shown is the observed number of events in data, non-QCD backgrounds are subtracted.

(a) Highest- $p_T$   $\tau_h$  candidate(b) Second-highest- $p_T$   $\tau_h$  candidate

**Figure 13.3:** Comparison of the transverse momentum of the two highest- $p_T$   $\tau_h$  candidates for events where the two highest- $p_T$   $\tau_h$  candidates have same and opposite charge. Shown is the observed number of events in data with electroweak and top background subtracted.

The shape of the  $m_{vis}$  distribution agrees in regions AB and CD, where AB (CD) refers to a combination of regions A and B (C and D). Figure 13.4 shows that the two distributions agree within the statistical uncertainties.



**Figure 13.4:** Comparison of the shape of the  $m_{vis}$  distribution in regions with **medium** (regions A and B) and **loose** but not **medium**  $\tau_h$  identification (regions C and D). Shown is the observed number of events in data with electroweak and top background subtracted.

# Chapter 14

## Validation of the Combined Background Estimation

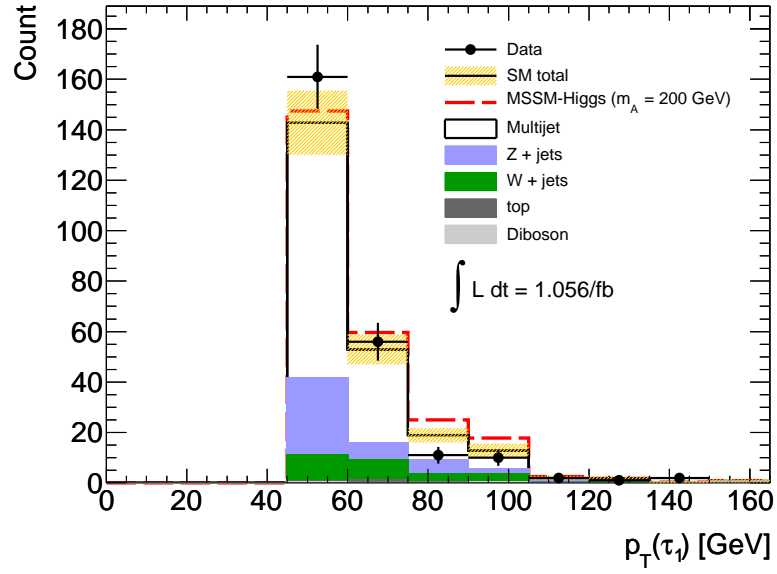
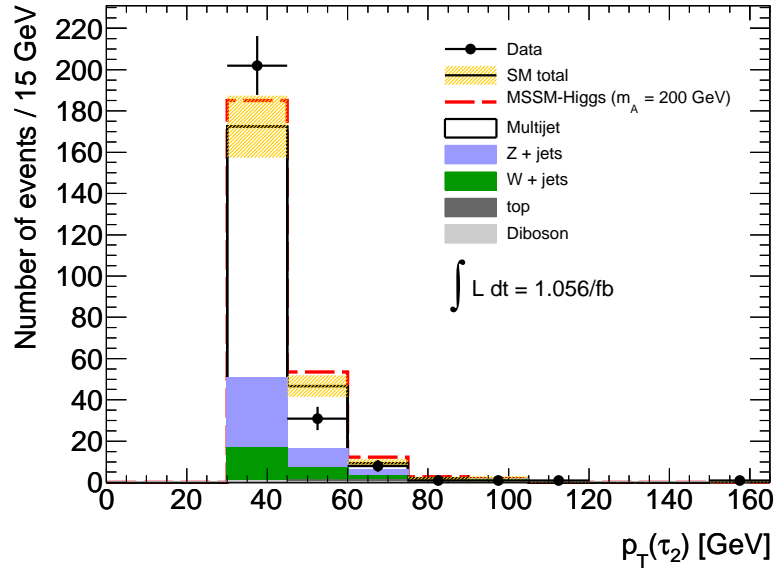
Distributions of the kinematic variables from events and inputs for  $\tau_h$  identification inputs are shown for both Monte Carlo simulation and real data in this chapter. At this stage, all the events used in the analysis have been selected. If the Monte Carlo models the data well, and the QCD background is correctly predicted the distributions should show agreement.

The following variables are considered:

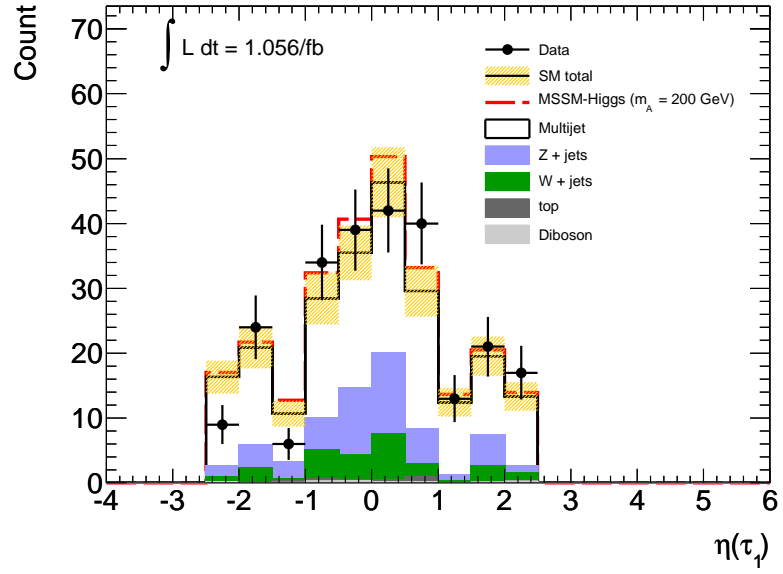
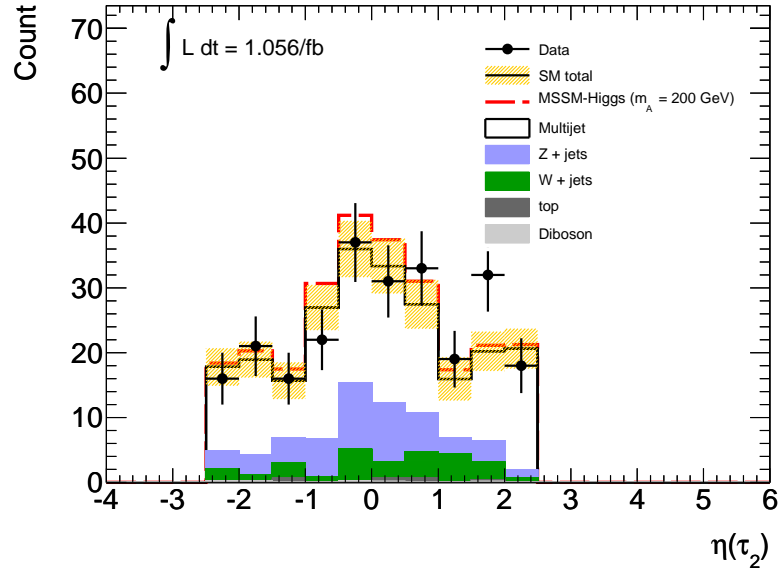
- transverse momentum  $p_T$  of the two highest- $p_T$   $\tau_h$  candidates (Figure 14.1)
- pseudorapidity  $\eta$  of the two highest- $p_T$   $\tau_h$  candidates (Figure 14.2)
- azimuthal angle  $\phi$  of the two highest- $p_T$   $\tau_h$  candidates (Figure 14.3)
- charge of the two highest- $p_T$   $\tau_h$  candidates (Figure 14.4)
- number of tracks of the two highest- $p_T$   $\tau_h$  candidates (Figure 14.5)
- electromagnetic radius of the two highest- $p_T$   $\tau_h$  candidates (Figure 14.6)
- missing transverse energy  $E_T^{Miss}$  (Figure 14.7)

- difference in azimuthal angle of the highest and second highest- $p_T$   $\tau_h$  candidate  $\Delta\phi$  (Figure 14.8)

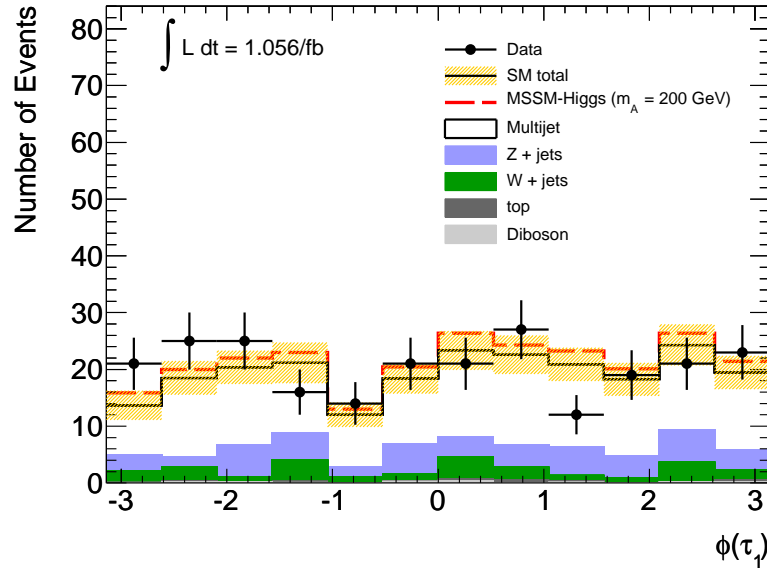
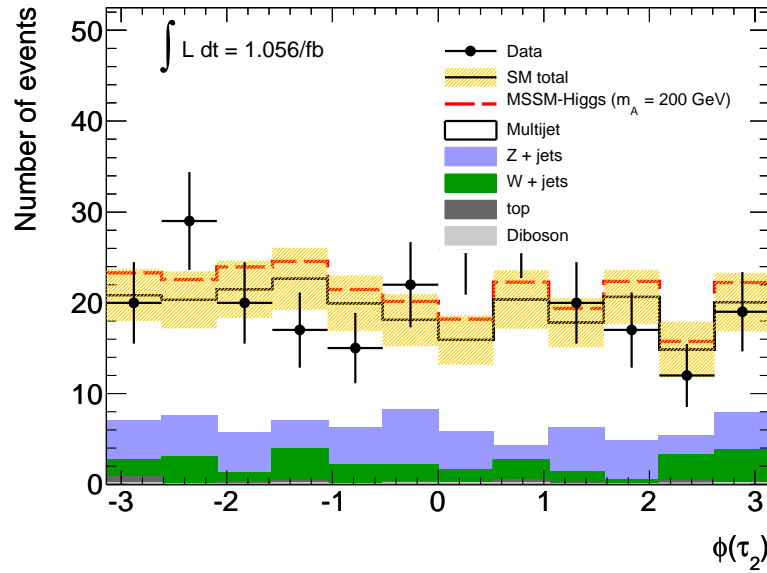
In all distributions shown, the potential signal from a Higgs, with  $m_A = 200$  GeV, and  $\tan\beta = 20$  is also shown with a red dashed line.

(a) highest  $p_T$   $\tau_h$  candidate(b) second highest  $p_T$   $\tau_h$  candidate

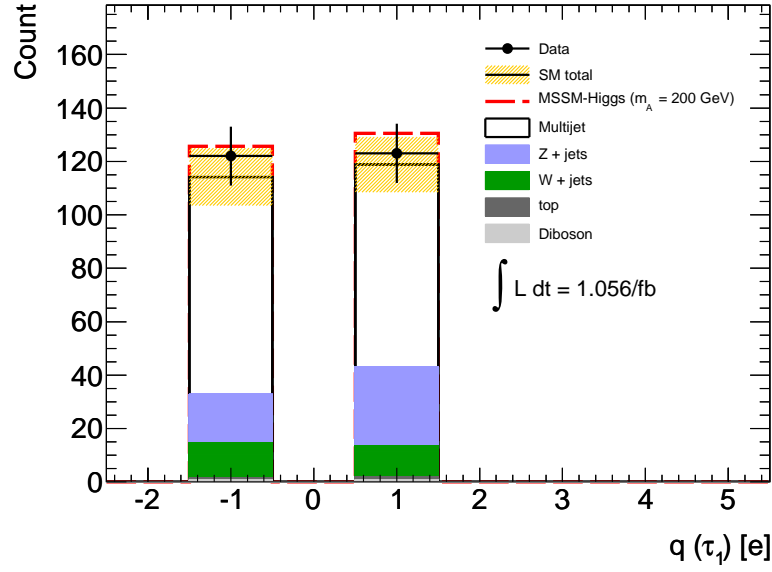
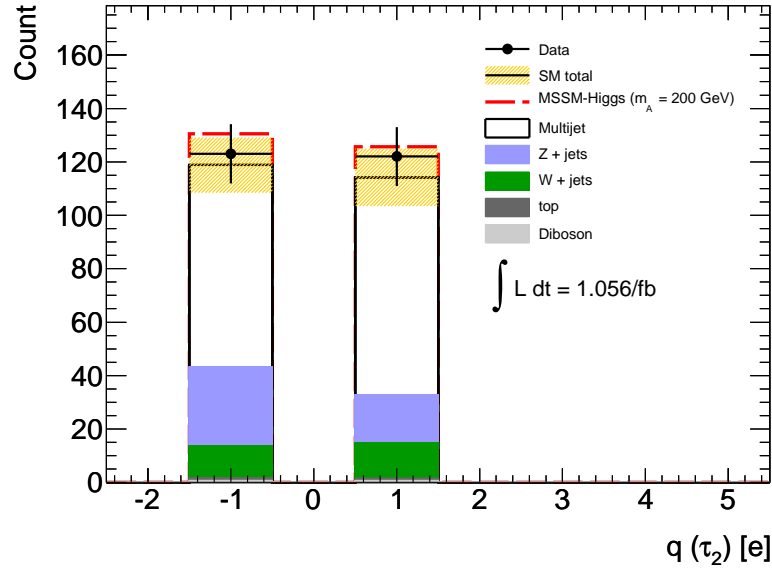
**Figure 14.1:** *Transverse momentum of the highest and second highest- $p_T$  hadronic tau candidate.*

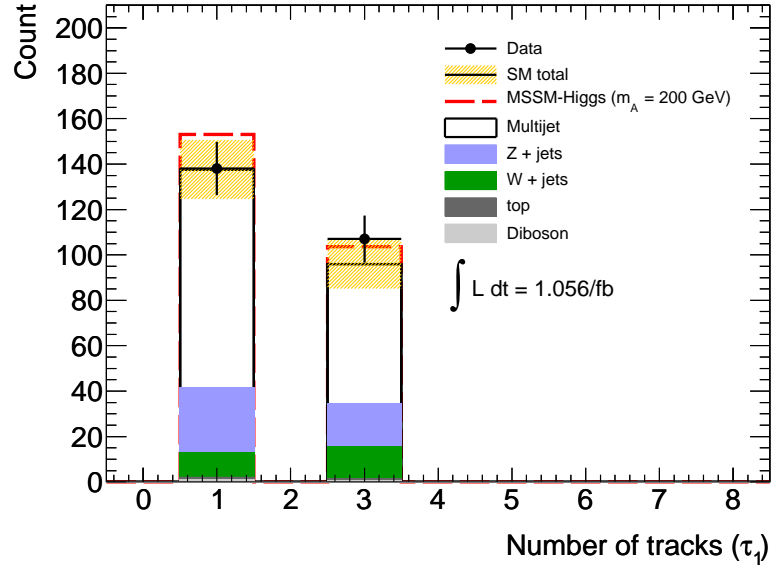
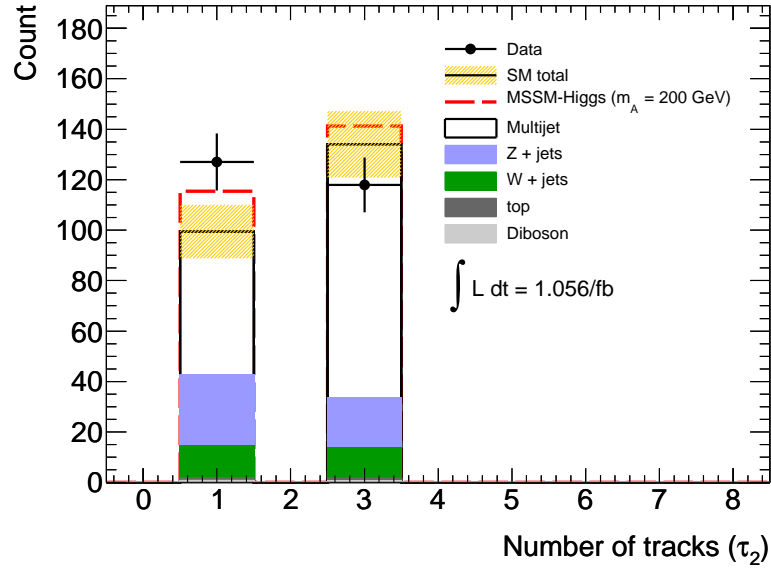
(a) highest  $p_T$   $\tau_h$  candidate(b) second highest  $p_T$   $\tau_h$  candidate

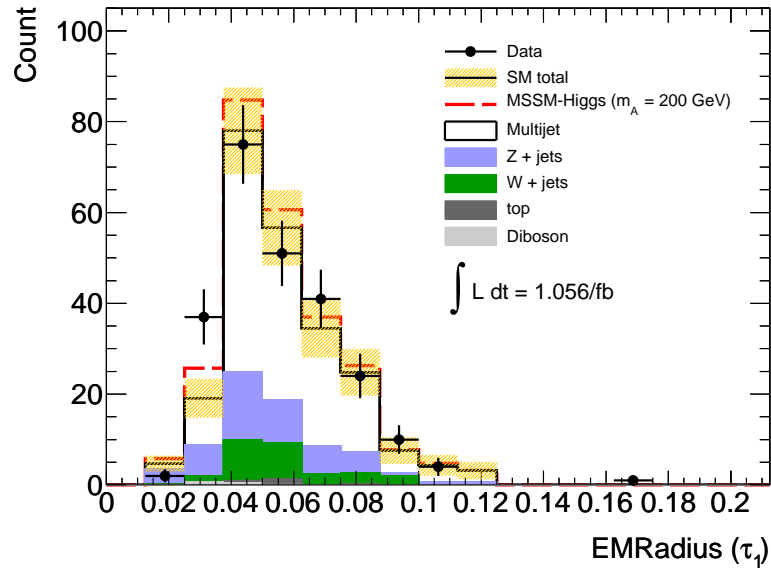
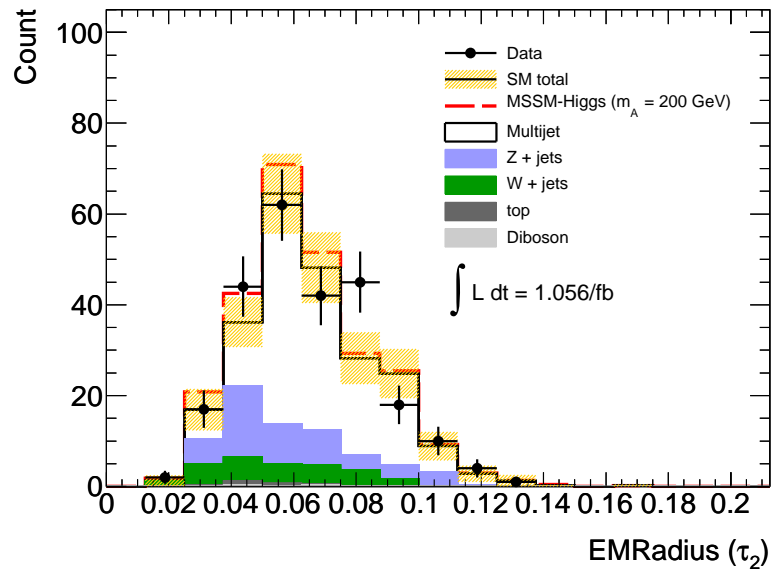
**Figure 14.2:** Pseudorapidity,  $\eta$ , of the highest- $p_T$  and second highest- $p_T$  hadronic tau candidate.

(a) highest  $p_T$   $\tau_h$  candidate(b) second highest  $p_T$   $\tau_h$  candidate

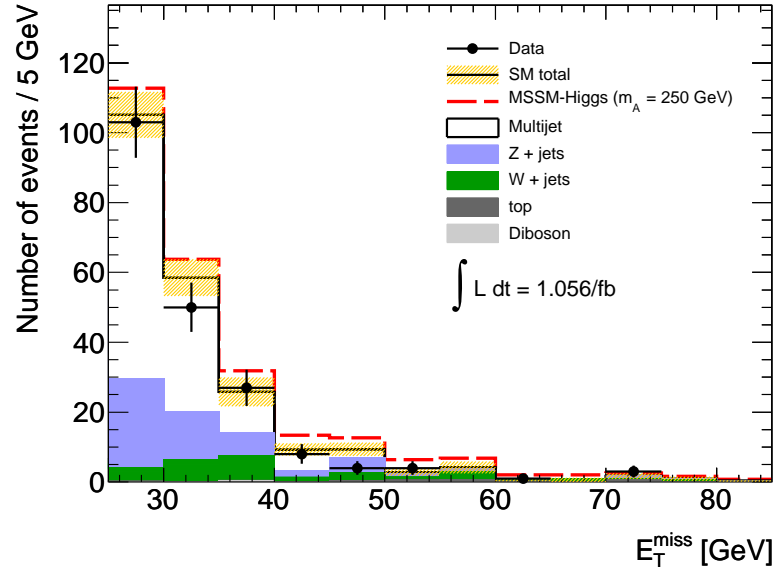
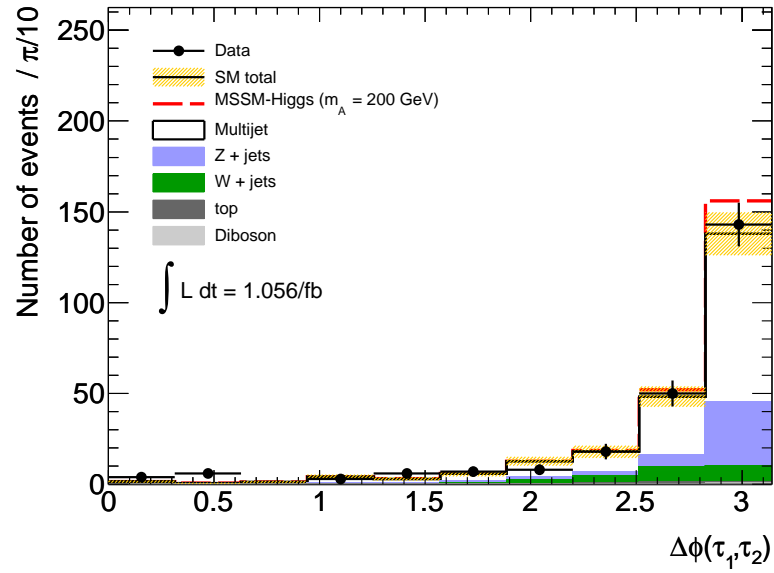
**Figure 14.3:** Azimuthal angle,  $\phi$ , of the highest- $p_T$  and second highest- $p_T$  hadronic tau candidate.

(a) highest  $p_T$   $\tau_h$  candidate(b) second highest  $p_T$   $\tau_h$  candidate**Figure 14.4:** Charge of the highest- $p_T$  and second highest- $p_T$  hadronic tau candidate.

(a) highest  $p_T$   $\tau_h$  candidate(b) second highest  $p_T$   $\tau_h$  candidate**Figure 14.5:** Number of tracks of the highest and second highest- $p_T$  hadronic tau candidate.

(a) highest- $p_T$  hadronic tau(b) second highest- $p_T$  hadronic tau

**Figure 14.6:** *Electromagnetic radius of the highest and second highest- $p_T$  hadronic tau candidate.*

Figure 14.7: *Missing transverse energy.*Figure 14.8: *Distance in azimuthal angle of the highest and second highest- $p_T$  hadronic tau candidate,  $\Delta\phi$ .*

Overall there is a good agreement between the distributions from data and from Monte Carlo, confirming that the background estimation methods work well. Now systematic uncertainties are considered, after which we're ready to consider the visible mass from the  $\tau_h\tau_h$  signal.

# Chapter 15

## Systematic Uncertainties

The systematic uncertainties considered in this analysis are determined using Monte Carlo. The data driven QCD background calculation is also affected by the systematic uncertainties on the Monte Carlo, since the non-QCD background contribution is subtracted from the observed number of events in data in each control region. The systematic uncertainties on the QCD background are evaluated by varying the non-QCD background per control region by their associated systematic uncertainties.

### 15.1 Monte Carlo cross-section predictions

#### 15.1.1 Signal cross-section

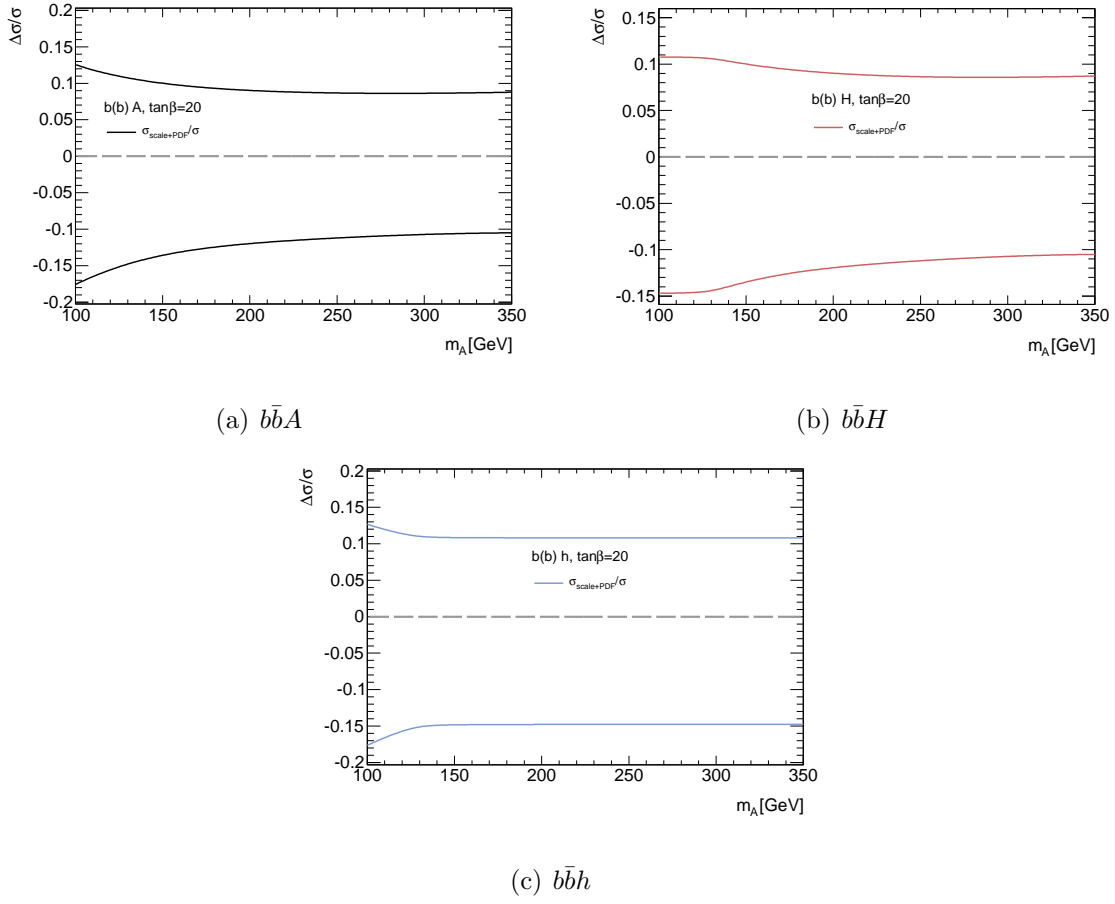
An uncertainty on the production and decay cross section ( $\sigma_{CS}$ ) is assigned for both signal and background Monte Carlo. The signal  $\sigma_{CS}$  uncertainty is both  $m_A$  and  $\tan\beta$  dependent. The Monte Carlo for the signal processes,  $b\bar{b}A/H/h \rightarrow \tau_h\tau_h$  and  $gg \rightarrow A/H/h \rightarrow \tau_h\tau_h$ , have been described in Section 7.1.1. For both production mechanisms, the cross section uncertainty is obtained by adding together linearly the parton distribution function (PDF) and  $\alpha_S$  scale uncertainties. (PDFs are described further in Appendix C). Table 15.1 gives the total uncertainty for  $m_A$  values of 200 and 300 GeV, with  $\tan\beta=20$ .

The final cross section uncertainty does not exceed 20% for signal sample generated for different values of  $m_A$ .

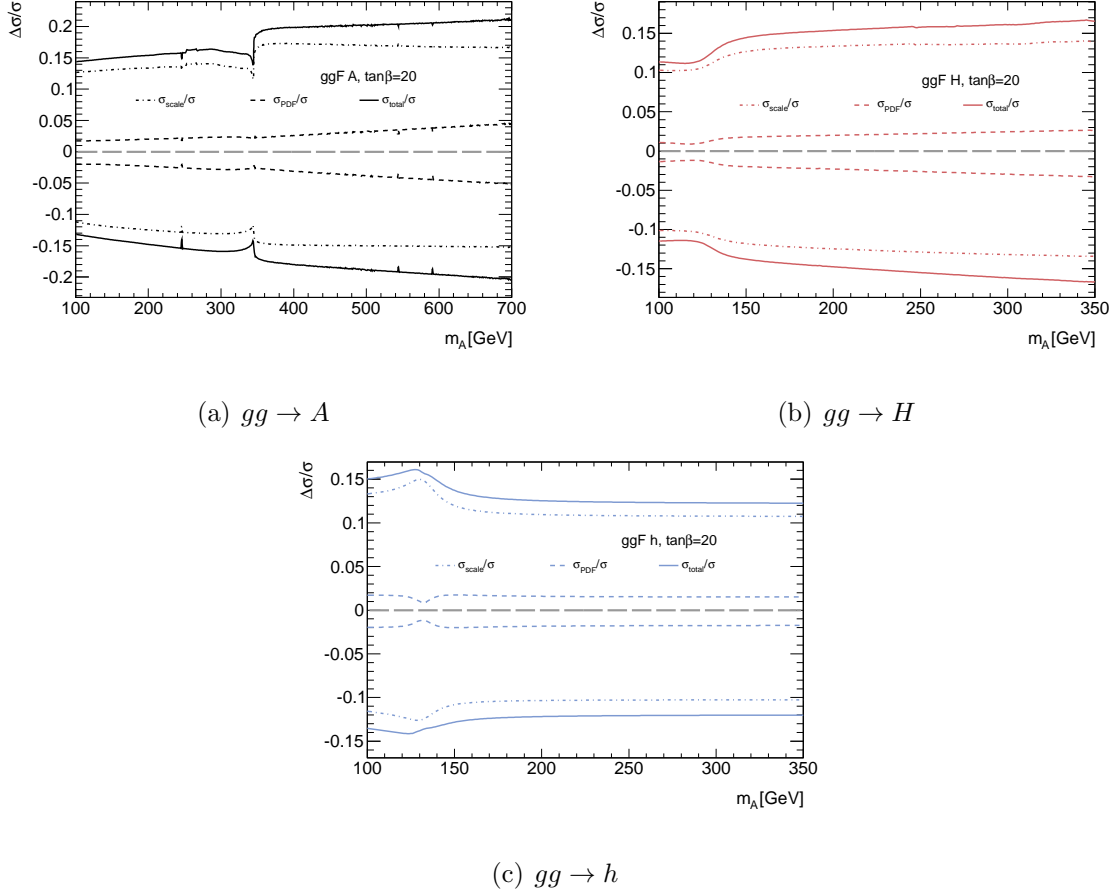
**Table 15.1:** *Production cross-section uncertainty for the signal with  $m_A$  of 200 GeV and 300 GeV, with  $\tan\beta = 20$ .*

	$m_A = 200 \text{ GeV}$	$m_A = 300 \text{ GeV}$
$b\bar{b}A$	+9.0%/-12.0%	+8.6%/-10.7%
$gg \rightarrow A$	+15.4%/-14.8%	+16.1%/-15.9%

Figures 15.1 and 15.2 show uncertainties on production cross section of  $b\bar{b}A/H/h$  and  $gg \rightarrow A/H/h$  processes as a function of  $m_A$  for  $\tan\beta = 20$ . For  $b\bar{b}A/H/h$  the total uncertainty is displayed, for  $gg \rightarrow A/H/h$  the PDF and scale uncertainty are shown independently as well as the combined uncertainty. In both cases the scale uncertainty dominates.



**Figure 15.1:** *Uncertainties on production cross section of  $b\bar{b}A/H/h$  process as a function  $m_A$  for  $\tan\beta = 20$*



**Figure 15.2:** *Uncertainties on production cross section of  $gg \rightarrow A/H/h$  process as a function  $m_A$  for  $\tan\beta = 20$ . PDF and scale uncertainties are shown separately, as well as the total uncertainty*

### 15.1.2 $Z$ and $W$ Background cross-section

The  $Z$  and  $W$  backgrounds are estimated from ALPGEN generated samples. Uncertainties on the production cross-section were assigned by varying generator parameters, whilst the shapes of the distributions coming from these samples have been verified with the embedding technique described in Chapter 12. Table 15.2 lists the varied parameters for the ALPGEN  $W/Z$ +jet samples with the associated uncertainties.

**Table 15.2:** *Systematic uncertainty for the  $Z$ +jets and  $W$ +jets production cross section from varying ALPGEN generator parameters.*

Sources	$Z$ +jets	$W$ +jets
MLM matching condition ( $p_T$ 20 GeV)	27.0 %	26.4 %
MLM matching condition ( $\Delta R$ 0.4 )	-15.5 %	-15.7 %
Matrix element parton 12 GeV	2.96 %	0.14 %
Matrix element parton $\Delta R > 0.4$	6.11 %	0.13 %
Different scale definition (iqopt2)	1.12 %	0.43 %
Different PDF set (MRST2001J)	5.94 %	5.23 %
$Q^2$ Scale (double)	0.36 %	0.33 %
$Q^2$ Scale (half)	0.45 %	0.78 %
kt factor (double)	0.37 %	0.44 %
kt factor (half)	0.37%	0.31 %

The dominant uncertainties come from MLM [86] matching conditions. ALPGEN generates a hard scattering process with an output that includes quarks and gluons. The showering of the partons is simulated in a different stage of Monte Carlo production, a more detailed description of which is given in Appendix B. If we wish to consider additional jets in the ALPGEN samples, these are included as partons radiated from

the hard scatter. This approach introduces the risk of adding a theoretically equivalent parton to one already generated in the hard scatter. Such double counting is avoided using MLM matching. This approach rejects generated events where the additional parton is contained in a region of phase space that is already treated by the hard scatter. In ATLAS this phase space is defined by jet  $p_T$  and the distance between two jets. By default the additional jets in ALPGEN have  $p_T > 15$  GeV, and must be separated from the hard scatter products by  $\Delta R < 0.7$ . Changing these conditions introduces the largest uncertainties on the ALPGEN generated  $Z$  and  $W$  samples.

The cross-section uncertainties on other Monte Carlo backgrounds were assigned by varying the PDF set. Since the event yield from the other background were small, any uncertainty on the cross-section leads to a small effect over all.

## 15.2 $\tau_h$ identification efficiency

The measured  $\tau_h$  identification efficiency is consistent with the efficiency predicted by Monte Carlo. The systematic uncertainties for  $\tau_h$  identification efficiency have been evaluated following the recommendations from Reference [79] and the uncertainties summarized in Table 9.4. For reconstructed  $\tau_h$  decays matched to a generator tau within  $\Delta R < 0.4$  we apply a scale factor to the over all event. The scale factor is  $1 \pm 0.085$  for  $\tau_h$  candidates passing **medium** identification requirements. The difference of the event yield based on this fluctuated event weight and the nominal event yield is used as a systematic uncertainty. Results are shown in Table 15.3.

## 15.3 Trigger efficiencies for $\tau_h$ decays

For the  $\tau_h$  trigger selection efficiency, scale factors were derived for each  $\tau_h$  trigger with a  $Z \rightarrow \tau\tau$  tag-and-probe method. These scale factors were applied to  $\tau_h$  candidates that cause the EF\_tau29\_medium1 or EF\_tau20\_medium1 trigger accept and are matched to

a hadronic  $\tau$  decay at generator level within  $\Delta R < 0.2$ . The systematic uncertainty is evaluated by varying the multiplicative scale factors according to their uncertainties documented in Table 11.2. The difference of the event yield based on this adapted event weight and the nominal event yield is used as a systematic uncertainty. Results are shown in Table 15.3.

## 15.4 Trigger fake rates of $\tau_h$ candidates

Mis-identified jets reconstructed as taus that trigger events are treated with fake rate scale factors, derived from  $W \rightarrow \mu\nu_\nu + \text{jet}$  events. A scale factor is applied to events where an object that fires the EF\_tau29\_medium1 or EF\_tau20\_medium1 trigger is identified as a medium  $\tau_h$  candidate, and is not matched to a hadronic tau decay at generator level within  $\Delta R < 0.2$ . A systematic uncertainty on the event yield is estimated by varying the scale factor due to the uncertainties documented in Table 11.4. Results are shown in Table 15.3.

	$b\bar{b} \rightarrow \tau_h \tau_h$	$gg \rightarrow \tau_h \tau_h$	W + jets	Z/ $\gamma^*$ + jets	Top	Di-boson	QCD
Trigger efficiency	$\pm_{12}^{14}\%$	$\pm_{13}^{14}\%$	$\pm 6.6\%$	$\pm 13\%$	$\pm 12\%$	$\pm_{13}^{14}\%$	$\pm_{0.5}^{0.6}\%$
$\tau_h$ id. eff	$\pm_{15}^{16}\%$	$\pm_{15}^{16}\%$	$\pm_{7.7}^{7.1}\%$	$\pm_{15}^{16}\%$	$\pm_{13}^{14}\%$	$\pm_{15}^{16}\%$	$\pm_{0.7}^{0.7}\%$
Jet $\rightarrow \tau_h$ fake rate	0%	0%	$\pm 20\%$	$\pm 0.9\%$	$\pm_{4.7}^{4.9}\%$	$\pm 0.4\%$	$\pm 1.2\%$

**Table 15.3:** *Systematic uncertainties from the variation of trigger,  $\tau_h$  identification efficiency and the combined trigger and jet  $\rightarrow \tau_h$  fake rate.*

## 15.5 Tau energy scale

The  $p_T$  of the tau candidates is scaled according to the uncertainties obtained in an independent study, with the results documented in Reference [79]. The scaling is dependent

on:

- Number of tracks associated to  $\tau_h$
- $p_T$  of the  $\tau_h$  candidate
- $\eta$  of the  $\tau_h$  candidate

The uncertainties vary between 3.5% and 5%, and are given in Table 15.4 for relevant single and multi-prong taus.

$p_T$	$0. <  \eta  \leq 1.3$	$1.3 <  \eta  \leq 1.6$	$1.6 <  \eta  \leq 2.5$
single-prong: $> 30$ GeV	3.5%	5%	4.5%
multi-prong: 30 - 40 GeV	5.5%	5.5%	5.5%
multi-prong: $> 40$ GeV	4.5%	5%	5%

**Table 15.4:** *Single and multi-prong  $\tau_h$  energy scale systematics.*

## 15.6 $E_T^{Miss}$ uncertainties

Acceptance of an event in this analysis depends on  $E_T^{Miss}$ , which is constructed from topological clusters. To address uncertainties in the  $E_T^{Miss}$ , procedures that have been implemented in previous analyses and documented in [62, 63, 64] were used. The systematic uncertainties from topological cluster energy, discussed below, will also effect the tau energy scale, and these effects have been taken into account in the values in Table 15.4.

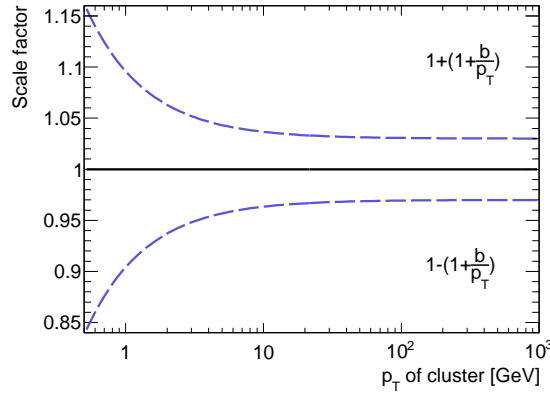
### 15.6.1 Topological cluster energy scale

Uncertainty on the topological cluster energy scale have been derived using energy/momentum ( $E/p$ ) studies from single charged hadrons [33, 80, 81, 82]. An  $E/p$  study considers the reconstructed momentum from the track, and the deviation in the reconstructed energy

of associated topological clusters. This deviation is at maximum 10% for a  $p_T$  of 500 MeV, and approximately 3% at higher  $p_T$  values. Clusters within  $|\eta| < 3.2$ , where  $\tau_h$  objects are expected from  $W/Z$  and Higgs decays, are scaled according to the following empirically derived factor:

$$1 \pm a \left( 1 + \frac{b}{p_T \text{ GeV}} \right) \quad (15.1)$$

Here, equation 15.1,  $a$  indicates the jet energy scale deviation at high momentum, and  $b$  is fixed at 2.2 to give a 10% scale factor at 500 MeV. The empirical functions, with  $a = 3\%$ , are shown in Figure 15.3.



**Figure 15.3:** Functions used to scale topological cluster energy.

In the forward region of the detector,  $|\eta| \geq 3.2$ , the energy scale uncertainty is estimated by data to correspond to  $a = 10\%$  [80]. Topological clusters in the forward region are scaled by this amount. However, the effect on event acceptance is small since most selected events do not have large transverse energy in this region. The results on the accepted number of events are summarized in Table 15.5 for signal ( $m_A = 200$  GeV and  $\tan \beta = 20$ ),  $W + \text{jets}$ ,  $Z/\gamma^* + \text{jets}$ , top, di-boson and QCD background contributions.

### 15.6.2 $E_T^{miss}$ resolution

The  $E_T^{Miss}$  resolution has been measured as  $0.49\sqrt{\sum E_T}$  in minimum bias events [83]. This resolution is degraded in the presence of high  $p_T$  jets [81]. To understand the effect

		$b\bar{b} \text{ A} \rightarrow \tau_h \tau_h$	$gg \rightarrow \text{A} \rightarrow \tau_h \tau_h$	W + jets	Z/ $\gamma^*$ + jets	Top	Di-boson	QCD
$ \eta  < 3.2$								
$a$	$b$							
0.03	2.2	11%	6%	34%	63%	12%	26%	-1.2%
-0.03	2.2	-10%	-6%	-21%	-23%	-13%	-12%	1.5%
$ \eta  \geq 3.2$								
$a$								
0.10		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
-0.10		-0.7%	-0.1%	-1.4%	-0.3%	0.0%	2.5%	0.1%

**Table 15.5:** *Relative variation of acceptances due to systematic uncertainty of the topological cluster energy scale for signal, electroweak and top backgrounds.*

of the  $E_T^{Miss}$  resolution on the acceptance of this analysis a gaussian smearing was applied on the  $x$  and  $y$  components of the  $E_T^{Miss}$ . The smearing parameters were chosen to allow the factor  $\alpha$  in  $\alpha\sqrt{\sum E_T}$  to cover the uncertainty produced by high  $p_T$  jets in an event. Table reftab:Sys:ERes shows the acceptance uncertainty for  $\alpha$  of 0.5 and 0.55. The larger uncertainty, from  $\alpha = 0.5$  was used.

		$b\bar{b} \text{ A} \rightarrow \tau_h \tau_h$	$gg \rightarrow \text{A} \rightarrow \tau_h \tau_h$	W + jets	Z/ $\gamma^*$ + jets	Top	Di-boson	QCD
$\alpha \text{ [GeV}^{1/2}]$								
0.50		-0.7%	0.9%	-3.3%	-2.5%	0.0%	2.2%	-1.3%
0.55		-0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

**Table 15.6:** *Effect on event acceptance by varying  $E_T^{Miss}$  resolution.*

## 15.7 Electron and Muon associated uncertainties

Any mis-measurement of the electron (muon) reconstruction, identification, or energy calculation will influence the analysis event selection at the electron (muon) veto. Since the contribution of events from all  $Z \rightarrow ee$ ,  $W \rightarrow e\nu$ ,  $Z \rightarrow \mu\mu$  and  $W \rightarrow \mu\nu$  is small, as

shown in Table 10.1, we expect this effect to be small also.

### 15.7.1 Reconstruction and Identification efficiencies

Other electron and muon based studies have derived scale factors applicable to Monte Carlo [61], to reproduce what is seen in data. For electrons two independent scale factors are provided for reconstruction and identification. Each of these depends on the electron  $p_T$  and  $\eta$ . With muons, a combined reconstruction and identification scale factor binned in muon  $p_T$  and  $\eta$  is given. To understand the effect these scale factors have on an event **veto** in the presence of an electron or muon, we consider events with one or no electrons (muons) separately. Here a reconstructed electron or muon refers to the criteria used in Chapter 10. For each event with one reconstructed electron (muon), we weight the event under the assumption that this is fake electron (muon):

$$\text{Event weight} = \begin{cases} 1 - \text{Scale factor}(p_T, \eta) & \text{one electron (muon) was reconstructed} \\ 1 & \text{no electron (muon) was reconstructed} \end{cases} \quad (15.2)$$

The Scale factor used in equation 15.2 refers to either the product of the electron identification and reconstruction factors, or the combined muon factor, all of which are less than 3% for all bins. In the case of Monte Carlo with one generated electron (muon), the effect of an electron (muon) veto is well reproduced in reconstruction. The cut efficiency is correctly described using this event weight procedure. However, with two truth generated electrons (muons), the scale factor dependent event weight does not model the electron (muon) veto well. This is primarily because this veto is very effective against backgrounds with more than one true electron (muon). These are completely suppressed, within the statistical limitations of the available simulated data. Therefore, the systematic effect on such backgrounds is assumed negligible. For events with one reconstructed electron or muon, a systematic uncertainty is assigned by the difference

in event acceptance with and without the event weight given in equation 15.2. Any uncertainty on the scale factors themselves are ignored. Table 15.7 summarizes the result for the electron and muon identification uncertainty.

	$b\bar{b} \text{ A} \rightarrow \tau_h \tau_h$	$gg \rightarrow \text{A} \rightarrow \tau_h \tau_h$	W + jets	Z/ $\gamma^*$ + jets	Top	Di-boson	QCD
Electron rec. eff.	$\pm 0.1\%$	$\pm 0.1\%$	$\pm 0.7\%$	$\pm 0.4\%$	$\pm 0.3\%$	$\pm 0.7\%$	$\pm 0.2\%$
Electron id. eff.	$\pm 0.1\%$	$\pm 0.1\%$	$\pm 0.3\%$	$\pm 0.3\%$	$\pm 0.1\%$	$\pm 0.8\%$	$\pm 0.1\%$
Muon eff.	$\pm 0.1\%$	$\pm 0.1\%$	$\pm 0.3\%$	$\pm 0.4\%$	$\pm 0.0\%$	$\pm 0.8\%$	$\pm 0.1\%$

**Table 15.7:** *Effect on analysis acceptance by the systematic uncertainties from electron and muon reconstruction and identification.*

### 15.7.2 Energy scale

There are also scaling and smearing factors for the energy or momenta of electrons or muons provided from other analyses. These factors allow the Monte Carlo electron  $E_T$  distribution and muon  $p_T$  distribution to match the distributions from data. No acceptance changes were seen by the application of these scale/smearing factors, and any systematic uncertainty here is neglected.

## 15.8 Luminosity in data

The uncertainty on the integrated luminosity recorded by ATLAS is 3.7% [65].

## 15.9 Summary of Systematic Uncertainties

All individual systematic uncertainties considered are shown in Table 15.8. The relative uncertainty is given for individual sources, and a combined absolute uncertainty on the number selected events is given in the final row.

A large uncertainty in the  $Z$  and  $W$  acceptance comes from the hadronic  $\tau$  energy scale. This occurs because of the shape of the tau  $E_T$  distribution coming from these sources. The generator level  $p_T$  forms a landau distribution, peaked at approximately 30 GeV, with most statistics falling within the 15 to 60 GeV range. Any systematic uncertainty on the  $E_T$  scale shifts this entire distribution to the right or left. Since this analysis cuts on  $\tau$   $p_T$  at 45 and 30 GeV, where a large number of events from  $Z$ s and  $W$ s are expected, any shift in the  $p_T$  spectrum greatly effects the acceptance.

	$b\bar{b} \text{ A} \rightarrow \tau_h \tau_h$	$gg \rightarrow \text{A} \rightarrow \tau_h \tau_h$	W + jets	Z/ $\gamma^*$ + jets	Top	Di-boson	QCD
Central values [events]	17.9	5.2	25	48	2.0	1.4	157
Statistical unc. [events]	$\pm 1.2$	$\pm 0.3$	$\pm 5$	$\pm 5$	$\pm 0.9$	$\pm 0.3$	$\pm 18$
	$\pm 7$	$\mp 0.4$					
$\sigma_{CS}$	$\pm 9\%$	$\pm 9\%$	$\pm 20\%$	$\pm 14\%$	$\pm 9\%$	$\pm 7\%$	$\mp 1.4\%$
Luminosity	$\pm 3.7\%$	$\pm 3.7\%$	$\pm 3.7\%$	$\pm 3.7\%$	$\pm 3.7\%$	$\pm 3.7\%$	$\mp 0.3\%$
Trigger efficiency	$\pm 13\%$	$\pm 13\%$	$\pm 6.6\%$	$\pm 13\%$	$\pm 12\%$	$\pm 13\%$	$\mp 0.5\%$
$\tau_h$ id. eff.	$\pm 15\%$	$\pm 15\%$	$\pm 7.4\%$	$\pm 15\%$	$\pm 13\%$	$\pm 15\%$	$\mp 0.7\%$
Jet $\tau_h$ misid.	0	0	$\pm 20\%$	$\pm 0.9\%$	$\pm 4.8\%$	$\pm 0.4\%$	$\mp 1.2\%$
Energy scale	$\pm 10\%$	$\pm 6\%$	$+34\%/-21\%$	$+63\%/-23\%$	$\pm 12\%$	$+26\%/-12\%$	$\mp 1\%$
Muon and elec. rec./id.	$\pm 0.2\%$	$\pm 0.2\%$	$\pm 0.8\%$	$\pm 0.6\%$	$\pm 0.3\%$	$\pm 1.3\%$	$\mp 0.2\%$
Total syst. unc. [events]	$\pm 5$	$\pm 1.4$	$\pm 10$	$+33/-16$	$\pm 0.5$	$\pm 0.4$	$\mp 4$

**Table 15.8:** All systematic uncertainties for the electroweak, top, and di-boson backgrounds, the signal based simulation, as well as the data-driven estimation of QCD background. Statistical uncertainty is given for the Monte Carlo samples, and is propagated to the QCD background. Individual systematic uncertainties are quoted as relative values, and the combined uncertainty is given as an absolute value. Uncertainties with similar lower and upper variations are quoted symmetrically.

# Chapter 16

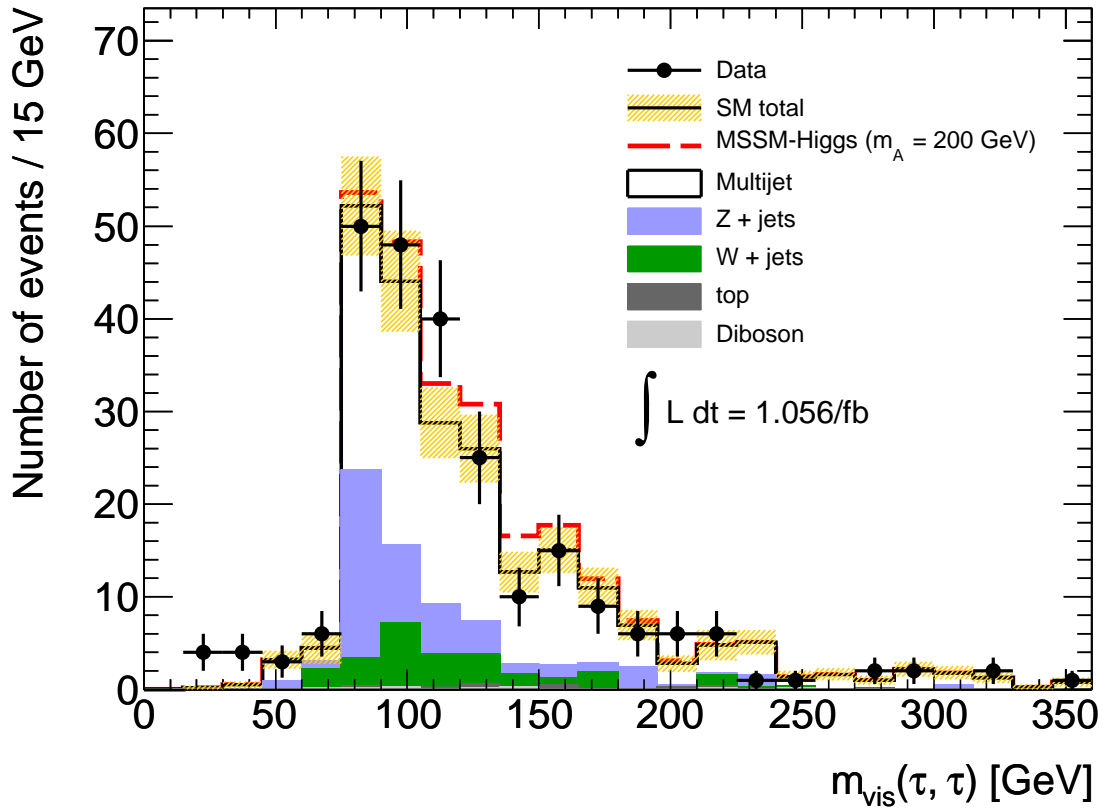
## Results

The visible mass from the two hadronic taus in selected events is shown in Figure 16.1, along with the signal expected from a Higgs with  $m_A = 200$  GeV and  $\tan\beta = 20$ . No excess of events above the expectation is observed, and we set exclusion limits for the production of neutral MSSM Higgs bosons as function of the parameters  $m_A$  and  $\tan\beta$ .

The expected number of events for signal (s) and background (b), and the observed number of events (N) for each bin in the visible mass distribution are combined in a likelihood function,  $\mathcal{L}(\mu, \theta)$ . The  $\mu$  term is a signal strength parameter that multiplies the expected signal in each bin. The value  $\mu = 0$  corresponds to the background only hypothesis, and  $\mu = 1$  is the nominal signal plus background hypothesis [85]. If we take  $\mu > 1$ , we are enhancing the theoretical predications for the neutral MSSM Higgs bosons cross-sections. Signal and background predictions depend on systematic uncertainties that are parameterised by nuisance parameters  $\theta$ , which in turn are constrained using Gaussian functions. The likelihood is then defined as:

$$\mathcal{L}(\mu, \theta) = \prod_j \text{Poisson}(N_j | \mu s_j + b_j) \prod_{\theta} \text{Gaussian}(\theta | 0, 1), \quad (16.1)$$

where  $j$  labels each bin in the visible mass distribution. A ratio of likelihoods is used to



**Figure 16.1:** Visible mass of the two leading  $\tau_h$  candidates for backgrounds, signal and data. Also shown is a  $m_A = 200$  GeV Higgs Boson in the  $m_h^{\text{max}}$  scenario with  $\tan \beta = 20$ .

construct a test statistic given by:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \frac{L(\mu, \hat{\vec{\theta}}(\mu))}{L(0, \hat{\vec{\theta}}(0))} & \hat{\mu} < 0 , \\ -2 \ln \frac{L(\mu, \hat{\vec{\theta}}(\mu))}{L(\hat{\mu}, \hat{\vec{\theta}})} & 0 \leq \hat{\mu} \leq \mu , \\ 0 & \hat{\mu} > \mu . \end{cases} \quad (16.2)$$

Here,  $\vec{\theta}$  represents the nuisance parameters,  $\hat{\mu}$ ,  $\hat{\vec{\theta}}$ , and  $\hat{\vec{\theta}}(\mu)$  are the Maximum Likelihood Estimators (MLE) of  $\mu$ ,  $\vec{\theta}$ , and the nuisance parameters evaluated at  $\mu$ , respectively. From the test statistic, the p-value for signal plus background, using some  $\mu$ , is calculated from:

$$p_{s+b} = \int_{\tilde{q}_{\mu, obs}}^{\infty} f(\tilde{q}_\mu | \mu, \hat{\vec{\theta}}(\mu)) d\tilde{q}_\mu \quad (16.3)$$

The p-value for the background only hypothesis is given by:

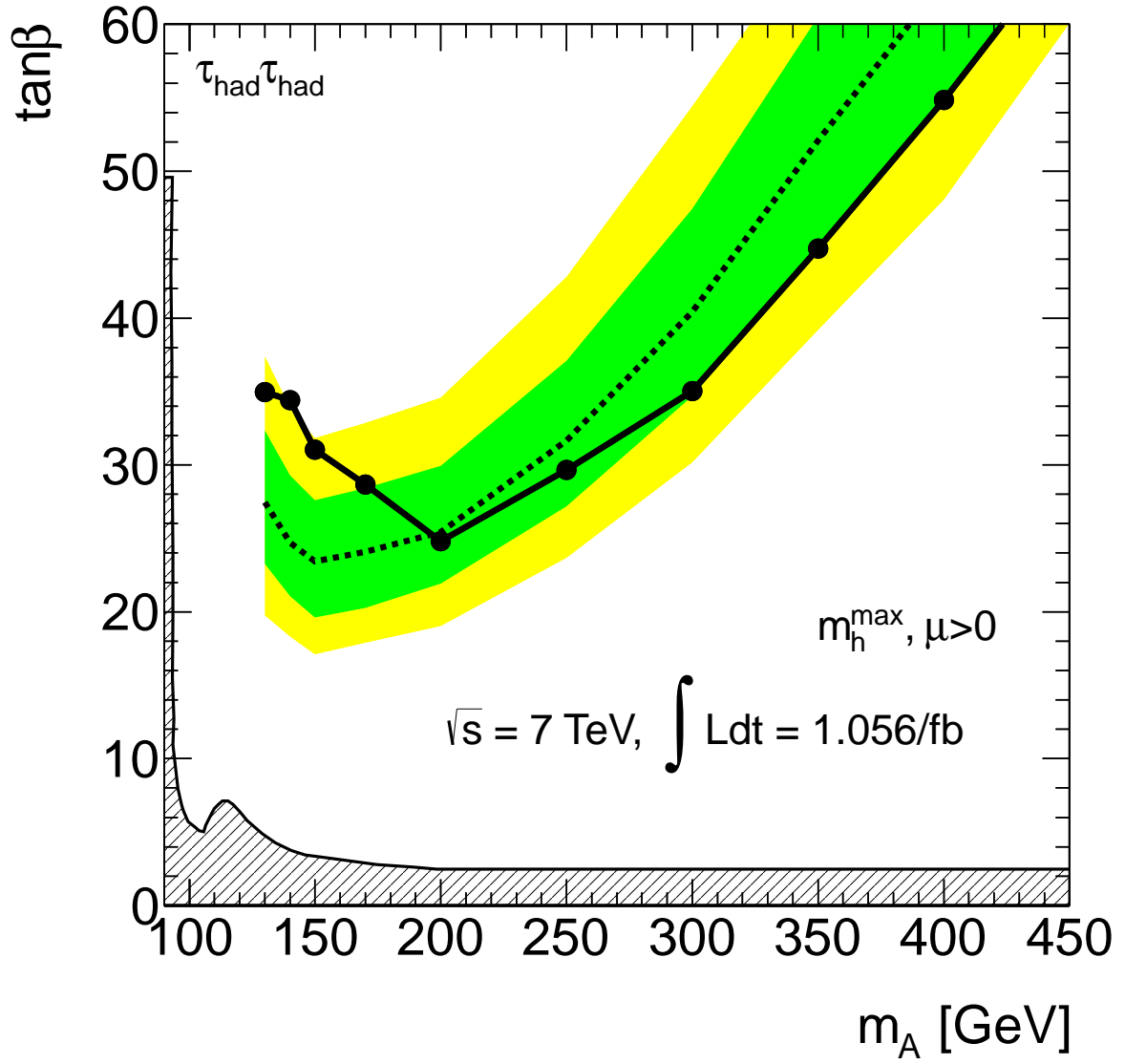
$$p_b = \int_{-\infty}^{\tilde{q}_{\mu=0, obs}} f(\tilde{q}_{\mu=0} | \mu = 0, \hat{\vec{\theta}}(\mathbf{0})) d\tilde{q}_\mu \quad (16.4)$$

A point in the  $(\tan \beta, m_A)$  phase space is excluded if the signal hypothesis is rejected at the 95% Confidence Level, using the  $CL_S$  method. The  $CL_S$  method protects against excluding models when there is little or no sensitivity, by considering the p-value for both the background plus signal and background only hypotheses. A signal model is regarded as excluded if:

$$\frac{p_{s+b}}{1 - p_b} < 0.95 \quad (16.5)$$

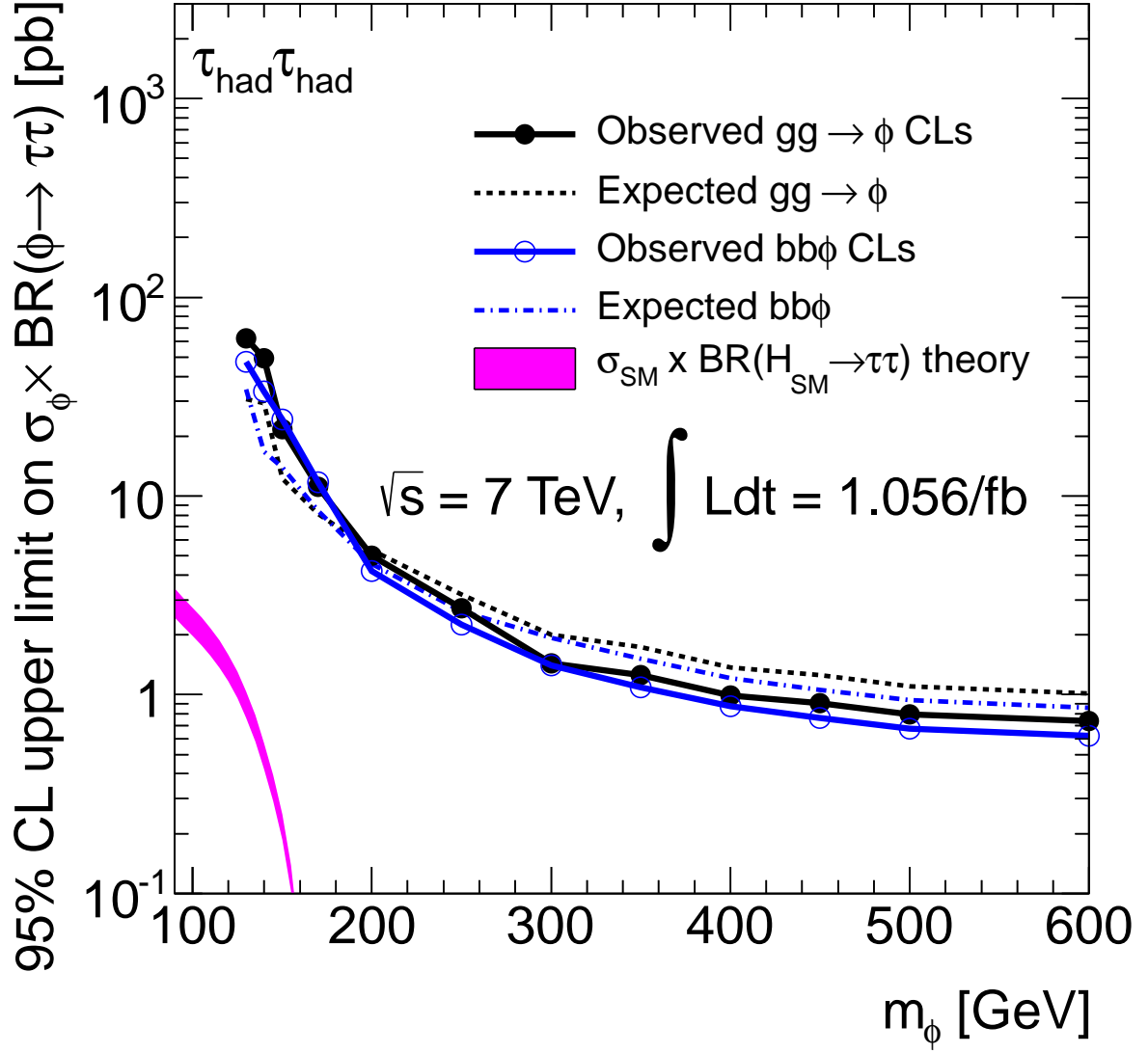
The exclusion, set using  $\mu = 1$ , for the  $m_h^{max}$  scenario, in the  $m_A$  and  $\tan \beta$  phase space is shown in Figure 16.2. The expected and observed 95% confidence-level limits are shown as solid and dashed black lines, respectively. The green and yellow bands correspond to the  $1\sigma$  and  $2\sigma$  error bands. The obtained limit excludes regions of parameters space beyond the existing limits from previous experiments at the Tevatron [1].

Additionally, it is possible to set a production cross-section limit on a model independent Higgs, denoted  $\phi$ , that can decay to two taus. Figure 16.3 shows this limit for



**Figure 16.2:** Limit in  $m_A$  and  $\tan\beta$  phase space for MSSM Higgs derived in this analysis. The grey shaded area shows the limit from LEP.

$b$  quark association (blue line) and gluon fusion (black line) separately. Expected limit and observed limit are shown as solid and dash lines respectively, the area above the lines is excluded. For comparison, the cross-section for a Standard Model Higgs decaying to two taus is shown in magenta.



**Figure 16.3:** Limit for cross-section of a generic Higgs boson, denoted  $\phi$ , decaying to two hadronic taus as derived in this analysis.

# Chapter 17

## Conclusions

The MSSM addresses some of the major theoretical concerns associated with the SM. The Higgs sector contains five physical Higgs bosons, which have different particle couplings than a SM Higgs. The tau decay channel is one of the most promising in the search for an MSSM Higgs, and since taus decay hadronically in the majority of cases, the search for  $A/H/h \rightarrow \tau_h \tau_h$  is valuable.

This thesis has presented the first such search at the LHC, using proton-proton collisions with a center-of-mass energy of  $\sqrt{s} = 7$  TeV, recorded with the ATLAS detector. With  $1.056 \text{ fb}^{-1}$  of collected data 245 selected events were observed, consistent with the expected background of  $256 \pm 26$  events. As no excess of events was seen, exclusion limits for MSSM Higgs production were derived as a function of  $m_A$  and  $\tan\beta$  in the  $m_h^{max}$  scenario.

The ATLAS detector continues to take data, and at the time of writing the LHC center-of-mass energy has increased to  $\sqrt{s} = 8$  TeV. Recent discovery of a new boson at 125-126 GeV has stirred excitement, and the increased data will help gain a complete understanding of the particle, including coupling to taus. We have seen from this analysis that tau final states become difficult to probe at low resonant masses because of a dominant irreducible  $Z \rightarrow \tau\tau$  background. Studies to better understand taus in the  $p_T$

range produced from  $Z$  decays are in progress, and necessary for a full understanding of the newly discovered boson. Such studies and additional data will also extend our exploration of the MSSM phase space, and in time we will see if the theory can survive the observed reality.

# Appendices

# Appendix A

## Run numbers and data periods

A subset of 2011 data with all **ATLAS** sub-detectors functioning well are used. There are 84 run numbers:

178044, 178047, 178109, 179725, 179739, 179771, 179804, 179938, 179939, 179940, 180122, 180124, 180139, 180144, 180149, 180153, 180164, 180212, 180225, 180241, 180242, 180309, 180400, 180448, 180481, 180614, 180636, 180664, 180710, 180776, 182013, 182161, 182284, 182346, 182372, 182424, 182449, 182450, 182454, 182455, 182456, 182486, 182516, 182726, 182747, 182766, 182787, 182796, 182879, 182886, 182997, 183003, 183021, 183038, 183045, 183054, 183078, 183079, 183081, 183127, 183129, 183130, 183216, 183272, 183286, 183347, 183391, 183407, 183412, 183426, 183462, 183544, 183580, 183581, 183602, 183780, 183963, 184022, 184066, 184072, 184074, 184088, 184130, 184169.

Each run consists of 50 ns bunch trains. Runs are combined in data taking periods with common parameters such as beam intensity and trigger menus. Generally, the later data periods have higher beam intensity and the triggers are altered to constrain event write out. In particular, trigger performance is evaluated separately between Periods **B<sub>2</sub>** - **E** and **F** - **H** because of  $\tau_h$  Level 1 trigger changes to accommodate rising instantaneous

luminosity.

The data periods referenced in this analysis are listed in table A.1.

Data Period	Included runs	Total number of runs
B <sub>2</sub>	178044 - 178109	3
D	179710 - 180481	23
E	180614 - 180776	5
F	182013 - 182519	17
G	182726 - 183462	28
H	183544 - 184169	13

**Table A.1:** *ATLAS data periods and runs used in this analysis.*

During data period **E** 6 front end boards on the A-side of the LAr Barrel calorimeter lost operational function. The localized region of lost output is referred to as the LAr hole, and is treated in this analysis (see Chapter 10). Tau identification performance in data has been analyzed using data between periods **B**<sub>2</sub> to **G** as detailed in Section. 9.3.

# Appendix B

## Monte Carlo Generators

The main MC generators used in this analysis are PYTHIA [66], HERWIG [70], and SHERPA [68]. Each of these is a general purpose generator capable of simulating all components in an event. All three use leading order calculations, which at times is of limited accuracy for effective comparison with data. Additional specialized generators can be interfaced with the general purpose generators to replace or improve a part of the simulation chain. In this analysis, these additional generators are: gg2WW [73], ALPGEN [69], and MC@NLO [67]. The next-to-leading order calculations for the loop induced gluon-fusion process  $gg \rightarrow W^*W^* \rightarrow \text{leptons}$  are available in gg2WW. While this is greatly beneficially in  $H \rightarrow WW$  searches, the effect is negligible for this analysis as the improvement is applied to a very small diboson background. MC@NLO includes a full NLO treatment for certain events with up to one extra parton, this generator is used whenever the hard interaction is available. The ALPGEN generator is a collection of LO generators for a variety of processes with up to six associated jets. This is used for  $W$  and  $Z$ + jets events, in order to better model the influence of additional jets, which are abundant in the QCD dominated background at the LHC. For these events separate samples of  $W/Z$ + 0, 1, 2, 3, 4, 5 partons are generated, which are summed together to get a complete  $W/Z$ + jets set of events. Each of these subprocesses includes

additional jets as hard radiation that are processed in the parton showering phase of the generation. Parton showering and hadronization is simulated with HERWIG [70] for MC@NLO, ALPGEN and gg2WW, and by PYTHIA itself for the PYTHIA generator.

In parton showering the radiation from incoming and outgoing coloured objects is approximated by DGLAP splitting functions [87] together with Sudakov form factors. [71] for both PYTHIA and HERWIG. Splitting functions allow one quark or gluon to split into two coloured partons, depending on the number of flavours and the momentum fractions carried away by the two outgoing functions. Only three splitting functions exist for QCD at leading order, these account for  $q \rightarrow qg$ ,  $g \rightarrow gg$  and  $g \rightarrow q\bar{q}$ . The Sudakov form factors give the probability of evolving from an initial time  $t_0$  to a later time  $t$  without branching [71]. Both initial and final state radiation are part of the parton showering step.

A double counting complication arises in some regions of phase space, as emissions are accounted for using both matrix elements and parton showering. ALPGEN, the generator where this is most relevant as we consider it for all  $W/Z$ + jets processes, using a scheme called MLM matching [86] to overcome the problem. It uses a veto on events that contain emissions from parton showers in the region already covered by the matrix elements. In ATLAS the phase space for MLM matching is defined by jet transverse energy and the radial distance between two jets ( $\Delta R$ ).

### B.0.1 QCD interlude

The theory and modelling of the electroweak processes is fairly well understood. QCD on the other hand is more complicated when generating events. In theory, QCD explains the interaction between quarks and gluons through three colour charges. The strength of the interaction is described by a running coupling that depends on the four momentum transfer,  $Q^2$ . The coupling decreases logarithmically, with a proportionality that depends

on  $Q^2$  and the QCD energy scale,  $\Lambda_{QCD}$ , as given by Equation B.1:

$$\alpha_S(Q^2) \propto \frac{1}{\ln(Q^2/\Lambda_{QCD}^2)}. \quad (\text{B.1})$$

This is referred to as asymptotic freedom. The QCD energy scale is not defined, but determined experimentally to be  $\sim 200$  MeV. A consequence of asymptotic freedom is that interactions between quarks and gluons becomes weak at high energy. In the region where  $Q^2 \gg \Lambda_{QCD}^2$ , it is possible to apply perturbative techniques to the theory. In this energy range the quarks and gluons can be thought of as free. Experimentally, however, isolated quarks have not been observed. We have only seen the results of QCD interactions in the production of familiar mesons and baryons. This is due to a concept called colour confinement, which is why we only observe colourless particles. The process through which ‘free’ quarks and gluons produce a jet of colourless particles is not well understood. The mechanism is modelled through hadronization and fragmentation.

Hadronization describes the formation of colourless hadrons. PYTHIA uses a string fragmentation model, and HERWIG uses a cluster fragmentation model. In string fragmentation the colour field is represented as a string between a quark-antiquark pair. As partons move apart, the string breaks creating new  $q\bar{q}$  pairs. Hadrons are formed from all the available quarks at a preset low enough energy. In cluster fragmentation gluons remaining after parton showering are split into quark antiquark pairs. All available quarks and antiquarks are combined into massive clusters of colour singlets (ie: colourless particles), that are allowed to decay into two hadrons [70]. After all final hadrons have been formed, the unstable ones are decayed into stable particles. The decay branching ratios are taken from experimental data as much as possible.

The underlying events are simulated using JIMMY [72] for HERWIG, and PYTHIA uses its own internal mechanism. Both models use the idea that other than hard interactions, the 2→2 QCD process dominates, because it is the lowest order in  $\alpha_S$ . **ATLAS** specific details for the data period used in this thesis can be found in reference [21].

TAUOLA [77] and PHOTOS [78] are used to model the tau lepton decay and additional photon radiation from charged leptons in the leading-log approximation for all generators except for SHERPA.

# Appendix C

## Parton Distribution Functions

A parton distribution function (PDF) is the probability density for finding a parton with longitudinal momentum fraction  $x$  at momentum transfer  $Q^2$ . As QCD is non-perturbative for partons bound in a proton, PDFs cannot be obtained through perturbative calculations. The known PDFs are instead obtained empirically by fits to data. Figure C.1 shows the PDFs from MRST2008 [88], for  $Q^2 = 10 \text{ GeV}^2$  and  $Q^2 = 10^4 \text{ GeV}^2$ .

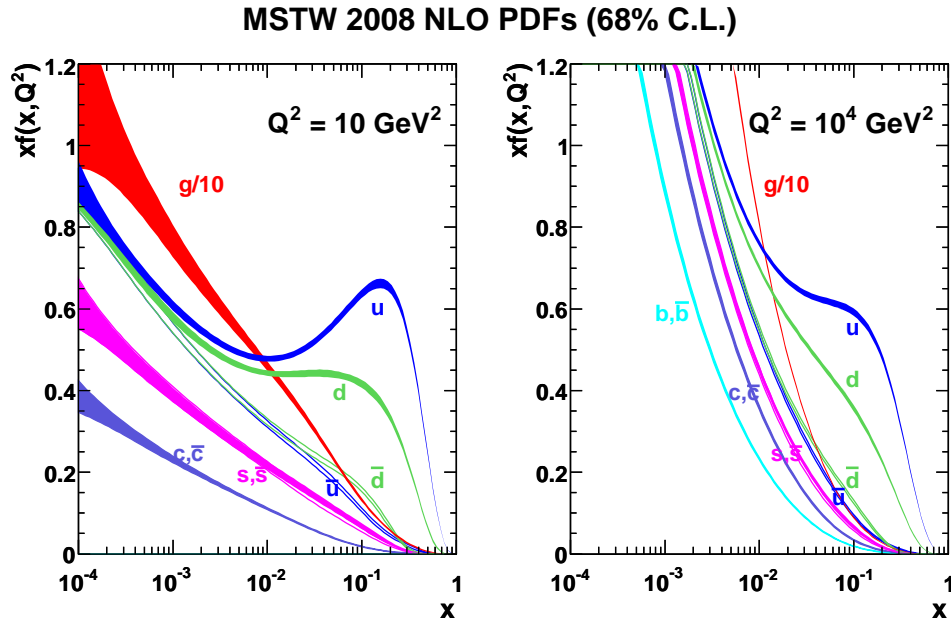


Figure C.1: *MSTW 2008 NLO PDFs at  $Q^2 = 10 \text{ GeV}^2$  and  $Q^2 = 10^4 \text{ GeV}^2$ .*

In proton-proton collisions at the LHC, the computation of cross-sections require the initial energies of the colliding partons. As these are inferred from emperical data, there is an associated uncertainty that must be considered.

## Appendix D

# Signal cross-sections predicted from Monte Carlo

This appendix lists all mass points generated for the signal (CP odd  $A^0$  Higgs boson).  $b$ -associated production samples were generated with SHERPA. Gluon fusion samples were generated with MC@NLO. All samples are generated for  $\tan\beta = 20$ .

**Table D.1:** *b-associated production signal MC samples and cross section including branching ratios. BR is a branching ratio of  $A/H \rightarrow \tau\tau \rightarrow \tau_h\tau_h$ , which is  $BR(A/H \rightarrow \tau\tau) \times BR^2(\tau \rightarrow \tau_h)$  ( $=0.648$ ).*

Process	$m_A$ (GeV)	$\sigma_{A/H} \times$ BR( $A/H \rightarrow \tau\tau \rightarrow \tau_h\tau_h$ ) (fb) for $\tan\beta = 20$ with 5FS
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	100	20546.2/34.8
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	110	15229.5/92.1
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	120	11489.2/370.8
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	130	8801.9/3738.0
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	140	6834.9/6327.8
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	150	5371.6/5239.4
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	170	3422.9/3393.9
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	200	1846.5/1847.0
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	250	747.8/755.5
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	300	335.4/344.2
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	350	155.7/166.9
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	400	80.8/86.3
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	450	44.1/44.9
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	500	25.1/24.7
$b\bar{b}A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	600	8.7/8.7

**Table D.2:** *gluon fusion produced signal MC samples and cross section including branching ratios. BR is a branching ratio of  $A/H \rightarrow \tau\tau \rightarrow \tau_h\tau_h$ , which is  $BR(A/H \rightarrow \tau\tau) \times BR^2(\tau \rightarrow \tau_h)(=0.648)$ .*

Process	$m_A$ (GeV)	$\sigma_{A/H} \times$ BR( $A/H \rightarrow \tau\tau$ ) (fb) for $\tan\beta = 20$
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	110	11560.5/849.0
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	120	7486.9/1102.7
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	130	4990.6/2977.3
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	140	3410.6/3314.7
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	150	2382.0/2340.8
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	170	1227.8/1211.6
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	200	504.5/506.3
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	250	141.0/147.3
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	300	46.3/51.7
$gg \rightarrow A, A \rightarrow \tau\tau \rightarrow \tau_h\tau_h$	350	16.8/20.4

# Appendix E

## Electroweak and top Monte Carlo

This appendix lists all the generated background samples used in this analysis. The production of  $W$  and  $Z$  bosons in association with jets via strong interactions was simulated with the ALPGEN [69] generator, which employs the MLM matching scheme [86] between the hard process, calculated with leading-order matrix elements for up to 5 jets, and the parton shower. The  $t\bar{t}$ , single-top ( $t/s$ -channels,  $Wt$ ) and di-boson ( $WW$ ,  $WZ$ ,  $ZZ$ ) processes were generated using MC@NLO [67]. For all the event samples described above the parton shower and hadronisation were simulated with HERWIG [70] and the activity of the underlying event with JIMMY [72]. The loop-induced  $gg \rightarrow WW$  processes were generated using gg2WW [73]. QCD jet events were generated using PYTHIA [66].

**Table E.1:** *W+jets background samples and cross-sections*

Process	$\sigma \times \text{BR}$ (pb)	k-factor	Generator
$W \rightarrow e\nu + \text{Np0}$	6913.3	1.20	ALPGEN
$W \rightarrow e\nu + \text{Np1}$	1293.0	1.20	ALPGEN
$W \rightarrow e\nu + \text{Np2}$	377.1	1.20	ALPGEN
$W \rightarrow e\nu + \text{Np3}$	100.9	1.20	ALPGEN
$W \rightarrow e\nu + \text{Np4}$	25.3	1.20	ALPGEN
$W \rightarrow e\nu + \text{Np5}$	6.9	1.20	ALPGEN
$W \rightarrow \mu\nu + \text{Np0}$	6935.4	1.20	ALPGEN
$W \rightarrow \mu\nu + \text{Np1}$	1281.2	1.20	ALPGEN
$W \rightarrow \mu\nu + \text{Np2}$	375.3	1.20	ALPGEN
$W \rightarrow \mu\nu + \text{Np3}$	101.1	1.20	ALPGEN
$W \rightarrow \mu\nu + \text{Np4}$	25.7	1.20	ALPGEN
$W \rightarrow \mu\nu + \text{Np5}$	7.0	1.20	ALPGEN
$W \rightarrow \tau\nu + \text{Np0}$	6835.8	1.20	ALPGEN
$W \rightarrow \tau\nu + \text{Np1}$	1276.8	1.20	ALPGEN
$W \rightarrow \tau\nu + \text{Np2}$	376.6	1.20	ALPGEN
$W \rightarrow \tau\nu + \text{Np3}$	100.8	1.20	ALPGEN
$W \rightarrow \tau\nu + \text{Np4}$	25.7	1.20	ALPGEN
$W \rightarrow \tau\nu + \text{Np5}$	7.0	1.20	ALPGEN

**Table E.2:**  $Z/\gamma^*$  background samples and cross-sections ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ )

Process	$\sigma \times \text{BR}$ (pb)	k-factor	Generator
$Z \rightarrow ee$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np0	661.9	1.26	ALPGEN
$Z \rightarrow ee$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np1	133.3	1.26	ALPGEN
$Z \rightarrow ee$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np2	40.3	1.26	ALPGEN
$Z \rightarrow ee$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np3	11.2	1.26	ALPGEN
$Z \rightarrow ee$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np4	2.7	1.26	ALPGEN
$Z \rightarrow ee$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np5	0.8	1.26	ALPGEN
$Z \rightarrow \mu\mu$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np0	657.7	1.26	ALPGEN
$Z \rightarrow \mu\mu$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np1	132.8	1.26	ALPGEN
$Z \rightarrow \mu\mu$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np2	39.6	1.26	ALPGEN
$Z \rightarrow \mu\mu$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np3	11.1	1.26	ALPGEN
$Z \rightarrow \mu\mu$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np4	2.8	1.26	ALPGEN
$Z \rightarrow \mu\mu$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np5	0.8	1.26	ALPGEN
$Z \rightarrow \tau\tau$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np0	657.4	1.26	ALPGEN
$Z \rightarrow \tau\tau$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np1	133.0	1.26	ALPGEN
$Z \rightarrow \tau\tau$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np2	40.4	1.26	ALPGEN
$Z \rightarrow \tau\tau$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np3	11.0	1.26	ALPGEN
$Z \rightarrow \tau\tau$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np4	2.9	1.26	ALPGEN
$Z \rightarrow \tau\tau$ ( $40 \text{ GeV} < m_{\ell\ell} < 2 \text{ TeV}$ ) + Np5	0.7	1.26	ALPGEN

**Table E.3:**  $Z/\gamma^*$  background samples and cross-sections ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ )

Process	$\sigma \times \text{BR}$ (pb)	k-factor	Generator
$Z \rightarrow ee$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np0	3051.62	1.22	ALPGEN
$Z \rightarrow ee$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np1	87.87	1.22	ALPGEN
$Z \rightarrow ee$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np2	41.10	1.22	ALPGEN
$Z \rightarrow ee$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np3	8.46	1.22	ALPGEN
$Z \rightarrow ee$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np4	1.84	1.22	ALPGEN
$Z \rightarrow ee$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np5	0.46	1.22	ALPGEN
$Z \rightarrow \mu\mu$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np0	3051.62	1.22	ALPGEN
$Z \rightarrow \mu\mu$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np1	87.87	1.22	ALPGEN
$Z \rightarrow \mu\mu$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np2	40.95	1.22	ALPGEN
$Z \rightarrow \mu\mu$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np3	8.41	1.22	ALPGEN
$Z \rightarrow \mu\mu$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np4	1.85	1.22	ALPGEN
$Z \rightarrow \mu\mu$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np5	0.46	1.22	ALPGEN
$Z \rightarrow \tau\tau$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np0	3054.68	1.22	ALPGEN
$Z \rightarrow \tau\tau$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np1	86.48	1.22	ALPGEN
$Z \rightarrow \tau\tau$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np2	41.53	1.22	ALPGEN
$Z \rightarrow \tau\tau$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np3	8.37	1.22	ALPGEN
$Z \rightarrow \tau\tau$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np4	1.84	1.22	ALPGEN
$Z \rightarrow \tau\tau$ ( $10 \text{ GeV} < m_{\ell\ell} < 40 \text{ GeV}$ ) + Np5	0.46	1.22	ALPGEN

**Table E.4:**  $t\bar{t}$  and single top background samples and cross-sections

Process	$\sigma \times \text{BR}$ (pb)	k-factor	Generator
$t\bar{t}$ (no fully hadronic decays)	164.570	1	MC@NLO
$t\bar{t}$ (fully hadronic decays)	164.570	1	MC@NLO
single top : $t$ -channel $W \rightarrow e\nu$	6.342	1	MC@NLO
single top : $t$ -channel $W \rightarrow \mu\nu$	6.342	1	MC@NLO
single top : $t$ -channel $W \rightarrow \tau\nu$	6.342	1	MC@NLO
single top : $s$ -channel $W \rightarrow e\nu$	0.4252	1	MC@NLO
single top : $s$ -channel $W \rightarrow \mu\nu$	0.4252	1	MC@NLO
single top : $s$ -channel $W \rightarrow \tau\nu$	0.4252	1	MC@NLO
single top : $Wt$ -channel	13.102	1	MC@NLO

**Table E.5:** *WW background samples and cross-sections*

Process	$\sigma \times \text{BR}$ (pb)	k-factor	Generator
$q\bar{q}' \rightarrow W^+W^- \rightarrow e^+\nu e^-\nu$	0.52371	1	MC@NLO
$q\bar{q}' \rightarrow W^+W^- \rightarrow e^+\nu\mu^-\nu$	0.52371	1	MC@NLO
$q\bar{q}' \rightarrow W^+W^- \rightarrow e^+\nu\tau^-\nu$	0.52371	1	MC@NLO
$q\bar{q}' \rightarrow W^+W^- \rightarrow \mu^+\nu\mu^-\nu$	0.52371	1	MC@NLO
$q\bar{q}' \rightarrow W^+W^- \rightarrow \mu^+\nu e^-\nu$	0.52371	1	MC@NLO
$q\bar{q}' \rightarrow W^+W^- \rightarrow \mu^+\nu\tau^-\nu$	0.52371	1	MC@NLO
$q\bar{q}' \rightarrow W^+W^- \rightarrow \tau^+\nu\tau^-\nu$	0.52371	1	MC@NLO
$q\bar{q}' \rightarrow W^+W^- \rightarrow \tau^+\nu e^-\nu$	0.52371	1	MC@NLO
$q\bar{q}' \rightarrow W^+W^- \rightarrow \tau^+\nu\mu^-\nu$	0.52371	1	MC@NLO

**Table E.6:**  $WZ$  and  $ZZ$  background samples and cross-sections

Process	$\sigma \times \text{BR}$ (pb)	k-factor	Generator
$W^+Z \rightarrow \ell\nu qq$	1.736	1	MC@NLO
$W^+Z \rightarrow \ell\nu \ell\ell$	0.167	1	MC@NLO
$W^+Z \rightarrow qq\ell\ell$	0.522	1	MC@NLO
$W^+Z \rightarrow \tau\nu\ell\ell$	0.083	1	MC@NLO
$W^+Z \rightarrow \ell\nu\tau\tau$	0.083	1	MC@NLO
$W^+Z \rightarrow \tau\nu\tau\tau$	0.042	1	MC@NLO
$W^+Z \rightarrow qq\tau\tau$	0.261	1	MC@NLO
$W^-Z \rightarrow \ell\nu qq$	0.977	1	MC@NLO
$W^-Z \rightarrow \ell\nu \ell\ell$	0.094	1	MC@NLO
$W^-Z \rightarrow qq\ell\ell$	0.294	1	MC@NLO
$W^-Z \rightarrow \tau\nu\ell\ell$	0.047	1	MC@NLO
$W^-Z \rightarrow \ell\nu\tau\tau$	0.047	1	MC@NLO
$W^-Z \rightarrow \tau\nu\tau\tau$	0.023	1	MC@NLO
$W^-Z \rightarrow qq\tau\tau$	0.147	1	MC@NLO
$ZZ \rightarrow \ell\ell qq$	0.5299	1	MC@NLO
$ZZ \rightarrow \ell\ell\ell\ell$	0.0255	1	MC@NLO
$ZZ \rightarrow \ell\ell\nu\nu$	0.1517	1	MC@NLO
$ZZ \rightarrow \ell\ell\tau\tau$	0.0255	1	MC@NLO
$ZZ \rightarrow \tau\tau\tau\tau$	0.0064	1	MC@NLO
$ZZ \rightarrow \tau\tau\nu\nu$	0.0758	1	MC@NLO
$ZZ \rightarrow \tau\tau qq$	0.2649	1	MC@NLO

# Bibliography

- [1] D. Benjamin *et al.* [Tevatron New Phenomena and Higgs Working Group Collaboration], “Combined CDF and D0 Upper Limits on MSSM Higgs Boson Production in tau-tau Final States with up to 2.2 fb<sup>-1</sup>,” arXiv:1003.3363 [hep-ex].
- [2] V. M. Abazov *et al.* [D0 Collaboration], “Search for neutral Minimal Supersymmetric Standard Model Higgs bosons decaying to tau pairs produced in association with  $b$  quarks in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV,” Phys. Rev. Lett. **107**, 121801 (2011) [arXiv:1106.4885 [hep-ex]].
- [3] S. L. Glashow, “Partial Symmetries of Weak Interactions,” Nucl. Phys. **22**, 579 (1961).
- [4] R. Brout and F. Englert, “Spontaneous symmetry breaking in gauge theories: A Historical survey,” In \*Jerusalem 1997, High energy physics\* 3-10 [hep-th/9802142].
- [5] Peter Higgs, *Broken Symmetries and the Masses of Gauge Bosons*. Phys. Rev. Lett. **13**, 508 (1964).
- [6] M. H. Ahn *et al.* [K2K Collaboration], “Measurement of Neutrino Oscillation by the K2K Experiment,” Phys. Rev. D **74**, 072003 (2006) [hep-ex/0606032].
- [7] M. C. Sanchez *et al.* [Soudan 2 Collaboration], “Measurement of the L/E

- distributions of atmospheric neutrinos in Soudan 2 and their interpretation as neutrino oscillations,” *Phys. Rev. D* **68**, 113004 (2003) [hep-ex/0307069].
- [8] M. Roos, “Dark Matter: The evidence from astronomy, astrophysics and cosmology,” arXiv:1001.0316 [astro-ph.CO].
- [9] G. Aad *et al.* [ATLAS Collaboration], “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Phys. Lett. B* **716**, 1 (2012) [arXiv:1207.7214 [hep-ex]].
- [10] S. Chatrchyan *et al.* [CMS Collaboration], “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys. Lett. B* **716**, 30 (2012) [arXiv:1207.7235 [hep-ex]].
- [11] C. P. Burgess and G. D. Moore, “The standard model: A primer,” Cambridge, UK: Cambridge Univ. Pr. (2007)
- [12] **Particle Data Group** Collaboration, J. Beringer *et al.*, “Review of particle physics,” *Phys. Rev.* **D86** (2012) 010001.
- [13] I. J. R. Aitchison, “Supersymmetry in Particle Physics. An Elementary Introduction,” Cambridge, UK: Univ. Pr. (2007) 222 p
- [14] M. Carena, S. Heinemeyer, C. E. M. Wagner, and G. Weiglein, *Suggestions for benchmark scenarios for MSSM Higgs boson searches at hadron colliders*, *Eur. Phys. J. C* **26** (2003) 601–607, hep-ph/0202167.
- [15] Lyndon Evans and Philip Bryant (editors), *LHC Machine*, 2008 JINST 3 S08001
- [16] T. Chiarusi and M. Spurio, “High-Energy Astrophysics with Neutrino Telescopes,” *Eur. Phys. J. C* **65**, 649 (2010) [arXiv:0906.2634 [astro-ph.HE]].

- [17] J. Abraham *et al.* [Pierre Auger Collaboration], “Measurement of the energy spectrum of cosmic rays above  $10^{18}$  eV using the Pierre Auger Observatory,” *Phys. Lett. B* **685**, 239 (2010) [arXiv:1002.1975 [astro-ph.HE]].
- [18] The ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [19] Don Groom, *Atomic and Nuclear Properties of Materials for more than 300 materials*, Available at <http://pdg.lbl.gov/2012/AtomicNuclearProperties/>.
- [20] The ATLAS Collaboration, *Trigger performance*, in *ATLAS Detector and Physics Performance Technical Design Report*, pp. 347–400. CERN/LHC C/99-14/15, CERN, Geneva, 1999.
- [21] The ATLAS Collaboration, *ATLAS Monte Carlo Tunes for MC09*, ATL-PHYS-PUB-2010-002.
- [22] The ATLAS Collaboration, *Charged particle multiplicities in pp interactions at  $\sqrt{s} = 0.9$  and 7 TeV in a diffractive limited phase space measured with the ATLAS detector at the LHC and a new PYTHIA 6 tune*, ATL-CONF-2010-031, June, 2010.
- [23] I. Koletsou [ATLAS LAr Collaboration], “The ATLAS liquid argon calorimeter at the LHC ,” *Nucl. Instrum. Meth. A* **628**, 351 (2011).
- [24] W.E. Cleland and E.G. Stern, *Signal processing considerations for liquid ionization calorimeters in a high rate environment*, *Nucl. Instrum. Meth. A* **338** (1994) 467.
- [25] Gavrilenko, I., *Description of Global Pattern Recognition Program XKalman*, ATL-INDET-97-165.
- [26] The ATLAS Collaboration, *Charged-particle multiplicities in pp interactions at*

- $\sqrt{s} = 900 \text{ GeV}$  measured with the ATLAS detector at the LHC , Phys. Lett. B **688** (2010) 21 [arXiv:1003.3124]
- [27] Wasicki, C., *Track and vertex reconstruction of the ATLAS Inner Detector in the high multiplicity LHC environment*, ATL-COM-PHYS-2012-886.
- [28] W. Lampl et al., *Calorimeter Clustering Algorithms : Description and Performance*, ATL-LARG-PUB-2008-002.
- [29] T. Barillari et al., *Local Hadronic Calibration*, ATL-LARG-PUB-2009-001.
- [30] M. Cacciari, G. P. Salam and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, JHEP **04** (2008) 063.
- [31] M. Cacciari and G. P. Salam Phys. Lett. B **641** (2006) 57.
- [32] M. Cacciari, G. P. Salam and G. Soyez, *FastJet*, <http://fastjet.fr/>.
- [33] The ATLAS Collaboration, *Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at  $\sqrt{s} = 7 \text{ TeV}$  and measured with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2010-056, CERN, Geneva, Jul, 2010.
- [34] The ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, CERN-OPEN-2008-020, arXiv:0901.0512.
- [35] The ATLAS Collaboration, *Measurement of the  $W \rightarrow \ell \nu$  and  $Z/\gamma^* \rightarrow \ell \ell$  production cross sections in proton-proton collisions at  $\sqrt{s} = 7 \text{ TeV}$  with the ATLAS detector*, JHEP **1012** (2010) 060, arXiv:1010.2130 [hep-ex].
- [36] The GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 - a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [37] The ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, ATLAS-SOFT-2010-01-004, submitted to Eur. Phys. J. C., arXiv:1005.4568.

- [38] S. Schael and others, LEPH Collaboration and DELPHI Collaboration and L3 Collaboration and OPAL Collaborations and LEP Working Group for Higgs Boson Searches, *Search for neutral MSSM Higgs bosons at LEP*, Eur. Phys. J. **C47** (2006) 547.
- [39] M. Spira, *HIGLU: A Program for the Calculation of the Total Higgs Production Cross Section at Hadron Colliders via Gluon Fusion including QCD Corrections*, hep-ph/9510347 (1995) .
- [40] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, Phys. Rev. Lett. **88** (2002) 201801, hep-ph/0201206.
- [41] R. V. Harlander and W. B. Kilgore, *Higgs boson production in bottom quark fusion at next-to-next-to-leading order*, Phys. Rev. D **68** (2003) 013001.
- [42] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka (Eds.), et al., *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, arXiv:1101.0593 [hep-ph].
- [43] S. Dittmaier, M. Kramer, 1, and M. Spira, *Higgs radiation off bottom quarks at the Tevatron and the LHC* , Phys. Rev. **D70** (2004) 074010, arXiv:hep-ph/0309204.
- [44] M. Frank et al., *The Higgs Boson Masses and Mixings of the Complex MSSM in the Feynman-Diagrammatic Approach*, JHEP **0702** (2007) 047, hep-ph/0611326.
- [45] R. Harlander, M. Kramer, and M. Schumacher, *Bottom-quark associated Higgs-boson production: reconciling the four- and five-flavour scheme approach*, CERN-PH-TH/2011-134, FR-PHENO-2011-009, TTK-11-17, WUB/11-04.
- [46] MSSM Neutral Higgs WG, *MSSM Neutral Higgs Tool*, [https://twiki.cern.ch/twiki/bin/view/LHC Physics/MSSMNeutral](https://twiki.cern.ch/twiki/bin/view/LHC%20Physics/MSSMNeutral).

- [47] ATLAS Top WG, *Top MC2009*,  
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopMC2009>.
- [48] S.Allwood-Spires et al., *Monte Carlo samples used for top physics*,  
ATL-PHYS-INT-2010-132.
- [49] J.M.Butterworth et al., *Single Boson and Diboson Production Cross Sections in pp Collision at  $\sqrt{s} = 7$  TeV*, [arXiv:1012.4686 \[hep-ex\]](#).
- [50] The ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons*, ATLAS-COM-CONF-2011-057.
- [51] *Tau Reconstruction and Identification Performance in ATLAS*,  
ATLAS-CONF-2010-086, Oct, 2010.
- [52] *Measurement of the Mis-identification Probability of  $\tau$  Leptons from Hadronic Jets and from Electrons*, ATLAS-CONF-2011-113, July, 2011.
- [53] The ATLAS Collaboration, *Observation of  $Z \rightarrow \tau_h \tau_l$  Decays with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2011-010.
- [54] The ATLAS Collaboration, *Measurement of the  $W \rightarrow l\nu$  and  $Z/\gamma^* \rightarrow ll$  production cross sections in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, JHEP **12** (2010) 060.
- [55] The ATLAS Collaboration,  *$Z \rightarrow \tau\tau$  cross section measurement in proton-proton collisions at 7 TeV with the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2012-006.
- [56] The ATLAS Collaboration, *Expected electron performance in the ATLAS experiment*, ATLAS-PUB-2011-006.

- [57] The ATLAS Collaboration, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of  $\sqrt{s} = 7$  TeV*, ATLAS-CONF-2010-038.
- [58] ATLAS EGamma WG, *LAr Cleaning And Object Quality*,  
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/LArCleaningAndObjectQuality>.
- [59] ATLAS Physics Analysis Tools, *Pile up Reweighting Tools*,  
<https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PileupReweighting>.
- [60] J. N.Moeser, M.Schmitz, *Estimation of  $Z \rightarrow \tau\tau$  background in VBF  $H \rightarrow \tau\tau$  searches from  $Z\mu\mu$  data using an embedding technique*, Tech. Rep. ATL-PHYS-INT-2009-109, CERN, Geneva, December, 2009.
- [61] C. Collard [ATLAS Liquid Argon Collaboration], “Electronic calibration of the ATLAS LAr calorimeter,” Nucl. Instrum. Meth. A **623**, 246 (2010).
- [62] A. Ahmed et al., *Supporting Document: Measurement of the  $W$  cross section and asymmetry in the and muon decay channels at  $\sqrt{s}=7$  TeV*, Tech. Rep. ATL-PHYS-INT-2010-116, CERN, Geneva, Sep, 2010.
- [63] The ATLAS Collaboration, *Observation of  $W\tau\nu$  Decays with the ATLAS Experiment*, Tech. Rep. ATLAS-CONF-2010-097, CERN, Geneva, Nov, 2010.
- [64] The ATLAS Collaboration, *Measurement of the  $W \rightarrow \tau\nu$  Production Cross Section in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV with the ATLAS Experiment*, Tech. Rep. ATLAS-COM-PHYS-2011-778, CERN, Geneva, Nov, 2010.
- [65] The ATLAS Collaboration, *Updated Luminosity Determination in  $pp$  Collisions at  $\sqrt{s}=7$  TeV using the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2011-011, CERN, Geneva, Mar, 2011.

- [66] T. Sjostrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026.
- [67] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, hep-ph/0204244.
- [68] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007.
- [69] M. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001.
- [70] G. Corcella et al., *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010.
- [71] J. C. Collins, “Sudakov form-factors,” Adv. Ser. Direct. High Energy Phys. **5**, 573 (1989) [hep-ph/0312336].
- [72] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton Interactions in Photoproduction at HERA*, Z. Phys. **C72** (1996) 637.
- [73] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer, *Gluon-induced W-boson pair production at the LHC*, JHEP **12** (2006) 046.
- [74] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys. Rev. D **78** (2008) 013004.
- [75] J. Pumplin, *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **0207** (2002) 012.
- [76] A. Sherstnev and R. S. Thorne, *Parton Distributions for Leading Order Generators*, Eur. Phys. J. **C55** (2008) 553.

- [77] S. Jadach, J. H. Kuhn and Z. Was, *TAUOLA - a library of Monte Carlo programs to simulate decays of polarized  $\tau$  leptons*, Comput. Phys. Commun. **64** (1990) 275.
- [78] E. Barberio, B. V. Eijk and Z. Was, *Photos - a universal Monte Carlo for QED radiative corrections in decays*, Comput. Phys. Commun. **66** (1991) 115.
- [79] TauWG.  
<https://twiki.cern.ch/twiki/bin/view/atlasprotected/tausystematicssummerconf2011>.
- [80] The ATLAS Collaboration, *In-situ pseudo-rapidity inter-calibration to evaluate jet energy scale uncertainty and calorimeter performance in the forward region*, Tech. Rep. ATLAS-CONF-2010-055, CERN, Geneva, Jul, 2010.
- [81] The ATLAS Collaboration, *ATLAS Calorimeter Response to Single Isolated Hadrons and Estimation of the Calorimeter Jet Scale Uncertainty*, Tech. Rep. ATLAS-CONF-2011-028, CERN, Geneva, Mar, 2011.
- [82] The ATLAS Collaboration, *Reconstruction and Calibration of Missing Transverse Energy and Performance in Z and W events in ATLAS Proton-Proton Collisions at 7 TeV*, Tech. Rep. ATLAS-CONF-2011-080, CERN, Geneva, Jun, 2011.
- [83] The ATLAS Collaboration, *Performance of the Missing Transverse Energy Reconstruction and Calibration in Proton-Proton Collisions at a Center-of-Mass Energy of 7 TeV with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-057, CERN, Geneva, Jul, 2010.
- [84] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, [arXiv:1007.1727 \[physics.data-an\]](#).
- [85] G. Cowan, “Topics in statistical data analysis for high-energy physics,” CERN Yellow Report CERN-2010-002, 197-218 [[arXiv:1012.3589 \[physics.data-an\]](#)].

- [86] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. **C53** (2008) 473, [arXiv:0706.2569](#).
- [87] V. N. Gribov and L. N. Lipatov, “Deep Inelastic  $ep$  Scattering In Perturbation Theory,” *Yad. Fiz.* **15**, 781 (1972) [*Sov. J. Nucl. Phys.* **15**, 438 (1972)].  
L. N. Lipatov, “The Parton Model And Perturbation Theory,” *Sov. J. Nucl. Phys.* **20**, 94 (1975) [*Yad. Fiz.* **20**, 181 (1974)]. G. Altarelli and G. Parisi, “Asymptotic Freedom in Parton Language,” *Nucl. Phys. B* **126**, 298 (1977).
- [88] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, “Parton distributions for the LHC,” *Eur. Phys. J. C* **63**, 189 (2009) [[arXiv:0901.0002 \[hep-ph\]](#)].